

Research and Development for Life Support Robots that Coexist in Harmony with People

Nobuto Matsuhira, Hideki Ogawa
Takashi Yoshimi, Fumio Ozaki and Hideaki Hashimoto
*Corporate R&D Center, Toshiba Corporation
Japan*

Hiroshi Mizoguchi
*Dept. of Mechanical Engineering, Tokyo University of Science
Japan*

1. Introduction

Robot technologies have been developed for robots used in special environments such as industrial robots, maintenance robots for nuclear plants and robots used in space. Recently, the scope of robot applications has been expanded to include the medical and welfare fields, maintenance and patrolling of buildings, and home-use robots. Although some robots have already been developed, work on practical robots for this field has only recently begun.

Toshiba's development of robot technologies has also reflected this trend (Matsuhira & Ogawa, 2004). Robots to support human life in the home, at public facilities, or outside of factories are called "life support robots". In this field, it is essential that robots understand the human intent and interact with humans and the environment. Thus, human-centric technologies are important not only for robots but also for all machines that humans use. On the other hand, Toshiba defined "humancentric technology" as technology that contributes to human life and society. Robots are representative of humancentric technology as shown in Fig. 1. Such robots will interact with humans by the integration of hearing, vision, and motion.

Toshiba has already developed a conceptual model of a robotic information home appliance "ApriAlpha™" to provide an interface between human and networked appliances (Yoshimi, et al., 2004). The robot is considered to be applicable to various fields. Based on ApriAlpha™, two types of robots offering improved performance, namely, ApriAlpha™ V3 and ApriAttenda™, are being developed. ApriAlpha™ V3 features an enhanced omnidirectional auditory function and ApriAttenda™ is a person-following robot. These two robots have been developed as part of the next-generation robot project of the New Energy and Industrial Technology Development Organization (NEDO) and exhibited to verify the performance in the prototype robot exhibition in June at Aichi EXPO 2005. Toshiba and Tokyo University of Science collaborated on the development of ApriAttenda™.

The remainder of this paper is organized as follows. Chapter 2 explains the problems of life support robots. Chapter 3 explains ApriAlpha™ in terms of its performance as a robotic information appliance and the network communication technology it employs. Chapter 4

explains the sharp ear robot ApriAlpha™ V3, which features an enhanced auditory function. Chapter 5 explains the person-following robot ApriAttenda™, which is designed to accompany a person and serve him/her anywhere. Chapter 6 explains the navigation technology implemented in ApriAttenda™ for map building and performing self-localization. Chapter 7 explains the importance of environmental design, including information and physical environments with universal design, to realize working robots in daily life.

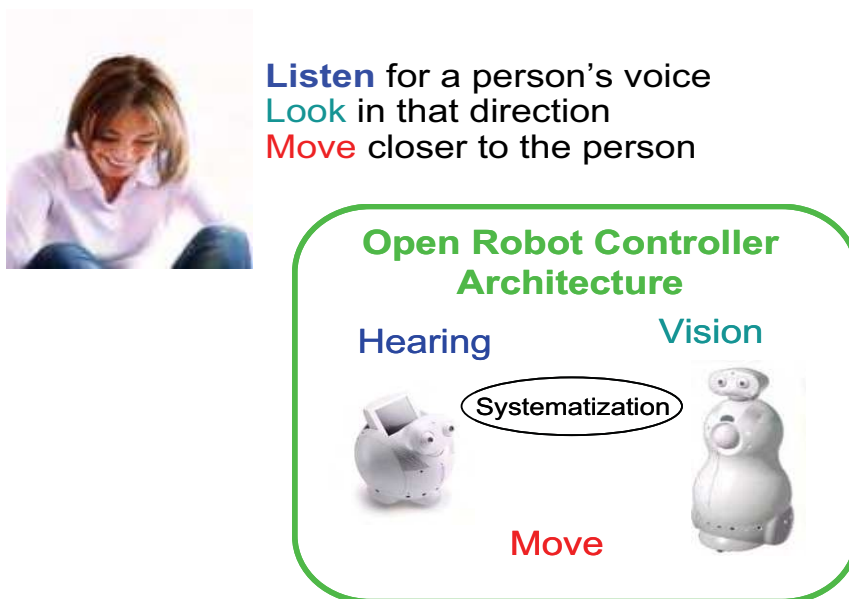


Fig. 1. Concept of human and robots with human-centric technology.

2. Subjects of Life Support Robots

In recent years, Japan and certain other countries have experienced a declining birthrate and a growing population of elderly people, the widespread introduction of information technology, and a growing preoccupation with security, including in the home. In particular, it is forecast that the numbers of skilled workers will decrease from 2007 onward in Japan. These developments pose difficult problems, which robot technologies are expected to ameliorate provided reliable and safe robotic technologies become available for applications in everyday life. However, there are many problems to be solved. In industrial plants and nuclear facilities, suitable environments are prepared in advance for robots. Parts are designed for easy handling and easy assembly procedures for robots. Working environments and assembling parts are prepared for robots. On the other hand, in the field of everyday life, the following problems remain because of the variation of users' requirements.

- a) Working environment and required tasks,
- b) Image processing environment including background scene and lighting, and
- c) Auditory processing environment including background noise such as TV are not constant.

Furthermore, robots are desired that can be used easily by anybody. To solve these problems, it is necessary to develop simple robots first for limited environments and tasks. Next, multifunctional robots should be developed by defining tasks step by step. And environmental design and knowledge embedded in the environment will be collaborated with robots, as necessary.

In daily life, not only the technologies for motion control of manipulation and mobility to perform the actual tasks, but also human machine interface technologies are important, such as for interaction between humans and robots by image and voice. A robot cannot perform services unless it communicates with a human to ascertain where he/she is and what he/she wants. Furthermore, it is natural and reliable for a robot to accompany the human when the robot performs services. Accordingly, a robot with omnidirectional auditory function and a robot with visual tracking function for a specified person have been developed as human interface technology.

Since a robot is a system, even if the image or voice processing function works well alone, the overall performance of the robot may be insufficient. Seamless integration of all functions, not only vision and auditory functions but also motion, is required. Regarding systemization, so far, the open robot controller architecture has been adopted to integrate auditory, vision, and motion functions as shown in Fig. 1. Robots working in everyday life need integration of vision, hearing, and motion. Vision and hearing are the interface functions enabling the robot to recognize the person and the commands. The voice and sound source position detected by hearing, and environment, obstacles, and human faces will be detected by vision, and the inputs' information is considered in order to plan for the robot motion. For that reason, we developed the auditory function and the person-following function and navigation function using the vision system as core technologies of human-friendly robots as described below. First, the basic function of ApriAlpha™ is explained.

3. Robotic Information Appliance “ApriAlpha™” with Network Ability

ApriAlpha™ is a conceptual model of a robotic information appliance. It recognizes the face and voice of the user and replies to the user by voice, and moves by wheels in the home. The robot was developed to perform the functions of house sitting and intelligent remote control for home appliances. Here, ApriAlpha™ means an advanced personal robotic appliance type alpha.

3.1 Open robot controller architecture

For ApriAlpha™, the framework of open robot controller architecture was adopted, integrating various robot technologies with a distributed object technology based robot controller that we are currently developing (Ozaki, 2003). An expandable and easy-to-use robot controller can be realized. The distributed object technology can deal easily with various objects for various functions distributed on a network. When executing the desired object, it is unnecessary to be conscious of the object's place on the network. The desired object can be executed by issuing a command message in a standard format. Various robot technologies and/or interfaces with peripheral equipment are included in the system using HORB developed by the National Institute of Advanced Industrial Science and Technology (AIST) (<http://horb.aist.go.jp/>). Various modules of ApriAlpha™ are easily combined as shown in Fig. 2. By upgrading these modules or adding new modules, the performance of the robot will be improved or a new robot will be realized.

3.2 Networked robot

Fig. 3 shows the cooperation system with networked appliances (Matsuhira et al., 2005), (Ueno et al., 2004). A refrigerator, an air conditioner, and a cooking range are connected to the LAN through the access point of Bluetooth™. The server of a question-answer system is also connected to the LAN with the UPnP™ protocol. These appliances are recognized automatically by UPnP™ technology. The program of the robot controller can be described bilaterally in this framework, independent of the category of the apparatus, which includes robots or home appliances. Depending on the user's requests, the robot provides the latest news via the internet, controls the temperature of an air conditioner via the LAN, checks the ingredients in the refrigerator, displays a recipe, and sets the cooking range appropriately.

(1) Agent Technology

Flipcast™, developed by Toshiba, is a lightweight agent technology. The application of searching news requested by the user has been developed using Flipcast™. When the user asks, "What is today's news?" or "Tell me the sports news." to ApriAlpha™, ApriAlpha™ searches the internet in order to provide the latest news corresponding to the categories of economy, general news, and international news. The script using Flipcast™ combines news acquisition and an analysis function implemented on the server of the National Institute of Informatics (NII) and the voice recognition and synthesis functions of ApriAlpha™. Since this application is very simple, the script can be made easily for information search and information is acquired on the internet depending on the user's request.

(2) Question-answer system

Multi-modal help, using a question-answer system developed to support the user, enabling him/her to make full use of home electric appliances and AV equipment, is connected to the LAN. This system can offer information suitable for the user, searching technology using multi-modal knowledge content consisting of graphical images, audio assist, and text images.

(3) Cooperation with Networked Home Appliances

Numerous communication protocols for home networks have already been proposed such as ECHONET™ for home electric appliances, IEEE1394 for AV equipment, wired and wireless LAN, and telephone line. In particular, ECHONET™ can connect home appliances in a home network and operate appliances and manage their conditions. It is too difficult for a home robot to support all of these protocols used in home appliances. A new framework has been developed for cooperation between a home robot and networked appliances (Morioka et al., 2004). Within this framework, all of the protocols used in the home appliances concerned are combined and structured with UPnP™ technology. Using the UPnP™ protocol, all features of the original protocols are provided via the interface of the framework.

In Fig. 3, when the user asks ApriAlpha™, "Is there any news?", ApriAlpha™ searches the latest news and reads it out by using the agent function. When the user says, "I feel rather hot.", ApriAlpha™ controls the temperature of the air conditioner automatically. When the user asks, "What is in the refrigerator? Tell me a recipe suitable for the ingredients.", ApriAlpha™ displays a suitable recipe using the question-answer system for recipes. The user can ask questions about the recipe during preparation and cooking. Furthermore, ApriAlpha™ can input the recipe to the cooking range and notify the user when the cooking is finished. In the future, by engaging in simple conversations with robots that support their household affairs, people will be able to enjoy comfort and peace of mind, having been liberated from the necessity of executing troublesome operations with remote-control devices.

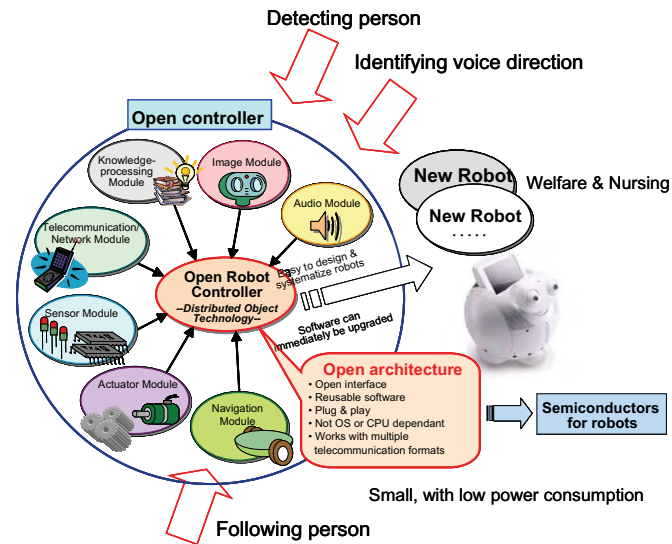


Fig. 2. Open robot controller architecture. New functions are easily added.

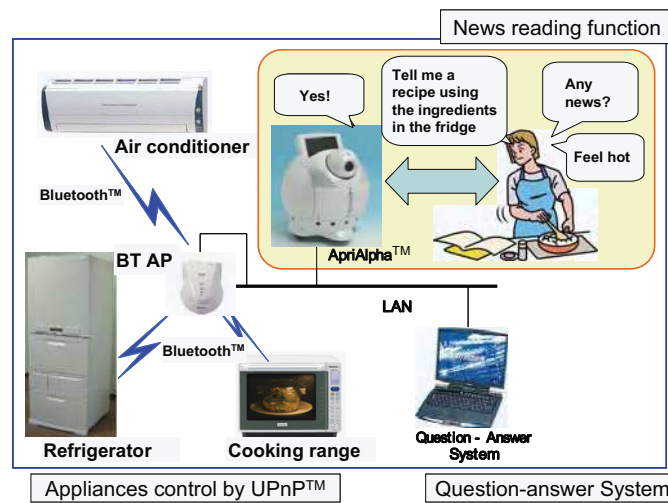


Fig. 3. Collaboration between ApriAlpha™ and home appliances.

4. Sharp Ear Robot “ApriAlpha™ V3”

4.1 The purpose of the sharp ear robot

In general, various noises exist in the home besides the orders spoken by the user to a robot. These include speech, the sound of the TV and other sound sources shown in Fig. 4. Thus, home robots

require a function enabling them to hear each voice from any direction and recognize the contents among the noises encountered in daily life. Therefore, research on an auditory function to recognize the voices around a robot has been performed actively (Nakadai et al., 2003), (Brandstein, 1999), (Asano, 2004), (Brandstein & Ward, 2001). We have realized a high-performance auditory function for sound source localization by the development of signal processing technology to detect a voice from any direction around a robot and to assume the voice source direction.

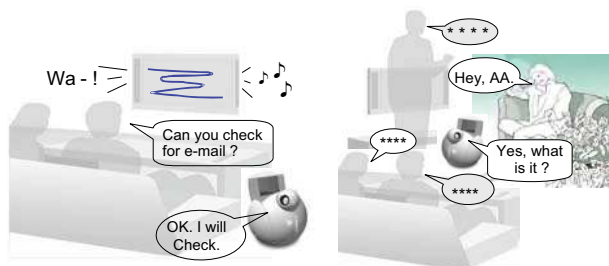


Fig. 4. Images of ApriAlpha™ in a living room.

4.2 Auditory signal processing for omnidirectional auditory function

Fig. 5 shows the block diagram of auditory signal processing to perform the omnidirectional auditory function. Two pairs of microphones from among the multiple microphones are selected and perform an auditory signal processing.

First, to assume the sound source direction, ΔT , which is the arrival time differences between two microphones, is used. If it is assumed that the wave from the sound source to the microphone is the plane wave, gained ΔT denotes the distance between the sound source and the two microphones. The angle θ between the microphone and the sound source direction is given as follows:

$$\theta = \sin^{-1}(\Delta T/d); \quad (4-1)$$

here, d is the distance between two microphones. Auditory signals are transformed by fast Fourier transformation (FFT) to phase difference and intensity data in the frequency region. Then, the number of the sound sources and their directions at each moment are assumed by the phase difference analysis, and the voices are detected as the time series of voice stream for each sound source. Here, Hough transformation was used to assume the sound source directions (Okazaki, 2000). Next, by matching the voice streams whose frequency ingredient and voice length are similar, the space position of a sound source is presumed. When the sound source positions are identified, noise reduced voice stream abstraction, by adapted array processing, is possible and each voice stream is recognized by the voice recognition engines. Here, array processing and recognition engines, which have already been developed, are used (Amada, 2004), (http://www3.toshiba.co.jp/pc/lalavoice/index_j.htm (in Japanese)). Thus, an auditory function that recognizes voices from any direction has been realized.

4.3 Developed ApriAlpha™ V3

To verify the developed auditory function, ApriAlpha™ V3 was developed based on ApriAlpha™ as shown in Fig. 6 and Table 1. It incorporates a prototype auditory signal processing board with an 8 ch parallel input interface. In the robot, auditory signals are input from six microphones mounted on the surface of the robot's body. Owing to smooth fusion with auditory function and mobile function, the robot can change direction quickly

whenever a voice is heard and approach the user. Fig. 7 shows the voice data processing result. This is an example of the result which detected the voice stream when two persons speak simultaneously from the different directions. One person stands at 20 degrees left side of the robot and the other stands at 40 degrees right side. From the directions where persons stand, voice streams were correctly extracted.

The developed robot performed a demonstration and verified the capability of omnidirectional auditory functions in an actual noisy exhibition hall at Aichi EXPO both in the “NEDO prototype robot exhibition” and the “robot exhibition at the Robot Station”. In the prototype robot exhibition shown in Fig. 8, a demonstration lasting about 10 minutes was performed 3 times per day, and two kinds of demonstrations were carried out with respect to a robot located at the center of a circular table with a diameter of 2 meters. In each demonstration, three people called the robot continuously from different positions around the table. One is a demonstration of household-electric-appliance control and information retrieval. Speaking in turn, the people order the robot to turn the air-conditioner and the electric light ON/OFF, and order the robot to retrieve the weather information from the Internet via wireless LAN and provide it. The robot could hear the orders and execute them without interruption. In the other demonstration, a robot is the chef of a sushi bar. The robot turns to the person who orders sushi from the circumference of a small stage, and recognizes the directions and responds by uttering the name of the type of sushi. Here, the number of voice commands is limited to about 10 words/phrases in order to verify the sound source localization. In the exhibition hall, the noise level at the circumference reached 80-85 dB at the start of the demonstration on the main stage of the Morizo-Kiccoro Messe. However, it was verified at our demonstration stage that the developed auditory function could work when the noise level at the circumference was about 80 dB. Moreover, ApriAlpha™ V3 could respond to a visitor’s voice at our local demonstration. To improve the auditory function, vision will be combined to search for the sound source localization. For example, the robot turns to the direction from where sound was detected to search for a face, and if a face is detected the localization performance is improved. When a voice command is not detected clearly, the robot can move to the position where the voice is detected more easily. So, sensing and motion should be combined to execute the task, rather than concentrating on the development of a dedicated sensing method. Of course, from the viewpoint of practical use, other methods of inputting commands should be considered, such as a touch panel.

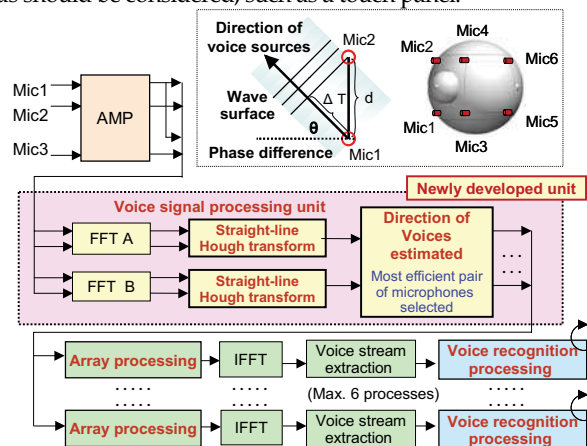


Fig. 5. Block diagram of auditory signal processing.



Fig. 6. Sharp ear robot ApriAlpha™ V3.

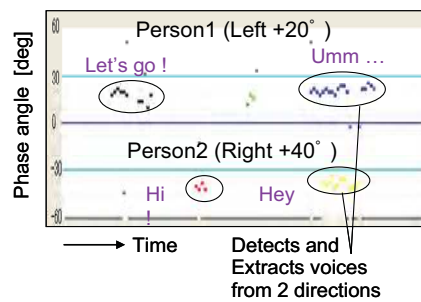


Fig. 7. Data processing results of ApriAlpha™ V3.



Fig. 8. Demonstration at Aichi EXPO.

ApriAlpha™ can recognize sequential orders from surrounding users.

Dimensions	Diameter: 380 mm Height: 430 mm
Weight	About 10 kg
User Interface	Microphone: 6 Speaker: 2 CCD Camera: 2 LCD with a touch panel
Motion	2 Independent drive wheels Face: rotate joint Eye: pan and tilt
Communication	Wireless LAN: IEEE802. 11 a/b
Power source	Lithium-ion battery: about 2 hours continuous motion

Table 1. Specifications of ApriAlpha™ V3.

5. Person-Following Robot “ApriAttenda™”

5.1 The purpose of the person-following robot

It is necessary for a robot working in the environment of everyday life not only to move to a specified place following a predetermined trajectory, but also to move with flexibility. To perform the motion, a person-following function that involves finding the specified person and following him/her is essential. Such a robot is expected to take care of an infant and or an elderly person, control home electric appliances, offer information services such as the news, and carry baggage as it follows the user in a shopping center as illustrated in Fig. 9. The robot can support everyday life from the viewpoints of safety, security and practicality.

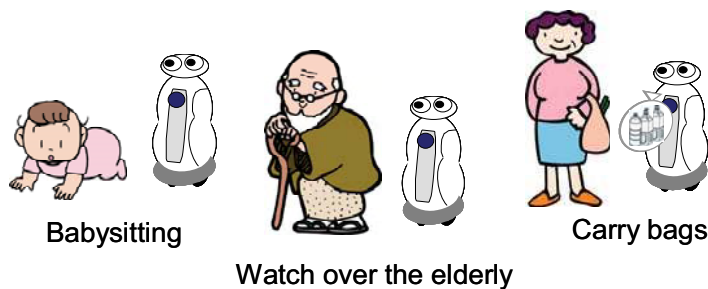


Fig. 9. Image of ApriAttenda™ used in everyday life.

5.2 Image processing of person-following robot

The basic functions involved in following a person are as follows: finding the specified person, following at his/her pace, avoiding obstacles, and resuming contacting when the robot misses him/her. Person-following robots developed until now use various types of cameras for detecting a target person, and some of them use other sensors (Schlegel et al., 1998), (Sidenbladh et al., 1999), (Kwon et al., 2005), (Cielniak et al., 2005), (Kobilarov et al., 2006). Our newly developed robot uses a stereo vision system. The combination of the vision system and the motion control system makes it possible to realize a robust person-following function. The control system from the viewpoints of function modules is shown in Fig. 10. For the person following function, Target Detection Module and Motion Control Module, were constructed. Two camera images of the target person including cluttered backgrounds are captured concurrently by the newly developed versatile multimedia front-end processing board (MFeP) (Sato et al., 2005), and sent to the Target Detection Module. At the Target Detection Module, the target person is detected by the newly developed image processing algorithm, and the result (distance and direction data of the target person from the robot) is sent to the Motion Control Module through the network. At the Motion Control Module, the two wheels and the head of the robot are controlled to follow the target person smoothly. The Target Detection Module is running on Windows PC and the Motion Control Module is running on Linux PC, because the Windows PC has many image processing applications, and the robot motion control requires real-time processing. The frame rate of the image processing system is about 15 fps, and the control cycle of the motion control system is 1kHz. Regarding ApriAttenda™'s systemization, the open robot controller architecture as shown in Chapter 3 has been adopted to easily integrate the Target Detection Module and the Motion Control Module.

To find the person, a shape of a human back and shoulder for visual tracking of the target person was used (Hirai & Mizoguchi, 2005). A new algorithm has been developed to recognize and abstract the region of the person from an image containing a complicated scene such as that shown in Fig. 11. The point at which the texture changes is abstracted automatically as a feature point. The distance from each feature point is measured by the stereo vision system, and then the person area is detected by the distribution of the distance and the history of the motion of feature points. Furthermore, by combining the information of the color and texture of the clothes the target person wears, the person area is identified without fail. A robust method against the lighting and change of scene has been developed by utilizing these variable information and importing and updating refresh feature points of the person detection area. In the following motion, the data captured by an ultrasonic sensor is imported during the person detection processing in parallel. When an obstacle is found by the ultrasonic sensor on a trajectory, the robot continues to follow the person by the vision sensor while the robot's body avoids the obstacle. Furthermore, in the case that the robot fails to detect and follow the person, searching motion starts to increase the searching area. A re-finding function is implemented that matches between the registered image data and the input image data of the clothes' texture.

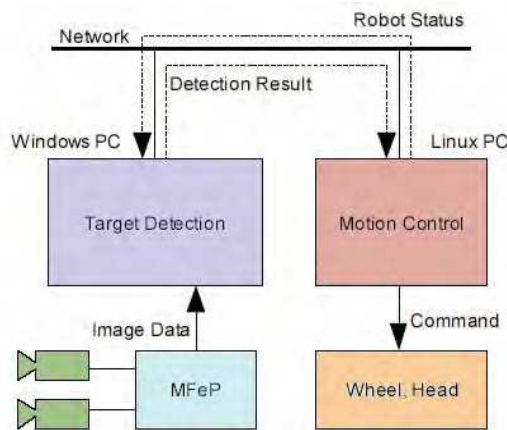


Fig. 10. Control system of ApriAttenda™.

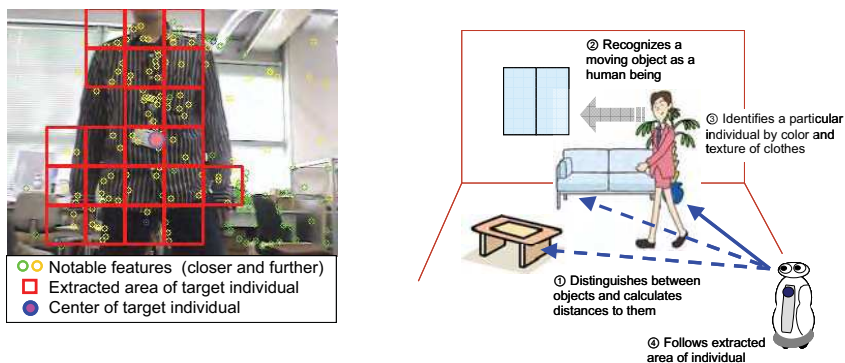


Fig. 11. Image processing algorithm of ApriAttenda™.

5.3 Developed ApriAttenda™

Fig. 12 shows a scene in which ApriAttenda™ is following a person. The specifications of ApriAttenda™ are shown in Table 2. The robot is 450 mm in diameter, 900 mm in height, and 30 kg in weight, and the robot shape is designed attractive and unintrusive. The robot finds the person by means of two CCD cameras on its head. Motion control with a newly developed inertia absorbing mechanism in the driving part realizes stable and smooth person-following motion.

At Aichi Expo, a robot could follow a person who walked freely in the demonstration space in front of many spectators as shown in Fig. 13. ApriAttenda™ calculates the distance between the target person and itself using stereo vision and follows him/her with the appropriate speed to keep the distance constant. Sometimes the robot failed to find the person, but once he/she moved into sight of the robot, the robot followed him/her again. In another demonstration, the robot could follow a child with a stuffed mascot "Kiccoro". The mascot was the target to be followed by the robot because anyone could join the robot-following demonstration by holding the mascot. This demonstration was also successful in the case that spectators were closer to the robot as shown in Fig. 13(b). Here, this robot verified the performance of person finding and tracking only by vision. But if the robot can hear the voice and recognize the face, or move to the position where it can search the voice and face more easily, the reliability of finding the person is drastically increased. So, the system integration in robots to perform the task robustly is very important as described in Chapter 2.



Fig. 12. Person-following robot ApriAttenda™.

Dimensions	Diameter: 450 mm Height: 900 mm
Weight	About 30 kg
User Interface	Microphone: 2 Speaker: 2 CCD camera: 2 LCD with a touch panel
Motion	2 Independent drive wheels Face: rotate 2 Joints Eye: pan and tilt
Communication	Wireless LAN: IEEE802. 11 a/b
Power source	Lithium-ion battery about 2 hour continuous motion

Table 2. Specifications of ApriAttenda™.



(a) Demonstration stage



(b) Searching for the mascot.

Fig. 13. Demonstration of ApriAttenda™ at Aichi EXPO.

6. Navigation of the Mobile Robot

Although ApriAttenda™ has a function for person searching and following, if the robot is ordered to watch or bring something, a guiding function is necessary to enable the robot to move freely in a house or any other building. For such a function, the map of a building is required. To build the map, a service-person or a user must input the map manually, teach the route using by a remote control device, or use the following function of the robot as described above. Target markers on the selected places and the function of local obstacle avoidance with ultrasonic sensors and so on are used for the robot to move to the desired position autonomously. In any event, target positions are input in advance. With regard to building a map, a great deal of research has been done on simultaneous localization and mapping (SLAM). Recently, the field of robotic SLAM is dominated by probabilistic techniques (Thrun, 2002). In this chapter, a novel map building and localization approach is explained. The map building stage is based on shape similarity comparison, which was originally developed for computer graphics applications. And the localization stage is based on an improved Monte Carlo localization. In the following parts of this Chapter, map building and localization are discussed, respectively. The experimental results are given based on the person-following robot ApriAttenda™ (Zhu et al., 2005).

6.1 Map building based on shape similarity

To acquire a map, the robot must memorize the perceived objects and features the robot has passed by, then merge the corresponding objects in consecutive scans of the local environment. Here in our approach, the information from LRF (Laser Range Finder) sensor on the robot and odometry is used to build a global map.

(1) Map building process

In the developed algorithm, the global map is built iteratively as the robot moves. We denote the scan and the global map at time t by S_t and G_t , respectively. Each scan S_t and global map G_t is composed of polylines. To create the global map G , we start with the first global map G_1

being equal to the first scan S_1 . Assuming at time $t-1$ the global map G_{t-1} has been created, G_{t-1} and S_t are used to create G_t in the following steps:

- a) S_t is converted into a cyclic, ordered vector of polylines. At this stage, noise and unreliable information in the scan reading should be excluded.
- b) A virtual scan VS_t is extracted from the global map G_{t-1} . VS_t is the part of the global map, which is assumed to be seen by the robot at position t .
- c) Sampling points from VS_t and S_t , then searching the corresponding points between VS_t and S_t .
- d) Repeatedly transforming S_t with different translation and rotation offset. By calculating the shape similarity value between VS_t and S_t , an optimal transformation matrix can be found, which will assign a minimal value to the shape similarity between two scans.
- e) Finally, the aligned S_t is merged into G_{t-1} to create G_t .

Repeating the above steps iteratively for each new scan, we are able to create a global map that remembers all the objects and features along the path the robot has traveled.

Here, the first step in this approach is excluding the noise from the scan, and grouping the valid scan points into cyclic ordered polylines. To make the representation more compact and to cancel noise, we employ a technique called DCE (discrete curve evolution) (Latecki & Lakamper, 1999) to achieve this goal.

(2) Matching map by shape similarity comparison

Before aligning a scan to the global map, not all the features in the scan and the global map should be compared. Here to save the computation cost, a virtual scan VS_t from the global map G_{t-1} is generated to calibrate the scan S_t . Since VS_t and V_t may have some information, which the other doesn't have, preprocessing is necessary to find out the corresponding points before the shape similarity comparison. Since we attempt to use the known information, which is assumed to be correct, to align the new scan, only the points with corresponding points will be used in shape similarity comparison. The points in candidate set VS_t are treated in the same way as well. Matching scans against the global map within the context of a shape-based representation is naturally based on shape matching. In order to find out how similar two shapes are to each other, we must know how different they are to each other.

For comparing two 3D models, the most commonly used error metric is L_2 norm. Here we would like to formulate shape-based analogs of L_2 norm. If $P(S_1)$ and $P(S_2)$ denote all the points in shape S_1 and S_2 , we can select two sets $X_1 \subset P(S_1)$ and $X_2 \subset P(S_2)$ containing k_1 and k_2 sample points, respectively. Thus, the sampled error metric between these two shapes $E_{avg}(S_1, S_2)$ is the average square distance from the sample points in X_1 and X_2 to shape S_2 and S_1 . In the application for comparing two scans, not all the points in the candidate point set are treated in the same way; a larger penalty coefficient is given for the feature points in similarity comparison, since mismatching of the feature points will result in a worse alignment for two scans. We let S_t^T be the scan transformed from S_t by $(\epsilon x, \epsilon y, \epsilon \theta)$. By iteratively transforming S_t with different $(\epsilon x, \epsilon y, \epsilon \theta)$, an optimal transformation vector $(\Delta x_t, \Delta y_t, \Delta \theta_t)$ could be found, which will assign a minimal value to the shape error metric. Therefore, it could be described as an optimization problem.

$$\text{Goal:} \quad \min(E_{avg}(VS_t, S_t^T))$$

$$\text{Where:} \quad S_t^T = T(S_t, \epsilon x, \epsilon y, \epsilon \theta)$$

$$\begin{aligned}
\text{Constrains:} \quad & \min(x) < \varepsilon x < \max(x) \\
& \min(y) < \varepsilon y < \max(y) \\
& \min(\theta) < \varepsilon \theta < \max(\theta)
\end{aligned} \tag{6-1}$$

The odometry error of the robot at time t will be the optimal solution of the above optimization problem.

(3) Merging maps

Merging is the task of combining similar line segments taken from the sensor reading S_t and the previous global map G_{t-1} to form a new global map G_t . During the merging process, we need to accurately identify objects in each scan, look for their corresponding objects in the existing global map, update the position of the object, remove moving objects, and connect separated objects into a single object whenever it is possible. Fig. 14 shows the map building result in Toshiba Science Museum, 1st floor hall (14 m x 18 m). The robot navigation trace from the odometry is displayed by the red line and the calibrated robot trace is displayed by the blue line.

6.2 Localization problems

The localization problems are of many different types (Thrun, et al., 2001). In the case of position tracking in Fig. 15, the initial robot position is known, and the problem is to compensate incremental errors in a robot's odometry. In a global localization problem, robot is to determine the position from the scratch. In the case of a global localization problem, the error in the robot's estimate cannot be assumed to be small. In the case of a kidnapped robot condition, a well-localized robot is moved to some other place without the robot's knowledge. The kidnapped robot problem is often used to test a robot's ability to recover from catastrophic localization failures. Unlike other algorithms, the Monte Carlo localization (MCL) approach solves a position tracking problem, a global localization problem and a kidnapped robot problem in a robust and efficient way.

(1) Monte Carlo localization approach

MCL uses Bayes filter to estimate the posterior distribution of robot poses based on sensor data in a recursive manner. According to Bayes filter, the future data is independent of past data if we have knowledge of the current state—an assumption typically referred to as the *Markov assumption*. In this case, the dynamical system is a mobile robot and its environment, the state is the robot's pose therein (a two-dimensional Cartesian space and the robot's heading direction), and measurements may include range measurements, camera images, and odometry readings.

If x denotes the state, x_t is the state at time t , and $d_{0..t}$ denotes the data starting at time 0 up to time t . Then, probability density of robot position is $p(x_t) = p(x_t | d_{0..t})$. For mobile robots, *perceptual data* o (for *observation*) such as laser range measurements, and *odometry data* a (for *action*), carry information about robot motion. Thus probability density of robot position becomes

$$p(x_t) = p(x_t | o_t, a_{t-1}, o_{t-1}, a_{t-2}, \dots, o_0) \tag{6-2}$$

where a_{t-1} refer to the odometry reading that measures the motion that occurred in the time interval $[t-1; t]$.

If initial knowledge is not available, it is typically initialized by a *uniform distribution* over the state space. To derive a recursive update equation, Eq.(6-2) can be transformed by Bayes rule and Markov assumption to (Dellaert et al.,1999).

$$p(x_t) = \eta p(o_t | x_t) \int p(x_t | x_{t-1}, a_{t-1}) p(x_{t-1}) dx_{t-1} \quad (6-3)$$

This equation is the recursive update equation in Bayes filter. Thus if we know the state of the robot at previous step, action or movement after the previous step and the observation or measurement information at the present state, we can easily determine the present state without knowing all the previous steps taken by the robot. Together with the initial belief, it defines a recursive estimator for the state of a partially observable system. This equation is the basis for MCL algorithms used in this work.

As mentioned above, to implement (6-3), two conditional densities are required: the probability $p(x_t | x_{t-1}, a_{t-1})$, which we will call *next state density* or simply *motion model*, and the density $p(o_t | x_t)$, which we will call *perceptual model* or *sensor model*. By simplifying, we can express these models as $p(x' | x, a)$, and $p(o | x)$, respectively.

(2) Probabilistic models for localization

The motion model, $p(x' | x, a)$, is a probabilistic generalization of robot kinematics. In this work, each pose comprises a robot's two-dimensional Cartesian coordinates and its heading direction (orientation, bearing). The value of a may be an odometry reading or a control command, both of which characterize the change of pose. The conventional kinematic equations, however, describe only the *expected* pose x' that the ideal, noise-free robot would attain starting at x , and after moving as specified by a . Of course, physical robot motion is erroneous; thus, the pose x' is uncertain. To account for this inherent uncertainty, the probabilistic motion model $p(x' | x, a)$ describes a posterior density of x' . In MCL the margin of uncertainty depends on the overall motion. For the MCL algorithm we need a *sampling model* of $p(x' | x, a)$ that accepts x and a as an input and generates random poses x' distributed according to $p(x' | x, a)$.

For the perceptual model or measurement model, $p(o | x)$, mobile robots commonly use range finders, such as ultrasonic transducers (sonar sensors), laser range finders, and camera images. A laser range finder and camera data are considered for localization in this work. In this work, first, the value of mean and standard deviation a noise-free sensor would generate is determined. Then the marker distance from the current robot position is calculated. Finally, the information of laser beams and marker distance is integrated into a single density value.

As noticed by other authors (Thrun et al., 2001), (Dellaert et al., 1999), the basic particle filter or MCL performs poorly if the number of particles is small, and if the number of particles is large, the processing time will increase and real-time or near real-time application will be impossible. Considering these problems, we kept the number of the particles to 100 in this case.

(3) Experimental studies

In this work, we took actual data with localization error from the robot movement in the Toshiba Science Museum in Fig. 15(a). The experimental scene is shown in Fig. 16. We used a similar map of the museum in this work. The actual map is more complex than the map we used in this simulation work. Generally, MCL algorithm works better with a comparatively more complex map such as that in Fig. 15(a) rather than with a simplified one, because in the simplified map there are many similar places and mean and standard deviations of LRF scan lines at these positions are close. In actual maps fewer places are similar and the LRF scan data are also different for different places. So mean and standard

deviations will also vary at different places and as a result localization works better in such a situation.

For localization, the robot has to be sure whether it should start position tracking steps or kidnapped condition recovery steps. If the ideal probability (calculated using LRF scan mean, standard deviation and the distance from the specific marker based on the map) and the actual probability (LRF and marker information collected by robot) differ significantly (e.g. more than 20%), the robot will start kidnapped recovery steps. Otherwise it will continue position tracking.

The process starts with collecting odometry information, LRF scan and marker distance. If the LRF using map at a position according to the odometry differs significantly from the actual LRF scan properties (mean and standard deviations), it starts kidnapped recovery steps. Otherwise it continues position tracking steps.

In the case of kidnapped recovery situation, the robot distributes particles [100 in this example] in the map randomly as shown in Fig. 15(b). Though in the actual map the area close to boundary is empty, the robot center cannot move very close to the wall or other obstacle because of the robot's size. This is considered in Fig. 15(b) where the robot will move only in the center area. The particles are distributed in this area to ensure efficient use of resources, because particles in the red area cannot represent the robot position and it is unnecessary to distribute particles in the area where the robot cannot go.

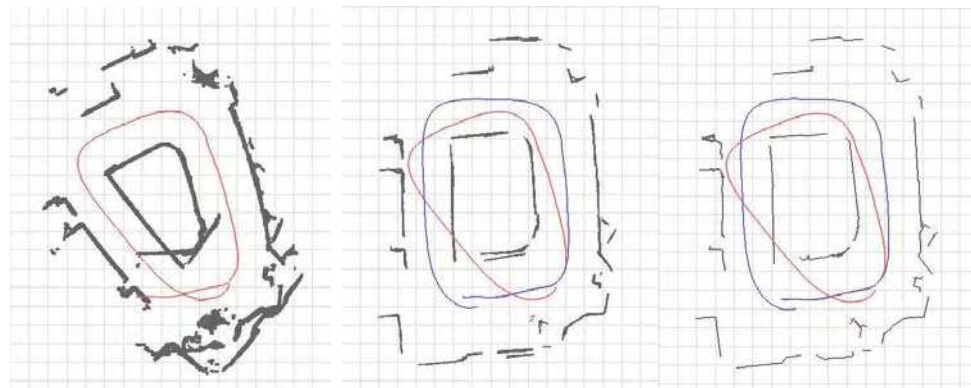
In this work we considered only those LRF scan lines that would reflect from walls or obstacles and reflect from a range of 20 mm to 4095 mm. In this case the LRF scan data is considered reliable for determining the robot position in the map. After random distribution, the robot determines the ideal LRF scan mean and standard deviation at each location of the particles. It also determines the marker position according to the map with which the robot has been provided. Then it starts comparing the mean, standard deviation and marker distance according to the actual data and the data determined using the map. The closer the mean of a particle to the actual condition, the higher is the probability for the robot to be at that particle position in Fig. 15(c).

If the robot fails to find a particle among the randomly distributed particles in one cycle it starts the whole cycle again until it finds a close position. During this time, instead of distributing all the particles randomly in the whole map, it uses the mean and standard deviations of the resultant particles at the end of each cycle to distribute particles for the next cycle.

For position tracking steps, the robot knows its initial position. So when it moves, it distributes the particles in the direction it is moving with position information by odometry. Then, it determines the best particle to decide its position using LRF and marker information.

(4) Result and discussion

The robot determined its position in both position tracking and kidnapped condition within 100 mm of the actual position. In order to increase accuracy of the position, it will be necessary for the robot to repeat the same steps several times, which will take more time. This is a subject of ongoing research regarding map building and localization. If we make the map and set the target markers in the building for self-localization of robot manually, the robot can move in the building. So, the challenge is to make a more intelligent robot.

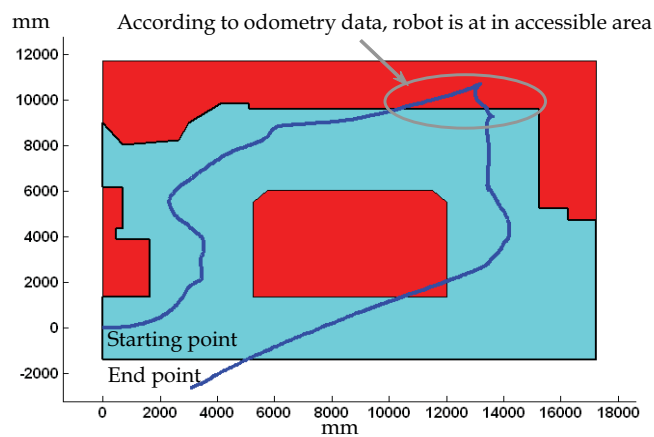


(a) Raw Scan Data of Toshiba Science Museum.

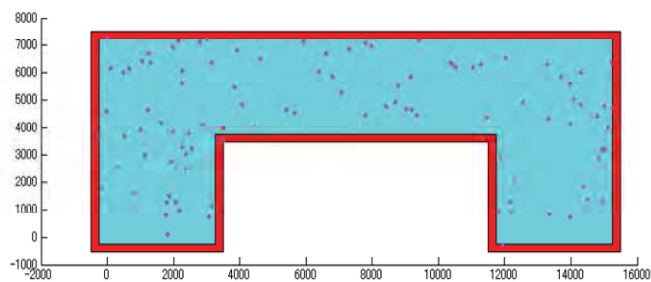
(b) Map of Toshiba Science Museum without Merging Process

(c) Feature Map for Toshiba Science Museum.

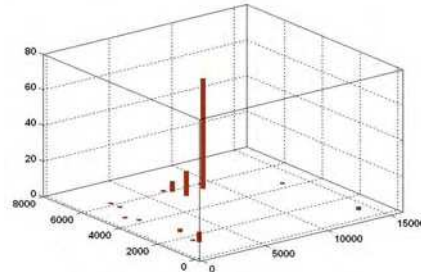
Fig. 14. Example of Map Building for Toshiba Science Museum.



(a) Localization error of ApriAttenda™ in the Toshiba Science Museum
Actually starting and end points are same, but odometry has different information.



(b) Random distribution of particles in the map for kidnapped recovery.



(c) Particle distribution after localization process.

Fig. 15. Example of Self-localization in Toshiba Science Museum.



Fig. 16. Scene of Experiment in Toshiba Science Museum.

7. To Realize Life Support Robots

The functionality of ApriAlpha™ as an information interface is upgraded so as to correspond to the needs in an environment where networked information appliances will be used. The person-following robot is intended to achieve coexistence with humans in terms of motion, not in terms of information. From now on, technology development will be continued with a view to performing actual tasks.

7.1 Enhancement by information network technology

To support everyday life, integration with an information network or sensor network is the key to offering services superior to those that an individual robot acting alone is capable of providing, as illustrated in Fig. 3 and discussed in Chapter 3. Such a robot is an active device in the network environment. Robots also need external sensors mounted in the environment. Thus, robots and information network need to collaborate with each other. In the near future, home-use robots are expected to be the core of home network systems. Also, home-use robots are expected to progress from those capable of simple functions such as cleaning to networked robots, and finally to multifunctional robots, as indicated in Fig. 17. Here, the same progression is considered not only for home-use robots but also for general service robots.

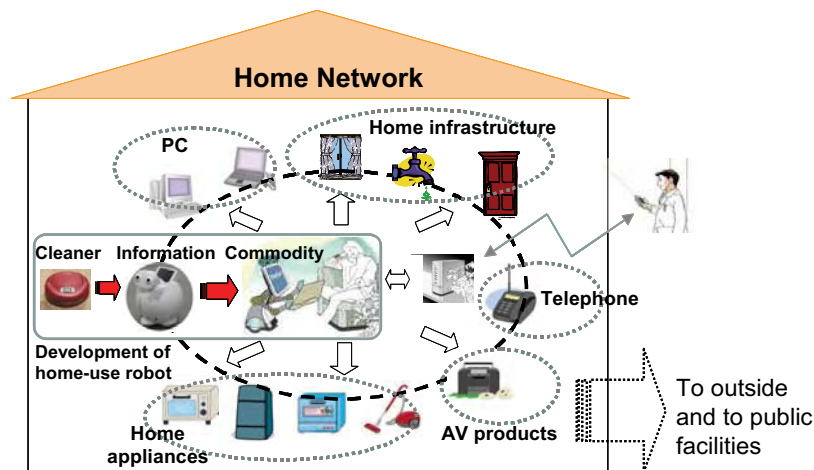
7.2 Enhancement by environmental design

Once robots enter widespread use in everyday life, it will be necessary not only to improve their ability, but also to improve their working environment. Information networking has become widespread. There is a need to define a common, friendly interface for both human and robots,

e.g. a barrier-free house. Here, we call it a universal design with robots (UDRob™) (Matsuhira et al., 2004), (Wada, 2004). Designing the environment both from the information side and the physical side, the realization of the life support robots will be facilitated as shown in Fig. 18. UDRob™ is based on the hypothesis that “an environment that is easy for anybody to use is also one in which it is easy for a robot to move around in and work.” The improvement of the environment will greatly contribute to the performance of work by the robot, for example, by providing a physical environment where the robot can move easily with few disturbances, common mechanical interface for handling objects by gripping, and an information environment where the robot knows where it is and what is there. Although it is difficult to improve a robot’s performance in terms of mobility, handling and visual recognition, these items are covered by environmental design, including the object interface.

Mobility - if a wheelchair can move in the house a robot can also move. Handling - if the gripping interface which anyone can easily grasp is included in the concept of universal design, the robot can also grasp it easily. Vision - if a marker is designed clearly for a person who has poor eyesight, the robot can also recognize the marker easily. These are the concepts of UDRob™. Of course, RFID tags help provide information to the robot. Thanks to RFID tags, the robot knows the contents of a cabinet even if the robot cannot open it. But this method is unsuitable for humans because tags’ information is invisible to human. So the UDRob™ concept includes tags with visible markers, such as simple icons of stationery, medicine, or tools, which are easily understood by both humans and robots.

The robot’s recognition ability by means of onboard computers is limited. The embedded knowledge in the environment via the network system helps the robot to execute tasks in daily life. For example, if an RFID tag has an object’s properties such as weight, shape, length and knowledge of how the object should be handled by a robot, the robot can execute the task through the RFID tag. Many sensors are distributed in the environment to enhance robot performance through collaboration via information networking. This is why networked robots (Hagita et al., 2005) and ubiquitous robots (Kim et al., 2005) are currently under development.



Robots support everyday life as a cornerstone of the network

Fig. 17. Development of home use robot as the core of the home network.

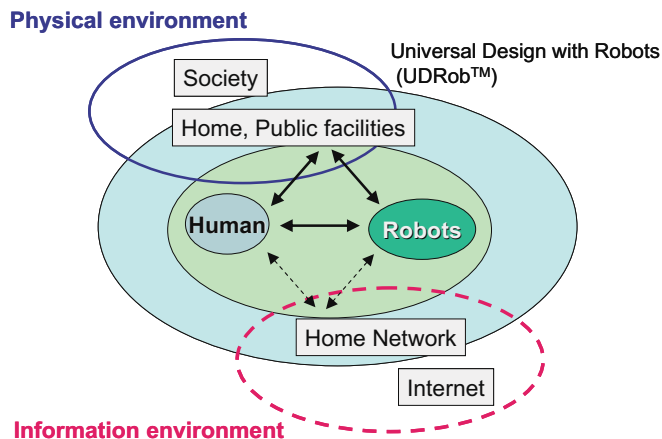


Fig. 18. Infrastructure improvement between information and physical environments.

8. Conclusions

The ongoing development of life support robots is presented by introducing the newly developed sharp ear robot ApriAlpha™ V3 and the person-following robot ApriAttenda™ from the viewpoint of human interfaces and mobile intelligence. In the future, by making full use of advanced network technology, home-use robots are expected to be at the core of home network systems and the widespread adoption of robots in everyday life is expected to be greatly facilitated by improvements in their working environment. Showing the concept of UDRob™, the environmental design including objects should be considered from the perspectives of both robots and humans. To realize life support robots, it is important to demonstrate what the robot can do in terms of actual tasks. The authors believe intelligent robots are the next technology whose development will decisively change the way people live.

Other important issues are standardization of the robot's interface and safety problem. The activities of RSi (Robot Service Initiative) contribute to the common interface of information service such as weather forecast or news for service providers (Narita, 2005). In the OMG (Object Management Group) meeting, such an interface is also discussed (Kotoku, 2005), (Mizukawa, 2005). On the safety problem, the discussion on safety is an ongoing topic from the Aichi Expo. These activities will be fruitful in near future.

9. Acknowledgements

ApriAlpha™ V3 and ApriAttenda™ were developed through entrustment by the New Energy and Industrial Technology Development Organization (NEDO). The authors are grateful to everyone involved in the NEDO prototype robot project. Map build and localization were carried out as a 100-day project of the Centers of Excellence (COE) Program at Tokyo Institute of Technology in collaboration with Toshiba Corporation. Especially, PhD students Mr. Jiang Zhu and Mr. Nafis Ahmad yielded good results. Furthermore, these developments were carried out by the big efforts of our colleagues Robotics Group people of Humancentric Laboratory in Corporate R&D center, Toshiba Corporation. The authors express our thanks for all.

10. References

- Yoshimi, T.; Matsuhira, N.; Suzuki, K.; Yamamoto, D.; Ozaki, F.; Hirokawa, J. and Ogawa, H. (2004). Development of a Concept Model of a Robotic Information Home Appliance, ApriAlpha, *Proc. of the 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2004)*, TA1-G4, Sendai, Japan, Sept. 28 – Oct. 2, 2004.
- Matsuhira, N.; Ozaki, F.; Ogawa, H.; Yoshimi, T. and Hashimoto, H. (2005). Expanding Practicability of ApriAlpha in Cooperation with Networked Home Appliances, *Proc. of IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO '05)*, Nagoya, Japan, June 2005.
- Kim, B.K.; Miyazaki, M.; Ohba, K.; Hirai, S. and Tanie, K. (2005). Web Services Based Robot Control Platform for Ubiquitous Functions, *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA2005)*, Barcelona, Spain, April 2005.
- Matsuhira, N. and Ogawa, H. (2004). Trends in Development of Home Robots Leading Advanced Technologies, *Toshiba Review*, Vol.59, No.9, pp.2-8, 2004.
- Nakadai, K.; Hidai, K.; Mizoguchi, H.; Okuno, H. G. and Kitano, H. (2003). Real-Time Human Tracking by Audio-Visual Integration for Humanoids, *JRSJ* Vol.21, No.5, pp.57-65, 2003. (in Japanese)
- Brandstein, M. (1999). Time-delay estimation of reverberated speech exploiting harmonic structure, *J. Acoust. Soc. Am.*, 105, 5, pp.2914-2929, 1999.
- Asano, F. (2004). Separation of Sound, *Journal of SICE*, Vol.43 No.4 325-330, 2004. (in Japanese)
- Brandstein, M. and D. Ward, D. (2001). *Microphone Arrays*, Springer Verlag, Berlin, 2001.
- Okazaki, A. (2000). *Hajimete no Gazo-shori*, Kyogyo-chosa-kai, 2000. (in Japanese)
- Amada, T. and Kanazawa, H. (2004). Microphone Array Technique for Speech Recognition, *Toshiba Review*, Vol.59, No.9, pp.42-44, 2004 (in Japanese) http://www3.toshiba.co.jp/pc/lalavoice/index_j.htm (in Japanese)
- Sato, H.; Hashimoto, H.; Ozaki, F. and Suzuki, K. (2005). A Versatile Multimedia Front-end Processing Board for Handling Visual and Acoustic Signals Concurrently, *Proc. of the 2005 JSME Conference on Robotics and Mechatronics (Robomec2005)*, 2P1-N-053, Kobe, Japan, June 2005. (in Japanese). <http://horb.a02.aist.go.jp/horb-j/> (in Japanese)
- Matsuhira, N.; Ogawa, H.; Yoshimi, T. and Mizoguchi, H. (2005). Development of Life Support Robots that Coexist in Harmony with People, *Proc. of the 36th International Symposium on Robot (ISR2005)*, TU 4H-5, Tokyo, December 2005.
- Schlegel, C. et al. (1998). Vision Based Person Tracking with a Mobile Robot, *Proc. of the 9th British Machine Vision Conference*, pp.418-427, 1998.
- Sidenbladh, H. et al. (1999). A Person Following Behaviour for a Mobile Robot, *Proc. of the 1999 IEEE Int. Conf. on Robotics and Automation (ICRA1999)*, pp.670-675, Detroit, USA, May 1999.
- Kwon, K. et al. (2005). Person Tracking with a Mobile Robot using Two Uncalibrated Independently Moving Cameras, *Proc. of the 2005 IEEE Int. Conf. on Robotics and Automation (ICRA2005)*, pp.2888-2894, Barcelona, Spain, April 2005.
- Cielniak, G. et al. (2005). Quantitative Performance of a People Tracking System on a Mobile Robot, *Proc. of ECMR05 (2nd European Conf. On Mobile Robots)*, Sep. 2005.
- Kobilarov, M. et al. (2006). People tracking and following with mobile robot using an omnidirectional camera and a laser, *Proc. of the 2006 IEEE Int. Conf. on Robotics and Automation (ICRA2006)*, pp.557-562, 2006.

- Ozaki, F. (2003). Open Robot Controller Architecture, Workshop on Middleware Technology for Open Robot Architecture, 2003 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics (AIM2003), 2003.
- Hirai, N. and Mizoguchi, H. (2003). Visual Tracking of Human Back and Shoulder for Person Following Robot, *Proc. of the 2003 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics (AIM2003)*, pp.527-532, 2003.
- Zhu, J.; Ahmad, N.; Nakamoto, H. and Matsuhira, N. (2005). Map Building and Localization of Home Robot using Feature Matching and Monte Carlo Approach, *Proceedings of the 2006 JSME Conference on Robotics and Mechatronics*, Waseda, Japan, May 26-28, 2006.
- Thrun, S. (2002). Robotic Mapping: A Survey. Exploring Artificial Intelligence in the New Millennium, Morgan Kaufmann, 2002.
- Latecki, L. J. and Lakamper, R. (1999). Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution, *Computer Vision and Image Understanding*, 73, 1999, p 441-454.
- Thrun, S.; Fox, D.; Burgard, W. and Dellaert, F. (2001). Robust Monte Carlo Localization for mobile robots, *Artificial Intelligence*, 128(2001), p99-141.
- Dellaert, F.; Fox, D.; Burgard, W. and Thrun, S. (1999). Monte Carlo Localization for Mobile Robot, *Proc. of the 1999 IEEE Int. Conf. On Robotics & Automation*, Detroit, Michigan, 1999.
- Yoshimi, T.; Nishiyama, M.; Sonoura, T.; Nakamoto, H.; Tokura, S.; Sato, H.; Ozaki, F.; Matsuhira, N. and Mizoguchi, H. (2006). Development of a Person Following Robot with Vision Based Target Detection, *Proc. of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2006)*, Beijing, China, Oct. 9-15.
- Ueno, K.; Hasegawa, T.; Ohsuga, A. (2004). Agent Enabling Robots to Join Ubiquitous World, *Toshiba Review*, Vol.59, No.9, pp.45-48, 2004. (in Japanese)
- Morioka, Y.; Aizu, H.; Tajika, Y. (2004). Framework for Cooperation between Networked Appliances and Home Robot, *Toshiba Review*, Vol.59, No.9, pp.49-52, 2004. (in Japanese)
- Hagita, N.; Kuwabara, K.; Tokuda, H.(2005). Introduction to Network Robots, *Proc. of Workshop on Network Robot Systems, IEEE International Conference on Robotics and Automation (ICRA2005)*, Barcelona, Spain, pp.1-4, (2005-04).
- Wada, T. (2004). Designing Home Robots, *Toshiba Review*, Vol.59, No.9, pp.15-19, 2004. (in Japanese)
- Narita, M.; Makiko Shimamura; M. (2005), A Report on RSi (Robot Services Initiative) Activities, 2005 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO'05), WAO-1-4, Nagoya, Japan, June 2005.
- Kotoku, T. et al. (2005), "Robot middleware and its standardization in OMG", Kotoku, T. et al. (2005), Robot middleware and its standardization in OMG, in *Proc. of The 2nd International Conference on Ubiquitous Robots and Ambient Intelligence (URAmI2005)*, pp.116-119, Daejeon, Korea, November 2005.
- Mizukawa, M. (2005), Robot Technology (RT) Trend and Standardization, 2005 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO'05), WAO-1-1, Nagoya, Japan, June 2005.



Mobile Robots: towards New Applications

Edited by Aleksandar Lazinica

ISBN 978-3-86611-314-5

Hard cover, 600 pages

Publisher I-Tech Education and Publishing

Published online 01, December, 2006

Published in print edition December, 2006

The range of potential applications for mobile robots is enormous. It includes agricultural robotics applications, routine material transport in factories, warehouses, office buildings and hospitals, indoor and outdoor security patrols, inventory verification, hazardous material handling, hazardous site cleanup, underwater applications, and numerous military applications. This book is the result of inspirations and contributions from many researchers worldwide. It presents a collection of wide range research results of robotics scientific community. Various aspects of current research in new robotics research areas and disciplines are explored and discussed. It is divided in three main parts covering different research areas: Humanoid Robots, Human-Robot Interaction, and Special Applications. We hope that you will find a lot of useful information in this book, which will help you in performing your research or fire your interests to start performing research in some of the cutting edge research fields mentioned in the book.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Nobuto Matsuhira, Hideki Ogawa and Hiroshi Mizoguchi (2006). Research and Development for Life Support Robots that Coexist in Harmony with People, Mobile Robots: towards New Applications, Aleksandar Lazinica (Ed.), ISBN: 978-3-86611-314-5, InTech, Available from:

http://www.intechopen.com/books/mobile_robots_towards_new_applications/research_and_development_for_life_support_robots_that_coexist_in_harmony_with_people

INTECH
open science | open minds

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

© 2006 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.