

Sensor network for structuring people and environmental information

S. Nishio, N. Hagita, T. Miyashita, T. Kanda, N. Mitsunaga and M. Shiomi
ATR Intelligent Robotics and Communication Laboratories
Japan

T. Yamazaki
National Institute of Information and Communications Technology
Japan

1. Introduction

Recently numbers of studies have been performed to make robots behave and interact with people more friendly and effectively. For example, nonverbal human-like behaviors such as eye-contact or maintaining suitable distance are known to be important in forming social relationships with people (Yoshikawa et al., 2006; Huettenrauch et al., 2006; Walters et al., 2005). To provide services effectively and behave humanlike, robots needs to obtain detailed information on where people exist or what they have been doing. Moreover, robots can provide services to people in two different ways: people find and approaches robots, or robots approach person that needs assistance. When we look for services in the current world, shops, information kiosks or wall maps in public spaces are typical examples of the first type. In order to receive the service, people need to find out where the service is provided and have to approach the place. On the contrary, there had been few examples for the second form. Lifeguards at the swimming pool may be a rare example. When someone drawn is found, s/he dives in to help the person. Although the latter form is much more convenient, why can they be rarely seen in our everyday life? This is because services of this form need monitoring and recognition of people requirements which is still a difficult task. However, with the progress in ubiquitous sensing systems, this convenient form of service is becoming real.

Several research projects have been running under this intent. One typical example is the *aware home* (Kidd et al., 1999)⁰, where sensors are installed in the residence and making the home a smart, intelligent space for supporting life. However, due to the high cost required in building such system on individual houses, it is hard to imagine every home having these facilities. Instead, public space such as shopping malls, stations or airports are more likely to be equipped with these sensing systems. However, public spaces show typical natures that cannot be seen in individual homes such as: 1. they are typically crowded with people, 2. the area is wide, and 3. most of the places are designed for certain purpose. The first two shows that sensing is much more challenging in these places. But by the last feature, there is a

possibility that robots may perform better utilizing the background knowledge specific to each environment.

Several robotic projects in public spaces have been performed, that provides guidance or shop recommendations. For example, in the *Robotics* project (Jensen et al., 2005), numbers of mobile robots were placed in the national expo for 5 months. Here the robots cooperated with each other, in order to locate themselves or find persons that need help, using the predefined knowledge on several areas in the expo pavilion. However, as you can experience by walking in a crowded station, it is not so easy to recognize what's happening in a place not nearby. With the help from the sensors equipped in the environment, this can be made easier. Also, when considering robotic services in real, knowledge on spatial area shall be better acquired mostly automatically, in order to follow the changes in the environment.

Also from the aspect of developing robotic services, raw position information such as coordinates may not be sufficient. For example, when making a navigation services that find a person in lost, what is required most is the status or behavior of the person, not the coordinate s/he resides. This is also useful for providing services appropriate in each context or situation the target person is in. The same is also true for spatial knowledge as described previously. The abstract, symbolic information is better suited for easily describing scenarios for human-robot interaction or robotic services (Fig. 1). Several research efforts have begun to obtain this abstract information from sensing results (Kanda, et al, 2007; Gerkey et al., 2003).

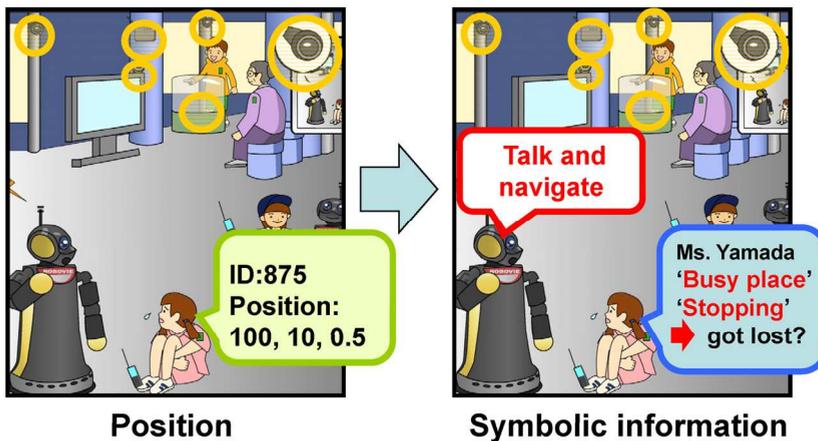


Fig. 1. Human-robot interaction with and without symbolic information.

Another issue related to the development of robotic services lies on their high cost in preparing a robotic system. Currently, developers need to prepare mostly every element by themselves that are required in composing a robot, such as localization, human identification or speech recognition. What is required is a common framework or platform system that can be used for building robots easily. Several projects have been conducted to prepare such common frameworks (Kranz et al., 2007; Thompson et al., 2007), but there seems to be nothing standard yet.

Our intention is to prepare an infrastructural system that senses human behavior and spatial nature in the environment. This system will provide robots information obtained from the environment in a standard manner. In past, factories have been designed so that industrial robots can perform effectively, albeit of their limited sensing ability. However, it is not likely that public spaces are going to be built in the same way suitable for robots. Thus, instead of structuring the environment as in factories, by utilizing the advanced sensing and pattern recognition techniques, we intent to make the information extracted from the environment to be structured and to be easily processed in robotic systems. Our server system is based on four-layer architecture, the *Structuring model of Environmental Information* (Figure 2). Measurements from sensors are first processed in the *sensor layer* and are then integrated to produce uniform positioning results in the *segment layer*. At the same time, positions are accumulated and processed to produce symbolic information on human behavior and space in the *primitive layer*. Both the coordinates and symbolic information produced are provided to robots, enabling robotic services to be designed and run based on both numerical and symbolic descriptions. In past, there have been several attempts to combine robotic services with environmental sensing systems (Thompson et al., 2007). There were also studies on sensing systems that produce abstract symbolic information (Jensen et al., 2003). But as far as we know, no such integrated system exists which can be used in the real world and which provides standard access methods for robotic services in general.

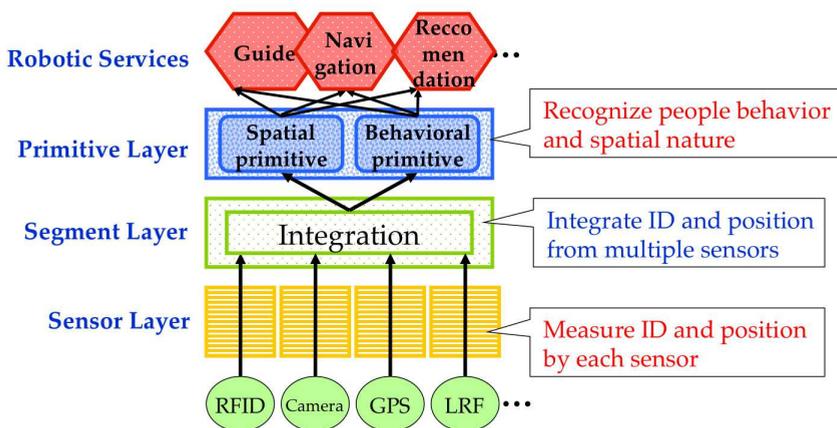


Fig. 2. Four-layer architecture of structured environmental information

In the following, we will first describe the first three layers in our server platform: robust sensing methods for human positioning in real environments, integration of positioning results by individual sensors and methods for recognizing human behaviors and spatial nature. Based on the proposed architecture and these technologies, we have built two trial fields in public spaces and have been providing these to researchers for testing purpose. In the last section, we describe these two trial fields and briefly show some robotic service experiments held using the proposed server features.

2. The Sensor Layer: Measuring Human Position

Human positioning is one major field of research. Various methods and sensing devices have been actively developed (Hightower & Borriello, 2001). One typical and frequently used device is the Global Positioning System (GPS). Many cellular phones nowadays are equipped with GPS receivers, and are widely used for location-based services such as navigation or restaurant recommendation. Although these systems are sufficient for rough navigation, the accuracy of positioning is still not enough for their use in crowded public spaces. Moreover, for robotic use, much accuracy is required for approaching a person or for interacting movements such as making gestures or keeping eye contacts.

Finding people by on-board sensors installed on robots are another field that is actively studied (Gockley et al., 2007; Fransen et al., 2007). However, this is effective only after the robots have approached close enough to the person in target. Thus, the issue still remains on how to find the person who needs help in the public environment. On the contrary, by installing and utilizing sensors in the environment, we can not only obtain a more accurate observation, but also be able to degrade the cost required in robot development. By accessing the positioning results from the environment, developers can get rid of positioning part and can concentrate on service development.

Environmental sensing has been long studied in the field of computer vision or pattern recognition. Unfortunately, most of these methods do not perform robustly in the real environment, where people behavior or conditions such as lightening is not under control. In order to implement a robust real-time human positioning, we decided to obtain and integrate results from four sensors of different modalities.

2.1 LRF tracking of human position

Laser range finder (LRF) is one of the most commonly used sensors in the field of robotics. It can quickly and precisely scan the environment. However, in a public environment where many obstacles or person exists, the sensing is often obscured by occlusions. Our solution was to perform tracking using a network of multiple LRFs. Each LRF was mounted at average human torso height, approximately 85 cm high from the ground. A particle filter is used to track the position and velocity of each people, by using a model of a torso and two arms (Dylan et al., 2007). Fig. 3 shows the raw sensor data during the walking motion and how the model is fitted, and the results of tracking in real field.

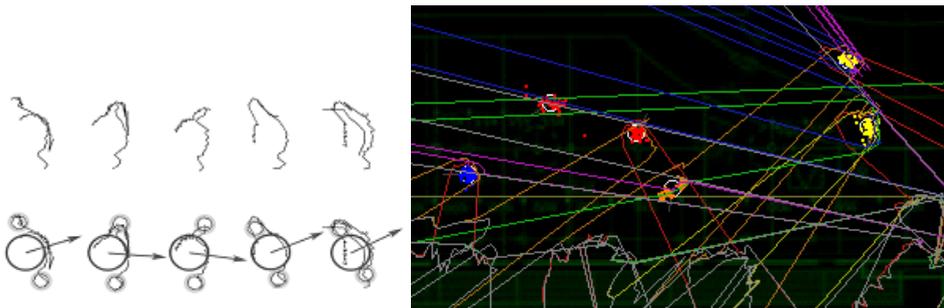


Fig. 3. Estimating pedestrian arm and torso movement from LRF output.

2.2 Multi-camera and LRF based human tracking

One of our goals for human positioning was to obtain the position of the human head. The head information is required for robots to effectively communicate with people, such as having conversation or maintaining eye contacts. This is done by integrating a head tracking system using multi-viewpoint images captured by a set of distributed cameras and the LRF based tracking system described above. Based on the 2D torso position estimation by the LRF tracker, head position is estimated by applying a human head model to edge images obtained (Matsumoto et al., 2008). As described previously, our purpose is to settle up this positioning system in a semi-outdoor public space, where there exists numbers of obstacles and the lighting condition frequently changes. This is a crucial condition for applying image processing techniques. Our solution was to combine both vision-based method and non-vision sensing method, in order to obtain detailed positioning estimation of people in the real environment. Fig. 4 shows an example of the tracking results.



Fig. 4. Tracking results by the combination of multi-camera and LRF. The grey circles indicate head tracking result and the yellow dot clouds show hypotheses. The green dots on the bodies indicate base LRF tracking results.

2.3 Human tracking by RFID tag and GPS

Radio-frequency identification (RFID) tags and Global Positioning System (GPS) receivers carried by people are also used for obtaining their IDs and positions. As for the RFID tags, position estimation are done by a particle filter using the measured signal strength of active RFID tags (Shiomi et al., 2005). Although the positioning result obtained by RFID tags and GPS receivers are low both in accuracy and precision, these estimations are useful for associating identities (ID) to precise positioning results obtained from LRF or camera trackers. Moreover, GPS receivers can be used to track people in the wide outdoor environments where other sensors are not suitable. As people often move between indoor and outdoor in public spaces, continuous and seamless positioning in both fields are useful for keeping track of each person, especially for capturing their intention.

3. The Segment Layer: Integration of Position Estimations

In the second layer, the segment layer, estimation results from individual trackers in the first layer are integrated to produce a final estimation of each person in the environment. That is, the 2D / 3D positions and IDs are integrated to compose a final 3D position of a human head accompanied with ID information (

Fig. 5). We currently use a simple nearest-neighbor method to integrate the resulting position and ID measurements, making concerns for error distribution of each of the initial estimations. Table 1 shows the positioning accuracy in the two trial fields and Fig. 6 shows the result of the integrated position estimation. We have found that our system is able to continuously track more than 20 people in real time. This was tested in real environments where various conditions such as number of people or illumination changes frequently.

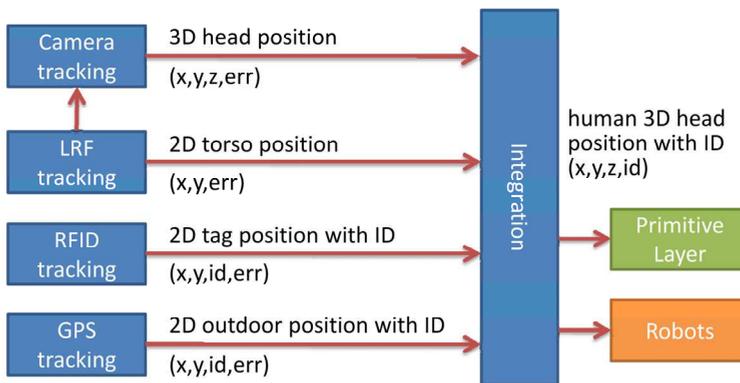


Fig. 5. Integration of individual tracking results in the segment layer. By combining estimation results from the four trackers, the segment layer outputs human head 3D positions with IDs.

Sensor type	Keihanna environment		UCW environment	
	sensor #	RMS	sensor #	RMS
LRF (x,y)	6	49 mm	6	58 mm
Camera (x,y,z)	16	13 cm	16	12cm
RFID (x,y)	17	287 cm	9	249 cm
GPS (x,y)	-	10 m	-	10-100 m

Table 1. Accuracy of position estimation in the two trial fields.

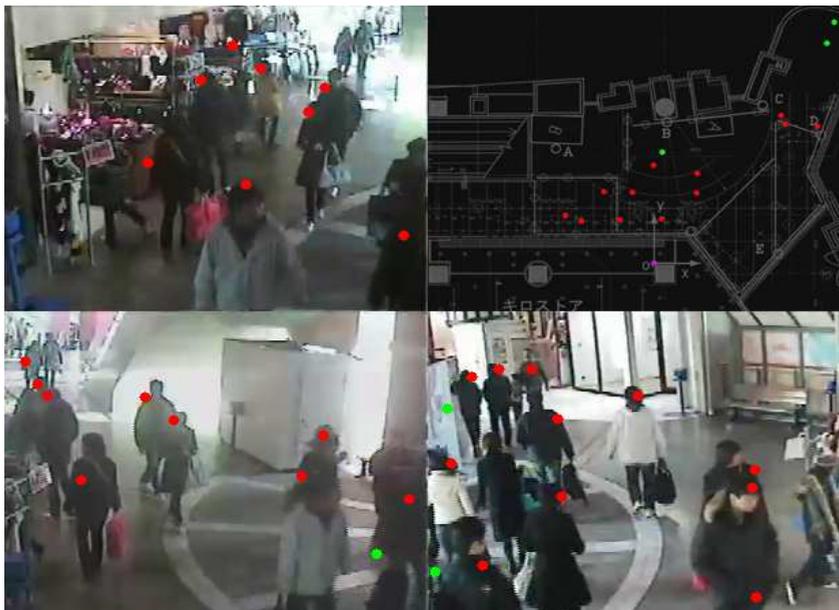


Fig. 6. Example of the integrated position estimation. The red dots show the head estimation results and the green dots show the torso estimation where head was not detected or when person was out of camera tracking area.

4. The Primitive Layer: Recognition of Human Behaviors and Spatial Natures

In order to efficiently design robotic services suitable for each environment, developers often need symbolic descriptions on how each person is behaving and what kind of attribute each area in the environment holds. Also, by knowing the typical behavioral patterns of people in the environment, robots will be able to provide context-aware services suitable for each person. These recognitions can be done by accumulating and learning people trajectories in the environment.

By obtaining symbolic information based on each environment, developers can easily adapt their own robotic services to various places. We call these types of symbolic information *primitives*. Using the measured human position history, the platform server calculates and provides the primitives to the robots. In the following, the three types of primitives are described: *local behavioral primitive*, *global behavioral primitive* and *spatial primitive*.

4.1 Local behavioral primitive

One sort of information useful in designing robotic services is the short-time trend in people movement, such as walking or running. This kind of symbolic information is called *local behavioral primitive*. Recognition of local behavioral primitives is performed by a simple classifier based on support vector machine (SVM) using the obtained people trajectories. Features such as the velocity or degree of curvature of the trajectory are used for training the classifier.

In the current implementation, two types of local behavioral primitives are computed, each with its own classifier. The first is for the *style* of walking (straight, turning right/left, wandering, u-turn). The durations of trajectories used were 5.1 seconds. This means that observation of at least 5.1 seconds is required to decide the style of walking. We tested the classification accuracy by leave-one cross validation test which resulted in 82.5 % accuracy. Another is about the *type* of walking (running, busy/idle walking, stopping, waiting). We trained the classifier with 166 labeled data. The durations of trajectories used are 4.9 seconds. Leave-one cross validation test resulted in 87.8 % accuracy. Fig. 7 shows recognition results in a real environment.

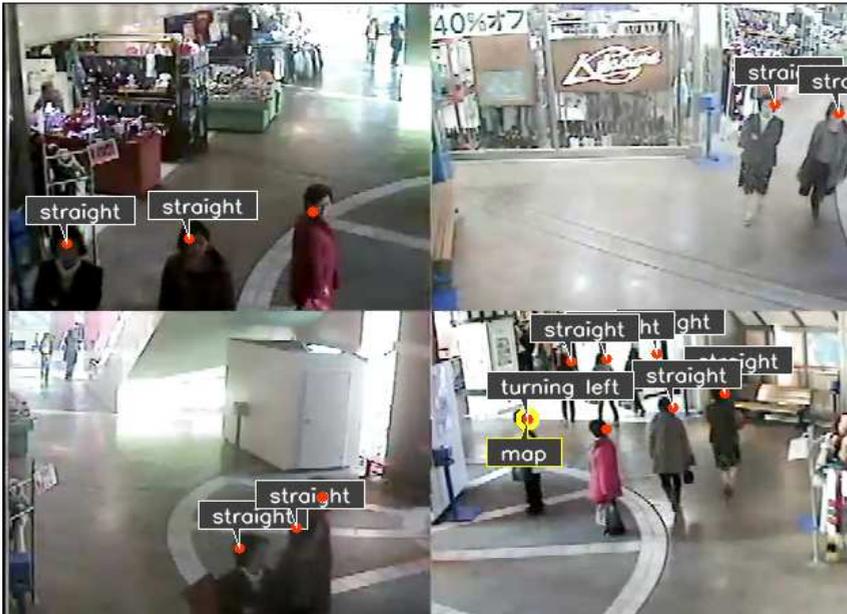


Fig. 7. Primitive recognition results. The box above a person indicates the detected local behavioral primitive (style of walking) and the one below shows the spatial primitive.

4.2 Global behavioral primitive

Global behavioral primitive shows what kind of typical visiting pattern the person is following. The difference in calculating between global and local behavioral primitives mainly exists on the length of the trajectory required. However, as the length of the trajectory increases, we can assume that the pattern holds more abstract, semantic information on the intent of each person.

In an experiment was held in a science museum, where visitor trajectories were collected through RFID tag tracking (Kanda et al, 2007). By analyzing more than 8,000 moving patterns acquired, 5 typical visiting patterns were extracted. As we can think that these patterns, *i.e.* global behavioral primitives, indicate topic of interest for each visitor, robots can easily provide services or information adapted to each person.

4.3 Spatial primitive

A *spatial primitive* indicates the nature of an area. How people have been typically behaving in the area, or what kinds of events have frequently occurred there in the past. This can be manually specified by system designers. Fig. 7 shows an example of manually specified spatial primitive. Here, the 'front_shop' indicator shows that the person in target is in the front of a shop. Otherwise, spatial primitives can be calculated based on visitor trajectory histories as global behavioral primitive (Kanda et al, 2007).

5. The Kansai Environment: Trial Field for Robotic Services

Using these technologies, we have built two trial fields for robotic services: the *Kansai Environment*. One is located at the entrance of an office building (NICT Keihanna building) and another in a shopping mall (Universal City Walk Osaka, UCW for short), both in the Kansai district of Japan. These two locations were chosen due their difference in nature of the place and in the people behavior. The first location is mainly aimed for business or exhibition purpose and the second is for shopping. As the underlying systems in both fields are common, here we will mainly describe about the UCW field. The measured positioning accuracy for both fields is shown in Table 1. The only difference was, due to high buildings in the surroundings, we were not able to use GPS in the UCW field.

The shopping mall where the UCW field is settled is located between a train station and an amusement park (Universal Studio Japan). Every day thousands of people walk through the mall for shopping, dining or just for the sake of passing for their way to the park. Fig. 8 shows sensor alignments in UCW trial field.



Fig. 8. Sensor alignments in the Kansai Environment (UCW).

These trial fields were intended to be a common environment for testing various robotic service implementations, assisted by the environmental server functionalities. Thus, a common, standard interface was required for the server system. For this purpose, we have used the Robotic Localization Service standard (Nishio, et al., 2009). The access interface,

including the outputs from the segment and primitive layer, was implemented to follow this standard. Also, a repository server was prepared to exchange the various metadata definition such as 3D coordinate reference systems or primitive symbolic coordinate systems used for the output.

5.1 Basic robotic service in the Kansai Environment

In order to see the effectiveness of the platform server system, we have conducted a robotic service experiment at the UCW field. A humanoid robot (Robovie-III) was used. This robot is equipped with several sensors, such as a LRF, an omni-directional camera and ultrasound distance sensors. However, none of them were used in this experiment. Instead, the robot completely relied on the information provided from the platform server and never used its own sensing system for interacting with people.

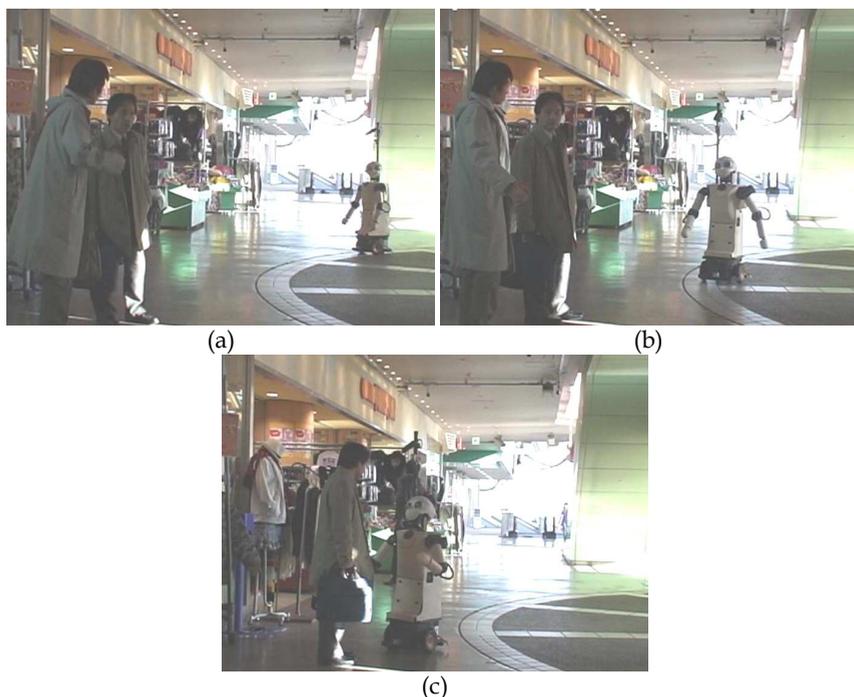


Fig. 9. Scenes from the robotic service experiment using the Kansai Environment (UCW). (a) Server system detects a person standing in front of a shop. (b) Robot starts approaching the people based on the positioning results from the server. (c) The robot starts shop guidance service, while keeping eye contact with the target person. The robot's eye contact motion is also performed based on the head position received from the server.

Fig. 9 shows a sample sequence from the experiment. When the platform server detects a person with a local behavioral primitive 'stopping' and with a spatial primitive 'shop front' or 'map', Robovie starts approaching the person using the estimated position from the platform server. After the robot reaches nearby the person, it begins providing route

guidance or shop information, depending on the spatial primitive it has received. While providing information, Robovie also makes gestures and keeps eye contact with the person, using the head position estimation given from the environmental server.

Although the robotic service designed is quite a simple one, the fact that it worked out quite stably and effectively seems to indicate the validity of our idea in decreasing the robot development cost. At the same time, this also shows the possibility of developing robotic applications independent to the sensing ability of each robot.

5.2 Using the Kansai Environment

From June 2008 to March 2009, we made the Kansai Environment available to public. Interface specification was released on the web, and robot researchers were invited to use the environment for trying the environmental server functionality with their robots free of charge. Since its release, eight research organizations or projects used the field.

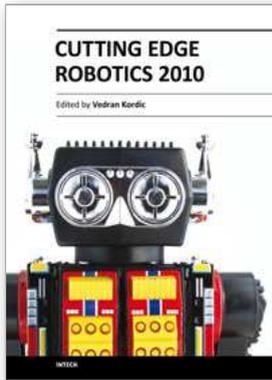
6. Acknowledgement

The authors would like to thank Sumisho Urban Kaihatsu Co. Ltd. and Robot Laboratory for their corporation in constructing the Universal City Walk Osaka trial field. This work was supported in part by the Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

7. References

- Fransen, B.; Morariu, V.; Martinson, E.; Blisard, S.; Marge, M.; Thomas, S.; Schultz, A. & Perzanowski, D. (2007). Using Vision, Acoustics, and Natural Language for Disambiguation, *Proceedings of ACM/IEEE Int. Conf. Human-robot interaction*, pp. 17-24
- Glas, D.; Miyashita, T.; Ishiguro, H. & Hagita, N. (2007). Laser Tracking of Human Body Motion Using Adaptive Shape Modeling, *Proceedings of IEEE/RSJ Int. Conf. Intelligent Robots and System*, pp. 602-608
- Gockley, R.; Forlizzi, J. & Simmons, R. (2007). Natural Person Following Behavior for Social Robots, *Proceedings of ACM/IEEE Int. Conf. Human-robot interaction*, pp. 17-24, 2007.
- Hightower J. & Borriello, G. (2001). Location Systems for Ubiquitous Computing, *IEEE Computer*, vol. 34, no. 8, pp. 57-66
- Huettenrauch, H.; Eklundh, K.; Green, A. & Topp, E. (2006). Investigating Spatial Relationships in Human-Robot Interaction, *Proceedings of IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 5052 - 5059
- Gerkey, B.; Vaughan, R. T. & Howard, A. (2003). The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems, *Proceedings of Int. Conf. Advanced Robotics*, pp. 317-323
- Jensen, B.; Tomatis, N.; Mayor, L.; Drygajlo, A. & Siegart, R. (2005). Robots Meet Humans - Interaction in Public Spaces, *IEEE Trans. Industrial Electronics*, Vol. 52, No. 6, pp. 1530-1546

- Jensen, B.; Philippsen, R. & Siegwart, R. (2003). Narrative Situation Assessment for Human-Robot Interaction, *Proceedings of IEEE Int. Conf. Robotics and Automation*, pp.1503-1508
- Kanda, T. ; Shiomi, M.; Perrin, L.; Nomura, T.; Ishiguro, H. & Hagita, N. (2007). Analysis of People Trajectories with Ubiquitous Sensors in a Science Museum, *Proceedings of IEEE Int. Conf. Robotics and Automation*, pp.4846-4853
- Kidd, C.; Orr, R.; Abowd, G.; Atkeson, C.; Essa, I.; MacIntyre, B.; Mynatt, E.; Starner, T. & Newstetter, W. (1999). The Aware Home: A Living Laboratory for Ubiquitous Computing Research, *Proceedings of Second Int. Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, pp.191-198
- Kranz, M.; Maldonado, A.; Rusu, R. B.; Hoernler, B.; Rigoll, G.; Beetz, M. & Schmidt, A. (2007). Sensing Technologies and the Player-Middleware for Context-Awareness in Kitchen Environments, *Proceedings of Fourth Int. Conf. Networked Sensing Systems*, pp. 179-186
- Matsumoto, Y.; Wada, T.; Nishio, S.; Miyashita, T. & Hagita, N. (2008). Scalable Multi-People Head Tracking for Robotic Services Combining Multiple Sensors, *Proceedings of 5th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI2008)*, pp. 581-586
- Nishio, S.; Lee, J.; Yu, W.; Sakamoto, T.; Noda, I.; Tsubouchi, T. & Doi, M. (2009). A New Standard for Robotic Localization Service, In: *Cutting Edge Robotics 2009*, Kordic, V., Lazinica, A. and Merdan, M. (Eds.), xxx-xxx, IN-TECH, ISBN978-3-902613-46-2, Vienna
- Shiomi, M.; Kanda, T.; Kogure, K.; Ishiguro, H. & Hagita, N. (2005). Position estimation from multiple RFID tag readers, *Proceedings of 2nd Int. Conf. Ubiquitous Robots and Ambient Intelligence*
- Thompson, S.; Kagami, S.; Sasaki, Y.; Nishida, Y.; Enomoto, T. & Mizoguchi, H. (2007). Combining Ubiquitous and On-Board Audio Sensing for Human-Robot Interaction, *Proceedings of ACM/IEEE Int. Conf. Human-robot interaction*, pp. 209-216
- Walters, M.; Dautenhahn, K.; Boekhorst, R.; Koay, K.; Kaouri, C.; Woods, S.; Nehaniv, C.; Lee, D. & Werry, I. (2005). The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment, *Proceedings of Int. Workshop Robot and Human Interactive Communication*, pp.347-352.
- Yoshikawa, Y.; Shinozawa, K.; Ishiguro, H.; Hagita, N. & Miyamoto T. (2006). The effects of responsive eye movement and blinking behavior in a communication robot, *Proceedings of IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 4564 - 4569.



Cutting Edge Robotics 2010

Edited by Vedran Kordic

ISBN 978-953-307-062-9

Hard cover, 440 pages

Publisher InTech

Published online 01, September, 2010

Published in print edition September, 2010

Robotics research, especially mobile robotics is a young field. Its roots include many engineering and scientific disciplines from mechanical, electrical and electronics engineering to computer, cognitive and social sciences. Each of this parent fields is exciting in its own way and has its share in different books. This book is a result of inspirations and contributions from many researchers worldwide. It presents a collection of a wide range of research results in robotics scientific community. We hope you will enjoy reading the book as much as we have enjoyed bringing it together for you.

How to reference

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

S. Nishio, N. Hagita, T. Miyashita, T. Kanda, N. Mitsunaga, M. Shiomi and T. Yamazaki (2010). Sensor Network for Structuring People and Environmental Information, Cutting Edge Robotics 2010, Vedran Kordic (Ed.), ISBN: 978-953-307-062-9, InTech, Available from: <http://www.intechopen.com/books/cutting-edge-robotics-2010/sensor-network-for-structuring-people-and-environmental-information>

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

© 2010 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.