AISL-TUT @Home League 2017 Team Description Paper

Shuji Oishi, Jun Miura,
Kenji Koide, Mitsuhiro Demura, Yoshiki Kohari, Soichiro Une, Liliana Villamar Gomez, Tsubasa Kato, Motoki Kojima, and Kazuhi Morohashi

Department of Computer Science and Engineering
Toyohashi University of Technology, Japan
http://www.aisl.cs.tut.ac.jp/robocup

Abstract. This paper introduces an overview of the AISL-TUT RoboCup @Home team. Although this is the first time for us to participate RoboCup @Home competition, we have been conducting research projects on various intelligent systems, such as intelligent robots, which can operate autonomously in complex real environments. Toward RoboCup@Home 2017, we've integrated our technologies on HSR developed by TOYOTA so that it can provide users with various services in daily life.

1 Introduction

Active Intelligent Systems Laboratory (AISL) at Toyohashi University of Technology (TUT) was founded in 2007. Since then, we have been conducting research in service robotics domain.

Person detection, tracking, and identification are important functions of service robots to properly interact with and support the user. To realize such functions, we developed, for example, LIDAR-based human detection and tracking [1], multiple feature-based human identification [2], Radio signal- and LIDAR-based person localization [3], person tracking with body orientation estimation [4, 5] and illumination invariant face recognition [6].

Planning a safe and efficient robot motion is also important, especially in a dynamic environment with many people. We developed an on-line motion planner [7], which combines a kinodynamic randomized search with a new estimated arrival time-based potential, and applied it to several person following robots[2, 3, 8]. For realizing a more comfortable robotic attendant, it is effective to choose its position adaptively according to the user's state such as standing, walking and sitting. We developed several methods including a viewpoint planner for choosing cost-effective positions for watching a freely-walking person [9] and an adaptive robot position planning for dealing with the change of the user's state [10].

Motion planning requires the understanding of the environment and many SLAM and place recognition methods have been adopted. We dealt with indoor environment recognition including an exploratory mapping[11] and a physical state monitoring[12].

Human-robot interaction (HRI) aspect is also important. We have been dealing with HRI issues in the context of human robot collaboration. Using a real humanoid robot, we have developed methods for collaborative remote object search [13], collaborative assembly [14], programming-by-demonstration for collaborative assembly[15], and robot-to-human teaching[16].

Toward RoboCup@Home 2017, we've integrated our technologies on TOYOTA HSR. RoboCup@Home competition has played an important role as an benchmark of service robots developed by robotics researchers all over the world, and provided opportunities to verify their performances in real environment. We would like to participate the competition in order to test and improve our technologies, and also to contribute to the RoboCup@Home league itself.

The rest of this paper is organized as follows. In section 2, the HSR hardware is introduced to show its specifications. In section 3, we describe our software implementations for carrying out tasks in the RoboCup@Home competition including human detection, path planing, object detection, and so on. Section 4 demonstrates how our HSR performs actual tasks in the real environment. Finally, in section 5, we conclude this paper and discuss the future work.

2 The HSR Platform

Human Support Robot (HSR)[17] has been developed by TOYOTA Motor Corporation as a platform for developing robot systems to assist, for example, the elderly or the disabled living alone. The hardware and software are well designed to provide services in daily life so that it can improve Quality of Life. This section gives a brief explanation of the HSR hardware.

2.1 Extendable Arm and Flexible Hand

HSR is designed to help people at home and has object handling functions such as fetching things and picking up stuffs on the floor. While a single arm equipped with HSR is folded tightly when HSR moves around, the arm extends along with its body for reaching out objects far from itself. A flexible gripper that has two fingers and a suction pad is attached to the arm that enables HSR to grasp objects or to lift light and thin items.

2.2 Ranging/Imaging Devices and Microphone array

A 2D laser scanner is mounted on HSR to measure the geometric structure of the environment. This allows HSR to create 2D maps and to safely navigate without collision. RGB-D camera and four digital cameras are also equipped so that HSR can perceive the surrounding environment for recognizing persons and objects.

Besides, a microphone array with four microphones is put on the top of HSR and is capable of localizing sound sources for speech recognition as described later.

3 Software

This section introduces the major software components essential for performing the fundamental tasks in RoboCup@Home. All of the software has been developed using Robot Operating System (ROS).

3.1 Person Tracking and Path Planning

For person tracking, we first extract leg-like clusters in laser range data by finding local minima in the distance histogram. Next, leg clusters are detected among them by calculating features, such as the length, the mean curvature, and the variance ratio by PCA, and classifying them with Support Vector Machine (SVM)[2]. These two steps are applied to each laser scan, and the robot tracks the target person's legs position using Unscented Kalman Filter (UKF).

In order to elaborate the state of the target person, we extend the state variables in UKF using torso shape data data so that it can estimate not only the position but also the body orientation of the person by comparing the input torso shape data with the model data, which is 360-degree torso shape data collected in advance[5]. Based on the pose and orientation information, we have developed an adaptive attendance robot that plans the appropriate attending position considering the change of the user's state[10]. We also extracts multiple features of the target person to follow including the clothing color/texture and the face to identify him/her correctly[2].

When HSR goes to a destination, it plans global and local paths to move without collision even in dynamic environments with the navigation package[18]. Moreover, we have been implementing our path planning algorithm[7] as an alternative local path planner to obtain in real time a shorter and safer path in a highly dynamic situations by utilizing a randomized path search.

3.2 Object Recognition

Object recognition is essential to handle stuffs at home. For example, when a robot working at home is asked to fetch something in the fridge, the robot should find the target object among the others in it.

We have implemented object recognition functions in HSR. For general object recognition, we have developed YOLO[19]-based object recognition system. For specific object recognition such as particular bottles and cans, we have developed a method based on image matching[20] using several features, such as Scale Invariant Feature Transform (SIFT) and Binary Robust Independent Elementary Features (BRIEF). The 3D position of each detected object is broadcasted so that HSR can grasp and handle them.

Combining the general and specific object detections, the HSR will be able to find the target object robustly and efficiently.

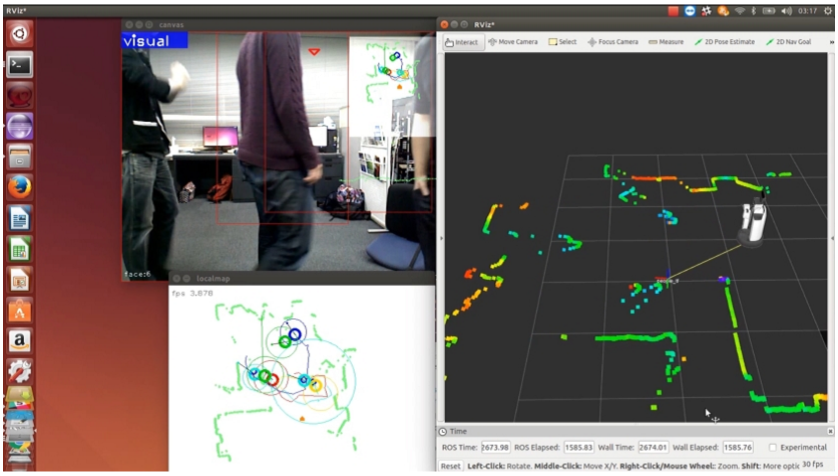


Fig. 1. Leg cluster detection

3.3 Speech Recognition

For recognizing human voice, we first use HARK[21] to perform sound source localization and separation. The separated sound is then recognized with Google Speech API[22]. Morphological analysis is also applied to the extracted texts voice with Stanford NLP[23] to obtain the meaning.

By extracting words and the words' relationships in the recognized speech, HSR understands what the target person means and execute the corresponding actions.

4 Experiments and results

4.1 Simple Tasks

People detection As described in 3.1, the HSR first extracts leg clusters in range data acquired with the 2D laser range finder by finding local minima in the distance histogram and detecting true ones with SVM[2] (Fig.1).The target person's position is tracked by UKF and published as ROS tf.

Object detection & manipulation Our YOLO-based object recognition system outputs a 2D position on a camera image of each object, such as apple, bottle, and oranges. Since the 3D positions are also available by referring to the corresponding depth image, the system publishes the tf. Specific object recognition based on image feature matching runs as well on HSR to find target objects (Fig.2).

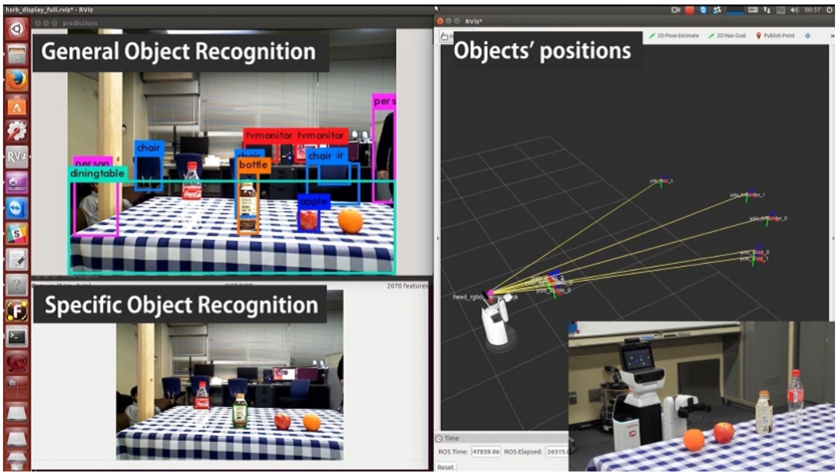


Fig. 2. General and specific object detection

Speech recognition Assuming beverage delivery task, HSR received an order from the operator and brought it. It recognized what the operator said with the system described in 3.3, and extracted words corresponding to the following categories: Action, Object, and Place (Fig.3).

4.2 Activity Recognition

We estimate the state of the target person for adaptive attendance to provide appropriate service. The robot estimates the body orientation of the target person using torso shape data[5] by extending the above-mentioned people tracking[2] (Fig.4(a)). Based on the position and orientation information, the robot judges the target person's state. We had a Hidden Conditional Random Fields (HCRF) learn the best discriminative structure from 5-frame consecutive features consisting of the walking speed, the distance and orientation to the nearest chair[10]. The HCRF successfully recognized the state of the person in real time based on same length of the consecutive features, and the robot moved to the appropriate position according to the state (Fig.4(b))

Note that the robot shown in the qualification video had 2 laser range finders at different heights to measure leg and torso shapes respectively, the proposed activity recognition method is available on HSR by using Xtion on it as an alternative of the upper laser range finder.

5 Conclusion

This paper describes the overall framework and major features of our technologies and implementations for RoboCup@Home 2017 SPL competition. The proposed human tracking and path planning methods allow HSR to follow the target

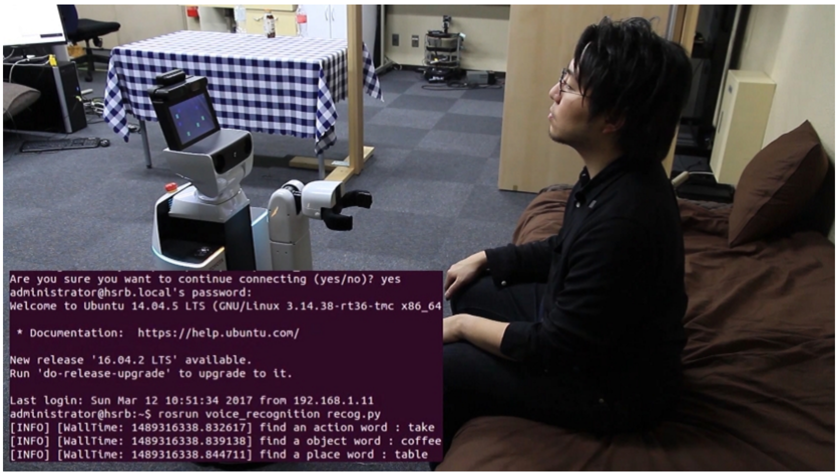
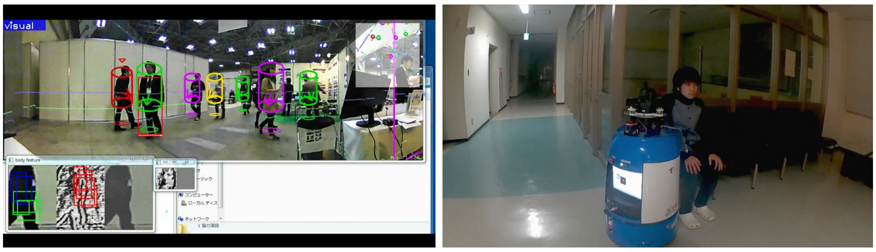


Fig. 3. Beverage delivery via speech recognition



(a) Body orientation estimation (b) Adaptive attendance

Fig. 4. Activity recognition

persons. The YOLO-based object recognition is also implemented for handling items in daily life. Besides, verbal communication function has been built in it that consists of sound source localization, speech recognition, and morphological analysis. On the other hand, the well-designed hardware on HSR such as the extendable arm, the flexible hand, and the omni-wheel enables HSR itself to execute a variety of tasks required in daily life based on the perception.

Toward the competition, we continue to improve the performance of each function to enable our HSR to deal with complex tasks and situations.

Bibliography

References

1. K.Kidono, T.Miyasaka, A.Watanabe, T.Naito, and J.Miura. Pedestrian recognition using high-definition lidar. In *2011 IEEE Intelligent Vehicles Symp*, pages 405–410, 2011.
2. K.Koide and J.Miura. Identification of a specific person using color, height, and gait features for a person following robot. *Robotics and Autonomous Systems*, 84(10):76–87, 2016.
3. K.Misu and J.Miura. Specific person tracking using 3d lidar and espar antenna for mobile service robots. *Advanced Robotics*, 29(22):1483–1495, 2015.
4. I.Ardiyanto and J.Miura. Partial least squares-based human upper body orientation estimation with combined detection and tracking. *Image and Vision Computing*, 32(11):904–915, 2014.
5. M.Shimizu, K.Koide, I.Ardiyanto, J.Miura, and S.Oishi. Lidar-based body orientation estimation by integrating shape and motion information. In *IEEE Int. Conf. on Robotics and Biomimetics*, pages 1948–1953, 2016.
6. B.S.B. Dewantara and J.Miura. Optifuzz: A robust illumination invariant face recognition system and its implementation. *Machine Vision and Applications*, 27(6):877–891, 2016.
7. I.Ardiyanto and J.Miura. Real-time navigation using randomized kinodynamic planning with arrival time field. *Robotics and Autonomous Systems*, 60(12):1579–1591, 2012.
8. M.Chiba J.Satake and J.Miura. Visual person identification using a distance-dependent appearance model for a person following robot. *Int. J. of Automation and Computing*, 10(5):438–446, 2013.
9. I.Ardiyanto and J.Miura. Visibility-based viewpoint planning for guard robot using skeletonization and geodesic motion model. In *IEEE Int. Conf. on Robotics and Automation*, pages 652–658, 2013.
10. Y.Kohari S.Oishi and J.Miura. Toward a robotic attendant adaptively bahaving according to human state. In *Int. Symp. on Robot and Human Interactive Communication*, pages 1038–1043, 2016.
11. Y.Okada and J.Miura. Exploration and observation planning for 3d indoor mapping. In *IEEE/SICE Int. Symp. on System Integration*, pages 599–604, 2015.
12. S.Kani and J.Miura. Mobile monitoring of physical states of indoor environments for personal support. In *IEEE/SICE Int. Symp. on System Integration*, pages 393–398, 2015.
13. K.Chikaarashi J.Miura, S.Kadekawa and J.Sugiyama. Human-robot collaborative remote object search. In *Int. Conf. on Intelligent Autonomous Systems*, 2014.
14. J.Miura H.Goto and J.Sugiyama. Human-robot collaborative assembly by on-line human action recognition based on an fsm task model. In *HRI2013 Workshop on Collaborative Manipulation: New Challenges for Robotics and HRI*, 2013.
15. H.Goto T.Hamabe and J.Miura. A programming by demonstration system for human-robot collaborative assembly tasks. In *IEEE International Conference on Robotics and Biomimetics*, pages 1195–1201, 2015.
16. K.Yamada and J.Miura. Ambiguity-driven interaction in robot-to-human teaching. In *Int. Conf. on Human-Agent Interaction*, pages 257–260, 2016.
17. Partner robot family. `http://www.toyota-global.com/innovation/partner_robot/family_2.html`.

18. navigation - ros wiki. http://wiki.ros.org/navigation.
19. Yolo: Real-time object detection. https://pjreddie.com/darknet/yolo/.
20. find_object_2d - ros wiki. http://wiki.ros.org/find_object_2d.
21. Hark wiki. http://www.hark.jp/.
22. Google cloud platform. https://cloud.google.com/speech/.
23. Stanford nlp. https://github.com/stanfordnlp.

6 Team Information

Team Name: AISL-TUT

Contact Information:

Shuji Oishi
Active Intelligent Systems Laboratory (AISL, Miura Laboratory) Department of Computer Science and Engineering, Toyohashi University of Technology C2-503 Building C, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi 441-8580, Japan

contact email: oishi@cs.tut.ac.jp

Fig. 5. HSR Robot

Website: http://www.aisl.cs.tut.ac.jp/robocup

Team Members: Shuji Oishi, Jun Miura, Kenji Koide, Mitsuhiro Demura, Yoshiki Kohari, Soichiro Une, Liliana Villamar Gomez, Tsubasa Kato, Motoki Kojima, Kazuhi Morohashi

Table 1. Hardware Specifications

Drive system	Omnidirectional moving mechanism
Robot Sensors	Laser measuring range sensor, IMU, Magnetic sensor
Gripper sensors	Potentiometer, gripping force sensor, Wide-angle camera
Head sensors	RGB-D sensor, Stereo camera, Wide-angle camera, Microphone array
Arm sensors	Absolute type joint angle encoder, 6-axis force sensor
Body	430mm diameter, 1,005-1,350mm height, 37kg weight
Hoisting	Telescope mechanism, Weight compensation mechanism
Max payload and speed	1.2kg, 0.8km/h
Max incline	5°
Display	7.0 inch size, 1024 x 600 resolution
CPU	4th Gen Intel Core i7 (16GB RAM, 256GB SSD)

Table 2. Software Specifications

Operating System	Ubuntu 14.04
Middleware	ROS Indigo
Localization	HSR API
Navigation	Randomized Path Planner[7] and HSR API
Arm Control	HSR API
Object Recognition	YOLO and Find object 2D
Speech Synthesis	Sound play ROS
Speech Recognition	HARK AND Google Speech API
Natural Language Understanding	Stanford NLP