ZJUCyber
Team Description for ROBOCUP 2013

Rong Xiong, Junnan Wang, Zenglin Hong, Qianshan Li, Penghui Yin, Yingying Xu, Xiaokai Sun, Longbin Shen, Yue Wang, Haoyuan Pang

State Key Laboratory of Industrial Control Technology, Zhejiang University
Zheda Road No.38, Hangzhou, Zhejiang Province, P.R.China
jnwang@iipc.zju.edu.cn
`http://www.nlict.zju.edu.cn/robocupathome/zjupanda.htm`

Abstract. This paper is the description of team ZJUCyber for the RoboCup@Home 2013. We report the works we have done which allows us to perform competently in the @home league, including the hardware architecture, software architecture and main algorithms used to implement various service robotics applications.

1 Introduction

The ZJUCyber is a team of Zhejiang University devoting to RoboCup@Home league, which was called ZJUPanda before, supported by Robot Laboratory of Institute of Cyber-Systems and Control (CSC), Zhejiang University. The team ZJUCyber which consisted of Master and PhD students was founded in 2010 and has participated three times in RoboCup@Home since 2010. We have got the 16th place and the 7th place in the RoboCup @Home in Singapore and Istanbul respectively, but we didn't get a satisfactory result in Mexico due to some hardware problem. Figure 1(d) and figure 1(e) show the robots which participated in RoboCup@Home 2010 and 2011, figure 1(f) is the current robot we use, the hardware was redesigned before Robocup2012, we will continue to use this version in the Robocup2013, and we are looking forward to its more stable.

The main research objective of our laboratory is developing intelligent mobile robotics especially intelligent service robots which can be used in regular home environments. Figure 1 shows some robot prototypes we developed in recent years. Our research interests include SLAM, path planning, environment modeling, 3D environment reconstructing, face recognition, object recognition, Human Robot Interaction and so on. For RoboCup@Home 2013, we improve the efficiency and stability of our algorithm based on the same hardware in Robocup2012, like object recognition, navigation and arm control. The details will be described later.

The rest of this paper is organized as follows. Section 2 will briefly present the mechanic design and hardware of our robot platform. Section describes the software architecture and main algorithms such as mapping, localization, navigation and vision will be described in Section 3.2 and section 3.3 respectively. Section 4 concludes the paper.

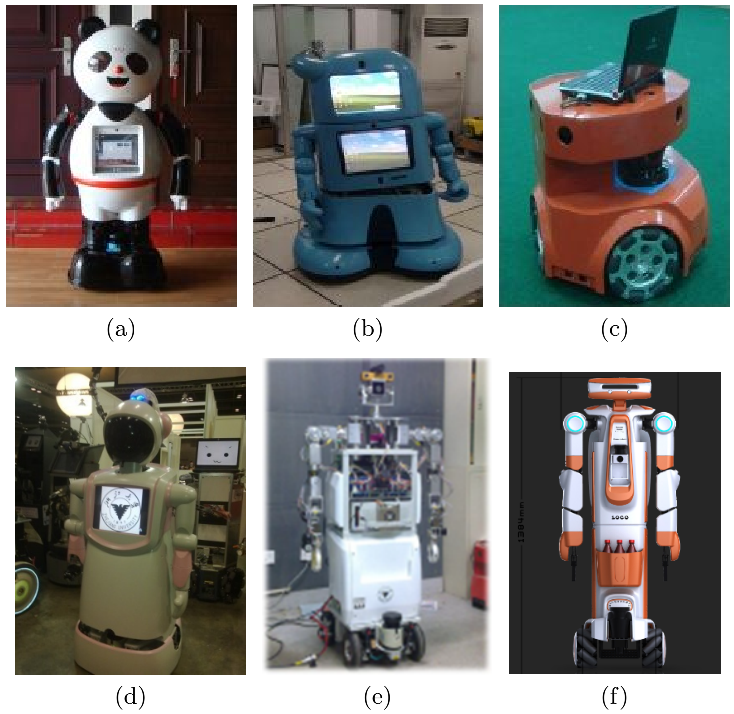


Fig. 1. Robots that developed in our laboratory: (a) Panpan, a shopping guide robot, (b) Haibao, a exhibition hall service robot, (c) Xiaohong, a mobile robot Platform for research and teaching (d) ZJUPanda for @Home 2010, (e) ZJUPanda for @Home 2011, (f) ZJUPanda designed for @Home 2012, and also ZJUCyber for @Home 2013

2 Robot Platform

The robot that will be used in the RoboCup @Home 2013 is based on the mobile platform of our previous service robots, see figure 1. The configuration of the new developed ZJUCyber is shown in figure 2 in detail. The the platform is equipped with four mecanum wheels for omnidirectional moving.

For environment perception, a Microsoft Kinect RGB-D camera is attached to robot's head for the detection and recognition of objects,faces, persons and gesture,as well as extracting the structure of the environment. A laser range finder(SICK LMS 100) is installed on the underpan of the robot for SLAM, obstacle avoidance and path planning, and anothre laser range finder (URG-30LX) is installed in the middle of robot with 1 DOF which is used for person tracking, 3D obstacle avoidence and environment reconstruction. A microphone and a speaker are installed for mainly speech-based human-robot interaction.

Besides, the robot was equipped with a pair of arms with 7 DOFs, a pair of two-finger hands for grasping, a touch screen for debugging, and a pan-tilt platform for the control of cameras mounted on it. The arms driven by DC motors are installed on the base of the robot not only used for manipulating things but also for showing the liquid movements in a robot dance performance.

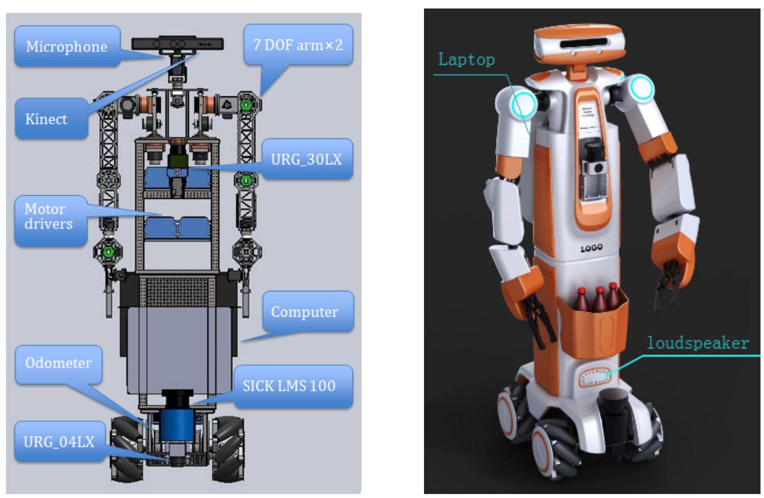


Fig. 2. Hardware configuration of ZJUCyber

The sensors and actuators are connected to two computers for perception and decision. A laptop (Dell Latitude E6420ATG) is used for image processing, image understanding, path planning, speech recognition, speech synthesizing and decision making, while another IPC with Intel Core I3 CPU, and 4G RAM is used for SLAM, navigation, and motion control. ALl the modules are communicating through the Ethernet.

3 Software Architecture and Main Algorithm

The main purpose of building the software architecture of ZJUCyber is to provide a platform or framework, which is robust, extensible and run-time dynamic for the development of separate modules such as localization and vision. So each independent module can focus on its algorithm logic while ignoring the irrelevant factors such as data acquisition or how and when the module will be invoked. All the algorithms are involved in the architecture as modules.

3.1 Software architecture

The entire software architecture is demonstrated in figure 3.

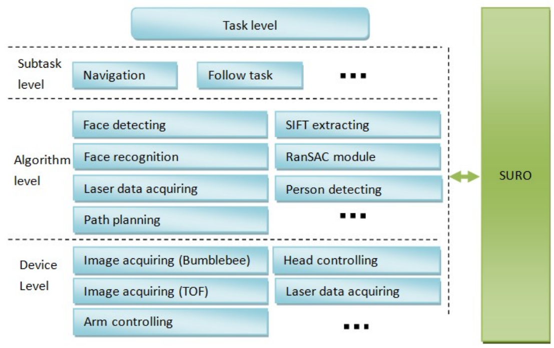


Fig. 3. Software architecture

The software system is divided into 4 levels: Task level, Subtask level, Algorithm level and Device level. The three lower levels are packaged into independent modules. The device level encapsulates the direct access to the device hardware and provides robust, exception-safe and thread-safe interface for data buffering and acquisition. Different device types, such as odometer, laser and camera are managed with a unified abstract interface to make the system extensible. The algorithm level contains a set of "algorithms" such as localization, path planning, face detecting etc.. Each algorithm can fetch data from device level and generate algorithm outputs. The subtask level includes some basic tasks of the robot, like follow task, navigation etc.. The task level invokes the modules in the lower three levels to complete some complex tasks, like the tasks of @Home. In the task level, we invoke all modules in a state machine.

SURO is the core of the software architecture, which manages all the modules in the lower three levels and the data exchanging between these modules within a computer or between different computers. The architecture of SURO is shown in figure 4.

3.2 Localization and Navigation

SLAM (Simultaneous Localization And Mapping) is generally regarded as one of the fundamental problems in the pursuit of building truly autonomous mobile robots. This part of our work was supported by the National Nature Science Foundation of China (Grants No. NSFC: 60675049), the National 863 plan (Grants No. 2008AA04Z209). Several papers have been published.

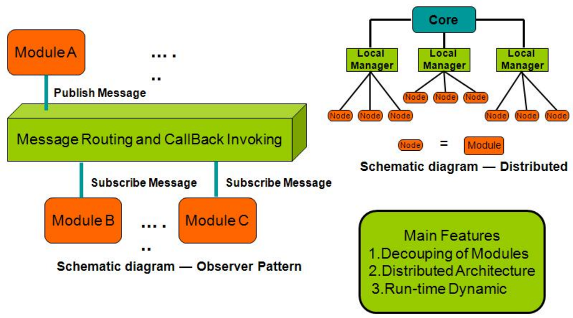


Fig. 4. Architecture of SURO

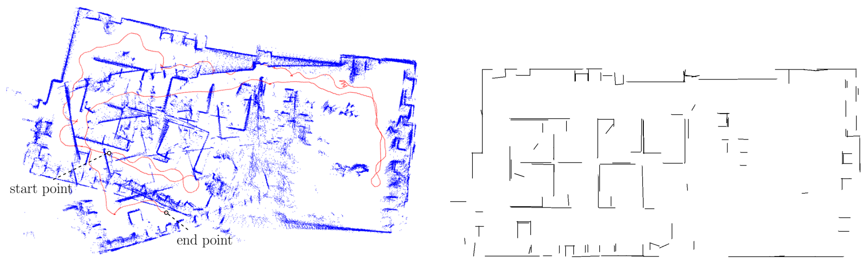
We solve the SLAM problem on our robot using a incremental mapping algorithm combined particle filter with dot-line congruence. In the approach, each particle carries an individual segment-features map of the environment. Both the motion and the observation information are considered in the importance function of the particle filter by using the dot-line congruence method to estimate the pose of robot. The weight of the particle is updated according to the congruence between current measurement and segment features in previously-built map. The wrong particles resulted from mis-matching or error accumulation are filtered with selective resampling.

As commonly used, we denote the data available for mapping the form of $d^t = \{z^t, u^t\}$, where z^t is sequence of sensor measurements, $z^t = \{z_1, \ldots, z_t\}$, u^t is the odometry measurements, $u^t = \{u_1, \ldots, u_t\}$. In statistical terms, mapping is the problem of finding the most likely map given the data d^t. We present the posterior of the map and the robot trajectory in a factored form the same with other Rao-Blackwellized particle filters.

$$p(x^t, m|z^t, u^t) = p(m|x^t, z^t) \cdot p(x^t|z^t, u^t). \quad (1)$$

Where x^t is the robot trajectory and m the most likely map. The difference is that the map we estimated is a segment-features map and we use dot-line congruence to evaluate the robot trajectory. As shown in the figure 5, our method is robust and can be used in various of indoor environments.

Map Interpretations is the way we explain the environments. Three different interpretations are used for different purpose. Occupancy grids is the map used for robot, which is efficient for localization and navigation. But the grid map is not intuitional for human operation. So we make other picture by hand to interpret the environment to people by words or logo. The last interpretation is especially for our task engine. We assign the name to each room in the map in order that the robot can understand where we want them to go by speech. The properties of certain objects are also recorded in this interpretation. For



(a) (b)

Fig. 5. (a) Map obtained in an office environment from sensor data using raw odometry. (b) Final vector map.

example, the type of the door or refrigeratory direct the robot how to open them, the stairs tell the robot to move away from them during navigation and a dock will leads the robot to automatic charge itself.

Localization Monte Carlo Localization(MCL) is used for our robot based on occupancy grid map. This method is approve to be very efficient in localization problem. We improve the original method by apply the method on GPU. A nVidia GeForce 9400GT Graphic Card is used to process this procedure. The time we used to evaluate 2000 particles' weight only cost less than 50ms.

Furthermore, each particle in our method process a scan registration so as to solve the kidnap problem. The scan registration is indeed a maximization search. Benefit from the fast GPU computing, we adopt a big search step. Usually, in a region of 5m*5m, 50 particles is enough to find the correct robot pose.

Path Planning We use Quadtrees[8,3] to extended the original probabilistic roadmap method(PRM)[5,2] for robot path planning. Instead of the probabilistic roadmap generation step, we use a quadtree to split the map and generate the road maps in a regular and recursive way.

After generating the road map, we use a A^* algorithm to search the best path between the robot and the target. Experimental results demonstrate our approach is effective and robust.

3.3 Vision System

Vision system including image preprocessing, face detection and recognition, object recognition and localization, person detection, and some auxiliary modules like RanSAC and so on. In this section, the modules of face and person detection, face recognition, will be described in detail.

(a) (b) (c)

Fig. 6. Results of face alignment

Face and Person Detection In the system, person and face are detected in the same way, using Haar-like features based cascade classifiers[7], which can be implemented using a state-of-the-art tool available in OpenCV. the only deference between them is that the face is detected based on color images, while the person is detected by finding the head-shoulder part of the person, based on range images captured by kinect.

Face Recognition Regarding to the application of service robot, face recognition module is designed considering the lack of face samples and the variations of illumination, expression and head pose. The face image is first aligned and then recognized.

For face alignment, a novel method is proposed to align face image proposed method learns a 3D face shape model comprised of 31 facial features and a texture model for each facial feature from a 3D face database. The 3D face shape model and the texture models are incorporated with Markov Random Field(MRF)[4] to model the texture similarity constraints, pairwise and high order constraints of the facial features, and an optimization algorithm is employed to get the optimal solution. Experiments results are shown in figure , results shows capability of adapting head pose variations of the method.

Face image is recognized based on MRF with single face sample. The gallary face image is divided into blocks which are used as the nodes of the MRF. The corresponding blocks in the test face image are estimated With the result of face alignment, and then refined by minimizing the potential function of the MRF, which means considering both the block similarity and neighbor smoothness. The recognition algorithm works under small face pose variants within about 45 degrees, and it is robust to small illumination variations due to the illumination invariant block descriptor (The texture is described by Histogram of Gradient, HOG[6]).

3.4 Person Tracking

Person tracking is essentially important for a service robot. We solve the problem in a probabilistic framework. First, person candidates are detected using kinect and laser range finder, and the multi-person tracking module based on

SJPDAFs(Sample-based Joint Probabilistic Data Filters[1]) is applied to filter out the error detections and give the tracks of the persons, according to equation (2).

$$\omega_{i,n}^{k} = \alpha \sum_{j=0}^{m_k} \{ \sum_{\theta \in \Theta_{ji}} [\alpha \gamma^{(m_k - |\theta|)} \prod_{(j,i) \in \theta} \frac{1}{N} \sum_{n=1}^{N} p(z_j(k)|x_{i,n}^{k})]\} p(z_j(k)|x_{i,n}^{k}) \quad (2)$$

The method combines the advantages of particle filters, owning to the ability to represent multi-modal state densities and the efficiency in assigning the measurements to individual targets.

3.5 Speech Synthesis and Recognition

For speech module, speech recognition and voice synthesis product from iFlyTek are adopted in ZJUCyber.

4 Conclusion

In this paper we presented the ZJUCyber RoboCup@Home team. We reported on our robot platform, software architecture and important approaches used to implement various service robotics applications in the @home league, among which map building, path planning and vision strategy were described in detail. With all the modules, ZJUCyber is enabled to accomplish many service tasks. In feature work, the perception like 3D environment reconstruction, image understanding under unstructured environment etc. will be emphasized.

References

1. D. Fox D. Schulz, W. Burgard. People traking with mobile robots using sample based joint probabilistic data association filters. *The International Journal of Robotics Research*, 22:99C116, 2003.
2. L. E.Kavrak, P.Svestka, J. C. Latombe, and M.Overmars. Probabilistic road maps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*, pages 566–580, 1996.
3. Samet H. Neighbor finding techniques for images represented by quadtrees. *Computer Graphics and Image Processing*, 18:37–57, 1982.
4. W. T. Freeman J. S. Yedidia. Understanding belief propagation and its generalizations. In *International Joint Conference on Artificial Intelligence*, 2001.
5. L. E. Kavraki and J.-C.Latombe. Randomized preprocessing of configuration space for fast path planning. In *IEEE International Conference on Robotics and Automation*, pages 2138–2145, 1994.
6. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:90C110, 2004.
7. M. Jones P. Viola. Rapid object detection using a boosted cascade of simple features. In *In IEEE Computer Society Conference on Computer Vision and and Pattern Recognition*, 2003.
8. H. Samet. An overview of quadtrees, octrees, and relatedhierarchical data structures. *NATO AS1 Series*, 40, 1988.