The b-it-bots RoboCup@Home 2013 Team Description Paper

Rhama Dwiputra, Matthias Füller, Frederik Hegger, Nico Hochgeschwender, Jan Paulus, Sven Schneider, Azden Bierbrauer, Elizaveta Shpieva, Iryna Ivanovska, Niranjan V. Deshpande, Adam J. Gaier, Alexander Hagg, José M. Sánches Loza, Alexey Y. Ozhigov, Paul G. Ploeger, and Gerhard K. Kraetzschmar

Bonn-Rhein-Sieg University of Applied Sciences
Department of Computer Science
Grantham-Allee 20, 53757 Sankt Augustin, Germany

Abstract. This paper presents the b-it-bots RoboCup@Home team and its mobile service robot called *Jenny* – a service robot based on the Care-O-bot 3 platform manufactured by the Fraunhofer Institute for Manufacturing Engineering and Automation. In this paper, an overview of Jenny control architecture and its capabilities is presented. The capabilities refers to the added functionalities which are the results of individual research and development projects (R & D) and master theses carried out within the Bonn-Rhein-Sieg University of Applied Science.

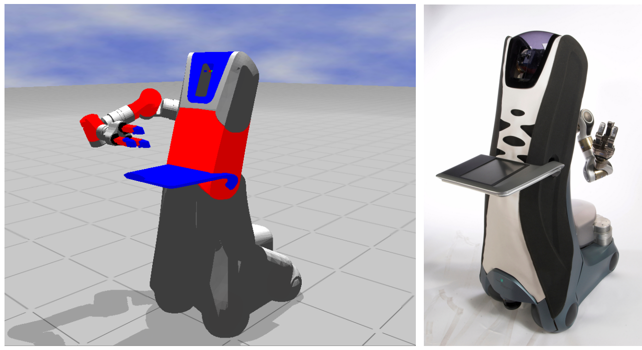
1 Introduction

The b-it-bots RoboCup@Home team at Bonn-Rhein-Sieg University of Applied Sciences (BRSU) has been established in 2007. The team consists of Bachelor, Master and PhD students, who are advised by two professors. The results of several research and development (R&D) as well as Master theses projects had already been successfully integrated into a well-functioning robot control software system. Through this kind of graded course modules our RoboCup@Home team is strongly interwoven with the Master by Research course in Autonomous Systems, which is offered at the BRSU[1]. During the last years our team ranked among the top five in the RoboCup@Home league with several placements on the podium. Our main research interests include mobile manipulation, environment modeling, computer vision and human robot interaction (HRI).

2 Robot Platform

In 2011, the primary robot platform for the b-it-bots team @Home competition was changed from "Johnny Jackanapes", a differential drive robot based on the Volksbot platform, to "Jenny", the omni wheeled robot Care-O-bot 3 (see Figure

[1] For more information, see `http://www.inf.h-brs.de/MAS`.



(a) (b)

Fig. 1. "Jenny" aka Care-O-bot 3 moving around in a simulated environment (a) and in real world (c) ©Fraunhofer IPA

1). The Care-O-bot has been developed by the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA) in Stuttgart, Germany[2] . The Care-O-bot 3 is equipped with a 7 DoF manipulator, a three finger hand and a omnidirectional platform, which is powered by 8 motors (2 motors per wheel: 1 for rotation axis, 1 for drive). The sensor head contains two AVT Pike 145 C Cameras, which are used for stereo processing, a MESA Swissranger 4000 time-of-flight camera, and a Kinect camera. Two SICK S300 laser scanners and one Hokuyu URG-04LX laser scanner are used for mapping and navigation.

3 Functionalities

In this section, the added functionalities for the robot platform is presented. The functionalities are the results of R & D projects, Master theses, and other projects. In most cases, the state-of-the-art and available best practice solution was investigated in the early phase of each project. This activity will be followed by implementation and integration of the applicable algorithms to the robot platform. Finally, a novel or adapted approaches is developed to improve the robustness in performing tasks within an uncertain domestic environment.

3.1 Object Categorization [1]

In service robotics, tasks without the involvement of objects are barely applicable, like in searching, fetching or delivering tasks. Service robots are supposed to capture efficiently object related information in real world scenes while for instance considering clutter and noise, and also being flexible and scalable to memorize a large set of objects. In order to detect object related information in cluttered domestic environments an object detection method is proposed that

[2] For more information, see `www.care-o-bot-research.org` and `http://www.care-o-bot.de/`

copes with multiple plane and object occurrences like in cluttered scenes with shelves. Further a surface reconstruction method based on Growing Neural Gas (GNG) [2] in combination with a shape distribution-based descriptor is proposed to reflect shape characteristics of object candidates. Beneficial properties provided by the GNG such as smoothing and denoising effects support a stable description of the object candidates which also leads towards a more stable learning of categories. Based on the presented descriptor a dictionary approach combined with a supervised shape learner is presented to learn prediction models of shape categories. Experimental results, of different shapes related to domestically appearing object shape categories such as cup, can, box, bottle, bowl, plate and ball, are shown. A classification accuracy of about 90% and a sequential execution time of lesser than two seconds for the categorization of an unknown object is achieved which proves the reasonableness of the proposed system design.

3.2 People Detection

People awareness helps a mobile service robots to operate safely in coexistence with humans and react on their movements and actions. Two different approach have been developed and investigated for people detection in domestic environment.

People detection by acoustic clues For humans speaking is a very easy way to communicate. So the spoken word is probably also the most natural ways to interact with a service robot. For speech recognition we employ the very mature speech SDK SAPI5.4 from Microsoft and some handcrafted grammar of keywords to understand the commands of the user. However, there are more ways to interact using voice. We propose an extended use of acoustic clues. Take for example a user, who calls for the robot from some distance and asks for assistance either in an apartment or an restaurant. Then it would be natural to turn the attention to the speaker and reassure for the understood command. Our sound localization component localizes sound sources through a microphone array with four microphones and a M-Audio fast track as interface. The beamform technique is used to calculate the direction of the sound source relative to the robot based on the time shift in the input signal of the individual microphones. The system performs very well in both quiet and noisy environment.

3D People detection [3] The 3D people detection uses the recently available *Microsoft Kinect* sensor which combines the advantages of LRFs (fast, accurate), monocular cameras(color information), and TOF cameras (3D information). The preliminary segmentation is based on a top-down/bottom-up technique which yields the capability of detecting partially occluded person, e.g. behind a desk or cupboard. The information gained from the local surface normals enable the system to detect a person in various poses and motions, i.e. sitting on other objects, bended to the front or side, walking fast/slow. As final machine learning technique, a Random Forest classifier is applied which outperformed the opponents AdaBoost and SVM. The presented approach is able to detect people up to a distance of 5 meters with a detection rate of 87.29% for standing and 74.94% for sitting people.

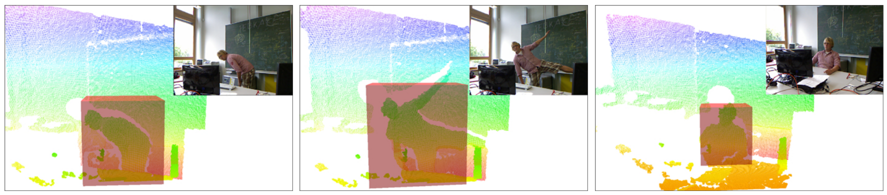


Fig. 2. Detections for various pose configurations.

3.3 Human Robot Interaction

Human robot interaction (*HRI*) is a multidisciplinary field aiming to find ever faster and more intuitive manners of communication between humans and robots. Mostly humans express their intentions via speech, gestures, expressions and sounds. Domestic service robots must be aware of those intentions and also be able to understand them. In this section we present three human robot interaction modules endowed into our domestic service robot *Jenny*. For some scenarios the different modules have been combined to increase the robustness of our robot.

Facial expression recognition Facial expression recognition (FER) is an unique feature of our entire human robot interaction system. Currently, our robot can recognize the human expressions neutral, joy, surprise, sadness, fear, anger and disgust. Our current system is based on the extraction of Gabor features at different orientations and scales. The features are first extracted from a normalized face image and then forwarded to a multi-classification stage. The number, location, orientation and scale of the Gabor filters to use is determined in the training phase via Adaboost as suggested in [4][5]. Our system was trained and tested with images of the Cohn-Kanade AU-Coded Facial Expression Database [6]. Accuracy of 87.84% was achieved for the subset 1 of expressions when SVM classifiers were used. Multi-classification with SVM was carried out using both the Error Correcting Output Coding scheme and the "one-versus-one" scheme. The system showed to perform very fast, in a person-independent manner and fully automatic (i.e., no manual settings are necessary in the testing phase).

Gesture recognition [7] For the interaction with a robot, pointing gestures could indicate objects and locations. Our person independent pointing gesture recognition application is able to cope with different skin colors, variable and complex backgrounds, skin colored regions in the background, and with various

Fig. 3. Four example images correctly classified by the system. From the Cohn-Kanade AU-Coded Facial Expression Database [6], with permission.

lighting conditions. For the visual input and the necessary depth information, a stereo vision camera is used. First the face of the user is detected and tracked with the CAMSHIFT algorithm, then the skin color histogram is extracted and the backprojection image is calculated. After background subtraction and morphological operations, only the face and the hands keep visible. The pointing target is the nearest object in the line-of-sight (in 3D space) between the head and the pointing hand. For pointing with the fist, the average recognition rate is 90.28% for a distance up to 216 cm (head to pointing target). When pointing with the fingertip, the recognition application has an average recognition rate of 81.25% - 87.5% for near distances (head to pointing target between 109 cm and 116 cm) and an average recognition rate of 70.92% for a distance of up to 254 cm (head to pointing target). In Figure 4 the first processing steps are exemplified.

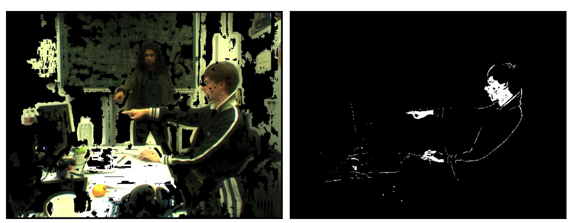


Fig. 4. Processing steps: histogram backprojection and deletion of background

Haptic Interface [8] The manipulator of the robot platform can be utilized to function also as a haptic interface for human user. This functionality have been developed and implemented in two scenarios, *guidance* and *cooperative transportation*. In *guidance*, a user can control the robot's base movement trough interacting with the manipulator wrist or gripper. In the second scenario, *cooperative transportation*, the robot can carry an object together with a human user and follow the movement direction. Both scenarios are shown in figure 5.

Through the use of smoothing filter and PID controller, the feature is able accommodate noisy input and produce a steady movement. The feature is developed so that it can be applied in almost all possible configuration. Specifically

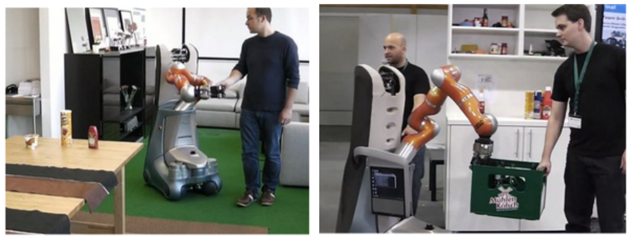


Fig. 5. Manipulator as a haptic interface

for *guidance* scenario, a user trial have been performed and the results shows that the functionality is intuitive and compatible for different type of users.

4 Conclusion

In this paper the robot platform of the b-it-bots RoboCup@Home team and its added functionalities is presented. Furthermore, some novel features of the robot, namely object categorization and an improved set of human robot interaction modalities like sound source localization were introduced. The integration of the functionalities in one robot to another was and is still an exhaustive exercise. In our current EU FP7 funded project BRICS (Best Practice in Robotics)[3] we are exploring first steps towards an improved software development methodology in robotics. Among others we applied the component-oriented development approach defined in BRICS for our software development which turned out to be very feasible when several heterogenous components are composed into a complete system.

Acknowledgement

We gratefully acknowledge the continued support of the team by the b-it Bonn-Aachen International Center for Information Technology, the Bonn-Rhein-Sieg University of Applied Sciences and the Fraunhofer Institute for Manufacturing Engineering and Automation in Stuttgart for the help and support in providing us the Care-O-bot as a research platform for the RoboCup@Home competition. In particular, we would like to thank Ulrich Reiser, Florian Weisshardt, Alexander Bubeck and Christian Connette for their support.

References

1. Mueller, C., Hochgeschwender, N., Kraetzschmar, G.K., Ploeger, P.G.: 3D Object Shape Categorization in Domestic Environments. Technical report, Bonn-Rhein-Sieg University, Germany (2011)

[3] For more information, see `www.best-of-robotics.org`

2. Mueller, C., Hochgeschwender, N., Ploeger, P.G.: Surface Reconstruction with Growing Neural Gas. In: Proc. of the Workshop on Active Semantic Perception and Object Search in the Real World held at the Conference on Intelligent Robots and Systems (IROS), San Francisco, USA. (2011)
3. Hegger, F., Hochgeschwender, N., Kraetzschmar, G.K., Ploeger, P.G.: People Detection in 3D Point Clouds Using Local Surface Normals. In: Proc. of the 16th RoboCup International Symposium, Mexico City, Mexico. (2012)
4. Shen, L., Bai, L.: AdaBoost Gabor feature selection for classification. In: Proc. of Image and Vision Computing NewZealand, Akaroa, New Zealand (2004) 77–83
5. Bartlett, M.S., Littlewort, G., Lainscsek, C., et.al.: Machine learning methods for fully automatic recognition of facial expressions and facial actions. (2004) 592–597
6. Kanade, T., Cohn, J., Tian, Y.l.: Comprehensive database for facial expression analysis. In Proc. of International Conference on Automatic Face and Gesture Recognition (2000) 46–53
7. Breuer, T., Hartanto, R., Ploeger, P.G.: Gesture-based HRI for a Mobile Manipulator (2009)
8. Dwiputra, R., Hochgeschwender, N., Paulus, J., Kraetzschmar, G.K.: Haptic Interface for Mobile Manipulator (2011)