Konrad and Suse, Two Robots Guiding Visitors in a University Building

Ronny Stricker, Steffen Müller, Erik Einhorn, Christof Schröter, Michael Volkhardt, Klaus Debes, Horst-Michael Gross

Ilmenau University of Technology, Cognitive Robotics Lab, 98693 Ilmenau, Germany

Abstract. This paper presents an overview of the hard- and software architecture of a mobile visitor information system for the Konrad Zuse building of the School of Computer Science and Automation at the Ilmenau University of Technology. Two mobile robots serve as mobile information terminals with capabilities for generating way descriptions and guiding the visitor to the points of interest (labs, meeting rooms, offices, employees) in the building. The paper focuses on the constraints resulting from the challenging environment in this multi-floor building, as well as on the integration aspects of various skills for navigation and human-robot interaction. Besides first experience with the system, the further development is outlined as well.

1 Introduction

The long-term research project Zuse-Guide aims at developing two autonomous interactive service robots called "Konrad" and "Suse" which are to autonomously guide and to tour visitors within the new Konrad Zuse building of the Ilmenau University of Technology, hosting the School of Computer Science and Automation on 4 floors since June 2011. The service robots are supposed to provide an intuitive interaction to potential users. Service tasks considered in this project are to autonomously navigate in the building and guide the visitors, to provide information about the different labs and to tour visitors around in the building, giving them a deeper insight into the research topics and exhibits of the institutes and labs. During their operation, the robots wait next to the main entrance doors at the ground floor and offer their service when a potential user approaches. The user can select the different functions using a touch screen-based graphical dialog system which is commented by a speech synthesis system to provide a natural and intuitive user experience. The robots also offer to guide the visitor to the destination of his/her choice and inform about news for the different institutes and labs. When the robots have finished guiding the user around, they return to their waiting position to welcome the next visitor.

The main part of the visitor information system are the two guiding robots (see Fig. 1). The robot platform is a SCITOS A5 produced by Metralabs GmbH Ilmenau, which originally has been developed for the application as shopping assistant [GBS$^+$08].

Besides the main sensors and devices used in the shopping robot setup, the robot is equipped with an RFID antenna at the bottom, able to detect tags burrowed in the floor (for emergency stop in front of the stairways). Additionally, for operating the elevator
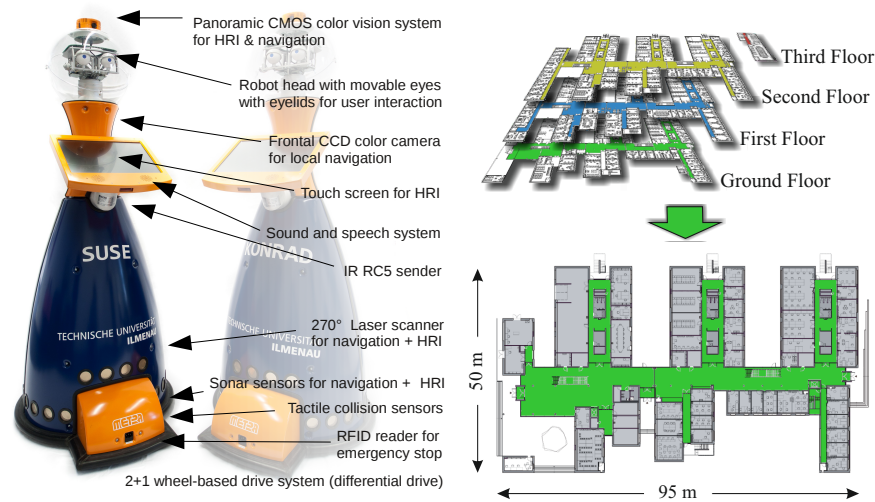
Fig. 1. Left: sensors and interaction devices of the robot platforms. Right: Different floors of the four-level Zuse building with a detailed view of the ground floor (operation area for the guide robots highlighted in color). The first and second floor have a similar shape, whereas the third floor only comprises two lecture rooms.

an infrared sender has been mounted that allows remote calling of the elevator and the "manual selection" of the target story.

This paper summarizes the project objectives as well as the current state of implementation which already allows for a continuous guidance service, however, with a limited spectrum of functionality. Since this project is still ongoing work, the paper concentrates on the aspect of the robot infrastructure and the functional interaction of the different software modules. Therefore, we will give an overview of the constraints and requirements arising from the challenging scenario in the following section. After that, we will give an insight into the infrastructure of the building and the administrative software modules in section 3 and section 4. A more detailed overview of the different software modules running on the robot platform is given in section 5. The paper concludes with some findings on the first days of operation and an outlook.

2 Challenges and Constraints of the Scenario and State-of-Art

In the last years, diverse mobile service robots employed as tour guides for exhibitions and museums have been presented [BJ05]. Among them are such well known robots as Rhino, Minerva, and Sage, the exposition guide RoboX, or the robots Mona/Oskar at the Opel sales center at Berlin. Also Fraunhofer IPA developed robots for the museum for communication Berlin [SGTJ01]. Usually, these robots guide visitors to a set of exhibits while offering related information, and thus have a similarity to our application. However, none of those systems is working in a multi-level environment and is involved in such a complex infrastructure like a school of computer science with hundreds of employees, umpteen of offices and labs, and a highly dynamic, continuously changing building layout.

In our operational environment a high degree of dynamic is present. Staff working in the labs and students are moving and working in the public areas. They are changing furniture and exhibits, and experimental setups are occupying the hallways, resulting in a quite crowded environment at some times. All this is demanding a dynamic and situation-aware navigation.

The Konrad Zuse building has more than 200 rooms (labs, lecture rooms, offices) and accommodates more than 225 people. In total the building has an area of more than 11,000 m^2 on 4 floors, of which the hallways the robots operate in are about 5,000 m^2 (Fig. 1).

More challenges arise from that multi-level architecture of the building. Not only must the robots be able to navigate over different floor levels, but even more important, the stairways are dangerous obstacles which have to be avoided very robustly. Multi floor navigation with incorporation of elevator usage is known from mobile transport systems [NK08],[TGVS01], which usually are centrally controlled or using fixed pathways for navigation. Since our robots are supposed to navigate fully autonomously, they have to be able to control and use the elevators of the building, and to recognize the current floor they are operating on, requiring multi-level localization.

For unsupervised continuous operation, docking stations enable the robots to autonomously recharge their batteries. Such self-sustaining operation can also be found at the shopping assistants of our former project [GBS+09].

Since the robots operate in a large area where it might take some time to find them when they are not at their initial waiting position, a convenient monitoring and remote control interface is required.

Besides these technical aspects also the human-machine interaction plays an important role in our scenario. People have to be interested in using the information services and communication should be adaptive and situation aware. Various studies regarding this topic again come from the domain of entertainment and tour guide robots, e.g. in [FBE+09].

3   Operational Area

As already described in the introduction, the public area of a large office building comes along with numerous challenges. This section will introduce the environmental conditions and necessary installations to enable a safe operation of autonomous robots.

Resulting from the four floor architecture, the central aisle shows a couple of stairs, that are the most critical areas for the robots. Downward stairs pose the risk of fall, while passing below an upward stair might result in a collision with the robot's head. The limited vertical range of the laser range scanner and sonar sensors does not allow to perceive both types of obstacles. In order to prevent crashes even when a faulty localization leads the robot to the proximity of these obstacles, the robot is equipped with an RFID reader, closely coupled to the drive system. RFID tags burrowed in the floor at the critical areas (in front of and beneath stairs) can deactivate the drive in less than one second. Tags have been placed in two and three rows respectively to guarantee at least one detection when passing over them even at high speed. In addition to the RFID based stair detection, we are preparing a second and independent system based on the detection of magnetic tapes which are burrowed in front of the stairs to meet
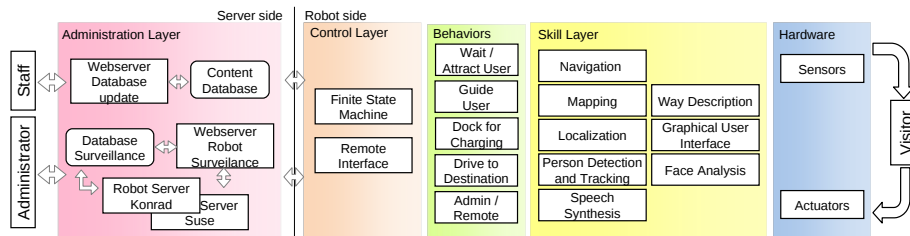
Fig. 2. System architecture, consisting of mobile robots part (right) and server side (left)

the safety requirements of public buildings and the criteria of the German Technical Inspection Agency (TÜV). As long as the installation of these is not finished, for safety reasons we are limited to operate on the ground floor at the moment. Therefore, the user is guided to the stairs or to one of the elevators if the target is located at one of the upper floors.

A second requirement is the ability to operate elevators. In order to enable the robots to control the elevators with only minimal changes to the elevators themselves, each of the 7 elevator doors was equipped with an RC-5 receiver (a protocol for consumer infrared remote control) that is connected to the call button of the existing elevator control electronics. In the same manner, two RC-5 receivers are installed in each of the two elevator cabins connected to the buttons for selecting the destination floor. In effect, robots interact with the elevator nearly similar to humans, and no change in the existing elevator controllers was necessary.

The third essential infrastructure is a wireless network installation that is covering the entire building. The wireless connection is used for security monitoring and for database access while browsing via the GUI only. Therefore, a temporary loss of connection during a guide tour is not critical.

4   Administrative Backend

A very important issue when running robots fully autonomously addresses the surveillance and remote administration. Therefore, the robots are integrated in a framework which enables the communication with the robots. This section gives a brief introduction into the used communication framework and explains the different software components involved to build up the background for the information system.

As shown in Fig. 2, the central administrative component is a server hosting a MySQL database and a webserver. Furthermore, on that host a robot server is running for each of the robots, which is the server side communication endpoint for the robot. Software on the robots as well as the robot server are based on the MIRA middleware [MIR], providing inherent means for sharing robot data between distributed modules.

The administration server has two main functions. On the one hand, it offers a website interface for maintaining the knowledge base comprising room data, institute and lab information, exhibits and people working in the building. Authorized staff can edit department specific news pages and change assignments of people to rooms and exhibits. In the future, it will also be possible to book the robot for tours or welcoming guests at the main entrance at a certain time via this website. All this data is stored in one database, that is accessed by the robots each time a user requests the respective information. On the other hand, there is an administration website for remotely observing

and controlling the state of the robots. A floor plan indicates the current position, connection state and internal state of the robots, and allows to send the robots to arbitrary target positions.

5   Software architecture

Relying on the administrative infrastructure, the robots each operate an application that is also based on the MIRA middleware framework with the aim to guarantee the main requirements to a modern robot control architecture, like modularity, extensibility, efficiency, customisability, reusability, and rapid application development. Similar to the former software architecture used for the shopping robots [GBS$^+$08], the robotics related methods and skills have been abstracted from the application itself resulting in a flexible layered system architecture (Fig. 2). In this architecture, the bottommost Hardware Layer encloses the hardware (sensors and actuators) and the operating system. The low-level sensor information is processed in the next level, the Skill Layer, which provides a set of necessary, robotic-specific basic skills. These skills are only weakly coupled and do not need any coordination since they are operating in parallel using exclusive resources. The Skill Layer covers the whole spectrum of navigation skills including map building and localization, as well as person detection and tracking. It is planned to also integrate face analysis [MG08] in order to get information on the interest, attention or moods of the current user and enable an adaptation of the robot's interaction behavior. Speech synthesis and the GUI, as well as the path description generation are also placed in this layer and provide their services to the higher application levels. Above the independent skills, there is a finite state machine representing the application control which makes use of the basic features provided by the skills and is orchestrating the behaviors. These behaviors are exclusive units each representing an individual control loop for accomplishing the different task-oriented navigation and interaction functions of the robot.

5.1   Localization and Navigation

Since our operational area has four different floors, localization and path planning can not be performed using a single occupancy grid map. Instead, we use a hybrid, hierarchical topological map for path planning, which allows us to model the elevators as transitions between the different floors. On the coarsest level of this graph-based map, each node represents a single floor of the building. Each node is further subdivided into sub-nodes that represent the aisles of each floor, etc. On the finest level, the leaf nodes contain metric occupancy maps. The path planning starts on the coarsest level using a general Dijkstra algorithm and is iteratively refined up to the metric occupancy maps, where we finally apply the computationally more expensive E* path planning algorithm [PS05]. This hierarchical approach combines fast path planning on a topological map and the ability of dynamic replanning that is supported by the E* algorithm.

For localization, we use a particle filter based Monte-Carlo-Localization, that primarily operates on the metric occupancy maps stored within the leaf nodes of the hybrid topological map. For the estimation of the current floor, the robot uses an embedded accelerometer, that allows to estimate the vertical movement of the robot while using the elevators.

Due to the variety of the different behaviors, simple navigation tasks like driving to a given position are not sufficient. Instead, additional tasks, like searching and following a person, or autonomous docking to a docking station, must be supported. These different tasks must be handled without reconfiguration or restart, since all of these navigation behaviors are part of a large application. To meet all of these requirements, we have developed a modular navigation architecture [EL10]. Such a modular navigator requires a suitable interface allowing to specify complex tasks. Therefore, we introduced a task based system which allows us to define navigation jobs consisting of several sub-tasks and their corresponding parameters, such as the goal point to drive to, the preferred driving direction of the robot, the accuracy for reaching the goal, etc [GSM$^+$11]. To fulfill a navigation task, the navigator must complete each sub-task simultaneously. The sub-tasks are supported by a set of separate software modules - called objectives - that are specialized for certain tasks. The objectives are realized as software plugins, allowing new objectives to be added easily when new tasks are necessary without changing other parts of the navigator. The motion planner itself is based on the Dynamic Window Approach (DWA) [GSM$^+$11]. It combines the output of the objectives to select the most appropriate motion command that is then sent to the robot's motor controllers.

5.2 Person Tracking

The robot needs to know the position of people in its close environment. This is useful for attracting people passing by and for facing the detected persons with the pan tilt head during interaction or a guiding tour. The developed probabilistic multi-hypothesis people tracking system is based on a 6D Kalman Filter that tracks the position and velocity of people assuming an uncertain random acceleration [VMSG11]. The person tracker processes the detections of different, asynchronous observation modules – namely a laser-based leg detector, a face detector, a motion detector, and an upper-body shape detector. The leg detector is based on the boosted classifier approach of [AMB07]. The face detection system utilizes the well-known face detector of Viola&Jones [VJ02]. The motion detection cue is only active when the robot is standing still and utilizes a fast and simple image difference approach [MSSG06]. Finally, we apply an upper body shape detector based on Histograms of Oriented Gradients (HOG) [FMJZ08].

5.3 Interaction and Control

Based on the person tracking and navigation skills, the central state machine can operate closely coupled to the graphical user interface (GUI). The state machine comprises 20 states where each state is associated with one of the behaviors (see Fig. 2). Transitions between states are triggered by navigation events, person tracking events, GUI interaction, or via the administration remote interface. The behaviors are independent control loops which are realizing either standing still and facing the detected persons with the pan tilt head, or any driving activities. Examples are driving to a desired target with or without observing the presence of a person following, docking to the charging station and resting there, or driving remote-controlled in the admin behavior.

One central part of the search functionality is the generation of natural language path descriptions from arbitrary start points to each destination. This is based on a hand-crafted labeled directed graph, that holds sentence parts on each edge. For a way

description, the start and end node are selected based on the current position and orientation and a shortest path is generated through the graph. To find a description finally, the sentences on the way edges are concatenated by using a generic set of fill words and a random start phrase.

6 Results and Outlook

Our long-term trials have been started at the end of April 2012. The first three weeks of operation included an open house presentation during which a lot of experimental setups were occupying the hallways, and where large groups of people were passing by. Even under such circumstances, the robots were able to offer their services without problems. During the whole day, we have registered more than 150 interactions with visual and verbal path description, leading to more than 60 successful guiding events per robot. The large gap between the number of interactions and the successful guiding events is caused by the fact that some users were satisfied with the plain way description and did not start a guidance tour.

For a regular office day, when the robot typically operates for about 6 hours before it drives to its docking station, the numbers are lower. Therefore, we counted about 30 guiding tour events with 60 interactions per robot.

Table 1. Average scores (0=worst, 1=best) for the evaluation of service offered by the robot.

Question	Average value
Did you like my driving skills?	0.70
Could you use my touch screen intuitively?	0.83
Did you find my service useful?	0.61/0.84

After the service has been offered by the robot, the user is asked to answer three questions (Table 1). To this end, the GUI provides a slider for every question, which can be moved freely between a sad looking smiley and a happy smiley. The selected slider position is interpreted as a value between 0 (worst) and 1 (best). Having a closer look at the results, we noticed that some users were very unhappy with the provided usefulness of the service. This can be easily explained, as due to the navigation restriction to the ground floor these users were guided to the nearest elevator or stairs only (just a few meters away), since their goal was located on one of the upper floors. The users which were guided all the way to the target room were quite happy with the service, resulting in an average value of 0.84. Therefore, we expect these values to improve significantly as soon as the robot is allowed to drive to the upper floors.

Continuing our work, we are planning to extend the functionality of the robot step by step. The short-term objective is to enable the usage of all four floors and the elevators, as well as improving the continuous tracking of people during driving. However, the long-term objectives are to provide a person specific and adaptive guidance system, which can provide user-adaptive tours. The tasks arising from this objective comprise advanced user estimations, like gender, age, or interest in continuing the interaction. Furthermore, the robot should be more entertaining by providing some kind of small talk and advertise exhibits if they are passed on the way to the destination.

References

[AMB07] K.O. Arras, O.M. Mozos, and W. Burgard. Using boosted features for the detection of people in 2d range data. In *Proc. Int. Conf. on Robotics and Automation*, pages 3402–3407. IEEE, 2007.

[BJ05] L. Mayor A. Drygajlo R. Siegwart B. Jensen, N. Tomatis. Robots meet humans - interaction in public spaces. In *IEEE Trans. Industr. Electronics 52*, pages 1530–1546, 2005.

[EL10] E. Einhorn and T. Langner. Pilot - Modular Robot Navigation for Real-World-Applications. In *Proc. 55th Int. Scientifc Colloquium, Ilmenau*, pages 382–393, 2010.

[FBE+09] F. Faber, M. Bennewitz, C. Eppner, A. Gorog, C. Gonsior, D. Joho, M. Schreiber, and S. Behnke. The humanoid museum tour guide robotinho. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 891–896. IEEE, 2009.

[FMJZ08] V. Ferrari, M. Marin-Jimenez, and A. Zisserman. Progressive search space reduction for human pose estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[GBS+08] H.-M. Gross, H.-J. Böhme, Ch. Schröter, St. Müller, A. König, Ch. Martin, M. Merten, and A. Bley. Shopbot: Progress in developing an interactive mobile shopping assistant for everyday use. In *IEEE Int. Conf. on Systems, Man and Cybernetics (IEEE-SMC 2008)*, pages 3471–3478. IEEE, 2008.

[GBS+09] H.-M. Gross, H.-J. Böhme, Ch. Schröter, St. Müller, A. König, E. Einhorn, Ch. Martin, M. Merten, and A. Bley. Toomas: Interactive shopping guide robots in everyday use - final implementation and experiences from long-term field trials. In *IEEE/RJS Intern. Conf. on Intelligent Robots and Systems*, pages 2005–2012. IEEE, 2009.

[GSM+11] H.-M. Gross, Ch. Schröter, S. Müller, M. Volkhardt, E. Einhorn, A. Bley, Ch. Martin, T. Langer, and M. Merten. I'll keep an eye on you: Home robot companion for elderly people with cognitive impairment. In *IEEE Int. Conf. on Systems, Man and Cybernetics, Anchorage, USA*, pages 2481–2488. IEEE, 2011.

[MG08] Ch. Martin and H.-M. Gross. A real-time facial expression recognition system based on active appearance models using gray images and edge images. In *8th IEEE Int. Conf. on Face and Gesture Recognition*, paper no. 299. IEEE, 2008.

[MIR] Website: www.mira-project.org.

[MSSG06] Ch. Martin, E. Schaffernicht, A. Scheidig, and H.-M. Gross. Sensor Fusion using a Probabilistic Aggregation Scheme for People Detection and People Tracking. *Robotics and Autonomous Systems*, 54(9):721–728, 2006.

[NK08] K. Niechwiadowicz and Z. Khan. Robot based logistics system for hospitals-survey. Technical Report 30, Mälardalen University Sweden, 2008.

[PS05] R. Philippsen and R. Siegwart. An Interpolated Dynamic Navigation Function. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 3782–3789, 2005.

[SGTJ01] R.D. Schraft, B. Graf, A. Traub, and D. John. A mobile robot platform for assistance and entertainment. *Industrial Robot: An International Journal*, 28(1):29–35, 2001.

[TGVS01] N. Tschichold-Gürman, SJ Vestli, and G. Schweitzer. The service robot mops: First operating experiences. *Robotics and Autonomous Systems*, 34(2):165–173, 2001.

[VJ02] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 57(2):137–154, 2002.

[VMSG11] M. Volkhardt, S. Mueller, Ch. Schroeter, and H.-M. Gross. Playing Hide and Seek with a Mobile Companion Robot. In *Proc. 11th IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, pages 40–46, 2011.