Proceedings of the ASER '03
1st International Workshop on Advances in Service Robotics
Bardolino, Italy, March 13-15, 2003

Design, Implementation and Exploitation of a New Fully Autonomous Tour Guide Robot

R. Siegwart[†], K.O. Arras[†], B. Jensen[†], R. Philippsen[†], N. Tomatis[‡]

[†]Autonomous Systems Lab, EPFL
Swiss Federal Institute of Technology Lausanne
CH-1015 Lausanne
roland.siegwart@epfl.ch

[‡]BlueBotics SA
PSE-C
CH-1015 Lausanne
nicola.tomatis@bluebotics.com

Abstract

This paper presents the genesis of a tour guide robot, which has been built from the scratch based on the experience of the Autonomous Systems Lab. The production of 11 of those machines has been realized by a spin-off of the lab: BlueBotics SA. The goal was to maximize the autonomy and interactivity of the mobile platform while ensuring high robustness, security and performance. The result is an interactive moving machine named RoboX. RoboX can operate in human environments and interacts by seeing humans, talking to and looking at them, showing icons and asking them to answer its questions. The complete design of mechanics, electronics and software is presented in the first part. Then, as extraordinary test bed, the Robotics exhibition at Expo.02 (Swiss National Exhibition) permits to establish meaningful statistics over five months (from May 15 to October 20, 2002) with up to 11 robots operating at the same time.

1. Introduction

Interacting robots have to acquire the attention of the visitors. As tour guide robots they further have to move around in the environment autonomously in order to fulfill their main goal: give the visitors the pre-defined tour. The environment is known and accessible, but a general approach requiring no environmental changes is better suited for a commercial product. For the same reason a fully-autonomous and self-contained robot is preferable. Other fundamental characteristics are long term reliability and minimal human supervision.

2. Related Work

The tour-guide robot task can be decoupled in two separate issues: navigation and interaction.

Navigation: A limited number of researchers have demonstrated autonomous navigation in exhibitions or museums [4], [14], [17], [9] and [18]. Furthermore, most of these systems have still some limitations in their navigation approaches. For instance *Rhino* [4] and *Minerva* [17] have shown their strengths in museums for one week, 19 kilometers and two weeks, 44 kilometers respectively. However, their navigation has two major drawbacks: it relies on off-board resources, and due to the use of raw range data for localization and mapping it is sensible to environmental dynamics. *Sage* [14], *Chips*, *Sweetlips*, *Joe* and *Adam* [18], use a completely different approach for permanent installations in museums: the environment is changed by adding artificial landmarks to localize the robot. This approach performed well, as shown with a total of more than half a year of operation and 323 kilometers for *Sage* [14] and a total of more than 3 years and 600 kilometers for *Chips*, *Sweetlips*, *Joe* and *Adam* [18]. However their movements, but for *Adam*, are limited to a predefined set of unidirectional safe routes in order to simplify both localization and path-planning. Another robot permanent installation which is operating since March 2000, is presented in [9]. Three self-contained mobile robots navigate in a restricted and very well structured area. Localization uses segment features and a heuristic for matching.

Interaction: Human-centered and social interactive robotics is a comparatively young field in mobile robotic research. However, several experiences where untrained people and robots meet are available. The analysis of the first public space experience with *Rhino* [4] underlines the importance to improve human-robot interfaces in order to ease the acceptance of robots by the visitors. In [17] *Minerva* attracted visitors and gave tours in a museum. It was equipped with a face and used an emotional state machine with four states to improve interaction. The *Mobot Museum Robot Series* [14] and [18] focused on the interaction. Robustness and reliability was identified as an important part of a public robot. The permanent installation at the *Deutsches Museum für Kommunikation* in Berlin [9], uses three robots which have the task to welcome visitors, offer them exhibition-related information and to entertain them.

The system presented here is designed to offer enhanced interactivity and autonomous navigation with a completely self-contained robot and without requiring changes of the environment. Furthermore it is intended to work permanently with minimal supervision.

3. Product Design

The specification of the mobile platform is:

- Highly reliable and fully autonomous navigation in unmodified human-environments crowded with hundreds of humans.
- Bidirectional multi-modal interaction based on speech (English, German, French and Italian), facial expressions and face tracking, icons (LED matrix), input buttons and robot motion.
- Safety for humans and objects all the time.
- Minimal human intervention and simple supervision.

The esthetic of the robot has been designed in collaboration with exhibition makers, industrial designers and scenographers. The result of the design of both hardware and software is RoboX: a mobile robot platform ready for the real world.
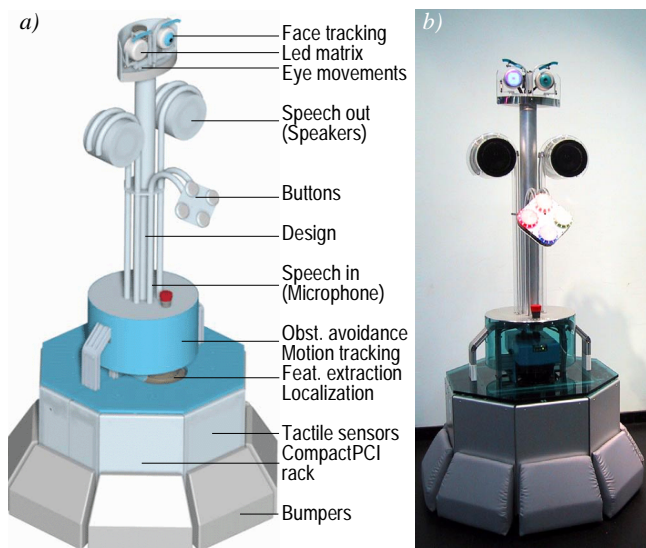

Figure 1: a) Functionality of the tour guide robot RoboX. b) An image of RoboX 9.

3.1 Mechanical Design

The lower part of the robot (base) consists mainly in the batteries, the CompactPCI rack with two control computers, the laser range sensors (two SICKs LMS-200), the bumpers and the differential drive actuators with harmonic drives. The base (figure 3) has an octagonal shape with two actuated wheels on a central axis and two castor wheels. In order to guarantee good ground contact of the drive wheels, one of the castor wheels is mounted on a spring suspension. This gives RoboX an excellent manoeuvrability and stability against tipping over in spite of its height of 1.65 m.

The upper part of the robot incorporates the interaction modules of the robot. The face includes two eyes with two independently actuated pan-tilt units and two mechanically coupled eyebrows. The left eye is equipped with a color camera for face tracking. The right eye integrates a LED matrix for the display of symbols and icons. The eyebrows further underline facial expressions with a rotational movement. Behind the face, a gray scale camera pointing to the ceiling is mounted for localization purposes.

The central input device for establishing a bidirectional communication with the humans are four buttons that allow the visitors to select the language, reply to questions the robot asked, and to perform other types of actions. The robot can further be equipped with a directional microphone matrix for speech recognition even though this seems too challenging in the very noisy environment of an exhibition.

3.2 Electronic Design

The control system (figure 2) has been designed very carefully by keeping in mind that the safety of the humans and the robot has to be guaranteed all the time. It is composed of a CompactPCI rack containing an Intel Pentium III card and
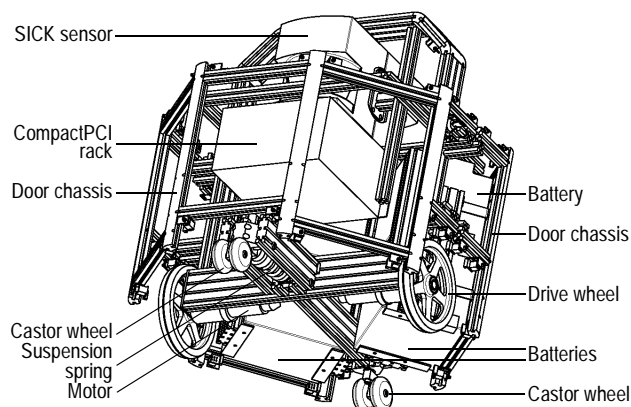

Figure 3: Mechanical design of RoboX base.

a Motorola PowerPC 750 card. The latter is connected by the PCI backplane to an analogue/digital I/O card, a Bt848-based frame grabber, an encoder IP module and a high bandwidth RS-422 IP module. Furthermore a Microchip PIC processor is used as redundant security system for the PowerPC card (figure 2).

Navigation is considered as safety-critical and is therefore running on the hard real-time operating system XO/2 [3] installed on the PowerPC. This processor has direct access to the camera looking at the ceiling, the two SICK sensors, the tactile plates and the main drive motors. It communicates with the interaction PC through Ethernet via an on-board switch.

Interaction is not considered safety-critical. It is running under Windows 2000 on an industrial PC. This allows using commercial off-the-shelf (COTS) software for speech synthesis and recognition, and makes scenario development easier. The PC has direct access to the eye camera, the eye and eyebrow controller, the input buttons, the microphone and the two loudspeakers.

The whole robot is connected by a radio Ethernet to a supervision computer which allows to track the status of the robot at any time on a graphical interface. However, RoboX operates fully autonomously, the wireless connection is thus not security relevant.
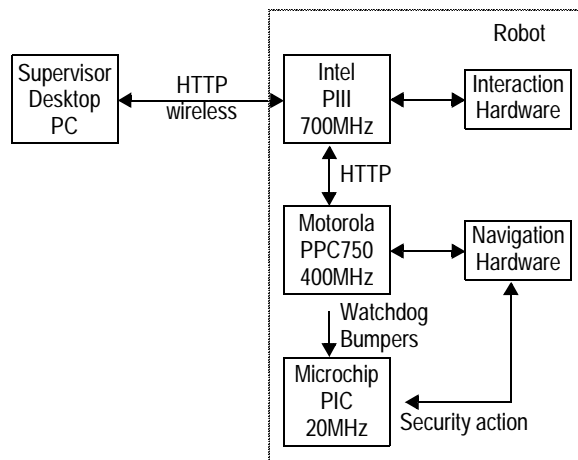

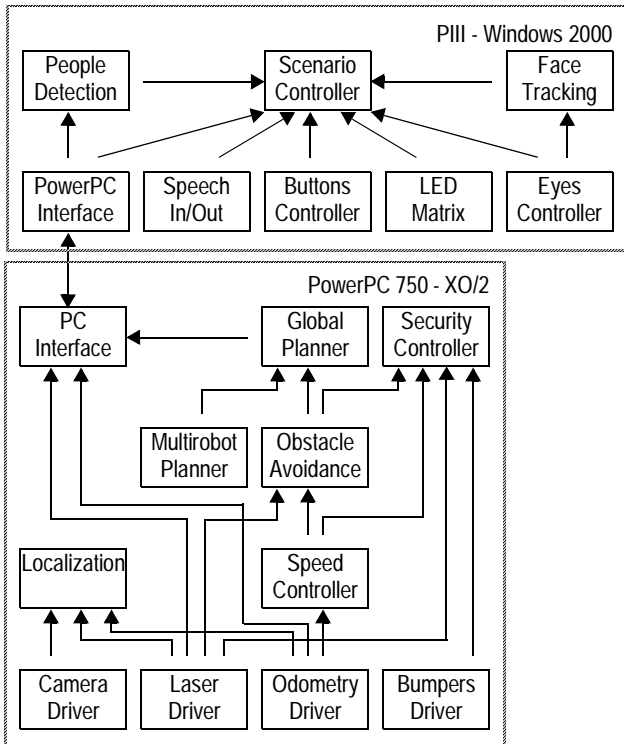*Figure 2: Simple scheme of the electrical design*

Figure 4: Software architecture on the distributed embedded system.

3.3 Software Design

As explained in the section above, the robot is composed of both an Intel Pentium and a Motorola PowerPC system. The software has been designed without taking into account this fact based on the functionality which was to be developed. However, as soon as the implementation started, the objects have been assigned to one of the two distributed embedded systems. For hardware related objects (mainly sensor drivers) the choice was obvious. For the others, their relevance to safety has been evaluated: due to the hard real-time characteristics of XO/2, all the time-critical objects in relation with the security have been implemented on the Power-PC. Objects requiring COTS components have been implemented on the Windows machine because of their wider availability (f.e. MBrola for speech out, small FireWire camera in the eye for face tracking, etc.).

The resulting object distribution is represented in figure 4. Tasks whose failure could cause injuries to people or damage objects required special attention during design. Software watchdogs are therefore implemented for the speed controller, the obstacle avoidance, the laser driver and the bumpers driver (figure 4). Failure of one of these tasks is detected by the security controller which then either restarts the failed task or stops the robot, turns on the alarm blinker and sends an e-mail to the maintenance. This permit to centralize the control of the security and to refer to a single object if a problem occurs. Furthermore, the security controller generates a watchdog signal on a digital output permitting to know if both the operating system and the security controller are still running.

The above mentioned software permits to have a consistent

control system running on the PowerPC. However, this isn't enough to guarantee the security of the robot and its surrounding. Even in case of failure of the electronics or problems on the operating system of the PowerPC, the robot must remain un-dangerous. For this, the robot has a third processor: a Microchip PIC (figure 2). The software running on it checks the watchdog generated by the security controller, awaits acknowledgements from the security controller for each bumper contact and guarantees that the pre-defined maximal speed is never exceeded. If one of these conditions is not respected the redundant security software running on the PIC safely stops the robot (it shortcuts the phases of the motors) and puts it in emergency mode (acoustic alarm).

The central object of the interaction subsystem is the scenario controller which accesses all the other objects. A *scenario* is a sequence of tasks from all modalities (speech, face expression, motion, LED matrix, etc.). A sophisticated tour-guide scenario consists of several small scenarios which are played by the scenario controller. The software for scenario creation is a stand-alone application with a user-friendly graphical interface.

4. Application Software

In this section, the most innovative techniques implemented on RoboX are briefly presented.

4.1 Navigation

Map

The map of the environment is a graph-like structure with nodes representing $[x, y, \theta]^T$ positions the robot has to reach in order to perform a certain task. This graph is therefore used for path-planning. Furthermore it contains the information about all the features in the environment. This permits to calculate which feature is visible from the current position of the robot.

Path Planning

On RoboX, three path planning algorithms are used. They work on different levels of abstraction and take sensor readings into account in varying degrees. The topmost layer is the graph-based global planner. It is based on the above mentioned graph structure where nodes are locations of interest (e.g. a showcase, a docking station) and edges denote traversability between locations. The planner employs a depth-first search and generates a length-optimal path. Since the path is global and no sensor readings are taken into account, dynamic path modification cannot be treated on this level.

The second layer of path planning uses the NF1 navigation function in a local grid around the robot [13]. It can thus take into account the current sensor readings and is not limited to nodes of the a-priori map. However, the paths generated by NF1 have a very poor geometry, consisting of linear segments that lie on angles which are multiples of 45°. Another disadvantage is their tendency to graze obstacles.

Smoothing the path and adapting it to dynamic surroundings is done in the third layer of path planning. It is based on the elastic band [15]. The initial plan, generated by the NF1, evolves toward a smoother curve (a list of via points) as long

as the elastic band does not "snap". In case dynamic obstacles move in such a way that the minimum clearance along the path cannot be maintained, or if the path lengthens beyond a reasonable amount, the NF1 is called upon again to re-initialize the path.

Obstacle Avoidance

The actual real-time obstacle avoidance task is based on the dynamic window method [8], which allows to:

- Take into account the actuator limits of the robot (speed which could result in later collisions are not allowed, motion commands never exceed the robot's speed or acceleration limits).
- Take into account the "exact" robot shape as represented by a convex polygon (extension to any polygon can be done by decomposition).

In comparison to the original dynamic window publications [8], two adaptations have been made:

- Instead of using the distance travelled before hitting an obstacle, the time until collision is used. This solves a singularity when the robot is turning on the spot (any collisions would seem instantaneous because the distance travelled seems zero). It also means the robot will choose more clearance when travelling at higher speeds.
- The objective functions for speed, heading, and clearance are calculated on the actuator phase space (v_l, v_r) instead of the usual (v, ω). Actuator limits are thus more directly taken into account.

The dynamic window task is executed with a frequency of 10Hz. It is part of the time- and security-critical processes on RoboX. Special attention has therefore been paid to optimize its execution time, which should be short and predictable. Both issues are addressed by the use of look-up-tables [16]: Their fixed size give an upper bound to the number of operations, and the intensive calculations (intersecting circles with line segments) can be done once at the initialization step. Adopting look-up-tables means large memory usage, especially when storing floating point values. This problem has been addressed by compressing the tables, using a Lloyd-Max quantizer. The compression is handled transparently, a fact that was facilitated by the object oriented design philosophy underlying our navigation software.

Localization

While autonomous guided vehicles (AGVs) usually employ for their navigation expensive and inflexible environment modifications such as floor tracks or retro-reflective beacons, nowadays localization approaches are ready for *unmodified* environments.

The new localization system [2] employed here, takes advantage of experience from earlier work [1] gathered over a five year period and more than hundred kilometers travelled distance. The method is a global feature-based multi-hypothesis localization using the Kalman filter as estimation framework. It overcomes limitations of the single-hypothesis Kalman filter [6], since the data association problem is explicitly addressed. The robot cannot get lost anymore, as it was possible before, while preserving typical advantages of feature-based approaches, such as very high localization ac-

curacy and an efficient implementation.

The technique which provides this property is a constrained-based search in an interpretation tree [5], [2]. This tree is spanned by all possible local-to-global associations, given a local map of observed features and a global map of model features. The same search is consistently employed for hypothesis generation and pose tracking.

4.2 Interaction

SOUL

SOUL aims at composing the scenarios like a theater or a music composition. It allows to combine different basic behaviors with synthesized speech, motion, sensory inputs and much more, and to supervise its execution.

In general we distinguish between static scenarios that are usually a fixed sequence of the tour, and dynamic scenarios that can be considered behaviors triggered by special events, e.g. if the visitors are blocking the way of the robot. If a dynamic scenario is triggered, SOUL will interrupt the current static scenario and execute a corresponding exception scenario telling the visitor that it is aware of his actions, before resuming the tour. In this sense, dynamic scenarios are more appealing for a lot of visitors, because they demonstrate awareness of the robot.

The main input for composing the scenarios are: robot location and state (goal reached / not reached), blocked pathway (from obstacle avoidance module), bumper contact (8 bumpers surrounding the robot), 4 input buttons for visitor interaction, face tracking [10][11], (figure 5), people-motion tracking [11] (through laser sensor, figure 6b), speech input (on two robots), other robots' locations (through multi-robot coordination system), emergency button and battery status.

The output of the scenario uses: synthesized speech in English, French, German, and Italian using Mbrola [7] and LAIPTTS [12], where pitch, rate and volume of the speech output can be selected directly in the SOUL interface, robot motion (definition of next goal location, expressional motions), illumination of four input buttons (e.g. green for "yes", red for "no", or different colors for language selection), facial expressions composed using the eye and eyebrows motion and the LED matrix display, control of environment through domotic system.

Figure 5: Visitors at the Expo.02 seen by the robot. Skin colored regions show a light border. Dark borders indicate clusters that passed the heuristic filters. The face tracking algorithm is able to maintain the visitor Ids even with a moving camera and resulting out of focus images.

Composition of Scenarios

Figure 6 depicts the layout of the exhibition. Presentation-stations are defined near particular objects in the exhibitions. Presentation-stations may comprise several goal nodes, as is the case for the welcome point, thus tours can start simultaneously. Fourteen presentation-stations were located all over the exposition space (figure 6a). Finally there are goodbye stations close to the exit. Each station corresponds to one scenario in the SOUL system, providing visitors with the necessary explanatory or entertaining information. Tours can be created by a succession of several presentation stations. Two stations photo and poet robots are not included in any of the tours and are permanently occupied with a dedicated robot. The remaining 9 tour-guide robots can reach all the other stations. In order to avoid having several robots presenting at the same goal node, each robot can ask the position of all the other robots at any time. This allows each robot to dynamically select the next goal, based on visitors' preferences and availability of the stations.
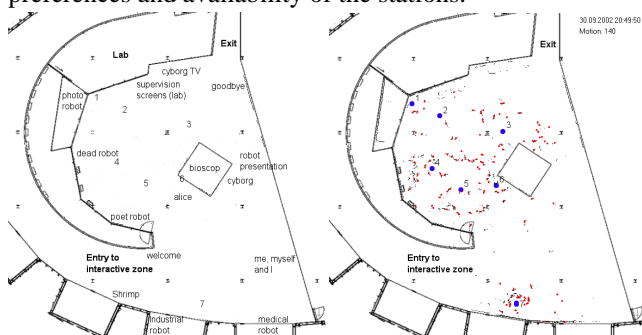

Figure 6: a) Scheme of the 315m2 exhibition area with the presentation stations of the tour. b) Snapshot of the exposition with seven robots (black dots) and the visitors as (point-clusters) found by the motion detection algorithm.

5. Experiments: *Robotics* at Expo.02

The whole RoboX development started in February 2001 after the decision of developing 10 robots for the *Robotics* exhibit at the Swiss National Exhibition in Neuchâtel. The Swiss National Exhibition takes place once in about 40 years. The current edition, Expo.02, was from May 15 to October 20, 2002. It was a major national event with 37 exhibitions opened ten and a half hours per day, seven days per week. Within the thematic *nature and artifice*, *Robotics* was intended to show the increasing closeness between man and machine. The central visitor experience was the interaction with up to 11 autonomous, freely navigating tour-guide robots on a surface of about 320 m^2.

5.1 Statistics and Reliability

By October 20, 2002, after 159 days of operation, the eleven Robox have accumulated:
- 13.313 hours of operational time,
- 9.415 hours of motion,
- 3.315 km travel distance,
- 686'000 visitors.

This is to our knowledge the biggest installation of interactive personal robots ever.

Starting with the opening of the exhibition on May 15, 2002, the robots operate fully autonomous, 7 days a week, 10.5 hours a day at the beginning, then 11 hours a day and 12 hours for the last 3 weeks. Due to some delays in the development, the software was still in the test phase during the first weeks of the exhibition (figure 7). Therefore we started with a mean time between failure (MTBF) of less then 1 hour. The MTBF starting from the fourth week up to the end of the exhibition is 4.6 hours, which results in an average of on problem per robot and day. The large majority of failures where caused by software problems in the interaction system (PC running Windows 2000) and were therefore not safety critical. The navigation system (PowerPC running real-time system XO/2) had since the beginning a MTBF in the range of 20 to 40 hours and has finally stabilized to around 60 hours. The robot hardware, apart from some problems with the amplifiers for the drive motors, was highly reliable from the beginning. We did not encounter problems with serious vandalism. However, visitors seemed to be strongly attracted by the foam bumpers and the emergency button that were very frequently tested by them.

5.2 Performance of Interaction System

The final interaction software SOUL was operational since July 1, 2002. With five stand-by scenarios, robots ran out of those scenarios less than once a week. Visitors stayed between 10 and 45 minutes with the robots. We tried to control this by changing the tour length from two to ten stations without noticing an impact on the visitor's stay. People just moved on to the next robot or even stayed with the current one. Here, enhanced environmental information, like motion information of the visitor or face recognition might help creating more convincing scenarios. We found that visitors quit a robot approximately after four stations, which was the actual tour length. The average number of visitors after installing the global supervisor (visitor flow control and robot coordination) rose slightly to around 4.600 per day. This makes it hard to prove a quantitative effect on the visitor flow. However, observation of the crowd shows that visitors appreciated having the choice to go to a station. This adds a little interactive element to the tour.

5.3 Visitors' Experience and Feedback

The visitor's experience is in general very positive, with more than 83% of the visitors rating the exhibition as very good or good and less than 5% of the visitors rating the exhibition as bad. However, we encountered various problems with the first concept of the exhibition:
- Guiding visitors at public exhibition by a robot is somewhat difficult, because playing and interacting with the robot seems to be more attracting to the visitors. Some visitors were also not very patient and not willing to follow the instructions of the robot.
- Due to the large number of visitors and robots sharing the exhibition hall, it was sometimes difficult to understand the artificial voice of the robot.
- The basic goal of the exhibition was to experience the increasing closeness between man and machine and not to present technical details of the robot. However, plenty

of visitors were eager to get some insight on the robots. We therefore added a station at which the robot was presenting his functionality with the help of a PowerPoint presentation.

- Some visitors were disappointed about the performance and intelligence of the robot. Their overrated expectations are mainly coming from science fiction movies and popular science presentations not really reflecting the state-of-the-art.

In order to quantify the visitors' appreciation and perception of the exhibition we made a questionnaire, which has been filled in by 206 visitors. Some principal results are summarized below:

- In a first question the visitors were asked to judge on the general experience, like amusement value or interactivity of the robot. Multiple answers were possible. 60% of the visitors judged the exhibition interesting and 40% amusing, and only 4% were bored by the exhibition. However, only about 12% of the visitors perceived the robots as really interactive. This might be due to the fact that we had to limit the interactivity in order to reach a reasonable visitor flow. However, natural and easy to understand multi-modal interaction is still an open research issue, which need further investigation.
- The functionality of the 4 input buttons was easily learned by 66% of the visitors through their interaction with the robot, 21% learned by imitation of other visitors. Only about 13% needed support from the staff or did not understand the functionality at all.
- Around 75% of the people were able to follow the spoken explanations of the robot very well. This is also verified by the fact that 76% of the visitors learned from the robot that it is using a laser for navigation and detection of its environment.

The appearance and character of the robot was appreciated very much by the visitors.

This also explain the fact that over 70% of the people would not hesitate to ask the robot for information or help if it would offer his service in a supermarket or railway station.
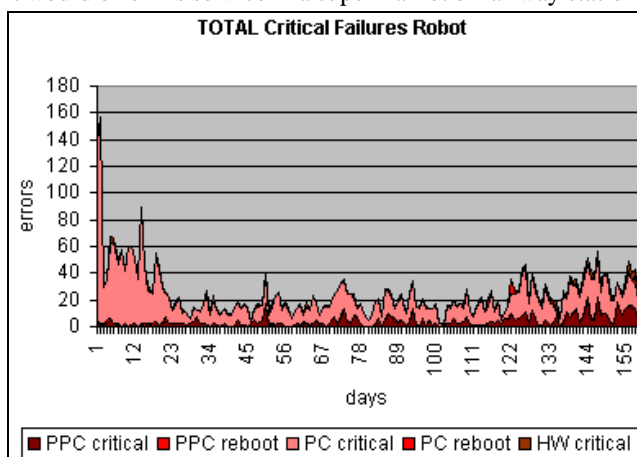

Figure 7: In case of critical errors the robot requires human intervention to continue its task. Sometimes errors require the reboot of the navigation system. Hardware errors also cause critical interventions.

6. Conclusions and Outlook

This project represents a milestone in the field of mobile robotics: for the first time tour-guide robots are produced (11 robots) and used for long time (five months) as real products instead of prototypes as in former projects. The paper presents their characteristics first, then goes into details about the used innovative approaches. The experiments section is dedicated to the results at the *Robotics* exposition in Neuchâtel, Switzerland, which is to our knowledge the biggest installation of interactive personal robots ever.

References

[1] Arras, K. O., N. Tomatis, B. Jensen, and R. Siegwart (2001). "Multisensor On-the-Fly Localization: Precision and Reliability for Applications." Robotics and Autonomous Systems 34(2-3): 131-143.

[2] Arras, K. O., J. A. Castellanos, and R. Siegwart (2002). Feature-Based Multi-Hypothesis Localization and Tracking for Mobile Robots Using Geometric Constraints. IEEE International Conference on Robotics and Automation, Washington DC, USA.

[3] Brega, R., N. Tomatis, K. Arras, and R. Siegwart (2000). The Need for Autonomy and Real-Time in Mobile Robotics: A Case Study of XO/2 and Pygmalion. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000), Takamatsu, Japan.

[4] Burgard, W., A. B. Cremers, et al. (1999). "Experiences with a Interactive Museum Tour-Guide Robot." Artificial Intelligence 00(1999): 1-53.

[5] Grimson W.E.L., Lozano-Pérez (1987). "Localizing Overlapping Parts by Searching the Interpretation Tree." IEEE Trans. on Pattern Analysis and Machine Intelligence 9(4): 469-82.

[6] Crowley, J. L. (1989). World Modeling and Position Estimation for a Mobile Robot Using Ultrasonic Ranging. IEEE International Conference on Robotics and Automation, Scottsdale, AZ.

[7] Dutoit, T., et al, The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes. ICSLP'96.

[8] Fox, D., W. Burgard, et al. (1997). "The Dynamic Window Approach to Collision Avoidance." IEEE Robotics & Automation Magazine: 23-33.

[9] Graf, B., R. D. Schraft, et al. (2000). A Mobile Robot Platform for Assistance and Entertaiment. International Symposium on Robotics, Montreal, Canada.

[10] Hilti, A., I. Nourbakhsh, B. Jensen, and R. Siegwart (2001). Narrative-level Visual Interpretation of Human Motion for Human-robot Interaction. IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, Hawaii.

[11] Jensen, B., G. Froidevaux, X. Greppin, A. Lorotte, L. Mayor, M. Meisser, G. Ramel, The interactive autonomous mobile system RoboX. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002, Switzerland.

[12] Keller, E. et al, A timing model for fast French. York Papers in Linguistics, University of York(96) pp.53-75.

[13] Latombe, J.-C. (1991). Robot motion planning. Dordrecht, Netherlands, Kluwer Academic Publishers.

[14] Nourbakhsh, I., J. Bodenage, et al. (1999). "An Effective Mobile Robot Educator with a Full-Time Job." Artificial Intelligence 114(1-2): 95-124.

[15] Quinlan, S. and O. Khatib (1993). Elastic bands: connecting path planning and control. IEEE International Conference on Robotics and Automation.

[16] Schlegel, C. (1998). Fast local obstacle avoidance under kinematic and dynamic constraints for a mobile robot. IEEE International Conference on Intelligant Robots and Systems, Victoria, B. C., Canada.

[17] Thrun, S., M. Beetz, et al. (2000). "Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva." International Journal of Robotics Research 19(11): 972-99.

[18] Willeke, T., C. Kunz, et al. (2001). The History of the Mobot Museum Robot Series: An Evolutionary Study. Florida Artificial Intelligence Research Society (FLAIRS), Florida.