To appear in IEEE Robotics and Automation Magazine, special issue on "EU-funded Projects in Robotics", in press.

TOURBOT and WebFAIR: Web-operated Mobile Robots for Tele-Presence in Populated Exhibitions

Panos Trahanias[1,2], Wolfram Burgard[3], Antonis Argyros[1],
Dirk Hähnel[3], Haris Baltzakis[1,2], Pattrik Pfaff[3], Cyrill Stachniss[3]

[1]Foundation for Research and Technology - Hellas (FORTH)
Institute of Computer Science
GR-711 10 Heraklion, Crete, Greece

[2] Department of Computer Science
University of Crete
GR-714 09 Heraklion, Crete, Greece

[3] Computer Science Department
University of Freiburg
Freiburg, Germany

Abstract

This paper presents developments and results from two recent projects, namely TOURBOT and WebFAIR, funded by the Commission of the European Union (EU). Within these projects techniques that facilitate mobile robots to be deployed as interactive agents in populated environments, such as museum exhibitions or trade shows, have been developed. The mobile robots can be tele-operated over the Internet and, this way, provide remote access to distant users. The developed robotic systems have been installed and operated in the premises of various sites. Use of the above techniques, as well as appropriate authoring tools, has resulted in drastic reduction of the installation times. Additionally, the systems were thoroughly tested and validated in real-world conditions. Such demonstrations ascertain the functionality and reliability of our methods and provide evidence as of the operation of the complete systems.

1 Introduction

Mobile robotic technology and its application in various sectors is currently an area of high interest. Research in this field promises advanced developments and novelties in many aspects. Applications of mobile robotic technology in public spaces can be found in a field that we informally term “robots in exhibitions”. In this context, robots can offer alternative ways for interactive access to exhibition spaces.

Two recent EU funded projects, namely TOURBOT (www.ics.forth.gr/tourbot) and WebFAIR (www.ics.forth.gr/webfair), address the above goal. TOURBOT started January 2000 and ended successfully February 2002; it pursued the development of an interactive tour-guide robot able to provide individual access to museums’ exhibits over the Internet. The results of TOURBOT were demonstrated through the installation and operation of the system in the real environment of the three museums that participated in the project as well as other interested organizations. WebFAIR started December 2001 and will end May 2004. WebFAIR builds on TOURBOT results and extends relevant developments to the more demanding environments of trade shows. Additionally, WebFAIR introduces tele-conferencing between the remote user and on-site attendants and employs a multi-robot platform, facilitating thus simultaneous robot control by multiple users. Factual information about the two projects is summarized in Table 1.

In order to cope with the requirements imposed by the specific application areas of TOURBOT and WebFAIR, a number of scientific, technological and technical issues need to be addressed. In this paper, we present selected techniques developed in the above mentioned projects. Among them are techniques for mapping large environments, an obstacle-avoidance technique that relies on laser-vision fusion to detect objects that are invisible to the laser scanner, and an approach to filter out range measurements coming from moving persons in the process of map construction.

A very important factor for the acceptance of a robotic tour-guide by the broader public is the degree to which the system smoothly interacts with both web and on-site visitors. For this reason, in both projects special emphasis was put in the development of appropriate user interfaces. In this paper, we also describe aspects of the developed interfaces, such as a speech interface for on-site users and a flexible web-interface with enhanced visualization capabilities for remote users. Additionally, we report on the demonstration events that took place in the framework of TOURBOT and argue on the drastic reduction of the system set-up time that was achieved.

Table 1. Factual Data of the TOURBOT and WebFAIR Projects.

Project: TOURBOT, Contract: IST-1999-12643	
Funding: Fifth Framework Programme, Information Society Technologies (IST)	
Start Date: Jan. 1, 2000	End Date: Feb. 28, 2002
Budget: 1,718,150 Euro	EU Contribution: 1,088,790 Euro
Consortium - Participant Roles	
Foundation for Research and Technology - Hellas (Greece)	Co-ordinator / Technology Provider
University of Freiburg (Germany)	Technology Provider
Foundation of the Hellenic World (Greece)	Technology Broker / End User
University of Bonn (Germany)	Technology Provider
THEON Robotic Systems (Greece)	Technology Provider
Deutsches Museum Bonn (Germany)	End User
Byzantine and Christian Museum of Athens (Greece)	End User
Project: WebFAIR, Contract: IST-2000-29456	
Funding: Fifth Framework Programme, Information Society Technologies (IST)	
Start Date: Dec. 1, 2001	End Date: May 31, 2004
Budget: 2,331,142 Euro	EU Contribution: 1,269,970 Euro
Consortium - Participant Roles	
ACTION Public Relations Hellas LTD (Greece)	Co-ordinator / End User
Foundation for Research and Technology - Hellas (Greece)	Scientific Co-ordinator / Technology Provider
RATIO Consulta SPA (Italy)	Technology Provider
IDEASIS Ltd (Greece)	Technology Provider
University of Freiburg (Germany)	Technology Provider
Polytechnic University of Madrid (Spain)	Technology Provider
Ente Fiere Castello di Belgioioso e Sartirana (Italy)	End User

2 Robots in Exhibitions

Although the notion of “robots in exhibitions” is relatively new, there have been various other research efforts targeting at systems with goals that are similar to the goals of TOURBOT and WebFAIR. In this section we present a brief overview of such approaches and we report on the main characteristics of the approach adopted in TOURBOT and WebFAIR systems.

2.1 Related work

Over the last decade, a variety of service robots were developed that are designed to operate in populated environments. Example cases are robots that are deployed in hospitals [14], museums [5, 17, 26], trade-fairs [20], office buildings [1, 24], and department stores [7]. In these environments the mobile robots perform various services, e.g., deliver, educate, entertain [22] or assist people [21, 15].

In addition, a variety of Web-based tele-operation interfaces for robots have been developed over the last years. Three of the earlier systems are the Mercury Project, the “Telerobot on the Web”, and the Tele-Garden [8, 9, 25]. These systems allow people to perform simple tasks with a robot arm via the Web. Since the manipulators operate in prepared workspaces without any unforeseen obstacles, all movement commands issued by a Web user can be carried out in a deterministic manner. Additionally, it suffices to provide still images from a camera mounted on the robot arm after a requested movement task has been completed. The mobile robotic platforms Xavier, Rhino and Minerva [24, 5, 26] could also be operated over the Web. Their interfaces relied on client-pull and server-push techniques to provide visual feedback of the robot’s movements; this includes images taken by the robot as well as a java-animated map indicating the robot’s current position. However, their interfaces do not include any techniques to reflect changes of the environment. 3D graphics visualizations for Internet-based robot control have already been suggested by Hirukawa et al. [13]. Their interface allows Web users to carry out manipulation tasks with a mobile robot, by controlling a 3D graphics simulation of the robot contained in the Web browser.

2.2 The TOURBOT-WebFAIR approach

The goal of the TOURBOT project was the development of an interactive TOUr-guide RoBOT (TOURBOT) able to provide individual access to museums’ exhibits and cultural heritage over the Web. TOUR-

BOT operates as the user's avatar in a museum, by accepting commands over the Web that direct it to move in its workspace and visit specific exhibits. The communication network is, thus, effectively extended by the introduction of interactive, mobile robots as terminal nodes. The imaged scene of the museum and the exhibits is communicated over the Internet to a remote visitor. As a result, the user enjoys a personalized tele-presence in the museum, being able to choose the exhibits to visit, as well as the preferred viewing conditions (point of view, distance to the exhibit, resolution, etc). At the same time, TOURBOT is able to guide on-site museum visitors providing either group or personalized tours.

The successful course of TOURBOT and the vision to introduce corresponding services to the taxing case of trade fairs, resulted in launching WebFAIR. The latter, currently under development, was additionally endorsed by experts in the organization and promotion of large trade shows. Besides the TOURBOT functionality, which is now offered in a more demanding environment, WebFAIR additionally introduces tele-conferencing between the remote user and on-site attendants and employs a multi-robot platform, facilitating thus simultaneous robot control by multiple users.

A tele-operated interactive tour-guide robot requires a high degree of autonomy since it operates in a populated environment in which humans are also present. Therefore, both projects opted for the development of a safe and reliable navigation system. The robotic avatars are equipped with a series of state-of-the-art sensors that allow it to acquire information about its environment. The main sensor modalities are laser range-scanners, ultrasound sensors and cameras. The navigation system uses this sensory information to adapt the robot's internal model of the environment so as to plan the robot actions.

In order to realize the TOURBOT and WebFAIR systems, the consortium developed a multimedia Web interface that allows people to interact with the tour-guide system over the Internet. Furthermore, an on-board interface for interaction with on-site visitors of the exhibition site has been developed. Using the Web interface, people are able to tele-control the robot and to specify target positions for the system. Camera controls are used to choose the part of the exhibition the user wants to inspect in more detail. The robotic tour-guide possesses a multimedia information base providing a variety of information about the exhibition at various levels of detail. Thus, the system serves as an interactive and remotely controllable tour-guide, which provides personalized access to exhibits with a large amount of additional information.

With respect to user interfaces, the TOURBOT and WebFAIR systems use video streams to convey observed information to the user. Additionally, they provide online visualizations of their actions in a virtual 3D environment. This allows the users to choose arbitrary viewpoints and leads to significant reductions of the required bandwidth.

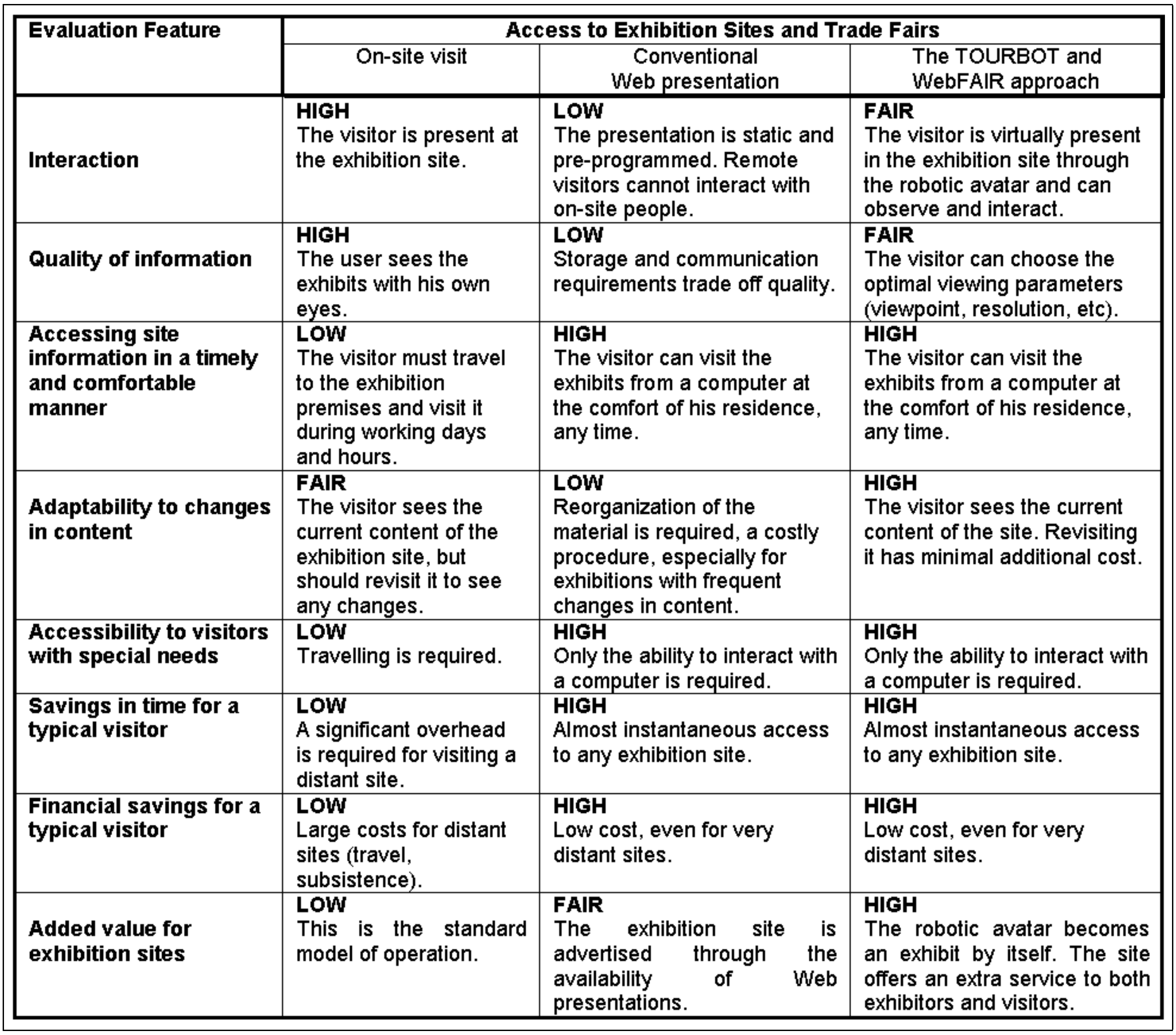
The use of tour-guide robots has several advantages for the exhibition visitor, as well as for the provider of such a service. First, the TOURBOT and WebFAIR approach allows a more effective use of exhibitions. A Web-based tour-guide robot can serve and educate people day and night, since a remotely controlled robot is not limited to the opening hours of an exhibition. Compared to static or even remotely controllable cameras for viewing parts of an exhibition, tour-guide robots can provide significantly more interaction capabilities. The mobility of the robots allows Web visitors to choose a wide variety of viewpoints and get the feeling of “being present” in the exhibition. This can be particularly advantageous for people with restricted mobility. Moreover, increased interaction capabilities with the exhibits themselves can be offered to the user, which may be useful when visiting a science or technology museum. Therefore, a scenario such as “press the red button to watch ...”, and the corresponding service offered to the user, is easily implemented. In addition to this increased interactivity, the robotic agent can deliver high-resolution images over the Web, thus being extremely beneficial to professionals and specialists.

The TOURBOT and WebFAIR concept provides visitors the ability to individually exploit the expertise stored in the tour-guide robot, which can react flexibly to the requirements of the visitors. It can, for example, offer dedicated tours on temporary focuses of the exhibition or, alternatively, give overview tours. Thus interactive tour-guide robots allow an exhibition a more flexible service and increase their offer of information. Furthermore, robots improve the flexibility for the planning of tours, especially when deployed in larger numbers. This allows people, even if they belong to a larger group, to be served individually. Additionally, robotic tour-guides contribute to the positive image of an exhibition, since they emphasize up-to-date technology and future-orientation. The latter is very important, especially for technology- and science-oriented museums.

Table 2 summarizes a comparison of the main features of the proposed approach –within TOURBOT and WebFAIR– for accessing exhibition sites and trade fairs with traditional on-site visits and conventional Web presentations. As evidenced from this table, there is indeed ground in pursuing the goals set forth in our projects.

Figure 1 provides a conceptual overview of the TOURBOT and WebFAIR systems. Many of the

Table 2. Access to Exhibition Sites and Trade Fairs.



Evaluation Feature	Access to Exhibition Sites and Trade Fairs		
	On-site visit	Conventional Web presentation	The TOURBOT and WebFAIR approach
Interaction	**HIGH** The visitor is present at the exhibition site.	**LOW** The presentation is static and pre-programmed. Remote visitors cannot interact with on-site people.	**FAIR** The visitor is virtually present in the exhibition site through the robotic avatar and can observe and interact.
Quality of information	**HIGH** The user sees the exhibits with his own eyes.	**LOW** Storage and communication requirements trade off quality.	**FAIR** The visitor can choose the optimal viewing parameters (viewpoint, resolution, etc).
Accessing site information in a timely and comfortable manner	**LOW** The visitor must travel to the exhibition premises and visit it during working days and hours.	**HIGH** The visitor can visit the exhibits from a computer at the comfort of his residence, any time.	**HIGH** The visitor can visit the exhibits from a computer at the comfort of his residence, any time.
Adaptability to changes in content	**FAIR** The visitor sees the current content of the exhibition site, but should revisit it to see any changes.	**LOW** Reorganization of the material is required, a costly procedure, especially for exhibitions with frequent changes in content.	**HIGH** The visitor sees the current content of the site. Revisiting it has minimal additional cost.
Accessibility to visitors with special needs	**LOW** Travelling is required.	**HIGH** Only the ability to interact with a computer is required.	**HIGH** Only the ability to interact with a computer is required.
Savings in time for a typical visitor	**LOW** A significant overhead is required for visiting a distant site.	**HIGH** Almost instantaneous access to any exhibition site.	**HIGH** Almost instantaneous access to any exhibition site.
Financial savings for a typical visitor	**LOW** Large costs for distant sites (travel, subsistence).	**HIGH** Low cost, even for very distant sites.	**HIGH** Low cost, even for very distant sites.
Added value for exhibition sites	**LOW** This is the standard model of operation.	**FAIR** The exhibition site is advertised through the availability of Web presentations.	**HIGH** The robotic avatar becomes an exhibit by itself. The site offers an extra service to both exhibitors and visitors.

problems in the realization of complex systems, such as TOURBOT and WebFAIR, can be solved by employing standard techniques and methods. In other cases innovative solutions have to be devised so as to provide the required functionality in a robust and efficient manner. A detailed description of the developed systems is beyond the scope of this paper. Instead, we emphasize on specific techniques developed in the context of the TOURBOT and WebFAIR projects for solving important scientific and technical problems towards realizing the corresponding robotic systems. We specifically report on techniques for

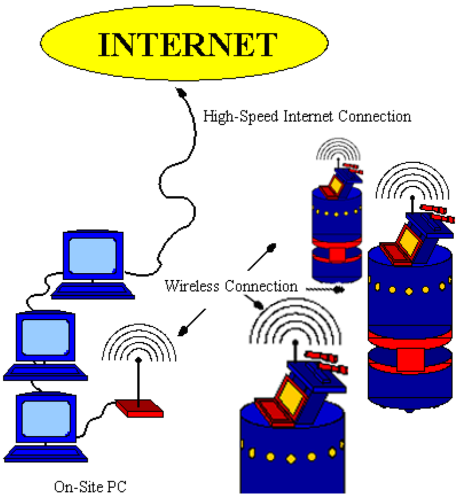


Figure 1. Conceptual overview of the TOURBOT and WebFAIR systems.

environment mapping, obstacle-avoidance, and an approach to filter out range measurements coming from moving persons in the process of map construction. We also present important aspects of the user interfaces of the developed systems.

An important aspect of EU funded projects is that they do not only finance important technical developments, but also permit the study of the maturity of technologies as well as the level of their acceptance by the broader public. Towards this end, extensive validation and demonstration trials have been performed within the TOURBOT project. They have ascertained that the developed techniques have resulted in a considerable reduction of the system set-up time, compared to predecessor systems of lower complexity. Moreover, they have facilitated the assessment and validation of the robustness and effectiveness of the system under real conditions. Finally, they have provided useful data regarding public response and attitude towards TOURBOT. This paper also reports on the experiences gained from these trials. Similar trials are also planned within the WebFAIR project for the first half of 2004.

3 Mapping

In order to navigate safely and reliably, mobile robots must be able to create suitable representations of the environment. Maps are also necessary for human-robot interaction since they allow users to direct the robot to places in the environment. Our current system uses two different mapping techniques depending on the characteristics of the environment in which the robot is deployed. In environments with a complex shape our system employs dense grid maps and operates according to a discrete Hidden Markov Model. The model is very efficient in terms of computational performance, has been proven quite robust with respect to range measurement errors and can operate regardless of the structure of the environment workspace. Its main drawback lies in the achieved localization accuracy, which is limited by the size of the grid. In order to overcome this limitation, an alternative localization method has been developed that involves a hybrid model, i.e. a combination of a discrete (Hidden Markov) model and a continuous (Kalman-filter) model [4]. This combination allows the use of Markovian dynamics for global localization and Kalman-filter based tracking for achieving increased accuracy.

Both methods offer very promising alternatives, each one with its own merits. In environments with no clearly defined structure (walls, corridors, corners, etc), the former method is more adequate, at the price of slightly decreased localization accuracy. When the environment structure becomes evident, the latter method can be employed, to render increased localization accuracy. Both mapping approaches are described in the remainder of this section.

3.1 Grid-based Mapping

This mapping technique employs occupancy grid maps. In probabilistic terms the goal of map learning is to find the map and the robot positions which yield the best interpretation of the data d_t gathered by the robot. Here the data $d_t = \{u_{0:t-1}, z_{1:t}\}$ consists of a stream of odometry measurements $u_{0:t-1}$ and perceptions of the environment $z_{1:t}$. The mapping problem can be phrased as recursive Bayesian filtering for estimating the robot positions along with a map of the environment:

$$\begin{aligned} & p(x_{1:t}, m \mid z_{1:t}, u_{0:t-1}) = \alpha \cdot p(z_t \mid x_t, m) \cdot \\ & \int p(x_t \,|\, x_{t-1}, u_{t-1}) p(x_{1:t-1}, m \,|\, z_{1:t-1}, u_{0:t-2}) dx_{1:t-1} \end{aligned} \tag{1}$$

In probabilistic mapping and localization it is typically assumed that the odometry measurements are governed by a so-called probabilistic motion model $p(x_t \mid x_{t-1}, u_{t-1})$ which specifies the likelihood that the robot is at state x_t given that it previously was at x_{t-1} and the motion u_{t-1} was measured. On the other hand, the observations follow the so-called observation model $p(z_t \mid x_t, m)$, which defines for every possible location x_t in the environment the likelihood of the observation z_t given the map m.

Unfortunately, estimating the full posterior in eq 1 is not tractable in general. One popular approach is to restrict observations to landmark detections, and to use an Extended Kalman filter for solving the SLAM-problem. Other researchers apply scan matching to compute at each step in time the most likely position of the robot and then update the map under the assumption that indeed the robot is at that particular location. Whereas this approach has the advantage that it yields accurate results in many situations, its major disadvantage lies in the greedy maximization step. When the robot has to close larger loops, this approach suffers from registration errors during loop closures and therefore tends to fail in large environments. Alternative techniques maintain a full a posterior about the position of the vehicle. To efficiently represent this posterior about robot trajectories and maps, Murphy and colleagues [16] have presented a technique using Rao-Blackwellized particle filters. The key idea of this approach is to solve the recursive Bayes filter update by the following equation:

$$\begin{aligned} p(x_{1:t}, m \mid z_{1:t}, u_{0:t-1}) = \\ p(m \mid x_{1:t}, z_{1:t}, u_{0:t-1}) p(x_{1:t} \mid z_{1:t}, u_{0:t-1}) \end{aligned} \tag{2}$$

In Rao-Blackwellized mapping, a particle filter is used to represent potential robot trajectories $x_{1:t}$. Each sample of the particle filter possesses its own map, which is conditioned on the trajectory of that particle. The importance weights of the samples are computed according to the likelihoods of the observations in the maximum likelihood map constructed using exactly the positions this particular particle has taken. The key advantage of this approach is that the samples approximate at every point in time the full posterior over robot poses and maps.

The systems described in this paper also apply a Rao-Blackwellized particle filter to estimate a posterior of the path of the robot. Our approach differs from previous techniques in that it transforms sequences of laser range-scans into odometry measurements using range-scan registration techniques [12]. These highly accurate odometry data is then used as input to the particle filter. Since the scan matching yields odometry estimates that are an order of magnitude more accurate than the raw wheel encoder data, our algorithm requires less particles than the original approach. Simultaneously, the transformation of sequences of scans into odometry measurements reduces the well-known particle deprivation problem [27], since also the number of resampling operations is reduced. In several experiments [10] it has been demonstrated that our approach results in an improved ability to map large environments.

Figure 2 shows the map obtained by our algorithm using data gathered in the entrance hall of the computer science building at the University of Freiburg. The size of the building is 15×70 meters. Also shown is the trajectory of the robot. As can be seen from the figure, the robot had to close several loops during the mapping process. Furthermore, the resulting map is highly accurate.

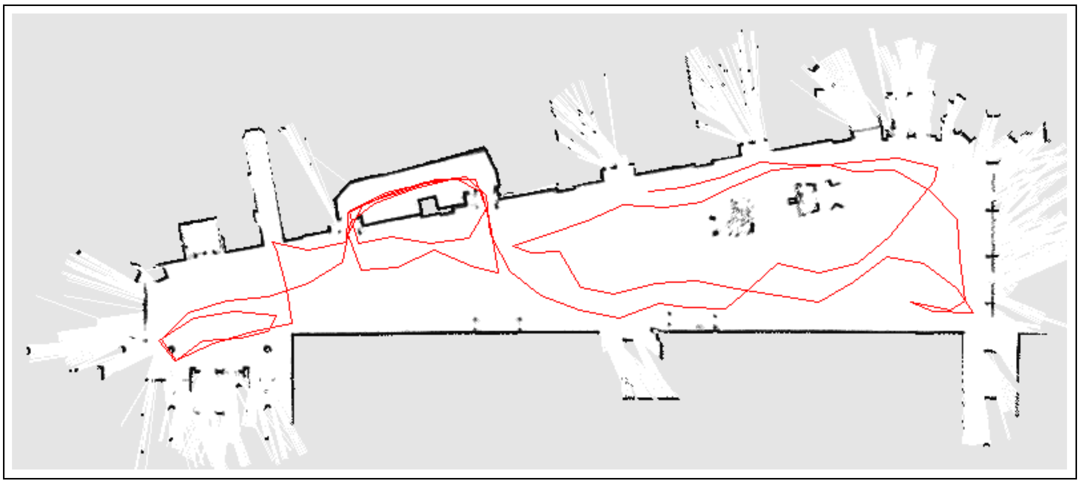


Figure 2. Map generated using our Rao-Blackwellized mapping algorithm. The size of the environment is 15×70 meters. The robot had to close several loops during the mapping process.

3.2 Feature-based Mapping

In the case of structured environments, localization accuracy can be increased by constructing and employing feature based maps of the environment. Our feature-based mapping algorithm utilizes line segments and corner points which are extracted out of laser range measurements. We apply a variant of the Iterative-End-Point-Fit algorithm to cluster the end-points of a range scan into sets of collinear points. Corner points are then computed at the intersections of directly adjacent line segments [3]. Data association (feature matching) during mapping is performed based on a dynamic programming string-search algorithm [4]. This algorithm exploits information contained in the spatial ordering of the features. Simultaneously, the dynamic programming implementation furnishes it with computational efficiency.

Existing approaches to simultaneous mapping and localization usually store all map features and the pose of the robot in a single complex random variable. This approach is computationally demanding since its space and time complexity grows quadratically in the number of features. In contrast, our approach treats map features as parameters of the dynamical system according to which the robot's state evolves. Therefore, the problem is reformulated as to simultaneously determine the state and the parameters of a dynamical system; that is, a learning problem, that is solved via a variant of the EM-algorithm.

In the mapping context, the E-step is responsible for calculating the state of the robot at each point t in time. Since the approach described here is an off-line mapping technique, all past and future observations are available and can be used to estimate the state at time t. The problem of estimating variables given both past and future observations is denoted as "smoothing". A very popular method for performing smoothing is the Raugh-Tung-Striebel smoother [19]. This algorithm consists of two steps. The first step (forward step) is the Extended Kalman Filter (EKF) forward recursions. The second step is a backward recursion. It starts with the final measurement and recursively estimates maximum a-posteriori estimates for the previous states.

To detect and close loops during mapping, our algorithm relies on the global localization capabilities of a hybrid method based on a switching state-space mode [4]. This approach applies multiple Kalman trackers assigned to multiple hypotheses about the robot's state. It handles the probabilistic relations among these hypotheses using discrete Markovian dynamics. Hypotheses are dynamically generated by matching corner points extracted from measurements with corner points contained in the map. Hypotheses that cannot be verified by observations or sequences of observations become less likely and usually disappear quickly.

Our algorithm (see Figure 3) iterates the E- and M-step until the overall process has converged or until a certain number of iterations has been carried out. Our current system always starts with a single hypothesis about the state of the system. Whenever a corner point appears in the robot's measurements, new hypotheses will be created at corresponding positions. On the other hand, hypotheses that cannot be confirmed for a sequence of measurements typically vanish. The resulting map always corresponds to the most likely hypothesis.

The right image of Figure 4 shows a typical map of an exhibition site resulting from this process; the map corresponding to the input data is shown in the left image. It has been computed within one hour on a standard personal computer. It can be verified that our algorithm was able to correctly eliminate the odometry error. Furthermore, during mapping the robot had to close serval cycles.

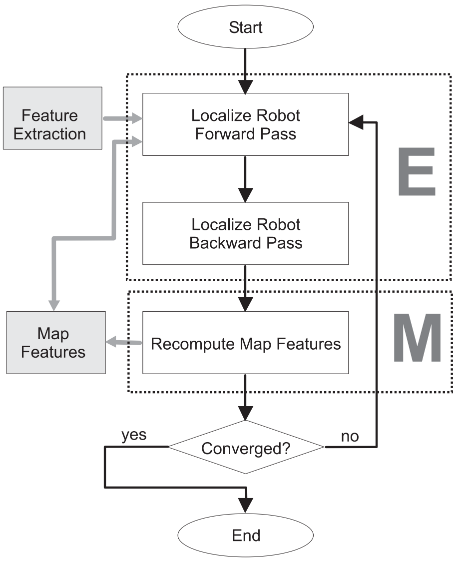


Figure 3. Flow-gram of the Iterative Mapping Algorithm.

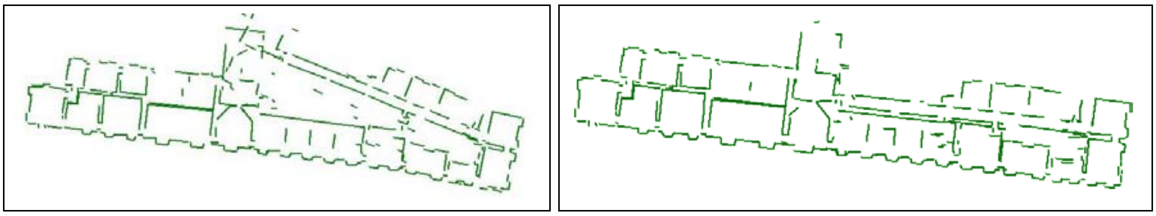


Figure 4. Line feature maps of an exhibition site: Original data (left image) and map generated by our algorithm (right image).

3.3 Mapping in Dynamic Environments

Whereas most of contemporary mapping methods are able to deal with noise in the odometry and the sensor data, they assume that the environment is static during mapping. In public spaces such as museums or trade fairs, the environment cannot be regarded as static during the robot installation time. When a robot maps its environment, dynamic objects can lead to serious errors in the resulting maps, such as spurious objects or misalignments due to localization errors. Therefore, the question of how to determine measurements corrupted by dynamic objects is an important precondition for building accurate maps in dynamic environments. The systems described in this paper address this issue through an approach that employs an EM-technique. In the E-step we compute a probabilistic estimate regarding measurements that correspond to static objects. In the M-step we use these estimates to determine the position of the robot and the map. Figure 5 shows an application of this approach to data acquired by a mobile robot in the Byzantine Museum in Athens. The figure shows in black all points classified as static. The points determined as reflected by dynamic objects are depicted as orange/grey. As can be seen from the figure, the dynamic objects are correctly identified as dynamic. Additionally, experimental results provided evidence that the resulting maps are more accurate than maps constructed without considering dynamic beams.

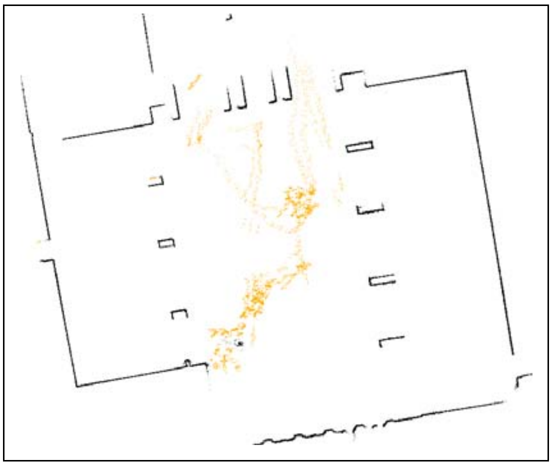


Figure 5. Map of the populated exhibition hall of the Byzantine Museum in Athens. In this map the measurements labeled as dynamic are shown in orange/grey.

4 Fusion of Laser and Visual Data

Laser scanners have proven to be very reliable sensors for navigation tasks, since they provide accurate range measurements in large angular fields and at very fast rates. However, laser range scans are 2D representations of a 3D world. Thus, the underlying assumption is that laser profiles accurately model the shape of the environment also along the vertical dimension which is invisible to the laser. Although this assumption is justified in many indoor environments, various objects such as chairs, tables, and shelves usually have a geometry that is not uniform over their height. Accordingly, not all aspects of such objects are visible in laser-range scans. The capability to correctly identify objects in the vicinity of

Figure 6. Left: image captured at time $t-1$; middle: image captured at time t; right: results of the evaluation process projected on the image captured at time t.

the robot, however, is particularly important for robots in real-world environments such as exhibitions, since it is a basic precondition for the safety of the robot and the exhibits.

One potential solution to this problem could be to exploit 3D laser scanners, which unfortunately are very expensive. The alternative solution is to utilize additional sources of information, such as vision, to infer 3D information. Our system follows exactly this approach. It uses the 2D structure acquired with the laser-range scanner and, based on this, it computes a $2\frac{1}{2}$D representation by introducing vertical planar walls for the obstacles in the 2D map. We then exploit camera information to (a) validate the correctness of the constructed model and (b) qualitatively and quantitatively characterize inconsistencies between laser and visual data wherever such inconsistencies are detected.

The employed method [2] operates as follows. At time $(t-1)$ the robot acquires a laser range scan s_{t-1} and an image i_{t-1}. Based on s_{t-1} the robot builds a $2\frac{1}{2}$D model of the environment. The same process is applied at time t resulting in i_t and s_t. Based on the world model derived at time $(t-1)$ and the motion of the robot u_{t-1}, the image i_{t-1} is back-projected to the reference frame of image i_t, resulting in the image $\hat{i}_t$. Images $\hat{i}_t$ and i_t are identical in areas where the $2\frac{1}{2}$D model is valid but differ in areas where this model is invalid. To identify obstacles not detected by the 2D range scans, we perform a local correlation of the intensity values in $\hat{i}_t$ and i_t. The inconsistencies between the laser and visual data are then converted to real word coordinates along epipolar lines. Finally, they are accumulated in a 2D occupancy map.

Figures 6 and 7 illustrate a typical application example in the corridor environment at ICS/FORTH. In this case the robot travels along a corridor with several objects that are invisible to the laser scanner,

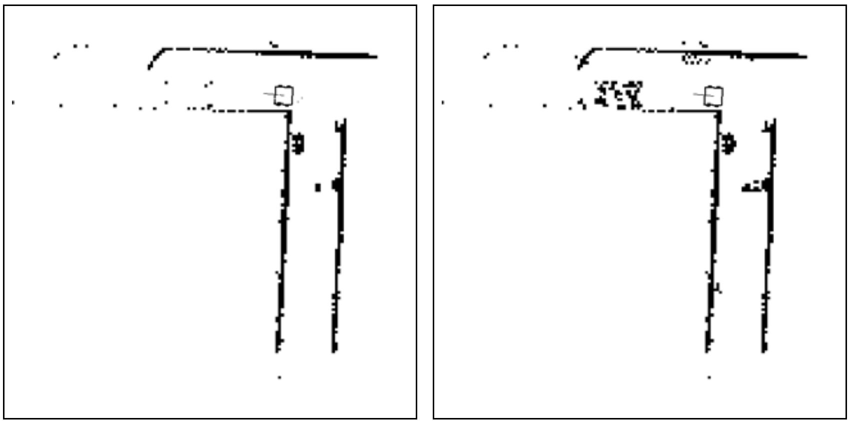


Figure 7. Occupancy grid maps computed based on the fusion of vision and laser data. The left image shows the map computed from the laser-range data. The right image shows the resulting map after combining vision and laser data.

such as tables, fire extinguishers, and a wall cabinet. Figure 6 (left and middle image) shows the images grabbed by the robot at time $t-1$ and at time t. The same figure (right image) shows the results of the obstacle detection technique. Regions with inconsistencies are marked with crosses. It can be verified that objects that cannot be detected with the range finder only, are successfully detected by the proposed technique. Finally, Figure 7 shows the occupancy grid maps obtained without considering visual information (left image) and obtained after integrating vision and laser data. The map generated by our algorithm provides a more accurate representation of the environment which can be used to prevent the robot from colliding with obstacles not visible in the range scans.

5 Interfaces

Robots in museums and exhibitions should be able to interact with on-site visitors in a natural way and to allow distant visitors to feel like being present in the site. Thus, the employment of intuitive human-robot interfaces is of paramount importance to the acceptance and the success of the overall system. The interfaces should be tailored to the type of user; clearly there are similarities as well as important differences between distant and on-site users.

5.1 Web Interface

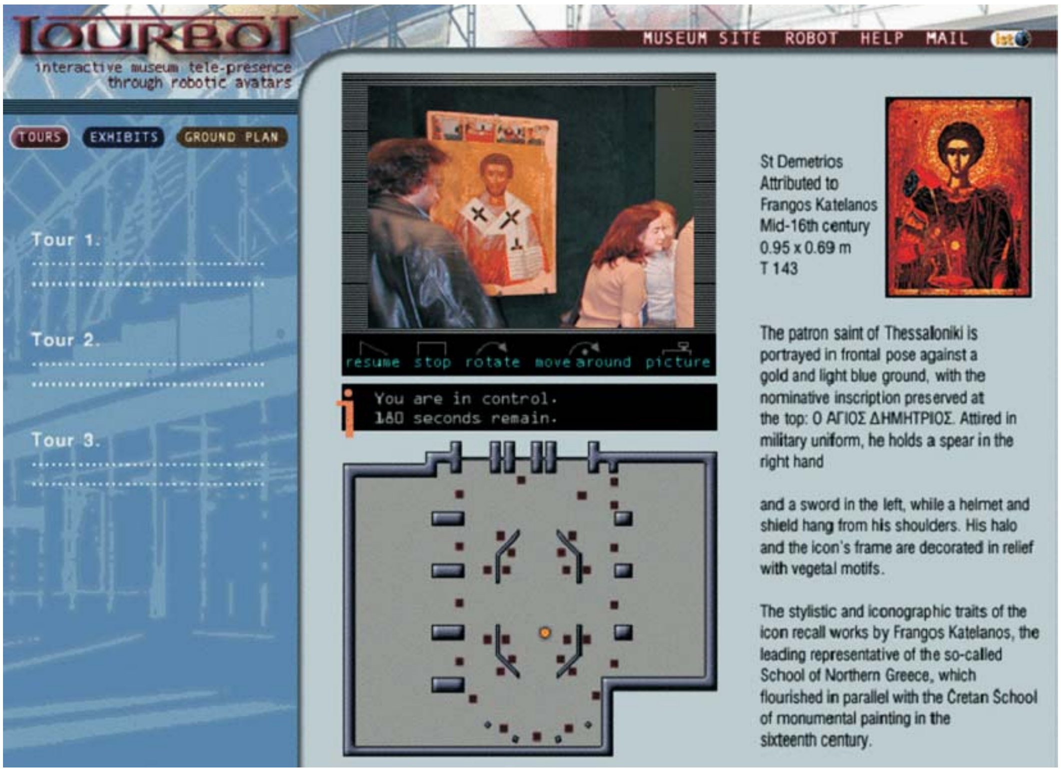
The developed web-interface has been designed to provide enhanced functionality and ease of use. Compared to interfaces of previous systems such as Xavier, Rhino and Minerva [23], it allows personalized control of the robot(s) with a number of desirable features. Instead of image streams that are updated via server-push or client-pull technology, it uses a commercial live streaming video and broadcast software [28] that provides continuous video transmissions to transfer images recorded with the robot's cameras to the remote user. Additionally, web-users have a more flexible control over the robot. They can control the robot exclusively for a fixed amount of time which generally is set to 10 minutes per user. Whenever a user has control over the robot, he/she can direct it to arbitrary points in the exhibition. The user can select from a list of predefined guided tours or direct the robot to visit particular exhibits or locations in the exhibition. At each point in time, the user can request a high-resolution image grabbed with the cameras maximal resolution. Furthermore, the interface allows the control of the pan-tilt unit of the robot. Thus, the user can look in any user-defined direction. Finally, the user may request the robot to move around an exhibit in order to view it from all possible directions.

The control page of the interface is depicted in Figure 8. The left side contains the predefined tours offered to the user as well as the list of exhibits that are known to the robot. The center shows the live-stream as well as a Java applet animating the robot in a 2D floor-plan. This map can also be used to directly move the robot to an exhibit or to an arbitrary location in the exhibition. Between the map and the live-stream, the interface includes control buttons as well as a message window displaying system status messages. The right part of the interface shows multi-media information about the exhibit including links to relevant background information.

5.1.1 Enhanced Visualizations

Once instructed by a Web user, the robot fulfills its task completely autonomously. Since the system also operates during opening hours, the robot has to react to the visitors in the museum. This makes it impossible to predict the robot's course of action beforehand. Therefore, it is highly important, to visualize the environment of the robot and the moving people therein, so that the web user gets a better understanding of what is going on in the museum and why the robot is carrying out the current actions.

A typical way of providing information to the users is video streams, recorded with static or robot-



You are in control.
180 seconds remain.

St Demetrios
Attributed to
Frangos Katelanos
Mid-16th century
0.95 x 0.69 m
T 143

The patron saint of Thessaloniki is portrayed in frontal pose against a gold and light blue ground, with the nominative inscription preserved at the top: Ο ΑΓΙΟΣ ΔΗΜΗΤΡΙΟΣ. Attired in military uniform, he holds a spear in the right hand

and a sword in the left, while a helmet and shield hang from his shoulders. His halo and the icon's frame are decorated in relief with vegetal motifs.

The stylistic and iconographic traits of the icon recall works by Frangos Katelanos, the leading representative of the so-called School of Northern Greece, which flourished in parallel with the Cretan School of monumental painting in the sixteenth century.

Figure 8. Web interface of the TOURBOT system for exclusive control over the robot.

mounted cameras. This, however, has the disadvantage of limited perspectives and high bandwidth requirements. For these reasons, we developed a control interface, which provides the user with a virtual reality visualization of the environment including the robot and the people in its vicinity. Based on the state information received from the robot and our tracking algorithm, our control interface continuously updates the visualization. Depending on the level of detail of the virtual reality models used, the Internet user can obtain visualizations, whose quality is comparable to video streams. For example, Figure 9 shows two sequences of visualizations provided during the installation of the system in the Deutsches Museum Bonn in November 2001 along with images recorded with a video camera and with the robot's on-board camera. Within the graphics visualization, people are shown as avatars. As can be seen, the visualization is almost photo-realistic and the animated avatars capture the behavior of the people in the scene quite well.

Compared to the transmission of video streams, the graphics-based visualization highly reduces the bandwidth requirements of the control interface. TOURBOT's standard web interface used a single

Figure 9. The enhanced 3D visualization allows arbitrary view-points. The left sequence shows the real and the virtual view through the robot's cameras. The right images show the robot guiding three people through the museum and a bird's eye view of the scene.

video stream to transmit images of 240 by 180 pixels in size with a frame rate of about 5 Hz. This still required a bandwidth of about 40kBit/s. Compared to that, the graphics-based visualization only needs about 1kBit/s to achieve the same frame rate, if we assume that 7 people are constantly present in the robot's vicinity. It has the additional advantage, that the bandwidth requirement is independent of the image size. The graphics-based solution, therefore, allows for more detailed visualizations. Beyond the bandwidth savings, the graphics-based visualization offers an increased flexibility to the Internet user. Virtual cameras can be placed anywhere and the viewpoints can even be changed at run-time, as illustrated in the right image sequence of Figure 9. Our current prototype implements these ideas. It uses Open Inventor models of the robot and of the environment for the 3D rendering. On start-up, the control interface connects to the robot via TCP/IP and after downloading the model, the visualization component receives state information from the robot and starts rendering the scene accordingly.

5.2 On-board Interface

Besides the web interface that is used by remote users, the robots have several means for communicating and interacting with the on-site visitors. A reduced version of the web interface is also displayed in a touch screen appropriately mounted at the rear side of the robot. Through this touch screen, the on site visitors may instruct the robot to guide them to specific exhibits. One of the main differences between the web and the on-board interface is that video streams and enhanced visualizations are not provided, since they are not actually required by the on-site visitors. Instead, the on-board interface makes extensive use of synthetic speech. To enhance the communication with users in the museum, the robots are equipped with a speaker-independent speech interface. We employ a commercially available speech system [18] that detects simple phrases. The input of the user is processed and the parsed phrase is used to generate corresponding actions. To improve the recognition rate, the software allows the definition of contexts, i.e., sets of phrases that are relevant in certain situations. Depending on user input or depending on the task that is currently carried out, the system can dynamically switch between the different contexts. The current system includes 20 different phrases, that can be used to request information about the robot, the exhibition site, or even the time and the weather. In several installations in populated environments we figured out that the overall recognition rate is approximately 90%. Figure 10 shows a scene in which a person interacts with the robot Albert during the Hannover trade fair in 2001. Here the person asked several questions about the robot and requested information about the time (*who are you?, where are you from?, what are you doing here?*). Depending on the input of the user the robot can dynamically generate speech output. The text to be spoken is converted into audio files that are played back.

Another very important aspect of the on-board interface is the capability of the robots to alter the facial expressions of their mechanical heads based on their internal status. Currently, there are three different facial expressions of the robot, namely “happy”, “neutral” and “angry”. These facial expressions are implemented by modifying the shape of the eyebrows and the mouth. Combined with a variety of voice messages, the robot uses these expressions to inform the on-site visitors regarding its internal status. For

Figure 10. Person interacting with Albert during a Hannover trade fair demonstration.

example, if the path of the robot is not obstructed by the on-site visitors, the robot appears happy. In the opposite case, the robot's "mood" changes progressively in time. Moreover, the head of the robot is controlled so that it looks towards the direction of intended motion. This way, on-site visitors adapt their motion so as not to obstruct its path.

6 System Installation and Demonstration

In the framework of the TOURBOT project a number of demonstration trials were undertaken in the premises of the participating museums. More specifically, the TOURBOT system has first been developed and fully tested in the laboratory environment. Following that, and in order to acquire performance data from actual museum visitors, the system has been installed and demonstrated in the three museums that participated in the project consortium. These demonstrations were combined with relevant events in order to publicize and disseminate the results of the project to professionals and the broader public. Factual information of these events is as follows:

- Foundation of the Hellenic World, Athens, Greece, May 28–June 2, 2001. Exhibition: "Crossia, Chitones, Doulamades, Velades - 4000 Years of Hellenic Costume." The exhibition area comprised 2000 square meters. During the trial the robot operated approximately 60 hours covering a distance of 14 kilometers. More than 1200 web users observed the exhibition through TOURBOT.

A typical situation, in which the robot Lefkos guides visitors through the museum is shown in Figure 11(a).

(a) (b)

(c) (d)

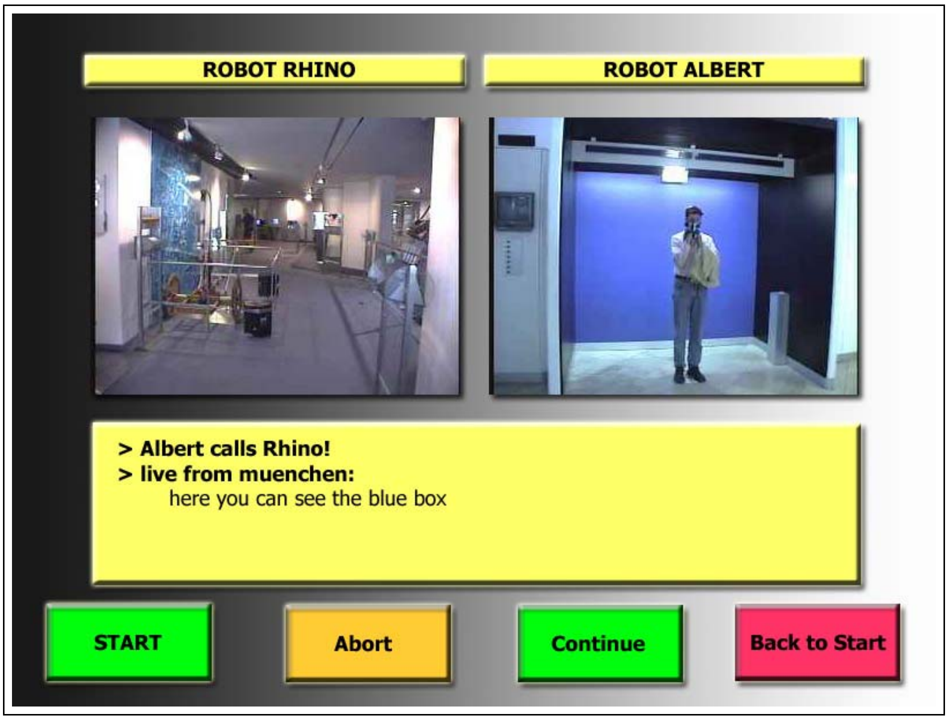
Figure 11. (a) Robot Lefkos operating in the exhibition of the Foundation of the Helenic World. (b) Robot Rhino operating in the Deutsches Museum Bonn. (c) Robot Lefkos operating in the Byzantine and Christian Museum. (d) Robot Albert interacting with a person at the Heinz Nixdorf MuseumsForum. This picture is curtesy of Jan Braun, Heinz Nixdorf MuseumsForum.

- Deutsches Museum Bonn, Bonn, Germany, November 6–11, 2001 (see Figure 11(b)). Exhibition: “Part of the permanent exhibition, highlighting scientific achievements that were awarded the Nobel Prize.” The exhibition area in which the robot moved comprised about 200 square meters. The system operated about 60 hours, covering a distance of 10 km. Approximately 1900 web visitors had a look around the museum via the robot.

- Byzantine and Christian Museum, Athens, Greece, December 3–7, 2001 (see Figure 11(c)). Exhibition: “Byzantium through the eyes of a robot.” The exhibition area comprised about 330 square meters. During the trial the robot operated 40 hours, covering a distance of 5.3 kilometers. The number of web users was small in this trial, due to the following fact. Since the first day of the trial at the Byzantine and Christian Museum, a large number of (on-site) visitors were coming to the exhibition. This forced the TOURBOT team to the decision to devote significantly more time of the system to on-site visitors as opposed to web visitors.

Additionally, TOURBOT was installed and operated for a longer period of time (Oct. 2001–Feb. 2002) at the Heinz Nixdorf MuseumsForum (HNF) in Paderborn, Germany (see Figure 11(d)). This was in the framework of the special exhibition “Computer.Gehirn” (Computer.Brain) with a focus on the comparison of the capabilities of computers/robots and human beings. Recently (June 2002), TOURBOT was introduced for one week in the Museum of Natural History of the University of Crete, Heraklion, Greece.

Within the WebFAIR project, a recent event, namely RoboTalk, has been realized. The goal of RoboTalk was to demonstrate that two robots installed in two different exhibitions can be coordinated so that the users can, at the same time, be guided through the two different exhibitions. In this particular case the mobile robot Rhino was deployed in the Deutsches Museum Bonn and the robot Albert was installed in the Deutsches Museum in Munich. The event lasted four days, from June 19 to June 22, 2003. The task of the robots was to guide the visitors through the different exhibitions. At the same time the robots were connected via the Internet so that visitors in one exhibition could participate in the tour given by the other robot. Whenever one robot arrived at an exhibit, it communicated this to the other robot. The latter, regardless of the motion command under execution, would stop and provide the video and audio data received from the other robot. This way, visitors of one museum could also see exhibits presented by the other robot at the remote site. Figure 12 shows the user interface of the robots during this event. In addition to information about the exhibition the robot is installed in, it includes a window showing the video stream obtained from the other robot. Figure 13 shows two particular trajectories of Rhino and Albert during the event. Whereas the upper image shows the trajectory taken by Rhino in Bonn, the lower image depicts the path of Albert in Munich. The points in time when the two robots were synchronized are indicated by the numbers. In this particular situation Albert stopped three times



> Albert calls Rhino!
> live from muenchen:
here you can see the blue box

Figure 12. Screen-shot from the GUI.

in Munich to present exhibits shown by Rhino.

6.1 Installation Time

The large number of test installations of the TOURBOT system required sophisticated tools for the setup of the overall system. The most crucial part of the overall procedure is the generation of the navigation map. However, based on the techniques described earlier in this paper, the overall mapping process could in all cases be accomplished within several hours. To avoid that the robot leaves its desired operational space or collides with obstacles that cannot be sensed, we manually create a second map, depicting such obstacles. This map is then fed to the collision avoidance module, thus preventing the robot from moving into the corresponding areas.

A further time consuming process is the generation of the multimedia-content that is presented to the user for each exhibit. The TOURBOT system includes a generic Multimedia database including HTML-pages, images, audio, and video sequences. Material in the database can be changed and/or edited using available software tools. Furthermore, the robot is equipped with a task specification that defines where

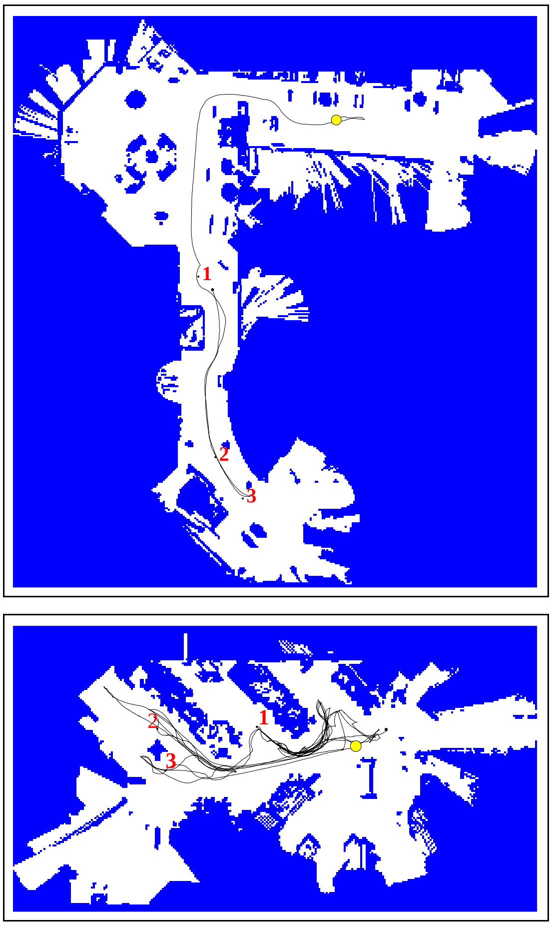


Figure 13. Synchronization of the tours given by Rhino and Albert in the Deutsches Museum Bonn and the Deutsches Museum Munich.

the designated exhibits are and which content has to be presented.

Most of the multimedia information pertinent to the exhibits can be obtained directly from the exhibition sites, since pictures, text and other relevant material are often already contained in existing Web presentations.

The whole setup can therefore be accomplished in less than two days. This is an enormous speed-up compared to previous tour-guide systems. Figure 14 shows the time required to install the Rhino and Minerva systems [5, 26] in comparison to that of the TOURBOT system. As can be seen, the TOURBOT system requires significantly less time than Rhino and Minerva. Our experience with tour-guide robots in exhibition sites suggests that 3D models of exhibitions' premises are generally not available. The

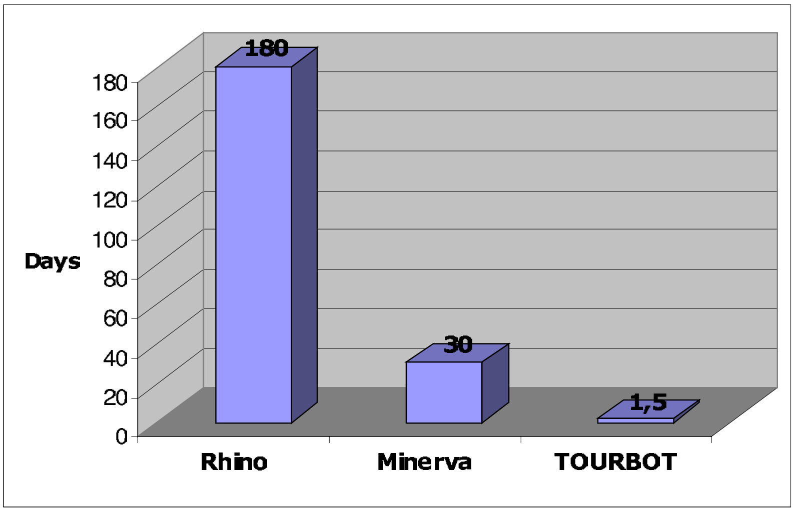


Figure 14. Time required to install the different tour-guide systems Rhino, Minerva, and TOURBOT.

automatic generation of such models with the mobile robot itself is a subject of ongoing research [11].

6.2 Reactions of the Public

An interesting study was carried out in the context of the above trials regarding the visitor-robot interaction patterns. The observation of the interaction between visitors and the robot in the first two trails (at the Foundation of the Hellenic World, Athens, Greece, and the Deutsches Museum Bonn, Bonn, Germany) led us to consider carrying out a more detailed study on the human-robot-interaction. This study was planned and carried out during the third trial in the Byzantine and Christian Museum, Athens, Greece. Based on the observations during the previous trials, the human-robot-interaction was qualitatively classified into four behaviors, namely watching the robot, moving towards the robot, touching the robot and using the touch screen. The frequency of the occurrence of these behaviors for the cases of adults and children is summarized in Figure 15. Considering these results one may come to some preliminary conclusions regarding the human-robot-interaction:

- Most visitors turn/move to the robot when they see it, be it immediately or after watching it for a while. This indicates that the robotic platform(s) attract the interest and are appealing to the visitors.

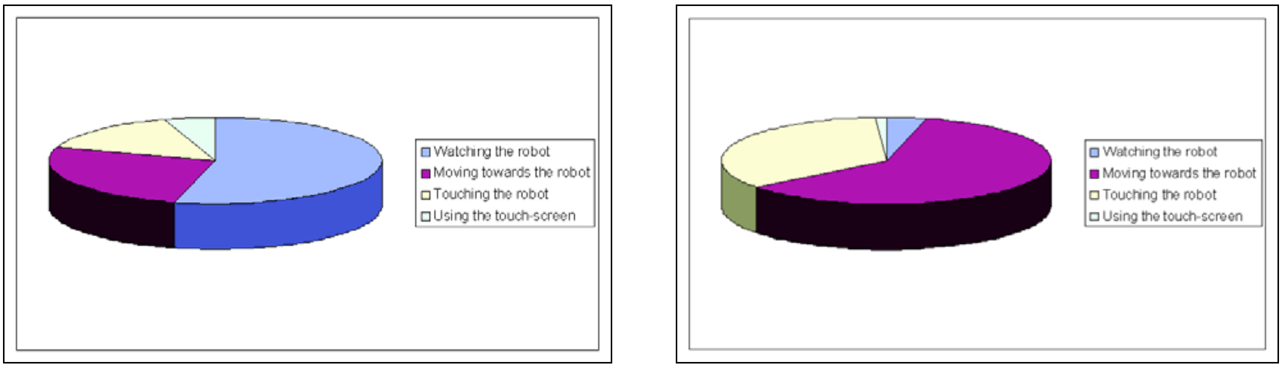


Figure 15. Reactions of Visitors to the Robotic Tour-guide. (a) Adult visitors, (b) Children.

- The smaller counts that were observed in the fourth behavior (using the touch-screen) are justified in many ways. A major reason for that is that only one person has the possibility to use the touch-screen at any point in time.
- Comparing adults and children it can be stated that children are uninhibited and natural in their behavior. Significantly more counts relate to the behavior “Moving towards the robot” than to “Watching it”. On the contrary, adults tend to exhibit a more “reserved” behavior towards the robot.
- The performance in the correct use of the robot is better in adults than in children.

7 Conclusions

In this paper we presented a number of techniques that are needed for realizing web-operated mobile robots. These techniques include effective map building capabilities, a method for obstacle avoidance that is based on a combination of range and visual information and advanced web- and on-board robot interfaces. In addition to video streams our systems provide high-resolution virtual reality visualizations that also include the people in the vicinity of the robot. This increases the flexibility of the interface and simultaneously allows a user to understand the navigation actions of the robot.

The techniques described in this paper have been successfully deployed within the EU-funded projects TOURBOT and WebFAIR which aim at the development of interactive tour-guide robots, able to serve web- as well as on-site visitors. Technical developments in the framework of these projects have resulted in robust and reliable systems that have been demonstrated and validated in real-world conditions.

Equally important, the system set-up time has been drastically reduced, facilitating its porting to new environments.

Our current research extends the navigation capabilities of the robotic systems by addressing obstacle avoidance in the cases of objects that are not visible by the laser scanner, 3D mapping, mapping in dynamic environments, predictive navigation, and multi-robot coordination [6]. Moreover, in the context of the above projects additional issues are addressed. They consider (a) how to adapt this technology in order to fit the long-term operational needs of an exhibition site, (b) how to evaluate the robotic system in terms of its impact to the main function and objectives of the exhibition site (financial impact, accessibility, marketing and promotion, impact on visitor demographic, etc.), and (c) how to evaluate the content and educational added value to museum and exhibition visitors, and generate a feedback to the technology developers in order to improve in the future the robotic avatars and adapt further to the needs of the users.

8 Acknowledgments

This work has partly been supported by the by the IST Programme of Commission of the European Communities under contract numbers IST-1999-12643 and IST-2000-29456. The authors would also like to thank the members of the IST-projects TOURBOT and WebFAIR for helpful comments and fruitful discussions.

References

[1] H. Asoh, S. Hayamizu, I. Hara, Y. Motomura, S. Akaho, and T. Matsui. Socially embedded learning of office-conversant robot jijo-2. In *Proceedings of IJCAI-97*. IJCAI, Inc., 1997.

[2] H. Baltzakis, A. Argyros, and P. Trahanias. Fusion of laser and visual data for robot motion planning and collision avoidance. *Machine Vision and Applications*, 15:92 – 100, 2003.

[3] H. Baltzakis and P. Trahanias. Closing multiple loops while mapping features in cyclic environments. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2003.

[4] H. Baltzakis and P. Trahanias. A hybrid framework for mobile robot localization: Formulation using switching state-space models. *Autonomous Robots*, 15(2):169 – 191, 2003.

[5] W. Burgard, A. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114(1-2), 1999.

[6] W. Burgard, P. Trahanias, D. Haehnel, M. Moors, D. Schulz, H. Baltzakis, and A. Argyros. Tele-presence in populated exhibitions through web-operated mobile robots. *special issue on On-line Robots, Autonomous Robots*, 15(3):299 – 316, 2003.

[7] H. Endres, W. Feiten, and G. Lawitzky. Field test of a navigation system: Autonomous cleaning in supermarkets. In *Proc. of the IEEE International Conference on Robotics & Automation (ICRA)*, 1998.

[8] K. Goldberg, S. Gentner, C. Sutter, J. Wiegley, and B. Farzin. The mercury project: A feasibility study for online robots. In K. Goldberg and R. Siegwart, editors, *Beyond Webcams: An Introduction to Online Robots*. MIT Press, 2002.

[9] K. Goldberg, J. Santarromana, G. Bekey, S. Gentner, R. Morris, J. Wiegley, and E. Berger. The telegarden. In *Proc. of ACM SIGGRAPH*, 1995.

[10] D. Hähnel, W. Burgard, D. Fox, and S. Thrun. An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2003.

[11] D. Hähnel, W. Burgard, and S. Thrun. Learning compact 3d models of indoor and outdoor environments with a mobile robot. In *Fourth European workshop on advanced mobile robots (EUROBOT'01)*, 2001.

[12] D. Hähnel, D. Schulz, and W. Burgard. Map building with mobile robots in populated environments. In *Proceedings of the Conference on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland, 2002.

[13] H. Hirukawa, I. Hara, and T. Hori. Online robots. In K. Goldberg and R. Siegwart, editors, *Beyond Webcams: An Introduction to Online Robots*. MIT Press, 2002.

[14] S. King and C. Weiman. Helpmate autonomous mobile robot navigation system. In *Proc. of the SPIE Conference on Mobile Robots*, pages 190–198, 1990.

[15] G. Lacey and K. Dawson-Howe. The application of robotics to a mobility aid for the elderly blind. *Journal of Robotics and Autonomous Systems (RAS)*, 23:245–252, 1998.

[16] K. Murphy. Bayesian map learning in dynamic environments. In *Neural Info. Proc. Systems (NIPS)*, 1999.

[17] I. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, and A. Soto. An affective mobile educator with a full-time job. *Artificial Intelligence*, 114(1-2), 1999.

[18] http://www.novotech-gmbh.de/.

[19] H. Rauch, F. Tung, and C. Striebel. Maximum likelihood estimates of linear dynamic systems. *American Institute of Aeronautics and Astronautics Journal*, 3(8):1445–1450, 1965.

[20] D. Rodriguez-Losada, F. Matia, R. Galan, and A. Jimenez. Blacky, an interactive mobile robot at a trade fair. In *Proc. of the IEEE International Conference on Robotics & Automation (ICRA)*, 2002.

[21] C. Schaeffer and T. May. Care-o-bot - a system for assisting elderly or disabled persons in home environments. In *Assistive technology on the threshold of the new millenium*. IOS Press, Amsterdam, 1999.

[22] R. Schraft and G. Schmierer. *Serviceroboter*. Springer Verlag, 1998. In German.

[23] D. Schulz, W. Burgard, D. Fox, S. Thrun, and A. Cremers. Web interfaces for mobile robots in public places. *IEEE-Magazine on Robotics and Automation*, 2000.

[24] R. Simmons, R. Goodwin, K. Haigh, S. Koenig, and J. O'Sullivan. A layered architecture for office delivery robots. In *Proc. of the First International Conference on Autonomous Agents (Agents)*, 1997.

[25] K. Taylor and J. Trevelyan. A telerobot on the World Wide Web. In *Proceedings of the 1995 National Conference of the Australian Robot Association*, 1995.

[26] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot Minerva. *Journal of Robotics Research*, 19(11), 2000.

[27] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan. The unscented particle filter. Technical Report CUED/F-INFENG/TR 380, Cambridge University, Department of Engineering, 2000.

[28] http://www.webcam32.com/.