# Robox, a Remarkable Mobile Robot for the Real World

Kai O. Arras, Nicola Tomatis, Roland Siegwart

Autonomous Systems Lab
Swiss Federal Institute of Technology Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

**Abstract.** $\nu$ In this paper we present Robox, a mobile robot designed for autonomous operation in a mass exhibition environment. Robox has unique multimodal interaction capabilities and a novel approach to localization using multiple Gaussian hypotheses. What makes Robox one of a kind is on the one hand its design and the variety of functionalities united in one platform and on the other hand the scale of the Expo.02 project where Robox has been deployed.

Here, we adopt an experimental view of the task. After the problem specification of mass exhibitions, we outline system integration aspects: mechanical design, safety, software and hardware architecture and interaction modalities. Finally, seen as an enabling technology for robots in exhibitions, we present the localization approach in more detail.

Building on former experience with feature-based Kalman filter localization we address the data association problem whose neglect was found to be the predominant reason for localization failures. Multiple hypotheses are generated by a constraint-based search in the tree of local-to-global associations, given a local map of observed features and a global map of the environment. As soon as hypotheses are available they get tracked with an algorithm relying on the same interpretation tree technique. By track splitting under geometric constraints, location ambiguity can be represented not only globally but also locally, thus forming a consistent framework for global Kalman filter localization. The experiments demonstrate significantly improved robustness at modest computational costs.

The raison d'être of Robox is the Robotics pavilion at the Swiss National Exhibition Expo.02. There, ten Roboxes guided more than half a million visitors through the exhibition, eleven hours per day, seven days per week, from May 15 to October 20, 2002.

## 1 Introduction

Progress in the application of estimation and decision theory combined with advances in sensor and embedded system technology enable robots today to navigate reliably in highly dynamic real world environments. This allows to enter new application domains such as public spaces. Over the past years, the number of robots that have been deployed in museums, trade shows and exhibitions has been grown steadily [1]. Far from the research laboratory, these installations have demonstrated high degrees of reliability, some of them over several years of operation. Robots in public spaces face a particularly

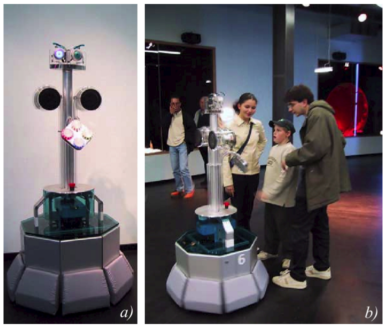

**Fig. 1.** Robox (a), Robox Nr.6 with visitors in the pavilion at Expo.02 (b)

difficult environment – for navigation and for human-robot interaction. A limited number of researchers have addressed this challenge [6,17,16,11,14]. *Rhino* [6] and *Minerva* [17] were both deployed in a technical museum during one or two weeks respectively. The robots had the task to give guided tours to visitors. They were not truly autonomous as localization was running off-board. The robot *Kapros* [14] served as an avatar for remote visitors of an art museum during two weeks. Unlike these short-term projects, [16] describes the permanent installation of the autonomous tour-guide robot *Sage* in a museum of natural history which gave rise to three successor installations. Altogether the robots exceeded a total operational time of seven years. For localization, the environment was modified by adding artificial landmarks (color patches). A multi-robot installation which is operational since March 2000 is presented in [11]. Three self-contained mobile robots, also as a permanent installation in a museum, have the task to welcome visitors, offer them exhibition-related information and to entertain them. Their navigation area is restricted and well structured. Localization uses segment features and a Kalman filter for pose estimation. Finally, the 72-robot installation at the World Fair Expo 2000 in Hannover, Germany, was the first application of mobile robots in a mass exhibition. The vehicles however were very low-tech. Localized and controlled from external infrastructure, their served as freely moving swarm entities forming a hugh interactive art installation during the six months of Expo 2000 (there is no publication to the knowledge of the authors).

## 2 Problem Statement

The Swiss National Exhibition takes place about every 40 years with *Expo.02* as the most recent issue (May 15–October 20, 2002). It is a major national happening with 37 exhibitions and a rich event program. The *Robotics* pavilion, one of these 37 exhibitions, is intended to show the increasing closeness between man and robot technology. The overall requirements can be summarized as follows

- Interaction with an autonomous, freely navigating and socially interactive robot as the central visitor experience of the exhibition
- Five hundred persons per hour visitor flow capacity (visitor flow is a major concern in mass exhibitions)
- Autonomous operation during eleven hours per day, seven days per week
- Robot tasks include tour-giving, taking pictures of visitors and entertaining

Technically this implies

- Navigation in unmodified, highly populated environments with visitors and other freely navigating robots
- Bidirectional multi-modal interaction via easy-to-use, intuitive yet robot-typical interaction modalities. Speech output in four languages: French, German, Italian and English
- Adaptive multi-robot coordination scenarios in function of the number of visitors and their interests
- Long-term robustness, minimal manual intervention, minimal supervision
- Safety for visitors and robots at all times
- Development of ten robots within tight budgets and schedules

## 3 The Mobile Robot Robox

Striving for maximal flexibility and based on a long-standing experience in robot design, system integration and navigation [3,5], we concluded that building a robot from scratch is the best to do. The outcome is shown in figure 1 and 2 and described in more detail in [18].

### 3.1 Hardware

The lower part of the robot (octagonal base) contains a CompactPCI rack, two SICK LMS 200 laser range finders, a redundant security circuit, the batteries, tactile plates with foam bumpers, a gray scale camera interfaced by a Bt848 framegrabber (mounted upwards looking behind the robot face), and the analogue and digital I/O periphery hardware. The two main CPUs are a PowerPC 750 (G3) at 380 MHz and a Pentium III at 700 MHz. The robot has a symmetric differential drive kinematics, actuated by two EC motors via 1:50 harmonic drives and a castor wheel on each side, one on a spring suspension. This gives Robox an good maneuverability and stability in spite of its height of 1.65 m. The batteries provide autonomy for about twelve hours.

The upper part of the robot incorporates and interfaces the interaction modalities. The face has five degrees of freedom: two eyes with independently actuated pan-tilt units and two mechanically coupled eyebrows. The right eye is equipped with a Firewire color camera, the left eye integrates an LED

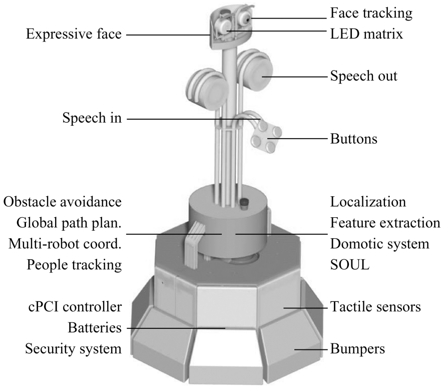

**Fig. 2.** Functionalities of Robox

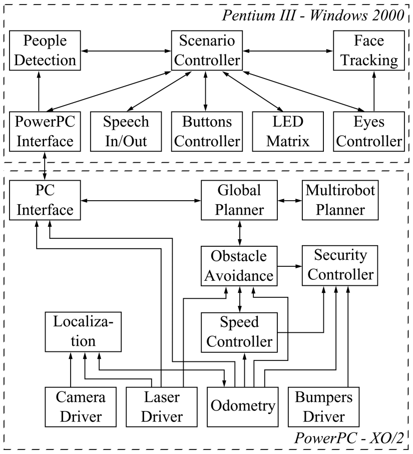

**Fig. 3.** Software architecture

matrix for display of static and animated icons. Below the face there are two loudspeakers for speech output and sound playback. The central input device for a bidirectional communication are four individually colored, capacitive buttons that allow language selection, response to questions and other types of interaction. Two of the ten Roboxes are also equipped with a directional microphone matrix for experiments on speech recognition.

Each robot is in connection to its own off-board PC via a 2.4 GHz radio Ethernet link. The PC serves as a supervision and logging tool, displaying and storing important state information in real time. All ten robots share the bandwidth of this network. The radio link is also used by the domotic system of Robox. It allows to communicate with networked devices in the environment. In the Expo.02 pavilion, the robots can switch on and off light sources, take pictures, trigger a flash or remote control other robots via infrared.

Safety for people and robots is guaranteed by true hardware redundancy. A separate circuit (Microchip PIC) receives a watchdog signal from a software security controller (which in turn receives a signal from the circuit), supervises the tactile plates and monitors the maximal vehicle speed. In case of a failure or the absence of a signal, the robot is stopped and set into an emergency mode. Hardware redundancy was the prerequisite for Robox to be officially classified as a safe machine by the respective Swiss professional association.

### 3.2 Software

Hardware relation, safety and availability were the criteria for the mapping of software units onto the hardware. The navigation software is considered safety- and timing-critical. It runs on the PowerPC under the hard real-time operating system XO/2 [5] due to the functional and temporal guarantees this operating system provides. The interaction software, not considered safety-critical, is mapped onto the Pentium under Windows 2000. So, we take advantage of a standard software environment and a wide availability of drivers, libraries, and COTS (commercial off-the-shelf) components. Finally, for hardware related objects such as drivers, the choice is given by the board to which the periphery is connected (e.g. the LMS 200 to the PowerPC, the eye camera to the PC). The resulting diagram is shown in figure 3.

The highest level program of Robox is a scenario controller which has access to all objects. A scenario implements decision-tree based sequences containing the controls for all modalities (speech, face expression, motion, LED matrix, etc.). An interactive tour for instance, consists of several small scenarios, emergency scenarios and settings for a behavioral state machine, played and managed by the scenario controller. The software for scenario creation (a stand-alone application with a user-friendly graphical interface) and the especially developed Scenario Object Utility Language (SOUL) allow to develop and adapt new scenarios quickly and to implement complex tours and interaction sequences also by an untrained user (further details in [13]).

## 4 Localizing Robox: a Global EKF Approach

Among the many components of Robox, localization belongs to the key elements for a project like Expo.02. Exhibitions are by their nature highly dynamic, cluttered, and possibly hostile. We must be prepared for occluded sensors, collisions, wheel slippage and kidnapping.

Clearly, a robot doing (single-hypothesis) extended Kalman filter (EKF) pose tracking can loose its track and go lost – especially in the above situations. The typical reason for this is the inability to represent and reduce data association ambiguity. Robust localization cannot ignore the data association problem. In this approach, building on former experience with feature-based EKF localization, we address the data association problem on a discrete non-Bayesian feature-to-feature level.

Unlike POMDP or Markov approaches [15,6,17] where locations are generated before they get evaluated by the exteroceptive sensors (as a grid or a set of particles), our approach to localization turns this process around: locations are generated as a direct consequence from sensory information. Features tell us *when* and *where* to place a location hypothesis. This allows to maintain always as many hypotheses as necessary and as few as possible. The technique for hypothesis generation is a constraint-based search in an interpretation tree [12,9,7,2]. This tree is spanned by all possible local-to-global associations, given a local map of observed features $L = \{l_i\}_{i=1}^{p}$ and a global map of model features $G = \{g_j\}_{j=1}^{m}$. Besides track formation, we show that with the same technique, robust tracking with track splitting under geometric constraints can be realized – thus forming a consistent framework for global and local EKF localization.

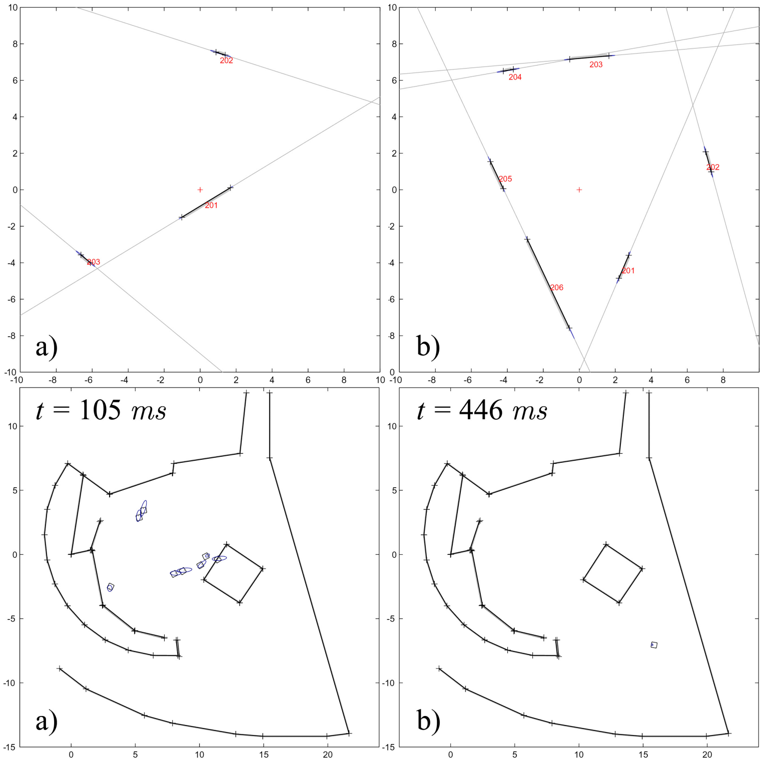

**Fig. 4.** Hypothesis generation. Given the local maps in a) and b), hypotheses are generated at locations where the local map fits into the global map. In a) there are 15 hypotheses (with their 95% error ellipse), in b) there is a single hypothesis; the robot is instantaneously localized. $t$ denotes the execution time.

### 4.1 Hypotheses Generation

We briefly outline the approach (refer to [2] for a more complete presentation): The search space for hypothesis generation is the space of all possible associations of the observed features $l_i$ and the model features $g_j$. The space has the structure of a tree with $p$ levels and $m+1$ branches [12]. $p$ is the number of observed features in $L$, $m$ the number of model feature in $G$. The extra branch (called star branch) allows correct associations in the presence of outlier observations (false positives) and thus accounts for environment dynamics and map errors. During tree traversal, statistically feasible *pairings* $p_{ij} = \{l_i, g_j\}$ are sought given all uncertainties associated to the features. A pairing says that the observed feature $l_i$ and the model feature $g_j$ denote the same physical object in the environment ($g_j$ is called an interpretation of $l_i$). Geometric constraints from the features are applied into the formation of pairings. They determine their statistical compatibility on a significance level $\alpha$. Although the problem is of exponential complexity, the geometric constraints reduce enormously the space to be explored.

They can be classified into two categories:

**Location Independent Constraints** *Unary constraint.* We accept the pairing $p_{ij}$ if $l_i$ and $g_j$ are of the same type, color, size or any other instrinsic property. Examples: the length of the observed segment $l_i$ is equal (or smaller) than the length of the model segment $g_j$.

*Binary constraint.* Given a valid pairing $p_{ij}$ we will accept the pairing $p_{kl}$ only if the two local features $l_i$ and $l_k$ are compatible to the two global features $g_j$ and $g_l$. Examples: $l_i$ and $l_k$ are lines with the intermediate angle $\phi_{ik}$. Then, the pairing $p_{kl}$ is considered compatible if $\phi_{ik} = \phi_{jl}$. With point features, for instance, the distances $l_i$-$l_k$ and $g_j$-$g_l$ must correspond.

**Location Dependent Constraints** The above tests do not involve the robot position $L_h$. Once this is known, a further class of constraints can be applied.

*Visibility constraint.* This constraint only applies to features from the map. It tests whether $g_j$ is visible from the robot position $L_h$. Example: lines or segments can be seen only from one side. If the robot is behind a wall, one of the two lines modeling the wall is invisible. With sensor specific parameters, the visibility constraint rejects features which are not detectable, for instance, because they are farer away than a maximal perception radius.

*Rigidity constraint.* A pairing $p_{ij}$ is considered compatible if $l_i$ and $g_j$, transformed into the same coordinate system given $L_h$, coincide (are at the same position). This is what happens in the matching step of any EKF localization cycle. Usually, $g_j$ is transformed into the frame of $l_i$.

*Extension constraint.* A pairing $p_{ij}$ is considered compatible if $l_i$ and $g_j$, transformed into the same coordinate system given $L_h$, fully overlap. Example: an observed segment $l_i$ must be completely contained in the transformed $g_j$ seen from the location $L_h$.

**The Search Algorithm** The constraints allow to discard whole subspaces (subtrees) from the search each time when an incompatible pairing is found at the root of such a subtree. With the uncertainties associated to local and global features, all decisions make use of the Mahalanobis distance on a significance level $\alpha$.

Tree traversal is implemented as a recursive back-tracking search algorithm generate_hypotheses described in [7,2]. The strategy is to first find a minimal number of valid pairings with location independent constraints such that a location estimate can be determined in order to apply location dependent constraints, too. Each time when the algorithm reaches the bottom of the tree, that is, all observed features could have been assigned to a model

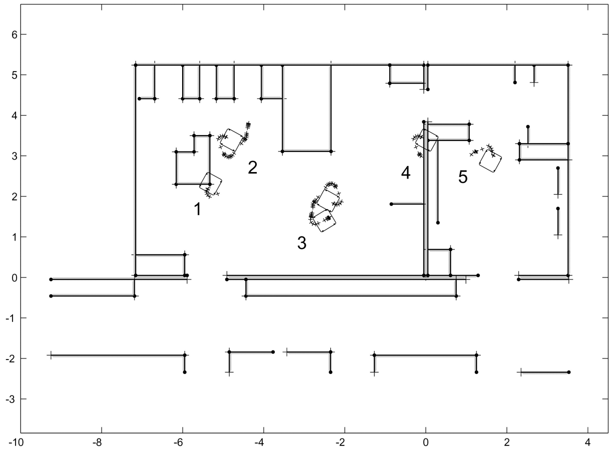

**Fig. 5.** Multi-hypothesis tracking. Starting with five hypotheses, track #3 turns out to be the true one after the last track (#2) was rejected at 1.89 m distance

feature or to the star branch, we have a valid robot location hypothesis. The pairings which support the hypothesis are put together in a *supporting set* $S_h = \{\{l_1, g_{j_1}\}, \{l_2, g_{j_2}\}, ..., \{l_p, g_{j_p}\}\}$ and thereby constitute a location hypothesis $h = \{S_h, L_h\}$. All hypotheses together form the set of robot location hypotheses $H = \{h_i\}_{i=1}^{n}$.

**Estimating the Robot Location from $S_h$** With the supporting set, the $(x, y, \theta)$-pose of the robot is not yet known. This is what the extended information filter (EIF) does. Given a supporting set with all associated uncertainties, it estimates the robot location and its covariance in the least square sense. The difference between the EIF and the EKF is that the former is the batch estimator formulation of the latter (which is recursive). This is needed, because, for hypothesis generation, there is no a priori knowledge on the robot location which means formally that the state prediction covariance, usually called $P(k|k+1)$, is infinite. With the EIF, this can be properly expressed as $P^{-1}(k+1|k) = \mathbf{0}_{3\times3}$ since covariance matrices are represented in the information matrix form, that is, by their inverse.

Figure 4 shows two examples of hypothesis generation in the Expo.02 environment. With multiple discrete hypotheses, to be localized is simply expressed as having a single hypothesis.

### 4.2 Multi-Hypothesis Tracking

The main reason for lost situations during tracking is incorrect data association. This occurs typically when there are several statistically feasible pairing candidates for an observation. Choosing the closest one – the most widely applied strategy called nearest neighbor standard filter –, leads to filter inconsistency and mostly to filter divergence if it was the wrong one.

Here we will follow another strategy. As soon as there is association ambiguity, that is, there is no guarantee anymore for the correct association to be found, we re-generate hypotheses locally. This property has the algorithm track_hypothesis in [2] which, given a predicted location, a local and a global map, splits up into multiple offspring hypotheses if statistical compatibility with several supporting sets can be established at that location. It has the identical structure than the algorithm for hypothesis generation but employs location dependent constraints *only* and does *not* recur with a refined position estimation. In this manner the algorithm finds all supporting sets in the vicinity of the initially predicted location.

After a hypothesis has been tracked, track_hypothesis yields three cases: *(i)* hypothesis confirmation (this is the 'normal' case), the location is refined by the EKF, *(ii)* hypothesis rejection and *(iii)* hypothesis splitting. Track rejection takes place when the predicted location is not supported anymore by location dependent constraints on the level $\alpha$. When track splitting occurs, the location of each track get newly estimated and the best one is taken. Best in a goodness-of-fit sense, expressed by the joint Mahalanobis distance.

Figure 5 shows an experiment how the robot converges towards the true location after a short trajectory. Note that by geometry only (or geometric falsification respectively), false tracks get rejected quickly. No free-space information is needed. Localization cycle time on Robox is about 10 Hz.

## 5 Conclusions

In this paper we present the mobile robot Robox, a platform designed for operation in mass exhibitions and deployed at the Swiss National Exhibition Expo.02. Its interaction capabilities, appearance and overall reliability make Robox a real success – scientifically, technically, from an exhibition maker's point of view and, of course, from a visitor's point of view. At the time of this writing, after 112 days of operation (of 159 total), the Robox family was up and running for 9,000 hours, interacted with more than 500,000 visitors and drove a overall distance exceeding 2,500 km. Already with these (preliminary) numbers, Robox surpasses the scope of previous exhibition projects and belongs to the most prominent specimens of its kind.

Considered as an enabling technology for robots in exhibitions, the localization technique was presented in more detail. From experiments and operation experience at Expo.02, we conclude that the presented global EKF approach is practical and exhibits the degree of robustness which is required in highly dynamic environments. With the results for the average computational effort for both hypothesis generation and tracking, the results further suggest that the typical efficiency of the feature-based paradigm could have been retained.

## References

1. Arras K.O., Burgard W. (editors), "Robots in Exhibitions", Workshop Proceedings, 2002 IEEE/RSJ Int. Conf. on Robots and Intelligent Systems, Lausanne, Switzerland.
2. Arras K.O., Castellanos J.A., Siegwart R., "Feature-Based Multi-Hypothesis Localization and Tracking for Mobile Robots Using Geometric Constraints", 2002 IEEE Int. Conf. on Robotics and Automation, Washington DC, USA.
3. Arras K.O., Tomatis N., Jensen B., Siegwart R., "Multisensor On-the-Fly Localization: Precision and Reliability for Applications", Robotics and Autonomous Systems, 34(2-3), 2001.
4. Bar-Shalom Y., Li X.-R., Estimation and Tracking: Principles, Techniques and Software, Artech House, 1993.
5. Brega R., Tomatis N., Arras K.O., Siegwart R., "The Need for Autonomy and Real-Time in Mobile Robotics: A Case Study of XO/2 and Pygmalion", 2000 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Takamatsu, Japan.
6. Burgard W., Cremers A.B., Fox D., Hhnel D., Lakemeyer G., Schulz D., Steiner W., Thurn S., "Experiences with a Interactive Museum Tour-Guide Robot." Artificial Intelligence 00(1-53): 2-53, 1999.
7. Castellanos J.A., Tardos J.D., Mobile Robot Localization and Map Building: A Multisensor Fusion Approach, Kluwer, 1999.
8. Cox I.J., Leonard J.J., "Modeling a Dynamic Environment Using a Bayesian Multiple Hypothesis Approach," Artificial Intelligence, 66(2), p. 311-44, 1994.
9. Drumheller M., "Mobile Robot Localization Using Sonar", IEEE Trans. on PAMI, 9(2), p. 325-32, 1987.
10. Fox D., Burgard W., Thrun S., "Markov Localization for Mobile Robots in Dynamic Environments", Artificial Intelligence Research, vol.11, p.391-427, 1999.
11. Graf B., Schraft R.D., Neugebauer J., "A Mobile Robot Platform for Assistance and Entertaiment", Int. Symposium on Robotics, Montreal, Canada, 2000.
12. Grimson W.E.L., Lozano-Prez, "Localizing Overlapping Parts by Searching the Interpretation Tree", IEEE Trans. on PAMI, 9(4), p. 469-82, 1987.
13. B. Jensen, Froidevaux G,, Greppin X., Lorotte A., Mayor L., Meisser M, Ramel G., "The interactive autonomous mobile system Robox", 2002 IEEE/RSJ International Conference on Intelligent Robot and Systems, Lausanne, Switzerland.
14. Maeyama S., Yuta S., Harada A., Experiments on a Remote Appreciation Robot in an Art Museum, 2000 IEEE/RSJ International Conference on Intelligent Robot and Systems, pp.1008-1013, Takamatsu, Japan.
15. Nourbakhsh I., Powers R., Birchfield S., "DERVISH, an office-navigating robot", AI Magazine, 16(2), p. 53-60, 1995.
16. Nourbakhsh I., Bodenage J., Grange S., Lutz R., Meyer R., "An Affective Mobile Robot Educator with a Fulltime Job." Artificial Intelligence 114(1-2): 95-124, 1999.
17. Thrun S., Beetz M., Bennewitz M., Burgard W., Cremers A.B., Dellaert F., Fox D., Hhnel D., Rosenberg C., Roy N., Schulte J., Schulz D., "Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva". International Journal of Robotics Research 19(11): 972-99, 2000.
18. Tomatis N., Terrien G., Piguet R., Burnier D., Bouabdallah S., Siegwart R., "Design and System Integration for the Expo.02 Robot," Workshop on Robots in Exhibitions, 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Lausanne, Switzerland.