

Multi-Robot Human-Interaction and Visitor Flow Management

B.Jensen, G.Froidevaux, X.Greppin, A.Lorotte, L.Mayor, M.Meisser, G.Ramel, R.Siegwart

Autonomous Systems Lab - Swiss Federal Institute of Technology - CH-1015 Lausanne

{bjoern.jensen,gilles.froidevaux,xavier.greppin,antoine.lorotte,laetitia.mayor,guy.ramel,roland.siegwart}@epfl.ch

Abstract -- In this paper we address the task of human-robot interaction in public mass exposition with several autonomous robots at a time. This implies questions regarding multi-robot control and interaction management with respect to social and commercial aspects of such an exposition.

Multi-robot and interaction management is addressed with respect to visitor density and visitor flow. Human-robot interaction is modeled using the SOUL environment.

Concluding we will present and discuss results from the Swiss national exhibition Expo.02 in the time from 15.05.02 to 20.10.02, with over 10'000 hours of total robot operation time and more than 600'000 visitors.

1. Introduction

Public space experiences in recent years are proof of a remarkable progress in mobile robotics. This enabled the operation of a public mass exposition with ten autonomous mobile robots running simultaneously during the Swiss national exhibition Expo.02.

Having several identical robots serving as tour-guide and main attraction of an exposition during a five-month period created a special situation. Men and machine operating in the same space make reliable and safe robot operation mandatory. Ten and a half hours per day, seven days a week over the exposition period imposed high demands on robotics hardware. In addition to this, visitor flow and fun factor of an exposition are important to operators and financiers of a public mass exposition.

Developing the interactive part for the exposition meant always taking into account the demand for visitor flow and entertainment. These criteria translate more or less directly into guided tour and unconstrained interaction. We propose the SOUL (Scenario Object Utility Language) system [8] as a solution for controlling guided tour and interaction.

The fact of having several robots at disposal makes them readily available for the visitors, but requires a resource management for the exposition space. The autonomous nature of our robot evokes the question of centralized or distributed system architecture



Figure 1: RoboX interacting with people visiting Expo.02.

Closely related with the multi-robot control, we try to support the natural visitor flow direction from entry to exit by constraining the displacement of the robot.

In general, high visitor density and a rapid visitor flow constrain interaction. Since these parameters are external, we seek a system allowing for a maximum of interaction under the current conditions.

Concluding, we evaluate these elements under real world conditions based on experience gained at the Expo.02 (figure 1).

2. Related work

We will look at mobile robot experiences in public spaces, arguing that the mobility of the platform and the direct presence of both human and robot render interaction particularly interesting. Therefore it is important to improve human robot interfaces [1] to help visitors interacting with mobile robots.

Face and emotional state machines were found useful elements for tour-guide-robots [2]. The Mobot Museum Robot Series [3,4] focused on the interaction.

Robustness and reliability was identified as an important part of a public robot. Several experiences with the museum robots showed further that the visitors do not always behave cooperatively with the robot and switch between seeing it as a simple machine or a tour-guide. Another permanent installation is at the “Deutsches Museum für Kommunikation” in Berlin, where three robots welcome the visitors and invite them to play with a ball [5].

Summarizing, we can state that the development of public robots has to take into account the differences in visitors’ behavior. First of all, the robot needs to sense the presence of visitors in order to react appropriately. We may distinguish if the robot is seeking an interaction or if it is already giving a tour and interacting with someone else [4].

3. RoboX

During Expo.02, the time which visitors can spend with RoboX is rather limited. We decided to use intuitive means of communication in order to use this time as efficiently as possible. The design of the robot uses common features for communication, situating its appearance somewhere between anthropomorphic and machine. The face of RoboX is the intended source of communication helping the visitors to feel more comfortable when communicating with the robot.

Collaborative interaction will mainly take place between one visitor and the robot, but we anticipate that a certain audience of other visitors will follow this interaction.

For good visibility, we designed RoboX (figure 2) to be of approximately average visitor’s height. The robot consists of a mobile base with an interactive top, making the face easy to look at.

Two differentially driven wheels located at the center of the robot allow on the spot turns. Two castor wheels, one at its back and one, with a suspension at its front, ensure the stability of the mobile base. Obstacle avoidance and reliable localization [6] ensure that the robot knows at all times its position and does not collide with visitors or parts of the exposition.

As an additional means of security, touch sensitive plates and foam bumpers ensure that the robot stops if it runs into anything. Two SICK Laser scanners mounted at knee height provide environmental information for navigation and interaction. A camera mounted in one of the robot’s eyes provides additional information for the interaction. Furthermore, the mobile base houses motor controllers, batteries for 10h autonomy, a PowerPC 750 clocked at 400 MHz dedicated for navigation and obstacle avoidance and a Pentium III running at 700 MHz, 128 MB RAM on Windows 2000 for all interaction tasks.

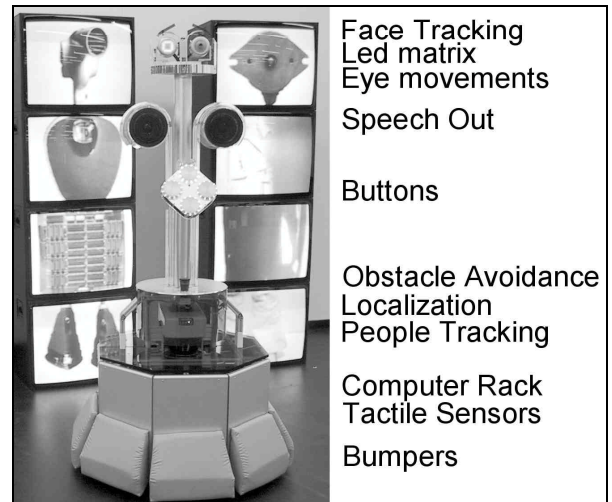


Figure 2: RoboX in front of an exhibit at the Swiss National exhibition, with a description of its main elements.

Both computers can communicate with each other over a 10 Mbit/sec local Ethernet and with a central computer over wireless interfaces to allow monitoring the state of the robot. Technical details are in [7].

4. Interaction at Expo.02

Interaction of visitors with several robots in a public exposition is a complex task. First of all we will present how interaction between RoboX and a visitor is realized. We will distinguish static and dynamic elements which help making each tour of the robot individual. By taking into account dynamic elements we aim at making the robot conscious of its environment.

Since RoboX is giving a tour it will stop at several stations and supply information related to a certain part of the exposition. With the several RoboXs running at the same time, we faced the problem of multi-robot coordination to avoid having several robots intending to go to the same place at the same time.

Finally we will present how parameters like visitor flow and visitor density are taken into account to provide the most of interaction under the current conditions of the exposition.

4.1 SOUL

We will briefly present SOUL, the program controlling the interaction on RoboX. It aims at combining elements of a guided tour with human-robot interaction. The tour the robot is giving presents a certain amount of information on several parts of the exposition. They are practically constant during the period of the exposition. Static scenarios can easily represent this information.

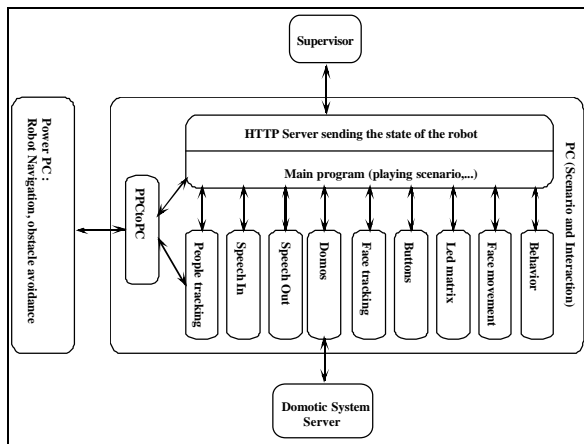


Figure 3: Structure of the interactive system. The supervisor is a separate computer allowing the operator to monitor of the robot's operation.

A SOUL scenario is the succession of robot actions such as speaking, moving and similar actions for a limited amount of time. Intelligent appearance can hardly be achieved by repeating these scenarios over and over again. Therefore, we use changing presentations and methods of adaptive behavior to avoid repetition. One way to avoid repetitive behavior is to provide several alternatives of the presented text and actions. The tools available to the SOUL system for creating such scenarios are shown in figure 3.

In addition to this permutation approach, we aimed at having a robot responding to events which can occur during a tour. Such events can be visitors blocking the robot, hitting its bumpers, playing with the buttons without being asked to, or pressing the emergency button. From the point of view of interaction one can see these signals as a certain acceptance of the robot by the visitor. From the point of view of a guided tour however, they are exceptions and are treated as such by SOUL. Technically SOUL will interrupt the current scenario and execute a corresponding exception scenario telling the visitor that it is aware of his actions, before resuming the tour. RoboX will treat one exception at a time.

4.1.1 SOUL sensors

RoboX is using several sensors and algorithms to achieve awareness of its environment. Simple switches detect events like visitors pressing the emergency button, the interactive buttons, or hitting the bumpers. The obstacle avoidance provides information when visitors are blocking the robot. In addition the robot is aware of visitor presence in its surrounding by means of face and motion tracking [8,9].

Motion Tracking

The main goal of motion tracking is to distinguish between moving and static elements of the environment. Computing differences between

successive scans and clustering the dynamic elements does this. Kalman-filter based tracking establishes a relationship between dynamic elements over time and allows to trace a visitor as explained in detail in [9].

Face Detection

In the left eye of the robot face a color camera is located. Our image processing uses the Intel Image Processing Library and detects and follows skin colored regions. Using visual servoing, the robot can look at the person it is interacting with. Especially with groups of people it is important to express who the robot is addressing.

The main steps of face detection and tracking are:

1. **Skin color detection:** among the different color spaces we chose the RGB space. Green and blue values are normalized using the red channel. This partially cancels differences in illumination. Fixed ranges for blue, green, and brightness values are accepted as skin color. Erosion and dilation are performed on the resulting binary image to remove small regions.
2. **Contour extraction and filtering:** the binary image is clustered and the contour of each cluster is extracted. We apply heuristic filters to suppress skin color regions that are not faces. These filters are based on rectangular areas, their aspect ratio, and the percentage of skin color and on the morphology of the skin color region.
3. **Tracking:** the system tries to update the positions of already tracked regions, based on the position of clusters in the current image. Clusters that remain unassigned to previous tracks are added and tracked until they leave the camera's field of view.



Figure 4: People seen from the robot. Skin colored regions show a light border.

Information gathered from face tracking is used in several parts. Together with the motion tracking it helps to verify the presence of visitors. It is used to look at the interlocutor. Finally, it triggers the behavior engine, which is explained later on.

4.1.2 SOUL expression

There are three interfaces available to communicate with the visitor. To express itself RoboX is using synthesized speech in English, French, German and Italian using Mbrola [9] and LAIPTTS [10].

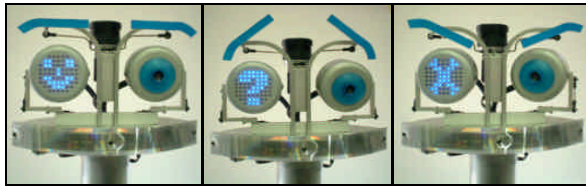


Figure 5: Three facial expressions. From left to right: happy, surprised, and angry.

The interactive buttons can be illuminated to indicate the mode they are in (language choice, yes/no, etc.). For visitors, the most expressive part remains the face (figure 6) imitating several grimaces, emphasized by a small LED matrix mounted in one of the eyes to display symbols and short animations.

4.1.3 Behavior component

Our aim was to create individual tours according to the visitor's action, but so far their action affects the tour only shortly by starting the appropriate exception scenario. With the behavior component presented in [11], RoboX started to accumulate impressions during a tour and to adapt its behavior accordingly.

For the representation of this internal state we chose the Arousal-Valence-Stance affect space [12], because of its three dimensional representation which is very intuitive to use. In this space (figure 7), six basic expression regions are defined as: sadness, disgust, joy, anger, surprise, fear and a reference expression that can be considered as a calm state.

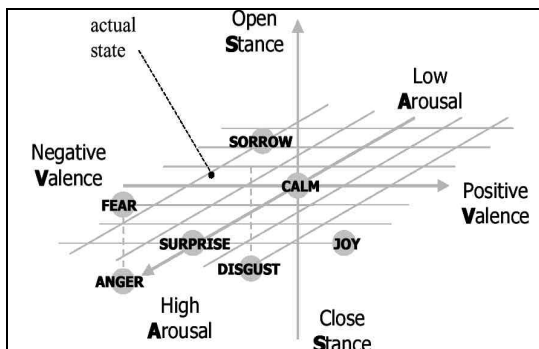


Figure 6: Six basic expressions and the neutral expression in the AVS space.

The internal state is mainly communicated using the synthesized voice and face movements, in some cases symbols are shown on the LED screen. Figure 6 shows where the internal states are located in the AVS space.

4.2 Multi-robot coordination

Figure 8 shows the layout of the exposition. Presentation stations are defined near particular objects in the expositions. There are several places where robots welcome visitors, thus tours can start

simultaneously. There are fifteen presentation stations all over the exposition space. Finally, there are goodbye stations close to the exit. Each station corresponds to one scenario in the SOUL system, providing visitors with explanatory or entertaining information. Tours can be created by a succession of several presentation stations. Two stations are not in any of the tours but permanently occupied by a dedicated robot. They have tasks of taking pictures from the visitors and presenting a slide show.

Working with multiple robots makes resource allocation an important point. In order to avoid having several robots presenting the same object an assignment has to be made at a certain moment.

In the beginning we solved this problem by assigning several stations exclusively to one tour which was operated by one robot all day. The tours were designed to have robots working spatially separated in order to avoid collisions among robots.

With ten robots operating the exposition this was no longer feasible, since it would result in tours of one or two stations only and thus quasi-static mobile robots. Improved obstacle avoidance allowed the robots to see each other and to avoid collisions. This enabled a dynamic assignment of stations to a robot for the duration of its presentation. The station is released thereafter and can be used by other robots. This is modeled by a list of all stations and their state. Stations are free until reserved by a robot. The robot can choose among the free stations in order to avoid deadlocks. Care has to be taken that robots decide successively to avoid several robots choosing the same goal.

Regarding the complexity of the communication scheme we opted for an additional central instance gathering data from all the robots and providing it on demand. Thus multi-robot coordination in our case is based on local decisions by each robot. After terminating a presentation, but before proposing a next station to the visitor, the robot asks the state of all exposition stations from the global instance. This request blocks the global supervisor until the robot reserves a specific station. The decision, which station to reserve is based on the free stations, the list of stations included in this tour and the stations already visited. The first free and unvisited station in the tour list is reserved and then presented to the visitor

4.3 Visitor density and interaction

Expo.02 is a mass exposition with several thousands visitors per day. During the preparation of this project we anticipated up to 500 visitors per hour, which assuming a 15 minutes stay inside the exposition results in 125 visitors enjoying the robots at the same time. Visitor behavior cannot be anticipated. To ensure a functioning of the exposition at all times four exposition modes were defined:

1. *Wait for visitor*: with few visitors, robots wait for one to come close enough before starting to talk and ask him which station he likes to see.
2. *Visitor's choice*: more visitors, the robot can ask permanently whether the visitor wants to go to a station without talking to anyone in particular.
3. *Robot's choice*: even more visitors, the robot will decide what is the next station and go there without asking.
4. *No move*: too many visitors for the robot to move, each robot will stay at one station and present it permanently.

The exposition mode is defined manually by the staff. It is included in the data provided by the global supervisor, so every time the robot requests the state of the exposition's stations it receives an update of the state and can adapt accordingly. Figure 10 shows how this is taken into account by the SOUL system:

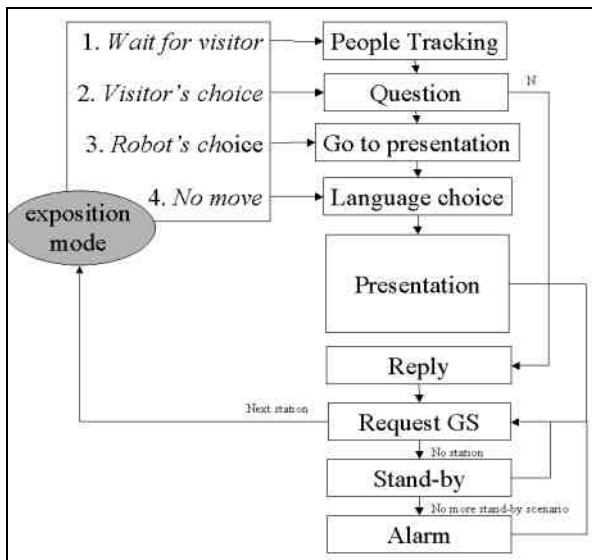


Figure 7: Structure of the SOUL sequence for a typical presentation station. GS stands for Global supervisor.

Depending on the exposition mode the scenario starts either with people tracking (*wait for visitor*), the question “Do you want to see ...?” (*visitor's choice*), the robot moving to the station (*robot's choice*) or directly with the language choice (*no move*).

These blocks are executed successively except if the visitor declines to go to a station. In this case SOUL jumps directly to the reply block commenting in some way the visitor's decision.

The request to the global supervisor is executed either after the reply block or after the presentation of a station. It provides all empty stations at this time, the choice is made as explained in the paragraph above.

If no empty station is available or all empty stations have already been visited during this tour, the robot

starts one of several stand-by scenarios. These are presentations, which are not located at a specific place in the exposition. The robot talks about itself, sings, or makes funny faces, while waiting for a presentation station being released by another robot.

After the stand-by scenario, the robot requests once again the exposition's state to find a free presentation station. If one is found the next scenario is run. Otherwise the robot continues to play stand-by scenarios and to request the global supervisor until either a presentation station is available or it has run out of stand-by scenarios. In the latter case, the global supervisor will give an alarm and the staff can interact. Starting the robot with another tour may solve this problem. To avoid having several robots giving the same presentation, a station remains blocked by one robot until it starts moving on to the next station.

4.4 Visitor flow

We estimated the average visit to 15 minutes in order to meet the visitor flow requirements. Previous test in our lab [8] proved it difficult for the robot to make visitors leave. In general, their interest span is not directly related to the duration of a tour.

Visitor flow is channeled by two factors: the number of stations the robot visits, and leading the visitor constantly closer to the exit throughout the tour. The robot visits S stations, where S can range from 1 to 15. Then it executes the goodbye scenario, which is located near the exit. By this proximity we aim at encouraging visitors to leave. The goodbye scenario is special in the way that it resets the list of stations visited during a tour and sets the counter of stations visited back to zero.

Throughout the exposition a tour will always lead visitors closer to the exit. This eases navigation and helps maintaining the visitor flow. Technically this is realized by a list of possible next presentation stations. Each presentation scenario is assigned an individual list, containing only stations to support the direction of the main visitor flow. When requesting exposition state from the global supervisor, the robot will seek only stations which it has not yet visited and are closer to the exit than it is currently.

5. Results

In the period from 15.05.02 to 17.07.02, an average number of 4427 people visited the exposition every day. The minimal number of visitors was 2299 the maximum was 5473. The average number results in a visitor flow of 422 persons per hour on 315 m² exposition space with up to ten robots in operation. This corresponds to a load of 84.3% percent of the planned maximal flow of 500 visitors. The maximum flow corresponds to a load of 104%.

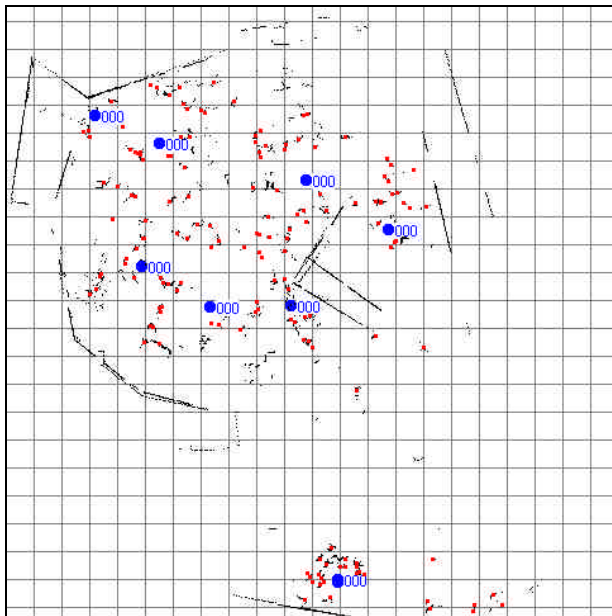


Figure 8: Exposition Robotics, 20th sept. 2002 at 14.20h: 8 robots running and approximately 120 visitors. Motion detection on all robots senses 156 moving objects sensed. Small circles corresponds to motion sensed by the robot, big dots represent the positions of the robots.

The global supervisor system is operational since 01.07.02. So far, the exposition mode *visitor's choice* was active approximately 95% the mode *robot's choice* 5% of the time. We experienced ten days with more than 5000 visitors and even in this crowded environment robots managed to move to their goal in a reasonable time, so that the mode *no move* was never used. Up to date the mode *wait for visitor* was never used, since interested visitors surround the robots most of the time. Figure 11 shows a snapshot of the exposition with such situations.

With currently three stand-by scenarios, alarms of a robot running out of those scenarios occurred approximately once a week. With two additional stand-by scenarios we aim at reducing this rate further.

Visitors stay between 10 and 45 minutes with the robots. The average stay is 17 minutes. 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 move on to the next robot or even stay 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 is the actual tour length. The average number of visitors during the first 17 days of operation of the global supervisor rose slightly to 4576 per day. This makes it hard to prove a quantitative effect on the visitor flow. However, observation of the crowd shows that visitor appreciate

having the choice to go to a station. This adds a little interactive element to the tour.

6. Conclusion

During over 10'000 hours of operation, 600'000 visitors interacted with the robots in the time from 15.05.02 to 20.10.02. SOUL seems to provide an appealing compromise of a guided tour and unconstrained interaction. For the last two and a half months the exposition was running with a multi-robot resource control scheme taking into account the visitor density and supporting visitor flow. Quantitative parameters like visitor flow and density meet the planned parameters. By enhancing environmental perception, we aim at creating even more convincing human-robot interaction.

References

- [1] Burgard W., et al, "Experiences with an interactive museum tour-guide robot". AI (114), 1999, pp. 3-55.
- [2] Thrun S., et al, "MINERVA: A Second-Generation Museum Tour-Guide Robot". In: Proc. ICRA, USA; 1999, vol.3; pp.1999-2005.
- [3] Willeke T., et al, "The History of the Mobot Museum Robot Series: An Evolutionary study". In: Proc. FLAIRS 2001, Florida.
- [4] Nourbakhsh I., et al, "An Affective Mobile Educator with a Full-time Job". AI, 114(1-2), October 1999, pp. 95-124.
- [5] Graf, B., et al, "Konzeption dreier Roboter zur Unterhaltung der Besucher eines Museums". In: VDI-Berichte 1552, VDI-Verlag, Düsseldorf, 2000.
- [6] Arras, K., et al "A Navigation Framework for Multiple Mobile Robots and its Application at the Expo.02 Exhibition". Workshop: Robots in Exhibitions, IEEE/RSJ.IROS, 2002, Switzerland.
- [7] Tomatis N., et al, "Design and System Integration for the Expo.02 Robot". Workshop: Robots in Exhibitions, IEEE/RSJ.IROS, 2002, Switzerland.
- [8] Jensen, B., et al, "The interactive autonomous mobile system RoboX". In: Proc. IEEE/RSJ.IROS, 2002, Switzerland.
- [9] Jensen, B., et al, "Narrative Situation Assessment for Human-Robot Interaction", In: Proc. IEEE ICRA, 2003, Taiwan.
- [9] Dutoit, T., et al, "The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes". In: Proc. ICSLP'96.
- [10] Siebenhaar-Röllli, B. et al, "Phonetic and Timing Considerations in a Swiss High German TTS System". In: E. Keller, G. Bailly, A. Monaghan, J. Terken & M. Huckvale, "Improvements in Speech Synthesis", pp. 165-175, Wiley & Sons, 2001.
- [11] Mayor L., et al, "Improving the expressiveness of mobile robots". In: Proc. ROMAN(2002), Berlin, Germany.
- [12] Ekman P., R. Davidson, "The Nature of Emotion: Fundamental Questions", Oxford University Press, New York, 1994.