Deliverable: D5.4

Demonstrator Evaluation Report

Consortium

UNIVERSITEIT VAN AMSTERDAM (UvA)

IDMIND - ENGENHARIA DE SISTEMAS LDA (IDM)

UNIVERSIDAD PABLO DE OLAVIDE (UPO)

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE (ICL)

UNIVERSITEIT TWENTE (UT)

Grant agreement no. **288235**

Funding scheme **STREP**

DOCUMENT INFORMATION

Project

Project acronym:	FROG
Project full title:	Fun Robotic Outdoor Guide
Grant agreement no.:	288235
Funding scheme:	STREP
Project start date:	1 October 2011
Project duration:	36 Months
Call topic:	ICT-2011.2.1 Cognitive Systems and Robotics (a), (d)
Project web-site:	www.frogrobot.eu

Document

Deliverable number:	D5.4
Deliverable title:	Demonstrator Evaluation Report
Due date of deliverable:	M36 - 30 September 2014
Actual submission date:	4 November 2014
Editors:	UPO
Authors:	All Partners
Reviewers:	All Partners
Participating beneficiaries:	All Partners
Work Package no.:	5
Work Package title:	Integration, Evaluation and Demonstrator
Work Package leader:	IDMind
Work Package participants:	All Partners
Estimated person-months for deliverable:	10
Dissemination level:	Public
Nature:	Report
Version:	1
Draft/Final:	Draft
No of pages (including cover):	63
Keywords:	Evaluation

Contents

1. Introduction

Two main test sites have been considered in the FROG project. The Lisbon Zoo and the Royal Alcázar in Seville.

In the second year of the project, an initial version of FROG was deployed in the Lisbon Zoo, and in particular from September 18 to September 22, 2013, for the second review meeting of the project. During these experiments, several different subsystems of FROG were tested and evaluated: such as the platform itself, localization and navigation, person detection and localization and affective computing. Only some of them had been integrated at that point.

In the final phase of the project, the full FROG robot, with all its subsystems integrated, has been deployed in the Royal Alcázar in Seville. Two deployment sessions were performed:

- From June 16 to June 27, 2014, with the first week devoted to final integration of the modules
- From September 15 to September 26, 2014, with the first week devoted to the final setup

Throughout those weeks the robot performed a set of full missions, guiding the tourists that were visiting the monument during those days. Between September and June, the main difference was in the setup of the some of the Points of Interest of the tour.

After those missions, some of the guided users were interviewed to have feedback about the user experiences. Furthermore, on September 24, 2014, a workshop with tour guides took place, involving a tour with the robot; and, on September 25, 2014, a Peer Group Workshop with colleagues from other EU related projects .

This document summarizes the outcomes from all these experiments, as well as the feedback obtained from these different evaluations.

The structure of the document is as follows: first, some relevant aspects and lessons learnt for the different individual modules during these experiment sessions are described. They are followed by the evaluation of the FROG robot as a whole. This last part includes the feedback from the users workshops. The document ends with conclusions.

2. FROG Robot Platform

The design of the robot platform had to find a good trade-off between the different requirements related with on-board technology, operating conditions and safety issues, and smoothly integrate them into an appealing design. The complete robot solution is presented in Figure 2.1.

Fig. 2.1. FROG: the Fun Outdoor Robotic Guide

The development and assembly of the FROG robot platform has been divided in two phases. The first phase included the platform base mechanics with the motors, batteries and low-level electronics. The resulting platform base can be adapted to serve different applications. A second phase, which specifically targets FROG's use case scenarios, included the installation of high-level devices mounted over an upper structure and the design and production of an outer shell.

2.1. Robot Platform Base Main Features

Regarding the development of the robot platform, FROG's Description of Work (DoW) targeted three main features for the platform:

- maximum velocity: 1.5 m/s
- payload capacity: 40 to 50 Kg
- power autonomy: 4 to 6 hours

Deliverables D1.2, D1.3 and D1.4 provide a complete description of the developed robot platform. An assembled platform base is shown in Figure 2.2 and its main technical specifications can be summarized as follows:

- Kinematics: Differential Four-Wheel Skid Steering
- Weight without batteries: 39 Kg
- Weight with batteries
    - When using 8 x 12V/17Ah lead free batteries: 87 Kg

- When using 8 x 12V/17Ah LiFePO4 batteries: 56 Kg

- Payload Capacity: 40 Kg
- Maximum Velocity: 1.6 m/s (low-level software limited to 1.0 m/s)
- Acceleration: 1 m/s2 (low-level programmed)
- Emergency Stop Acceleration: 3.3 m/s2 (low-level programmed)
- Batteries Autonomy: 2 to 8 hours
    - Using a minimal configuration with only one computer, the robot can have an autonomy of 8 hours. But with all the installed equipment (e.g., the final FROG solution includes four embedded computers, see D1.4) the autonomy gets close to 2 hours.

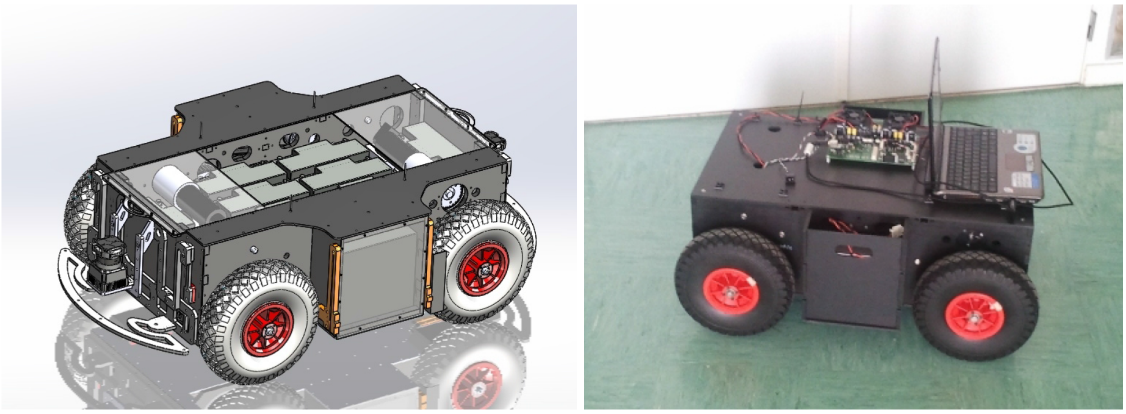

Fig. 2.2. Robot platform base

2.2. Robot Kinematics Performance

Regarding the platform kinematics, the first integration meeting demonstrated two facts. The first one was that the use of a two-wheeled configuration, with a castor wheel, had some drawbacks as the robot faced some difficulties when the castor wheel (with no traction) went through small gaps in the ground. The second one, was that the payload resulting from all the on-board technology from the different partners and the inherent battery power, demanded a more stable platform with a bigger footprint. Based on this, IDM opted for the development of a 4-wheel skid steering platform.

FROG had to deal with various terrain types in its scenarios. The robot must maintain a good performance on all of these terrains and while moving from one to the other. Safety issues raised by the terrain types for the visitors that may follow the FROG are not a matter that the FROG project has to deal with. However, the fact that these are sites that draw many visitors does mean that these terrains are well maintained. Both sites have outdoor areas that may be wet due to rain, cleaning operations or irrigation of the plants.

The outdoor and indoor areas that FROG visits on its missions at the Royal Alcázar in Seville and in the Lisbon Zoo present the following types and conditions:

2.2.1. Mission terrains – Royal Alcázar, Seville

At the Royal Alcázar, none of the surfaces that the FROG robot has to traverse are loose laid. All outdoor paving materials are set in concrete. The gaps between bricks or stones are filled with concrete. Some grouting is flat, some has profiling lines as if drawn with a finger in wet cement. Indoors, some marble or tiled floors are grouted with special materials. Other (marble) tiles are laid so close together and what little grouting there is is so highly smoothed that, to the touch, there is only one surface.

Both indoor and outdoor surfaces are kept scrupulously clean – most of the work being done in the early morning before visitors arrive and maintained to a high standard during the day. The route performed by the robot is now presented.

After moving from the stainless steel docking station, the robot turns and leaves the shop with its wooden laminate floor passing over a stainless steel strip and a brick doorstep, onto a chequered courtyard of natural flagstones and cobbles set in concrete (Figure 2.3)

Fig. 2.3. Left: FROG robot on its stainless steel docking station, surrounded by wood laminate floor. Right: FROG making the transition from wood laminate over a brick doorstep onto flagstone/cobble chequered courtyard.

From there the mission moves into the Lion's Courtyard, moving from a chequer of flagstones and cobbles in concrete, across bricks laid in concrete and onto smooth red bricks laid in concrete in a herringbone pattern (Figure 2.4).

Fig. 2.4. Left: transition from flagstones and chequer flagstones / cobbles in concrete, over dark red bricks laid

in line in concrete, to herringbone laid red bricks in concrete. Right: overview red herringbone laid bricks in concrete.

The herringbone laid smooth red bricks continue through the arches (down an incline of 3°) into the Hunting Courtyard – and area of herringbone laid dark red brick fields between smooth white stone slab borders (Figure 2.5)

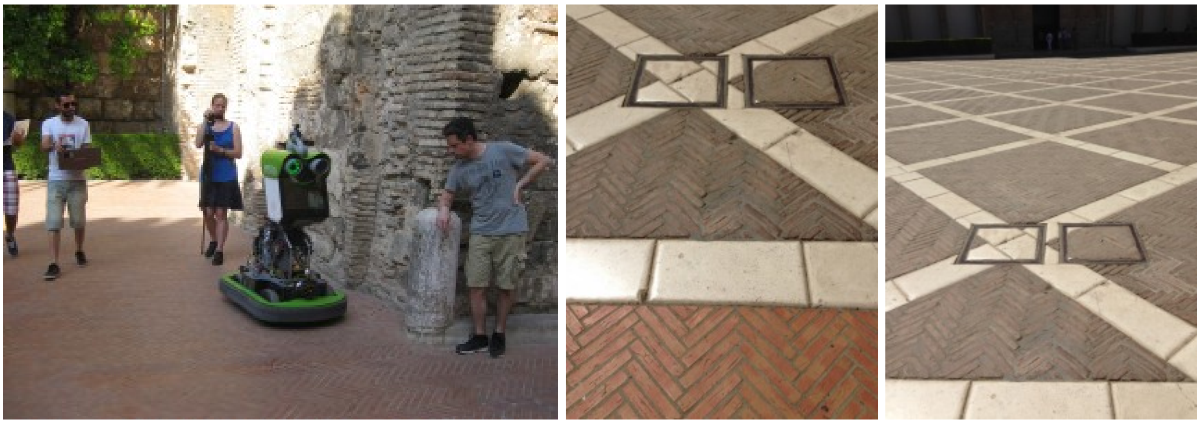

Fig. 2.5. Left: through the triple arched wall. Middle: transition from herringbone laid smooth red bricks in concrete to fields of rougher dark red bricks herringbone laid in concrete between smooth white stone slab dividing lines. Right: fields of rougher dark red bricks herringbone laid in concrete between smooth white stone slab dividing lines – including 2 custom-made drain covers.

On leaving the Hunting Courtyard, the mission continues between curved curbs and up a 6° incline into a covered walkway (Figure 2.6).

Fig. 2.6. Left: one of the curbs before the covered walkway. Middle and right: slightly inclined transition from fields of rougher dark red bricks herringbone laid in concrete between smooth white stone slab dividing lines to fields of cobbles laid in concrete between flagstone slabs.

Looking down the performed route (Figure 2.7).

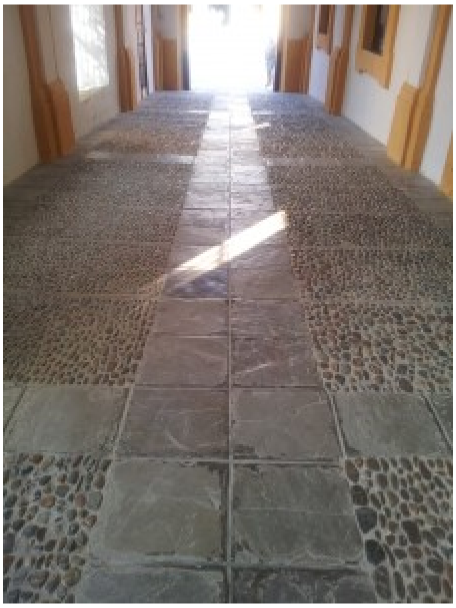

Fig. 2.7. Stone flags / fields of cobbles in concrete.

The route continues through the Crossing Courtyard with a transition from patterned cobblestones in concrete over row of stone flags onto herringbone pattern laid brick with deep-set gate wheel tracks with wooden fillers (Figure 2.8)

Fig. 2.8. Left: overview of herringbone laid brick Crossing Court showing gate wheel gutters with wooden fillers. Right: air vent grate in raised brick surround along herringbone laid brick path.

The route now passes from outdoors onto the close-laid marble floor in Tapestry Room (Figure 2.9). Please note that this door is set in a larger door and that the whole is opened for the robot thus removing this wooden door sill.

Fig. 2.9. Transition from the Crossing Courtyard's yellow brick onto 3-coloured marble tiles.

From the Tapestry Room the route continues to the Vault Room with a transition from marble tiles to marble tiles with glazed ceramic tile pattern (Figure 2.10).

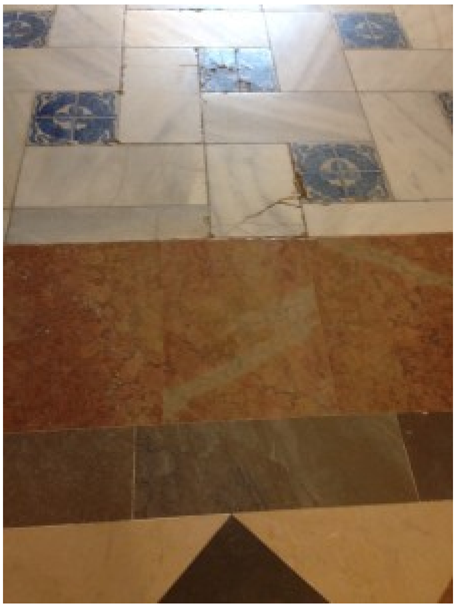

Fig. 2.10. Transition from marble tiles to marble tiles with glazed ceramic tile pattern

The route leaves the Vault room via a wheelchair ramp to mount a step in marble tile floor with glazed tile pattern (Figure 2.11).

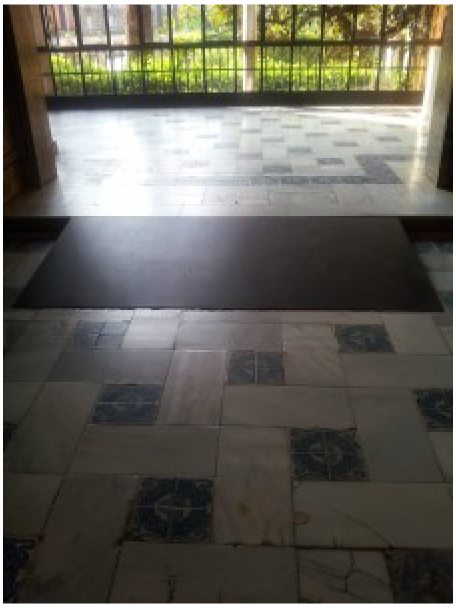

Fig. 2.11. Wheelchair ramp with anti-slip surface.

The route goes outdoors again towards the end point of the mission with a transition from marble with glazed tile pattern, over flagstones surround, to chequer: flagstone/cobbles in concrete (Figure 2.12).

Fig. 2.12. Left: transition from marble to chequer. Right: finalizing the scenario for the end of the mission.

From here the robot returns to the starting point through a covered walkway with natural flagstones (Figure 2.13).

Fig. 2.13. Covered walkway with natural flagstones.

2.2.2. Mission terrains – Lisbon Zoo

The Lisbon Zoo mission begins on large slabs set in concrete with deep grouting, close to the Valley of the Tigers entrance (Figure 2.14).

Fig. 2.14. Large slabs set in concrete with deep grouting

The robot then goes through the Valley of the Tigers complex (Figure 2.15).

Fig. 2.15. Transition from stone slabs to smooth concrete (painted red).

When FROG leaves the Valley of the Tigers it passes over a section of interlocking brick paving, a grate covering a drainage channel onto a bitumen path with fine aggregate (Figure 2.16).

Fig. 2.16. Looking back towards the Valley of the Tigers – brick paving, grate, bitumen with fine aggregate.

From here the mission carries on to the Birds. This section of the route stays on bitumen with fine aggregate and has some slight inclines (Figure 2.17).

Fig. 2.17. Bitumen with fine aggregate and deep-set drain cover.

The route continues past the Sea Lions, staying on bitumen with fine aggregate (Figure 2.18).

Fig. 2.18. Passing the Sea Lions – bitumen with fine aggregate.

The next presentation point along the route, The Gorilla's Belvedere, is on slightly inclined interlocking paving bricks (Figure 2.19). The FROG stops close to this corner.

Fig. 2.19. The route follows around to the left – interlocking brick paving.

And after the corner the path climbs the slight incline, passing from interlocking paving bricks to Portuguese pavement over separating wide concrete bands (Figure 2.20, left). The route continues then through a long stretch of Portuguese pavement under a pergola (Figure 2.20, right).

Fig. 2.20. Left: transition from interlocking brick paving, over concrete bands, to Portuguese pavement. Right: a long stretch of Portuguese pavement under a pergola

After the pergola walk, the route follows around to the right over interlocking bricks, through impressive gate posts (Figure 2.21), on towards the Primate Temple.

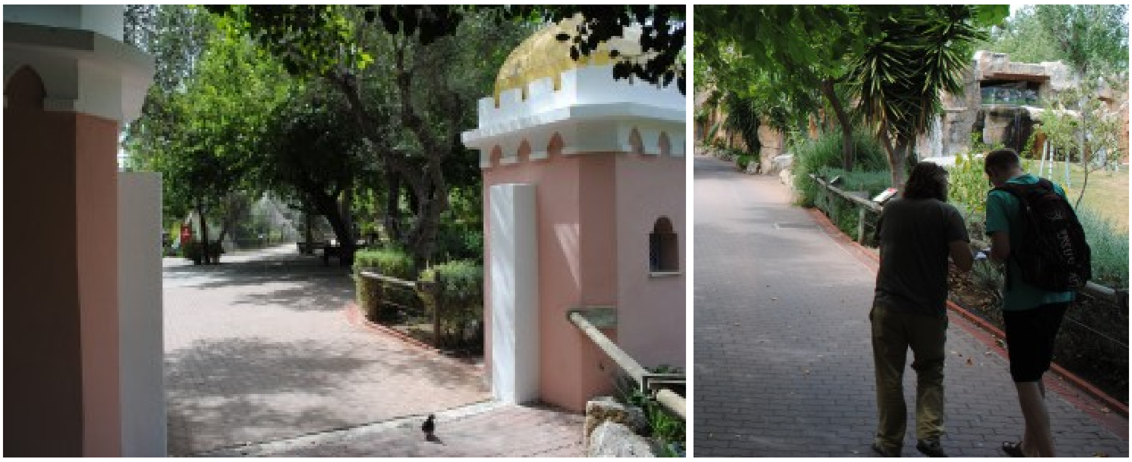

Fig. 2.21. Left: through the gate. Right: on towards the Primate Temple – interlocking bricks.

The routes continues through the Primate Temple (Figure 2.22).

Fig. 2.22. Large stone slabs laid in concrete with wide grouting and a hairpin bend at the top of the incline

Once inside the Temple, the surface is smooth concrete painted red (Figure 2.23).

Fig. 2.23. Inside the Primate Temple

On leaving the Primate Temple the route winds over a wooden separator onto large stone slabs set in concrete, quickly crossing a decorative line of glazed bricks, a drainage grate set in concrete and once again onto interlocking bricks (Figure 2.24).

Fig. 2.24. At the end of the mission – a wide variety of surfaces in an S-bend in the path

The route now follows the path downwards and the robot makes its way back to the starting point (Figure 2.25)

Fig. 2.25. Inclined ramp on the way back.

2.2.3. Overall Kinematics Performance

As can be seen from the above descriptions, the FROG had to deal with different terrains along its route. The platform has proved to be very stable and effective during the different tests. The platform is making use of four puncture-proof foam-filled tyres made of polypropylene rim (see Figure 2.26, left), which had a good behaviour over almost all the terrains. Over terrains with higher friction properties, as the bitumen with fine aggregate, we noticed an increase on the motor power consumption and also on the motor surface temperatures. These types of terrains necessitate the use of wheels with a smoother surface and consequently less grip.

Some tests were also made with two omnidirectional wheels mounted on the back axis of the robot (see Figure 2.26, right). This configuration facilitates the rotation of the robot as the robot behaves almost as a traditional two-wheel differential robot. This looks like a good alternative for use cases in which the robot has to repeatedly make turns on itself and/or in the presence of a high friction terrain. These tests were made using UT's FROG robot replica at a later stage of the project. The consortium decided not to include this solution in FROG because the project was getting close to its end, without time of major changes, and this would impact the kinematics of the robot and consequently the navigation module. Moreover, the omnidirectional wheels are a bit noisy when the robot moves on a terrain with irregularities (e.g., Portuguese cobblestone pavement) and much of the FROG's route in the Alcázar is over a terrain with irregularities as depicted, for example, in Figures 2.4 or 2.8.

Fig. 2.26. Left: foam-filled tyre on polypropylene rim. Right: omnidirectional aluminium wheel with bearing rollers.

2.3. Low-Level Safety Loops

Deliverable 1.4 contains an extended description of all low-level safety measures included in FROG. For human and robot safety purposes the DoW proposes the creation of fast and reactive low-level safety loops. With that purpose two different sensors were included in the platform and connected directly to the low-level electronics hardware: a ring of switches associated with a foam bumper and a ring of sonars. The conjunction of these two systems should provide a very fast and reactive safety layer to avoid or stop the robot: the sonars should create a minimum safety distance around the robot, allowing the robot to avoid or stop in the presence of an obstacle/human; and the bumper as the last safety system to hard-stop the robot in case of a collision.

Both were implemented and evaluated. With that purpose, the platform base with a mock-up upper structure was controlled with a joystick while using the sonars and bumpers to stop the robot (see Figure 2.27). A series of tests at different velocities has been performed to determine the minimum reaction distance that allowed the robot to stop safely before the obstacle.

Fig. 2.27. Safety layer evaluation

With the velocity set to 1 m/s on a horizontal flat floor with low adherence, the robot was able to stop in less than 15 cm. The safety stopping distance for the sonars was then set to 30 cm from the robot. The method was tested and validated in the robot while driving it with the joystick and the robot was able to stop without hitting the obstacles using the sonars. To test the bumpers the sonars were disconnected and the robot was able to stop almost immediately after the collision. To increase the safety of the whole system a large soft foam bumper was placed around the robot (see Figure 2.28). This foam bumper was connected to 12 switch sensors which are able to detect collisions in eight different directions.

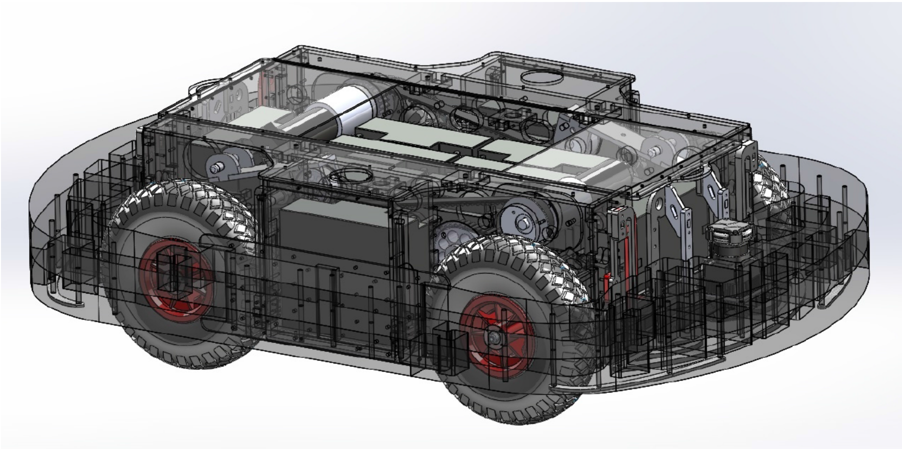

Fig. 2.28. Foam bumper

Low-level safety sensors were integrated in the robot and used with the navigation algorithms. After some trials, the partners realized that the use of the sonar sensors in the crowded use case scenarios was almost impossible as the robot was being too “cautious” and stopping even when it had enough space to move. Moreover the Navigation model needed to be well tuned to be able to have an optimal performance. Whenever the sonars' low-level safety algorithms were being computed a deviation arose between the real robot and the model, which caused the robot to lose its correct position on the map. After the first integrated tests it was decided that the use of sonars in the low-level safety loops was creating more problems rather than solving the ones that they were intended to avoid.

Therefore it was decided to remove the sonars from the low-level safety loops and include a tilted laser scanner in the front neck of the robot. This would allow the detection of obstacles close to the front of the robot. This laser was integrated in the obstacle avoidance navigation algorithms performed by the Navigation computer. The sonars are still being used but only to detect obstacles in the dead zones of the navigation lasers.

Regarding the bumpers, they are included in the low-level loops allowing to hard-stop the robot in case of collision. To inform the Navigation computer about an actuation of the bumper a hard-stop status bit has been included. This status bit is set every time the low-level control loop gets the control of the robot. To recover the control of the robot, the Navigation computer is able to control the time that this status bit is enabled after a collision, from 0 s to 255 s. If the hard-stop time is set to 0 s the robot will not stop in case of a collision, or, it can also be used to remove the robot from a collision.

2.4. Protection From Environmental Conditions

Regarding the protection from environmental conditions, the development of the robot platform was targeting the following specifications from the DoW:

- the ingress of dust would not be entirely prevented, but it could not enter in sufficient quantity to interfere with the satisfactory operation of the robot;
- the robot should be also protected against water falling and accidental projections of water;

- the targeted IP reference code was IP54.

The design of the platform and its outer shell took these specifications into consideration and the final system was constructed to comply with them. Some small compromises had to be taken regarding the need of access to the installed equipment by the partners during integration and test meetings. These preclude efficient isolation of some points (e.g., camera apertures).

Moreover, the robot is not supposed to work in the rain but in case it starts raining, the robot outer shell gives it enough protection so that it can safely return to its shelter. A rain sensor was installed in the robot head to detect this event.

3. Navigation and Localization

3.1. Localization

The localization module developed in the project is in charge of supporting the navigation as well as the Augmented Reality (AR) components of the robot. Different tests were performed in both scenarios considered: the Lisbon Zoo and the Royal Alcázar of Seville. These scenarios consist of partially non-planar and GPS-denied areas, due to the existence of ramps and mixed indoor and outdoor regions. A map-based localization algorithm is thus used in FROG. Deliverable D2.1 describes both the mapping and localization procedures, and analyzed and benchmarked them using these tests.

The mapping process is significantly more complex in the Lisbon Zoo than in the Royal Alcázar of Seville, owing to irregularities of the Zoo, the absence of well-defined structures and the presence of numerous ramps that may be detected by horizontal lasers and interfere negatively in the generation of the map, which is based on the robot's odometry and laser measurements. The resulting maps are shown in Figs. 3.1, overlaying CAD maps of both scenarios with good accuracy.

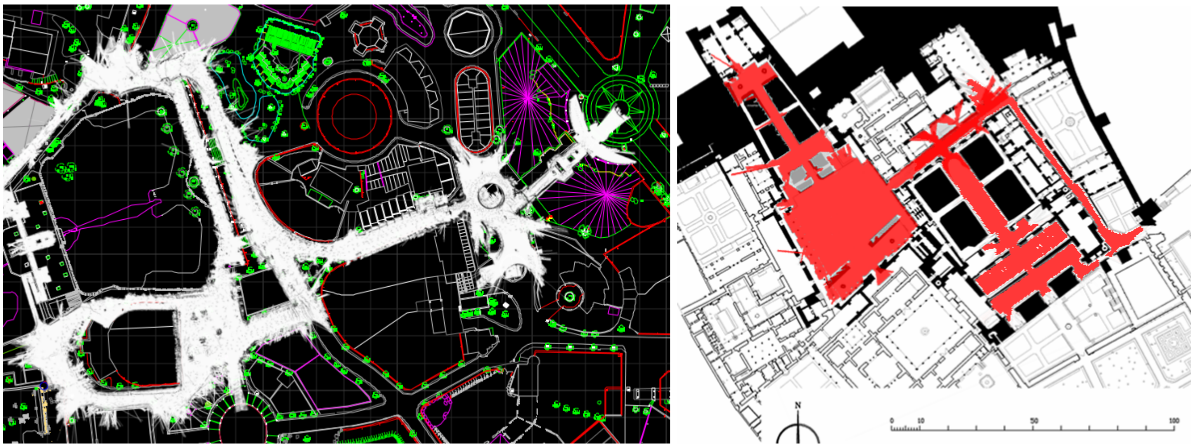

Fig. 3.1. Left: 2D occupancy map of the Lisbon Zoo overlaid on a CAD drawing of the Zoo. Right: Occupancy map of the Royal Alcázar overlaid on a planimetry of the complex.

3.1.1. Localization Accuracy

The accuracy of the localization algorithms was evaluated in the Lisbon Zoo by deploying some ARUCO markers [1] (see Fig. 3.2) on the navigable area. Theses markers are quite easily detected with low time of processing and error rates (as related in [1]).

In Deliverable D2.1 it is explained how these markers were deployed in the Lisbon Zoo and their ground truth pose obtained. Once ground truth pose of markers was obtained it was used for testing localization accuracy. We recorded data during autonomous navigation, and processed offline the divergences in localization of markers by comparing the local position of the detected markers (in the robot reference) and their estimated relative pose based exclusively on robot's pose in the map (see Fig. 3.2).

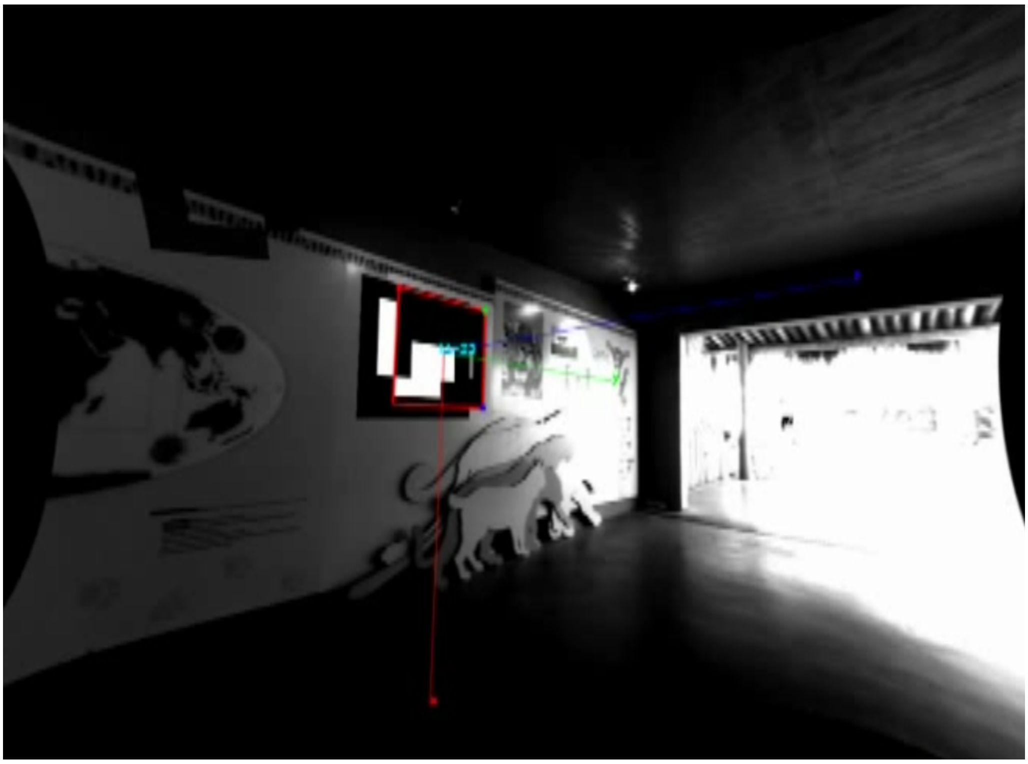

Fig. 3.2. Real ARUCO marker detected (red lines) and estimated marker based on robot's pose (reprojected on image).

The estimated parameters to validate accuracy are: absolute error in position (meters), absolute error in orientation (radians) and absolute error in pixels (in image). These parameters were chosen in order to test not only augmented reality content displayed on the robot's screen (pixels and radians) but also the content projected by the back projector (radians and meters) over walls in special POIs. The localization accuracy is summarized in Table 3.1.

	Small markers		Big markers	
Experiment	Position error (m)	Angle error (rad)	Position error (m)	Angle error (rad)
Lisbon Zoo	0.4055	0.1235	0.8193	0.106

Table 3.1: Global medium error for the different markers in Lisbon Zoo. The angle error is considered as the angle difference between real marker and estimated one as seen by the robot.

The augmented reality content was fully working for the final demonstration realised in Royal Alcázar of Seville during the third year. This allowed us to make real tests online with real content and real POIs where special AR content is played (which is based only in robot's pose). The tests realised were done successfully, and the robot was localized with enough accuracy for playing the content. The estimated error in orientation was less than 5 degrees and in position less than 15cm. The worst error in pixels was observed at the Tapestry Room, with an error of 129 pixel (10% of the image width) in the screen when the distance between the robot and the wall is around 7 m., as can be seen in the screen capture of Fig. 3.3.

Fig. 3.3. AR content displayed in the FROG screen in the Tapestry Room of the Royal Alcázar of Seville. The green marker determines the reprojected image over the real one.

3.1.2. Localization Robustness

During final the experimental sessions, the localization system was working properly 23 out of the 26 full tour experiments (these experiments are summarized in Section 7). And in the other 3 experiments, the Inertial Measurement Unit (IMU) used to estimate the angular velocities stopped working and the fail-safe mode of the localization mode prevented continuation of the experiment, and the system was working properly until the sensor fault was detected. The faulty IMU was replaced during the second series of experiments. The localization module was also supporting all the tests performed besides these full-tour experiments, with no failures.

3.2. Navigation

To assess the performance of the motion planner we did some repeatable experiments in a controlled area, obtaining some metrics to compare with other algorithms. The experiments evaluated the developed approach by comparing a classic local planner [2] with the same planner augmented with the learned costs and the costs based on Proxemics Theory [3]. The results shows that the augmented local planner with social costs improves the basic local planner and the planner based on Proxemics. This assessment is described in Deliverable 2.2, *Path planning and execution component for efficient and human-aware navigation*.

Now, we will describe the set of experiments that were performed in June and September of 2014 in the Royal Alcázar of Seville, where the final demo of the FROG project took place. In this case, the last version of the navigation stack, including the learned social costs previously obtained, was commanding the robot.

In the experiments, the robot was performing the complete tour, navigating to all the Points of Interest and showing the proper content. In June, the guided tour had an average distance of 460 meters approximately, including the return path to the initial point, which involves a slightly different route. During 8 days of continuous testing, the tour was performed on average twice a day, which gives a total of almost 7,4 kms of autonomous navigation in a very crowded scenario. A summary of the data collected along these tours is presented in the Tables 3.2 and 3.3. The total duration of the tour in minutes, the total distance travelled by the robot (meters), and the average distance between the robot and the pedestrians (meters), are shown. Some data about the person distances is missing as the information from the person detection module was not recorded in some experiments (but was used along them).

Date	Total time (min)	Total distance (m)	Pedestrians mean dist (m)
06/19/14 15:02	24.70	223.3	3.42
06/20/14 09:14	25.30	260.0	---
06/20/14 17:52	21.92	258.23	2.65
06/25/14 09:51	22.88	231.16	---
06/26/14 11:53	26.60	257.91	2.09
06/26/14 17:35	27.73	275.46	2.25
Average	24.85 ± 2.19	251.01 ± 19.70	2.60 ± 0.59

Table 3.2. Metrics obtained in the experiments in June of 2014 without including the return route to the initial point.

Date	Total time (min)	Total distance (m)	Pedestrians mean dist (m)
06/21/14 12:40	38.03	460.22	2.23
06/22/14 16:52	37.39	453.4	2.16
06/23/14 11:55	37.08	468.25	2.3
06/23/14 16:46	35.93	453.77	2.24
06/24/14 12:49	36.02	465.44	---
06/25/14 12:03	38.00	489.03	---
06/25/14 18:08	36.51	458.4	2.36
06/26/14 11:08	36.02	434.97	2.2
06/27/14 14:40	37.54	470.03	---
Average	36.95 ± 0.85	461.50 ± 14.71	2.25 ± 0.07

Table 3.3. Metrics obtained in the experiments in June of 2014 including the return route to the initial point.

In September of 2014, another set of experiments were performed in the Royal Alcázar again. That time, the tour was slightly different from the tour performed in June, as new points of interest were included. The tour had an average distance of 500 meters. The person tracking with lasers was also improved and integrated (see Section 4), and we have included a new data column expressing the total number of pedestrians detected by the lasers-based tracker. These values are higher than the real number of pedestrians due to some false positives and persons re-entering in the detection area.

In Table 3.4, the data of the navigation tests and the tests in which the tour was completed or nearly completed are shown. The differences in total time and distance traveled between some tests are due to different conditions in terms of persons densities or slight modifications of the route.

Date	Total time (min)	Total distance (m)	Pedestrians mean dist (m)	Pedestrians detected
09/18/14 10:22	40.09	440.22	2.89	1881
09/22/14 11:20	53.49	558.74	2.52	3599
09/22/14 18:41	59.22	685.33	2.51	2942
09/23/14 10:25	53.57	540.96	2.47	3108
09/23/14 15:18	46.59	492.84	2.37	2246
09/23/14 17:33	39.59	461.94	2.37	2202
09/24/14 09:49	57.02	456.52	3.01	3283
09/24/14 16:57	43.96	510.28	2.53	2799
09/25/14 10:27	40.25	390.46	2.46	2628
09/25/14 16:41	49.63	592.93	2.62	3104
09/26/14 11:00	35.98	354.47	2.49	1464
09/26/14 18:47	56.16	520.76	2.05	3918
Average	47.96 ± 7.94	500.45 ± 89.84	2.52 ± 0.24	2764.50

Table 3.4. Metrics obtained in the experiments in the Royal Alcázar in September of 2014.

Regarding the robustness of the navigation stack, only in 1 case of the 26 missions performed a manual intervention was required to recover the robot from a blocking situation (see Section 7 for an explanation).

3.2.1. Docking Evaluation

The docking procedure was presented in Deliverable 2.2. Figure 2.3 shows the docking station and part of the undocking maneouver. In the undocking phase, FROG moves out from the charging station, and navigates towards the first Point of Interest of the tour, to wait for interested visitors. For the docking operation, FROG is commanded a final approach waypoint, and then the docking behavior is actived.

During the final experiments, in 25 out of the 26 full tests performed, the robot undocked automatically from its charging station at the beginning of the tour (and in the remaining experiment the charging station was not deployed). Furthermore, in 11 of those 25 experiments, after the end of the tour the robot automatically docked into the charging station (42.3% of tests were thus correctly done with complete undocking-docking procedure). In 7 of the remaining 14, the docking procedure was not tested (either because the docking place was not ready - 2 cases- or because the robot was commanded to the lab. space at the Alcázar for mission debriefing). Finally, in 2 cases the docking did not take place because a wrong final approach was commanded from the high-level state machine and in 5 occasions the full mission was affected by hardware issues before the end, preventing to perform the docking manoeuvre.

4. Person detection and localization

The person detection and localization module based on stereo vision has been evaluated and the results reported in D3.1. This module has been integrated and run during the experiments performed in the Royal Alcázar.

Furthermore, this module has been complemented with a person detection and tracking module employing laser data (described in D2.2), to increase the field of view of the person estimation modules, following the reviewers' comments in the second year review.

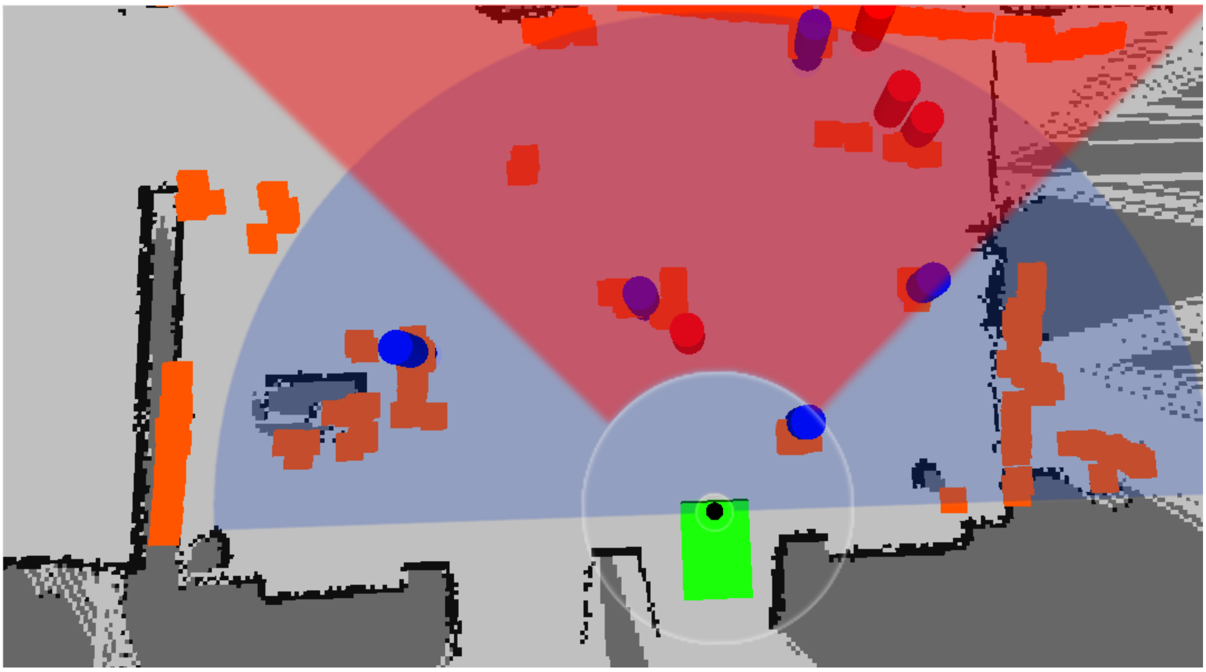

Fig. 4.1. People detected by the stereo-based module (red) and laser-based module (blue). Regions of detection are shown in the respective colour.

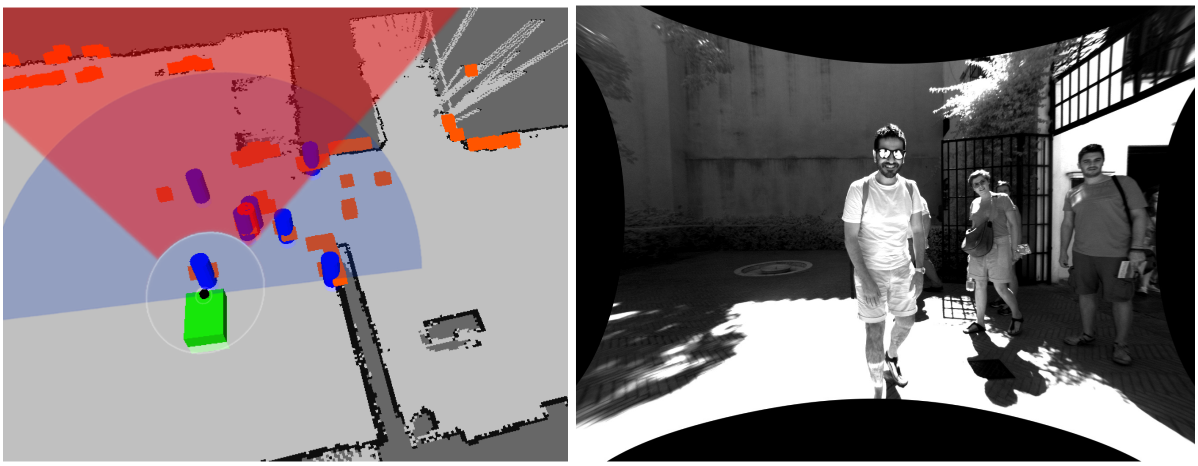

Fig. 4.2. Left: People detected by the stereo-based module (red) and laser-based module (blue), regions of detection are shown in the respective colour. Right: Camera capture.

Figures 4.1 and 4.2 show two examples of the outcome of both algorithms. The camera-based method can detect persons between 2 and 25 meters, and its horizontal field of view is approximately 90º. On the other hand, the laser-based method can detect persons between 50 cm. and 10 m. and has a 180º field of view (each laser).

The laser-based method typically detects more persons than the camera-based method, given its larger field of view; but also it presents a higher false-detection rate. On the other hand, the camera-based algorithm has an extremely low false positive rate. Combining both algorithms allows us to have a better person detector, due to the extremely low level of false positives in the camera-based detector and the important number of detections of laser method with a low level of false positive and false negative.

The following table summarizes the mean number of pedestrians detected for every message received from the different modules (the message includes all the persons tracked at that particular time) for a subset of the experiments in June. As commented, the laser-based detector has a higher index of detections.

Date	Duration (min)	Distance (m)	Mean pedestrian detected by laser*	Mean pedestrian detected by camera*
06/21/14 12:40	38.03	460.22	6.077/msg	3.49/msg
06/22/14 16:52	37.39	453.4	4.108/msg	3.31/msg
06/23/14 11:55	37.08	468.25	5.54/msg	**
06/23/14 16:46	35.93	453.77	5.42/msg	3.51/msg
06/24/14 12:49	36.02	465.44	5.42/msg	5.102/msg
06/25/14 12:03	38.00	489.03	5.64/msg	3.94/msg
06/25/14 18:08	36.51	458.4	3.093/msg	3.57/msg
06/26/14 11:08	36.02	434.97	5.39/msg	2.98/msg
06/27/14 14:40	37.54	470.03	2.92/msg	**

Table 4.1. *Mean values by sensor measurement (persons tracked per message). **The camera-based detector was not running in these experiments.

The laser-based detector works at the rate of 10Hz, while the camera-based detector is able to work at the same rate as the stereo camera pair: 5Hz (with a maximum mean effective rate of 3 messages per second for the experiments analysed above).

As an example, Figures 4.3 and 4.4 show the number of detections of both detectors by time for data collected on 06/24/14 at 12:49.

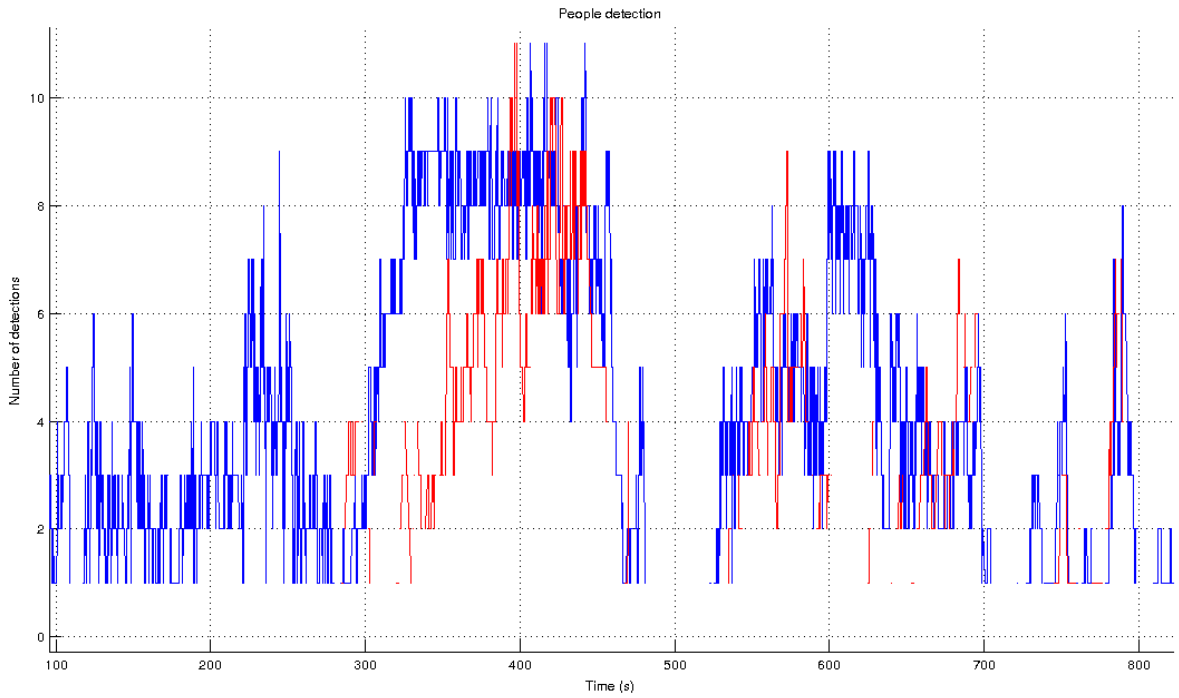

Fig. 4.3. Data collected on 06/24/14 12:49. Blue: Laser detections by time. Red: Stereo pair camera detections by time.

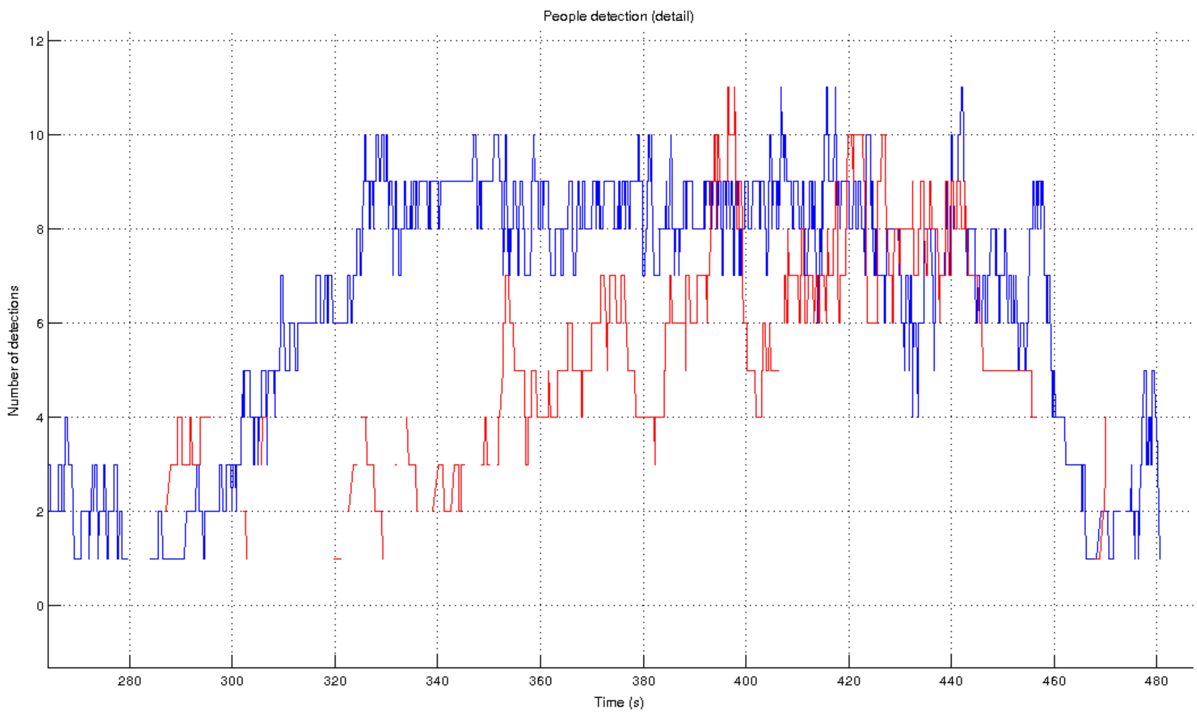

Fig. 4.4. More detailed comparison for data collected on 06/24/14 12:49. Blue: Laser detections by time. Red: Stereo pair camera detections by time.

Consistent people tracking

The two following figures (Fig. 4.5 and 4.6) show the stability and consistency of the improved laser-based people detector. Fig. 4.5 shows the 'Person ID' detected at each time while Fig. 4.6 shows with more detail how tracked people maintain their own ID while being in the range of detection of the algorithm.

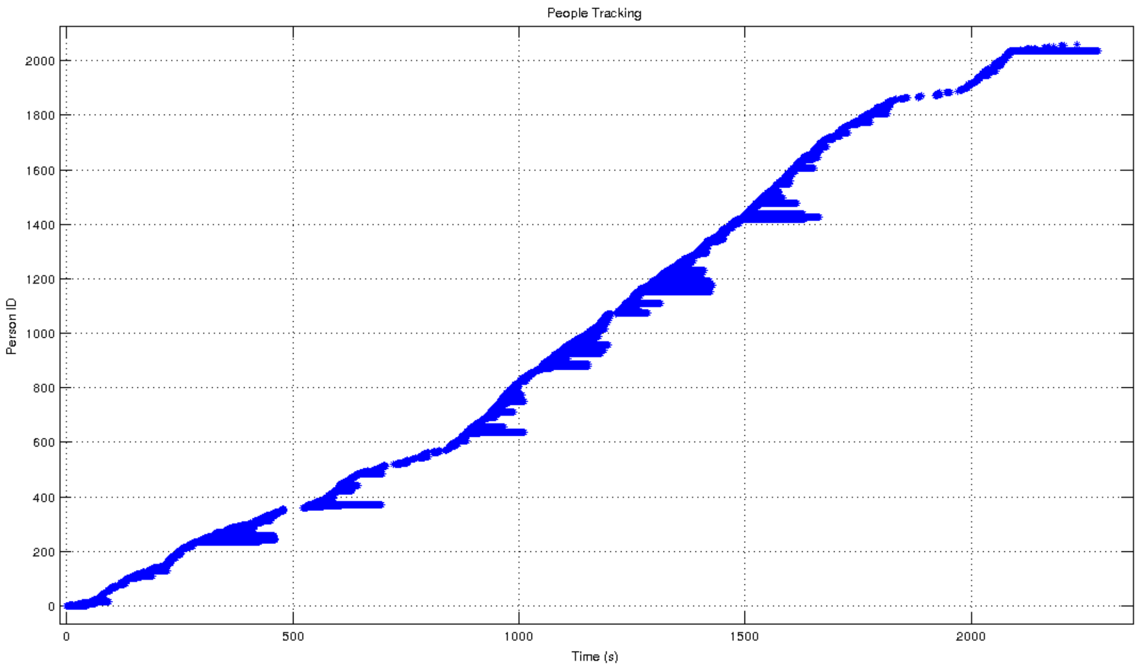

Fig. 4.5. People ID tracked by time for data collected on 06/24/14 12:49.

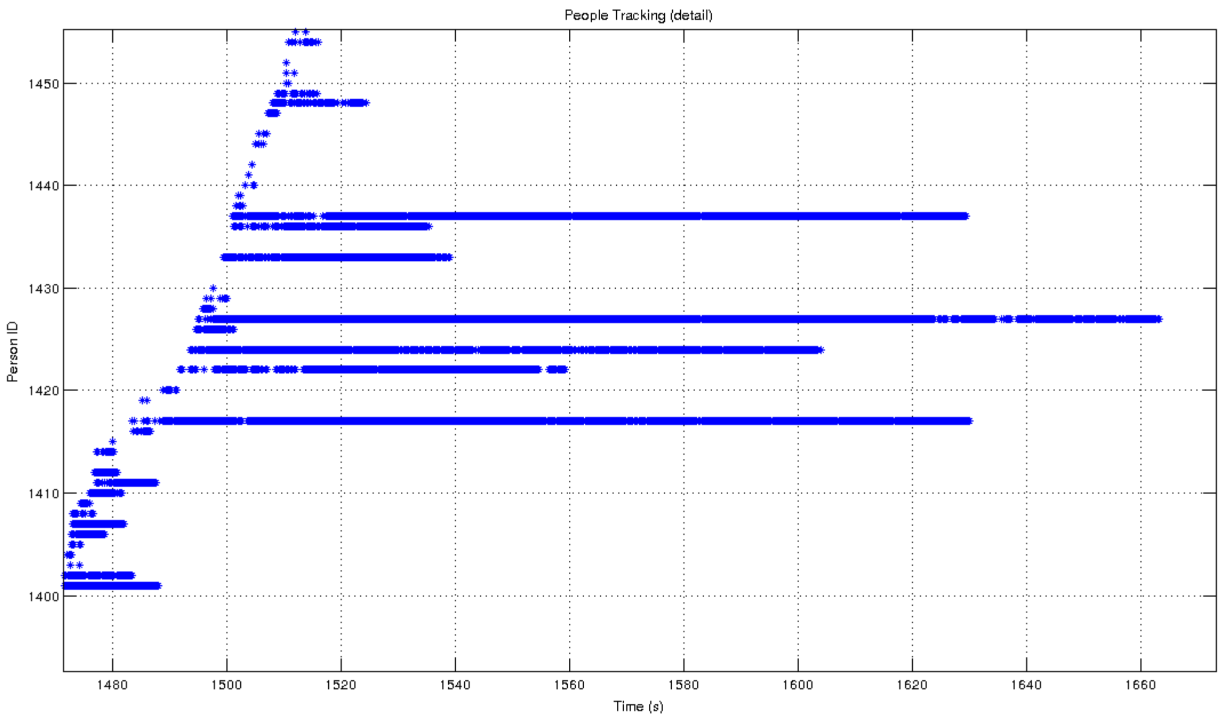

Fig. 4.6. More detailed People ID tracked by time for data collected on 06/24/14 12:49.

5. Analysis of human affective signals

The human affective signal analysis module is further divided into three main components: the facial landmark tracker, the interest level predictor, and the face verification component. A nod / shake and agreement / disagreement detection component has also been implemented, as according to the original design. Nonetheless, this component was later removed from the final version of the module because its output was not utilised by subsequent modules within the FROG system.

The facial landmark tracker serves as the basis for the other two components. It outputs an accurately tracked 3D face mesh of the tourist interacting with FROG. The registered and normalized 3D locations of the facial landmarks are then concatenated into the feature vector for the interest level predictor and the face verification component. The interest level predictor uses a support vector machine cascade trained over annotated FROG data to recognise whether the tourist is not interested, interested, or highly interested in the content displayed on FROG's screen. The face verification component is used to give each tourist a consistent ID throughout the tour. Due to the camera's limited field-of-view and occasional tracking failures, it is not uncommon for the facial landmark tracker to loose track of a tourist's face and later find the same face again during a single interactive session. Therefore, we developed the face verification component to check whether a newly tracked face belongs to a new tourist or a tourist tracked earlier. Specifically, the face verification component keeps a small database storing up to 10 face templates belonging to the tourists FROG has interacted with recently. Whenever a new face is tracked, the component attempts to align the new face to each of the stored templates using the alignment error to recognise which person the face belongs to or if the face belongs to a new tourist when the minimum alignment error exceeds a predefined margin.

The face-tracking algorithm has been tested on the Multi-PIE database, the XM2VTS database, and the FROG dataset consisting of 15 video clips recorded during the tours. All test results are described in detail in section 3.3 of the D3.1 documentation. The results on the Multi-PIE and XM2VTS databases show that our proposed DRMF method outperformed other state-of-the-art face tracking algorithms in terms of accuracy measured by shape RMS error.

Results on the FROG data show our method performed satisfactorily in most cases. Nonetheless, tracking failures were also observed when our component was operating under extremely challenging illumination conditions, such as when the camera is pointing toward the direction of the sun. Related POI locations were then refined to eliminate these difficult cases.

The interest level prediction algorithm has also been tested on the aforementioned FROG dataset (with human-annotated per-frame interest level values as ground truth). The test results can be found in section 3.6 and chapter 4 of the D3.2 documentation.

The face verification algorithm has been tested on the FRGC database, Multi-PIE database, FERA database and the aforementioned FROG datasets. High recognition rates (>80%) have been achieved in all experiments. Detailed description and discussion of the test results can be found in subsection 2.4.2 of D4.2 documentation.

Though not running in real-time during the FROG experiments in June and September, the proposed method for nod / shake and agreement / disagreement detection has been tested on the Canal 9 Database of Political Debates dataset and the FROG dataset. All results are presented in section 3.3 and subsection 3.4.1 of the D4.2 documentation.

The human affective signal analysis module has been deployed on FROG and has been running continuously during the June and September, 2014, experiments. We now provide a summary of our findings gathered from these two experiments and how we have improved the module based on these findings.

During the experiments, the human affective signal analysis module performed satisfactorily at the three locations (POI0, POI1, and POI2) where it was used. As an example, Fig. 5.1 shows the results obtained from the tour taken on 09/23/2014. Note that thanks to the face verification procedure, the first tourist was able to keep the same ID (face 4) at the first two locations (he did not follow the tour to POI2).

Fig. 5.1. Face analysis results obtained during the tour performed on 09/23/2014.The images, from left to right, were captured when the FROG robot was at POI0, POI1, and POI2, respectively. An interest level of 1.0 means 'highly interested' and 0.5 means 'interested'.

Not immediately apparent from the figure, one earlier engineering challenge we faced was the large illumination condition variation at different locations and at different times of the day. We tried a simple centre-weighted auto-exposure algorithm but failed to obtain good results. This is because the tourist's face usually occupies only a small portion of the scene and the face region is not necessarily in the centre of the image. Hence, we later customised the auto-exposure algorithm to focus on the face region as given by the output of the facial landmark tracker and only reverse to use global brightness estimator when no face is visible in the scene. Because of this, it can be observed from Fig. 5.1 that the background regions were either under-exposed or over-exposed but the brightness of face regions were kept constant.

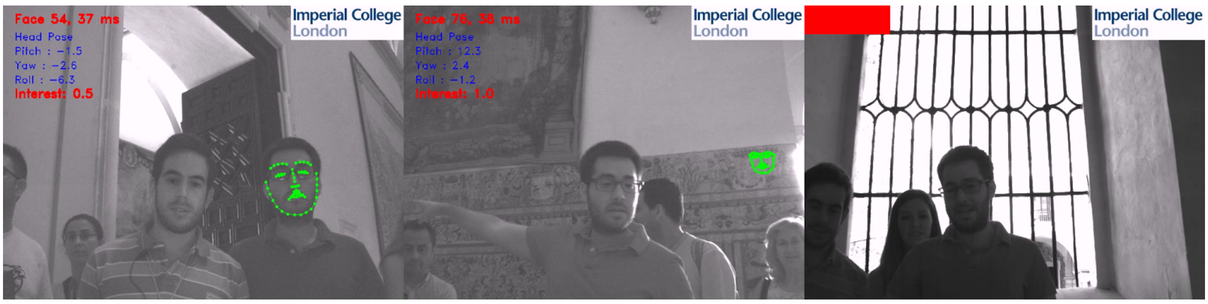

Fig. 5.2. Facial landmark tracking failures occurred during the tour performed on 09/23/2014.The images, from left to right, were captured when the FROG robot was at the POI4, POI5, and POI7, respectively.

Albeit the improved camera control method, the facial landmark tracker could still fail in extremely difficult environments, as shown in Fig. 5.2. Because POI4 and POI5 are located inside poorly-illuminated halls, the camera was forced to use a long exposure time (up to 30ms) and a large gain value (up to 10dB), causing the captured video to become blurry (because of motion blur) and noisy, thus decreasing the accuracy of our algorithm. For instance, due to the poor image quality, the tracked face mesh around the mouth region was incorrect in the left frame and the face detector produced a false-positive in the middle frame. At POI7 (the right frame), because of the large contrast between the faces and the bright background (the window), face detection failed. Note that the module's output was not utilised at these POIs because of the content design. Hence the errors shown here had no impact on the system's perceived performance. For future applications, a better image acquisition device may be used to capture higher quality video data in these hard cases for accurate facial analysis.

During the experiments, we also discovered an issue with the interest level predictor. Due to the nature of the machine learning techniques we used, the accuracy of the interest level predictor largely depends on the quality of the annotated training data. The resulted predictor would be unable to produce accurate output if the training examples were unrepresentative and / or correctly labelled. This is why our module sometimes reported 'no interest' (0) while the tourist was actually interested in the tour, as in the cases depicted in Fig. 5.3.

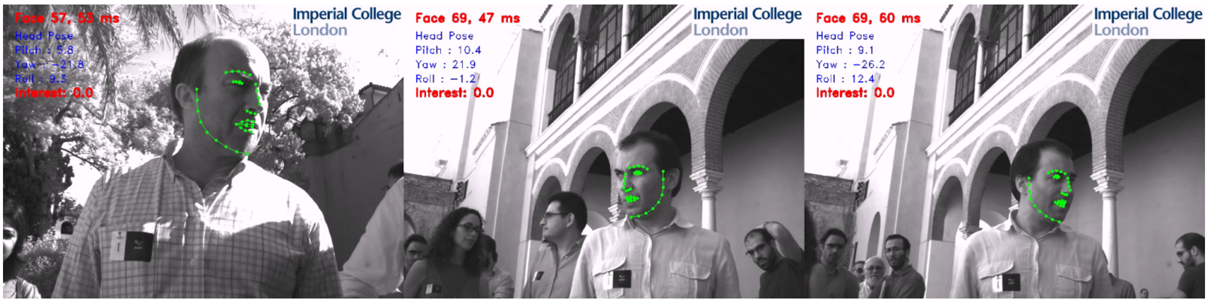

Fig. 5.3. Interest level prediction error occurred during the tour performed on 09/25/2014.The images, from left to right, were captured when the FROG robot was at the POI1, POI2, and POI2, respectively.

Our interest level predictor was trained on a dataset captured before the FROG content was finalised. At that time, the assumption was that FROG would mostly display content on its screen. Hence the tourist would be considered interested in the tour if he / she is interested in the content displayed on the screen and vice versa. However, the current version of FROG also uses its snorkel occasionally to point at different objects during the tour. In these cases, the tourist's attention would be attracted by the the object pointed at. The predictor would then report a 0 interest level because the tourist is not looking at the screen. However, this result would be incorrect because the interest level toward the displayed content is no longer equivalent to the interest level toward the tour. Realising this, we advised the FROG state machine to disregard the interest level prediction result when the snorkel is pointing at an object of interest. Since video information displayed through FROG's screen still consists of a significant part of FROG's content, the tourist's interest level recognised during the video playback sessions could be considered a reliable estimation of his / her interest level toward the tour. Hence in most cases the FROG state machine was able to make correct decisions based on this estimation. Nevertheless, simply retraining the predictor using a newer and more representative dataset would very likely further improve the predictor's accuracy thus improve FROG's overall performance.

6. FROG State Machine

In this section we will present the two different versions of the State Machine and the Front End of FROG. These versions were developed between June and September 2014. They were prepared using different simulators for navigation and communication with the RosBridge. During two integration sessions the State Machine and Front End were tested in real-life.

6.1 June 2014

The first version of the FROG State Machine was developed for the demo in June 2014. This version of the State Machine was developed using two simulators provided by UPO. One simulator was simulating recorded sensor data of the FROG, such as position and detected persons. This simulator was also used to test the communication via the RosBridge. The second simulator was simulating the navigation of the FROG. Navigation commands were executed and the navigation responses were published via the RosBridge. A complete tour through the Royal Alcázar could be simulated using the navigation simulator and the first version of the Front End. A simulated tour consists of navigating along the different POIs and starting the content at the POIs.

The first version of the State Machine was tested in June 2014 at the Royal Alcázar. FROG gave 16 complete and autonomous tours during two weeks. During these two weeks the State Machine was updated and small bugs were fixed. The updates included the docking and undocking of the FROG to/from its docking station and the integration of the data from the ICL human affective signals analyser at POI2.

6.2 September 2014

The two simulators were used to prepare the second version of the FROG tour and the State Machine. Point of Interest 3 and Point of Interest END moved to other locations and the camera in the pointer armplayed a more prominent role during the story told at the POIs. The content of the tour was updated to support three different languages (English, Spanish and Dutch). Data from the ICL human affective signals analyser will now be used during the introduction of the FROG and at POIs 1 and 2 and the volume of the audio is based on the location of the FROG and more examples of augmented reality were added at the different POIs. The updated State Machine and Front End was tested in Twente using the navigation simulator and the Campus Robot.

During the integration meeting in September 2014 the new State Machine and Front End were tested at the Royal Alcázar. Ten complete and autonomous tours were given during two weeks (see Section 7.1 for more details about the tours). Four new function waypoints were integrated for navigation purposes. During these two weeks the new /PeopleUPO topic for laser based person detection was integrated into the State Machine. Information obtained from this topic is used at several moments during the tour. It is used to look at a person or people with the pointer arm camera, to detect whether FROG has lost its visitors (or not) when FROG arrives at one of the POIs and to approach potential candidates for a guided tour at POI0. The integration of the new topic into the State Machine and the approaching of candidates for a tour were tested using a third simulator provided by UPO. The content of the tour was updated with the support of subtitles for the voice-over of the guide because the quality of the audio recording is not always good enough. The tested projection of augmented reality elements (using the simulator) had to be fine-tuned in real life at the different POIs for better precision. Two different tours for the press event were developed during the meeting. One indoor (POI 4 and POI 5) for rainy weather and one longer version (POI 2, POI 4 and POI 5) for dry weather.

The State Machine and Front End can be copied one-to-one to the real robot after development on the simulator. When all the software is copied to the FROG both the State Machine and the Front End could be executed on the FROG interaction PC. The IP address of the RosBridge should be changed to the local FROG network and local path to the controllers (eyes, pointing arm and projector), the Front End and the server of the web interface should be adapted. Starting the State Machine will start all the controllers and the Front End before starting the tour.

User evaluation of the tours was studied in June, 2014, with the goal gaining insight into the user experience of a guided tour with the FROG robot in the Royal Alcázar. Details about the number of tours, participants and the results of these user evaluations studies will be presented in Section 7.2.

7. Overall FROG system

In this section, the evaluation of the full FROG robot system, integrating all the functionalities is presented. Results from the June and September 2014 deployments are described.

The experiments consisted in performing the full FROG tour, which is described in Fig. 7.1. The tour consists of 7 Points of Interest (POIs). There were some slight variations on positions of the POIs between June and September, but the main structure of the tour is the same.

Figure 7.1. Map of the route of the FROG tour

Before beginning the tour, the robot has to undock from its charging station. After the last POI, the robot may go back to the docking station to charge itself.

First of all, we will describe a numerical evaluation of the missions performed, indicating the number of completed missions and the issues if any for the particular mission. Then, the user evaluation will be presented.

7.1. Overall evaluation of robot deployment

The next tables summarize the missions performed in the Royal Alcázar. All the missions began on the robot docking station. In many of them, after the tour was completed, instead of moving back to the charging station, the robot was moved to the lab. space in the Alcázar to spend the night or to facilitate the analysis of the logged data.

Some data and remarks collected in the experiments performed in June of 2014 in the Royal Alcázar are presented in Table 7.1. In this session, the code was frozen on Thursday 19, and experiments were conducted until Friday 27.

Date	Total time (min)	Total distanc e (m)	Localization ok?	Navigation ok?	Docking and undocking?	Hw issue	Remarks
06/19/14 15:02	24.70	223.3	Yes	Yes	Undocking	No	Tour completed but the robot didn't go back to the charging point. Docking place not ready.
06/20/14 09:14	25.30	260.0	Yes	Yes	Undocking	No	Tour completed but the robot didn't go back to the charging point. Docking place not ready.
06/20/14 17:52	21.92	258.23	Yes	Yes	Undocking	Yes	Tour completed. Going back to the final charging point, the frontal laser failed.
06/21/14 12:40	38.03	460.22	Yes	Yes	Undocking	No	Tour completed. The docking maneuver was not performed due to a wrong approaching final position.
06/22/14 16:52	37.39	453.4	Yes	Yes	Undocking	No	Tour completed. The docking maneuver was not performed due to a wrong approaching final position.
06/23/14 11:55	37.08	468.25	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
06/23/14 16:46	35.93	453.77	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
06/24/14 12:49	35.01	465.44	Yes	Yes	Undocking and	No	Tour completed, including returning to the

					docking		charging station. No problems.
06/24/14 17:01	---	----	Yes	Yes	Undocking and docking	No	No problems, but sensors and odometry data weren't recorded.
06/25/14 09:51	22.88	231.16	Yes	Yes	Undocking	Yes	The IMU sensor failed and the tour had to be aborted.
06/25/14 12:03	38.00	489.03	Yes	Yes	Undocking	Yes	Tour completed. Going back to the final charging point the IMU sensor failed.
06/25/14 18:08	36.51	458.4	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
06/26/14 11:08	36.02	434.97	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
06/26/14 11:53	26.60	257.91	Yes	Yes	Undocking	No	Tour completed successfully without returning to charging station.
06/26/14 17:35	27.73	275.46	Yes	No	Undocking	No	Tour completed with manual intervention without returning to charging station.
06/27/14 14:40	37.54	470.03	Yes	Yes	Undocking and docking	No	Tour completed. No problems.

Table 7.1. Evaluation data of the experiments performed in June 2014 in the Royal Alcázar of Seville.

In the table above, a failure in the localization module is declared if manual intervention is required due to wrong position estimates. Equivalently, a failure in the navigation component is declared if manual intervention is required to prevent wrong navigation behaviors (like the robot getting stuck).

As can be seen in Fig. 7.2 and Table 7.1, the single time we experimented a problem with the navigation system was the day 06/26/14 at 17:35, when the main patio was occupied with tables and tablecloths for a special gala dinner at the Royal Alcázar. Normally this patio is free and there is no permanent obstacles. The robot was able to negotiate the (previously unmapped) obstacles of the place, even though the remaining space was quite narrow, reaching the corresponding POI in the square. When leaving the tables zone, the robot was avoiding a table, and, due to the wind, the table clothes began flapping, coming into the robot safety buffer zone, and even touching the robot. The robot thus, stopped, and the wind kept the clothes coming in and out of the safety zone. A short manual intervention was required to bring the robot out of the situation. The rest of the tour was completed autonomously.

Fig. 7.2. The main patio was plenty of tables for a special gala at the Royal Alcázar.

The missions performed in September are presented in the Table 7.2. Again, the first week was devoted to integrate and test the new configuration of the tour. Then, the system was fixed on Friday the 19th. Due to weather conditions (with heavy rain during the 20th and 21th and sporadic rain along the whole week), less missions were performed this time. The docking system worked properly all the time. The main problems were related to the inertial sensor (IMU), which was finally replaced on the 25th.

Date	Total time (min)	Total distance (m)	Localization ok?	Navigation ok?	Docking and undocking?	Hw issue	Remarks
09/18/14 10:22	40.09	440.22	Yes	Yes	Neither	No	The tour test started and finished in different points.
09/22/14 11:20	53.49	558.74	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
09/22/14 18:41	59.22	685.33	Yes	Yes	Undocking	Yes	The IMU sensor failed and the tour had to be aborted.
09/23/14 10:25	53.57	540.96	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
09/23/14 17:33	39.59	461.94	Yes	Yes	Undocking and docking	No	Tour completed, including returning to the charging station. No problems.
09/24/14	43.96	510.28	Yes	Yes	Undocking	Yes	Powering problems with

16:57							lasers and IMU.
09/25/14 10:27	40.25	390.46	Yes	Yes	Undocking	No	Tour completed. No problems.
09/25/14 16:41	49.63	592.93	Yes	Yes	Undocking and docking	No	The robot was teleoperated in the last part of the tour due to low batteries status.
09/26/14 11:00	35.98	354.47	Yes	Yes	Undocking	No	Tour completed. No problems.
09/26/14 18:47	56.16	520.76	Yes	Yes	Undocking	No	European Researchers' Night. Tour completed. No problems.

Table 7.2. Evaluation data of the experiments performed in September 2014 in the Royal Alcázar of Seville.

In all these experiments, the person detection and affective computing modules were running and integrated, and the State Machine was coordinating the execution of all the modules.

To summarize, the results obtained in the experiments performed in June and September were satisfactory. The localization and navigation systems seem to be robust enough to guarantee a proper performance. The docking system was integrated successfully and the robot was able to start and finish a tour autonomously. The person detection and human affective signals provided social feedback for FROG, and the State Machine integrated all the modules with the interactive content to perform the tours. The problems encountered were mainly due to hardware issues, like the malfunction of the IMU sensor. The percentages of success obtained in June and September are presented in the Table 7.3.

Number of missions	% successful tours
26	85

Table 7.3. Percentages of successful missions performed in the June and September experiments.

7.2. End-user evaluations

7.2.1. User Experience Evaluation with FROG robot

Technical innovations are very important for the development of a robotic outdoor tour guide. However, at the end of the project, the robot has to be able to guide visitors through a museum and the FROG's success will depend on its abilities. Therefore, besides developing and applying state-of-the-art technology for this robot, we aim to create the best visitor experience possible. To gain an understanding into how people experience a robot guided tour, we observed and interviewed visitors to the Royal Alcázar who followed (parts of) a tour given by the FROG robot. Evaluations of user experience with the FROG robot were performed in the last week of June 2014. The FROG robot gave 16 autonomous tours in seven days, four of which were with recruited visitors and the other 12 were given to naïve visitors of the Royal Alcázar. Our study focused on the four tours with recruited visitors. People participating in these tours evaluated the user experience of the robot throughout the tour. However, we also evaluated the tours of non-recruited visitors. For these tours, observations were made during the tours and interviews after the tour had ended. People that left the tour before it had ended were also interviewed and additional results of these interviews gave important insight into the shortcomings of the robot-guided tour. We found that participants especially liked the interactivity of the tour, many of the participants stated that they would like to see more interactivity in the tour. Especially the activities at the first Point of Interest could be changed, because that one was very static and not representative for the rest of the tour. Participants did not really like the repetitive utterances of the robot during the transition from one Point of Interest to the next. Some of the participants suggested the use of sound or more different sentences during the transitions. Also, during transitions participants missed some feedback about where the robot was going or where it wanted to stop. Most participants thought the robot was a bit slow and still moved a bit jerkily. And a lot of participants mentioned that the robot tours would be great for children and young people, who want to know something about the site, but do not want to follow a tour guide. Also some participants stated that the robot looked as if it was designed for children, so they were unsure whether they were supposed to follow it.

7.2.2. Goal of the Study

This user evaluation study was performed to get insight into the user experience of a guided tour with the FROG robot in the Royal Alcázar in Seville, Spain.

7.2.3. How was the Study performed?

To gain insight into how people experienced the robotic tour guide, visitors of the Royal Alcázar who followed the robot for a complete tour or for just a few stops were observed during the tour and interviewed afterwards about their behaviours and experiences with the robotic tour guide.

During the week that the FROG robot was active in the Royal Alcázar, about 16 tours were given to both (groups of) naïve visitors as well as recruited visitors. Naïve visitors were people who visited the Royal Alcázar by chance and who joined one of the FROG tours spontaneously. In this deliverable we will report responses by visitors who joined a tour, and participated in a short interview. As well as these naïve visitors, we recruited people to follow the whole tour and participate in a longer interview. In this report, we will refer to these people as recruited participants. A total of eight participants were recruited in four groups; two groups comprised of two people (a Dutch man and a Dutch woman), one group comprised of one Spanish woman (with a baby) and one group comprised of three Spanish students. In this study, we will focus on the evaluation of the tours with the recruited participants, since these gave us most information about the user experience during a tour. However, in some cases, the information obtained from our observations or from the short interviews with naïve visitors gave us additional information, especially about the reasons why people left the robot before the tour was ended.

The robot gave two or three tours a day and always started in the Lion's Courtyard. Here the robot waited to pick up visitors. These could be the recruited participants or naïve visitors. Also when FROG gave the tour, visitors could join or leave the robot whenever they wanted. The tours were given autonomously by FROG. However, the researchers followed the tour from a distance to check the log files and the like. Some of the participants were not aware of the fact that the robot was totally autonomous, because one of the researchers walked with the robot (at a distance of at least 10 meters) with a remote control just to stop the robot in any case of emergency.

The recruited participants were asked to think aloud during the tour. In each group, one of the recruited participants wore a small microphone, which was able to pick up the speech of the whole group if they were close to each other. None of the participants had been in the Royal Alcázar before and each of them expressed liking to hear the information given by the robot. Therefore, the researcher kept her distance while the robot was giving information about an item and only joined the participants when the robot and the group were moving to a next item. During these walks, the researcher asked questions to gather first reactions from the group. The microphone that one of the participants wore was recording at all times, which allowed us to include remarks made between people in the group in the analyses as well. The recruited visitors were asked to follow the whole tour, even if there would be a point where they would like to leave.

At the end of the tour, a semi-structured interview was held with the whole group of recruited visitors, so they would be able to complement each other's remarks. In each interview several topics were addressed, such as their experience of joining the robot tour, the things they liked best about the robot or the tour, what they would improve in the robot or the tour, and how they experienced the behaviour of the robot, the length of the stops and the entire tour, the information given, the use of the eyes, arm pointer, screen and projector, the way the robot moved to the next point, how they felt about sharing the robot with strangers, why they were talking to the robot, and if they understood the meaning of the face tracking. As the interview was semi-structured, there was no specific order of the topics and participants were able to focus a lot on what they thought was important. The interviews were held in Dutch or English.

During the 24/7 period that the robot was active in the Royal Alcázar, small changes were made to the behaviour of the robot to improve the tour iteratively. Therefore, the experience of the tour differed slightly between participants. The iterative approach did not largely change the tour but some people did not see some features of the tour that were added at a later point in time. For example, a window with face recognition feedback was added to the robot. The subjects that visitors talked about during the interviews were somewhat influenced by this. However, the interviews were semi-structured to begin with and in all cases focused on those points that the visitors thought were most remarkable. The content, the length of the tour and the locations of the Points of Interest were the same for all visitors.

The recruited participants had either no or little experience with robots. The three Spanish students were from a computer science course, and had some background knowledge about the navigation of the FROG robot. However, their experience with real robots was limited. Naïve visitors who participated in the interview had various experiences with robots. Some of them were technicians or robotics lecturers, others had no previous experience with robots. Most participants did have previous experience with human guides or audio guides and could therefore compare the tour given by the FROG robot to some extent.

7.2.4. What did a Tour look like?

A tour given by the FROG robot started in the Lion's Courtyard (see the end of this chapter for a map and Fig. 7.3), close to the visitors' entry. The robot stood at a certain place and uttered with a mechanical, synthesized 'robot voice' "looking for interested visitors" and once a visitor was within a range of 2.5 meters, the robot said "do you want to take a tour with me? Please touch yes or no on the screen." When visitors chose no, the robot started to look for interested visitors again. When the visitors pushed yes, the robot started the tour. The robot started with an introduction, consisting of a movie on the screen and a voice-over, about what visitors could expect. Subsequently, it invited the people to follow the tour.

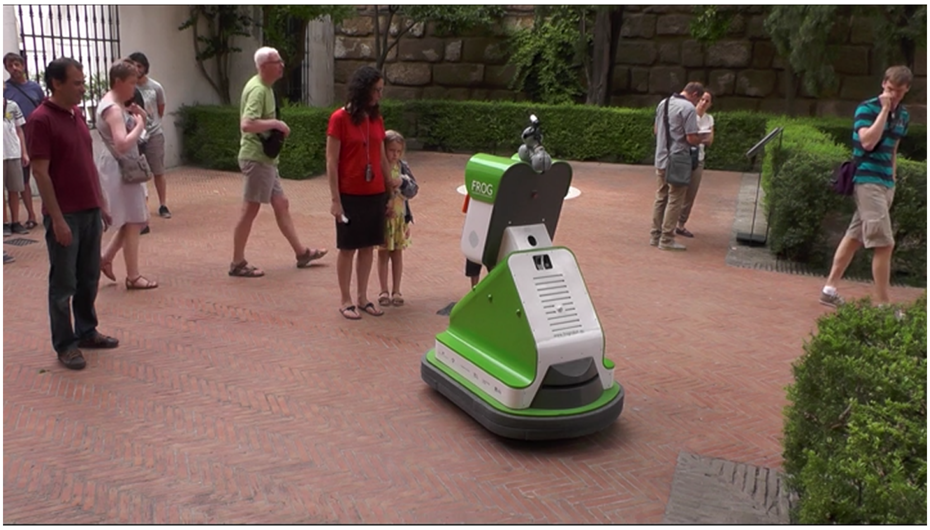

Figure 7.3. Girl who is a bit frightened of the robot

The tour started when the robot drove to the first point, still in the Lion's Courtyard, close to the triple arch wall and next to the Hall of Justice (Fig. 7.4). Upon arrival, the robot gave some information about the development of the courtyard over the years. For this, mainly the screen (showing pictures and movie) and the pointer were used.

Figure 7.4. Robot pointing to Hall of Justice and visitors looking in the direction indicated

When the robot was finished at the first Point of Interest, it drove to the next point. While moving, the robot repeated in its 'robot voice' "please follow me" and "driving to next location". The visitors who were following the robot were expected to walk next to or behind the robot (Fig 7.5.). On the screen of the robot a small map, a smile and the text the robot uttered were visible to inform visitors that encountered the robot from the opposite direction. (This was the same every time the robot traveled to the next Point of Interest).

Figure 7.5. Following the robot to the first Point of Interest

The second stop for the robot was in the Hunters' Courtyard, the main square in the Royal Alcázar. Here, based on the time, the robot chose to which side it would stand to find shadow. At this point, the robot turned towards the visitors and started the story about the buildings that are visible from the main square (Fig. 7.6). During the story some information was given on the screen, and by using the arm pointer, the robot pointed to several Points of Interest. At this Point of Interest the robot extended the story with a second part when it rated the visitors as interested. For this, it used the face recognition data of Imperial College London. When visitors were rated as 'not interested', the robot finished at this Point of Interest and proceeded immediately to the next Point of Interest.

Figure 7.6. The robot telling the story at the main square

The transition to the third Point of Interest (Fig. 7.7) was quite long (about 3.5 minutes just following the robot) and moved through a narrow hallway and another courtyard. At the third Point of Interest, the robot used the projector to project the information about the Lisbon earthquake and its implications for the buildings in this area, on the wall. As the wall was yellow and there was a bench in front of the wall, the projection was not easy to see. Also, due to technical problems, the screen showed the same pictures but rotated 90 degrees clockwise. At this point the robot also used its arm pointer.

Figure 7.7. Ready to start at third Point of Interest

After this Point of Interest, the robot went inside. In the first hall in the Gothic Palace, the Tapestry Room, the robot gave information about the tapestries and focused on one tapestry in particular (Fig. 7.8); the tapestry with the map. As this map is made upside down according to our conventions, during the story, the robot added land names and flags on the screen using augmented reality and it flipped the map around on the screen. At this point there was a short conversation between the voice-over and the robot and also the arm pointer was used.

Figure 7.8. Discussing that the map is upside down

At the fifth Point of Interest, the robot gave information about the Vault Room; a hall with yellow tiles (fig. 7.9). After giving the information, the robot proposed a quiz to the visitors. It showed a picture of some tiles on the screen, with a part of it left out. The visitors had to identify the tiles in the room and to answer which part was missing. To help the visitors, the robot's arm pointer was indicating the direction in which people could find the tiles.

Figure 7.9. The quiz attracts a lot of people

After the quiz, the robot proceeded to the last Point of Interest. To get there the robot had to go through a very narrow hallway. In the hallway the robot therefore asked the visitors if it could pass. Sometimes, passing the narrow hallway lead to problems, when people who wanted to pass in the opposite direction got interested in the robot and stopped, making it more difficult for the robot and the people following the robot to pass. However, in all cases the robot was able to reach the last Point of Interest, except for the last case, when the ramp was removed to clean the hall and FROG was not able to jump up the step.

At the last Point of Interest, the robot showed some of its fun capabilities, such as showing augmented reality using the camera in the pointer and showing fun content on the screen, and it gave the visitors the chance to take a picture before it drove away (Fig. 7.10). At this point it used the camera on the pointer to focus on faces of different visitors and used augmented reality to show the faces and something extra on the screen. After a while, the robot showed a last video in which a voice-over explained that there is much more to see at the site and that human tour guides would be able to give even more information. After this, the robot told the visitors that it had to recharge and would drive back to its docking station. During the travel it uttered "please do not follow me" and "I'm driving to my docking station."

Figure 7.10. Visitors taking pictures at the end of the tour

When the robot was recharged again, it could start a new tour.

7.2.5. Map of the tour through the Royal Alcázar

The tour took about 25 minutes, depending on the number of obstacles and the groups of people it would encounter in small hallways, and whether or not the story at the second Point of Interest was extended for interested visitors. The route the FROG drove is drawn in the map in figure 7.1, p34. The red line is the route FROG drove during the tour, with the stops where it gave a presentation are indicated in green. The orange line is the route FROG drove when it had ended the tour and drove back to the starting point.

1. Starting point in the Lion's Courtyard
2. First Point of Interest in the Lion's Courtyard
3. Point 3a; Second Point of Interest in the Hunters' Courtyard before 12.00 hrs.
   Point 3b; Second Point of Interest in the Hunters' Courtyard after 12.00 hrs.
4. Third Point of Interest in the Courtyard of the Cruise
5. Fourth Point of Interest inside the Gothic Palace
6. Fifth Point of Interest inside the Vault Room
7. Last Point of Interest and end of tour close to the entrance to the gardens

7.2.6. Data collection and analysis

7.2.6.1. Data collection

The data used for this user evaluation was collected during the FROG meeting in June 2014 in Seville. All tours were recorded with video cameras that were placed in spots where they captured the robot while giving information about the place. In the recordings the visitors are visible from their front and therefore most of the time the robot is visible with its back. However, due to sound it is clear what the robot is doing at these moments. When the robot drove, the camera was repositioned. This resulted in only some of the actions of the visitors following the robot being captured (right after the robot had finished the story and right before the robot started the new story), but some actions in the middle of a transition were not recorded.

One of each group of recruited participants wore a voice recorder, so that the remarks they made to each other about the robot and the tour were recorded for analysis. Also the interviews with the recruited participants were recorded using the voice recorder. The interviews with the participants that participated in the short interviews were recorded with the voice recorder as well.

After a tour with the recruited participants the researcher wrote down notes of the interview and some noteworthy observations. Also, during the short interviews the researcher took notes.

7.2.6.2. Data analysis

The data collection led to a variety of data sources. For this user experience report we have mainly used the interviews with the recruited participants and the short interviews. However, the field notes and a look back to a replay of the recordings were also used.

The interviews with the recruited participants and the voice recordings made during the tour were transcribed on a detailed level. The short interviews were also transcribed. These text files were coded and analyzed using the computer software Atlas.ti (http://atlasti.com/) [4]. This software helps to analyze qualitative data. Basically, all quotes coded with the same code were put together and the researcher read all quotes carefully, searching for commonalities and noteworthy experiences. These findings were combined in a story about that specific subject, presented as a paragraph in the result section. Also one or two quotes are presented in each paragraph to illustrate the visitors' experiences in their own words.

7.2.7. Results

7.2.7.1. User experiences of the tour

In general, participants and recruited participants liked to be guided by the robot because they found it an easy way to obtain information about the site. Four of the recruited participants said that you could get the same information the robot gave yourself from books. However, the experience of following the robot was much more fun to them than reading the books would have been. People were enthusiastic about being guided by the robot, and claimed that they would definitely tell others about it. Most of the participants mentioned that the robot could make a visit to a tourist site nicer and more interesting, especially for children.

“And we were flabbergasted, we thought that the man with the controller was controlling the robot…”

According to most participants the length of the tour was ok, it could even be a bit longer (for example, one extra stop). Also, people liked interacting with the robot, for example the way they did during the quiz, best. Some of the participants mentioned that the robot seemed to search for contact with the visitors by looking at them and using the arm pointer. Others also stated that they did not like the fact that the robot did not turn towards them after moving to a next Point of Interest, which could indicate that they expected more social behaviour from the robot. Participants stated that the behaviour of the robot was clear to them, however, in the beginning they had to get used to it.

Participants told us that the robot was like a magnet for people, it attracted so many people in the site, a lot for just taking a picture and some actually joined the tour. The recruited participants stated that they did not mind when others joined the tour, but that it could be a problem when the new people stood in front of them or talked too loudly while the robot was explaining about the Point of Interest. From our observations, we could see that visitors also seemed to want to touch the robot, and one recruited participant mentioned that if all visitors wanted to touch the robot, it could be possible that the robot would be damaged very soon.

“This is really nice, it should do this in more places, somewhere in between”

7.2.7.2. Appearance

“but it is just like StarWars, isn’t it? R2-D2 is that cute as well”

Participants were mainly positive about the appearance of the robot and used words such as friendly, nice and cute to describe the robot. Three of the recruited participants stated that the robot would do really well in a zoo, but that it fits less at the Royal Alcázar. However, the fact that the robot was designed to look like an animal or like machine was not a problem for them. They indicated that a robot designed like a person would be scary. Also, some participants mentioned that the robot looks as if it was designed for children, which made them unsure about following it. On the other hand, a few participants mentioned that the robot might frighten children, because its movements and the movements of the arm pointer are mechanical and jerky. One of the children we interviewed was scared in the beginning, because the robot was so big, but later on she enjoyed the tour, although she still found it rather exciting. Other children ran towards the robot, touched the robot and even seemed to want to climb on it.

7.2.7.3. Content

People liked the content that the robot provided, because it gave the basics of the history and architecture of the Royal Alcázar, something they would not get when going around the site without a guide or audio-guide. As the tour is not very long, participants stated that they did not have to concentrate that long while they definitely got to know the essentials of the site. Many participants stated that younger visitors would like these kind of tours better compared to tours given by tour guides or audio-guides. They also mentioned that it would be great for children, because the robot is more fun than a guide and keeps the attention of the children for at least some time.

“The things it said about the location, the history of Seville, of this site being special”

Of all the Points of Interest, participants most liked the Point of Interest where the quiz was held, because it was very interactive. According to the participants, that level of interactivity might also be repeated in other Points of Interest. Especially at the first Point of Interest this would be a good option because at that point, the information that was given was very static, which is not representative for the rest of the tour. Participants expressed that they liked seeing the pictures and movies that were shown on the robot's screen, rather than just listening to an (audio) guide. This also helped children and youngsters to keep their attention on the robot and the story. The participants mentioned that the change in and combination of modalities (screen, sound, arm pointer, projector) was fine and easy to understand. Some participants stated that they liked to find information in English that the robot provided, because otherwise there is little information in English available for free at the Royal Alcázar.

"It is the essence that he tells"

7.2.7.4. Robot voice/voice-over

"It seems to be a robot made for talking to, you know?"

Soon after the start of the tour the participants understood that the robot had a 'robot voice' that (repetitively) uttered what it did and a 'voice-over' that gave the information at each Point of Interest. Some participants stated that the 'robot voice' could be less repetitive, because it was the same all the time and became boring after a while. Some participants gave the option to use different sentences that have the same meaning or to use sound instead of text.

Even though the participants knew that the robot would not hear or understand them, they often talked to the robot. Remarkable was that participants seemed to have even more intention to talk to the robot, when it said "please follow me" or "location reached" or one of the other repetitive sentences using the 'robot voice'. Probably, this is caused by the fact that in these instances, the robot is more directly addressing the visitors or asking them to do something. The 'voice-over' function was clear to the participants. During the tours, participants did not talk or ask questions of the robot when the 'voice-over' was on. At these moments, people did make comments to one another about the content the robot gave. The fact that the robot would not understand people when they talked to it was not a problem. All people were happy to use the touch-screen for interaction.

One thing that confused the participants a bit was that there was a third person in the robot that gave the introduction and the closing presentation. Participants stated that they preferred the 'robot voice' and the 'voice-over' and not the 'robot voice' and two 'voice-overs'. It would be better to have the introduction and closing presentations with either the 'robot voice' or with the 'voice-over,' but participants were not unanimous about which type of voice to use.

"The only problem I saw is that the message of "please follow me" is so repetitive"

"Maybe a bit more, and instead of saying "Loading information data" all the time, put it… put something different on that word, maybe play a little tune"

7.2.7.5. Screen

It was obvious to all visitors that the screen was used to explain the site. For the first two tours there were problems with the illumination of the screen, so that visitors had difficulty seeing the screen, while the sun was reflecting on the screen. Later on, the illumination problems were solved and the screen was bright enough to see even in full sunlight. Pictures used on the screen should be of high quality, which was, for example, not the case for the pictures of the theater at Point of Interest one in the Lion's Courtyard.

During transitions between Points of Interest, feedback was given on the screen, while the visitors following usually walked behind the robot. Because of this participants stated that they had the impression they missed out on information, especially about where the robot was going to stop next.

7.2.7.6. Pointer

"He is targeting something, isn't it?"

All participants stated that they liked the arm pointer, because it helped them to focus on the Point of Interest more than for example an audio-guide does. Only one recruited participant stated that an arrow on the screen would be more clear. Most visitors discovered the function of the pointer already at the first stop, while for some it took more time. When one of the visitors reacted to the pointer, most of the time, the others also started to see the pointer and looked at the details it targeted. Some participants stated that they did not see the direction the pointer indicated during the quiz, while this was intended to help them.

The pointer was often mentioned as the main modality that created the interaction between robot and people. According to three recruited participants the pointer could be better, by making its movements less mechanical, and more organic.

"I think the antenna was right. I really liked it, because it surprised me the first time"

7.2.7.7. Projection

Some participants indicated that they were not aware of the projection on the wall at the third Point of Interest, because the information was also shown on the screen of the robot, but flipped 90 degrees. Also, people did not expect that the robot would give the information using a different modality, as the screen had already been used at the first two Points of Interest. The projection was on a narrow yellow wall with a bench in front of it, people did not expect the robot to use the surroundings. One recruited participant stated that the projection was useful for larger groups, as more visitors can see the information than when the information was only given on the screen.

Due to technical limitations, the timing at the Point of Interest where the projector was used was different than those at the other places. Recruited participants made remarks of being confused by these moments of changed timing, or by the sound of the beamer turning on or by the flipped image on the screen. They would make remarks like "that seemed to go wrong, I think," or "for some reason I thought it was finished already, but then it took some time before it started with the guiding again."

7.2.7.8. Length of the tour

All participants stated that the length of the tour was OK; probably one location more would have been possible, but no more than that. Also the time spent at each location was fine according to most participants. Some would have liked the information at each Point of Interest to be a little longer. On the other hand, due to the fact that the information was short and the complete tour was not that long, participants recommend the robot for guiding children.

“I would like to know the length of the tour”

At the last Point of Interest it was unclear to most recruited participants that the tour had ended. This was caused by several things; first, the robot finished the tour in the middle of the Royal Alcázar. Some people stated that if the robot would finish close to the beginning, they would recognize that part and would easily understand that the tour had ended. Second, at the last Point of Interest, the robot plays special content, after which he states that the tour has ended. The length of the special content took some minutes, so people waited to go to the next point, but then the robot stated the tour was finished. Therefore, if these two actions were to be reversed, people would understand more easily that the tour has ended and then play a bit more with the robot and take pictures.

7.2.7.9. Navigation

“Ok, to which side…”

At the beginning of the tour, participants walked carefully behind the robot, because they needed to get used to the robot’s movements. In the interviews they stated that the behaviour of the robot was clear to them. However, most recruited visitors mentioned during the whole tour that they were unsure where the robot was going or if they were standing in its way when it started driving. Participants stated that the robot drove a bit jerkily and stopped every now and then, so they had to watch out not to be hit by the robot or bump into the robot. Due to this, some participants chose to walk next to the robot.

“You need to get used to follow the robot, it is jerky and so off course, so you need to get used to it, because where does it go?”

Participants mentioned that the robot was a bit slow while driving. One of the recruited participants said that he would like to physically push the robot a bit to go faster, some participants mentioned that for larger groups the speed might be alright, because larger groups tend to move a bit slower. When the robot was close to the Point of Interest, the robot needed a lot of time to find a place to settle. Also, at the Point of Interest, participants would appreciate if the robot could turn towards them before turning to the final position to start the story, because some participants felt ignored by the robot.

When the robot was moving, it gave too little feedback to the followers. A map and some text were displayed on the screen, but for those people who followed the robot and walked next to or behind the robot that information was not visible. Therefore, it was sometimes unclear to the participants where the robot was going. To obtain this information, visitors sometimes walked to the front of the robot to look at the screen. One recruited participant thought that having the robot driving backwards would improve this point.

“I felt best walking beside it”

7.2.7.10. Face recognition

In the later tours, when the robot stopped at a Point of Interest it showed a small window with the faces of the visitors who were close. The lines for emotion recognition were also visible in this window. From inquiry with the participants who saw the emotion recognition window we learned that they recognized themselves on the screen, and they assumed they were in the field of view of the robot. Nevertheless, they did not have any ideas on what the face tracking exactly was for. Also the robot did not give any clue about this, which would have helped to understand the meaning of the window.

Tracking faces to obtain data about the visitors' interest was sometimes difficult, when there was a group of people around the robot but none of them stood close enough to perform face recognition analyses. From our observations we learned that visitors gave each other room to look at the robot. Standing in front of the robot and blocking the view of other people, even if they joined later, would conflict with social rules. In these cases none of the people in the group stood close enough to track their faces.

7.2.7.11. Shortcomings

> *"We have limited time and don't want to miss something"*

From all visitors in the Royal Alcázar, just a few followed a tour given by the robot. Even though not all visitors can be guided by the robot, some more could be attracted if a few minor changes were made to the robot. For example, from our observations we saw that people hesitated to press "yes" when the robot asked if they wanted to follow a tour. This might be because they did not want to have the responsibility for starting the tour without knowing exactly what would happen. At the beginning of the tour and when visitors joined during the tour, visitors missed information about the length of the tour and the route of the tour. Not knowing how long the tour would take and where the robot would stop and what it would do was a problem for some visitors. Especially when visitors only had a limited amount of time, they wanted to see the whole Royal Alcázar, without missing anything. Because it was not clear from the beginning how long a tour would take and where the robot would go, people could decide not follow the tour or to leave the tour before it ended.

Also some people hesitated following the FROG because they obviously did not speak the language (English) well and probably did not understand what was expected of them.

7.2.7.12. Improvements

The participants mentioned several improvements for the robot. Some are quite small and easy to implement, but will nevertheless make large improvements to the user experience of the robot. Other improvements may be more difficult to implement, or will not be as great an improvement for all visitors. We derived requirements for improvement from the remarks as stated below. In some cases, several improvements mentioned by participants are combined into one requirement.

- The robot should speak Spanish and have English subtitles, so more people will be able to join the guided tour.
- During the quiz all text should be on the screen and it should be ordered in such a way that the text is not blocked by the eyes of the robot.
- At the last Point of Interest, the robot should first state that the tour has ended, and then play the music and other special content.
- The robot should first organize the orientation of the visitors before it turns to start the story at a Point of Interest.
- The text uttered with the 'robot voice' should be less repetitive, some different sentences or sound could be used instead.

- Add a kind of FAQ, so people can chose themselves what they would like to hear more about.
- Add an option that the robot can wait for some time before it travels to the next Point of Interest, so visitors have time to look around a bit more.
- At the start, allow people to give their name to the robot, so that the robot can speak to those visitors in particular all the time. A picture of the visitor could also be used to the same ends.
- The robot could identify people with colours that they would receive at the entrance. In this way it would follow the same visitors all the time, in this way reversing who follows whom. The robot follows visitors and visitors can choose what they want to hear.

7.2.8. Further development of FROG

Based on the results presented above we would advise to make some alterations to FROG that will largely improve the user experience of a guided tour given by FROG.

1. The speed of the robot, should be a bit faster during the transitions.
2. The utterances in 'robot voice,' need redesigning because they become repetitive and annoying.
3. The information on the screen should be easily visible without being blocked by the eyes of the robot.
4. The social navigation of the robot can be improved in such a way that the robot does not give the impression to visitors that it is ignoring them by turning away from them or not turning towards them.
5. The screen should have a higher illumination, so that it will also be visible in full sunlight.
6. More moments of interaction should be included in the tour, this will increase people's liking of the robotic tour guide.
7. The projector should be used at an earlier stage in the tour, because people did not expect a new modality that far in the tour.
8. At the last Point of Interest it should become clear that the tour has finished.
9. The design of the robot should be altered a bit to ensure that it looks as a guide for adults as well as one for children, because participants recommended the FROG often for children.
10. The face recognition feature should be made more flexible, because it is difficult to change people's social behaviour.
11. Give people options for language, because visitors who do not speak English avoided the robot.

After the June 2014 evaluation some changes for the robot were agreed for the next iteration of the FROG robot and for the sessions in September 2014. The changes are stated below.

To solve point 2, the FROG robot should use a different, less mechanical voice, so the text is easier to understand and it will give more feedback about its state. For example, FROG would not only utter that it is "navigating to next location" but when it becomes close, it would also say "almost there" and the utterances would be made less repetitive.

To solve point 3, the information in the screen would be ordered in another way, so the eyes would not block the visibility anymore.

To solve point 5, the screen would be used with its maximum illumination at all points, so that the screen would be visible in full sunlight as well.

To solve point 6, the content of FROG would be made more interactive and fun. This would be done by adding more interactive features starting from the first point of the tour and by having the tour guide and FROG interacting every now and then.

Point 7 would be partially solved. The projector would not be used earlier in the tour. However, the location of the stop in which the projector is used is closer to the previous point and the projection would be on a white wall without any other distractions close by.

To solve point 8, in the last stop the robot would state that the tour is finished before it would show the special content. Also the stop would be closer to the starting point, so people would recognize they have been there already and can take the opportunity to explore the site at a slower pace.

To solve point 11, on the interface of the robot three flags would be placed (Spanish, English and Dutch). Visitors could opt for one of those three flags to hear the guide (voice-over) explain the site in the language of their choice. The guide would have a Spanish accent when speaking English and Dutch, to reinforce the illusion the guide is from Spain.

7.2.9. Conclusions

From the user evaluations with invited participants and naïve visitors of June 2014 and later sessions, we have derived several suggestions for improving the user interaction design, which inform the final design of the FROG robot behaviours. Some of the findings point toward more general user experience issues that can be expected in outdoor cultural heritage settings. These include the speed of the robot, motion behaviours to walk along or to engage the visitor and feedback about the robot's state. Other improvements specifically concern how visitors of the Royal Alcázar can experience the site and learn more about the location.

The user evaluations show that the FROG project succeeded in guiding people in the Royal Alcázar and engaging them in a novel interactive experience. The robot was successful in engaging users, guiding them from one location to the next, engaging them in interaction, adapting the information it provides to their level of interest and allowing users to explore the information in a novel way using augmented reality.

Future projects that wish to move beyond state of the art in interaction design of guide robots in populated settings should specifically focus on:

- The robot's motion behaviours (to move more gracefully but also to indicate it's awareness of the visitors and allow easy walking alongside the robot. This impacts control techniques and navigation software)
- Recognition of the socio-emotional state of groups of users from facial feature recognition (taking into account all members of a group to build a model of affect and engagement).
- Alternative models of interaction design (the current design involves a virtual guide using the robot as a 'sidekick' in offering an interactive tour, future guide robots should explore alternative models. For instance, focusing on an ongoing natural social dialogue between a human guide and robot to explore social collaboration)

The user experiences evaluated offer a qualitative view on how a guide robot can inform visitors. Also, our sample is representative yet limited in size. We realize the context dependency of our findings but argue that many of the results can be used to inform the design of guide robots' behaviour in general. As far as we know, this is one of the first 'in-the-wild' studies where actual visitors explore a cultural heritage site on location, as part of a normal routine day. Therefore, we expect to use these findings not only for the further improvement of the interaction design of FROG but also to inform related projects where robots are deployed for guide services, such as, the FP7 framework project SPENCER.

7.3. Peer Group Workshop

On 25 September 2014, coordinators of 4 EU projects using robots (EASEL, SPENCER, Cargo-ANTS, MonarCH) were invited for a Peer Group Workshop. Part of the program was a full tour with the FROG robot in the Royal Alcázar in Seville. The consortium had asked the guests to treat this as an informal review and to put the robot through its paces and to be free with any criticism. Their comments are those of experts in the field but without knowledge of the Description of Work of the project.

The main Peer Group's Comments on FROG after tour were the following:

- narrative and interaction – it seems to be an intense interaction at the beginning then it just seems to run its program.
- narrative and interaction - it is also not clear who is the agent telling the story, having the robot different voices and "incarnations" – not very interesting
- motion - what is visible of the robot is 80% motion - but the way FROG moves does not fill the users with confidence to follow it (The navigation behaviour was degrading with time during the tour. Later we discovered that the tape on the wheels, used to reduce the wheels' friction, was totally worn through. The batteries were also running very low partly due to long delays while the guests tested the social navigation and obstacle avoidance and partly as the batteries were not fully charged to start with.)
- The robot needs better intention display.
- Perceptual capacity of the robot –we cannot suppose, or dare to suppose, that people know what the robot intends.
- FROG's facial features recognition is great (though not really necessary in FROG).
- At no point did I have the idea that it was being used to alter the system or FROG's behaviour through the perceptual feedback

Furthermore, some suggestions for improvements were made:

- the robot could reassure at every stop whether there are still visitors present
- it would be helpful to make more explicit – end of data file
- explore interactive environment
- more interaction capabilities, with questions like: "can I see you? Please check"
- recognize last 5 people when they come back

- Why not be a bit more pushy when the robot is on the move (and blocked by visitors)

7.4 Local Tour Guides Workshop

A third workshop was organized. The audience in this case was composed of Tour Guides. The goal of the workshop was to learn from the experiences of the guides with FROG and obtain guidelines from the guides for improvement of the robot's application at the Site. Also, we wanted to inform the guides of the Royal Alcázar and of the Lisbon Zoo about FROG and the objectives of the project.

An informal yet structured workshop was developed to study the experiences and gain feedback from 4 guides of the Royal Alcázar and 1 guide from the Lisbon Zoo.The workshop included a full autonomous tour given by FROG. In this section, the minutes of the workshop are presented. The subjects we prepared to discuss during the workshop are indicated by the paragraph headings.

7.4.1 Introduction to the Participants of the workshop

The workshop was led by Daphne Karreman and Vanessa Evers. No other FROG-partners were present during the workshop to ensure guides were able to talk freely and critically.

Guides that were present:

> **Guide 1** (m) has been a tour guide at the Lisbon City Zoo for eight years. According to him, being a tour guide is the best job in the world, because he can talk with children and people about the animals he loves and tell people how they can contribute to conservation of the animals. Guide 1 also participated in the observation studies of guide behaviour in the first year of the project (see D4.1b).
>
> **Guide 2** (f) guides people in Seville and the Royal Alcázar. She studied Art History in Seville and started as a tour guide 14 years ago, just after finishing her degree. Guide 2's main motivation is that she loves art. Guide 2 also participated in the observations of guide behaviour in the first year of the project.
>
> **Guide 3** (f) has been a guide for 2 years. Before, she worked as an doctor in the hospital, which was not a nice job, so she decided to study tourism. Now she feels lucky to work as a guide. She wanted to join the workshop, because she has some concerns. She waits at the entrance of the Royal Alcázar for tourists waiting a tour guide. This is an insecure job. Guide 3 had to leave early.
>
> **Guide 4** (f) is also a guide who offers her guide services at the entrance of the Alcázar. She arrived a bit later and had to leave early.
>
> One participant, the manager (m) of the audio guides company that rents audio guides in the Royal Alcazár joined the workshop. However, he had to leave soon after the start of the workshop because of technical issues with the audio guides at the town hall.

7.4.2 First impression of the robot

It became clear early on in the workshop that the tour guides had discussed the FROG robot amongst themselves. The guides who participated at the workshop therefore, also represented the thoughts and opinions of other guides.

After introductions, we asked the guides whether they had encountered the robot and what their first impressions of the robot were. From their responses it became clear that 2 of the guides were a bit worried about their jobs and that the robot might take away business. The first time Guide 3 saw FROG, she said it seemed a bit like ET. Also, she pointed out that the volume of the robot is very high. When she was with a group in an indoor room of the Royal Alcázar, it was difficult to speak to her audience because the robot was too loud. Guide 2 liked the robot; she thinks it can give the information in an interactive way. Guide 1 found the robot a good idea from an educational point of view. He thinks it is an easy way to offer information to people who would not read information offered at the animal exhibits. He also thought it gives information complimentary to human guides. Guide 1 explained that in the Zoo, the animals are not always active. If an animal remains unseen, a tour guide can only talk about the animal. However, the robot can show videos of the interesting behaviours of the animal that cannot be seen at that moment. Guide 3 added that for children the robot is great, because it can keep their attention for much longer time.

To inform the guides, a short presentation was given about the FROG project specifically and EU-research projects in general. Afterwards, we invited the guides to join a tour with the robot.

7.4.3 Impression of the robot after a guided tour

The tour guides then followed a demo-tour offered by the robot. Many of the other project partners were present during the tour, so that the guides could ask questions about the different aspects of the technology.

Figure 7.11. Impression of the FROG tour with the guides

After the tour was completed, we returned to the workshop room. Here, we asked the guides for their impression of the robot now they had experienced a full tour and we asked them for positive and negative feedback concerning the robot’s behaviour and information as well as the tour itself.

Guide 2 repeated that a negative point was the loud volume. She also thought another negative point is that that people can join the robot without paying; Because if you pay for the robot, you know it cannot answer all your questions. Also busy places will be difficult for the robot. Guide 4 agreed with this, she also mentioned that it might be difficult for the robot to send away people who did not pay to follow the tour. If you apply it in places for free, such as the Zoo, it might be useful when a robot tour is offered at specific hours that are advertised so that people know when they can join. Guide 3 said that the robot is fantastic for children. In the Royal Alcazár, tours by the robot in the gardens would be good, because there it would not be a problem if it were really loud.

7.4.4 About Guiding

In the next part of the workshop, we asked the guides about their jobs, what tasks they like best and what tasks they would like to do less. All the guides reported that they like their jobs because most of the people they guide around are grateful. People come to visit and follow a tour in their leisure time; they are relaxed and have fun. All guides mentioned that they always learn from their visitors. For example, when the visitors ask something, and they do not know the answer, they will look it up and use it in the following tours. This way they are always changing and improving their tours, as well as adapting their tours to the interests of the people that are in group.

The difference between a good and a not so great tour is usually the people that comprise the group. Some people join a tour, but do not show any interest. In these cases the guides try to finish the tour in the best way they can. Guide 2 added that, of course it is not always the visitors that make a tour go less well; sometimes a guide has a bad day. Guide 4 told us that she feels from the beginning if a tour will work, and also Guide 3 told that she feels “in her skin” when a tour is going well. Guide 2 mentions that she had good children groups, but also wonderful senior groups. It does not matter what kind of group they have, as long as the people in the group are interested. The guides agreed with a chuckle that the worst people to guide are foreign exchange students that have been partying all night and are expected to take the tour rather than joining out of interest; usually these students are too tired to show any interest. Guide 4 adds that also family groups with children and parents are sometimes difficult to guide. When parent do not look after the children or will not stop them from bad behaviour, the guide has to police the children as well as tell an interesting story to the parents. For Guide 1, the long walks between animal enclosures with young children are difficult as they are too small to walk these lengths. In these tours, he just goes to a few of the emblematic animals (bears or elephants). The guides do not report to have favorite subjects to tell about. As Guide 4 explains, this depends on the visitors; she will tell more about the subjects the visitors are interested in, so the focus can be for example more on history or architecture or the garden.

7.4.5 Possible ways to use the FROG in their own Tours

Guide 4 and Guide 3 especially reported some concerns about the robot as a guide, as it could take visitors, and the they would not be able to guide them anymore. However, all of the guides saw the opportunity to use the robot during their own tours. One of the main functions they liked was that the robot can be used to show how things were in the past in pictures, videos, projections or augmented reality. For example, in the rooms of the Royal Alcázar is no longer any furniture and visitors often ask what it used to look like. One of the guides mentioned that it would be nice if the robot could join on the tour and during these moments, they could ask the robot to show what it looked like, Also, the differences between the different architectural styles could be explained by FROG, by showing pictures and video’s of the different styles, as people often find it difficult to spot and understand the differences.

The robot could be used very well to guide groups of children, or groups that consist of children and parents. In this case, the guide could give information to the parents, while the robot entertains the children with an informative game, a video or a more simple explanation. Currently, the guides do this themselves. Guide 4 explains: “I often give children an assignment to search for animals in the gardens, while I tell the parents about what they see in the gardens”. Guide 1 also sees the advantages in the Zoo to take the FROG with him to give different information to different visitors in the group.

Figure 7.12. Impression of the discussion table in the workshop rooms

At this point in the workshop, Guide 3 and Guide 4 had to leave. We continued the workshop with Guide 2 and Guide 1.

Guide 2 explains that the best place for the FROG to give tours in the Royal Alcázar is probably in the 16th century area. This part of the Royal Alcázar links to history that is well known. It gives an impression of the people that were in Seville at the time and the events surrounding the discovery of America. Because, most people already know about the basics concerning Christopher Columbus' travels to America it would be nice if FROG could give some more information. In the Zoo, according to Guide 1, the tour of FROG and his own tour could complement each other. If FROG would tell about 10 animals, Guide 1 could tell about other animals. Finally, in the Royal Alcázar, FROG could give themed tours, for instance focusing specifically on architecture, art or history.

According to the guides a robotic tour guide FROG should be able to show pictures and videos. Guide 2 states that she thinks the re-creation of rooms is a good idea, to show people how things used to look. Guide 1 adds that games are also interesting to implement, because from an educational point of view, participation of visitors in the story is a consolidation of knowledge. In any case the robot should be fun, as people join the robot tour in leisure time. Or, as Guide 2 adds, timeslots for different kind of tours could be used so that people can choose themselves what tour to follow, either fun or more serious and factual. The fact that people pay attention to the robot instead of the exhibits will be a matter of time according to the guides. They think that after 15 minutes people will pay less attention to the robot and more to the Points of Interest. Guide 2 especially liked the re-creation of the "Crossing Courtyard" using the projection. For the Zoo there should be something changed during the 'dead-times' walking from one animal enclosure to the next, because these distances are much longer than the those in the Royal Alcázar.

The guides laughed when asked what things a guide is better at than a robot. Yes, of course, they feel. Guide 1 told us that he can interpret the behaviours and the smells of the animals in real time, and a robot would not be able to do that, yet. Guide 2 added that it is easier for a robot to tell about the buildings, as these are not alive, but that the robot cannot yet answer all questions that visitors ask. Especially at the Royal Alcázar, questions can be about many different concepts art, history, architecture, birds, plants, politics and so on. However, technology is improving, and probably more and more will be possible in the future.

The guides expressed their concerns about having tour guide robots. Robots like FROG could take visitors that would otherwise take a tour with a tour guide. However, in general guides are concerned about all kinds of changes in their industry. Their jobs are irregular and insecure. Income depends on the number of visitors and people choosing to take a tour. Changes have happened before. For instance, they are used to the audio guides by now, but 10 years ago, they were not happy with the introduction of this technology. According to Guide 2, it will take time to adapt to a robot tour guide, as it did for previous changes.

7.4.6 What is the best idea to implement in FROG

During the workshop, many thoughts and ideas on what the FROG could do were mentioned. We asked the guides which idea that was mentioned they felt would be the best idea to implement on FROG. Guide 2 stated that she liked the robot's ability to show pictures, videos and recreating what a place looked like in the past. Also, adding games would help visitors to be interested. Using these technologies, FROG could mix the explanations in a way that tour guides are not able to. Guide 1 mentioned two main things that would be interesting to implement in FROG. First, augmented reality to show what exhibits looked like in the past. Currently, the zoo favors the enclosures to be as natural as possible. For instance only water separates people from the primates. In the past, there were walls and cages, and that difference should be shown to make people aware of these changes in the way they deal with animals. Second, it would be nice if people could see themselves as if they were the animal in the enclosure. If one of the zookeepers would walk around an exhibit (without the animals in) and film it as if they were the animals, the video could be used to show to the visitors and they would have the perspective of the animal. Guide 1 feels this would create more understanding for the animal and its habitat.

7.4.7 First thing to implement in FROG

After asking them which idea they found best, we then asked the guides which of the ideas they felt should be implemented first. Guide 2 immediately said that the volume of FROGs explanations should be less loud. But also, in the case that the robot is offered as a private tour guide, it should be able to recognize which visitors belong to the tour and which do not. It should send away the people who have not paid. Guide 1 would definitely want to change something about the 'dead-times' in between animal enclosures when the visitors follow the robot to the next location. Otherwise, he feels the robot will loose people during the tour.

With these remarks we closed the session and thanked the guides for their participation in the workshop. We learned the guides' concerns for the robot, but also the guides gave us some fruitful ideas for further development of the robot. If the Lisbon City Zoo or the Royal Alcázar would further develop FROG to become a tour guide in their site, it would be great if the robot could give tours together with the guides or if the guides could use the robot in their tours, for example to recreate how things used to look.

8 Conclusions and Lessons Learnt

As described in the document, the FROG robot has been deployed in the Royal Alcázar for 4 weeks in 2 different months, 2 of the weeks devoted to guiding people “in the wild”.

The project has ended, thus, having a complete robot, with the technologies of the different partners integrated; we have completed nearly 16 hours of autonomous tours, 11 kilometres, 22 successful guided tours.

From the operational point of view, 85% of the tried tours were successful (22 out of 26). Nearly all failures were due to hardware issues (sensor failures).

From the user point of view, the user evaluations have shown that people were enthusiastic about the robot guide, both adults and children. This was confirmed by the acceptance the robot enjoyed in the city, with many visitors asking for it, and more than 40 persons (including 15 children) following the robot during the European Researchers’ Night. The users enjoyed the tour given by FROG.

The interactivity of the robot is deemed to be a key feature, the users would have liked to see even more interactivity on the tours. Users were also amazed by the fact that the robot autonomously navigated the full tour with no intervention. But, while most of the time it was clear the behaviour of the robot, the readability of its motion should be improved. Furthermore, its behaviour should be smoothed a bit.

Also, some of the capabilities of the robot are not apparent to the users, like its affective computing capabilities. The experiments have also shown that the behaviour of the persons with respect to the robot made it difficult for the affective computing camera to get information, and this should be considered for future designs.

The robot appearance is perceived as friendly

The workshop with the guides helped to, first, discuss their concerns about the use of robots for these activities; in particular, as a threat for their jobs. All through the project, some guides expressed these concerns whenever they saw the robot in action. During the workshop, all guides saw the opportunity to use the robot during their own tours. They have seen the added value offered by some of its capabilities, like Augmented Reality, or interactive games. And some guides even considered the possibility of having a robot as co-worker in some places.

References

[1] Rafael Munoz-Salinas. ARUCO: a minimal library for augmented reality applications based on OpenCV. http://www.uco.es/investiga/grupos/ava/node/26.

[2] B. Gerkey and K. Konolige. Planning and control in unstructured terrain. In Workshop on Path Planning on Costmaps, Proceedings of the IEEE International Conference on Robotics and Automation, 2008.

[3] Edward T. Hall. The Hidden Dimension. Anchor, October 1990.

[4] Friese, Susanne. *Qualitative data analysis with ATLAS. ti*. Sage, 2014