In Proceedings of the 10th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Nashville, TN, December 2010

Intuitive Multimodal Interaction and Predictable Behavior for the Museum Tour Guide Robot Robotinho

Matthias Nieuwenhuisen, Judith Gaspers, Oliver Tischler, and Sven Behnke

***Abstract*— Deploying robots at public places exposes highly complex systems to a variety of potential interaction partners of all ages and with different technical backgrounds. Most of these individuals may have never interacted with a robot before. This raises the need for robots with an intuitive user interface, usable without prior training. Furthermore, predictable robot behavior is essential to allow for cooperative behavior on the human side. Humanoid robots are advantageous for this purpose, as they look familiar to persons without robotic experience. Moreover, they are able to resemble human motions and behaviors, allowing intuitive human-robot-interaction.**

In this paper, we present our communication robot Robotinho. Robotinho is an anthropomorphic robot equipped with an expressive communication head. Its multimodal dialog system incorporates body language, gestures, facial expressions, and speech. We describe the behaviors used to interact with inexperienced users in a museum tour guide scenario. In contrast to previous work, our robot interacts with the visitors not only at the exhibits, but also while it is navigating to the next exhibit. We evaluated our system in a science museum and report quantitative and qualitative feedback from the users.

I. Introduction

In the last decades, tremendous progress has been made in improving the core functionalities of autonomous mobile robot systems. For many years, the main focus was on improving the hard skills, namely planning, safe navigation, and reliable perception. Meanwhile, human-robot-interaction and the appearance of robots to humans has become an active research area. So far, virtually all working complex robotic systems are found in industrial domains or research institutes. Here, only few specialized persons operate these robots, utilizing complex user interfaces.

The vast majority of persons have never interacted with a robot so far. Hence, they are not used to the presence of robots. Persons of different ages and technical backgrounds raise the need for interfaces that are intuitive and easy to use. Deploying robots as a museum tour guide gives a good opportunity to gather experience in human-robot-interaction. It also has the advantage that the robot itself can even be seen as an exhibit. The domain of a museum usually is a bounded, but highly dynamic environment. Many users that never interacted with a robot before, including children and elderly persons, will try to control the robot. The interaction time in a museum is usually short, limited to the museum visit. Hence, an interface that is intuitive to use without prior introduction is essential.

M. Nieuwenhuisen, O. Tischler, and S. Behnke are with the Autonomous Intelligent Systems Group, Computer Science Institute VI, University of Bonn, Germany. J. Gaspers is now with the Semantic Computing Group, Cognitive Interaction Technology - Center of Excellence (CITEC), University of Bielefeld, Germany

Fig. 1. Our museum tour guide robot Robotinho explains an exhibit at the *Deutsches Museum Bonn* to a group of children. Visitors listening to explanations usually block the robot's path to its next destination. Clearly articulating the desired motion direction is essential for successful navigation.

Our museum tour guide robot Robotinho, shown in Fig. 1, has an anthropomorphic appearance. This facilitates users to predict the robot's motions and behaviors. Additionally, it is equipped with an expressive communication head to display the robot's mood.

Sudden movements of the robot can startle museum visitors. To avoid this, it is necessary to communicate the robot's intentions either verbally or non-verbally in a natural way.

In a museum, robots not only have to communicate with visitors. Safe and reliable navigation in dynamic environments is essential. Also interaction with visitors attending the tour or strolling around in the proximity of the tour is an important skill.

The evaluation of service robots is difficult. Basic skills, like collision-free navigation, accurate localization, and path planning, may be evaluated in a laboratory using objective and comparable metrics. Evaluating factors like the appearance to and interaction with people not familiar with robots requires operating the robots in public environments. Robotinho was evaluated before in different scenarios:

- Conducting cellists of the Berlin Philharmonic. This required extensive use of body language.
- Explaining exhibits in a static scenario using its multimodal interaction skills.
- Guiding visitors to exhibits in a corridor at the University of Freiburg.
- Communication robot at RoboCup@Home 2009. Our team NimbRo won the innovation award for "Innovative robot body design, empathic behaviors, and robot-robot cooperation".

Details of the previous evaluations are given in [6] and [10]. Extending our prior work, we developed new behaviors to interact with museum visitors during the navigation to the next exhibit. In this paper, we describe our approach to this mobile interaction. We also report our experiences made in a science museum and present quantitative results from questionnaires.

The remainder of this paper is organized as follows: After a discussion of related work, we briefly describe the robot's hard- and software design in Sec. III. Sec. IV details our approach to intuitive multimodal human-robot interaction. Our system to track people in the robots vicinity is explained in Sec. V. In Sec. VI, we present the techniques used to guide visitors through a museum. We report the results of our evaluation in a science museum in Sec. VII. The paper concludes with a discussion of the conducted experiments.

II. Related Work

The idea of enhancing the museum experience by the use of robots has been pursued by several research groups. Wheeled robots, for example, have already been deployed as museum tour guides or on large fairs [11], [15], [18], [3]. The main focus in these systems, however, was reliable, collision-free navigation. The researchers of these projects did not emphasize natural, human-like interaction.

Humanoid robots offer the potential to realize intuitive and more human-like interaction. Bischoff and Graefe [2] installed the robot Hermes for a long-term experiment in a museum. Hermes possesses an upper body with arm that is mounted on a wheeled base. However, the robot does not have an animated face and its multimodal interaction capabilities are limited.

Dautenhahn et al. [5] evaluated how to approach and navigate in the presence of humans to avoid having the users feel uncomfortable. This implies being visible to the individual most of the time and approaching them from the front. In our work, we focus on predictive behavior by clearly articulating the robot's intents.

Recently, Shiomi et al. [14] studied if tour guiding robots should drive forward or backward, facing the visitors, during navigation to keep them interested in the tour. In contrast, we focus on the gaze direction of the robot.

The tracking of persons using laser-range sensors and cameras has been investigated, e.g. by Cui et al. [4], Schulz [13], and Spinello et al. [16]. In contrast to these works, our approach uses laser-range finders (LRFs) on two heights to detect legs and trunks of persons and also utilizes a static map of the environment to reject false positive hypotheses.

III. Robot Design

The hardware design for our robot Robotinho focused on low weight, sleek appearance, and dexterity ([6],[10]). These are important features for a robot that interacts with people in daily life environments.

The robot is fully autonomous. Robotinho has an anthropomorphic appearance, which supports human-like multimodal communication. For use as a communication robot, we equipped Robotinho with an expressive 15 DOF head, depicted in Fig. 2.

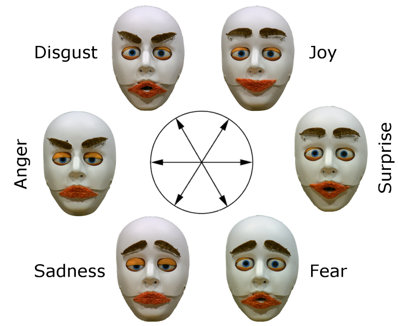

Fig. 2. Six basic facial expressions generated by Robotinho's expressive head.

Fig. 4. Robotinho directs the audience's attention to distinctive parts of an exhibit using pointing gestures.

Although Robotinho is able to walk, we placed it on a wheeled base that can drive omnidirectionally. This ensures faster and safer movement in dynamic environments. On its base Robotinho is about 160cm tall, simplifying the face-to-face communication with adults.

As tasks like speech recognition, vision, and localization require substantial computational power, we distributed these modules to two distinct computers. Most of the computing tasks related to the dialog system are performed by one PC, whereas the other PC is dedicated to localization and navigation, using the robotic framework Player [7].

IV. Multimodal Interaction

A. Gestures

Our robot Robotinho performs several human-like gestures. These include subconscious gestures, like questioning gestures, arm movements to ask for clearance in front of it to navigate, and intentional gestures. Intentional gestures are used to attain or direct the attention of the communication partners. Examples of gestures to attain attention are come-closer and waving gestures. Fig. 4 shows Robotinho explaining different parts of an exhibit, pointing to the currently explained part.

B. Speech

In fully autonomous operation, we recognize speech using a commercial speech recognition system [9]. This system is speaker-independent and uses a small vocabulary grammar

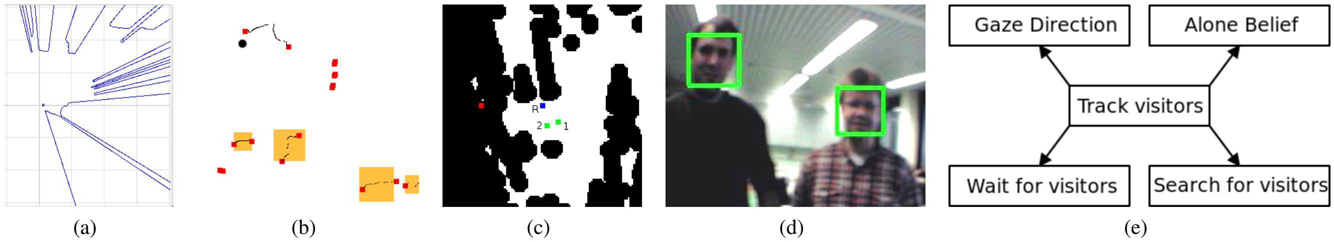

Fig. 3. Person tracking: Laser-range measurements (a) of two LRFs are used to detect trunk and leg features. (b) depicts segments of a laser scan (black lines with red boxes) identified as trunk candidates. The orange boxes correspond to the estimated positions and size of the accepted trunk features. (c) Detections (green boxes (1,2) and red box) close to or inside static obstacles are rejected (red box) using a map of the environment and the robot's position (blue box (R)). The likelihood of face detections is increased, if a range measurement falls within an angular threshold (d). The fused measurements and the raw laser measurements are used to keep eye-contact with visitors, to wait or search for guided visitors, and to calculate the belief that the robot is alone (e).

which is changed corresponding to the dialog state. In semi-autonomous mode an operator can select recognition results using a wireless connection. Qualitatively human-like speech is synthesized online by a text-to-speech system.

C. Emotional Expressions

In addition to the generated speech, humans emphasize their utterances by the use of emotional expressions. Humans are trained to quickly appreciate emotions and interpret the communication partner's behavior accordingly. Robotinho can express emotions by facial expressions and by emotional speech synthesis. Fig. 2 shows the six basic emotional states that we use to compute the current facial expression, following the notion of the Emotion Disc [12]. For emotional speech synthesis, we adjust the parameters pitch, speed, and volume of our speech synthesis system [1].

V. People Awareness and Tracking

To involve visitors actively in conversations, it is necessary to know their whereabouts. We detect and track persons using fused measurements of vision and LRFs. The two cameras are used to detect faces, using the approach from Viola & Jones [19]. In the laser-range measurements, we detect trunks and legs, associating and tracking them using a multi-hypothesis-tracker with Hungarian data association [8]. The sensor fusion allows us to reject detected faces in regions without corresponding range measurements. Face detections z_t^c and tracks gained by laser-range measurements z_t^l are associated according to their angular distance within a threshold. The resulting state estimate that also incorporates the prior belief is given as

$$p(x_{1:t}|z_{1:t}^c, z_{1:t}^l) = p(z_t^c|x_t, z_t^l)\overline{bel(x_t)}. \qquad (1)$$

Our measurement model $p(z_t^c|x_t, z_t^l)$ is implemented as a lookup table. The belief update is performed by linear Kalman filtering [20]. Laser tracks are updated independently before the sensor fusion. Fig. 3 illustrates our person tracking algorithm.

The field of view (FOV) of the LRF in Robotinho's neck is 240°. However, the robot can turn its upper body and with it the LRF. To track people in a larger opening angle while moving, it gazes alternatingly into the direction of different known tracks to update their belief. As we are mainly interested in the tracking of groups, we do not need accurate tracks of every visitor. Thus, we give priority to a human-like looking behavior and only look at one random track per time segment. The lengths of the time segments are randomly chosen within an interval. Thus, the current gaze direction α at time t depends on the active time interval T_i and the number of tracks currently maintained in the set of tracks L. If no tracks exist, Robotinho turns the upper body and its head to explore the space behind it. Here, the turning direction is determined by the last known track positions. Be α_d the angle of the driving direction, α_{l_t} be the angle of track l at time t and $\alpha_{\max}$ be the maximum possible turn angle we arrive at

$$l_t = \begin{cases} l \in L_t, & \text{if } \|L_t\| > 0 \\ l_{t-1}, & \text{otherwise} \end{cases},$$
$$\alpha = \begin{cases} \alpha_d, & \text{if } t \in T_1 \wedge \|L_t\| > 0 \\ \alpha_{l_t} & \text{if } t \in T_2 \wedge \|L_t\| > 0 \\ \operatorname{sgn}(\alpha_{l_t})\alpha_{\max} & \text{if } \|L_t\| = 0 \end{cases}$$

to calculate the gaze direction.

Similar to the validation of face detections by LRF measurements, the validation of hypotheses achieved by laser-range measurements with face detections is possible. We neglect this as the FOV of the LRFs is much larger than the FOV of the cameras and a few spurious hypotheses are not disadvantageous to our purpose. However, we filter most of the spurious measurements by comparison with the static map, as shown in Fig. 3c.

VI. Skills for a Tour Guide

In a museum tour guide scenario, a robot has to navigate in a dynamic environment. Most of the people in a museum have never interacted with a robot before and hence can't predict the robot's reactions. Thus, the robot has to indicate its intentions and abilities in an intuitive human-like manner. For instance, many people may not be aware of the possibility to communicate with the robot. Also unexpected actions of the robot, e.g. the sudden start of movements, may startle visitors.

Individuals may feel uncomfortable, if a robot drives past them too close. We use the tracked positions of persons to

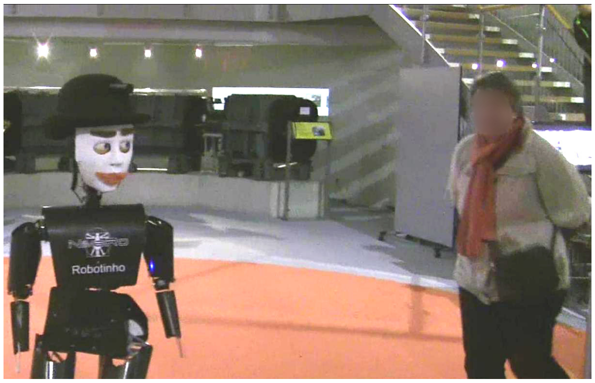

Fig. 5. Robotinho attends guided visitors during navigation. Alternately it looks into the driving direction and to the persons following it. If the guided persons are falling back, the robot turns around and request them to catch up.

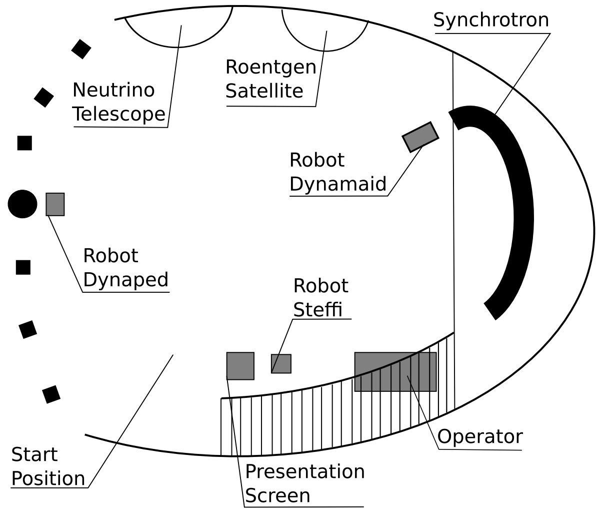

Fig. 6. Robotinho operated in the central area of the museum. It explained three permanent exhibits and three of our robots.

influence the repulsive forces of obstacles in our local path planner. If possible, a path closer to static than dynamic obstacles is chosen.

Our robot has an attentional system, reacting to the presence of humans. It shows interest in multiple persons in its vicinity, by alternating looking at them. As laser-range finders offer only a 2D slice of the environment looking into people's faces relies on visual detections. However, we use LRF detections to direct Robotinho's attentional system to newly arrived individuals. After being alone, Robotinho addresses newly detected persons by offering a tour or asking them to come closer, backed by a come-closer gesture.

We represent every exhibit as the 3D-position of the exhibit, an optional list of 3D-positions of parts of the exhibit, a preferred robot pose and the explanations Robotinho shall give. The poses are used for navigational purposes and the object positions are necessary to perform pointing gestures.

Performing gestures using inverse kinematics enables Robotinho to point to exhibits from any pose. In addition to pointing to an exhibit from not exactly reached poses, we use this to point to the next exhibit in the tour before starting to drive. This makes the behavior of the robot much more predictable to the visitors.

In addition to natural interaction during the explanation of exhibits through gestures, facial expressions, and speech, we are convinced that interaction with the visitors during transfers between exhibits is essential. A good tour guide has to keep visitors involved in the tour. Otherwise, it is likely that visitors leave the tour. Hence, our robot looks alternatingly into its driving direction and to its followers. Looking into the driving direction shows the visitors that it is aware of the situation in front of it and facilitates the prediction of its movement. Looking into the direction of the guided visitors is both necessary to update the robots belief about their positions and to show interest in the persons it interacts with. Following the approach described in the previous section, Robotinho gives its attention to a random visitor, if the positions of persons are known. Otherwise, it looks over its shoulder to find its followers (see Fig. 5).

If Robotinho is uncertain about the whereabouts of its followers, or if they fall back, its head and upper body are turned and a come-closer gesture backed by a verbal request to follow the guide is performed. Additionally, it can turn its base to look and wait for the visitors. If this is successful, our robot indicates verbally that it became aware of the presence of its followers and continues the tour.

The dynamic nature of a museum environment can cause stalls in the robot movement. If persons step into the safety margins around the robot or the path is globally blocked by persons standing around the robot, Robotinho asks for clearance. This request is backed by an angry facial expression. This has been shown before to be more effective then the clearance request alone [18].

VII. Evaluation

A. Scenario

We evaluated our museum tour guide robot Robotinho in the *Deutsches Museum Bonn*, a public science museum. The permanent exhibition focuses on research and technology in Germany after 1945. Hence, most visitors are interested in and are open-minded to new technologies. Furthermore, a number of workshops for children is offered. Consequently, a broad range of age groups was present in the museum.

Robotinho operated in a central area of the museum. The permanent exhibits residing in this area are a synchrotron, a neutrino telescope and parts of a Roentgen satellite. We set up three more exhibits: Our anthropomorphic service robot *Dynamaid* [17], our TeenSize soccer robot *Dynaped* [21], and our KidSize soccer robot *Steffi*. Fig. 6 shows the area of the museum we used and the placement of our exhibits in it.

From a starting position, Robotinho looked for visitors in its vicinity. If visitors showed interest, Robotinho explained itself and showed some of its skills, like performing gestures and showing emotions. The decision whether persons are showing interest is calculated by evaluating an *alone belief*. The alone belief is determined by incorporating the person awareness algorithm described in section V. Afterwards, Robotinho offered two different tours, one explaining the three permanent exhibits and one explaining the other robots.

TABLE I

"DO YOU LIKE ROBOTS IN GENERAL?"

(in %)	1 (not at all)	2	3	4	5	6	7 (exceedingly)	μ
adults	1	3	4	7	21	**34**	26	5.5
children	0	0	0	10	15	26	**47**	6.1

TABLE II

"HOW DOES THE ROBOT APPEAR TO YOU?"

	(in %)	1 (not at all)	2	3	4	5	6	7 (very)	μ
Adults	appealing	0	2	6	8	25	**32**	24	5.5
	polite	0	1	2	2	7	36	**50**	6.3
	friendly	0	1	0	3	19	**37**	**37**	6.1
	attentive	1	7	3	12	22	**31**	21	5.3
	manlike	6	6	**25**	19	24	14	2	4.1
	attractive	2	6	3	14	27	**33**	12	5.1
	clumsy	16	14	14	16	**18**	16	3	3.7
Children	appealing	7	5	0	5	17	23	**41**	5.5
	polite	2	2	0	0	7	23	**64**	6.3
	friendly	0	0	2	4	12	29	**51**	6.2
	attentive	2	5	5	7	7	20	**52**	5.1
	manlike	5	5	7	12	15	**30**	23	5.3
	attractive	0	2	2	7	15	27	**45**	6.0
	clumsy	**42**	12	5	17	10	12	0	2.8

TABLE III

"HOW DID YOU PERCEIVE THE COMMUNICATION WITH THE ROBOT?"

	(in %)	1 (not at all)	2	3	4	5	6	7 (very)	μ
Adults	intuitive	6	6	11	21	**23**	17	12	4.5
	easy	0	4	11	8	22	25	**27**	5.3
	artificial	1	3	13	21	**33**	10	16	4.8
	manlike	6	6	**25**	19	24	14	2	4.0
	convenient	1	6	4	20	21	**29**	15	5.1
	cumbersome	16	**32**	8	18	15	5	3	3.1
	labored	13	**25**	17	22	13	3	3	3.2
Children	intuitive	9	6	6	6	15	**28**	**28**	5.1
	easy	8	2	2	13	13	**32**	27	5.3
	artificial	20	7	10	12	15	**23**	10	4.1
	manlike	5	5	7	12	15	**30**	23	5.1
	convenient	0	0	7	2	20	20	**48**	6.0
	cumbersome	**33**	7	10	5	23	7	12	3.5
	labored	**31**	21	5	15	10	5	10	3.1

A possibility to continue with the other tour was given after the first tour finished. Finally, the robot said goodbye and asked the visitors to fill out a questionnaire. The overall duration of both tours, including the introduction and end, was about 10 minutes.

The experiments were performed on two consecutive weekends in January 2010. A video summarizing the museum tours is available on our website (NimbRo.net).

B. Results

After every tour, the visitors were asked to fill out a questionnaire containing questions about the general appearance of the robot, the verbal and non-verbal communication skills of the robot and the tour itself. Most questions could be answered on a scale from one to seven. Robotinho performed 40 tours and 129 questionnaires were filled out. The total number of persons (partially) attending the tour, but not filling out a questionnaire was not separately counted.

In the remainder of this section, we will aggregate the results into positive (marks 6-7), average (3-5) and negative (1-2) answers, if not stated otherwise.

First of all, most persons replied that they like robots in general. Over 70% of the children ($<$ 15 years) and over 60% of the polled adults ($\geq$ 15 years) answered with the two top marks. Only 5% of the adults and none of the children answered with the lowest marks (cf. Table I).

More than three quarter of the persons say that the robot appeared friendly and polite to them. Nearly as many children think the communication with the robot is convenient. 45% of the adults gave a positive answer and only 7% gave a negative answer (cf. Table II).

Robotinho appeared very attentive to 72% of the children and 52% of the adults (cf. Table II). 63% of the children and 48% of the adults felt that the robot paid them adequate attention (cf. Table IV). Reasons for high marks given as free text are that Robotinho gazes at persons and reacts on blocked pathes. The main reason for tour attendance stated by the visitors is that the tours are performed by a robot.

The average of the answers on how friendly, manlike, appealing, and polite the robot appeared shows a trend to correspond to the question if persons generally like robots. The same applies to how pronounced the verbal communication skills (cf. Table V) and how intuitive and convenient the communication at all was rated.

In the museum, we had the experience that in groups of visitors often contradicting answers were given at the same time. Also some visitors reacted only by nodding or head shaking to yes/no questions of the robot. Many children asked the robot interposed questions, not only about the tour, but also about other exhibits. For these reasons, we manually selected the appropriate phrase to be recognized, if the robot could not recognize the phrase automatically.

Generally, children appraised the robot to be more humanlike. Consequently, the mean of the answers given by children to the questions regarding to human-likeliness of the appearance and communication (cf. Tables II, III) is in both cases more than one mark higher than the mean of the adult's answers. In contrast to adults, many children have no reservation against the robot. Hence, groups of children were often surrounding the robot closely while adults mostly stood back. Also, adults were often more observant, waiting for the robot to progress by itself. This may be induced by the learned expectation that natural interaction with machines is mostly not possible.

At the beginning of our evaluation period, we omitted to let Robotinho announce its next destination during tours. This worked quite well in situations with only a few visitors, but broke down when visitors surrounded the robot. As the visitors were not aware about the robot's intention, Robot-

TABLE IV

"DO YOU THINK, THE ROBOT ATTENDED YOU ADEQUATELY?"

(in %)	1 (not at all)	2	3	4	5	6	7 (highly)	μ
adults	0	1	6	21	21	**40**	8	5.2
children	0	4	0	12	19	24	**39**	5.8

TABLE V
"HOW PRONOUNCED DID YOU EXPERIENCE THE VERBAL COMMUNICATION SKILLS / THE NON-VERBAL INTERACTION SKILLS OF THE ROBOT, E.G. EYE-CONTACT?"

	not at all						very	
(in %)	1	2	3	4	5	6	7	μ
non-verbal (adults)	3	4	6	7	27	**29**	20	5.2
non-verbal (children)	2	2	2	10	17	12	**51**	5.8
verbal (adults)	1	9	7	16	**29**	23	12	4.8
verbal (children)	0	0	5	20	15	22	**37**	5.7

inho had to ask multiple times for clearance, until the visitors stepped several meters back. We observed that indicating the robot's intention by announcing the next exhibit and pointing to it causes the visitors to look at the next exhibit and to open a passageway into the right direction.

The importance of interaction with visitors during the navigation from one exhibit to the next became clear in our experiments. As the environment where we performed the tours is clearly arranged, many persons stayed back during tours and watched the robot from a distance. The majority of these persons, and some visitors only strolling around in the vicinity of the robot, followed the request of Robotinho to come closer again. In one situation, even a large group of visitors sitting at the boundary of the exhibition area stood up and went to the tour guide after its request.

VIII. CONCLUSIONS

Although our communication robots were successfully evaluated before in different static and mobile scenarios, the mobile evaluations in the past took all place in non-public environments. In this paper, we evaluated Robotinho in a dynamic tour guide scenario in a public science museum. Our robot interacted with a large number of users who were unfamiliar with robots.

For this purpose, it was necessary to extend the multimodal interaction skills of Robotinho with new behaviors to communicate its intended navigation goals and to keep track and interact with visitors during the navigation.

The results of our evaluation show that the visitors are generally open-minded and interested in robot use. Many persons gave high marks for the interaction capabilities and typical human attributes like politeness and friendliness. In nearly all tours, the tour explaining our robots was chosen first by the communication partners and most of them proceeded with the second tour afterwards.

A major and almost unsolved problem in public environments is speech understanding. It is not only difficult to recognize speech independent of the speaker in noisy environments, but also the robot has to resolve ambiguities due to different answers given simultaneously by a group of visitors.

The anthropomorphic appearance and the multimodal interaction system induce high expectations in the communication abilities of the robot. Children see the robot as very human like, using terms like illness if the robot is shut down and asking lots of general questions to the robot. Generally, the robot was well received by the museum's visitors. At the beginning, we neglected to communicate the robot's navigation goals, leading to disruptions in the tours until the visitors guessed the correct robot intentions. We observed a substantial improvement in the navigation between exhibits after activating the pointing to the next exhibit.

In future work, we aim at integrating the communication abilities of Robotinho into our domestic service robot Dynamaid [17].

ACKNOWLEDGMENTS

The authors wish to thank Dr. Andrea Niehaus and her team at *Deutsches Museum Bonn* for providing the location and their support before and during our tests.

This work has been supported partially by grant BE 2556/2-3 of German Research Foundation (DFG).

REFERENCES

[1] M. Bennewitz, F. Faber, D. Joho, and S. Behnke. *Humanoid Robots: Human-like Machines*, chapter Intuitive Multimodal Interaction with Communication Robot Fritz. Intech, Vienna, Austria, 2007.

[2] R. Bischoff and V. Graefe. Demonstrating the humanoid robot HERMES at an exhibition: A long-term dependability test. In *Proc. of IROS; Workshop on Robots at Exhibitions*, 2002.

[3] W. Burgard, A. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence 114(1-2):3-55*, 1999.

[4] J. Cui, H. Zha, H. Zhao, and R. Shibasaki. Tracking multiple people using laser and vision. In *Proc. of IROS*, 2005.

[5] K. Dautenhahn, M. Walters, S. Woods, K. Koay, C. Nehaniv, A. Sisbot, R. Alami, and T. Simeon. How may I serve you?: a robot companion approaching a seated person in a helping context. In *Proc. HRI*, 2006.

[6] F. Faber, M. Bennewitz, C. Eppner, A. Grg, C. Gonsior, D. Joho, M. Schreiber, and S. Behnke. The Humanoid Museum Tour Guide Robotinho. In *Proc. of Humanoids*, 2009.

[7] B. Gerkey, R. T. Vaughan, and A. Howard. The Player/Stage project: Tools for multi-robot and distributed sensor systems. In *ICAR '03*.

[8] H. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1):83–97, 1955.

[9] Loquendo S.p.A. Vocal technology and services. loquendo.com, 2009.

[10] M. Nieuwenhuisen, J. Stückler, and S. Behnke. Intuitive Multimodal Interaction for Domestic Service Robots. In *Proc. ISR/ROBOTIK '10*.

[11] I. Nourbakhsh, C. Kunz, and T. Willeke. The mobot museum robot installations: A five year experiment. In *Proc. of IROS*, 2003.

[12] Z. Ruttkay, H. Noot, and P. ten Hagen. Emotion Disc and Emotion Squares: Tools to explore the facial expression space. *Computer Graphics Forum*, 22(1):49–53, 2003.

[13] D. Schulz. A Probabilistic Exemplar Approach to Combine Laser and Vision for Person Tracking. In *Proc. of RSS*, 2006.

[14] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. A larger audience, please!: encouraging people to listen to a guide robot. In *HRI '10*.

[15] R. Siegwart, K. Arras, S. Bouabdallah, D. Burnier, G. Froidevaux, X. Greppin, B. Jensen, A. Lorotte, L. Mayor, M. Meisser, et al. Robox at Expo. 02: A large-scale installation of personal robots. *Robotics and Autonomous Systems*, 42(3-4):203–222, 2003.

[16] L. Spinello, R. Triebel, and R. Siegwart. Multimodal detection and tracking of pedestrians in urban environments with explicit ground plane extraction. In *Proc. of IROS*, 2008.

[17] J. Stückler and S. Behnke. Integrating Indoor Mobility, Object Manipulation and Intuitive Interaction for Domestic Service Tasks. In *Proc. of Humanoids*, 2009.

[18] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, et al. Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva. *The International Journal of Robotics Research*, 19(11):972, 2000.

[19] P. Viola and M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In *Proceedings of IEEE CVPR*, 2001.

[20] G. Welch and G. Bishop. An introduction to the Kalman filter. *University of North Carolina at Chapel Hill, Chapel Hill, NC*, 1995.

[21] T. Wilken, M. Missura, and S. Behnke. Designing falling motions for a humanoid soccer goalie. In *Proc. of the 4th Workshop on Humanoid Soccer Robots, Int. Conf. on Humanoid Robots*, 2009.