In Proceeding of The 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Osaka, Japan, 2010.

Intuitive Multimodal Interaction for Service Robots

Matthias Nieuwenhuisen, Jörg Stückler, and Sven Behnke
Computer Science Institute VI, Autonomous Intelligent Systems
University of Bonn, Germany

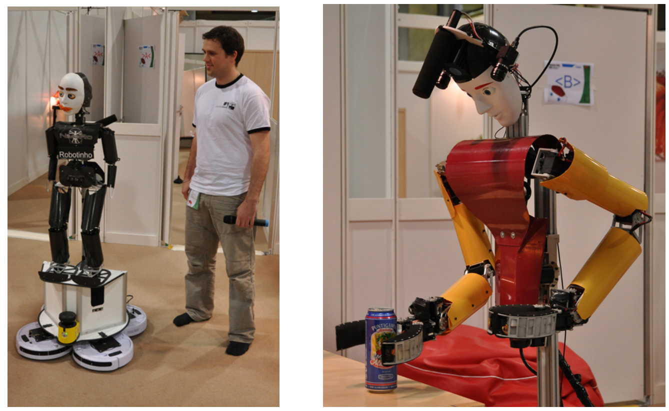
***Abstract*—Domestic service tasks require three main skills from autonomous robots: robust navigation in indoor environments, flexible object manipulation, and intuitive communication with the users. In this report, we present the communication skills of our anthropomorphic service and communication robots Dynamaid and Robotinho. Both robots are equipped with an intuitive multimodal communication system, including speech synthesis and recognition, gestures and mimic. We evaluate our systems in the @Home league of the RoboCup competitions and in a museum tour guide scenario.**

***Index Terms*—anthropomorphism, multimodal human-robot-interaction; service robotics**

I. Introduction

Advanced domestic service robots should make use of multiple modalities such as speech, facial expressions, gestures, body language, etc. to interact with people. If successful, this approach yields a user interface that leverages the evolution of human communication and that is intuitive to naive users, as they have practiced it since early childhood.

In this report, we present our robots Robotinho [1] and Dynamaid [2] (see Fig. 1). Robotinho is a full-body humanoid robot that has been equipped with an expressive communication head and an omnidirectional drive. Dynamaid has an anthropomorphic upper body, a movable trunk, and is able to drive omnidirectionally as well. The anthropomorphic appearance of our robots supports human-like multimodal communication.



Robotinho Dynamaid

Fig. 1: Our robots are about $160cm$ and $180cm$ tall. For the use as communication robot, we equipped Robotinho with an expressive 15 DOF head. The eyes are movable cameras. Additionally, Robotinho can animate its jaw, mouth, eyelids and eyebrows. Dynamaid's upper body consists of two anthropomorphic arms, a movable head, and an actuated trunk.

II. Dialogue System

Both robots are equipped with a multimodal dialogue system. They can perceive persons in their environment, recognize and synthesize speech. Robotinho can express its emotional state and shifts its attention between different persons. Our robots perform human-like arm gestures during the conversation and also use pointing gestures generated with eyes, head, and arms to direct the attention of their interaction partners towards objects in their environment (see Fig. 2a).

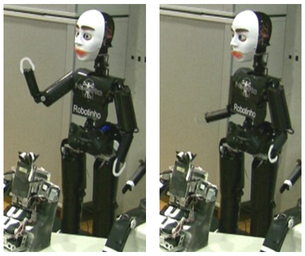
To interact with human users, a robot requires situational awareness of the persons in its surrounding. For this purpose, our robots detect and keep track of nearby persons using fused measurements of two laser-range finders. Robotinho additionally detects and tracks people in the environment using the images of its two cameras [3]. Performing tasks for humans requires the ability to recognize people. To this end, we utilize a face enrollment and identification system. Face descriptors are learned during the first interaction with individuals. These descriptors are used to identify persons during further operation.

Robotinho shows interest in multiple persons in its vicinity and shifts its attention between them so that they feel involved into the conversation. To determine the focus of attention of the robot, we compute an importance value depending on the distance and position of persons. The robot focuses its attention on the person who has the highest importance, which means that it keeps eye-contact with this person. While focusing on one person, from time to time, our robot also looks into the direction of other people to involve them into a conversation and to update its belief. Dynamaid also gazes at tracked people by using its pan-tilt neck. Turning towards interaction partners is distributed over three levels [4]: the eyes, the neck, and the trunk. While the eyes are allowed to move quickly, the neck moves slower, and the trunk follows with the slowest time constant.

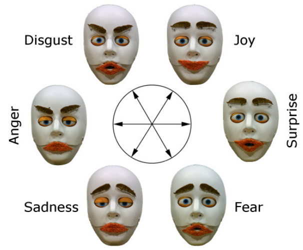
To communicate the robot's mood, we use a face with animated mouth and eyebrows to display facial expressions. The robot's mood is computed in a two-dimensional space, using six basic emotional expressions (joy, surprise, fear, sadness, anger, and disgust). Here, we follow the notion of the Emotion Disc [5]. This technique allows continuous changes of the facial expression. Fig. 2b shows how we implemented the six basic expressions for Robotinho. In combination with facial expressions, we modulate the parameters of our speech synthesis system to express the robot's mood.

III. Tour Guide

To gain more experience about the interaction of unexperienced users with our robots, we are currently evaluating our



(a) Our robots perform several human-like gestures, e.g., a greeting gesture (left figure), single- and both-handed come-closer gestures and inquiring gestures (right figure).



(b) To communicate the robot's mood, Robotinho is equipped with an expressive head. Emotions are computed using six basic expressions.

Fig. 2: The communication with humans is supported by gestures and facial expressions.

system in a museum tour guide scenario. In addition to natural interaction with spectators during the explanation of exhibits through gestures, mimic, and speech, we are convinced that interaction with people 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 alternating into the driving direction and to his followers. Looking into the driving direction makes clear that it notices people in front of it. Looking into the direction of the guided visitors is both necessary to update the robots knowledge about their positions and to show interest in the people it interacts with. If the positions of people are known, Robotinho gives its attention to a random visitor. Otherwise, it looks over its shoulder to find its followers.

If Robotinho is uncertain about the whereabouts of its spectators, 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 spectators. If this is successful, our robot indicates that it became aware of the presence of the spectators and continues the tour.

At an exhibit, Robotinho turns towards the spectators. The pointing gestures are parameterized such that Robotinho can perform the correct gestures independent of the pose relative to an exhibit.

IV. System Evaluation

Benchmarking robotic systems is difficult, because it is not easy to reproduce results. In recent years, robot competitions, such as the DARPA Grand and Urban Challenges and RoboCup, play an important role in assessing the performance of robot systems, because they provide a common test ground.

The RoboCup competitions include the @Home league, where robots have to perform service tasks in a domestic environment. Natural interaction with humans is demanded in all tests. Our Team Nimbro [6] participated 2009 for the first time in the @Home league. We used our communication and service robots in the *Introduce* task. Robotinho introduced itself, Dynamaid, and the team members to the audience and interacted with humans in a natural way to demonstrate its human-robot-interaction skills. Both at RoboCup German Open 2009 and at RoboCup 2009, the juries awarded our robots the highest score of all robots in this test.

In the *Who-is-Who* test a robot has to detect three persons, approach them, ask for their names, remember their faces and recognize them again when they leave the apartment. Dynamaid was awarded the second highest score in this test at RoboCup 2009. Both robots together were awarded the second highest score in the *Open Challenge*, where Robotinho gave a home tour to a guest while Dynamaid delivered a drink. Videos of our RoboCup performance in these, and the tests not mentioned here for brevity, are available on our website[1].

V. Conclusions and Future Work

While much research is performed in the field of domestic service robotics, most research groups focus on improving mobility and manipulation. We are convinced that multimodal communication between service robots and their human users will be key to the success in this application domain.

We evaluated our systems at the RoboCup German Open 2009 competition, where we reached the second place. At the RoboCup 2009 world championship we reached the third place and won the innovation award for "Innovative robot body design, empathic behaviors, and robot-robot cooperation". The communication skills of our robots were well received by the high-profile juries.

In the near future, we plan to equip our robot Dynamaid with an expressive communication head as well and to integrate gesture-recognition [7] and head-posture-estimation systems [8] that we developed outside our robotic platform.

Acknowledgment

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

References

[1] F. Faber, M. Bennewitz, C. Eppner, A. Görög, C. Gonsior, D. Joho, M. Schreiber, and S. Behnke, "The humanoid museum tour guide Robotinho," in *Proc. of 9th IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids), Paris*, 2009.

[2] J. Stückler, M. Schreiber, and S. Behnke, "Dynamaid, an anthropomorphic robot for research on domestic service applications," in *Proc. of the 4th European Conference on Mobile Robots (ECMR)*, 2009, pp. 87–92.

[3] M. Bennewitz, F. Faber, D. Joho, S. Schreiber, and S. Behnke, "Integrating vision and speech for conversations with multiple persons," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2005.

[4] F. Faber, M. Bennewitz, and S. Behnke, "Controlling the gaze direction of a humanoid robot with redundant joints," in *Proc. 16th Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2008.

[5] Z. Ruttkay, H. Noot, and P. ten Hagen, "Emotion Disc and Emotion Squares: Tools to explore the facial expression space," *Computer Graphics Forum*, vol. 22, no. 1, pp. 49–53, 2003.

[6] S. Behnke, J. Stückler, and M. Schreiber, "NimbRo @Home 2009 team description," in *RoboCup 2009 @Home League Team Descriptions, Graz, Austria*, 2009.

[7] T. Vatahska, M. Bennewitz, and S. Behnke, "Feature-based head pose estimation from images," in *Proc. of the IEEE-RAS Int. Conference on Humanoid Robots (Humanoids)*, Pittsburgh, USA, 2007.

[8] T. Axenbeck, M. Bennewitz, S. Behnke, and W. Burgard, "Recognizing complex, parameterized gestures from monocular image sequences," in *Proc. of IEEE/RSJ Int. Conf. on Humanoid Robots (Humanoids)*, 2008.

[1] http://www.NimbRo.net/@Home