# Intuitive Multimodal Interaction with Communication Robot Fritz

Maren Bennewitz, Felix Faber, Dominik Joho and Sven Behnke
*University of Freiburg, Computer Science Institute*
*Germany*

## 1. Introduction

One of the most important motivations for many humanoid robot projects is that robots with a human-like body and human-like senses could in principle be capable of intuitive multimodal communication with people. The general idea is that by mimicking the way humans interact with each other, it will be possible to transfer the efficient and robust communication strategies that humans use in their interactions to the man-machine interface. This includes the use of multiple modalities, such as speech, facial expressions, gestures, body language, etc. If successful, this approach yields a user interface that leverages the evolution of human communication and that is intuitive to naïve users, as they have practiced it since early childhood.

We work towards intuitive multimodal communication in the domain of a museum guide robot. This application requires interacting with multiple unknown persons. The testing of communication robots in science museums and on science fairs is popular, because the robots encounter there many new interaction partners, which have a general interest in science and technology. Here, we present the humanoid communication robot Fritz that we developed as successor to the communication robot Alpha (Bennewitz et al., 2005). Fritz uses speech, facial expressions, eye-gaze, and gestures to interact with people. Depending on the audio-visual input, our robot shifts its attention between different persons in order to involve them into an interaction. He performs human-like arm gestures during the conversation and also uses pointing gestures generated with eyes, its head, and arms to direct the attention of its communication partners towards the explained exhibits. To express its emotional state, the robot generates facial expressions and adapts the speech synthesis.

The remainder of the chapter is organized as follows. The next section reviews some of the related work. The mechanical and electrical design of Fritz is covered in Sec. 3. Sec. 4 details the perception of the human communication partners. Sec. 5 explains the robot's attentional system. The generation of arm gestures and of facial expressions is presented in Sec. 6 and 7, respectively. Finally, in the experimental section, we discuss experiences made during public demonstrations of our robot.

## 2. Related Work

Many research groups world-wide work on intuitive multimodal communication between humanoid robots and humans. Some example projects are the Leonardo robot at MIT



Source: Humanoid Robots: Human-like Machines, Book edited by: Matthias Hackel
ISBN 978-3-902613-07-3, pp. 642, Itech, Vienna, Austria, June 2007

(Breazeal et al., 2004), Repliee Q2 at Osaka University (Matsui et al., 2005), and BARTHOC at Bielefeld University (Spexard et al., 2006).

Several systems exist that use different types of perception to sense and track people during an interaction and that use a strategy to decide which person gets the attention of the robot.

Lang et al. apply an attention system in which only the person that is currently speaking is the person of interest (Lang et al., 2003). While the robot is focusing on this person, it does not look to another person to involve it into the conversation. Only if the speaking person stops talking for more than two seconds, the robot will show attention to another person.

Okuno et al. also follow the strategy to focus the attention on the person who is speaking (Okuno et al., 2002). They apply two different modes. In the first mode, the robot always turns to a new speaker, and in the second mode, the robot keeps its attention exclusively on one conversational partner. The system developed by Matsusaka et al. is able to determine the one who is being addressed to in the conversation (Matsusaka et al., 2001). Compared to our application scenario (museum guide), in which the robot is assumed to be the main speaker or actively involved in a conversation, in their scenario the robot acts as an observer. It looks at the person who is speaking and decides when to contribute to a conversation between two people.

The model developed by Thorisson focuses on turn-taking in one-to-one conversations (Thorisson, 2002). This model has been applied to a virtual character. Since we focus on how to decide which person in the surroundings of the robot gets its focus of attention, a combination of both techniques is possible.

In the following, we summarize the approaches to human-like interaction behavior of previous museum tour-guide projects. Bischoff and Graefe presented a robotic system with a humanoid torso that is able to interact with people using its arms (Bischoff & Graefe, 2004). This robot also acted as a museum tour-guide. However, the robot does not distinguish between different persons and does not have an animated face. Several (non-humanoid) museum tour-guide robots that make use of facial expressions to show emotions have already been developed. Schulte et al. used four basic moods for a museum tour-guide robot to show the robot's emotional state during traveling (Schulte, et al., 1999). They defined a simple finite state machine to switch between the different moods, depending on how long people were blocking the robot's way. Their aim was to enhance the robot's believability during navigation in order to achieve the intended goals. Similarly, Nourbakhsh et al. designed a fuzzy state machine with five moods for a robotic tour-guide (Nourbakhsh et al., 1999). Transitions in this state machine occur depending on external events, such as people standing in the robot's way. Their intention was to achieve a better interaction between the users and the robot. Mayor et al. used a face with two eyes, eyelids and eyebrows (but no mouth) to express the robot's mood using seven basic expressions (Mayor et al., 2002). The robot's internal state is affected by several events during a tour (e.g., a blocked path or no interest in the robot).

Most of the existing approaches do not allow continuous changes of facial expression. Our approach, in contrast, uses a bilinear interpolation technique in a two-dimensional state space (Ruttkay et al., 2003) to smoothly change the robot's facial expression.

## 3. Mechanical and Electrical Design of the Robot Fritz

Our humanoid robot Fritz has originally been designed for playing soccer in the RoboCup Humanoid League TeenSize class (Behnke et al., 2006). The left part of Fig. 1 shows him as goalie in the TeenSize Penalty Kick final at RoboCup 2006. Fritz is 120cm tall and has a total weight of about 6.5kg. His body has 16 degrees of freedom (DOF): Each leg is driven by five large digitally controlled Tonegawa PS-050 servos and each arm is driven by three digital Futaba S9152 servos.


Body Head

Figure 1. Humanoid robot Fritz. Left: Goalie at RoboCup 2006. Right: Communication head

For the use as communication robot, we equipped Fritz with a 16DOF head, shown in the right part of Fig. 1. The head is mounted on a 3DOF neck. The eyes are USB cameras that can be moved together in pitch and independently in yaw direction. Six servo motors animate the mouth and four servos animate the eyebrows.

The servo motors are controlled by a total of four ChipS12 microcontroller boards, which are connected via RS-232 to a main computer. We use a standard PC as main computer. It runs computer vision, speech recognition/synthesis, and behavior control.

## 4. Perception of Communication Partners

To detect and track people in the environment of our robot, we use the two cameras and a stereo microphone. In order to keep track of persons even when they are temporarily outside the robot's field of view, the robot maintains a probabilistic belief about the people in its surroundings.

### 4.1 Visual Detection and Tracking of People

Our face detection system is based on the AdaBoost algorithm and uses a boosted cascade of Haar-like features (Viola & Jones, 2004). Whenever a new observation is made it must be determined to which person, that has already been detected by the robot, the newly detected face belongs. To solve this data association problem, we apply the Hungarian Method using a distance-based cost function. We use a Kalman filter to track the position of a face over time. Fig. 2 shows three snapshots during face tracking. As indicated by the differently colored boxes, all faces are tracked correctly.

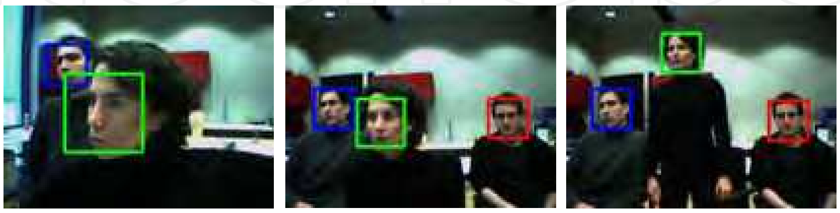


Figure 2. Tracking three faces

To account for false classifications of face/non-face regions and association failures, we apply a probabilistic technique. We use a recursive Bayesian update scheme (Moravec & Elfes, 1985) to compute the existence probability of a face. In this way, the robot can also estimate whether a person outside the current field of view is still there. When a person's probability falls below a threshold, the robot makes a reconfirming gaze to this person.

### 4.2 Speaker Localization

In addition to the visual perception of the persons around the robot, we implemented a speaker localization system that uses a stereo microphone. We apply the Cross-Power Spectrum Phase Analysis (Giuliani et al., 1994) to calculate the spectral correlation measure between the left and the right microphone channel. Using the corresponding delay, the relative angle between the speaker and the microphones can be calculated.
The person in the robot's belief that has the minimum distance to the sound source angle gets assigned the information that it has spoken. If the angular distance between the speaker and all persons is greater than a certain threshold, we assume the speaker to be a new person, who just entered the scene.

## 5. Attentional System

It is not human-like to fixate a single conversational partner all the time when there are other people around. Fritz shows interest in different persons in his vicinity and shifts his attention between them so that they feel involved into the conversation. We currently use three different concepts in order to change the robot's gaze direction.

### 5.1 Focus of Attention

In order to determine the focus of attention of the robot, we compute an importance value for each person in the belief. It currently depends on the time when the person has last spoken, on the distance of the person to the robot (estimated using the size of the bounding box of its face), and on its position relative to the front of the robot. The resulting importance value is a weighted sum of these three factors. In the future, we plan to consider further aspects to determine the importance of persons like, for example, waving with hands.
The robot focuses its attention always on the person who has the highest importance, which means that it keeps eye-contact with this person. Of course, the focus of attention can change during a conversation with several persons.

### 5.2 Attentiveness to a Speaker

If a person that is outside the current field of view, which has not been detected so far, starts to speak, the robot reacts to this by turning towards the corresponding direction. In this way, the robot shows attentiveness and also updates its belief about the people in its surrounding.

### 5.3 Gazes outside the Focus of Attention

Since the field of view of the robot is constrained, it is important that the robot changes its gaze direction to explore the environment and to update its belief about it. Our robot regularly changes its gaze direction and looks in the direction of other faces, not only to the most important one. This reconfirms that the people outside the field of view are still there and involves them into the conversation.

### 5.4 Example

Fig. 3 illustrates an experiment that was designed to show how the robot shifts its attention from one person to another if it considers the second one to be more important. In the situation depicted here, person 2 was talking to the robot. Since person 2 had the highest importance, the robot initially focused its attention on person 2 but also looked to person 1 at time steps 10 and 21, to signal awareness and to involve him/her into the conversation. When looking to person 1 at time step 21, the robot then noticed that this person had come very close. This yielded a higher importance value for this person and the robot shifted its attention accordingly.

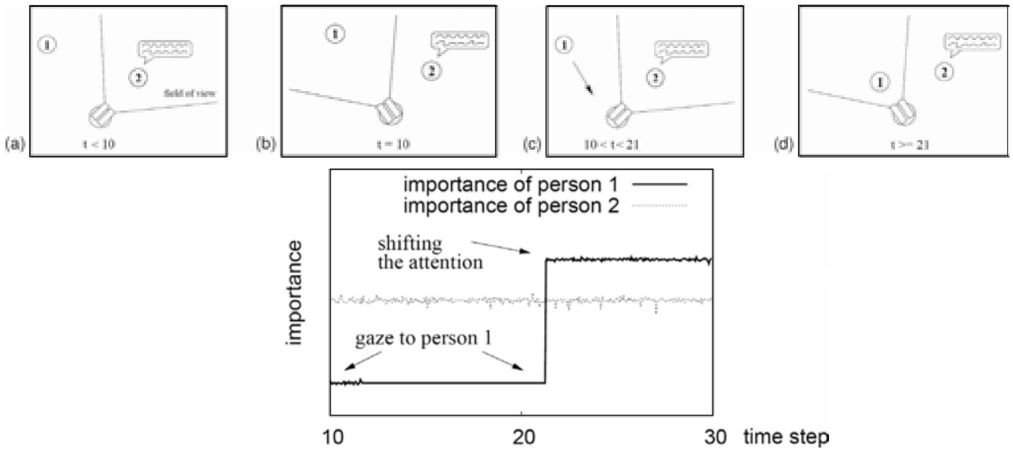



Figure 3. The images (a) to (d) illustrate the setup in this experiment. The lower image shows the evolution of the importance values of two people. See text for a detailed explanation

## 6. Arm and Head Gestures

Our robot uses arm and head movements to generate gestures, and to appear livelier. The gestures are generated online. Arm gestures consist of a preparation phase, where the arm moves slowly to a starting position, the stroke phase that carries the linguistic meaning, and a retraction phase, where the hand moves back to a resting position (MacNeill, 1992). The stroke is synchronized to the speech synthesis module.

### 6.1 Symbolic Gestures

Greeting  Inquiring 

Figure 4. Fritz performing two symbolic gestures with its arms

Symbolic gestures are gestures in which the relation between form and content is based on social convention. They are culture-specific.

- **Greeting Gesture:** The robot performs a single-handed gesture while saying hello to newly detected people. As shown in the left part of Fig. 4 it raises its hand, stops, and lowers it again.
- **Come Closer Gesture:** When the robot has detected persons farther away than the normal conversation distance (1.5-2.5m), he animates the people to come closer. Fig. 5 shows that the robot moves both hands towards the people in the preparation phase and towards its chest during the stroke.
- **Inquiring Gesture:** While asking certain questions, the robot performs an accompanying gesture, shown in the right part of Fig. 4. It moves both elbows outwards to the back.
- **Disappointment Gesture:** When the robot is disappointed (i.e., because it did not get an answer to a question), it carries out a gesture to emphasize its emotional state. During the stroke it moves both hands quickly down.
- **Head Gestures:** To confirm or disagree, the robot nods or shakes its head, respectively.

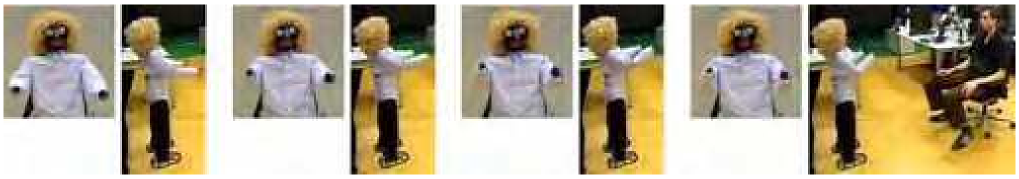


Figure 5. Fritz asks a person to come closer

### 6.2 Batonic Gestures

Humans continuously gesticulate to emphasize their utterances while talking to each other. Fritz also makes small emphasizing gestures with both arms when he is speaking longer sentences.

### 6.3 Pointing Gestures

To draw the attention of communication partners towards objects of interest, our robot performs pointing gestures. While designing the pointing gesture for our robot, we followed the observation made by Nickel et al. that people usually move the arm in such a way that, in the poststroke hold, the hand is in one line with the head and the object of interest (Nickel et al., 2004). This is illustrated in Fig. 6.

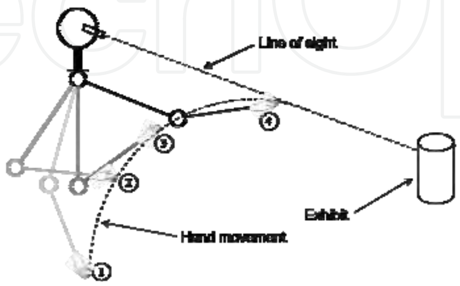



Figure 6. Side view of the arm movement during a pointing gesture

When the robot wants to draw the attention to an object, it simultaneously moves the head and the eyes in the corresponding direction and points in the direction with the respective arm while uttering the object name.

### 6.4 Non-Gestural Arm Movements

While standing, people typically move unconsciously with their arms and do not keep completely still. Our robot also performs such minuscule movements with its arms. The arms move slowly, with low amplitude in randomized oscillations.

## 7. Expression of Emotions

Showing emotions plays an important role in inter-human communication. During an interaction, the perception of the mood of the conversational partner helps to interpret his/her behavior and to infer intention. To communicate the robot's mood we use a face with animated mouth and eyebrows that displays facial expressions and also synthesize speech according to the current mood. 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 (Ruttkay et al., 2003).

### 7.1 Facial Expressions

Fig. 7 shows the six basic facial expressions of our robot. As parameters for an expression we use the height of the mouth corners, the mouth width, the mouth opening angle, and the angle and height of the eye-brows.
To influence the emotional state of our robot, we use behaviors that react to certain events. For example, if no one is interested in the robot, it is getting more and more sad, if someone then talks to it, the robot's mood changes to a mixture of surprise and happiness. Each behavior submits its request in which direction and with which intensity it wants to change the robot's emotional state. After all behaviors submitted their requests, the resulting vector is computed by the sum of the individual requests. We allow any movement within the circle described by the Emotion Disc.

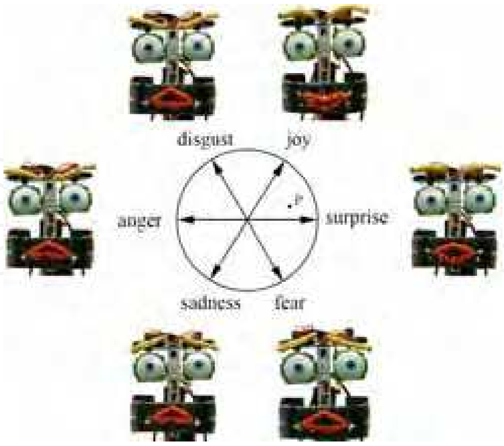



Figure 7. The two-dimensional space in which we compute the robot's emotional state. The images show the six basic facial expressions

The parameters $P'$ for the facial expression corresponding to a certain point $P$ in the two-dimensional space are calculated by linear interpolation between the parameters $E'_i$ and $E'_{i+1}$, the adjacent basic expressions:

$$P' = l(p) \cdot \big( \alpha(p) \cdot E'_i + (1 - \alpha(p)) \cdot E'_{i+1} \big). \quad (1)$$

Here, $l(p)$ is the length of the vector $p$ that leads from the origin (corresponding to the neutral expression) to $P$, and $\alpha(p)$ denotes the normalized angular distance between $p$ and the vectors corresponding to the two neighboring basic expressions. This technique allows continuous changes of the facial expression.

### 7.2 Emotional Speech Synthesis

In combination with facial expressions, we use emotional speech to express the robot's mood. Most speech synthesis systems do not support emotional speech directly; neither does the Loquendo TTS system that we use. However, in this system, we can influence the parameters average pitch, speed, and volume and thereby express emotional speech.
Cahn proposed a mapping of emotional states to the relative change of several parameters of a speech synthesis system (Cahn, 1989). She carried out experiments to show that test persons were able to recognize the emotion category of several synthesized sample sentences. In the mapping, she used the same six basic emotions that constitute the axes of the Emotion Disc. We use her mapping for the parameters average pitch, speech rate and loudness to set the parameters average pitch, speed and volume of our speech synthesizer.
The mapping of emotional states to the relative change of the speech parameters can be seen in Tab. 1. Let **M** be such a mapping matrix, and **e** be an emotion intensity vector of the six basic emotions. Then we can compute the three speech parameters as a vector **s**, as follows:

$$\mathbf{s} = \mathbf{d} + \mathbf{S}\,\mathbf{M}\,\mathbf{e}. \quad (2)$$

The three-element vector **d** contains the default values for the parameters and **S** is a diagonal matrix used to scale the result of the mapping, thereby allowing for an adaptation of the mapping to the characteristics of the synthesizer system. The emotion intensity vector contains only two non-zero entries, $l(p)\cdot\alpha(p)$ and $l(p)\cdot(1 - \alpha(p))$, that correspond to the influence factors of the two adjacent basic expressions of the current mood (see Fig. 7 and Eq. 1).

| | joy | surprise | fear | sadness | anger | disgust |
|---|---|---|---|---|---|---|
| pitch | -3 | 0 | 10 | 0 | -5 | 0 |
| speed | 2 | 4 | 10 | -10 | 8 | -3 |
| volume | 0 | 5 | 10 | -5 | 10 | 0 |

Table 1. Mapping of emotions to the relative change of the speech parameters (**M** in Eq. 2)

Emotions influence many more characteristics of speech, e.g. breathiness, precision of articulation, and hesitation pauses. Hence, the three parameters used in our system can only roughly approximate emotional speech. In spite of these limitations, we experienced that even such simple adjustments can, in conjunction with facial expressions, contribute to the emotional expressiveness.

## 8. Public Demonstrations

To evaluate our system, we tested our communication robots Alpha and Fritz in two public demonstrations. In this section, we report the experiences we made during these exhibitions. We chose a scenario in which the robot presents four of its robotic friends. We placed the exhibits on a table in front of the robot. Our communication robot interacted multimodally with the people and had simple conversations with them. For speech recognition and speech synthesis, we used the Loquendo software. Our dialog system is realized as a finite state machine. Fig. 8 illustrates a simple version. With each state, a different grammar of phrases is associated that the recognition system should be able to recognize. The dialog system generates some small talk and allows the user to select which exhibits should be explained and to what level of detail.

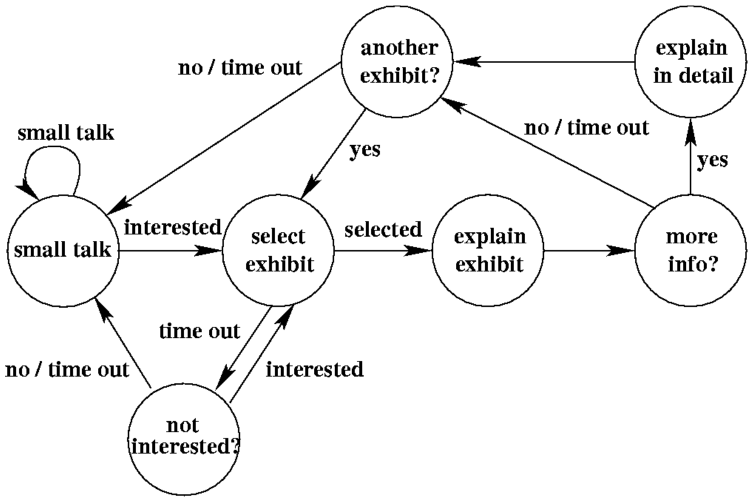



Figure. 8. Finite state machine controlling the flow of the dialog

### 8.1 Two-Day Demonstration at the Science Fair 2005 in Freiburg

The first demonstration was made using the robot Alpha, the predecessor of Fritz. We exhibited Alpha during a two-day science fair of Freiburg University in June 2005. In contrast to Fritz, Alpha did not use emotional speech and performed pointing gestures with his arms but not any other human-like gestures.

At the science fair, we asked the people who interacted with the robot to fill out questionnaires about their interaction-experiences with Alpha (see (Bennewitz et al., 2005) for more details). Almost all people found the eye-gazes, gestures, and the facial expression human-like and felt that Alpha was aware of them. The people were mostly attracted and impressed by the vivid human-like eye movements. To evaluate the expressiveness of the pointing gestures, we carried out an experiment in which the people had to guess the target of the pointing gestures. The result was that 91% of the gestures were correctly interpreted.

However, one limitation that was obvious is that speech recognition does not work sufficiently well in noisy environments, even when using close-talking microphones. To account for this problem, in our current system, the robot asks for an affirmation when the speech recognition system is not sure about the recognized phrase.

### 8.2 Three-Day Demonstration at the Science Days 2006 in the Europapark Rust

Figure 9. Fritz presenting its robot friends to visitors at the Science Days

In October 2006, we exhibited Fritz for three days at the Science Days in the Europapark Rust (see Fig. 9). Since the people at the previous exhibition were most attracted by the human-like behavior, we augmented the number of arm gestures as explained in Sec. 6. In general, the gestures served their purpose. However, the *come closer* gesture did not always have the desired result. In the beginning of the interaction, some people were still too shy and barely wanted to come closer to the robot. This effect is not uncommon even for human museum guides starting a tour. As soon as the visitors became more familiar with the robot, their shyness vanished and they choose a suitable interaction distance by themselves.

In contrast to the exhibition of Alpha, where toddlers often were afraid of the robot and hid behind their parents, we did not observe such a behavior with Fritz. This is probably due to the different sizes and appearances of the robots. The kids found Fritz apparently very exciting. Most of them interacted several times with the robot. At the end, some of them knew exactly what the robot was able to do and had fun in communicating with Fritz.

When there were people around Fritz but nobody started to talk to the robot, Fritz proactively explained to the people what he is able to do. While speaking, he performed gestures with his head and arms so that after the explanation the people had a good idea about the capabilities of the robot. This idea resulted from lessons learned of the first exhibition where people often did not know about what the robot is actually able to do and what not.

Due to the severe acoustical conditions, speech recognition did not always work well. The affirmation request helped only if the correct phrase was the most likely one. Hence, for the next exhibition, we plan to employ an auditory front-end that focuses on the fundamental frequency of the speaker, in order to separate it from background noise.

A video of the demonstration can be downloaded from http://www.NimbRo.net.

## 9. Conclusions

In this chapter, we presented our humanoid communication robot Fritz. Fritz communicates in an intuitive, multimodal way. He employs speech, an animated face, eye-gaze, and gestures to interact with people. Depending on the audio-visual input, our robot shifts its attention between different communication partners in order to involve them into an interaction. Fritz performs human-like arm and head gestures, which are synchronized to the speech synthesis.

He generates pointing gestures with its head, eyes, and arms to direct the attention of its communication partners towards objects of interest. Fritz changes its emotional state according to the number of people around him and the dialog state. Its emotional state is communicated by facial expressions and emotional speech synthesis. We tested the described multimodal dialog system during two public demonstrations outside our lab. The experiences made indicate that the users enjoyed interacting with the robot. They treated him as an able communication partner, which was sometimes difficult, as its capabilities are limited.

The experienced problems were mainly due to perception deficits of the robot. While speech synthesis works fairly well, robust speech recognition in noisy environments is difficult. This is problematic, because the users expect the robot to understand speech at least as well as it talks. Similarly, while the robot is able to generate gestures and emotional facial expressions, its visual perception of the persons around it is limited to head position and size. To reduce this asymmetry between action generation and perception, we currently work on head posture estimation from the camera images and on the visual recognition of gestures.

## 10. Acknowledgment

Funding for this project is provided by the DFG (Deutsche Forschungsgemeinschaft), grant BE 2556/2-1,2.

## 11. References

Behnke, S.; Schreiber, M.; Stückler, J.; Strasdat, H. & Bennewitz M. (2006). NimbRo TeenSize 2006 team description, In: *RoboCup 2006 Humanoid League Team Descriptions*, Bremen, Germany

Bennewitz, M.; Faber, F.; Joho, D.; Schreiber, M. & Behnke, S. (2005). Towards a humanoid museum guide robot that interacts with multiple persons, *Proceedings of the IEEE/RSJ International Conference on Humanoid Robots (Humanoids)*, pp. 418-423, Tsukuba, Japan

Bischoff, R. & Graefe, V. (2004). Hermes – a versatile personal robot assistant. *Proc. IEEE – Special Issue on Human Interactive Robots for Psychological Enrichment*, 92(11):1759-1779

Breazeal, C; Brooks, A.; Gray, J.; Hoffman, G.; Kidd, C.; Lee, H.; Lieberman, J.; Lockerd, A. & Mulanda, D. (2004). Humanoid robots as cooperative partners for people. *International Journal of Humanoid Robots*,1(2):1-34

Cahn, J. (1989). *Generating expression in synthesized speech*, Master thesis, Massachusetts Institute of Technology

Giuliani, D.; Omologo, M. & Svaizer P. (1994). Talker localization and speech recognition using a microphone array and a cross-powerspectrum phase analysis, *Proceedings of International Conference on Spoken Language Processing (ICSLP)*, pp. 1243-1246, Yokohama, Japan

Lang, S.; Kleinehagenbrock, M.; Hohenner, S.; Fritsch, J.; Fink, G. & Sagerer, G. (2003). Providing the basis for human-robot-interaction: A multi-modal attention system for a mobile robot, *Proceedings of IEEE International Conference on Multimodal Interfaces*, pp. 28-35, Vancouver, Canada

MacNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*, University of Chicago Press

Matsui, D.; Minato, T.; MacDorman, K. F. & Ishiguro, H. (2005). Generating natural motion in an android by mapping human motion, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1089-1096, Edmonton, Canada

Matsusaka, Y.; Fujie, S. & Kobayashi, T. (2001). Modeling of conversational strategy for the robot participating in the group conversation, *Proc. of the European Conference on Speech Communication and Technology (Eurospeech)*, pp. 2173-2176, Aalborg, Denmark

Mayor, L.; Jensen, B.; Lorotte, A. & Siegwart, R. (2002). Improving the expressiveness of mobile robots, *Proceedings of IEEE International Workshop on Robot and Human Interactive Communication (ROMAN)*, pp. 325-330, Berlin, Germany

Moravec, H. & Elfes, A. (1985). High resolution maps from wide angle sonar, *Proceedings of IEEE International Conference on Robotics & Automation (ICRA)*, pp. 116-121, St. Louis

Nickel, K.; Seemann, E. & Stiefelhagen, R. (2004). 3D-Tracking of heads and hands for pointing gesture recognition in a human-robot interaction scenario, *Proceedings of International Conference on Face and Gesture Recognition (FGR)*, pp. 565-570, Seoul, Korea

Nourbakhsh, I.; Bobenage, J.; Grange, S.; Lutz, R.; Meyer, R. & Soto, A. (1999). An affective mobile robot educator with a full-time job. *Artificial Intelligence,* 114(1-2):95-124

Okuno, H.; Nakadai, K. & Kitano, H. (2002). Social interaction of humanoid robot based on audio-visual tracking, *Proceedings of IEEE International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE)*, pp. 725-735, Cairns, Australia

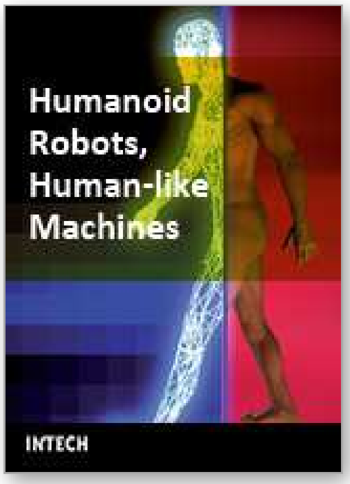
Ruttkay, Z.; Noot, H. & Hagen, P. (2003). Emotion Disc and Emotion Squares: Tools to explore the facial expression space. *Computer Graphics Forum,* 22(1):49-53

Schulte, J.; Rosenberg, C. & Thrun, S. (1999). Spontaneous short-term interaction with mobile robots in public places, *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 658-663, Detroit, Michigan

Spexard, T.; Haasch, A.; Fritsch, J. & Sagerer G. (2006). Human-like person tracking with an anthropomorphic robot, *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1286-1292, Orlando, Florida

Thorisson, K. R. (2002). Natural turn-taking needs no manual: Computational theory and model, from perception to action, In: *Multimodality in Language and Speech Systems*, pp. 173-207, Kluwer Academic Publishers

Viola, P. & Jones, M.J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision,* 57(2):137-154




**Humanoid Robots, Human-like Machines**
Edited by Matthias Hackel

ISBN 978-3-902613-07-3
Hard cover, 642 pages
**Publisher** I-Tech Education and Publishing
**Published online** 01, June, 2007
**Published in print edition** June, 2007

In this book the variety of humanoid robotic research can be obtained. This book is divided in four parts: Hardware Development: Components and Systems, Biped Motion: Walking, Running and Self-orientation, Sensing the Environment: Acquisition, Data Processing and Control and Mind Organisation: Learning and Interaction. The first part of the book deals with remarkable hardware developments, whereby complete humanoid robotic systems are as well described as partial solutions. In the second part diverse results around the biped motion of humanoid robots are presented. The autonomous, efficient and adaptive two-legged walking is one of the main challenge in humanoid robotics. The two-legged walking will enable humanoid robots to enter our environment without rearrangement. Developments in the field of visual sensors, data acquisition, processing and control are to be observed in third part of the book. In the fourth part some "mind building" and communication technologies are presented.