Proceedings of the 2006 IEEE/RSJ
International Conference on Intelligent Robots and Systems
October 9 - 15, 2006, Beijing, China

Design and Evaluation of Emotion-Display EDDIE

Stefan Sosnowski, Ansgar Bittermann, Kolja Kühnlenz, and Martin Buss
Institute of Automatic Control Engineering (LSR)
Technische Universität München
D-80290 München, Germany
Email: sosnowski@mytum.de, bittermann@tum.de, kolja.kuehnlenz@ieee.org, m.buss@ieee.org

***Abstract*— This paper focuses on the development of EDDIE, a flexible low-cost emotion-display with 23 degrees of freedom. Actuators are assigned to particular action units of the facial action coding system (FACS). Emotion states represented by the circumplex model of affect are mapped to individual action units. Thereby, continuous, dynamic, and realistic emotion state transitions are achieved. EDDIE is largely developed and manufactured in a rapid-prototyping process. Miniature off-the-shelf mechatronics components are used providing high functionality at low-cost. Evaluations conducted in a user-study show that emotions can be recognized very well. Further experiments show that additional features adapted from animals have significant but small influence on the display of the human emotion 'disgust'.**

I. Introduction

Robots are likely to become an important part in our daily lives. As they have been mainly specialized tools, designed for a single task without interactions with humans until now, current research opens new areas of application like care-taking or reception tasks. Due to this, a new generation of robots has to appeal to persons of various age, sex, income or education.

Interaction between humans is heavily influenced by non-verbal communication. Future systems in robotics will adapt such non-verbal forms of communication in order to realize an intuitive man-machine interaction. The benefits are higher efficiency, security and robustness, as well as easier adaption to, use of and teaching-process for new systems for beginners. Another important factor is social integrity [1].

In the context of the above mentioned research fields a variety of robotics heads mimicking human facial expressions have been developed. Prominent examples are Kismet [1], Leonardo [2], Aryan [3], Saya [4], and Roberta [5] with up to 32 degrees of freedom (DOF) (Leonardo) for displaying facial expressions. Used drives reach from dc servos (e.g. Kismet) over dc servos with spring coupling (e.g. Leonardo) to pneumatic or hydraulic cylinders (e.g. Roberta).

In this paper, EDDIE, a mechatronical emotion-display with 23 DOF is presented. In order to achieve dynamic, continuous, and realistic emotional state transitions, the two dimensional emotional state-space based on the circumplex model of affect is directly mapped to joint space corresponding to particular action units of the facial action coding system (FACS). The display is largely developed and manufactured in a rapid-prototyping process. Miniature off-the-shelf mechatronics are used providing high functionality while extremely low-cost. The display is evaluated in a user-study. The impact of the non-human elements ears and crown not covered by FACS on the recognized emotions is evaluated.

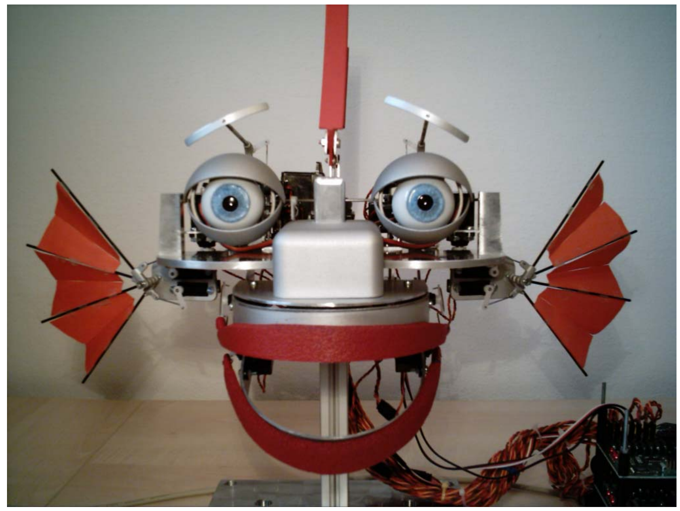


Fig. 1. Emotion-display EDDIE, emotional state: Surprised.

The paper is organized as follows: In Section 2 background and design criteria are given; in Section 3 the implementation of EDDIE is shown; the evaluation in a user-study is described in Section 4; conclusions are given in Section 5.

II. Related Work and Design Criteria

A. Modeling emotions

1) Discrete emotional states: The existence of universal emotions, being represented and interpreted equally in the whole world, is said to be assured. Fridlund, Ekman and Oster affirm in their literature research in 1987, that six basic emotions are clearly identified in a multitude of different cultural groups. Joy, surprise, fear, sadness, anger and disgust, therefore, are considered universal [7].

2) Two- and three-dimensional representation: One of the main models for emotions is the verbal semantic differential scale developed by Mehrabian and Russell [8]. Emotional states are represented within a three-dimensional space spanned by valence, arousal, and dominance. A simplification is the circumplex model of affect shown in the user interface in Figure 4. Influencing concepts of this model are located on a circle [9]. There are two components of influence: valence (pleasure), the horizontal dimension, and arousal (excitement), the vertical dimension. Every mood-describing expression can be defined according to its pleasure- and excitement-component [10]. However, neglecting dominance makes it more difficult to distinguish between certain emotional states. Russell's studies come to the conclusion, that emotional expressions have the mentioned structure in every culture [11]. This two-dimensional state space enables smooth and believable transitions between the emotional expressions.

B. How to display emotions

There exist many different communication channels for the expression of emotions, verbal by speech or non-verbal by a certain posture, gesture or facial expression. However, gesture and posture - in contrast to facial expression- show no specific emotions, but demonstrate the intensity of the actual emotion. Even by vocalization, man can not transfer such a diversity of emotional expressions, particularly as speech can be under better deliberate control than facial expressions [13].

1) Facial action coding system: In order to get an objective distinction among traceable changes in facial expressions, the facial action coding system (FACS) by Ekman and Friesen codes facial expressions based on face muscles [12].

For this purpose FACS defines smallest movable units, the so called action units. The limitation on visually distinguishable changes in faces leads to the definition of 44 action units [13]. The current construction covers 13 of the total 21 action units needed for the basic emotions. For an even better match, the integration of a flexible skin would be required.

2) Animal anatomy: Many species have a relatively low number of facial muscles compared to humans. To intensify certain emotional expressions most species have evolved a special gesture or possess body parts with additional signaling. EDDIE's crown is inspired by a cockatoo, whose most striking attributes are movable feathers on his head. The folding mechanism of the ears is inspired by a dragon lizard (Chlamydosaurus kingii).

C. Appearance

1) Virtual or real: Breazeal concludes that for a socially intelligent robot it is better to be situated in the real world. People are used to interact with embodied creatures and have evolved communication skills, which need a body for expression. They rely on having a reference point to refer their communication to and to know from where feedback comes from. Two advantages of a physically implemented robot compared to a computer animation are: for animated objects, an intelligent control is expected, whereas flat pictures seem to be the result of playback of stored sequences (TV, animated sequences in computer games). The other advantage is that a three-dimensional robot can be viewed from all sides, without a viewer having to stand directly in front of the robot.

2) Uncanny valley: Mori's theory of the uncanny valley states that as a robot increases in humanness, there is a point where the robot is not fully similar to humans but the balance between humanness and machine-likeness is uncomfortable for a user.

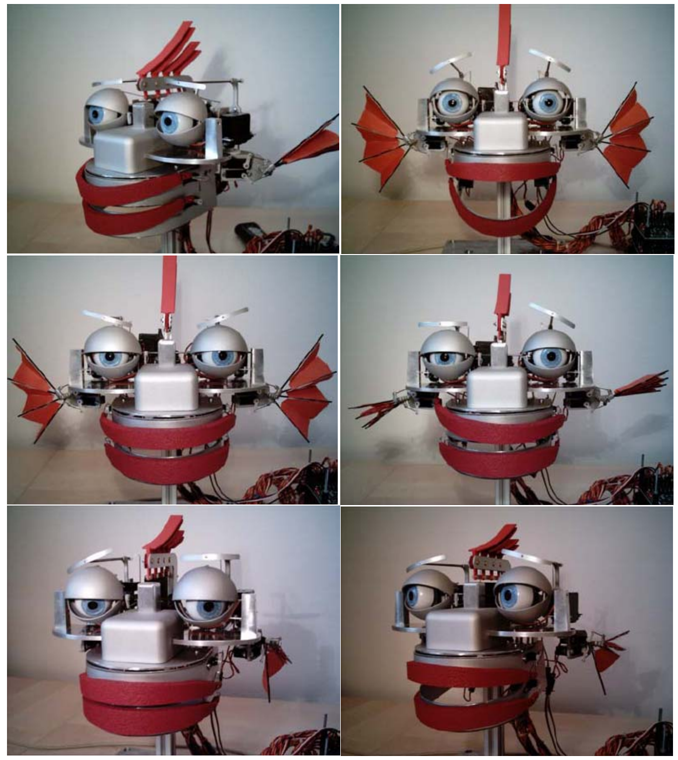


Fig. 2. EDDIE, emotional states from top left to lower right: Joy, surprise; anger, disgust; sadness, fear (See companion video, http://www.lsr.ei.tum.de/).

There is a reasonable degree of familiarity that should be achieved and maintained in humanoid robots [14]. Because of this theory, the current design does without elements covering the mechanical parts. The display should be clearly recognizable as a machine. This is necessary for a user not to have wrong expectations on the emotional abilities of the robot, but to perceive the possibilities a machine offers. A certain degree of humanness is also necessary for a user to feel comfortable acting socially with the robot. Finally, there has to remain a share in looking like a product, making it comfortable for people to use the robot [15].

3) Scheme of childlike characteristics: According to Konrad Lorenz (1935) the sight of children and certain young animals activates inborn reactions the viewer is not able to subdue. Due to this, Lorenz speaks of an inherent scheme. Thus, the sight of a child automatically activates sympathy and evokes a care motivation [16]. This scheme does not only work for children, but also for young animals and can be ported on machine-creatures, as the growing number of toys using this scheme shows. The sympathy effect for the robot is very important for the acceptance of a user, as it is lowering the inhibition threshold to interact with the robot. For the robot design, we try to establish a good balance between adult and childlike attributes.

4) The naming: Giving a name to a robot should not be underestimated. A good choice for a nickname, sounding nice and friendly, can enforce a sympathy effect given by the scheme of childlike characteristics and build a stronger emotional bond toward the machine-creature. Also a name has an anthropomorphic effect. If a machine has a name, an own personality is easier simulated. Due to this, the emotion display presented in this paper is named EDDIE, as an abbreviation for *Emotion Display with Dynamic Intuitive Expressions*.

D. Perception

1) Visual sensors: Vision based systems are able to identify faces, to measure head position and gaze and to read facial expressions [1]. To prepare the emotion display for an autonomous interaction with a user, firewire cameras are integrated into the eyes forming a stereo pair with a baseline of 12cm and 30fps at a resolution of 640x480 pixels.

2) Audio sensors: Speech recognition and speaker identifying are well investigated. Even the recognition of emotions in speech become interesting [1]. For this purpose we included a microphone located at the center of the head. Our next steps in development will be to teach EDDIE simple spoken commands.

III. Implementation

A. Mechanics

1) CAD Development: All mechanical parts are modeled in CATIA P3 V5R11 (Dassault Systems). This secures that all parts fit and are manufactured exactly, and the final appearance and function, the quality of displayable emotions and the overall picture, can easily be evaluated. CAD modeled freeform parts are manufactured by rapid-prototyping, the rest is crafted in aluminium. Aluminium as a basis material has some advantages: it is quite lightweight, in contrast to steel easy to work on and gives a high-quality finish. Moreover the machinelike appearance is achieved. The rapid-prototyping parts consist of PA3200GF, a glass reinforced plastic. The white plastic is coated in a metallic aluminium finish to conform it to the aluminium parts.

2) Motors: For actuation commercial miniature servo motors are chosen: Atom Mini (3Ncm, Jamara), C1081 (12Ncm, Graupner), C5191 (56Ncm, Graupner). Despite their strength, these are compact, low cost and easy to obtain. With their internal logic and the PWM-signal, they can easily be controlled.

3) Eyes: The eyes are based on two gimbal-mounted concentric rings. Both rings consist of PA3200GF. A miniature servo, mounted in a bracket at the outer ring, moves the inner ring in the horizontal of the outer ring. Both eyes are tilted vertical by a shared servo. The rotation axis of the outer rings is held in place by three aluminium holders. The camera is attached to the inner ring (see Figure 3). The eyes are each covered with a pair of eyelids, made of PA3200GF, each eyelid is separately moved by a servo. The lower eyelid is smaller than the upper, just like it is at the human eye (70° aperture angle of the spherical segment instead of 90°). Integration complexity is raised by the lower eyelids, but identification rate of facial expressions is increased in tests according to Breazeal [1].

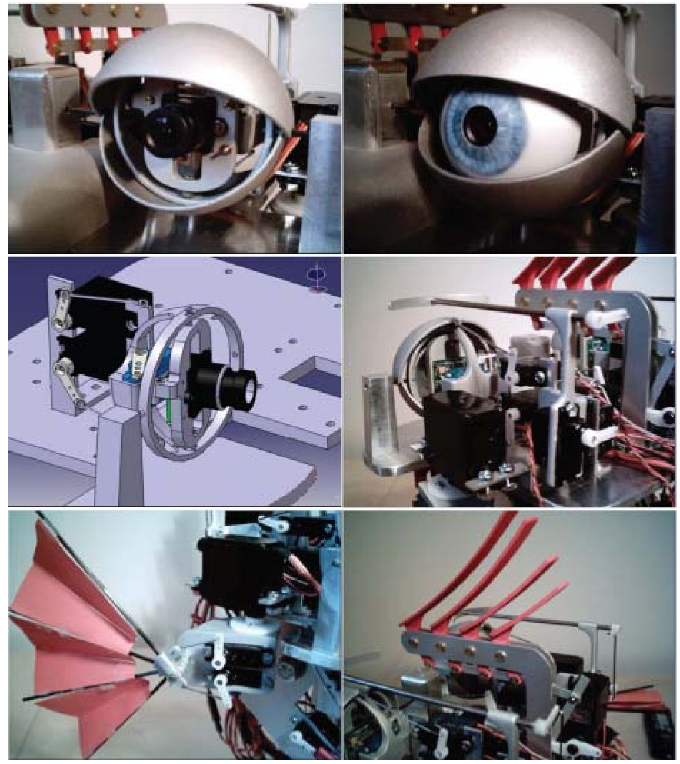


Fig. 3. Mechatronics from upper left to lower right: left eye; left eye with iris cover; CAD-drawing of right eye gimbal; eye-brow mechanics; ear mechanics; crown mechanics.

Camera covers are halves of table tennis balls, painted by airbrush with an iris including structure and the iris muscle. The pupil is a hole of 10mm diameter. Combination with the dark camera lens results in a natural looking eye.

4) eyebrows: The brows have two degrees of freedom each. A servo rotates a mounting in which a second servo is integrated (see Figure 3). The first servo is for up-down-movement, the second rotates a rod. At the end of the rod the virtual brow is mounted.

5) ears: The ears are composed of a folding and a rotating unit (see Figure 3). In the folding unit four rods, which are supported in a swivel joint and connected by a membrane, are moved by two servos. With it the ears can be tilted up or down or be streched completely. The rotating unit is moved by a single servo and can position the ears righted or parallel to the side of the head. The membrane is made of red paper segments.

6) Mouth and jaw: The jaw consists of a rigid upper jaw, similar to a human's upper jaw, and a separate lower jaw, which can be moved by a servo. Both jaw parts have semicircles as frontal areas, to appear more organic. In each jaw part two servos are integrated, which can draught the lips at the corner of the mouth up or down. These lips are made of foam, colored red. The red color increases the signal effect and the similarity to human lips. A test with blue lips showed a significantly lower effect on the whole appearance (subjective impression). Foam has as advantage that the lips get enough

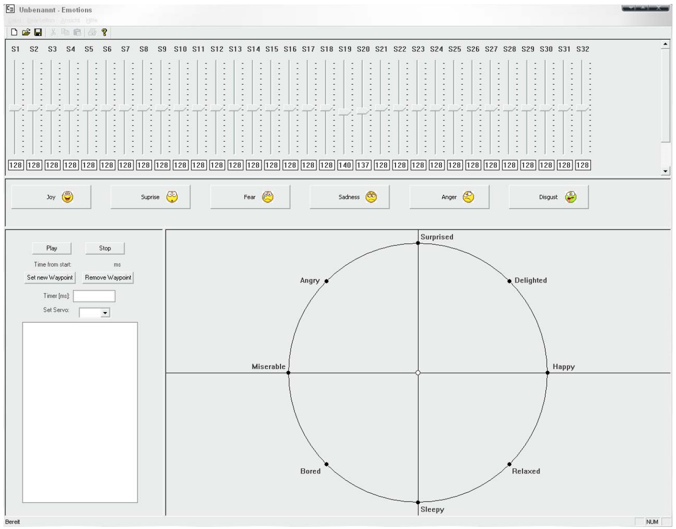


Fig. 4. Graphical user-interface.

TABLE I
DEGREES OF FREEDOM

parts	ears	eyes	eyelids	brows	jaw	lips	crown	total
DOF	6	3	4	4	1	4	1	23

volume and are flexible. Even the contour can be adapted to human lips.

7) crown: The crown is composed of a L-shaped holder, giving the necessary height above the rest of the head, a servo and four feathers. The feathers are pivoted and are moved simultaneously by a lever mechanism (see Figure 3). Hence, the same attack angle can be assured quite easily.

The feathers copy the feathers of a lesser sulphur-crested cockatoo (cacatua sulphurea sulphurea). Their lenght is arranged in a level that all ends are located above another when closed. The feathers can support facial expressions in their effect by being positioned either up or down. The red color brings out their signal effect.

8) Cost: The costs for the mechatronical system are approximately 2400 EUR.

B. Control unit

The basic system architecture is shown in Figure 7. Processing the input from the graphical user interface (GUI) (see Figure 4) the motor command generator generates appropriate desired positions for the servo motors based on heuristics mapping the emotional state-space to joint space corresponding to 13 action units. Example mappings are shown in Figure 5 and Figure 6 for the two servos of each ear. A microcontroller generates PWM signals from the motor commands transmitted via RS232 driving the servo motors. The motor current is evaluated by the microcontroller and the signal is fed back to the host PC.

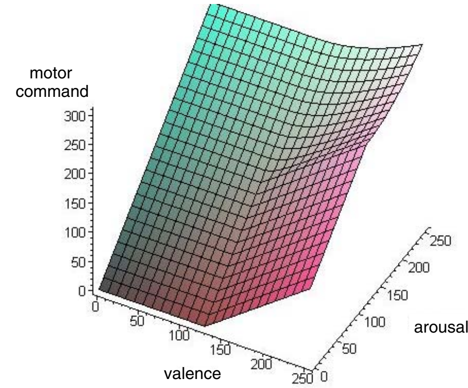


Fig. 5. Plot of motor command despos[upper_servo_ear_left_right].

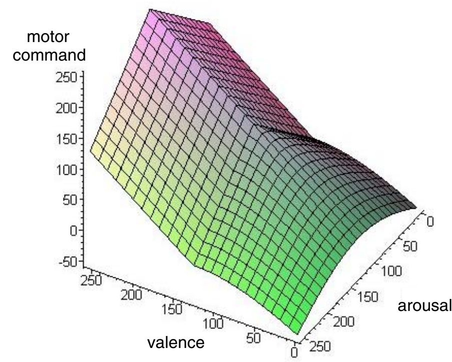


Fig. 6. Plot of motor command despos[lower_servo_ear_left_right].

C. Control software

The control software is programmed in Microsoft Visual-C++ and provides a user a graphical interface shown in Figure 4. The emotion display can be controlled in three ways. By a set of buttons for the discrete basic emotions, by the circumplex model of affect or by a control panel which sets each servo motor separately. Within the circular model, a pointer can be dragged continuously generating corresponding expressions and continuous transitions. On this, the reader may also refer to the companion video and to additional information presented at the authors' web-site (http://www.lsr.ei.tum.de/).

1) Control of separate servo movements: The control panel for the servos has 32 sliders, setting each servo in a range of

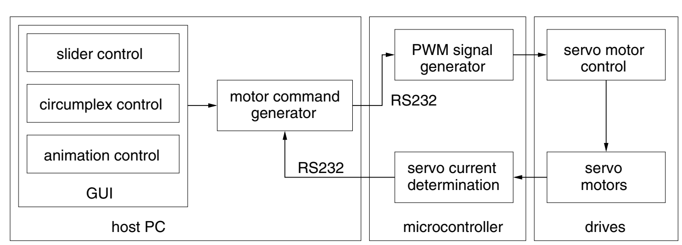
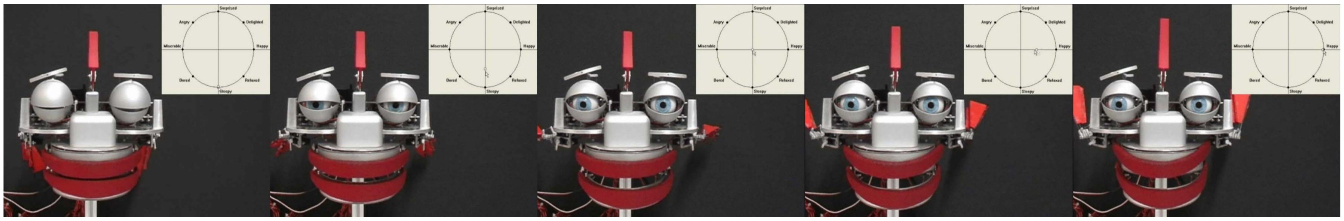
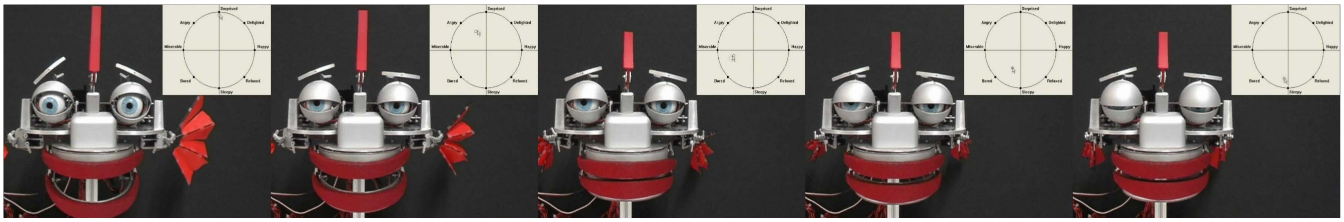


Fig. 7. Control architecture.



a)



b)

Fig. 8. Dynamic emotional state transitions commanded using the graphical user interface for a) a transition along the main axes and b) a free curve within the circumplex model (See companion video, http://www.lsr.ei.tum.de/).

0 to 255. All servos are updated within 20ms. The number of sliders is scalable, enabling upgrades.

2) Control of the basic emotions: Six buttons trigger default position settings for the basic emotions. There is no soft transition between these states, EDDIE changes emotions in a sudden. Simultaneously, the 2D-view is updated. The definition for the default settings was determined by tests to maximize identification of the shown facial expressions.

3) Computation of the position in 2D-view: The 2D-view is based on the circumplex model of affect. A white point shows the current state of the display. This point can be set to any position inside the circle by a mouse click to change the state in a sudden or can be dragged by holding the mouse button down to change the emotional state continuously. This is exemplarily shown in Figure 8 for a transition along the main axes (Figure 8a) and for a free curve within the circumplex model (Figure 8b).

The desired motor positions, determined in reference to the 2D-coordinate system, are computed for every single servo motor. As a purely linear mapping of a position of a servo motor to one of the axes of the circumplex model of affect (valence and arousal) would give no satisfying result, the area enclosed by the circle is divided into four quadrants, which are equal to the quadrants of a cartesian coordinate system. In every quadrant, the desired position is computed by a different equation. The functions are continuous at the cutting edges.

The variables $valence$ and $arousal$ vary in a range from 0 (= miserable / sleepy) to 255 (= happy / surprised), the desired motor positions $despos$ equally range from 0 to 255. Variable $radius$ is between 0 (= middle of the circle) and 1 (= on the circle).

IV. Evaluation

A preliminary user-study with 24 participants has been conducted. The study consists of a multiple-choice-test, in which people should match six shown facial expressions to 10 given answers. The 10 given answers are the same Breazeal used in her evaluation for Kismet to assure comparable results [1].

Eight children between the ages of five to eight and 16 adults between the ages of 25 to 48 participated, seven of whom are female and 17 are male.

The study consists of a total of 32 questions. Participants have to choose their best guess for a shown emotion. In a following question, they can choose other possible emotions which could match the shown expression. On a scale from 1 to 10 they can rate their best guess (1 very unsure if choice matches, 10 very sure). EDDIE is shown to the adults in groups of three and to the children one by one.

Evaluations for the projects Kismet, Aryan and Feelix were conducted in a similar way [3], [17]. The comparison of the results is shown in Table II.

The results show that the basic emotional states of EDDIE are recognized better than those of Feelix, but not as good as those displayed by Kismet. Considering the simple heuristic mappings from emotion-space to action units and joint-space of the current version of EDDIE these results are promising. Some shortcomings of the current version potentially impairing the recognizability of emotional states are a weak recognizability of the eyebrows and degrees of freedom for head motions are not yet integrated in contrast to, e.g., Kismet, which also have a strong impact on emotion expression [18]. Another important issue are the additional elements ears and crown which are not covered by FACS and heuristically controlled in the current version of EDDIE. These aspects will be adressed in the next paragraph.

A. Usage of animal-like features to display human emotions

In an experimental study with 30 participants (15 females, 15 males) the influence of two animal-like features (crown of a cockatoo and ears of a dragon lizard) was explored. In a 2x2 MANOVA design (1. factor crown, 2. factor ears) it was analyzed if these two factors would shift the observed six basic

TABLE II

IDENTIFICATION RATE OF DIFFERENT ROBOT FACES IN COMPARISON

	Eddie	Kismet	Aryan	Feelix
anger	54 %	76 %	94 %	40 %
disgust	58 %	71 %	–	–
fear	42 %	47 %	41 %	16 %
joy	58 %	82 %	–	60 %
sadness	58 %	82 %	–	70 %
surprise	75 %	82 %	71 %	37 %

emotions in the emotional space. Each factor was realized in four conditions (from fully stilted to dismantled). After each MANOVA a post-hoc test (Scheffe procedure) was conducted. All six basic emotions were displayed with every combination of the two factors. Afterwards the participants rated each displaced emotion on the verbal semantic differential scale [8]. Results suggest that the crown has no influence on the observation of displayed emotions. The ears however influence only the emotion 'disgust' as seen in Table III.

TABLE III

INFLUENCE OF STILTED EARS ON OBSERVED EMOTION 'DISGUST'

	conditions	measured difference	p-value	effect (R^2)
valence	1 AND 3	0.845	0.005	0.130
	1 AND 4	0.842	0.006	0.130
arousal	1 AND 3	1.0911	0.006	0.126

conditions: 1 = fully stilted, 2= half stilted, 3=not stilted, 4=dismantled

But as the effect-values show (R^2 ranges between 0 and 1 and can be seen as an performance parameter) the effect of the ears on the displayed emotions is relatively small.

V. CONCLUSION

A novel flexible low-cost emotion-display for dynamic human-robot interaction is presented. A heuristic emotion to joint-space mapping based on FACS provides smooth emotion state transitions. Results of the evaluation regarding the recognizability of emotions in a first user-study are satisfying. Based on the current heuristic mappings basic emotion states can be recognized better than those of Feelix, but not yet as good as those of Kismet. Regarding the introduction of additional animal-like features it has been shown that only the lizard-ears have an significant but small influence on the display of the human emotion 'disgust'.

Current research is concerned with an improvement of the design and of the mappings from emotions to joint-space. The appearance of the eye-brows will be enhanced with furlike covers, additional degrees of freedom for the head platform and additional degrees of freedom for the mouth will be introduced. In the video proceedings a companion video is included; see http://www.lsr.ei.tum.de for additional videos and information.

ACKNOWLEDGMENTS

The valuable comments by the anonymous reviewers are highly appreciated.

REFERENCES

[1] C. L. Breazeal, *Designing Sociable Robots*, The MIT Press, 2002.
[2] C. L. Breazeal, *Robotic Life - Sociable Robots*, http://robotic.media.mit.edu/projects/Leonardo/Leo-learning.html, 2005.
[3] H. Mobahi, *Building an Interactive Robot from Scratch*, Bachelor Thesis, 2003.
[4] A. Faiola, *Humanoids With Attitude: Japan Embraces New Generation of Robots*, The Washington Post, 2005.
[5] F. Hara, *Personality Characterisation of Animate Face Robot through Interactive Communication with Human*, Proceedings of the 1998 International Advanced Robotics Program (IARP98), 1998.
[6] C. R. Darwin, *The expression of the emotions in man and animals*, John Murray, 1872.
[7] J. Altarriba, D. M. Basnight, T. M. Canary, *Emotion representation and perception across cultures*, Online Readings in Psychology and Culture, 2003.
[8] A. Mehrabian and J. A. Russell, *An approach to environmental psychology*, Cambridge: MIT Press, 1974.
[9] D. R. Carney and C. R. Colvin, *The circumplex structure of emotive social behavior*, Northeastern University, 2005.
[10] J. A. Russell, *A circumplex model of affect*, Journal of Personality and Social Psychology, 1980.
[11] J. A. Russell, *Pancultural aspects of the human conceptual organizations of emotions*, Journal of Personality and Social Psychology, 1983.
[12] P. Ekman and W. V. Friesen, *Facial Action Coding Consulting*, Psychologist Press, 1977.
[13] *e-learning-course "about facs"*, Saarland University, http://www.uni-saarland.de/fak5/orga/Kurs/home.htm, 2003.
[14] J. Reichard, *Robots: Fact, Fiction and Prediction*, Penguin Books, 1978.
[15] C. F. DiSalvo, F. Gemperle, J. Forlizzi, S. Kiesler, *All Robots Are Not Created Equal: The Design and Perception of Humanoid Robot Heads*, Carnegie Mellon University, 2002.
[16] I. Eibl-Eibesfeldt, *Grundriss der vergleichenden Verhaltensforschung - Ethnologie*, Munich, 1987.
[17] L. D. Canamero, J. Fredslund, *How does it feel? Emotional Interaction with a Humanoid LEGO Robot*, Socially Intelligent Agents: The Human in the Loop, Papers from the AAAI 2000 Fall Symposium, 2000.
[18] H. G. Wallbott, *Bodily expression of emotion*, European Journal of Social Psychology, 28, 1998.