The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Design and Testing of a Hybrid Expressive Face for a Humanoid Robot

Daniel Bazo, Ravi Vaidyanathan, Alexander Lentz, and Chris Melhuish

***Abstract*—The BERT2 social robot, a platform for the exploration of human-robot interaction, is currently being built at the Bristol Robotics Laboratory. This paper describes work on the robot's face, a *hybrid face* composed of a plastic faceplate and an LCD display, and our implementation of facial expressions on this versatile platform. We report the implementation of two representations of *affect space*, each of which map the space of potential emotions to specific facial feature parameters and the results of a series of human-robot interaction experiments to characterize the recognizability of the robot's archetypal facial expressions. The tested subjects' recognition of the implemented facial expressions for *happy*, *surprised*, and *sad* was robust (with nearly 100% recognition). Subjects, however, tended to confuse the expressions for *disgusted* and *afraid* with other expressions, with correct recognition rates of 21.1% and 52.6% respectively. Future work involves the addition of more realistic eye movements for stronger recognition of certain responses. These results demonstrate that a hybrid face with *affect space* facial expression implementations can provide emotive conveyance readily recognized by human beings.**

I. INTRODUCTION

RESEARCH in assistive robotics, in particular biomimetic robots with humanoid characteristics, has experienced prolific growth in recent years. However, there remain significant issues to be addressed related to human-robot cooperation. In order for such robots to function, they must perform physical tasks within the personal-space of a human. Unlike most Human-Computer Interaction (HCI) applications, this is typified by the shared manipulation of objects and even direct contact (e.g. moving an infirm person). However, human beings exhibit a complex myriad of verbal and non-verbal cues which affect one-another's behavior that are critical to cooperative execution of tasks. For example, if two humans are moving a heavy object, they will affect each other through inflection of speech, facial expressions, gestures with the limbs, gross body movements, and a range of other reactions. If a robot is to work with a human and possess human form, then it becomes very important that its motions are 'human-like', and that it uses the full range of human communication channels. Enabling such reactions to assistive robots with humanoid form, almost by default, necessitates an approach based on natural communication (verbally, through facial expressions, or through gesture), perception and understanding of intention, and cognition supporting interaction. Humans integrate these challenges effortlessly; an assistive robot in this scenario must exploit such mechanisms in a similar manner.

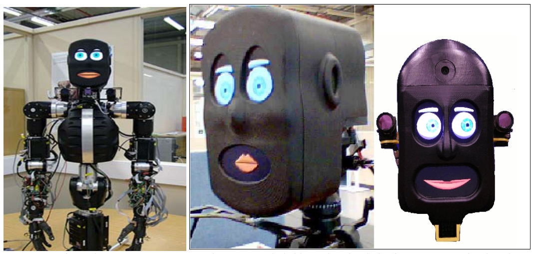
Fig. 1. Left: the BERT2 robot; Middle: hybrid face used during the experimental work; Right: second version of BERT2's hybrid face with integrated stereo vision system for human gaze tracking and wide-angle web-camera in the forehead.

The second generation Bristol-Elumotion Robotic Torso (BERT2) robot (Figure 1) is currently under development at the Bristol Robotics Laboratory (BRL). BERT2 is a social robot with an expressive face meant to help researchers design intelligent systems capable of ensuring mutual trust, safety, and effective cooperation with human beings.

This paper describes a *hybrid face* implemented on BERT2 face. First, we discuss the motivation for a hybrid-face robot and the implementation and degrees of freedom of the computer-graphic facial features; second, we give two different mathematical representations for the affective space defined by the robot's facial expressions; finally, we present experimental results with human end-users characterizing the recognizability of the BERT2 facial expressions as a quantitative measure to verify the overall efficacy of the hybrid face.

A. Related Work

Facial expressions are well-recognized as critical to the functionality of social robots. Edsinger, O'Reilly, and Breazeal [1] posit that the face of a robot implies a social contract between itself and its users, drawing upon the emotions and expectations that people have for each other and for machines [1]. According to Schiano, Ehrlich, Rahardja, and Sheridan, facial expressions are the primary method of communication of affective information [2], and roboticists Canamero and Fredslund stress the importance of an expressive face to promote natural, believable interactions [3]. Fukuda, Jung, Nakashima, Arai, and Hasegawa, point out a potential benefit of robots capable of following human social conventions: a decrease in the amount of training a human requires before he or she can interact productively with the robot [4].

Manuscript submitted March 10, 2010 revised July 15, 2010.

Danny Bazo is with the Department of Media Arts and Technology at the University of California, Santa Barbara, CA 93106 USA (dannybazo@gmail.com).

Ravi Vaidyanathan is with the Bristol Robotics Laboratory, Department of Mechanical Engineering, University of Bristol, Bristol, BS8 1TR, UK the US Naval Postgraduate School, Monterey, CA, USA, 93943 (r.vaidyanathan@bristol.ac.uk)

Chris Melhuish and Alexander Lentz are with the Bristol Robotics Laboratory [melhuish; lentz]@brl.ac.uk, Bristol, UK, BS16 1QD

978-1-4244-6676-4/10/$25.00 ©2010 IEEE

Several robots, (e.g. [1], [5] [6]), have been designed to have child-like or intentionally 'cute' facial features in order to increase their likeability (or to decrease human expectations of the robot's capabilities). Other researchers, attempting to replicate human adult appearance, have created androids which can pass for human in still photographs [7], [8]. Some researchers focus on replicating the mechanics of facial expressions, using robotic heads capable of the same range of facial expression as a human [9], [10]. Blow turns to illustration techniques to classify this range of robot morphology, using a scale that spans realism, iconicity, and abstraction [11].

1) *Range of Expressions*

The most commonly used design methods for exploring the expressive potential of a robot's face have drawn from the field of psychology. Ekman's [12] introduced a 'Facial Action Coding System (FACS)' that led to a theory implying a categorical nature for expressions, suggesting all expressions are combinations of the basic expressions of *happiness, sadness, anger, fear, disgust, and surprise*. Russell and Fernandez-Dols argue an alternative theory of facial expressions, proposing a space of facial expressions spanning a two-dimensional space defined by *valence* and *arousal* axes [13]. Experiments by Schiano et al. attempted to illuminate this discussion by asking human research subjects to classify emotions seen in both images of human actors and robot faces. Using MDS analysis, the researchers argued that Russell's two-dimensional affect space model, augmented with a third dimension, yielded the most robust results [14]. Recent work by Kleinsmith, Silva, and Bianchi-Berthouze supports a three-dimensional representation of affect space, which they call *action tendency* [15] while stressing that dimension's role in disambiguating the expressions for surprise, fear, disgust, and anger—a problem which has been seen in the recognition of both robotic and human faces [2], [3, [4], [16]. Recognizability of Ekman's basic expressions is a common test used to gauge the abilities of an expressive robot face [2], [3], [16], [17].

The three-dimensional, continuous [14] affect space preferred by Schiano has been utilized by Breazeal to represent the emotions of her Kismet robot [6]. In that wor, the advantage of the third affect axis, *stance*, was used to separate fear, anger, and disgust in affect space. This usage is consistent with the action tendency dimension proposed by Kleinsmith, et al. in [15]. Breazeal's stance dimension is defined by two expressions: "openness" in the positive direction, and "closedness/sternness" in the negative direction. Furthermore, anger, fear, and disgust expressions are given their own locations in affect space, each with negative valence and high arousal.

Finally, alongside the research into two- and three-dimensional representations of affect space in robots, the categorical approach to facial expression creation is still practiced. For example, Gockley's Valerie robot represents its emotional state as some combination of all of its basis emotions [18], as does Nourbakhsh's Sage robot [19].

B. *Relation to Previous work*

The BERT2 robot, currently being built at BRL as part of the Cooperative Human Robot Interaction Systems (CHRIS) project, is intended to facilitate the study of interactions between humans and social robots. The complete system makes use of a torso with two arms and hands, an eye-tracking camera system, microphones, and a video camera to develop methods of safe cooperation between humans and robots. The BERT2 face is a key part of the social abilities of the robot, allowing recognizable facial expressions to foster affective exchanges between BERT2 and a human partner. The face of the robot is a *hybrid* face robot, in that it combines a digital face with a static human visage-like structure. It is designed to provide the flexibility of a digital countenance with some of the benefits of a full-featured, fully actuated face. Ideally it will afford some of the benefits in human-reaction and trust of such robots, but without the complexity in control or actuation of a full facial robot. All functionalities of BERT2's head have been fully integrated into a communication infrastructure via YARP [23] and all animation states can be accessed via remote procedure calls (RPC) from external software modules.

Valerie the Roboceptionist, the creation of Gockley, et al., shares an element with the BERT2 face: the use of a display screen with rendered computer graphics rather than a mechanical face [18], [20]. However, Gockley notes that the use of a display screen can make detecting the precise direction of Valerie's gaze difficult to discern [18]. This may lead to complications during interactions with multiple humans. Minato and Imai posit that accurate perception of gaze direction is vital to social interaction [21], [22]. In BERT2, this has been addressed by using a "hybrid face" design. The LCD display screen on which features are shown is partially covered by a plastic faceplate, with molded contours meant to emulate the brows, nose, and chin of a human face.

We have attempted to extend the representation of three-dimensional affect space as proposed by Breazeal in [6] by making changes to this formulation, in the hopes of creating a more compact mathematical representation:

First, the third affect dimension (called *stance*, adopting Breazeal's nomenclature) is defined by the expressions for *anger* and *fear*, under the hypothesis that sternness and openness can be expressed by some other combination of arousal, valence, and stance. Second, a *neutral* expression defines the center of the affect space coordinate system. Third, disgust has no unique location in affect space, under the assumption that disgust can be shown effectively by some other static or time-varying combination of arousal, valence, and stance, or else with body or head gestures.

Finally, though not formally explored in this work, the BERT2 face has a feature not present in most social robots; its pupils capable of dilation and contraction. The effect of the inclusion of realistic pupillary response on recognizability of facial expressions is a topic of future research.

II. IMPLEMENTATION

A. Degrees of Freedom

The BERT2 face consists of four facial features: eyebrows, eyelids, eyeballs, and mouth, with a total of thirteen degrees of freedom (Fig. 3). These degrees of freedom are: left and right eyebrow angle B_{al} and B_{ar}; left and right eyebrow vertical height B_{hl} and B_{hr}; left and right eyelid openness L_l and L_r; eye pitch and yaw E_p and E_y; pupil size P; mouth corner vertical height M_h; mouth width M_w; top lip openness M_t; and bottom lip openness M_b.

Taken together, these thirteen values characterize the current facial expression in the facial expression state $\vec{e}(t)$ for a given time t.

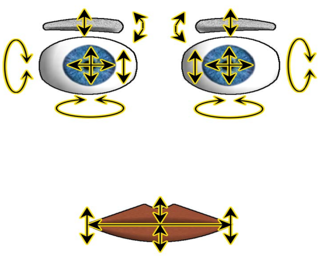

Fig. 3. The thirteen degrees of freedom of the animated BERT2 face.

B. Realism Features

In order to add realism to the BERT2 face, an added-randomness feature can be enabled. This feature simulates the subtle twitching and constant motion that characterizes a real human face. The mouth corner height, eyebrow angle and height, and eye pitch and yaw are all varied from the desired input signal by adding noise to the desired value. We hypothesized that this effect causes human viewers to perceive a dynamic, expressive character in the face, adding to the recognizability, likeability, and realism of the BERT2 face.

Another optionally active feature of the BERT2 face is the random blinking of the eyes in a lifelike manner. Blinking is controlled by a random number generator, which generates a random integer $r \in \{0, 1, \ldots, \mu_{blink} - 1\}$, where $\mu_{blink} \in \{1, 2, \ldots\}$ is a chosen blinking factor. If $r = 0$ (an event with a $\frac{1}{\mu_{blink}}$ probability of occurring), then the BERT2 face eyes simultaneously close then open at a high speed.

The random number r is generated once every 200ms when this feature is active (corresponding to a maximum blinking rate of five blinks per second, similar to the typical human limit). Thus a very small value of μ_{blink} (such as 1) will result in a maximum of five blinks per second.

III. AFFECT SPACE REPRESENTATIONS

A. Categorical Affect Space

In the categorical representation of the facial expression state, it is desired to characterize the robot's facial expression as being some combination of basis expressions. Typically, these faces are taken from the psychology literature. For this program, a set of basis expressions (set B) was adapted from Breazeal's work in [10].

In this implementation, these basis facial expressions are: *happiness, sadness, anger, fear, surprise, tiredness, sternness*, and *disgust*. Additionally, there is a *neutral* expression. Although in this implementation there are eight basis expressions, this technique is generalizable for any number of basis expressions and will be discussed in the general case (using the variable n to represent the number of basis expressions).

The set of basis expressions B consists of n basis expressions $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$, each of which is a vector containing thirteen values (within the ranges given in Section 3.1), one for each of the BERT2 face's thirteen degrees of freedom. There is also a *neutral* expression: $\vec{b}_N$.

To create a face based on a combination of basis expressions, one selects the vector of weights

$$\vec{w} = [w_1, w_2, \ldots, w_n],$$

where each element $w_i \in [0, 1]$ represents the amount by which the corresponding basis expression $\vec{b}_i$ should contribute to the final facial expression $\vec{e}$.

We assume that the defining characteristics of each basis expression are its *variances from the neutral expression*, so the difference $\vec{b}_i - \vec{b}_N$ is used to get the contribution $\vec{e}_i$ for each weighted basis expression:

$$\vec{e}_i = (\vec{b}_i - \vec{b}_N)\vec{w}_i.$$

Then, let the *basis difference matrix*

$$\hat{B} = \left[\hat{\vec{b}}_1\, \hat{\vec{b}}_2 \ldots \hat{\vec{b}}_n\right] = \left[\vec{b}_1 - \vec{b}_N \,\middle|\, \vec{b}_2 - \vec{b}_N \,\middle|\, \ldots \middle|\vec{b}_n - \vec{b}_N\right].$$

Hence, each $\vec{e}_i = \hat{\vec{b}}_i w_i$.

The final expression $\vec{e}$ is the sum of the weighted basis difference expressions added to the neutral expression:

$$\vec{e} = \left(\sum_{i=1}^{n} \vec{e}_i\right) + \vec{b}_N = \left(\sum_{i=1}^{n} \hat{\vec{b}}_i w_i\right) + \vec{b}_N = \hat{B}\vec{w} + \vec{b}_N.$$

Using this method, any weighted combination of basis expressions is achievable. For the eventual implementation of more sophisticated human-robot interactions, this abstraction is desirable, as shown in work by Gockley in [23] and Nourbakhsh in [24].

B. 3-Dimensional Affect Space

Based on the work of Russell, Schiano, and others, there is a movement towards lower-dimensional representation of affect space [2], [13]. Russell claimed originally that a two-dimensional model would suffice, and labeled the two axes of his affect space *arousal* and *valence* [13]. Since his work was published, statistical analysis by Schiano and others has shown that three dimensions, rather than two, are capable of capturing the vast majority of facial expressions [2], [15]. There is little agreement on whether the axes of such a space actually correspond to emotional parameters or some combination of physical parameters, but some useful conjectures have been put forth. Breazeal opts to combine past and recent work on the Kismet robot, whose affective space axes correspond to *arousal, valence,* and *stance*.

Research on facial expression recognition in humans has

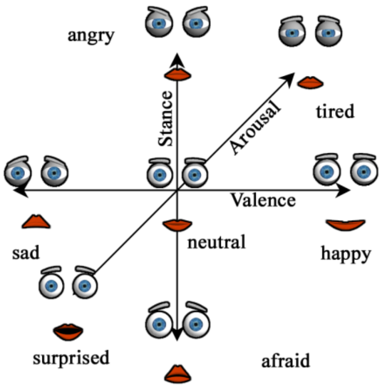

Fig. 4. The 3-dimensional affect space defined by the Arousal, Valence, and Stance axes, and examples of basis expressions.

shown that people have difficulty identifying fear, surprise, anger, and disgust. Breazeal's choice of the stance dimension serves primarily to alleviate this discrimination problem by separating fear from anger, giving them both a slightly negative valence and high arousal. However, this results in not only a larger set of bases for affect space, but in increased complexity in the function mapping affect space variables to facial expression.

In this implementation, an experimental approach was taken by the authors in which the stance axis itself is characterized by the facial expressions for anger (in the positive direction) and fear (in the negative direction) (Fig. 4). This decision was made in order to simplify the affect space representation posited by Breazeal, and to examine several hypotheses. First: that disgust, an expression which is typically difficult to identify in a static face, can be expressed in a more recognizable fashion through postural and vocal mechanisms, and thus does not need to be a basis expression of affect space. Second: expressions can show anger or fear at the same time as other, meaning that both expressions deserve a role as basis expressions. Third: that the open and closed/stern stances used by Breazeal to characterize the stance dimension can be expressed by a combination of arousal, valence, and stance, when stance is characterized by the fear and anger expressions.

The three axes of the three-dimensional affect space are characterized by six basis expressions, grouped in three pairs of opposites: happiness is opposite sadness on the valence axis, surprise is opposite tiredness on the arousal axis, and anger is opposite fear on the stance axis. Because of this opposing relationship, it is assumed that it is undesirable (or even impossible) for a face to display two opposite emotions simultaneously. For example, it is difficult to imagine a face that simultaneously expresses happiness and sadness.

Each opposing basis expression is placed at $+1$ or -1 on its appropriate axis, meaning that the total useable affect space is a cube with sides at $+1$ and -1 on each of its axes.

Thus, at any given time, a maximum of three expressions will be contributing to the current expression. A linear contribution law is applied, with movement along any affect axis corresponding to a linear increase in the contribution of the basis face being approached along that axis. This time, unlike under the categorical paradigm, there are only six basis expressions $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_6$ making up the basis set B; again with the addition of the neutral expression $\vec{b}_N$. Once again, because the importance of each basis expression is its variance from the neutral expression, we can form the set of basis differences

$$\hat{B} = \left[\vec{b}_1 - \vec{b}_N \mid \vec{b}_2 - \vec{b}_N \mid \ldots \mid \vec{b}_6 - \vec{b}_N\right].$$

Given a desired location in affect space $\vec{x} = [\alpha\ \beta\ \gamma]^T$ where $\alpha, \beta, \gamma \in [-1, +1]$, the corresponding expression-space representation $\vec{e}$ is given by

$$\vec{e} = \max(\alpha, 0)\,\hat{\vec{b}}_1 + \max(-\alpha, 0)\,\hat{\vec{b}}_2 + \max(\beta, 0)\,\hat{\vec{b}}_3 + \max(-\beta, 0)\,\hat{\vec{b}}_4 + \max(\gamma, 0)\,\hat{\vec{b}}_5 + \max(-\gamma, 0)\,\hat{\vec{b}}_6 + \vec{b}_N,$$

where the maximum function max is used to negate the contribution of a basis expression when the current position $\vec{x}$ is closer to its opposite basis expression. Note that the basis difference expressions are assumed to be ordered in B such that $\vec{b}_1$ and $\vec{b}_2$, $\vec{b}_3$ and $\hat{\vec{b}}_4$, and $\vec{b}_5$ and $\vec{b}_6$ are, respectively, opposites.

IV. EXPERIMENT

In order to perform some primary testing of the functionality of the BERT2 face, an experiment was conducted using human subjects. The experiment as designed primarily to provide validation of the face software's functionality, and to give indications of what types of features and design decisions may be useful in the continuing construction of the BERT2 robot.

A. *Expression Recognition*

1) *Procedure*

The first experiment was an expression recognition task common to robotic facial expression research [2], [3], [16], [17]. Human subjects were given a list of eight emotions (happiness, sadness, anger, sternness, surprise, disgust, fear, and tiredness) and were seated facing the BERT2 head. Subjects were then shown a series of facial expressions of approximately four seconds each. After each facial expression, the face display screen was turned off momentarily and the subject was asked to select which emotion best matched the presented face (a forced-choice). Additionally, each face had zero, one, or two of the following conditions applied: *realism* and *transition*. Several instances of each condition were shown to all subjects.

Realism features were comprised of the random twitches and blinking described in Section II.B. The *transition* condition specified whether the face was a static expression (which did not vary the expression being shown at any given time), or a dynamic expression (which transitioned smoothly from a neutral face to the presented face over two seconds). Ten subjects of various ages and cultural backgrounds took in the experiment, which was carried out at BRL. Subjects included five adult men, four adult women, and one ten-year-old girl. Two male subjects and one adult female subject had previous experience in robotics research.

2) *Results and Discussion*

The results from the experiment are shown in Tables I - II. The values on each row represent, for a single facial expression, the percentage of responses of all of the forced-choice facial expression options as chosen by subjects.

Under all conditions, the expressions happy, surprised, and sad showed nearly perfect recognition. However, other expressions showed some confusion between one another. In particular, stern and tired were often confused with one another, angry was confused with stern, afraid was confused for surprised and to a lesser degree sad, and disgusted suffered from a general identification difficulty.

A possible reason for the confusion between stern and tired is their similarity: both expressions feature partially-closed eyelids and low mouth width. Though stern is characterized by a slightly downward-curving mouth and a slightly knitted brow, these differences were apparently too subtle for subjects to robustly disambiguate the two expressions from one another. This problem was worsened in all conditions with animation present, whose effect of slowly closing the eyelids most likely contributed to increased confusion.

The expressions for stern and angry have similar eyebrow and mouth shapes, though stern featured a less extreme eyebrow angle and slightly closed eyelids. Furthermore, it is possible that subjects perceive sternness as a form of anger. Both expressions communicate displeasure and aggression--active aggression in case of anger and passive aggression in the case of stern.

The expression for afraid shared the raised eyebrows and opened eyelids of surprise, as well as the downward-curving mouth shape of sad, which may explain their confusion. The disgusted expression was problematic for subjects. The positions of the facial features were inspired by the expression of disgust in Breazeal's Kismet robot [17], but subjects did not associate the eyelid asymmetry or downward-curving mouth with the emotion of disgust. The presence of one raised eyebrow complicated matters, with three subjects informing the experimenter that the expression appeared quizzical—an option not available for selection. Another possibility is the dissimilarity of the disgusted expression as used in this experiment and the expression of disgust in the general population as described by Ekman [12], which features a wrinkling of the nose. This movement is not possible with the current configuration of the BERT2 face, so an alternative was chosen.

The confusion of disgust, fear and surprise is common in facial recognition literature [2], [3, [4], [16], and suggests that facial expression alone is not sufficient to robustly communicate the entire range of human emotion. This is to be expected, however, as humans do not simply make static faces to express emotion-body language, context, and time-varying expressions all play a significant role in the meaning of a particular facial expression.

The presence of animation and realism decreased recognition rates for stern, angry, and tired, while increasing the recognition rates for surprise and afraid. One hypothesis for the decrease in recognition of some expressions is that the speed of animation (a two-second interpolation) was uncharacteristic for certain expressions. For example, the slow transition to angry led to increased confusion with sternness, potentially because anger is typically characterized as a violent, active emotion (part of the so-called fight-or-flight response), and thus with a rapid change of expression. The slow transition also increased confusion of stern for tired--of which the latter is significantly more associated with slow movement. The slow movement of the closing eyelids, regardless of the difference in eyebrow position for those two expressions, may have led to the increased confusion.

The improvement of recognition of surprise under all conditions with animation present may be due to the nature of surprise as a transitory emotion--humans typically do not spend long amounts of time with an expression of surprise on their face. Thus, the transition from neutral to surprised may have accentuated the non-static quality of the emotion of surprise.

V. CONCLUSIONS

Cooperation between humans and robots in a shared workspace requires trust—humans achieve this trust among one another partially by displaying and reading facial expressions, and so the benefits of a robot capable of the same types of communication are obvious. At the same time, the optimal balance between realistic appearance and iconic appearance in a social robot is not precisely known.

The BERT2 face, a hybrid face consisting of both a plastic faceplate and an LCD graphics display, attempts to address both of these issues while serving as a platform for the general exploration of human-robot interaction. It features two types of mathematical representations of the space of affective potential: a *categorical* representation in which any facial expression can be represented as a linear combination

TABLE I
OVERALL EXPERIMENTAL EXPRESSION CONFUSION MATRIX
(% TIMES REPORTED OUT OF TOTAL PRESENTED)

	Hap	Ste	Ang	Dis	Sur	Afr	Sad	Tir
Hap	98.7	0	0	0	1.3	0	0	0
Ste	0	62.8	7.7	5.1	1.3	0	1.3	21.8
Ang	0	30.8	64.1	5.1	0	0	0	0
Dis	0	25.0	13.2	18.4	13.2	10.5	5.3	14.5
Sur	0	0	0	0	93.4	6.6	0	0
Afr	0	0	1.3	0	37.3	44.0	16.0	1.3
Sad	0	0	0	0	0	0	100	0
Tir	1.3	18.8	6.3	1.3	1.3	0	5.0	66.3

TABLE II
CORRECT RESPONSES WITH AND WITHOUT REALISM FEATURES
(% CORRECT RESPONSES FOR GIVEN FEATURES)

	Hap	Ste	Ang	Dis	Sur	Afr	Sad	Tir
Tot	98.7	62.8	64.1	18.4	93.4	44.0	100	66.3
Ani	100	75	65	16.7	94.7	27.8	100	65.0
Real	95.2	60	65	15.0	90.0	47.4	100	65.0
Both	100	42.1	55.6	21.1	100	52.6	100	70.0
No	100	73.7	70.0	21.1	89.5	47.4	100	75.0

of any of a set of basis expressions, and a *3-dimensional* representation, in which three orthogonal axes defined by six basis expressions define affective space.

As an attempt to verify the functionality of the BERT2 face, and to gain insights into mathematical representations of affective potential, an experiment was carried out in which human subjects were asked to identify the BERT2's facial expression given a list of potential expressions. Subjects repeatedly correctly identified the expressions for happy, surprised, and sad, and repeatedly incorrectly identified the expressions for disgusted and afraid. The presence of animation and realism in general decreased recognition rates for all expressions except surprise. The current basis expressions used on the BERT2 robot were adapted from past robotics research, in particular [2], [3], [16], [17]. Results show considerable promise for the platform as many significant facial expressions were recognized by human beings while interacting with the robot.

VI. FUTURE WORK

In order to improve the unsatisfactory recognition rates of the basis expressions, an experiment carried out by Schiano in [2] could be repeated, wherein subjects were presented with a neutral expression and were asked to adjust the individual facial features themselves, in order to conform to each of the desired basis expressions. In particular, humans who are to work at length with the completed BERT2 robot could perform this experiment, increasing the likelihood of later recognition of the expressions they themselves programmed.

The expressions for stern and afraid should be modified to further differentiate them from angry and surprised respectively. The expression for disgust should be completely remodeled, perhaps by attempting to simulate the nose-wrinkling motion described by Ekman.

The inclusion of pupillary response may help to disambiguate between problematic expressions. To the best of our knowledge, this feature has not been extensively studied in relation to recognition of robotic facial expressions.

Finally, the speed of animation should be varied in order to more closely match the expected speed for a given expression. One method would tie the speed of expression interpolation to the affective dimension of arousal, decreasing interpolation speed with low-arousal states and increasing interpolation speed with high-arousal states.

VII. ACKNOWLEDGMENTS

This work has been partially funded by the European Commission under the Robotics and Cognitive Systems, ICT Project CHRIS (FP7-215805) and the authors gratefully acknowledge this.

REFERENCES

[1] Edsinger, A., O'Reilly, U.M., and Breazeal, C., Personality through faces for humanoid robots. In: Proceedings of the 9th IEEE International Workshop on Robot and Human Interactive Communication, 2000, pp. 340-345.

[2] Schiano, D., Ehrlich, S., Rahardja, K., and Sheridan, K., Face to interface: facial affect in (hu) man and machine. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 193-200.

[3] Canamero, L. and Fredslund, J., I show you how I like you—can you read it in my face? IEEE Transactions on Systems, Man and Cybernetics, Part A, 2001, 31, no. 5, 454-459.

[4] Fukuda, T., Jung, M., Nakashima, M., Arai, F., and Hasegawa, Y., Facial expressive robotic head system for human-robot communication and its application in home environment. Proceedings of the IEEE, 2004, 92, no. 11, 1851-1865.

[5] Scheeff, M., Pinto, J., Rahardja, K., Snibbe, S., and Tow, R., Experiences with sparky, a social robot. In: Proceedings of the Workshop on Interactive Robot Entertainment, 2000.

[6] Breazeal, C., Affective interaction between humans and robots. Lecture Notes in Computer Science, 2001, 582-591.

[7] Hanson, D., Exploring the aesthetic range for humanoid robots. In: Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, pp. 39-42.

[8] Minato, T., Shimada, M., Itakura, S., Lee, K., and Ishiguro, H., Does gaze reveal the human likeness of an android? International Conference on Development and Learning, 2005, 106-111.

[9] Berns, K. and Hirth, J., Control of facial expressions of the humanoid robot head ROMAN. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3119-3124.

[10] Jaeckel, P., Campbell, N., and Melhuish, C., Facial behaviour mapping{from video footage to a robot head. Robotics and Autonomous Systems, 2008, 56, no. 12, 1042 -1049.

[11] Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, C., and Lee, D., Perception of robot smiles and dimensions for human-robot interaction design. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication, pp. 469-474.

[12] Ekman, P. and Friesen, W., The Facial Action Coding System. Palo Alto: Consulting Psychologists Press, 1978.

[13] Russell, J. and Fernandez-Dols, J., The Psychology of Facial Expression. New York: Cambridge University Press, 1997.

[14] Schiano, D., Ehrlich, S., and Sheridan, K., Categorical imperative NOT: facial affect is perceived continuously. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 49-56.

[15] Kleinsmith, A., Silva, P., and Bianchi-Berthouze, N., Grounding affective dimensions into posture features. In: Proceedings of the First International Conference on Affective Computing and Intelligent Interaction, pp. 263-270.

[16] Bartneck, C., Reichenbach, J., and Van Breemen, A., In your face, robot! the influence of a characters embodiment on how users perceive its emotional expressions. In: Proceedings of the Design and Emotion 2004 Conference.

[17] Breazeal, C., Emotion and sociable humanoid robots. International Journal of Human-Computer Studies, 2003, 59, no. 1-2, 119-155.

[18] Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A.C., and Wang, J., Designing robots for long-term social interaction. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1338-1343.

[19] Nourbakhsh, I. Bobenage, J., Grange, S., Lutz, R., Meyer, R., and Soto, A., An affective mobile robot educator with a full-time job.

[20] Gockley, R., Simmons, R., and Forlizzi, J., Modeling affect in socially interactive robots. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 558-563.

[21] Minato, T., Shimada, M., Itakura, S., Lee, K., and Ishiguro, H., Does gaze reveal the human likeness of an android? International Conference on Development and Learning, 2005, 106-111.

[22] Imai, M., Kanda, T., Ono, T., Ishiguro, H., and Mase, K., Robot-mediated round table: analysis of the effect of robots gaze. In: Proceedings of 11th IEEE International Workshop on Robot and Human Communication (ROMAN2002), pp. 411-416.

[23] P. Fitzpatrick, G. Metta, and L. Natale, "Towards long-lived robot genes," Robot. Auton. Syst., vol. 56, no. 1, pp. 29–45, 2008. [Online Available: http://portal.acm.org/citation.cfm?id=1327705