I Show You How I Like You: Human-Robot Interaction through Emotional Expression and Tactile Stimulation

Dolores Cañamero* Jakob Fredslund*

May 31, 2000

Abstract

We report work on a LEGO robot capable of displaying several emotional expressions in response to physical contact. Our motivation has been to explore believable emotional exchanges to achieve plausible interaction with a simple robot. We have worked toward this goal in two ways. First, acknowledging the importance of physical manipulation in children’s interactions, interaction with the robot is through tactile stimulation; the various kinds of stimulation that can elicit the robot’s emotions are grounded in a model of emotion activation based on different stimulation patterns. Second, emotional states need to be clearly conveyed. We have drawn inspiration from theories of human basic emotions with associated universal facial expressions, which we have implemented in a caricaturized face. We have conducted experiments on both children and adults to assess the recognizability of these expressions.

1 INTRODUCTION

In this paper we introduce Feelix\textsuperscript{2} (Figure 1), a 70cm-tall humanoid-looking LEGO robot that displays different facial emotional expressions in response to tactile stimulation.

\textsuperscript{1}LEGO-Lab, Dept. of Computer Science, University of Aarhus, Aabogade 34, DK–8200 Århus N. E-mail: lola@daimi.au.dk, chili@daimi.au.dk.

\textsuperscript{2}FEELIX: FEEL, Interact, eXpress.
Feelix is a descendant of Elektra [8], a mobile, also “humanoid” LEGO robot first exhibited at the FIRA Robot World Cup in Paris in 1998. People, in particular children, found it very natural to interpret the happy and angry expressions of Elektra’s smiley-like face; however, the interaction humans could have with it to elicit these expressions was not so natural, as it consisted in inserting color LEGO parts into its chest. Our motivation for building Feelix—Elektra’s body with a new head and feet—was twofold. First, we aimed at a more plausible interaction with the robot. We wanted to focus on the interaction itself, and therefore we did not want it to be influenced by the robot performing a particular task. For this, we decided to exploit the potential that robots, unlike computer simulations, offer for physical manipulation, as this plays an important role in children development and in human interaction in general. Interaction with Feelix is therefore through tactile stimulation rather than through other sensory modalities that do not require physical contact, such as vision. Occasionally, we have observed as a side-effect humans mirroring the emotion they want to elicit in Feelix in their own faces and in the nature of the presses they apply on the feet. Our second motivation was to achieve a richer interaction so that a wider range of interaction patterns giving rise to a wider range of emotional responses (both in the robot and, by empathy, in the human) was possible. However these emotional responses must be clearly recognizable, and therefore we limited ourselves to implementing the ones known in the emotion literature as “basic emotions”. People were not only able to recognize the expressions quite successfully (see Section 4), but in many cases they even mimicked Feelix’ expression with vocal inflection while commenting on the expression (“ooh, poor you!”), “look, now it’s happy!”). People thus seem to empathize with the robot rather naturally.

Our models of emotional interaction and expression are inspired by psychological theories about emotions in humans. This makes Feelix not only a robot very suitable for entertainment purposes, but also a proof-of-concept that these theories can be used with a synthetic approach that nicely complements the analytic perspective for which they were conceived.
2 THE MODEL OF EMOTIONS

The emotions we have implemented in Feelix correspond to the ones known as “basic emotions”. The use of this term is still highly controversial among students of human emotions (see [11] for a good account of this controversy), as researchers do not agree neither in the number and subset of emotions that can be considered as basic (classifications range from two to nine), nor in what sense they are so. As a general characterization, we could say that basic emotions seem to be universally found across societies in humans (some authors extend this universal character to other mammals), have particular manifestations associated with them (facial expressions, behavioral tendencies, physiological patterns), have adaptive, survival-related roles that evolved through their value in dealing with situations which are or were fundamental in life, can be seen as prototypical emotional states of a number of emotion families (e.g., rage and anger belong in the same family, anger being the more standard or prototypical case, while rage corresponds to a highly intense anger), and can be taken as building blocks out of which other, more complex emotions form. In our case, the main reason for adopting this hypothesis of a subset of discrete basic emotions is the ease with which their facial expressions are recognized even by children.

Some researchers prefer to characterize emotions in terms of continuous dimensions, rather than as discrete categories. The two most commonly used dimensions (names may vary) are valence (positive/negative) and activation or arousal (calm/excited). These views are not incompatible, and in our model we use a combination of them. In fact, basic emotions can be easily placed in an emotional space defined by these dimensions (see for instance [12]), although two dimensions alone are not enough to distinguish among all the basic emotions—intense fear and anger, for example, are both characterized by negative valence and high arousal. A third dimension, potency (powerfulness/powerlessness), is sometimes added.

2.1 Facial expression of emotions

The particular subset of basic emotions that we have adopted is the one proposed by Ekman [6]—anger, disgust, fear, happiness, sadness, and surprise—with the exception of disgust, since this emotion does not make sense for the kind of interactions that humans can have with Feelix. The reason for choosing this classification is that its main criterion to define emotions as basic is their having distinctive (universal) facial expressions [6].
As for facial primitives to express emotions, we have largely adopted the features concerning eyebrows and lips (the only elements that Feelix can use to express emotions) described in [9]. This coding system is inspired by Ekman’s Facial Action Coding System [5].

Concerning an observer’s perception of emotional expressions, we have adopted the hypothesis proposed by De Bonis [4] that the upper and lower parts of the face function as the building blocks at the basis of emotion perception, rather than the building blocks of finer granularity postulated by other authors. We have thus taken the most telling feature of each part of the face—eyebrows and lips—to express emotions in Feelix, making the hypothesis that these two features should allow humans to recognize its emotional expressions. Feelix’ face is thus closer to a caricature than to a realistic model of a human face. Moreover, also following [4], we have assumed one of the parts of the face to be dominant for some emotions (e.g., the upper part in fear and sadness, the lower part in happiness and disgust); in those cases, the perception of the dominant part expressing a given emotion should be enough for a human observer to recognize Feelix’ expression. In other cases, the perception of the expression in both parts of the face is needed for its recognition.

2.2 Emotion activation

Emotions are complex phenomena that involve a number of related subsystems and can be activated by any one (or by several) of them. Elicitors of emotions are for example grouped in [10] under the categories of neuro-chemical, sensorimotor, motivational, and cognitive. Some of these elicitors are emotion-specific, but emotions also show a certain degree of generality [14] (e.g., of object, of time) that accounts for the fact that a person can experience the same emotion under different circumstances and with different objects. But if emotions show this generality, what accounts for the activation of different affects? As we already mentioned, activation theories that only take into account the arousal and valence properties of emotions, are not able to fully account for their differential activation. To overcome this problem, Tomkins [14] proposed three variants of a single principle:

---

3In [4], emotion perception is a two-step process of probabilistic nature in which expressive features correspond to separate configurations of the upper and the lower part of the face. Due to technical constraints detailed in Section 3, we didn’t integrate enough expressive features in Feelix face to be able to implement this probabilistic model. We therefore decided to take the most expressive feature of each part of the face.
• **Stimulation increase.** A sudden increase in the level of stimulation can activate both positive (e.g., interest) and negative (e.g., startle, fear) emotions. The degree of suddenness accounts for the differential activation of these emotions.

• **Stimulation decrease.** A sudden decrease in the level of stimulation only activates positive emotions such as joy.

• **High stimulation level.** A high level of stimulation sustained over time only activates negative emotions such as distress or anger.

The fact that this model is based on different patterns of stimulation makes it particularly suited for emotion activation in our robot, since we were aiming at physical interaction based on tactile stimulation. However, an analysis of the possible interactions that humans could have with Feelix revealed two cases that the model did not account for. On the one hand, the model did not propose any principle to activate negative emotions such as sadness or boredom. On the other hand, happiness is only considered in the sense of relief resulting from stopping a very high (and therefore annoying) level of stimulation, i.e., as the cessation of a negative stimulus. Happiness, however, can also be produced by positive stimuli, such as gentle interaction. We have therefore refined the model by postulating two more principles:

• **Low stimulation level.** A low level of stimulation sustained over time produces only negative emotions such as sadness.

• **Moderately high stimulation level.** A moderately high stimulation level produces positive emotions such as happiness.

One of the drawbacks of this model lies in the fact that, being of general nature rather than associated with specific stimuli, some of the patterns can activate more than one emotion; another element is therefore needed to differentiate among emotions activated by the same general pattern. In the simulated creatures presented in [2], this was achieved by means of a synthetic physiology that allowed to associate to basic emotions a set of physiological parameters specific to each of them (e.g., fear is characterized by high heart rate and low skin temperature, versus interest that is accompanied by low heart rate). Since we wanted to do all the computations on board, simulating a rich enough physiology was not appropriate. We therefore used a simpler discrimination criterion based on the intensity
with which each pattern occurs (e.g., the stimulation increase is higher in the case of fear than in that of surprise). This criterion was sufficient to characterize the subset of emotions we used in Feelix.

3 EMOTIONS IN FEELIX

Feelix makes the different facial emotional expressions (see Figure 2) by means of two eyebrows and two lips. The eyebrows are controlled using an angle sensor and one motor and move symmetrically. They are two slightly bent LEGO parts that resemble the bent shape of human eyebrows, and are attached at their long end to a shaft around which they rotate. Each lip is controlled by an angle sensor and a motor, and therefore they can move up and down independently. A lip is a flexible rubber tube that can curve both ways. The mouth can be made narrow or wide by symmetrically moving its corners inwards or outwards by means of another motor.

The face is controlled by a LEGO Mindstorms RCX™ computer (www.legomindstorms.com). An RCX has a Hitachi H8/300 CPU and 32K RAM, and it has three input ports and three output ports. It can have limited communication with other RCXs or with a PC via an infra-red channel. Each angle sensor is connected to an input port, one sharing its port with the touch sensor. To control the four degrees of
freedom (DOF) of Feelix’ face by means of only three output ports, we have arranged the four motors in two pairs, eyebrows/mouth width and upper lip/lower lip. A fifth motor switches control between these two pairs of motors. Due to this switching between pairs of motors, eyebrows and mouth cannot move simultaneously when making a facial expression. Taking advantage of this asynchrony, we decided to start expressing an emotion first in the corresponding dominant part of the face, then in the non dominant one, hoping that, at least for some expressions—happiness, sadness, and surprise—it would be possible to guess the emotion being displayed before the expression is complete.

With more motors, it would have been possible to build an even more expressive face, either by increasing the number of DOF of the existing elements (e.g., moving the eyebrows up and down, moving the corners of the mouth asymmetrically up and down) or by adding other expressive elements (e.g., eyelids). In that case, however, the face would have been much heavier, requiring a bigger body, and its increased complexity would have required more than one RCX to control it, which would have had a negative impact on its performance in case of noisy communication between RCXs. By limiting the DOF to four, the face can distinctively display the five basic emotions we chose while being controlled by only one RCX.

A second RCX controls the interaction with humans and communicates with the RCX controlling the face. We wanted the interaction to be as natural as possible, and since for this project we are not using Feelix as a mobile robot—the human is sitting in front of it so as to better observe the face—the feet seemed to be the best location for tactile stimulation, as they are protruding and easy to touch. We built two special feet for Feelix using touch-friendly (smooth, large, and rounded) LEGO parts. Underneath each foot is a binary touch sensor—pressed or not-pressed.

### 3.1 Displaying emotional expressions

Each emotional state has an associated distinctive prototypical facial expression (shown in Figure 2), characterized as follows:

- **Anger**: raised eyebrows, moderately open wide mouth with upper lip curved downwards and straight lower lip.

---

4When we talk about raised or lowered eyebrows, it is in fact their external ends that are raised or lowered, since the internal end of each eyebrow is attached to a shaft and therefore eyebrows can only rotate on that axis.
Fear: lowered eyebrows, moderately open wide mouth with upper lip bent downwards and lower lip slightly curved upwards.

Happiness: straight eyebrows, closed wide mouth bent upwards.

Sadness: very lowered eyebrows, closed mouth bent downwards.

Surprise: highly raised eyebrows, very open narrow mouth.

These expressions are slightly modified by moving the expressive feature in the corresponding half of the face—eyebrows or mouth—when the emotion reaches a very high intensity.

3.2 Implementing the model of emotion activation

Interaction with Feelix is only through tactile stimulation on the feet, causing the touch sensors to be pressed. To distinguish between different kinds of stimuli, we use duration and frequency of presses. As for duration, it was enough to define three types of stimuli to implement the general stimulation patterns in our model: short (less than 0.4 seconds), long (up to five seconds), or very long (over five seconds). Frequency is calculated on the basis of a minimal time unit of 2 seconds that we call chunk. When a chunk ends, information about stimuli is analyzed and a message encoding Feelix’ current emotional state and its intensity is sent to the RCX controlling the face, so that the emotional expression can be updated if necessary.

Although it is possible to combine two expressions in Feelix’ face, we have as for now adopted a winner-takes-all strategy based on the level of emotion activation to select and display the emotional state of the robot. Emotions are assigned different intensities calculated on the grounds of stimulation patterns. At the end of each chunk, the emotion with the highest intensity determines the emotional state of Feelix. However, for this emotion to become active and get expressed, its intensity has to reach a certain threshold. By setting this threshold higher or lower, we can give Feelix different temperaments—i.e., make it more “extroverted” or “introverted”. When a new emotion becomes active, it temporarily inhibits all the other emotions by resetting their intensities to 0.

Emotion intensities are calculated by an update function that depends on time, and reflects some of the distinctive features of basic emotions, namely quick onset and brief duration [6]. The intensity of the active emotion increases with appropriate stimulation depending on how long this emotion has been active. Intensity
increases fast within a period of about ten seconds after the onset of the emotion until it reaches a high level; the increase is then close to 0 for a period of forty seconds; the increase is negative thereafter until the intensity drops below the activation threshold. An emotion of which the intensity has just dropped below the threshold will therefore be easily reactivated if the same type of stimulation persists. This is to reflect the fact that emotions have a limited, short duration\(^5\) (from a few seconds to few minutes in humans), although they can be repeatedly reactivated. All emotions increase their intensities with stimulation except sadness, which is produced when Feelix gets no attention. A time decay function makes emotion intensities decrease when Feelix is not being stimulated. For sadness, this function applies only after a long period of inactivity, when its intensity has reached its highest level, to reflect again the fact that emotions have a short duration. When no emotion is active—i.e. above the activation threshold—Feelix displays a neutral face.

We mapped the general stimulation patterns from our model into tactile stimulation patterns as follows.

- **Stimulation increase** is achieved by frequent short presses on any of the feet. This pattern can give rise to two emotions, surprise and fear. *Surprise* is produced by a less intense increase, i.e., one or two short presses after a period of inactivity or low/moderate activity. Surprise cannot reoccur within a short period of time. We therefore have an inhibition mechanism that suppresses the activation of surprise for some time after an episode of this emotion. *Fear* is produced when the increase is more intense, needing more frequent short presses to become active.

- A **sustained high stimulation level** overwhelms Feelix and produces anger. Very long presses, lasting three or more chunks, or many frequent short presses increase the intensity of anger.

- A **moderate level** of stimulation that neither overstimulates nor understimulates Feelix produces happiness. This level is achieved by gentle, long (but not too long) presses. The model we presented in Section 2.2 also incorporates a pattern for happiness in the sense of relief, i.e., the cessation of

---

\(^5\)Affective states with a long duration are called moods. They are much less intense and have different elicitors and much weaker associated (behavioral, physiological, etc.) manifestations; they make more likely the onset of specific emotions. Emotional states with an intensity below the activation threshold can be seen as moods in Feelix.
a too high, overwhelming stimulation level. We have not implemented this pattern in Feelix yet.

- Sadness is produced by a sustained low level of stimulation. As we already mentioned, in Feelix’ case this corresponds to a lack of (or very little) interaction for a long period.

The amount of stimulation required to change Feelix’ emotional state and its expression depends on the intensity of the currently active emotion—the more intense the emotion, the more stimulation is needed for a change to happen.

This model of emotion activation is implemented by means of a timed finite state machine (FSM) with three states: Idle, Pressed, and Released. Three timers are used to measure the chunks, the duration of a stimulus, and 5-second periods (the minimal length of a very long stimulus), respectively. During each chunk the FSM will circle between the Pressed and Released states, counting the number of short and long stimuli. The duration of a very long stimulus is calculated across chunks, by counting the number of 5-second periods that the stimulus lasted. The FSM returns to the Idle state and awaits the next stimulus either at the end of a chunk (i.e. after two seconds) or when a stimulus longer than two seconds ends. Upon returning, the intensities of the different emotions are updated according to the number of short, long, and (when appropriate) very long stimuli.

4 EXPERIMENTS AND RESULTS

To evaluate the recognizability of the facial expressions of Feelix we designed three tests. The first one is a free test—no list of emotion adjectives or any other cues are provided—in which subjects are asked to label a sequence of five expressions performed by Feelix: anger, sadness, fear, happiness, surprise. The second test is a multiple-choice one in which subjects are asked to label the same sequence of expressions, but this time they are given a list of nine emotion descriptors including four extra ones: disgust, anxiety, pride, worry. In addition, to test whether subjects can recognize the valence of the emotion, for each emotion they are asked whether they think the expression is elicited by something Feelix likes or by something it does not like. As a form of control, we also designed a free test where subjects are asked to label emotional expressions from pictures of human faces, namely anger, sadness, happiness, fear, contempt, surprise, and disgust.
We have conducted two suites of experiments. The first suite was conducted on four groups of children (5–6 children each) aged 9–10 years. Due to time constraints, only the first two tests were performed in this case. In the second suite we performed the three tests on a group of twenty adults (students and employees at the Department of Computer Science in Aarhus University), ages 24 to 57. Answers were considered to be correct when the subjects used the same descriptors we have employed or very close synonyms.

The results we obtained were surprisingly close to those commonly reported in the emotion literature on recognition of facial expressions of basic emotions in cross-cultural studies (see for instance [3] for an overview of these studies), in particular in the free tests. In the robot tests, the emotions that were most easily recognized were anger, happiness and sadness, while the results for fear were very poor (slightly better in the multiple-choice test than in the free one, unlike the other emotions). Recognition of surprise was only slightly over 50% in adults, less than 40% in children. Results in the multiple-choice test were worse (substantially so for children) than in the free one for anger, happiness, and sadness, while they were better for surprise and fear. Words thus seem to be confusing for emotions that are intuitively very easy to recognize; these emotions are also the ones that have better recognition results in cross-cultural studies (over 90% in some cases), and which are listed in most classifications of basic emotions. Fear is also commonly considered as one of the core basic emotions, but its recognition in humans is usually more difficult (figures turn around 65%, which is pretty close to the 60% we obtained in the test on human faces). Our poor results with the fear expression might indicate that more expressive resources are needed to display this emotion, in particular the ability to move the inner eyebrow up and down, but this would require one more DOF, and therefore more RCX. We plan instead to associate sound to the facial expressions as an additional means of conveying emotion. Comparing the two free tests, results on sadness (85% success on both Feelix’ face and the human face for adults, 76% on Feelix’ face for children) and anger (70% success on the human faces and 65% on Feelix’ face for adults, 71% for children) were basically the same, with results for happiness being nearly so (95% success on the human face, 80% on Feelix’ face for adults and 86% for children).

Average recognition of emotional expressions was 71% for adults and 66%

\footnote{These figures exclude results for fear in the robot tests and for contempt in the human faces one, since all subjects agreed that these expressions very bad (results were close to 0%). Their inclusion lowers figures by about 10 points.}
for children in the free test on Feelix’ face, 66% for adults and 53% for children in the multiple-choice test on Feelix’ face, and 78% (only for adults) in the test on pictures of human faces.

Valence recognition was very high (close to 100%) in all cases except for the always controversial case of surprise, which was attributed a negative valence by about 2/3 of the subjects, a positive one by the rest. This matches findings from studies on human faces (see [6] for an account) that surprise is perceived differently than other emotions, not defining an exclusive category (e.g., it is not always distinguished from fear, as it often happened in our tests) and without a clear valence.

We have not yet performed any formal tests to evaluate the plausibility of the emotion activation patterns in the interaction with Feelix. So far we have only observed people spontaneously interacting with the robot, or trying to guess the stimulation patterns we used to elicit different emotions in it. Some interaction patterns (those of happiness and sadness) seemed to be very natural and easy to understand, while others presented more difficulty (e.g., it takes more time to learn to distinguish between the patterns that cause surprise and fear, and between those that produce fear and anger). However, these informal observations do not allow us to draw any conclusive results. We are currently designing more formal experiments to further investigate this aspect.

5 RELATED WORK

Very little work has been carried out using emotions and their facial expressions in the context of human-robot interaction. Cynthia Breazeal has developed Kismet as a testbed for learning social interactions in situations involving an infant (the robot) and her caretaker (a human). Kismet is a head with active stereo vision and configurable facial features—eyelids, ears, eyebrows, and a mouth. Humans can interact with it either by direct face-to-face exchange or by showing it a toy. In [1], Breazeal reports on some experiments where Kismet uses nine different facial expressions to manipulate its human caretaker into satisfying its internal drives—a social drive, a stimulation drive, and a fatigue drive. Since the focus of this work is on social learning, no experiments to test the expressive aspects of the interaction have been reported so far. However, given that we have taken almost opposite working hypotheses concerning expressive facial features (a rather sophisticated, configurable set of features in her case, versus a minimalist one in ours), it would be highly interesting to perform common tests on human recognition of the robots’
expressions and to compare our results.

Minerva, developed by Sebastian Thrun, is a mobile robot that gives guided tours to visitors of the Smithsonian’s Museum of American History [13]. It displays emotional states—neutral, happy, sad, and angry—using a caricaturized face and simple speech. Emotional states arise as a consequence of travel-related interaction (e.g., anger results from its inability to move due to the presence of people), and their expressions aim at affecting this interaction towards achieving the robot’s goals—traveling from one exhibit to the next one, engaging people’s attention when describing an exhibit, and attracting people to participate in a new tour. Although [13] reports very successful interactions that the authors attribute to empathetic feelings in people, it also states that emotions in Minerva are purely a means to an end and not an integral part of the robot’s architecture and interface.

A surprising experiment, although using a computer instead of a robot, was conducted by Clark Elliott [7] to test the computer’s ability to express emotions by having humans recognize them. The computer used both caricaturized facial expressions and voice inflection to convey different emotional states while saying sentences devoided of emotional content. As a control, he had an actor say the same sentences and express the same emotions. It turned out that humans performed substantially better when recognizing the emotions expressed by the computer (70% of success) than those expressed by the actor (50% of success). Elliott suggests that these results might be partly due to the use of caricaturized expressions.

6 CONCLUSIONS AND FUTURE WORK

We have presented early work on Feelix, a humanoid-looking LEGO robot capable of displaying several emotional expressions in response to direct physical stimulation. Feelix implements two complementary models drawn from the psychological literature on human emotion—one concerning “universal” facial expressions of basic emotions, the other postulating a principle for emotion activation based on general stimulation patterns that we however associate to discrete basic emotions. We have conducted some experiments to assess how well humans can recognize emotional expressions in Feelix’ face. Our results match quite well results reported in the literature on emotion recognition from pictures of human faces. They also show that the “core” basic emotions of anger, happiness, and sadness are most easily recognized, whereas fear was mostly interpreted as anxiety, sadness, or surprise. This result might be due to the need of additional expressive
features. We are currently working with a composer of the Musikhuset in Aarhus to associate sound patterns to the different expressions, in order to emphasize the emotion conveyed. In addition, we intend to implement the emotion activation model using other sensory modalities. Special ears are currently being developed to enable Feelix to respond to auditory stimulation (e.g., claps). Finally, to obtain a more sound analysis of Feelix’s emotional expressions and a better comparison with studies of recognition of emotional expressions in human faces, we plan to analyze the robot’s expressions in collaboration with psychologists in De Bonis group at the University of Paris-XI.

ACKNOWLEDGMENTS

We are grateful to Henrik Lund, Ole Caprani, Brian Mayoh, and the students at the LEGO-Lab for their insightful suggestions and their cheerful support. This research is funded by the LEGO-Lab, Dept. of Computer Science, University of Aarhus.

References


