Identity Emulation (IE)

Bio-inspired Facial Expression Interfaces for Emotive Robots

D. Hanson, G. Pioggia, S. Dinelli, F. Di Francesco, R. Francesconi, D. De Rossi

D. Hanson, University of Texas at Dallas, USA
G. Pioggia, D. De Rossi, Center "E. Piaggio", University of Pisa, Italy
F. Di Francesco, R. Francesconi, Institute of Clinical Physiology, Pisa, Italy
S. Dinelli, National Health Service, Rome, Italy

Abstract

Our facial expression robots use biomimetic structures, aesthetic design principles, and recent breakthroughs in elastomer material sciences to enact a sizable range of natural humanlike facial expressions. Applications of this class of human-robot interface devices will rise in relevance as humans and robots begin to have more sociable encounters in the coming years. The Identity Emulation project also seeks to imbue these robots with several forms of interactive intelligence, including human-form and facial-expression recognition, and natural language interaction. It is anticipated that a multi-modal integration of mechanics, sociable intelligence, and design aesthetics will yield the most effective Human Computer Interface robots.

1. Introduction

The nonverbal expression of the emotions, especially in the human face, has rapidly become an area of intense interest in computer science and robotics. Exploring the emotions as a link between external events and behavioural responses, artificial intelligence designers and psychologists are building the foundational principles which now appear critical to the physical embodiment of artificial intelligence. The Identity Emulation project at UTD and the FACE project (facial automaton for conveying emotions) at Center Piaggio in Pisa, Italy are seeking in parallel to achieve more naturalistic interface devices for human-AI interactions. In the AI community, interest in emotions and in their expression and communication has developed in a growing number of applications: animated characters, expressive interfaces, music-dance-drama applications, systems of human-computer emotional interaction, and expressive and social robots, intended in more or less "naturalistic" terms under the aesthetic respect, etc.

Special emphasis is given in this paper to nonverbal expressions of emotions. Some examples of different approaches to this area of research will be considered, with no pretense of completeness, but just intending to indicate some examples of interesting different strategies in the field of emotional communication between robots and humans. Furthermore, we will consider methods and means for advancing expressive-emotive AI systems, and methods by which we are developing Electro-Active Polymer (EAP) artificial muscle actuators to assist in emulating paralinguistic communications.

One IE robot built by the authors is now actively deployed as a test platform for EAP actuators at JPL's Non-Destructive Evaluations and Advanced Actuators (NDEAA) Lab [Bar-Cohen, 2002], (see FIGURE 1). Advanced EAP materials with properties more like biological tissue may lead to tremendous leaps in biomimetic technology. Conversely, considering the difficulty of matching natural facial expressions, communicative biomimesis is proving a helpful application for researchers by serving as the high-bar of achievement standards.

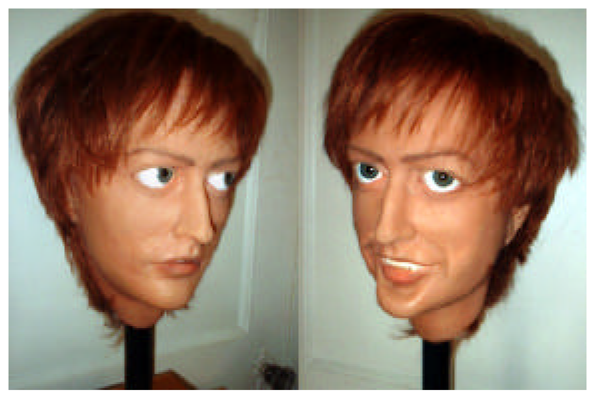

FIGURE 1: Andy-Roid robot, testing artificial muscles at NASA's Jet Propulsion Lab. Sculpted and built by David Hanson, with contribution from Giovanni Pioggia.

In the quest to humanize computers, emulating the subtlety of human physical expression certainly ranks as one of the most ambitious and challenging tasks. Much of the technical and aesthetic groundwork for this goal has been established in the entertainment special effects industry and the classical fine arts (these techniques are discussed later in this section). Though techniques available today are substantial, they can't yet match the expressive power of the natural human face. Much work remains to effectively render human nonverbal emotional expression in all its complexity.

It is clear that the first front of intelligence occurs within our input and output devices. Our facial expressions deliver much of our emotions and thoughts (the two differ only as data types) to our community, facilitating socially organized intelligence (a.k.a. civilization). A large portion of our brain responds to the immediacy of this input and output; in fact it is well demonstrated that the brain requires regular interaction with an external world to function properly. Human and animal intelligence requires the push and pull of a sensory and social world to facilitate its emergence and maintenance. One may assume that likewise, the more effective a computer's integration with the external world (including in no trivial part, sociable civilization), the more effective the computer intelligence will be as it emerges and evolves. In effect: powerful facial expressions will prove critical to the development of intelligent, sociable robots.

2. The Semiotics and development of emotional expressions in humans

The human face and its kinesics are among the most powerful mediators in emotional elicitation and involvement. Innately, our emotional facial expressions weave together the human emotional experience, and in so doing, create the fabric of social bonds. If robots can master the "emotional language" of facial expressions, it will put people feel at ease with the robots' presence, and add tremendous value to sociable robotic applications. The functionality of such robots will spring from humanity's innate paralinguistic system and its physical mechanisms.

There is an immense body of scientific studies of the expressions of humans and their associated emotions. Darwin was the first to advance the hypothesis that expression of various primary emotions are universal and easily recognizable as specific physiological patterns, and that these are these innate and common to humankind. In The Expression of the Emotions in Man and Animals ", Darwin argued the same state of mind is expressed throughout the world with remarkable uniformity." [Darwin and Ekman, 1998/1872]. A large amount of research has amassed since, which has reliably confirmed this hypothesis for basic emotions (both in terms of mimic patterns and of recognition), also on a transcultural basis. [Izard 1971, Eibl-Ebesfeldt 1973; Ekman 1989]. This research went along with the Ekman's studies to construct a system for detailed description of emotional expressions, the Facial Action Coding System (FACS). At the beginning FACS was one of the sources used by computer science and robotic researchers to explore the field of emotional expression and communication.

A large amount of ethological and psychological researches about newborns and babies revealed the developmental evolution of mimic expressions, in terms that are by now largely shared by many different authors [Anolli, 1995]. The field has been deeply explored also thanks to audiovisual recording, which permits analytical, dynamic and contextual accurate observation. Clear reactions of pleasure, disgust, discomfort and response to threatening or disorienting stimuli, appear at birth in the newborn as an innate response to specific environmental stimuli. In the following weeks manifestations of interest and selective attention towards new stimuli and the human face appear. Successively, (between second and twelfth month) the emotional system develops various patterns of social interaction, including social smile, happiness, sadness, anger and rage, whereas, in the second part of the first year, fear and wariness appear. A third phase develops since the beginning of the second year, when expressions of shame, guilt and shyness, emerge, followed by contempt, and, later, by mixed emotional expressions. These are basically learned by the rearing and socialization experiences the baby has been and is going through. The studies on infants in fact also revealed a substantial similarity in the mimic and motor configurations between the child's expression and the adult's. Such is the basis of the beginning emotional dialogue between the adult and the infant, with the adult attributing meaning and emotional intentionality to the motoric and mimesic manifestations of the child. This dialogue evidently has many implications: the expression of emotions is in fact an important mediator of social exchanges and has a crucial role in learning and personal development.

Many labs around the world are now addressing the problem of emulating human paralinguistics and affect in AI and robotics, with a wide variety of approaches.

3. Standing robotic approaches to nonverbal expression in human-robot interaction

The basic systems of human paralinguistic communicationas are well understood. The challenge, then, is to emulate these systems technologically. While this challenge falls largely into the domain of controls and perception, the physical manifestation of robots' paralinguistic communication is an undertaking in aesthetics, mechanics, materials, and manufacturing. It should be obvious, however, that an effective sociable-robot will be an integrated design: a hybrid of AI, mechanism, and aesthetic design. In support of the emerging technology of affective computing—the AI of sociable robots—researchers are working to create Human Computer Interface (HCI) devices capable of displaying the emotional inner state of a machine, by mimicking nonverbal expressions native to human communications. Researchers have developed a wide diversity of directions in conceiving biologically inspired robots, of which sociable agents constitute a very important subset.

Pioneer work on this subject was carried on at MIT in the mid-1990's by Breazeal and Scassellati [2000]. These authors designed a robot (Kismet) to interact socially with human "parents". In human dyads infant-parent, emotions and drives play an important role in generating meaningful interactions, regulating these interactions to maintain an environment suitable for the learning process, and assisting the caretaker in satisfying infant's drives. To achieve a similar dynamic, the authors designed a framework that integrates perception, attention, drives, emotions, behaviour selection and motor acts. The aim is to

> "…design a robot that is biased to learn how its emotive acts influence the caretaker in order to satisfy its own drives. Toward this end, we endow the robot with a motivational system that works to maintain its drives within homeostatic bounds and motivates the robot to learn behaviors that satiate them. Further we provide the robot with a set of emotive expressions that are easily interpreted by a naive observer…"

The robot responds with expressive displays, which reflect an ever changing motivational state and give the human cues on how to satisfy its drives, neither overwhelming nor under-stimulating it. In this workframe, the "emotions" of the robot are always active, but their intensity must exceed a threshold level before they are expressed externally. They serve two functions: 1. They influence the emotive expression of the robot by passing activating energy to the face motor processes. 2. They play an important role in regulating face-to-face exchanges with the caretaker.

As the authors evidence, the organization and operation of their framework is heavily influenced by concepts from psychology, ethology and developmental psychology, as well as the applications of these fields to robotics [Brooks, Ferrell, Irie, Kemp, Marjanovic, Scassellati & Williamson, 1998]. In this remarkable architecture the aesthetic aspects of emotional expressions are not in the foreground and are at the moment reduced to meaningful but aesthetically essential elements.

Inspired by Kismet, a similar affective computing system is now employed at the Sony Computer Science Lab Paris to endow the Sony SDR-4X with emotional responsiveness. Using AIBO dog-robots, this affective model is being tested as the emotional framework for automatic speech acquisition using biologically inspired language games [Steels, 2001], [Steels, 2002]. Other labs addressing sociable and Affective Computing include MIT Media Lab's Synthetic Characters division, lead by Bruce Blumberg, and CMU's Affective Computing Lab,

Another integrated approach can be found in the Gesture and Narrative Language Research Group at MIT, directed by Justine Cassell, which incorporates narrative studies, linguistics, natural language processing and affective computing into full affective AI agents which are represented by CGI avatars in narrative interactions with people [Cassell, 2002]. These sociable avatars also respond to gestural and other paralinguistic cues integrated into the spoken language interactions; such multimodal communications are a powerful feature which is characteristic to natural biological communications.

We can find a different approach to the social exchange between humans and robots for instance in the work of Camurri and Coglio [1998]. Their robot is thought to act in various contexts, one of which was the Music Atelier of the Children's Town in Genoa. This architecture includes five components: input, output, reaction, rational and emotional. Emotional component holds the agent's emotional state, whose temporal evolution is governed by emotional stimuli from input, reactive and rational components. Here the "emotional state" consists in coordinates of a point in some "emotional space", partitioned into zones (8 sectors of mood, each with intensity levels), whose bords can be fuzzy. Physically the Robot is an Active Media Pioneer 1: it communicates with children in a play environment, through lights, style of movement, speech, sound and music, with variations related to its " emotional state". Also in this case the aesthetic aspects are very essential, based on a kind of "credibility" of the agent, to which in fact children playfully respond.

A different but also very interesting approach to the creation of agents emotionally and socially meaningful to humans is research aimed at endowing robots with more naturalistic mimic expressions, to which human are very sensitive. Researches of this type touch the possibility of elaborating structures and actuators apt to reduce the distance from the subtle to the refined textures of human mimesics. It is true that by now researchers are working on the basis of a narrow number of basic expressions (this is already a big challenge). But human emotional expressions have indeed an extraordinary complexity, as evidenced by the great descriptive work made by Ekman and others; this is also evident from the complexity of facial expressions' musculo-skeletal physiology. So, advancements in plasticity and accuracy also appear relevant in terms of possible social interaction between AI agents and human agents. Such refinement can be important in reducing effects of "expressive rumor", and creating a meaningful and interesting context in this type of exchange. Entering in this subject it is useful exploring the physical basis of human face expressive potential.

While anthropologists have largely described the basic semiotics of paralinguistic communications, Ekman points out that there exist distinct, large codices of subtle expressions which are culturally dependent, and have not yet been scientifically formalized. Such culture-specific subtleties will be highly relevant to nonverbal communications robotics, and their implementation need not necessarily await the rigor of scientific formalization. Such complex problems in aesthetics and communications biomimetics are already solved frequently by animation artists, relying on natural human talents for communication instead of scientifically formalized principles.

The heuristics of artistic talents can be considered a "black box" problem-solver for tasks of aesthetic biomimesis. The trained artist simply extends the natural human facility for communication into devices and materials (see FIGURE 5 for an example of aesthetic biomimesis). This black-box approach has proven essential to getting sociable and communications biomimesis applications to market, as demonstrated in the design of biomimetic toys, film special effects, video games, and sundry devices [Frauenfelder, 2002]. In addition to artist's static design work (sculpture, painted surfaces, etc.), artists are also integral to animating biomimetic applications. The power of biomimetic aesthetics in the hands of animatronics artists is easily to observed in major motion pictures, such as the "Jurassic Park" films, "Mighty Joe Young", and of particular relevance, the film "AI".

FIGURE 2: "Mask", a 5' tall self-portrait by artist Ron Mueck shows the sophistication of aesthetic biomimesis, which could enhance robots. Photo courtesy of Mark Feldman and 8am.com.

Stan Winston Studios, Walt Disney Imagineering, Jim Henson Creature Shop, among dozens of other "animatronics" (themed animation robots) shops, already apply nonverbal facial and body language regularly in robotically animated narrative arts. Additionally, computer simulations of facial expressions are becoming increasingly sophisticated and will be useful for flat-display based affective computing. As the manifest purposes of animatronics, aesthetics and narrative steer the technical implementation of these robots. To this end, animatronic robots have pushed the threshold of communicative biomimesis as few academic endeavors have. Integrating visual and narrative artistry, mechanics, elastomer skins, electronics, and controls (increasingly empowered by Artificial Intelligence), animatronic robots now show a portfolio of novel innovations that could be quite pertinent to communicative and expressive robotics development. In addition to the craft and art that animatronics offers to sociable robots, the entertainment market also represents the leading, most proven sector for deployed biomimetic robot applications. The IE robots spring from the foundation of knowledge and skills of film-effects animatronics.

Leonardo, the new sociable robot of MIT's Artificial Intelligence Lab exemplifies the merger of film special effects artistry and the AI community. Designed by Cynthia Breazeal with Stan Winston Studios, and contructed by Stan Winston Studios, this robot infuses the knowledge of the animatronics industry into one of the leading sociable robotics research labs. Leonardo, a furry creature, with a highly simplified face, intentionally avoids the visual paradigm of the realistic human face. This is useful, since humans are so attuned to the human face, they are particularly demanding in the qualities of its representation. Yet there is much to be said for striving to achieve greater realism in robotic faces. Human facial systems communicate huge quantities of information very quickly [Darwin and Ekman, 1872/1996], and robots will be greatly furthered by emulating the subtle extremes of these systems. This is the ambition of the Identity Emulation and FACE endeavors.

Since our senses are so finely attuned to healthy movements generated by highly evolved biological facial tissues, it has proven challenging to achieve full aesthetic biomimesis using standard engineered materials alone. Scientists, engineers and artists are now attempting to define a new, more biomimetic mechanical engineering, based on wet, soft, and viscoelastic materials.

4. Myo-Physiology of Emotional Expressions in the Face

To emulate the appearance of dynamic facial expressions, it can be useful to consider the mechanisms by which our flesh effects these expressions. Figure 1 shows the muscles responsible for facial expression [Ratner, 2000].

Facial muscles can be divided in five groups: epicranius muscle, muscles of the eyelids, muscles of the lips, auricularis muscles and muscles of the nose. In humans, all aforementioned muscles are controlled via the facial nerve. First discussion will be focused on the first three muscle groups, the primary contributors to facial expressions.

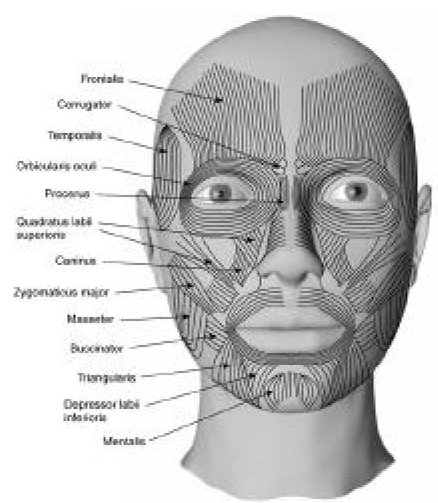

FIGURE 3: Facial muscles that produce expressions [Ratner 2000]

The muscles of the eyelids include the orbicularis oculi muscle and the corrugator supercilii muscle. The orbicularis muscle has an elliptic ring shape around the palpebral rima. It is possible to note orbital, palpebral and lachrymal parts. The orbital part is the most developed and forms almost complete ring around the eyelids. From the lateral corner of the eye some little fasciculus comes down to the cheek and to the zygomaticus muscle. The palpebral part is contained into the thickness of the upper and lower eyelids; its fibers take origin from the medium palpebral ligamentum. The lachrymal part is located more deeply; it is divided in two fasciculae, where respectively the upper and lower eyelids come to the palpebral part of the muscle. The orbicularis muscle by its contraction performs the closing of the palpebral rima, directs the tears to internal angle of the eye and dilating the lachrymal sac helps their downflow. The corrugator supercilii muscle is a thin plate with lower concavity locat ed nearby the eyebrow. It starts from the medium limits of the supercilii arch up to the derm; its location is deeply into the frontalis muscle and into the orbitalis part of the orbicularis oculi muscle. Its contraction moves down the skin of the eyebrow.

The muscles of the lips include the zygomaticus muscle, the buccinator muscle, the triangular muscle, the square muscle of the lower lips and the orbicularis oris muscle. The zygomaticus muscle starts from the lateral surface of the zygomatic bone up to the derm and to the labial mucous nearby the commissure, where fibers start to the orbicularis oris muscle. By its action it moves up and down the labial commissure. The buccinator muscle is a fleshy plate that constitutes a large region of the cheek. It starts nearby the molar teeth, its fasciculus come up to the labial commissure, where fits deeply into the derm and into the mucous. In proximity of the commissure insertion some fibers intersect so that some upper fasciculus come to the upper lips, while others come to the lower lips; these fasciculus contribute to the formation of the orbicularis oris muscle of the mouth. The buccinator muscle contracting shifts behind the labial commissure and sticks the cheeks and the lips to the dental alveolus arch helping the chewing phase. The triangular muscle is located below the labial commissure, it starts from the external face of the jaw, its fasciculus come up inserting in part into the derm of the commissure and in part into the upper lips where form part of the orbicularis oris muscle. Its contraction moves down the labial commissure. The muscle square of the lower lips is located more deeply with reference to the triangular muscle. It starts nearby this last one up to insert into the derm and into the membrane of the lower lips. Its movement consists of shifting below and laterally the lower lips. The orbicularis oris muscle is one of the mainly constituents of the lips, it is an elliptical ring placed around to the buccal rima and has an external and an internal part. The above muscular systems have been greatly simplified in the first of our IE robots.

5. Physical Emulation of Emotional Expressions

Since we intended the Identity Emulation robot to be a work of fine art in its application (the latest version being a self portrait of co-author David Hanson, see FIGURE 4), the first stage of IE's design involved the narrative and design conceptualization. Several questions were asked: how should the robot look? What movements and gestures are imperative to convey the identity of the subject? In addressing these questions, a series of sketches were produced, using both classical rendering tools of pencil, paints, and paper, and digital design tools of photoshop and Maya 3-D. Because our goal is to push the threshold of the robotic technology, we chose a realistic form. Nevertheless, as the author David Hanson rendered the form by eye and with his hands, expressive distortions characteristic to the artistic process added style to work of sculpture.

FIGURE 4: Robotic Self Portrait by David Hanson.

When building an expressive robot, one must have as a foundation, a strong visual and conceptual design. The tools of classical art can be highly useful for this task. Consider the visual appearance of Michelanglo's Pieta: enormously full of life, almost super-human in its expression; now imagine it furthered by effective robotic animation. Such a piece would be overwhelming in its expressive presence. The traditions of classical art and its training remain alive and indispensable throughout the world of design, including the animationa and animatronics (themed robotics) industry.

During the design phase, one must ponder and devise the story of the piece, the purpose and function of the robot's interaction with people, and the complex cloud of implications of its visual and narrative concept. These considerations should also shape the design of the intelligence systems (including the "personality" of the robot), in accord with the visual and mechanical design. Given the complexity of this task, an iterative visual design process is useful if not indispensable. The process of mechanical, electronics and controls systems design should be conducted in tandem with the visual and conceptual designs. For this reason, members of a design team will ideally be polymaths, capable of both visual art and technical design. Later, this section will examine techniques for mechanically emulating facial expressions, and techniques for controlling them. For now, next considered will be the physical execution of the external visual design, starting with the sculpture.

To realize the design as a sculpted three-dimensional form, the maximally subtle sculpting medium is desirable. At present, physical sculpting is faster and more responsive than any haptic sculpting interface system. Virtual sculpting is reaching maturity as a medium, but for now clay can be used more quickly, cheaply, and simply. Oil-based clays are the best choice, since they will not dry out and can be worked, reworked, and recycled. Before sculpting a large of complex form, an armature of wood, steel, orother rigid material, will be built to support the weight of the clay. The smaller IE faces shown in this article were sculpted sans armature. The 150% self-portrait, though, was rough-hewn from rigid urethane foam, a common insulation material; fine detail was then added using oil-clay.

FIGURE 5: The robot's mold.

To translate the sculpted animatronic figure into the rubber skin, a mold (negative impression) of the sculpture is made (see FIGURE 5). Usually this mold will have an inner surface of an elastomer (in our case silicone), surrounded by a shell, which is also called a "mother mold". This shell is composed of a hard, rigid material such as fiberglass, rigid urethane plastic, or classically, plaster reinforced with burlap. The rubber inner-layer will lock into the shell with keys, or raised bumps. The shell will be composed of several interlocking parts to avoid undercuts (regions where the sculpture or cast will lock in the mold); these parts will be held together with bolts.

To mimic the supple flexibility of human and animal skin, the robot's skin will be cast in one of various chemical elastomers. Urethanes and silicones are the most widely used elastomers for facial expression robotics, and each offers unique benefits. Urethanes, showing characteristics that actually lie somewhere between classic elastomers and viscous materials [Pioggia, 2002], exhibit properties that are quite like the natural visco-elastic behavior of animal tissue. In particular, urethanes will function like both springs and dampeners, with very quick responses, even when the particular urethane is very soft. Skin-Flex urethanes, manufactured by BJB Industries, are the gold standard of special effects urethanes. Easily painted and pigmented, Skin-Flex elongates to 1000% with a shore hardness as low as 3 to 5 durometer, shore A (as soft as human skin). The drawback of urethanes lies in their tendency to degrade with exposure to UV radiation. Silicones, on the other hand, are not sensitive to UV radiation and are very environmentally stable. While silicones are easily pigmented, they are not easily painted once the silicone's cross-linking is complete. And silicone as soft human skin silicone begins to exhibit kinetic induction—a delayed reaction between actuation and effective movement of the skin surface. In short, urethanes work better than silicones, but aren't as tough and don't last as long.

To effect the expressive movement, anchors in the skin provide mechanical connection-points for attaching the expressions' actuation mechanisms, while gaps in the skull permit actuators housed inside the skull to effectuate the skin movement. Anchors were composed of loose-weave fabric cast directly into the skin. Such fabric spreads force over a larger skin-area, while allowing the elastomer to retain its strength by cross-linking through the voids in the fabric's weave. Because the fabric inhibits the elongation properties of the elastomer in the anchor-zone, the size and placement of the anchors was critical to allow the skin to stretch into expression properly.

Several methods exist to impart movement to the anchors. The most common method is to attach linear actuators directly to the anchors; a similar approach attaches gimbols to the anchors, thus effectuating two axes of motion (three if the gimbal sits on an extending boom). In the author's experience, a third method works better than the previous two: composite fiber cables are cast into the skin, in lines that correlate to the contraction vectors of natural muscles. At one end, each cable attaches to an anchor; at the other, the cable slides through the skull to reach an actuator. Then, when actuated, the cable slides freely through the skin, transmitting force directly to the anchor. This method is more biologically accurate, in that it effectuates motion from the interior of the skin-mass, as do biological muscles; also, it allows layered vectors of actuation, mimicking the layers of natural facial musculature. To embed the cables and anchors, a first layer of 10 durometer shore A rubber (recommended: Skin-Flex special-effects urethane from Burman Industries) is painted into the face's mold in 2-3 layers, and is pigmented to suit the artistic design. While still tacky, the rubber actuation elements are placed upon this painted epiderm, and the force transmission elements are run through the skull's bushing and slots. At this time the skull is placed and registered, and all openings are plugged with wax to prevent leakage. Next, a special mix of elastomer and "foam-rubber" is poured to fill the skin-form, and allowed to set. This material, developed at the University of Texas by coauthor David Hanson in the spring of 2002, approximates both the elongation and the compression characteristics of human skin, enabling maximally lifelike movement and easy actuation by servos small enough to house in the skull cavity (FIGURE 6).

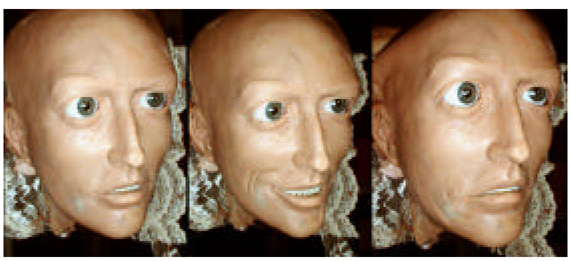

FIGURE 6: Facial Expression robot using elastomer-foam blend and embedded cable drives. Built by author David Hanson at the University of Texas at Dallas.

This material overcomes the one notable complication resulting from the dissimilarity of solid elastomers and liquid-filled animal tissue. A liquid-filled cellular matrix like human skin is as easily compressible as it is elongated, whereas elastomers require considerably greater force to compress than they do to elongate. This means that either enlarged actuators are required (rendering electrical actuation prohibitively heavy and bulky), or the skin-architecture needs to be modified. Several methods for such modification previously existed. The first embedded liquid or air-filled pouches in the skin, but this is complicated and costly. The second method involved the strategic reduction of skin-thicknesses, but this results is a skin which will not move like natural flesh.

The novel foam and rubber mix, which the author chooses to call "F'rubber", elongates to 600% (as opposed to standard urethane foam's 150%), and has compresion characteristics identical to those of the original urethane foam rubber. The resulting facial expressions, shown in FIGURE 6 are actuated with hobby servos producing mere19 oz.-inches of torque at 6 volts. This material promises to enable a powerful new class of highly expressive low-power mobile robots.

Once poured, the skin is attached to a skull-like substrate. In the process of sculpting this skull, the author achieved anatomically correct proportions relative to the original portrait by utilizing forensics data [Helmer, 1998] as reference for skin thicknesses. Notating these thicknesses with depth markers, the skull was sculpted in reverse (in negative) inside the mold of the sculpture of the face. The resulting variable skin thickness enables more lifelike dynamic skin distortions, more closely resembling natural facial expression. The skull-sculpture was then itself molded, and thin composite-plastic casts were made for the final robot (see FIGURE 7). The skull also serves as an inner-core to the mold so that when cast, the skin's inner and outer surfaces will conform tightly to both the exterior artistic design and the inner mechanism. Moreover, the skull also serves as the mechanical frame for mounting actuators and sensors.

Mechanically, the Identity Emulation heads are built using standard bearings, u-joints, and other common machine-components. If using conventional skin materials, one would need to use pneumatic or hydraulic actuators, which provide much greater force and speed. Neither pneumatics nor hydraulics are well suited for mobile robotics, however, since these actuators require large, off-board compressors to pressurize their air or fluid. Thanks to the highly plastic characteristics of the skin materials, mere hobby servos are used to actuate the expressions. To further enable mobile paralinguistic robotic interface devices, Electro-Active Polymer (EAP) Actuators may provide the next step of improved power densities for mobile animatronics [Hanson and Pioggia, 1999], as described in section 7. of this paper. During the interim, the IE project seeks to decrease the weight by sttempting to simplify the mechanical system, thinning the skin, and utilizing lightweight composite materials

When the mechanism and skin are assembled, the figure is dressed out for final display, and the AI controls are implemented and tested.

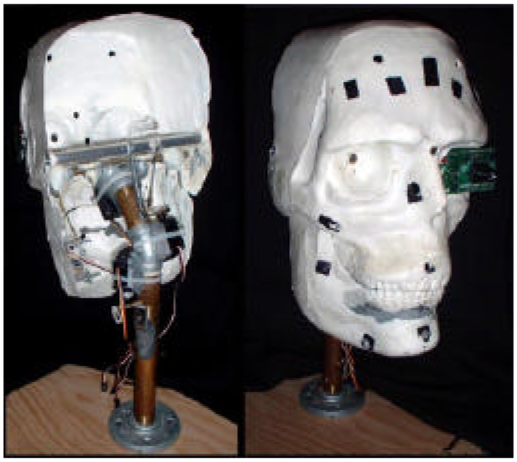

FIGURE 7: Skull and mechanical frame of the IE self-portrait, under tests using the CMUcam object-tracking algorithm.

6. Facial Controls: Moving Towards an Integrated Humanoid AI system

A major goal of both the IE at the University of Texas at Dallas, and the FACE project at the University of Pisa, is the integration of a complete multimodal intelligence and perceptual control system to govern the facial robots in sociable interactions. At this point, the IE robot effectively tracks colored blobs using the CMUcam system, and the robot can also lip-sync to prerecorded speech. We are working on implementing rudimentary natural language interaction using Carnegie Mellon's Sphinx, Automatic Speech Recognition (ASR) software, head-tracking, and facial expression recognition using the Eyematic SDK. The robot servos are currently controlled using Oopic microcontrollers. The driver of the FACE android consists of a shielded box containing seven cards, one supply for the EAP actuators up to 10 kV and one supply (5 V) for the logic controls. The unit is designed to in a modular configuration to allow flexible operation and future modifications. Each card contains a programmable microprocessor, four drivers with each having a standard RS232 serial interface. The protocol is organised using three bytes: the first for the address (0..255), the second one for the desired position (0..255) and the last one for the information of moving rate (0..255). The baud rate is 57.6 Kbaud with a command execution rate of 520 μs, all the actuators are driven using 12.5 ms signals. The block diagram for a single channel is shown in figure 2. The microprocessor acquires information about the desired deformation for the actuator including rise time and activation forces. An electronic comparator (A) and an encoder is added to provide a closed loop control, where at the moment the encoder is under development and the control is performed in open loop.

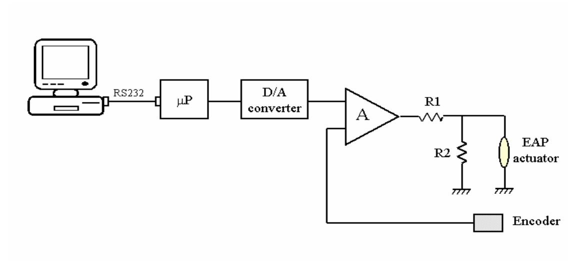

FIGURE 8: Block diagram of a single EAP channel

The actuators are controlled using various force amplitudes to mimic the dynamics of the human muscles. The aim is to model the operation of EAP actuators into the control system in terms of surface shape modifications that are needed to create the desired facial expressions. For this reason, human expressions are acquired by motion capture sequences, while perceptions and expressions are acquired directly from a human model. The acquired, stored patterns are used to manage the actuators to effectuate expressions and provide information input for the driver unit. A matrix of data of different expressions is acquired (for example: happiness, disgust and uncertainty) and chewing phases, wherein the columns are the actuators and the rows represent the scan rate.

For future improvements in the emulation of nonverbal expressions, following issues will be considered and addressed. The complexity of modes of human facial expression largely spring from the microscopic structural and physical characteristics of the organic tissues which compose the skin. Conventional elastomers cannot match the extremely soft yet tough characteristics of the human skin. Nor can it match the shearing effect among layers, intimately affected by differing Young's modulii of skin and flesh's various internal tissues. So it should be clear that this elaborate web of liquid-saturated cellular tissue would be tricky at best to emulate with the elastomers that are the staple of animatronics and bio-inspired robotics.

Notice that in a human face, wrinkles lie dormant all over, invisible until an expression is enacted. Notice also how opposing movements evoke radically different types of folding and bunching of skin. Clearly, these subtle qualities will be nearly impossible to attain without significant advances in technology. The use of organic modeling software and iterative Rapid Prototyping design cycles will facilitate increasingly elaborate and precise manufacture of facial tissue emulation. Such innovation, however, will not be trivial. The emerging tools of Smart Materials, MEMS manufacturing, Electroactuated Polymer (EAP) Actuators, nanotech, and other emerging technologies, may come to offer the means to realize such a lofty ambition.

In the short-term, however, smaller innovations certainly can and will enable friendlier and more endearing robots. Improved nonlinear control algorithms, the use of visco-elastic dampening [Full, 1999], and further study of the physiological actions of nonverbal expression using advanced medical imaging, all will improve the technology of robotic emulation of nonverbal expression. The most profound short-term improvement in robotic nonverbal expression, though, likely will come from a breakthrough in the field of EAP actuation research.

7. Developing EAP Actuators for Emotional Expression Emulation

EAP materials have characteristics attractive to the emulation of emotional expressions, offering the potential to more effectively model living creatures at significantly lower cost [Hanson and Pioggia, 2001]. Visco-elastic EAP materials could provide more lifelike aesthetics, vibration and shock dampening, and more flexible actuator configurations. Moreover, multiple studies have shown that the visco-elastic properties of animal tissue are key to locomotion and general stability [Full, 2000; Full and Koditschek, 1999; Dickinson, et al, 2000; Full, et al, 1998]. In these regards, EAP is the only actuation technology to date that earns the moniker "artificial muscle".

EAP offers many improvements over traditional actuation technologies, both in function and form. Functionally, EAP promises higher power densities than any other kind of actuator [Baughman, et al, 2000], even higher than that of the workhorse of heavy industry: hydraulic actuation. Being electric, EAP actuators would do away with bulky and fault-prone hydraulic pumps, and would eradicate the mess associated with hydraulic failure. Given the generally moderate cost of polymers, EAP actuators are likely to be more affordable than rival actuators. EAP actuators' unitary mechanical form allows them to be packed into smaller volumes. Volume would be further conserved by EAP's flexibility, and conserved even further should EAP's potential to be cast to form be realized. Such flexibility will explode the options for mechanical configuration of actuators. Absence of gearing and transmission reduces points of possible failure. EAP can act as dampener, insulating a mechanism from shock and damaging resonant vibrations. Elastic qualities will allow for spring-mass energy cycling, improved efficiency and stability in locomotion and gesture, and a generally more lifelike appearance.

Aesthetically, EAP's affinity to biological tissue is very appealing. EAP should allow more physiologically accurate biomimetic design; the movement of the resulting animation would look more inherently organic. Occam's razor supports this assertion; structurally rigid mechanisms certainly represent the long-route to simulating animal soft-tissue. EAP's softness and elasticity can be foreseen to eliminate the distracting surface-distortion caused by rigid actuator elements. Compliance and dampening characteristics of EAP closely correspond with those of biological muscle, simplifying controls and leading to more gesturally accurate animation. EAP actuators can be flat or bundled, like real muscle. Such formations could be imbedded directly into rubber skins—which are currently in widespread use for special effects. By layering these artificial muscles, the complex musculature of the human face could be emulated. Clearly such an array of EAP actuators would render much more subtle and intricate expressions than other contemporary actuation technology can.

No previous actuation technology has presented so many potential benefits to robotics and animatronics character animation. If EAP actuation proves practical, it will certainly revolutionize biomimetic robotics in all its forms.

Generally, useful characteristics of EAP actuators include low voltage, high strain, high power density, robustness and durability, and the support of fully developed adjunct technology (fasteners, drivers etc.). For practical use, the technology must also be reliable, affordable, and easy to use.

But even with effective EAP actuators, other obstacles need also be overcome. Utility of EAP in robotics requires a significant rethinking of mechanical engineering, which is conventionally based on principles of rigid structures—gears, u-joints, and the like. Control and feedback of these inherently compliant materials also pose a significant challenge [ibid].

In short, EAP and its adjunct technologies must mature considerably to actualize EAP's potentially enormous offerings to emotionally expressive robotics. Development of an anthroid face using EAP actuators is, arguably, the most important step in the introducing EAP technology to robotic nonverbal expression emulation.

8. Conclusion and Future Work

Initiated in February of 2002, the IE project has produced three functioning prototypes in the relatively short time of 7 months, and has shown a series of materials advances that has enabled a great reduction in actuator size and power requirements. It is anticipated that better mechanization of the universal visual-language of facial expressions will unlock panoply service and entertainment applications, from toys to comforting companions for the elderly. Even in a military scenario, wherein a robot must communicate swiftly with human soldiers, the power of emotive communications cannot be over-esteemed. Next step objectives of the IE project include the refinement of the manufacturing process, and more effective inclusion of multimodal AI perceptual systems, including facial biometrics, head-tracking, and facial expression recognition.

9. Acknowledgements

David Hanson would like to acknowledge Holly Yanco, Ian Horswill, AAAI, DARPA, UTD, Yoseph Bar-Cohen, Cynthia Breazeal, Thomas Linehan, David Najjab, Andy Blanchard , Kristen Nelson, Elaine Hanson, Paul McCarthy, and Michael Dobry.

References

Anolli, L. "Lo sviluppo emotivo precoce", in Di Blasio P. (ed), Contesti relazionali e processi di sviluppo, R. Cortina Editore, Milano, pp.147-188, 1995.

Bar-Cohen Y., (Ed.), EAPAD Worldwid Web Hub, Newsletter, Spring 2002

Bar-Cohen, Y., (Ed.), Proceedings of the SPIE's Electroactive Polymer Actuators and Devices Conf., 6th Smart Structures and Materials Symposium, Volume 3669, ISBN 0-8194-3143-5, pp. 1-414, 1999.

Bar-Cohen Y., (Ed.), Proceedings of the SPIE's Electroactive Polymer Actuators and Devices Conf., 7th Smart Structures and Materials Symposium, Vol. 3987, ISBN 0-8194-3605-4 pp. 1-360, 2000.

Baughman R.H., C. Cui, A. A. Zakhidov, Z. Iqbal, J. N. Barisci, G. M. Spinks, G. G. Wallace, A. Mazzoldi, D. De Rossi, A. G. Rinzler, O. Jaschinski, S. Roth and M. Kertesz, "Carbon Nanotube Actuators," Science, Vol. 284, pp. 1340-1344, 1999.

Breazeal (Ferrell), C. and Scassellati, B., "Infant-like Social Interactions Between a Robot and a Human Caretaker". To appear in Special issue of Adaptive Behavior on Simulation Models of Social Agents, guest editor Kerstin Dautenhahn, 2000.

Brooks, RA, "Intelligence without Reason." MIT AI Lab Internal Report, 1991.

Camurri A., Coglio A., "An Architecture for emotional Agents", IEEE Multimedia, oct-dec 1998, pp. 24-33.

Cassell, J. "Embodied Conversational Agents", AI Magazine, Volume 22 No. 4, winter 2001.

Chai, J., Horvath, V., Nicolov, N., Stys, M., Kambhatla, N., Zadrozny, W., Melville, P., " Natural Language Assistant: A Dialogue System for Online Product Recommendation", AI Magazine, Volume 23, Number 2, Summer 2002.

Darwin, C., Ekman, P. (Ed.), The Expression of the Emotions in Man and Animals, Oxford University Press, New York, (1998/1872).

Dickinson, M.H. Farley, C.T., Full, R.J., Koehl, M. A. R., Kram R., and Lehman, S., "How animals move: An integrative view." Science 288, (2000), pp. 100-106.

Eibl-Ebesfeldt, I., (1973), "The expressive behavior of the deaf-and blind-born", in Von Cranach, M., Vine,I., (Eds) Social communication and movement, Academic Press, New York.

Ekman,P., "The argument and evidence about universals in facial expressions of emotion", in Wagner,H., Manstead, A., (Eds), Handbook of psychophysiology, John Wiley, London, 1989.

Frauenfelder, Mark, "Baby Robots Grow Up", ID, Industrial Design magazine, May 2002

Full, R.J and Meijer. K., "Artificial muscles and natural actuators from frogs to flies". SPIE, 1999.

Hanson D. and Pioggia G., "Entertainment Applications for Electrically Actuated Polymer Actuators", in Electrically Actuated Polymer Actuators as Artificial Muscles, SPIE PRESS, International Society of Optical Engineers, Washington, USA, Vol. PM98, Ch. 18, March 2001.

Hanson D., Pioggia G., Bar-Cohen Y., De Rossi D., "Androids: application of EAP as artificial muscles to entertainment industry", Proc. SPIE's Electroactive Polymer Actuators and Devices Conf., 7TH Smart Structures and Materials Symposium, Newport Beach, USA, 2001.

Pioggia G., Di Francesco F., Geraci C., Chiarelli P., De Rossi D., "An human-like android face equipped with EAP artificial muscles to endow expressivity", Proc. SPIE's Electroactive Polymer Actuators and Devices Conf., 7TH Smart Structures and Materials Symposium, Newport Beach, USA, 2001.

Ratner P., "Mastering 3D animation", Allworth Press, New York, 2000.

Izard, C. E., "The Face of Emotions", Appleton-Century-Crofts, New York, 1971.

Lamb, S. M., Pathways of the Brain, "The Neurocognitive Basis of Language", Amsterdam & Philadelphia: John Benjamins Publishing Co., 1999.

Helmer, R. "Computer Modelling of Tissue Depth and Density," ISDN 034340607599, 1998

Izard, C. E., "The Face of Emotions", Appleton-Century-Crofts, New York, 1971.

Marjanovic, M., Scassellati B, and Williamson M. "Self-Taught Visually-Guided Pointing for a Hu-manoid Robot". Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior (SAB-96),Cape Cod,MA,September 1996.

Olazàbal V., Sansinena, J.M., and Bar-Cohen, Y., "Versatile Actuators with Linear/Bending Movements in Air using a Series of Polypyrrole and Solid Polymer Electrolyte Trilayers" Proceedings of the SPIE's Electroactive Polymer Actuators and Devices Conf., 8th Smart Structures and Materials Symposium, 2001 (to be published).

Pelrine R., R. Kornbluh, Q. Pei, and J. Joseph, “High Speed Electrically Actuated Elastomers With Strain Greater Than 100%,” Science, Vol. 287, 2000 pp. 836-839.

Pioggia G.,Hanson, D. Di Francesco F., Geraci C., Chiarelli P., De Rossi D., “An human-like android face equipped with EAP artificial muscles to endow expressivity”, Proc. SPIE’s Electroactive Polymer Actuators and Devices Conf., 7TH Smart Structures and Materials Symposium, Newport Beach, USA, 2001.

Steels, Luc and Kaplan Frederic “AIBO’s first words. The social learning of language and meaning” SONY Computer Science Laboratory – Paris (2) VUB Artificial Intelligence Laboratory – Brussels steels@arti.vub.ac.be 2002

Tomkins S. S., "Determinants of Affects", AIG-I, Ch. 8, 248-258. New York: Springer, 1962. Williamson M. “Robot Arm Contro Exploiting Natural Dynamics,” Doctoral Dissertation, Department of Electrical Engineering and Computer Science, M.I.T. 1999

Zhang Q. M., V. Bharti, and X. Zhao, “Giant electrostriction and relaxor ferroelectric behavior in electron-irradiated poly(vinylidene fluoride-trifluorethylene) copolymer,” Science, Vol. 280, pp.2101-2104, 1998.