Proceedings of 2005 International SenseMaker Workshop on Life-Like Perception Systems

7th –8th April 2005
School of Computing and Intelligent Systems
University of Ulster, Magee Campus, Derry, NI, UK

Sensory Fusion and Emerging Behaviour in an Anthropomorphic Robot as a Man-Machine Interface

G. Pioggia, M. Ferro, S. Casalini, C. Caudai, A. Ahluwalia, D. De Rossi

Interdepartmental Research Centre "E. Piaggio"
Faculty of Engineering – University of Pisa
ITALY
E-mail: giovanni.pioggia@ing.unipi.it

Abstract – **Scientists involved in developing human-machine interfaces have recently begun to address complex problems such as biological believability and how to design machines which can perceive, learn and make choices. Once this ambitious target has been reached, dynamic human-machine interactions will have to be studied focusing in particular on possible changes in the interactive structure and thus on any manifestations of emerging behaviours. The FACE (Facial Automaton for Conveying Emotions) project addresses both issues. FACE is a life-like artifact intended as a believable human-machine interface that is able to engage in non-verbal communication by imitating and learning the emotional behaviour of an interlocutor. This paper outlines the biomimetic systems it is equipped with, the main focus being on neural control architecture which consists of a sensory-motor map that can permit sensory fusion plus a neurocontroller able to navigate within a simplified behavioural space. The neurocontroller is based on recent discoveries of the role of astrocytes in cognitive processes.**
Keywords – **human-machine interface, life-like artifact, imitation, learning, sensory fusion, neurocontroller**

I INTRODUCTION

Human beings have an extraordinary capacity for social interaction. Not only can humans communicate their moods, emotions, and reactions, but they can also interpret and predict those of an interlocutor. In the last few years, thanks to the development of artificial intelligence, we are increasingly interacting with digital autonomous agents implemented on powerful computers that are able to respond with emotions that mimic our own. There are agents simply with a text interface and agents with a graphic interface which is seemingly human. Despite the amazing graphics of some of these, inevitably the complex dynamics and responses of living organisms are lost and thus no non-verbal communication can take place. The human eye is not tricked by such agents and in fact humans tire more quickly of the graphic rather than the text interface. Many users do not find such implementations believable and are able to predict responses after a very short time.

Few agents that are physically present have so far been developed. In the physical world, compared with the digital world, it is very hard to reach a satisfactory sense of reality. The world of robotics has in any case decided to accept this challenge by drawing in from other sciences such as behavioural psychology, biology, philosophy, and anatomy, and by recognizing the great importance of emotional states on non-verbal expression.

More and more embodied agents are now being created. Some researchers are developing synthetic agents that interact emotionally with their environment. Sony have implemented in AIBO, the robot dog, an example of emotional software. The Life Group of Media Lab at the Massachusetts Institute of Technology (MIT) is using an experimental platform called Leonardo [1,2]. MIT's target is to get Leonardo to interact socially with humans. Leonardo is thus capable of facial mimicry, gesture and of manipulating objects. Leonardo can map the expression of an interlocutor and create a replica: essentially the idea is to imitate human expression. However, MIT's goal is focused on the

cognitive, affective and behavioural architecture and not on biological believability.

In order to develop a biomimetic robotic face, researchers at the Science University of Tokyo are attempting to reach greater fidelity with the real world [3]. A complex underlying architecture hosts shape memory alloy based actuators which move an elastomeric skin and allow six different expressions to be generated. Via a video camera located in the left eye, this prototype can identify such human emotions as surprise, fear, happiness and disgust. However, due to the muscular architecture that is non biomimetic, natural movements and their believability are inevitably undermined.

As these attempts clearly highlight, the fundamental elements in biomimetic robotics are the materials and how they move, and the sensory inputs and how these are controlled. As a result, the use of biologically-inspired paradigms are taking on an increasingly important role. However, to create a real mechanical android, it is not enough to evolve and assemble these technological blocks. Believability and the capacity of non-verbal communication of a life-like artefact that acts as a human-machine interface in a social environment are fundamental characteristics.

II   FACE: AIM WITHIN AN ENVIRONMENT

FACE (Facial Automaton for Conveying Emotions) [4] is a facial automaton with an anthropomorphic passive body developed at the Interdepartmental Research Centre "E. Piaggio" of the University of Pisa. In designing the physical features of an android, the environment it is placed in and its function and objective are of prime importance. FACE is placed in an environment in which the only stimuli come from the non-verbal communication of a human placed in front of it. Its aim is to be a credible human-machine interface that can establish non-verbal communication through learning and imitating the emotional behaviour of the interlocutor.

As pointed out by Jaqueline Nadel, the process of imitation is innate to humans, and place a crucial role in distinguishing between actions arising from within or actions induced by others. Moreover, imitation establishes a reciprocal nonverbal communication process in which the roles of imitator and model are continuously exchanged. We believe that a truly biomimetic approach to implement human-like facial dynamics and behaviour is through a process of imitation-based learning [5]. In order to establish imitation-based communication, FACE is human-like and possesses facial mimicry thanks to a muscular architecture. The android perceives stimuli via an artificial vision system and a biomimetic proprioception system. The signals that come from the sensory systems are fused within a sensory-motor map and processed by a neurocontroller. The data provided by the neurocontroller reach the muscular architecture via the sensory-motor map. How efficient FACE is in reaching its objective thus depends not only on the control system but also on the efficiency of the sensory and actuating apparatus that it is equipped with. Thus the biomimetic features themselves and how they are integrated take on a key role in terms of FACE's believability and powers of non-verbal communication.

III   FACE: PHYSICAL FEATURES

FACE can express and modulate the six basic emotions in a repeatable and flexible way via an artificial muscular architecture and servomotors. This process can be controlled thanks to an artificial skin consisting of a 3D latex foam equipped with a biomimetic system of proprioceptive mapping. The sensing layer responds to simultaneous deformations in different directions by means of a piezoresistive network which consists of a carbon rubber mixture screen printed onto a cotton lycra fabric. These sensors are elastic and do not modify the mechanical behaviour of the fabric. This structure allows the expression required to be achieved by means of a trial and error process. The artificial skin covers an artificial skull which is equipped with an actuating system. The head is fabricated by means of life-casting techniques and aesthetically represents a copy of the head of a subject, both in shape and texture. The artificial muscular architecture permits movement of the skin and confers human-like dynamics to it.

FACE is able to analyse the emotional reactions of individuals through optical analyses of facial expressions, to track a human face over time and to automatically store all data [6,7]. FACE's eyes are realised using animatronic techniques and their expressiveness is achieved through an artificial muscular structure surrounding the orbital region. It "sees" differently from man, using stereoscopic vision over frequency rather than over space. A three-dimensional contouring apparatus, equipped with a section for data analysis, rebuilds an internal representation of a portion of the world before it. It is based on the measurement of the average curvature that has been calculated using a 3D contouring system based on an out-of-phase sinusoidal fringe pattern projection [12,13]. This solution enables very fast image acquisition in conditions of normal day-light ambient illumination, providing complete information about shape and texture of a human face. Currently FACE surveys the curvature of the three dimensional scene once per second.

IV   FACE'S NEUROCONTROLLER

The aim of the FACE project is to develop a biomimetic machine which is believable not only concerning the materials used and the movements but also through its behaviour. By behaviour we

mean an emerging form of interaction with the environment FACE is engaged with. The problem we are currently setting ourselves is that of realizing a neural structure capable of creating its own representation of the surrounding environment in order to make it possible for innovative behaviours to emerge. These could derive from an associative memory through which it may be possible to navigate within a behavioural space. These characteristics are typical of some areas of the central nervous system like the hippocampus, upon which the architecture for the neurocontroller of FACE will be based. The current hippocampus models make use of a preformed topology of artificial neurons with varying levels of complexity, like *Integrate And Fire* or *Leabra* [8], interconnected between themselves, whose learning process depends on parameters linked to the epochs of presentation of the training set. This method creates a dichotomy between learning and acting, with different times and procedures which impede a continuous learning process. This led us to abandon the idea of realising a neurocontroller based solely on a group of neurons in various states of connection. Furthermore, preformism impedes the topological and geometrical structure from developing in an adaptive manner. Moreover, the current neural models do not include the role of glia cells and in particular those of the astrocytes. As has been recently demonstrated, the glia modulates the neural communication achieving a two-dimensional continuum in which calcium ion waves influence synaptic communication [9]. The glia are the centre of spontaneous activity induced by the continuous rhythm of the oscillations of ions at specific frequencies which influence the coordination and control of neural cells. The complex and dense branching which extends from each astrocyte defines a three-dimensional space, thereby defining an anatomical domain of influence. It is our intention to consider the group of the domains of influence as a single continuous domain, as first suggested by Beurle [10]. We propose therefore to develop a neurocontroller made up of a group of neuro-entities placed inside a continuous volume of connected astrocyte cells using an epigenetic topology. The learning process of this type of architecture will be based on imitating predefined stereotypical behaviours which can be represented in terms of FAPs (Fixed Action Patterns) and will allow navigation within a simplified behavioural space. FACE will therefore be able to continually learn, to adapt and evolve in function of the environment in which it is placed and to maintain spontaneous activity open to any innovative and intelligible behaviours arising which may then be interpreted.

a) Sensory-motor map
A sensory-motor map which acts as a buffer for the flow of information coming in from the sensors and out to the actuators has been developed. The neurocontroller will be interfaced with the outside world through the sensory-motor map. With this strategy sensorial fusion is gained enabling an abstraction with respect to the specific technology of the transducers used. Signals coming from the sensors are gathered in parallel and are encoded following a standard protocol. The encoded information is received by a specific filter for each sensor, which then sorts them to the sensory-motor map. For each actuating system a mirror image architecture has been reproduced with respect to the one described for the sensors. The information available in the sensory-motor map is encoded by a filter using the same standard protocol to then be sent to a specific interface for each actuator which will pilot its specific hardware system. The efficiency of the filtering and buffering processes over the data coming from the sensors and over the data directed to the actuating devices is delegated to appropriate driver interfaces, which are redefined according to the particular device's technology. The framework takes care to dispatch the transducer data to the control system in an efficient way, through an indexing operation during the initialization step. The presence of dynamic structures implies an opportune resource management, so the framework offers an optimized interface for enumeration and direct access requests. This architecture allows setting up a communication language between the neurocontroller and the various devices of FACE. This guarantees an increased flexibility thanks to the presence of interfaces performing the function of interpreters for the specific hardware and filters which specify the way the neurocontroller "senses" and "communicates" the information. Figure shows the flow of information to and from the neurocontroller.

b) Testing and Applications
During the first set of applications, the sensory-motor map was used to supply information to an automatic system in order to recognize facial expressions [6,7]. The system consisted of a hierarchical neural network connected directly with the area of the sensory-motor map relative to the artificial vision system of FACE. The network architecture was based on a biomimetic parallel self organising classifier followed by a predictive map which permits the classification to be appropriately quantified. The *Integrate And Fire* model was used as the base neuronal entity. The network was able, after appropriate supervised training, to recognize the facial expressions of the human interlocutor in front of FACE.
Within the framework of the FACET project (FACE as a Tool for autism), the sensory-motor map was

connected to a group of devices and communication protocols designed to permit therapeutic interactions among an autistic patient, a therapist and FACE [11].
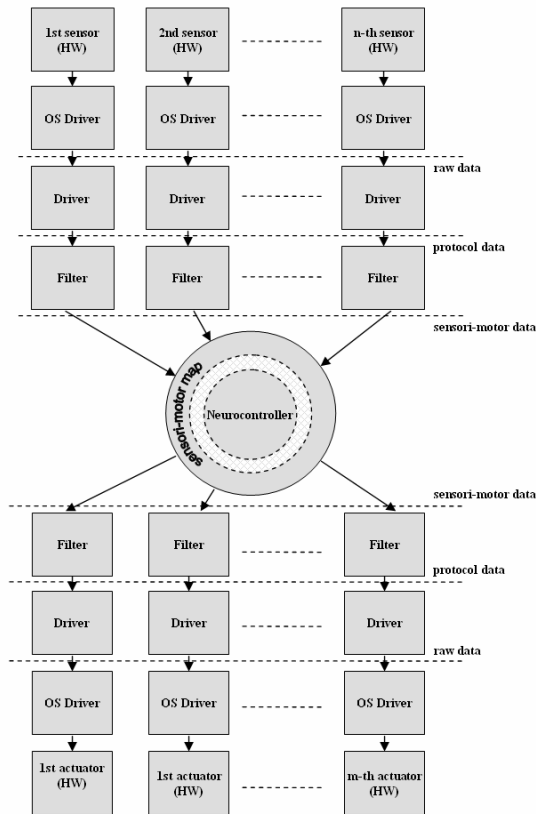

Figure – Flow of information to and from the neurocontroller. Data related to specific hardware are translated and dispatched to and from the neurocontroller

V CONCLUSIONS

We have described FACE, an ongoing project of the research group at the Interdepartmental Research Centre "E. Piaggio" of the University of Pisa to develop a believable android both in appearance and behaviour. The aim of the android is to act as a human-machine interface for non verbal communication within a simplified environment. We have described the biomimetic principles followed in designing the materials and the control system of FACE. The android possesses a neuro-control system and a sensory-motor map. The sensory-motor map is able to achieve sensory fusion through abstraction from the hardware devices used. However, the neurocontroller inevitably relies on the efficiency with which these devices interact with the environment. The architecture of the neurocontroller was inspired by the structure of the hippocampus and allows navigation within a simplified behavioural space. This architecture enables the development of models and simulations of entities and processes that are in line with the recent discoveries on the role of astrocytes in cognitive processes.

REFERENCES

[1] C. Breazeal, G. Hoffman, A. Lockerd. "Teaching and Working with Robots as a Collaboration" *Autonomous Agents and Multi-Agent Systems*, submitted.

[2] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, B. Blumberg. "Learning from and about Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots". *Artificial Life*, L. Rocha and F. Almedia e Costa (eds.), 2004.

[3] F. Hara, H. Akazawa, H. Kobayashi. "Realistic facial expressions by SMA driven face robot". *Proceedings of Robot and Human Interactive Communication, 10th IEEE International Workshop*, pp. 504–511, 2001.

[4] G. Pioggia, A. Ahluwalia, F. Carpi, A. Marchetti, M. Ferro, W. Rocchia, D. De Rossi. "FACE: Facial Automaton for Conveying Emotions". *Applied Bionics and Biomechanics*, 1(2):91-100, 2004.

[5] J. Nadel. "Early imitation and the emergence of a sense of agency". *Proceedings of the Fourth International Workshop on Epigenetic Robotics*, Genova, Italy, 2004.

[6] G. Pioggia, M. Ferro, J. Kozlowski, A. Marchetti, F. Di Francesco, A. Ahluwalia, D. De Rossi. "Automatic facial expression recognition by means of a neural approach". *Proceedings of the 2nd International Symposium on Measurement, Analysis and Modeling of Human Functions ISHF2004*, Genova, Italy, 2004.

[7] M. Ferro, G. Pioggia, F. Di Francesco, D. De Rossi. "A Hierarchical Neural Network for Automatic Facial Expression Recognition". *Pattern Recognition*, submitted.

[8] R. C. O'Reilly, Y. Munakata. "Computational Exploration in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain". *Bradford Book, MIT Press, Cambridge*, 2000.

[9] M. Zonta, G. Carmignoto. "Calcium oscillations encoding neuron-to-astrocyte communication". *Journal of Physiology*, 96, 193-198, 2002.

[10] R.L. Beurle. "Properties of a mass of cells capable of regenerating pulses". *Phil. Trans. R. Soc.* A240, 55–97, 1956.

[11] G. Pioggia, R. Igliozzi, M. Ferro, A. Ahluwalia, F. Muratori, D. De Rossi. "An Android for Enhancing Social Skills and Emotion Recognition in Autistic Patients". *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, submitted.

[12] J. Kozlowski, G. Serra, "A new modified PLL method for fringe pattern demodulation", *Opt. Eng.*, 36, 2025-30, 1997.

[13] J. Kozlowski, G. Serra, "The Complex Phase Tracing method for fringe pattern analysis", *Appl. Opt.*, 38, 2256-62, 1999.