The FACE of Autism

Daniele Mazzei Member, IEEE, Lucia Billeci, Antonino Armato, Nicole Lazzeri, Antonio Cisternino, Giovanni Pioggia, Roberta Igliozzi, Filippo Muratori, Arti Ahluwalia, Danilo De Rossi,

Abstract— People with autism are known to possess deficits in processing emotional states, both their own and of others. A humanoid robot, FACE (Facial Automation for Conveying Emotions), capable of expressing and conveying emotions and empathy has been constructed to enable autistic children and adults to better deal with emotional and expressive information. We describe the development of an adaptive therapeutic platform which integrates information deriving from wearable sensors carried by a patient or subject as well as sensors placed in the therapeutic ambient. Through custom developed control and data processing algorithms the expressions and movements of FACE are then tuned and modulated to harmonize with the feelings of the subject postulated by their physiological and behavioral correlates. Preliminary results demonstrating the potential of adaptive therapy are presented.

I. INTRODUCTION

It is well known that autistic people have difficulty in processing emotional and social information and often they behave in an out-of context manner due to this disability [1]. Various types of technology from "low" to "high" tech have been used to try to improve social dysfunction of ASD (Autistic Spectrum Disorders) subjects. Low technology approaches do not involve any type of electronic operated device and they usually treat ASD disabilities focusing on comprehension [2], communication [3] or social skills [4]. Mid technology approaches are "simple" electronic devices requiring limited advancements in technology and they mostly use Voice Output Communication Aids (VOCAs) [5]. More recently several computer technology tools used in teaching and therapy of ASD have been developed. There is reported evidence that the use of computers with autistic children helps to improve focused attention, attention span, in-seat behaviour, fine motor skills, generalization skills, etc. Possibly due to their predictability and contextuality, their controllability by the child and their lack of confusing social messages [6][7][8]. Although they profit of the benefits afforded by computer technology, they are often based on exactly the same concepts as lower-tech solutions. Most recently, innovative technological tools based on new key concepts have been developed.

R. Picard [9] [10] seminal work on affective computing has been recently focused on enabling methodologies and technologies to help people with communication impairments, like ASD subjects. [11]. Further developments of computer software are virtual reality systems as a medium to provide a simplified but exploratory interaction environment for children with ASD [12] [13].

In recent years, robots have been proposed for the treatment and rehabilitation of people with cognitive disorders. The driving force was the emergence of epigenetic robotics, a fusion between psychology and robotics. A robotic cognitive system grows as a result of interaction with the social and physical environment [14]. Recent research has shown that certain subjects with cognitive deficits perceive and treat robots not as machines, but as their artificial partners [15] [16]. Based on these observations several robotic artefacts have been used to engage proactive interactive responses in children with ASD.

C. Plaissant carried out promising studies using the Pet Robot [17], F. Michaud and his team are investigating the use of mobile robots and characters testing several robots with different shape, colour and functions [18]. In the pioneer work of K. Dautenhahn [19] [20] [21], robotic dolls, as Robota [22], are used as therapeutic or educational toys specifically for children with autism. H. Kozima builds a child-like humanoid, Infanoid, and a small creature-like robot, Keepon, by which they investigate human social development, especially related to interpersonal communication [23] [24].

Human-like robots which embody emotional states expression, empathy and non-verbal communication have also been proposed for autism therapy [25]. They can be thought of as a sort of robot based affective computing. Our group has pioneered this approach through the use of a life-like android FACE (Facial Automation for Conveying Emotions) which presents emotional information through non-verbal communication. People with autism are known to have great difficulties in inferring and analysing emotional states, both their own and of others. S. Baron-Cohen has defined mindblindness as the inability to develop an awareness of what is in the mind of another human [8]. Our hypothesis is that adaptive therapy using a robot endowed with the ability to sense, adapt and respond to a patient's postulated emotional and mental states will enable autistic subjects to learn empathy and gradually enhance their social competence. In particular, the therapy could help autistic subjects to interpret emotional states of an interlocutor, through familiarity and contextual information presented in a stepwise and controlled manner.

The FACE robot is capable, in its present embodiment, of mimicking a limited set of facial expressions, which are more easily accepted by autistic patients because of their simple and stereotypical nature. This aspect also leads to the issue of believability, an argument which is largely explored in the Japanese robotic community under the umbrella of “the uncanny valley” paradigm [26]. In this scenario, after an initial training period in which the patient familiarizes himself with the robot, we create different emotional “mood
states” in the android modulating its behaviour influencing the expression selection and intensity as function of the subjects postulated emotional states. Our paradigm is oriented towards social cognitive training helping ASD subjects to postulate the interlocutor’s emotional states through an interactive scenario based on a therapist guided questioning and imitation.

As the process of imitation plays a crucial role in distinguishing between actions arising from within and actions induced by others, it is an important part of the therapeutic process. It paves the way of the comprehension of others’ intentions by establishing a reciprocal non-verbal communication process in which the roles of imitator and reference are continuously exchanged. Moreover, imitation plays a fundamental role for the emergence of proprioception, of the perception of the external world and of the ability to act our own actions as well as those of others. True adaptive therapy entails control and modulation of the expressions and non-verbal actions of FACE by feedback from the emotional reactions from the subjects themselves, postulated by capturing and analysing their physiological and behavioural correlates through sensorised wearables. Using this closed looped approach the path towards learning and processing emotional understanding is orchestrated by a therapist, but finely tuned by the ASD subject's responses. In adaptive therapy the treatment is automatically synchronized to the subject’s emotional needs and learning pace without being overly intrusive or insistent. The range of autism spectrum disorders, as well as other developmental deficits is huge, and it is likely that individuals respond to behavioural and psychological therapy differently, much as they respond to drugs.

We are therefore using FACE in a structured therapeutic environment in which the subject’s behavior and responses are monitored using a multiplicity of sensors and then processed and fed back to the android to modulate and modify its expressions. The integrated sensing, monitoring, processing and emotionally responsive android-based therapeutic platform is termed FACET (FACE Therapy). Physiological signals (HR, HRV, RR, EMG and EDA) and eye gaze are acquired using a comfortable, unobtrusive sensorised shirt and a cap containing integrated cameras and mini gyroscopes.

To exploit the interactivity and multifunctionality of the robot, a three step adaptive therapeutic set-up was programmed and implemented. Each of the three therapeutic stages are based on a framework that includes various phases: subject-robot familiarization, subject recognition of the robot expressions, shared attention and name calling. In the first step, the subject is led through a sitting with the help of a trained therapist, who concentrates on inducing spontaneous behaviour and reaction, imitation, and emotion recognition while the robot generates a pre-programmed series of expressions. Sessions performed with patients in accordance to this preliminary stage have confirmed the hypothesis that FACE can increase emotional responsiveness in subjects with autism. A further finding was that autistic children rarely showed fear of the robot, while control subjects were uncomfortable with some of the programmed expressions manifested by FACE [27].

In the second stage the therapeutic set up involves FACE and a therapist who can intervene and decide which expression the robot should show during the patient’s interrogation and exploration. Robot-subject interaction is direct and spontaneous. During the session, the subject wears the eye-tracking cap and the sensorised shirt. Images from cameras in the room and inside the robot eyes complete the set of data to be treated and stored. Post-processing of data shows that they are correlated with the emotional participation of the patient and his/her interaction with FACE [28].

In the third stage the adaptive therapeutic set up involves FACE and the therapist operates both as supervisor and observer. However, the expressions and movements of the robot during the patient’s interrogation and exploration are selected and tuned by the robot control system taking into account and processing the data coming from the sensors and adjusting the quality and intensity of facial expressions in order to fit the subject’s emotional state postulated by the physiological, behavioral and environmental signals.

In this paper we present the results of the first 2 stages of development of the FACET platform and describe the design and evolution of the final phase which will culminate in the adaptive therapy platform.

II. MATERIALS AND METHODS

The FACET set-up in which the android guided therapy takes place includes a room equipped with motorized cameras, directional microphones, and other acquisition systems as shown in fig. 1.

A. Face Robot Hardware

FACE is an android used as emotion conveying system. It consists of a female face made of Flubber™, a skin-like silicone based rubber patented by HansonRobotics. Android faces produced by D. Hanson have been used in other robots, with their own software architectures, like the Ibn Sina Robot [29], the Javier Movellan’s robot at UCSD [30], and the INDIGO project [31]. FACE is actuated by 32 servo motors that move the artificial skin through cables inserted in the face according to the human anatomy. The motor cables act as tendons moving FACE skin and allowing human facial expressions to be re-created. FACE servo motors are all integrated in the android skull except for the 5 neck servos that allow pitch, roll and yaw movement of the head.

The android has a CCD camera in the right eye used for face tracking of the subject through an OpenCV based face tracking algorithm [32].

B. Face Control

The entire FACET system behaviour is controlled by custom made software responsible for monitoring the environment, the subject, and the robot. A number of subsystems control the different features of the system with the goal of combining reactive and deliberate behaviours. For example, the android should blink its eyes while
performed other tasks. System modules are integrated using a framework for programming robots called Robotics4.NET [33] whose purpose is to provide a robust communication infrastructure among software modules, called roblets, implementing autonomous behaviours. The underlying metaphor is the human nervous system: XML messages, corresponding to neural communications, are sent and received by roblets to and from the body map, where the state of the body is collected and used as a global perception by the program coordinating all the activities, a sort of brain.

The environment parameters are perceived through a number of different sensors (fig. 1), in part mounted on the android, in part on the surrounding environment. Nevertheless, from the software standpoint these are part of a single body even if they are connected to different computing systems. Roblets correspond to body organs, and embed autonomous reactive behaviors, possibly combining perception and actuation. This architecture guarantees signals synchronization.

![Fig 1: Connection scheme of the FACET platform. In orange are indicated the different roblets that communicate with the FACE control unit where the brain and the body map are hosted (in blue). The supervision and therapist control is allowed through the FACE configurator roblet (in green).](image)

A relevant aspect of the control software has been the development of a framework for controlling the 32 servo motors actuating FACE, responsible for defining facial expressions of the android. It is difficult to capture emotions in single expressions and combine them in an appropriate way. We organized the subsystem in layers implementing high level operations reduced into basic commands. At the very bottom motor values are normalized in the range 0-1 so that facial expressions can be reliably stored over time. Motor configurations can be saved into XML files and later reloaded. In fact these files represent a still facial expression. Face movements are achieved by means of interpolation of known positions. This is a standard approach in the context of 3D animation that has greatly influenced the design of our control system, and it enables forward and backward compatibility with well-known graphic programs. This first abstraction layer, designed to decouple software from specific hardware, is used by another layer whose purpose is to receive requests for facial expressions adaptation and combine them appropriately. The control software is inherently concurrent and different behavioural modules are expected to send requests for facial expression adaptation without having to care for possible conflicts. This is the most important abstraction provided by the software layer since it is responsible for mixing reflexes such as eye blinking or head turning to follow the patient with more deliberate actions. Each created expression request has an associated priority, which is used to blend conflicting commands, and when it should be performed. The blending and animation process is necessary to resolve conflicting expressions that may be requested frequently.

C. Sensorized Shirt

The sensorized shirt is based on e-textiles and has been developed in collaboration with Smartex Srl, Prato, Italy. It gathers, computes and transmits HR, HRV, RR, skin conductance, skin temperature and respiratory rate, all of which are known to be bodily correlates of emotional states [6]. The shirt integrates textile sensors within a garment together with on-body signal conditioning and pre-elaboration, as well as the management of the energy consumption and the wireless communication systems.

Three key points make up the sensing shirt; these are the fabric electrodes based on interconnecting conductive fibers, a piezoresistive network and a wearable wireless communication unit [34]. Electrodes and connections are interwoven within the textile by means of natural and synthetic conductive yarns. The shirt allows physiological signal acquisition with a minimal subject discomfort and total unobtrusiveness which are of paramount importance when dealing with autistic children.

D. Eye Tracking

There is a growing body of research that makes use of eye-tracking technology to study attention disorders and visual processing in ASD. Atypical gaze patterns were already described for individuals with ASD when presented with social scenes and faces [35] [36]. For instance, Klin et al. [37][38] pointed out that, in social environments, individuals with autism show reduced eye-region fixation time in favour of an increased focus on mouths and objects. Reduced attention to the face but not to the actions of a demonstrator to be imitated has been found by Vivanti et al. [35] in a group of children with autism Gaze tracking is thus a critical and useful indicator of a subject’s interest and emotional involvement during a therapeutic session with FACE. To be acceptable to children, particularly those with social difficulties, the tracking technology should be unobtrusively and ecologically used in social experimental paradigms. To this end, we developed a gaze tracking cap, HATCAM which is a wearable device that was specifically designed to investigate early attention disorders in infants.

The HATCAM device (figure 1 and detail in figure 2) was designed to be wearable, comfortable and allow eye and head tracking. Basically HATCAM consists of a child-sized cap or head band with a brim, on which a small rectangular mirror is fixed directed towards the wearer’s eyes. An opening in the brim directs the reflection from the mirror to a small video camera attached to the top of the cap. Thus the direction of the pupils with respect to the subject’s head is constantly monitored and recorded. At the same time a 3 axis inertial platform maintains information on the orientation of the head, and together the 2 sets of data
provide information on eye gaze within the framework of the therapeutic setting through a purposely developed algorithm [39].

Although more sophisticated instruments may be used to obtain accurate and repeatable data on pupillary motion, most of them require long lasting calibration procedures which may completely spoil the child’s collaboration. In accordance with this limitation we developed a fast and easy calibration procedure which is intended to be the best trade-off between accuracy and feasibility. In particular, the algorithm, use images of both subject eyes that allow calibrating the system using few calibration points than classic single eye algorithms [40] [41].

III. RESULTS
A. PHASE 1

In order to obtain a preliminary evaluation of the behaviour of a child with autism when exposed to a home-built version of FACE (shown in figure 3) with a restricted set of emotional expressions, we set up a preliminary experiment in which the reactions of two children, one a normally developing child and the other with autism, were monitored.

During the session with FACE, after a preliminary explorative phase, the autistic child (7 years) attributed the robot with emotion (sadness), and did not show any sign of fear as confirmed by the therapist furthermore, the heartbeat monitoring do not shows any significant change. This experiment suggested that autistic children can develop positive “social” interactions with an expressive system, possibly because the range of actuated emotional states is limited, reproducible and easy to process.

The control child (8 years) was much more uncomfortable with FACE. He was very interested and expressed a positive reaction at the beginning, but as soon the robot’s facial movements progressed, the child became uneasy as also indicated by the change of his heartbeat analysis.

B. PHASE 2

Following the initial evaluation in the first phase, the robot was improved aesthetically as well as functionally by employing improved elastomers and increasing the degrees of freedom in the movement of its facial skin. To evaluate the behaviour of autistic children in therapist guided sessions with FACE, the reactions of 4 subjects (3 male and 1 female), between 7 and 20 years old, were monitored and compared as reported in[27] [28]. The children were diagnosed as autistic subjects using two specific diagnostic instruments: ADI-R (Autism Diagnostic Interview Revised) and ADOS-G (Autism Diagnostic Observation Schedule Generic). Experiments were carried out in order to study the interaction with FACE during twenty minute sessions. Three specific robot-subject interactions were explored:

- spontaneous behaviour of the participants and their reactions to therapist presses in correlation with the time course of the physiological and behavioural data
- the focusing of the attention towards FACE’s eye movements
- the spontaneous ability of imitation of the expressions of the android

The evaluation of the treatment was performed by analysing the recorded sessions through the CARS (Childhood Autistic Rating Scale) evaluation using 8 relevant items from the scale. In particular, we observed that the CARS score decreased or remained the same for all items as regards 2 subjects after the therapy session. Only subject 4 (the oldest, with lowest IQ and highest ADOS rating) showed an increase of 0.5 points for listening, fear and verbal communication. More importantly, all the subjects demonstrated a decrease in the score of emotional response in the CARS scale of between 1 and 0.5 points, while imitation improved in 3 out of 4 children, so implying a marked improvement in these areas after interacting with FACE.
C. PHASE 3

The current version of FACE is more sophisticated and more believable than the previous 2 versions (figure 5). It is endowed with face tracking, such that it automatically follows the subject’s head movements, and auto blinking routine allowing attention sharing to be conducted in a more natural fashion. Furthermore, the 32 degrees of freedom of the facial skin enables the implementation a wide range of expressions and movements. Clinical trials are on-going.

A. Setting up mood states

The robot’s “mood state” is its base line behaviour which influences the automatic selection of expressions during the therapy. The Operator or the therapist could also manually select expressions through two dedicated GUI (Graphical User Interface). These manually and automatically selected expressions are modulated in intensity by the response of the subject (Physiological signal, movement, cameras, facial expressions and voice analysis).

In order to maintain a stable and strong robot-subject empathic link it is important to prevent what we call the “Joker Effect”. The Joker Effect is the emphatic misalignment typical of sociopaths who are not able to regulate their behaviour to the social context. The Joker effect as in the Batman movie consists in the generation of unsuitable expressions for the therapeutic context and could induce fear or discomfort in the subject.

With the integration of the adaptive control algorithms FACE will choose autonomously its expressions from a pre-determinate set according to its “mood state”. The Expressions in accord with the “mood state” have higher probability to be chosen by the robot control algorithm.

The subject state is internalized by the robot through signal processing algorithms based on the robot’s signals (Fig. 1) whose output is a value in the range -1 to 1 [42] [43] [44] [45]. Negative values indicate that the subject is nervous and not confident otherwise positive values indicate a positive emotional state.

FACE’s “mood state” can be coupled to one of the six basic “Ekman” emotions [46] [47] which are rated form -3 to +3 as shown in fig. 6. Expressions are automatic selected on the basis of their closeness to the “mood state”.

In order to prevent the Joker Effect selected expressions are not fully actuated but they can be modulated in intensity from 0 to 1 or avoided. The final expression intensity is modulated according to the subject state postulated by the physiological signals merging the motor position of the selected expression with that of the neutral expression. Sensors will be also used to implement reactive behaviours. For instance the robot eye-camera will be used to track the subject’s face and to turn the robot’s neck and eyes to keep the subject in view as a real person would do.
Furthermore a 3D simulator of FACE is being developed in order to study and design facial expressions. Algorithms which change the 3D mesh according to what the real android does with a specific configuration of the servo motors are being developed and implemented as a virtual FACE editor (figure 7). The editor will allow therapists to prepare expressions and facial behaviours more effectively even off-line. The FACE editor will be extended with a vision module to acquire facial expressions from a 3D camera.

![Figure 7: The FACE editor in the 3D simulator mode.](image)

V. CONCLUSIONS

In this paper we describe how face, together with a therapeutic setup comprising sensorised shirt, video cameras and eye tracking hat is being used in adaptive therapy for autistic children such that they can learn empathy. By our aim is to furnish ASD subjects with appropriate tools to enable them to enhance their social cognitive skills.

Although this work is focused on autism and mindreading, the FACET system could have a wide range of applications in the study of non-verbal communication and expression.

REFERENCES