T-LIFE: Therapeutic LearnIng of Facial Emotions

Verónica Orvalho
Instituto de Telecomunicações
Instituto Superior Técnico,
Torre Norte, Av. Rovisco Pais
Lisboa, Portugal
veronica.orvalho@fc.up.pt

Miguel T. Coimbra
Instituto de Telecomunicações
Instituto Superior Técnico,
Torre Norte, Av. Rovisco Pais
Lisboa, Portugal
mcoimbra@dcc.fc.up.pt

Miguel Sales Dias
Microsoft Portugal
Ed. qualidade, C1 e C2, Av.
Prof. Dr. Silva, Tagus Park
Porto Salvo, Portugal
Miguel.Dias@microsoft.com

Mel Slater
Universidad de Barcelona
Campus de Mundet, Ed.
Teatre, P. V. d'Hebron 171
Barcelona, Spain
melslater@gmail.com

Diego Gutierrez
Universidad de Zaragoza
CPS (Edificio Ada
Byron)Maria de Luna, 1
Zaragoza, Spain
diegog@unizar.es

Julián Flores
U. de Santiago de Compostela
Ed.Monte da Condesa s/n.
Campus Sur
S. de Compostela, Spain
julian.flores@usc.es

ABSTRACT

The face is the key element to convey emotion and plays an important role in verbal and non-verbal communication. Many efforts have been done to teach people with Autism Spectrum Disorders (ASD) to recognize facial expressions with varying results [2], but none focused on using real time facial synthesis. Most methodologies use Paul Ekman's approach [1] based on photographs of facial expressions. Besides having severely limited interactivity, they fail to reproduce the dynamics of a facial expression: far from being a still image, it is the voluntary and involuntary contraction of muscles that produce different facial movements. These movements convey emotions from one individual to another, enabling non-verbal communication. Thus, we need to weigh in an additional teaching method that allows facial motion. Then, it is necessary to study techniques that will allow real-time facial synthesis. T-LIFE is designed to assist people with ASD to recognize facial expressions in a playful way. The key technological contributions of this project are: a real-time facial markerless motion capture system, a new facial expression analyzer and classification method and an immersive interaction model.

Keywords

autism spectrum disorder, facial expression recognition, facial emotion recognition, facial analysis

1. INTRODUCTION

T-LIFE aims to improve the ability of socially and emotionally impaired individuals to recognize and respond to emotions conveyed by the face by means of Therapeutic Learning Tools.

We argue that current technological advances in character animation can substantially improve the way we teach people with Autism Spectrum Disorders (ASD) to recognize facial expressions. Our approach introduces a novel and sophisticated interaction model that enables patients to learn by imitating the avatars' movements. Our work includes a thorough field test with therapists and patients to validate our methodology, which will be performed by APPDA, FPCE/UP and Stanford University School of Medicine.

A substantial part of our research is targeted towards 3D avatars capable of teaching individuals with social phobia, anxiety and ASD, how to improve personal interactions. The main outcome will be a comprehensive system that analyzes the facial behavior of a human and enables avatars to respond to it in real-time. The avatars are 3D models with underlying anatomical behavior, and have a wide range of visual styles: human, cartoon or fantasy creature. We shall develop a prototype, which will run on PCs and game consoles, that builds upon the patient-therapist relationship: it is designed to be used both by the therapist and the patient, but also by parents and teachers. The prototype will include a set of exercises to help the patient improve his interactions with other people.

In previous work where people interact with a virtual character, the virtual character representations have been quite simple, and little or no attention was paid to facial expression [4]. Here we intend to incorporate the full range of facial expressions based on a geometric and muscle model, thus giving a much greater repertoire of possible expressions. Moreover, in the context of our solution, it is possible to feedback the patientst' own facial expression as captured on a webcam, by mapping it into a virtual character and allowing them to see a representation of themselves in a virtual mirror. As they look into this mirror and change their own facial expressions, they will see their virtual mirror image change accordingly. This can, with the help of the trainer, be used as part of a feedback loop, so that the patient can learn to control facial expressions in a playful way, without induction of stress. For example, suppose that the patient looks in the mirror and sees himself as a famous cartoon

character which nevertheless reflects their own facial movements; this is something to be enjoyed rather than a difficult process of learning.

The technology behind the prototype is based on the facial synthesis of 3D characters. The main research challenges arise from the synchronization and realism problems, the support for the reusability of facial components, and the need for an avatar-user interaction model with real time response. Traditional techniques to achieve high quality facial animation include keyframe animation and motion capture based on facial markers. These solutions are not suitable for our approach, because keyframe animation is very laborious and time-consuming and motion capture requires the user to wear markers in the face, which is unpleasant and " unnatural" . To overcome these problems, we will develop a markerless facial motion capture system that uses low cost hardware, like a webcam, to capture the information, and create a sophisticated facial rig capable of reproducing the subtleties of a face through animation. A rig is a set of controls that allows animating a character. This is not a trivial task, as we are using low resolution input data to drive the face model animations. Thus, this rig becomes the foundation of the system pipeline. It will ensure that the characters animations follow a consistent artistic style, to ease the process of recognizing facial expressions and emotions. The PI has successfully demonstrated in her SIGGRAPH07 work [3] that retargeting of facial rigs is possible and it is key to obtain high quality animations. As far as we know, there is no facial markerless motion capture system that runs in real-time and produces cinematographic quality results.

2. THE TEAM

To carry out this project and achieve its goals, we have gathered an interdisciplinary team, as well as external consultants, which include experts in the study of presence in virtual environments - Dr. Slater -, in facial character animation - Dra. Orvalho - in speech synthesis and recognition - Dr. Sales Dias - in rendering and computational photography - Dr. Gutierrez -, in human-computer interaction - Dr. Julián Flores -, in pattern recognition and computer vision - Dr. Coimbra - and in Autism Spectrum Disorder - Dr. Fienstein, Dra. Freitas and Dra. Queirós. We expect to include in the near future experts from Germany, Switzerland and the UK on geometric modeling and deformation, game design and global illumination effects for interactive applications. Our current industrial partner is Microsoft.

3. RESEARCH PROBLEMS

We shall address the following research problems: **1.** develop a markerless facial motion capture system that uses low cost hardware, like a webcam, and is still capable of reproducing in real-time cinematographic quality facial animations. As far as we are aware, this type of system doesn't exist. **2.** create a facial expression analyzer and classifier system that provides facial detail information that cannot be captured using current motion capture systems. We will explore novel computer graphics algorithms. **3.** explore different user interaction models to define them most adequate tangible user interface to allow an immersive behavior when interacting with people that suffer of ASD. Current approaches lack of an efficient HCI methodology. **4.** study and propose a novel interactive learning environment that will improve social and communication skills by training interpersonal awareness through facial recognition. **5** deploy a prototype that enhances the ability to observe and recognize emotions through an interactive experience and becomes a key reference in this field.

4. CONCLUSION

There are four important potential repercussions of the T-LIFE project: technological, scientific, clinical and social.

Technological: We aim to become a key reference in the field of facial animation in real-time, as our technology can speed up the character animation process by capturing a performer facial movements directly to multiple 3D models. It will enable creating animations on the fly by scripting facial behaviors, using our facial expression analyzer and classifier system. Then, non-experienced users will be able to create high quality animations. Having a person simply tell the system how an expression is, by describing it in a high level language is immensely empowering.

Scientific:We aim to set the foundation of a new generation of facial synthesis algorithms by exploring the emerging domain of markerless facial motion capture, and the research of novel solutions on the field of facial expression recognition.

Clinical: We will develop a prototype that can be used for validating this new approach for helping autistic people in identifying and expressing emotions. The novelty of this approach is the interactivity of the process, as opposed to previous methods using still images.

Societal: Integration of socially impaired people, which will help them become a contributing member of a advanced society.

5. ADDITIONAL AUTHORS

Additional authors: Cristina Queirós (F. Psicologia, U. Porto, email: `cqueiros@gmail.com`), Monica Oliveira (F. Psicologia, U. Porto, email: `monicaqoliveira@gmail.com`) and Paula Freitas (APPDA Norte, email: `pmfreitas@icbas.up.pt`).

6. REFERENCES

[1] P. Ekman and W. Friesen. Unmasking the face. 1975.

[2] O. G. et al. Facilitating emotional understanding and face processing in young children with asd, using animations of vehicles with faces. 2007.

[3] V. Orvalho. *Reusable facial rigging and animation.* PhD thesis, UPC, 2007.

[4] X. Pan and M.Slater. A preliminary study of shy males interacting with a virtual female. *In PRESENCE 2007.*