Improving Communication Skills of Children with ASDs through Interaction with Virtual Characters

Abirached, B., Aggarwal, J., Tamersoy, B., Zhang, Y.
Dept. of Electrical & Computer Engr.
University of Texas
Austin, Texas, USA
yanz@ischool.utexas.edu

Fernandes, T., Miranda, J., Orvalho, V.
Instituto de Telecomunicações
Faculdade de Ciências da Universidade do Porto
Porto, Portugal
veronica.orvalho@dcc.fc.up.pt

Figure 1. Overview of the game pipeline. From left to right: (1) facial rig, (2) and (3) facial expressions the user needs to recognize, (4) screen shot of the game mode “recognize the expression” and (5) a young child playing with the game.

Abstract—This article presents the LIFEisGAME project, a serious game that can help children with ASDs to recognize and express emotions through facial expressions. The game design tackles the main experiential learning cycle of emotion recognition: watch and recognize, learn by doing, recognize and mimic, generalize or knowledge transfer to real life. We briefly describe the technology behind the character animation pipeline centered on the creation of a generic rig. Then, we detail the facial expression analyzer that uses Active Appearance Models. Last, we describe the user study experiment using the game mode “recognize the expression”.

Keywords - autism spectrum disorder; emotions; serious games; facial animation; emotion recognition; HCI; consumer health informatics

I. INTRODUCTION
Autism spectrum disorders (ASDs) are a group of developmental disabilities that span a continuum of symptoms and behaviors. Individuals with ASDs often struggle with significant communication, social and behavioral challenges; however, each individual has a unique set of characteristics and severity of symptoms. The difficulties experienced by individuals with ASDs are often described as consisting of a “triad of impairments” including social and communication impairments and a tendency towards inflexibility in thinking, language, and behavior [1]. Additionally, individuals with ASDs may have difficulty understanding and making correct inferences about the cognitive and emotional states of others [2].

The impact of this disorder is considerable. The United States Center for Disease Control (CDC) estimates that one in every 110 children in the U.S. has an ASD [3]. The high prevalence rate of ASDs means that understanding the impact of autism on the greater society and developing appropriate and effective intervention tools and strategies has become increasingly important.

Although ASDs affect individuals differently, they often experience common issues, including difficulties maintaining eye contact and comprehending non-verbal cues such as facial expressions. Facial expressions are a primary means of conveying human feelings and emotions, which play a critical role in accomplishing normal social interactions. The difficulties that individuals with ASDs have in processing and understanding emotions expressed through facial expressions, especially complex and nuanced emotions, seriously hinders their ability to participate in social life. An intervention tool, such as a game, that facilitates their emotion-recognition skills could significantly improve the social interactions of children with ASDs.

Several considerations must go into designing a game to assist children with ASDs in recognizing and understanding facial expressions of emotion. The first challenge in developing an interactive system to distinguish emotions lies in recapitulating actual facial expressions: those generated by contraction of facial muscles that result in temporarily deformed facial features. To this end, any system that classifies and analyzes facial expressions must determine the location of facial features, their intensity, and their dynamics. When adding emotion recognition, contextual information, such as characters and setting, must
also be considered and incorporated into the system design [4].

Our project, LIFEisGAME attempts to apply a serious game approach to teaching children with ASDs to recognize facial emotions using real-time automatic facial expression analysis and virtual character synthesis. Most of the current means of teaching children emotions are non-interactive, and the effectiveness of these existing pedagogical means is questionable. Meanwhile, most existing training programs have not systematically focused on teaching emotion recognition, but instead were incorporated as part of group social-skills interventions [5]. In order to achieve better learning outcomes, we intend to design an interactive game to engage the children and help them learn emotions in a fun way. In this article, we will outline an overarching view of the game we are developing, including the pedagogical modes in the game, the technologies that enable the game, and some preliminary user testing results. Future research and development concerning the game will be discussed at the end of this article.

II. GAME DESIGN: PEDAGOGICAL MODES

The goal of the game is to help children with ASDs recognize emotions through facial expressions. Based on the existing literature about the characteristics of autistic children and a survey of psychologists specializing in autism therapy, therapists, and parents of autistic children, we defined the following four game modes. (see figure 1)

1. **Recognize the expression.** In this mode, the player is presented with a sequence of random facial expressions and required to identify a specific (pre-selected) expression from the set. Each session is time-limited, as determined by a therapist, to maximize player attention.

2. **Build a face.** This mode asks the player to construct a facial expression on a 3D avatar to match a defined emotion. To build the expression, the player can drag and drop specific facial features, such as the eyes or mouth, to a canvas, until the 3D avatar is completed.

3. **Become your avatar.** This mode has three levels. Level one is a free-play mode, where the avatar simply mimics the player’s facial expressions. Level two is an expression training mode, where the player attempts to achieve a target facial expression. Level three is a follow-up mode, where the player must follow an avatar’s expressions.

4. **Live a story.** In this mode, the player is presented with a story (e.g. a video or comic). The player is required perform the expression that corresponds to the situation depicted in the story.

The 3D characters in the game, including the two characters in the figure, are intentionally made to be likable and agreeable to children. Therefore, the avatars possess some attractive features such as healthy skin and symmetrical faces and bodies. Also, some avatars have child-like traits to appear warmer and more trustworthy for the player [12].

The four modes match the modes outlined in the experiential learning cycle (Kolb, 1984). In the first mode, children are encouraged to watch and recognize facial expressions. In the second mode, they are encouraged to learn by doing, that is, actively experiment with different possibilities of constructing a desired facial expression. In the third mode, children are not only encouraged to recognize and mimic, but also to concretely experience how to make the expression with their own faces. Finally, the children are encouraged to generalize or transfer their knowledge of facial expressions to real-life situations, requiring them to understand each emotion. The four modes differ in their interactivity and engagement. Children will be encouraged to play these modes in a sequential manner, but customizations could be made to allow them to begin with any mode.

III. TECHNOLOGIES ENABLING THE GAME

A. Real-time Virtual Character Synthesis

Each face is unique, causing the primary difficulty in animation. LIFEisGAME aims to create a virtual representation of the player’s face and map his or her facial movements and behaviors onto a virtual face. Achieving this goal requires facial synthesis in real-time with cinematographic quality images, which we have the additional challenge of capturing during routine play—i.e. not in a laboratory setting where players could wear markers. At the same time, the game must support different styles of characters varying in shape and appearance. The key technology behind the facial animation system will be the definition and deployment of a sophisticated rigging pipeline and a motion capture technique. Rigging is the process of taking a static, inanimate computer model and transforming it into a character that an artist can manipulate to create animations [8]. Thus, an automatic rigging process becomes crucial to allow the animation of 3D avatars that are created on the fly (e.g. avatars representing a specific person) and to guarantee results with cinematographic quality.

In LIFEisGAME, we study the creation of a new automatic animation pipeline. The research provides a significant advance over traditional games’ passive approach. In LIFEisGAME, dynamic avatars embodied from the players themselves can participate in game interactions and scenarios.

B. Real-time Facial Expression Analysis

Children with ASDs can be overly sensitive towards outside involvement. Hence, facial motion capture of the player, which is the primary input of the system, should be done
without any invasive sensors or markers. This is achieved by Active Appearance Models (AAMs) [9].

AAMs are non-linear, generative, and parametric models, which are most commonly used for modeling facial expressions. An AAM is trained for each player for robust tracking results. Training involves manually landmarking a small number of images of the player, which is achieved easily with our landmarking tool. During the tracking phase, a 2D mesh is automatically fitted to the image of the player. From consecutive 2D meshes in time, a more informative 3D mesh may be generated for each frame using structure-from-motion. These frame-by-frame 3D meshes are later used for facial expression recognition and virtual character synthesis.

Even though automatic facial expression recognition is a widely researched topic, current state-of-the-art approaches are far from being general or robust. Our work in automatic facial expression recognition is in progress and we are investigating approaches where both holistic and local, shape-and-appearance based features are used in recognition.

The primary contribution of LIFEisGAME in this area will be to develop a real-time, markerless facial motion capture system, which is robust against challenges such as extreme lighting conditions and low-quality input streams.

IV. PRELIMINARY USER STUDY AND FEEDBACK

In order to evaluate our game design, we conducted a user study. Nine participants, seven boys and two girls, ranging from four to eleven years old, with ASDs, participated in the study. Their ASD diagnoses varied. Six were identified as having high-functioning autism or Asperger’s syndrome. Two were in the middle of the spectrum. The testing sessions took place at the children’s home or cafes designated by the parents. Every child was accompanied by at least one parent, mostly mothers.

In the sessions, participants were asked to play two testing versions of the game with only the first game mode, “Recognize the expression”, implemented. The first version of the game begins by asking players to select an avatar. After making this selection, the player must choose one of six basic emotions. The available emotions are based on Ekman and Friesen’s [10] six primary and cross-cultural emotions: happiness, sadness, fear, disgust, surprise, and anger. After choosing the emotion, the player must recognize and correctly identify the emotion on the avatar, which randomly cycles through facial expressions representing the six emotions. Players also select from three levels of difficulty. The second version of the game differs in that players are given the option to play additional modes where particular facial features, such as the eyes or mouth, may be covered on the avatar.

Participants were instructed to play as long as they wanted. After ending play, the children were interviewed about what they liked and disliked about the game and how we could make the game more enjoyable. The parents were also interviewed with similar questions. In situations where the children did not wish to talk, we only interviewed the parents. The game play sessions were video recorded, and the interviews were audio recorded.

The testing results suggest that overall, the children responded favorably to the game. When selecting avatars, all of the male participants preferred to play with the young-boy avatar, while the girls expressed a preference for the young-girl avatar. The majority of male participants requested to play with the alien-like avatar.

Several parents discussed the importance of game context. They expressed an interest in a game that included storylines involving social scenarios. In terms of feedback, all of the participants enjoyed the auditory feedback; however, several children deliberately made wrong selections because they preferred the wrong-answer feedback. This suggests that the game design should allow for customization, including the ability to turn off or adjust colors and sounds, or provide a less engaging wrong-answer response.

Additionally, in the testing version of the game, players had to choose the correct facial expression of an emotion, which is selected, in the current game design, from a series of images. One issue we observed was that children selected the correct answer by matching rather than by recognizing the expression. In order to reinforce learning and avoid matching, the correct answer image should be changed to text or represented by a different face. Changing the correct answer image not only has the ability to reinforce a player’s learning, but also enhance a player’s ability to generalize the expression to other faces.

V. DISCUSSION

LIFEisGAME aims at providing children with ASDs a fun game that can effectively help them recognize and express emotions through facial expressions. The game approach is important given the fact that games can stimulate competition in the player, who attempts to beat previous scores. The cooperation and competition game techniques allow the player to become more immersed in the game, in contrast to traditional facial emotion recognition applications.

The user testing suggests that the majority of our participants favored the game. Nevertheless, the game design needs to take into consideration the individuality of each child, allowing them to customize settings such as characters, color and sounds. The need for customization is echoed by the results of our survey with psychologists,
parents, and therapists, who suggested that children would like to create their own avatars and usually have very specific, but changing, interests (e.g. football, dinosaurs).

The user study also suggested that, in addition to human avatars, children want to play with different types of avatars. This result is consistent with our survey with psychologists, parents, and therapists, who suggested that acceptable characters could be humans, animals, or aliens. Furthermore, the animals and aliens should be cartoonish, while the human characters may be either cartoonish or realistic. This is due to the fact that children with ASDs have difficulty recognizing boundaries between the real and the virtual worlds. In future character designs, we will also consider more subtle aspects of the avatars, such as the agreeableness and dominance, which are conveyed through posture and demonstrated through actions. We also hope to take advantage of the face and body of the 3D avatars by giving hints to the player and transmitting information through the expressiveness of the character.

The purpose of the game is to help children with ASDs learn to recognize and understand facial expressions of emotions; therefore, how the game reinforces learning is an important issue to consider. Mechanisms need to be developed in order to avoid behaviors such as making mistakes intentionally to hear certain sounds and simply matching expressions with the pre-selected image. Furthermore, we envision that the game could function as a therapeutic intervention, especially for younger children in the simpler game modes. This could potentially facilitate cooperation between children and their therapists, which would aid in fostering the child’s learning.

Technologically, to support the development of the more advanced modes outlined in the game mode section, we are developing a novel sketching control system. This system is inspired by the way people draw and would allow a stroke to define the shape of an object, reflecting the user’s intentions [11]. Our method will create a real-time simple control system where facial deformation is sketched, which will significantly speed up the creation of facial expressions. This technology will enable patients and therapist to quickly and easily create new facial expressions, without the need of artistic or technical skills. We will perform future user studies to verify the game design enabled by this technology and improve the usability and user-friendliness of its implementations.

ACKNOWLEDGMENT

The authors would like to thank Cristina Queiros for her valuable feedback on the psychology of emotion recognition and Pamela Davis for feedback on the design of facial expressions. This work is partially supported by Instituto de Telecomunicações and the UT Austin/Portugal Program CoLab grant (FCT) UTA-Est/MAI/0009/2009 (2009).

REFERENCES