LIFEisGAME: An approach to the utilization of serious games for therapy for children with ASD

Tiago Fernandes[1,5], Samanta Alves[2], José Miranda[3,5], Cristina Queirós[2], Verónica Orvalho[1,4]

[1] Instituto de Telecomunicações, Lisboa, Portugal,
[2] Faculdade de Psicologia da Universidade do Porto, Porto, Portugal,
[3] Instituto Politécnico da Guarda, Porto, Portugal,
[4] Faculdade de Ciências da Universidade do Porto, Porto, Portugal,
[5] Faculdade de Engenharia da Universidade do Porto, Porto, Portugal,
tfernandes@fe.up.pt, veronica.orvalho@gmail.com

Abstract. This article presents the LIFEisGAME project, a serious game that will help children with ASDs to recognize and express emotions through facial expressions. The game design tackles one of the main experiential learning cycle of emotion recognition: recognize and mimic (game mode: build a face). We describe the technology behind the game, which focus on a character animation pipeline and a sketching algorithm. We detailed the facial expression analyzer that is used to calculate the score in the game. We also present a study that analyzes what type of characters children prefer when playing a game. Last, we present a pilot study we have performed with kids with ASD.

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

1 Introduction

Understanding emotions is a key element in social interaction because it enables individuals to accurately recognize intentions of others and promotes appropriate responses [1]. But what comes naturally for some, others struggle to achieve it, like how a person with ASD (Autism Spectrum Disorders) can recognize emotions. Impairments in social interactions in ASD are frequently observed as a limited use of nonverbal behaviors, including eye gaze, facial expressions, and a lack of social and emotional reciprocity [1].

Individuals with ASD are also less likely to attend to faces [2] and are impaired in face discrimination tasks when compare with typically developed children [3][4][5]. Most recently, technology has been playing a more active part in promoting facial recognition and helping individuals with ASD to understand emotions. Some examples come from games as “Mind reading” [6] or the animation series “Transporters DVD” (www.transporters.com). However, these tools do not explore the maximum potential of interactive applications.

Some authors believe that individuals with autism "may not learn to recognize emotional expressions in real time during live social situations, because emotions are fleeting and do not repeat in an exact fashion, which may reduce the number of opportunities to systematically learn from repetition" [7]. For Ambadar, Schooler and Cohn [8] dynamic emotion displays facilitate recognition, particularly for more subtle facial expressions.

Computer training and multi-technology have been shown to be successful for teaching emotional skills to children with autism [7][9][10]. Accurate recognition and interpretation of facial expressions help individuals decide when to make socially acceptable statements and provide guidance in determining approach or withdrawal strategies in interpersonal transactions [1]. But when creating a computer game to help children with ASD to understand the world of emotions we need to build it in such a way that is motivating and takes into account the particularities of this disorder.

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 games in the pedagogical way 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 [11]. In order to achieve better learning outcomes, we designed an interactive game to engage the children and help them learn emotions in a fun way. We use faces of characters, faces of the participant and faces of people the participant is familiar with to help them engage with the learning process. In this article, we will outline an overarching view of two modes of the game we developed, 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.

2 Game Design Overview

Based on the learning cycle defined in [19], we have outlined four different pedagogical modes to be included in LIFEisGAME: The first mode is "Recognize the expression", where the player is encouraged to identify a pre-selected expression in an avatar showing emotions in a random order; "Build a Face" is the second mode, here the player needs to build a certain expression on a 3D avatar; In the third mode "Become your avatar" the player uses its own facial expression to control the avatar expressions; finally, in "Live the Story" the player is presented with a story and must perform the correct facial expression in certain moments of the plot.

In this study, we developed a game prototype based on "Build a Face". It includes two game versions and runs on a touch-screen computer. The game starts with a list of avatars that the user can choose from. This avatar becomes the character the player will use during the game. The first version we call "free-play mode" in which the player controlled the expressions of a 3D avatar by drawing on a sketching zone on

the right side of the screen. After the player draws an expression, the player can drag the avatar to a timeline on the bottom of the screen. After drawing several expression the player can press the "Play" button the expressions are interpolated and an animation is automatically generated (see Fig. 1a).

Figure 1 – Build a Face. a) Free-play and mode. b) Matching an expression by sketching.

In the second version of this prototype, and after choosing the avatar, the player must choose some photos with facial expressions from the game database or take a photo of itself with the webcam. The player can choose up to a maximum of five photos to play with. After choosing the expressions to play with, the player can start the game. The objective is to bring the avatar's expression, by sketching, as close to the expressions in the images. On the bottom of the screen, there is a timeline with all the expressions the user will play with and a time bar that indicates how much time the player has left (see Fig. 1b). The score is calculated by the similarity between the avatar sketch and the photo expressions, the number of movements required and the remaining time. For more information on the learning cycle and the pedagogical game modes implemented in LIFEisGAME, please refer to our previous work in [20].

3 Technologies Enabling the Game

To achieve the game modes depicted above, we need some supporting technologies. The main technological impact of this project will be mostly in real-time virtual character synthesis and real-time facial expression analysis.

3.1 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 [13]. 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.

3.2 Real-time Facial Expression Analysis

In order to perform the facial expression analysis for the current game prototype, we used the solution presented on [18] was used. The algorithm takes a photo as input and enhances the contrast to ease the calculations. Then, after a skin color conversion and segmentation step, the algorithm finds the largest connected region on the image. The original algorithm used a threshold value to determine if a certain pixel color was either skin or not. However, this solution was not sufficiently flexible and would not work for different skin colors and, for example, for people with blond hair. This is due to the fact that the difference between skin and hair colors is very small. Therefore, it was necessary to implement an adaptive threshold algorithm to increase the number of cases in which it is able of detecting a face. At this point, the algorithm checks if the image obtained can be a face, i.e. if the proportions and size are sufficient and valid (see Fig. 2). If the image is a face, then the algorithm may proceed. Otherwise, the threshold is lowered until a given minimum and the process is repeated.

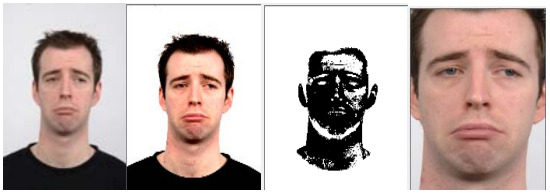

Figure 2 – Face detection algorithm. From left to right: 1) the original input photo; 2) after enhancing contrast; 3) after skin color conversion and segmentation; 4) after finding largest connected region.

If the previous steps were able of finding a possible face in the image, the next stage consists in performing an image binarization. Consequently, the process of detecting the actual face is easier and we can remove the hair, neck, etc. from the image. After this step, the facial components are extracted, namely the eyes and mouth. The algorithm will then obtain the Bézier curves for each of these facial parts along with its parameters. After these calculations, the parameters of the Bézier curves obtain from each facial part are separately matched against a database. If there is a consensus from at least two parts, then that is probably the emotion in the image.

For example, if the result for left eye is “Sad”, for right eye is “Disgust” and for the mouth is “Sad”, the final result for the person in the image is probably sad.

4 User Preference Study

Together with the evidence presented above it was also necessary, as a requirement of our game, to create an experiment that would help to better understand children’s characters preferences. In this experiment we would assess children´s preferences between a group of characters with different characteristics by showing them examples (images) of such characters and by asking them which one they preferred above all. The different types of characters would be 3D (complex cartoons), 2D (simple cartoons), photorealistic characters, (female, male), children (male, female) or animals and familiar and unfamiliar characters.

Based on this experiment our goals were:

a) To identify appealing characteristics,
b) Identify preferred type of character and
c) Get a hierarchy of preferences.

4.1 Subjects

Our sample was composed by 145 mainstream children with ages varying from six to 12 years old (M=8.34, SD= .995); in a total of 70 males and 75 females. The 145 participants were students of two State Primary Schools and were distributed between the second and the fourth year of school (M=2.85, SD=.739). All participants were Portuguese speakers. This was a convenience sample because protocols of collaboration between the school’s participants and the present project were made. All children´s parents were aware of this collaboration and consented the experience.

4.2 Procedure

Trying to answer to the previous mention questions: Do children prefer 3D (complex cartoons), 2D (simple cartoons) or photorealistic characters, do they favour adult (female, male), children (male, female) or animals and is there a difference when talking about familiar and unfamiliar characters? - Through Google search engine a set of images was selected that represented each one of the types of characters available to use in the game. All images selects are identified by their internet link to preserve author privacy.

The characters chosen varied in three types of format:

a) Photorealistic,
b) simple cartoon characters or 2D characters and
c) complex cartoons or 3D characters.

All characters could also be children (both genders), adults (both genders) or animals. An extra variable considered was familiar and unfamiliar characters know to children by the media to evaluate how this would affect their choices.

The best set of representative images was selected by three judges, expertise in facial recognition. A PowerPoint slide show with selected images was presented to the participants collectively in their class groups, or when these groups were too small in more than one class group together. In each slide show, in a total of 17, children could see four characters, each one number from one to four (see Fig. 3). All characters showed the same emotion (happiness) and all were unknown by the media, except when this variable was considered. In each slide just one variable was assessed, for example, preferences between photorealistic and complex/3D cartoons (see Table 1).

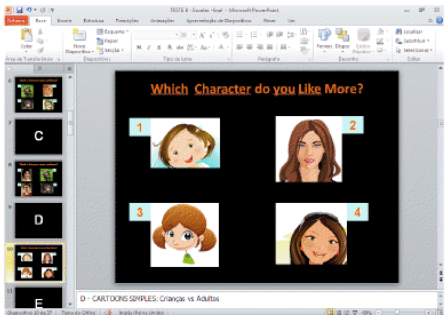

Figure 3 – Example of slideshow (image links: http://www.shutterstock.com)

Gender variable was controlled for each slide and sometimes the same variables were repeated but characters changed from female or male to assess results in both genders. Other aspects like images positioning, part of body exposed and size of image was also taken into consideration to avoid answer contamination. Each child had an answering sheet, were he/she had to mark their answer and other data like age, and school year was also added. All questionnaires were confidential. The Instructions presented on the PowerPoint slideshow were: Which character do you like more?

4.3 Results

The SPSS software, version 18.0 was use for the statistical analysis of the data. For each slide was calculated the valid percentage for each character. For the photorealistic characters, children seem to prefer animals and the same result is found when talking about 2D (simple cartoons), 3D (complex cartoons) or even familiar versus unfamiliar characters. When comparing preferences between 2D and 3D cartoons, children seem to choose 3D characters, except when talking about animals, in this case photorealistic characters are preferred instead of 3D cartoons. In terms most voted characters first we find familiar animals (70.1%), in second 3D unfamiliar

animal cartoon (66.2%), in third 2D animal cartoon (62.8%) and lastly photorealistic animal (58.6%) (see Table 1).

Table 1. User Preference Study Results

Slide	Choices	Results (%)
A	Photorealistic characters – Adults Vs. Children	Children (57.9)
B	Photorealistic characters – Animals Vs. Children	**Animals (58.6)**
C	Photorealistic characters - Adults Vs. Animals	Animals (42.1)
D	2D Cartoon – Children Vs. Adults	Adults (60.0)
E	2D Cartoon - Children Vs. Animals	**Animals (62.8)**
F	2D Cartoon - Adults vs. Animals	Animal (37.2)
G	3D Cartoon – Children Vs. Adults (males)	Children (50.3)
H	3D Cartoon – Children Vs. Adults (female)	Adults (35.2)
I	3D Cartoon – Children Vs. Animals	**Animals (66.2)**
J	3D Cartoon – Adults Vs. Animals	Adults (36.6)
K	3D Cartoon Vs. Photorealistic (children)	3D (38.6)
L	3D Cartoon Vs. Photorealistic (adults)	**3D (56.6)**
M	3D Cartoon Vs. Photorealistic (animals)	Photorealistic (46.2)
N	Familiar Vs. Unfamiliar (male/children)	Familiar (55.2)
O	Familiar Vs. Unfamiliar (female/children)	Familiar (56.6)
P	Familiar Vs. Unfamiliar (adults)	Familiar (55.2)
Q	Familiar Vs. Unfamiliar (animals)	**Familiar (70.1)**

5 Game Testing and Validation

A pilot experiment was made with two children, both male and with Asperger's syndrome. However, we have made two sessions with one child and three with other to test how they reacted when playing the game in different occasions. The children

were accompanied by their therapist. In this study, we focused on a “Build a Face” prototype as presented above.

These experiments pointed to several important issues. Although the children reacted positively to the game in the first session, one of them was more resilient to play in the second session. However, this could also be due to the fact that this child was submitted to an intensive speech therapy session minutes before our second session of testing which made him become very tired. Also, this second session wasn’t preceded by a plan made by the therapist to prepare him for the game tasks that we were presenting. Nevertheless, this result points once again to the necessity of a context in the game and to the possibility of customization.

We also noticed that the children enjoyed the touch-screen interface, but at some times, experienced some problems due to the small size of the curve handles. Since these children were used to the utilization of computers, they quickly changed to the mouse to overcome this difficulty. However, this is clearly a point to take into consideration, since it may be a problem not only for children with high-functioning autism, but mostly for children with a lower functioning level.

6 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.

In terms of character preferences, we can observe that there is a sense of hierarchy; children seem to prefer first familiar characters, then 3D, 2D cartoons and lastly photorealistic characters, being animals the most voted in all categories. It is obvious that the media plays an important part in children’s preferences. When we look at similar traits between the most voted animal characters, we can observe a Gentile/kind posture, presence of big expressive eyes and that the animals seem to be common to children like dogs, monkeys…

Human characters share some features observed in the animals characters, like the gentile posture and big expressive eyes and when looking at the shape of the faces - we can detect babyface characteristics. In addition, looking at the most voted children character, he has some female features like the long hair and a small nose.

Looking at this analysis, we can confirm some of the findings in Radbound Database [15] like the importance of masculinity/femininity of faces and also the baby face characteristics. Thinking about the game it seems important to have a variety of characters, especially animals, but thinking about the therapeutic purpose of the game in promoting the development of facial recognition and emotions understanding it is important to focus on adults and children characters.

As a limitation to our study, we have the conscience that these results are not obtained in an ASD population, but this the first of several experiments to test the characters likings. Although, being the characters that ASD children are most exposed

to are thought for the general population, this experiment is a good starting point. Plus we do not won't to over expose the ASD population that we have available to cooperate with our project. Later on, once the characters are designed character's attractiveness can be tested by ASD children. But one characteristic seems to be coherent and that is the necessity of exaggerate facial expressions, in particular the eyes. Like stated in [16], individuals with autism may need more exaggerated facial gestures to be able to interpret the emotional state of others correctly. As a next step, we are preparing game design document templates to provide to therapists so they can create their own teaching therapy.

In the other hand, the prototype testing suggests that the 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 [17]. 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.

References

[1] Bal, E., Harden, E., Lamb, D., Van Hecke, A.V., Denver, J.W., & Porges, S.W. (2009).Emotion recognition in children with Autism Spectrum Disorders: Relations to eye gaze and autonomic state. Journal of Autism and Developmental Disorders, 40, 358-370. doi: 10.1007/510803-009-0884-3

[2] Tanaka, J.W., Wolf, J.M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., … Schultz, R.T. (2010). Using computerized games to teach face recognition skills to children with autism spectrum disorder: The Let's Face It! program. Journal of Child Psychology and Psychiatry, 2258, 1 – 12.

[3] Behrmann, M., Avidan, G., Leonard, G.L., Kimchi, R., Luna, B., Humphreys, K., et al. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, 44, 110–129.

[4] Tantam, D., Monaghan, L., Nicholson, H., & Stirling, J. (1989). Autistic children's ability to interpret faces: A research note. *Journal of Child Psychology and Psychiatry*, 30, 623–630.

[5] Wallace, S., Coleman, M., & Bailey, A. (2008). Face and object processing in autism spectrum disorders. *Autism Research*, 1, 43–51. 299-316.

[6] Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). Mind Reading: the interactive guide to emotions. London: Jessica Kingsley Limited (www.jkp.com/mindreading/).

[7] Golan, O., Baron-Cohen, S., Ashwin, E., Granader, Y., MacClintock, S., Day, K., at al. (in press). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders.*

[8] Rump, K.M., Giovennelli, J. L., Mishew, R.N. & M.S. Strauss (2009). The Development of Emotion Recognition in Individuals With Autism. *Child Development*, 80(5), 1434-1447.

[9] Silver, M. & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger Syndrome to recognize and predict emotions in others. *Autism*, 5

[10] Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or High Functioning Autism to recognise complex emotions using interactive multimedia. *Development and Psychopathology*, 18, 591-617.

[11] T. Barry, L. Klinger, L. Lee, N. Palardy, T. Gilmore, and S. Bodin, "Examining the effectiveness of an outpatient clinic-based social skills group for high-functioning children with autism," *J. of Autism and Developmental Disorders*, vol. 33, no. 37, pp. 685-701, 2003.

[12] P. Eckman and W. Friesen, "Constants across cultures in the face and emotion," *J. Personality and Social Psychology*, vol. 17, no. 2, pp. 124-129, 1971.

[13] R. Falk, D. Minter, C. Vernon, G. Aretos, L. Modesto, A. Lamorlette, N. Walker, T. Cheung, J. Rentel-Lavin, and H. Max, "Art-directed technology: anatomy of a Shrek2 sequence," in Proc. of SIGGRAPH '04: SIGGRAPH 2004 Course Notes, p. 13, 2004.

[14] I. Matthews and S. Baker, "Active appearance models revisited," *Int. J. of Computer Vision*, vol. 60, num. 2, pp. 135-164, 2004.

[15] Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H., Hawk S.T. and Van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, 24(8), 1377-1388.

[16] Rutherford, M. D., & McIntosh, D. N. (2007). Rules versus prototype matching: Strategies of perception of emotional facial expressions in the autism spectrum. *Journal of Autism and Developmental Disorders, 37* 187–196.

[17] J.C. Miranda, X.A. Blanco, A.A. Sousa, D. Gutierrez, J. Orvalho, and V. Orvalho, "Painting on canvas: a facial sketching control system," poster/demo session at ACM SIGGRAPH/Eurographics Symp. on Computer Animation, New Orleans, LA, 2010.

[18] Human Emotion Detection from Image, http://www.codeproject.com/KB/GDI-plus/HumanEmotionDetection.aspx, 2010.

[19] Kolb, D. A., "Experiential Learning", Englewood Cliffs, NJ.: Prentice Hall, 1984.

[20] Abirached, B., Aggarwal, J., Costa, T., Fernandes, T., Orvalho, V., Tamersoy, B., Zhang, Y., "Improving Communication Skills of Children with Autism Spectrum Disorders (ASDs) through Interaction with Virtual Characters", 2011.