Facial virtual tracking: a system to mirror emotions

Pedro Santos[1], Vinícius Silva[2], Filomena Soares[1,2], Alberto Simões[3]

[1] Industrial Electronics Dept, School of Engineering, University of Minho, Portugal
[2] R&D Centre Algoritmi, School of Engineering, University of Minho, Portugal
[3] IPCA, Barcelos, Portugal
(a74504, a65312)@alunos.uminho.pt, fsoares@dei.uminho.pt, asimoes@ipca.pt)

Abstract. The capacity for social interaction, communication and recognition of emotions is a characteristic that the human being possesses and that allows him/her to be socially included. However, some have difficulties in expressing and interpreting emotional states, which can contribute to their marginalisation in society. A particular case is children with Autism Spectrum Disorder. These children have difficulties in social interaction and manifest repetitive patterns. In this work, a serious game is being developed in which an avatar (ZECA avatar) is able to interact with the child, through challenges and training of certain facial movements that will be validated by the system, facilitating imitation and recognition of emotions (happiness, sadness, anger, surprise and fear). The tests performed in the laboratory environment allowed to conclude the adequacy of the game for promoting emotional states in a friendly way.

Keywords: Serious Game, Facial expressions, Emotions, Human Computer Interaction.

1 Introduction

The human being has the capacity to mimic several face expressions that characterize an emotional state. This is done in a conscious way according to what the human is feeling at a given moment. In addition, he/she can interpret these emotions through the facial expressions made by others, which help to promote and understand the interaction. In fact, for humans to be able to interact, it is essential that they understand what the partner is feeling to comprehend his/her reactions. In the face, it is reflected what we feel, and from this, anyone can interpret a message, verbal or non-verbal. This message can be passed in the form of gestures and expressions. However, automatic reading of these expressions can be a very complex task. This is the case of children with Autism Spectrum Disorder (ASD) [1].

The authors of the present paper have been working on promoting emotional and social skills of children with ASD by using a humanoid robot ZECA as a mediator [2, 3]. The five emotional states considered are: happiness, sadness, anger, surprise and fear. Three activities were designed: emotions imitation and recognition and storytell-

adfa, p. 1, 2011.

ing. The results obtained so far allow to conclude that the use of robotic platforms in therapeutic scenarios with children with ASD may be a successful tool [4]. The experimental set-up of this system includes the robot (approximately height: 17 cm, weight: 1 kg), computer and cameras to register the session for further analysis of children behavior. The sessions take place in a school environment, requiring the system to be portable.

It would be interesting to reinforce emotional skills acquisition in other experimental configurations and tools, especially more lightweight. A serious game can be a suitable tool to fulfill this need. In [5] the authors present a study on using a serious game with a playware object to promote emotional skills in children with ASD. The playware object is a physical/interactive tool for the user to play the game, acting as a game controller. It has six buttons, each displaying an emoji with a different facial expression, and communicates with the game via Bluetooth. The game includes the three activities designed for the ZECA robot - imitation, recognition, and storytelling - and includes six facial expressions - happiness, sadness, anger, surprise, fear, and neutral. The preliminary tests allowed to conclude that the playware object and the serious game had a high level of approval from the children and their therapists.

Following this trend, the goal of this work is to develop a serious game that replicates the activities designed with the ZECA robot in order to promote the ability to associate facial expressions with emotions. The ZECA avatar challenges the child to understand and interpret emotional states. For this, the user's face will be monitored in order to verify the execution of his/her facial movements and to support improvements, if they are not executed correctly. This way, the user interacts with ZECA avatar by performing facial movements, identifying emotional states, and receiving feedback for improvement of possible incorrect movements/challenges. The system automatically registers the response time and the number of correct/incorrect answers.

The final goal of the work presented in this paper is to compare the results obtained by using the ZECA robot and the ZECA avatar.

This paper is organized in five sections. Section 2 presents the related works available in the literature; section 3 details the developed serious game; section 4 explains the facial expressions recognition process; section 5 shows the preliminary results obtained so far; and the final remarks and future work are addressed in section 6.

2 Related Work

The use of robots for interaction with children with ASD has been widely used and tested [6-10]. One of such robots is ZECA (Portuguese acronym for Zeca Engaging Children with Autism) from Hanson Robotics [11]. This robot allows to promote the interaction with the children providing the communication and the development of capacities of emotion understanding. Another more recent example is the QTROBOT [12] (launched in 2018) created by LuxAI to assist children in interpreting facial expressions and promoting social interaction.

Currently the design of applications to support these children is an area that has been increasing. These applications are designed to promote the development of certain capabilities in a didactic way. Livox is one of those applications [13] that contains several modes of games where the child receives stimulus to develop the capacity of communication. Using its database with over 20,000 images, audio and video features, the user can choose an item and Livox produces the audio or video that is associated with it. The association of words with images is a good way of training children who have difficulties in distinguishing emotions. Sono Flex [14] is one of the applications that is based on this concept. TippyTalk [15] allows a person with verbal disabilities to communicate by translating the pictures into text messages, which are sent to a family member or to a friend. This allows the child to communicate and express a desire, need or feeling.

For all these applications to be well accepted by the user, it is necessary that the Graphical User Interface (GUI) is captivating and intuitive so that children do not have difficulties in using them. As such, the use of games to support the development of communication and interaction skills in children with ASD is one of the potential solutions to use. In fact, some tests have already been conducted with children [16], where significant progress has been made in their behavior regarding relationships with colleagues and teachers, for example.

However, Human-Computer Interaction (HCI) is a concept that, although widely used, requires some advances in order to understand the user's reactions and to adapt the machine's behavior accordingly. It is important that this interaction is increasingly fluid and dynamic, preventing the human at any time from feeling that on the other side there is a machine programmed for certain predefined and static tasks. Through the face, the human being can transmit several expressions [17] and it is with this capacity that he/she can communicate with others allowing him/her to be accepted and introduced in society. Thus, it is important that children with ASD perceive each emotion and relate it to characteristic facial expressions.

As children with ASD have some cognitive and social problems, it is important to create a support in the sense of improving their abilities to make their relationship with others easier and more fluid [18]. For this, the creation of tools or games capable of aiding in the learning of emotions and understanding of the interlocutor's mood becomes a relevant solution. These applications become an important base of support for enabling children to develop capacities in a didactic way through constructive entertainment.

3 Serious Game Development

This section presents the purpose, the developed scenarios, the architecture and the design of the serious game that allows to mirror the user facial expressions onto the avatar's face in real-time by recognizing emotions and emotional states in a predefine context.

3.1 Game Purpose

The goal of the serious game is to have a friendly and an easily portable tool to promote the recognition and imitation of emotional states in children with ASD. The game must replicate the activities designed with the ZECA robot in the triadic interactions (child, robot, and researcher) performed in a school environment [3, 4]. In total, the system is capable of recognizing up to five emotions: happiness, sadness, surprise, anger and fear.

Taking the target group into account, the game interface should be as simple as possible, with minimum information and colour, in order to avoid distracting the children. The game activity is then dedicated to promoting the understanding and recognition of emotional states in children with ASD.

It is worth pointing out that the proposed game is to be tested in a Portuguese Special Education Unit, part of a first cycle school. For that, all the game scenarios (and voice commands) were implemented in Portuguese language.

3.2 Game Scenarios

The game is composed by three levels with increased difficulty: facial movements, facial expressions, and storytelling.

The goal of the first game level is to start introducing the child to the facial expressions related to emotional states. So, the game asks the child, by voice and through phrases written in the game scenario, to help the ZECA avatar to perform simple facial movements. Fig. 1 presents (from left to right) the commands read out by the game for the execution of the facial movements “frown”, “raise the eyebrow” and “open the mouth”.

a) “frown” b) “open the mouth” c) “raise the eyebrow”

Fig. 1. Challenges of the first level facial movements (in Portuguese).

These facial movements performed by the child are acquired and validated by the system. If the movement corresponds to the cue, it is mirrored in real-time in the avatar. As a positive feedback, the ZECA avatar performs a dance. In case the child executes no movement (during a pre-defined time), or the movement does not correspond to the cue, the game moves to other challenge within the current level until a sequence of different movements is completed. The system registers the number and the time to perform the movements.

In the second level, based on learning facial expressions through a model, the difficulty is slightly increased. In this level there is a text box, with the name of an emotion, and an animation of ZECA performing the corresponding emotional expression.

The game then challenges the child to imitate that expression. As he/she correctly responds to the task, the image is decreased, and the text box is increased until the animation (the model) disappears and there is only the textbox with the emotion written in it. Thus, the child must mimic the emotional state written in the textbox without having the model as reference. Fig. 2 shows the learning by model sequences in the game level for "surprise". The system registers the number of correct answers and the response time. This functionality allows to infer the performance of the child along the sessions.

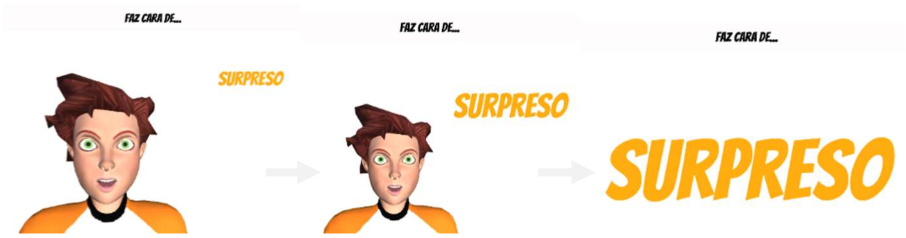


Fig. 2. Learning by model: challenges of the second level facial expressions, "surprise" (in Portuguese).

In the last level, storytelling, there is a visual and audio cue to find the emotional state. The avatar tells a story and the game shows a scenario where ZECA expresses the correspondent emotion. The child must identify ZECA's emotional state in the scene. This task is particularly difficult for children with ASD, as they rarely are able to identify their peers' emotional state. However, this level is still under development.

3.3 Game Architecture and Design

Fig. 3 presents the system workflow. The camera acquires the facial features, the computer processes the information and communicates with the serious game.

Concerning the facial tracking of the user, the open source library OpenFace was used [19]. This toolkit is capable of tracking up to 72 facial landmarks, head pose estimation, facial action unit recognition, and eye-gaze estimation only using the RGB information of a webcam.

A GUI was developed in C# to extract the user's facial Action Units (AUs) and to visualize the user's facial cues. The extracted AUs are then sent to the game and the user's facial AUs are mirrored in real-time onto the avatar's face.

The game was developed in Unity 3D software, a powerful platform to create 2D or 3D games.

In order to develop the serious game, an avatar (Fig. 3, right) was developed using the Adobe Cloud with Fuse [20] and Mixamo [21]. The physical design of the avatar was inspired by the actual humanoid robot ZECA [3]. The facial blend shapes of the avatar's face were animated using Mixamo.

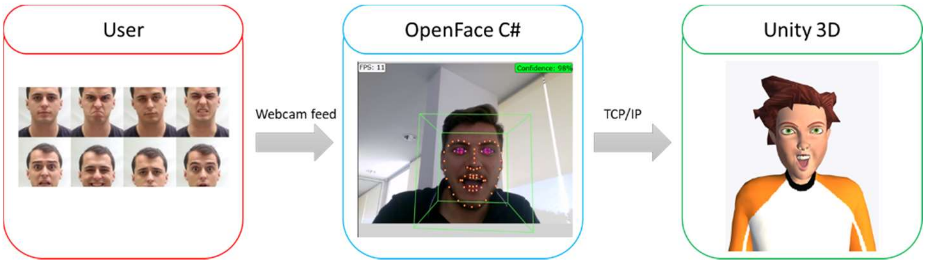


Fig. 3. System workflow.

Concerning the game design, the game was developed bearing in mind the characteristics of the target group (children with ASD) and their impairments. Thus, the game environment was kept as simple as possible to maintain the user focused on the main game character. Fig. 4 shows the system software flowchart describing the main procedure of the game. Firstly, a TCP/IP connection to the GUI is established. Then, the game starts and, while the user face is detected, the AUs are sent to the game.

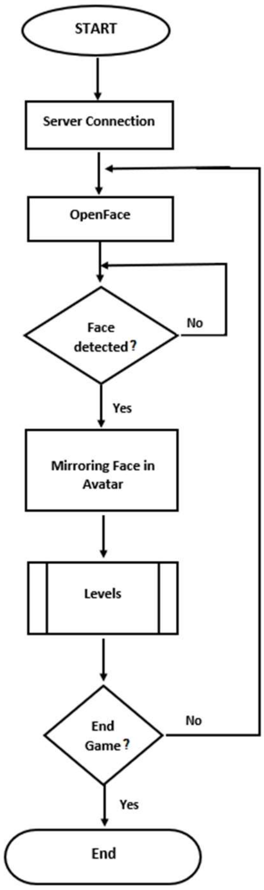


Fig. 4. Game flowchart.

4 Facial Movements and Emotion Recognition

Typically, emotions can be characterized as negative (sadness, anger, disgust or fear), positive (happiness or surprise) or neutral. The Facial Action Coding System (FACS) is the earliest method for characterizing the physical expression of emotions [17]. FACS describes several categories of head and eye positions and movements as well as the visually distinguishable facial activity. It associates the action of the muscles to the changes in facial appearance. An Action Unit (AU), which are actions performed by a muscle or a group of muscles, is the basic measurement used by the FACS. Following this idea, OpenFace is used to extract 18 facial AUs (such as blink, mouth open, among others), and the head pose angles (pitch, roll, and yaw). Fig. 5 and Table 1, show the features used as input for training the model.

A dataset was created with 32 typically developing children aged between 6 and 9 years old and 30 adults. Then, to detect the user facial expressions, a Support Vector Machine with a Radial Basis Function kernel was trained, achieving a K-fold cross-validation (with K = 5) accuracy of 89%. Table 2 shows the confusion matrix obtained for this model.

In Table 3 the precision and recall results are addressed for each class. The best performances are achieved for the recognition of 'Happy' and 'Surprise'. This model (adapted from previous work, Silva et al. 2016 [11]) can detect the six basic emotions plus neutral.

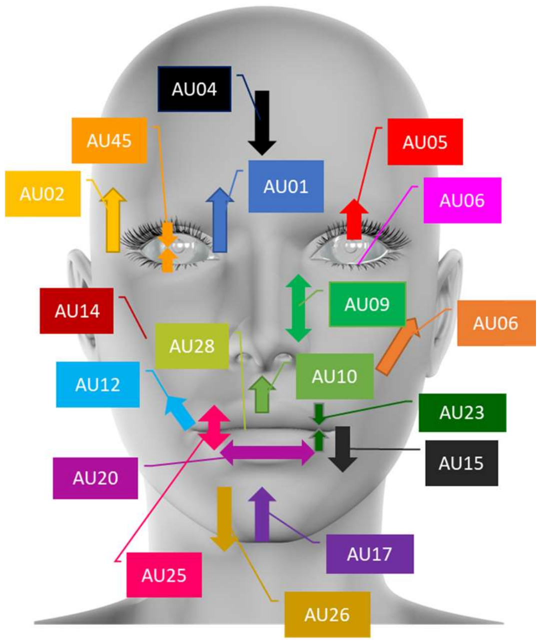


Fig. 5. Facial AUs used as input for the model (adapted from https://www.freepik.com/free-photo/3d-head_1020655.htm).

Table 1. Facial AUs and description.

AU code	AU description
AU01	Inner Brow Raiser
AU02	Outer Brow Raiser
AU04	Brow Lowerer
AU05	Upper Lid Raiser
AU06	Cheek Raiser
AU07	Lid Tightener
AU09	Nose Wrinkler
AU10	Upper Lip Raiser
AU12	Lip Corner Puller
AU14	Dimpler
AU15	Lip Corner Depressor
AU17	Chin Raiser
AU20	Lip Stretcher
AU23	Lip Tightener
AU25	Lips Part
AU26	Jaw Drop
AU28	Lip Suck
AU45	Blink

Table 2. Facial AUs and description.

	Happy	Sad	Scared	Anger	Surprise	Disgusting	Neutral
Happy	**95%**	0%	2%	1%	0%	2%	0%
Sad	1%	**88%**	2%	0%	0%	2%	4%
Scared	1%	2%	**86%**	1%	6%	3%	3%
Anger	1%	2%	2%	**88%**	0%	4%	1%
Surprise	1%	0%	3%	0%	**92%**	0%	1%
Disgusting	1%	1%	1%	2%	0%	**87%**	1%
Neutral	1%	6%	5%	7%	1%	2%	**89%**

Table 3. Facial AUs and description.

	Precision	Recall
Happy	94.9%	94.8%
Sad	88.3%	91.1%
Scared	86.3%	84.0%
Anger	88.3%	90.1%
Surprise	92.2%	94.4%
Disgusting	87.4%	93.1%
Neutral	89.2%	80.7%

5 Preliminary Results

This section presents the results obtained with the proposed system for the real-time mirroring of the user facial AUs. The first level of the game was tested in a laboratorial setting. The game was tested with 12 adults in different environment configurations, such as distance to the camera, light conditions and camera configurations in order to test and optimize the game performance.

The ZECA avatar successfully mimicked the user´s movements and the system correctly registered the user performance for each level. As an example, Fig. 6 presents the text file where the game outputs are saved: type of movement (left column), correct answer (middle column) and time to answer (right column). If the child does not perform any movement, the system registers "Timeout" in the last column.

data - Bloco de notas

Ficheiro Editar Formatar Ver Ajuda

Movimento	Resposta correta	Tempo de resposta
Baixar Cabeça	Não	TimeOut
Levantar Cabeça	Não	TimeOut
Inclinar a Cabeça Dir	Não	TimeOut
Sorrir	Não	TimeOut
Inclinar a Cabeça Esq	Não	TimeOut
Franzir sobrolho	Sim	00:00:03
Olhar para a Dir	Sim	00:00:04
Piscar olhos	Sim	00:00:05
Abrir a boca	Não	TimeOut
Lev/ Sobrancelha	Sim	00:00:10
Olhar para a Esq	Sim	00:00:05
Novo Jogo		
Baixar Cabeça	Sim	00:00:02
Levantar a Cabeça	Sim	00:00:07
Inclinar a Cabeça Dir	Sim	00:00:10
Sorrir	Sim	00:00:03
Inclinar a Cabeça Esq	Sim	00:00:03
Franzir sobrolho	Sim	00:00:03
Olhar para a Dir	Sim	00:00:03
Piscar olhos	Sim	00:00:02
Abrir a boca	Sim	00:00:02
Lev/ Sobrancelha	Sim	00:00:02
Olhar para a Esq	Sim	00:00:10

Fig. 6. Text file with the results of two trials of the first level of the game: type of movement (left column), correct answer (middle column) and time to answer (right column) (in Portuguese).

Fig. 7 shows some of the preliminary results, where the avatar mimics the model's facial movements: frown, blink, raise the eyebrow and open the mouth.

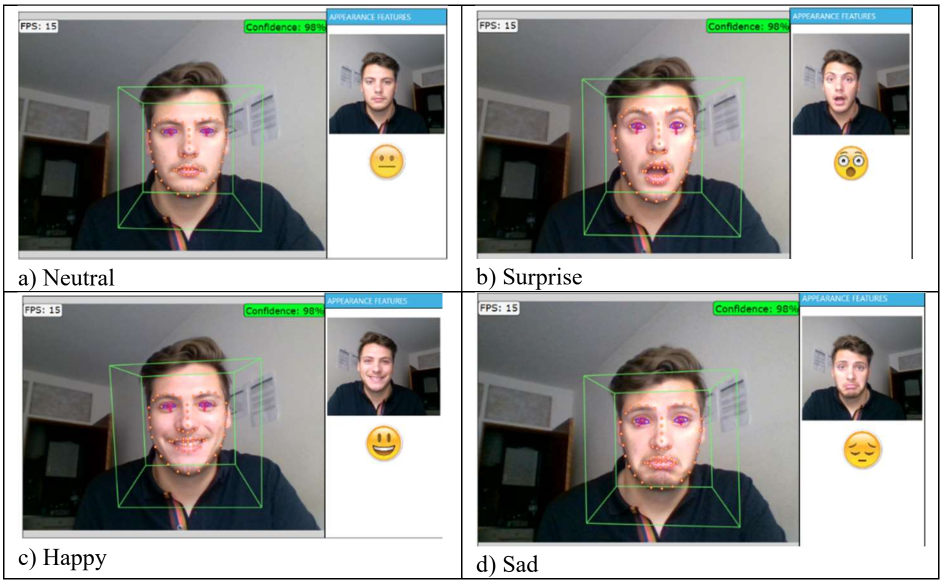
Before testing game level 2, the emotion recognition system presented in section 4 was tested in laboratorial environment. Fig. 8 presents the neutral, surprise, happiness and sadness recognition results with 98% confidence.

The next steps include the testing of level 1 and 2 in a school environment with children with Autism Spectrum Disorder.

a) frown b) blink

c) raise the eyebrow d)) open the mouth

Fig. 7. Test results of the first level of the game, facial movements, in laboratorial environment (in Portuguese).



a) Neutral b) Surprise

c) Happy d) Sad

Fig. 8. Test results of emotion recognition system in laboratorial environment using the approach presented in section 4 (98% Confidence).

6 Final Remarks and Future Work

In this work, a serious game was developed to promote the ability to perform characteristic movements (smiling, opening the eyes or opening the mouth) related to some emotions, such as being content, surprised, among others. Regarding the game implementation, an avatar (ZECA Avatar) is able to interact with the child, proposing challenges and training certain facial movements.

The system allows real-time monitoring of the user's face, verifying certain patterns related to emotions. If these are not being detected, the application can aid in the training of movements, thus improving the ability to express emotions.

The serious game is being developed in Unity software. This is a free access game engine, having libraries optimized for creating 2D and 3D games and for containing an intuitive and accessible interface. For the monitoring and follow-up of the child's face it is applied the OpenFace library (open-source library) that was developed and optimized for facial recognition with neural networks.

Future work considers optimizing and testing the system in a school environment with children with autism spectrum disorder. With this game, children are expected to promote their cognitive abilities and improve their ability to communicate and socially interact with peers.

Acknowledgments

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2019. Vinicius Silva also thanks FCT for the PhD scholarship SFRH/BD/ SFRH/BD/133314/2017.

References

1. C. Martins, Face a face com o Autismo-será a Inclusão um mito ou uma realidade?, Escola Superior de Educação João de Deus, Lisboa, 2012 (in Portuguese).
2. S. Costa, F. Soares, A. P. Pereira, C. Santos, A. Hiolle, A Pilot Study using Imitation and Storytelling Scenarios as Activities for Labelling Emotions by Children with Autism using a Humanoid Robot, The Fourth Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, IEEE ICDL-EPIROB 2014, Genova, Italy, 13-16 October, 2014.
3. V. Silva, F. Soares, J. S. Esteves, Mirroring and recognizing emotions through facial expressions for a RoboKind platform, IEEE ENBENG'2017 - 5th Portuguese Meeting in Bioengineering, Coimbra, Portugal,16-18 February, 2017.
4. S. Costa, F. Soares, C. Santos, A. P. Pereira, A. Hiolle, V. Silva, Socio-Emotional Development in Children with Autism Spectrum Disorders using a Humanoid Robot, Interaction Studies, (in press).
5. J. Azevedo, V. Silva, F. Soares, A. P.Pereira, J. S. Esteves, An Application to Promote Emotional Skills in Children with Autism Spectrum Disorders, In Göbel S. et al. (eds). Lecture Notes in Computer Science, vol 11243, pp 282-287, Springer, Cham, 2018, Serious Games, JCSG 2018 - Joint Conference on Serious Games - 9th Int. Conf. on Serious Games Development & Application, Darmstadt, Germany, 7-8 November, 2018.
6. K. Dautenhahn, A. Billard, Games children with autism can play with Robota, a humanoid robotic doll. Universal Access and Assistive Technology, pp. 179–190, 2002.
7. K. Dautenhahn, C.L. Nehaniv, M.L. Walters, B. Robins, H. Kose-Bagci, N.A. Mirza, M. Blow, KASPAR - a minimally expressive humanoid robot for human-robot interaction research. Applied Bionics and Biomechanics, 6(3–4), pp. 369–397, 2009.

8. H. Kozima, M.P. Michalowski, C. Nakagawa, Keepon. International Journal of Social Robotics, 1(1), pp. 3–18, 2008.
9. M.A. Miskam, S. Shamsuddin, M. Samat, H. Yussof, H. Ainudin, A. Omar, Humanoid robot NAO as a teaching tool of emotion recognition for children with autism using the Android app. 2014 International Symposium on Micro-NanoMechatronics and Human Science, MHS 2014, pp. 2–6, 2015.
10. H. Freitas, P. Costa, V. Silva, A. P. Pereira, F. Soares, and J. Sena Esteves, Using a humanoid robot as the promoter of the interaction with children in the context of educational games, International Journal of Mechatronics and Applied Mechanics, pp. 282-288, 2017.
11. V. Silva, F. Soares, and J. Sena Esteves, Mirroring Emotion System – On-line Synthesizing Facial Expressions on a Robot Face, 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Lisboa, Portugal, 18-20 October, 2016.
12. LuxAI, "QTrobot - LuxAI.com," Retrieved from: http://luxai.com/qtrobot/. December 6, 2018.
13. Livox, "Livox: para pessoas com deficiências e transtornos de aprendizagem Retrieved from: http://www.livox.com.br/pt/quem-somos/#sobre. December 6, 2018 (in Portuguese).
14. Tobii Dynavox logo, "Sono Flex for Communicator 5 - Tobii Dynavox." Retrieved from: https://www.tobiidynavox.com/software/content/sono-flex- for-communicator-5/, December 7, 2018.
15. TippyTalk - Comunicação Instantânea para Pessoas Não Verbais. (n.d.). Retrieved from http://www.tippy-talk.com/, January 16, 2019 (in Portuguese).
16. V. Ferreira, Utilização das tecnologias em crianças com perturbações do espectro autista em contexto da Prática de Ensino Supervisionada, Universidade Católica Portuguesa, Centro Regional de Braga, Faculdade de Ciências Sociais, Braga, 2013 (in Portuguese).
17. P. Ekman and E. L. Rosenberg, What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), Oxford University Press, 2005.
18. C. Darwin, The expression of emotions in man and animals. New York, D. Appleton anD Company, 1897.
19. T. Baltrušaitis, M. Mahmoud, and P. Robinson, Cross-dataset learning and person-specific normalisation for automatic Action Unit detection, in Facial Expression Recognition and Analysis Challenge, IEEE International Conference on Automatic Face and Gesture Recognition, 2015.
20. Fuse -> Adobe. Fuse (Beta). Retrieved April 30, 2019, from https://www.adobe.com/products/fuse/features.html.
21. Mixamo -> Adobe. Mixamo. Retrieved April 30, 2019, from https://www.mixamo.com/.