Workshops at 18th International Conference on Intelligent Environments (IE2022)
H.H. Alvarez Valera and M. Luštrek (Eds.)

doi:10.3233/AISE220064

Towards an Augmented Reality Based System for Monitoring and Handling of Emotions

Janson ROJAS [a,1], Javier NAVARRO [b], Víctor ZAMUDIO [a], David GUTIERREZ [a], Carlos LINO [a], Faiyaz DOCTOR [c] and Elizabeth NOLASCO [a]

[a] *Division of Research and Postgraduates Studies Tecnológico Nacional de México/IT de León*

[b] *Centro Estatal de Tamizaje Oportuno Hospital de Especialidad - Materno Infantil de León*

[c] *School of Computer Science and Electronic Engineering, University of Essex*

Abstract. The past decade has seen a rapid development of Augmented Reality (AR) systems and applications in many research areas including education, medicine and psychological treatments. This has been due to the growing evidence of the potential of AR systems on supporting learning tasks among other benefits. However, very little is known about the potential impact of using AR systems for influencing the emotional state of users in the treatment of cognitive impairment. In this paper, we build upon our previous work on developing a tool based on augmented reality to stimulate the emotions of patients suffering with dementia through audiovisual content (image, video and audio) [1]. Here, we focus on exploring the effects of stimulation of emotions in a controlled experiment where users of the AR system were presented with different types of multimedia. The data was collected in different sessions and analysed using a 2-factor *ANOVA* test to compare and determine the differences/changes in users' emotions. Our findings suggest that there is a significant relationship between the emotions that the user feels before and after having contact with audiovisual stimuli, thus, showing the potential for being used as efficient tools for improving the emotional state of cognitive impairment sufferers.

Keywords. Affective Computing, Augmented Reality, Emotional health, Emotions.

1. Introduction

Recent advances in Artificial Intelligence have pointed out that emotions play a key role on human daily tasks such as learning, communication and decision making by influencing human perception and rational thinking [2]. In this context, Affective Computing (AC) [3] is a relatively new area of research that aims to create "sensitive" systems by providing computers with the ability to recognise and express emotions. Thus, efforts have been made to develop human-computer interaction technologies capable of

[1]Corresponding Author: Victor Zamudio, Division of Research and Postgraduate Studies of the National Technological Institute of Mexico/IT de León, León; Mexico, E-mail: vic.zamudio@leon.tecnm.mx

addressing and identifying human behaviour and its emotional state in different circumstances. For example, the use of new tools to help people with autism understand and operate in the socio-emotional world around them [4], the focus on affective games that focuses mainly on the detection and recognition of emotions of the players and in the adaptation of the game's responses to these emotions [5], a method for measuring affective states through motor behaviour parameters from standard input devices (mouse and keyboard) [6], seeking to provide training, help in tasks related to education and provide psychological treatment, among others.

The existing body of research on identifying/transmitting emotions during human-computer interactions has established focusing on different approaches, including: a) *facial micro expressions* that can help to distinguish how a person feels where, most of the current implementations to recognise facial expressions are based on the subsequent Facial Action Coding System of Paul Ekman [10]; *Oral Communication* (voice), which is based on adjusting particular characteristics adjusted during voice synthesis to control the affect expressed through a computer generated voice, using an emotion classifier at the word level the segment can be classified automatically, it can be improved by adding information about tone and power. [11]; measuring of *physiological responses* such as heart rate, blood pressure, pulse, pupillary dilation, breathing, temperature, etc. For example, the emergence of sensors and wearable devices as mechanisms for the acquisition of physiological data of people in their daily lives has made possible the research in the recognition of emotional patterns, for the improvement of the user experiences in diverse contexts. [12]; *Written communication*, that is, either handwritten text (or typed through a computer/smartphone), can also be analysed using NLP techniques in order to extract meaning including sentiment and emotion.

In turn, these studies aim to create intelligent environments with which humans can interact while computers improve their understanding of emotions. The current evolution of Information and Communication Technologies (ICT) is leading to the development of heterogeneous technologies, including smart (computer-based) devices and objects, that are highly interconnected and cooperate through the Internet, creating technology-enriched environments where interaction occurs at various levels between humans, devices and autonomous agents [14]. An example of this is illustrated in [1] in which a non-pharmacological Alzheimer's therapy based on AR is evaluated in the context of an immersive environment. This is achieved using IBM Watson services for processing users' responses to specific questions on the emotional state. The outputs of such processing is a vector of emotion related values associated to the user/input data which is finally presented to the therapists as a means of assessing their emotional state into the whole therapy.

Another example is the work reported in [15, 16] in which, the authors report on an intelligent environment that allows users to interact and review the emotional state of people through oral communication between the user and an intelligent agent (TJBot). The case reported in [17] illustrates two users, one receiving a visual stimuli selected from a video dataset and the other noting the emotions perceived from the second user in which a quicksort based algorithm on the computer categorises videos yielding positive emotions on the user.

In order to achieve the goal of developing HCI technologies capable of addressing and identifying human emotions, AC has made use of different approaches and technological tools including Machine Learning based models [7] and more recently, *Aug-*

mented Reality (AR) based tools [8,9]. AR is a relatively new technology that has shown great promise in different interdisciplinary research areas due to its positive effects on increasing on-content understanding, memory preservation and learning motivation [9]. However, despite these benefits little has been explored in other research scenarios involving stimulating/handling human emotions. This indicates a need to explore the potential effects of AR technologies used in therapeutic scenarios such as treatment of dementia patients in which emotions can play a key role and/or can be prone to rapid mood changes difficult to deal with.

The aim of this paper is two-fold, first, we set out to describe an AR-based platform developed to provide audiovisual stimuli and allow users to describe their self-perceived emotions through sliders linked to 6 emotions[2]. Second and most important, The main objective of this study is to investigate whether such type of systems have the potential to influence on users' emotional state by analysing the data gathered from different sessions (PRE and POST) and using different stimuli. This data is in turn inspected in order to determine whether there are significant differences between the users' emotions before and after having contact with the multimedia content based on augmented reality. The paper first provides background on the *Augmented Reality* technologies and the developed *EmotionApp* in 2. Section 3 describes the experiments carried out to gather data from a group of participants with the goal of being analysed to inspect for suggestive findings of the potential of this tool in *emotion-centred* scenarios. That is followed by Section 4 where, such experimental data is analysed using a statistical approach, providing implications of the results obtained. Finally, conclusions of this work are presented in Section 5.

2. Background

This section aims to present the main concepts on which our proposal is built upon, hence, Section 2.1 in which further detail is provided on Augmented Reality technologies and the basic elements for developing an AR tool. Finally, Section 2.2 describes the developed AR based tool used to provide multimedia stimuli to users.

2.1. Augmented Reality

Augmented reality is defined as that additional information obtained from observing an environment, captured through the camera of a device that previously has specific software installed. Additional information identified as augmented reality can be translated into different formats. It can be an image, a carousel of images, an audio file, a video, or a link [19]. In order to use these technologies the following elements are required:

- A Device with camera integrated.
- A software in charge of making the necessary transformations to provide the additional information (e.g. Vuforia, AR Foundation, etc.).
- A trigger or activator of information such as image, physical environment (landscape, urban space, observed environment), marker, object or QR code.

[2] For further details on the previous work we refer the reader to [1].

2.2. Augmented Reality Application: EmotionApp

Emotion App is an Android application which allows users to view multimedia content (image, audio and videos) based on augmented reality and allows to collect information on 6 emotions (joy, sadness, anger, surprise, fear and disgust) these emotions are measured in a range of values between 0 and 1. There are 3 preloaded objectives (images) which allow the user to focus on them with the smartphone and be able to observe the content in augmented reality. The targets are set in a distraction-free environment and the multimedia content is loaded. When starting the application it is asked to provide a degree of affinity for each of the 6 emotions. Then, different audiovisual stimuli are presented to the user and at the end of the interaction a second self assessment of emotions is requested to track for possible changes.

Having presented the main tools of the AR system we move on to describe the experimental settings used for its assessment.

3. Experiments

As it was earlier discussed, little is currently known about the potential of the use of AR tools in the context of AC applications in which it is desired to identify and influence on users' emotions. In this section, we focus on describing the first set of tests of the presented application *EmotionApp*, aiming to collect information on 6 emotions perceived before and after by a group of participants undergoing a session of interaction with the AR system.

For these experiments, AR was integrated using Unity and Vuforia in order to present copyright-free audios and videos of a duration of approximately 3 to 5 minutes. Thematic images were embedded on short videos (50 seconds) to be displayed for 10 seconds each. Furthermore, users' information on their emotional state (6 basic emotions) was collected *PRE* and *POST* AR stimuli through *Sliders* representing their degree of affinity on a scale from 0 to 1 as depicted in Fig. 1.

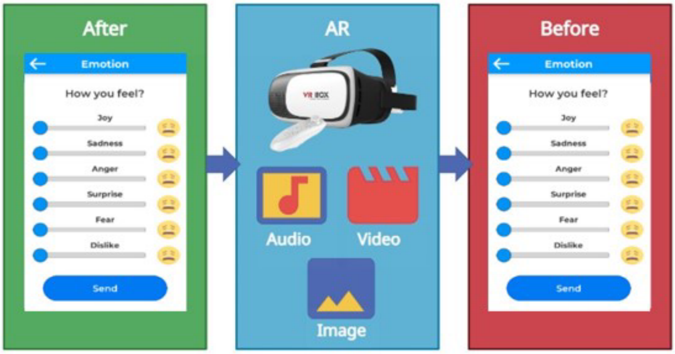

Figure 1. Application usage diagram.

The experiment was set up as follows: it consisted of 3 interaction sessions (as shown in Fig.2) performed throughout 3 different days in which we looked at the self-reported users' emotions when presented to multimedia AR sequences (audio, image and video). It was hypothesised that a rise on positive emotions could be observed after the first 2 phases, while more diverse emotions could be found after phase 3. The phases (sessions) were organised as follows:

Phase 1

- Audio: classical music with *Classic* label.
- Video: beach scenery with *Beach* label.
- Images: different landscapes with *Scenery* label.

Phase 2

- Audio: electronic music with *Electronic* label.
- Video: animated comic shorts with *Funny* label.
- Images: aesthetically pleasing animals with *Cute* label.

Phase 3

- Audio: hard rock music with *Rock* label.
- Video: a bonfire with *Fire* label.
- Images: skulls and blood covered hands with *Horror* label.

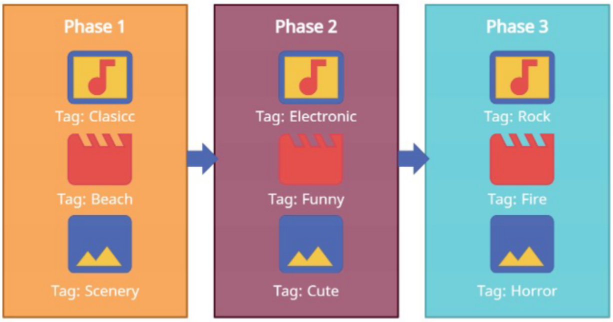

Figure 2. Phases of the experiment.

Each phase of the study consists of 15 participants, 9 were women and 6 were men aged between 20 and 65 years. Due to Covid-19 pandemic related limitations, (non-diagnosed) participants were recruited to participate including people within the set age range and without any medical diagnosis acknowledged. The average age of participants was 41.13 years (SD = 14.43). Criteria for selecting the subjects were as follows:

1. No sight problems acknowledged.
2. Not having illnesses preventing from using augmented reality technologies.
3. Compromising to use the AR tool in a quiet environment.

During 3 days, all the users interacted with the application which showed a multimedia sequence, this with the aim of observing the behaviour of the user's emotions regarding the multimedia label. Also every day before and after using the application their perception on the basic emotions was recorded. Finally, the data was prepared for analysis using *ANOVA* test. The results and discussions on this test will be presented on the next section.

4. Preliminary Results

As a result of the experiment described in last section, the information on the 3 phases was collected, that is, the self reported affinity with the 6 emotions before and after being presented with a stimuli. In order to illustrate the gathered data, the box plots shown in Figs. 3, 4, 5 depict the information of the emotions of the 15 users, where the *X* axis corresponds to the emotions (Joy, Sadness, Anger, Surprise, Fear and Disgust), whereas the *Y* axis represents the intensity perceived on a scale from 0 to 1.

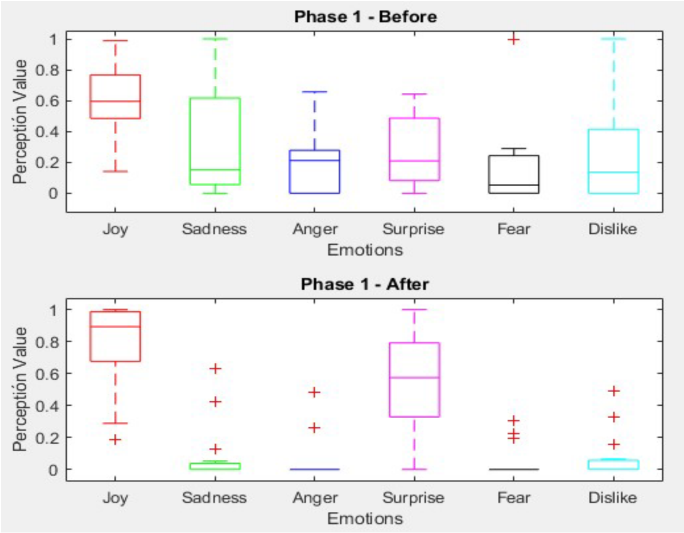

Figure 3. Boxplot Diagram Phase 1.

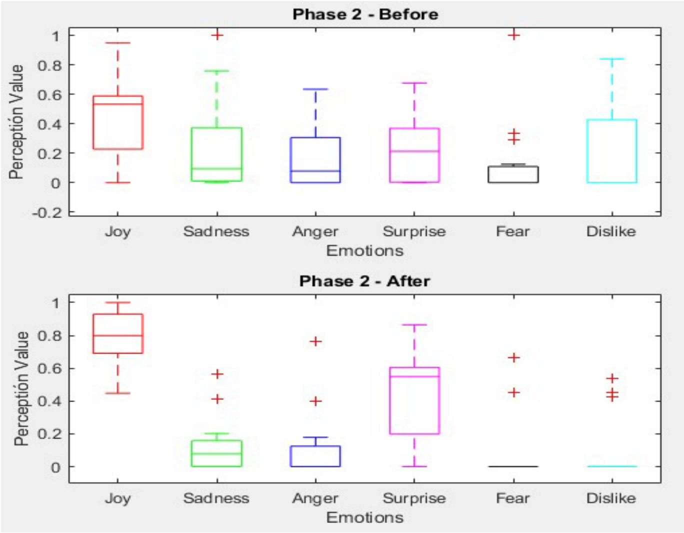

Figure 4. Boxplot Diagram Phase 2.

As can be seen in Fig. 3, the first phase, in the graph above the emotions of the user before the stimuli that are how the user normally feels, while in the graph below we can observe the emotions after seeing the multimedia stimuli and we can observe successive increases in the intensity of the emotions Joy and Surprise, while the other 4 emotions decreased. In Fig. 4, we have phase 2 in which the same effect occurs as in phase 1, the emotions of *Joy* and *Surprise* increased and the emotions of Sadness, Anger, Fear and Disgust decreased. Looking at the data from the 3rd phase depicted in Fig 5, it can be observed that when exposed to this different type of multimedia, the emotion of *Joy* decreased while the other emotions showed a pronounced decrease.

Figure 5. Boxplot Diagram Phase 3.

To provide further insight into the strength of influence on emotions during these sessions, a Two-way ANOVA test was performed to analyse the differences (changes), so that 3 null hypothesis were tested:

- H_0^A: Time factor (PRE and POST) for interacting with AR have no significant effect on users' emotional state. 30
- H_0^B: There are not significant differences regarding the emotion factor (Joy, Sadness, Anger, Surprise, Fear, Disgust).
- H_0^{AB}: There are not significant differences with regards to the interaction of the two factors.

As can be seen from the F-scores and their associated critical values shown in Table 1, only the hypotheses of factors B and interaction AB can be rejected, since their critical value for F is lower than their F-scores $F^B(5,168) = 23.62, p < .05$ and $F^{AB}(5,168) = 5.9$ $p < .05$, which suggests that there are significant differences in emotions and in the interaction between emotions and the time factor in the participants.

	SS	df	MS	F	Critical value for F
Factor A	0.0899	1	0.0899	1.4196	3.8974
Factor B	7.4827	5	1.4965	23.6188	2.2679
Interaction AB	1.8698	5	0.3739	5.9021	2.2679
Error	10.6449	168	0.0633		
Total	20.0875	179			

Table 1. Two-way ANOVA test results for Phase 1 where *SS* is the Sum of Squares, *df* is the degrees of freedom, *MS* is the Mean of Sum of Squares, and *F* is the F-score.

Similar results can be appreciated in Table 2 as in Table 1 since only the hypotheses of the factors B and the interaction AB can be rejected. That is, their critical value for F is lower than their F scores ($F^B(5,168) = 21.3$ $p < .05$ and $F^{AB}(5,168) = 4.46$ $p < .05$), basically, what the information allows us to infer is that H_0^B and H_0^{AB} are rejected and there are significant differences between the emotions and the time factor between them.

	SS	df	MS	F	Critical value for F
Factor A	0.0596	1	0.0596	1.0252	3.8974
Factor B	6.2002	5	1.2400	21.3024	2.2679
Interaction AB	1.3006	5	0.2601	4.4687	2.2679
Error	9.7796	168	0.0582		
Total	17.3402	179			

Table 2. Phase 2 ANOVA test.

Finally, in Table 3 we can see that for the 3 factors the hypotheses are rejected as all F scores are greater than their Critical values for rejection ($F^A(5,168) = 30.33$ $p < .05$, $F^B(5,168) = 11.97 p < .05$ and $F^{AB}(5,168) = 22.42, p < .05$), As we can see, the 3 hypotheses are rejected, which tells us that there are significant differences between the 2 factors and the interaction between them.

	SS	df	MS	F	Critical value for F
Factor A	1.1321	1	1.1321	30.3315	3.8974
Factor B	2.2345	5	0.4469	11.9726	2.2679
Interaction AB	2.1323	5	0.4264	11.4252	2.2679
Error	6.2709	168	0.0373		
Total	11.7699	179			

Table 3. Phase 3 ANOVA test.

As can be seen in the previous tables, there are differences between the data related to emotion PRE and POST and, since in all cases the significance value is less than 0.05 (5%), it gives us a 95% reliability that there are significant differences in the tests with the 15 users.

5. Conclusions and Future Work

In this investigation, the aim was to assess the potential effect of an AR based tool on the emotional state of users. To evaluate this, we designed a set of sessions with the aim of collecting data on self reported assessments of 6 basic emotions during 3 sessions. The sessions were designed to present different types of multimedia material to the users while asking for the users' input before and after session. As confirmed by the analysis, there was an evident change in the emotions, before and after having contact with the stimuli, since in the 3 cases the users self-perceived that their emotions improved (or worsened) with respect to the type of multimedia that they were presented at the time. Likewise, the graphs show a clear change in these emotions, depending on the multimedia sequences that the user was watching, they raised or lowered them. In turn, the 2-factor *ANOVA* test revealed that in the first 2 sessions significant differences were found only in emotions and in the interaction of emotions and the before and after each phase. Moreover, the third session presenting more diverse material showed significant differences on each factor studied, namely, time factor (pre and post interaction), emotion factor (differences on each individual emotion) and lastly, the interaction of these two factors together.

These experiment results suggest that, either positively or negatively, their emotions were stimulated by the content displayed by the AR tool. In summary, these results have provided insights to explore in a subsequent experimental phase reviewing the effects of automated tools to generate personalised multimedia sequences on the emotional state of people diagnosed with cognitive impairment.

For future work, we will focus on expanding and embedding this into an intelligent environment, including smart algorithms for selecting sequences of multimedia content by evaluating soft computing techniques which allow the users to improve their emotional state and thus, helping to have a better Quality of Life. Finally, we will aim to investigate the effect of *Emotional Congruity* [20] in the development of the intelligent algorithm.

Acknowledgement

We would like to thank the people who supported the tests to carry out the work sessions. Furthermore, we would like to acknowledge the support of CONACYT and TecNM / ITL for this research.

References

[1] Ronquillo L, Zamudio V, Gutiérrez-Hernández D, Lino C, Navarro J and Doctor F, "Towards an automatic recommendation system to well-being for elderly based on augmented reality", 2020 16th International Conference on Intelligent Environments (IE), 2020, pp. 126-131, doi: 10.1109/IE49459.2020.9155010.

[2] Cambria E, Affective Computing and Sentiment Analysis, in IEEE Intelligent Systems, vol. 31, no. 2, pp. 102-107, Mar.-Apr. 2016, doi: 10.1109/MIS.2016.31.

[3] Picard R W, Affective Computing. M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321, page 1–16, Nov 1995.

[4] Baron-Cohen S, Affective Computing and Autism, Annals of the New York Academy of Sciences, vol. 1093, no. 1, pp. 228–248, Feb. 2017.
[5] Hudlicka E, AFFECTIVE COMPUTING FOR GAME DESIGN, 2008. Available: http://www.imgd.wpi.edu/speakers/0910/Hudlicka_GAMEON_TUT_01.pdf.
[6] Zimmermann P, Guttormsen S, Danuser B & Gomez P (2003). Affective Computing—A Rationale for Measuring Mood With Mouse and Keyboard, International Journal of Occupational Safety and Ergonomics, 9:4, 539-551, DOI: 10.1080/10803548.2003.11076589
[7] Watson Studio - Visión general, Ibm.com, 2021. https://www.ibm.com/mx-es/cloud/watson-studio.
[8] Zhang J, Kamioka E, & Tan P X. (2019). Emotions detection of user experience (Ux) for mobile augmented reality (mar) applications. International Journal of Advanced Trends in Computer Science and Engineering, 8(1.4 S1), 63-67.
[9] Cipresso P, Giglioli I A C, Raya M A, & Riva G. (2018). The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Frontiers in psychology, 9, 2086.
[10] Ekman P and Friesen W, Facial Action Coding System. Consulting Psychologists Press, 1977.
[11] Ries K and Waibel A, Activity Detection for Information Access to Oral Communication, DTIC, 2022. https://apps.dtic.mil/sti/citations/ADA458656
[12] Santamaria-Granados L, Munoz-Organero M, Ramirez-Gonzalez G, Abdulhay E, and Arunkumar N, Using Deep Convolutional Neural Network for Emotion Detection on a Physiological Signals Dataset (AMIGOS), IEEE Access, vol. 7, pp. 57–67, 2019, doi: 10.1109/access.2018.2883213.
[13] Phan D-A, Matsumoto Y, and Shindo H, Autoencoder for Semisupervised Multiple Emotion Detection of Conversation Transcripts, IEEE Transactions on Affective Computing, pp. 1–1, 2018, doi: 10.1109/taffc.2018.2885304.
[14] Burzagli L, Emiliani P L, Antona M, Intelligent environments for all: a path towards technology-enhanced human well-being. Univ Access Inf Soc (2021). https://doi.org/10.1007/s10209-021-00797-0
[15] Cruz-Parada P, Zamudio Rodriguez V M, Lino Ramirez C, Identification and Analysis of Emotions in a Game Based Therapy for Patients with Cognitive Impairment, Iospress.nl, pp. 506–515, 2018, doi: 10.3233/978-1-61499-874-7-506.
[16] Manríquez Santos J J, Zamudio Rodriguez V M, Lino Ramirez C, Navarro Barrón J, Baltazar R, Gutiérrez Hernández D A and Doctor F. (2018). Towards an Innovative Architecture to Monitor and Handle Emotions in Work Scenarios. Iospress.nl, pp. 516–525, 2018, doi: 10.3233/978-1-61499-874-7-516.
[17] Baveye Y, Dellandréa E, Chamaret C and Chen L, LIRIS-ACCEDE: A Video Database for Affective Content Analysis, in IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 43-55, 1 Jan.-March 2015, doi: 10.1109/TAFFC.2015.2396531.
[18] We W, Yale Center for Emotional Intelligence, Yale Center for Emotional Intelligence, 2014. https://ycei.org/what-we-do.
[19] Blázquez Sevilla A, Realidad Aumentada en Educación, Madrid: Universidad Politécnica de Madrid, Gabinete de Tele-Educación, 2017
[20] Schmid P C and Schmid Mast M, Mood effects on emotion recognition, Motivation and Emotion, vol. 34, no. 3, pp. 288–292, Jun. 2010, doi: 10.1007/s11031-010-9170-0.