The PlayMancer Database: A Multimodal Affect Database in Support of Research and Development Activities in Serious Game Environment

Theodoros Kostoulas[†*], Otilia Kocsis[†], Todor Ganchev[†], Fernando Fernández-Aranda[‡], Juan J. Santamaría[‡], Susana Jiménez-Murcia[‡], Maher Ben Moussa[#], Nadia Magnenat-Thalmann[#], Nikos Fakotakis[†]

[†]Wire Communications Laboratory, University of Patras, 26500 Rion-Patras, Greece
[‡]Department of Psychiatry, University Hospital of Bellvitge and Ciber Fisiopatologia Obesidad y Nutricion (CIBEROBN), Instituto de Salud Carlos III, Spain
[#]MIRALab - CUI, University of Geneva, CH-1227 Carouge, Switzerland
[*]tkost@wcl.ee.upatras.gr

Abstract

The present paper reports on a recent effort that resulted in the establishment of a unique multimodal affect database, referred to as the PlayMancer database. This database was created in support of the research and development activities, taking place within the PlayMancer project, which aim at the development of a serious game environment in support of treatment of patients with behavioural and addictive disorders, such as eating disorders and gambling addictions. Specifically, for the purpose of data collection, we designed and implemented a pilot trial with healthy test subjects. Speech, video and bio-signals (pulse-rate, SpO_2) were captured synchronously, during the interaction of healthy people with a number of video games. The collected data were annotated by the test subjects (self-annotation), targeting proper interpretation of the underlying affective states. The broad-shouldered design of the PlayMancer database allows its use for the needs of research on multimodal affect-emotion recognition and multimodal human-computer interaction in serious games environment.

1. Introduction

Initially videogames have been conceived for entertainment purposes. However, during the last few years a new wave of video games, designed for a primary purpose other than pure entertainment, had appeared the serious games. The targets of these serious games have included education (Barab, 2005; Gee, 2003; Squire, 2003) or military training (Bergeron 2008). In the same way, several naturalistic studies have shown that serious games are useful for children, adolescents and young adults to improve self-esteem and knowledge about different illnesses, increase adherence to treatment, improve problem solving skills (Beale et al., 2007; Coyle et al., 2005).

The use of new technologies has been applied for a range of mental illnesses (Zarate et al., 1997; Myers et al., 2004; Botella et al., 2004; Difede et al., 2007), including depression (Griffiths 2004), alcohol abuse (Saitz 2004) or posttraumatic stress (Lange 2001). Recently, publications reflecting the possible benefits of some videogames appeared (Schott & Hodgetts, 2006; Griffiths, 2004), for example Russoniello (2009) found that playing a video game can increase mood and decrease stress.

The enhancement of the human computer interaction with affects and emotions is considered the tomorrow technology, and as such various research projects have been promoted, dealing with the multiple facets of the affective interaction. HUMAINE aims to lay the foundations for European development of systems that can register, model and/or influence human emotional and emotion-related states and processes in "emotion-oriented" systems. CALLAS aims to design and develop a framework based on a plug-in multimodal architecture to interpret and process emotional aspects in real-time for easy and fast development of applications for art and entertainment. In ElderGames advanced visualisation and interaction interfaces are implemented for IST-based games, which are used as preventive and therapeutic tools by elderly people to improve cognitive, functional and social skills.

Furthermore, significant research efforts towards creating multimodal affect databases have been reported. Honig et al. (2007) describe the DRIVAWORK database, which contains physiological plus audio and video recordings of participants in a simulated car-drive environment in different stress levels. The Augsburg database of biosignals, containing physiological recordings of four induced emotions by listening to music is described in (Wanger et al., 2005). More recently, Katsis et al. (2008) describe preliminary results evaluating physiological data obtained from ten subjects in simulated racing conditions.

The present paper reports on the efforts towards establishing a unique multimodal database, referred to as the PlayMancer database, which was created in support of research and development activities aiming at the establishment of a platform for rapid development of serious games. The reminder of this work is organized as follows. In Section 2 we present the objectives of the PlayMancer project and the motivation behind the database design. Section 3 details the methodology followed for implementing the database. Section 4 reports statistical information on the data collected and Section 5 describes the annotation procedures. Finally, Section 6 concludes this work.

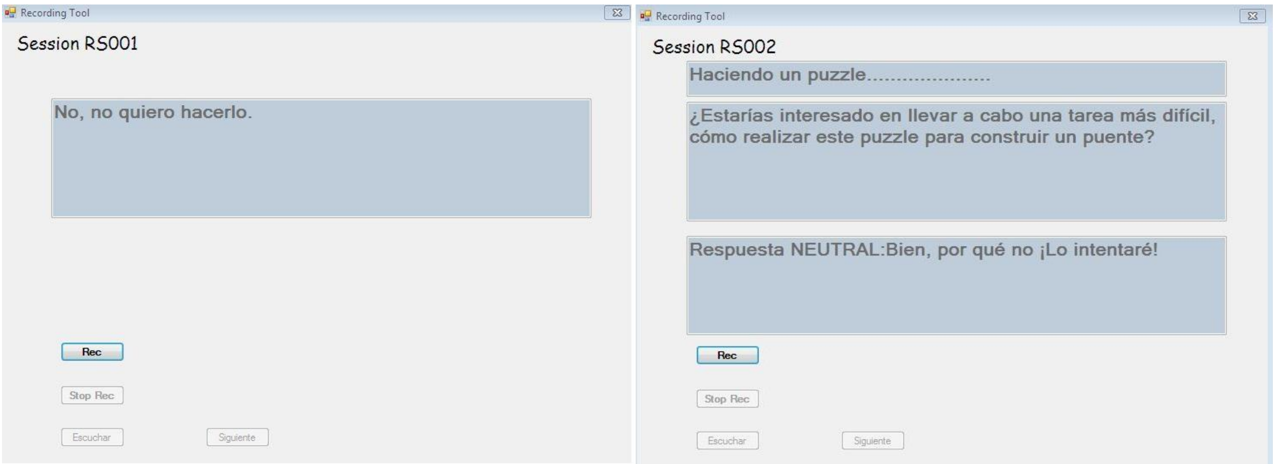

Figure 1: Graphical User Interface for recording sessions RS001 and RS002

2. Objectives

The main objective of the PlayMancer project (Jiménez-Murcia et al., 2009) is to research and develop a purposely-designed platform, which will facilitate the development of serious games. The PlayMancer platform architecture integrates various existing or newly developed components, such as a game engine, a spoken interaction interface (speech recognition, understanding and synthesis), affect-emotion recognition components (speech-based, facial, biosignal-based), a motion tracking system and a specialized graphical user interface (GUI) for therapists (Kocsis et al, 2009). Two use cases, related to development of serious games in support of the physical rehabilitation of patients, and in support of treatment of patients with behavioural and addictive disorders are meant to validate the usefulness of the developed technological framework.

The PlayMancer video-game to be adopted for chronic mental disorders (mainly eating disorders and behavioural addictions) treatment, introduces the player to an interactive scenario where the final goal is to increase his general problem solving strategies, self-control skills and control over general impulsive behaviours. After using the game, specific targeted attitudinal, emotional and behavioural changes are expected by the subject. Each task will permit access to one or several types of resources which will facilitate and improve the game character's, and hence the player's, relaxation techniques and planning skills. The game will encourage the player to learn and develop new confrontation strategies.

3. Methodology

Successful modelling of the affective states defined by the PlayMancer user requirements imposes sufficient data to be collected. For this purpose a pilot trial was conducted. It included three different data collection sessions:

- The first recording session (RS001) targeted collection of spoken audio database for the domain adaptation of the acoustic and language models for Spanish speech recognizer.
- The second recording session (RS002) targeted collection of spoken audio data for (a) acted emotion for Spanish language, and (b) acoustic model for Spanish speech recognizer.
- The third recording session (RS003) targeted collection of spoken audio data, video and physiological data (Pulse Rate and SpO2) for naturally expressed emotion for Spanish language.

During the first and the second recording sessions the participants uttered prompts presented to them through the graphical user interface (GUI) seen in Figure 1. The GUI consists of textboxes displaying the content to be uttered and four buttons – two for start and stop of the recording process, one for listening/validating the recording completed, and finally one for proceeding to the next recording/saving the finished capture. All recording sessions took place in a room of 12 square meters with low reverberation properties.

During the first recording session, no interaction scenarios were envisaged. The content of the prompted utterances was related to the vocabulary expected within the user's scenarios, according to current specifications of the PlayMancer game.

In the second recording session, the participant answered spontaneously to scenario-specific questions provided, by acting upon requested emotions (anger, boredom, joy, neutral, sadness, and surprise). For each scenario, suggested answers were presented; though each participant was encouraged to use her/his own words/expressions to answer (Figure 1).

In the third recording session, the participants were asked to interact with three video games (Global Conflict: Latin America, Dragon´s Lair 3D: Return to the Lair, and Trackmania Nations Forever). Each participant was playing for 20 minutes, and the different types of videogames were assigned in a balanced form (Latin Square).

Before each recording session the participants were encouraged to express freely, using speech, whenever they felt so. During each recording session the participant was left alone in the room. After the third recording session, the subjects were asked to proceed with the annotation session in a different room.

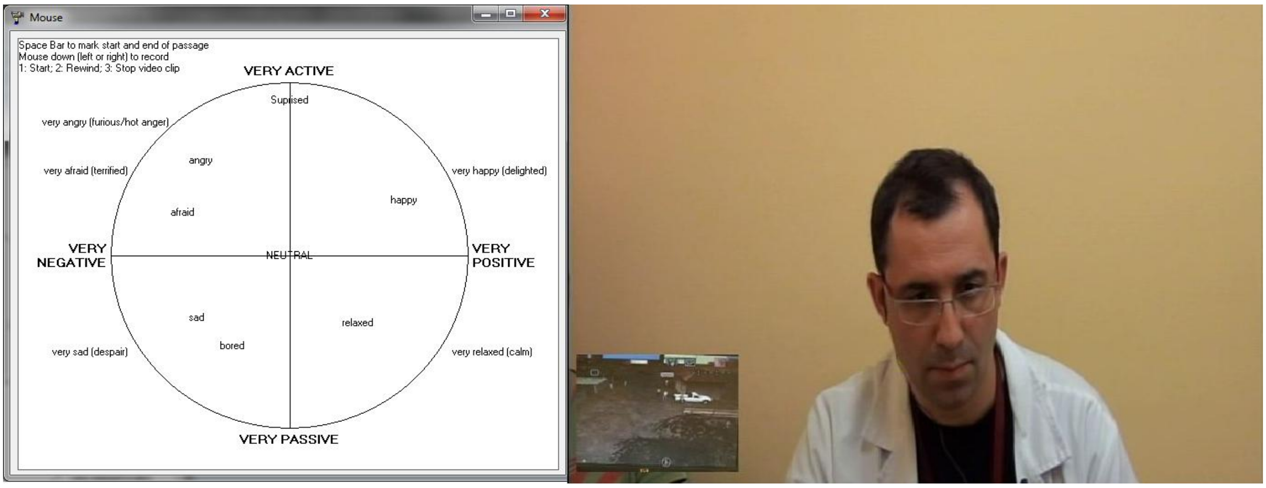

Figure 2: FeelTrace with customized emotion circle and example caption of video annotation material

4. Recordings

Volunteered psychiatric-healthy subjects participated in this study. The Ethics Committee of the University Hospital of Bellvitge (Barcelona, Spain) approved this study and informed consent was obtained from all participants.

The collected data consist of recordings from 18 people: eight males and ten females. The age and gender distribution of the test subjects are shown in Table 1. The age reported by the test subjects was in the range between 18 and 43 years old, with mean value of 29.3.

Table 1: Distribution of participants over age groups and genders

Age	**Males**	**Females**	**Total**	**%**
18-30	5	6	11	61.11
31-45	3	4	7	38.89
Total	8	10	18	100

The number of the test subjects with respect to their region of origin and gender is shown in Table 2.

Table 2: Number of the participants with respect to their region of origin and gender.

Region	**Males**	**Females**	**Total**	**%**
Catalonia	7	7	14	77.80
Sweden	0	1	1	5.55
Valencia	1	0	1	5.55
Argentina	0	1	1	5.55
Italy	0	1	1	5.55
Total	8	10	18	100

For all the recording sessions, the speech signal was captured using a conventional lapel microphone with sampling rate of 44.1 kHz, single channel, resolution 16 bits. All audio files were stored in .WAV format. The camera recordings consist of both audio and video captures saved in MPEG-2 format, with a Sony HDR-XR520VE video camera in a 720 x 576 resolution. The peripheral signals pulse-rate and oxygen saturation recorded from the test subjects were captured using the commercial device Mobi5 and transferred to a computer by the PortiLab software[1]. The bio-signals were saved as PortiLab binary format and converted to tab-separated files.

Realization of RS001 and RS002 resulted in total duration of speech recordings of approximately 2 and 4 minutes (84 and 132 items) collected per session, respectively. In RS003, approximately 60 minutes of audio-video and bio-signals were captured per session.

Consequently, the total duration of speech recordings collected during the first two recording sessions equals to approximately 110 minutes (3888 items). Further, the total duration of audio-video and bio-signal data collected is approximately 18 hours.

5. Annotation

The subjects were asked to self annotate their own affective states based on three types of information that were recorded during the game sessions, using the software FeelTrace (Cowie et al, 2001): The game screen when the subject was playing, her/his facial expressions recorded by camera at the time of playing and the vocals recorded by the microphone (please refer to Figure 2).

FeelTrace is an emotion annotation software based on a two-dimensional emotion model, containing emotion labels positioned by the dimensions activation and evaluation. This software has been chosen among other annotation tools because of its convenient interface and ease of use.

The annotation is conducted by continuously moving the mouse pressed around the circle. For the needs of PlayMancer project, we created a partially customized circle including:

- the emotions anger (angry), boredom (bored), joy (happy), neutral, sadness (sad), and surprise (surprised)
- high intensities of the emotions anger and sadness.

[1] Product of the company TMSi (http://www.tmsi.com/)

For time alignment, the annotators were instructed to move the mouse cursor to one of the emotions appearing in the circle when that emotion appears (based on their intuition) and move away from that label when the emotion ends or another emotion appears. In case there is no emotion, the annotators were instructed to move the mouse cursor to the label "NEUTRAL" in the middle of the circle.

6. Summary

The present work outlines the design and implementation of a multimodal affect database in support of research and development activities in serious game environment. A pilot trial, which consisted of three distinct recording sessions, was conducted for the collection of domain specific data. Speech, video and bio-signals were captured and annotated. The broad-shouldered design of the PlayMancer database allows its use for the needs of research on multimodal affect-emotion recognition and multimodal human-computer interaction in serious games environment.

7. Acknowledgements

This work was supported by the PlayMancer project (FP7-ICT-215839-2007), which is co-funded by the Seventh Framework Programme of the European Commission and CIBER (initiative of Instituto Carlos III). The authors wish to thank the European Commission as well as all members of the project consortium for their support.

8. References

Barab, S. A., Thomas, M, Dodge, Carteaux, R., and Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. Educational Technology Research and Developmen. 53(1), pp. 86--108.

Beale, I.L., Kato, P.M., Marin-Bowling, V.M., Guthrie, N., Cole, S.W. (2007). Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. In *J. Adolesc. Health*, 41(3), pp. 263--270.

Bergeron, B.P. (2008). Learning & retention in adaptive serious games. In *Stud Health Technol. Inform*. 132, pp. 26--30.

Botella, C., Villa, H., García Palacios, A., Quero, S., Baños, R.M., Alcaniz. M., (2004). The use of VR in the treatment of panic disorders and agoraphobia. *Stud. Health. Technol. Inform.*, 99., pp. 73--90.

CALLAS: Conveying Affectiveness in Leading-edge Living Adaptive Systems. URL: *http://www.callas-newmedia.eu*

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J. (2001). Emotion recognition in human-computer interaction. In *IEEE Signal Processing Magazine*, 18 (1), pp. 32--80.

Coyle, D., Matthews, M., Sharry, J., Nisbet, A., Doherty, G. (2005). Personal Investigator: "A therapeutic 3D game for adolescent psychotherapy". *Journal of Interactive Technology & Smart Education*, 2(2), pp. 73--88.

Difede, J., Cukor, J., Jayasinghe, N., Patt, I., Jedel, S., Spielman, L., Giosan, C., Hoffman, H.G. (2007). Virtual reality exposure therapy for the treatment of posttraumatic stress disorder following September 11, 2001. *J. Clin. Psychiatry*, 68(11), pp. 1639--1647.

ElderGames: Development of High Therapeutic Values IST-based Games for Monitoring and Improving the Quality of Life of Elderly People. URL: *http://www.eldergames.org*

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *New York: Palgrave Macmillan*.

Griffiths, K.M., Christensen, H., Jorm, A.F., Evans, K., Groves, C. (2004). Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br *J. Psychiatry*. 185, pp. 342--349.

Griffiths, M. (2004). Can videogames be good for your health? *J. Health. Psychol*., 9(3), pp. 339--344.

Honig, F., Batliner, A., Nöth, E. (2007). Real-time recognition of the affective user state with physiological signals. In *Affective Computing and Intelligent Interaction, Doctoral Consortium*, pp. 1--8.

HUMAINE: Research on Emotions and Human-Machine Interaction. URL: *http://emotion-research.net*

Jiménez-Murcia, S., Fernández-Aranda, F., Kalapanidas, E., Konstantas, D., Ganchev, T., Kocsis, O., Lam, T., Santamaría, J.J., Raguin, T., Breiteneder, C., Kaufmann, H., Davarakis. C. (2009). Playmancer project: a serious videogame as an additional therapy tool for eating and impulse control disorders. In *Stud. Health Technol. Inform. 2009*, 144: pp. 163--166,

Katsis, C.D., Katertsidis, N., Ganiatsas, G., Fotiadis, D.I. (2008). Toward Emotion Recognition in Car-Racing Drivers: A Biosignal-Processing Approach. In *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol 38, No 3, pp. 502--512.

Kocsis, O., Ganchev, T., Mporas, I., Papadopoulos, G., Fakotakis, N. (2009). Multi-modal system architecture for serious gaming. In *Artificial Intelligence Applications and Innovations III, Series: IFIP Advances in Information and Communication Technology*, Vol. 296, Iliadis, Lazaros, Vlahavas, Ioannis, Bramer, Max (Eds), pp. 441--447.

Lange, A., van de Ven, J.P., Schrieken, B., Emmelkamp, P.M. (2001) Interapy, treatment of posttraumatic stress through the Internet: a controlled trial. In *J. Behav. Ther. Exp. Psychiatry*. 32(2), pp 73--90.

Myers, T.C., Swan-Kremeier, L., Wonderlich, S., Lancaster, K., Mitchell. J.E. (2004). The use of alternative delivery systems and new technologies in the treatment of patients with eating disorders. *Int. J. Eat. Disord*., 36(2), pp. 123--143.

PlayMancer: A European Serious Gaming 3D Environment. URL: *http://www.playmancer.eu*

Russoniello, C.V., O'Brien, K., Parks, J.M. (2009) EEG, HRV and Psychological Correlates while Playing Bejeweled II: A Randomized Controlled Study. *Stud. Health Technol. Inform.* 144, pp. 189--192.

Saitz, R., Helmuth, E.D., Aromaa, S.E., Guard, A., Belanger, M., Rosenbloom, D.L. (2004). Web-based screening and brief intervention for the spectrum of alcohol problems. *Prev. Med.* 39(5), pp. 969--975.

Schott, G., Hodgetts. D. (2006). Health and Digital Gaming, The Benefits of a Community of Practice. *J. Health. Psychol.*, 11(2), pp. 309--316.

Squire, K., (2003). Video games in education. *International Journal of Intelligent Simulations and Gaming* (2)1 n.p.

Wagner, J., Kim, J., Andre, E. (2005). From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In *Int. Conf. on Multimedia & Expo (ICME 2005)*, pp. 940--943.

Zarate, C.A., Weinstock, L., Cukor, P., Morabito, C., Leahy, L., Burns, C., Baer. L. (1997). Applicability of telemedicine for assessing patients with schizophrenia: acceptance and reliability. *J. Clin. Psychiatry*, 58(1), pp. 22--25.