DYNAMIC MULTIMEDIA CONTENT DELIVERY BASED ON REAL-TIME USER EMOTIONS
Multichannel Online Biosignals Towards Adaptative GUI and Content Delivery

Vasco Vinhas, Luís Paulo Reis and Eugénio Oliveira

FEUP - Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias s/n, Porto, Portugal
DEI - Departamento de Engenharia Informática, Rua Dr. Roberto Frias s/n, Porto, Portugal
LIACC - Laboratório de Inteligência Artificial e Ciência de Computadores, Rua do Campo Alegre 823, Porto, Portugal

Keywords: Affective Computing, Emotion Assessment, Biosignals, Multimedia, Interfaces.

Abstract: Recently topics such as affective computing and multichannel multimedia distribution have gained the attention and investment of both industry and academics. The proposed system joins these domains so that ubiquitous system can be potentiated by means of online user emotion assessment based on real-time user's biosignals. It was used IAPS as a emotional library for controlled visual stimuli and biosignals were collected in real-time - heartbeat rate and skin conductance - in order to online assess the user's emotional state through Russell's Circumplex Model of Affect. To improve usability and session setup, a distributed architecture was used so that software models might be physically detached. The conducted experimental sessions and the validation interviews supported the system's efficiency not only in real-time discrete emotional state assessment but also considering the emotion inducing process. The future work consists in replicating the success in multi-format multimedia contents without pre-defined emotional metadata.

1 INTRODUCTION

Emotional state assessment constitutes a transversal research topic that has captured the attention of several knowledge domains. In parallel with this reality, both the multimedia industry and ubiquitous applications have gained a crescent academic and industrial significance and impact. Having this in consideration, one shall refer the integration of the enunciated domains as an opportunity to potentiate each of the areas and explore mutual synergies. The present project intends to perform automatic real-time discrete user emotional state assessment and with this information, and by following a flexible emotional policy, deliver the next multimedia content appropriately. The authors believe that the success of such system would enable the intention of developing a fully automatic affective system that would be able to provide the user the exact multimedia content that he or she would like best to be presented with, in terms of emotional content. This is considered to be a major breakthrough with immediate practical applications not only to multimedia content providers but also to videogame industry, marketing and advertisement and even medical and psychiatric procedures. The outcome of such project would also be useful, mainly considering the emotion assessment engine, to greatly enhance ubiquitous computing through user interfaces immediate adaptation to the user's emotional profile.

This document is structured as follows: In the next section, the current state of the art, considering emotion induction and classification is presented; in section 3, the project's global architecture and functionalities are depicted; and its experimental results are illustrated through section 4; finally in the last section, conclusions are withdrawn and the most significant future work areas are identified.

2 STATE OF THE ART

This section is reserved to refer and detail state of the art regarding emotion representation, assessment and induction in a cross-cultural way.

2.1 Emotion Representation and Induction Methods

Until a recent past, researchers in the domains related to emotion assessment had very few solid ground standards both for specifying the emotional charge of stimuli and also a reasonable acceptable emotional state representation model. This issue constituted a serious hurdle for research comparison and conclusion validation. The extreme need of such metrics led several attempts to systematize this knowledge domain.

Considering first the definition problem, Damásio states that an emotional state can be defined as a collection of responses triggered by different parts of the body or the brain through both neural and hormonal networks (Damásio, 1998). Experiments conducted with patients with brain lesions in specific areas led to the conclusion that their social behaviour was highly affective, together with the emotional responses. It is unequivocal to state that emotions are essential for humans, as they play a vital role in their everyday life: in perception, judgment and action processes (Damásio, 1994).

One of the major models of emotion representation is the Circumplex Model of Affect proposed by Russell. This is a spatial model based on dimensions of affect that are interrelated in a very methodical fashion (Russel, 1980). Affective concepts fall in a circle in the following order: pleasure, excitement, arousal, distress, displeasure, depression, sleepiness, and relaxation - see Figure 4B. According to this model, there are two components of affect that exist: the first is pleasure-displeasure, the horizontal dimension of the model, and the second is arousal-sleep, the vertical dimension of the model. Therefore, it seems that any affect stimuli can be defined in terms of its valence and arousal components. The remaining variables mentioned above do not act as dimensions, but rather help to define the quadrants of the affective space. Although the existence of criticism concerning the impact of different cultures in emotion expression and induction (Altarriba, 2003), Russell's model is relative immune to this issue if the stimuli are correctly defined in a rather universal form. Having this in mind, the circumplex model of affect was the emotion representation abstraction used in the proposed project.

Regarding induction methods, in order to analyse biometric data that contains a discrete set of emotional states, it is essential to create and define an experimental environment that is able to induce a subject in a specific and controlled emotional state.

It is a common practice to use an actor as one possible approach to human beings emotions' simulation (Chanel, 2005). As the actor predicts specific emotions, outside aspects as facial expression or voice change accordingly. However, the physiological responses will not suffer any variations, which lead to one of the biggest disadvantages of this approach, as the gathered biometric information does not represent the real emotional state of the actor. An alternative method, adopted in this study, is the use of multimedia stimuli (Chanel, 2005). These stimuli contain a variety of contents such as music, videos, text and images. The main advantage of this method resides in the strong correlation between the induced emotional states and the physiological responses, as the emotions are no longer simulated.

2.2 International Picture Affective System

To assess the three dimensions of pleasure, arousal, and dominance, the Self-Assessment Manikin (SAM) (Lang, 1980) was used. A graphic figure depicting values along each of the three dimensions on a continuously varying scale is used to indicate emotional reactions as depicted in Figure 1. Each picture in the IAPS (Lang, 2005) is rated by a large group of people, both men and women, for the feelings of pleasure and arousal that the picture evokes during viewing.
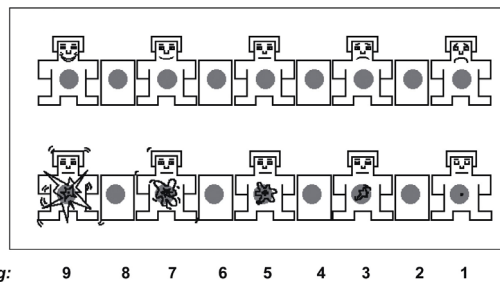
| Scoring: | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |

Figure 1: The self-assessment manikin (SAM)(Lang, 1980).

The IAPS library was developed to provide ratings of affect for a large set of emotionally-evocative, internationally accessible, color photos that include contents across a wide range of semantic categories (Bradley, Lang) so that cultural and intrinsic variables could be, as much as possible, discarded from the evaluation.

2.3 Equipment Solutions

Emotions assessment needs reliable and accurate communications with the subject so that the results are conclusive and the emotions correctly classified. This communication can occur through several channels and is supported by specific equipment. The invasive methods are clearly more precise, however more dangerous and will not be considered for this study. On the other hand, non invasive methods such as EEG, fMRI, GSR, oximeter and others have pointed the way towards gathering together the advantages of inexpensive equipment and non-medical environments with interesting accuracy levels.

Due to the medical community scepticism, EEG, in clinical use, it is considered a gross correlate of brain activity (Ebersole, 2002). In spite of this reality, medical research studies (Aftanas, 1997) have been trying to revert this scenario by suggesting that increased cortical dynamics, up to a certain level, are probably necessary for emotion functioning and by relating EEG activity and heart rate during recall of emotional events. Similar efforts, but using invasive technology like Electrocorticography, have enabled complex BCI like playing a videogame or operating a robot (Leuthardt, 2004). Some more recent studies have successfully used just EEG information for emotion assessment (Ishino, 2003). These approaches have the great advantage of being based on non-invasive solutions, enabling its usage in general population in a non-medical environment. Encouraged by these results, the current research direction seems to be the addition of other inexpensive, non-invasive hardware to the equation. Practical examples of this are the introduction of GSR and oximeters by Takahashi (Takahaski, 2004), Kim (Kim, 2008) and Chanel (Chanel, 2005). For this study, the Oxicard oximeter and the ToughtStream GSR equipments shall be used.

3 PROJECT DESCRIPTION

3.1 Architecture

The main principle of the proposed system resides in the fact of its distribution capability as each of its main modules can be allocated and run in different machines - in spite of being possible to concentrate all functional units in one computer. As illustrated in Figure 2, the three key modules are the biosignals equipment data collector, the emotional picture database and the interpretation and visualization control module.
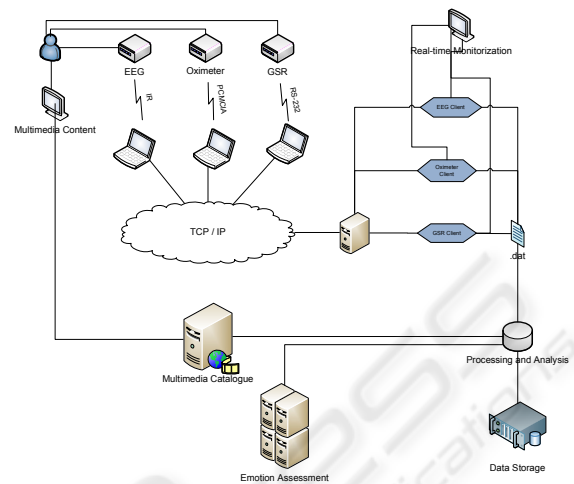

Figure 2: System's Global Architecture.

This architecture independence enables physical distribution through several machines, although this feature can be overridden by placing all software units in a single entity, and simultaneously allows the system's usage in several environments as it is easy to setup a new experimental session with this degree of freedom as long as there is a TCP/IP network connecting all modules.

The database unit consists in the relational module depicted in subsection 3.2 and intends to constitute a replica of the IAPS library but in a relational database environment. The current model is stored in the campus's Oracle 10g server but the system is considered to be database agnostic. This unit's intention is not only IAPS emulation but also to constitute both a biosignals record repository - by gathering all data provided during experimental sessions - and a refined base station for emotion classification. The system's second unit consists in the biosignals collection module that by itself constitutes a distributed system. Its main characteristics are based on the hardware set composition flexibility as the aggregation unit supports several equipments. The instantiated module is an adaptation of the Multichannel Emotion Assessment Framework (Teixeira, 2008) as, for this particular project, it is used the oximeter data and the GSR equipment.

Each of this hardware equipments are physically attached to the user by means of minimal invasive techniques: the oximeter sensor just plugs around any user's hand finger and the GSR dry electrodes make contact with the user's left or right hand palm

through a wrist-band lookalike. The used oximeter, Oxicard, connects to a personal computer through a PCMCIA and the GSR, ThoughtStream, uses a standard RS-232 connection. For each of the equipments, there were designed and developed distinct drivers both with TCP/IP communication capabilities in order to efficiently collect and further distribute the data to one or multiple client applications. This driver network capability was properly exploited in this project in order to enable data collection in one computer and data processing in another. Finally, one shall mention the interpretation and visualization unit. This module is responsible for accessing the available data provided by the collector unit and from this determine the most likely user's emotional state and from the previously defined emotional policy, the user's given baseline and the historical data determine and extract, from the database, the next picture to present.

3.2 Database Model

The key function of the developed database resides in the replication of the IAPS offline file system based picture collection. As described in section 2, this library consists in thousands of pictures with emotional metadata, namely valence, arousal and dominance values. The first undertaken action was to design a database relational model to accommodate such collection and further load it with the data - emotional metadata and picture representations as BLOBs. In spite of the importance of this, it is possible to understand, through Figure 3, that the database model was designed to store other information rather than IPAS emulation.

Having this in mind, the first extra enhancement resides in user information record as it is important to know who is performing the session as each user as different emotional state triggers, value definitions, and biosignals baselines. Using this user identification, It was designed session models in order to accommodate all biosignals data provided by each of the used equipments, both for offline analysis and also online relative real-time emotional state assessment. Another issue that needs to be addressed is the session emotional policy. In other words, what shall the system do in what concerns next picture retrieval: corroborate current emotional state by choosing a similar content; contradict current emotional state by picking an antagonist picture or simple lock a desired state such as joy, sadness, excitement, neutral, amongst others.
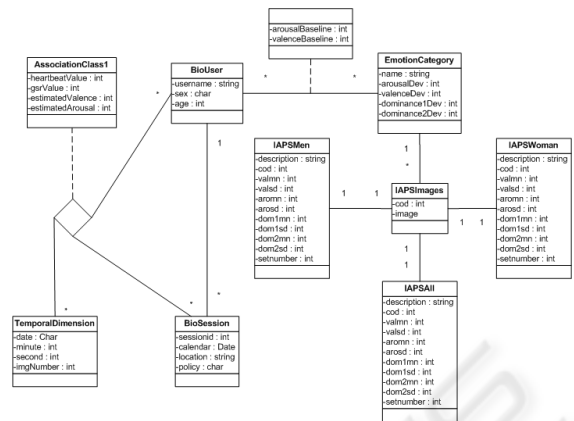

Figure 3: Database Model.

3.3 Calibration & Assessment

The multimedia content visualization tool, Figure 4A, its graphical user interface is designed to be as simplest as possible in order not to maximize the user's immersion sensation. Following this principle, for default, the application runs in full screen mode stretching the selected picture from the database to whole screen resolution. The picture is displayed for six seconds, interval that when ended triggers the emotion assessment procedure so that the next picture is selected and downloaded.
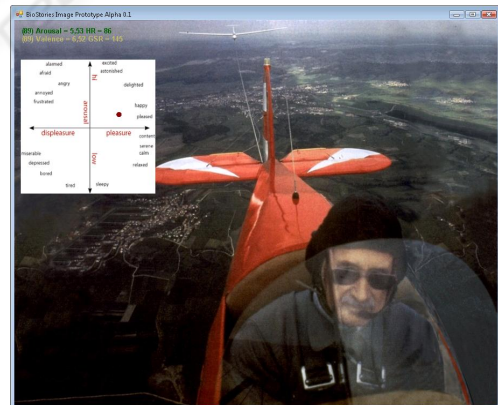

Figure 4A: Application Running Screenshot.

Mainly for debugging/analysis purposes, therefore it is not intended for user's interaction and even perception, there were added two additional interface controls. The first is the textual information regarding both the heartbeat rate collected by the oximeter and the skin conductance value provided by the GSR equipment. Also in this textual form, located in the top-left corner of the screen it is possible to acknowledge the actual number of

pictures that match the current emotional state. The second added control is the emotional bi-dimensional space considering valence and arousal values, commonly named as the circumplex model of affect. As more visible in Figure 4B, in this model it is possible to directly map arousal in the Y axis and valence In the X axis, the heartbeat rate was used as arousal indicator where high rates mean high arousal levels and for valence control the skin conductance was used as high conductance values generally mean higher levels of transpiration and therefore tension and displeasure, on the other hand, low levels of transpiration tend to point no stress indicators, thus act as pleasure indicators.
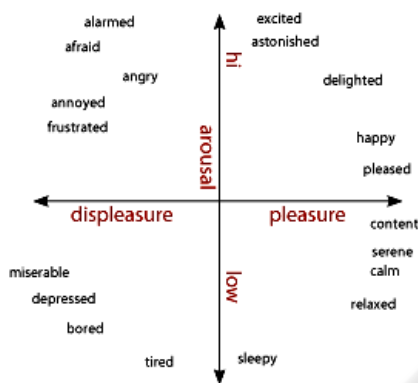

Figure 4B: Circumplex Model of Affect.

By asking the users, at their first application usage, to indirectly define their emotional baseline by, after equipment data collection enabling, to point what is their most accurate current emotional state it was possible to define, for each individual a fully adapted affect model and bidimensional space. Although, in the first interaction, there is only one defined point of affect baseline and the positioning in the emotional state domain is effectuated by default, there is the possibility, as partially depicted in the previous subsection, to the user define multiple baseline points do that space navigation become more accurate as it is done with more user-specific information. In what concerns the *next picture* policy, this action is conducted based on the session's defined policy and if this is not fixed to one particular state, it is inferred the current user's emotional state - that can be optionally visualized through the red dot drawn over the circumplex model - and the scaled to 1 to 9 valence and arousal values are extracted. If the defined policy is to corroborate the current state, it is defined a 10% tolerance area around the emotional point and a random picture is retrieved from the database that

matches. If the policy is to contradict the current state, the search domain is reversed around the neutral point and again a random content amongst the eligible ones is selected to presentation. In the remaining cases of fixed policy, a typical emotional category baseline is gathered from the database and a similar tolerance area is defined. In spite of the remaining process being known, if the user's emotional state does not converge to the desired one, the tolerance area is continually shifted to more extreme valence and/or arousal values.

4 RESULTS

In order to produce metrics to evaluate the conducted process, two interviews were conducted to users after the experimental session's end. The experimental sessions were composed of six series of ten pictures, and the analyzed sample was composed by twenty-five undergraduate students with no prior knowledge of the project's characteristics.

Table 1: Emotion Assessment Confusion Table.

		Automatic Assessment			
		1st Quadrant	2nd Quadrant	3rd Quadrant	4th Quadrant
Users	1st Quadrant	**20%**	3%	2%	4%
	2nd Quadrant	6%	**11%**	2%	1%
	3rd Quadrant	1%	2%	**15%**	5%
	4th Quadrant	6%	4%	4%	**14%**

In the first interview, it was asked to the users to point what was the most significant emotional state for each of the ten pictures that they have just seen. The results are shown in Table 1. Considering the emotion induction process, a similar interview process was performed and the extracted data is condensed in Table 2.

Table 2: Emotion Induction Confusion Table.

		Automatic Assessment			
		1st Quadrant	2nd Quadrant	3rd Quadrant	4th Quadrant
Users	1st Quadrant	**20%**	3%	2%	4%
	2nd Quadrant	6%	**11%**	2%	1%
	3rd Quadrant	1%	2%	**15%**	5%
	4th Quadrant	6%	4%	4%	**14%**

It is clear that most of users tend to locate themselves in the first quadrant. The system achieves better classification results in assessing emotional states located in the first and third quadrants - 61% and 65%. In this case, the system was used to put into practice a fixed emotion induction policy with fixed emotional states. It was performed, for each user, with an appropriate time

interval, four sessions, each one to each particular quadrant. Through the data analysis it is possible to state that induction towards the first and third quadrant is more effective and induction towards the second quadrant was not successful - this is particularly due to distinct individual reactions to the presented content.

5 CONCLUSION AND FUTURE WORK

The distributed architectural paradigm proved to be robust and effective, preserving modularity. It was achieved an immersive interface that capably was able to retrieve biosignals data and access the picture database. Secondly, the automatic emotion assessment following the enunciated state distribution through Russell's model, according to the performed interviews, showed to achieve success rates of 65% - in a four hypothesis situation. On the other hand, the emotion induction, by means of IAPS library usage and valence/arousal values, was particularly successful with hit rates of 70-80% for three of the four quadrants. Considering the above mentioned results, the authors are interested in further exploiting this approach by refining emotional state assessment through adding biosignals, such as respiratory movements and electromyography to therefore perform information fusion to axis movement. Another development considering emotional assessment was the fully comply with the third dimension represented by dominance - that has not been subject of study in the presented project.

Orthogonally, the authors have identified project's extensions, in order to enhance the whole system's applicability in several practical domains. The first improvement should be the constitution of a multimedia database composed not only by pictures with a significant metadata layer so that it would be possible to build, in real-time, a dynamic storyline. This feature would enable flexible storytelling based on audience emotion non-intentional feedback. This type of systems would have vast applicability in all entertainment industry, marketing and advertisement as well as user interfaces enhancement. Its appliance would also be possible and even desirable in medical, especially psychiatric, procedures namely in phobia treatment emotional response assessment.

REFERENCES

Aftanas, L. I., el al. 1997. Nonlinear forecasting measurements of the human eeg during evoked emotions. *Brain Topography*. pp. 155-162.

Altarriba, J., Basnight, D. M., & Canary, T. M. 2003. *Emotion representation and perception across cultures*. In W. J. Lonner, D. L. Dinnel, S. A. Hayes, & D. N. Sattler (Eds.), Readings in Psychology and Culture (Unit 4, Chapter 5), Center for Cross-Cultural Research, Western Washington University, Bellingham, Washington USA.

Chanel, G., et al. 2005. Emotion Assessment: Arousal Evaluation using EEG's and Peripheral Physiological Signals. *University of Geneva, Switzerland: Computer Science Department*, pp 530-537, vol 4105/2006.

Damásio, A. R. 1994. *Descartes error: Emotion, reason and human brain*. Europa-América.

Damásio, A. R. 1998. *Emotions and the Human Brain*. Iowa, USA: Department of Neurology.

Ebersole, J. (2002). Current Practice of Clinical Electroencephalography. *Lippincott W&W.*

Ishino, K. (2003). A feeling estimation system using a simple electroencephalograph. *Journal of Neural Engineering*, pp. 63-71.

Kim, J. & André, E (2008). *Multi-Channel BioSignal Analysis for Automatic Emotion Recognition*. Biosignals 2008.

Lang, P. J. 1980. *Behavioral treatment and bio-behavioral assessment: Computer applications*. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–137). Norwood, NJ: Ablex.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. 2005. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. University of Florida, Gainesville, FL.

Bradley, B. C. & Lang, P.G. Studying emotion with the *International Affective Picture System* (IAPS). In J. A. Coan and J.J. B. Allen (Eds). *Handbook of Emotion Elicitation and Assessment*. Oxford University Press.

Leuthardt, E. 2004. A braincomputer interface using electrocorticographic signals in humans. *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4204-4209..

Russell, J. A. 1980. *A circumplex model of affect*. Journal of Personality and Social Psychology, 39, 1161-1178.

Takahaski, K. 2004. Remarks on Emotion Recognition from Bio-Potential Signals. *2nd International Conference on Autonomous Robots and Agents* pp 1148-1153, vol 3.

Teixeira, J., Vinhas, V. et al. 2008. Multichannel Emotion Assessment Framework: Gender and High-Frequency Electroencephalography as Key-Factors. *ICINCO 2008.*