Proceedings of the Third International Symposium on AI & Games, Daniela M. Romano and David C. Moffat (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

A Flexible Bio-Affective Gaming Interface

Jorge Arroyo-Palacios[1], Daniela M. Romano[1]

Abstract. Affective bio-feedback can be an important instrument to enhance the game experience. Several studies have provided evidence of the usefulness of physiological signals for affective gaming. However, due to discrepancies in the findings of psychology studies, the pattern matching models employed are limited in the number of emotions they are able to classify. This paper presents a bio-affective gaming interface (BAGI) that can be used to customize a game experience according to the player's emotional response. Its architecture offers important characteristics for gaming such as: flexibility, scalability and multi-category discrimination in real time. These features are important not only because they allow the classification with pattern matching and machine learning models but also because they make possible the reusability of previous findings and the inclusion of new models to the system.

In order to prove the effectiveness of BAGI two different types of neural networks have been trained to recognize emotions. After that, they were incorporated into the system to customize in real-time the computer wallpaper, according to the emotion experienced by the user. The best results were obtained with a probabilistic neural network with accuracy results of 84.46% on the training data and 78.38% on the validation for new independent data sets.

1.- INTRODUCTION

Physiological signals offer a promising medium to interact in a natural and intuitive way with the game environment. In addition of being reliable, sensible and provide real time feedback, physiological signals offer an insight into human's physical and mental state which can be used to enrich the game interaction. Common physiological measures already used in gaming or Human Computer Interaction (HCI) include: cardiovascular, electrodermal, muscular tension, ocular, skin temperature, brain activity and respiration measures. The game industry has already seen their potential and several research studies have been conducted to investigate the best way to use them.

One of the earliest commercial examples of a bio-adaptive interface used in commercial games can be seen in "Tetris 64"[2], released for the Nintendo 64 platform. In "Tetris 64" a sensor is placed at the player's ear to monitor the heart rate, the game speeds up or slows down depending on the player's cardiovascular activity. A more modern commercial example is "Relax & Race"[3] by Vyro games. In "Relax & Race" the game takes the form of a race between two characters: stress and the player. The game senses the electrodermal activity (EDA) of the player and increases the speed of the character as the player relaxes.

There are also several research studies that explore physiological signals in gaming. Toups et al. [1] for example use electrodermal and electromyographic (EMG) activity as an indication of increased attention, effort and stress in team game play. In their game, the computer opponents sense the physiological signals and pursue players with higher activation levels. One interesting feature of this study was that the physiological information was provided to the team coordinator thus he could modify the game strategy. Another study that also uses EDA and EMG activity was carried out by Predinger et al. [2] where the emotional response of a character was adapted according to physiological responses of the player in a virtual card game.

Gilleade et al. [3] proposed that the identification of emotions, through physiological signals, can provide means to the Artificial Intelligence (AI) of the game for automatically: i) assist (e.g when the player is frustrated), challenge (e.g. when the player is bored), emote (to enhance the game experience of the player). In their studies they developed a game that monitors the cardiovascular activity to increase or decrease the number of threats in a game.

Educational games and tutoring systems have also employed the use of physiological signals to provide affective feedback to the system and adapt according to the student emotional state [4].

Research in brain activity has lead to the creation of brain-computer interfaces (BCI) which have been used as a game interface control for an immersive 3D gaming environment [5] and for moving a ball on a table in Brainball [6]. Ko et al. presents in [7] the emotion recognition capabilities for games of three currently available portable BCI devices: NIA[4], Emotive EPOC[5] and Mind Set[6].

Respiration measures have also been used to interact with games. Arroyo and Romano [8] developed a simple balloon game where the player blows up a virtual balloon with his chest expansion and contraction while breathing. Table 1 presents a summary of some games that recognize emotions by the use of physiological signals.

Game	Signal	Detect	Emotion	Action
Tetris 64	HR	Increase	Anxiety	Game slows down (normal mode) Game speeds up (reverse mode)
		Decrease	Relaxation	Game speeds up (normal mode) Game slows down (reverse mode)

[1] Dept. of Computing Science, Univ. of Sheffield
Email: {j.arroyo, d.romano}**@dcs.shef.ac.uk**.
[2] http://uk.ign64.ign.com/articles/154/154081p1.html
[3] http://www.vyro-games.com/products/relaxandrace.php
[4] http://www.ocztechnology.com/products/ocz_peripherals/nia-neural_impulse_actuator
[5] http://www.emotiv.com/
[6] http://www.neurosky.com/

Relax & Race	SR	Increase	Relaxation	Character slows down
		Decrease	Anxiety	Character speeds up
PhysiRogue [1]	EDA, EMG	Increase	Attention, effort, stress.	Opponents pursue players with highest activation levels
Card game [2]	EDA, EMG	Increase	High arousal, negative valence	If player makes progress agent shows: -joy (empathic mode) -fear (negative empathic mode)
Similar to Missile Comm and [3]	HR	Increase	Over stimulation	Reduce number of threats.
		Decrease	Lack of engagement	Increase number of threats
Prime Climb [4]	EMG, EDA, HR	Specific changes	Reproach, shame and joy	Depending on emotion, agent can intervene in game.
Brain ball [6]	Brain activity	Specific changes	Relaxation	Ball is moved
Blow up balloon [8]	RSP	Increase and decrease	None	Virtual balloon is blown up
Relax to Win [9]	SR	Increase	Relaxed	Character slows down
		Decrease	Anxious	Character speeds up

HR = Heart rate SR = Skin Resistance
RSP = Respiration

Table 1.- Use of physiological signals in games.

There are two important observations from the literature reviewed:

i) Most of the emotion recognition systems (ERS) follow an ad hoc strategy (i.e. they provide a solution designed for a specific problem or task, non generalizable, and which is not easily adaptable to other purposes).

ii) There are very few multi-category discrimination emotion recognition systems implemented in real-time. All the emotion recognition systems of the games presented in table 1 are aimed to identify the presence or absence of a particular emotion, or to discriminate among small sets of opposite emotions.

Nowadays, games are being produced for a great variety of domains including: entertainment, training, medical therapies, education and socialization. Users are more demanding in their expectations, and current games involve more photographic realism, cinematic sequences and new interaction methods. Emotions recognition systems can contribute to enhance the emotional experience, however their current construction strategy is expensive, time-consuming, limited (in discrimination capabilities) and does not allow the reusability of tools and previous findings.

In this paper the architecture, implementation and validation of a flexible multi-category emotion recognition system is presented. Specifically, this paper contributes to the research efforts in affective gaming by presenting:

- Implementation of a real-time multi-category emotion recognition system.
- A flexible and scalable architecture that facilitates the integration of findings from other studies.
- Real-time classification results from two machine learning algorithms.
- A simple application that adapts the desktop wallpaper according to the emotional state of the user.

2.- VARIABLE PARAMETERS IN EMOTION RECOGNITION SYSTEMS

2.1.- Set of emotions.

Depending on the type of application or game, different sets of discrete emotions need to be considered. For instance, adventure/action games might need the identification of emotions such as: enjoyment, frustration, anxiety, boredom. In learning/training games shame might need to be considered too (as suggested in [4]). Social games such as Second Life[7] could find useful to include in the recognition emotions such as: love, anger and sadness. For horror games such as Dead Space[8], FEAR[9] and Resident Evil[10], fear and anxiety might need to be recognized.

Besides the use of discrete emotions, researchers have also proposed the use of dimensional models for representing emotions in games (e.g. [10]). In this way a set of emotion for a game could be composed by the following 4 affective states: 1) positive valence – high arousal, 2) positive valence – low arousal, 3) negative valence – high arousal, and 4) negative valence – low arousal.

2.2.- Measures and features to analyze

Common physiological measures already used in gaming or HCI include: cardiovascular, electrodermal, muscular tension, ocular, skin temperature, brain activity and respiration measures. Each of these measures can be more or less convenient depending on the type of game and emotions to be recognized. From these measures, several features can be calculated from various analysis domains including: statistics, frequency domain, geometric analysis, multiscale sample entropy and sub-band spectra. Reviewing some relevant studies in HCI it can be seen that they have privileged the use of different physiological measures over others, and have extracted different numbers of features to analyze. For instance, Kim and Andre [11] extracted a total of 110 features from 4 sensors: EMG, ECG, SC and RSP; Picard et al. [12] extracted 40 feautres from EMG, BVP, RSP and SC; Nasoz et al. [13] used GSR, temperature and HR signals.

2.3.- Times to analyze

Another variable parameter among the emotion recognition systems in the literature concerns to the time window used to

[7] http://secondlife.com/
[8] http://www.visceralgames.com/us/page/deadspace
[9] http://www.whatisfear.com/
[10] http://www.residentevil.com/index.php

analyze the physiological signals. For instance, Picard et al. [12] used a time window of 100 seconds from the end of each emotion segment to analyze all the signals in their first data set; Kim and Andre [11] segmented their data in samples of 160 seconds by taking the middle part of each signal; Nasoz et al. [13] used a time frame of 7 seconds when the emotion was most likely to be experienced.

2.4.- Approaches for physiological pattern classification

2.4.1.- Pattern matching

The emotion recognition systems in the games listed in table 1 are implemented following a pattern matching approach. Following this strategy, the system detects the presence of the constituents of a rigidly specified pattern in the physiological data. The classification process can be based on existing models describing the relationship between psychology and physiology, like for example the model proposed by Cacioppo & Tassinary [14] or use empirical findings from literature, for example, Stemmler [15]. The dilemma is to decide which model and physiological definitions, the system should be based on.

Table 2 illustrates some patterns of heart rate associated to emotions from different studies that a computer system could use to discriminate a set of three emotions composed by fear, joy and anger.

Author	Signal	Fear	Sadness	Happiness
Christie [16]	Heart rate	Increase	Decrease	Decrease
Ekman et al. [17]	Heart rate	Increase (high)	Increase (high)	Increase (low)
Frederickson [18]	Heart rate	Increase	Decrease	No significant difference

Table 2.-Patterns associated to emotions from different studies.

A vast number of studies in psychology document the autonomic responses to emotions. However not all of them followed the same methodology. Unfortunately, the differences are present in critical elements of the study such as: model and set of emotions investigated, stimuli used to evoke the emotions, physiological measures, features extracted, and type of analysis on the data (already highlighted in [19]). As a result some discrepancies remain among their findings, as it can be observed in Table 2 and, in more detail, in the compilation of physiological responses to specific emotions in [20, 21].

Despite the discrepancies among psychological studies, in the gaming literature pattern matching techniques have demonstrated to be practical especially for: i) determining the presence or absence of a particular emotion, or ii) discriminating small sets of opposite affective states. Additionally pattern matching offers important characteristics for a real time interaction such as simple implementation and very fast computation.

2.4.2.- Pattern recognition

Another approach that can be used to implement emotion recognition for gaming is pattern recognition. It implies the training of a system to recognize a set of affective states corresponding to a particular emotional model. The machine learning model receives as input labelled sets of physiological data and then it is trained to detect the patterns associated to each emotion. The model is then able to classify new sets of physiological data into specific emotions. As a result, a physiological definition for different affective states has been obtained through empirical data. The advantage of this approach is that it offers ecological validity on the emotions recognized (the settings during the pattern recognition process can be identical or similar to the classification settings) and provides good classification rates among several emotions (as showed in several offline HCI studies [11-13, 22-24]).

Despite the advantages of this approach it has not been fully exploited by the game industry, perhaps due to concerns that might arise as consequence of:

- The time that a pattern recognition process takes (depending on the processing power, algorithms and quantity of data, it can take up to several hours or days.
- Difficulties to reuse the system for other purposes or to reuse the finding from other studies.
- Few studies show good results from user independent and real-time implementations (e.g. [25]).
- Lack of guided principles for its application
- Lack of physiological affective databases to train the systems.

In the research presented in this paper some of these concerns have been addressed.

2.5.- Importance of flexibility in ERS

There is a wide range of applications where emotion recognition systems can be very useful however the development of such system for a particular situation is complex, multifaceted and very time-consuming. It requires taking into consideration theoretical, methodological, technical and some other aspects.

The ad hoc strategy followed so far has proven good classification results; nevertheless, the reusability of systems and findings is very difficult and the comparison of results among experiments carried out in different labs is almost impossible.

The flexibility in an emotion recognition system would facilitate its use in different type of applications, allow the reusability of findings and facilitate the comparison of results. It would be also possible to use the classification strategy that best suit to our application. Moreover, the research community is continuously proposing new theories and providing new discoveries. A flexible architecture on the set of emotions, physiological signals, time windows and classification strategy would also allow scalability for the inclusion of advances in emotion research.

Table 3 presents some machine learning based emotion recognition studies where different parameters were considered for the classification of emotions.

Study	Emotions	Time window	Measures	Analysis technique
Kim & Andre [11]	High arousal, low arousal, positive valence, negative valence	160 seconds	EMG, ECG, SC, RSP (110 features)	LDA, Binary tree model

Picard et al. [12]	No emotion, anger, hate, grief, platonic love, romantic love, joy, reverence.	100 seconds	EMG, BVP, RSP, SC (40 features)	SFFS, FP, SFFS-FP
Nasoz et al. [13]	Sadness, anger, fear, surprise, frustration, amusement	7 seconds	EDA, ST, HR	KNN, DFA, MBP
Kim et al. [22]	Fear, neutral, joy	50 seconds	ECG	K-means based expectation maximization
Haag et al. [23]	Valence and arousal	2 seconds	ECG, BVP, SC, EMG, RSP, ST	NN trained with resilient propagation
Kim et al. [24]	Sad, stress, anger, surprise	50 seconds	ST, EDA, ECG, BVP	SVM

ST = Skin Temperature
LDA = Linear Discriminant Analysis
MBP = Marquardt Backpropagation
SFFS= Sequential Floating Forward Selection
SC = Skin Conductance
NN = Neural Network
SVM = Support Vector Machine
FP = Fisher Projection

Table 3.- Different parameters in offline emotion recognition studies

3.- BAGI – DESIGN AND IMPLEMENTATION

One of the most important characteristics of BAGI is its architecture which allows: flexibility, scalability, multi-category discrimination and real time affective feedback. The innovative aspect resides in the classification module which provides two important characteristics: i) flexibility to receive different configurations in the parameters mentioned in section 2, and ii) ability to incorporate different models obtained through a machine learning process for a real time multi-category discrimination. Figure 1 presents the architecture of BAGI.

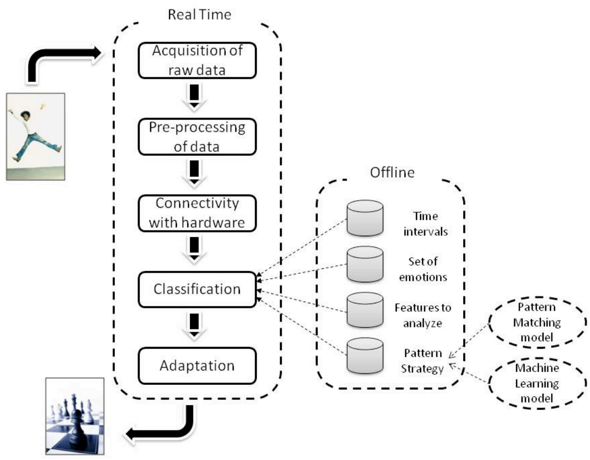


Figure 1.- Architecture of BAGI

The following subsections briefly present the modules involved in BAGI, with a more detailed description of the classification module.

3.1.- Acquisition of physiological data

The equipment selected for the acquisition of the physiological data was Procomp Infiniti by Thought Technology[11]. It offers important characteristics such as: user-friendly sensors, real time feedback, and connectivity with other applications. The Procomp Infiniti is an eight channel, multi-modal device for real time computerized biofeedback and data acquisition. It has 8 sensor inputs with two channels sampled at 2048 samples/second (s/s) and six channels sampled at 256 s/s.

3.2.- Pre-processing

The pre-processing module focused in two main tasks: i) the application of noise reduction algorithms to clean the signals, and ii) the construction of vectors with the values of the different physiological measures.

Although there are several techniques that could be applied to clean the signals and facilitate the analysis, simple and fast algorithms that could contribute to the effectiveness of the system without compromising its real-time capabilities had to be considered. For this reason, the noise in the signals was reduced using a moving average filter. A moving average filter smoothes the data by replacing the data points of a signal by averaging the value of a selected number of d adjacent points, Equation (1) illustrates a 3-point smooth:

$$S_j = \frac{Y_{j-1} + Y_j + Y_{j+1}}{d} \quad (1)$$

where S_j is the j[th] point in the smoothed signal, Y_j the j[th] point in the original signal, and d is a positive integer called *smooth width* or *averaging factor*. The higher the number of d, the less variability will be observable in the resulting signal.

This filter was computed directly on the signals by using the smoothing algorithm function provided by Biograph Infiniti software.

The construction of vectors was carried out to facilitate the analysis and classification process. The first step was the extraction of 5 statistical features from the signals: mean, standard deviation, coefficient of variation, maximum and minimum.

The next step involved the normalization of the signals, a common practice in the emotion physiology literature, as a consequence of the physiological differences among individuals. The relation of change (ρ) was obtained by comparing the extracted features of each signal during the stimuli (fstimuli) with the extracted features of each signal during baseline (fbaseline):

$$\rho = \frac{f_{\text{stimuli}}}{f_{\text{baseline}}} \quad (2)$$

[11] http://www.thoughttechnology.com/proinf.htm

In this way, the normalization facilitates the identification of physiological changes from the baseline to one particular emotional state: If $\rho > 1$ it means an increase in the signal, if $\rho < 1$ it means a decrease in the signal and if $\rho = 1$ it means no change in the signal.

Finally, each vector is composed of one emotional label and the normalized values corresponding to the signals considered for the analysis.

3.3.- Connectivity with hardware

The connection to the Procomp equipment was actually done through the Biograph Infiniti with the aid of a Software Development Kit provided by Thought Technology. There were three components involved in the implementation of the connectivity with Biograph Infiniti: C++ classes, XML file and Connection Instrument.

3.4.- Classification

First, at the initial setup, the classification module receives the intervals of time, set of emotions, features to analyze and the pattern strategy for classification. For the pattern strategy this module can receive models from two types of sources: i) existing rigid pattern matching models from literature and ii) models obtained through a machine learning process. The system allows the user to browse through different folders and select the data that BAGI will use for the classification.

Once the system started, the classification module receives in real-time the relation of change obtained by the pre-processing module. Finally with the information provided at the initial configuration the module classifies the physiological patterns and gives as an output one emotion.

BAGI allows the incorporation of new sets of emotions, intervals of times and pattern matching models to the system. The only constraint is that this information must be in a specific but simple format. A simple text editor can be used for this purpose. A brief description of the implementation of the flexibility features of the system in the classification module is provided below.

3.4.1.- Different set of emotions

In BAGI the user specifies the set of emotions to classify; this set can vary in the number and labels of affective states. In order to do this, a simple text file with the emotions can be created, and saved into the system. Figure 2 presents two sets of affective states that can be used for classification.

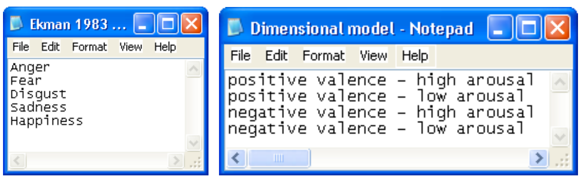


Figure 2 .- Two sets of emotions for the BAGI system

3.4.2.- Flexible time window

Following the same strategy, BAGI enables the user to specify the time window for the analysis. The user can create a file in simple text editor with the time intervals, indicating the starting and ending time (in milliseconds) for each section to analyze. The system reads the file and let a timer know about the intervals, the first two values in the file represent the start and ending time for the baseline and the following pairs of values represent the start and ending time of the sections to analyze.

The calculated values of the first interval are stored as the baseline for a user and at the end of the following intervals the system processes the physiological signals and gives as an output an emotion. Figure 3 presents an example of a file containing the time intervals of a session.

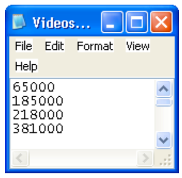


Figure 3.- Time intervals for the analysis of signals

3.4.3.- Pattern Matching models

Different pattern matching models can be used with BAGI. They just need to follow a specific format as for example the one illustrated in Table 4.

	CH1		CH2	
ID	HR		SC	
Anger	1.2	2	0.9	2
Fear	1.2	2	1.2	5
Disgust	0.5	1	1.2	2
Sadness	0.5	1	0.9	1.1
Joy	0.9	1.1	1	1.3

Table 4.- Example of pattern matching model

The values presented in table 4 correspond to normalized values obtained by comparing a feature of a signal during the stimuli with the same feature during the baseline, as explained previously in equation (2).

The two values in each channel (CH1 and CH2) in figure 4 represent the ranges in the physiological signals that will characterize a particular emotion. For instance, the pattern of fear is specified in this model as 20% to 100% increase in the HR signal, and a 20% to 500% increase in the SC signal.

The pattern matching files can be easily created in a program like for example Microsoft Excel and saved with a .csv extension.

3.4.4.- Multi-category discrimination in real-time

The key to achieve effective and real-time emotion recognition is the ability of the system to incorporate information obtained from an offline process. In this way, it is possible to use machine learning models for multi-category discrimination and use them for real-time classification.

Different types of neural networks were trained, wrapped into *.dll* and *.h* files and integrated in the system. In this way, BAGI calls trained neural networks and performs real-time classification.

Following this procedure new pattern recognition models can also be included in the system.

3.5.- Game adaptation

The adaptation module receives as an input the emotion provided by the classification module and matches the emotion with a predefined action in the game. After that it establishes a connection and, if required, instructs the game to perform any adjustment.

4.- EXPERIMENTAL VALIDATION OF BAGI

The validation of BAGI consisted in three stages. First the multi-category discrimination in real-time was tested. For this purposes an experiment was carried out to collect physiological and emotional data, to train our machine learning model and classify physiological signals in real-time.

In order to test the process of game adaptation, a computer application that customizes the computer desktop wallpaper according to the emotion of the user was developed.

Finally, in order to validate the flexibility of BAGI a series of black box tests were carried out to evaluate different: set of emotions, times to analyze the signals, physiological measures and pattern strategies.

The three stages of the validation are described in the following subsections.

4.1.- Real-Time Multi-category Discrimination

As a first step on the validation of the multi-category discrimination in real-time of BAGI, a series of experiments to gather physiological and emotional data were conducted. After that, machine learning models were trained and implemented to achieve real-time emotion classification.

The target emotions to be recognized by the system in this experiment included: relaxation, joy, sadness and fear. They were selected because each of the emotions are localized in one different quadrant in the circumplex model of emotions [26]. Figure 4 illustrates the target emotions in the valence and arousal dimensions.

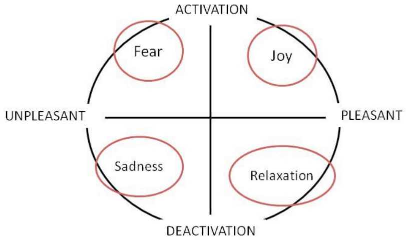


Figure 4.- Target emotions in valence and arousal dimensions

The physiological measures considered for this study included cardiovascular, electrodermal and respiration measures. From these measures a total of 17 signals were extracted using the Biograph Infiniti software.

To create a database of physiological responses to emotional stimuli, highly evocative video clips from movies were used.

Video clips have been widely used as stimuli in Psychology studies (e.g. [27, 28]) and in HCI (e.g. [13]). They have shown effectiveness to elicit emotional responses and offer good control for the experiment. Films also, have the desirable properties of being already categorized and being dynamic rather than static. There can be found several sources of information about movies and plenty websites where thousands of users vote and rate films and video clips from different genres (e.g. The Internet Movie Database, Yahoo movies, Youtube).

In order to select the most evocative video clips for the participants, the following procedure was undertaken. First, three types of sources that could provide titles of highly evocative video clips for our four target emotions were considered: i) previous studies; ii) websites with massive rates of video clips; and iii) recommendations from our research group and one media expert. Some criteria used for the selection include: self explanatory clips, highly evocative, short dialogues, length of clips. At the end, a corpus of 30 clips that fulfilled these conditions was gathered. Finally a pilot study was carried out with a group of 17 participants to select the best 12 video clips (3 clips for each target emotion).

4.1.1.- Experiment setup

At the beginning of the session the participant was briefed about the purpose of the experiment and signed a consent form. The electrodermal and cardiovascular sensors were placed in the fingers and the respiration sensor around the chest of the participant. During the experiment the participant was sat down in front of a desk with a questionnaire. Two minutes of baseline were recorded from the participant while he was watching a relaxing video, then 12 more video clips were projected in a 10ft by 8ft screen with the lights off. It was decided to use this big screen, instead of a common PC monitor, in order to get greater emotional responses. Studies recompiled in [29] support that more emotional responses are obtained when larger screens are used to present the stimuli.

The videos had an average of 3 minutes length each. After each video presentation the participants had 30 seconds to report his emotional experience in the questionnaire. The session lasted for an hour in total and there was just one participant per session. The session was video recorded to corroborate the emotions reported with the facial expressions of the participants. A total of 61 participants volunteered for this experiment.

4.1.2.- Feature selection

There were different motivations to carry out a feature selection process in our study. One of the most important was due to a restriction from the Biograph Infiniti software when connecting with third-party applications in real-time. In other words, the software that comes with the Procomp equipment, can create up to 8 connections through which only 8 signals can be passed from the Biograph Inifiniti to the emotion recognition system. Having extracted a total of 17 signals for analysis, it was necessary to select 8 signals that discriminate better the target emotions. Additionally, the feature selection process can improve the performance of the leaning model and recognition system by:

- Enhancing the generalization capabilities.
- Speeding up the learning and classification process.
- Improving the model interpretability.

The feature selection process was carried out offline, using a stepwise algorithm to select the best 8 features. Table 5 presents the 8 most important variables obtained from our analysis.

No.	Variable	Importance
1	Raw signal from BVP sensor	100.00
2	Amplitude from BVP	97.900
3	Respiration Period	81.486
4	Heart Rate	67.175
5	Raw signal from respiration sensor	34.112
6	Skin Resistance	23.702
7	Low Frequency % power	18.646
8	Ratio LF/HF	18.281

Table 5.- Variables obtained from feature selection analysis.

4.1.3.- Training results

With the selected variables, machine learning models were designed to automatically recognize the physiological patterns from the emotions experienced by participants. A supervised learning approach was followed to deduce a function from our training data and predict a class label. Two types of models were trained to recognize the patterns of the target emotions: i) Neural Network (NN) trained with a backpropagation algorithm and ii) Probabilistic Neural Network (PNN).

From the 61 participants some interruptions in the recording of the physiological data were observed in 2 participants. Although the information could be recovered and reconstructed, it was decided to discard it from the analysis to reduce any external bias.

A total of 708 training vectors were constructed from the 59 participants and 12 video clips considered for the analysis. The video clips were grouped in 4 categories, according to the target emotions, and there were 177 vectors for each category. Each vector contained the emotion reported by the participant and the normalized values for the 8 physiological signals presented in Table 6.

From the training process we achieved a total classification accuracy of 62.42% with the backpropagation NN and 84.46% with PNN. Tables 6 and 7 present the confusion matrices of the NN and the PNN.

Actual category	Predicted category			
	Relaxation	Joy	Sadness	Fear
Relaxation	**100 (56.49%)**	28	35	14
Joy	15	**137 (77.4%)**	17	8
Sadness	28	18	**93 (52.54%)**	38
Fear	13	13	39	**112 (63.27%)**

Table 6.- Confusion matrix for the backpropagation NN

Actual category	Predicted category			
	Relaxation	Joy	Sadness	Fear
Relaxation	**151 (85.31%)**	15	8	3
Joy	14	**156 (88.13%)**	3	4
Sadness	14	6	**139 (78.53%)**	18
Fear	5	3	17	**152 (85.87%)**

Table 7.- Confusion matrix for the PNN

In order to assess the generalization capabilities of our predictive models on new independent data sets, a leave one out cross validation was used. From the evaluation of the trained models a total classification accuracy of 54.94% on the NN trained with backpropagation and a 78.38% on the PNN was obtained. The confusion matrices of the validation of these two models are presented in tables 8 and 9.

Actual category	Predicted category			
	Relaxation	Joy	Sadness	Fear
Relaxation	**102 (57.62%)**	26	33	16
Joy	27	**123 (69.49%)**	16	11
Sadness	36	21	**65 (36.72%)**	55
Fear	17	15	46	**99 (55.93%)**

Table 8.- Confusion matrix for the validation of the backpropagation NN

Actual category	Predicted category			
	Relaxation	Joy	Sadness	Fear
Relaxation	**143 (80.79%)**	18	13	3
Joy	22	**145 (81.92%)**	5	5
Sadness	18	8	**127 (71.75%)**	24
Fear	8	8	21	**140 (79.09%)**

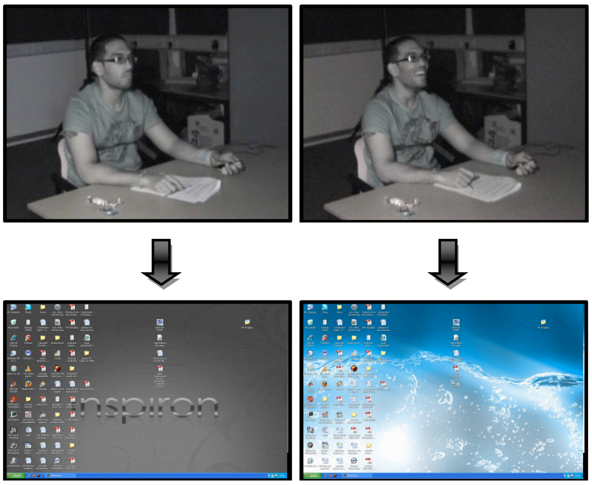
Table 9.- Confusion matrix for the validation of the PNN

4.2. Real-Time bio-affective feedback

For testing the real-time recognition and adaptation capabilities a very simple application that adapts the desktop wallpaper according to the emotion recognized by the system was built. The replay capability of Procomp was used to simulate the real-time acquisition of physiological signals from 59 participants of our experiment. Then, the BAGI was connected to the application that customizes the desktop wallpaper. Two types of models for real time classification were tested: one pattern matching model based on the findings from Ekman et al. [17] and one machine learning model trained with the PNN. Both models were able to successfully classify physiological signals in real-time and adapt the computer wall paper to the emotion recognized. Table 10 presents a description of the wallpapers assigned to each emotion and an example of BAGI performing real-time adaptation is presented figure 5.

Emotion recognized	Wallpaper colour
Relaxation	White wallpaper
Joy	Blue wallpaper
Sadness	Gray wallpaper
Fear	Yellow wallpaper

Table 10.- Customization of the wallpaper



a) b)

Figure 5.- a) participant at beginning of experiment; b) BAGI customizing wallpaper according to emotion

Figure 5 a) shows the default computer wallpaper when the participant is starting the experiment; figure 5 b) illustrates how BAGI customize the computer wallpaper as soon as it recognizes joy.

4.3.- Flexible Architecture

The flexibility of BAGI has been tested using the black box testing method. The detailed tests carried out on BAGI are presented in table 11.

Test	Details
Set of emotions	System tested with 3 different sets of emotions with different number of emotions and labels.
Time intervals	System tested using 3 different time intervals for the analysis of the signals: 10, 120 and 180 seconds.
Physiological measures	System tested using 3, 5 and 8 different physiological measures.
Pattern matching models	System tested with 3 different pattern matching models.
Machine learning models	System tested with 1 NN trained with backpropagation algorithm and 1 PNN.

Table 11.- Flexibility black box tests on BAGI

The results obtained from the black box tests demonstrated the flexibility of BAGI to work with different affective sets, time intervals, measures and psycho-physiological models.

5.- CONCLUSION AND FUTURE WORK

At the present time, when the game industry is looking for more immersive interactive techniques and ways to improve the artificial intelligence, affective recognition systems can play an important part as an instrument to enhance the game experience.

Physiological signals have been already used to recognize emotions in games; nevertheless the pattern matching approach commonly followed is not practical for the recognition among several emotions.

The feasibility of the use of physiological signals with machine learning techniques has already been proved by several offline HCI studies. However the implementation of such systems in a real-time application is not a straightforward task.

In this study a user independent real-time emotion recognition system that could be used for gaming was presented. The architecture of BAGI facilitates the reusability of previous findings and the integration of new models.

An experiment to gather emotional and physiological data was carried out with 61 participants. The results from the questionnaires showed that the video clips used as a stimuli effectively evoked the four target emotions in the most of the participants.

From the three sensors used to record the physiological signals a total of 17 signals were extracted. And due to the real-time connectivity limitations of the equipment with third-party computer applications the best 8 signals that discriminate our data were selected.

Two machine learning models were trained to recognize the physiological patterns of 4 emotions, having the best results with a PNN with classification accuracy of 84.46% on the training data and 78.38% when it was cross-validated. The two machine learning models were implemented on the system for a real time classification. Finally, a simple application that customizes the desktop wallpaper to the emotional state of the user was developed. The results from our study provide evidence of the feasibility of the use of BAGI as an interface for gaming.

Although the training models can be claimed to apply just to classify emotions experienced while watching video clips, the methodology used in this study can be used to train models for different types of stimuli. And the architecture of BAGI allows the inclusion of these new trained models to perform a real-time classification.

An interesting future study would be the utilization of a model trained with stimuli to classify emotional responses evoked by other type of stimuli (e.g. use a model trained with video clips to classify emotions evoked by pictures). There are already some differences in the physiological patterns, for the same emotions using different stimuli, in the psychology literature; however there is no empirical evidence in the literature of machine learning models attempting to classify emotions evoked by different types of stimuli.

Our future work will involve the training of a machine learning model in a game context and its inclusion to BAGI. Hopefully, with the affective feedback provided by BAGI it will be possible to reinforce the AI of the game and enhance the emotional experience.

ACKNOWLEDGEMENTS

J.A.P. would like to thank CONACYT and SEP for their financial support.

REFERENCES

[1] Z. O. Toups, R. Graeber, A. Kerne, L. Tassinary, S. Berry, K. Overby, and M. Johnson, "A Design for Using Physiological Signals to Affect Team Game Play," in *Augmented Cognition International*, 2006.

[2] H. Prendinger, C. Becker, and M. Ishizuka, "A study in Users' Physiological Response to an Empathic Interface Agent," *International Journal of Humanoid Robotics,* vol. 3, pp. 371-391, 2006.

[3] K. M. Gilleade and A. Dix, "Affective Videogames and modes of Affective Gaming: Assist Me, Challenge Me, Emote Me," in *Digital Games Research Association (DiGRA)*, 2005.

[4] C. Conati, "Probabilistic Assessment of User's Emotion in Educational Games," *Applied Artificial Intelligence,* vol. 16, pp. 555-575, 2002.

[5] E. C. Lalor, S. P. Kelly, C. Finucane, R. Burke, R. Smith, R. B. Reilly, and G. McDarby, "Steady-State VEP-Based Brain-Computer Interface Control in an Immersive 3D Gaming Environment," *EURASIP Journal on Applied Signal Processing,* vol. 19, pp. 3156-3164, 2005.

[6] S. I. Hjelm and C. Browall, "Brainball - using brain activity for cool competition," in *1st Nordic Conference on Human-Computer Interaction*, 2000.

[7] M. Ko, K. Bae, G. Oh, and T. Ryu, "A Study on New Gameplay Based on Brain-Computer Interface," in *Digital Games Research Association (DiGRA)*, 2009.

[8] J. Arroyo-Palacios and D. Romano, "Exploring the Use of a Respiratory-Computer Interface for Game Interaction," in *IEEE Consumer electronics Society Games innovation Conference*, London, 2009, pp. 154-159.

[9] D. Bersak, G. McDarby, N. Augenblick, P. McDarby, M. Daragh, M. Brian, and R. Karkun, "Intelligent Biofeedback using an Immersive Competitive Environment," in *Designing Ubiquitous Computing Games Workshop*, Atlanta GA, 2001.

[10] C. Peter and A. Herbon, "Emotion representation and physiology assignments in digital systems," *Interacting with Computers,* vol. 18, pp. 139-170, 2006.

[11] J. Kim and E. Andre, "Emotion Recognition Based on Physiological Changes in Music Listening," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30, pp. 2067-2083, 2008.

[12] R. W. Picard, E. Vyzas, and J. Healey, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, pp. 1175-1191, 2001.

[13] F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals using wireless sensors for presence technologies," *International Journal of Cognition, Technology and Work,* vol. 6, pp. 4-14, 2004.

[14] J. T. Caciopo and L. G. Tassinary, "Inferring Psychological Significance from Physiological Signals," *American Psychologist,* vol. 45, pp. 16-28, 1990.

[15] D. Stemmler, "The autonomic differentiation of emotions revisited: Convergent and discriminant validation," *Psychophysiology,* pp. 617-632, 1989.

[16] I. C. Christie, "Multivariate Discrimination of Emotion-Specific Autonomic Nervous System Activity." vol. MSc Thesis: Virginia Polytechnic Institute and State University, 2002.

[17] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science,* vol. 221, pp. 1208-1210, 1983.

[18] B. L. Fredrickson, R. A. Mancuso, C. Branigan, and M. M. Tugade, "The Undoing Effect of Positive Emotions," *Motivation and Emotion,* vol. 24, pp. 237-258, 2000.

[19] J. Arroyo-Palacios and D. Romano, "Towards a Standardization in the Use of Physiological Signals for Affective Recognition Systems " in *Measuring Behavior 2008*, Maastricht, The Netherlands, 2008, pp. 121-124.

[20] R. B. Zajonc and D. N. McIntosh, "Some Promising Questions and Some Questionable Promises," *Psychological Science,* vol. 3, pp. 70-74, 1992.

[21] J. T. Cacioppo, D. J. Klein, G. G. Berntson, and E. Hatfield, "The Psychophysiology of Emotion," in *Handbook of Emotions*, M. Lewis and J. M. Haviland, Eds., 1993, pp. 119-142.

[22] T. Kim, D. Shin, and D. Shin, "Towards an Emotion Recognition System based on Biometrics," in *International Joint Conference on computational Sciences and Optimization*, 2009, pp. 656-659.

[23] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System," in *Lecture Notes in Computer Science. Affective Dialog Systems.* vol. 3068/2004: Springer Berlin / Heidelberg, 2004.

[24] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Medical & Biological Engineering & Computing,* vol. 42, pp. 419-427, 2004.

[25] E. Leon, G. Clarke, V. Callaghan, and F. Sepulveda, "A user-independent real-time emotion recognition system for software agents in domestic environments," *Engineering Applications of Artificial Intelligence,* vol. 20, pp. 337-345, 2007.

[26] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology,* vol. 39, pp. 1161-1178, 1980.

[27] J. R. Averill, "Autonomic response patterns during sadness and mirth," *Psychophysiology,* vol. 5, pp. 399-414, 1969.

[28] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition and Emotion,* vol. 9, pp. 87-108, 1995.

[29] M. Lombard and T. Ditton, "At the Heart of It All: The Concept of Presence.," *Journal of Computer Mediated Communication,* vol. 3, 1997.