Towards Affective Camera Control in Games

Georgios N. Yannakakis · **Héctor P. Martínez** ·
Arnav Jhala

Abstract Information about interactive virtual environments, such as games, is perceived by users through a virtual camera. While most interactive applications let users control the camera, in complex navigation tasks within 3D environments users often get frustrated with the interaction. In this paper, we propose inclusion of camera control as a vital component of affective adaptive interaction in games. We investigate the impact of camera viewpoints on psychophysiology of players through preference surveys collected from a test game. Data is collected from players of a 3D prey/predator game in which player experience is directly linked to camera settings. Computational models of discrete affective states of *fun*, *challenge*, *boredom*, *frustration*, *excitement*, *anxiety* and *relaxation* are built on biosignal (heart rate, blood volume pulse and skin conductance) features to predict the pairwise self-reported emotional preferences of the players. For this purpose, automatic feature selection and neuro-evolutionary preference learning are combined providing highly accurate affective models. The performance of the artificial neural network models on unseen data reveals accuracies of above 80% for the majority of discrete affective states examined. The generality of the obtained models is tested in different test-bed game environments and the use of the generated models for creating adaptive affect-driven camera control in games is discussed.

Keywords Camera control · player experience modeling · skin conductance · blood volume pulse · neuro-evolution · preference learning

1 Introduction

There is a growing demand for adaptive interactive experiences in virtual environments, such as games, that provide highly customized affective experience to individual users. Affective modeling research in interactive entertainment environments to date has focused on

Georgios N. Yannakakis · Héctor P. Martínez
Center for Computer Games Research
IT University of Copenhagen
E-mail: {yannakakis, hpma}@itu.dk

Arnav Jhala
University California, Santa Cruz
E-mail: jhala@cs.ucsc.edu

various properties of the environment like narrative engagement (Aylett et al., 2006), inter-action with characters (Paiva et al., 2002; Conati et al., 2003; Aylett and Paiva, 2006), level features and game mechanics (Pedersen et al., 2009), and physical interaction (Yannakakis and Hallam, 2008). While these characteristics of virtual environments are important for tailoring a user's experience in the game environment, a vital component of communication between the system and the user is a virtual camera. The camera could be one of the primary components of a closed-loop control system using affect as the feedback signal (i.e. affective loop), since it determines the way users perceive the world and receive visual feedback from their actions.

The camera viewpoint in games and learning environments is carefully chosen by designers in order to maximize communication of information about the state of the world to the user. However, this goal of designers often conflicts with user's agency, where a user controls her own viewpoint. Thereby, it is important to have a dynamic camera system that adapts the viewpoint to improve user's experience during different states of the game. The very first step towards achieving affect-driven camera control in games is to efficiently identify and synthesize the affective state of the user and discover the relationships between camera control and affect.

This paper investigates the effect of camera viewpoints to the psychophysiological state of players and evaluates the efficiency of preference learning and automatic feature selection (Yannakakis et al., 2009; Yannakakis, 2009) for the design of affective preference models of players. Such models are trained to learn the association between virtual camera control parameters, physiological signals of players, and self-reported affective states. Our attempts for constructing computational models of affect presented in this paper rely on three main hypotheses:

- there is an unknown function between biosignal features, camera control parameters and subject reported emotional preferences that a machine learning algorithm is able to approximate;
- the above-mentioned function is approximated better by non-linear (rather than linear) models; indicating its non-linear characteristics; and
- biosignal features which are significantly correlated with emotional preferences are not likely to be picked by automatic feature selection in the process of constructing non-linear models of affect.

Towards validating those three hypotheses, physiological signal data and expressed emotional responses are collected through a game survey experiment. The discrete affective states investigated are *fun*, *challenge*, *boredom*, *frustration*, *excitement*, *anxiety* and *relaxation* and are expressed as pairwise preferences (namely, *emotional preferences*) using forced-choice questionnaires. The physiological signals recorded are *heart rate*, *blood volume pulse* and *skin conductance*. The test-bed game used is a simple three-dimensional prey/predator game in which camera viewpoint affects gameplay performance.

The notion of affective camera control as a necessary component towards affective games (Hudlicka, 2009) was introduced by Martinez et al. (2009) in which an initial investigation of the linear impact of different camera viewpoints on the physiology and expressed emotions of users is presented. The study presented here builds upon and extends the work by Martinez et al. (2009) in several dimensions. First, non-linear computational models of affect are constructed to approximate the function between biosignal data and reported emotional preferences using neuro-evolutionary preference learning (Yannakakis et al., 2009; Yannakakis, 2009) — i.e artificial neural networks trained via artificial evolution on preference data. Linear relationships, reported by Martinez et al. (2009), between

statistical features (e.g. mean and standard deviation) extracted from physiological signals — namely, *physiological features* — and self-reported affective states are solely used as an initial analysis of the data. Second, three different mechanisms of automatic feature selection (n best, sequential forward and perceptron feature selection) are applied to select minimal physiological feature sets as artificial neural network (ANN) inputs that maximize the affective models' performance. Feature selection proves to be essential for both selecting minimal feature sets and distinguishing the appropriate features that contribute to the prediction of each affective state. The ANN models obtained are accurate predictors of the emotional preferences of the players; 3-fold cross validation accuracy lies between 71.71% for challenge up to 85.55% for frustration. Third, three controllable features of the camera, namely *distance*, *height* and *frame coherence*, are embedded to the ANN models as an intermediate step towards affective camera control. Camera controllable features — when embedded to the affective model — constitute components of the non-linear interplay between player characteristics (via physiology) and reported emotions. Thus, using the obtained accurate models of affect one could design real-time adaptive mechanisms for adjusting camera controls to accommodate the desired affective game experience.

All three hypotheses stated earlier are validated by the results obtained: the non-linear models built are accurate predictors of affect relying on biosignal information and camera control parameters; the simple non-linear models built are better estimators of affect than the linear models of affect for most affective states; and only a small number of significantly correlated biosignal features are picked as inputs for the non-linear (ANN) models built.

The paper is organized as follows: Section 2 focuses on the test-bed game considered in the experiments presented in this paper. Section 3 and Section 4 are, respectively, concerned with the experimental protocol followed and the biosignal features extracted after data collection. Neuro-evolutionary preference learning and automatic feature selection as a method for constructing affective models is presented in Section 5. Section 6 and Section 7 analyze the main results and investigate the generality of the obtained models, respectively. The paper concludes with an overview of the state-of-the-art on affective modeling and camera control (Section 8); a discussion on open research questions as well as the generality and limitations of the approach (Section 9); and a summary of main results and conclusions (Section 10).

2 The Maze-Ball Game

2.1 The Game

Maze-Ball is a three-dimensional prey/predator *PacMan*-like game (see Figure 1). The player (prey) controls a ball which moves inside a maze hunted by 10 red-colored opponents (predators), namely *zombies*, moving around the maze. The goal of the player is to maximize her score by gathering as many gold tokens, scattered in the maze, as possible while avoiding being touched by the zombies in a predefined time window of 90 seconds. The 90 second play-time window is designer-driven and attempts to maintain a good balance between sufficient gameplay interaction and the player's cognitive load. We need to minimize memory-dependent effects of post-experience on questionnaire items and the total time required for the experience to run; on the contrary the game should provide sufficient interaction for the requested affective states to be elicited.

The purpose of using Maze-Ball for our experiments is two-fold: first, it consists of a minimal interface for an enjoyable game (arrow keys for controlling the character) and a
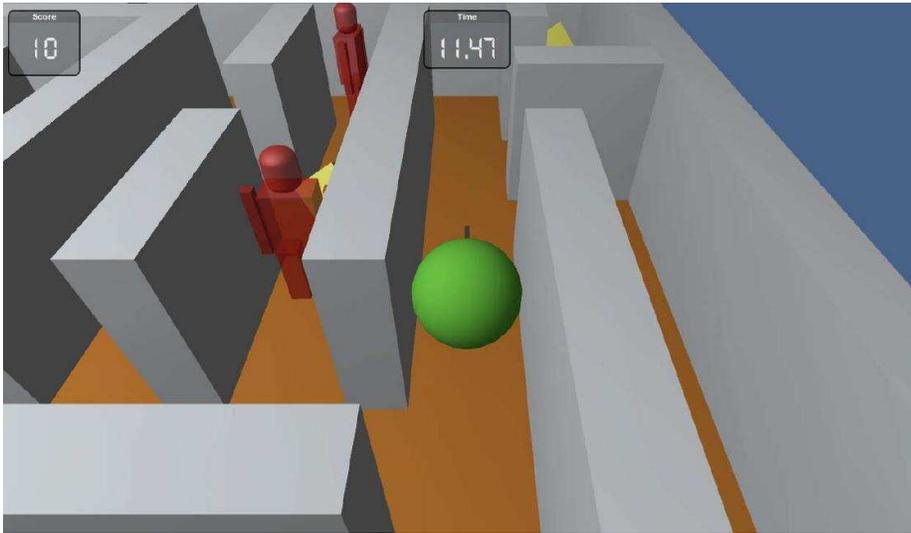
Fig. 1 Screen-shot of Maze-Ball.

simple visual environment. Single-hand game control via the keyboard allows for the unobtrusive placement of biofeedback sensors to the free hand which is essential for our purpose. Second, there is a direct effect of the amount of information available to the player about the world — via camera viewpoint — on her movement strategy. For instance, in the top-down view of the full maze the player has complete global information about the world for planning out the path along the maze. This viewpoint, however, may not be optimal for controlling the character's local movement as the character takes up only a small fraction of the entire screen. A close view, such as the first person view, makes moving the character and avoiding enemies easier but strategically moving along the maze harder.

2.2 The Camera Controller

A dynamic camera controller maintains the position, orientation, and the field-of-view (FOV) of the camera in a graphical world. Maintaining the camera position and orientation amounts to finding and maintaining the 3D coordinates (x, y, z) for the location and rotation angles for orientation $(pitch, yaw, roll)$ to satisfy viewing constraints imposed on the camera by the game design and the environment. Constraint-based techniques have been extensively used in virtual camera control systems (Christie et al., 2005b).

The camera system implemented for our experiments is based on a weighted Constraint Satisfaction Problem (CSP) solver framework for satisfying view constraints at each frame (Bourne et al., 2008). The CSP representation used in the camera framework of Maze-Ball contains fewer variables in order to make it interactive and efficient. The camera variables used are: *distance*, (D), *height*, (H), and *frame coherence*, (F_c). The distance variable's values are constrained to maintain a relative distance relationship with the target. Height variable's values are constrained to maintain a fixed height relative to the target. Frame Coherence variable's values are constrained to maintain smooth motion across frames and avoid erratic camera movements. We define a *camera profile* in the context of our game as a
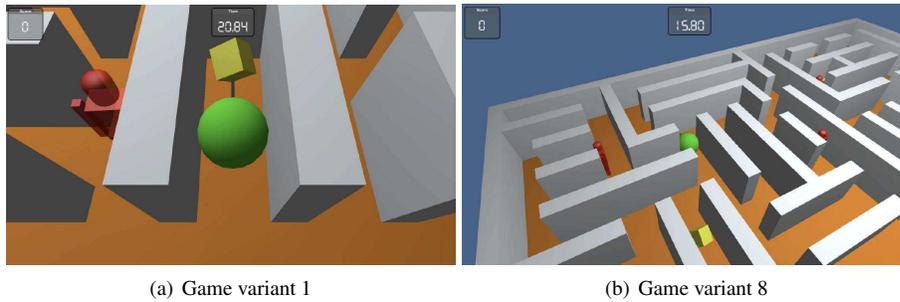
(a) Game variant 1   (b) Game variant 8

Fig. 2 Two example game variants of Maze-Ball: (a) *Low* height, *Low* distance and *Low* frame coherence (b) *High* height, *High* distance and *High* frame coherence.

triplet of distance, height, and frame coherence values. The constraint values are expressed in the range $[0, 1]$.

In this work, we are interested in evaluating the effect of camera viewpoints resulting from different camera profiles to the psychophysiological state of the players. For the same physical setting of the game, we vary camera profiles while keeping the game design, level design, and game mechanics unaltered. Eight game variants are implemented in Maze-Ball by varying distance, height and frame coherence. For each of the three camera control variables, two states ('*High*' and '*Low*') are selected[1]. Two example game variants are illustrated in Figure 2.

3 Experimental Methodology

Physiological signals and subject's emotional preferences were acquired for Maze-Ball via the following game survey experiment. Thirty six subjects (males: 80%) aged from 21 to 47 years (mean and standard deviation of age equal 27.2 and 5.84, respectively) participated in the experiment. Each subject played a predefined set of eight games for 90 seconds each (see Figure 3); the games differ in the levels of *distance*, *height* and *frame coherence*. For each completed pair of games A and B, subjects report their emotional preference using a 4-alternative forced choice (4-AFC) protocol:

– game A [B] was/felt more E than game B [A] game (*cf.* 2-alternative forced choice);
– both games were/felt equally E or
– neither of the two games was/felt E.

Where E is the affective state under investigation and contains *fun*, *challenging*, *boring*, *frustrating*, *exciting*, *anxious* and *relaxing*. The selection of these seven states is based on their relevance to computer game playing with parameterized camera positioning. The first five have been previously used in game-related user studies (Mandryk and Atkins, 2007) while the last two are included for maintaining a uniform covering of the arousal-valence appraisal space.

The number of experiment participants is determined by $C_2^9 = 36$, this being the required number of all combinations of 2 out of 9 game variants. The experimental protocol

---

[1] The *Low* and *High* values selected for distance, height and frame coherence are respectively 2.5 and 6; 6 and 15; and 0.01 and 0.35
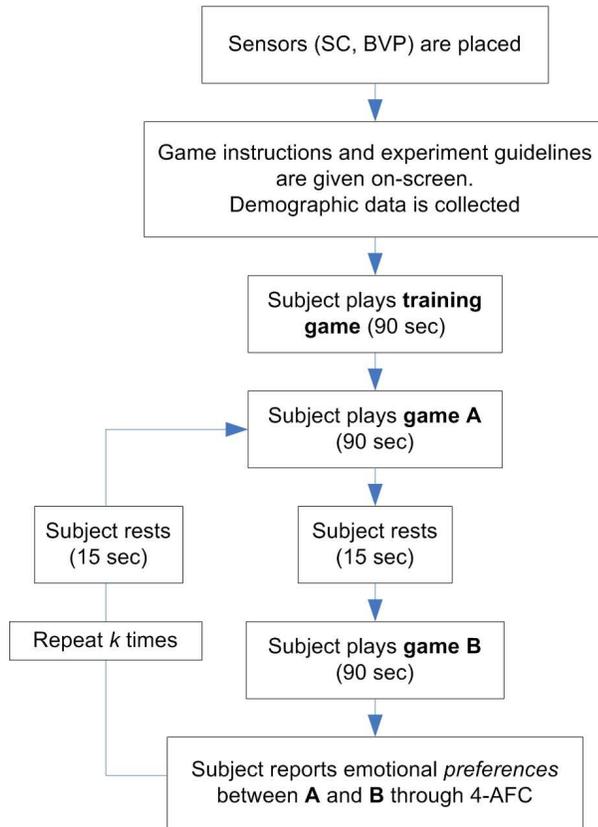
Fig. 3 The experimental protocol followed.

is designed in such a way that 4 preference instances should be obtained for each pair of the 9 game variants played in both orders (2 preference instances per playing order). Given that, a number of 8 games ($k = 4$ — see Figure 3) is required to be played by each subject resulting to $36 \cdot 4 = 144$ game pair preferences. Note that the affective modeling analysis presented in this paper is based on the 8 game variants described in Section 2.2 and in their corresponding 112 preference instances — this number excludes the 32 game pairs played in total containing the ninth variant: $144 - 32 = 112$. The ninth game variant defines a control experiment utilized to test for generality of the obtained models. That game variant controls for the impact of walls in the game by introducing a Maze-Ball game variant with much higher visibility; data obtained via the ninth game variant are analyzed in Section 7.

Blood Volume Pulse (BVP), b, and Skin Conductance (SC), s, were collected in real-time at 32 Hz. Heart rate (HR), h, is computed every 5 seconds by extrapolating the inter-beat time intervals detected in the BVP signal (see Figure 5). A number of statistical features of those signals is extracted and presented in Section 4 below. These three signals are selected because: (1) affective states have been reported to have a direct influence to those signals and (2) they are correlated to sympathetic arousal (see (Picard et al., 2001) among others). Measurement units for HR and SC are, respectively, heart beats per minute (bpm) and micro-Siemens (μS) whereas BVP is a relative measure of blood vessel pressure. For
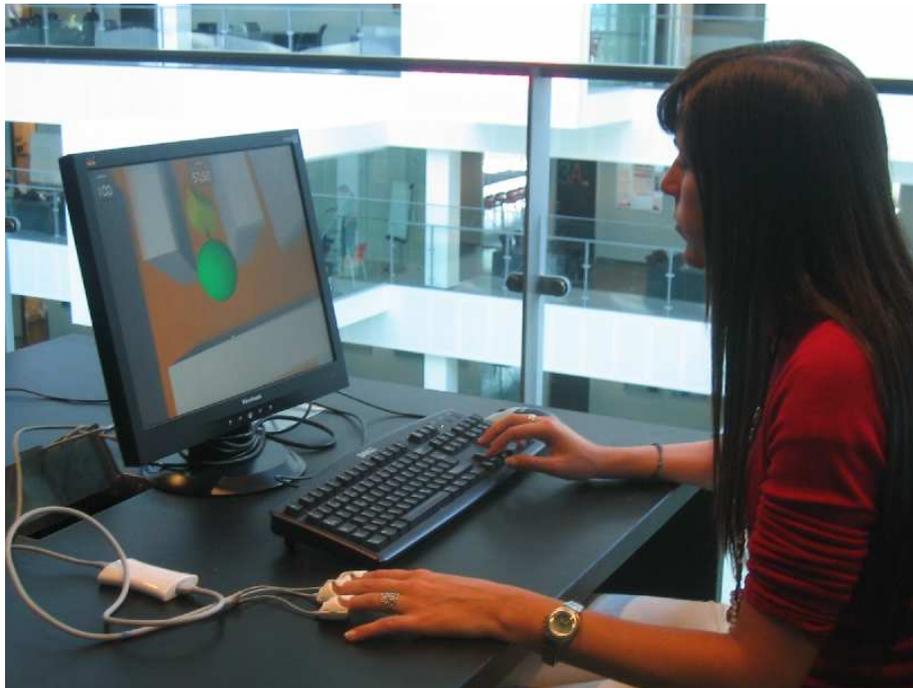
Fig. 4 The setup of the experiment. The Maze-Ball game is viewed on screen; the IOM bio-sensing hardware is placed on the subject's left hand in this picture.

measuring the above-mentioned signals the IOM biofeedback device (IOM Biofeedback Equipment) is used consisting of three sensors (two electrodes for SC and one photo sensor for BVP) placed on the subject's fingertips as seen in Figure 4. By using small and accurate commercial apparatus like the IOM biofeedback device in the least intrusive way we attempt to minimize (psychological) experiment effects caused by the presence of recording devices.

Note that subjects are not interviewed but are asked to fill in a comparison questionnaire, minimizing interviewing effects. The 4-alternative forced choice (4-AFC) protocol is used since it offers several advantages for subjective emotion capture: we believe it minimizes subjects' subjective notions of scaling and allows a fair comparison between the answers of different subjects while also making explicit the "no preference" cases concealed by 2-AFC. The 4-AFC and 2-AFC protocols have been successfully utilized to provide data for building accurate computational models of reported emotional preferences (Yannakakis et al., 2008; Yannakakis and Hallam, 2008).

4 Features Extracted

This section lists the statistical features extracted from the three physiological signals monitored. Some features are extracted for all signals while some are signal-dependent as seen in the list below. The choice of those specific statistical features is made in order to cover a decent amount of the HR, BVP and SC signal dynamics proposed in the majority of previous studies in the field (Picard et al., 2001; Goldberger et al., 2001; Yannakakis and Hallam, 2008).
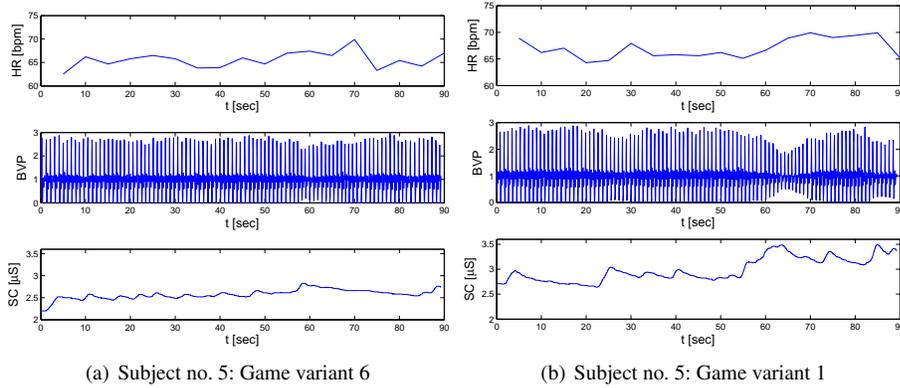
(a) Subject no. 5: Game variant 6   (b) Subject no. 5: Game variant 1

Fig. 5 Example of the HR, BVP and SC signals obtained in a pair of Maze-Ball games: a subject (no. 5) plays a game of *High* height, *Low* distance and *High* frame coherence (a) and then a game of *Low* height, *Low* distance and *Low* frame coherence (b). The subject expressed a fun, boredom and relaxation preference for the game variant 6 whereas expressed a challenge, excitement, frustration and anxiety preference for game variant 1.

- **All signals** ($\alpha \in \{h, b, s\}$)**:** Average $E\{\alpha\}$, standard deviation $\sigma\{\alpha\}$, maximum $\max\{\alpha\}$, minimum $\min\{\alpha\}$, the difference between maximum and minimum signal recording $D^{\alpha} = \max\{\alpha\} - \min\{\alpha\}$, Pearson's correlation coefficient R_{α} between raw α recordings and the time t at which data were recorded and autocorrelation (lag equals 1) of the signal ρ_1^{α}.
- **HR** (h)**:** Initial, h_{in}, and last, h_{last}, HR recording; time when maximum h occurred $t_{\max}\{h\}$, time when minimum h occurred $t_{\min}\{h\}$ and the difference $D_t^h = t_{\max}\{h\} - t_{\min}\{h\}$; the slope, L_h, of linear regression on the signal (linear least squares fitting is used); and the parameters β and γ of the quadratic regression ($h_Q(t) = \beta t^2 + \gamma t + \delta$) on the signal which respectively quantify the curvature and the rotation angle with respect to the x-axis of the quadratic curve. QR factorization is used to fit the data.
- **BVP** (b)**:** Average inter-beat amplitude $E\{IBAmp\}$, mean of the first and second differences of the raw BVP (δ_1^b and δ_2^b respectively), and normalized BVP (δ_1^{Nb} and δ_2^{Nb} respectively) according to the suggestion of Picard (Picard et al., 2001), and mean of the absolute values of the first and second differences of the BVP signal (Picard et al., 2001) ($\delta_{|1|}^b$ and $\delta_{|2|}^b$ respectively). Moreover, given the inter-beat time intervals (RR intervals) of the BVP signal the following Heart Rate Variability (HRV) parameters were computed:
  - HRV - time domain: the standard deviation of RR intervals $\sigma\{RR\}$, the fraction of RR intervals that differ by more than 50 msec from the previous RR interval $pRR50$ and the root-mean-square of successive differences of RR intervals RMS_{RR} (Goldberger et al., 2001).
  - HRV - frequency domain: the frequency band energy values derived from power spectra obtained using discrete Fourier transformation; energy values are computed as the integral of the power of each of the following four frequency bands (see (Goldberger et al., 2001) among others): Ultra Low Frequency (ULF) band: $[0.0, 0.0033]$ Hz; Very Low Frequency (VLF) band: $(0.0033, 0.04]$ Hz; Low Frequency (LF) band: $(0.04, 0.15]$ Hz and High Frequency (HF) band: $(0.15, 0.4]$ Hz.

– **SC** (s): All extracted features used for the HR signal. Additional features include the mean of the first and second differences of the raw SC (δ_1^s and δ_1^s respectively) and the normalized SC (δ_1^{Ns} and δ_2^{Ns} respectively), and the mean of the absolute values of the first and second differences of the SC signal ($\delta_{|1|}^s$ and $\delta_{|2|}^s$ respectively).

Even though there appear to be correlations between signal dynamics and emotional preferences by direct speculation of the signals (see Figure 5) a statistical analysis of the extracted features above will better determine the extent to which each statistical feature has an impact on reported emotions (see Section 6).

5 Affective Modeling Method

The method followed for designing accurate affective models utilizes machine preference learning (see Section 5.1 below) to approximate the function between the reported affective states, the physiological features described earlier in Section 4 and the three camera controllable features — distance, height and frame coherence. Feature selection, as described in Section 5.2, is employed for each training attempt to choose the appropriate input vector of the trained emotional preference model.

5.1 Neuro-evolutionary Preference Learning

As previously mentioned, we assume there is an unknown mapping between physiological features of the player, camera profile parameters and reported emotional preferences. In this paper, neuro-evolutionary preference learning is utilized to approximate that function.

Other preference learning approaches are possible for this problem. Bayesian learning and support vector machines are amongst the most popular preference learning mechanisms; however, comparative studies have already demonstrated the advantages of the use of neuro-evolution over those techniques for constructing models of cognitive and affective states (Yannakakis, 2009; Yannakakis and Hallam, 2008; Yannakakis et al., 2009). Thus, we believe that a non-linear function such as an artificial neural network (ANN) is a good choice for approximating the mapping between reported emotions and input data. Simple single-layer perceptrons (SLPs) — employing non-linear logistic activation functions — are utilized for learning the relation between controllable and physiological features (ANN inputs) — selected from feature selection schemes presented in Section 5.2 — and the value of the investigated emotional preference (ANN output). The main motivation for using single-layer perceptrons, instead of multi-layer perceptrons, is that we want to be able to analyze the trained function approximator and discuss the underlying physical meaning behind the non-linear relationships obtained; e.g. see discussion in Section 6.3. While an multi-layer perceptron (MLP) can potentially approximate the function investigated with a higher accuracy, a function represented as a single-neuron ANN is easily interpretable by humans.

The single neuron employed uses the sigmoid (logistic) activation function; connection weights take values from -5 to 5 to match the normalized input values that lie in the [0, 1] interval. Since there are no prescribed target outputs for the learning problem (i.e. no differentiable output error function), ANN training algorithms such as back-propagation are inapplicable. Learning is achieved through artificial evolution by following the preference learning approach presented in (Yannakakis, 2009; Yannakakis and Hallam, 2008; Yannakakis et al., 2009).

A generational genetic algorithm is implemented, using a fitness function that measures the difference between the subject's reported emotional preferences and the relative magnitude of the corresponding model (ANN) output. More specifically, the logistic (sigmoidal) function $g(E, \epsilon) = 1/(1 + e^{-\epsilon E})$ is used where $E = e(A) - e(B)$ is the difference of the ANN output values (investigated emotion) between game A and game B; $\epsilon = 30$ if A is preferred ($A \succ B$) and $\epsilon = 5$ when B is preferred ($A \prec B$). Both the sigmoidal shape of the objective function and its selected ϵ values are inspired by its successful application as a fitness function in neuro-evolution preference learning problems (Yannakakis et al., 2009; Yannakakis and Hallam, 2008; Yannakakis et al., 2008).

A population of 1000 individuals is used, and evolution runs for 100 generations. A probabilistic rank-based selection scheme is used, with higher ranked individuals having higher probability of being chosen as parents. Reproduction is performed by uniform crossover, followed by Gaussian mutation with a 5% probability. All SLP preference learning experiments reported in this paper consist of 20 trails for each emotional preference and each feature selection mechanism.

5.2 Feature Selection

Automatic feature selection (FS) is an essential process towards distinguishing those features that can assist in predicting the investigated affective states. It is our desire to maintain the physical meaning of the features extracted for our analysis.

Therefore data pre-processing techniques such as *principal component analysis* are not applicable. It is also desired that the affective models constructed are dependent on a minimal number of features that yield the highest prediction accuracy. The primary reasons for minimizing the feature subset are improvement of model expressiveness (interpretability); reduction of problem dimensionality; and reduction of computational effort in training and real-time performance.

FS is utilized to find the feature subset that yields that most accurate model and save computational effort of exhaustive search on all possible feature combinations. Note that none of the methods presented are guaranteed to find the optimal feature set since neither of them search for all possible combinations (they are all variants of hill-climbing). Using the extracted features described earlier the *n best individual feature selection* (nBest), the *Sequential Forward Selection* (SFS), and the *Perceptron Feature Selection* (PFS) schemes are applied and compared.

The performance of a feature set selected by automatic FS is measured through the average classification accuracy of the model in three independent runs using 3-fold cross-validation. In our case, FS algorithms select the input feature set for the SLP model. The three FS mechanisms utilized have been previously described and compared in (Pedersen et al., 2009; Yannakakis, 2009); their main components are presented here for brevity. The nBest FS mechanism ranks the features used individually in order of model performance; the chosen feature set of size n is then the first n features in this ranking. SFS is a bottom-up search procedure where one feature is added at a time to the current feature set. The feature to be added is selected from the subset of the remaining features such that the new feature set generates the maximum value of the performance function over all candidate features for addition. PFS is a single top-down perceptron pruning approach (Mejia-Lavalle and Arroyo-Figueroa, 2006) modified to match preference learning problems. The perceptron's initial input vector has the size of the number of features examined and it is trained to match emotional preferences using artificial evolution (see Section 5.1). Features whose

corresponding connection weight absolute values are below a threshold are eliminated. The reader is referred to (Yannakakis, 2009) for further details on the three FS mechanisms.

Since we are interested in the minimal feature subset that yields the highest performance we terminate selection procedure when an added feature yields equal or lower validation performance to the performance obtained without it.

6 Results

Herein we test if the order of playing the test-bed game affects reported emotions and provide an analysis on the effects between emotional preferences expressed and physiological signal features extracted. The impact of the dissimilar game variants of Maze-Ball on emotional responses is also highlighted.

6.1 Order of Play

To check whether the order of playing Maze-Ball game variants affects the user's judgement of emotional preferences, we follow the order testing procedure described in (Yannakakis et al., 2008) which is based on the times that the subject prefers the first (K) or the second (J) game in both pairs. Given K and J the *order of play* test statistic is as follows: $r_c = (K - J)/N$, where N_s is the number of samples considered. The statistical analysis presented in Table 1 shows that the order of play does not significantly affect the emotional preferences of users. The insignificant order effects also, in part, demonstrate that effects such as a user's possible preference for the very first game played and the interplay between reported emotions and familiarity with the game are statistically insignificant.

Reported affective state	r_c	p-value
Fun	-0.150	0.214
Challenge	-0.222	0.121
Frustration	-0.066	0.427
Anxiety	-0.222	0.121
Boredom	-0.111	0.407
Excitement	-0.117	0.303
Relaxation	0.052	0.435

Table 1 Order of play correlation coefficients (r_c) and corresponding trinomially distributed p-values for all investigated affective states.

6.2 Statistical Analysis

An analysis for exploring statistically significant correlations between subject's expressed preferences and recorded physiological signal features is presented here. Correlation coefficients are obtained through $c(\mathbf{z}) = \sum_{i=1}^{N_s}\{z_i/N_s\}$, where N_s is the total number of game pairs where physiological signals were properly recorded and subjects expressed a *clear preference* for one of the two games (e.g. $A \succ B$ or $A \prec B$) and $z_i = 1$, if the subject
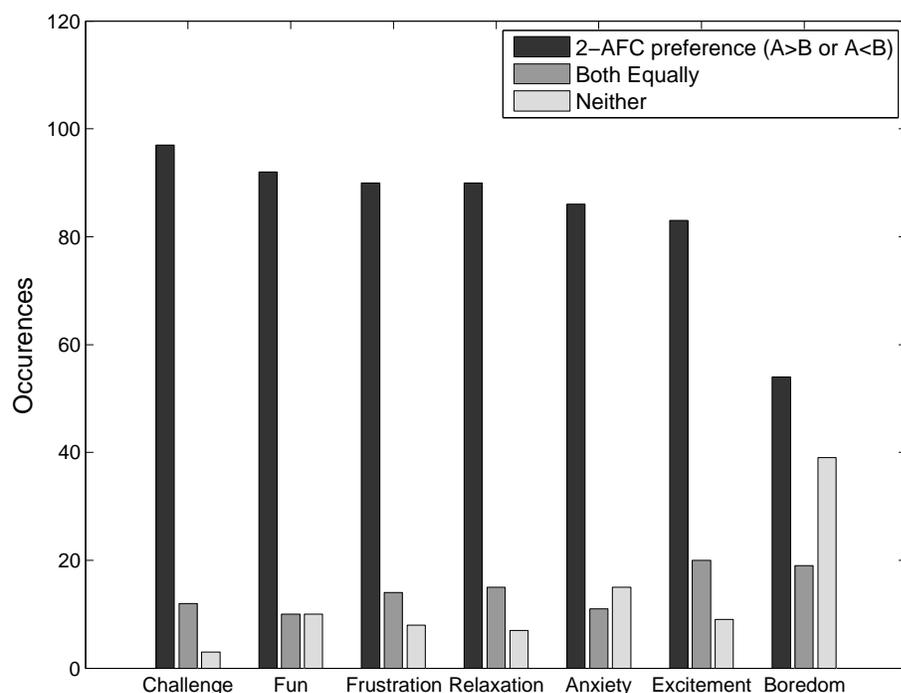
Fig. 6 Expressed emotional preferences

preferred the game with the larger value of the examined feature and $z_i = -1$, if the subject chooses the other game in the game pair i. Note that, N_s is 97, 92, 90, 90, 86, 83 and 54 respectively, for reported challenge, fun, frustration, relaxation, anxiety, excitement and boredom (see Figure 6).

The variation of the N_s numbers above indicates, up to a degree, the difficulty in expressing a clear emotional preference on different game variants. The percentage of $A \succ B$ and $A \prec B$ selection occurrences over all 112 preference instances for different affective states vary from 86.6% (challenge) to 48.2% (boredom) — see Figure 6. These percentages provide some first evidence that camera viewpoints (game variants) have a dissimilar impact on the affective states investigated. For instance, challenge, fun and frustration appear to be very much affected by varying camera viewpoints whereas boredom, on the contrary, does not appear as an emotion which is directly affected by the game. The amount of "no preference" ("both equally and "neither") occurrences for an affective state also indicates, in part, the appropriateness of the selected state for the test-bed game investigated. Note that, the high number of "neither" selection instances for boredom suggests the success of the Maze-Ball game design per se.

Significant correlations are observed between average HR and reported fun preferences (see Table 2). This effect is consistent with HR effects on fun preferences found through physical interactive game experiments (Yannakakis et al., 2008; Yannakakis and Hallam, 2008) showing the generality of the $E\{h\}$ effect on fun preferences across dissimilar game genres and modes of interaction. Moreover, the difference between minimum and maximum HR, D^h, and the last (s_{last}) and the maximum (max$\{s\}$) recording of the SC signal appear

Fun		Challenge	
$E\{h\}$	**0.347**	$E\{IBAmp\}$	**−0.257**
$\min\{h\}$	**0.260**	s_{in}	0.237
D^h	0.239	$\min\{s\}$	0.216
s_{last}	0.217	$t_{\min}\{s\}$	0.216
$\max\{s\}$	0.195	$\delta^b_{|2|}$	−0.196
Frustration		**Boredom**	
$t_{\min}\{h\}$	**0.244**	$\min\{h\}$	**−0.333**
HF	0.222	**Excitement**	
$\sigma\{RR\}$	0.200	$\min\{h\}$	0.204
RMS_{RR}	0.200	**Relaxation**	
Anxiety		$E\{h\}$	0.244
h_{last}	−0.233	$\min\{s\}$	−0.244
$\max\{h\}$	−0.209	R_s	0.200

Table 2 Statistically significant (p-value $< 5\%$) correlation coefficients ($c(\mathbf{z})$) between reported emotions and individual physiological features. Bold values denote highly significant (p-value $< 1\%$) effects.

to correlate significantly with reported fun. All three features indicate increased sympathetic arousal and appear reasonable correlates of fun.

The minimum HR recording correlates significantly with three emotions. It is positively correlated to fun and excitement and negatively correlated to boredom indicating heightened fun and excitement as well as decreased boredom when minimum HR is increased. Moreover, minimum HR appears to affect reported fun and boredom more than excitement.

Challenge, together with fun, yields the largest number of statistically significant effects with physiological features; the most interesting among them include the minimum recording of the SC signal, $\min\{s\}$, and the average inter-beat amplitude, $E\{IBAmp\}$, extracted from BVP. Suppressed $E\{IBAmp\}$ and increased minimum SC correlate to more challenging games and both indicate heightened sympathetic arousal.

Anxiety appears to be high when the last recording of the HR signal (h_{last}) and the maximum h ($max\{h\}$) are low. Both effects suggest that a decreasing HR correlates highly to high levels of reported anxiety. For anxiety, one would expect effects that indicate high arousal but the linear effects obtained in this experiment show the opposite.

For reported frustration preferences it appears that the further in the game the minimum HR recording ($t_{\min}\{h\}$) is highly correlated to the frustration of the user. Player frustration is also increased when the uniformity of inter-beat time intervals is lowered (high $\sigma\{RR\}$ and RMS_{RR}) and the changes in HR (high energy of HF) become quicker.

Average HR appears correlated to reported relaxation — on top of its highly significant effect to fun — showing a potential relationship between fun and relaxation via the $E\{h\}$ feature. Moreover, relaxation is anti-correlated to minimum SC indicating suppressed sympathetic arousal. The positive correlation to the correlation coefficient of SC (R_s) suggests increased relaxation with respect to linear temporal changes of the SC signal.

6.3 Non-Linear Models of Affect

The correlations calculated above provide linear relationships between individual biosignal features and reported emotions. However, these relationships are most likely more complex than can be captured by linear models. Given that linearly separable problems are extremely rare the aim of the analysis presented below is to construct non-linear computational models

for reported emotions and analyze the relationship between the selected features and expressed preferences. Furthermore it is our hypothesis that features which are significantly correlated with a specific affective state are not necessarily picked by automatic feature selection while designing inputs of the non-linear model of the affective state. Previous studies on user modeling via non-linear functions have demonstrated this inconsistency between significantly correlated and selected features (Yannakakis et al., 2008; Yannakakis and Hallam, 2008; Yannakakis et al., 2009). Results below validate this hypothesis once more: the obtained non-linear models are successful predictors of affective states without necessarily relying on significantly correlated features.

A designer cannot predict the affective state of the player from the available *controllable* features if those are not embedded in the derived models. Since controllable features (such as camera control parameters) are those that we can vary, and therefore those that can be optimized by artificial evolution or other optimization techniques, we need to be able to predict emotions, at least partly, from controllable features. Closing the affective loop via camera control defines the ultimate aim of this study; however, such a goal is not possible without controllable features embedded in the affective models constructed. For that purpose, all three controllable features are initially *forced* into the input vector of the model without feature selection. Then feature selection builds on the three-controllable features input model to choose physiological features that can assist in predicting the investigated affective state.

The inclusion of all controllable features into the model gives the designer all the flexibility the parameter space offers (i.e. *representation completeness*) to effectively tailor player experience by generating personalized camera control for the player. By enforcing features not necessarily chosen by feature selection one runs the risk of constructing models of lower accuracy since the hill climbing process of feature selection may initiate from undesired local minima. However, we choose to take the risk of lower performance of affective models which nevertheless will generate a more complete control parameter set for the camera.

Driven by the objectives just mentioned we evolve weights for non-linear single layer perceptrons (SLPs) as described in Section 5.1. The rationale behind the use of SLPs instead of multi-layer perceptrons (MLPs) is the *expressiveness* of the former. Using simple non-linear models (rather than more complex MLPs) allows for a clearer observation of the player physiological characteristics and camera control parameters that contribute to each affective state investigated. This discussion is vital for the deeper understanding of the unknown underlying function between statistical features of physiology, controllable in-game parameters and reported emotions. Multi-layer perceptrons have also been employed but not considered in this study. The resulting MLP networks do not generate significantly higher performances for all affective states which cannot counter-balance the expressiveness of the SLP.

The average performance (3-fold cross validation) of 20 trials of each feature selection algorithm is depicted in Figure 7(a) and the corresponding average number of features selected is illustrated in Figure 7(b). It is apparent that SFS, being a better search mechanism, outperforms nBest and PFS in selecting high performing feature sets of small size. As seen in results below, SFS selects small feature sets of size three to five to form the input vector of accurate SLP models.

Table 3 shows the highest performance (P) of an SLP obtained via SFS. Prediction accuracy of the models is above 80% for the majority of emotions (reaching 85.55% for frustration). Accuracy for fun, is just below 80%, while model performances for challenge and anxiety are 71.71% and 72.62%, respectively. The variation in model accuracy can be explained, in part, due to the variation of the play order test-statistic values which sug-
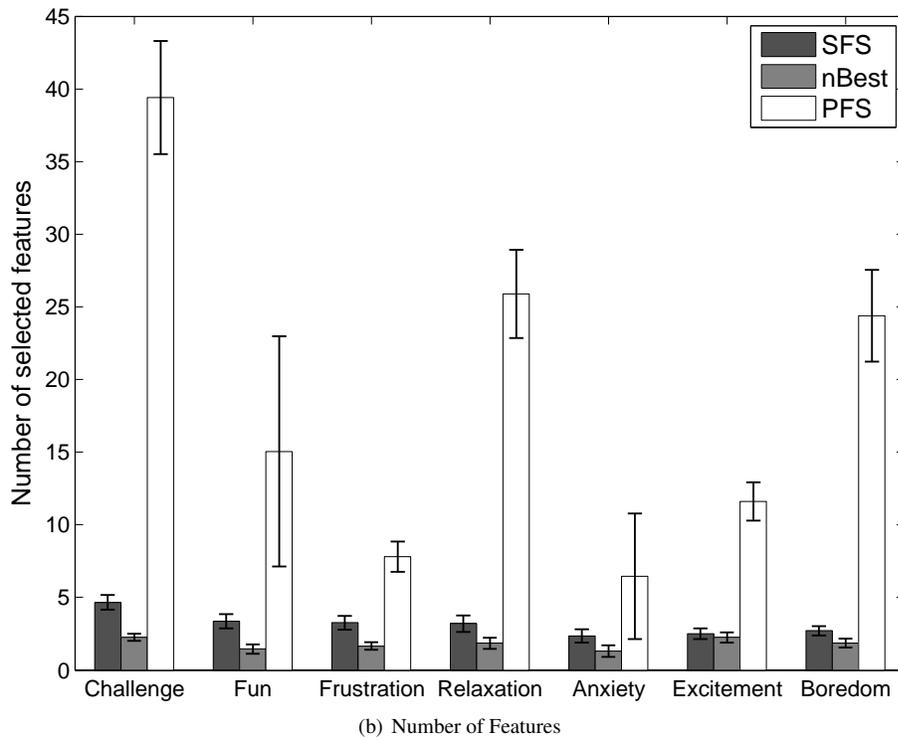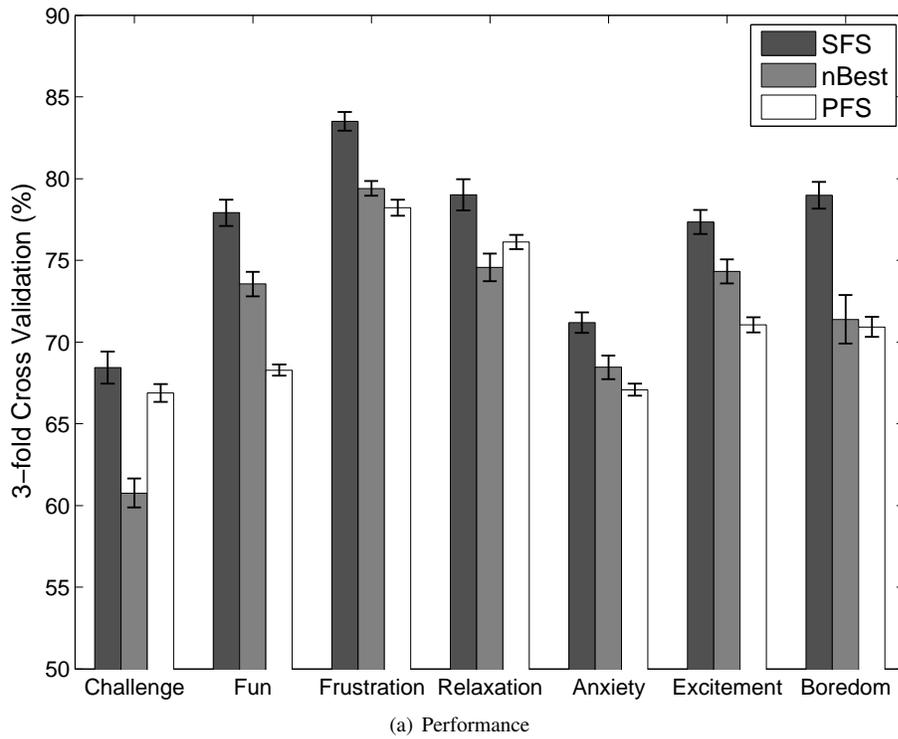
(a) Performance

(b) Number of Features

Fig. 7 Average performance and number of features of 20 trials of the three feature selection mechanisms for all seven affective states investigated. Confidence intervals (95%) of the averages are also illustrated in the figures.

	Challenge	Fun	Frustration	Relaxation	Anxiety	Excitement	Boredom
P	71.71%	78.49%	85.55%	82.22%	72.62%	81.48%	83.33%
P_c	44.44%	67.74%	72.22%	71.11%	65.48%	61.73%	50.00%
P_p	70.71%	65.59%	63.33%	70.00%	69.05%	66.67%	70.37%
P_{ps}	67.68%	64.52%	58.89%	60.00%	58.33%	38.27%	62.96%
P_l	70.71%	78.49%	83.33%	81.11%	72.62%	77.78%	79.63%
P_r	48.74%	52.42%	48.33%	50.17%	49.82%	50.86%	48.43%

Table 3 Preference models performance. Highest performing (via 3-fold cross validation) ANNs built on selected physiological features (P_p) via SFS; on the significantly correlated physiological features (P_{ps}) as those are presented in Table 2; on the controllable features (P_c); and on both selected (via SFS) physiological and controllable features (P). The best performance of linear ANN models built on selected (via SFS) physiological and controllable features (P_l) and the average performance of 20 ANNs with random connection weights (P_r) are presented for comparison purposes.

		Challenge		Fun		Frustration	
Physio	δ_2^s	-3.66	$E\{h\}$	3.09	$pRR50$	-2.79	
	$\delta_{|1|}^b$	-2.44	D^s	-2.24	$\sigma\{s\}$	1.50	
	$\sigma\{s\}$	-2.16	$\min\{s\}$	1.23	D_t^h	-1.08	
	L_h	2.02	VLF	-0.12	γ	0.17	
	D^s	-0.90					
H		-3.15		3.35		-1.98	
D		1.81		-1.01		0.27	
F_c		-1.14		-3.99		5.0	

	Relaxation		Anxiety		Excitement		Boredom	
Physio	$pRR50$	2.73	L_h	-2.87	δ_1^{Nb}	4.41	ρ_1^s	1.87
	h_{last}	2.75	$E\{s\}$	1.76	h_{last}	-3.26	β	-1.65
	δ_1^{Ns}	1.07	$t_{\min}\{h\}$	-0.54	$pRR50$	2.21	δ_2^{Ns}	-1.18
	R_s	0.30			ρ_1^s	-1.68	$\min\{s\}$	0.70
H	4.94		-4.33		0.09		1.27	
D	1.09		-1.93		-2.12		-1.19	
F_c	-3.89		4.16		-4.99		5.0	

Table 4 Learning from preferences: physiological (Physio) and controllable (D, H and F_c) features and corresponding connection weights for highest performing ANNs. Features are presented in descending order of their corresponding absolute connection values.

gest complexity of reporting the specific emotion within the Maze-Ball game. Challenge and anxiety reveal the highest, nevertheless non-significant, order of play correlation coefficients (see Table 1) among all reported affective states. On the other hand, frustration and relaxation have very small order of play correlation values which assist genetic search when building the SLP models. All evolved networks built on physiological and controllable features performed better than networks with random weights (P_r) — which reached chance level prediction accuracy — and networks built solely on controllable (P_c) or physiological features (P_p).

The difference between the SLP performance and the performance of random networks indicates the robustness of the SLP models generated; the difference between P and the performance of SLP models containing only controllable features (P_c) shows that subjectivity via individual player characteristics (e.g. physiology) is necessary for a more accurate approximation of reported affective states. It is also apparent from the performance of SLP models built on the three controllable features that camera parameters affect some emotional preferences (e.g. frustration) more than others (e.g. challenge). The difference

between P and the performances of networks containing only physiological features (P_p) suggests that the inclusion of game information into the model can help in preference prediction. Finally, the difference between the performance of SLPs employing a non-linear sigmoidal activation function and the SLPs employing a linear activation function (P_l) shows that even simple non-linear functions perform equally well to or better than linear approaches for modeling the affective states investigated. In particular, SLPs employing the sigmoidal activation function generate, on average, significantly higher performances than the performances of linear perceptrons in frustration (p($T \leq 6.74$) = $1.0 \cdot 10^{-7}$)[2] and boredom (p($T \leq 2.34$) = 0.02) and perform better, yet not significantly, in challenge (p($T \leq 1.78$) = 0.08) and excitement (p($T \leq -0.72$) = 0.48). Those effects validate, in part, the hypothesis that the function between camera parameters, player physiology and reported affect incorporates non-linear characteristics that only a non-linear model can approximate.

The weights of the highest performing SLP networks obtained via SFS are presented in Table 4. Very few features found to be significantly correlated with affective states (see Table 2) are picked from automatic feature selection as inputs to the non-linear SLP models. Specifically, average HR ($E\{h\}$) in fun; last HR recording (h_{last}) in anxiety; and correlation coefficient of the SC signal (R_s) in relaxation are the only three features that are both significantly correlated with their corresponding emotions and chosen by feature selection. No such feature exists for frustration, boredom, excitement and challenge. In addition, it is observed that automatic feature selection of physiological features (P_p) outperforms the accuracy of SLP models trained on the statistically correlated physiological features (P_{ps}), as those are illustrated in Table 2. The aforementioned findings validate the hypothesis that the physiological features which are significantly correlated with affective states are not likely to be picked by automatic feature selection in the process of constructing non-linear models of affect.

6.3.1 Challenge

Challenge is increased by lower camera position heights and faster transition between frames as shown by the corresponding connection weights of H and F_c. It is also worth noticing that D has a positive, but nevertheless smaller, impact to challenge suggesting that higher values of camera distance increase challenge.

There are five physiological features embedded in the challenge SLP model; the most physiological features selected by SFS among all affective states. According to the corresponding connection weights higher average 1-step absolute BVP differences ($\delta^b_{|1|}$) lead to lower reported challenge. Higher SC variability through the standard deviation and the 2-step differences of the SC signal lead to lower reported challenge. Challenge also increases with respect to the slope of the HR signal.

The one existing study investigating the impact of camera viewpoints on reported challenge focuses on the available information provided by the game (Schwartz et al., 2009) and not on physiological signals. The generated cross-validation performance in that study is 71.88% but the three camera controllable features are not embedded in the model therefore no comparative discussion can be made at this point.

---

[2] P-values are obtained through a t-test on the performances of the 20 independent ANN runs. The significance value is set to 5%.

6.3.2 Fun

Average heart rate (indicating heightened sympathetic arousal) appears to have a positive impact to reported fun. On the other hand the SC features related to fun prediction suggest that heightened SC signals (high $min\{s\}$) without too much deviation (low D^s) contribute to higher reported fun. The above-mentioned relationships appear commonsensical shaping the interplay between indicators of sympathetic arousal and fun. The average heart rate effect validates its appropriateness as a positively related predictor of fun showcased in several other studies (Yannakakis et al., 2008; Yannakakis and Hallam, 2008).

Furthermore, as derived from the SLP model, the higher the camera height and the lower the distance and frame coherence camera parameters, the higher the level of reported fun.

6.3.3 Frustration

According to the SLP model, frustration is increased with decreased $pRR50$ which is a HR variability indicator strongly dependent on vagal tone. Vagal tone, in turn, corresponds to parasympathetic activity of the Central Nervous System (CNS) (Massina et al., 1999). This relationship suggests that increased frustration occurs during parasympathetic inhibition as showcased in several medical studies (e.g. see (Brosschot and Thayer, 1998)). The standard deviation of SC has a positive impact to the predicted frustration value. Moreover, frustration is increased when the time difference between maximum and minimum HR is decreased.

Among the controllable features, it appears that slower frame-to-frame coherence and lower camera height have a positive impact on increased frustration. It is also worth noticing that no significantly correlated feature is selected for the non-linear SLP model of frustration.

6.3.4 Relaxation

Unsurprisingly, higher distance and height values combined with fast frame-to-frame transition (low F_c values) contribute positively to higher relaxation. Reasonably, low parasympathetic activity, inferred via low values of $pRR50$, contributes to high levels of reported relaxation. The somewhat surprising relationship is the one between last HR recording and relaxation; the connection weight suggests that higher values of h_{last} contribute to higher relaxation values.

6.3.5 Anxiety

The connection weights of the physiological features of the SLP model of anxiety are relatively small. It is worth noticing that average SC has a positive relationship with anxiety whereas, surprisingly, the linear slope of the HR signal is negatively related to anxiety.

Frame coherence and distance have respectively a large positive and negative impact to anxiety. Similar frame coherence effects are observed in the frustration and boredom models. The negative effects of both the height and the distance parameter of the camera are somewhat expected since low values of both generate a rather unpredictable and stressful gameplay.

6.3.6 Excitement

The higher the one-step variation within the BVP signal showcased by increased values of δ_1^{Nb}, the higher the excitement. Unexpectedly, however, excitement is increased during heightened parasympathetic activity as showcased via its positive relationship with the $pRR50$ feature. Puzzling is also the effect of low excitement caused due to increased last HR recordings. One would expect that feature selection would either pick features corresponding to sympathetic arousal for excitement or neuro-evolution would provide opposite signs to the connection weights of $pRR50$ and h_{last}. The negative relationship of ρ_1^s and excitement suggests that lower autocorrelation values which imply more randomness in the SC signal generate higher excitement.

While lower values of frame coherence and distance generate higher excitement, height has a rather small positive impact to excitement. The combination of such parameter values generates a fast and rather unpredictable gameplay experience which is reported as exciting by subjects.

6.3.7 Boredom

The most notable relationship is that between boredom and frame coherence in which higher values of F_c (slower frame-to-frame transitions) affect boredom positively. Moreover, sensibly, the lower the two-step variation within the SC signal showcased by decreased values of δ_2^{Ns}, the higher the boredom. Furthermore lower curvature values, β, of the HR quadratic model — which indicate a trend of decreasing HR values as the game progresses — are associated with higher reported boredom. Boredom increases when the SC signal is easier to predict with a time-series model as implied by stronger autocorrelation values of the SC signal.

7 Generality

This section presents an attempt to test the generality of the affective models across games of different game design properties. For this purpose a dissimilar maze game named *Maze-Ball-Low* was designed and played in the game survey experiment. The game is a clone of Maze-Ball with a very important difference: walls have a very low height value; walls are still visible though defining the shape of the maze (see Figure 7). The absence of walls as viewing obstacles in Maze-Ball generates a completely new playing experience with respect to available visual information. It is, therefore, expected that camera control on the Maze-Ball-Low game has a dissimilar impact to player psychophysiology.

The reader is reminded that there are eight games, in total, played by each subject during the main game survey experiment. The Maze-Ball-Low variant (named as variant 9 in Section 3) is evenly placed in either of the eight places of the game sequence to minimize any potential order effects. Each subject is asked to express her emotional pairwise preferences between the Maze-Ball-Low variant and a Maze-Ball variant. Both game variants have the same camera profiles (i.e. camera control parameter set).

To investigate the extent to which the predictive models of emotional preferences computed using the data from Maze-Ball generalizes to a dissimilar game the best performing evolved SLPs of Table 4 are presented with and evaluated on the unseen data and preferences
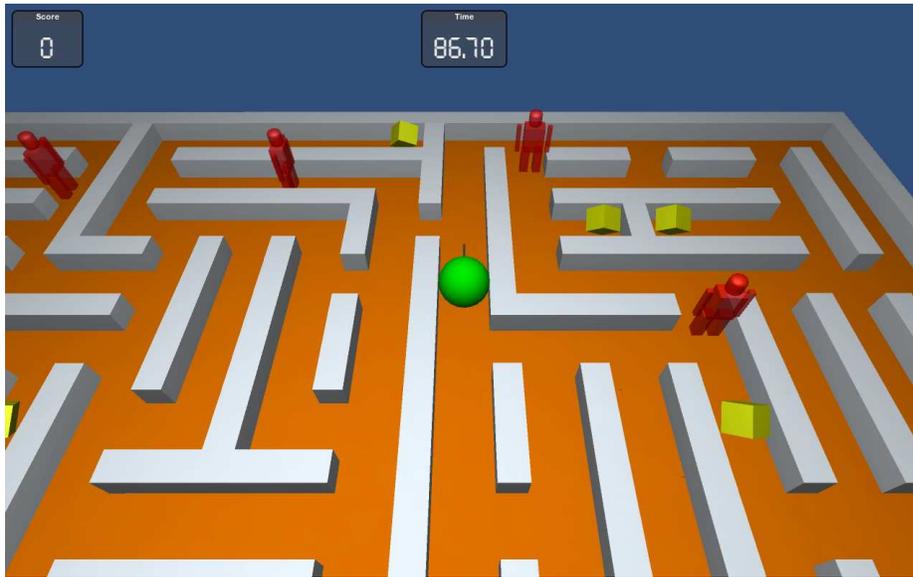
Fig. 8 Snapshot of the Maze-Ball-Low game.

of the Maze-Ball-Low game. Table 5 shows a comparison between the average total classification accuracy on Maze-Ball (as previously presented in Table 4 and on Maze-Ball-Low of all three evolved SLPs.

	Maze-Ball-Low	ΔP
Challenge	55.95	15.76
Fun	58.02	20.47
Frustration	68.18	17.37
Relaxation	61.54	20.68
Anxiety	52.77	19.85
Excitement	51.66	29.82
Boredom	48.18	35.15

Table 5 Classification accuracy (%) on unseen preference data of the Maze-Ball-Low (P_L) game and performance difference (%) from the accuracy obtained on the Maze-Ball game ($\Delta P = P - P_L$). Random SLP performance reaches chance level.

It is apparent that the performances obtained on the Maze-Ball-Low game are lower compared to the percent accuracies on the validation data of the Maze-Ball game. Even though for some affective states (anxiety, challenge, excitement and boredom) performance reaches chance level on the unseen data of the Maze-Ball-Low game the SLP models predict fun and relaxation with accuracies around 60% and frustration with accuracy of 68.18%. It therefore appears that the generalizability of physiological signal features varies across different affective states. Note that the values of the three controllable features for the two games compared are identical (both Maze-Ball and Maze-Ball-Low games have the same camera profile); thus, the prediction is relying solely on the physiological features selected.

Models appear to generalize up to a good degree for fun, frustration and relaxation in a completely different game where the impact of walls on player experience is substantial. Model performances also indicate that these three discrete affective states can be predicted well via physiological signals across dissimilar games. The prediction of the remaining four states could be supported by additional modalities of input such as generic in-game parameters and keyboard events. Preliminary results showcase the effectiveness of multimodal inputs resulting to increased validation performance of the SLP models.

8 Background

8.1 Affective Games

The concept of affective gaming has recently attracted significant attention within the affective computing community (Hudlicka, 2009). Measurements of physiological quantities have been used extensively for emotion recognition of users while playing games (Fairclough, 2009). Correlations between physiological signals — galvanic skin response (GSR), jaw electromyography (EMG), respiration and cardiovascular measures — and reported user experiences in computer games have been examined in various genres of games (Mandryk and Atkins, 2007; Ravaja et al., 2006; Hazlett, 2006; Rani et al., 2005; McQuiggan et al., 2006). However, this study focuses on the construction of non-linear models of pairwise emotional preferences of players using stochastic optimization. Previous studies showcase the effectiveness of preference learning on building models of fun preferences in physical interactive games (Yannakakis et al., 2008; Yannakakis and Hallam, 2008; Yannakakis et al., 2009) and models of player experience (Pedersen et al., 2009, 2010) for the purpose of evaluating level content in platform games.

Camera control can be viewed as an important component of the overall affective game engine envisioned by Hudlicka (2009); however, to the best of the authors' knowledge, the role of the camera in affective interaction has not been investigated yet. This paper is novel in that it introduces the concept of affective camera control and investigates the impact of camera parameter settings on psychophysiology of players.

8.2 Camera Control

One central problem in the field of automatic camera control is the selection of viewpoints in 3D space in real-time in a way that communicates information present in a 3D environment effectively. Previous work on intelligent camera control has focused mainly on the frame-by-frame graphical placement of the camera to satisfy given cinematic constraints (Bares et al., 2000; Nicholas et al., 2001; Christianson et al.). Less attention has been paid to informing the placement of the camera over time based on the context of unfolding events (Amerson et al., 2005; Jhala and Young, 2005) within 3D environments and the affective communication of these events through the camera.

The problem of optimizing camera parameters has received significant attention from the research community (Christie and Olivier, 2006). Earliest approaches (Ware and Osborne, 1990; Blinn, 1988; Gleicher and Witkin, 1992) focused on interactive 3D camera movement within the parameterized space of camera variables such as location, rotation, and field-of-view. Optimization approaches have used techniques like evolutionary algorithms (Burelli and Yannakakis, 2010a,b), artificial neural networks (Pickering, 2002), par-

ticle swarm optimization (Burelli et al., 2008), and constraint satisfaction solvers (Christie and Normand, 2005) for improving the performance and quality of camera compositions. These approaches have applied different mathematical representations for constraint satisfaction to select feasible 3D volumes, optimizing search within selected volumes based on high-level constraints.

Recent approaches to camera control (e.g. (Jhala and Young, 2005, 2006)) have started investigating the motivation behind camera placement that is influenced by the underlying content and communicative affective goals. Such approaches, mainly oriented towards narrative communication systems, consider the underlying context of the narrative to generate planned action sequences for the camera. Affective adaptation of presentation is a key aspect involved in the design of such camera control systems. None of the current approaches have addressed the importance of a virtual camera in the affective loop for human interaction. Evaluation of existing camera systems has either been on runtime performance of algorithms (Christie et al., 2005a), or user-modeling based on specific interaction models in interactive scenarios (Bares and Lester, 1997). The affective modeling approach presented in this paper could be used to evaluate different camera systems for their efficacy in modulating affective experience for users.

9 Discussion

For the experiments presented in this paper, we only defined three controllable features of the camera. The primary reason for restricting ourselves to three features is the combinatorial explosion of the required number of subjects given the experimental methodology followed. Random sampling of the hyperplane of the generated game variant space should be considered for such an attempt. Even though distance, height and frame coherence encapsulate the dynamics of camera control well, a number of additional controllable features could easily be designed, given an adequate number of experimental subjects; for example the center of view and angle of camera shot.

One of the limitations of the experimental protocol proposed is post-experience. Users report affective states after playing a pair of games which might generate memory dependencies in the reports. Effects such as order of play and game learnability might also be apparent and interconnected to memory. The experimental protocol, however, is designed to test for order of play effects which, in part, reveal memory (report consistency over different orders) and learnability effects, if any. Results showcase that reported challenge and anxiety are affected (yet not significantly) by the order of play which in turn reveals potential memory and/or learnability effects for these two particular affective states. The insignificant, yet notable, order of play effect on challenge and anxiety appears to be generic over dissimilar games and study purposes. Significant order of play effects on reported challenge and anxiety are demonstrated in a similar study involving the control of the level generator for the popular platform game Super Mario Bros (Pedersen et al., 2010). In that study and the study presented here challenge and anxiety are the states most difficult to predict with the use of ANNs.

Forced report (4-AFC) provides viable data for a machine learning algorithm but it does not necessarily capture the dynamics of the experience. On the other hand, free emotional report could potentially provide more genuine response but it is harder to analyze and requires a laboratory experimental setup, which is not desired given the aims of this study. There exist solutions for both testing emotional models over different time windows as introduced in (Yannakakis and Hallam, 2009) and capturing the association between gameplay dynamics

and emotional responses via e.g. recurrent neural networks. Both include future directions of this research prior to adjusting the camera behavior for maximizing player experience.

Building computational models of affect defines the intermediate step towards adaptive camera control since one (e.g. a game designer) can potentially use such models for adjusting camera internal controls (*distance*, *height*, *frame coherence*) in real-time to maximize the model's output value (e.g. fun). Note that, the constraint-based camera controller used here can support the transition between any camera profile designed. Furthermore, the methodology of our experiments is useful to game designers in verifying the effect of their choice of viewpoints on the desired affective states of the players.

10 Conclusions

This paper introduces the first step toward affective camera control in games. An efficient experimental protocol was designed and the quantitative impact of subject's reported emotions on HR, BVP and SC signal statistics was investigated through a prey/predator 3D action game (Maze-Ball). Player performance in the Maze-Ball game is directly affected by camera viewpoint parameters. Statistical analysis on the data collected reveals that the order of playing Maze-Ball variants does not appear to significantly affect any affective state preferences expressed by the subjects. Furthermore, the statistical effects obtained provide some first insights for the physiology of reported emotions and the impact of camera viewpoint on the psychophysiological state of the player. Several features indicating sympathetic activity of the heart appear to correlate with reported emotions, in participants of our experiments. Highly significant effects are found between average and minimum heart rate and fun; minimum heart rate and boredom; average inter-beat amplitude and challenge; and time of the minimum HR recording and frustration.

According to the methodology proposed, we do not attempt to infer values for appraisal dimensions (e.g. arousal and valence) through HR, SC and BVP (Mandryk and Atkins, 2007) nor built affective models on annotated data streams; instead, we assume there is an unknown mapping (unknown number of dimensions and search space) between camera viewpoint, player characteristics (derived from physiology) and reported discrete affective states. Furthermore, given the variation of the notion of scaling across subjects, we believe that subjective affective models can be built more accurately on reported data derived from pairwise emotional preferences. Based on work by Yannakakis and Hallam (2008) and Yannakakis et al. (2008), we assume that the mapping between physiological characteristics and reported emotional preferences is non-linear. For this purpose, neuro-evolutionary preference learning (Fürnkranz and Hüllermeier, 2005) was utilized for the construction of non-linear computational models (i.e. artificial neural networks) of reported emotions.

Results show that construction of accurate predictors of reported affective states is possible using neuro-evolutionary preference learning. Single layer perceptrons whose input vector is designed by automatic feature selection manage to capture the biosignal dynamics associated to reported emotions. Only very few significantly correlated biosignal features are included as inputs in the single-layer perceptron models obtained and these models vastly outperform the models trained on only significantly correlated features. These findings validate the hypothesis of the inappropriateness of linear relationships as being accurate indicators (or estimators) of affect for most affective states investigated.

Tests performed to check the generality of the obtained models across dissimilar games show that models are generic up to a degree for specific affective states. In particular, fun, frustration and relaxation are predicted well in a game that generates dissimilar playing

experiences to Maze-Ball with respect to visual information available. Results from this analysis suggest that biosignal features are good predictors of affect for fun, frustration and anxiety independently of important gameplay features like wall height. However, wall height in Maze-Ball appears to affect challenge, anxiety, excitement and boredom in a fashion which cannot be captured by physiological indices; additional modalities of input seem to be required for better predictions of those affective states.

The ANN models incorporate all three controllable features of the camera which is a choice forced by the designer. The exact relationship between camera control, player's physiology and affective state is therefore embedded in the ANNs. To effectively close the affective loop one could vary camera controls and observe changes in the output of the model by computing the partial derivatives of the output (e.g. frustration) with respect to the controllable input (e.g. distance) as proposed in (Yannakakis and Hallam, 2009). Gradient information of the ANN model will solely provide an indication of whether the associated controllable features should be increased or decreased for the affective predictor output to be set to a pre-defined value. Evidently, the designer (or an automated process like reinforcement learning) may then choose *how often* and by *how much* the parameters of the camera should change during interaction for eliciting the desired playing experience.

Acknowledgements

References

Dan Amerson, Shaun Kime, and R. Michael Young. Real-time cinematic camera control for interactive narratives. In *Advances in Computer Entertainment Technology*, page 369, 2005.

Ruth Aylett and Joao Dias Ana Paiva. An affectively driven planner for synthetic characters. In *Proceedings of ICAPS*, pages 2–10. AAAI, 2006.

Ruth Aylett, Sandy Louchart, Joao Dias, Ana Paiva, Marco Vala, Sarah Woods, and Lynne Hall. Unscripted narrative for affectively driven characters. *IEEE Computer Graphics and Applications*, 26(3):42–52, 2006.

William Bares and James Lester. Cinematographic user models for automated realtime camera control in dynamic 3d environments. In *User Modeling: In Proceedings of the sixth international conference, UM97*, 1997.

William H. Bares, Scott McDermott, Christina Boudreaux, and Somying Thainimit. Virtual 3d camera composition from frame constraints. In *ACM Multimedia*, pages 177–186, 2000.

James Blinn. Where am i? what am i looking at? *IEEE Comput. Graph. Appl.*, 8(4):76–81, 1988.

Owen Bourne, Abdul Sattar, and Scott D. Goodwin. A constraint-based autonomous 3d camera system. *Constraints*, 13(1-2):180–205, 2008.

Jos F. Brosschot and Julian F. Thayer. Anger inhibition, cardiovascular recovery, and vagal function: A model of the link between hostility and cardiovascular disease. *Journal Annals of Behavioral Medicine*, 20(4):326–332, 1998.

P. Burelli and G. N. Yannakakis. Combining Local and Global Optimisation for Virtual Camera Control. In *Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games*, pages 403–401, Copenhagen, Denmark, August 2010a. IEEE.

P. Burelli and G. N. Yannakakis. Global Search for Occlusion Minimization in Virtual Camera Control. In *Proceedings of the 2010 IEEE World Congress on Computational Intelligence*, pages 2718–2725, Barcelona, Spain, July 2010b. IEEE.

Paolo Burelli, Luca Di Gaspero, Andrea Ermetici, and Roberto Ranon. Virtual camera composition with particle swarm optimization. In *Smart Graphics*, pages 130–141. Springer-Verlag, 2008.

David B. Christianson, Sean E. Anderson, Li wei He, David Salesin, Daniel S. Weld, and Michael F. Cohen. Declarative camera control for automatic cinematography. In *AAAI/IAAI, Vol. 1*, pages 148–155.

Marc Christie and Jean-Marie Normand. A semantic space partitioning approach to virtual camera composition. *Computer Graphics Forum*, 24(3):247–256, 2005.

Marc Christie and Patrick Olivier. Camera Control in Computer Graphics. pages 89–113. Eurographics Association, 2006.

Marc Christie, Rumesh Machap, Jean-Marie Normand, Patrick Olivier, and Jonathan Pickering. Virtual camera planning: A survey. In *Smart Graphics*, pages 40–52, 2005a.

Marc Christie, Rumesh Machap, Jean-Marie Normand, Patrick Olivier, and Jonathan Pickering. Virtual camera planning: A survey. In *Smart Graphics*, pages 40–52, 2005b.

C. Conati, R. Chabbal, and H. Maclaren. A Study on Using Biometric Sensors for Detecting User Emotions in Educational Games. In *Proceedings of the Workshop "Assessing and Adapting to User Attitude and Affects: Why, When and How?". In conjunction with UM03, 9th International Conference on User Modeling*, pages 60–65, Pittsburgh, PA, U.S.A., 2003.

Stephen H. Fairclough. Fundamentals of physiological computing. *Interact. Comput.*, 21 (1–2):133–145, 2009.

Johannes Fürnkranz and Eyke Hüllermeier. Preference learning. *Künstliche Intelligenz*, 19 (1):60–61, 2005.

Michael Gleicher and Andrew Witkin. Through-the-lens camera control. In *Computer Graphics*, pages 331–340, 1992.

Jeffrey J. Goldberger, Sridevi Challapalli, Roderick Tung, Michele A. Parker, and Alan H. Kadish. Relationship of heart rate variability to parasympathetic effect. *Circulation*, 103: 1977–1983, 2001.

Richard L. Hazlett. Measuring emotional valence during interactive experiences: boys at video game play. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1023–1026, New York, NY, USA, 2006. ACM Press.

Eva Hudlicka. Affective game engines: Motivation and requirements. In *Proceedings of the 4th Int. Conf. on the Foundations of Digital Games*, pages 299–306. ACM, 2009.

IOM Biofeedback Equipment. Wild divine. `http://www.wilddivine.com/`.

Arnav Jhala and R. Michael Young. A discourse planning approach to cinematic camera control for narratives in virtual environments. In *AAAI*, pages 307–312, 2005.

Arnav Jhala and R. Michael Young. Representational requirements for a plan based approach to virtual cinematography. In *Artificial Intelligence in Interactive Digital Entertainment (AIIDE)*, Marina Del Rey, CA, 2006. AAAI, AAAI Press.

R. L. Mandryk and M. S. Atkins. A Fuzzy Physiological Approach for Continuously Modeling Emotion During Interaction with Play Environments. *International Journal of Human-Computer Studies*, 65:329–347, 2007.

H. P. Martinez, A. Jhala, and G. N. Yannakakis. Analyzing the Impact of Camera Viewpoint on Player Psychophysiology. In *Proceedings of the Int. Conf. on Affective Computing and Intelligent Interaction*, pages 394–399, Amsterdam, The Netherlands, September 2009. IEEE.

M. M. Massina, B. Derkenne, and G. von Bernuth. Correlations between Indices of Heart Rate Variability in Healthy Children and Children with Congenital Heart Disease. *Cardiology*, 91(2):109–113, 1999.

Scott McQuiggan, Sunyoung Lee, and James Lester. Predicting User Physiological Response for Interactive Environments: An Inductive Approach. In *Proceedings of the 2nd Artificial Intelligence for Interactive Digital Entertainment Conference*, pages 60–65, 2006.

Manuel Mejia-Lavalle and Gustavo Arroyo-Figueroa. Power System Database Feature Selection Using a Relaxed Perceptron Paradigm. In *Proceedings of 5th Mexican International Conference on Artificial Intelligence, LNCS*, pages 522–531. Springer Berlin/Heidelberg, 2006.

Halper Nicholas, Helbing Ralf, and Strothotte Thomas. A camera engine for computer games: Managing the trade-off between constraint satisfaction and frame coherence. In *proceedings Eurographics*, volume 20, pages 174–183, 2001.

Ana Paiva, Gerd Andersson, Kristina Hook, Dario Mouraand Marco Costa, and Carlos Martinho. SenToy in FantasyA: Designing an Affective Sympathetic Interface to a Computer Game. *Personal and Ubiquitous Computing*, 15(4):378–389, 2002.

Chris Pedersen, Julian Togelius, and Georgios N. Yannakakis. Modeling Player Experience in Super Mario Bros. In *Proceedings of the IEEE Symposium on Computational Intelligence and Games*, pages 132–139, Milan, Italy, September 2009. IEEE.

Chris Pedersen, Julian Togelius, and Georgios N. Yannakakis. Modeling Player Experience for Content Creation. *IEEE Transactions on Computational Intelligence and AI in Games*, 2(1):54–67, 2010.

Rosalind W. Picard, Elias Vyzas, and Jennifer Healey. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(10):1175–1191, 2001.

Jonathan Pickering. *Intelligent Camera Planning for Computer Graphics*. PhD thesis, University of York, 2002.

P. Rani, N. Sarkar, and C. Liu. Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th International Conference on Human Computer Interaction*, 2005.

Niklas Ravaja, Timo Saari, Marko Turpeinen, Jari Laarni, Mikko Salminen, and Matias Kivikangas. Spatial Presence and Emotions during Video Game Playing: Does It Matter with Whom You Play? *Presence Teleoperators & Virtual Environments*, 15(4):381–392, 2006.

M. Schwartz, H. P. Martinez, G. N. Yannakakis, and A. Jhala. Investigating the Interplay between Camera Viewpoints, Game Information, and Challenge. In *Proceedings of Artificial Intelligence and Interactive Digital Entertainment (AIIDE'09)*, Palo Alto, CA, October 2009. AAAI Press.

Colin Ware and Steven Osborne. Exploration and virtual camera control in virtual three dimensional environments. *SIGGRAPH*, 24(2):175–183, 1990.

Georgios N. Yannakakis. Preference Learning for Affective Modeling. In *Proceedings of the Int. Conf. on Affective Computing and Intelligent Interaction*, pages 126–131, Amsterdam, The Netherlands, September 2009. IEEE.

Georgios N. Yannakakis and John Hallam. Entertainment Modeling through Physiology in Physical Play. *International Journal of Human-Computer Studies*, 66:741–755, October 2008.

Georgios N. Yannakakis and John Hallam. Real-time Game Adaptation for Optimizing Player Satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(2):121–133, June 2009.

Georgios N. Yannakakis, John Hallam, and Henrik Hautop Lund. Entertainment Capture through Heart Rate Activity in Physical Interactive Playgrounds. *User Modeling and User-Adapted Interaction, Special Issue: Affective Modeling and Adaptation*, 18(1-2): 207–243, February 2008.

Georgios N. Yannakakis, M. Maragoudakis, and John Hallam. Preference Learning for Cognitive Modeling: A Case Study on Entertainment Preferences. *IEEE Systems, Man and Cybernetics; Part A: Systems and Humans*, 39(6):1165–1175, November 2009.