Fuzzy Affective Player Models: A Physiology-Based Hierarchical Clustering Method

Pedro A. Nogueira[1,3,4], Rúben Aguiar[3], Rui Rodrigues[2,3], Eugénio Oliveira[1,3], Lennart E. Nacke[4]
[1] Artificial Intelligence and Computer Science Laboratory, [2] INESC-TEC, [3] University of Porto, Portugal
[4] HCI Games Group, University of Ontario\ Institute of Technology, Ontario, Canada
{pedro.alves.nogueira, rui.rodrigues, eco}@fe.up.pt, r.aguiar9080@gmail.com, lennart.nacke@acm.org

Abstract

Current approaches to game design improvements rely on time-consuming gameplay testing processes, which rely on highly subjective feedback from a target audience. In this paper, we propose a generalizable approach for building predictive models of players' emotional reactions across different games and game genres, as well as other forms of digital stimuli. Our input agnostic approach relies on the following steps: (*a*) collecting players' physiologically-inferred emotional states during actual gameplay sessions, (*b*) extrapolating the causal relations between changes in players' emotional states and recorded game events, and (*c*) building hierarchical cluster models of players' emotional reactions that can later be used to infer individual player models via fuzzy cluster membership vectors. We expect this work to benefit game designers by accelerating the affective play-testing process through the offline simulation of players' reactions to game design adaptations, as well as to contribute towards individually-tailored affective gaming.

1 Introduction

Over the past two decades, videogames have propelled many of the breakthroughs in computer graphics, artificial intelligence, and interaction techniques, among others. These advances have led us to the modern photorealistic virtual worlds filled with believable character behaviours and precise physics systems that bring players to an ever-growing level of immersion and enable games to elicit a wide range of emotions.

Although the search for reasons that lead people to play (Gilleade, Dix, & Allanson, 2005; Ryan, Rigby, & Przybylski, 2006) and why it's a pleasurable experience (Ermi & Mäyrä, 2005) are subjects that have been studied over the years, it is only now, with the stagnating improvements on audio-visual realism, that the gaming industry has started to focus on the players' affective experience.

It is a commonly accepted truth that video games must provide an engrossing experience by transporting the player from the real to the virtual world. Understanding and improving on this ability to deeply engage the player seems to be the key to producing better gaming experiences (Ermi & Mäyrä, 2005). This search for better levels of user experience embodies the philosophy behind Affective Computing; detecting and using human emotions as inputs to steer the overall experience towards a desirable one.

Along with improving the existing gameplay experience, predicting players' emotional reactions can give game developers a powerful tool to more finely-tune the original experience before releasing a game and accelerate the game design process - ultimately resulting in not only improved, but also more quickly produced titles.

Our main contribution focuses on a method to model player's emotional reactions to game stimuli based on fuzzy memberships to player clusters obtained through a bootstrapped hierarchical clustering algorithm. Our clustering approach initially attempts to approximate individual players' emotional response functions to game events based on a set of observed emotional reactions. Individual player model pairs are then compared to find groups of players that share similar reactions, through the hierarchical clustering algorithm. Using the clusters' more robust models, we compute players' fuzzy membership vectors to the found clusters, based on a simple distance function.

2 Player Modelling

Player modelling has been a popular topic in game research, mainly because of the substantial advantage that knowing in advance how a player might chose to act provides in terms of potential future game adaptations.

More recent approaches have taken this concept and attempted to use them for modelling player experience (Yannakakis & Togelius, 2011). The most straightforward way to do this is through player's questionnaire responses.

Although this process may create very accurate models (Shaker, Yannakakis, & Togelius, 2010), the considerable presence of experimental noise (derived from human error in self-judgment, memory, etc.) and the intrusiveness of the method can lead to difficulty in analysing the data. The work by (Tognetti et al., 2010) shows how self-reports can be successfully used to capture aspects of player experience. Similarly, in (Shaker et al., 2010) player experience models were built on an emotional basis using crowd-sourced data on player actions and level design features.

A less intrusive approach relies on using gameplay data (e.g. how many times a certain action was performed) in an attempt to build these models (Etheredge, Lopes, & Bidarra, 2013). The main assumption is that player actions and preferences are linked to player experience, making it possible to infer players' emotional states by studying their interaction patterns. This approach is the least intrusive one, thus becoming a candid possibility for real world usage. However as stated in (Yannakakis & Togelius, 2011), the models are often based on several strong assumptions that relate player experience to gameplay actions and preferences, resulting in a low-resolution, often unsatisfactory or over simplistic, model of players' affective experience.

Prior work on physiological player modelling has achieved promising results in predicting subjective player experience reports using simple physiological metrics (Lankes et al., 2012; Martinez, Garbarino, & Yannakakis, 2011; Vachiratamporn et al., 2013). Game narratives have also been shown to be dynamically adaptable in response to players' physiological state (Gilroy & Porteous, 2012), proving this technology can be applied to various facets of the gaming experience. We hypothesize that the success obtained by these approaches may be due to their usage of a more objective and continuous data source that arguably lowers the amount of experimental noise and provides a richer data source.

While, as in the aforementioned work, we employ physiological metrics, our method is input independent, abstracting the emotional data as an *n*-dimensional waveform and game events as class labels (see Section 4).

3 Testbed Horror Game

Vanish is a survival horror videogame where players must navigate a network of procedurally generated maze-like tunnels to locate a set of key items, before being allowed to escape (Fig. 1). During gameplay, the player must avoid a creature that continuously stalks him. Several visual and audio events (e.g. lights failing, steam or water pipes bursting or the creature distant cries) also occur, in order to keep the player engaged. These events, along with creature encounters, deaths and locating new items are automatically logged by the game and constitute the set of considered game events. The dataset used in this study was extracted from 72 gameplay sessions from 24 participants, ranging 3 different game versions varying in level layout, pacing, game mechanics and difficulty. The dataset comprised over 30 hours of logged gameplay events and physiological data (Fig. 3). Before playing, players underwent a physiological calibration session and a brief game tutorial before playing the game in an isolated environment.

Figure 1. Screenshot of a creature encounter event on a Vanish gameplay session.

4 Data Collection & Extraction

As we have previously discussed, our aim is to model players' individual emotional response functions to a set of game events $G=\{g_1,...,g_k\}$, given an initial (henceforth referred to as prior) emotional state λ_p, such that $\lambda_p \in \Lambda$, the set of considered emotional states (see Section 5 for a formal definition).

One of the major issues in building physiological-based models is acquiring enough data from (unbiased) gameplay sessions. In this section we describe how the physiological data was interpreted as emotional states and how the emotional reactions to each game event were extracted.

Emotional State Detection

We estimate players' emotional states in terms of the arousal-valence dimensions (Russel, 1980) through their physiological readings. As physiological data varies considerably between individuals, we employ the regression-based approach proposed by (Nogueira, Rodrigues, Oliveira, & Nacke, 2013b) to properly scale these readings using data gathered during a pre-game calibration session. We then apply the grounded rules proposed by (Mandryk & Atkins, 2007) to convert the normalised readings into arousal and valence.

Arousal was derived from skin conductance (SC) and heart rate (HR) data, while valence was derived from facial electromyography (EMG) measured at two sites and HR as a fall back. Regarding sensor placement, SC was measured at the player’s left index and middle fingers using two Ag/AgCL surface sensors snapped to Velcro straps; HR

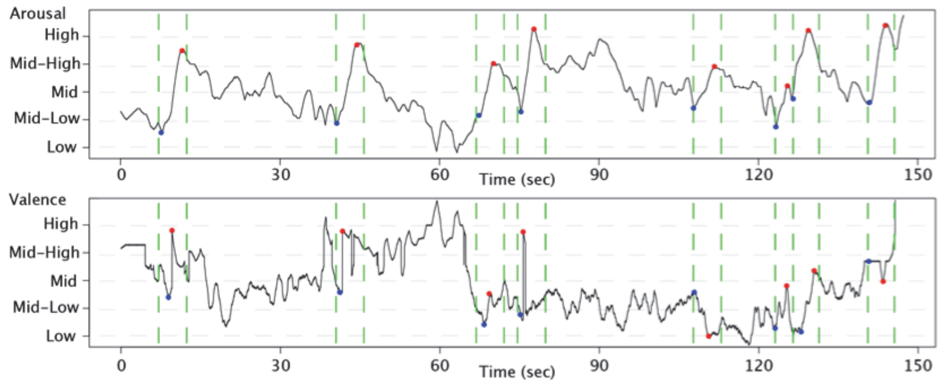

Figure 2. Emotional signal waveform with annotated reactions to game events. Arousal and valence levels are shown discretized into 5 levels for illustrative purposes. Blue and red circles denote the player's prior and response emotional state, respectively. Green dotted lines represent the event associated time window T (~5 seconds in this illustrative example).

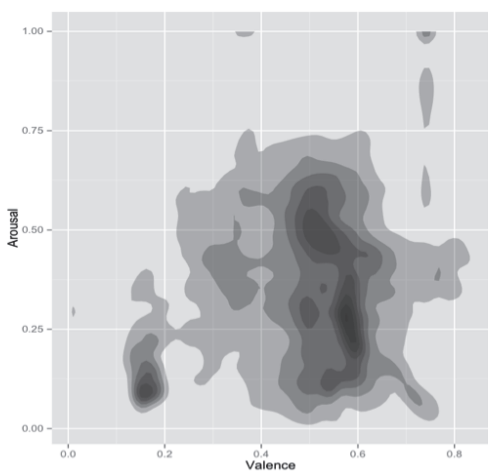

Figure 3. Density plot of the normalised elicited emotional spectrum over all 72 gameplay sessions.

was derived from BVP, measured by a clip-on sensor at the left thumb; and facial EMG was measured at the zygomaticus major and the corrugator supercilii muscles.

While our method is not reliant on physiological data – any continuous *n*-dimensional waveform is acceptable (see definitions 1 and 2) – we describe the data collection and treatment process to assure our approach is reproducible.

Definition 1 (*Emotional State*). An emotional state $\lambda \in \Lambda$ is defined as a *n*-tuple $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$, where each element represents a dimension of the considered emotional theory representing the emotional state.

In our case $\lambda = (A, V)$, where:

- A is the observed value in the arousal space, with $A \in \mathbb{R} \mid A \in [-1, 1]$
- V is the observed value in the valence space, with $V \in \mathbb{R} \mid V \in [-1, 1]$.

Definition 2 (*Emotional State Waveform*). The emotional state waveform W is defined as the tuple $W = (w_A, w_V)$, where w_A and w_V are the continuous, uniformly sampled emotional state classifications for arousal and valence, respectively and are of the form $w_x=[w_{x1}, w_{x2}, \ldots, w_{xn}]$.

Emotional Reaction Extraction

The processed physiological data produced two emotional state waveforms w_A and w_V with a 1:1 mapping for each study participant over a wide emotional spectrum (see Fig. 3). The waveforms, along with the game's event log metadata, were then synchronized and used to extract player's emotional reactions in a '*through-to-peak*' fashion.

Definition 3 (*Emotional Reaction*). Consider a specific instance of a game event g_i and its corresponding timestamp g_i^t on the emotional state waveform w. Consider also the time interval T associated with this specific event g_i, such that $T_g = [max(g_{i-1}^t, g_i^t\text{-}\alpha), min(g_i^t+\beta, g_{i+1}^t)]$, where both α and β are parameterisable event horizon variables (in this paper α=2 and β=8 to account for a baseline value prior to the game event and the delays in physiological responses (Stern, Ray, & Quigley, 2001) and event logging). We define an emotional reaction as game event g_i and the pair of local maxima or minima (m, m') of each emotional state dimension, taken from the prior $T_m = [max(g_{i-1}^t, g_i^t\text{-}\alpha), g_i^t]$ and subsequent $T_{m'} = [g_i^t, min(g_i^t+\beta, g_{i+1}^t)]$ time intervals that exhibit the highest distance between them. Both m and m' are extracted from a set of candidate peaks M such that:

$$M = \left\{ \begin{matrix} t \in T\colon W(t) = \frac{d_W(t)}{dt} = 0, \\ |W(t) - W(g_i^t)| \geq \varphi \end{matrix} \right\} \qquad \text{(Eq. 1)}$$

Where φ is a minimum absolute local variability threshold, such that $\varphi = (\mu_{e_i} + 2\sigma_{e_i})$, with μ_{e_i} and σ_{e_i} denoting the mean and standard deviation values of the considered AV dimension in the considered time interval, respectively. The maximum 10-second window imposed on T by α and β was specifically designed for this particular study by having in mind: *a)* the response delays of the physiological data used in the emotional classification method (up to 5 seconds for SC), *b)* the time the stimuli usually takes to be perceived – between 1 to 2 seconds due to the lag between the game's telemetry system logging the event and the time it was actually triggered *in-game*, and *c)* the time the emotional response may take to fully manifest itself – in average approximately 1 second, from empirical analysis.

Thus, an emotional reaction $\mathscr{r}$ is defined by the triggering game event instance g_i, the prior emotional state λ_p and the response emotional state λ_r formed by the conjunction of the local maxima or minima pairs (m, m') of each emotional state dimension: $\mathscr{r} = (\lambda_p, g_i, \lambda_r)$. An illustrative example of several emotional reactions identified on a participant's emotional state waveform is presented in Fig. 2.

Unfortunately, two subjects didn't have their reactions recorded due to hardware failures. A third one isn't present due to a logging system malfunction. Furthermore, events that occurred very sparsely (e.g. dying) or registered no emotional reactions were not considered as their inclusion would artificially inflate the results. After this filtering process, the used dataset registered over 1,160 emotional reactions, over 12 of the original 15 gameplay events.

5 Affective Player Models

As we have previously discussed, our aim is to model players' individual emotional response functions to a set of game events, given a prior emotional state. More formally, these models should obey the generic function $\boldsymbol{\Phi}$ (1):

$$\Phi : \Lambda X G \rightarrow \vec{v} \mid \sum_{l=1}^{len(\Lambda)} v_l = 1 \ \wedge \ v_l \in [0,1] \qquad \text{(Eq. 2)}$$

Where function Φ receives a prior emotional state λ_p and a game event g, and outputs a weight vector $\vec{v}$ that contains the probabilities of observing a transition to each of the possible emotional states Λ ($|\Lambda| = |\vec{v}|$), if the considered game event g is performed at the prior emotional state λ_p.

In this section we describe how the extracted reactions are used to approximate players' emotional reaction functions to game events. We start by defining this approximation process (Definition 4) and then describe how the created models can be used to cluster player pairs through a bootstrapped hierarchical clustering approach.

Approximated Player Models

Definition 4 (*Approximated Player Model*). Let $\mathcal{R} = (\mathscr{r}_1, \mathscr{r}_2, \ldots, \mathscr{r}_n)$ be the set of emotional reactions extracted for a single player p and $\mathcal{R}_g = (\mathscr{r}_1^g, \mathscr{r}_2^g, \ldots, \mathscr{r}_k^g)$ be the sub-set of emotional reactions to a particular game event g, such that $\mathcal{R}_g \subset \mathcal{R}$. An approximated player model (APM) can be defined as the approximation (i.e. conditional expectation) function of the independent variables - the player's response emotional state λ_r - in regards to the dependent variables - the player's prior emotional state λ_p and triggering game event g. Given that we logically assume players' reactions to be independent to game events (i.e. each game event influences the player differently), game events are regressed separately, which also reduces the overall model complexity and training time. Moreover, since no assumptions can be made on the form of the players' underlying emotional reaction function, as no theoretical framework exists, a non-linear model, as in previous works would seem advisable (Martinez et al., 2011; Shaker et al., 2010). Thus, the player's reaction to a specific game event can be modelled as a multivariate non-linear regression function based on his emotional state prior to the triggering game event (3):

$$\begin{aligned} Y = \beta_0 + \beta_{11}X_1 + \cdots + \beta_{1n}X_1^n + \cdots + \\ \beta_{p1}X_p + \ldots + \beta_{pn}X_p^n + \varepsilon_i \end{aligned} \qquad \text{(Eq. 3)}$$

Where $X = (X_1, \ldots, X_p)$ denote the dimensions of the prior emotional state, n denotes the regression's polynomial order (in this paper p=2 and n=3) and Y represents the predicted value for the dimension of the response emotional state being modelled. A player's APM is thus defined as a set of k=nm regression surfaces that represent his emotional reaction functions to each of the individual n game events over the m dimensions of the AV space (4). In this paper, n=15 and m=2.

$$APM = M_{n,m} = \begin{bmatrix} Y_{11} & \cdots & Y_{1m} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{nm} \end{bmatrix} \qquad \text{(Eq. 4)}$$

Due to their high expressiveness, polynomial regression models are known to easily *overfit* the available data (albeit to a much lesser degree than other popular machine learning techniques such as neural networks). Thus, upon an initial analysis of our dataset, we decided to adopt a supervised stepwise regression scheme capped at third order polynomials. Another common issue with machine learning techniques is the rapidly increasing degree of uncertainty when extrapolating (i.e. the error involved in making predictions outside of the training dataset's value range quickly rises). To counter this issue we applied a tapering function to each built model, restricting it from predicting values outside $\mu \pm 2\sigma$ of the dependent variable. This prevents the model from predicting illogical response values due to simple extrapolation errors.

Distance Matrices

In order to cluster player's according to the similarity of their emotional reactions, we require a inter APM distance metric (Definition 5).

Definition 5 (*Inter APM Distance*). Let M and M' be two approximated player models, whose distance is given by:

$$\delta(M, M') = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} d(M_{ij}, M'_{ij})}{|M|} \qquad \text{(Eq. 5)}$$

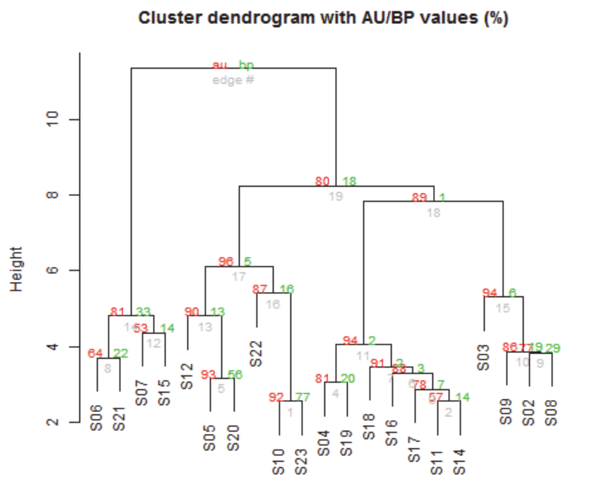

Figure 4. Bootstrapped hierarchical clusters. AU p-values are shown in red, BP values in green.

Where $d(Y,Y')$ represents the distance between two regression surfaces, which, in turn, is given by:

$$d(Y,Y') = \frac{\sum_{i=1}^{b\Delta^{-1}} \sum_{j=1}^{b\Delta^{-1}} \left| e^{Y(i\Delta, j\Delta) - Y'(i\Delta, j\Delta)} \right|}{|b\Delta^{-1}|^2} \qquad \text{(Eq. 6)}$$

Where Δ is the sampling granularity of the continuous regression surface generated by the model Y ($\Delta = 0.1$ in this paper). Although various metrics can be employed in this step (e.g. Euclidean, Manhattan), not to mention numerous transforms (e.g. sigmoid, logarithmic), an exponential function allowed us to easily differentiate between increasingly dissimilar models in a non-linear fashion appropriate for the clustering approach.

Put differently, the inter APM distance represents the average similarity of two players across equivalent game events and emotional response dimensions. Notice that since we have no supporting evidence that any game event poses a higher influence on the game's affective experience, all elements of the APM matrices were equally weighted (5).

Upon computing the inter APM distances for each approximated player model pair, they are organized into a distance matrix that is fed to the clustering algorithm.

Bootstrapped Hierarchical Clusters

To find player clusters, we applied a hierarchical clustering algorithm to the distance matrix obtained from computing the inter APM distance for each player pair. In order to obtain an initial estimate on the clusters' significance, we applied a multi-scale bootstrap resampling process, which allowed us to obtain Approximately Unbiased (AU) *p*-values for each identified cluster (see Fig. 4) (Suzuki & Shimodaira, 2006). These values represent the confidence that a particular cluster is supported by the data, as opposed to a random sampling error effect. Formally, a cluster with an AU *p*-value of *x* the null-hypothesis "*the cluster does not exist*" is rejected with a significance level $s = 100 - x$. As such, high AU values lead us to the belief that the cluster would be stably observed with an increasing number of observations.

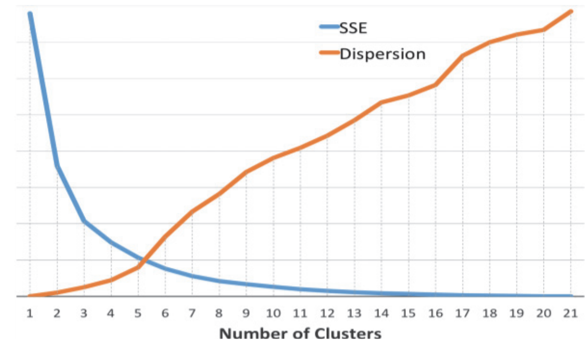

Figure 5. SSE and dispersion values for increasing cluster counts.

Cluster Analysis & Validation

While the AU *p*-values provided by the bootstrap resampling process indicate that the created clusters are well supported by the data (see Fig. 4), they do not offer a tangible proposition as to the optimal number of clusters.

Common approaches are to cut the tree at the largest links or to manually analyse the created clusters to determine the most relevant ones (Etheredge et al., 2013; Holmgård, Togelius, & Yannakakis, 2013). However, more objective and less domain knowledge dependant techniques exist. Two of the most widely accepted ones are the Sum of Squared Error (SSE) and Dispersion coefficients (7-8), which measure the within cluster cohesion and between cluster separation, respectively.

$$SSE = \sum_i \sum_{\in Ci} d(M, Ci)^2 \qquad \text{(Eq. 7)}$$

$$Separation = \sum_i |Ci| d(C, Ci)^2 \qquad \text{(Eq. 8)}$$

Since cluster cohesion (inverse SSE) and dispersion naturally increase as more clusters are created, a stopping criterion is required. Popular choices include zero-crossings and inflexion points, as these denote critical points where further splitting the data results in diminishing returns. We chose the inflexion points as both curves had noticeable inflexion points at the same cluster number (Fig. 5).

The created clusters were also manually cross-matched with the demographics data reported by the players; genre preference (whether players liked horror games or not), gamer type (casual or hardcore) and sex. Results were encouraging and showed clear divergences between clusters. For example, we found that *C1* contained only male hardcore players, while *C4* was made up of mainly softcore players that disliked horror games. However, this does not mean that, for example, *C1* contains all male hardcore players, which suggests the collected demographics are able to roughly outline the player characteristics but additional information would be necessary to characterize the found player groups.

6 Extrapolating Individual Player Models through Fuzzy Memberships

While the identified player clusters allow us to examine how particular player types would react under any given game state, they also remove the possibility of individually predicting players' reactions. Using players' approximated models instead could easily solve this issue, but doing so would not allow the reuse of the hierarchical model for new players. To address this issue, we model each player according to a fuzzy membership vector, expressed through his relative distance to each cluster (see Definition 6). This is a common approach in soft clustering methods (e.g. fuzzy c-means) to allow a certain degree of uncertainty when making predictions and also to differentiate between members of the same cluster. Since we wanted our model to be capable of representing unseen players without rebuilding the cluster models, we decided to compute the membership function outside the clustering process. Figure 6 presents the obtained player-cluster membership vectors.

Definition 6 (*Fuzzy Player Model*). Let $C=\{C_1,C_2,...,C_n\}$ be the set of clusters as identified by the hierarchical clustering algorithm. Also, consider a cluster C to be represented by the average of the APM models $\{M_1,M_2,...,M_n\}$ associated with it. A player with an APM M belongs to each cluster C in inverse proportion to his distance to the cluster model (9) after being normalised over all existing clusters (10):

$$I(M, C_i) = \frac{\sum_{i=1}^{|C|} d(M, C_i)}{d(M, C)} \quad \text{(Eq. 9)}$$

$$\rho_i = \frac{I(M, C_i)}{\sum_{j=1}^{|C|} I(M, C_i)} \quad \text{(Eq. 10)}$$

Thus, the final player model P can be expressed in terms of its fuzzy cluster membership tor $P = [\rho_1, \rho_2, \dots, \rho_n] \mid n = |C|$ that determines the weight each individual cluster response should be given when predicting the player's emotional response.

7 Discussion

The results presented by the clustering approach reveal that besides the selected clusters, most other possible clusters are heavily supported by the data (see AU/BP p-values in Fig. 6). Besides indicating that we were able to successfully differentiate between players based on their emotional reaction functions, it also suggests that re-interpreting the

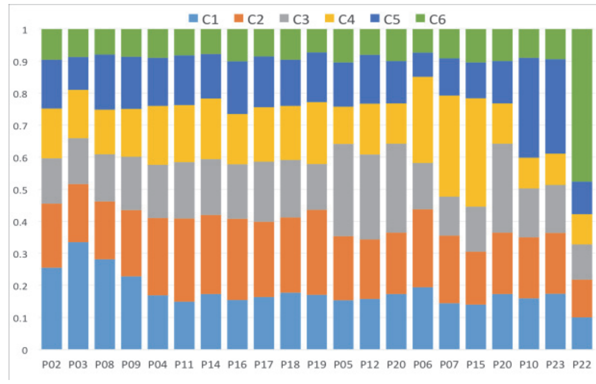

Figure 6. Individual player membership vectors. Columns represents players and shaded areas the individual cluster membership probability for clusters C1 to C6.

clustering structure at different levels might lead to clusters possibly representing very dissimilar player groups.

Game designers can, for example, examine which clusters are highly significant and then inspect the models to see how large chunks of the gamer population are reacting to each game event and under which emotional conditions.

Another potential application would be to use the created models to simulate how hypothetical changes to the game's design (both at a fundamental and punctual levels) would impact each player/player group's affective experience. This would require at least a high-level simulator for the game on which the players' behaviour could be replicated but could pose an invaluable asset for accelerating the game design and testing process. We are pursuing this goal also in order to validate the effectiveness of this type of approach on affective player modelling (Nogueira, Rodrigues, Oliveira, & Nacke 2013a; Nogueira 2013).

While we use psychophysiological related data to infer players' emotional states, our method is not dependent on it and can be easily adapted to a varying number of emotional theories, using dissimilar dimensions. Despite further validation still being required, given that our method models player reactions to game events individually, it should scale well to other game genres with dozens or even hundreds of events (e.g. MMORPGs).

Due to the added robustness offered by the clustering method and our fuzzy membership approach, our method should be able to both accurately estimate new players' cluster memberships and improve the clustering structure. Besides this, the relatively low computational cost involved in building these models also means that recomputing them on-the-fly is feasible as more data becomes available. This means that they are able to converge towards players' current true affective reaction models as they become used to (and possibly less emotionally influenced) by their continued exposure to the game.

References

Ermi, L., & Mäyrä, F. 2005. Fundamental components of the gameplay experience: Analysing immersion. In *Digital Games Research Association Conference: Changing Views - Worlds in Play*.

Etheredge, M., Lopes, R., & Bidarra, R. 2013. A Generic Method for Classification of Player Behavior. In *IDPv2 2013 - Workshop on Artificial Intelligence in the Game Design Process*.

Gilleade, K. M., Dix, A., & Allanson, J. 2005. Affective Videogames and Modes of Affective Gaming : Assist Me , Challenge Me , Emote Me. In *DIGRA - Digital Games Research Association*, 1–7.

Gilroy, S., & Porteous, J. 2012. Exploring passive user interaction for adaptive narratives. *Intelligent User Interaction*, 119–128.

Holmgård, C., Togelius, J., & Yannakakis, G. 2013. Decision Making Styles as Deviation from Rational Action A Super Mario Case Study. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

Lankes, M., Hochleitner, W., Hochleitner, C., & Lehner, N. 2012. Control vs. complexity in games: comparing arousal in 2D game prototypes. In *Proceedings of the 4th International Conference on Fun and Games,* 101–104.

Mandryk, R., & Atkins, M. 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, *65*(4), 329–347. doi:10.1016/j.ijhcs.2006.11.011

Martinez, H. P., Garbarino, M., & Yannakakis, G. N. 2011. Generic Physiological Features as Predictors of Player Experience. In *Proceedings of the 2011 Affective Computing and Intelligent Interaction Conference*, 267–276.

Nogueira, P. A. 2013. Towards Dynamically Crafted Affective Experiences Through Emotional Response Player Modelling. In *Proceedings of the Ninth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*.

Nogueira, P. A., Rodrigues, R., Oliveira, E., & Nacke., L. E. 2013a. Guided Emotional State Regulation: Understanding and Shaping Players' Affective Experiences in Digital Games. In *Proceedings of the Ninth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, 51–58.

Nogueira, P. A., Rodrigues, R., Oliveira, E., & Nacke, L. E. 2013b. A Hybrid Approach at Emotional State Detection: Merging Theoretical Models of Emotion with Data-Driven Statistical Classifiers. In *Proceedings of the 2013 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 253 – 260.

Russel, J. A. 1980. A Circumplex Model of Affect. *Personality and Social Psychology*, *39*(6), 1161–1178.

Ryan, R., Rigby, C., & Przybylski, A. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, *30*(4), 344–360.

Shaker, N., Yannakakis, G., & Togelius, J. 2010. Towards automatic personalized content generation for platform games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 63–68.

Stern, R. M., Ray, W. J., & Quigley, K. S. 2001. *Psychophysiological recording* (2nd ed.). New York: Oxford University Press.

Suzuki, R., & Shimodaira, H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, *22*(12), 1540–1542.

Tognetti, S., Garbarino, M., Bonarini, A., & Matteucci, M. 2010. Modeling Enjoyment Preference from Physiological Responses in a Car Racing Game. In *IEEE Conference on Computational Intelligence and Games*, 321–328.

Vachiratamporn, V., Legaspi, R., Moriyama, K., & Numao, M. 2013. Towards the Design of Affective Survival Horror Games: An Investigation on Player Affect. In *Affective Computing and Intelligent Interaction (ACII)*, 576–581.

Yannakakis, G. N., & Togelius, J. 2011. Experience-driven Procedural Content Generation. *Transactions on Affective Computing*, *2*(3), 147–161.