Towards Understanding the Nonverbal Signatures of Engagement in Super Mario Bros

Noor Shaker[1] and Mohammad Shaker[2]

[1]IT University of Copenhagen, Rued Langaards Vej 7, 2300 Copenhagen, Denmark,
[2]Independent Researcher, Damascus, Syria
nosh@itu.dk,mohammadshakergtr@gmail.com

Abstract. In this paper, we present an approach for predicting users' level of engagement from nonverbal cues within a game environment. We use a data corpus collected from 28 participants (152 minutes of video recording) playing the popular platform game *Super Mario Bros*. The richness of the corpus allows extraction of several visual and facial expression features that were utilised as indicators of players' affects as captured by players' self-reports. Neuroevolution preference learning is used to construct accurate models of player experience that approximate the relationship between extracted features and reported engagement. The method is supported by a feature selection technique for choosing the relevant subset of features. Different setup settings were implemented to analyse the impact of the type of the features and the position of the extraction window on the modelling accuracy. The results obtained show that highly accurate models can be constructed (with accuracies up to 96.82%) and that players' nonverbal behaviour towards the end of the game is the most correlated with engagement. The framework presented is part of a bigger picture where the generated models are utilised to tailor content generation to a player's particular needs and playing characteristics.

Keywords: Player experience modelling, affect recognition, nonverbal behaviour, affect modelling, facial expressions, neuroevolution preference learning.

1 Introduction

Analysing users' interaction with a digital interface has been the focus of many research studies. One of the main motivations behind investing in this area is to keep the user engaged during the interaction. This direction has drawn significant attention from both the Affective Computing community and Human-Computer Interaction. Motivated by the current trend towards understanding and modelling users' interaction and ultimately generating immersive experience, this paper investigates the relationship between players' nonverbal behaviour during an in-game interaction and their experienced affective states elicited as a result of this interaction. We envision a game where the proposed approach can be employed to effectively capture players' affects as they interact with a game and efficiently use this information to alter the content according to a specific player's needs consequently providing a more engaging experience.

In games, such as in many other applications, the engagement value of content is crucial to its success. While players' skill and expertise vary over time, and in order to accommodate for the differences between players in terms of playing characteristics, behaviour and preferences, the game should be able to efficiently detect these differences and automatically adjust its content so that a personalised engaging experience can be created.

In this study we investigate the effectiveness of different visual and facial expression features as indicators of player's level of engagement while playing the popular 2D platform game *Super Mario Bros.* A dataset of 28 subjects was collected for this study and experiments were conducted with several settings to analyse the impact of the feature type, feature extraction and selection on the classification accuracy. Models of player experience are constructed as classifiers and neuroevolutionary preference learning is implemented as a modelling approach.

The main findings can be summarised as follows: (1) players' nonverbal behaviour carries rich information about the interaction experience and consequently they can be used to accurately predict the engagement value of content, (2) the importance of visual cues differs along the game session; players' visual reactions towards the end of the game is found to be the most relevant for predicting engagement.

This paper is structured as follows: section 2 summarises the previous research efforts; section 3 provides detail information about the dataset collected and the tool used; section 4 describes the modelling methodology followed to relate features extracted from players' visual behaviour to reported engagement; section 5 illustrates the different settings for feature extraction and the experiments conducted to construct models of player experience; finally, the experimental results are presented and discussed in section 6.

2 Related Work

There are many studies in the literature dealing with the problem of user state estimation during digital interaction. Recent research in computer vision techniques has discussed a number of methods incorporating notions such as body and head movements [2], eye gaze [10] and facial expressions [9] as indicators of user affective states with varying accuracies depending on the settings, application and the modality used. Furthermore, several psycho-physiological signals such as heart rate and skin conductance are found to correlate with affective states in response to visual stimuli [25, 14].

Recently, there has been an increasing attention given to detect users' emotional states while interacting with a game. Players' performance captured through recorded actions and interactions with game events is usually utilised as an indicator of player's affect [18, 20, 5]. Physiological signals are used in some studies to investigate the relationship between this modality, player behaviour and affects [16, 26]. Measuring affect using physiological signals usually requires specialised hardware, which is often expensive and hard to calibrate. As a result, related approaches may be efficient in terms of recognising player affect, but are extremely problematic to deploy in mass scales and for commercial uses. On the other hand, affect estimation approaches based on video sequences taken from low-end cameras utilise hardware that most gamers already possess and do not impose any additional requirements.

The use of verbal and nonverbal affective communication has led to advances in realising real-time multi-modal affective interactions between agents and learners in Intelligent Tutoring System and Learning Companion technologies [4, 11]. In games, facial expression has been used for investigating affects with promising results [6, 8]. Features extracted from multiple modalities, including players' actions and visual reactions, have been investigated [22, 24, 1].

The increased diversification of users demographics, needs, skills and preferences has increased the importance of experience personalisation. In the domain of games, player experience modelling [27] studies that rely on single or multiple modalities of user input (see [23, 26, 15, 11] among many) have provided some initial benchmark solutions towards achieving such a goal. The use of computational models of engagement, among other affective states, has been investigated in several studies to predict the appeal of specific piece of content to a particular player [16, 26, 22, 20] .

In this paper, we take a further step in this direction by investigating nonverbal cues extracted through different settings. We build on our previous attempts on modelling player experience [22, 24] and we extend it through: (1) investigating new types of visual and facial expression features (while in [22] we experimented with different attributes of the head movement such as the average, fluidity and energy, in this paper we focus on the bias of the head in the x- and y- coordinate, we extract new features related to the eye and the mouth and we utilise several facial expression states); (2) experimenting with different settings for feature extraction and (3) analysing the importance of the portion of the session segment from which the best features can be extracted for predicting engagement.

3 Experimental Set-up and Protocol

In the following sections we present the procedure followed to collect the dataset and the tool used to process the data gathered.

3.1 Material and setup

In order to study the relationship between users' interaction within a game environment and their affective states, we conducted a survey experiment where users were asked to play a clone of the popular game *Super Mario Bros* while interaction and self-assessments of a number of affective states were gathered.

The gameplay in our testbed game constitutes of moving the main character, *Mario*, through a 2D environment while collecting rewarding items, killing enemies, and avoiding obstacles. The character can achieve these objectives by doing simple actions such as running, jumping, shooting and ducking. The main goal of the game is to reach the end of the level with minimum lose of lives and maximum number of collected rewards.

The study was conducted in a lab environment in Denmark and Greece. Data was collected from users who played individually in a room where lighting conditions were typical of an office environment and interruption was kept to minimum. The system is composed of a laptop where the software was installed, equipped with a High Definition camera (Canon Legria S11) for video recording. A detailed description of the dataset used can be found in [12, 21].

3.2 Study Procedure

For the purpose of this study, data from 28 participants (21 male; player age varied from 22 to 48 years) was recorded. Subjects are seated in front of a computer screen for video recording. Each session consists of playing at least a pair of two games followed by a post-experience game survey design to collect subject self-reports of *engagement*. Self-assessments were collected using the 4-Alternative Forced Choice (4-AFC) questionnaire protocol proposed in [29]. According to this protocol, after completing a pair of two games A and B, the subject is asked to report the preferred game given four options: A is preferred to B; B is preferred to A; both are preferred equally; neither is preferred (both are equally not-preferred). Each participant played three pairs on average resulting in a total of 66 valid pairs and 152 minutes of gameplay. For detailed information about the experimental protocol followed, the reader may refer to [21, 12].

3.3 Tools: SHORE Engine

SHORE[1] (Sophisticated High speed Object Recognition Engine) [13] is a recognition engine that enables the detection of objects and faces. SHORE enables feature extraction from both images and videos. The list of features that can be extracted by SHORE includes the position of the face, eyes, nose and mouth, information about whether the eyes or the mouth are open or closed and recognition of facial expressions (the set of recognised expressions includes: happiness, sadness, anger and surprised). All of the above mentioned features can be extracted in realtime. SHORE engine processes a video using 5 frames per seconds. We use the C++ interface of the engine. All the videos recorded were post-processed with the engine to extract visual and facial expression features as will be described in the following sections.

[1] http://www.iis.fraunhofer.de/de/bf/bsy/fue/isyst.html

4 Method

The first step towards creating a more engaging experience is to effectively recognise whether the player is enjoying an piece of content. One way to achieve this goal is to construct models of player experience derived from the in-game interaction where content is annotated with user experience tags. Several machine learning techniques have been utilised to build such models including Support Vector Machine and Bayesian Inference [30]. In a comparison study, Neuroevolutionary Preference Learning (NPL)) [7, 30] showed capability of constructing models with accuracies surpass those achieved by the other methods in a similar settings to the one at hand [30]. Therefore in this paper, NPL is used to approximate the function between nonverbal features and reported affective preferences.

In the following sections, we describe the procedure followed to build the player experience models using NPL.

4.1 Feature Extraction

The first step towards understanding the relationship between reported affects and players' nonverbal reactions while interacting with a game is to extract representative features that can be utilised as indicators of players' behaviour in reaction to game events. This work is motivated by previous attempts where expressive nonverbal behaviour is linked to user engagement; namely eye gaze and smiles were found to be correlated with user engagement and interest while interaction with a robot [4, 19].

In this study, we track head location and changes in the properties of the eyes and the mouth. More specifically, we calculate the averages and standard deviations of the followings: head position on the $x-$ and $y-$axis, the closeness of the left and right eyes and the openness degree of the mouth (the full list of features and their description are presented in Table 1). We call this set of nonverbal features *visual reaction* features (VR). The use of these features is inspired by earlier work where visual features were found to be related to arousal [3].

Along with VR features, we employ another set of *Facial expression* (FE) features. This set includes the average and standard deviation degrees of the following states: happiness, sadness, anger, and surprise.

When constructing the Player Experience Models (PEMs) we hope that these features will correlate with reported affects, i.e. we aim at exploring content-behaviour relationships that can be utilised to predict players' emotional state of engagement from his/her visual cues. For example, the player might experience a moderate degree of happiness and a steady head movement when the game starts, these states, however, might change during the course of gameplay and a state of surprise combined with a sudden movement of the head might arise as the player encounters an enemy. A game that comprises these two states, among some others, can then be annotated as being highly engaging by the PEMs.

A total of 18 features were extracted (averages and standard deviations of all the above mentioned features) as can be seen in Table 1. We experimented with different settings of the frames from which these features are extracted as will be seen in Section 5.

Table 1. Visual reaction and facial expression features extracted from the data recorded. Average and standard deviation values are extracted for each feature.

Feature	Description
Visual reaction	
h_x	Bias of head location on the x-axis compared to the location of the head in the first frame
h_y	Bias of head location on the y-axis compared to the location of the head in the first frame
$leftEye_c$	Closeness degree of the left eye
$rightEye_c$	Closeness degree of the right eye
$mouth_o$	Openness degree of the mouth
Facial expression	
H	Degree of happiness
S	Degree of sadness
A	Degree of anger
U	Degree of surprise

4.2 Player Experience Modelling

As mentioned earlier, neuroevolutionary preference learning is chosen as a modelling approach in our study. In NPL, a genetic algorithm evolves an Artificial Neural Network (ANN) so that its output matches the pairwise preferences in the dataset. The input of the ANN is a set of extracted features.

As a preprocessing step, all features extracted are uniformly normalised to [0,1] using standard max-min normalisation. These features are then used as inputs for feature selection and ANN model optimisation. The steps followed to construct the models can be summarised as follows:

- Feature selection: We use Sequential Forward Selection (SFS) [28] to select the relevant subset of features for predicting reported engagement [30]. The simplest form of ANN, that is a single-layer perception was implemented to achieve this.
- Feature space expansion: The feature subset derived from the first phase is used as the input to small multi-layer perceptron (MLP) models of one two-neuron hidden layer and SFS selects additional features from the remaining set of features based on their classification accuracy.
- Optimising topology: In the last phase, the topology of the MLP models is optimised. This process starts with a small two hidden-neuron MLP and the network topology gradually increases up to two hidden layers consisting of

10 hidden neurons each. The network that achieves the best accuracy is then chosen as the final model.

The quality of a feature subset and the performance of each MLP is obtained through the average classification accuracy in three independent runs using 3-fold cross validation across ten evolutionary trials. The data was partitioned into folds such that the likelihood of a subject appearing in more than one fold is minimised. We use a population of 100 individuals and we run evolution for 20 generations. A probabilistic rank-based selection scheme is used, with higher ranked individuals having higher probability of being chosen as parents. Finally, reproduction is performed via uniform crossover, followed by Gaussian mutation of 1% probability. All of these parameters were chosen as a result of a tuning process.

5 Experiments

In order to analyse the relationship between the extracted features and players' reported engagement and to investigate the significance of the type of features and the importance of context information, we construct models on three different settings for feature extraction as follows (a summary can be seen in Table 2):

- VR_{all}: The features used as input to these models are VR features extracted from the full sessions, i.e. the average and standard deviation values of all features presented in Table 1 are calculated from all frames for each game played (note that each game session consists of a maximum of three trails). This setting is considered in order to minimise the effect of habituation where the player becomes more familiar with the game over time.
- VR_{events}: As players' expressivity appears to increase during certain events, we considered features extracted during certain gameplay events as described below:
  - When the player loses a life.
  - When the player kills an enemy by stomping on it.
  - When the player starts or ends a critical move: jump, duck, run, and move left or right.
  - When the player interacts with a game object.

  These features are calculated for a period of 3 frames before and after the corresponding events.
- All_{wind}: In this category, VR and FE feature values are calculated from the players' last attempt only in an attempt to minimise the habituation effect. Similar to VR_{events}, the values are considered only when specific game events occur. We also introduced a windowing factor: the last trail is partitioned into three equal-sized windows and the values of all features are calculated from each window separately resulting in a total of $18 * 3 = 54$ features. Three separated models are then constructed for each window permitting investigation of the part of the session that carries the most

significant information for predicting reported affective state of engagement. More specifically, the models constructed from information extracted from the first window encompass details about players' behaviour at the beginning of the game and the importance of the initial behaviour and content on the experienced engagement. The significance of the outcome of the game on the level of engagement felt, however, can be captured through the models constructed from the third window.

Table 2. The settings for the three types of models constructed.

	Feature extraction frames	Game session	Input features
VR_{all}	All frames	Three trails	VR
VR_{events}	Three frames before and three after specific gameplay events	Three trails	VR
All_{wind}	Three frames before and three after specific gameplay events	Last trail	VR + FE

6 Results and Analysis

Several models are constructed from the different types of features and the various settings of feature extraction. For each feature input space, models with different subset of features as selected by the feature selection methods are constructed. The models also differ in their complexity as depicted by the various network topologies obtained for the best performing models. Table 3 presents the different models constructed. Each column stands for a model build from a specific set of features (VR, FE or both) as input. The features selected using SLPs and the corresponding best and average performance over 50 runs are presented for each model (row 2-4). The models are further improved by selecting features using simple MLP (rows 5-7). Finally, using the selected subset of features, the models were optimised for best prediction accuracy and the obtained topologies (row 8) and performance are presented (row 9-10). In general, models of very high accuracies are obtained with a performance up to 96%.

As can be seen in Table 3. The best accuracy obtained is the one from the models constructed from the last window in the session using both VR and FE as input features. It is worth noticing however that only VR features are selected with no FE features. These models have a moderate network size consisting of one hidden layer of eight neurons. An interesting observation is that all the features are selected through a single-layer perceptron indicating a simple relationship between the features and the experienced level of engagement. Three features only out of the 18 extracted features are selected, two of which are related to the closeness of the eyes and the third captures the shift of the head along the x-axis. This can be understood as an implicit indication of the importance of the outcome of the game (whether it is winning or losing) which appears to be reflected in players' visual reactions as closing the eyes and posing the head. The results align with our observation of the video recordings which showed a common pattern of closed eyes and head poses when winning and losing. Example instances of such cases are presented in Figure 1.

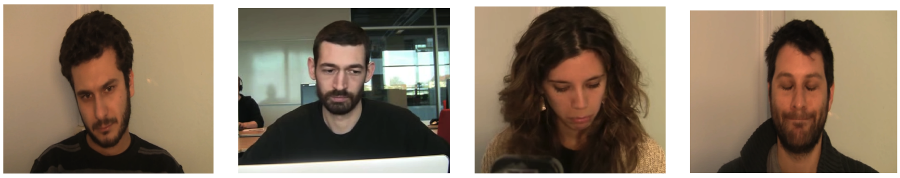
Players' FE appears to be an important predictor of engagement in the middle of the game. Specifically, the level of recognised happiness and sadness are found to be correlated with experienced engagement.

High prediction accuracies are also obtained when constructing models from VR features only as inputs, both when calculated during specific game events and when all frames are considered. More features and more complex topologies, however, are observed in the latter case.

The worst accuracies obtained are the ones for the models constructed from features from the first window in the sessions. This implies that players' visual reactions at the beginning of the game is trivial for predicting the level of engagement felt compared to their visual behaviour towards the end of the game.

Table 3. Features selected from the set of visual reaction and facial expression features for predicting engagement. The table also presents the corresponding average and best performance values obtained from the ANN models' and the best models' ANN topologies. The ANN topologies are presented in the form: number of neurons in the first hidden layer−number of neurons in the second hidden layer.

	VR_{all}	VR_{events}	All_{wind}		
			1^{st} window	2^{nd} window	3^{rd} window
SLP features	$mouth_o\_avg$ re_c_avg h_y_std	$mouth_o\_avg$ $rightEye_c\_avg$	h_x_avg h_y_avg	$leftEye_c\_std$ H_avg S_std H_std	h_x_avg $rightEye_c\_avg$ $leftEye_c\_avg$
$SLP_{perf}\%$	65.77	64.44	41.21	68.78	70.63
$SLP_{max}\%$	69.69	71.42	48.48	80.30	77.77
MLP features	$rightEye_c\_std$ $leftEye_c\_avg$ h_x_avg	$mouth_o\_std$	S_avg	-	-
$MLP_{perf}\%$	82.12	84.28	60.90	-	-
$MLP_{max}\%$	92.42	92.06	66.66	-	-
ANN Topology	8_2	2_6	6_0	4_0	8_0
$MLP_{opt}\%$	81.57	83.14	62	84.03	86.09
$MLP_{opt\_max}\%$	92.42	93.65	72.72	93.93	96.82



(a) (b) (c) (d)

Fig. 1. Example instances from the video recordings of subjects playing the testbed games. (a) and (b) show head pose and eyes condition when winning the game while (c) and (d) corresponds to visual reactions when losing.

7 Conclusions and Future Directions

In this paper we empirically investigated whether users' visual reactions and facial expressions can be efficiently utilised as indicators of engagement when playing a computer game. To facilitate such analysis, a large dataset of video recordings from 58 players playing the 2D game *Super Mario Bros* was collected. The dataset consists of gameplay sessions annotated with players' self-assessment of engagement. Eighteen representative visual reaction and facial expression features are then extracted from the gameplay session for each subject. These features form the input for classifiers that are trained to predict reported engagement from the extracted features. The experiments conducted included several settings for feature extraction, the type of features and the position of the extraction window. The results obtained show that models of very high accuracies (around 96%) can be constructed using visual reaction features. The results also highlight the importance of some of the features compared to others. When studying the best segment for feature extraction that yields the highest accuracy, the results show that the most accurate models can be built from visual reaction features extracted from the last portion of the game. This finding signifies the importance of the content presented towards the end of the game and indicates that the level of engagement felt correlates with the outcome of the game.

Our findings align with studies reported in the literature that concluded that nonverbal channels carry informative affective cues [11, 17]. However, although models of high accuracies are constructed, understanding these models is not trivial due to the complex nature of neural networks. Therefore, investigating other, more expressive, approaches such as Bayesian Inference or SVM constitutes an important future direction towards understanding the link between nonverbal behaviour and affects. Moreover, the models built are average models across all participants, therefore it is interesting to investigate whether different results could be obtained after clustering players according to their expressiveness and/or demographics .

The study presented in this paper is the first step towards generating user-adapted content. The player experience models constructed can be ultimately used to predict the appeal of a piece of content to a specific player given her visual behaviour. Based on this information, adjustments can be made to the content in real-time so that the game will become more engaging.

Another future direction is to investigate the use of fused features from multiple modalities such as features representing content and players' in-game actions. Previous attempts showed that models constructed from fused features are usually more accurate than those built from a single modality [22, 24, 16, 1].

References

1. Abadi, M.K., Staiano, J., Cappelletti, A., Zancanaro, M., Sebe, N.: Multimodal engagement classification for affective cinema. In: Humaine Association Conference on Affective Computing and Intelligent Interaction. pp. 411–416. IEEE (2013)
2. Asteriadis, S., Tzouveli, P., Karpouzis, K., Kollias, S.: Estimation of behavioral user state based on eye gaze and head poseapplication in an e-learning environment. Multimedia Tools and Applications 41(3), 469–493 (2009)
3. Caridakis, G., Castellano, G., Kessous, L., Raouzaiou, A., Malatesta, L., Asteriadis, S., Karpouzis, K.: Multimodal emotion recognition from expressive faces, body gestures and speech. In: Artificial intelligence and innovations 2007: From theory to applications, pp. 375–388. Springer (2007)
4. Castellano, G., Pereira, A., Leite, I., Paiva, A., McOwan, P.W.: Detecting user engagement with a robot companion using task and social interaction-based features. In: Proceedings of the 2009 international conference on Multimodal interfaces. pp. 119–126. ACM (2009)
5. Conati, C., Maclaren, H.: Empirically building and evaluating a probabilistic model of user affect. User Modeling and User-Adapted Interaction 19(3), 267–303 (2009)
6. D, Mello, S.K., Craig, S.D., Graesser, A.C.: Multi-method assessment of affective experience and expression during deep learning. Int. J. Learn. Technol. 4(3/4), 165–187 (2009)
7. Fürnkranz, J., Hüllermeier, E.: Pairwise preference learning and ranking. In: Machine Learning: ECML 2003, pp. 145–156. Springer (2003)
8. Grafsgaard, J.F., Boyer, K.E., Lester, J.C.: Toward a machine learning framework for understanding affective tutorial interaction. In: Intelligent Tutoring Systems. pp. 52–58. Springer (2012)
9. Ioannou, S., Caridakis, G., Karpouzis, K., Kollias, S.: Robust feature detection for facial expression recognition. Journal on Image and Video Processing (2) (2007)
10. Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. International journal of human-computer studies 66(9), 641–661 (2008)
11. Kapoor, A., Burleson, W., Picard, R.: Automatic prediction of frustration. International Journal of Human-Computer Studies (8), 724–736 (2007)
12. Karpouzis, K., Shaker, N., Yannakakis, G.N., Asteriadis, S.: The platformer experience dataset. IEEE Transactions on Computational Intelligence and AI in Games (2014)
13. Küblbeck, C., Ernst, A.: Face detection and tracking in video sequences using the modifiedcensus transformation. Image and Vision Computing 24(6), 564–572 (2006)
14. Lisetti, C.L., Nasoz, F.: Using noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP Journal on Applied Signal Processing 2004, 1672–1687 (2004)

15. Mandryk, R., Inkpen, K., Calvert, T.: Using psychophysiological techniques to measure user experience with entertainment technologies. Behaviour & Information Technology (2), 141–158 (2006)
16. Martinez, H., Yannakakis, G.: Mining multimodal sequential patterns: A case study on affect detection. In: Proceedings of the 13th International Conference in Multimodal Interaction, ICMI 2011, Alicante. ACM Press (November 2011)
17. McDaniel, B., DMello, S., King, B., Chipman, P., Tapp, K., Graesser, A.: Facial features for affective state detection in learning environments. In: Proceedings of the 29th Annual Cognitive Science Society. pp. 467–472 (2007)
18. Pedersen, C., Togelius, J., Yannakakis, G.N.: Modeling player experience for content creation. IEEE Transactions on Computational Intelligence and AI in Games 2(1), 54–67 (2010)
19. Peters, C., Asteriadis, S., Karpouzis, K., de Sevin, E.: Towards a real-time gaze-based shared attention for a virtual agent. In: Workshop on Affective Interaction in Natural Environments (AFFINE), ACM International Conference on Multimodal Interfaces (ICMI08) (2008)
20. Shaker, N., Yannakakis, G.N., Togelius, J.: Towards Automatic Personalized Content Generation for Platform Games. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE) (2010)
21. Shaker, N., Asteriadis, S., Yannakakis, G.N., Karpouzis, K.: A game-based corpus for analysing the interplay between game context and player experience. In: Affective Computing and Intelligent Interaction, pp. 547–556. Springer (2011)
22. Shaker, N., Asteriadis, S., Yannakakis, G.N., Karpouzis, K.: Fusing visual and behavioral cues for modeling user experience in games. Cybernetics, IEEE Transactions on 43(6), 1519–1531 (2013)
23. Shaker, N., Togelius, J., Yannakakis, G.N., Weber, B., Shimizu, T., Hashiyama, T., Sorenson, N., Pasquier, P., Mawhorter, P., Takahashi, G., Smith, G., Baumgarten, R.: The 2010 Mario AI championship: Level generation track. IEEE Transactions on Computational Intelligence and Games (2011)
24. Shaker, N., Yannakakis, G.N., Togelius, J.: Crowd-sourcing the aesthetics of platform games. IEEE Transactions on Computational Intelligence and AI in Games (2013)
25. Soleymani, M., Lichtenauer, J., Pun, T., Pantic, M.: A multimodal database for affect recognition and implicit tagging. Affective Computing, IEEE Transactions on 3(1), 42–55 (2012)
26. Tognetti, S., Garbarino, M., Bonanno, A.T., Matteucci, M., Bonarini, A.: Enjoyment recognition from physiological data in a car racing game. In: Proceedings of the 3rd international workshop on Affective interaction in natural environments. pp. 3–8. ACM (2010)
27. Yannakakis, G.N., Togelius, J.: Experience-Driven Procedural Content Generation. IEEE Transactions on Affective Computing (2011)
28. Yannakakis, G.N., Hallam, J.: Entertainment modeling through physiology in physical play. Int. J. Hum.-Comput. Stud. 66, 741–755 (October 2008)
29. Yannakakis, G.N., Maragoudakis, M., Hallam, J.: Preference learning for cognitive modeling: a case study on entertainment preferences. Trans. Sys. Man Cyber. Part A 39, 1165–1175 (2009)
30. Yannakakis, G.N., Maragoudakis, M., Hallam, J.: Preference learning for cognitive modeling: a case study on entertainment preferences. Trans. Sys. Man Cyber. Part A 39, 1165–1175 (November 2009)