

A Game-based Corpus for Analysing the Interplay between Game Context and Player Experience

Noor Shaker¹, Stylianos Asteriadis², Georgios N. Yannakakis¹, and
Kostas Karpouzis²

¹ IT University of Copenhagen, Rued Langaards Vej 7, 2300 Copenhagen, Denmark

² National Technical University of Athens, 157 80 Zographou, Athens, Greece
{nosh, yannakakis}@itu.dk, {stias, karpou}@image.ntua.gr

Abstract. Recognizing players’ affective state while playing video games has been the focus of many recent research studies. In this paper we describe the process that has been followed to build a corpus based on game events and recorded video sessions from human players while playing *Super Mario Bros*. We present different types of information that have been extracted from game context, player preferences and perception of the game, as well as user features, automatically extracted from video recordings. We run a number of initial experiments to analyse players’ behavior while playing video games as a case study of the possible use of the corpus.

Keywords: game-based corpus, player behavior, player’s affective state, player experience

1 Introduction

Many theories exist regarding why we play games and what makes computer games engaging and immersive environments [4, 8, 9, 11, 2]. Studying player’s behavior and recognizing players’ affective state while playing video games has been the focus of attention of recent research. Closing the affective loop [10] is one of the ultimate aims of the research carried out in the field of affective computing [16]. Closing the affective loop within games context entails recognizing/modeling players affective state and identifying game events that elicit certain players’ behavior. The next step is the generation of player-adaptive content by incorporating players’ emotions into the game in a closed-loop manner using models of player emotion built from the interaction between the player and the game [19, 27].

Recognition of emotion, such as entertainment, is considered a hard problem mainly because of the multi-modal nature of emotion which makes sensing, recognizing and modeling emotion a hard task [6]. Traditional evaluation methods such as subjective and objective techniques have been adopted to detect affective states [13, 14, 24, 19, 20]. Although subjective reporting is a good approach for

affective capturing and modeling, there are limitations to this methodology [13]. Other input sources, including biofeedback and additional context-based game metrics, could be used for further analysis but do not supplant the self-reports.

Physiological measures have been used to assess users emotional experience when engaged with HCI systems [12, 3, 26]. Physiological data is a very powerful source for assessing affective state as it can continuously and objectively provide a quantitative metric of user experience [12]. Although physiological measurement devices have been used extensively within the affective computing research for emotion recognition [17] and despite the great efforts that have been devoted to make these devices wearable, their use in commercial computer games is still limited. Facial expressions [7] and head movements [1] are rich non intrusive sources of information regarding the issue of capturing the emotional or behavioral state of a person while interacting with a machine or undertaking certain tasks. A lot of work has been done in recent bibliography for modeling such cues in a variety of environments. For example, Asteriadis et al. in [1] investigate the issue of using head movements, as well as gaze patterns in order to achieve a mapping between these and a child's state towards learning material, aiming at the adaptation of the learning process to the child's needs. Smith et al. [22] propose a scheme for mapping gaze and eye blinks to driver state, in the context of driving conditions, utilizing Finite State Automata. Sidner et al. in [21] investigate those head movements that imply high degree of engagement with a robot utilizing head nods, gaze patterns, as well as head gestures declaring user familiarization with the environment.

The focus and main contribution of this paper is the design of a game-based corpus derived from the popular Super Mario Bros platform game. The corpus contains information about the game context, different game events, players' preferences and perception of the game, as well as visual feedback of players' faces, taken with video recordings. The design of this corpus is an important step towards a more in-depth analysis of the relationship between game context, players' behavior and players' affective state. Findings out of the presented corpus are expected to play a significant role in the game adaptation and personalization procedure. We are not aware of any other similar corpus that has been designed within games context.

Together with the presentation of the corpus, in this study, we present initial findings for correlating a set of game events, a set of features extracted from the corresponding video recordings and/or questionnaires; the results obtained here, are an indicate of the validity of our corpus and the findings that can be extracted from it. The idea of correlating human behavior with interaction context has been studied before in a variety of problems, aiming especially at interpreting attention levels of a user towards specific events. Thus, the same behavioral cue can be interpreted differently under different conditions. For example, although one person might appear to be attentive towards his or her computer screen, his real levels of engagement are low if there is nothing being exposed on it. Furthermore, a person sitting in front of a computer monitor, at the start of an interaction should be expected to be "excused" if he or she is looking around. As



Fig. 1. Snapshot from Super Mario Bros game.

the authors in [21] state, this is not necessarily a sign of indifference, but could declare a high degree of engagement, as it can be translated into the person’s curiosity to familiarize himself with the environment. Following these ideas, the presented dataset and the accompanying information aim at correlating what is actually taking place in the game with people’s reactions and their perception of game quality. The results obtained by such an analysis can be further employed for game adaptation for personalization which is expected to increase the level of engagement.

2 Testbed Platform Game

The testbed platform game used for our study is a modified version of Markus Perssons Infinite Mario Bros (see Fig. 1), which is a public domain clone of Nintendo’s classic platform game Super Mario Bros. The original Infinite Mario Bros and its source code is available on the web ³.

The gameplay in Super Mario Bros consists of moving the player-controlled character, Mario, through two dimensional levels. Mario can walk, run, duck, jump, and shoot fireballs. The main goal of each level is to get to the end of the level. Auxiliary goals include collecting as many coins as possible, and clearing the level as fast as possible.

While implementing most features of Super Mario Bros, the stand out feature of Infinite Mario Bros is the automatic generation of levels. Every time a new game is started, levels are randomly generated. In our modified version, we concentrated on a few selected parameters that affect gameplay experience.

3 Dataset Design

To assess the players’ affective state during play, the following experiment was set up: We seated 36 volunteers (28 male) in front of a computer screen for video recording. Players’ age varied from 23 to 39 years old, while experiments were

³ <http://www.mojang.com/notch/mario/>

carried out in Greece and Denmark. Lighting conditions were typical of an office environment, and for capturing players' visual behavior, a HD camera was used.

We designed a game survey study to collect subjective affective reports expressed as ranking and pairwise preferences of subjects playing different variants (levels) of the test-bed game by following the experimental protocol proposed in [24]. A detailed description of the procedure followed is described.

- An introduction scene presents the game to the player and contains information about the procedure that will be followed. The player is being told that during the session she will play two short games and will be asked to answer a few questions about her game experience.
- Then, a demographics scene is presented which is used to collect the demographics data.
- The player is introduced to the keys that can be used to control *Mario* and their functionality.
- The player is then informed that her game sessions will be video recorded and analysed.
- After these introductory steps the player is set to play the first game (game *A*). The player is given three chances to complete the short game level of Super Mario Bros. If she fails in the first trial the game is reset to the starting point and the player is set to try again. The game ends either by winning one of the three trials or by failing the third one.
- After finishing game *A* a Likert questionnaire scheme is presented to the player. The player is asked to express her emotional preferences of the played game across the three different emotional states (engagement, frustration and challenge). The questionnaire is inspired by the game experience questionnaire (GEQ) [18] according to which a likert scale from 0 to 4 represents the strength of the emotion (4 means “extremely”; 0 means “not at all”).
- A second short game (game *B*) is then presented to the player and the player is set to play. The player is given three chances (i.e. Mario lives) to complete the level and the same rules apply as in game *A*.
- After finishing game *B* the GEQ questionnaire is presented to the player (as in game *A*).
- After completing a pair of two games *A* and *B*, the player is asked to report the preferred game for the three emotional dimensions through a 4-alternative forced choice (4-AFC) questionnaire protocol (i.e. *A* is preferred to *B*, *B* is preferred to *A*, both are preferred equally, neither is preferred (both are equally unpreferred)) [23]. Please note that the questionnaire presented to the players is the following: “Which game was more *x*” where *x* is one of the three emotional states under investigation.
- The player then has the choice to either end the session or to continue. In the latter case, a new pair of two games is presented and the procedure is repeated.

Each participant played three pairs on average. The game sessions presented to players have been constructed using a level width of 100 Super Mario Bros

units (blocks), about one-third of the size usually employed when generating levels for Super Mario Bros game in previous experiments [15, 19]. The selection of this length was due to a compromise between a window size that is big enough to allow sufficient interaction between the player and the game to trigger the examined affective states and a window which is small enough to set an acceptable frequency of an adaptation mechanism applied in real-time aiming at closing the affective loop of the game [5].

4 Recorded Variables

The following sections describe the data that has been included in the corpus.

4.1 Controllable Features

The level generator of the game has been modified to create levels according to the following six controllable features:

- The number of gaps in the level.
- The average width of gaps.
- The number of enemies.
- Enemies placement: The way enemies are placed around the level is determined by three probabilities which sum to one; around horizontal boxes; Around gaps and random placement.
- The number of powerups.
- The number of boxes.

Two states (low and high) are set for each of the controllable parameters above except for enemies placement which has been assigned three different states allowing more control over the difficulty and diversity of the generated levels. The selection of these particular controllable features was done after consulting game design experts, and with the intent to cover the features that have the most impact on the investigated affective states [15, 19]. Other features of the levels have been given fixed values such as the number of cannon and flower tubes = 1, the type of background = over ground, the number of coins = 7, the number of coins hidden in boxes = half the total number of boxes and the number of stairs around the gaps = half the number of gaps.

The analysis presented in this paper is based on 36 players playing 120 game pairs chosen randomly by varying the above-mentioned controllable parameter values. Readers may refer to [20] for more detailed information about these features and the level generation process.

4.2 Gameplay Events

While playing the game, different player actions and their corresponding timestamps have been recorded. These events are categorized in different groups according to the type of the event and the type of interaction with the game objects.

- Winning: This event is generated when the player wins.
- Losing: This event is generated when the player loses. An extra attribute is associated with this event to define the type of object that causes the death. This attribute can take one of the following values (turtle, goompa, flower, cannon ball and gap).
- Interaction with game items: This event is generated when the player collects items or interacts with intractable game objects. The event has an attribute that defines the type of the object that can take one of the values: free coins, empty rock, coin block/rock and power-up rock/block.
- Interaction with enemies: This event is generated when the player kills different types of enemies using different possible actions. The event has two attributes that defines the type of the action performed (stomp, shoot fire balls or unleash a turtle shell) and the type of the enemy killed (red/green/armored turtle, goompa, flower or cannon ball)
- Changing Mario mode: Mario can be in one of the following modes; little, big and fire. Whenever Mario mode is changed, an event is generated with the information about the type of the new mode.
- Changing Mario state: An event is generated whenever Mario changes his state between: moving left/right, jumping, running and ducking.

Note that for each of the above-mentioned events the associated time within which this event occurred is also recorded. Also note that for the last two events two time-stamps are saved marking the start and the end of the event.

4.3 Player Experience

Player experience is measured through rankings presented to the player after each game, and a 4-alternative forced choice questionnaire, presented to the player after playing a pair of games with different controllable features. The questionnaire asks the player to report the preferred game for three user states: *engagement*, *challenge* and *frustration*. The selection of these states is based on earlier game survey studies [15] and our intention to capture both affective and cognitive/behavioral components of gameplay experience [25]. Moreover, we want to keep the self-reporting as minimal as possible so that experience disruption is minimized.

4.4 Demographics

Each subject has been asked to give information about the following:

- Nationality, gender and age.
- Whether or not the player plays video games (yes and no).
- The weekly time spent playing video games (0-2 hours per week; 2-5 hours per week; 5-10 hours per week and 10+ hours per week).
- Whether or not the player has played Super Mario before (yes and no).

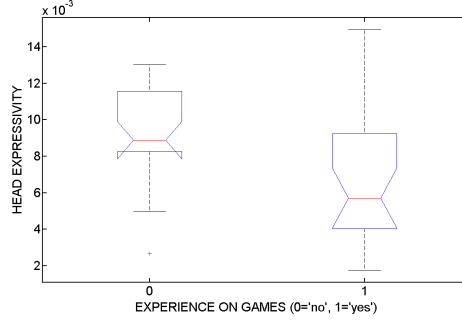


Fig. 2. Experienced game players’ and non-experienced players’ average head expressivity during game sessions.

5 Experiments

5.1 Correlating player movements with self reported data

Discovering game events that trigger particular player reactions is expected to highlight instances of personalized preferences or lack of interest. In the following, we present an initial experiment that has been conducted to demonstrate how data from the corpus can be used to analyze this relationship.

A series of head movement parameters have been extracted [1] (please note that these parameters are not part of the data included in the corpus, but rather, features extracted from the video recordings). More specifically, the parameters extracted for each player were: Horizontal H_h and Vertical V_h components of head pose, the horizontal component of eye gaze H_e , head roll H_r and the fraction of the inter-ocular distance with regards to a frame where the algorithm starts. This last indicator is a variable standing for player movements back and forth.

Experiments have shown that all players, regardless of their sex, nationality or experience, tend to adopt high head expressivity (here, defined as the absolute derivative of head rotation) when they lose. More specifically, all 36 players pose fast head movements, on average almost twice (1.8 times) as fast as their head rotation speed during the whole session.

Furthermore, analyzing the players’ video sessions has shown that head expressivity is less intense, as experience on game playing increases. Figures 2 and 3 show box plots of average head expressivity per player and reported experience on game playing, as well as average head expressivity and amounts of hours spent on video game playing on a weekly basis.

The respective one-way anova tests also illustrate the fact that head expressivity is an indicator of how experienced one is on games ($F(1, 34) = 4.27$, $p = 0.049$) or the amount of hours they spend on game-playing, on a weekly basis ($F(3, 32) = 3.46$, $p = 0.028$).

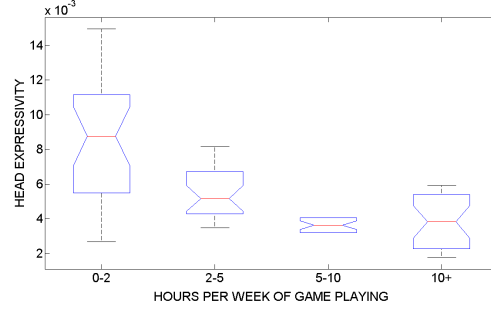


Fig. 3. Average head expressivity as a function of the hours spent on game playing on weekly basis.

	E	Not E	F	Not F	C	Not C
E			44%	56%	69%	30%
F	44%	56%			76%	24%
C	69%	29%	78%	22%		

Table 1. Within-game percentage of agreement between self-reported user states: Engaging (E); Frustrating (F) and Challenging (C).

5.2 Analysis of self reported data

On another front of the analysis we attempt to make some initial experiments in order to see how self-reports relate to each other and derive conclusions about the interplay between self-reported, engagement, challenge and frustration in Super Mario. Table 1 presents the pairwise comparison of the percentage of the preferred games across the three emotional states examined. The data presented in the table has been extracted from the players' answers to the preference questionnaires in which the players showed a clear preference for one game over the other. Based on these results it appears that, 30% of the players were engaged in a game that is not challenging while 56% of them found a frustrating game not engaging. These findings suggest that a game can be engaging if it contains the right level of challenge and frustration. It seems that, on average, a game that is too frustrating for the player is not engaging and so is a game with too high or too low level of challenge. This means that the challenge level in the game should meet each individual players' skill. The results from the relationship between engagement and the level of challenge show a consistency with Csikszentmihalyi's theory of *flow* [4].

6 Conclusions and future work

In this paper, we presented a dataset built for capturing player behavior and preferences during gameplay. The participants covered the whole gamut of experience levels, and attention was paid so that players were from a multinational background. Ground truth of the dataset comes in the form of visual feedback of players' behavior, answers to questionnaires and demographical information.

Initial results from experiments that demonstrate the use of the corpus suggest that there is a significant affect of players' expertise on average head expressivity. Additional findings indicate that self-reported engagement, challenge and frustration are linked, in part, to the theory of flow. The work presented in this paper is the first important step to our future work that will focus on a more in-depth analysis of the relationship between behavioral, game events, game parameters and experience variables. Successfully defining this relationship constitutes a very important step in any adaptation mechanism applied aiming at closing the affective loop in games.

Acknowledgments. We would like to thank the participants in the experiments. This research was supported, in part, by the FP7 ICT project SIREN (project no: 258453) and by the Danish Research Agency, Ministry of Science, Technology and Innovation project *AGameComIn*; project number: 274-09-0083.

References

1. Asteriadis, S., Tzouveli, P., Karpouzis, K., Kollias, S.: Estimation of behavioral user state based on eye gaze and head pose - application in an e-learning environment. *Multimedia Tools and Applications*, Springer 41(3), 469 – 493 (2009)
2. Calleja, G.: *In-Game From Immersion to Incorporation*. MIT Press, USA (2011)
3. Conati, C.: Probabilistic assessment of users emotions in educational games. *Applied Artificial Intelligence* 16, 555–575 (2002)
4. Csikszentmihalyi, M.: *Flow: The Psychology of Optimal Experience*. Harper Perennial (Mar 1991)
5. Höök, K.: Affective loop experiences - what are they? In: *PERSUASIVE. Lecture Notes in Computer Science*, vol. 5033, pp. 1–12. Springer (2008)
6. Hudlicka, E.: Affective computing for game design. In: *GAMEON-NA'08: Proceedings of the 4th Intl. North American Conference on Intelligent Games and Simulation*. pp. 5–12. Montreal, Canada (2008)
7. Ioannou, S., Caridakis, G., Karpouzis, K., Kollias, S.: Robust Feature Detection for Facial Expression Recognition. *EURASIP Journal on Image and Video Processing* 2007(2), 1–23 (2007)
8. Isbister, K., Schaffer, N.: *Game Usability: Advancing the Player Experience*. Morgan Kaufman (2008)
9. Koster, R.: *A theory of fun for game design*. Paraglyph press (2004)
10. Leite, I., Pereira, A., Mascarenhas, S., Castellano, G., Martinho, C., Prada, R., Paiva, A.: Closing the loop: from affect recognition to empathic interaction. In: *Proceedings of the 3rd international workshop on Affective interaction in natural environments*. pp. 43–48. *AFFINE '10*, ACM, New York, NY, USA (2010)

11. Malone, T.: What makes computer games fun? (abstract only). In: Proceedings of the joint conference on Easier and more productive use of computer systems. (Part - II): Human interface and the user interface - Volume 1981. pp. 143–. CHI '81, ACM, New York, NY, USA (1981)
12. Mandryk, R.L., Inkpen, K.M.: Physiological indicators for the evaluation of co-located collaborative play. In: Proceedings of the 2004 ACM conference on Computer supported cooperative work. pp. 102–111. CSCW '04, ACM, New York, NY, USA (2004)
13. Pagulayan, R.J., Keeker, K., Wixon, D., Romero, R.L., Fuller, T.: The human-computer interaction handbook, chap. User-centered design in games, pp. 883–906. L. Erlbaum Associates Inc., Hillsdale, NJ, USA (2003), <http://portal.acm.org/citation.cfm?id=772072.772128>
14. Pedersen, C., Togelius, J., Yannakakis, G.N.: Modeling player experience in super mario bros. In: CIG'09: Proceedings of the 5th international conference on Computational Intelligence and Games. pp. 132–139. IEEE Press, Piscataway, NJ, USA (2009)
15. 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)
16. Picard, R.W.: *Affective Computing*. Cambridge, MA: The MIT Press (1997)
17. Picard, R.W., Vyzas, E., Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 1175–1191 (2001)
18. Poels, K., IJsselstein, W.: Development and validation of the game experience questionnaire. In: FUGA Workshop mini-symposium. Helsinki, Finland (2008)
19. 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). AAAI Press (October 2010)
20. Shaker, N., Yannakakis, G.N., Togelius, J.: Feature Analysis for Modeling Game Content Quality. In: *IEEE Transactions on Computational Intelligence and AI in Games*. IEEE Press (2011)
21. Sidner, C., Lee, C., Kidd, C., Lesh, N., Rich, C.: Explorations in engagement for humans and robots. *Artificial Intelligence* 166(1-2), 140–164 (2005)
22. Smith, P., Shah, M., da Vitoria Lobo, N.: Determining driver visual attention with one camera. *IEEE transactions on intelligent transportation systems* 4(4), 205–218 (2003)
23. Yannakakis, G.N.: Preference Learning for Affective Modeling. In: Proceedings of the Int. Conf. on Affective Computing and Intelligent Interaction. pp. 126–131. IEEE, Amsterdam, The Netherlands (September 2009)
24. 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)
25. Yannakakis, G.N., Togelius, J.: Experience-driven Procedural Content Generation. *IEEE Transactions on Affective Computing* pp. – (2011), (in print)
26. Yannakakis, G., Hallam, J.: Real-time adaptation of augmented-reality games for optimizing player satisfaction. In: *Computational Intelligence and Games, 2008. CIG '08. IEEE Symposium On*. pp. 103–110 (dec 2008)
27. Yannakakis, G., Hallam, J.: Real-time Game Adaptation for Optimizing Player Satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games* 1(2), 121–133 (June 2009)