Does your profile say it all? Using demographics to predict expressive head movement during gameplay

Stylianos Asteriadis[1], Kostas Karpouzis[1], Noor Shaker[2], and Georgios N. Yannakakis[2]

[1] Image, Video and Multimedia Systems Lab, National Technical University of Athens, 9 Iroon Polytechniou str, GR-157 80, Athens, Greece.
[2] IT University of Copenhagen, Rued Langaards Vej 7, 2300 Copenhagen, Denmark
{stiast,kkarpou}@image.ntua.gr
{nosh,yannakakis}@itu.dk

Abstract. In this work, we explore the relation between expressive head movement and user profile information in game play settings. Facial gesture analysis cues are statistically correlated with players' demographic characteristics in two different settings, during game-play and at events of special interest (when the player loses during game play). Experiments were conducted on the Siren database, which consists of 58 participants, playing a modified version of the Super Mario. Here, as player demographics are considered the gender and age, while the statistical importance of certain facial cues (other than typical/universal facial expressions) was analyzed. The proposed analysis aims at exploring the option of utilizing demographic characteristics as part of users' profiling scheme and interpreting visual behavior in a manner that takes into account those features.

Keywords: interaction personalization, user profile, game adaptation, facial expression recognition, facial gestures recognition

1 Introduction

Video games is a flourishing industry for more than three decades now, with revenues surpassing even those of the movie and music industries [9]. Due to their high popularity and the computational demands they are associated with, video games would always introduce leading technologies and pioneer methods in the field of Human-Computer Interaction. Another reason for this is that fact that game players can be immersed when playing a game they enjoy, a state which usually makes them more expressive, thereby providing more non-verbal cues, which an adaptive HCI system should be able to capture and process so as to provide a truly personalized experience. Today's technologies have reached a point where new add-ons, most notably hand-held controllers or cameras, can boost the game-play experience, altering and guiding game content and evolution following affect-dependent strategies. To this aim, using *context* and *behavior*-related parameters to elicit information regarding current player state (and, consequently, obtain hints about her/his needs regarding interaction evolution) are of primary importance for constructing personal behavioral and interaction-related models and guiding game adaptation in order to achieve maximum engagement and flow levels [1], [4].

1.1 Related Work

There is an abundance of works presented in bibliography, dealing with the problem of interpreting user behavioral characteristics within the framework of game-play and, in general, Human-Computer Interaction. Recent advances in computer vision under non pretending conditions have allowed the proposal of techniques incorporating notions such as body and head movements [1], eye gaze (with eye gaze still necessitating specialized hardware [7]) and facial expressions [6]. Typical works are those reported in [3] and [10], where the authors utilize Bayesian networking on gaze, postural and contextual data for detecting user engagement with a robot companion [2] posing various expressions. The use of context itself is also a very important factor for predicting one's current state within the frame of game-play. Within this view, mouse pressure and accelerometers positioned on the player's back are used in [5]. The authors also employ information coming from chair sensors and mappings are created to player's self reports, as well as actual game difficulty levels. A Gaussian process classification is employed in [8] for detecting moments of frustration in a person-independent scenario. Children were asked to deal with a problem-solving activity, and a multiple-sensor setup was installed. The authors in [12] have conducted an experiment on 10 people, measuring the amount of their finger pressure for creating mappings of arousal to difficulty levels.

Affect induction is an essential part of this field, since most games can be tweaked in order to make the player experience more expressive and, thus, produce multimodal data which can be analyzed and classified. Scheirer [17] frustrates the user on purpose in a general HCI framework, in order to produce and record rich affective data and Katsis [18] put this approach to use in the context of car racing games, a popular game paradigm. Wang and Marsella [19] produced EVG (Emotion eVoking Game), a dungeon role playing game used to induce emotions related to discrete emotion labels (boredom, surprise, joy, anger and disappointment). This approach was first tested on the popular PacMan paradigm by Kaiser et al. [20], where Scherer's appraisal dimensions are also taken into account [21], [22], catering for a richer affective representation than the usual discrete labels.

In the context of game control, an increasing number of studies utilize physiological signals as a means to integrate affect and go beyond the usual keyboard/mouse game control metaphor. Istance et al. [23] extended the PacMan paradigm with the concept of frustration induction, which is measured using EEG sensors, while Saari et al. [24] introduce adaptation in the context of a first person shooter (FPS) game, where player affect is measured via psychophysiological measures. Tijs et al. [25] use self-reported affect in the 2D space of valence vs. arousal space and mainly physiology-based emotion-related features to distinguish between a boring, frustrating and enjoying game mode while Rani et al. [26], and Mandryke et al. [27] experiment with challenge during gameplay. Chanel et al. [28], [29] and Levillain et al. [30] investigate player experience as an indicator towards game adaptation.

Within this framework, the presented work goes one step further into player profile construction using demographics along with visual behavior. This work utilizes personalized or feature-dependant manners to interpret visual behavior with regards to cognitive or emotional states. In particular, we examine the impact of age and gender to facial expressivity and head motion. The presented analysis is intended to place certain criteria on mapping vision to behavior, aiming to interaction adaptation for maximizing engagement or learning effect.

The structure of the rest of the paper is the following: Session 2 gives an overview of the data acquisition procedure, related to game and visual features. Section 3 gives the analytical results regarding statistically important features for female and male participants, as well as different age groups. Discussions and conclusions follow at the end of the paper.

2 Data acquisition

2.1 Game environment

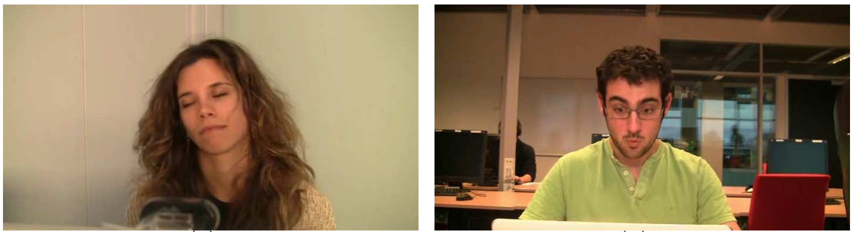
For extracting correlations between visual features (facial gestures) and interaction content, we used a modified version of Super Mario, which provides timestamped output of game-related events (player jumping or changing direction, loss, level end, power up use, etc.) (Fig. 1). For collecting data, players from Greece and Denmark were seated in front of a computer screen, in order to play the testbed game. Volunteers played from four to ten games each, with each game lasting about one minute, depending on player's performance. While playing the game, different player actions and their corresponding timestamps were recorded. At the end of each session, the players were asked to report demographic data including age, nationality and expertise in playing video games. A more analytical description of the protocol followed for collecting data can be found in [11].

In this study, the dataset was extended, reaching a total of 58 participants (30 female, 28 male) corresponding to about 4.7 hours of game play material. People's upper body was recorded while they were playing, and certain facial gestures were measured at certain events (here, whenever Super Mario was killed, Fig. 2), as well as, throughout game play of a whole stage per player.

2.2 Facial Gestures extraction

A number of facial cues (based on detecting and processing prominent facial feature points) were extracted during interaction and their relations to game events,

Fig. 1. Snapshot from Infinite Mario Bros, showing Mario standing on horizontally placed boxes surrounded by different types of enemies



(a) (b)

Fig. 2. Facial expressions of players at moments of losing

as well as information stored in the player's profile, were explored. Initially, the face is detected based on a fusion of techniques: Viola-Jones [15] followed by an ellipse fitting algorithm [16] are used for a precise estimate of face location. A skin detection step was also used here, to get the precise contour of the face [13] and the upper, lower, left and rightmost points were extracted. At each frame, eye centers and eyes midpoints are detected using the methodology described in [16], and an area expanded above each eye is used for locating two points on each eyebrow, based on local minima in luminance. Finally, mouth bounding box is extracted using Haar features. Based on the features above, seven parameters were measured and further utilized and correlated to user characteristics (Fig. 3, 4):

- The euclidean distance of the eyes midpoint with regards to its initial position at the beginning of the game session was extracted and its first derivative was considered as the *overall head movement.*
- Average distances of eyebrows points from the eye centers was considered as well, as an indicate of surprise or anger.
- As a measure of player movements on the z-axis, the inter-ocular distance was also tracked and its first derivative was calculated.

Fig. 3. Typical examples of play instances. The two players react in a different way to the same event.

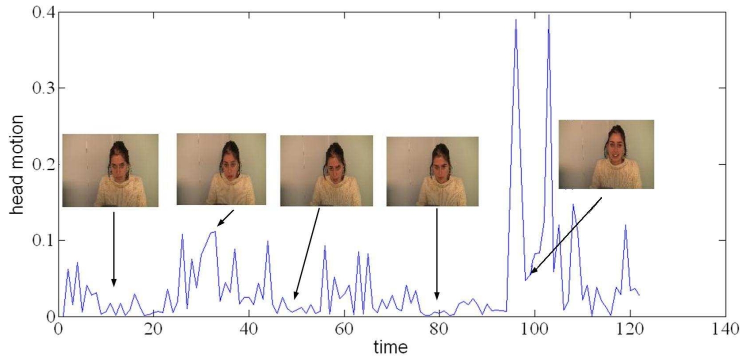


Fig. 4. Player visual behavior during gameplay. In this session, Super Mario was killed in seconds $\simeq 32$ and $\simeq 100$.

- Eye centers position with regards to facial vertical boundaries, as an indicate of gaze was also considered. To measure the overall activation of this measurement, the first derivative was calculated.
- Horizontal head movements' first derivatives would also account for overall movement on this axes.
- Similar, vertical head movements were considered.
- Pre-defined areas around the mouth position were considered and the average saturation values per frame were calculated. Changes in these values would account for lip activity detection.

3 Experiments

Linear regression analysis was performed to check the dependence of facial cues on age. Homoscedasticity, normal error distributions and Durbin-Watson statistic have been examined and were met.

3.1 Age, Gender and Facial Expressivity

General expressivity of facial cues was found to have statistically important correlations with age, for the entire set of players. More specifically, eye gaze directionality and head movements appear to have higher values as age increases (Fig. 5(a), 5(b), 5(c), 5(d)).

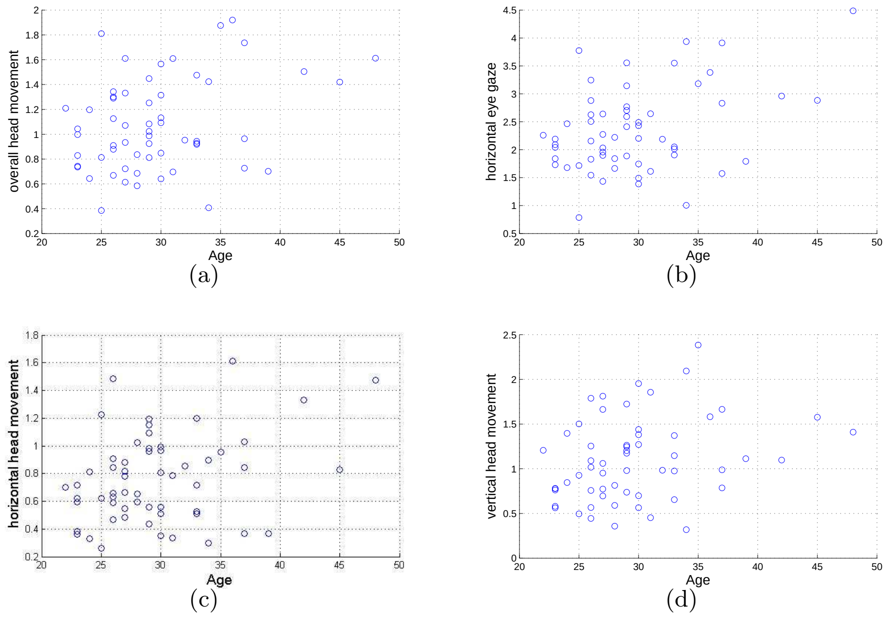


Fig. 5. a) Overall head movement (r=0.32, R^2=0.11, F=6.6, p=0.013), b) eye gaze directionality (r=0.4, R^2=0.16, F=10.6, p=0.002, c) horizontal head movements (r=0.32, R^2=0.10, F=6.5, p=0.014 and d) vertical head movements throughout game sessions and their relations with age (r=0.29, R^2=0.09, F=5.3, p=0.026;

Averaged expressivity throughout whole game sessions, however, is highly related to age, especially for female participants, and appears to increase as age increases. The features that appear to have statistically important correlations with age were the Overall head expressivity (r=0.43, R^2=0.18, F=6.3, p=0.018), horizontal eye gaze (r=0.50, R^2=0.25, F=9.5, p=0.005) and head movement (r=0.46, R^2=0.21, F=7.6, p=0.01). The above are depicted in Figures 6(a), 6(b) and 6(c).

In the case of male participants, facial cues do not appear to significantly depend on age, apart from that of lip movements (r=0.46, R^2=0.21, F=7.0, p=0.014) which, as seen in Figure 7, increases as age increases.

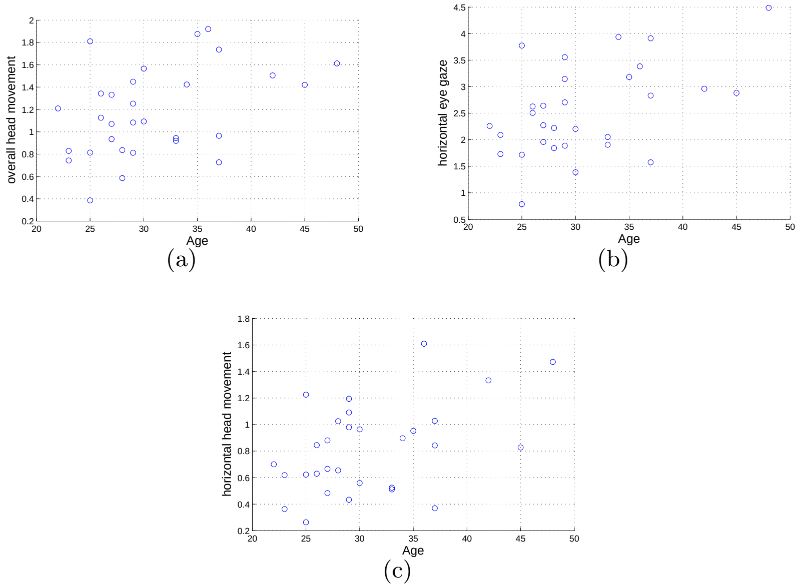


Fig. 6. a) Overall head movement in relation to age for women in the dataset; b) Horizontal eye gaze in relation to age for women in the dataset; c) Horizontal head movement in relation to age for women in the dataset

3.2 Age, Gender and Facial Expressivity during critical events

Most of the facial cues appeared to have higher expressivity for the total of the population when players would lose. For instance, analysis of variance for the overall head movements shows that they take much larger values (1.73±0.66) during moments of losing than normal game play (1.07±0.37) (p<<0.01). Similar is the case for horizontal eye gaze, and vertical and horizontal components of head pose.

As in the case of general game play durations, changes of eye gaze directionality increase as age increases. Figure 8(a) shows the scatter plot of the above parameters (r=0.32, R^2=0.10, F=6.3, p=0.015).

Similar (Figure 8(b)), horizontal movements of head appear to have strong correlations with age at instances when Super Mario is killed (r=0.43, R^2=0.19, F=12.7, p=0.0007).

Most of the expressivity, however, appears to be coming from women, at least during moments of failure. Furthermore, this was more evident as age would increase. In particular, eye gaze patterns appear to have strong correlation with age in women (Figure 9(a)), (r=0.55, R^2=0.30, F=12.07, p=0.0017).

Figure 9(b) shows the dependency of head horizontal movements on age, for women (r=0.58, R^2=0.34, F=14.4, p=0.0007).

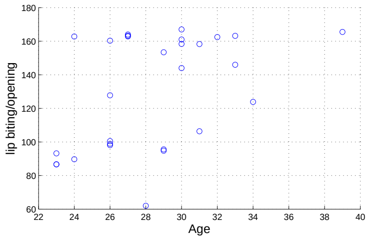


Fig. 7. Dependence of lip movements on age for male participants.

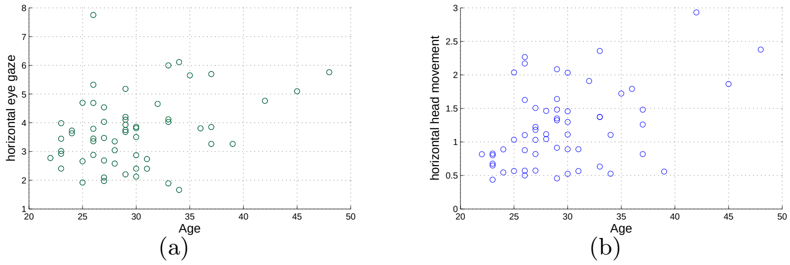


Fig. 8. a) Eye gaze and b) horizontal head movements for all players in relation to age, during instances when player loses (Super Mario is killed).

Similar, Figure 9(c) is illustrative of the dependency of overall head movement (translational and rotational movements towards any direction) for women (r=0.39, R^2=0.15, F=4.9, p=0.035).

For males, the only facial cue that was dependent on age (with p <0.05) was lip movements (biting, opening, etc.), with the older ones showing higher chances of posing such expressions (Figure 10, r=0.46, R^2=0.21, F=6.8, p=0.015).

4 Discussion and Conclusions

The presented analysis showed that there exist statistically significant correlations between visual expressivity and player demographic characteristics. In particular, it was shown that, taking into account statistically important facial gestures, players tend to be more expressive as age increases; the above is more evident among female game players, with highly statistically significant measurements referring mainly to eye movements and horizontal movements of the head. The above conclusions allow to interpret, in a personalized (or based on profile-dependent characteristics) manner, reactions people tend to adopt when interacting with games. For instance, a person reacting differently than expected (based on her/his profile characteristics), might be experiencing certain cognitive/emotional states, and special attention should be given. On the contrary, if a player for whom, based on her/his profile, certain reactions are expected, not

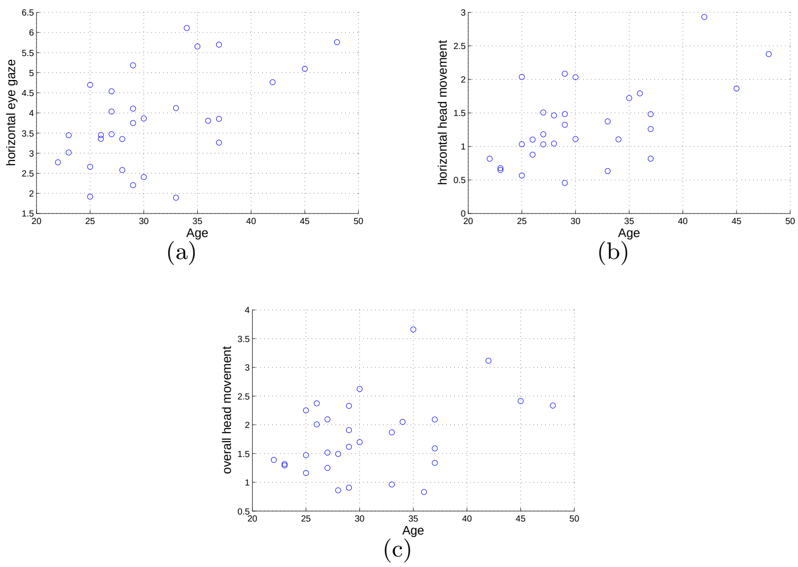


Fig. 9. a) Eye gaze, b) Horizontal head movements and c) Overall head movement for female participants as a function of age, for events of losing.

a lot of importance will be placed into interpreting her/his state, when these reactions are actually observed. These observations can lead to the construction of more sophisticated models, based on the analysis of a series of visual features, as well as contextual and user-related features.

In this research, we have focused on certain player features (age and gender) and game conditions (general play, moments when the player suddenly loses). In the near future, personalization (and classification to prototypical profiles) will be further expanded by taking into account self reports and game context (e.g. difficulty levels). Already, initial results on the Super Mario database [14] have shown that, based on a fusion of visual reactions with certain events and player

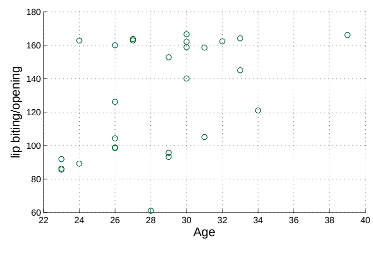


Fig. 10. Lip movements of male participants, as functions of age, for the events of losing.

self-reported experience, it is possible to estimate user cognitive or emotional states. The above is very useful in circumstances where personalized interaction environments are required, aiming at maximizing engagement or expected outcome (e.g. learning effect).

Acknowledgments. This research was supported by the European FP7 ICT project SIREN (project no: 258453).

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. van Breemen, A.J.N., Yan, X., Meerbeek, B.: icat: an animated user-interface robot with personality. In: AAMAS. pp. 143–144 (2005)
3. 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, New York, NY, USA (2009)
4. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, New York, NY (1991)
5. Van den Hoogen, W., IJsselsteijn, W., de Kort, Y.: Exploring behavioral expressions of player experience in digital games. In: Proceedings of the Workshop on Facial and Bodily Expression for Control and Adaptation of Games ECAG. pp. 11–19 (2008)
6. Ioannou, S., Caridakis, G., Karpouzis, K., Kollias, S.: Robust Feature Detection for Facial Expression Recognition. EURASIP Journal on Image and Video Processing (2) (2007)
7. Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. Int. J. Hum.-Comput. Stud. 66, 641–661 (2008)
8. Kapoor, A., Burleson, W., Picard, R.W.: Automatic prediction of frustration. Int. J. Hum.-Comput. Stud. 65, 724–736 (2007)
9. Reuters: Video-game sales overtaking music. `http://articles.moneycentral.msn.com/Investing/Extra/VideoGameSalesOvertakingMusic.aspx`
10. Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P.W., Paiva, A.: Automatic analysis of affective postures and body motion to detect engagement with a game companion. In: Proceedings of the 6th international conference on Human-robot interaction. pp. 305–312. HRI '11, ACM, New York, NY, USA (2011)
11. Shaker, N., Asteriadis, S., Yannakakis, G., Karpouzis, K.: A game-based corpus for analysing the interplay between game context and player experience. Affective Computing and Intelligent Interaction pp. 547–556 (2011)
12. Sykes, J.: Affective gaming: measuring emotion through the gamepad. In: CHI 2003: New Horizons. pp. 732–733. ACM Press (2003)
13. Kovac J, Peer P, Solina F (2003) Human skin colour clustering for face detection. In: IEEE International Conference on Computer as a Tool, vol 2
14. Asteriadis S, Shaker N, Karpouzis K, Yannakakis G.N. (2012) Towards Players Affective and Behavioral Visual Cues as drives to Game Adaptation. In: LREC Workshop on Multimodal Corpora for Machine Learning, Istanbul

15. Viola PA, Jones MJ (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, vol 1, pp 511–518
16. Asteriadis S, Nikolaidis N, Pitas I (2009) Facial feature detection using distance vector fields. Pattern Recognition 42(7):1388–1398
17. J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard, "Frustrating the user on purpose: a step toward building an affective computer," *Interacting with Computers*, vol. 14, no. 2, pp. 93–118, 2001.
18. C. D. Katsis, N. Katertsidis, G. Ganiatsas, and D. I. Fotiadis, "Toward emotion recognition in car-racing drivers: A biosignal processing approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 38, no. 3, pp. 502–512, 2008.
19. N. Wang and S. Marsella, "Introducing evg: An emotion evoking game," in *IVA*, 2006, pp. 282–291.
20. S. Kaiser, T. Wehrle, and S. Schmidt, "Emotional episodes, facial expressions, and reported feelings in human-computer interactions," in *Proceedings of the Xth Conference of the International Society for Research on Emotions*. Würzburg: ISRE Publications, 1998, pp. 82–86.
21. N. Lanctôt and U. Hess, "The timing of appraisals." *Emotion*, vol. 7, no. 1, pp. 207–12, 2007.
22. C. van Reekum, T. Johnstone, R. Banse, A. Etter, T. Wehrle, and K. Scherer, "Psychophysiological responses to appraisal dimensions in a computer game," *Cognition & Emotion*, vol. 18, no. 5, pp. 663–688, 2004.
23. H. O. Istance, A. Hyrskykari, S. Vickers, and T. Chaves, "For your eyes only: Controlling 3d online games by eye-gaze," in *INTERACT (1)*, 2009, pp. 314–327.
24. T. Saari, M. Turpeinen, K. Kuikkaniemi, I. Kosunen, and N. Ravaja, "Emotionally adapted games - an example of a first person shooter," in *HCI (4)*, 2009, pp. 406–415.
25. T. J. W. Tijs, D. Brokken, and W. IJsselsteijn, "Creating an emotionally adaptive game," in *ICEC*, 2008, pp. 122–133.
26. P. Rani, N. Sarkar, and C. Liu, "Maintaining optimal challenge in computer games through real-time physiological feedback," in *Proceedings of the 1st International Conference on Augmented Cognition, Las Vegas, NV*, 2005.
27. R. L. Mandryk and M. S. Atkins, "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies," *Int. J. Hum.-Comput. Stud.*, vol. 65, no. 4, pp. 329–347, Apr. 2007.
28. G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun, "Boredom, engagement and anxiety as indicators for adaptation to difficulty in games," in *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*, ser. MindTrek '08. New York, NY, USA: ACM, 2008, pp. 13–17.
29. G. Chanel, C. Rebetez, M. Bé andtrancourt, and T. Pun, "Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 41, no. 99, pp. 1–12, 2011.
30. F. Levillain, J. O. Orero, M. Rifqi, and B. Bouchon-Meunier, "Characterizing player's experience from physiological signals using fuzzy decision trees." in *2010 IEEE Symposium on Computational Intelligence and Games (CIG)*, G. N. Yannakakis and J. Togelius, Eds. IEEE, 2010, pp. 75–82.