Towards Automated Player Experience Detection With Computer Vision Techniques

Chek Tien Tan
Games Studio, Faculty of
Engineering and IT, University
of Technology, Sydney
chektien.tan@uts.edu.au

Yusuf Pisan
Games Studio, Faculty of
Engineering and IT, University
of Technology, Sydney
yusuf.pisan@uts.edu.au

CHI 2012, May 5–10, 2012, Austin, TX, USA.
ACM xxx-x-xxxx-xxxx-x/xx/xx.

Abstract
There has been an increasing number of quantitative methods to measure and evaluate player experiences. However, current methods either focus on telemetry approaches, which are insufficient to capture real life responses, or psychophysiological methods, which are intrusive and more suited to controlled laboratory environments. This paper presents the position that computer vision techniques can provide a less intrusive and more versatile solution for automatic evaluation of game user experiences. A conceptual framework to automatically infer flow intensity is presented and a work-in-progress study is included to demonstrate the feasibility of this research direction.

Keywords
player experience, playtesting, facial expression analysis

ACM Classification Keywords
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces: Evaluation/methodology; I.2.1 [**Applications and Expert Systems**]: Games.

General Terms
CHI 2012, workshop, Game User Research

Introduction

This paper presents a novel perspective using computer vision techniques to automatically infer gameplay experience metrics. The motivations for this research is derived from the state of current research and practice in the area of player experience analysis, which is broadly categorized into qualitative and quantitative methods.

Qualitative methods involve the collection and analysis of subjective data and common techniques include direct observations, interviews and think-aloud protocols. These methods are usually employed by practitioners in formal playtest sessions which requires players to be put into artificial play environments [12]. This inhibits true play experiences, as the players might not be totally at ease when someone is watching or questioning them. Players might not be able to properly self-articulate their play experiences concurrently during gameplay and might not even remember important experiences when post interviews are performed. The sessions also often require a lot of time and resources to conduct and analyse.

Quantitative methods work on objective data with common approaches like telemetry and psychophysiology. Telemetry primarily deals with the logging of player in-game interactions and several studies have been done with it [16, 8, 9, 5]. The advantage of Telemetry over qualitative methods is that it is non-disruptive and that it can continuously capture gameplay statistics. However, the data is greatly limited by the in-game actions available to the player and events in the game world. Hence these "virtual observations" do not capture full experiences and might not even represent the true experiences of the player in real life.

Psychophysiology deals with inferring psychological states from physiological measurements that commonly include electrodermal activity (EDA), electromyography (EMG), electrocardiogram (ECG), electroencephalography (EEG), body temperature and pupil dilations. In game-related psychophysiological research, a number of works [7, 10, 11, 16, 3] have also been performed. An important observation is that several physiological data are found to be strongly correlated to user self-reports on gameplay metrics. Similar to telemetry, physiological measurements are able to non-disruptively capture player experiences continuously. In addition, physiological data represent the real life experiences of the player. Unfortunately, most current approaches deal with expensive specialized equipment that are obtrusive, which are usually only viable in controlled laboratory settings.

Motivated by the shortcomings outlined above, we propose a computer vision-based technique to infer player experiences automatically based on facial expressions analysis [6]. It can be also viewed as a type of psychophysiological approach, which seemed to be ignored in player experience research. Facial expression recognition is a mature domain with techniques that boast a high level of accuracy and robustness [1]. Being a video-based approach, it is non-obtrusive compared to current physiological approaches. This allows for a more authentic play experience and enables data collection in non-laboratory settings. In addition, current technological advancements favors this approach. The advent of motion detection game consoles like Microsoft's Kinect and Nintendo's 3DS, video and voice feeds are naturally incorporated into gameplay. For other games, webcams are also relatively cheap and prevalent in most mobile computing devices nowadays.

To the best of our knowledge, no work has been performed to evaluate the feasibility of facial expressions

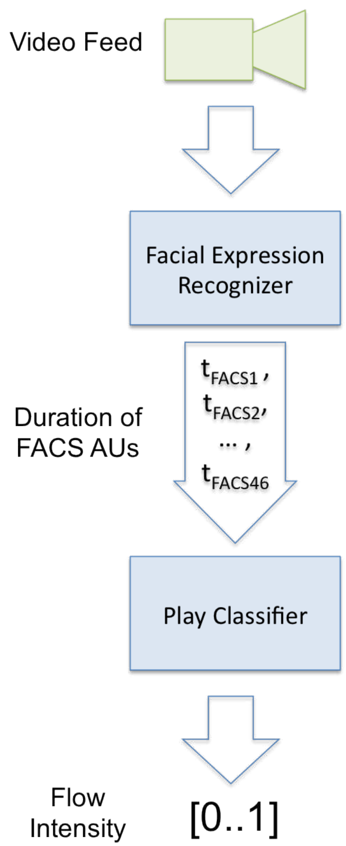

Figure 1: Overview of the automated play experience detector framework. The player's captured video is fed into the face recognizer which outputs FACS AUs[4]. These features are then input into the play classifier which determines the intensity of flow [13].

analysis as a basis to infer gameplay experience metrics like flow, immersion and presence. Research in non-game domains have shown good results for inferring other kinds of user experience metrics [15, 6]. As an initial study, this paper proposes a method to automatically infer whether a player attains the flow state, as prior psychophysiological approaches [10] have shown promising correlations between the measured physiological data and self-reported flow experience. Focusing on only flow tightly contains our research before we move on to other metrics. This paper sets up the groundwork for our planned experiments.

Current Work

The proposed framework is a supervised learning system that maps facial expression features to gameplay experience metrics. An overview of our framework is as shown in Figure 1.

For the duration of play for each player, the system will continuously detect the duration of each Facial Action Units (AU) as defined in the Facial Action Coding System (FACS) [4]. The usage of FACS, as opposed to common expressions like happy, sad and anger, ensures we have a rich and detailed input that can potentially describe a larger range of gameplay experience metrics. While it can be argued that EDA and ECG approaches can capture more detailed responses, they can only be used in laboratory settings on small numbers of players. Since FACS is based on detailed muscle twitches, subtle expressions can be captured (like a visual lie detector).

The duration value for each AU will then be inputs into the play classifier in which flow intensity will be determined as the output. For the play classifier, at this proof of concept stage, we will utilize a feed-forward Artificial Neural Network (ANN) due to its success in prior work involving facial expressions [2].

After the game, the player will be required to fill in the Flow State Scale (FSS) [13] questionnaire so as to determine an aggregated flow intensity score during the game. This score will act as the supervision information used to train the classifier. As training convergence is reached, the system can then be used reliably to automatically infer flow intensity results.

As an initial proof of concept, we target Microsoft's Kinect games as this seems to be currently most suitable platform genre. It is the least intrusive due a video and depth map already being recorded as part of gameplay. In addition, depth data from the Kinect can contribute to enhanced facial expression recognition as evident in [14].

We are currently developing a Kinect game with game balance mechanics that can be easily modified. Once that is completed, the automated learning system will be incorporated within the game. The game will then be played by about 100 participants, who will complete the FSS questionnaire. The game balance will be modified to collect a wide range of play experiences. The training stage is naturally built into the process, and the system will learn continuously until convergence and then performs maximally. The results will then be plotted for convergence analysis as well as to cater for further iterations of the system. The results will also be analysed against the questionnaire findings to assess the correlations.

Conclusion

We have proposed an approach, building on current research, to propose a non-disruptive and non-obtrusive user experience analysis method that can be readily employed to automatically detect flow. It can easily integrate into existing games to provide feedback even after they have been launched. Using facial expressions, we can capture a rich and continuous representation of real life player responses in authentic play situations, enabling a finer-grained analysis of gameplay experiences. We believe it will have significant impact for game companies who would like real life user experience data even after launching their games.

In order to meaningfully use the data to modify game design, a planned future work is to incorporate telemetry data into the analysis system, so as to match the user experience metrics with specific game events temporally.

References

[1] J. M. Buenaposada, E. Muñoz, and L. Baumela. Recognising facial expressions in video sequences. *Pattern Analysis and Applications*, 11(1):101–116, Oct. 2007.

[2] M. N. Dailey, G. W. Cottrell, C. Padgett, and R. Adolphs. EMPATH: a neural network that categorizes facial expressions. *Journal of cognitive neuroscience*, 14(8):1158–73, Nov. 2002.

[3] A. Drachen, L. Nacke, G. Yannakakis, and A. Pedersen. Correlation between heart rate, electrodermal activity and player experience in First-Person Shooter games. In *Proc. SIGGRAPH Symposium on Video Games*, pages 49–54, 2010.

[4] P. Ekman, W. V. Friesen, and J. C. Hager. *FACS*, volume 48. A Human Face, 2002.

[5] A. Gagne, M. Seif El-Nasr, and C. Shaw. A Deeper Look at the use of Telemetry for Analysis of Player Behavior in RTS Games. In *Proc. ICEC*, 2011.

[6] M. Ghijsen. Facial Expression Analysis for Human Computer Interaction. *Affective Computing, IEEE Transactions on*, 2(3):147–161, 2004.

[7] R. L. Mandryk, M. S. Atkins, and K. M. Inkpen. A continuous and objective evaluation of emotional experience with interactive play environments. *Proc. CHI '06*, page 1027, 2006.

[8] B. Medler, M. John, and J. Lane. Data cracker: developing a visual game analytic tool for analyzing online gameplay. In *Proc. CHI '11*, pages 2365–2374. ACM, 2011.

[9] D. Moura, M. Seif El-Nasr, and C. D. Shaw. Visualizing and Understanding Players ' Behavior in Video Games : Discovering Patterns and Supporting Aggregation and Comparison. In *Proc. SIGGRAPH*, pages 2–7, 2011.

[10] L. Nacke and C. Lindley. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proc. Future '08*, pages 81–88, 2008.

[11] L. E. Nacke, M. N. Grimshaw, and C. a. Lindley. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers*, 22(5):336–343, Sept. 2010.

[12] C. Tan and A. Johnston. Towards a Nondisruptive, Practical, and Objective Automated Playtesting Process. In *Technical Reports of AIIDE '11*, pages 25–28. AAAI Press, 2011.

[13] G. Tenenbaum, G. J. Fogarty, and S. A. Jackson. The flow experience: a Rasch analysis of Jackson's Flow State Scale. *Journal Of Outcome Measurement*, 3(3):278–294, 1999.

[14] T. Weise, S. Bouaziz, and H. Li. Realtime performance-based facial animation. *ACM Transactions on Graphics*, 30(4), 2011.

[15] B. Zaman and T. Shrimpton-Smith. The FaceReader: Measuring instant fun of use. In *Proc. NordiCHI '06*, number October, pages 457–460. ACM, 2006.

[16] V. Zammitto, M. Seif El-Nasr, and P. Newton. Exploring Quantitative Methods for Evaluating Sports Games. *Signals*, pages 8–11, 2010.

Authors' Bio

Dr. Chek Tien Tan is a Lecturer at the University of Technology, Sydney (UTS) and holds a Ph.D. in Computer Science from the National University of Singapore (NUS). He is also currently the co-director of the Games Studio at UTS. Chek has served as conference chair for GAMEON 2011, program chair for Interactive Entertainment 2012, and is part of the academic advisory board for the Sydney International Animation Festival. With the industry, Chek currently runs a Kinect Lab with Microsoft and has worked with Ubisoft whilst at the DigiPen Institute of Technology.

Dr. Yusuf Pisan is an Associate Professor at the University of Technology, Sydney (UTS) and holds a Ph.D. in Artificial Intelligence from Northwestern University. He is the director of the Games Studio at UTS. Yusuf has served as a general conference chair and program chair for a number of international conferences. He founded the Australasian Conference on Interactive Entertainment. He was appointed to head the International Game Developers Associations (IGDA) taskforce to design a curriculum for game studies. He is also a founding member of the Sydney Chapter of ACM Special Interest Group on Graphics (SIGGRAPH).