Automatic Emotional Reactions Identification:
A Software Tool for Offline User Experience Research

Pedro A. Nogueira[1], Vasco Torres[2], and Rui Rodrigues[2]

[1] LIACC – Artificial Intelligence and Computer Science Lab., University of Porto, Portugal
[2] INESC TEC and Faculty of Engineering, University of Porto, Portugal
{pedro.alves.nogueira,rui.rodrigues}@fe.up.pt,
vascoptorres@gmail.com

Abstract. Current affective response studies lack dedicated data analysis procedures and tools for automatically annotating and triangulating emotional reactions to game-related events. The development of such a tool would potentially allow for both a deeper and more objective analysis of the emotional impact of digital media stimuli on players, as well as towards the rapid implementation of this type of studies. In this paper we describe the development of such a tool that enables researchers to conduct objective *a posteriori* analyses, without disturbing the gameplay experience, while also automating the annotation and emotional response identification process. The tool was designed in a data-independent fashion and allows the identified responses to be exported for further analysis in third-party statistical software applications.

Keywords: Player modelling, digital media, emotional reaction identification.

1   Introduction

User experience research studies generate large amounts of data that are *a)* difficult to manually analyse and *b)* require expert personnel to interpret. Furthermore, although there is some work on using biometric data for game research, there is no standard for automatic analysis or interpretation of this data [1, 2]. Therefore, there is a current need of tools capable of automatic psychophysiological data classification and emotional reaction identification. To address this issue we have developed a tool for the automated analysis of psychophysiological-measured emotional reactions to in-game stimuli. The developed tool allows game researchers to import and synchronise gameplay session videos with their corresponding psychophysiological readings [3]. Using the recorded gameplay videos allows players and game researchers to revisit past game events, thus eliminating the need for in-game interruptions. Also, using the emotional classification signal, the system is then able to automatically isolate the emotional responses to each annotated event via a two-dimensional peak detection algorithm. Throughout this paper we succinctly describe the developed tool, as well as report on its accuracy in detecting emotional responses.

2   Developed Tool

2.1   Requirement Analysis

Prior to starting the tool's development we held a small focus group with psychology researchers, which led to the following system requirements:

1) Provide an easily interpretable measure of the volunteer's emotional state
2) Provide a real-time and synchronised view of the volunteer's gaming session from both an audio-visual and psychophysiological perspective
3) Allow free manipulation of the experiment's rate of time passage
4) Allow for a simple and straightforward annotation of relevant events, with as few clicks and parameter selection as possible
5) Present time markers for each of the annotated events, as well as the ability to quickly cycle through and edit them
6) Allow for subjective data to be included for each event, if necessary
7) Automatically compute which events triggered emotional reactions
8) Incorporate a save/load feature for resuming the annotation process in relatively large data collections and posterior analysis/verification.

2.2   Tool Development

Given the modular nature of our framework, we choose to divide it into various independent components, so that any future additions or changes could be performed in an expeditious manner. These components are described in the following paragraphs.

Emotion Recognition Module (ERM)
The first step in the annotation process is determining a simplified, although relevant to our needs, image of the user's current emotional state (*ES*). Following requisite 1, we adopted one of the most popular dimensional (i.e. quantifiable) interpretations of emotions – Russell's Arousal and Valence dimensional theory [4].

ERM uses a two-layered system based on four distinct physiological sensor inputs (skin conductance, heart rate and facial electromyography at the cheek and brow muscles). The first classification layer uses several regression models to normalize each of the sensor inputs both across participants and experimental conditions, while also correlating each one to either Arousal or Valence. The second classification layer then employs a residual sum of squares (RSS) weighting scheme to merge the computed regression outputs into one optimal Arousal/Valence prediction (see Fig. 1) [3].
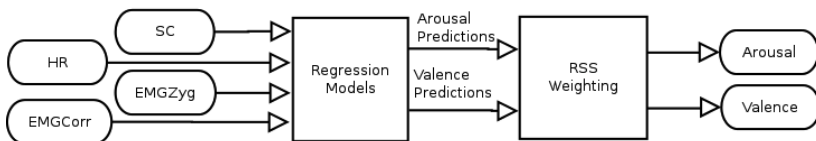
Fig. 1. High-level overview of the employed emotional classification system [3]

Details on how the method was designed, developed and validated can be found elsewhere [3]. The current paper only presents a high-level overview of the system as a part of the whole emotional reaction annotation process. Using this methodology, the authors have achieved convincing accuracy ratings – 85% for Arousal and 78% for Valence –, while at the same time successfully predicting these states in a continuous manner and without requiring lengthy calibration or parameter tuning procedures. Given these properties we deemed the method appropriate for our needs and incorporated it in the designed tool.

Event Annotation Module (EAM)

Since a considerable percentage of our requisites (55%) were related with how to visualise and annotate the recorded material, we devoted a great deal of attention to the development of this module. To fulfil requirements 1 and 2, we decided to combine a custom video player with a time series graph drawing library (ZedGraph). Timestamps were logged for both the gameplay videos and the physiological data, which were then used to synchronise the emotional classification and gameplay video streams. The video player component was designed to allow the user to quickly skip through the video using a simple slider or to accelerate the video through a fast-forward and backwards button (requisite 3).

Fig. 2. The add event window, super-imposed on the developed tool

Regarding the event annotation process itself (requisites 4-5), we decided to limit the user input to the barest essentials in both terms of actions and input required. To insert a new event, the user needs only to perform a right-click on the video player window or right-click on the emotional classification time plot and choose "Add new event" (see Fig. 2). This will add a new event at the current time and bring forth a pop-up form where the user can choose which event took place and any subjective commentary deemed relevant (requisite 6).

Emotional Response Identification Module (ERIM)

The final component in our tool is the emotional response identification module, which is responsible for performing the basic triangulation between the recorded events and the ensuing responses in the Arousal and Valence dimensions (requisite 7). The triangulation process was automated by developing a simple local peak detection algorithm. Let $c=[c_1,c_2,\ldots,c_n]$ be the continuous, uniformly sampled emotional state

classification signal for a dimension of the AV space. Furthermore, consider the signal to be smoothed using the unimodal kernel with compact connected support and unit action $w\Upsilon(t) \geq 0$, and $\Upsilon > 0$ bandwidth parameter through the following process:

$$c_\gamma(t) = w_\gamma(t) * c(t) = \int_{-\infty}^{\infty} w_\gamma(t-s)y(s)\, ds$$

The peak detection process occurs in parallel for both dimensions and is contained in a standalone iteration for each event e_i, within a time region $\Omega = [max(T(e_{i-1}),T(e_i)-\alpha,\ min(T(e_i)+\beta,\ T(e_{i+1})))]$, where T is the mapping function between an event and its corresponding timestamp. Moreover, both α and β are parameterisable event horizon variables (in this paper $\alpha=2$ and $\beta=10$, as determined by an empirical analysis of the available data). For each iteration of the peak detection process, the smoothed signal $c\Upsilon(t)$ is taken and the maximal peak $|p|$ is extracted from a set of candidate peaks $\tilde{P}$:

$$\tilde{P} = \left\{ t \in \Omega : \dot{c}_\gamma(t) = \frac{d c_\gamma(t)}{dt} = 0, \left| c_\gamma(t) - c_\gamma(T(e_i)) \right| \geq \varphi \right\}$$

Where φ is a minimum absolute local variability threshold, $\varphi = (\mu_{e_i} + 2\sigma_{e_i})$, such that μ_{e_i} and σ_{e_i} denote the mean and standard deviation values of the considered AV dimension in the processed event's time region. A light validation of the proposed algorithm using a total of 210 gameplay-related events randomly sampled from all the participants revealed that the algorithm was able to identify local maxima and minima with an approximate success rate of ~93%. Although further validation is required, this seems to hint that our approach is adequate for our described needs. In future work we will present a larger validation of the algorithm and more in-depth description of the system's interface.

3   Conclusions

Current UX research methods are unable to perform in-game evaluations without disrupting – and thus potentially contaminating – the gameplay experience. Moreover, emotional state classification methods are difficult to integrate in these studies due to their complex development nature and technical skillset. The methodology presented in this paper has the potential to contribute to a wider accessibility of emotional response studies by, not only easing the aforementioned issues, but also by removing the necessity of developing standalone emotional state detection systems – which in itself contributes to a standardisation and comparability of the annotation process.

References

1. Matias Kivikangas, J., Nacke, L., Ravaja, N.: Developing a triangulation system for digital game events, observational video, and psychophysiological data to study emotional responses to a virtual character. Entertainment Computing 2, 11–16 (2011)
2. Mirza-babaei, P., Nacke, L.E., Fitzpatrick, G., White, G., Mcallister, G., Collins, N.: Biometric Storyboards: Visualising Game User Research Data. In: CHI 2012 Extended Abstracts on Human Factors in Computing Systems, pp. 2315–2320 (2012)
3. Nogueira, P.A., Rodrigues, R., Oliveira, E.: Real-Time Psychophysiological Emotional State Estimation in Digital Gameplay Scenarios. In: 14th Conference on Engineering Applications of Neural Networks (EANN) (to appear)
4. Russel, J.A.: A Circumplex Model of Affect. Journal of Personality and Social Psychology 39, 1161–1178 (1980)