Affective games: a multimodal classification system

Salma Hamdy
David King

This conference paper was presented at the 19th annual European Game-ON Conference (GAMEON'2018) on Simulation and AI in Computer Games, Abertay University, Dundee, UK, 18-20 September 2018

AFFECTIVE GAMES: A MULTIMODAL CLASSIFICATION SYSTEM

Salma Hamdy and David King
Division of Computing and Mathematics
SDI, Abertay University
DD1 1HG, Dundee,
Scotland
E-mail: {s.elsayed, d.king}@abertay.ac.uk

KEYWORDS
Affective games, physiological signals, emotion recognition.

ABSTRACT

Affective gaming is a relatively new field of research that exploits human emotions to influence gameplay for an enhanced player experience. Changes in player's psychology reflect on their behaviour and physiology, hence recognition of such variation is a core element in affective games. Complementary sources of affect offer more reliable recognition, especially in contexts where one modality is partial or unavailable. As a multimodal recogntion system, affect-aware games are subject to the practical difficulties met by traditional trained classifiers. In addition, inherited game-related challenges in terms of data collection and performance arise while attempting to sustain an acceptable level of immersion. Most existing scenarios employ sensors that offer limited freedom of movement resulting in less realistic experiences. Recent advances now offer technology that allows players to communicate more freely and naturally with the game, and furthermore, control it without the use of input devices. However, the affective game industry is still in its infancy and definitely need to catch up with the current life-like level of adaptation provided by graphics and animation.

1. INTRODUCTION

Affective computing (AC) (Picard 1997) is the science that aims to design and develop emotionally intelligent machines. Such automated systems should process and interpret human emotions via analysing sensory data. An affective model cannot be generic as applications vary in emotion models, available information, input devices and user requirements. For example, health care systems may require intrusive sensors to collect very reliable data, while e-learning and games may not demand such optimality and may or may not require additional controllers (Szwoch and Szwoch 2015). Overall, an affect recognition system is typically a trained classifier and, regardless of the application or input, includes components of traditional supervised classification (Fairclough, 2009). Human-Computer Interaction (HCI) applications further require system adaptation according to the predicted user emotion. This extends to affective games (AGs) where the goal is to increase engagement by explicitly or implicitly altering the game in response to players' emotions.

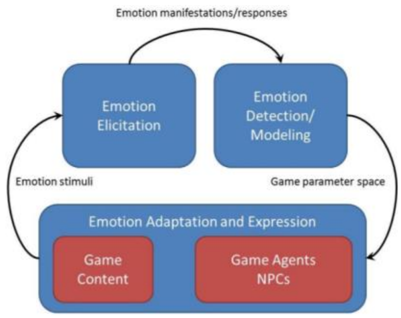

Figure 1. The realisation of the affective loop in games (Yannakakis and Paiva 2014).

Though a typical diagram of the affective loop in games (Fig. 1) does not reflect how the system infers the emotions, it implies the need to classify or estimate the response received from users (Novak et al. 2012). Very few affect-ware games truly reflect the concept of the full circle and are rather developed for academic research purposes. Commercial affective loops engage players' emotions through gameplay and other content in the development stage based on a representative player model (Adams 2014). This is problematic since individual players often differ from the average model, in addition to the rich spectrum of emotions experienced by players, which could change from sessions to session even for the same player, making it almost impossible to predict. However, this is likely to change with the advances in affect recognition techniques and input devices, allowing the capture of various information channels and more reliable predictions.

Fortunately, there are a variety of traditional classifiers that fit the task of emotion recognition and a large number of software libraries that make these classifiers available. Also, several emotion models and databases have been developed and standardised to an extent. Hence, the issue to consider often is what affective channels to acquire information from, and how to properly process them. Most attempts address the face as the main source of affect, while others involve speech, bodily and physiological signals. The latter has recently gained attention in the gaming context with the growth of affordable wearable technology and the well-established psychophysiological correlation (Christy and Kuncheva 2014). As observed by (Picard et al. 2001), physiological responses are translated to discrete psychological (emotional) states by a supervised classification pipeline. Furthermore, it is believed that commercial game publishers will start considering "psychophysiological hardware" in their next generation of game consoles (Valve Steam Box 2013).

A multimodal architecture was presented in (Hamdy 2016) as a generic model for affect-aware machines. It suggest a more reliable prediction by fusing different types of input information. This can naturally be extended to games and hence, a typical closed AG loop would include: multimodal emotion acquisition, modelling and identification of the collected signals via machine learning or statistical methods, and reflecting the decision back into the game engine to subsequently alter the game, ultimately taking into account the strength and type of the recognised affect (Christy and Kuncheva 2014). Variations of game adaptation includes dynamic difficulty adjustment (DDA), audiovisual content alterations, and affect-aware NPCs.

This paper discusses the external part of the AG loop as a multimodal recognition system, and reviews the different sources and methods of collecting affect information from players. In section 2, a number of modalities used as input to affective systems are discussed along with examples from the literature that employ these in games. Section 3 analyses these sources of affect in gaming context, and highlights relevant design issues of AGs as multimodal classifiers. Conclusions are presented in section 4.

2. AFFECTIVE INFORMATION

Attempting to improve classification tasks, it is recommended that multiple types of input from different modalities or different features from the same modality be combined (Gunes and Piccardi 2005). Hence, identifying psychological states from user biometrics requires that different types of measurements be provided simultaneously to allow one to verify the others (Drachen et al. 2010). The commonly used modelling approach for categorising emotions from mono- or multi-modal input is based on the arousal (high-low) and valence (positive-negative) dimensions in terms of the collected information. In addition to the apparent facial expressions and body movement, AGs often use monitoring modalities (Giakoumis et al. 2009) produced by the autonomic nervous system reflecting cardiovascular, electrodermal, or electrical activity in the human brain.

2.1 Behavioural

Vision channels hold the most informative data as humans tend to convey their feelings in a visual sense. The non-intrusive properties of cameras make vision-based systems more practical especially with the rapid advances in hardware and computer vision technology (Szwoch and Szwoch 2015). It is well-established that facial expressions and emotion mutually influence each other, hence the majority of affect recognition systems focus on face. Several features like Face Action Units or facial landmarks have been studied and benchmarked to model primary and secondary emotions in terms of selected dimensions. However, this is the least explored category in the literature with a limited number of studies addressing facial expression recognition in the context of games.

NovaEmötions (Mourão and Magalhães 2013) is a multiplayer game where players score by acting an emotion through facial expressions. The captured emotions are labelled using a multiclass Support Vector Machine (SVM) and the player with the closest expression wins a round. Authors claim the face images were captured in a novel and realistic setting despite the purpose being "act out an emotion" rather than spontaneously reacting to a stimulus. However, the experiment released a novel facial expression dataset of several emotions. Three AGs with linearly increasing difficulty were developed in (Bevilacqua et al. 2018) to investigate the relation between facial actions and heart rate, and player's emotional states. Expectedly, participants retained a neutral face for longer periods of time during the boring game parts. The study concluded that fusing the two cues is more likely to detect the emotional states. Authors in (Asteriadis et al. 2012a) used images of human faces and expressions in an attempt to assess the emotional state of a player. Player frustration and engagement as well as the challenge imposed by gameplay were used to alter the game in response. Other examples were previously discussed in (Hamdy and King 2017) to develop AGs through emotional NPCs that can believably respond to a player's facial affect.

Some affective expressions are reflected better through the body than the face. Cameras and motion detection devices enable the development of posture tracking techniques to construct models of body movement. The most common technique to capture motion is a suit with visibly trackable markers where posture is reconstructed by observing the subject with a camera and analysing the imagery. This is a well established technique widely used in film animation and could easily be functional in games. Alternatively, markerless optical systems are available with no special equipment needed, like Microsoft's Kinect.

A simple yet very effective five-dimensional representation of body expressions was introduced in (Caridakis et al. 2010) and proved to have a strong association with how humans perceive emotions in real environments, making them strong candidates for affective HCI systems including games. In (Savva and Berthouze 2011), a motion capture system was attached to subjects playing a Wii tennis game to identify their affective states from non-acted body movements. The most dominant motions were used with a neural netrowk (NN) classifier to identify eight emotions. Similarly, Kleinsmith et al. (2011) represented postures as rotations of the joints and assessed players in Wii sports games after winning or losing a point. Distance between body joints was used in (De Silva and Berthouze 2004) to recognise four basic emotions. Interestingly, the acted dataset of postures was labelled by observers from different cultures. The research in (Kapoor et al. 2004) examined non-acted postures through a multimodal system of facial expressions, body postures, and game state information. They reported the highest recognition accuracy from posture, although a limited description of the body was used. A system was proposed in (Gunes and Piccardi 2009) to identify emotions using a Hidden Markov Model (HMM) and a SVM to fuse facial and body cues to identify user affect. However the database did not include any real body pose information and was of a single subject.

Other vision-based modalities of player input that have been explored use pupilometers and gaze tracking, which are argued to be implausible within commercial development due to unreliability (sensitivity to distance, light and screen lamination) (Yannakakis et al. 2016). However, eye tracking is able to reveal information on attention from the duration of fixation, and hence is a good candidate for sensing player's engagement (Bradley et al. 2008; Xu et al. 2011)

Speech is one of the important behavioural modalities for detecting emotions. However, compared to facial expressions, emotions may not be captured as clearly in voice. In terms of vocal emotional dimensions, arousal is reflected by voice intonation and acoustics and has the strongest impact on speech, hence, can distinguish emotions better. Valence on the other hand is reflected by spoken words and is much harder to estimate from voice (Guthier et al. 2016).

Automatic speech recognition (ASR) is currently available on most low resource devices, smart phones, and game consoles, but mostly focus on the recognition of some context-dependant keywords. Although this is limited, it is a robust feature against possible interferences from game sounds, music and NPC voices in natural gaming environments. Hence, there is the trade-off of including a "heavy" continuous ASR engine in the game, or limiting the emotion analysis to a few affective words (Schuller 2016). It is important to note that even lower accuracies of ARS modules are proved to be sufficient to identify emotions from a word in a consistent context (Metze et al. 2010). In addition to words, nonverbal expression of emotions like laughter or groans convey a lot of information about the speaker's affective state, and can also be handled by the ASR engine.

Games have been used as means of eliciting emotions for data collection in speech research implying the rich spectrum of affect present in or by games (Schuller 2016). However, it is argued that a player is less likely to want to speak to the game (Jones and Deeming 2008) and only a few games truly made use of the ability to recognise emotion from speech.

A voice activated game for identifying four attitudes from childrens' speech was presented in (Yildirim et a. 2011). Spontaneous dialog interactions were carried out with computer characters and acoustic, lexical, and contextual features were captured. Interestingly, results showed that the selected features have varying performance with different assessed affective states and that fusion of all three cues significantly improved classification results.

Authors in (Jones and Sutherland 2005) developed a game with an acoustic recognition system to identify player's emotions from affective cues in speech and alter the behaviour of the game NPC accordingly. This was extended to a system capable of capturing 40 acoustic features from voice to assess five emotions where the character is better able to overcome obstacles based on the emotional state of the player.

In (Kim et al. 2004), affective speech and physiological signals were collected from players to elicit certain reactions in a pet NPC. A pre-selected set of features were used with a simple threshold-based classifier. Results showed improved accuracy when the two affective channels are combined. In a slightly different perspective, Rudra and Tien (2007) proved the feasibility of recognising voice emotions of a game character. Arbitrary utterances from the artificial Pidgin language was classified using a SVM to identify neutral and anger states of the NPC.

The work in (Alhargan et al. 2017) combined eye tracking with speech signals in a game that elicits controlled affective states. A SVM was used to classify emotions based on arousal and valence. Recognition results revealed eye tracking outperforming speech in affect detection and, when fused at decision level, the two modalities were complementary in interactive gaming applications.

2.2 Physiological

The Cognitive-motivational-emotive model assumes the human emotional state consisting of both affect and physiological response (Szwoch and Szwoch 2015). Visible affect can be controlled to an extent, but it is almost impossible for the average person to control physiological reactions. Furthermore, studies show that experienced players tend to stay still and speechless while playing (Asteriadis et al. 2012a), hence, the need for other forms of affect expressions. However, it is argued that in games, with enough practice, players' skills could allow them to control their physiological responses, converting an AG loop into a straight-forward biofeedback (Nacke et al. 2011).

The most popular biometric signals used in adaptive player-centric games are summarised in Table 1 with their correlation to emotion and feasibility in practice (Christy and Kuncheva 2014; Bontchev 2016; Garner 2016).

Table 1: The Common Physiological Signals for Affect Detection

Signal	Measurement/tool	Features
Electrodermal Activity (EDA) or Galvanic Skin Response (GSR) or Electrical Skin Response (ESR)	Electrical conductivity of the skin surface. A band between two fingers on either hand.	Reliable indicator of affective arousal like stress and anxiety. Simple and low cost. Common alone or combined with other techniques. Widely used for affect detection including in games. Easy to adapt well into games controllers. Suffers latency. Unsuitable for games with hand controllers, unless sensors are attached to the controller.
Electromyography (EMG)	Electrical activity from muscles. Non-invasive electrodes.	Vary across subjects and cultures. Need to be placed at various body locations.
Electroencephalography (EEG)	Electrical signals from the brain. Non-invasive electrodes.	Used in various contexts and superior for games due to portability, ease of use, temporal resolution and affordability. Able to detect presence of emotions and identify the discrete classes. Excellent for examining attempts to conceal or pretend emotions. Spatial resolution is relatively low and may be insufficient for complex emotion detection.
Respiration	Breaths speed. Respiration belt or sensors.	Not as accurate as other signals. Mainstream applications could be hindered as sensors are embedded into clothing.
Blood volume	Heart rate and blood oxygen. Optical technology clip.	Good indicator of affective arousal like stress and anxiety.
Temperature	Body temperature. Contact and contactless sensors.	Related to specific emotional states. Has been used in games. Sensitive to movement causing inaccuracies in collected data.

Typically, measurement of such signals requires standard hardware sensors borrowed from psychological research, which could be expensive and not suitable for active game play. Wearable sensors were introduced by (Picard and Healy 1997) for AC applications, which can be embedded in clothes or glasses making them fitting for AGs. Seamless sensors, where the user should not be aware of any interaction, come into contact with the body for a limited time through classical interfaces like mouse and keyboard. Some research attempts to incorporate traditional game controllers with such sensing ability. Scheirer et al. (2002) proposed a system that combines physiological data with behavioural data, namely mouse-clicking patterns, to build a HMM classifier of affective classification.

In (Christy and Kuncheva 2013), a fully functional mouse was designed with GSR and HR frequency measuring capability for capturing clean physiological signals from the player in real time. Rarely, an adaptive AG may use keyboard pressure as indication of changes in player effort or emotion during gameplay (Tijs et al. 2008). The game ‘‘Rush for Gold’’ (Bontchev and Vassileva 2016) used the GSR of the player to assess their arousal level and alter game componenets accordingly.

Attempts have been made to commercially produce multimodal affect-aware games. Companies had to limit their trials within the capabilities of existing sensing technology and emotion recognition algorithms. However some were involved in manufacturing the necessary hardware and software to include affective elements in their games. Christy and Kuncheva (2014) provide a tabulated historical survey of AC, specifically psychophysiological system developments, and industry trends with respect to producing commercial AGs. Though unsuccessful, some “retro” systems employed the affective concept in player input since the early 80’s with custom tailored sensing equipment. More recent AGs that made it to the commercial environment are found in (Kotsia et al. 2016).

3. MULTIMODAL AFFECTIVE GAMES DESIGN

Integrating AC into games involve interdisciplinary fields; signal processing, machine learning, and input from psychology. The sensing part of the AG loop is about fitting a supervised classifiers component into the design and development of the game while maintaining real-time performance and adaptation. Below, we discuss the feasibility of the modalities in section 2 in games context.

3.1 Modalities

Face and Body

Classification is one of the difficult tasks to automate, and when visual data is involved, this becomes more problematic. Facial and body movement prove very rich and useful for examining emotion expressions but are computationally expensive and time consuming (Kaplan et al., 2013). Camera-based modalities are highly within reach and do not require expensive equipment. However, the majority of vision-based affect detection systems cannot operate well in real-time (Zeng et al. 2009) and often require a well-lit environment that is not always available or preferred by gamers, in addition to posing privacy issues. Fortunately, this can be resolved to an extent by the advances in computer vision and hardware, and the increasing number of available vision-based emotion detection software. A rich collection of databases exist of facial expressions for primary and secondary emotions. However, due to the difficulty of obtaining natural emotions in experimental settings, only few databases exist that show spontaneous emotions. Real expressions could differ greatly from posed ones in terms of facial geometry and timing. This deems the majority of exiting datasets unsuitable for real-time generalisation especially in game environments where natural emotion is key. Furthermore, the validity of vision-based affect is highly subjective since observations vary between cultures, races and social environments (Jack et al., 2012). On another level, open space or collaborative games may require several cameras posing even further challenges of stereo-vision, real-time detection, and handling several occlusions due to space limitations and presence of several people.

Speech

As with visual cues, speech is a highly accessible real-time and unobtrusive modality, yet it is only applicable for games controlled by speech which are not that common. That is why few games up to this point make use of the ability to recognise emotion from speech, in addition to environmental audio posing additional challenges (Schuller 2016). Speech signals may not require as much processing power as visual cues and it is an advantage that sound recognition has been employed in HCI for quite sometime, and with reliable performance. A rich number of affective speech resources are available although only a few cover different age groups with realistic spontaneous emotions. However, this is still missing for many languages and cultures. Similar to facial expressions, speech emotions are obtained by recording performing actors to acquire intense clean samples avoiding background noise that accompanies ordinary voice samples. The content is often scripted and meaningless for emotion detection as opposed to natural speech where some emotions appear more than others depending on mood. This increases the generalisation error of the trained detector in real environments. Furthermore, the validity of voice-based labelling in realistic recordings is highly subjective and prone to disagreement (Guthier et al. 2016). It is also worth noting that most datasets and ASR systems focus on verbal content rather than “animated” sounds like laughter and sighs, which seem to be the more common in a game environment.

Physiological signals

Even though the core technology for physiological signals is well founded and developed, hardware for affective gaming is still not widely available. Furthermore, most of these signals respond to other external/internal factors such as subject’s health, physical condition, temperature, etc., deeming them unfit for the usual computer usage, not to mention gaming. Since video games mostly require active players, affective input to AG must be comfortable and intuitive as the sensors should not hinder player enjoyment of the game. AG would benefit best from seamless contact sensors, hence the need for them to be populated outside testing labs and into affordable consumer devices (Picard 2010). Nevertheless, great development is witnessed for non-intrusive low-powered sensors for remote collection of physiological and behavioural data from people. In addition, for existing hardware, real-time collection can be done through comfortable affordable wristbands and stored on local devices for further processing (Yannakakis et al. 2016). Furthermore, several computer manufacturers are considering embedding physiological sensors into game controllers (Szwoch and Szwoch 2015); Valve and Sony have implied that EDA and HR could soon be incorporated into standard controllers (Christy and Kuncheva 2014). A major leverage physiological signals have over other modalities as Table 1 indicates, is that they have been widely used in AC and games research and proved reliable indicators/classifiers of real-time emotions. Also, they do not lack generalised data, and are robust across the populous.

3.2 Model

Affect recognition systems, being a multimodal type of classifier, are more likely to incorporate methods applied in machine learning (ML) applications. Acquiring rich amounts of data from different affective signals seems appealing as it helps improve recognition and complement situations where some signals are not available. However, collecting physiological signals is subject to standard pre-processing and noise removal methods. Moreover, incompatibility, dimensions, and fusion of the collected signals present further challenges. Research in (Al Osman and Falk 2017; Calvo and D'Mello 2010) analyses automatic multimodal affect recognition and the challenges imposed by the need to acquire, process and fuse different types of data. Games pose additional challenges with respect to the ML model components, as many factors affect the collected data that not even carefully designed environments can eliminate without affecting player experience (Yannakakis et al. 2016).

Input signals

Most relevant sensors used in the gaming context are highly intrusive affecting the quality of gathered data. In addition, the fast-paced rich data from games may reflect rapid movement and quick alteration in emotions which may not be accurately captured or may even be missed. In general, physiological responses are affected by factors like mood, age, health in addition to external elements. When recording, it is often needed to offset the signal before modelling to calibrate the interaction model and eliminate subjective biases (Picard et al. 2001). This means a user will be recorded for a short resting time before any interaction, which may not be feasible for players. Nevertheless, it could be a suitable start for AG to exploit player dependent classifiers for better prediction. The tutorial level usually used to familiarise the player with the game, controls, and characters, can be exploited to calibrate the system to expressions of the specific player. This can also dynamically train AI companions to be accustomed with this player's forms of affect, hence more aware and believable in their responses. Surely, this raises feasibility issues and poses more constraints regarding system resources, game design and adaptation.

Features

Due to the rich affective interaction and the varying types of emotions experienced in games, the produced signals are complex and non-trivial to sample. Some extra features may need to be engineered for better distinction of displayed emotion. Standard extraction methods may suffice for AC applications, but for games, research shows that other complex methods such as sequence mining and deep learning offer richer representations of affect in games (Yannakakis et al. 2016). To reduce computational effort of training and real-time performance, it is best if the model is based on a minimal number of features that yield the highest prediction accuracy. Dimensionality reduction methods like principla component analysis (PCA) and Fisher's linear discriminant analysis (LDA) are all applicable, but current work in AG focussed so far on sequential forward selection, sequential backward selection and genetic search-based feature selection (Martínez and Yannakakis 2010). Another important issue to consider is the sampling rate. Most studies use an event-based approach where important game events determine the response time window that features are extracted from (Ravaja et al. 2006; Kivikangas et al. 2011).

Modelling (Classification)

Mapping features to emotions primarily depends on the representation model of emotions. If classes or annotated states are used to model player's affect, any of the traditional ML algorithms can be used to build an affective classifier. On the other hand, if a pairwise preference (rank) format is used, the problem becomes a preference learning (Yannakakis 2009). Dynamic models of player behaviour can be used to infer emotion in real time and induce appropriate emotions during gameplay (Bontchev 2016), however, Novak et al. (2012) conclude that the majority of adaptive physiological systems use static data fusion methods. The practical challenges result in emotions being identified with a widely varying accuracy (51%–92% according to (Nicolaou et al. 2011)) over the literature. Nevertheless, it is fair to say that a margin of error is allowed in games as an entertainment media (Christy and Kuncheva 2014). If the AG convinces players it is recognising and interacting with their emotions, then occasional misclassifications should not have a significant impact on the player's experience. This can relax the design constraint put on the system, especially for commercial products.

4. CONCLUSION

Although game developers have traditionally focused their efforts on improving the graphic quality of games, speculations is that the advancements in graphics will plateau, forcing them to discover new ways of adding attraction to their games. This is expected to open a commercial perspective for AG (Christy and Kuncheva 2014; Lara-Cabrera and Camacho 2018), which are basically classification systems with a variety of biometrics preferred as input.

Behavioural affective inputs are highly accessible but add the traditional challenges associated with audiovisual data processing and hence, require robust algorithms with higher generalisation level. Besides, with games being a global entertainment industry, cross-cultural and social experiences influence on emotions must be addressed (Kleinsmith et al. 2006; Sauter et al. 2010). Ethical implications arise when the game requires to audiovisually record players consistently, which are barely addressed in the literature.

Physiological signals offer a commonly acceptable alternative. Contact-based sensors produce a wider range of reliable, objective and quantitative data (Guthier et al. 2016). However, most existing biometric sensors are rather impractical and highly intrusive for interactive applications and some are still very costly for a broad use in gaming. Also, wearable devices can obscure a significant part of the face/body and influence players to exhibit unusual behaviour, even subconsciously, which may affect interaction and subsequent actions. In such a context, information from different channels is required.

Studies show that behavioural and physiological signals can be used to model players state continuously during interactive gameplay without interruptions, making the gathered data more temporally reliable as opposed to post-game interviews and questionnaires (Mandryk et al. 2006). Although it is evident from the literature that combining modalities of different types increases classification precision, novel methods for modelling/predicting interactions are required and efficient fusion of multimodal data remains an open problem. Nevertheless, while reliable recognition seems required, independent of external factors or personality profiles, Christy and Kuncheva (2014) suggest that AG should not exclusively rely on accuracy of emotion recognition. Clever game design can reimburse misclassifications for an uninterrupted game experience.

The most obvious way to represent emotion computationally is as labels for a limited number of discrete emotion categories. This scheme is easy to implement, but may be too general to be useful. Samples of affective data are often obtained from laboratory experiments with limited context, mostly of acted postures or stereotypical expressions (Kotsia et al. 2016). Picard (2000) highlighted the common emotions experienced or expressed around computer games, and the significance of systems that can recognising such affect from players. This can narrow the gap in HCI with development of more user-centred systems (Hudlicka 2003) when trained on emotions more likely related to gamers. However, emotion recognition is mostly done to standard predefined classes as spontaneity is an extra challenge (Kotsia et al. 2016). The experimental research is often done in heavily controlled environments limiting its chances of being deployed in practice, and results of AG research conducted in commercial settings are rarely published.

According to (Borod et al. 1998), the valence hypothesis suggests that there is a difference between processing and displaying positive and negative emotions. Hence, it may be obvious not to treat all basic emotions equally as it is less likely that all emotions will occur with the same probability in every day life. This however, could be slightly different for games as the genre, content or level are most likely intended to elicit particular affective states. In relevance, one thing to consider with affect-aware games is signal habituation (Sokolov 1963). Getting too familiar with the stimulus, such that bodily reactions tend not to be triggered as much, is a phenomenon commonly observed with experienced gamers or people who spend a lot of time on the same game or level. Successful interaction design should be dynamic enough to offer ranges of stimuli and keep the game exciting (Garner 2016).

It also worth noting that the majority of research in AG addresses single player scenarios. Physical space limitations are understandable, in addition to the added complexity of having to track and process biometrics of multiple players in a virtual environment, while keeping up system performance. Moreover, modelling multiplayer free interaction and how it influences their subsequent emotions is still a novel field of research (Kotsia et al. 2016). Although emotion recognition requires to be a real-time application with reasonable resources and ability to run on local platforms, it is a huge advantage to be able to distribute the recognition between the local console and a server (Schuller 2016).

Although home consoles do not by default incorporate biometrics, research shows that interest in biofeedback applications is growing and it is anticipated that in ten years, biometrics within games will become mainstream (Garner 2016). This move should inspire the game industry to consider design and development of AG loops in their products. It is argued that the future of affective gaming lies in more sophisticated, smaller, sensorless, noise-free devices (Kotsia et al. 2016; Christy and Kuncheva 2014). Fitting affective input devices and fast reliable pattern recognition algorithms in a game, while maintaining the desired game adaptation, is the biggest challenge for AG, especially in products affordable to the average player.

REFERENCES

Adams E. 2014. *Fundamentals of Game Design*. New Riders Publishing, USA.

Al Osman H. and T.H Falk. 2017. “Multimodal Affect Recognition: Current Approaches and Challenges”. In *Emotion and Attention Recognition Based on Biological Signals and Images*, S.A. Hosseini (Eds.). IntechOpen, 59-86.

Alhargan A.; N. Cooke; and T. Binjammaz. 2017. “Multimodal Affect Recognition in an Interactive Gaming Environment using Eye Tracking and Speech Signals”. In Proceedings of the 19th ACM International Conference on Multimodal Interaction (Glasgow, UK, Nov.13-17). ACM, N.Y., 479-486

Asteriadis S.; K. Karpouzis; N. Shaker; and G.N. Yannakakis. 2012. “Does your Profile Say it All? Using Demographics to Predict Expressive Head Movement During Gameplay”. In *20th Conference on User Modeling, Adaptation, and Personalization (UMAP 2012) Workshops* (Montreal, Canada, Jul.16-20) User Modeling Inc. 11.

Asteriadis S.; N. Shaker; K. Karpouzis; and G.N. Yannakakis. 2012. “Towards Player’s Affective and Behavioral Visual Cues as Drives to Game Adaptation”. In *LREC Workshop on Multimodal Corpora for Machine Learning* (Istanbul, Turkey, May 22). LREC, 6-9.

Bevilacqua F.; H. Engström; and P. Backlund. 2018. “Changes in Heart Rate and Facial Actions During A Gaming Session with Provoked Boredom and Stress”. *Entertainment Computing*, No.24 (Jan), 10-20.

Bontchev B. and D. Vassileva. 2016. “Assessing Engagement in An Emotionally-Adaptive Applied Game”. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (Salamanca, Spain, Nov.2-4). ACM, N.Y., 747-754.

Bontchev B. 2016. “Adaptation in Affective Video Games: A Literature Review”. *Cybernetics and Information Technologies* 16, No.3 (Sep), 3-34.

Borod J.C.; B.A. Cicero; L.K. Obler; J. Welkowitz; H.M. Erhan; C. Santschi; I.S. Grunwald; R.M. Agosti; and J.R. Whalen. 1998. “Right Hemisphere Emotional Perception: Evidence Across Multiple Channels”. *Neuropsychology* 12, No3 (July), 446-458.

Bradley M.M.; L. Miccoli; M.A. Escrig; and P.J. Lang. 2008. “The Pupil as A Measure of Emotional Arousal and Autonomic Activation”. *Psychophysiology*, 45 No4 (July), 602-607.

Calvo R.A. and S. D'Mello. 2010. “Affect Detection: An Interdisciplinary Review of Models, Methods, and their Applications”. *IEEE Transactions on Affective Computing* 1, No1 (Jan), 18-37.

Caridakis G.; J. Wagner; A. Raouzaiou; Z. Curto; E. Andre; and K. Karpouzis. 2010. “A Multimodal Corpus for Gesture Expressivity Analysis”. In *Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality Workshop* (Malta, May 18). LREC, 80.

Christy T. and L.I. Kuncheva. 2013. “Amber Shark-fin: An Unobtrusive Affective Mouse”. In *The Sixth International Conference on Advances in Computer-Human Interaction* (Nice, France, Feb.24-Mar.1). IARIA XPS Press, 488-495.

Christy T. and L.I. Kuncheva. 2014. “Technological Advancements in Affective Gaming: A Historical Survey”. *GSTF Journal on Computing (JoC)* 3, No4 (Apr), 32-41.

De Silva P.R. and N. -Berthouze. 2004. “Modeling Human Affective Postures: An Information Theoretic Characterization of Posture Features”. *Computer Animation and Virtual Worlds* 15, No3-4 (July), 269-276.

Drachen A.; L.E. Nacke; G.N. Yannakakis, G.N.; and A.L. Pedersen. 2010. “Correlation Between Heart Rate, Electrodermal Activity and Player Experience in First-Person Shooter Games”. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (L.A., California, July 28-29) ACM, N.Y., 49-54.

Fairclough S.H. 2008. “Fundamentals of Physiological Computing”. *Interacting with Computers* 21, No.1-2 (Nov), 133-145.

Garner T.A. 2016. “From Sinewaves to Physiologically-Adaptive Soundscapes: The Evolving Relationship Between Sound and Emotion in Video Games”. In *Emotions in Games*. K. Karpouzis and G.N. Yannakakis (Eds.). Springer International Publishing, Switzerland, 197-214.

Giakoumis D.; A. Vogiannou; I. Kosunen; D. Devlaminck; M. Ahn; A.M. Burns; F. Khademi; K. Moustakas; and D. Tzovaras. 2009. “Multimodal Monitoring of the Behavioral And Physiological State of The User in Interactive VR Games”. In *Proceedings of Multimodal Interfaces eNTERFACE’09* (Genoa, Italy, July13-Aug.7), 17.

Gunes H. and M. Piccardi. 2005. “Affect Recognition from Face and Body: Early Fusion Vs. Late Fusion”. In *IEEE International Conference on Systems, Man and Cybernetics* (Waikoloa, HI, USA, Oct.12). IEEE, 3437-3443.

Gunes H. and M. Piccardi. 2009. “Automatic Temporal Segment Detection and Affect Recognition from Face and Body Display”. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, No1 (Feb), 64-84.

Guthier B.; R. Dörner; and H.P. Martinez. 2016. “Affective Computing in Games”. In *Entertainment Computing and Serious Games*, R. Dörner et al. (Eds.). Springer International Publishing, 402-441.

Hamdy S. 2018. “Human Emotion Interpreter.” In *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016. IntelliSys 2016. Lecture Notes in Networks and Systems* 15, S. Kapoor et al. (Eds.). Springer, 922-931.

Hamdy S. and D. King. 2017. “Affect and Believability in Game Characters – A Review of The Use of Affective Computing in Games”. In *GameOn’17 – The Simulation and AI in Games Conference* (Carlow, Ireland, Sep.6-8). Eurosis, 90-97.

Hudlicka E. 2003. “To Feel or Not to Feel: The Role of Affect in Human–Computer Interaction”. *International Journal of Human-computer Studies* 59, No.1-2 (July), 1-32.

Jack R.E.; O.G. Garrod; H. Yu; R. Caldara; and P.G. Schyns. 2012. “Facial Expressions of Emotion Are Not Culturally Universal”. *Proceedings of the National Academy of Sciences* 109, No.19 (May), 7241-7244.

Jones C. and A. Deeming. 2008. “Affective Human-robotic Interaction”. In *Affect and Emotion in Human-Computer Interaction. Lecture Notes in Computer Science* 4868, Peter C. and Beale R. (Eds.). Springer, 175-185.

Jones C.M. and J. Sutherland. 2005. “Creating an Emotionally Reactive Computer Game Responding to Affective Cues in Speech”. In *Proceedings of the 19th British HCI Group Annual Conference* 2. Springer-Verlag, 1-2.

Kaplan S.; R.S. Dalal; and J.N. Luchman. 2013. “Measurement of Emotions”. In *Research Methods in Occupational Health Psychology: Measurement, Design, and Data Analysis*. R.R. Sinclair et al. (Eds.). Taylor & Francis Group, FL, 61-75.

Kapoor A.; R.W. Picard; and Y. Ivanov. 2004. “Probabilistic Combination of Multiple Modalities to Detect Interest”. In Proceedings of the 17th International Conference on *Pattern Recognition* (ICPR 2004) (Cambridge, UK, Aug.26). IEEE. 969-972.

Kim J.; N. Bee; J. Wagner; and E. André. 2004. “Emote to Win: Affective Interactions with a Computer Game Agent”. *Lecture Notes in Informatics (LNI)* 50, No.1 (Sep), 159-164.

Kivikangas J.M.; G. Chanel; B. Cowley; I. Ekman; M. Salminen; S. Järvelä; and N. Ravaja. 2011. “A Review of the Use of Psychophysiological Methods in Game Research”. *Journal of Gaming and Virtual Worlds* 3, No.3 (Sep), 181-199.

Kleinsmith A.; P.R. De Silva; and N.B -Berthouze. 2006. “Cross-cultural Differences in Recognizing Affect from Body Posture”. *Interacting with Computers* 18, No.6 (June), 1371-1389.

Kleinsmith A.; N.B. -Berthouze; and A. Steed. 2011. “Automatic Recognition of Non-acted Affective Postures”. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 41, No.4 (Aug), 1027-1038.

Kotsia I.; S. Zafeiriou; G. Goudelis; I. Patras; and K. Karpouzis. 2016. “Multimodal Sensing in Affective Gaming”. In *Emotions in Games*. K. Karpouzis and G.N. Yannakakis (Eds.). Springer International Publishing, Switzerland, 59-84.

Lara-Cabrera R. and D. Camacho. 2018. “A Taxonomy and State of the Art Revision on Affective Games”. In press to appear in *Future Generation Computer Systems*. (Jan), North-Holland. [Available at DOI: 10.1016/j.future.2017.12.056].

Mandryk R.L.; M.S. Atkins; and K.M. Inkpen. 2006. “A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, Apr.22-27). ACM, N.Y., 1027-1036.

Martínez H.P and G.N. Yannakakis. 2010. “Genetic Search Feature Selection for Affective Modeling: A Case Study on Reported Preferences”. In *Proceedings of the 3rd International Workshop on Affective Interaction in Natural Environments* (Firenze, Italy, Oct.29) ACM, N.Y., 15-20.

Metze F.; A. Batliner; F. Eyben; T. Polzehl; B. Schuller; and S. Steidl. 2010. “Emotion Recognition Using Imperfect Speech Recognition”. In *Proceedings of INTERSPEECH*, ISCA, 478–481.

Mourão A. and J. Magalhães. 2013. “Competitive Affective Gaming: Winning with a Smile”. In *Proceedings of the 21st ACM International Conference on Multimedia* (Barcelona, Spain, Oct.21-25). ACM, N.Y., 83-92.

Nacke L.E.; M. Kalyn; C. Lough; and R.L. Mandryk. 2011. “Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, Canada, May7-12). ACM, 103-112.

Nicolaou M.A.; H. Gunes; and M. Pantic. 2011. “Continuous Prediction of Spontaneous Affect from Multiple Cues and Modalities in Valence-arousal Space”. *IEEE Transactions on Affective Computing* 2, No.2 (Apr), 92-105.

Novak D.; M. Mihelj; and M. Munih. 2012. "A Survey of Methods for Data Fusion and System Adaptation Using Autonomic Nervous System Responses in Physiological Computing". *Interacting with Computers* 24, No.3 (May), 154-172.

Picard R.W. 1997. *Affective Computing*. The MIT Press, Cambridge, MA.

Picard R.W. and J. Healey. 1997. "Affective Wearables". *Personal Technologies* 1, No.4 (Dec), 231-240.

Picard R.W. 2000. "Toward Computers that Recognize and Respond to User Emotion". *IBM systems journal* 39, No3.4, 705-719.

Picard R.W; E. Vyzas; and J. Healey. 2001. "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, No.10 (Oct), 1175-1191.

Picard R.W. 2010. "Emotion Research by the People, for the People". *Emotion Review* 2, No.3 (June), 250-254.

Ravaja N.; T. Saari; M. Salminen; J. Laarni; and K. Kallinen. 2006. "Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation". *Media Psychology* 8, No.4 (Nov), 343-367.

Rudra T; M. Kavakli; and D. Tien. 2007. "Emotion Detection from Male Speech in Computer Games". In *TENCON 2007-2007 IEEE Region 10 Conference* (Taipei, Taiwan, Oct.30-Nov.2). IEEE, 1-4.

Sauter D.A.; F. Eisner; P. Ekman; and S.K. Scott. 2010. "Cross-cultural Recognition of Basic Emotions through Nonverbal Emotional Vocalizations". *Proceedings of the National Academy of Sciences* 107, No.6 (Feb), 2408-2412.

Savva N. and N.B. -Berthouze. 2011. "Automatic Recognition of Affective Body Movement in a Video Game Scenario". In *Intelligent Technologies for Interactive Entertainment. INTETAIN 2011. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 78*. Springer, Berlin, 149-159.

Scheirer J.; R. Fernandez; J. Klein; and R.W. Picard. 2002. "Frustrating the User on Purpose: A Step Toward Building an Affective Computer". *Interacting with Computers* 14, No.2 (Feb), 93-118.

Schuller B. 2016. "Emotion Modelling via Speech Content and Prosody: In Computer Games and Elsewhere". In *Emotions in Games*. K. Karpouzis and G.N. Yannakakis (Eds.). Springer International Publishing, Switzerland, 85-102.

Sokolov E.N. 1963. "Higher Nervous Functions: The Orienting Reflex". *Annual Review of Physiology* 25, No.1 (Mar), 545-580.

Szwoch M. and W. Szwoch. 2015. "Emotion Recognition for Affect Aware Video Games". In *Image Processing & Communications Challenges 6. Advances in Intelligent Systems and Computing* 313, R. Choraś (Eds.). Springer, 227-236.

Tijs T.J.; D. Brokken; and W.A. IJsselsteijn. 2008. "Dynamic Game Balancing by Recognizing Affect". In *Fun and Games. Lecture Notes in Computer Science* 5294, P. Markopoulos et al. (Eds.). Springer, Berlin, 88-93.

Xu J.; Y. Wang; F. Chen; H. Choi; G. Li; S. Chen; and S. Hussain. 2011. "Pupillary Response Based Cognitive Workload Index Under Luminance and Emotional Changes". In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, Canada, May7-12). ACM, N.Y. 1627-1632.

Yannakakis G.N. 2009. "Preference Learning for Affective Modeling". In *The 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII 2009)* (Amsterdam, Netherlands, Sep.10-12) IEEE. 1-6.

Yannakakis G.N. and A. Paiva. 2014. "Emotions in Games". In *The Oxford Handbook of Affective Computing*, R. Calvo et al. (Eds.). Oxford University Press, 459-471.

Yannakakis G.N.; H.P. Martinez; and M. Garbarino. 2016. "Psychophysiology in Games". In *Emotions in Games*. K. Karpouzis and G.N. Yannakakis (Eds.). Springer International Publishing, Switzerland, 119-137

Yildirim S.; S. Narayanan; and A. Potamianos. 2011. "Detecting Emotional State of A Child in A Conversational Computer Game". *Computer Speech & Language* 25, No.1 (Jan), 29-44.

Zeng Z.; M. Pantic; G.I. Roisman; and T.S. Huang. 2009. "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, No.1 (Jan), 39-58.

WEB REFERENCES

Exclusive Interview: Valve's Gabe Newell on Steam Box, biometrics, and the future of gaming. 2013. Available from: https://www.theverge.com/2013/1/8/3852144/gabe-newell-interview-steam-box-future-of-gaming. [Accessed May 2018].

AUTHORS BIOGRAPHY

SALMA HAMDY obtained her degrees in computer science from Ain Shams University in Egypt. She is a lecturer at Abertay University in Scotland. Her research interests include computer vision, AI, and affective computing.
Email: s.elsayed@abertay.ac.uk.

DAVID KING is a lecturer in Maths and Artificial Intelligence at Abertay University teaching on the Computer Games Technology and Computer Games Application Development Programmes. Email: d.king@abertay.ac.uk.