Emotion in Interactive Storytelling

Huiwen Zhao
School of Animation and Digital Arts
Communication University of China
Beijing, China
(86) 15652375838
zhaohuiwen@163.com

ABSTRACT
Emotion, as one of the major factors in creating engaging user experience, has been widely studied in interactive storytelling. Effort has been made mainly in two approaches, which are modeling emotions for believable virtual characters and capturing user emotions to generate affect-based interactive narrative. This paper provides an overview of emotion research in interactive storytelling in terms of these two approaches. In the end, current studies are concluded and potential future research is indicated.

Keywords
Emotion modeling, computational model of emotion, emotion recognition, affect-based interactive storytelling

1. INTRODUCTION
In recent decades, the convergence of traditional and interactive media has led to a promising field in digital entertainment, which is interactive storytelling. Since the first interactive fiction *Colossal Cave Adventure* was published in 1976, researchers have been working on storytelling using the computer as a new medium, whereby the user can have an impact on the evolution of the story. Therefore, interactivity, as "the computer's ability to take in voluntary or involuntary user input and to adjust its behavior accordingly" [1], opens up new horizons for storytelling.

A number of interactive storytelling concepts and systems have been proposed in an effort to create a dramatic and engaging narrative experience. Emotion, as one of the major factors in creating engaging user experience [2] and attracting users to interact with the story [3], has been widely studied in both academia and interactive entertainment industry. In particular, research efforts have been made in terms of two approaches: modeling emotions for believable virtual characters and capturing user emotions to generate affect-based narrative. Driven by the potential of modeling believable virtual characters to build empathic interactions, there has been significant expansion in recent years which applies psychological theories of emotion to developing interactive storytelling systems to create coherent and dramatic interactive stories [4]. A number of computational models of emotion were therefore designed to generate autonomous emotions for believable characters. On the other hand, user emotions have been considered significant to optimize user experience by adapting the narrative to user emotional responses. Therefore, recognition of user emotion and dynamic construction of affect-based narrative represent promising but challenging areas in interactive storytelling [5]. Although a number of techniques have been explored to recognize user emotion (e.g. [6][7][8]), few attempts have been made to incorporate these techniques into interactive storytelling systems and adapt stories to user emotional responses [9]. Therefore, two main questions remain to be answered: how to apply the methods of emotion recognition to interactive storytelling systems? How to adapt the story to user emotional states?

Given these considerations, this paper attempts to provide an overview of current interactive storytelling systems based on two approaches of emotion application: computational models of emotion for virtual characters and affect-based narrative generation. However, instead of giving an exhaustive number of examples, this paper attempts to highlight representative work that relates to these two approaches. At the beginning, an introduction of two well-developed computational models of emotion is presented. What follows is a review of interactive storytelling systems with respect to techniques and methods of user emotion recognition and narrative adaptation to user emotion. Finally, current studies of emotion in interactive storytelling are considered which indicates areas of potential research for the future.

2. EMOTION MODELING FOR VIRTUAL CHARACTERS
Emotion is one of the most central experiences of a human being [10]. It is subtle and has many facets. Therefore, different approaches have been developed to study emotions. In particular, inspired by the shift of research focus in psychology during the 1960s and 1970, a cognitive approach has been regarded as the central paradigm of studying emotions. Most cognitive theories of emotion examine the cognitive experience of emotion in terms of two processes: appraisal and coping. Appraisal is defined as the individual's interpretation of their relationship with the environment or events. It is informed by the cognitive process. Coping refers to the individual's cognitive responses to the appraised significance of events [11]. It consists of cognitive and behavioral responses which can be perceived internally or externally by individuals in the form of physiological and behavioral change. In general, appraisal is considered as the central process of eliciting emotions. Most appraisal theorists converge on the idea that appraisal is the cause of emotion [10] [12] [13] [14]. Early on considerable research was conducted in

interactive storytelling with the focus on modeling emotions for characters through appraisal. However, recently research has tended to concentrate on how coping strategies direct the storyline. Based on the cognitive understanding of emotion, two well-developed computational models of emotion are applied to interactive storytelling.

2.1 EMA

Marsella and Gratch [11] designed a computational model of emotion-EMA (EMotion and Adaptation) to simulate naturalistic emotional experience in a computational environment. EMA focuses on the dynamics of emotion across a range of eliciting situations which is developed on the basis of Smith and Lazarus' [13] cognitive theory of emotion.

In the structure of EMA [11], a character's interpretation of the character-environment relationship is represented as a conjunction of propositions. Eight appraisal variables, *relevance, perspective, desirability, likelihood, expectedness, causal attribution, controllability and changeability* are used to evaluate each proposition. All appraisal variables are maintained in a data structure called the appraisal frame. There are multiple appraisal frames in EMA, and each appraisal frame is responsible for evaluating one proposition. As a result, the evaluation leads to multiple emotion labels (*e.g.* hope, fear, joy, surprise etc) and intensities. In order to select one emotion as the individual's explicit emotional state, a higher-level notion, *mood*, is introduced as the aggregation of all emotional intensities with the same emotion label across various appraised events. The final affective state is determined by the most recently accessed appraisal frame with the highest mood intensity.

Appraisal is a fast, parallel and automatic process. The appraised significance of events leads to two types of coping behaviors in EMA: emotion-focused and problem-focused coping activities. To achieve these two types of coping activities, a series of coping strategies are employed in terms of their impact on the individual's *attention, beliefs, desires or intentions*. Problem-focused coping activities are concerned with attention-related or intention-related coping strategies, such as taking action or seeking information. Their execution is preferable when the controllability of the situation is high and sometimes the changeability is high. However, emotion-focused coping activities are more likely to influence the agent's psychological state, such as beliefs and desire. Their execution is preferable when the controllability and changeability are low. Multiple coping strategies can be applied towards a given circumstance. However, they are executed in sequence. The consequences of coping also exert an influence on the environment and alter the character's interpretation of character-environment relationship, which leads to a dynamic and cyclic emotional appraisal and coping process.

2.2 FearNot!

FearNot! (Fun with Empathic Agents Reaching Novel outcomes in Teaching) is a narrative based interactive storytelling system which is designed for anti-bullying education. It allows the user (children aged 8-12) to explore what happens in situations without placing themselves personally in a threatening environment. In FearNot!, the user plays the role of an 'invisible friend' who gives advice to the victim (virtual character) in order to influence the victim's behaviors. FearNot! attempts to create an empathic relationship between the user and the victim to get the user involved. To achieve this, it is important to model believable characters which can be affected by the event happening to them and react to the event in terms of their emotional states and personality automatically.

Basically, there are five modules to work through in FearNot![15][16]: *sensors, appraisal, emotional state, coping* and *effectors*. *Sensors* are used to perceive the environment, including events, objects etc in the world. After receiving the perceived information, the character *appraises* its significance based on the goals, intentions and plans. The consequence of appraisal triggers the appropriate *emotional state* which leads to a set of *coping* strategies. The coping strategies are turned into voluntary behaviors and carried out through *effectors*. The central part of the architecture is appraisal and coping. In particular, there are two distinct levels of both appraisal and coping in FearNot!: the reactive level provides a fast mechanism to appraise and react to a given event; while the deliberative level allows for a much more complex appraisal, and coping behavior takes a longer time.

The emotion theory adopted in FearNot! is that of Ortony, Clore and Collins (OCC)[10]. Accordingly, appraisal at the reactive level is implemented by a set of appraisal variables triggered by the event, agency and object. It generates most types of OCC emotions (*e.g.* Fortune of Others emotions, Well-being emotions, Attribution emotions, Attraction emotions [10]), and leads to the character's action tendencies. At the deliberate level, goals are introduced to appraise events together with appraisal variables. Therefore, prospect-based emotions, *e.g.* fear and hope, tend to be generated at the deliberate level. Two types of goals are included in FearNot!: active-pursuit goals which are goals that the character actively attempts to pursue, and interest goals which are those that the character wants to happen but have not happened yet. In addition, five emotion attributes are assigned for each emotion type: *valence* defines whether the emotional response is positive or negative; *target* denotes the name of the agent or object to which the emotion is directed; *cause* states the event or action which caused the emotion; *intensity* is the intensity of emotions; and *time-stamp* records the moment in time when the emotion is evoked.

The emotional state, resulting from the appraisal process, leads to coping. At the reactive level, action tendency is triggered as an impulsive reaction to a particular emotional state without considering any oriented goals, such as crying when the victim is bullied. A set of rules are used for action selection. These rules consist of preconditions and elicited emotions to execute the particular actions. At the deliberate level, two kinds of coping, which are borrowed from EMA are defined: problem-focused coping and emotion-focused coping. Problem-focused coping involves actions and plans to achieve goals, while emotion-focused coping alters the character's interpretation of the environment. Emotion-focused coping usually happens when the problem-focused coping has a low chance of success, and is used to lower strong negative emotions. Three strategies have been used for emotion-focused coping: *acceptance, denial* and *mental disengagement*. *Acceptance* is selected when the current plan fails to achieve the goal and the character does not attempt to improve it. *Denial* is selected when the chance of experiencing a threat is very low and the character ignores its effect. *Mental disengagement* is selected when the *acceptance* strategy is applied and the character lowers the importance of goals. The action selection and coping mechanism are implemented in a planner in FearNot!.

This section has introduced two computational models of emotion- EMA and FearNot!. These two models have been evaluated as effective approaches to create believable and empathic virtual characters [11][17]. However, when they are applied to an interactive storytelling system, emotions modeled by EMA and FearNot! only consider the character's pre-authored personality, including his/her goals and desires, but fail to consider the influence of the user's interaction on the character's emotional states and the development of the story.

3. AFFECT-BASED INTERACTIVE STORYTELLING

In recent decades, interactive storytelling has achieved great development, especially in the field of digital games. The number of game players and the demographic diversification have been boosted, which leads to a growing need for a *closed affective loop* [18]: producing enriched emotional experience, recognizing user emotion and directing the narrative according to user emotion. In particular, affective gaming which aims at creating a new type of game experience by adapting the game to the player's emotions has become a research focus in recent years [19]. In order to increase the user's level of enjoyment and engagement, it is necessary to allow the user to express his/her emotional states and dynamically adapt the character's actions and storylines along with the user's emotions. Given this consideration, a number of attempts have been made to incorporate user emotion into interactive storytelling systems.

3. 1 User Emotion Recognition

Detecting user emotion is the first step to incorporate user emotion into interactive storytelling. Techniques to capture and recognize user emotion have been studied and applied in different areas in human-computer interaction. As emotion is one type of user experience, methods of user experience evaluation can also be applied to recognize user emotion, which has been categorized into three approaches: a subjective approach, an objective approach and a game-based interaction approach [18].

The subjective approach captures emotions via directly asking the user questions about his/her emotional responses to the story. Two types of questions, open questions and closed questions are often used. Specifically, open questions allow the user to give any responses about her/her experience without restriction, while closed questions constrain the user's responses to a set of choices which are provided by the author, and the user is only allowed to select from these choices.

The advantage of open questions is that it gives the freedom to the user who can use any word or phrase to express his/her experience. However, it may end up with various user responses which add difficulties to emotion recognition and narrative adaptation. Roberts and his colleagues [6] suggest using explicit and subtle questions to elicit categorical emotional responses and assigning them with emotional labels. In their storytelling system, open questions, such as "How do you feel about…?" or "what would you like to do in response to…?" are inserted into the narrative process at several specific decision-making points. The user's responses to these questions are categorized using emotional labels such as 'satisfied' or 'disgusted'.

Compared with open questions, closed questions are easy to analyze user emotional states. However, one problem with closed questions is the user may not find his/her emotional responses in the choice list that the author provides, which may confuse the user and not reflect his/her actual emotions [20].

The subjective approach to emotion recognition is straightforward and relatively easy to implement. However, since questions are usually inserted in the process of storytelling, the subjective approach can be intrusive and violate user engagement in the story [18][6]. In addition, a subjective approach relies on the user's retrospection to capture his/her emotions, but the user's memory may decline after interacting with the story for a while [18][20].

The objective approach captures user emotion based on behavioral, expressive and physiological modality of emotion. In particular, user behavioral and expressive responses are the most visible and noticeable approach to recognize emotion. For example, a smile face indicates the user is happy while fast tempo and higher vocal pitch indicates the user is in fear. In addition, physiological responses of the user also facilitate emotion recognition, such as galvanic skin response (GSR), eye movement (EOG), muscle movement (EMG) and heart rate variability (HRV) [20]. Since user emotion can be recognized from multiple modalities, the key question of the objective approach is to decide which modality is appropriate to apply in interactive storytelling systems.

Facial expressions reveal useful information to capture emotion. Recent research has reported a high accuracy to recognize certain emotions by capturing user facial expressions [7]. However, facial recognition relies on facial markers to extract facial features and requires the user to stay in a fixed position, which is obtrusive to user experience and inconvenient to apply to interactive storytelling systems [21]. The same problem also arises with physiological methods [21]. In this sense, speech, as a natural and real-time efficient approach to emotion recognition, has been used in interactive storytelling systems [18].

CALLAS (Conveying Affectiveness in Leading-edge Living Adaptive Systems) is an affect-based interactive project which detects user emotion from vocal expression. It provides a platform for interactive installation and interactive storytelling systems, such as *E-Tree, EmoEmma*. The core component of CALLAS is EmoVoice [22], a system for vocal feature extraction and classification. In particular, acoustic features rather than semantic features are captured and analyzed based on pitch, energy, pause, voice quality from user speech [8]. In addition, a classifier is trained using a database which has stored a set of emotion-inducing utterances recorded beforehand. For example, three emotional classifications have been incorporated in the interactive storytelling system *EmoEmma*: Neutral, PositiveActive and NegativePassive [23]. It has also revealed that two or three emotional classifications are suitable for real-time emotion recognition [23]. Due to the high uncertainty of user vocal input in an interactive storytelling system, strategies for coping with non-prototypical emotions are required. Vogt et al. [22] have found that interest, transport and ludic pleasure are the most frequently reported emotions by users. Therefore, a limited set of emotions related to pleasure and arousal are suggested to be trained as the base and a "garbage" class can be added for all other non-prototypical emotions.

Another approach to emotion recognition using natural language is *Façade* where the user can type his/her emotional responses *i.e.* surface text to influence the character's behavior and storyline [24]. Different from CALLAS, *Façade* focuses on the pragmatic

effects of user natural language input, *i.e.* how does user language input influence the story world. Therefore, emotion recognition in *Façade* cannot be isolated from storytelling. Instead, it is related to *story beats*, which is the smallest unit to construct the narrative in *Façade*, and also, it has a direct impact on the evolution of the story. Specifically, the surface text from user interaction is mapped into *discourse acts* based on a set of rules. There are currently 24 sets of discourse acts in *Façade*, such as *DAAgree? Char,* which means the user agrees with a character [24]. According to the complexity of sentences the user types in, surface text can be mapped to intermediate meaning representations based on rules and then combine all representations to produce the final discourse acts. For example, if the user types in "I agree with Grace", the surface text "I agree with" is mapped to the intermediate meaning representation "iAgree" and "Grace" is mapped to "iCharacter". These two intermediate representations are then combined to produce the discourse act *DAAgree? Char*. Rules play a significant role in deciding how to map surface text to discourse acts. Template rules and key word rules are often used in *Façade*. However, these two types of rules are easy to result in promiscuous mapping and false positive which means the system will recognize ungrammatical and meaningless user input as a discourse. Therefore, the system gives a tradeoff that the character can eke the meaning and give some interesting responses [24].

In addition to using natural language input, either vocal speech or surface text to capture user emotion, gaze-based techniques are also considered as a practicable way to recognize user emotion, especially in terms of socio-emotive aspects, such as interest and attention. A recent study has suggested using a simple web camera to process user head pose and eye gaze in real time [25]. Therefore, it allows the user interaction to be non-intrusive by sensors and also allows the user to gaze inside or outside the screen to be calculated. Technically, head pose is determined by calculating the displacement of the point on the middle of the inter-ocular line when the head is rotating. In order to make sure the calculation is valid, the fraction between the inter-ocular distance and the vertical distance between the eyes and the mouth is also captured, and only if it changes along with rotation, the detection of head pose is valid. Similarly, the eye gaze is estimated by the displacement of the iris center with regards to the points around the eye when the user is looking at different directions. Based on the results from gaze detection, a number of metrics have been defined to assess the user's interest and engagement level. Every object including the outside screen is attached to the *virtual attention object (VAO).* The level of attention of VAO refers to the history of how much and when the user's gaze is on the VAO. By comparing the attention paid to the objects which are relevant to user interaction and the objects which are not involved in user interaction, the user's level of interest is revealed.

In addition to the subjective and objective approaches, according to cognitive theories of emotion [13], user actions and behaviors during interaction are related to their emotional states through affecting their cognitive focus and processing patterns. In this case, any element from the interaction between the user and the storytelling system can also be used to identify user emotion. In a study conducted by Zimmermann et al.[7], all mouse and keyboard actions are recorded without obtruding user interaction in terms of nine aspects: number of mouse clicks per minute, average duration of mouse clicks, total distance of mouse movements, average distance of a single mouse movement, number and length of pauses in mouse movement, number of "heavy mouse movements" (more than 5 changes in direction in 2 seconds), maximum, minimum and average mouse speed, keystroke rate per second, average duration of key keystroke. Meanwhile, other methods, questionnaires and physiological measurements including respiration, pulse, skin conductance level and corrugators activity, are also used to detect user emotion in their study. By comparing results from different methods, the interaction-based approach to emotion recognition is proved as a simple but effective alternative to recognizing user emotion [7].

3. 2 Affect-based Narrative Adaptation

Once user emotion has been detected, a further question that needs to be answered is how to adapt the narrative to user emotion? Hudlicka [26] suggest that the adaptation of narrative depends on the specific context and objectives for the interactive storytelling system. Gilleade and his colleagues [19] proposed three levels of design heuristics for digital games: assist me, challenge me and emote me, which means the narrative should respond sympathetically to player frustration, alter the challenges to increase engagement and provoke the player- intended emotions. In particular, three approaches have been found in current interactive storytelling systems to adapt the narrative to user emotion.

First, narrative is adapted in order to elicit or avoid one specific author-defined emotion.

Blom and Beckhaus [27] presented an emotional storytelling model as an extension of current interactive storytelling systems. It parameterized the user's emotions and compared them with the author's ideal emotional value; the outcome was used to guide the story development. Therefore, two core components were introduced into interactive storytelling systems. One was the Emotion Tracking Engine (ETE) which was used to keep track of the user's emotional states. Another one was the Emotional Path Graph (EPG) which was a time dependent graph showing the ideal emotional experience for the user from the author's perspective. Blom and Beckhaus applied these two components to improve the interactive drama *Façade*. The story of *Façade* centers on the relationship of one couple. The user plays the role of the couple's friend and is invited to the couple's house for dinner [24]. However, when the user arrives at their house, the couple is having an argument. Through the user's interaction, the relationship between the couple changes. In particular, tension is chosen as the core emotion of *Façade* and is experienced by users. Moreover, the whole storyline is divided into a set of story segments. Each story segment is given an expected tension value. By tracking the difference between EPG and ETE, the story engine selects the best story segment to play in order to enhance or impair the user's current tension level.

A similar attempt to direct the story based on user emotion was made by Roberts et al. [6]. They inserted several questions into the course of storytelling and authored the story based on the user's emotional responses. When selecting a video to present to the user, if the system has a goal of eliciting emotion e_i, two kinds of user emotional states are tracked: the average emotional response of the users who have already seen the video (V_i), and the emotional response of a particular user (P_i). The system reasons how the user's emotional reactions are likely to be different from those expected. Therefore, if $V_i \approx P_i$, the video

elicited emotion e_i would be presented, if $V_i \neq P_i$, the system will search and determine if there is another video that may be better at evoking emotion e_i. This approach integrates the user's emotional responses to guide the development of a story. One disadvantage is that it uses questions to ascertain the user's emotion, which interrupts the story flow and is likely to negatively influence user experience [9].

Second, narrative is adapted based on real-time user emotional states.

Cavazza and colleagues [28, 9] introduced an emotional planner to improve their original character-based interactive storytelling system which allows user emotion to have an impact on virtual characters' feelings and behaviors. The story is based on Gustave Flaubert's *Madame Bovary*. Instead of describing a character's emotions at a cognitive level, a detailed description of characters' emotions by Gustave Flaubert has been used. Therefore, a number of the character's emotional states and related actions have been preserved in the system, such as *pride-of-having-a-lover; emboldened-by-love* and *feels-hatred-for-Charles.* The user's emotional input can be translated into modifications of the character's emotional states, which further alters the character's behavior and influence the narrative. EmoVoice is an important component to recognize user emotion through vocal speech. As noted in Section 3.1, it relies on the acoustic signal from the user's voice rather than semantic meaning. Based on combinations of valence and arousal, five emotional categories have been recognized: *NegativeActive, NegativePassive, Neutral, PositiveActive* and *PositivePassive* [9]. According to the narrative context and the character's expectation, the five categories can be interpreted differently to influence the character's emotional states. For example, when the character has a high expectation, the user's NegativeActive emotional input can be interpreted as feelings of disappointment with a high intensity. Therefore, a communicative action related to the character's expectation will be decided whether or not it should be executed accordingly.

Different from Cavazza et al.'s study, Zhao et al. [20] proposed an emotion-driven interactive storytelling approach from a cognitive perspective of understanding emotion. In particular, Smith and Lazarus' *cognitive-motivational-emotive* structure of emotion has laid a theoretical foundation for their research [13]. According to Smith and Lazarus, emotion evokes action tendency which is an urge to respond to the stimulator in a particular way. The nature of the emotion decides the nature of the action tendency, for example, to attack in anger, flee or avoid in anxiety. Furthermore, action tendency can be translated into corresponding actions based on a set of coping strategies. Therefore, in interactive storytelling, the user's emotional response to the story leads to a set of action tendencies. In certain narrative context, these action tendencies are turned into a specific character's action to interact with other characters. As emotion appraisal and coping is a dynamic process, the character interacts with other characters, which changes the relationship between the character and the environment and leads to new emotions, thereby, new actions for the character are produced [20].

Third, narrative is adapted to a specific user experience which is related to user emotion.

Emotion is an important dimension of user experience, especially for some conception, such as engagement [29][30][20]. Therefore, detecting user emotion and adapting the narrative accordingly helps to optimize user experience in interactive storytelling.

Yannakakis and Togelius [18] propose an experience-based procedural content generation approach which aims to produce game content and interactive narrative in real time according to user experience. To achieve that, four components have been incorporated in an interactive storytelling system: player experience modeling, content quality, content representation and content generator. The player experience modeling aims to assess user experience based on data collected from the user in terms of affective response, cognitive response and playing style. The player experience modeling is then used to evaluate the quality of digital games concerning suitability for use in games and capacity for reaching the desired emotional states. Once the aforementioned two steps have been done, a bottom-up search mechanism driven by computational heuristics of player experience will be executed to check the game content. Ideally, the content generator should be able to identify *if, how much and how often* content should be generated for a particular player based on his/her experience and even identify whether or not the player likes the notion of adaptation [18]. In this case, user emotion is part of the user experience together with user cognitive response and playing style. Therefore, user emotion does not play an explicit role in determining the evolution of the story, but in an implicit way.

4. CONCLUSION

Emotion research in interactive storytelling has attracted plenty of attention in recent decades. Two main approaches have been reviewed in this paper, which is modeling computational emotions for virtual characters and detecting user emotion to adapt the evolution of the story.

Sophisticated techniques from artificial and computational intelligence have been used to model autonomous characters with a focus on the characters' emotional states and responses. A number of computational models of emotion have been developed on the basis of cognitive theories of emotion which argue that emotion arises from two basic processes: appraisal and coping. Specifically, EMA is developed based on Smith and Lazarus' theory while FearNot! is based on Ortony, Clore and Collins' theory. Therefore, the main difference between EMA and FearNot! is different types of emotions can be generated because different appraisal process with different appraisal variables have been incorporated. Although a number of studies [4] [31] [32] have confirmed these models are critical to create believable synthetic characters and help the user feel empathized, the character's emotions are modeled only from the appraisal of character-environment relationships, *i.e.* narrative context without considering the user emotional responses to the story. Therefore, the user's emotions as a result of witnessing the character's emotions, however, cannot influence the character's emotional states. In this sense, the user can only gain an emotional experience through empathizing with the characters rather than through interacting with the stories.

In addition to computational models of emotion, research on affect-based interactive storytelling aims to take user emotional responses to the story as a driving force to influence the evolution of the story. To achieve that, user emotion recognition is the first step. Although techniques and methods to recognize user emotions have been widely developed, a set of requirements are needed for applying these techniques and methods to interactive storytelling. First, the method to recognize user emotion should be unobtrusive; second, in addition to prototypical emotions, non-

prototypical emotions can also be recognized in real time; third, user emotion is detected for the reason of generating affect-based narrative, so the techniques and methods should be suitable for drama management. After capturing the user emotion, mechanisms for adapting the narrative were also reviewed in this paper. Basically, current studies make efforts in three approaches. The first is to adapt the narrative in order to elicit or avoid an author-defined emotion. In this case, user real-time emotion serves as an "assistant" to help the author manage the story rather than serving as a driving force to move the story forward and motivate the user's interactions. Therefore, interactive storytelling systems in this approach follow a goal-oriented narrative mode *per se*. The evolution of the story aims to achieve a pre-authored goal *i.e.* to evoke a specific emotion, rather than pursuing a 'pure' hedonic experience by engaging the user in a dramatic social interaction without offering clear goals [33]. In contrast, the second approach aims to generate the affect-based narrative based on user real-time emotions without considering any pre-authored goal. Basically, two mechanisms are used in this approach for narrative adaptation: when the user is detected to have negative emotional responses, the narrative will be adapted to generate positive emotions; while when the user has positive emotional responses, the system will try to keep these positive emotions for the user by selecting plots which can elicit similar emotions. However, one question arising from this approach is whether or not it can give user a dramatic narrative experience? Since suspense is important for creating dramatic experience [34], does keeping user emotions positive help to engage the user with the story? This therefore inspires the need for further research to evaluate current affect-based interactive storytelling approaches, and to explore more affect-based storytelling approach which takes the user's real-time emotional responses as a driving force to move the story forward and optimize user experience.

5. ACKNOWLEDGMENTS

Many thanks to my PhD supervisors, Jian J Zhang, Siné McDougll and James Pope. They inspire my research interest in this area and encourage me to explore further. I also owe my thanks to Leigh McLoughlin who helped me do the final proofreading and correct some spelling mistakes.

6. REFERENCES

[1] Ryan, M. L., 2006. Avatars of story. Minneapolis: University of Minnesota Press.

[2] Bernhaupt, R., Boldt, A., Mirlacher, T., Wilfinger, D., Tscheligi, M., 2007. Using emotion in games: emotional flowers . In Proceedings of the international Conference on Advances in Computer Entertainment Technology (Salzburg, Austria, June 13 - 15, 2007). ACE '07, vol. 203. ACM Press, New York, NY, 41-48.

[3] Lazzar, N. 2004. Why we play games: Four keys to more emotion without story. http://www.xeodesign.com/whyweplaygames/xeodesign_whyweplaygames.pdf.

[4] Hall, L., Woods, S., Aylett, R., Newall, L., Paiva, A., 2005. Achieving empathic engagement through affective interaction with synthetic characters. In Proceedingsof the First International Conference on Affective Computing and Intelligent Interaction. Vol. 3784, Springer-Verlag, pp. 731–738.

[5] Yannakakis, G., Karpouzis, K., Paiva, A., Hudlicka,E., 2011. Emotion in Games. In *Proceedings of the 4th international conference on Affective computing and intelligent interaction*, volume part II, p. 497-497.

[6] Roberts, D. L., Narayanan, H., Isbell, C. L., 2009. Learning to Influence Emotional Responses for Interactive Storytelling. US: Association for the Advancement of Artificial Intelligence.

[7] Zimmermann, P., Guttormsen, S., Danuser, B., and Gomez, P. , 2003. Affective computing - a rationale for measuring mood with mouse and keyboard. Inter. J. of Occ. Saf. and Ergo. 9, 4 (2003), 539-551.

[8] Vogt, T., André, E. and Bee, N. 2008. EmoVoice – A framework for online recognition of emotions from voice. In *Proceedings of Workshop on Perception and Interactive Technologies for Speech-Based Systems*, Springer, Kloster Irsee, Germany, (June 2008).

[9] Cavazza, M., Pizzi, D., Charles, F., Vogt, T., André, E., 2009. Emotional input for character-based interactive storytelling. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, May 10-15, 2009, Budapest, Hungary

[10] Ortony, A., Clore, G., Collins, A., 1988. The cognitive structure of emotions. Cambridge University Press.

[11] Gratch, J., Marsella, S., 2006. Evaluating a computational model of emotion. Journal of Autonomous Agents and Multiagent Systems (Special issue on the best of AAMAS 2004), 11(1), pp. 23-43.

[12] Frijda, N. H., 1988. The laws of emotion. American Psychologist, 43, 349–358.

[13] Smith, C. A., Lazarus, R., 1990. Emotion and adaptation. In L. A. Pervin (Ed.), Handbook of personality: Theory and research (pp. 609–637). New York: Guilford Press.

[14] Ellsworth, P. C., Scherer, K. R., 2003. Appraisal processes in emotion. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), Handbook of the affective sciences (pp. 572–595). New York: Oxford University Press.

[15] Dias, J., Paiva, A., 2005. Feeling and Reasoning: a Computational Model for Emotional Agents. Paper presented at the Proceedings of 12th Portuguese Conference on Artificial Intelligence, EPIA 2005.

[16] Aylett, R., Louchart, S., Dias, J., and Paiva. A., 2005. FearNot! an Experiment in Emergent Narrative. In *Proceedings of the 5th International Conference on Intelligent Virtual Agents*, 2005.

[17] Louchart S., and Aylett, R., 2007. Building synthetic actors for interactive dramas. In *Proceedings of the AAAI Fall Symposium on Intelligent Narrative Technologies*, pp. 63–71.

[18] Yannakakis, G. N., and Togelius, J., 2011. Experience-driven procedural content generation. In *IEEE Transactions on Affective Computing,* 2011.

[19] Gilleade, K. M., Dix, A. and Allanson. J., 2005. Affective videogames and modes of affective gaming: Assist me, challenge me, emote me ACE. In *Proceedings of the 2005*

International Conference on Changing Views: Worlds in Play, 2005.

[20] Zhao, H., 2012. Emotion-driven Interactive Storytelling. PhD Dissertation, Bournemouth University, UK.

[21] Hudlicka, E., 2008. Affective Computing for Game Design. In *Proceedings of the 4th International North American Conference on Intelligent Games and Simulation* (GAMEON-NA), Montreal, Canada.

[22] Vogt T, André E, Wagner J, Gilroy S, Charles F, Cavazza M., 2009. Real-time vocal emotion recognition in artistic installations and interactive storytelling: experiences and lessons learnt from CALLAS and IRIS. In: *Proceedings of the ACII*, Amsterdam, September 10–12, pp 670–677

[23] Charles, F., Lemercier, S., Vogt, T., Bee, N., Mancini, M., Urbain, J., ... & Cavazza, M., 2007. Affective Interactive Narrative in the CALLAS Project. In *Proceedings of 4th international conference of Virtual storytelling: using virtual reality technologies for storytelling,* ICVS 2007, Saint-Malo, France.

[24] Mateas, M. and Stern, A. 2004. Natural Language Understanding in Façade: Surface Text Processing. In *Proceedings of Technologies for Interactive Digital Storytelling and Entertainment,* Darmstadt, Germany, June 2004.

[25] Peters, C., Asteriadis, S., Karpouzis, K., and de Sevin. E., 2008. Towards a real-time gaze-based shared attention for a virtual agent. International Conference on Multimodal Interfaces (ICMI), Workshop on A®ective Interaction in Natural Environments (AFFINE), Chania, Crete, October 2008.

[26] Hudlicka, E., 2003. To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Interaction*, 59, 1–32.

[27] Blom K., Beckhaus S., 2005. Emotional Storytelling. In: *IEEE Virtual Reality 2005 Conference*, Workshop "Virtuality Structure ", Bonn: IEEE, p. 23-27.

[28] Cavazza, M., Charles, F., Mead, S. J., 2002. Character-Based Interactive Storytelling. IEEE Intelligent Systems, 17 (4), 17-24

[29] Ryan, M. L., 2008. Interactive Narrative, Plot Types, and Interpersonal Relations. In *Proceedings of the 1st Joint International Conference on Interactive Digital Storytelling.*

[30] Peters, C., Castellano, G., & de Freitas, S., 2009. An exploration of user engagement in HCI. In *Proceedings of the International Workshop on Affective-aware Virtual Agents and Social Robots* (AFFINE'09), ICMI-MLMI'09, Boston, MA, USA, November 6, 2009. ACM, New York, NY.

[31] Mcquiggan, S. W., Lester, J. C., 2009. Modelling affect expression and recognition in an interactive learning environment. International Journal of Learning Technology. 4(3- 4): 216-233.

[32] Pizzi, D., Cavazza, M.A.:Affective storytelling based on characters' feelings., 2007. In Intelligent Narrative Technologies: Papers from the AAAI Fall Symposium, 111-118.

[33] Knickmeyer, R., and Mateas, M. 2005. Preliminary Evaluation of the Interactive Drama Façade. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 2005). Portland OR, April 2-7. New York: ACM Press, pp.1549 – 1552.

[34] Aristotle, 2300BC. Poetics. Translated: S. H. Butcher. Introduction: Francis Fergusson. New York: Hill and Wang, 1961.