

Exploring Passive User Interaction for Adaptive Narratives

Stephen W. Gilroy, Julie Porteous, Fred Charles, Marc Cavazza

Teesside University, School of Computing

Middlesbrough, TS1 3BA, UK

{s.w.gilroy, j.porteous, f.charles, m.o.cavazza}@tees.ac.uk

ABSTRACT

Previous Interactive Storytelling systems have been designed to allow active user intervention in an unfolding story, using established multi-modal interactive techniques to influence narrative development. In this paper we instead explore the use of a form of passive interaction where users' affective responses, measured by physiological proxies, drive a process of narrative adaptation. We introduce a system that implements a passive interaction loop as part of narrative generation, monitoring users' physiological responses to an on-going narrative visualization and using these to adapt the subsequent development of character relationships, narrative focus and pacing. Idiomatic cinematographic techniques applied to the visualization utilize existing theories of establishing characteristic emotional tone and viewer expectations to foster additional user response. Experimental results support the applicability of filmic emotional theories in a non-film visual realization, demonstrating significant appropriate user physiological response to narrative events and "emotional cues". The subsequent narrative adaptation provides a variation of viewing experience with no loss of narrative comprehension.

Author Keywords

Passive interaction, interactive storytelling, physiological, entertainment computing.

ACM Classification Keywords

H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities.

General Terms

Design, Experimentation.

INTRODUCTION

In Interactive Storytelling (IS) the goal is the development of systems whereby the evolution of a narrative and its

presentation to a user, typically through the use of 3D graphics, are influenced in real-time by user interaction. Most approaches have assumed that users will want to *actively* engage with the interactive system. This can be seen in systems such as Façade [23], where users actively take on the role of a dinner party guest; FearNot! [4], where users take on the role of an "invisible friend"; and Gunslinger [15], where users are immersed in a virtual environment. However, another paradigm would consist of seamlessly adapting the (unfolding) of the narrative to the perceived user experience.

Considering a user's affective experience allows a richer interaction with an IS system, tapping into the empathic processes leveraged in films, games and plays. Game-like interactive experiences are commonly utilized when integrating affective concepts, with notions of "playing" a character or intervening like a ghost, manipulating the characters' interactions to affect their emotional relationships [7]. The challenge is to utilize techniques that draw on filmic properties in ways that preclude traditional interactivity, in order to create compelling affective experiences that also allow consideration of the consequences of emotional responses.

We propose that one way of achieving this is *passive* interaction, whereby implicit indicators of affective responses can be used to adapt narratives and 3D visualizations in such a way that they foster diverse affective and empathic experiences within an IS system. Filmic emotional theories such as those proposed by Smith [34] and Tan [35] appear to give credence to this approach using cinematographic techniques on film, and the success of 3D animation such as *Toy Story* clearly suggests that being computer-generated does not preclude the elicitation of affective experience [17]. Smith proposes "emotional cues" that prime viewers into accepting certain prototypical conceptualizations of film-based affective experiences, and we reason that translation of cues and cinematographic techniques into an adaptive 3D visualization and interactive narrative can be facilitated by passive interaction techniques that allow us to assess the consequences of such priming.

In this paper we outline a framework for capturing certain forms of affective responses of users of IS systems accessible through physiological signals, passive and undirected, and translating them into concepts that can form the basis of an IS system which adapts narratives compatible with translation of cinematic experiences into non-filmic 3D graphics. For example, two easily captured physiological measurements, surface electromyography (EMG) and galvanic skin response (GSR) signals can be interpreted as indicators of users' affective response [5, 19, 32] to realizations of emotional “cues” in an IS system. Responses that indicate increases and decreases in physiological arousal are used to inform appropriate narrative progression and pacing. Valenced affective responses identify potential empathic relationships between viewer and characters, and the combination of both used uncover the most effective way to present a story, in order to foster additional responses in the viewer to provide a rich empathic and affective experience.

Our approach to passive interaction has been fully implemented in a prototype system, called PINTER (Passive INTERaction). The system is presented here, along with initial experimental results with real users. PINTER features a medical drama (*Arztserie*) as a narrative setting, chosen because it is a popular genre that has been shown to elicit audience reactions to both dramatic events and

character relationships (for example, the television series ER [6]).

RELATED WORK

In the area of entertainment computing, and video games in particular, there has been considerable recent interest in the use of physiological input. For instance, physiological signals have been interpreted as an indication of users' emotional state and used to trigger reactions [31], taken as an indicator of user interest in order to increase game-play challenge during periods of relative inactivity [1], and used for continuous modeling of user emotions during game interaction [22].

User studies have demonstrated the potential for “physiological games” [11], brain-computer interaction as an added modality in game design [26], truly affective video games [12] and user preference for direct physiological control in games [25]. In addition, it has been shown that skin conductance level can be used to predict user actions in games [20]. However, most computer games remain essentially task-based, the player needing to achieve specific objectives and/or complete discrete “levels”. We posit that this constant user involvement and the resultant emotional response may be of a different nature than is found in IS—one that is more task and performance related and less aesthetic. In the area of IS user emotions can arise more readily from the filmic aspects rather than the task-

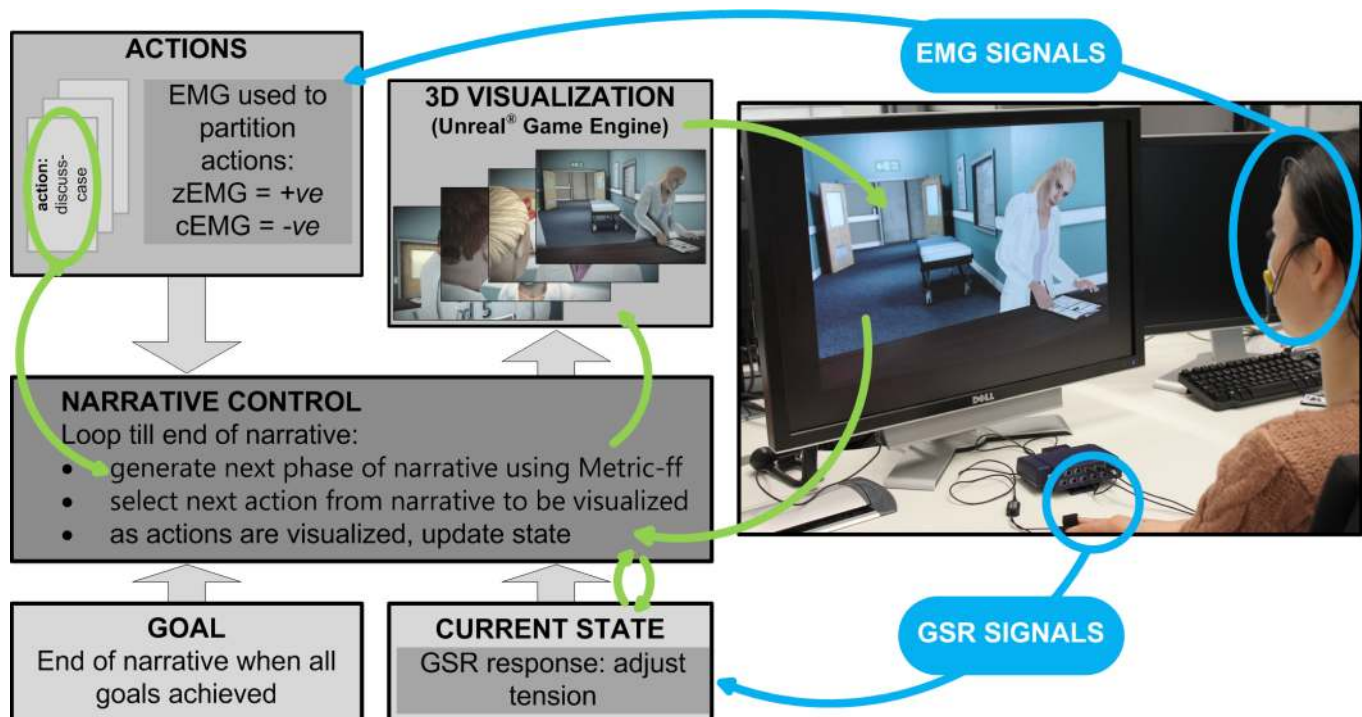


Figure 1. Overview of Passive Interaction in PINTER. An inner narrative control loop incrementally selects actions for the narrative based on the current actions, goal and state of the narrative world. They are visualized using 3D graphics and the state of the world updated to reflect the result of visualization. The users' physiological responses to the visualized narrative, measured using EMG and GSR, are fed back to the module controlling narrative generation. Positive (negative) EMG response indicates user preference (dislike) for certain actions while GSR response triggers change to the current state of the tension level.

challenge one, exhibiting a different pace to that found in games, where progress is always user-driven.

A number of approaches have emerged to identify the characteristic properties of films that match the viewer's affective experience. ELVIS [24] used a range of physiological data to categorize the most effective film summaries (by a combined *entertainment value* measure), using moving average samples of physiological data, in a unified temporal model.

In IS, story dramatization can influence both affective response and story understanding via information presentation. As an example, Jhala and Young developed an approach for building camera directives for 3D animations of a sequence of story events [16]. Their motivation was to convey “dramatic effects”, which appear to be a type of affective conceptualization, through the use of a subset of Van Sijl's [36] cinematic conventions (concentrating on camera shot-based approaches). The results of their experiments showed a significant effect of the choice of shot on viewer comprehension.

Of those IS systems that feature some element of user interaction, they all adhere to the same interaction paradigm which requires the user to actively engage with the system, through the use of established multi-modal interactive techniques. For example, the anti-bullying application, FearNot! [4], and the interactive dramas Façade [23], Thespian [33], Gunslinger [15] and EmoEmma [7] all require *active* participation on the part of the user.

The starting point for our work was to challenge this dominant IS paradigm, according to which user interaction is active and the objective for our work was to explore the potential for passive interaction in interactive narratives. Our working hypothesis being that users' physiological responses to a narrative and its presentation can be used to drive narrative adaptation and evolution.

SYSTEM OVERVIEW

Figure 1 gives an overview of the passive interaction system implemented in PINTER. At the start of an interactive session the narrative control component selects initial narrative content and style parameters for presentation to the user, using real-time 3D visualization. This narrative control loop continues until a narrative endpoint is reached, incrementally adjusting the next part of the narrative at each step according to the user reaction. Internal variables are constantly updated to reflect changes to characters, plot and so on as the narrative unfolds. Users are equipped with EMG and GSR sensors as an interaction device while watching the interactive narrative, which feed information back to the narrative control component to determine the way in which the narrative evolves, in terms of both its content and presentation style, thus interaction is incorporated in an affective “loop”.

GSR has been shown to be a correlate of human arousal and has a quick response (onset of 1.0 to 3.0 seconds) with a

longer decay period (returning to a baseline over several seconds), thus events that induce arousal will have a reinforcing effect, and also events widely spaced will not impact on arousal affect. This property is used as a reflection of the pacing of the narrative, judging when the level of narrative tension has been successfully raised, when it should be released to avoid over-saturation, and when more dramatic/different events are needed for a better arousal response. The key passive response evaluated in PINTER was the attempt to match user arousal with a cyclic pattern of narrative tension, by increasing the narrative tension level in response to sustained low levels of physiological response from the user and to decrease the narrative tension level if sustained high levels of user response are detected.

EMG information is a more instantaneous response than GSR and it does not have a decay component. The EMG information is used to partition the action set as an indication of user preference: positive user responses to narrative actions leads the system to prefer those actions and negative responses leads the system to avoid them. PINTER uses a model of congruency; repeated similar responses to categories of actions increase or decrease the perceived preference for that type of action.

The success of the affective loop, shown in Figure 2, relies on the ability of the system to trigger appropriate physiological responses through the narrative that is presented, interpret the physiological responses of the user and use them to adapt the presented narrative in ways that appropriately reflect the user responses. These are discussed in the following sections.

INDUCING PHYSIOLOGICAL RESPONSES

As mentioned in the previous section, the affective loop in PINTER requires appropriate techniques for inducing user physiological responses in a passive filmic context. Our

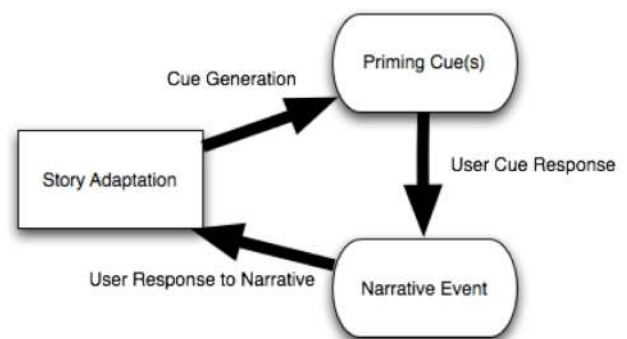


Figure 2. “Smithian” affective loop. Cues are generated to foster an emotional “priming” response using a variety of filmic techniques. This enhances the affective reaction to key narrative events. Cues and events can be combined narrative actions, or cues can stand alone.

Table 1. Example action and goal. The action preconditions specify the required tension level and relative status of the doctors in order for this action to be selected for a narrative. The dimensions intensity and dominance are set to high for use during visualization. One effect of the action is to show more of the doctors' relationship. The goal specifies the amount that must be shown about the doctors' relationship.

Action	Goal
<pre>(:action confrontation :parameters (?d1 ?d2 - doctor ?l location) :precondition (and ... (≥ (tension) high)) (boss-of ?d1 ?d2)) :effect (and ... (= (intensity) high) (= (dominance ?d1 ?d2) high) (increase (shown-relation ?d1 ?d2) 1))</pre>	<pre>(:goal (and ... (≥ (shown-relationship benton macnair) 5)) (...))</pre>

proposed method is inspired by Smith's emotional filmic theory [34]. Smith relates emotions to the narrative structure of films and placed emphasis on style of presentation in particular. He further argues that the emotional loop is made of multiple (and redundant) mood-inducing emotional cues, which make the viewer more likely to interpret further cues on the basis of their current mood. According to Smith it is the emotional associations provided by non-diegetic aspects, such as the mise-en-scène, cinematic idioms, color, sound, lighting and music that are crucial to triggering these filmic emotions.

Smith's theory was developed for film and not 3D graphics but our expectation was that such emotional cues would also be applicable in such a non-filmic context. Consequently we designed PINTER to utilize affective filmic principles, inspired from Smith's theory, in order to show that the theory did transfer to this context.

So that the narrative generation component of the PINTER system was able to reason about affective information, our earlier plan-based approach to narrative generation [28] was extended to include these aspects. The resulting modified narrative generator featured an implementation of the planner Metric-FF (chosen since it can plan with numeric values, a useful feature for handling physiological responses and tension levels). In adopting a plan-based approach we are following the dominant approach in IS, although taking a weaker representational stance on modeling narratives as plans. With this approach the planner generates narrative actions that can be staged in various ways according to the affective loop and affective categories. As an example, in PINTER's narrative setting of a medical drama, the actions include a confrontation between doctors resulting from pressures of work. The effects of this action include showing the viewer something of the doctors' personalities, their relationships, the working atmosphere, and so on.

The system maintains an explicit representation of the current narrative tension level and uses this to select actions

Table 2. Mapping of mood dimensions to cinematic techniques.

Dimension	Cinematographic Technique
Intensity	shot distance, lighting, point of view shots, pacing of cuts between shots environmental atmosphere, and music
Dominance	relative character position and camera angle
Menace	camera angle and lighting
Suspiciousness	camera angle, shot distance and character demeanor

and presentation style, which means that different categories of actions are brought in and out of consideration depending on the current tension level. For example, higher tension story line segments in PINTERs' medical narrative might include conflict between characters, deception, confrontation and tragic consequences for patients, whereas less dramatic story line segments could feature professional conduct, a calm hospital atmosphere and support amongst colleagues. Goals can relate to the resolution of patient cases or ensuring that aspects of a characters personality, relationships and medical style have been shown to the audience. This still ensures the existence of a baseline story, possibly made of emerging sub-plots to do with patient cases and relationships between staff, rather than a simple juxtaposition of events based on their dramatic properties. The action in Table 1 has the effect of showing something of the relationship between doctors and hence contributes to the goal shown in the table.

The actions for our medical interactive narrative were modeled to include non-diegetic aspects inspired by Smith's description of emotional cues [34] and idiomatic cinematographic techniques as discussed in [3, 37]. The action set included both *narrative* actions, which contribute to the goals of the narrative and *non-narrative* actions, which make no narrative contribution and act solely as emotional mood cues. In order to induce different levels of physiological response in the narrative the effects of actions included non-diegetic information that could be used during the visualization of the action. This information was expressed using a number of mood dimensions of relevance to the hospital domain. These dimensions were labeled Intensity, Dominance, Menace and Suspiciousness and the way they were mapped to cinematographic techniques during visualization is outlined in Table 2. Figure 3 illustrates the use of some of these cinematographic techniques in a number of narrative segments of different levels of narrative tension.



Figure 3. Example cinematic idioms applied to the hospital drama visualizations. PINTER reasons about the current level of narrative tension and selects the next segment of the narrative accordingly. It sends the narrative action along with relevant mood dimensions for visualization. For example, the high tension segment to do with deceptive behavior utilizes dramatic camera angles, lighting and character demeanor to heighten emotional cues.

PASSIVE INTERACTION AND PHYSIOLOGICAL RESPONSE

It is clear that *film* has the potential to engage viewers in a wide range of affective experiences, and creators use techniques such as those described earlier to foster empathy for characters and emotional investment in narrative development, without explicit intervention from the viewer.

We define Passive Interaction (PI) as all interaction with a system not performed intentionally through a dedicated input device. In traditional computing settings this might incorporate data from motion sensing, eye-tracking, facial recognition, and more. An application of this is as side-channel information to facilitate better experience in other interactions [14]. Certain ambient computing applications take up the idea of passive interaction in domestic and entertainment settings, where explicit sporadic inputs might break a viewer's engagement, suspension of disbelief, or concentration on other tasks, so they rely on extra “back channels” of information from the user. Of particular relevance to the system described in this paper are instances where responses are affective in nature as we are interested in utilizing mood- and emotion-based filmic characterizations. Game-like interactive scenarios come with a pre-conceived expectation of explicit interaction, to the point where periods of inactivity in games (i.e., cut scenes), are being subverted to restore an interactive experience [8]. It also has been suggested that these breaks in interactivity are still needed to invoke particular emotional engagement that involves low-arousal affective patterns [38], a feat ably performed in a filmic setting. In

Entertainment Computing installations, and more latterly, domestic applications such as television, multi-modal affective information not directly related to interaction can be used to characterize the empathic viewer response to the experience. This can range from detection of implicit, uncontrolled emotional expressions (especially of an Ekmanian type) to explicit expressions of affective response to an experience [21].

Physiological signals appear an ideal candidate for an input in passive interaction system, being always available, and obtainable without further direct action on the part of the user. “Lean back” experiences, as found in traditional filmic settings, avoid problems associated with physical activity, talking or other interactions interfering with the recording/collection of physiological data. With PINTER we aim to show that such signals can be used effectively to implement passive interaction for movie-like experiences in interactive media.

Temporal Properties of Physiological Data

Previous uses of physiological data as compound proxies of affective response have integrated a variety of signal types into a single processing model. This includes conceptualizing physiological input in terms of the temporal properties of emotions and mood [22, 24]. Generally, physiologically-based systems are aiming to reconstruct a discrete emotion within a window of consideration of input.

The two types of physiological input used in PINTER are the two most strongly associated with dimensions of

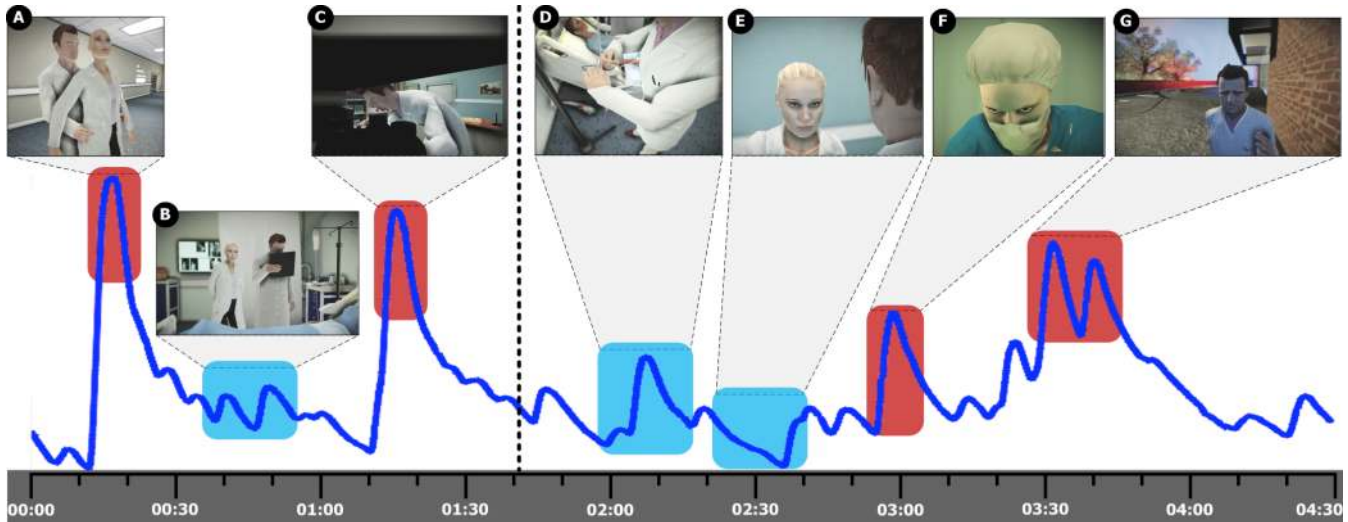


Figure 4. Physiological Response to the Interactive Narrative. A GSR signal trace is shown in blue representing the users' arousal response to narrative tension. The left-hand side of the trace shows an expected response following an alternation of higher and lower tension actions: harass-colleague (A), introduce-patient (B), get-treatment-deceptively (C). On the right-hand side is shown further adaptive behavior when GSR response does not increase following intended high tension events give-treatment-deceptively (D) and confront-behavior (E). PINTER adapts the story to include more high tension narrative actions, with more dramatic events, and GSR response picks up again, with emergency-treatment (F) and falling back into a cycle of tension with one last dramatic peak as the antagonist doctor attempts to run-away (G).

ffective response—GSR for arousal and EMG for positive and negative valence in the *zygomaticus major* and *corrugator supercilii* muscles in the face, respectively [5, 27]. GSR is associated with a changing overall level of arousal (skin conductance level or SCL), as well as short-term changes (skin conductance responses or SCRs). These changes are not instantaneous, but have an onset of 1.0-3.0 seconds (allowing for a SCR magnitude of sufficient level to establish) and return to baseline over a number of seconds. This can be contrasted to EMG, where there is always a low level of electrical activity in muscles, but activation events occur immediately, either as spikes or bursts of electrical activity much greater than the baseline level [2].

With PINTER, instead of a single compound temporal rationalization of physiological input, we are utilizing these divergent temporal properties of EMG and GSR input to distinguish between the medium- and long- term properties of dramatic tension and pacing and the more immediate positive and negative affective response to particular types of scene or shots. Adding more physiological signals to the interaction does not necessarily guarantee a better conceptualization of true affective response to the experience. For instance, Heart Rate correlation to arousal and valence aspects of emotion is both less clear-cut and less significant as well, and furthermore, its response time does not fit well in this division of labor between EMG and GSR, with too much lag and required processing for instantaneous responses, but too much inconsistency over longer periods of time for long-term tension assessment

(due to naturally varying non-affective physiological demands).

Galvanic Skin Response to the Interactive Narrative

PINTER uses SCR magnitude and frequency to calculate an arousal measure over a 30 second moving window. More SCRs of a higher magnitude indicate a higher level of arousal. We have developed an algorithm to detect SCRs in real-time without needing baseline measures. This avoids the need to handle drift in baseline SCL; these can otherwise be hard to detect when considering SCL only. The rationale behind this adaptation was the avoidance of long periods of sustained levels of arousal (either high or low)—in keeping with following a form of idealized Aristotelian curve over the course of a narrative.

Passive interaction should involve the continuous changing level of the arousal measure matching this idealized pacing curve of dramatic narrative events. If this arousal measure anticipates or lags behind the desired narrative pacing, the planner adapts the narrative actions as described in the next section. PINTER's response to a sustained period of low GSR response to low tension narrative segments is to select narrative content and presentation aimed at raising GSR response and similarly after sustained periods of high GSR response to high tension narrative segments to lower it.

If GSR still does not react in response to opposing tension narrative events, more extreme mood cues are employed and a variety of optional actions are introduced to attempt to find one that restores the natural pacing. Our experiments have shown that variation in narrative actions eventually

introduces actions that the user will respond to appropriately. Figure 4 shows an example GSR trace where PINTER is adjusting dramatic tension in narratives in order to manage user arousal levels to match a model of pacing. It can be seen that arousal in terms of GSR activity starts to lag behind the anticipated pacing of major narrative events. This leads to the planner preferring more intense and varied “high-tension” actions that foster a user response and a return to the normative cycle of pacing.

Facial Electromyography Response to the Interactive Narrative

The two EMG physiological signals used in PINTER stand for discrete positively or negatively valenced passive interaction events. We are not employing a unified emotional model, so these events are not associated with an emotion concept other than valence. Higher levels of EMG signal in an event indicate a stronger valenced reaction.

We would expect a strong EMG reaction to certain mood cues within actions, for which EMG signals are used to determine the success of such cues. For the majority of the rest of dramatic narrative we aim to avoid an abundance of negative reactions to narrative events, and EMG signals are interpreted in terms of the congruence of the expected affective responses given the current discourse, and the consistency of these reactions to action categories. In other genres, such as horror, more negative reactions may be a more desirable alignment with the narrative. Therefore, both positive and negative EMG events are taken into account when deciding what categories of narrative events to show (e.g., character development, conversations, patient scenes).

Consistency and strength of EMG reactions determine how strongly preferences to particular categories of action are perceived. For example, if conflicts between characters produced a negative EMG response then the system adaptation would be to exclude this type of action from the narrative unless needed for restoration of pacing. Alternatively if the EMG response was positive then the system would prioritize the selection of this “conflict” type action over others (e.g., flirting or joking around).

Figure 5 shows an example set of both EMG and GSR measurements against key scenes in an example narrative generated by PINTER. Two strong spikes of *zygomaticus* EMG activity can be seen which are aligned with mood cues within narrative actions and precede subsequent SCRs in the GSR activity, illustrating the difference in temporal nature of the signals.

EVALUATION

We designed a proof-of-concept study to investigate our proposed model of passive interaction. The study involved 23 subjects (7 male, 16 female), aged between 19 and 54, and the aim was to investigate the following properties of the PINTER system:

- Physiological signal levels were of a level and nature greater than random fluctuations watching neutral filmic material. (i.e., that passive interaction is not just an artifact of natural background activity).
- The narrative adaptations provided by PINTER, visualized by idiomatic cinematographic techniques provoked appropriate PI responses even though presented in non-filmic media.
- The narratives produced, whilst diverse, are also causally consistent and understandable to viewers.

Procedure

Subjects were shown 2 minutes of baseline video of a nature scene, of a kind rated as neutral in the IAPS picture set [18], while baseline physiological signal measurements were taken. They were then shown the visualization of narrative generated and adapted in real-time starting from a randomized narrative tension level. The interactive sessions lasted between 235 and 330 seconds, depending on the particular actions shown. Subjects then answered a questionnaire based on QUEST graphs [10, 13], generated from the particular narrative graph produced by PINTER during their interaction.

Physiological signals were normalized against the respective baseline per subject, and GSR measurements

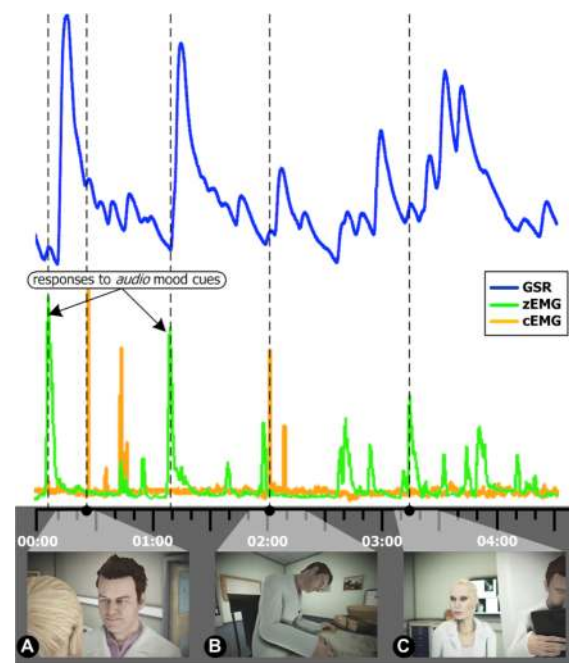


Figure 5. Combination of affective responses in story adaptations. Both zEMG and GSR are “primed” by affective audio cues, but the overriding response to actions flirt-with-colleague (A) and overwhelmed-with-work (B) are negative (cEMG), with little arousal response. When PINTER selects scenes involving the main patient story (e.g., reflect-on-successful-treatment-arrogantly (C)), more positive zEMG and responsive GSR results.

Table 3. Mean test(baseline) EMG activity values in peak/s. Starred values are significantly different from baseline.

Threshold	zEMG	cEMG
Low (< 20 μ S)	1.72 (0.51)*	0.79 (0.09)*
Med (< 30 μ S)	0.93 (0.28)*	0.21(0.05)*
High (< 40 μ S)	0.93 (0.21)*	0.10 (0.03)
Ex High (< 50 μ S)	0.48 (0.16)	0.05 (0.02)

were de-trended to remove effect of drifting baseline SCL over an interactive session as a whole.

Physiological Activity Compared to Baseline

GSR normalized means for baseline and test conditions show that test conditions invoke a greater average SCL than a neutral, non-dramatic baseline filmic video. A paired t-test of normalized mean GSR between baseline and test conditions across all subjects showed a significant difference ($p < 0.01$, $t = -8.8$, $df = 22$, $s.d. = 0.235$).

Both cEMG and zEMG showed difference in activity level in tests against baseline. In general there was more *zygomaticus* EMG activity than *corrugator* EMG. Mean values don't give much insight into EMG traces, due to the sporadic nature of EMG activity spikes. Instead, EMG was characterized as the number of detectable peaks in a trace, normalized for time, to give a measure of peak/s. This was evaluated at a range of threshold levels from close to the average noise floor to above what would normally be considered a significant detection level of a spike event. Lower thresholds showed better distinction between test and baseline, but obviously included more of the noise. Higher thresholds exclude noise, and still show significant difference (t-test $p < 0.05$) until you get to the highest threshold (50 μ S), at which point too many valid peaks are being rejected. This is illustrated in Table 3 (baseline means are in parentheses).

Whole Session GSR

Approximately 30% of whole session GSR across all subjects was matched to cue response, 50% to narrative action and 20% unattributable. This was calculated allowing a 5 second window for response to a cue or action start.

Three examples of tension adaptation to GSR response from the experiment are shown in Figure 6. Adaptation is shown as a red curve characterizing the internal tension representation for the actions that were selected. Normative attempted tension response is an alternative of high and low tension. Subject 18 (top) required a sustained high tension level at the mid-point of the session to re-invoke a GSR response. Subject 4 (middle) followed a fairly successful cyclic tension curve and response, although preferences and narrative causality necessitated a period of sustained higher tension between the 2:30 and 3:00 mark. Subject 13 (bottom) involved a slow response to start with, but once

the story was better adapted to pacing and preferred actions, user response followed the cycle of tension level, invoking a GSR arousal response during the dramatic finale.

These examples are representative of the kinds of experiences encountered across subjects. Generally response followed higher tension actions, and actions characterized as lower-tension engendered less response. While the characteristic cycle of tension is similar across subjects, the user response and hence experience and adapted narrative can be seen to be very individual.

Narrative Understanding

We also explored how well users could understand the interactive narratives that they viewed. To do this we used the QUEST model whereby narratives are represented as conceptual graphs that can be used to rate the relative quality of comprehension questions [10, 13, 16]. After viewing a presentation of a narrative, users are asked to assign goodness of answer (GOA) values to question-answer pairs and then the correlation with the predicted quality given by the QUEST graphs is assessed. For example, a sample pair from our experiments was:

Q: *What was the outcome for the patient?*

A: *She made a good recovery after initial treatment.*

The GOA of this pair depends on the content of the narrative that was shown to the user, i.e. whether the patient did or did not recover. For each user five questions and five answers were randomly selected from the QUEST model and subjects were asked to rate the GOA for each question

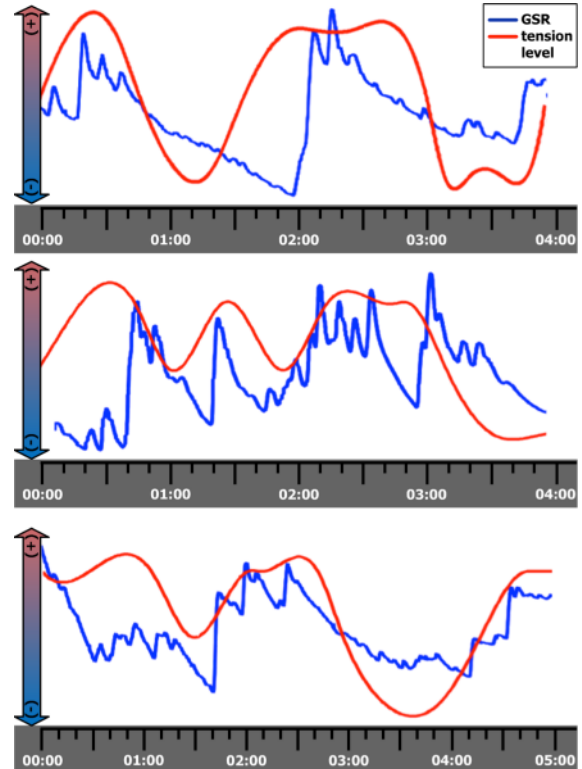


Figure 6. Example adaptations to subject GSR response in experimental subjects.

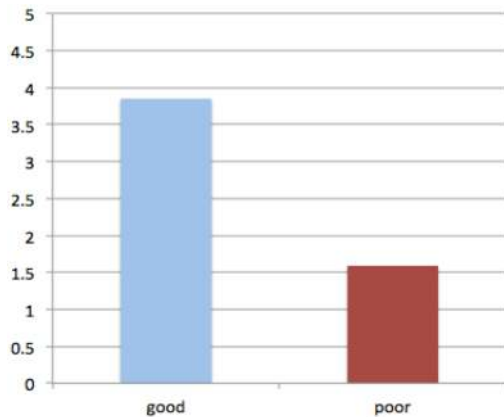


Figure 7. Bar chart showing mean Goodness of Answer for “good” and “poor” question/answer pairs. As expected, poor pairs had significantly lower GOA scores.

with a value from 1 (very bad) to 5 (very good). User responses were compared against measures of reachability and arc distance in the QUEST graph for their narrative. We set expected values for the GOA with 5 (very good) for those with arc distance 1, 4 for those with arc distance 2, and so on, with 1 (very bad) expected for question-answer pairs that were unreachable in the QUEST graph. Riedl and Young [29] partition their question space into “good” and “bad” question-answer pairs, where in a rating system of 1–5 “good” pairs have a system predicted GOA greater than the neutral, and “bad” pairs are those with a predicted GOA of less than neutral. Their general rule of thumb was understandability was indicated if the user scores for “bad” pairs were low and those for “good” pairs were high. For narratives generated by PINTER, it can be seen that this is indeed the case, as shown in Figure 7.

CONCLUSIONS

The motivation for the work presented in the paper was to demonstrate that viewers’ physiological responses to the narrative could be used to drive story adaptation in a passive feedback loop. Our expectation was that filmic theories of emotion provided a suitable mechanism to induce physiological responses in viewers, despite being developed for high quality films. To show this we designed an interactive narrative using 3D graphics that featured emotional cues characteristics of these filmic theories.

Our experiments confirmed that emotional cues were applicable in a non-filmic context, with physiological responses to emotional cues significantly over the baseline. In addition to inducing physiological responses we were also able to show that narratives could be adapted on the basis of these user responses whilst preserving viewer story comprehension. This research opens the way to new paradigms for IS that could enjoy better acceptability and avoid the bottleneck of active interaction modalities such as speech and natural language input.

ACKNOWLEDGEMENTS

This work has been funded (in part) by the European Commission under grant agreements IRIS (FP7-ICT-231824) and CEEDs (FP7-ICT-258749). Visual content developed by Matthew Laverick and Catherine Dixon, Teesside University. Story ideas provided by Terry Borst, Santa Fe University of Art and Design. Character models purchased from aXYZ Design (<http://www.axyz-design.com>).

REFERENCES

1. Ambinder, M., Biofeedback in Gameplay: How Valve Measures Physiology to Enhance Gaming Experience. In *Procs. Game Developers Conference (GDC)*, Citeseer (2011).
2. Andreassi, J. L. 2007. *Psychophysiology: Human Behavior and Physiological Response*. Psychology Press.
3. Arijon, D. 1991. *Grammar of the Film Language*. Silman-James Press.
4. Aylett, R., Dias, J. and Paiva, A., An Affectively-driven Planner for Synthetic Characters. In *Procs. 16th Int. Conf. Automated Planning and Scheduling (ICAPS)*, AAAI Press (2006), 2-10.
5. Bradley, M. M., Codispoti, M., Cuthbert, B. N. and Lang, P. J. 2001. Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing. *Emotion*, 1 (3), (2001), 276-298.
6. Bradley, S. D. 2007. Examining the Eyeblick Startle Reflex as a Measure of Emotion and Motivation to Television Programming. *Communication Methods and Measures*, 1 (2007), 7-30.
7. Cavazza, M., Pizzi, D., Charles, F., Vogt, T. and André, E., Emotional Input for Character-Based Interactive Storytelling. In *Procs. 8th Int. Joint Conf. Autonomous Agents and MultiAgent Systems, (AAMAS)*, IFAAMAS (2009), 313-320.
8. Cheng, P., Waiting for Something to Happen: Narratives, Interactivity and Agency and the Video Game Cut-scene. In *Procs. Digital Games Research Conference (DIGRA)*, Citeseer, (2007), 15-24.
9. Cheong, Y.-G. and Young, R. M., Narrative Generation for Suspense: Modelling and Evaluation. In *Procs. 1st Int. Conf. on Interactive Digital Storytelling (ICIDS 08)*, Springer (2008), 144-155.
10. Christian, D. and Young, M., Comparing Cognitive and Computational Models of Narrative Structure. In *Procs. Proc. of the 19th Nat. Conf. on Artificial Intelligence*, AAAI Press (2003), 385-390.
11. Fairclough, S. H., Psychophysiological inference and physiological computer games. In *Procs. Brainplay '07: Workshop on Brain-Computer Interfaces and Games. Advances in Computer Entertainment (ACE 2007)*, ACM (2007).

12. Gilleade, K., Dix, A. and Allanson, J., Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me. In *Procs. Digital Games Research Conference (DIGRA)*, Citeseer (2005), 16-20.
13. Graesser, A. C. and Hemphill, D. 1991. Question answering in the context of scientific mechanisms. *Journal of Memory and Language*, 30 (2), (1991), 186-209.
14. Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C. and Want, R., Squeeze hold tilt me! An exploration of manipulative user interfaces. In *Procs. SIGCHI Conf. on Human Factors in computing systems*, ACM Press. (1998), 17-24.
15. Hartholt, A., et al., At the Virtual Frontier: Introducing Gunslinger, a Multi-Character, Mixed-Reality, Story-Driven Experience. In *Procs. 10th Int. Conf. Intelligent Virtual Agents (IVA)*, Springer (2009), 500-501.
16. Jhala, A. and Young, R. M. 2010. Cinematic Visual Discourse: Representation, Generation, and Evaluation. *IEEE Trans. on Computational Intelligence and AI in Games*, 2 (2), (2010), 69-81.
17. Joly, J. Can a Polygon Make You Cry? <http://www.jonathanjoly.com/front.htm>
18. Lang, P., Bradley, M. and Cuthbert, B. 1997. International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention* (1997).
19. Lang, P. J. 1995. The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*, 50 (5), (1995), 372-385.
20. Laufer, L. and Németh, B., Predicting user action from skin conductance. In *Procs. 13th Int. Conf. on Intelligent User Interfaces (IUI)*, ACM (2008), 357-360.
21. Lee, C.-H. J., Chang, C., Chung, H., Dickie, C. and Selker, T., Emotionally reactive television. In *Procs. 12th Int. Conf. on Intelligent User Interfaces (IUI)*, ACM (2007), 329-332.
22. Mandryk, R. L., Atkins, M. S. and Inkpen, K. M., A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments. In *Procs. 24th Int. Conf. on Human Factors in Computing Systems (CHI)*, ACM Press (2006), 1027-1036.
23. Mateas, M. and Stern, A., Façade: An Experiment in Building a Fully-Realized Interactive Drama. In *Procs. Game Developers Conference*, Citeseer (2003), 82.
24. Money, A. and Agius, H. 2010. ELVIS: Entertainment-Led Video Summaries. *ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP)*, 6 (3), (2010).
25. Nacke, L., Kalyn, M., Lough, C. and Mandryk, R., Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction. In *Procs. Conf. on Human Factors in Computing Systems (CHI)*, ACM Press (2011), 103-112.
26. Nijholt, A., Bos, D. O. and Reuderink, B. 2009. Turning Shortcomings into Challenges: Brain-Computer Interfaces for Games. *Entertainment Computing*, 1 (2), (2009), 85-94.
27. Partala, T. 2005. *Affective Information in Human-Computer Interaction*. PhD Thesis. Department of Computer Sciences, University of Tampere, Finland.
28. Porteous, J., Cavazza, M. and Charles, F. 2010. Applying Planning to Interactive Storytelling: Narrative Control using State Constraints. *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 1 (2), (2010), 1-21.
29. Riedl, M. and Young, R. M., An Objective Character Believability Evaluation Procedure for Multi-Agent Story Generation Systems. In *Procs. 5th Int. Conf. on Intelligent Virtual Agents (IVA)*, Springer (2005), 278-291.
30. Riedl, M. O. and Young, R. M. 2010. Narrative Planning: Balancing Plot and Character. *Journal of Artificial Intelligence Research*, 39 (2010), 217-267.
31. Sakurazawa, S., Yoshida, N. and Munekata, N., Entertainment feature of a game using skin conductance response. In *Procs. Int. Conf. on Advances in Computer Entertainment Technology (ACE)*, ACM (2004), 181-186.
32. Sánchez-Navarro, J. P., Martínez-Selva, J. M., Torrente, G. and Román, F. 2008. Psychophysiological, Behavioral and Cognitive Indices of the Emotional Response: A Factor-Analytic Study. *The Spanish Journal of Psychology*, 11 (1), (2008), 16-25.
33. Si, M., Marsella, S. and Pynadath, D. V., Thespian: using multi-agent fitting to craft interactive drama. In *Procs. 4th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*, IFAAMAS (2005), 21-28.
34. Smith, G. M. 2003. *Film Structure and the Emotion System*. Cambridge University Press.
35. Tan, E. S. 1995. Film-induced affect as a witness emotion. *Poetics*, 23 (1-2), (January 1995), 7-32.
36. Van Sijll, J. 2005. *Cinematic Storytelling*. Studio City, CA: Michael Wiese Productions.
37. Vineyard, J. 1999. *Setting up your shots: great camera moves every filmmaker should know*. Studio City, CA: Michael Wiese Productions.
38. Zagalo, N., Torres, A. and Branco, V. Passive Interactivity, an Answer to Interactive Emotion. In *Entertainment Computing (ICEC)*, Springer, (2006) 43-52.