

2018-07-03

Unconsciously interactive Films in a cinema environmenta demonstrative case study

Kirke, Alexis

<http://hdl.handle.net/10026.1/11022>

10.1080/14626268.2017.1407344

Digital Creativity

Taylor & Francis (Routledge)

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Unconsciously Interactive films in a Cinema Environment - a demonstrative case-study

Anon, Anon, Anon, Anon, Anon, Anon, Anon

Anon

Anon

Corresponding author: Anon.

ABSTRACT

‘many worlds’ is a short narrative live-action film written and directed so as to provide multiple linear routes through the plot to one of four endings, and designed for showing in a cinema environment. At two points during the film, decisions are made based on audience biosignals as to which plot route to take. The use of biosignals is to allow the audience to remain immersed in the film, rather than explicitly selecting plot direction. Four audience members have a bio-signal measured, one sensor for each person: ECG (heart rate), EMG (muscle tension), EEG (‘brain waves’) and Galvanic Skin Response (perspiration). The four are interpreted into a single average of emotional arousal. ‘many worlds’ was the first live action linear plotted film to be screened in a cinema to the general public utilizing multiple bio-sensor types. The film has been shown publically a number of times, and lessons learned from the technical and cinematic production are detailed in this paper.

Keywords: bio-signal monitoring, immersive interaction, interactive cinema, movies, emotion

1. INTERACTIVE CINEMA

Most films are fixed – they are not changed by the audience during the film, and they in fact do not change once they are distributed. Interactive cinema allows the audience to influence or change elements of the film, so their experience can be different to others’. (Hales 2005) surveys cinematic interaction. Before the wide availability of digital technology, this was done manually. Nielson and other data tracking/rating companies have analyzed film and television reactions for decades (Kana-zawa et al. 2001) but they have not used it to interact in real-time with the media.

This paper documents the design and implementation of an engine for real-time detection of multi-modal biosignal responses from an audience in order to drive live editing of a film and its soundtrack. This generates streaming video for the purpose of audience affective manipulation whilst they watch the narrative of an algorithmic short film written and directed by the first author: ‘many worlds’. A key element of the film is that at fixed points in the plot the audience’s arousal level will be sampled and if it is below a pre-determined threshold, a more intense version of the next scene (as pre-defined by the director) will be selected.

The film ‘Kinoautomat’ (Činčera, 1967; Willoughby 2007) and the other films in this paragraph involve the audience consciously and explicitly selecting which narrative route to take, unlike ‘many worlds’. ‘Kinoautomat’ involved a moderator who would appear in front of the screen nine times during the film showing. The moderator asks the audience which of two choices they want to be followed for the next scene and there is a vote. Then the next scene is shown. The increasing availability of digital technology led to other solutions as found in the film ‘I’m Your Man’. Some mainstream cinemas were actually fitted with controllers for the film to allow people to vote on the main character’s decisions (Bejan 1992; Grimes 1993). However the approach did not take off. A more modern approach to interactive cinema can be found with those utilizing online video. For example ‘The Outbreak’ is an interactive film which is viewed online in a browser (Lund et al 2008;

Wong 2008). The user can click to select actions at certain plot points. Discussions of interactive cinema more generally can be found in (Beacham 1995), (Lunenfeld 2004), and (Vesterby et al. 2006). The area of interactive audiovisual entertainment and games also has obvious overlaps with video games and in particular with the history of laser disk games (Fahs 2008) and some innovative highlights like *Half Life* (Bates 1998) and *Dear Esther* (Shaw 2012).

The above examples, and most interactive cinema, has involved the audience consciously selecting film behavior. As observed in (Tikka 2006) and later (Gilroy et al. 2012), this can have the effect of reducing the immersion in the story. This led to Tikka et al. (2006) proposing the concept of enactive cinema where the audience do not consciously choose story directions. Tikka used heart-rate, breathing and movements of the audience to change the experience of an installation called 'Obsession' (Tikka 2005). This work also proposed a framework for basing various dimensions of the cinema experience on audience non-conscious inputs. Tikka et al. (2012) have also investigated the relationship between enactive cinema and neuroscience. Another pioneer in this non-consciously controlled form of story-telling has been Cavazza's group which has used eye gaze (Bee et al. 2010) and emotional neurofeedback (Cavazza et al. 2014) in narratives. Cavazza has also linked his interactive media research to non-entertainment applications. A key element in Cavazza's work is the use of "character" agents who inhabit the narrative space and have internal states and interactions. Vesterby et al. (2006) used eye gaze to control the direction of a two minute video clip. The downside with eye gaze is that the technology is still relatively costly and physically large, compared to bio-signals, a technology whose cost and size has been reducing more quickly, and which is now being incorporated into – for example – wearable items (Mosenia et al. 2017).

Castermans et al. (2012) measured audience bio-signals while they sat in a cinema to see if their emotional reactions could be detected. The results indicated that such detection was possible. Kierkels et al. (2008) attempted to detect audience interest during movie scenes using various bio-signals but could not quantify the precise nature of 'interest'. Gilroy et al. (2012) have also attempted to address this in a simple computer-generated graphical drama using single viewers at a workstation. A related study is found in (Giakoumis et al. 2011) which attempted to detect 'boredom' in people playing a video game.

The power of both eye-gaze and bio-sensor approaches is that not only can they maintain peoples' immersion, but they might potentially increase it by reactively manipulating plot, editing, or adjusting soundtrack elements in response to the audience dynamically. 'many worlds' is an attempt to extend previous work in this area. The work done in most interactive films differs to 'many worlds' by requiring conscious selection. The work done in non-conscious interactive films is extended by 'many worlds' into a public cinema environment with a larger number of sensors. 'many worlds' is also a new artistic experiment in the field of non-conscious interactive films that directly links the topic of the film story (how the observer can unconsciously affect the observed in quantum mechanics), with the mode of the film showing (in which the viewer can unconsciously affects the scenes viewed).

2. AFFECTIVE DETECTION

The 'many worlds' engine utilizes a system for detecting one dimension of emotion. The various models of emotion proposed by affective sciences offer complex, and still evolving, representations which can be used to map musical features to mood and vice versa. The dimensional approach to specifying emotion utilizes an n-dimensional space made up of emotion 'factors'. Any emotion can be plotted as some combination of these factors. The 2-Dimensional 'circumplex' model of affect (Russell 1980), with emotion comprised of valence and arousal, is often utilized in emotional evaluation for music (Schubert 1999; Mattek 2011; Doppler et al. 2011; Rubisch et al. 2011) and video (Xu et al. 2008; Sun et al. 2009; Hazer et al. 2015). In many emotional music creation systems (Kirke 2011) these dimensions are used. In the model used in 'many worlds', emotions are plotted on a

graph with the first dimension being how positive or negative the emotion is (valence), and the second dimension being how physically excited the emotion is (arousal). For example 'Happy' is high valence high arousal affective state, and 'Stressed' is low valence high arousal state.

Self-reporting arousal on such a model (Vuoskoski et al. 2011; Eerola et al. 2010) presents problems for the presentation and development of responsive, immersive media — and particularly as in this case, responsive immersive cinema — in that they force the interruption of any narrative. An alternative is to use a range of biosensors to estimate arousal responses (Tikka et al. 2006; Trost et al. 2011; Rossignac-Milon et al. 2010) from the cinema audience, unobtrusively. Then by making the film automatically responsive to these measurements, a real-time experience can be achieved without interrupting the audience's viewing. Such a real-time approach has much in common with previous work done in biofeedback in gaming (Dekker et al. 2007; Kuikkanaemi et al. 2010; Cederholm et al. 2011; Silva et al. 2014).

'many worlds' attempts this with a system that does not utilize valence, focusing on the measurement of arousal as a time-based vector. Important factors of the movie experience (beside emotions) fall outside of what such a system can take into account. Aspects of the viewers' cognitive processes, aesthetic dimensions, evaluative reactions, etc. — which are central parts of the movie experience — fly under the radar of the system. However arousal was chosen for this initial implementation because most biosensor research in the past has been more successful in detecting emotional arousal than emotional valence (Sammler et al. 2007).

As has been mentioned, in emotional measurement, arousal is what distinguishes Happy from Relaxed, and Angry from Depressed. It measures the physical activity of the emotion. So if a watcher is feeling positive about a film, an arousal-maximizing strategy will make them feel Happy rather than Relaxed, or Angry rather than Depressed. This is obviously a fairly blunt instrument but provides a first in-road into implementing emotion-control strategies.

The arousal vector is involved in a constant feedback loop, as ongoing arousal is continuously 'pinged' in real-time within the limits of a preset buffer. This vector is evaluated at various time values (at timestamp 00:03:28 and then either at 00:07:00 or 00:07:24), mapping the arousal and time value to a video selection, creating a range of possible narrative routes through the film for the audience. The entities involved in this process are time and a high-level arousal estimate (at a lower level, raw bio-signal data), with the relationship between these entities determined by the director in order to sustain or increase audience arousal whilst watching the film.

3. PRODUCTION

3.1 Story and Pre-production

Given the planned bio-signal reactivity of the film, the writing of the 'many worlds' script had a number of stages. A concept had already been developed concerning a person putting themselves in a box as a form of Schrodinger's Cat experiment (Schrodinger 1935). This was combined with the idea of a film that could be passively changed by some of the audience members affecting through their biosensor readings the film they observe : "the observer affects the observed". This was to implement the motivating creative concept that the key story element of the film (Schrodinger's cat observation experiment) was conceptually related to the way in which the film was designed to be watched — with some audience members who were observing the film, and were actually affecting the film in real time. In the Schrodinger's cat experiment, the observer affects the result purely by being an observer, without having to take any direct action on the cat or its surroundings. With the film, some of the film viewers affect the film without taking direct action on the film, but by passive readings taken from biosensors connected to them as they observe the film.

The first part of the process involved writing a normal short film story. Readers and viewers are used to experiencing stories written as an organic whole (Smiley et al. 2005). Thus it was consid-

ered that if a multipart multi-branching story was attempted from the outset, such a unity would be hard to achieve. So the process was to first write a single story, and then to develop it into a multi-path experience.

In summary, here is what the audience experience at the heart of the story: two students Charlie and Olivia arrive at the apartment of Connie on her birthday. They find Connie, a physics student, has sealed herself in a coffin-sized box with a cyanide gas-capsule connected to a Geiger counter. At any time a large enough burst of cosmic rays in the atmosphere could trigger the cyanide and kill Connie; in fact it could already have happened. Charlie – also a physics student – realizes Connie is performing a twisted version of a famous quantum physics experiment about the nature of reality, but one that was never meant to be performed in real-life. During the continuation of the film – through clips from their phones and a "mysterious" camera observing the room – the audience learn the true reason for the experiment.

After writing the core story various background elements crystalized. These elements had an impact on writing the alternative story routes. Charlie is a narcissist who is dating Connie and treating her very badly, driving a troubled person to the edge of her sanity. Olivia, Connie's best and most trusted friend, slept with Charlie as a one-off a few months before the events of the story. But – unbeknownst to them – Connie found out. Connie became obsessed and depressed by the idea that her best friend was now also her greatest enemy. Charlie's betrayal also made her believe that her beloved must despise her. In a bizarre attempt to rationalize these contradictions Connie creates a lethal scenario. Charlie and Olivia clearly don't realize the psychological damage they have done to Connie, hence their walking into the scenario.

Connie has set-up the situation so she is sealed in an airtight box, unable to see or hear what is going on in the room. The following information is only revealed during "twists" of certain endings of the film. The room itself is the "box", and the gas and detector – which Charlie and Olivia think is connected to the box, in fact feeds into the air of the room. Thus Charlie and Olivia are in the "box". In an analagous way to Schrodinger's cat being alive and dead at the same time in it's box, Connie wants to come to terms with her friends contradictory attitudes to her, to make them both friends and enemies at the same time. This, together with Connie's desire to punish them, makes her create a Schrodinger's Box for her closest companions.

Once the background became clearer, three more versions were written. This involved identifying possible split points in the script. The first split point came from the fact that when Charlie and Olivia arrive at Connie's house, Connie does not answer the door. Charlie has a key (because he is her boyfriend). Olivia says it is rude of them to walk in. So the first split is: whether Olivia allows Charlie to pressure her to come into the house straight away, or otherwise walks in a few minutes later. Another narrative split occurs just before Charlie tries to make Olivia drink vodka as part of an experiment he is using to demonstrate the Schrodinger's cat experiment (which he thinks Connie is trying to perform on herself). If Olivia allowed Charlie to pressure her to enter the house with him, she takes a drink, otherwise she refuses to take a drink. These two splits lead to four possible endings. In two of the four endings – the ones that follow on from her coming into the house with Charlie and taking a drink - Olivia dies and Charlie lives or dies. In the other two endings, Olivia survives and Charlie lives or dies. Figure 1 shows four screenshots, one from each of the four possible final endings in the film.

The story is thus driven by Olivia's strength of character in relation to the obnoxious Charlie. She has one main opportunity to determine to resist Charlie's behavior. The route the film takes depends on whether she takes this opportunity, and thus which of the endings happen. Olivia thereby controls the narrative direction. The results of the four endings emulate the result of an unentangled two Schrodinger box experiment, where Olivia and Charlie are the two cats: either neither, or one, or both die, as seen in Figure 3.

[Figure 1 suggested location]

3.2 Production

While attempting to transition from four detailed story descriptions to four formal scripts, it became clear how hard it would still be to write full scripts based on the multiple stories. Furthermore the writer had become aware of two feature film projects which were based on guided-improvisation by the actors involved (Goldstein 2013; Doto 2009). He investigated these projects and decided to utilize guided improvisation. The two main actors who were finally chosen both had experience of improvising together. Because of this approach, there was a longer process of rehearsal with the actors than would normally be found for a short film. When viewing cuts of the film it was clear that one of the endings was significantly weaker than the others.

During editing, it took a while to explain what was required of the professional editor. He had never worked in such a multi-path film before, but developed a format for working with standard editing tools to edit a multi-route film.

3.3 Soundtrack

A soundtrack was composed by the writer / director for each of the film clips. The electronic composition was generated with binaural beats. These involve two pure tones with frequencies that are slightly different. This creates the psychoacoustic effect of a beating slowly modulating their frequencies. The apparent frequency of modulation increases as the difference in pure tone frequencies increases. Although there has been work suggesting that binaural beats can affect mood (Owens et al. 1998), they are used here as an aesthetic choice by the composer, not with any scientific claims of mood manipulation. The use of such an abstract soundtrack is not so unusual. One key example of how audiences are becoming more used to such soundtracks is the sparse sub-bass soundtrack found in the mainstream feature film *Paranormal Activity 2* (Williams 2010).

The composition for ‘many worlds’ was done intuitively based on scene drama, not based on the arousal mark-ups. However an interesting structure emerged, as shown in Figure 2 (the branching composition of this figure will be explained in more detail later). Given that film Clip 1.2.1 was pre-defined by the director as having a higher arousal than Clip 1.2.2, it was found in post-analysis that the soundtrack had a maximum higher energy peak for the clip whose arousal was marked up as being higher (i.e. Peak 0.105 > Peak 0.052). Similarly with Clip 1.1.1 being marked up by the director as higher arousal than Clip 1.1.2, the peak energy of the soundtrack turned out to have a higher energy in the higher arousal-marked clip (i.e. Peak 0.0975 > Peak 0.0920). This was a working accident and not deliberately designed into the soundtrack.

[Figure 2 suggested location]

4. SYSTEM OVERVIEW

Four sensors are used to monitor participating four audience members physiological reactions in real-time. Each audience member wears one of the four sensors. These responses are combined in an affective estimation algorithm to give a moving average value for audience arousal. This moving average value is compared with an arousal threshold at various decision points in the narrative. The results of this comparison give the control data that maps the next part of the narrative the audience will watch, seamlessly creating an edit ‘on-the-fly’. Previous computer music research has made use of similarly collected bio-signal data as control inputs for music with emotional correlations. Such affective correlations to the selected bio-signals are well documented in literature (Tikka et al. 2006; Le Groux et al. 2009; Salimpoor et al. 2009). Our system broadly comprises three sections: Biosignal

metering, Arousal estimation, and Video editing (arousal synchronous narrative selection). These sections are explained in more detail below.

Four biosensors were utilized, all of which have implicated in detecting affective arousal:

1. Electrocardiograph (EKG), indicating mean heart rate from the participant above calibration threshold, averaged over 2-10 beats (Salimpoor et al. 2009)
2. Electromyograph (EMG), indicating muscle tension from the right forearm of the participant, as a mean within each buffer(n) (Hoehn-Saric et al. 1997)
3. Electroencephalograph (EEG), using three electrodes to indicate frontal brain activity, filtered to give only the alpha region using a band-pass 8-12kHz two-pole filter (Schmidt et al. 2001; Owens et al. 1998). As in (Owens et al. 1998) the natural logarithm of the alpha data was calculated and multiplied by -1
4. Galvanic skin response (GSR), giving a normalized value for perspiration on the left wrist and forefinger of the participant (Salimpoor et al. 2009)

Much research has indicated that GSR (Kim et al. 2008) is almost linearly related to emotional arousal. Similarly there have been multiple studies that relate EEG asymmetry to valence, and EEG frontal power to arousal (Schmidt et al. 2001; Reuderink et al., 2013). Heart rate has been linked to emotional positivity and negativity (Ekman et al. 1983). In particular Kim et al. (2008) supports that higher heart rate was linked to increased fear and anger, both high arousal emotions. Thus a detection of increased EEG frontal power, skin conductivity and heart rate, are indicative of increased emotional arousal. Muscle measurements (ECG) have also been used to detect tension (Kim et al. 2008).

4.1 Bio-signal metering

Sensor responses are digitized and passed to Max/MSP as raw data in real-time. Each data stream is calibrated to remove background noise using adjustable maximum and minimum input level outliers with EEG and GSR responses, and a simple noise-gating threshold for EKG and EMG responses. The responses from each sensor were then passed to an affective estimation routine to determine an instantaneous audience arousal value with which to carry out video selection.

4.2 Arousal estimation routine

Affective arousal is estimated from the four biosensors as a moving average. It was decided to combine the four inputs into a single measurement, and deal with the difference in their dynamics during a manual calibration stage. Given that different viewers would give different ranges of measurement during a viewing, a calibration stage was required, and the aim was to deal with the different dynamics of the inputs for each viewer. The output from each sensor is normalized before being summed across a nominal buffer, as shown below, where $A(n)$ = estimated arousal for buffer (n).

$$A(n) = \frac{\Sigma EEG(n)+EKG(n)+EMG(n)+GSR(n)}{n}$$

The values of the four inputs are normalized between 0 and 1, to aid in consistent manual calibration, however they are not transformed in any other way. This means that in essence four non-linear processes with different dynamics are being linearly combined. To mitigate against the effects of this, each process was adjusted during calibration with the four audience members. Thus the faster moving dynamics of the muscle tension monitor can be reduced in effect so it does not overwhelm the slower moving GSR. Similarly any single measure can be prevented from saturating the sum. Another issue is that there is a difference between muscle tension and the other three measures. Frontal EEG, GSR and heart rate have all been related to higher arousal. Whereas the muscle tension

measure is in this case based on a less standard concept, which has not been tested, which is the idea of arm tension caused by gripping - for example a seat arm. It is planned in future work to investigate the advantages of non-linearly transforming and combining the four inputs.

Results from the arousal estimation algorithm are compared with a pre-determined arousal threshold (AT), set separately during calibration for each viewing, in order to generate a control message for selection of video playback in the video mapping portion of the code. The arousal threshold was set for each viewing by observing the movement of combined arousal after placing the sensors on the viewer(s) while they were seated in the viewing area (the cinema or laboratory). Once sensors were attached, the combined arousal value was observed over a period of time to get a sense of its range and dynamics. Based on this, a threshold value was set.

4.3 Video editing: arousal-synchronous narrative selection

The first iteration of the Jitter-based video playback engine was designed in order to switch between three different narratives ‘on-the-fly’ by direct comparison of arousal values with the pre-determined arousal threshold (AT). In the finished system, video timecode is also used as a mapping entity such that time and arousal are mapped to video selection and playback, creating an arousal-synchronous method of video narrative selection. Seven clips in total are used in this system, as illustrated in Figure 3.

[Figure 3 suggested location]

Figure 3 shows that the 7 clips present four possible ‘routes’ for the audience, through two branches or ‘split’ points based on timecode values at the end of video clips, timestamp 00:03:28 (end of Clip 1) and, depending on which clip is selected, timestamp 00:07:00 (end of Clip 1.1) or 00:07:24 (end of Clip 1.2). The arousal buffer, (n), is reset after each of the split points as part of the affective estimation algorithm. This real-time detection of arousal allows the filmmaker to select narrative according, and in direct response to, the audience’s arousal. This allows the film to adapt to the audience, and the filmmaker to discretely target the induction of arousal in the audience, maintaining or increasing arousal through the narrative. The choice of trying to increase audience arousal was an artistic decision by the filmmaker (writer / director). The alternative would have been to either minimize the audience’s arousal, or create some alternating pattern of low then high arousal or vice versa. This in itself highlights the role of the director/author in this decision. Not only is the author deciding how the segments are written, but also how the storytelling technology reacts to the biosensors. A decision to minimize arousal, or to create some alternating arousal pattern, would be the authoring of a different film, despite the fact that the segments were the same segments. The labelling of segments as being higher or lower arousal was also based on the author’s judgement, as discussed below. A different decision on which segments were higher or lower arousal would have led to a different film. In fact it may be useful to talk of a “film-meta”. The film could be described as what the audience experiences in a viewing. The film-meta is the combination of the script, together with the author’s desired arousal pattern and the author’s perception of arousal levels of each segment of script.

This was a subjective process and part of the artistic decision making, as there are no agreed methodologies for measuring such ‘plot arousal’. The core strategy is as follows: if a sample of the audience are not aroused enough, they will be presented with a higher arousal clip at the next split point. If they show enough arousal, they get shown a more calming scene instead. Thus the arousal of story scenes was estimated by the director, as follows. The story involves three lead characters, and takes place in two locations: one outside and one inside in a single room. Clips 1.1 and 1.2 in Figure 3 are differentiated by action taking place with one or with two people respectively. The two-person clip was considered to have a higher arousal due to their interactions. Clip 1.2.1 was considered to be higher arousal than Clip 1.2.2. In Clip 1.2.1 Charlie kills Connie violently, whereas in Clip

1.2.2 Charlie and Olivia die in less violent circumstances. Similarly Clip 1.1.1 was judged higher arousal than Clip 1.1.2. In Clip 1.1.1 Charlie dies, whereas in Clip 1.1.2 Charlie and Olivia survive. During editing it became clear that the ending in which all characters survive (Clip 1.1.2) was narratively weaker than the others, but reshoots were not possible.

5. RESULTS

5.1 Laboratory Environment

Although 'many worlds' was shown in multiple cinema screenings, and at a screening room, data was not collected during these showings. Hence an experiment was done with 6 viewers: five males / one female, ages 20-25. Each wore all four sensors. The bio-sensors were calibrated for each participant by setting the threshold separately. This was done manually by observing their bio-signal activity for a few minutes after they first put on the sensors. By the time of these experiments a significant amount of informal experience had been gained by the experimenter in various applications with the final software. The viewers watched the film in a laboratory environment on a computer monitor, with the system adapting to their bio-signals. After completion of a full film viewing, the participants watched the other three endings which they had not yet seen. The experiment was not repeated multiple times with a single participant, as narrative impact would change over multiple viewings in unpredictable ways.

After watching the film the participants filled out a questionnaire anonymously. The key questions were:

- 1. Which ending (or type of ending) did you find had the highest physical intensity?*
- 2. Which ending (or type of ending) did you find had the lowest physical intensity?*
- 3. Did you find Olivia's agreement to drink the vodka a more intense narrative experience than Charlie's drinking of it? Or no difference?*
- 4. Do you have any memory / comments on the soundtrack or its impact?*
- 5. Do you have any general comments?*

The first thing ascertained was whether the writer / director's mark-ups of ending intensities had any relationship to those experienced by the viewers. Figure 3 shows the result. The Figure shows that 4 of the 5 who responded agreed with the writer/director's interpretation of the ending with least arousal (ending 1.1.2). 3 of the 5 agreed that 1.2.1 was the highest arousal ending. Another viewer said that 1.1.1 was the highest arousal, whereas the director said that 1.1.1 was the second highest arousal ending. So only 1 respondent out of 5 entirely disagreed with the writer/director's interpretation of arousal. In fact – respondent 3 takes the entirely opposite view to the writer / director, but for a specific reason. They said: "Strangely, I found ending 1.1.2 to be the highest. I believe this was a result of the acting in the other endings being weak, which made them feel a bit silly and took away from the intensity of the situation" and "Ending 1.2.1, just felt the acting and the filming shot made it feel silly."

The viewers were also asked if they perceived a difference in arousal between the decision point 1 selections (Clips 1.1 and 1.2). None of them did. This raises the question of the validity of the

whole selection at decision point 1 (at Clip 1 end). There were also no strong feelings about the impact of the soundtrack.

Figure 4 details the biosensor responses at the key decision point where one of the four endings is selected (decision point 2), and at the point just before the credits roll (story end). In Figure 4, two of the six participants (3, 5) were below their arousal threshold at decision point 1 (the end of Clip 1) and so were shown Clip 1.2 in an attempt to raise their arousal. At decision point 2 (after Clip 1.1) participant 2 was below their arousal threshold and so was shown Clip 1.1.1 to try to raise their arousal. At decision point 2 the rest of the participants were judged to have a sufficiently high arousal, hence the clips they were shown. Note, that as mentioned earlier, each participant had their threshold arousal set during calibration.

Figure 5 also shows which endings the participants considered to be the highest and lowest arousal (by verbal response). Only participant's 4 and 6 verbal responses concurred clearly with the ending selected by the system. They were both shown the director-judged low arousal ending, and both agreed it was the lowest arousal ending. In terms of the functionality of the bio-sensor readings results can be suggestive only due to the number of participants. However average arousal readings did increase across all viewers, between decision point 2 and the end credits, though only a small amount (13%).

[Figure 4 suggested location]

[Figure 5 suggested location]

5.2 Public Cinema Environment

Prior to the laboratory experiments, 'many worlds' was premiered to a live cinema audience at the Peninsula Arts Contemporary Music Festival, UK, on February 23rd 2013. Further showings were later given at BBC Research and Development at MediaCity UK, London Free Film Festival, the Walt Disney Company cinema in London, and at PrintScreen Festival in Tel Aviv, Israel. Footage from the premier is available (Al Jazeera 2013).

Although no data was collected during the cinema / screening room showings outside of the laboratory environment, it was noted that different endings were selected across the different showings, i.e. the same ending did not always occur. One key element to note is that after the weakest ending was chosen at the premiere (Clip 1.1.2, Charlie and Olivia live) it was decided to lower the probability of this ending occurring by changing threshold values at decision point 2. However, the differences in endings shown at the public and screening room presentations was not solely because of this fact, due to the variety of endings experienced.

In terms of sensor usage in the cinema environment, it became clear that the length of sensor leads needed to be extended. Some of the four audience volunteers were found to have to sit in a less comfortable position because of the lead lengths. In later showings this was corrected and the volunteers were much more comfortable. Some people also found the EEG headset uncomfortable, and one person found the muscle tension monitor uncomfortable. The sensors most amenable to calibration were heart rate, and also the muscle tension monitor, as the audience member could be asked to directly flex the area of muscle involved. EEG was the most difficult to calibrate because of the noisiness of the data and its artifacts, in fact a key addition to the system in future would be an artifact removal algorithm. GSR was also difficult to calibrate quickly because it was such a slow moving signal. The heart rate sensor, as well as being simple to calibrate, was the simplest to use. The downside was there was sometimes a false triggering of a heartbeat, so a suitably long averaging window needed to be used to filter these out.

6. CONCLUSIONS AND FUTURE WORK

This paper has described the motivation, pre-production, production, and technical development behind a multi-biosensor live action film 'many worlds', designed for and shown in a cinema environment. Issues of story, direction, and biosensor interpretation have been discussed. Additionally an artistic motivation concerning the analogy of story content and mode of viewing has been highlighted. Small scale laboratory tests support that directors may be able to predict which endings an audience member would find most "arousing", but were inconclusive concerning how appropriately the bio-sensor metering was selecting them. The system described is capable of playing back full HD video and synchronous audio whilst monitoring and calculating the arousal estimate in real-time.

In many ways the work so far could be viewed as a pilot project to examine how to run more extensive tests. These could include, given the difference in responses of the sensors, tracking each sensor separately with their own thresholds and create a multi-linear or multi-non-linear decision process at each decision point. Furthermore it would be helpful to analyze the biosensor data in terms of the seen plot for each participant, in order to understand how the viewers responded, in addition to the subjective questionnaires. It became clear while running tests how difficult it was to investigate such a system. How does one test four versions of a narrative on a single person? It is not possible because each test will affect the participant's future experience of the story. It is also hard to compare tests across subjects because of personal reactions to a narrative based on the viewer's character and life-experience. A possible future approach would be to create a set of clustering experiments. In these, a much larger set of viewers are shown a specific narrative unrelated to the primary test narrative. This clustering narrative is designed to test user responses to narrative more broadly, and allow a clustering of such users into multi-viewer clusters cluster1, cluster2, cluster3, etc. Once this cluster-

ing is completed, one can attempt to test a single film with multiple narrative paths, such as the film in this paper. Each path is tested on *different* users, but within *the same* cluster. The idea being, that people in the same cluster are roughly comparable in their responses. But even this approach moves much of the problem down to assumptions and designing of the clustering narrative. Another approach would be through a simple pre-questionnaire where one could ask for film preferences (e.g. “Do you like films where people cheat on each other? Do you like horror movies?”) then the same results could be assigned to different interest groups. This individual impact of narrative highlights that a larger sample is required to cover a range of ages and gender, which are known to react differently to films.

The issue of averaging different signal types was discussed earlier. However this highlights another issue that can be addressed in future work. Each participant in the public cinema showings did not have all four sensors. Given the experience of the work done so far, future work includes giving all four sensors to each participant, and then implementing a decision system: for example majority voting.

In conclusion, this paper has presented an approach to interactive filmmaking that involves the use of arousal detection through multiple bio-sensors, which leads to a dual-decision process for the director concerning: (a) which segments of the film are expected to induce what level of emotional arousal, and (b) what the desired emotional arousal arc is for the film. Additionally, in this case, the writer/director also created a story that had strong conceptual parallels with the story-telling technology. In the Schrodinger’s cat experiment (a core element of the story), the observer affects the result purely by being an observer, without having to take any direct action on the cat or its surroundings. With the film, some of the film viewers affect the film without taking direct action on the film, but by passive readings taken from biosensors connected to them as they observe the film.

Acknowledgments

‘many worlds’ was commissioned by Peninsula Arts Contemporary Music Festival 2013, Sensing Memory.

7. REFERENCES

1. Al Jazeera. 2013. UK movie audience gets wired into plot. <http://www.aljazeera.com/video/europe/2013/02/201322472548983827.html> [accessed 9 June 2016].
2. Bates, J. 1998. Half-Life. IGN. Available from <http://uk.ign.com/articles/1998/11/26/half-life-5> [accessed 7 October 2014].
3. Beacham, F. 1995. Movies of the Future: Storytelling With Computers. *American Cinematographer*, April:4-12.
4. Bee, N., Wagner, J., André, E., Vogt, T., Charles, F., Pizzi, D. and Cavazza, M. 2010. November. Discovering eye gaze behavior during human-agent conversation in an interactive storytelling application. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, 9. ACM.
5. Bejan, R. 1992. *I’m Your Man*. New York, USA.
6. Bubic, A., von Cramon, D. and Schubotz, R. 2010. Prediction, cognition and the brain. *Frontiers in Human Neuroscience*. 4:25.

7. Castermans, T., Duvinage, M. and Riche, N. 2012. Emotive Cinema, *QPSR of the numediart research program*, 5(1):7-14.
8. Cavazza, M., Charles, F., Aranyi, G., Porteous, J., Gilroy, S.W., Raz, G., Keynan, N.J., Cohen, A., Jackont, G., Jacob, Y. and Soreq, E., 2014, March. Towards emotional regulation through neurofeedback. In *Proceedings of the 5th Augmented Human International Conference*, 42, ACM.
9. Cederholm, H., Hilborn, O., Lindley, C., Sennersten, C. and Eriksson, J. 2011. The aiming game: using a game with biofeedback for training in emotion regulation. In *Proceedings of DiGRA 2011 Conference*, Utrecht, Netherlands.
10. Cincera, R. 1967. Kinoaautomat. Czech Republic.
11. Cohen, A. 2010. Music as a source of emotion in film. In *Handbook of Music and Emotion*, eds. P. Juslin and J. Sloboda, 879-908. Oxford: Oxford University Press
12. Dekker, A. and Champion, E. 2007. Please biofeed the zombies: enhancing the gameplay and display of a horror game using biofeedback. In *Proceedings of 2007 DiGRA*, 550-558, Tokyo, Japan.
13. Doto, B. 2009. Quiet Earth Interviews SUBHYSTERIA'S creative team. *Quiet Earth*. Available from <http://www.quieterth.us/articles/2009/01/20/Quiet-Earth-Interviews-SUBHYSTERIAS-creative-team> [accessed 11 October 2014]
14. Doppler, J. Rubisch, J., Jaksche, M. and Raffaseder, H. 2011. "RaPScoM: towards composition strategies in a rapid score music prototyping framework" In *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, 8-14, ACM.
15. Eerola, T. and Vuoskoski, J. 2010. A comparison of the discrete and dimensional models of emotion in music, *Psychology of Music*. 39(1):18-49.
16. Ekman, P., Levenson, R.W. and Friesen, W.V. 1983. Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616), 1208-1210.
17. Fahs, T. 2008. The Lives and Deaths of the Interactive Movie. IGN. Available from <http://uk.ign.com/articles/2008/03/04/the-lives-and-deaths-of-the-interactive-movie> [accessed 7 October 2014].
18. Giakoumis, D., Tzovaras, D., Moustakas, K. and Hassapis, G. 2011. Automatic Recognition of Boredom in Video Games Using Novel Biosignal Moment-Based Features. *IEEE Transactions On Affective Computing*. 2(3):119-133.
19. Gilroy, S., Porteous, J., Charles, F. and Cavazza, M. 2012. "PINTER: Interactive Storytelling with Physiological Input". In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, 333-334, Lisbon, Portugal, ACM.
20. Goldstein, G. 2012. Review: 'Goodbye Promise' captures Hollywood heartbreak. *Los Angeles Times*. Available from <http://articles.latimes.com/2012/jun/01/entertainment/la-et-goodbye-promise-capsule-20120601> [accessed 11 October 2014].
21. Grimes, W. 1993. When the film audience controls the plot. *New York Times*. Available from <http://www.nytimes.com/1993/01/13/movies/when-the-film-audience-controls-the-plot.html> [accessed 7 October 2014]
22. Hazer, D., Ma, X., Rukavina, S., Gruss, S., Walter, S. and Traue, H. C. 2015. Emotion Elicitation Using Film Clips: Effect of Age Groups on Movie Choice and Emotion Rating. In *Pro-*

- ceedings of 2015 International Conference on Human-Computer Interaction*, 110-116, Springer International Publishing.
23. Hoehn-Saric, R., Hazlett, R., Pourmotabbed, T. and McLeod, D. 1997. Does muscle tension reflect arousal? Relationship between electromyographic and electroencephalographic recordings. *Psychiatry Research*. 71(1):49–55.
 24. Le Groux, S. and Verschure, P. 2009. “Neuromuse: Training your brain through musical interaction” in *Proceedings of the International Conference on Auditory Display*, Copenhagen, Denmark.
 25. Hales, C. 2005. Cinematic interaction: From kinoautomat to cause and effect. *Digital Creativity*, 16(1): 54-64.
 26. Kanazawa, M. T., & Funk, J. P. 2001. Racial discrimination in professional basketball: Evidence from Nielsen ratings. *Economic Inquiry*, 39(4): 599-608.
 27. Kierkels, J. and Pun, T. 2008. “Towards detection of interest during movie scenes”, In *Proceedings of the 2008 PetaMedia Workshop on Implicit, Human-Centered Tagging*. London, UK.
 28. Kim, J. and André, E. 2008. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 30(12), 2067-2083.
 29. Kirke, A. 2011. Application of Intermediate Multi-agent Systems to Integrated Algorithmic Composition and Expressive Performance of Music, PhD Thesis, Plymouth University.
 30. Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I. and Ravaja, N. 2010. The influence of implicit and explicit biofeedback in first-person shooter games. In *Proceedings of the 2010 SIGCHI Conference on Human Factors in Computing Systems*, 859-868, ACM.
 31. Lund, L., Lund C. 2008. *The Outbreak*. Los Angeles, USA.
 32. Lunenfeld, P. 2004. *The Myths of Interactive Cinema*’ In *Narrative Across Media: The Languages of Storytelling*, Edited by: Ryan, M. Nebraska: University of Nebraska Press.
 33. Marshall, S. and Cohen, A. 1988. Effects of musical soundtracks on attitudes toward animated geometric figures Music Perception. 6(1):95–112.
 34. Mattek, A. 2011. “Emotional Communication in Computer Generated Music: Experimenting with Affective Algorithms” In *Proceedings of the 26th Annual Conference of the Society for Electro-Acoustic Music in the United States*. Miami, Florida.
 35. Mitry, J. 1997. *The aesthetics and psychology of the cinema*. Indiana University Press.
 36. Mosenia, A. Sur-Kolaty, S.U.S.M.I.T.A., Raghunathan, A. and Jha, N. 2017. Wearable Medical Sensor-based System Design: A Survey. *IEEE Transactions on Multi-Scale Computing Systems*, PP(99): 1
 37. Owens, J., Justine, E. and Marsh, G. 1998. Binaural auditory beats affect vigilance performance and mood. *Physiology and Behavior*. 63(2):249-52.
 38. Pantelopoulos, A. and Bourbakis, N. 2008. “A survey on wearable biosensor systems for health monitoring”. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4887-4890, Vancouver, BC, IEEE.

39. Reuderink, B., Mühl, C. and Poel, M. 2013. Valence, arousal and dominance in the EEG during game play. *International journal of autonomous and adaptive communications systems*, 6(1), 45-62.
40. Rossignac-Milon, M. 2010. Affective Computing: A Survey Internship Report. MIRALab, University of Geneva, Switzerland.
41. Rubisch, J., Doppler, J. and Raffaseder, H., 2011. "RAPSCOM - A Framework For Rapid Prototyping Of Semantically Enhanced Score Music" In *Proceedings of Sound and Music Conference*, Padova, Italy.
42. Russell, J. 1980. A circumplex model of affect. *Journal of personality and social psychology*. 39(6):1161.
43. Salimpoor, V., Benovoy, M., Longo, G., Cooperstock, J. and Zatorre, R. 2009. The Rewarding Aspects of Music Listening Are Related to Degree of Emotional Arousal. *PLoS ONE*. 4(10):e7487. doi:10.1371/journal.pone.0007487.
44. Sammler, D., Grigutsch, M., Fritz, T. and Koelsch, S. 2007. Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*. 44(2):293–304.
45. Schmidt, L. and Trainor, L. 2001. Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition and Emotion*. 15(4):487–500.
46. Schrödinger, E. 1935. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* 23:807-812; 823-828; 844-849.
47. Schubert, E. 1999. Measuring Emotion Continuously: Validity and Reliability of the Two-Dimensional Emotion-Space. *Australian Journal of Psychology*. 51(3):154–165.
48. Shaw, P. 2012. GDC: Creating Dear Esther's Eerily Desolate World, Wired, March. Available from <http://www.wired.com/2012/03/dear-esther-gdc/> [accessed 7 October 2014].
49. Smiley, S., Bert, N.A. 2005. Playwriting: The structure of action. Yale University Press.
50. Silva, G. A. D., Nogueira, P. A. and Rodrigues, R. 2014. Generic Shooter 3000: a realistic first person shooter powered by biofeedback. In *Proceedings of the 1st ACM SIGCHI annual symposium on Computer-human interaction in play*, 371-374, ACM.
51. Sun, K., Yu, J., Huang, Y. and Hu, X. 2009. An improved valence-arousal emotion space for video affective content representation and recognition. In *Proceedings of 2009 IEEE International Conference on Multimedia and Expo*, 566-569, IEEE.
52. Tanaka, S., Samejima, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., and Doya, K. 2006. Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Networks*. 19:1233-1241.
53. Tikka, P. 2005. Enactive Cinema Installation 'Obsession'. Museum of Contemporary Art Kiasma, Helsinki.
54. Tikka P., Vuori, R., & Kaipainen, M. 2006. Narrative logic of Enactive Cinema: Obsession. *Digital Creativity Journal*, 17 (4): 234 – 238.
55. Tikka, P. Väljamäe, A., de Borst, A., Pugliese, R., Ravaja, N., Kaipainen, M. and Takala, T. 2012. Enactive cinema paves way for understanding complex real-time social interaction in neuroimaging experiments. *Frontiers of Human Neuroscience*. 6:298.
56. Trost, W., Ethofer, T., Zentner, M. and Vuilleumier, P. 2011. Mapping Aesthetic Musical Emotions in the Brain. *Cerebral Cortex*, 22(12):2769-83.

57. Willoughby, I., 2007. Groundbreaking Czechoslovak interactive film system revived 40 years later. Radio Prague, Available from <http://www.radio.cz/en/section/panorama/groundbreaking-czechoslovak-interactive-film-system-revived-40-years-later> [accessed 7 October 2014].
58. Vesterby, T., Voss, J., Hansen, J., Glenstrup, A., Hansen, D., and Rudolph, M. 2005. Gaze-guided viewing of interactive movies. 16(4):193-204.
59. Vuoskoski, J. and Eerola, T. 2011. Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*. 15(2): 159–173.
60. Williams, T. 2010. Paranormal Activity 2, Los Angeles, USA.
61. Wong, C. 2008. Available from <http://flipthemedia.com/2008/10/interactive-films-back-from-the-dead/> [accessed 7 October 2014]
62. Xu, M., Jin, J. S., Luo, S. and Duan, L. 2008. Hierarchical movie affective content analysis based on arousal and valence features. In *Proceedings of the 16th ACM international conference on Multimedia*, 677-680, ACM.

Figure 1. Four screenshots, one from each of the four possible final endings in the film. The top left shot is of Connie and Charlie in Clip 1.1.1, and the other shots are of Olivia and Charlie in Clips 1.1.2, 1.2.1 and 1.2.2 moving clockwise around the figure.

Figure 2. Parallel soundtrack structure using binaural beats sounds. Peaks indicate the maximum sample peak in the soundtrack in a clip.

Figure 3. Illustrating arousal threshold and ‘split’ points – editing decisions are made at predetermined timecodes (00:03:28 and then either at 00:07:00 or 00:07:24) by comparing the estimated arousal from the four biosensors to an arousal threshold and selecting a bipolar route through to four separate narratives.

Figure 4. Routes taken by different participants through the film 'many worlds' depending on the estimate of their arousal at each of the two split points (after Clip 1 and then Clip 1.1 or 1.2).

Figure 5. How the Director and Participants verbally judged clip arousal, vs what clip participants were shown by the system.