An Implementation of Affective Adaptation in Survival Horror Games

Vanus Vachiratamporn, Koichi Moriyama, Ken-ichi Fukui, and Masayuki Numao
The Institute of Scientific and Industrial Research
Osaka University, Osaka, Japan
Tel.: +81-6-6879-8426
Fax: +81-6-6879-8428
Email: {vanus, koichi, fukui, numao}@ai.sanken.osaka-u.ac.jp

Abstract—In this research, we investigated two important aspects of affective gaming, which are the recognition of the player’s affective states and the adaptation of the game based on the player’s current affective state, by using survival horror games as an experimental environment. In the previous work, we analyzed the affective states of players collected through our own affect annotation tool during their gameplay and constructed affective-state prediction models that predicted their affective states from their brainwave and heart rate signals collected concurrently. In this work, we developed an affective survival horror game based on the previous findings and conducted another experiment to compare between affective and non-affective versions of the game. In the affective version, the timing of horror events were implicitly changed based on the player’s affective state. The result showed that the implicit change of timing made no significant difference in the evaluation of the players, but the game itself shows good potential for future research in the affective gaming field.

I. INTRODUCTION

The concept of affective gaming refers to the form of gameplay where the player’s current affective state is used to manipulate gameplay [1]. Two main components are necessary for developing an affective game, i.e., emotion recognition and affective adaptation [2]. Most researchers in game research study, however, only work on one component at a time. For example, some researchers worked on recognizing emotion with no implementation on the game [3]–[5], whereas other researchers made use of physiological signals to manipulate the game parameters directly without an awareness of emotion of the player [6], [7]. In order to bridge the gap between the two components and develop the complete affective gaming system, the emotion model related to the game context has to be identified and the game has to be able to recognize emotion of the player in real-time.

We choose the survival horror game genre as the game environment because it has an interesting and unique emotion set. Survival horror is loosely defined as an action-adventure game with horror-themed atmosphere and story. Most research on survival horror games tried to find a way to elicit more fear intensity from players [8]–[10]. However, as far as our knowledge is concerned, there is no research that tries to work on recognizing horror-related emotion before, even though knowing emotion of the player in real-time could benefit the game on how it can twist the gameplay to scare the player much more.

In the previous work [11], [12], we collected the affective states of players playing a survival horror game through our own affect annotation tool and constructed affective-state prediction models that predicted their affective states from their brainwave and heart rate signals collected concurrently.

Based on the findings of the previous work, this work developed an affective survival horror game using the affective-state prediction model for real-time affective state recognition. This game adjusts itself in the way that it will scare players the most. We conducted an experiment to compare between affective and non-affective versions of the game.

This paper consists of five sections. Section II reviews several existing relevant works. Section III shows our previous work that analyzed the affective states of players playing a survival horror game. Section IV is the main contribution of this paper that introduces the affective survival horror game we developed and shows the result of the experiment we conducted to compare between affective and non-affective versions. Finally, Section V concludes this paper.

II. RELATED WORK

In this section, we review several existing works related to the measurement of gameplay experience and the adaptation of games using physiological sensors.

A. Measuring Gameplay Experience

Several researchers have been trying to measure the gameplay experience of players, which is normally derived from self-report questionnaires using various tools and methods. Nacke and Lindley [3] created three separate levels of a first-person shooter (FPS) game designed to assess boredom, immersion, and flow experiences. They used game experience questionnaires (GEQ) for subjective measurement of various gaming experiences (e.g., immersion, flow, challenge, and tension) and used facial electromyography (EMG) and galvanic skin responses (GSR, also known as skin conductance) for objective valence and arousal measures, respectively. They showed that the responses from EMG and GSR were significantly different over the three levels, whereas GEQ had significant differences only on challenge and tension. Drachen et al. [13] collected heart rate (HR), GSR, and GEQ of players across three different FPS games. They showed that (i) a high HR was indicative of tensed and frustrated emotions; (ii) a low HR was indicative of positive affect, flow, feeling of competence, immersion and low levels of challenge; and (iii) GSR was...
correlated to negative affect. Similarly, Martinez et al. [4] also made use of HR and GSR features to predict reported affective states across two dissimilar games (predator/prey and racing games) with artificial neural network models. They showed that the affective models trained on one game could predict the reported affective states of the other game and suggested using the average HR and the two-step GSR variation for affect prediction. Jennett et al. [14] investigated immersion of players in games and whether immersion could be measured quantitatively through a series of experiments. Their overall findings showed that it could be measured subjectively by questionnaires and objectively by completion time and eye movements, and also that it could happen even when the players were in negative emotions.

Self-report questionnaires at the end of a session can only assess the affect of players over a whole game experience. For the questionnaires, the game has to be limited to a short duration because the player has to remember his/her affective experience. Furthermore, only one specific emotion can be effectively assessed at a time. For example, as mentioned above, Nacke and Lindley [3] needed three game levels for assessing boredom, immersion, and flow experience. These constraints do not sit well with the dynamic nature of interactive gaming experience.

There are only a few researchers who are interested in modeling emotions of players continuously throughout the game. Mandryk and Atkins [5] used a fuzzy logic model to transform physiological data (HR, GSR, and EMG) to the arousal-valence space and then to the game-related emotions while the player was playing the game. The fuzzy rules were created based on the psychophysiological findings in which arousal is correlated to GSR and HR and valence is correlated to EMG and HR, and the model was then compared with the subjective reports. With the continuous measurement of emotions in real-time, the transition of affect caused by game elements (e.g., sound and images) can be analyzed individually. More importantly, it opens up the possibility of adapting those elements in real-time based on the player’s current affective state.

B. Real-time Game Adaptation Using Physiological Sensors

As well as using physiological sensors to assess the gameplay experience of players, some researchers are interested in using the sensors to adapt the game. Kuikkaniemi et al. [6] investigated the influence of explicit and implicit biofeedback mechanisms on the experience of playing an FPS game. In the explicit biofeedback case, the player was informed that his/her current physiological state directly influenced the game. In the implicit biofeedback case, on the other hand, the player did not know that the game was influenced by his/her physiological states. They showed that, by explicitly informing the players how their GSR and respiration changed parameters of the game, they were more immersed and positively affected. Similar study was done by Nacke et al. [7] where physiological sensor inputs were used to augment traditional game control directly (EMG, respiration, temperature, and gaze) and indirectly (electrocardiography (EKG) and GSR). They showed that the participants preferred the direct physiological control over the indirect control, and suggested that the direct physiological control should be mapped to intuitive actions in the game while the indirect physiological control should be used as a “dramatic” device that alters the game world.

III. INVESTIGATION ON PLAYER AFFECT

We need to first identify what kind of emotions players feel while they are playing survival horror games. Different games contain different sets of emotions and the target emotion to which the games want to lead the players might be different as well. For example, eliciting intense fear from players is the main entertainment factor of survival horror games. In this section, we review our previous work [11], [12] in which we analyzed the affective states of players playing a survival horror game and constructed affective-state prediction models that predict their affective states from electroencephalography (EEG) and EKG collected concurrently.

A. Survival Horror Affect Model

Our survival horror affect model is based on the fear framework of Garner and Grimsay [15] and the definition of anxiety, suspense, and fear by Toprac and Meguid [10]. In a game context, these three emotions can be distinguished by the interaction and concreteness of how a threat is perceived. Prior to the confrontation with a threat, the player feels “anxiety” if he/she thinks that the threat is near or is going to appear soon but does not know or cannot imagine how it is going to come out. The player is in “suspense” if he has a strong feeling that the threat is going to appear soon. On the other hand, “fear” is an emotional response to the threat or an attempt to cope with the situation after the confrontation. Figure 1 shows the affect model that we used for defining the affective state of players while they are playing a survival horror game. The affects are separated into two groups, pre-fear affects and post-fear affects. Anxiety and suspense belong to the pre-fear affects as they are the affective states before the confrontation. We consider one more state called “neutral” in the pre-fear affects, which covers the case where the player has no feeling of uncertainty towards the threat at all. After the confrontation, the player transitions to the post-fear affects which are categorized as three levels of fear experienced, i.e., low, medium, and high. After the player is able to escape from the threat, he/she will transition back to the pre-fear affects and anticipate the new confrontation. Note that the player may fluctuate among the pre-fear affects before the confrontation and among the post-fear affects after the confrontation.

B. Constructing Affective-State Prediction Model

We developed an affect annotation tool (AAT) where the participant can annotate his/her affect continuously by
watching his/her recorded gameplay and facial expression simultaneously (see Fig. 2). EEG, EKG, and keyboard-mouse activities (KMA) were collected while the participant was playing a survival horror game. EEG is the recording of electrical activities of the brain and we used the Emotiv EPOC\(^1\), which is a high resolution and wireless portable EEG system, for recording EEG data. EKG measures the electrical activities generated by the heart. We recorded the EKG signals with the Procomp5 Infiniti encoder and EKG-Flex/Pro sensor\(^2\) by attaching three electrodes on the participant’s chest and abdomen.

Fig. 2. Screenshot of our affect annotation tool (AAT) [12]. The left screen shows the recorded gameplay video and the right screen shows the video of facial expression of the participant at that moment. In this screenshot, unfortunately, the gameplay video is not shown well because it is dark. After the gameplay, the participant watches them simultaneously and indicates his/her affective states at that time by using the bars below the screens.

The game we used is “Slender: The Eight Pages (STEP)\(^3\)”. In STEP, the player is situated in the woods in the night and has to collect eight pages of paper scattered around the woods, while being pursued by a threat called “Slender Man”. As each page is collected, the Slender Man will pursue faster and teleport more frequently. The game ends when the Slender Man comes in contact with the player or when the player’s sanity drops to zero by looking at the Slender Man for too long.

Eleven participants took part in this experiment. Since STEP has a very short play time and unpredictable scary events, the participants can play the game multiple times while still getting scared. Each participant repeated a gameplay and annotation for three times, thereby providing us with three sets of data per participant.

Using the collected affective states with AAT, we constructed prediction models that predict the pre-fear and post-fear affects from EEG, EKG, and KMA, respectively. Ten-fold cross validation, which is a usual method to evaluate classifiers in machine learning, showed that the prediction model trained from EKG features with C4.5 decision tree [16] obtained the best result; the f-measure values were 0.92 for pre-fear affects and 0.85 for post-fear ones. The used features for this model were HR, inter-beat interval (IBI), the mean of HR in 20 seconds with 1-second sliding windows ($\mu_{\text{IBI}}$), the standard deviation of IBI in 20 seconds with 1-second sliding windows ($\sigma_{\text{IBI}}$), and the total number of interval differences of successive heart beats greater than 50 milliseconds (NN50), which are all calculated from the EKG signal.

![http://www.emotiv.com/](http://www.emotiv.com/)


![http://parseproductions.com/slender/](http://parseproductions.com/slender/)

C. Transition Likelihood

With the continuous self-reported affective states collected with AAT, we can investigate how the affective states of the participants transitioned from a pre-fear affect to a post-fear one. As the objective of survival horror games is to elicit the highest fear from players, knowing when the players are likely to transition into high-fear state can be good information for designing the affective adaptation.

To measure the likelihood of transition from a pre-fear affect to a post-fear one, we used the transition likelihood function introduced by D’Mello et al. [17]. The likelihood function, presented in Eq. 1, serves as a better measurement than probability because it takes the base rate of transition into account. All pairs of affect transitions from a pre-fear affect $F_{\text{pre}}$ to a post-fear affect $F_{\text{post}}$ are counted and used to compute the transition likelihood,

$$L(F_{\text{pre}} \rightarrow F_{\text{post}}) = \frac{\Pr(F_{\text{post}}|F_{\text{pre}}) - \Pr(F_{\text{post}})}{1 - \Pr(F_{\text{post}})} \quad (1)$$

where $F_{\text{pre}}$ is neutral, anxiety, or suspense and $F_{\text{post}}$ is low-fear, mid-fear, or high-fear. $\Pr(F_{\text{post}})$ is the transition probability from any $F_{\text{pre}}$ to the $F_{\text{post}}$ and $\Pr(F_{\text{post}}|F_{\text{pre}})$ is the conditional probability of transition from the $F_{\text{pre}}$ to the $F_{\text{post}}$. Transition likelihood $L$ returns a value ranging from $-\infty$ to 1 where $L > 0$ indicates a likely transition with increasing likelihood as it approaches to 1, $L = 0$ indicates that the transition is equal to chance, and $L < 0$ means that the transition is less likely to occur compared to the base rate of transition into the $F_{\text{post}}$.

Table I shows the result of transition likelihood values for each pair between $F_{\text{pre}}$ and $F_{\text{post}}$ with the positive transition likelihood values highlighted. It indicates that the participants were likely to transition into high-fear if they confronted with the threat while experiencing suspense.

<table>
<thead>
<tr>
<th>$F_{\text{pre}}$/$F_{\text{post}}$</th>
<th>Low-Fear</th>
<th>Mid-Fear</th>
<th>High-Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>0.06</td>
<td>0.22</td>
<td>−0.26</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.09</td>
<td>0.10</td>
<td>−0.16</td>
</tr>
<tr>
<td>Suspense</td>
<td>−0.05</td>
<td>−0.12</td>
<td>0.14</td>
</tr>
</tbody>
</table>

This result suggests that keeping players in suspense beforehand is the best way to maximize the effectiveness of producing a scary event. This result also supports Perron’s study [18] that showed many successful examples of forewarning techniques used in horror movies and horror games. It suggests that long anticipation of a harmful confrontation (i.e., suspense) is more disturbing than short anticipation (i.e., surprise).

IV. AFFECTIVE SURVIVAL HORROR GAME DEVELOPMENT

As mentioned above, keeping players in suspense beforehand is the best way to maximize the scariness of a scary event. In the case of affective survival horror games, the pre-fear affect can be used as an input so that the game will detect if the previous forewarning technique was enough to keep the player in suspense. If not, the game can delay the scary event and make use of other game elements such as
sound and a distorted vision to try to elicit suspense before actually showing the scary event to the player.

In this work, we developed a new game named “Slender Affect” (SA) using the Unity\(^4\) game engine (Version 4.1.2f1). It is a STEP-like game, but we did not use STEP itself because we did not have a control over the game and could not make any modification we needed. We set two goals for this game development, which were

- To create a STEP-inspired game that is adaptable, more simple, and shorter while keeping the same level of enjoyment and scariness as the original game, and
- To compare affective and non-affective versions of the game.

As shown in the previous section, our previous work [11], [12] indicated that the prediction model trained from EKG features with C4.5 decision tree was the best. Therefore, in this work, we use EKG signals and the model to predict the affective states of the player and control the game.

We first introduce the overall process from EKG signals to affective-state prediction and game control. Then, we introduce SA and its implementation policies that make the game more friendly for players who are not familiar with games, and introduce two versions of SA; one is the affective version and the other is the non-affective version. Finally, we explain the experiment we conducted and show its result.

A. Overall Process

Figure 3 summarizes the overall process of collecting and predicting the affective states of the player in real-time. From EKG sensors attached to the participant, EKG raw signals are delivered to the feature extraction method. The feature extraction method outputs EKG features that are used in the prediction method. The prediction method uses the EKG features and the prediction model constructed from the data of the previous experiment. In the previous experiment, the features were normalized by the minimum and maximum values of the entire session that are not known in advance. Therefore, we gradually update the minimum and maximum values for the normalization. Finally, the predicted affective states are used for game control. The overall process has to be done in real-time.

![Fig. 3. The overall process for predicting the affective states of the player and using them for game control in real-time.](image)

However, unfortunately, due to the restriction of the EKG device we used on extracting EKG features in real-time, we were not able to use HR, IBI, and NN50 for real-time prediction, which have to be calculated every time a new heart beat is detected. Therefore, we constructed the prediction model again using only long-time features, i.e., \(\mu_{HR}\) and \(\sigma_{HR}\) with C4.5 decision tree as a classifier. Ten-fold cross validation, again, showed that the f-measure values become 0.86 for pre-fear affects (reduced from 0.92) and 0.75 for post-fear affects (reduced from 0.85).

B. Game Design

Most of the STEP design concept still remains. However, we aimed to reduce the difficulty of the game so that it is easier for non-gamers to collect more pages. In the previous experiment, we found a problem that many players spent a lot of time walking randomly while they could not find or found only few pages, which consequently reduced the chance of encountering with the Slender Man. Hence, the map of SA was designed in a way that eight pages of paper are placed on eight areas instead of placed randomly on ten areas in STEP, and that the first page can be found easily at the start of the game. Figure 4 shows the map of SA.

![Fig. 4. The map of Slender Affect. Triangles indicate the start points of each gameplay and circles indicate the positions where the pages are placed.](image)

The Slender Man is controlled by a simple rule-based system. Main inputs to the system are the distance between the Slender Man and the player, and the visibility of the Slender Man. The distance is divided into three ranges: pursuit, damage, and capture. The basic rules of the Slender Man are as follows:

- The Slender Man roams randomly if the player is out of the pursuit range.
- The Slender Man will move straight to the player direction if in the pursuit range. The player hears stomping sound with probability 0.5.
- If the player is in the damage range, the player’s sanity will start decreasing. The player can know it because static noise is shown up on the screen and becomes stronger as the sanity level decreases.
- If the player is in the capture range or the player’s sanity drops down to zero, the player loses the game.

\(^4\)http://unity3d.com/
If the player is facing to the Slender Man, which is the scary event of this game, the Slender Man will stop moving, but the player’s sanity will start decreasing.

The Slender Man can teleport to the player if the player is not looking at it.

The teleportation is the key event of SA. There are three types of it:

- The Slender Man appears in front of the player after static noise covers the screen for 0.2 seconds (called a forced scary event or FSE). FSE really occurs only when the player’s sanity is full.
- The Slender Man changes its position randomly behind the player after static noise covers the screen for 0.2 seconds (called a fake scary event or FKSE).
- The Slender Man changes its position randomly behind the player without noise (called simple teleportation or ST).

Two versions of the game, affective and non-affective ones, were created. The difference is mainly in the rules of teleportation. In both versions, teleportation is triggered when the player collects a page (except for the first one) or 60 – 7n seconds elapses since the previous trigger, where n is the number of collected pages. In the non-affective version, the teleportation rules are as follows:

- When triggered, FSE, FKSE, and ST are chosen with probability 0.3, 0.1, and 0.6, respectively.
- When the player is turning to the Slender Man, ST occurs with probability 0.15.

In the affective version, on the other hand, the rules are as follows:

- When triggered, FSE is chosen if the player is in suspense. If he/she is in anxiety, FKSE and ST are chosen with equal probability. If in neutral, FKSE and ST are chosen with probability 0.25 and 0.75, respectively.
- In addition, after the first page is collected or 4 minutes elapses since start, FSE is forcibly triggered if the player is continuously in suspense for 8 – 0.6n seconds. If he/she saw the Slender Man in the latest 30 seconds, however, this duration is doubled.
- When the player is turning to the Slender Man, ST occurs if he/she is in neutral. That is, the player does not see the Slender Man at all when in neutral.

In summary, based on the findings in the previous experiment, the affective version makes use of the player’s affective state to delay the scary event when the player is in neutral, and force the scary event when the player is in suspense.

We remove the ability to turn on/off flashlight from STEP so players do not have to worry about the battery of flashlight and the control becomes more simple. Instead, we add flashlight malfunctions, i.e., flickering for 15 seconds and blackout for 8–15 seconds. These malfunctions are activated every 60 seconds with probability 0.33 each. In addition, in the affective version, the game elements that create tension will be activated when the player is continuously in neutral for 38.5 – 2.5n seconds. They are the flashlight blackout, static noise, stomping sound, and FKSE, which are randomly chosen.

The rules and parameters described above were based on the results of preliminary experiments with three players who were out of the experiment. Also, the preliminary experiments showed that the flow of the game was destroyed (e.g., FSE happened too much frequently) if the game was controlled by the EKG features collected every second as in the previous work. Therefore, we changed the frequency of EKG feature collection to every 5 seconds in the experiment.

C. Experiment Procedure

The main objective of this experiment is to compare how much fear and fun the players felt while they are playing the affective and non-affective versions of SA. With the record of affective states of the players and game event logs, we can investigate the transition likelihood of affective states caused by each game element as well.

First we set up the tools used for collecting data. The EKG sensor position image will be shown to the participant and he/she has to attach the sensor on his/her body according to the image. There is a chance that EKG sensors are not placed properly and the signal might not be clear enough for detecting the participant’s heart beats. So, we first test capturing EKG signals and see whether HR can be calculated or not. If not, we need to inform the participant to adjust the placement of EKG sensors and perform the test again.

After the sensor is attached properly, we inform the participant that

- he/she has to play a survival horror game while his/her heart rate signal, facial expression, and gameplay video are being recorded,
- the game will start with a tutorial stage where he/she can learn the basic control of the game,
- after the tutorial stage, he/she has to play the game for 4 times,
- before each game starts, he/she needs to sit back and rest for 2 minutes,
- a couple of questions will be shown up on the screen between each game, which he/she has to answer by pressing the number on the keyboard, and
- while playing the game, he/she should not move his/her body if not necessary.

On the other hand, we do not tell the participant

- how we use his/her heart rate signal,
- that there are affective and non-affective versions of the game, and
- which questions are going to be asked between the games and after the whole session.

When the participant does not have any more questions and is ready to play the game, we run the game and start the overall
process of Fig. 3 that will send the participant’s affective state to the game in real-time. The game starts with the tutorial stage. The main objective of the tutorial stage is to teach the participant how to control using keyboard and mouse and to give him/her a chance to try out the actual control in the real setting while keeping him/her from seeing the actual game that he/she is going to play. Another objective is to tell the background story of the game. This part is necessary to raise the participant’s anticipation toward the scary things in the game and to cause every participant to have the same feeling before playing the actual game. At the end of the tutorial stage, the participant has to answer two questions asking

- how much experience the participant has in FPS game control (no experience, tried but not familiar, or familiar), and
- whether the participant likes horror movies/games (no, neutral, or yes).

After finishing the tutorial stage, the participant goes to the series of rest, play, and evaluation for 4 times. Before starting each game, the participant has to sit back and rest for 2 minutes. The feature extraction method starts recording the minimum and maximum values of heart rate (HR) and interbeat interval (IBI) during the rest period and the recorded values are used for normalization throughout the game. The participant plays the affective (A) version twice and the non-affective (N) version twice. We predefined two play orders for this experiment, NAAN and ANNA, one of which is chosen randomly for each participant. The order was predefined so that each participant plays both NA and AN pair and the random order was used to reduce the order effect. Whether the participant won or lost the game, he/she has to rate the fun, fear, and difficulty of that game with a 5-point scale from least to most. After finishing all the 4 games, the participant also has to rate the overall fear, fun, and difficulty of the games in the post-questionnaire with the 5-point scale. We also ask the participant if he/she recognized any different settings among the 4 games and if he/she recognized anything that changed or reacted to his/her emotions.

D. Participants

Twelve participants (eleven males and one female) aged between 22 and 36 years (mean = 25.42, SD = 3.75) took part in the main experiment. Six of them are Thai, four are Japanese, one is Filipino, and the last one is Korean. There was one participant who considered himself as a gamer, four participants were not, and seven participants considered themselves as casual gamers. Six participants did not like horror movies nor games, four participants felt neutral, and two participants liked to watch horror movies and/or play horror games. There were four participants who did not know the FPS type controls and four participants who knew the controls but were not familiar with them. There were two participants who had played STEP before and other two participants who knew STEP but had not played it before. Six participants played in the NAAN order and the other six played in the ANNA order.

E. Results

From the post-questionnaire, the average of overall fear ratings was 4.33 (SD = 0.65) and the average of overall fun ratings was 3.5 (SD = 0.91). When comparing the values with the previous experiment, overall fear ratings of STEP were significantly higher than overall fear ratings of STEP ($p < 0.05$) while there was no statistical difference between overall fun ratings. From the individual gameplay ratings at intermission, the average of fear ratings was 3.65 (SD = 1.00) and the average of fun ratings was 3.52 (SD = 1.01), similar to the previous experiment in which fear ratings were higher than fun ratings. These results show that SA has comparable enjoyment and scariness levels with STEP.

Next we look into each gameplay separately. There was one gameplay in which the participant did not get any pages and headed toward the Slender Man because the participant did not want to play anymore. We remove that pair of the gameplay out (both affective and non-affective), so there are 23 affective gameplay and 23 non-affective gameplay in total. Table II shows the average and standard deviation of fear, fun, and difficulty ratings, the number of collected pages, and the duration in seconds of individual gameplay, with respect to the repetition of play. There are no significant differences in the ratings. Although some participants stated that they felt less fear in the latter plays, repeating the game was not the main factor that influenced the ratings as the result shows that the average fear rating went up at the fourth gameplay. The increasing number of collected pages indicates the improvement of skills owing to the repetition of gameplay.

### Table II. Average and standard deviation of the ratings, the number of collected pages, and the duration of individual gameplay, with respect to the repetition of play.

<table>
<thead>
<tr>
<th>Play</th>
<th>Fear (SD)</th>
<th>Fun (SD)</th>
<th>Difficulty (SD)</th>
<th>Pages (SD)</th>
<th>Duration (s) (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>3.42 (0.90)</td>
<td>3.33 (0.89)</td>
<td>3.67 (0.99)</td>
<td>3.08 (1.88)</td>
<td>433 (349)</td>
</tr>
<tr>
<td>2nd</td>
<td>3.92 (0.79)</td>
<td>3.5 (1.17)</td>
<td>4.17 (0.84)</td>
<td>3.25 (2.09)</td>
<td>340 (181)</td>
</tr>
<tr>
<td>3rd</td>
<td>3.46 (0.93)</td>
<td>3.82 (0.98)</td>
<td>4.00 (1.00)</td>
<td>3.81 (1.54)</td>
<td>348 (206)</td>
</tr>
<tr>
<td>4th</td>
<td>3.82 (1.33)</td>
<td>3.46 (1.04)</td>
<td>4.18 (0.75)</td>
<td>4.27 (1.42)</td>
<td>452 (235)</td>
</tr>
</tbody>
</table>

As the repetition of gameplay did not show any significant differences, we divided the results by the version (Table III). Overall, the participants spent more time and collected more pages in the non-affective version, and they rated the version with higher fear, fun, and difficulty. Although the differences are not significant, these results indicate that there might be some balance issues between the two versions that influenced the participants to collect more pages and have longer play time on the non-affective version. The number of FSEs the game activated can be a good indicator of this issue because FSE increases the chance that the player is captured. The game activated can be a good indicator of this issue because FSE between the two versions, which might shorten the play time in the affective version.

Another issue we found on the game is that none of the participants noticed that there were two different versions of the game. It was our design choices that we did not make the explicit differences between the two versions because we wanted the participants to play and evaluate the game with no bias. However, the different timing of scary events might
TABLE III. AVERAGE AND STANDARD DEVIATION OF THE RATINGS, THE NUMBER OF COLLECTED PAGES, AND THE DURATION OF INDIVIDUAL GAMEPLAYS, WITH RESPECT TO THE VERSION.

<table>
<thead>
<tr>
<th>Version</th>
<th>Fear</th>
<th>Fun</th>
<th>Difficulty</th>
<th>Pages</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective</td>
<td>3.50</td>
<td>3.36</td>
<td>3.96</td>
<td>3.29</td>
<td>361</td>
</tr>
<tr>
<td></td>
<td>(0.96)</td>
<td>(0.85)</td>
<td>(1.00)</td>
<td>(1.73)</td>
<td>(283)</td>
</tr>
<tr>
<td>Non-Affective</td>
<td>3.91</td>
<td>3.82</td>
<td>4.13</td>
<td>3.67</td>
<td>413</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(1.05)</td>
<td>(0.80)</td>
<td>(1.93)</td>
<td>(208)</td>
</tr>
</tbody>
</table>

not have as enough impact on the overall evaluation of each gameplay as we assumed.

Next we looked into the transition likelihood of the affective states of the participants in the non-affective game. We recorded pre-fear and post-fear affects of the participants, which were obtained from their EKG features with the affective-state model, before and after each game element was shown. Then, we used the data to calculate the transition likelihood using Eq. 1. Table IV shows the transition likelihood of FSE that is the controllable scary event trigger and causes the transition from a pre-fear affect to a post-fear one. The result shows that

- To mid-fear, neutral and anxiety had the positive transition likelihood,
- To high-fear, only suspense had the positive transition likelihood, and
- To low-fear, suspense had the positive transition likelihood, too.

It is similar to the previous experiment (Table I) when transitioning to mid-fear and high-fear, but it is completely opposite when transitioning to low-fear. Analyses of transition show that the ratio of transition from neutral and anxiety to high-fear was much higher than that in the previous experiment, whereas that from suspense was not so different. Then, Pr(high-fear) becomes relatively large and Pr(low-fear) becomes relatively small. As a result, \( L_{\text{neutral} \rightarrow \text{low-fear}} \) and \( L_{\text{anxiety} \rightarrow \text{low-fear}} \) become small whereas \( L_{\text{suspense} \rightarrow \text{low-fear}} \) becomes large. It is not clear why the transition from neutral and anxiety to high-fear increased, but it may relate to the fact that the overall fear ratings of SA were significantly higher than those of STEP. Also, it may be due to the frequency change of EKG feature collection, which may have deteriorated the prediction model itself, flawed the post-fear affect detection due to delayed prediction, and/or simply reduced available data for the likelihood analysis.

TABLE IV. Transition likelihood caused by forced scary events

<table>
<thead>
<tr>
<th>( F_{\text{pre}} / F_{\text{post}} )</th>
<th>Low-Fear</th>
<th>Mid-Fear</th>
<th>High-Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-0.18</td>
<td>0.21</td>
<td>-0.05</td>
</tr>
<tr>
<td>Suspense</td>
<td>0.33</td>
<td>-0.38</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table V shows the transition likelihood between pre-fear affects when game music changed to a more intense one, triggered after the number of collected pages became an odd number. Mostly, the participants stayed in the same state after the music changed.

Table VI shows the transition likelihood between pre-fear affects when the flashlight started flickering. The participants were likely to stay in the same state after the flickering occurred. The transitions from neutral to suspense mostly happened only on the first time flickering occurred. After the participants figured out that the flickering did not have any relation to the appearance of the Slender Man, they transitioned from suspense to anxiety.

TABLE VI. Transition likelihood caused by flickering flashlight events

<table>
<thead>
<tr>
<th>Before/After</th>
<th>Neutral</th>
<th>Anxiety</th>
<th>Suspense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>0.08</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-0.06</td>
<td>0.33</td>
<td>-0.29</td>
</tr>
<tr>
<td>Suspense</td>
<td>-0.21</td>
<td>0.05</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Unfortunately, many game elements were not occurred as often as the music and flashlight malfunctions, so it is difficult to make any good interpretation without more data. Nevertheless, these results show an example on how the effectiveness of game elements can be examined individually using the real-time affect detection tool. This information can be later used to improve the game in various ways.

V. Conclusion

In this work, we illustrated the methodology for developing an affective survival horror game. Based on the previous work in which we analyzed the affective states of players with our affect annotation tool (AAT) and constructed the affective-state prediction model from EKG signals, we developed an affective survival horror game named “Slender Affect (SA)”. It is similar to “Slender: The Eight Pages (STEP)”, a famous survival horror game in the world, and the result of experiment shows that it has comparable enjoyment and scariness levels with STEP.

Two versions of SA, affective and non-affective, were constructed. The affective version uses the affective-state prediction model to predict the player’s current affective state and uses the state to change the timing of scary events. The non-affective version controls the scary events regardless of the player’s affective states.

We conducted an experiment to compare between the affective and non-affective versions of SA. Twelve participants took part in the experiment. In the experiment, each participant played both versions two times each and rated the fun, fear, and difficulty of games with a 5-point scale.

The result shows that the ratings of the affective version was less than those of the non-affective version, although the differences were not significant. It may be because the frequency of the scary event in the affective version was more than that in the non-affective one and it shortened the play time of the affective version. Also, none of the participants noticed that the timing of scary events was controlled by their affective states. It was from our design choices, but the different timing of scary events might not have as enough impact on the overall evaluation of each gameplay as we assumed.
Using SA, we can individually analyze the transition likelihood of the affective states of participants caused by a specific game element. This information can be later used to improve the game in various ways.

As a future work, SA itself should be improved. In particular, we have to remove the unbalance of FSE activations between the affective and non-affective versions. Also, we would like to analyze the result from various points of view. For example, the result may depend on features of the participants such as gender, ethnicity, game experience, and horror preference. In order to analyze such effects, we have to conduct a massive experiment in which a lot of participants take part.

ACKNOWLEDGMENT

This work was partly supported by JSPS Core-to-Core Program, A. Advanced Research Networks.

REFERENCES


