Sonancia: a Multi-Faceted Generator for Horror

Phil Lopes
Intitute of Digital Games
University of Malta
Msida, Malta
Email: louis.p.lopes@um.edu.mt

Antonios Liapis
Intitute of Digital Games
University of Malta
Msida, Malta
Email: antonios.liapis@um.edu.mt

Georgios N. Yannakakis
Intitute of Digital Games
University of Malta
Msida, Malta
Email: georgios.yannakakis@um.edu.mt

***Abstract*—Fear and tension are the primary emotions elicited by the genre of horror, a peculiar characteristic for media whose sole purpose is to entertain. The audience is often lead into tense and fearful situations, meticulously crafted by the authors using a narrative progression and a combination of visual and auditory stimuli. This paper presents a playable demonstration of the Sonancia system, a multi-faceted content generator for 3D horror games, with the capability of generating levels and their corresponding soundscapes. Designers can also guide the level generation process, by defining an intended progression of tension, which the level generator and sonification will adhere to.**

I. Introduction

Digital games can be defined as a synthesis of different types of content, such as sound, visuals and level architecture, which provide different interactive entertaining experiences for players [1]. Procedural content generation (PCG) systems have often concentrated on the level architecture facet of creativity [2], however more recently various efforts have been made for procedurally generating playable experiences that blend a variety of different faceted content such as: audio and gameplay [3]; level architecture and gameplay [4]; and audio, level architecture and visuals in earlier studies of the Sonancia generation [5]. The orchestration of audiovisual content can augment and provide meaningful experiences as digital games are primarily an interactive audio-visual activity. This is especially true in the survival horror genre, where the effective use of lighting, audio and labyrinths provide tense and frightening experiences [6].

This paper showcases a playable demonstration of Sonancia, a multi-faceted procedural content generator for the horror genre. Sonancia can generate multi-faceted levels by first evolving the level's layout, lighting, enemy positioning and 3D diegetic audio placement. Once a structure has been created, the system will then pick and allocate background audio assets within the level (i.e. sonification) for the creation of its soundscape. Sonancia targets the survival horror genre as a specific case study for the concepts of multi-faceted blending, due to the genre's heavy reliance on the audio and visual facets [6]. It also follows a specific game progression scheme based on tension, where the focus is to lead the audience into uncomfortable and fearful situations. Sonancia attempts to simulate this scheme by generating levels through a *tension frame*, allowing designers to define the intended rise and fall of tension guiding the level generation process, which will adhere to the defined frame. Sonification will also take into account the tension frame by proxy through the level generation. It allocates background audio assets by following the level's tension progression and a crowdsourced ranking annotation attributed to each background audio asset available in the Sonancia library. The methodologies used in this demo are detailed in [5], while the autonomous generation of levels using dramatical tropes are described in [7]. In this demo, however, several features are added to the evolutionary blending component of Sonancia, including the placement of lighting and 3D audio (not to be confused with background audio) within the level layout. For demonstration a play-through of the demo can be found here[1] and downloaded here[2].

II. Sonancia

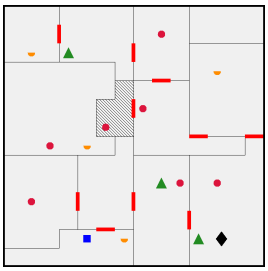
The generated game consists of a 3D first-person survival horror game, taking place in a dark underground dungeon. These dungeons consist of multiple rooms that are interconnected through doorways. Each dungeon contains an ancient statue that players must reach in order to complete the level, acting as the objective. These dungeons also host several monsters that will run after the players, if they are in their line of sight. If the player gets hit more then three times by a monster, they will be defeated and the level will restart. Players can not directly fight back, but they can run away and use stealth to get past and progress through the level.

A. Level Generation

Sonancia generates levels using a genetic algorithm without recombination, and a human or machine defined framing of tension. Due to spatial constraints, this paper will briefly describe the algorithm, but for the interested reader a more detailed description can be found in [5].

The level structure in the genotype consists of an array of integers that represent a tile within the level and what room it belongs to. Mutations will shift walls, divide rooms and move, place or remove monsters, lights, doors, and 3D audio in rooms, if possible. The level generation currently applies the following constraints: than one type of the same object (i.e. monster, light or 3D audio) may be placed in the same room; rooms must have at least 5 tiles; and a path between the start and objective must exist.

[1] http://goo.gl/2RQDG6
[2] http://goo.gl/jcXcBu



(a) Example 2D level

(b) 3D Level of 1a.

Fig. 1: Figure 1a depicts a 2D illustration of a dungeon, consisting of rooms and interconnecting doors. Thin black lines represent walls, while thick red lines represent doors. Monsters are represented as green triangles, the objective is represented as a blue square, lights are circles, 3D audio are half circles, and the black diamond is the player position in Fig. 1b. The darker room is the player's starting room. Figure 1b consists of gameplay footage of the level represented in Fig. 1a.

The frame represents the intended progression of the rise and fall of tension that the level generator must adhere to. More specifically, a frame consists of a line graph where the y-axis represents a numerical representation of tension, while the x-axis consists of a specific room in the level. This frame is then used as the fitness function for the level generator, which informs the optimal number of rooms for the level, and how monsters, lights and 3D audio will be distributed throughout the level. For example, adding a monster or a 3D audio (to a lesser degree) raises the tension of that room; adding lights decreases the tension slightly as it is less dark compared to other rooms, allowing players to orient themselves better. Sequences of rooms that do not increase in tension will continuously suffer a tension decay, in order to simulate the player relaxation after a stressful event.

B. Sonification

Level sonification consists of the selection and allocation of *background audio* assets that will loop during gameplay. Two constraints are enforced by the sonification algorithm: only one background audio piece may be allocated per room, to avoid cacophony; and audio pieces may not repeat in the same level. Sonification will also select and position *3D audio* assets within the rooms, which were defined by the previous level generation process.

Background Audio is allocated according to the characteristics of a generated level, such as the distribution of monsters, lighting and which rooms the 3D audio assets were placed. More precisely it follows the tension that was obtained through the generation process (the actual tension progression), and not the previously defined tension frame. This allows sonification to more closely adapt a soundscape that follows the exact characteristics of the level, instead of the conceptual specifications defined by the tension frame.

All background audio assets in the library (40 in total) are ranked based on human-annotated tension preferences. The *global order* of sound tension is derived through the pairwise preference test statistic [8], which is used by the sonification algorithm to select which sounds are better suited for a specific section of the level based on both the tension progression of the generated level and the annotated audio piece. For the interested reader a video demonstrating some examples of background audio is available here[3].

3D Audio consists of audio cues that play only once, when players trigger them within the environment. Although the level generation specifies these sounds should be located in, the sonification system places the audio trigger event within the room itself, by calculating the mid-point of the shortest path between two doors of the room. If a room contains only one door, the trigger is placed at the room's centre. For the interested reader a video demonstrating examples of 3D audio is available here[4].

III. Future Work

Future work will include enhancements to the sonification system such as the creation of a data-driven model of tension, allowing the system to automatically annotate and more accurately choose between which audio assets in comparison to others. Extensive user testing will also be conducted, for both validating and improving the multi-faceted game content generation algorithms developed.

Acknowledgment

This work was supported, in part, by the FP7 Marie Curie CIG project AutoGameDesign (project no: 630665) and the Horizon 2020 project CrossCult (project no: 693150).

References

[1] A. Liapis, G. N. Yannakakis, and J. Togelius, "Computational game creativity," in *Proceedings of the International Conference of Computational Creativity*, vol. 4. Springer, 2014, pp. 71–78.

[2] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, "Search-based procedural content generation: A taxonomy and survey," *Conference on Computational Intelligence and Games*, vol. 3, no. 3, pp. 172–186, 2011.

[3] A. K. Hoover, W. Cachia, A. Liapis, and G. N. Yannakakis, "Audioinspace: Exploring the creative fusion of generative audio, visuals and gameplay," in *Proceedings of the EvoMusArt conference*. Springer, 2015, pp. 101–112.

[4] M. Cook and S. Colton, "Multi-faceted evolution of simple arcade games." in *CIG*, 2011, pp. 289–296.

[5] P. Lopes, A. Liapis, and G. N. Yannakakis, "Targeting horror via level and soundscape generation," in *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*, 2015.

[6] I. Ekman and P. Lankoski, "Hair-raising entertainment: Emotions, sound, and structure in silent hill 2 and fatal frame," *Horror video games: Essays on the fusion of fear and play*, pp. 181–199, 2009.

[7] P. Lopes, A. Liapis, and G. N. Yannakakis, "Framing tension for game generation," in *Proceedings of the International Conference on Computational Creativity*, 2016.

[8] G. N. Yannakakis and J. Hallam, "Ranking vs. preference: a comparative study of self-reporting," in *Affective computing and intelligent interaction*. Springer, 2011, pp. 437–446.

[3] https://youtu.be/dtcV6v5fHHA

[4] https://youtu.be/FUhzm6UiudI