Sound Similarity as a Tool for Understanding Player Experience

Applying Similarity Matrix to Gameplay Performance Segmentation

Raphaël Marczak, Gareth Schott, & Pierre Hanna

Transactions of the Digital Games Research Association
2016, Vol. 2, No. 3, pp. 221-242
ISSN 2328-9422
http://todigra.org

ABSTRACT

Analytical accounts chronicling engagement with digital games can always benefit from empirical data outlining the patterns of behavior produced by different players as they engage with the same game, or similar sequence within a game. This paper presents an extension to a novel method, termed *feedback-based gameplay metrics*, which exploits the audio and visual output of an activated game to produce accounts of player performance. This paper offers an account of an affiliated method, based on similarity matrices, which is derived from the same measurement process and that has yet been applied to the interests of game studies (over design oriented research) to determine the similarity or diversity within encounters with particular games. This paper introduces

the method and illustrates its potential applications in the analysis of performance.

Keywords

Similarity Matrix, Sound Processing, Sound similarity, Player Experience, Gameplay Performance Segmentation.

INTRODUCTION

Trying to understand the specific experience that represents playing a videogame has been a core area of research in game studies for more than a decade now. This is notably challenging because a videogame is "both an object and process that must be played, [and] playing is integral, not coincidental" (Aarseth, 2001). Thus, to understand a player's experience, it is necessary to be able to assess the way a player goes about fulfilling the need of the system to be activated in play. Numerous approaches have been designed that not only account for the manner in which players actually engage with a game system, but also for the rationale behind their actions and interactions. A large range of works variously address how players engage in games. These works include more theoretically-oriented approaches, such as Gordon Calleja's (2011) work on player involvement, which speculates as to what constitutes the main factors explaining why players continue to engage with game systems. The literature also includes more methodologically-oriented approaches, such as studies that accurately trace and log the different interactions between the player and a game system (Drachen, Thurau, Togelius, Yannakakis, & Bauckhage, 2013; Kim et al., 2008). Some approaches blend theory and method, such as the analysis of how flow theory (Csikszentmihaly, 1990) might be translated to game systems (Nacke & Lindley, 2008).

When hours of play experience need to be understood, summarized and/or visualized, several approaches have been suggested to automatically process and analyze the play sessions. This is for instance the case of *gameplay metrics* (Drachen, Seif El-Nasr, & Canossa, 2013), which are time-stamped quantitative data about player interaction automatically logged by the game system while activated; *player modelling* (Yannakakis, Spronck, Loiacono, & André, 2013) which focusses on understanding players in order to create a computer and mathematical model that can be used within the game system to improve the experience of play; or *biometric storyboard* (Mirza-Babaei, Nacke, Gregory, Collins, & Fitzpatrick, 2013), which displays biometric signals along with other core measures in order to propose an exhaustive representation of a play session.

The notion of experience can be expended to other measures and perspectives. In this paper, we suggest the use of *similarity matrix* for automatically summarizing and visualizing a play performance through the detection of segment of plays that carry strong similarities. More precisely, this paper seeks to demonstrate how the production of a similarity matrix, based on a sound analysis of audio outputs from game play can be used in order to perform a segmentation of a gameplay performance to express the manner players engage with selected games. We use the term performance (see Laurel, (1993), in order to insist on the fact that we focus on *gameplay* as relative to specific activations of the play – meaning that various player performances with the specific game are segmented – rather than the *game* as an absolute entity. This particular method of *gameplay performance segmentation* seeks to emphasize the extent to which individual player experiences with linear structures conform or diverge. It is important to specify that *gameplay performance segmentation* does not override the *gameplay segmentation* notion as previously defined by Zagal et al. (2008), which represents “the manner in which a game is broken down into smaller elements of gameplay” (Zagal et al., 2008, p. 178). Gameplay performance segmentation must be seen as a continuation of gameplay performance: after identifying gameplay elements using gameplay segmentation, it then becomes possible to focus on the evolution of each determined segment using a gameplay performance segmentation approach.

A similarity matrix represents a meaningful approach to segmentation for both the assessment and representation of the similarities between different documents or similarities within the same document (termed auto-similarity matrix). A similarity matrix can be employed for a large variety of modalities, such as textual documents (Choi, 2000), visual documents (Cooper & Foote, 2001) and musical documents (Hanna, Robine, & Ferraro, 2008). However, while similarity matrices have been successfully applied within computer science, they have yet to be employed within game studies to aid understanding and assessment of player experience. Having made this point, an approach that seeks to assess player experience through the automatic analysis of audio-visual streams has recently emerged in the form of *feedback-based gameplay metrics* (Author et al., 2013; Author et al., 2014; Author, 2015). This method exploits the audio and video feedback streams produced by a game once it has been activated and recorded, to process it as data in order to describe the manner in which a player engages with a specific game system. As similarity matrices can be produced based on sound or video data, this paper outlines an exploration into the potential of this form of analysis as a component of feedback-based gameplay metrics. The main contribution of feedback-based gameplay metrics is that they can be captured from any game, whether the source code is available or not, thus offering access to a wider range of games. Moreover, feedback-based gameplay metrics represents a post-processing method, allowing an analyst to explore the data however they wish and as many times as they wish. Currently, however, feedback-based gameplay metrics requires a pre-analysis stage in order to elicit the significant elements of the game to be processed by the method. What similarity matrices offer this mode of game metrics gathering is a means of exploiting the sound stream produced by the game play performance without the need for any pre-analysis.

In this paper, three different usages of similarity matrix are illustrated, using three different games in order to also demonstrate the broad nature of this approach. The first one is dedicated to a comparison of two different gameplay performances produced by separate individuals playing the same game. The second example illustrates the detection of repetitions from within the same performance (that is, which sections of the game are replayed and experienced more than once by the same player). The last example illustrates how it is possible to compare a performance by exploiting the game's soundtrack to study player progression. Before illustrating the creation and analysis of similarity matrices applied to understanding gameplay, it is important to outline what a similarity matrix entails.

Similarity Matrix

A similarity matrix is a mathematical entity consisting of a rectangular array in which each entry describes the degree of similarity between the element represented by the current row, and the element represented by the current column. In the case of media documents, similarities are computed for each sub-units of a document, with every sub-unit of another document. For a textual document for example, a sub-unit can be a word or sentence; for video document, a sub-unit can be a frame; and for an audio document, a sub-unit can be a sound sample. In a similarity matrix, the columns represent the ordered sequence of consecutive sub-units from one document, and the rows represent the ordered sequence of consecutive sub-units from a second document (or the same one in the case of auto-similarity matrix). Each entry of the matrix at the intersection of a row and a column contains the similarity score between the two sub-units represented by the matching row and column.

Each similarity estimation is generally a score between 0 (no similarity) and 1 (identical). Once the matrix has been completely filled, all sub-units of a document have been compared with all the sub-units of the second one; and a score has been given for each comparison. That means that a similarity matrix represents an exhaustive comparison process between two documents (or inside the same one in the case of auto-similarity). It then becomes possible to look for the highest scores in the matrix in order to extract similar sections of documents. A similarity matrix may therefore be used to determine the degree of linearity and freedom designed into a game, the agency of the player to determine how they progress and whether this yields a quite different experience from within the range of experiences available by the game, the nature or style of play employed by players (e.g. explorative or instrumental and goal driven), or the degree of repetition contained within a game experience.

One of the strengths of similarity matrices is their ability to be easily displayed as an image, visualizing the data so that similar sections of documents can be quickly and easily identified. Indeed, a low similarity score can be represented by a white dot, and a high similarity score can be represented by a black dot. Then, similar sub-units of documents can be immediately spotted. Moreover, because the rows and columns represent ordered consecutive sub-units of documents (that is, words in the order in which they appear in a textual document; frames from the beginning to the end of a video document; or samples from the beginning to the end of an audio document), it is also easy to detect not just similarities between sub-units, but similarities between consecutive sub-units (such as a full sentence or paragraph, or a long sequence inside an audio-visual document). For that, all that is required is to identify black diagonals inside the similarity matrix image representation. Indeed, a black diagonal means that a sequence of consecutive sub-units is fully identical with another sequence. In the following sections, we will focus on the diagonals to interpret the different similarity matrices that have been produced using gameplay audio.

Figure 1 illustrates what a similarity matrix is and what a similarity matrix image looks like. In this figure, Document A has been cut into 20 sub-units and Document B into 15 sub-units. The 2nd sub-unit of Document B is similar to the 8th of Document A for instance, and the 17th of Document A is similar to the 6th of Document B. More than that, it is possible to highlight similar sequences by looking for diagonals. For example, sub-units 3 to 4 in Document A are similar to sub-units 11 to 14 in Document B. If the two documents of Figure 1 represent sounds for instance, with each sub-unit being one second of sound, it would be possible to say that the sequence from 8 seconds to 10 seconds in Document A sounds the same as the sequence from 2 seconds to 4 seconds in Document B (upper diagonal); the sequence from 17 seconds to 18 seconds in Document A sounds the same as the sequence from 6 seconds to 7 seconds in Document B (middle diagonal); and that the sequence from 3 seconds to 6 seconds in Document A sounds the same as the sequence from 11 seconds to 14 seconds in Document B (lower diagonal).

An example of similarity matrix, as currently used for media structural analysis, can be seen in Figure 2, when applied to the understanding of musical structure (Hanna et al., 2008). By comparing a musical creation (*Minuet* part of the *Water Music Suite No1 in F* by *Handel*) with itself (self-similarity matrix, meaning that the rows and the columns represent the same document), it is possible to quickly characterize the structure of the musical piece, in terms of its major themes. Diagonals indicate that parts of the musical piece on the vertical axis are detected as similar to other parts of the same piece on the horizontal axis, thus highlighting on the Figure 2 example the general ABA structure of a minuet.

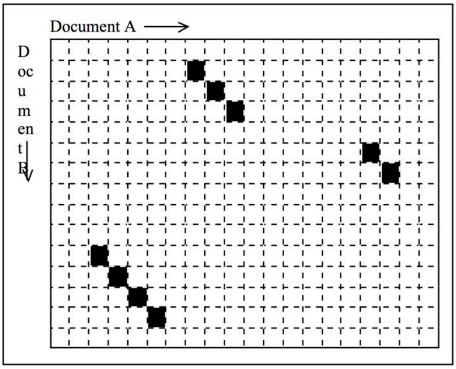

Figure 1: Schematized version of a similarity matrix, illustrating the similarities between two documents. Both documents A and B are segmented in sub-units, and each sub-unit of A (columns) is compared with all the sub-units of B (rows). A black dot represents a similarity, while a white dot represents dissimilarity. By identifying diagonals, it is then possible to characterize contiguous units of A similar to contiguous units of B; thus similarity between full sequences.

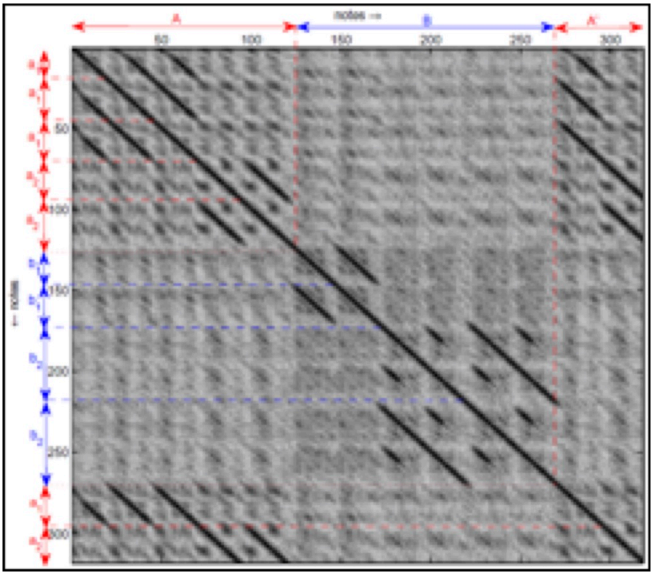

Figure 2 (Hanna et al., 2008): Example of a similarity matrix applied for musical structure analysis. By comparing a musical creation (Minuet part of the Water Music Suite No1 in F by Handel) with itself (self-similarity matrix), it is possible to quickly characterize the structure of the musical piece, in terms of its major themes. Here, the diagonals suggest an ABA structure, representative of the usual minuet structure.

In this paper, and in the following sections, we produced a similarity matrix by using the audio streams generated by games from players' interaction with the game system. To achieve that, we recorded the gameplay footage using a screen-capture software system *FRAPS* (Beepa, 2013), then we discarded the video stream in order to obtain an audio file. The audio stream was then cut in small sub-units of several milliseconds, and each unit was translated into a chroma representation (Serra, Gomez, Herrera, & Serra, 2008). A chroma, or Harmonic Pitch Class Profile (HPCP) is the frequency distribution of a portion of music in terms of the 12 usual semitones of the equal tempered scale. By using the chroma representation instead of the raw sound stream, we ensure that small noises will not have a strong influence on the similarity result. Then, each computed chroma of one sound is compared with the computed chromas from a second stream (in terms of distance, the shorter the distance, the greater the similarity), generating the similarity matrix. Finally, in order to generate the similarity matrix image, a threshold value is selected, under which a dot would be white (dissimilar), and over which a dot would be black (similar). The three examples of similarity matrices presented below are all based on this approach using sound streams and chroma representations.

GAMEPLAY PERFORMANCES: SIMILARITY

The first similarity matrix introduced in this paper has been produced using two performances derived from the game *Max Payne 3* (Rockstar Studios, 2012), by two different players. In this third-person shooter game, the player controls Max, a security guard in charge of the security of a wealthy and famous family. The narrative has a strong role during the gameplay, through two main mechanisms: Max Payne thinking aloud to inform the player about what is happening and what Max recalls in conjunction with the current action; and cut-scene explaining further to the player the context in which he/she interacts. These cut-scenes, sometimes included suddenly between player's actions (i.e., not uniquely between two levels), can be long, and are recognizable through their specific sonic atmosphere. Moreover, the cut-scenes' order of appearance is linear, as they always appear in the same order regardless of the player activations. Then, identifying the cut-scene is a way of identifying a player's progression.

Figure 3 shows a similarity matrix produced using two soundtracks (truncated after 80 minutes for readability purposes) recorded during game sessions with two different participants. Each dotted square represents 5 minutes of play. As explained in the previous section, the patterns to look for are the diagonals, as they represent contiguous sequences of similarity. Figure 4 is an annotated version of Figure 3, highlighting patterns that are interesting and worthy of discussion.

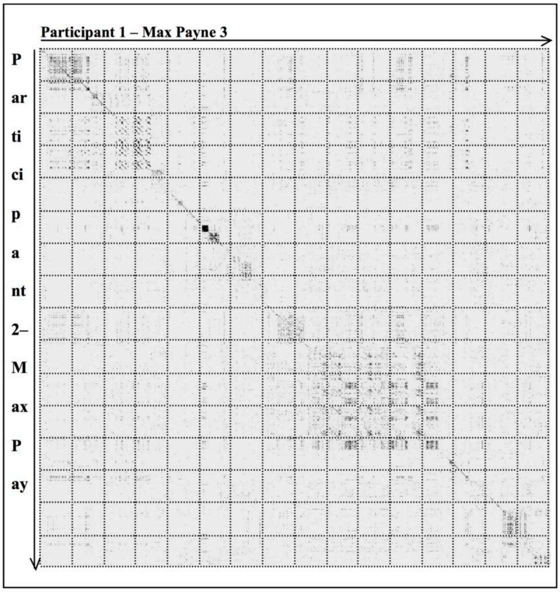

Figure 3: Similarity matrix using the audio of two gameplay performances with the game Max Payne 3. See Figure 4 for a more detailed description of this matrix.

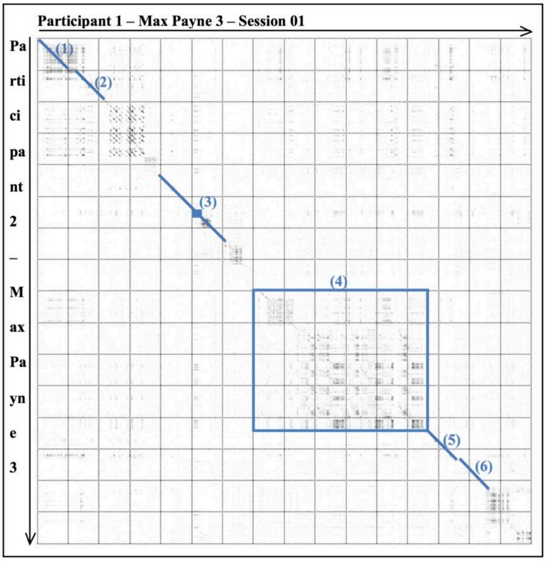

Figure 4: Figure 3 with annotations highlighting the patterns of interest in the similarity matrix

In Figure 4, (1) represents the introduction cut-scene, lasting for around 5 minutes and unskippable, that crossfades directly into the main menu (the cut-scene continues in background, looping on Max drinking and smoking, sitting at a table). (2) represents the first game chapter introduction cut-scene, which is played when the player exits the main menu. It is interesting to focus on the break in the line between (1) and (2), because it indicates different player engagements with the game. Indeed, this break indicates that Participant 1 spent around one minute more in the main menu, while Participant 2 obviously went directly into the game action. For Participant 1, it was important to customize the game to his/her preferences before starting the actual game (probably in order to match preferred control), while Participant 2 did not want to lose any time configuring the gameplay to come. (3) represents the cut-scene between chapter 1 and chapter 2, that Participant 1 achieved after twenty minutes of play, and Participant 2 after twenty-two minutes. This very close duration indicates that both participants had the same level of skill on this level, or that the level is designed to not offer significant latitude to players. The first level of *Max Payne 3* is actually a tutorial, therefore it is reasonable to expect players to take a similar amount of time to complete this level. The developer is likely to have had a desire to keep this level interesting, diverse and not too challenging. The square on the (3) line represents a moment in the cut-scene when the music looped, making all the units in the area similar. (4) highlights the most difficult sequence of chapter 2, and both participants 1 and 2 obviously died a number of times during this sequence. As the game soundtrack loops when a player dies (the music starts again from the beginning, and Max is speaking again to himself in order to recall his current state to the player), the more the diagonals that are present inside the similarity matrix, the more the players had to redo the sequences. Finally, (5) and (6) represent the cut-scene between chapter 2 and 3, which becomes semi-interactive half-way. The player can die during the semi-interactive sequence (they can only move, but not shoot). Participant 2 obviously did not die, as (5) and (6) are vertically continuous, but Participant 1 died once, explaining the horizontal break between (5) and (6). (6) ends when the player regained full interactivity, and participants 1 and 2 played differently from this point onwards, and this, therefore, ends the diagonal.

Thanks to such similarity matrices comparing different player engagements with the same game sequence, it is possible to appreciate the distinct strategies of players (like diagonals (1) and (2) in Figure 4 showing a difference between players who need a customization stage, and players who want to go straight into the action), whilst also demonstrating that difficulty and challenge levels will produce a less fluid experience, causing some players to engage more in some environments rather than others that may have an impact on their motivation, enjoyment and length of game play session (when self-determined outside of research contexts).

GAMEPLAY REPETITION AND AUTO-SIMILARITY

Similarity matrices can also be employed in order to detect repetitions from within a performance. In this case, the similarity matrix is termed auto-similarity matrix. In this paper, we propose to use a performance from the game *Battlefield 3* (EA Digital Illusions CE, 2011) in order to demonstrate the usefulness of auto-similarity matrices applied for the study of gameplay performance. The first-person shooter game *Battlefield 3* uses a less musically induced atmosphere than *Max Payne 3*. However, the death sequences and the loading screen following death are recognizable by their highly specific sound background. When comparing a performance with itself, the repetitive moments in a performance with the game *Battlefield 3* are highlighted, and likely to represent death sequences for this specific game.

Figure 5 shows the auto-similarity matrix generated using a 45-minute session with the game *Battlefield 3*. In auto-similarity matrices, the main diagonal must be discarded, as it represents a sub-unit compared with itself. Moreover, auto-similarity matrices are symmetrical using this main diagonal. In Figure 5 several short diagonals can be distinguished, aligned on the same row or column. This means that all these diagonals represent exactly the same sound. Figure 6 is a zoomed version of the bottom right corner of Figure 5, in order to have a better view of these diagonals.

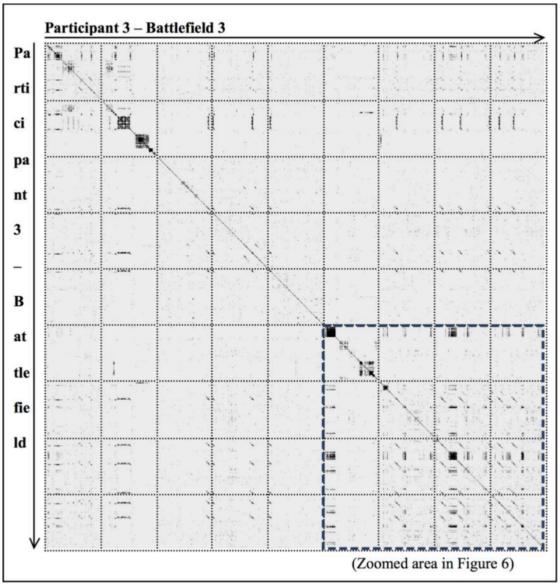

Figure 5: Auto-similarity matrix based on a performance with the game Battlefield 3.

Figure 6 actually represents one of the most difficult sequences in *Battlefield 3*, where the player is asked to protect a bridge from numerous enemies. During this sequence, six deaths can be distinguished at time 1890, 2134, 2233, 2390, 2443 and 2521 seconds, by counting them vertically or horizontally, as showed by the blue lines. But actually, such a matrix representation can highlight more than death screens with *Battlefield 3*. Figure 6 illustrates that a diagonal actually accounts for more than the actual death screen. When dying, a full sequence is repeated: the death screen, the loading screen (with a specific background music) and, importantly, the beginning of the bridge sequence, initiated by the same incoming radio message sent by a member of the team.

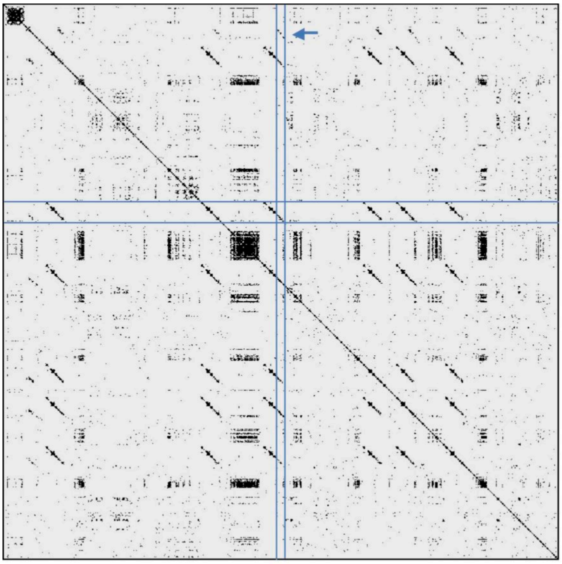

Figure 6: Zoom of the bottom right corner of Figure 5, displaying numerous diagonals indicative of death sequences in Battlefield 3 (see Figure 7)

Knowing that, the top diagonal in Figure 6 (pointed to by a blue arrow), which is shorter and matches the end of all the death diagonals in this section, actually represents the first instance of the radio message, without any prior death. It is then possible to locate the first time the player entered the difficult section, around 1863 seconds. Thanks to the auto-similarity matrix, it is possible to not only quickly identify difficult sequences when the player is forced to play again after dying, but it is also possible to detect the exact entry point of the difficult section of the game.

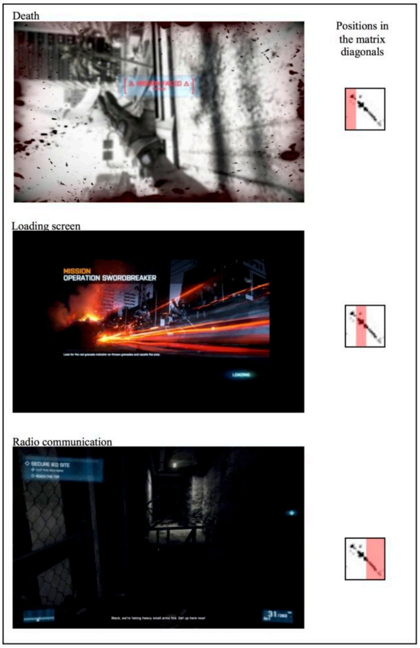

Figure 7: The three sequences constituting a death diagonal in the self-similarity matrix: the actual death screen, the following loading screen, and the restart of the same mission.

Soundtrack SimilarityThe final similarity matrix compares a performance of the game *Super Hexagon* (Terry Cavanagh, 2012) with the original soundtrack of the game. Indeed, in this challenging game, where the main goal consists of surviving for as long as possible, a progressively lively music score accompanies the game play, adding to the intensity. Each time the player loses, the music suddenly stops and is restarted at a random position (anywhere from the beginning to the middle of the score) when the player restarts the game. By comparing the audio of a performance with the original soundtrack, it is possible to have some idea of the player's level of skill. It is important to note, however, that due to the repetitive nature of the music, some square noises appear on the matrix, which can complicate detection of the diagonals. However, it is still possible to gauge the player's skill level.

Figure 8 for instance illustrates a highly skilled player who succeeds in surviving for long periods without losing. The matrix in Figure 8 is actually a 4-minute performance of a player who played the same level 3 times, indicating that each performance lasted roughly only one minute.

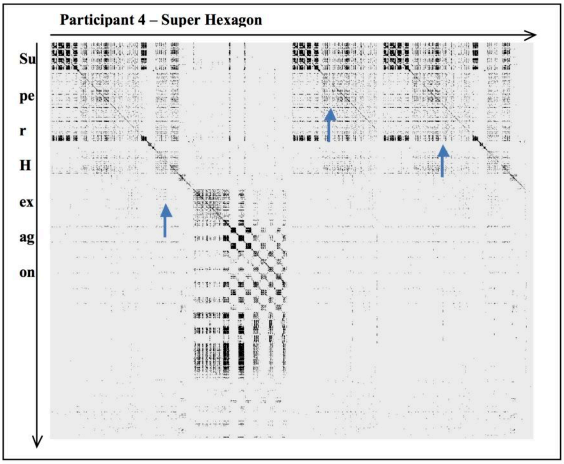

Figure 8: Skilled player interacting with Super Hexagon, surviving more than one minute in each session (played three times in four minutes)

Figure 9, on the other hand, shows a 3-minute performance by a beginner. During this performance, not less than 9 diagonals can be distinguished, indicating that the player never actually survived more than 30 seconds.

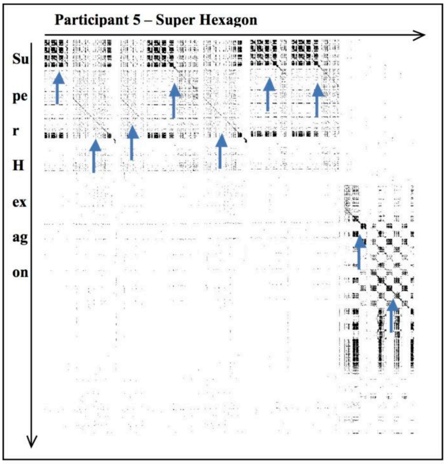

Figure 9: Beginner interacting with Super Hexagon, unable to survive more than 30 seconds (nine "try again" in three minutes).

Thanks to this similarity matrix representation, it is possible, at a glance, to have a clear idea of the skill level of a player, by studying the number of deaths and the repetition of sequences.

DISCUSSION AND CONCLUSION

This paper showed that similarity matrices, commonly used in computer science for segmenting audio-visual material, can actually also be applied to game studies for the analysis of gameplay performances, and understanding of the player experience. Similarity matrices offer both a means of analysis and a means of visualization, in order to ease the work of game studies researchers interested in exploring a particular dimension of player experience with particular games. This paper has promoted and illustrated the value and use of similarity matrices for the analysis of games, based on recorded audio footage, through three different applications: comparison of performances, intra-performance repetitions detection and player progression assessment by using the audio footage and the game's original soundtrack.

Moreover, the outcomes derived from similarity matrices analysis can be applied beyond a mono-modal consideration of performances (i.e., where only similarity matrices are considered on their own or in isolation to describe player experience), and can be used in conjunction with other modalities that provide measures of player experience. The modalities of interest include, but are not limited to, biometry and keystroke. For instance, it would be valuable to study the player's controller inputs while redoing similar sequences, in order to assess if the player is reacting similarly or using a different strategy (approach similar to the one published in a previous DiGRA conference (Author et al., 2013)). It would also be interesting to map the detected similarity outcomes with biometric research (Mirza-Babaei et al., 2013; Author et al., 2014) in order to assess whether similar sequences produce similar bodily responses.

However, this paper is only an introduction of what similarity matrices can bring to the understanding of player experience. Indeed, numerous improvements can be made in the future. For instance, the video stream similarity can also be assessed in conjunction with the sound stream, thus reducing the amount of noise inside the matrices when the sound is looped. Moreover, it would be highly valuable to be able to automatically detect the diagonals through the use of image processing algorithms based on the similarity matrix image. For instance, an algorithm that can automatically count the number of distinct diagonals would also automatically classify beginner from skilled players in the *Super Hexagon* example presented above.

ACKNOWLEDGMENTS

- The Royal Society of New Zealand, Marsden Grant for funding 'Videogame Classification: Assessing the experience of play' (10-UOW-024).
- The University of Waikato, International Scholarship Grant, for funding 'Quantitative Data and Gameplay Segmentation Applied

to the assessment of Players’ Experience with Digital Games’.

- The LaBRI, for inviting and hosting Raphaël Marczak during the redaction of the current paper.
- BIBLIOGRAPHY

Aarseth, E. (2001). Computer Game Studies, Year One. *Game Studies, 1*(1).

Calleja, G. (2011). *In-Game: From Immersion to Incorporation*. Cambridge, MA: The MIT Press.

Choi, F. Y. Y. (2000). Advances in Domain Independent Linear Text Segmentation. In *NAACL 2000 Proceedings of the 1sr North American chapter of the Association for Computational Linguistics Conference* (pp. 26–33).

Cooper, M. L., & Foote, J. T. (2001). Scene boundary detection via video self-similarity analysis (Vol. 3, pp. 378–381). Presented at the Image Processing, 2001. Proceedings. 2001 International Conference on.

Csikszentmihaly, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper Perennial.

Drachen, A., Seif El-Nasr, M., & Canossa, A. (2013). Game Analytics – The Basics. In *Game Analytics*. Retrieved from http://andersdrachen.files.wordpress.com/2012/08/thebasics.pdf

Drachen, A., Thurau, C., Togelius, J., Yannakakis, G. N., & Bauckhage, C. (2013). Game Data Mining. In *Game Analytics*. Retrieved from http://andersdrachen.files.wordpress.com/2012/08/gamedatamining_intro.pdf

Hanna, P., Robine, M., & Ferraro, P. (2008). Visualisation of Musical Structure by Applying Improved Editing Algorithms. In *Proceedings of the 2008 International Computer Music Conference*.

Kim, J. H., Gunn, D. V., Schuh, E., Philips, B., Pagulayan, R. J., & Wixon, D. (2008). Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems (pp. 443–452). Presented at the SIGCHI Conference on Human Factors in Computing Systems, New York, USA. http://doi.org/10.1145/1357054.1357126

Laurel, B. (1993). *Computers as Theatre*. Addison-Wesley Professional.

Marczak, R., & Schott, G. (2015). Audio Visual Analysis of Player Experience: Feedback-Based Gameplay Metrics. In P. Lankoski & S. Björk (Eds.), *Game Research Methods: an overview* (pp. 207–230).

Marczak, R., Schott, G., Hanna, P., & Rouas, J.-L. (2013). Feedback-Based Gameplay Metrics. In *Proceedings of FDG 2013*.

Mirza-Babaei, P., Nacke, L. E., Gregory, J., Collins, N., & Fitzpatrick, G. (2013). How Does It Play Better? Exploring User Testing and Biometric Storyboards in Games User Research. In *Proceedings of CHI 2013*. Paris, France.

Nacke, L. E., & Lindley, C. A. (2008). Flow and Immersion in First-Person Shooters: Measuring the Player's Gameplay Experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 81–88).

Schott, G., Marczak, R., Mäyrä, F., & Vught, J. V. (2013). DeFragging Regulation: From putative effects to "researched" accounts of player experience. In *Proceedings of DiGRA 2013*.

Schott, G., Marczak, R., & Neshausen, L. (2014). <Exploring> the <Cause> of Game (Derived) <Arousal>: What Biometric Accounts of Player Experience Revealed. In *Proceedings of DiGRA 2014*.

Serra, J., Gomez, E., Herrera, P., & Serra, X. (2008). Chroma Binary Similarity and Local Alignment Applied to Cover Song Identification. *IEEE Trans. on Audio, Speech and Language Processing*, 1138–1152.

Yannakakis, G. N., Spronck, P., Loiacono, D., & André, E. (2013). Player Modeling. In *Artificial and Computational Intelligence in Games* (Vol. 6, pp. 45–59).

Zagal, J. P., Fernandez-Vara, C., & Mateas, M. (2008). Rounds, Levels, and Waves: The Early Evolution of Gameplay Segmentation. *Games and Culture*, *3*(2), 175–198.

Game references

Rockstar Studios (2012), *Max Payne 3*. [PC], Rockstar Games.
EA Digital Illusions CE (2011), *Battlefield 3*. [PC], Electronic Arts.
Terry Cavanagh (2012), *Super Hexagon*. [PC], Independent.