An Analysis of Repetition in Video Game Music

Samuel J. Hunt

Creative Technologies Laboratory
UWE Bristol
Samuel.hunt@uwe.ac.uk

Abstract. Video game music, unlike other forms and genres of music, is comparatively young in its development, and undoubtedly heavily influenced by the hardware it was originally played on. This study observes whether the amount of repetition in video game music has been affected by limitations in hardware and physical storage size. Musical structure often derives from repetition and is one of many crucial musical elements for defining structure. We hypothesise that early videos games will contain above average repetition due to hardware limitations. Early video game soundtracks were minimally encoded, utilising hardware based chip-tune synthesis, whereas modern games can support *Hollywood style* soundtracks. This research analyses 21,391 pieces of video game music across a range of various video game consoles. In conclusion our original hypothesis was mostly disproven, and in fact the repetition structure discovered across different platforms and generations remained more-or-less consistent. There were however improvements in song length and the number of instrument tracks within a song. Overall this paper presents an initial informal analysis of repetition in video game music.

Keywords: Video game music, music analysis, music structure

1 Introduction

Despite its relative immaturity, video game hardware and audio has rapidly progressed over the last 60 years. Early hardware had primitive, if any sound, with some systems only able to emit a single beep (Chang, Kim, & Kim, 2007). Modern systems are employing dynamic and evolving soundtracks, immersive audio, and complex surround sound mixing (Hutchings & McCormack, 2019). Music analysis is a vast subject and there are many feasible research objectives for analysing video game music. However, one characteristic that perhaps captures musical structure holistically is repetition. Rahn (1993) states most musical structure derives from repetition - without repetition there is no structure and the music itself will be little more than random. The main objective of this research is to observe how repetition within video game music has progressed over the various generations of hardware. Other musical elements such as key, harmony and rhythm could also have been analysed and compared, however, in line with other work in the field (see section 2) and to keep things within the scope of this paper we focus on a singular element (repetition).

2 Literature Review and Video Game Hardware

Musical structure happens at different hierarchal levels. Lerdahl and Jackendoff (1996) consider music to be in one of a four-level hierarchy; motifs, phrases, periods, and larger sections. Within these levels repetition occurs in different ways, for example a single bar of music may be repeated several times in succession, whereas an entire section may be repeated more than once, such as a chorus section. Classical musical forms, such as ternary form (ABA), emphasises a full repeat of the first A section, after the B section. Humans are excellent pattern recognisers and even those without formal training are competent enough at recognising repeating musical sequences and motifs (Chai, 2005).

Discovering repeating patterns within music automatically using computer programs is not a novel problem, and researchers have gone about solving this in different ways. This research focuses on analysing digitally sequenced music (i.e. MIDI). Some notable research has attempted this on recorded audio (Lu, Wang, & Zhang, 2004), and with the recent developments in machine learning, interest in this has increased (Jhamtani & Berg-Kirkpatrick, 2019). Hsu, Liu, and Chen (2001) present two methods for discovering repeating patterns in music, one using a correlative matrix and the other using a tree type data structure. Their study focuses on finding patterns of music at the smallest structural level. Other methods include using a dictionary-based compression algorithm (Shih, Narayanan, & Kuo, 2001). Both of these methods are a little complex for the purpose of this study. Another simpler method used by the authors (Hunt, Mitchell, & Nash, 2019) in previous work for analysing repetition, is to simply split a piece of music into smaller segments based on bar lines and then compare these against one another. Although some nuanced details are lost by this process it permits a simple automated approach that can cover a large dataset. We employ the same methods in this research and discussed further in section 3.

Collins et al. (2008) has thoroughly explored the relationship between technical systems and their musical use in video game music. An example of her work that had similar research objectives to the work in this paper, is exploring the link between hardware limitations and the use of loops in 8-bit video game hardware (Collins, 2007). She suggests that there is some correlation between the provision of music (use of repetition) and the storage capacity of the games cartridges, but in conclusion notes *"that rather than being the consequence of the limited memory available on the systems, loops were, at least in part, an aesthetic that grew as the games became more popular and more complex"*. Noting that creative composers have invented ways in which to overcome or even to aestheticize those limitations. In summary the research was limited by focusing solely on 8-bit generation hardware. With influence of Colins' work, we conduct similar research but for a wider range of hardware.

A full technical description of the vast history of video game audio hardware is outside of the scope of this research, however, we can broadly categorise each piece of hardware into one of 3 groups using work by (Chang et al., 2007). These are; chip tune and 8-bit (Group A), synthesis and sampling (Group B), and pre-recorded and streaming music (Group C).

Chip tune and 8-bit systems (A) synthesised music in real-time using primitive on-board hardware based sound generators, similar to analogue synthesisers. Synthesis and sampling systems (B) moved away from discrete waveform generators into more general purpose synthesisers. Wavetable and FM synthesis were heavily utilised given their small memory footprint. Pre-recorded and streaming based systems (C) differ in that music could be composed and recorded externally and then played back in real time. Fritsch (2013) notes that in early video games the role of the composer often fell on the programmers themselves given that forms of storage and playback were not well formalised. Music, for the first two categories, required extensive programming for in game playback. Whereas with pre-recorded and streaming based systems the role of the composer and the sound programmer were fully decoupled.

A game's storage medium may have affected the music and audio utilised within the game. Early hardware used cartridge based media with limited storage capacity, whereas disk based media provided ample storage. For example, a Nintendo 64 cartridge provided a maximum of 64 mb whereas its closest competitor, Sony's Playstation, had 660 mb with its CD based storage (Sakamura, 1999).

There are 8 defined generations of video game consoles (Kemerer, Dunn, & Janansefat, 2017), with music from generations 3-6 used in this research. The first two generations contained minimal, if any, music as the hardware was primitively simple. The most recent generations have seen little advancements in terms of audio hardware implementation (Chang et al., 2007), but have emphasised improvements in immersive audio (surround sound formats) and dynamic generative soundscapes (Hutchings & McCormack, 2019) which create challenges for obtaining and analysing such music. Furthermore, the dataset (discussed shortly) used in this study contained few examples of music for generations 7-8.

The video games consoles whose music was analysed for this research are listed in table 1 (see appendix). The Sega Megadrive and Nintendo 64 both supported disk based media through hardware add-ons (Sega CD and N64 DD) however games using those add-ons were removed from the datasets to eliminate additional variables. For this study we only focused on home video game consoles, neither handhelds or PCs were considered.

3 Methodology and Software

To analyse the music needed for the project, the interactive generative music environment (IGME) was used to compute the repetition analysis (Hunt, Mitchell, & Nash, 2018). The same authors (Hunt et al., 2019) have completed similar work using IGME but focused on algorithmic representation of music. The dataset for the project was obtained from the online VGMusic database[1] which contains a large database (over 30,000) of MIDI files categorised by console.

Individual songs are imported into IGME as MIDI files and then analysed. IGME only supports the analysis of files in a 4/4 time signature. The music is

[1] https://www.vgmusic.com/

segmented into smaller clips for analysis, for the purpose of this study (and as a limitation of the software) the smallest clip size is 1 bar. If notes are tied across a bar line then they are simply segmented at the next possible (non-note crossing) bar line, leaving a phrase larger than a single measure (we subsequently found this to only account for a small minority of pieces). Therefore, a clip for analysis is defined as 1 or more measures of music. A clip size average score of 1, implies that all the music in the piece could be split uniformly into 1 bar measures. Splitting into smaller units (i.e. less than a single measure) requires more complex analysis on both segmenting and subsequent analysis. Some notable work (Temperley, 2007) has explored options for this (albeit in a non-automated way), as well as the more general purpose repetition analysis methods discussed previously. Note length and onsets are quantised to the nearest 1/32 note. Other details are lost by the process, such as dynamic markings. Without making these modifications the complexity of doing this analysis would be implausible.

The repetition analysis attempts to find and group measures of music that have identical content. So that within each bar, every note has the same note length, onset time and note number; therefore, comparing every clip against every other clip on the time line. In the case of parallel duplication (two tracks doubling the same music), this increases the overall amount of repetition found as these two clips would be identical. Should two tracks be completely identical in their entirety, one is removed before starting the analysis processes. Although our definition of repetition excludes measures that are almost identical (for example a variation on a theme), it is difficult to analyse and categorise such measures, this could be an area for future research.

4 Results

Year	Generation	System	Sample Size	Average Repetition Score	*Repetition Standard Deviation*	Average Length in bars	Average Number Of Tracks	Average Unit Size
1983		NES	4085	60.27	28.77	48.44	6.21	1.24
1986	3rd	Sega Master System	1582	55.88	31.75	38.13	4.45	1.16
1988		Sega Megadrive	2348	59.48	31.56	59.93	8.74	1.4
1990		NeoGeo	230	66.75	28.26	65.85	8.02	1.23
1991	4th	Super Nintendo	6518	62.67	27.09	63.93	8.58	1.43
1994		PlayStation	2578	63.75	24.19	75.45	9.58	1.4
1994		Sega Saturn	430	63.7	26.96	63.68	9.58	1.29
1996	5th	Nintendo 64	1755	68.23	23.78	72.68	8.65	1.26
1998		Sega Dremcast	287	64.64	25.52	82.72	10.26	1.5
2000		PlayStation 2	793	60.85	25.13	75.2	9.86	1.44
2001	6th	Nintendo Gamecube	785	64.46	25.94	67.18	9.42	1.32

Fig. 1. Repetition analysis results, grouped by system.

Our original hypothesis stated older game music would contain more repetition than modern equivalents, given the hardware limitations discussed previously. The research analysed 11 groups of MIDI files (1 per platform), using the method described previously, with the results in figure 1. The repetition score is given as a mean for all music analysed per system as well as a standard deviation of the mean. The average song length in bars, the average number of MIDI tracks (after removing duplicate tracks) and the average clip size are also given.

The overall results, show that the repetition score varies little between systems. Taking an overall mean average, we find a repetition value of 62.79 and a standard deviation of 3.49. This would suggest that on average 62.79% of the musical clips in the music are a repeat of a previous clip.

One value that changes significantly over time is the average song length as well as the number of permissible tracks. It would appear that advancements in a game's storage size allowed for longer song lengths, and advancements in audio hardware increased polyphony. One interesting observation is that when comparing the Nintendo 64 and Playstation, they have similar values for repetition and song length. We hypothesised that Playstation music would have longer song lengths given how its storage capacity was 10x that of the Nintendo. Although the Playstation has marginally longer songs and higher track counts, it appears that these storage limitations had little effect on a songs structure.

5 Conclusion

In conclusion we found little variance in musical repetition between video game consoles and their subsequent hardware. Although older hardware itself did not specifically contribute to repetition, it limited the amount of songs that a individual game could contain. Such an observation would however require future study. Collins (2007) whose research was similar to that discussed here, observed that despite some quite rigid technological constraints provided only *'loose pressures'* on the development of audio for 8-bit video game hardware.

The style of music composed for a video game is hugely dependent on the type of game and is something overlooked in this study. Future work may focus on the differences between such genres although this would require significant organisation of the dataset. A limitation of using MIDI files is that two independent music lines or voices might be encoded on a single MIDI track, this may therefore have an adverse effect on the results. Is it however considerably difficult to separate voices once they are encoded without tediously analysing and manually separating the content. Also worth noting, is the smaller timbre characteristics of the audio when played on its respective physical hardware undoubtedly creates small nuances in the structure of the music. These are not sufficiently captured by MIDI encoding. Although repetition is an important musical characteristic it does not describe all the complexities of musical structure. Future work may also include additional analysis processes, such as: rhythm, use of key, instrumentation, polyphony, harmony and others.

References

Chai, W. (2005). *Automated analysis of musical structure* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Chang, K., Kim, G., & Kim, T. (2007). Video game console audio: evolution and future trends. In *Computer graphics, imaging and visualisation (cgiv 2007)* (pp. 97–102). Bangkok, Thailand: IEEE.

Collins, K. (2007). In the loop: Creativity and constraint in 8-bit video game audio. *Twentieth-century music*, *4*(2), 209.

Collins, K., et al. (2008). *Game sound: an introduction to the history, theory, and practice of video game music and sound design*. Mit Press.

Fritsch, M. (2013). History of video game music. In *Music and game* (pp. 11–40). Newyork: Springer.

Hsu, J.-L., Liu, C.-C., & Chen, A. L. (2001). Discovering nontrivial repeating patterns in music data. *IEEE Transactions on multimedia*, *3*(3), 311–325.

Hunt, S., Mitchell, T., & Nash, C. (2018). A cognitive dimensions approach for the design of an interactive generative score editor. In *Fourth international conference on technologies for music notation and representation, montréal, canada.* (pp. 1–8). Montreal, Canada: TENOR.

Hunt, S., Mitchell, T., & Nash, C. (2019). Automating algorithmic representations of musical structure using igme: The interactive generative music environment. In *Innovation in music 2019* (Vol. 1, pp. 1–18). University of West London: Innovation In Music.

Hutchings, P., & McCormack, J. (2019). Adaptive music composition for games. *IEEE Transactions on Games*, 1 –1.

Jhamtani, H., & Berg-Kirkpatrick, T. (2019). Modeling self-repetition in music generation using generative adversarial networks. In *International conference on machine learning.* Long Beach, California: International Conference on Machine Learning.

Kemerer, C. F., Dunn, B. K., & Janansefat, S. (2017). Winners-take-some dynamics in digital platform markets: A reexamination of the video game console wars. *University of Pittsburgh, PA*.

Lerdahl, F., & Jackendoff, R. S. (1996). *A generative theory of tonal music*. Cambridge, Massachusetts: MIT press.

Lu, L., Wang, M., & Zhang, H.-J. (2004). Repeating pattern discovery and structure analysis from acoustic music data. In *Proceedings of the 6th acm sigmm international workshop on multimedia information retrieval* (pp. 275–282). New York NY USA: ACM SIGMM.

Rahn, J. (1993). Repetition. *Contemporary Music Review*, *7*(2), 49–57.

Sakamura, K. (1999). Entertainment and edutainment. *IEEE Micro*, *19*(6), 15–19.

Shih, H.-H., Narayanan, S. S., & Kuo, C.-C. (2001). Automatic main melody extraction from midi files with a modified lempel-ziv algorithm. In *Proceedings of 2001 international symposium on intelligent multimedia, video and speech processing. isimp 2001 (ieee cat. no. 01ex489)* (pp. 9–12). Hong Kong, China: IEEE.

Temperley, D. (2007). *Music and probability.* Cambridge, Massachusetts: Mit Press.

Appendix

Table 1. A table summarising the various video game consoles used in this study.

Generation	**Console**	**Audio group Type**	**Media Type**	**Year of release**
3rd	Sega Master System	B	Cartridge	1986
	Nintendo entertainment system	A	Cartridge	1983
4th	Super Nintendo	B	Cartridge	1991
	Sega Megadrive	B	Cartridge	1988
	SNK NEOGeo	B	Cartridge	1990
5th	Playstation	C	Disk	1994
	Sega Saturn	C	Disk	1994
	Nintendo 64	C	Cartridge	1996
6th	Playstation 2	C	Disk	2000
	Sega Dreamcast	C	Disk	1998
	Nintendo Gamecube	C	Disk	2001