Thesis for a Ph.D. Degree in Engineering

A Visual Analysis Methodology for Music Compositional Processes with Sound Resynthesis

February 2019

Graduate School of Science and Technology
Keio University

Yuriko Takakura

Abstract

A Visual Analysis Methodology for Music Compositional Processes with Sound Resynthesis

With the advent of computer technologies, composers have begun to use computers for their compositions. Among the many compositional techniques, sound analysis and synthesis is a fundamental process to create new timbres. In this thesis, the process of sound analysis and synthesis is referred to as "resynthesis." Computer technologies have also changed the actions of composers during their compositions. Because they may forgo writing sketches and scores on paper, some studies point out that it is difficult to analyze the compositional process of music created with computers.

The aim of this thesis is to propose a visual analysis methodology for analyzing compositional processes made with sound resynthesis by complementally utilizing two approaches: sketch study and a visual analysis. First, a sketch study to clarify the compositional process of two symphonies by Toshiro Mayuzumi (1929–1997) is presented. Mayuzumi is widely known for his compositions using the sound resynthesis of temple bells. This method can be regarded as an ancestor to that of the spectral school. In this case study, the relation between the two symphonies is specifically clarified, and the possibilities and limitations of sketch studies are generally discussed. Second, the difficulties in the research of a compositional process to trace the processes of computer usage in music composition are pointed out, and a system, *Spectrail*, that can visually analyze the compositional process made with sound resynthesis is proposed. This system allows the users to visually analyze the history of a composition with a pixel-oriented spatial substrate by using data taken from AudioSculpt, which is a commonly used software for sound resynthesis. The users who are primarily targeted by this system are musicologists and composers. The system is evaluated empirically with actual pieces and with interviews by composers from the Institut de Recherche et Coordination Acoustique/Musique (IRCAM).

Throughout the research, it has been shown that the present visual analysis approach is effective in clarifying the compositional process with sound resynthesis in a comprehensible way. The visual analysis provides precise information on the chronological order of the compositional process together with an overview of sound changes. In addition, sketch studies provide additional information on the concepts of the composition and/or the structure of the piece.

Acknowledgments

First, I would like to express a special appreciation to my supervisor Prof. Issei Fujishiro for his continuous support of my Ph.D. study. When I was a student of the Graduate School of Letters, I hoped to enter his laboratory during my second master's course, and then I changed my major to information visualization. He kindly welcomed me and gave me precious advice, especially on how computer science and musicology can collaborate together. After entering his laboratory, I noticed that it was difficult to understand the basic concepts of technical words in computer science. He kindly explained these again and again. Without his help, this thesis would not have been possible. He also advised my future career on several occasions and suggested me various career paths I could take.

I would like to express my deepest appreciation to Prof. Takayuki Ito, Prof. Yuriko Kojima, and Prof. Hiroaki Saito, who as the committee members gave me insightful comments and helped me widen my research from various perspectives. At the joint meetings, Prof. Ito also gave me much advice on how to visualize the target. Prof. Kojima taught me much about computer music. She and her students kindly supported me when I conducted an empirical experiment during the master course of the Graduate School of Science and Technology. Prof. Saito gave me advice when entering the second master course.

I owe my deepest gratitude to Dr. Nicolas Donin for hosting me at the Equipe Analyse des Pratiques Musicales in the Institut de Recherche et Coordination Acoustique/Musique (IRCAM), Paris, and it was a great honor for me to study under his supervision. I would also like to express my thanks to Dr. Diemo Schwarz, who is my co-supervisor at IRCAM, Dr. Charles

Picasso, who is a developer of AudioSculpt, Dr. Mikhail Malt, who is an engineer and composer, and the members of the Equipe Analyse des Pratiques Musicales.

A special thanks also go out to the four composers who kindly collaborated with me during my research, Mr. Luciano Leite Barbosa, Mr. Alessandro Ratoci, Mr. Frédéric Le Bel, and Mr. Maurilio Cacciatore. They composed new pieces for my research and answered many questions about their compositions.

I would also like to express my thanks to the family of Toshiro Mayuzumi and staff of the TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University for their kind assistance in finding Mayuzumi's autograph sketches, "Campanology Documents."

My grateful thanks also go out to my supervisors in the Graduate School of Letters, Prof. Hisao Nishikawa and Prof. Akira Hirano, for giving me valuable opportunities and advice both on research and on my career.

I would particularly like to thank every member of Fujishiro Laboratory, Nishikawa Seminar for encouraging me and helping me in various ways. Especially, I am grateful for considerable efforts of Mr. Masanori Nakayama, Dr. Ken Nagao, Mr. Yasunari Ikeda, and Mr. Malik Olivier Boussejra. They helped me at all times when developing *Spectrail* and writing papers.

I would like to offer my special thanks to the RAs, mentors, and professors in the Program for Leading Graduate School for "Science for Development of Super Mature Society" for their invaluable and warm support.

Finally, I would like to show my greatest appreciation to my parents, Mr. Toru Takakura and Ms. Masako Takakura for their unconditional support and encouragement. My thanks also go to my sister, Ms. Mayuko Takakura, and my grandparents, Mr. Yoshio Takakura, Ms. Kayoko Takakura, Mr. Takayuki Ishikawa, and Ms. Ayako Ishikawa, for their continuous support.

Table of Contents

Abstract i

Acknowledgements ii

1 Introduction 1

1.1 Development of Computer Music . . . 3

1.1.1 Exploring new timbres . . . 3

1.1.2 Process of sound resynthesis . . . 4

1.2 Compositional Process Study in Musicology . . . 7

1.2.1 Research areas of musicology . . . 7

1.2.2 Study of the compositional process . . . 7

1.2.3 Difficulties in analyzing computer music . . . 8

1.3 Aim and Contributions of Thesis . . . 8

1.4 Organization of Thesis . . . 9

2 Related Work 11

2.1 Putting the Research into the Study Perspective . . . 13

2.2 Musicological Study for Computer Music . . . 14

2.2.1 Sketch studies . . . 14

2.2.2 Analysis with graphical interfaces . . . 14

2.2.3 Analysis with sound descriptors . . . 15

2.3 Music Visualization . . . 17

2.4 Provenance Visualization . . . 19

3 A Sketch Study: The Compositional Process of Two Symphonies by Toshiro Mayuzumi **21**

3.1 Introduction . . . 23

3.2 Outline of Campanology Documents . . . 25

3.3 *Nirvana Symphony* . . . 29

3.3.1 Main chords of *Nirvana Symphony* . . . 29

3.3.2 Analysis of Campanology Documents . . . 31

3.3.3 Compositional process of the *Nirvana Symphony* . . . 40

3.4 *Mandala Symphony* . . . 41

3.4.1 Main chords of *Mandala Symphony* . . . 41

3.4.2 Analysis of Campanology Documents . . . 43

3.4.3 Compositional process of the *Mandala Symphony* . . . 50

3.5 Comparison between *Nirvana Symphony* and *Mandala Symphony* . . . 51

3.6 Possibilities and Limitations of Sketch Studies . . . 52

4 A Visual Analysis Approach: *Spectrail* **54**

4.1 System Outline . . . 56

4.1.1 Requirement analysis . . . 57

4.1.2 Object software: AudioSculpt . . . 58

4.1.3 Input data . . . 59

4.1.4 Spatial substrate and users of *Spectrail* . . . 61

4.2 Functions for Analysis . . . 61

4.3 Implementation . . . 67

4.4 Limitations of *Spectrail* . . . 72

5 Empirical Evaluation of *Spectrail* **73**

5.1 Outline of Empirical Experiment . . . 75

5.2 Subjects . . . 76

5.2.1 Composer A 76
5.2.2 Composer B 78
5.2.3 Composer C 78
5.2.4 Composer D 78
5.3 Visual Analysis 82
5.3.1 Viewpoints of analysis 82
5.3.2 Compositional process 82
5.3.3 Repeated pattern 84
5.3.4 Length of treatments 85
5.3.5 Sound descriptor 86
5.3.6 Summary of the analysis 88
5.4 Interview with the Composers 91
5.4.1 Original sound file 91
5.4.2 Evaluation of the visual analysis 91
5.4.3 Sketches 92
5.4.4 Possibility to use *Spectrail* for compositions 92
5.4.5 Summary and discussions 94

6 Conclusion **95**
6.1 Conclusion 97
6.2 Future Research 99
6.2.1 Taking into account the actual composition 99
6.2.2 Provenance of analysis by *Spectrail* 99
6.2.3 Adding new data for detailed analysis 100
6.2.4 Extending to multimodal analysis 100

Publications **101**

References **102**

List of Figures

1.1 Composition with sound resynthesis. The orange rectangles show the process with computers and the blue the process with acoustics. 6

1.2 Interface of *Spectrail* 10

2.1 The research is at an intersection of these three fields: musicology, music visualization, and provenance visualization 13

3.1 Arrangement of three orchestra groups 23

3.2 Cover of the Campanology Documents (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 28

3.3 Main chords of the *Nirvana Symphony* 29

3.4 Document No. 3-1 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 33

3.5 Document No. 5 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 35

3.6 The right side of the chord table in Document No. 8 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 37

3.7 Document No. 4 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 39

3.8 The tone series in "figure one." The main chords are organized with the fundamental tone series 42

3.9 Main chords of the *Mandala Symphony* 42

3.10 Eight tone series in Document No. 7 44

3.11 The left side of Document No. 7 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 48
3.12 The right side of Document No. 7 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University) 49
3.13 Utilizing data in the software. It is important to visually analyze the data in the software to see the compositional process. The data will be one of the main resources to study the compositional process in music 53
4.1 Data flow in *Spectrail*. As its input, *Spectrail* takes an SDIF file and audio files from AudioSculpt. 56
4.2 A screenshot of AudioSculpt interface with treatments 58
4.3 Internal format of an SDIF file 59
4.4 Visualizing the time intervals between treatments as the distances between treatment blocks. The history of the composition is represented from the bottom-up . 62
4.5 Visualizing the difference in values for the sound descriptors 63
4.6 Listening to the sound in each treatment 65
4.7 Details of the SDIF data 65
4.8 Overview and zoom 68
4.9 Filter mode 69
4.10 Extracts 70
4.11 Change in the MFCC 71
5.1 Evaluation method 75
5.2 Compositional process of **Composer A** 77
5.3 Compositional process of **Composer B** 79
5.4 Compositional process of **Composer C** 80
5.5 Compositional process of **Composer D** 81
5.6 Compositional process of **Composers A** and **B** 83
5.7 Compositional process of **Composers B** and **D** 84

5.8 Treatments used by **Composers B** and **C** 85

5.9 Change of the sound descriptors in the compositional process of **Composer A** . 87

5.10 Change of the sound descriptors in the compositional process of **Composer B**. . 87

5.11 Change of the sound descriptors in the compositional process of **Composer D** . 87

5.12 Pitch and loudness are changed by `Transposition` and `TimeStretch` 89

5.13 The differences of the sound descriptors are small, and it corresponds to where the composer uses tiny short treatments 89

5.14 The descriptors are changed drastically where the repeated pattern "`Transposition` + `Gain`" is used 90

5.15 Sketch by **Composer A** 93

List of Tables

3.1 Data of Campanology Documents 26

3.2 The number strings in Document No. 7 45

4.1 List of treatments and their abbreviations 60

5.1 Basic information about each composition 76

Chapter 1

Introduction

With the advent of computer technologies, composers have begun to use computers for their compositions. Among the many new compositional techniques, sound analysis and synthesis is a fundamental process for creating new timbres, and in this thesis, the process of sound analysis and synthesis is referred to as “resynthesis.” Indeed, computer technologies have changed composers’ actions during their compositions as well. Because composers may forgo writing sketches and scores on paper, some studies point out that it is difficult to analyze the compositional process of music created with computers. The aim of this thesis is to propose a methodology for analyzing the compositional process by using a visual analysis and sketches. In this chapter, the background and aim of the research are presented.

1.1 Development of Computer Music

1.1.1 Exploring new timbres

Since the twentieth century, composers have begun to explore new timbres in their compositions. Atonal music is one of the new trends, that does not have tonality and harmonies. Twelve-tone music is a representative example of atonal music. In this music, the hierarchy of twelve tones brought about by tonality collapses, and each twelve-tone in an octave is treated coequally [Griffiths, 2001]. This trend has also led to total serialism, which systematically organizes every element in the music.

Composers have also created new instruments and styles of performances. For example, the prepared piano, which changes the timbre of the piano using several kinds of non-musical materials such as metals and woods, was invented by John Cage in 1940. Cluster, a group of adjacent notes sounding simultaneously, was also created as a new style of performance [Cluster, 2001]. These trials have led to creating a new notation system, such as graphic notation, because these new timbres cannot be represented in the traditional musical score.

Because of the interest in new timbres, composers have begun to use electronic sounds in their musical pieces [Stauffer, 2001]. After computers became to be more commonly used, composers also began to use sound technologies [Marston, 2001]. The paper by Moorer [Moorer, 1977] summarizes the contemporary sound synthesis technologies used in a composition, and he identifies two reasons why the sound synthesis with a computer becomes generalized: one is that composers expect to extend their means of expression, and the other is that they can synthesize sounds precisely. He classifies the techniques of sound synthesis into three categories. The first one is a sound synthesis method that directly synthesizes the sounds in terms of waveforms. The second is a so-called “analysis-based synthesis” that uses the digitization and analysis of existing natural tones. The third kind of synthesis is “musique concrète,” which Pierre Schaeffer (1910–1995) created in 1948.

Computers can generate a variety of sounds compared with these traditional ones: however, it is difficult to write the sound with an ambiguous pitch because traditional Western music is based on a discrete pitch system. However, not all compositions use computers, and composers sometimes use paper sketches to create ideas about their compositions.

1.1.2 Process of sound resynthesis

Among the new types of compositional techniques, sound analysis and synthesis is one of the fundamental processes of composition. In the current thesis, a process of creating and synthesizing sounds with sound analysis is referred to as "resynthesis." This term does not indicate a specific technique but rather shows the process of sound synthesis through analyzing the characteristics of sounds.

In musique concrète, composers record some "concrete" sounds and then reorganize them, as shown in Figure 1.1(a). "Musique concrète" is a kind of electroacoustic music that uses natural sounds, not electronically generated tones, as the original sounds for the synthesis. For example, recordings of machinery, running water, and/or musical instruments are transformed and intermixed to form a composition [Latham, 2011].

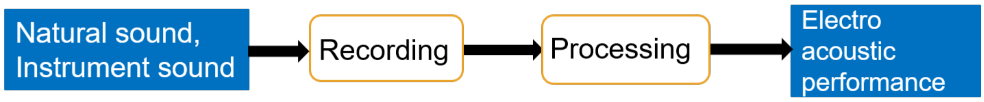
These methods were mainly studied and developed at the research centers for electronic music in Köln, Paris, New York, and Milan. These kinds of new trends also have an influence on other countries. In Japan, the NHK electronic music studio was established in the 1940s. In 1953, *X・Y・Z for musique concrète* by Toshiro Mayuzumi premiered and is known as the first musique concrète piece in Japan [Loubet et al., 1997]. The work consists of *X* with metallic sounds, *Y* mainly with voices, and *Z* with instrumental sounds.

Live electronics, which is another example of the music using sound resynthesis, has appeared. Live electronics is a genre of electronic music that is associated with acoustic instruments, as shown in Figure 1.1(b). The acoustic instrumental sounds performed in real-time are recorded, and live performers create the electronic sounds by means of synthesizers, ring modulators, and other electronic devices [Montague, 1991].

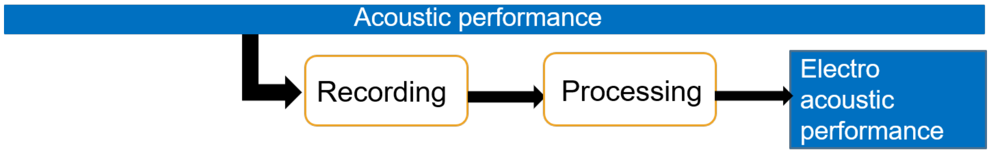
Yet another example is spectral music, whose roots lie in the Institut de Recherche et Coordination Acoustique/Musique (IRCAM). Starting in the 1970s, spectral music refers to music composed mainly in Europe, where the acoustic properties of sounds are used as the basis of the compositional materials [Rose, 1996] [Eldénius, 1998] [Fineberg, 2000a] [Fineberg, 2000b] [Grisey and Fineberg, 2000] [Anderson, 2001], as shown in Figure 1.1(c).

Spectral composers use various sound materials that encompass any sounds, from musical sounds to unpitched ones, in their pieces. Examples of the sound materials used by these composers include abstract harmonic series, natural sounds, and mathematical models of sound materials. Tristan Murail (1947–), who is known as a co-founder of the spectral school [Smith and Murail, 2000], analyzed bell sounds and orchestrated his representative piece *Gondowana* (1980) based on the harmonics of these sounds. Gérard Grisey (1946–1998) used the result of analyzing a trombone sound within his representative musical piece, *Les Espaces Acoustiques* (1974–1985).

Klingbeil [Klingbeil, 2008] summarizes composers' methods when using spectral analysis and synthesis. He explains numerous kinds of methods for spectral modeling, analysis, and synthesis and also refers to graphical user interfaces as a way to work with these methods.



(a) A type of compositional process in musique concrète



(b) A type of compositional process in live electronics

(c) A type of compositional process in spectral music

Figure 1.1: Composition with sound resynthesis. The orange rectangles show the process with computers and the blue the process with acoustics.

1.2 Compositional Process Study in Musicology

1.2.1 Research areas of musicology

Musicology includes many research areas, such as music history, acoustics, psychology, music education, and so forth. A musicologist, Guid Adler, founded a journal titled "Vierteljahrsschrift für Musikwissenschaft" in 1885 and presented his paper "Umfang, Methode und Ziel der Musikwissenschaft" as the first paper in the journal with the intention to organize the research areas and methods of musicology [Adler, 1885]. He classified the wide research areas of musicology into historical and systematic areas.

The main methods in musicology are analyzing scores, sketches, and composers' remarks. The primary resources, such as autographs, articles, letters, diaries, and programs for concerts, are important materials for musicology. These resources are used to study compositional processes and/or to examine how the music was enjoyed by people.

1.2.2 Study of the compositional process

One of the areas of musicology is compositional process research, which aims to understand the process of composing musical pieces. The compositional process includes the processes of creating ideas for the composition, organizing the music, and writing down scores.

When studying compositional processes, musicologists analyze sketches, composers' remarks, and the musical scores of works. In traditional compositions, sketches are basically written on paper [Marston, 2001], and analysts acquire the resources from physical archives.

Radocy et al. [Radocy and Boyle, 1997] summarize the types of compositional process studies as "analyzing sketches, examining the composer's discourses, observing composers." These methods are classified into resource studies and observations. Donin et al. [Donin and Traube, 2016] also mention the areas of studying the compositional process. They point out that the targets include the art of record productions, sound engineering theories and practices, improvisations, and so forth. There are studies in the arts education field that help to clarify the compositional process when using computers [Rose, 1996] [Crawford and Gibson, 2017].

1.2.3 Difficulties in analyzing computer music

As described above, since the twentieth century, composers have begun to use computers. However, this situation contains many issues when trying to study these pieces. Some studies point out that it is difficult to analyze the compositional process of music created with computers. According to Couprie [Couprie, 2016], one of the difficulties in analyzing these musical pieces is that they do not have any visual supports; therefore, listeners need to understand complex relations among parts, moments, and units without visual support, and this can be hard to discover through simple listening, which is motivation in creating a visual interface for analyzing these musical pieces. Wiggins [Wiggins, 2007] argues that not all of the compositional processes include notations, especially in the case where the composer uses graphical user interfaces for the composition. He also mentions that the compositional process not only includes notes, but also some other elements, such as experience beyond the score.

1.3 Aim and Contributions of Thesis

In this thesis, compositional process studies are introduced with two approaches—sketch analysis and visual analysis.

First, a sketch study used to explore the compositional process of two symphonies by Toshiro Mayuzumi is discussed. Mayuzumi is widely known for his composition with the sound resynthesis of temple bells and the method can be seen as an ancestor of the spectral school. In this study, the relation between these symphonies is clarified, and the possibilities and limitations of sketch studies are discussed. Second, a system—*Spectrail*—to visually clarify the compositional process of sound resynthesis is proposed. This system allows users to visually analyze the history of a composition with a pixel-oriented spatial substrate (see Figure 1.2) using data taken from AudioSculpt, which is a representative software for sound analysis and synthesis; the users can analyze the methods and sound changes in the compositional process. The primarily targeted users of this system are musicologists and composers. The system is empirically evaluated with the pieces and with interviews by some composers.

1.4 Organization of Thesis

The remaining part of this thesis is organized as follows: In the next chapter, the related work is summarized. In Chapter 3, the results of sketch studies are presented by examining the compositional processes of the two symphonies composed by Toshiro Mayuzumi, [Publications–1,4,5] and the limitations of sketch studies when studying computer music are discussed. In addition, the necessity for a visual analysis system is proposed. Following these discussions, the basic architecture and primary functionalities of the proposed visualization system, *Spectrail*, are described in Chapter 4 [Publications–2,3,6]. Then, the evaluation results using the data provided by four composers are described, and the effectiveness and limitations of *Spectrail* are clarified through the interviews with those composers in Chapter 5. Chapter 6 concludes this thesis and discusses future research.

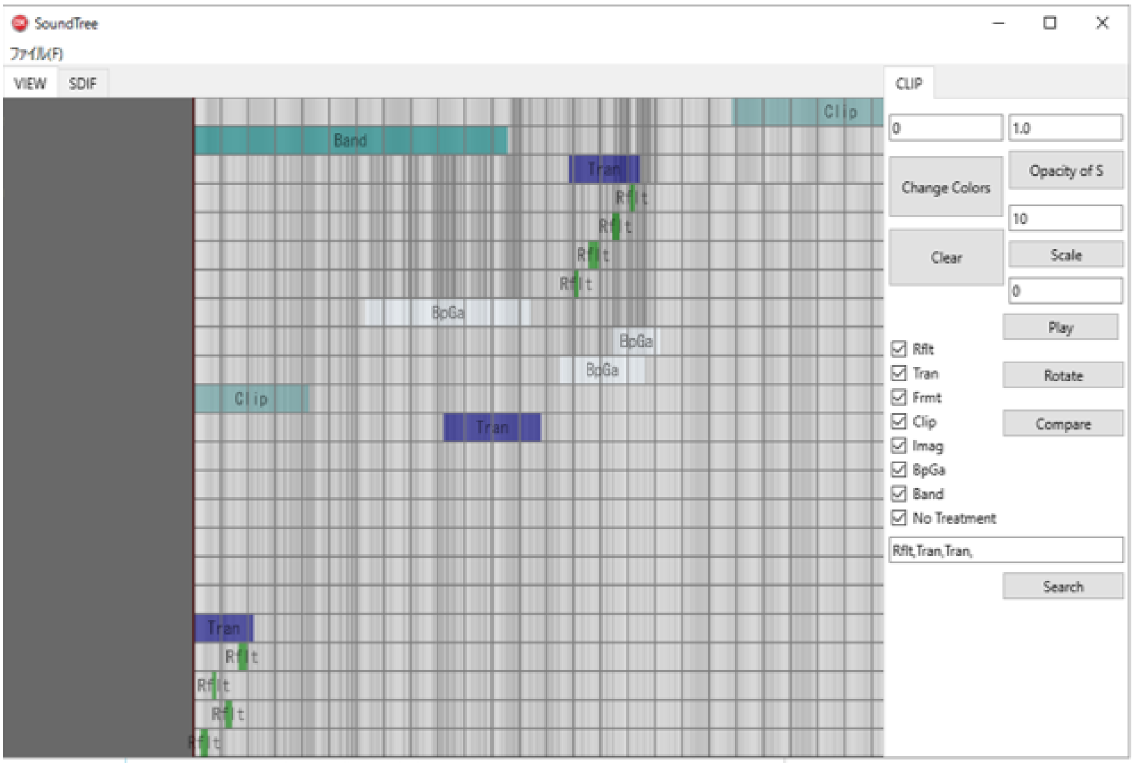


Figure 1.2: Interface of *Spectrail*

Chapter 2

Related Work

The related work of this thesis can be found at the intersection of three domains—musicological study for computer music, music visualization, and provenance visualization. As mentioned in Chapter 1, the aim of the current thesis is to propose a new approach for analyzing the compositional processes made with sound resynthesis by combining two approaches: sketch studies and visual analysis.

The related work of musicology is discussed with a focus on sketches, visual interfaces, and sound descriptor studies to analyze computer or electroacoustic music. Music visualization research is considered in terms of which element of music the research focuses on. Because musical pieces commonly consist of many elements, such as melodies, harmonies, rhythms, timbres, and so on, it makes difficult to understand the organization of the piece. As the provenance research, that chapter introduces some examples which visualize the processes of creation.

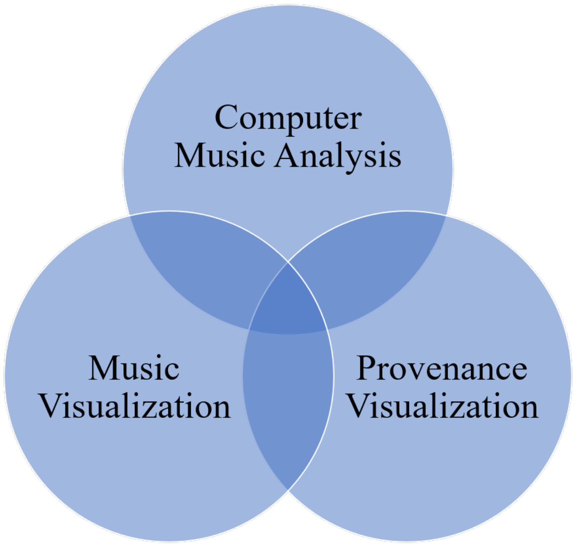


Figure 2.1: The research is at an intersection of these three fields: musicology, music visualization, and provenance visualization

2.1 Putting the Research into the Study Perspective

The current research first proposes a method to visually analyze the compositional process of computer music. In the next section, the previous research in the three fields of musicology, music visualization, and provenance visualization is summarized, as shown in Figure 2.1.

In *Spectrail*, the compositional processes of musical pieces are focused on, whereas the primary targets of most previous works on music visualization are only completed pieces. The current study aims at visualizing a specific provenance—the compositional process of computer music—and can also be regarded as an initial attempt at managing the provenance of *time-series events* in the music visualization field.

2.2 Musicological Study for Computer Music

Computer music is widely studied in many research fields. In this section, the analysis methods of sketch studies, graphical interfaces, and sound descriptors are presented.

2.2.1 Sketch studies

As mentioned in Chapter 1, sketches are used to study the compositional processes of music. Some studies on spectral music are focused on in this section because the compositional method is relevant to Mayuzumi's sketch study in Chapter 3.

Spectral music has been widely studied in musicology [Harvey, 2000] [Pressnitzer and McAdams, 2000], and there are many studies that have intended to clarify compositional processes by using analyzing sketches. The paper [Hirs and Gilmore, 2009] on the compositional process of Gondowana by Tristan Murail points out that the composer mainly used frequency modulation and interpolation of amplitude envelopes or duration curves. In the piece, the frequencies of the bell sounds are calculated by the composer himself. The compositional process of Grisey's *Les Espaces Acoustiques* is clarified by analyzing sketches by Féron [Féron, 2011].

2.2.2 Analysis with graphical interfaces

Many visual interfaces that are used to analyze musical pieces, especially for electroacoustic music, have been developed. Because these kinds of music may not have any scores or any other graphical representation, novel visual representations are needed to analyze these kinds of musical pieces.

A book that focuses on the analysis of electroacoustic music titled *Expanding the Horizon of Electroacoustic Music Analysis* edited by Emmarson and Landy was published in 2016 [Emmerson and Landy, 2016]. It includes graphical interface studies to visually analyze electroacoustic music. The editors point out that the nine benchmarks for analyzing electroacoustic music are "representation, materials, listening behavior, the behavior of materials, ordering, space, performative elements, intention/reception, social, emotional and meaning-related aspects, elements specific to a given genre or piece." Based on these benchmarks, some studies have proposed a novel method for analyzing electroacoustic music [Young, 2016] [Minsburg, 2016]. Kendall [Kendall, 2016] analyzes the aural experience with the five mental layers model, which is used when people listen to electroacoustic music.

In the book by Emmerson et al. [Emmerson and Landy, 2016], several graphical interfaces that can be used to analyze electroacoustic music are presented as well. One of the representative systems is "EAnalysis," which was first proposed by Couprie [Couprie, 2004] [Couprie, 2016]. The system provides information about the electroacoustic music, including pitches and other sound descriptors. In the system, the X-axis represents time position and duration, the Y-axis corresponds to the pitch or frequency range, and the morphology of the shape represents the amplitude of the sound, and some previous visualization methods such as arc diagrams are also utilized. This system visualizes electroacoustic music from multiple aspects, whereas it does not include a guideline to analyze the musical piece.

Park et al. [Park et al., 2009] map in a 3D graph a result of sound analysis using three types of sound descriptors. In the system, twenty-six descriptors are used to analyze the sounds in a musical piece.

Malt [Malt, 2015] develops a system to analyze computer music with sound descriptors. He represents sound changes in a musical piece by placing a particular focus on the brightness of the sound and then analyzing Xenakis's musical pieces. The system uses only one descriptor to show the changes of the sound in the musical piece with a line graph.

2.2.3 Analysis with sound descriptors

Sound descriptors are developed mainly in the music retrieval field and are widely used for music analysis and composition. A comprehensive study of sound descriptors is provided by Peeters [Peeters, 2004] with his proposals of fifty-seven descriptors and a taxonomy of them. Peeters also published a paper in 2011, and it classifies the sound descriptors as temporal, spectral, and energetic properties of sounds [Peeters et al., 2011].

Malt et al. [Malt and Jourdan, 2008] propose a library to interactively compute descriptors. The sound descriptors proposed by them include spectral centroid, spectral spread, energy, spectral flux, and spectral flatness. He points out that descriptors are rarely used in contemporary music compositions because composers do not have knowledge of the relationships between descriptors and the pertinent perceptual characteristics of the sound. One descriptor is not sufficient to characterize a complex "sound state." Recent studies show how the composed functions of descriptors are more effective in recognizing the characteristics of a given sound signal.

There are several known studies and systems aiming at composing musical pieces with sound descriptors. Indeed, Schwarz et al. [Schwarz and Schnell, 2010] propose a method to generate sounds using descriptors. AudioSculpt has some functions to see the loudness, spectral decrease, sharpness, spectral rolloff, spectral centroid and spectral spread of a sound [AudioSculpt, 2019]. Lartllot et al. propose the MIR toolbox [Lartillot et al., 2007] for MatLab system, which includes some descriptors such as attack time, attack slope, attack leap, decay time, decay leap, decay slope, duration, zerocross, rolloff, brightness, centroid, spread, skewness, kurtosis, flatness, entropy, MFCC, and roughness regularity. Cannam et al. [Cannam et al., 2006] develop a system, Sonic Visualiser, to visualize low- and mid-level features from musical audio data.

2.3 Music Visualization

Music remains in time and space, but it does not have any visual representations. The most representative visualization for music is a musical score. It provides a lot of information about music such as pitch, rhythm, tempo, and dynamics in a qualitative manner.

Although musical scores are widely used, they also have several issues when it comes to understanding the organization of the musical piece. One of the issues is that musical scores do not provide an overview of a musical piece, and this makes difficult to understand the features of it. Indeed, these problems make it hard for beginners to understand musical pieces. Moreover, recently, people enjoy streaming music on the Internet and they have to choose a musical piece which they intend to listen to among the enormous number of pieces. This situation has made it more important to provide information about the musical piece to select what users intend to listen to among the many available pieces.

Another issue is that a musical score also has some limitations in expression. For example, pitches that can be written on a musical score are basically twelve in one octave; otherwise, the pitches are consecutive. In addition, composers cannot define timbres precisely because they can only use some verbal descriptions to express them.

Studies on visualization include visualizing the structure of a single musical piece and visualizing multiple pieces. Most of the research on a single piece aims at revealing the structure of the musical piece and focus on elements of the music such as tonality, melody, and harmony.

Research on visualizing the melodies of the musical pieces includes the study conducted by Li [Li, 2016], Snydal et al. [Snydal and Hearst, 2005], and de Prisco et al. [de Prisco et al., 2016]. Snydal et al. visualize the melodies of improvisations in jazz session. They represent the changes of the pitches in the melody with lines.

There are also some studies that focus on repeated patterns. The study by Hayashi et al. [Hayashi et al., 2011] proposes a system, Color Score, that shows the same motifs in a musical piece with the same color. There are also some studies that focus on repeated patterns. Puzoń et al. [Puzoń and Kosugi, 2011] propose a system using the pulse code modulation format. Foote et al. [Foote and Cooper, 2001] compute the similarity of a sound with MFCC (mel frequency cepstrum coefficients) and depict the similarity in grayscale based on the values. For example, if the similarity is high, it will be depicted in white, and if the similarity is low, it turns gray or black. This study is extended by Wolkowicz et al. [Wolkowicz et al., 2009]. The system, “Arc Diagrams,” proposed by Wattenberg [Wattenberg, 2002], where some repeated patterns used in the musical piece are visually identified with some arcs.

Visualization studies focusing on the tonality and harmonies of the piece include the system proposed by Sapp [Sapp, 2005], which provides an overview of the transition of the tonality in music. Another system is presented by Moudirossian et al. [Mardirossian and Chew, 2007] and represents the change of tonality based on the tonal pitch space theory proposed by Lerdahl. In this system, a musical piece is segmented by users, and the tonality in the segment is highlighted. Bergstrom et al. [Bergstrom et al., 2007] propose the isochords system based on the Tonnetz theory proposed by musicologist Karl Wilhelm Julius Hugo Riemann and this system visualizes the progress of chords with paved triangles. Malandrino et al. [Malandrino et al., 2015] also develop a system to show the tonality and harmonics in music.

The studies that attempt to show the mood of the performance include the study by Hiraga et al. [Hiraga and Matsuda, 2004], which visualizes the atmosphere of music with a colored square. Herremans et al. [Herremans and Chew, 2016] create visualization to show the tensions of music, where the tensions are represented with cloud diameter.

Some of the studies focus on the instruments in orchestral music. Chan et al. [Chan et al., 2010] visualize orchestral music, focusing on the roles of the instruments. Cuiha et al. [Ciuha et al., 2010] represent the structure of music by mapping the type of instruments in depth. In this system, the same tones are represented as the same point to avoid visual clutters.

Cantereira et al. [Cantareira et al., 2016] introduce a system—MoshViZ—to show the multiple elements in music. They take Midi data as the input, and the horizontal axis shows the time in the sound, and the vertical axis represents the pitch. In the overview mode, abstracted pitch movement is represented, and in the detailed view, the system provides precise information about the sound.

Several visualization studies peer into the relationship between multiple musical pieces. A system developed by Muelder et al. [Muelder et al., 2010] uses graphs to visualize the relationship between the musical pieces. Their method is based on physical acoustic similarity. Juhasz [Juhasz, 2011] proposes a method to classify the musical pieces with a self-organizing cloud algorithm. The input data of the system are pitch movements in melodies of 32,000 folk songs that are represented with lines. Zhu et al. [Zhu and Lu, 2005] create a system based on a pre-user study to explore the impression from the musical pieces. They map the music collection with two axes–one is for the brightness (Bright-Dark), and the other is for the tempo (Fast-Slow). Each musical piece is mapped as a small rectangle in two-dimensional space.

These related work only targeted at completed pieces whereas the target of research in the current thesis is the compositional process.

2.4 Provenance Visualization

Provenance visualization can also be regarded as another key aspect in the current study because it concentrates on the history of sound processing. The term "provenance" has been used in a variety of ways to describe the history of visualization research [Rio and da Silva, 2007] [Ragan et al., 2016] [Walker et al., 2013]. Comprehensible provenance visualization systems include VisTrails [Bavoil et al., 2005] [Callahan et al., 2006] [Silva et al., 2007], which has been widely applied to scientific research fields [Morisette et al., 2013] [Zhang et al., 2013], and VIDELICET [Fujishiro, 2007]. Both systems visualize the modular structure of a visualization application and support how to design effective visualization workflows for a given goal. Jankun-Kelly et al. [Jankun-Kelly et al., 2007] present the P-set model to describe a process of visualization exploration. Anand et al. [Anand et al., 2010] propose an interactive provenance graph browser for visualizing data dependency. Brown et al. [Brown et al., 2012] present a prototype implementation, DisFunction, that can interact with a visualization to define a distance function.

There are other systems that can support design or sensemaking [Andrews et al., 2010]

[North et al., 2011] [Xu et al., 2015]. Kurlander et al. [Kurlander and Feiner, 1988] introduce an editable graphical history as well. The main method to visualize provenances is timelines [Brehmer et al., 2017]. Recently, Doboš et al. [Doboš et al., 2014] propose an approach for visualizing the history of constructing 3D models with timelines.

Other studies introduce several types of designs to visualize specific provenances. The system by Chen et al. [Chen et al., 2012] visualize provenances using networks, studying the provenance of a satellite imagery processing pipeline and a large-scale computer network testbed. The work by Borkin et al. [Borkin et al., 2013] is one case where provenances are visualized with a tree diagram, and the effectiveness of the system is demonstrated by some user studies. Yoon et al. [Yoon et al., 2013] propose a system that visualizes the history of code changes using timelines. Some studies propose analysis approaches that combine multiple visualization methods. Dunne et al. propose a system to analyze networks by combining familiar graphs and charts [Dunne et al., 2012]. Heer et al. [Heer et al., 2008] represent visualization states with a graphical interface.

Chapter 3

A Sketch Study: The Compositional Process of Two Symphonies by Toshiro Mayuzumi

This chapter presents a sketch study to clarify the compositional process of Toshiro Mayuzumi's two symphonies: *Nirvana Symphony* (1958) and *Mandala Symphony* (1960). Mayuzumi is known as a composer who composed his works based on the analysis of temple bell sounds, and these two symphonies are his representative works. When he composed these symphonies, he analyzed the sounds of temple bells and orchestrated them. This process is similar to the compositional process of spectral music, which started in the 1970s. His compositional processes are clarified using his autograph sketches, the Campanology Documents, which are for the first time introduced in this research. Based on the research, the effectiveness and limitations of a sketch study are discussed, and the necessity for a visual analysis approach is shown.

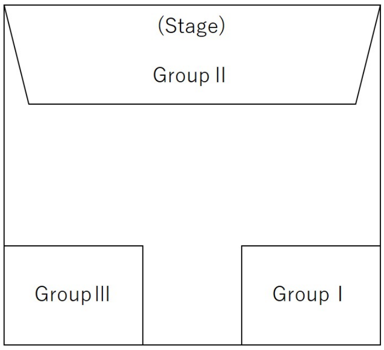


Figure 3.1: Arrangement of three orchestra groups

3.1 Introduction

Premiering in the 3rd "Three composers' society (Sannin-no-kai)" concert in 1958, *Nirvana Symphony* is a work consisting of six movements. The odd movements entitled "Campanology I-III" are performed only by the instrumental orchestra, and "SHURENNENJINSHU (首楞厳神咒)" (2nd movement), "MAKABON (摩訶梵)" (4th movement), and "Finale" (6th movement) are played with a male chorus in addition to the instrumental orchestra. The large orchestra accompanied by a male chorus is divided into three groups and arranged to surround the audience. The first orchestra group includes woodwind instruments, while the second has string instruments, woodwind instruments, brass instruments, percussions, and a male chorus, and the third consists of brass instruments. The score, published by Edition Peters, shows two types of seating plans for the orchestra [Mayuzumi, 1969]. One is for concert halls with balconies, and the other is for venues with a single story (see Figure 3.1). Mayuzumi explains the acoustic effect provided by these three groups of orchestras using the term "Campanology effect" in the program notes of the premiere.

Mandala Symphony premiered at the 4th "Sanninnokai (Three Composers' Society)" concert in 1960. The work has two movements: the first movement is titled "KONGŌKAI Mandala (金剛界曼荼羅)," and the second is "TAIZŌKAI Mandala (胎蔵界曼荼羅)." The latter symphony is played only with an instrumental orchestra.

Although these works have been widely studied soon after their premieres, most of the scholars evaluate these symphonies as a fusion of Eastern and Western music [Herd, 1989] [Heifetz, 1984]. There are also some previous studies on the compositional methods for composing this kind of music. Shimizu [Shimizu, 2007] [Shimizu, 2010] analyzed these symphonies from the viewpoint of how Mayuzumi used tone series and Buddhist music pointing out that the twelve-tone series is used for the chord formation in these pieces. However, no research has clarified the compositional process by using the primary materials.

Based on these previous works, a survey of primary resources is conducted. Most of the primary resources, including Mayuzumi's autograph sketches, were archived in the TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University after his death. Throughout the survey, a series of primary resources titled the Campanology Documents was discovered in the archive. Including his handwritten sketches, the Campanology Documents are a series of primary resources related to Mayuzumi's multiple pieces that are based on temple bell sounds.

In this chapter, the compositional processes of the *Nirvana Symphony* and *Mandala Symphony* are examined using the Campanology Documents, and the relation between the two works is clarified. Section 3.2 gives an overview of the Campanology Documents. In Sections 3.3 and 3.4, the main chords of the *Nirvana Symphony* and *Mandala Symphony* are described by referring to the remarks by Mayuzumi, respectively, and the process used to create the main chords is presented through an analysis of the Campanology Documents. Finally, in Section 3.5, the processes for creating the main chords between the two works are compared.

3.2 Outline of Campanology Documents

The Campanology Documents consist of eight documents that include seven of Mayuzumi's handwritten sketches plus a copy of a book that he might have referred to during his composition. TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music numbered these documents from No. 1 to No. 8. These numbers are based on the order according to which the materials overlap and do not correspond to the creation time of each material. In the current thesis, each document is referred to with these numbers. The basic information of each document is tabulated in Table 3.1.

Table 3.1: Data of Campanology Documents

Number	Material	Size (mm)	Writing instrument	Remarks
1	One music sheet	397*287	Black pencil	"Campanology 資料 (Campanology Documents)"
2	Six pages of stapled notepaper with "NHK" mark	290*203	Black, red, and blue pencils	"Oct. 26 1966"
3	Two pages of graph paper	290*203	Black marker, red pencil, ball-point pen	
4	A copied paper with cardboard	351*248		"EL[E]KTRONISHE MUSIK SKALA"
5	One music sheet	454*281	Black and red pencils, black marker	
6	One music sheet	454*281	Black and red pencils, ball-point pen	34 steps musical sheet, pasted with 6 steps musical sheet
7	One music sheet	397*284	Black and red pencils, black marker	
8	One music sheet	394*286	Black pencils, black marker	

In Document No. 1, the title and signature "Campanology Documents // Toshiro Mayuzumi" is written with large letters, and it can be regarded as the cover sheet of the Campanology Documents, as shown in Figure 3.2. Documents No. 2, No. 3, No. 5, No. 6, No. 7, and No. 8 are sketches, and Document No. 4 is a copy of a book. Document No. 2 consists of six sheets, Document No. 3 has two sheets, while Documents No. 4 through No. 8 only have one sheet. Document No. 2 is written on a paper with the NHK (Japan Broadcasting Corporation)'s logo, Document No. 3 is on grid papers, while the other handwritten documents are written on music papers. From the next section, the process of creating the main chords of the *Nirvana Symphony* and *Mandala Symphony* is presented using these Campanology Documents.

This image is unavailable on the Internet
due to copyright protection.

Figure 3.2: Cover of the Campanology Documents (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

3.3 *Nirvana Symphony*

3.3.1 Main chords of *Nirvana Symphony*

In the *Nirvana Symphony*, the three chords shown in Figure 3.3 are used as the main chords, which were presented with Mayuzumi's own handwriting in the concert program at the premiere. To distinguish these chords from the Western traditional harmonic theory, Mayuzumi called these chords "GO-ON" in the premiere program. Three or more accumulated sounds are referred to as the more general term "chord" to avoid complication of wording. These three chords are called "main chord 1," "main chord 2," and "main chord 3," which are played at the beginning of the *Nirvana Symphony* and repeatedly used in the first and fifth movements. Three types of main chords always appear in the same place in the symphony.

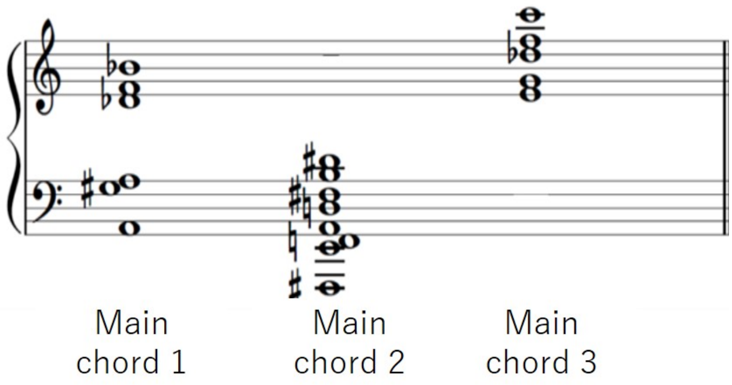


Figure 3.3: Main chords of the *Nirvana Symphony*

Main chord 1 is played by orchestra group 2, main chord 2 by orchestra group 3, and main chord 3 by orchestra group 1. Mayuzumi explained the compositional process of these main chords in the program of the premiere [Mayuzumi, 1958].

With the assistance of the NHK, I obtained some recorded temple bell sounds broadcast on New Year's Eve and some results of analyzing each temple bell sound. (snip) After taking the average of the frequencies of each overtone, I tried to create an acoustical effect as temple bell sounds with the NHK Symphony Orchestra. (snip) As a result of these various experiments, I composed a symphony titled "Campanology" based on the sounds of seven temple bells and three HANSHOs (small temple bells), and it premiered through broadcasting in November last year. Seven types of bells include the bells of Todaiji, which is the largest bell in Japan with 9 SHAKU 1 SUN, Byodo-in Temple, Honenin Temple, Zenrinji Temple, Myoshinji Temple, Kekoji Temple, and Daiunji Temple.

私は *NHK* の厚意を得て，毎年大晦日に放送される除夜の鐘の録音テープを手に入れ，その一つ一つの特徴ある鐘の音を音響分析してもらった．（中略）分析の結果の近似値をとって平均率（原文ママ）に直し，今度は *NHK* 交響楽団に協力してもらって，鐘の音的オーケストレィションの実験を繰返した．（中略）これらさまざまな実験の結果，私は *7* 種類の梵鐘と *3* 種類の半鐘の音をもととして，「カンパノロジイ」という曲を試作し，昨年 *11* 月に放送発表した．*7* 種類の梵鐘とは，口径 *9* 尺 *1* 寸わが国最大の鐘である東大寺の鐘，以下，平等院，法然院，禪林寺，妙心寺，華光寺，大雲寺の鐘である．

In these remarks, Mayuzumi mentions which temple bells he used as the sound materials for the *Nirvana Symphony*. He said that he used seven temple bell sounds, including Byodoin, Honenin, Zenrinji, Myoshinji Kekoji, Daiunji and Todaiji, and three HANSHO sounds. In the following sections, the precise compositional process of the *Nirvana Symphony* will be clarified based on Mayuzumi's remarks and an analysis of the Campanology Documents.

3.3.2 Analysis of Campanology Documents

Document No. 3

Document No. 3 includes two documents written on grid paper. A black pen, a ballpoint pen, and a red pencil are used as writing instruments. In the current thesis, the two documents in Document No. 3 are called Document No. 3-1 and Document No. 3-2 according to the order in which the documents overlap. Document No. 3-1 is clearly related to the *Nirvana Symphony*, whereas any no relation can be found in Document No. 3-2. Document No. 3-1 is shown in Figure 3.4. At the top of the document, there are the title and the author of a book "JIKKEN ONKYO GAKU (Experimental Acoustics)" by Keiji Yamashita (included in "ONKYO KAGAKU (Acoustical Science)" edited by Shuji Yagi and published by Ohmsha). Keiji Yamashita (1899–1969) was an acoustical physicist and a professor of Kyoto Imperial University. The Yamashita paper "Experimental Acoustics" [Yamashita, 1948] is presented in the third chapter of the book titled "Acoustical Science" published by Ohmsha in 1948 although it is unclear why Mayuzumi obtained the paper. Hereinafter, the paper will be referred to as the "Yamashita paper." Document No. 3-1 has a table that summarizes the data of the overtones' frequencies of the six bell sounds and three HANSHO sounds ((a), (b), (c)). The temple names, "Byodoin," "Honenin," "Zenrinji," "Myoshinji," "Kekoji," "Daiunji," and the three HANSHOs are also found in the table. A description about the "Todaiji" bells' sound is also shown at the bottom of the document, which mentions that estimating the fundamental tone of Todaiji temple bell sound based on the ratio between the diameter and frequency of a fundamental tone of Byodoin temple bell, the fundamental tone of the Todaiji temple bell (caliber length = 2.76 m) should be 44 c/s. This frequency corresponds to Fis_2 tone. These descriptions are perfectly consistent with Mayuzumi's remarks in his program note. This implies that Document No. 3-1 must be regarded as a sketch for the *Nirvana Symphony*. Document No. 3-1 also has data on "diameter," "ratio between the overtones' frequencies," "additional tones for creating a beat," "volume changes," and "average ratio between the overtones' frequencies." After comparing these descriptions with the Yamashita paper, it is apparent that some of them are not included.

The data quoted from the Yamashita paper is "diameters of temple bells," "overtones' frequencies," and "the ratio between the overtones' frequencies." In the Yamashita paper, the data of "diameter of bells" are shown as "Table 2," "overtones' frequencies" as "Table 3," and the "ratio between the overtones' frequencies" as "Table 4." Comparing these data with the content of Document No. 3-1, after comparing these descriptions with the Yamashita paper, it is apparent that some of them are not included. On the other hand, in the Yamashita paper, there is no description about the "additional tones for creating a beat," "volume changes," and the data on Todaiji temple bell sounds. The "additional tones for creating a beat" is mentioned in the descriptions on the temporal volume changes of the temple bells' overtones, which is written at the bottom of Document No. 3-1. In the descriptions, the "additional tones" are explained as "adding a tone near the second or third overtone is effective for creating a beat, however, the beat is inadequate at 700 c/s." The frequency of the "beat (about five cycles difference with the regular overtones)" is also shown in the table of Document No. 3-1 with the terms "additional tones for creating a beat" or simply "beat." The "additional tones for creating a beat" is added near the second, third, and fourth overtones of "Byodoin," the third overtone of "Myoshinji" and "Kekoji," and the difference between the frequencies of these overtones and the "additional tones for creating a beat" is 2 to 5 c/s (Hz).

The "dynamics" are indicated by graphs and comments at the bottom of the document. The graph shows the temporal volume change of the first to sixth overtones of the bell sound. According to this graph, the volume of the first overtone rapidly drops, and the other overtones gradually decrease. Although the overtones' frequencies of the Todaiji temple bell are not directly mentioned in the Yamashita paper, it can be guessed that Mayuzumi referred to Yamashita [Yamashita, 1948] on the ratio between the diameter of the temple bells and the frequency of the bell sounds: "The frequency of the bell corresponds mostly to the inverse proportion of their diameter."

This image is unavailable on the Internet
due to copyright protection.

Figure 3.4: Document No. 3-1 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

Document No. 5

A slitted music paper is used for Document No. 5 (see Figure 3.5). This document includes the names of the temples and some chords based on the overtones of the temple bell sounds (hereinafter "bell chords"). Each bell chord is written using two staves, and the names of the temples are written on the left side of each staff. The temple names are written as "Byodoin," Honenin," "Zenrinji," "Myoshinji," "Kekoji," "Daiunji," "(a)," "(b)," "(c)," and the "average of overtones' frequencies." The chord of "average of overtones' frequencies" corresponds to Yamashita's description—"the average ratio between the frequencies of overtones." Some alphabets are written with the temple bells' names to distinguish each of the bell chords.

The bell sounds are written in two ways. The chords written on the left are called "Bell chord 1," and the chords on the right are called the "Bell chord 2." The tones included in the "Bell chord 1" and "Bell chord 2" are almost the same. However, the notation methods are different between the two types of chords. First, the "additional tones for creating beat" appearing in Document No. 3-1 are not added to the "Bell chord 1" but are rather added to the "Bell chord 2." For example, in the "Bell chords" of Byodo-in Temple, "additional tones for creating beat" are not added to "Bell chord 1," but added to "Bell chord 2." "Bell chord 2" has a gis tone near the second overtone and a cis^1 tone near the third partial tone for creating the "beat." Second, to describe the microtones, "Bell chord 1" describes them with words such as "a little lower." On the other hand, "Bell chord 2" is noted with microtones. For example, in the highest tone of the Bell chord of Honenin, the marks of the microtones representing a tone lower than the semitone are used for "Bell chord 2," but "Bell chord 1" only has some descriptions about the pitches.

This image is unavailable on the Internet
due to copyright protection.

Figure 3.5: Document No. 5 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

Document No. 8

Document No. 8 is written on a two-sided folded music paper, and the sketches are written on both sides. The right side of the surface has the title of the document: "CAMPANOLOGY // TOSHIRO MAYUZUMI." On the inside of the documents, sketches of the tone series are written on the left side and a table of chords on the right side.

On the left of the inner side, the original forms of the tone series and their inversions are written from the first through the twelfth steps. The tone series are transposed on semitone upper as going downward in the score. Although clefs are not written in this document, if it is written with a G clef, the fundamental tone series of the *Nirvana Symphony* first appeared from the 36th bar of the first movement (a^1, b^1, e^1, f^1, fis^1, gis^1, g^1, cis^2, d^2, es^2, c^2, h^1). In the table of chords on the right side (see Figure 3.6), ten types of chords are transposed so as to be the fundamental tones aligned with C.

It can be argued that the chords in Document No. 8 are created with the pitch transformation of the "Bell chord" in Document No. 5. Comparing the table of chords with the main chords, main chord 1 is based on the sound of Byodoin's bell. The chord is written without any microtones, added A tone, and is not transposed. Their pitches are the same as that of Byodoin's bell sound. The main chord 2 consists of the "average of overtones," transposing the fundamental tone to be the Fis_1 tone, with the E tone added to the second tone from the bottom. Main chord 3 is a transposition of the partial tones of the HANSHO (c), which is accomplished by changing the fundamental tone to e^1, and its g^2 tone, which is the second tone from the bottom of the chord, is lowered one octave. According to this comparison, each main chord is selected from the chord table in Document No. 8 and is used by adding some sounds and changing the pitches. The reason why he added some tones when he creating the main chords is not clear, but Mayuzumi may have changed the sound from the perspective of orchestration. In addition, the chords based on the bell sounds are not only used in the odd number movements of the *Nirvana Symphony*, but they are also used in the even number movements.

This image is unavailable on the Internet
due to copyright protection.

Figure 3.6: The right side of the chord table in Document No. 8 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

Document No. 4

In this section, Document No. 4, which Mayuzumi referred to when he made the main chords (see Figure 3.7), is presented. Document No. 4, titled "EL[E]KTRONISCHE MUSIK SKALA," is a copy of a book attached to cardboard, however the title of the book cannot be made clear.

The table is written on a musical score with ten steps. The title—"EL[E]KTRONISCHE MUSIK SKALA"—is written at the top of the score, and the description "C_1=32, a^5=432," which indicates the frequencies of the lowest pitch, C_1, and a^5, is written under the title. The first step of the score includes the explanation of the ratio between the frequency of C and the other tones. In the second and following steps, the chromatic tone series and the frequencies of each tone are written.

Although the reason behind why Mayuzumi used this document is unclear, it can be estimated that he used this document for writing "Bell chords" in Document No. 5.

This image is unavailable on the Internet
due to copyright protection.

Figure 3.7: Document No. 4 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

3.3.3 Compositional process of the *Nirvana Symphony*

According to the descriptions in the Campanology Documents, the compositional process of the *Nirvana Symphony* is as follows: Mayuzumi obtained the data on the frequencies of the temple bell sounds and made Document No. 3 while also creating Document No. 5 based on Document No. 3. In Document No. 5, each temple bell sound is written with microtones. Document No. 8 was made based on Document No. 5, and the three main chords were created from Document No. 8.

There is a difference between Document No. 5 and Document No. 8 in terms of using microtones. In Document No. 5, Mayuzumi uses microtones to represent the temple bell sounds. Mayuzumi might have tried to play the temple bell chords with the NHK Symphony Orchestra after making Document No. 5. Regarding the microtones, Mayuzumi mentioned that he first tried to use microtones; however, he could not compose such a huge symphony with them, and the orchestra could not play them correctly [Mayuzumi et al., 1959]. Because it is difficult to play microtones, Mayuzumi did not use them when he created Document No. 8 although he used microtones in Document No. 5.

Although the chords used in the *Nirvana Symphony* clearly correspond to the temple bell's sound frequency data presented in the Yamashita paper, the data provided by the NHK are not found. However, it cannot be denied that Mayuzumi asked the NHK to analyze the temple bell sounds. For example, Document No. 3 has a description of the acoustic characteristics of a temple bell sound that is not found in the Yamashita paper. In addition, there is a possibility that the other documents might have information about the *Nirvana Symphony.* More research is needed in this area to deeply analyze the relationship between the sound analysis results provided by the NHK and Yamashita paper.

3.4 *Mandala Symphony*

3.4.1 Main chords of *Mandala Symphony*

Next, the compositional process of the main chords in the *Mandala Symphony* is discussed using the Campanology Documents and Mayuzumi's remarks in the program for the premiere [Mayuzumi, 1960].

> *During the composition of "Nirvana Symphony," I found that there are differences between the numbers which the tones appeared in the overtones of the bell sounds and created the tone series in "figure one" according to how often the tone appears." When I analyzed several temple bells' sounds and arranged them on the same fundamental tone (in this case, C tone), tone series as shown in "figure one" is appeared.*
>
> 「涅槃交響曲」の作曲中，梵鐘の部分音構成を研究しているうちに，私は非常に興味深い発見をした。数十種類の古今東西にわたる梵鐘の部分音の平均値を取ってみたところ，第一図の様な結果を生んだのである。すなわち，いくつかの梵鐘を音響分析して，仮りに（原文ママ）それらを或る一つの同一の基音（この場合 *do*）上に，現れてくる部分音の頻度の順に並べてみると，第一図の様になるのである（この場合，部分音が第何次部分音であるかは問題でなく，専ら，同一基音上の同一部分音を対象とする）.

This "figure one" is shown in Figure 3.8. This tone series consists of ten tones, and these tones do not include C and G tones. In the program notes for the premiere, Mayuzumi mentioned that the constituent tones of the tone series correspond to the average frequencies of the temple bells' overtones. In the current thesis, these two chords are called "main chord A" and "main chord B," as shown in Figure 3.9.

Figure 3.8: The tone series in "figure one." The main chords are organized with the fundamental tone series

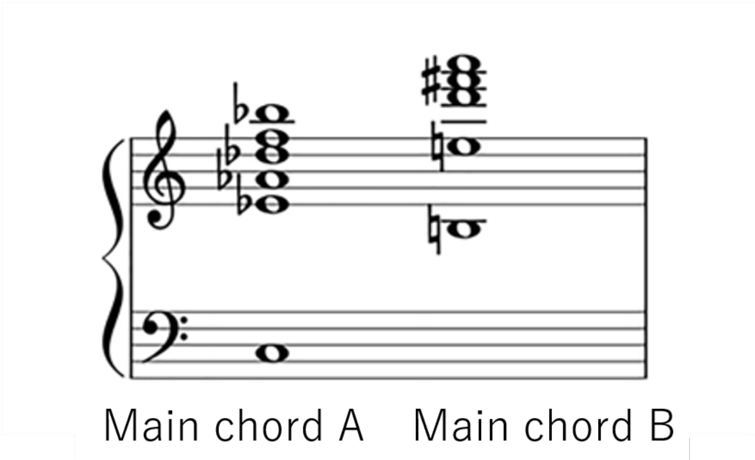


Figure 3.9: Main chords of the *Mandala Symphony*

3.4.2 Analysis of Campanology Documents

Document No. 7

This document was created with two folded musical papers. The details of the compositional process are written on the outside of the document although the other side also has some sketches. These sketches are written with a black pencil and a fiber pen. There is a table with ten types of chords.

The fundamental tones of the ten chords are C. These chords are noted with temperament. Some of the chords are also written with microtones. There are pitch names on the top of the table, and the constituent tones of each chord are aligned. For example, because the constituent tones of the chord on the top of the table are "c, h, dis^1, gis^1, c^2, f^2, a^2," there are two notes in column "C," one note in "Do," "Re ♯," "Fa," "Sol ♯," "La," "Si." Compared with Document No. 5, these ten chords are "average of frequencies," "Byodoin," "Honenin," "Honenin," "Zenrinji," "Myoshinji," "Kekoji," "Daiunji," and "three types of HANSHO." The chords with microtones are found in the chords of "Byodoin" and "Honenin." There are three number strings under the table. The numbers in the first number string indicate the number of identical pitches in the seven "Bonsho" chords. For example, because cis is included in four Bonsho chords (Byodoin, Honenin, Zenrinji, Myoshinji), the number in the column "cis" becomes "four." The numbers in the second number string indicate the numbers in "HANSHO" harmonics, and those in the third show the sum of the first and second number strings, as shown in Table 3.2.

Eight kinds of tone series described in Figure 3.10 are written in Document No. 7, and the two tone series are clearly related to the number strings. In the current thesis, as shown in Figure 3.11, the two types of these tone series written on the left side are called tone series 1 and 2. The notes of the tone series 1 are gis^1, cis^1, f^1, ais^1, h^1, d^1, dis^1, e^1, fis^1 and a^1, and comparing this tone series and Table 3.2, the order of the tones is determined according to how large the number is. The Roman numerals "I-II-III-IV" are also found, and these depend on the order of the numerical values written in the first row of Table 3.2. The constituent notes of

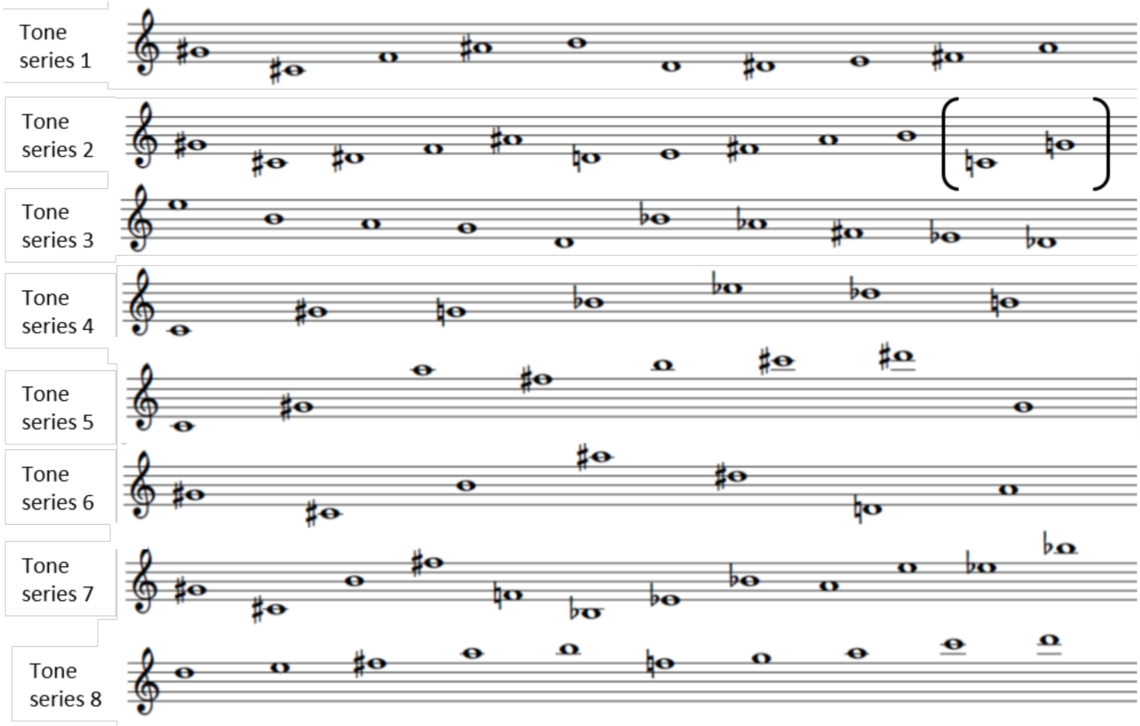


Figure 3.10: Eight tone series in Document No. 7

the tone series 2 are gis^1, cis^1, dis^1, f^1, ais^1, d^1, e^1, fis^1, a^1 and h^1. In this tone series, the order of the tones is determined according to the sum of the bonsho and the HANSHO partial tone, which corresponds to the third row in Table 3.2. If the number of the tone is the same, (meaning the number indicated in the third row of Table 3.2 is identical), the sounds are arranged from low to high. In addition, in the parentheses written at the end of the tone series, the C tone and the G tone are written. The C tone is used as a tone, and the G tone is not contained at all in the bell chord. “Augmented 5th, augmented 1st, minor 3rd, perfect 4th, minor 7th, major 2nd, major 3rd, major 4th, major 6th, and major 7th” are written under each sound of the tone series, indicating the pitch interval between the C tone and the component of the tone series.

The tone series 1 and 2 are created based on “frequencies of overtones found as a result of acoustic analysis” written by Mayuzumi in the program for the premiere. Tone series 3 (e^2, h^1, a^1, g^1, d^1, b^1, as^1, fis^1, es^1 and des^1) is written under tone series 1 and 2. Although C and G tones are not used in tone series 1 and 2, and a C tone is not used in tone series 3 either, the

Table 3.2: The number strings in Document No. 7

	C	Cis	D	Dis	E	F	Fis	G	Gis	A	Ais	H
Temple bell	2	4	2	2	2	3	2	0	5	1	3	3
HANSHO	0	2	1	2	1	1	1	0	2	2	1	0
Total	2	6	3	4	3	4	3	0	7	3	4	3

F and G tones are included in the constituent tones of tone series 3. In addition, a description "tone series for harmony (Harmonie TEKI ONRETSU)" is written under tone series 2, implying that tone series 2 is expected to be used as a chord. As described above, a total of eight kinds of tone series are written in Document No. 7. Hereafter, the features of the remaining five types of tone series are described. Five types of tone series are written on the right side of Document No. 7 (Figure 3.12). These five tone series are sequentially called tone series 4, tone series 5, tone series 6, tone series 7, and tone series 8.

The constituent tones of tone series 4, which were written on the left side of Document No. 7, are "c^1, gis^1, g^1, b^1, es^2, des^2, and h^1," and those of tone series 5 are "c^1, gis^1, a^2, fis^2, h^2, cis^3, dis^2, and g^1." The pitch intervals between the side-by-side tones in tone series 4 are c^1 → augmented fifth → Gis^1 → augmented first → g^1 → minor third → b^1 → perfect fourth → es^2 → major second → des_2 → diminish thirds → h^1. In addition, the interval between the sixth tone and the seventh tone is a major second if the des (the sixth note) is regarded as cis^2, which is the enharmonic of des. The pitch intervals between each constituent tone of tone series 5 are c^1 → augmented → gis^1 → one octave minor second → a^2 → minor third → fis^2 → perfect fourth → h^2 → major second → cis^3 → major second → dis^3 → one octave and augmented fifth → g^1, and the interval between the second and third tones is one octave and augmented first if the gis^1 is regarded as as^1. Thus, the pitch intervals between each tone in the former seven tones of tone series 4 and 5 are equal if the enharmonic is taken into account. Both of the pitch intervals between the constituent tones in the tone series 4 and 5 are "augmented fifth," “augmented first,” “minor third,” “perfect forth,” “major second,” and “major second.” These pitch intervals almost correspond to the intervals between the c and the constitute tones of tone series 2 (“augmented first, augmented first, minor third, perfect fourth, minor seventh, major second, major third, augmented fourth, major sixth, and major seventh”). Thus, the intervals between the constituent tones of tone series 4 and 5 are determined in a similar way as tone series 2.

In the document, tone series 4 and 5 are deleted with a large cross mark and they are not used in the *Mandala Symphony.* Although the intentions of Mayuzumi are not clear, there is a description “melody-like tone series” under tone series 4 and 5. According to the description, the tone series are likely to be created for making melodies in the piece.

Tone series 6 through 8 are written on the right side of tone series 4 and 5. The constituent sound of tone series 6 is “gis^1, cis^1, h^1, ais^2, dis^2, d^1, and a^1,” those of tone series 7 are “gis^1, cis^1, h^1, fis^2, f^1, b, es^1, b^1, a^1, e^2, es^2, and b^2,” and those of the tone series 8 are “d^2, e^2, fis^2, a^2, h^2, f^2, g^2, a^2, c^3, d^3.” Because the feature of tone series 6 and 7 are similar to that of series 1 and 2, these tone series are based on the bell sounds. In addition, the first five tones of tone series 8 are the same as the last five tones of tone series 2, and the last five tones of the tone series 8 correspond to tone series 2 when transposed to the minor third higher.

In tone series 1 through 8, tone series 2 corresponds to the tone series shown in the program for the premiere. It can be said that tone series 2 is the fundamental tone series of the second movement “Taizokai Mandala.” Further, many kinds of tone series were created based on the frequencies of bell sounds. According to these results, it is possible to say that the tone series shown in “Figure one” is just one of the tone series created based on the bell overtones. In Document No. 7, the process to apply the bell overtones to the *Mandala Symphony* is presented. On the right side of the document, a tree diagram, in which two kinds of sounds branch out from one sound, is written on the upper side of tone series 4 and 5. Although clefs are not written in the tree diagram, the following regularity can be seen if the head sound of the tree diagram is regarded as a c^2 tone. Whereas some of the pitch intervals between the constitute tones of the tone series are different because of the enharmonic, the pitch intervals between the tone before the branch of the tree diagram and the tone after the branch are "augmented fifth," "augmented first," "minor third," "perfect fourth," "minor seventh," and "major second." On the right side of this tree diagram, there is one more tree diagram that shows "minor third" pitch intervals. The pitch intervals "augmented fifth," "augmented first," "augmented third," "perfect fourth," "minor seventh," "major second," and "major third" correspond to the pitch intervals between the C sounds and the constitute sounds of the fundamental tone series.

This image is unavailable on the Internet
due to copyright protection.

Figure 3.11: The left side of Document No. 7 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

This image is unavailable on the Internet
due to copyright protection.

Figure 3.12: The right side of Document No. 7 (TOYAMA Kazuyuki Memorial Archives of Modern Japanese Music of Meiji Gakuin University)

3.4.3 Compositional process of the *Mandala Symphony*

The compositional process of the *Mandala Symphony* is clarified here. Mayuzumi counted the number of pitches included in the constituent sounds of the six kinds of bell overtones and three types of HANSHO overtones, hence creating the tone series based on the numbers of the pitches. The relationship between tone series 2 and the main chords shown in Figure 3.9 is that main chord A is made from the first half of tone series 2 and main chord B is from the last five sounds of tone series 2, meaning that the compositional process of these main chords is concretely clarified. In addition, to compose the main chords and tone series of the *Mandala Symphony*, Mayuzumi focuses not only on how often the twelve tones are used in the temple bell sounds, but also the pitch intervals between the tone series and the C sound.

3.5 Comparison between *Nirvana Symphony* and *Mandala Symphony*

Comparing Document No. 8 (document for the *Nirvana Symphony*) and Document No. 7 (document for the *Mandala Symphony*), it became clear that both of these documents were created based on Document No. 5. In addition, as mentioned in Section 3.3, Document No. 5 is based on the data quoted from the Yamashita paper. According to these facts, although the *Nirvana Symphony* and the *Mandala Symphony* are based on the same bell sound, the process behind creating the main chords is different. In the *Nirvana Symphony*, the main chords are directly based on the temple bell overtones. On the other hand, the main chords of the *Mandala Symphony* are based on the regularity of the constituent tones in bell overtones. Mayuzumi created the tone series based on the regularity of the bell overtones, and the main chords are based on the tone series. Moreover, in the *Mandala Symphony*, the chords are created by dividing the tone series into two groups and organizing the chords. This is the characteristic method of Mayuzumi, which is also found in the *Phonology Symphonic* (1957) and other pieces already pointed out in the previous research [Shimizu, 2010]. However, in the *Nirvana Symphony*, although some attempts to modify the bell chords are found, the pitch intervals between the constitutional tones of the chord itself reflect the overtones of the bell sounds directly. According to these analyses, the *Nirvana Symphony* is a work directly utilizing the regularity of the bell overtones, while the *Mandala Symphony* is regarded as an attempt to apply them to his conventional composition method.

Based on the analysis, both the *Nirvana Symphony* and *Mandala Symphony* were created using the same bell sounds. Mayuzumi tried to expand his expression by analyzing the regularities of the bell sounds from multiple aspects and applied them to his own conventional technique.

3.6 Possibilities and Limitations of Sketch Studies

The current study shows that sketch studies are effective in clarifying what a composer did in the compositional process. In addition, sketches include much information that would not have appeared in the completed pieces, so these studies are also effective in discovering the trials not included in the completed piece.

However, it is difficult to clarify all of the compositional processes only with sketches. For example, it cannot be determined why Mayuzumi chose the main chords of the *Nirvana Symphony*, even though he created many more chords based on the temple bell sounds. Because sketches also have much information not included in the completed piece, it is difficult to analyze the chronologies in a sketch.

In addition to this, it is also difficult to take an overview of the changes of the sound with a sketch analysis. As mentioned in Chapter 1, it is necessary to comprehend the precise sound characteristics in computer music [Couprie, 2016]. Most types of computer music do not have any visual materials as a score, and this situation makes it difficult to comprehend the compositional process of computer music. Hence, a novel visual interface to analyze these pieces should be provided.

There are some studies that have traced the compositional process of such pieces. Bernardini [Bernardini, 2016] presents an analysis to clarify the compositional process of Giacinto Scelsi's tape music. This research clarifies that the compositional process counted the number of stops and restarts while recording a composition. It provides another approach for studying the compositional process although the method still depends on the physical materials (see Figure 3.13).

In the next chapter, a visual analysis system—*Spectrail*—that can be used to analyze the compositional process will be proposed.

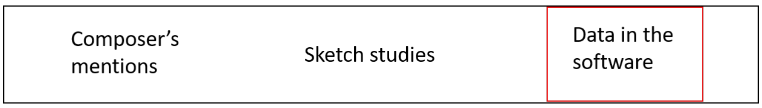


Figure 3.13: Utilizing data in the software. It is important to visually analyze the data in the software to see the compositional process. The data will be one of the main resources to study the compositional process in music

Chapter 4

A Visual Analysis Approach: *Spectrail*

In Chapter 3, the possibilities and limitations of sketch studies were discussed. This chapter provides an outline of the system, *Spectrail* (see Figure 4.1(a)), which visualizes the compositional process with AudioSculpt, a system for sound resynthesis. *Spectrail* takes a sound description interchange format (SDIF) file and audio files extracted from AudioSculpt to analyze the compositional processes. This chapter introduces *Spectrail*'s visualization substrate and various functions for the visual analysis of the compositional process with reference to "TTT (type-by-task taxonomy)" advocated by Shneiderman [Shneiderman, 1996].

4.1 System Outline

This section provides an overview of *Spectrail*, which has been developed as an accompanying system for AudioSculpt, a system for sound resynthesis.

The input/output relationship between AudioSculpt and *Spectrail* is shown in Figure 4.1. The target software of *Spectrail* is AudioSculpt [AudioSculpt, 2019], which was developed by the IRCAM (see Figure 4.1(c)).

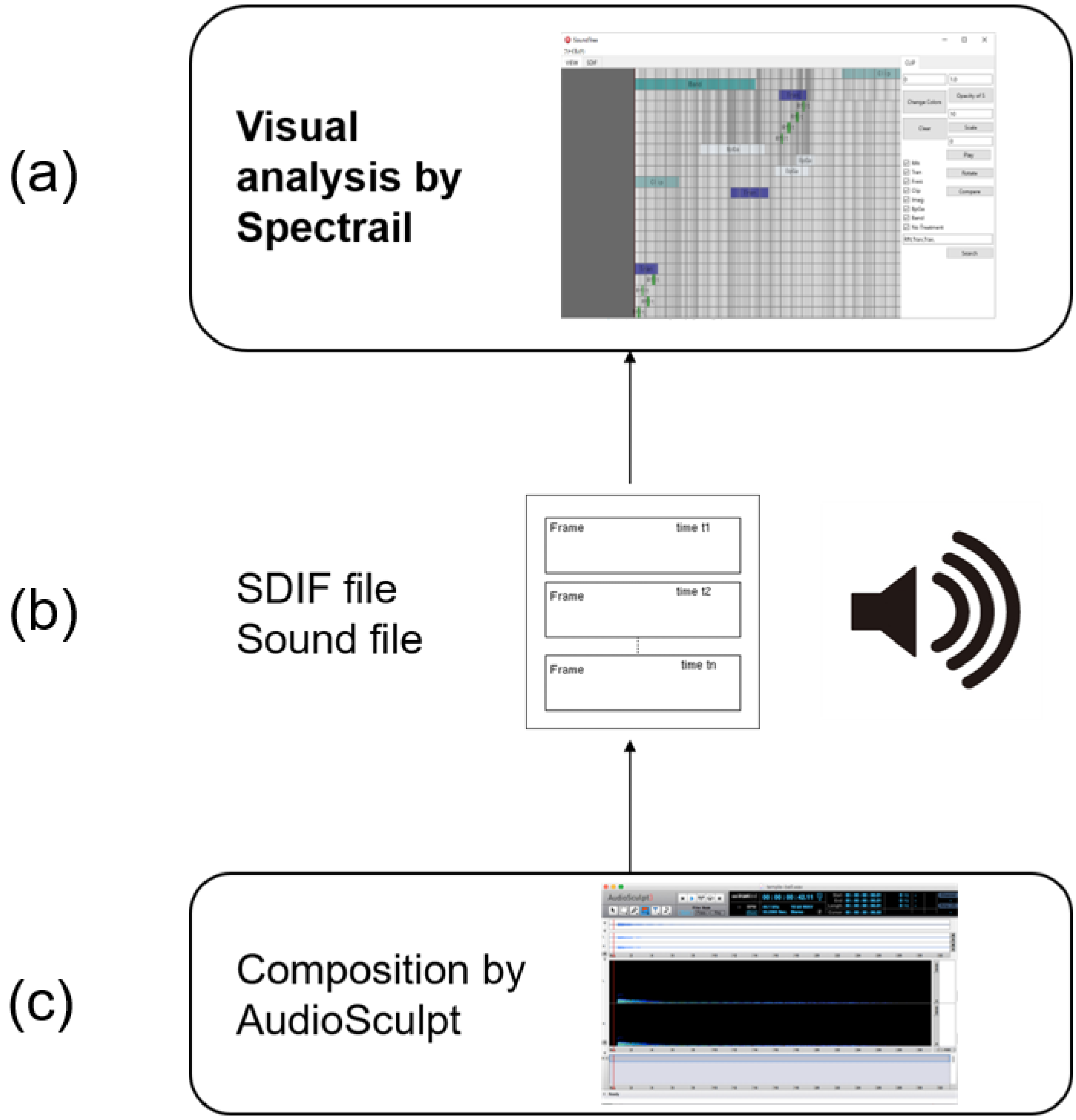


Figure 4.1: Data flow in *Spectrail*. As its input, *Spectrail* takes an SDIF file and audio files from AudioSculpt.

4.1.1 Requirement analysis

The requirements for *Spectrail* were analyzed according to the limitations of sketch studies mentioned in Chapter 3 and the objectives of the compositional process study.

The objective of studying compositional processes is twofold. The first is to clarify how the musical piece was composed. The second is to analyze how the sound has changed. To satisfy these two objectives, the following five requirements were defined when designing *Spectrail*:

1. To clarify the order of the manipulations in the compositional process;
2. To find how the composer chooses the treatments;
3. To analyze where the composer is focusing;
4. To find how the composer makes sections; and
5. To track the change of the sounds during the composition.

Figure 4.2: A screenshot of AudioSculpt interface with treatments

4.1.2 Object software: AudioSculpt

The current research focuses on the process of sound resynthesis using AudioSculpt, and *Spectrail* takes full advantage of the functions "treatment" of the system. Among many software systems that enable spectral analysis and synthesis [Audacity, 2019] [Adobe, 2019], AudioSculpt has many more functions, including `Transposition`, which allows a composer to transpose a sound without changing its length and losing sound quality, and `Formant Filter`, which creates a series of second-order resonant band filters.

Through the use of AudioSculpt, composers can "sculpt" a sound by directly editing its spectral contents. The interface of AudioSculpt displays the spectrogram of a sound, and this system allows the users to visually process the sounds by means of their spectrogram, returning an up-to-date result of the spectral analysis of the sounds, as shown in Figure 4.2.

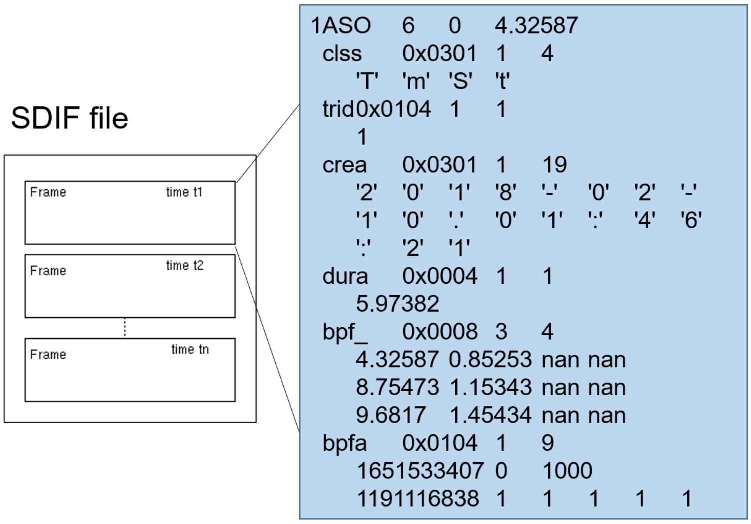


Figure 4.3: Internal format of an SDIF file

4.1.3 Input data

Spectrail takes two types of input data, as shown in Figure 4.1(b). One is an SDIF [SDIF, 2019] file, which is an intermediate product of AudioSculpt, and the other is audio files. The SDIF was co-developed by the IRCAM and the University of California, Berkeley and includes sound processing data in the sequence of the sound. An SDIF file includes data that show the order of sound processing, as shown in Figure 4.3. Numerous types of data items on the sound synthesis, the data on timestamps, the types, and the durations of treatments are utilized in *Spectrail*.

Table 4.1: List of treatments and their abbreviations

Name of Treatment	Abbreviation
`Transposition`	Tran
`Time Stretch`	Tmst
`Formant Filter`	Form
`Gain Envelope`	BpGa
`Rectangular Surface Filter`	Rflt
`Spectral Clipping Filter`	Clip
`Harmonic Filter`	Gsim
`Freeze`	Frze
`Reverse/Repeat`	Revs
`Image Filter`	Imag
`Spectral Break Point`	Brkp
`Filter from Analysis (Surface Filter)`	Surf
`Multi Band Filter`	Band

4.1.4 Spatial substrate and users of *Spectrail*

After extracting the input SDIF file, each of the treatments is represented by a colored block. These are referred to as the "treatment blocks." Each of the treatment blocks has its own abbreviation denoted by unique four characters, as listed in Table 4.1. The color of a treatment block represents its type. The left side of each block corresponds to the start time of the treatment of the sound material, while the width of a block is the duration of the treatment. In the system, the history of creating sounds is represented by a vertical sequence of treatment blocks. In addition, spectrograms may be superimposed on demand on the back of these treatment blocks.

The primarily targeted users of *Spectrail* are composers and musicologists. The approach of *Spectrail* is intended to pave the way for composers to analyze, develop, and share their compositional methodologies. Composers can review and manage their compositional processes after the compositions. On the other hand, musicologists may use *Spectrail* independently for analyzing the compositional processes of the musical pieces composed by composers.

4.2 Functions for Analysis

The proposed system provides multiple functions on the spatial substrate when trying to study the compositional process and the requirements of visualization advocated in Shneiderman's TTT [Shneiderman, 1996].

Composers target the creation of new unique sounds with spectral analysis, whereas musicologists aim to analyze compositional processes. The difference is described by giving expositions of eight functions according to the classification proposed by Shneiderman. Note that these functions meet the above five requirements mentioned in Section 4.1.1.

1) *Overview* (R1, R3, R4): *Spectrail* provides an overview of the history of a computer usage by a composer. This view makes it possible for both composers and researchers to understand the complete picture of the sound-processing history. In the overview mode, spectrograms are averaged and gray-scaled to avoid visual clutter.

In addition, users can see the time distance between the treatments in the selected date,

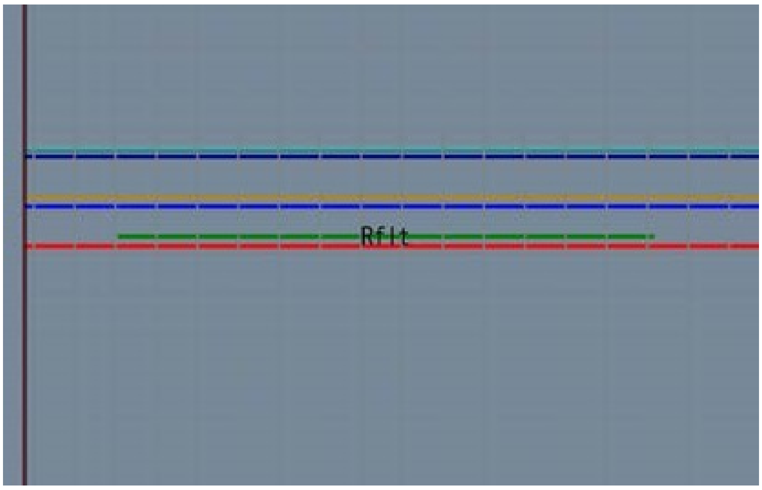


Figure 4.4: Visualizing the time intervals between treatments as the distances between treatment blocks. The history of the composition is represented from the bottom-up

as shown in Figure 4.4, where the vertical axis shows the time distance between the treatments and the horizontal the time of the sound. Users can see where the composer took his or her time when composing the piece.

2) *Zoom* (R5): The zoom mode provides the details of the spectrograms. The spectrograms are depicted with full colors. Yellow shows the loudest sound, and the sound will be weaker if the colors turn to red or blue. Because the spectrograms were adopted to visualize stereo-audio files, the upper half of each spectrogram corresponds to the right channel and the lower half to the left channel. The system can also locate a particular timing of the sound to extract some treatments provided at the same timing. Both composers and researchers can understand the effects of certain treatments at the specified timing.

3) *Filter* (R2): *Spectrail* can filter some particular treatment blocks. Users can use this function by manipulating the corresponding checkboxes. This function makes it possible for composers to consider their styles of sound processing, while musicologists can understand the unique treatments of the composer and comprehend the history of sound processing in more depth.

In the information visualization field, a horizontal axis is normally used for elapsed time.

Figure 4.5: Visualizing the difference in values for the sound descriptors

In this system, users can interchange these two axes depending on which axis the user is focusing on.

4) *Extract* (R3): The system enables the extraction of some processing patterns that the composer uses frequently. Users can understand which treatment pattern is mainly used by the composer.

5) *View Relations* (R5): Another function is provided to show the mutual relationships among the treatments, for example, whether a pattern of treatments shown in one place can also be observed in another. If users input the data created by a different composer, they will be able to understand the history of the sound processing of other composers and compare their processes.

To analyze the compositional process in more depth, the visualization results to visualize the sound changes are created. Audio descriptor is a or multidimensional parameter(s) characterizing a particular aspect of the sound signal, reducing a particular dimension of this signal to one or more digital parameters.

The system provides a function to observe the changes in sonorities by calculating the feature quantity, mel-frequency cepstrum coefficients (MFCC), of each sound created by sound processing and plotting the results on a graph. The horizontal axis of this graph is the progress of sound processing, while the vertical axis is the difference between the MFCC and the original sound. The differences are based on the distance matrices of the MFCC. If users click the vertices of the graph, the corresponding treatment block will be highlghted.

To describe sound changes, *Spectrail* has a function to visualize the sound changes by computing sound descriptors. Pitch, loudness, and timbre are significant aspects [Giannakis and Smith, 2012], and noisiness is also an important perceptual element [Berthaut et al., 2010] for describing the features in computer music. In addition, it is important to see how the harmonics are changed with treatments because they are drastically changed through some treatments, such as `Clip` filters and `Bandpass Filters`. According to this, five kinds of sound descriptors were chosen for *Spectrail* to visualize the sound changes, as shown in Figure 4.5. The descriptors include pitch (depicted with blue), spectral centroid (red), spectral spread (green), spectral flatness (yellow), and loudness (purple). This visualization shows the difference of the sound descriptors between the steps of the composition. The saturation of each square represents the difference in the values for the sound descriptors. The larger the difference is, the higher the saturation will be.

Spectral centroid, spread, and flatness are spectral features that describe the characteristics of the spectrum. Previous studies have shown that the spectral centroid, spectral spread, and spectral flatness correspond to the brightness of the sound, the distribution of the spectrum, and the noisiness of the sound, respectively [Peeters, 2004]. The difference of the sound descriptors is calculated as follows:

- 1. Divide the audio into some small frames for 0.05 seconds;
- 2. Compute sound descriptors for each frame;
- 3. Extract the different frames; and
- 4. Compute the means of the extracted sound frames.

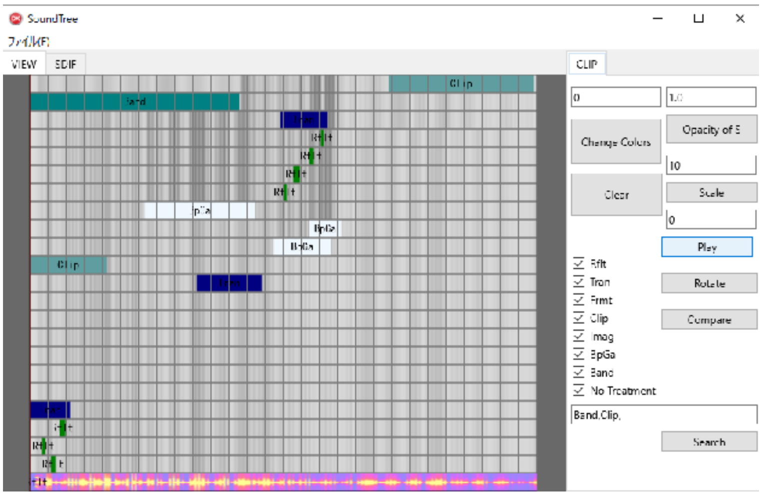


Figure 4.6: Listening to the sound in each treatment

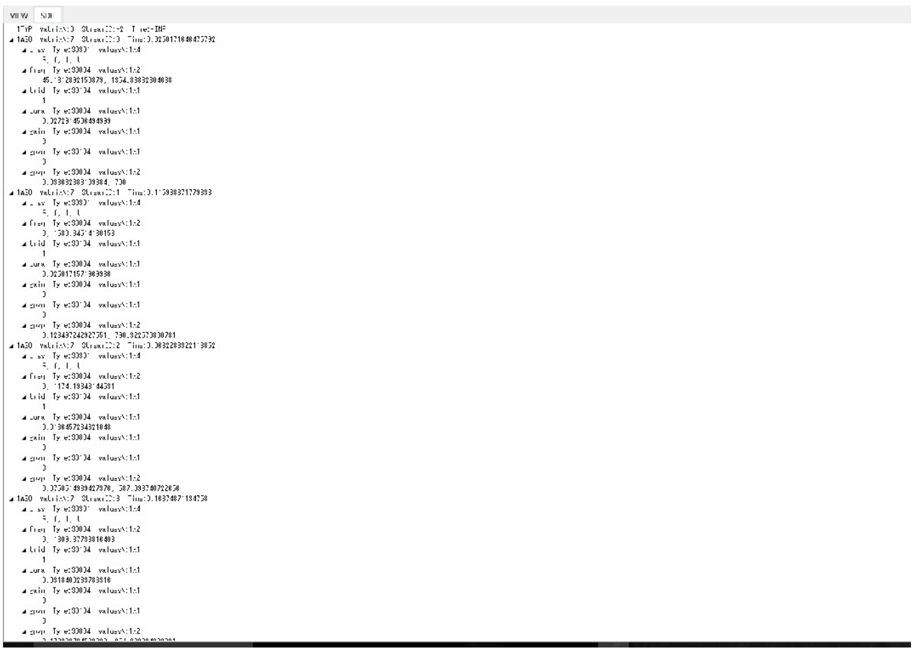


Figure 4.7: Details of the SDIF data

6) *Details on Demand* (R2, R5): More functions are available that can allow users to better understand the transformation of the sound. Users can check the effects of each treatment on the spectrogram. In addition, users can listen to an actual sound that reflects the effects provided by the treatments. While the sound is playing, the color of the spectrogram corresponding to the playing sound turns to a full color, as shown in Figure 4.6. Users can also see the detailed data in SDIF files (see Figure 4.7).

4.3 Implementation

The current prototype system has been developed with Delphi and is running on the following environment:

- CPU: Intel Core i7 2.40 GHz
- GPU: Intel HD Graphics 5500
- Memory: 8 GB
- OS: Microsoft Windows 10's 64-bit.

Figure 4.8 (a) depicts the history of processing the first 2.788 seconds of a recorded temple bell sound with a total of twenty-three treatments. Furthermore, the sound changes drastically in the latter part of this sound processing because the colors of the spectrograms turn darker.

Figure 4.8 (b) illustrates a two-times wider view of the highlighted part of the left figure. Users can take a closer look at the transformation of the spectrogram reflecting the effect of filters. This result shows that the high-pitched harmonies are deleted by the `Bandpass Filter`.

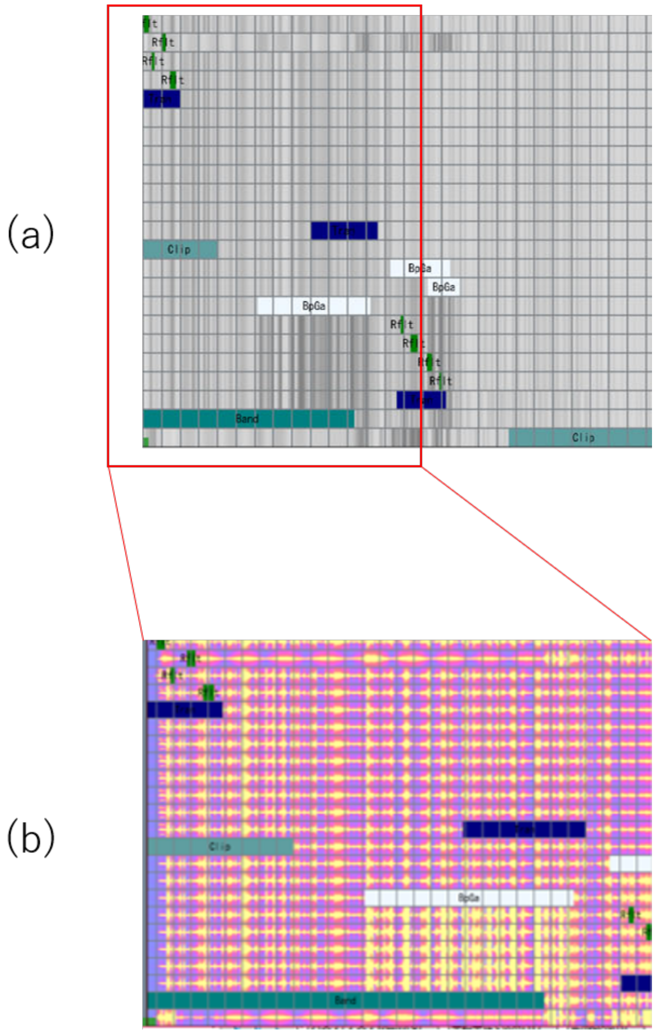


Figure 4.8: Overview and zoom

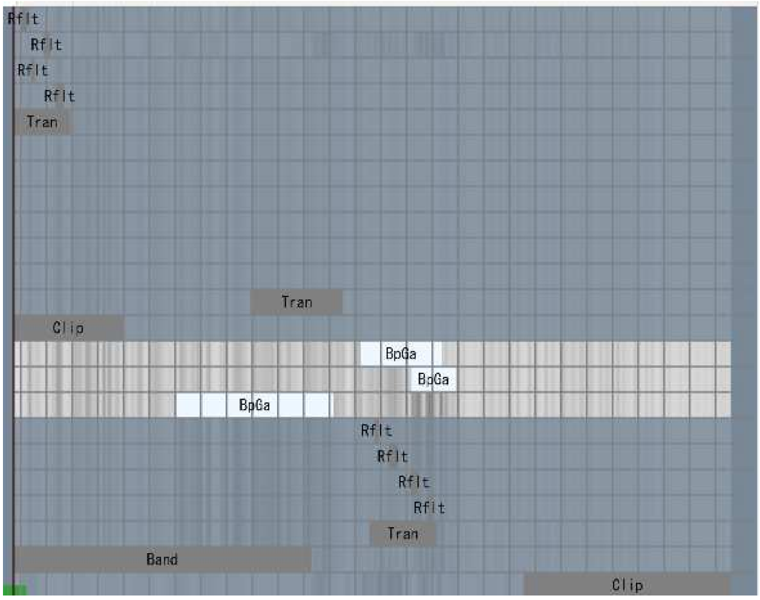


Figure 4.9: Filter mode

Figure 4.9 represents the portion extracted where the `Gain Filters` were applied. The `Gain Filter` strengthens the sound. It turns out that `Gain Filters` were used intensively in the middle of the sound, implying that the sound was emphasized there.

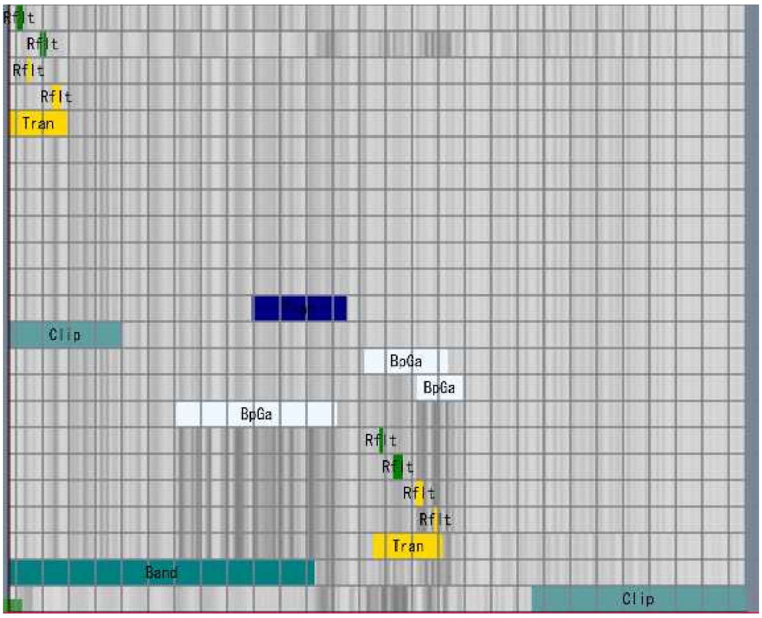


Figure 4.10: Extracts

Figure 4.10 shows that `Rectangular Surface Filters` and `Transposition` are extracted by using the "Extract" function to choose particular treatment patterns. The `Rectangular Surface Filter` is a distinguishing filter of AudioSculpt, while `Transposition` is one of the most effective treatments by which the pitch of the sound can be changed. With this function, a pattern of treatment—using `Transposition` after utilizing several `Rectangular Surface Filters`—is repeated in the first and latter parts of the processing. Users can see that this pattern is a unique feature of this specific sound processing.

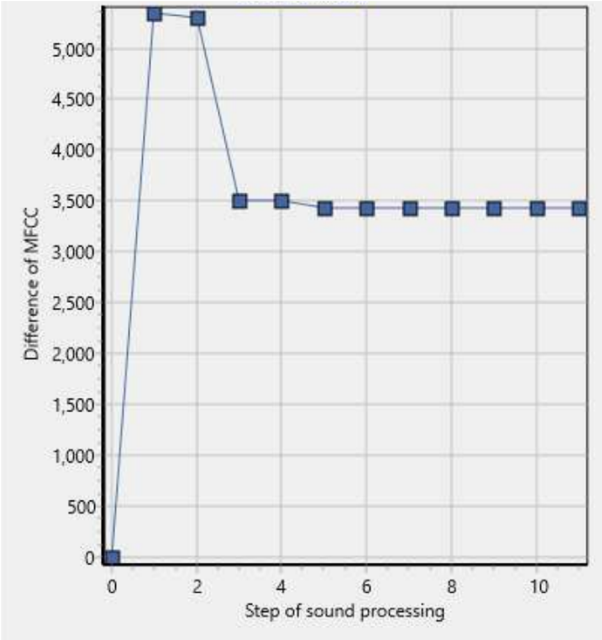


Figure 4.11: Change in the MFCC

Figure 4.11 shows the changes in sonority as the sound was processed. The data on the vertical axis represent the differences of the MFCC from the original sound. The slope of this graph becomes steeper in the first part of the compositional process. This provides proof that the sound changed drastically by the first treatment.

4.4 Limitations of *Spectrail*

This visualization approach can be applied to other software systems such as OpenMusic [OpenMusic, 2019]. However, *Spectrail* needs to be extended so that it could be used by these other systems. It is possible to visualize sound changes if the sound files can be obtained. However, the current system does not simultaneously visualize multiple compositional processes when composers use numerous software systems.

In addition, *Spectrail* can visualize the process of manipulating software systems for compositions; however, the current system does not deal with processes such as making a concept of the composition because the process to create ideas contains a number of ambiguous points.

Chapter 5

Empirical Evaluation of *Spectrail*

In Chapter 4, the visual analysis system, *Spectrail*, was presented. In this chapter, the results of an empirical experiment for *Spectrail* are described. The evaluation experiment was conducted with four composers who studied at the IRCAM. The four composers who participated in the empirical evaluation experiment are as follows:

- **Composer A**: Luciano Leite Barbosa;
- **Composer B**: Alessandro Ratoci;
- **Composer C**: Frédéric Le Bel; and
- **Composer D**: Maurilio Cacciatore.

The piece by **Composer D** is titled *Tape on Piano—Sketch for Audio Sculpt*.

The actual analysis was performed with two aspects—one is how the composer used computers, and the other is how he changed the sounds during the composition. The evaluation also includes interviews with the composers.

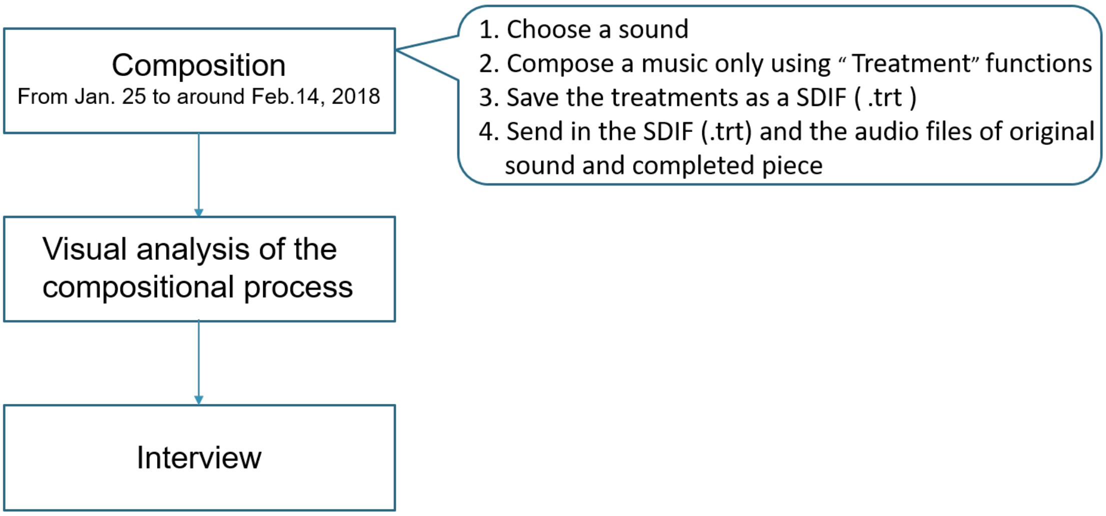


Figure 5.1: Evaluation method

5.1 Outline of Empirical Experiment

This section explains the procedure that was used for the evaluation (see Figure 5.1). Four composers who studied at the IRCAM were asked to create a piece with AudioSculpt version 3.4.6. First, the composers chose a sound file and then composed a piece only using the “treatment” functions of AudioSculpt. The original sound file for the composition was chosen by each of the composers. After the composition, they saved the data on treatments as an SDIF file and submitted it together with the audio files of the original sound and completed piece.

The compositional processes are visually analyzed using the SDIF files. The visualization results are evaluated by also using the composers' interviews. The interviews aim at exploring how the visualization results correspond to the actual compositions and also verifying how the composers used *Spectrail* for their compositions.

Table 5.1: Basic information about each composition

Composer	Number of Treatments	Number of Rendering	Original Sound File
Composer A	52	6	Guitar
Composer B	100	3	Guitar
Composer C	8	0	Continuous sound
Composer D	47	0	VHS tape on piano strings, ebow on piano, metal objects

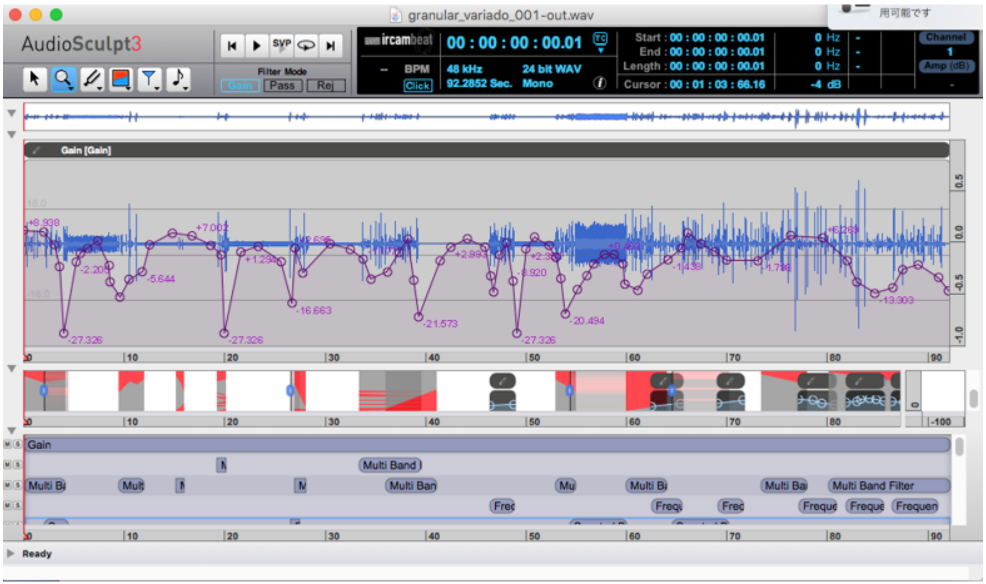
5.2 Subjects

All four composers studied in the CURSUS program at the IRCAM, which includes some classes for using IRCAM software (AudioSculpt, Max, OpenMusic, and the like). They had gone through abundant experiments with compositions using AudioSculpt. In this section, the basic information of the provided pieces is presented (see Table 5.1).

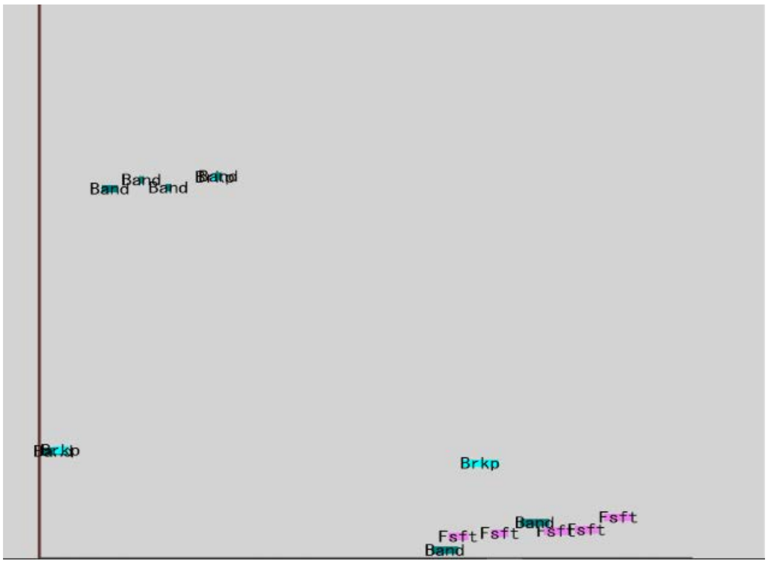
5.2.1 Composer A

Composer A used a guitar performance with pizzicato, which continues for two minutes as the original sound of the piece. The total number of treatments is fifty-four, and he rendered the piece six times during the composition. Before the second rendering, **Composer A** used many treatments, and after that, he used only a few. Figure 5.2(a) shows a screenshot of the AudioSculpt interface after the first rendering.

With the visualization shown in Figure 5.2(b), it can be seen that he first focused on the latter part of the original sound and mainly used frequency shifts, and then, he composed the former part mainly with `Multi Band` filters.



(a) A screenshot of the AudioSculpt interface after the composition



(b) Visualization of compositional process

Figure 5.2: Compositional process of **Composer A**

5.2.2 Composer B

Figure 5.3(a) shows the interface of AudioSculpt after the composition by **Composer B**.

The visualization results of **Composer B**'s compositional process is shown in Figure 5.3(b). It can be estimated that before the first rendering, he focused on the part between twenty-five minutes and thirty-two seconds of the original sound file, before the second rendering between zero and seventy-seven seconds, and before the third rendering from twenty-seven to one hundred and thirty-one seconds.

5.2.3 Composer C

Figure 5.4(a) shows the AudioSculpt interface after the composition by **Composer C**. He did not render the piece before completing it and composed only seven manipulations.

Composer C first used `TimeStretches` for an entire sound. **Composer C** used the `Transposition` and `Gain` filter for the entire sound, and then used some `Clip` filter, as shown in Figure 5.4(b).

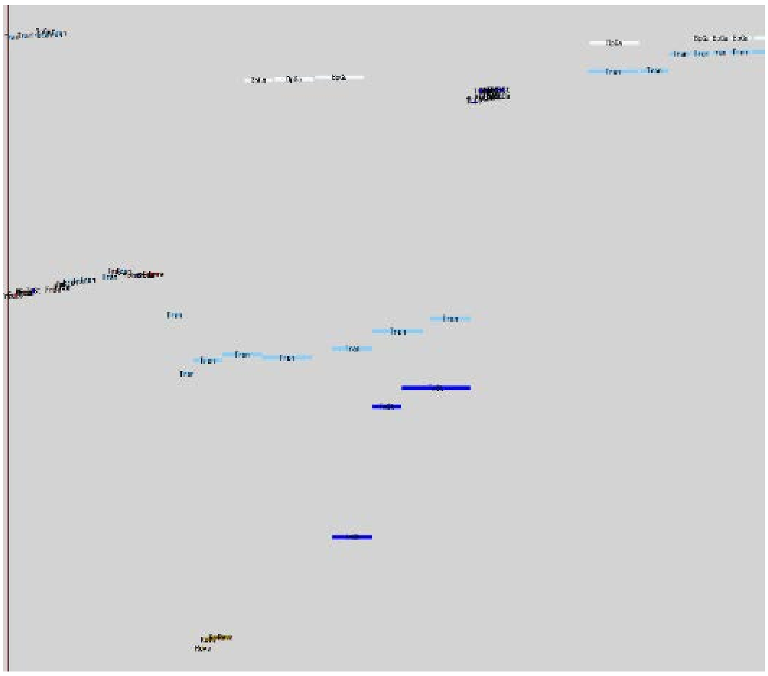
5.2.4 Composer D

Composer D used forty-nine treatments for his composition and did not render the piece before completing. Figure 5.5(a) shows a screenshot of the AudioSculpt interface after a composition by **Composer D**.

According to the visualization, he took time between the eleventh and twelve treatments and between the thirty-sixth and thirty-seventh treatments. The compositional process can be divided into three steps.

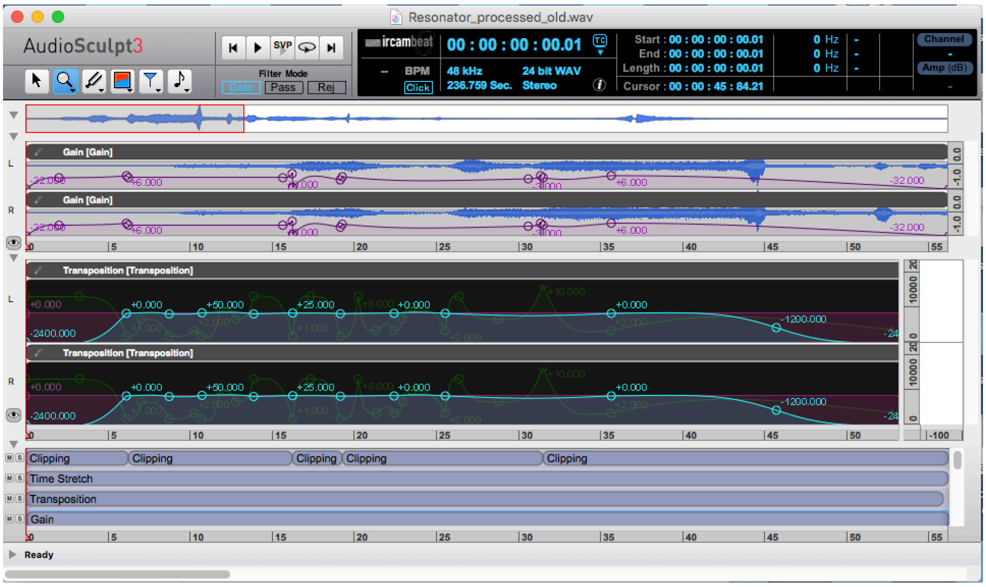
Figure 5.5(b) represents the order of the treatments by **Composer D**. This composer composed the piece from the first to last of the sound file. In the composition, he first used the treatments over the entire original sound, and after that, he used small treatments.

(a) A screenshot of the AudioSculpt interface after the composition

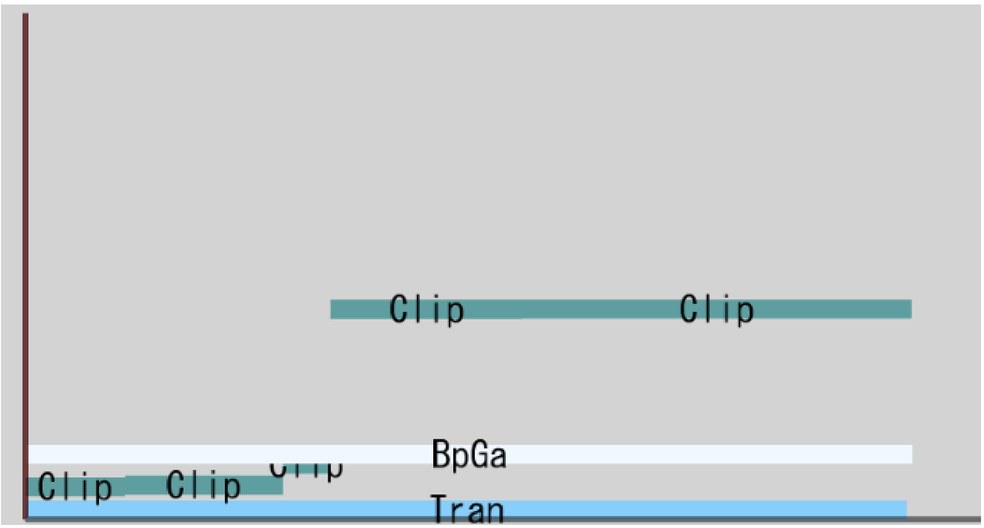


(b) Visualization of compositional process

Figure 5.3: Compositional process of **Composer B**

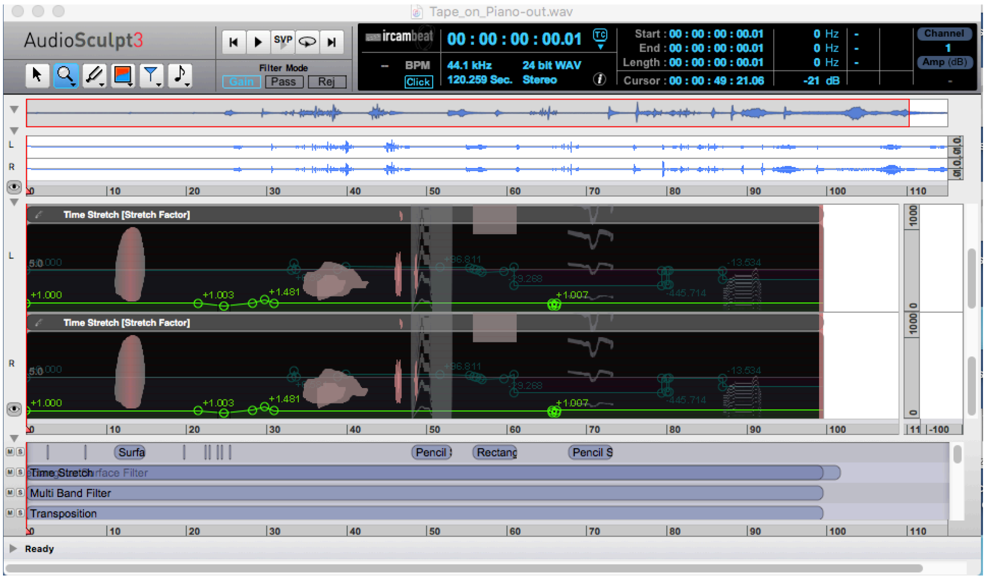


(a) A screenshot of the AudioSculpt interface after the composition

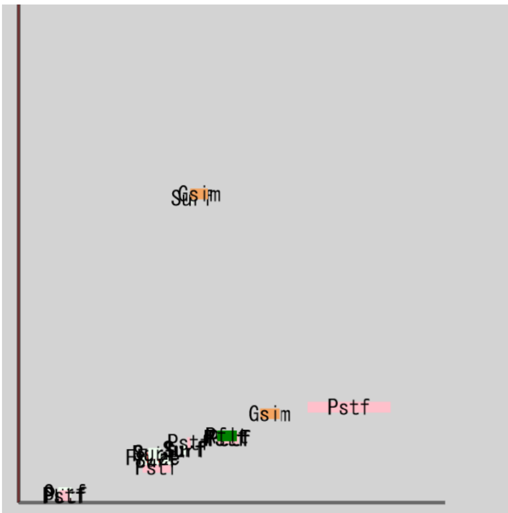


(b) Visualization of compositional process

Figure 5.4: Compositional process of **Composer C**



(a) A screenshot of the AudioSculpt interface after the composition



(b) Visualization of compositional process

Figure 5.5: Compositional process of **Composer D**

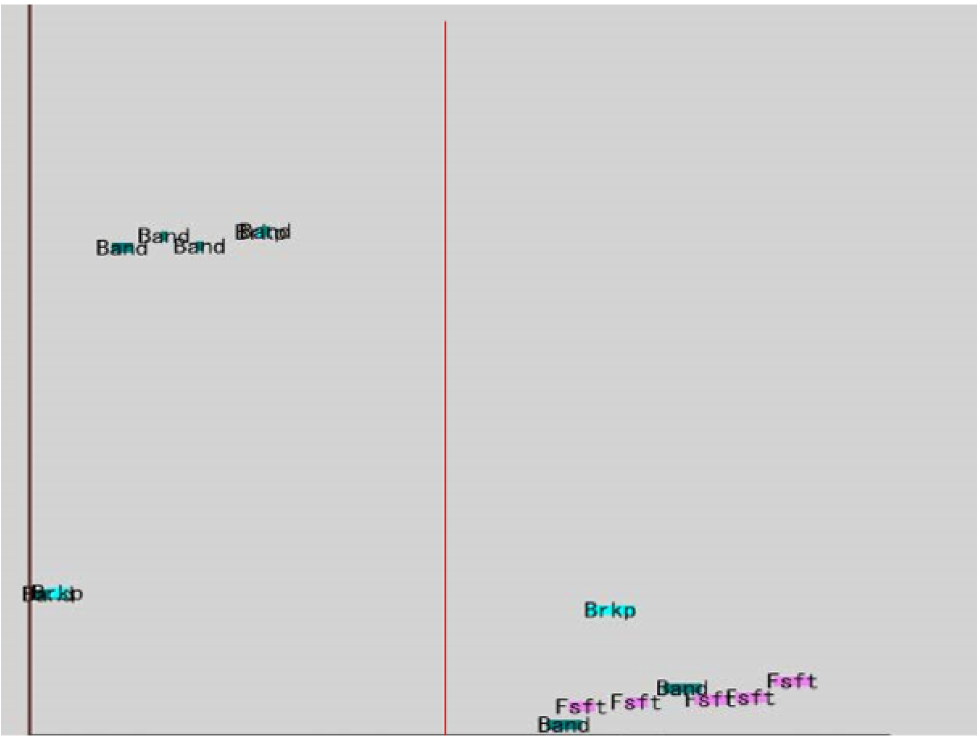
5.3 Visual Analysis

5.3.1 Viewpoints of analysis

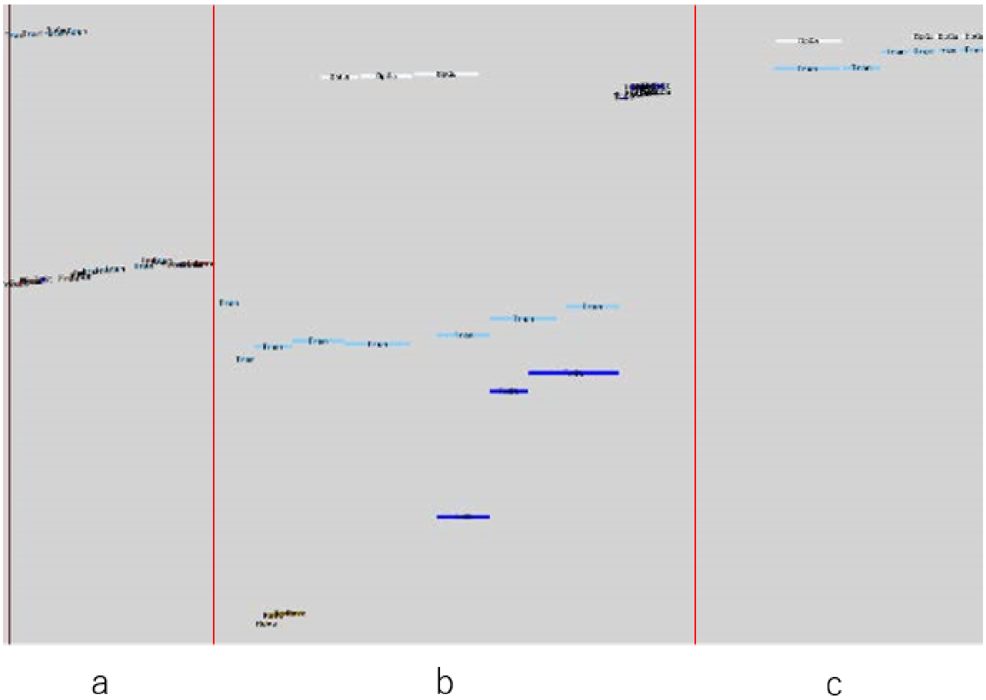
In the current research, the compositional processes are analyzed in two ways. One is how to use the treatments, and the other is how the sound changes in the compositional process. The processes are analyzed by looking at the five requirements mentioned in Chapter 4. Using these viewpoints, where the composer focused in the original sound file. Because many composers create their pieces based on a structure that they have planned before starting to compose, it is important to know where they focused by looking at the original sound file can be found. In addition, repeated treatment patterns indicate the intentions of the composer. From the viewpoint of the acoustic changes, an overview of them in the entire compositional process is provided. Hence, it can be seen how the composer used computers.

5.3.2 Compositional process

Figure 5.6 compares the compositional processes of **Composers A and B**. **Composer A** composed the piece from the end of the original sound file (see Figure 5.6(a)). On the other hand, in the piece by **Composer B**, it is presumed that the composer divided the piece into three parts (a, b, c), which can be seen by the concentration of the treatments. **Composer B** first used many treatments in the middle part (b). After that, the composer composed in the former part (a) and finally added the treatment in the last part (c), as shown in Figure 5.6(b).

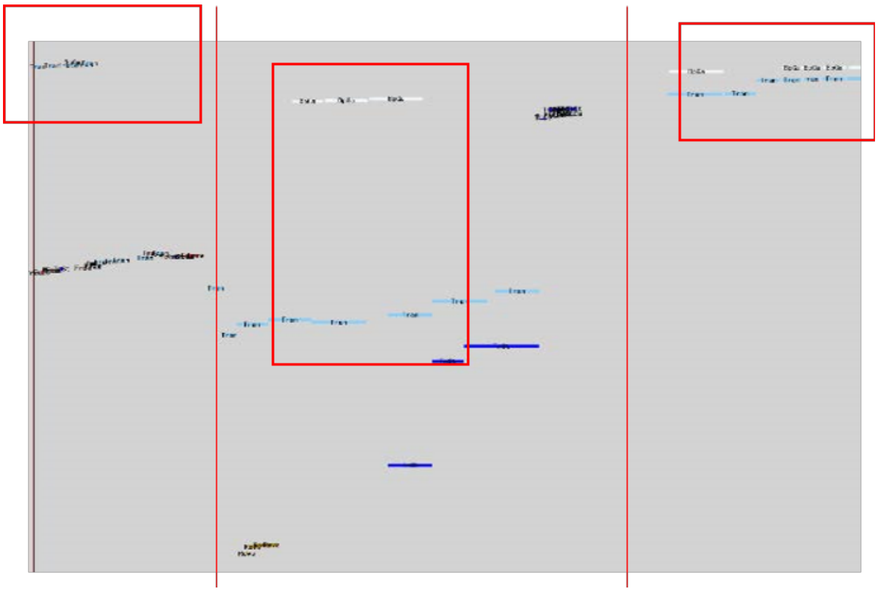


(a) **Composer A**

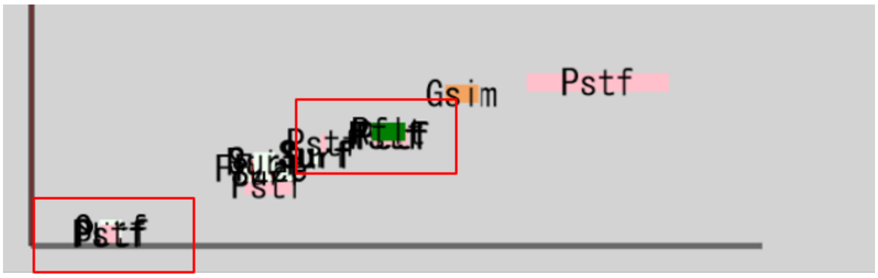


(b) **Composer B**

Figure 5.6: Compositional process of **Composers A** and **B**



(a) **Composer B**

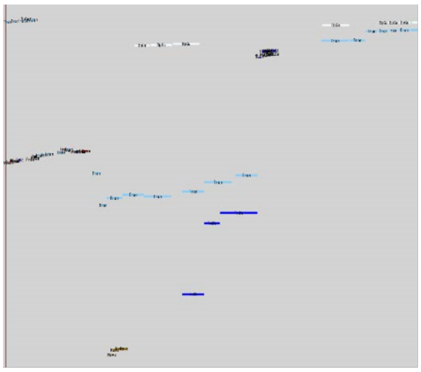


(b) **Composer D**

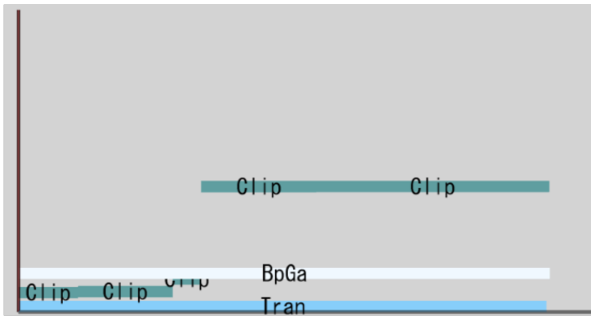
Figure 5.7: Compositional process of **Composers B** and **D**

5.3.3 Repeated pattern

Next, the repeated patterns of the treatments used by the composers are analyzed. It can be seen that the composers were paying attention to the acoustic effects of the repeated treatments. Figure 5.7(a) shows that **Composer B** used a repeated pattern that consists of `Gain` filters after `Transpositions`. This pattern can also be seen in **Composer D**'s compositional process. In addition, **Composer D** repeatedly used `Spectral Snippets` and `Gain` filters, by which the composer could copy and paste the sound spectral to another place, as shown in Figure 5.7(b).



(a) Short treatments



(b) Long treatments

Figure 5.8: Treatments used by **Composers B** and **C**

5.3.4 Length of treatments

Figure 5.8 compares **Composer B**'s and **Composer C**'s treatments in terms of their granularity. These visualization results show that **Composer B** composed the piece using small treatments (see Figure 5.8(a)), whereas **Composer C** used the `Transposition` and `Gain` filters from the beginning to the end of the original sound file (see Figure 5.8(b)). This indicates that **Composer C** composed the piece with two main treatments and that he manipulated the parameters of these treatments.

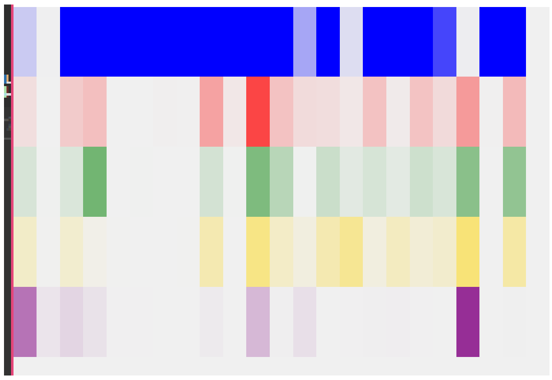
5.3.5 Sound descriptor

Overview

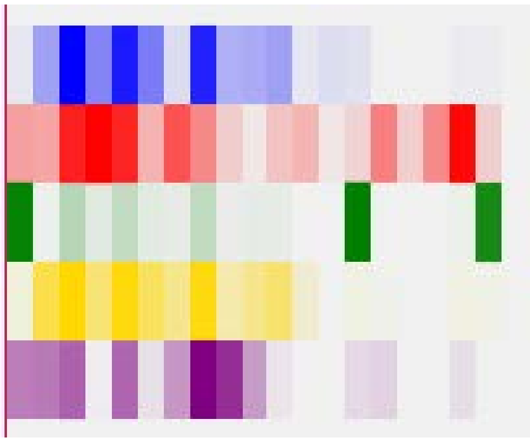
Figure 5.9(a) shows the sound changes in **Composer A**'s compositional process before the first rendering, and Figure 5.9(b) shows this before the second rendering. Before the first rendering, the pitches were notably changed. After the second rendering, all of the spectral features had changed, and then the spectral centroid, which indicates the brightness of the sound, became higher.

Figure 5.10 shows the sound changes in **Composer B**'s compositional process. In the first part of the composition, the treatments changed the pitches and loudness. Then, the spectral centroid and spectral spread were higher, and in the last part of the composition, all of the descriptors had changed drastically.

Figure 5.11 shows the sound changes in **Composer D**'s compositional process. It shows that some of the treatments in the middle part of the composition changed the original sound drastically.



(a) Before the first rendering



(b) Before the second rendering

Figure 5.9: Change of the sound descriptors in the compositional process of **Composer A**

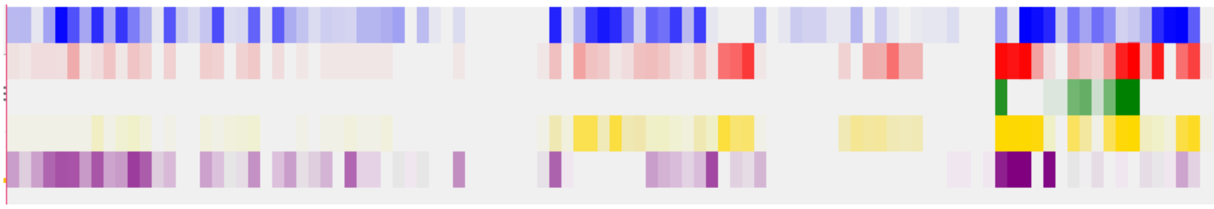


Figure 5.10: Change of the sound descriptors in the compositional process of **Composer B**.

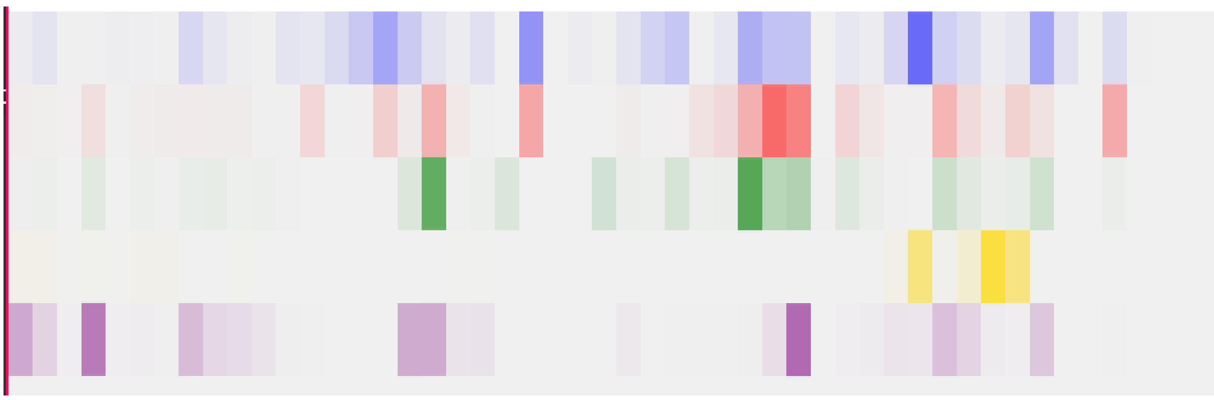


Figure 5.11: Change of the sound descriptors in the compositional process of **Composer D**

Relation between the treatments and sound changes

The relation between the sound descriptors and treatments is discussed by taking **Composer B**'s compositional process as an example. First, the pitch and loudness were changed drastically in the first part of the composition. Figure 5.12 shows that the changes come from `Transposition` and `TimeStretch`.

Figure 5.13 shows where the differences of the sound descriptors are small, and it corresponds to where the composer used short treatments. In addition, in Figure 5.14, the descriptors are changed drastically where the repeated pattern "`Transposition` + `Gain`" was used.

5.3.6 Summary of the analysis

First, the compositional processes of the four composers are visualized, and it turns out that the composers used the same treatments repeatedly. For example, **Composer A** used `Frequency Shift` in the later part of the original sound file, while **Composer D** repeatedly utilized `Spectral Snippet`.

Furthermore, the granularity of the treatments and the order of where the composer focused in the original sound are different. **Composers C** and **D** used short treatments after applying long ones. On the other hand, **Composers A** and **B** created their pieces using many short treatments. Also, **Composers B** and **C** were applying treatments from the beginning of the original sound file.

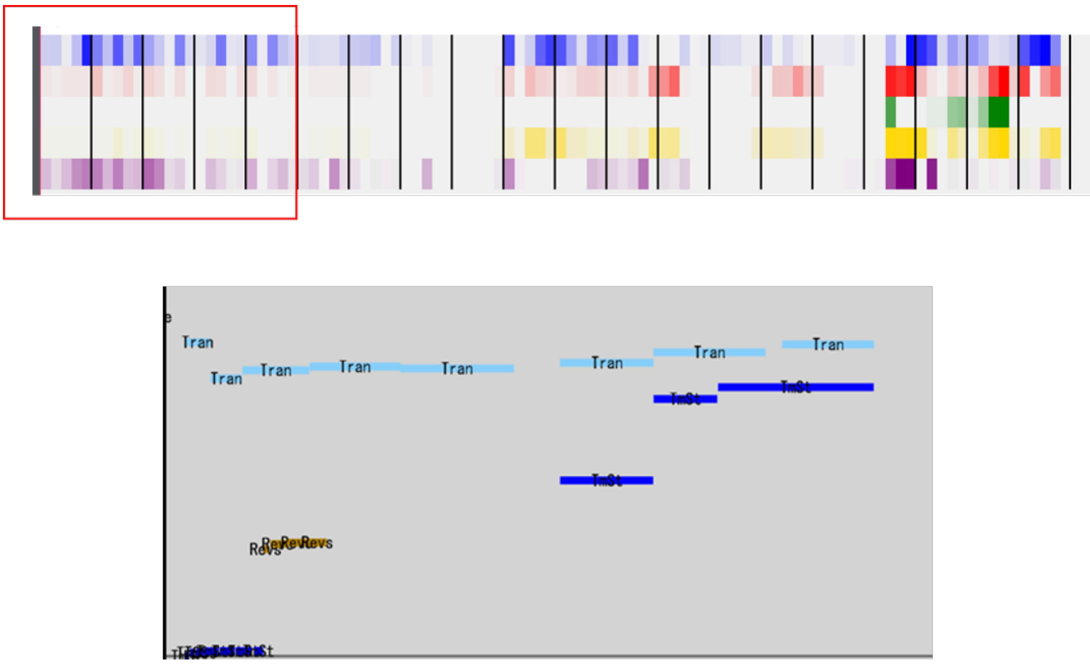


Figure 5.12: Pitch and loudness are changed by `Transposition` and `TimeStretch`

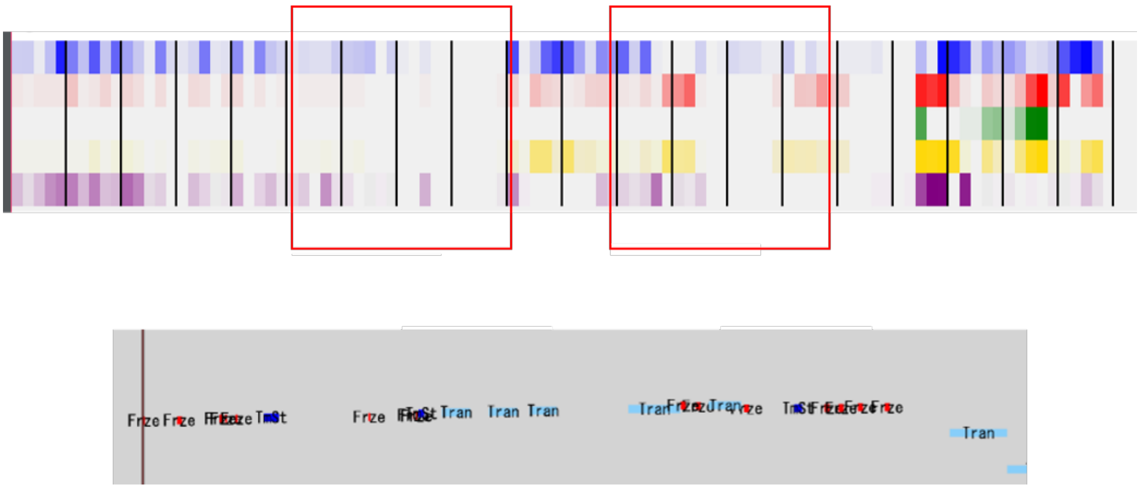


Figure 5.13: The differences of the sound descriptors are small, and it corresponds to where the composer uses tiny short treatments

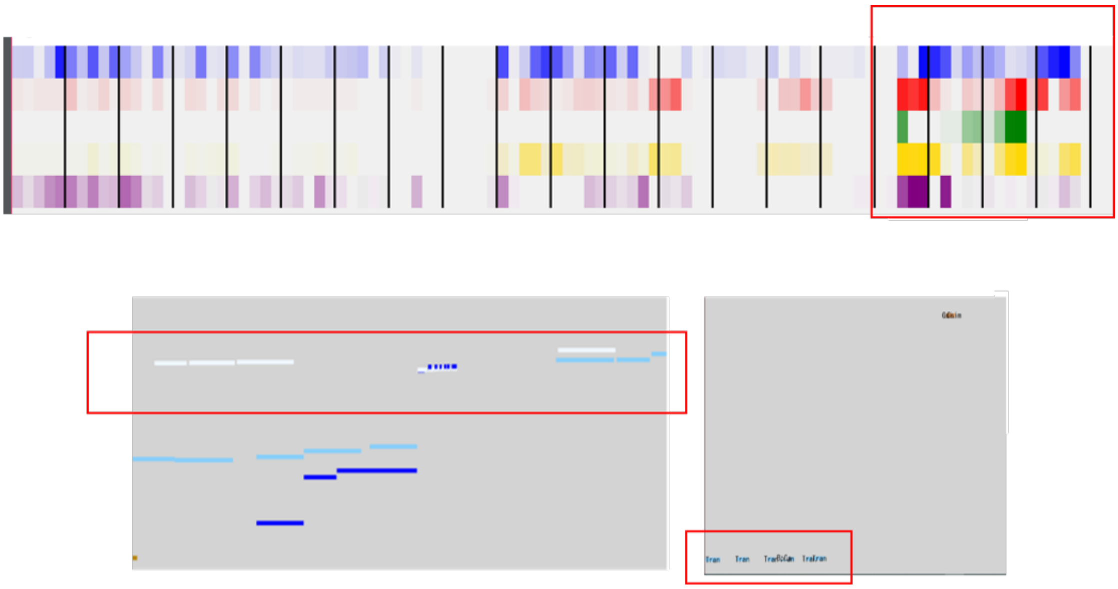


Figure 5.14: The descriptors are changed drastically where the repeated pattern “Transposition + Gain” is used

5.4 Interview with the Composers

To evaluate the effectiveness of the *Spectrail* system, interviews were conducted with **Composers A** and **B** because these two composers used similar original sound files for their compositions.

The aim of this interview was to verify the visualization results and to clarify how the composers may use the visualization results for future compositions. First, the two composers were asked a question about their compositional processes to identify common points and differences between the visualization and the compositional process. Then, to better understand how *Spectrail* works in comprehending the compositional process, along with the future direction that *Spectrail* should take when applying their compositions, the composers were asked whether they made sketches or not. The following parts describe the remarks provided by the composers.

5.4.1 Original sound file

First, the two composers were asked why they chose the sound file. **Composer A** answered that the reason behind it was because he was interested in the contrast between improvisation and composition, and **Composer B** also tried to make a contrast between electronic and acoustic sounds. Because the granularity of the original sound was high, **Composer B** was able to make a rhythmic contrast by stretching the sound.

5.4.2 Evaluation of the visual analysis

Composer A first used `Freezes` to make contrasts because the granularity of the original sound was high. Then, the composer selected a `Gain` filter because he intended to emphasize (and normalize) the details of the sounds using the filters. After that, the composer utilized some frequency shifts. In the original sound, the frequencies were uniform, and the treatment made the variations occur in the frequency domain. He also mentioned that he listened to the original sound file many times and then found what he was interested in and processed the sound.

Composer B mentioned that he always composes in three phases for his compositional process. First, he stretched the complex sound and used `Freezes` in the first section of the piece because he was interested in the grains of the frozen sound. Second, he used `Transpositions` to explore the relationship between the electronic and acoustic sound. Then, the composer used `Transpositions` and transformed the sound. Third, he used `Gains` to emphasize the resonances of the sound that was changed with `Transposition`. **Composer B** referred to the sections because he made three sections and a coda in the piece, but **Composer A** denied making any sections.

According to what they said, the order of treatments in the visualization corresponds to their compositional process, and this reflects their intentions for the compositions. **Composer B** made three sections in the piece, which are also shown in the visualization.

5.4.3 Sketches

In the interview, the composers were asked whether they made sketches during the composition. Although **Composer A** sometimes makes sketches for a larger piece, he directly works with a computer when he composes electroacoustic pieces. For the piece, what he wrote on paper is only the types and effects of the treatments, as shown in Figure 5.15.

Composer B does not write sketches nor scores for electroacoustic music, either. Instead, he composes the pieces by taking advantage of spectrograms.

5.4.4 Possibility to use *Spectrail* for compositions

Composer B mentioned that he was interested in the three-staged change of the sound that occurred in the visualization of the sound changes. This visualization allowed him to find that the feature in the sound changes corresponded to his compositional process because he rendered the piece three times.

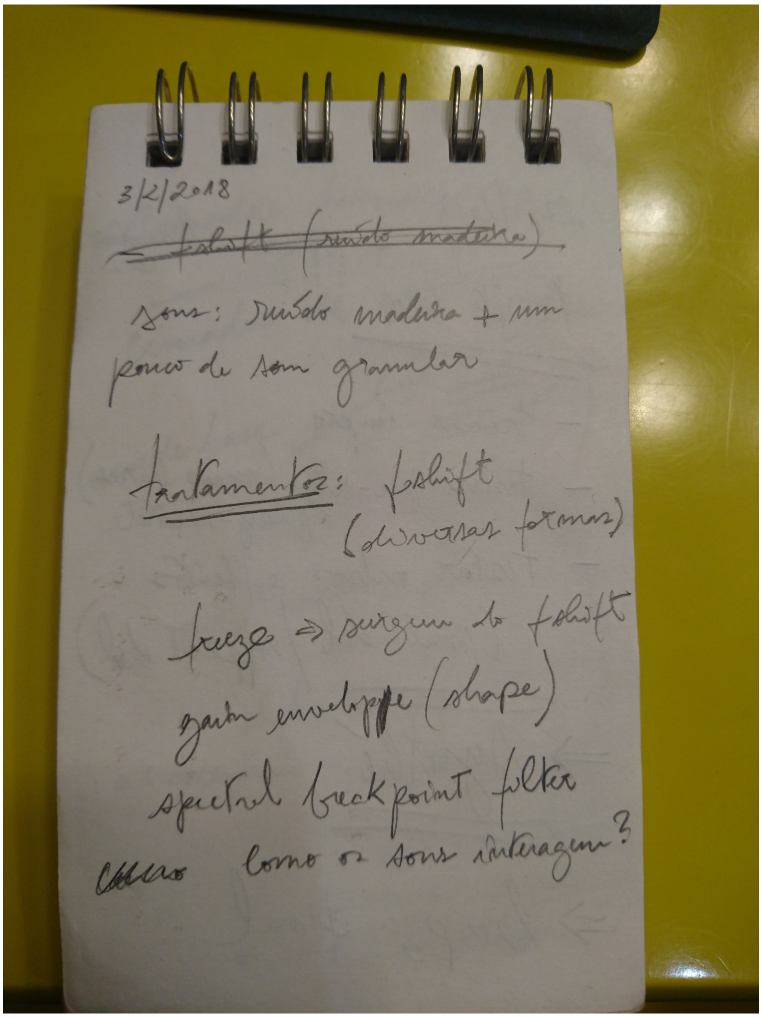


Figure 5.15: Sketch by **Composer A**

5.4.5 Summary and discussions

The interviews show that the visualizations effectively represent the order of using treatments in the compositional processes and that this reflects the composer's intentions in the compositions. This suggests that *Spectrail* could provide a kind of "computational sketch" for creating music, even when composers do not make sketches during composition. The interviews also clarified that *Spectrail* allows composers to find methods of sound processing in their compositions.

Composer A mentioned that he sometimes makes sketches or draws pictures to compose the pieces during his usual compositional process. Composers sometimes draw the structure of the piece as a picture. The pictures indicate the impression of the sounds and/or the overall structure of the piece. However, the pictures do not show how the composer created these kinds of sounds. *Spectrail* would provide additional information on how the composer created the sounds during their composition.

In Chapter 3, it was clarified that Mayuzumi composed the two symphonies from the same original sound file through the use of different methods. In this case, *Spectrail* would be useful to represent the difference in the methods of composition used among the composers.

Chapter 6

Conclusion

In the present thesis, a sketch study for analyzing Mayuzumi's compositional process and a visual analysis approach—*Spectrail*—were presented. Throughout these studies, the possibility of a complementary analysis using a sketch study and visual analysis was presented, and then, this chapter concludes the thesis and provides future extensions of this research.

6.1 Conclusion

In the current thesis, first, the relation between the two symphonies, *Nirvana Symphony* and *Mandala Symphony*, both composed by Toshiro Mayuzumi, was clarified through a sketch analysis. The case study clarified that both of these symphonies are based on the same temple bell sounds, whereas the compositional method used for each is different. In the *Nirvana Symphony*, Mayuzumi used the pitch relations between the overtones of temple bell sounds. On the other hand, in the *Mandala Symphony*, he made a tone series based on the number of tones with the same pitch names. Based on the results, the possibilities and limitations of sketch studies were discussed. Sketch studies can clarify what a composer did in the compositional process and show his/her trials not included in the completed piece. However, it is difficult to analyze the chronologies in a sketch and to trace the changes in sounds during the composition.

Based on the above sketch analysis, *Spectrail*, a system that accompanies AudioSculpt to trace the process of creating sounds was presented. On the pixel-oriented spatial substrate, the system made it possible to visually analyze the compositional processes by using dedicated interactive manipulations. This system allowed the users to examine the compositional processes of musical pieces and can also be regarded as an initial attempt at managing the provenance of time-series events in the music visualization field.

The current study provided an overview of sound processing history with the transformations of spectrograms and extracting unique sound processing patterns in the process. The study also showed how the sound changed during the compositional process. By using *Spectrail*, composers can better understand their own compositional processes in an objective way, and musicologists can comprehend and visually analyze the history of sound processing.

To evaluate the system, empirical experiments were conducted with four composers from the IRCAM, and the study included interviews with two composers. According to the results, the styles of the composers in the compositional processes could be clarified with the visualization, and it was also shown that the composers did not make sketches in these small pieces. In the interview, one of the composers mentioned that he could see the sound changes and understand the feature of his compositional process using the visualization. This suggested that the visualization was able to achieve an essential goal of music visualization.

This research showed that it is effective to visually analyze the process of computer usage in a composition. A visual analysis provided information on the chronological order of the compositional process precisely, giving an overview of the sound changes. In addition, sketch studies would provide additional information on the ideas of the composition and/or the structure of the piece.

6.2 Future Research

To extend the research, it is important to take into account the actual composition and add new data for analysis.

6.2.1 Taking into account the actual composition

In the experiments, composers composed the pieces using short sound files. However, in a real situation, most composers use longer sounds. In addition, composers do not assume that others will observe their compositional process.

For more detailed studies, *Spectrail* has to be improved to interpret more functions on AudioSculpt because the system currently covers only around twenty percent of the functions provided by AudioSculpt. For example, one of the main functions of AudioSculpt, "partial synthesis," is not treated in the present system.

Furthermore, as pointed out by **Composer B**, *Spectrail* should also be applied to other software, such as OpenMusic [OpenMusic, 2019], which is a well-known software system used by computer music composers. In addition, there are various composition methods, even if just focusing on pieces such as synthesizing with many sin curves. It would also be possible to apply *Spectrail* to a work composed by integrating multiple software systems that appeared in provenance management.

6.2.2 Provenance of analysis by *Spectrail*

The current research showed that the composers used *Spectrail* to look back at their processes after their compositions. However, it would be possible to visualize how musicologists and composers analyze the compositional process by looking at log data of using *Spectrail*. This trial would be helpful in designing future systems for music analysis.

6.2.3 Adding new data for detailed analysis

In the current version of *Spectrail*, composers' behavior between the treatments could not be visualized. In the interview, a composer mentioned that he listened to the sound files many times. This suggests that the number of times when playing the sound with the software is an effective method for understanding the composer's behaviors. In addition, the lifelog data of the composer would have important information in showing what the composer was thinking about during the composition.

6.2.4 Extending to multimodal analysis

In the interview, one of the composers mentioned that he sometimes draws a picture during his composition to understand the structure of the completed piece. In fact, many composers create pictures or some graphs during their composition. This suggests that it would be effective to directly analyze the relationship between these visual materials and the musical pieces through a unified user interface.

Publications

[Publications-1, 2018] Yuriko Takakura, "A Comparison of the Compositional Process between the Nirvana Symphony and the Mandala Symphony: An Analysis of the Campanology Documents," *Ongakugaku*, Vol. 63, No. 2, pp. 61–77, March 2018 (in Japanese).

[Publications-2, 2018] Yuriko Takakura, Masanori Nakayama and Issei Fujishiro, "Visual Analysis for the Compositional Process of Composers in Spectral School," *IIEEJ Transactions on Image Electronics and Visual Computing*, Vol. 6, No. 1, pp. 22–30, June 2018.

[Publications-3, 2017] Yuriko Takakura, "Visual Analysis System for Compositional Processes of Composers in Spectral School," in *Proceedings of the 5th IIEEJ International Workshop on Image Electronics and Visual Computing*, pp. 5B-2-1–5B-2-6, March 2017.

[Publications-4, 2015] Yuriko Takakura, "A Study on ‘Campanology-Effect’ in Nirvana Symphony (1958)— ‘Campanology documents’ in the Collection of the Archives of Modern Japanese Music," in *Proceedings of the Student Session in the 66th National Convention of Japanese Society for Aesthetics*, pp. 63–72, March 2016 (in Japanese).

[Publications-5, 2016] Yuriko Takakura, "The Compositional Process of Nirvana Symphony (1958) and Manda Symphony (1960)—The Comparison Based on ‘Campanology Documents' in the Collection of the Archives of Modern Japanese Music," in *Proceedings of the 36th Conference of the East Japanese Branch of the Musicological Society of Japan*, April 2016 (in Japanese).

[Publications-6, 2017] Yuriko Takakura, Masanori Nakayama, Yuriko Kojima and Issei Fujishiro, "Visual Analysis Approach for the Process of Sound Processing by Composers in Spectral School," in *Proceedings of SIG Technical Reports*, Vol. 2017-MUS115, No. 36, pp. 1–6, June 2017 (in Japanese).

References

[Adler, 1885] G. Adler, "Umfang, Methode und Ziel der Musikwissenschaft," *Vierteljahrschrift für Musikwissenschaft*, vol. 1, pp. 5–20, 1885.

[Adobe, 2019] Adobe, "Adobe Audition CC," [Online]. Available: http://www.adobe.com/jp/products/audition.html, (accessed February 8, 2019).

[Anand et al., 2010] M. K. Anand, S. Bowers, and B. Ludäscher, "Provenance browser: Displaying and Querying Scientific Workflow Provenance Graphs," in *Proceedings of IEEE 26th International Conference on Data Engineering*, pp. 1201–1204, 2010.

[Anderson, 2001] J. Anderson, "Spectral music," [Online]. Available: http://www.oxfordmusiconline.com/subscriber/article/grove/music/50982, 2001.

[Andrews et al., 2010] C. Andrews, A. Endert, and C. North, "Space to Think: Large High-Resolution Displays for Sensemaking," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 55–64, 2010.

[Audacity, 2019] Audacity, "Audacity," [Online]. Available: http://www.audacityteam.org/, (accessed February 8, 2019).

[AudioSculpt, 2019] AudioSculpt, "Audiosculpt Ircam Anasynthe," [Online]. Available: http://anasynth.ircam.fr/home/english/software/audiosculpt, (accessed February 8, 2019).

[Bavoil et al., 2005] L. Bavoil, S. P. Callahan, P. J. Crossno, J. Freire, C. E. Scheidegger, C. T. Silva, and H. T. Vo, "Vistrails: Enabling Interactive Multiple-View Visualizations," in *Proceedings of IEEE Visualization*, pp. 135–142, 2005.

[Bergstrom et al., 2007] T. Bergstrom, K. Karahalios, and J. Hart, "Isochords: Visualizing Structure in Music," in *Proceedings of Graphics Interface*, pp. 297–304, 2007.

[Bernardini, 2016] N. Bernardini, "Quantitative Considerations about the Tapes of Giacinto Schelsi," *Texts and Beyond: The Process of Music Composition from the 19th to the 20th Century*, ser. Ad Parnassum Studies. UT Orpheus, pp. 43–50, 2016.

[Berthaut et al., 2010] F. Berthaut, M. Desainte-Catherine, and M. Hachet, "Combining Audiovisual Mappings for 3D Musical Interaction," in *Proceedings of International Computer Music Conference*, New York, United States, pp. 357–364, 2010.

[Borkin et al., 2013] M. A. Borkin, C. S. Yeh, M. Boyd, P. Macko, K. Z. Gajos, M. Seltzer, and H. Pfister, "Evaluation of Filesystem Provenance Visualization Tools," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2476–2485, 2013.

[Brehmer et al., 2017] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner, "Timelines Revisited: A Design Space and Considerations for Expressive Storytelling," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 9, pp. 2151–2164, 2017.

[Brown et al., 2012] E. T. Brown, C. E. J. Liu, Brodley, and R. Chang, "Dis-function: Learning Sistance Functions Interactively," in *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pp. 83–92, 2012.

[Callahan et al., 2006] S. P. Callahan, J. Freire, E. Santos, C. E. Scheidegger, C. T. Silva, and H. T. Vo, "Vistrails: Visualization Meets Data Management," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*. New York, NY, USA: ACM, pp. 745–747, 2006.

[Cannam et al., 2006] C. Cannam, L. Christian, S. Mark, and P. B. Juan, "The Sonic Visualiser: A

Visualisation Platform for Semantic Descriptors from Musical Signals," in *Proceedings of 7th International Conference on Music Information Retrieval*. pp. 324–327, 2006.

[Cantareira et al., 2016] G. D. Cantareira, L. G. Nonato, and F. V. Paulovich, "Moshviz: A Detail+Overview Approach to Visualize Music Elements," *IEEE Transactions on Multimedia*, vol. 18, no. 11, pp. 2238–2246, 2016.

[Chan et al., 2010] W. Chan, H. Qu, and W. Mak, "Visualizing the Semantic Structure in Classical Music Works," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 1, pp. 161–173, 2010.

[Chen et al., 2012] P. Chen, B. Plale, Y. Cheah, D. Ghoshal, S. Jensen, and Y. Luo, "Visualization of Network Data Provenance," in *Proceedings of International Conference on High Performance Computing*, pp. 1–9, 2012.

[Ciuha et al., 2010] P. Ciuha, B. Klemenc, and F. Solina, "Visualization of Concurrent Tones in Music with Colours," in *Proceedings of the 18th ACM International Conference on Multimedia*, pp. 1677–1680, 2010.

[Cluster, 2001] Cluster, "Cluster," [Online]. Available: http:////www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000005992, 2001.

[Couprie, 2004] P. Couprie, "Graphical Representation: an Analytical and Publication Tool for Electroacoustic Music," *Organised Sound*, vol. 9, no. 1, pp. 109–113, 2004.

[Couprie, 2016] ——, "Eanalysis: Developing a Sound-Based Music Analytical Tool," *Expanding the Horizon of Electroacoustic Music Analysis*, L. Landy and S. Emmerson, Eds. Cambridge University Press, pp. 170–194, 2016.

[Crawford and Gibson, 2017] T. Crawford and L. Gibson, Eds., *Modern Methods for Musicology: Prospects, Proposals, and Realities*, Routledge, 2017.

[de Prisco et al., 2016] R. de Prisco, D. N. Lettieri, D. Malandrino, G. Pirozzi, Zaccagnino, and R. Zaccagnino, "Visualization of Music Plagiarism: Analysis and Evaluation," in *Proceedings of 20th International Conference Information Visualisation*, pp. 177–182, 2016.

[Doboš et al., 2014] J. Doboš, N. J. Mitra, and A. Steed, "3D Timeline: Reverse Engineering of a Part-Based Provenance from Consecutive 3D Models," *Computer Graphics Forum*, vol. 33, no. 2, pp. 135–144, 2014.

[Donin and Traube, 2016] N. Donin and C. Traube, "Tracking the Creative Process in Music: New Issues, New Methods," *Musicae Scientiae*, vol. 20, no. 3, pp. 283–286, 2016.

[Dunne et al., 2012] C. Dunne, N. H. Riche, B. Lee, R. Metoyer, and G. Robertson, "Graphtrail: Analyzing Large Multivariate, Heterogeneous Networks While Supporting Exploration History," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1663–1672, 2012.

[Eldénius, 1998] M. Eldénius, *Formalised Composition on the Spectral and Fractal Trails*, Göteborg : Göteborg University, School of Music and Musicology, 1998.

[Emmerson and Landy, 2016] S. Emmerson and L. Landy, *Expanding the Horizon of Electroacoustic Music Analysis*, S. Emmerson and L. Landy, Eds. Cambridge University Press, 2016.

[Féron, 2011] F.-X. Féron, "The Emergence of Spectra in Gérard Grisey's Compositional Process: from Derives (1973–74) to Les Espaces Acoustiques (1974–85)," *Contemporary Music Review*, vol. 30, no. 5, pp. 343–375, 2011.

[Fineberg, 2000a] J. Fineberg, "Spectral Music," *Contemporary Music Review*, vol. 19, no. 2, pp. 1–5, 2000.

[Fineberg, 2000b] ——, "Guide to the Basic Concepts and Techniques of Spectral Music," *Contemporary Music Review*, vol. 19, no. 2, pp. 81–133, 2000.

[Foote and Cooper, 2001] J. Foote and M. L. Cooper, "Visualizing Musical Structure and Rhythm via Self-Similarity," in *Proceedings of International Computer Music Conference*, 2001.

[Fujishiro, 2007] I. Fujishiro, "Visualization Design and Lifecycle Management," in *Proceedings of Dagstuhl Seminar on Scientific Visualization*, 2007.

[Giannakis and Smith, 2012] K. Giannakis and M. Smith, "Imaging Soundscapes: Identifying Cognitive Associations between Auditory and Visual Dimensions Kostas," *Musical Imagery*, pp. 161–180, 2012.

[Griffiths, 2001] P. Griffiths, "Serialism," [Online]. Available: http:////www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000025459, 2001.

[Grisey and Fineberg, 2000] G. Grisey and J. Fineberg, "Did You Say Spectral?" *Contemporary Music Review*, vol. 19, no. 3, pp. 1–3, 2000.

[Groth and Streefkerk, 2006] D. P. Groth and K. Streefkerk, "Provenance and Annotation for Visual Exploration Systems," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1500–1510, 2006.

[Harvey, 2000] J. Harvey, "Spectralism," *Contemporary Music Review*, vol. 19, no. 3, pp. 11–14, 2000.

[Hayashi et al., 2011] A. Hayashi, T. Itoh, and M. Matsubara, "Colorscore—Visualization and Condensation of Structure of Classical Music," in *Proceedings of 15th International Conference on Information Visualisation*, pp. 420–425, 2011.

[Heer et al., 2008] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala, "Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1189–1196, 2008.

[Heifetz, 1984] R. J. Heifetz, "East-West Synthesis in Japanese Composition: 1950–1970," *The Journal of Musicology*, vol. 3, no. 4, pp. 443–455, 1984.

[Herd, 1989] J. A. Herd, "The Neonationalist Movement: Origins of Japanese Contemporary Music," *Perspectives of New Music*, vol. 27, no. 2, pp. 118–163, 1989.

[Herremans and Chew, 2016] D. Herremans and E. Chew, "Tension Ribbons: Quantifying and Visualising Tonal Tension," in *Proceedings of the International Conference on Technologies for Music Notation and Representation*, vol. 2, pp. 8–18, 2016.

[Hiraga and Matsuda, 2004] R. Hiraga and N. Matsuda, "Graphical Expression of the Mood of Music," in *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 2035–2038, 2004.

[Hirs and Gilmore, 2009] R. Hirs and B. Gilmore, *Contemporary Compositional Techniques and OpenMusic*. IRCAM, 2009.

[Jankun-Kelly et al., 2007] T. J. Jankun-Kelly, K. Ma, and M. Gertz, "A Model and Framework for Visualization Exploration," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 2, pp. 357–369, 2007.

[Juhasz, 2011] Z. Juhasz, "Low Dimensional Visualization of Folk Music Systems Using the Self Organizing Cloud," in *Proceedings of International Society for Music Information Retrieval Conference*, pp. 299–304, 2011.

[Kendall, 2016] G. S. Kendall, *Listening and Meaning: How a Model of Mental Layers Informs Electroacoustic Music Analysis*. Cambridge University Press, pp. 31–57, 2016.

[Klingbeil, 2008] M. Klingbeil, "Spectral Analysis, Editing, and Resynthesis: Methods and Applications," Ph.D. dissertation, Columbia University, 2008.

[Kurlander and Feiner, 1988] D. Kurlander and S. Feiner, "Editable Graphical Histories," in *Proceedings of IEEE Workshop on Visual Languages*, pp. 127–134, 1988.

[Lartillot et al., 2007] O. Lartillot, P. Toiviainen, and T. Eerola, "A Matlab Toolbox for Music Information Retrieval," in *Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V.*, pp. 261–268, 2007.

[Latham, 2011] A. Latham, *The Oxford Companion to Music*, version: 2011 ed. Oxford University Press, 2011.

[Li, 2016] J. Li, "Music analysis through visualization," in *Proceedings of the International Conference on Technologies for Music Notation and Representation*, pp. 220–225, 2016.

[Loubet et al., 1997] E. Loubet, C. Roads, and B. Robindor, "The Beginnings of Electronic Music in Japan, with a Focus on the NHK Studio: the 1950s and 1960s," *Computer Music Journal*, vol. 21, no. 4, pp. 11–22, 1997.

[Malandrino et al., 2015] D. Malandrino, D. Pirozzi, G. Zaccagnino, and R. Zaccagnino, "A Color-Based Visualization Approach to Understand Harmonic Structures of Musical Compositions," in *Proceedings of 19th International Conference on Information Visualisation*, pp. 56–61, 2015.

[Malt, 2015] M. Malt, "Une Proposition pour l'Analyse des Musiques Électroacoustiques de Xenakis à Partir de l'Utilisation de Descripteurs Audio," in *Iannis Xenakis, La musique Électroacoustique*, M. Solomos, Ed. L'Harmattan, pp. 159–198, 2015.

[Malt and Jourdan, 2008] M. Malt and E. Jourdan, "Zsa.Descriptors: a Library for Real-Time Descriptors Analysis," in *Proceedings of 5th Sound and Music Computing Conference*, Berlin, Germany, pp. 134–137, 2008.

[Mardirossian and Chew, 2007] A. Mardirossian and E. Chew, "Visualizing Music: Tonal Progressions and Distributions," in *Proceedings of the International Society of Music Information Retrieval*, pp. 189–194, 2007.

[Marston, 2001] N. Marston, "Sketch," [Online]. Available: http:////www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000042828, 2001.

[Mayuzumi, 1958] T. Mayuzumi, *Nirvana Symphony, 3rd SANNINNOKAI program*, 1958 (in Japanese).

[Mayuzumi, 1960] ——, *Mandala, 4th SANNINNOKAI program*, 1960 (in Japanese).

[Mayuzumi, 1969] ——, *Nirvana Symphony*, Edition Peters, 1969 (score).

[Mayuzumi et al., 1959] T. Mayuzumi, F. Koizumi, and M. Shibata, "About Nirvana Symphony," *Philharmony*, vol. 31, no. 3, pp. 34–43, 1959 (in Japanese).

[Minsburg, 2016] R. Minsburg, *Some Ideas Concerning the Relationship between Form and Texture.* Cambridge University Press, pp. 102–120, 2016.

[Montague, 1991] S. Montague, "Live Electronics Introduction," *Contemporary Music Review*, vol. 6, no. 1, pp. 85–88, 1991.

[Moorer, 1977] J. A. Moorer, "Signal Processing Aspects of Computer Music: A Survey," in *Proceedings of the IEEE*, vol. 65, no. 8, pp. 1108–1137, 1977.

[Morisette et al., 2013] J. T. Morisette, C. S. Jarnevich, T. R. Holcombe, C. B. Talbert, D. Ignizio, M. K. Talbert, C. T. Silva, D. Koop, A. Swanson, and N. E. Young, "Vistrails SAHM: Visualization and Workflow Management for Species Habitat Modeling," *Ecography*, vol. 36, no. 2, pp. 129–135, 2013.

[Muelder et al., 2010] C. Muelder, T. Provan, and K. Ma, "Content Based Graph Visualization of Audio Data for Music Library Navigation," in *Proceedings of IEEE International Symposium on Multimedia*, pp. 129–136, 2010.

[North et al., 2011] C. North, R. Chang, A. Endert, W. Dou, R. May, B. Pike, and G. Fink, "Analytic Provenance: Process+Interaction+Insight," in *Proceedings of CHI '11 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '11. New York, NY, USA: ACM, pp. 33–36, 2011.

[OpenMusic, 2019] OpenMusic, "IRCAM Forumnet OpenMusic," [Online]. Available: http://forumnet.ircam.fr/product/openmusic-en/, (accessed February 8, 2019).

[Park et al., 2009] T. H. Park, Z. Li, and W. Wu, "Easy Does It: The Electro-Acoustic Music Analysis Toolbox," in *Proceedings of the 10th International Society for Music Information Retrieval Conference*, pp. 693–698, 2009.

[Peeters, 2004] G. Peeters, ""A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project"," pp. 1–25, 2004.

[Peeters et al., 2011] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals," *The Journal of the Acoustical Society of America*, vol. 130, no. 5, pp. 2902–2916, 2011.

[Pressnitzer and McAdams, 2000] D. Pressnitzer and S. McAdams, "Acoustics, Psychoacoustics and Spectral Music," *Contemporary Music Review*, vol. 19, no. 2, pp. 33–59, 2000.

[Puzoń and Kosugi, 2011] B. Puzoń and N. Kosugi, "Extraction and Visualization of the Repetitive Structure of Music in Acoustic Data: Misual Project," in *Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services*, ser. iiWAS '11. New York, NY, USA: ACM, pp. 152–159, 2011.

[Radocy and Boyle, 1997] R. E. Radocy and J. D. Boyle, *Phychological Foundations of Musical Behavior*, 3rd ed. Charles C Thomas Publisher, 1997.

[Ragan et al., 2016] E. D. Ragan, A. Endert, J. Sanyal, and J. Chen, "Characterizing Provenance in Visualization and Data Analysis: An Organizational Framework of Provenance Types and Purposes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 31–40, 2016.

[Rio and da Silva, 2007] N. D. Rio and P. P. da Silva, "Probe-it! Visualization Support for Provenance," in *Proceedings of the 3rd international conference on Advances in visual computing*, pp. 732–741, 2007.

[Rose, 1996] F. Rose, "Introduction to the Pitch Organization of French Spectral Music," *Perspectives of New Music*, vol. 34, no. 2, pp. 6–39, 1996.

[Sapp, 2005] C. S. Sapp, "Visual Hierarchical Key Analysis," *Transactions of the ACM Computers in Entertainment*, vol. 4, no. 44, pp. 1–19, 2005.

[Schwarz and Schnell, 2010] D. Schwarz and N. Schnell, "Descriptor-Based Sound Texture Sampling," in *Proceedings of the 7th Sound and Music Computing Conference*, pp. 510–515, 2010.

[SDIF, 2019] SDIF, "SDIF Sound Description Interchange Format," [Online]. Available: https://sdif.sourceforge.net/, (accessed February 8, 2019).

[Shimizu, 2007] Y. Shimizu, "On the Chorus of Nirvana symphony by Toshiro Mayuzumi with Special Reference to the Relationship with Buddhist Chants," *Harmonia*, vol. 37, pp. 15–33, 2007 (in Japanese).

[Shimizu, 2010] ——, "Serial Technique in Toshiro Mayuzumi's 'Campanology Effect'," *Journal of the Musicological Society of Japan*, vol. 56, no. 1, pp. 26–37, 2010 (in Japanese).

[Shneiderman, 1996] B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," in *Proceedings of IEEE Symposium on Visual Languages*, pp. 336–343, 1996.

[Silva et al., 2007] C. T. Silva, J. Freire, and S. P. Callahan, "Provenance for Visualizations: Reproducibility and Beyond," *Computing in Science Engineering*, vol. 9, no. 5, pp. 82–89, 2007.

[Smith and Murail, 2000] R. B. Smith and T. Murail, "An Interview with Tristan Murail," *Computer Music Journal*, vol. 24, no. 1, pp. 11–19, 2000.

[Snydal and Hearst, 2005] J. Snydal and M. Hearst, "Improviz: Visual Explorations of Jazz Improvisations," in *Proceedings of the CHI'05 Extended Abstracts on Human Factors in Computing Systems*, pp. 1805–1808, 2005.

[Stauffer, 2001] S. L. Stauffer, "Composing with Computers," *Bulletin of the Council for Research in Music Education*, no. 150, pp. 1–20, 2001.

[Walker et al., 2013] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. Nguyen, D. Stephens, B. Wong, and Y. Zheng, "An Extensible Framework for Provenance in Human Terrain Visual Analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, pp. 2139–2148, 2013.

[Wattenberg, 2002] M. Wattenberg, "Arc Diagrams: Visualizing Structure in Strings," in *Proceedings of IEEE Symposium on Information Visualization*, pp. 110–116, 2002.

[Wiggins, 2007] J. Wiggins, *Compositional Process in Music*, *International Handbook of Research in Arts Education*, vol. 16, no. 1, pp. 453–470, 2007.

[Wolkowicz et al., 2009] J. Wolkowicz, S. Brooks, and V. Keselj, "Midivis: Visualizing Music Pieces Structure via Similarity Matrices," in *Proceedings of the International Computer Music Conference*, pp. 53–56, 2009.

[Xu et al., 2015] K. Xu, S. Attfield, T. J. Jankun-Kelly, A. Wheat, P. H. Nguyen, and N. Selvaraj, "Analytic Provenance for Sensemaking: A Research Agenda," *IEEE Computer Graphics and Applications*, vol. 35, no. 3, pp. 56–64, 2015.

[Yamashita, 1948] K. Yamashita, *Experimental Acoustics*. Ohmsha, pp. 85–130, 1948 (in Japanese).

[Yoon et al., 2013] Y.-S. Yoon, B. A. Myers, and S. Koo, "Visualization of Fine-grained Code Change History," in *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing*, pp. 119–126, 2013.

[Young, 2016] J. Young, "Forming Form," *Expanding the Horizon of Electroacoustic Music Analysis*, L. Landy and S. Emmerson, Eds. Cambridge University Press, pp. 58–79, 2016.

[Zhang et al., 2013] J. Zhang, P. Votava, T. J. Lee, O. Chu, C. Li, D. Liu, K. Liu, N. Xin, and R. Nemani, "Bridging Vistrails Scientific Workflow Management System to High Performance Computing," in *Proceedings of IEEE World Congress on Services*, pp. 29–36, 2013.

[Zhu and Lu, 2005] J. Zhu and L. Lu, "Perceptual Visualization of a Music Collection," in *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 1058–1061, 2005.