Vol.6 No.2

先端芸術音楽創作学会　会報

今号のコンテンツ

研究報告

五度圏に基づいた確率的複雑化と偏向化による情動的旋法を用いた音楽生成システム

SOUND GENERATOR SYSTEM WITH EMOTIONAL MUSICAL MODE BASED ON PROBABILISTIC COMPLEXITY AND BIAS ON THE CIRCLE OF FIFTH

大村英史, 池田大樹, 吉池卓也, 金喜晴, 栗山健一
Hidefumi Ohmura, Hiroki Ikeda, Takuya Yoshiike, Yoshiharu Kim, Kenichi Kuriyama
独立行政法人 国立精神・神経医療研究センター 精神保健研究所 精神保健研究部
Department of Adult Mental Health, National Institute of Mental Health
National Center of Neurology and Psychiatry

概要

音楽は，期待の実現や裏切りによって情動を引き起こすと言われている．筆者は，このような期待の実現・裏切りを情報理論に基づいてリズムを生成するシステムを開発したきた．本発表では，この方法論ピッチ（五度圏）に適用させ，確率的分布の複雑性と偏向性を定量的にコントロールし旋法を生成する手法を提案する．そして，この手法を組み込んだ情動的サウンド生成システムを紹介する．

It is say that deviations and implications of expectation when listening to music arouses emotions. Following this perception, we developed the system which generated emotional rhythm using deviations and implications of expectation based on information theory. In this paper, I propose the method that is generating musical modes controlling probabilistic complexity and bias on the circle of fifth. I show the system generating emotional system with the method.

1. はじめに

音楽は，人間の情動に働く．普段の生活でも，多くの人間が情動コントロールに使っている [1]．音楽の人間への影響がわかれば，健常な人間の情動作用だけでなく，精神疾患の患者への応用も可能になる．さらに，抽象芸術全般や雰囲気といった事象なども，音楽と同様な性質を持ったものであると，筆者は考えており，応用可能である．これらを，定量化および定式化することができれば，上記にあげたような様々な分野の事象に活用が可能である．

本研究では，このような抽象的な事象の定量化を目指し，まずは，音楽における物理的特性の一つである周波数の関係性に着目し，定量的なコントロールを試みた．

2. 音楽の特性

2.1. 二つの時間的特性

音は空気の振動である．一秒間に振動する回数を周波数としてあらわすと，この値が音の高さが決定する．この音の高さを音高（pitch）という．一方，振動している（発音している）期間を音価（value）という．音価は一般的に，相対的な長さで表す．基準となる長さ（小節）に対する比で表すのが一般的である．両者とも，時間的な特徴量である．

音の特徴として，音色がある．音色は音の波形として考えることができる．任意の音を周波数解析すると，構成要素の正弦波を確認することができる．逆に考えると，音色は，異なる時間的特徴量を保持した複数の正弦波から成り立っており，音高の時間的特徴量の合成として見なせる．

また，音楽の構成要素の和音は，音高の異なる複数の音が同時に響くおとであり，音色と同様，音高の時間的特徴量の合成として見なすことができる．

音楽の構成要素の律動（リズム）は，基準となる時間的長さ（たとえば小節）で作られるパルスであり，その中に音を配置する（たとえば裏拍）ことで複雑性が増す，音価の時間的特徴量の構造である．

このように概観していくと，音楽の構造は，音高と音価の二つの特徴量に帰着することができ，時間的な関係

性によって成立している．

筆者はいままで音価に注目し，定量的にリズムを生成するシステムを開発した [2]．そこで，本研究では，もう一方の音高（pitch）に注目し，定量的に音高を出力するモデルを提案し，システムに実装する．

2.2. 12 音階と 5 度圏

現在，世界中で最も多く用いられている音階は 12 音階である．この音階は，ピタゴラス音律や三分損益法と呼ばれる方法で得ることができ，基本的にはある音の周波数を 1.5 倍することで他の音を得ていく．12 音階は 1 オクターブ（2 倍の周波数）内に 12 音を配置する．ある音の周波数を 1.5 倍を 12 回繰り返して得られる音が，2 倍して得られる音を 7 回繰り返して得られる音を等しいとして考えることによって作られる．しかし，1.5^{12} と 2^7 は，わずかな誤差（ピタゴラスコンマ）が生じるためため，実際は近似的な 12 分割であり，これを修正するための方法はいくつか存在する．1 オクターブ（2 倍の周波数）内に 12 音を配置する 12 音階はあくまでも近似であるが，1.5 倍（5 度）と 2 倍（オクターブ）でシンプルに構造化できるだけでなく，12 が多くの約数を持つため，多元的な関係を持った構造の構築が可能となる．

本研究では，12 音階を踏襲し，オクターブの関係は同一と見なし 5 度の関係と 4 度の関係（5 度の逆）が最も近い関係であると見なす．つまり，5 度圏における両隣を最も近い関係だと見なす．

3. モデル化

3.1. エントロピー

Meyer によると，音楽における情動喚起は，音楽的期待からの逸脱によって生じる [3]．このような逸脱は，頻度の少ない事象として表現可能であり，情報理論でいうところのエントロピー [4] として表現可能である．

情報理論では，どの程度意味のある情報が伝達したのかを定量的に求めることができる．音楽は，時間変化と共に様々な構造を形成する音を聴取者が享受することにより，情動喚起を生じさせる．このような音の配置の伝達を，情報理論をもとにモデル化することにより情動喚起のための音楽を生成することが可能になると，私たちは考える．

情報理論は，事象 i が起きた際の情報量を次式で定義する．

$$I = -\log p_i$$

ここでは，イベント i が生じる確率を p_i とする．そして，n 個の事象がそれぞれ，$p_1, p_2, \cdots p_n$ で生じるとき情報量の期待値は次式で求められる．

$$H = -\sum_{i=0}^{n} p_i \log p_i$$

この値が大きくなると，多くの情報が得られることになる．つまり，不確実性が高まり複雑になると言い換えることができる．この値をエントロピーまたは平均情報量と呼ぶ．

12 音の各音に，出現確率を与えることで，上記のエントロピーを求めることが可能になる．

3.2. ガウス関数による各音の確率

12 音に各音の確率を与えるために，正規分布であるガウス関数を用いる．この理由は，2.2 で説明したように，音高には関係性があるからである．関係性が近いほど，音同士親和性が高く，離れると親和性が低くなると考える．

ガウス関数は以下の式で表される．

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

これは平均 μ，分散 σ^2 の正規分布を表すグラフである．x の値は各音に相当する．μ と σ の値を変化させることで各音の出現確率を計算する．

3.3. 複雑性

ガウス関数の分散 σ^2 を小さくすと，曲線が扁平となる．すると，各音の確率がおおよそ等しくなり，一様分布に近づく．このとき，どの音が発音されるか予想することが困難で有り，複雑性が上昇している．一様分布になるとエントロピーは最大になる．一方，分散 σ^2 を大きくすると特定の音のみ発音されることになり，聴取者は次の音を予測することが容易になる．

システムでは，ガウス関数の面積が 95% に収まるような範囲で音が鳴るように設定されている．

3.4. 偏向性

発音範囲が決まったら，μ の値により中央値をずらす．中央値は，最も頻度の出現頻度の高い音であり，主音として見なすことができる．右（正）方向にずらした場合，ガウス関数が 4 度の方向にずれる．左（負）方向にずらした場合は 5 度の方向にずれる．

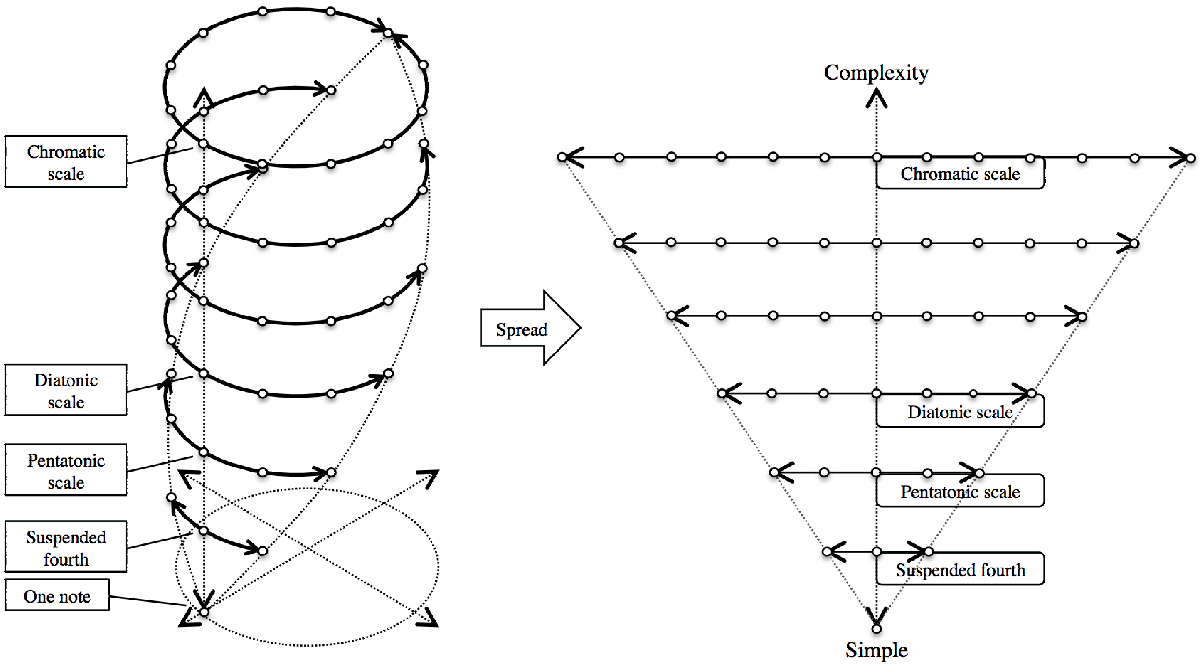

図 1. 複雑性の変化による選択される音

4. 実装

実装は，html 上に javascript を用いて行った[1]．音の出力には Web Audio API を用いている．また，各パラメータをチューニングには midi 機器を利用できるように Web Midi を用いている．Google chrom で動作確認を行ってある．初期の Main ページで，音量や単位時間あたりの音の密度などが調整できる．音の密度設定に従って，ランダムで音が出力される．左下の Play/Stop ボタンでシステムが起動/停止する．Setting ページでは，複雑性と偏向性をコントロールできる．また，手動でも各音の出現確率や，利用の可/不可がコントロールできる．

5. 考察

複雑性と偏向性をコントロールしてシステムを動かすと興味深い結果が得られる．複雑性および偏向性それぞれにおいて考察を行う．

5.1. 複雑性の考察

複雑性をコントロールしていくと図 1 のような変化をする複雑性を最小にすると，一音だけが選択される．この時はある音のオクターブだけが発音する．この値を大きくすると 3 音が選択され，5 度と 4 度の音が発音される．この音は，コードの suspended 4th の響きと類似している．さらに大きくすると，5 音が選択され，ペンタトニックスケールのような響きとなる．さらに大きくすると，7 音が選択され，ダイアトニックスケールのような響きとなる．順に大きくしていくと 9 音，11 音となり，最終的に 12 音となり，クロマティックスケールと同等の響きが得られる．

複雑性を変化させることにより，選択される音が増え，複雑なサウンドが出力されていることがわかる．

5.2. 偏向性の考察

複雑性のコントロールにより，発音される音が決定されたら，中央値を変化させて，偏向性を変化させる．興味深いことに，5 度（負）の方向へ中央値を変化させた場合，明るい響きに変化していく．一方，4 度方向へ変化させた場合，暗い響きに変化していく．

例えば，7 音のときに教会旋法と比較すると，中央値を変化させない場合ドリア旋法のような響きであるが，5 度の方向へ中央値を移動させると，ミクソリディア，イオニア（長調），リディア旋法のような響きに変化していく．各モードとも明るい響きである．一方，中央値を 4 度の方向へ変化させた場合，エオリア（短調），フリギア，ロクリア旋法のような響きに変化していく．こちらは，暗い響きである．

5 音の場合も同様で，5 度方向へ変化させると，メ

[1] https://sites.google.com/site/hidefumiohmura/home/program/entropymode

ジャーペンタトニックや呂音階やヨナ抜き音階のような明るい響きになり，4 度方向へ変化させると，マイナーペンタトニックや陽音階やのような暗い響きになる．

5.3. 情動的な旋法

以上をまとめると，複雑性と偏向性を変化させると，出現する音の組み合わせが変化して，情動的な響きの変化を得られることがわかる．Meyer の言うところの，期待の逸脱は複雑性（エントロピー）の変化に対応していると考えられる．一方で，長調・短調のような明るい・暗いと言った情動的変化は偏向性を変化させることで得られることがわかる．

提案したモデルでは，設定した二つのパラメータを変化させることで，旋法のような響きが得られたが，これらはすべて連続的に変化させることが可能だ．このような旋法は名称が付いており離散的だと考えられているが，2 次元の連続的空間にマッピングできる可能性が示唆される．

6. おわりに

音の関係性により，出力される音を複雑性と偏向性の二つのパラメータでコントロールするモデルを提案した．このモデルを実装したシステムで出力されるサウンドを聴いてみると，実際の旋法と類似した響きが得られた．この結果から，旋法で定義されるような音の響きは，二つの特性によって作られる空間にマッピングすることができ，連続的である可能性が示唆された．このような定量的なコントロールは音楽が人間にどのような影響を与えているか，より厳密に調べることが可能となる．

今回は，5 度の連続した関係のみ扱ったが，琉球音階や都節音階のような 5 度で連続していない音階も存在する．今後は，このような音階も考慮できるようなモデルも検討していきたい．

7. 参考文献

[1] T. DeNora: Aesthetic agency and musical practice, *New directions in the sociology of music and emotion*, Oxford University Press, 2001.

[2] H. Ohmura, T. Shibayama, T, Shibuya, T, Takahashi, K. Okanoya, and, K. Furukawa, “Modeling of Melodic Rhythm Based on Entropy toward Creating Expectation and Emotion,” *Proceeding of Sound and Music Computing 2013 (SMC)*, pp.69-73, 2013.

[3] L.B. Meyer, *Emotion and meaning in music*, Chicago: University of Chicago Press, 1956.

[4] C. E. Shannon, *The Mathematical Theory of Communication*, The University of Illinois Press 1949. (植松訳「通信の数学的理論」ちくま学芸文庫，2009)

8. 著者プロフィール

大村英史 (Hidefumi Ohmura)

2002 年東京農工大学工学部機械システム工学科卒業．2009 年東京工業大学大学院博士課程修了．博士 (工学)．同年 JST, ERATO, 岡ノ谷情動情報プロジェクト研究員．および，理化学研究所脳科学総合センター生物言語チーム客員研究員，2014 年国立精神神経医療センター精神保健研究所流動研究員，現在に至る．音楽によって生じる人間の情動の分析・応用に関する研究や，人工的雰囲気生成の研究に従事．知能情報ファジィ学会，人工知能学会，認知科学会，音響学会など会員．

Research Report

ALARM/WILL/SOUND: CAR ALARM SOUND PERCEPTION EXPERIMENTS AND ACOUSTIC MODELING REPORT

Alexander Sigman
College of Music and Performing Arts
Department of Composition
Keimyung University
Daegu, South Korea

Nicolas Misdariis
Sound Perception and Design
Research and Development
STMS-IRCAM-CNRS-UPMC
Paris, France

ABSTRACT

This article outlines the salient phases, goals, and results of the sound perception and design research component of *alarm/will/sound*, a multidisciplinary musical research project carried out in the context of the IRCAM IRC *(Interface Recherche-Création)* Musical Residency Research program. After the rationale for and motivations behind the project are presented, the following research milestones are described: 1) a sound perception experiment testing source typicality of a sub-category of sounds within the corpus; 2) an acoustic descriptor space in which a subset of the stimuli employed in the typicality experiment were situated; 3) the construction of synthetic auditory warnings from sound-sources within the descriptor space, prototypical environmental sound envelopes, and inter-onset intervals (IOI's) derived from extant car alarms; and 4) the design and results of a second experiment pertaining to levels of repulsion vs. attraction to the synthetic auditory warnings. Finally, short-, mid-, and long-term objectives and directions for the project are discussed.

1. INTRODUCTION

alarm/will/sound [1] is a tripartite collaboration between composer Alexander Sigman, IRCAM Sound Perception and Design (SPD) team researcher Nicolas Misdariis, and Stuttgart-based product designer/visual artist Matthias Megyeri. This project was begun in January 2013 under the IRCAM IRC *(Interface Recherche-Création)* Musical Research Residency program, and is still in progress at present (September 2014).

Taking as a point of departure the proven ineffectiveness of current audible car alarm systems as deterrents [1] and the relative lack of research into and development of audible car alarm design compared to other sound-emitting components of vehicles (e.g., the audio system, engine, horn, turn signal, or door), we have sought to produce innovative modified car alarm prototypes. The design of these prototypes would be informed by musical, artistic, scientific, and industry expertise, as well as sound perception research and acoustic modeling.

As an often-ignored and predictable source of noise pollution, the car alarm as an auditory warning device raises a host of intriguing questions of a sociological nature. How may the essential functionality of the audible car alarm be defined? To whom is the alarm directed: the potential perpetrator, the car owner, or the public? Studies summarized in the report cited above have indicated that even the most high-end audible alarms require a maximum of ten minutes for professional car thieves to disable, and in most instances fail to prevent break-ins. The owner may or may not be within earshot of his/her respective car alarm when it is activated. However, given the homogeneity of car alarm emissions, the alarm may not be detected in time for the car owner to intervene. As a result of the sheer number of false alarms, as well as the aforementioned element of aural annoyance (among other factors), members of the more public have a greater documented tendency to flee from or simply ignore an activated alarm than to proceed towards the source.

In the presence of a car alarm, wherein lies the boundary between the public space of the vehicle's physical environment and the private territory of car? An audible alarm designates a boundary beyond the car's physical perimetera "grey zone" that is often creatively explored, as is made evident by the countless videos of car alarm dance routines posted to Youtube. [2]

Our approach to audible car alarm system design has been guided and constrained by a fundamental question: wherein lies the alarm's identity? Does it consist in the hardware components and context (i.e., embedded in the automobile), or in the sounds that it emits and the interaction protocol by which it operates? If the latter, is it possible to transform the alarm system from a mechanism of (ineffective) deterrence into one of engagement? That is to say: through the expansion and customization of its sonic vocabulary and potential modes of human-machine interaction, could this device be repurposed as a sort of virtual instrument that the passerby (or car owner) learns to manipulate, with the help of audio-visual feedback? By the same token: could the alarm gain sensitivity to more physical parameters than simply physical proximity? Could temporal variation in these physical parameters trigger time-varying sonic responses? At this point, it is worth mentioning that the aim of the study is notat least, in the short termto address mnemonic and learning issues attached to auditory warning stimuli. Even if these mechanisms could radically change the way people react to alarms, our investigation do not pertain directly to how

[1] Also referred to as "*a/w/s*" in this paper.

[2] Here is a particularly well-choreographed example: http://www.youtube.com/watch?v=1Li3mNl2-EM (accessed 22 September 2014).

people integrate over time the information induced by an alarm. This is a question that would require a long-term experimental paradigm to accurately and rigorously investigate.

Indeed, it was not our intent purely to focus upon enhancing the car alarm's deterrence effectiveness. Non-audible devices such as the Lojack system [3] have been associated with a high documented vehicle recovery rate. Rather than entirely replacing an existing system in service of security enhancement, the focus has been placed on expanding and enhancing this system's functionality and sonic potential.

Despite the seemingly unique nature of the project and the collaborative model underlying it, aspects of *a/w/s* are in fact extensions of Matthias Megyeri's work on domestic security systems [4] and Alexander Sigman's compositional interests in the influence of sonic phenomena in physical environments on the aesthetics of the composer/sound artist and the impact that a composer/sound artist may have on transforming the physical environment (and by extension, the human behaviors therein). [5]

The IRCAM Sound Design and Perception research team's involvement with the automobile industry has assumed the form of a partnership with Renault on sound design for electric vehicles (in collaboration with composer Andrea Cera), a follow-up study on electric vehicle detectabilty in urban environments [2], and an earlier study on car horn sound quality [3]. Research topics in the domain of human-machine interaction has included the influence of audio features on perceived urgency and its application to car interior Human-Machine Interfaces [4] and the influence of naturalness of auditory feedback of an interface on perceived usability and pleasantness [5]. In addition, Sigman's background in Cognitive Science and timbre perception has been relevant to the project's collaborative model. It is thus hoped that both the artistic and research outcomes of *alarm/will/sound* will contribute not only to the understanding and development of vehicle alarm systems specifically, but also to the design and classification of auditory warnings in general.

2. PROJECT PHASES AND GOALS

For practical purposes, the project was divided into four primary phases: three research and production phases (see Figure 1) and one presentation and user experience documentation phase. The first phase was devoted to the production and characterization of a corpus of potential alarm sounds. Subsequently, a sound perception experiment in sound source identifiability was designed and conducted on a significant portion of the corpus, in order to focus on acoustic properties, rather than exclusively to sound causality. A subset of the stimuli used in this experiment was then placed within an acoustic features descriptor space. Phase III has consisted of constructing synthetic auditory warnings via the integration of the source-sounds within the descriptor space, prototypical auditory warning temporal morphologies, and the inter-onset intervals (IOI's) of real car alarm sounds. These synthetic warnings are currently being tested for their respective capacities for repulsion vs. attraction in a second experiment. Concurrently, Matthias Megyeri is developing hardware designs for the eventual prototypes, which will be exhibited as interactive installations in public and gallery spaces during the final phase of the project.

Given the breadth of *a/w/s*, we have decide to focus in the present article on the sound corpus elaboration and characterization process of Phase I, the source typicality experiment and acoustic feature modeling completed during Phase II, and the synthetic auditory warning construction and deterrence vs. engagement experiment of the third phase. All of these project achievements pertain to the creation of semantic and acoustic classification of the alarm prototype sounds. The classification methodologies employed will be critical to both subsequent research and to exhibiting the prototypes in an interactive art installation context.

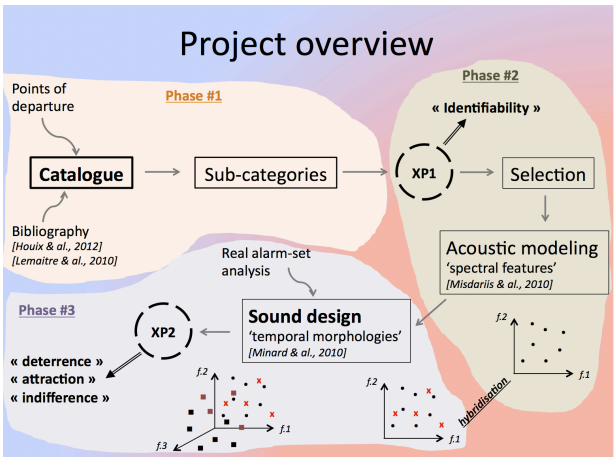

Figure 1. The three primary project phases. "XP1" and "XP2" refer to "Experiment 1" and "Experiment 2," respectively.

3. BACKGROUND

Besides the IRCAM Sound Perception and Design team studies mentioned above, several important researches have informed each stage of the project. Initially, a survey of historical audible alarm patents and current standards was made. Sound corpus construction was guided both by classic twentieth century approaches to timbral classification (e.g., Pierre Schaeffer's *Traité des objets musicaux* [6]), as well as more recent studies in environmental sound categories (e.g., Houix et al, 2012 [7]). Formal components were also extracted from Olivier Claude's 2006 thesis *La recherche intelligente des sons* [8], where taxonomies of natural, animal, human, and object/machine sounds are proposed, such that sounds within these taxonomies are organized into limited sets of morphological, causal (physical), and semantic sub-categories.

In Ballas (1993) [9], acoustic, ecological, perceptual, and cognitive factors that influence the identification of

[3] http://www.lojack.com/Home (accessed 22 September 2014).

[4] E.g., *Sweet Dreams Security*, a commercial line of security products developed and distributed by Megyeri: http://www.sweetdreamssecurity.com/sweetdreamssecurity.html (accessed 22 September 2014).

[5] Sigman's *VURTRUVURT* (2011) for prepared violin and live electronics and *down the bottle* (2012) for bass flute, installation, and live electronicsboth members of the *VURT* cyclereflect these interests. Scores and recordings to both works may be found on the composer's website: http://lxsigman.com/media/audio.htm (accessed 22 September 2014).

current environmental sounds were evaluated. As is explained in Section 5.2, this study was particularly relevant to the sound causality confidence rating protocol utilized in Experiment 1.

The acoustic modeling step was largely informed by several studies done on the definition of an exhaustive set of acoustic featuresat first dedicated to musical sounds (Peeters et al, 2002 [10]) and the identification of some of these features that are best suitable for describing similarities and differences of environmental sounds.

The auditory warning construction was grounded on the standard template defined by Patterson (1990) [11] and, among others, the direction taken by Edworthy (2011) [12] to extend the conception of the inner structure of such signals was considered. Moreover, our own auditory warning design process was also based on prototypes of morphological profiles found out by Minard et al. (2010) [13] from an environmental sound corpus.

Finally, existing auditory design in automobiles (Yamauchi et al, 2004 [14], Kuwano et al 2007 [15]), and approaches to the synthesis of new auditory warnings in military helicopters (Patterson 1999 [16]) and intensive care units (Stanton and Edworthy, 1998 [17]) were also taken into account in this process.

4. SOUND CORPUS TAXONOMY AND SOURCES

The first phase of the project entailed the elaboration and characterization of a sound corpus to apply to the modified car alarm prototypes. As is presented in Figure 2, the sound corpus taxonomy consists of three primary categories: individual sounds, ”auditory scenes,” or sound complexes, and real car alarm sounds (i.e., the standard repertoire of six auditory warnings typical of audible car alarm systems). The Individual Sound category is further divided into 1) Synthetic/Electroacoustic; 2) Vocal; 3) Film Danger Icons; and 4) Industrial/Mechanical Sounds. Further subdivisions were made along semantic/contextual andparticular at the lowest levels of the taxonomyacoustic lines.

Among the non-synthetic sounds in all three primary categories, the majority were mined from existing sound databases (e.g., SoundIdeas, Blue Box, Auditory Lab [6] and freesound.org). Under the Auditory Scenes rubric, a series of field recordings of public spaces in Parisstreets, the Forum Les Halles shopping concourse, the Centre Pompidou, metro stations, and train car interiorswere compiled in February 2013 by Alexander Sigman and Matthias Megyeri. It is intended that the collection of field recordings be expanded over time to include further site-specific entries.

Synthetic individual sounds were generated and edited using such synthesis software as AudioSculpt, [7] Pure Data (Pd), [8] SuperCollider, [9] and the Python-based concatenative synthesis program Audioguide. [10]

[6] http://www.psy.cmu.edu/ auditorylab/website/index/home.html (accessed 22 September 2014).

[7] http://anasynth.ircam.fr/home/english/software/audiosculpt

[8] http://puredata.info/

[9] http://supercollider.sourceforge.net/

[10] AudioGuide was developed by composer Ben Hackbarth, and is obtainable from his website: http://www.benhackbarth.com/audioGuide/doc.html (accessed 22 September 2014).

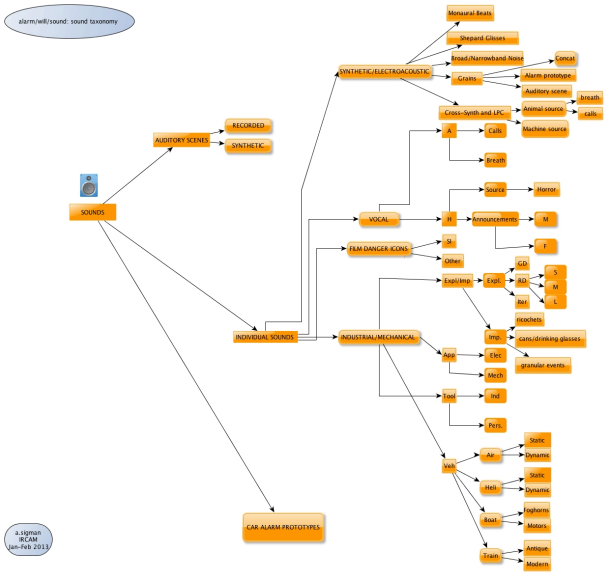

Figure 2. Sound corpus taxonomy, constructed in January-February 2013.

5. EXPERIMENT 1: SOURCE IDENTIFIABILITY OF INDUSTRIAL/MECHANICAL SOUNDS

5.1. Experimental Objectives

The first sound perception experiment was designed in the interest of determining levels of source identifiability of sounds within the corpus. Based upon the results of the experiment, it would be possible to construct an abstractness-iconicity scale across the corpus, as well as to determine the salient semantic and acoustic attributes of the sounds using empirical data. However, given the size and scope of the catalogue, the selected stimuli were limited to a subset of the Industrial/Mechanical category. This category was chosen due to a) the number of subcategories and entries; and b) the range of source abstractness and ecological context relative to other Individual Sound categories.

5.2. Methods and Materials

In order to obtain data from a broad range of subjects over a relatively short period of time (ca. one month), the experiment was conducted in an online, crowd-sourced format. [11] Subjects were asked to listen to each stimulus, provide a brief description of sound causality, and indicate a confidence rating of sound causality identification on a 1-5 Likert scale (see Figure 3). [12] The stimuli could be played back any number of times, but could not be paused and resumed mid-file. Thirty-nine stimuli were presented in MPEG-3 format. Every sub-category of the Industrial/Mechanical category was represented by at least one stimulus.

[11] The experiment may be found at the following URL: http://recherche.ircam.fr/equipes/pds/projects/asigman/causality/ src/EvaluationStartTrial.php (accessed 22 September 2014).

[12] This experimental protocol was based on Ballas (1993), who found a correlation between the measure of confidence of sound causality and the more laborious causal uncertainty measure (Heu) [9]

All subjects were required to complete a trial session consisting of three practice trials prior to being directed to the experiment, in order to determine judgment stability for each participant. Subjects could not advance to the next trial until the "Sound Source Description" field was filled out. Once the experiment was completed, subjects were requested to submit a questionnaire pertaining to the subjects' professional and educational backgrounds, audio equipment used during the experiment, and acoustics of physical environment in which the subjects were located at the time of participating in the experiment. In addition, feedback regarding the experiment was solicited. The whole listening test lasted approximately thirty minutes. Subjects were not paid for their participation.

From an objective point of view, the use of an online procedure can appear to be either a blessing or a curse: as mentioned above, it allows a large number of participants representing a broad range of professional backgrounds and geographic locations to complete the experiment in a limited period of time. On the other hand, it may induce "noisy," unreliable data mainly due the inability to directly manage listener fatigue or lack of concentration within this protocol. The results presented below in section 5.3 should be examined in light of these limitations.

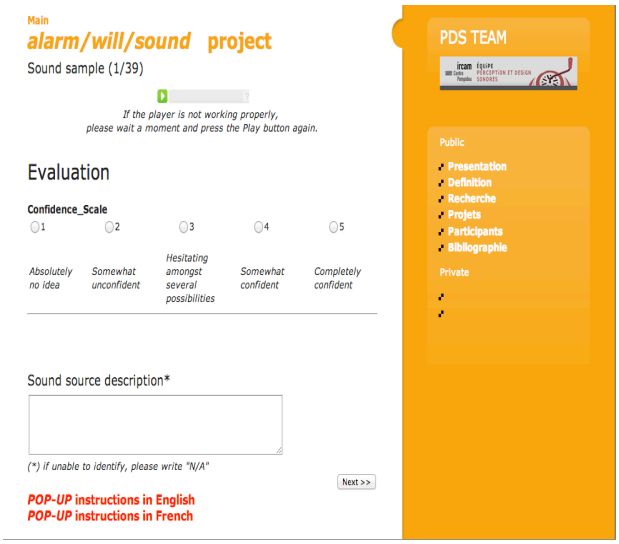

Figure 3. Experiment 1 user interface.

5.3. Results

Of the ca. 100 visitors to the experiment website, twenty-four subjects began the experiment. Of this subject pool, only fifteen subjects completed all trials. Data collected from the nine subjects who did not reach the end of the experiment was excluded from the analysis.

Figure 4 indicates the confidence ratings of the fifteen subjects across the thirty-nine stimuli. The stimuli are indicated on the x-axis from left to right in order of confidence rating (from low to high). A significant difference t-test was applied to the lowest and highest mean confidence ratings in order to locate the threshold between iconic (strongly identifiable) and non-iconic sounds. Stimuli and their respective mean confidence ratings are listed (alphabetically) in Figure 5.

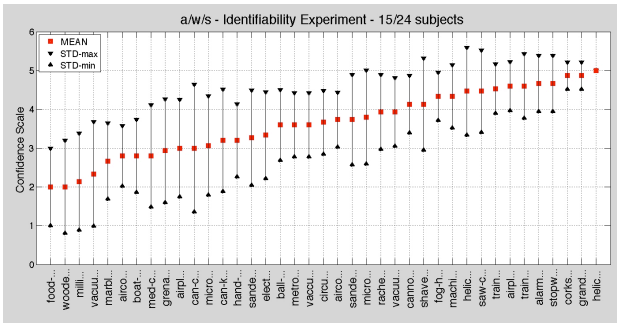

Figure 4. Confidence rating means and standard deviations for thirty-nine stimuli indicated from low (left) to high (right).

Amongst the sound source descriptions provided by the subjects (in the field below the confidence scale on the experiment interface), the responses ranged in confidence and specificity. (In one case, for instance, a subject identified the brand of metronome of one of the stimuli.)

5.4. Discussion

The relatively high attrition rate amongst potential subjects (100 to twenty-four to fifteen) suggests two factors that discouraged participants from continuing with the experiment: 1) the duration required to complete the task (45-60 minutes) and 2) the fatiguing nature of the sounds presented. There was also no available "Save and Continue" option, so the experiment had to be conducted in one sitting. Experiment 2 requires a shorter completion time (ca. 15-20 minutes), and consists of stimuli of shorter durations.

The individual differences in confidence ratings and specificity of responses may be attributed in part to differences in experience with mechanical/electrical tools and industrial equipment. The sounds that caused the most confusion either had contextually-obscure sources (e.g., wooden gears or milling machine), impulsive (e.g., grenade blast, marbles in vase, or a ball-ricochet) or were steady-state and drone-like in nature, with motor sources (e.g., air conditioner, microwave oven, boat motor, or food processor). The iconic sounds were quite context-specific (e.g., grandfather clock bells, antique train, or machine gun), time-varying (e.g. helicopter, airplane, or train passing), or common (e.g., corkscrew opening a bottle or alarm clock). Interestingly, a few sounds that were segmented into two stimuli produced different confidence scale ratings for each segment. The sound "airco-off" (air conditioner off) had a mean confidence rating of 2.8, while "airco-on" was correlated with a mean rating of 3.73. Similarly, "airplane-end" had a mean rating of 3.0, while the rating "airplane-beginning" fell to the right of the iconicity threshold, at 4.6. As otherwise drone-like, steady-state, motor-produced sounds, the onsets of "airco-on" and "airplane-beginning," as well as "sander-on," may have provided more spectral cues than the respective offsets. By contrast, the termination of "microwave-oven-end" (confidence rating = 3.8) seems to have given more cues than "microwave-oven-begin" (confidence rating = 3.07).

Despite the lower levels of participation than expected, these results did enable us to construct an abstraction-iconicity scale and to determine salient semantic charac-

Stimulus	Mean Confidence Rating
1 airco_off	2.80
2 airco_on	3.73
3 airplane_beginning	4.60
4 airplane_end	3.00
5 alarm_clock_bell	4.67
6 ball_ricochet	3.60
7 boat_motor_edited	2.80
8 can_crushed_edited	3.00
9 can_knocked_over	3.20
10 cannon_shot	4.13
11 circular_saw	3.67
12 corkscrew	4.87
13 electric_screwdriver	3.33
14 fog_horn	4.33
15 food_processor_off	2.00
16 grandfather_clock_bells	4.87
17 grenade_blast	2.93
18 hand_mixer_off	3.20
19 helicopter_hovering	4.47
20 helicopter_passing	5.00
21 machine_gun_3_iteration	4.33
22 marbles_in_vase	2.67
23 med_clock_ticks	2.83
24 metronome	3.60
25 microwave_oven_begin	3.07
26 microwave_oven_end	3.80
27 milling_machine_on	2.13
28 rachet	3.93
29 sander_off	3.27
30 sander_on	3.73
31 saw_cutting_pipe	4.47
32 shaver_middle	4.13
33 stopwatch_beep	4.67
34 train_antique	4.53
35 train_rail_noise	4.60
36 vaccuum_end	3.60
37 vacuum_begin	3.93
38 vacuum_cleaner_in_motion	2.33
39 wooden_gears_excerpt	2.00

Figure 5. The thirty-nine stimuli (listed alphabetically) and their respective mean confidence ratings.

teristics of the sounds tested. It was concluded that the sounds falling to the right of the iconicity threshold would not function effectively in the car alarm context, as they were too closely associated with specific sources, which may very well exist within a vehicle's immediate environment (e.g., trains/airplanes/helicopters and clock chimes).

6. ACOUSTIC MODELING: PSC-HNR DESCRIPTOR SPACE

The next step in the corpus characterization process was to compute perceptually relevant acoustic descriptors. The objective was to construct an acoustic descriptor space in which to situate the corpus sounds, thereby enabling us to trace relative distances between constituent sounds, as well as between extant sounds and new entries.

Several approaches were considered, including Mel Frequency Cepstral Coefficient (MFCC) attached to a Gaussian Mixture Model (GMM) and a Multi-Dimensional Scaling (MDS) analysis. It was ultimately decided that weighted-mean Perceptual Spectral Centroid (PSC) and Harmonicity-to-Noise Ratio (HNR)two acoustic parameters included in the IrcamDescriptor 2.7 toolbox [18]would be employed based upon a 2010 study by Misdariis et al. on metadescription and modeling of environmental sounds such as interior car sounds, air conditioners, car horns, and car doors [19]. This choice of descriptors was made due to the fact that, in the Misdariis et al. study, it has been shown that these two features appear to be significant in the perceptual description of environmental sounds. Moreover, this study also indicated that, in a first-level description, considering the mean value of the featureseven if sounds are objectively non-stationarycorresponds to a perceptually relevant process (see correlation values between acoustic features and perceptual dimensions in [19]). That is the primary reason that we chose in the present study to remain on this first level and construct our Acoustic Features Space upon mean values of the features.

A two-dimensional PSC-HNR space was calculated and populated by the thirty-one Industrial/Mechanical sounds tested in Experiment 1 that fell below the iconicity threshold. Moreover, in order to fill empty or underrepresented zones of the space, hybrid sounds were constructed via the constant cross-synthesis of pairs of one-second samples of extant sounds in AudioSculpt. The resulting acoustic features space is illustrated in Figure 6.

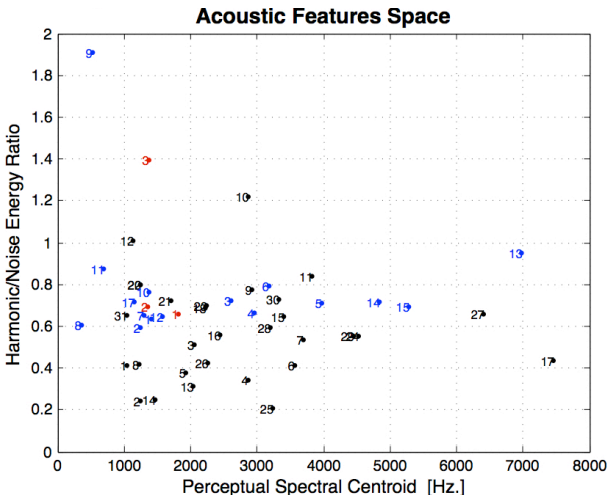

Figure 6. Perceptual Spectral Centroid (PSC)-Harmonicity-to-Noise Ratio (HNR) acoustic features space containing industrial/mechanical source sounds (black dots), actual car alarm sounds (red dots), and hybrid sounds (blue dots).

7. SYNTHETIC AUDITORY WARNING CONSTRUCTION

Synthetic auditory warnings were constructed by combining: a) selected industrial/mechanical source sounds and hybrids placed within the acoustic descriptor space described above; b) five typical environmental sound envelopes; and c) the inter-onset intervals (IOI's) of the six alarms comprising a standard car alarm repertoire.

Of the set of stimuli shown in Figure 6, six original sounds and three cross-synthesized hybrids that lie at the extremes and center of the descriptor space were selected. In the Minard, et al. study [14] mentioned previously, six perceptually distinct environmental sound morphologies were devised and tested: 1) stable; 2) decreasing; 3) increasing; 4) pulse-train; 5) single impulse; and 6) rolling (see Figure 7). Given the iterative nature of these alarms, the "stable" category was excluded, as this inhibits recognition of new iterations in an ecological or experimental context.

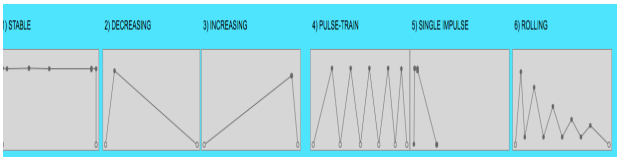

Figure 7. Six environmental sound envelopes, rendered as breakpoint functions (BPF's) in Max/MSP.

As is presented in Figure 8, the inter-onset intervals of six standard car alarms were calculated (in Matlab [13]).Given the similarity in IOI of alarms 1 and 2, 3 and 5, and 4 and 6 (respectively) to each other, each pair of IOI's was averaged, producing three distinct IOI durations: 300 ms. (short), 536 ms. (medium), and 2082 ms. (long).

Since the combination of nine sources times five envelopes times three IOI's would still produce too many stimuli to be realistically tested in a sound perception experiment, and several combinations would create contradictions among the parameters (and by extension, lose perceptual salience in an auditory warning context), further exclusions were made on an intuitive (but empirical) basis. For instance, new onsets of sounds with a decreasing envelope and a short IOI were deemed imperceptible. On the opposite end of the scale, single-impulse envelopes with long IOI's would lose the qualities of an auditory warning, given the latency between onsets. Similarly, impulsive and granular source-sounds (e.g., wooden gears, or a toppled tin can) would not effectively be paired with long IOI's.

Once these restrictions were applied, it was possible to generate an array of contrasting stimuli to be employed in Experiment 2. This was achieved via the Max 6 patch. In the patch, the six environmental sound envelopes are rendered as breakpoint functions (BPF's), as is indicated in Figure 7. The user first chooses from amongst the nine possible source sounds mentioned above. This sound is modulated with one of the six BPF's, selected from a drop-down menu. The BPF duration and triggering/looping rate may be altered by clicking on one of nine IOI durations: six corresponding to the IOI's of the standard car alarms, the three others corresponding to the aforementioned averaged inter-onset intervals. In addition, it is possible to trigger a Morphology Sequencer, which cycles through the six BPF's at the current IOI rate. The Morphology Sequencer loops until it is deactivated.

[13] http://www.mathworks.com/products/matlab/

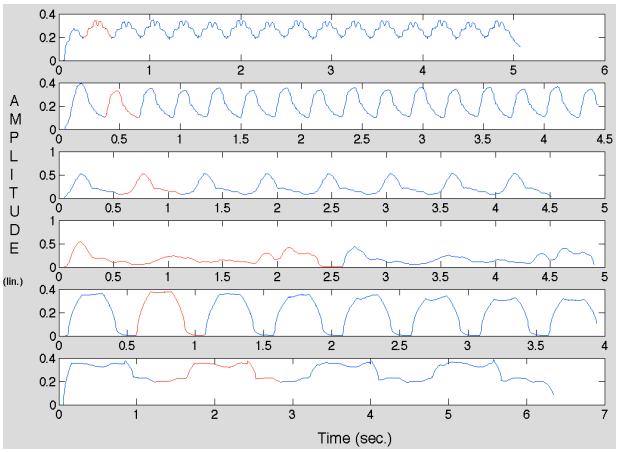

Figure 8. Six car alarm inter-onset intervals (IOI's). For each profile, the blue line represents the whole profile and the red line represents the elementary profile used for computing standard values of IOI's.

8. EXPERIMENT 2: ATTRACTION VS. REPULSION TO SYNTHETIC AUDITORY WARNINGS

As was the case for Experiment 1, Experiment 2 is being conducted at the present time via an online, crowd-sourced format on the IRCAM Sound Perception and Design research team site. [14] Subjects are presented with synthetic auditory warning stimuli, and asked to rate the level of attraction or repulsion to the sound on a three-point scale (attraction/indifference/repulsion). In order to facilitate this task, the instructions to the experiment include metaphors to other sensory modalities (e.g., attraction vs. repulsion to odors).

In the interest of limiting the required completion time for the experiment to 15-20 minutes, forty-five synthetic auditory warnings, rather than all possible combinations of the sources, temporal morphologies, and IOI's described above are employed as stimuli. These stimuli will be selected on the basis of contrast of spectral and temporal characteristics. It is hoped that the results of this experiment will enable us to determine which synthetic auditory warnings would be more effective as deterrents, and which could be utilized as auditory feedback in the context of user interaction with the alarm prototypes.

9. FUTURE WORK

Goals for the short-term include completing Experiment 2 and analyzing the results. Thereafter, we will focus on the interactivity component of the project, and determine the optimal hardware and software development environments in which to pursue this. Once our alarm prototypes have reached a sufficient stage of development, they will be exhibited primarily in site-specific public-space contexts, but also in gallery spaces. The establishment of collaborative relationships with industry partners is in progress, and should be confirmed within the next year or so.

[14] http://recherche.ircam.fr/equipes/pds/projects/asigman/deterrence/src/EvaluationStart.php (accessed 22 September 2014).

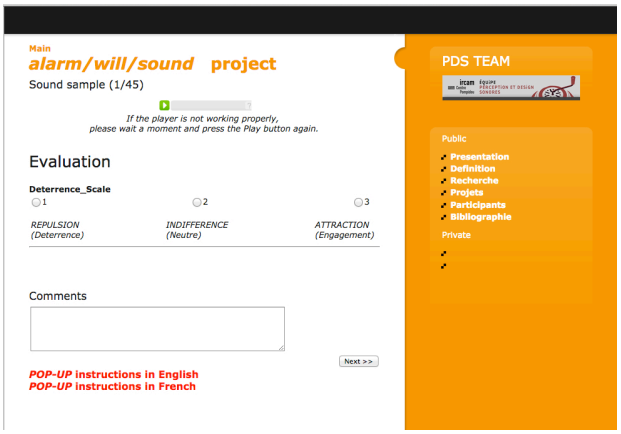

Figure 9. Experiment 2 user interface.

When the prototypes are presented to the public, the interaction models described previously will be featured. Exhibition visitors will be able to trigger alarm sounds remotely via control stations or mobile devices, in a searchable catalog organized and characterized based upon the results obtained from the two experiments and the constructed acoustic descriptor space. User experiences with the various modes of interaction will be documented via video/camera images, data collected from user input at control stations, and survey responses. Ultimately, it is intended that users will be able to upload their own recordings, which may then be edited and indexed according to criteria derived from the experiments.

10. REFERENCES

[1] F. Friedman, A. Naparstek, and M. Taussig-Rubbio, "Alarmingly Useless: The Case for Banning Car Alarms in New York City," *Transportation Alternatives*, 2003.

[2] N. Misdariis, A. Gruson, and P. Susini, "Detectability study of warning signals in urban background noises: a first step for designing the sound of electric vehicles," in *Proceedings of Meetings on Acoustics (POMA)-ICA 2013*, Montreal, 2013, pp. 1-9.

[3] G. Lemaitre, P. Susini, S. Winsberg, B. Letinturier, and S., McAdams, "The sound quality of car horns: Designing new representative sound," *Acta acustica united with Acustica*, vol. 95, no. 2, pp. 356-372, 2009.

[4] C. Suied, P. Susini, and S. McAdams, "Evaluating Warning Sound Urgency With Reaction Times," *Journal of Experimental Psychology: Applied*, vol. 14, no. 3, pp. 201-212, 2008.

[5] P. Susini, N. Misdariis, G. Lemaitre, and O. Houix, "Naturalness influences the perceived usability and pleasantness of a user interface's sonic feedback," *Journal on Multimodal User Interfaces*, vol. 5, no. 3, pp. 52-80, 2012.

[6] P. Schaeffer, *Traité des objets musicaux*. Paris, France: Seuil, 2002.

[7] O. Houix, G. Lemaitre, N. Misdariis, P. Susini, and I. Urdapilleta, "A Lexical Analysis of Environmental Sound Categories," *Journal of Experimental Psychology: Applied, vol. 18, no. 1*, pp. 52-80, 2012.

[8] O. Claude, *La recherche intelligente des sons.* (Masters Thesis.) Universit de Provence, Aubagne, France, 2006.

[9] J.A. Ballas, "Common Factors in the Identification of an Assortment of Everyday Sounds," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 19, no. 2, pp. 250-267, 1993.

[10] G. Peeters and X. Rodet, "Automatically selecting signal descriptors for Sound Classification," *International Computer Music Conference Proceedings*, Gteborg, Sweden, 2002, pp. 455-458.

[11] R.D. Patterson, "Auditory warning sounds in the work environment," in *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 1990, pp. 485-492.

[12] J. Edworthy, E. Hellier, K. Titschener, A. Naweed, and R. Roels, "Heterogeneity in auditory alarm sets make them easier to learn," *International Journal of Industrial Economics*, vol. 41, no. 2, pp. 135-146, 2011.

[13] A. Minard, N. Misdariis, O. Houix, and P. Susini, "Catégorisation des sons environnementaux sur la base de profiles morphologiques," *10ème Congrès Français d'Acoustique*, Lyon, France, 2010, pp.1-6.

[14] K. Yamauchi, J. Choi, R. Maiguma, M. Takada, and S. Iwamiya, "A Basic Study on Universal Design of Auditory Signals in Automobiles," *Journal of Physiological Anthopology and Applied Human Science*, vol. 23, no. 6, pp. 295-298, 2004.

[15] S. Kuwano, S. Namba, A. Schick, H. Höge, H. Fastl, T. Filippou, and M.Florentine, "Subjective impression of auditory danger signals in different countries," *Acoustical Science and Technology*, vol. 28, no. 5, pp. 360-362, 2007.

[16] R.D. Patterson and A.J. Datta, "Extending the Domain of Auditory Warning Sounds: Creative Use of High Frequencies and Temporal Asymmetry," in *Human Factors in Auditory Warnings*, N. Stanton and J. Edworthy (Eds.), Brookfield, VT, USA: Ashgate, 1999, pp. 73-88.

[17] N. Stanton and J. Edworthy, "Auditory affordances in the intensive treatment unit," *Applied Ergonomics*, vol 29, no. 5, pp. 389-394, 1998.

[18] N. Misdariis, A. Minard, P. Susini, G.Lemaitre, S. McAdams, and E. Parizet, "Environmental Sound Perception: Metadescription and Modeling Based on Independent Primary Studies," *EURASIP Journal on Audio, Speech, and Music Processing*, January 2010, Article No. 6.

[19] G. Peeters, A large set of audio features for sound description (similarity and classification), CUIDADO project Ircam technical report, 2004.

11. AUTHORS' PROFILES

Alexander Sigman

Alexander Sigman's award-winning instrumental, electroacoustic, multimedia, and installation works have been featured on major international festivals, exhibitions, institutions, and venues across Europe, Asia, Australia, and the US. In June 2007, Sigman was Composer-in-Residence at the Musiques Dmesures festival in Clermont-Ferrand, France. Subsequently, he was awarded residency fellowships by the Akademie Schloss Solitude (Stuttgart, Germany), the Djerassi Foundation, and the Paul Dresher Ensemble Artists Residency Center. In 2013-2014, he undertook a musical research residency at IRCAM.

Sigman completed his doctorate in Music Composition at Stanford University in 2010. Prior to Stanford, he obtained a BM in Music and a BA in Cognitive Sciences from Rice University. Further postgraduate studies were undertaken at the University for Music and the Performing Arts Vienna, as well as the Institute for Sonology of the Royal Conservatory in The Hague (Netherlands). He is currently Assistant Professor of Composition at Keimyung University in Daegu, South Korea. More information may be found here: www.lxsigman.com.

Nicolas Misdariis

Nicolas Misdariis is a research fellow and the co-head of IRCAM Sound Perception and Design (SPD) team. He is graduated from a college of university level, in 1993, specializing in professional training on mechanics (CESTI-SupMeca). He made a specialization in acoustics within the Acoustical Laboratory of Maine University (LAUM, Le Mans). Since 1995, he has worked at Ircam as researcher where he has taken part in several projects concerning different fields of research dealing with sound science and technology. In 1999, he contributed to the constitution of the IRCAM Sound Design team where he has mainly developed works related to sound synthesis, diffusion technologies, environmental sound and soundscape perception, auditory display, and interactive sonification. Since 2010, he is also a lecturer within the Master of Sound Design at the Fine Arts school at Le Mans.

研究報告

新しいコンピュータ音楽言語LCとその３つの特徴
LC: A NEW COMPUTER MUSIC LANGUAGE AND ITS THREE CORE FEATURES

西野 裕樹
シンガポール国立大学
Keio-NUS Cute Centre

小坂 直敏
東京電機大学
情報メディア学科

中津 良平
シンガポール国立大学
IDM Institute

概要

This paper gives a brief overview of the three core features of LC, a new computer music language we prototyped: (a) prototype-based programming, (b) mostly-strongly-timed programming, and (c) the integration of objects and manipulations for microsound synthesis in its sound synthesis framework. After addressing three problems in existing computer music languages, which are (1) insufficient support for dynamic modification, (2) insufficient support for precise timing behavior and other features with respect to time, and (3) difficulty in microsound synthesis programming, we describe how the three core features of LC correspond to these problems, together with several code examples. Since the design of LC is significantly motivated by the problems that hinder creative practices in computer music of our time, such a language design can contribute to both academic research and creative practices, at least as a design exemplar.

本稿は著者らがプロトタイピングした新しいコンピュータ音楽用言語であるLCについて概説するものである．LCは (a) prototype-based programming(b) mostly-strongly-timed programming とその他時間に関連する機能，および (c) マイクロサウンド・シンセシスの為のオブジェクト及び操作の音響合成フレームワークへの統合という３つの特徴を有する．本稿では，まず既存のコンピュータ音楽言語における (1) 動的変更のサポートの不十分さ，(2) 正確な時間的振る舞いと時間に関する特徴の不十分さ，(3)　マイクロサウンド・シンセシスのプログラミングの難しさの３つの問題を挙げた後，LCのそれぞれの特徴がどのようにこの３つの問題に関わっているかを記述する．LCの設計はこのように現代のコンピュータ音楽の創造的実践を妨げる問題によって動機づけられているため，そのような言語設計は コンピュータ音楽における研究と実践の両面にたいして，すくなくとも デザインの一例として有益であろう．

1. はじめに

本稿では著者らが開発した新しいコンピュータ音楽用プログラミング言語LCの概要を述べる．コンピュータ音楽は様々なコンピュータ技術の発展やプログラミング言語の研究の進展などに影響されつつ，研究され続け現在に至っているが，一方でこのような外部的な要因だけではなく音楽的・創造的な要求や創作の現場で生じた問題も，コンピュータ音楽言語の発展に大きな役割を果たしてきた．同言語の設計過程において著者らは，このような音楽的・創造的な要求もその発展に寄与して来たという視点から，まず現在のコンピュータ音楽の創作活動においてプログラミング言語側の不備により起っていると考えられる問題を同定し，それを新しい言語の設計の際に解決すべき課題と捉えた．

本稿での以降の部分では，まず現在の既存のコンピュータ音楽言語における主要な問題，(1) 動的変更のサポートの不十分さ，(2) 正確な時間的振る舞いと時間に関する特徴の不十分さ，(3)　マイクロサウンド・シンセシスのプログラミングの難しさという３つの問題をあげた後，LCの３つの特徴，(a) prototype-based programming，(b) mostly-strongly-timed programming とその他時間に関連する機能，および (c) マイクロサウンド・シンセシスの為のオブジェクト及び操作の音響合成フレームワークへの統合の３つの特徴に関連させ解説する．なお，本論文は過去に公表された著者らによる論文に記述されている研究の日本向けの要約を意図して書かれたものである．

2. 既存のコンピュータ音楽における３つの問題

2.1. 動的変更のサポートの不十分さ

例えば，近年注目されている Live-coding [4] の演奏では，単にプログラムをステージ上で実行するだけではなく，既に実行中のプログラムの作曲アルゴリズムを変更するなどの動的な変更がコンピュータ音楽のシステムに加えられる事が多い．Kaltenbrunner [12] らの reacTable システムでは，音響生成モジュールのネットワークが演奏中に動的に変更される．

このようにシステムの動的な変更は作曲アルゴリズムと音響生成の両レベルのにおいて必要とされている．しかしながら多くの既存のコンピュータ音楽言語においては，少なくとも作曲アルゴリズムや音響生成レベルの両方もしくはどちらかにおいて動的変更が不可能であったり，様々な制限がかせられている事が多い．動的な変更に対するサポートの不十分さは，コンピュータ音楽言語の設計において，重視すべき問題となってきている．

2.2. 正確な時間的振る舞いと時間に関する特徴の不十分さ

時間的な振る舞いの正確さはリズムのレベルだけではなく，音響合成のレベルにおいてマイクロサウンド・シンセシスなどを行う際にも重要な要素であるが，サン

プル・レートの単位での正確な時間的振る舞いを保証するインタラクティブかつリアルタイムなコンピュータ音楽言語は，いまだそれほど多くなく[1]，多くのコンピュータ音楽言語やシステムで時間的な振る舞いの不正確さが問題となっている．

コンピュータ音楽システムにおいてサンプルレート単位（もしくはコントロールレート単位のものもあるが）で，論理時間にもとづく正確な時間的振る舞いを実現するときには，一般的にはいわゆる the ideal synchronous hypothesis[2] に基づく設計で実現される．これはコンピュータ音楽システムの実装においては，ある時間にスケジュールされたタスクが全て処理が終わるまで論理時間を進めずにブロックし，するべき処理がない場合に限って論理時間が進められ，出力サンプルの計算を行うというかたちで解釈されている．しかし，リアルタイム音響生成を滞りなく行うには，次の出力サンプルを実時間上のデッドラインに間に合うように生成し出力デバイスへと送らねばならないという制限がかせられているため，このような synchronous なアプローチのみに基づくコンピュータ音楽システム上で，時間のかかるタスクが実行された場合，音響処理のデッドラインに間にあわず音響生成が一時的にサスペンドすることになる．

また，正確な時間的振る舞いを保証する言語においても，start-time constraint や execution-time constraint など，時間に関する好ましい特徴が考慮されていなかったり，考慮されていても不十分なものとなっている．FORMULA [1] 等，過去の MIDI シンセサイザとコンピュータで構成された言語・システムの一部においては仕様として含まれていたものであるが，現状でサンプル・レート単位の正確な振る舞いとともに，このような機能が十分な形で実装されているものはない．

このため sychronous な振る舞いによるサンプルレート単位での正確な時間的振る舞いを保ちつつ，このような時間のかかるタスクをどうあつかうかとうことは，現代のコンピュータ音楽言語に提示された問題であり，同様にそこに時間に関する好ましい特徴をどのように組み込みかというのもひとつの考慮すべき課題である．

2.3. マイクロサウンド・シンセシスのプログラミングの難しさ

ユニット・ジェネレータのコンセプトが，それが発明された当時に主流であったアナログ回路による音響合成をもとに行われた抽象化であると見方は，その発案者である Mathews らの記述をみても妥当なものであると言える．[3] 実際にデジタルにマイクロサウンド・シンセシス [8] が行われたのは，ユニット・ジェネレータの登場より遥かに遅い．アナログ回路による音響合成を抽象化した側面が非常に強いユニット・ジェネレータのコンセプトと，小さな音の粒子が互いに重なり合い全体の音響を構成するマイクロサウンド・シンセシスのコンセプトにはかなりの開きがある．このようなギャップは，何らかのマイクロサウンド・シンセシスの技法をユニット・ジェネレータ言語で実装する際にユーザのコードのレベルにおいては，abstraction inversion[2] という，より高レベルの抽象を組み合わせて，低レベルの抽象を表現するという，ソフトウェア開発におけるアンチ・パターンの一つとして現れ，ユーザに対してエキスパート・レベルのプログラミング・スキルを必ずしも期待できないコンピュータ音楽言語においては，マイクロサウンドの領域における創造的な探求を阻害し好ましくないと考えられる．このようなユニット・ジェネレータ言語にマイクロサウンド・シンセシスの技法の実装におけるプログラミングの困難さは解決されるべき課題である．

[1] 例えば ChucK [10] や LuaAV [9] が等がある

[2] “Ideal systems produce their outputs synchronously with their inputs” and “all computation and communications are assumed to take zero time (that is, all temporal scopes are executed instantaneously)” [3, p.360]

[3] 例えば，Mathews らはユニット・ジェネレータについて “conceptually similar functions to standard electronic equipment used for electronic sound synthesis” を持つと記述している [6, p.15].

3. LC における３つの特徴

3.1. プロトタイプ・ベース・プログラミング

既存言語における動的変更のサポートの不十分さにを考慮しするため，LC はプロトタイプ・ベースとしてデザインされた．プロトタイプ・ベース・プログラミング [5] は，オブジェクト指向に属するプログラミング・コンセプトである．クラスというテンプレートからオブジェクトを生成するクラス・ベースの言語と違い，プロトタイプ・ベースの言語においては，まずオブジェクトを無から (ex nihilo) もしくは他オブジェクトのクローニングにより生成し，その後そのオブジェクトにたいして自由に属性やメソッドの追加や変更を，他のオブジェクトに影響を与えずに行える．このような特性によってプロトタイプ・ベースの言語は実行時における動的な変更により柔軟に対応ができる．

LC では，作曲アルゴリズムに関わる部分において，一般のプロトタイプベース言語でのオブジェクトと同様の役割を果たす *Table* を用意しているが，音響合成に関わる部分においては，ユニット・ジェネレータ言語におけるパッチの役割をはたす別のオブジェクト *Patch* を用意している．これは両者の間で想定されるオブジェクトの動作が違うことによる．例えば，オブジェクトをクローンするとき，一般のプロトタイプ・ベース言語のオブジェクトでは，浅いコピー (shallow copy) が行われ，そのオブジェクトが参照しているオブジェクトまではコピーされず，単におなじ参照の値をもつことが多い．また，delegation の機能等は必須とされている．一方，音響合成においてあるパッチをコピーする際は，そのパッチが参照しているユニット・ジェネレータやサブパッチもコピーしなければ意味がなく深いコピー（deep copy）が必要であるが、一方で delegation の機能は特に必要とされない．この様な要求される性質の違いから別のオブジェクトをそれぞれ用意した．Figure 1 と Figure 2 にそれぞれ Table オブジェクト，Patch オブジェクトの使用例を示す．

3.2. Mostly-strongly-timed programming とその他時間に関する機能

既に述べたようにサンプルレート単位の正確な動作を synchronous な振る舞いによって論理時間上では保証ができる一方，時間のかかるタスクによる次回の音響生成の実時間上のデッドラインに間にあわなくなり，音響生成に支障が生じる事態が発生しうる．この問題は「次回の音響処理のデッドラインの前に，全てのタスクの処理がおわり論理時間を進める準備ができていること」という，the ideal synchronous hypothesis の実装時における解釈に違反してしまうことにより起るものなので，単純に synchronous な振る舞いのみを行うシステムでは回避が不可能である．

prototype-based programming example (1)

```
01:  //create a Table object
02:  var obj = new Table();
03:
04:  //then, attach values and functions to its slots.
05:  obj.name = "John";
06:  obj.age = 34;
07:  obj.print = function(var self){
08:    println("name :" ..  self.name ..
09:            ", age:" ..  self.age);
09:  };
10:
11:  //calling the method.
12:  obj.print(obj);
13:  //alternatively, use —> operator
14:  //to abbreviate `self'.
14:  obj->print();
```

The output from the above example (1)

```
name :John, age:34
```

図 1. A Table object example in LC.

prototype-based programming example (2)

```
01:  //create a Patch object.
02:  var p = new Patch();
03:
04:  //store unit-generator objects to its slots.
05:  p.src = new Sin∼(440);
06:  p.dac = new DAC∼();
07:
08:  //build a connection between ugens.
09:  //then, compile the patch to update the graph
10:  p->connect(\src, \defout, \dac, \defin);
11:  p->compile();
12:
13:  //start playing the patch immediately.
14:  p->start();
15:
16:  //wait for 1 second and change the frequency.
17:  now += 1::second;
18:  p.src.freq = 880;
19:
20:  //swap a sine wave osc with a phasor.
21:  var tmp = p.src; //store a sinewave osc to tmp.
22:  p.src = new Phasor∼(440);
23:  p->compile();
24:
25:  //restore a sinewave osc after 1 sec.
26:  now += 1::second;
27:  p.src = tmp;
28:  p->compile();
```

図 2. A Patch object example in LC.

そこで，LC では synchronous programming の一種である Wang らが提案した strongly-timed programming [10] を拡張し，論理時間にたいて synchronous な振る舞いと asynchronous な振る舞いの間で明示的にコンテキストをスイッチできる mostly-stronlgy-timed programming を提案し実装した．mostly-strongly-timed なプログラムでは，スレッドが asynchronous なコンテキストに有る場合，音響生成のデッドラインに間にあうように明示的な論理時間の進行が行われずとも，スケジューラは任意のタイミングでスレッドをプリエンプト可能であり，時間のかかる処理が音響生成に支障をもたらすことを避けられるものとなる．一方で，asynchronous コンテキストではこのように明示的な論理時間の進行が指定されていなくとも，論理時間が進みうるということになる．このようにサンプルレート単位の正確な動作を synchronous コンテキストにおいて行い，時間のかかるタスクは asynchronous コンテキストに行うという，プログラム側での明示的コンテキストの切り替えにより，時間的な振る舞いの正確さを保持しつつ，時間のかかるタスクもバックグラウンドで実行可能とした．Figure 3 のコード例に示すように sync と async の２つの予約語によりコンテキストのスイッチが明示的に行われる．

また LC には，start-time constraint, execution-time constraint, timed message communication など，時間に関連した機能も実装されており，これらも論理時間においてサンプル単位での正確な振る舞いが保証されている．

A mostly-strongly-timed programming example

```
01:  //create an array with 16 elements.
02:  var wsarray = new Array(16);
03:
04:  //load 16 sound files (snd0.wav-snd15.wav)
05:  //and then perform waveset analysis.
06:  //this can be a time-consuming task and
07:  //may suspend real-time DSP temporarily.
06:  for (var i = 0; i < 16; i += 1){
08:    LoadSndFile(i, "snd" ..  i ..  ".wav");
09:    wsarray[i] = ExtractWavesets(i);
10:  }
11:
12:  //`async' block switches to the asynchronous
13:  //context.  The current thread can be preempted
14:  // even without explicit advance of logical time.
15:  async {
16:    //the below code performs the same task again,
17:    //but doesn't suspend real-time DSP this time.
18:    for (var i = 0; i < 16; i += 1){
19:      LoadSndFile(i, "snd" ..  i ..  ".wav");
20:      wsarray[i] = ExtractWavesets(i);
21:    }
22:
23:    //switching to the synchronous context.
24:    //`sync' and `async' can be nested freely.
25:    sync {
26:      //perform some task that requires
27:      //precise timing behaviour here.
28:    }
29:    //the async context (lines 15- ) is recovered.
30:  }
31:  //the end of the asyn context (lines 15-30).
```

図 3. A mostly-strongly-programming example in LC.

3.3. マイクロサウンド・シンセシスの為のオブジェクト及び操作の音響合成フレームワークへの統合

前述のように本研究では既存のユニットジェネレータ言語におけるマイクロサウンド・シンセシスのプログラミングの困難さを abstraction inversion によるものとか分析している．この分析に基づき，マイクロサウンドに直接該当するオブジェクトと関連する操作を直接表現できるライブラリを用意した．Mostly-storngly-timed programming による簡便な論理時間の制御に加え，これらのオブジェクトやライブラリによる簡潔なマイクロサウンドの操作とスケジューリングが可能となり，実現したいタスクよりも高レベルの抽象を組み合わせてプログラムを書く必然性がなくなり，対応する抽象を直接つかって、マイクロサウンド・シンセシスの様々な技法を簡潔にプログラミングすることが可能となった．

実例として，waveset harmonic distortion を行うプログラムを Figure 5 に示す．Waveset harmonic distortion とは，Figure 4 に図示されているように，基音となるもとの Waveset[4] のハーモニクスに重み付けしたのちに，もとの Waveset の上に重ね合わせるものである [8, p.207]．このコード例ではスペースの関係から第２ハーモニクスまでを利用したが，LC ではマイクロサウンドと関連した操作を直接表現できるオブジェクトとライブラリ関数・メソッドを用意した結果，単純で理解しやすいコードとして記述できるようになっている．同様のコードを SuperCollider 等で記述した場合この約２倍の長さとなり[5]，*Pbind* その他複数の SuperCollider 特有の知識を必要とするパターン・オブジェクトを利用

[4] a waveset is defined as a “distance from zero-crossing to a 3rd zero-crossing” whereas wavecycle is defined as“wavelength of sound, where clearly pitched” [11, p.50].

[5] 具体的なコード例は [7, pp.49-50] に示されている．

するためコードも複雑になりユーザにも相応の知識が要求される．

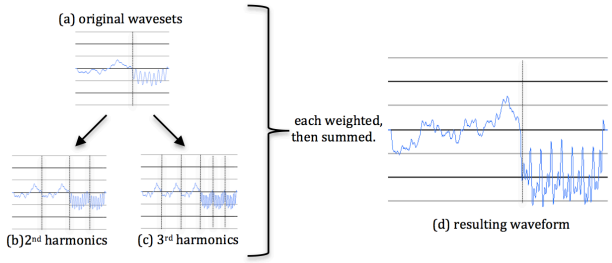

図 4. A pictorial representation of waveset harmonic distortion.

```
01:  LoadSndFile(0, "/sound/sample1.aif");
02:  var wavesets = ExtractWavesets(0);
03:
04:  //weights for 2nd harmonics
05:  var weight2 = 0.5;
06:
07:  for (var i = 0; i < wavesets.size; i +=1 ){
08:    //create the 2nd harmonics
09:    //from the original waveset.
10:    var ws = wavesets[i];
11:    var harm2 = ws->resample(ws.size / 2)
12:                  ->amplify(weight2);
13:
14:    //schedule the original waveset for output.
15:    PanOut(ws, 0.0); //0.0 is the center (stereo).
16:    //then, the 2nd harmonics (two in series)
17:    PanOut(harm2, 0.0);
18:    PanOut(harm2, 0.0, offset:harm2.dur);
19:
20:    now += ws.dur;
21:  }
```

図 5. A waveset harmonic distortion example in LC.

4. 結論

本稿では，既存のコンピュータ音楽言語における３つの問題，(1) 動的変更のサポートの不十分さ，(2) 正確な時間的振る舞いと時間に関する特徴の不十分さ，(3)　マイクロサウンド・シンセシスのプログラミングの難しさを提示・分析し，これを新しいプログラミング言語の設計開発の機会としてとらえ設計された LC 言語において，それらを解決する３つの特徴，(a) prototype-based programming，(b) mostly-strongly-timed programming とその他時間に関連する機能，および (c) マイクロサウンド・シンセシスの為のオブジェクト及び操作の音響合成フレームワークへの統合の概要を述べた．言語設計に関する問題については様々な解決が存在するであろうが，LC におけるこれらの特徴はコンピュータ音楽における創作支援および言語設計の研究にたいして，少なくともデザインの一例として有益であると考える．

5. その他

本稿ではページ数の制限から，より深く議論することはできなかったが，公表済みの論文（[7] などが `http://nus.academia.edu/HirokiNishino` からダウンロード可能である）においてより詳細に記述されているため，必要であればそちらを参照していただきたい．

6. 参考文献

[1] Anderson, D. P. and Kiuvila, R. “Formula: A programming language for expressive computer music”. *Computer Vol.24 (7)*, 1991.

[2] Baker, T. “Opening up ada-tasking”. *Proceedings of the 4th International Workshop on Real-time Ada Issues*, 1990.

[3] Burns, A. and Wellings, A. J. *Real-time systems and programming languages: Ada 95, Real-time Java and Real-time Posix*, Addison Wesley, 2001.

[4] Collins, N. et al. “Live coding in laptop performance.” *Organised Sound Vol.8 (03)*, 2003.

[5] Noble. J et al. *Prototype-based Programming: Concepts, Languages and Applications*, Springer-Verlag, 1998.

[6] Matthews, M.V. et al. *The Technology of Computer Music*. The MIT Press, 1969.

[7] Nishino, H. *LC: A Mostly-strongly-timed Prototype-based Computer Music Programming Language that Integrates Objects and Manipulations for Microsound Synthesis*. Ph.D. thesis, National University of Singapore, 2014.

[8] Roads, C. *Microsound*, 2004.

[9] Wakefield, G., Smith, W., and Roberts, C. “LuaAV: Extensibility and Heterogeneity for Audiovisual Computing.” *Proceedings of the Linux Audio Conference*, 2010.

[10] Wang, Ge. *The chuck audio programming language. a strongly-timed and on-the-fly environ/mentality*. Ph.D. thesis, Princeton University, 2008.

[11] Wishart, T. *Audible Design: Appendix 2*, Orpheus Books LTD, 1994.

[12] Kaltenbrunner, M. et al. “Dynamic patches for live musical performance.” *Proceedings of NIME*, 2004.

創作ノート

INTERACTIVE PROCESSES BETWEEN SOUNDS AND IMAGES IN VISUAL MUSIC
ヴィジュアル・ミュージックにおける サウンドとイメージ間のインタラクティヴ・プロセス

Wilfried JENTZSCH
Composer/Visual Music Artist, Electronic Music Studio HfM Dresden (i.R)
作曲家/ヴィジュアル・ミュージック作家、元ドレスデン国立音楽大学/同大電子音楽スタジオ

概要

Audiovisual composition becomes more significant in our today's artistic creation. Visual artists on the one hand, and composers of electroacoustic music on the other hand, are more and more involved in the visualization of music. The term Visual Music distinguishes it from other audio-visual works in the way to give images and music the same importance. What does it mean, the same importance in both media?

Maura McDonnell wrote as follows:

"A visual music piece uses a visual art medium in a way that is more analogous to that of music composition or performance. Visual elements are composed and presented with aesthetic strategies and procedures similar to those employed in the composing or performance of music."[1]

Andrew Hill, composer and visual music artist, has given a definition about technology-based Visual Music relating to electroacoustic music:

"An electroacoustic audio-visual music work could be defined as a cohesive entity in which audio and visual materials are accessed, generated, explored and configured, primarily currently with the use of computer-based electronic technology, in the creation of a musically informed audio-visual expression. Electroacoustic audio-visual music works explore the possibilities that the combination of their two time-based media (sound and moving image) allow."[2]

In this paper I will present three conceptions of my audiovisual compositions and discuss about some remarks revealed through them.

音響映像作品 (audiovisual composition) の作曲は今日の私達の芸術創作においてより一層意義深いものとなって来ている。一方では映像作家、他方では電子音響音楽作曲家が、ますます音楽の映像化に関わるようになっている。ヴィジュアル・ミュージックという用語は、イメージと音楽とに同等の重要さを与えるという在り方によって他の音響映像作品と一線を画している。双方のメディアにおいて同等の重要さを持つということはどういう意味なのだろうか。

ヴィジュアル・ミュージック作家 Maura McDonnell は次のように書いている：

"A visual music piece uses a visual art medium in a way that is more analogous to that of music composition or performance. Visual elements are composed and presented with aesthetic strategies and procedures similar to those employed in the composing or performance of music."[1]

作曲家でヴィジュアル・ミュージック作家の Andrew Hill は電子音響音楽に関連させてテクノロジーをベースとしたヴィジュアル・ミュージックについて次の定義付けをしている：

"An electroacoustic audio-visual music work could be defined as a cohesive entity in which audio and visual materials are accessed, generated, explored and configured, primarily currently with the use of computer-based electronic technology, in the creation of a musically informed audio-visual expression. Electroacoustic audio-visual music works explore the possibilities that the combination of their two time-based media (sound and moving image) allow."[2]

この記事では私自身のオーディオヴィジュアル作品から三つの構想について紹介し、それらの制作を通して現れた幾つかの特記すべき点について議論していく。

1. 背景

1978 年から 1981 年にかけて私はパリの CEMAMu (Centre d'Etude de Mathématiques et Automatiques Musicales) のメンバーだった。Iannis Xenakis が UPIC (Unit é Polyagogique Informatique CEMAMu) を開発しており、私もグラフィックとサウンドについての研究に関わった。Xenakis による UPIC システムの定義は下記の通りである：

"UPIC est un syst è me informatique conversationnel unique, permettant de composer de la musique en dessinant sur une table à dessin é lectrique de grandes dimensions, connect é à un mini-ordinateur."[3]

90 年代始め、ドレスデン音楽大学電子音楽スタジオ所長として私は GraphicComposer の開発を指揮した [4]。GraphicComposer は Macintosh コンピューターのために開発された、MIDI を介してグラフィックをサウンドにトランスフォームするソフトウエアであり、マルチメディア・プレゼンテーションや作曲の創作ツールとして作曲家達によって使われた。GraphicComposer にはその本質的な部分である様々なトランスフォーメーションの機能が搭載されていた。最も重要なのは『ミラー』『回転』『ストレッチ』『傾斜』『屈曲』、そしてこれらの機能を水平線と垂直線に対して設定できる選択肢があった。

これらのソフトウエアは両方とも音楽を違った視点 - グラフィカルな構造 - から作り出すツールだった。この構想から、創作への新たなアイデアが生まれた- 例えば日本の文字を音楽に '置き換える' ことだ。この（音とイメージ間の）二重規律的創作は結局私をヴィジュアル・ミュージックに導くことになった。

2. 私自身のオーディオヴィジュアル作品

次のオーディオヴィジュアル作品の例で、三つの構想を紹介したい。

- Imaginäre Landschaft (Imaginary landscape) [5]
- Kyotobells [6]
- Particle world

2.1. Imaginäre Landschaft (1980/2009)

UPIC を使って音楽的構造を作り出すためには、マグネット式ペンシルでエレクトリックボードに描く。次のパラメーターがグラフィカルに定義されなくてはならない。

- 音色としての波型
- 相対的エネルギーの変化としてのアンプリチュード
- 絶対的エネルギー (pppp から ffff まで)
- 65Hz から 2093Hz までの連続した周波数帯域を持つ X 軸上に決定されたピッチ
- 継続時間 Y 軸 (進行方向は左から右に向かう)
- 絶対的継続時間 (1 から 60 秒)

これらのすべてのパラメーターはコンピュータによって計算され、スピーカーによって結果を聴くことが出来る。この創作の出発点は木々や根や葉といった自然から着想したグラフィックスである。これらはグラフィックボード上に描かれることを介して周波数と時間の二次元的要素を持つサウンドへと翻訳された。全てのページから作られたサウンドは最終的に作品『Paysages A 938』を作るためにミックスされた。この作品は、五人の歌手のためのヴォーカルパートがテープとオーバーラップする電子音響のために作曲された。テキストは老子の Tao te king『道徳經』の一節に基づいている。この作品はケルンの西ドイツ放送による委嘱で、1980 年のヴィッテン現代室内楽音楽祭で初演された。ヴォーカルパートはコレギウム・ヴォカーレ・ケルンによる。

それから 28 年後、オリジナルのグラフィックス（静止画）に時間構造を与えてヴィジュアル・ミュージック作品を作ろうというアイデアが生まれた。この音楽の元々の性格は Max/MSP による音処理により時間的に修正された。白黒のグラフィックスという元々の性格はトランスフォーメーションに関わらず終始一貫してそのまま変わらない。一方、様々に異なる明るさとパースペクティヴでイメージが動くことはこの作品にとって重要な点である。オリジナルのイメージはどんどん抽象化していき最後にはオーディオパートのノイズ・グレインに呼応するようにグレインの雲となる。オリジナルのグラフィックスは木を象徴しており、その中にある線は継続する周波数レンジの glissandi を作り出している (図.1)。もう一つの例は Evolve という機能を使って作られたオリジナル・イメージからランダム・イメージのエヴォリューションであり、この機能では 'ランダム' と '繰り返し回数' の総計を決定することができる（図.2)。これは素材イメージ全体が乱気流を経て最後にはカオス的状態となるというプロセスである。Evolution は開始点と最終点の間を補填するタイムラインによってコントロールされたシークエンスと性格づけることが出来る。トランスフォーメーションはソフトウエア Studioartist によって行なった。

音楽とイメージは、同一の素材から作り出されているので、密接に関係している。グラフィックスは電子音響を作り出し、同じグラフィックスが動く映像を作り出した。両メディアが『結婚』して目と耳のための新しい感覚受容を作り出したのである。

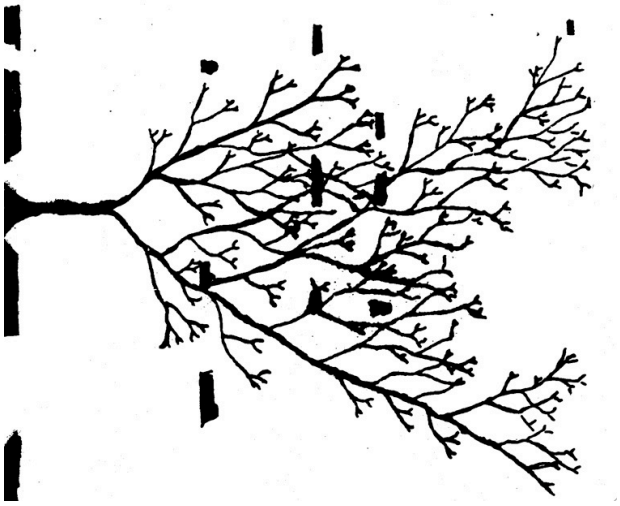

図 1. tree

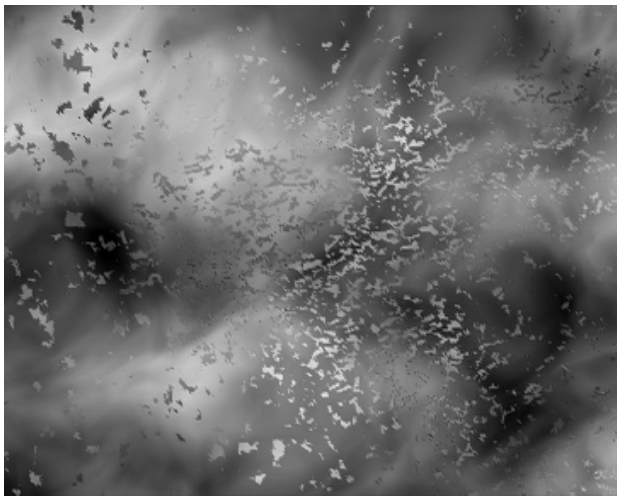

図 2. random

2.2. Kyotobells (1994/2006)

1993 年、私は京都郊外の或る小高い丘にあるお寺を訪ねた。そこで木に吊るされた幾つかの風鈴の美しい響きを聞いた。風が吹いていて、風鈴は長い余韻を響かせていた。それらは高い音低い音、それぞれ異なる音の高さで、風のエネルギーがそれらを意のままに奏で、時折無音の沈黙の時に遮られてはまた響かせていた。私はその音楽を長い間聴き入っていた。様々な音色、音の厚み、仕草、そして空間での動きなど、脳裏に次々と新しいサウンドが現れた。私は今でもこの瞬間をよく覚えている。それは私の新しい電子音響音楽作品 Kyotobells の創作の始まりだった。

この作品は四つの小さな風鈴の音によって構成されている。曲はひとつの音で始まり、様々な数と音高の鈴が作り出すアンサンブルは次第に音楽を濃厚にしていく。

図 3. 風鈴

glissandi が空間を開き、豊かな濃厚なひびきが柔らかに始まり、静止した性格を作り出し、最終的にはノイズに至る。これはひとつの鈴のハーモニック音からノイズへの発展であり、最後にはまたハーモニック音に戻る。単音とノイズの間を、加工処理された音の様々な段階がその二つの構成要素の橋渡しをしている。

この作品の最初のヴァージョンは京都市 1200 年記念のために委嘱され、京都市国際会議ホールにて初演された。その 10 年近く後、この作品にヴィジュアル部分を作るアイデアが生まれた。

青色と白色の四角形がヴィジュアルパートをジオメトリックに構成する。サウンドのアンプリチュードとスペクトラが形と色の明度を調整し、この四角形をトランスフォームする。すべてのトランスフォーメーションはインタラクティヴに行なわれ、そのもっとも影響力のあるパラメータはアンプリチュードである。インタラクティヴのプロセスは、プログラムを通して自動的に行なわれ、親密でいきいきとしたサウンドとイメージの関係を作り出す。一方で、同じプロセスの継続的な繰り返しは、すぐに飽きられてしまうという危険を帯びている。そのため、

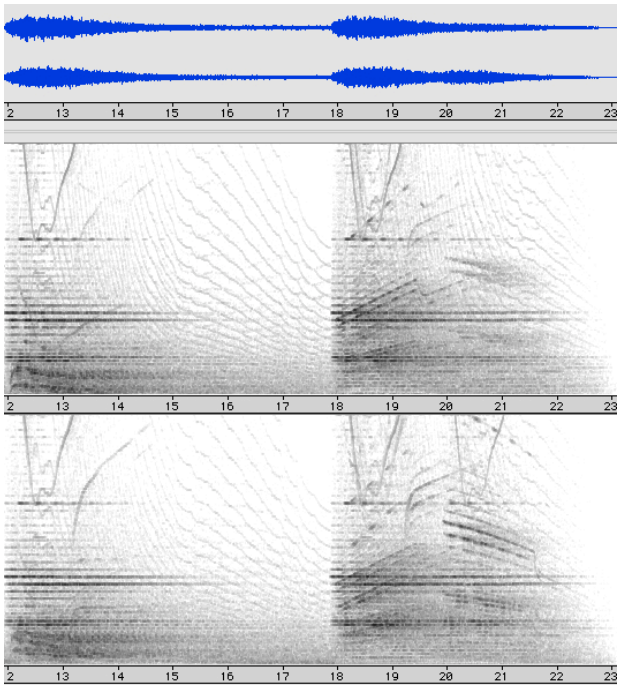

図 4. sonogram of bell glissandi

- 自動的にサウンドの特徴がヴィジュアルのパラメーターをコントロールする
- 手動によりサウンドの特徴がヴィジュアルのパラメーターをコントロールする設定をする

の二つの種類のインタラクティヴの方法を使った。

二つのインタラクティヴのモード、プログラム自体を通しての自動コントロールとアーティストによる手動でのコントロール、はリアルタイムでのアクションとリアクション間のおもしろいプレイである。このため（ヴィジュアル部分の作成は）最初から最後までノーカットで作られた。幾つかの違ったヴァージョンを作成し、その中でもっとも面白いとおもわれるテイクを最終ヴァージョンとして採用した。

最も重要な手動コントローラーは、

- fade in/out
- 四角形の filled in
- 波型による構造のトランスフォーメーション (bending、twisting)
- スピード変化 (slow/fast) とランダム

ヴィジュアルパートは Geißというソフトウエアによって制作され、2007 年の Visual Music Marathon Boston で初演された。

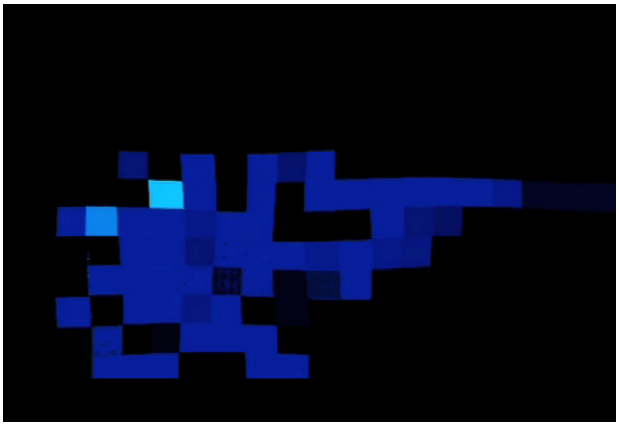

図 5. 四角形のアンサンブル

2.3. Particle World (2013)

このオーディオヴィジュアル作品ではイメージと音楽が Particle Synthesis という聴視感覚の最も小さいユニットのレベルで互いに親密に関連している。ヴィジュアルのパーティクルの数は 1800 〜 25000/sec で、エヴォリューション速度、オパシティ (不透明度)、複雑度といった補助パラメーターによって構造を修正出来る。パラメーターのひとつの値から別の値への間は、インターポレーション（補填）によって計算される。ユニークな音源素材は中国のリュートで、グラニュラーシンセシスによりトランスフォームされ、スペクトラル・エクストラクションとスペクトラル・コンプレッションにより音加工した。ハーモニック部分とノイズ部分の二つの新しいサウンドはスペクトラ抽出により得たものである。アンプリチュードは高い周波数帯域と低い周波数の二つの周波数帯域において別々に計算され、キーフレームに変換される。キーフレーム Z における結果値はイメージの中心に置かれた『カメラポジション』の動きを生み出す。

両メディアはインタラクティヴな関係にある。サウンドの特徴は、オーディオのアンプリチュードが 3D 回転する (x,y,z 軸を持つ）カメラの動きをコントロールしているパーティクルに直接反応する。このインタラクティヴィティは (三次元イメージとしての) 空間と時間において密度と色彩を変化させながら複雑な動きを作り出す。

この作品はオーディオとヴィジュアルの相互協力関係により活き活きとした表現を造り出そうという実験だ。それは先端デジタル技術を用いることによってのみ可能である。Let’s “See the sound (さあ音を見よう)“.

この作品は After Effects (AE) プラグインによって制作された。

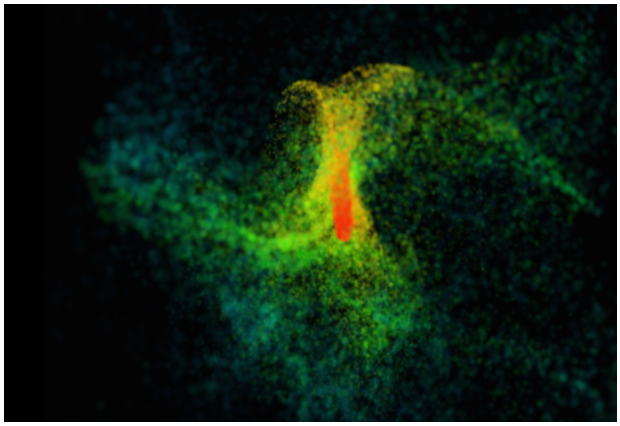

図 6. particle world

3. まとめ

この記事の冒頭の「二つメディアにおける同等の重要さとは何か」という質問に戻って、それに対してサウンドとイメージ加工処理における似通った方法（手法）は美学的な見地からして親密な関係を作り出すと言うことが出来るだろう。オーディオヴィジュアル作品の創作に塾達すればするほど動画とサウンドは類似性によって関連づけられることを明確に理解することが出来る。例えば、自分の創作経験を通して下記の幾つかの点に気がついた。

- サウンドと動画における時間の流れ
- 両メディアにおけるイヴェントの速度
- 色彩と音色の間の表現
- サウンドのアンプリチュードとイメージの明度のバランス
- イヴェントの密度
- イメージとサウンドにおける緊張と弛緩の関係
- 音楽とイメージにおける構造の似通った性格特徴 (例えば パーティクルとグレイン)
- 両メディア間のインタラクティヴィティ

ヴィジュアル・ミュージックのような新しいメディアアートは、イメージと音楽創造の新しい構想を必要としている。私の音楽においては、それは “ビートとメロディ” でもないし、また伝統的なハーモニーでもない。グラフィックソフトウエア (UPIC や GraphicComposer) により、自然にインスピレーションを得て音楽と動画における新しい構造を発案してきたが、形を持った画像（例えば木のような）は、秩序から無秩序へ、そして最終的に全くあたららしい異なった秩序に到達するトランスフォーメーションのプロセスの出発点として（またはその逆のプロセス）使われただけである。この音楽とサウンドにおけるプロセス中心の創作は聴取者の受容に緊張感を作り出す。目で聴き、耳で見ること、それが私のヴィジュアル・ミュージック創作の意図である。

（日本語訳：石井紘美）

4. 参考文献

[1] Maura McDonnell. The Program of Visual Music Marathon Boston 2007

[2] Andrew Hill, 2010. http://cec.sonus.ca/econtact/12_4/hill_reception.html

[3] Definition by Iannis Xenakis (CEMAMu)

[4] Wilfried Jentzsch, KlangArt-Kongreß Osnabrück 1995, Musik und Neue Technologie 1

[5] Imaginäre Landschaft. http://vimeo.com/27490864

[6] Kyotobells. http://vimeo.com/55464696

5. 著者プロフィール

Wilfried JENTZSCH (ヴィルフリート・イェンチ)

1941 年生まれ。ドレスデン音楽大学、ベルリン芸術アカデミーにて作曲を、ケルン音楽大学にて電子音楽を学ぶ。1976 年から 1981 年までパリ・ソルボンヌ大学にてイアニス・クセナキスに師事、音楽美学の分野で博士号を取得。同時期に IRCAM と CEMAMu にてデジタル音響合成の研究を指揮する。1993 年から 2006 年までドレスデン国立音楽大学作曲科教授および同大電子音楽スタジオ所長。スイス・ボスヴィル国際作曲コンクール、パリ Art&Informatique、ブルージュ IMEB、ZKM World Music Days にて優勝、入賞。作品はワルシャワの秋 (4 回)、ヴィッテン音楽祭 (4 回)、ダルムシュタット音楽週間などの主要な現代音楽祭にて、またブールジュ音楽祭 (5 回)、Musica Viva リスボン, EMUfest ロ一マ, MusicAcoustica 北京、IRCAM、GRM、ZKM、Musiques&Recherches ブリュッセルなど、多くの電子音響音楽祭や研究機関にて入選、招待されている。イェンチの Visual Music 作品は ZKM、VM Marathon ボストン&NY、メルボルン Cinema 祭、北京 MusicAcoustica、ローマ EMU 祭、モントリオール Cinema Nouveau など世界各地のフェステイヴァルで紹介され、彼自身も VM のキューレーターとして活動している。ブルージュ IMEB、ブリュッセル Musiques&Recherches、ZKM、GRM、オハイオ州 Capital University Colombus にて客員作曲家。現在ケルン郊外に在住。ISCM 会員、DegeM 創立会員。

創作ノート

ツィルコニウムを使った三次元空間アクスマティック作品『梁塵譜』
ACOUSMATIC COMPOSITION “RYOJINFU” IN 3D WITH THE ZIRKONIUM

石井 紘美
Hiromi ISHII
作曲家/国際芸術アカデミー・ハイムバッハ
International Kunstakademie Heimbach

概要

現在の電子音響音楽において実際的な音響空間をどのように構築するのかという問題は作曲構想の初期段階から関わってくる。著者は劇場や先端技術のエキジビションにおける音楽と効果音の制作、および現場での音響監督を経験したことから、電子音響音楽の空間構築への興味をスタートさせ、その後自身の作品において様々な構成の多チャンネルアクスマティック作品を作曲してきた[1]。

そして２０１３年にはZKMでの二度目の客員作曲家滞在を通じ、同センターで開発された3Dサウンドディフュージョン・システム・ツィルコニウム Zirkonium v0.9.38beta を使い三次元空間によるアクスマティック作品『梁塵譜』Ryojinfu を制作する機会を得た。この記事では、その作曲の過程を解説する。

How to structure the real acoustic space is a question which arises in early stages of an acousmatic composition. The author started her interest in sound space-designing through some work-experiences as a composer for theater performances, and as a sound director for exhibitions of cutting-edge technology . This interest has been developed later in composing acousmatic works with various multi-channel conceptions. 2013 the author has been invited for her second residency at ZKM to compose Ryojinfu, the 3D acousmatic work, using the 3D sound diffusion system Zirkonium v0.9.38beta invented by ZKM. This article explains its compositional process.

[1] 石井のアクスマティック作品系列に Dreaming Stones (ダブル・クアドラフォニック)、Summer Grasses (6チャンネル)、Himorogi シリーズ（日本の楽器と多チャンネル)、Ginn シリーズ特に Ginn-Tokio 2006 (ダブル・クアドラフォニック)、SHINRA (オクタフォニック）がある。

1. 多チャンネル音響空間に関わる歴史的背景

電子音響音楽の空間はそれを作り出す技術の発展の上に成り立っている。モノラルのミュジック・コンクレート初期からステレオへ、そして多チャンネルへと空間に対する認識は広まった。とくに日本では1970年の大阪万博のドイツ館、鉄鋼館でのイヴェントにより多チャンネル音響空間は一般社会においても広く経験されるところとなった。フランスのGRMでは1970年代初期にフランソワ・ベルによりスピーカーアンサンブルAcousmoniumが考案され、以後フランスにおける新たな電子音響音楽の形として広まった。また、ベルギーのレオ・キュッペルは、1977年ローマにおけるThe sound cupola in Romeを皮切りとして360度を15度ごとのパラメーターとするMIDIベースのサウンド・ドーム・システムを実現させた。

しかしながら、これらの三次元空間は概ねコンサートでのサウンドディフュージョンに留まり、これを作曲過程で作品の表現要素（パラメーター）として組み込むまでには至っていない。その理由として第一に、作曲家がその創作過程で初段階からこのような大スピーカーアンサンブルを使うことはほとんど出来ないこと、第二に、音響空間や音像移動を決定しても、それを記憶させることが出来るようなタイムラインを持ったソフトウェアが無かったことが挙げられる。

2. ZIRKONIUM と KLANGDOM

2000年以降、ZKMで開発され続けているツィルコニウム Zirkonium は、こういった問題点を解決し、作曲過程の初段階からコンサートでのサウンドディフュージョンまでを作曲家が一貫した方式で音響空間作曲できるシステムとして考案された。開発チームのラマクリシュナン C. Ramakrishnan、グロスマン J.Großmann、

ブリュマー L. Brümmer は、NIME06 で発表された論文『ZKM Klangdom』の中で電子音響音楽の歴史の中で多チャンネル音響空間へのアイデアの発達は総括するとアクスマティック的アプローチとシミュレーション的アプローチの２つに分けられると解説する。そしてシミューレーション的アプローチの例として、IRCAM の SPAT、Ambisonic、Wave Field Synthesis を挙げている (Ramakrishnan 2006)[2]。

ZKM が開発している Zirkonium システムは主として ZKM のクラングドーム Klangdom（または Kubus) を使用空間として開発されて来た。クラングドームは「クアドラフォニック、5.1、オクタフォニック、16 チャンネルなどのスタンダードな形式の音楽を上演する為に常時良く整えられている。過去には客席の上部にスピーカーが吊るされることもあったが、音響に没頭することを可能ならしめるのに限界があった。さらにはチャンネル数の増加によりマルチスピーカー環境をコントロールするにはミキシング・コンソールとは別の楽器が要求されると思われた。この切なる欲求が Zirkonium というプロジェクトに繋がった」(Ramakrishnan 2006)。

Zirkonium は２００６年にはほぼ完成されたが、初代プログラマーのラマクリシュナン Ramakrishnan の仕事を、現在 D. ヴァグナー D. Wagner が引き継ぎ、さらなる開発を続けている。

クラングドームでは現在 43 の MeyerUPJ-1P スピーカーが設置されており、これとは別に 23 のスピーカーによるスタジオ・ミニドームがある。クラングドームはコンサートホールなので、作曲家はここで作曲の初段階から仕事を始めるわけにはいかないが、ある程度の形が出来た段階でミニドームを使って作品を仕上げることが出来る。その『ある程度の形が出来る』までのもっと初期の段階では、作曲家はヘッドフォンを使い Zirkonium システムをインストールしたコンピュータでシミュレーションによるサウンドを聴きながら作曲を進める。つまり、この段階であれば自宅でも制作可能ということになる。

3. 作品『梁塵譜 (RYOJINFU)』の作曲

3.1. 着想・リテラリーな素材と構成への利用

著者は、日本の伝統音楽に基づく或は関連させた電子音響音楽の作曲をライフワークとしている。本作品もその系列に位置する[3]。

図 1. Klangdom と制作作業中の著者

また、著者は音と視覚的イメージをほぼ同時に着想するので、そのうちのどちらかを発展させ作品化する、或は同時進行的に制作し音響映像作品とするなどの方法を取っている。またリテラリーな着想をも伴ってそれをおよその音楽展開の方向作りに利用することもある。

Ryojinfu (梁塵譜）は、故柴田南雄氏の創作アイデアへの協力者であった柴田純子夫人から伺った『梁塵秘抄』の話とたまたま録音した穀物のグレインサウンドが結びついてきっかけとなった。『梁塵秘抄』は後白河天皇により編纂された今様集のタイトルである[4]。

『梁塵秘抄』は「漢の虞公（ぐこう)、斉の韓娥（かんが）はひとを感動させる名手であった。彼らの歌声のひびきに梁の塵も感動し、静かに舞い上がって三日の間は落ちて来ないといわれた」という中国の故事にちなんで後白河天皇が付けたタイトルである。後白河は一方で年を重ねるごとに頻繁に熊野詣でをするようになっていった。出家した年の熊野詣での折り、今様を歌っていると突然麝香の香りが漂い、風もないのにご神体の鏡が揺れ動き、まるで誰かが通り過ぎて奥へ入って行ったかのように思われたという。この後白河自身の霊異な体験の逸話と前述の中国の故事から、Ryojinfu の『歌声とそれに呼応する霊達の動き』という音楽的な着想が生まれた。また、今様を歌ってさえいれば幸せだったという後白河天皇の仏教心とそれに反した策謀と戦の人生を対照させ、全く異なる二つの世界のあつれきの高まりを空想し、その空想を音楽的展開に利用した（実際の作曲作業

[2] “The ZKM Klangdom”, NIME06. C. Ramakrishnan, J.Großmann, L. Brümmer.

[3] 著者の作曲において関連の作り方は作品により様々であり、単に日本の楽器や音楽を使うというだけではない。例えば尺八の音色分析を楽曲構成に応用した尺八のための Mixed music 作品『Wind Way』、音素材は全く音美学や音文化に関わらない騒音を用いながら、尺八楽や能楽におけるひびきの構造と音の美学を応用した acousmatic 作品『Summer Grasses』などがある。

[4] 後白河は母の影響で少年時代から今様に熱心であり、その甲斐あって当代一の今様の名手となったが、政治抗争のため第四皇子でありながら天皇に即位を余儀なくされたといわれる。しかし、策謀家としての才を発揮して院政を敷き、権力を誇示した。一方で後白河は大変仏教心の厚い人物であったともいわれている。今様とは今風、当世風の意味であり、雅楽のメロディーに歌詞を付けて歌うという当時流行っていた歌謡のこと。

は音の聴覚的印象に基づいて純音楽的に行なわれた)。

2013 年 3 月 ZKM クラングドームにおける著者の別の 8 チャンネル・アクスマティック作品初演の際、Zirkonium を使用するうちに、当時構想中だった Ryojinfu には Zirkonium の持つ空間性が最適の表現手段であると確信した。

3.2. 音素材と音加工

前節で述べた着想と楽想展開に基づき、典型的な今様の始まり方であるかけ声「やあー」の部分、米・豆・粟などの混ぜ合わせによるグレインサウンド、仏教の式典のライブ録音の一部（読経と雑音・騒音の混ざり合い）を主に音素材に応用した。「やあー」の声はオリジナルとしては使われず、他の素材とともに Audiosculpt を始め、Convolution、Mutation、Evolution、Fusion などを使い加工し、単声、合唱、また風のような音に、少年の声のように変容する。声を素材としたこれらの加工音は主にゆっくりと全空間を使っての音像移動をする。また穀物グレインサウンドは名手の歌声に反応する音楽的な塵をイメージしており、二種類の Granular Synthesis プログラムと GRM Fusion を主に使って加工、空間において鳥のように、または妖精のように自由自在に素早く動き回る。また、グレインサウンドの『ぶつかり合う音』としての性格は曲の後半において『ぶつかる重々しい騒音』『破壊的で暴力的な騒音』に移行拡大し、それは激しさを増す心理的葛藤（これも一種のぶつかり合い）を暗示する。これらの騒音は主に低い位置の２スピーカー間で音像移動する。終結部の少年のような声（ピッチシフトではなく、篳篥との Convolution による）は後白河の回顧（懐古）を暗示するように現れ、繰り返しながら天空を舞い、音像を広げながら消えて行く。

3.3. 作品の工程

前出の Ramakrishnan、Großmann、Brümmer のペーパーにより例として挙げられているシステムのうち、著者は過去に 4 チャンネル、8 チャンネル、8ch 以上のマルチチャンネルシステム、Acousmonium、Ambisonic、Zirkonium により自作品を上演した経験を持っている(ホールやスピーカーはそれぞれ異なる)。これらの聴取経験と過去の職業的音響経験をもとに、今回 ZKM の Zirkonium システムを用いて『Ryojinfu』を作曲したが、その経緯は、大もとの楽想の段階から大きく異なっていた。

著者は Zirkonium によるサウンドディフュージョンを経験して、他のシステムと大きく異なる特徴は空間の３次元性であると了解した。Zirkonium は、座った聴衆のおよそ耳の高さとする第一リングを最も低いリングとし、その上方にスピーカーを結ぶ架空の線による半球空間を作る。ZirkoniumDom43 のスピーカー・セッティングを例に取ると、半球内の仰角を５段階に分け、それぞれを『リング』と称している。仰角は z で表され、底辺は z 値はゼロ、天頂は 1 となる。半球の底辺（およそ座った聴取者の耳の高さ）の円周に １４個のスピーカーを配置した第一リング、１４個のスピーカーによる第二リング、更に上方に８個のスピーカーによる第三リング、６スピーカーによる第四リング、そして真上に一つのスピーカー（リングではないのだが、第五リングと呼ぶ）からなる構造を持つ。Zirkonium によって与えられた、この『高低』という第三次元の軸をどのように活かすことが出来るかという点は、創作上の大きな課題であった。

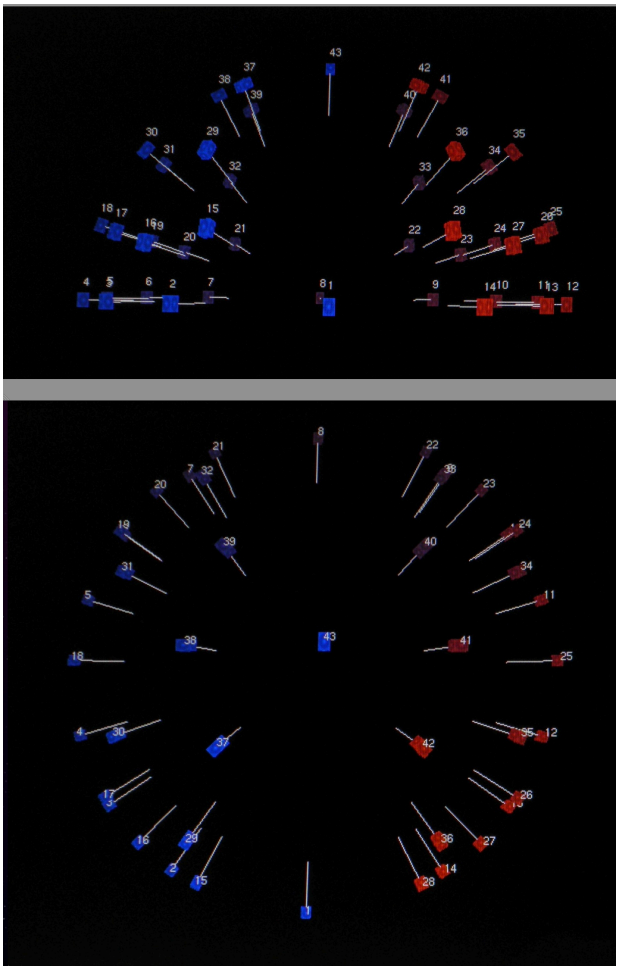

図 2. Zirkonuim のスピーカー配置図。後ろ視点（上）、真上視点（下）

幾つかのアイデアを挙げると、作品冒頭では、『やあー』という声を加工したサウンドが前方左右からやや中央よりに集まり、上方へ立ち上る。続いてグレインサウンドが回転しながら舞いあがる。これらの動きは Zirkonium でなければ出来ない表現である。このよ

うに、音源の音加工とその音像移動へのアイデアは作曲の初期段階からほぼ同時に存在した。

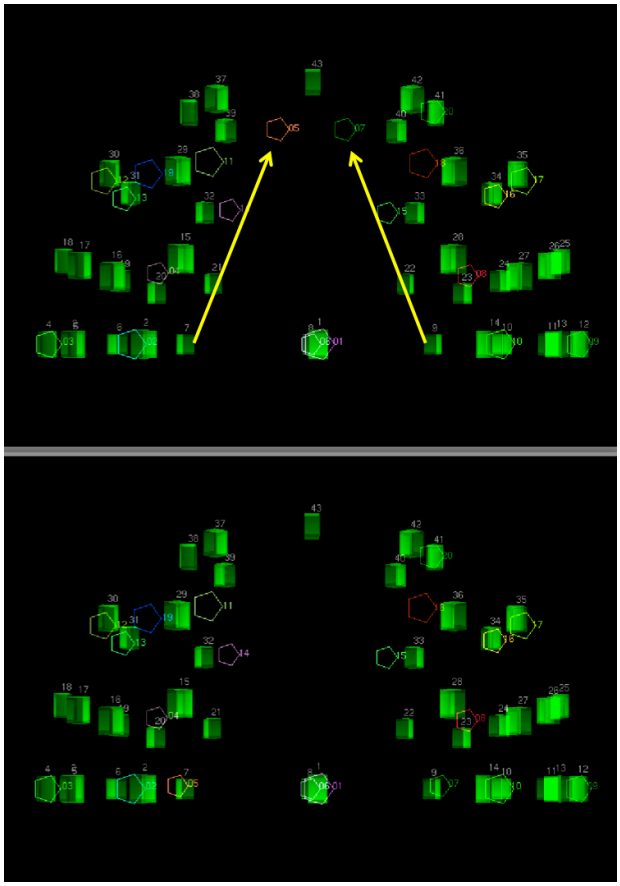

図 3. Ryojinfu 冒頭のサウンドの動き

作業の工程としては、音源の選定、その音源から特定の思い描かれたような加工音を様々な処理により作り出す（ミクロ作曲段階）、これらを使っての ProTools での時間的構成化（マクロ作曲の段階）、ProTools での作業が終了した段階（口語的表現をすれば作品の筋が一本通った段階）で各トラックを予定のチャンネル数のサウンドファイルとして mix し bounce した（ここまではステレオスピーカー Genelec 8040A を使用）。これを Zirkonium セッションににインポート、空間音像移動や音像位置設定を行なって行く（この段階ではヘッドフォン SennheiserHD600 を使用）。この過程の前半までを自宅スタジオにて行なった。ZKM のスタジオ・ミニドームでは、すでに準備されているセッションを Dom23_hp という 23 個のスピーカーを用いたセッティングに移植し、実際に２３スピーカーの音を聴きながら作業を進めた。ちなみにヘッドフォン段階での Zirkonium セッションでもヴァーチャルモード・Dom23_hp を使ったので移行に問題は生じなかった。

4. PROTOOLS セッション中の問題

作業過程で問題となった点に ProTools セッションまでの音加工により作り出されるサウンドファイルをモノにするかステレオにするかという疑問があった。通常モノやステレオのソースファイルは音加工の段階でモノ、ステレオ、4 ch など様々な形態に仕上げて PT セッションにインポートしているのだが、Zirkonium にインポートされたサウンドファイルはモノであっても幾つにも割り振ることが可能である。このため、PT にインポートする段階でどのようなファイルにするのか、またセッション中にどのようなファイルにするのかという決定は、そのサウンドファイルを最終的に幾つ”音源として”持ちたいのかという質問になる。例えばサウンドを左右別々の方向へ移動させるのであれば、ステレオファイルが有効になるが、ただ左右２つのスピーカーから再生サウンドを得るという意味で”ステレオとして”固定して置きたいのであれば、モノファイルを２スピーカー間に位置付ければよい。

また、ある”声部”を形成するサウンドが幾つかのサウンドファイルのミックスであるような場合、Zirkonium 内で再び分けるということは不可能だ。幾つかのサウンドが終始同じ動きをするのか、そうでないのかという点を PT セッションでの bounce 時に正確に決めなくてはならない。

一体幾つのサウンドファイルを音源とするのかも、この段階で空間デザインのプランとともに決定する必要がある。システムが 23 個のスピーカーを持つからといって音源となるサウンドファイルも２３無くてはならないのかというとそうではない。スピーカー数と音源数は一致する必要がない。しかしながら、Acousmonium のサウンドディフュージョンにおいてステレオで完成された作品を多くのスピーカーに割り振ってもやはり音楽構造的にはステレオサウンドの延長形になるように、もとの音源数が少なければ Zirkonium を使ってディフュージョンしてもあまりその醍醐味が出て来ない。一方、逆に音源数が多い場合（例えば 43 のサウンドファイル）、それらをどう音像移動させればよいだろうか。43 のスピーカーシステムでさえ全てのスピーカーが基本的に’塞がっている’ことになるため、移動しようとしてもあまり身動きが取れないだろう。また作業過程でミニドームを使うなら、23 以上の音源数は同様の理由でやりにくい、ということになるだろう。そのような作品では Zirkonium の優れた特性である event 機能を活かすことが難しい。

このような現実的な背景から、後の Zirkonium セッションでの自由な音像移動を可能にする『あそび』を確保しておくために、著者は PT セッションでの最終ファ

イル数を 20 前後になるように想定した。そして最終的にはファイル数は２０となった。

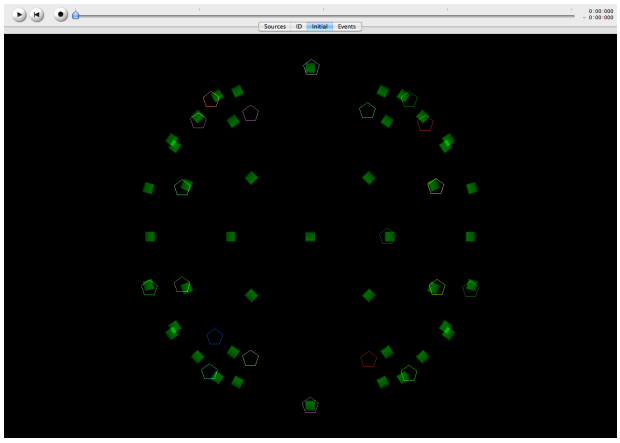

図 4. 43 のスピーカー位置と Ryojinfu の冒頭の 20 の音源位置

5. ZIRKONIUM セッション

Zirkonium 上では、まず preference の『Setup』にてシステムの形態を選び、『リアル』か『ヴァーチャル』かを決定する。『リアル』では実際に同数のアウトプットからスピーカーへ音信号が送られるため、オーデヴィオ・インターフェースのアウト数が合致しなければならない。『ヴァーチャル』の場合はヘッドフォン使用を想定しているので、実際のアウト数は２となる。重要なのは、スピーカー形態（位置は固定）と音像位置（位置は可変）があり、サウンドはスピーカーにダイレクトに送られるのではない、つまり直結していないという点だ。サウンド・ルーティングは「『スピーカーシステム』上の『音像位置』上に送られる」(図.4 参照)。

その音像位置を時間領域で動かす為には『タイム・イヴェント』リストを作成する必要がある。このタイム・イヴェント機能が Zirkonium を画期的な存在にしている。リストの作成はかなり根気を要する。また明確な音像移動の想像力が要求される。タイム・イヴェント・ページでのスピーカー及び音像位置の画像は、マウスで引っ張ると視点を移動させることが出来る。ある一定の動きのリストを作成する場合に、X、Y、Z の３つの次元の視点に時間軸を加えた動きとして吟味することが出来るので、このややこしい作業をかなり助けてくれる。

6. ZIRKONIUM セッション中に表出した問題

6.1. シークエンスソフトと Zirkonium の共同作業性

Ryojinfu の作曲過程は、前後左右上下の三次元空間に時間軸を加えた四次元での音像移動を 20 の独立した音

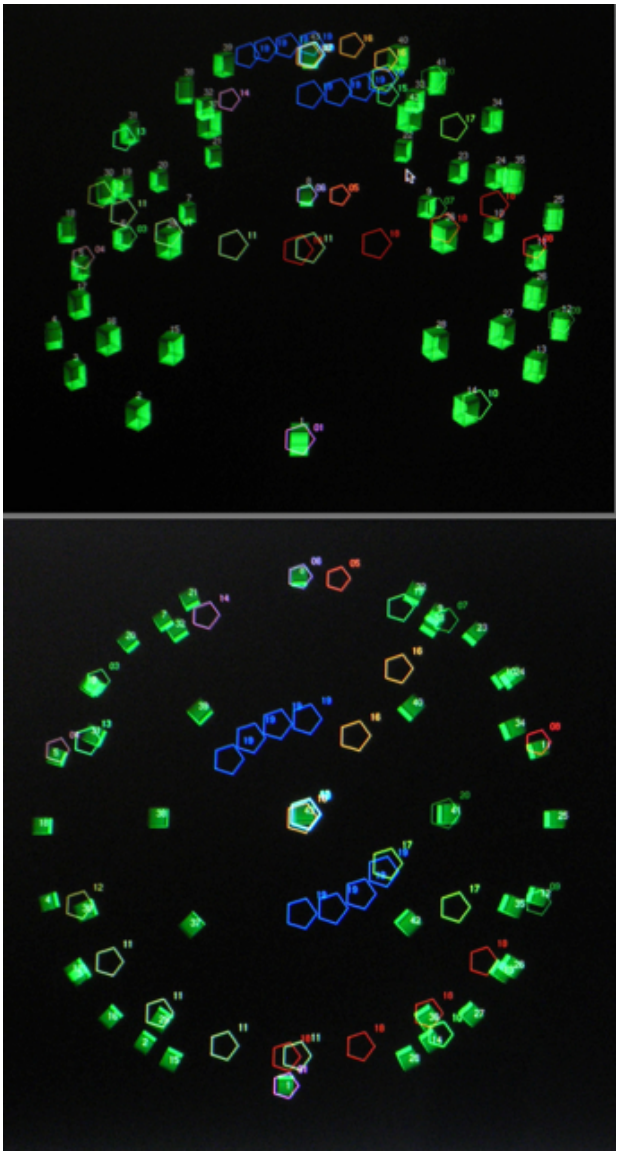

図 5. サウンド配置例。後方上視点 (上）とその真上視点 (下)

源に関して平行して想像し続けるという初めての体験であった。このため一緒に動いて欲しくないサウンドを一緒にバウンスしてしまっているなど、PT セッション中での音の動きへの決定の甘さ、「想像」不明確さが幾度が露呈し、三度ほど PT セッションに戻りファイルをミックスしなおすこととなった。ミックスした２つのサウンドが一緒に動かない方が良い、また、２つのサウンドの音量バランスが次第に変わる、或はそのバランスを修正したいなど様々な不都合が起こる。これはヴァーチャル音響空間とリアル音響空間の違いでもあり避けることが難しい。シークエンス・ソフトと Zirkonium を行ったり来たりせずに両方同時に立ち上げリンクさせ作業する、というアイデアはこういう不都合から生じたのかと思われる。カナダ・モントリオール大

学の Robert Normandeau は Zirkonium を使った音響空間作曲を行なっている作曲家の一人であるが、Nuendo と Zirkonium を独自のプラグインでリンクさせリアルタイムで Zirkonium を使った音響空間を操作する試みをしている。この方法なら従来のシークエンサーソフト上でのマルチチャンネル作曲の作業の際に Zirkonium にもアクセスすることが出来る。難点は操作した動きはリアルタイムのみであり、記録としてイヴェントリストに残らない点だ。空間での音響を吟味構築したい場合は使うことが出来ない。

6.2. イヴェントリスト入力の困難さ

前述のように、空間移動を主なテーマとしながら、イヴェントリストへの入力はひとつひとつ数値で行なわなくてはならない。数値としての入力は最も確かな方法ではあるが、リアルタイムでないため作曲的想像への集中はしばしば分断される。また時々リストを走らせてみないと結果が分からない難しさがある。現プログラマーのヴァーグナーは、マウスで動きを描くことの出来るグラフィック機能のプラグインを試作中であり、もっと感覚的に入力が出来るデヴァイスが期待されている。しかしながら、正確さに関してはやはり我慢強く数値として入力する方が良いという意見もあり、グラフィック機能の付加に入力方法に選択肢が生まれると理解するのが適当と思われる。

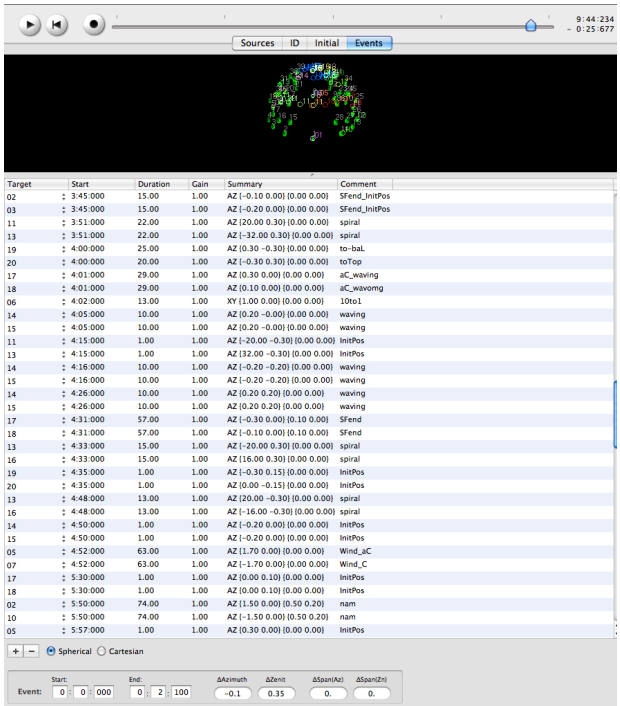

図 6. イヴェントリストの一部

7. まとめ

7.1. Zirkonium で可能になったこと

よく知られているサウンドディフュージョン・システムのうちに GRM の Acousmonium、IMEB の Cybernephone があるが、どちらもステレオを基本のアイデアとする、いわばエクステンデッド・ステレオである。Cybernephone では上下パノラマ移動のアイデアも組み込まれ、独自のミキシング・コンソールも開発されていたが、音像表現はすべてマニュアル操作であった。

例えばディフュージョン・パフォーマンスでの主たる操作はヴォリューム・フェーダーによる各サウンドの'手動' の音量加減であるが、それに基づく音像移動はマスキング効果やハース効果に基づいている。Zirkonium で設定可能な『音源が一秒間に 10 回転しながら上昇する』とか『20 の音源がバラバラに空間を動き回る』というような動きは十指を駆使してもフェーダー操作によるリアルタイムのパフォーマンスでは不可能であり、また音源の発生地点を『一点から面へ次第に広がるように設定する』という音響表現なども不可能だ。

Leo Küpper のサウンド・ドームなどの過去に存在したアクスマティック的アプローチによる多チャンネル・システムとデジタルベースのシミュレーション的アプローチによるシステムとの根本的な違いは、時間上の音の変化をデータとして記憶することが出来るかどうかという点だ。

IRCAM の SPAT は、シミュレーション的システムの最初のものであり、4 チャンネル、8 チャンネルの空間での音像表現を作曲の表現としてその作業過程に取り入れ、音場処理したサウンドをレコーディングすることができる画期的なプログラムだった。しかし一度に読み込める音源数は一つで、多数の音源の動きをデータとして記憶することは出来なかった。

これに対して Zirkonium では多数の音源の複雑な動きをタイムライン上にデータとして記録することが可能であり、それらは同時に再生される。例えば、2 つの音源が半球の上方をぐるぐる回って激しく動いている間に他のステレオ音源が前方左右から耳の高さを後方へ向かって矢のように過ぎ去り、その間に第三の音源が次第に音像をマルチファイしながら後ろから前へ広がって行く、といったようなことも可能になってくる。これらはすべてタイムライン上に『いつ（開始時点）』『いつからいつまで（持続時間）』『どのように（形態）』というデータとして記録される。作曲家はこれらを繰り返し聴き、色々試し、やり直しては吟味して最も良いと思われる動きや響きを最終決定することが出来る。作曲家は完全に新しい絵の具とカンバスを与えられた、といっても過言

ではない。

細かい音像移動や音像表現をひとつひとつ吟味し定着させる行為は、“テープ作品” の延長にある。音楽がライブでないことへの批判は、それが『いつでもどこでも同じである』ことだが、音響空間を重要な音楽要素とするからにはただ再生すればよいというものではないことは過去十数年間に渡って議論されてきた。その場その場で変わる響きに対処する『サウンドディフュージョン』或いは『ディフュージョン・パフォーマンス』といった概念もそういった議論を通じて生まれて来た。Zirkonium を使ったからといってディフュージョン・パフォーマンスが不要になるわけではないが、イヴェントリストを使った作品なら手動での操作は構築された音響空間設定とのバランスを綿密に繊細に計算した最小限の操作にとどめるのがよいだろう。

7.2. 今後の展開によせて

これまで述べて来たように、Zirkonium によって ‘創作に’（サウンドディフュージョンにではなく）使用可能な音響空間が耳の高さの『面』から頭上に広がる三次元の半球体へと劇的に広がった。デジタルベースの仮想空間であるので、スピーカーの位置に関わらず半球の内部面をくまなく滑らかに音像を移動させることが可能だ。また、従来のステレオやバイノーラルなどと組み合わせて２つのサウンドファイルを同時に動かせば、半球の内部を縦横無尽に動くことが出来ることになる。しかし一方で、著者が以前から取り組んでいるような『遠いサウンド』―『遠音』は日本人の音の美学でもある―を作り出すことは Zirkonium だけでは不可能だ。同システムのサウンドはあくまで半球内に留まっている。何十メートルの離れた所に音源が位置するかのような音像を作り出す技術は既に存在しているが、そういった仮想音響空間をも創作の可能性として取り入れることに関して、Zirkonium でも今後 Ambisonic を組み合わせるなど検討がなされているそうだ。

もう一つの点は、Zirkonium が ZKM の Klangdom と密着したシステムであることだ。IRCAM など幾つかの専門機関は既に Zirkonium を導入しているが、まだ一般的に普及してはいない。このため作品は ZKM など限られた venue でのみ上演可能ということになる。

ソフトとしての Zirkonium は、その『ハコ』としてのホールの空間性と、音楽再生行為の間に介在する。Zirkonium は『作曲家が指定すべき作品の構成要素』となった『空間性』をさらに積極的に活用する[5]。スピーカー位置や音響機器の決定や選択、また『ハコ』としてのホール内部構造などへの要望だけでなく、むしろ『空間性』がそれらに決定的に左右されないよう恒久性を持たせることで創作の初段階からディフュージョン（最終段階）までを一環化させ、空間性を積極的に操作するためのパラメーター化のデヴァイスとして存在する。

『専売特許』と『普及』の兼ね合いの難しさはあるが、今後の acousmatic 音楽の発展を考えた場合、その方向は 3D 空間へ広がっていくだろうと予想される。日本でも映画の効果サウンドだけでなく芸術音楽制作の分野で電子音響音楽作曲家が関われるようになること、またそのような施設や環境が作られることを期待したい。

著者はこの制作後、インフォ・ヴァージョンとしてのステレオ版を作成する必要があった。Zirkonium の 3D 空間をどのようにステレオ音像に収めるのが良いのか、はなはだ苦心した。上下軸の動きは取り除かれ、前後間の平行移動も失われる。空間における音響構築を重要な要素とする作品において、ステレオ版はどれだけオリジナル版の表現を伝えうるのだろうか。このネガティヴ作業を通じて、3D 音響空間が所有する表現の豊かさを改めて実感することになった。

[5] 電子機器を使って音響や音楽的構造を創作する場合の再生環境は、作曲家が指定すべき作品構成要素となったのである。（水野みか子　2004）

8. ACKNOWLEDGEMENT

本制作プロジェクトを承諾し、客員作曲家滞在を助成してくださったカールスルーエ ZKM、ZKM 音楽音響部門 IMA 所長の Ludger Brümmer 氏、技術スタッフの Götz Dipper 氏、David Wagner 氏、また事務スタッフの Till Kniola 氏に感謝致します。また『梁塵秘抄』を知るきっかけを与えてくださった柴田純子さんに感謝いたします。

9. 参考文献

[1] Ramakrishnan, C. Zirkonium. Karlsruhe : ZKM Institut für Musik und Akustik.

[2] Barth, J. 2011. Zirkonium Reference. Karlsruhe : ZKM Institut für Musik und Akustik.

[3] Ramakrishnan, C., Großmann, J., and Brümmer, L. 2006. The ZKM Klangdom. NIME06.

[4] Küpper. Léo, 1997. Analysis of the spatial parameter Psychoacoustic measurements in sound cupolas. In Composition/Diffusion en Musique Electroacoustique. Actes III Vol.3, pp.289-314. Academie Bourges, Editions MNEMOSYNE.

[5] 水野みか子　2010　「1970 年大阪万博のシュトックハウゼン」『先端芸術音楽創作学会会報』vol.2No.3 pp.21-26

[6] 柳田益造　2010　「1970 年大阪万博のシュトックハウゼン ―西ドイツ館スナップショット ―」『先

端芸術音楽創作学会会報』vol.2 No.3 pp.13-20
[7] 吉川英史　1977　『日本音楽の歴史』大阪：創元社.
[8] 新訂標準音楽辞典『今様』の項　1995　東京：音楽之友社
[9] 旺文社日本史事典『後白河天皇』の項　2000　旺文社
[10] 後白河天皇　『梁塵秘抄』[馬場光子監修『梁塵秘抄口伝集』]　講談社学術文庫
[11] 安藤彰男、濱崎公男、小野一穂、中山靖茂、松井健太郎、大久保洋幸、田島利文、杉本岳大、大出訓史 2009　「高臨場感音響システム」『 NHK 放送技術研究所研究史’00-’09』pp.109-120. http://www.nhk.or.jp/strl/publica/kenkyuushi/ken-j.html
[12] 水野みか子　2004　「音楽と空間 5 ― 戦後日本音楽における三つの空間 ―」　http://www.jia-tokai.org/sibu/architect/2004/0401/ongaku/htm

10. 著者プロフィール

石井 紘美 (Hiromi ISHII)

博士（PhD. 電子音響音楽作曲/音響美学）。ドレスデン音楽大学上級課程にて W・イエンチに電子音響音楽を師事。ドイツ音楽家資格試験に卓抜にて合格後、英国から奨学金を得て 2001 年よりシティ大学にて S・エマーソン、D・スモーリーに師事、『日本伝統音楽との関係における電子音響音楽作曲』のテーマで博士研究。CYNETart、フロリダ EMF、英国 SAN・EXPO966、北京 MusicAcoustica、リスボン Musica Viva、Musiques&Recherches、オランダ・ガウデアムス、ローマ EMU、アイスランド Punto y Raya、NYCEMF など様々な音楽祭や音楽週間にて作品が入選、上演され、また西ドイツ放送、中部ドイツ放送、ベルリン・ドイツ放送などで放送されている。2006 年 ZKM 奨学金を得て客員作曲家、2013 年再度客員作曲家。WERGO より CD『Wind Way 風の道』が出版される。ヴィジュアル・ミュージック作家としては、音楽映像の両メディアを同時進行的に制作する手法を取り、キューレーターとしても活動している。ケルン在住。

連載

SUPER COLLIDER チュートリアル (5)
SUPER COLLIDER TUTORIALS (5)

美山 千香士
Chikashi Miyama
ケルン音楽舞踏大学
Hochschule für Musik und Tanz Köln

概要

本連載では、リアルタイム音響合成環境の SuperCollider(SC) の使い方を、同ソフトを作品創作や研究のために利用しようと考えている音楽家、メディア・アーティストを対象にチュートリアル形式で紹介する。
SuperCollider(SC) is a realtime programming environment for audio synthesis. This article introduces SC to musicians and media artists who are planning to utilize the software for their artistic creations and researches.

1. 今回の目標：バスを理解する

当連載ではこれまで SC によるシンセサイザーやエフェクトの作り方、MIDI 機器との連携、メロディーの演奏の仕方などを紹介してきた。これらを組み合わせてより大規模なプログラムを制作する時に要となるのが**バス** (Bus) という機能である。今回はこのバスについて、例を挙げつつ紹介していく。

2. バスを使う意味

本連載の第 2 回で学習したように、SC では様々な Ugen を組み合わせた一連のオーディオ処理プロセスを **SynthDef** で定義する事で、自分だけの「楽器」を作る事ができる。例えば以下のリスト 1 では Saw で生成したノコギリ波にエンベロープを施し、さらにローパス・フィルタとリバーブをかける一連の処理を SynthDef を用いて *combined* という名前で定義し、その後、定義した SynthDef を **Synth** を用いて実行している。

リスト 1. Synth の例

```
1 SynthDef("combined",{
2   arg freq=440, amp=0.5;
3   e=Env([0,1,1,0],[0.1,0.5,0.4]);
4   g=EnvGen.ar(e,doneAction:2);
5   a=Saw.ar(freq,amp*g);
6   l=LPF.ar(a,700);
7   r=FreeVerb.ar(l,0.5,0.8);
8   Out.ar(0,[r,r]);
9 }).load;
10
11 Synth("combined");
```

しかし、このプログラムには 2 つの問題がある。1 つ目の問題は、doneAction:2 が **EnvGen** に与えられているため、エンベロープが終了した時点（開始から 1 秒後）でこの Synth は解放され、FreeVerb の付加した残響音が不自然に途中で切れてしまい、最後まで聞こえない事。

リスト 2. Synth を用いたアルペジオの演奏

```
1 Routine({
2   [60,64,67].do{
3    arg pt;
4    Synth("combined",["freq",pt.midicps]);
5    0.1.wait;
6   }
7 }).play
```

2 つ目は、この SynthDef を用いて、例えばリスト 2 のようにド・ミ・ソのアルペジオを、第 2 回のチュートリアルで勉強した **Routine** を使って演奏した場合、Saw や EnvGen のように個々の音の波形を生成しエンベロープを施す Ugen だけでなく、ローパス・フィルターとリバーブのように音を加工する Ugen もこの SynthDef の定義に含まれているため、図 1 のように各 Synth に個別のローパス・フィルター、リバーブが生成されてしまい、無駄な CPU 負荷が生じる事である。

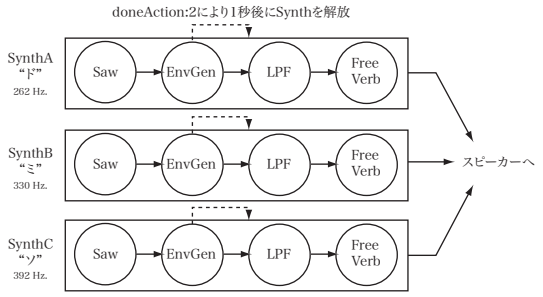

図 1. SynthDef *combined* の問題点

ここで、もし、図 2 の示すように、1 つ 1 つの音を作り出す「生成用」の SynthDef と、生成した音を加工する「加工用」の SynthDef を個別に定義し、複数の「生成用」の Synth が出力したオーディオ信号を単一の「加工用」の Synth に送信し、「加工用」の Synth で全ての「生成用」の Synth の音をまとめて加工することができれば、より効率のよい処理が可能となるはずである。しかし、これを実現するためには、ある Synth から別の Synth に音を「送る」メカニズムが必要となる。そのメカニズムこそが今回のテーマである**バス**である。

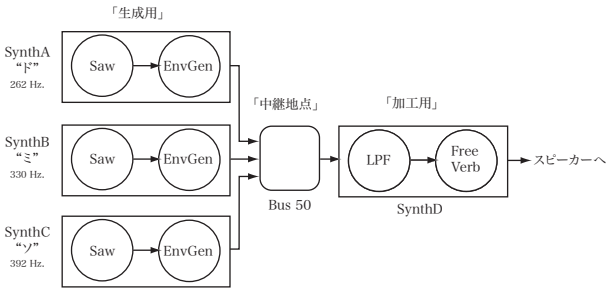

図 2.「生成部」と「加工部」の分離

3. バスとは何か？

バスとは DAW やデジタルミキサーなどでよく使われる用語で、一般的には「信号経路」という意味であるが、SC でのバスは簡単に言えばオーディオ信号の**中継地点**を意味する。これを用いることにより、例えばある Synth で生成した信号を一度バス（＝中継地点）に送り、その後、他の Synth がその中継地点から音を受け取り、音を加工し、スピーカーに送るというような事が可能となる。バスは非常に柔軟であり、1 つのバスに複数の Synth から同時に音を送る事も、複数の Synth が 1 つのバスから同時にオーディオ信号を取得することも可能である。

リスト 3 はリスト 1 をバスを利用して書きなおしたものである。先ほど 1 つの SynthDef、*combined* として定義していたものを、バスを用いると、「生成用」の Synth、*generate* と「加工用」の Synth、*process* に分割して定義することができる。

リスト 3. バスの例

```
1  SynthDef("generate",{
2    arg freq=440,amp=0.5;
3    e=Env([0,1,1,0],[0.1,0.5,0.4]);
4    g=EnvGen.ar(e,doneAction:2);
5    a=Saw.ar(freq,amp*g);
6    Out.ar(50,a);
7  }).load;
8
9  SynthDef("process",{
10   i=In.ar(50);
11   l=LPF.ar(i,700);
12   r=FreeVerb.ar(l,0.5,0.8);
13   Out.ar(0,[r,r]);
14 }).load;
```

リスト 3 から分かるように、バスにオーディオ信号を送るには、**Out** を、バスからオーディオ信号を取得するには **In** を用いる。SC ではバスを複数同時に使用する事ができるため、個々のバスには番号が与えられている。例えば、SynthDef、*generate* では、ノコギリ波の信号を Out.ar(50, a) でバスの 50 番に送っている。そして SynthDef、*process* では 50 番に送られた信号を In.ar(50) を用いて取得し、それをローパス・フィルターとリバーブで加工している。Out.ar はスピーカーにオーディオ信号を送る Ugen として、既にチュートリアルの第 2 回で一度利用しているが、ここではスピーカーにではなくバスにオーディオ信号送るために使用している。この Out のバスとオーディオ入出力との兼ね合いについては後述する。

バスを用いて 2 つ以上の Synth を連携させる時の注意点として、必ずオーディオ信号を「受け取る側」そして「送る側」の順番で実行しなくてはならない事があげられる。例えば、リスト 3 の例は、*process* が信号を受け取る側、*generate* が送る側であるので、リスト 4 のように *process*、*generate* の順番で実行する。これをもし逆の順番で実行した場合には音が聞こえてこないだけでなく、エラーも出力されない。故に、バスを使う場合に Synth を実行する順番に細心の注意を払う必要がある。これは SC サーバー内での「ノード」と呼ばれる構造に関連した事項であるが、詳細は次回以降に扱う。

リスト 4. Synth 実行の順番

```
1 Synth("process");
2 Synth("generate");
```

In

指定したバスからの信号を取得する。.ar でオーディオ信号、.kr でコントロール信号を得る。

.ar(*bus*, *numChannels*)

.kr(*bus*, *numChannels*)

bus･･･ 取得するバスの番号。複数取得する場合は最初の番号。

numChannels･･･ 取得するチャンネルの数。

リスト 5 はリスト 2 のアルペジオの演奏をバスを用いて行ったものである。

リスト 5. バスを用いたアルペジオの演奏

```
1 Synth("process");
2 Routine{
3   [60,64,67].do{
4     arg pt;
5     Synth("generate",["freq",pt.midicps]);
6     0.1.wait;
7   }
```

```
8 }.play
```

リスト 2 とリスト 5 をそれぞれ実行して、ステータス・バーで使われている Ugen の数や CPU の使用率を比較してみると、図 3 の示すように、明らかに分割した方が使用している Ugen の数も少なく、また CPU 使用率も低い事がわかる。このように、SynthDef をバスを用いて「生成部」と「加工部」に分割する事で、余計な CPU 負荷を生じさせる事なく音の生成と加工が可能となる。無論このような小規模なプログラムの場合は CPU に与える負荷の違いは大きなものではないが、例えば、より複雑なエフェクトを施す SynthDef を定義し、それを用いて 100 音を同時に生成するといった場合には、大きな差が生まれる。

図 3. リスト 2(上) と 5(下) のパフォーマンス比較

また、リスト 5 では、doneAction:2 により FreeVerb がエンベロープの終了とともに解放されないため、残響効果を最後まで聞き取ることができる。

4. バスの信号を可視化する

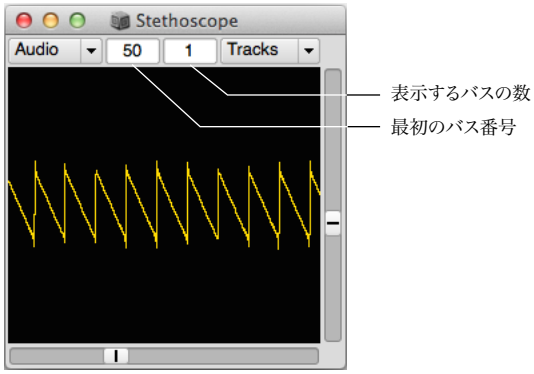

図 4. scope によるバスの可視化

バスを複数同時に使用するようになると、Synth 間の連結が複雑になり、指定したバスに任意の信号が送られているかを確認する必要が生じる。これを行うには、リスト 6 のようにバスにオーディオ信号を送る Synth を実行し、その後、SC サーバーに.scope メッセージを送る。すると、図 4 のようにバスの信号をスコープに表示させる事ができる。この時.scope メッセージの第 1 引数は表示するバスの数、第 2 引数は表示を開始するバスの番号となる。故に、.scope(1,50) を送ると 50 番のバス 1 チャンネルが表示される。また表示されるスコープ・ウインドウの GUI を操作する事によって、表示するバスとその数を任意に変更する事も可能である。

リスト 6. バスをスコープで確認

```
1 Synth("generate");
2 s.scope(1,50);
```

5. バスとオーディオ・チャンネルの関係

リスト 3 では 2 つの Synth、*generate* と *process* を繋ぐために 50 番のバスを利用した。このバス番号は、ユーザーが基本的に任意に決めてよいが、幾つかの点に留意する必要がある。

まず、デフォルトでは、バスは 128 個、つまり 0 から 127 までしか使用できない。もし、それ以上のバスを利用したい場合は、サーバーを起動する前にリスト 7 の要領でバスの数を変更する必要がある。

リスト 7. オーディオ・バスの数の変更

```
1 s.options.numAudioBusChannels=256;
```

また 3 や 7 など低いバス番号を使う時も注意が必要である。それには以下のような理由がある。

SC サーバーは使用しているハードウェアの構成、つまりコンピュータのマイク・ライン入力や HDMI 入出力などの数に関わらず、デフォルトで 8 チャンネルのオーディオ入力と 8 チャンネルのオーディオ出力、計 16 チャンネルのオーディオ入出力を確保する。現在、確保されているオーディオ入出力チャンネルの数を確認するには、リスト 8 のように、SC サーバーにメッセージを送る。すると、設定を変更していない限り、現在確保されている入出力オーディオ・チャンネル数である「8」が 2 度ポスト・ウインドウに表示される。

リスト 8. 入出力チャンネルの確認

```
1 s.options.numOutputBusChannels.postln;
2 s.options.numInputBusChannels.postln;
```

しかし、例えば筆者の環境ではオーディオ入力は 2 チャンネル、出力も 2 チャンネルしかないため、デフォルトのままでは余剰な入出力チャンネルが確保されている事になる。無論、これまでのように、この設定のままで SC サーバーを使用しても問題はないが、ハードウェアの入出力チャンネル数に合わせて、SC の確保する入出力チャンネル数を、リスト 9 のように変更する事も可能である。

リスト 9. 入出力チャンネルの変更

```
1 s.options.numOutputBusChannels=2;
2 s.options.numInputBusChannels=2;
```

リスト 9 では入出力のチャンネル数をそれぞれ 2 チャンネルに設定し直している。これにより SC サーバーは起動時に入力・出力ともに 2 チャンネル、合計 4 チャンネルのオーディオ入出力チャンネルを確保する。このサーバーを起動する前に確保したオーディオ入出力チャンネル数はバスと密接な関係があり、SC サーバーを起動した時に、N 個のオーディオ出力と M 個のオーディオ入力が確保されている場合は、バスの「0」から「N-1」が自動的にオーディオ出力に送られ、「N」から「M-1」が自動的にオーディオ入力からの信号を得る。

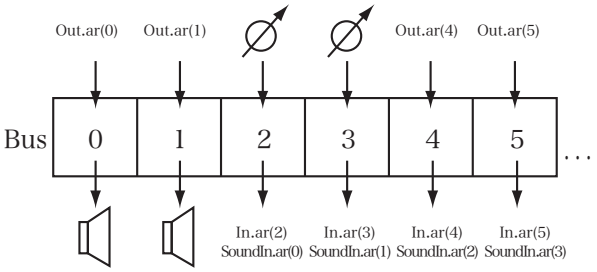

図 5. リスト 8 実行後のオーディオ入出力とバスの関係

例えば、リスト 9 を実行してオーディオ入出力チャンネル数をそれぞれ 2 に設定すると、図 5 のようにバスの 0 と 1 チャンネルはオーディオ出力に送られ、2 から 3 チャンネルはオーディオ入力からの信号を受けとる事となる。故に、リスト 9 を実行し、サーバーを立ち上げた後に、リスト 10 のようなコードを実行すると、入力信号が出力にバイパスされる（直接送られる）こととなる[1]。

リスト 10. オーディオ入出力のバイパス

```
1 SynthDef("bypass",{
2   Out.ar(0,In.ar(2, 2));
3 }).load;
4
5 Synth("bypass");
```

このように、低い番号を持つバスは、ハードウェアのオーディオ入出力と接続している可能性があるため、現在何チャンネルのオーディオ入出力が SC サーバーに確保されているのかを念頭におき、Synth 間でのオーディオ信号の受け渡しに使うバスとしてこれらを使用しないように留意する必要がある。

コンピュータにオーディオ・インターフェースなどの機器を繋ぎ、より多くのオーディオ入出力を使う時は、リスト 9 の要領で、入出力チャンネル数を増やす必要がある。このように、ハードウェアの増設などによって、入出力チャンネル数の設定を変更しなければならない時、例えば、オーディオ出力数を 32 チャンネルに設定

[1] 注：ハウリングを防止するため、ラップトップでこのプログラムを実行する際はヘッドフォンを着用の事。

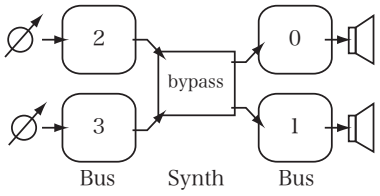

図 6. リスト 10 のバスと Synth の関係

した場合には、オーディオ入力をとるために、In の第 1 引数を 32 に変更しなくてはならない。このように In を用いてマイクやラインからのオーディオ入力を得ていると、ハードウェア構成を変更した際にプログラムを変更しなくてはならないため、オーディオ入力を得るには前回のチュートリアルで紹介した **SoundIn** が推奨される。SoundIn は In とほぼ同じ機能を提供する Ugen であるが、In と違いチャンネルに 0 番を指定すると、オーディオ出力のチャンネル数の設定に左右されることなく、最初のオーディオ入力が接続されているバスからのオーディオ信号を得る事ができる。

6. 2 種類のバス

.kr と.ar については第 2 回のチュートリアルで学習した。.kr はコントロール・レート、.ar はオーディオ・レートの意味で、コントロール・レートで作成した信号は直接スピーカーに送って聞くことは不可能だが、他の Ugen のパラメータに代入し、そのパラメータを動的に変化させる事ができる。例えば、リスト 11 では SinOsc で作られたコントロール・レートの波形 (LFO) が 2 つ目の SinOsc の周波数をコントロールすることで、ビブラートのような効果を得ている。

リスト 11. コントロール信号の例

```
1 SynthDef("vibrato",{
2   l=SinOsc.kr(5,0,100);
3   x=Saw.ar(l+440,0.5);
4   Out.ar(0,[x,x])
5 }).load;
6
7 Synth("vibrato");
```

SC のバスにはオーディオ信号を送るオーディオ・バス以外にもこのようなコントロール信号を送ることのできるコントロール・バスがある。コントロール信号をバスに送るためには、Out.ar の代わりに Out.kr を用いる。リスト 12 はリスト 11 のプログラムをコントロール・バスを用いて 2 つの Synth に分割して記述したものである。

リスト 12. コントロール・バスの例

```
1 SynthDef("lfo",{
2   l=SinOsc.kr(5,0,100);
```

```
  Out.kr(0,l);
}).load;

SynthDef("saw",{
  arg freq;
  l=In.kr(0);
  x=Saw.ar(l+freq,0.5);
  Out.ar(0,[x,x]);
}).load;
```

この SynthDef をリスト 13 の要領で実行すると、Synth、*lfo* で、振幅が 100 で周波数が 5 Hz. の正弦波が作られ、その信号がコントロール・バスの 0 番に送られる。そして、SynthDef、*saw* は 0 番のコントロール・バスから信号を取り出し、ノコギリ波を生成する Saw.ar の周波数にその信号と引数 freq を加算したものを適用し、リスト 11 と同様のビブラートのかかったノコギリ波を生成する事ができる。

リスト 13. リストの実行

```
Synth("saw",["freq",440]);
Synth("lfo");
```

前述したように、複数の Synth が同一のバスから信号を得ることも可能である。これを利用して例えば、リスト 14 のように複数の Synth を同じバスに接続し、2 つの Synth のパラメータを単一の LFO を同期させるなどという利用が考えられる。

リスト 14. LFO の同期

```
Synth("saw",["freq",440]);
Synth("saw",["freq",540]);
Synth("lfo");
```

このリストの 2 つ *saw* はどちらもコントロール・バスの 0 番を参照しているため、2 つの Synth によって奏される音の周波数は異なるが、ビブラートが同期している。

コントロール信号を扱うコントロール・バスはオーディオ信号を扱うオーディオ・バスと完全に分離しており、相互に干渉する事はない。また、オーディオ・バスと異なり、コントロール・バスはデフォルトで 4096 チャンネルまで使う事ができる。

7. バス番号を引数を用いて指定する

今までの例では、Out、In に与える引数は全て、50 や 0 などの定数であった。しかし、実際のプログラミングでは、より SynthDef を柔軟に活用するために、Out 及び、In のバス番号を SynthDef 内で引数 (arg) を介して指定するのが一般的である。例えばリスト 15 はリスト 12 に変更を加え、Out と In のバスをそれぞれ引数で指定できるようにしたものである。

リスト 15. バス番号を引数で指定

```
SynthDef("lfo",{
  arg outBus;
  l=SinOsc.kr(5,0,100);
  Out.kr(outBus,l);
}).load;

SynthDef("saw",{
  arg inBus,freq;
  l=In.kr(inBus);
  x=Saw.ar(l+freq,0.5);
  Out.ar(0,[x,x]);
}).load;
```

SynthDef をこのように定義した場合、リスト 16 のように、コントロール信号の送信先、または受信元のバスを実行時に決定する事ができるため、定義した SynthDef をより柔軟に用いる事ができる。

リスト 16. 引数を介した実行時のバス番号の指定

```
Synth("saw",["inBus",50,"freq",440]);
Synth("lfo",["outBus",50]);
```

8. バスの接続をリアルタイムに変更する

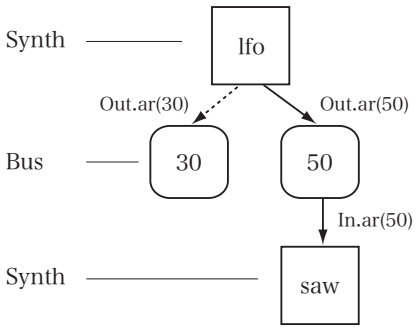

図 7. 実行時のバス接続変更

バスの接続は Synth を実行した後も任意に変更することが可能である。リスト 17 ではリスト 16 と同様に LFO を 50 番のバスに送り、*saw* がそれを受け取って、ビブラートのかかったノコギリ波が生成される。リスト 16 との唯一の違いは最初の Synth、*lfo* が変数 x に代入されている点である。

リスト 17. Synth「lfo」を x に代入

```
Synth("saw",["inBus",50,"freq",440]);
x=Synth("lfo",["outBus",50]);
```

その後、x に set メッセージを送ることにより、outBus を他の番号に変更すると、2 つの Synth 間のバスによる接続が中断され、ビブラートが停止する (リスト 18)。

リスト 18. outBus の変更

```
x.set("outBus",30);
```

しかし、もう一度リスト 19 のように再度バスを 50 番変更した場合、両 Synth が再接続され再びノコギリ波にビブラートがかかる。

リスト 19. outBus を再び戻す

```
1 x.set("outBus",50);
```

このように、SC ではバスの接続はいつでも自由に変更する事ができる。これを応用すれば、例えば、ライブパフォーマンスの際に、バスの接続を任意に変更し、音を生成する Synth のパラメータを様々な LFO でコントロールしたり、素材音に次々に多様なエフェクトを施していく事も可能である。

9. まとめ

今回は、SC のバス機能について学習した。バスによる Synth の連携は便利であるが、あらゆる側面で有効な手段とはいえない。バスの使用により SynthDef を分割しすぎると、Synth の実行の順番の問題や、どのようにバスを繋ぐかを事前に入念に設計する手間がかかることとなる。故に、バスをプログラムに利用する際は、その利用によって、CPU 負荷が大幅に軽減されるなどの明白なメリットがあるかどうかを事前に十分に検討する必要がある。次回は、このバス機能に密接に関連した Synth の計算順序に関わる「ノード」という概念について学習する。

10. 参考文献

[1] *SuperCollider*, http://supercollider.sourceforge.net(アクセス日 2014 年 9 月 29 日)

[2] Wilson, S., Cottle, D. and Collins, N. (eds), *The SuperCollider Book* The MIT Press, 2011

11. 著者プロフィール

11.1. 美山 千香士 (Chikashi Miyama)

作曲家、電子楽器創作家、映像作家、パフォーマー。国立音楽大学音楽デザイン学科より学士・修士を、スイス・バーゼル音楽アカデミーよりナッハ・ディプロムを、アメリカ・ニューヨーク州立バッファロー大学から博士号を取得。作曲・コンピュータ音楽を莱孝之、エリック・オニャ、コート・リッピ氏らに師事。Prix Destellos 特別賞、ASCAP/SEAMUS 委嘱コンクール 2 位、ニューヨーク州立大学学府総長賞、国際コンピュータ音楽協会賞を受賞。2004 年より作品と論文が国際コンピュータ音楽会議に 15 回入選、現在までに世界 19 カ国で作品発表を行っている。 2011 年、DAAD(ドイツ学術交流会) から研究奨学金を授与され、ドイツ・カールスルーエの ZKM で客員芸術家として創作活動に従事。近著に「Pure Data-チュートリアル＆リファレンス」(Works Corporation 社) がある。現在ドイツ・ケルン音楽大学講師。スイス・チューリッヒ芸術大学コンピュータ音楽・音響研究所 (ICST) 研究員。バーゼル造形大学のプロジェクト「Experimental Data Aesthetics」プログラマ。2013 年に松村誠一郎氏と日本語の Pure Data ポータル Pure Data Japan(http://puredatajapan.info) を創設。公式ウェブサイト：http://chikashi.net

会告

■先端芸術音楽創作学会運営体制

運営委員

事務局
会長：小坂直敏 (東京電機大)
副会長：古川聖 (東京芸大)
広報 (研究会)：寺澤洋子 (筑波大)
広報 (イベント)：有馬純寿 (帝塚山学院大)，柴山拓郎 (東京電機大)
広報 (Web)：喜多敏博 (熊本大)
会計：森威功 (東京芸大)
会報：安藤大地 (首都大)

一般
石井紘美 (英国シティ大), 今井慎太郎 (国立音大),
今堀拓也, 川崎弘二,
中村滋延 (九州大), 西野裕樹 (シンガポール国立大),
沼野雄司 (桐朋学園大)，　水野みか子 (名古屋市立大),
高岡明 (玉川大), Cathy Cox(玉川大),
Michael Chinen

在外国メンバー
Mara Helmuth(University of Cincinnati, U.S.A.),
Karen Wissel(Growth in Motion, Inc., U.S.A.),
Mark Battier(Sorbonne, France)

■電子ジャーナルへの投稿を歓迎します

原稿は原稿執筆要領に沿って書いていただき，編集委員まで送付して下さい．また，詳細については編集委員までお問い合わせ下さい．
編集委員：安藤大地 dandou[at]sd.tmu.ac.jp

原稿は以下のカテゴリに分類されます．

- 原著論文 研究論文．査読を経て採録されたものが掲載されます．
- 研究報告 研究の予稿．査読はなし．通常の学会の研究会の予稿に相当．
- 会議報告 国際会議等の参加報告．
- 解説 既に知られている重要な技術，概念，研究動向を読者にわかりやすく伝える記事．
- 連載 何回か継続して綴られる原稿．解説や報告などさまざまな区分が個々の原稿にはあるが，全体を連載として区分する．
- インタビュー 作曲家，音楽家へのインタビュー．
- 書評 読者へ紹介したい単行本の感想，評論など．
- 報告 自身のあるいは研究室の活動報告など．
- 作品解説 自作品の哲学，用いているシステム紹介，音楽理論などを作品の中で特筆すべき内容を解説する．プログラムノートを発展させ，より学術的にしたもの．

このほかのカテゴリも必要に応じ，作成したいと考えています．上記に当てはまらないものは編集委員にご相談下さい．

■研究会の参加費について

非会員の場合は研究会参加が有料となりました．参加費500 円を徴収いたします．

■研究会への発表募集

研究会の講演発表を募集します．詳細は別途ご案内いたします．内容は電子ジャーナルの原稿のカテゴリと同等なものです．また，それ以外の話題がありましたら，運営委員までご相談ください．

■第 21 回研究会

日時：2014 年 9 月 28 日（土）　13:30 -
会場：筑波大学 東京キャンパス文京校舎 1F 120 講義室

プログラム

○ 1 件目
講演タイトル：五度圏に基づいた確率的複雑化と偏向化による情動的旋法を用いた音楽生成システム
発表者：大村 英史（独立行政法人 国立精神・神経医療研究センター）

○ 2 件目
講演タイトル：ALARM/WILL/SOUND: Car Alarm Sound Perception Experiments and Acoustic Modeling Report
発表者：Alex Sigman（Keimyung University）

○ 3 件目
講演タイトル：新しいコンピュータ音楽言語 LC とその 3 つの特徴
発表者：西野 裕樹（シンガポール国立大学）

○ 4 件目
講演タイトル：ヴィジュアル・ミュージックにおけるサウンドとイメージ間のインタラクティヴ・プロセス
発表者：ヴィルフリート・イェンチ

○ 5 件目
講演タイトル：ツィルコニウムを使った三次元空間アクスマティック作品『梁塵譜』
発表者：石井 紘美（国際芸術アカデミー・ハイムバッハ）