

平成 29 年 6 月 16 日現在

機関番号：32657

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26370111

研究課題名(和文) 構造的音色を持つ音合成方式

研究課題名(英文) Sound synthesis system for structured timbre

研究代表者

小坂 直敏 (OSAKA, NAOTOSHI)

東京電機大学・未来科学部・教授

研究者番号：20366389

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：コンピュータ音楽の中で音色合成技術は新たな表現と直結するため、重要な技術分野である。これまでサウンドモーフィングやサウンドハイブリッドは個別技術として検討されてきたが、これに「音の音」と呼ぶ新たな概念を加え、全体を構造的音色と呼ぶ概念を提唱した。音の音は、一つの音色の合成に別の音色を使う方法だが、サウンドハイブリッドと異なる点は、階層的な構成をとり、下位の音色を明確に識別できる音色として定義している。

この考えに基づき、2つの合成例を製作した。一つは朗読音声、ストリングスの合成音で達成すること。もう一つは、NMF(非負値行列因子分解)を応用して、音声をヴァイオリンや尺八の音で合成した。

研究成果の概要(英文)：In computer music, timbre synthesis is a very important technical area since it directly devotes to the new expression. So far sound morphing and sound hybridization have been studied isolatedly. We add "sound by sounds" and defined a new concept of structured timbre for all of three. 'Sound by sounds' used many other timbres to synthesize one particular timbre. The difference from sound hybridization is that it is defined to have a hierarchical structure and the timbre of the lower layer is identified with high intelligibility.

Based on this idea, two synthesis examples are studied: one is a speech is expressed in terms of string sound. The other is as an application of NMF (Non Negative Matrix Factorization), speech is expressed in terms of violin and string sound.

The former is example was introduced in a real computer music concert. In the second example, if the intelligibility of speech is higher, the recognition of lower layer instrumental sound becomes weaker.

研究分野：楽音合成

キーワード：音色 音色合成 構造的音色 音の音 サウンドモーフィング サウンドハイブリッド

1. 研究開始当初の背景

コンピュータ音楽の技術の中で音合成は発展の初期からある代表的な分野である。これまでに確立された音合成法は、残響やフィルタのようにアナログの時代から用いられていたエフェクトのデジタルへの焼き直しのほか、音色を一つの音から別の音へ連続的に変化させていく**サウンドモーフィング**や、ある音色に別の音色の一部を転写しつつも全体として一まとまりの音(音脈)の体裁を保つ**サウンドハイブリッド**が提案されてきた。

これらの技術は、技術側からの興味により発展してきたもので、相互に音楽的価値や創作的価値が整理されてきたわけではない。これらはより体系的に整理され、また新たな音色を開拓する際も、このような大きな枠組みの下で検討する必要がある。

サウンドモーフィングや、サウンドハイブリッドは、一つの音、という概念を極限まで拡張しているもので、これを**拡張音脈**と定義した。また、これとともに、**音の音**という新しい合成手法を定義し、これらをまとめて**構造的音色**と定義する。音の音は、一つの音色を、別の音色で表現するもので、両者は階層的な関係がある。文字の例では、222222222222 は、「2の一」で、2を用いて、その上位で漢字の一を表現している。同様に、「音の音」とは、楽音を寄せ集めて音声など、別の音を作る試みである。端的には、複数の楽音による一つの擬音である。

楽音、環境音の他の音による表現の検討は、茂出木による MIDI 音の楽音によるオート符[1]が知られている。しかし、その手法は MIDI 音源に限定された表現である。本研究は、信号音の処理を対象としている。

2. 研究の目的

本研究では、「音の音」エフェクトを中心として、このエフェクトを用いる音楽作品の制作と、また合成音の製作方法について検討する。

3. 研究の方法

この目的のために、以下の検討を行った。

(1) 「音の音」エフェクトを含む楽曲を制作し、構造的音色、および同エフェクトを啓蒙する。

(2) 「音の音」エフェクトの合成方式について検討する。具体的に、音声の母音をフルートなどの楽音で表すために、NMF(非負値行列因子分解)を用いて、楽音を基底に分解し、音声をこれらの基底で表現することにより、音の音を実現する。

通常は、合成技術が確立してから、これを応用した作品制作を行うのが研究の流れであるが、オーケストラ作品による音楽制作の実現を検討した際スケジュールの関係で初年度となったため、作品制作と合成音技術の検討は別々に分けた。合成技術が確立していない中では、これまでのサウンドハイブリッ

ド技術などを組み合わせ、同エフェクトの紹介、および、認知に努めた。

4. 研究成果

(1) 音の音エフェクトを持つ楽曲の制作

音楽作品制作においては、既存の手法を用いて音の音を造ることを試みた。まず、オーケストラ音を下位層の音色とし、これにより、音声を表現しようと試みた。

一方、生のオーケストラで音声を実現するのは困難である。同様の意図の作品である、Jonathan Harvey のオーケストラ作品 “Speaking’s” を見ても、音声の明瞭性が達成されているわけではない。

そこで、まず、音声の音韻性を犠牲にして韻律のみを表現した音色を造った。オーケストラ音として、ストリングス(弦楽)を選んだ。また、筆者が「いろは歌」を朗読し、この音声からピッチ分析を行い、これを離散化して音列素材とした。

音楽作品とするために、この音列に4声の和声を付与した。オーケストラ作品「音の音」では、この譜面により、ストリングスによる主テーマとして制作(作曲)された。この譜面の制作過程を図1に示す。

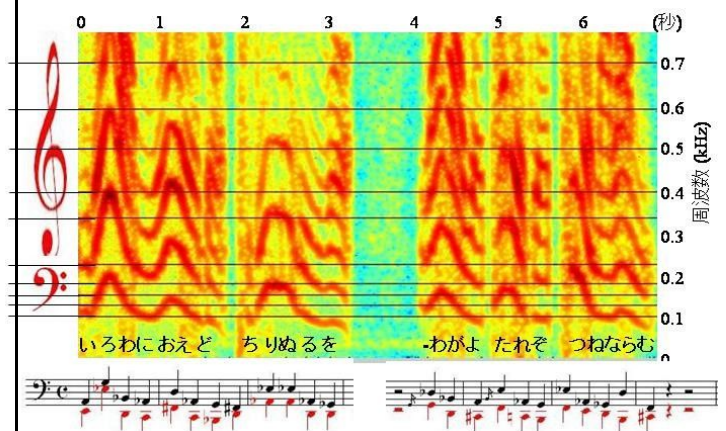


図1 「いろは歌」のソナグラム(声紋)と表示とそこから起こした譜面

ピッチ(音高)は最下段のヘ音記号内の曲線で、他はそのハーモニクス。譜面は上図の完全な書き起こしではなく、数回の朗読の平均の部分もある。譜面上声部(黒色)が書き起こした旋律で、ソナグラム上の連続線を離散化したものである。下声部はつけられたヘテロフォニー声部。

なお、このように、ピッチを指定せずに話された音声から音列を紡いでいく手法は Steve Reigh の Different Trains などでも使われている。

次に、別の作品「ハイブリッド・コラージュ」[作品]において、「音の音」エフェクトを付与した。まず、先のいろは歌の朗読音声を LPC 分析により、音韻スペクトルと音源とに分解し、音源部を、先の作品のストリングスの収録音を用いて、クロス合成すること

により音韻を付与した。

上記処理のみではサウンドハイブリッドであり、ストリングスの響きが弱い。サウンドハイブリッドと音の音との明確な差異は、後者は下位層の音色が明確に聞こえることである。そこで、ここに音源とした収録音を再度混合することにより、音の音エフェクトを達成した。

(2)NMF を用いた音の音エフェクト合成法

ここでは人間の音声を複数の楽音で構成した音、つまり「楽音による音声の合成」を試みた。以後、表現される音(音声)を**目的音**、目的音を構成する音(複数の楽音)を**要素音**と呼ぶこととする。音声の音韻性を保ちつつこれを楽音により表現する目的で、その合成法について検討した。

(2.1)NMF による音響信号分析

非負値行列因子分解 (Non-Negative Matrix Factorization; NMF) は、非負値のみからなる観測行列を、元のデータが持つ特徴パターンを示す基底行列と、その重み付けを示す係数行列の2つに近似的に分解するアルゴリズムである[2]。近年、音楽情報処理において自動採譜、雑音除去、音源分離など幅広く応用されている。

音響信号を扱う際、振幅スペクトログラムを観測行列とし、NMF により周波数スペクトルを表す基底と、その時変レベルを表すアクティベーションとに分解する。

本研究では分解時に乖離度を一般化 Kullback-Leibler ダイバージェンス (KL divergence) 基準で最小化する NMF アルゴリズム[3]を用いる。

(2.2)再合成アルゴリズム

NMF による処理後の振幅スペクトルから音響信号に再合成する際、適切な位相復元をする必要がある。本研究では Griffin らの LSEE-MSTFTM[4]により振幅スペクトログラムからの位相復元を行った。

(2.3)音の音の合成アルゴリズム

以下の学習、推定、再合成の3つの手順で音合成を行う。図2に学習と推定におけるNMFによる分解の様子を示す。

学習：要素音を NMF に適用し、要素音を構成する周波数スペクトルを抽出する。(図2、上) 1音あたり4基底ですべての要素音に対し NMF 反復計算を行い、得られた周波数スペクトル行列を学習データとして保存する。

推定：目的音である女性音声のパワースペクトログラムに対し、学習データを用いて NMF の適用をする。学習データで周波数スペクトル行列を固定し、アクティベーションのみ推定を行う。(図1、下)

再合成：推定により得られたパワースペクトログラムを LSEE-MSTFTM を用い音楽信号を合成する。

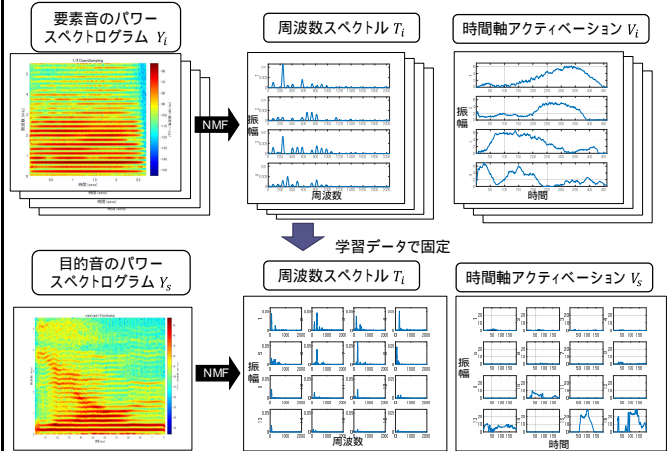


図2 NMF を用いた学習と推定

表1 使用楽器種と奏法

楽器名	奏法	ファイル名
バイオリン	ノーマル	151VNNOF
	ノンビブラート	151VNNVM
	Sul Ponticello	151VNPFCF
ビオラ	ノーマル	161VLNOF
	Sul Ponticello	161VLPCF
チェロ	ノーマル	171VCNOF
	Sul Ponticello	171VCPCF
コントラバス	ノーマル	181CBNOF
	Sul Ponticello	181CBNOF
尺八	むら息	353SHMIF
	ノーマル	353SHNOF
	ビブラート(強)	353SHVIF
ピブラフォン	ハードマレット/ノーマル	041VIHNF
	ハードマレット/ペダル	041VIHPF
	ハードマレット/ビブラート	041VIHVF

表2 合成音1~5の要素音の組み合わせ

	要素音 楽器	要素音数	基底数
合成音	弦楽器(Vn, Vl, Vc, Cb)	9	36
合成音	弦楽器(Vn, Vl)	4	16
合成音	尺八	4	16
合成音	ピブラフォン	3	12
合成音	尺八、ピブラフォン	7	28

表3 「音の音」エフェクト合成条件

学習反復計算	500回
推定反復計算	1000回
FFTポイント数	4096
分析窓	Modified hamming window[4]
窓幅	256
シフト幅	64
LSEE-MSTFTM更新回数	200回

(2.4) 合成条件

楽音による音声の合成にあたり、目的音は女声の母音「いえあおう」の音声データ、要素音には RWC 研究用音楽データベース：楽器音より、弦楽器(バイオリン、ピオラ、チェロ、コントラバス)、尺八、ビブラフォンの単独音、数奏法を用いた。使用した要素音の一覧を表 1 に、合成する際の要素音の組み合わせを表 2 に示す。

目的音・要素音のいずれもピッチは A2(220Hz 近辺)、すべてサンプリング周波数は 11025Hz とした。その他条件は表 3 に示す。

(2.5) 結果と考察

目的音のソナグラムを図 3 に、合成結果例として合成音 1、3 のソナグラムを図 4、5 に示す。合成音 1、5 のように、目的音を表す要素音を多く用意するほど音韻性(音色変化)が表れたが、その楽器らしさは損なわれた。一方、要素音が少ないと楽器らしさは保たれるものの、音韻性に劣る結果となり、音韻性と楽音性はトレードオフの関係といえる。

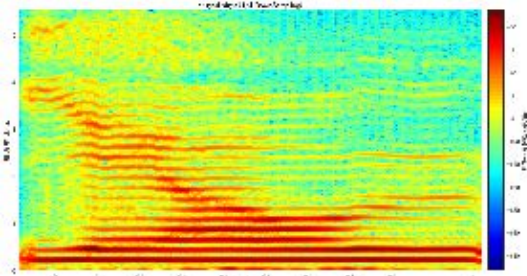


図 3 目的音(女声)のソナグラム

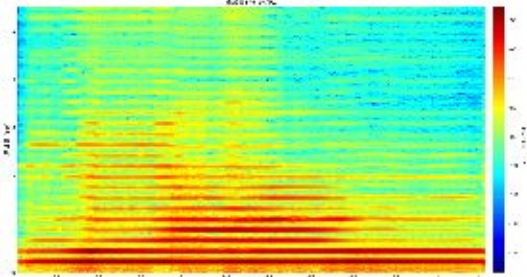


図 4 合成音 1(要素音：弦楽器)のソナグラム

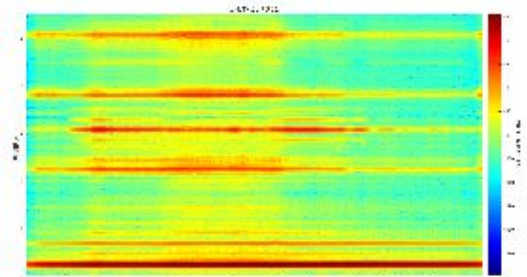


図 5 合成音 4(要素音：ビブラフォン)のソナグラム

(3) まとめ

新たな音色の粹組みとして、**構造的音色**という概念を提唱し、サウンドモーフィングとサウンドハイブリッドを**拡張音感**としてその一部に位置づけ、新たに**音の音**というエフェクトについて検討した。現実に存在しない新たな音色の開拓であるため、1) 製作された合成音の特性も検討しつつ、2) その音楽的価値を問うために、作品にも反映させた。両社は 1) の技術応用として 2) を行うことが望ましいが、実施上の利便から、両検討を独立して行った。

「音の音」エフェクトを組み入れた楽曲制作を行った。作品では、サウンドハイブリッドを入れ子にしてストリングスによる「いろは歌」朗読音声が付与した。

また、新たな「音の音」合成音技術を NMF を用いて提案した。いずれの検討も、目的音(上位音)を明確に(音声では明瞭性)行わせようとするほど、要素音の知覚が難しくなり、逆も成り立つ。音の音という新たなエフェクトとその合成音方式を提案し、これらの特性を明らかにした。

したがって、音声に関する「音の音」エフェクトは、音韻性を多少に犠牲にして楽音(要素音)の知覚を明確にする、という適切な知覚点を求めること「音の音」の音質向上であり、これが今後の検討課題である。

また、目的音を音声以外としたより広範な「音の音」エフェクトについて発展させていく。

参考文献

- [1] 茂出木敏雄, “音響信号の平均律音階に基づく汎用解析ツール「オート符」の開発”, 電気学会・電子情報システム部門誌, Vol.123-C, No.10, pp. 1768-1775 (2003).
- [2] D. D. Lee and H. S. Seung, “Learning the parts of objects with nonnegative matrix factorization,” *Nature*, vol. 401, pp. 788-791 (1999).
- [3] P. Smaragdis and J. C. Brown, “Non-negative matrix factorization for polyphonic music transcription,” *WASPAA*, pp. 177-180 (2003).
- [4] D. W. Griffin and J. S. Lin, “Signal Estimation from modified Short-Time Fourier Transform”, *ASSP-32*, April 1984, pp. 236-242 (1984).

5. 主な発表論文等

[学会発表](計 8 件)

柴田理央、小坂直敏, 「音の音 NMF を用いた音の異音表現とそのエフェクト応用」, 情報学会全国大会 7L-05, 名古屋大学, 2017.3.18.

作品

aotoshi Osaka, Present-Day Jakuchu Series: Butterfly Pictures “Mourning Cloak”, for violin, video and electro-acoustics, NWEAMO 2016 (in JSSA

Music Festival 2017), Jun Kurumisawa (Video), Gregory Walker (Violin), 公園通りクラシックス(東京渋谷), 2016.11.12【招待】

小坂直敏, 今若冲蝶の図より「オオムラサキ」 ヴァイオリン、μgic、電子音響のための、主催 先端芸術創作音楽学会、Sonic Arts Projects Vol.4 No. 1, ドイツ文化会館 OAG ホール(東京赤坂), 木村まり (Violin), 2016.9.30

小坂直敏, 「ハイブリッド・コラージュ」 ピアノと電子音響のための、NYCEMF (New York City Electro-acoustic Music Festival)2016, Abron Art Center, Playhouse (New York) Julia Den Boar(piano), 2016.6.16【審査あり】

小坂直敏, 同作品, 主催 東京電機大学未来科学部、Media Project Vol. 13, 現在の電子音響音楽六景、すみだトリフォニー小ホール(東京錦糸町), 大井浩明(Piano)、2015.10.16.

小坂直敏, 「Morphing Collage」 ピアノとコンピュータのための、NYCEMF (New York City Electro-acoustic Music Festival)2015, Abron Art Center, Playhouse (New York) 小坂紘未(piano), 2015.6.27 【審査あり】

小坂直敏, 同作品, 夢想田音楽館 Taiwan Sonic Arts Festival 2015 会場: 夢想田音楽館(台南市内) 2015.2.28 【招待】

小坂直敏, 「音の音 オーケストラのための」 オーケストラプロジェクト 2014, 管弦楽: 東京交響楽団, 指揮: 大井剛史、東京芸術劇場コンサートホール(東京池袋)、2014.12.1

[その他] (計3件)

啓蒙を目的とした音楽企画

小坂直敏, JSSA 音楽祭 2016 拡がりの始まり 以下 a)-e)5 件の総合企画。(c)は直接企画(=)

主催: JSSA(先端芸術音楽創作学会)、後援: JSEM(日本電子音楽協会), JFC(一社 日本作曲家協議会), JSCM(日本現代音楽協会)

a) 高岡明, NWEAMO フェスティバル 2016 東京, [公園通りクラシックス(渋谷)], 2016.12.12/13

b) 森威功, Asia Computer Music Project Concert - Sonic Arts Project Vol. 4 No. 2【洗足学園音楽大学ブラックホール 1F ビッグマウス】 2016.10.1

c) 小坂直敏, Sonic Arts Project Vol.4 No. 1, 演奏システムと音楽, [OAG ホール(東京赤坂)], 2016.9.30

d) 水野みか子, Electroacoustic Concert in Nagoya, featuring petites forms [名古屋市立大学芸術工学部 大講義室] 2016.9.17

e) 高岡 明, Amphibian Music/Video Series New York 教徒コンサート [ゲーテ

インスティトゥート・ヴィラ鴨川] 2016.7.2

小坂直敏, Media Project Vol. 13 現在の電子音響音楽六景 主催: 東京電機大学 未来科学部 講演: 日本現代音楽協会、(一社)日本作曲家協議会、先端芸術音楽創作学会[すみだトリフォニー小ホール] 2015.10.16.

小坂直敏, 藤原嘉文, 中村滋延, 柳田孝義 オーケストラプロジェクト 2014 2014 プロジェクト 35 年の歩みを経て未来への新たな胎動, 主催および場所 オーケストラプロジェクト [東京芸術劇場コンサートホール(池袋)] 2014.12.1.

ホームページ等

<http://www.srl.im.dendai.ac.jp/people.otsuka/>

6. 研究組織

(1) 研究代表者

小坂直敏 (OSAKA NAOTOSHI)

東京電機大学・未来科学部・教授

研究者番号: 20366378