

図形の視覚的印象に基づいた音色生成によるインタラクティブな音楽体験の提案

林田航[†] 望月茂徳[†]

概要: インタラクティブな音楽インスタレーションは、入力インターフェースの開発と共に発展してきた。本研究では、言語学における音象徴の存在を下敷きに、図形の入力から取得した、凸性の欠陥や輪郭の頂点に占める鋭角の割合などの特徴量に基づき音色を決定するシステムを開発する。このシステムを用いた体験型作品を展示しアンケート調査を行ったところ、74.7%の体験者が図形と音色に関連性を感じる事が出来た。

1. 背景と目的

本研究は、インタラクティブな音楽インスタレーションの制作を行うことを目的とする。その手段として、図形の視覚的印象と音色の関連付けを提示する体験型インスタレーション作品の制作、及びその実装のための入力インターフェース開発を行う。

1.1 インタラクティブな音楽インスタレーション

インタラクティブな音楽インスタレーションは、インタラクティブ性を実装するために様々な入力インターフェースによって開発されている。岩井俊夫による、トラックボールで描いた線を使ってピアノが演奏される作品『映像装置としてのピアノ』[1]は、インタラクティブアートにおける「操作の入力/体験の出力」が、音楽における「記譜/演奏」に対応づけられている。これはインタラクティブな音楽インスタレーションの制作における基本的な考え方の一つであると言える。また、机に埋め込まれたインターフェースの操作によって演奏される音楽が変化する作品『テーブルの上の音楽』[2]は、映像をインターフェースにプロジェクションマッピングすることにより、マーカーの直感的な操作を提示している。ponboksによる、『置く位置で高低の変わるひらがな楽器』[3]は再帰性反射材を使用した、赤外線カメラによる入力を行っている。これにより投影映像の影響を受けずに画像処理を可能にしている。

1.2 音色の生成について

音楽の聴取において、音楽の構造よりもむしろ音の響きそのものに注目するという視点が存在している。ジョン・ケージの『4分33秒』は、演奏者がステージに上がり、4分33秒間何の音も出さずに演奏を終了するという作品である。この作品では、沈黙の中で聞こえる、生活空間を取り巻く音が、作曲された音楽の響きに劣らない価値を持っていることを示すことが試みられている[4]。佐々木は、音響派という日本の音楽ジャーナリズムの用語について、その前提には音の響きへの意識的なこだわりと、音を「発すること/奏でること」から「聴くこと/耳を澄ますこと」への重心の移動もしくは展開があるとしている[5]。これらから、音楽の聴取において音の響きを意識することが喚起

されており、一般に行われていると考えられる。

そういった音の響きを重視して、音作りをすることにおいて、シンセサイザは重要な手段である。シンセサイザでの音作りは音響技術の進歩と共に確立してきた。1960年代にはアナログシンセサイザが楽器として一般に知られるようになり、モーグ III で制作された W.カルロスの『スイッチト・オン・バッハ』、富田勲の『雪は踊っている(Snowflakes are Dancing)』などのレコードがヒットを収めその地位を確立した。1980年代に入ると音づくりにデジタル技術が使われるようになり、現在では音づくりはコンピューターに象徴されるデジタル機器を用いて行われている[6]。アンビエント・ミュージックの作曲家として著名なブライアン・イーノは自身の楽曲制作について、「音作りこそが新しい音楽の独自性であり、それこそが作曲上の関心の中心となるべきだと示したかった」と説明している[7]。

1.3 音象徴が視覚に与える影響

音自体が印象を持っているという考えを音象徴(sound symbolism)、または音感覚性(phonesthesia)と言う。Ramachandran と Hubbard によって命名された「ブーバ/キキ効果」の実験[8]は、2つの図形(図1)を見せ名前を選ばせるものであり、95%の被験者が左の直線で描かれた図形がキキであると回答している。この結果は、図形を構成する線の方向の変化と、音韻的抑揚の類似によるものであると考察される[8]。

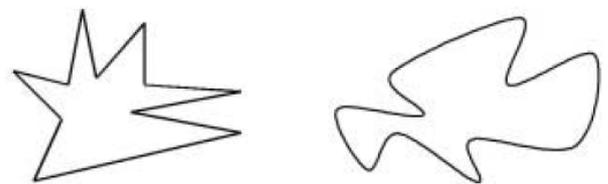


図1 ブーバとキキ [8]から引用

音象徴に関する実験では、多くが文字を使って言語音を視覚的に呈示する形で行われてきた。平田、浮田による実験[9]では、明度の違うひらがなの画面表示による刺激、ストランビーニらによる実験[10]では紙に印刷した無意味語

のリストによる刺激が用いられた。このことを課題とした板垣と小林による実験[11]の結果からは、聴覚的刺激によっても音象徴が発生し、音象徴が視覚の大きさ判断に影響を与えることが示唆される。

2. 作品『形象に傾聴する』の制作

2.1 作品概要/体験の方法について

本研究では、音楽の聴取においても聴覚刺激による音象徴のようなものが存在すると仮定し、図形の視覚的印象と音色の関連付けを提示するシステム開発を行った。具体的には、画像処理によって入力画像内の輪郭を検出し、その輪郭から得られる情報をソフトウェアシンセサイザのパラメータに代入することによって音を生成する。また検証方法としては、粘土を用いて卓上の操作キャンパス(図2)に図形を描画することによって、音楽と映像に変化が起こる体験型作品として実装した。体験者は、キャンパスの描画内容を変化させることによって、リアルタイムに音楽と映像の変化を鑑賞することが出来る。

音楽は、周期的なリズムのドラムトラックと、操作によって変化する主旋律の、二つの要素から構成される。図形が描画されているとき、予め設定したリズムで、図形の特徴量から生成された音が演奏される。また、複数の図形が描画されているとき、それぞれの図形の特徴量から生成された音が順番に演奏される。音の生成アルゴリズムについては、2.2にて詳細に記述する。

映像については、音声のタイミングに合わせて、描画した図形が光る映像が自動生成され、操作キャンパスに投影される。



図2 操作キャンパスへの操作の様子

2.2 音色合成のアルゴリズム

本作品では画像処理によって入力画像内の輪郭を検出し、その輪郭から得られる情報をソフトウェアシンセサイザのパラメータに代入することによって音の生成を行っている。画像からの輪郭の検出と情報の抽出にはOpenCV4.5.2を使用し、パラメータを渡すシンセサイザにはNativeInstrument社のMassiveを使用した。

輪郭から得られる情報には様々なものが考えられるが、今

回使用した輪郭から得られる情報と、それらを反映させる音の要素は表1に示した通りである。今回は、図形の重心座標や面積といった基本的な情報に加え、図形の視覚的印象を数値に起こすための「デコボコ度」、及び「ギザギザ度」という指標を独自に定義して制作を行った。本研究の目的には、作品に用いたこれらの指標の妥当性を検討することも含まれる。

表1 輪郭から得られる情報と反映させる音の要素

輪郭から得られる情報	情報を反映させる音の要素
重心のX座標	定位(音の聴こえる方向)
重心のY座標	音高(音程 ピッチ)
面積	音量(ボリューム)・音価(音の長さ)
デコボコ度	音色の複雑度
ギザギザ度	音色の鋭さ

2.2.1 デコボコ度

点集合について、点すべてを含む最小の凸多角形のことを、凸包と言う[12]。輪郭において凸性の欠陥とは、輪郭の頂点について凸包を考える際に失われる、輪郭の内側への凹みのことである。本研究ではデコボコした図形を判断する指標として、輪郭の頂点における凸性の欠陥を採用し、その数を「デコボコ度」と定義する。

「デコボコ度」を音に反映させる考え方としては、凸性の欠陥数に合わせて、用意した音色を切り替えるという方法を採用した。異なる複雑さを演出した3種類の音色を制作し、これらを図3の通りに対応づけを行った。最も単純な音色は矩形波と正弦波の中間のもの、中程度に複雑な音色はフィルターのレゾナンスによる倍音を付加したもの、最も複雑な音色は前述のフィルターの遮断周波数を時間的に変化させたものを使用した。



図3 凸性の欠陥の例

表2 凸性の欠陥数と対応するオシレーター

凸性の欠陥数	対応するオシレーター
0以上, 1以下	最も単純な音色
2以上, 4以下	中程度に複雑な音色
5以上	最も複雑な音色

2.2.2 ギザギザ度

本研究では、ギザギザした図形を判断する指標として、

輪郭の頂点に占める鋭角の割合を採用し、「ギザギザ度」と定義する。輪郭に含まれるすべての頂点に対して、鋭角か否かを判断し、(鋭角数)/(頂点の全数)で鋭角の割合を算出する。

「ギザギザ度」を音色に反映させる考え方としては、上述の鋭角割合に比例して波形フォルマントを移調するという方法を採用した。フォルマントとは、波形をフーリエ級数展開することで得られるスペクトルエンベロープにおけるピークのことであり、倍音成分の含まれ方を表す[9]。

図4に共通のオシレーターにおけるギザギザ度が最小の場合(図4a)とギザギザ度が最大の場合(図4b)の合成波形のスペクトルエンベロープを示す。



図 4a ギザギザ度が最小の場合のスペクトル分布

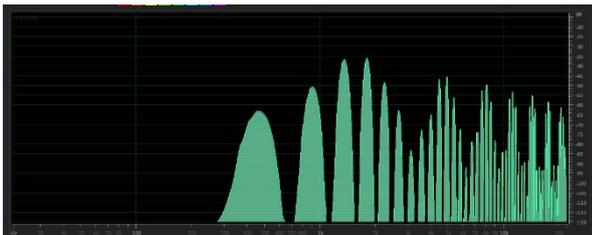


図 4b ギザギザ度が最大の場合のスペクトル分布

2.3 作品全体のシステム構成

図5に、体験者が描画行為による入力を行い、映像及び音声のフィードバックが出力される一連の流れを示す。システムの開発はプログラミングソフトウェアであるTouchDesignerで行った。

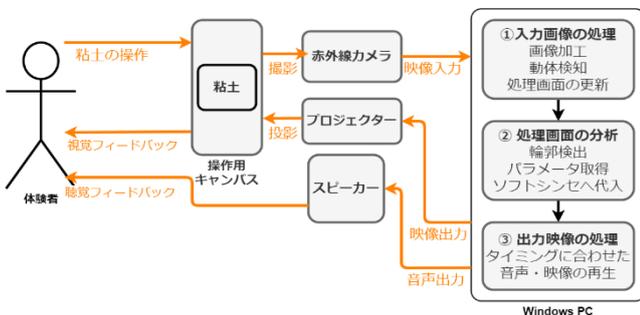


図 5 作品全体のシステム構成

映像の入力と出力は操作キャンバスに対して垂直真上に固定したデバイスから行う。映像の入力には、赤外線カメラを使用した。これははっきりとしたエフェクトを投影しつつ粘土への操作にのみ動体検知を行うためである。映像

の出力は、小型のプロジェクターを用いて操作面に真上から投影する。これは、入力映像上の操作用キャンバスの歪みを抑え、また投影に写り込む体験者の影を削減するためである。音声は2チャンネルステレオで出力するため、2つのスピーカーを操作用キャンバスの左右に設置した。



図 6 操作キャンバスの垂直真上のデバイス



図 7 映像の投影

3. 展示

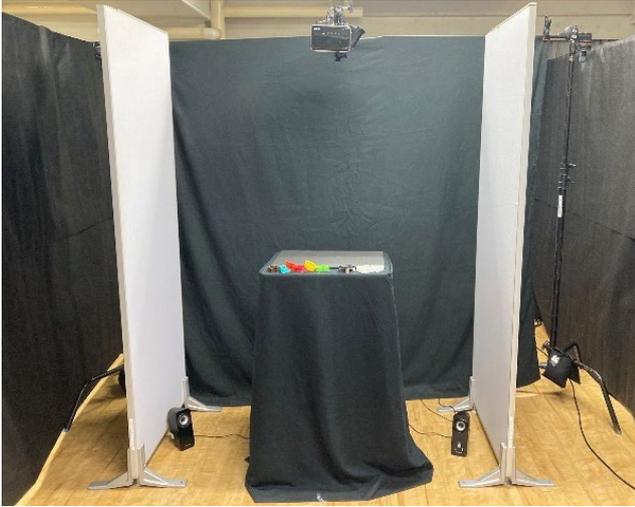


図 8 作品の展示の様子

作品は、11月10日に立命館大学以学館多目的ホールにて、対面型展示を行った。その際に回収した82件のアンケート結果を元に考察を行う。

4. 結果

4.1 音色と図形の結びつきを感じた経験はあるか

この設問は、体験者の経験から、一般的に音色と図形の形状の結びつきがどれくらい意識されているのか確かめることを意図して設定した。結果は以下の図9に示す通りである。

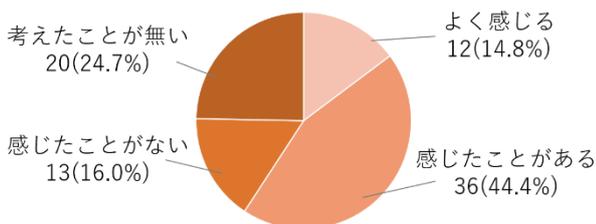


図 9 音色と図形の結びつきを感じた経験はあるか

4.2 図形と再生される音色に関連性を感じたか

この設問は、作品によって図形の視覚的印象と音色の関連付けを提示できたか確かめるために設定した。結果は以下の図10に示す通りである。

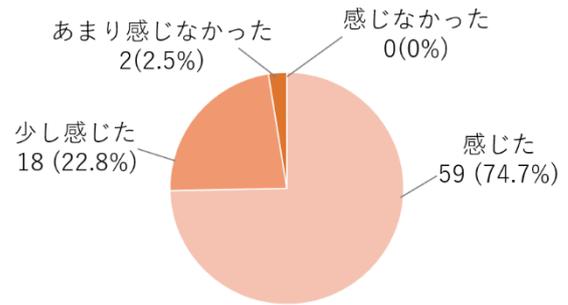


図 10 図形と再生される音色に関連性を感じたか

また、この設問に付け加えて、使用した図形の情報それぞれについて、どの程度図形と関連性を感じたか、感じなかった場合を1、最も感じた場合を6として、6段階評価で回答を求めた。結果は以下の図11に示す通りである。

回答項目 / 評価	X座標	Y座標	面積	デコボコ度	ギザギザ度
1	1	1	0	0	0
2	2	0	0	3	1
3	9	5	2	7	5
4	16	8	9	15	11
5	22	24	17	22	24
6	30	43	54	33	37

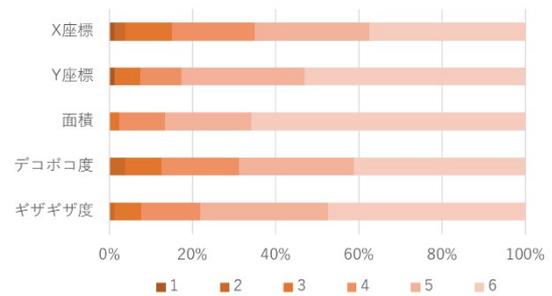


図 11 個別6段階評価

4.3 入力によって音色を操作していると感じたか

この設問は、インタラクションデザインの観点で、操作性の評価を確かめるために設定した。結果は以下の図12に示す通りである。

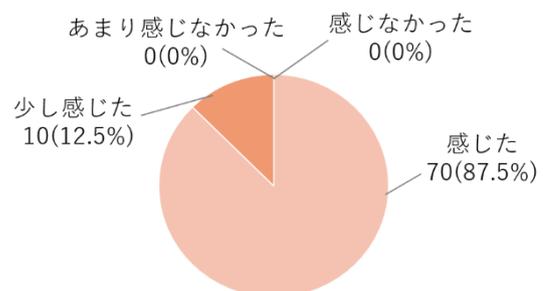


図 12 入力によって音色を操作していると感じたか

4.4 その他 意見や感想(自由記述)

この設問は、アンケートの設問に関わらず体験者の自由な意見を求めるために設定した。結果を分類し、以下に記述する。

- 体験の感想 : もよもやしていた感覚が体験できて面白かった。音楽を作れて楽しかった。新鮮な体験で面白かった。インタラクティブだと感じた。音と図形の対応関係が分かりやすかった。
- 改善点 : 大きさ以外の違いが分からなかった。「音価」はどちらかというと「音の鋭さ」に関係するものなのかなと感じた。
- 良い点 : 直感的な操作性。リアルタイムの映像投影。どんな図形でも体験可能である。変化するのが音程だけではない。
- 発展案 : 図形の演奏される順番の操作。もっと大きな面積で大人数でできるような形へのアップデート。聴覚が不自由な方に対して音楽を視覚的に伝えられる。ライブパフォーマンス・幼児教育などへの応用。

5. 考察

4.2の結果より、回答者の74.4%が図形と音色に関連性を感じたと回答していることから、図形の視覚的印象と音色の関連付けを提示する音声合成アルゴリズムの開発という点で達成感があった。また、4.1で「感じたことがない」「考えたことがない」と回答した人の全員が、「感じた」「少し感じた」と回答している。このことから、音色と図形の結びつきを感じたことが無かった回答者の多くが作品を通して音色と図形の結びつきを感じる事が出来たと考えられる。

また4.2の個別6段階評価ですべてのパラメータの評価について6と回答した人数が一番多く、5、4と順に減っていく結果となった。また、2.2で独自に定義した指標についても、デコボコ度で86.4%、ギザギザ度で92.3%が4以上の肯定的な評価を付けた。このことから、今回使用した図形の特徴量と反映させるパラメータの組み合わせにある程度の妥当性が考えられる。また、4.4で改善点として、音の変化が分からなかったり、他のパラメータに変更した方がいいという指摘も見られたことから、今回定義した図形の特徴量と反映させるパラメータの組み合わせには、まだ検討の余地が残っている。

また4.3で「あまり感じなかった」「感じなかった」と回答した人が居なかったことから、インタラクションデザインという観点からは、直感的な体験を支えるインターフェース開発に成功したと言える。今回制作した、音色を再生

するシステムは、一定のリズムで図形から生成した音を鳴らすというものであった。入力による音の変化を、音楽への操作として捉えると、4.4の発展案として挙げたように、今回操作を加えなかったリズムに対する操作というアプローチについても考える余地が残されている。

参考文献

- [1] 岩井俊夫「映像装置としてのピアノ」(2022年12月16日確認) https://www.ntticc.or.jp/pub/ic_mag/ic016/lyon/lyon_i6_j.html
- [2] 岩井俊夫「テーブルの上の音楽」(2022年12月16日確認) https://www.ntticc.or.jp/en/feature/1999/+Works/composition_j.html
- [3] ponboks「置く位置で高低の変わるひらがな楽器」(2022年12月16日確認) <https://ponboks.com/prototype/hiragana/>
- [4] Paul Griffiths 『Cage』 Oxford University Press, 1981(ポール・グリフィス, 堀内宏公(訳)『ジョン・ケージの音楽』, 青土社, 2003)
- [5] 佐々木敦, 『テクノイズ・マテリアリズム』, 青土社, 2001
- [6] 岩崎真, 小泉宣夫『サウンドシンセシス』, 講談社, 2011
- [7] Brian Eno 『A Year With Swollen Appendices』 Faber & Faber, 1996(ブライアン・イーノ, 山形浩生(訳)『A YEAR』, 株式会社 PARCO 事業局出版部, 1998)
- [8] Ramachandran, V.S. & Hubbard, E.M. 『Synaesthesia: A window into perception, thought and language』, Journal of Consciousness Studies 8 巻 12 号 pp.3-34, 2001
- [9] 平田佐智子・浮田潤, 『音韻象徴・音象徴と身体性一言語認知過程研究をベースとした実験的アプローチの提案』, 関西学院大学人文論究 58 巻 2 号 pp.49-63, 2008
- [10] ストランビーニ ニコーラ, 備瀬優, 矢野雅貴, 坂本勉『大きさと関連する音象徴について』, 2012年度日本認知科学会第29回大会発表論文集 pp.790-796, 2012
- [11] 板垣沙知, 小林耕太『Cross-modal Effect of an Aurally Presented Phoneme on the Judgment of a Visual Object's Size』, 同志社大学ハリス理化学研究報告 57 巻 4 号 pp.239-243, 2017
- [12] Jack Sklansky 『Finding the convex hull of a simple polygon』, Pattern Recognition Letters 1 巻 2 号 pp.79-83, 1982