

# 子供に向けた平易な演奏体験を提供する インタラクティブな音・映像インスタレーションの制作

須藤美帆<sup>†</sup> 赤池龍成<sup>†</sup> 森涼華<sup>†</sup> 張天弛<sup>†</sup> 鈴木彩音<sup>†</sup>

**概要:** 本研究では、子供を中心とした体験者が視覚や身体の動きを通じて音を楽しめるインタラクティブな音・映像インスタレーション「らくガッキ」を開発し、2024年のMaker Faire Tokyoで展示を行い、アンケート調査と入力データの収集を実施した。視覚情報と音色の対応や身体的自由度を高めた操作が、体験価値を向上させる可能性が示唆された一方、独自特徴量「複雑度」に偏りが見られる課題が確認された。

## 1. 目的

子供を中心とした体験者の楽しさにつながる要素仮説を実現する手段として、インタラクティブな音・映像インスタレーションの制作を行い、目的とする。

著者らは、子供を中心とした体験者にとっての「楽しさ」につながる要素として、以下3点を仮説とした。

- 視覚情報を基にして音を対応付けること: 体験者が描画した図形形状を基に音を対応付けること
  - 身体を使った自由度の高い演奏方法
  - 他の体験者の絵にどのような音が対応付けられているか確認できること: 結果の多様性が確認できること
- 仮説を実現する手段として、視覚的印象と音色を関連付けて提示する、体験型の入力図形に基づいた音色発生装置「らくガッキ」[1]を開発・実装した。そして、2024年9月に開催されたMaker Faire Tokyo 2024[2]（以下MFT24）にてデモ展示を行い、体験・アンケートの収集を行った。

## 2. 関連研究

本研究では、仮説1の内容に対し、独自の手法を検討・実装した。そのため、本節では、仮説1に関する関連研究を記す。

### 2.1 視覚情報と音の心理的關係性

視覚情報と音の心理的關係性については、Tyan-Yu WyuとAlan Leeら（2017）が、視覚的形状と音の属性（音圧や音高）の認知的な一致關係を検証した[3]。この研究では、「見た目の大きさと音圧」および「鋭角の形状は大音量・高音、湾曲した形状は小音量・低音」といった特徴が示されている。視覚刺激としては、人間、車、動物（自然物）、円錐、立方体、球体（人工物）の6種類が用いられた。さらに、西田沙織ら（2009）は、円、正三角形から六角形、星、10本のとげを持つ形状、曲線的な複雑形状、直線的な複雑形状の計9種類の図形を用い、これらに対応する聴覚刺激を調査した[4]。この研究では、振幅と時間を軸に取った音の特徴（周期的・非周期的）を6パターン設定し、計11種類の音を対応付けて検証した。その結果、非周期的な音は複雑な図形と結びつきやすく、周期的な音はシンプル

で滑らかな形状と関連する傾向が示されている。

### 2.2 入力動作と音の關係

入力動作と音の關係については、Konstantina MargiotoudiとFriedemann Pulvermüller（2020）が、人がペンで円や三角を描く際、その動きと音の特徴が形状の特性（丸い、鋭い）と類似していることを明らかにした[5]。また、形状に対応する音色の付与については、林田航ら（2023）が、粘土の面積を音量や音価に、へこみの特徴量を音色の複雑さに、輪郭の鋭角の割合を音色の鋭さに関連付けた研究を行った[6]。この研究では、これらの対応關係が参加者の体験において形状と発音の結びつきを形成することが示されている。

### 2.3 本研究の位置づけ

これらの先行研究を踏まえ、本研究では、図形形状の「滑らかさ」「形状の複雑さ」の二軸をとり、これらに対応する音の印象を著者らの判断により適切に設定した。特に、規定された図形だけではなく、体験者が自由に描いた図形に基づいた音の対応付けを実現することを目指した。

## 3. 制作方法

### 3.1 体験概要

制作した「らくガッキ」は、一筆書きで描いた図形に音色やキャラクタパーツを付与し、手に追従するポインタ（カーソル）で図形に触れることで音を出すインタラクティブなプロトタイプである。具体的には、ペンタブで描いた図形の特徴量を基にキャラクタ化し、モニタ上に表示した楽器オブジェクトを、カメラによる手のトラッキングを用いて演奏操作できる体験型装置として実装した。以下図1に全体のシステム構成を示す。

図1のように、全体のシステム構成は体験者が図形を描画する入力部（図2）と、楽器オブジェクトを操作する演奏部（図3）の大きく2つに分けられる。

なお、体験用装置として、入力部に使用するペンタブ1台、演奏部に使用するモニタ2台をWindows PC 1台に接続し、これらを一体の装置として構成した。入力部の実装については3.2にて、演奏部の実装については3.3にて記述する。なお、入力部、演奏部のソフトはそれぞれをUnity[7]アプリケーションとして制作し、併せてアプリケー

ション間でデータを移動させる仕組みを実装した。

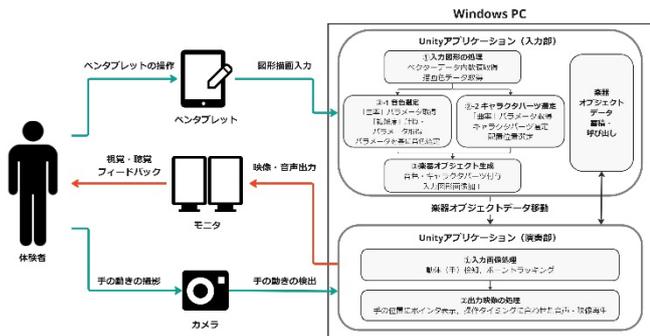


図1 「らくガッキー」全体のシステム構成



図2 (左図) 入力部：ペンタブレットにタッチペンを用いて操作を行う

図3 (右図) 演奏部：大型のモニタの前にカメラを設置しハンドトラッキングで操作を行う

### 3.1.1 体験仮説と実装の関連について

本制作は、第1節で述べた体験仮説を基に実装方法を検討した。

仮説1では、描いた図形の特徴量に基づき音色やキャラクターパーツを選定するシステムを実装し、視覚と音の対応を直感的に体験できる仕組みを構築した。

仮説2では、ハンドトラッキングによって手の動きに応じた音の発生を実現し、自由度の高い演奏体験を提供した。

仮説3では、他者が描いた図形から生成した楽器オブジェクトを複数同時に表示し、体験者同士が相互にオリジナリティを感じられる仕組みを導入した。

## 3.2 入力部

本制作の入力部では、ペンタブレットで一筆書きの図形を描画し、その図形をベクターデータとして扱う。図形の特徴量「曲率」と「複雑度」を計算し、対応する音声ファイルやキャラクターパーツを選定して付与することで、図形をキャラクタ化し、楽器オブジェクトとして再構成する。

### 3.2.1 図形形状と音の対応

図形形状と音の対応付けについて、図形形状を分類する軸となる尺度を決め、続いて尺度の印象と合う音の選定を行った。なお、図形形状と音の印象対応については、2節にて触れた先行研究を踏まえ、著者らの判断により適切と考えられる形で対応付けを行った。

用意する音声ファイルは、子供をメインターゲットとすることから、教育現場で使用される音楽制作Webアプリケーションカトカトーン[8]内の音色ラインナップを参照し、

著者らが面白いと感じた音色の種類を選定。その後、類似の音声ファイルを音楽制作ソフトウェア Garage Band[9]を用いて作成した。

本制作では、曲線的/直線的を対とした5段階尺度：滑らかさ(以下尺度イ)と、複雑/単純を対とした3段階尺度：形状の複雑さ(以下尺度ロ)の二軸で、合計15種類の音から対応するものを選定する仕組みを採用した。図4は本制作に向け著者らが作成した図形形状と音色の対応を示す。

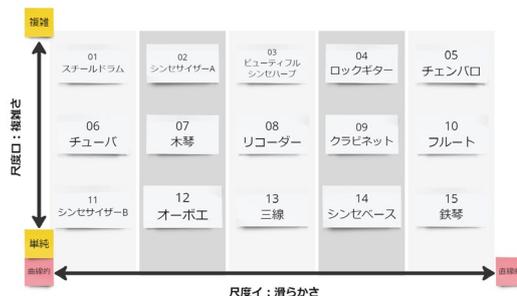


図4 図形形状と音色の対応表

横軸を滑らかさ、縦軸を複雑さとし、5×3の15の判定に区分して音を対応させた。

### 3.2.2 尺度に対応する特徴量計測手法の選定

尺度に対応する特徴量計測手法として、尺度イは図形のベクターデータ内の「曲率」の値を、尺度ロは図形のベクターデータ内の数値を基に計算する独自特徴量数値「複雑度」を用いた手法を実装した。

複雑度を算出する手法としてフローAとフローBの2つのアプローチを用いた。フローAでは、入力線における曲がり部分の割合を考慮し、外枠の角に丸み(R)を加えた基準線を設定し、その基準線長と入力線の総線長との差分から複雑度を算出する。一方、フローBでは、入力図形を囲う最小の外枠(Bounding Box)の幅と高さを基に、最も単純な形状として直線や円弧を想定し、基準線長を導出する。この基準線長と入力線の総線長を比較し、その差分を基に複雑度を算出する。これら2つのフローにおいて得られた複雑度の数値を比較し、大きい方の値を最終的な複雑度とする。以下図5にて、複雑度の算出フローの概要を示す。最終的な複雑度は閾値に基づき、単純、中間、複雑の3段階に分類される。

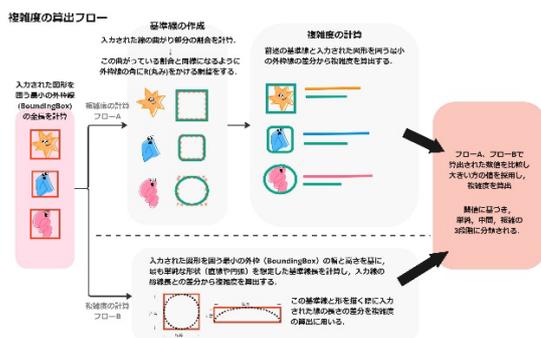


図5 独自特徴量数値「複雑度」の算出フロー

### 3.2.3 音色選定のフロー

本制作では、入力図形に対応する音を割り当てるために、「曲率」および独自特徴量「複雑度」を基に音色選定を行った。

まず、曲率の値に応じて5段階（レベル0~4）に分類する。具体的には、曲率が0.2未満の場合はレベル1、0.2以上0.4未満でレベル2、0.4以上0.6未満でレベル3、0.6以上0.8未満でレベル4、0.8以上の場合はレベル5とする。次に、複雑度の値に応じて3段階（レベル1~3）に分類する。具体的には、複雑度が0.4未満の場合はレベル1、0.4以上0.7未満でレベル2、0.7以上の場合はレベル3とする。曲率レベルと複雑度レベルが決定した後、これらの組み合わせに基づき、音色番号を発行する。最後に、発行された音色番号に対応する音色名を取得し、入力図形の音として割り当てる。

なお、分類するための閾値について、曲率は5段階で単純に分割し、複雑度は独自に開発を行ったため、ペンタブレットを用いた図形入力と数値計算の結果を確認していくことで調整した。

### 3.2.4 キャラクターパーツの選定

図形形状の特徴量を視覚的にも感じられるよう、入力図形に付与するキャラクターパーツの選定手法を検討・実装した。具体的には、入力された図形の特徴量「曲率」数値を基に、キャラクターパーツを図6のように選定する。

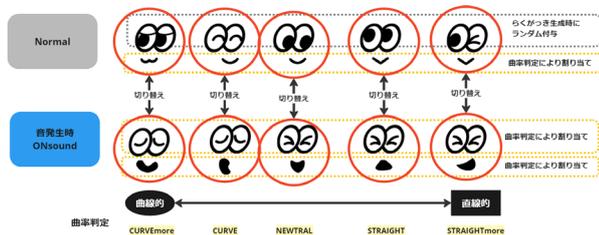


図6 曲率数値の判定に基づくキャラクターパーツ選定手法

なお、この仕組みは、目や口のパーツの形状から、尺度イの対応感を視覚的にも感じられるようにすることをねらいとする。

## 3.3 演奏部

本制作の演奏部では、モニター2台とWebカメラを用い、体験者の手をトラッキングして画面内にポインタを表示する。楽器オブジェクトは画面に4つ表示され、ポインタを重ねると対応する音が発生する。演奏は両手で操作可能で、複数人で同時に体験できる。また、他者が描いた図形から生成された楽器オブジェクトも表示され、相互にオリジナリティを感じられる仕組みを実装した。

### 3.3.1 ハンドトラッキング手法とポインタの実装

本制作では、操作に用いるハンドトラッキングの実現に、Googleが提供するオープンソースライブラリMediaPipe[10]をUnity向けプラグインとして使用し、リアルタイムで手の位置や指の動きを検出するシステムを構築した。

MediaPipeで取得できる手のボーン上座標に演奏用のポインタを追従させるため、図7のようにポインタが表示されるよう実装した。

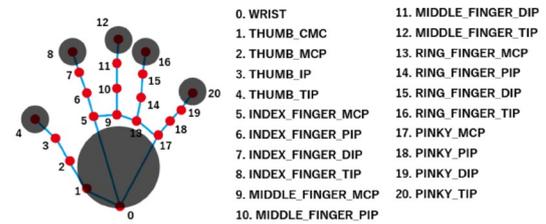


図7 MediaPipeで取得できる手のボーン上座標とポインタ表示位置の対応

### 3.3.2 音階の分布設定

演奏操作において、身体をより動かしたくなる仕掛けとして発生する音の音階操作を検討・実装した。具体的には、ポインタが楽器オブジェクトに重なるごとに一音鳴り、その音階は、楽器オブジェクトの基となる図形の描画線内の座標を基に割り当てる仕組みを実装した。図8は、図形の描画線を基にした音階の割り当て方を示す。

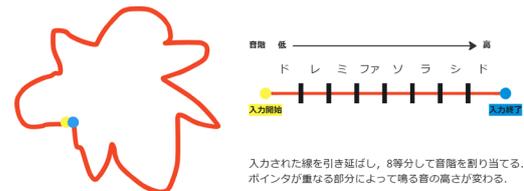


図8 図形形状と音色の対応

なお、この仕組みは、一筆書きで入力する図形形状に特別な制約を設けることなく、すべての楽器オブジェクトから8音階の音を発生させることを可能にするのをねらいとする。

### 3.3.3 視覚的なフィードバック

演奏操作において、音が聞こえづらい条件でも演奏を楽しめたり、身体をより動かしたりすることを促す仕掛けとして、視覚的なフィードバックを検討・実装した。具体的には、音発生時のキャラクターのモーション・音符エフェクトのモーションを、図9のように実装した。



図9 音発生時のモーション：発音に合わせて、キャラクターが小さく拡大縮小し、表情が変化する。ランダムな方向に音符パーティクルが出現する。

## 4. 展示方法

### 4.1 MFT24 でのブース展示について

MFT24 では、以下条件でブース展示を行った。

日時：計 2 日間展示を行った。

<1 日目>2024 年 9 月 21 日 12:00~18:00

<2 日目>2024 年 9 月 22 日 10:00~17:00

ブース構成：3.1 にて示した体験装置を 2 セットブースに設置した。その様子を図 10 に示す。なお、1 日目の実施にて、体験者の児童の手が小さい場合にトラッキング精度が低かったため、2 日目はカメラの設置方法を変更した。1 日目はモニタ上部の装飾パネルにカメラを組み込んで設置、2 日目はモニタ中央部の高さの三脚にカメラを設置した。以下図 11 に 1 日目のカメラ設置の状態、図 12 に 2 日目のカメラ設置の状態を示す。



図 10 展示ブースの概観：大型モニタ 2 枚、ペンタブレット 1 台からなる体験装置を 2 セット設置。カメラは 2 日目の配置になっている。



図 11 (左図) 1 日目のカメラ設置場所：モニタ上部に設置  
図 12 (右図) 2 日目のカメラ設置場所：カメラを三脚に取り付け、モニタ中央部に設置

### 4.2 データ回収について

1 節にて記載した仮説 3 点の考察材料として、以下 2 つの方法でデータ回収を行った。

#### 4.2.1 体験者へのアンケート

ブース展示時に図 13 のようなアンケートボードを用いて実施し、体験者に以下 2 つの質問を行った。

Q1 「どのくらい楽しかったですか」

Q2 「どこが楽しかったですか」

Q1 は「とても楽しかった」「まあ楽しかった」「ふつう」「イマイチ」の 4 段階で、該当する評価の欄にシールを貼ってもらうことで回答を得た。Q2 は付箋紙へ自由記述をする形式で回答を得た。あらかじめ体験の流れを「絵をかく

ところ」「かいた絵に顔がつくところ」「かいた絵にあわせて音がつくところ」「らくガッキーが動いているところ」「体を動かして音を鳴らすところ」といった 5 つに分けて記載し、それを自由に見ながら記入した後、該当箇所に付箋を貼る回答形式とした。

本研究では Q2 で得られた自由記述を基に、1 節で挙げた仮説の内容が子供を中心とした体験者の楽しさに貢献したかを考察する。



図 13 アンケートボード：上部に Q1 回答部、下部に Q2 回答部を配置

### 4.2.2 入力図形データの回収

本制作では、展示会で入力された図形データを記録するため、Unity アプリケーションに Excel 操作を可能にするライブラリ NPOI[11]を実装した。体験を妨げないよう配慮し、(1)入力図形の幅、(2)入力図形の高さ、(3)入力図形の曲率、(4)入力図形の複雑度、(5)選出された音色名、(6)入力図形のビジュアル、(7)(6)が楽器オブジェクトに変化した際のビジュアルといった 7 項目を Excel 形式で記録できる仕組みを構築した。

本研究では、(3)(4)(5)の結果から、ユーザが入力する形の特徴量の傾向について分析することで、3.2.3 の音色選定フローにて、それぞれの尺度イ、ロの判定が適切なものであるかを考察する。

## 5. 結果

### 5.1 体験者へのアンケート結果

#### 5.1.1 Q1 結果

Q1 の回答数は 204 件であった。結果を図 14 に示す。

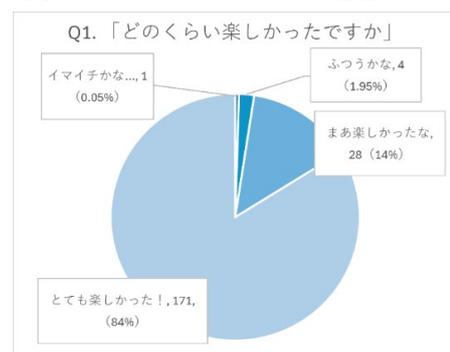


図 14 Q1 「どのくらい楽しかったですか」 回答結果

### 5.1.2 Q2 結果

Q2 の回答数は 188 件（うち記述ナシ 1 件）であった。回答内容のうち一部を体験の流れごとに分け、以下に記す。

#### 【絵をかくところ】65 件

「自分でかいたものが楽器になるところがおもしろいと思った。」「I like it because it can train us to bw more creative...」「自分でいろんな形をつくれて楽しかった。」

#### 【かいた絵に顔がつくところ】19 件

「身体動かせるのが面白かったです!」「書いた絵をベースにカワイイ顔がついたのが意外でオドロいた!」「かいたえがいきものみたいになるところ」

#### 【かいた絵にあわせて音がつくところ】29 件

「他の人のも演奏できてハッピー♪」「形によって音が違うので、色々ためたくなる」「形によって、それぞれの音がつくのが面白い」「絵に合わせて音がなるというアイデアがおもしろかった!」

#### 【らくガッキーが動いているところ】12 件

「自分の書いた絵が画面で動いて音がつく所が良かった!」「自分の書く絵が実際に画面に出て遊ぶことができる」「自分が使った絵が動くのがすごい!!楽しかった!」

#### 【体を動かして音を鳴らすところ】50 件

「動きに合わせて音楽が鳴るところ♪」「自分の動きに合わせて音が鳴るところ」「じぶんのらくガッキーがおした手にはんのうしてくれてたのしかった!」「おとがいっぱいあって楽しかった」

#### 【その他】(欄外貼付) 13 件

「絵を描くとか身体を動かすとかかんたんに音楽を楽しめるのが良かった!」「いろんなかたちがあったのしかった」

### 5.2 入力図形データの回収結果

入力図形データは 367 件回収した。この数にはスタッフの入力データを含まず、各データに日程（いつ描かれたかを示す情報）を新たに追加した。スタッフのデータ選別や日程の判別は、著者らの記憶と開始/終了時に撮影した写真記録を基に行った。以下図 15 に、Excel ファイルの実例を示す。

回収したデータ内「曲率」「複雑度」それぞれの統計量を、以下表に示す。サンプル数は 360 件であるが、装置の特性を十分に把握し、音色選定フローの評価を行う上で妥当であると考えられる。なお、複雑度が 8 以上のデータ 7 件は外れ値として除外している。

表 1 入力図形データの統計量 (n=360)

	曲率	複雑度
平均値	0.60	0.94
中央値	0.63	0.45
標準偏差	0.24	1.28
範囲	0.99	7.65

項目(3)(4)のデータより入力図形の曲率と複雑度の分布を図 16、なお、図内の外れ値のデータについては、図 17 に示す。項目(5)のデータより音色ごとの選出回数を集計結果を図 18 に示す。

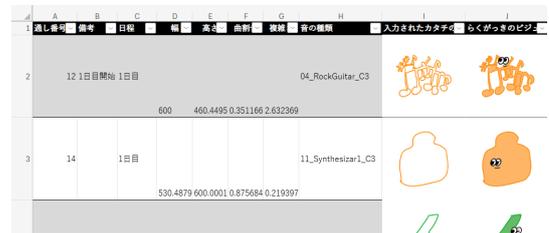


図 15 入力図形データを回収した Excel ファイル

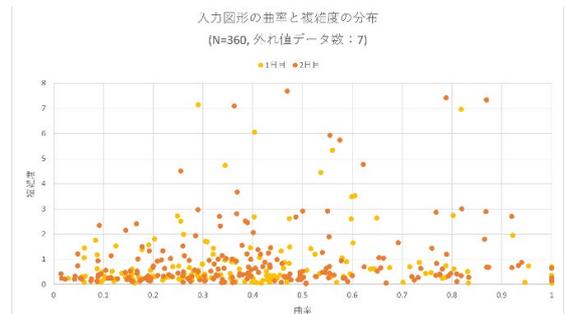


図 16 入力図形の曲率と複雑度の分布

曲率	複雑度	入力図形のビジュアル	曲率	複雑度	入力図形のビジュアル	曲率	複雑度	入力図形のビジュアル
0.09421328	39.11452		0.1703703	17.93274		0.04054798	15.71594	
0.1994248	19.79177		0.5904337	18.2097				
0.1810074	19.22279		0.1448712	15.9822				

図 17 複雑度において外れ値のデータ

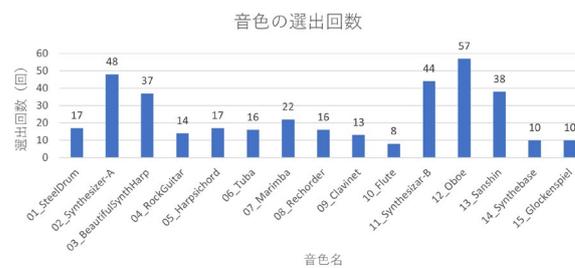


図 18 音色の選出回数

## 6. 考察

### 6.1 子供を中心とした体験者の楽しさに対する貢献点

5.1.2 に記した Q2（自由記述アンケート）の結果より、本制作の子供を中心とした体験者の楽しさに対する貢献点を、1 節で挙げた仮説 3 点に基づいて考察する。

仮説 1「視覚情報を基にして音を対応付けること：体験者が描画した図形形状を基に音を対応付けること」に対して、下記のような回答内容から、体験者の楽しさに貢献したことが示唆される。

「自分でかいたものが楽器になるところがおもしろいと思った。」「形によって、それぞれの音がつくのが面白い」「絵

に合わせて音がなるというアイデアがおもしろかった！」  
 「自分の書いた絵が画面で動いて音がつく所が良かった！！」

仮説2「身体を使った自由度の高い演奏方法」に対して、  
 下記のような回答内容から、体験者の楽しさに貢献したことが示唆される。

「身体動かせるのが面白かったです！」「動きに合わせて音楽が鳴るところ♪」「自分の動きに合わせて音が鳴るところ」

仮説3「他の体験者の絵にどのような音が対応付けられているか確認できること:結果の多様性が確認できること」  
 に対して、下記のような回答内容から、体験者の楽しさに貢献したことが示唆される。

「他の人の演奏できてハッピー♪」「形によって音が違うので、色々ためしたくなる」「いろんなかたちがあったのしかった」

以上より、本制作には仮説3点の内容が認められ、これらの要素が子供を中心とした体験者の楽しさに貢献したことが示唆される。

## 6.2 音色選定フローの評価

5.2の結果から、音色選定フローの評価について考察する。図19は、図16に示した分布の傾向を図4と照らし合わせて見るべく調整した散布図である。具体的には、x軸「曲率」を逆転させ、特にプロットの密集が確認された複雑度が2以下の範囲に調整した。また、図20は、音色ごとの選出回数と閾値を図4に加筆したものである。

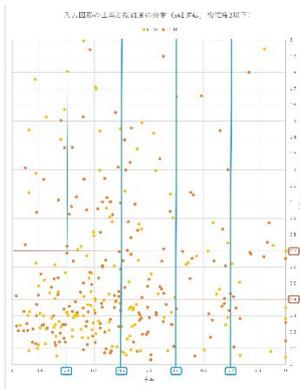


図19 (左図) 入力図形の曲率と複雑度の分布 (x軸逆転, 複雑度2以下)

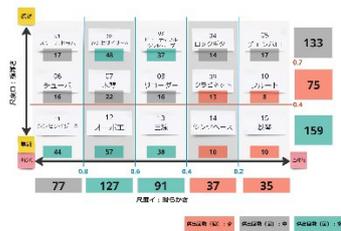


図20 (右図) 音色の選出回数と閾値を加筆した図形形状と音色の対応図

この図から、尺度イの判定フローでは、曲率が0.4以上0.6未満のレベル3区間、0.6以上0.8未満のレベル4区間で得られる音色の選出回数が多く、0.2未満のレベル1区間、0.2以上0.4未満のレベル2区間では少ないため、レベル1,2区間の閾値範囲を数値が大きい方に拡大し、レベル4区間の上限値をそのままにレベル3,4区間の閾値範囲を縮小すると、音色の選出回数の偏りが解消されることが示唆される。

また、尺度ロの判定フローでは、複雑度が0.4未満のレベル1区間で得られる音色の選出回数も多く、0.4以上0.7未満のレベル2区間では少ないため、レベル2区間の閾値範囲を数値が低い方に拡大すると、音色の選出回数の偏りが解消されることが示唆される。

## 7. 今後の展望

今後の展望として、まずは5.1.2の自由記述アンケートの結果に対し質的分析を行い、子供たちにとって「楽しさ」につながる要素を探ることが挙げられる。

続いて、6.2にて得られた考察に加え、入力図形の図形形状の傾向を分析し、計算方法や閾値を改善することで、より多様な楽器オブジェクトの出力を可能にすることが挙げられる。

**備考** 本研究で開発した内容は、特許出願を行っている。

**謝辞** 本研究の制作物の開発・実装に関し、共同制作者である株式会社日南のプロジェクトメンバーの皆様には、短い期間にも関わらず、新しい体験を実現するべくご協力いただいたことに厚く御礼申し上げます。また、MFT24にて体験して下さった皆様にも感謝申し上げます。

## 参考文献

- [1] “描いた絵が“楽器”になり、指の動きで音を出すことができる「らくガッキ」”. <https://design.jvckenwood.com/event/2024/kodomodesign/index.html>, (参照 2024-12-16)
- [2] “Maker Faire Tokyo 2024 | Make: Japan”. <https://makezine.jp/event/mft2024/>, (参照 2024-12-20).
- [3] Wu, TY., Lee, A. The Study of Sound and Shape Effects on Design. AHFE 2017, 2018, Advances in Intelligent Systems and Computing, vol 586, p.123-130.
- [4] 西田沙織, et al. 視聴覚統合における刺激音の時間-周波数的特徴と視覚刺激の形状特徴の検討-人間の知覚特性に基づく音楽の可視化システムに向けて. 研究報告音楽情報科学 (MUS) 2009.13 (2009-MUS-79) (2009), p.65-70.
- [5] Margiotoudi, K., Pulvermüller, F. Action sound–shape congruencies explain sound symbolism. Sci Rep 10, 12706 (2020). <https://doi.org/10.1038/s41598-020-69528-4>, (参照 2024-12-18)
- [6] 林田航,望月茂徳. 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. 情報処理, 2023, インタラクション 2023 論文集, 2023, p.348-352.
- [7] “Unity”. <https://unity.com/ja>, (参照 2024-12-16).
- [8] “カトカトーン”. <https://katokatone.com/>, (参照 2024-12-16).
- [9] “Garage Band”. <https://www.apple.com/jp/ios/garageband/>, (参照 2024-12-16).
- [10] “MediaPipe”. <https://github.com/google-ai-edge/mediapipe>, (参照 2024-12-16).
- [11] “NPOI”. <https://www.nuget.org/packages/npoi/>, (参照 2024-12-18)