

筑波大学 情報学群 情報メディア創成学類

卒業研究論文

実物体と音を関連付けた
音楽制作システムの構築

渡辺 絵美子

指導教員 三末 和男 志築 文太郎 田中 二郎

2011年2月

概要

音楽を聴くことは、誰でも気軽に楽しむことができる。しかし、音楽を制作することに対しては、作曲に関する知識が必要であったり、まずどこから始めればよいかわからないというような敷居の高い印象が存在する。そこで、本研究ではこの敷居の高さを軽減し、音楽に素養のない場合であっても手軽に音楽制作を楽しむことのできるシステムの構築を目標とする。

本研究で開発したシステムは、テーブル上で好きな実物体を自由に並べることによって音楽を制作していくというものである。このシステムでは、テーブル上の自由な実物体を直接手で触れることで、制作している音楽の全体像の把握や細部の修正が容易になるインターフェースを採用する。また、実物体と音の関連付けに人間の感性を利用し、関連付けを理解しやすくするとともに制作の新たな刺激となるものを提案する。システムが生成する音楽については、聴こえが良くなるように音程の調整を行い、音楽制作の支援を目指す。

このようなシステムの設計と実装を行い、実際にシステムを使用して行った評価により、操作手法の有効性を示し、物と音の関連付けに関する有効性の考察を行った。

目次

第1章	はじめに	1
1.1	音楽制作の難しさ	1
1.2	本研究の目的	1
1.3	本論文の構成	1
第2章	開発する音楽制作システムへの要求	2
2.1	専門知識を必要としないこと	2
2.2	音楽の制作を支援すること	2
2.3	楽しさを提供すること	2
第3章	実物体を利用する音楽制作システム	3
3.1	設計方針	3
3.2	インタフェースと操作方法	3
3.2.1	テーブルの設計	4
3.2.2	システム使用の流れ	5
3.3	音楽らしさを得るための設計	5
3.3.1	使用する音階	5
3.3.2	和音に関する音程	6
3.4	音の適切なマッピング	6
3.4.1	マッピングの方針	6
3.4.2	実物体と音を持つ属性	7
3.4.3	色調に対する表現語	7
3.4.4	音色に対する表現語	8
3.4.5	色調と音色のマッピング	8
第4章	音楽制作システムの実装	9
4.1	システムの構成	9
4.2	使用する音源	10
4.3	カメラによる画像の取得	10
4.3.1	テーブル全体の画像の取得	10
4.3.2	塊検出	12
4.4	実物体と音のマッピング	12

4.4.1	位置	12
4.4.2	大きさと長さ	13
4.5	タイムライン処理	14
4.5.1	タイムライン表示に伴う画像の処理	15
4.6	色と音色のマッピング	15
4.6.1	色相と音色のマッピング	15
4.6.2	色調と音色のマッピング	16
4.7	「音楽らしさ」を得るための調整	17
4.7.1	使用する音階の制限	17
4.7.2	和音に関する音程調整	18
第5章	ケーススタディ	20
5.1	結果	21
5.1.1	システムの使用例	21
5.1.2	アンケート調査結果	21
5.1.3	得られたコメント	23
5.2	考察	24
5.2.1	物と音の関連付けについて	24
第6章	関連研究	25
6.1	視覚的音楽インタフェース	25
6.2	タンジブルテーブルトップインタフェースを用いた作曲システム	25
6.3	色彩と音楽を対応付けたシステム	25
第7章	まとめ	26
	謝辞	27
	参考文献	28

目次

3.1	構築するシステムのイメージ	4
3.2	テーブル上での音の高さ	4
3.3	テーブル上で演奏される順	5
3.4	自然的長音階	6
3.5	自然的短音階	6
4.1	システムの構成	9
4.2	カメラ画像	10
4.3	背景画像	11
4.4	2値化画像	11
4.5	色情報の扱い	12
4.6	テーブルと座標の対応	13
4.7	長い物体への対応	13
4.8	タイムライン処理	14
4.9	画像を書き換える部分	15
4.10	HSV 色相環	16
4.11	音の鳴らし方	19
5.1	評価で使用した実物体	20
5.2	絵を描いていた例	21
5.3	曲線を描いていた例	22
5.4	塊を作っていた例	22
5.5	調査結果	22

第1章 はじめに

1.1 音楽制作の難しさ

音楽を楽しむということは、一般的な趣味として広く受け入れられており、音楽を聴くだけであれば特別な知識や経験を必要としない。

しかし、音楽を制作する場合はその知識や経験が必要となることが多く、初心者にはどこから始めればいいのかわからない。また、楽器の演奏経験があっても、具体的にどのような組み合わせでいけばよいかかわからないということが考えられる。このように、音楽制作に対して敷居の高い印象が存在する。

1.2 本研究の目的

音楽に素養のない人であっても、わかりやすく手軽に音楽制作を楽しむことのできるシステムを構築することを目的とし、本研究では、音楽の構成要素となるひとつのフレーズをシステムによって生成する。

ここで言うフレーズとは、並べることである程度のメロディとなり得るような単音の集合を指しており、本論文では、単音を並べてフレーズを作っていくことを音楽制作と定義する。

1.3 本論文の構成

本論文では、まず第2章で、目的を達成するためにシステムが要求される要素を述べる。第3章でシステムの設計や実装のアイデアを述べ、第4章でこのシステムの実装について説明する。第5章でケーススタディの内容と結果について考察し、第6章で関連研究を挙げる。最後に第7章でまとめる。

第2章 開発する音楽制作システムへの要求

第1章で挙げた音楽制作の難しさや敷居の高さを踏まえ、開発するシステムに求められることを整理すると、以下の3項目が挙げられる。

2.1 専門知識を必要としないこと

初めに述べたように、音楽を制作することに対して敷居の高い印象が見受けられる。楽譜を書いて制作したり、楽器で旋律を奏でて制作したりというような、特定の知識を用いた音楽の制作方法ではなく、音楽に素養のない者でも可能な、特別な知識を必要としないインタフェースと操作方法によって制作できることが重要であると考えられる。先行研究としては、自由に操作できる手のような、人間の自然性を重視したアプローチを用いた音楽知育メディアの試みが行われている [1]。

2.2 音楽の制作を支援すること

ただ音を並べただけでは、不協和音の発生や音の不自然な流れにより、「音楽らしさ」が得にくくなり、音楽とは感じにくい。このことも敷居の高さの要因として考えられる。そのため、特別な知識を持たずに音を並べた状態であっても、「音楽らしさ」を感じられるものを生成したい。このような、音楽に素養がなくても手軽に音楽が制作できることの実現が求められる。

2.3 楽しさを提供すること

楽しめることは、音楽制作のきっかけにもなる重要な要素である。前述した「音楽らしさ」を得られることは、システムを用いて音楽を制作し、これが奏でられる楽しさに繋がると考えられる。また、考えることも楽しさのひとつであると考えられる。物をどのように並べるとどのような音楽が出来上がるか、どのような物でどのような音が鳴らせるか、といったことを考え学習する楽しさも期待する。

第3章 実物体を利用する音楽制作システム

ここでは、前章で挙げた3つの要求に対しての設計方針を述べる。

3.1 設計方針

インタフェースと操作方法、音楽らしさを得るための調整、音の適切なマッピングを設計方針として考える。前章で挙げたそれぞれの要求に対する設計方針は次のようになる。

専門知識を必要としないことについては、特殊な操作をする必要がないという点、また、音楽制作に関する専門知識も必要としないという点で、インタフェースと操作方法や、音楽らしさを得るための設計について考える。

音楽の制作を支援することについては、生成する音楽の響きを向上させるためと、物と音の結び付きをわかりやすくするために、音楽らしさを得るための設計と、音の適切なマッピングについて考える。

楽しさを提供することについては、システムを楽しく使用できるようなインタフェースと操作方法や、物と音の結び付きを考える学習効果から得られる楽しさに繋がるような音の適切なマッピングについて考える。

3.2 インタフェースと操作方法

タンジブルテーブルトップインタフェース(図3.1)を採用する。このテーブル上において、実物体(オブジェクト)の置かれた位置を音の高さや鳴るタイミングの情報として利用することができる。また、「オブジェクトが置いてある=その位置に音が存在している」ということになるため、音の位置関係や全体の流れが目で見えて把握できる。

このインタフェースにおいては、それぞれの音を表すオブジェクトとなる実物体をテーブル上に置き、直接手で触れて操作する。音の追加はテーブル上に新しいオブジェクトを置くことで行える。音を修正するためには手でオブジェクトをずらしたり、交換したりすればよい。音を削除したい場合も、オブジェクトをそのまま取り除けばよいので、操作が容易になるといえる。

また、直接手で触れる操作は、使用者が自分の思うように動かせる自然な操作方法であるため、操作に関するストレスもなく楽しんで使用できると考える。

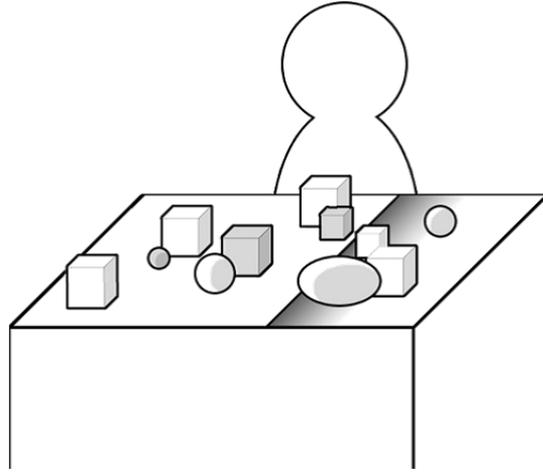


図 3.1: 構築するシステムのイメージ

3.2.1 テーブルの設計

テーブルの奥行きが音高に対応する(図 3.2)。オブジェクトが奥にあるほど音高が高く、手前にあるほど低くなる。一般的な楽譜においては、下から上へ音高が高くなっていく。このことを考慮し、このシステムでは、テーブルの手前側を楽譜の下部、低音に対応付ける。

演奏の際は、一番左に置いてあるオブジェクトから、左から右の方向へ順に音が鳴っていく(図 3.3)。これは、楽譜や横書きの文章は一般に左から右の方向へ読まれることを考慮したものである。

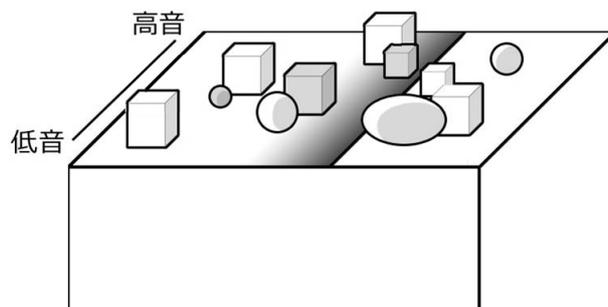


図 3.2: テーブル上での音の高さ

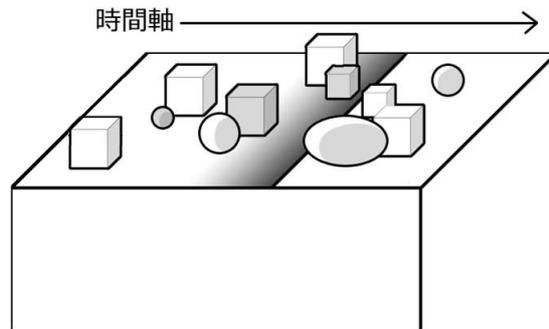


図 3.3: テーブル上で演奏される順

3.2.2 システム使用の流れ

使用者はテーブル上に実物体を並べていく。この実物体ひとつひとつが音を表すオブジェクトとなる。システムはオブジェクトの情報を取得し、この情報に応じて鳴らす音を決定する。

テーブルには左から右へ流れるタイムラインが表示されており、オブジェクトとタイムラインが交差したとき、そのオブジェクトに対応する音を鳴らす。タイムラインはテーブルの右端まで到達すると、再び左端から右方向へ流れることで鳴らす音をループさせる。

3.3 音楽らしさを得るための設計

システムではテーブル上に自由にオブジェクトを置くことが可能だが、その位置をそのまま音の情報に利用すると、音の位置関係や重なりを考慮していないため、響きの良いものを生成しにくい。そこで、音の位置を修正したり、使用する音を制限したりすることによって、「音楽らしさ」を得ることを考える。

3.3.1 使用する音階

1 オクターブ中の 12 音すべてを使用すると、音を適当に配置した場合に響きが不自然なものになりやすい。そこで、1 オクターブ中で使用する音の数を減らすことを考える。

人が暖かい・熱いと感じる色のまとまりを暖色系、寒い・冷たいと感じる色のまとまりを寒色系という。ここから、暖色系は明るいイメージ、寒色系は暗いイメージも持つと考える。

一般に、長音階は明るい・楽しい雰囲気の特徴として持ち、短音階は暗い・悲しい雰囲気の特徴として持つと言われる。ここで、暖色系が多い場合は自然的長音階、寒色系が多い場合は自然的短音階を使用し、実際に鳴らす音の高さをそれぞれの音階における範囲に制限する。また、どちらにも属しない中間色が最も多かった場合は、今回は自然的短音階を使用した。

図 3.4 と図 3.5 に、今回使用した音階を示す。



図 3.4: 自然的長音階



図 3.5: 自然的短音階

3.3.2 和音に関する音程

横方向が同じ位置で、テーブルの手前と奥にオブジェクトを並べた場合、音が重なって鳴るため和音が発生する。和音は音楽の響きに影響を与えると考えられるが、不協和音がいくつも発生したとき、「音楽らしさ」は少ないと感じてしまうだろう。このため、和音について、良い響きを得られると考えられるパターンを選択し、音程をそのパターンに合わせるようシステム側で調整する。

ある音について、和音の基盤となる根音からの音程が完全一度(音高が同じ)、完全八度(オクターブ)、完全五度、完全四度である順によく調和すると言われている。本研究ではこの関係を調整に利用する。

3.4 音の適切なマッピング

それぞれの実物体に対して人間が持つイメージを、音の音色や大きさといった属性に反映させる。例えば、柔らかい物なら柔らかい音、のような反映の仕方がある。どんな物でどんな音が出るかが想像でき、音程だけに留まらない、音楽制作に関する新たなアプローチとする。

3.4.1 マッピングの方針

物に対して用いられる表現語と、音に対して用いられる表現語について考慮する。表現語とは主に形容詞によって表されるものである。方針としては、物と音に共通して用いられる語を探し、マッピングの手がかりとする。ここでは、共通した語で表せるもの同士のマッピ

ングが適切であると考え。特に音色に関しては、「澄んだ」「柔らかい」等の共感的表現が用いられることが多い[4]ため、表現語を用いたマッピングが可能である。

本研究では、色から表現語、表現語から音色と通してマッピングを行う。

3.4.2 実物体と音が持つ属性

まず、実物体と音が共通して持つ属性として、大きさと長さを挙げる。それぞれ、大きい物・音、長い物・音、のように共通した表現語を用いて表すことができるため、これらの属性については、大きい物に対しては大きい音、長い物に対しては長い音、のようにマッピングが可能である。

3.4.3 色調に対する表現語

実物体が持つ属性として、形・色・素材などが挙げられる。今回は、この中で視覚的印象とイメージが結びつきやすいと思われる色に着目した。

ここでは、色に対して用いられる表現語について考える。PCCS¹において、色調は表3.1のように決められており、それぞれの色調に対するイメージも挙げられている(表3.2)。本研究では、各色調に対するイメージを音色のイメージと照らし合わせることにする。

表 3.1: 色調の例

	低彩度	中彩度	高彩度
高明度	Pale	Light	Bright
低明度	Dark grayish	Dark	Deep

表 3.2: 色調に対するイメージの例

Pale	軽い、淡い
Light	澄んだ、爽やかな
Bright	明るい、華やかな
Deep	深い、伝統的な
Dark	暗い、丈夫な
Dark Grayish	重い、固い

¹日本色研配色体系：Practical Color Co-ordinate System(<http://www.sikiken.co.jp/pccs/>)

3.4.4 音色に対する表現語

音色を言葉で表現するにあたり、音色の3因子として「美的因子」「明るさ因子」「量的因子」の存在が知られている [3]。表 3.4 は各因子に対する表現語の例である。音色を表現する際に用いる表現語に関する研究においては、実験に使用する音色表現語を選択する際に、尺度として使用頻度が用いられている [2]。この評価により、音色表現語として使用頻度の高い表現語が存在することがわかる。表 3.3 に、選択された 50 語のうち上位 15 語を示す。

これらの表現語について考えると、音色を表す主な要素となる 3 因子は、使用頻度の多い表現語で表せることがわかる。ここから、人間が音色に対して抱くイメージは主にこれらの表現語で表せるものであると推測できる。

表 3.3: 音色表現語として使用頻度の高い 15 の表現語

1 明るい	11 豊かな
2 柔らかい	12 固い
3 きれいな	13 こもった
4 澄んだ	14 迫力のある
5 響きのある	15 伸びのある
6 暗い	
7 力強い	
8 軽い	
9 深みのある	
10 厚みのある	

表 3.4: 音色の 3 因子とその表現語

美的因子	美しい, 澄んだ, 潤いのある
明るさ因子	軽やかな, 華やかな
量的因子	響く, 迫力のある

3.4.5 色調と音色のマッピング

音色に対する表現語を色調に対するイメージと比較すると、共通して使われる単語が存在することがわかる。そのため、色調と音色をマッピングすることはある程度の自然さを得られると考える。本研究では、色調の明暗と音色の明暗を対応させる。ここでは、音色の明暗は同種の楽器における高音楽器、低音楽器という種類に置き換えることとする。

第4章 音楽制作システムの実装

4.1 システムの構成

システムは、実物体を並べるためのテーブルと、実物体を認識するためのカメラから構成される。音が鳴るタイミングの目安となるタイムラインを表示させる必要があるため、ディスプレイを水平に置き、その上にアクリル板を載せたものをテーブルとして使用した。次に、テーブル全体がカメラの撮影範囲に収まる高さにカメラを固定し、テーブルを真上から撮影した。

カメラで撮影した画像を PC で処理し、PC から音を鳴らす。テーブルはディスプレイとして PC と繋がっており、テーブルの背景とタイムラインを出力する。



図 4.1: システムの構成

カメラは Logicool HD Pro Webcam C910、ディスプレイは DELL E248WFPs を使用した。24 型のディスプレイの画面をカメラで撮影するために、ディスプレイ表面から約 40cm 離れた高さにカメラを固定した。

4.2 使用する音源

システムの実装はJavaで行った。Java Sound API¹のMIDI機能を使用して音を鳴らす。MIDIでは、音高 (pitch)、音量 (velocity)、音色 (program) を数値で入力することで鳴らす音を決めることができる。これらの入力情報はシステムによって各オブジェクトに応じて決定される。1チャンネルに1音ずつ、最大16チャンネルまで鳴らすことが可能である。

今回は、デフォルトのJava Sound Synthesizerというソフトウェア音源を用いた。インタフェースはSynthesizerを使用し、ノートのオン・オフのみを用いて演奏する。

4.3 カメラによる画像の取得

テーブル全体の様子を認識するために、まずカメラを使用しテーブル全体を画像として取得する。次に、テーブルに置かれた実物体を認識するために、この画像に対して処理を行う。そして、処理した画像に塊検出を行うことで、テーブル上に存在する実物体の情報を得る。

4.3.1 テーブル全体の画像の取得

カメラを使用して、テーブル全体を1枚の画像として取得する(図4.2)。ここでのカメラの解像度は320x240に設定した。

また、予め背景となる画像(図4.3)を保存しておく。カメラから取得した画像と背景画像の差分を取り、背景画像と異なる部分を白、背景画像と一致する部分を黒として2値化した画像を作成する(図4.4)。差分を取る際は、両方の画像をグレースケール化し、各ピクセルの値を0-255に変換してから処理を行った。

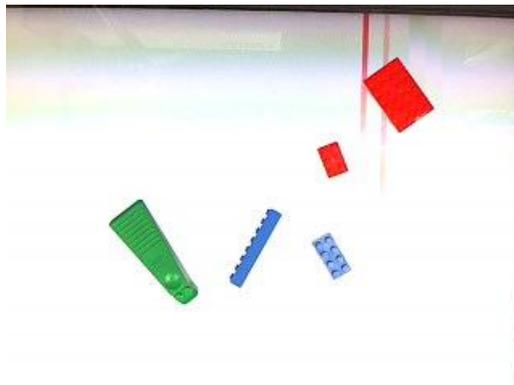


図 4.2: カメラ画像

¹<http://www.oracle.com/technetwork/java/index-139508.html>



图 4.3: 背景画像

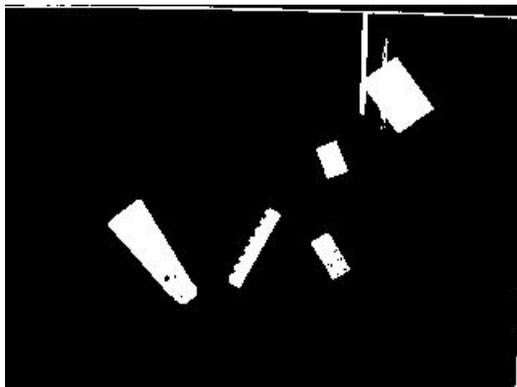


图 4.4: 2 值化画像

4.3.2 塊検出

実物体それぞれの情報を取得して利用するために、画像から塊を検出し、この塊を実物体として扱う。

2 値化画像 (図 4.4) について、BlobDetection library²を使用し、塊検出を行う。塊の情報として取得できるもののうち、塊の中心座標、最大・最小座標、面積の値を利用する。これらの値は全てカメラから取得した画像の範囲内に収まるものである。また、塊の中心座標を図 4.2 のようなカラー画像の座標と照らし合わせ、塊の中心座標の位置にある色をその塊の色と考え、その塊が示す実物体の色として扱う。

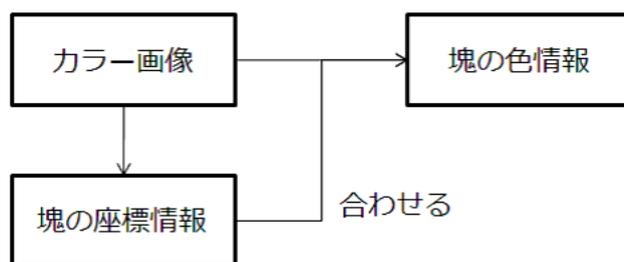


図 4.5: 色情報の扱い

4.4 実物体と音のマッピング

4.4.1 位置

テーブル全体に対応する音階は、ピッチ 48-72、2 オクターブ、24 音とする。テーブルを垂直方向に 24 分割し、テーブルの最も手前の領域をピッチ 48、最も奥の領域をピッチ 72 と設定する。

テーブルと画像の対応付けは、テーブルの手前を画像下部、奥を画像上部とする。画像も縦方向に 24 分割しておき、例えばピッチ 48 の領域は画像最下部の領域とする。

画像内における塊の中心座標のうち、テーブル上での左右の位置にあたる x 座標の値 (0-320) はそのまま使用する。テーブルの奥行きにあたる y 座標の値 (0-240) について、24 分割された領域のうちどのピッチ領域に属するかを求める。オブジェクトが手前にあるほど音が低く、奥にあるほど音が高くなるという設計に基づき、y 座標の値が 240 のときピッチ 48、0 のときピッチ 72 のように対応させる。ここで求めたピッチを、そのオブジェクトの音高として音情報の入力に使用する。

²<http://v3ga.net/processing/BlobDetection/>

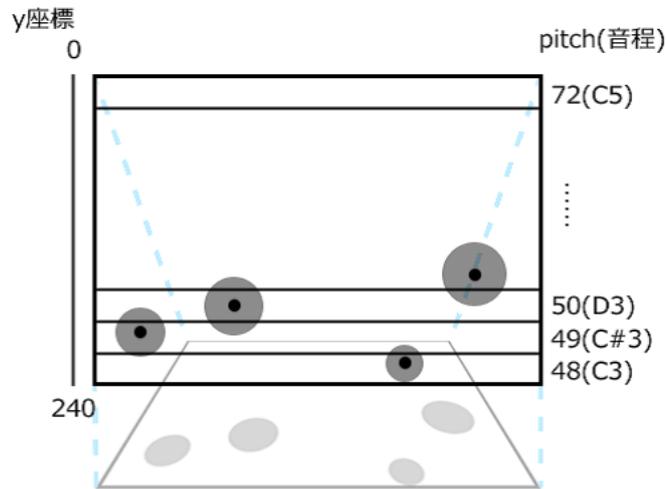


図 4.6: テーブルと座標の対応

4.4.2 大きさと長さ

塊の面積を音の音量に反映させる。ここでは閾値を用いて、面積の大きさに応じて音量を n 段階に分ける。今回の実装においては3段階の音量とし、面積が1000(実際の大きさは約 $5\text{cm} \times 3\text{cm}$ 程度) 以上の場合は最大音量である127に設定した。また、面積が100以下で小さい場合は音量を50に設定し、それ以外の場合は90に設定した。

次に、塊の横方向の大きさを長さとして利用する。今回の実装においては長音を使用しなかったため、ある長さを超えた物体については左端、中央、右端のそれぞれの位置について音を鳴らす設定とした。

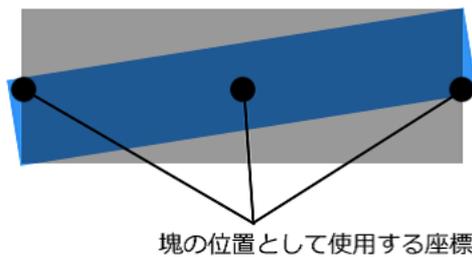


図 4.7: 長い物体への対応

4.5 タイムライン処理

画面を横方向に n 分割する。今回はフレームの幅を 10px とし、カメラ画像の幅が 320px であったため 32 分割となった。

現在のフレーム位置を保存しておく。30 ミリ秒ごとにフレーム位置は右へ移動する。フレーム位置が移動すると、まず塊検出を行い、検出された塊を順にリストに入れて保存しておく。このリストを順に見ていき、現在のフレーム位置に塊が存在したとき、すなわち塊の中心座標がフレームの範囲内にあった場合に音を鳴らす処理を行う。このフレームで鳴る音の出力が終わると、フレーム位置は右に移動する。

フレーム位置が右端まで到達した場合は、次のフレーム位置を左端に戻して同様に処理を繰り返す。タイムラインは、現在のフレーム位置に描画する。

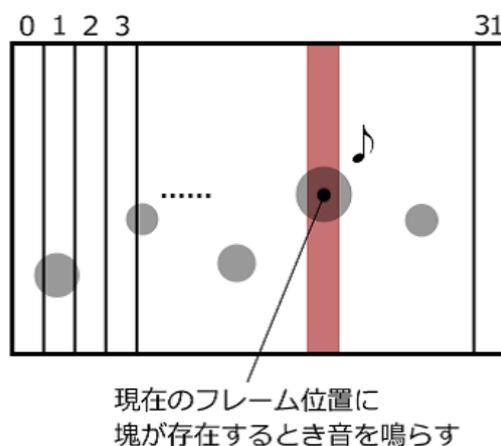
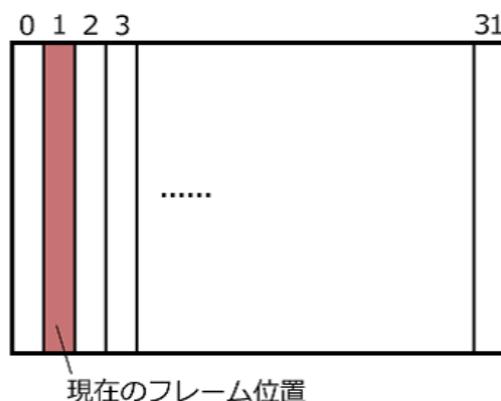


図 4.8: タイムライン処理

4.5.1 タイムライン表示に伴う画像の処理

本システムではテーブルにタイムラインが表示され、これをカメラで撮影することになるため、画像にタイムラインが写り込んでしまい、認識の妨げとなる。これを防ぐために、認識に使う画像はタイムライン部分を取り除いたものを使用する。

システムで使用する2値化画像は、50ミリ秒毎に更新されている。このときのタイムラインの表示位置から左右を少し開けた部分のみを更新し、タイムラインが存在する箇所は50ミリ秒前の画像を使用する。実物体は同じ位置にあり、タイムラインの位置が異なるのみであるため塊自体の情報は変化しないと考え、このような処理とした。

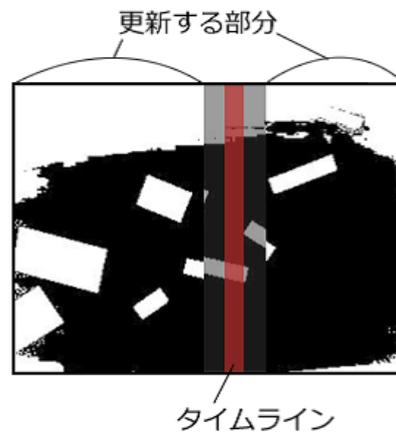


図 4.9: 画像を書き換える部分

4.6 色と音色のマッピング

音と色の対応については、過去に多くの報告がなされている。これらの報告によれば、細部までの一致は見られないが、傾向として一致する部分が見られるようである。

本研究での音と色の対応は、まず大まかな色相と大まかな楽器の音色を対応させ、次にそれぞれの色相について、色調に応じて細かい楽器の音色を決定する手法を取る。

4.6.1 色相と音色のマッピング

楽器の音色と色の対応については、「細部に関しては必ずしも一致しないが、金属楽器の高い音は赤の輝き、木管楽器は青系統の色、弦楽器は黄や緑系統の色と対応するという点では一致する」という報告がある [5]。本研究では、この報告に基づき、色相と音色のマッピングを行う。

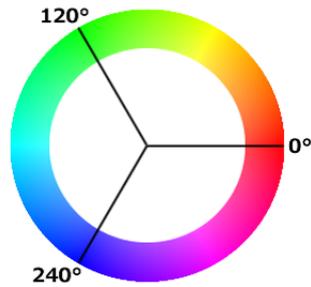


図 4.10: HSV 色相環

HSV 色空間上の色相環 (図 4.10) について、 $0^\circ - 30^\circ$ と $260^\circ - 360^\circ$ を赤、 $30^\circ - 60^\circ$ を黄、 $60^\circ - 160^\circ$ を緑、 $160^\circ - 260^\circ$ を青のグループとする。各グループについて、大まかな楽器の音色を表 4.1 のように決める。

表 4.1: 色相と音色のマッピング

色相	楽器
赤	金管楽器
青	木管楽器
黄	打弦楽器 (ピアノ)
緑	擦弦楽器 (ヴァイオリン等)

4.6.2 色調と音色のマッピング

音の高さと色の対応については、音高が上がると色の明度が上がる傾向が見られる [6]。色に関しては、派手な、動的な等の活動性の高さや彩度の高さに相関関係がある [7]。この傾向と色調に対するイメージから、色調と音色のマッピングを行う。

HSV 色空間における彩度と明度を見る。これらは $0-1$ の値をとる。ここでは、彩度について $0-0.33$ を低彩度、 $0.33-0.66$ を中彩度、 $0.66-1.0$ を高彩度とし、明度について $0-0.5$ を低明度、 $0.5-1.0$ を高明度として、全ての色をこれらを組み合わせた 6 つの色調に分ける (表 4.2)。なお、ここでの色調は第 3 章の色調に関する表 3.1 と表 3.2 を対応させたものである。

表 4.2: 6 つの色調

	低彩度	中彩度	高彩度
高明度	軽い色調	澄んだ色調	明るい色調
低明度	重い色調	暗い色調	深い色調

前述した各色のグループについて、軽いイメージの色調には高音楽器、重いイメージの色調には低音楽器というように、色調に応じて細かい楽器の音色を決める。楽器の種類は各色につき4種類とした。

表 4.3: 赤色系における楽器の対応付け

	低彩度	中彩度	高彩度
高明度	Muted Trumpet	Trumpet	Brass Ensemble
低明度	Tuba	Tuba	Trombone

表 4.4: 黄色系における楽器の対応付け

	低彩度	中彩度	高彩度
高明度	Electric Piano 1	Electric Piano 2	Electric Piano 2
低明度	Grand Piano	Bright Piano	Electric Grand Piano

表 4.5: 緑色系における楽器の対応付け

	低彩度	中彩度	高彩度
高明度	Tremolo	Viola	Violin
低明度	Contrabass	Cello	Cello

4.7 「音楽らしさ」を得るための調整

4.7.1 使用する音階の制限

カメラから得た画像について、各ピクセルを走査しどのような色が最も多いかを調べる。ここでは、HSV 色空間上の色相環について、0°–60°(赤–黄)、330°–360°(赤紫–赤)を暖色系、160°–250°(青緑–青紫)を寒色系、これ以外を中間色とした。

暖色系が多ければ自然長音階、寒色系・中間色が多ければ自然短音階を使用する。各音階において使用される音は決まっており、システムでは使用される音に対応するピッチのみを使用する。オブジェクトの位置から求めたピッチが、使用するピッチのリストに含まれている場合はそのまま入力に使用する。求めたピッチが使用するピッチのリストに含まれていなかった場合は、ピッチの値を下げて使用するピッチにあてはめる。表 4.7 に、例として自然長音階の場合に使用するピッチのリストを示す。

表 4.6: 青色系における楽器の対応付け

	低彩度	中彩度	高彩度
高明度	Piccolo	Flute	Clarinet
低明度	Baritone Sax	Alto Sax	Soprano Sax

表 4.7: 自然長音階の場合

ピッチ	音高	ピッチ	音高
72	C5	59	B3
71	B4	57	A3
69	A4	55	G3
67	G4	53	F3
65	F4	52	E3
64	E4	50	D3
62	D4	48	C3
60	C4		

4.7.2 和音に関する音程調整

現在鳴らそうとしている音に対応する塊と、ひとつ前に鳴った音に対応する塊について、その座標位置を比較する。この2つの塊について、x座標がほぼ同じ位置であった場合について考える。人間が物を同じ位置に置いたと思っても、システム側で物の位置とみなす座標はズレている場合が多いと考えられるため、この場合のx座標の幅には少し余裕を持たせておく。

x座標がほぼ同じ位置であるとき、y座標が離れていた場合は和音の関係であると見なす。1つのチャンネルで同時に鳴らせる音は一音なので、和音を作るために現在の音はひとつ前の音とは異なるチャンネルで鳴らすよう設定する。次に、y座標もほぼ同じ位置であった場合は、同じ物体を指していると思なし、鳴らさないこととする。

和音の関係であると見なした場合は、現在の音の音程調整を行う。ひとつ前の音のピッチと現在の音のピッチを比較し、完全五度もしくは完全四度の関係にあった場合はそのまま鳴らす。これら以外の場合は、現在の音をひとつ前の音のオクターブに設定して鳴らす。

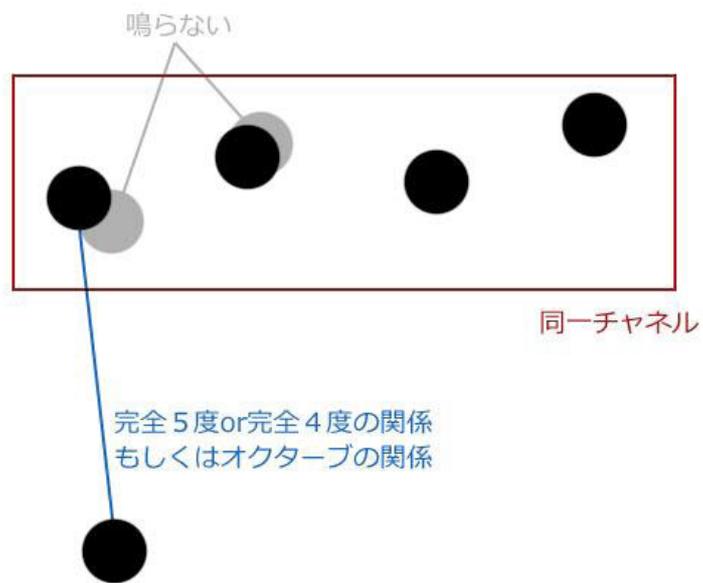


図 4.11: 音の鳴らし方

第5章 ケーススタディ

開発したシステムの有効性を評価するため、評価実験を行った。今回は音楽経験のある大学生3名と音楽経験のない大学生1名を被験者とした。ここでの音楽経験は、学校の授業を除いた、部活や習い事等による経験を指す。

被験者には自由にシステムを使ってもらい、どのように使用したかの観察を行った。ここで使用する実物体には、6色のブロック等の小型の玩具(図5.1)を用いた。

ケーススタディでは、被験者に物体をどう置くべきか指定せずに、どのように物を置いていくかを見た。システムの使用を開始してからしばらくの間は、物と音のマッピングに関する説明は行わなかった。



図 5.1: 評価で使用した実物体

ケーススタディ以外に、わかりやすさや楽しさという観点でのシステムの有効性を評価することを目的として、4が最高となる4段階でのアンケート調査を行った。表5.1にアンケートの設問を示す。また、システムを使用しての感想やコメントも収集した。

表 5.1: アンケートの設問

Q1	操作方法は分かりやすかったか
Q2	操作は簡単だったか
Q3	物の位置と音の関連付けは分かりやすかったか
Q4	物の色と音の関連付けは分かりやすかったか
Q5	生成される音楽は聴こえがよいと感じたか
Q6	システムを使用して楽しいと感じたか

5.1 結果

5.1.1 システムの使用例

複数の被験者が、テーブルをキャンバスに見立て、複数色のブロックを並べて絵を描いていた(図 5.2)。絵になっている部分は物体が多く配置されるため、この部分で音が多く鳴ることによって、この絵の部分が鳴っているのだろうと推測できたようであった。

一色のブロックを右上がりの曲線上に配置する例も見られた(図 5.3)。この被験者は、並べた通りに音程が上がっていくと面白いと思って配置したと述べた。また、物体を同じ高さで横一列に並べたり、大きな塊を作ってどう鳴るかを見ている者もいた(図 5.4)。

一度物を並べてからは、その状態を基本として物を取り除いたり増やしたりして操作する場合が多かった。



図 5.2: 絵を描いていた例

5.1.2 アンケート調査結果

表 5.2 と図 5.5 に調査結果を示す。なお、被験者 D のみが音楽経験なしと答えた。

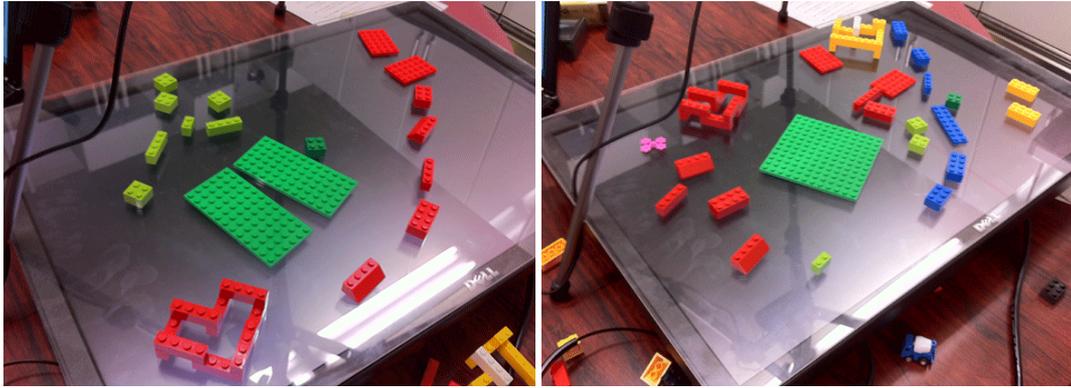


図 5.3: 曲線を描いていた例

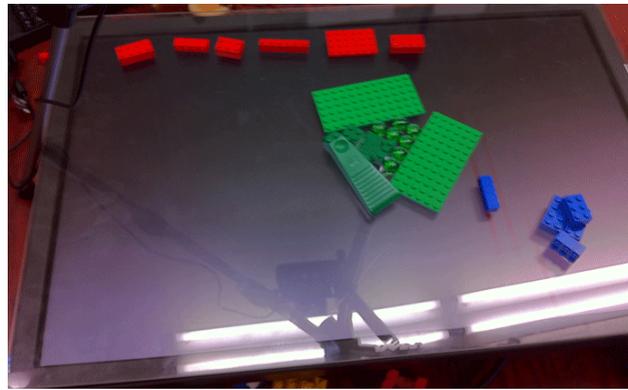


図 5.4: 塊を作っていた例

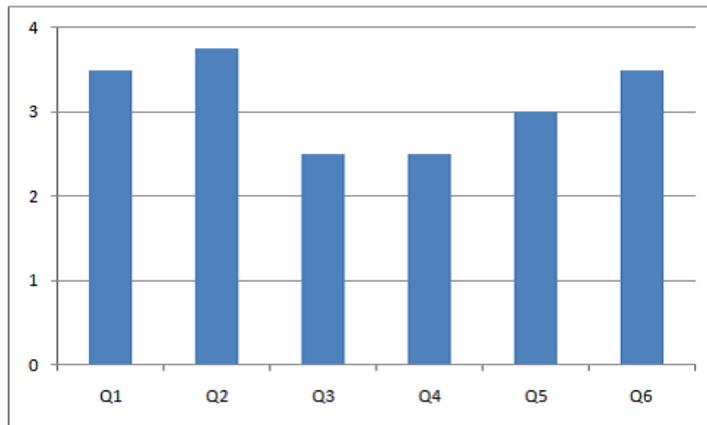


図 5.5: 調査結果

表 5.2: 調査結果

被験者	Q1	Q2	Q3	Q4	Q5	Q6
A	4	3	4	2	3	4
B	3	4	2	3	3	3
C	3	4	2	2	3	3
D	4	4	2	3	3	4
平均	3.5	3.8	2.5	2.5	3	3.5

4段階評価のうち、3以上を良い評価が得られたものとする、操作方法の理解しやすさと容易さ、システムの楽しさについては、良い評価を得られた。生成される音楽の聴こえのよさについても、やや良い評価を得られた。物と音の関連付けに関しては、個人差が見られたが、やや理解しにくかったという評価になった。また、音楽経験の有無による回答の差は見られなかった。

5.1.3 得られたコメント

システムを使用した被験者から、以下のコメントが得られた。

操作について

- 何も置いていない状態からどう物を置いたらいいか迷った
- 操作自体は分かりやす過ぎるくらいだと思う

物と音の関連付けについて

- 黒いブロックを置いたときに音が変わったのが分かった
- 音の高さの変化はあまり分からなかった
- どの物体がどの音を発しているか分かりやすいとなお良かった
- 物体を取り除いて初めてその物体がどの音を出していたかが分かった

生成される音楽とシステム全体について

- 金管系の音がないときの音楽がヒーリングミュージックのようでいい感じだった
- パーカッションが入るとノリのいい音楽も作れそう
- 狙っている音楽を作るのには向いていないと思うが、適当に置いていってどんな音楽になるか楽しむ用途には向いていると思う

5.2 考察

実際に使用する様子を観察すると、絵を描いた例が多かった。これは、自由に物を並べて音楽を作ると言われても、まずどうすればいいのか分からなかったと考えられる。ある程度物を並べてからは、これを基準にして物の配置を変更していく例がよく見られたので、最初に何をすればよいかという部分については、あらかじめ作例を示しておくなどの配慮が必要だったと考える。

アンケート Q5 の聴こえのよさは、不協和音などで変な響きになっていなかったかという部分を尋ねた。その回答からやや良い評価が得られたため、音程調整はある程度有効であったと考えられる。

音楽経験の有無は評価の結果に影響を及ぼさず、操作方法についてはアンケート Q1、Q2 の結果から良い評価を得られ、分かりやすさを目指すという点では有効だったといえる。システム全体については、楽しさを得られるかどうかについて、アンケート Q6 の結果から有効であったといえる。

5.2.1 物と音の関連付けについて

物の位置と音の関連付けについては、アンケート調査やコメントで個人差が見られた。分かりやすかったと答えた被験者は、システムを使用する際にどう鳴るかを考えてブロックを配置していた。同色のブロックを曲線状に並べたとき、テーブルの手前側にブロックが存在したときは低い音が鳴り、奥側にブロックが存在したときは高い音が鳴っており、自分の予想と一致したために分かりやすかったと感じたのだと思われる。分かりにくかったと答えた被験者も、同じ位置でブロックを置いたり取り除いたり繰り返すことで、この位置でどの音が鳴っていたかを後から理解してはいたようであった。

物の色と音の関連付けについて気にかけていた被験者は見られなかったため、有効であったとはいえない。また、色によって音色が異なることを被験者が自力で気付いたかどうかは今回の評価では分からなかった。色と音色の関連付けが妥当であったかどうかについては、色相と音色については妥当であったかの評価はできなかったが、黒のような重い色調と低音楽器の音の割り当ては判別しやすかったと思われる。

関連付けが分かりやすいかどうかについては、被験者が予想したものと合致することが重要だと考えられる。この点では、位置と音の高さのマッピングは有効であったといえる。

システムが物体を認識する精度がよくなかったために、関連付けの部分が分かりにくくなっていた部分も大きかったと思われるので、認識の精度を上げるのが課題として残った。

第6章 関連研究

6.1 視覚的音楽インタフェース

誰でも直感的、視覚的に音楽を演奏することのできるような、視覚的音楽 (visible music) インタフェースの研究が行われている。TENORI-ON[8] は、LED スイッチによって光と音が直結しており、演奏する音楽の構造を直感的、視覚的に理解しながら作曲できる視覚的音楽インタフェースである。

6.2 タンジブルテーブルトップインタフェースを用いた作曲システム

Audiopad[9] や The reacTable[10] は、テーブル上でオブジェクトを直接動かすことによって、音楽を合成、変化させる作曲システムである。本研究は、明確なフレーズを生成するために予め決められた音を出す点で異なっている。

The Music Table[11] は、AR マーカーをテーブル上に並べ、その位置や傾きに応じて音を鳴らす作曲システムである。位置によって音を決める点は本研究と同じだが、本研究では音に付加する大きさ等の属性の決定に実物体の大きさ等の情報を利用している。

Noteput¹は、音符の形をしたオブジェクトを五線譜の表示されたテーブル上に置き、楽譜の形にすることで演奏するシステムである。本研究は、五線譜をシステムから取り払い、楽譜の知識を必要としないよう設計した点で異なる。

また、どの研究もマーカーを利用しているが、本研究はマーカーの付加されていない実物体を利用することで、システムに使用する物体を限定することがない点が特徴といえる。

6.3 色彩と音楽を対応付けたシステム

色彩と音楽の共通性に関しては、古くから議論されてきており、これらに対応付けたシステムも開発されてきた。Ziplayer[12] は、音楽と絵画をそれぞれ解析し、音程に色彩を割り当てることで音楽と絵画を融合した表現手法を持つメディアプレイヤーである。

また、人間の感情をラベルとして色と音に結び付けることで、色に対応したドラムフレーズを生成するシステムも開発されている [13]。

本研究では、人間の持つ印象を、表現語を介すことで色と音を結び付けた。

¹<http://www.jonasheuer.de/index.php/noteput/>

第7章 まとめ

本論文では、初めに音楽制作が持つ敷居の高さを指摘し、開発する音楽制作システムの目的を、音楽に素養のない場合も手軽に音楽制作を楽しめることとした。次に、開発する音楽制作システムへの要求を整理し、専門知識を必要としないこと、音楽の制作を支援すること、楽しさを提供することの3つを挙げ、この要求に基づいてシステムの設計と実装を行った。

開発したシステムは、テーブル上に自由な実物体を並べていくことで音楽制作を行うものである。テーブル上で実物体を手で触れて動かすという操作方法によって操作に関する専門性を取り除き、物の位置と音の対応については、奥行きに対応した音高や右方向の時間軸のような、理解しやすいと思われる対応付けを行った。物の色と音の関連付けについては、人間の持つイメージを反映させるために、色や音色に対する表現語をマッピングの手がかりとして用いた。

評価はシステムを実際に被験者に使用してもらうことで行った。評価の結果、操作手法は分かりやすさの点で有効であり、システムによって楽しさを得られたことが分かった。色相と音色のマッピングに関しては有効だという評価は得られなかったが、色調と音色のマッピングにはある程度の有効性が見られた。

今回は実物体の色のみをマッピングに用いた。人間の持つイメージは質感による部分も大きいと思われるので、今後はテクスチャと音色のマッピングも考慮したい。また、カメラと画像処理による実物体の認識が甘く、評価に影響を及ぼしていたので、認識精度の向上も目指したいと思う。

謝辞

本論文を執筆するにあたり、指導教員である三末和男先生、志築文太郎先生、田中二郎先生、ならびに高橋伸先生には、丁寧なご指導と適切な助言をいただきました。心より感謝申し上げます。また、インタラクティブプログラミング研究室の皆様にも多くの貴重なご意見をいただきました。評価実験も快く引き受けてくださり、心より感謝致します。特に、NAIS チームの皆様には研究全般にわたり大変お世話になりました。本当にありがとうございました。

参考文献

- [1] 前川督雄, 蓼沼眞, 萩田紀博, 非西欧音楽スタイルに学ぶ音楽知育メディアの試み, 情報処理学会研究報告 音楽情報科学 2003(16), pp. 61–66, 2003.
- [2] 上田和夫, 音色の表現語に階層構造は存在するか, 日本音響学会誌 44(2), pp. 102–107, 1988.
- [3] 安倍幸司, 小澤賢司, 鈴木陽一, 曽根敏夫, 音色表現語, 感情表現語及び音情報関連語による環境音評価, 日本音響学会誌 54(5), pp. 343–350, 1998.
- [4] 難波精一郎, 音色の定義を巡って, 日本音響学会誌 49(11), pp. 823–831, 1993.
- [5] 三雲真理子, 小谷里美, 管楽器の音色の色彩的イメージ, 音楽心理学の研究, pp. 186–222, ナカニシヤ出版, 1996.
- [6] 長田典子, 岩井大輔, 津田学, 和氣早苗, 井口征士, 音と色のノンバーバルマッピング: 色聴保持者のマッピング抽出とその応用, 電子情報通信学会論文誌. A, 基礎・境界 J86-A(11), pp. 1219–1230, 2003.
- [7] 佐藤昌子, 皆川基, 吉川研一, 形状と色彩の感情効果に関する研究(第2報): その1. 単色の感情効果とその色の幾何学文様に配色した場合の感情に及ぼす色面積の影響, 日本色彩学会誌 20(2), pp. 41–55, 1996.
- [8] Yu Nishibori, Toshio Iwai, TENORI-ON, NIME '06, pp. 172–175, 2006.
- [9] James Patten, Ben Recht, Hiroshi Ishii, Audiopad: a tag-based interface for musical performance, NIME '02, pp. 1–6, 2002.
- [10] Sergi Jorda, Gunter Geiger, Marcos Alonso, Martin Kaltenbrunner, The reacTable: a tangible tabletop musical instrument and collaborative workbench, TEI '07, pp. 139–146, 2007.
- [11] 牧野真緒, Berry Rodney, 阿部明典, 樋川直人, 鈴木雅実, The Augmented Composer Project : The Music Table, 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎 103, pp. 9–14, 2003.
- [12] 福本麻子, 塚田浩二, 蔡東生, 安村通晃, Ziplayer: 絵画, 文章, 音楽を組み合わせたメディアプレイヤーの提案, インタラクション 2006 予稿集, pp. 79–81, 2006.

- [13] 金箱淳一, Color による作曲支援システムの構築, 岩手県立大学ソフトウェア情報学部卒業論文, 2006.