

## 音楽演奏における情緒構造のフローテクスチャベース可視化

末 永 和 史<sup>†</sup>      藤 代 一 成<sup>†‡</sup>

<sup>†</sup>東北大学 大学院 情報科学研究科

<sup>‡</sup>東北大学 流体科学研究所

**要旨:** 音楽の演奏は、演奏者によってテンポ、音の強弱などが異なる。この相違は、演奏中にさまざまな曲想や感情を表現するために生じるものである。このような演奏の情感表現を、音楽演奏の情緒構造とよぶことにする。本研究では、そのような情緒構造を対象とした可視化を試みる。演奏情報としてMIDI (Musical Instrument Digital Interface) データを用い、音量、音価などの音楽の論理構造を表す2次元ベクトル場を、情緒構造を各種属性に反映したLIC (Line Integral Convolution) によって可視化する。基準演奏のLICテクスチャとの相違から、相対的に対象演奏の情緒構造を顕在化できると考えられる。簡単な楽曲のMIDIデータを用いて、提案手法の有効性を検証する。

## Flow Texture-Based Emotional Visualization of Music Performances

Kazufumi SUENAGA<sup>†</sup>      Issei FUJISHIRO<sup>†‡</sup>

<sup>†</sup>Graduate School of Information Sciences, Tohoku University

<sup>‡</sup>Institute of Fluid Science, Tohoku University

**Abstract:** Music performance varies drastically according as players in terms of tempo, sound level, and so on. This difference stems from various factors such as player's interpretation of music motifs and emotional transition during the performance. In this study, we attempt to visualize such emotional structures of a performance as well as the logical structure of the underlying musical composition. After a two-dimensional vector field is derived from parameter analysis of a given MIDI (Musical Instrument Digital Interface) dataset, the field is visualized densely by using colored LIC (Line Integral Convolution). We are allowed to understand the emotional structure of the target performance by comparing its resultant texture from that for the standard performance. The effectiveness of this approach is illustrated with application to practical MIDI datasets.

### 1 背景と目的

音楽の演奏は、同一楽曲であっても、演奏者によってテンポ、音の強弱などが異なる。この相違は、演奏者が演奏の中にさまざまな曲想や感情を表現するために生じるものである。楽譜上に記された音符の音価や音程などの論理構造は演奏表現の基本であるが、上述のような情緒構造も重要であり、これが演奏のよし悪しを最終的に決定すると言っても過言ではない。

Bresinらによって、コンピュータを用いて半自動的に演奏に表情付けを行うシステムの研究が知られている [1]。6つの基本感情それぞれに対応する演奏パラメータを決定し、演奏データにそれを適用することで、表情豊かな演奏を作り出している。

音楽の表現方法としては無論聴覚的表現が一般的であるが、コンピュータグラフィックス (CG) を使って、視覚的に音楽の理解を助長する音楽可視化 (music visualization) も研究されてきた。可視化により、

ユーザは対象の楽曲を聴くことなく、一瞥するだけでその構造を理解できる。それによって、異なる演奏の比較等も容易になる。また、音楽の知識の浅いユーザを、読譜の煩わしさから解放することもできる。

BRASS[2][3]は、楽曲学習の支援を目的とし、楽譜の効率的な表示を目指したシステムである。スコア全体をフォーカス+コンテキストを用いて、単一ウィンドウ内に表示することで、楽曲全体の構造の把握を容易にしている。また、音楽データの自動表情付けのシステムとして、tutti[4]がある。音楽データのもつ音高などの詳細情報を隠すことによって、誰でも簡単に音楽の表情付けが楽しめる音楽データのブラウジング・編集システムである。

音楽演奏の可視化に関する研究もなされている。comp-i[5][6]は、演奏情報を与えるMIDIデータを、3次元空間内に円柱の直径・高さ・彩度で離散的に表現し、音楽の知識が少ないユーザにも、視覚的に

わかりやすい音楽表現を提供している。また、平賀らは聴解力 (listening comprehension) を高めるための研究 [7] を行った。楽譜上の音符と演奏された音符を比較し、シテイスケープ法を利用することで、部分的に演奏の情緒表現の可視化を試みている。

comp-i のような演奏情報の可視化は、幾何学的なオブジェクトを用いて可視化し、音楽の論理構造の可視化を目指していた。それに対し本研究では、音楽演奏の情緒構造に着目し、テクスチャをベースとした可視化を試みる。人間は、ある模様を見たときに、本来秩序のない模様は何らかの意味付けをしようと試みる。その解釈は人によってさまざまであり、一意には決まらない。音楽演奏の情緒構造を考えると、人によってその演奏から感じ取るものは異なる。時系列イベントデータとしての音楽演奏の情緒構造の可視化を行うには、フローテクスチャを用いて表現し、同様の効果を狙うのが効果的と考えられる。

本研究では、フローテクスチャ生成技法として、線積分量み込み法 (Line Integral Convolution: LIC) [8] を用いる。先行研究で開発した MIDI-LIC 変換 [9] を拡張し、音楽演奏における情緒表現を LIC テクスチャの各種属性で表現することで、ユーザに情緒的な何かを誘起させるような画像の生成を目指す。

本論文は次節以降、以下のように構成されている。次節で、音楽演奏の情緒構造を示すパラメタの抽出方法について述べる。そして、3 節で音楽演奏の情緒可視化を行う拡張 MIDI-LIC を提案する。続く 4 節で提案手法の有効性を事例を用いて検証し、最後に 5 節で本論文をまとめ、今後の課題に言及する。

## 2 音楽演奏の情緒パラメタの抽出

音楽演奏の情緒構造を表すために、さまざまな形容詞が使われる。本研究では、そのような形容詞をいくつか抽出し、それに対応する MIDI パラメタを考える。

### 2.1 情緒語の抽出・分類

演奏の情緒を表す形容詞として、Band Journal [11] より、2005 年に発行された号をすべて調査することで、136 語を抽出し、それらを意味の近接度によってグループ化した。音楽演奏の情緒構造を顕著に表す代表的な情緒語 10 個を選び出し、対となる意味をもつ語を 1 つのグループとし、最終的に 5 つのグループに帰着させた：

- **音量に関するグループ**  
静かな  $\longleftrightarrow$  力強い
- **テンポに関するグループ**  
ゆったりとした  $\longleftrightarrow$  活気のある
- **フレージングのなめらかさに関するグループ**  
はねるような  $\longleftrightarrow$  なめらかな
- **ダイナミクスに関するグループ**  
平坦な  $\longleftrightarrow$  色彩豊かな
- **明るさに関するグループ**  
暗い  $\longleftrightarrow$  明るい

### 2.2 情緒語と MIDI パラメタの対応

前項で分類した各情緒語と対応する MIDI パラメタを考える。まず、音量・テンポは MIDI データがもつ音量・テンポと対応させる。ダイナミクスは音量差を算出する。次にフレージングのなめらかさは、注目する音の直前の音が消音してから注目する音が発音されるまでの時間間隔に対応させる。これが短ければ音のつながりが良くなめらかに聞こえ、長ければ流れの悪い演奏になる。最後に演奏の明るさであるが、これは楽曲の調性に対応させる。西洋音楽において調性は全部で 24 調存在し、各調によって与える印象が異なる。調性は楽曲の雰囲気決定する重要な要素であり、長調の曲は比較的明るく、短調の曲は比較的暗いものとなる。ここで、調性は楽曲自身もつ楽曲の雰囲気を決定するパラメタであり、演奏によって変化するものではないことに注意する。

これらのような音楽演奏の情緒構造を表すパラメタは、楽曲単位、または楽曲中のフレーズ単位で決まる。このパラメタを示すレーダーチャートを PQP (Performance Quadratic Prism) とよぶことにする。PQP の例を図 1 に示す。ここで、音楽演奏の明るさに対応する調性は演奏によらないものなので、軸方向を他の 4 つとは別方向としている。

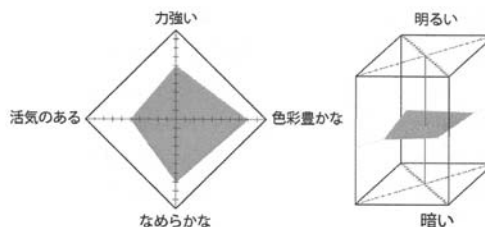


図 1: Performance Quadratic Prism

### 3 拡張 MIDI-LIC 変換

MIDIデータをLICに変換する手法は、MIDI-LIC変換 [9] を基本にする。さらに、LICの拡張表現に、音楽演奏の情緒構造を示すパラメタのマッピングを施し最終的に出力する。図2にその処理の流れを示す。まず、SMFに格納されたMIDIデータを2次元ベクトル場に変換する(ベクトルマッピング)。次に、入力MIDIデータからPQPのパラメタを算出する。そして、得られたベクトル場とホワイトノイズ画像を入力し、PQPのパラメタを反映させたLICを用いて最終的に2次元の静止画に変換する(パラメタマッピング)。

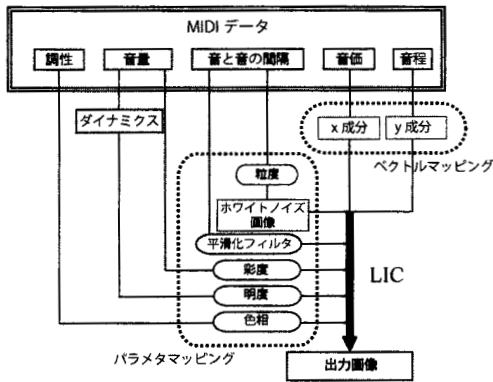


図 2: 拡張 MIDI-LIC 変換

#### 3.1 ベクトルマッピング

SMFに格納されているMIDIデータのうち、扱うデータはノートオン、ノートオフ、ピッチとする。ベクトルマッピング方法を図3に示す。(a)は出力画像の $x-y$ 平面の定義で、(b)は個々のベクトルの定義である。出力画像の $x$ 軸方向に時間を定義することで、音楽の時間に対する流れを表現する。 $y$ 軸方向には $x$ 軸上と同等のベクトルを並べる。各ベクトルの $x$ 成分に音価を定義する。 $y$ 成分には、注目しているベクトルに対応する音の一つ前の音と比較したときの音高の変化量を対応させる。各フレーズの最初の音はひとつ前の音と比較することができないため、その $y$ 成分を零とする。

#### 3.2 パラメタマッピング

2節で分類した音楽演奏の情緒語に対応する情緒構造を視覚的に表現できるような可視化手法を考える。

本研究では、代表的フローテキストチャ生成技法であるLICの拡張表現を用いることで可視化を行う。

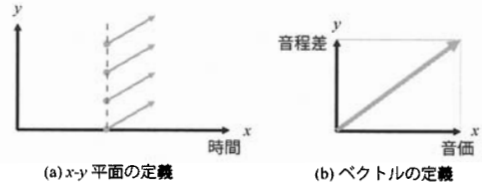


図 3: ベクトルマッピング

#### 3.2.1 LIC

LICは、1993年にCabralとLeedomによって提案された、2次元流れ場を可視化するテキストベース技法である。入力テキストの各ピクセル値を局所流線に沿ってぼかすことにより、流れ場の密な可視化を行う。この技法ではまず、2次元ベクトル場を与え、それと同様の大きさのホワイトノイズ画像を入力する。出力画像の各ピクセル値を求めるために、2次元ベクトル場の対応するセル上を通過する一定の長さの流線を構成する。それをホワイトノイズ画像上にマッピングし、この流線上に1次元のカーネル関数を定義して、流線が通過する入力画像のピクセルと畳み込みを行うことで出力画像の各ピクセル値を算出する。

LICによってベクトルの方向は表現できるが、その向きや大きさは表現できない。しかし、ホワイトノイズ画像の粒度の変化による表現 [10] やアニメーションによる拡張表現を用いることでその問題を解消する手法も知られている。音楽演奏の情緒構造を、このような拡張表現に変換することで、効果的に可視化することができる。

#### 3.2.2 各パラメタと拡張表現の対応

- **調性**: 楽曲の調性に関しては各調に対する色相を対応させる。ここで、色と音の共感覚 [12] を利用する。共感覚とは、ある一つの感覚の刺激によって別の知覚が不随意的に引き起こされる現象をいう。ここでは、色と音の共感覚をもつユーザーがある調の楽曲を聴いたときに感じる色を利用する。これにより、共感覚をもたないユーザーも似たような感覚を受けることができる。A.Skryabinの色聴表 [13] 等で試みられている色づけを統合した、調と色の関係を図4に示す。
- **音量**: 音量をLICテキストチャの色の明度に対応させる。楽曲全体での音量の平均を計算し、それに対して着目する音の音量が大きければ大き



いほど明度を大きくし、小さければ明度を小さくする。

- **テンポ:** 基準演奏よりも早い、または遅いテンポで演奏されている場合、本研究のベクトルマッピング手法を用いると、基準演奏の出力画像に対して  $x$  軸方向に縮小・拡大された形で出力されることになる。よってそれが演奏のテンポの速度を示す。
- **音符間の間隔:** 着目する音の発音とその直前の音の消音の間隔が正の値をとるとき、それを LIC の粒度に対応させる。負の値をとるとき、すなわち直前の音が鳴り終わる前に次の音が発音された場合、その重なり部分に線形平滑フィルタをかけることでフレージングのなめらかさを表現する。
- **ダイナミクスレンジ:** 楽曲の各フレーズに対して、ダイナミクスの差を計算し、その値に対して出力画像の対応フレーズ部分の彩度を変化させる。ダイナミクスが大きければ彩度は高く、小さければ低くする。

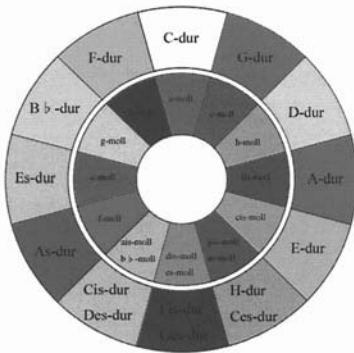


図 4: 調と色の関係

MIDI パラメタと情緒語の対応関係、また、それに対応するマッピング方法を表 1 にまとめる。

#### 4 システム実装と実験

標準的な PC(OS:Windows XP Professional, CPU: Pentium 4 3.00GHz, RAM:2.00GB) 上で言語に C++, ライブラリに MaxMidi Toolkit, グラフィクスパッケージに OpenGL を用いて、拡張 MIDI-LIC 変換を実装した。

実験には、「かえるの歌」(ドイツ民謡), 「トロイメライ」(シューマン作曲) の実演奏から得た SMF を用いた。

表 1: 情緒語とマッピング方法

	軸の詳細	MIDI パラメタ	マッピング方法
1	静かな ← 力強い 0 ← 1	音量	明度
2	ゆったりとした ← 活気のある 0 ← 1	テンポ	出力画像の横幅
3	はねるような ← なめらかな 0 ← 1	ノートオフ・オンの期間	粒度 平滑化フィルタ
4	平坦な ← 色彩豊かな 0 ← 1	ダイナミクスレンジ	彩度
5	暗い ← 明るい 0 ← 1	調性	色相

さらに、拡張 MIDI-LIC 変換で得られた出力結果に SMF を同期させ、アニメーション機能を実装した。アニメーション化することにより、音符と出力画像の対応関係がわかりやすくなり、曲の構造を把握しやすくなった。

#### 4.1 「かえるのうた」を用いた基本実験

PQP のパラメタを適用せず、MIDI-LIC 変換だけを行ったリファレンステクスチャを図 6(a) に示す。楽曲全体としての流れを崩すことのないようにするため、休符部分は元のホワイトノイズテクスチャを表示している。音高の変化が少なく、音価の長い部分については  $x$  軸に近い流れとなり、逆に音高の変化が大きく、音価の短い部分では、 $y$  軸に近い流れになる。また、休符によって元のホワイトノイズが表示されることで、フレーズの切れ目を確認でき、全部で 4 つのフレーズから構成されていることが容易にわかる。

次に色相を変更した出力結果を図 6(b) に示す。ト長調に移調したもので、図 4 に従って全体に緑の色相をマッピングした。ハ長調の単純な感じと比べ、緑色から受ける癒しの感覚が、ト長調の静かで優美な感覚と合っていることがわかる。

明度を変更した出力結果を図 6(c) に示す。明度の違いを明確に示すため、第 1 フレーズでは全体的にクレッシェンド、第 2 フレーズではデクレッシェンドをしている。図 6(a) に比べ、各音符での音量の違いがわかる。

彩度を変更した出力結果を図 6(d) に示す。ダイナミクスレンジの差による出力の違いを示すため、フレーズごとにダイナミクスレンジの幅を大きくしてある。彩度が高くなることで、全体的に鮮やかな感じを受ける。

粒度を変更し、平滑化フィルタを適用した出力結

果を図 6(e) に示す。第 1 から第 3 フレーズまでは、それぞれ違った粒度になるように出力している。第 4 フレーズで、音符の重なりによる平滑化フィルタの効果を示す。粒度が粗くなるにつれて、流れが悪くなるような印象を受け、演奏のなめらかさの度合いを感じ取れる。また、平滑化フィルタをかけることで、音の重なり具合を良く把握することができる。

最後に、PQP のもつすべてのパラメタを変更した統合結果を図 6(f) に示す。各音符での音量の違いや、フレーズごとの演奏の鮮やかさ、なめらかさを把握することができる。

#### 4.2 より複雑なデータによる可視化実験

演奏の違いによる出力結果の違いを示すため、「トロイメライ」を 2 人の演奏者が演奏した MIDI デー

タに本提案手法を適用した出力テクスチャを図 7 にそれぞれ示す。「トロイメライ」はへ長調の曲であるが、この調のもつ単純で牧歌的な印象がオレンジ色の色相によって表されている。A の演奏の方がテンポが速く、それが出力画像の横幅の違いに表れている。A の演奏に比べ、B の演奏の方が彩度が高く、全体的に鮮やかである。実際の演奏では B の方がダイナミクスレンジが大きく、色彩豊かな演奏である。また、B の方が粒度が高く荒々しい演奏に感じる。実際の演奏も A の方がフレーズングがひじょうになめらかで流れるような演奏なのに対し、B は少しはねるような演奏で、両者の演奏の違いを的確に反映していると考えられる。トロイメライは全 6 フレーズでなる曲であるが、各フレーズの最初の音

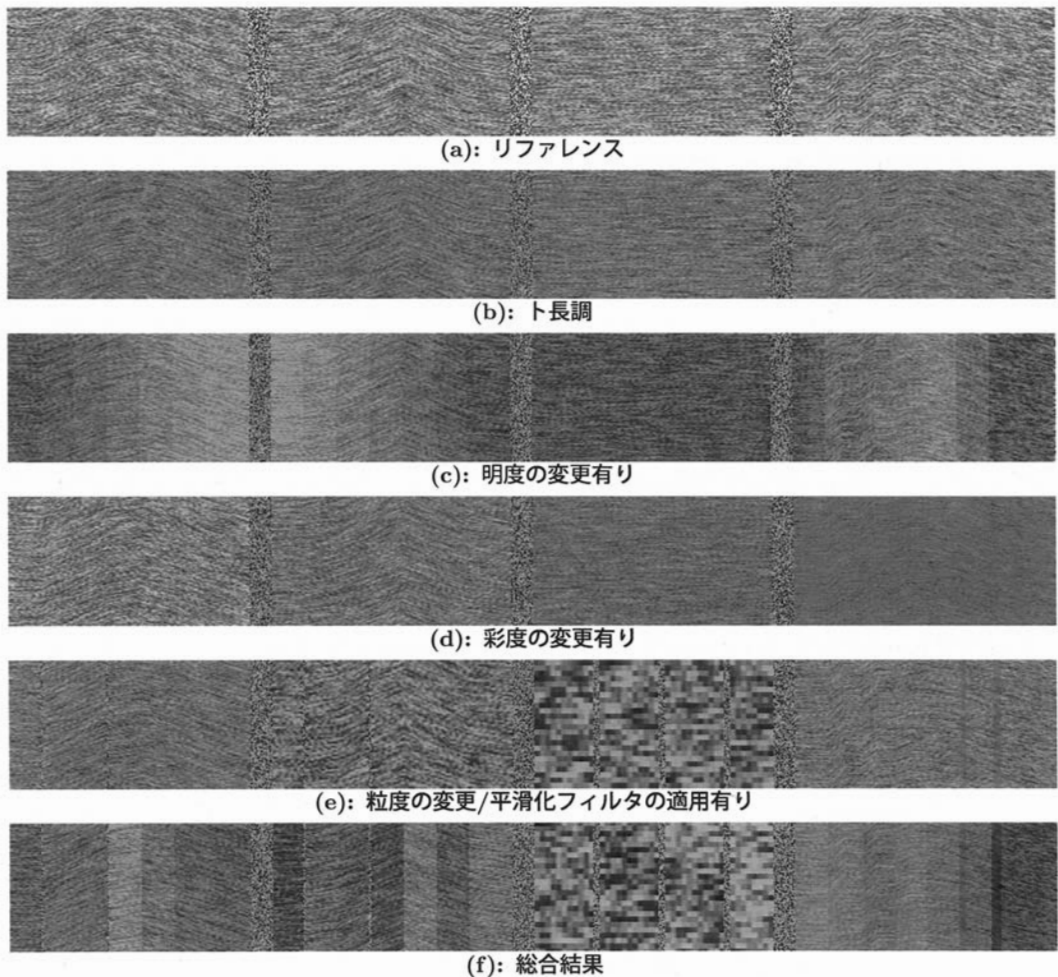
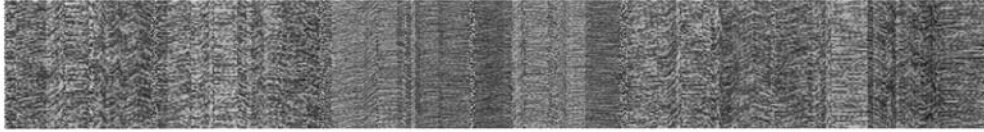
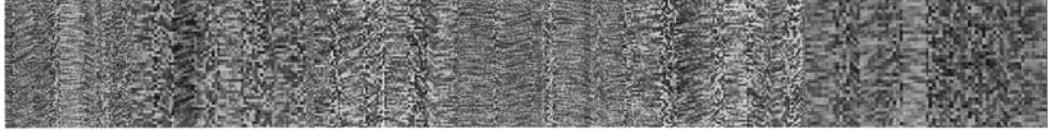


図 6: 「かえるの歌」の演奏における情緒構造の可視化結果



(a) : 演奏者 A



(b) : 演奏者 B

図7: 「トロイメライ」の演奏における情緒構造の違い

の明度が低く、音が小さいのがわかる。これは各フレーズが弱起で始まるためであると考えられる。

## 5 まとめと今後の課題

音楽演奏の情緒構造を表すパラメタを対応するMIDIデータから算出し、そのパラメタを反映させたLICテクスチャに変換することで、音楽演奏の情緒構造を可視化する手法を提案し、その実装を行った。また、実際に演奏者による出力結果の違いを見る実験を行い、その効果を検証した。

より効果的な音楽演奏の情緒可視化 (emotional visualization) 手法の確立に向けて、今後以下のような課題を解決していきたい：

- **フォーカス+コンテキストの実装** : 楽曲の長さによっては、出力結果が横に伸び、ウィンドウ内をスクロールすることでしか全体を把握することができないため、楽曲の全体構造の把握が難しい。フォーカス+コンテキスト表示を適用することで、これを解消する。また、コンテキスト表示された部分が時間軸方向に圧縮され、テクスチャが変化することで、新たな情緒を誘起できると考える。
- **和音への対応** : 現在の拡張MIDI-LIC変換では、単音での演奏の可視化しか行えず、和音に対応していない。これが解消できれば、より複雑な楽曲等が扱える。また、オーケストレーションの違いによる出力の差を見ることも可能になる。

## 謝辞

本研究の一部は、平成19年度科学研究費補助金基盤(B)18300026の支援によって実施された。

## 参考文献

- [1] R. Bresin, A. Friberg : "Emotional coloring of computercontrolled music performances," *Computer Music Journal*, Vol.24, No.4, pp.44-67, 2000.
- [2] F. Watanabe, R. Hiraga, I. Fujishiro : "BRASS: Visualizing scores for assisting music learning," In *Proceedings of 2003 International Computer Music Conference*, pp.107-114, 2003.
- [3] 渡辺ふみ子, 藤代一成, 平賀留美 : 「デジタルスコアによる楽曲学習支援インタフェース」, 情報処理学会論文誌, Vol.45, No.3, pp.710-718, 2004年
- [4] 渡辺 綾子, 藤代 一成 : 「tutti:音楽データのブラウジング編集のための3次元インタラクティブインタフェース」, インタラクティブシステムとソフトウェア IX, 近代科学社, pp.33-38, 2001年
- [5] R. Miyazaki, I. Fujishiro, R. Hiraga : "comp-i: A system for visual exploration and editing of MIDI datasets," In *Proceedings of 2004 International Computer Music Conference*, pp.157-164, 2004.
- [6] 宮崎 麗子, 藤代 一成, 平賀 留美 : 「comp-i:MIDIデータの視覚検索システム」, 情報処理学会論文誌, Vol.45, No.3, pp.739-742, 2004年
- [7] R. Hiraga, N. Matsuda : "Visualization of music performance as an aid to listener's comprehension," In *Proc. AVI'04*, pp.103-106, 2004.
- [8] B. Cabral, C. Leedom : "Imaging vector field using line integral convolution," In *Proc. ACM SIGGRAPH93*, pp.263-270, 1993.
- [9] 渡辺 絵理, 藤代 一成 : 「MIDIデータの画素指向可視化」, 第67回情報処理学会全国大会, 3R-3, 2005年
- [10] M.-H. Kiu, D.C. Banks : "Multi-frequency noise for LIC," In *Proc. IEEE Visualization'96*, pp.121-126, 1996.
- [11] *Band Journal*, 音楽之友社, 2005年
- [12] H. Han, J. Yamashita, I. Fujishiro : "3D haptic shape perception using a 2D device," In *ACM SIGGRAPH2002 Conference Abstracts*, p.135, 2002.
- [13] 野原 泰子 : 「A. スクリャービンの《プロメテウス》作品60 色光ピアノパートに基づく構造と解釈の研究」, 日本音楽学会機関紙, 第47巻, 3号, 2001年