

comp-i : MIDIデータの3次元可視化

宮崎 麗子 藤代 一成

お茶の水女子大学 大学院人間文化研究科

〒 112-8610 東京都文京区大塚 2 - 1 - 1

reiko@imv.is.ocha.ac.jp, fuji@is.ocha.ac.jp

MIDI 楽曲の編集に使われるシーケンスソフトには二つの問題点が考えられる。一つは、曲の再生・編集のためのウィンドウが独立した状態でしか表示されないことである。このため、一度に表示したいウィンドウの枚数が増えると、重なって隠れたウィンドウの内容が見えなくなってしまう。また、個々のウィンドウサイズを小さくすると、情報そのものの認識が困難になる。もう一つは、各ウィンドウが現在再生・編集している部分だけしか表示できないため、曲全体の把握が困難になることである。そこで本稿では、このようなシーケンスソフトにおけるウィンドウの独立性による問題を解決するため、3次元仮想空間内に MIDI データを可視化するシステム comp-i (Comprehensible MIDI Player - Interactive) を提案する。

キーワード: MIDI , ブラウジング , 情報可視化 , 仮想空間 .

comp-i: 3D Visualization of MIDI Datasets

Reiko Miyazaki Issei Fujishiro

Graduate School of Humanities and Sciences
Ochanomizu University

2-1-1 Otsuka, Bunkyo-Ku, Tokyo 112-8610, Japan

reiko@imv.is.ocha.ac.jp, fuji@is.ocha.ac.jp

MIDI (Musical Instrument Digital Interface) is known as a designated protocol for communicating control information on musical composition. So-called sequence software systems, which are commonly used to edit MIDI-encoded music, possess two kinds of problems due to interactions with MIDI data through multiple independent windows. In this article, we address the problems by developing a prototype system, called “comp-i(Comprehensible MIDI Player - Interactive),” which provides a novel type of 3D virtual space, where the users are allowed to interactively understand the global structures and local features embedded in a given MIDI datasets.

Keywords: MIDI , browsing , information visualization , virtual space .

1 背景と目的

音楽の演奏に関する情報を送受信するための代表的な通信プロトコルの一つとして、MIDI (Musical Instrument Digital Interface) が知られている [1][2]。MIDI 規格に対応したコンピュータとソフトウェアがあれば、誰でも MIDI 楽曲を楽しむことができる。一般的にはシーケンスソフトとよばれる専用のソフトウェアを使って曲を編集する。

従来のシーケンスソフトは、エディットウィンドウとよばれる楽曲の再生・編集専用のウィンドウを提供している。エディットウィンドウには、楽譜形式のスタッフウィンドウや音程を棒グラフ状に表示するピアノロールウィンドウ、各種 MIDI データの数値を表示するリストウィンドウなど、ユーザの目的に合わせてさまざまな表示方法が用意されている。一般的に、ユーザは提供された各種ウィンドウの中から、自分の編集目的に合ったウィンドウを次々に選択・表示して編集作業を行う [3]。しかしこの編集方法に従うと、ひとつのウィンドウに表示されるパラメタの種類は限定されるため、複数のウィンドウを見比べなければイベントの総合的な把握が困難になる。また、各ウィンドウは現在注目している部分の局所的な表示しかできないため、曲全体の把握が難しい。

そこで本研究では、これらの問題を解決するための MIDI データの可視化技法を検討する。そして 3 次元仮想空間に MIDI データを可視化するシステム comp-i (Comprehensible MIDI Player - Interactive) を開発し、ウィンドウの独立性と、楽曲の局所的表示による問題の解決を目指している [4]。

ところで、楽曲の 3 次元可視化という点では、すでにいくつかの研究がなされている [5][6]。しかしこれらの研究に伴って開発されたシステムでは、局所的な表示しかできないという問題と、3 次元可視化による透視投影に起因する、詳細度表示における曖昧性の問題がある。そこで comp-i では、音楽の 3 次元可視化という点で、既存の研究からの新規性を図るためにも、楽曲を大局的に表示する機能と、詳細度表示の曖昧性をなくす機能を加える。

多次元化は情報可視化の代表的な手法のひとつであるが、3 次元空間の操作は往々にして困難であり、特に入力・編集作業には効果的なインタフェースの設計が肝要である。我々は、3 次元空間内で MIDI データを直観的に可視化し、ユーザとのインタラクションによる情報探索機能をもたせ、さらにその操作性を損なわないような可視化結果の表示方法や入力インタフェースの仕様を検討する。

また comp-i は、3 次元音を用いることによって臨場感を生成する。音源からの距離感やパンの聴覚情報と可視化結果を対応させることにより、音場のシュミレーション機能をもたせる。

本稿では、2 節で我々の提案するシステム comp-i の機能について説明する。3 節では実装した comp-i のプロトタイプを紹介し、その

効果を考察する。そして 4 節では、可視化例を用いて、楽曲のもつ階層構造の可視化技法について検討を加える。最後に 5 節で本稿をまとめ、今後の課題にふれる。

2 comp-i システム

comp-i は、Standard MIDI File (SMF) に格納された MIDI データを入力し、可視化する。MIDI データは、MIDI ケーブル上を流れる MIDI メッセージと、メモリやハードディスク等に記録される MIDI イベントの二つに区別される。comp-i は、この二つの情報の中から主な要素を選び可視化する。

2.1 MIDI データの可視化方法

可視化対象は、さまざまな MIDI データのうち、MIDI メッセージである全チャンネルのノートオン・ノートオフと、メタイベントであるセットテンポとする。チャンネルとは音源にあたるもので、それぞれのチャンネルに異なる音源を割り振ることができる。ノートオンとノートオフは、それぞれ発音と消音のメッセージであり、チャンネル、音程、音量のパラメタを含む。セットテンポからはテンポのパラメタを取り出し可視化する。

3 次元空間の x 軸に時間、 y 軸に音程、 z 軸にチャンネルをそれぞれマッピングする。また現在の再生箇所を示すために、 x 軸方向に移動するスキャン面を配置する。 x 軸には、小節にあたる区間の節目にポイントノードを配置し、小節番号を表示する。 z 軸にはレイヤ構造をもたせ、1 レイヤに 1 チャンネルを対応させる。奥に表示されたチャンネルの様子も認識するために、各レイヤは半透明で表示する。一つの音はレイヤ上にノートオンからノートオフまで円柱で表し、複数のチャンネルを見分けるために、各レイヤの円柱に異なる色相を対応させる。

各パラメタの可視化方法を以下に示す：

- 音程：SMF に格納された音程の値 (0 ~ 127) をそのまま座標値に直す方法と、入力した SMF の最高音程と最低音程を座標の最大値と最小値に対応させる方法の二つが考えられる。前者は複数の SMF の比較に、後者は空間を有効に使った音程差の明確な表現にそれぞれ向いている。
- 音量：円柱の太さを音量に対応させ、音量が多いほど円柱を太く表示する。音程を明確にするため、円柱は半透明で表示し、その中心に実線を通す。
- テンポ：HSV 変換を用いて、テンポの変化に彩度を対応させ、テンポが速いほど円柱を鮮やかに表示する。これにより、テンポが遅くなると円柱はより白く表示される。明度については、3 次元空間内の光源からの距離によって設定した値とは異なる表示をしてしまうため、採用しない。

2.2 機能

comp-i では、情報探索、再生、編集、音場の生成といった 4 つの大きな機能をもつ。ユーザは、comp-i の提供するさまざまなインタラクション機能により、可視化された 3 次元空間内で必要な情報を探索して、編集作業を行う。comp-i には再生機能もあるので、視覚的な情報に加え、もちろん聴覚的にも楽曲を確認することができる。また 3 次元音を用いることにより、可視化された 3 次元空間内に、従来のシーケンスソフトにはない高い臨場感を生成する。また、入力した MIDI 楽曲全体を表示し、その階層構造を可視化する技法も検討しており、これについては 4 節で述べる。

2.2.1 情報探索

視覚的情報検索システムとしての必要条件を満たすかどうかを検証するため、comp-i の情報探索機能を Shneiderman の提唱する *Visual Information Seeking Mantra* [7] に従って分類する：

- **Overview** : すべてのイベントを把握するために、3 次元空間全体をとらえる位置に視点を設定できる。
- **Zoom** : 注目したい対象の詳細を認識するために視点を近づけことができる。また、特定のチャンネルに注目する場合は、対応するレイヤを手前に表示できる。
- **Filter** : 特に必要な情報を抽出するために、他のパラメタ、チャンネルの可視化表現を、単純化または非表示に切り替えることができる。また、ユーザは各レイヤを自由に並べ替えることができる。これにより、特定された複数のチャンネルの関連に注目したい場合は、対応するレイヤを隣り合わせに並べて、比較しやすいように表示できる。
- **Details-on-Demand** : デフォルトでは、仮想空間の底面とスキャン面にスケールを表示できる。ユーザの要求に応じて、時間を表す x 軸と音程を表す y 軸にスケールを表示できる。このとき、すべてのチャンネルに対してスケールを表示すると、クラッタリングの原因になってしまうため、ユーザの指定したチャンネル上にだけスケールの付いたレイヤを表示できる。また、マウスとキーボードで注目するチャンネルやパラメタを指定すると、各パラメタの数値を表示できる。

2.2.2 再生

曲を再生するとスキャン面を表示し、現在の再生箇所を示す。スキャン面は半透明で表示されているので、スキャン面と交わっている円柱の色が変化し、どの音が鳴っているのかがとらえやすい。

再生方法には、任意の小節番号から再生するモードを提供する。また、チャンネルごとに音の

オン・オフを切り替えるモードを提供し、Filter 機能の効果を聴覚的に拡張する。

2.2.3 編集

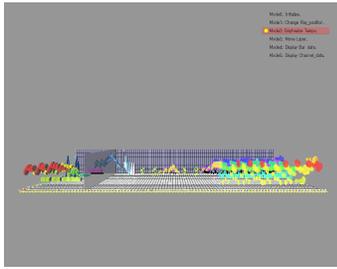
3 次元空間は、一般的に奥行き感のある透視投影で描画されることが多い。このため、楽曲を 3 次元空間内に可視化すると、表示される情報の詳細度が曖昧になりがちである。特に詳細な情報を制御しなければならない編集作業では、3 次元空間における情報の操作性を改善しなければならない。そこで comp-i では、投影方法を切り替えることによってこの問題を回避する。ユーザは、通常透視投影で描画された 3 次元空間内で情報探索を行い、編集作業に移行する場合に、投影方法を平行投影に切り替えることができる。これにより、透視投影では曖昧性であった情報も、平行投影によって正確な値を認識することができる。

投影方法の変換とともに、編集機能においてはインタフェースの設計も重要である。comp-i ではパラメタの入力方法として、可視化されたオブジェクトをマウスによって操作する方法と、テキストボックスから値を入力する方法を提供する。オブジェクトを直接操作すると、直観的に音量やテンポを変更することができ、前後のイベントとの脈絡も考えやすい。また、MIDI のパラメタ値についての知識がないユーザにとっても、この方法であれば編集作業の効率が良い。一方、MIDI について知識があるユーザにとっては、パラメタ値を直接入力の方が効率の良いこともある。特に音程のパラメタは正確な値の入力を必要とするので、この方法が向いている。

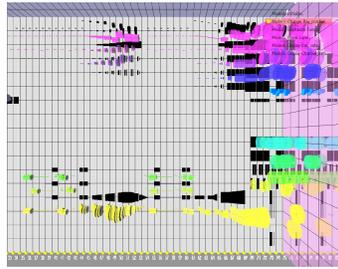
想定される comp-i を用いた編集スタイルについて説明する。ユーザは、情報探索機能によって編集したいパラメタ、およびオブジェクトを見つける。そしてそのオブジェクトを直接マウスで変形させたり移動させることによって各パラメタ値を編集したり、編集したいオブジェクトをマウスで指定してテキストボックスからパラメタ値を入力する。ユーザはこれらのインタラクションによって、リアルタイムで編集結果を確認できる。ここで、透視投影による 3 次元空間表示の曖昧性を回避したい場合は、いつでも平行投影へ切り替えられる。

2.2.4 音場の生成

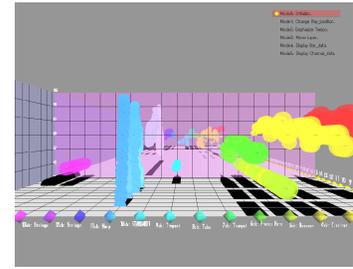
3 次元空間における可視化の応用として、comp-i は 3 次元音を用いてレイヤとの聴覚的な距離感を生成する機能をもつ。この機能を使って、手前に表示したレイヤに対応するチャンネルほど、近くで聴いている感覚を生成し、情報探索の Zoom 機能の効果を聴覚的に増強することができる。また、パンとよばれる音の定位（ステレオ再生時の左右位置）を調整する MIDI メッセージを視覚的に確認するために、視点とスキャン面の位置を考慮した可視化を行う。



(a) 全体



(b) 真上からの投影



(c) 真横からの投影

図 2: comp-i の仮想空間内における視点移動

3 プロトタイプの実装

3.1 実装環境

ハードウェアプラットフォームには IntergraphZx10(RAM:523MB, CPU:PentiumPro III¹933MHz, OS:WindowsNT4.0)を用いた。また3次元音生成機器としてRSS-10システム²を用いた。RSS-10は距離感,方向感による3次元音場生成のほかに,直接音と壁や床などで反射した反射音も併せて聞くことで,さらなる距離感を得ることができる。また,空間内の残響感を出すこともできる。これらによりドップラー効果やフランジング効果を実現でき,より現実的な音場の生成を行なうことができる。ソフトウェアライブラリにはWorldToolKit (WTK) Release 9^{TM3},3次元音の制御はRSS-10コントロールライブラリ⁴を用いた。これはWTKのアプリケーションプログラムで使用されることを前提としたC言語ライブラリである。

3.2 comp-iの可視化空間

16チャンネルから構成されるSMFを読み込み,3次元空間に音程,チャンネル,音量,テンポの各パラメータを可視化するプロトタイプの機能を説明する。

仮想空間の時間軸方向には,小節線を表すポイントノードと小節番号を表示する。また,奥行き方向には,各チャンネルを操作するために,正四面体のチャンネルノード配置する。チャンネルノードは,デフォルトではすべて回転しているが,任意のモードによって特定のチャンネルが指定された場合は,対応したチャンネルノードだけを回転させる。また,テンポに円柱の彩度を対応させているため,各チャンネルノードの色はデフォルトの彩度で表示し,彩度の変化を比較できるようにする。

スキャン面とレイヤ,仮想空間の底面にはそれぞれスケールを表示し,各円柱のチャンネル,音程,時間の各パラメータの認識を助ける。仮想空間の底面には,各円柱をy軸方向に投影し

た影を描画し,底面から離れて表示された円柱のチャンネル,時間,音量差の情報の損失を防ぐ(図1)。

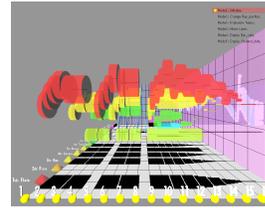


図 1: 影の描画

また,スキャン面を3次元空間におけるスピーカとして定義しているため,視点をスキャン面に近づけると,聴覚的にも近くで聞いているような感覚を生成する。

3.3 視点移動による情報探索

comp-i 起動時は,仮想空間全体がディスプレイに収まる位置に,視点を設定してある(図2(a))。

ユーザはキーボードにより左右,上下,奥行き方向に視点を移動し,マウスを使って視線方向を自由に調節できる。また,キーボードからの入力により,あらかじめ指定された位置(仮想空間全体の斜め上,真正面,真上,真横)に視点を移動できる。例えば,視点を真上に移動すると,各チャンネルの時間軸方向の変化の様子がよくわかる(図2(b))。このとき,音量と発音・消音のタイミングだけが認識できるので,音程,つまり調に関係なく,同じパターンを見つけることができる。また,視点を真横に移動すると,16チャンネルすべてを操作するのに便利である(図2(c))。

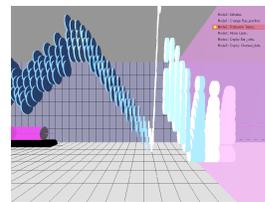


図 3: Zoom 表示

その他,特定の小節にアクセスするために,小節の節目を表すポイントノードをクリックすると,視点がその近くまで移動しMIDIイベントの詳細を確認できる(図3)。

3.4 インタクション機能

comp-i では,2つの再生モードと,5つの表示モードを提供する。これらのモードは,画面の右上の位置にヘッドアップディスプレイとして常に表示されている。現在選択されているモードには赤いプレートを表示する。

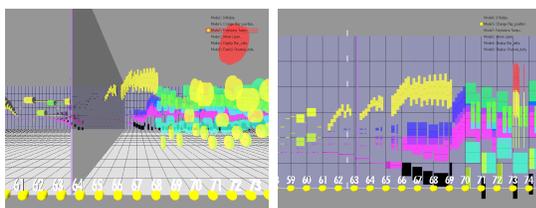
¹PentiumProはIntel社の商標である。

²RSS-10はローランド株式会社の商標である。

³WorldToolKitはSense8社の商品登録である。

⁴RSS-10コントロールライブラリは旭エレクトロニクス株式会社から提供される。

- **Initialize:** デフォルトの再生モード。曲を再生すると、 x 軸の正の方向にそって、現在再生されている場所を示すスキャン面が移動する。このスキャン面の移動速度は、テンポ変化に対応している。
- **Change Play Position:** 任意の小節からの途中再生モード。ポイントノードをクリックすると、対応する小節から曲を再生できる。
- **Emphasize Tempo:** テンポ変化の強調表示モード。デフォルトでは RGB 変換を用いて円柱を表示するが、HSV 変換の表示に切り替え、彩度をテンポに対応させる。
- **Move Layer:** スケールの付いたレイヤを任意のチャンネル上に移動する表示モード。これにより、特定のチャンネルの音程変化の詳細を認識できる。
- **Display Bar Data:** 各小節の詳細情報表示モード。任意の小節に対応したポイントノードをクリックすると、小節ごとに演奏しているチャンネル番号、対応する音源名、テンポなどの詳細情報を表示する。詳細情報を表示するプレートは、仮想空間の表示を妨げないように半透明で表示する。また、選択された小節に対応した、仮想空間上の底面にもこのプレートと同じ色をつける。
- **Display Channel Data:** 各チャンネルの詳細情報表示モード。任意のチャンネルノードをクリックすると、選択されたチャンネルノードだけが回転し、対応するチャンネルの音源名、音程範囲、音量範囲などの詳細情報を表示する。ここでも詳細情報を表示するプレートは、半透明で表示する。
- **Change Projection:** 3次元空間の投影方法を変換するモード(図4)。透視投影から平行投影に変換するときは、任意の視点からの投影を表示できる。しかし平行投影では、視点を奥行き方向に移動しても投影結果は変わらない。そこで、キーボードからの入力によって、投影されるクリッピング面の範囲を変更して表示する空間の範囲を調整することで、この問題を回避する。

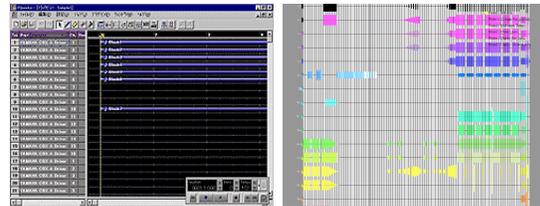


(a) 透視投影 (b) 平行投影

図4: 投影変換

3.5 シーケンスソフトとの比較

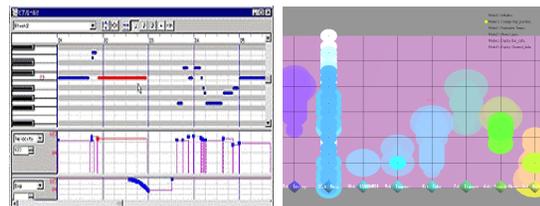
図5(a)、図6(a)に広く普及しているシーケンスソフトの一つであるYAMAHAのXGworks Ver4.0を挙げ、comp-iのプロトタイプと比較する。comp-iは、直交投影の軸選択により、シーケンスソフトの各エディットウィンドウの役割を果たす。仮想空間を真上から見ると、全チャンネルの演奏状態を示すトラックビューウィンドウと同じ役割を果たす。しかしトラックビューウィンドウでは、どのチャンネルでいつ演奏が行われているか、というくらいの曖昧な情報が得られないが、comp-iでは、実際の明確な発音状況に加え、音量とテンポも同時に確認できる(図5)。



(a) トラックビュー (b) 真上からの平行投影

図5: シーケンスソフトとの比較1

また、座標空間を真横から見ると、音程と音量を示すピアノロールウィンドウと同じ役割を果たす。しかしピアノロールウィンドウでは、ウィンドウを二つに区切って音程とボリュームを示すため、二つの表示を見比べなければならないが、comp-iでは円柱の半径の変化から、同時に複数のチャンネルの音程と音量の両方を、より直観的に確認できる(図6)。



(a) ピアノロール (b) 真横からの平行投影

図6: シーケンスソフトとの比較2

4 楽曲構造の表示

曲を理解する上で重要なのは、部分的な音の認識だけではない。曲全体を認識し、楽曲構造を理解することも必要である。そこでcomp-iでは、MIDI楽曲全体の認識のためにfocus+context情報を表示する。これにより、ユーザの注目している部分を際立たせて表示し、ユーザは曲全体をとらえながら、注目部分の詳細情報を確認することができる。

特に楽曲の大局的な表示方法として、帯状の仮想空間をリング状に表示する方法を検討中である。これには、スキャン面の帯上における位置からよりも、リング上における角度からのほうが、注目箇所の楽曲全体に対する割合がわかりやすいという利点がある。また、チャンネル数

は1つのMIDIコンピュータインタフェースで高々16なのに対し、入力される曲はより長くなることあり得るので、縦横両方向に偏りのないリング状の構造の方が、ディスプレイを有効に利用できると思われる。

次に楽曲構造について述べる。ここで言う楽曲構造とは以下のことを示す：

1. 楽譜上の音符の集まり (occurrence) に音楽的な意味が存在
 - 音の速さのゆらぎ (agogics)
 - 音の大きさの変化 (dynamics)
2. 具体的な音の列に依存しない (context free) 演奏ルール
3. 音楽の階層構造において上位階層が下位階層へ与える演奏上の影響

comp-i では、この中から特に1と3の楽曲構造を可視化する機能を検討する。楽曲構造1については、comp-i で用いられる可視化方法によってある程度の認識は可能になる。音の早さのゆらぎは、円柱の彩度の変化から、音の大きさの変化は、円柱の太さから確認できる。

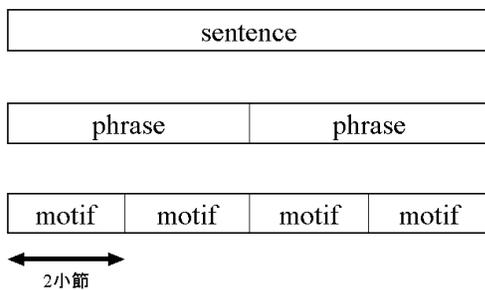
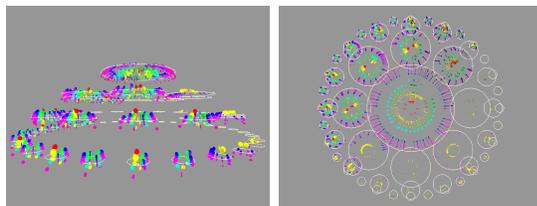


図7: 楽曲の階層構造

また、楽曲のもつ階層構造の可視化については、階層構造の代表的な可視化技法であるConeTrees[8]の応用を検討している。楽曲は一般的に複数のセンテンスで構成され、センテンスは楽曲の最小単位であるモチーフから構成される。また、モチーフの上位階層にはフレーズが存在する(図7)。comp-iでは、ひとつの楽曲をセンテンス、フレーズ、モチーフの各階層ごとに表示する機能をもたせ、楽曲構造単位の類似性や、上位階層から下位階層への音楽解釈の影響を可視化する(図8)。



(a) 横からの投影 (b) 上からの投影

図8: 楽曲階層構造の可視化

5 まとめと今後の課題

シーケンスソフトにおけるウィンドウの独立性と楽曲の局所的表示による問題を解決するため、3次元仮想空間にMIDIデータを可視化するシステムcomp-iを提案した。また機能の一部を実現したプロトタイプを実装した。

今後の課題を以下にあげる。

- 可視化データ型の拡大：持続音系楽器の抑揚を表現するため、エクスプレッションとよばれるMIDIメッセージを可視化する。また、パーカッションについては、SMFへのデータの格納方法が一般的な楽器の音源とは異なるため、他の表示方法を検討する。
- 楽曲構造の表示：楽曲の大局的認識のために、階層構造の可視化技法をさらに検討し、オペレーション機能を再設計する。
- 音場の生成：情報探索のZoom機能を増強するために、3次元音を用いてレイヤとの聴覚的な距離感を生成する。また、パンのMIDIメッセージをより直観的に認識するために、3次元音による聴覚情報と、スクリーン面の位置情報を同期させる。

謝辞

本研究に対し、貴重なコメントを頂いている文教大学情報学部の平賀 瑠美先生に深く感謝いたします。

参考文献

- [1] 中島安貴彦(編,著): MIDIバイブル I -MIDI 1.0規格 基礎編-, リットーミュージック出版編集部編, 1997年11月
- [2] 中島安貴彦(編,著): MIDIバイブル II -MIDI 1.0規格 実用編-, リットーミュージック出版編集部編, 1998年3月
- [3] <http://www.yamaha.co.jp/product/syndtm/read/xgworks/>
- [4] 宮崎, 藤代: 「comp-i -MIDIデータの可視化-」, 第64回情報処理学会全国大会, 3E-03, 2002年3月
- [5] Sean M. Smith, Glen N. Williams: "A Visualization of Music," In *Proc. of IEEE Visualization '97*, October 1997, pp.499-503.
- [6] Hans G. Kaper, Sever Tipei: "Manifold Compositions, Music Visualization, and Scientific Sonification in an Immersive Virtual-Reality Environment," In *Proc. of International Computer Music Conference*, October 1998, pp.339-405.
- [7] Shneiderman, B.: *Designing the User Interface Strategies for Effective Human-Computer Interaction*, Addison-Wesley, 3rd edition, 1998, Chapter 15.
- [8] G. G. Robertson, et al.: "ConeTrees: Animated 3D Visualizations Hierarchical Information," In *Proc. CHI'91, ACM Conference on Human Factors in Computing System*, New York, pp.189-194, 1991.