

## tutti : 音楽データのブラウジング・編集向け 3次元空間インタフェース

渡辺 綾子      藤代 一成<sup>†</sup>

お茶の水女子大学 大学院人間文化研究科

<sup>†</sup> お茶の水女子大学 理学部 情報科学科

〒112-8610 東京都文京区大塚 2-1-1

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

オーディオデータのブラウジングは、その時間的特性から文書データのブラウジングよりも困難である。人間のもつカクテルパーティ効果と空間的位置の記憶の特性を利用して、音声データを効率よくブラウジングするための空間的インタフェースとして Dynamic Soundscape (DS) システムが知られている。本稿では、この DS アプローチを拡張して、音楽データを効果的にブラウジングするだけでなく編集も可能にするシングルユーザインタフェースシステム tutti (Total Utility for Time Travel Immersion) を提案する。

キーワード: 音楽データ, ブラウジング, 編集, ユーザインタフェース, カクテルパーティ効果, 空間的位置記憶

## tutti:A 3D Spatial Interface for Browsing and Editing Sound Data

Ayako Watanabe      Issei Fujishiro<sup>†</sup>

Graduate School of Humanities and Sciences

<sup>†</sup>Department of Information Sciences

Ochanomizu University

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

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

Browsing audio data is not so easy as browsing printed documents because of the temporal nature of sound. “Dynamic Soundscape” (DS) is known as a spatial interface for temporal navigation of audio data by taking advantage of human abilities of cocktail party effect and memory of spatial location. In this article, we extend the DS approach to propose a system called “tutti (Total Utility for Time Travel Immersion)”, which provides the users with a unified interface to browse and edit sound data effectively.

**Keywords:** Sound data, browsing, editing, user interface, cocktail party effect, spatial location memory.

## 1 はじめに

医療技術や宇宙開発，建築，映画，芸術など幅広い分野で，バーチャルリアリティ(Virtual Reality：VR)システムが利用されている．これらのシステムでは，人間の五感である視覚，聴覚，触覚などを刺激することにより，現実感を生成する試みが行われている．なかでも立体的な映像を用いて視覚に訴える場合が多い．しかし近年では，高い臨場感を実現するために，視覚情報に加え，聴覚情報である3次元音響の重要性が増してきている．

視覚の有効範囲は視線方向に限定されるのに対し，聴覚は全方向的である．また，人間は左右の耳に到達する音の時間差や周波数の違い，音圧の減衰や反射音などから，音源の方向，移動感，距離感などを感じとっている．このように，聴覚は視覚とは全く異なった側面をもち，それ特有の特性が存在する．視覚提示技術を中心としたVRシステムにおいて，総合的な臨場感を高める情報として用いている聴覚情報は，視覚情報を補助するという役割を担っている．

聴覚特有の特性を利用し有効なインタラクションを実現しているシステムとして，Kobayashiらによって報告されている Dynamic Soundscape (DS) システムがあげられる [1]．これは，オーディオデータである音声データの時間軸を空間軸にマッピングし，効率よくブラウジングを行うインタフェースである．本研究は聴覚の特性を活用するインタフェースを目指し，音楽データを対象としてDSを拡張したシステム tutti (Total Utility for Time Travel Immersion) を提案する [2]．

tuttiはその拡張機能の一つとして，音楽データ編集機能をもつ．有名な画家が描いた絵画作品は鑑賞によって接することができるのに対し，過去の偉大な作曲家が残した音楽作品には，鑑賞の他に演奏することで触れ合える素晴らしい特徴がある．本システムの拡張機能である音楽編集機能は，曲の表情付けを可能とする．ユーザはこの機能を通じて音楽と能動的に関わることができる．本システムは，対象者を演奏経験

者に限定せず，楽器の演奏経験がない人や楽譜の読み方を習得していない人も操作可能なユーザインタフェースを目指す．編集できる音楽要素は，音色，テンポ，音の強弱の3つとし，これらの組合せにより曲の表情付けを行う．

本システムが，音楽データのブラウジングと編集機能をもつことは先に述べたが，これらの操作を別のウィンドウで行うと作業効率が落ちてしまう．音楽の表情付けは，音楽を聴きながら行うのが最良である．そこで，我々の先行研究 [3] でも利用した，WIM(World-in-Miniature)[4, 5] のパラダイムを採用し，単一のウィンドウで音楽データのブラウジングと編集が行えるユーザインタフェースの開発を目指す．

本稿では，2節で関連研究であるDSについてふれたのち，3節で我々の実装したシステム tuttiの音響空間と機能を紹介し，WIMの概念，利用方法を説明する．4節でシステムを考察，最後に5節で本稿をまとめ，今後の課題にふれる．

## 2 Dynamic Soundscape

オーディオデータのブラウジングは，文書データのブラウジングよりも困難である．これは，オーディオデータが時間的特性をもつ時系列データであることに起因する．オーディオデータの内容すべてを理解するには，最初から最後まで聞く必要がある．また早送りや巻き戻しの間，再生不可能な場合その間の情報は得られない．以上よりKobayashiらは，オーディオデータの効果的なブラウジングの有用性を指摘し，DSの研究に至っている．

今日，情報科学の分野で扱われているオーディオデータには音声と音楽の2つがあげられる．音声データは，人間が日常のコミュニケーションの手段として利用している言語の情報であり，一度の試聴で提示したい多くの情報が伝わる．DSは，その音声データのブラウジングを目的としている．

一方音楽データの場合，一度の試聴で曲全体

を口ずさむことは不可能である．また，表情付けを行う場合は，曲の深い理解が必要である．よって，音楽データの効果的なブラウジングも重要であると考え，音楽データを対象として DS を拡張する．

### 3 tuttiシステム

はじめにシステムの実装環境を述べる．次にシステムの概要として，音楽データのブラウジングと編集機能について説明し，我々の先行研究 [3] でも採り入れた WIM パラダイム [4, 5] の基本概念を説明し，tutti での利用方法を述べる．

#### 3.1 実装環境

プラットフォームには IntergraphZx10 (CPU:PentiumProIII<sup>1</sup>933MHz, RAM:523MB, 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<sup>2</sup> を用いる．RSS-10 は距離感，方向感のほかに，直接音と反射音を合わせることで現実的な音場の生成を実現できる．ソフトウェアライブラリには World-ToolKit (WTK) Release 9<sup>TM3</sup>，3 次元音の制御は RSS-10 コントロールライブラリ<sup>4</sup> を用いる．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．音楽データは，スタンダード MIDI ファイル (SMF) を使用する．

#### 3.2 tutti の音響空間

図 1 に tutti の音響空間を示す．3 次元オブジェクトである球をワイヤフレームで表示し，外枠を構成する．その結果，内外から見る事が可能となる．音源を表すスピーカーは，ソリッドな球として表示する．また，経線と緯線の交点でかつスピーカーの軌道となる緯線上に，目

印となる小さな球を配置する．以後，これをポイントノードとよぶ．ユーザのリスニングポジションは球の原点とする．球上のすべての点から等距離である原点は，3 次元音響の構成に好影響を与える．

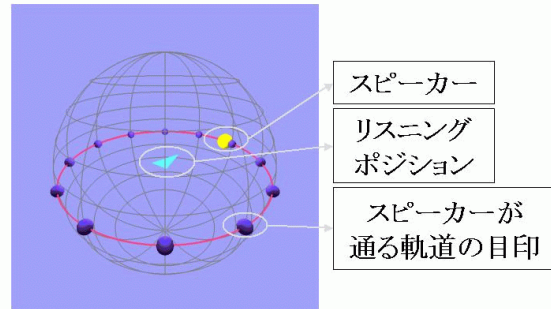


図 1: tutti の音響空間

#### 3.3 音楽データのブラウジング

tutti を起動すると，図 2 のように赤道上に 1 つのスピーカーが生成される．スピーカーは 1 つの音楽データを再生しながら，赤道上进行して回転する．そのとき重要なことは，回転しながら時間軸から空間軸へのマッピングを行うことである．軌道上で再度聞きたい部分を指し示すと，そこに新たなスピーカーが生成され，その場所に応じたオーディオデータを再生起点として回り始める (図 3)．オリジナルのスピーカーは，新しいスピーカーが生成された後も，音量を下げて再生を続ける．これにより，ユーザは複数のスピーカーから，単一の音楽データの異なる部分を聞くことができる．スピーカーの選択はユーザによって行われ，再びオリジナルのスピーカーに注意が向けられれば音量は大きくなる．最も注目されているスピーカー以外は小さな音量で再生される．また早送りを行った場合，オリジナルのスピーカーがスキップされたオーディオデータを再生してくれるため，情報の欠落を防ぐことができる．

このシステムは DS と同様に，人間の 2 つの

<sup>1</sup>PentiumPro は Intel 社の商標である．

<sup>2</sup>RSS-10 はローランド株式会社の商標である．

<sup>3</sup>WorldToolKit は Sense8 社の商品登録である．

<sup>4</sup>RSS-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．

能力を利用している．一つはカクテルパーティ効果 (cocktail party effect) とよばれるもので，人間は2つ以上の音が混在しているときに，関心のあるものを選択的に聴取できる能力をもつ．ユーザは tutti の音響空間に複数個提示されているスピーカーからオーディオデータを聞き分け，聴取したいものを選択することができる．もう一つは空間的な位置の記憶 (spatial location memory) である．あるイベントが生じたとき，そのイベントは空間的な位置という視覚的特性と結合して記憶される．ユーザは，「20秒前」といった時間的表現を使う代わりに，「この辺りで聞いたフレーズ」などの空間的な記憶を頼りに，オーディオデータに効果的にアクセスできる．

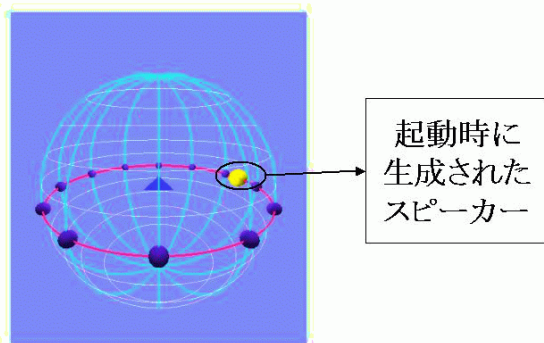


図 2: 起動時のシステム画面

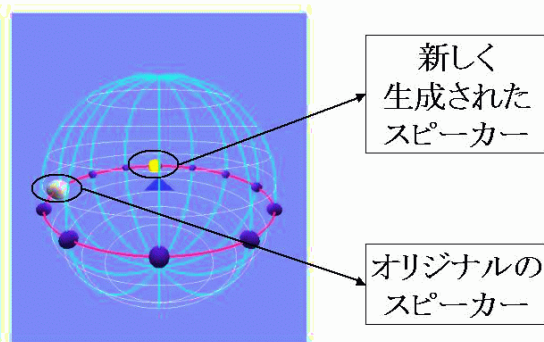


図 3: スピーカー生成時のシステム画面

### 3.4 音楽データの編集機能

軌道の上に配置されているポイントノードをマウスで操作することで，3つの音楽要素の編集を行う．

- 音の強弱：スピーカーの軌道を変形させて，音の強弱を変化させる．軌道は3次元ベジエ曲線で記述されているので，連続的な軌道変形操作が可能である．軌道を球の中心に近づければ音量が大きくなり，遠ざければ小さくなる．このとき，MIDIのノートオンメッセージのベロシティは変化させず，3次元音だけで強弱を実現する．スピーカーの軌道は連続なので，なめらかなクレッシェンド，デクレッシェンドが実現できる (図 4(a)) ．
- テンポ：ワイヤーフレームの経線の幅と SMF のテンポを関連付け，経線の幅の伸縮によってテンポを変化させる．伸ばすとテンポが遅くなり，縮めるとテンポが速くなる (図 4(b)) ．ただし，音の強弱の編集によって経線の幅に影響が生じた場合は，テンポの変化はないものとする ．
- 音色：ワイヤーフレームで表示されている球の緯度ごとに音色を対応づける．スピーカーがポイントノードを通過するときに，その緯度に対応したプログラムチェンジメッセージを送信する．ユーザはポイントノードの緯度，つまり高さを設定してスピーカーの軌道を変えることで，音色を変更できる (図 4(c)) ．

### 3.5 WIM の導入

#### 3.5.1 WIM とは

多くの VR 環境では，1つの画面が与えられ，1つの視点に限られている．このような仮想環境で環境全体を把握するのは困難であり，ユーザが即座に到達できる範囲も限られている．これらの問題を考慮して提案されたのが，WIM

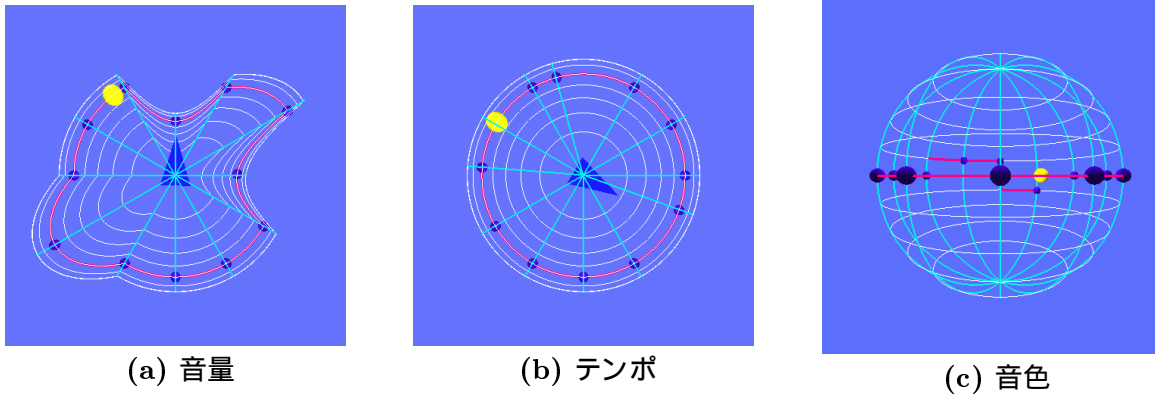


図 4: 音楽データの編集

である。WIM とは、仮想環境全体の 3 次元ミニチュアの中で、仮想空間におけるナビゲーションや操作に利用される。WIM にはユーザの化身であるアバタがおかれる。ディスプレイには、WIM とアバタから見た仮想環境が同時に表示される。ユーザが WIM の中のアバタやオブジェクトを動かすと、連動して仮想環境の視界やオブジェクトも動く。WIM により、ユーザはアバタやオブジェクトを容易に移動でき、また自由に仮想環境を再構成することができる。一般に、仮想空間の一箇所に没入した場合、全体像の認識が難しくなる。WIM を導入することによって、ユーザが仮想空間に没入しながらも、同時に環境全体を把握できる利点がある [4, 5]。

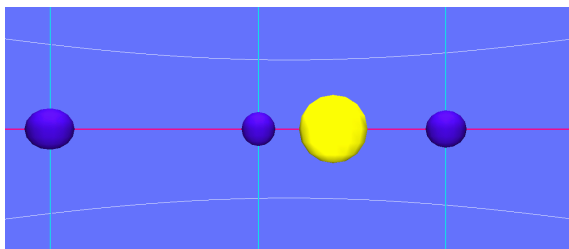


図 5: アバタからの視界

### 3.5.2 WIM を採り入れた音響空間の表示方法

本システムでは、リスニングポイントである球の中心にアバタを設置する。アバタの視線方向

を変更することはできるが、アバタ自身を平行移動することは許されていない。

そのため、アバタから見た仮想環境は局所的で、球を形成していることさえ把握できない (図 5)。そこで、WIM を用いて大局的な視点を与え、タスクのスループット向上を図る。図 6 の上部には、真上、斜め上、真横の 3 方向からの WIM が表示されている。

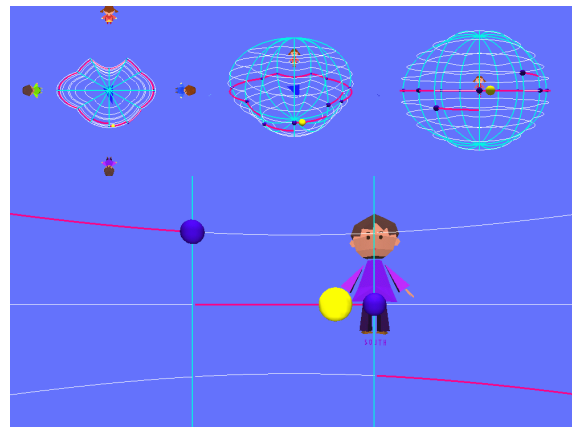


図 6: tutti における WIM

## 4 システムの考察

日常生活は音楽も含め音で溢れている。その多くは「音楽を聴く」といった受け身による関わりにとどまっています、能動的な関わりに至っていない。我々が提案したシステム tutti は、楽器の演奏、作品の制作ほどの能動性はないが、作品に積極的に近づき、表情付けを行うことで、

感情の表出を促すことを目的としている．tutti の使用方法として，主な次の3つをあげる．

- 教育支援：tutti は楽譜などの表示を用いず，初心者でも直感的に操作可能なインタフェースを提供している．tutti を通じて音楽に対する興味，関心を高め，表現活動に対する意欲を育成したい．
- 編曲支援：編集可能な音楽要素は3つとしているため，高度な編曲は難しい．しかし，曲の表情付けを行うことで創造的な活動の支援になると考えられる．出力形式に楽譜が加われば有効である．
- 演奏の比較：球をワイヤフレームで表示しているため，変形具合が視覚によって理解できる．演奏者による演奏方法の比較，また自分の表情付けと他者との比較により，音楽に対する創造性がさらに高まることを期待できる．

## 5 まとめと今後の予定

本稿では，音楽データを対象としてブラウジングと編集を行うためのシングルユーザインタフェース tutti を提案した．

今後は，システムのユーザテストを行い，改良を行う予定である．以下の点を検討項目としてあげておく．

- 音楽データのブラウジング機能
  - 聴覚情報として扱っている3次元音響は，自然界音響の模倣であって同等ではないため，ユーザが音源位置を正確に把握しているかどうかの基礎的な検討を重ねる．
  - 音声データはトピックごとに分割ができ，空間的位置との関連付けがしやすい．空間にマッピングされたオーディオデータにアクセスするためには，この点が重要となってくる．本研究で対象としている音楽データの

空間的位置との関連付けについて検討する．

- 同一話者の音声データが複数提示されると，それぞれの音声データの解釈が困難になることが報告されている．よって，音楽データを複数提示した場合も，聞き取りの難易度を評価しなければならない．
- 音楽データの表情付け機能
  - 音楽編集機能による表情付けで，音楽性を損なってしまう場合があるかどうかを考慮する．また，他の音楽要素を採り入れるべきか，その場合はどのように実現できるかを検討していく．

## 参考文献

- [1] M. Kobayashi, C. Schmandt: “Dynamic Soundscape: Mapping Time to Space for Audio Browsing,” In *Proc. ACM SIGCHI Human Factors in Computer Systems Conference*, March 1997, pp. 194–201.
- [2] 渡辺, 渡辺, 藤代, 竹島: 「tutti:音楽データのためのブラウジング・編集システム」, 第62回情報処理学会全国大会, 7Q-05, 2001年3月
- [3] 渡辺, 藤代, 竹島: 「WIMを用いた3次元仮想音響制御」, 第60回情報処理学会全国大会, 6Z-02, 2000年3月
- [4] R. Stoakley, M. Conway, R. Pausch: “Virtual Reality on a WIM,” In *Proc. ACM SIGCHI Human Factors in Computer Systems Conference*, May 1995, pp. 265–272.
- [5] R. Pausch, et al.: “Navigation and Locomotion in Virtual Worlds via Flight into Hand-Held Miniatures,” In *Proc. SIGGRAPH 95*, August 1995, pp. 399–400.