

# フレーズと音楽プリミティブの相互関係の可視化による 旋律創作支援の試み

西本一志<sup>†</sup> 間瀬健二<sup>†</sup> 中津良平<sup>†</sup>

本論文では、特に即興演奏における新たな旋律の創造の支援を目的とした、旋律データの可視化手法と、それに基づく旋律創作ルール抽出と新しい旋律創作のためのヒントの獲得手法を提案する。本手法では、旋律を構成するフレーズおよび各音楽プリミティブの相互の関連を空間的な構造として表現し可視化する。この空間を利用者が分析することにより、様々な旋律創作のルールやヒントを得ることができる。プロトタイプシステムを演奏例に適用した実験から、静的な状況での旋律分析や、実時間状況での即興演奏における演奏構成や新たなアイデアの考案における、本手法の有効性が示唆された。

## Melody Creation Support by Visualization of Relationships among Phrases and Musical Primitives

KAZUSHI NISHIMOTO,<sup>†</sup> KENJI MASE<sup>†</sup> and RYOHEI NAKATSU<sup>†</sup>

In this paper, aiming at supporting melody creation in improvisation, we propose a visualization method for melody data, an extraction method for rules of melody composition, and an exploring method for hints to create novel melody. In the visualization method, relationships among phrases and musical primitives are represented as a spatial structure. By analyzing the space, a user can obtain various rules and hints for melody composition. We also show experiments with the prototype system which was applied to several sample performances. From the results, it is suggested that the methods effectively support people to extract the rules and the hints through the static as well as dynamic analysis of improvised melodies.

### 1. はじめに

筆者らは現在、音楽における人間の創造性を計算機によって触発し支援する技術の研究を進めている。筆者らが目指しているものは、機械による創造ではない、芸術などに代表される創造的な行為において、その行為主体はあくまで人間であり、機械はそのための道具であるべきであるというのが筆者らの立場である。しかしながら、それらの道具はより使いやすく、より賢い、気の利いたものであることが望ましいのは当然であろう。本論文で示す旋律創作支援手法の研究は、このような方向性に沿ってなされている。したがって本研究は、自動作曲や各種演奏ルールの自動獲得などの研究とは、共通する要素は多数存在するが、その目指すところが大きく異なる。

本論文で示す手法は、特にジャズなどでよく見られ

る即興演奏における旋律創作を対象とし、将来的には一般的な作曲における旋律創作をも処理対象として視野に含めた、旋律創作過程の支援を目標としている。取り扱う音楽の形式は、一般的な西洋軽音楽に見られる、ある固定のコード進行の上で基本的に単一の主旋律（以下、旋律と略す）が演奏される形式の音楽である。また取り扱う旋律は、単音旋律とする。なお、旋律の創作にあたって考慮すべき要素はきわめて多岐にわたり、しかもそれらは相互に影響を及ぼしあうため、そのすべてを取り扱う手法を構築することは容易ではない。そこで本論文では、旋律を構成する音の音楽的機能（2章参照）の組合せと、それらの音の音高の時系列的な動き（3.1.1項参照）のみを処理対象とする。すなわち本手法では、まずこれらの要因に関連するいくつかの音楽プリミティブを定義する。そして、旋律をいくつかのフレーズに分解し、各フレーズが持つ音楽プリミティブとそれらの重みを求める。こうして得られたフレーズ・重み付き音楽プリミティブ集合に対して統計処理を施すことにより、各フレーズおよび音

<sup>†</sup> 株式会社エイ・ティ・アール知能映像通信研究所  
ATR Media Integration & Communications Research  
Laboratories

楽プリミティブの相互関係を2次元の空間的位置関係として可視化し、利用者に提示する。

したがって、本手法はまず単純には単音旋律データの分析手法あるいは可視化手法である。さらに、提示された空間を観察し分析することにより、旋律創作のために有益なルールの抽出や、さらにはまったく新しい旋律創作のためのヒントの獲得を、利用者が効率的に行うことができるようにすることを目指している。ここで、旋律創作のためのルールとは、ある旋律を「良い」旋律たらしめている $\star$ 、音の機能および音高の時系列的動きの組合せと関連した、旋律データ中にそもそも内在している要因のことで定義する。たとえば、対象の旋律で用いられている、ある2種類の機能の音の組合せなどが、その要因となる。また、旋律創作のためのヒントとは、ある旋律自体には含まれていない（したがって、分析結果にも陽には現れない）が、その分析結果から間接的に示唆される、新しい旋律を作るための手がかり情報のことと定義する。たとえば、自作の旋律を分析してみた結果から見てとれる、その旋律には含まれていない、まだ試されていない音の機能の組合せなどである $\star\star$ 。旋律データの可視化表現によって、これらのルールやヒントの抽出は、元の旋律データそのままを見て（聞いて）判断するよりも容易となるだろう。このように本手法は、旋律創作における人間の知識獲得支援ならびに発想支援のための手法でもある。また、この手法を複数の演奏者によるかけあい即興演奏にリアルタイムに適用し、バンド全体による演奏における旋律創作ルールおよびヒントの獲得を支援することも考慮している。これにより、各演奏者が創作する旋律の特徴やそれらの相互関係を音以外の視覚的モーダルで伝えあうことになる。この意味で本手法は音楽における CSCW システムとしての性格もあわせ持つ。

筆者らは、すでに同様の分析・可視化手法を自然言語による対話に適用したシステムを開発しており、その試用経験から、この手法が、単なる分析結果の提示を超えた、発想の触発効果を持つことを確認している<sup>14)</sup>。本研究は、この自然言語対話に対する発想支援

の手法を音楽に適用することにより、音楽における創造性を支援しようとする試みである<sup>11)</sup>。

本論文は、以下の章で構成される。2章では、音の機能分析について概説する。3章では、統計手法を応用し、フレーズと音楽プリミティブの相互関係を空間構造化することによる旋律データの可視化手法について説明する。さらに、この手法によって可視化された空間から読み取れる情報について考察する。4章では、前2章で説明した手法に基づき構築したプロトタイプシステムについて説明する。5章では、プロトタイプシステムを用いての適用例と分析例を示し、本手法の有効性について考察する。6章では、関連研究について概観する。7章はまとめである。

## 2. 音の機能に基づく分析

本章では音の機能とその求め方について簡単に述べる。詳細は文献12)を参照されたい。

音の機能とは、なんらかの音楽的環境  $E_i$  のもとで、ある音  $s_j$  が人に特定の情動的影響を与える作用のことである。なお、 $s_1$  と  $s_2$ 、あるいは  $E_1$  と  $E_2$  とを異なった音ないし音楽的環境とする要因としては、本論文では音高のみを考慮する。ごく単純な例を示せば、Cの音が鳴っている環境では、Cに対して完全5度の音程となるGの音は「完全協和」という安定感を感じさせる機能を持ち、Cに対して長2度音程となるDの音は、「不協和」という不安定感を感じさせる機能を持つ。

ただし、ある音楽的環境のもとでのある音の持つ機能は、時代や音楽の種類によって異なり、必ずしも一定ではない。そこで本研究では、処理対象に1940年代中期に始まったBe Bopと呼ばれる形式のジャズ（以下、単に「ジャズ」と呼ぶ）を取りあげ、楽曲のコードとコード進行のみを、その楽曲の音楽的環境とする。本研究では、Berklee理論<sup>18)</sup>と呼ばれる音楽理論を応用して、音の機能分析を行う。Berklee理論は、そもそもBe Bopスタイルのジャズに基づいて構築されたが、ジャズ以外の多くの現代西洋軽音楽もこの理論を基礎としている、重要な理論である。

ここでは、Berklee理論による分析を、きわめて簡略に説明する。まず、楽曲中のドミナント7thのコードに注目し、そのコードの根音の完全5度下の音をその箇所周辺の主音であると仮定し、その周辺の他のコードと照合して調性（主音および長調/短調）を決定する。調性が決定されたら、次は各コードをその部分の主音に対する相対音高表記で書き直す。ついで、Berklee理論で経験的に与えられている、相対音高表

$\star$  なお、旋律の良し悪しの判断を下すのは、あくまで人間である。本手法が何らかの基準情報を持っていて、自動的に良し悪しの判断を下すのではない。

$\star\star$  当然、そのような要因は非常に多数存在し、そのすべてが有効なヒントとなるわけではない。実際にはごく一部の要因が有効なヒントとなる可能性を持つにすぎないであろう。また、このような要因の有効/無効のふい分けをするのは、ルール抽出の場合と同様にあくまで人間であり、本手法が自動的にそのふい分けを行うのではない。

記でのコードに対する使用可能な音列（これをアヴェイラブル・ノート・スケールと呼ぶ）データを参照し、各コードに対するアヴェイラブル・ノート・スケールを決定する。この結果、各スケールの根音から数えていくつめの音にあたるかによって、各音のその部分における機能を決定できる。各音の機能的名称は、根音を1度音とし、以下根音に近いものから順に上へ2度音、3度音…とする。

文献12)では、基本的にオクターブ中にある12の音全部が個々に異なる機能を持つものとして扱っているが、本研究では作品の特徴や旋律創作ルールの把握を容易にするために、機能の類似性に基づき、以下の5つの機能グループにまとめて扱う。

**無彩色コード音** 1度音と5度音。強い安定感を表現する機能を持つが、調性感（長調、短調の別などを感じさせる機能）に乏しい。

**有彩色コード音** 3度音と7度音。コードの機能（調性感や終止感など）を強く表現する機能を持つ。

**テンション音** アヴェイラブル・ノート・スケール中の音で、無彩色コード音、有彩色コード音、アヴォイド音のいずれでもない音。浮遊感や緊張感を表現する機能を持つ。

**アウト音** アヴェイラブル・ノート・スケールに含まれない音。本来のコードから外れた、強い浮遊感を表現する機能を持つ。

**アヴォイド音** Berklee理論において、アヴェイラブル・ノート・スケール中でAvoid Noteに指定される音。本来のコードが別のコードに変化したような印象を与える機能を持つ。

このように本手法では、音高を環境（コード進行）に対する相対的表現としての音の機能に記述し直す。この結果、異なる環境における異なる旋律が与える情動的作用を、相互に比較できるようになる。

### 3. フレーズの空間構造化による楽曲構造分析

本研究では、旋律はフレーズの時系列連鎖で構成され、各フレーズはその中に含まれる音楽プリミティブとその重みによって特徴づけられるものとして旋律をモデル化する。ここに、フレーズとは「文章の句読点に相当するような段落感で区切られるひとまとまりの楽句」<sup>1)</sup>のことであり、その長さは一般に不定で一概に規定できない。しかし、どこからどこまでを1つのフレーズと見なすかを機械的に判断することは、現状では困難である。そこで本研究では、便宜的に4小節を1つのフレーズとして扱うことにする。現代の一般的な西洋軽音楽では、これはさほど無理な設定ではない。

また、音楽プリミティブとは、演奏を構成するすべての基本的な音楽要素のことである（たとえば、コード、ハーモニー、ピッチ、音量、テンションなど）<sup>7)</sup>。

旋律データは、MIDI (Musical Instrument Digital Interface) データとして記述されたものを用いる。すなわち、Note On (発音)/Note Off (消音)、MIDI ノート・ナンバー (音高)、および velocity (音の強さ) によって旋律情報を記述する。また音長は、MIDI realtime message の Timing Clock の Tick 数で記述する。四分音符1つあたりの Tick 数は24である。現在の実装では、pitch bend や modulation などは考慮していない。また、1つの旋律中ではテンポは一定であるとするとする。

以下本章では、まず本研究で使用した音楽プリミティブについて説明する。ついで、フレーズ・重み付き音楽プリミティブ集合を用いた、旋律の可視化手法と、可視化された空間の利用法を説明する。

#### 3.1 音楽プリミティブ抽出

本研究で使用する音楽プリミティブは、輪郭属性と色彩属性とに分類される。以下、それぞれの属性に含まれる音楽プリミティブとその重み付け方法について説明する。なお、各プリミティブの重みは、音楽認知に関連する心理的重要度を表すものではなく、単に各フレーズを相互に差別化するための指標にすぎないことに注意されたい。

##### 3.1.1 輪郭属性

輪郭属性とは、演奏されるフレーズにおける音高の時系列的動きから得られる属性である。本研究では、以下の9種類のプリミティブを定義する。それぞれのプリミティブにおける音高、音長の扱い、および各プリミティブの重みの求め方もあわせて示す。

(1) 音数：1フレーズ中で演奏された音の数。個々の音の長さは不問。したがって、1フレーズに含まれる MIDI Note On メッセージの数が、本プリミティブの重みとなる。

(2) 平均音高：1フレーズ中のすべての音の音高の、各音の長さを考慮した加重平均値。ある音の長さは、その音の発音時刻 (MIDI Note On メッセージ発行時刻) から、次の音の発音時刻までの MIDI Timing Clock の Tick 数である。したがって、それら2つの音の間に実際には休符 (無音部) が存在した場合でも、その無音部は、その無音部直前の音が伸ばされている

\* 実際には、実装したシステムでは処理途中でテンポを変えることは可能である。しかし、本手法では音長は timing clock の tick 数で扱っているため、テンポの変更によって絶対時間での音長が変化しても、同じ tick 数の音は同じ音長として扱われる。

ものとして扱っている。なお、フレーズ冒頭が休符である場合は、その部分は平均音高の計算に加味しない。フレーズ全部が休符である場合の本プリミティブ重みは0とする。

(3) 音高分散：1フレーズ中のすべての音の音高の、各音の長さを考慮した分散。音の長さ、および無音部の扱いは平均音高の場合と同じ。

(4) 代表繰返し長：正規化された音高の推移の自己相関を計算した結果、最も強い自己相関係数が得られた周期。ここに音高の推移とは、MIDI Note Onメッセージのみを処理対象とし、どの音高の音がどの時点で発音されたかを示す、音高の時間推移曲線（実際には折れ線）のことである。ある音の音高  $p_i$  の正規化音高  $\bar{p}_i$  は、次式で与えられる。

$$\bar{p}_i = \frac{p_i - p_{av}}{\sigma_p}$$

ここに、 $p_{av}$ 、 $\sigma_p$  は、それぞれそのフレーズの平均音高とその標準偏差（音高分散の平方根）である。

(5) 極長周期脈動強さ：音高の推移を3次スプライン補間し、高速フーリエ変換を施して得たパワースペクトルの、2分音符長以上のパワーの積分。音高の推移の定義は、代表繰返し長の場合と同じ。

(6) 長周期脈動強さ：同じく4分音符以上2分音符未満のパワーの積分。

(7) 中周期脈動強さ：同じく8分音符以上4分音符未満のパワーの積分。

(8) 短周期脈動強さ：同じく16分音符以上8分音符未満のパワーの積分。

(9) 極短周期脈動強さ：同じく16分音符未満のパワーの積分。

### 3.1.2 色彩属性

色彩属性とは、フレーズ中の各音の音機能に基づく属性である。本研究では、2章で示した音の5つの機能グループを、そのまま5つのプリミティブとして扱う。各プリミティブの重みは、フレーズ中でのそれらの音の長さ（Note On から Note Off までの tick 数）と強さ（MIDI velocity 値）の積の総和とする。

### 3.2 フレーズ・音楽プリミティブの相互間系の可視化手法

上記のモデルに基づき、各フレーズ相互の関係を空間構造として可視化する。この可視化には、双対尺度法<sup>13)</sup>を応用した概念可視化手法<sup>14)</sup>を適用した。双対尺度法とは、複数の数量化属性で構成されたオブジェクト集合が与えられたときに、オブジェクト集合と属性集合にそれぞれ得点数量を与えることによって、オブジェクトどうしの属性共有性と属性どうしの共起性

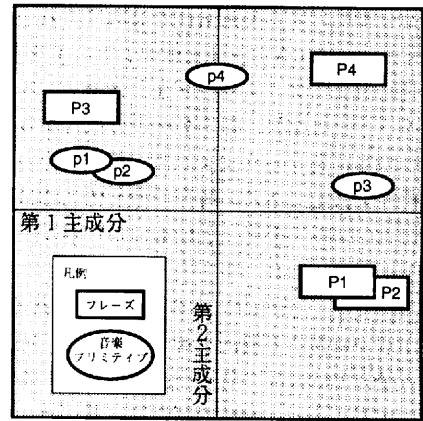


図1 可視化された空間の例

Fig. 1 A sample of visualized space.

を空間における相対的な位置関係として表現する手法である。本研究では、個々のフレーズをオブジェクト、個々の音楽プリミティブをオブジェクト属性、音楽プリミティブの重みを属性値として双対尺度法を適用した。これによって、各フレーズと各音楽プリミティブとの関連性が1つの空間上（本研究では2次元空間）に表現される。

得られる空間の、ごく単純な例を図1に示す。図中、長方形のアイコンがそれぞれ1つのフレーズに対応し、楕円形のアイコンがそれぞれ1つの音楽プリミティブに対応する。図の水平軸は第1主成分、垂直軸は第2主成分に対応する。座標の原点は、図の中心である。この図において、たとえばフレーズP1とP2は非常に接近して配置されている。このようなフレーズは、各音楽プリミティブを同じような重みで持っており、一般に類似したフレーズである。逆に、フレーズP1とP3は相互に離れて配置されている。したがって、この2つのフレーズは、相互にあまり類似していないフレーズであると見なせる。一方、音楽プリミティブについて見れば、接近して配置されているp1とp2は、1つのフレーズ中では同じような（一般には重い）重みを持ち、逆に離れて配置されているp1とp3は、1つのフレーズ中では一方の重みが重いときにはもう一方の重みは軽いという、排他的な重み関係を持つことを示している。また、p4はフレーズP3とP4の間に配置されているが、これはp4がP3とP4と同じような比較的重い重みで含まれていることを示している。

### 3.3 可視化された空間からのルールやヒントの獲得方法

本空間の分析によって得られる旋律創作のルールやヒントの例をあげる。たとえば、既存の名演奏を本手法で分析することにより、その名演奏を名演奏たらし

めている原因となっている音楽的特徴（ある2つの音楽プリミティブが同時に重い重みで使用されている、など）を探り当てることができる可能性がある。こうして得られた音楽的特徴（名手の旋律創作ルール）を応用することで、自分の演奏の質を向上することができるだろう。また、自分の演奏に対して本手法による分析を行うことで、自分の演奏の特徴（自分の旋律創作ルール）を見いだすことができるようになる。音楽は時間軸上の芸術であり、絵画などのようにじっと全体を眺めて分析することは難しいため、特に初心者の場合、録音した結果を繰り返し聞くことができる場合でも、自分の演奏の全体的な内容を自分で把握することは一般に難しい。そこで、本手法による分析を行うことにより、自分が見逃していたような演奏の特徴を発見するための一助となろう。一方、逆に自分の演奏に欠けている、それまでは気づいていなかったような旋律創作のヒントに気づかされる可能性もある（ある2つの音楽プリミティブを同時に重い重みで使ったことがないが、これらを同時に重い重みで使うことにより、新しい可能性が得られるのではないかなど）。また、空間上でフレーズアイコンが配置されていない空白の領域があった場合、そのような領域には、自分の思いつかないような、新しいフレーズ作りのアイデアが潜んでいる可能性がある。具体的には、空間の基底や、その空白を取り巻く音楽プリミティブやフレーズを考慮することによって、空白に含まれるようなフレーズのヒントを得ることができるだろう。さらに、ジャズなどでよく行われる、複数の演奏者によるかけあい即興演奏を分析することにより、お互いの演奏者が演奏するフレーズが、相互にどのような関係を持つかをひと目で見てとることができるようになる。これは、音楽作品としての演奏の全体的構造を意図的に構築するために有効な情報となろう。

#### 4. 実験システムの構成

##### 4.1 システムの概要

図2に実験システムの構成の概略を示す。本システムでは、クライアントシステムを演奏者の数だけ準備することにより、複数の演奏者による旋律データのリアルタイム入力にも対応可能としている。この図では、2人の演奏者が旋律データをリアルタイム入力する場合の構成を示しているが、MIDI楽器とクライアントシステムを追加することで3人以上の演奏にも対応できる。以下、図2の構成のシステムを用いて、2人の演奏者が4バース形式の掛け合い即興演奏<sup>☆</sup>を行う場合の処理の流れの概略を説明する。もちろんクラ

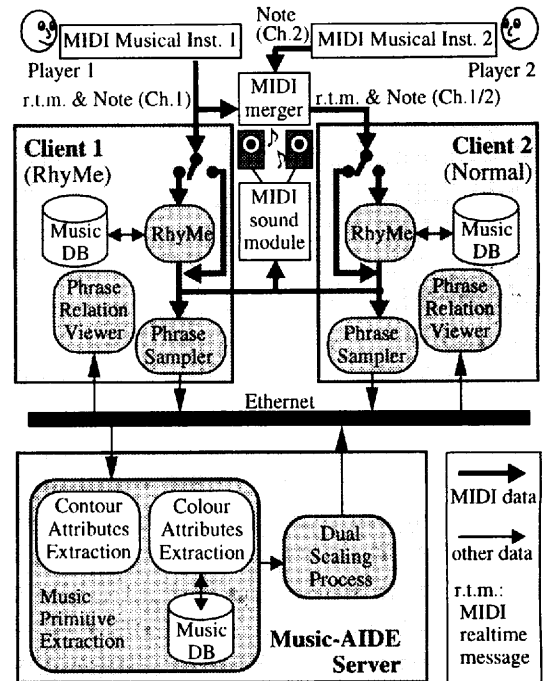


図2 実験システムの構成図

Fig. 2 Structure of prototype system.

イアントを1つだけ用いれば、1人の即興演奏を処理することや、あらかじめ打ち込んでおいたデータを静的状況で入力して解析することも可能である。

まず、演奏開始に先だって、処理対象とする楽曲のコード進行を2章に示した手法に基づき分析し、各コードにおけるアヴェイラブル・ノート・スケールを求め<sup>☆☆</sup>、得られたデータをクライアントおよびサーバのMusic DBに蓄積しておく。演奏開始時に曲名を指定することにより、当該曲目のアヴェイラブル・ノート・スケール・データがMusic DBから取り出され、サーバのMusic Primitive ExtractionモジュールならびにクライアントのRhyMeモジュールに与えられる。さらに、全演奏者数および各クライアントの演奏者が何番目に演奏するかを設定する。本実験例では、Client 1の演奏者が先に、Client 2の演奏者が後に演奏をするものとする。また、各クライアントに接続されているMIDI楽器からの演奏データは、それぞれ異なるMIDIチャンネルに出力されるように設定する。この図の場合、Client 1の楽器からの演奏データはチャンネル1（図中Ch.1）に、Client 2の楽器からの演奏

<sup>☆</sup> 4バースとは、演奏者が4小節ずつ交代して短い即興演奏を連続して行う演奏形式である。

<sup>☆☆</sup> 簡単なコード進行の場合には、この処理を自動的に行えるツールも用意している。

データはチャンネル 2 (図中 Ch.2) に出力される。システム全体の同期を制御する MIDI realtime message (図中 r.t.m.), すなわち演奏開始・終了信号, および MIDI Timing Clock は, すべてある単一の Client (本実装例の場合は Client 1) の MIDI 楽器のみが発信するように設定する。MIDI merger は, 入力されたすべての MIDI データをマージして 1 つの MIDI data のストリームとして出力する。したがって, 本実装例の場合, MIDI merger からの出力は, Client 1 の楽器からの MIDI realtime message とチャンネル 1 の演奏データ, および Client 2 の楽器からのチャンネル 2 の演奏データが, 1 つのストリームとしてまとめられたものとなり, これが Client 2 に入力される。各クライアントは, 指定されたチャンネル上の演奏データと MIDI realtime message のみを用いてそれぞれの処理を進める。本実装例の場合, Client 1 はチャンネル 1 の演奏データと MIDI realtime message を用い, Client 2 は, チャンネル 1 の演奏データも受信するがこれは使用せず, チャンネル 2 の演奏データのみと MIDI realtime message を用いる。

Client 1 の楽器から演奏開始信号が発信されると, 全クライアントは MIDI Timing Clock 信号に同期して処理を進める<sup>\*</sup>。MIDI sound module は演奏開始信号によって自動伴奏を開始する。現在の実装では, MIDI sound module には KORG 社製キーボード i1 を使用しており, 自動伴奏は i1 が内蔵するバックシンセシス機能を利用している。

前奏が終了すると, 第 1 番目の演奏者が演奏を開始する。第 1 番目の演奏者の演奏は, Client 1 の Phrase Sampler モジュールで所定期間 (今の場合は 4 小節) だけ旋律データとしてサンプリングされる。所定期間が終了すると, Phrase Sampler モジュールは採取した旋律データをサーバに送信する。同時に, Client 2 に処理権が渡され, 第 2 の演奏者の演奏とそのサンプリングが開始される。いずれのクライアントにおいても, Phrase Sampler モジュールに入力される旋律データは, 同時に MIDI sound module にも入力され, 実際の音として出力される。なお, クライアントには初心者の即興演奏を支援するための RhyMe<sup>12)</sup> と呼ぶ機能が備えられている。この機能を使用する場合は, RhyMe モジュールから出力される旋律データをサンプリングする。

サーバでは, まず Music Primitive Extraction モジュールで, 各クライアントから送られてきた旋律データから重み付き音楽プリミティブを抽出する。得られた重み付き音楽プリミティブに基づき, Dual Scaling Process モジュールが個々のフレーズおよび音楽プリミティブの関連性を 3 章で示した方法によって 2 次元空間構造として表現する。この結果が各クライアントに送信され, Phrase Relation Viewer モジュールが空間構造を表示し, 演奏者に提示する。したがって, 新たなフレーズが投入されるたびに, 双対尺度法に基づく空間の基底計算とそれに基づく各フレーズおよび音楽プリミティブの再配置が行われる。このため, 空間構造は新フレーズ投入のたびに一般に大きく変化する。空間構造の再計算と再表示には, 若干の遅延が生じる。遅延の量は, 処理すべきフレーズ数が増えるに従って大きくなるが, 5 章で示す 20 フレーズ程度の処理の場合, 最大で 2 秒程度の遅延となる。

利用者に提示される空間上で, フレーズは長方形のアイコンで表示される。各アイコンには, そのフレーズを入力した演奏者の名前と, その演奏者による何番目のフレーズかを示す通し番号が表示される。また, アイコンの色は演奏者によって異なるものことができる。これは視認性を向上させるためである。また各音楽プリミティブには, 図 3 に示す図柄のアイコンを採用した。これは, 後の実験でも示すように, この空間分析において音楽プリミティブの配置の把握がとりわけ重要となるので, 特にリアルタイム演奏に本システムを適用する場合における視認性を向上させるためである。実際に図柄アイコンの採用により, 少し慣れれば演奏中にでも提示される空間構造の把握は十分可能となった。なお図柄は, 各音楽プリミティブの特徴をよく表すように, 筆者がデザインした。

#### 4.2 即興演奏支援機能 RhyMe

RhyMe モジュールは, 与えられたアヴェイラブル・

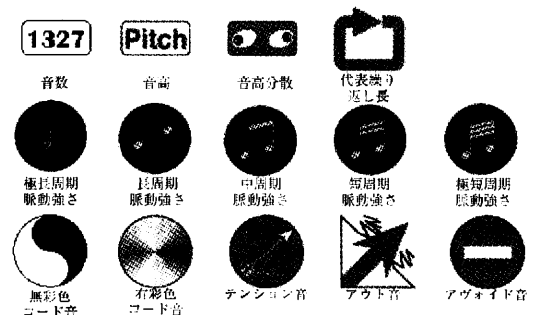


図 3 各音楽プリミティブのアイコン

Fig. 3 Icons of the musical primitives.

<sup>\*</sup> 本実装例では, システム利用者が音として出力される演奏を聞くことによって, 全体の演奏に同期して楽器演奏を行うことを想定している。このため, Client 2 の楽器には MIDI realtime message を供給していない。

ノート・スケール・データに基づき、本機能使用クライアント付属の楽器の演奏インタフェースを「音機能固定マッピング」の考え方に基づく音配列にする。従来の一般的楽器では、ある一定の演奏ポジションからは、つねに一定の音高の音が出力される。たとえば、ピアノの鍵盤の「ド」のキーを押せば、つねにCの音が出る。これに対し、演奏インタフェースの一定のポジションにはつねに一定の機能を持つ音をマッピングするのが音機能固定マッピングである。この手法を用いることにより、初心者でも比較的容易に任意の色彩感を持つフレーズを演奏することが可能となる。

## 5. 実験と考察

プロトタイプシステムを用いていくつかの適用実験を行った。ここでは、既存の演奏データを用いた静的な状況での適用例と、2人の演奏者による即興演奏における実時間状況での適用例を示し、それぞれの例について、どのようなルールやヒントを可視化された空間から読み取れるかを示す。

### 5.1 静的状況での適用例

本実験では、既存の名演奏を分析することにより、良い旋律を創作するためのルールを抽出した例を示す。使用した演奏事例は、アルバム *Somethin' Else* に収録されている、ジャズのスタンダードナンバー「枯葉」の演奏における、Miles Davis によるトランペットのソロ即興演奏と、Julian “Cannonball” Adderley によるアルト・サックスのソロ即興演奏である。旋律データは、文献3)にある書き起こされた楽譜に基づき作成した。このため、実際の演奏における表情づけなどは、完全には再現されていない。また、velocity はすべての音について一定値(100)とした。

図4にMiles Davisによる64小節(16フレーズ)の演奏を2次元空間構造化したものを、図5にCannonballによる64小節(16フレーズ)の演奏を2次元空間構造化したものを、また図6には、この両者によるフレーズ全部によって1つの空間を構築したものを示す。

これらの図から、図4では第1主成分は主に無彩色コード音、テンション音、有彩色コード音の3つのプリミティブで特徴づけられ、第2主成分はアウト音とわずかに無彩色コード音で特徴づけられていることが即座に読み取れる。したがって、Miles Davisのこの枯葉の演奏におけるフレーズ作りの特徴は、コード音とテンション音という、基本的で理論に沿った音の使い方にあることが分かる。この結果をジャズギターを勉強している者に見せたところ、特に有彩色コード音

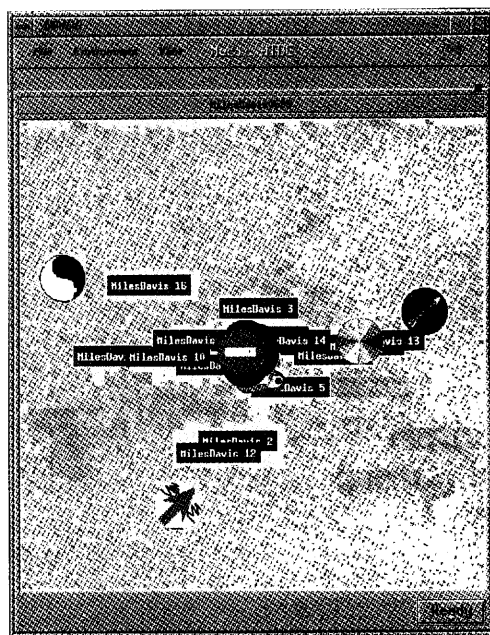


図4 Miles Davisによる枯葉の即興演奏に基づく空間構造  
Fig. 4 Spatial structure based on Miles Davis's improvisation in "Autumn Leaves."

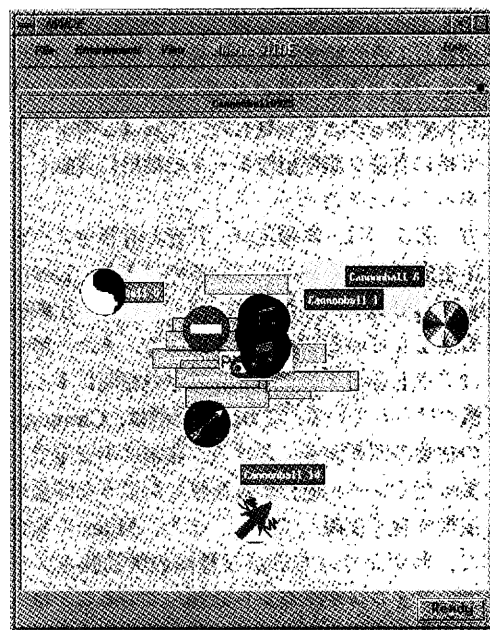


図5 Cannonballによる枯葉の即興演奏に基づく空間構造  
Fig. 5 Spatial structure based on Cannonball's improvisation in "Autumn Leaves."

と無彩色コード音の負の相関(水平方向の両端にアイコンが配置されていること)に興味を示した。従来そのような音使いは聞いたことも考えたこともなかったという。そして、即座にいくつかのフレーズを試しに

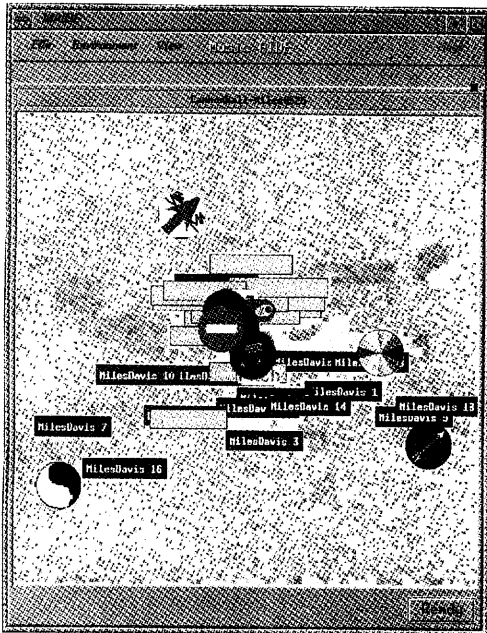


図6 Miles DavisとCannonballの両者による枯葉の即興演奏に基づく空間構造

Fig. 6 Spatial structure based on Miles Davis's and Cannonball's improvisation in "Autumn Leaves."

即興演奏してみた結果、特に無彩色コード音の5度音を有彩色コード音と同時に使用しない方が「イキナ」フレーズとなることを発見した。これはこのような分析が直接に有益な旋律創作ルールの抽出に貢献した好例であるといえるだろう。

一方、図5では、無彩色コード音と有彩色コード音が第1主成分に強く寄与しているのはMiles Davisの場合と同じであるが、さらにアヴォイド音も比較的大きな貢献をしていることが読み取れる。また、第2主成分はアウト音とテンション音の正の相関によって主に特徴づけられている。これらの特徴は、Cannonballのこの枯葉の即興演奏における、非常に不安定で、時にはまったくコード進行から逸脱したような印象を与える演奏を良く反映している。また、Milesの場合に比べて、わずかではあるが5種類の脈動強さプリミティブのアイコンが原点から分散していることは、ゆるやかな動きをするフレーズや速い動きをするフレーズなどが様々に入り混じっていることが示されている。これらの結果から、Cannonballの旋律創作ルールの一端が把握できる。

しかしながら、これら2つの図は、それぞれ水平軸、垂直軸が異なるプリミティブで特徴づけられ、かつその尺度も異なるため、これらをバラバラに並べて比較しても両者の演奏の違いは曖昧にしか見えてこない。

そこで両者の演奏を1つの空間上に表現することにより、両者の差を浮彫りにする。図6において、第1主成分は無彩色コード音、テンション音、有彩色コード音に特徴づけられており、第2主成分はアウト音とアヴォイド音、および一部の輪郭属性プリミティブに特徴づけられている。そして、明確にMilesの演奏は水平方向に分布し、逆にCannonballの演奏は垂直方向、特に中央より上方に分布している。この結果から、Milesは理論に沿った素直な音使いによる演奏の特徴づけをしているのに対し、Cannonballは緊張感の高い音によるフレーズ作りをしていることが、より明らかとなる。

以上のような分析から、Miles風のフレーズ作り、あるいはCannonball風のフレーズ作りをするにはどのような音使いをすればよいかという旋律創作におけるルールの一端を抽出することが可能となる。

## 5.2 実時間状況での適用例

本実験では、複数の演奏者によるリアルタイムな演奏にこのシステムを適用することにより、自分(たち)自身の旋律創作ルールの抽出するとともに、さらにそこから新たな旋律創作のためのヒントを獲得できる可能性について示す。

実験では、試作したシステムを用いて、2人の演奏者による4バース形式の即興演奏実験を行った。各演奏者は、それぞれのクライアントシステムのコンソールに表示される、フレーズの投入にともなって逐次構築更新される空間構造を見ながら演奏を行う。演奏者の1人は、ジャズ・ピアノの経験がかなり豊富である。もう1人は筆者の1人であり、ジャズ理論はひとつと理解しているが演奏経験は浅く、また鍵盤楽器の演奏にはあまり習熟していない。そこで、前者はNormal Clientを使用し、後者はRhyMeクライアントを使用した。演奏曲目はAll the things you areである。この曲は、1コーラス36小節(9フレーズ)で構成されている。演奏実験は、1回2コーラスを全部で8回行った。

図7に、ある回の演奏における2コーラス(18フレーズ)演奏終了時のPhrase Relation Viewerの画面を示す。図中、色の薄い方のアイコンがNormal Client使用者の演奏、色の濃いアイコンがRhyMe Client使用者の演奏である。この例において、第1主成分を特徴づけるものは、アウト音、有彩色コード音、アヴォイド音であった。また、第2主成分を特徴づけるものは、代表繰返し長さとテンション音であった。

この図から読み取れるのは、RhyMeクライアント側の演奏者は、コード音を用いてコード進行に沿った



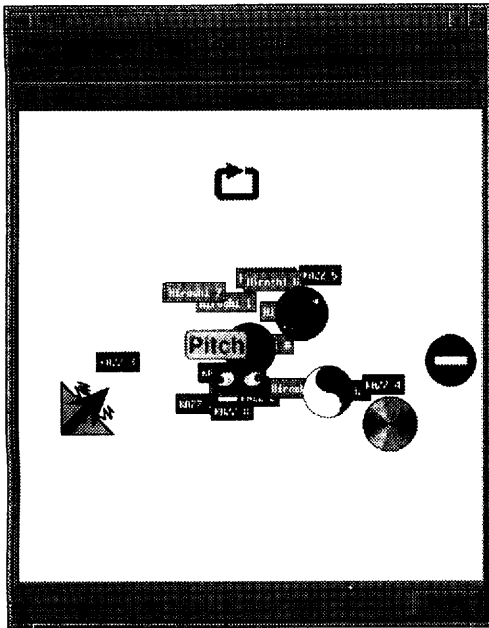


図7 2コーラス演奏終了時の空間構造

Fig. 7 Spatial structure after 2-chorus performance.

色彩感を出そうとしてはいるが、しかしアヴォイド音も非常に多用する場合があります<sup>\*</sup>、その分はずれ感の非常に強いフレーズも作っているということである。また、先の Miles や Cannonball の場合と違い、有彩色コード音と無彩色コード音の正の相関が強いというのも特徴である。一方、Normal Client 側の演奏者はアヴォイド音やアウト音をほとんど使用していない。また、明瞭な繰返しパターンが含まれるフレーズが存在することも示されている。

演奏実験終了後、Normal Client 側被験者に対しインタビューを行った。得られた感想を列挙する。

- (1) 演奏の特徴がよく分かる。構成された空間を観察することによって「客観的視点」からのフィードバックが得られ、有益である。
- (2) 空間の軸がどんどん変化して意味を考える余裕がないため、演奏中に実時間的にこの空間を見て意味を理解することはやや難しい。
- (3) 特定の音楽プリミティブを意識して配置結果を予想しながら演奏しても、予想とまったく異なる配置になることがある。

第1の感想は、本研究が目指す重要な効果の1つであり、その有効性が示唆されたものといえる。これにより、たとえば、空間構造を参照しながら、相手と類

似したパターンのフレーズ作りをして、全体に同じような色合いを持つ演奏にしよう、あるいは逆にまったく異なるフレーズ作りをして、対照的な色合いを持つ演奏にしよう、というような演奏の全体構造を設計できる。しかもそのためにはどのようなプリミティブに重点を置けばよいかに関する示唆を、この空間から得ることができる。

被験者の第2, 3の感想は、基本的に Normal Client 側の被験者がこの空間の見方に慣れていなかったことに起因している。しかし、この空間の読み方に慣れている RhyMe クライアント側の演奏者は、演奏中に空間構造を比較的容易に把握でき、それを演奏に反映させようとする事ができた。したがって、本システムは、実時間状況でも有効に活用することは可能と思われる。また第3の感想は、たとえば意識して大きな繰返しパターンがある演奏をしたとき、この被験者はその演奏のフレーズアイコンと、代表繰返し長さのアイコンの位置関係だけに興味を持ってしまっていた。しかし、本システムではつねに14種類すべての音楽プリミティブを考慮して空間を構成するため、演奏者が無意識に行ってしまった演奏特徴の影響により、予想していなかった思いがけない位置にその演奏のフレーズアイコンが配置される場合がある。しかし、これはまさに演奏者本人が見落としていた演奏の特徴のシステムによる顕在化であり、これも本手法の重要な効果の1つである。そのことを空間構造から読み取れるようになれば、つねに幅広い視野で自分の演奏を見直し、その結果を次の自分の演奏に反映させていくことが可能となるだろう。

### 5.3 問題点

現状のシステムの第1の問題は、輪郭属性の特徴抽出が現状ではうまくいっていないとはいいい難いことである。図4～図7を見れば分かるとおり、輪郭属性は主成分に対してほとんど寄与していない。現在の代表繰返し長や脈動強さプリミティブは、たとえば2小節連続的に音高が単調上昇し、その後2小節連続的に単調下降するような、非常に単純な音高の動きパターンを想定している。しかしながら、実際の演奏でそのような明確なパターンが出現することはきわめて稀であり、このような分析では個々のフレーズを差別化できるほどの違いを得ることができていない。

第2の問題は、旋律の創作には音高だけではなく、リズムも非常に重要な要素となる。しかし、現状リズムに関するデータをどのように取り扱えばよいのかについての解は得られていない。これは、より具体的な形で旋律創作に関するルールを獲得するために非常に

<sup>\*</sup> これは、現在の RhyMe ではアヴォイド音のポジションが浮動しており、演奏者はどこにアヴォイド音があるかを把握できないことによる。

重要な課題であるので、今後研究を進めたい。

## 6. 関連研究

本研究と関連する研究分野としては、音楽における発想支援・CSCW、自動作曲を目的とした機械学習による各種ルールや知識の自動獲得技術、音楽情報の可視化表現の研究などがあげられよう。

まず、音楽における発想支援、CSCW技術の関連研究について見る。後藤ら<sup>5)</sup>は、人間による音楽演奏の特徴に応じて様々に異なる振り付けの踊りを踊るCGダンサーを演奏者にリアルタイムに提示することによって、視覚的側面からの情報伝達を強化した、複数の演奏者による演奏システムを構築している。しかし、このシステムは音楽とCGを同時生成するマルチメディアアート作成装置としての性格が強く、音楽的に意味のある示唆をCGダンサーの踊りから得ることは難しいと思われる。かけあい即興演奏のcollaborationとしての側面に注目し、自然言語対話における話者交代(Turn-taking)の考え方を持ち込むことにより、適切なタイミングでかけあい演奏に参加してくる自動演奏システムも提案されている<sup>17)</sup>。しかし、このシステムは演奏者に対して積極的な旋律創作に関する示唆を与えるものではない。文献4)や市販ソフトKoan Pro V2<sup>20)</sup>は、いわゆるアルゴリズム作曲<sup>10)</sup>によって自動生成される旋律の「意外性」による、作曲に対する発想支援を試みている。これは、具体的な旋律を提示する点で、我々の手法と支援の方法が異なるが、今後本研究でも、たとえば利用者が指定した空間上の1点に対応する旋律を自動生成することによる、同様の支援手段を組み込みたい。また、発想支援とは多少異なるが、音楽に関する知識獲得を支援する教育システムとして、既存の音楽作品に対する解釈の仕方を、GUIを用いて指導するシステムも提案されている<sup>2)</sup>。この手法は、本論文で提案した手法などで作曲した旋律をさらに吟味するためなどに有効と思われる。

音楽に関する各種ルールや知識の自動獲得に関する研究のうち、旋律生成にかかわるルールを取り扱う研究としては、たとえば対位法による二声からなる楽曲の作曲における、対旋律作曲ルールの自動獲得システム<sup>19)</sup>や、作曲作業時の作曲家の振舞いをプロトコル解析することによって、作曲ルールを言語化することなく、非言語情報のままで獲得しようとする研究<sup>8)</sup>などがある。また梅本ら<sup>16)</sup>は、多数の楽曲データを分析し、音長パターン、コード進行、音高分布それぞれについての知識ベースを構築している。これらの研究では、ルールや知識を獲得するのは機械であり、獲得さ

れたルールや知識を利用するのもまた機械であり、それらのルールや知識を人間が利用することをあまり考慮していない点で本研究と異なる。

音楽情報の可視化表現にかかわる研究としては、筑波大学における計算機音楽プロジェクトPsycheにおける統合演奏視覚化システムがある<sup>6)</sup>。このシステムは、既存の楽曲(主にクラシックの作品)を様々な演奏家が演奏する際の演奏表現(音の速さのゆらぎや音の大きさ)を視覚化することを主目的としている。またTodd<sup>15)</sup>らは、画像処理のエッジ検出技術を適用することで、GTTM<sup>9)</sup>における楽譜データを用いた楽曲のグルーピング分析と同様の分析を音楽の音響データから行い、これを可視化して提示する手法を提案している。いずれも本研究とは視覚化する対象の情報が異なるが、今後総合的に演奏表現も含めた旋律創作演奏支援を行っていく場合に、これらの手法は有効と思われる。

## 7. おわりに

本論文では、特に即興演奏における新たな旋律の創造の支援を目的とした、旋律データの可視化手法と、それに基づく旋律創作ルール抽出と新しい旋律創作のためのヒントの獲得のための手法を提案した。本手法では、まず旋律を構成するフレーズおよび各音楽プリミティブの相互の関連を空間的な構造として表現する。この空間を利用者が分析し、たとえば各プリミティブの関係等を吟味することにより、様々な旋律創作のルールやヒントを得ることができる。さらに、構築したプロトタイプシステムを演奏例に適用した実験から、静的な状況での旋律分析や、実時間状況での即興演奏における演奏構成や新たなアイデアの考案に有効であることが示唆された。今後はさらにリズムなどに対する処理手法を考案し組み込むことにより、より実用的なシステムを構築したいと考えている。

謝辞 本研究の機会を与えてくださった(株)ATR 知能映像通信研究所の酒井保良会長に感謝いたします。また、双対尺度法処理ライブラリをご提供くださった学術情報センターの杉本雅則氏、空間構造の分析について有益なご示唆をくださった(株)ATR 知能映像通信研究所の角康之氏、実験にご協力くださった通商産業省工業技術院大阪工業技術研究所の渡邊洋氏、神戸大学大学院文化科学研究科の馬田一郎氏、システム実装にあたってご協力くださった豊橋技術科学大学の照井一由氏に感謝いたします。

## 参 考 文 献

- 1) 浅香 淳 (編) : 音楽中辞典, 音楽之友社 (1994).
- 2) Baker, M.J.: Design of an Intelligent Tutoring System for Musical Structure and Interpretation, *Understanding Music with AI*, pp.467-489, AAAI Press (1992).
- 3) 藤井英一 (編) : 完全コピージャズスコア, Vol.2, ヤマハミュージックメディア (1995).
- 4) Gerbrich, J., Randula, P. and Růžička, R.: Computer Program for Composition and Automatic Notation of Contemporary Instrumental and Vocal Music, *Computers and Artificial Intelligence*, Vol.10, No.5, pp.487-503 (1991).
- 5) 後藤真孝, 村岡洋一 : 音楽に踊らされる CG ダンサーによるインタラクティブパフォーマンス, コンピュータソフトウェア, Vol.14, No.3, pp.20-29 (1997).
- 6) 平賀瑠美, 五十嵐滋, 松浦陽平 : 統合演奏視覚化システム, 情報処理学会論文誌, Vol.38, No.11, pp.2391-2397 (1997).
- 7) 金森 務, 片寄晴弘, 新美康永, 平井 宏, 井口 征士 : ジャズセッションシステムのための音楽認識処理の一実現法, 情報処理学会論文誌, Vol.36, No.1, pp.139-152 (1995).
- 8) Laske, O.E.: The Observer Tradition of Knowledge Acquisition, *Understanding Music with AI*, pp.259-289, AAAI Press (1992).
- 9) Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, MIT Press (1983).
- 10) 長嶋洋一 : アルゴリズム作曲, *bit* 別冊コンピュータと音楽の世界, pp.402-418, 共立出版 (1998).
- 11) 西本一志, 角 康之, 間瀬健二 : 音楽創造性支援システム Music-AIDE, 計測自動制御学会システム工学部会第 19 回研究会 “発想支援ツール” 講演論文集, pp.1-8 (1997).
- 12) 西本一志, 渡邊 洋, 馬田一郎, 間瀬健二, 中津良平 : 創造的音楽表現を可能とする音楽演奏支援手法の検討—音機能固定マッピング楽器の提案, 情報処理学会論文誌, Vol.39, No.5, pp.1556-1567 (1998).
- 13) 西里静彦 : 質量データの数量化—双対尺度法とその応用, 朝倉出版 (1982).
- 14) 角 康之, 小川竜太, 堀 浩一, 大須賀節雄, 間瀬健二 : 思考空間の可視化によるコミュニケーション支援手法, 信学論, Vol.J79-A, No.2, pp.251-260 (1996).
- 15) Todd, N.P. and Brown, G.J.: Visualization of Rhythm, Time and Metre, *Artificial Intelligence Review*, Vol.10, No.3/4, pp.253-273 (1996).
- 16) 梅本あずさ, 内山幹乃扶, 河合敦夫, 椎野 努 : 音楽理論と経験的知識を整合活用した作曲支援システム, 情報処理学会研究報告音楽情報科学 23-11, pp.57-62 (1997).
- 17) Walker, W.F.: A Computer Participant in Musical Improvisation, *Proc. CHI97*, pp.123-130 (1997).
- 18) 渡辺貞夫 : *Jazz Study*, エー・ティ・エヌ (1996).
- 19) Widmer, G.: The Importance of Basic Musical Knowledge for Effective Learning, *Understanding Music with AI*, pp.467-489, AAAI Press (1992).
- 20) アイズ : *Koan Pro V2*.  
(平成 10 年 5 月 29 日受付)  
(平成 10 年 12 月 7 日採録)

西本 一志 (正会員)



1987 年京都大学工学部工学研究科機械工学専攻修士課程修了。同年松下電器産業 (株) 入社。現在 (株) ATR 知能映像通信研究所第二研究室客員研究員。エージェントによる人の創造的活動の支援の研究に従事。人工知能学会, 言語処理学会, International Computer Music Association 各会員。博士 (工学)。

間瀬 健二 (正会員)



1979 年名古屋大学工学部電気工学科卒業。1981 年同大学院修士 (情報) 課程修了。同年日本電信電話公社 (現在 NTT) 入社。1988~1989 年米国 MIT メディア研究所客員研究員。1995 年より (株) ATR 知能映像通信研究所第二研究室長。工学博士。コンピュータグラフィックス, 画像処理とそのヒューマンインタフェース, コミュニケーション支援への応用が主な研究テーマ。IEEE, ACM, 電子情報通信学会各会員。

中津 良平 (正会員)



1971 年京都大学大学院工学研究科修士課程修了。同年日本電信電話公社 (現 NTT) 武蔵野電気通信研究所入所。現在 (株) ATR 知能映像通信研究所代表取締役社長。マルチメディア技術を応用した通信方式の研究などに従事。工学博士 (京大)。IEEE, 電子情報通信学会, 日本音響学会, 人工知能学会, 情報考古学会, 日本 VR 学会各会員。