

## テンション・パラメータを用いた協調型自動演奏システム： JASPER

和 気 早 苗<sup>†,\*</sup> 加 藤 博 一<sup>†</sup>  
才 脇 直 樹<sup>†</sup> 井 口 征 士<sup>†</sup>

人間と機械による合奏において、人間の即興演奏に合わせて出力が適宜変化する、協調型自動演奏システム“JASPER”(JAm Session PartnER)について述べる。本研究では合奏での人間・機械間のコミュニケーションを実現するために、計算機内に「聴く→感じる・考える→演奏する」という人間の即興演奏過程を模倣した演奏家モデルを構築した。また、テンション・パラメータという感覚量を定義し、それを用いることで、演奏に現れる感性情報を抽出し処理することを試みた。システムへの入力はMIDI楽器からの即興演奏である。その演奏から人間の演奏者のテンション値を計算する。これは人間の演奏における緊張感を表すものである。演奏者のテンション値を考慮し、システムの持つさまざまなルールによってシステムのテンション値が計算される。出力はドラムとベースの演奏であり、システムのテンション値に適合するパターンが複数の候補より選択される。実際にシステムとの合奏を行った演奏者からは「とても楽しく演奏できる」「計算機が自分の演奏にตอบสนองすることは非常に興味深い」といった評価が得られた。本研究は演奏に含まれる感性情報を計算機で扱おうとする試みの一つである。合奏という形でのコミュニケーションを研究しシステムを構築することで、人間・機械間における協調作業、そして感性情報を含むコミュニケーションの可能性を見いだすことができた。

### Cooperative Musical Partner System Using Tension- Parameter: JASPER (Jam Session Partner)

SANAE WAKE,<sup>†,\*</sup> HIROKAZU KATO,<sup>†</sup> NAOKI SAIWAKI<sup>†</sup> and SEIJI INOKUCHI<sup>†</sup>

This paper proposes JASPER (Jam Session Partner), a cooperative musical partner system. JASPER performs improvisation with a human musician. The purpose of this research is to realize communication of subjective information, such as human emotion, between the user and a computer. A performance model is constructed in the computer to simulate the human improvisation processes such as [listen→feel→perform]. Also introduced is a new concept we call Tension Parameter, which represents the realization of emotive information in performance. The performance style of JASPER is similar to that of a jam session. In a jam session, it is very easy for performers to signify their emotions or intentions and to communicate with each other through their performance. Performers are able to convey feelings of exuberance or depression in music. This system copes with such an emotive range through the Tension Parameter. In experiments, users indicated that they enjoyed a jam session with the system very much. Also, most of them were very interested in the possibilities of this system as a new way of using a computer. This system provides several possibilities of interactive operation and communication. Most significant, it allows the transfer of subjective information, such as human emotion, between the user and a computer.

#### 1. ま え が き

近年コンピュータはわれわれの生活に欠かせない存在となりつつある。その適用分野も多岐に渡るが、現在

のところあくまでも人間が何かを行うときの‘道具’として扱われている場合がほとんどである。しかし、コンピュータはそのような‘道具’としての役割のみでなく、人間の‘パートナー’ともなりうるものであろう。本研究の目的は人間同士で行うような協調作業を人間・コンピュータ間で行わせることである。協調作業の題材としては、音楽、特に合奏を取り上げることとした。従来開発されてきた楽器(道具)としてのコンピュータとは異なり、演奏者である人間の共演者

† 大阪大学基礎工学部システム工学科  
Department of Systems Engineering, Faculty of  
Engineering Science, Osaka University

\* 現在 NEC 関西 C&C 研究所  
Presently with Kansai C&C Research Laboratory,  
NEC Corporation

(パートナー)としてふるまう協調型自動演奏システムを構築した。

音楽を題材とした人間とコンピュータ間の協調型システムはこれまでもいくつか研究されている。Dannenberg<sup>1)</sup>や Vercoe<sup>2)</sup>は人間の演奏するメロディにコンピュータの演奏する伴奏をうまく追従させる自動演奏システムを開発した。また森田らはデータグローブを用いることによって自動演奏をコントロールすることのできる、指揮に従うコンピュータミュージックシステムを構築した<sup>3)</sup>。しかしこれらは協調型システムであっても、演奏者間における双方向の意志の伝達が人間とコンピュータの間で実現されているとはいえない。なぜなら、それらでは人間からコンピュータへのメッセージの送信とそれに対するコンピュータからの応答は実現されているが、コンピュータからのメッセージの送信とそれに対する人間の応答は実現されていないからである。

本論文では、ジャズやポップスといった“異なった楽器を担当する演奏者による分散協調演奏”の形態において、演奏者とコンピュータが相互に影響を与えながら演奏を進めていくシステムを提案する。つまり、指揮システムのようにタイミングやプログラムチェーン\*などの物理的パラメータを直接与えるのではなく、ユーザも一演奏者として感情や意志を表現し、コンピュータがそれを解釈してコンピュータ自身の演奏方法を変化させるものである。状況によってはコンピュータの方から演奏に変化を与えることも考慮する。このような研究例としては、人間とコンピュータによるかけあい演奏\*\*を可能としたニューロ・ミュージシャン<sup>4)</sup>があげられる。ニューロ・ミュージシャンではニューラルネットワークを利用した自己適応的モデルを採用しているが、本システムではシステム内部に演奏家の即興演奏モデルを明示的に構築する方法を採用している。さらに、テンション・パラメータという一次元パラメータを用い、演奏に含まれる感情表現のうち緊張感(落ち着いたとか盛り上がったというような言葉で表現される感覚量)という主観量を取り扱う。コンピュータはこのテンション・パラメータを用い、処理を進める。

また、これまでの自動演奏・自動伴奏システムはあらかじめ楽譜情報の与えられた楽曲を演奏するもので

あった。これに対して、今回提案するシステムにおいては即興演奏を取り上げているため、次に演奏される音やリズムをあらかじめ知ることはできない。本システムは、演奏を通じて人間とコンピュータがリアルタイムでコミュニケーションをとり、協調しながらこれまでにない曲を創造する、ということへのアプローチでもあり、自動演奏システムの新たな分野を開拓するものである。さらに、ここでのコミュニケーションにおいては論理的・客観的な情報を交換するものではなく、緊張感という主観的な感情表現をコミュニケーションすべき情報としている。

本研究の従来からの音楽情報処理研究との差異の1つに、研究目標を挙げることができる。つまり、従来研究の目標は音楽的質の追及や有効な道具(ツール)の開発であった。本研究では道具としてのコンピュータシステムではなく、人間のパートナーとしてのシステムを目指すものである。さらにその中で重要となるのは、利用者が楽しめるシステムいわば楽しさの追求であり、音楽的質の追及は副次的なものとして考えた。

ところで、楽しさや快適さというような主観的な要素を情報処理システムで取り扱うための研究はヒューマン・インタフェースの分野などで行われているものの、確立された強力な方法論は現在のところ存在しない。近年になり、このような主観的要素を含み持つ感性という枠組みを対象に、その情報学的研究が行われるようになった<sup>5)</sup>。それらの研究においては、人間に内在する感性を知能(知性)と対比させ、

知能=知識+推論

という枠組みに対し

感性=感性情報+感性情報処理

と対応付け、従来の人工知能研究、知識処理研究における方法論からの模索が行われている。これらの研究では、知識とは情報の人間の知性に働きかける側面であり、感性情報とは情報の人間の感性に働きかける側面であると定義している。また、感性情報処理とはこの感性情報を工学的に取り扱う、つまり処理することであると考えている。

そこで、本研究をこの感性情報処理の枠組みで見直すと、以下のように解釈することができる。“楽しさ”という主観的評価は、人間の知性より感性に重点のおかれた評価と考えられる。音楽活動において、楽しさを感じることができる感性情報として、演奏における感情表現の1つである緊張感を1つの軸として選択し

\* 電子楽器で音色などを切り替える情報。

\*\* 交互に演奏すること。ジャズなどでは2人以上の演奏者が4小節もしくは8小節といった単位ごとに交互に即興演奏を行うことをいう。

た。その情報を交換することでお互いの発信する緊張感の変化を時間変化と共に相互に合わせていくという協調作業の舞台を自動演奏システムという情報処理システムの上で考える。このとき、この協調作業が達成できれば、利用者はそこに“楽しさ”を感じるであろうと仮定し、それを本システムの目的としている。

## 2. 協調型自動演奏システム“JASPER”の概要

ここでは、人間とコンピュータによる緊張感のコミュニケーションを通しての即興演奏を可能とした協調型自動演奏システム“JASPER”(JAm Session Partner)の概要について述べる。本システムで取り上げた演奏形態は、ドラム、ベース、ピアノという3人によるジャム・セッションである。ジャム・セッションの特徴は演奏における制約が極端に少ないことであり、事前に決められているのは大まかなコード進行のみで、細部は演奏者の即興で構成される。このようなことから、ジャム・セッションは感情や意志を表現しやすく、また互いの音を通じての演奏によるコミュニケーションがとりやすい形式であるといえる。今回、コード進行はブルース進行\*とした。本システムにおいてはシステム側がドラムとベースを、人間はピアノを受け持つことにより、システム内の疑似ドラマー(ドラム奏者)、疑似ベーシスト(ベース奏者)、および人間のピアニストという3人によるジャム・セッションを試みる。システムの開発においては、音程を考慮にいれなくてもよいドラムを中心に研究を進めた。ベースについては音程という要素が新たに加わるため音楽的質の向上にはドラムとは異なる処理を行うべきだと思われるが、今回は同様な処理を用いた。

システムにはあらかじめ1小節単位の演奏パターンを複数持たせておき、1小節ごとにそれらの中から最も適当であると思われるパターンを選択し出力することにより演奏を変化させる。これは“即興演奏は、利用できる無限の選択肢の中から各自の特色をもって即興のレパートリーを選ぶことにより創作され、限られた時間内にある結果を生み出すためには数多くの手続きや方法が‘ただちに利用できる’状態にあることが必須である”という即興演奏についての Sloboda の仮説<sup>6)</sup>に基づくものである。実際はパターンの選択であっても数多くのパターンを持たせることでその組み

合わせは十分多数となり、人間の創作に近い演奏が可能であると思われる。特にドラムについては音程を考慮にいれる必要がないということや1小節内の大きな変化はほとんどおこらないという理由から、1小節単位のリズムパターン・データベースは有効である。

### 2.1 人間・システム間の相互関係

システムへの入力はい MIDI キーボードからの自由な演奏であり、出力は‘人間演奏者の演奏の緊張感を考慮したドラムとベースの演奏’である。この演奏は人間の耳を介し、逆に人間演奏者への入力となる。人間演奏者はそれに対して即興演奏を行い、これが再びシステムへの入力となる。このように出力される互いの音によりコミュニケーションを行い合奏を進める。

### 2.2 システム構成

システムは聴取部、思考部、演奏部という3つに分割することができる。

聴取部は人間の耳にあたる部分であり、MIDI キーボードの演奏データから音数、ペロシティ(強さ)、ピッチ(音高)を抽出する。

思考部は疑似演奏家の頭に該当し、演奏パターンを決定する。この思考部は“追従過程”と“自己主張過程”という2つの思考過程を持つ。これは合奏において、共演者の演奏にあわせようとする行為と、自分は今こう演奏したいという意志に従って自分の主張を表現するという2つの行為をモデル化したものである。

演奏部は人間の手足に相当し、思考部により決定されたパターンを実際の音として出力する。

### 2.3 テンション・パラメータ

人間は音楽を鑑賞することで感情に変化を受ける。その要因としては楽曲の持つメロディ、コード進行、歌詞などさまざまなものが考えられる。われわれは、これまでにコード進行による影響に注目して、感情表現をパラメータ化する研究を行ってきた<sup>7)</sup>。そこでは、曲の雰囲気表現する形容詞と具体的なコード進行を結び付ける‘感性パラメータ’を定義した。

今回題材とするジャム・セッションにおいては、8小節または12小節というような短い単位のコード進行が繰り返される。このような制限の中で、演奏者はさまざまな演奏技巧を用いて多面的な感情表現を行うが、その1つに緊張感という気分の盛り上がりや沈み込みを表す感覚量がある。本研究では、この演奏の緊張感をパラメータとして取り扱う。このパラメータをここでは‘テンション・パラメータ’と呼ぶ。

\* 12小節を1サイクルとした音楽のコード進行。ジャズやロックなどその応用範囲は広い。

## 2.4 処理のタイミング

音楽という時系列データをリアルタイムで取り扱う場合、その処理のタイミングや処理時間などの時間管理が非常に重要となる。本システムにおいてはテンポを一定とし、演奏パターンを1小節単位で切り替える。そこで、小節内の最後に処理時間を設けその時間内で次小節の演奏パターンを決定する。処理時間は短いほどよいが、今回は計算機の能力上 250 ms とした。

## 2.5 演奏パターン・データマップ

本システムは1小節単位の演奏パターンを利用する。すべてのパターンはそのビート(4ビート、8ビート等)によってグループ分けされ、おのおのにはテンション値の上限と下限が付けられている。図1(a)は、このようなテンション値を持つパターンをビートごとにグループ化し、縦軸にテンション値をとった平面上に並べたもので、演奏パターン・データマップと呼ぶ。図1(b)にドラム演奏パターンの例を示す。最終的に計算されたシステムのテンション値と各演奏パターンのテンション値のマッチングをとることにより演奏パターンを決定する。

## 3. 思考部 (テンション・パラメータを用いた処理)

前章でシステムの概要について述べたが、本章ではその思考部におけるテンション・パラメータを用いた処理方法を述べる。図2は演奏家モデル、つまり思考部における処理構成を示したものである。まず、聴取部によって得られた演奏データからその演奏に含まれる緊張感を表すヒューマン・テンション値を導き出す。このヒューマン・テンション値とシステム側の演奏に盛り込みたい緊張感を表すシステム・テンション値の双方を利用することで、次小節に出力する演奏パターンを決定する。

これらの処理では、現在の演奏データ以外に4種類のデータを利用する。ヒューマン・テンション抽出フィルタは、人間の演奏とその緊張感の対応を表すものである。個性データベースは、システムの演奏に個性を与えるルールの集合であり、自己主張過程において重要な役割を持つ。演奏パターン・データマップは2.5節で述べたものである。遷移データは、現在までの人間の演奏データ、ヒューマン・テンション値およびシステム・テンション値、演奏パターンなどの遷移

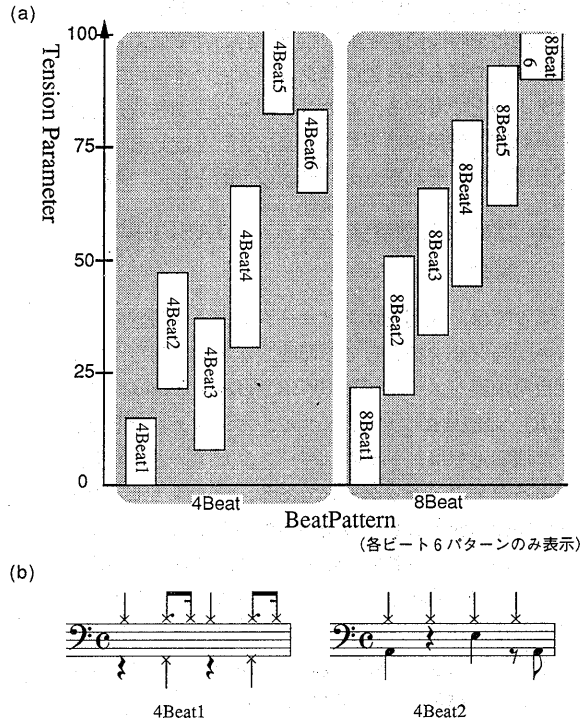


図1 (a) 演奏パターン・データマップ  
(b) ドラム演奏パターンの例  
Fig. 1 (a) Output Pattern Data-Map.  
(b) Drum Play Pattern.

状況が格納され、各処理において利用される。

また、2.2節で述べた追従過程と自己主張過程は、処理全体を通して随時実現されるものである。

### 3.1 ヒューマン・テンション値の計算

人間とシステムによる協調演奏の実現には、まず人間の演奏における緊張感、すなわちヒューマン・テンション値の計測を行う必要がある。ここで、音楽のような時間軸上の対象をリアルタイムで扱う場合、時間遅れは常に意識しなければならない問題である。2.4節で述べたように、本システムでは1小節を処理の単位としているので、現在のヒューマン・テンション値のみを用いて次小節の演奏を決めたのでは、緊張感の追従に少なくとも1小節の時間遅れが生じる。言いかえると、人間の演奏の現在の緊張感がわかって、それを用いて決定される演奏パターンが出力される頃にはすでにその緊張感に変化しているであろう。そこで、全体の楽曲構成の中での現在の演奏位置やこれまでの演奏の変化状況なども考慮に入れることによって、次小節での人間の演奏の緊張感を予測する。つまり、ヒューマン・テンション値の計算は、'現ヒュー

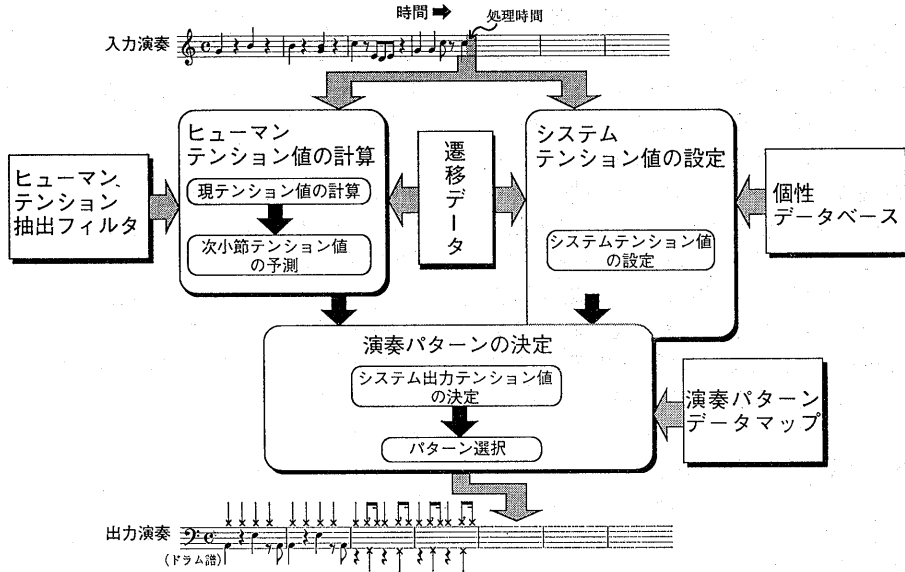


図 2 演奏家モデル  
Fig. 2 Performer model.

マン・テンション値の計算’と‘次小節ヒューマン・テンション値の予測’という2段階の処理で行う。

以下複数のパラメータが取り扱われるが、図 3 にテンション・パラメータを整理したものを示す。変数  $n$  は小節番号を指し、 $f(n)$  とは第  $n$  小節目の最後の処理時間内に計算される  $f$  の値である。関数間の矢印は、計算時の値の参照関係を示す。また、本章内で使用されるルールや数値のいくつかは、4 章でその詳細を述べることにする。

### 3.1.1 現ヒューマン・テンション値の計算

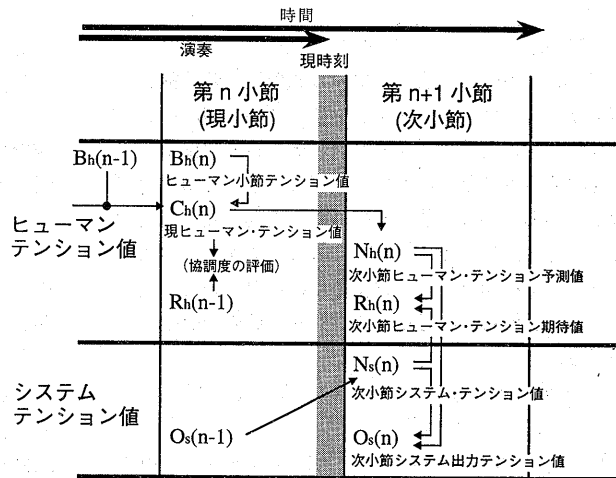
ここでは入力データから人間の演奏の現時点でのテンション値を導きだす。

#### (1) ヒューマン小節テンション値

音楽が演奏される中、システムは毎小節最後の処理時間において小節ごとのテンション値である第  $n$  小節テンション値  $B_h(n)$  を計算する。 $B_h(n)$  の導出には、

- 音数
- 平均ベロシティ
- 平均ピッチ

の 3 つの特徴量を用いる。音数とは 1 小節内に弾かれる音の総数である。ベロシティとは



- |        |   |
|--------|---|
| human  |   |
| $B_h$  | ... a bar ... 小節のテンション値   |
| $C_h$  | ... current ... 現在のテンション値   |
| $N_h$  | ... next ... 次小節のテンション予測値<br>: 次に相手(人間)はこう演奏するだろう                 |
| $R_h$  | ... require ... 次小節におけるテンション期待値<br>: 次に相手(人間)にはこう演奏してほしい          |
| system |   |
| $N_s$  | ... next ... システムのテンション値<br>: 次に自分(システム)はこう演奏したい                  |
| $O_s$  | ... output ... システムの演奏として表現されるテンション値<br>: 次に自分(システム)はこう演奏することにしよう |

図 3 テンション・パラメータ  
Fig. 3 Tension Parameter.

音が弾かれる強さであり、ピッチとは音の高さである。MIDI 規格ではベロシティは 0~127 の離散数値で、またピッチは C3 を 60 としてそれから半音上がるとに 1 を加えるという数値で表される。特徴量としてこれらの 1 小節内での平均値を用いる。まず、これら 3 つの数値をヒューマン・テンション抽出フィルタの中で定義される特徴量正規化関数によって正規化する。この関数はピアノ演奏者の演奏データとその時の演奏者の主観的な緊張感の関係を事前の演奏評価実験により求めたものである。変換後の数値は、大きいほど緊張感が高いことを示し、演奏者の主観的な観点において 0 近傍が静かな演奏、50 近傍が標準的な演奏、100 近傍が盛り上がった演奏に対応する。

これら 3 つの特徴量（音数、平均ベロシティ、平均ピッチ）の正規化後の各数値（ $n_r, n_v, n_p$ ）に要素ごとの重み（ $a_n, a_v, a_p$ ）を付け、式(1)によって小節ごとに  $B_h(n)$  を求める。

$$B_h(n) = (a_n \cdot n_r + a_v \cdot n_v + a_p \cdot n_p) / (a_n + a_v + a_p) \quad (1)$$

これら重みもヒューマン・テンション抽出フィルタの中に記述される。

### (2) 現ヒューマン・テンション値

$B_h(n)$  を用いて、現時点での人間の演奏の緊張感を表す現ヒューマン・テンション値  $C_h(n)$  を決める。 $B_h(n)$  がそのまま現ヒューマン・テンション値であるとも考えられるが、ここでは演奏における緊張感の変動は、小節という単位と比較し長い周期で変化を見せると仮定する。そこで式(2)のように過去 4 小節のヒューマン小節テンション値を用い平滑化を行うことにより、小節ごとの微小なテンション値の変動の除去を行う。

$$C_h(n) = (4 \cdot B_h(n) + 3 \cdot B_h(n-1) + 2 \cdot B_h(n-2) + B_h(n-3)) / 10 \quad (2)$$

### 3.1.2 次小節ヒューマン・テンション予測値の計算

ここでは、過去の演奏の遷移状況から人間の演奏の次小節テンション値を予測する。現在、次の 2 種類のルールによる予測を行っている。

#### (1) クレシェンド、デクレシェンドからの予測

過去 3 小節のベロシティの変化状況より、現在がクレシェンド傾向であるかデクレシェンド傾向であるかを判断する。式(3)におけるクレシェンド値  $t_c$  は現在のベロシティ傾向に応じて決まる値であり、現ヒューマン・テンション値に加算される。現在がクレ

シェンド傾向であれば、次の小節もクレシェンドが期待されるとしてクレシェンド値は正数を取り、逆にデクレシェンド傾向であれば負数をとる。

#### (2) 現在の位置情報からの予測

通常の音楽の場合 4 小節が 1 つのまとまりとなりやすい。この 4 小節のまとまりを楽節と定義する。楽節における 4 小節目は他よりも自然に盛り上がるという傾向があり、ドラム演奏においても多くの場合、'フィルイン' といわれる 1~3 小節目より派手なパターンが用いられる。そこで、楽節内 4 小節目のテンション値の予測においては位置値  $t_r$  を加える。

これら 2 種類のルールを用いて式(3)により次小節ヒューマン・テンション予測値  $N_h(n)$  を求める。

$$N_h(n) = C_h(n) + t_c + t_r \quad (3)$$

### 3.2 システム演奏パターンの決定

システムの出力は、人間の演奏の緊張感と個性データベースに基づいたシステム自身の緊張感を考慮に入れ決定される。個性データベースの役割は、システムの個性によって演奏に変化を与え人間の演奏者との双方向のコミュニケーションを実現することである。

#### 3.2.1 リードの判断

合奏においては多くの場合、演奏者の中のいずれか 1 名がソリストとなり、演奏をリードしていく。ソリストは自分の意志に基づいて演奏の流れを決めることができる。それに対し、バック（ソリスト以外の演奏者）はソリストのリードにあわせながら演奏を進める。本システムにおいても、このような演奏の基本姿勢をモデル化する。ジャムセッションのような自由度の大きい合奏の場合、誰がいつリードをとるかはあらかじめ決められてはいない。曲の流れと演奏者間のコミュニケーションを通して、リアルタイムに誰がリードをとるのが判断されなければならない。この判断を行うルールは、個性データベースに記述しておく。具体的なルールについては 4 章で述べる。

図 4 に演奏パターン決定までの処理の流れおよび演奏の状態を示す。以下、この流れに従って説明を行う。

#### 3.2.2 人間のリード（システムはバック）

リード判断において人間がリードをとる（システムはバック）と判断された場合の処理について述べる。

##### (1) 次小節システム・テンション値の設定

まず、システム自身の演奏に対する方針を決定するために、次小節でのシステム自身の演奏に盛り込みたい緊張感（次小節システム・テンション値  $N_s$ ）を設

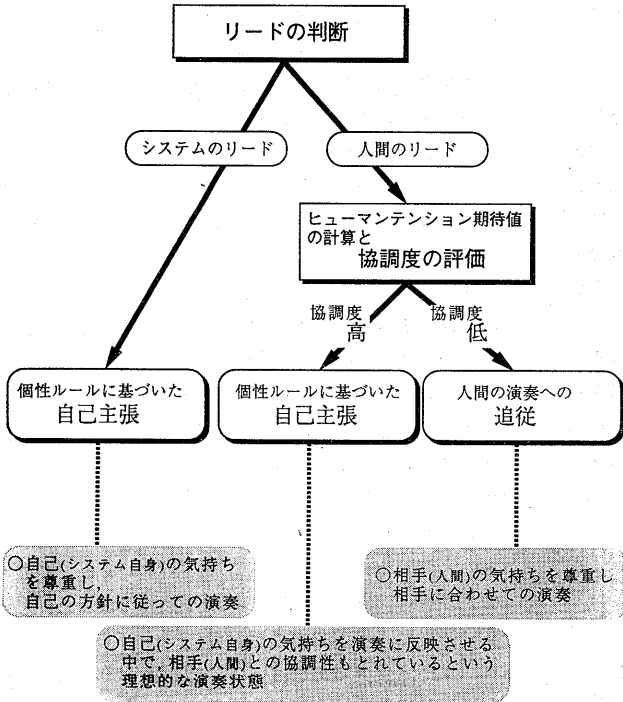


図 4 演奏の状態

Fig. 4 The styles of Ensemble.

定する。これは、システムの個性の一部であり、この判断のためのルールは個性データベースに記述される。現在、演奏における‘単調さの解消’を主な目的として、以下に示すルールを登録している。

○  $N_s$  は先行する  $O_s$  の値に応じて、式(4)のように設定する。

$$N_s(n) = O_s(n-1) + c \quad (4)$$

ただし、 $c$  は以下の条件で決定する。

- システム出力テンション値  $O_s$  が、 $n_1$  小節の間、一定値  $t_1$  より小さければ、 $P_{1A}(n_1)$  の確率で、 $c = c_{1A}$  ( $c_{1A}$  は定数)、とする。

- システム出力テンション値  $O_s$  が、 $n_2$  小節の間、一定値  $t_2$  より大きければ、 $P_{2A}(n_2)$  の確率で、 $c = c_{2A}$  ( $c_{2A}$  は定数)、とする。

- 上記2つがいずれも適用されなければ、 $c = t_p$  ( $t_p$  は楽節内の位置によって決まる値)、とする。

$O_s(n-1)$  は  $n-1$  小節目に導出された次小節システム出力テンション値で、現在の小節のシステムの演奏の緊張感を表す。 $N_s(n)$  は次小節でのシステム・テンション値である。この  $N_s(n)$  を決定するためのルールを追加・調整することによりシステムの個性を変えることができる。

(2) 次小節システム出力テンション値の決定

実際に演奏として表現される緊張感には、共演者（人間）の緊張感も関与してくる。今、人間がリードをとっているとき、システムは人間の演奏にあわせる必要があり、したがって人間の演奏の緊張感を十分考慮に入れなければならない。

〈協調度の評価〉

現時点において人間とシステムが協調できているかどうかを評価する部分である。この判断には、次小節ヒューマン・テンション期待値  $R_h$  を用いる。このパラメータは、次小節における人間の演奏の緊張感としてシステムが期待する値である。この値の算出は後述する。前小節の処理で求められた  $R_h(n-1)$ 、つまり現小節における人間の演奏に対し期待した値と実際の現小節ヒューマン・テンション値  $C_h(n)$  を比較する。この値が等しいということは、システムが人間に期待したものと実際の人間の反応が等しいということであり、協調できていると判断できる。そこで、次のルールを基に協調関係の存在を調べ、それぞれ以下の a) b) の処理を行う。

- $R_h(n-1)$  と  $C_h(n)$  の差が 10 以内であれば協調関係にあると見なす。

a) 自己主張に基づく協調演奏

協調できていると判断された場合、システムの立案した方針どおりの演奏を行うことにより、次小節においても協調演奏ができると考えられる。そこで、先に求めた  $N_s(n)$  を  $O_s(n)$  とする。また、 $R_h(n)$  も  $O_s(n)$  と同一の値に設定する。これは、本論文での協調作業が、お互いのテンションを等しく保ち演奏するというを意味しているからである。

b) 人間の演奏への追従

人間との協調ができていないと判断されたとき、システムは演奏を‘人間の演奏への追従’という形に修正する。このとき、 $O_s(n)$ 、 $R_h(n)$  を共に 3.1.2 項で予測した  $N_h(n)$  とする。

3.2.3 システムのリード

リードの判断において、システムがリードをとると判断された場合、システムはシステム自身の自己主張（個性データベースに記述）に基づいて演奏を進めて

いく。つまり、 $N_i(n)$  がそのまま  $O_i(n)$  と  $R_i(n)$  に代入される。

### 3.2.4 演奏パターンの選択

2.5 節で述べたように、システムはあらかじめ1小節単位のドラムパターンを持つ。これらのパターンにはそれぞれテンション値の上限と下限が付けられている。この値は、緊張感の程度をすべてのパターンについて一対比較法（心理実験）を用いて設定したものである。この演奏パターン・データマップを用いて最終的な演奏パターンが決定される。

$O_i(n)$  の値とデータマップ上のパターンを比較することによって対応のとれた演奏パターンが出力される。 $O_i(n)$  がデータマップ上で、どのパターンとも対応がとれなかった場合は、最短距離に位置するパターンを選択する。また逆に、対応のとれるパターンが複数あれば次のルールが適用される。

○楽節内最初の小節の演奏パターンを決めるとき ( $n \bmod 4 = 0$  のとき)

→乱数を用いてランダムに決定。

○上記以外の小節の演奏パターンを決めるとき

→現小節の演奏パターンを参照し、選択肢の中に同じものがあれば必ずそのパターンを選択する。なければランダムに決定。

これは基本的な演奏ルール「4小節はひとまとまりとなりやすく、このまとまり（楽節）の中では大きな変化はおこらない」を実現するものであり、自然な演奏の流れのために重要な役割を果たす。このようにして1小節ごとに次小節の演奏パターンが選択され、演奏部によって逐次出力される。

ベース・パターンに関しては、音程を考慮しなければならない。演奏パターンをコードを変更できる形式で管理し、選択された演奏パターンを出力される小節番号に対応するコードに変換した後に出力する。また、常にすべてのパターンを選択可能にしておいた場合、ベースライン（ベースの旋律）に不自然さが生じる。そこで、あらかじめ自然な演奏にするという評価で演奏パターンの遷移グラフを作成しておき、その遷移グラフを参照、次に遷移可能なパターンを限定し、その中でドラムの場合と同様な処理を行う。

以上の演奏パターンの選択処理に先立ち、ビート変更ルールが遷移データを基に、演奏パターン・データマップ中のどのビート・グループを用いるかを決定する。ただし、ビートの変更はどの小節においても即座にできるものでなく、特定小節からビート変更の導入

フレーズが始まり、続いてビートが変更するという流れをとるため、ビート変更の決定が行われる時刻と実際に変更が演奏に現われる時刻は異なる場合もある。このとき導入フレーズなどの特定の演奏パターンが利用されるが、それらもまた1つのグループをつくり、通常のパターンと同様に演奏パターン・データマップ上で管理されている。同様なものに演奏終了を決定するルールとその演奏パターンがある。このルールは、演奏終了のキーとなるフレーズを常に監視し、検出したときに演奏終了の意志決定を行うルールである。この場合もビート変更と同様に、終了決定の時刻と終了フレーズの開始時刻は一致しない。

ビート変更や演奏終了の決定に関する処理も、人間とシステムがコミュニケーションを通じて決定すべきだと考えられるが、このようなメッセージの交換はテンション・パラメータと同様な定式化はできない。実際の人間同士のジャムセッションにおいては、これらのメッセージは演奏以外に身振りなどによるところも多い。また、事前の打ち合わせによって決められたり、ビート変更であればドラム担当者にその意志判断がまかせられる場合も多い。そこで、本システムにおいては、ビート変更に関してはシステムの個性データベースの中にそのルールをシステムの個性として登録することにより、ドラム担当であるシステム側にその判断をまかせた。つまり、人間の演奏者からビート変更を直接的に指示できないようにした。演奏終了に関しては、実際の演奏に先立ち人間の演奏者から終了のキーとなるフレーズを入力してもらい、そのフレーズをそのまま個性データベースに登録するようにした。つまり、人間の演奏者の演奏終了の意志を直接システムに伝えることができるようにした。

## 4. システムの構築

3章で述べた考えを実際にシステムとして動作させるためには、多くの値を決定しなければならない。この章では、実験で使用した具体的なルールや数値の決定方法について述べる。

### 4.1 特徴量正規化関数の決定法

ピアノ演奏者に対し、「静かな演奏 (Low Tension)」、 「標準的な演奏 (Middle Tension)」、 「盛り上がった演奏 (High Tension)」の3種類の指定を行い、各48小節の演奏を記録する。1小節での音数、ベロシティの平均、ピッチの平均をその中央の24小節の演奏データから求める。表1に、分析結果の例を示す。この分



表 1 演奏の分析例

Table 1 An example of the analysis of performance.

Tension	音数	平均ベロシティ	平均ピッチ
Low Tension	9	47	58
Middle Tension	15	67	65
High Tension	19	74	70

析結果に対し、Low Tension の値が 0, Middle Tension の値が 50, High Tension の値が 100 となるような近似関数を求める (式 (5)). この近似関数を逆に利用することにより、演奏の特徴量をテンション・パラメータに変換できると考える。

$$\begin{aligned} n_r &= 0.35n^2 - 28.92 \\ v_r &= 0.07 \exp(v/10) - 7.20 \\ p_r &= 8.33p - 483.33 \end{aligned} \quad (5)$$

また、式 (1) の重み係数に関しては各特徴値のばらつきを考慮に入れ、経験的に設定した。表 1 の演奏データに対しては、 $a_n=3$ ,  $a_v=4$ ,  $a_p=1$  と設定した。

#### 4.2 $N_h(n)$ の予測式で用いられる値

式 (3) の  $N_h(n)$  の計算で用いられる  $t_c$ ,  $t_f$  の値は経験的に決定した。

$t_c$  の決定にあたって、現在がクレシェンド傾向かどうかの判断であるが、過去 3 小節の平均ベロシティの変化が単調増加であり、かつその変化量が 12 (MIDI におけるベロシティ値) 以上であれば、現在がクレシェンド傾向であるとした。逆に、過去 3 小節のベロシティの変化が単調減少で変化量が 12 以上であればデクレシェンド傾向であるとした。

$t_c$  および  $t_f$  の値を以下に示す。

クレシェンド値：

$$t_c = \begin{cases} 20 & \text{クレシェンド傾向のとき} \\ -20 & \text{デクレシェンド傾向のとき} \\ 0 & \text{otherwise} \end{cases}$$

位置値：

$$t_f = \begin{cases} 20 & \text{楽節内 4 小節目の } N_h \text{ 決定時} \\ & (n \bmod 4 = 3 \text{ のとき}) \\ 0 & \text{otherwise} \end{cases}$$

#### 4.3 リードの判断ルール

演奏の緊張感を基にリードを判断することを試みた。人間が一定以上の時間静かな演奏を続けた場合、人間はリードすることを放棄したと考える。また、人間が一定以上の時間盛り上がった演奏を続けた場合、人間がリードの譲渡を要求していると考え。そこで、現在次のようなルールを利用している。

○  $C_h(n)$  がブルース進行 2 サイクル (24 小節) の間、

30 より小さければ、以後システムのリードと見なす。

○  $C_h(n)$  がブルース進行 1 サイクル (12 小節) の間、50 より大きければ、以後人間のリードと見なす。

#### 4.4 $N_s(n)$ の決定ルール

$N_s(n)$  の決定 (式 (4)) で用いるルールは、現在単調さの解消を目的に登録している。このルールで用いられる数値は、システムの個性となるもので利用者の嗜好で調整することが好ましいと考えられるが、今回の実験では以下の値を利用した。

$$\begin{aligned} n_1 &= 12 (1 \text{ サイクル}) \text{ のとき } P_{iA}(n_1) = 0.4 \\ n_1 &= 24 (2 \text{ サイクル}) \text{ のとき } P_{iA}(n_1) = 0.7 \\ n_1 &= 36 (3 \text{ サイクル}) \text{ のとき } P_{iA}(n_1) = 1.0 \\ t_i &= 30, c_{iA} = 30 \\ n_2 &= 12 (1 \text{ サイクル}) \text{ のとき } P_{A1}(n_2) = 0.6 \\ n_2 &= 24 (2 \text{ サイクル}) \text{ のとき } P_{A1}(n_2) = 0.8 \\ n_2 &= 36 (3 \text{ サイクル}) \text{ のとき } P_{A1}(n_2) = 1.0 \\ t_A &= 50, c_{A1} = -30 \end{aligned}$$

$$t_p = \begin{cases} 20 & \text{楽節内 4 小節目の } N_s \text{ 決定時} \\ & (n \bmod 4 = 3 \text{ のとき}) \\ -20 & \text{楽節内 1 小節目の } N_s \text{ 決定時} \\ & (n \bmod 4 = 0 \text{ のとき}) \\ 0 & \text{otherwise} \end{cases}$$

#### 4.5 演奏パターン・データマップの作成法

演奏パターン・データマップは、大阪大学軽音楽部在籍の音楽経験 1 年以上の 20 名を対象に、一対比較法により作成した。2 小節繰り返しのパターンを刺激とし、2 種類のデータのすべての組み合わせをランダムな順序で被験者に聞いてもらい、以下のアンケートに答えてもらった。

あなたは、キーボーディスト (ギタリスト) で、セッションに参加しています。あなたにソロがまわってきました。ソロの途中あなたの気分は盛り上がり、激しいパターンを弾きました。あなたはドラム (ベース) にも盛り上がったパターンを叩いてほしいと思います。どちらのパターンを叩いてほしいですか。

一対比較法では各刺激の選択確率より、刺激間の相対距離とそれぞれの心理尺度値を導きだすことができる。本システムにおいては演奏パターンに幅を持つ尺度を与えることを考え、心理尺度値をもとに相対距離を絶対距離に置き換え、その最大最小値を心理尺度幅と定義し利用する。図 5 に 6 種類のドラムパターンの心理尺度幅を示す。この数値を 0~100 の値に変換し

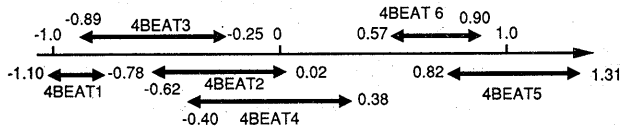


図 5 心理尺度幅  
Fig. 5 Scale width of patterns.

たものが演奏パターン・データマップとなる。実際の演奏実験には、ドラムパターンとして4ビートを15種類、8ビートを16種類、その他ビート変更、ブレイク、終了のための特殊パターンを10種類用いた。ベース用としては、4ビートを24種類、8ビートを24種類、特殊パターンを7種類用いた。

4.6 ビート変更ルール

ビート変更は人間の演奏を考慮せず、システムの個性で行った。具体的には、以下のルールを用いた。

- 2サイクル (24小節) 間、同一のビートで演奏されていれば、4割の確率でビートを変更する。
- 3サイクル (36小節) 間、同一のビートで演奏されていれば、6割の確率でビートを変更する。
- 4サイクル (48小節) 間、同一のビートで演奏されていれば、8割の確率でビートを変更する。
- 5サイクル (60小節) 間、同一のビートで演奏されていれば、必ずビートを変更する。

5. システムの評価

JASPER はこれまでに述べた方法で、MacintoshIIcx 上に構築された。

システムの評価を行うにあたり、以下の3種類の設定に対して5人のバンド経験者にシステムを実際に使用してもらい、その感想および人間同士の演奏との違いを聴取した。被験者はすべて大学の軽音楽部部員とOBで、音楽経験は10年程度、アマチュア・バンド活動が5年程度の経験を持つ。また、計算機システムについての専門的知識を有するのは1人である。各実験条件はあらかじめ被験者に知らされていた。演奏時間に関しては条件を与え

なかった。

【実験1】

人間の演奏に対してのシステムの応答を調べるために、自己主張過程を排除して、実験を行った。つまり、図4において常に人間をリーダーとみなし、協調度も低いと判断するようにルールを変更した。ビートも4ビートに固定にした。また、特徴量正規化関数は4人の演奏データの平均値から求めた。表1、式(5)がそれを表す。図6は演奏を採譜したものである。開始直後の12小節のみを示す。

【実験2】

実験1と同様、自己主張過程を排除した実験ではあ

Figure 6 shows a musical score for a performance. It consists of three staves: Piano (Human), Bass (System), and Drum (System). The Piano part is written in treble clef with a key signature of one flat and a 4/4 time signature. The Bass and Drum parts are written in bass clef. The score shows the first 12 measures of the performance, with various musical notations including notes, rests, and drum patterns.

図 6 演奏例の楽譜  
Fig. 6 A score of a performance.

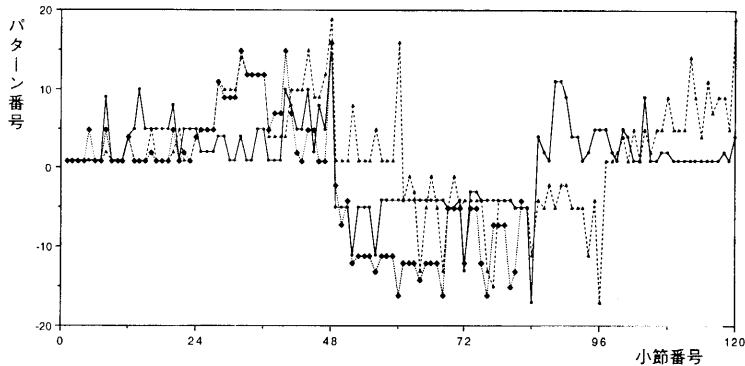


図7 ドラムパターンの遷移 (正が4ビート, 負が8ビートを表す)

Fig. 7 Flow of the selected drum pattern (positive value: 4 beat, negative value: 8 beat).

るが, 特徴量正規化関数は各個人のデータのみを用いて求めた. 一人の被験者の例を式(6)に示す.

$$\begin{aligned} n_r &= 7.14n - 42.86 \\ v_r &= 0.05 \exp(v/10) - 2.54 \\ p_r &= 5.00p - 300.00 \end{aligned} \quad (6)$$

また, 式(1)の重み係数は,  $a_n=5$ ,  $a_v=4$ ,  $a_p=1$  とした. 式(5)と比較して大きく異なるのがわかる.

#### [実験3]

実験3は, 実験2に自己主張過程を組み込んだものである. ビート変更のルールも加えた. 図7に3人の演奏者の実験におけるドラムパターンの遷移を示す. 縦軸はドラムパターンの番号に対応するが, 正数で4ビートのパターンを負数で8ビートのパターンを表している. また, パターン番号の値は20まで表現されているが, 実際に選択可能なパターン数はそのうち, 4ビートで15種類, 8ビートで16種類である. 3人の演奏データのうち1つは82小節までで, 83小節目から終了フレーズに移行し, 次の小節で終了している. この図よりシステム自体もバリエーションの多い演奏を行っているのがわかる.

以下, 被験者から聴取した感想をまとめる.

#### 〈演奏内容について〉

- 実験1においては少し不満が残るが, 実験2においては人間の演奏にかなりよく追従できている.
- 時間遅れはほとんど気にならない.
- 過去4小節のデータを用いることにより, ドラムパターンのつながりが滑らかに行われる. ただ, 急激な変化がほしいこともある.
- 実験3のように自己主張が入るとより演奏が面白くなる. 変化がつき飽きにくいいため, 現実の合奏に近づき長時間楽しめる.

#### 〈協調型システムについて〉

- 今までにない対話的なシステムだけに面白い.
- こちらの演奏に反応し出力が変わるので, こちらもそれにより刺激を受ける.
- このようなことがリアルタイムでできるとは, 計算機が身近に感じられる.

このように, 人間・機械間のコミュニケーションをリアルタイムで実現するシステムとしてかなり良い評価が得られた.

結果として, 演奏を“計算機との共演”という立場で見た場合は興味深い, 面白いといった良い評価が得られたが, “実際の人間との演奏”と比較すると演奏者個人の演奏特徴への適応等の点でまだ差があるといえる. ただ, 演奏時間についてはなにも指示を与えなかったにもかかわらず, どの被験者も1セッションにつき平均15分という長時間にわたってシステムとの合奏を行ったこと, そして楽しいという意見が得られたことは大きく評価できる.

## 6. 問題点と展望

今回の実験では, 演奏と緊張感の関係における個人差がかなり大きく, この関係の調整次第でシステムの評価が大きく異なった. そこで, 演奏者個人個人の演奏の特徴に対応するために, ヒューマンテンション抽出フィルタの調整を自動化するものとして‘リハーサル機能’や演奏中に演奏者の特徴を学んでいく‘学習機能’などを検討しなければならない.

また, 本システムを音楽的にも質の高いものにするためには, 緊張感以外にも多くの情報のコミュニケーションを実現させなければならない. 緊張感についても現在, 音数, 平均ベロシティ, 平均ピッチの3種類

の情報から求めているが、これら特徴値のみで緊張感が完全に計測できると思われぬ。さらに異なる特徴量を導入することが考えられる。そのためには、実際のジャム・セッションにおいて、どのようなメッセージがどのような手段で伝達されているのかを解析することが、最も重要であろう。

## 7. むすび

人間と機械によるジャム・セッションを可能とする協調型自動演奏システム“JASPER”を構築した。評価実験においては一様に‘かなり楽しく演奏できる’という意見が得られた。

入力演奏に対して出力演奏をただ対応づけるのではなく、システム内に人間の即興演奏過程をモデル化した演奏家モデルを構築したことは非常に有効であった。また、合奏を演奏者同士のコミュニケーションであるという観点から見ることで、演奏に表れる感性情報に注目し、それを抽出することを試みた。つまり、相手（人間）の演奏者の感情表現である緊張感を知覚し自分（システム）の緊張感へと結びつける過程を実現するものとして、テンション・パラメータを提案した。このテンション・パラメータの導き方や処理のためのルールを拡張することで、システムの演奏を人間の演奏により近づけることが可能であろう。

また、今回は一台の計算機システムと人間であるピアノ奏者の二者による対話に注目して研究を進めたが、今後複数の人間と複数の計算機システムによる合奏も考えたい。

謝辞 本研究を行うにあたり、多くの助言・協力をいただきました。大阪大学井口研究室の皆様および軽音楽部の皆様に心より感謝いたします。

## 参 考 文 献

- 1) Dannenberg, R.: An On-Line Algorithm for Real-Time Accompaniment, *Proceedings of the International Computer Music Conference 1984*, pp. 193-198 (1984).
- 2) Vercoe, B.: The Synthetic Performer in The Context of Live Performance, *Proceedings of the International Computer Music Conference 1984*, pp. 199-200 (1984).
- 3) Morita, H., Hashimoto, S. and Ohteru, S.: A Computer Music System that Follows a Human Conductor, *Computer*, Vol. 24, No. 7, pp. 44-53 (1991).
- 4) Nishijima, M.: Teaching Musical Style to a Neural Network, *Proceedings of the Com-*

*puter World '90*, pp. 195-201 (1990).

- 5) 文部省重点領域研究「感性情報」A班編:「感性情報・感性情報処理とは何か」論集 (1993).
- 6) Deutsch, D. (寺西立年, 大串健吾, 宮崎謙一監訳): 音楽の心理学(下), 西村書店, pp. 588-609 (1982).
- 7) 才脇直樹, 片寄晴弘, 井口征士: 感性情報を利用したミュージック・エディタ, 1991年電子情報通信学会春季全国大会講演論文集, 分冊6, p. 242 (1991).

(平成5年2月19日受付)

(平成6年3月17日採録)



和気 早苗 (正会員)

昭和44年生。平成4年大阪大学基礎工学部制御工学科卒業。同年日本電気(株)入社。現在、同社関西C&C研究所にて、ヒューマンインタフェース関係の研究開発に従事。



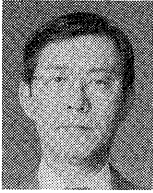
加藤 博一 (正会員)

昭和38年生。昭和61年大阪大学基礎工学部制御工学科卒業。昭和63年同大学院修士課程修了。現在、同大学基礎工学部システム工学科助手。パターン認識、画像処理、ヒューマンインタフェース等に興味を持つ。



才脇 直樹 (正会員)

昭和40年生。平成元年大阪大学基礎工学部制御工学科卒業。平成3年同大学院修士課程修了。現在、同大学大型計算機センター豊中地区データステーション助手。脳波解析、信号処理、音楽情報処理等の研究に従事。



井口 征士（正会員）

1940年生。1962年大阪大学工学部電気工学科卒業。1964年同大学院修士課程修了。1965年基礎工学部制御工学科助手。1969年助教授。1984年教授（現システム工学科）。専門はパターン計測，画像処理，音響信号処理。最近「感性情報」に興味を持つ。著書「三次元画像計測」「感性情報処理」など。電子情報通信学会，計測自動制御学会，IEEE，ICMA 各会員。

---