

エージェントを用いた音楽セッションシステム

才脇直樹 川端純一¹⁾ 小田晃弘²⁾ 寺嶋知織 仲谷美江 西田正吾

大阪大学大学院 基礎工学研究科

¹⁾ 現在 川崎重工業株式会社

²⁾ 現在 NTT西日本株式会社

近年、コンピュータが個々の人間の個性や独自性を尊重しつつ、人間をサポートすることが求められている。本研究では、協調システムのなかでも特に合奏システムに注目しエージェントを用いて人間とコンピュータの協調の実現を目指している。合奏時にエージェント間でやりとりされるコミュニケーション情報に関して、これまで心理実験などから得られる人間の音楽認知特性を十分に考慮した例は多くない。しかし、音楽をどのように分析したり、どのように感じるかといった音楽認知特性を知ることはより人間らしい協調を実現するために必要不可欠である。そこで、本稿ではエージェントの仮想感性ともいえる「心理的ポテンシャル」と人間の音楽認知特性を知るための「注目度」という観点を提案し、プロトタイプ合奏システムを構築して結果について評価検討を加える。

Construction of Session System with KANSEI Agent

Naoki SAIWAKI, Chiori TERASHIMA, Mie Nakatani and Shogo NISHIDA,
Graduate School of Engineering Science, Osaka University
Jun'ichi KAWABATA,
Kawasaki Heavy Industries, Ltd.
Akihiro ODA
NTT WEST, Ltd.

Abstract: This paper describes the cooperative performance system on musical ensemble with mental model. To realize more intimate cooperation, we take notice of the information on feelings, such as user's personal taste and mental state. In this model, emotional factors of agents are expressed by "mental potential," which is estimated by the relation between agent's behavior and the situation. And Attention level is proposed for considering the characteristics of communication between agents and human.

The prototype system was developed for evaluation and it was confirmed that users were adequately satisfied with playing with agents.

1. はじめに

初期のコンピュータは、人間の機能を代行したり支援したりして、作業の信頼性、安全性、効率をあげるための道具であった。しかし、近年では、人間の作業を支援するだけでなく、人間の心をいやしたり、楽しませたりするパートナーとして登場することが多くなってきた¹⁾。人間のパートナーになるためには、論理的な思

考や計算能力だけではなく、優しさ、楽しさ、柔軟さなど、いわゆる感性的なコミュニケーションが必要になってくる。感性コミュニケーションの研究では、うなずきなどの動作や簡単な返事で人間と会話するロボットの開発^{2) 3) 4)}や、音楽や会話における“間”の研究⁵⁾などがある。本研究では、音楽分野に注目し、人間とジャズセッションを楽しむコンピュータ（以下エー

エージェントと呼ぶ)を開発した。ここでは、感性情報の中でも特に合奏のときに有効なパラメータを提案し、伴奏するだけでなく、人間との掛合いをするシステムを開発したので報告する。

2. セッションにおける感性協調モデル

人間の意図をくみとって知的に対応するエージェントの研究はこれまでも行われている⁶⁾⁷⁾。ほとんどのエージェントは情報検索やスケジューリングなど人間の日常業務のサポートを目的としている。エージェントは、人間による入力(テキスト情報や操作履歴など)から目的や意図を理解し、その意図や指示を満足する方法を提供する。

一方、ここで扱うジャズセッションは、複数のプレイヤーの共同作業だが、ゴールや方向性が明確にあるわけではない。開始時に決まっているのはコード進行と楽器の種類だけである。どんなメロディにするのか、誰がリードをとるのか、どこで盛り上げるのか、などは演奏中にダイナミックに変わっていく。しかも、演奏中はプレイヤー同士が言葉でコミュニケーションすることはなく、相手の演奏の強さやリズムの変化から判断してリアルタイムに合わせていかなければならない。さらにセッションを面白くするためには、相手の演奏にただ合わせるだけでなく、単調な演奏が続いたらリズムを変えてみたり、次は自分がリードをとるという自己主張をするなど、各プレイヤーが個性を持っていることが望まれる。このような場合のエージェントには、セッションに関する基礎的な知識の他に、セッションに変化をつけるためのパラメータ(感性)が必要になってくる。

以上の概念に基づいて感性協調プロセスをモデリングする。セッションには、相手にあわせて演奏する協調的な側面と、自分の心理状態に応じて演奏する自己表現の側面がある。これらは、次の3つの概念で表される。

(1) 並列な状況認識及び各々の満足度
人やそれぞれのエージェントが客観的に見てど

のような状況にあるかを外部から一元的に評価するのではなく、各々が異なった視点から自分自身がどのような状況におかれていると認識しているのか。

(2) 各々の意図

各々が、認識した状況をどのように変化させることを望んでいるのか。

(3) 各々の状況操作

(2)で望む変化を引き起こすために、それぞれがどのような意思表示手段を用いるか。

人間やエージェントがそれぞれ異なる意図に基づいて行動を決定する以上、相互のコンフリクトは避けられない。しかし、各々の満足度がバランスよく増大するように、各々が自身の知識と観点に基づいて繰り返し状況の改善を試みる事で協調が実現される。以上の概念に基づいて構築した感性協調モデルをFig.1に示す。

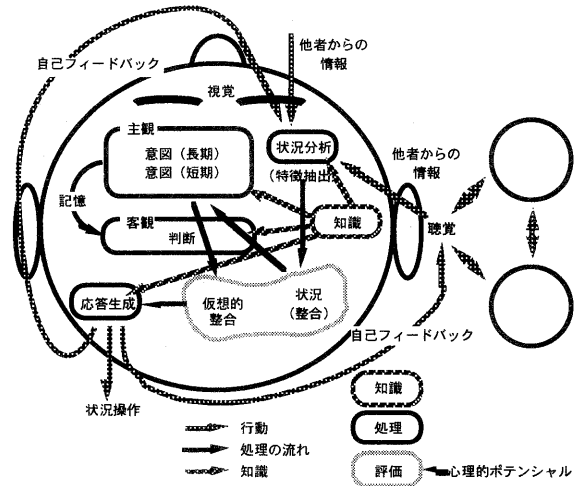


Fig.1. Model of cooperation with feelings.

まず、外部から来る他者のふるまいに関する情報と、自己のふるまいとの相互関係を分析し、その特徴を抽出することを状況認識と定義する。その状況に対して、次にどのような応答をする

かを心理的ポテンシャルを用いて各々が主観的に決定する。主観的意図の決定に際しては、現在の状況に対する他者の次の行動が予測され考慮される。そして各々のふるまいが仮想的に整合された結果をもたらすような応答が生成される。ここで状況分析、主観的意図決定、客観的行動判断、応答生成の処理は各々独自の知識に基づいて行われる。

知識は明示的なルールベースとして存在するわけではなく、一つ一つは単純な最小限の機能に対応するモジュールとしてそれぞれシステムに埋め込まれ、いくつかが集まって分析や生成の機能を実現し、さらにそれらが集まって全体として一つのエージェントとして振る舞う。これは、他のエージェントや人間から与えられる環境、例えば音楽の場合は演奏の状況、を柔軟に獲得し、又逆に心理的ポテンシャルからフィードバックされる情報によって機能を動的に変更するのに適していると考えたためである。本論文では、音楽を対象としてこのような記述を実現するプログラミング環境 MAX を用いた⁸⁾。

この感性協調モデルは、最終的に一意で静的な解（状況）に安定するわけではなく、刻々変化しながら局所的に協調の実現した安定状況が出現することが予測される。つまりリアルタイム性やインタラクションを持つことが特徴であり、これらが個人差や状況依存性を吸収してシステムを安定させつつ、次の変化の引き金になるというダイナミックな協調を行うところが従

来提案してきた音楽協調型システムとは大きく異なるところである^{9) 10)}。

感性協調における心的状況は、各々のエージェントの中に仮想的に構築される。本論文ではその一手法として、自身の振るまい（意図）と状況の関係から算出される「心理的ポテンシャル」を提案する。感性協調モデルにおける状況評価や整合のとれた意思決定は、心理的ポテンシャルを用いて行われる。心理的ポテンシャルは「ストレス」と「盛り上がり」の二つの側面から、予め与えられた知識に基づいて計算される。ストレスは状況に対する満足度であり、状況が予想されたもの、すなわち起こる確率が高いものならばストレスは小さくなり、逆に予想外で確率の低いものでは大きくなる。盛り上がりはテンションの高低を表すもので、理論的に盛り上がりと考えられるときにプラスされ、盛り下がるときにマイナスされる。この二つの組合せで心理状態が表現され、それを具体的に計算した値が心理的ポテンシャルである。

一方、以下のようにして注目度の導入を試みた。人間は、音が大きくなれば「うるさい」とか、短音階ならば「かなしい」といったような心理的影響を受ける。これらから合奏時に何らかのコミュニケーション情報がやりとりされていると考えられる。一方が目立っているときには他方が控えめであったり、両方激しい演奏を行っている状態で急に一方が穏やかになる事によって、他方の演奏を際立たせたりといった合奏戦略によって演奏がよりダイナミックで自然なものに聞こえる。これは、合奏に参加している人間がどの楽器の演奏にどの程度注目させられるかを制御していると言い換えることもできる。本稿では、「聴き手が、演奏者の音や演奏パターンに注目させられる度合いを表す指標」を注目度と定義する。そして、あらかじめ心理実験から注目度が実際にどのように変化するかを調べ、エージェントに持たせることを試みた。

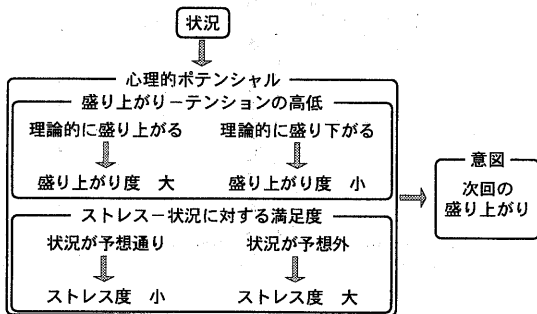


Fig.2. Mental Potential.

例えば、

- ・ベースやドラムパターンの注目度は、リズムや音程変化によって変化する。
- ・メロディパターンの注目度はベースパターンとドラムパターンの絶対的な注目度の合計で決まるのではなく、両者の相互関係で決まる。
- ・ベースとドラムの個別の注目度変化についてはメロディパターンの時系列変化の影響が強い。

などが知識としてシステムに与えられた。実際のセッションで注目を集めるパートが時間の経過により次々と移り変わっていくように、エージェントが状況を判断して自主的に注目度を制御できるような心理的モデルを持つことによって、より自然な合奏を再現できると考える。

3. エージェントの構成

各エージェントは Fig. 3 のように三部から構成されている。

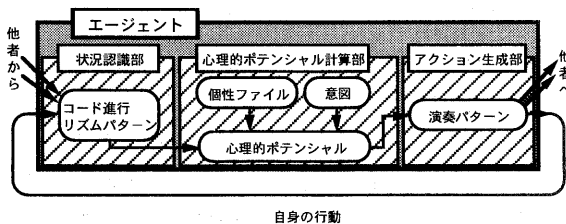


Fig. 3. Architecture of the agent

(1) 状況認識部

ここではエージェントに入力された情報の特徴を抽出することで、そのエージェントがおかれている状況を認識する。合奏における状況を表すものにはコード進行やリズムパターンなどがある。

(2) 心理的ポテンシャル計算部

状況認識部で得られた状況をもとに心理的ポテンシャルの値が計算される。合奏における心理的ポテンシャルは、例えばリズムを変えると

きの人間のフィーリングの役割を果たす。心理的ポテンシャルの計算におけるその他のパラメータの値、各要素の重み付けを「個性ファイル」と呼び、エージェント毎に異なったものが与えられている。それらの値は音楽理論をもとに決定する。個性ファイルの違いにより同じ状況に対する応答に違いが生じるため、この部分でエージェントの個性が実現される。心理的ポテンシャルによって、予測可能で落ちついた状況が長く続いた場合は状況を打開しようとする意図が強まり、逆に予測しにくく活動的な状況が続いた場合は状況を安定させようとする意図がエージェントに働くようになる。

心理的ポテンシャルのフィールドは他のエージェントや人間からの入力情報にも基づいて常に変動しているため、状況によってはエージェント自身の意図的な操作が加わらなかった場合でもポテンシャル値が変動してエージェントのアクションを促進したり、逆に抑制したりする。

(3) アクション生成部

エージェントが次にとるアクションは、この部分で心理的ポテンシャルを閾値として用いてリアルタイムで決定、生成される。閾値により生成されるリズムパターンやフレーズが変わる (Fig.4).

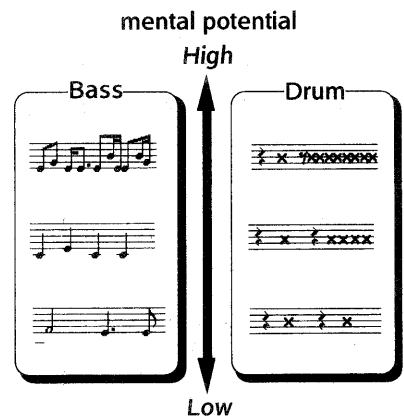


Fig.4. Example of performance pattern

4. システム構成

エージェントの有効性を検証するため、人間とエージェントが合奏を行うプロトタイプシステムを構築した。システムの概要を Fig. 5 に示す。ここではピアノ、ドラム、ベースという3人によるセッションを取り上げ、人間がシンセサイザを演奏し、二つのエージェントはドラム（エージェント1）とベース（エージェント2）を担当する。そして合奏時における人間-エージェント間、エージェント-エージェント間の協調を目指す。

エージェントは人間の演奏からコードとコード進行を抽出することで人間の心理状態を認識する。その結果に応じて出力をリアルタイムに生成し、それがエージェントの演奏となって人間に与えられる。エージェントどうしの場合、ベースはドラムの演奏からフィルインの長さを抽出し、それに応じて自身のストレスを解消し、出力を生成する。またドラムはベースの演奏から1小節内の音数を抽出し、それに応じた出力を生成し演奏する。

エージェントは人間の演奏からコードとコード進行を抽出することで人間の心理状態を認識する。その結果に応じて出力をリアルタイムに生成し、それがエージェントの演奏となって人間に与えられる。エージェントどうしの場合、

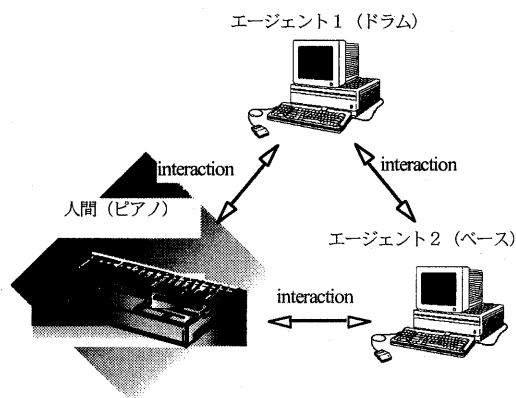


Fig. 5. Prototype System

ベースはドラムの演奏からフィルインの長さを抽出し、それに応じて自身のストレスを解消し、出力を生成する。またドラムはベースの演奏から1小節内の音数を抽出し、それに応じた出力を生成し演奏する。

5. 実験と結果

前節で構築したシステムを用いて評価実験を行った。人間はシンセサイザ演奏の初心者3名、上級者3名の計6名を被験者とした。その結果、

(1) 初心者3名はエージェントとの合奏を楽しむことができた。しかし、客観的に見て音楽的に良いか悪いかの判断はできず、主観的にインタラクティブな変化が感じられれば良い、という程度の判断基準である。

(2) 演奏に熟練した3名は、最初のうちはエージェントとの合奏を楽しんでいたが、10分ぐらいで全員興味を失ってしまった。なぜならエージェントの演奏パターンが予測できるようになるからである。

(3) 被験者の演奏レベルが低い場合、生成される演奏パターンを豊富にすれば、それだけで興味の持続時間を延ばすことができる傾向にあった。しかし、上級者ではインタラクションの質を重視する人が多く、音楽的文脈を分析できる被験者にとっては、洗練度の低さが幼稚に感じられるようであった。などが挙げられる。

6. むすび

実験結果を踏まえた上でより良い協調を実現するため、

(1) 音楽生成機構の見直し

今回は確率とルールを音楽生成に用いたが、創発的要素の導入や生成する音楽の構造について検討したい。

(2) 状況を表すパラメータの増加

例えば音量やテンポなどの要素を増やすことで、より複雑な状況をシステムに認識させることができる。

(3) 動的に変化する個性の実現

個性ファイルを状況からのフィードバックでリアルタイムに改変できるようにすれば、変化する心理状態やさまざまな個性をより適切に表現することができる。また、ドラムとベースにそれぞれ全く異なった個性を持たせた場合や微妙な差を持たせた場合などで合奏がどのように変化するかを調べることによって、システムの問題点について別の観点から検討することができると思われる。

(4) 評価方法の検討

今後、音楽の構造分析などを含む定量的な評価法を確立する必要がある。但し、これは(1)との関連も大きいと考えられる。などについて取り組んでいきたい。

謝 辞

本研究は一部、日本学術振興会未来開拓学術研究推進事業「インタラクシオンによる相乗効果を用いた感性創発世界の構築」による支援を受けている。

参考文献

- (1) 柴田崇徳：人の心を癒すメンタルコミットロボット，日本ロボット学会誌，Vol.17, No.7, pp.943-946 (1999)
- (2) 藤田雅博：ペット型ロボットの感性表現，日本ロボット学会誌，Vol.17, No.7, pp.947-951 (1999)
- (3) M. Okada, S. Sakamoto. and N. Suzuki, "Muu: Artificial Creatures as an Embodied Interface", SIGGRAPH 2000, the Emerging Technologies: Point of Departure, 2000
- (4) 渡辺富夫, 大久保雅史, 中茂睦裕, 檀原龍正：InterActor を用いた発話音声に基づく身体的インタラクシオンシステム，ヒューマン・インタフェース学会論文誌，Vol.2, No.2, pp.127-135 (2000)
- (5) 中村敏江：“間”の感性，井口征二（編）感性情報処理，pp.151-169，オーム社（1994）

(6) Pattie Maes, "Agents that Reduce Work and Information Overload", Communication of the ACM, Vol. 37, No.7, pp31-40, 1994

(7) M. R. Genesereth, S. P. Ketchpel, "Software Agents", Communication of the ACM, Vol.37, No.7, pp48-53, 1994

(8) 才脇直樹, 川端純一, 西田正吾：心理的ポテンシャルを用いた感性協調型合奏システム，ヒューマンインタフェース学会論文誌 Vol.2, No.1, pp.47-57 (2000)

(9) A. Oda, N. Saiwaki and S. Nishida, "Cooperative Performance System with Information on feelings", Proceedings of the 1999 IEEE International Conference on Systems, Man, and Cybernetics(SMC'99), pp. IV 236-241, Tokyo, Japan, October, 1999.

(10) 才脇直樹, 小田晃弘, 寺嶋知織, 西田正吾, 川端純一：感性的情報を用いたセッションシステムの構築，第39回計測自動制御学会学術講演会 CD-ROM 予稿集 (2000)