

感情的意図を持ったセッションを行うシステムの実現に向けて

大野 孝紘 岡 夏樹
(京都工芸繊維大学)

音楽によるインタラクションの最たるものがセッションであるが、これをコンピュータで実現しようとするとき、感性的情報の授受を実現させることは極めて重要である。これを実現するための手法がいくつか提案されているが、まだまだ制約の大きいものである。インタラクションにおいて不可欠な相手の意図を汲み取ること、特に音楽インタラクションを行う上では感情的な意図を汲み取ることが必要であると考えられる。本稿では、まず我々が「気の合うエージェント」と呼んでいる、ユーザの意図を把握して行動を決定するエージェントシステムの必要性と一実現法を述べ、それがセッションシステムにおいても有効ではないかということ提案する。

Suggestion to Reading Emotional Intention for Musical Sessions

OHNO Takahiro OKA Natsuki
Kyoto Institute of Technology
<http://www-vox.dj.kit.ac.jp/>

Session is the most typical interaction by music. However some studies aimed to realize a session with a musician in a computer, this is not fully accomplished yet. We believe that it is necessary to understand the emotional intentions of the partner in order to realize a musical session with a human player. The main point of it is that the computer guesses the intention in a probabilistic way from the sound he made by using a Bayesian Network.

1. はじめに

音楽によるインタラクションの最たるものがセッション(特にジャムセッション)である。音楽のような感情的なものを含む媒体によって能動的なインタラクションするとき、そこには音楽的意図(曲の流れや編成)の他に感情的な意図(例えば「悲しい感情を持っている相手を楽しくさせよう」など)が含まれる。

コンピュータの中に音楽家を作り、人間の演奏者とのセッションを実現するシステムは多数提案、実装されているが、それらは多くの制約の上に成り立っている。

セッションシステムは

1. 音響情報の取得
2. 音楽の解釈
3. 演奏の生成

と、大きく3つの処理によって構成される。本稿では、まず我々が提案している「気の合うエージェント」の実現に向けての研究を紹介し、続いていくつかのセッションシステムの研究を紹介したあと、「気の合うエージェント」を用いて「音楽の解釈」の面からその制約を軽減するための提案と考察を行う。

2. 「気の合うエージェント」

2.1. 「気の合うエージェント」とは

知的エージェントシステムとは周りの状況を判断し自律的に行動を決定するシステムのことであるが、人間を目標として研究され、人間に近づけようとするのが行われている。さらに最近ではこれに感情を持たせようという動きもある。しかし、この動きが進んだところ

で人間の手を離れて行動していくわけではない。なぜならエージェントの使命はそもそも文字通り人間の代理人として行動することであり、人間と意思の疎通をはかる、いわばインタフェースとしての役目も大きいからである。人間のように振舞わせることで、人間が人間と接するようにシステムを使える、即ち人間にとって最も自然なインタフェースとなりうる。この代表例がロボットであり、ロボットは形や物理的動作も人間に近づけようとしている。

しかし、このようにエージェントシステムが人間に近づいていくことは利点ばかりではない。現在は人間と機械には大きな隔たりが存在するが、それが無くなったとき、人間は機械と、機械としてではなく感情ある人間として接することになる。これは人間同士に存在する問題がそのまま人間と機械の間にも存在しうることを意味している。その大きなひとつとして、人間とエージェントが気が合うかどうかの問題となると考えられる。例えば、価値観や感性が違うなど意思の疎通がうまくいかなければ、コミュニケーションどころか使う気さえ起きなくなる。もともと人間が機械を使いやすくするために人間に近づけたのであるはずが、これでは本末転倒である。

そこで、常にユーザが発する動作や言葉などからユーザの意図を理解し、「気の合った」状態を保つエージェントシステムの実現が必要となる。これを我々は「気の合うエージェント」と呼んでいる。

2.2. 「気の合うエージェント」に対する要件と構成

ユーザがエージェントの感情に訴えかける意図を持ったインタラクションにおいて、「気の合うエージェント」システムに求められる要件を調べる実験を行った[1]。

まず、被験者に図1のようなアンケート用紙を渡す。これには Lang の 2次元感情モデルを基にした代表的な4感情 (Joyful, Enraged, Sad, Relaxed) が配されている。そして、「あなたの親友の感情が現在 {Joyful, Enraged, Sad, Relaxed} です。この曲を聴いたときになると思う感情の場所に丸をつけて下さい。」と言い、各感情に対して 20 曲聴かせ、延べ 80 曲についてこれを行った。このアンケート結果を基に、元の感情から各音楽によって遷移すると感じた量として Joyful→Relaxed といったようなベクトル的な感情変化を得る。

(^>.<) (^^)

(T~T) (T~T)

図1 アンケート用紙

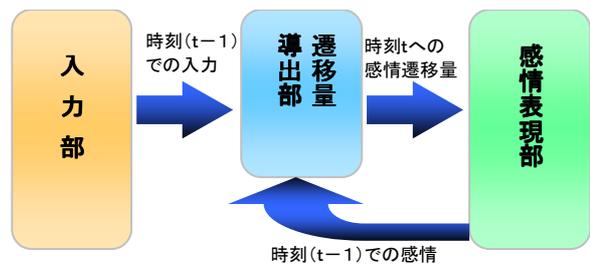


図2 感情遷移モデル

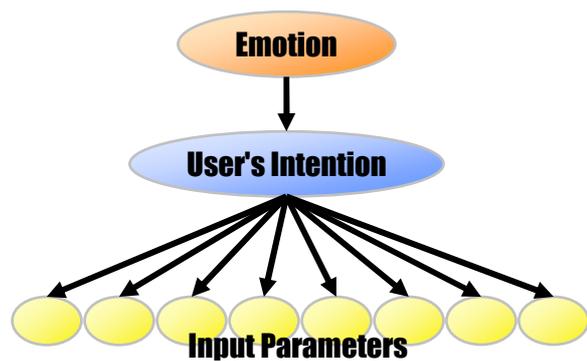


図3 遷移量導出部

この実験の結果から読み取れることとして、まず個人差がある。同じ音楽、同じ元の感情からの遷移方向が被験者によってほぼ異なった。すなわち、ユーザが感じる音楽の感情的作用はユーザによって様々である。次に、被験者は相手の感情状態によって異なる感情遷移を予想していることが分かった。異なった感情状態でも同一の音楽に対して同じ感情遷移先を示した結果は極めて少数であった。

以上の結果を「気の合うエージェント」システムに組み込む感性モデルについて言えば次のようになる。

- エージェントはユーザ毎に違った感性を持たなければならない
- エージェントの感情の遷移を決めるとき、外的刺激だけではなくその時点での感情も無視できない

これらよりベクトル的な遷移量を導出して感情が遷移する図 2 のような感情遷移モデルを考えることができる。同図において「感情表現部」はエージェントの感情空間を持つところで、感情を動かす外的刺激（ユーザからの入力）から感情を遷移させる特徴量を取り出すのが「入力部」、その両者からの値を考慮して感情の遷移量を決定するのが「遷移量導出部」である。

さらにこの「遷移量導出部」は、確率的に遷移量を決定するベイジアンネットワーク[2]を用いた図 3 のような構成が最も自然であると思われる。ユーザごとの差異はこの CPT（条件付確率テーブル）に依る。

3. 感性情報を取り入れたセッションシステムの研究事例

コンピュータの中に音楽家を作り、人とセッションをするシステムは多数研究されている。それぞれ個性豊かなアイデアで人間のミュージシャンとやっているようなセッションを実現しようとしている。

人間同士でのジャムセッションにおいて、ミュージシャンは演奏から生じる感覚に従って演奏を行う。インタラクティブなセッションにおいて、感性的な情報が有効であるということはここで言及するまでもない。そこで、1.で述べた「音楽の解釈」に感性情報を取り入れる為の研究事例について、どのように感性情報を扱っているかを中心に紹介する。

3.1. JASPER とその関連研究

文献[3]において JASPER というセッションシステムが紹介されている。このシステムではピアノを人間が、ベースとドラムをシステムが担当しブルースのコード進行にのって即興的な演奏をする。

このシステムでは人間の演奏の抑揚の変化を「テンション・パラメータ」と呼ばれる感性量として導入している。まず、入力された演奏データのうち音数、ベロシティ、ピッチを用いて各小節ごとの演奏者のテンション値を求め、そして最近のテンション値の平均からその時刻の演奏者のテンション値を決定する。

このパラメータは、まずリード奏者が人間であるのかシステムであるのかの判定に用いられる。そして人間がリードをとっているときは、求めたテンション値に「クレシェンド・デクレシェンドの情報」及び「現在の小節位置」を加味して次のテンション値の予測を行い、それを

基に（即ちユーザの演奏意図に従って）予め決められた小節単位の演奏パターンから選択する形で出力を行う。また、このシステムには「個性データベース」と呼ばれるテンション・パラメータの推移則を記述したルール群が実装されている。システムがリード演奏を行うときやバックアップであっても単調さを避けるためにこのデータベースを用いて次のテンション値を導き、演奏パターンを選択する。

文献[4]では JASPER を発展させたシステムが考案されている。このシステムでは、テンション・パラメータ以外の、「盛り上げる」「コードを強調する」「テーマを再現する」などの多様なソロ奏者の演奏意図を抽出している。これらの検出のために音響情報の取得においてもさらに多様な特徴を扱っている。また演奏の生成についても、あらかじめ与えられた基本演奏データを状況に応じて変化させるという、ジャムセッション的な手法の提案をしている。

3.2. 金森らの研究

文献[5]においてセッションシステムの聴取部分、特にコード進行の認識についての精緻な研究が紹介されている。このシステムでも JASPER と同様ピアノを人間がベースとドラムをシステムが担当し、スタンダードジャズの即興演奏を行う。

入力された演奏を基に、そのコード名を決定する。候補をすべてあげ、過去のコード進行との連続性を考慮して現在のコードを決定することによって（特定のジャンルのみを扱ってはいるが）非常に高い精度のコード名の認識を実現している。

こうして得られたコード進行をもとに、トニック、ドミナントといったコードの機能と局所的な調性、転調が生じた場合はその調の距離を求める。得られた「コード名・進行」「コードの機能」「調性」「転調の距離」の4つの情報を分析して、「トニックによる安定感」「ドミナントによるトニックへの期待感」「コードの多義性による緊張感」などといった13種類の形容語をもった感性情報を抽出している。これらは音楽的情報と感性形容語が主観的判断によって対応付けられたものではあるが、これだけ多様な感性情報を適切に演奏出力に結びつけることができれば、音楽性と表情豊かなジャムセッションが期待できる。しかし、残念ながらその生成の仕方については文献[6]においてほとんど触れられていない。

4. 「気の合うエージェント」を使ったセッションシステム

前項で 2 種類のセッションシステムの実現例を挙げたが、双方とも導入された感性量は音楽の進行とかなり密接な関係にある。即ち、ユーザの感情的意図の導入を行っているが、読み取っているユーザの意図は、音楽に対するテンションや音楽による感情表現をどう持つていくか、といった音楽に向けたものだけであり、システムに向けたものではないということが言える。例えば JASPER において、予測テンション値または個性データベースから発した音楽によってユーザが返してきた音は、人間はシステムとインタラクションを行っているにも関わらず JASPER 自身に感性的影響を与えない。それでは感情的に影響を与え合うインタラクションとしては不十分であると考えられる。

これは両者とも感性量からのみ演奏を創出していることに原因があると考えられる。我々は音楽的な処理とは別に感情情報を取り入れることを行うべきであると考え、ここに「気の合うエージェント」の考え方を導入して、感情的なセッションシステムが実現できるかどうか考察する。

このシステムにおいては、コード進行や盛り上がりなどの音楽的な処理と、感性的な処理は別に扱うことになる。即ち、「音楽の解釈」を「音楽的理解」と「感性情報処理」に分ける。前者は様々な音楽理論に基づき、後者には「気の合うエージェント」を用いる。

まず、取得した適当な音響情報（ベロシティやピッチなど）からユーザのシステムに向けた感情的意図を推論する。ユーザはシステムの出す音楽を聴いているので、時点毎のシステムの感情を推測しているはずである。これにより、システムに投げかける音楽に込められているユーザの感情的意図とはシステムの感情を現在の感情から次の感情へと遷移させるというものである（例えば「sad→joyful」）。これを推論してシステムの感情を変化させる。ベイジアンネットの CPT の学習を行うことで、ユーザはエージェントと気が合った状態、即ち思い通りにエージェントの感情を変化させることができる。

演奏の創出にはシステムも感性的意図を持つべきである。そこで、まずユーザの感情をニューラルネットなどのなんらかの推論法により得、自らの感情と合わせてシステムのユーザ

に向けた感情的意図を決定する。さらに音楽的なコードの流れやメロディーなどを担当する部分からの情報とあわせてシステムの感性的音楽を創出する。

以上が提案する感情的意図を持ったセッションを行うシステムであるが、これについて考察を行う。まず、システムは感情（空間）をもつことにより、ユーザからの刺激により変化する内部状態を持つことになる。これにより、ユーザは相手を意識したより自然で豊かなコミュニケーションを行うことができる。また、システムの感情は演奏創出の一部に影響を及ぼすのみであり、基本は成果を上げているものも多い感情を取り入れていないセッションシステムや自動作曲システム[6]と同等の仕組みを用いることができるので成果は十分期待できるものとする。

5. まとめと今後の課題

本稿では人間とシステムの感情的意図を持った音楽によるインタラクションシステムの一構成法を提案した。

ただ、提案は「音楽の解釈」の一部分であるに過ぎないため、今後ほかの部分についても議論を進め、一日も早く実装してこのシステムの有効性を示したい。

参考文献

- [1] 大野孝紘：『ユーザの意図を意識した感情遷移システムの提案 - 「気の合うエージェント」の実現に向けて -』, 京都工芸繊維大学卒業研究報告書, 2003.
- [2] 2003 年ベイジアンネットセミナー (BN2003) <http://www.bn2003.org/>
- [3] 和気早苗, 加藤博一, 才脇直樹, 井口征士：テンション・パラメータを用いた協調型演奏システム - JASPER -, 情報処理学会論文誌, Vol.35, No.7, pp.1469-1481, 1994.
- [4] 日高伊佐夫, 後藤真考, 村岡洋一：ジャズの独奏の変化に対応する自動伴奏システム, 情報処理学会研究報告, 95-MUS-9, Vol.95, No.19, pp.7-12, 1995.
- [5] 金森, 片寄, 新見, 平井, 井口：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.
- [6] 長嶋 洋一ら編：「コンピュータと音楽の世界 - 基礎からフロンティアまで」, 共立出版, 1999.