

研究者依頼用報告書

1. 研究課題名

ドレミっち： 成長する仮想演奏者の構築

2. 氏名

浜中 雅俊

3. 研究のねらい

本研究では、ユーザと即興演奏することにより、次第に演奏が上達するなど、音楽的に成長する仮想演奏者ドレミっちの実現を目指し、次の3点について研究を行った。第一に、高度な音楽的活動を可能とする仮想演奏者を実現するため、計算機上へ音楽家の音楽知識の実装を試みた。第二に、仮想演奏者があたかも初心者から熟練者に成長するような機能を実現するため、メロディのモーフィング手法を構築した。第三に、初心者が複数の仮想演奏者の演奏を聴き分けるための能動的音楽鑑賞インタフェースについて研究を行った。このような仮想演奏者システムの実現は、初心者にも音楽を楽しませるアミューズメント性をもつだけでなく、プロの演奏者にとっても音楽制作の生産性をあげる技術の一つとして期待が持たれる。

4. 研究成果

1) 音楽理論 GTTM の計算機上への実装

音楽家の音楽知識を計算機上へ実装する第一歩として、音楽家の音楽知識を体系化したものである音楽理論の計算機上への実装を行った。その際、音楽知識を計算機上に形式的に記述する観点から、音楽理論として GTTM (Generative Theory of Tonal Music) を採用した。そして、GTTM を計算機実装用に拡張した exGTTM を提案し、複数の調節可能なパラメータを導入することで、音楽そのものに内在する曖昧性を積極的に認めつつ、音楽理論の曖昧性を解消することを可能にした。exGTTM を計算機上へ実装した楽曲分析器 ATTA (Automatic Time-span Tree Analyzer) は、46 個の調節可能なパラメータを持ち、パラメータの値を変更することで音楽の様々な解釈の結果が出力される。さらに、様々な解釈の可能性のうち、人間は出来る限り安定な解釈を優先すると仮定することで ATTA のパラメータを自動で最適化する自動楽曲分析器 FATTA (Full Automatic Time-span Tree Analyzer) を実現した (図 1)。FATTA の実現によって、メロディの重要な部分から装飾的な部分を分離するメロディの簡約や、メロディ、コード進行の予測など音楽家が行う高度な音楽的活動を計算機上で実現することが可能となった。

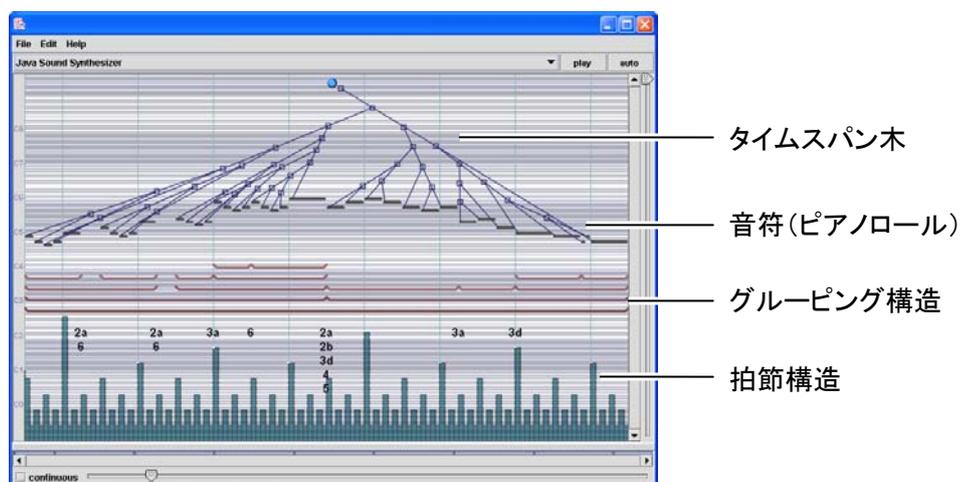


図 1: 自動楽曲分析器 FATTA を用いた楽曲構造分析の結果

- 音楽理論 GTTM を実装する目的

音楽というメディアの認識や表現は曖昧なため、ユーザの思い通りに作曲させたり演奏させることは一般に困難である。市販の楽譜エディタやシーケンサが操作できる対象は、音符、休符、和音名など曖昧性の低い表層的な構造に限定されている。我々の研究の最終的な目的の一つは、専門家の持つ音楽知識を計算機上に形式的に記述することで音楽知識の乏しいユーザを支援し、メロディ、リズム、和声といった高次の音楽的な構造を適切に操作できる音楽システムを実現することである。そのための第一歩として、本さきがけ研究では、音楽理論 GTTM の計算機上への実装を行った。

- 音楽理論として GTTM を採用した理由

音楽理論は、音楽に関する知識、経験、技能を利用して楽曲を分析、解釈する方法論を我々に与える。音楽は様々な側面から分析、解釈することができるため、これまで多くの音楽理論が提案されており、楽曲の分析や解釈に用いられる様々な音楽的な概念が抽出され、様々な手順が議論されてきた。

我々が音楽理論として GTTM を採用した理由は、GTTM は音楽が備える多様な側面を包括的に表象しているからである。音楽知識の乏しいユーザを支援し音楽的な構造を適切に操作するという我々の目的と照らし合わせると音楽の持つメロディ、リズム、和声という3つの側面に対して一貫性のある操作を実現する必要があり、実装する音楽理論として GTTM が適切だと考えた。

- 音楽理論 GTTM の実装の困難さ

GTTM は他の音楽理論と比べて比較的厳密なルールで記述されており、音楽を形式化する上で最も有望であると考えられるが、それでもなお、実装を実現する上で以下のような困難な問題が複数存在した。

(1) 楽曲分析における曖昧性の問題

一般に GTTM を含む音楽理論の分析結果は唯一に決まらないことが多い。その理由としては (i) 音楽理論自体に曖昧性があるために分析結果が曖昧になることと、(ii) 楽曲の解釈自体曖昧性が内在していることで、分析結果も曖昧になること、の2つが考えられる。

(2) アルゴリズムの欠落

GTTM のルールには、局所的な構造からより高次の構造を生成するボトムアップ方向に働くルールと大局的な構造をより低次の構造に分解するトップダウン方向に働くルールが混在している。しかしその両者をどのように組み合わせると適切な階層構造を生成するかに関するアルゴリズムが欠落している。

(3) ルールの競合

ルールを適用する際、ルール間での優先度が決まっていないため競合が起きることがある。

- exGTTM の提案と ATTA の実現

GTTM を計算機上へ実装する上での課題は、以下の2つに大別される: (1)すでにルールや概念として GTTM に定性的に記述されているが、計算機上のプログラムとして実現するために定量的に記述しなおすこと、(2)ルールや概念として明示的に記述されていないが、計算機上のプログラムとして実現するために必要なアルゴリズムやパラメータを発見し補うこと。

まず(1)に関して、パラメータの存在は示されているがその値が不明な場合は、パラメータの値域を定義することにした。音楽の解釈は曖昧なので、一般に、人間が GTTM に従って分析した結果は唯一に決まらない、よって我々の分析器が、人間の分析結果全体を包含すれば十分であると考え、そのような分析結果を生成できるように値域を設定した。

(2)に関して同様に、その値を調整すれば人間の分析結果を包含できるような新しいパラメータとその値域を導入した。

具体的には、楽曲の正しい解釈は複数あるという前提のもとで、以下に述べる2つの方針に

基づき、計算機上で実行可能となるよう GTTM の理論を拡張した exGTTM を設計した。まず、1 目の方針は、パラメータを導入して曖昧さを出来限り排除することである。exGTTM を計算機上に実装した ATTA では、それらのパラメータを手動で動かすことによって、楽曲の多様な解釈をすべて出力できるような分析器を実現した。exGTTM に導入したパラメータは以下の 3 つのカテゴリーに分類できる。まず第 1 のカテゴリーは、GTTM でその存在は明らかであったものの具体的な値が与えられていなかったパラメータである。例えば、あるルールが成立するか否かを、 D_{rule} というパラメータであらわし、ルールが成立すれば 1 に、不成立ならば 0 に対応付けた。第 2 のカテゴリーは、GTTM ではその存在自体が暗黙であったものを明示化したパラメータである。この例として、GTTM の各選好ルールの強さを重みづけするパラメータが挙げられる。GTTM ではルールの適用に際して競合が起きることは認識されており、何らかの方法で解決しなければならなかった。このパラメータを導入することによって、ルールの強さの制御が可能となり、競合を解消することが可能となった。第 3 のカテゴリーとして、GTTM ではその存在自体が議論されていなかったパラメータがある。この例として、並列性に関するルールにおいてフレーズ間で並列性がどの程度で成立するかを表すパラメータが挙げられる。GTTM では並列性について詳細な定義が与えられていなかった。2 目の方針は、詳細な定義が無いまま GTTM で用いられている用語に直観的でわかりやすい定義を与えることである。そして、局所的な構造に関するボトムアップなルールと、大局的な構造に関するトップダウンのルールを組み合わせ、階層的な構造を獲得する手法を提案した。

もし exGTTM が、楽曲の正しい分析の結果を出力できない場合、我々は新たなパラメータを exGTTM に追加することで、正しい分析結果が出力できるように試みた。このようにして、外在化およびパラメータ化を繰り返し、最終的に、グルーピング構造分析に関する 15 個のパラメータ、拍節構造分析に関する 18 個のパラメータ、タイムスパン簡約に関する 13 個のパラメータを導入した。

- FATTA の実現

ATTA で正しい分析結果を得るためには、46 個の調節可能なパラメータを手動で適切に設定しなければならず、多大な労力がかかるという問題があった。そこで、ATTA の分析の結果得られるタイムスパン木を用いて、タイムスパン木の安定度および拍節構造の安定度を評価するアルゴリズムを与えることで、ATTA の各パラメータの値を自動的に最適化する手法を提案した。その際用いるタイムスパン木の安定度を定義するにはいろいろな尺度が可能であるが、調とコード進行の安定性という観点に立ち、GTTM と同著者による音楽理論 Tonal Pitch Space に基づき定義した。一方、拍節構造の安定度は、タイムスパン木と整合性の高い拍節構造ほど安定であると考え、タイムスパン木と拍節構造の整合性の高さを安定度として定義した。

- 基盤技術としての意義

第一に、GTTM に基づく音楽的な構造の自動獲得できるようになったこと自体が意義である。従来、GTTM に基づく音楽の意味構造の分析や、演奏の表情付け、音楽要約システムの研究がなされてきたが、それらのシステムの動作には GTTM の分析結果が必要であり、音楽家が作成した GTTM の分析結果を用いていた。本研究により、GTTM に基づく音楽的構造の獲得が自動化され、それらの音楽システムへ応用されることが期待される。

第二に、メロディの簡約、モーフィング、予測など、音楽家が行う高度な音楽的活動が音楽理論に基づき計算機上で実現できることを示したことである。

第三に、音楽知識の乏しいユーザへ音楽理論に基づいた支援の可能性を示したことである。次頁で述べる、予測ピアノおよびメロディモーフィング手法はその一例である。

図2は、メロディの簡約の例である。図のメロディAの上にある木構造は、FATTAによる分析の結果得られたタイムスパン木である。タイムスパン木のレベルBより下にある枝の音符を省略するとメロディBのようになる。さらに、レベルCよりも下にある枝の音符を省略するとメロディCのようになる。

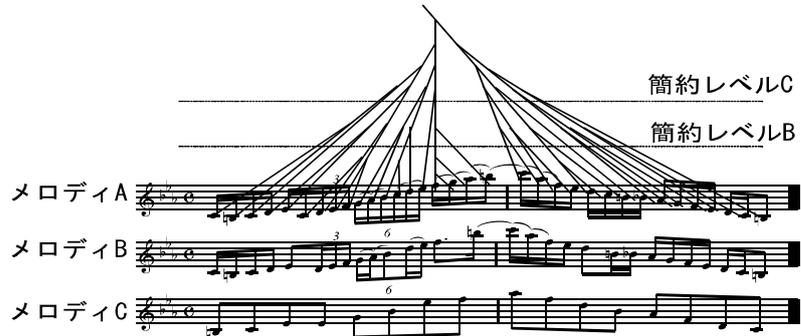


図2: タイムスパン木を用いたメロディの簡約

図3は、FATTAに基づくメロディの予測手法を用いて、演奏者の次の音を予測し天板上に結果を表示する「予測ピアノ」というシステムである。天板上の表示は時間と共に縦スクロールし、明るいところほど次の音が鳴りやすい位置である。素人でも明るく表示されている位置の鍵盤を押していくことで即興演奏が可能となる。

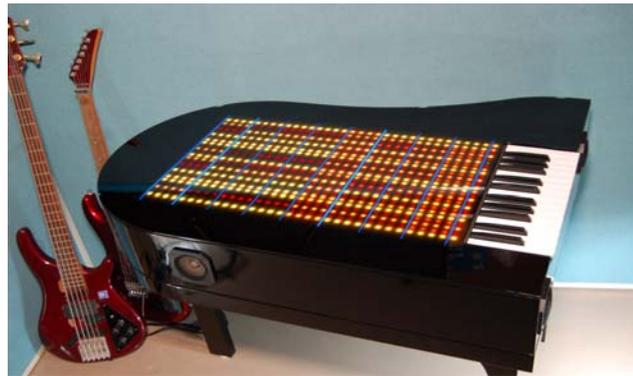


図3: 予測ピアノ

2) メロディモーフィング手法の構築

あるメロディと別のメロディの間にある複数のメロディをある尺度のもとで順序付けて生成するメロディモーフィング手法を構築した(図4)。



図4: メロディモーフィング手法

メロディモーフィング手法は、音符や休符などメロディの表層的な構造を操作できない音楽初心者が自分の意図を反映させたメロディを生成させる際に有効である。たとえば、ユーザがメロディAの一部を修正し何らかのニュアンスを付加したいとき、ユーザはそのようなニュアンスを持つメロディBを知っているとす。このような場合、メロディモーフィング手法を用いると、ユーザは「メロディAにメロディBのニュアンスを付加せよ」と指示するだけで、ユーザの意図を簡易かつ的確に反映したメロディの生成が可能である。メロディモーフィングには、上記の簡易かつ的確という利点の他に、システムの入力と出力の因果関係の理解が比較的容易、システム操作が簡便という利点がある。

3) 能動的音楽鑑賞インタフェースの構築

音楽初心者でも複数のパートの聴き分けを容易とするシステム、「サウンドスコープヘッドフォン」を構築した。サウンドスコープヘッドフォンは、直感的な操作で複数パートのミキシングを自由に変更できる能動的音楽鑑賞インタフェースである。頭を上下左右に振ったり、手を耳に近づけて耳を済ませるようなポーズをとるなど、人間が音を聴くときに自然に行う動作をヘッドフォンに搭載した地磁気センサ、傾斜センサ、距離センサで検出することで音楽用ミキサーのコントロールを可能とした(図5)。

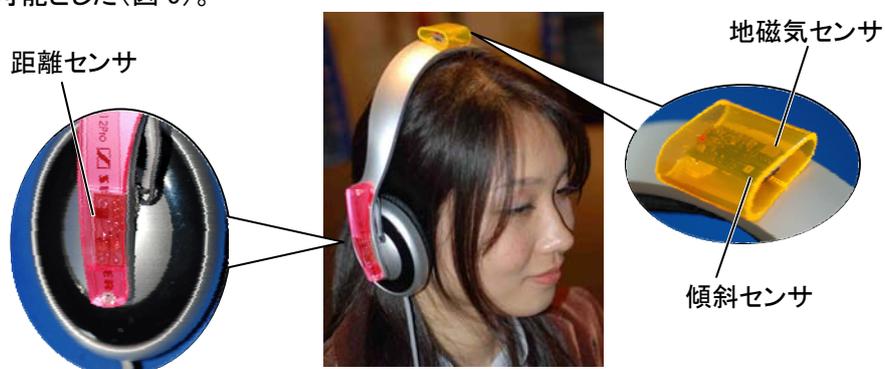


図5: サウンドスコープヘッドフォン

図6は、サウンドスコープヘッドフォンの展示風景で、各パートの楽器を周囲に配置し、音量に応じて楽器を照らす明るさを変化させることで、視覚的にも音量の変化を確認できるようにしたものである(論文2参照)。



図6: サウンドスコープヘッドフォンの展示風景

5. 自己評価

音楽理論 GTTM のタイムスパン木の自動獲得は、理論が考えられてから 20 年以上困難とされてきた問題であり、その実現は音楽情報処理の研究分野に新たな展開を与えるもので、非常に大きな貢献ができたと考えている。当初実装を行っていた、GTTM に基づく音楽分析器 ATTA(Automatic Time-span Tree Analyzer)は、予定通り約 1 年で完成することができた。しかし、ATTA はユーザが手作業で 46 個のパラメータを調節する困難な作業が必要であったため、さらに 1 年間改良を進め、パラメータを自動で最適化する FATTA(Full Automatic Time-span Tree Analyzer)を構築した。分析器の構築に、当初の予定より 1 年長くかかってしまったが、予想以上の成果を上げることができた。そして、音楽家が行う高度な音楽的活動の例として、メロディの予測およびメロディのモーフィングの 2 つを考え、それらを FATTA に基づき実現することで、FATTA の有用性を示すことができた。今後、FATTA を用いて作曲・編曲など様々な音楽的操作を実現していく予定である。

サウンドスコープヘッドフォンは、一般の人が能動的に音楽を楽しむためのデバイスとして構築したものであるが、文化庁メディア芸術祭協賛展など展示の機会を多く得ることができたため改良を重ねることができた。子供やお年寄りなど、当初は想定していなかった年齢層のユーザでも容易に使用できるデバイスを構築することができた。

本研究でやり残した課題は、これまで開発してきたシステムを統合し、実際に仮想演奏者を実現することである。その実現には、複数の演奏者から得られる情報の統合など難しい問題を解決する必要がある。今後、それらの問題を解決し、自律的に演奏する仮想演奏者の実現に向けた研究に邁進していくことを予定している。

6. 研究総括の見解

音楽はデジタルメディアにとって重要な対象である。今回の研究では、音楽家の音楽知識を体系化した GTTM 理論を、PC 上に実装した自動楽曲分析器の開発を目指した。

音符、休符のように曖昧性がないものは、既に市販の楽譜エディタなどで計算機に実装されている。しかし、旋律、リズム、和声のような高次の構造に対しては、GTTM が音楽理論としてあったが、計算機上には実装することが困難であった。旋律・リズム・和声は GTTM によって分析しても、複数の解釈が可能になるという曖昧性があることなどがその理由である。今回、曖昧性を積極的に認め、パラメータを調整し、複数の解釈のうち人に安定な解釈を優先するというルールをもうけた。そしてこれらを PC 等の計算機上に実装したシステム (FATTA) を完成した。さらに成果を応用しメロディ予測やメロディモーフィングのシステムまで完成した。このことは高く評価されるし、論文が ICMC(コンピュータ音楽国際会議)において受賞されるなど、専門家にも評価されるものである。また、サウンドスコープヘッドフォンはこうした技術を一般の人に分かりやすい形で提示しものであり、多くの展示会でも好評であった。

今後はこうした成果を他の人に具体的に使うための工夫(公開方法やインターフェース改良など)を行い、「ドレミっち」のネーミングとともにこの分野の基盤技術となっていくことが期待される。

7. 主な論文等

(1) 論文(原著論文)発表

A. さきがけ研究者が主導で得られた成果

(1) 論文(原著論文)発表

1. Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, Implementing “A Generative Theory of Tonal Music”, *Journal of New Music Research*, Vol. 35, No. 4, pp. 249–277, 2006.
2. 浜中雅俊, 李 昇姫: サウンドスコープヘッドフォン, 日本バーチャルリアリティ学会論文誌, Vol. 12, No. 3, pp. 295–304, 2007.

3. 浜中雅俊, 平田圭二, 東条敏: 音楽理論 GTTM に基づくグルーピング構造獲得システム, 情報処理学会論文誌, Vol. 48, No. 1, pp. 284-299, 2007.
4. Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: FATTA: Full Automatic Time-span Tree Analyzer, Proceedings of the 2007 International Computer Music conference (ICMC2007), Vol. 1, pp. 153-156, August 2007.
5. Masatoshi Hamanaka, Seunghee Lee: *Music Scope Headphones: Natural User Interface for Selection of Music*, Proceedings of the 2006 International Conference on Music Information Retrieval (ISMIR2006), pp. 302-307, September 2006.

(2) 特許出願

発明者: 浜中雅俊、池月雄哉
 発明の名称: ミキシング装置及び方法並びにプログラム
 出願人: 科学技術振興機構
 出願日: 2005年6月21日
 出願番号: 特願 2005-18772

発明者: 浜中雅俊
 発明の名称: ネットワークを用いた遠隔多地点合奏システム
 出願人: 科学技術振興機構
 出願日: 2005年8月3日
 出願番号: 特願 2005-225878

発明者: 浜中雅俊、東条 敏
 発明の名称: 楽曲のタイムスパン木の自動分析方法および自動分析装置
 出願人: 科学技術振興機構
 出願日: 2005年10月3日
 出願番号: 特願 2005-289732

発明者: 浜中雅俊、李 昇姫
 発明の名称: 音源選択装置
 出願人: 科学技術振興機構
 出願日: 2006年9月29日
 出願番号: 特願 2006-269623

発明者: 浜中雅俊
 発明の名称: モーフィング楽曲生成装置及び該装置用プログラム
 出願人: 科学技術振興機構
 出願日: 2008年2月5日
 出願番号: 特願 2008-25374

(3) 受賞

JNMR (Journal of New Music Research) Distinguished Paper Award for Best Paper in the International Computer Music Conference 2005 (H17.9)

(4) 招待講演

ISMIR2007 チュートリアル開催(タイトル: Techniques for Implementing the Generative Theory of Tonal Music)