

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 23 日現在

機関番号：14301

研究種目：若手研究(A)

研究期間：2013～2016

課題番号：25700036

研究課題名(和文)音楽音響信号に対する音楽操作の蓄積と再利用の実現に関する研究

研究課題名(英文)Research on accumulation and reuse of musical operations

研究代表者

浜中 雅俊 (Hamanaka, Masaotshi)

京都大学・医学(系)研究科(研究院)・研究員

研究者番号：30451686

交付決定額(研究期間全体)：(直接経費) 8,200,000円

研究成果の概要(和文)：これまで音楽家が楽譜に行ってきた高次の音楽操作を一般ユーザに可能にすることを目指す。それを可能とするため我々は音楽理論GTTM (A generative theory of tonal music)に着目している。GTTMの分析の結果求まるタイムスパン木は、楽曲の要約や表情付け、生成、メロディモーフィング、作曲支援に用られる。しかし有効に利用するためには、高精度なGTTM分析器が必要である。本研究では、Deep learningに基づくグルーピング・拍節構造器を構築したところ音楽家に分析結果と高い精度で一致していた。今後、deep learningに基づくタイムスパン木分析器を構築する。

研究成果の概要(英文)：Our goal is to create a system that will enable a musical novice to manipulate a piece of music, which is an ambiguous and subjective media, according to his or her intentions. The main advantage of analysis by a GTTM is that it can acquire tree structures called time-span trees. A time-span tree provides a summarization of a piece of music, which can be used as the representation of an abstraction, resulting in a music retrieval system. It can also be used for performance rendering and reproducing music. The time-span tree can also be used for melody prediction and melody morphing. These systems need a GTTM analyzer that enables us to output the results obtained from analysis that are the same as those obtained by musicologists. In this study, we developed grouping structure and metrical structure analyzer based on deep learning. Experimental result shows that the analyzer shows high performance. We plan to implement time-span reduction analysis on the bases of deep learning.

研究分野：音楽情報知能

キーワード：音楽理論GTTM 音楽構造分析 グルーピング構造 拍節構造 タイムスパン木 メロディモーフィング

1. 研究開始当初の背景

音楽というメディアの認識や表現は曖昧なため、音楽知識が乏しいユーザが思い通りに計算機に作曲させたり演奏させたりすることは一般に困難である。本研究の最終的な目標は、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。

音楽知識が乏しいユーザが操作可能な音楽システムを実現する上で重要なのは、1) 音楽をいかにして操作するのか、2) ユーザの意図をいかにして反映するのか、の2つであると我々は考えている。その際注意すべきなのは、操作対象の抽象度を上げると操作は容易になる反面、ユーザの意図を反映しにくくなる可能性がある点である。たとえば、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、和音名などあいまい性の低い表層的な構造に限定されている。したがって、音楽知識が乏しいユーザがそれらの構造を適切に扱うことは困難である。一方、Gragebandなどは、システムがあらかじめ多くのループ素材を用意することで、それを組み合わせるといったシンプルな操作のみで作曲を行うことができるが、作った曲のメロディの一部を修正したいと考えた場合には、手で音符や休符など表層的な構造を操作する必要があるため、音楽知識が乏しいユーザがその意図を反映することは困難である。

そこで、我々は音楽家の音楽操作を一般ユーザでも可能とするツールを構築することを考えた。そのようなツールが実現すれば、音楽初心者の音楽制作を支援するだけでなく、ゲームや映像の挿入歌など多くの楽曲を制作しなければならない音楽の専門家にとっては、ツールを用いて音楽制作を効率的に行うことが可能となる。図1はそのようなシステムのイメージである。元のメロディ（黒く反転）に対して編集を加えたメロディが複数表示され、ユーザはそのなかから好みのものを選択することで複雑な編曲作業を省略することができる。

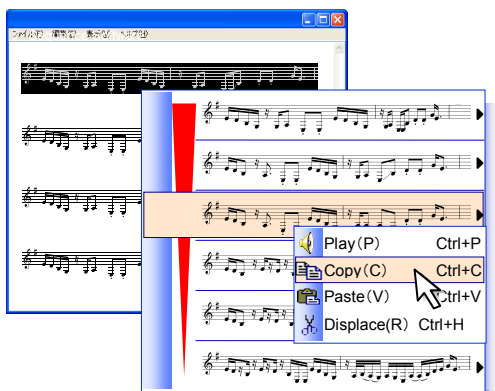


図1: 作曲・編曲支援システム

2. 研究の目的

本研究では、これまで音楽家が楽譜に対して行ってきた高次の音楽操作を一般ユーザに可能にすることを目指している。そのようなことを可能とするため我々は音楽理論 GTTM (A generative theory of tonal music) に着目している。

音楽理論 GTTM の特徴は、音楽が備える多様な側面を包括的に表象しているという点である。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標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つの側面に関して一貫性のある操作を実現する必要があると考える。たとえば、楽曲を2つに分割するという単純な操作を考えたとき、着目する音楽的な構造によってその操作の実現は異なってくるが、装飾が付いた楽曲とそうでない楽曲に対して2つに分割する箇所は本質的に同じであることが望ましい。GTTM では、メロディの区切りを表現するグルーピング構造とリズムや韻律を表現する拍節構造をもとに、メロディや和声を本質的な部分と装飾的な部分に区別するタイムスパン木を抽出する手順が提案されている。GTTM に従えば、メロディ、リズム、和声という3つの側面に関して一貫性のある操作の実現が期待できよう。

一方、「ユーザの意図をいかにして反映するのか」に関して、メロディを生成する場合を考える。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時、ユーザはそのようなニュアンスを持つメロディ B を知っているとする。この時、ユーザがシステムに対して「メロディ A にメロディ B のニュアンスを付加せよ」と指示できれば、ユーザの意図を簡易かつ的確にシステムに伝達することができよう。ここで、メロディ A にメロディ B のニュアンスを付加して少しずつメロディ B に近づけていく操作はモーフィングと呼ばれる。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もある。

GTTM の分析の結果求まるタイムスパン木は、楽曲の要約や表情付け、生成、メロディモーフィング、作曲支援に用いることができる(図2)。これらを実現するシステムを有効に利用するためには、GTTM に基づく楽曲分析を精度高く自動で行う必要がある。

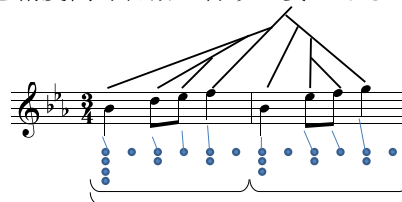


図2: タイムスパン木

3. 研究の方法

これまで多くの音楽理論が提案されてきたが GTTM は、複数のルールによって緻密に構成されており、計算機上への実装が最も期待されてきた音楽理論である。

我々はこれまで音楽理論 GTTM に基づく楽曲分析器を構築してきた。我々は、2004年にグルーピング構造分析器および拍節構造分析器を構築し、それらを統合したタイムスパン木分析器 ATTA (Automatic Time-span Tree Analyzer) を 2005年に構築した。ATTA は、46個の調節可能なパラメータを持ち、それらを適切に調整すればグルーピング構造では8割弱、拍節構造では9割強の正解率、タイムスパン木では、6割の正解率であった。ATTA は GTTM のルールの優先順序を適切に適用すれば、正解率の高いグルーピング構造や拍節構造を獲得できることを示した点で意義があるが、その一方で、パラメータを調整するためには、音楽的な知識が必要で、音楽家でなければ扱うことが困難であった。

2007年に構築した FATTA は、タイムスパン木の安定性に関するルールに基づきタイムスパン木の安定性を定義し、その安定性が高くなるように ATTA のパラメータを自動調整するシステムであった。拍節構造分析の正解率は9割程度であったが、グルーピング構造分析およびタイムスパン簡約の正解率は5割弱で、音楽家の手作業による修正を行わなければ分析結果を利用することは難しかった。

誤りのない完全な分析結果を出力する分析器を構築することが困難であったことから、2009年に構築した Interactive GTTM analyzer では分析器と手動による編集をシームレスに行うことを可能にしたものであった。Interactive GTTM analyzer は現在でも分析データの蓄積のために使用されており無料でダウンロードすることができる。

2008年に構築した σ GTTM では、決定木を用いた統計的学習により自動で局所的グルーピング構造の検出を可能にしていた。FATTA より性能が高かったものの、パラメータ調整後の ATTA よりは低い性能であった。

4. 研究成果

2014年に構築した σ GTTMII では、複数の学習済みの決定木を用意し、それらを手動で切り替えることで、ATTA を超える性能を実現した。しかし、適切なグルーピング境界を選択するためには、音楽知識が必要で、音楽家でなければ難しい作業であった。

2015年に構築した σ GTTMIII では、音楽家による楽曲分析結果を確率文脈自由文法 (Probabilistic context-free grammar, PCFG) に基づき統計的に学習することでタイムスパン木の自動分析を可能とした。 σ GTTMIII では、GTTM データベースに収めら

れている 300 曲を教師データとして PCFG の生成規則と生成確率を学習することで、タイムスパン木分析器として最も高い性能を示すことに成功した。また、同じく PCFG に基づくタイムスパン木分析器である pGTTM では、教師なし学習を可能とした。これら、 σ GTTMIII および pGTTM の特長は、タイムスパン木の先端付近では拍節の影響が強く、根の近くでは和声の影響が強くなるなど、階層によるコンテキストの違いを学習可能とした点である。

2016年に構築した Deep learning に基づく楽曲分析器では、さらに高い精度の分析結果を得ることが可能であった。

音楽家による分析正解データの整備も行ってきた。拡張した GTTM によりポリフォニーを分析した結果のデータベースの公開は世界でもこれまでにないため、クロスチェックの実施したのち1年以内の公開を目指している。

5. 主な発表論文等

[雑誌論文] (計 1 件)

①金森光平, 星野准一, 浜中雅俊: 類似楽曲の決定木学習に基づく音楽理論 GTTM のグルーピング構造検出システム, 電子通信学会論文誌, 査読有, Vol.J100-D, No.1, pp.129-139, 2017.

[国際会議発表] (計 13 件)

①Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: “deepGTTM-II: Automatic Generation of Metrical Structure based on Deep Learning Technique”, 13th Sound and Music Conference (SMC2016), , 査読有, pp.221-249, 2016.

②Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: “deepGTTM-I: Local Boundaries Analyzer based on Deep Learning Technique”, 13th International Symposium on Computer Music Multidisciplinary Research (CMMR2016), , 査読有, pp.8-20, July 2016.

③ Eita Nakamura, Masatoshi Hamanaka, Keiji Hirata, Kazuyoshi Yoshii: “Tree-Structured Probabilistic Model of Monophonic Written Music Based on the Generative Theory of Tonal Music”, Proc, 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), , 査読有, pp.276-280, 2016.

④Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: “Sigma GTTM III: Learning based Time-span Tree Generator based on PCFG”, Proceedings of The 11th International Symposium on Computer Music

Multidisciplinary Research (CMMR 2015), , 査読有, pp.303-317, June 16-19, 2015.

⑤Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: “Structural Similarity based on Time-span Sub-trees”, Proceedings of The 5th International Conference on Mathematics and Computation in Music (MCM2015), 査読有, June 2015.

⑥Kouhei Kanamori, Masatoshi Hamanaka, Junichi Hoshino: “Method to detect GTTM Local Grouping Boundaries based on Clustering and Statistical Learning”, Proceedings of the 42nd International Computer Music Conference (ICMC) joint with the 13th Sound & Music Computing conference (SMC), 査読有, October 2014.

⑦Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: “Musical Structural Analysis Database Based on GTTM”, Proceedings of ISMIR 2014, , 査読有, pp.325-330, October 2014.

⑧Sakurako Yazawa, Masatoshi Hamanaka, Takehito Utsuro: “A Novel Approach to Separation of Musical Signal Sources by NMF”, The 12th IEEE International Conference on Signal Processing, 査読有, pp.610-615, October 2014.

⑨ Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka: “Algebraic Mozart by Tree Synthesis”, Proceedings of Joint ICMC and SMC 2014, 査読有, pp.991-997, September 2014.

⑩Sakurako Yazawa, Masatoshi Hamanaka, Takehito Utsuro: “Melody Generation System based on a Theory of Melody Sequences”, International Conference on Advanced Informatics: Concepts, Theory and Applications, 査読有, pp347-352, August 2014.

⑪Sakurako Yazawa, Masatoshi Hamanaka: “Extension of Implication-Realization Model for Subjective Melodic Similarity”, BKN25 Milestones in Music Cognition, 査読有, July 2014.

⑫Masatoshi Hamanaka, Keiji Hirata Satoshi Tojo: “Time-Span Tree Analyzer for Polyphonic Music”, Proceedings of The10th International Symposium on Computer Music Multidisciplinary Research (CMMR 2013) , 査読有, pp.886-893, October 2013.

⑬ Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka: “Cognitive Similarity Grounded

by Tree Distance from the Analysis of K.265/300e”, Proceedings of The 10th International Symposium on Computer Music Multidisciplinary Research (CMMR 2013) , 査読有, pp.415-430, October 2013.

[学会発表] (計 14件)

① 浜中雅俊, 平田圭二, 東条敏: “deepGTTM-II: ディープラーニングに基づく拍節構造分析器”, 情報処理学会 音楽情報科学 研究会 研究報告 2016-MUS-112, Vol.2016, No.5, pp. 1-8, March 2016.

②金森光平, 浜中雅俊, 星野准一: “クラスタリングと機械学習を用いた音楽理論 GTTM に基づく楽曲構造分析システム”, 情報処理学会 音楽情報科学 研究会 研究報告 2016-MUS-110-18, Vol. 2016, No.18, March 2016.

③浜中雅俊: “σGTTM III の構築”, JSAI2015 人工知能学会全国大会第 29 回, 2C5-OS-21b-1, May 2015.

④竹内 星子, 浜中 雅俊, 星野 准一: “音楽理論 GTTM に基づく映画の構造化手法”, 情報処理学会 ヒューマンコンピュータインタラクション研究会, 2015-HCI-162, No. 6, 8 pages, March 2015.

⑤榎部義幸, 瀧田寿明, 浜中雅俊, 矢澤櫻子, 星野准一: “楽曲構造解析への統計的機械学習法の適応”, 信学技報, vol. 114, no. 502, IBISML2014-89, pp. 31-38, March 2015.

⑥浜中雅俊: “音楽理論 GTTM に基づく音楽構造解析研究用データベース”, JSAI2014 人工知能学会全国大会第 28 回, 1K5-OS-07b-5, May 2014.

⑦竹内星子, 浜中雅俊: “音楽理論に基づく映画の構造化”, JSAI2014 人工知能学会全国大会第 28 回, 1K5-OS-07b-4, May 2014.

⑧矢澤櫻子, 浜中雅俊: “音楽理論に基づいたメロディ生成システム”, JSAI2014 人工知能学会全国大会第 28 回, 4G1-5, May 2014.

⑨國澤佳代, 浜中雅俊, 星野聖: “ピアノ演奏時の手指画像を用いた接鍵検出法”, 情報処理学会全国大会, 2014-IPJSJ-76, 5Q-5, 2 pages, March 2014.

⑩竹内星子, 浜中雅俊: “音楽理論に基づく映画の要約映像生成手法”, 情報処理学会全国大会, 2014-IPJSJ-76, 5ZD-7, 2 pages, March 2014.

⑪浅川智瑛, 浜中雅俊: “スペクトル包絡保存に基づく NMF による音源分離”, 情報処理学

会 音楽情報科学研究会, 2014-MUS-102,
No.16, 6 pages, February 2014.

⑫金森光平, 浜中雅俊: “クラスタリングと統計的学習に基づく音楽理論 σ GTTM II:局所的グルーピング境界の検出”, 情報処理学会 音楽情報科学研究会, 2014-MUS-102, No.2, 7 pages, February 2014.

⑬矢澤櫻子, 浜中雅俊: “主観的類似度を反映した暗意実現モデルの拡張”, 情報処理学会 音楽情報科学研究会, 2014-MUS-102, No.1, 5 pages, February 2014.

⑭瀧田寿明, 櫛部義幸, 矢澤櫻子, 星野准一, 浜中雅俊: “深層学習に基づく入力楽器音響信号のMIDI音色パラメータ値推定”, 第13回情報科学技術フォーラム(FIT2014), 2014.

[図書] (計 3件)

① Masatoshi Hamanaka, Keiji Hirata: “Implementing Methods for Analyzing Music Based on Lerdahl and Jackendoff’s Generative Theory of Tonal Music”, In Computational Music Analysis, David Meredith (Ed.), pp. 221-249, Springer, 2016.

②Keiji Hirata, Masatoshi Hamanaka: “An Algebraic Approach to Time-Span Reduction”, In Computational Music Analysis, David Meredith (Ed.), pp. 251-270, Springer, 2016.

③Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka: “Cognitive Similarity grounded by tree distance from the analysis of K. 265/300e”, In Sound, Music, and Motion, 10th International Symposium, CMMR 2013, Marseille, France, October 15-18, 2013. Revised Selected Papers, Lecture Notes in Computer Science (LNCS) Vol. 8905, pp. 589-605, Springer, 2014.

[その他]

ホームページ URL

<http://gttm.jp/>

6. 研究組織

(1) 研究代表者

浜中 雅俊 (HAMANAKA, Masatoshi)

京都大学・医学研究科・研究員

研究者番号: 30451686