

令和 5 年 6 月 18 日現在

機関番号：32617

研究種目：若手研究

研究期間：2019～2022

課題番号：19K20301

研究課題名（和文）音楽メロディのベクトル表現の実現とその有効性の検証

研究課題名（英文）Implementing Vector-Based Representation for Music Melodies and Evaluation of Its Effectiveness

研究代表者

平井 辰典（Hirai, Tatsunori）

駒澤大学・グローバル・メディア・スタディーズ学部・講師

研究者番号：70780542

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では、音楽のメロディを計算機で扱うための分散表現手法の確立とその評価を目的として、メロディのモデル化手法、評価手法、アプリケーションを提案した。メロディのベクトル化の実現を始め、メロディを対象とした客観的かつ定量的な評価尺度の提案、それを応用したメロディの接続コスト算出モデル等を提案した。また、提案モデルを応用した検索インタフェースや創作支援システムを開発し、その有効性の評価を通じてモデルの評価を行った。

さらに、従来の12平均律による音楽のメロディのみならず、12平均律外の微分音についても計算機で扱えるようにするために、微分音のメロディを入力可能なインタフェースについても提案した。

研究成果の学術的意義や社会的意義

本研究では、これまでメロディを扱う上で主流だった記号論的アプローチに対し、大規模なデータに基づく深層学習によってモデル化するというデータ駆動型のアプローチによってより効果的にメロディを扱うための手法を提案した。さらに、提案したモデル化手法を応用し、その有効性を図るために検索や創作支援などの様々なアプリケーションを開発し、その一部については誰もが利用可能なオープンソースソフトウェアとして公開した。このように、基礎研究から始まりながらも、その成果を誰もが恩恵を受けられるアプリケーションの形で公開し、メロディに関する一貫した研究成果を上げることができたという点で意義がある研究であったと言える。

研究成果の概要（英文）：In this study, we proposed a method for establishing and evaluating a distributed representation method for dealing with musical melodies on computers, with the aim of proposing methods for modeling melodies, evaluating them, and implementing applications. Starting with the realization of vectorizing melodies, we proposed an objective and quantitative evaluation criteria for melodies, a model for calculating connection costs of melodies using the evaluation scale, among other things. Furthermore, we developed melody search interfaces and creative support systems for melody using the proposed models, and evaluated their effectiveness. In addition, not only to handle melodies based on the traditional 12-tone equal temperament, but also to handle microtonal melodies outside the 12-tone equal temperament, we proposed interfaces that could handle the input of microtonal melodies in order to make them manageable on computers.

研究分野：音楽情報処理

キーワード：音楽情報処理 分散表現 メロディ検索 音楽生成

## 1. 研究開始当初の背景

音楽のメロディは、楽曲を代表する最も重要な要素であり、鼻歌ではメロディを口ずさみ、カラオケでもメロディを歌う。そのような重要な要素であるメロディを計算機で扱う上では様々な課題があり、これまでに長きにわたり研究が行われてきて様々な手法が提案されてはいるが、確固とした手法は未だに提案されていない。

メロディを計算機で扱うことを目的としたこれまでの研究における代表的なアプローチはメロディを記号と見なす記号論的アプローチである。メロディの特徴を考慮するための代表的な記号論的アプローチの手法である **IRM** は、メロディを3音ずつの音符でグループ化し、それらの音程に応じたシンボルをグループ毎に付与する。この手法では、メロディを3音毎に分割するため、グループを跨いだ音の音程などについての関係性を記述できないなどの欠点がある。このように、記号論的アプローチではメロディの詳細は量子化された記号情報で表現されてしまうため、楽曲同士の関係を詳細に表現できるような数値的な類似度の算出には不十分である。一方で、類似度算出等の数値的処理には、信号処理によるアプローチがとられることが多い。信号処理によるアプローチでは音声波形を対象とした処理を行うことで、楽曲の音色等の類似度を用いた音楽ジャンル推定や類似度算出などのタスクで成果が挙げられている。しかし、混合音信号である音楽においてメロディのみを扱うには、音源分離やメロディ推定等の困難な前処理を経る必要があり、メロディに関する信号処理的アプローチはある程度対象を限定しなければ困難である。

音符を基準にメロディを扱う記号論的アプローチをとりながら、信号処理的アプローチと同等の数値表現を実現できれば、より効果的に計算機でメロディを扱うことができると期待される。本研究では、メロディを計算機で扱うための新たなアプローチとしてメロディをベクトル化するためのモデル化手法を提案する。特に、研究開始当初にはそれほど多くの例がなかった大規模なメロディデータを対象とした深層学習のアプローチを採用することで、近年成果を上げている機械学習技術の音楽のメロディへの応用手法について研究する。

## 2. 研究の目的

本研究では、音楽のメロディをベクトルとして表現するための分散表現手法を実現するために、音楽と同様に数値的に表現することが困難であった自然言語の「単語」をベクトル化する手法として大きな成果を上げた **word2vec** を音楽へと応用する。自然言語処理分野で成功している手法を参考に、機械学習に基づくメロディのベクトル化手法を確立し、それを応用することでメロディの検索やメロディの創作支援などのアプリケーションの実現を目指す。また、メロディベクトルの活用により、従来の主流アプローチである確率モデルに基づく自動作曲手法とは異なる新たなアプローチによる自動作曲や作曲支援技術の可能性についても検証する。

また、メロディのモデル化を通して音楽の検索や創作支援などに関する様々なアプリケーションを実装する。その後、実装したアプリケーションの評価を通じてメロディのベクトル表現が音楽情報処理分野の各研究タスクに対してどのような有効性を示すものなのかを検証する。

## 3. 研究の方法

メロディをベクトルとしてモデル化する分散表現の実現のためには、大規模なメロディデータを対象とした機械学習によるアプローチが必要となる。これまでの研究で扱ってきた大規模な **MIDI** データを対象とした機械学習のアプローチを踏襲し、その対象をメロディとして分散表現を獲得するためのモデルの構築を図った。

さらに、当初想定していた研究の方法として、実現したメロディベクトルの有効性を図るために、人を対象としたメロディベクトルの評価を予定していた。しかし、本研究計画実施期間の大半の期間がコロナ渦の影響を受けてしまった都合から、大人数を対象とした評価実験の実施が困難な状況となってしまった。そこで、研究のアプローチを変え、なるべく人を対象とせずとも有効性が評価できるようなメロディに関する評価指標について検討した。具体的には、客観的な指標によって精度を数値化可能なメロディの接続の有無に関する評価指標を検討する。また、それを応用して実現した様々なアプリケーションなどの要素に対する主観評価実験を行い、当初の計画で想定していた大人数を対象とした評価実験を行わずにメロディのモデル化の精度や有効性を評価した。

分散表現獲得のためのメロディのモデル化やメロディのフレーズのモデル化の他にも、メロディ素片同士の接続の自然さを評価するための深層学習モデルなど、メロディを対象とした多面的なモデルを構築し、その評価を行った。本研究では、メロディという対象を計算機によって様々な角度から扱うための方法について検討する。

#### 4. 研究成果

##### (1) メロディベクトルの実現

本研究の最も基本的な研究成果として、楽曲のメロディ全体のベクトル化を実現するためのモデル化手法の確立が挙げられる。これまでの研究では、メロディを短い素片に分割したフレーズ単位でのベクトル化が実現していたが、本研究では `doc2vec` という文章全体の分散表現を実現する手法をメロディへと応用した。これにより、楽曲一曲単位でメロディ全体の特徴を反映したベクトル表現を獲得する手法を実現した。このベクトルを応用することで、楽曲検索を含むメロディの特徴に基づく様々なアプリケーションの開発が可能となった。その一例として、図1のメロディ探索インタフェースでは、メロディのベクトルに基づきメロディを二次元平面へとプロットし、一つ一つのメロディを聴きながらユーザがそのメロディを気に入ったかどうかをフィードバックしながら探索できるものとなっている。メロディベクトルの距離が近いメロディ同士はこの平面上でも近い位置にプロットされるため、ユーザがメロディ平面内を探索しながら、好みのメロディと好みでないメロディをラベリングしながら気に入ったメロディを探すことができるインタフェースとなっている。図1では、ユーザが付与した複数のラベルデータを基に、ユーザが気に入ったメロディの近くに位置するメロディの領域をヒートマップによって平面上に可視化した例である。

メロディベクトルは、楽曲の音楽的な特徴を低次元の数値データで効率的に表現したものであり、その実現方法は一通りではない。本研究課題で実現した手法の他にも、今後、新たなメロディのベクトル化手法についても研究を進めていく予定である。その実現により、本研究課題でこれまでに開発してきたメロディベクトルの応用事例をさらに高精度に実現できることが期待される。

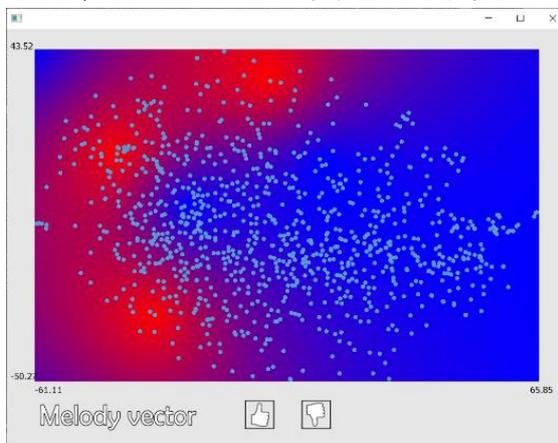


図1 メロディ平面の探索

##### (2) メロディを対象とした機械学習モデルにおける定量的な評価尺度の提案

本研究の過程でメロディ同士の類似度を算出する尺度をいくつか提案してきたが、それが人間の聴感上の感覚と一致するかどうかを評価することは困難な課題であった。このような人の主観に関わる要素が関係する音楽のモデル化において、その妥当性を一貫性のある指標によって評価することはこれまでも大きな課題であった。多くの場合、人を対象とした評価実験を行うなどのアプローチが行われてきたが、多くの人数を集めての主観評価実験の実施が難しい情勢となったことも影響し、本研究ではメロディを評価するための新たな定量的な指標を提案した。

本研究で新たに提案したメロディに関する定量的な評価指標は、メロディを1小節程度の素片に分割し、その素片同士が元々接続していたメロディ同士なのかを判定可能な機械学習モデルを構築するアプローチである。このモデルによって、任意の2つのメロディ素片に対してそれらが元々接続していたメロディ同士であったのかを推定し、その正解率を精度として判定するというものである。これにより、メロディの接続の有無を推定するタスクを通じて、異なるアーキテクチャのネットワーク同士の精度を同一の定量的な指標によって比較することを実現した。

図2は、本研究で構築した、メロディの素片同士の接続の妥当性(接続コスト)を算出する機械学習モデルの一例である。RNN モジュールの箇所に LSTM や BiLSTM などの時系列データの処理に適したニューラルネットワークを挿入したモデルを提案し、その性能を比較・検証した。このモデルによって、任意のメロディのペアを入力することで、そのメロディ同士の接続の妥当性が判定できる他、任意の入力メロディに対して、最も接続が自然な別のメロディ素片を探索することもできる。

ここで提案したメロディに関する定量的な評価尺度の他にも、定量的かつ客観的にメロディを評価可能な手法が存在することが考えられる。今後、新たな客観的評価手法についても研究を進めていくことにより、定性的な評価が主であったメロディに関する情報処理技術を、データのみに基づいて評価可能な形とし、より一層研究を発展させていきたいと考えている。

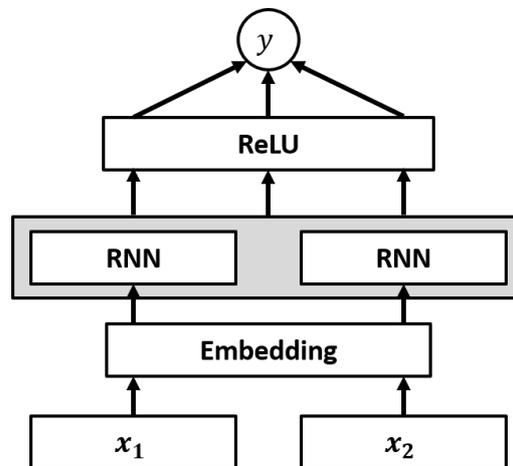


図2 メロディの接続コストを算出するモデル

### (3) メロディのモデル化に基づくメロディ創作支援インタフェースの実現

図2のモデルによって算出可能なメロディ同士の接続コストを創作支援インタフェースへと応用した。図3は、ピアノロール（コンピュータ上で音符を入力するためのインタフェース）上にユーザがメロディを入力すると、そのメロディに自然に繋がるような後続のメロディフレーズの候補を複数提案してくれるインタフェースである。これにより、ユーザが音楽制作に途中で行き詰った場合でも、自然に繋がるような続きのメロディの候補が推薦される。

このインタフェースに関して特筆すべき点は、後続メロディ候補として提示されるメロディが自動生成されたメロディではなく、すべてデータセット内に含まれる既存のメロディであるという点である。従来の機械学習に基づくメロディ生成技術とは異なり、メロディのすべてを自動生成するのではなく、既存のメロディデータセットの中から接続が自然な素片を検索し、推薦するという形での創作支援となっている。これは、機械による創作ではなく、人間による創作を機械によって支援するというアプローチであり、創作支援技術の中でも機械と人間の役割のバランスがより人間に傾いていると言える。今後、生成AI技術が創作の現場で本格的に活用されることが予想される中で、本インタフェースではユーザ主体の音楽創作支援の一形態を示している。

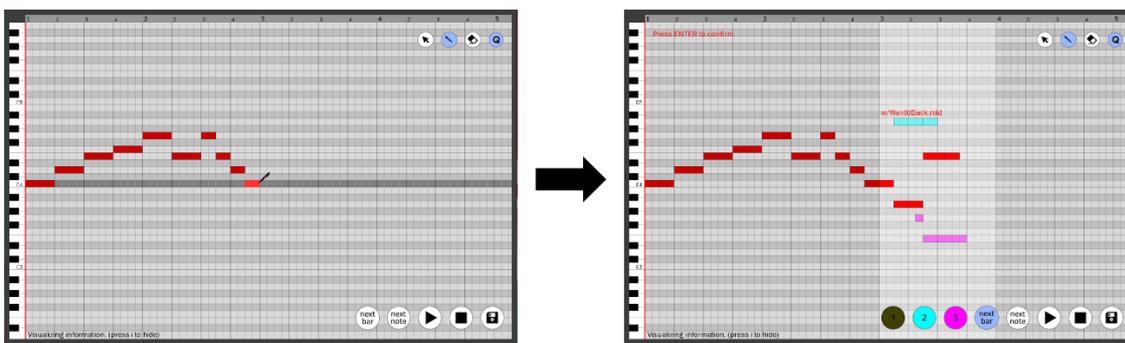


図3 接続コストを用いたメロディの創作支援インタフェース

### (4) 12平均律の枠を超えたメロディを計算機で扱うための基本的なインタフェースの開発

西洋音楽に由来する現在の一般的な音楽では12平均律というドレミを基本とした1オクターブあたり12種類の音高のみを使用する音律が採用されている。一方で、音楽のメロディには12平均律以外の音も使われることがあり、それらの12平均律外の音は微分音と呼ばれている。本研究の当初の適用範囲は暗黙的に12平均律に限定されていたが、対象をより広範な音楽へと拡張させるために、12平均律の枠に収まらない微分音のメロディを計算機で扱うための基本的なインタフェースを開発した。

具体的には、ピアノロールという従来のメロディ入力のためのインタフェースを微分音への入力へと対応させた音律のデザインが可能なメロディ入力インタフェースを開発した(図4)。このインタフェースでは、ユーザが自由に音律をデザインできる。これは、1オクターブあたりの鍵盤の数とそれぞれの鍵盤に対応する音高を自由にデザイン可能なソフトウェア上のピアノに対応する。本研究では、ユーザがデザインした任意の音律を使用したメロディを作成可能とした。

これにより、12平均律以外のより多様な音高を含むようなメロディを計算機によって扱う技術が実現した。微分音を対象とした音楽情報処理技術についてはまだ十分に研究が進んでいな

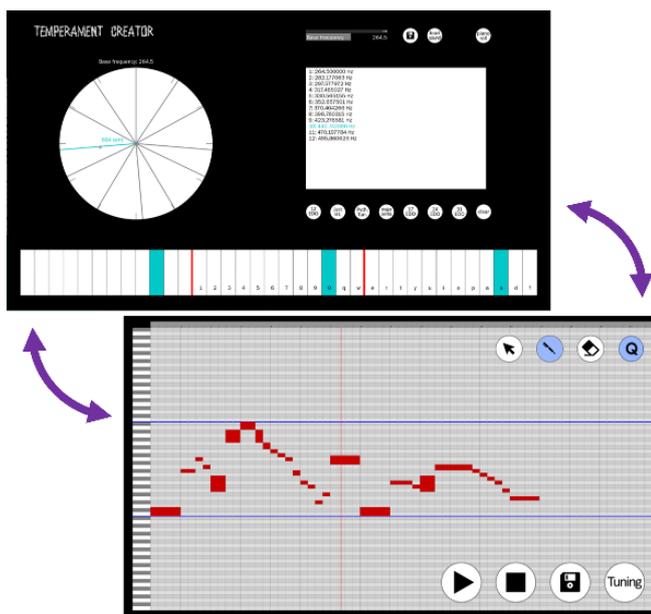


図4 12平均律の枠を超えたメロディを扱えるインタフェース（上：音律デザインインタフェース、下：任意の音律によるメロディが入力可能な拡張ピアノロールインタフェース）

い領域である。本研究成果を足掛かりとして、今後、各種音楽情報処理技術を微分音にも対応させられるような技術基盤の構築に取り組んでいくことを検討している。特に、微分音のメロディのモデル化については、本研究課題では達成ができておらず、その実現に向けて、今後、微分音楽曲データセットの整備をはじめとした研究基盤の構築を行っていきたい。

#### (5) メロディ生成技術に関する VST プラグインの開発

本研究課題の研究成果を一般的な音楽制作ソフトウェアでできるようにすることを目的として、後続メロディ生成に関する技術を VST プラグインとして実装した。本研究課題の成果の中でも、音楽制作において有益な後続メロディを生成するための機械学習モデルに関する技術を VST プラグイン化し、オープンソースソフトウェアとして公開した。本研究では、学習済みのモデルや学習データそのものを公開しないというポリシーで研究成果を公開してきたため、この VST プラグインの開発にあたっては、データそのものの

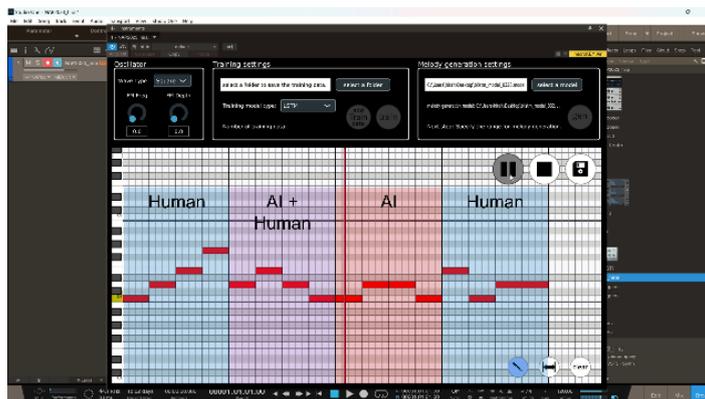


図5 メロディ創作支援インターフェース  
の VST プラグイン化

公開は行っていない。その代わりに、ユーザが自身のメロディを学習データとして自由に追加することができるような新たな学習の仕組みを実装した。ユーザがそれまでの音楽制作中に作成してきたメロディを学習データとして追加していき、一定量のメロディが貯まったところで LSTM や BiLSTM による後続メロディを生成するためのモデルを学習するという流れでの使用を想定している。これによって、本 VST プラグインでは、誰かが作成した生成モデルを活用するのではなく、ユーザのオリジナルのメロディ生成モデルを構築し、活用できる仕組みとなっている。これにより、AI がユーザのコントロールを離れたデータを使って意図しないコンテンツを生成するようなことなく、ユーザの裁量の範囲内で動作する便利なツールとして使えることを目指している。本プラグインは一般的な音楽制作ソフトウェア上で動作し、最低限の動作環境の構築さえ行えば誰もが本研究成果を活用できるものとなっている。

このプラグインで構築した後続メロディ生成モデルを使用した音楽制作では、ユーザが途中で打ち込んだメロディに対して、その続きとなるメロディを、自身のメロディを学習したオリジナルの生成モデルにより生成可能としている。生成されたメロディに対してユーザがさらに修正を加えたり、その続きをさらに生成したりといった、人と AI とのインタラクティブなメロディ制作が可能となった。

本研究課題において VST プラグインを開発して公開した事例はこの一件だけであるが、今後、さらに研究成果を VST プラグインとして実装して公開していくことを検討している。これにより、研究成果が実際に音楽制作の現場で活用可能なアプリケーションとして社会に還元されていくことを目指す。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 0件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Tatsunori Hirai	4. 巻 31
2. 論文標題 A Method for Calculating Melody Concatenation Cost based on BiLSTM	5. 発行年 2022年
3. 雑誌名 Journal of Global Media Studies	6. 最初と最後の頁 55-64
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Tatsunori Hirai	4. 巻 29
2. 論文標題 Comparison of Deep Learning Models for Melody Generation	5. 発行年 2022年
3. 雑誌名 Journal of Global Media Studies	6. 最初と最後の頁 1-16
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計8件（うち招待講演 0件/うち国際学会 1件）

1. 発表者名 長坂らも, 平井辰典
2. 発表標題 微分音楽曲データセットの制作とその評価
3. 学会等名 情報処理学会第85回全国大会
4. 発表年 2023年

1. 発表者名 平井辰典
2. 発表標題 微分音を含む任意の音数の音律のデザインとそれを用いたメロディ打ち込みインタフェースの提案
3. 学会等名 音学シンポジウム2022（情報処理学会研究報告2021-MUS-134(41)）
4. 発表年 2022年

1. 発表者名 平井辰典
2. 発表標題 量子コンピュータによるメロディ生成の試み - 遺伝的アルゴリズムを用いた量子回路デザイナー
3. 学会等名 音学シンポジウム2022 ( 情報処理学会研究報告2021-MUS-134(12) )
4. 発表年 2022年

1. 発表者名 Tatsunori Hirai
2. 発表標題 Redesigning a piano roll: A melody input interface that can play microtones with an arbitrary number of keys
3. 学会等名 Sound and Music Computing Conference (SMC 2022) ( 国際学会 )
4. 発表年 2022年

1. 発表者名 平井辰典, 澤田 隼
2. 発表標題 接続コストに基づく後続メロディ候補の提示によるメロディ打ち込み支援インタフェース
3. 学会等名 情報処理学会研究報告2021-MUS-132(13)
4. 発表年 2021年

1. 発表者名 平井辰典, 澤田 隼
2. 発表標題 LSTMを用いたメロディ素片間の接続コストの算出
3. 学会等名 音学シンポジウム2021 ( 情報処理学会研究報告2021-MUS-131(41) )
4. 発表年 2021年

1. 発表者名 平井 辰典
2. 発表標題 メロディを対象とした生成Deep Learningモデルの比較
3. 学会等名 情報処理学会研究報告2021-MUS-130(15)
4. 発表年 2021年

1. 発表者名 平井 辰典
2. 発表標題 Doc2vecの応用によるメロディ検索
3. 学会等名 音学シンポジウム2020 (情報処理学会研究報告2020-MUS-127(27))
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関