

令和 5 年 6 月 5 日現在

機関番号：12102

研究種目：基盤研究(C) (一般)

研究期間：2020～2022

課題番号：20K12119

研究課題名(和文)クラシックピアノ楽曲の演奏表情モデル生成のための演奏情報解析

研究課題名(英文) Performance Information Analysis for Creating Performance Expression Model of Classical Piano Music

研究代表者

水谷 哲也 (Mizutani, Tetsuya)

筑波大学・システム情報系・講師

研究者番号：70209758

交付決定額(研究期間全体)：(直接経費) 1,600,000円

研究成果の概要(和文)：クラシックピアノ楽曲の演奏表情モデル生成のための演奏情報解析を行うとともに、解析の際に必要な技術について研究を行った。具体的には以下のとおりである。1. 楽曲演奏情報と楽譜情報から演奏表情を抽出するデータ整形システムのプロトタイプを設計した。2. 音声データから音の鳴り始め(オンセット)を正確に検出することに関して計算量を減らす方法についての研究及び堅牢性と正確性に関する研究を行った。3. GTTMやTPSといった精密ではあるが計算量や深い音楽的知識が必要となる音楽モデルに変わる、暗意実現モデルに基づいた演奏表情の解析を行なった。

研究成果の学術的意義や社会的意義

本研究ではクラシックのピアノ音楽に関して、演奏表情を主に強弱と緩急の変化で表現されるものととらえ、その演奏表情モデルを解析した。演奏表情モデルを明らかにすることで、音楽鑑賞および演奏への理解が深まる。また、本研究で特に重要視した「オンセット検出」は、演奏表情を導出するための基本的概念であり、この検出が瞬時に正確に容易に行うことができれば豊かな表情のついた自動演奏などの応用範囲が格段に広がる。

研究成果の概要(英文)：I have analyzed performance information for the generation of performance expression models of classical piano music and studied the technologies required for the analysis. Specifically, I have designed a prototype of a data cleansing system for extracting performance expressions from music performance information and music score information. And I have studied methods to reduce the computational complexity of accurately detecting the onset of sound in waveform data, as well as the robustness and accuracy of the system. I have also analyzed performance expressions based on implicit realization models instead of musical models such as GTTM and TPS, which are precise but computationally expensive and require in-depth musical knowledge.

研究分野：音楽情報科学

キーワード：音楽情報学 演奏表情 演奏分析 オンセット検出 暗意実現モデル クラシック音楽

1. 研究開始当初の背景

本研究の理論的基盤となる、音楽の表情分析および自動演奏生成については広く研究されている。楽曲および演奏の構造的分析は **GTM, TPS** といった認知科学的音楽理論で詳細に分析されている。これらは人による音楽の理解・認知を構造的・階層的に分析するための音楽理論である。この理論の根底には「音楽鑑賞に深い理解のある聴取者は楽曲を構造的・階層的に聴く」という仮定が存在し、そのような音楽の直観を忠実に反映した分析結果を得ることを目的としている。

注意すべきこととして、これらの理論はあくまで音楽の聴取者の理解のモデル化であり、演奏の表情付けのための理論ではない。

研究代表者は、過去の研究において、クラシックにおけるピアノ楽曲を対象に、演奏表情モデルの確立およびそれに基づいたコンピュータによる表情豊かな音楽演奏に関する諸問題は、人間の演奏者による演奏と、あらかじめ自動生成された演奏表情を与えられた計算機による演奏との実時間協調演奏システムの構築において特に顕著に現れることを示した。

すなわち、あらかじめ適切な表情のある演奏データを計算機に与えておけば実時間の協調演奏に際して計算機の演奏と人間の演奏とを同期させるために時間的微調整を行うだけで表情豊かで滑らかな協調演奏が比較的容易に実現できるが、逆にあらかじめ計算機に与えてあるデータの演奏表情が不適切であればシステムをどのように調整しても協調演奏は成功しないこと明らかにした。このように、協調演奏システム構築の上でも適切な演奏表情の重要性が確認された。

ここで、クラシック音楽を題材として選んだのは、演奏に関しての評価(いわゆる「いい演奏」の判定基準)がある程度安定しているからである。

また、ピアノ音楽を題材として選んだのは、演奏表情のパラメータが他の楽器に対して少ないからである。ピアノは「打鍵」する鍵(=音の高さ)、「打鍵」の時刻、「打鍵」の速さ(=音の大きさ)、「離鍵」の時刻およびペダルのタイミングでほぼ全ての演奏表情を表すことができる。特に「打鍵」は音声信号上では音の鳴り始め(オンセット)として検出可能な物理的実体があり、パイオリンやフルートなどを典型例とする他の楽器の音声とは異なる。逆にそれらの楽器では各音の発音中にその音に表情を付加することができるが、格段に情報が多くなり分析が困難になる。

これら一連の研究を通じ、楽曲構造分析と演奏表情との間に密接な関係があることも明らかになった。楽曲分析手法としては上記の認知科学的音楽理論が世界的に広く行われているが、この理論はあくまでも音楽理解のためのものであるため、この理論を用いた表情豊かな演奏表情の研究はほとんど行われていない。

2. 研究の目的

本研究課題では、表情豊かな音楽演奏の解析および生成という研究テーマに関して、演奏表情モデルの構築を目標とした。

楽曲構造を中心とした楽譜情報から演奏表情の傾向を抽出するシステムの開発および実際の演奏表情モデルの提案を行うことを目的とした。演奏データの収集は、既存の音楽演奏データベースを用いるのみならず、**CD**などの波形情報から演奏を抽出する。そのため、波形情報と楽譜情報より楽曲の演奏表情を得るシステムの構築を行う。

演奏表情の生成および演奏表情モデルの確立に当たっては、深層学習や回帰分析などの人工知能的アプローチを使用するものの、ただ単に力づくで演奏を合成するのではなく、どの要素がどのようにどの演奏表情に影響するかという点を重点的に解析する。これにより音楽理論的にも数理的に明快な演奏表情モデルが構築できるものと考えられる。

3. 研究の方法

演奏データと楽曲情報との関連性の研究を深化させる。重回帰分析における説明変数に相当する楽曲情報における演奏表情パラメータを増加させる。そのために表情記号のような演奏表情パラメータ同士の関係を新たに演奏表情パラメータとして採用するなどの方法を用いる。

さらに、適用させる演奏データの追加のために、既存の演奏データベース内のデータの他にクラシック音楽 **CD** やインターネットなどから得られる演奏データを対応する楽譜とともに収集することを目的に、波形データからピアノ音のオンセット情報の検出方法について深く研究を行う。

4. 研究成果

令和2年度は**CD**や**YouTube**などから得られる音声データなどの楽曲演奏情報と**MusicXML**の形式で表現される楽譜情報から演奏表情を抽出するデータ整形システムを構築した。このシステムを用いることにより、既存の音楽演奏表情データベースなどに頼ることなく演奏表情が収集できる。このシステムで得られるデータにより演奏表情解析を行った。また先行研究との解析結果の比較も行った。実験の結果、いくつかの演奏記号が演奏表情であるテンポ変化に影響を与えることが分かった。また先行研究との解析結果の比較では決定係数および偏回帰係数で同

じような傾向の結果が得られた以上を踏まえて本研究で構築したデータ整形システムは有用性があるという結論に至った。

令和 3 年度はこれまで楽曲の認知モデルとして楽曲構造の解析に用いた **GTTM** および **TPS** とは異なる観点での音楽理論である暗意実現モデル(**Implication-Realization Model**)に基づき演奏表情の解析を行なった。本手法では楽曲の各声部を連続する 3 音の音符からなる組の列としてとらえ、その高低の関係で数種類のシンボルに抽象化し、それらが楽曲を演奏する際の意識的・無意識的な情動を表現していることを解析する。従来の解析は精密であるが一方では音楽理論を正確に扱う必要があるのに対して本手法での楽曲の解析は非常に容易である。この手法を用いた演奏の解析の結果を従来の手法による解析と比較したところ、おおむね同程度の結果を得ることができた。

令和 4 年度は **CD** や **YouTube** から得られる音声データからオンセットを正確に検出することに関して、2 方面から研究を行なった。オンセットを正確に検出する方法は従来から行われているが、従来の方法では正確に検出できる反面、計算量が莫大であり、低性能の計算機環境で実行することが困難である。そこで、複数の手法とそれらを融合した **RNN** を利用し、さらにネットワークの学習とピークピッキングを行う新しい手法を提案した。提案手法は、処理速度が半分になる代わりに、従来手法より性能がわずかながら向上しており、精度と速度のトレードオフを考慮した新たな選択肢となり、別の高速なオンセット検出関数を用いることによりさらに精度を上げることが期待できる。一方、検出したオンセットと実際の音声信号との誤差の解析を行なった。3 種類のシンセサイザーとの実演奏データを用いて、音声収録条件やペロシティ情報の事前情報がない場合を模擬した。その結果、実演奏の音量情報とシンセサイザーの選択により、ほとんどの場合、平均誤差が約 2 倍となることがわかった。また、位相に基づくオンセット検出と従来のフレームワークを組み合わせた実用的な試みも行い、異なるタイプのオーディオ特徴を組み合わせることの利点が示された。研究期間全体としては、既存の演奏データベースによるデータを分析することで、実演奏のオンセット情報が正確に行われているという仮定の元でクラシックピアノ楽曲の演奏表情モデル生成のための演奏情報解析は一定レベルで行われた。一方、既存のデータを用いない新たなデータを追加する目的でオンセット検出の方法やさまざまな環境下におけるオンセット検出の精度などの研究成果を上げることができた。

研究期間全体としては、既存の演奏データベースによるデータを分析することで、実演奏のオンセット情報が正確に行われているという仮定の元でクラシックピアノ楽曲の演奏表情モデル生成のための演奏情報解析は一定レベルで行われた。一方、既存のデータを用いない新たなデータを追加する目的でオンセット検出の方法やさまざまな環境下におけるオンセット検出の精度などの研究成果を上げることができた。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Mizutani Tetsuya, Hasegawa Kei	4. 巻 1
2. 論文標題 A Data Cleansing System for Musical Expression Analysis	5. 発行年 2020年
3. 雑誌名 The 8th International Conference on Computer and Communications Management (ICCCM 2020)	6. 最初と最後の頁 111-116
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3411174.3411176	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Mizutani Tetsuya, Sasaki Shigefumi	4. 巻 1
2. 論文標題 A Linear Regression Analysis of Musical Expressions using the Implication-Realization Model	5. 発行年 2021年
3. 雑誌名 The 9th International Conference on Computer and Communications Management (ICCCM 2021)	6. 最初と最後の頁 85-91
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3479162.3479175	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Xuan Mobai, Mizutani Tetsuya	4. 巻 1
2. 論文標題 Peak Picking Multiple Onset Detection Functions Using Recurrent Neural Networks	5. 発行年 2022年
3. 雑誌名 The 10th International Conference on Computer and Communications Management (ICCCM 2022)	6. 最初と最後の頁 31-36
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3556223.3556228	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Ding Yifan, Mizutani Tetsuya	4. 巻 1
2. 論文標題 Audio Feature Extraction for DTW-based Audio-to-Score Alignment	5. 発行年 2022年
3. 雑誌名 The 10th International Conference on Computer and Communications Management (ICCCM 2022)	6. 最初と最後の頁 214-220
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3556223.3556255	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 4件）

1. 発表者名 Mizutani Tetsuya and Sasaki Shigefumi
2. 発表標題 A Linear Regression Analysis of Musical Expressions using the Implication-Realization Model
3. 学会等名 ICCCM '21: The 2021 9th International Conference on Computer and Communications Management (国際学会)
4. 発表年 2021年

1. 発表者名 Mizutani Tetsuya and Hasegawa Kei
2. 発表標題 A Data Cleansing System for Musical Expression Analysis
3. 学会等名 ICCCM '20: 2020 The 8th International Conference on Computer and Communications Management (国際学会)
4. 発表年 2020年

1. 発表者名 Xuan Mobai and Mizutani Tetsuya
2. 発表標題 Peak Picking Multiple Onset Detection Functions Using Recurrent Neural Networks
3. 学会等名 ICCCM '22: the 10th International Conference on Computer and Communications Management (国際学会)
4. 発表年 2022年

1. 発表者名 Ding Yifan and Mizutani Tetsuya
2. 発表標題 Audio Feature Extraction for DTW-based Audio-to-Score Alignment
3. 学会等名 ICCCM '22: the 10th International Conference on Computer and Communications Management (国際学会)
4. 発表年 2022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
---------	---------