

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 7 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2015～2016

課題番号：15K16054

研究課題名(和文)演奏者の音楽性を捉える統計的音楽モデルと演奏・伴奏の自動化

研究課題名(英文) Statistical Music Modelling and Automation Techniques for Expressive Performance and Accompaniment Generation Capturing Performers' Musicality

研究代表者

中村 栄太 (Nakamura, Eita)

京都大学・情報学研究科・特別研究員 (PD)

研究者番号：10707574

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：多様なスキルレベルや演奏表現を含む演奏者の音楽性を捉える統計モデルに基づく演奏モデルを構成し、情報処理技術への応用を研究した。テンポ変動・発音時刻のゆらぎ・消音時刻のゆらぎ・演奏誤りなどを含む詳細な演奏モデルを構築し、自動伴奏・楽譜と演奏のアラインメント・自動採譜をはじめとする複数の応用へとつなげることに成功した。また演奏表情付け・自動採譜・音楽アラインメントなどの研究に必要な、楽譜情報とマッチングがとれた演奏データを収集整備し、今後の研究での利用を可能にした。ジャーナル論文4件、査読有り国際会議論文15件、その他の学会発表38件など多くの成果が得られた。

研究成果の概要(英文)：Statistical music performance models have been constructed to capture various skill levels and characteristics of performance expressions and applied to information processing techniques. Performance elements such as tempo variation, onset and offset time fluctuation, and performance errors have been incorporated in the constructed performance models, and they have been successfully applied to multiple techniques including automatic music accompaniment, music alignment, and music transcription. In addition, performance data with accessible references to the score information, which are necessary in the research fields of expressive performance generation, music transcription, and music alignment, have been collected and set up as database, which were made available for future studies. Achievements include four journal papers, 15 refereed international conference papers, and 38 other conference presentations.

研究分野：音楽情報学

キーワード：音楽情報処理 自動伴奏 自動作曲 音楽アラインメント 演奏モデル

## 1. 研究開始当初の背景

近年、知能情報学の一分野として音楽情報学の研究が増えている。作曲・演奏などの音楽活動を音楽データの情報処理と捉え、この数理モデル化及び音楽技能の自動技術化を目指すものである。研究手法としては、統計モデルやニューラルネットなどを用いた機械学習によるモデル・手法の構築が盛んである。同じく人間の知能に関わる時系列を扱う自然言語処理や音声処理などと共通したモデルが適用されることもしばしばあるが、こうした分野に比べ研究者が利用できる音楽コーパスの量と種類が限られている。一方で、近年の情報処理技術の進化により、従来よりも遥かに効率的なデータ収集・整備が可能な状況にある。

こうした背景もあり、これまで音楽情報学の研究では、人間の音楽知能・技能の一般的な側面を対象とすることが多い。つまり一般的な音楽らしさを記述する作曲モデルや平均的な演奏の特徴を捉える演奏モデルに興味集中している。しかし実際には、音楽には多様なスタイル・音楽性が存在する。演奏を例にとれば、演奏者にはプロからアマまで多様な演奏スキルを持つものがおり、また演奏表情の表現も時代や作曲家に合わせた多様なスタイルが存在する。こうした多様なスタイル・音楽性を記述するモデルが構築できれば、音楽スタイルに一致した演奏生成やユーザの音楽性に呼応した伴奏生成ができる自動伴奏技術など、汎用性の高い応用も可能となると考えられる。

## 2. 研究の目的

本研究では、演奏の音楽情報処理に注目し、統計モデルに基づき演奏者のスタイルや音楽性を記述する詳細な演奏モデルの構築とその応用技術の開発に取り組む。また演奏に関わる研究でモデル学習や評価に必要な多楽曲多演奏者の演奏データの収集と整備を行う。具体的な項目は以下の通りである。

(1) 音楽表現で重要である、テンポ・発音時刻のゆらぎ・音長・装飾音などの演奏要素と、これに加えて実際の演奏で現れる誤りや弾き直しや弾き飛ばしを確率的に記述する統計モデルを構成する。また、このモデルのパラメータや構造により音楽性を表現する枠組みを構築する。特に音楽特有の性質として、複雑な時間構造や同時に複数の声部が進行する多声部構造を記述するモデルの枠組みを構築する。

演奏モデルは、演奏信号を楽譜へと変換する自動採譜問題においても重要な要素である。この際には楽譜における「音楽らしさ」を記述する楽譜モデル(音楽言語モデル)との統合が必要である。この点からのモデル構成技術の開拓も取り組む。

以上のモデルのパラメータをデータから学習する手法、そしてモデルの変数を入力デ

ータから推定する推論手法の開発を行う。

### (2) 情報処理技術への応用

構成したモデルを以下の情報処理に適用し、モデルの検証および応用技術開発を行う。

① 楽譜追跡・演奏のアラインメント：楽譜と演奏のマッチングを自動で行う技術。特に、装飾音や演奏誤りを含む演奏の扱いの改善を行う。

② 演奏誤り認識：楽譜を参照として、演奏データの中の演奏誤りの検出(分類を含む)を行う。

③ 自動伴奏：人間の演奏に対して自動で同期してコンピュータが伴奏・合奏を行う。特に、装飾音や演奏誤りに対応可能とする。また、演奏者の技術に合わせて楽譜の簡略化や音符の補完ができる技術への応用を行う。

④ 自動採譜：演奏信号から楽譜情報を推定する。特に、ビート単位で離散化された時間単位での発音(オンセット)位置と消音(オフセット)位置の認識技術の開発。

### (3) 演奏データベース構築

多楽曲多演奏者の演奏データの収集と整備を行う。研究のためのデータには、電子楽譜データと演奏データそして両者の音符のマッチングデータが必要である。楽譜データでは、最も一般的に用いられている MusicXML 形式による楽譜を扱う。演奏データでは、発音時刻・消音時刻・強弱・ペダル情報などの詳細な演奏情報が記録可能である MIDI 情報による演奏データを収集する。両者のマッピング情報を記録するデータとして、リスト形式による独自フォーマットを導入する。また演奏誤りの記述形式についても検討・提案を行う。

楽譜と演奏のマッピング情報は自動取得できないため、データ整備が必要である。これを効率的に行うため、上記演奏アラインメント技術による自動化技術とこの結果を人手による確認・修正をするために必要なアノテーションツールの開発を行う。

## 3. 研究の方法

(1) モデル構成は、従来より研究を進めている隠れマルコフモデル(Hidden Markov Model; HMM)を基本の枠組みとして進める。楽譜からの演奏信号の生成過程を、楽譜上の演奏進行の生成過程とその結果の実際の音符の生成過程に分けて考え、これらを HMM の遷移確率と出力確率を用いて記述する。音楽演奏の時間構造は、演奏の速度であるテンポにおおよそ比例する要素とトリルやアルペジオや前打音などの装飾音のようにテンポ依存性が弱い要素が存在する。またどちらも実際の演奏ではゆらぎやノイズを伴う。これらを統一的に記述するために、連続時間上の Markov 過程による記述を行う。

演奏誤りや弾き直し・弾き飛ばしは、HMM の遷移確率と出力確率を用いて記述する。

HMM では、和音を一状態として記述するが、演奏誤り認識を行う際には、特に音高誤りについてより精密な記述が必要であるため、音楽家の誤り認識の傾向に基づく改良モデルを構成する。

多声部音楽の演奏では、それぞれの声部間の同期が緩くなるという現象が知られている。これを記述するため、出力合流 HMM に基づく演奏モデルを構成する。これは各声部をそれぞれ HMM で記述し、その出力を合流させることで多声部音楽の演奏モデルを構成するものである。緩い同期を記述するため、テンポは複数声部間で共有するモデル化を行う。また従来は、推論手法としてごく粗い近似による最適化に基づくアルゴリズムしか知られていなかったが、演奏履歴を記述する変数の導入によりよりより近似が良く効率的な推論アルゴリズムの導出を行う。

装飾音の代表であるトリルの演奏では、演奏される音符数や速さが演奏ごとに異なる。HMM ではトリルでの音符生成を、対応する HMM 状態の自己遷移を用いて記述されるが、これではトリル全体の音長や音符数の期待値を記述できない。この問題を解決するため、自己回帰隠れセミマルコフモデル (HSMM) に基づくモデル拡張を行う。

以上のモデルは、テンポ変動の大きさや装飾音の速さ、演奏誤りの頻度などに対応するパラメータを含んでおり、これをデータに合わせて学習することにより、演奏者の音楽性が記述できると考えられる。一方で、演奏表現では演奏要素の詳細な記述が必要であり、楽譜要素の特徴による依存性も組み入れる必要がある。これを記述するモデルとして、楽譜要素をコンテキスト (文脈; 特徴量) として扱い、演奏要素を文脈ごとに分類を行いながら学習をする、文脈木クラスタリングに基づくモデル構成も検討する。

自動採譜を想定したモデルにおいては、楽譜モデルとして音符単位のマルコフモデル (音符マルコフモデル) とビート遷移に基づくマルコフモデル (拍節マルコフモデル) の二種類との結合を考え、定式化する。また多声部音楽に有効なモデルを構築するため、上記の出力合流 HMM に楽譜モデルを組み込んだモデルを構成する。

## (2) 情報処理技術への応用

① 楽譜追跡・演奏のアラインメントに関しては、上記 HMM の推論による定式化を行う。演奏誤りや装飾音を含む演奏データによりその精度を評価する。

また出力合流 HMM を用いた多声部音楽に有効なモデルは、シンプルな HMM に基づくものに比べて推論に多くの計算時間が必要であるという短所がある。これを補うために、演奏誤り認識技術との統合により、必要な箇所のみ出力合流 HMM を自動選択して適用する方法の定式化も行う。

② 演奏誤り認識：演奏アラインメントの結果に基づき、音高誤り・音符脱落・音符挿入を認識できるアルゴリズムを導出する。実際には、これを演奏アラインメント技術の一部として組み込むため、演奏アラインメントの評価は演奏誤り認識精度の評価を含む。

③ 自動伴奏：改良した楽譜追跡技術を自動伴奏技術に応用する。

④ 自動採譜：演奏 MIDI 信号から演奏情報の推定問題 (リズム採譜) を扱い、上記モデルの推論に基づく手法の開発を行う。

## (3) 演奏データベース構築

当初は、演奏データ収録を多く行う予定であったが、研究資金が不足していたため、演奏データ収録は一部に留め、残りの演奏データはインターネット上で収集を行う。楽譜データの収集も、一部は新規に作成し、残りはインターネット上で収集を行うことで効率化を行う。

データ整備は、上記演奏アラインメントをもちいた自動ステップとアノテーションツールを用いた人手による確認・修正ステップにより行う。アノテーションツールの開発では、演奏誤りの表示により効率的に修正箇所が発見できるよう工夫を行う。

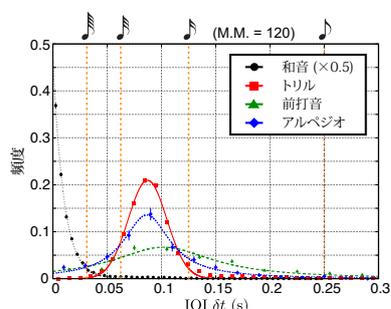
以上のデータとツールは、研究者を対象に無償での配布を行う。

## 4. 研究成果

研究成果は多岐に渡るが、以下では主要な成果をまとめる。論文・学会発表のリストおよび原稿は下記の研究代表者の個人 Web ページから閲覧可能である。

(1) 装飾音と演奏誤りを含む楽譜追跡と自動伴奏 [雑誌論文④・学会発表⑧など]

HMM に基づく装飾音と演奏誤りを含む演奏モデルの構成を行った。装飾音を演奏時の任意性で分類し、それぞれのタイプに対応する HMM 状態を構成した。また時間構造に関しては、テンポを潜在変数とする動的システムで記述し、発音間時間間隔 (IOI) の分布を和音や装飾音ごとに解析し、求めた (図)。



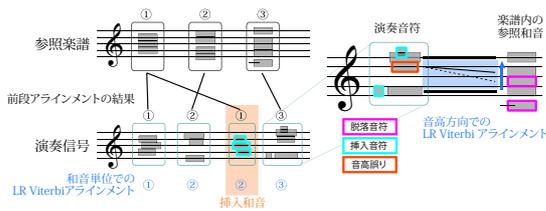
これにより、装飾音の取り扱いがないモデルに比べ最大 12%程度精度向上を確認した。また自動伴奏システムに組み込み、従来はしばしば問題であった装飾音を含む曲の自動伴奏で良好な結果が得られた (Web ページにデモ動画あり)。また HSMM を用い

た改良モデルによりトリルやトレモロが連続する楽譜において顕著な精度向上があった。これは現在世界最高レベルの精度である。

この他、文脈木クラスタリングに基づく演奏モデルの演奏者適応手法の定式化や演奏誤りモデルおよびピアノ運指モデルに基づく演奏難易度の定式化とピアノ楽譜の簡略化と自動伴奏技術への応用も行った。

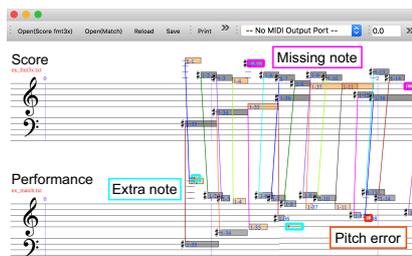
## (2) 演奏誤り認識と演奏アラインメント技術 [雑誌論文④・学会発表③など]

演奏誤り認識については、以下のように和音単位の LR Viterbi アラインメントと各和音ごとに音高方向での LR Viterbi アラインメントを組み合わせる方法を考案した。



これと HMM に基づく演奏アラインメントを組み合わせることで、自動アラインメントの推定誤りのほとんどは演奏誤りの周りで起こることが確認された。これに基づき、演奏誤りの周りで出力合流 HMM に基づくアラインメントを行うことで、高精度かつ高速なアラインメント手法を開発した。この手法によるアラインメント精度は複数のデータで 99% 以上であり、現在世界最高性能であることを確認した。

またこの知見に基づき、自動アラインメント結果と演奏誤り認識の結果を表示して、確認・修正が可能なツールを開発した (図)。



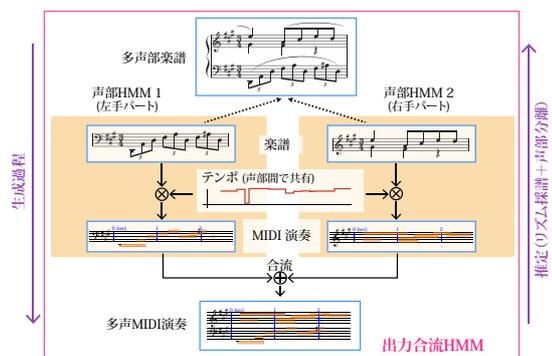
## (3) 演奏データベース構築と公開

上記の演奏アラインメントツールを用いて、演奏データベースの整備を行った。これまでに多様な作曲家・演奏家の音楽を含む、60 曲 (フレーズ) × 3 演奏者以上のデータを整備できた。このデータとアラインメントツールは他の研究者へ公開している。

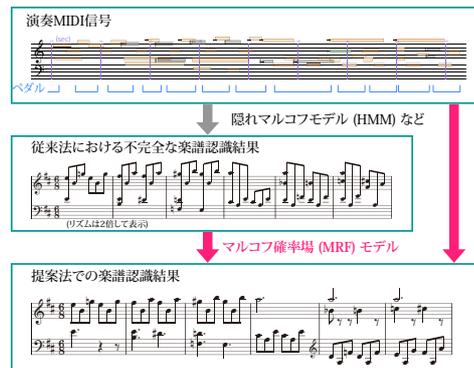
データ収集・整備は続いており、現在研究協力者と共同でさらなるデータベースの拡張のプロジェクトを進めている。このプロジェクトでは、フレーズ構造の情報を含めた楽譜・演奏データの収集も進めているが、データ整備の技術・ツールは本研究で開発されたものが用いられている。

## (4) 自動採譜への応用 [雑誌論文①・学会発表④⑤など]

研究期間の後半では、自動採譜技術に関して多くの進展があった。出力合流 HMM の自動採譜への応用では、声部間で共有されたテンポ変数の離散化と履歴変数を導入することで、計算コストを抑えながらほぼ完全な最適化が可能なアルゴリズムの導出に成功した。またこの結果、特に両手で異なるリズムが現れるポリリズムを含む演奏の採譜性能が大きく向上した。



また楽譜上での音長である音価を認識するには、消音時刻に対応する楽譜時間を推定する必要があるが、消音時間は演奏では大きな逸脱を含むため、演奏モデルのみでは高精度な認識は難しく、適切な楽譜モデルを構成し組み合わせることが必要であるとわかった。そこで認識した発音楽譜時刻の情報を基に消音楽譜時刻の予測分布が得られるモデルを構築し、音価認識に応用した (図)。この結果、従来法より高い精度約 75% の認識率を達成した (投稿論文は現在査読中)。



また音楽では同じフレーズが複数回繰り返されることが多いが、この構造を用いて採譜精度を向上できると考えられる。これを実現するため、階層ベイズマルコフモデルに基づく楽譜モデルと演奏モデルの統合を行い、実際に反復構造の認識と採譜性能の向上が得られることを確認した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

① E. Nakamura, K. Yoshii, S. Sagayama, “Rhythm Transcription of Polyphonic Piano Music Based on Merged-Output HMM for Multiple Voices,” IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 25, No. 4, pp. 794-806, 2017. 査読有 [arXiv:1701.08343]  
DOI: 10.1109/TASLP.2017.2662479

② Miato Ohkita, Yoshiaki Bando, Yukara Ikemiya, Eita Nakamura, Katsutoshi Itoyama, Kazuyoshi Yoshii, “Audio-Visual Beat Tracking Based on a State-Space Model for a Robot Dancer Performing with a Human Dancer,” Journal of Robotics and Mechatronics, Vol. 29, No. 1, pp. 125-136, 2017. 査読有  
DOI: 10.20965/jrm.2017.p0125

③ Tomohiko Nakamura, Eita Nakamura, Shigeki Sagayama, “Real-Time Audio-to-Score Alignment of Music Performances Containing Errors and Arbitrary Repeats and Skips,” IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 24, No. 2, pp. 329-339, 2016. 査読有 [arXiv:1512.07748]  
DOI: 10.1109/TASLP.2015.2507862

④ E. Nakamura, N. Ono, S. Sagayama, K. Watanabe, “A Stochastic Temporal Model of Polyphonic MIDI Performance with Ornaments,” Journal of New Music Research, Vol. 44, No. 4, pp. 287-304, 2015. 査読有  
DOI: 10.1080/09298215.2015.1078819

[学会発表] (計38件)

① Yusuke Wada et al., “An Adaptive Karaoke System That Plays Accompaniment Parts of Music Audio Signals Synchronously with Users' Singing Voices,” Proc. 14th Sound and Music Computing Conference (SMC), 2017/7/5, エスポー (フィンランド).

② Mitsuyo Hashida et al., “Constructing PEDB 2nd Edition: A Music Performance Database with Phrase Information,” Proc. 14th Sound and Music Computing Conference (SMC), 2017/7/8, エスポー (フィンランド).

③ 福田翼ほか, “楽譜簡略化と自動補完伴奏によるピアノ演奏練習支援システム,” 第114回情報処理学会音楽情報科学研究報告, Vol. 2017-MUS-114, No. 21, pp. 1-4, 2017/2/28, ヤマハ本社 (静岡県浜松市).

④ Eita Nakamura et al., “Rhythm Transcription of Polyphonic MIDI Performances Based on a Merged-Output

HMM for Multiple Voices,” Proc. 13th Sound and Music Computing Conference (SMC), pp. 338-343, 2016/9/3, ハンブルグ (ドイツ).

⑤ Eita Nakamura et al., “Rhythm Transcription of MIDI Performances Based on Hierarchical Bayesian Modelling of Repetition and Modification of Musical Note Patterns,” Proc. 24th European Signal Processing Conference (EUSIPCO), pp. 1946-1950, 2016/9/1, ブタペスト (ハンガリー).

⑥ Yasuyuki Saito et al., “Conversion from Standard MIDI Files to Vertical Line Notation Scores and Automatic Decision of Piano Fingering for Beginners,” Proc. 2nd International Conference on Technologies for Music Notation and Representation (TENOR), pp. 200-211, 2016/5/29, ケンブリッジ (英国).

⑦ Eita Nakamura et al., “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, pp. 276-280, 2016/3/22, 上海 (中国).

⑧ Eita Nakamura et al., “Autoregressive Hidden Semi-Markov Model of Symbolic Music Performance for Score Following,” Proc. 16th International Society for Music Information Retrieval Conference, pp. 392-398, 2015/10/28, マラガ (スペイン).

⑨ Eita Nakamura et al., “Automatic Piano Reduction from Ensemble Scores Based on Merged-Output Hidden Markov Model,” Proc. 41st International Computer Music Conference, pp. 298-305, 2015/9/29, デントン (米国).

⑩ Eita Nakamura et al., “Characteristics of Polyphonic Music Style and Markov Model of Pitch-Class Intervals,” Proc. 5th Mathematics and Computation in Music, pp. 109-114, 2015/6/22, ロンドン (英国).

[その他]

ホームページ等

研究代表者個人ページ (成果一覧など):

<http://eita-nakamura.github.io/index-ja.html>

演奏誤り認識・演奏データ整備ツール:

<https://anonym9382.github.io/demo.html>

自動採譜 (オンセットビート位置認識):

<http://anonymous4721029.github.io/demo.html>

自動採譜 (オフセットビート位置認識):

<https://anonymous574868.github.io/demo.html>

## 6. 研究組織

### (1)研究代表者

中村 栄太 (NAKAMURA, Eita)

京都大学・大学院情報学研究科・特別研究員 (PD)

研究者番号：10707574

### (4)研究協力者

吉井 和佳 (YOSHII, Kazuyoshi)

京都大学・大学院情報学研究科・講師

研究者番号：20510001

金子 仁美 (KANEKO, Hitomi)

桐朋学園大学・音楽学部・教授

研究者番号：00408949

渡邊 健二 (WATANABE, Kenji)

東京芸術大学・音楽学部・教授

齋藤 康之 (SAITO, Yasuyuki)

木更津工業高等専門学校・情報工学科・准教授

研究者番号：40331996

片寄 晴弘 (KATAYOSE, Haruhiro)

関西学院大学・理工学部・教授

研究者番号：70294303

橋田 光代 (HASHIDA, Mitsuyo)

相愛大学・音楽学部・講師

研究者番号：20421282