

多声MIDI演奏の楽譜追跡における 演奏の即興性のモデル化と自動伴奏への応用

中村 栄太¹ 山本 龍一² 齋藤 康之³ 酒向 慎司² 嵯峨山 茂樹¹

概要: 人間の演奏に対して自動的に伴奏を同期する, 自動伴奏において楽譜追跡は重要な問題である. 本稿では, 装飾音や即興的な楽句を含む楽曲・演奏に対して追従可能な楽譜追跡について述べる. 装飾音のモデル化を演奏生成を記述する隠れマルコフモデルを用いて行うとともに, さらに一般の即興的な楽句の構造を記述する方法について議論する. このモデルを用いた楽譜追跡アルゴリズムを示し, その有効性を実際の演奏を用いて評価する. また, このアルゴリズムを用いた自動伴奏システムを構築し, その動作を確認する.

キーワード: 楽譜追跡, 自動伴奏, 演奏の不確定性, 隠れマルコフモデル, テンポ推定.

Modeling Performance Indeterminacies for Polyphonic Midi Score Following and Its Application to Automatic Accompaniment

NAKAMURA EITA¹ YAMAMOTO RYUICHI² SAITO YASUYUKI³ SAKO SHINJI² SAGAYAMA SHIGEKI¹

Abstract: Score following plays an important role in automatic accompaniment, which is an automated performance of accompaniment in synchrony with human performances. This paper describes the score following capable of following performances with ornaments and improvised phrases. We construct a probabilistic model of ornaments based on hidden markov model and discuss a method of describing the structure of more indeterminate improvisational phrases. A score following algorithm based on the model is proposed and its effectiveness is evaluated using human-played performances. An automatic accompaniment system using the algorithm is built and its operation is tested.

Keywords: score following, automatic accompaniment, performance indeterminacies, hidden markov model, tempo estimation.

1. 導入

人間の演奏に対して, 機械が楽譜に基づく伴奏を, 自動的にリアルタイムで同期して演奏することは, 自動伴奏と呼ばれる. 自動伴奏は, 奏者数を軽減した合奏を可能とす

るほか, 合奏曲などの練習の際における効用も期待される. また, 電子音楽と生演奏の融合にも用いることができ, 新しい作曲や演奏形態の可能性を与えることができる. こうした背景から, 自動伴奏は Dannenberg [1] や Vercoe [2] の研究以来, 活発に研究されている (この分野の発展については, 例えば文献 [3] を参照).

自動伴奏を可能とするには, 人間による演奏と楽譜とのマッチングをリアルタイムに行う, 楽譜追跡が必要である. 実際の演奏には, 楽譜からは一意に定まらない不確定性や即興性が存在する. 特に, 楽曲中に現れる装飾音や古典派の協奏曲のカデンツァのような構造では, 演奏の即興性が

¹ 東京大学
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
² 名古屋工業大学
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
³ 木更津工業高等専門学校
KisarazuNCT, 2-11-1 Kiyomidai-Higashi, Kisarazu, Chiba 292-0041, Japan

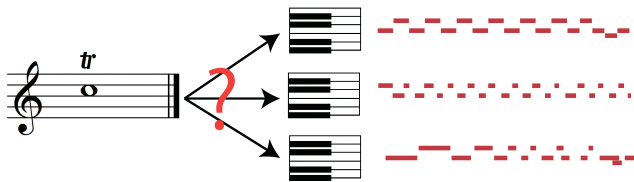


図 1 トリルの様々な演奏法。
Fig. 1 Various realizations of a trill.

大きく反映される。楽譜追跡は、このような演奏の不確定性に対して頑健である必要があるが、従来の楽譜追跡では、装飾音や即興的なフレーズを含む演奏に対する追従は十分に議論されていなかった [4], [5]。本稿では、これらを含む演奏に対しても追従可能な楽譜追跡を目標とし、議論する。

楽譜追跡には入力信号として、音響信号を用いるものと MIDI 信号を用いるものがある。音響信号による楽譜追跡には、Raphael [6] や Cont [4] により考案されたものがある。これらのシステムは特に単旋律楽器による演奏に対しては優れた性能を示すものの、ピアノなどの多声楽器の演奏に対しては依然として大きな困難を有している。このような楽器の演奏に対しては MIDI 信号を入力とする楽譜追跡が実際的である [5]。また、鍵盤楽器の場合には電子ピアノの普及により直に MIDI 信号を得ることが可能な場合も多い。本稿では、特に需要の大きいと考えられるピアノなどの鍵盤楽器の演奏に対する自動伴奏を目的とし、多声 MIDI 演奏を入力とする楽譜追跡を扱う。また、提案する楽譜追跡アルゴリズムの自動伴奏への応用も議論する。

2. 演奏における即興性と楽譜追跡

2.1 演奏における不確定要素

楽譜に基づいた音楽の演奏では、演奏行為の多くの部分が指定されているものの、同時に、直接指示されず奏者の判断に委ねられる部分もある。これは演奏の不確定性と呼ばれる。演奏の不確定性の代表的なものを以下に列挙する [8]。

- a) 音長や発音時刻の微妙な変動
- b) テンポ
- c) 強弱やアーティキュレーション
- d) 演奏誤りや弾き直し・弾き飛ばし
- e) 装飾音における音価や音数、音高
- f) 即興的な楽句

これらの、楽譜からは一意に予測できない演奏の不確定性は、奏者の意図や即興性によるものとそうでないものがあるが、それらをを分別することは一般的に難しい。これらの不確定性を含む演奏に対して自動伴奏を行うには、頑健に楽譜追跡を行いながら、必要な演奏要素の情報を抽出・推定することが要求される。上述した演奏の不確定性のうち a) から d) については、多くの文献で議論されている [3], [5], [7], [8]。

以下では、特に e) と f) に重点を置いて議論を進めていくことにする。まず、装飾音の例としてトリルを考える (図 1)。トリルでは、構成音の具体的な音価や音数が指定されていないため、様々な演奏法が可能である。慣習などにより、音高に関しても不確定性がある場合がある。このような様々な演奏に対して、構成音を一音一音追跡すると、意図されない演奏楽譜位置を推定してしまう危険性が高い。また、トリルの構成音の音長を用いてテンポ推定を行うと、変動が大きく、また意図されないテンポを推定してしまう可能性が高い。同様のことは、アルペジオや前打音など他の装飾音に対しても当てはまる。

この問題は、装飾音をその構成音一つ一つに注目するのではなく、全体として一つの演奏行為として捉えることにより解決できると考えられる。また、テンポ推定に関しても、全ての音符の情報ではなく、装飾音の開始位置など選択された情報をもとに推定することが重要であると考えられる。

次に、f) に挙げた即興的な楽句の代表例として、古典派の協奏曲のカデンツァがある。カデンツァは、基本的に、奏者が自由に演奏できる即興的なセクションである。このような楽句の演奏では、楽譜追跡は基本的に不可能であり、またテンポ推定も不可能に近く、また推定する必要性も小さい。ただし、即興的な楽句の開始と終了は楽譜追跡で重要となる。

2.2 楽譜追跡問題

以下では、音価をもとに指定された各音符の楽譜上での位置のことを楽譜時刻 (単位は、例えば、四分音符) と呼ぶことにし、変数として τ を用いることにする。これに対して、通常の時刻のことを実時刻 (または単に、時刻) と呼び、変数として t を用いることにする。また、各音符の発音開始点 (onset time) のことを、発音時刻や発音楽譜時刻と呼ぶことにする。

演奏される楽譜 (以下、演奏譜) は、発音楽譜時刻を与えられた音符の列であるので、 $X = \{(\tau_i, c_i)\}_{i=1}^I$ と表される。ここで、 I は演奏譜上の「音符」の個数であり、 i は各「音符」を示す添え字である。 τ_i は i 番目の「音符」の発音楽譜時刻、 c_i はその音価や音高、強弱などの各情報を表しているものとする。簡単な単旋律の楽譜の場合、 c_i は単音からなる音符を表すが、多声音楽の楽譜の場合、 c_i は同楽譜時刻に発音される音全部の情報を表しているものとする。また、上述したように装飾音は、その構成音を別々の音符とするより一つの「音符」と見なす方が適切であり、 c_i はそうした装飾音も表すものとする。以下、 c_i を演奏行為と呼ぶことにする。

奏者による演奏は、演奏イベントの時系列 $S = \{(t_m, s_m)\}_{m=1}^M$ で表される。ここで、 m は各演奏イベントを表す添え字であり、 t_m は演奏イベントの起こる時刻

(以下、演奏イベント時刻), s_m は演奏イベントの種類を表す. これを, 演奏イベント列と呼ぶことにする. 本稿では, MIDI イベントを演奏イベントとして用いるため, s_m は MIDI メッセージを表していると考えてよい. また, s_m は多声楽器の楽譜追跡にとって特に重要だと考えられる, 発音開始イベントの音高情報を表すものとする [8]. 演奏イベント列およびその長さ M はオンラインで演奏イベントが到達する度に更新していくものとする.

演奏譜 X と, 演奏開始から現在までの演奏イベント列 S が与えられた時の, 現在の演奏位置の推定問題は, X と S に対して, 現在の楽譜位置 i_M を推定するという問題として定義される. ここで, i_M は M 番目の演奏イベントに対応する楽譜上の演奏行為の番号を表している. その楽譜時刻は, τ_{i_M} である.

3. HMM による演奏生成のモデル化

3.1 演奏生成の確率モデルと楽譜追跡

前章で述べた, 不確定的要素を含んだ演奏は確率的に生じると解釈できる. このような演奏の生成過程を確率モデルとして記述できれば, 楽譜追跡の問題は, 与えられた演奏イベントの観測に対する確率的逆問題として捉えることができる. この方法は, 人間による楽譜追跡の方法を模倣していると考えられる. 以下では, この理念に立脚して楽譜追跡を議論する.

確率モデルに基づいた楽譜追跡の利点として, モデルに含まれる各パラメータに物理的意味を与えることが容易である点が挙げられる. さらに, 確率モデルに基づいたアルゴリズムでは, 各パラメータの値を機械学習によって最適な値に調整可能であるという点が長所である.

3.2 演奏 HMM

人間による演奏の過程は, まず楽譜上の演奏行為が意図され, その次にその意図に基づいて演奏行為が実行されて演奏イベントが発信される, という 2 段階に分けて考えることができるが, これは隠れ状態を持つ確率モデルで記述される. 演奏イベントと演奏イベント時刻が観測される確率 $p(t_m, s_m)$ は, 意図された演奏行為とその直前の演奏行為と演奏イベント時刻に依ると考えられるので, 条件付き確率 $p(t_m, s_m | i_m, i_{m-1}, t_{m-1})$ によって表される. この確率は, 演奏イベント時刻間の相対的な差にのみ依存すると考えられるので, 以下では

$$b_{i_{m-1}, i_m}(\delta t_m, s_m) = p(t_m, s_m | i_m, i_{m-1}, t_{m-1}) \quad (1)$$

と記すことにする. ただし, $\delta t_m = t_m - t_{m-1}$ は発音時刻間隔 (inter-onset interval, ioi) を表している.

また, 演奏行為が意図される確率は, 過去の演奏行為によって定まると考えられる. 過去の演奏に対する依存関係がマルコフ的であると仮定すると, これは $p(i_m, t_m | i_{m-1}, t_{m-1})$

によって表される. この確率は, 一般には δt_m に依存するが, 通常の演奏ではこの依存性は小さいと考えられるので, 以下, この依存性は取り扱わないことにする. 以上より, 演奏イベント列 $S = \{(t_m, s_m)\}_{m=1}^M$ が観測される確率は,

$$p(S) = \sum_Q p(S|Q)p(Q) \quad (2)$$

$$= \sum_{i_1, \dots, i_M} \prod_{m=1}^M a_{i_{m-1}, i_m} b_{i_{m-1}, i_m}(\delta t_m, s_m) \quad (3)$$

で与えられる. ただし, $Q = \{i_m\}_{m=1}^M$ は演奏行為の系列であり, $a_{i_{m-1}, i_m} = p(i_m, t_m | i_{m-1}, t_{m-1})$ と記した. 上式は, 演奏イベント生成が, 状態遷移確率 a_{i_{m-1}, i_m} と出力確率 $b_{i_{m-1}, i_m}(\delta t_m, s_m)$ を持つ隠れマルコフモデル (Hidden Markov Model, HMM) で記述されることを表している. このモデルを, 演奏 HMM と呼ぶことにする.

出力確率において, ioi の出力と演奏イベントの出力は独立であると仮定すると,

$$b_{i_{m-1}, i_m}(\delta t_m, s_m) = b_{i_{m-1}, i_m}^{(ioi)}(\delta t_m) b_{i_m}^{(evt)}(s_m) \quad (4)$$

と書ける. ここで, 演奏イベントの出力確率は, 以前の演奏に依らないと仮定した.

3.3 状態遷移確率と出力確率

状態遷移確率 a_{i_{m-1}, i_m} は演奏行為の演奏順序を確率的に表している. 誤りのない演奏の際には, $a_{i_{m-1}, i_m} = \delta_{i_{m-1}+1, i_m}$ である. しかし, これ以外の遷移を考えることにより弾き間違いなどを含んだ演奏のモデル化が可能である. また, 遠く離れた和音間の状態遷移確率に対しても 0 でない値を与えることにより, 弾き直しや大きな弾き飛ばしを含む演奏をモデル化できる.

具体的には, 弾き間違いによる一つの和音の脱落の確率は a_{i_{m-1}, i_m} において $i_m = i_{m-1} + 2$ の値で表される. また, ミスタッチによる音の挿入誤りの確率は $i_m = i_{m-1}$ の場合の値により与えられる. さらに, $i_m < i_{m-1}$ の場合の値は弾き直しの確率であり, $i_m = i_{m-1} + d$ で $d \gg 1$ の場合は大きな弾き飛ばしの確率を表している. a_{i_{m-1}, i_m} の具体的な値は, 演奏データから学習できるが, ここでは経験に基づいて設定することにする. 演奏 HMM の状態遷移トポロジーを図 2 に示す.

演奏イベントの出力確率は, 次章に述べるように, 各演奏行為に応じて定められる. また, 弾き間違いを含む演奏をモデル化するために, 正しい演奏で期待される演奏イベント以外の出力確率も正の値を持つものとする. ioi の出力確率は, テンポ推定の結果と演奏譜の音価情報をもとに設定することができる.

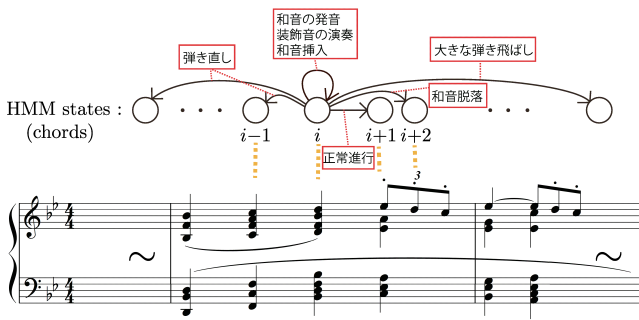


図 2 演奏生成の HMM の状態遷移トポロジー. i 番目の状態は演奏譜の i 番目の演奏行為 (和音, 装飾音など) に対応している.

Fig. 2 Topology of state transition probability for the performance HMM. The i 'th state corresponds to the i 'th "musical action" (chord, ornament etc.) of the performance score.

4. 演奏 HMM における和音および装飾音の記述

4.1 和音

演奏譜の i 番目の演奏行為が和音である場合を考える. 和音は, 楽譜上では同時に発音される多重音として記述される. しかし, 実際の演奏では各和音構成音は, 異なる時刻に順序不定な演奏イベントとして観測される. こうした和音の演奏は, 演奏 HMM において小さな ioi を出力する自己遷移の結果として記述できる. 和音構成音間での ioi は, 約 35 msec 以下に分布していることが知られている [9]. 以下では, 和音の演奏に対応する自己遷移に関する出力確率 $b_{i,i}^{(ioi)}(\delta t)$ は, $\delta t \leq 35$ msec に主に分布しているものとする. また, 演奏イベントの出力確率 $b_i^{(evt)}(s)$ は, s が和音構成音のときに正となる分布により表される.

4.2 トリル

通常のトリルは, 2 音の交互繰り返しとして演奏される. これは, 2 つの HMM 状態を用いて記述することも考えられるが, ここでは 1 つの状態の自己遷移によって表現することにする*1. これにより, 和音を含む shake や多重トリル, 3 つ以上の音のトリルに対しても, 同様な記述が可能である.

トリル内の演奏音間の ioi は, 演奏の状況により様々であるが, およそ $30 \lesssim \delta t \lesssim 200$ msec の範囲に分布している. 以下では, 多様な演奏に対応するため, トリルの演奏に対応する自己遷移に関する出力確率 $b_{i,i}^{(ioi)}(\delta t)$ は, 上記の範囲に主に分布しているとする. また, 演奏イベントの出力確率 $b_i^{(evt)}(s)$ は, s がトリル構成音のときに正となる分布により表される.

*1 これにより, 実際にはほとんど演奏されない同音連打を生成する確率も 0 ではなくなるが, 確率的逆問題を解く際には問題は生じにくいと考えられる.

4.3 アルペジオと前打音

アルペジオは構成音間の ioi が短い一連の音によって演奏される. 通常アルペジオの演奏は, 構成音の順序が定まっているが, 即興的に順序や音数が変化して演奏されることもある. ここでは, 多様な演奏に対応するため, アルペジオの構成音は順不同に現れるとし, また同音高の繰り返しも起こるものとする. これは, 1 つの HMM 状態の自己遷移による出力によって記述される.

前打音は, 1 つまたは複数の短い音が音価の指定された音符の前に演奏されるものであるが, ここでも多様な演奏を想定し, 後続 (和) 音を含めた全体を 1 つの HMM の自己遷移による出力によって記述することにする.

アルペジオ構成音間の ioi および前打音の音長は, 通常 $30 \lesssim \delta t \lesssim 100$ msec の範囲に分布している. 以下では, 対応する出力確率 $b_{i,i}^{(ioi)}(\delta t)$ は, 上記の範囲に主に分布しているとする.

4.4 多声音楽での装飾音の演奏

鍵盤音楽などの多声音楽では, 上述した装飾音は, 他の動きをする声部と同時に演奏されることがある. このような状況は, 装飾音と同時に演奏される他の声部の音を 1 つの HMM 状態にまとめることにより記述が可能である. こうして作られたモデルは, 実際の演奏には現れないような演奏も生成するが, 逆問題を解く過程である楽譜追跡では問題は生じにくいと考えられる.

5. 即興フレーズ

5.1 即興フレーズの必要性

前章では, 装飾音が HMM を用いて記述されることを説明した. この演奏 HMM では, 不確定性の高い装飾音が, より不確定性の低い通常の和音と統一的方法で記述することが可能であった. しかし, 同様な方法で装飾音よりもさらに不確定性の高い即興的な演奏を記述することは難しい. 例えば, 古典派の協奏曲におけるカデンツァのような, ほぼ完全な即興的なフレーズを HMM 状態を用いて記述しようとすると, ioi や音高の出力確率の情報量が高い状態となる. これでは, 逆問題を解く際にこの状態が最尤状態となることは必ずしも保障されない.

一方, 即興的なフレーズの楽譜追跡では, 個々の音を認識する必要はなく, フレーズの開始と終了を検知することのみが重要である. そこで, このような即興的なフレーズは, 特別な HMM 状態から出力される演奏イベント列で表されるとする. 以下では, このようなフレーズあるいはそれに対応した HMM 状態のことを即興フレーズと呼ぶことにする.

5.2 カデンツァと一般の即興フレーズ

即興フレーズの代表例として, 上述した協奏曲のカデン

ツァを考える。カデンツァを表す即興フレーズは、指定された演奏譜上の位置に置かれ、HMM 状態としては、ほぼ全ての音高を出力する状態として記述できる。カデンツァは通常、特定の和音あるいはフレーズによってその終了が示される。よって、対応する即興フレーズは、この和音以外の音高を自由に出力して構成されるものとしてよい。カデンツァの終了を示す和音やフレーズは、カデンツァの後に続く HMM 状態 (列) として表される。

同様に、一般の即興フレーズは、特定の HMM 状態に対応して置かれ、即興フレーズで用いられる音高の集合を指定することにより記述できる。そして、フレーズの終了は後続する HMM 状態列により表現される。また即興フレーズは、フレーズの楽譜時間長または実時間長によって指定されることもあるが、この場合には、フレーズの終了はこれらの情報により表される。

5.3 即興フレーズとしての装飾音

前章では、HMM 状態を用いた装飾音の記述を行ったが、装飾音を即興フレーズとして扱うことも可能である。これにより、さらに多様な演奏法を表現することが可能である。例えば、トリルはトリル構成音を構成音とする即興フレーズとすることにより、ioi の制限なしの記述が可能である。また、アルペジオや前打音も同様に即興フレーズとして表現可能である。この場合は、即興フレーズの長さを対応する音符の音価をもとに指定することが考えられる。

6. 不確定な演奏行為を含む演奏の楽譜追跡

6.1 演奏の楽譜位置推定

演奏の楽譜位置推定問題は、3 章に述べた演奏 HMM の逆問題として扱うことができる。具体的には、演奏譜 X に基づく現在までの演奏イベント列 S が与えられたとき、最も確からしい演奏行為の系列 $Q = \{i_m\}_{m=1}^M$ を求めることにより、現在の演奏位置の推定が可能である。Bayes の定理より、これは

$$\operatorname{argmax}_Q p(Q|S) = \operatorname{argmax}_Q [p(S|Q)p(Q)] \quad (5)$$

$$= \operatorname{argmax}_{i_1, \dots, i_M} \left[\prod_{m=1}^M a_{i_{m-1}, i_m} b_{i_{m-1}, i_m} (\delta t_m, s_m) \right] \quad (6)$$

を解くことにより求めることができる。

HMM の最尤推定は (前方) Viterbi アルゴリズムを用いて行うことができる。時刻 t_{M-1} における各状態 i_{M-1} での最尤値を

$$\hat{p}_{i_{M-1}} = \max_{i_1, \dots, i_{M-2}} \left[\prod_{m=1}^{M-1} a_{i_{m-1}, i_m} b_{i_{m-1}, i_m} (\delta t_m, s_m) \right] \quad (7)$$

とすると、

$$\hat{p}_{i_M} = \max_{i_{M-1}} [\hat{p}_{i_{M-1}} a_{i_{M-1}, i_M} b_{i_{M-1}, i_M} (\delta t_M, s_M)] \quad (8)$$

と帰納的に求まる。楽譜追跡では、最尤状態推定をリアルタイムで行う必要があり、演奏譜の長が大きくなると処理時間の増大が問題となる。これは、高速 Viterbi 探索アルゴリズム [8] を用いることにより大幅に改善される。

演奏 HMM に即興フレーズが含まれる場合には、上記最尤推定によって即興フレーズが最尤となった際、フレーズの終了まで新たな最尤状態更新は行わないものとする。

6.2 演奏のテンポ推定

自動伴奏を行うためには、演奏の楽譜位置の推定とともにテンポ推定を行う必要がある。テンポ推定は、現在までの直前の演奏イベント列をもとに行うことができる。

装飾音を含む演奏の場合、装飾音の構成音の音価が不確定であるので、これらをもとにテンポ推定を行うことは不適切であると考えられる。そこで、HMM 状態の遷移のときのみテンポの更新を行うことにする。具体的には、状態遷移が起きた時刻をもとに隣り合う発音点間の局所テンポ $r_{i_m} = (\tau_{i_{m+1}} - \tau_{i_m}) / (t_{i_{m+1}} - t_{i_m})$ を計算し、その移動平均を求めることによりテンポ推定を行う。

例えば、前・後打音の演奏は、奏者により発音のタイミングが異なりうる。また、他の場合においても特定の音の発音時刻をもとにテンポ推定を行うのが適切な場合が起こりうる。これらの場合に対応するため、テンポ推定に用いない HMM 状態の選択、および特定の演奏イベントを用いたテンポ推定が可能のように楽譜追跡アルゴリズムを設計した。

7. 楽譜追跡の評価と自動伴奏への応用

7.1 楽譜追跡の評価

提案する楽譜追跡アルゴリズムの有効性を確認するために、アルゴリズムの実装を行い、評価試験を行った。このアルゴリズムは、今回新たに導入した装飾音と即興フレーズに関する部分を除けば、文献 [8] で示されたものと若干の差を除いて同じである。

評価には、複数の楽曲に対する人間の演奏を MIDI 録音したものをを用いた。演奏はアマチュアのピアノ経験者により行われた。結果の一部を以下に示す。図 3 は、トリルを含む演奏の楽譜位置推定結果を示している。譜例は、Beethoven のピアノ協奏曲第 3 番第一楽章の一部である。ピアノロール図の縦線は発音点の位置、青線は位置推定が更新された位置を示している。推定された楽譜位置は、全て正しい位置であった。音価が不確定なトリルや発音時刻の変動がある演奏に対しても、正しく楽譜位置推定が行われたことが分かる。図 4 は、アルペジオを含む演奏の楽譜位置推定結果を示している。譜例は、Mozart の 2 台のピアノのためのソナタ第一楽章提示部終わりの部分の第一ピアノパートを示している。推定された楽譜位置は、全て正しい位置であった。アルペジオが左手のバスと重なってい

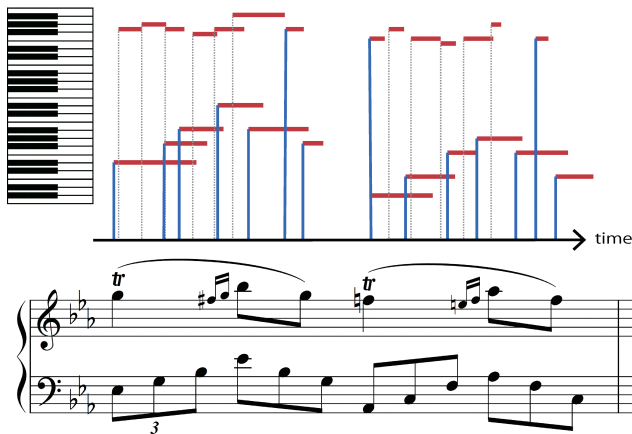


図 3 トリルを含んだ演奏の楽譜位置推定．ピアノロール図の縦線は発音点の位置，太実線（青）は位置推定が更新された位置を示している．

Fig. 3 Estimation of score location for a performance with trills. In the piano roll, the vertical lines show onsets, and the blue bold lines show onsets where the score location estimation is updated.

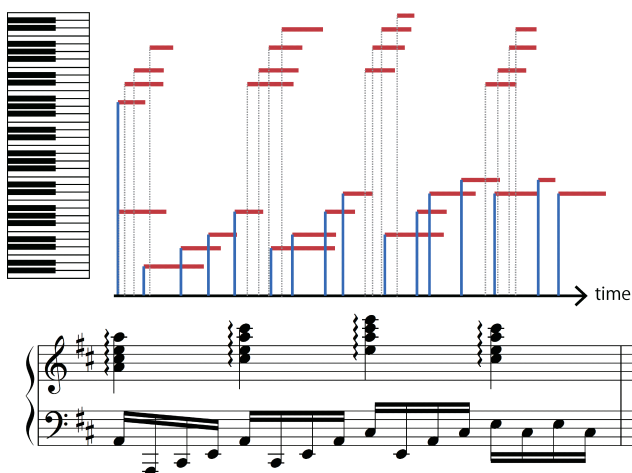


図 4 アルペジオを含んだ演奏の楽譜位置推定．

Fig. 4 Estimation of score location for a performance with arpeggi.

るところでも，推定位置が前後することなく，正しく位置推定が行われたことが分かる．

また，即興フレーズを実際のカデンツァにおいて用いた試験も行った．楽曲には上記 Beethoven の楽曲を用い，演奏は，カデンツァの直前から楽章の終わりまで行った．複数のカデンツァの演奏の場合にも問題なく楽譜追跡が行われ，カデンツァの開始と終了を検知できた．

7.2 自動伴奏への応用

提案する楽譜アルゴリズムを用いて自動伴奏システムを構築した．自動伴奏システムは，Eurydice [8], [10] の楽譜追跡アルゴリズムを上述のものに変更することにより実現した．前節で評価に用いたもの含む複数の楽曲に対して，

実際に自動伴奏をして，システムの動作を確認した．

いずれの楽曲でも，多くの場合に，装飾音などを含む楽句において，演奏者のテンポを反映しながら，安定した伴奏が得られた．

8. 結論

本稿では，装飾音や即興的楽句を含む楽曲・演奏に追従可能な楽譜追跡を議論した．装飾音は，演奏生成を記述した演奏 HMM により表現されることを示した．また，より不確定度の高い協奏曲のカデンツァのような即興的楽句の構造は，即興フレーズという状態を導入することにより表現されることを示した．

この確率モデルの逆問題を解くことにより，楽譜追跡アルゴリズムを設計した．アルゴリズムの有効性を確かめるため，実際の演奏を用いた評価を行った．評価の結果，装飾音や即興的楽句を含む演奏に対して，実際に追従可能であることが確認された．また，このアルゴリズムを用いた自動伴奏システムを構築し，その動作を確認した．

現段階では，装飾音などに対応する HMM 状態の構成は手動で行っている．今後の課題として，MusicXML や midi で書かれた楽譜からの装飾音を含む HMM 状態の構成を自動化することが挙げられる．また，テンポ推定に関しても，推定箇所の選択の自動化や推定精度の向上が課題である．

参考文献

- [1] R. Dannenberg, “An on-line algorithm for real-time accompaniment,” *Proc. ICMC*, pp. 193–198, 1984.
- [2] B. Vercoe, “The synthetic performer in the context of live performance,” *Proc. ICMC*, pp. 199–200, 1984.
- [3] N. Orio *et al.*, “Score Following: State of the Art and New Developments,” in *New Interfaces for Musical Expression*, 2003.
- [4] A. Cont, “ANTESCOFO: Anticipatory synchronization and control of interactive parameters in computer music,” *Proc. ICMC*, 2008.
- [5] D. Schwarz, Nicola Orio and N. Schnell, “Robust Polyphonic Midi Score Following with Hidden Markov Models,” *Proc. ICMC*, 2004.
- [6] C. Raphael, “Music Plus One: A system for expressive and flexible musical accompaniment,” *Proc. ICMC*, 2001.
- [7] 武田晴登, 西本卓也, 嵯峨山茂樹, “HMM による MIDI 演奏の楽譜追跡と自動伴奏,” 情報処理学会研究報告, MUS, pp. 109–116, 2006.
- [8] 中村栄太, 武田晴登, 山本龍一, 齋藤康之, 酒向慎司, 嵯峨山茂樹, “Eurydice: 弾き誤り・弾き直し・弾き飛ばしを含む演奏に追従する自動伴奏システム,” 投稿中．
- [9] 武田晴登, 西本卓也, 嵯峨山茂樹, “テンポ曲線と隠れマルコフモデルを用いた多声音楽 MIDI 演奏のリズムとテンポの同時推定,” 情報処理学会論文誌, 48(1), pp. 237–247, 2007.
- [10] 山本龍一, 中村栄太, 齋藤康之, 酒向慎司, 嵯峨山茂樹, “Eurydice: 弾き飛ばし・弾き直しを含む演奏に追従する自動伴奏システム,” 情報処理学会第 96 回研究発表会, デモセッション, 2012.