

# 音響信号によるリアルタイム楽譜追跡と能動的演奏支援システムに関する研究

代表研究者

酒向慎司

名古屋工業大学大学院工学研究科 助教

## 1 はじめに

協奏とは、複数人が同時に演奏を行う楽器演奏の形態であり、例えば、アンサンブル、吹奏楽、オーケストラなどがこれに該当する。人間同士の協奏においては、複数の楽器演奏者（以下、奏者）が、互いにずれが発生しないように演奏することが期待されるが、人による演奏にはテンポの揺らぎ、時には演奏誤り（音符の置換、挿入、脱落）が含まれるため、演奏ではお互いの楽器がうまく調和することが肝心である。

熟練した奏者であれば、相手の演奏を聴き、次にどのように演奏されるかをその演奏過程から予測し、相手の演奏に柔軟に合わせることができる。このような人の働きを機械で実現しようとする試みとして、奏者の演奏に追従して予め記録された伴奏を自動的に再生させる、自動伴奏の研究が盛んに行われている[1-6, 8]。このような自動伴奏は、協奏の機会が多く得られない奏者に対して奏者単独での仮想的な協奏を実現するほか、伴奏付きの演奏支援としての活用が期待される。特に演奏支援は楽器演奏のための練習に留まらず、音楽の演奏を媒介にしたリハビリテーションへの活用も期待できる。

人に合わせて演奏するという行為は、相手が楽譜上のどの位置を演奏しているのかを認識し次の演奏位置を予測する働きと、物理的に演奏を行う働きが統合されたものとして捉えられる。前者は人の音楽認知機能に関わる重要な働きであり、計算機で実現する試みは楽譜追跡と呼ばれる。

このような観点から本研究では、楽器演奏支援としての活用を想定し、テンポの揺らぎ、演奏誤りや弾き直しを含む、楽譜通りではない演奏に対しても頑健に楽譜追跡を実現することに重点を置く。楽譜追跡における既存研究では、MIDI 信号を入力とした演奏誤り、弾き直しにも頑健な楽譜追跡手法および自動伴奏システムは提案されている[4]。これらのシステムでは、MIDI 出力を備えたごく限られた楽器にしか対応できないという根本的な問題が存在していたが、近年では音響信号を入力とする自動伴奏システムも提案されている[5-6, 8]。しかし多声楽器で音響信号が複雑になりうる場合に頑健に追従可能で、かつ練習支援を想定した弾き直しを考慮したシステムは、楽譜追跡の難しさゆえにいまだに十分なもの存在しない。

このような背景から、本研究ではテンポの揺らぎ、演奏誤り、弾き直しなどの演奏の不確実性を考慮しつつ、音響信号が複雑になりうる多声楽器の演奏に対しても頑健に楽譜追跡を実現する、演奏の確率モデルに基づく楽譜追跡手法を提案するとともに、多声楽器にも対応可能な音響入力自動伴奏システムを開発した。

## 2 自動伴奏システムの構成

人同士の協奏では、奏者は互いの演奏を聴き、次にどのように演奏されるかを予測し、相手の演奏に合うように演奏していると考えられる。この考えに基づき、自動伴奏の問題を、特徴抽出、楽譜追跡、伴奏再生の3つに分けて考える。図1に、本研究で提案する自動伴奏システムの基本的な構成を示す。システムは、上記の3つの問題に対応する要素から構成されており、入力演奏の音響信号、出力は同期された伴奏のMIDI信号である。

特徴抽出部では、音響信号から楽譜追跡の手がかりとなる音高、発音の情報が抽出される。具体的には、音響信号を一定の時間幅で区切り、ウェーブレット変換に基づくクロマベクトルと Spectral flux に基づくオンセット特徴が算出される。これらの2つを合わせた特徴量を用い、楽譜追跡部ではある時刻での演奏位置と瞬時テンポが推定される。

伴奏再生部では、入力の演奏に同期した伴奏を遅延なく再生する。なお、システム前提条件として、演奏される楽譜は既知であり、MIDI データとして与えられるとする。以降では、本システムの核となる要素技術

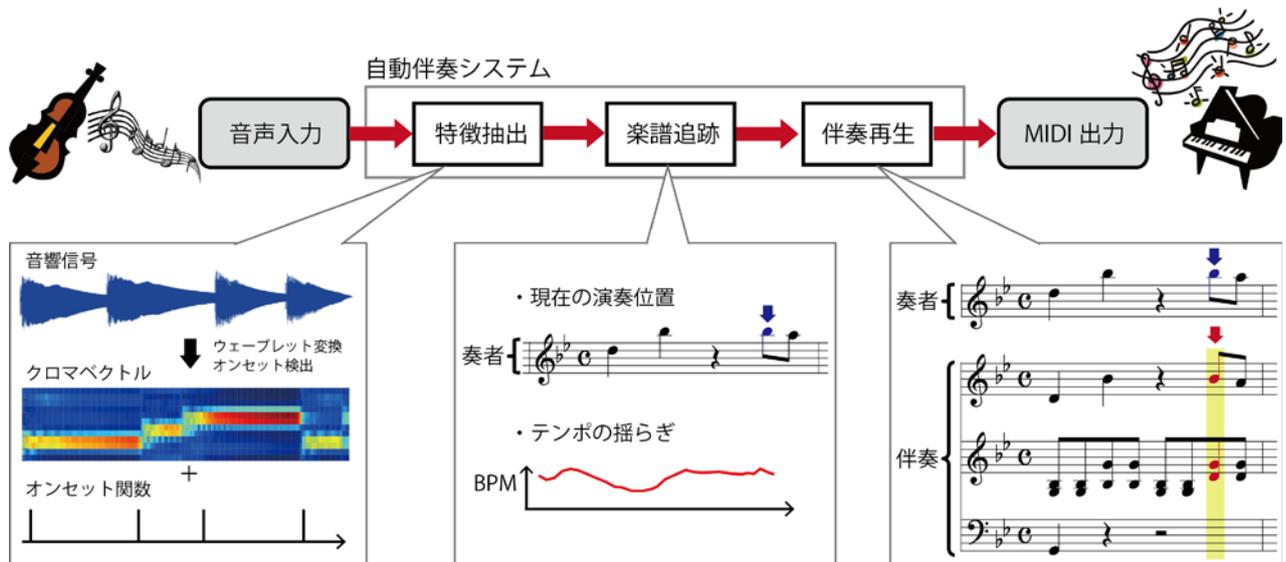


図1. 自動伴奏システムの基本構成である楽譜追跡と伴奏再生について述べる。

### 3 楽譜追跡手法

#### 2-1 楽譜追跡手法の問題設定

人の演奏は楽譜という明確な演奏指示がある場合でも確定的ではなく、楽譜の誤読や演奏誤りによって、音符の置換、挿入、脱落、弾き飛ばしや弾き直し等が含まれる。このような楽器演奏の過程は、楽譜上の和音（同時発音の音符の集合）を演奏位置の単位として、それらを遷移しながら演奏を生成していく過程と捉えられる（図2）。

楽譜追跡は、奏者の演奏に対して、楽譜の情報を元に楽譜上の演奏位置をリアルタイムで求める問題である。ここでは、伴奏再生のために演奏の速さを表すテンポの情報が必要となるため、楽譜追跡をある時刻の演奏位置と瞬時テンポを求める問題とする。自動伴奏においては、とくに音響信号を対象とする意味では、単旋律楽器のみならず多声楽器を含む様々な楽器においても頑健に追従可能であり、かつ練習の場合を想定して、演奏誤り、弾き直し、テンポの揺らぎを考慮することが望ましい。そのような楽譜追跡を実現するためには、以下の問題を解決する必要があると考えられる。

- (1) 演奏誤り、弾き直し、テンポの揺らぎなど、多くの演奏の不確実性が含まれる。
- (2) 特に多声音楽の場合に、追跡の手がかりとなる音高や発音時刻の不確実性が増加する。
- (3) 楽譜追跡は、リアルタイムという時間的な制約を満たし、かつ因果的なシステムでなければならない。

上記のうち、(1)および(2)に関しては、これまでも多くの研究で議論されてきた[14-19]。しかし、(3)については十分ではなく、著者らは新たに(3)の制約の元でも頑健な楽譜追跡手法を提案する。

#### 2-2 確率モデルに基づく楽譜追跡

本研究で提案する楽譜追跡手法の基本的な指針は、演奏位置の遷移とテンポの揺らぎの依存関係を確率モデルで記述することにある。確率モデルで表すことで、演奏誤りや弾き直しを見通しよくモデル化することができ、テンポの揺らぎ、音高や発音時刻の不確実性を確率分布の揺らぎとして捉えることができる。また、

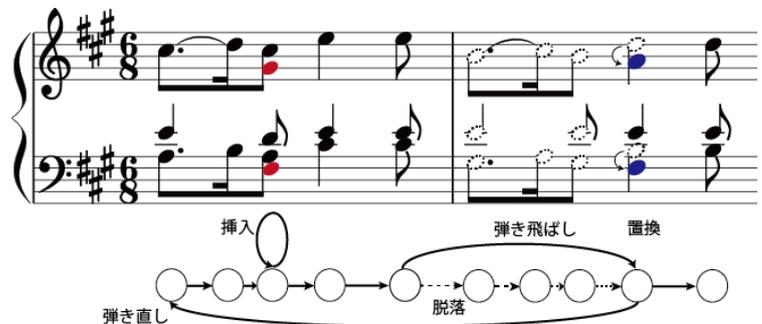


図2. 変動を考慮した演奏過程の概念

楽譜追跡の問題は確率モデルの事後確率最大化として定式化することができるほか、本研究では扱わないが確率モデルを記述するパラメータを、データから統計的に自動学習することも可能である。

このような指針に基づき、本研究で提案する確率モデルを図2に示す。提案モデルは、演奏位置の遷移を表すセグメントモデルと、テンポの変化を表すテンポモデルを統合したものに相当する。モデルは、音響信号から抽出する特徴量を $o$ 、演奏位置（和音単位） $s$ 、継続長 $d$ 、テンポ $r$ としたセグメンタル条件付き確率場（Conditional Random Fields; SCRF）として解

積することができる。演奏位置およびテンポの推定は、それぞれViterbi アルゴリズム、カルマンフィルタによって行うことで、リアルタイムという時間的な制約下でも効率的に計算できることが期待できる。

SCRF は、生成モデルである隠れマルコフモデル（HMM）の識別モデルに相当する条件付き確率場（CRF）を、セグメントモデルへと拡張したものである。SCRF の特徴として、複数フレームに渡る観測信号を一つの和音としてモデル化することで、従来法で多く用いられているHMMとは異なり、和音の継続長などの音楽の時間的構造を直接モデル化することができる。また、生成モデルの尤度を素性関数として組み込むことも可能である。

特徴量は、従来研究に習い音高と発音の情報を表すウェーブレット変換に基づくクロマベクトルと Spectral flux に基づくオンセットを用いる。特徴量とモデルの関係を表す観測モデルはそれらを元に設計し、テンポの時間変化を表すテンポモデルは、短時間にテンポが大きく変化しないと仮定することで線形動的システムを用いて設計する。

なお提案モデルは、Joder らによって提案されているCRFをベースにしたモデルに類似したものである[7]。Joder らによるモデルでは、テンポが離散変数であるのに対して提案モデルでは連続変数である点が大きく異なる。この違いは十分に議論する必要があるが、自動伴奏のようにテンポ変動の解像度が重要視される場合には、テンポが連続変数としてモデル化されるほうが望ましいと考えている。

## 2-2 遅延決定と未来予測に基づく 楽譜追跡

従来法の多くでは、リアルタイムという制約のためViterbi アルゴリズムや動的計画法に基づいた手法において、逐次的に最適解を出力する形に近似していた[20-21, 24]。

しかし逐次的に入力される信号系列において、時刻 $t$ における最適解が、時刻 $t+1$ においては最適解ではないといったことが良く起こりうる。特に、多声演奏や多重音演奏のように音響信号が複雑な場合にはその傾向が顕著となり、演奏位置の誤推定が問題となる。そこで、一定の遅延幅を考え、 $\alpha$ 秒だけ過去の演奏位置をViterbi アルゴリズムによって推定し、現在位置はテンポにより予測する頑健な楽譜追跡方法を考案した。

これは、一定の遅延を許容することで相対的に未来の情報も用いることで演奏位置を正確に推定することを目的としたものである。すなわち、演奏位置の推定が正確に行われていれば、そこから導出されるテンポの信頼性も高まり、結果的に遅延分の遅れを適切にカバーすることができるというアイデアに基づいている。この遅延決定アルゴリズムの基本的なステップは以下の通りである。詳細は[13]を参照されたい。

[Step 1] Viterbi アルゴリズムにより入力信号の演奏位置と発音時刻の系列を求める。

[Step 2] Step 1 の結果からカルマンフィルタによって局所的なテンポの推定を行う。

[Step 3] 遅延幅 $\alpha$ を一つ前の発音時刻までの時間と設定し、演奏位置の遷移の1次近似モデルにより現在の演奏位置を予測する。

上記のアルゴリズムを音響信号が入力される度に繰り返すことにより、リアルタイムで演奏位置とテンポを求めることができる。ここで、 $\alpha$ は遅延幅を与えるパラメータであり、実際には局所的なテンポが一定と見なせる程度の小さい値を持つと考え、本研究では実験的に設定する。

## 4 伴奏の同期再生

楽譜追跡の結果より得られる演奏位置とテンポから、伴奏の演奏位置とテンポを決定し、リアルタイムに同期された伴奏を再生する問題を伴奏再生とする。なお、伴奏の出力にはMIDI 信号を用いる。伴奏に音響信

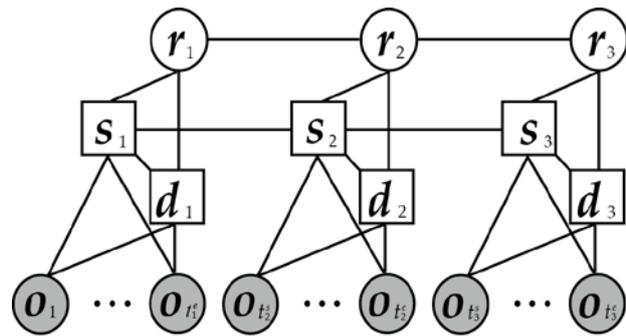


図3. 提案モデルの依存関係の概要

各ノードの形状は円（連続）、矩形（離散）を表し、ハッチングの有無は観測（あり）、隠れ変数（なし）を表す

号を用いる自動伴奏システム[6, 8]も存在し、それらはMIDI 信号では表現できない表情豊かな伴奏を可能にするほか、既存の音源をそのまま利用できる利点がある一方で、伴奏用の特定のパートだけの音響信号を予め用意しておく必要がある。

これには、ある音源から特定のパートだけを除去するマイナスインと呼ばれる技術が必要であり、一般的に困難である。MIDI であればそのような困難さはなく、各パートがトラック別に作成された MIDI データを用意するだけで自動伴奏を実現することができる。人同士が協奏を行う際には、協奏する相手の次の演奏のタイミングを予測して、自身の演奏を合わせようすると考えられる。ここでは、このような予測に基づいて、奏者の演奏に追従する伴奏再生手法を考える。

ある時点において、奏者と伴奏の演奏位置および瞬時テンポが一致し、かつ瞬時テンポが局所的に変化しないと仮定すれば、演奏位置の遷移の1次近似によって、伴奏の演奏のタイミングを予測することができると考えられる。しかし、演奏中のテンポには揺らぎが含まれるため、時間が経つにつれて奏者と伴奏の演奏位置にずれが生じることが予想される。したがって、本研究では、楽譜追跡において推定された奏者の演奏位置とテンポをリアルタイムで伴奏に同期することで、伴奏の演奏のタイミングを予測する方法を用いる。

計算機による MIDI 演奏であれば、伴奏の演奏をほぼ遅延なく再生することができるため、予測された位置に基づき伴奏の MIDI 信号を再生させることで、奏者に瞬時に追従する自動伴奏を実現することができる。弾き直しが行われた場合には、奏者と伴奏に演奏位置のずれが生じるが、そのような場合には伴奏を一時停止し、奏者の演奏位置とテンポを伴奏に同期することで対応させる。これにより、奏者の演奏に追従する自動伴奏を実現することができる。

実際には、奏者のテンポの揺らぎにより、奏者の演奏位置と予測される伴奏の演奏位置とテンポは一致しないことが考えられるが、本システムでは、楽譜追跡部で発音が検出されたときに、伴奏の演奏位置とテンポを奏者に同期する方法を取る。これにより奏者と伴奏の極端な演奏位置のずれを防ぐことができ、奏者に追従する自動伴奏を実現することができる。

また、実際には楽譜追跡の誤推定の可能性を考慮して、楽譜追跡が一定時間楽譜通りの演奏位置の遷移を出力していれば伴奏の MIDI 信号を再生させ、そうでなければ伴奏を一時停止することで、不自然な伴奏を回避することができる。なお、入力信号のフレーム毎に演奏位置とテンポは推定されるが、実際には楽譜追跡部によって発音が検出された場合に、演奏位置とテンポを伴奏再生部に対して出力する。

## 5 実験

### 5-1 スペクトルテンプレートの自動適応

本節では、多様な楽器に頑健に対応させるため、非負値行列分解(NMF)を導入し、スペクトルテンプレートを対象楽器に合わせて逐次的に更新する方法を検討した。過去の推定結果から、スパース NMF に基づいてスペクトルテンプレートを更新することで、図4のように標準的なテンプレート(左)から対象楽器の調波構造に特化したテンプレートへ変形させることで、様々な楽器に変化した時の性能低下が抑えられると考えられる。複数の楽器が含まれる RWC データベース(ジャズ音楽)を用いたシミュレーション実験では、スペクトルの逐次更新による改善が見られなかった。その原因として主に二つの理由が考えられる。一つ目はアライメント推定結果に基づいてテンプレートを更新するため、アライメントの精度によって適切なテンプレートに更新されない場合がある点、二つ目の問題として、過度にテンプレートを変形してしまうことによる汎化性の低下により、継続して行うアライメント性能が低下するという問題である。また、楽器ごとに調波構造を調節し、複

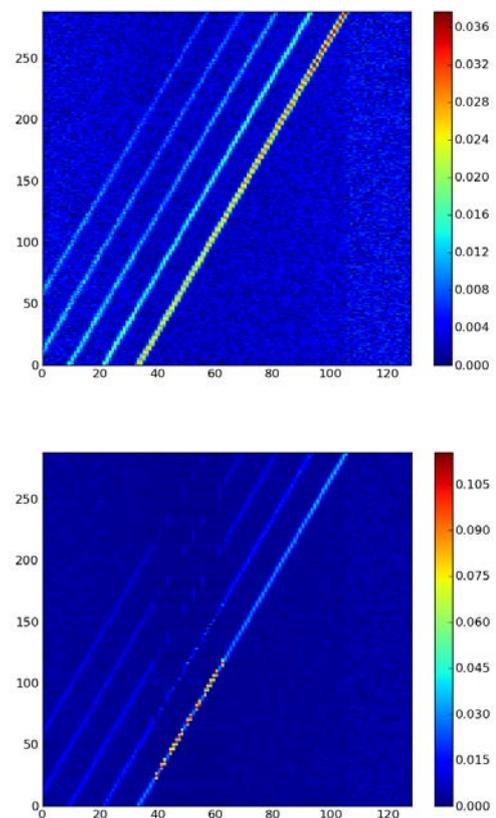


図4. スペクトルテンプレートの変形  
(上:更新前、下:更新後)

数のテンプレートを作成し、目的の楽器に合わせて変更する方法を試みたところ、単一のテンプレートとの比較では推定精度に優位な差は見られなかった。

テンプレートを更新するため、アライメントの精度によって適切なテンプレートに更新されない場合がある点、二つ目の問題として、過度にテンプレートを変形してしまうことによる汎化性の低下により、継続して行うアライメント性能が低下するという問題である。また、楽器ごとに調波構造を調節し、複数のテンプレートを作成し、目的の楽器に合わせて変更する方法を試みたところ、単一のテンプレートとの比較では推定精度に優位な差は見られなかった。他の音響信号での楽譜追跡の関連研究において、楽器の違いに対してテンプレートを逐次適応する手法の有効性が示されていることから、この課題については継続的に改善を行う必要があると考えている。

### 5-2 テンポ推定の高精度化

本節では LDS によるテンポ推定の状態方程式を 0 次近似モデルから 1 次近似モデルに拡張し、シミュレーション実験によって性能評価を行った。図 5 は、正解テンポ（赤線）に対してそれぞれのモデルで推定されたテンポ（緑・青の破線）をプロットしたものである。傾向として、0 次近似モデルは直接的に正解テンポに近づくように推定が行われていることに対し、1 次近似モデルは正解テンポの傾きに追従するように正解テンポに近づくような推定がなされることが分かった。RWC 音楽データベース（ジャズ音楽）による拍位置推定精度では、0 次近似モデル、1 次近似モデルの  $F$  値がそれぞれ 0.849、0.850 となり、数値的な有意差は認められなかった（表 1）。アライメントの推定精度が信頼できる場合には、1 次近似モデルのほうが自然なテンポ推定がなされていることが多く、逆にアライメント推定精度が信頼できない（推定に誤りが多い場合）は、逆に不自然さが目立つことも分かった。総合的に考えると、アライメントの推定精度に依存せずテンポ推定が行える点から、0 次近似モデルのほうが、様々な変動要因が含まれる実環境のロバスト性が期待できることが分かった。また、テンポに対する追従性（テンポ変化に追いつく速さ）の観点では、テンポ変化を許容する分散の値に依存することが分かった。分散を大きく設定することでテンポ変化の追従性は高くなるが、一方でエラーも増大する傾向があり、テンポの追従性とテンポ推定精度はトレードオフの関係にあることが分かった。これらのバランスを考慮して、ロバスト性の高い分散の値をシミュレーション実験によって設定した。

### 5-3 実演奏データベースを用いた評価実験

本節では、提案する楽譜追跡手法において多声演奏に対する頑健性を検証する。具体的な評価対象は、オンセット検出において遅延決定および未来予測の枠組みを用いることによる性能比較を行う。

実験に用いるデータとして、MAPS クラシック音楽ピアノデータベース[22]から 60 曲、RWC ジャズ音楽データベース[23]から 50 曲を用いる。評価基準はオンセット検出の適合率、再現率、 $F$  値を用いる。前者のデータベースはピアノの多声演奏で構成され、後者はピアノのみでなくギター、バイオリンなど多くの楽器演奏で構成される。オンセット検出の許容誤差は、100 ms、300 ms、500 ms とする。また、遅延決定を行うための遅延幅は 0.0 s、0.5 s、1.0 s とする。表 2 にオンセット検出の結果をまとめたものを示す。 $F$  値の観点から見れば、100 ms の許容誤差で、0.5 s の遅延幅において最も高い結果が得られた。MAPS と RWC-Jazz データベースの両方で 30%以上の向上が見られた。これは、遅延幅を考慮することでオンセットの誤検出が回避できたことを示している。音楽情報処理に関するアルゴリズムの評価コンテストである MIREX の Real-time Audio-to-score Alignment のタスクにおける標準基準である 300 ms の許容誤差では、MAPS と RWC-Jazz データベースでそれぞれ 11%、19%の向上が得られた。

また、許容誤差が小さいほど、遅延幅が大きい場合に良い結果が得られていることがわかる。この結果は、アプリケーションによって遅延幅が決定されるべきであるということを示唆している。自動伴奏のような同期の精密さが求められるアプリケーションでは、100 ms の許容誤差で最も良い結果を得た 0.5 s の遅延幅が望ましいと考えられる。

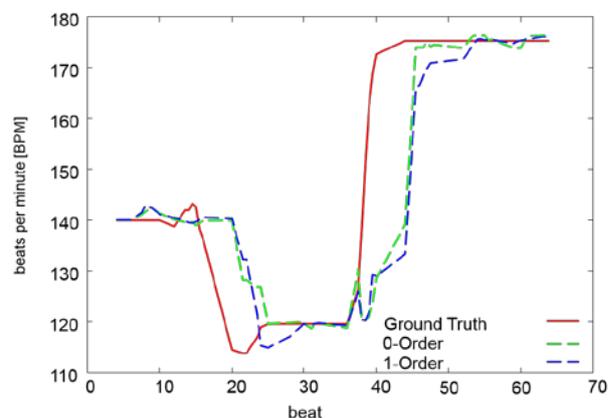


図 5. テンポ追従性能の評価

表 1. 実験結果 (2 種類のデータベースにおける適合率・再現率・F 値)

MAPS データベース												
基準	適合率				再現率				F 値			
	0.0	0.5 s	1.0 s	1.5 s	0.0 s	0.5 s	1.0 s	1.5 s	0.0 s	0.5 s	1.0 s	1.5 s
a												
	s											
100 ms	51.6	78.9	73.1	65.6	42.9	78.7	73.3	65.9	46.9	78.8	73.2	65.8
300 ms	89.0	91.8	92.7	92.3	73.1	91.5	93.0	92.7	80.3	91.7	92.8	92.5
500 ms	94.6	93.9	95.0	95.3	77.5	93.5	95.2	95.6	85.2	93.7	95.1	95.4

RWC-Jazz データベース												
基準	適合率				再現率				F 値			
	0.0	0.5 s	1.0 s	1.5 s	0.0 s	0.5 s	1.0 s	1.5 s	0.0 s	0.5 s	1.0 s	1.5 s
a												
	s											
100 ms	37.2	60.4	54.7	47.6	25.4	60.0	54.8	48.1	30.2	60.2	54.7	47.8
300 ms	74.7	79.1	80.3	79.1	49.4	78.2	80.3	80.0	59.5	78.7	80.3	79.6
500 ms	85.0	84.0	85.8	85.9	55.4	83.1	85.9	86.9	67.0	83.5	85.8	86.4

また、自動譜めくりのようなアプリケーションでは、同期の精密さは求められないため、より緩い許容誤差で高精度を得るような遅延幅が望ましいと考えられる。

大きな遅延幅は、小さい許容誤差において精度の悪化を示している。この結果は、遅延を考慮した過去の演奏位置推定とテンポを用いた未来予測のトレードオフの関係に起因していると考えられる。大きな遅延幅は、過去の位置から見てより未来の情報をアライメントに用いることができるため、過去の位置を高精度に推定できると考えられる。一方で、遅延幅の間でのテンポが一定であるという仮定により、現在位置の予測誤差は大きくなるのが予想される。

また、テンポ推定の結果は両データベースに対して、4%の許容誤差を考慮した場合には約 60%であり、それによる予測誤差の拡大に影響していると考えられる。

RWC-Jazz データベースに対する結果は、MAPS と比較して全体的に精度が低いことがわかる。この結果は、音響的な複雑度が影響していると考えられる。MAPS は、多声演奏ではあるがピアノ演奏のみである一方で、RWC-Jazz ではピアノだけでなく、ギター、クラリネット、またドラムスも含んでいるため、音響信号が複雑になり、観測モデルとの乖離が生じた結果である。遅延幅が 0 のベースラインでは、両データベースにおいて再現率は適合率よりも低い傾向にある。これは、多声演奏によって音響的な複雑度が増すほど、観測モデルと実際の信号の乖離が生じ、アルゴリズムを不安定にするからであると考えられる。しかし、再現率は遅延決定と未来予測の枠組みを用いることで大幅に改善されたことが見られる。これらの結果より、多声演奏に対する頑健性は示されたといえる。

## 6 まとめ

人と機械の協奏を実現する試みとして、多声楽器にも対応可能な音響入力による楽譜追跡手法について提案し、性能改善のためにスペクトルテンプレートの適応手法やテンポ推定精度の改善について取り組んだ。

今後の課題としては、楽譜追跡アルゴリズムの計算量の削減がある。本稿で提案した楽譜追跡手法は、不確実性を含む多声楽器の演奏にも頑健な一方で、楽譜に含まれる和音数に従って計算量が増加するため、曲の長さによってはリアルタイムでの動作が難しくなる問題があり、応用システムを開発する上で大きな障害となる。関連して、実用的な観点では、対象となる楽器音の区間を検出する、いわゆる音声認識における VAD (Voice Activation Detection) に類した機能なども検討する必要がある。

### 【参考文献】

- [1] R. Dannenberg, An on-line algorithm for real-time accompaniment, Proc. International Computer Music Conference (ICMC), pp.193-198, 1984.
- [2] B. Vercoe, The synthetic performer in the context of live performance, Proc. International Computer Music Conference (ICMC), pp.199-200, 1984.
- [3] J. Bloch and R. Dannenberg, Real-time computer accompaniment of keyboard performances, Proc. International Computer Music Conference (ICMC), pp.279-289, 1985.
- [4] 武田晴人, 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, vol.2006-MUS-66, no. 90, pp.109-116, 2006.

- [5] A. Cont, ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music, Proc. International Computer Music Conference (ICMC), 2008.
- [6] C. Raphael, Music plus one: A system for expressive and flexible musical accompaniment, Proc. International Computer Music Conference (ICMC), 2001.
- [7] C. Joder and et al., A conditional random field framework for robust and scalable audio-to-score matching, IEEE Transactions on Audio Speech and Language Processing (TASLP), vol. 19, no. 8, pp. 2385-2397, 2011.
- [8] 中村友彦, 水野 優, 鈴木孝輔, 中村栄太, 樋口祐介, 深山 覚, 嵯峨山茂樹, 音楽演奏の誤りや反復に頑健な音響入力自動伴奏, 日本音響学会 2012 年秋季研究発表会, pp. 931-934, 2012
- [9] 山本龍一, 酒向慎司, 北村正: 隠れセミマルコフモデルと線形動的システムを組み合わせた音楽音響信号と楽譜の実時間アライメント手法, 情報処理学会研究報告, vol. 2012-MUS-96(13), pp. 16, (2012).
- [10] 山本龍一, 酒向慎司, 北村 正, セミマルコフ条件付き確率場を用いた音楽音響信号と楽譜のアライメント, 日本音響学会 2012 年秋季研究発表会, pp. 935-936, 2012.
- [11] 山本龍一, 酒向慎司, 北村正, Ryry: 多声楽器に対応可能な音響入力自動伴奏システム, 情報処理学会シンポジウム インタラクシオン 3EXB-13, 2013.
- [12] R. Yamamoto, S. Sako, and T. Kitamura, {Real-time Audio to Score Alignment Using Semi-Markov Conditional Random Fields and Linear Dynamical System}, The Music Information Retrieval Evaluation eXchange (MIREX), 2012.
- [13] R. Yamamoto, S. Sako, and T. Kitamura, Robust On-line Algorithm FoReal-time Audio-to-score Alignment Based on A Delayed Decision and Anticipation Framework, Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2013.
- [14] S. Dixon, An On-line Time Warping Algorithm for Tracking Musical Performances, Proc. International Joint Conferences on Artificial Intelligence (IJCAI), pp. 1727-1728, 2005.
- [15] R. Dannenberg and Hu. N., Polyphonic audio matching for score following and intelligent audio editors, Proc. International Computer Music Conference (ICMC), pp. 27-34, 2003.
- [16] C. Raphael, Aligning music audio with symbolic scores using a hybrid graphical model, Machine Learning, vol. 65, no. 2-3, pp. 389-409, 2006.
- [17] A. Cont, A coupled duration-focused architecture for real-time music-to-score alignment, IEEE, Trans. on Pattern Analysis and Machine Intelligence (PAMI), vol. 32, no. 6, pp. 974-987, 2010.
- [18] T. Otsuka, K. Nakadai, T. Takahashi, K. Komatani, T. Ogata, and H.G. Okuno, {Design and implementation of two-level synchronization for interactive music robot, Proc. AAAI, 2010.
- [19] T. Otsuka, K. Nakadai, T. Takahashi, T. Ogata, and H.G. Okuno, Real-time audio-to-score alignment using particle filter for coplayer music robots, EURASIP J. Adv. Signal Process, 2011.
- [20] Cano P., Loscos A., and Bonada J., Score-Performance Matching Using HMMs, Proc. International Computer Music Conference (ICMC), pp. 441-444, 1999.
- [21] N. Orio and F. Dechelle, Score following using spectral analysis and hidden markov models, Proc. International Computer Music Conference (ICMC), pp. 151-154, 2001.
- [22] V. Emiya and et al, Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle, IEEE Transactions on Audio Speech and Language Processing (TASLP), vol. 18, no. 6, pp. 1643-1654, 2010.
- [23] M. Goto and et al, RWC Music Database: Popular, Classical, and Jazz Music Databases, Proc. International Society for Music Information Retrieval Conference (ISMIR), pp. 287-288, 2002.
- [24] K. Suzuki, Y. Ueda, R.S. A., N. Ono, and S. Sagayama, Real-time audio to score alignment using locally-constrained dynamic time warping of chromagrams, The Music Information Retrieval Evaluation eXchange (MIREX), 2011.

## 〈発表資料〉

題 名	掲載誌・学会名等	発表年月
Ryry: A Real-Time Score-Following Automatic Accompaniment Playback System Capable of Real Performances with Errors, Repeats and Jumps	Active Media Technology Lecture Notes in Computer Science, LNCS 8610	Aug., 2014
Violin Fingering Estimation According to Skill Level based on Hidden Markov Model	Proc. of ICMC SMC 2014	Sep., 2014