

# 自由なドラムス演奏の複数尤度を用いた HMM による自動伴奏

林耕平<sup>†1</sup> 高橋登紀夫<sup>†2</sup> 永田晃弘<sup>†2</sup> 嗟峨山茂樹<sup>†1</sup>

**概要:** これまで、人間の演奏に自動伴奏付けを行う研究は広く行われてきたが、奏者は与えられた楽譜に従って演奏を行うという前提があった。そこで、本稿ではドラムスという楽器における、より即興的な演奏に対して自動伴奏付けを行う手法について議論する。まず、自由なドラムス演奏をリズムパターンの同期遷移や接続によって行われるものとして近似し、HMM(Hidden Markov Model)を用いてモデル化を行う。モデルの確率的な逆問題を解くことで、演奏のリズムパターンや演奏箇所の推定を行う。具体的には複数の尤度計算とテンポ推定を行い、Viterbi アルゴリズムによってモデルの最尤状態を推定する。また、ここまでの手法を評価するために実験を行い、高い精度で演奏箇所の推定が行われることを示す。更に、本稿では自動伴奏付けの構想についても述べる。

## 1. はじめに

ドラムスは明確な音高を持たない複数の打楽器で構成されており、伴奏者としてのドラムス奏者は一般的にはリズムパターンを奏する役割を担う。ジャムセッション等の自由な演奏においては、ドラムス奏者はリズムやテンポを先導する役割を担うこともあり、他の奏者はそれに合わせて伴奏を行う。ドラムス奏者にとってはこのような演奏が醍醐味であるが、意図に沿った伴奏が可能な伴奏者がいなければ成立しない。そこで、このような伴奏生成を自動で行うシステムが実現できれば、ドラムス奏者にとって有用となり得る。

楽譜に基づいた人間の演奏に対して自動で伴奏付けを行う研究は、Dannenberg [1]や Vercoe [2]をはじめとして、以後盛んに行われてきた。楽譜を参照し、楽譜のどこを演奏しているか推定することは楽譜追跡と呼ばれ、同期した自動伴奏を行う上では不可欠である。東京大学のグループ [3]は演奏誤りや弾き直し、および任意箇所への弾き飛ばしに対応した楽譜追跡の手法を確立し、MIDI 入力の演奏における自動伴奏システム Eurydice [4]を実現させた。音響信号を入力とした楽譜追跡や自動伴奏の研究 [5-8]も行われており、MIDI 信号への変換が難しい管楽器なども自動伴奏の対象となった。

本稿では演奏の自由度を更に拡大し、楽譜に基づかないような自由なドラムス演奏への自動伴奏の実現を目指し、その手法について検討を行う。

## 2. ドラムス演奏のモデル化

### 2.1 自由なドラムス演奏の自動伴奏問題

前章で述べた自動伴奏システムを実現するためには、自由なドラムス演奏に同期した伴奏をリアルタイムで生成する必要があり、かつ伴奏内容は演奏内容に対応し音楽的に

妥当なものであることが望ましい。

しかし、自由なドラムス演奏を構成する要素には次のような様々な未知量が存在し、これらの変動は容易に予測できるものではない。

#### a) リズムパターン

音楽のジャンルによって演奏されるリズムパターンは異なり、打点が1つ増えるなどの微小な変動も異なるリズムパターンとして解釈できるため、無限のバリエーションが起り得る。

#### b) 演奏誤り

人間のドラムス演奏では叩き損ね、奏者の意図とは異なる楽器を鳴らしてしまう打ち間違いなどの演奏誤りが起り得る。

#### c) 演奏休止

奏者の判断で演奏を中断したとき、伴奏音が鳴り続けていれば任意のタイミングから復帰する可能性がある。そのため演奏は必ずしも連続的に行われるとは限らない。

#### d) テンポ変動

奏者の意図によってテンポの変動は任意に行われる。ただし、不安定な演奏にみられるテンポの揺らぎも、変動の一種として考えられる。

奏者の意図が反映された自動伴奏を実現するためには、これらの要素への対応が不可欠となる。そこで、次節より自由なドラムス演奏がどのように行われる現象であるか議論する。

### 2.2 自由なドラムス演奏の仮説

ジャムセッションなどの自由なドラムス演奏では、奏者の音楽的な経験により蓄積・記憶された多数のリズムパターン(リズム語彙)の内から意図したものが次々と選択され、連続的に演奏が行われると考えられる(有限語彙仮説)。

また、自由なドラムス演奏において、奏者はリズムパターンを断片的に繋ぎ合わせることで新たなパターンを生成

<sup>†1</sup> 明治大学大学院先端数理科学研究科  
Meiji University Graduate School of Advanced Mathematical Sciences

<sup>†2</sup> ローランド株式会社  
Roland Corporation

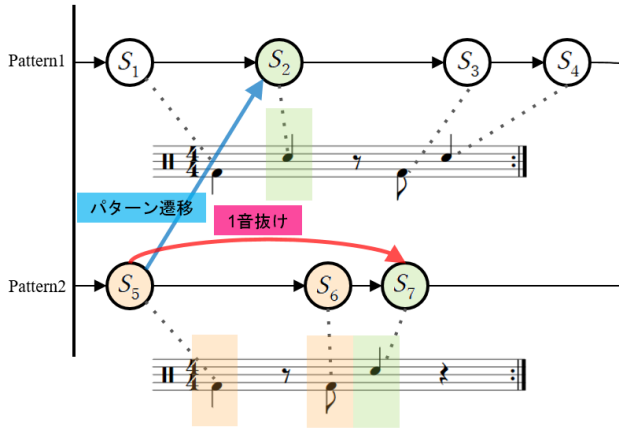


図1 ドラムス演奏モデルの状態遷移図

したり、任意のタイミングで異なるリズムパターンに遷移して演奏を行うことが可能である。このような遷移は共通の拍位置によって同期的に行われると考えられる(同期遷移仮説)。

本稿では、自由なドラムス演奏がこれら2つの仮説に基づいて行われるものとして扱う。

### 2.3 HMM を用いたドラムス演奏のモデル化

この仮説に基づいたドラムス演奏をモデル化し、入力信号からリアルタイムに演奏箇所を推定することで、自動伴奏付けを実現することが出来る。以下ではこのような自由なドラムス演奏のモデルを HMM によって記述する。本稿ではこのモデルを「ドラムス演奏モデル」と呼ぶ。

人間のドラムス演奏を観測すると、時刻情報とどの打楽器が叩かれたかという情報が得られる。そのためドラムス演奏は、直前の打点からの相対時刻を表す IOI と、発音楽器の種類の組み合わせで表現できる。ドラムス演奏モデルはこれらの出力を持つモデルとして記述できる。十分な種類のリズムパターンを譜面で表したものを、以下では「パターン譜」と呼び、これらのパターンの終点は始点に繋がるものとする(環状楽譜)。K個のパターン譜が存在するとき、iはk番目のパターン譜のn番目の音符を表すとすると、

$$i = \sum_{k=1}^K N_k + n \quad (1)$$

で表せる。ここで、 $N_k$ はパターンkにおける音符の数を表す。そこで、このようなリズムパターンの各音符を隠れ状態 $S_i$ とすると、 $S_i$ は楽譜時刻 $\tau_i$ と発音楽器 $c_i$ の情報を持ち、リズムパターン $k_i$ に属する。状態数の累計がI個であるとき、ドラムス演奏モデルの状態列を $S = \{(\tau_i, c_i)\}_{i=1}^I$ と表すものとする。ドラムス演奏は様々な状態遷移によって表される。

#### (ア) 順次進行モデル

同一パターンにおいて、演奏誤りがなければパターン譜

の通り演奏され、1つ先の状態へと順次移行する。以下、この遷移を遷移iと呼ぶ。

#### (イ) 打点脱落モデル

実際の演奏では奏者の誤りによって打点の脱落が起こる可能性がある。これを考慮し、打点の脱落を1音抜け遷移(以下、遷移iiと呼ぶ)として表す(図1)。

#### (ウ) パターン遷移モデル

遷移元の状態iが属するパターン $k_i$ と、遷移先の状態i'が属するパターン $k_{i'}$ において、 $k_i \neq k_{i'}$ となる場合、異なるパターンへの遷移とみなすことができる。特に、誤りのない演奏の場合、i'の楽譜時刻 $\tau_{i'}$ の直前の拍位置であるiからの遷移(以下、遷移iiiと呼ぶ)として記述できる(図1)。

#### (エ) その他のモデル

遷移i~iiiのどれにも該当しない遷移も、演奏休止からの復帰などによって起こり得る。ここではこのような遷移を遷移ivと呼ぶこととする。

これらの状態遷移の際、状態間の楽譜時刻の差 $\Delta\tau$ に従って IOI が出力される。ここで出力される IOI の系列を、 $X = x_1, x_2, \dots, x_{M-1}$ とする。Xの生成過程において、m番目に発音される打点の状態番号を $i_m$ とすると、このモデルによる、Xの生成確率 $p(X)$ は、

$$p(X) = \pi_{i_1} \prod_{m=2}^M b_{i_{m-1}, i_m}(x_{m-1}) a_{i_{m-1}, i_m} \quad (2)$$

で表せる。ここで、 $\pi_{i_1}$ は初期状態確率、 $b_{i_{m-1}, i_m}(x_{m-1})$ は状態遷移における $x_{m-1}$ の出力確率、 $a_{i_{m-1}, i_m}$ は状態遷移確率を表す。

一般的にドラムスは複数の打楽器で構成されるため、ドラムス演奏を表すには IOI の他に発音楽器の情報が必要である。そのため本稿ではドラムス演奏モデルを、状態遷移で起こる IOI 出力の他に、状態においても発音楽器が出力される HMM として記述する。

m番目の打点で出力される発音楽器を $y_m$ とし、状態iにおいて $y_m$ が出力される確率を $b'_i(y_m)$ と表記する。ここで、複数の打楽器が同時発音される場合も、その組み合わせを1つの発音楽器として扱うため、 $y_m$ は発音楽器の組み合わせを表す。出力される発音楽器の系列を、 $Y = y_1, y_2, \dots, y_M$ とすると、Yの生成確率 $p(Y)$ は、

$$p(Y) = \pi_{i_1} b'_{i_1}(y_1) \prod_{m=2}^M b'_{i_m}(y_m) a_{i_{m-1}, i_m} \quad (3)$$

で表される。

楽譜時刻 $\tau$ に対して演奏時刻を $t$ で表すと、ドラムスの演奏は、 $y_m$ が発音される時刻 $t_m$ を用いて演奏イベントの時刻系列 $Z = \{(t_m, y_m)\}_{m=1}^M$ と表すことができる。ここで、

$$x_m = t_{m+1} - t_m \quad (4)$$

であるため、 $x_m$ は $t_{m+1}$ の、 $t_m$ からの相対的な値を表す。そのため、 $t_1$ が既知であるとき演奏イベント列Zが生成され

る確率 $p(Z)$ は、

$$p(Z) = \pi_{i_1} b'_{i_1}(y_1) \prod_{m=2}^M b'_{i_m}(y_m) b_{i_{m-1}, i_m}(x_{m-1}) a_{i_{m-1}, i_m} \quad (5)$$

と表される。

### 3. 複数尤度を用いた楽譜追跡

#### 3.1 ドラムス演奏の楽譜位置推定

自由なドラムス演奏への自動伴奏を実現するためには、奏者の意図したリズムパターンと、その演奏箇所をリアルタイムで推定する必要がある。本稿ではこれらを推定する問題を演奏の楽譜位置推定問題と呼び、以下では前章で記述したドラムス演奏モデルを基に楽譜位置を推定する手法について述べる。

ドラムス演奏モデルの各状態はパターン $k_i$ に属し、楽譜時刻情報 $\tau_i$ を持っているため、楽譜位置推定問題は、ドラムス演奏モデルの逆問題として捉えられる。このような確率モデルにおいては、楽譜内の全ての箇所が現在の演奏の楽譜位置の確率は0でなく、その意味では楽譜内の全ての箇所が解の候補である。ここでは条件付き確率の最大化問題として、尤度最大となる状態の番号 $i_M$ を解として推定する。リアルタイムでの楽譜位置推定では、 $M$ の値が随時更新されていくため、最尤状態の時系列を $Q = \{(t_m, i_m)\}_{m=1}^M$ で表す。

演奏イベントの時系列 $Z = \{(t_m, y_m)\}_{m=1}^M$ が、ドラム演奏モデルの状態列 $S = \{(\tau_i, c_i)\}_{i=1}^I$ に基づいているとすると、最尤状態の時系列 $Q$ は、Bayesの定理より、

$$\begin{aligned} \operatorname{argmax}_Q p(Q|Z) &= \operatorname{argmax}_Q [p(Z|Q)p(Q)] \\ &= \operatorname{argmax}_{i_1, \dots, i_M} \left[ \pi_{i_1} b'_{i_1}(y_1) \prod_{m=2}^M b'_{i_m}(y_m) b_{i_{m-1}, i_m}(x_{m-1}) a_{i_{m-1}, i_m} \right] \quad (6) \end{aligned}$$

を求めることで解かれる。

#### 3.2 ドラムス演奏モデルの尤度

ドラムス演奏モデルの状態遷移や出力は確率的に行われるため、観測されたドラムス演奏からモデルの状態を推定する逆問題では、これらの確率によって遷移や状態に様々な尤度が付与される。以下では各尤度について議論する。

##### A) 状態遷移確率による尤度

ドラムス演奏モデルでは、同一パターン内で状態遷移 $i$ が繰り返されることが基本であるが、同期遷移仮説に従うと、任意の状態から他のパターンへ遷移が起こることも許容する必要がある。このような他パターンへの遷移尤度を $L_{i,j}^{\text{tran}}$ を用いて表す。

また、同一パターン内の1音抜け遷移は、奏者の演奏誤りによって生じる打点の脱落のモデル化であり、ここでは1音抜け遷移の尤度を $L_{i,j}^{\text{skip}}$ と表す。

##### B) IOIの状態遷移出力確率による尤度

ドラムス演奏モデルで出力されるIOIの値 $x$ は、状態間の楽譜時刻の差 $\Delta\tau$ に従って定められるが、 $x$ と $\Delta\tau$ は必ずしも一定の比をとるとは限らず、実際には確率分布をなすと考えられる。この要因として、奏者の技術的な制約による打点時刻の誤差やテンポ変動などが挙げられる。IOIの出力確率 $b_{i,j}(x)$ はこの分布に従って与えられるため、状態遷移には尤度 $L_{i,j}^{\text{IOI}}(x)$ が付与される。 $x$ が正規分布をなすと仮定すると、

$$L_{i,j}^{\text{IOI}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{x}{v} - \Delta\tau\right)^2}{2\sigma^2}\right) \quad (7)$$

で表される。ここで、 $\sigma^2$ は分散であり、この値が大きいかほど緩やかな概形となるため、不安定な演奏にどれだけ追従するかを定めるパラメータとしての意味をもつ。 $v$ は現在のテンポを表すが、本稿ではテンポの定義を楽譜時間に対する実時間の長さ $\frac{\Delta t}{\Delta\tau}$ とする。

##### C) 発音楽器の状態出力確率による尤度

人間によるドラムス演奏では、意図した打楽器と異なる打楽器を叩いてしまう演奏誤りが起こる可能性がある。そのため、奏者の意図した発音楽器と実際の演奏で発音される楽器は必ずしも一致するとは限らない。このような発音楽器の打ち間違い尤度を $L_i^{\text{tone}}(y)$ と表す。

#### 3.3 各尤度を用いた Viterbi 経路推定

それぞれの尤度は独立に与えられるものと仮定すると、最尤状態を各尤度の掛け合わせで求めることができる。これは式(6)の $Q$ と $Z$ の同時確率の最大化と同義である。また、 $i_M$ の推定は Viterbi アルゴリズムを用いることで効率的に行われる。 $i_{M-1}$ の最尤値 $\hat{p}_{i_{M-1}}$ は各尤度を用いて、

$$\hat{p}_{i_{M-1}} = \max_{i_1, \dots, i_{M-2}} \left[ \pi_{i_1} L_{i_1}^{\text{tone}}(y) \prod_{m=2}^{M-1} L_{i_{m-1}, i_m}^{\text{IOI}}(x) L_{i_m}^{\text{tone}}(y) L_{i_{m-1}, i_m}^{\text{tran}} L_{i_{m-1}, i_m}^{\text{skip}} \right] \quad (8)$$

と表せる。 $\hat{p}_{i_M}$ はこれを用いて、

$$\hat{p}_{i_M} = \max_{i_{M-1}} \left[ L_{i_{M-1}, i_M}^{\text{IOI}}(x) L_{i_M}^{\text{tone}}(y) L_{i_{M-1}, i_M}^{\text{tran}} L_{i_{M-1}, i_M}^{\text{skip}} \hat{p}_{i_{M-1}} \right] \quad (9)$$

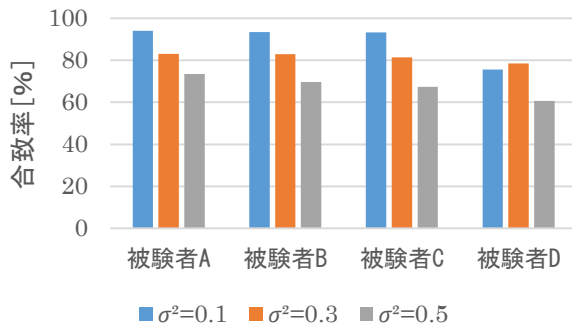
と表せることから、Viterbi アルゴリズムによって、帰納的に最尤な状態番号 $i_M$ を求めていくことができる。

#### 3.4 一次遅れ系平滑化によるテンポ推定

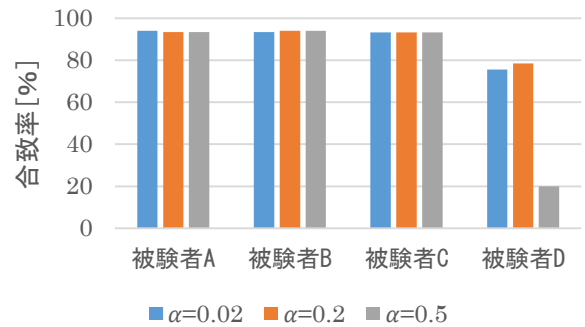
演奏に同期した自動伴奏を実現するには、楽譜位置情報の他にもテンポ情報が不可欠である。自由なドラムス演奏のテンポは奏者に委ねられ、常に変動が起こり得る。そのため演奏イベントが更新されるたびに奏者の意図したテンポを逐次推定する必要がある。

局所的なテンポ変動が起こり得ることから、現在のテンポは過去の局所的なテンポに依存すると考えられる。そこで、 $M(\geq 2)$ 番目の打点における局所的なテンポ $r_M$ を、

$$r_M = \frac{x_{M-1}}{\tau_{i_M} - \tau_{i_{M-1}}} \quad (10)$$



IOIの確率分布モデルの分散 $\sigma^2$ の3種の値設定時の合致率



テンポ一時遅れ平滑化の係数 $\alpha$ の3種の値設定時の合致率

図2 楽譜位置推定結果の比較

と定義する。 $r_M$ を推定テンポに逐次取り込むことで、反映が可能となる。本稿では、 $M \geq 2$ における推定テンポ $\hat{v}_M$ を、 $0 \leq \alpha \leq 1$ である係数を用いて次のように表す。

$$\hat{v}_M = \alpha r_M + (1 - \alpha) \hat{v}_{M-1} \quad (11)$$

ここで、 $\alpha$ は $r_M$ の反映度を表すパラメータであり、 $r_M$ の変動度合いに応じて調整することが望ましい。

#### 4. 楽譜位置推定の評価

前章で述べた楽譜位置推定法を評価するために、実験を行った。本実験では、楽器初心者から経験者までを含めた4名の被験者を集め、実際にリズムパターンの演奏をしてもらった。ただし、ドラムスの代用としてMIDIキーボードを用いて、鍵盤を叩くことを入力とした。

被験者には4/4拍子1小節のリズムパターンを6種類提示し、演奏順を定めたシナリオ(リズムパターンの接続・繰り返しを含む)に沿って演奏をしてもらった。演奏の自由度として、演奏誤り(音抜け・鍵盤の打ち間違い)と、初期テンポ(120BPM)からの自由なテンポ変動を許容するものとした。実験によって得られた入力データを用いて楽譜位置推定を行い、推定結果とシナリオの楽譜位置の合致率を精度として、パラメータ値による比較を行った。式(7)で用いた正規分布の分散値 $\sigma^2$ と、テンポ推定における係数 $\alpha$ をパラメータとした。

##### ① $\sigma^2$ の比較

まず、 $\alpha$ を固定し(ここでは $\alpha = 0.02$ とした)、 $\sigma^2$ を3通りの値(0.1, 0.3, 0.5)に定めて楽譜位置推定を行った(図2、左)。被験者A~Cでは $\sigma^2 = 0.1$ で最も高い合致率が得られた。これは被験者A~CのIOIによれが殆ど見られなかったことが要因であり、 $\sigma^2$ が小さいほど、打点の誤差に対する尤度 $L_{i,j}^{IOI}(x)$ の下がり幅が大きいため、正確な演奏では高い精度で推定結果が得られることを示している。被験者Dのみ $\sigma^2 = 0.3$ が最も高い合致率となったのは、被験者Dが楽器初心者であり、推定テンポに対する入力IOIの誤差が顕著に見られたためと考えられる。

##### ② $\alpha$ の比較

次に、 $\sigma^2 = 0.1$ に固定し、 $\alpha$ を3通りの値(0.02, 0.2, 0.5)に定めて楽譜位置推定を行った(図2、右)。 $\alpha = 0.02, 0.2$ ではどの被験者も高い合致率が得られたが $\alpha = 0.5$ では被験者Dのみ合致率が著しく低下した。これは、被験者Dの演奏の局所的なテンポ $r_M$ に揺らぎが見られ、推定テンポに大きく反映されてしまったためと考えられる。

①、②の結果より、最適なパラメータは奏者の特徴によって異なり、適切な値を定めることが高精度な楽譜追跡を行うために必要であるといえる。

#### 5. 伴奏再生と伴奏同期尤度

これまで述べた問題を解くことで、自由なドラムス演奏に対応、同期した自動伴奏が実現できる。具体的には、それぞれのパターン譜に対応した伴奏音源を用意し、得られた楽譜位置に対応した伴奏箇所を推定テンポで再生、随時更新することで行われる。

これまで、自由なドラムス演奏が独立して行われる前提で議論を進めたが、実際の演奏では伴奏音の影響を受ける可能性がある。例えば、演奏と伴奏が完全に同期していない場合、伴奏音は奏者にとってそれを修正するための指標であるため、特に打点時刻についてはその影響を無視できない。演奏に同期した伴奏付けを行うためには、このような伴奏音の影響を自動伴奏に反映する必要がある。

そこで、IOIの伴奏音の影響による尤度として、3.2節のB)で述べた尤度 $L_{i,j}^{IOI}(x)$ の他に、伴奏同期尤度 $L_{i,j}^{\text{back}}(d)$ を用いる。伴奏同期尤度を反映することによって、2.1節のc)で挙げたような、演奏休止からの復帰に対応した伴奏付けが可能となる。伴奏箇所に対応する楽譜時刻 $\tau_{\text{back}}$ と、遷移先の状態の楽譜時刻 $\tau_j$ の距離 $d = |\tau_j - \tau_{\text{back}}|$ が正規分布をとると仮定し、

$$L_{i,j}^{\text{back}}(d) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (12)$$

で表す。これを反映した最尤値を $\hat{p}_{i_m}$ とすると、

$$\hat{p}_{i_M}^t = \max_{i_{M-1}} \left[ L_{i_{M-1}, i_M}^{IOI}(x) L_{i_M}^{\text{back}}(d) L_{i_M}^{\text{tone}}(y) L_{i_{M-1}, i_M}^{\text{tran}} L_{i_{M-1}, i_M}^{\text{skip}} \hat{p}_{i_{M-1}} \right] \quad (13)$$

によって更新される。

## 6. 結論と展望

本稿では、自由なドラムス演奏への自動伴奏を実現するための手法と理論について検討した。リズムパターンの同期遷移と接続によって行われるドラムス演奏を HMM でモデル化し、このモデルに付与される様々な尤度を用いた演奏位置の最尤推定法と、局所的なテンポ変動を取り込んだテンポ推定法についても示した。評価実験によって、これらを用いた楽譜位置推定法が高精度で行われることを確認した。

今後の課題としては、構想段階であった伴奏再生の実現や、伴奏同期尤度の妥当性の評価などがある。また、本稿で検討した楽譜追跡法では、急激なテンポの変動が伴奏へすぐに反映されにくく、より同期性の高い楽譜追跡を実現することも課題の1つである。パラメータの更なる調整や、テンポ推定法に加速度を取り込むことで精度向上が見込める可能性がある。更に、演奏の入力にベロシティ情報などを加えたモデルについても検討していきたい。

## 参考文献

- [1] Roger Dannenberg, "An on-line algorithm for real-time accompaniment," Proceeding of International Computer Music Conference (ICMC) 1984, pp. 193-198, 1984.
- [2] Barry Vercoe, "The synthetic performer in the context of live performance," Proceeding of International Computer Music Conference (ICMC) 1984, pp. 199-200, 1984.
- [3] 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-066, pp. 109-116, 2006.
- [4] 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, "任意箇所への弾き直し・弾き飛ばしを含む演奏に追従可能な楽譜追跡と自動伴奏," 情報処理学会論文誌, Vol. 54, No. 4, pp. 1338-1349, 2013.
- [5] Arshia Cont, "ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music," Proceeding of International Computer Music Conference (ICMC), 2008.
- [6] Christopher Raphael, "Music plus one: A system for expressive and flexible musical accompaniment," Proceeding of International Computer Music Conference (ICMC), 2001.
- [7] 鈴木孝輔, 上田雄, 齋藤康之, 小野順貴, 嵯峨山茂樹, "HMM を用いた音響演奏の楽譜追跡による弾き直しに追従可能な自動伴奏," 情報処理学会研究報告, MUS, pp. 1-6, 2011.
- [8] 中村友彦, 中村栄太, 嵯峨山茂樹, "誤り・任意の弾き直し・引き飛ばしを含む演奏音響信号への高速な楽譜追跡," 情報処理学会研究報告, MUS, 2013.