

ジャムセッションシステムに向けた数理モデル化と 演奏特徴量の統計解析

保利武志[†] 中村和幸[†] 嗟峨山茂樹[†]

概要：ジャズのようなアドリブ演奏を基本としたセッションでは、それまでの演奏から、他の演奏者がどのような意図をもって演奏しているのかを推定して適切な演奏をする必要がある。システムがそのような機能を有するためには、実際の演奏で観測される様々な特徴量からそれらを推定する必要があり、従来の研究ではいくつかの閾値やパラメータ等をヒューリスティックに決定していた。本研究では実際の演奏 MIDI データのみから各楽器演奏者による楽器間の対応関係を統計的に学習可能なセッションシステムの数理モデルを提案する。そこで、演奏はあらゆる音楽事象を内包する高次元空間におけるトラジェクトリとして表現されるものであると考えた場合、これらを学習するためにはその連続性と高次元性が問題となるため、特徴量を選別し、また空間を分割、あるいは補間することで演奏を予測する必要がある。本稿では空間の分割手段としてのクラスタリングを焦点とし、k-means、GMM、NMF による3つのクラスタリング手法について、それぞれ N-gram 及び HMM による演奏予測精度によって比較評価した。その結果、NMF がクラスタリング手法としてもっとも適していることが示された。

キーワード：ジャムセッションシステム, HMM, k-means, GMM, NMF

1. はじめに

人間と計算機とが協調演奏を行う自動伴奏システム [1]では、リアルタイムに人間の演奏を追跡して楽譜とのマッチングを行う Score Following によって正確な演奏位置が推定される。しかし、この時人間の演奏は必ずしも楽譜通りに進行されるのではなく、演奏誤りや弾き直し、弾き飛ばし、テンポ変化等を考慮する必要があるため、これらを確率的に推定する手法がこれまで数多く提案されている。我々が開発した自動伴奏システム Eurydice[2]は広範囲に渡る小節間の弾き飛ばしにもフレキシブルに対応して人間の演奏に追従することができる。

そこで我々は次のステップとして、即興演奏にも対応できる自動伴奏システムの開発を目的とし、特にインタラクティブな即興演奏が頻繁に行われるジャズを対象としたジャムセッションシステムの開発に向けて、統計的なアプローチから学習可能なセッションシステムの数理モデリングを行った。

ジャムセッションで用いられる楽曲は主としてジャズのスタンダードである。一般にジャズスタンダードではイントロ、リードシート通りに演奏を行うテーマ部、各楽器を主としたアドリブパート、そして再度テーマ部が演奏された後のアウトロといった構成を取ることが多い。特にこのアドリブパートでは、ある演奏者は他の演奏者の音を聞き、その演奏意図を解釈して演奏を決定するというインタープレイによって様々な演奏が展開される。このようなセッションを実現するために、計算機は過去の演奏から、ある演奏が行われた時、次にどのような演奏が行われるのかを推定することによって、適切な演奏を出力する必要がある。 a

従来のジャムセッションに関する研究[3-6]では、事前に設計されたパラメータや、即興演奏の掛け合いを前提とした直前の演奏のみを対象とした入力データを用いて、人間の演奏に追従する形で演奏が静的に決定されていたため、計算機に個性を持たせることはできなかった。[7,8]はジャズのピアノトリオを対象としたジャムセッションシステムであり、テンションパラメータ[7]や主張度・主導率[8]といった、演奏中に動的に変化するパラメータを導入することにより、計算機に個性を持たせ起伏のあるセッションを実現していたが、入力された演奏に対する演奏出力の対応関係は演奏特徴量の閾値を主観的に設定することによって決定されていた。[9]はギタートリオを対象とし、その演奏の入出力関係を実際の演奏データから統計的に学習することによって、ある演奏者の振る舞いを模倣することを可能としたが、特定の被験者に対する心理実験結果を用いて学習を行っていたため、[7,8]と同様、パラメータやルール群が主観的に決定されていたといえる。

本研究では従来ヒューリスティックに決定されていたパラメータを統計的に学習し、実際の演奏データのみから演奏の予測及び楽器間の対応関係を決定してインタラクティブにセッションを行うシステムを提案する。

2. モデリング

2.1 システムの構成

本研究の目的は、可能な限り主観的、あるいはヒューリスティックな部分を排除した、統計的に学習可能なジャムセッションシステムの数理モデル化及びその実現である。Fig. 1 にシステムの全体像を示した。本稿ではジャズピアノトリオを対象とし、人間がピアノ、計算機がベース及びドラムスを担当することとした。また、入出

[†] 明治大学先端数理科学研究科
E-mail: {cs51003, knaka, sagayama}@meiji.ac.jp
〒164-8525 東京都中野区 4-21-1

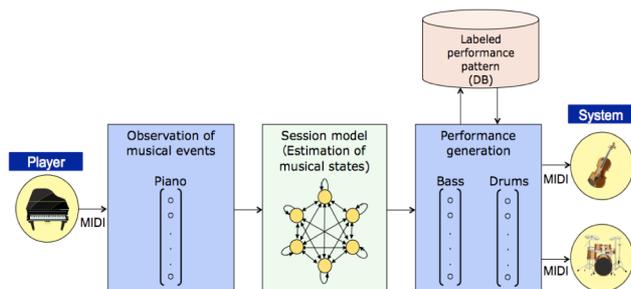


Figure 1 – システムの全体像

力は MIDI とした。これは、学習に際しては大量の演奏データが必要となり、またその精度も重要であるため、容易に正確な演奏データが得られることを期待してのものである。

2.2 セッションの数理モデル

演奏を数学的に表現するために、我々はまず音楽に関わるあらゆる情報を内包する高次元空間を定義した。このような理想的な高次元空間を音楽事象空間 Ω として定義すると、この空間は、例えば音それ自体に関するもの(音数、音量、音価等)やジェスチャー、アイコンタクト等様々な情報による軸によって表現されているとみなすことができる。すなわち、ある時刻における音楽情報を高次元特徴量ベクトルとして捉えて音楽事象空間 Ω で表現したとき、その演奏特徴量を毎時 Ω 内にプロットすると、その高次元時系列ベクトルによって表現されたトラジェクトリ T を用いて、演奏は (Ω, T) のように表すことができる。

次にセッションに拡張する。本稿ではジャズピアノトリオを対象としているので、本節においてもピアノトリオを例とする。個別に各楽器の演奏を音楽事象空間 Ω で見た場合、それぞれの演奏は同様に (Ω, T) として表されるが、セッションにおいては楽器間のインタラクションが行われ、また音楽理論等を基盤とした協調演奏や前の演奏からの連続性が考慮されるために、そのトラジェクトリ T は必ずしも自由な軌道を描くことができるわけではなく、確率的な偏りをもつ。すなわち、各楽器のトラジェクトリ T は互いの共起確率と時間的・音楽的ルールによる依存関係によって、確率的な偏りをもって変化すると考えられる。Fig. 2 はその理想的なピアノトリオのセッションモデルである。

実際に有限なデータからこれらを学習するためには、音楽事象空間 Ω の高次元性及びトラジェクトリ T の連続性、また実際に得られるデータのスパース性等が問題となる。そこで、本稿ではトラジェクトリ T を離散化して確率的に状態が遷移すると捉え、これを Hidden Markov Model (HMM) によってモデル化した。空間及びトラジェクトリを離散化することによって、確率的な偏りを持つ

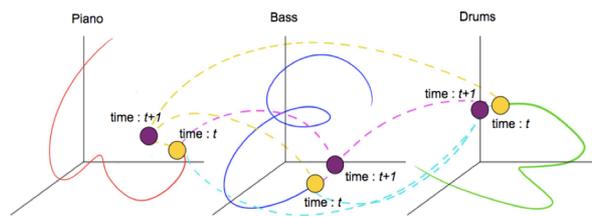


Figure 2 – ピアノトリオの理想的なセッションモデル

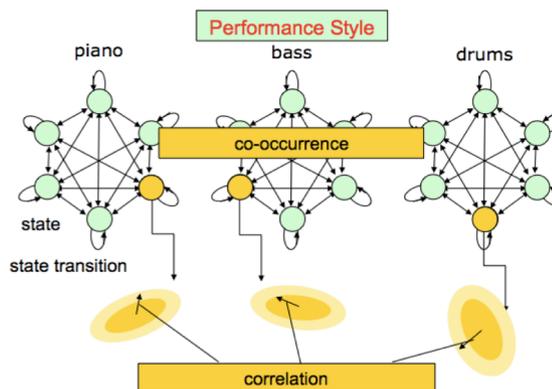


Figure 3 – ピアノトリオの離散化状態遷移モデル(HMM)

軌跡は Fig. 3 のように確率的状態遷移モデルとして表現できる。また楽器間のインタラクション、すなわち共起関係は隠れ状態の共起関係としてモデル化できる。ただし、状態数に依存してモデルの表現力が決定してしまうため、Fig. 3 のような隠れ状態数を 6 としたのみでは非常に粗いモデルとなる。そこで、各状態における特徴量ベクトルのセントロイドからの偏移の相関関係を導入することによって、より表現力を高めることとした。また、高次元性の問題に対しては、実際の演奏 MIDI データのみから得られる、ジャズセッションに特に関わりのあるパラメータに限定して抽出し、これをスタイルパラメータと定義することによって次元削減し演奏を表現した。

また、システムは学習とリアルタイムな推論の 2 つの機能を有する必要がある。まずシステムは上記のモデルに従って、大量の MIDI 演奏データから得られるスタイルパラメータを用いて各楽器間の応対関係及び、各楽器における演奏の予測をするための学習を行う。次に、その学習によって得られたパラメータを用いて、Fig. 1 のように、MIDI ピアノから得られたスタイルパラメータから、リアルタイムに遷移する全楽器の状態及び遷移を予測し、ルール群によってマッピングされた演奏データから、ベース及びドラムスの適した演奏を出力する。

2.3 スタイルパラメータ

低次元化のため、実際のジャズセッションによって得られた MIDI データのみから特にジャズセッションに関

わる特徴量として、以下のようにスタイルパラメータを定義した。具体的には、小節単位及び拍単位ごとに抽出できる演奏特徴パラメータに限定しており、その選定にあたってはジャズ理論やジャズ演奏家の助言を元に以下の特徴量を抽出、計算した。

- ピアノに関する特徴量
 - ・ ダイアトニックコードの構成音、テンションノート数、アヴォイドノート数、ブルーノート数
 - ・ 音域幅
- ベースに関する特徴量
 - ・ 音域幅
- ドラムスに関する特徴量
 - ・ ハイハットシンバル音数、スネアドラム音数、クラッシュシンバル音数、その他の音数
- 全楽器共通の特徴量
 - ・ 音数、同時音数、平均ベロシティ
 - ・ 上述の特徴量について、前小節または前拍の数で除した値
 - ・ 上述の特徴量について、合計数で除した値
 - ・ 裏拍の場合 1、そうでない場合は 0

以上の特徴量から、小節単位、拍単位ともにスタイルパラメータはそれぞれ 68 次元の時系列ベクトルとして定義された。

3. クラスタリング

音楽事象空間内の演奏(Ω, T)を HMM によって表現し実際の演奏データから学習可能なものとするために、空間及びトラジェクトリをクラスタリングによって離散化する。クラスタリングには以下の 3 手法: (1) *k*-means, (2) Gaussian Mixture Model (GMM), (3) Non-negative Matrix Factorization (NMF) を用い、演奏予測の推定精度によって比較検討を行った。

3.1 *k*-means

K-means はある観測データに対し、各クラスターのセントロイドを代表点として、その点から最も近いクラスのラベルを付与する手法である。

本研究では *k*-means 処理を行う前に、68 次元の各スタイルパラメータに対してさらに Principle Component Analysis (PCA) を用いて、寄与率 90%を閾値として次元削減を行った。

3.2 ガウス混合モデル(Gaussian Mixture Model, GMM)

GMM はある確率分布を複数の正規分布の重ね合わせによって表すモデルであり、EM アルゴリズムによって

尤度を最小化するようにパラメータ(平均, 分散, 混合率)の更新が行われる。

本研究では *k*-means 同様に PCA によって寄与率 90%を閾値として次元削減したベクトルに対して EM アルゴリズムを適用した。

3.3 非負値行列因子分解(Non-negative Matrix Factorization, NMF) [10]

近年, NMF もまたクラスタリングの手法として用いられている[11]. NMF は非負値行列 X を、同様に非負値行列である H と U の行列積によって近似する手法であり、以下のように表される。

$$X \approx HU$$

H を基底行列, U をアクティベーション行列とし、基底のインデックスを k として表すと、上式は、

$$X_{i,j} \approx \sum_k H_{i,k} U_{k,j}$$

と分解できる。 H と U の各要素に対してある値が与えられた時、元の非負値行列 X と行列積 HU の距離を最小化するように H , U の要素の値が更新される。その更新式は距離の定義によって異なり、一般にはユークリッド距離, Kullback-Leibler ダイバージェンス, 板倉斎藤距離のいずれかが用いられる。なお本研究では X と HU のフィッティングの程度から Kullback-Leibler ダイバージェンスを適用した。この場合、各要素は以下の式によって乗法更新される。

$$h_{ik} \leftarrow h_{ik} \frac{\sum_j \frac{x_{ij}}{\hat{x}_{ij}} u_{kj}}{\sum_j u_{kj}}, \quad u_{kj} \leftarrow u_{kj} \frac{\sum_i \frac{x_{ij}}{\hat{x}_{ij}} h_{ik}}{\sum_i u_{ik}}$$

なお、パラメータ更新の際、基底のスケール任意性を除去するため、各基底の L1 ノルムを 1 に揃えた。

$$\sum_i h_{ik} = 1$$

以上の更新式によって与えられた基底行列 H を用い、ある観測ベクトル x_j に対して最も大きいアクティベーションを与えた基底ベクトルのインデックス k をクラスラベルとして付与した。

$$c_k(x_j) = \arg \max_k u_{kj} \quad (1)$$

4. 評価実験

各クラスタリングの手法について, trigram 及び Hidden Markov Model (HMM)による演奏予測精度によって比較評価を行った。学習には YAMAHA Music Datashop [12] からジャズピアノ 13 曲を用い、正答率は交差検定によって算出した。

テストデータのクラスラベルはそれぞれ以下のように

与えた.

(i) *k*-means

学習時に得られた各クラスの重心ベクトルを用い、ユークリッド距離において最小となるクラスラベルを付与した.

(ii) GMM

学習時に得られた各クラスの平均及び分散を用い、最も高い確率を与えるクラスラベルを付与した.

(iii) NMF

学習時に得られた基底行列 H_{sample} を用いて求めた一般化逆行列 H_{sample}^+ を用いて、学習データ X_{sample} 及びテストデータ X_{test} のアクティベーション行列 U_{sample} , U_{test} を求め、(1)式からクラスラベルの付与を行った.

$$U_{sample} = H_{sample}^+ X_{sample}$$

$$U_{test} = H_{sample}^+ X_{test}$$

4.1 trigram

n 個の状態 (s_1, s_2, \dots, s_n) が与えられた時、その連鎖確率 $P(s_1, s_2, \dots, s_n)$ を N -gram によって表すと、

$$P(s_1 s_2 \dots s_n) = \prod_{i=1}^n P(s_i | s_{i-N+1} \dots s_{i-1})$$

となる. 特に $N=3$ のときを trigram といい、クラス遷移列 s_{i-2}, s_{i-1}, s_i を $N(s_{i-2}^i)$ と表した時、その連鎖確率は、

$$P(\omega_i | \omega_{i-2}^{i-1}) = \frac{N(\omega_i^{i-1})}{N(\omega_{i-2}^{i-1})}$$

で与えられる.

Fig. 4-6 は全楽器特徴量、楽器別それぞれについて、*k*-means, GMM, NMF によるクラスタリングによって trigram による演奏予測を行った時の正答率である. Fig. 4 は小節単位でスタイルパラメータを抽出して予測精度を求めたもの、Fig. 5 は拍単位で予測精度を求めたもの、そして Fig. 6 は小節クラス数を 10 と固定したときの、各小節クラス別の拍クラス遷移を見た階層構造における連鎖確率を求めて演奏予測を行った結果のグラフである. 小節単位において全楽器を用いた場合の予測精度は 3 つの手法に有意な差は見られなかったが、楽器別、あるいは拍単位、階層構造においては NMF によるクラスタリングにおいて最も高い精度を得られた.

4.2 Hidden Markov Model (HMM)

HMM はマルコフ性を持って確率的に遷移する観測できない各隠れ状態から、それぞれの確率分布に従って事象が生成されるとした統計的なモデリング手法であり、一種の動的計画法と EM アルゴリズムを用いた Baum-Welch アルゴリズムによってパラメータが求められる.

状態 $i - j$ 間の遷移確率を a_{ij} , 状態 i において時刻 t に

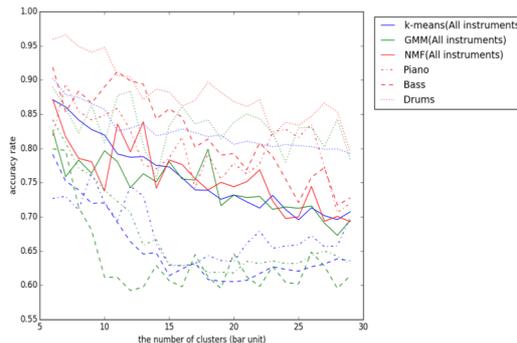


Figure 4 – 小節単位の演奏予測精度正答率(クラス数 6-29)

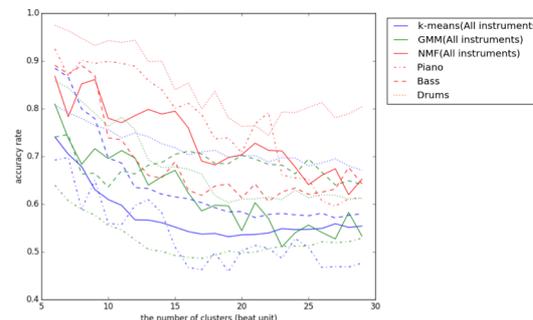


Figure 5 – 拍単位の演奏予測精度正答率(クラス数 6-29)

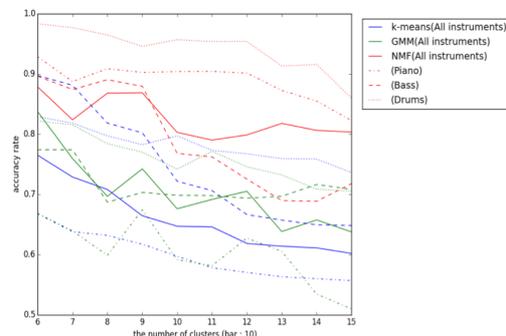


Figure 6 – 階層構造時の演奏予測精度正答率(小節クラス数 10, 拍クラス数 6-15)

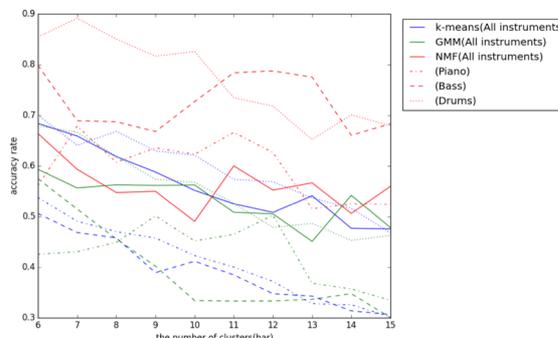


Figure 7 – HMM における小節単位の演奏予測精度正答率(クラス数 6-29, 隠れ状態数 6)

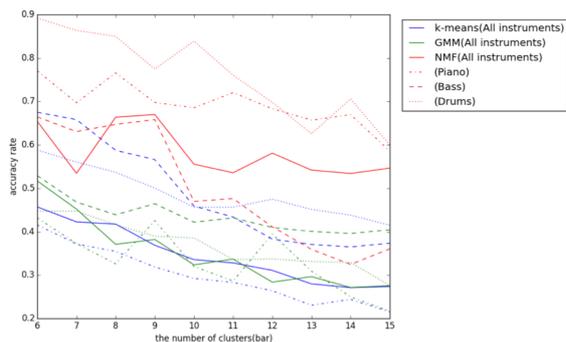


Figure 8 – HMM における拍単位の演奏予測精度正答率
 (クラス数 6-29, 隠れ状態数 6)

観測される出力信号を $b_t(o(t))$ とし, Forward・Backward アルゴリズムによって求められた前向き確率 $\alpha_t(i)$ と後ろ向き確率 $\beta_t(i)$ を用いて, E-step において,

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{Pr[O|\lambda]}$$

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{Pr[O|\lambda]}$$

を求めることにより, M-step において,

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_n(i, j)}{\sum_{i=1}^{T-1} \gamma_t(i)}$$

$$b_i(k) = \frac{\sum_{t=1, s.t. o(t)=k}^T \gamma_t(i)}{\sum_{i=1}^T \gamma_t(i)}$$

としてパラメータを更新する.

本研究では隠れ状態を演奏スタイルとし, それぞれの演奏はこの状態と楽器間の偏移の相関によって記述されるものであると定義しているが, 本節では前節同様全特徴量を用いた全楽器ベクトル及び楽器別のクラス遷移のみから, それぞれのクラスターリング手法について演奏予測を行った. HMM による演奏の予測には Forward アルゴリズムを用い, 全曲に対して交差検定を行うことで正答率を評価した.

Fig. 7-8 は隠れ状態数を 6 としたときの, 小節単位, 拍単位における予測精度である. trigram 時と同様に, 小節単位の全特徴量ベクトルに対して予測精度を求めた場合は, 各クラスターリング手法の正答率に有意な差は得られなかったが, 楽器別, あるいは拍単位においては NMF で最も高い予測精度が得られた.

trigram と HMM の両者においても NMF が最も高い精度を得られた理由として, 一つは学習時にいずれかのクラスラベルを明確に付与するハードなクラスターリング手法である k -means と異なり, 全クラスに属する可能性を残すソフトなクラスターリングを行っていることが考えられる. これは同様にソフトな学習を行っている GMM の方が k -means に比べて比較的良い精度を与えていることから推測できるが, GMM に対して NMF の方が高い精

度を得られたことの説明にはなっていない. これについては, 抽出したデータのスパース性が NMF に対して効果的に作用したと考えることができる. 例えば NMF は画像認識等でも用いられるが, 人間の顔写真を NMF によって分解すると, その基底には目や鼻, 口といった顔のパーツが抽出される. これらはそれぞれが独立ではないが, 完全に依存した関係にはないことと, NMF の非負値による加法性がマッチした結果であると考えられる. スタイルパラメータもまた同様にそれぞれの特徴量は独立ではないが完全な依存関係にはないことから, 演奏の状態をうまく表すことができた可能性があるが, これらはまだ仮説段階であり, その理由については今後も考察を続ける必要がある.

4.3 楽器間の共起確率

各楽器演奏者は互いにインタラクションを行い, またそれぞれの演奏を阻害しないように適した演奏を返す. すなわち, 各楽器の演奏間には共起関係が存在するはずであり, また我々のセッションモデルにおいてもそれは重要な要素である. TABLE1-4 は楽器間のクラスラベルについて共起関係を条件付き確率で示したものである. TABLE1-2 は NMF における各楽器による演奏のクラス番号について求めた表であり, TABLE3-4 は HMM で計算されたパラメータを用いて, Viterbi アルゴリズムによって隠れ状態間の共起確率を求めた表である. これらの表を見ると明らかに楽器間のクラスに共起しやすいものやそうでないものがあることがわかる. 従って, 演奏状態をより正確に追跡するためには, 楽器間の共起関係をより明らかにする必要がある, またそれは演奏の予測に有用であることが示唆された.

(iii) おわりに

本稿ではできるだけヒューリスティックを排除し, 実際の演奏データのみから統計的に学習可能なジャムセッションシステムの数理モデルを提案し, 特にクラスターリング手法を焦点とした比較検討を行った. 演奏特徴量としてスタイルパラメータを定義して抽出し, k -means, GMM, NMF の 3 つの手法について, trigram と HMM による演奏予測精度による比較を行った結果, NMF によって特に高精度な予測結果が得られた. また, 楽器間に共起関係があることを実際に条件付き確率によって具体的に数値で示し, その有効性を示唆した.

また, 本研究によって提案した数理モデルは, ジャムセッションシステムのみでなく, 相互作用を持つ複数の事象による状態遷移モデル(例えば, 社会や自然現象と人間の活動といったダイナミクスなモデル)としてもフレキシブルに適用できると期待している.

TABLE 1. – 条件付き確率
 (Piano-Bass / 小節単位, クラス数 6)

Cluster number (Piano / bar)	Cluster number (Bass / bar)					
	0	1	2	3	4	5
0	0	0	0	0	0	1
1	0.023	0.048	0.014	0.114	0.786	0.016
2	0.062	0	0.021	0.229	0.292	0.396
3	0.007	0.013	0.009	0.092	0.866	0.012
4	0.031	0.041	0.004	0.156	0.757	0.011
5	0.102	0.02	0	0.061	0.102	0.714

TABLE 2. –条件付き確率
 (Piano-Drums /小節単位, クラス数 6)

Cluster number (Piano / bar)	Cluster number (Drums / bar)					
	0	1	2	3	4	5
0	0	0	1	0	0	0
1	0.134	0.027	0.027	0.002	0.775	0.034
2	0.062	0.375	0.208	0	0.312	0.042
3	0.085	0.066	0.015	0.001	0.762	0.072
4	0.155	0.099	0.025	0.004	0.606	0.112
5	0	0.184	0.02	0	0.224	0.571

今後は HMM を内包した Dynamic Bayesian Network (DBN) による共起関係を反映した演奏予測モデルによる予測精度の向上を目指す。また、本研究では音楽事象空間 Ω 及びトラジェクトリ T を離散化することによって学習可能なモデルとしたが、Deep Neural Network (DNN) 等を用いた補間によるモデルの適用も検討する。

参考文献

[1] R. B. Dannenberg, “An On-line Algorithm for Real-time Accompaniment,” Proc. ICMC, 193-198, 1984.
 [2] 中村栄太, 武田晴登, 山本龍一, 斎藤康之, 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, No4, pp.1338-1349, 2013.
 [3] R. Rowe, “Machine Listening and Composing with Cypher,” Computer Music Journal, 16(1), 43-63, 1992.

TABLE 3. – 条件付き確率
 (Piano-Bass / 小節単位, HMM, 状態数 6)

State number (Piano / bar)	State number (Bass / bar)					
	0	1	2	3	4	5
0	0.02	0.367	0	0.143	0.041	0.429
1	0.143	0.014	0	0.146	0	0.708
2	0.021	0.125	0	0.146	0	0.708
3	0.183	0.033	0.05	0.033	0.655	0.05
4	0.146	0.166	0.099	0.13	0.439	0.019
5	0.127	0.137	0.095	0.064	0.565	0.012

TABLE 4. – 条件付き確率
 (Piano-Drums / 小節単位, HMM, 状態数 6)

State number (Piano / bar)	State number (Drums / bar)					
	0	1	2	3	4	5
0	0	0.041	0.204	0.245	0	0.51
1	0.033	0.048	0.054	0.057	0.73	0.078
2	0.042	0.604	0	0	0.104	0.25
3	0.067	0.117	0.05	0.033	0.617	0.117
4	0.011	0.157	0.122	0.127	0.296	0.287
5	0.005	0.049	0.073	0.076	0.699	0.098

[4] 青野裕司, 片寄晴弘, 井口征士: バンドライクな音楽アシスタントシステムについて, 情報処理学会研究報告音楽情報科学, Vol.94, No.103, pp.45-50, 1994.
 [5] Y. Aono, H. Katayose, and S. Inokuchi, “An Improvisational Accompaniment System Observe Performer’s Musical Gesture,” Proc. ICMC, 106-107, 1995.
 [6] N. Nishijima and K. Watanabe, “Interactive Music Composer on Neural Networks,” Proc. ICMC, 53-56, 1992.
 [7] 和気早苗, 加藤博一, 才脇直樹, 井口征士: テンション・パラメータを用いた協調型自動演奏システム: JASPER, 情報処理学会論文誌, Vol.35, No.7, pp.1469-1481, 1994.
 [8] M. Goto, I. Hidaka, H. Matsumoto, Y. Kuroda and Y. Muraoka, “A Jazz Session System for Interplay among All Players – VirJa Session (Virtual Jazz Session System),” Proc. ICMC, 346-349, 1996.
 [9] 浜中雅俊, 後藤真孝, 麻生英樹, 大津展之: Guitarist Simulator: 演奏者の振舞いを統計的に学習するジャムセッションシステム, 情報処理学会論文誌, Vol.45, No.3, pp.698-709, 2004.
 [10] D. D. Lee, H. S. Seung, “Advances in Neural Information Processing Systems,” 13, 556-562, 2000.
 [11] J. Kim and H. Park, “Sparse nonnegative matrix factorization for clustering,” Technical Report GT-CSE-08-01, Georgia Institute of Technology, 2008.
 [12] Yamaha, “Yamaha Music Datashop,” available from <https://yamhamusicdata.jp/>, (accessed 2015-10-14).