

## ベイジアンネットワークを用いた 自動コードヴォイシングシステム

北原 鉄朗<sup>†1,†2</sup> 勝占 真規子<sup>†1,\*1</sup>  
片寄 晴弘<sup>†1,†2</sup> 長田 典子<sup>†1,†2</sup>

本論文では、ベイジアンネットワークを用いたコードネームからの自動ヴォイシングシステムについて述べる。ヴォイシングは音楽的同時性や音楽的連続性を考慮しながらテンションや転回形を決定する必要があり、自動的に決定するのは容易ではない。この問題を解決するため、メロディやヴォイシング進行を考慮した事例学習型のコード・ヴォイシングモデルを構築する。このモデルでは、音楽的同時性と連続性を、「現在のコード」のヴォイシングを表すノードとメロディや前後のコードのヴォイシングを表すノードとの確率的な依存関係として表現する。このモデルにおいて最ももらしいヴォイシングを確率的推論によって導くことで、音楽的同時性と連続性を両方満たすヴォイシングを得ることができる。実際にジャズの楽譜から学習したヴォイシング推定モデルによって実験したところ、音楽的同時性と連続性を両方満たすヴォイシングが出力されることを確認した。

### Automatic Chord Voicing System Using Bayesian Network

TETSURO KITAHARA,<sup>†1,†2</sup> MAKIKO KATSURA,<sup>†1,\*1</sup>  
HARUHIRO KATAYOSE<sup>†1,†2</sup> and NORIKO NAGATA<sup>†1,†2</sup>

This paper describes an automatic chord voicing system using the Bayesian network. Automatic chord voicing is not easy because it is necessary to decide tension notes and inversions by taking into account both musical simultaneity and sequentiality. To solve this problem, we construct a chord voicing model based on the Bayesian network. This model represents musical simultaneity as probabilistic dependencies between voicing and melody nodes and sequentiality as probabilistic dependencies between current-chord and previous- or following-chord voicing nodes. This modeling makes it possible to infer the most likely voicings that have both simultaneity and sequentiality. Experimental results of chord voicing for jazz musical pieces showed that our system generated chord voicings that have appropriate simultaneity and sequentiality.

### 1. はじめに

音楽情報処理の課題の1つに自動編曲があり、その部分問題として和声付けがある。和声は音楽の3要素の1つであり、楽曲の持つ印象や雰囲気を決めるうえで重要である。そのため、これまで数多くの研究がなされてきた。川上らは、和音名を隠れ状態とする隠れマルコフモデルを用いて、観測された旋律の背後にある和声を推定する手法を提案した<sup>1)</sup>。平田らは、演繹オブジェクト指向データベースに基づく音楽知識表現体系に基づいてコード、ケーデンス、ヴォイシングを表現し、この枠組みの下でジャズ和声を探査・推論・生成するインタラクティブシステム「パーピープン」を実現した<sup>2)</sup>。同様の枠組みに基づいて、与えられたコード進行をジャズらしいコードにリハーモナイズするシステムを実現している<sup>3)</sup>。三浦らは、音楽理論の1つである和声法を計算機上に実装し、和声付けシステムAMORを開発した<sup>4)</sup>。江村らも、音楽理論に基づいたルールベースの自動編曲システムを実装した<sup>5)</sup>。そのほかにも、Pachet<sup>6)</sup>、Chemiller<sup>7)</sup>も和声に関する研究を行っている。

自動編曲において和声付けと相補的な関係にあるものとして、ヴォイシングがあげられる。ヴォイシングとは、コードネームで記されたコード進行に対して、具体的にどの音を演奏するかを決定するプロセスである。クラシック音楽では、通常ヴォイシングは作曲時に決定されているが、ポップスやジャズでは、メロディ譜(メロディとコードネームのみが書かれたシンプルな譜面)のみが与えられることが少なくない。その場合、演奏者がその場でヴォイシングをしながら演奏することが要求される。特に、ジャズの場合、メロディ譜にはテンションノートが書かれていないことが多く、適切なテンションノートを考えながら演奏するのは、楽器演奏初心者には難しい。そのため、メロディ譜の入力に対して、テンションノートの付与も含めてヴォイシングを自動で行ってくれるシステムがあれば、初心者でも演奏を気軽に楽しめると思われる。

しかしながら、ヴォイシングの研究事例は和声付けに比べて少ない。上であげた中でも、平田らのパーピープン<sup>2)</sup>/ハービー君<sup>3)</sup>と江村らのシステム<sup>5)</sup>だけであり、そのほかではヴォ

†1 関西学院大学理工学研究科

Graduate School of Science and Technology, Kwansai Gakuin University

†2 科学技術振興機構戦略的創造研究推進事業 CrestMuse プロジェクト

CrestMuse Project, CREST, JST

\*1 現在、大日本印刷株式会社

Presently with Dai Nippon Printing Co., Ltd.

イシングを扱った研究はほとんど見られない\*1。

ヴォイシングが難しいのは、テンションノートの付与、転回形の選択、構成音の省略といった問題を、これらの相互依存性を考慮に入れながら行わなければならないからである。ヴォイシングを行うにあたっては、メロディなどの他のパートとの兼ね合い（音楽的同時性）や前後の和音とのつながりの良さ（音楽的連続性）を十分に考慮する必要がある。ここであげた3つの問題は、いずれも同時性と連続性の双方に影響するため、単純なルールでは効果的なヴォイシングは困難である。

この問題に対する1つの解決方法は、音楽理論を活用することである。特にジャズでは音楽理論の体系化が進んでおり、ヴォイシングに対しても有用な技法を提供する。実際、江村らのヴォイシングシステム<sup>5)</sup>は、ジャズのための音楽理論を膨大なルールの集合として実装し、ジャズらしいヴォイシングを行うシステムを実現した。しかし、一般に音楽理論は計算機に実装するには抽象度が高く、音楽理論で記述されている各技法をルール化した場合、各ルールが競合することも少なくない。その場合でもルール間の優先順位は明確でなく、どのルールをどのように組み合わせれば全体として最も音楽的同時性や音楽的連続性を考慮できるのか、判断するのは難しい。

本研究では、ヴォイシングにおける音楽的同時性や音楽的連続性を確率的依存関係として表現し、最ももっともらしいヴォイシングを確率的推論によって導き出すというアプローチをとる。音楽的同時性は、メロディに関する事象とヴォイシングに関する事象の確率的依存性によって、音楽的連続性は、現在のコードのヴォイシングに関する事象と直前あるいは直後のコードのヴォイシングに関する事象の確率的依存性によって表現する。このような複数の確率事象に対する複数の依存関係を記述し、推論するモデルとして、本研究ではベイジアンネットワークを用いる。この2種類の依存関係をベイジアンネットワークとしてモデル化し、全体として最も尤度が高くなるヴォイシングを推論することで、音楽的同時性と音楽的連続性を同時に考慮したヴォイシングを求めることができる。

以下、2章ではヴォイシングにおける課題とアプローチについて議論する。3章ではヴォイシングを確率的に定式化し、ベイジアンネットワークに基づいたヴォイシングモデルの設計について述べる。4章では評価実験について報告し、最後に5章でまとめをする。

\*1 市販ソフトウェア「Band-in-a-Box 16」ではコード進行を入力してスタイル（音楽ジャンルに相当）を選ぶと自動的に編曲を行うことができる。ただし、我々が試用したところ、ヴォイシングに関しては後述の音楽的同時性や音楽的連続性を十分に考慮したものになってはいない。

## 2. ヴォイシングの課題とアプローチ

### 2.1 ヴォイシングにおける課題

コードネームからヴォイシングを決定する課題において困難とされるのは、コードネームに対しヴォイシングが1つに決まらないことである。たとえば、コードネーム“CMajor”が与えられた場合、基本構成音であるド、ミ、ソは容易に決まるが、テンション・省略音・転回形は複数の解候補がある。すなわち、以下のような課題があげられる。

- テンションノートの付与  
効果的なテンションノートを付与するためには、同時に発音するメロディを考慮しなければならない。コードネームに付与可能なテンションノートは、ある程度ルールで決められているが、その中からどれを適用するかは決定は難しい。メロディ音との不協和な半音衝突は、不適切なハーモニーとなるので避ける必要がある。
- 省略音の選択  
単にテンションノートを付与しただけでは、メロディや元の構成音と濁った響きを生む可能性がある。また、指の本数や大きさといった物理的な制約も存在する。こうした制限の下でできるだけ効果的な響きを作るため、構成音のいくつかを省略する必要性が出てくる。
- 転回形の決定  
音楽的文脈を考慮して転回形を決定するには、前後のヴォイシング進行を考える必要がある。転回形が利用される目的として、メロディを効果的に補う手法である「カウンターメロディ」がある。転回形をうまく利用し、伴奏の中にメロディの副旋律的要素を含ませることで音楽表現を豊かにすることが可能である。故意に低い音や高い音に移行して音楽に変化をつける手法もある。また、物理的にスムーズに手を動かし滑らかな遷移を保つため、なるべく近くの音に移動することも考えられる。このように、転回形は時系列の依存関係が強く、前後のヴォイシング遷移がどのようになっているのかを考慮しながら決定する必要がある。

Conklinの研究<sup>8)</sup>では、楽曲を同時性（simultaneity）と連続性（sequentiality）という2つの側面からとらえてパターン抽出を行っている。ヴォイシングにおいても同様に、音楽的同時性と音楽的連続性が重要である。ヴォイシング課題における音楽的同時性とは、同時に発音されるヴォイシングの響き、音楽的連続性とは、音楽的流れを感じさせる和音間のつながりを意味する。基本的に、テンションノートの付与や構成音の省略は前者に属し、転回

形の決定は後者に属する。しかし、実際に、テンションノートの付与を行うには音楽的同時性だけ考慮すればいいわけではない。テンションリゾルブという手法があるが、これはテンションによる音楽的文脈を考慮した緊張から解決へ導く手法であり、流れを考慮しながら決定しなければならない。すなわち、別々に決定するのではなく、互いに考慮しあいながら決定することが望ましく、それを行える枠組みが必要となる。

## 2.2 ヴォイスニングのアプローチ：ルールベースと事例ベース

ヴォイスニングを含む音楽生成のアプローチは、ルールベースと事例ベースに大別することができる<sup>2)</sup>。

前者は文字どおり音楽に関する知識や概念をルールの形で記述し、そのルールに基づいて音楽を生成する、一種のエキスパートシステムである。たとえば、テンションノートの付与、省略音の選択、転回形の決定のそれぞれに対して音楽的同時性を実現するためのルール群、音楽的連続性を実現するためのルール群、それらが競合したときのためのメタルール群を設計し、これらを順次適用することでヴォイスニングを行う手法が考えられる。音楽理論や和声学などの形で音楽に関する再利用可能な知識が膨大にあるため、これらを利用できることは大きなメリットである。しかし、前章で述べた、音楽理論のあいまいさやルール間の競合といった問題に加え、推論制御の困難さも指摘されている<sup>2)</sup>。とはいえ、音楽の専門家がルール群を十分に練って設計すれば品質の高い出力を得られることも事実で、江村らはジャズ理論に基づいたルールベースヴォイスニングシステムを実現している<sup>5)</sup>。ヴォイスニングのほかにも、和声付けや和声解析などで多くのルールベースシステムが提案されている<sup>9),10)</sup>。

一方、後者は既存の楽曲を利用して音楽の生成を行うアプローチ<sup>2),11),12)</sup>である。このアプローチは、ルールベースの次の問題点を解決することができる。

- ルールベースでは、網羅的なルール群を注意深く設計する必要があったのに対し、事例ベースでは音楽の知識や概念をすべて網羅的に記述しておく必要がない。
- 特定のジャンルの特徴が反映されたヴォイスニングを行いたい場合、ルールベースではジャンルごとにルールを設計する必要があるが、事例ベースではそのジャンルの事例曲集を集めるだけでよい。

事例曲の利用方法には、事例曲を確率モデルの推定に用いる方法（確率ベース）と事例曲を統計処理をせずに直接用いる方法（狭義の事例ベース）が考えられる。ハービー君<sup>3)</sup>やパーピーブン<sup>2)</sup>は後者の方法をとったシステムである。ハービー君では、入力されたコード進行に対して事例集から最長一致したコード進行を検索し、それに付与されていたヴォイスニングを入力コード進行の当該箇所付与する。この場合、最長一致したコード進行区間内では

音楽的連続性は満たされると予想されるが、別々の事例（楽曲断片）からヴォイスニングが付与された隣接するコード進行区間の間では、音楽的連続性は考慮されない<sup>\*1</sup>。音楽的同時性については、事例検索時にメロディの類似性を条件に入れることで原理的に対応可能であるが、ハービー君では扱われていない。パーピーブンでは、上述のうち1つめの問題点を解決するため、ケーデンスの階層的な連なりであるケーデンス木を定義し、この部分木に対してヴォイスニングを行い、ヴォイスリーディングの円滑さを表す指標も導入している。ただし、ケーデンス木を自動的に推定するのは難しいため、ユーザが入力する必要がある。一方、確率ベースのヴォイスニングには次のメリットがあると考えられる。

- 狭義の事例ベースでは異なる事例が次々と用いられることで一貫性のないヴォイスニングが出力される可能性がある一方、複数の事例が確率モデルという形で抽象化されることで汎化能力が生まれ、一貫性のあるヴォイスニングを得やすくなると考えられる、その際、汎化能力はモデルの表現能力（複雑さ）とトレードオフの関係にあり、設計によってそのバランスを図ることができる。

- ベイジアンネットワークなど、事象間の複雑な依存関係を記述し、推論するための技術が多数提案されており、それらを適用できる。

その一方で、次のようなデメリットもある。

- 大量の学習データが必要である。特に、事象間の複雑な依存関係に対応するにはモデルが複雑になり、限られたデータから精度良く学習するのは困難である。
- あるアーティストの個性を模倣しようとした場合、個性的な演奏は頻度としては少ない場合が多いため、確率分布に表れにくい。

以上の議論に鑑み、本研究では確率モデルに基づいた自動ヴォイスニング手法について検討する。特に、次の点に留意して検討を進める。

- 明示的なルールを記述せずにどこまで音楽的に妥当なヴォイスニングが可能かどうかを明らかにするため、モデルの設計段階における最低限のもの以外は、音楽知識は使わずにボトムアップな方法のみでヴォイスニングを行う。
- 特定のジャンルの事例集（学習データ）のみを用いることで、そのジャンルの特徴を反映させられることを確かめるため、特定のジャンルの事例集を用いる。
- 学習データが有限であることを考慮したモデルの設計を行う。

\*1 実際、平田らも「ケーデンス間のヴォイスリーディングに不自然な部分が残っている」と指摘している<sup>2)</sup>。

### 3. ベイジアンネットワークによるモデル化

以上の方針に従い、本章では、ベイジアンネットワークに基づく音楽的同時性と音楽的連続性の2点を考慮したヴォイシングモデルを提案する。

#### 3.1 問題設定

本研究では、メロディ譜が与えられたときに、ヴォイシングが付与された楽譜を生成することを目的とする。ジャンルはジャズ音楽、楽器は電子オルガンを対象とし、メロディを右手で、コードは左手で、ベースは左足で弾くことを前提とする。入力にはメロディとコードネーム列、出力は全コードに対応するヴォイシングとする。ここでヴォイシングとは、左手によるコード演奏のヴォイシングとベース音の選択の両方を指す。なお、コード間に挿入される経過和音、ベースラインにおける経過音は扱わないものとする。

#### 3.2 ヴォイシング問題の確率論的定式化

本研究では、メロディラインとコードネーム列が入力され、各コードに対して左手でおさえるべきヴォイシングとベース音を付与する問題を解く。いま、入力されたコードネーム列における  $i$  番目のコードネームを  $C_i$ 、そのコードが演奏される区間におけるメロディを  $M_i$ 、そのコードの左手ヴォイシングを  $V_i$ 、ベース音を  $B_i$  とすると、事後確率を最大にするヴォイシングは、

$$(\hat{V}_i, \hat{B}_i) = \operatorname{argmax}_{V_i, B_i} P(V_i, B_i | C_i, M_i)$$

として与えられる。ただし、この式では音楽的同時性は考慮されているが、音楽的連続性は考慮されていない。 $i$  番目のコードに対するヴォイシング  $(V_i, B_i)$  が直前のコードに対するヴォイシング  $(V_{i-1}, B_{i-1})$  に依存して決まるとすると、

$$(\hat{V}_i, \hat{B}_i) = \operatorname{argmax}_{V_i, B_i} P(V_i, B_i | C_i, M_i, V_{i-1}, B_{i-1})$$

となる。同様に、 $i+1$  番目のコードに対するヴォイシング  $(V_{i+1}, B_{i+1})$  の条件付き確率は

$$P(V_{i+1}, B_{i+1} | C_{i+1}, M_{i+1}, V_i, B_i)$$

と表される。本来、直前のコードだけでなく、より長い文脈に依存すると思われるが、ここでは簡単のため、直前のコードだけを取り上げる。一般に、ヴォイシングでは、単に直前のコードとのつながりを考慮するだけでなく、その先のコードと適切につながりかどうかを考慮して行っていると考えられる。そこで、1つ先のコードのヴォイシングも同時に推論し、両者の同時確率が高くなるようにヴォイシングを決定する。つまり、

$$\begin{aligned} & (\hat{V}_i, \hat{B}_i, \hat{V}_{i+1}, \hat{B}_{i+1}) \\ &= \operatorname{argmax}_{V_i, B_i, V_{i+1}, B_{i+1}} P(V_i, B_i | C_i, M_i, V_{i-1}, B_{i-1}) P(V_{i+1}, B_{i+1} | C_{i+1}, M_{i+1}, V_i, B_i) \end{aligned}$$

である。

#### 3.3 各確率変数の設計

次に、上式の各確率変数を設計する。本研究では、上式に基づいたヴォイシングモデルをベイジアンネットワークとして構築するが、ベイジアンネットワークの典型的な利用方法では、各確率変数を離散変数とし、各ノード(確率変数)に対して条件付確率表を作成する。そこで、本研究でも確率変数を離散変数として設計する。

コードネームに関する確率変数  $C_i$  は単に対象とするコードネームを値としてとるものとする。コードネームはルート音とコードの種類の組合せで表され、本研究では、3和音として major, minor, augmented, diminished, suspended 4th の5種類<sup>\*1</sup>、4和音として dominant 7th, major 7th, minor 7th, minor major 7th, minor 7th (b5)/half diminished, major 6th, minor 6th の7種類<sup>\*2</sup>を用いる。

左手ヴォイシングに関する確率変数  $V_i$  は、最高音、最低音、それ以外の3つの確率変数 ( $V_i^{\text{top}}$ ,  $V_i^{\text{btm}}$ ,  $V_i^{\text{mid}}$ ) に分解して考える。これは、ヴォイシングの音楽的連続性において、最高音と最低音の連続性が特に重要だからである。 $V_i^{\text{top}}$  と  $V_i^{\text{btm}}$  は音名を、 $V_i^{\text{mid}}$  は音名の組合せを値にとる。たとえばコード  $C_i$  のヴォイシングが “C-E-G-B” であれば、 $V_i^{\text{top}}$  = “B”,  $V_i^{\text{btm}}$  = “C”,  $V_i^{\text{mid}}$  = “EG” となる。 $V_i^{\text{mid}}$  が複数の音名の組合せを記述できるため、任意の音符数におけるヴォイシングに対応できる。ただし、とりうる値を有限個に抑えるため、学習データに出てきた音名の組合せのみをとりうる値とする。たとえば、電子オルガンの左手で演奏する場合には、原理的に  $V_i^{\text{mid}}$  は4つ以上の音が記述されることはないため、こういった場合は考慮に入れない(現実にはほとんどが2音以下である)。なお、少ないデータからでも学習できるように、オクターブは考慮しないものとする。

ベースに関する確率変数  $B_i$  は、本研究ではコード1つに対してベース音を1つだけ選び、ベースラインにおける経過音を考えないので、単純に音名を値にとることにする。左手ヴォイシングと同様に、オクターブは考えないものとする。

メロディに関する確率変数  $M_i$  については、メロディの音符数が一定ではないので、音符

\*1 ルート音がCであれば、左から C, Cm, Caug, Cdim, Csus4.

\*2 ルート音がCであれば、左から C7, CM7, Cm7, CmM7, Cm7(b5), C6, Cm6.

そのものを確率変数にすることはできない．そこで，当該コード区間において各音名が鳴っている時間の割合を表す 12 次元ベクトルを採用する．各次元において値が 1.0 に近ければ，そのコード区間において当該音名が長時間鳴っており，その音名と不協和な響きを生まないように考慮しなければならない．一方，この値が低ければ，その音名はメロディ中で使われていないか経過音として使われていると考えられ，不協和な響きを生む可能性のある音をヴォイシングにおいて避ける必要性は低くなる．なお，各次元の値は 0.0~1.0 (0.2 刻み) の 6 段階に離散化して表す．

### 3.4 データスパースネス問題の解決のためのモデルの簡略化

ペイジアンネットワークでは，原理的には任意の複雑さのモデルを構築することができるが，モデルが複雑になるほど，学習に大量のデータが必要になり，データスパースネス問題が起きやすくなる．ペイジアンネットワークの典型的な利用方法では，上述のとおり，離散化された確率変数間の確率的依存関係を条件付き確率表として表し，この表に基づいて確率の計算を行う．そのため，データスパースネス問題の回避においては，条件付き確率表のサイズを削減することが肝要である．

条件付き確率表のサイズは，当該条件付き確率に現れるすべての確率変数のとりうる値の個数の積で与えられる．そこで，以下では確率変数間の独立を仮定するなどして，条件付き確率における確率変数の個数を減らしていく．

まず，左手のヴォイシングとベース音の選択は，本来独立ではないが，独立と見なして別個のモデルを構築する．また，左手ヴォイシングにおける 3 つの確率変数  $V_i^{top}$ ,  $V_i^{btm}$ ,  $V_i^{mid}$  も独立であると見なす．つまり，

$$(\hat{V}_i^h, \hat{V}_{i+1}^h) = \operatorname{argmax}_{V_i^h, V_{i+1}^h} P(V_i^h | C_i, M_i, V_{i-1}^h) P(V_{i+1}^h | C_{i+1}, M_{i+1}, V_i^h),$$

$$(\hat{B}_i, \hat{B}_{i+1}) = \operatorname{argmax}_{B_i, B_{i+1}} P(B_i | C_i, M_i, B_{i-1}) P(B_{i+1} | C_{i+1}, M_{i+1}, B_i)$$

である ( $h = top, btm, mid$ ).

次に，ヴォイシングとメロディの依存関係を逆転させる．つまり，メロディに合うヴォイシングを考えるのではなく，メロディの背後には先にヴォイシングを含めたコード進行があり，それに依存する形でメロディが生成されたと考える\*1．具体的な定式化は，

\*1 このようにメロディがその背後にある和声から生成されたと考えるアプローチは，確率モデルを用いた和声付け<sup>1)</sup>や和声解析<sup>13)</sup>などでは一般的である．

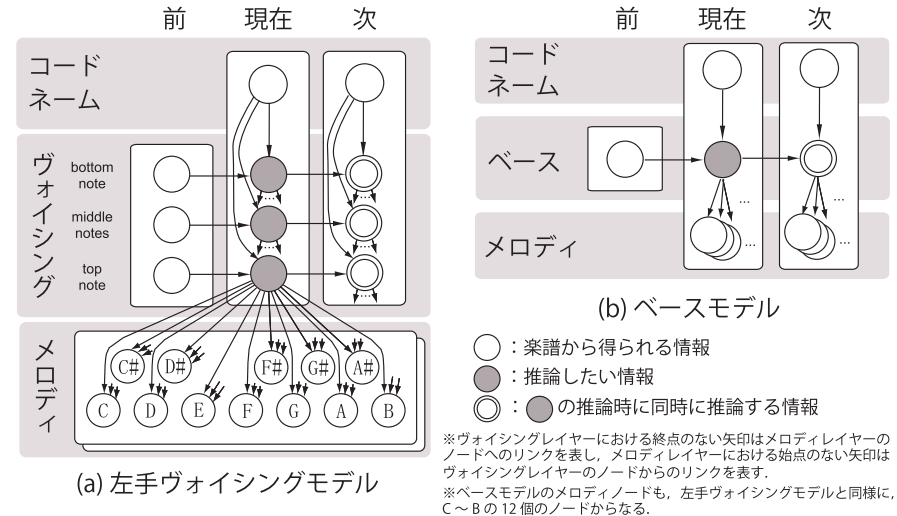


図 1 本研究で設計したヴォイシングモデル  
Fig. 1 Voicing model designed in this paper.

$$(\hat{V}_i^h, \hat{V}_{i+1}^h) = \operatorname{argmax}_{V_i^h, V_{i+1}^h} P(M_i | V_i^h) P(V_i^h | C_i, V_{i-1}^h) P(M_{i+1} | V_{i+1}^h) P(V_{i+1}^h | C_{i+1}, V_i^h),$$

$$(\hat{B}_i, \hat{B}_{i+1}) = \operatorname{argmax}_{B_i, B_{i+1}} P(M_i | B_i) P(B_i | C_i, B_{i-1}) P(M_{i+1} | B_{i+1}) P(B_{i+1} | C_{i+1}, B_i)$$

となる．このようにすることで，音楽的同時性と音楽的連続性を別々の条件付き確率で表すことができるので，個々の条件付き確率における確率変数の個数を抑えることができ，条件付き確率表のサイズが削減される．

### 3.5 ヴォイシングモデルの設計

以上の定式化に従って，ヴォイシングモデルを設計する．左手による演奏のヴォイシングのためのモデルを「左手ヴォイシングモデル」，ベース音の選択に用いられるモデルを「ベースモデル」と呼ぶ．

我々が設計した左手ヴォイシングモデルとベースモデルを図 1 に示す．各モデルは，メロディを表すノードからなるレイヤ，コードネーム列を表すノードからなるレイヤ，左手ヴォイシングまたはベース音を表すノードからなるレイヤの 3 レイヤ構成となる．これらのレイヤの間でノードがリンクによって結ばれることで音楽的同時性を表現する．また，各レイ

表 1 図 1 のモデルにおけるノード名とその状態

Table 1 Nodes and their states included in models in Fig. 1.

ノード名	状態	
コードネーム	ルート音 × コードの種類*	
ヴォイシング	top note	音名 (C~B)
	middle notes	音名 (C~B) の組合せ
	bottom note	音名 (C~B)
ベース	音名 (C~B)	
メロディ	各音名の相対的長さ (0 と 1 を 0.2 刻みにした 6 段階)	

\*コードの種類: major, minor, augmented, diminished, suspended 4th, dominant 7th, major 7th, minor 7th, minor major 7th, minor 7th (b5)/half diminished, major 6th, minor 6th の 12 種類.

は、現在のコード、直前のコード、次のコードに対応するノードに分かれる(ただし、推論に必要なノードは用意されない)。各ノードは 3.3 節で述べた確率変数に 1 対 1 で対応している。各ノードがとりうる値を表 1 にまとめる。

このモデルは、「現在のコード」を曲の進行に合わせてシフトさせながらヴォイシングを進めることを想定している。まず、楽曲の先頭のコードを「現在のコード」とする。直前のコードは存在しないため不定とする。メロディとコードネームのレイヤに、入力されたメロディ譜から得た値(エビデンスという)を代入し、推論を実行する。具体的には、エビデンスが設定されていないノードの確率分布の推定(離散変数なのでとりうるすべての値に対する確率値の計算)を、Pearl の確率伝播法<sup>14)</sup>に基づいて行う。「現在のコード」に対するヴォイシングだけでなく「次のコード」に対するヴォイシングのノードにもエビデンスが設定されていないため、これらも同時に推論していることになる。推論後、最も確率が高いヴォイシングを求め、「現在のコード」に対するヴォイシングとして確定する。「現在のコード」を 1 つ進め、以上と同じ処理を行ってヴォイシングを求める。ただし、「直前のコード」に対するヴォイシングには 1 つ前のステップで確定したヴォイシングを代入する。「直前のコード」に対するヴォイシングは確定事項であり、推論対象ではないため、直前のコードに対応するメロディレイヤやコードネームレイヤのノードは用意されない。

このモデルは、前述のように少量のデータからでも推定できるように簡略化されているため、音楽的同時性と音楽的連続性のすべてをモデル化しているわけではない。このモデルでは扱われていない音楽的同時性と音楽的連続性について以下にまとめる。

- 音楽的同時性

ヴォイシングの個々の音(左手最高音, 左手最低音, 左手中間音, ベース音)とメロ

ディの間の音楽的同時性はモデル化されているが、ヴォイシング内の同時性はモデル化されていない。

- 音楽的連続性

ヴォイシングの個々の音における音楽的連続性についてはモデル化されている(たとえば最高音どうしや最低音どうしのつながり方は考慮される)が、たとえば直前のコードの最高音と現在のコードの中間音の間の連続性については考慮されない。また、メロディも本来音楽的連続性が存在するが、そちらもモデル化されていない。ただし、本タスクではメロディは既知で推論対象ではないので、メロディ中の音楽的連続性の考慮の有無は推論には影響しない。

#### 4. システムの実装と評価

以上で述べたモデルに基づきコードヴォイシングを行うシステムを実装し、評価実験を行った。評価実験として、音楽の専門家によるコードごとの評価および音楽の非専門家による楽曲全体の印象評価を行った。そのほか、モデルを一部変えたときのヴォイシングの変化についても考察する。

##### 4.1 システムの実装

コードヴォイシングシステムを、図 1 に示したモデルを用いて構築した。モデル構築には、ベイジアンネットワーク構築システム BayoNet<sup>15)</sup> を利用し、学習データには、ジャズまたはジャズ風に編曲された電子オルガンの楽曲 30 曲(すべて長調)を用いた。入力データには、コードネームが 7th までで表記されているメロディ譜を与えた。出力は、各コードネームに対するヴォイシングとした。

ベイジアンネットワークの推論結果はあくまで各ノードに対する各音名の尤度であるので、それをもとに各コードのヴォイシングを確定させる必要がある。ここでは、演奏可能音域を設定し、その範囲での音高を割り当てた。左手ヴォイシングの演奏可能領域を C3~A#4、ベースは C2~G3 とした。最初のヴォイシング決定はあらかじめ設定しておいた基準音(左手ヴォイシング: E3, ベース: D2) から最も音高差の少ない音を最低音とし、他の音は、最低音が決まることで自動的に決定される。以降のヴォイシングも 1 つ前の bottom note を基準音とし、同じ手順で音高を決定した。演奏可能音域を超える場合は 1 オクターブ推移させ極力その範囲内に収めた。左手ヴォイシングの 3 要素の中で、同じ音が最尤であると推定された場合は、最低音, 最高音, 中間音の順で決定した。同じ音名の選択は禁止し、次にもっともらしいと判断した音を選択した。ベース音の決定において構成音以外の音

表 2 音楽専門家による音楽的同時性の評価結果  
Table 2 Evaluation of musical simultaneity.

評価	コード数	割合
美しい	82	9.0%
良い	719	78.8%
許容できる	63	6.9%
許容できない	48	5.3%

表 3 音楽専門家による音楽的連続性の評価結果  
Table 3 Evaluation of musical sequentiality.

評価	コードチェンジ数	割合
良い	180	25.8%
普通	407	58.3%
悪い	111	15.9%

が推測された場合には、次にもっともらしいと推測された音を結果とした。

#### 4.2 出力結果に対する音楽専門家の評価

本システムを用いた推定結果を音楽的同時性と音楽的連続性の2つの観点から、音楽専門家による評価実験を行った。評価者は、メロディ譜を見ながらそれに対するヴォイスニング結果を聴く。音楽的同時性を評価するため、1コードネームに対して、美しい、良い、許容できる、許容できないの4段階で評価する。コードネームからの音判断やイメージを避けるため、ヴォイスニングが存在する位置のみが分かるようにした。楽曲は全部で5曲(As Time Goes By, I Left My Heart In San Francisco, My One And Only Love, Misty, Satin Doll), 1人あたり304コードを評価した。評価者は、ジャズを専門とする音楽専門家3人である。

その結果を表2に示す。全体の94.7%のヴォイスニングが許容できる以上の範囲を占めた。この結果から、本システムではある程度音楽的同時性の感じられる妥当なヴォイスニングが推定され、適切にヴォイスニングできると分かった。しかし、約5%が許容できないと評価された。原因の1つとして、左手のヴォイスニングとベース音を別々に決定したため、不整合が生じたことが考えられる。

次に、コードチェンジ時のヴォイスニングの音楽的連続性を、良い、普通、悪いの3段階で音楽専門家2人が評価した。楽曲は、同様に5曲、1人あたり299コードチェンジである。その結果(表3)から、普通以上の評価を得たのは全体の84.1%であり、ほぼ妥当なヴォイスニングであるといえる。しかし、15.9%が悪いと評価された。この原因として、演奏可能音域を制限していることがあげられる。ヴォイスニングモデルでは、半音または全音で移動可能なヴォイスニングの尤度が高かったとしても、演奏可能音域を超えた場合にオクターブ移動する可能性があり、これによって不必要な音高の飛躍が発生したと考えられる。

#### 4.3 3種類のヴォイスニングに対する印象評価

ジャズの事例曲集を用いて出力したヴォイスニングにジャズの特徴が反映されていることを

確かめるため、以下の3種類のヴォイスニングに対して印象評価を行った。

- 本システムを用いたヴォイスニング(以下、システムヴォイスニング)  
メロディ譜のみを入力とし、ヴォイスニングの出力をシステムを用いて行う。
- ジャズの専門家によるヴォイスニング(以下、専門家ヴォイスニング)  
専門家が、電子オルガンで演奏することを前提に1コードネームに対して1ヴォイスニングを決定する。また、和音の構成音数の決定や代理和音の使用も可能とする。
- 単純なルールによるヴォイスニング(単純ヴォイスニング)  
コードネームの構成音をすべて用い、それ以外の音は使用しない。楽曲の最初のヴォイスニングは、システムと同様、基準音(左手ヴォイスニング: E3, ベース: D2)から最も近い音をbottom noteとし、middle notesとtop noteはbottom noteの決定後、順に音を重ねる。それ以降のヴォイスニングは、1つ前のヴォイスニングから音高差が最も小さくなるよう転回形を決定する。

楽曲はErroll Garner作曲の“Misty”を用いた。上記3種類のヴォイスニングを市販のソフトシンセサイザを用いて音響信号に変換し、評価者は、それを聴いて全15種類の形容詞対に対して7段階で評価した。この評価において、システムヴォイスニングの示す傾向が、単純ヴォイスニングの傾向よりも専門家ヴォイスニングのそれに近ければ、そのヴォイスニングはジャズらしいと判断することができる。音色は右手パートにフルート、左手パートにストリングス、ベースパートにアコースティックベースを用い、演奏に表情は付けずに変換処理を行った。ジャズ音楽の経験が5年以下で20~25歳の計25人が評価した。

15種類の形容詞対を独立変数とし、各々の項目の得点を従属変数として1要因の分散分析を行った。その結果、“個性のある—個性のない”、“自然な—不自然な”、のそれぞれにおいて有意差が見られた(順に $F(2, 48) = 3.25, p < .05$ ;  $F(2, 48) = 6.65, p < .01$ )。

“個性のある—個性のない”の評価得点の結果を図2に示す。この結果からシステムと専門家によるヴォイスニングでは、単純ヴォイスニングに比べて個性的であることが分かる。この原因の1つに、前者の2つのヴォイスニングに含まれているテンションが考えられる。テンションは、緊張感を付与する効果があり、またジャズ音楽の特徴を効果的に表す役割がある。単純ヴォイスニングには、テンションが含まれていないため、それが個性的でないという評価に表れたと考えられる\*1。

次に、図3に“自然な—不自然な”の評価結果を表す。この結果から、単純ヴォイスニング

\*1 ただし、テンションがこの理由であることは可能性の1つにすぎず、実験結果から明らかになったわけではない。

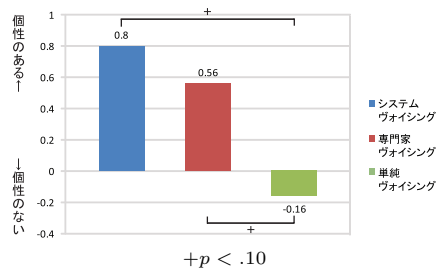


図2 “個性がある—個性がない”の評価

Fig. 2 Evaluations of “characteristic or not”.

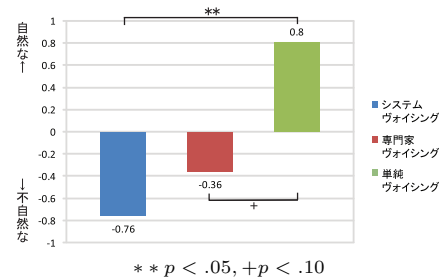


図3 “自然な—不自然な”の評価

Fig. 3 Evaluations of “natural or not”.

は自然と感じ、システムと専門家によるヴォイシングは不自然と感じることが分かる。これは、評価者自身がジャズ音楽に不慣れ、または親しみが無いため、テンションを多く含むジャズ音楽が不自然に聞こえたと考えられる<sup>\*1</sup>。実際、評価者25人のうち、ジャズ音楽を最もよく聴くと答えたのは2人で、ほとんどの評価者は、ふだん、ポップスやロックを聴いていると答えている。また、システムヴォイシングの方が音楽専門家によるそれより負の値が大きい。これは、前者には楽曲全体のうち数カ所だけ音楽的に不自然なヴォイシングが出力され、その印象が強かったからと考えられる。

#### 4.4 音楽的同時性・音楽的連続性に関する定性的な評価

本研究で用いたモデルのメロディ情報と前後のヴォイシング情報が、音楽的同時性や音楽的連続性の実現に寄与しているかどうかを確かめるため、評価を行った。図1で示したモデルから前後のヴォイシングに関するノードを削除したもの(モデルA)、メロディに関するノードを削除したもの(モデルB)、すべてのノードが有効なもの(モデルC)を用意し、同じ楽曲に対するヴォイシングの結果を比較した。なお、音楽的同時性や音楽的連続性を定量的に評価する尺度は存在しない<sup>\*2</sup>ため、定性的な考察による比較とした。

##### (1) モデルA：メロディ情報のみを考慮するモデル

出力結果には、メロディと衝突する音はなく、音楽的同時性の感じられるヴォイシングが出力された(図4)。しかし、ヴォイシングどうしのつながりはあまり感じられず、矢印で示すとおり、ほとんどのヴォイシングが上下に飛躍していることが分かっ

た。一時的なヴォイシングの飛躍は一般的にもみられるが、このようにつねに上下進行があるヴォイシングは、音楽的連続性があるとはいえない。

##### (2) モデルB：前後情報のみを考慮するモデル

出力結果を図5に示す。ヴォイシングどうしのつながりが比較的なめらかであることが分かる。bottom noteの推移はG-G-Bb-Bb-A、middle notesは、B-B-C-A-D、top noteはD-D-D-D-Eと完全5度以上の飛躍はみられなかった。しかし、四角で示したGm7とC7においては、ヴォイシングどうしの音の衝突やメロディ音との衝突が見受けられた。

##### (3) モデルC：メロディと前後情報の両方を考慮するモデル

このモデルでは、図6に示したとおりメロディと前後の情報の両者の利点を取り入れたヴォイシングが出力された。

以上のことから、本研究で提案したモデルでは、ヴォイシングにとって重要な音楽的同時性と音楽的連続性が考慮されているといえる。

#### 4.5 出力結果例と考察

例として楽曲“Misty”に対するヴォイシング結果を図7、図8に示す。1段目に与えたメロディとコードネーム、2段目が左手ヴォイシング、3段目がベースの推定結果である。

左手ヴォイシング結果は、20小節めの3拍めでベース音と不協和を生じた以外は、明らかに不協和であるものは見受けられず、全体的に、妥当なヴォイシングがされたものが多かった。しかし、他のパートと不協和を生じないものの、楽譜の指定と異なるコードのヴォイシングを出力するケースがあった。たとえば、4小節めの1拍めは楽譜ではFmと指定されているが、出力されたヴォイシングはF(正確にはFadd9)である。これは条件付き確率のスムージングによるものと考えられる。BayoNetではデータスパースネス問題に対する一対処法として条件付き確率のスムージングを自動的に行う。そのため、学習データに出現しなかった項目(たとえばコードネームがFmという条件下でのヴォイシングの音名がA)に対しても0よりも大きな確率値が与えられる。その場合、メロディとの共起確率によってはコードFmに対してAの音名が選ばれることもありうる。20小節めの3拍め、31小節めの4拍めも同様の理由によって指定されたコードではないコードのヴォイシングが出力されたと考えられる。

テンションノートの付与に関しては、9thや13thが付与されたものが多く生成された。全62コード中、30コードにテンションノートが付与されていた。このモデルを構築する際に使用したデータベースの中でも、テンションが使用されていた事例は約4割を占め、ほぼ

\*1 こちらも同様に可能性の1つにすぎず、実験結果から明らかになったわけではない。

\*2 たとえば音高の推移の小ささで音楽的連続性を評価することが可能だが、適切なタイミングで大きく音高が推移することはむしろ好ましいことであり、十分な評価尺度とはいえない。



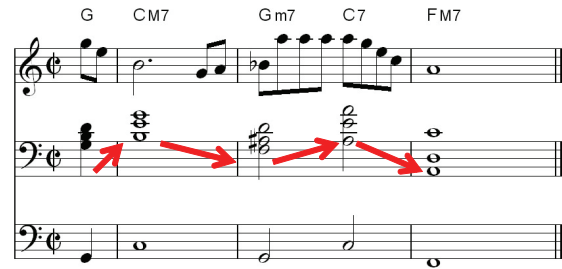


図 4 モデル A の出力結果例  
Fig. 4 Example of Model A's outputs.

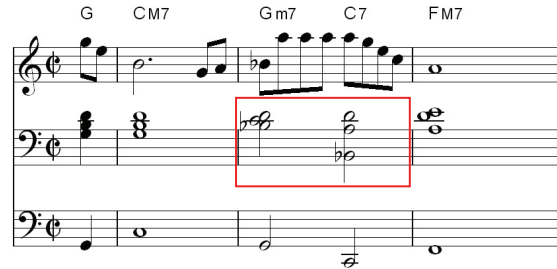


図 5 モデル B の出力結果例  
Fig. 5 Example of Model B's outputs.

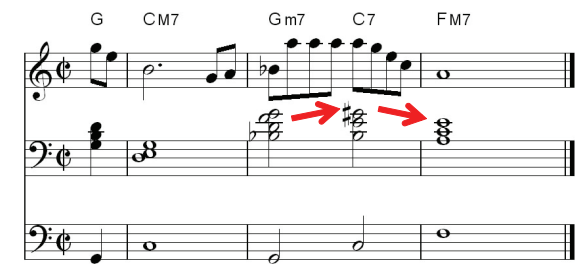


図 6 モデル C (提案モデル) の出力結果例  
Fig. 6 Example of Model C's outputs.

同等の割合であった。また、2つのテンションが同時に選択されている箇所もあった。4小節めのB♭Majorには9thと13thの含まれるヴォイシングが出力されている。理論書<sup>16)</sup>の中で、“4声以上で13thを用いるときは、9thを入れたほうが響きが豊かになる”といわれているヴォイシングである。また、すべてのヴォイシングが3和音または4和音で構成されており、実際に左手で演奏可能な範囲で出力されたことから、適切な省略音が選択されたと考えられる。

前後のつながりに関しては、とこところで半音による進行がみられた。2小節めのソ→ソ♯への進行や8小節めのF♯→F、15小節めのDm7→G7のヴォイシングによるC→Bなどがあげられる。ヴォイシングを3要素に分けることで、うまくbottom noteやtop noteが作用したと考えられる。しかし、これらは2連続のヴォイシングによる進行であり、3連続以上のヴォイシングによる効果的な進行は見受けられなかった。また、24小節めのDm7→G7は、演奏可能音域(C3~A♯4)を決定したため、半音で推移可能するとB2→D3→F3であるが、B2は音域外であるので1オクターブ上に移動して不自然な飛躍した進行になっている。

ベースに関しては、約95%がルート音を出力し妥当な結果を得た。20小節めのGdimのような経過をスムーズにするために置かれている比較的拍数の少ないコードでは、ルート音以外の音出力されることがあった。

#### 4.6 モデルの簡略化に関する考察

ベイジアンネットワークなどの確率モデルでは、モデルが詳細になるほど学習に必要なデータは増大する。有限の学習データを用いてモデルを学習できるようにするため、本研究では3.4節で述べたようにモデルを簡略化した。本節では簡略化の影響を定性的に考察する。

#### ● 左手ヴォイシングモデルとベースモデルの分割

4.2節の評価で許容できないと評価された5%のヴォイシングには、左手ヴォイシングとベース音がミスマッチを起こすケースが含まれていた。これは、左手ヴォイシングモデルとベースモデルの分割の弊害といえよう。しかし、約95%は妥当と判断されており、左手ヴォイシングとベース音が適切に組み合わせられたヴォイシングが出力されていた。たとえば、図7の1小節めではベース音“C”と左手ヴォイシング“D-E-G”によってCM7(正確にはCadd9)が表現されている。これは、左手ヴォイシングとベース音の依存関係を十分に考慮した事例を学習に与えているために、この依存関係を間接的に学習したからと考えられる。

#### ● 直前・直後のコードのみの考慮

4.2節の評価および4.4節における前後情報を用いないモデルとの比較から、直前・直後のコードを文脈として用いるだけで音楽的連続性を考慮できていることが分かった。ただ、同じメロディが繰り返されている箇所(たとえば1~6小節めと9~14小節め)では、1つめのコード以外はまったく同じヴォイシングだった。一方、4.3節の評価で用いた音楽専門家によるヴォイシングでは、このような箇所でもヴォイシングに変化をつけていた。楽曲を盛り上げるためにはメロディが同じであってもヴォイシングを変えることが望ましい。これを行うには、より大局的な文脈のモデル化が必要である。

#### ● ヴォイシングとメロディの依存関係の逆転

4.4節での議論により、ヴォイシングとメロディの依存関係を逆転させても、メロディとの兼ね合いを考慮した適切なヴォイシングを行えることが分かった。

以上より、一部で簡略化の影響が出たものの、一定の音楽的同時性と音楽的連続性を考慮

Chord voicings for measures 1-12 of "Misty":  
 1: G, CM7  
 2: Gm7, C7  
 3: FM7, Fm  
 4: Bb  
 5: CM7  
 6: Am  
 7: Dm7, G7  
 8: E7, A7  
 9: D7, G7  
 10: CM7  
 11: Gm7, C7  
 12: FM7, Fm, Bb

図 7 楽曲“Misty”でのヴォイシング結果  
 Fig. 7 Example of voicing with “Misty”.

Chord voicings for measures 13-28 of "Misty":  
 13: Gm7, C7, FM7, F6, F6, Gdim, F6  
 14: B7, D7, G7, C#dim, Dm7, G7, G7  
 15: CM7, Gm7, C7, FM7, Fm, Bb  
 16: CM7, Am, Dm7, G7, C, FM7, Em7, D#7, CM7

図 8 楽曲“Misty”でのヴォイシング結果(つづき)  
 Fig. 8 Example of voicing with “Misty” (cont'd).

することはできたといえる。

## 5. 議 論

### 5.1 本研究の到達点と残された課題

本研究では、確率モデルに基づく自動コードヴォイシングについて検討し、専門家による評価、非専門家による印象評価およびモデルに対する定性的評価を行った。その結果、

- 明示的なルールを記述することなく、音楽的同時性と音楽的連続性の観点から、一定の妥当性を持つヴォイシングを出力できること、
- 特定のジャンルのみの事例集を用いることで、そのジャンルの特徴を反映したヴォイシングが可能であること、
- 学習データが有限であることを考慮してモデルの簡略化を行っても、ほぼ妥当なヴォイシングが可能であること、

を明らかにすることができた。

しかしながら、これは、本システムのヴォイシングが既存のシステムのもの<sup>2),3),5)</sup>よりも優れていることをただちに意味するわけではない。今後、これらのシステムと性能比較をすることで、2.2節で議論した各アプローチのメリット・デメリットを検証する必要がある。その他の課題としては次のものなどがあげられる。

- トップダウン情報の導入  
音楽理論などの知見を条件付確率の形で記述することで、ベイジアンネットワークに取り込むことができる。これにより、事例と既存の音楽知識を統合した推論が可能になる。
- 他のジャンルへの適用  
今回提案したモデルを他のジャンルにも適用することで、モデルを変えることなく事例集だけを切り替えることでジャンルを切り替えられるのか検証する必要がある。

### 5.2 ベイジアンネットワークを用いた他の研究との比較

ベイジアンネットワークは、音楽が持つ階層性と時間的依存性を自然に記述できるため、過去の音楽情報処理研究においても用いられてきた。

柏野らの音楽情景分析モデル OPTIMA<sup>17),18)</sup> は、周波数成分、単音、和音の3つのレイヤからなるベイジアンネットワークを用いて音楽音響信号から単音・和音を抽出するものである。このネットワークでは、同時刻のノードがレイヤ間で結合されるほか、和音レイヤのノードが時間軸に沿って結合されている。各ノードは、単音形成などの各種ボトムアップ処理やトップダウン処理の仮説集合を表し、これらの処理結果が仮説の確信度の形で反映され

る。同時刻のノードのレイヤ間の結合（縦リンク）と時間軸に沿った結合（横リンク）の2方向の結合をベイジアンネットワークで表現している点は、本研究と共通である。しかし、目的が音響信号処理であるため周波数成分を表すノードが用意されている点、縦リンクが和音レイヤにのみ用意されている点が異なる。

川上らの和声付けの研究<sup>1)</sup>では、隠れマルコフモデル(HMM)を用いている。HMMはダイナミックベイジアンネットワークの特殊な形であり、観測できる出力記号を表すノードと隠れ状態を表すノードの2つのレイヤで構成される。この研究では、旋律を出力記号列、和音を隠れ状態と見なして、旋律の背後にある和音列を推定した。隠れ状態ノードと観測ノードの関係は上記の縦リンクに、隠れ状態のマルコフ過程は横リンクに相当する。ただし、HMMはネットワークがこの形態に固定されているので、ヴォイシングのために新たなレイヤを追加することは基本的にできない。

## 6. おわりに

本論文では、ベイジアンネットワークを用いた事例学習型の自動ヴォイシング手法を提案した。ヴォイシングにおいて重要な側面である音楽的同時性と音楽的連続性を、ベイジアンネットワークにおけるノード間の確率的依存関係としてモデル化することで、音楽的同時性と音楽的連続性をできるだけよく満たすヴォイシングを確率的推論として求めることを可能にした。

本研究では、メロディとコードネームを既知として、それに合う左手ヴォイシングとベース音を生成したが、何を既知にして何を推定するかを切り替えることで、様々なタスクに対応できると考えられる。今後は、本研究の成果をベースに、ヴォイシング以外の様々なタスクに対応できるようにし、音楽の汎用推論モデルの確立を目指していきたい。

## 参 考 文 献

- 1) 川上 隆, 中井 満, 下平 博, 嵯峨山茂樹: 隠れマルコフモデルを用いた旋律への自動和声付け, 情報処理学会研究報告, 99-MUS-34, pp.59-66 (2000).
- 2) 平田圭二, 青柳龍也: パーピーブーン: ジャズ和声を生成する創作支援ツール, 情報処理学会論文誌, Vol.42, No.3, pp.633-641 (2001).
- 3) 後藤真孝, 平田圭二: ハービー君: 演繹オブジェクト指向に基づいてジャズらしいコードにリハーモナイズするシステム, 情報処理学会研究報告, 96-MUS-16, pp.33-38 (1996).
- 4) 三浦雅展, 青山容子, 谷口 光, 青井昭博, 尾花 充, 柳田益造: ポップス系の旋律に

- 対する和声付与システム：AMOR，情報処理学会論文誌，Vol.46, No.5, pp.1176–1187 (2005).
- 5) 江村伯夫，三浦雅展，柳田益造：与えられた旋律に対するコード・ネーム付与システムと，テンション・ノートを考慮したヴォイシング・システムの統合，日本音響学会音楽音響研究会 (2007).
  - 6) Pachet, F.: Computer Analysis of Jazz Chord Sequences, Is Solar a Blues?, *Readings in Music and Artificial Intelligence*, Harwood Academic Publisher (2000).
  - 7) Chemiller, M.: Improvising Jazz Chord Sequences by Means of Formal Grammers, *Proc. International Computer Music Conference* (2001).
  - 8) Conklin, D.: Representation and discovery of vertical patterns in music, *Music and Artificial Intelligence: Lecture Notes in Artificial Intelligence*, Vol.2445, pp.32–42 (2002).
  - 9) Balaban, M., Ebcioğlu, K. and Laske, O. (Eds.): *Understanding Music with AI*, MIT Press (1992).
  - 10) Pachet, F.: Musical Harmonization with Constraints: A Survey, *Constraints Journal*, Vol.6, No.1, pp.7–19 (2001).
  - 11) 片寄晴弘，平田圭二，原田利宣，平賀瑠美，笠尾敦司：事例に基づくデザイン支援と評価基盤の構築，人工知能学会論文誌，Vol.18, No.1, pp.24–28 (2003).
  - 12) 平田圭二：アートからデザインへの転換について—音楽情報処理の事例から，人工知能学会全国大会，1B1-01 (2003).
  - 13) 諸岡 孟，西本卓也，嵯峨山茂樹：確率文脈自由文法を用いた和声学規則の表現と楽曲の自動和声解析，情報処理学会研究報告，2008-MUS-74, pp.77–82 (2008).
  - 14) Pearl, J.: *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann (1988).
  - 15) 株式会社数理システム：ペイジアンネットワーク構築支援システム BayoNet. <http://www.msi.co.jp/BAYONET/>
  - 16) 松田 昌：松田昌の音楽講座ポピュラーアレンジの基礎知識，ヤマハ (1986).
  - 17) 柏野邦夫，中臺一博，木下智義，田中英彦：音楽情景分析の処理モデル OPTIMA における単音の認識，信学論，Vol.J79-D-II, No.11, pp.1751–1761 (1996).
  - 18) 柏野邦夫，中臺一博，木下智義，田中英彦：音楽情景分析の処理モデル OPTIMA における和音の認識，信学論，Vol.J79-D-II, No.11, pp.1762–1770 (1996).

(平成 20 年 5 月 14 日受付)

(平成 20 年 11 月 5 日採録)



北原 鉄朗 (正会員)

2002 年東京理科大学工学部卒業．2007 年京都大学大学院情報学研究科博士後期課程修了．博士 (情報学)．日本学術振興会特別研究員 (DC2) を経て，現在，科学技術振興機構 CREST (CrestMuse プロジェクト) 研究員．音楽情報処理，聴覚的情景分析等に興味を持つ．電気通信普及財団第 19 回テレコムシステム技術学生賞，京都大学第 2 回総長賞等受賞．電子情報通信学会，人工知能学会，日本音響学会，IEEE 等各会員．



勝占真規子

2006 年関西学院大学工学部情報科学科卒業．2008 年関西学院大学理工学研究科情報科学専攻修士課程修了．同年大日本印刷株式会社に入社し，現在に至る．音楽情報処理，データマイニングに興味を持つ．



片寄 晴弘 (正会員)

1991 年大阪大学大学院基礎工学研究科博士課程修了．工学博士．イメージ情報科学研究所，和歌山大学を経て，現在，関西学院大学工学部教授．ヒューマンメディア研究センターセンター長．音楽情報処理，感性情報処理，HCI の研究に従事．科学技術振興機構さきがけ研究 21「協調と制御」領域研究者．科学技術振興機構 CREST「デジタルメディア (略称)」領域 CrestMuse プロジェクト代表研究者．電子情報通信学会，人工知能学会各会員．



長田 典子 (正会員)

1983 年京都大学理学部数学系卒業．同年三菱電機 (株) 入社．産業システム研究所等で産業計測機器の研究開発に従事．1996 年大阪大学大学院基礎工学研究科博士後期課程修了．2003 年より関西学院大学工学部情報科学科助教授，2007 年教授．専門は感性情報学，メディア工学．博士 (工学)．電子情報通信学会，IEEE 等各会員．