

音楽を理解するロボットの実現に向けて

リサーチアシスタント 米田隆一

情報理工学系研究科システム情報学専攻

概要

1 はじめに

音楽を理解するロボットの実現のための基盤技術として、演奏情報 (MIDI) および楽譜等のシンボリックな音楽情報を入力とする対旋律、和音、調などのラベルを付与する汎用的な手法を提案する。これらは、音楽情報検索、作曲や編曲の支援、楽器の練習支援などの場面で重要な役割を果たす。

2 従来技術の問題点

現在実現されている音楽解析技術においては、マルコフ連鎖などのモデルを適用することが多かった。しかし、MIDI、楽譜などの2次元的な情報 (縦の和音、横の声部進行等) は、マルコフ連鎖のようなモデルでは不十分であり、より広いコンテキストの重視が必須である。

3 提案手法

音楽情報解析においてマルコフ確率場 (最大エントロピーモデル) を用いて、音楽的知識を素性関数の設計という操作に還元することにより、汎用性が高く高性能な手法を実現する。

具体的には、以下の構成で音楽情報解析を行う。

1. マルコフ確率場を用いて、演奏情報および楽譜等における音階などの音楽情報を、周辺の文脈に依存する値としてモデル化を行う。具

体的には、式 (1) において V_c を素性関数 f_i とそれに対応する重み λ_i の積であるとみなす。素性関数 f_i は通常ラベル (和声や調など解析結果として得たい情報) の有無を表す二値関数を用いる。これは最大エントロピーモデルにおける確率分布関数の式 (2) と本質的に同一である。

2. マルコフ確率場モデルを用いて確率の積が最大となるラベル系列を探索する。具体的に、Maximum Entropy Markov Model (MEMM) の手法を用いる場合は、最大エントロピー法 (ME) の順次適用により実現する。最適解の探索にはビームサーチ等が適用できる。また、Conditional Random Fields (CRF) の手法を用いる場合は、ME の特殊形、かつ、HMM の一般化に相当する手法となり、入力系列そのものに関するグローバルな最適解が得られる。なお、式 (2) におけるパラメータ λ の推定には反復スケールング法が適用できる。

$$P(f) = \exp\left(\sum_{c \in \mathcal{C}} V_c(f)\right) / Z \quad (1)$$

$$p_\lambda(y|x) = \exp\left(\sum_i \lambda_i f_i(x, y)\right) / Z_\lambda(x) \quad (2)$$

4 効果

マルコフ確率場を (最大エントロピーモデル) を用いて、音楽的知識を素性関数の設計という操

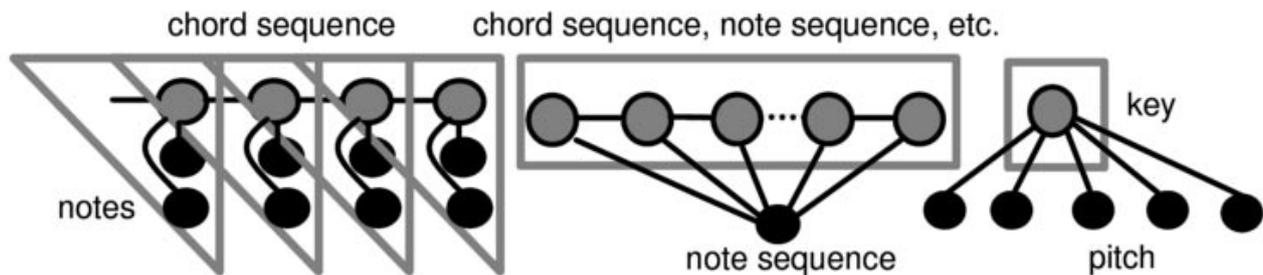


図 1:

作に還元し、広いコンテキストを柔軟に設計できる音楽情報分析を実現する。マルコフ確率場のグラフによる表現(図2)において黒のノードは入力、グレイのノードは出力、グレイの線で囲まれた領域は最適解を表す。和声解析は図1左の概念図のように MEMM を用いて、各声部の音高と前和音から最尤の和音を求める操作を繰り返し適用していく。対旋律付けおよび和声付けは図1中の概念図のように CRF を用いて、入力系列そのものから、グローバルな最適解を求める。調認識は図1右の概念図の構造を使用し、楽曲全体の音高頻度により調を決定する。それぞれの課題において MIDI 情報および humdrum データベースを入力とした実験を行った結果、対旋律付けで 60~70%、和声付けで 60~70%、ドミナント定型同定で 99%、和声解析で 75%、調認識で 82% の認識率となり、提案手法の有効性が示された。

本手法の特徴となる点は以下のとおりである。

1. 演奏情報および楽譜等における音階などの音楽情報に対して、対旋律付け、和声付け、ドミナント定型同定、和声解析、調認識などのラベル付与を行う課題において、マルコフ確率場を用いて周辺の文脈に依存する値として

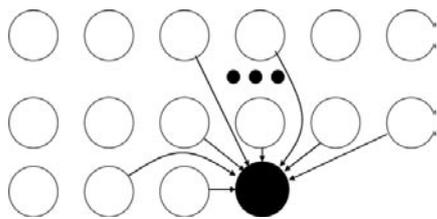


図 2:

モデル化を行う手法。

2. 1. において、マルコフ確率場モデルを用いて確率の積が最大となるラベル系列を探索する際に、Maximum Entropy Markov Model (MEMM) の手法を用い、最大エントロピー法 (ME) の順次適用により実現し、最適解の探索にビームサーチを適用する手法。
3. 1. において、Conditional Random Fields (CRF) と呼ばれる、最大エントロピーモデルの特殊形、かつ、HMM の一般化に相当するモデル化を行う手法。

5 ロボットの実現の見通し

本手法は、高機能な電子楽器や、音楽を理解し演奏を行うロボットやコンピュータを実現するための基盤技術である。和声解析や調認識、対旋律付け、和声付けなどは、初心者や協奏曲演奏者の練習支援、アマチュアの音楽演奏などを支援する楽器に利用できる。その他にも、ロボット、ソフトウェア、コンテンツ産業において利用可能である。

6 まとめ

案内ロボットのシステム開発において、音楽を理解するロボットの実現のための基盤技術として、マルコフ確率場モデルを用いた音楽情報解析手法を提案した。また、大学院の講義である実世界システム講究に参加してプロジェクト進行のための議論を進めた。