

関係調を考慮したパラメータ共有HMMに基づく 音響信号の自動和音認識の検討

A Study of Automatic Chord Recognition from Musical Acoustic Signal Based on Parameter-Shared HMM Considering Relative Key

杉山 雄一[†]
Yuichi Sugiyama

奥村 健太[†]
Kenta Okumura

酒向 慎司[†]
Shinji Sako

北村 正[†]
Tadashi Kitamura

1. はじめに

近年、計算機やネットワークの普及により、音楽コンテンツの多様化・大規模化が進んでいる。一方で多量の音楽を柔軟かつ効率的に検索する技術は十分とはいえず、音響信号の内容そのものを音楽的な側面で解析する技術が求められている。本稿では、楽曲の内容を表す情報の一つとして和音進行に着目し、音響信号の自動和音認識の問題を扱う。自動和音認識においては、隠れマルコフモデル(HMM)を用いた手法として、前後和音に依存した詳細な環境依存和音HMMによる自動和音認識手法[1]が提案されているが、環境依存モデルの分類の偏りにより局所的にモデルの過学習が起こり、結果としてモデルの汎用性が失われるという問題点があった。本稿では、関係調の知見に基づいてモデルパラメータの共有を柔軟に行う手法を提案し、その有効性を検討する。

2. 環境依存HMMによる和音認識

2.1 特徴量:クロマベクトル

和音には複数の転回形が存在し、音高の配置が異なる場合でも構成音が同じであれば同一の和音として認識される。そこで、短時間パワースペクトルをオクターブ間に帯域を分割し、音名ごとにオクターブ間で足し合わせた特徴量が有効であると考えられる。このような特徴量はクロマベクトルと呼ばれ、音響信号の和音認識でよく利用されている。本稿では式(1)のように1つの半音が1つの次元に相当する12次元のクロマベクトルを用いる。ここで、 $H(f, t)$ はスペクトログラムにおける周波数 f 、時間 t でのパワーを表し、 I は加算するオクターブの範囲を表す。このクロマベクトルの時系列ベクトルを演奏パターンとみなす。

$$c(k, t) = \sum_{f=0}^{I-1} H(12f + k, t) \quad (1)$$

2.2 HMMによる和音進行のモデル化

和声学に基づくと、和音を構成する音名の出現には一定の傾向があるといえるため、クロマ系列でも同様のことが言え、和音に依存して出現しやすい音名の組み合わせはクロマベクトルの分布によって表現できる。また、和音間の遷移の傾向においても和声学に基づく一定の規則があり、確率的に扱うことができる。これらの特徴を考慮した和音進行のモデルとして、一般的にHMMがよく用いられる。また調性音楽では、メロディは和音進行に基づいて作成されると考えることができるため、自動

和音認識はHMMを用いて、観測されたクロマ系列から和音進行を求める逆問題として定式化できる。本研究では、1つの和音が1つの状態に対応し、全ての和音へ遷移可能なergodicHMMで和音進行をモデル化する。また簡略化のため、全ての調を1つのモデルとして扱う。

2.3 環境依存HMMの利用

HMMを用いた和音認識は過去にも研究例があり、H.Papadopoulosら[2]は出力確率分布を直接学習せず、音楽理論に基づいて定めた。出力確率分布は単一の正規分布を仮定し、倍音の影響を考慮することで従来の手法よりも高い認識率を得ている。一方、内山ら[3]はクロマベクトルはその時刻の和音のみに依存して確率的に生成されると仮定し、各和音のクロマベクトルから出力確率分布を学習した。実際の演奏では和声内音が省略されることも多く、1つの和音から生成されるクロマのパターンは多様であると考えられる。そこで、ある当該和音の特徴が前後の和音にも影響されると仮定し、音声認識の分野で広く用いられるtriphoneモデルの考え方をいすることでモデルの詳細化を行う。

2.4 関係調に基づくHMMのパラメータ共有

上記のモデルの詳細化により生じる問題として、すべての和音連鎖において網羅的に学習データを用意することが難しく、学習データが存在したとしても十分な学習ができなければ、過学習によって汎用的な和音連鎖モデルを推定できないといった点が挙げられる。本研究では、既存の和音連鎖モデル間においてはパラメータを共有することで頑健なモデルを学習する一方で、未知の和音連鎖に対しても適切なモデルを選択可能な枠組みとして、決定木に基づいたクラスタリング手法を導入する。この手法に関し先行研究[1]では、クラスタリングの際に前後和音に関する分類基準を用いていた。和声学に基づくと、楽曲を構成する和音連鎖に頻出傾向が存在するため、上記の手法ではモデルのパラメータ共有がその傾向に左右され、結果的に過学習が起こりやすいパラメータが生じるといった問題があった。

上記の改善に有効な特徴として、和声学に基づく関係調という考え方がある。関係調とはそれぞれ12個ずつある長調・単調の中での和音間の近親度合いを表現したもので、これらを長澤ら[4]は和音間の編集距離の考え方をいって数量化している。多くの楽譜から和音間の編集距離の統計をとることで、和音連鎖の出現傾向を特徴化でき、その特徴をもとに前後和音に対する分類基準を定めクラスタリングを導入することで、柔軟なパラメータ共有が可能になると考える。

[†]名古屋工業大学, Nagoya Institute of Technology

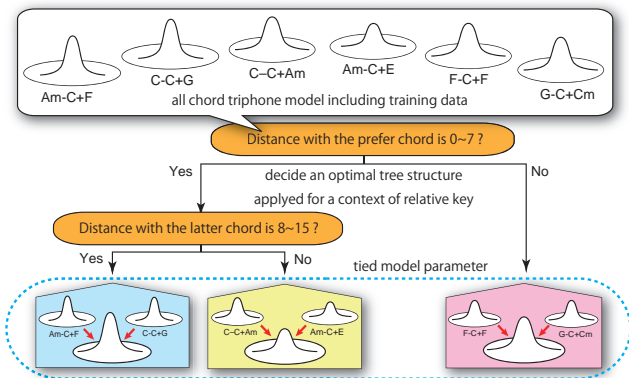


図 1: 関係調を分類基準とした決定木によるパラメータ共有

表 1: 実験データと分析条件

アーティスト	The Beatles
楽曲	“Please Please Me” の 14 曲
サンプリング周波数	11,025 Hz (モノラル)
フレーム長	100 msec
オクターブ数	55.0 Hz (A0) ~ 3,729.3 Hz (B♭6)
特徴ベクトル	12 次元クロマベクトル+Δ

3. 実験

3.1 実験条件

提案手法の評価実験として、楽曲に含まれる音響信号からクロマベクトルを抽出し、和音連鎖 HMM を学習する。楽曲は CD のデータをモノラル化し、ダウンサンプリングして用いた。表 1 に示す実験条件のもと、定 Q フィルタバンクを用いて時間周波数解析を行った。和音の語彙は major, minor の計 24 種類とし、それ以外の和音は第 3 音に着目して major, minor に近似した。各和音の出力確率は単一の多次元正規分布と定め、クロマベクトルの各次元間の相関を考慮しない対角共分散行列とした。楽曲 14 曲中 13 曲を学習データとして用い、各学習セットの中で出現する和音連鎖について HMM を学習し、次の分類基準に基づく決定木クラスタリングによりパラメータを共有する。

1. 前後の和音に関する分類基準 (従来法 [1])
2. 上記 + 前後和音に対する関係調に関する分類基準

その際、クラスタリングの停止基準を調整し、段階的に決定木の規模を変化させながらモデルを作成した。

3.2 モデルの評価

パラメータ共有の際に関係調に関する分類基準を導入した効果を確認するため、未学習データに対する認識率でモデルを評価する。先の 2 つの分類基準を用いた際のモデルの規模ごとの認識率を図 2 に示す。モデルの規模の下限は、前後環境によるモデル分類を全く考慮しない単独の和音モデルに相当し、規模の上限は学習データに出現した全ての和音連鎖ごとに完全に分類した場合に相当する。

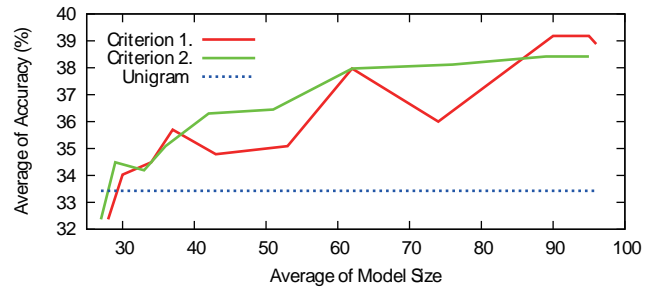


図 2: モデルの規模ごとの和音の認識率

クラスタリングの分類基準に関わらず、モデルの規模を小さくした場合はモデルが過剰に平滑化されるため認識率が下降し、モデルの規模を大きくした場合は各和音連鎖において詳細なモデル化がなされるため、認識率が上昇する傾向にあると考えられる。適度なモデル規模をとった際は、関係調に関する知見を導入した分類基準 2. において高い認識率が得られた。これは同じ粗さのモデルにおいても、関係調の知見をパラメータの分類基準に導入することにより、未学習データに対してより適合しやすいモデル化がなされていることが示されている。

4. むすび

HMM に基づく音響信号からの自動和音認識の手法において、先行研究 [1] では環境依存モデルの分類の偏りにより局所的にモデルの過学習が起こり、結果としてモデルの汎用性が失われるといった問題が存在した。それに対し本稿では、前後和音に対する関係調に関する分類基準により柔軟にモデルパラメータの共有を行う手法を提案し、より汎用性の高いモデルの構築を目指した。

分類基準別のモデルを用いての認識実験から、適度なモデル規模とした際、関係調の知見を用いてモデルを分類することで認識率が向上することが確認でき、本手法の有効性が示された。今後の課題としては、モデルの規模を一定の基準に基づき決定することが挙げられる。また、クロマベクトルの分布は楽曲のジャンルや使用楽器にも依存すると考えられるため、楽曲そのものを区別する分類基準の導入が有効であると考えられる。

参考文献

- [1] 伊藤 綾, 他, “パラメータ共有 HMM に基づく音響信号からの自動和音認識の検討”, 情報処理学会研究報告, 2010-MUS-88(4), 2010.
- [2] H. Papadopoulos *et al.*, “Large-scale study of chord estimation algorithms based on chroma representation and HMM”, *Proc. CBMI*, pp.53–60, 2007.
- [3] 内山 裕貴, 他, “調波音を強調したクロマに基づく音楽音響信号からの自動和音認識”, 日本音響学会春季研究発表会講演集, pp.901–902, 2008.
- [4] 長澤 槇子, 他, “ポピュラー音楽クラスタリングのための近親調を用いたコード進行類似度の提案”, 情報処理学会研究報告, 2007(37), pp.69–76, 2006.