

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*

内山裕貴, 宮本賢一, 西本卓也, 小野順貴, 嵯峨山茂樹 (東大院情報理工)

1 はじめに

近年、音楽の内容に基づく解析の研究が盛んである。この研究には、音楽情報検索 (MIR) や、自動採譜など様々な応用がある。本稿では音楽の内容に関する情報として、和音に着目し、音楽音響信号から自動和音認識を行う。特徴量に調波音を強調したクロマを用いる手法を提案し、評価実験により認識率が向上したことを示す。

2 調波音を強調したクロマによる和音認識

2.1 調波音を強調したクロマベクトル

和音は異なるオクターブで演奏されても、構成音が同一であれば同一の和音と認識される。このような和音の性質を上手く扱う特徴量として、クロマベクトル [1] があり、音楽音響信号からの自動和音認識でも使われている [2]。クロマベクトルは 12 個の半音名の強さであり、パワースペクトルを半音名ごとに複数オクターブ間で足し合わせることで得られる。

さて、本稿で主に対象とするポピュラー音楽は、一般に打楽器音のような非調波的な音を含むが、打楽器音はピッチを持たない音であり、音楽の調波的な要素である和音を認識する際に妨げとなる。そこで本稿では、宮本ら [3] の調波音・打楽器音の分離を用い、調波音を強調し打楽器音を抑制する。この手法は、「スペクトログラムにおいて、打楽器は時間方向に急峻に変化するが周波数方向には滑らかであり、調波音は逆に周波数方向に急峻であるが時間方向には滑らかである」という点に着目し、スペクトログラム上で、時間方向と周波数方向の滑らかさを用いて、調波音と打楽器音を分離する。本稿では、まず入力した音楽音響信号から調波音・打楽器音分離により調波音を抽出し、これをクロマ化した特徴量を用いる。

2.2 HMM の利用

和音は、一般に構成音がそのまま演奏されるわけではなく、和声内音の省略・非和声音の挿入があり、時刻ごとにみて和音を認識するのは難しい。そこで、和音は、時間ごとにランダムに出現するのではなく同じ和音がある程度の時間連続したり、和音間での遷移のしやすさには偏りがあったりするという性質を利

Table 1 和音認識の正解率

楽曲	混合信号	調波信号
RWC-MDB-P-2001 No. 05	61.7%	66.0%
RWC-MDB-P-2001 No. 20	51.3%	68.0%

用し、現在の和音を認識するのに周辺の時刻の和音を手がかりとして利用する。このように和音を扱うモデルとして川上らは HMM を用い [4]、音楽音響信号からの和音認識でもよく使われるモデルとなっている [2]。和音を、音楽に直接表れないが隠れて存在する状態と捉え、各和音の出力確率分布により演奏パターンが出力されるとみなすことで、和声内音の省略・非和声音の挿入を確率的に扱うことが出来る。また、和音の遷移はマルコフ過程に従うと仮定し、状態間の遷移確率によって和音の進行しやすさを反映できる。本稿でも、和音のモデルとして HMM が有効と考えこれを用いる。

3 評価実験

3.1 実験条件

本システムを、RWC 研究用音楽データベースより 2 曲 (RWC-MDB-P-2001 No.5, No.20) を用いて評価した。それぞれの楽曲からドラムを含む部分 30 秒を抜粋し、モノラル化したものを入力の音楽音響信号とした。サンプリング周波数は 44100Hz である。時間周波数解析は 27.5Hz から半音に 96 個 (8 オクターブ) の中心周波数を持つ定 Q フィルタバンクを用い、分析時間分解能は 0.1s とした。扱う和音は major、minor の 24 種類とし、正解データのラベルは手動で作成した。

出力確率分布は単一正規分布として、その平均と分散共分散行列は、正解ラベルを付与した別の 4 曲を用い、混合信号用パラメータと調波信号用パラメータをそれぞれ最尤推定した。状態遷移確率は市販のギタースコアから 8 曲を抜粋したデータから学習した。

3.2 結果と考察

評価実験の結果を表 1 に示す。いずれの曲でも、特徴量を調波音を強調したクロマとした場合は従来の

*"Automatic Chord Detection Using Harmonic Sound Emphasized Chroma from Musical Acoustic Signal," by Yuki Uchiyama, Ken-ichi Miyamoto, Takuya Nishimoto, Nobutaka Ono, Shigeki Sagayama, Graduate School of Information Science and Technology, The University of Tokyo.

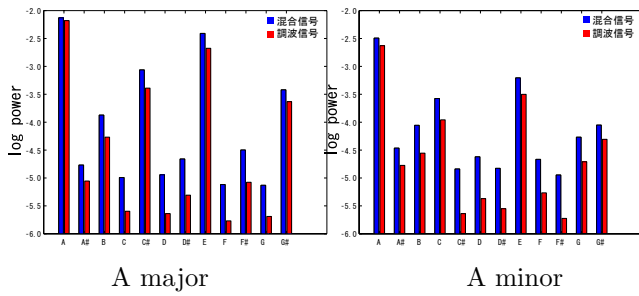


Fig. 1 出力確率分布の平均



Fig. 2 認識結果例 (No. 20)

クロマベクトルに比べて正解率が上がっている。また学習した、出力確率分布の平均を図1に示す。

混合信号に比べ、分離した調波信号はパワーが小さくなるが、和声内音 (A major では A、C#、E) ではその小さくなる程度が小さく、非和声音ではその程度が大きい。その結果、相対的に和声内音が強調されている。

認識結果 (図2) を見ると正解率が上がった主要素は、従来手法では和音の主音は正しいものの major と minor を誤認識した部分を、本手法では正しく認識したことである。major と minor は第3音が異なるだけであるため、認識誤りの原因となりやすい。本手法では、調波クロマベクトルにより和声内音が強調され非和声音が抑制されることによって、major と minor の認識精度向上につながったと考えられる。

4 打楽器音の利用の検討

前節まで調波音を用いる手法を述べたが、本節では打楽器音の利用について検討する。音楽音響信号からの和音認識では、和音名称に加えて和音境界も未知であり、前節のシステムでは和音境界誤りが和音認識誤りの原因の一つとなっていた。この問題を解決するために、和音境界は拍位置に置かれて打楽器音が重畳する確率が高いという仮説のもとに、打楽器音を和音境界の手がかりとして利用することを考える。

前述の調波音・打楽器音分離によって得られた打楽器音のパワーを特徴量の一つに用いることを考える。実際には拍でも打楽器が演奏されなかったり、拍以外でも打楽器が演奏されたりするので、調波音や前後の関係を確率的に考慮する必要があり、図3のような HMM が利用できる。HMM の状態として (C、major、打楽器有り)、(C、major、打楽器無し)、...

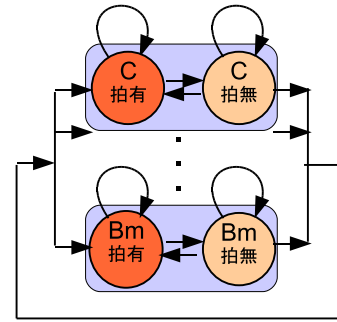


Fig. 3 打楽器音を考慮した HMM

のように和音 24 種と打楽器有り・無しによる合計 48 状態を考える。同一和音の拍有りの状態と拍無しの状態は left-right 型の HMM をなし、和音間ではどの和音にも遷移しうるエルゴディックな HMM とする。状態出力は調波クロマベクトルと打楽器特徴量の 13 次元とし、拍有りの状態では打楽器特徴量が大きく、拍無しの状態では打楽器特徴量が小さいような分布を用いることにより、打楽器有りの場合は拍の可能性が高く、和音が変わりやすいという現象が扱える。今後、実験検証する予定である。

5 おわりに

本稿では、音楽音響信号から自動和音認識の性能向上のために、特徴量に調波音を強調したクロマベクトルを用いる手法を提案し、評価実験により性能が向上することを示した。また、分離により得られた打楽器音を用いて、和音の区切りを捉える手法の検討をした。今後は、和音と打楽器音を用いる手法について評価を行い、更に和音認識性能を向上させる手法を検討する予定である。

謝辞 本研究の一部は、科学技術振興機構 CREST の補助を受けて行われた。

参考文献

- [1] T. Fujishima, "Real-time chord recognition of musical sound: A system using common lisp music," Proc. ICMC, pp. 464-467, 1999.
- [2] A. Sheh *et al.*, "Chord segmentation and recognition using em-trained hidden markov models," Proc. ISMIR, pp. 183-189, 2003.
- [3] 宮本賢一他, "スペクトログラムの滑らかさの異方性に基づいた調波音・打楽器音の分離," 日本音響学会春季研究発表会講演論文集, 2008.
- [4] 川上隆他, "隠れマルコフモデルを用いた旋律への和声付け," 平成 11 年電気関係学会北陸支部大会講演論文集, F-61, pp. 361, 1999.