

HMM を用いた旋律への自動和声付けと調性推定*

◎吉川 響[†] 川上 隆[‡] 中井 満[†] 下平 博[†] 嵯峨山 茂樹[†]
 (†北陸先端大 ‡キヤノン)

1 はじめに

与えられた音楽旋律に対する自動的な和声付けや調性認識処理は、従来の研究では和声学の知識処理や芸術に関する発見的な原理や規則に基づいたアプローチが多かった [1,2,3,4]。これに対し、我々は数理的なモデルに基づく自動和声付けのアプローチとして、連続音声認識の方法論を適用することを試みた [5]。今回はこの手法に関して、主観評価実験の結果を含めて述べる。

2 HMM を用いた旋律和声付けのモデル化

2.1 和声と旋律の関係

西洋音楽では作曲者は旋律の背後に、ある和声進行を想定して旋律を作る場合が多いと考えられる。本研究ではこの仮定に基き、事前に想定されている和声進行 H を、生成された旋律 M から遡って推定する逆問題として和声付けの問題を考える。ある和声進行からの旋律の生成過程を HMM により 2 段階に確率モデル化し、与えられた旋律を生成する遷移系列の中で最も尤度の高い系列を Viterbi 探索によって求める。このモデルでは、HMM の各隠れ状態は和声を持ち、和声は旋律を生成する。

2.2 和声進行のモデル

和声進行 $H = \{h_1, h_2, \dots, h_l\}$ (h_i は i 番目の和声) の確率 $P(H)$ は、和声の連鎖確率を用いて、

$$P(H) = P(h_1, h_2, \dots, h_l) \\ = \prod_{i=1}^l P(h_i | h_1, \dots, h_{i-1}) \quad (1)$$

と表される。あらゆる和声進行 H に関して統計を取り、 $P(H)$ を統計推定することは困難なので、以下の 2 つの方法で近似をする。

- 2 和声間遷移モデル

式 (1) の右辺の条件項の長さを制限し、2 和声間の Markov 遷移確率 (bigram 確率) で

$$P(h_i | h_1, \dots, h_{i-1}) \approx P(h_i | h_{i-1}) \quad (2)$$

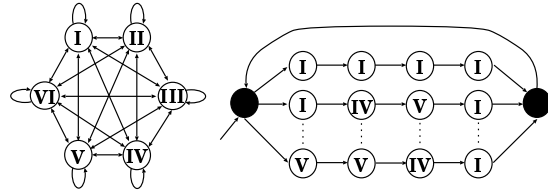
と近似するモデルであり、図 1(a) の通り、全ての状態間の遷移が可能なモデルとなる。

- 常套句接続モデル

短い和声進行の常套句の連鎖として、 $P(H)$ を近似する。これは、 $P(h_1, \dots, h_l)$ における l を一定にした常套句の連鎖で H を近似するモデルとなり、図 1(b) に示すようになる。

以上により、状態に和声を対応させて状態遷移ネットワーク中の状態遷移により和声進行を表現することができる。この状態 i から状態 j への遷移確率を HMM の記法に従って a_{ij} と表す。

* "Hidden Markov model applied to automatic harmonization and tonality estimation of given melodies" by Hibiki Yoshikawa[†], Takashi Kawakami[‡], Mitsuru Nakai[†], Hiroshi Shimodaira[†] and Shigeki Sagayama[†] († JAIST, ‡ Canon Inc.).



(a) 2和声間遷移モデル (b) 常套句接続モデル
 図 1: 2 種類の和声進行モデル

2.3 和声からの旋律生成の確率モデル

和声進行 H から旋律 M が生成される条件つき確率 $P(M|H)$ の統計推定は実際には容易ではないので、単純化したモデルで $P(M|H)$ を与える。

和声進行 $H = \{h_1, h_2, \dots, h_l\}$ から旋律 $M = \{n_1, n_2, \dots, n_T\}$ (n_t は M の t 番目の音符) が生成される確率は条件つき確率により

$$P(M|H) = \prod_{t=1}^T P(n_t | H; n_1, \dots, n_{t-1}) \quad (3)$$

と表せるが、これを単純化し、和声 h_i から単音 n_t が生成される unigram 確率 $P(n_t | h_i)$ で近似する。旋律の各音符は独立に和声 h_i から生成されるとする。音符長に関しては 8 分音符を基準長と考え、基準長の単音の生成確率のべき乗 $\{P(n|h)\}^{\frac{L}{T}}$ を L 分音符の生成確率とし、式 (3) の確率項の近似として用いる。

上記の unigram 確率の単純なモデルとして、ここでは和声 h から生成される音符 n として根音、和声音、非和声音の 3 種類のみを考慮し、その順に生成確率が高いと仮定する。なお、休符は和声に無関係に一定の生成確率を持つものとする。

和声 h から音符 n が生成される確率を HMM の記法に従って $b_h(n)$ と表し、HMM の状態出力確率として用いる。

2.4 旋律を生成する HMM

ある和声進行 H からの旋律 M の生成確率 $P(M|H)P(H)$ は、上記の 2 つの確率の積

$$P(M|H)P(H) = \prod_{t=1}^T a_{q_{t-1}q_t} \cdot b_{q_t}(n_t) \quad (4)$$

で表せる。 t は時刻、 T は総時間、 q_t は時刻 t における隠れ状態であり、状態 q_t の T 個の列が和声進行 H に対応する。

3 HMM による旋律の和声付け

3.1 旋律和声付けアルゴリズム

旋律 M が与えられたとき、それが和声進行 H から生成されたものである確率 $P(H|M)$ は、ベイズの定理により

$$P(H|M) = \frac{P(M|H)P(H)}{P(M)} \quad (5)$$

と表される。(5) 式の $P(M)$ は H によらないから、 $P(M|H)P(H)$ を最大にする H を Viterbi 探索することで最も尤もらしい和声進行を推定できる。

2和声間遷移モデル: G C G D G C F#dim G G
 常套句連接モデル: G C G D G C F#dim G G

図 2: コラールのモデルによる和声付け結果

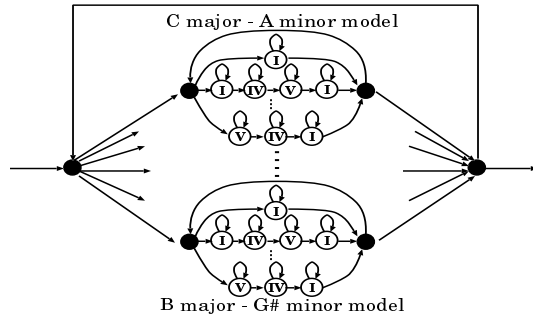


図 3: 転調モデルの概念

HMM の各パラメータは以下に示す意味を持つ。

- 状態 i : 和声 i
- 初期確率 π_i : ある和声 i から曲の始まる確率
- 遷移確率 a_{ij} : 和声 i から和声 j への進行確率
- 出力確率 $b_i(n)$: 和声 i から音符 n の生成される確率
- 入力系列 M : 与えられた旋律
 $M = n_1, n_2, \dots, n_T$

3.2 コラールで学習したモデルによる和声付け

J. S. Bach による単純四声体のコラールの中で 4/4 拍子の曲を対象にして、317 曲から遷移確率を、94 曲から和声常套句を統計的に求めた。P. E. Bach の旋律を入力として和声付け実験を行なった結果例を図 2 に示す。音符の上のコードネームが HMM によって出力された和声進行である。

4 HMM による調性のモデル化

4.1 転調を含む旋律の和声付けと調性認識

調ごとのモデルを作成し、各モデル間 (各調間) に遷移確率 (転調確率) を持つような HMM によって、転調を含む旋律に対する和声付けが可能になる。これを転調モデルと呼ぶ (図 3)。平行長短調は同一調として扱うため、実際にはハ長調 (イ短調)、…、ロ長調 (嬰ト短調) の 12 個のモデルを作成し、これらのモデル間に遷移確率を与えた。

転調モデルを用いると、一定調の旋律に対しては、各調のモデルに旋律を入力し、最も尤度が高いモデルの調を旋律の調性として推定できる。また、転調を含む旋律に対しては、転調モデルを用いて和声付けをした結果から、転調の検出を含めた調性推定ができる。

4.2 HMM による調性認識実験

童謡・民謡 13 曲、Mozart のピアノソナタ 9 曲、バロック音楽等 18 曲の合計 40 曲を入力として、調性認識実験を行なった結果を表 1 に示す。転調検出位置が誤った場合に不正解とする場合を認識率 A、転調の調性認識が正解した場合を認識率 B

表 1: 調性認識実験結果

入力旋律	認識率 A [%]	認識率 B [%]
一定調の旋律 (32 曲)	100.0	100.0
転調含む旋律 (8 曲)	62.5	100.0
合計 (40 曲)	92.5	100.0

表 2: 主観評価実験結果

順位	和声付け方法	平均評価
1	人間	3.08
2	HMM(常套句連接モデル)	3.24
3	人間	3.33
4	HMM(2 和声間遷移モデル)	3.36
5	人間	3.48
6	HMM(常套句連接モデル)	3.50
7	人間	4.06
8	人間	4.47
9	人間	4.52
10	市販ソフト	5.82

とした。一定調の旋律に対しては全ての曲に対して正しく調が推定できた。また、転調を含む旋律では 8 曲とも転調を検出して正しく調性を認識でき、うち 5 曲は転調位置も正しく検出した。

5 主観評価実験

図 2 の旋律に対して、市販ソフトによる和声付け 1 種類、HMM による和声付け 3 種類、人間による和声付け 6 種類をランダムに被験者 22 名 (男性 21 名、女性 1 名) に提示し、それぞれの和声付けの印象を [1. 非常に良い] - [7. 非常に悪い] の 7 個のカテゴリーで判断させ、評定尺度法を用いて評価した。各和声付けの評価の平均と、全体の順位を表 2 に示す。表中、2 種類の常套句連接モデルは、2.3 で示した単音の生成確率 $P(n|h)$ の設定が異なる。HMM の行った各和声付けの結果は概ね良好な評価が得られた。

6 まとめ

本研究では、連続音声認識の方法論を適用して、HMM によって入力旋律に自動で和声付けをする方法を提案し、和声付け実験の結果を示した。また、転調検出も同じ枠組で扱えることを示した。主観評価実験では、人間や市販ソフトによる和声付けにも劣らない評価を得た。今後の課題としては、旋律生成確率、和声進行の確率の近似・推定方法などに関して、モデルの精密化には大きな研究の余地がある。また、音楽に関するより高次の知識を組み合わせることも有効であると考えられる。

参考文献

- [1] 長嶋 洋一, 橋本周司, 平賀 譲, 平田 圭二: コンピュータと音楽の世界, bit 別冊, 共立出版株式会社, 1998.
- [2] 中西 正和: 計算機による作曲と編曲, 情報処理, Vol. 29, No. 6, pp. 608-612, 1988.
- [3] 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, 94-MUS-8, pp. 45-50, 1994.
- [4] 平井 重行, 金森 務, 平井 宏: ジャズの伴奏からの調性を含めた感情情報抽出, 情報処理学会研究報告, 94-MUS-8, pp. 1-6, 1994.
- [5] 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-2, pp. 59-66, 1999.