

HMM と音符連鎖確率を用いた旋律への自動和声付け*

菅原啓太 米田隆一 西本卓也 嵯峨山茂樹 (東大・情報理工)

1 はじめに

本報告では、与えられた任意の旋律に対する自動和声付けの問題を扱う。ここで扱う自動和声付けとは任意の旋律を入力としたときにその旋律に適合する和声を自動的に出力することである。自動和声付けの応用分野としては、専門的な知識や経験がない人のための作曲・編曲ツール、さらには自動伴奏、作曲家や時代やジャンル別の和声付けの数理的モデル化などが考えられる。

過去に報告されている、発見的な原理や規則に基づいた和声付け手法は、楽曲のスタイルごとに規則を変えなければならなかった [1]。一方、学習に基づく手法では、学習データを替えるだけで様々なスタイルの楽曲に適した和声付けが行え、また、作曲家の経験を反映した和声付けも行えたと考えられる。われわれはこのような展望のもとに、HMM (Hidden Markov Model, 隠れマルコフモデル) を用いた和声付け手法を提案した [2]。本報告では特に和声と複数の音符の連鎖の関係を考慮し、経過音・刺繍音・倚音・掛留音などの使用が許される和声外音 (以降装飾音と呼ぶ) を区別することによってより旋律に適合する和声付けを目指す。また、和声進行のモデルを改良することでより自然な和声進行の和声付けを試みる。

2 HMM を用いた和声付け

2.1 着眼点

以下次のような仮定を立て、それに基づいた自動和声付けの手法を導く。

1. 作曲者が想定する和声進行は確率的にモデル化できる
2. 旋律は想定された和声進行から確率的に生成されるとモデル化できる
3. 旋律への和声付け問題は与えられた旋律を生成した最も尤もらしい和声進行を推定する逆問題として解ける
4. 実際には和声進行が想定されていない旋律に対してもこの逆問題は適用できる

上記の仮定 1,2 は数理モデルとしての仮説であり、作曲の実際がその通りに行われる確証はないが、たとえば A. Schönberg の教科書 [3] には楽曲の作曲のための和声の設計などが述べられており、和声を設計してその中で旋律を作曲する過程が示唆されている。ある和声進行から旋律の生成される過程を、和声進行モデルと旋律生成モデルと呼ぶ 2 種類の確率モデルを統合した HMM で表現し (図 1)、和声付け問題を HMM における探索問題として解く。

2.2 定式化

和声進行 (和声系列) $H = (h(1), \dots, h(L))$ が生成される確率 $P(H)$ を、これらすべてが生起する同時確率として

$$P(H) = P(h(1), \dots, h(L)) \\ = \prod_{i=1}^L P(h(i) | h(1), \dots, h(i-1)) \quad (1)$$

と表現する。ここで $h(i)$ は和声進行 H の i 番目の和声 (chord) を表す。条件付き確率 $P(h(i) | h(1), \dots, h(i-1))$ の値を、有限長の n -gram モデルでモデル化する。これにより、既存の楽曲の和声進行の統計から学習推定することができる。また和声進行 $H = (h(1), \dots, h(L))$ から、旋律 $M = (n(1), \dots, n(K))$ が生成される確率 $P(M|H)$ を条件付き確率

$$P(M|H) = P(n(1), \dots, n(K)|H) \quad (2)$$

により表現する。ここで、 $n(i)$ は旋律 M の i 番目の音 (音高や音長などの情報を含む) を表す。旋律は各區間ごとに対応する和声があると考えられる。そこで、旋律はその背後の和声の影響のみを受けて確率的に生成されると仮定する。つ

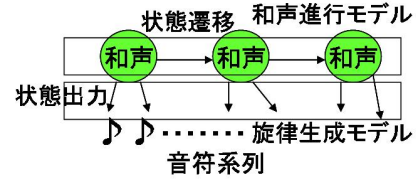


図 1: 和声進行から旋律を生成する HMM の概念図

まり、ある和声進行 H から旋律 M が生成される確率を、旋律の断片が、対応する和声進行の断片から生成される確率の積でモデル化する。

上記の二つのモデルは HMM として一つのモデルに統合することができる。旋律のみが与えられたとき、旋律の背後にある和声は直接には観測できず、マルコフ過程に従って遷移する。これが HMM の状態遷移に対応する。また、旋律は背後の和声から確率的に生成される。これが HMM の状態出力に対応する。そして与えられた旋律 M に対する適切な和声 H を推定する問題を、旋律 M を観測したときの事後確率を最大化する (maximum a posteriori) 和声 \hat{H} を求める問題とする。即ち、

$$\hat{H} = \operatorname{argmax}_H P(H|M) \quad (3)$$

と定式化する。ここで Bayes の定理

$$P(H|M) = \frac{P(M|H)P(H)}{P(M)} \quad (4)$$

より、式 (3) は

$$\hat{H} = \operatorname{argmax}_H P(M|H)P(H) \quad (5)$$

と表される。最も尤もらしい和声進行 \hat{H} は HMM の状態ネットワークの最適状態遷移系列として、Viterbi 探索を用いて効率的に求めることができる。このようにして、和声付け問題を旋律 M を生成したと考えられる和声進行 H を推定する逆問題として考える。

3 和声進行モデル

3.1 和声常套句間の bigram 確率の近似

本報告では楽曲の和声は短い和声常套句の連鎖であると考え、和声進行モデルを採用する。ここでは常套句として、1 つの和声で 2 拍に対応した 2 小節長のものを使う。使用した和声は 12 音階それぞれを主音とする 3 音和音 (長三和音・短三和音・増三和音・減三和音) 48 (12×4) 種類とし、今回は七の和音や変化和音は除外する。豊富な和声常套句を得るために、バッハの四声のコラール (BWV 1-438 に含まれるカンタータ、モテット、受難曲、オラトリオ、および C.Ph.E.Bach 編纂のコラール集などに含まれる) のうち、4 分の 4 拍子で冒頭に弱起をもたない楽曲 94 曲から学習した。

常套句を用いることで、局所的には楽曲らしい自然な和声進行が実現できると考えられる。しかし、従来の常套句連鎖モデル [2] では常套句間に遷移則を与えていないため、常套句間の不自然な接続が現れることもあった。そこで本報告では常套句間に遷移確率を与えることを試みる。常套句間の bigram 確率を統計学習して用いるためには多量の学習サンプルを必要とする。そこで常套句間の bigram 確率を近似するために、先行する常套句の最後の和声と後続する常套句の最初の和声との 2 和声間遷移確率を用いる (図 2)。ある和声進行 $H = (H'_1, \dots, H'_T)$ が出現する確率 $P(H)$ をある和声常套句 H'_i が出現する確率 $P(H'_i)$ と常套句間の遷移確率 $a_{H'_i, H'_{i+1}}$ を用いて次のように近似する。

$$P(H) = P(H'_1, \dots, H'_T) \approx \prod_{i=1}^T P(H'_i) a_{H'_i, H'_{i+1}} \quad (6)$$

3.2 和声常套句の出現確率値の補間

常套句数の不足による旋律との不適合を避けるために常套句の出現確率値を近似することで出現確率値が 0 でない常套句数を増やす。和声常套句 $H'(h(1), h(2), h(3), h(4))$ の出現確率値の近似には以下のような線形和を用いる。

* "Automatic Harmonization for Melodies based on HMMs Including Note-chain Probability" by Keita Sugawara, Ryuiti Yoneda, Takuya Nishimoto and Shigeki Sagayama (Graduate School of Information Science and Technology, The University of Tokyo).

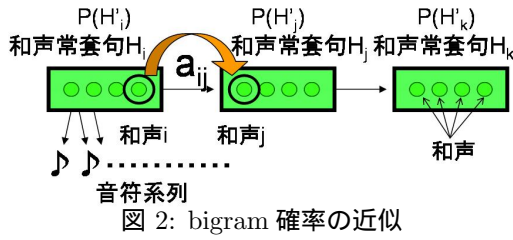


図 2: bigram 確率の近似

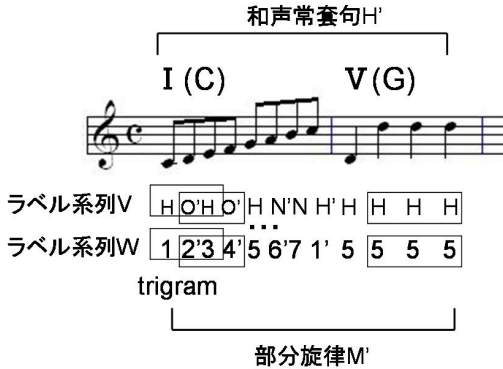


図 3: 音種連鎖・音高連鎖モデルでのラベル付け例

$$P(h(1), h(2), h(3), h(4)) \approx \alpha P(h(1), h(2), h(3), h(4)) + \beta P(h(1), h(2), h(3)) + \gamma P(h(2), h(3), h(4)) \quad (7)$$

ただし補間係数は $\alpha : \beta : \gamma = 48 : 1 : 1$ の条件の下で確率値を正規化して求める。これにより長調の楽曲から学習した 908 種類の常套句からは 4195 種類の確率値 0 でない常套句を、短調の楽曲から学習した 954 種類の常套句からは 4937 種類の確率値 0 でない常套句を得た。

4 旋律生成モデル

4.1 音種連鎖モデル

本報告では和声と単音の関係をモデル化する [2] のではなく、複数の音符の連鎖との関係をモデル化する。音種連鎖モデルは旋律を和声に対する音符の役割で分類し、それらの trigram 確率を用いて旋律生成確率を定義する。従来法で和声外音として扱っていた音符のうち、経過音・刺繍音・倚音・掛留音などの和声学的に使用が許されている音符に関しては区別して扱う。また、拍の強弱と音符の種類の関係も考慮するために、同じ役割の音符であっても表拍と裏拍では区別して扱う。

旋律生成確率の学習は以下のような手順で行う。

1. 学習したい和声を与えられた旋律の各音符に対して、和声との対応に応じてラベルを付与する。ここでは和声内音 (H)、装飾音 (O)、それ以外の和声外音 (N) の 3 種類のラベルを用いる。
2. 基本音符長の表拍・裏拍を区別する情報を付加する。(今回は表拍=強拍、裏拍=弱拍とする)
3. このようにして得られた 6 種類のラベル (音種ラベル) からなる系列の trigram 確率を学習する。

このようにして得られた trigram 確率を用いて以下のように旋律生成確率を求める。

1. 与えられた旋律に対して候補となる和声常套句を当てはめ、旋律の各音符をその拍の和声に対してラベル付ける。図 3 に例を示す。図中 ' が付されていないラベルは強拍を表し、付されているラベルは弱拍を表す。
2. 求めたラベル系列の trigram 確率の積をその常套句からの旋律生成確率とする。

求めたラベル系列を $V_i = (v_i(1), \dots, v_i(K_i))$ とすると、式 (5)、式 (6) より \hat{H} は次式ようになる。

$$\hat{H} = \operatorname{argmax}_{H'_1, H'_2, \dots, H'_T} \prod_{i=1}^T \left[\prod_{j=1}^{K_i} P(H'_i) a_{H_i, H_{i+1}} \cdot P(v_i(j) | v_i(j-2), v_i(j-1)) \right] \quad (8)$$

4.2 音高連鎖モデル

音高連鎖モデルでは、旋律の音符と和声の根音との音程を用いて旋律を分類し、それらの trigram 確率を用いて旋

I (Am) VII (G#dim) I (Am) II (Bdim) I (Am) I (Am) I (Am) I (Am)
 I (Am) VII (G#dim) I (Am) VII (G#dim) I (Am) IV (Dm) IV (Dm) V (E)
 I (Am) VII (G#dim) I (Am) VII (G#dim) I (Am) II (Bdim) VI (F) V (E)

I (Am) VII (G#dim) I (Am) II (Bdim) I (Am) I (Am) I (Am) I (Am)
 I (Am) VII (G#dim) I (Am) VII (G#dim) I (Am) IV (Dm) V (E) I (Am)
 I (Am) VII (G#dim) I (Am) VII (G#dim) I (Am) IV (Dm) V (E) I (Am)

図 4: 和声付け結果の例: 上段から川上らの常套句連続モデル・音種連鎖モデル・音高連鎖モデルの順 (旋律: Sholom Secunda 作曲 “ドナドナ” より 第 1~8 小節)

律生成確率を定義する。このラベル付けは音種連鎖モデルのラベル付けに比べヒューリスティックでないラベル付けである。音種連鎖モデルと同様に得られた (図 3) ラベル系列を $W_i = (w_i(1), \dots, w_i(K_i))$ とすると、 \hat{H} は次式のようになる。

$$\hat{H} = \operatorname{argmax}_{H'_1, H'_2, \dots, H'_T} \prod_{i=1}^T \left[\prod_{j=1}^{K_i} P(H'_i) a_{H_i, H_{i+1}} \cdot P(w_i(j) | w_i(j-2), w_i(j-1)) \right] \quad (9)$$

5 実験

以下のような実験条件において、和声付け実験を行った。

- 使用和声としては 3 音和音 48 種類を用いる。
- 使用した和声常套句は前述のように得た 4195・4937 種類を用い、2 拍で 1 和声とする。
- 和声常套句間の遷移確率に用いた 2 和声間遷移確率も常套句と同様にバッハのコラールより学習する。
- 旋律生成確率は童謡・唱歌・民謡など 60 曲から学習する。

また、次の 3 種類の手法の結果を比較した。

1. 川上の常套句連続モデル [2]
2. 音種連鎖モデル (提案手法 A): 式 (8)
3. 音高連鎖モデル (提案手法 B): 式 (9)

和声付け結果例を図 4 に示す。楽譜の上に 2 拍ごとに記した和声は、上段から、従来法・音種連鎖モデル・音高連鎖モデルの結果を示す。3、7 小節目の 3、4 拍目の和声では、従来法では第 2 音を和声内音とする和声付けがなされているが、提案手法 A・B では第 1 音と第 3 音を和声内音とする和声付けがなされている。これは第 2 音を経過音と捉え、和声付けがされたものであると考えられる。音高連鎖モデルでも音種連鎖モデルと同様に経過音を判別できていることから、音高連鎖確率は経過音などの装飾音とそれ以外の和声外音との出現頻度の違いを内包していると考えられる。

また 2、6 小節目の 3、4 拍目の和声と 3、7 小節目の 1、2 拍目の和声間では、従来法ではこの 2 和声間が不自然な和声進行になっているが、提案手法 A・B では多少自然な和声進行になっていることが確かめられた。これらの結果より、期待されていた装飾音への対応と和声常套句間の自然な接続が確かめられた。

6 おわりに

本報告では、音種連鎖確率と常套句間の bigram 確率の近似を用いて和声付けを行った。音符連鎖確率を用いることで、装飾音と他の和声外音の区別、拍の強さと現れやすい音符との関係を反映した和声付けが可能となり、和声常套句間遷移確率を与えることで常套句間での和声進行が多少自然になることが確認された。しかしながら和声付け結果はまだ十分とはいえない。今後の課題としては、旋律への適合性という点では今回は考慮していなかった音符長の情報の使用、楽曲らしい自然な和声進行の実現という点では各確率値の比率の最適化、常套句間の接続文法の作成などが挙げられる。

参考文献

- [1] 長嶋洋一他: “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] 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.” 平成 11 年情報処理学会音楽情報科学研究会, 99-MUS-34, pp.59-66, Feb 2000.
- [3] A.Schönberg 著, 中村太郎訳: “作曲法入門”, カワイ楽譜株式会社, 1966.