

隠れマルコフモデルを用いた旋律への自動和声付け

川上 隆 中井 満 下平 博 嵯峨山 茂樹
北陸先端科学技術大学院大学 情報科学研究科
〒923-1292 石川県能美郡辰口町旭台 1-1
URL: <http://www-ks.jaist.ac.jp/index-j.html>

あらまし 本稿では、隠れマルコフモデル (HMM) を用いて与えられた旋律に自動で和声付けを行なう手法を提案し、実験的検討結果を報告する。旋律は背後に隠れた和声進行から生成される、とする旋律生成の隠れマルコフモデルを提唱し、その逆問題として、与えられた旋律から背後の和声進行を最尤推定する。2種類の和声進行のモデル、いくつかの旋律生成のモデル、さらに N -best アルゴリズムによる和声付け複数候補抽出や、与えられた旋律の調性推定及び転調検出についても論じる。童謡や歌曲及びバッハのコラールから学習した和声進行確率モデルを用いた、実際に和声付け実験を行なった結果についても述べる。

キーワード ● 隠れマルコフモデル ● 和声付け ● 調性認識

Hidden Markov Model Applied to Automatic Harmonization of Given Melodies

Takashi Kawakami Mitsuru Nakai
Hiroshi Shimodaira Shigeki Sagayama
Japan Advanced Institute of Science and Technology
1-1 Asahi-dai, Tatsu-no-kuchi, Ishikawa 923-1292
URL: <http://www-ks.jaist.ac.jp/>

Abstract This paper proposes the use of Hidden Markov Model (HMM) for automatic music harmonization of given melodies and describes experimental results. The process of producing a melody is modeled by a hidden Markov model consisting of harmony transitions and emission of musical notes from harmonies. The underlying harmony is estimated as an inverse problem solved by a maximum likelihood estimation from the given melody. Harmony transition models, melody emission models, N -best candidate generation, tonality estimation and detection are also discussed. Experimental studies were made based on statistical analysis of harmony transitions in tunes and songs as well as in more than 300 chorales by J. S. Bach.

Key words ● Hidden Markov Model ● automatic harmonization ● tonality estimation

1 まえがき

与えられた音楽旋律に対して自動的に適切な和声付けを行なうことは、実用的にも技術的にも興味深い問題である。

実用的な面では、音楽の愛好者は多く、旋律の作曲には多くの人が興味を持つが、それに適切な和声付けをし編曲するにはやや専門的な音楽の知識や経験が必要なので、旋律に自動和声付けをして編曲ができる支援ツールには広い用途がある。作曲・編曲支援ツールとしては、楽曲データを入力すると、自動演奏や楽譜を清書する研究がされており、市販されているツールも多数ある [1]。

一方、技術的な面では、旋律への和声付けの問題

は和声学の知識処理や芸術に関連する知的な処理と考えられ、この種の問題に対してどのようなアプローチが可能であるか、という点で興味深い。従来の方法としては、1小節等の基本単位を設け、基本単位中にある旋律の音符が最も多く含まれる和音を判定し、曖昧さをなくすために終始形などの和声学の知識を用いる方法が広く行なわれている [2, 3, 4, 5]。また、入力された和音をジャズの和音らしく編曲し直すような研究も報告されている [6]。

また、和声付けには旋律の調性が深く関係すると考えられるが、調性認識に関してはある基本単位中の音符がどの調の構成音に最も多く含まれるかを判定する方法 [7] や、なるべく全音階的な枠組にあてはまる調を探す方法 [8] が提案されている。

表 1: 音声認識と自動和声付けの概念対応

	連続音声認識	自動和声付け
入力単位	文音声	楽曲
語彙	単語	和声常套句
隠れ状態	音響イベント	和声
観測値	スペクトル系列	音符系列

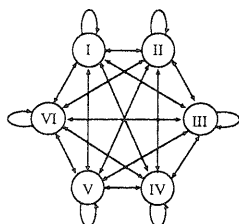


図 1: 2 和声間遷移モデルの例

これらの従来の研究では、発見的な原理や規則に基づいたアプローチが多かった。これに対し、本稿では数理的なモデルに基づく和声付けのアプローチを提案する。具体的には、連続音声認識の分野の方法論を旋律への自動和声付けの問題への適用を試みるものである。また、調性認識の定式化についても述べる。

この目的は、この種の問題に対し、確率計算に基づいた問題の定式化、統計的な学習、 N -best 解探索などの枠組を与えることである。このような数理的枠組と尤度 (あるいは事後確率) 最大化の原理で、複雑な問題を統一的な視点でどれほど扱えるかを解明することも、大きな関心の対象である。さらに、将来、楽曲や作曲家のスタイルを数理的にモデル化する足がかりとしたい。

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 (特に古典派から近代まで) においては、旋律の作曲時点と同時に、あるいは事前の和声設計 [9] の段階で、作曲家はその背後の和音を想定していると仮定できる場合が多い。分散和音形の旋律の場合はその典型である。

そこで本稿では、「旋律のみが与えられた場合、その背後に隠された和声進行を推定する問題」という発想で和声付けの問題を考える。すなわち、与えられた旋律を処理して適合する和声を決めるのではなく、その旋律を生成したと考えられる (事前に存在した) 和声を、遡って推定する逆問題と考える。これは、連続音声認識で成功しているモデル化

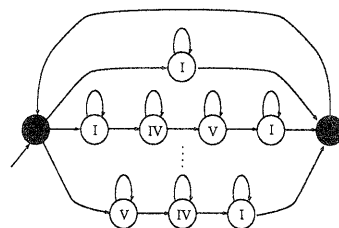


図 2: 常套句接続モデルの例

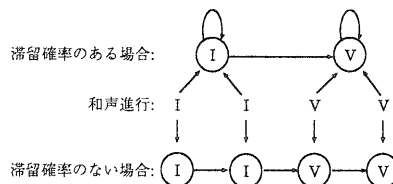


図 3: 滞留確率の有無の概念

であり、どのような単語を意図して発話すると入力音声のような観測が得られるかという問題定式化を行ない、発声の確率変動のモデルにより逆問題を解くことが、連続音声認識の問題の解法の主流になっている。その確率モデルには、隠れマルコフモデル (Hidden Markov Model, HMM) が多く使われている。

ここでは、旋律の和声付けの問題も HMM を用いて定式化する。ある和声進行から旋律の生成される過程を HMM を用いて 2 段階の確率モデル化し、尤度最大の原理によって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。和声付けの問題と音声認識は表 1 に示すように概念を対応付けることができる。

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

まず、和声進行 $H = \{h_1, h_2, \dots, h_l\}$ の確率 $P(H)$ を考える。 h_i は i 番目の和声である。 $P(H)$ は、(旋律に関係なく) 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。

この確率は楽曲の和声構造に関係し、理念的には作曲家のセンスや楽曲のスタイルなどを反映する。十分な量の同一スタイルの学習サンプルが得られれば $P(H)$ を統計推定ができると考えられが、あらゆる和声列 H に関して統計を取ることは困難である。

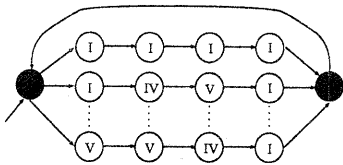


図 4: 滞留確率なしの常套句接続モデルの例

表 2: 楽曲データより学習した出力確率

音符の種類	確率
根音 (r)	0.32
和声音 (h)	0.28
非和声音 (n)	0.13

そこで、第一の方法として、右辺の条件項の長さを制限していわゆる N -gram 確率によって近似する。本稿では、 $N = 2$ の場合の 2 つの和声間の Markov 遷移確率 (bigram 確率) を用いる。第二の方法としては、短い和声進行 H の常套句の連鎖として、 $P(H)$ を近似する。これらはそれぞれ、音声認識の場合の音素連鎖 bigram 確率を用いた音声タイプライタ、および subword-based の連続単語音声認識に関連づけられる。

● 2 和声間遷移のモデル

2 つの和声間の確率 (2-gram 確率) モデルは、

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

と近似するモデルであり、図 1 に示すように、和声に対応付けられた状態 i から他の任意の状態 j に確率 $a_{i,j}$ で進行するモデルとなる。以降、この確率モデルを 2 和声間遷移モデルと呼ぶ。

● 和声進行常套句のモデル

和声進行の常套句の連鎖で曲が成立しているとすると、 $P(h_1, \dots, h_l)$ 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 2 に示すようになる。以降、この確率モデルを常套句接続モデルと呼ぶ。

また、常套句接続モデルにおいて各状態の滞留確率を 0 とするモデルも考える。このモデルは、和声の長さが確率的に変化するのではなく、モデルに組み込まれる。例えばある 2 小節の和声進行が $I \rightarrow I \rightarrow V \rightarrow V$ である場合、この和声常套句を表すモデルは、滞留確率が 0 でない場合には、図 3 上に示すように、滞留確率を持つ I と V の 2 状態で表され、滞留確率 0 の場合には同図下に示すように、 I の状態と V の状態それぞれ 2 個ずつ、合計 4 状態で表

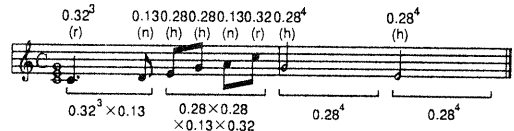


図 5: 旋律の生成される確率の計算例

されることになる。以降この確率モデルを滞留確率 0 の常套句接続モデルと呼ぶ。滞留確率 0 の常套句接続モデルの例を図 4 に示す。

以上により、状態に和声に対応させて (状態数が和声の種類より多い場合があり得る)、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} と書く。

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 $P(M|H)$ を定義する。この確率は、和声と旋律の関係を記述するもので、理想的には作曲家のセンスや楽曲のスタイルなどを反映する。統計的立場からは、十分な量の同一スタイルの楽曲のサンプルが得られれば $P(M|H)$ を統計推定できると考えられるが、実際には容易ではない。本稿は HMM を用いた和声付けの手法の提案自体を目的としているので、単純化したモデルにより $P(m|h)$ を与える。

旋律 m 中の第 i 番目の音符を n_i により表すと、和声列 $H = \{h_1, h_2, \dots, h_l\}$ から旋律 $M = \{n_1, n_2, \dots, n_k\}$ が生成される確率は条件つき確率により

$$P(M|H) = \prod_{i=1}^k P(n_i|H; n_1, \dots, n_{i-1}) \quad (3)$$

と書ける。多重の条件の場合の右辺の確率値の統計推定を行うのは容易でないので、たとえば trigram 確率値 $P(n_i|H; n_{i-2}, n_{i-1})$ により近似することが考えられる。

本稿ではさらに単純化して、和声 h_j から単音 n_i が生成される確率 (unigram 確率) $P(n_i|h_j)$ で近似する。旋律の各音符は独立に和声 h_j から生成されるとする。音声認識では、隣接フレーム間の相関を無視することが一般的であり、この近似はそのことに対応する。音符の長さに関しては、8 分音符を基準長にして音符の長さ L 分、単音の生成される確率のべき乗 $\{P(n|h)\}^L$ を音符の生成確率とし、式 3 の確率項の近似として用いることにする。

さらに、上記の unigram 確率の最も単純なモデルは、和声 h から生成される個々の音符 n の種類として、根音、和声音、非和声音の 3 種類のみを考え、その順に生成される確率が高いと考えるもので

ある。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。

和声進行 $H = \{h_1, h_2, \dots, h_l\}$ の可能な組合せは膨大になるので、ここでは基本和声長(一つの和声から1度に生成される旋律の単位)として、以下の2種類のモデルについて考える。

- 音符単位モデル: 1つの和声から音符が1つ生成されるとするモデル
- 2拍単位モデル: 1つの和声から2拍分の旋律がまとめて生成されるとするモデル

以上のように定義される和声 h から個々の音符 n が生成される確率をHMMの記法に従って $b_h(n)$ と書き、(状態 h)HMMの出力確率として用いる。

童謡・民謡・歌曲を対象に、62曲を用いて、これらの確率を学習した。結果を表2に示す。

2.4 旋律を生成するHMM

以上のような2つの確率モデルにより、ある和声進行 H から旋律 M が生成される確率を表現できる。旋律 M の生成確率 $P(M|H)$ は上記の2つの確率の積

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

で表せる。ただし、 t は(たとえば8分音符を単位とする)時間、 T は総時間、 q_t は時刻 t における隠れ状態であり、いずれかの和声に対応している状態 q_t の T 個の列は、和声列 H に対応する。

図5に旋律が生成される確率の計算例を示す。

この確率モデルは、既に述べたように極めて単純化したものである。これを高度化するには、和声音の種類(根音かそれ以外か)、非和声音の種類(経過音、補助音、繋留音など)、それらの小節内位置、拍との関係、旋律音形、楽曲スタイル、表現意図など、詳細にモデル化すべき要素は他に多くある。これらは、今後の研究課題である。

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

与えられた旋律への和声付けの問題を、旋律 M を生成した和声進行 H を推定する逆問題と考える。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、Bayesの定理により

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

と表される。この事後確率を最大にする H (および N -best の H) を探索することで、旋律への和声付

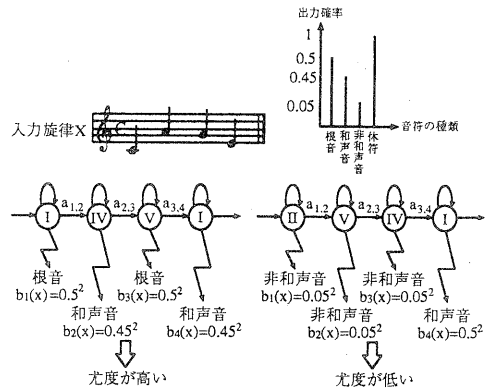


図6: HMMによる和声付けの概念

けを行う。これは、上式中の $P(M)$ は H によらないから、 $P(M|H)P(H)$ を最大にする H を Viterbi 探索することにより行える。

HMMにおいて Viterbi 経路を求めれば、最も尤もらしい和声進行を推定することができる。HMMの各パラメータは以下のような意味を持つ。

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

図6に示すように、旋律 $M = \{\text{ド, ラ, ソ, ファ}\}$ が入力された時、この旋律 M を生成する確率の最も大きい和声列 H を、隠れ状態列の Viterbi 経路探索により求めることができる。

3.2 童謡・民謡・歌曲データからのHMM学習と和声付け実験

童謡・民謡・歌曲を学習サンプルとして用い、2和声間遷移確率及び、和声常套句の統計を求めた。基本和声長が2拍のモデルには合計60曲を、基本和声長が音符のモデルには合計86曲を用いた。扱う和音は3和音のみとし、ここでは7th・6thの和音は3和音にして用いた(例: $C_7 \rightarrow C$)。

和声常套句の統計を求める際、実際の曲の中でどこからどこまでが1つの常套句であるのかを判断するのは難しいため、1つの常套句は2小節であると仮定して統計を求め、滞留確率が0でない場合は、基本和声長が2拍のときで157種類、基本和声長が音符のときで243種類の和声常套句が得られた。また、滞留確率を0とする場合は169種類の和声常套句が得られた。

表 3: 童謡・民謡・歌曲から学習したモデル

	和声進行モデル	基本和声長
モデル 1	2 和声間遷移モデル	2 拍
モデル 2	常套句連接モデル	2 拍
モデル 3	常套句連接モデル (滞留確率なし)	2 拍
モデル 4	2 和声間遷移モデル	音符
モデル 5	常套句連接モデル	音符

図 7: 中田 喜直「夏の思い出」の楽譜

音声認識の場合のモデル学習は、Baum-Welch アルゴリズムなどを用いて、状態の区分化を行わずに尤度最大化の原理によりモデルパラメータを同時に推定することが多いが、ここでは、状態遷移確率と状態出力確率を上記述べたように別々に求めている。これは、音声の場合は発話内容が与えられても時間対応が一意に定義できないのに対し、旋律と和声の場合は、サンプル楽曲において時間対応付けが明確であるためである。

和声進行モデルと基本和声長の組合せで表 3 に示す 5 種類のモデルを作成し、中田 喜直「夏の思い出」(図 7) の旋律を入力として和声付け実験を行なった。結果を図 8 に示す。

2 和声間遷移モデル・常套句連接モデルともに、基本和声長が音符単位の場合の和声付け結果は、モデル 5 の第 6 小節・モデル 4 の第 2・6 小節などに見られるように旋律の変化に伴って和声に変化しやすい反面、第 4・8・16 小節のように 1 拍で和音の変化する和声付けをすることもできている。

3.3 非和声音の種類を考慮した和声付け

和声学では、非和声音の中でも経過音・掛留音・先行音などは使用が許されている(但しこれらの呼称は書籍により異なる [10, 11, 12, 13])。そこでそれらの生成確率を、説明できない非和声音より高く評価する手法が考えられる。和声 H から音符 n が生成される時 n の条件

図 8: 中田 喜直「夏の思い出」の和声付け結果

表 4: 経過音・掛留音・先行音を考慮した出力確率

音符の種類	確率
根音	0.32(1 種類)
和声音	0.28(2 種類)
経過音・補助音・掛留音・先行音	0.012(9 種類)
非和声音	0.002(9 種類)

- 経過音の条件: 前後に隣接する音符がともに和声音である
- 掛留音の条件: 現在のフレームの最初の音符であり前の和音の和声音である
- 先行音の条件: 現在のフレームの最後の音符であり、次の和音の和声音である

により、確率 $P(n|H)$ を高めに補正することにする。これらの和声学で説明のつく非和声音をその他の非和声音と区別して出力確率の学習を行なった。結果を表 4 に示す。図 5 の例では、2 拍目のレの音や 4 拍目のラの音は和声音には含まれた非和声音のため経過音と判定し、確率を高く与える。

この方法で図 7 の旋律をを入力として和声付け実験を行なった。結果を図 9 に示す。用いたモデルは表 3 と同一である。

経過音・掛留音・先行音を考慮することで、モデル 1 の結果には 12、15、16 小節目に、モデル 5 の結果には、13、14 小節目に変化が見られる。しかし、それほど大きな変化であるとはいえない。また、モデル 2、モデル 3、モデル 4 の結果は、経過音等を考慮しない場合と全く同じであった。

モデル1:
C G F C F C

モデル5:
C G C F C F C G C

図 9: 経過音・掛留音・先行音を考慮した中田喜直「夏の思い出」の和声付け結果

表 5: バッハのコラールから学習したモデルの種類

	和声進行モデル	和声常套句長
モデル 1	2 和声間遷移モデル	-
モデル 2	常套句連接モデル	1 小節
モデル 3	常套句連接モデル	2 小節

経過音が2つ続く場合や、装飾のある掛留音などについても対処することができれば、基本和声長が音符の場合には、全体的に和音の変化が少なくなり、旋律の変化にともなう和音の不自然な変化をなくすることができるかと期待される。

3.4 バッハのコラールから学習したモデルによる和声付け

和声間遷移確率・和声常套句の学習に、同一スタイルの楽曲データを大量に用いれば、そのスタイルを反映した和声付け結果が得られることが期待される。そこで、J. S. Bach による単純四声体のコラール (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) の中で 4/4 拍子の曲を対象にして、317 曲から遷移確率の、94 曲から和声常套句の統計を求めた。バッハのコラールは基本的には 1 拍に 1 つの和声を持つスタイルであるため、1 拍に 1 つの和音が対応するものとして和声づけを行った。和声常套句モデルでは状態滞留確率を 0 とし、和声常套句の長さを 2 小節とする場合と 1 小節とする場合について統計を求めた。ただし、楽曲データから和声を決めることは必ずしも単純ではないので、和声の統計の精度には限界がある。

表 5 に示すモデルを用い、図 10 の旋律を入力として和声付け実験を行なった。L. v. Beethoven による旋律への和声付けの結果例を図 10 に示す。

童謡・民謡から学習したモデルでの和声付けでは

モデル2: C Am F G C C Bdim C Dm G Am C Dm E C G G

モデル3: C Am F G C C Bdim C Dm G Am C Dm E C G G

図 10: L. Beethoven: 交響曲第 9 番 第 4 楽章主題「歓喜の歌」と Bach のコラールのモデルによる和声付け結果

第1候補:
C G F C F C

第200候補:
C G F C F C

第400候補:
C G F C F C

図 11: N-best モデルを用いた和声付け結果

全く見られなかった減 5 度の和音が多く現れている。たとえば、vii° の和音である B_{dim} が使われるのは Bach の曲から学習したモデルの特徴が現れていると考えられる。このように、特定のスタイルの楽曲からモデルを学習することにより、そのスタイルの特性を持つ和声付けの可能性が示唆される。

3.5 和声付け N-best 候補の生成

和声付けのツールとしては、和声付け候補を複数提供し、使用者が好みの和声進行を選択することができることが望ましい。そこで、N-best アルゴリズム [14] を用いて、尤度最大の和声進行から、N 番目に尤度の高い和声進行まで求める。このモデルを N-best モデルと呼ぶことにする。

N-best モデルを用いて図 7 の旋律を入力として 400 番目の候補までの和声付け実験を行った。童謡・民謡から学習をした 2 和声間遷移モデルを用い、基本和声長は 2 拍とした。結果を図 11 に示す。

少数の候補の間には大きな相違が見られないことが多いため、複数の和声付け候補を提供する目的のためには多数 (例えば 100 個以上) の候補が必要で

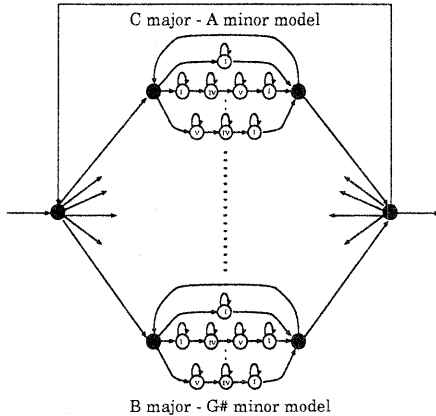


図 12: 転調モデルの概念

ある。その中から好みの候補を見つける作業は容易でないため、多数の候補の中から互いの差が大きい少数 (10 個程度) を選択する方法が今後の課題の一つである。

また、生成した多数の和声付け結果に対して、より高次の和声づけ評価値を与えて再評価することができる。これは、連続音声認識におけるマルチパス方式のアプローチに対応するものである。

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

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) を持つような大きな HMM を用いることで、転調を含む旋律に対する和声付けが可能になる。このモデルを転調モデルと呼ぶことにする。その概念を図 12 に示す。調は 24 種類存在するが、長調と短調の区別はせず、平行長短調は同一調として扱うことにしたため、実際にはハ長調 (イ短調)、変ニ長調 (変ロ短調)、…、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。

4.2 転調モデルを用いた和声付け

転調モデルを用いて J.S.Bach: 無伴奏 Violin Partita II の冒頭 16 小節 (図 13) の旋律を入力として和声付け実験を行なった。転調確率としては、転調する確率を 0.05、転調しない確率を 0.95 とし、モデルとしては表 5 のモデルを用いた。結果を図 14 に示す。

2 和声遷移モデル、常套句連接モデルともに、一時転調を含めて臨時記号の多く使われている中間部分では適切な調を見つけることができず、和声進行が旋律に良く適合していない部分が多く見られる。しかし、前半部分及び後半部分に関しては、旋律に適合した和声付けができていていると言える。また、経

Partita II (BWV 1004)



図 13: J. S. Bach: 無伴奏 Violin Partita II の冒頭 16 小節

表 6: 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
転調を含む旋律 (8 曲)	62.5	100.0
合計 (40 曲)	92.5	100.0

過音等を考慮することにより、派生音 (臨時記号のついた音符) を経過音とみなすことにより、部分的に改善することができた。より高次のモデルを導入して、経過音等の判断をよりの確なものにすることができれば、中間部分の不自然な部分により適切な和声付けが可能であろう。

4.3 HMM による調性認識

和声付け HMM を用いて転調の検出を含めた調性を推定することができる。すなわち、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することができる。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曲の調の推定をすることができる。本研究では平行長短調を区別していないが、和声付けされた和声列を後処理することにより長調と単調を区別することが可能である。

4.4 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、2 和声間遷移モデル (基本和声長: 2 拍) を用いて、転調の検出を含めた調性認識実験を行なった。これらの入力旋律中、途中で転調する旋律は 8 曲である。結果を表 6 に示す。表中で、転調検出位置が誤った場

モデル1: Dm	A	Dm	Gm	A	Dm
モデル2: Dm	Gm	A	Dm	Gm	A
モデル1(経過音等あり): Dm	A	Dm	Gm	A	Dm
1		2			
Dm		F		A#	Gm
Dm	Edim	F	F	A#	Gm
Dm	Gm	F	F	A#	Gm
	Edim	F	F		
3		4			
C	F	F	Dm	C#dim	Dm
C	F	Dm	A#	C#dim	Dm
C		F		C	F
5		6			
Dm	D	F		Esus4	C
Em	F	G	G	G	E
C		G	C		
7		8			
F		C	F	Esus4	C
Am		Am		Esus4	E
F	Bdim	E	Am	Esus4	E
	E		Am		
9		10			
F	C	A#	F	A#	
Am	C	A#	F	A#	
	F	G	C	A#	
11		12			
A#		E			
A#		E			
A#	F	E			
	F	E			
13		14			
F	C	Dm	A		
F	C	Dm	A		
F	F	Dm	A		
	F	Dm	A		
	C#dim	Dm			
	C#dim	Dm			
15		16			

図 14: J. S. Bach: 無伴奏 Violin Partita II 冒頭部分への和声付け結果

合に不正解とする場合を「認識率 A」に、転調の調性認識が正解した場合を「認識率 B」に示した。

これらのうち、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、転調を含む旋律に対しては、8 曲とも転調を検出して正しく調性を認識でき、そのうちの 5 曲については、転調位置も正しく検出した。

5 まとめと今後の課題

本稿では、連続音声認識の方法論を適用して、隠れマルコフモデルを用いて入力旋律に対して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。このような数理的な枠組で、楽曲スタイル(バッハのコラールなど)がある程度扱えることも示唆された。

本報告の目的は、音楽的な知識を極力用いずに数理的なモデルによりこの種の問題を扱い、統計学習が可能な統一的な枠組を提案することにあるので、性能評価は次の段階に残されている。技術的には、旋律生成確率、和声進行の確率の近似・推定方法、経過音・掛留音・先行音の判断、転調する旋律に対する和声付け方法、複数の候補を提供するモデルの使用方法などに関して、今後のモデルの精密化には大

きな研究の余地がある。また、音楽学、作曲法、和声学(さらに対位法や楽式論なども)の高次の知識を組み合わせることも有効であろうと考えられる。

今後の目標は、この枠組で楽曲のスタイルや作曲家の個性などが、モデルパラメータの学習によりどこまで表現できるかを追求すること、またその結果、(古典文学の研究における統計と同様に)音楽分析に利用できる統計モデルを確立すること、そして実用的な音楽作曲・編曲・教育のツール提供を可能とすることなどである。一方、和声付け結果の評価基準の確立は今後の大きな課題の一つであり、この点での学会のイニシアティブを期待するものである。

謝辞

本研究の一部は科学技術振興事業団による戦略的基礎研究推進事業 CREST プロジェクトの援助を受けた。本研究に関して有益な議論をしてくださった齋藤直樹氏を始めとする当研究室の学生諸氏およびバッハのコラールの和声統計抽出を手伝って下さった内海太陽氏に感謝します。

参考文献

- [1] XG Works V3.0 for Windows 95/98 取り扱い説明書, ヤマハ, 1998.
- [2] 長嶋 洋一, 橋本 周司, 平賀 譲, 平田 圭二 編: コンピュータと音楽の世界, bit 別冊, 共立出版株式会社, 1998.
- [3] 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.
- [4] 坪井 邦明: 音楽分析への AI 手法の応用, 情報処理, Vol. 29, No. 6, pp. 579-585, 1988.
- [5] 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.
- [6] 後藤 真孝, 平田 圭二: ハービー君: 演繹オブジェクト指向に基づいてジャズらしいコードにリハーモナイズするシステム, 情報処理学会研究報告, 96-MUS-16, pp. 33-38, 1996.
- [7] 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.
- [8] 吉野 巖, 阿部 純一: メロディの調を認定する過程の計算モデル, 情報処理学会研究報告, 93-MUS-3, pp. 23-30, 1993.
- [9] A. Schönberg: Models for Beginners in Composition, Schirmer, 1942. (A. シェーンベルク, 中村太郎訳, 作曲法入門 - 初めて作曲を学ぶ人のための曲の範例 -, カワイ楽譜, 1966.)
- [10] 石黒脩三: 和声学 (上), 全音楽譜.
- [11] J. J. Fux: Gradus ad Parnassum, 1725. (坂本良隆訳: 古典対位法, 音楽之友社, 1950.)
- [12] 島岡譲: 音楽の理論と実習 I, 音楽之友社, 1983.
- [13] 池内 他: 和声 - 理論と実習 - I, 音楽之友社, 1964.
- [14] Frank K. Soong, Eng-Fong Huang: A Tree-Trellis Based Fast Search for Finding the N Best Sentence Hypotheses in Continuous Speech Recognition, Proc. ICASSP 91, Vol. 1, pp. 705-708, 1991.