

隠れマルコフモデルに基づくピアノ運指の自動決定

米林 裕一郎[†] 亀岡 弘和^{††} 嵯峨山 茂樹^{††}

ピアノ曲演奏の運指を HMM を用いて自動決定するアルゴリズムを提案する。手の状態を「隠れ状態」とし、状態遷移から楽譜上の音符推移系列が出力されると考え、楽譜から Viterbi 探索により最尤状態遷移を求める。この手法により、指の独立性の難易度、黒鍵を含む鍵盤上での 2 次元的な指位置をモデル化できる。和音を含む両手の場合への拡張、音符長の考慮、学習、複数解の探索が可能で、本稿ではその一部について述べる。

Automatic Determination of Piano Fingering based on Hidden Markov Model

YUICHIRO YONEBAYASHI,[†] HIROKAZU KAMEOKA^{††}
and SHIGEKI SAGAYAMA^{††}

This paper proposes an algorithm for determining piano fingering automatically based on Hidden Markov Model. We assume here that states of hands are the states of a Hidden Markov Model and that a sequence of notes in a piano score are outputted in the process of the state transition of the Hidden Markov Model. By Hidden Markov Model, fingering determination is reduced to calculating the most likely state transition by means of Viterbi search. The algorithm proposed here can model the difficulty of moving a finger pair independently as well as positioning fingers two-dimensionally on a keyboard containing black keys. This algorithm not only can be applied to the case of playing chords or playing in the both hands, but can be enhanced to consider note length. Learning and multiple-solution search can be also easily realized.

1. はじめに

本稿では、隠れマルコフモデル (HMM) を用いてピアノ譜の運指を自動決定するアルゴリズムを提案する。

本稿は、高精度の運指決定を実現するだけでなく、ピアノ演奏の合理的な手の動きを定式化することを目標とする。ピアノ自習者への模範演奏の提示、曲の難易度判定や難易度をキーとした作品検索などへの利用が考えられるほか、ピアノ演奏ロボットの動作計画と考えればロボット指作業全般への応用も期待できる。

ピアノの運指決定に関する既存研究は片手の単旋律を対象とするものが多く、統計的な手法¹⁾、隣り合った音符間の運指コスト関数を定義する手法²⁾⁴⁾、rule-based な手法³⁾ などが見られる。片手の和音やポリフォニーを取り扱っているものもある⁴⁾。ただ、統計的な手法を用いる場合には運指情報がタグ付けされた

統計データが必要であり、rule-based な手法を用いる場合には適用ルールの競合解決が必要である。本稿で提案するアルゴリズムは、確率論的手法を用いる点の特徴である。

2. 運指の自動決定アルゴリズム

2.1 ピアノ演奏のモデル化

ピアノ演奏者が、楽譜あるいは MIDI データなどの形式で与えられた音符列を演奏する過程を考える。運指 S は、音符系列 N をどの指で弾くかというだけでなく、広義にはどのような手の位置/形で演奏するかという手指の状態遷移系列であり、さらには姿勢や肘の使い方など身体動作すべてを含み得る。一方、音符系列 N はピッチや音符長 (IOI) のほか、鍵の押下時間、音の強さ、音色など、楽譜で明示的/暗示的に示される広義の音符情報を含み得る。このような手指の状態遷移系列 S から音符系列 N を生み出す過程が演奏であると考えられる。

2.2 運指決定への HMM 導入

逆に、音符系列 N に対して理想的な演奏者が行な

[†] (株) アスケイドに勤務。

^{††} 東京大学大学院情報理工学系研究科

Graduate School of Information Science and Technology, The University of Tokyo

うであろう最も尤もらしい運指 S を推定することを運指決定と考えることができる。ここで、以降の節でのモデル近似の都合上、音符系列 N を音符長の系列 T 、ピッチの系列 Y 、それ以外の情報の系列 O の三つの系列に分解しておく ($N = (Y, O, T)$)。

運指で最も重要なのは、ある時点までの音符をそれぞれある手指状態で打鍵した後、次の音符を時間内に次の手指状態で無理なく打鍵できる確率と考えられる。各音符 $n_i = (y_i, o_i, t_i)$ での手指状態は直前の音符での手指状態とその音符長にのみ依存すると仮定すれば、音符系列 $N = (Y, O, T)$ に対して運指 S が使われる確率は、Bayes の定理により

$$p(S|(Y, O, T)) \propto p((Y, O, T)|S)p(S) \\ \approx \prod p((y_i, o_i, t_{i-1})|(s_i, s_{i-1})) \prod p(s_i|s_{i-1})$$

と近似できる。

以上のように考えると、運指を定式化するモデルとしては、演奏時の手指状態を「隠れ状態」と考え、その状態遷移の途中に音符推移が「出力文字列」として観測されるという Moore 型 HMM が適合する。

(ある音符から別の音符への) 音符推移すべてを表す集合を M とすれば、状態 s_1 から状態 s_2 へ遷移する際の出力確率関数は $b_{s_1, s_2} : M \rightarrow [0, \infty)$ で表される。この出力確率関数 b_{s_1, s_2} は、ある音符推移を演奏する上での運指遷移 $s_1 \rightarrow s_2$ の尤もらしさの違いを表すものである。例えばここで、 $s_1 =$ (右手薬指で打鍵している状態)、 $s_2 =$ (右手小指で打鍵している状態) とし、右手の薬指 小指という運指を考える場合、 $F\# \ G$ の演奏は自然であるが $A \ G$ の演奏は不自然である、という状況は以下の図 1 で表される。

2.3 HMM による運指決定の定式化

HMM を用いれば、運指決定は音符系列の背後に隠れた運指を推定するという、音声認識と同型の確率的な逆問題であると捉えられる。運指決定は、上式中の $p(S|(Y, O, T))$ を最大化する S を HMM の事後確率最大経路の Viterbi 探索により求める問題として定式化できる。

このように HMM を用いる利点として、運指決定のメカニズムを HMM のモデルパラメータの特徴や値により定性的・定量的に述べる点が挙げられる。

2.4 運指決定モデルへの近似導入

以上に述べたモデルは理想的には膨大なモデルパラメータを含み得るが、本稿では最初のステップとして少数のパラメータに絞るため以下の 3 つの近似を導入した。

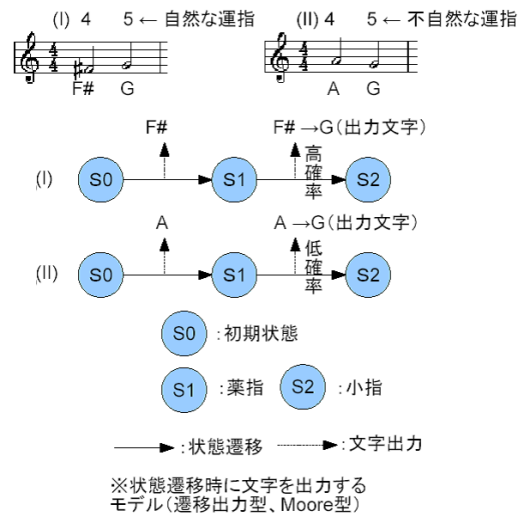


図 1 運指決定を定式化する HMM

2.4.1 近似 1: 手指状態・音符情報の簡略化

打鍵している指の番号(複数指で打鍵している場合はその集合)のみを「隠れ状態」と考え、さらに音符推移において O, T の成分を無視しピッチのみに着目した音符推移 (e.g., $F\# \ G\#$) を「出力文字列」とする。つまり、

$$p(S|(Y, O, T)) \\ \propto \prod p((y_i, o_i, t_{i-1})|(s_i, s_{i-1})) \prod p(s_i|s_{i-1}) \\ \approx \prod p(y_i|(f_i, f_{i-1})) \prod p(f_i|f_{i-1})$$

のような近似を行なう (f_i は i 番目の音符(群)を打鍵している指の番号(の集合)を表す)。2.2 節で述べた集合 M はピッチ推移すべてを表す集合となる。

この近似は、取り扱う入出力データを 2 つ(運指を決定する過程において第一の入力情報であるピッチ、決定すべき運指の第一義である打鍵指)に情報を限定することで、運指決定モデルの基本性能を確認するためのものである。

この近似の問題点としては以下が挙げられる：

- 運指を左右する要因である音符情報をピッチに限定することで、重要な要因を取りこぼす可能性がある(例えば音符長を考慮した運指決定ができないなど)。運指決定の精度への影響が予想される。
- 広義の運指情報(手の位置/形や姿勢や肘の使い方など)を決定できない。決定される運指情報が打鍵指に限られることで、運指情報をベースとした難易度判定などの応用可能性を損なう恐れがある。

前者の問題点については、出力確率関数 $b_{s_1, s_2} :$

$M \rightarrow [0, \infty)$ の定義域 M のパラメータに音符情報のパラメータを含めることで解決されると考える。後者については、可能な手指状態（広義の運指）の集合を何らかの離散集合として定義し、その中で状態遷移を定義することで解決されると考える。

2.4.2 近似 2: オクターブ位置の無視

集合 M において、ピッチのオクターブ位置の違いを考慮しない。例えば、鍵盤中央での D G と、高音部での D G は運指決定上同等であると近似する。

この近似は、出力確率関数 $b_{s_1, s_2} : M \rightarrow [0, \infty)$ の定義域 M を簡略化する効果を持つ。

一方でこの近似により、右手で低い音域を打鍵する（あるいは左手で高い音域を打鍵する）ことの困難性を運指決定モデルに反映できなくなるという問題が予想され、特に音符を右手/左手のどちらで打鍵するかを決定する場合には影響が大きい。

この問題への現実的な対策として、オクターブ位置により出力確率関数 b_{s_1, s_2} を補正する方法が考えられる。打鍵の容易な鍵盤の範囲は打鍵指（ $= s_i$ ）により異なることから、この補正は出力確率関数 b_{s_1, s_2} 毎に（個別に）行なわれるべきと思われる。

2.4.3 近似 3: 出力確率関数の簡略化

集合 M におけるピッチ（鍵）の推移を、ある鍵の打鍵位置から別の鍵の打鍵位置への二次元ベクトル \vec{a} として表す（黒鍵を含む鍵盤上での左右・奥行き方向の指移動を考慮するため）。この連続空間 M 上の出力確率関数 $b_{s_1, s_2} : M \rightarrow [0, \infty)$ は二次元 Gauss 分布に従うと近似する。

出力確率関数 b_{s_1, s_2} の二次元 Gauss 分布の例を図 2 に示す。この図は、右手中指（ $= s_1$ ）で F#（二次元座標平面の原点）を打鍵している場合、次の右手小指（ $= s_2$ ）での打鍵は A（二次元 Gauss 分布の中心）である確率が一番高いことを示している。

この近似は、隠れ状態遷移 $s_1 \rightarrow s_2$ が与えられた際に、対応する出力確率関数 $b_{s_1, s_2} : M \rightarrow [0, \infty)$ を具体的に定義するパラメータ数を低減する効果を持つ。

一方でこの近似により、出力確率関数 b_{s_1, s_2} の形が大きな制約を受け（例えば b_{s_1, s_2} を局所的に微調整できないなど）、パラメータ学習による精度向上の限界を生む可能性が考えられる。

この問題への現実的な対策として、Gauss 分布の代わりに折れ線型分布を採用するなど、出力確率関数 b_{s_1, s_2} の形の制約を緩める方法が考えられる。

2.5 モデルパラメータ値の設定方針

運指決定モデルの持つ構造やパラメータだけではなく、そのパラメータが持つ具体的な値についても、ピ

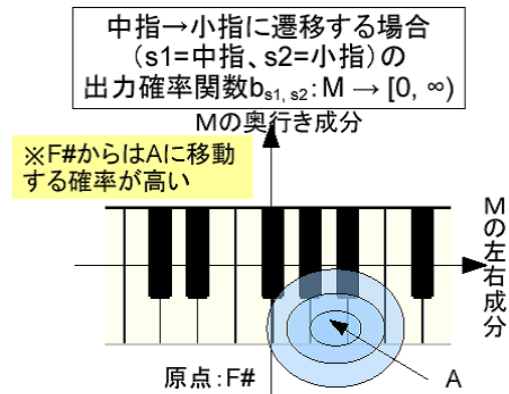


図 2 出力確率関数の二次元 Gauss 分布

アノ演奏時における手の動きを合理的に説明するものであることが望ましい。そこで、手の物理的な構造（「親指とそれ以外の指は広がりやすいが、人差し指～小指の間はそれに比べて広がりにくい」という性質や指の長さの制約など）や運指決定上の慣習（基本的には鍵間隔に指間隔を対応させるが、指くぐりが起こることもあるなど）を反映したパラメータ値を設定することが考えられる。

具体的には、モデルパラメータを学習によらずに以下の方針で設定することを考える：

- 出力確率関数の左右方向の中心は、指間隔に対応する鍵間隔に大体等しくする（例えば中指 小指は白鍵二つ分の間隔）。
- 出力確率関数の奥行き方向の中心は、手首から指先までの距離の差に大体等しくする。
- 親指と他の指が広がりやすいという傾向は、対応する出力確率関数の分散を比較的大きく設定することで表現する。
- 親指から他の指への遷移（あるいはその逆）では指くぐりが起こりやすいという傾向は、極大値を 2 箇所（通常の運指/指くぐりの運指に対応）で持つ出力確率関数で表現する。
- 指の独立性の理由により中指 薬指、薬指 小指などの運指が敬遠される傾向は、その遷移確率を（他の遷移に比べ）低くすることで表現する。

3. 評価実験

この章では、提案する運指決定アルゴリズムの動作確認のために行なった評価実験の内容を示す。

3.1 実験条件

最初のステップとして片手単旋律の楽曲（具体的には、バッハのインヴェンション第 1, 2 番の右手の単旋

律、音符数は約 620) を対象にし、運指決定モデルに 2.4 節の近似を導入した。

モデルパラメータの初期値は 2.5 節の方針により設定し、運指決定結果がより尤もらしくなるよう手動でパラメータ調節を行なった。

3.2 実験結果

以下の図 4 に示すように、多くの箇所でもっとも尤もらしい運指決定が行なわれた：

(A) 指間隔と鍵間隔の対応

(B) 適切な指くぐり

(C) 親指の黒鍵打鍵の難しさ

しかし、問題のある運指の箇所も見られた (図 5)：

(D) 音符長によっては同音連打を同じ指で弾きづらい

(E) 指間隔と鍵間隔が対応しているものの、手の動きが不安定

図 4 尤もらしい運指決定結果

図 5 近似に起因する問題箇所

3.3 結果の考察

(A) ~ (C) のような尤もらしい運指決定結果が得られたのは、2.5 節でのパラメータ設定方針によるものと考えられる。

一方、(D), (E) のような問題箇所は、2.4.1 節での近似に起因すると考えられる。具体的には：

(D) 音符系列のモデル (Y, O, T) から音符長の情報 T を除外したことによる。

(E) 手の動きを表すモデル要素が隠れ状態に不足している。

ただし (E) の問題については、2.2 節において手指状態は 2 つ以上前の音符の手指状態には依存しないと近似したことが理由である可能性もある。

手動調節されたモデルパラメータについては、2.5 節の方針から大きく外れていないと考えられる (表 1、表 2、表 3)。特に、指の独立性の理由により敬遠される運指は、その遷移確率を (他の遷移に比べ) 3 割程度低くすることで実現されている。

4. 運指決定モデルの考察

4.1 モデルの性質について

2.2 節において、現在の運指が将来の情報 (運指や音符) に依存するとは仮定されていない。しかし、3 章では過去と共に将来の情報が現在の運指に反映されたような実験結果が得られていることがわかる。

本稿の運指決定モデルは、各音符での手指状態は直前の音符での手指状態とその音符長にのみ依存するという最小限の仮説に基づくものであり、将来の情報への依存性は仮定されていない。実際、HMM において遷移確率 $p(s_i | s_{i-1})$ や出力確率関数 $b_{s_{i-1}, s_i} : M \rightarrow [0, \infty)$ は時刻 $i + 1$ 以降の情報に依存しない。それに関わらず将来の情報が現在の運指に反映されるように見えるのは、HMM の事後確率最大経路を探索する際に、曲全体として将来の運指や音符の反映された解が事後確率の高い解として得られるからだと考えられる。

このことから、本稿の運指決定モデルは直観的な最小限の仮説から運指決定のメカニズムを定式化したものであり、HMM の持つ機能や既存の成果 (Viterbi 探索、パラメータ学習など) を利用することで最大限の効果が得られていると言える。

4.2 モデルパラメータの学習について

モデルパラメータが学習によりどのような値に収束するかについての実験は未着手である。しかし、パラメータを手動調節した結果についての考察 (3.3 節) からは、2.5 節に述べたような手の物理的な構造や運指決定上の慣習を反映したパラメータ値に近くなることが予想される。さらに言えば、楽譜データに基づくパラメータ学習のほか、演奏者の手の物理的な構造を測定することによるパラメータ学習の可能性も考えられる。

5. 結論

本稿では、HMM を用いてピアノ譜の運指を自動決定するアルゴリズムを提案した。近似されたモデルで

		遷移後の指				
		親指	人差し指	中指	薬指	小指
遷移前の指	親指	(0.0, 0.0)	(42.0, 25.0)x0.81 (-23.0, 25.0)x0.19	(50.0, 30.0)x0.89 (-16.0, 30.0)x0.11	(85.0, 25.0)x0.91 (-16.0, 25.0)x0.09	(110.0, 0.0)x0.95 (-21.0, 0.0)x0.05
	人差し指	(-42.0, -25.0)x0.81 (23.0, -25.0)x0.19	(0.0, 0.0)	(23.0, 10.0)	(50.0, 0.0)	(82.0, -25.0)
	中指	(-50.0, -30.0)x0.89 (16.0, -30.0)x0.11	(-23.0, 10.0)	(0.0, 0.0)	(18.0, -10.0)	(57.0, -25.0)
	薬指	(-85.0, -25.0)x0.91 (16.0, -25.0)x0.09	(-50.0, 0.0)	(-18.0, 10.0)	(0.0, 0.0)	(20.0, -20.0)
	小指	(-110.0, 0.0)x0.95 (21.0, 0.0)x0.05	(-82.0, 25.0)	(-57.0, 25.0)	(-20.0, 20.0)	(0.0, 0.0)

※単位は(mm², mm²)。

※(a, b)は、2次元 Gauss 分布の中心の x/y 座標(左右/奥行き方向の成分)がそれぞれ a, b であることを示す。

※x(数値)は、2つの Gauss 分布を足し合わせる際の重みを表す。

表 1 出力確率関数のパラメーター一覧(2次元 Gauss 分布の中心)

		遷移後の指				
		親指	人差し指	中指	薬指	小指
遷移前の指	親指	(5.0, 30.0)	(900.0, 30.0)x0.81 (400.0, 30.0)x0.19	(900.0, 30.0)x0.89 (100.0, 30.0)x0.11	(900.0, 30.0)x0.91 (100.0, 30.0)x0.09	(900.0, 30.0)x0.95 (400.0, 30.0)x0.05
	人差し指	(900.0, 30.0)x0.81 (400.0, 30.0)x0.19	(5.0, 30.0)	(180.0, 30.0)	(200.0, 30.0)	(200.0, 30.0)
	中指	(900.0, 30.0)x0.89 (100.0, 30.0)x0.11	(180.0, 30.0)	(5.0, 30.0)	(190.0, 30.0)	(250.0, 30.0)
	薬指	(900.0, 30.0)x0.91 (100.0, 30.0)x0.09	(200.0, 30.0)	(190.0, 30.0)	(5.0, 30.0)	(200.0, 30.0)
	小指	(900.0, 30.0)x0.95 (400.0, 30.0)x0.05	(200.0, 30.0)	(250.0, 30.0)	(200.0, 30.0)	(5.0, 30.0)

※単位は(mm⁴, mm⁴)。

※(a, b)は、2次元 Gauss 分布の分散共分散行列 σ が $\sigma_{xx}=a$, $\sigma_{yy}=b$, $\sigma_{xy}=\sigma_{yx}=0$ であることを示す。

※x(数値)は、2つの Gauss 分布を足し合わせる際の重みを表す。

表 2 出力確率関数のパラメーター一覧(2次元 Gauss 分布の共分散)

		遷移後の指				
		親指	人差し指	中指	薬指	小指
指の総数	親指	0.20	0.20	0.20	0.20	0.20
	人差し指	0.20	0.20	0.20	0.20	0.20
	中指	0.21	0.21	0.21	0.16	0.21
	薬指	0.23	0.23	0.16	0.23	0.16
	小指	0.21	0.21	0.21	0.16	0.21

表 3 出力確率関数のパラメーター一覧(遷移確率)

実験を行うことにより、この手法の基本性能を確認すると共に、近似を軽減したより精密なモデルが必要であることを示した。

HMM を用いた当アルゴリズムの利点として、以下の点が挙げられる:

- 運指決定のメカニズムを HMM のモデルパラメータの特徴や値により定性的・定量的に述べることができる。例えば、隣り合った2つの音符を打鍵する2本の指の広がりやすさは HMM の出力確率関数の分布で表され、(2本の)指の独立性の難易度の違いは HMM の状態遷移確率の大きさの違いで表される。

- 推論過程・結果の正当性(尤もらしさ)を確率として表現できる。例えば、複雑なルールベースにおける適用ルールの競合といった問題は生じない。
- 運指を推定するだけでなく、推定された手の状態(手の形、位置など)をもとに運指の難易度を判定し、それをベースとした応用研究が可能である。(想定される応用については後述)

さらに、本稿で示したアルゴリズムの枠組みは以下のような拡張可能性を持つ:

- 片手の単旋律に留まらず、両手の和音やポリフォニーを含む場合、さらに左右の手への振り分けが

片手の単音旋律
(Chopin: Mazurka Op.59-1)

片手の和音旋律
(Ravel: Rigaudon)

片手の複旋律(ポリフォニー)
(J.S.Bach: Fuga from Well-tempered Clavier BWV 846)

両手の和音旋律
左右の割り振り既知・未知
(Brahms: Intermezzo Op.118-2)

図3 運指決定の様々なケース

未知の場合も取り扱える。譜例を図3に示す。

- 強拍のフォルツァンドは親指で叩く、あるいは、楽譜に既存の部分的な指使いのような例外ルールも、探索空間の限定により実現可能である。
- 運指が指定された楽譜によるパラメータ学習のほか、Baum-Welch アルゴリズムを用いて運指情報のない楽譜データから行うパラメータ学習も可能。さらに、演奏者の特性(手の大きさ、指の広がりやすさや独立性)に応じてパラメータ値を調整(学習)することによるカスタマイズも可能である。
- 複数の解を求める N -best 探索が可能である。

本稿の運指決定アルゴリズムの応用可能性としては、以下のようなものが考えられる:

- 難易度をベースとした応用
曲中の難しい箇所を検出することによる曲の難易度判定や練習計画の自動作成、難易度をキーとした楽曲データベース検索の実現など。あるいは、手の大きさに応じた難易度(子供向けの曲か否か、など)を判定する技術。
- 作曲・編曲(アレンジ)への応用
自動作曲された曲が演奏可能かどうかを判定する要素技術の研究。あるいは、演奏が容易かつ効果が大きい自動作曲・編曲(オケ譜からピアノへの編曲など)を行うための基礎。
- 本稿で提案した数理的な問題解決を、他の楽器にも応用する。例えばギターなど。

今後は、より精密なモデルの実装のほか、学習、和音・ポリフォニーを含む場合や両手を含む場合の定式化と実験検証を予定している。

参考文献

- 1) 野口賢治, 野池賢二, 乾伸雄, 野瀬隆, 小谷善行, 西村恕彦, “ n グラムの手法を用いたピアノ運指の推論,” 情報処理学会第52回全国大会講演論文集, Vol.2, pp.101-102, 1996.
- 2) Melanie Hart, Robert Bosch, Elbert Tsai, “Finding Optimal Piano Fingerings,” The UMAP Journal 21(2), pp.167-177, 2000.
- 3) 林田教裕, 水谷哲也, “楽曲構造に基づくピアノ運指ルールの論理表現,” 情報処理学会第65回全国大会講演論文集, 2003.
- 4) Alia Al Kasimi, Eric Nichols, Christopher Raphael, “Automatic Fingering System (AFS),” poster presentation at ISMIR, 2005.