

事例転写型自動演奏表情付けのための 効率的最適事例系列探索手法の検討

奥村 健太^{1,a)} 酒向 慎司^{1,b)} 北村 正^{1,c)}

概要: 演奏者が持つ特徴の再現性を重視した自動演奏生成システムの提案のため、未知の演奏指示に対して最適な表情を持つ事例を既知の実演奏から探索し、その表情の転写によって仮想演奏を生成する問題を考える。その際に論点となるのは、どのようにして最適な事例を探索するかと、その探索をどれだけ効率化できるかである。本稿では、実演奏における楽譜とその逸脱を関連付けた統計モデルの体系的構造を指標として活用する枠組みについて説明するほか、評価実験から探索対象とする事例群の制御にその体系的構造を活用することが仮想演奏の生成に最適な事例の効率的な探索に有効であることを示す。

A study of the method for efficient rendition of optimal performance expression in the case-based automatic performance rendering approach

KENTA OKUMURA^{1,a)} SHINJI SAKO^{1,b)} TADASHI KITMAURA^{1,c)}

Abstract: We consider a problem to explore cases with suitable expression for unknown performance direction from known performance cases, for the proposal of automatic performance rendering system that places importance in reproducibility of the characteristics of the performer. The issues to be solved are designing index for selecting optimal cases, and efficient search of such cases. In this paper, we describe the validity of proposed method for efficient search of the optimal cases for the rendition of the virtual performance of unseen composition by the utilization of systematic structure of the stochastic model that associates the artistic deviation and notated direction in the real performance.

1. はじめに

機械に人間らしい演奏をさせることを関心に、演奏表情付けを目的としたシステムが多数提案されている [1]。主に演奏行為のどの部分を自動化するかに依存し、多様な形態や手法が採られている。ここでは既存の演奏者個人の特徴に忠実な再現性を備えた仮想演奏の自動生成を目標とし、その特徴の生成規則を実演奏から自動的に獲得し、適切な使い分けにより演奏表情を自律的に制御するシステムを考える。そこで合理的なのは実演奏の素片から任意の事例を選択し、その特徴を転写する手法であるが、本来実演奏で連続していない事例の特徴を接続するため、その仮想演奏で自然性を保つことが比較的困難である [2-4]。その点は

複数の事例における特徴の統計量を転写することで緩和できるが、特徴量が平滑化されるため、演奏者の特徴の再現が困難となる [5-7]。以上から再現性と自然性はトレードオフの関係にあるといえるが、我々が重視するのは前者である。よって課題となるのは、(1) 演奏者の特徴を保持しつつ自然性を損なわない適切な事例の選択と、そのような事例を必要かつ十分な候補事例から探索するための、(2) 探索規模の増大に従い増加する計算コストの軽減である。

我々は実演奏事例の生成過程を説明する統計モデルを定義し、自動分類によって体系化したそれらの木構造に従って自動演奏生成を実現するアイデアを提案している [8]。それに基づくシステムが生成した仮想演奏は自動演奏表情付けシステムの合同評価環境^{*1}において自律型部門の第1位を獲得しており、課題(1)については主観評価から既に

¹ 名古屋工業大学 (Nagoya Institute of Technology)

a) k09@mmsp.nitech.ac.jp

b) sako@mmsp.nitech.ac.jp

c) kitamura@nitech.ac.jp

^{*1} Rencon 2013: Music Performance Rendering Contest for Computer Systems, <http://smac2013.renconmusic.org>

一定の成果を得ている。本稿ではまず提案済みのアイデアについて説明する。そして仮想演奏に最適な事例の探索効率を客観的に評価し、課題 (2) における有用性を検討する。

2. 提案手法

演奏者は与えられた楽譜から音楽的な構成や感情などの解釈を楽器の操作に反映することで実演奏に表情を付与しており、それらは音符と拍を尺度とした定量的な観測が可能とする。ここでは物理的な楽譜に加え、演奏指示内容上で実演奏に一致するように調整された論理的な楽譜を仮定し、演奏者による実演奏の生成過程を以下のように考える：

- (1) 物理楽譜の演奏指示系列 $S = (s_1, \dots, s_M)$ の解釈により実演奏の内容を仮想化する論理的変換を行い、論理楽譜上の演奏指示系列 $\hat{S} = (\hat{s}_1, \dots, \hat{s}_N)$ を生成*2。
- (2) \hat{S} に意図した表情を加えて楽器を制御する物理的変換を行い、実演奏系列 $\hat{R} = (\hat{r}_1, \dots, \hat{r}_N)$ を生成。

過程 (1) で生成される演奏指示に不一致な事例 [9] を仮想演奏でも再現するため、以下の情報により S の演奏指示のみからそのような事例の生成過程の説明を可能にする：

X^{PS} Pitch Shift. 演奏指示の音高からの相違を半音単位で記述する整数値。指示と同一の音高ならば 0。

X^{KS} Key Strokes. その事例と物理楽譜上で同一の演奏指示に基づき演奏された総事例数を示す整数値。通常は 1 として励起事例数を加算。欠落なら 0 とする。

以上から、システムは以下の過程で仮想演奏を生成できる：

- (I) 与えられた物理楽譜の演奏指示系列 S に対する最適事例系列 $V = (v_1, \dots, v_M)$ を事例の探索により生成。
- (II) v_m について x_m を参照し、 \hat{v}_n への変換に際する不一致事例系列の生成規則を適用して得られた仮想演奏系列 $\hat{V} = (\hat{v}_1, \dots, \hat{v}_N)$ の仮想演奏を生成。

尚、過程 (2) における表情は、 \hat{S} と \hat{R} の対照によって以下の各種逸脱量系列 $D = (d_1, \dots, d_N)$ として観測できる：

D^{AT} 音符単位の打鍵時刻の逸脱量（基準は拍長）。

D^{GR} 打鍵から離鍵までの継続長と記譜音価の比率。

D^{DR} 記譜強弱指示と実演奏の velocity (MIDI) の比率。

D^{BR} 楽曲単位に対する拍単位の局所的な BPM の比率。

これらは実演奏に固有な特徴の主な構成要素であり、演奏者の芸術的な意図の下で楽器の操作に付加されている [10]。

また、各種逸脱量の生成はそれ以前の演奏指示における逸脱傾向にも依存すると考えられるが、ここでは簡単のため、直前の逸脱量との差分量を逸脱量と同様に扱う。任意の演奏指示 \hat{s}_n に基づく実演奏事例 \hat{r}_n における、要素 F の逸脱量 d_n^F とその差分量 $d_n^{\Delta F}$ の算出式を以下に示す：

$$d_n^F = \begin{cases} \hat{r}_n^F - \hat{s}_n^F, & F = AT \\ \hat{r}_n^F / \hat{s}_n^F, & F = (GR, DR, BR) \end{cases}, \quad (1)$$

$$d_n^{\Delta F} = d_n^F - d_{n-1}^F, \quad F = (AT, GR, DR, BR). \quad (2)$$

*2 拍の場合は $M = N$ であるが、音符の場合は $M \simeq N$ となる。

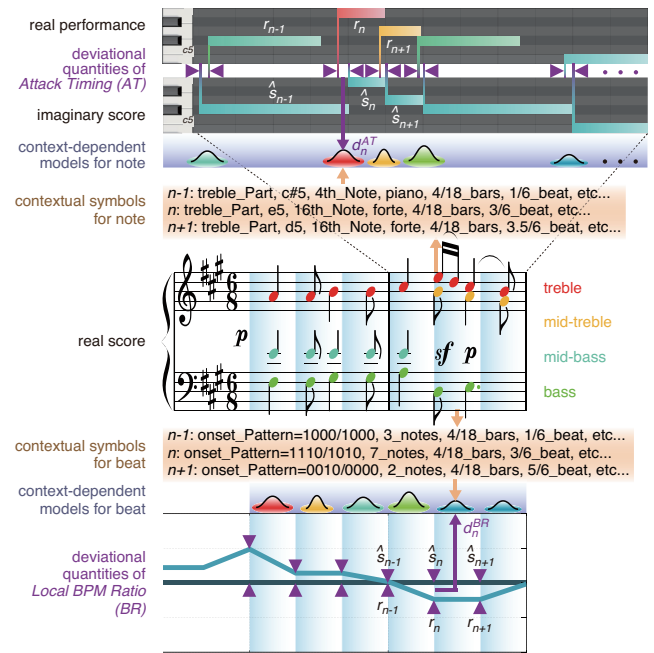


図 1 局所的な演奏指示の文脈に依存した逸脱量生成過程のモデル。

Fig. 1 Formation of context-dependent models.

2.1 楽譜情報に基づく芸術的逸脱の付加傾向の体系化

物理楽譜の情報のみから仮想演奏の生成に適切な事例を選択するには、実演奏事例を生成した演奏者がどのように論理楽譜の演奏指示と逸脱量を関連付けたかを一般的に説明する規則が必要となる。しかし解釈そのものに表現方法に依存した揺らぎがあり、その一般的な説明は演奏者本人にも困難である [11]。ここでは演奏者の高次な解釈そのものではなく、物理楽譜から一意に得られる情報を用い、それらの依存関係の近似的な説明を考える。楽譜から得られる局所的な演奏指示系列を文脈に見立て、演奏者の振舞いがそれに依存した傾向を持つとする仮定の下では、専門家による楽曲分析 [4] などの付加的な情報を用いることなく、両者の体系的な関連付けを可能とする手法がある [12]。楽譜の情報に基づいて実演奏の事例の分類を行うという点で、問題の本質はそれと共通である。

実演奏の生成について先述した過程に基づき、論理楽譜中の任意の演奏指示の文脈の下で出力された個々の逸脱情報の対応関係を記述するモデル（文脈依存モデル）を定義する。このモデルでは、演奏指示 \hat{s}_n による事例 \hat{r}_n において観測された G 種類の逸脱量および差分量 d_n の傾向を、次式に示す確率密度関数を持つ多変量正規分布とみなす：

$$P(d_n | \mu_n, \sigma_n) = \prod_{f \in F} P(d_n^f | \mu_n^f, \sigma_n^f) \\ = \left\{ (2\pi)^G \prod_{f \in F} |\sigma_n^f| \right\}^{-\frac{1}{2}} \exp \left[- \sum_{f \in F} \left\{ (2\sigma_n^f)^{-1} (d_n^f - \mu_n^f)^2 \right\} \right] \quad (3)$$

$$\begin{cases} F = (AT, GR, DR, \Delta AT, \Delta GR, \Delta DR), & G = 6, \text{ for note} \\ F = (BR, \Delta BR), & G = 2, \text{ for beat} \end{cases}$$

各種逸脱要素は演奏者個人の中で内部的な依存関係が考えられるが、その具体化は別途検討を要する問題として、こ

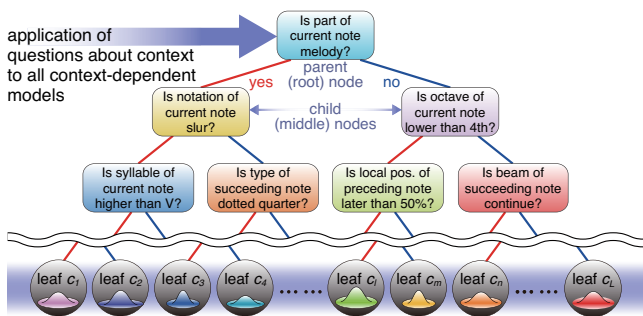


図 2 Tree-based clustering による文脈依存モデルの体系化.

Fig. 2 Systematization of context-dependent models.

ここでは各変量をそれぞれ独立とみなす. そして d_n の生成にはその周囲の局所的な文脈が寄与すると考え、 \hat{s}_{n-1} , \hat{s}_n , \hat{s}_{n+1} から得られる文脈情報の組合せ (図 1) を付与する.

実演奏から得られる事例は有限であるため、充分なモデル学習を行えない可能性があるだけでなく、あらゆる文脈情報に対応することも現実的に困難である. 学習データに含まれない未知な文脈情報を含む仮想演奏の生成には、有限な文脈依存モデルの何れかをその代替として使用可能にするための、体系的な共有規則が望まれる. そのような共有規則の体系は、Tree-based clustering [13] によって全ての文脈依存モデルを分類することで、決定木として構築できる (図 2). この木構造は、演奏において何らかの類似した傾向を持つ事例が楽譜中のどのような演奏指示の組合せによって生成されるのかを明示する. 分類された個々の文脈依存モデルは末端のリーフノードに配置され、それら以外の各中間ノードには分類基準となった質問が保持される. また、ルートノードから質問に従い木構造の中間ノードを辿ることで、文脈の内容が未知であっても、最終的にはそれに最も類似した事例のリーフノードに到達できる. よって、これを活用することが未知楽曲の文脈情報に対して適切な表情を持つ事例の特定にも有効であると考え.

2.2 仮想演奏生成のための最適事例の探索手法

木構造の最適化基準は学習データの文脈情報や逸脱傾向に依存して固定されるため、任意の未知楽曲の仮想演奏生成に最適な木構造を得ることは現実的に困難である. 一般的な木構造において、ルート付近のノードの構造は共通し、比較的高い汎用性を持つ傾向にある [12]. 一方リーフノードは特定の事例の逸脱傾向に依存し、その付近のノードはある程度類似した逸脱傾向を持つ傾向にある. 仮想演奏の生成に際して要求される文脈情報の自由度に対し、現実的に実演奏から得られる事例は著しく限られることを考慮すると、任意の文脈情報の組合せに対して木構造上で厳密に対応するリーフノードの逸脱傾向が必ずしも最適とは限らない. そこで、仮想演奏の生成に際して与えられる物理楽譜の演奏指示 $S = (s_1, \dots, s_M)$ において、 m 番目の演奏指示 s_m の文脈情報に従って木構造を順に辿った結果到

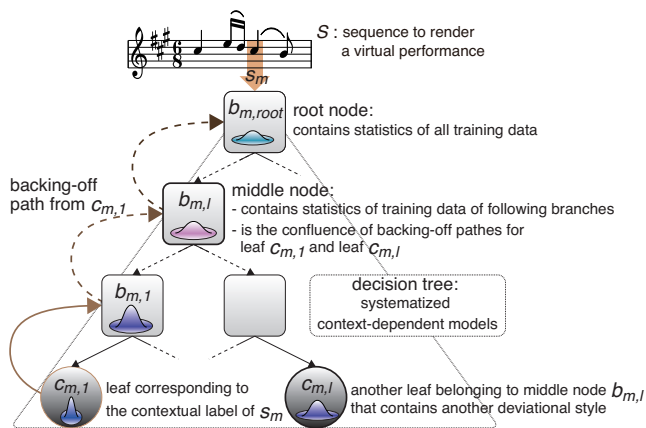


図 3 木構造のバックオフによる選択候補事例の拡張.

Fig. 3 Decision tree backing-off concept.

達したリーフノード $c_{m,1}$ を起点として、辿ってきた中間ノード群を遡る木構造のバックオフ [14] を行う. それにより、各中間ノードに属する他のリーフノードに相当する事例群を探索候補として段階的に拡張できる (図 3).

ここで任意の回数のバックオフが中間ノード $b_{m,l}$ に到達した際、任意のリーフノード $c_{m,l}$ に相当する事例が探索候補に追加されるものとする. その $c_{m,l}$ が $c_{m,1}$ よりも s_m に適したものであると仮定し、木構造のバックオフによって段階的に拡張された候補の中でその探索を行う. S の仮想演奏として最適な事例系列を $V = (v_1, \dots, v_M)$ とする. このうち v_m は、木構造のバックオフによって s_m 毎に拡張された L_m 個の候補事例群 $C_m = (c_{m,1}, \dots, c_{m,l}, \dots, c_{m,L_m})$ の中から最適な事例を選択することで獲得できる. この選択を s_m 毎に段階的に行えるものとする. その系列の探索には動的計画法 [15] を適用できる (図 4).

演奏指示 s_m に対する事例 $c_{m,l}$ の選択の相応しさを評価するため、 $c_{m,l}$ が実演奏として観測された際の各種要素 F についての逸脱量 $d_{m,l}^F$ とその差分量 $\Delta d_{m,l}^F$, そして中間ノード $b_{m,l}$ が持つそれらの統計量 (平均 $\mu_{m,l}^F$, 分散 $\sigma_{m,l}^F$) に基づく尤度を用いる. まず s_1 に対して $c_{1,l}$ を選択する場合の評価値を $h_1(c_{1,l})$ とする. 次に (s_1, s_2) に対して $(c_{1,k}, c_{2,l})$ の組合せを選択する場合の評価値を $h_2(c_{1,k}, c_{2,l})$, それ以降の評価値も同様に $h_m(c_{M-1,k}, c_{M,l})$ から求める. 各評価値の算出式をそれぞれ以下に示す:

$$h_1(c_{1,l}) = P(d_{1,l}^F | \mu_{1,l}^F, \sigma_{1,l}^F) = \prod_{f \in F} P(d_{1,l}^f | \mu_{1,l}^f, \sigma_{1,l}^f), \quad (4)$$

$$h_m(c_{m-1,k}, c_{m,l}) = P(d_{m,l}^F | \mu_{m,l}^F, \sigma_{m,l}^F) P(\Delta d_{m,l}^F | \mu_{m,l}^{\Delta F}, \sigma_{m,l}^{\Delta F}) \quad (2 \leq m \leq M). \quad (5)$$

ここで音符の場合 $F = (\text{AT}, \text{GR}, \text{DR})$, 拍の場合 $F = \text{BR}$ であり、 $\Delta d_{m,l}^F = d_{m,l}^F - d_{m-1,k}^F$ は $(s_{m-1,k}, s_{m,l})$ に対する事例の組合せとして $(c_{m-1,k}, c_{m,l})$ の選択を仮定した際の F についての差分量である. 式 (5) では $\Delta d_{m,l}^F$ を生じる選択の相応しさを $d_{m,l}^{\Delta F}$ の統計量 (平均 $\mu_{m,l}^{\Delta F}$, 分散 $\sigma_{m,l}^{\Delta F}$) に基づき考慮している. 以上から求める系列の探索

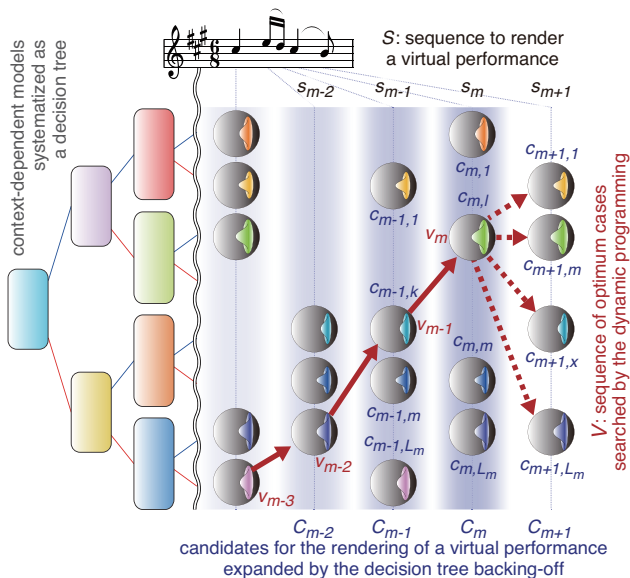


図 4 動的計画法による最適事例系列の探索。

Fig. 4 Selection of the sequence with optimum cases to render.

は式 (4), (5) に示す各評価値の総和を最大化する選択を各段階で行う問題といえる。その目的関数を以下に示す：

$$J = h_1(c_{1,l}) + h_2(c_{1,k}, c_{2,l}) + \dots + h_M(c_{M-1,k}, c_{M,l}) \rightarrow \max. \quad (6)$$

このバックオフは最終的にルートノードに到達し、木構造に含まれる全ての事例が探索候補となる。しかし木構造を参考にする上で $c_{m,l}$ から著しく遠い位置にある事例が最適となる可能性が高いとは考えにくいので、常に木構造中のあらゆる事例を対象とした探索が必要とは限らない。従って、探索候補とする事例の拡張を任意の規模に制御することによる、計算コストの面での探索の効率化が考えられる。任意の回数のバックオフで到達する中間ノード $b_{m,l}$ の統計量に基づく評価値 $\theta_{m,l}$ を次式で求める：

$$\theta_{m,l} = (b_m^{\max} - b_m^{\min})^{-1} \{P(d_{m,1} | \mu_{m,l}, \sigma_{m,l}) - b_m^{\min}\} \quad (0 \leq \theta_{m,l} \leq 1). \quad (7)$$

ここで b_m^{\max} , b_m^{\min} は、それぞれ木構造をバックオフした際に辿る各中間ノードについて $c_{m,1}$ に基づく尤度を求めた中での最大値、最小値である。事前に指定した閾値とこの評価値によって、各中間ノードでバックオフの継続・停止を判定する。 $\theta_{m,l} = 1$ に近い条件では、木構造上で $c_{m,1}$ に近い範囲の事例のみが探索候補となる。

3. 評価実験

未知楽曲の仮想演奏生成における木構造のバックオフの導入の有効性を検証するための評価を行った。文脈依存モデルの学習データとして用いた演奏者または部分的に共通な楽曲の実演奏の組合せ (表 1) に対してオープンとなる楽譜 (W. A. Mozart's Piano Sonata, the 1st movements

of K. 279, treble voice part) を用い、仮想演奏を生成した。文脈依存モデルの木構造を活用し、異なる規模の事例群を対象とした場合に最適となる事例系列を探索するため、式 (7) に示した基準を用いて事例の探索規模を制御している。

3.1 最適事例の選択における探索規模の影響

事例の探索規模を変化させた各場合で選択された事例が、学習データの全事例を探索対象とした場合 (all cases) とどの程度一致したかを調査した。各セットについて、音符の場合を図 5 に、拍の場合を図 6 に示す。各図の横軸は各セットの N^{note} または N^{beat} に対する探索規模の比率を表す。横軸の 100% はそれらの全事例を対象とした探索を、各グラフの左端はバックオフを行わず木構造上で対応した事例を直接転写した場合を意味する。

各図から、いずれのデータの組合せにおいても木構造上で近い範囲の極めて部分的な事例のみを探索範囲としながら、全事例を対象とした場合と完全に一致する選択結果を得られていることがわかる。特に P1 の約 2 倍の学習データ量を持つ P2 では、P1 よりも相対的にも絶対的にも少数の事例を対象として、全事例を対象とした場合と完全に一致する選択結果を得られている。よって、学習データに含まれない未知の楽曲に対する柔軟な対応には、その木構造の構築に用いる学習データとしてある程度以上の絶対量を確保することが望ましいと考えられる。

一方、学習データを限定した P1 では同一の楽曲の実演奏を学習データとした G1 に対し、特に音符の場合でより多数の事例の探索を要している。ここから、未知の楽曲の仮想演奏として望ましい結果を安定して得るには楽曲の組合せだけでなく、演奏者に依存した特性の相違も多分に考慮を要すると考えられる。しかしどちらの場合でも限られた学習データの約 20% から 50% 程度の事例で全事例の探索に一致する選択結果に到達できていることから、木構造のバックオフは学習データに含まれない未知の楽曲の仮想演奏の生成に最適な事例を選択するための探索対象を削減し、その効率化を可能にするといえる。尚、ここで用いた演奏者や楽曲の組合せのみからは、適切といえる学習データの量や組合せの一般化は困難である。その具体化には、多様な実演奏データの組合せで作分けした文脈依存モデルを用いての検証を要する。

ここで G1 での仮想演奏には、各探索規模において図 7 に示すような逸脱量が転写されている。これらの図からは、全ての探索規模において概ね類似した各種逸脱量の変動傾向が得られているように見える。しかし w/o backing-off の条件のみ、直前・直後の傾向に対して突発的かつ急峻、そして他の条件に対して逆向きの変動となっているような箇所が明らかに多い。これは正解系列の無い比較ではあるが、大幅かつ継続性の無い局所的な変動が演奏に自然性を与えるものであるとは一般的には考えにくい。そのような

表 1 文脈依存モデル学習のための実演奏データと仮想演奏生成のための楽譜データ。

Table 1 Data used to train models and render virtual performance.

Data set	G1	P1	P2
Real performer	G. Gould	M. J. Pires	
Data to train context-dependent models	W. A. Mozart's Piano Sonata, the 2nd and 3rd movements of K. 279 and the first movement of K. 310.	the 2nd and 3rd movements of K. 310, in addition to the left column.	
Amount of data	$N^{\text{note}} = 2305$, $N^{\text{beat}} = 396$.	$N^{\text{note}} = 2292$, $N^{\text{beat}} = 396$.	$N^{\text{note}} = 4767$, $N^{\text{beat}} = 893$.

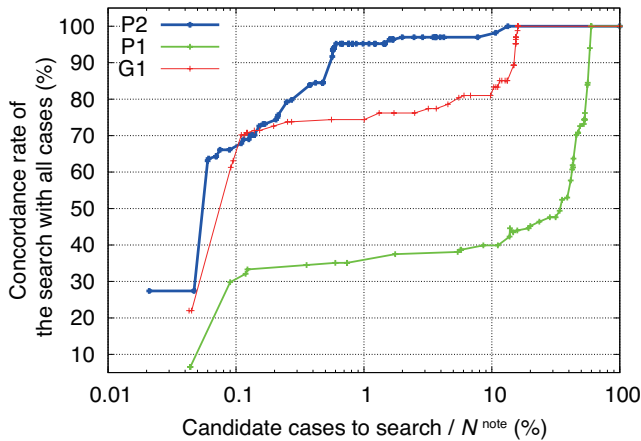


図 5 全事例探索に対する選択結果の一致率（音符）。
Fig. 5 Concordance rate of selected cases (note).

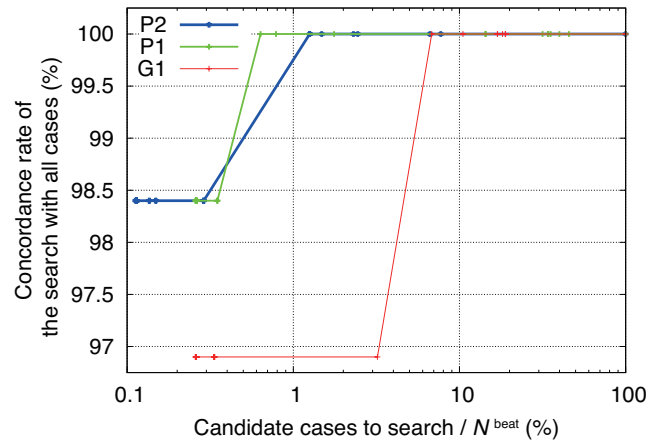


図 6 全事例探索に対する選択結果の一致率（拍）。
Fig. 6 Concordance rate of selected cases (beat).

変動を $\theta_{m,l} > 0.9$ のように比較的小規模な探索範囲の拡張でも解消できている点からも、最適事例探索の効率化における木構造のバックオフの有効性を確認できる。

3.2 仮想演奏と実演奏における逸脱傾向の比較

データセット G1 から学習した文脈依存モデルを用いて生成した仮想演奏について、G1 と同一の演奏者による実演奏と各種逸脱量を比較し、それらの相関係数を求めた結果を表 2 に示す。

表 2 から、全ての要素において正の相関が得られており、その傾向は D^{GR} において比較的明確である。その他の要素の値は強い相関を示すものとはいえないが、演奏の計測分析を行った既存研究^{*3} から、逸脱傾向が演奏者の同一性よりも楽曲の特性に強く制約されることが示されており、今回の条件はそれに該当する可能性がある。また、 D^{AT} 、 D^{GR} 、 D^{BR} は選択候補事例の探索規模が拡張されることで相関係数も向上しているが、 D^{DR} では逆の傾向となっている。ここで D^{AT} や D^{GR} は単一の音符の発音に比較的強く関与するのに対し、 D^{DR} や D^{BR} の変動がここで用いた局所的な文脈の単位よりも大域的な情報に制約される可能性がある。その点から、文脈依存モデルに付与する文脈情報の記述内容と適用範囲を逸脱要素間で統一した設計には改良の余地があるといえる。しかし文脈情報につい

表 2 生成された仮想演奏と同一演奏者による実演奏の相関係数。

Table 2 PPMCC between virtual and real performance.

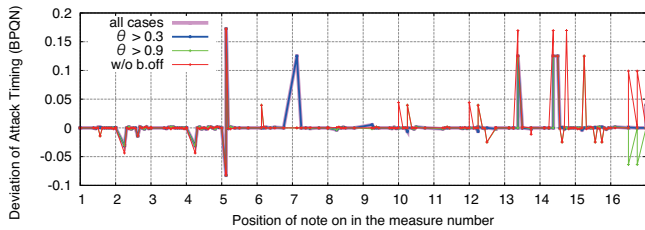
Search range of cases	Correlation coefficient of			
	D^{AT}	D^{GR}	D^{DR}	D^{BR}
all cases	0.221	0.314	0.119	0.205
w/o backing-off	0.196	0.303	0.227	0.117

ては、学習データに含まれない演奏指示に対応する上で一般性を保つ必要がある。特に D^{DR} の傾向は、強弱に関する具体的な演奏指示よりも音高などの一般的な情報に強く制約されることが示されており [16]、楽曲の特性への依存度が高い情報を文脈の記述に用いることは適当ではない。従って、文脈依存モデルの設計の改良を考える場合に検討すべき方法には、モデルに関連付ける文脈情報を単純に多様化させる以外のアプローチを要すると考えられる。

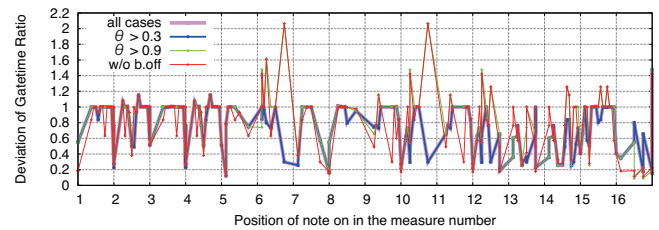
4. おわりに

演奏者の特徴の再現性を重視した仮想演奏の自動生成の一解法として、本稿では実演奏事例の生成過程を説明するモデルを体系化した木構造をバックオフによって最適な事例の探索を行う解法を提示し、その客観的な性能の評価を行った。事例の特徴の転写による仮想演奏の生成における課題の一つは、適切な事例の効率的な探索である。本手法で導入している木構造のバックオフでは、仮想演奏生成に最適な事例の選択候補として、事例群に木構造に従った段階的な優先順位を与えている。本稿の評価実験からは、それによって段階的に拡張された事例群を探索対象とした場

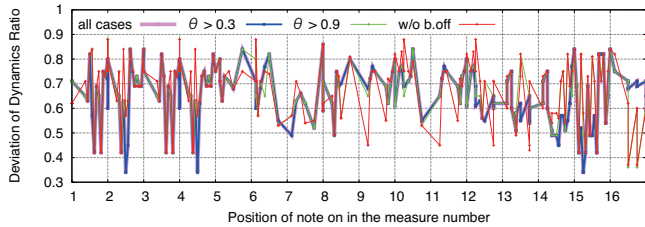
^{*3} CHARM Department of Music Royal Holloway, AHRC Research Centre for the History and Analysis of Recorded Music: Mazurka Project, <http://www.mazurka.org.uk>



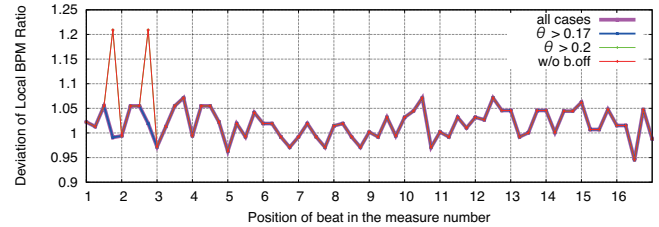
(a) Rendered expression of Attack Timing.



(b) Rendered expression of Gatetime Ratio.



(c) Rendered expression of Dynamics Ratio.



(d) Rendered expression of Local BPM Ratio.

図 7 事例の探索規模別に生成された逸脱量の軌跡 (逸脱要素別).

Fig. 7 Deviation quantities in virtual performances rendered for each search range.

合で、全ての事例から探索を行う場合と同様の探索結果を得られること、また、その傾向はより大規模な事例での探索を行う場合により効果的に作用することが示された。これにより本手法は、生成される仮想演奏の自然性に加え、その生成手法としての合理性も評価されたといえる。

尚、本稿では実演奏データベース^{*4}における構成の偏りから、対照比較の組合せに制約を受けている。仮想演奏の生成に実演奏を用いる自動演奏表情付け手法においてはそのデータ調達の困難さに普遍的な問題がある [5] が、演奏者の特徴を作り分けた仮想演奏の生成には、適切な学習データの量やその組合せについての検討が今後の課題となる。

参考文献

[1] 松原正樹, 深山 覚, 奥村健太ほか: 創作過程の分類に基づく自動音楽生成研究のサーベイ, コンピュータソフトウェア, Vol. 30, No. 1, pp. 101–118 (2013).
 [2] Suzuki, T., Tokunaga, T. and Tanaka, H.: A case based approach to the generation of musical expression, *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, pp. 642–648 (1999).
 [3] Hirata, K. and Hiraga, R.: Ha-Hi-Hun: Performance rendering system of high controllability, *Proceedings of Rencon Workshop in International Conference on Auditory Display*, Kyoto, Japan, pp. 40–46 (2002).
 [4] Tobudic, A. and Widmer, G.: Relational IBL in Music with a New Structural Similarity Measure, *Proceedings of the 13th International Congerence on Inductive Logic Programming*, Szeged, Hungary, pp. 365–382 (2003).
 [5] Grindlay, G. and Helmbold, D.: Modeling, Analyzing and Synthesizing Expressive Piano Performance with Graphical Models, *Machine Learning Journal*, Vol. 65, No. 2-3, pp. 361–387 (2006).
 [6] Flossmann, S., Grachten, M. and Widmer, G.: Expressive Performance Rendering with Probabilistic Models,

Guide to Computing for Expressive Music Performance (Kirke, A. and Miranda, E. R., eds.), Springer, pp. 75–98 (2013).

[7] Kim, T. H., Fukayama, S., Nishimoto, T. and Sagayama, S.: Statistical Approach to Automatic Expressive Rendition of Polyphonic Piano Music, *Guide to Computing for Expressive Music Performance* (Kirke, A. and Miranda, E. R., eds.), Springer, pp. 145–179 (2013).
 [8] 奥村健太, 酒向慎司, 北村 正: 芸術的逸脱の生成モデルにおける決定木のバックオフに基づく鍵盤楽器演奏の自動表情付け, 情報処理学会研究報告, Vol. 2013-MUS-99, No. 10, pp. 1–6 (2013).
 [9] Grachten, M., Arcos, J.-L. and López de Mántaras, R.: A case based approach to expressivity-aware tempo transformation, *Machine Learning Journal*, Vol. 65, No. 2–3, pp. 411–437 (2006).
 [10] Seashore, C. E.: *Psychology of Music*, Courier Dover Publications (1938).
 [11] Sloboda, J. A.: The Acquisition of Musical Performance Expertise: Deconstructing the “Talent” Account of Individual Differences in Musical Expressivity, *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. (Ericsson, K. A., ed.), Lawrence Erlbaum Associates, Inc., chapter 4, pp. 107–126 (1996).
 [12] 奥村健太, 酒向慎司, 北村 正: 楽譜に基づく鍵盤楽器演奏の統計的モデル化手法, 情報処理学会論文誌, Vol. 54, No. 4, pp. 1288–1301 (2013).
 [13] Odell, J. J.: The Use of Context in Large Vocabulary Speech Recognition, PhD Thesis, Cambridge University (1995).
 [14] Kataoka, S., Mizutani, N., Tokuda, K. and Kitamura, T.: Decision-tree backing-off in HMM-based speech synthesis, *Proceedings of the 8th International Conference on Spoken Language Processing*, Jeju Island, Korea, pp. 1205–1208 (2004).
 [15] Bellman, R. E.: *Dynamic Programming*, Princeton University Press (1957).
 [16] Grachten, M. and Widmer, G.: Explaining musical expression as a mixture of basis functions, *Proceedings of the 8th Sound and Music Computing conference*, Padova, Italy (2011).

^{*4} CrestMusePEDB version 2, <http://www.crestmuse.jp/pedb>