

楽譜情報のコンテキストクラスタリングに基づく 演奏者のモデリング手法

奥村 健太^{†1} 酒向 慎司^{†1} 北村 正^{†1}

本稿では、実演奏 MIDI データから演奏者の特徴を確率モデルとして学習するため、芸術的逸脱の要素を楽譜の情報を用いたコンテキストクラスタリングにより効率的に分類する手法を提案する。これによって演奏者が楽曲に対して独自に持つ解釈の内容をモデルに反映させ、その内容を観測することが可能となった。また、演奏者識別実験による検証でも提案手法の有効性を確認できた。

A technique for making musical performer's model with context-clustering of musical scores

KENTA OKUMURA,^{†1} SHINJI SAKO^{†1}
and TADASHI KITAMURA^{†1}

This paper presents about the method to classify elements of artistic deviations by context-clustering with information available from musical scores to learn musical players' model efficiently. The effectiveness of the proposal method was able to be confirmed by the player identification experiment.

1. はじめに

同じ楽曲を示す同じ楽譜からでも演奏者によって異なる演奏が生まれ、聴く人はその違いを知覚し、楽しむことができる。このような演奏の個人性を計算機上で数理的に扱うことができれば、演奏者の識別や個人性を備えた演奏生成などへの応用が期待できる。

そのような違いをもたらすものを観測し、目的を持って扱うためには楽器演奏における人間性を数的に表現する必要があり、その試みは長きにわたってなされている。1980年代には KTH model [Sundberg, 1983]¹⁾ や Todd model [Todd, 1985]²⁾ などのルールベースの手法が考案されたが、1990年代には Mazzola model [Mazzola, 1990]³⁾ や machine learning model [Widmer, 1995]⁴⁾ などのように数学的な知見からのアプローチがなされるようになった。近年では演奏中の発音タイミングの振る舞いを HMM によってモデル化した手法 [Hamaoka, 2003]⁵⁾、Gaussian Process によって楽譜と演奏者の関係を表現した手法 [Teramura, 2008]⁶⁾、その表現に離散 HMM を用いた手法 [Flossmann, 2009]⁷⁾ など、多様なアプローチが試みられている。

我々は演奏者の個人性に焦点を置き、それを特徴付ける要素の一つである芸術的逸脱を実演奏 MIDI データから取得し、その傾向を確率モデルによって学習する手法に取り組んできており¹⁰⁾、データの周辺環境を考えたモデル化が効果的であるという知見が得られている。演奏者は演奏に必要な様々な情報を楽譜から参照して演奏を行っており、演奏された全ての音符はそれぞれが演奏者の楽曲解釈を内包していると言える。逸脱はその演奏者の楽曲解釈によって変動の傾向を統制・制御されていると考えられ、モデル化に際してそれらの影響を取り込む必要がある。そこで本稿では逸脱の振る舞いや演奏者の楽曲解釈を捉える手法として、演奏者モデルの学習過程において楽譜から得られる情報を利用したコンテキストクラスタリングの導入を検討する。

2. 演奏者の特徴とモデル化

ここでは本稿で用いる手法の概要を述べる。芸術的逸脱とその抽出手法、モデル化について説明する。次に楽譜から得られる情報をコンテキストとして導入し、それを用いて演奏者独自の楽曲解釈をモデルに反映する手法について論じる。

2.1 芸術的逸脱とその要素

演奏者による演奏は、楽譜通りに演奏されたとしても厳密には楽譜通りではなく、各自の解釈による芸術的な表現を目的とした逸脱（ずれ）が生じている。このような逸脱は“Artistic Deviation”（芸術的逸脱）と呼称される⁸⁾。これはただ楽譜を機械的に再生するのみの演奏との決定的な違いであり、実演奏において現れる人間性への影響は大きい。

本研究で対象とする鍵盤楽器において演奏者が演奏に際して制御できるのは、打鍵の時間と強弱のばらつき、離鍵までの時間による音の長さのばらつき、ペダルによる弦の共振の変化である。以上の中で数量化の容易な MIDI による扱いが可能なものとして、本研究では

^{†1} 名古屋工業大学
Nagoya Institute of Technology

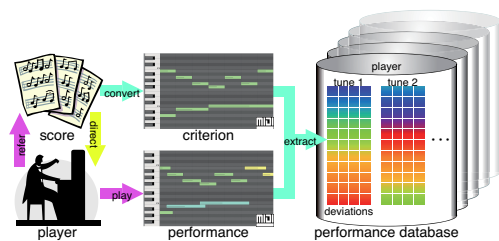


図 1 逸脱量の抽出
Fig.1 Extracting artistic deviations

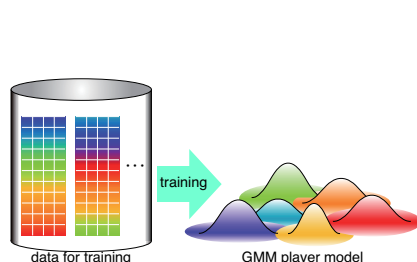


図 2 GMM による演奏者のモデル化
Fig.2 Player modeling with GMM

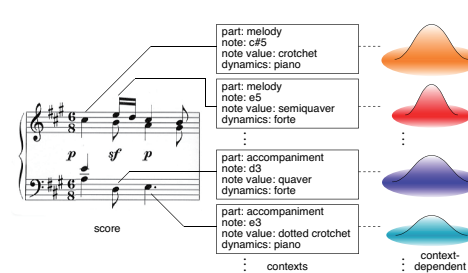


図 3 楽譜から得られるコンテキスト
Fig.3 Contexts available from scores

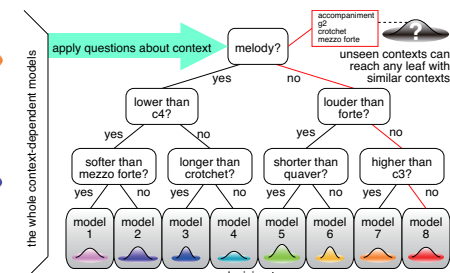


図 4 決定木に基づくコンテキストクラスタリング
Fig.4 Context-clustering based on decision tree

打鍵タイミング, 離鍵タイミング, 打鍵強度, そして発音時間の 4 種類の逸脱要素を扱う. 図 1 に示すように, 楽譜を厳密に MIDI データ化したものを正解データとし, 各音符について実演奏 MIDI データとの比較を行うことで逸脱量の抽出を行う. これによって音符ごとに逸脱量が定まる.

2.2 演奏者の特徴のモデル化

実演奏データにおける逸脱量から特定の演奏者を表現するためには, 「その人物の振る舞いの特徴」をモデル化する方法が必要になるが, これは容易に数量化が可能なものではない. 一方, 音声の認識や合成の分野では, 複雑な変動を伴う音響信号を扱うために有効な方法として様々な変動を確率的な揺らぎとして捉えた確率モデルがよく用いられている. 演奏者の個性が表れるものとして逸脱の振る舞いがあり, そのゆらぎを観測することで演奏者の個性を観測できるとした場合, これらを類似した問題として扱えると我々は考えた. 先行研究では混合正規分布モデル (Gaussian Mixture Model: GMM) を用いてそれぞれの演奏者が独自に持つ逸脱量の揺らぎをモデル化する手法を提案している⁹⁾. その手法におけるモデル化の過程を図 2 に示す. 実演奏データから 2.1 節で述べた手順に従って抽出した逸脱量を観測ベクトルとし, その傾向を混合分布として学習している.

2.3 コンテキスト依存モデルの導入

演奏者のモデル化において, 本研究では演奏者の個性のみではなく, その演奏が持つ音楽的な意味も包含することも目標としている. そのためには何が必要か. ここではある楽曲の実演奏データとその楽譜が与えられたとき, それらをどのように活用するかを考えたい.

まず, 演奏者の個性が現れるものとして逸脱の振る舞いがあり, これは何らかの音楽的な要因によって制御されていると考えられる. その要因の一つとして考えられるものに演

奏者個人が演奏する楽譜に対して独自に持つ解釈がある. 一方, 演奏者は楽譜を見て演奏を行っている. その結果なされる演奏は必ずしも厳密に楽譜通りではないということから, 楽譜から得た情報に自らの解釈を付加して演奏を行っていると言える. ここでは, それは演奏者が得ているのと同様に楽譜から得られる情報によって表現が可能なものとする. 以上から, 逸脱を制御する要因として楽譜の情報があることと, これを扱うことで演奏者が持つ楽曲の解釈を表現可能であることを仮定する.

尚, 演奏者の解釈が具体的にどのようなものかは演奏者自身にも説明が困難な抽象的なものであったりする場合が多く, 実際にそれをそのまま扱うことは困難であるが, 形式上簡単でも違いを区別できるようなものを扱えば近似的にそれを表現することが可能であると考えられる. 音符に連動して楽譜から容易に得られる情報として, 例えば単独の音符では音名, 音価, 強弱などがあり, 音符が前後で関係を持つことで成立する楽句などがある. 便宜上, ここではそれらをコンテキストと称する. 本稿では図 3 に示すように楽譜から各音符についてコンテキストを付与する.

2.4 決定木に基づくコンテキストクラスタリング

より多種のコンテキストを扱うことにより逸脱の制御要因である楽譜の情報をより詳細に扱えるようになると言えるが, 同時に学習データに関する弊害が生じる. 使用するコンテキストの種類が増加すればその組合せの数は指数的に増加し, モデルごとに確保できる学習データは著しく減少する. また, あらゆるコンテキストの組合せを学習データのみで賄うのは現実的に困難であるため, その中に存在しない組み合わせが必要となった場合にモデルの生成が不可能になる. 尚, 今回設定した条件ではコンテキストの組合せの総数はおよそ 2.9×10^{12} 種類にも及ぶ. このような量のコンテキストをそのまま扱うのは現実的ではない.

以上の問題を解決するため、ここでは決定木を用いたコンテキストクラスタリング¹¹⁾を導入する。この手法は主に音声合成の分野で音素ごとに振る舞いの異なる話者の特性を効率的に扱うために用いられている。クラスタリングでは、モデルはコンテキストに関する質問によって分割され、その結果は以下の手順で構築される二分木構造によって表される。

- (1) 全てのモデルをまとめ、ルートノードとする。
- (2) その時点で全てのリーフに対して予め用意された楽譜情報に関する質問を全て適用し、最小記述長基準により最尤なリーフと質問のセットを選択する。
- (3) 分割停止条件を満たしていれば分割を終了する。
- (4) リーフを二分し、新たなリーフを作成する。元のリーフはノードとして手順(2)の質問を保持する。
- (5) 手順(2)に戻る。

楽譜の情報を用いて決定木を構築していく際の過程を図4に示す^{*1}。決定木の各ノードは分割に用いた質問を保持するため、ルートノードからそれぞれの質問に従って木を辿ることで、未知のコンテキストに対しても類似する既知のコンテキストが属するリーフノードを一意に決定することが可能になる。

また、ここでコンテキストを扱う目的の一つに演奏者の楽曲解釈をコンテキストの組合せによって近似することが挙げられるが、その内容がいかに充実していてもその数が過剰になると意味を成さない。従ってそれらを演奏者の解釈として意味を持つ程度にまで分類する必要があるが、分類に際してどのような基準を用いればよいかをその都度人為的に決めることは非常に困難である。しかも、全く同一のコンテキストでもその背景や周囲の条件によって異なる振る舞いが発生しうる。演奏者それぞれの個性を考慮する上でも特定の場合における解釈の仕方を特定のものに決め打つことは妥当とは言えないため、その違いに柔軟に対処できる手法としてもコンテキストクラスタリングは有効である。これによってモデルが扱い易くなるとともに、演奏者の解釈を可視化し、モデルの次元での比較が可能になる。

2.4.1 最小記述長基準

クラスタリングでは最小記述長基準 (minimum description length criterion: MDL 基準) を用い、式1に示す目的関数 F が最小となるようにモデル M の共有構造を決定している。

$$F = -\log P(O|M) + KN \log \Gamma(S_0) \quad (1)$$

ここで $\Gamma(\cdot)$ は各クラスタの学習データ量、 K は特徴ベクトルの次元数、 N は分割後のクラ

*1 ここでは簡略化のため音符単体のみに対するコンテキストと質問を例示している。

表1 収録に使用した楽曲と楽譜
Table 1 Tunes and scores used for recording

ID	Composer	Title	Score
B1	J. S. Bach	Two Part Inventions, BWV 772-786	G. Henle Verlag, HN64
M1	W. A. Mozart	Twelve Variations, K. 265	Wiener Urtext Edition, 8
M2	W. A. Mozart	Piano Sonata No. 11 in A Major, K. 331	Wiener Urtext Edition, 227

スタの総数、 S_0 は決定木のルートノードを表す。クラスタを分割し N を増加させると観測系列における対数尤度 $\log P(O|M)$ を減少させることができるが、 $KN \log \Gamma(S_0)$ が増加するため、必ずしも F が最小になるとは限らない。両者にあるトレードオフの関係によって、モデルの汎化性とデータに対する追従性のバランスを保っている。

モデル M においてクラスタ S を質問によって S_1 と S_2 に分割する場合、分割前後の MDL 値の変化量は次式のようになる。

$$\Delta_q = \log P(O|MS_1) + \log P(O|MS_2) - \log P(O|MS) + K \log \Gamma(S_0) \quad (2)$$

分割は式2において、 $\Delta_q < 0$ を条件に行われる。ここで分割前後の尤度の変化量の閾値は $K \log \Gamma(S_0)$ となり、この値によって生成される決定木の規模が制御される。

3. 実験

先述の手法に従ってモデルの学習を行い、演奏者識別実験によりその評価を行う。また、得られた決定木からモデルが演奏者の解釈をどのように捉えているかを検証する。

3.1 実演奏 MIDI データの収集

芸術的逸脱を数量的に扱えるものとして得るため、ピアノ演奏を専門とする複数の演奏者に電子ピアノの演奏を依頼し、実演奏 MIDI データの収録を行った。収録に際して、楽曲は以下の点を考慮して選択した。

- 一般的に認知されたクラシック・ピアノ独奏楽曲であること
- 各楽曲の長さ（音符数）や難易度に差が少ないこと
- 作曲家ごとに一貫性が保たれていること

また、逸脱量の算出基準となる楽譜は各作曲家の作品を収めたものとして国内のピアノ演奏教育用途において一般的に用いられているものを使用している。後述の実験で用いる楽曲セットとしての ID を併記し、内容を表1に示す。演奏者にはこれらの楽譜に表記のある指示には従った上で、各自の解釈を持って演奏するように依頼している。また、変奏曲については変奏それぞれを1曲と見なしている。

3.2 コンテキストの設計

コンテキストクラスタリングにおいて、楽譜からどのようなコンテキストを使用するかは任意に決定できる。ここでは以下のように定義する。

3.2.1 音符に対する分類の定義

一般的な大譜表から音符について得られる情報として、音名や音価、強弱などがある。今回は各音符に付加できる情報として以下の4種類のコンテキストを定め、それぞれについて一定数の分類を設定した。

声部： 2種類（主旋律とそれ以外を大譜表の上下段および演奏する手の左右などから判断）

音名： 88種類（A0=21, ..., C4=60, ..., C8=108）

音価： 16種類（装飾音符：0, 三十二分音符：1, ..., 四分音符：8, ..., 全音符：15）

強弱： 8種類（*ppp*：0, ..., *mp*：3, ..., *mf*：4, ..., *fff*：7）

尚、音符が前後に関係を持ち楽句となった場合も考慮するため、当該音以外にもその先行音・後続音についても声部以外で当該音と同様に分類を行っている。従ってこれらの組合せの総数は $2^1 \times 88^3 \times 16^3 \times 8^3$ 種類となるが、モデル学習に使用できるデータが限られているため、現実的にその全てが同時に生じることにはほばない。

3.2.2 コンテキストの組合せを分類する質問の定義

今回決定木の作成において使用したコンテキストに関する質問を以下に示す。これも音符の分類と同様に、音符が前後に関係を持ち楽句となった場合も考慮するため、当該音の情報のみではなく、その先行音・後続音を持つ情報についても定めている。

- { 当該 } 音符の声部値 (melody)
- { 先行・後続 } 音符との距離 (distance) の { 範囲・該当・大小 }
- { 先行・当該・後続 } 音符の音名 (notenumber) の { 範囲・該当・大小 }
- { 先行・当該・後続 } 音符の音価値 (gatetime) の { 範囲・該当・大小 }
- { 先行・当該・後続 } 音符の強弱値 (velocity) の { 範囲・該当・大小 }
- { 先行・当該・後続 } 音符の属するオクターブ (octave) の { 範囲・該当・大小 }

3.3 モデル学習

コンテキストクラスタリングを行うことにより、少量のデータによる学習でも汎化性の高いモデルが得られると考えられる。今回は3.1節で収録した実演奏データから、as, jn, mt, rs, ty の5名の演奏者のデータを使用し、B1, M1, M2 それぞれについて学習データが各楽曲セット全体の約1/2および約1/4になるように楽曲単位で楽曲セットを2分割または4分割し、モデル学習を行った。

3.4 演奏者識別実験

3.4.1 演奏者の識別

先述の5人の演奏者の実演奏データに含まれる逸脱の傾向を演奏者ごとにモデル化したGMMモデルセットに対して、演奏者全員の実演奏データから抽出した逸脱量を観測ベクトルとして与える。それぞれの尤度を比較し、最も尤度が高いテストデータとモデルの基になった演奏者とが一致した場合を正解とする。これをテストデータの数だけ行い、それらの平均正解率を識別率とする。

3.4.2 実験条件

未知データに対する客観的な評価を行うため、テストセットとして3.3節で定めた学習データと同一の楽曲セット内で学習に使用していない楽曲をテストデータとする。内訳は以下になる。

B1 1/4 data: BWV772-775, 776-778, 779-782, 783-786

B1 1/2 data: BWV772-778, 779-786

M1 1/4 data: Theme and var.1-3, var.4-6, var.7-10, var.11-12 of K.265

M1 1/2 data: Theme and var.1-6, var.7-12 of K.265

M2 1/4 data: Theme and var.1, var.2-3, var.4 and 6, var.5 of K.331

M2 1/2 data: Theme and var.1-3, var.4-6 of K.331

この実験ではコンテキストクラスタリングの効果を検証するため、コンテキストに依存しないモデル化手法との比較を行う。具体的な手法として以下の2種類を試行する。両者はいずれも学習時に音符の分類を行うが、楽曲の音楽性を考慮するか否かが異なる。

従来法 GMMの混合数を変化させてモデル学習を行う。楽譜の情報は参照しない。

提案法 観測ベクトルに楽譜から得られる情報を与え、コンテキストクラスタリングによるモデル学習を行う。決定木の大きさを制御することでリーフ数を変化させるが、コンテキストクラスタリングの効果のみを観測するため、GMMの混合数は1で固定とする。

3.4.3 実験結果

各テストセットにおける識別結果の平均値を図5に示す。学習データが少ないほど提案法が有効にはたらし、その差が広がっていることがわかる。モデルの空間内よりも楽譜の情報を参考にクラスタリングを行う方が未知データへの対応をうまく説明する上で有効であると言える。

次に、学習データ量を楽曲セットの約1/2とした場合における両手法の混同表を表2に示す。特に従来法において、多くのセットでも識別率の高い演奏者と低い演奏者が分離して

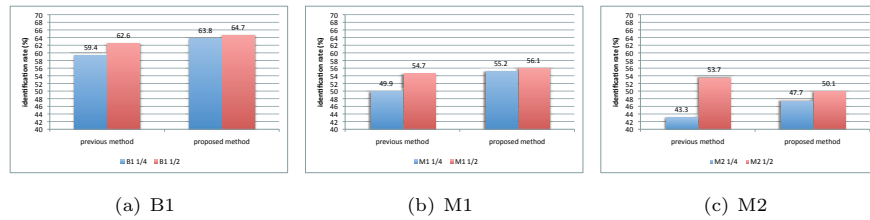
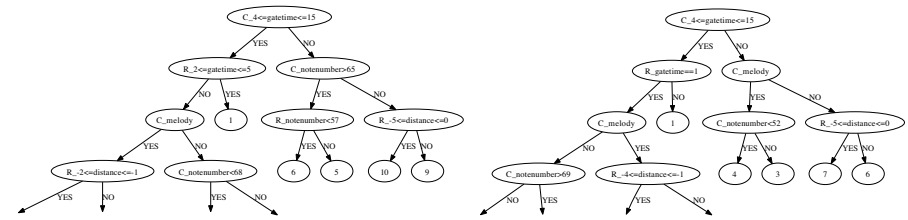


図 5 各テストセットにおける識別率 (%)
Fig. 5 Identification rate for each test set (%)

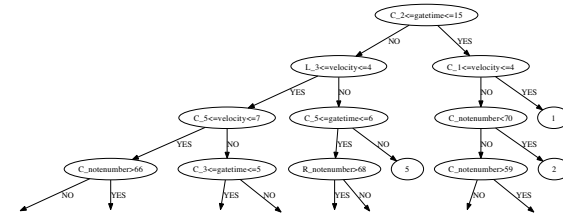
表 2 学習データを 1/2 としたテストセットにおける混同表
Table 2 Confusion matrix for each half-trained test set

Test set	Model	Previous method					Proposed method				
		as	jn	mt	rs	ty	as	jn	mt	rs	ty
B1 1/2	as	10	1	1	2	1	12	0	0	2	1
	jn	0	7	1	6	1	0	6	1	6	2
	mt	1	0	5	1	8	2	0	6	2	5
	rs	0	0	1	13	1	0	0	0	15	0
	ty	0	0	2	0	13	0	0	1	0	14
M2 1/2	as	4	3	2	4	0	8	0	3	2	0
	jn	1	4	0	6	2	1	4	0	4	4
	mt	3	0	7	2	1	2	0	8	2	1
	rs	2	1	3	7	0	2	0	2	9	0
	ty	1	0	0	0	12	1	0	0	1	11
M2 1/2	as	2	3	2	4	0	4	1	0	2	0
	jn	1	1	0	6	2	1	3	0	2	1
	mt	3	0	5	2	1	2	0	5	0	0
	rs	2	1	3	2	0	0	1	0	6	0
	ty	1	0	0	0	6	1	0	1	0	5

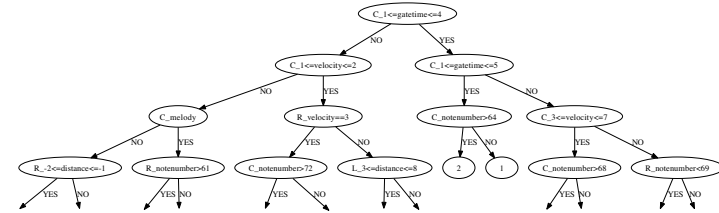
いることがわかる．後者は本人よりも前者として識別されているような場合が多く見られ，その特徴が別の演奏者に吸収されていると言える．一方提案法では，楽曲 M1 や M2 でそのような傾向が解消され，従来法では本人と識別されていなかった演奏者が本人として識別された例が増加している．各演奏者のデータに音楽的な意味付けがなされている効果が出ていると考えられるが，楽曲 B1 では両手法の間で差異が少ない．ここから楽曲セット B1 は演奏者間で解釈の差の出にくい傾向があるのではないかと考えられるが，これはコンテキストクラスタリングの際に生成された決定木を観測することで確認できる．



(a) Data set B1, player jn (b) Data set B1, player rs



(c) Data set M2, player jn



(d) Data set M2, player rs

図 6 各データセットにおける決定木の例
Fig. 6 Decision trees for each data set

3.5 決定木の例

B1, M1, M2 の各楽曲セット全曲をそれぞれ学習データとして演奏者モデルを作成した．3.2.2 節の質問を用いた楽譜情報のコンテキストクラスタリングによって作成された決定木の中から，ここではそれぞれ代表として jn と rs の 2 名分の例を図 6 に示す*1．尚，図中では質問内容をノード名として，音符の { 先行・当該・後続 } を { L・C・R } で表す．また，

*1 ここでは 4 段目までを表示し，それより下位のノードは省略する．

その他の数値の意味については 3.2.1, 3.2.2 節で定義している。これらの決定木を見ると、各演奏者が持つ振る舞いの傾向の違いがそれぞれの決定木のノードに選択された質問の違いとして現れている。

データセット B1 を学習データとした場合の決定木は、両演奏者でその形状が酷似しており、同じ箇所のノード間でもほぼ同様の質問が選択されている。従って B1 の楽曲は少なくとも両者にとってはほぼ同様の解釈による演奏がなされたものと考えられる。しかし、コンテキストクラスタリングの考え方としては同一と見なせるモデルがマージされるかどうかはクラスタリングによって判断されるべきであるため、今回設定した分類基準や質問はこの楽曲における演奏者間の違いを表現するには十分なものではなかったと言える。

その他のデータセットを学習データとした場合は B1 ほどの類似性はなく、その形状やノードの内容で演奏者間に細かな違いが見られる。特に下位に位置するノードで選択された質問は演奏者ごとに異なり、このようなコンテキストの組み合わせによって演奏者ごとに異なる細かな反応の違いが説明できると考えられる。ここから、この手法を生成モデルとして利用できる可能性が挙げられる。

また、上位のノードで選択された質問には異なる演奏者間で類似したものも見られるため、そのような演奏指示が各演奏者の振る舞いを分岐する大本の基準となっていることがわかる。これは、演奏者個人の特性というよりも楽器を演奏する上で基本的な感覚に近い暗黙の常識と言えるものではないかと推測できる。ここから、多様な楽曲に対応したモデルの汎用化が期待できる。

更に得られた知見として、同じ演奏者でも楽曲が異なれば生成される決定木の構造が異なることが挙げられる。ここから、傾向の異なる楽曲を同時に複数用いてモデル学習を行った場合でも楽曲による違いに対応したモデルを生成できる可能性があると考えられる。

4. ま と め

本稿では、演奏者の持つ逸脱の変動要因を楽譜情報を用いて効率的に分類するクラスタリング手法について検討した。演奏者識別実験による検証の結果、少量のデータによるモデル学習においてその識別率からコンテキストクラスタリングの導入による効果が確認できたが、現状では従来法を安定して上回る結果とはなっておらず、傾向が類似した演奏者を明確に分離できるまでには至っていない。また、楽譜の演奏指示などにも未対応のものがあること、演奏者間に相違がない決定木が生成される場合があることなど、楽曲の持つ音楽性への追従はまだ充分ではない。

更に、現実的な問題として同一の演奏者でも楽曲によって逸脱の傾向が異なることから、どのような楽曲でも個人性を表現できることが必要であり、様々な楽曲から汎用なモデルを生成することが課題である。その解決のため、今後は扱うコンテキストの内容をより充実させ、演奏や楽曲が持つ音楽性についての考慮を深めていく必要がある。

謝辞 本研究で使用している実演奏データベースのデータ収録に際しご協力をいただいたピアノ演奏者の方々に感謝する。

参 考 文 献

- 1) J. Sundberg, et al.: *Musical performance. A synthesis-by rule approach*, Computer Music Journal, 7, 1983.
- 2) N. P. M. Todd: *A model of expressive timing in tonal music*, Music Perception, 3, 1985.
- 3) G. Mazzola: *Geometrie Der Töne*, Elemente der Mathematischen Musiktheorie. Basel: Birkhäuser Verlag, 1990.
- 4) G. Widmer: *A machine learning analysis of expressive timing in pianists' performances of Schumann's "Träumerei"*, Proc. KTH Symposium on Grammars for Music Performance, 1995.
- 5) M. Hamanaka, et al.: *A Learning-Based Jam Session System that Imitates a Player's Personality Model*, Proc. IJCAI, August 2003.
- 6) K. Teramura, et al.: *Gaussian Process Regression for Rendering Music Performance*, Proc. ICMP, August 2008.
- 7) S. Flossmann, et al.: *Expressive performance rendering: introducing performance context*, Proc. SMC, July 2009.
- 8) C. E. Seashore: *Psychology of Music*, Dover Publications, New York, 1967.
- 9) 奥村 健太, 酒向 慎司, 北村 正: GMM に基づく演奏者の音楽的逸脱のモデル化, 電子情報通信学会, March 2009.
- 10) 奥村 健太, 酒向 慎司, 北村 正: 楽譜情報のコンテキストクラスタリングに基づく実演奏データからの演奏者モデリング手法の検討, 音響学会秋季研究発表会, September 2009.
- 11) J. J. Odell: *The Use of Context in Large Vocabulary Speech Recognition*, Ph.D thesis, Cambridge University, 1995.