

楽譜を考慮した演奏者の統計的モデル化手法の 改善と演奏者識別による評価

奥村 健太^{†1} 酒向 慎司^{†1} 北村 正^{†1}

個性を備えた演奏は、楽譜の演奏指示に対して演奏者の意図による逸脱が加わることで生じると考えられる。我々はそのような逸脱の振る舞いが確率的な偏りを持つと仮定し、実演奏とその楽譜に基づいて両者の依存関係を統計学習する手法を提案してきた。従来法では楽譜から扱える演奏指示の種類が限られており、演奏指示の限られる楽曲には柔軟な対応ができないことが課題となっていたが、本稿では音符単位で付与できる演奏指示情報を拡充し、より楽曲に適合したモデルの生成について検討した。また、演奏者モデルの木構造についてもより効率的な生成手法を検討し、得られた演奏者モデルに未知の演奏における演奏者を識別させる評価を行った。これによって提案手法の有効性を検討し、今後の課題についても述べる。

Improvement in stochastic performer modeling reflects musical score, and its evaluation by performer identification

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

A unique performance is caused because deviating by the performer's intention joins the instruction of the score. We assumed that have bias with a stochastic behavior of such deviating. And we have already proposed a technique to train dependence of both as stochastic models from performance and its score. In a past proposal, performance instructions dealt from the score were not enough, and being not able to do flexible correspondence became a problem in music from which the performance instruction was limited. In this paper, the generation of the model with high adaptability to music was examined. Moreover, a more efficient generation technique was examined about the tree structure of the performer model, and the evaluation to identify the performer in an unseen performance was done to the trained performer model. The effectiveness of the proposal technique is examined, and future tasks are described from results.

1. はじめに

同じ楽曲を示す同じ楽譜からでも演奏者の違い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。個々の演奏が持つ個人性が何に由来するかを追求することは、何が演奏に人間らしさを与えているのかを知ることに繋がる。そのような問題を扱う試みとしては、1980年代のルールベースの手法¹⁾²⁾に始まり、1990年代には数学的な見地からのアプローチ³⁾⁴⁾がなされるようになり、近年では統計モデルの導入など、更に多様な手法が考案されている⁵⁾⁶⁾。

同一の楽譜からなされる演奏が演奏者によって異なったものとなるのは、楽譜の演奏指示に演奏者独自の表現が加わるためである。ここで楽譜を単純に発音する場合と演奏者による音楽的解釈が加わった演奏との比較を考えると、両者の違いは芸術的表現を目的とした逸脱⁸⁾として観測され、演奏者の個人性はここに凝縮されていると言える。逸脱が個人性を表すからには、その振る舞いには何らかの偏りがあると考えられる。我々はそのような逸脱の振る舞いが確率的な偏りを持つと仮定し、既に実演奏 MIDI データとその楽譜に基づいて両者の依存関係を統計学習する手法を提案してきた。これまでに得られた知見によって、演奏者の個性をモデル化する手法としての可能性が示されている⁷⁾。

本稿では、このモデル化手法の改善と、演奏者の識別による性能の評価に取り組む。任意の演奏が誰のものであるか、あるいは誰の演奏にどの程度近いものであるかを定量的に評価できれば、一般的な識別システムなどの他に、模範演奏と何がどの程度違うのかななどを明示できるなど、演奏者の育成支援に寄与する技術としての用途も期待できる。

2. 演奏者の統計的モデル化手法

先行研究では、混合正規分布モデル (Gaussian Mixture Model, GMM) を用いて実演奏における逸脱量の発生傾向を統計モデルとして学習する手法を提案している。概要を図 1 に示す。この手法は鍵盤楽器を対象としており、演奏者が楽器を物理的に制御可能な要素 (音符単位の打鍵、離鍵、押鍵長、打鍵強度) を電子ピアノから得られた実演奏 MIDI データとその演奏の際に用いた楽譜から抽出する。各要素における逸脱量を観測し、それらの平

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

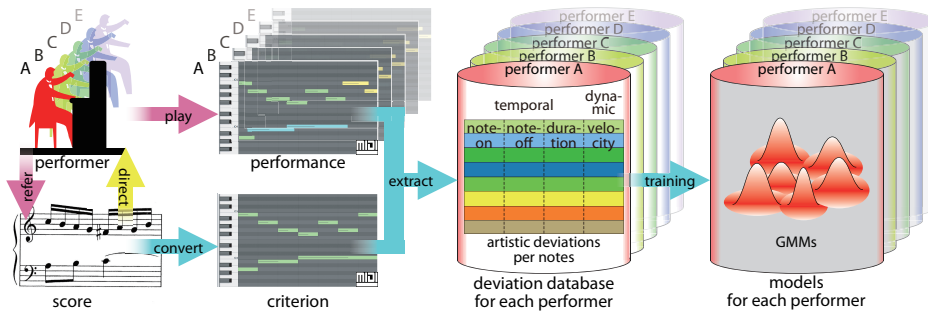


図 1 統計モデルによる演奏者のモデル化
Fig. 1 Stochastic performer modeling

均および分散の傾向を統計学習することで演奏者の個性をモデル化している。以降の本節では、この手法における課題および改善のための手法の提案を行う。

尚、同様の手法で人間の個性を扱う話題として、音声関連の研究が先行している。音声を表す音響信号も逸脱の振る舞いと同様に複雑な変動を伴うが、その変動を確率的な揺らぎとして捉えるために GMM のような統計モデルが頻繁に用いられている。それらを用いた手法は多数の用途において実績があり、その一例として、話者の話し方や声質などからそれが誰の声を識別する問題を話者の個性をモデル化することによって解くものがある。本稿でモデルの評価手法とする演奏者識別もそれと同型の問題と捉えることができる。

2.1 コンテキスト依存モデルの導入

GMM に基づくモデル化手法では、学習データの各次元における音符毎の逸脱量がどのように振る舞うかを任意の個数の正規分布を線形結合したものとして学習している。この手法はデータ系列の複雑な動態を自動的に学習可能である点では有効であったが、モデルの学習において参照されるのは学習データとなる逸脱量の数値のみであるため、各データが音楽的にどのような意味を持つためにそのような振る舞いをしているのかは考慮されていない。

学習データは楽譜に従って発音された音符の系列であり、それぞれが音楽的な意味を持つことは自明である。一般的な例として、楽譜上で f と指示された音符のを演奏した場合と p と指示された音符を演奏した場合とを考えた場合、両者の逸脱量の間にはそれを理由とした発音のタイミングにおける差異がある他、その強弱指示があった場合の平均からどれだけ更に強弱方向に逸脱するかについても何らかの個性が影響すると想定される。従って、逸脱を扱う際にはデータの数値のみを参照するのではなく、どのような条件で発された音符

がどのような振る舞いをするかを音楽的な知見に基づいて分類するべきであると言える。

逸脱の発生には複数の要因が複雑に関与していると考えられる。ここでは逸脱が生じる過程を次のように考える。

- 一般的に楽譜に記された楽曲を演奏する際、演奏者は楽譜から得た情報に自らの楽曲解釈などの意図的な要因を付加して演奏を行っている。
- そこで発生した逸脱の傾向は、演奏者自身の身体的な特性などによる非意図的な要因が加わることで更に制御される。

後者の影響による逸脱の制御は比較的単純な傾向に基づくと考えられる。また、一般的に演奏技術の上達は非意図的な要因による意図的な演奏表現への干渉を低減することを意味する。従ってこれによる影響は誤差と言える範囲に近付くと言えるため、ここでは明示的には取り上げないことにする。一方前者は人為的に与えられるものであり、一般的に演奏技術の上達に従って変化幅が広がるため、こちらが逸脱の発生に与える影響は大きいと考えられる。しかし、その基準は演奏者自身にも説明し難いような抽象的なものである場合が多い。従って、本研究では演奏の解釈そのものを扱うのではなく、解釈に影響を与えているであろう楽譜中の情報を演奏情報中の逸脱と関連付けることにする。

ここでは楽譜から一意に得られる情報を用いることを考える。そのようなものの例として、単独の音符では音名、音価、強弱等があり、それら音符が前後で関係を持つことで成立する楽句等がある。これらの情報が演奏における解釈に寄与する文脈のようなものを形成すると考え、便宜上コンテキストと称してそれらを導入する。音符単位で予め音名のみによる粗い分類でモデル学習を行った上でコンテキストの組合せを付与し、その後はそれぞれの組合せ毎に固有のモデル（コンテキスト依存モデル）と見なしてパラメータを更新する。

2.2 コンテキストの拡充

コンテキストは実演奏と照合される楽譜の MIDI データから抽出する。本稿で以降の実験に使用するコンテキストの一覧を表 1 に示す。

これまでで使用していたコンテキストの種類は音名 n の他、 D , L , M のような比較的基本的なもののみである。このため、先行研究では楽譜中に演奏指示情報の種類が少ない楽曲などではストリーム毎の木構造の差異が明確にならず、学習データを複数のストリームに分割する効果が十分に得られていない可能性が提示されていた。クラスタ分割の際に適用できる質問がより多種類の観点から選択できれば、より学習データにおける逸脱を分類する基準として相応しい決定木を得られると考えられる。ここでは楽譜から一意に得られるという前提はそのままで考えうるほぼ全ての情報を使用することで、ストリーム毎に選択される質問

表 1 コンテキストの定義
Table 1 Definition of contexts

| Context | description | resolution, range or example |
|---------|-----------------|---|
| n | note number | 88 notes |
| A | time signature | 2/2, 3/4, 4/4, 6/8, ... |
| B | key signature | 12 keys with major or minor scale |
| C | beat per minute | from 30 to 240 |
| D | part | melody, polyphony, accompaniment, ... |
| E | global position | amount of passed time in the tune (%) |
| F | local position | amount of passed time in the measure (%) |
| G | interval | distance from previous or to next note |
| H | on to on | interval between note on and note on by midi clock |
| I | off to on | interval between note off and note on by midi clock |
| J | scale degree | unison, minor 2nd, ..., major 7th |
| K | range | 8 octaves |
| L | note value | 16 lengths (breve, minim, crotchet, quaver, ...) |
| M | dynamics | as note on velocity in 8 levels |
| N | accidental | whether accidental is added or not |
| O | key color | white or black |
| P | chord tone | top, middle, bottom or individual |
| Q | pedal action | whether each 3 pedals are stomped at note on timing |

に明確な違いが出るような決定木の生成を目指す。尚、今回の定義では楽譜内のあらゆる音符に対してユニークなコンテキストの組合せが定まるため、モデル学習時に出現した組合せの総数は学習に使用した音符数と同等となる。その分類基準は延べ約 25000 種類の質問となり、クラスタ分割の度にその中から最尤なものが選択される。

2.3 コンテキスト依存モデルの Tree-based クラスタリング

楽譜から得られる情報をより詳細に表現するためには多様なコンテキストを扱う必要がある。しかし、使用するコンテキストを増加させると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が著しく減少する。また、学習データのみで想定される全てのコンテキストの組合せを網羅することは難しい。演奏者の識別を行う場合、学習時には未知な楽曲の演奏を識別する需要も考えられるため、有限なデータであらゆるデータに対応可能なモデル学習手法が必要である。また、コンテキストの導入の目的は演奏者の楽曲解釈をその実演奏における逸脱動態と関連付けることにある。従って、膨大なコンテキストの組合せをそのような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るのが適切であるかをその都度人為的に決めることは非常に困難である。

そのような問題に対処するため、本研究では Tree-based クラスタリング⁹⁾を用いてコン

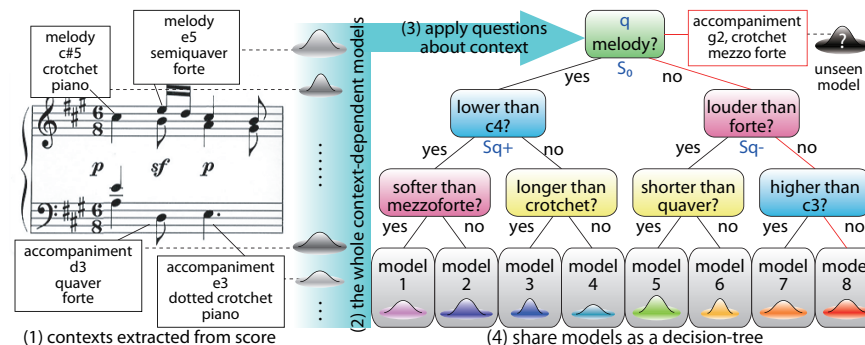


図 2 コンテキスト依存モデルと決定木に基づくクラスタリング
Fig. 2 Context-dependent models and clustering based on decision tree

テキスト依存モデルを分類する。本手法の概要を図 2 に示す。尚、ここで導入するコンテキスト依存モデルでは当該する音符だけでなく先行、後続の音符が持つ情報までほぼ同様に考慮するため、コンテキストの定義によっては類似した条件で発生した音符もそれぞれほぼ唯一のものと思なすことができる。楽譜上では同一の演奏指示が与えられた音符でも実演奏においてはその周囲の条件によって異なる振る舞いが発生しうるが、本手法はそのような違いにも必要に応じて個別の対応が可能となる。

2.3.1 MDL 基準に基づく Tree-based クラスタリング

先述の Tree-based クラスタリングでは、最小記述長 (Minimum Description Length, MDL) 基準¹⁰⁾に基づいてクラスタの分割を行う。親クラスタ S に対して質問 q を適用し、子クラスタ S_{q+} , S_{q-} に分割する際、分割前後の記述長の変化量 Δ_q は次式で表される。

$$\Delta_q = \frac{1}{2} \{ \Gamma(S_{q+}) \log |\Sigma_{S_{q+}}| + \Gamma(S_{q-}) \log |\Sigma_{S_{q-}}| - \Gamma(S) \log |\Sigma_S| \} + K \log \Gamma(S_0) \quad (1)$$

ここで、 $\Gamma(S)$ はクラスタ S の学習データ量、 Σ_S はクラスタ S の共分散行列、 S_0 は決定木のルートノード、 K はクラスタ S の学習データの次元数である。分割は式 (1) において、 $\Delta_q < 0$ を条件に行われる。また、分割前後の尤度の変化量の閾値は $K \log \Gamma(S_0)$ であり、この値によって生成される決定木の規模が制御される。

2.3.2 各ストリームへの Tree-based クラスタリング

本稿のモデル学習では学習データを 2 つのストリームでモデル化し、時間要素 (temporal) ストリームと強弱要素 (dynamic) ストリームで個別に Tree-based クラスタリングを行う。

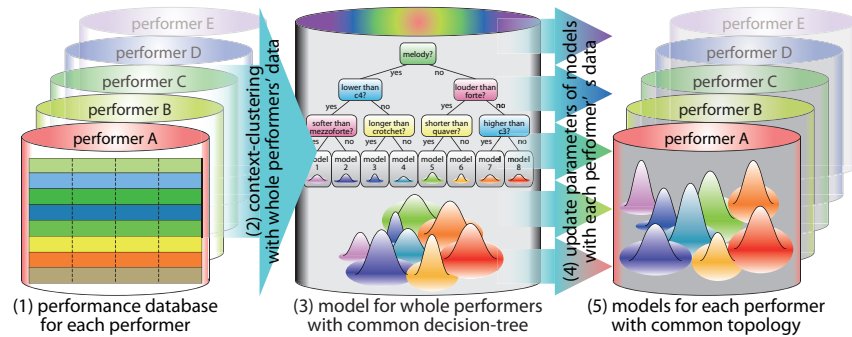


図 3 モデルの木構造の学習

Fig. 3 Training tree structure of the models

学習データの各次元の相関が小さい場合、学習データを複数のストリームでモデル化し、各ストリーム毎に Tree-based クラスタリングを行うことで、モデルを部分的な結びつきで共有化できる。学習データ内の 4 つの次元について、打鍵、離鍵、押鍵長は時間軸上で表現されるが、打鍵強度は時間には直接的な関係が無く、表現上の尺度が異なる。特に前者の逸脱量は演奏者に、後者の逸脱量は楽曲に依存した分布を持つという傾向が先行調査で明らかになっている。また、同時に学習データ内のストリーム分割を複数検討した結果、先述の分割手法がモデルの性能向上に最適となるという知見が得られており、本稿ではそれに倣う。

2.4 木構造を共有する演奏者モデル

今回の演奏者モデル生成の概要を図 3 に示す。演奏者の識別を目的とした演奏者モデルの生成を考えた場合、未知の演奏者のデータに対して学習済みの各演奏者のモデルとの尤度比較を行うことで学習済みの演奏の中では誰に尤も近いかを判別することになる。特定の個人性を備えた演奏を生成するためのモデルを考える場合は個人に特化した詳細な木構造を持つモデルが望ましいが、異なる木構造を持つモデル同士では平等な条件での比較が困難になる。今回の場合はモデル間での比較が評価の主体となるため、同時に比較されるモデルは同一の木構造を持ち、パラメータの値のみが異なるような条件を設定する。

3. 実験

先述の手法に従ってモデルの学習を行い、従来手法からの改善点を検証する。また、演奏者識別実験によってモデルの性能を評価する。

表 2 収録に使用した楽曲と楽譜
Table 2 Tunes and scores used for recording

| ID | composer | title of collection | score | notes |
|----|--------------|-------------------------------------|------------------------|-------|
| B1 | J. S. Bach | Two Part Inventions, BWV 772-786 | G. Henle Verlag, HN64 | 8838 |
| M1 | W. A. Mozart | Twelve Variations, K.265 | Wiener Urtext Ed., 8 | 6150 |
| M2 | W. A. Mozart | Piano Sonata No.11 in A Maj., K.331 | Wiener Urtext Ed., 227 | 8166 |

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演奏者 5 名による実演奏 MIDI データを収集した。楽曲は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、楽曲間で長さ（音符数）や難易度に差が少ないこと、作曲家毎に一貫性が保たれていることなどを考慮して選択している。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。それらの内容を楽曲集毎にまとめ、表 2 に示す。各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。よって以降の実験に使用する全ての実演奏には楽譜に対する明らかな演奏ミスなどは存在しない。音符数は表 2 中の数と一致しており、異なるのは音符毎の逸脱量のみとなる。

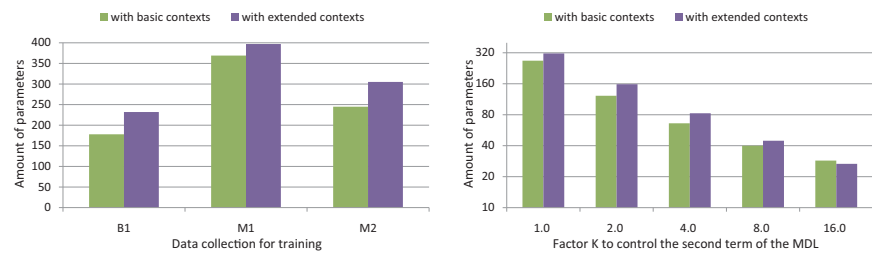
3.2 モデルの木構造

B1, M1, M2 の各楽曲集をそれぞれ学習データとしてモデルの学習を行い、木構造を生成した。各ストリームに含まれる学習データの次元数を D 、各ストリームで生成された決定木のリーフノードの数を L とすると、今回の場合モデルが持つパラメータ数 P は次式によって得られる。

$$P = (D_{temporal} \times L_{temporal}) + (D_{dynamic} \times L_{dynamic}) \quad (2)$$

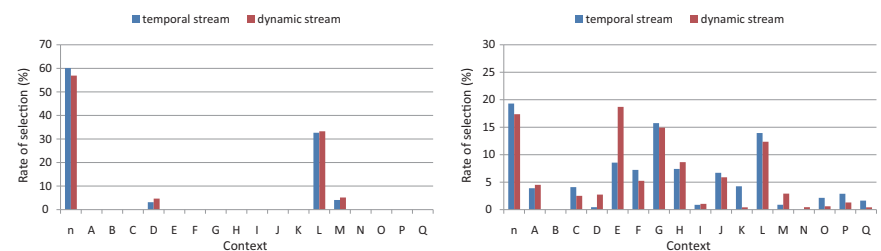
コンテキスト拡充前後でのモデルのパラメータ数の変化を図 4 に示す。図 4(a) から、学習データによって規模の違いはあるが、コンテキストの拡充によって全ての場合でモデルのパラメータ数が増加し、木構造が複雑化していることがわかる。従って過学習の懸念があるため、後述の識別実験では式 (1) の第 2 項、 $K \log \Gamma(S_0)$ で表される MDL 基準の閾値を変化させることによって木構造の規模を調整しながら性能の評価を行うこととする。各場合における木構造の規模は図 4(b) のようになる。モデルのパラメータ数は概ね閾値の係数 K の逆数倍で変化しているが、両者の大小関係は $K = 16.0$ で逆転することが確認できる。

次に、コンテキスト拡充前後で得られたモデルの木構造における質問の選択頻度を、全体 (a, b) と上位 10% (c, d) の 2 通りで図 5 に示す。拡充後の質問は特定の要素に集中するこ

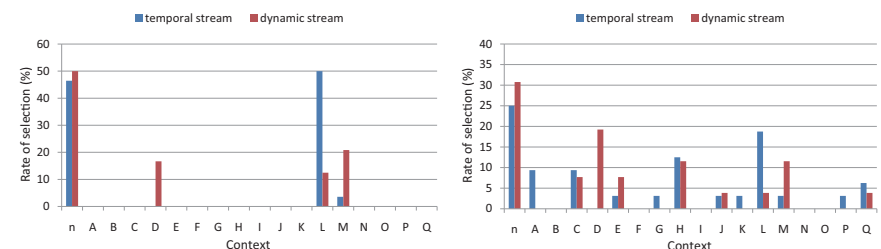


(a) Amount of parameters for each data for training ($K = 1.0$) (b) Amount of parameters for each thresholds for MDL criterion

図 4 モデルのパラメータ数の変化
Fig. 4 Amount of parameters for each models



(a) All questions with basic contexts (b) All questions with extended contexts



(c) Essential questions with basic contexts (d) Essential questions with extended contexts

図 5 コンテキストに関する質問の選択頻度 ($K = 1.0$)
Fig. 5 Selection frequency of question concerning contexts

となく広く分散しており、特に C, E, H など、拡充前のものに近い頻度で選択されているものもある。それらについては拡充前に選択可能であった質問よりも拡充後の質問の方が適当と判断されたと言え、拡充分のコンテキストが逸脱動態の分類基準として有効なものであったと考えられる。以上から、扱うコンテキストを多様化させたことで、ストリーム毎に Tree-based クラスタリングを行う本来の意図に近付いたと言える。

尚、拡充済コンテキストを用いたモデル学習で得られた決定木の一部を図 6 に示す。いずれも $K = 16.0$ の場合のもので、質問の内容は表 1 に準じる。併記の”L, C, R” は質問の対象となるコンテキストの位置（先行, 当該, 後続）を示し、関係するコンテキストによってノードの配色を変えてある。先述の通り、どの場合でもストリーム毎に質問の選択傾向が偏っており、いずれかのストリームにしか出現しない質問が見られる。しかし、同じストリームでも学習データによって木構造の形状は異なる。特に楽曲集 B1 と他 2 種の相違は大きく、学習データの影響が強く反映されることが伺える。本稿での演奏者モデルは演奏者全員に共通の木構造を持たせたが、共通の木構造のモデルを比較するという観点から、学習データの特徴が異なっても各ストリームのモデルは近い木構造を持つことが理想である。コンテキストの拡充によってモデルの学習データへの依存度が高まったことで、この目的の達成には学習データの組合せを吟味することが重要になっていると考えられる。また、モデルに反映させる情報として今回有効であったコンテキストを選別し直すことに加え、楽曲の特徴を越えて演奏者の個人性のみを捉えるような要素を考慮していく必要がある。

3.3 演奏者識別実験

モデルの学習と識別時の提示に使用する楽曲の組合せを表 3 に示す。尚、演奏者の候補は 5 人であるため、ランダムに選択した場合の識別率は 20% となる。学習に使用するデータ量を曲集全てから約 1/2, 1/4 へと制限しながら楽曲集や作曲者についてクローズまたはオープン条件を試行するが、学習に使用した楽曲データは識別時に提示しない。学習済みの各演奏者モデルに未知の楽曲データを与えて尤度を比較、モデルとデータに関して演奏者が最尤となった場合を正解とし、各セットの全試行における平均正解率を識別率とする。

今回は以上の手順で以下の 2 通りの実験を行う。

- (a) 木構造の規模の変化による性能の変化を検証する。Tree-based クラスタリングによるクラスタ分割の停止基準は図 4(b) を参考に、 $K = 1.0$ から 16.0 までの範囲で試行する。
- (b) 参考比較として木構造を演奏者毎に共通化することによる性能の変化についても検証する。先行研究で生成した演奏者毎に個別の木構造を持つモデルを使用するため、試行は拡充前のコンテキストのみによる Tree-based クラスタリングを $K = 1.0$ で分割停止

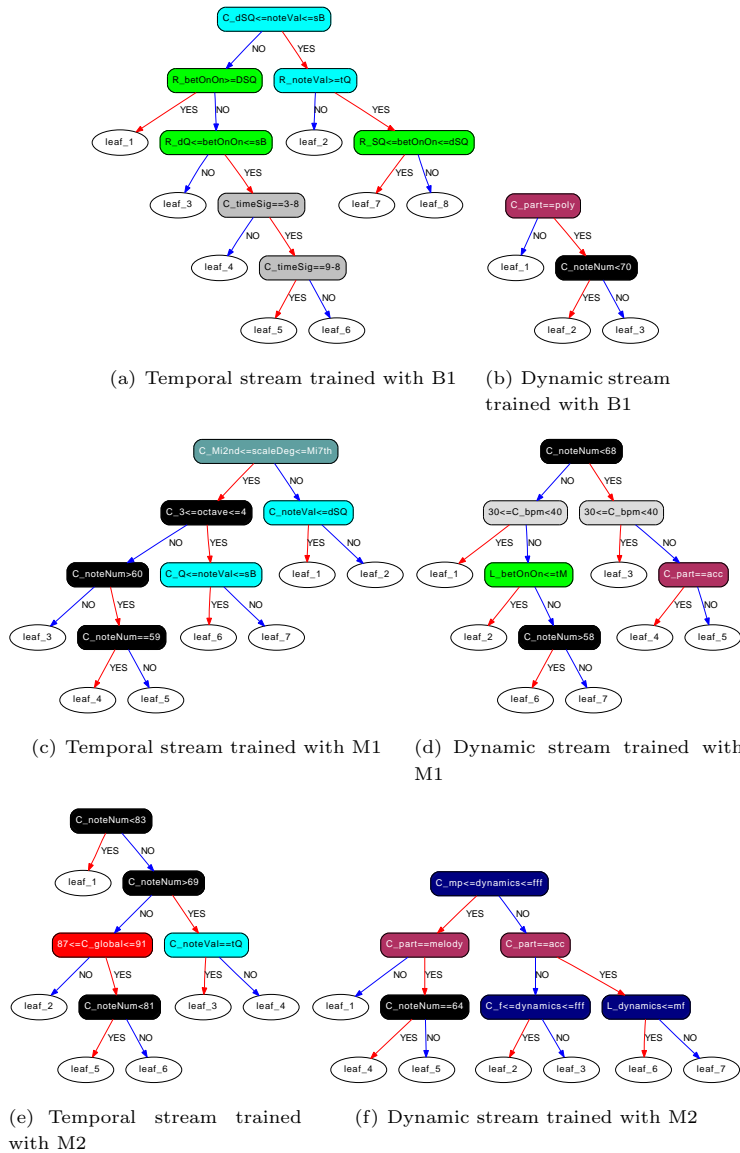
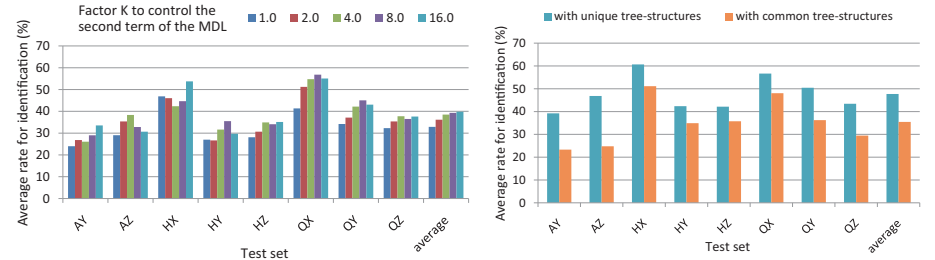


図 6 各楽曲集でのモデル学習における決定木の例
Fig. 6 Examples of decision-trees for model-training with each collections

表 3 演奏者識別評価用テストセット
Table 3 Test sets for evaluation by performer identification

| Set | data for training | data for test | collection | composer |
|-----|------------------------|---|------------|----------|
| AY | all of each collection | all of same composer's other collection | open | close |
| AZ | all of each collection | all of other composer's collections | open | open |
| HX | 1/2 of each collection | unseen 1/2 of same collection | close | close |
| HY | 1/2 of each collection | all of same composer's other collection | open | close |
| HZ | 1/2 of each collection | all of other composer's collections | open | open |
| QX | 1/4 of each collection | unseen 3/4 of same collection | close | close |
| QY | 1/4 of each collection | all of same composer's other collection | open | close |
| QZ | 1/4 of each collection | all of other composer's collections | open | open |



(a) Varying size of common tree-structures with extended contexts
(b) Varying tree-structures with basic contexts ($K = 1.0$)

図 7 演奏者識別実験結果
Fig. 7 Results for performer identification

した場合には限定する。

3.4 識別性能の評価

識別実験の結果を図 7 に示す。先述の条件に従い、(a) では共通の木構造を持つモデルで木構造の規模を変化させた場合を比較、(b) では演奏者毎に個別の木構造を持つモデルの場合と共通の木構造を持つモデルの場合を比較している。

3.4.1 (a) 木構造の規模による性能の変化

まずテストセットごとの傾向を見ていく。セット X では楽曲集をクローズとした条件で識別を行っており、ランダム選択に対して概ね 2 倍から 3 倍の識別率となっている。同一の楽曲集の場合、演奏上で類似した振る舞いが多く、楽譜の演奏指示上でも類似したものが多くなる。そのためこの条件での結果が良好となるのは想定通りである。これに対して楽曲

集や作曲家をオープンにしたセット Y や Z の識別率は明らかに低く、場合によってはランダムに選択した場合に近い結果となっている。一般的に同一の演奏者でも楽曲が異なれば演奏スタイルを変えるのは当然と考えられるが、そのような違いがあっても残るような演奏者の個人性を今回生成したモデルで捉えるのは困難であったと考えられる。

次に木構造の規模について見ていく。全体的に木構造をある程度単純にすることで結果が改善されていく傾向がある。また、セット H と Q の結果について比較すると学習データ量が多いほど識別率が下がっている傾向があることから、学習データが多い、または木構造の規模が大きいモデルは過学習となっている可能性が高い。しかしその傾向が全ての場合で確かなのは $K = 4.0$ までであるため、どの程度の規模の木構造や学習データ量の設定がこの問題に最適であるかについてはより詳細な調査が必要となる。

3.4.2 (b) 木構造の生成手法による性能の変化

各演奏者モデル間で入力データに対する尤度を比較する際には各個のモデルの差異がより大きい方が有利なのは明らかであり、共通の木構造を持つとしたモデルでは個別の木構造を持つモデルには及ばないと想定されたが、図 7(b) から、モデルの木構造によって明確な違いが出る結果となった。

尚、図 7(b) の共通木構造の場合と (a) の $K = 1.0$ の場合と比較すると、特にテストセットとしてデータの類似度が高い X、または学習データ量が多い H において、コンテキストの拡充による性能の低下が見られる。この傾向は木構造の単純化によって緩和されているため、先述の検証と同様に過学習の影響と考えられる。よって本稿で行ったコンテキストの拡充が識別率としての結果に明確に貢献できるものであるかについて、この結果からはポジティブな判断はできない。また、木構造の生成手法が異なる場合でもコンテキストの追加による性能の変化があれば同様に反映されることが考えられるため、両者の性能差を覆すのは容易ではない。演奏者の識別を考える場合、モデルが個人に特化することを優先する必要はないため、共通の木構造である程度の性能を保障できればそれで問題はない。しかし今回得られた結果からは実用を想定できる性能ではないため、現時点では演奏者毎に個別の木構造を持たせることはモデルの性能上不可避な条件となると考えられる。

4. ま と め

本稿では、実演奏に含まれる楽譜からの芸術的逸脱の動態に楽譜の情報を反映させて統計学習を行い、獲得した演奏者モデルを演奏者の識別に用いて性能の検証を行った。

モデル化手法の改善として、従来から懸案となっていた楽譜情報の考慮内容を拡充した。

これによって、ストリーム毎に生成するモデルの木構造の構築において楽譜の演奏指示の情報量による干渉を低減できるようになった。しかし、演奏者識別による評価の結果から、コンテキストの拡充によって学習データへの依存性が高まり、識別性能に影響を与えることがわかった。現時点では未試行の条件も多いため断言はできないが、演奏者によって既になされた楽曲解釈を柔軟に扱うことと、未知の演奏の内容をその中のいずれかと同定することは、今回用いた手法のみで同時に達成が可能な課題とは言えない可能性がある。

以上から、本手法を演奏者識別に適用するにはモデルの木構造について検討の余地が多く残されており、今後も継続してより適したモデルの木構造を探るなど、更なる検討を要する。また、今回扱ったコンテキストは楽譜から自動的に得られるローカルな情報が多いため、楽曲や演奏者のプロフィールにも関わるグローバルな情報も考慮するなど、まだ捉えきれていない情報も取り入れて検討を進める必要がある。

謝辞 本研究の一部は国立大学法人名古屋工業大学「平成 22 年度学内研究推進経費」の支援を受けて行った。

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 Musikthorie. 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) S. Flossmann, et al.: *Expressive performance rendering: introducing performance context*, Proc. SMC, July 2009.
- 7) 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演奏者のモデリング手法, IPSJ-SIGMUS-83-11, December 2009.
- 8) C. E. Seashore: *Psychology of Music*, Dover Publications, New York, 1967.
- 9) J. J. Odell: *The Use of Context in Large Vocabulary Speech Recognition*, Ph.D thesis, Cambridge University, 1995.
- 10) K. Shinoda, T. Watanabe: *MDL-Based context-dependent subword modeling for speech recognition*, A. Acoustical Society Japan (E), vol. 21, no. 1, pp.70-86, 2000.