

## 事例ベースの演奏表情生成手法に関する研究

鈴木泰山, 徳永健伸, 田中穂積  
東京工業大学 情報理工学研究科 計算工学専攻  
{taizan, take, tanaka}@cs.titech.ac.jp

我々は事例ベースの手法を用いた演奏表情の自動生成の研究を行なっている。この手法は、入力として演奏対象曲の楽曲情報と演奏スタイルなど設定を与えると、あらかじめ用意された演奏データ集の中から対象曲に類似した曲の演奏データを検索し、それらの演奏に見られる演奏表情のパターンを対象曲に適用するというものである。演奏表情パターンを獲得する際には、演奏スタイルの設定などによって個々の事例の重要度を評価するので、演奏スタイルの設定に応じて多様な演奏表情を生成することが可能となる。本稿では事例ベースの演奏表情生成手法の概念と特徴を述べた上で、実際の生成の手順について簡単に説明する。

## Example Based Approach for Production of Expressive Performance

Taizan Suzuki, Takenobu Tokunaga, Hozumi Tanaka  
Department of Computer Science  
Graduate School of Information Science and Engineering  
Tokyo Institute of Technology

We have been developing an example-based approach for the production of expressive musical performance. In this approach, the production system uses music score and performance situation parameters inputs to search a previously-arranged musical performance data set for similar musical sentences, and retrieves examples as sentence performance data. The system then process abstracts a common musical expression pattern from the retrieved examples, and applies this pattern to subject sentence. Since, the significance of each example is evaluated based on the similarity of the sentence and performance situation parameters, this approach can produce varying expressive performances. In this paper, we describe the concept and the merits of this example-based approach, and briefly propose a method for the generation process.

### 1 はじめに

人間が音楽の演奏を行なう時には、テンポを変化させたり特定の音符を強くならすなど様々な演出を加えることで自然な演奏を行っている。楽曲の演奏中にみられるこれらのテンポや音の強さの芸術的逸脱のことを演奏表情と呼ぶ。このような演奏表情は表情豊かで聴者を魅了するような音楽演奏を行なう上で必要不可欠なものである。しかしながら、その与え方や変化量などは楽譜には殆んど記載されておらず事実上全てが演奏者の裁量に

委ねられているといつてよい。演奏表情の与え方は非常に自由であり、演奏者はその知識や経験、趣向に基づいて独自の演奏表情を与えている。自然な音楽演奏における人間の思考活動は直観や経験によるところが多く、またその多くは無意識に行なわれているので、生成の過程や根拠が不明確である。

このような演奏表情を計算機に自動的に生成させ、表情豊かで自然な演奏を行なう手法は過去にも様々なものが提案されている。以前は人間が与えた演奏表情のルールを利用して演奏表情を生成すると

いう手法が中心であったが [5, 6, 7, 9], 最近是人間の演奏データから演奏表情の生成ルールを学習し, それに基づいて未知曲に対する表情付けを行なう手法が主流である. こういったルールの学習を伴った演奏表情の生成手法としては, Widmer が提案した音楽構造との関連づけることで学習する方法 [8] や, 白川, 小田らのエキスパートシステムとニューラルネットワークを組み合わせた手法 [1, 2], 片寄らの重回帰分析を用いた手法 [4] などあげられる.

本研究では演奏表情の自動生成の手法として事例ベースによる手法の提案を行なう. この手法は人間の演奏データのなかから対象曲に似た曲や曲の部分の演奏データを検索し, それらの演奏に見られる演奏表情を対象曲に適用するというものである.

## 2 事例ベースの概念

近年の学習を伴った演奏表情の自動生成の研究の殆んどは, 学習などの手法こそ根本的に違うが, 人間の演奏データの多くに共通してみられる演奏表情の法則性を発見しそれらから汎用の演奏表情生成ルールを獲得しようとするという点で共通している (図 1). 演奏データの分析から演奏表情の生成を行なう過程でルールという抽象的な記述を経由しているのど, いずれも広い意味でのルールベースの手法であるといことができる.

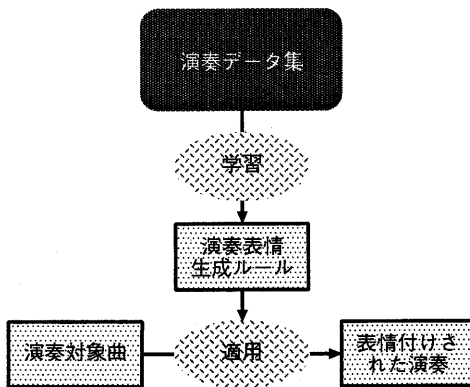


図 1: ルールベースの演奏表情生成

これに対して, 本論文で提案する手法では汎用的な演奏表情の生成ルールの学習を行わず, 参考事例として検索された人間の演奏データを直接的に利

用し演奏表情の生成を行なっている点で先述の広義のルールベースによる手法とは異なる (図 2).

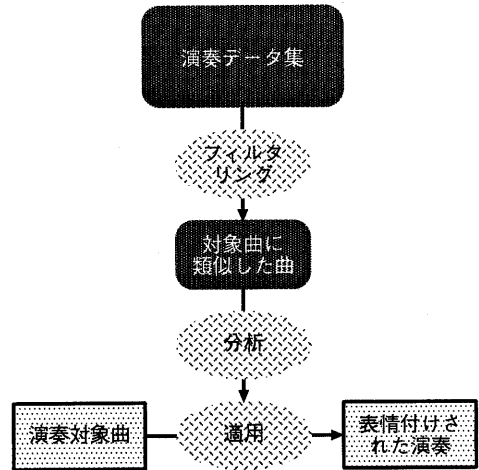


図 2: 事例ベースの演奏表情生成

事例ベースの手法とルールベースの手法はその長所や短所において対極的な性質を持っている場合が少なくない. 演奏表情の生成というタスクに限ってみると, ルールベースの手法では, 一般に学習された演奏表情の生成ルールを人間の演奏表情の生成メカニズムの分析などの音楽学や認知科学などの研究に応用しやすい点や, どのような曲に対してもとりあえず表情付けを行なうことができるという利点があるが, その一方で汎用のルールとして学習を行なうことによって個々の演奏の特徴的な表情づけが失われてしまう点や, 一部の手法ではパラメータによるルールの制御が困難であるため入力に対して一意な演奏表情の生成しか行なえないなどの欠点に伴う. これに対して事例ベースの手法では, 演奏表情の生成規則の解明は非常に困難であり, また参考データが不十分な場合には有効な演奏表情生成を行なうことができないなどの欠点があるが, ルールという形での一般化を行なわないため, 個々の演奏事例の特徴的な部分や例外的な事例を有効に利用することができることや, 参考事例の検索条件を変更することにより一つの曲に対して多様な演奏表情の生成を容易に試みることができるなどの利点がある.

一般に事例ベースの手法では, ルールベースによる手法に比べて取り扱う情報量が非常に多くなるため, 計算機の処理能力が要求されるといわれ, 計算機

の処理能力の限界から、事例ベースの手法は困難であったが、近年の計算機技術の目覚ましい発展にともなって、事例ベースによる手法を計算機上に実装するのさほど困難ではなくなりつつある。

### 3 事例ベースの演奏表情生成

#### 3.1 演奏表情生成関数

演奏表情の生成とは、表情付けされていない演奏データに演奏表情を付与し表情豊かな演奏データに変換をする作業である。演奏表情の与え方は一意ではなく、同じ曲でも演奏時の状況によって演奏表情の与え方が変化する。この演奏時の状況とは、演奏者の経験や経歴、音楽の趣向、演奏時の周囲の状況や心理状態、演奏スタイルなどの様々な要素からなる。以下ではこの演奏表情の与え方を左右する要素を“演奏状況”と呼ぶことにする。

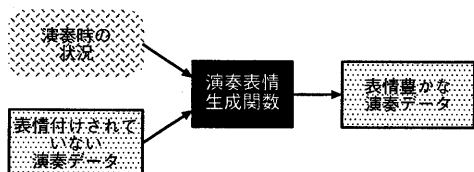


図 3: 演奏表情生成関数

演奏表情の生成を行なうシステムは、演奏データと演奏状況を入力として受けとり、演奏データを出力するブラックボックスと考えることができる。このブラックボックスを演奏表情生成関数と呼ぶ(図3)。ルールベースによる演奏表情生成手法ではたいてい演奏表情生成ルールが演奏表情生成関数の中心になっており、ルールの学習は演奏表情生成関数のチューニングという役割を果たしている。

事例ベースの手法での演奏表情生成関数は、あらかじめ収集された人手による演奏データの中から対象曲に似た曲の演奏事例を検索し、その演奏事例に見られる演奏表情の特徴を分析し演奏対象曲に適用するという機能を持つ。演奏状況に応じて多様な演奏表情を生成できるようにするためには、個々の演奏データには演奏状況情報も記録しておく必要がある。

最も理想的な事例ベースの演奏表情生成関数は、世の中に存在する全ての曲について、演奏状況ごとに実際の演奏データを蓄えているようなデータベース

であろう。このようなデータベースが存在したとすれば、データベース中から演奏状況が入力と一致するような対象曲の演奏事例を検索し出力するだけで演奏表情生成関数を実現することが可能である。しかしながらこのようなデータベースは無限の大きさが必要となるため実現は不可能である。そこで有限の事例からでも任意の曲に対する表情付けが可能となるようなメカニズムを導入する必要がある。

基本的には有限の事例集、つまり演奏データ集の中から対象曲に似ている曲の演奏データを検索し、その演奏データに見られる演奏表情を対象曲に適用することになる。しかしながら、曲全体が対象曲に類似しているというような曲が演奏データ集のなかにある確率は、曲の長さの指数オーダーに反比例して低くなる。逆に演奏対象曲中のごく一部分だけを取り出し、同じように演奏データ集の曲の各部分と比較したとすると、対象となる旋律の長さがくなる上にサンプルの絶対的な数が増えるため、類似した旋律が発見できる可能性が格段に向上すると考えられる。このことは曲全体が類似しているという例は殆んどないが曲中の一部分が別の曲のある部分に似ているという例やよくあるということからも容易に推測できる。

また、曲全体が類似している例を検索するにしても、細かい部分の類似性には拘らず、コード進行や構成といった大局的な特徴だけが類似しているが例の方が、制約が緩い分、類似した事例を発見できる可能性が高くなると予想される。

従って曲全体が類似しているものを演奏データ集から検索するのではなく、曲の部分ごとに制約を緩めて類似している演奏データを検索し、それらを組み合わせ対象曲に適用するのが望ましい。

#### 3.2 演奏表情の段階的生成

ここでは、曲の各部分ごとの参考事例を集約し対象曲の演奏表情を生成するために演奏表情の生成の階層性について考える。

まず最初に、次のような二つの旋律があったとする。

この二つの旋律は音楽的に非常に類似しているといつてよいだろう。各音符の長さは前述の楽譜と全く同じになっているし、音の高さの変化も同じである。唯一の違いは前者がFからAの音で構成されているのに対し、後者はGからBで構成されており一音高くなっている点である。



図 4: 旋律 1



図 5: 旋律 2

さて、この二つの旋律を同じ演奏状況でそれぞれ独立して人間が演奏したとする。状況や条件などによっても違って来るが、旋律内で最も高い音を強く鳴らしたり、あるいは最後の音を長めにのぼし前半をやや速めに演奏するなど、なんらかの表情付けをした演奏が行なわれるであろう。いずれにせよこの二つの旋律の演奏表情は非常にかよったものになることが予想される。



図 6: 旋律 3

次に図 6 のような旋律を考える。この旋律は旋律 1 と旋律 2 を連結したものである。この旋律を先ほどと同じ条件で人間が演奏したとする。前述の 2 つの旋律の例から類推すれば、前半と後半をいずれも先の旋律 1 や旋律 2 と同じように演奏することが予想されるが、実際にはそうならないことが多いであろう。いずれの部分も単独で演奏した場合の雰囲気や少なからず残しているものの、それぞれ旋律 1 や旋律 2 を単独で演奏したときとは多少変化した演奏になる。一般的には前半より後半の方がテンポを速くし音も強くしたような演奏になるであろう。

旋律 3 の演奏表情が旋律 1 や旋律 2 と異なる原因は、旋律 3 の前半部分を演奏している時と後半部分を演奏している時の演奏状況の違いにあると考えられる。前半部分を演奏しているときには演奏者はその後

続く後半部分にどうつなげるかを考えながら演奏しているし、後半部分に移ってからは逆に前半部分の続きであることを意識しながら後半部分を演奏している。このような心理状態の変化は旋律 1 や旋律 2 を単独で演奏した時にはなかったものである。つまり曲全体を演奏している時の状況が全く同じであったとしても、それぞれの部分を演奏しているときには旋律 1 や旋律 2 を単独で演奏しているときとは微妙に演奏状況が違っているということである。

このようなことを考慮すると旋律 3 の演奏表情には次のような要素が含まれていると考えられる。

1. 旋律 1 の基本的な演奏表情
2. 旋律 2 の基本的な演奏表情
3. 旋律 1 と旋律 2 が接続したことによる演奏表情

1. と 2. は旋律の前後の接続関係を一切無視して単独でその旋律を演奏したと仮定した場合の演奏表情である。それに対して 3. は二つの旋律が接続することによって生じる演奏表情で、個々の旋律を単独で演奏した時には出現しないものである。旋律 3 の演奏は前半部分に 1, 後半部分に 2 を適用した後に全体に 3 を適用した結果ということである。

さて先述の演奏表情生成関数の定義では入力も出力も演奏データなので、演奏表情生成関数の出力をまた演奏表情生成関数に入力することも可能である。従って旋律 3 の演奏表情生成関数は図 7 のように、演奏表情生成関数の合成関数として記述することができる。

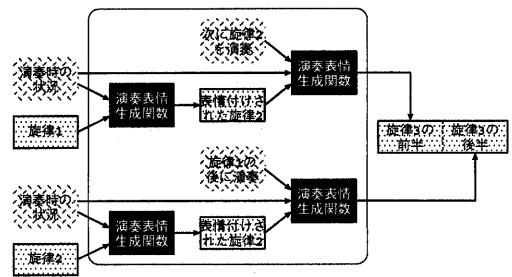


図 7: 旋律 3 の演奏表情生成関数

実際の曲の場合にはこの旋律の前後にさらに別の旋律が接続しそれによって新たな演奏表情が与えられることになるだろうし、また旋律 1 や旋律 2 の基本的な演奏表情も 4 つの音符の接続によって演奏表情のない単音に表情付けしたものになっている。殆

んどの楽曲はフレーズ構造といったような階層をもったツリー構造になっていると考えられるので、こういった曲の各部分の接続関係による演奏表情の変化は曲のフレーズ構造の階層の深さに応じて何度も適用されることになる。

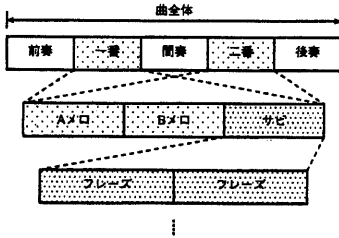


図 8: フレーズの階層構造

これらをふまえ、本手法ではまず対象曲をフレーズ構造に基づいて部分旋律に階層的に分割する。そして、それぞれの部分旋律ごとに演奏表情生成関数を作成し、それぞれの関数を合成することで曲全体の演奏表情生成関数を実現する。

### 3.3 演奏データ集

本手法で利用する演奏データは以下の情報を持っている。

- 楽曲に関する情報
  - 楽譜情報
  - フレーズ構造
  - それぞれの部分旋律の特徴
- 演奏データ
  - 分析済みの演奏表情
- 演奏状況情報
  - 演奏者
  - 演奏スタイル
  - 曲想の主観的評価

楽曲情報は、楽譜情報だけでなくフレーズ構造や部分旋律ごとの特徴などもあらかじめ登録しておく。これにより生成時の計算量を低減することが可能である。演奏データも同様にあらかじめ比率形式で分析しておく [10]。演奏状況情報は、演奏状況に応じて多様な演奏表情を生成する上で必要不可欠なので、これらもあらかじめ演奏データと一緒に登録しておく必要がある。

### 3.4 事例ベースの演奏表情生成関数

ここでは、曲中の各部分の演奏表情生成関数の動作について述べる。なお、フレーズ構造はあらかじめ与えられているものとする。演奏表情生成関数の動作は以下の通りである (図 9)。

1. 演奏データ集から対象に類似した曲の演奏データを検索する
2. それぞれの演奏事例に重み付けする
3. 重みを考慮して演奏表情パターンを求める
4. 対象に適用する

まず最初に、演奏データ集の中から部分旋律を順次とりだし対象曲の部分旋律と比較する。取り出した部分旋律が対象曲の部分旋律に類似している場合は、その演奏データを参考事例に追加する。部分旋律の類似度は各旋律の特徴の近さから求める。旋律の特徴は、リズムや、拍位置、旋律パターン、コード進行などといった要素からなる。演奏データ集の中の各部分旋律の特徴と対象曲の特徴とを比較し、共通する特徴要素を多く持つもの部分旋律を、対象曲に類似した旋律とする。

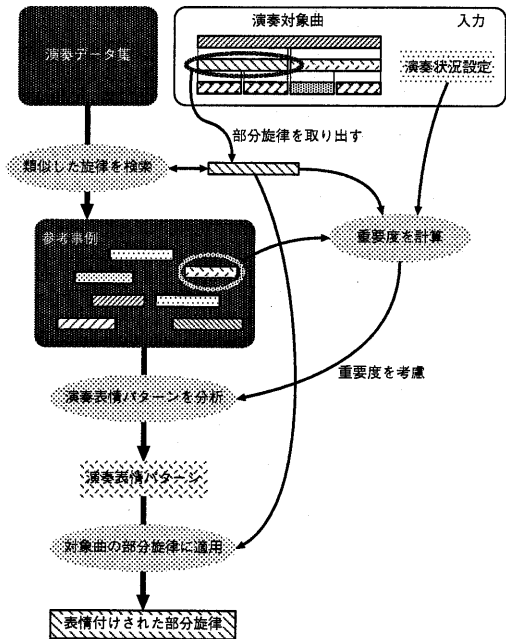


図 9: 部分旋律の演奏表情生成

次に、各事例の重要度を計算する。検索で得られた事例の中には、対象曲と全くおなじ曲とある程度似ている曲が混在しているが、これらの事例を対等に扱うのはあまり得策ではなく、対象曲に近いものほど有効な事例として取り扱うべきである。また、事例の検索の段階では演奏状況を特には考慮していないので、演奏状況があまり近くないような事例も含まれている。演奏状況に応じた演奏表情を生成するためには演奏状況が近くない事例をふるい落とす必要がある。そこで、それぞれの事例についてどの程度有効であるかを表す重要度を計算する。この重要度は以降の演奏表情パターンの計算を行なう時の重みとして利用する。

個々の事例の重要度の計算には以下の二種類の情報を利用する。

- 旋律の類似度
- 演奏状況の近さ

旋律の類似度については、先の類似事例の検索で述べた類似度の評価に加えて、前後に接続する旋律の類似度も評価の対象となる。演奏状況の近さは、入力として設定された演奏状況と検索された事例に付与されている演奏状況情報を比較し、その近さから算出する。

各事例の重要度を計算したら、次に演奏事例に共通して見られる演奏表情パターンの分析を行なう。演奏表情パターンは演奏対象曲の構成要素ごとに抽出する。重要度が高い事例は、目標とする演奏表情に近い表情付けが行なわれているものと考えられるので、事例中に見られる演奏表情を重要度に基づいて傾斜平均し、演奏表情パターンとする。演奏表情パターンは、各音符のテンポや音の強さを比率で表したものである。

最後に、ここで得られた構成要素ごとの演奏表情パターンを演奏対象曲に適用することで、演奏表情が付与された演奏データが作成される。

## 4 おわりに

本論文では事例ベースによる演奏表情の自動生成手法の提案を行なった。この手法ではルールの学習を行わず人間の演奏データから直接的に対象曲の演奏表情生成を行なうため、個々の演奏の特徴的な部分を対象曲に生かすことができ、また一つの曲に対して様々な表情付けを試みるができるという

特徴をもっている。

現在のシステムは、旋律の類似度の評価や重要度の計算を非常に単純な方法で行なっているが、これらの部分は生成される演奏表情の質に多大なる影響を及ぼすと考えられる。そこで、旋律の特徴要素の共通性や相違性と旋律の類似性との関係などを詳細に分析した上で、旋律の類似性や重要度の客観的な評価方法について検討する必要がある。

一般的な楽曲に対してこのような演奏表情生成手法がどの程度有効であるかを確認するためには大規模で良質な演奏データ集が必要となる。しかしながら現在このような用途に利用するための標準的な演奏データ集が構築されたという報告はあまりない。システムの評価などの諸問題を考えると、演奏表情生成の研究者が共同利用できるような大規模演奏表情データ資源の収集、整備が必要であると思う。

## 参考文献

- [1] 白川健一, 小田安彦, 熊谷俊行, 梶川嘉延, 野村康雄, “演奏者情報を加味したピアノの自動演奏システムの構築: ~音符属性に着目した個人的解釈ルールの生成~”, 情報処理学会研究報告, 95-MUS-10, pp. 1-6, 1995.
- [2] 小田安彦, 白川健一, 村上豊, 梶川嘉延, 野村康雄, “演奏者情報を加味したピアノの自動演奏システムの構築: ~ニューラルネットワークによる局所部分に於ける演奏特徴の抽出~”, 情報処理学会研究報告, 95-MUS-10, pp. 7-12, 1995.
- [3] 山内敏彰, 野池賢二, 野瀬隆, 乾伸雄, 小谷善行, 西村怒彦, “Gaを用いたピアノ演奏の局所テンポ解析”, 情報処理学会研究報告, 96-MUS-14, pp. 1-6, 1996.
- [4] 上符裕一, 青野裕司, 片寄晴弘, 井口征士, “演奏ルールの抽出について”, 情報処理学会研究報告, 96-MUS-15, pp.79-84, 1996.
- [5] 田口友康, “自動演奏と演奏モデル:演奏の芸術性を目指して”, bit 別冊, コンピュータと音楽, 共立出版, pp.42-53, 1987.
- [6] 野池賢二, 西岡大祐, 高田正之, 小谷善行, “曲の構造情報から表情付けを行なう自動演奏”, 情報処理学会第44回全国大会, 6U-3, 1992.
- [7] 彌富あかね, 五十嵐滋, “音楽構造分析を用いたピアノ演奏の表情付け”, 情報処理学会第50回全国大会, 7S-6, 1995.
- [8] Gerhard Widmer, “Modeling the rational basis of musical expression”, Computer Music Journal, Vol. 19, No. 2, pp. 76-96, 1995.
- [9] L. Fryden and J. Sundberg, “Performance rules for melodies., origin, functions, purposes”, Proc. of the 1984 ICMC, pp. 221-224, 1984.
- [10] 鈴木泰山, 徳永健伸, 田中穂積, “表情情報の操作を目的とした音楽の構造モデルに関する研究”, 情報処理学会研究報告, 96-MUS-15, pp.67-72, 1996.