

# PEVI : Scapeplot を利用したクラシック音楽の 演奏表情検索・分析インタフェース

三木 翔太<sup>†1,a)</sup> 馬場 隆<sup>†1,b)</sup> 片寄 晴弘<sup>†1,c)</sup>

**概要:** 本稿では、Scapeplot を利用したクラシック音楽の演奏表情検索・分析システムを提案する。Scapeplot は、複数の階層における演奏間の類似関係を一度に可視化する方法として提案された。しかし、最も類似度の高い演奏しか可視化できないことや、どの演奏表情特徴が影響して高い類似度を示しているかまでは可視化できないという問題点がある。ユーザが Scapeplot 上の点をポイントすることでこれらの不可視な情報を別領域に可視化すること、また、演奏表情特徴の重みを変更できることを可能にし、Scapeplot の問題点を解決することで、Scapeplot をインタフェースとして利用した演奏表情を動的に検索・分析できるシステムを実現した。

## PEVI: An Interface Based on the Scapeplot for Retrieving Classical Music Interpretations

MIKI SHOTA<sup>†1,a)</sup> BABA TAKASHI<sup>†1,b)</sup> KATAYOSE HARUHIRO<sup>†1,c)</sup>

**Abstract:** In this paper, a novel interface based on the Scapeplot for finding expressions of classical music is presented. The Scapeplot was proposed to visualize the similarity between performances in multiple layers at the same time, but it can show only the most similar performance, so we cannot know which musical feature effects strongly on the similarity from observing the Scapeplot. We solved these problems by displaying invisible information in another area and enabling to change weights of musical features, and construct a retrieve system for finding expressions of classical music by using the Scapeplot as an interface.

### 1. はじめに

ネットワーク社会の発展とニーズに支えられ、音楽検索技術は大きな進展を遂げてきた。音楽の検索方法には、協調フィルタリングと内容検索の二通りがあるが、特に異なる奏者が同じ楽曲を演奏する機会が少ないポピュラー音楽では、協調フィルタリングを利用した楽曲の検索や推薦の実用サービスが展開されている [1]。しかし、クラシック音楽ではその性質上、ポピュラー音楽に比べて音楽検索技術の恩恵に預かりにくい状況が起こっている。クラシック音楽では、一つの楽曲に複数の演奏が存在し、その演奏表情や

解釈の違いを聞くという楽しみ方があるが、現状で、その楽しみ方をサポートしうる実用の音楽検索システムはほとんどない。前澤らは、ユーザの指揮動作から取得したテンポ情報を利用して解釈の検索を行う Query-by-Conducting [2] を提案した。Query-by-Conducting を用いれば、似たテンポ表現の演奏を検索できるが、音量表現を含めた他の演奏上の特徴が利用できない、という課題がある。

一方で、可視化研究の文脈において、演奏表情を扱ったものがいくつか提案されている。Dixon らの The PerformanceWorm [3]、C.Sapp らの Scapeplot [4][5] がその代表である。The PerformanceWorm では、テンポと音量の時系列経緯を 2 次元のグラフで提示する。Scapeplot では、楽曲間の演奏情報の差異を階層的に比較し、当該の階層において最も似ている演奏名を表示する。Scapeplot は、クラシック音楽の音楽検索応用に向けての可能性を有すると

<sup>†1</sup> 関西学院大学  
Kwansei Gakuin University

a) s.miki@kwansei.ac.jp

b) takashi-baba@kwansei.ac.jp

c) katayose@kwansei.ac.jp

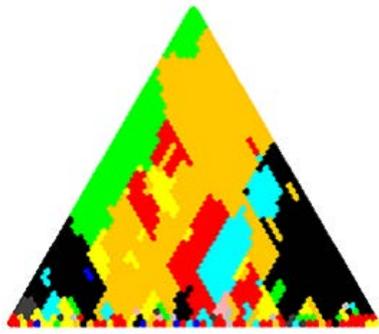


図 1 Scapeplot の例

Fig. 1 An example of the Scapeplot.

考えられるが、1) 類似度 2 位以下の楽曲が直接的にはわからない、2) 特徴量の重み設定を含め、演奏表情検索にむけてのインターフェースが提供されていない、という課題がある。本研究では、これらの課題を解決するものとして、Scapeplot の画面自体を検索用のインターフェースとして利用した音楽検索インターフェース PEVI について提案する。

## 2. 検索インターフェースとしての Scapeplot

図 1 に Scapeplot の例を示す。Scapeplot は三角形の形をしており、底辺は時間軸、左の頂点は曲の始まり、右の頂点は終わりを表す。底辺から特定の範囲が選択された時、その範囲を底辺とする新たな正三角形が描かれる。その正三角形の上の頂点は選択範囲において分析対象と最も類似度が高い演奏を示す色で着色される。

Scapeplot は、楽曲間の階層レベルの類似演奏を表示するユニークな可視化表現であるが、結果の静的な表示しか行わないため、当該レベルでの類似度 2 位以降の演奏が何かが明にはわからない [5]。また、複数の演奏表情特徴を考慮して類似度を算出した時、どの演奏表情特徴が効いて類似度が高くなったのかよくわからないという課題がある。

クラシック音楽では、同一の楽譜から異なる解釈による様々な演奏表情が生まれる。各演奏を特徴付ける要因として、テンポ変化、音量変化、拍内表情、音色等が挙げられる。中でもテンポと音量の特徴は演奏表情に大きな影響を与え、また、比較的正確かつ容易に取得できることから演奏表情の研究にて頻りに利用される [6]。階層構造も演奏表情において重要な要因であり、演奏の局所構造と大局構造でそれぞれ違う特徴を示す場合がある。ゆえに、クラシック音楽の演奏表情検索や分析においては、1. 多数の演奏を一齐比較できること、2. 複数の演奏表情特徴を扱い、それぞれの重みを設定できること、3. 階層構造を扱うことができること、の三点が求められる。

PEVI では、可視化の手段だけではなく、検索用のインターフェースとして Scapeplot を利用し、Scapeplot のみでは不可視な情報をリアルタイムで別領域に可視化する。これにより、動的に演奏表情の検索、分析ができる。また、演

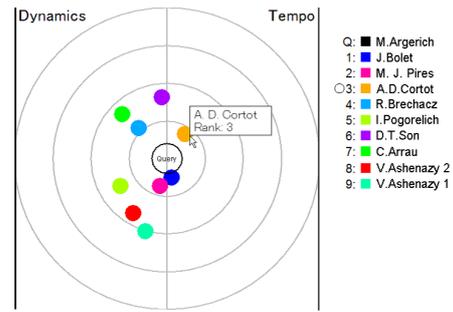


図 2 各演奏のプロットの例

Fig. 2 Target performances plotted on a plane and a ranking list of similarities of the target performances.

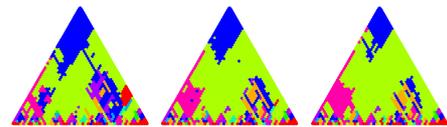


図 3 演奏表情特徴の重みの違いによる Scapeplot の変化。テンポと音量の重みの割合は左から 1:1, 1:2, 1:3。

Fig. 3 The change of the scapeplot by the difference of the ratio of musical features weights.

奏表情について、文献 [3][4][5] においては、テンポと総体的な音量しか扱われていなかったが、加えて、拍内表情 [7] も取り扱えるようにし、各特徴の重みを変更できるようにする。階層構造は Scapeplot それ自身によっても可視化されているが、Scapeplot のみでは不可視であった情報も可視化するため、より詳細に演奏の階層構造を知ることができる。このようにして、上記の 3 つの条件を確保する。

## 3. システム

本システムの使用の流れ、および実装について述べる。PEVI がユーザに提供する主なディスプレイ画面は、Scapeplot を表示する領域と、クエリと各ターゲットの関係を可視化する領域 (図 2 参照) である。以下、ユーザが分析対象とする演奏をクエリ、クエリと比較する演奏群をターゲットと呼ぶ。ユーザがクエリを選択すると、クエリと同一楽曲で省略等の条件が等しいデータベース内の全演奏がターゲットとして選択される。全タイムスケールにおけるクエリと各ターゲットの類似度を算出し、Scapeplot が表示される。この Scapeplot は類似度が最も高い演奏や階層構造を示すだけでなく、インターフェースとしても機能する。ユーザが Scapeplot 上の任意の点を選択すると類似度を比較する範囲と階層が決定し、クエリと全ターゲットの類似関係が即座にもう一つの表示領域に描画される。各演奏は Scapeplot 上での色と同じ色の円で表される。直径は各演奏の拍内表情の類似度に比例する。各ターゲットはクエリとの類似度が高い演奏ほどクエリの近くに描画される。ユーザが Scapeplot 上でポイントをドラッグすると、演奏関係の描画はリアルタイムで更新され、各演奏表情の関係

の推移を動的に見ることができる。各演奏を再生して聴き比べることも可能である。

### 3.1 データベース

演奏表情データベース CrestMusePEDB[8] より演奏表情情報を取得している。CrestMusePEDB は拍内表情までもが記録されているという特徴があり、現在も録音データから手作業でのラベル付けにより構築が進められている。バッハ、モーツァルト、ショパン等のピアノ曲を対象に、名演奏の演奏表情が、テンポと音量の推移は拍単位、MIDI レベルで取得され、MusicXML 形式で記述されている。機械演奏からの各音のズレは演奏表情として捉えられる。

### 3.2 類似度の計算

演奏の類似度を表すためにピアソンの積率相関係数を用いている。相関係数  $r$  は次のように表される。

$$r = \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})}{\sqrt{\sum_n (x_n - \bar{x})^2 \sum_n (y_n - \bar{y})^2}} \quad (-1.0 \leq r \leq 1.0) \quad (1)$$

$x$  および  $y$  は、クエリとターゲットの内の一つの演奏における四分音符単位でのテンポまたは音量変化のデータ列を表す。 $\bar{x}$  と  $\bar{y}$  は、 $x = \{x_i\}$  と  $y = \{y_i\}$  の相加平均である。これらのデータ列がテンポ変化を表す場合は各拍での BPM が、音量変化を表す場合は各拍のベロシティが記録される。CrestMusePEDB は拍内表情として、拍内の一音一音に対して音の開始点や音量のダイナミクスの値が記録されている。これらの標準偏差を求めることで、その拍内での各音のズレ具合を算出し、拍内表情として使用している。ユーザは拍内表情の類似度を算出する際、テンポのみを考慮するか、音量のみを考慮するか、または、両方を考慮するかを選択することができる。

$r = 1.0$  の時、二つの演奏表情は完全な一致を示し  $r = 0.0$  に近いほど類似関係が低いことを表す。 $r$  が負値の時は負の相関を示し、例えば、ある演奏のテンポが早くなっている箇所では他の演奏は遅くなっていることを意味する。

### 3.3 演奏のプロット

二通りの方法で演奏をプロットし、類似関係を可視化する。

まず、二次元座標平面にプロットする方法がある。 $x$  軸がテンポ、 $y$  軸は音量の変化の類似度を示し、クエリは第一象限にプロットされる。例えば、クエリに対してテンポ変化が負の相関を、音量変化が正の相関を示す演奏は第二象限にプロットされる。 $r_t$ 、 $r_d$  をテンポ変化、音量変化の相関係数、 $W_t$ 、 $W_d$  をテンポ変化、音量変化の重みとすると、演奏がプロットされる座標  $(x, y)$  は以下で表される。

$$x = \frac{r_t W_t}{W_t + W_d}, y = \frac{r_d W_d}{W_t + W_d} \quad (2)$$

もう一つの演奏をプロットする方法は、クエリを常に中

表 1 演奏表情特徴の重みの違いによる各演奏の類似度の変化。  
Table 1 Differences of the Pearson correlation values (r-values) by changes of the weights of musical features.

Rank	The ratio of weights of tempo to dynamics					
	1 to 1		0 to 1		1 to 0	
	pianist	r-value	pianist	r-value	pianist	r-value
1st	J.Bolet	0.81	I.Pogorelich	0.86	J.Bolet	0.81
2nd	M.Argerich	0.73	J.Bolet	0.79	M.Argerich	0.80
3rd	D.T.Son	0.65	R.Brechacz	0.71	D.T.Son	0.60
4th	R.Brechacz	0.58	D.T.Son	0.69	A.D.Cortot	0.54
5th	A.D.Cortot	0.58	M.Argerich	0.65	R.Brechacz	0.45
6th	I.Pogorelich	0.53	V.Ashenazy 2	0.65	C.Arrau	0.29
7th	V.Ashenazy 2	0.46	A.D.Cortot	0.61	V.Ashenazy 2	0.27
8th	C.Arrau	0.40	V.Ashenazy 1	0.53	V.Ashenazy 1	0.24
9th	V.Ashenazy 1	0.38	C.Arrau	0.51	I.Pogorelich	0.20

央に描画する方法である。演奏のプロットの例を図 2 に示す。クエリと各ターゲットの距離は相関係数の値に反比例し、テンポの方が音量より相関が高い場合ほど右側に、音量の方がテンポより相関が高い場合ほど左側に描画される。テンポと音量両方の類似度が同じ場合、センターライン上に描画される。クエリと各ターゲットの距離を  $d$ 、各ターゲットのセンターラインからの距離を  $s$  とすると、 $d$  と  $s$  は演奏全体の類似度  $r$  を用いて以下で表される。

$$d = |r - 1.0| \quad (3)$$

$$s = W_t(r_t + 1.0) - W_d(r_d + 1.0) \quad (4)$$

Scapeplot 上の選択点が変わると演奏のプロットはリアルタイムで更新される。

### 3.4 演奏特徴の重みの変更

各演奏表情特徴の重みをスライダーにより変更できる。重みが変わると同時に Scapeplot とプロットされた演奏は更新される。例として、 $W_t$  と  $W_d$  の割合が変更された時の Scapeplot の変化を図 3 に示す。拍内表情の重みを  $W_{wb}$ 、類似度を  $r_{wb}$  とすると、各演奏表情特徴の重みを考慮した演奏全体の類似度  $r$  は以下のように定める。

$$r = \frac{r_t W_t + r_d W_d + r_{wb} W_{wb}}{W_t + W_d + W_{wb}} \quad (5)$$

演奏が二次元座標平面上にプロットされる時、重みの変化に伴いクエリは第一象限内を動く。テンポの重みが 0 の時  $y$  軸上に、音量の重みが 0 の時  $x$  軸上にプロットされる。

## 4. システムの試用

ショパンの前奏曲第 7 番イ長調の様々な奏者の演奏を 10 用意した。内、二つは同一奏者の異なる演奏である。Maria J. Pires の演奏をクエリにし、本システムを試用した。

### 4.1 不可視情報の可視化

同一奏者の演奏は非常に類似するという仮定の元、それら

の演奏を見付けることを試みた。二つの演奏は Scapeplot からは判断できなかった。そこで、Scapeplot 上でポイントをドラッグすると、ポイントの位置に関わらず常に二つ揃って移動する演奏を見付けた。これは、これらの演奏は楽曲のどの部分や階層においても非常に類似した演奏表情を持つことを示す。これらの演奏の奏者を確認すると、同一奏者の演奏であった。可視化による多数の演奏の一斉比較で、同一奏者の異なる演奏を見付け出すことができた。

Scapeplot 上をドラッグしている時、階層により演奏のプロットの分布に異なる傾向が見られた。Scapeplot の上の頂点付近、つまり大局構造比較の時、演奏は一ヶ所に集まるようにプロットされ、全演奏が近い類似度を示した。一方、Scapeplot の底辺付近、局所構造では演奏は散らばってプロットされる傾向があった。これは、大局構造に比べ局所構造の方が奏者の判断が容易である可能性を示唆する。他の演奏表情特徴も考慮出来れば、大局構造での各演奏特徴もより詳細に区別できるかもしれない。

#### 4.2 演奏表情特徴の重み操作

テンポと音量の変化の重みを 0 対 1 に設定した。中間部では Pogorelich の演奏はクエリである Pires の演奏に最も類似していたが、この部分に対してテンポと音量の重みを 1 対 0 にすると、Pogorelich の演奏は離れ、類似度は最も低くなった。各演奏表情特徴の重み操作により、演奏を演奏表情特徴に分解した比較が可能となった。表 1 に各演奏の異なる重みの割合での類似度を表す。

各演奏表情特徴について気に入った演奏を見付けても、奏者が異なる場合が考えられる。その時、例えばテンポは Pogorelich の演奏に似ているが音量は Pires に似ている、というように、ユーザは複数の演奏を混ぜ合わせたいと思うかもしれない。混ぜ合わせた演奏を聞くことができれば、ユーザは本当に気に入った演奏を得られると考えられる。

### 5. 検討

ユーザは本システムを用いて、単なる演奏の類似度比較だけでなく、どの演奏表情特徴がどの部分で似ているか知ることができる。そのため、例えば、弟子と師匠間の演奏の類似を調べることで音楽研究者による演奏家の系譜調査に役立つと考えられる。さらに、Scapeplot 上をドラッグしたり演奏表情特徴の重みを変化させることで演奏表情特徴の推移をリアルタイムで動的に見ることは、プロットの変化の推移から、その演奏表情を特徴付ける重要なカギの発見となり得る。ある演奏が他の演奏に急速に近づく時、それは演奏表情の急激な変化を示す可能性がある。本システムはクラシック音楽初心者にも有用であると考えられる。演奏表情の聴き比べはクラシック音楽の楽しみ方の一つであるが、初心者には難しい。演奏表情の可視化は、演奏表情をより深く理解し比較する助けになると考えられる。

本稿では、CrestMusePEDB を対象として、PEVI の動作について紹介したが、音源分離技術 [9][10] は急速な発達を遂げており、近い将来、音響信号中より一音毎の抽出がより容易になると予想される。録音データから演奏表情の完全自動取得ができれば、現存する全録音データが本システムの検索対象となる。また、自分の演奏とプロの名演奏の比較することで自身の演奏解釈を固めていくといった使用法など、提案システムの利用はより進むと予想される。

今後の課題として、演奏の類似度比較に音色を反映することが挙げられる。音色の利用は、ギター等音色変化に富む楽器の演奏表情比較や、オーケストラ等複数楽器による演奏の楽器パート毎の演奏表情調査に有用と考えられる。

### 6. おわりに

本稿では、Scapeplot を利用したクラシック音楽の演奏表情検索・分析システム PEVI を提案した。Scapeplot 自体をインタフェースとして利用することで、最も類似度の高い演奏しか可視化できない Scapeplot 本来の問題点を解決し、演奏表情を動的に検索・分析できるインタフェースを実現した。今後の課題として、音色等他の演奏表情特徴を類似度計算に反映することがあげられる。

#### 参考文献

- [1] L. Barrington, R. Oda, and G. Lanckriet, "Smarter than genius? human evaluation of music recommender systems," in Proc.ISMIR, 2009, pp. 357-362.
- [2] A. Maezawa, M. Goto, and H. Okuno, "Query-by-conducting: An interface to retrieve classical-music interpretations by real-time tempo input," in Proc.ISMIR, 2010, pp. 477-482.
- [3] S. Dixon, W. Goebel, and G. Widmer, "The performance worm: Real time visualisation of expression based on langner's tempo-loudness animation," in Proc.ICMC, 2002, pp. 361-364.
- [4] C. Sapp, "Comparative analysis of multiple musical performances," in Proc.ISMIR, 2007, pp. 497-500.
- [5] C. Sapp, "Hybrid numeric/rank similarity metrics for musical performance analysis," in Proc.ISMIR, 2008, pp. 501-506.
- [6] M. Grachten and G. Widmer, "Who is who in the end? recognizing pianists by their final ritardandi," in Proc.ISMIR, 2009, pp. 51-56.
- [7] H. Katayose and K. Okudaira, "iFP: A music interface using an expressive performance template," in International Federation for Information Processing (IFIP), 2004, pp. 529-540.
- [8] M. Hashida, T. Matsui, and H. Katayose, "A new music database describing deviation information of performance expressions," in Proc.ISMIR, 2008, pp. 489-494.
- [9] N. Yasuraoka et al., "Musical sound separation and synthesis using harmonic/inharmonic gmm and nmf for phrase replacing system," in Information Processing Society of Japan, vol. 52, no. 12, 2011, pp. 3839-3852.
- [10] M. Nakano et al., "Bayesian nonparametric spectrogram modeling for music signal analysis," in IPSJ SIG Technical Report, vol. 91, no. 6, 2011, pp. 1-8.