

聴感による旋律の演奏表情の類似性評価

-ピアノロール画像の比較による旋律類似性評価結果との比較-

日野達也[†] 鈴木泰山^{††} 野池賢二
徳永幸生[‡] 杉山精^{‡‡}

我々は、事例に基づく演奏表情生成システムにおいて旋律が似ている事例の実演奏を転写する手法を用いている。そのため生成に用いる事例の演奏表情が対象曲に相応しいものがある必要がある。つまり、対象曲のある演奏データにおける演奏表情と事例の演奏表情が類似していなければならない。しかし、旋律の類似性と演奏表情の類似性がどの程度一致するかは明らかとなっていない。そこで、実演奏データのある対象曲を用いてまず旋律類似性評価を行い、旋律類似性の高かった事例に対して対象曲の実演奏との演奏表情の類似性を聴き比べによって評価する実験を行った。その結果、旋律の類似性と演奏表情の類似性が一致する場合とそうでない場合があった。また、旋律類似性評価手法における問題点も明らかとなった。本稿では実験結果と、旋律類似性と演奏表情の類似性に関する考察について述べる。

Similarity Evaluation of Performance Expression by Listening

-Comparing with a Phrase Similarity Evaluation Using Piano roll Image-

Tatsuya Hino[†] Taizan Suzuki^{††} Kenzi Noike
Yukio Tokunaga[‡] and Kiyoshi Sugiyama^{‡‡}

The Case-based expressive performance rendering system generates a performance by copying human performances. So, expression of performance has to be appropriate for target score using for rendering. But a relation is not clear between phrase similarity and similarity of performance expression. To evaluate similarity of performance expression, we conducted an experiment that compare target performance data to sample by listening. The result of experiment shows that similarity of performance expression conform partially to phrase similarity. And some problem of phrase similarity evaluation turned out. In this paper, we report the result of experiment of the evaluation of performance expression and relation between phrase similarity and similarity of performance expression.

1. はじめに

現在の音楽情報科学の主要なテーマのひとつに、コンピュータによる表情の付いた演奏の自動生成がある[1][2]。コンピュータによって表情付けされた演奏をコンテスト形式で評価する試みである Rencon が 2002 年から定期的に開催されるようになり、今までに演奏表情生成システムが数多く発表されている[3][4][5][6][7]。

我々はこれまでに、事例に基づく推論手法を用いた演奏表情生成システム“Kagurame Phase-III”の構築を行った。事例に基づく演奏表情生成では人間によって行われた演奏を事例として、その演奏に見られる演奏表情を転写することで演奏データを生成する。Kagurame Phase-III では、表情付けの対象曲と類似している楽曲の演奏を事例として用いている。表情付けの対象曲との類似性を評価する際には、それぞれの楽譜情報から生成したピアノロール画像を比較する。ピアノロール画像を用いることで楽譜の特徴量を抽象化することなく比較できるため、より多くの特徴を反映した類似性評価が可能となる[11]。また、音高は異なるが概形は類似している旋律は演奏表情の生成に有効な事例であると考え、画像を音高方向にずらしながら比較することで音高によらない旋律の概形の類似性を考慮している[12]。

事例に基づく演奏表情生成ではどのように参照する事例を選択するかが生成結果に大きく関わる。Kagurame システムでは楽譜上で旋律が似ていれば演奏表情も似ているという仮定に基づいて演奏表情を生成している[9][10]。しかし、旋律類似性が高い事例の実演奏データにおける演奏表情が対象曲への表情付けに適しているかは明らかでない。

そこで本稿では、入力楽譜の旋律断片に対して類似度が高いと判断された演奏事例の旋律断片の実演奏を用いて、旋律類似性が高い事例の演奏表情が入力楽譜に対して適切かどうか調査を行った。具体的には、実演奏データが存在する楽曲を対象曲としてシステムによる旋律類似性評価を行い、類似性が高いと判断された事例の旋律断片の実演奏データと対象曲の旋律断片の実演奏データを聴き比べて演奏表情の類似性を評価する聴取実験を行った。聴取実験では 1 つの対象曲の旋律断片に対して事例の旋律断片を 2 つずつ聴き比べ、より対象曲に類似しているものを選ぶ、一対比較法を用いた。また、実験によって得られたデータを統計的に処理し、定量評価を行った。

[†] 芝浦工業大学大学院 工学研究科
Graduate School of Engineering, Shibaura Institute of Technology

^{††} 株式会社ピコラボ
Picolab Co., LTD

[‡] 芝浦工業大学 工学部 情報工学科

Department of Information Science and Engineering, College of Engineering, Shibaura Institute of Technology

^{‡‡} 東京工芸大学
Tokyo Polytechnic University

2. 事例に基づく演奏表情生成システム

本章では, Kagurame システムの概要と Kagurame Phase-III で用いている旋律類似性評価手法について説明する.

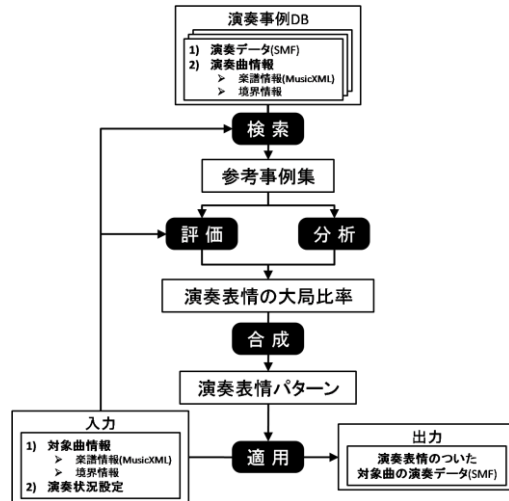


図 2.1 Kagurame システムの構成

2.1 システムの概要

Kagurame システムの構成を図 2.1 に示す. Kagurame では演奏表情生成に関する知識として演奏事例 DB を用いる. DB 中の個々の演奏事例は演奏データと演奏曲情報からなる. 演奏データには人間による演奏を標準 MIDI 形式ファイル(SMF)で与える. 演奏曲情報は演奏データに対応する楽譜情報と, 音楽的なまとまり(旋律断片)を示す境界情報からなり, 楽譜情報は MusicXML で与え, 境界情報は独自フォーマットの XML ファイルで与える. 対象曲情報は演奏曲情報と同様の形式で与える. 境界情報は, 個々の演奏事例と対象曲を旋律断片に分割する際に用いる. 図 2.2 に示すように対象曲や演奏事例を様々な長さの旋律断片に分割し, 旋律断片を対象に演奏事例を検索することで事例データのスパースネス問題に対応し, 事例を効率よく活用している.

入力が与えられると, まず, 対象曲を構成する旋律断片ごとに類似した演奏事例の旋律断片を演奏事例 DB から検索する. これによって, 対象曲の旋律断片ごとに類似した旋律断片の集合が得られる. この集合を参考事例集とする.

次に参考事例集に含まれる事例(参考事例)について重要度を評価する. 重要度は,

対象曲の演奏表情を生成する際に各事例がどの程度参考になるかを表すスコアである. 参考事例の重要度は, 対象曲と事例との旋律断片の類似性から決定する. したがって, 対象曲の断片と参考事例の断片とが類似しているほど重要度が高くなる. また, 参考事例の重要度の評価とともに参考事例の演奏表情を分析する. 演奏表情は演奏データと楽譜とのずれという形で取り出すことができる. Kagurame システムでは, 演奏表情をテンポや音の強さなどの絶対的な数値ではなく, それを変化量の比率に変換した相対的な数値で扱っている. この相対的な変化量を演奏表情の大局比率と呼んでいる.

それぞれの参考事例に対して, 重要度の評価と演奏表情の分析を行うと, 重要度でスコア付けされた演奏表情の大局比率の集合が得られる. この集合における個々の演奏表情の大局比率を, 重要度で加重平均して合成し, 対象曲の演奏表情を生成する. 最後に, 生成した演奏表情を対象曲の楽譜情報に適用して, 演奏表情のついた対象曲の演奏データの SMF ファイルを作成する.

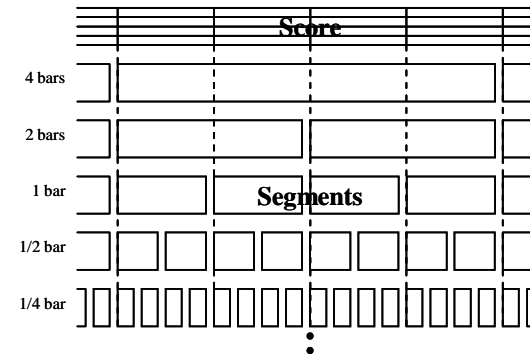


図 2.2 旋律断片への分割

3. 画像比較による旋律類似性評価

従来のシステムで用いている旋律類似性評価式において, リズムパターンの評価では旋律の長さによらず 16 分割したり, 旋律中の音の分布の評価ではオクターブの違いを無視したり, 音高の評価では旋律中の音高の平均値を用いるなど旋律の特徴量を簡略化, 抽象化してしまっていた. そのため特に長い旋律に対して類似する旋律を選出できないといった問題がある.

そこで Kagurame Phase-III では旋律の特徴量をより多く用いた類似性評価を実現するために楽譜画像の比較による旋律類似性評価を用いる. 比較する画像にはピアノロール画像を用いる. 五線譜は演奏記号が多く煩雑であり, 小節ごとの幅や段の配置が

譜面によって大きく異なることが多いなどの理由から画像比較による類似性評価は困難だと考える。一方ピアノロール画像では楽曲によって異なるのが音列の配置のみであるため画像比較に適しているといえる。

3.1 ピアノロール画像の比較手法

表情生成の対象曲の旋律断片の集合を $S=\{s_1, s_2, s_3, \dots, s_m\}$ 、演奏事例 DB の旋律断片の集合を $T=\{t_1, t_2, t_3, \dots, t_n\}$ とする。 S と T の直積 $S \times T = \{(s_1, t_1), (s_1, t_2), (s_1, t_3), \dots, (s_m, t_n)\}$ を求め、すべての旋律断片の対 (s_i, t_j) について s_i と t_j の類似度を画像比較によって評価する。比較に用いるピアノロール画像と対応する五線譜を図 3.1 に示す。ピアノロール画像はグレースケール画像であり、音符は音価に応じたグラデーションで表現する。画像の高さは 88 ピクセル、幅は楽曲の長さに応じた幅となり、音符を表す矩形の高さは 1 ピクセル、幅は 4 分音符を 32 ピクセルとしている。音符の拍位置に相当する画素を白 (RGB(255,255,255))、音符の終端を黒 (RGB(1,1,1)) としている。音符が無い部分の画素は黒 (RGB(0,0,0)) である。ピアノロール画像の比較によって旋律の類似性を求めるには対象曲画像と事例画像とで対応する画素の明度の差分を利用して求める。対象曲画像の各画素の明度を L_T 、事例画像の各画素の明度を L_S として相違性 D_{image} を式(1)によって求める。 D_{image} が 0 に近いほど旋律が類似しているため、 D_{image} の値の符号を反転して式(2)指数関数に適用して重要度 R_{image} を求める。



図 3.1 比較に用いるピアノロール画像と対応する五線譜

$$\text{相違性 } D_{image}(s_i, t_j) = \frac{\sum |L_T - L_S|}{\sum L_T + \sum L_S} \quad (1)$$

$$\text{重要度 } R_{image}(s_i, t_j) = e^{-D_{image}} \quad (2)$$

3.2 音高の違いを考慮した画像比較

3.1 で述べた手法では旋律中の拍位置と音高がともに一致しないと類似性が低いと評価してしまう。しかし、例えばある上昇系の旋律に対して音高は異なるが同じ音程で上昇している旋律は類似性が高いといえる。このように音高は異なるが概形は類似している旋律は演奏表情の生成に有効な事例として扱うことが好ましい。そこでピアノロール画像を比較する際に、音高方向つまり画像の縦方向にずらして比較を繰り返すことで音高が異なっても旋律の概形が類似している事例の類似性を適切に評価

できると考える。画像をずらして比較するため、相違性を式(3)によって算出して旋律の音高の違いを加味する。 $Height$ は画像の高さ(pixel)であり、 $Overlap$ は画像の重なっている部分の高さ(pixel)となる。 $Overlap$ の最大値は $Height$ と等しくなるが、最小値は $Height$ の半分とした。この手法では画像を縦方向に 1 ピクセルずつずらしてその都度重なっている部分の明度の差を計算し、最も明度の差が少ない位置で類似性を評価する。したがって 1 つの旋律断片に対して何回も明度の差を計算するため、対象曲が長くなるほど、また演奏事例 DB に収録される事例が増えるほど類似性評価に時間がかかってしまう。

$$\text{相違性 } D_{Shift} = D_{image} \cdot \frac{Height}{Overlap} \quad (3)$$

4. 聴取による演奏表情の類似性評価実験

Kagurame システムでは、楽譜が似ていれば演奏表情も似ているという仮定のもとで演奏表情を生成している。しかし、旋律の類似性が高い事例の実演奏データにおける演奏表情が対象曲への表情付けに適しているかは明らかでない。そこで、実演奏データが存在する楽曲を対象曲としてシステムによる旋律類似性評価を行い、類似性が高いと判断された事例の旋律断片の実演奏データと対象曲の旋律断片の実演奏データを聴き比べて演奏表情の類似性を評価する聴取実験を行った。

4.1 評価用演奏データ

演奏表情の類似性評価は、対象曲の旋律断片の長さを 2,4,8 小節とした 3 パターンについて行った。表 4.1 に示すそれぞれの曲を対象曲として Kagurame による旋律類似性評価を行い、長さごとに類似性の高い事例が選ばれている対象曲の旋律断片を演奏表情の類似性評価の判断基準(Target)とした。

表 4.1 旋律類似性評価に使用した楽曲

曲名	作曲者
Etude No.23, Op.25-11 冒頭30小節	Chopin
Prelude Op.28 No.15 冒頭27小節	Chopin
Prelude Op.28 No.20 冒頭13小節	Chopin
Nocturne No.9, Op.32-1 冒頭8小節	Chopin
Piano Sonata K.279 1st Mov. 冒頭8小節	Mozart
Piano Sonata K.310 3rd Mov. 冒頭32小節	Mozart
Piano Sonata K.331 1st Mov. 冒頭16小節	Mozart
Piano Sonata K.331 2nd Mov. 冒頭18小節	Mozart
Piano Sonata K.545 1st Mov. 冒頭12小節	Mozart
Piano Sonata K.545 2nd Mov. 冒頭16小節	Mozart

表 4.2, 表 4.3, 表 4.4 は, 長さ 2,4,8 小節それぞれの対象曲の演奏に対する演奏表情の類似性評価に使用した演奏データの楽曲名と使用した部分である。演奏データは CrestMusePEDB[8]に収録されているものと, ピアノ歴 17 年のアマチュア演奏者によるものである。それぞれのデータは SMF で与えられており, 実験時には MP3 形式に変換したデータを使用する。Target は Kagurame システムによる旋律類似性評価の対象とした旋律断片であり, a~e は Target に対して旋律類似性が高いと評価された旋律断片の演奏データである。なお, a から順に Kagurame システムの旋律類似性評価による類似性が高い旋律断片となっている。

表 4.2 長さ 8 小節の対象曲と演奏表情の類似性評価に使用した事例一覧

	曲名	使用した部分	演奏者
Target	Prelude Op.28 No.15	第20~27小節	Vladimir Ashkenazy
a	Piano Sonata K.310 3rd Mov.	第1~16小節	Maria J. Pires
b	Etude No.23, Op.25-11	第1~8小節	Vladimir Ashkenazy
c	Nocturne No.9, Op.32-1	第1~8小節	アマチュア
d	Prelude Op.28 No.20	第1~8小節	Vladimir Ashkenazy
e	Piano Sonata K.545 1st Mov.	第5~12小節	Maria J. Pires

表 4.3 長さ 4 小節の対象曲と演奏表情の類似性評価に使用した事例一覧

	曲名	使用した部分	演奏者
Target	Nocturne No.9, Op.32-1	第1~4小節	アマチュア
a	Piano Sonata K.545 1st Mov.	第1~4小節	Maria J. Pires
b	Prelude Op.28 No.20	第5~8小節	Vladimir Ashkenazy
c	Piano Sonata K.310 3rd Mov.	第25~32小節	Maria J. Pires
d	Prelude Op.28 No.15	第24~27小節	Vladimir Ashkenazy
e	Etude No.23, Op.25-11	第23~26小節	Vladimir Ashkenazy

表 4.4 長さ 2 小節の対象曲と演奏表情の類似性評価に使用した事例一覧

	曲名	使用した部分	演奏者
Target	Piano Sonata K.545 1st Mov.	第9~10小節	Maria J. Pires
a	Prelude Op.28 No.15	第26~27小節	Vladimir Ashkenazy
b	Etude No.23, Op.25-11	第23~24小節	Vladimir Ashkenazy
c	Prelude Op.28 No.20	第7~8小節	Vladimir Ashkenazy
d	Piano Sonata K.310 3rd Mov.	第21~24小節	Maria J. Pires
e	Nocturne No.9, Op.32-1	第1~2小節	アマチュア

4.2 実験手法

実験は一対比較法を用いて行う。この手法では, 被験者は例えば表 4.2 の a~e の 5 つの演奏データから 2 つを取り出して聴き比べ, どちらの演奏表情が Target の演奏表情に類似しているかを判断する。1 つの Target につき 5 つの事例からつくられる 10 組すべての組み合わせについて聴き比べを行う。さらに今回は類似していると判断した理由を「雰囲気が似ている」「リズム感が似ている」「音の強さが似ている」「その他」から 1 つ以上選択するようにした。「その他」を選んだ場合は理由をさらに記述する。

以上のような手法で表 4.2, 表 4.3, 表 4.4 のそれぞれの Target と事例の演奏データにおける演奏表情の類似性評価を行った。

なお, 被験者は本学の学生 11 名で, 特別な音楽指導の経験等はない。

5. 実験結果

表 4.2, 表 4.3, 表 4.4 の演奏データを用いた実験結果を表 5.1, 表 5.3, 表 5.5 に示す。それぞれの表中の(a,b)のセルは, a と b の演奏表情を比べた際に a の演奏表情の方が Target の演奏表情に類似していると判断した人数を表す。比率は各セルの値を被験者数で割った値の行ごとの平均値であり, 値が大きいかほど他の演奏に比べて演奏表情が Target に似ていると評価されたことを表す。「(比率)-0.5」, 「絶対値」, 「標準正規分布表より」は, 順位間の差の幅をみるために行った処理である。また, 表 5.2, 表 5.4, 表 5.6 は表 4.2, 表 4.3, 表 4.4 に示したそれぞれの演奏データの調と演奏表情値である。演奏表情値は演奏データの元となった SMF ファイルから求めた Velocity の平均値と BPM の平均値である。

5.1 長さ 8 小節の旋律に対する演奏表情の類似性評価結果

表 5.1 よりピアノロール画像の比較による旋律類似性の上位 3 つの順位が聴取評価による演奏表情の類似性の順位と異なる結果となった。特に, 旋律類似性評価において 3 番目に類似性が高いと評価された c の Nocturne No.9, Op.32-1 第 1~8 小節が聴取による演奏表情の類似性評価において最も類似性が高いと評価された。c の演奏表情のほうが似ていると判断した理由には「雰囲気が似ている」が最も多かった。表 5.2 より Target と c の調はともに長調であり, Velocity の平均値も他に比べて近いことからこのような結果になったと考えられる。一方, 同じ長調である e の旋律は聴取評価においても類似性が低い結果となった。さらに, d の演奏表情の類似性評価も旋律類似性評価と同様に低い結果となった。d,e の演奏表情値を Target と比較すると Velocity の平均値は 2 倍近く異なり, BPM の平均値では d は約 0.5 倍, e は 2 倍以上となっている。このことから d と e については旋律類似性と聴感上の演奏表情の類似性が一致したといえる。

表 5.1 長さ 8 小節の旋律に対する演奏表情の類似性評価結果

	a	b	c	d	e	合計	比率	順位	(比率)-0.5	絶対値	標準正規分布表より
a		6	0	8	8	22	0.400	2	-0.100	0.100	-0.26
b	5		1	8	8	22	0.400	2	-0.100	0.100	-0.26
c	11	10		11	11	43	0.782	1	0.282	0.282	0.78
d	3	3	0		6	12	0.218	4	-0.282	0.282	-0.78
e	3	3	0	5		11	0.200	5	-0.300	0.300	-0.85

表 5.2 表 5.1 の実験に使用した演奏データの調と演奏表情値

	調	平均Velocity	平均BPM
Target	変二長調	32.81	66.2
a	イ短調	45.34	240.0
b	イ短調	75.53	91.4
c	口長調	49.61	91.4
d	ハ短調	75.09	35.6
e	ハ長調	75.41	147.7

5.2 長さ 4 小節の旋律に対する演奏表情の類似性評価結果

表 5.3 より、長さ 4 小節の旋律に対してもピアノロール画像の比較による旋律類似性の順位と聴取評価による演奏表情の類似性の順位が異なる結果となった。特に旋律類似性評価では 4 番目に類似性が高いと評価された d の Prelude Op.28 No.15 第 24~27 小節が、聴取による演奏表情の類似性評価では最も類似性が高いと評価された。表 5.4 より聴取評価に用いた各旋律の調をみると、Target は口長調、a はハ長調、b はハ短調、c はイ短調、d は変二長調、e はイ短調となっている。一般に長調は明るい印象を与え、短調は暗い印象を与える。そのため聴取評価では Target と同じ長調である a と d が他に比べて類似していると評価されたと考えられる。また、この結果から現在の画像比較による旋律類似性評価では長調か短調かの区別がつけられていないと考えられる。

表 5.3 長さ 4 小節の旋律に対する演奏表情の類似性評価結果

	a	b	c	d	e	合計	比率	順位	(比率)-0.5	絶対値	標準正規分布表より
a		11	10	5	11	37	0.673	2	0.173	0.173	0.44
b	0		4	0	11	15	0.273	4	-0.227	0.227	-0.60
c	1	7		0	11	19	0.345	3	-0.155	0.155	-0.40
d	6	11	11		11	39	0.709	1	0.209	0.209	0.55
e	0	0	0	0		0	0.000	5	-0.500	0.500	-5.00

表 5.4 表 5.3 の実験に使用した演奏データの調と演奏表情値

	調	平均Velocity	平均BPM
Target	口長調	48.49	96.0
a	ハ長調	38.06	137.1
b	ハ短調	56.88	34.3
c	イ短調	48.29	240.0
d	変二長調	29.77	68.6
e	イ短調	73.77	120.0

5.3 長さ 2 小節の旋律に対する演奏表情の類似性評価結果

表 5.5 より、長さ 2 小節の旋律に対してもピアノロール画像の比較による旋律類似性の順位と聴取評価による演奏表情の類似性の順位が異なる結果となった。特に、旋律類似性評価では最も類似性が高いと評価された a は演奏表情の類似性評価では最も類似性が低いと評価された。また、c についても旋律類似性と演奏表情の類似性が異なる結果となった。表 5.6 より a の演奏表情値は他に比べて Target と大きく差があることから聴取による評価は妥当であるといえる。c についても BPM の平均値に大きな差があり、演奏表情の類似性は低いといえる。a や c の旋律類似性が高かった要因として、Target や c の旋律に音符が少ないため、式(1)による画像の相違性の求め方が適切でない可能性がある。

表 5.5 長さ 2 小節の旋律に対する演奏表情の類似性評価結果

	a	b	c	d	e	合計	比率	順位	(比率)-0.5	絶対値	標準正規分布表より
a		0	6	0	0	6	0.109	5	-0.391	0.391	-1.23
b	11		11	6	11	39	0.709	1	0.209	0.209	0.55
c	5	0		1	3	9	0.164	4	-0.336	0.336	-0.98
d	11	5	10		11	37	0.673	2	0.173	0.173	0.45
e	11	0	8	0		19	0.345	3	-0.155	0.155	-0.40

表 5.6 表 5.5 の実験に使用した演奏データの調と演奏表情値

	調	平均Velocity	平均BPM
Target	ハ長調	82.32	160.0
a	変二長調	30.74	68.6
b	イ短調	75.38	120.0
c	ハ短調	65.71	32.0
d	イ短調	42.33	240.0
e	口長調	48.40	96.0

6. 考察

実験結果から、ピアノロール画像の比較による旋律類似性評価と聴取による演奏表情の類似性評価では結果が必ずしも一致しないことが明らかとなった。

長さが8小節の旋律を Target とした演奏表情の類似性評価では、最も類似性が高かった旋律が Target と同じ長調であったことから、現在の旋律類似性評価手法では調の類似性評価が十分に出来ていないことが考えられる。また、旋律類似性評価において上位3つの旋律の間に大きな差が見られなかったことから長調であるか短調であるかを考慮することで、旋律類似性評価結果が演奏表情の類似性評価結果に近づく可能性がある。一方、Target と同じ長調でも演奏表情の類似性評価が低い旋律もあった。これは平均 Velocity や平均 BPM の値で Target と大きく差があることによると考えられる。また、旋律類似性評価でも他に比べて類似性が低くなっており、旋律類似性と演奏表情の類似性が概ね一致しているといえる。

長さが4小節の旋律を Target とした演奏表情の類似性評価においても、演奏表情の類似性の高い2つの旋律が Target と同じ長調であった。一方は旋律類似性評価においても他に比べて類似性が高い結果となっており、旋律類似性と演奏表情の類似性が一致した場合だといえる。しかし、もう一方は旋律類似性において評価結果が低くなっていた。これは8小節の場合と同様に長調であるか短調であるかを十分に評価出来なかったためだと考えられる。さらに、旋律類似性評価では2番目と3番目だった旋律が演奏表情では類似性が逆転していた。この2つについて Target と Velocity の平均値を比較すると、旋律類似性評価で3番目となっていた旋律の平均 Velocity が Target とほぼ同じであったため、演奏表情の類似性が逆転したと考えられる。しかし、この2つの演奏表情の類似性評価結果の差は他と比べて小さく、旋律類似性評価でもほとんど差がないことから、旋律類似性と演奏表情の類似性が一致する傾向があるといえる。

長さが2小節の旋律を Target とした演奏表情の類似性評価においては、旋律類似性評価が最も高かった旋律が演奏表情の類似性評価で最も低くなっていた。演奏表情値を比べると特に平均 Velocity において他の旋律よりも Target の演奏表情値との差が大きいがわかる。この旋律に対する旋律類似性評価が高かった原因は、旋律中に現れる音符が少なく、画像比較の際に求めている明度の差が小さくなってしまったことであると考えられる。特に、今回は Target の旋律にも音符が少なかったためその傾向が顕著に現れたと言える。

全体を通して、極端に演奏表情の異なる事例については順序が逆転しなかったことから旋律の類似性と演奏表情の類似性に関連があるといえる。ただし、ピアノロール画像の比較による旋律類似性評価では旋律の特徴を抽象化しないために調の類似性をうまく評価出来ていないことがわかった。また、音符の少ない旋律についても現在の画像比較手法が適切でないことがわかった。

7. おわりに

本稿では、旋律類似性と演奏表情の類似性の関係を明らかにするために、旋律類似性評価が高い旋律の演奏データを用いて聴取による演奏表情の類似性評価を行った。

その結果、旋律類似性と演奏表情の類似性が必ずしも一致しないこともあったが、概ね旋律の類似性と演奏表情の類似性には関連があるといえる。また、ある程度一致する傾向があることから、旋律が類似していれば演奏表情も類似しているとして演奏表情を生成することは有用であるといえる。

今後は、旋律類似性評価手法の問題点を改善し、演奏表情の類似性と旋律類似性との関連について明らかにしたい。そして画像比較による旋律類似性評価を用いて生成される表情付き演奏の質を向上させていきたい。

参考文献

- 1) 平賀瑠美: 音楽の表情付け, bit 別冊, コンピュータと音楽の世界, pp.270-282, (1998)
- 2) Widmer, G: Inductive learning of general and robust local expression principles, Proceedings of the 2001 International Computer Music Conference, International Computer Music Association, pp.322-329, (2001)
- 3) 野池賢二, 片寄晴弘: 演奏表情生成システム COPER における音楽グループ境界に基づく事例検索方式とペダル情報生成の試みについて, 情報処理学会研究報告, Vol.2009-MUS-81, No.24, 2009
- 4) 鈴木泰山, 徳永健伸, 田中穂積: 事例に基づく演奏表情の生成, 情報処理学会論文誌, Vol.41, No.4, pp.1134-1145 (2000)
- 5) 寺村佳子, 前田新一: 統計的学習によるテンポの変動を考慮したピアノ演奏模写, 情報処理学会報告, Vol.2010-MUS-84, No.12, pp.1-6, 2010
- 6) Sebastian Flossmann, Maarten Grachten, and Gerhard Widmer: Experimentally Investigating the Use of Score Features for Computational Models of Expressive Timing, Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10), pp.218-223 (2008)
- 7) 橋田光代, 片寄晴弘, 平田圭二, 鈴木健嗣, 北原鉄朗: Rencon Workshop2010: 演奏表情付けコンテスト, 情報処理学会研究報告, Vol.2010-MUS-86, No.14, pp.1-5, 2010
- 8) 橋田光代, 松井淑恵, 北原鉄朗, 片寄晴弘: ピアノ名演奏の演奏表現情報と音楽構造情報を対象とした音楽演奏表情データベース CrestMusePEDB の構築, 情報処理学会論文誌, Vol.50, No.3, pp.1090-1099, 2009
- 9) 鈴木泰山, 金子雄介, 徳永幸生: 事例に基づく演奏表情生成アルゴリズムの分析, 情報処理学会研究報告, 2005-MUS-59, Vol.2005, No.14, pp.49-54 (2005)
- 10) 金子雄介, 鈴木泰山, 徳永幸生: 事例に基づく演奏表情生成システムにおける演奏類似性と試聴評価, 情報処理学会研究報告, 2005-MUS-59, Vol.2005, No.14, pp.43-48 (2005)
- 11) 日野達也, 鈴木泰山, 野池賢二, 徳永幸生, 杉山精: 事例に基づく演奏表情生成システムにおける旋律類似性評価内容の視覚化, 情報処理学会研究報告, Vol.2009-MUS-80, No.8, pp.1-6, 2009
- 12) 日野達也, 鈴木泰山, 野池賢二, 徳永幸生, 杉山精: ピアノロール画像の比較による旋律類似性評価手法の検討-音高の違いを考慮した旋律概形の比較-, 情報処理学会研究報告, Vol.2010-MUS-86, No.17, pp.1-6, 2010