

事例に基づく演奏表情生成システムにおける 旋律と和声の構造を考慮した旋律断片自動生成

清水 厚志[†] 鈴木 泰山[‡] 野池 賢二^{††} 金子 雄介^{‡‡} 徳永 幸生[†] 杉山 精[†]
[†] 芝浦工業大学大学院 工学研究科 電気電子情報工学専攻 [‡] 株式会社ピコラボ
^{††} 株式会社トランス・ニュー・テクノロジー ^{‡‡} 株式会社日本総合研究所

1. はじめに

事例に基づく演奏表情生成システムの一つ "Kagurame Phase-II"[1][2][3]では、演奏表情生成の過程において楽譜の類似度を評価するために入力楽譜と演奏事例 DB の楽譜を旋律断片に分割する。そして、旋律断片類似性評価式によってそれぞれの旋律断片を比較し、算出した類似度に基づいて演奏事例 DB の実演奏情報を入力楽譜に適用する。

本報告では、楽譜の旋律と和声の構造を考慮した旋律断片自動生成アルゴリズムを提案する。また、本アルゴリズムと従来のアルゴリズムについて演奏表情生成の出力結果を分析し、比較検討した結果について述べる。

2. 事例に基づく演奏表情生成システム

本研究で用いた事例に基づく演奏表情生成システム "Kagurame Phase-II" の構成を図1に示す。

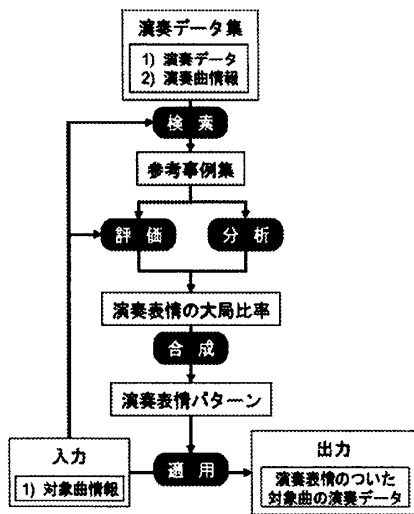


図1 “Kagurame Phase-II”の構成

Kagurame が類似事例を検索するときに必要な情報のひとつに、楽曲の境界情報がある。これは、楽曲の階層的な「まとまり」を示す情報であり、これを基に楽曲を階層的な旋律断片に区切り、それを検索単位として類似フレーズを探し出す。

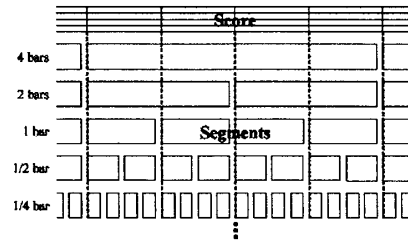


図2 旋律断片への分割

現在は、楽譜情報として MusicXML を使用している。あらかじめ与えた境界情報を用いて小節単位の旋律断片を生成し、小節より小さい旋律断片は、小節内の音符の出現位置に応じて自動生成するという手法をとっている[2]。

しかし、この手法では、旋律断片の構造が小節の区切りに強く依存するため、小節をまたぐ旋律の流れがうまく表現できない。

また、右手と左手の音符を区別せずに旋律断片を生成しているため、現在の旋律断片の構造は、楽曲の和声の構造を正確に表現しているとは言えないと考えられる。

3. 楽譜の記号を利用した旋律断片自動生成

一般的な西洋音楽の楽譜では、「音符のまとまり」を示す記号が使用されている。これらの記号は一つのフレーズを表していることが多く、旋律断片を生成する上での指標として有用であると考えられる。

音符のまとまりを示す記号には以下のようなものがある(図3)。

- (1) スラー：いくつかの音符を弧でくくり、音と音とを滑らかにつなげて演奏することを表す演奏記号
- (2) 連桁(れんこう)：8分音符より細かい音符が連なる時に符尾をつなげて書いたもの



図3 スラーと連桁

Automatic Music Piece Segmentation with Structure of Melody and Harmony in a Case-Based Performance Rendering System

[†] Atsushi SHIMIZU
[‡] Taizan SUZUKI
^{††} Kenzi NOIKE
^{‡‡} Yusuke KANEKO
[†] Yukio TOKUNAGA
[†] Kiyoshi SUGIYAMA

[†] Department of Electrical Engineering and computer Science, Graduate School of Engineering, Shibaura Institute of Technology
[‡] PicoLab Co., Ltd.
^{††} Trans New Technology, Inc.
^{‡‡} The Japan Research Institute, Limited.

これらの記号は、MusicXML では音符を表す `note` 要素の子要素として定義されている (図4)。

```

<note>
  <pitch>
    <step>D</step>
    <octave>6</octave>
  </pitch>
  <duration>1</duration>
  <voice>1</voice>
  <type>eighth</type>
  <stem>down</stem>
  <staff>1</staff>
  <beam number="1">begin</beam>
  <notations>
    <slur type="start" number="1"/>
  </notations>
</note>

```

図4 MusicXML のスラーと連桁を表すタグ

これらの情報を用いて MusicXML から旋律断片の木構造を構築するように、アルゴリズムの改良を行った。

- ① 右手部分と左手部分それぞれについて音符を走査し、各音符ごとに旋律断片を生成する。
- ② 連桁を一つのみとみなし、旋律断片の部分木を構築する。
- ③ スラーの始点から終点までに含まれる旋律断片を内包する旋律断片の部分木を生成する。
- ④ ①から③で生成した旋律断片の部分木を結合する。
- ⑤ 以降、部分木の結合を繰返し、最終的に楽譜全体を内包する木構造をボトムアップで構築する。

従来のアルゴリズム[2]と改良したアルゴリズムによって生成された旋律断片の比較を図5、図6に示す。

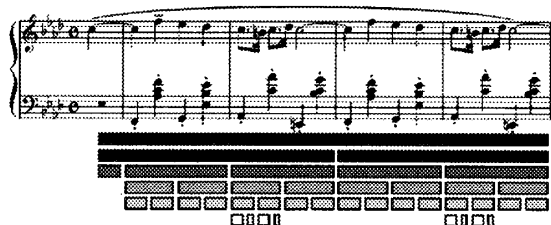


図5 従来のアルゴリズムによる旋律断片

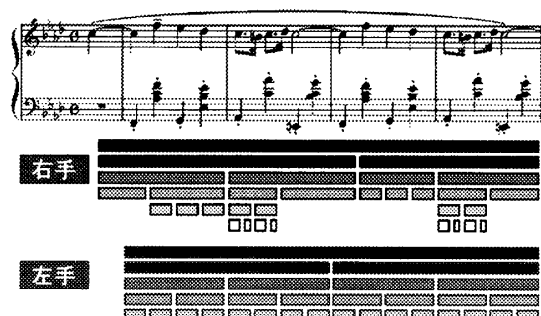


図6 改良したアルゴリズムによる旋律断片

4. 生成結果の比較

改良した旋律断片自動生成アルゴリズムの有効性を評価するために、改良前と後のそれぞれのシステ

ムが生成した演奏を用いて、Closed-Test による定量評価を行った。また、評価には以下の曲を用いた。

- ・メヌエット BWV Anh.114(J.S.Bach) 冒頭 32 小節
- ・Nocturne Op.55-1(F. Chopin) 冒頭 16 小節

システムが生成した演奏と人間の演奏との相異度 *ratio* を、以下の式(1)により算出する。

$$ratio = \frac{1}{n} \sum_{i \leq n} \left| \ln \frac{E_{kp2}(i)}{E_h(i)} \right| \quad (1)$$

$E_{kp2}(i)$ はシステムが生成した i 番目の音符の演奏表情の値であり、 $E_h(i)$ は人間の演奏の i 番目の音符の演奏表情の値である。また、 n は楽曲中に含まれる音符の総数である。相異度が小さいほど、実演奏とシステムの演奏に近いことを表す。

改良前の演奏と改良後の演奏の音の強さ (Velocity) に対する相異度 *ratio* を表1に示す。改良後の方が、*ratio* の値が低くなっていることがわかる。

表1 相異度 *ratio* の比較

	Anh.114	Op.55-1
改良前	0.089	0.083
改良後	0.060	0.076

5. まとめ

本報告では、一般的な楽譜で使用されている音符のまとまりを表す記号を利用し、楽譜の旋律と和声の構造を考慮した旋律断片自動生成アルゴリズムの構築を試みた。また、本アルゴリズムと従来のアルゴリズムについて演奏表情生成の出力結果を分析した結果について述べた。

今後は、対象曲と異なる曲を DB に用いた Open-Test 形式で生成を行い、さまざまな組み合わせで生成した演奏を定量的、定性的に評価することで、旋律断片自動生成アルゴリズムの改良を進めていきたいと考えている。

参考文献

- [1] 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, pp1134-1145, 2000.
- [2] 清水厚志, 原田杏奈, 鈴木泰山, 野池賢二, 金子雄介, 徳永幸生: 事例に基づく演奏表情生成システムにおける MusicXML からの旋律断片自動生成, 情報処理学会研究報告 音楽情報科学, 2006-MUS-68, Dec2006.
- [3] 原田杏奈, 清水厚志, 鈴木泰山, 野池賢二, 金子雄介, 徳永幸生: 事例に基づく演奏表情生成システムにおける複旋律を考慮した表情生成式, 情報処理学会第69回全国大会 6N-7, Mar2007.
- [4] 平賀瑠美, 平田圭二, 片寄晴弘: 蓮根, 目指せ世界一のピアニスト, 情報処理, Vol.43, No2, pp.136-141, 2002.