

# 事例に基づく演奏表情生成システムにおける 視覚的表現を用いた旋律類似性評価手法の検討

日野 達也<sup>†</sup> 鈴木 泰山<sup>‡</sup> 野池 賢二 徳永 幸生<sup>†</sup> 杉山 精<sup>†</sup>

<sup>†</sup>芝浦工業大学 工学部 情報工学科 <sup>‡</sup>株式会社ピコラボ

## 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"<sup>[1][2][3]</sup>では、演奏表情生成の過程において楽曲の類似度を評価するために表情生成の対象曲と演奏事例 DB の楽曲を階層的な旋律断片に分割する。そして、旋律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 の実演奏情報から抽出した演奏表情を対象曲に適用する。

現在我々は、本システムにおける演奏表情生成過程の根幹である「楽曲の類似性の評価」をより洗練するために、従来の Kagurame の数値的な比較だけではなく、ピアノロールなどの画像を用いた視覚的な比較の導入を考えている。

本報告では、その画像を用いた視覚的な比較の有用性の検証として、旋律類似性評価式によって評価している楽曲の特徴量を画像で視覚的に表現することによって、その妥当性の検討を行った。また、この検討を通して、楽曲の特徴量の視覚的な表現が有効であることを確認した。

## 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。

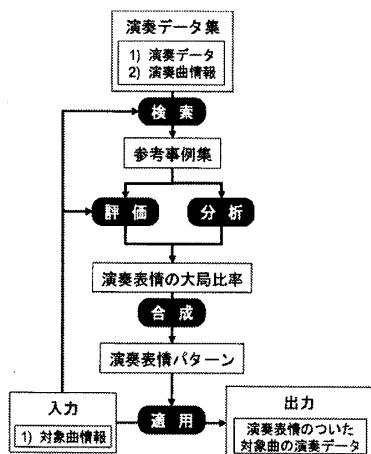


図 1 “Kagurame Phase-II” の構成

A Study of Phrase Similarity Evaluation Method using Visual Representation in A Case-Based Performance Rendering System

<sup>†</sup> Tatsuya HINO

<sup>‡</sup> Taizan SUZUKI

Kenzi NOIKE

<sup>†</sup> Yukio TOKUNAGA

<sup>†</sup> Kiyoshi SUGIYAMA

<sup>†</sup> Department of Information Science and Engineering,  
Shibaura Institute of Technology

<sup>‡</sup> PicoLab Co., Ltd.

Kagurame が類似事例を検索する際に必要とする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事例を探し出す。

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  $t_j$  の類似度を評価する。旋律の類似性を評価する際には、以下の 4 種の旋律類似性評価式を用いる。

- 1) Length Evaluator : 旋律の長さの類似性  $D_l(s_i, t_j)$
- 2) KeyRate Evaluator : 音高の類似性  $D_k(s_i, t_j)$
- 3) Rhythm Evaluator : リズムの類似性  $D_r(s_i, t_j)$
- 4) Harmony Evaluator : 音階の類似性  $D_h(s_i, t_j)$

これらの 4 種の旋律類似性の値をさらに式(1)により正規化して合計し、  $s_i$  に対する  $t_j$  の重要度  $R(s_i, t_j)$  とする。  $R$  が大きいほど、類似度が高い演奏事例の旋律断片であることを表す。

$$R(s_i, t_j) = e^{-(A_l D_l + A_k D_k + A_r (1-D_r) + A_h (1-D_h))} \quad (1)$$

ただし、  $A = (A_l, A_k, A_r, A_h)$  は重み付け変数であり、それぞれ 0 以上の値をとる。この変数を増加させると演奏表情の生成に用いられる参考事例の数が絞られ、生成される演奏表情が変化する。

## 3. 楽曲構造の類似性評価

Kagurame では類似性評価によって算出した重要度に基づいて事例を順位づけし、演奏表情の生成を行う。したがって、類似性評価の精度を高めることは人間らしい演奏を生成する上で重要である。しかし、類似性の判断を定量的に定義することは難しく、評価内容の妥当性を検証することは容易ではない。そこで、類似性評価で用いる旋律類似性評価式が評価する楽曲の特徴量をピアノロールなどの形式で画像化し、視覚的な表現の基で定性的な評価内容を検討することを試みる。特徴量の画像化により評価内容を人間が直観的に理解することができ、その妥当性を検証しやすくなることが期待できる。また、これにより画像処理手法を利用した類似性評価の洗練も、今後は視野に入れることが可能となる。ここでは、現在の Kagurame で利用している 4 種の評価式のうち、Rhythm Evaluator と Harmony Evaluator を例に、画像による視覚的な表現を検討したので、その結果を以下で述べる。

Rhythm Evaluator は、旋律断片を 16 個の枠に分割し、それぞれの枠を各枠内で発音が始まる音符の音価の和によって重み付けした要素数 16 のベクトルを用いた評価を行う。これを視覚的に表すために、枠の重みによって色付けをし、横方向に時間軸を取って 16 本の縦縞で表現した画像を図 2 に示す。これにより時間軸方向の音の分布を評価していることが視覚的にわかる。

Harmony Evaluator は、旋律断片中の音符の音高分布をオクターブの差異を考慮せずに、各音高を音価の和によって重み付けした要素数 12 のベクトルを用いた評価を行う。これを視覚的に表すために音高の重みによって色付けをし、12 本の横縞で表現した画像を図 3 に示す。これにより音高分布を評価していることが視覚的にわかる。

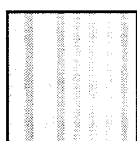


図 2 Rhythm Evaluator による旋律断片の特徴量を表した画像

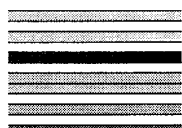


図 3 Harmony Evaluator による旋律断片の特徴量を表した画像

#### 4. 結果



図 4(a) 対象曲の旋律断片の五線譜

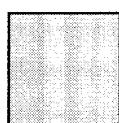


図 4(b) Rhythm Evaluator による対象曲の特徴量画像

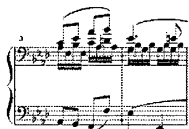


図 5(a) 重要度第 1 位の事例の五線譜

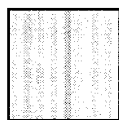


図 5(b) Rhythm Evaluator による図 5(a) の事例の特徴量画像



図 6(a) 重要度第 10 位の事例の五線譜

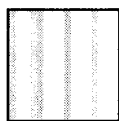


図 6(b) Rhythm Evaluator による図 6(a) の事例の特徴量画像

図 4 は対象曲の旋律断片であり、図 5,6 は図 4 の対象曲に対して検索された事例の旋律断片である。それぞれの図の(a)は五線譜による表現、(b)は Rhythm Evaluator における旋律断片の特徴量を画像で表現したものである。図の(b)に注目して比較すると、図 5 の事例のほうが図 6 の事例よりも対象曲に類似している。さらに図の(a)においても、16 分音符の分布の仕方などから図 5 の事例のほうが対象曲に類似していると判断できる。したがって、それぞれの図の(b)は Rhythm Evaluator の評価内容を表現できているといえる。



図 7(a) 対象曲の旋律断片の五線譜

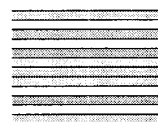


図 7(b) Harmony Evaluator による図 7(a) の旋律断片の特徴量画像



図 8(a) 重要度第 1 位の事例の五線譜

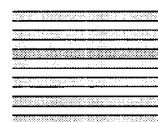


図 8(b) Harmony Evaluator による図 8(a) の旋律断片の特徴量画像



図 9(a) 重要度第 2 位の事例の五線譜

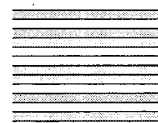


図 9(b) Harmony Evaluator による図 9(a) の旋律断片の特徴量画像

図 7 は対象曲の旋律断片であり、図 8,9 は図 7 の対象曲に対して検索された事例の旋律断片である。それぞれの図の(a)は五線譜による表現、(b)は Harmony Evaluator における旋律断片の特徴量を画像で表現したものである。それぞれの図の(b)を比較すると、横縞の現れ方から図 8 の事例のほうが対象曲に類似しているといえる。図の(a)においても、階名ソが多く使われている点や、旋律断片中の音高の分布の仕方から図 8 のほうが類似していると判断できる。したがって、それぞれの(b)は Harmony Evaluator の評価内容を表現できているといえる。

#### 5. 結論

本報告では、本システムにおける演奏表情生成過程の根幹である「楽曲の類似性の評価」をより洗練するために、4 種の旋律類似性評価式のそれぞれの評価結果を画像によって視覚的に表現し、それを用いて評価内容の妥当性の検討を行った。また、この検討を通して、視覚的な表現を用いた検討が評価を洗練する上で有効であるという見通しを得た。

今後は、旋律の類似性評価により効果的な画像化の手法や、画像同士での比較による旋律類似性評価手法の研究に取り組む。

#### 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, pp.1134-1145, 2000.
- [2] 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表情生成式の最適化, 情報処理学会研究報告 音楽情報科学, 2004-MUS-58, Dec 2004.
- [3] 清水厚志, 鈴木泰山, 野池賢二, 金子雄介, 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事例検索結果の調査とその評価, 情報処理学会第70回全国大会, 5X-5, (Mar 2008)