

事例に基づく演奏表情生成システムにおける 表情生成式の改良と評価

金子 雄介[†] 鈴木 泰山^{††} 徳永 幸生[†]

[†] 芝浦工業大学 情報工学科 ^{††} 株式会社 デュオシステムズ

1. はじめに

本稿では、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改良と評価について述べる。事例に基づく演奏表情の自動生成では、事前に用意した演奏データ集の中から対象曲に類似した事例を検索し、それらに見られる演奏表情を対象曲に適用することで表情つき演奏データを生成する。したがって、類似事例の検索や演奏表情の適用時に用いる旋律類似性評価式をいかに適切に設定するかが、システム性能の鍵となる。本研究では、旋律類似性評価式の最適化のために、評価式のパラメータと生成された演奏表情との関係を分析した。また、聴取実験を行い、演奏の類似度と人間の聴感との関係について考察した。

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”[1]の構成を、図1に示す。

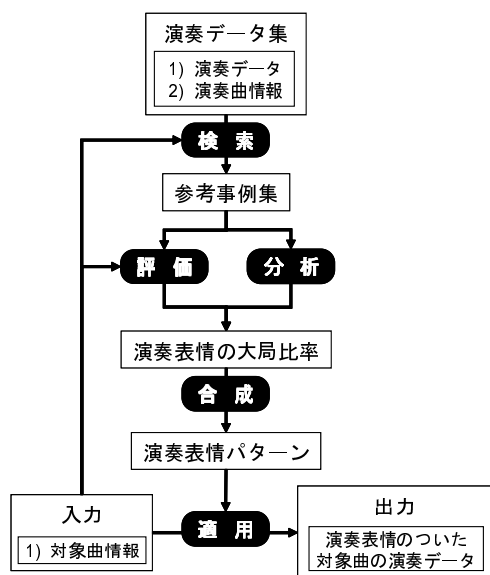


図1 “Kagurame Phase-II”の構成

本システムでは、演奏データ集から対象曲に類似する旋律を検索し、その旋律の重要度を評価する際に、

Improvement and Evaluation
of Expression Generation Algorithm
in a Case-Based Performance Rendering System

[†] Yusuke KANEKO (m103177@sic.shibaura-it.ac.jp)

^{††} Taizan SUZUKI (taizan@duo.co.jp)

[†] Yukio TOKUNAGA (tokunaga@sic.shibaura-it.ac.jp)

[†] Department of Information Science and Engineering,
Shibaura Institute of Technology

^{††} Duo Systems Co., Ltd.

以下の4つの旋律類似性評価式を用いている。

- 1) LengthEvaluator：旋律の長さの類似性 D_l .
- 2) KeyRateEvaluator：音高の類似性 D_k .
- 3) RhythmEvaluator：リズムの類似性 D_r .
- 4) HarmonyEvaluator：音階の類似性 D_h .

これらの4つの旋律類似性評価式をさらに式(1)により正規化して合計し、旋律の類似性とする。

$$D = A_l D_l + A_k D_k + A_r (1 - D_r) + A_h (1 - D_h) \quad (1)$$

D は0以上の値であり、0に近いほど類似している。

$A = (A_l, A_k, A_r, A_h)$ は重み付け変数であり、それぞれ0以上の値をとる。この変数を変化させることで、それぞれ異なる演奏表情をもつ演奏が生成される。 A の値が小さいときは、複数の参考事例をもとに演奏表情が生成される。一方、 A の値が大きくなると、最高重要度の参考事例のみをもとに演奏表情が生成される。

そこで、各評価式ごとに重み付け変数 A の値を変化させて演奏を生成し、評価式のパラメータと生成された演奏表情との関係を分析した[2]。

3. 演奏表情の比較

システムが人間の演奏にどの程度近い演奏を生成できたか評価するため、システムが生成した演奏と人間の実際の演奏との相異度 $ratio$ を、以下の式(2)より算出する。

$$ratio = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{E_{kp2}(i)}{E_h(i)} - 1 \right| \quad (2)$$

$E_{kp2}(i)$ はシステムが生成した演奏データの i 番目の音符の演奏表情の値であり、 $E_h(i)$ は人間の実演奏の 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人間の実演奏とシステムの演奏に近いことを表す。2つの演奏が全く同一であるとき、相異度は0となる。

4つの評価式をそれぞれ単独に適用して生成した演奏と人間の実演奏を比較し、演奏表情がどの程度近似しているか調べる。個々の評価式ごとに重み付け変数 A を変化させ、 A 値と相異度 $ratio$ との関係をグラフ化する。相異度は、 A が大きくなるにつれて小さくなり、ある A 以上の値で一定値となる。このため、グラフの理想曲線はL字型となる。

4. 旋律類似性評価式と相異度

本実験では、演奏の音の強さにおける表情生成に着目して分析を行った。演奏家の演奏データ集を用いてシステムが生成した演奏と、演奏家の実際の演奏とを比較し、システムがどの程度人間の演奏に似た演奏を生成できたかを評価する。そのため演奏データ集と対象曲には、同じ演奏家が作曲、演奏したものを用いた。

演奏データ集には、坂本龍一のアльбом“BTTB”[5]の制作で使用された坂本の実演奏 SMF ファイルの中から5曲選び、それぞれの楽曲の冒頭16小節を採用した。また対象曲には、同アルバム収録曲“sonatine”の冒頭16小節を採用した。なお、対象曲は演奏データ集に含まれない。

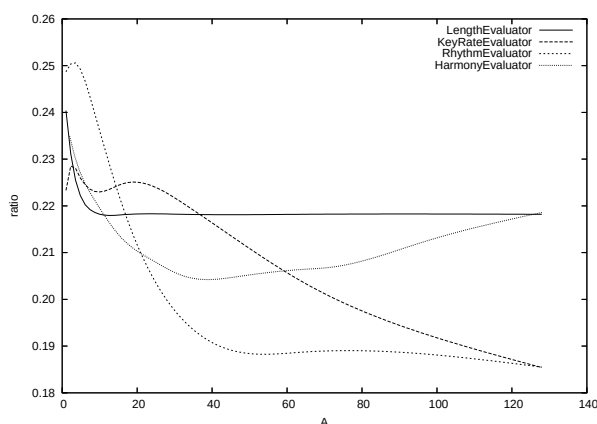


図2 評価式パラメータと相異度との関係

図2は、生成演奏の評価式パラメータと相異度との関係をグラフ化したものである。 D_l は理想曲線に高い精度で近似しているため、頑健性の高い評価式といえる。 D_r も理想曲線に近似しており、 D_l と比べて相異度が一定となるときにA値が大きい。 D_k は、Aの値が小さいときに相異度が振動しているが、おおむね理想曲線から逸脱していない。 D_h は他の曲線と異なり、Aの値が増加するにつれて相異度が増加する傾向がみられ、理想曲線から逸脱している。このことから、 D_h は算出方法を改善する必要があると考えられる。

5. 試聴実験と結果

演奏評価には、システムの性能を示す定量的評価とともに、試聴による定性的評価が不可欠である。そこで生成演奏の試聴実験を行い、相異度を用いた定量評価と人間の感性評価との関係についてオピニオン評価実験を行った[3]。

被験者は工学部の学生30名である。試聴実験の方法は、「蓮根」方式[4]を参考に設定した。被験者は演奏を聴き、どの程度見本の演奏に似ているかを示す“近似度”(5が似ている)、どの程度自然な演奏に聴こえるかを示す“自然度”(5が自然)をそれぞれ5段階

評価で回答用紙に記入する。また、それぞれの演奏について、印象や感想を自由に回答してもらった。実験は、130人収容の教室でCDプレーヤーから流した演奏を試聴する全体実験と、ヘッドフォンを用い被験者が任意に演奏を試聴するヘッドフォン実験の2種類を行った。試聴演奏の順序や回数は、全体実験では決められているのに対し、ヘッドフォン実験では被験者の自由とした。試聴実験の結果を表1に示す。

表1 試聴実験の結果

	近似度		自然度	
	全体	ヘッドフォン	全体	ヘッドフォン
相異度 0.127	3.73	3.10**	3.07	2.57**
相異度 0.235	3.07	3.13	3.37	3.53
相異度 0.359	2.80	2.97	2.20	2.63*

*は5%水準で有意、**は1%水準で有意。

全体実験では、相異度が高くなるほど、近似度、自然度ともに低くなっている。この結果から、演奏表情の定量比較のために用いた相異度と、人間の試聴による定性比較から得られた近似度は相関関係があることがわかる。一方、ヘッドフォン実験では、演奏評価に差がみられなかった。また相異度0.127の演奏の評価結果から示されるように、同じ被験者群が同じ演奏を聴いた場合でも、試聴環境によって評価結果が異なることが明らかになった。

6. おわりに

本稿では、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改良と評価について述べた。

生成された演奏の定量分析をもとに、旋律類似性評価式のパラメータと生成された演奏表情との関係を考察した。また試聴実験から、定量値である相異度と人間の試聴評価による演奏近似度との関係を考察した。

今後は、計算機による演奏の定量評価手法について引き続き検討を行う。また、旋律類似性評価式の重み付け定数の最適値を、楽曲の構成ごとに調査する。

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, 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] 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音楽情報科学, 2005-MUS-59, Feb 2005.
- [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, No.2, pp.136-141, 2002.
- [5] 坂本龍一: BTTB (限定版), WPC6-8524, ワーナーミュージック・ジャパン, 1998.