

テンポの類似性を取り入れた演奏表情生成の検討

A Study of Generating Music Expression in Tempo Similarity

柴崎 正浩†1
Masahiro Shibasaki

鈴木 泰山†2
Taizan Suzuki

徳永 幸生†1
Yukio Tokunaga

杉山 精†3
Kiyoshi Sugiyama

1. はじめに

近年、コンピュータによる情緒や抑揚のある人間らしい演奏の自動生成システムが数多く提案されている。その中で、我々は事例に基づく演奏表情生成システム“Kagurame Phase-III”^[1]の構築を進めている。本システムは、表情付けを行う対象曲と類似した楽曲の人間による演奏を対象曲に転写することで生成する。楽曲の類似性評価は楽譜情報から生成したピアノロール画像を比較することで行う。一方、我々のこれまでの実験から、テンポが異なると音が強く聴こえたり発音の長さが短く聴こえたりすることが明らかとなった。つまり、テンポは演奏の印象に大きな影響を与える因子であると考えられる^[2]。しかし、これまでのピアノロール画像による比較手法では、旋律の概形しか比較できない。従って、ピアノロール画像の類似性のみでは、適切な事例を選択できないことが示唆された。そこで、テンポの類似性を事例選択の指標に取り入れることで、より人間らしい演奏表情の生成が可能になると考えた。

本稿では、ピアノロール画像の類似性のみで生成した演奏とテンポの類似性を取り入れて生成した演奏の差分を試聴実験で明らかにし、テンポの類似性を取り入れることで、より人間らしい演奏の生成が可能になることを示す。

2. システムの概要

我々が構築を進めている“Kagurame Phase-III”の構成を図 1 に示す。本システムは、楽曲を様々な長さの階層的な旋律断片に分割し、旋律断片ごとに類似性の評価を行う。類似性の評価で用いるピアノロール画像の例を図 2 に示す。

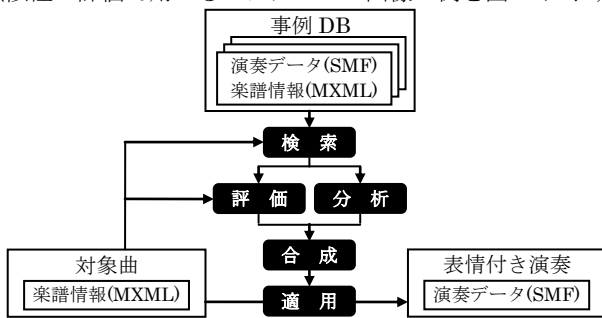


図 1 “Kagurame Phase-III” の構成



図 2 類似性の評価で用いるピアノロール画像の例

ピアノロール画像を用いることで、音列を楽譜通りに比較できる。縦方向は音高、横方向は拍位置を表す。音符は白で表記し音価に応じた矩形を取り、音符の拍位置から終端までグラデーションをかける。これは、音符の拍位置か

らの距離を判定するためである。対象曲画像と幅が異なる事例曲画像は、対象曲画像と同じ幅に伸縮させてから評価する。音高が異なっても音列の概形が類似していれば有効な事例であると考え、画像を音高方向に上下動させ、最も音符の矩形が一致する場合で評価する。

ピアノロール画像の類似性の評価は式(1)(2)を用いる。まず、画像相違度 D_{image} を式(1)より算出する。対象曲画像の各画素の明度を L_t 、事例曲画像の各画素の明度を L_s とする。 L_t と L_s の差分に、音符の拍位置の類似性を重視するために L_t と L_s の高い方の値をとる $Max(L_t, L_s)$ により重み付けをする。これを全ての画素で行い加重平均をとる。式(2)の画像類似度 R_{image} は画像相違度 D_{image} を 0 から 1 に正規化した値である。 w は画像伸縮と上下動分の重みである。ピアノロール画像の類似性のみで演奏を生成する場合は、画像類似度 R_{image} が高い事例ほど演奏表情を強く転写する。

$$D_{image} = \frac{\sum |L_t - L_s| \cdot \text{Max}(L_t, L_s)}{\sum \text{Max}(L_t, L_s)} \quad (1)$$

$$R_{image} = \left(1 - \frac{D_{image}}{255} \right)^w \quad (2)$$

3. テンポの類似性を取り入れた評価手法の提案

テンポが酷似した楽曲同士の演奏表情の差分は小さいと考え、楽譜上のテンポ指示の値の差分を正規分布式に適用させたテンポ類似度 R_{tempo} を算出することとした。ただし、楽曲によって正規分布の適切なプロファイルは未知であることから、式(3)(4)(5)の 3 種類について検討した。本稿ではそれぞれを A 案、B 案、C 案とし、テンポは 4 分音符を 1 拍とした速さとする。対象曲のテンポを T_t 、事例曲のテンポを T_s とし、 $Max(T_t, T_s)$ は T_t と T_s の高い方の値をとる。次に、算出したテンポ類似度 R_{tempo} を画像類似度 R_{image} と掛け合わせて式(6)の楽譜類似度 R_{score} を算出する。テンポの類似性を取り入れて演奏を生成する場合は、楽譜類似度 R_{score} が高い事例ほど演奏表情を強く転写する。

$$R_{tempo} = e^{-\frac{(T_t - T_s)^2}{200}} \quad (3) \text{ A 案}$$

$$R_{tempo} = e^{-\frac{(T_t - T_s)^2}{T_t + T_s}} \quad (4) \text{ B 案}$$

$$R_{tempo} = e^{-\frac{(T_t - T_s)^2}{\text{Max}(T_t, T_s)}} \quad (5) \text{ C 案}$$

$$R_{score} = R_{tempo} \cdot R_{image} \quad (6)$$

4. 試聴実験

ピアノロール画像の類似性のみで生成した演奏とテンポの類似性を取り入れて生成した演奏の差分、及び A 案、B 案、C 案でそれぞれ生成した演奏の差分を明らかにするための試聴実験を行った。

4. 1. 実験手法

人間による演奏を見本演奏、本システムで生成した演奏を評価演奏として、演奏表情の類似性をマグニチュード推定法で評価する。評価方法は演奏表情付けコンテスト

†1 芝浦工業大学, Shibaura Institute of Technology

†2(株)ピコラボ, PicoLab Co.,Ltd.

†3 東京工芸大学, Tokyo Polytechnic University

“Rencon” で用いられている、演奏の自然性と豊かさを指標として評価する。自然性は評価演奏の演奏表情が人間の演奏のように付加されているかの指標を表し、豊かさは演奏表情の付加の度合いを表す。被験者は図 3 の評価軸上その類似性を記入して回答する。見本演奏を評価軸の右端として回答で記入された位置までの距離を評価値とする。従って、評価演奏が見本演奏と同一である場合の評価値は 0 となり、見本演奏と異なるほど評価値は大きくなる。評価値の最大値は 6 とした。



図 3 回答で用いる評価軸

被験者は音楽経験者と未経験者各 6 名ずつとした。

試聴で用いるピアノ曲は Beethoven Piano Sonata Op.13 No.8-2 冒頭 8 小節(テンポ 30), Mozart Piano Sonata K331-1 冒頭 8 小節(テンポ 60), 坂本龍一 lorenz 冒頭 16 小節(テンポ 120)の 3 曲とし、3 曲ごとにピアノロール画像の類似性のみで生成した演奏と A 案, B 案, C 案によるテンポの類似性を取り入れて生成した演奏の 4 種類ずつを評価演奏とした。また、その 4 種類ずつの評価演奏のテンポを見本演奏と同一にした演奏も評価演奏とした。これは、テンポが異なると演奏の印象に大きく影響するためである^[2]。従って、評価演奏は 3 曲が 8 種類ずつで合計 24 種類となる。

事例曲を表 1 に示す。()内の数値はテンポを表す。

表 1 事例曲

Chopin Etude Op.10-4 (176)	Chopin Prelude Op.28 No.15 (72)
Mozart Piano Sonata K545-1 (160)	Mozart Piano Sonata K545-2 (60)
Chopin Nocturne Op.32-1 (120)	Chopin Trauer Marsch (60)
Chopin Nocturne Op.55-1 (120)	Chopin Prelude Op.28 No.4 (60)
Burgmuller AveMaria (100)	Chopin Prelude Op.28 No.20 (40)
Chopin Prelude Op.28 No.7 (88)	Chopin Etude Op.10-3 (30)

4. 2. 実験結果

実験結果を図 4~6 に示す。図中の「A 案」は A 案によるテンポの類似性を取り入れて生成した演奏、「画像のみ」はピアノロール画像の類似性のみで生成した演奏を表す。「A 案」は A 案によるテンポの類似性を取り入れて生成した演奏のテンポを見本演奏と同一にした演奏を表す。

4. 2. 1. Beethoven の結果

Beethoven Piano Sonata Op.13 No.8-2 における実験の結果を図 4 に示す。自然性から A 案, B 案, C 案によるテンポの類似性を取り入れて生成した演奏が最も自然であることが示された。また、画像のみでの豊かさは音楽経験者と未経験者で評価が大きく異なった。

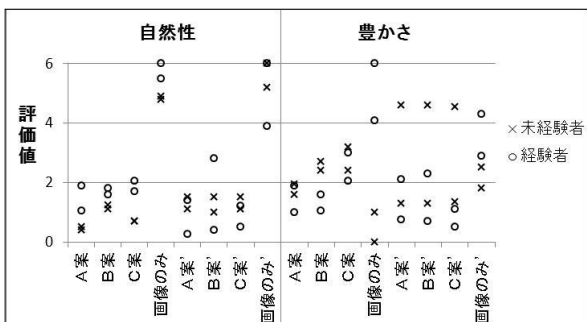


図 4 Beethoven Piano Sonata Op.13 No.8-2 の結果

4. 2. 2. Mozart の結果

Mozart Piano Sonata K331-1 における実験の結果を図 5 に

示す。自然性から C 案によるテンポの類似性を取り入れて生成した演奏が最も自然であることが示された。

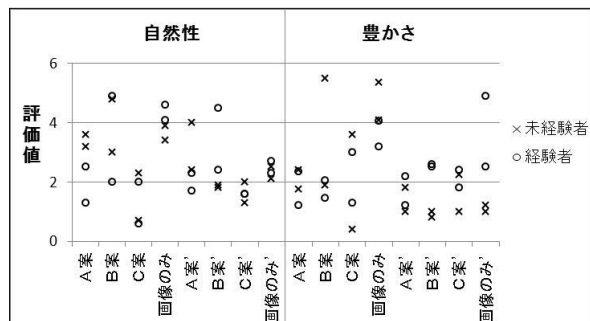


図 5 Mozart Piano Sonata K331-1 の結果

4. 2. 3. 坂本龍一の結果

坂本龍一 lorenz における実験の結果を図 6 に示す。自然性から A 案, B 案, C 案によるテンポの類似性を取り入れて生成した演奏が最も自然であることが示された。しかし、テンポを見本演奏と同一にすると不自然で単調な演奏になることが示された。また、A 案, B 案での豊かさは音楽経験者と未経験者で評価が大きく異なった。

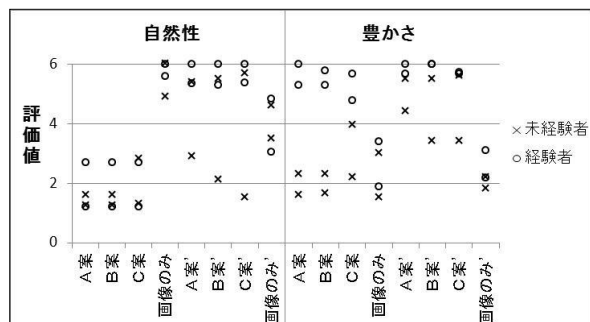


図 6 坂本龍一 lorenz の結果

4. 3. 考察

自然性の評価から、ピアノロール画像の類似性のみで生成した演奏よりテンポの類似性を取り入れて生成した演奏が、人間らしい演奏であることが明らかとなった。しかし、生成した演奏のテンポを見本演奏と同一にすると、不自然で単調な演奏になることがある。これは、人間が演奏する際には、テンポに合った演奏表情を付加するためであると考えられる。また、音楽経験者と未経験者で豊かさの解釈が異なることが明らかとなった。これは、経験者は音楽経験で得た知識も踏まえて豊かさを評価したため、未経験者と差が生じたと考えられる。

5. おわりに

本稿では、ピアノロール画像の類似性のみで生成した演奏とテンポの類似性を取り入れて生成した演奏の差を明らかにするための試聴実験を行った。その結果、本システムでは、テンポの類似性を事例選択の指標に取り入れることで、より人間らしい演奏の生成が可能になることが明らかとなった。今後は、より人間らしい演奏の生成を目指し、テンポの類似性の評価手法を洗練する。

-参考文献-

- [1]日野達也, 他: 聴感による旋律の演奏表情の類似性評価-ピアノロール画像を用いた旋律類似性評価結果との比較-, 情報処理学会研究報告, 2011-MUS-89, (2011)
- [2]柴崎正浩, 他: 聴感による演奏表情の評価特性の分析, 情報処理学会第 74 回全国大会, 3S-2, (2012)