

言葉による印象を反映した演奏表情の生成と その音楽表現学習支援への応用

鈴木 未央*

Musical Expression Generation Reflecting Impression and Application to Learning Support of Musical Representation

Mio SUZUKI

Abstract — This paper describes a learning support system of musical representation using teacher's example of musical expression that is generated based on impression expressed by an adjective with the application of our musical expression generation method. The system evaluates learner's musical performance comparing with teacher's example using a Kansei space and fuzzy rules expressing the relationship between musical expression and impression. The system presents advice by text to a learner for improvement of learner's musical representation. The experimental results show that the practice using the proposed system is effective for learning musical representation.

Key words : Musical Expression, Impression, Interaction, Learning Support System

1. はじめに

演奏表情とは、演奏によってある印象や感情を表現するために、楽譜に記述されている速度記号や強弱記号といった指示記号から逸脱した演奏のことを指す。同じ楽曲でも演奏ごと、演奏者ごとに印象が異なって聞こえるのは、演奏時の状況や演奏家ごとの表現の違いが影響し、演奏表情が変化しているためである。我々はこれまでにピアノ演奏を対象として、言葉で表された印象を反映する演奏表情の生成手法を提案している[1]。提案した手法では、演奏表情と印象の関係を感性空間とファジィルールによって表現し、印象を反映した演奏表情が生成できることを示している。さらに、生成された演奏表情に印象が十分に反映されていない場合には、ユーザとの対話により演奏表情の修正を行う。本論文では、我々が提案した演奏表情生成手法をピアノ演奏における音楽表現の学習支援に応用することを考える。

演奏生成システムを用いた音楽の学習支援に関する研究では、家庭において親子でピアノの練習をする

際、ピアノの演奏経験がない親が先生代わりとして子どもの練習を見るときにコンピュータによって演奏の音程を自動的に正しい音程に補正することで、両者の練習に対する意欲を高める効果があることがわかっている[2]。本論文では、教師が学習者に対して演奏と言葉によるアドバイスをを用いて音楽表現を教えるというレッスンの形を前提とし、演奏表情生成手法を応用した音楽表現の学習支援システムを提案する。

2. 演奏表情の要素

本論文ではピアノ演奏による演奏表情を想定する。演奏表情の要素は、テンポ、音量、音の長さの3つと定義する。要素はそれぞれ表1に示すパラメータで表現する。

3. 音楽表現学習支援システムの概要

学習支援の流れを図1に示す。支援システムは、演奏表情生成手法によって生成された演奏表情を印象を表す言葉とともに学習者へ提示する。ここで提示さ

* 釧路高専情報工学科

表1 演奏表情のパラメータ

要素	パラメータ	値	意味
テンポ	TempoBase	[40, 208]の整数	演奏全体の基準となるテンポ
	TempoRange	[0.0, 0.6]の実数	テンポの変化幅
	TempoVar	constant, decrease, increase	テンポの変化の仕方を示すフラグ
音量	VelocityBase	[16, 127]の整数	演奏全体の基準となる音量
	VelocityRange	[0.0, 0.7]の実数	音量の変化幅
	VelocityVar	constant, decrease, increase	音量の変化の仕方を示すフラグ
音の長さ	LengthSign	-1, 0, 1	音を長く演奏するか短く演奏するかを示す
	LengthBase	[1.5, 10.0]の実数	楽譜からどれだけ長く・短くなるかを示す
	LengthRange	0.0	音の長さの変化幅
	LengthVar	constant	音の長さの変化の仕方を示すフラグ

れる演奏表情を本論文では手本表情と呼ぶ。学習者は手本表情を聞いて練習を行い、演奏を記録する。システムは記録された学習者の演奏から演奏表情パラメータを求める。学習者の演奏表情パラメータと手本表情のパラメータとの比較と、感性空間上において学習者の演奏表情があると思われる範囲の推定を行い、その結果にもとづき、学習者の演奏のよかった点、改善のためのアドバイスを文章で提示する。学習者は提示された手本表情とよかった点、アドバイスを参考にして演奏表情の練習を続ける。

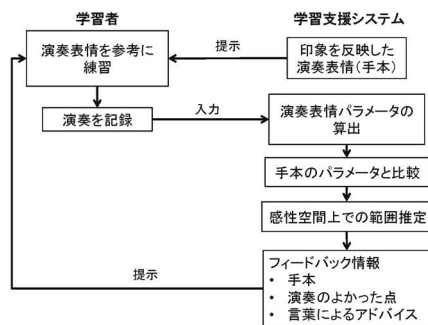


図1 音楽表現学習支援の流れ

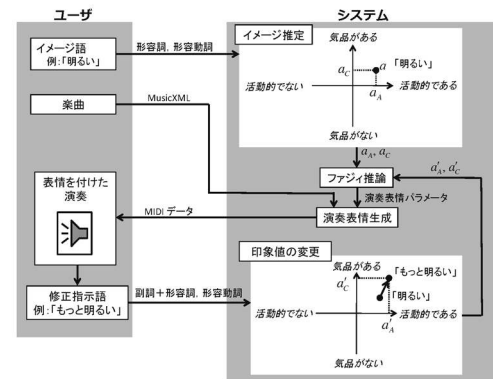


図2 演奏表情生成の流れ

を実験により求め、感性空間を構築する。実験では、実験のために生成した演奏表情を実験参加者へ提示し、5段階SD(Semantic Differential)法で印象を評価してもらう。評価結果を因子分析し、因子を抽出する。本論文では、抽出された第1因子を「活動性因子」、第2因子を「気品因子」と名付け、[-1.0, 1.0]の範囲で改めて定義し、この2因子を軸とした2次元空間を演奏表情の生成に用いる感性空間とする。

3. 1. 演奏表情の生成

図2に演奏表情生成の流れを示す。教師は演奏表情を付与したい楽曲フレーズと、演奏表情で表現したい印象を「激しい」「元気な」などの形容詞または形容動詞で入力する。楽曲フレーズは、楽譜表記フォーマットであるMusicXML形式で入力される。入力された印象の感性空間(以後、単に感性空間と呼ぶ)上での座標値をイメージ推定処理で求め、座標値から演奏表情パラメータ値をファジィ推論により求める。求められた演奏表情パラメータ値と楽曲フレーズから演奏表情のMIDIデータを作成し、教師に提示する。なお、本手法で生成される演奏表情はメロディ部分の演奏表情である。

3. 1. 1. 感性空間の構築

本論文では、演奏表情が人に与える印象の因子構造

3. 1. 2. イメージ推定処理

感性空間上での印象の座標値を求めるために、イメージ推定処理を行う。イメージ推定処理では共立共起表現と印象因子による印象推定の方法[3]を用いる。まず、入力された印象と感性空間の因子を構成する形容詞対で共立共起表現を作成する。共立共起表現とは二つの形容詞または形容動詞をつなげて用いる表現のことである。作成された共立共起表現をキーワードとしてWeb検索を行い、検索件数から印象と感性空間を構成する因子に含まれる形容詞の印象との類似度を求める。これを因子に含まれる形容詞対全てに対して行い、印象の座標値を求める。

3. 1. 3. ファジィ推論によるパラメータ決定

イメージ推定処理で得られた座標値からファジィ推論を行い、演奏表情パラメータ値を求める。演奏表

情パラメータ値の推論に用いるファジィルールは、前件部を活動性因子と気品因子に対する印象の座標値、後件部を演奏表情パラメータ値とした、以下に示すような形とする。

```
if “活動性因子軸の座標値  $a_A$  is  $LI$ ”
and “気品因子軸の座標値  $a_C$  is  $VIE$ ”
then “基準テンポ  $TempoBase$  is  $Allegro$ ”
```

テンポ、音量の変化方法を示すパラメータ、演奏時に音を長くするか短くするかを示すパラメータである $TempoVar$, $VelocityVar$, $LengthSign$ の値を推論するファジィルールは演奏表情とイメージの定性的な関係[4]から作成する。その他のパラメータ値を推論するルールは、重回帰分析の説明変数を演奏表情の因子得点、目的変数を各演奏表情パラメータとして、重回帰係数をもとに作成する。

3. 1. 4. 対話による演奏表情の修正

生成、提示された演奏表情に教師が満足できなかった場合、演奏表情の修正を行う。どのくらい演奏表情の印象を変化させたいかを「少し元気な」のように、程度を表す副詞と入力した印象の肯定形もしくは否定形を組み合わせた修正指示語として入力する。用いる副詞は「少し」「とても」「だいぶ」の3語である。教師の評価にもとづいて感性空間上で演奏表情の修正を行うために、本論文では群知能の一種で、評価関数を人間の評価に置き換えたInteractive Particle Swarm Optimization [5](以下、IPSOとする)を用いる。IPSOでは、位置と速度を持つ個体群が解空間を移動し、最適な位置を探索する。各個体が持つ位置と速度は式(1)、式(2)で計算される。

$$\mathbf{x}_j(t+1) = \mathbf{x}_j(t) + \mathbf{v}_j(t) \quad (1)$$

$$\mathbf{v}_j(t+1) = w(t)\mathbf{v}_j(t) + c_1r_1(\mathbf{p}_j(t) - \mathbf{x}_j(t)) + c_2r_2(\mathbf{g}(t) - \mathbf{x}_j(t)) \quad (2)$$

$\mathbf{x}_j$ は個体 j の位置、 $\mathbf{v}_j$ はその速度を表す。式(2)の $\mathbf{p}_j$ は個体 j が t 世代目までで最もよい評価を得た位置、 $\mathbf{g}$ は個体群の中で最もよい評価を得た位置である。 r_1 と r_2 はそれぞれ[0.0, 1.0]の乱数、 w は慣性係数である。 c_1 , c_2 は学習係数であり、 c_1 は各個体が最もよい評価を得た位置に近づく重み、 c_2 は個体群の中で最もよい評価を得た位置に近づく重みを表す。本論文では、修正指示語や、教師による演奏表情の評価を w , c_1 , c_2 と対応付ける。また、IPSOにおける個体が演奏表情に対応し、個体の位置 $\mathbf{x}_j$ が演奏表情 j の座標値を表す。

次に、教師が演奏表情を修正する手順について述べる。1回目の修正では、感性空間上で、修正指示語の副詞に対応する領域内でランダムに座標値を決定する。これを修正した演奏表情の座標値 $\mathbf{x}$ とする。修正

された演奏表情に対して、教師は印象を反映しているかを「-1:反映していない」「0:どちらでもない」「+1:やや反映している」「+2:反映している」の4段階で評価する。2回目以降の修正では、前世代の演奏表情に対する教師の評価と、入力された修正指示語から、演奏表情の位置 $\mathbf{x}_j$ と速度 $\mathbf{v}_j$ を更新し、演奏表情を修正する。以上の手順を教師が満足できる演奏表情が生成されるまで繰り返す。

3. 2. 演奏表情パラメータの比較

提案する音楽表現学習支援システムは学習者の演奏から演奏表情パラメータ値を算出し、手本表情のパラメータ値と比較することで学習者の演奏が手本に近いかどうかを判断する。算出する演奏表情パラメータ値は $LengthVar$, $LengthRange$ を除いた8つとする。算出された演奏表情パラメータ値が手本表情のパラメータ値と同じファジィ集合に属していた場合、手本表情に一致していると判断する。パラメータ値の比較後、学習者の演奏表情のよかった点と演奏表情の要素についての改善点を文章で提示する。例えば、パラメータ $VelocityBase$ が手本表情よりも小さい値で、一致していなかった場合、「現在の音量は52(p)です。音量は今より大きくしましょう」というアドバイスが生成される。

3. 3. 感性空間上での学習者の演奏表情の範囲

学習者の演奏から算出した演奏表情パラメータ値と演奏表情生成で用いられるファジィルールから、感性空間上で学習者の演奏があると思われる範囲を推定する。まず、学習者の演奏表情パラメータ値と手本表情のパラメータ値が属するファジィ集合が一致しない場合、一致していなかった学習者のパラメータ値それぞれについて、属するファジィ集合のラベルが後件部に当てはまるファジィルールから感性空間の活動性因子と気品因子の範囲を求める。次に、一致していなかったパラメータ値ごとに求められた範囲を重ねあわせ、学習者の演奏表情が存在すると思われる範囲の絞り込みを行う。そして、推定された感性空間上の範囲に、手本表情が完成するまでに生成された演奏表情の座標値が含まれているかを求める。座標値が含まれている場合、該当する演奏表情に与えられた教師の評価値、評価値ごとの演奏表情の数、入力された修正指示語を用いて、学習者の演奏表情についてアドバイスを生成する。例えば、推定された範囲内の演奏表情が明るい印象を「やや反映している」と評価されていた場合には「明るい印象をやや反映できています」というアドバイスが生成される。

4. 評価実験

提案した音楽表現学習支援システムの有効性について、実験により検証する。実験参加者は大学生、大学院生の男女7名で、いずれの参加者も過去にピアノ演奏の経験がある。実験では、実験者が生成した手本表情2曲について、参加者が電子ピアノと学習支援システムを用いて練習を行う。練習は、システムから手本表情に反映された印象が参加者の演奏にも反映されていると判断されるか、システムによる評価が10回行われるまで続ける。以下は本実験の評価項目である。

- Q1. アドバイスは練習の役に立ちましたか？
- Q2. アドバイスの内容は適切だと感じましたか？
- Q3. アドバイスはわかりやすかったですか？
- Q4. アドバイスをもらって練習することで、手本に近い演奏ができるようになったと感じましたか？
- Q5. このシステムを使うことで、音楽表現を学べそうだと感じましたか？
- Q6. このシステムを使って、音楽表現の練習をしたと感じましたか？

項目のQ1からQ4は1曲の練習が終わるごとに、Q5とQ6は実験全てが終了した後、それぞれ7段階の尺度で回答する。手本として用いる楽曲は、エリック・サティ作曲のグノシエンヌ第1番、L.v.ベートーヴェン作曲のエコセーズ ト長調 WoO.23の2曲であり、それぞれ楽曲中の1フレーズから2フレーズを練習対象とする。グノシエンヌには「かたい」、エコセーズには「軽やかな」印象を反映させる。

実験の結果を図3に示す。図3は評価項目に対する実験参加者の評価の割合である。Q1, Q2, Q4に関してはそれぞれ、85.7%, 92.9%, 91.7%の肯定的な評価を得ていることがわかる。これは、参加者がシステムから提示されたアドバイスが音楽表現の練習に有用であったことを示している。また、Q5とQ6についてもそれぞれ、71.4%, 85.7%の肯定的な評価を得ている。このことから、音楽表現を学習するのに、提案した学習支援システムが効果的であることがわかる。一方で、Q3については肯定的な評価が50.0%、否定的な評価が35.7%であった。これは実験参加者がシステムから提示されたアドバイスがわかりにくかったことが考えられる。これは、学習支援システムが手本表情との比較のみの評価を行っており、学習者自身のこれまでの演奏との比較を行っていないため、練習の際に、前回の演奏からどのように改善すべきかが想像できなかったと考えられる。そのため、より効果的な学習支援を行うためには、学習者自身のこれまでの演奏にも注目した比較が必要である。

5. おわりに

本論文では演奏表情生成手法を用いた音楽表現学

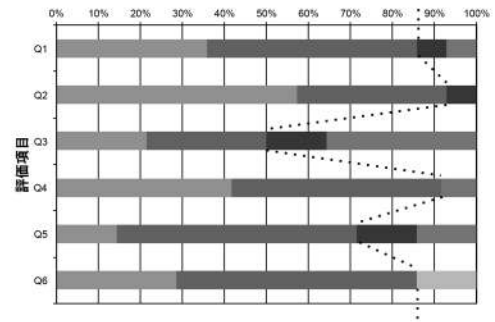


図3 アドバイスとシステムに対する評価結果

習支援システムを提案し、実験によってその有効性を検証した。学習支援システムは、学習者の演奏から演奏表情パラメータ値を算出し、定義されたファジィ集合のうちのどれに属するかを求め、手本表情と比較を行うことで、学習者の演奏のよい点とアドバイステキストを生成、提示する。また、ファジィルールと学習者のパラメータ値から、感性空間上で学習者の演奏表情があると思われる範囲を推定し、その範囲に手本表情が完成するまでに生成された演奏表情の評価値から、学習者の演奏表情に印象が反映されているかについてもアドバイステキストを生成、提示する。実験の結果から、学習支援システムによって音楽表現の学習に効果が見られること、また、実験参加者もその効果を体感していることがわかった。一方で、提示されるアドバイスがわかりにくいと評価されることもあった。現在は学習者の演奏表情と手本表情の比較のみであるため、学習者の演奏表情同士の比較を行うことで、よりわかりやすいアドバイスを提示できると考えられる。

参考文献

- [1] M.Suzuki and T.Onisawa. Interactive Generation of Musical Expression Reflecting User's Impression. In *Proc. of 1st International Symposium on Affective Engineering 2013*, pp.231-236, 2013.
- [2] C.Ohshima, K.Nishimoto and K.Suzuki. Family Ensemble: A Collaborative Musical Edutainment System for Children and Parents. In *Proc. of the 12th Annual ACM International Conference on Multimedia*, pp.556-563, 2004.
- [3] 盛多亮, 鬼沢武久. 形容詞・形容動詞の印象を用いたフレーズアニメーションの生成. *日本知能情報ファジィ学会誌*, Vol.22, No.1, pp.121-134, 2010.
- [4] Hans-Peter Schmitz, 演奏の原理. シンフォニア, 1977.
- [5] J.Madar, J.Abonyi and F.Szeifert. Interactive Particle Swarm Optimization. In *Proc. of 5th International Conference on Intelligent Systems Design and Application*, pp.314-319, 2005.