

## クラシック音楽の理解支援を目指して： 箇所表現を利用した内容記述間対応付け

栗林拓<sup>†</sup> 浅野泰仁<sup>†</sup> 吉川正俊<sup>†</sup>

京都大学大学院情報学研究科<sup>†</sup>

### 1. はじめに

クラシック音楽の楽曲は、一般には、「聴きどころ」が理解し難いものとして捉えられることが多い。音楽の専門家でない者やクラシック音楽に関する詳細な知識を持たない者にとっては尚更である。なぜなら、クラシック音楽は、他のジャンルの音楽に比べ、理論に基づいた形式が確立され、様々な予備知識を持つことで楽曲の特徴をつかむことが容易になるものであり、また、楽曲の規模も他ジャンルよりも比較的大きいものが多いため、専門的な知識を持たない者が独力で理解を得ることが困難だからである。

楽曲の内容記述とは、楽曲の内容、構造に関する客観的な記述、また、楽器名などを用いて、曲の具体的な部分を言葉で説明する記述のことである。内容記述を得ることで、例えば楽曲を鑑賞する際に、現在聴いている楽曲の部分が、楽曲の理論的な構造のどの部分に相当し、どういった意味を持つのか、などといった、予備知識なしに独力では理解することが困難なことを理解する手助けになる。

楽曲の内容記述はクラシック音楽の理解支援において非常に重要であるが、それを利用するためには、現状ではいくつか問題点がある。まず、第一の問題点として、内容記述を書籍や一般のウェブ検索から効率的に手に入れることは困難である。『名曲事典』(音楽之友社)[1]など、曲の内容記述を行っている書籍はあるが、例えば『名曲事典』は管弦楽曲に限られピアノ曲は載っていないなど、ウェブに比べ曲数に限界がある。一方で、ウェブ上の情報は玉石混淆であり、特定のウェブサイトから、あるいはウェブ検索から効率的に収集することが困難である。また、第二に、楽曲の理解支援を実現するために、単に内容記述を収集するだけでなく、その内容を更に分かりやすく提示することが重要である。具体的には、内容記述が曲のどの部分を解説し

ているか、どういった表現が曲のどこに対応付くかを提示することが役に立つ。本稿では、第二の点に対する解決手法を提案する。

第一の点に対して、我々はウェブから内容記述を効率的に収集する手法を先行研究で提案している[2][3]。クラシック音楽に関するウェブページは、楽曲に関する内容記述、楽譜、あるいはCDに関するページなど、特徴を持っている。我々は、それらのページに対する8種類の分類を提案し、それらの分類をラベルとしてL-LDA(Labeled LDA)に学習させる。そして、そのモデルを用いて、新たに得た記述が内容記述を含むかどうかのスコア付け、再ランキングを行うことで、楽曲の内容記述を段落単位で効率的に取得する。

第二の点に対して、我々は、内容記述文が、曲のどの部分を解説しているかを示す箇所表現という考えを提案する。内容記述文は、曲全体の説明をしているか、若しくは曲の部分を解説しているものである。後者の場合に、箇所との対応付けを決定する表現を箇所表現と定義する。例えば、“The first theme is played by the solo flute.”という文があるとき、“first theme”が指す箇所と“solo flute”が指す箇所が対応づくことを読み取ることができる。箇所表現を獲得することで、どういった表現が曲に対応付くかが分かり、曲と内容記述の対応付けの実現に利用することができると考えられる。本研究では、得られた内容記述から、曲中の具体的な箇所を表す表現を、ブートストラップ手法により獲得する手法を提案する。また、得られた内容記述文同士に対して、DTW(Dynamic time warping)を用いて対応付けを行う手法を検討する。

### 2. 提案手法

#### 2.1. 箇所表現の獲得

箇所間の対応関係を得るために、まずは、どういった表現(名詞)が箇所を指し示すものであるかを特定する必要がある。内容記述文の性質として、箇所間の関係を表すような動詞は、他の場所で現れても箇所間の関係を表すというこ

**Towards Understanding Support of Classical Music:  
Alignment of Content Descriptions Utilizing Passage  
Expressions**

Taku KURIBAYASHI<sup>†</sup>, Yasuhito ASANO<sup>†</sup>, Masatoshi YOSHIKAWA<sup>†</sup>

<sup>†</sup>Dept. of Informatics, Kyoto University

表 1：簡所表現獲得

	All	L-LDA	word2vec (0.128)	word2vec (0.3)	L-LDA && word2vec	L-LDA    word2vec
精度	0.232	0.323	0.2387	0.3341	0.292	0.331
再現率	1	0.808	1	0.801	0.948	0.808
F 値	0.377	0.462	0.384	0.471	0.447	0.469

とが考えられる。即ち、例えば簡所表現 2 つを主語と目的語として持つ動詞は、簡所間の関係を表すのに用いられる動詞である可能性が高いと考えられる。また、そういった動詞の主語、目的語の名詞は簡所表現である可能性が高いと考えられる。

そのために、本研究では、シード集合から対応する名詞及び動詞を繰り返し獲得するという手法を取った。ナイーブ手法として単純なブートストラップ手法、改良手法として LDA のスコア、また、word2vec[4]を用いたシード単語とのコサイン類似度を利用しノイズを除去する手法を提案する。具体的には、LDA に関しては内容記述らしさが 0 の場合を除外、word2vec に関しては閾値を 0.128 と 0.3 の 2 種類を設定し、また、LDA との組み合わせでは、双方が残す場合のみ残す (L-LDA && word2vec) と一方が残す場合には残す (L-LDA || word2vec) の 2 種類を提案する。

## 2.2. 内容記述間対応付け

DTW とは、2 つの時系列データのアラインメントを取るための手法である。本研究では、得られた内容記述の段落の各文の意味ベクトルを、各単語の word2vec ベクトルの平均により求め、各段落をその系列とすることで、段落の文同士のアラインメントを取る手法を検討している。

また、2.1 節の手法で得られた簡所表現を活用する手法として、word2vec のベクトルに加え、文中に出現する簡所表現を利用する手法を 2 種類検討している。1 つ目は、簡所表現の個数だけ次元を持ち、出現する簡所表現に対応する要素だけ 1、残りは 0 の値を持つベクトルを word2vec に付け加える。2 つ目は、出現する簡所表現それぞれの意味ベクトルを持ち、文同士の距離を計算する際に、距離を最小化するような簡所表現のベクトルを採用する。

## 3. 評価実験

簡所表現の獲得に関する評価実験を以下のように行った。

23 曲に関する内容記述文 2300 段落に対し、2.1 節で提案したブートストラップ手法を適用し、計 63 の名詞を獲得した。シード単語は、12 の名詞と 29 の動詞とした。各手法により、この 63 語を絞り、どれだけ有効な単語を残せたかを評価した。

評価は次のように行った。まず、2300 段落の中で 63 語がそれぞれ現れる文のうち、各 10 文ずつを収集した。それらの文の中で、その単語が簡所表現の役割を果たしているかどうかラベル付けを行い、その割合により、すべての文で何文簡所表現の役割を果たしているかを推定した。例えば melody という単語は 122 文に現れ、収集した 10 文のうち 8 文で簡所表現の役割を果たしていたため、全体では 122 文中 97.6 文で簡所表現の役割を果たしていると推定した。各手法が残した単語に対しこれを計算し、F 値により評価した結果が表 1 である。各手法が何も絞らない手法に比べ良い結果を残し、word2vec (閾値 0.3) が F 値で最も良い結果を残していることが分かる。

## 4. まとめ

本研究では、クラシック音楽の理解支援で重要である内容記述に着目した。そして、内容記述を効率的に収集する手法を提案した先行研究を踏まえ、内容記述を活用するために重要である簡所表現を獲得する手法を提案した。また、得られた内容記述同士のアラインメントを取る手法について検討した。

評価実験により、提案する手法により、簡所表現としてより良い単語の集合が得られたことを検証した。

今後の課題としては、ここで提案したアラインメント手法を用い、実際にユーザに分かりやすい形で複数の内容記述を提示することが挙げられる。

[1] 属啓成(1969)『名曲事典』(音楽之友社)

[2] T. Kuribayashi, Y. Asano, M. Yoshikawa. Ranking method specialized for content descriptions of classical music. Poster Proc. of WWW2013.

[3] 栗林拓, 浅野泰仁, 吉川正俊. クラシック音楽の内容記述のウェブからの収集手法. DEIM2014.

[4] T. Mikolov, K. Chen, G. Corrado, J. Dean. Efficient estimation of word representations in vector space. ICLR 2013.