

## 楽譜データに基づく楽曲の特徴量抽出の研究

米倉弥伸<sup>†1</sup> 清水宏泰<sup>†2</sup> 藤井章博<sup>†3</sup>

**概要**：楽曲の分析やそれに基づく自動作曲は多種多様な取り組みがある。その中で本研究を行う理由として次の2点を挙げる。1つは、機械学習によって類似の楽曲が生成できるとしてもそこに至るプロセスを明確に捉えておきたいと考えていることである。もう1つとして、分析の成果を「知識グラフ」という構造化したデータとして蓄積し、これを発展させることで人間の創作者による作曲活動の支援に資するシステムの実現の可能性を検討したいという点である。このような発想に基づき、機械学習に応用させる目的で、楽曲の旋律・調・音名・音価の関係を分析するために、構造化データを用いることを試みた。分析対象としてMusicXML形式で記述したバッハの無伴奏チェロ組曲第1番～第6番の全36曲を扱い、N-gram、クラスタリングなどの自然言語処理の手法を適用した。これにより1つの組曲を通じて現れる特徴や、曲種ごとの類似性等の特徴量を抽出することができた。

**キーワード**：MusicXML, 自然言語処理, RDF

### Research of extracting the characteristics of music based on score data

Yonekura Hisanobu<sup>†1</sup> Simizu Hiroyasu<sup>†2</sup> Fujii Akihiro<sup>†3</sup>

#### 1. はじめに

コンピュータによる楽曲の分析は、作曲家の創作的活動を何とかしてコンピュータ上で再現しようという試みの一部である。近年、機械学習においては大量の楽曲データをニューラルネットワークに学習させ作曲家の傾向を再現する試みが成功し、例えばバッハの楽曲を学習したネットワークからその楽曲の傾向を有するメロディーを創作できるようになっている。一方、楽曲の特徴は、例えばバロック時代の舞曲に供する楽曲であれば、メヌエットなどの舞曲の様式に即した楽曲の構成がなされており、作曲者の創意図の前提となっている。音楽という創作的活動をコンピュータで再現しようとする試みにおいては、機械学習の有効性を認識しつつ、創作の前提となっている作曲様式などの特徴をコンピュータで制御可能な形で把握することの重要性も失われていないと考える。

楽曲の分析やそれに基づく自動作曲は多種多様な取り組みがあり、[1]基本的な文脈自由文法やベイズ推定に基づく分析や生成の提案が多く存在する。総合的なシステムとしては、Orpheus[2]やGoogle Doodle[3]などが有名である。そのような状況に立ってあえてこのような研究を行う理由として次の2点を挙げる。1つは、楽曲の特徴は関心を寄せ

る視点が様々考えられるため、機械学習によって類似の楽曲が生成できるとしても、そこに至るプロセスを明確に捉えておきたいと考えていることである。具体的には、舞曲曲という制約を持つ様式の特徴量を把握することができるかどうか検証を行いたいという点である。もう1つの観点として、分析の成果を「知識グラフ」という構造化したデータとして蓄積し、これを発展させることで人間の創作者による作曲活動の支援に資するシステムの実現の可能性を検討したいという点である。すなわち、音楽学の分野で「アナリーゼ」と呼ばれる行為を支援し、演奏者や作曲者が創造性を発揮するための支援を行う方法を検討したいということである。

本研究では、このような発想に基づいて、楽譜の構造のみに着目して楽曲の特徴を抽出し、機械学習に応用させることを検討する。MusicXML形式の構造化データを利用し、複数の共通した様式から構成される組曲について分析を行った。具体的には各楽曲の旋律・調・音名・音価の関係を分析し、1つの組曲を通じて現れる特徴や、曲種ごとの類似性等の特徴量を抽出することができるかについて調査した。これにより数理的構造のみから楽曲をどの程度示すことができるかについて検討した。

#### 2. 分析対象楽曲

本研究では、分析対象楽曲としてバッハが作曲した無伴奏チェロ組曲第1番～第6番を編曲した「ギターのための無伴奏チェロ組曲全曲集」[4]に収録されている全36曲を扱った。表1に分析対象楽曲の一覧を示す。

<sup>†1</sup> 法政大学大学院理工学研究科  
Graduate School of Science and Engineering, Hosei University  
<sup>†2</sup> シミズウェブワークス  
Shimizu web Works  
<sup>†3</sup> 法政大学理工学部  
Faculty of Science and Engineering, Hosei University

表 1 分析対象楽曲一覧

第 1 番 (ト長調)	第 2 番 (ニ短調)	第 3 番 (ハ長調)
前奏曲	前奏曲	前奏曲
アルマンド	アルマンド	アルマンド
クーラント	クーラント	クーラント
サラバンド	サラバンド	サラバンド
メヌエット	メヌエット	ブレー
ジーク	ジーク	ジーク
第 4 番 (変ホ長調)	第 5 番 (ハ短調)	第 6 番 (ニ長調)
前奏曲	前奏曲	前奏曲
アルマンド	アルマンド	アルマンド
クーラント	クーラント	クーラント
サラバンド	サラバンド	サラバンド
ブレー	ガヴォット	ガヴォット
ジーク	ジーク	ジーク

分析対象とする組曲は、古典的な舞曲曲を数種集めたものに前奏曲をつけたものであり、第一曲に前奏曲、第二曲にアルマンド、第三曲にクーラント、第四曲にサラバンド、第五曲にメヌエット、あるいは他の舞曲曲、第六曲にジークとなっている。解説書による舞曲曲それぞれの特徴を表 2 に示す[5]。

表 2 楽曲の特徴

前奏曲	即興的な自由な形式なものが多い
アルマンド	4 分の 4 拍子で中庸な速度をもった舞曲曲
クーラント	4 分の 2 拍子で力強い舞曲曲
サラバンド	4 分の 3 拍子、あるいは 2 分の 3 拍子による極めて緩やかな舞曲曲
メヌエット	4 分の 3 拍子で典雅な舞曲曲
ジーク	8 分の 3 拍子、あるいは 8 分の 6 拍子、または 8 分の 12 拍子などの極めて迅速で活発な舞曲曲

### 3. 分析のためのデータ構造

#### 3.1 考え方

本研究では、大きく 2 つの観点で分析を行う。1 つは、音符の連続を時系列情報として統計的に分析する方法である。その上で、2 つ目として「主題」や「展開」を検討するために楽曲ごとに 1 つ以上の小節からなる「フレーズ」を定義し、それらの和声的關係性等を扱う。

#### 3.2 主旋律データの準備

初めに、「KAWAI スコアメーカー9pro」を用いてスキャナで読み込んだ紙媒体の楽譜を MusicXML 形式でエクスポートした。この時、読み込んだ後の楽譜に認識ミスがあれ

ばエクスポートする前に修正を行った。その後、「python」のプログラムを用いて MusicXML データから主旋律データを取り出す。主旋律データについては、後の分析で利用するために音名と音価で 1 つのタプルとし、1 小節ごとで 1 つのリストとなるツールを開発した。また、本研究では 1 拍分を音価 1 としている。

#### 3.3 MusicXML

MusicXML は XML をベースとした楽譜記述言語である。MusicXML 形式の記述例を図 1 に示す。

```
<note>
  <pitch>
    <step>A</step>
    <octave>3</octave>
  </pitch>
  <duration>4</duration>
  <voice>2</voice>
  <type>quarter</type>
  <stem>down</stem>
  <notations/>
</note>
```

図 1 MusicXML での記述例

図 1 に示したように、note タグは子として pitch タグや duration タグなどを持ち、pitch タグは子として step タグや octave タグを持つなど MusicXML は木構造を持つ要素の集合として表記される。また、step タグは音名、octave タグはオクターヴ、duration タグは音価などそれぞれのタグが楽譜の構成要素を示す。

#### 3.4 音符ごとの属性の付与

本研究では、連続する音符を区切って、メロディーの 1 つの単位を「フレーズ」と定義して分析対象とする。フレーズを厳密に定義することは一般には困難であり、これをソフトウェアで実現することは本研究の重要な課題の 1 つであると考えている。ここでは、舞曲曲であるという特徴を考えて、4 小節を 1 つのフレーズと定義している。

これらの音符とフレーズに関して、RDF (Resource Description Framework) の考え方に基づいて、属性となる情報をすべて共通の 3 項組のデータとして蓄積し、多様な問い合わせに対処できるように準備する。具体的には以下の 4 つを 3 項組のデータとして蓄積する。

- (1) 音階、音価の情報
- (2) 所属するフレーズとその中での位置
- (3) 各フレーズの和音
- (4) 音楽学での主題とその展開の別

組曲第 1 番前奏曲での上述したデータ構造を表 3 に示す。

表 3 組曲第 1 番前奏曲での 3 項組データ例

p1	chord	chA
p2	chord	chD
p3	chord	chE7
p1	is	B1
p2	is	B1'

RDF は、トリプルと呼ばれる構造によって関係を示すものであり、トリプルは事物の関係を主語、述語、目的語の関係をういた表現を行う。表 3 では各行が 1 つのトリプルとなっている。RDF トリプルによる表現の例を図 5 に示す。

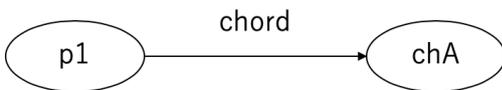


図 5 RDF トリプルによる表現の例

図 5 では、「p1」が主語、「chA」が目的語、「chord」が述語となっていて、「p1 (フレーズ 1) は A の chord である」を示すものとなる。

また、和音に関してもフレーズの分割同様重要な研究課題と認識しているが、ひとまず既存の人手による和音の情報を参照した。フレーズは「テーマ (B\*)」「テーマの展開 (B\*)」「それ以外 (E)」のどれかに属すると定め、1 楽曲中に 3~4 のテーマが現れると想定している。例として組曲第 1 番前奏曲では最初の 1 小節が最初のテーマであり、続く 3 小節はその展開と捉えている。この判別には音階と音価からなる文字列の Levenshtein 距離を計算することにより定めている。あるテーマに対して距離が 0.3 以下の場合には「新しいテーマ」であるか「それ以外」であるとし、テーマであるかどうかは後に類似の (距離 0.3 以上) フレーズが複数回するかどうかを基準に人手で判定した。

## 4. 分析

### 4.1 構成音の分析

1 つ目の分析として、各楽曲の構成音の音階、音価について、それぞれの音がどの程度の割合を占めるかの分析を行い、また、得られた分析結果が解説書による特徴とどの程度一致するかについて分析した。分析結果の例として、各組曲の前奏曲の音価についての分析結果を図 2 に、ジグの音価についての分析結果を図 3 に示す。

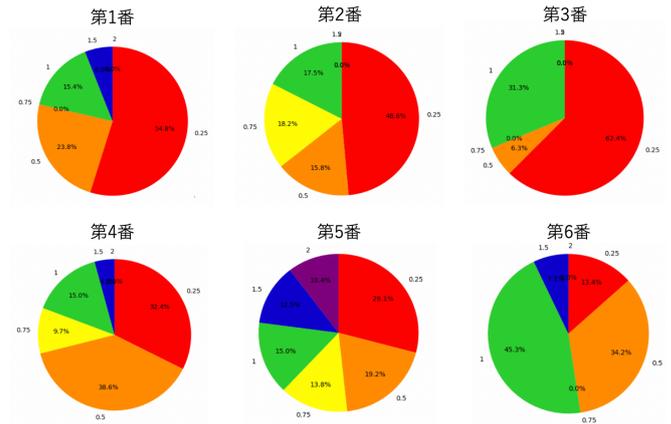


図 2 各組曲の前奏曲の音価の分析結果

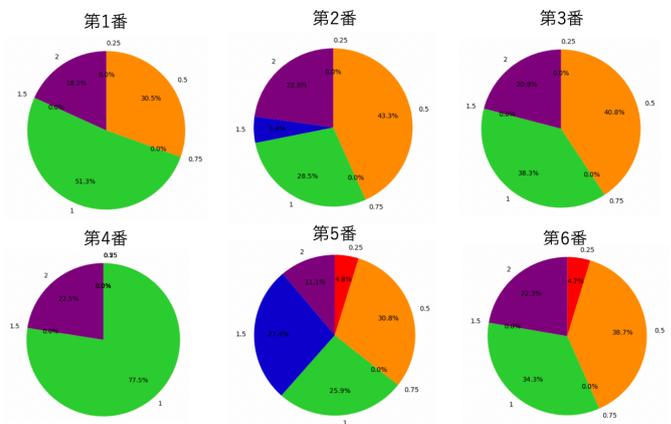


図 3 各組曲のジグの音価の分析結果

音階については、組曲第 1 番のジグや組曲第 6 番のサラバンドとガヴォットを除いて割合が最も高い音と低い音で 2 倍の差がつくような大きな偏りは見られなかった。このことから、主音や属音のように重要な音ほど割合が高いということではなく、組曲全体を通じてそれぞれの音が満遍なく使われていると考えられる。

音価については、図 2 に示したように前奏曲では音価の小さい音が多く使われているものもあれば音価が大きい音が多く使われているものもあるという結果となった。また、サラバンドでは他の楽曲と比べて音価の大きい音の割合が高い結果となった。これらのことから解説書による特徴と同じ結果を得ることができたと考えられる。一方でジグについては音価の大きい音が高い割合となる結果となり、解説書による特徴と逆の結果となった。これは、8 分音符が拍子の基準であるジグと、4 分音符が拍子の基準であるその他の楽曲共に 1 拍分を音価 1 として分析したためだと考えられる。

