

音楽知識に基づく音高・音長の組合せ特徴量を用いた MIDIデータからの作曲家判別

長谷川 隆, 西本 卓也, 小野 順貴, 嵯峨山 茂樹
東京大学大学院情報理工学系研究科

本稿では、作曲家判別に有効な音符特徴を探求することを目的として、MIDI のクラシック楽曲を対象に、音楽知識に基づいた特徴として、和音、音の広がり・厚さ、音程、リズム、及びこれらの n -gram である音程・リズム列を提案する。また、判別に用いる特徴量の選択方法として、線形判別係数を用いた寄与率を用いる方法を提案する。更に、提案する手法を用いて作曲家判別評価実験を行い、本稿で提案する特徴が作曲家の楽曲を表す指標の一部であると考えられることを示す。最後に、それらの特徴量を用いて作曲家の年代、主要活動地域の関係が示唆されることを示す。

Composer Identification from MIDI Data by Combination Features of Pitch and Duration based on Musical Knowledge

Takashi Hasegawa, Takuya Nishimoto, Nobutaka Ono and Shigeki Sagayama
Graduate School of Information Science and Technology, The University of Tokyo

This paper describes features based on musical knowledge which are effective for composer identification from classical MIDI files. The features consist of chords, expanse and thickness of notes, musical intervals, rhythms, and n -grams of intervals and rhythms. This paper also describes methods for selecting significant features using contribution ratio derived from coefficients of linear discriminants. A cross validation shows the features are part of indicators which describe composer characteristics. Finally, relations among periods and areas which involved composers are suggested using them.

1. はじめに

知識科学が対象とする情報は論理的客観的知識だが、人間の精神活動には「感じるプロセス」も含まれる。このような感性情報の抽出や表現を追求する感性情報処理の研究分野が確立されている¹⁾。また、近年の音楽配信、大容量携帯オーディオの普及によりデジタル化された音楽が大量に生成・蓄積・流通されるようになった。このような状況の下、音楽の分類・検索のための重要なメタデータの1つである感性表現に対応する特徴量を抽出する音楽解析^{2)~4)}が研究されている。これらの研究では、感性表

現として形容詞のみが考慮されている。しかし、感性情報としては、雰囲気等を表す形容詞に加えて「クラシック風の」、「ブラームス風の」などのジャンルや作曲家等を表す感性表現も含まれる⁵⁾。従って、ジャンルや作曲家に対応する音楽の特徴量を抽出する音楽解析も、感性情報処理の重要なテーマである。

また、感性表現の中で、作曲家は、形容詞やジャンルに比べ、楽曲に対応して客観的で一意なラベ

ル付が可能であるため、特徴の探究に適している。更に、作曲家の特徴(idiom)は、時代(ロマン派風等)や地域(フランス風等)の感性表現に対応する特徴(dialect)を階層的に内包すると考えられ⁶⁾、これらの探究にも役立つので、求められた特徴は、音楽の分類・検索に用いることができるだけでなく、音楽学的特徴構造の解明に役立つと考えられる。更に、特徴を用いた自動作曲への応用も考えられる。また、作曲家による音楽解析によって抽出された特徴量を用いて楽曲の作曲家を予測する作曲家判別は、特徴量の有意性を確認するための指標となる。

音楽をある基準に従って判別する様々な手法の精度を評価する MIREX (Music Information Retrieval Evaluation eXchange)⁷⁾では、評価項目に、雰囲気に加え、「ジャンル」や「作曲家」が判別基準として含まれる。この中で、作曲家判別に適用された手法は幾つかあるが、いずれの場合も他の判別基準と同じ手法が用いられていた。しかし、作曲家判別は、雰囲気やジャンルの判別と比較し、以下に挙げる特徴がある²⁾。

- ① 音響的に類似すると考えられる同編成の楽曲でも別作曲家の作品を判別しなければならない。一方、音響的に大きく異なると考えられる異編成の楽曲でも、同作曲家の作品は同じと判別しなければならない。
- ② クラシック音楽におけるモーツァルトとベートーベンの判別等、ジャンルに比べより細かく判別する必要がある。
- ③ 音響の特徴から直接捉えることが困難と考えられる作曲家の様式の特徴を把握する必要がある。

従って、作曲家判別に関しては、音響特徴に加え、音楽学の知見を用い、楽譜情報に相当する音

符情報から抽出した特徴(以下、音符特徴と記す)が有効であると考えられる。MIREX においても、音響情報から音符情報を抽出し、同情報を組み合わせた音符特徴を用いて判別を行った例があり、音響情報のみを扱った場合より判別精度が高くなったという報告がある⁸⁾。しかし、作曲家判別に対して、音符特徴としてどのような特徴が適切なのか検討した例はほとんど無い。

以上から、作曲家判別に有効な音符特徴を探求することを本研究の目的とする。本報告では、音符情報に相当する MIDI 情報を入力とし、ジャンルはクラシックとした。

2. 従来の研究

従来から音符情報を入力とした作曲家判別の研究例は幾つか存在する。Cilibrasi らは音符の相対音高、平均音量を特徴とし、特徴量ベクトルの情報量を用いた類似度で楽曲を階層木にクラスタリングした⁹⁾。また、Hardoon らは 16 分音符毎の音高集合を特徴とし、HMM と SVM を用いて 2 作曲家の楽曲の判別を行った¹⁰⁾。しかし、いずれの手法も特徴の有効性に関する検討は行っていない。Pollastri らは、音程と音長比を特徴とし、HMM を用いた判別を行った¹¹⁾。その際、特徴として採用する値の範囲や量子化の有無・粒度に関する検討を行っており、音程は量子化無しで±10半音の範囲、音長比は3値量子化で最良となり、メロディーのみを手で抽出した MIDI 情報を入力として 5 作曲家の判別で 42%の判別精度を得ている。

本報告では、第 3 節にて音楽の 5 要素に基づく特徴の検討並びに同特徴の MIDI データからの抽出方法に関して、第 4 節にて判別手法と特徴量の選択方法に関して述べる。次に第 5 節にて試行実験結果を、第 6 節にて結果に関する考察を述べ、最後に第 7 節にて結論を述べる。

3. スタイル・アナリシスに基づく特徴

作曲家判別に限らず音楽判別を行う場合、通常、楽曲データからの特徴抽出→判別分析という処理を

1 楽曲に対する形容詞付は個人差が大きく、ジャンル付もラベルの境界が曖昧な場合、カテゴリが時代毎に変化する場合等があり一意対応が困難である。

2 これらの特徴は、作曲判別において特に考慮すべきであるが、雰囲気やジャンルの判別においても有用である。

行うので、特徴が適切に選択されなければ、強力な判別分析を用いても精度は向上しない。そこで以下に、特徴の選択に関して議論する。

3.1. 特徴の選択

第 1 節にて述べたように作曲家判別には、音楽学的知見を用いた特徴が有効と考えられる。このような知見の一つであり、楽曲様式を組織的に取り扱う、ラルーらのスタイル・アナリシス¹²⁾では、以下に示す 5 つの要素を考慮する必要があると述べている。

- ① サウンド：音色
- ② ハーモニー：和声のような、音の垂直的結合の連続に関連した現象
- ③ メロディー：音高の結合によって形成された線的輪郭
- ④ リズム：音の持続時間の列
- ⑤ グロウス：上記①～④を統合する概念であり、音楽の進行と形態

この中で「サウンド」は音響情報に関連する。MIDI でも Timbre 情報を用いれば音色を扱うことができるが、本研究は最終的に音響情報を入力することを目指すため、本報告ではサウンド情報は取り扱わない。「ハーモニー」は、同時発音音高に関連するので、和音だけでなく、音の広がり・厚さも含まれる。「メロディー」は、所謂旋律のみでなく、音高の時間的結合一般が含まれる。「グロウス」は、本来音楽構造全般を包含するが、本報告では音楽の進行のみに着目し、特徴の n -gram を考慮した。

以上から、本報告では表 1 に示す特徴量を考慮する。

ここで、音高並びに音長そのものはメロディーやリズムを表す特徴ではなく、これらの値を相対的に変

化させても音楽的印象の変化はほとんど無く、感性情報として作曲家判別に寄与するとは考え難い。一方、和声に関しては、作曲家判別において重要と考えられるが、後述するノイズの問題等が存在するため、今後の課題とし、本報告では取り扱わない。

3.2. 特徴の詳細

(1) 和音

音の垂直的結合を表す特徴の 1 つとして、同時発音されている異なる音階の重なりである和音が考えられる。同特徴は音高のオクターブ間の相違や構成音の音高の上下で印象の変化が少ない。含まれる音階を表現する方法として Chroma Vector が知られているが、音高の平行移動、更に転回形を同等に扱うことを考慮し、以下のように定義した。

【特徴 1】ある時刻に同時発音されている音の階名差列の中で、辞書式順序で最初になる列

例えば、(C4, E4, G4, C5)の 5 音が同時発音されている場合、階名で(C, E, G)となり、CE, EG, GC の階名差は半音数で各々(4, 3, 5)、辞書式順序で最初であるのは(3, 5, 4)となる。従って、長 3 和音等、同種の和音は転回形を含めて 1 つの特徴に纏められる。また、要素数が 0 の場合は休符、1 の場合は単音若しくはユニゾン、2 の場合は二重音であるが、いずれも本特徴に含まれる。

特徴量は、当該和音が出現する入力楽曲全体に対する時間割合である。

なお、前後の音の重なりにより、和音として非常に多くの音高が得られる場合があり、ノイズとなるが、本報告では 4 和音以下のみを特徴とすることにより対処した。

(2) 音の広がり・厚さ

音の垂直的結合に関連する別の特徴として、音の広がり・厚さが考えられる。前者は同時発音されている音高差が大きい場合に大きくなる性質を、後者は音数が多い場合に大きくなる性質を持つ必要がある。これらの特徴を纏めて音高の標準偏差で表す方法¹⁴⁾も考えられるが、両者は独立に考慮すべきで

表 1 考慮した特徴量

	Harmony	Melody	Rhythm
uni-gram	和音	音高	音長
	音の広がり・厚さ		
bi-gram	和声	音程	リズム
n-gram		音程列	リズム列

あるため、以下の2つの特徴を定義した。

【特徴 2】音の広がり:ある時刻に同時発音されている最高音と最低音の音程

【特徴 3】音の厚さ:同時発音数

特徴量は、値が出現する入力楽曲全体に対する時間割合で重みづけした平均と標準偏差である。

(3) 音程

【特徴 4】音高と1音符前の音高の差

多重音の場合は、より高い音が知覚され易いことを考慮し、最高音高の音高差とした。また、音程が不変の場合には、(リズムを除き)音符を繋げた場合と印象に大きな違いが生じないと考えられるため、音程 0 は特徴に含めず、値の範囲は頻度が高い±2オクターブとした。

特徴量は全音程に対する割合とした。

(4) リズム

【特徴 5】量子化後の音長比

音長の僅かな相違は、感性情報としては同一と考えるべきであるため、入力された音長に Quantize (量子化)を施した。しかし、全ての音長を与えられた最小音長の整数倍に合わせる方法では連符情報が失われるため、3 連符までの情報が得られる図 1 の方法用いた。

また、3:1と1:3等、逆数の関係にある音長比は、それが繰り返される場合には同じリズムを表すため、同一特徴量とみなした³。

特徴量は全リズムに対する割合である。

(5) 音程列、リズム列

【特徴 6】音程列、リズム列共、連続する特徴の列

特徴量の数が列の要素数増加に従い指数関数的に増加しスパースになるため、要素数 2、すなわち音高・音長の trigram のみとした。

また、特徴量は、Association Rule で用いられる以下の量のいずれかとした。

$$\text{Support: } S(f_i, f_{i+1}) = P(f_i \wedge f_{i+1}) \quad (1)$$

3 予備実験で、逆数関係の特徴量が類似の値となることを確認している。

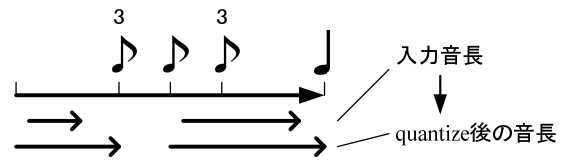


図 1 Quantize

$$\text{Confidence: } C(f_i, f_{i+1}) = \frac{P(f_i \wedge f_{i+1})}{P(f_i)} \quad (2)$$

$$\text{Lift: } L(f_i, f_{i+1}) = \frac{P(f_i \wedge f_{i+1})}{P(f_i)P(f_{i+1})} \quad (3)$$

ここで f_i は i 番目の音程若しくはリズム、 $P(f_i)$ は入力楽曲全体の音程若しくはリズムに対する f_i の割合、 $P(f_i \wedge f_{i+1})$ は入力楽曲全体の音程列若しくはリズム列に対する $\{f_i, f_{i+1}\}$ の割合である。

4. 判別分析

4.1. 判別分析手法

以下に挙げる理由により、正準判別分析が特徴量の探究に適していると考えられる。

- ① 本研究は、「楽曲は作曲家特有の音楽的特徴を有していて、特徴量空間上で作曲家の作品は多変量ガウス分布を形成する」という仮定に基づいている。
- ② 線形判別係数と特徴量の統計量によって特徴量の判別における寄与率が求められるため、特徴毎の有効性を検討できる。

4.2. 特徴量の選択手法

3.2 で述べた特徴量全てを判別器の入力とすると、次元の呪い¹³⁾に陥り、精度が低下するため、以下に示す寄与率を指標にして特徴量を選択した。

正準判別分析では、 N_{LD} 個の下記の判別関数の値によって判別を行う。

$$f_i = \sum_{j=1}^{N_f} a_{ij} x_j + c_i \quad (1 \leq i \leq N_{LD}) \quad (4)$$

ここで a_{ij} は i 番目の判別関数の j 番目の特徴量に対する係数、 x_j は判別されるデータの j 番目の特徴量、 c_i は定数である。従って、 $|a_{ij} x_j|$ が相対的に

大きな値の場合、判別に大きな影響を及ぼしていると考えられる。

以上から、以下の値を寄与率と定義し、その値が大きい特徴量を判別器への入力とした。

$$C_j = A_j \sigma_j \quad \left(A = \max_i |a_{ij}| \right) \quad (5)$$

ここで、 σ_j は j 番目の特徴量の標準偏差である。寄与率が大きい場合、データが大きく変動している上に判別関数の係数が大きいので、判別に大きく寄与すると考えられる。

寄与率は、判別関数への影響の大きさを直感的に把握し易いが、特徴量採用の客観的基準が定まらないため、試行錯誤を行う必要がある。

5. 実験評価

上述の特徴を用いて作曲家判別を行った結果を以下に述べる。

5.1. 対象データ

RWC 研究用音楽データベース(クラシック音楽、音楽ジャンル)¹⁴内のクラシック音楽で、一人の作曲家に対し4曲以上収録されている6作曲家の判別を試みた。各作曲家に対応するファイル数を100以上とするため、同データベースに収録されているMIDIデータ47曲に23曲加え、約30秒に分割したファイルを対象とした(表2)。

5.2. 実験方法

対象データから3.2に記載した特徴量を抽出、4.2記載の手法によって用いる特徴量を選択し、交差検定を行った。交差検定では、各々のファイルを、異なる楽曲由来のファイルを教師データとして判別した。指標は正解率(正しく判別されたファイル数/全ファイル数)とした。

表2 データ詳細

Composer	曲数	ファイル数
Bach	16	117
Mozart	10	157
Beethoven	10	138
Chopin	13	102
Brahms	12	229
Ravel	9	113
Total	70	856

特徴量の選択に関しては、特徴を、特徴量数が少ない順(音の広がり・厚さ、音程、リズム、和音、音程・リズム列)で加えながらその都度選択を行った。

5.3. 結果

表3に評価結果を示す。各行が実際の作曲家で、列に示した作曲家に判別されたファイル数を示す。全体で47.5%の正解率が得られた。作曲家毎の正解率ではBrahmsが69%と最も高く、BachとRavelで50%を超えた一方、Beethovenは18.8%と低かった。

6. 考察

6.1. 特徴の選択

5.3で示した結果より、本報告で示した音高・音長の組合せ特徴が、作曲家の特徴を表す指標の一部であると考えられる。実際、特徴を追加する毎に正解率は向上した。今回取り扱わなかった和声や4-gram以上の組合せを加えることにより、更に正解率が向上されることが期待できる。

6.2. 作曲家間の関係

正解率が低かったBeethovenに関して、同じドイツ系で古典派のMozartとロマン派のBrahmsへの誤判別が多い。同氏の音楽が、古典派の影響を受け、かつロマン派の新しいエポックの門戸を開いた¹⁵とされる事との対応が示唆される。

次に、各作曲家の特徴重心点を、第1判別関数と第3判別関数の2軸にマッピングした結果(図2)、及び同図の座標軸を回転した結果を示す(図3)。ドイツ系作曲家(Bach, Mozart, Beethoven, Brahms)が右側に年代順に並び、フランス系の作曲家(Chopin, Ravel)が左側に配置され、年代・地域特徴との対応が示唆される。

表3 評価結果

	Bach	Mozart	Beethoven	Chopin	Brahms	Ravel	正解率
Bach	60	21	5	5	17	9	51.3%
Mozart	26	62	39	12	17	1	39.5%
Beethoven	13	38	26	9	48	4	18.8%
Chopin	6	16	20	35	19	6	34.3%
Brahms	11	23	23	5	158	9	69.0%
Ravel	24	5	3	3	12	66	58.4%
全ファイル数	856		正解数	407		全正解率	47.5%

7. 結論

7.1. まとめ

本稿では、作曲家判別に有効な音符特徴を探索することを目的として、MIDIのクラシック楽曲を対象に、音楽知識に基づいた特徴を考案すると共に、特徴量の選択方法を検討した。また、作曲家判別評価実験を行い、結果から作曲家の年代・主要活動地域が特徴として表現されていることを示唆した。

7.2. 今後の展望

データ数を増やすと共に、今回検討対象としなかった和声特徴量の検討を行い、作曲家判別精度向上を進める。また、作曲家の年代、地域別等の特徴を更に探求する予定である。

また、音響入力の作曲家判別に向けた特徴量並びにその抽出方法に関する検討を始める。

参考文献

- 1) 井口征士, “感性情報処理が目指すもの,” 情報処理, Vol.35, No.9, pp.792-800, Sep. 1994.
- 2) 大串健吾, “音楽演奏における感性情報の抽出と分析,” 重点領域「感性情報処理」成果報告書, pp.135-140, Nov. 1995.
- 3) 武藤誠他, “音楽認知モデルによる感性情報抽出,” 情処研報, MUS-36, pp.25-30, Aug. 2000.
- 4) 平江遼他, “感性に基づくクラシック音楽の分類,” 音響学会誌, Vol.64, No.10, pp.607-615, 2008.
- 5) 長谷川隆他, “マルチメディアに対する感性合成処理のコンセプトと実験システムの評価,” 情処論文誌, Vol.38, No.8, pp.1517-1530, Aug. 1997.
- 6) Leonard B. Meyer, *Style and Music*, The University of Chicago Press, 1989.
- 7) MIREX 2008, http://www.music-ir.org/mirex/2008/index.php/Main_Page
- 8) Thomas Lindy *et al.*, “Audio Music Classification using a Combination of Spectral, Timbal, Rhythmic, Temporal and Symbolic Features,” MIREX 2008, Sep. 2008.
- 9) Rudi Cilibrasi *et al.*, “Algorithmic Clustering of Music Based on String Compression,” *Computer Music Journal*, Vol. 28, No. 4, pp.19-67, 2004.
- 10) David Hardoon *et al.*, “Using Fisher Kernels and Hidden Markov Models for the Identification of Famous Composers from their Sheet Music,” PASCAL, 2005.
- 11) Emanuele Pollastri *et al.*, “Classification of Melodies by Composer with Hidden Markov Models,” WEDELMUSIC'01, pp.88-95, 2001.
- 12) ヤン・ラルー他, *スタイル・アナリシス*, 音楽乃友社, 1988.
- 13) C.M.ビショップ, *パターン認識と機械学習*, シュプリンガー・ジャパン, 2007.
- 14) Masataka Goto, “Development of the RWC Music Database,” ICA 2004, pp.I-553-556, Apr. 2004.
- 15) 標準音楽辞典, 音楽乃友社, 2000.

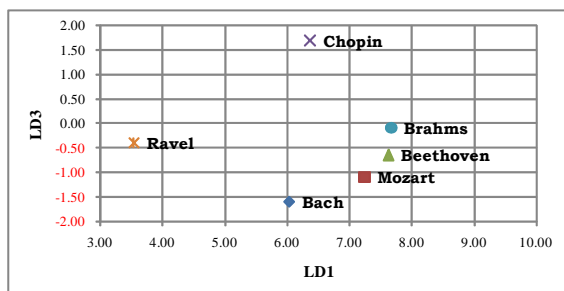


図 2 作曲家特徴量の関係

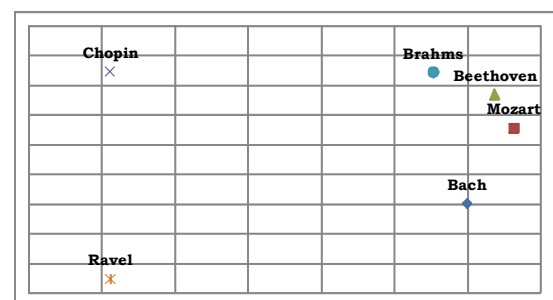


図 3 作曲家特徴量の関係(2)