

音響情報から抽出される作曲家らしさ認識のための 音楽特徴量の提案と妥当性評価

長谷川 隆^{1,a)}

受付日 2014年3月13日, 採録日 2014年9月12日

概要: 本論文では、音楽から受ける作曲家らしさの印象と作曲家の時代/地域の印象を同時に説明し定量的に測定できる工学的手法を目指して、音楽学における様式分析手法の1つである LaRue らの総合的様式分析において論じられている様々な定性的特徴の意味を解釈することにより、一般の演奏家による音楽音響から抽出可能な情報を用いた特徴量を提案する。正準判別分析の作曲家推測精度を求めることにより、提案した特徴量群による特徴空間上で同作曲家の楽曲が近接して配置されることが示された。また、同特徴空間上で、作曲家特徴重心点の1座標が作曲家の年代と相関し、他の座標が地域を表すことが示された。さらに作曲家が未知の作品特徴点、時代/地域が共通した作曲家の特徴点に近接して配置されることが示された。以上から、提案した特徴群は「作曲家らしさ」と作曲家の時代/地域の印象を同時に説明する尺度として妥当性を持つと考えられる。

キーワード: 作曲家らしさ, 音楽特徴量, 様式分析, 音楽学, 感性情報処理

Proposal of Musical Features Extracted from Audio for Composer-characteristics Recognition, and their Feasibility Evaluation

TAKASHI HASEGAWA^{1,a)}

Received: March 13, 2014, Accepted: September 12, 2014

Abstract: The purpose of this paper is to establish a technique to estimate composer-characteristic impression, and his or her chronological and geographical impression simultaneously from music audio signals. For that purpose, musical features which can be extracted from audio performed by musicians are proposed by interpreting various qualitative features stated in Style Analysis by LaRue et al., one of musicological analysis methods. Composer discrimination accuracy evaluated with canonical discriminant analysis showed points of music by the same composer in the feature space of proposed features are placed nearby. In addition, specified coordinate of composers' centroids in the space are shown to correlate to their generations, and points of composers having similar geographical backgrounds are also shown to be allocated nearby in the space with the other coordinates. Therefore, the proposed musical features are presumed to be valid for estimating composer-characteristic impression and the chronological and geographical impression.

Keywords: composer-characteristic impression, musical feature, style analysis, musicology, kansei information processing

1. はじめに

クラシック音楽愛好家は「バッハが好き」とか「ドビュッシーが好き」等、自分の音楽の好みを作曲家名で表現する

¹ 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology,
The University of Tokyo, Bunkyo, Tokyo, 113-8656, Japan

^{a)} tapi.jp@yahoo.co.jp

場合が多いことから明らかなように、「モーツァルト風」等の特定の「作曲家らしさ」、すなわち、ある作曲家が作曲した楽曲と類似した印象を与える楽曲を自動検索できれば有用であろう。また、未知の作品に対しても、類似した印象を持つ作品を作曲する作曲家を求められることが望ましい。たとえば、「ブラームス好き」の愛好家にとって、「ブラームス風」の印象を受ける未知の作曲家の楽曲が検索で

されば、新たな好みの楽曲を発見でき、有用であろう。しかし、各作曲家は多種多様な曲を作曲しているため、それらに共通する作曲家らしさはかなり主観的な感覚である。本論文では、CD や音楽配信等の音響情報から得られる客観量から、主観的な感覚である「作曲家らしさ」という印象の尺度を与えることを目的とする。

このような尺度は、下記の条件を満たす印象特徴空間を構築することにより得られる。

(1) 同作曲家作品特徴点の近接配置

(2) 近い印象を与える作曲家特徴点の近接配置

一方、Meyer は、楽曲は作品の特徴 (intraopus)、作曲家の特徴 (idiom)、時代/地域が共通の作曲家集合の特徴 (dialect) の 3 階層から構成される作曲戦略・規則に従って創作されていると述べている [1]。したがって時代や地域が共通の作曲家は同じ dialect に従って創作されており、互いに類似した印象を受けることになる。たとえば、20 世紀初頭 (時代) におけるフランス (地域) の作曲家は、「印象主義音楽」という類似した印象を受ける音楽を作曲している。したがって、同作曲家作品特徴点に加え、時代/地域が類似した作曲家特徴点が近接して配置される印象特徴空間が構築されていれば (1), (2) の条件を満たすと考えられる。

作曲家の特徴と時代の特徴を同時に反映した印象特徴空間の構築を試みた研究例として、Huovinen らは、楽譜情報から抽出した楽曲中の音符集合と音階との類似度を特徴量とし、印象特徴空間上で同一作曲家の作品特徴点が近接して配置されると同時に、作曲家特徴重心点が時代の経過に従い並んで配置されていることを示した [2]。また、長谷川ら [3], [4], [5] は、LaRue らの総合的様式分析 [6] を基にして、同手法に記載されている特徴に対応する特徴量を楽譜情報から抽出し、組み合わせて構築した印象特徴空間において、同一作曲家の作品特徴点、および時代/地域が共通している作曲家の特徴重心点が、同時に近接して配置されていることを示した。これらの研究ではいずれも、楽譜情報から特徴量を求めていた。しかし、「作曲家らしさ」の検索は一般に、CD や音楽配信等の音響情報に対して行われるので、音響情報から特徴量を抽出できることが求められる。

以上から、本論文では、一般の演奏家による音楽音響に対して、同一作曲家の作品特徴点が近接して配置され、同時に時代/地域が共通しているために近い印象を与える楽曲の特徴点が近接して配置される印象特徴空間を構築することを目的とする。

以下では、2 章で印象を定量化する基礎理論として楽譜における特徴が網羅された音楽分析手法に関して述べ、3 章で同分析手法を音響に適用するための音響情報に対する前処理に関して述べる。次に、4 章で作曲家の特徴と時代/地域の特徴を同時に抽出できる音楽特徴量を提案し、5 章

で音楽特徴量を用いた印象特徴空間構築手法に関して述べる。最後に 6 章で、同特徴量を用いて構築した印象特徴空間において、前記 (1), (2) の条件を満たすことを、実データを用いた評価により示す。

2. 基礎となる音楽分析手法

1 章で述べた印象特徴空間上では、近い印象を与える楽曲特徴点は互いに近接して配置されるので距離が小さく、各座標値の差も小さい。したがって、同座標値は楽曲の印象に関連する特徴量から求められる必要がある。一方、類似の印象を受ける同作曲家の作品でも楽曲の種類 (管弦楽/室内楽/声楽等) により音響のスペクトル形状が大きく異なるので、同形状の特徴のみではこのような特徴量として不十分である。そこで、視聴者が印象としてとらえる音楽的な特徴 (音楽特徴量) を音楽音響情報から抽出する必要がある。

一方、音楽学の分野では、楽譜情報を用いて、個々の作曲家の音楽的特質を見極めたり、「印象主義」「ワーグナー主義」のような、複数の作曲家に共通する特徴を客観的に理解したりすること等を目的とした様式分析と呼ばれる分析手法があり、「作曲家らしさ」や時代/地域特有の印象という複数楽曲共通の特徴把握もその目的に含まれる。様式分析手法の 1 つである総合的様式分析 [6] には、Sound, Harmony, Melody, Rhythm の 4 つの音楽的要素の観察視点と、これらのすべてを総合する過程である Growth の視点で、複数楽曲にまたがる共通特徴の要素となる楽譜に基づく音楽的特徴が網羅的にあげられている。しかし、同手法で述べられている特徴群は音楽専門家による楽譜の分析を想定しているため、定性的であり、そのままコンピュータによる分析に使えるものではない。

そこで、総合的様式分析で楽譜情報に基づき定性的に述べられている特徴に関する記載内容を解釈し、「作曲家らしさ」を定量化するための音楽特徴量を求める。ただし、「作曲家らしさ」は楽曲の断片を視聴するだけで感じることができるので、短い分析範囲から求められる特徴を対象とし、楽曲全体の特徴はその分布形状を表す平均、分散等の統計量を用いる。また、総合的様式分析における Growth は、音楽の動きとその記憶に基づく形態から構成されるため、比較的長い時間範囲が必要であるので、他の 4 つの音楽的要素の特徴のみを対象とする。以上を考慮した総合的様式分析における観察視点ごとの定量化可能な特徴 (以下、様式特徴と記す) を表 1 に示す。

なお、一般に音楽音響には作曲家の特徴に加えて演奏者の特徴も含まれているが、楽曲は作曲家が創作した楽譜に基づき演奏されているので、音響の中に含まれる楽譜に関連する情報に基づく特徴量を用いれば楽曲そのものが与える印象を抽出できると考えられる。

表 1 様式特徴

Table 1 Features in Style Analysis.

観察視点	様式特徴	
Sound	音域	
	Sound Texture	音の厚さ
		パートの交代 / 重複
音色		
Harmony	Harmonic Color	和音頻度
		不協和音 乖離 / 密集
	調性	音の垂直的結合の連続
Melody	音程	
	音階	
Rhythm	Surface Rhythm	
	Level	

3. 音楽特徴抽出のための前処理

音楽音響に対して楽譜情報に基づく様式分析を適用するためには、音響情報から楽譜に相当する情報を抽出する必要がある。音響情報から楽譜情報を抽出する技術は採譜処理として研究されているが、一般的な音楽音響から精度良く楽譜情報を抽出するのは困難である。しかし、一般に音楽を鑑賞する場合、必ずしも音符を個々に認識しているのではなく、楽譜に関連する音高・音名や音価の順序や組合せ等の楽譜関連情報を認識していると考えられ、総合的様式分析でもそのような情報を基にした特徴が論じられている。そこで以下に、特徴量を求める際に基礎となる楽譜関連情報の抽出に用いることができる、すでに知られている前処理に関して述べる。

3.1 Pitch Spectrum

音楽特徴量を求める際には重要な情報の1つである、時間 t における音高 p の音の大きさを表す Pitch Spectrum は以下のように定義できる。

$$\text{Pitch Spectrum: } P'(t, p) = \log P_p(t, p)$$

$$P_p(t, p) = \sum_{f_{p-\frac{1}{2}} \leq f < f_{p+\frac{1}{2}}} w_p(f) P(t, f)$$

ここで p は MIDI における Note Number に等しいピッチ番号、 $P(t, f)$ は周波数 f におけるパワー、 f_p は p に対応する周波数で以下の式で求められる。

$$f_p = 440 \cdot 2^{\frac{p-69}{12}}$$

なお、上記定義では $A4 = 440 \text{ Hz}$ としているが、パワーを求める際にハミング窓関数を使ったフーリエ変換を用いるので、調律の違いは無視できる。さらに、 $w_p(f)$ は音高ごとのパワーを求めるための窓関数で、本論文では三角窓を用いる。本前処理により、音高に関する情報を得ることができる。

3.2 Specmurt 分解

Specmurt 法 [7] とは、入力音響スペクトル $v(x)$ を近似

的に以下に示す条件に従う 2 成分に分解する手法であり、各成分の条件と分解後の成分との差異を最小にする繰返し最適化演算により求める。

$$v(x) \approx h(x) * u(x)$$

ここで x は対数周波数 (底は任意)、 $*$ は畳み込み演算で、 $u(x)$ の条件は各ピッチに相当する対数周波数のみにパワーを持つ基本周波数分布であること、 $h(x)$ の条件は $x = \log n$ ($n = 1, 2, 3, \dots$) に正の値を持つ分布で $h(0) = 1$ であることである。一般に音楽音響情報には複数のピッチで演奏された倍音構造を持つ楽音の組合せが含まれる。したがって、 $h(x)$ は周波数の整数倍におけるパワー比率なので倍音構造を表し、 $u(x)$ は倍音抑制された楽音のピッチ情報を表す。そこで $u_t(x)$ 、 $h_t(x)$ をそれぞれ時間 t における $u(x)$ 、 $h(x)$ とすると、Specmurt U と Specmurt H を以下のように定義できる。

$$\text{Specmurt U: } U(t, p) = \log(u_t(\log f_p))$$

$$\text{Specmurt H: } H_s(t, m) = h_t(\log m)$$

Specmurt U は音色成分を除去したうでの時間 t における音高 p の音の大きさを表すので、演奏されている各音符の音高組合せに近い分布が得られ、一般に音響からの採譜処理に用いられている。一方、同時に求めることができる Specmurt H は時間 t における第 m 倍音の基音に対する音の大きさの比を表すので、時間 t に演奏されている楽器群の平均的音色を表すと考えられる。そこで本論文では、音色を表す特徴量を抽出する前処理として Specmurt H を用いる。したがって、本前処理により、音高や、音価を求める基になる音高の変化、さらに音色に関する情報を得ることができる。

3.3 Chroma

和音・音階・調性等、複数の音名間の差異が意味を持つ特徴の場合、時間 t における音名 n の音の大きさを表す Chroma が重要な情報の1つと考えられる。

$$\text{Chroma: } C_H(t, n, Y) = \log \left(\sum_{p \equiv n \pmod{12}} (Y(t, p)) \right)$$

$$\text{Chroma Vector: } \vec{c}(t, Y) = (C_H(t, n, Y))$$

ここで音名 n は、C (ハ) から B (ロ) までの音名を 0~11 の数値で表した値である。また Y は、Chroma の基になる分布であり、 P_p または u_t が考えられる。さらに、Chroma Vector は音名ごとの Chroma を成分とする 12 次元ベクトルである。本前処理により、同時発音されている音名組合せに関する情報を得ることができる。

3.4 HPSS 分解

HPSS (Harmonic/Percussive Sound Separation) [8] は

入力音響スペクトルを、時間-周波数平面上でのパワーが時間方向に滑らかな（一定の音高が持続している）H成分と、周波数方向に滑らかな（特定の音高にパワーが集中せず打楽器のような音である）P成分に分解する手法で、条件に近づくように繰返し最適化により各成分を求める。

$$P(t, f) = H_h(t, f) + P_h(t, f)$$

ここで $H_h(t, f)$ は楽器持続音のような調波的成分、 $P_h(t, f)$ は打楽器音や楽器発音部分のような成分である。

本前処理の $H_h(t, f)$ における持続音高の変化や $P_h(t, f)$ のパワーピークを用いることにより、音価を求める基となる楽音の区切りを求めることができる。

3.5 Tail Segment

Tail Segment [2] は、与えられた N_c 種類の音名が演奏音中に現れる時間範囲の集合であり、前述の Chroma においてパワーが大きい音名を計数することにより求める。

$$T_S(k, X) = [t_{k,\min}, t_{k,\max}]$$

$t_{k,\min}, t_{k,\max}$ は、以下の条件を満たして、かつ他の範囲に含まれない時間範囲のそれぞれ下限ならびに上限である。

$$\text{count} \left(\bigcup_{t \in T_S(k, X)} P_N(t, Y) \right) \geq N_c$$

$$P_N(t, Y) = \{n | C_H(t, n, Y) > C_M(t, Y)\}$$

ここで $\text{count}(A)$ は集合の要素数を求める関数を表す。また、 $C_M(t, Y)$ は音名が演奏されているか否かを判断する閾値である。

一般に分散和音や五音音階等、用いられている音名の種類が少ない場合、与えられた数の音名が現れるまでに多くの時間を要するので、Tail Segment の範囲は長くなる。したがって本前処理により、使用音名数に関する情報を得ることができる。

4. 音楽特徴量

以上で述べた前処理により抽出された楽譜関連情報を用いて、表 1 に示した様式特徴に対応して求めることができる音楽特徴量を用いる記号とともに表 2 に示す。ここで、 p は音高、 t は時間、 i, j は音域番号、 C は和音、 d は半音数で表した音高差、 k は Tail Segment 番号、 K は音階、 d_w は全音階差 [5]、 R は音価比である。また、 X, Y, Z は音楽特徴量を求める際に用いる、3章で示した前処理で求められた分布を表し、 X は P' または U を、 Y は P_p または u_t を、 Z は H_h または P_h を表す。

(1) Pitch Distribution

楽曲中で選択されている音域、好んで使われる音域、極端な高音域や低音域が使われているかを表す特徴であり、楽曲中の音高の分布形状の特徴と考えられるので、分布形

表 2 音楽特徴量

Table 2 Musical features.

音楽特徴量		様式特徴
(1) Pitch Distribution	$P_p(p, X)$	音域
(2) Sound Expanse	$S_E(t, X)$	音の厚さ、乖離/密集
(3) Sound Accumulation	$A_E(t, X), A_G(t, X)$	パートの交代 / 重複
(4) Sound Texture Correlation	$S_T(i, j, X)$	和音頻度
(5) Chord Type	$C_c(t, C, Y)$	不協和音
(6) Dissonance	$D_d(t, Y)$	調性
(7) Partial Intervals	$\Delta_{pr}(d, X), \sigma_{pr}(d, X)$	調性、音階
(8) Tail Segment Length	$L_m(k, Y)$	音の垂直的結合の連続
(9) Interval Class Vector	$V_v(k, d, Y), V_r(k, d, Y)$	音程
(10) Key Profile Correlation	$K_c(k, Y)$	Surface Rhythm
(11) Parallel Intervals	$\Delta_{pr}(d, d, Y)$	Level
(12) Pitch Difference	$D_d(d, X)$	音色
(13) Rhythm	$R_r(R, Z), N_r(Z)$	
(14) Rhythm Level	$R_r(Z)$	
(15) Timbre	$H_r(t, m), H_o(t), H_e(t)$ MFCC	

状を代表する特徴量として一般に広く使われている、下式から求められる $P_D(p, P')$, $P_D(p, U)$ の分布形状を代表する値である平均、分散、歪度、尖度として定義できる。

$$P_D(p, X) = \sum_t X(t, p)$$

次に、音符の配置の特徴である Sound Texture に関連する特徴量を述べる。まず、「音の厚さ」は、同時に演奏されている音の音高に関する特徴で、これらの音高が広い音域にわたっている場合、音が厚く感じられ、音域が同じでも、音の重なりが多い方が、音が厚く感じられる。したがって、「音の厚さ」は、同時発音域 (Sound Expanse) と音の重なり (Sound Accumulation) に着目することで定量化が可能である。どちらの特徴量も、平均値だけでなく楽曲中のバラツキや偏りが楽曲の特徴と考えられるので、各特徴量の分布形状を代表する値によって定量化できる。

(2) Sound Expanse

同時に演奏されている音の広がりを表す特徴なので、一般に値のバラツキ度合いを表す量として知られている分散を用いて、下式に示す同時発音高の分散である $S_E(t, P')$, $S_E(t, U)$ の平均、分散、歪度、尖度として定量化できる。

$$S_E(t, X) = \frac{\sum_{p=p_{\min}}^{p_{\max}} (p \cdot X(t, p) - \bar{X}_p(t))^2}{p_{\max} - p_{\min} + 1}$$

$$\bar{X}_p(t) = \frac{\sum_{p=p_{\min}}^{p_{\max}} p \cdot X(t, p)}{p_{\max} - p_{\min} + 1}$$

ここで p_{\max}, p_{\min} はそれぞれ最大/最小ピッチである。

(3) Sound Accumulation

音高の集中度合いを表す特徴なので、同時発音高の偏りを表す値として、一般に広く用いられているエントロピーである $A_E(t, P')$, $A_E(t, U)$ とスパース性を測る量である $A_G(t, P')$, $A_G(t, U)$ の平均、分散、歪度、尖度で定量化できる。

$$A_E(t, X) = - \sum_{p=p_{\min}}^{p_{\max}} \frac{X(t, p)}{X_{\Sigma}(t)} \log \left(\frac{X(t, p)}{X_{\Sigma}(t)} \right)$$

$$A_G(t, X) = 1 - \sum_{p=p_{\min}}^{p_{\max}} \left(\frac{X(t, p)}{X_{\Sigma}(t)} \right)^2$$

$$X_{\Sigma}(t) = \sum_{p=p_{\min}}^{p_{\max}} X(t, p)$$

(4) Sound Texture Correlation

Sound Texture に関連する特徴量で、各楽器パートが交代して演奏されているか重複して演奏されているかを表す特徴である。しかし、音響情報からパートを精度良く分離することは困難である。一方、パート間の交代の典型例としての高音楽器と低音楽器の掛け合い、パート間の重複の典型例としての (solo に対する) tutti 等は、音域をいくつかに分割し、各々の音域間の相関として定量化できる。そこで、下式から求められる $S_T(i, j, P')$, $S_T(i, j, U)$ を特徴量として提案する。

$$S_T(i, j, X) = \frac{S(i, j, X)}{\sqrt{S(i, i, X)S(j, j, X)}}$$

$$S(i, j, X) = \sum_t (X_{\Sigma(i)}(t) - \bar{X}_{\Sigma(i)})(X_{\Sigma(j)}(t) - \bar{X}_{\Sigma(j)})$$

$$\bar{X}_{\Sigma(i)} = \frac{1}{T} \sum_t X_{\Sigma(i)}(t)$$

$$X_{\Sigma(i)}(t) = \sum_{p=p_{i, \min}}^{p_{i, \max}} X(t, p)$$

ここで i, j は音域番号、 T は入力音響スペクトルの時間軸方向データ数、 $p_{i, \max}, p_{i, \min}$ はそれぞれ、音域 i の最大/最小ピッチである。

(5) Chord Type

一般の和音認識では、短い時間における Chroma と和音のテンプレートとの類似度で和音を認識するが、本特徴量は楽曲における調とは独立な和音パターンの楽曲全体における頻度を表す特徴なので、音響の Chroma と、あらかじめ用意された和音パターンの Chroma に相当する Chord Profile との類似度である、下式に示す $C_T(t, C, P_p)$, $C_T(t, C, u_t)$ を特徴量として提案し、その t に対する平均、分散で定量化する。

$$C_T(t, C, Y) = \max_i (\theta(\text{rot}_i(\vec{c}(t, Y)), \vec{C}_C))$$

$$\theta(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \cdot \|\vec{y}\|}$$

ここで $\text{rot}_i(\vec{x})$ はベクトルの各成分を左へ i 成分回転する関数を表し、Chord Profile は和音 C に含まれる音名成分が 1, その他の音名成分が 0 の 12 次元ベクトルである PCS (Pitch Class Set) [9] である。

類似度の平均値が大きい場合には C が音響全体にわたって使われていたことを表し、平均値が小さく、分散も小さ

い場合には C があまり使われていないことを表し、分散が大きい場合には、音響の一部で使われていたことを表す。

次に、不協和音に関連する特徴量を述べる。音程ごとの協和度合いを順位付けした協和順位を用いて不協和度合いを定量化した Dissonance だけでなく、特定音程から構成される和音の利用が「作曲家らしさ」となる場合も考えられるため、各音程の含有率の大きさをそのまま定量化した Partial Intervals も特徴量として定義すべきである。

(6) Dissonance

同時に響く 2 つの音響が協和しない度合いである不協和音を表す特徴だが、音程の協和と不協和の境界は時代ごとに異なるうえ、文化によって境界自体が曖昧である場合もある。そこで、長谷川ら [5] により提案された協和度合いを順位付けした値である $o_c(d)$ を不協和度合いとして用い、二重音の Chord Profile との類似度を不協和度合いで重みづけした合計である下式の $D_S(t, P_p)$, $D_S(t, u_t)$ を特徴量として提案し、その平均と分散で定量化する。

$$D_S(t, Y) = \sum_{d=1}^6 o_c(d) C_T(t, D_d, Y)$$

ここで、 d は半音数で表した音程で、音程の転回を同一視すると短二度 (1) から増四度 (6) の値になる。また、 D_d は音程 d に対応する二重音である。

(7) Partial Intervals

順序に無関係に多次元化した、協和感を作り出している各音程の含有比率なので、同時に演奏されている楽音における各音程の含有率を表す特徴であり、Pitch に対する自己相関で定量化できる。そこで、下式に示す $\Delta_{PT}(d, P')$, $\Delta_{PT}(d, U)$ を特徴量として提案する。

$$\Delta_{PT}(d, X) = \frac{1}{T} \sum_{t, p} X(t, p) X(t, p + d)$$

ただし、自己相関では大音量部分の特徴量が支配的になるが、小音量部分も大音量部分と同様の印象を受ける場合も考えられるので、下式から求められる、音量で規格化した特徴量である $\delta_{PT}(d, P')$, $\delta_{PT}(d, U)$ も加えるべきである [5]。

$$\delta_{PT}(d, X) = \frac{1}{T} \sum_t \frac{\sum_p X(t, p) X(t, p + d)}{(X_{\Sigma}(t))^2}$$

本特徴量により、特定の音程を多用する楽曲の特徴をとらえることができる。

(8) Tail Segment Length

調性の複雑度合いを抽出するため、3.5 節で示した Tail Segment の長さを表す特徴であり、下式に示す $L_{TS}(k, P_p)$, $L_{TS}(k, u_t)$ を特徴量として提案し、その平均、分散、歪度、尖度で定量化する。

$$L_{TS}(k, Y) = t_{k, \max} - t_{k, \min}$$

(9) Interval Class Vector (ICV)

使われている各音間の音程の種類を用いて調性を表す特徴 [2] であり, Tail Segment 内の各音程出現頻度で調性の種類が, その変化度合いで調性の変化の特徴が求められる [5]. Tail Segment 中の音程の存在は Chroma Vector と音程分回転した Chroma Vector の内積で求められる. そこで, 下式に示す $V_{IC}(k, d, P_p)$, $V_{IC}(k, d, u_t)$ を特徴量として提案し, その k に対する平均, 分散で定量化する.

$$V_{IC}(k, d, Y) = \frac{\vec{C}_\Sigma(Y) \bullet \text{rot}_d(\vec{C}_\Sigma(Y))}{L_{TS}(k, Y)}$$

$$\vec{C}_\Sigma(Y) = \sum_{t \in T_S(k, Y)} \vec{c}(t, Y)$$

さらに, (7) と同様の理由で, 規格化した特徴量である $v_{IC}(k, d, P_p)$, $v_{IC}(k, d, u_t)$ も加えるべきである.

$$v_{IC}(k, d, Y) = \frac{\vec{c}_\Sigma(Y) \bullet \text{rot}_d(\vec{c}_\Sigma(Y))}{L_{TS}(k, Y)}$$

$$\vec{c}_\Sigma(Y) = \sum_{t \in T_S(k, Y)} \frac{\vec{c}(t, Y)}{\|\vec{c}(t, Y)\|}$$

(10) Key Profile Correlation

調性の中で用いられる音階の特徴なので, 音階のテンプレートとの類似度で求められる. そこで, 入力された音響とあらかじめ用意された音階 \vec{C}_K との類似度である下式に示す $K_P(K, P_p)$, $K_P(K, u_t)$ を特徴量として提案する.

$$K_P(K, Y) = \max_i(\theta(\text{rot}_i(\vec{c}(t, Y)), \vec{C}_K))$$

ここで \vec{C}_K は, 長谷川ら [5] により示された, 一般に用いられる主な音階の PCS である.

(11) Parallel Intervals

平行五度, 平行三度等, 音の垂直的結合の連続の特徴を抽出するため, 同じ全音階差 [5] の連続度合いを表す特徴であり, 隣り合う時刻で同じ全音階差の異なる二重音が含まれる割合なので, t と $t+1$ における二重音 D_d の Chord Type 特徴量との類似度で求められる. そこで, 下式に示す $\Delta_{PR}(t, d_w, P_p)$, $\Delta_{PR}(t, d_w, u_t)$ を特徴量として提案し, その t に対する平均, 分散で定量化する.

$$\begin{aligned} \Delta_{PR}(t, d_w, Y) &= \sum_{d_1, d_2 \in D(d_w)} C_T(t-1, D_{d_1}, Y) C_T(t, D_{d_2}, Y) \end{aligned}$$

ここで, $D(d_w)$ は全音階差が d_w となる音程の集合である.

(12) Pitch Difference

音程, すなわち隣り合う楽音間の音高差の頻度を表し, 音の動きをとらえる特徴であるが, 一般の音楽音響で同時に演奏されている楽音すべてをパートごとに精度良く分離することは困難である. さらに, 複数の楽音の動きをとらえようとすると, 時間経過の中でそれらの楽音がどのような組合せで対応しているのかを精度良く認識することも困

難である. 一方, 支配的な音高の動きは, パワーが大きい音高の時間的変化を計数することにより求められる. そこで, 下式に示す $D_P(d, P')$, $D_P(d, U)$, $D_P(d, H_h)$ を特徴量として提案する.

$$D_P(d, X) = \frac{n_D(d, X)}{\sum_i n_D(i, X)}$$

$$n_D(d, X) = \text{count}\{p_D(X, t) - p_D(X, t-1) = d\}$$

$$p_D(X, t) = \arg \max_p(X(t, p))$$

ここで, d は半音数で表した音高差である. また, 音高差を求める際, 分布として P' , U に加えて, アタック音のノイズ除去のために H_h を用いる.

次に, リズムに関する特徴量を述べる. リズム特徴は音価を基に計算されるが, 一般の音楽音響から音価を精度良く認識することは困難であり, 複雑な語法等の認識はできないので, 単純な音価比, テンポ, 音価種類数を特徴量とする.

(13) Rhythm

音符の音価組合せである Surface Rythm の特徴として音価比の出現頻度, およびテンポを表す特徴であり, 音高の変化や音量の変化を音の区切りとして求められる隣り合う音価の比と, 一定時間内の音の区切り数により求めることができる. そこで, 下式に示す $R_H(R, H_h)$, $R_H(R, P_h)$, $N_R(H_h)$, $N_R(P_h)$ を特徴量として提案する.

$$R_H(R, Z) = \frac{n_R(R, Z)}{\sum_r n_R(r, Z)}$$

$$N_R(Z) = \frac{\text{count}(T_D(k, Z))}{T}$$

$$n_R(R, Z) = \text{count}\left\{Q\left(\frac{T_D(k, Z)}{T_D(k-1, Z)}\right) = R\right\}$$

ここで音価比 R はあらかじめ複数定められた 1 以下の単純な分数で, $Q(r)$ は $\min(r, 1/r)$ を前記の単純な分数比に量子化する関数を表す. k 番目の音価 $T_D(k, Z)$ は, 前処理された入力信号である $x(t, Z)$ を平滑化し, 局所的に最大値となる時刻の集合を求め, それらを境界と考える. ここで, $x(t, Z)$ は下式で表され, 用いる分布 Z により, H_h の場合には音高の変化, P_h の場合にはパワーである.

$$x(t, H_h) = \frac{\sum_f |H_h(t, f) - H_h(t-1, f)|}{\sum_f H_h(t, f)}$$

$$x(t, P_h) = \sum_f P_h(t, f)$$

(14) Rhythm Entropy

Rhythm Level, すなわち音価の種類数を表す特徴である. 音価は二分音符, 四分音符, 八分音符等, 等比級数的

に定められているので、対数音価のパラツキ（分散）で定量化できる．そこで、下式に示す $R_E(H_h)$, $R_E(P_h)$ を特徴量として提案する．

$$R_E(Z) = \frac{1}{N_{TD}} \sum_k (\log(T_D(k, Z)) - \overline{\log(T_D)})^2$$

(15) Timbre

音響情報には楽譜情報に加えて音色情報が含まれ、有用な特徴となりうる．音色には、音響全体の音色と、演奏されている楽器の音色が考えられる．いずれの音色も時刻ごとに変化するので、その変化度合いも表す統計量を用いた特徴量が必要である．

音響全体の音色を表すスペクトルは、ジャンルの一種である楽曲の種類（管弦楽曲/協奏曲/室内楽曲/声楽曲等）によって、同じ作曲家でも大きく異なるが、他の特徴量との組合せによって、たとえばジャンルごとの「作曲家らしさ」をとらえるために必要となる．音響全体の音色の特徴は、振幅スペクトル形状の中の倍音構造に表れるので、一般に用いられている、振幅波形形状の対数周波数における特徴を表す MFCC の t に対する平均、分散で定量化できる．

一方、演奏されている楽器音色の特徴は、楽音ごとの倍音の減衰形状に表れるので、Specmurt H が時刻ごとの倍音構造を表すことから、 $H_s(t, m)$ と、下式の倍音減衰率 $H_G(t)$ と減衰に対する残差 $H_E(t)$ を特徴量として提案し、その分布形状を表す平均、分散、歪度、尖度により定量化する．また、楽曲全体の平均的楽器音色である $H_s(t, m)$ の t に対する平均値の減衰率 $H_G(t)$ ならびに残差 $H_E(t)$ も特徴量として提案する．

$$H_G(t) = \frac{N_{H_s} \sum_m m H_s(t, m) - \sum_m m \sum_m H_s(t, m)}{N_{H_s} \sum_m m^2 - \left(\sum_m m\right)^2}$$

$$H_E(t) = \frac{1}{N_{H_s}} \sqrt{\sum_m (H_s(t, m) - m H_G(t) - H_B(t))^2}$$

$$H_B(t) = \frac{\sum_m m^2 \sum_m H_s(t, m) - \sum_m m \sum_m m H_s(t, m)}{N_{H_s} \sum_m m^2 - \left(\sum_m m\right)^2}$$

ここで、 $H_G(t)$, $H_B(t)$ はそれぞれ、 m を x , $H_s(t, m)$ を y として求めた回帰直線 $y = ax + b$ における傾き a と y 切片 b である [10]．また、 $H_E(t)$ は下式で表される回帰直線の残差平方和 S_e の二乗根を倍音数 N_{H_s} で割った値である．

$$S_e = \sum_i (y - \hat{y}_i)^2 = \sum_i (y - ax_i - b)^2$$

5. 印象特徴空間

作曲家を目的変数、音楽特徴量を説明変数として判別分析を行うことにより、同一作曲家の作品特徴点が近接して

配置されるように最適化された印象特徴空間各座標値への、音楽特徴量からの変換関数が求められる．同変換関数が音楽特徴量に対して単調写像であれば、音楽特徴量の類似した楽曲が近接して配置されるので、前記印象特徴空間上で、近い印象を与える楽曲が近接して配置されることが期待できる．以上から、線形の判別分析手法である正準判別分析を用いて音楽特徴量を線形変換することにより印象特徴空間が得られると考えられる．ただし、統計処理における次元の呪い問題を考慮し、特徴量を取舍選択する必要がある．選択アルゴリズムとしては、全探索より短い処理時間で高いパフォーマンスが得られる Pudil らにより提案された SFFS (Sequential Floating Forward Search) を用いた [11], [12]．

6. 印象認識評価

6.1 評価の目的と評価項目

演奏家による音楽音響から本論文で示した特徴量を抽出して構築された印象特徴空間において、「作曲家らしさ」の印象が近い作品の特徴点が近接して配置されていることを確認することを目的とする．そのために、以下に示す評価を行った．

- (A) **特徴量予備評価**：作曲家ごとに抽出された特徴量に関し、すでに知られている作曲家特徴との比較を行った．
- (B) **同一作曲家作品近接配置評価**：同一作曲家の作品特徴点が近接して配置されていることを確認するために、印象特徴空間を用いて作曲家予測を行い、判別精度を求めた．
- (C) **類似作曲家の近接配置評価**：類似した作曲家の作品特徴点が近接して配置されていることを確認するために、時代ならびに地域が類似する作曲家の作品特徴重心点の印象特徴空間上での相対位置を視覚化した．
- (D) **未知作品の類似作曲家への近接配置評価**：印象特徴空間構築で用いた作曲家以外の作曲家（以下、「未知作曲家」と記す）の作品特徴点が、その作曲家と類似の作曲家の作品特徴点群に近接して配置されていることを確認するために、未知作曲家の作品に関して、印象特徴空間を用いた作曲家予測を行い、時代/地域が

表 3 音楽特徴量算出のためのパラメータ

Table 3 Parameters for musical feature calculations.

パラメータ	値
フレーム長	743 ms
フレームシフト長	372 ms
Specmurt H 倍音数	10
Tail Segment N_c	5, 7
Sound Expanse p_{min}	21
Sound Expanse p_{max}	108
Sound Texture 音域数	3
Sound Texture $p_{1,min}$	50
Sound Texture $p_{2,min}$	79

表 4 評価データ

Table 4 Data for evaluations.

Composer	時代	地域	管弦楽	協奏曲	室内楽	声楽曲	データ数	学習用	テスト用	Time (hms)
Bach	Baroque	Germany	25	51	199	158	433	390	43	25:59:37
Handel	Baroque	Germany	25	44	134	56	259	234	25	11:17:41
Vivaldi	Baroque	Italy		111	30	39	180	162	18	9:16:05
Haydn	Classic	Germany	97	39	111	94	340	306	34	27:13:50
Mozart	Classic	Germany	173	32	189	35	429	387	42	35:28:05
Beethoven	Classic	Germany	62	17	182	48	309	279	30	33:57:40
Elgar	Romantic	Britain	37	7	11	28	83	75	8	10:03:38
Chopin	Romantic	East Europe		9	83	19	111	48	5	8:52:15
Liszt	Romantic	East Europe	11	10	32		53	171	18	9:04:01
Tchaikovsky	Romantic	East Europe	83	16	72	18	189	146	16	16:53:59
Dvorak	Romantic	East Europe	54	11	67	30	162	82	9	19:05:33
Rimsky-Korsakov	Romantic	East Europe	31	3	6	51	91	135	14	7:17:34
Berlioz	Romantic	France	30			119	149	43	4	13:16:24
Frank	Romantic	France	7	2	29	9	47	130	14	7:30:17
Saint-Saens	Romantic	France	29	22	44	49	144	79	8	13:15:50
Bizet	Romantic	France	27		3	57	87	118	13	5:45:55
Faure	Romantic	France	10	5	77	39	131	154	17	10:09:54
Schubert	Romantic	Germany	30		81	60	171	167	18	19:52:08
Mendelssohn	Romantic	Germany	59	15	69	42	185	202	22	17:09:26
Schumann	Romantic	Germany	26	9	123	66	224	156	17	15:22:52
Brahms	Romantic	Germany	33	13	98	29	173	108	12	23:15:06
Mahler	Romantic	Germany	72		1	47	120	144	16	16:38:58
Rossini	Romantic	Italy	16		33	111	160	130	14	13:31:53
Verdi	Romantic	Italy	14	4	4	122	144	153	17	10:54:01
Puccini	Romantic	Italy	12		5	153	170	100	11	9:40:13
Grieg	Romantic	North Europe	25	3	31	16	75	68	7	6:27:15
Sibelius	Modern	East Europe	62	3	17	42	123	111	12	11:14:32
Rachmaninov	Modern	East Europe	13	12	32	32	89	81	8	10:09:40
Debussy	Modern	France	13		108	58	178	161	17	11:42:29
Satie	Modern	France	4		169	33	206	186	20	8:12:46
Ravel	Modern	France	34	4	56		94	85	9	7:15:15
		31名	1,114	442	2,096	1,660	5,309	4,791	518	445:54:52

共通の作曲家に判別されることを確認した。

- (E) 音楽特微量ごとの寄与率評価：4章で述べた音楽特微量が印象特徴空間構築に寄与していることを確認するために、各音楽特微量の寄与率を求めた。

6.2 評価方法

- (1) 評価データから本論文で述べた音楽特微量を抽出した。抽出の際に用いたパラメータ値を表 3 に示す。ここでフレーム長とフレームシフト長は 3.1 節の $P(t, f)$ ならびに 3.2 節の $v(x)$ を求める際のパラメータである。また、Sound Texture 特微量において、 $p_{0,\min} = p_{\min}$, $p_{0,\max} = p_{1,\min}$, $p_{1,\max} = p_{2,\min}$, $p_{2,\max} = p_{\max}$ である。
- (2) (A) のために、抽出した特微量の作曲家重心座標値を正規化した。
- (3) (B) のために、評価データを学習用データとテスト用データに分け、正準判別分析を用いて学習用データの特微量から特徴空間を構築し、テスト用データの特微量を用いて予測を行い、以下に示す値を求めた。
 - ① 正解率：正しい作曲家を予測したデータ数の全データ数に対する割合
 - ② 上位含有率 [5]：予測結果の事後確率が 2 位以内ならびに 3 位以内に正解が含まれるデータ数の全データ数に対する割合
- (4) 正準判別分析を用いて評価データすべての特微量から特徴空間を構築した。
- (5) (C) の評価と、構築した印象特徴空間と楽譜情報から特微量を抽出した場合 [5] の特徴空間とを比較するた

めに、文献 [5] と同様に作曲家の判別得点重心に対して、ユークリッド距離を用い、ウォード法で階層的クラスタ分析を行った。

- (6) (C) の時代に関する評価のために、本論文で述べた印象特徴空間上の時間に関係すると考えられる軸の作曲家特徴重心点座標と作曲者の没年との関係を求めた。さらに、各音楽特微量の時代に対する寄与度合いを評価するために、同軸の判別関数の音楽特微量 f に対応する係数 a_f と同特微量の標準偏差の積を寄与度 c_f として求めた。判別関数は係数を重みとした特微量の線形和で求められるので、係数と特微量の変化度合いである標準偏差との積は判別関数に対する影響度を表すと考えられる。
- (7) (C) の地域に関する評価のために、(5) で述べた軸を除く軸が形成する空間上の座標を、一次変換、投影、多次元尺度法により平面上の点に変換した。なお、ここで用いた変換は非可逆なので、一意に各音楽特微量の寄与度を求めることはできない。
- (8) (D) のために、未知作曲家の作品の作曲家予測を行い、(a) 時代、(b) 地域が共通の作曲家に判別される割合を求めた。
- (9) 最後に (E) のために、印象特徴空間形成に対する各音楽特微量の寄与を確認するために、特微量 f の寄与率 C_f を以下のように定義し算出した。

$$C_f = \frac{A - A_{-f}}{A}$$

ただし、 A は特微量すべてを用いた場合の正解率、 A_{-f} は特微量 f のみを除いた場合の正解率である。

表 5 未知作曲家のテストデータ

Table 5 Data for unknown composer evaluation.

作曲家	曲数
Chedeville	28
Delius	10
Vaughan-Williams	5
Walton	9

したがって寄与率は、特微量 f を省略して同様の実験を行ったときの精度低下率を表し、同値が大きい特微量は作曲家らしさの判別に大きく寄与していると考えられる。

6.3 評価データ

評価に用いたデータはすべて、RIFF waveform 形式で保存されたサンプリング周波数 44.1kHz, 16 bit の職業演奏家による音楽音響データである。印象空間構築ならびに作曲家判別精度評価には、表 4 に示す後期バロックから 1900 年初頭までの作曲家の 31 作曲家, 5,309 曲のすべての部分を用いた。各作曲家の時代, 地域, 楽曲の種類ごとの曲数と合計曲数, 6.2 節 (3) における学習用データ数とテスト用データ数, 総時間長を示す。

また、(D) のための作品として、印象空間構築には用いられていない以下にあげる作曲家の作品を用いた。

- (a) 時代に関する評価: Vivaldi, Bach, Handel と同じバロック時代の作曲家として新たに Chedeville の作品
- (b) 地域に関する評価: Elgar と同じイギリスの作曲家として新たに Delius, Vaughan-Williams と Walton の作品

各作曲家とその楽曲数を表 5 に示す。

7. 評価結果と考察

(A) 特微量予備評価

2 章で述べたような短い分析範囲から抽出可能な特徴の中で知られている作曲家の特徴は少ないが、以下にそれらの特徴の例と、抽出された正規化特微量との比較を示す。なお、正規化特微量は、正の場合その作曲家の特徴であることを表し、値の大きさは、全作曲家の特微量の標準偏差に対する割合を表す。

- (1) 減七の和音は主にバロックで用いられ、Bach が特に頻繁に使ったことが知られている [13]。減七に相当する Chord Type 特微量の頻度特微量は、5 章で述べた印象特徴空間構築の際に SFFS により選択されていないが、和音を構成する特徴的音程である減五度の Partial Interval 特微量が Bach の場合、正の値 1.55 を示している。他の作曲家では時代の進行とともに値が小さくなる傾向があり、正である作曲家はバロック、古典の作曲家ならびに Tchaikovsky, Dvorak, Faure, Schubert, Mendelssohn, Brahms, Satie である。

表 6 正解率と上位含有率

Table 6 Accuracy and dominant inclusion ratio.

正解率	65.3%
2位含有率	75.3%
3位含有率	81.3%

- (2) モーツァルト五度として知られる平行五度は Mozart が好んで用いたことが知られている [17]。平行五度に対応する Parallel Interval 特微量は (1) と同様に選択されていないが、和音内の五度音程頻度を表す五度の Partial Interval 特微量は Mozart の場合、正の値 2.09 を示している。他の作曲家で 2 以上の値を示しているのは Vivaldi と Satie のみである。

以上の例から、多重和音や垂直的結合の連続のような複合的な特徴は抽出されない場合もあるが、部分的な特徴は抽出できていることが示唆される。

(B) 同一作曲家作品近接配置評価

評価結果を表 6 に示す。31 名の作曲家判別分析で 65.3% の精度が得られ、81.3% の楽曲に関して推測結果の上位 3 位までに正解が含まれている。同結果から、同一作曲家の楽曲の特徴点が印象特徴空間上で近接して配置されていると考えられる。

次に、作曲家ごとの予測結果の曲数と正解率を表 7 に示す。作曲家は年代, 地域別に並べてあり、各行が各作曲家の楽曲を表し、各列が予測された作曲家を表す。Bach~Beethoven がバロック~古典派, Elgar~Grieg がロマン派, Sibelius~Ravel が近代の作曲家である。この中で、近代の作曲家の楽曲を、時代として隣接していないバロック~古典派の作曲家と予測する誤認率は 3.0% であり、逆は 0.5% のみである。さらに、バロック~古典派の作曲家であることを判別する正解率は 88.5% に達する。

したがって、本論文で構築した印象特徴空間は、時代の差異を表現できていることを示唆するとともに、特にバロック~古典音楽に有効であると考えられる。

(C) 類似作曲家の近接配置評価

まず、階層クラスタ分析結果を図 1 に示す。互いに折れ線で結ばれた 2 つの作曲家もしくはクラスタが、階層的にクラスタを形成していることを表す。同図にアルファベットで示したクラスタに属する作曲家に関し、以下に、それらの共通性と楽譜情報から特微量を抽出した場合 [5] のクラスタとの比較を列挙する。

- A) 後期バロックから古典派の作曲家がクラスタを形成している。特に共通の様式を持っているバロックの Vivaldi, Bach, Handel と、古典派の Haydn, Mozart が文献 [14], [15], [16] でそれぞれ下位クラスタを形成していることから、本論文で提案した特微量がそれらの様式をとらえることができたと考えられる。ただし、文献 [5] では Beethoven も古典派のクラスタに属していたが、本研究では B) のドイツ系ロマン派に含まれ

表 7 作曲家ごとの予測結果と正解率

Table 7 Detailed estimation results and accuracy for each composer.

作曲家	予測結果																			Accuracy												
	Bach	Handel	Vivaldi	Haydn	Mozart	Beethoven	Elgar	Liszt	Tchaikovsky	Dvorak	Rimsky-Korsakov	Berlioz	Franck	Saint-Saens	Bizet	Faure	Schubert	Mendelssohn	Schumann		Brahms	Mahler	Rossini	Verdi	Puccini	Chopin	Grieg	Sibelius	Rachmaninov	Debussy	Satie	Ravel
Bach	34	2	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	79.1%
Handel	1	19	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	76.0%	
Vivaldi	3	0	14	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	77.8%	
Haydn	1	2	1	23	5	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	67.6%	
Mozart	0	0	0	4	33	1	0	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	78.6%	
Beethoven	0	0	1	0	1	20	0	0	0	0	0	0	0	0	1	0	0	1	2	0	0	0	1	3	0	0	0	0	0	0	66.7%	
Elgar	0	0	0	0	0	5	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	62.5%	
Liszt	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.0%	
Tchaikovsky	0	0	0	1	1	0	0	8	0	0	0	0	1	1	0	1	0	2	0	0	0	0	0	0	2	1	0	0	0	0	44.4%	
Dvorak	0	0	0	0	0	0	0	10	0	1	0	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0	62.5%	
Rimsky-Korsakov	0	0	0	0	0	0	0	1	5	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	55.6%	
Berlioz	0	0	0	0	0	0	0	2	0	8	0	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	57.1%	
Franck	0	0	0	0	0	0	0	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50.0%	
Saint-Saens	0	0	0	0	0	2	1	0	1	0	1	0	0	4	0	1	0	1	0	1	0	0	1	0	0	0	1	0	0	0	28.6%	
Bizet	0	0	0	1	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87.5%	
Faure	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	8	1	0	0	0	0	0	0	0	0	1	0	0	0	0	61.5%	
Schubert	0	0	0	0	0	2	0	0	2	0	0	0	0	0	1	9	1	2	0	0	0	0	0	0	0	0	0	0	0	0	52.9%	
Mendelssohn	0	0	0	1	1	1	0	0	2	0	0	0	2	0	0	1	8	0	1	0	0	1	0	0	0	0	0	0	0	0	44.4%	
Schumann	0	0	0	0	1	1	0	0	1	2	1	0	0	0	0	0	0	0	14	1	0	0	0	0	0	0	0	0	1	0	63.6%	
Brahms	0	0	0	0	1	2	0	0	0	2	0	0	0	0	0	0	0	0	0	11	0	0	0	0	1	0	0	0	0	0	64.7%	
Mahler	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	6	1	0	0	0	0	0	1	0	0	0	50.0%	
Rossini	0	0	0	0	1	3	0	0	0	0	1	0	0	1	0	0	0	0	1	0	8	1	0	0	0	0	0	0	0	0	50.0%	
Verdi	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	10	1	0	0	0	0	0	0	0	0	0	71.4%	
Puccini	1	0	0	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0	1	1	0	10	0	0	0	0	0	0	58.8%	
Chopin	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	7	0	0	0	1	1	0	63.6%	
Grieg	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	85.7%	
Sibelius	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	2	1	0	0	0	0	6	1	0	0	0	50.0%	
Rachmaninov	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	4	0	0	0	50.0%	
Debussy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	11	2	1	64.7%	
Satie	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	18	0	0	90.0%	
Ravel	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	5	0	55.6%	

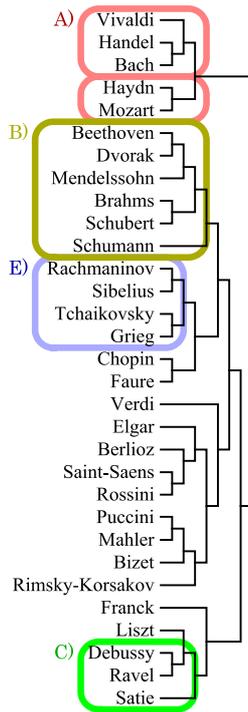


図 1 作曲家に関するクラスタ分析結果

Fig. 1 A result of cluster analysis.

る。Beethoven はロマン派の傾向もある [16] ためと考えられるが、本研究で新たに加えた音色の特徴が特にロマン派の傾向を示すことを示唆する。

B) Mahler を除くドイツ系（ドイツ，オーストリア）ロマン派の作曲家と Beethoven がクラスタを形成している。ドイツ・ロマン派の音楽も共通の様式を持って

る [14] ことから、本論文で提案した特徴量がそれらの様式をとらえることができたと考えられる。また、文献 [5] も Beethoven を除き本研究と同様に、ドイツ系ロマン派がクラスタを形成している。

C) 共通の様式を持っているフランス印象派 [14] の Debussy と Ravel がクラスタを形成していることから、本論文で提案した特徴量によってフランス印象派の特徴をとらえることができたと考えられる。さらに同時代のフランス人である Satie がクラスタを形成している。いずれも文献 [5] の結果と同様である。

E) 文献 [5] と同様に、Sibelius が、音楽を勉強していた頃の祖国フィンランドがロシアの支配下にあったため影響を受けた Tchaikovsky [17] とクラスタを形成している。さらに、Tchaikovsky と同じロシアの作曲家である Rachmaninov と、Sibelius と同じ北欧の作曲家である Grieg が同じクラスタに含まれる。

クラスタ分析では、時代と地域々々に対する包括的評価はできないが、バロックと古典、ドイツ系ロマン派、フランス印象派、ロシア・北欧系ロマン派の作曲家がそれぞれ印象特徴空間上で近接して配置されていることが分かる。また、新たに追加されたイタリアの作曲家を含むクラスタを除き類似したクラスタを形成していることから、楽譜情報から特徴量を抽出した場合 [5] の特徴空間と本印象特徴空間は作曲家の位置関係が類似していると考えられる。

次に、作曲家の没年を X 軸、印象特徴空間の第 1 軸（第 1 判別関数）の特徴重心点座標を Y 軸として作曲家をプロットした図を図 2 に示す。同軸が作曲家の没年と正の

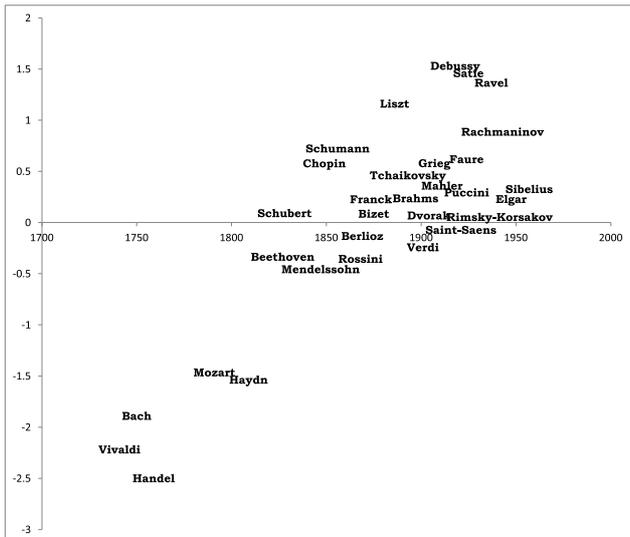


図 2 作曲家没年と第 1 判別関数値との関係

Fig. 2 Relation between death year and the 1st discriminant value.

表 8 第 1 判別関数の寄与度

Table 8 Contributions in the 1st discrimination function.

特徴量	寄与度
ICV $d=4$ ($N_c=7$) Normalized Average	2.26
Pitch Distribution Skewness	2.09
Pitch Distribution Average	1.78
ICV $d=3$ ($N_c=7$) Normalized Average	-1.77
ICV $d=5$ ($N_c=7$) Normalized Average	-1.44
ICV $d=2$ ($N_c=7$) Normalized Average	1.43
ICV $d=4$ ($N_c=5$) Normalized Average	0.86
ICV $d=4$ ($N_c=7$) Normalized Average	-0.69
MFCC #2 Average	0.55
ICV $d=3$ ($N_c=7$) Normalized Variance	0.55
Chord Type Dual $d=1$ Average	0.47

相関を示している（相関係数 0.845）ので、この軸が、作曲家の時代を表していると考えられる。また、音楽特徴量に対する寄与度を、寄与度絶対値が最大寄与度の 20%までの特徴量に関して表 8 に示す。ここで、正の値は時代とともに特徴量が増加することを、負の値は減少することを表す。文献 [2] で示されたとおり、音符集合と音階との類似度である ICV 特徴量が大きく寄与していることが、音響から特徴量を抽出した本研究でも示された。また、短 2 度の Chord Type 特徴量が正の値であるのは、時代とともに不協和音程である短 2 度が使われるようになってきた [17] ことに対応すると考えられる。次に、ロマン派の作曲家の特徴重心点を、評価方法 (7) で示した方法で変換した配置を図 3 に示す。ドイツ・オーストリア系の作曲家を□、東欧・ロシア・北欧系を◇、イタリア系を△、フランス系を○、イギリス系を×で表している。ドイツ・オーストリア系の作曲家が右、東欧・ロシア・北欧系が中央上、イタリア系が中央下、フランス系が左に集まっていることが分かる。

作曲家の判別分析によって最適化された特徴空間で、以上で示したように時代と地域が類似した作曲家特徴点が近接して配置されていることが示されたことから、本論文で

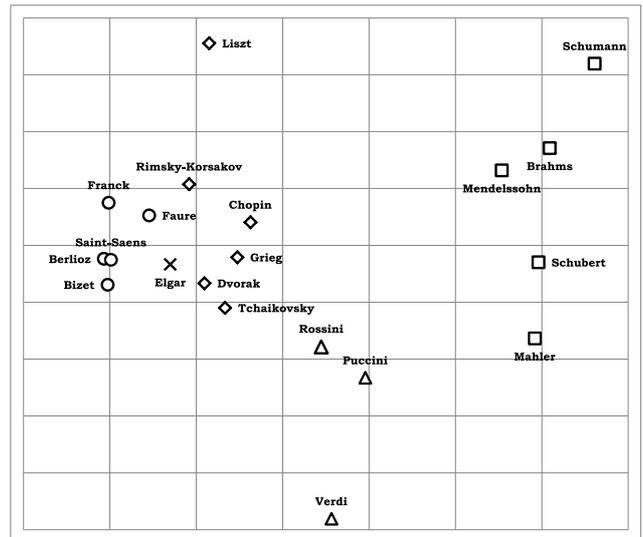


図 3 印象空間上の作曲家重心点の配置

Fig. 3 Composer centroids in impression feature space.

提案した音楽特徴量で Mayer [1] が述べたような時代や地域の特徴を含めた階層的な作曲家の特徴をとらえることができたと考えられる。

(D) 未知作品の類似作曲家への近接配置評価

評価結果を表 9 に示す。各列に示した作曲家の楽曲の中で、印象特徴空間構築に用いた各行の作曲家に予測された楽曲数が記載されている。また、Chedeville の場合には同じバロックの作曲家である Vivaldi, Bach, Handel であると予測された楽曲数が、Delius, Vaughan-Williams と Walton の場合には同じイギリスの作曲家である Elgar であると予測された楽曲数が、それぞれ灰色で示され、最下行にそれらの楽曲数の全楽曲数に対する割合が含有率として示されている。31 作曲家の中で、Chedeville の場合には (a) 時代が共通しているバロックの作曲家に 86% 判別され、イギリスの作曲家の場合には (b) 地域が共通している Elgar の楽曲に 57% 判別されている。以上から、印象特徴空間上で、未知の作品の特徴点が時代/地域が共通した作曲家の作品に近接して配置されていることが分かる。

したがって、本論文で提案した音楽特徴量を用いた既知の作曲家の分析により、時代や地域の特徴を含めた印象をとらえることができることが示唆される。

(E) 音楽特徴量ごとの寄与率評価

音楽特徴量ごとの寄与率を表 10 に示す。本論文で定義したすべての特徴量が寄与していることが分かる。特に、音色を表す Timber, 使われている和音の頻度を表す Chord Type, 調性を表す Interval Class Vector, 音高推移の特徴を表す Pitch Difference, 同時に発音されている音の音程の特徴を表す Partial Intervals が印象特徴空間で重要な役割を果たしていることが分かる。

以下に、寄与率が 10% 以上の音楽特徴量に関して、詳細に述べる。なお、楽譜情報から特徴量を抽出した場合 [5]

表 9 未知作曲家の予測結果

Table 9 Estimation results for unknown composers.

Composer	Chedeville	Delius	Vaughan-Williams	Walton
Bach	12			
Handel	8			
Vivaldi	4			
Haydn				
Mozart		1		
Beethoven				
Schubert				
Mendelssohn				
Schumann				
Brahms				
Mahler		1		
Chopin				
Liszt				
Tchaikovsky				1
Dvorak		1		1
Rimsky-Korsakov				
Rossini				
Verdi				
Puccini				
Berlioz				
Franck				
Saint-Saens		3		2
Bizet	1			
Faure				
Grieg	1			
Elgar		4	4	5
Debussy	1			
Satie				
Ravel				
Sibelius				1
Rachmaninov				
含有率	86%			57%

表 10 特徴量ごとの寄与率

Table 10 Contribution of each feature.

Feature	寄与率
(15) Timbre	27.27%
(5) Chord Type	12.47%
(9) Interval Class Vector	12.38%
(12) Pitch Difference	11.55%
(7) Partial Intervals	9.77%
(10) Key Profile Correlation	6.20%
(1) Pitch Distribution	5.82%
(4) Sound Texture Correlation	3.16%
(11) Parallel Intervals	3.01%
(8) Tail Segment Length	2.10%
(3) Sound Accumulation	1.82%
(13) Rhythm	1.77%
(2) Sound Expanse	1.73%
(6) Dissonance	0.77%
(14) Rhythm Entropy	0.19%

(以下、楽譜特徴量と記す)においても、音符情報に含まれない Timber 特徴量を除く以下に列挙する特徴量の寄与率が高く、類似の傾向が見られる。

(15) Timber

Timber の寄与率のうち、15.8%は音響全体の特徴を表す MFCC 特徴量により、残りが楽器音色特徴量の寄与率である。主な特徴量(寄与率 > 0.7%)を表 11 に示す。第 2, 3 倍音平均, 第 5 倍音の分散, 平均倍音の減衰率と残差, 倍音残差の平均が寄与している。楽曲に使われている楽器の組合せにより平均倍音の減衰形状が変わり、特に形状の違いが第 2, 3 倍音に表れることから、これらの特徴量が

表 11 主な楽器音色特徴量の寄与率

Table 11 Outstanding contributions of timbre.

特徴量	寄与率
Harmonics #2 Average	0.94%
Harmonics #3 Average	0.94%
Harmonics #5 Variance	0.75%
Average Harmonics Gradient	0.73%
Harmonics Error Average	0.73%
Average Harmonics Error	0.71%

表 12 主な Chord Type 特徴量の寄与率

Table 12 Outstanding contributions of chord type.

特徴量	寄与率
Minor Average	0.79%
Diminish 7th Variance	0.75%
Dual $d=2$ Average	0.73%
7th Average	0.66%
Sus4 Average	0.64%
Dual $d=1$ Average	0.62%
Diminish Variance	0.60%
Minor 7th Average	0.56%
Major Variance	0.54%

表 13 主な ICV 特徴量の寄与率

Table 13 Outstanding contributions of ICV.

特徴量	寄与率
$d=1 (N_c=7)$ Normalized Average	1.01%
$d=4 (N_c=7)$ Normalized Average	0.83%
$d=2 (N_c=7)$ Normalized Average	0.75%
$d=5 (N_c=7)$ Normalized Average	0.71%
$d=5 (N_c=5)$ Normalized Average	0.69%
$d=3 (N_c=7)$ Normalized Average	0.62%
$d=6 (N_c=7)$ Normalized Average	0.51%

寄与していると考えられる。

(5) Chord Type

主な特徴量(寄与率 > 0.5%)を表 12 に示す。短三和音, 属七, Sus4, Minor 7th, および長二度と短二度二重音との類似度平均, Diminish (7th), 長三和音との類似度分散が寄与している。二度音程を含む二度二重音, 7th, Sus4, および減三和音を含む Diminish (7th) が寄与していると考えられる点で、楽譜特徴量と類似した結果が得られた。

(9) Interval Class Vector

主な特徴量(寄与率 > 0.5%)を表 13 に示す。いずれも規格化された特徴量であり、 $N_c = 7$ における ICV すべての要素と $N_c = 5$ における四度音程の出現頻度が寄与している。ICV は調性の特徴であり、全音階等の多くの音階が七音階から構成されていることから、 $N_c = 7$ の特徴量が寄与していると考えられる。また、小音量部分の特徴も印象に影響を与えることから、基準化された特徴量が寄与していると考えられる。

(12) Pitch Difference

主な特徴量(寄与率 > 0.5%)を表 14 に示す。1 度, 短 2 度, 短 3 度, 5 度, 短 7 度が寄与している。本特徴量に関しては、主に跳躍音程が寄与していた楽譜特徴量と異なる。これは、楽譜特徴量がパートごとの音程を抽出したのに対し、本特徴量では支配的な音高の動きのみを抽出して

表 14 主な Pitch Difference 特徴量の寄与率

Table 14 Outstanding contributions of pitch difference.

特徴量	寄与率
$d=0$	0.68%
$d=+3$	0.58%
$d=+7$	0.58%
$d=+1$	0.54%
$d=+10$	0.51%

いたためと考えられる。

8. まとめ

本論文では、クラシック音楽の嗜好検索において重要な「作曲家らしさ」に関する検索を実現するために、作曲家の作品に近い印象を与える楽曲の特徴点が近接して配置されるような特徴空間の尺度を、音楽音響から得られる客観量から構築することを目的として、複数の作曲家に共通する定性的特徴が網羅的に述べられている LaRue らの総合的様式分析に記載されている内容に基づき、音響から抽出可能な情報を用いた 15 の特徴量を提案した。また、これらの特徴量から構築される印象特徴空間の妥当性を確認するため、以下に述べる評価を行った。まず、作曲家判別精度の評価を行い、後期バロック～18 世紀初頭の 31 作曲家、5,309 曲の判別において 65.3% の正解率を得、81.3% が予測結果第 3 位までに正解が含まれることが分かった。以上から、同一作曲家の作品特徴点が近接して配置されていることを確認した。次に、印象特徴空間における 1 つの座標軸が作曲家の年代と相関している（相関係数 0.845）ことを、前記座標軸以外の軸の作曲家特徴重心点座標を一次変換、投影、多次元尺度法により平面上の点に変換することにより地域が共通する作曲家特徴点が印象特徴空間上で互いに近接していることを確認した。さらに、印象特徴空間構築に用いた 31 作曲家以外の作曲家の作品に対して作曲家予測を行い、バロックの作曲家の作品の 86% が 3 名のバロックの作曲家のいずれかに、3 名のイギリスの作曲家の作品の 57% がイギリスの作曲家である Elgar に予測された。以上から、印象特徴空間構築に用いられなかった作曲家の作品も、印象特徴空間上で、時代や地域が共通する作曲家の作品に近接して配置されていることを確認した。

Meyer によると、時代や地域が共通の作曲家は互いに類似した印象を受けるので、作曲家の判別で良好な結果が得られただけでなく、時代/地域が共通する作曲家が印象特徴空間上で近接して配置されたことから、本論文で提案した印象特徴空間を用いて「作曲家らしさ」の印象認識が可能となると考えられる。

今後は、編曲等、複数の作曲家が関係する楽曲の特徴、同作曲家の生涯における特徴の変遷等、「作曲家らしさ」をさらに詳細化した研究、および、本論文における分析結果を事前知識として用い、「作曲家らしさ」以外の主観性・曖昧性をともなう形容語等の印象認識の研究に取り組みたい。

参考文献

- [1] Meyer, L.B.: *Style and Music*, The University of Chicago Press (1989).
- [2] Huovinen, E. and Tenkanen, A.: Bird's-Eye Views of the Musical Surface: Methods for Systematic Pitch-Class Set Analysis, *Music Analysis*, 26/i-ii, pp.159–214 (2007).
- [3] 長谷川隆, 西本卓也, 小野順貴, 嵯峨山茂樹: 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, 情報処理学会研究報告, MUS-79, pp.47–52 (2009).
- [4] 長谷川隆: 情報処理向け音楽特徴量抽出による作曲家の年代・地域性の検討, お茶の水音楽論集, Vol.13, pp.36–46 (2011).
- [5] 長谷川隆, 西本卓也, 小野順貴, 嵯峨山茂樹: 楽譜情報からの作曲家らしさ認識のための音楽特徴量の提案, 情報処理学会論文誌, Vol.53, No.3, pp.1204–1215 (2012).
- [6] ヤン・ラルー, 大宮真琴: スタイル・アナリシス, 音楽之友社 (1988).
- [7] 亀岡弘和, 齊藤翔一郎, 西本卓也, 嵯峨山茂樹: Specmurt における準最適共通調波構造パターンの反復推定による多声音楽信号の可視化と MIDI 変換, 情報処理学会研究報告, 2004-MUS-56, pp.41–48 (2004).
- [8] Ono, N., Miyamoto, K., Kameoka, H. and Sagayama, S.: A Real-time Equalizer of Harmonic and Percussive Components in Music Signals, *ISMIR*, pp.139–144 (2008).
- [9] Forte, A.: *The Structure of Atonal Music*, Yale University Press (1973).
- [10] 林 周二: 統計学講義, 丸善株式会社 (1981).
- [11] Jain, A. and Zongker, D.: Feature Selection: Evaluation, Application, and Small Sample Performance, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.19, No.2, pp.153–158 (1997).
- [12] Pudil, P., Novovicova, J. and Kittler, J.: Floating Search Methods in Feature Selection, *Pattern Recognition Letters*, Vol.15, No.11, pp.279–283 (1994).
- [13] Motte, Diether de la: *Harmonielehre*, Bärenreiter – Verlag, Kassel (1976). 吉田雅夫 (監修), 滝井敬子 (訳): 大作曲家の和声, シンフォニア (1980).
- [14] Burkholder, J.P., Grout, D.J. and Palisca, C.V.: *A History of Western Music*, W.W. Norton & Company Inc. (2010).
- [15] Bukofzer, M.F.: *Music in the Baroque Era*, W.W. Norton & Company Inc. (1947).
- [16] Rosen, C.: *The Classical Style*, Faber and Faber (1997).
- [17] 岡部博司: 新訂 標準音楽辞典, 音楽之友社 (1991).



長谷川 隆 (学生会員)

1988 年東京大学大学院工学系研究科情報工学専攻修士課程修了。1988 年 (株) 日立製作所中央研究所入社。2006 年東京大学大学院情報理工学系研究科システム情報学専攻博士課程入学。音楽情報処理, 感性情報処理, コンテンツ配信, レコメン, 保全計画最適化の研究に従事。人工知能学会, 感性工学会, 日本音楽知覚認知学会各会員。