

譜面情報による楽曲の特徴付けの検討

鈴木 真人, 杉山 雅英

会津大学 コンピュータ理工学部
〒965-8580 会津若松市一箕町

あらまし 本報告では、譜面情報に含まれる楽曲の特徴量、特に音高(旋律情報)に着目した特徴量、特徴ベクトルの作成方法、ベクトル間の距離について述べ、それらの有効性を演奏楽器(フルート/ヴァイオリン)分類問題における分類性能を指標として評価した。その結果、音高の二階差分を特徴量とし、差分の値の出現確率をベクトルとし、ベクトル間の距離としてKL情報量を用いた時、演奏楽器の分類実験は、87.5%の高い性能を持つことを示した。ここで音高の二階差分は音高の凸性に対応しており、その有効性が示された。また主音による音高正規化法を述べ、その有効性を評価し、同程度の分類性能を持つことを示した。

キーワード: 音楽特徴量, 特徴抽出

Study on Characterization of Music based on Music Score Information

M. Suzuki, M. Sugiyama

The University of Aizu
Aizu-Wakamatsu, Fukushima 965-8580

This paper describes the extraction methods of feature in music score, especially height of notes (pitch, melody information) and definition of feature vectors and distances between two feature vectors. The classification experiments were carried out in order to evaluate the effectiveness of proposed features, where classification means to classify music score by instrument specification; flute and violin. The experimental results show that proposed feature extraction using output probability vector of 2nd difference for pitch and KL information as distance achieves 87.5% flute/violin classification accuracy. Here, 2nd difference for pitch corresponds to convex property of pitch sequence. a pitch normalization by tonic is proposed and is shown that this normalization technique has the same classification capability.

Key words: Music feature, Feature extraction

1 まえがき

我々は、例えばクラシック音楽を聴くときでも、パッサリの作曲であるとか、何時代の楽曲であるかなどを聞き分けることができる。また、演奏される楽器によっても曲の印象が変わったり、その楽器特有の編曲法を感じることができる。本報告では、楽曲の特徴付けを目的として効果的な楽曲特徴量について検討する。楽曲を扱う場合、音響信号を処理して特徴付けする方法もあるが、本報告では楽譜に表現される楽曲の特徴を扱うことにする。楽譜には楽譜を構成する音符などで音高、音長、相対的な音の強さなどを表現するが、本報告では音高に限定して、音高の出現頻度や推移から特徴量の抽出を行ない、楽曲の特徴量の抽出を検討する。

自然言語をコンピュータ処理するために、大別して、二通りの研究手法がある。一つは文法・規則に基づく手法であり、もう一つは確率統計に基づく手法である [1]。楽曲が音楽作成文法に基づいて生成されているものとするれば、楽曲（譜面情報）を扱う場合も、同様に二通りの研究手法を取ることができる。楽曲のメロディ、リズム、ハーモニーが音楽の 3 大要素として知られている。これらの要素を組み合わせて楽曲を作成する場合、自然言語と同様に文法に従って作曲することになると考えられる。この音楽文法を明示的に取り扱うことは自然言語処理の前者の方法に対応し、音符を確率統計的に処理し特徴付けし取り扱う方法が後者の方法に対応する。現在の自然言語処理における後者の手法を強力さとその成功を考慮して、本報告では後者の手法を用いることにする。確率統計的に大量な楽曲データを処理することにより有効な特徴量を検討することから、楽曲における有効な特徴量を探し出すことを目的としている。

自然言語においては言語を構成する記号 (alphabet) があり、文法はその記号列の生成に関する規則を与えている。譜面情報を自然言語と同様に取り扱うためには、まず譜面を構成するための記号体系を規定し、さらにその記号列に存在する文法を推定することになる。自然言語処理における文法の推定は、記号の出現確率および記号の出現に関する条件付確率 (N -gram) が用いられる。従って、本報告では譜面を構成するための記号体系、その上での文法の推定、また推定文法の有効性の評価を行なう。

楽曲特徴付けに関する従来の研究 [2]、[3]、[4] などにおいても、主として音高による特徴抽出（出現確率、エントロピー、一階差分など）が行なわれてきたが、本報告では特に音高の二階差分に着目し、二つの演奏楽器（フルート、ヴァイオリン）の自動判別可能性を通じてその有効性を検討する。

以下の節は次のように構成されている。第 2 節では楽曲特徴ベクトルの抽出方法とその特徴量の違いを計量す

るための距離について述べる。第 3 節では特徴量の有効性を評価するための実験手法、資料について述べる。第 4 節では実験結果とその考察について述べる。

2 楽曲特徴ベクトルとベクトル間距離

2.1 実験に使用する楽曲データ形式

楽曲の表現形式として MIDI を用いることとした。MIDI data file は楽曲データの表現形式として広く世界に普及しており、インターネット上で様々なジャンルの楽曲データを容易に入手できる。例えば <http://www.ses.ele.tue.nl/midi/> には、多数の MIDI 関連ホームページへのリンクがある。

2.2 MIDI から plain text への変換

MIDI data 形式から plain text に変換して楽曲情報を数値化することが可能である。MIDI data 形式から plain text への変換には、音楽記述言語 PMML [5] の変換ツールを利用した¹。但し、楽曲の特徴付けに直接関係しない MIDI イベントを出力せずに音符情報だけを取り出した。

このコードを `note.pml` とすると、以下のコマンドで、発音時刻、高さ、強さ、継続時間をフィールドとするレコードが出力される。

```
PMML を用いた変換
% pmm1 -i note.pml -l <MIDI ファイル名>
```

3.1 節で述べる BWV1013 の第一楽章を plain text に変換した結果の一部を表 1 に示す。

表 1: PMML を用いた変換の例

発音時刻	音高	強さ	継続時間
120	76	100	120
240	81	100	120
360	80	100	120
480	81	100	120
600	84	100	120
720	81	100	120
840	76	100	120
960	69	100	116
1080	76	100	116
1200	81	100	116

ここで、発音時刻は一小節を 1920、音高は 60 を中央の C4 (261.63 Hz) とし、半音を 1 とする。強さは 0 (無音) から 125、継続時間は一小節を 1920 としている。ここで $1920 = 2^6 \times 3 \times 10$ で全音符、2分音符、4分音符、8

¹ 詳細は <http://cglsun4.u-aizu.ac.jp/pmml/> を参照。

分音符、16分音符および3連符などを表現することが可能である。

2.3 楽曲の特徴量

楽曲特徴量として以下が考えられる。

1. 音高
2. 音長
3. 音の強さ
4. コード進行・和声

本報告では、メロディ、リズム、ハーモニーの3要素の中でも特に音高の推移(メロディ)に着目する。音高の推移は最も「作曲家らしさ」や「演奏楽器らしさ」に関係する要素であると考えられる。

音高を用いた特徴量として、次のようなものが考えられる。

- 絶対的な音高
- 正規化された音高
- 音高の一階差分/二階差分

ある曲が移調されても同一の曲に聞こえるのは、音の間の相互関係が保存されるからである。従って、音の絶対的な高さが固有の特徴を表すのではなく、連続する音高の間の相対関係を考慮することが必要である。即ち、特徴量は「移調不変」でなければならない。 s を移調のためのシフト値とする時、連続する r 個の絶対的音高 x_1, x_2, \dots, x_r に対する特徴抽出関数 $f(x_1, x_2, \dots, x_r)$ は以下の性質を持つことになる。

$$f(x_1 + s, x_2 + s, \dots, x_r + s) = f(x_1, x_2, \dots, x_r) \quad (1)$$

簡単化して $f(x_1, x_2, \dots, x_r)$ を x_1, x_2, \dots, x_r の1次形式

$$f(x_1, x_2, \dots, x_r) = \sum_{i=1}^r w_i x_i,$$

とすると、式(1)を満たすために以下の関係式が導かれる。

$$\sum_{i=1}^r w_i = 0. \quad (2)$$

$r = 2$ の場合、 $w_1 = 1, w_2 = -1$ 、 $r = 3$ の場合 $w_1 = 1, w_2 = -2, w_3 = 1$ などがその例であり、これらはそれぞれ2, 3階の差分に対応している。そこで本報告では、楽曲中の連続する音高の差分及び正規化された音高を扱うこととする。

「移調不変」な特徴量を扱うことの利点として、楽曲から推定された統計値の累積の合成(和)を算出すること

ができることである。楽曲 S_1, S_2 から抽出された「移調不変」な特徴量 $f(x)$ の値の累積を F_1, F_2 とする時、接続した新たな楽曲 $S_0 = S_1 + S_2$ の特徴量の累積 F_0 に関して以下の式が成り立つ。

$$F_0 = F_1 + F_2. \quad (3)$$

これは楽曲の接続だけでなく、楽曲を部分分割して得られた特徴量の値の累積と、全体を一括して計算した累積と一致することになる²。

例えば音高の一階差分を特徴量とする時、一階差分の値の異なり数(これを N とする)が、いわゆる、音高という言葉で構成する記号集合(alphabet)の大きさとなる。二階差分の値の異なり数は $2 \times N$ となり、一般に r 階差分の値の異なり数は $r \times N$ となるので、記号集合が大きくなることになる。一般的に記号数が多くなればその記号に関する統計量の推定に必要な母集合は拡大する。

音高の差分を扱うため、実験材料となる楽曲は単旋律のもの、もしくは、3.2で述べる単音化ルールを適用して得られた単旋律を使用する。

2.4 特徴ベクトルの抽出

音高の一階差分と二階差分は、音高の時系列 X に対して次式で計算される。

$$X = (x_1, x_2, x_3, \dots, x_n, x_{n+1}, x_{n+2}, \dots),$$

- 一階差分

$$\Delta_n = x_{n+1} - x_n,$$

- 二階差分

$$\Delta_n^2 = \Delta_{n+1} - \Delta_n = x_n + x_{n+2} - 2x_{n+1}.$$

$\Delta_n = (0, \pm 1, \pm 2, \dots)$ の値をとる。各々の記号(数値)の楽曲中での出現回数を求めて、出現確率を推定して特徴ベクトルとする。二階差分に対する出現確率を図1に示す。

2.5 特徴ベクトル間の距離

各楽曲毎に求めた出現確率の特徴ベクトルを p, q とする。

$$p = (p_1, p_2, p_3, \dots, p_n)$$

$$q = (q_1, q_2, q_3, \dots, q_n).$$

ここで、 p_i, q_i は記号 i の出現確率であり、 $p_i \geq 0, \sum p_i = 1$ である。 p, q 間の差異を計量化するために特徴ベクトル間の距離を算出する。本報告では Euclidean distance (ユークリッド距離) と Kullback-Liebler distance (KL 情報量) とを比較する。以下で各々の距離について述べる。

²厳密には接続部分で差分値が異なるので誤差が生じる。

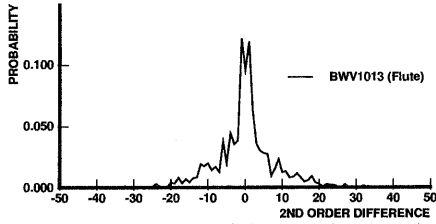


図 1: BWV1013 の二階差分に対する出現確率

2.5.1 Euclidean distance

Euclidean Distance を用いたベクトル間の距離 D_E は以下の式 (4) で定義される。

$$D_E(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (4)$$

Euclidean Distance $D_E(\mathbf{p}, \mathbf{q})$ は距離の公理である以下の 3 つの性質を持つ。

$$\begin{cases} D_E(\mathbf{p}, \mathbf{q}) \geq 0 \\ D_E(\mathbf{p}, \mathbf{q}) = D_E(\mathbf{q}, \mathbf{p}) \\ D_E(\mathbf{p}, \mathbf{q}) \leq D_E(\mathbf{p}, \mathbf{r}) + D_E(\mathbf{r}, \mathbf{q}) \end{cases} \quad (5)$$

2.5.2 KL (Kullback-Liebler) distance

KL 情報量 を用いたベクトル間の距離 (D_{KL1}) は以下の式 (6) で定義される。

$$D_{KL1}(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n p_i \log \frac{p_i}{q_i}} \quad (q_i \neq 0) \quad (6)$$

ここで、KL 情報量は対称性を持たない ($D_{KL1}(\mathbf{p}, \mathbf{q}) \neq D_{KL1}(\mathbf{q}, \mathbf{p})$) ので、 D_{KL1} を用いて $KL2$ を以下の式 (7) で定義する。

$$D_{KL2}(\mathbf{p}, \mathbf{q}) = D_{KL1}(\mathbf{q}, \mathbf{p}) \quad (7)$$

一般に \mathbf{q} のいくつか次元の成分には 0 を含んでいる可能性があるため付録 (A) で述べるフロアリング処理を施す。

3 実験資料と評価実験

3.1 実験資料

評価実験には表 2 に示す J.S.Bach 作曲による五つの組曲を用いた。以降では各々を BWV 番号で呼ぶこととする。BWV1013 はフルートで演奏される楽曲であり、BWV1001 から BWV1004 はバイオリンで演奏される楽曲である。これらに対してそれぞれ 3.2 節のルールに基

づき、単音化処理を行なう。表から組曲ごとに楽章の数も音符の数も異なっていることが分かる。図 2,3,4, 5 に BWV1001, BWV1002, BWV1003, BWV1004 から得られた二階差分の出現確率 (ヒストグラム) を示す。

表 2: 実験資料

楽曲名	楽章	音符数
無伴奏フルートソナタ パルティータ イ短調 BWV1013	1 - 4	2743
無伴奏ヴァイオリン ソナタ第 1 番 ト短調 BWV1001	1 - 4	7202
無伴奏ヴァイオリン パルティータ第 1 番 ロ短調 BWV1002	1 - 8	7634
無伴奏ヴァイオリン ソナタ第 2 番 イ短調 BWV1003	1 - 4	3617
無伴奏ヴァイオリン パルティータ第 2 番 ニ短調 BWV1004	1 - 5	4713

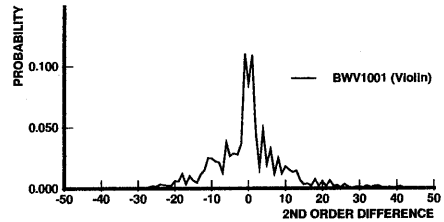


図 2: BWV1001 の二階差分に対する出現確率

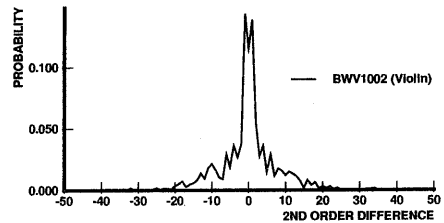


図 3: BWV1002 の二階差分に対する出現確率

3.2 単音化のルール

本報告では単旋律の楽曲のみを扱うこととしたが、ヴァイオリンで演奏される楽曲の場合、ソロ演奏であっても複数の音が同時に発音することがある。そこで複数の音

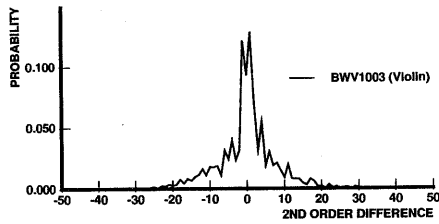


図 4: BWV1003 の二階差分に対する出現確率

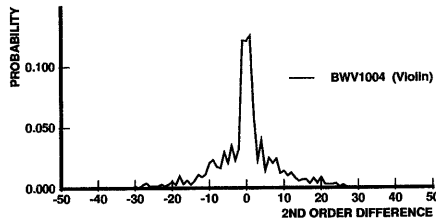


図 5: BWV1004 の二階差分に対する出現確率

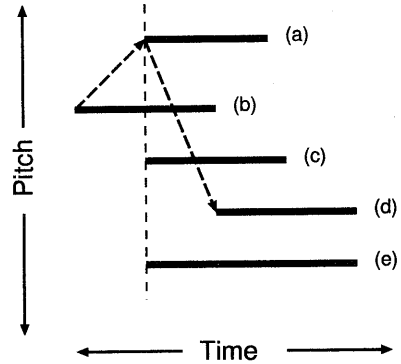


図 6: 単音化処理の概念

- R が1に近い ⇔ 正の相関が強い
- R が-1に近い ⇔ 負の相関が強い
- R が0に近い ⇔ 無相関

が同時発音する場合は、その中の最も高い音を取り、単音化処理を行なう。

図 6 に処理の概念を示す。まず (b) が発音され、次に (a), (c), (e) が同時に発音されている。この時、(a), (c), (e) の中で最も高音である (a) を取る。最後に (d) が発音される。従ってこの場合、音高は (b), (a), (d) の順で発音されたとする。音高の継続長 (終止点) については考慮しない。

3.3 評価方法

楽曲毎に推定された特徴ベクトルを以下に述べる二つの方法で評価する。

3.3.1 相関係数による評価

各特徴量に対する各距離の相関を調べるため、相関係数 (R) を求める。相関係数は以下のように定義される。 x_n, y_n には、各々の距離の計算結果が入る。

$$R(x, y) = \frac{\sum \hat{x}_n \hat{y}_n}{\sqrt{\sum \hat{x}_n^2} \sqrt{\sum \hat{y}_n^2}} \quad (8)$$

但し、以下の式で \hat{x}_n, \hat{y}_n は定義される。

$$\hat{x}_n = x_n - \frac{1}{N} \sum x_n, \quad \hat{y}_n = y_n - \frac{1}{N} \sum y_n.$$

この時、 R は $-1 \leq R \leq 1$ であり、以下の性質がある。

3.3.2 カテゴリ分類による評価

図 7 に示すように、距離の計算結果からカテゴリ分類におけるカテゴリ内の accept 率とカテゴリ間の reject 率との平均が最大となる値に threshold を設定し、カテゴリ内/カテゴリ間への分類率を計算する。ここでカテゴリとは、演奏楽器がフルートと指定される楽曲とヴァイオリンと指定される楽曲の二つである。

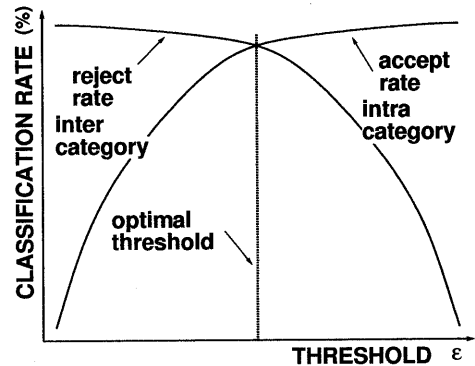


図 7: カテゴリ分類による評価の概念

4 評価実験結果

4.1 距離の計算結果

一階差分、二階差分を特徴量とし、Euclidean distance、KL 情報量を用いたベクトル間の距離を表 3 から表 6 に示す。二階差分に対する KL 情報量を用いた計算結果 (表 6) が、他に比べて極端に大きな値であることがわかる。

表 3: ベクトル間の距離 (1st-diff, Euclid)

距離: Euclidean distance、特徴量: 一階差分

	1013	1001	1002	1003	1004
1013	—	0.0701	0.0625	0.0629	0.0836
1001		—	0.0804	0.0889	0.0704
1002			—	0.0510	0.0760
1003				—	0.0657
1004					—

表 4: ベクトル間の距離 (1st-diff, KL)

距離: KL 情報量、特徴量: 一階差分

	1013	1001	1002	1003	1004
1013	—	0.3086	0.2352	0.2356	0.3047
1001	0.4130	—	0.3130	0.3069	0.2444
1002	0.2874	0.2911	—	0.2117	0.2539
1003	0.2620	0.2785	0.2039	—	0.2340
1004	0.3924	0.2399	0.2917	0.2650	—

表 5: ベクトル間の距離 (2nd-diff, Euclid)

距離: Euclidean distance、特徴量: 二階差分

	1013	1001	1002	1003	1004
1013	—	0.0487	0.0473	0.0393	0.0467
1001		—	0.0629	0.0467	0.0521
1002			—	0.0474	0.0371
1003				—	0.0413
1004					—

表 6: ベクトル間の距離 (2nd-diff, KL)

距離: KL 情報量、特徴量: 二階差分

	1013	1001	002	1003	1004
1013	—	0.2903	0.2292	0.2293	0.2672
1001	0.3980	—	0.3040	0.2920	0.2818
1002	0.2573	0.2685	—	0.2123	0.2219
1003	0.2845	0.2630	0.2314	—	0.2365
1004	0.3607	0.2718	0.2551	0.2763	—

4.2 相関係数による評価

4.1 節の結果から、一階差分、二階差分に対する Euclidean distance と KL 情報量を用いた距離値の間の相関係数を表 7 に示す。

表 7: 相関係数の値の比較

計算方法 (差分)		相関係数
a.	Euclid (1) Euclid (2)	0.74
b.	Euclid (1) KL1 (1)	0.84
c.	Euclid (1) KL2 (1)	0.48
d.	KL1 (1) KL2 (1)	0.72
e.	Euclid (2) KL1 (2)	0.71
f.	Euclid (2) KL2 (2)	0.39
g.	KL1 (2) KL2 (2)	0.29
h.	KL1 (1) KL1 (2)	0.90

この結果から、以下のようなことがわかる。

- 同じ計算方法での一階差分と二階差分は相関が強い。(a, h)
- Euclid と KL1 は相関が強い。(b, e)
- Euclid と KL2 は相関が弱い。(c, f)
- KL1 と KL2 は、一階差分では相関が強いが二階差分では相関が弱い。(d, g)

4.3 カテゴリ分類による評価

4.1 節の結果を元に、各計算方法での分類率を表 8 に示す。Euclidean distance は一階差分、二階差分ともに分類率が低く、二階差分の KL2 による計算が比較的高い分類率を示している。これは確率ベクトル間の距離計算には KL 情報量が適していることを示している。ここで得られた結果は表 7 に示した相関係数による評価と対応している。

音高の一階差分は、旋律の増減情報を表している。それに対し、二階差分は旋律の凸性情報を表している。表 8 の結果は、KL2 において、旋律の増減情報よりも旋律の凸性情報の方が特徴量として有効であることを示している。実験資料数が 10 と少ないので、今後、資料数を増やしこの結果の成立を確認する必要がある。

4.4 楽曲分割の効果

フルート曲である BWV1013 の楽章の組合せによってサンプルを増やすため、音符数を変化させた時の出現確率ベクトルの収束の様子を調べる。BWV1013 は 4 つの楽章から構成されてるので、この 4 つの楽章の組合せ ($14 = {}_4C_1 + {}_4C_2 + {}_4C_3$ 通り) に対応するそれぞれの特徴ベクトル (二階差分) を求め、それらと 4 楽章全てを

表 8: 分類率の比較

計算方法	threshold	分類率 (%)	
一階差分	Eucild	0.05107	58.3
	KL1	0.23406	66.7
	KL2	0.29173	75.0
二階差分	Eucild	0.03716	58.3
	KL1	0.22193	66.7
	KL2	0.27640	87.5

用いて得られる出現確率ベクトルとの距離 (KL 情報量) を図 8 に示す。この結果から、およそ 1500 (3 楽章分) の音符数があれば表 8 に示した最適 threshold の値よりも小さくなる事が分かる。

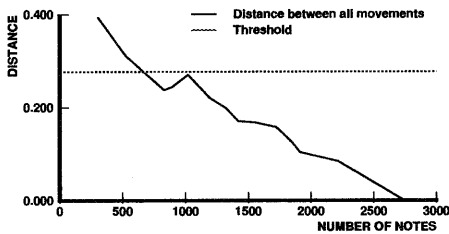


図 8: BWV1013 の音符数増加に伴う収束の様子

図 8 の結果に基づき、フルートの楽曲の 4 楽章の中の 3 楽章の一つの楽曲として四種の特徴ベクトルを算出し、カテゴリ分類を行なう。ここで特徴量は、二階差分であり、距離は KL 情報量を用いた。分類結果を表 9 に示す。

KL2 において、フルートから見たヴァイオリンとの判別は threshold を 0.2854 に設定することで分類率 81.3% を示した。また、ヴァイオリンから見たフルートとの判別は threshold を 0.2764 に設定することで分類率 93.6% を示している。

4.5 主音による正規化法とその評価

差分による「移調不変」特徴量の抽出以外に、各楽曲の主音をもとに音高を正規化する方法について述べる。主音による正規化は、差分を取る手法と同様、音高の相互関係を保存することができるのに加え、音高情報を残すことができる。

楽曲の調を以下の手順で推定する。

1. 12 音階のヒストグラムを作成
2. 長調・短調の template と比較
3. 最も適合する調を確定

次に、音高の平均から最も近い主音 (x_T) を求め、その主音を用いて正規化する ($x_n - x_T$)。

楽章をひとまとめにして一括して正規化を行い、出現確率を求め算出した距離を表 10 及び表 11 に示す。その分類率を表 12 に示す。主音による正規化法に対する分類性能は距離への依存性が差分による正規化法よりも小さく、高い性能が得られることが分かる。

表 10: 主音正規化法を用いたベクトル間の距離
距離: Euclidean distance

	1013	1001	1002	1003	1004
1013	—	0.2341	0.1976	0.2103	0.0924
1001		—	0.0681	0.0656	0.1702
1002			—	0.0377	0.1278
1003				—	0.1418
1004					—

表 11: 主音正規化法を用いたベクトル間の距離
距離: KL 情報量

	1013	1001	1002	1003	1004
1013	—	1.2712	0.8715	1.0308	0.3732
1001	1.4614	—	0.4041	0.3548	0.8915
1002	1.0896	0.4588	—	0.2178	0.5754
1003	1.1650	0.3807	0.1649	—	0.6420
1004	0.5053	1.0482	0.7037	0.8703	—

表 12: 主音正規化法を用いた分類の評価

計算方法	threshold	分類率 (%)	
主音による 正規化	Euclid	0.17021	87.5
	KL1	0.64205	79.2
	KL2	0.87035	79.2

5 むすび

「移調不変」な特徴量として音高に対する差分および主音による正規化法を提案し、演奏楽器の判別による評価実験を通してそれらの有効性を示した。旋律の一階差分情報より二階差分情報を特徴量とする方が有効であること、特徴ベクトル間の距離の計算には、Euclid distance より KL 情報量の方が有効であり、さらに KL1 より KL2 が有効であることを示した。また、二階差分を特徴量とし、KL2 で距離計算をしたときの、フルート、ヴァイオリンのカテゴリ分類率は、87.5% を得た。本報告では単旋律の楽曲のみ扱っているため、数多く作成されている MIDI データではあるが、実験資料に使用できるものは少なかった。そのため、同一作曲家による楽曲の中での演奏楽器の判別問題を検討した。特徴量として二階差分を用いることにより、高い性能で演奏楽器の判別ができた

表 9: 分割したベクトルとの距離

	f123	f234	f134	f124	1001	1002	1003	1004
f123	—	0.2321	0.1976	0.1000	0.3266	0.2483	0.2455	0.2949
f234	0.2215	—	0.2871	0.1799	0.2906	0.2182	0.2475	0.2425
f134	0.1853	0.2851	—	0.1262	0.3153	0.2887	0.2641	0.3178
f124	0.1261	0.2121	0.1665	—	0.3000	0.2490	0.2532	0.2852
1001	0.4387	0.4035	0.4388	0.3998	—	0.3040	0.2920	0.2818
1002	0.2843	0.2544	0.3324	0.2694	0.2685	—	0.2123	0.2219
1003	0.3128	0.2853	0.3317	0.3008	0.2630	0.2314	—	0.2365
1004	0.4109	0.3360	0.4253	0.3715	0.2718	0.2551	0.2763	—

ことは、作曲者の分類や時代同定等への応用の可能であると考えられる。

本報告では音高の推移情報(メロディ)にのみ着目して特徴抽出を行なった。今後の課題として、音長や音の強弱(リズム)、コード進行や和声(ハーモニー)なども考慮に入れて特徴抽出法を検討する。例えばリズムを考慮する場合、ある一定の音長でメロディを正規化する手法などが考えられる。特徴ベクトルの次元を増やすことは、より多くのサンプルデータを必要とする。また判別対象として、クラシック音楽ではなく、単旋律化しやすく資料数が多い歌謡曲(ポップス)を用いることや、波形から得られる楽曲情報[6]に基づく特徴付けや分類を行うことも検討したい。

謝辞

研究を進める上で pmml の御教示いただいた、会津大学 西村憲博士、また有益な助言をいただいた会津大学滞在研究員 W.Martens 博士に感謝いたします。

参考文献

- [1] 北研二, 中村哲, 永田昌明, "音声言語処理 - コーパスに基づくアプローチ", 森北出版株式会社, p.1 (1996).
- [2] 松田稔, 秋山好一, 森和義, "日本の楽曲の基本的特徴 - 音高について -", 音学誌, Vol.50, No.11, pp.897-905 (1994).
- [3] 松田稔, 秋山好一, 森和義, "楽曲の特徴量 - 平均推移エントロピー -", 音学講論, 2-7-4 (1995-9).
- [4] 松田稔, 秋山好一, 森和義, "線形予測による旋律パターンの判別分析", 音学講論, 2-5-17 (1996-9).
- [5] 西村憲, "音楽記述言語 PMML の概要", 情処学会研究報告, 97-MUS-21-10 (1997-7).

- [6] 木下, 田中, 音楽情景分析システム OPTIMA, コンピュータソフトウェア, Vol.15, No.1, pp.59-66 (1998-01).

A KL 情報量計算におけるフロアリングの手順

$q = (q_1, q_2, q_3, \dots, q_n)$ における $q_i = 0$ となる q_i を $q_i^* = e^{-\alpha}$ と置き換え、 $q_i = 0$ の個数を M とすると、

$$q_1 + q_2 + \dots + q_n = 1 + M \times e^{-\alpha}.$$

ここで増加した分を $q_i \neq 0$ に振り分ける。

$$\Delta = \frac{M e^{-\alpha}}{n - M},$$

とおくと、

$$q_i^* = q_i - \Delta.$$

$q_i \neq 0$ の時の最小の q_i を q_{\min} とすると、

$$q_{\min} - \Delta > 0,$$

とならなければならない。つまり、

$$\begin{aligned} q_{\min} &> \Delta = \frac{M e^{-\alpha}}{n - M}, \\ \frac{n - M}{M} \times q_{\min} &> e^{-\alpha}, \\ \log \left(\frac{n - M}{M} \times q_{\min} \right) &> -\alpha, \\ \alpha &> -\log \left(\frac{n - M}{M} \times q_{\min} \right). \end{aligned}$$

よって $\alpha = 1 - \log \left(\frac{n - M}{M} \times q_{\min} \right)$ とおけば、十分である。

以上の操作により、KL 情報量を計算する時は次のように q^* を与え、 $D_{KL}(p, q^*)$ として計算する。

$$q_i^* = \begin{cases} e^{-\alpha} & (q_i = 0) \\ q_i - \Delta & (q_i \neq 0) \end{cases}$$