

# 楽曲属性を反映する特徴量の動的選出に関する研究

大高 宏樹<sup>1,a)</sup> 岡田 龍太郎<sup>1,b)</sup> 北川 高嗣<sup>2,c)</sup>

**概要:** 自動作曲システムでは、既存楽曲を分析して得た楽曲の特徴量を用いる手法が主流である。特徴量はシステム毎に独立しているため、本研究では特徴量を統一的に扱うべく、特徴量を数多く定義した。そして、特定の楽曲属性らしさを反映する特徴量を動的に選出するモデルを作成した。このモデルを用い、楽器とジャンルを対象に特徴抽出の実験を行った結果、定義した特徴量の、各楽曲属性らしさに関するランキングが得られた。ランキング上位の特徴量を用いることにより、各楽曲属性を反映させた楽曲の自動生成が可能になる。

## A Study of How to Select Features of The Musical Attribute Dynamically

HIROKI OTAKA<sup>1,a)</sup> RYOTARO OKADA<sup>1,b)</sup> TAKASHI KITAGAWA<sup>2,c)</sup>

**Abstract:** When one write a program of automatic composition system, music features are generally used, which were extracted from some existing music. Music features are different on each system. In this paper, to consolidate features, we define a lot of music features. Then, we make a model which select dynamically music features of the musical attribute. We experiment with this model for features-extraction of an instrument and a genre. We got ranked features of how much music attribute's identity the feature express. By using higher ranked music features, we can compose a music which has features of the musical attribute.

### 1. はじめに

自動作曲システムでは、既存の楽曲を分析し有用な楽曲特徴量を解析しそれらを元に作曲システムを構築するという手法が主流である。この時、あるアルゴリズム作成者は特定のジャンルに着目し、また別のアルゴリズム作成者は特定のアーティストに着目するなど、システム構成に用いられる楽曲特徴量はアルゴリズム作成者によって様々である。そのため、例えばジャンルとアーティストを組み合わせたいという要求があった場合、二つの手法をそのまま取り入れることはできず、また別のシステムを作成する必要

がある。

そこで、これらの楽曲特徴量を統一的に扱うことにより、楽器やジャンルやアーティスト等の楽曲情報を組み合わせ、ユーザの嗜好に合わせた楽曲の生成が可能になる。このモデルを実現するために、使用する特徴量を限定せず、特徴量を数多く定義し、それらを目的に応じて使い分けたいという要求がある。

本研究では、特定の楽曲属性を反映する楽曲特徴量を、計算機上で動的に選出するモデルを提案する。ここで楽曲属性とは、楽器やジャンルやアーティスト等の楽曲情報を指す。特に本稿では、楽曲属性の中で楽器情報及びジャンル情報に着目し、楽器及びジャンル毎の特徴を反映する楽曲特徴量を、計算機を用いて選出する。具体的には、まず、各楽曲属性の特徴を反映すると考えられる音楽的指標（以下、楽曲特徴量）を数十種類定義する。次に、楽曲属性毎に研究対象の MIDI データを用意し、それらから定義した楽曲特徴量を抽出する。そして楽曲属性毎に、楽曲と楽曲属性に関する教師付きクラスタリング及びその性能評価を

<sup>1</sup> 筑波大学大学院 システム情報工学研究科 コンピュータサイエンス専攻

Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba

<sup>2</sup> 筑波大学大学院 システム情報系

Division of Information Engineering, Faculty of Engineering, Information and Systems, University of Tsukuba

a) otaka@mma.cs.tsukuba.ac.jp

b) rytrkd@gmail.com

c) takashi@cs.tsukuba.ac.jp

特徴量毎に行い、特徴反映の有効性に関する楽曲特徴量のランキングを得る。この結果から、自動作曲に適用する場合の有効性について検討する。

## 2. 関連研究

楽曲から特徴を抽出するためには、楽曲の特徴量を定めなければならない。

中山らは、ユーザの好みの音楽を効率よく検索するシステムを構築するために、楽曲の特徴量を定義し、楽曲分類を行う際に有効な特徴量について検討した [1]。この研究では、200 曲の MIDI 楽曲から抽出した特徴量を利用して K-means 法による楽曲のクラスタリングを行い、その結果と被験者が分類した結果を比較して、定義した楽曲特徴量の音楽分類に対する有効性について検討している。

また、辻らは楽曲特徴量として旋律の音高変化に着目した [2]。楽曲の局所的な音高推移を局所パターンとして抽出し、それらの特徴量を感性語による楽曲検索システムに適用し、局所パターンと感性語との相関性について検討した。

本研究では、これらの研究で用いられた楽曲の特徴量を参考にし、楽曲の特徴量を独自に定義し、楽器毎の特徴を抽出することができる特徴量について検討する。

また、音楽心理学者 Hevner は、楽曲の構成要素として調（長調／短調）、旋律（上昇／下降）、音高（高／低）、和声（単純／複雑）、テンポ、リズム（固定／流動）を定義しており [4], [5], [6]、我々は Hevner の研究成果を用いた自動作曲システムを実現している [7]。

## 3. 楽曲の特徴抽出

本研究では、まず楽曲の特徴量を独自に定義する。次に、研究対象として用意する楽曲から各特徴量を抽出する。そして特徴量毎に楽曲を、解析に用いた楽曲属性の種類分のクラスタへクラスタリングする。最後に、クラスタリングの結果に対し、性能を評価する。クラスタリングの性能評価の結果、性能が良ければ、その特徴量においては楽曲属性の種類毎に上手く楽曲を分類できることになる。すなわち、その特徴量は各楽曲属性の特徴を反映していると考えられる。

本章では、楽曲の特徴量およびクラスタリングの性能評価手法について述べる。

### 3.1 楽曲特徴量の定義

本研究では、以下の 22 種類の値を、楽曲の特徴量として定義する。それぞれの特徴量の式を示す。まず、式中における変数及び関数を、以下のように定義する。

- $n$  番目の主旋律の音高:  $Mnote(n)$
- $n$  番目の主旋律の音符の長さ:  $Mndur(n)$
- 主旋律の音符数:  $M$

- $n$  番目の音符の音高:  $Note(n)$
- $n$  番目の音符の長さ:  $Ndur(n)$
- $n$  番目の音符の発音開始時刻:  $Nstart(n)$
- $n$  番目の音符の発音終了時刻:  $Nend(n)$
- 全音符数:  $N$
- $\Delta Mnote(n) = Mnote(n+1) - Mnote(n)$
- $U(k) = \begin{cases} 1 & (k > 0) \\ 0 & (k \leq 0) \end{cases}$
- $D(k) = \begin{cases} 1 & (k < 0) \\ 0 & (k \geq 0) \end{cases}$
- $\delta(k) = \begin{cases} 1 & (k = 0) \\ 0 & (k \neq 0) \end{cases}$
- $N'start(n) = \begin{cases} lag & ((lag = Nstart(n) \bmod 120) < 60) \\ lag - 120 & ((lag = Nstart(n) \bmod 120) \geq 60) \end{cases}$
- $N'end(n) = \begin{cases} lag & ((lag = Nend(n) \bmod 120) < 60) \\ lag - 120 & ((lag = Nend(n) \bmod 120) \geq 60) \end{cases}$

楽曲の主旋律を対象に、以下の値を抽出する。

- 上昇音割合 =  $\frac{\sum_{n=1}^{M-1} U(\Delta Mnote(n))}{M-1}$
- 下降音割合 =  $\frac{\sum_{n=1}^{M-1} D(\Delta Mnote(n))}{M-1}$
- 水平音割合 =  $\frac{\sum_{n=1}^{M-1} \delta(\Delta Mnote(n))}{M-1}$
- 上昇音高平均 =  $\frac{\sum_{n=1}^{M-1} U(\Delta Mnote(n)) \Delta Mnote(n)}{\sum_{n=1}^{M-1} U(\Delta Mnote(n))}$
- 下降音高平均 =  $\frac{\sum_{n=1}^{M-1} D(\Delta Mnote(n)) \Delta Mnote(n)}{\sum_{n=1}^{M-1} D(\Delta Mnote(n))}$
- 上昇音高分散 =  $\frac{\sum_{n=1}^{M-1} (\Delta Mnote(n) - (\text{上昇音高平均}))^2 U(\Delta Mnote(n))}{\sum_{n=1}^{M-1} U(\Delta Mnote(n))}$
- 下降音高分散 =  $\frac{\sum_{n=1}^{M-1} (\Delta Mnote(n) - (\text{下降音高平均}))^2 D(\Delta Mnote(n))}{\sum_{n=1}^{M-1} D(\Delta Mnote(n))}$
- 主旋律音高平均 =  $\frac{\sum_{n=1}^M Mnote(n) Mndur(n)}{\sum_{n=1}^M Mndur(n)}$
- 主旋律音高分散 =  $\frac{\sum_{n=1}^M (Mnote(n) - \text{主旋律音高平均})^2}{M}$
- 16 分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 120)}{M}$
- 8 分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 240)}{M}$
- 4 分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 480)}{M}$
- 付点 4 分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 720)}{M}$
- 2 分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 960)}{M}$

- 付点2分音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 1440)}{M}$
- 全音符割合 =  $\frac{\sum_{n=1}^M \delta(Mndur(n) - 1920)}{M}$

次に、楽曲の全音高を対象に、以下の値を抽出する。

- 全音高平均 =  $\frac{\sum_{n=1}^N Note(n)Ndur(n)}{\sum_{n=1}^N Ndur(n)}$
- 全音高分散 =  $\frac{\sum_{n=1}^N (Note(n) - \text{全音高平均})^2}{N}$
- 全音高ズレ分散 =  $\frac{\sum_{n=1}^N (N'start(n))^2 + (N'end(n))^2}{2N}$
- 調性外楽音割合 =  $\frac{\text{スケールノート以外の楽音の総数}}{N}$

また、楽曲のメタデータから、以下の値を抽出する。

- テンポ = 楽曲のメタデータに設定されているテンポデータ

最後に、楽曲の主旋律を対象に、辻らが提案した局所パターン [2] を抽出する。局所パターンでは、楽曲の旋律上の順番に並んだある3つの楽音に注目し、音高の変化の度合いを17のパターンに分類する(図1)。これらの結果を局所パターンとして全楽音から抽出し、各パターンの平均出現割合を求めたものを、楽曲の特徴量とする。局所パターンの分類では、例えばパターン1の場合は、1番目の音と2番目の音の上昇方向の音高差が、2番目の音と3番目の音の上昇方向の音高差よりも大きい場合を表している。

		$\Delta Mnote(n+1)$						
		+		0	-			
$\Delta Mnote(n)$	+	1 	2 	3 	4 	5 	6 	7 
	0	8 			9 	10 		
	-	11 	12 	13 	14 	15 	16 	17 

図1 局所パターンの分類

### 3.2 クラスタリングの性能評価手法

クラスタリングの結果に対し、Purity 及び Entropy を用いて各クラスタリング結果の性能評価を行う。

#### 3.2.1 Purity

Purity(純度)とは、ある正解のクラスタをどの程度含むかという指標に対し、各クラスタのデータ数による重み付き平均を取った値である。結果は0~1の値を取り、値が大きいほどクラスタリングの結果がよい。Purityの式を以

下に示す。

$$P_i = \frac{1}{|C_i|} \max_h |C_i \cap A_h|,$$

$$\sum_{i=1}^k \frac{|C_i|}{N} P_i = \frac{1}{N} \sum_{i=1}^k \max_h |C_i \cap A_h|.$$

#### 3.2.2 Entropy

Entropyとは、各クラスタのエントロピーに対し、クラスタのデータ数による重み付き平均を取った値である。結果は0~1の値を取り、値が小さいほどクラスタリングの結果がよい。Entropyの式を以下に示す。

$$E_i = - \sum_{h=1}^k P(A_h|C_i) \log P(A_h|C_i),$$

$$\sum_{i=1}^k \frac{|C_i|}{N} E_i = \sum_{i=1}^k \frac{\sum_{j=1}^k x_{ij}}{N} E_i.$$

## 4. 実験

本章では、楽器単体で演奏されている楽曲および特定のジャンル毎に演奏された楽曲から、3.1節で定義した特徴量を抽出する。次に、特徴量毎にK-means法によるクラスタリングを行う。K-means法では、Hartigan-Wongアルゴリズムを用いる。クラスタリングの結果に対し、3.2節で述べた手法を用いて性能評価する。そして、各楽曲属性の特徴を反映していると考えられる特徴量について検討する。

### 4.1 実験環境

#### 4.1.1 研究対象データ

本研究では、音高や音長などの楽曲の音楽的特徴を取りやすいというMIDIデータの利点から、MIDIデータを持つ楽曲を利用する。楽器対象の特徴抽出実験では、MIDIデータを持つAppleLoopsを用いる。ジャンル対象の特徴抽出実験では、Web上に公開されている特定のジャンル毎に演奏されたMIDIデータを収集し、それらを用いる。

AppleLoopsとは、多様な楽器毎に存在する楽曲のパーツとなる短いフレーズ(以下、ループ)であり、ループシーケンサーに用いられる楽曲である。AppleLoopsには、波形データのみループと、MIDIデータを含むループが存在するが、前述したMIDIデータの利点から、MIDIデータを持つAppleLoopsを研究対象とする。本研究ではそれらのループの中で、Bassを107曲、Electric Pianoを35曲、Guitarを13曲、Organを23曲、Pianoを64曲、Synthesizerを94曲の6種類の楽器、計336曲のループを利用し、特徴抽出及びクラスタリングを行う。

ジャンル対象の特徴抽出実験では、アメリカ民謡を35曲、ヨーロッパ民謡を46曲、日本民謡を40曲、沖縄民謡を27曲、ロシア民謡を21曲の5種類のジャンル、計169

曲の楽曲を Web<sup>\*1\*</sup>から収集し、特徴抽出及びクラスタリングを行う。

#### 4.1.2 解析データの選定

本節では、AppleLoops において、特徴量抽出に用いる楽曲データの選定を行う。

##### (1) クォンタイズ処理

クォンタイズとは、入力した MIDI データのタイミングを修正する処理のことである。本研究の音長に関する特徴量において、最短の音長を 16 分音符と定義したため、各楽音の発音時間を 16 分音符の整数倍の値で扱いたいという要求がある。しかし、AppleLoops における MIDI データでは、ユーザが作曲の際に行う表情付けやリアルタイム録音を行う際に、発音のタイミングにズレが生じる場合がある。そのため、それらの特徴抽出を行う前にデータを整理する必要がある。AppleLoops の MIDI は、4 分音符の発音時間を 480 としている。本研究では各楽音に対して、楽音の発音開始時刻と発音終了時刻が 16 分音符の分解能 (120) の整数倍に収まるように、各楽音に対し、前後 60 の範囲で発音開始時刻及び発音終了時刻の切り上げ・切り下げ処理を行う。次に、特徴量定義に用いた発音時間に各楽音の発音時間を収めるため、特徴量における発音時間の、長さが隣り合う発音時間の中間となるタイミングを元に、各楽音の発音開始時刻及び発音終了時刻の切り上げ・切り下げ処理を行う。

##### (2) 主旋律推定

本研究で定義した特徴量には楽曲の主旋律を解析対象としているものがある。そのため、Piano 等の伴奏が含まれるループに関しては、特徴抽出を行う前に主旋律を推定する必要がある。本研究では、同時発音している楽音の中で音量が最も大きい楽音（音量最大の楽音が複数ある場合はその中で音高が最も高い楽音）を主旋律と定義し、特徴抽出前に主旋律推定を行う。この主旋律推定法は、文献 [3] を参考にしている。主旋律推定の結果、主旋律を構成する楽音が 2 つ以下になる楽曲が 5 曲存在し、それらのデータは局所パターンの特徴量を抽出する際に解析データから除いている。

## 4.2 実験

### 4.2.1 クラスタリング結果

本節では、楽曲属性毎に次の操作を行う。

\*1 「MIDI ダウンロード (童謡 唱歌 叙情歌)」  
<<http://www.5b.biglobe.ne.jp/~pst/douyou-syouka/12dl/dl.htm>>, 2012 年 9 月 7 日。

\*2 「音楽研究所」  
<[http://www.asahi-net.or.jp/~hb9t-kttd/music/Japan/Studio/Midi/Traditional/ryukyuu\\_index.html](http://www.asahi-net.or.jp/~hb9t-kttd/music/Japan/Studio/Midi/Traditional/ryukyuu_index.html)>, 2012 年 9 月 7 日。

まず、3.1 章で定義した特徴量を、用意した全楽曲から抽出する。ただし、全音高ズレ分散の特徴量に関しては、打ち込みによる MIDI 作成時には発音タイミングのズレが生じないことから、AppleLoops のみ特徴抽出を行う。

次に特徴量毎に、全楽曲を対象に K-means 法を用いて、楽器は 6 クラスタへ、ジャンルは 5 クラスタへクラスタリングを行う。クラスタリング結果の例として、次節で示すクラスタリングの性能評価において、楽器対象時の概ね上位に位置する全音高平均、概ね下位に位置する 2 分音符割合、及びジャンル対象時の概ね上位に位置する局所パターン、概ね下位に位置する主旋律音高平均の正誤表を、表 1-4 に示す。以降の表で、C1~C6 はクラスタであり、B: Bass, E: Electric Piano, G: Guitar, O: Organ, P: Piano, S: Synthesizer, US: アメリカ民謡, EU: ヨーロッパ民謡, JA: 日本民謡, OW: 沖縄民謡, RU: ロシア民謡である。

表 1 全音高平均 (楽器)

	C1	C2	C3	C4	C5	C6
B	0	10	0	43	41	13
E	7	8	14	3	3	0
G	0	10	3	0	0	0
O	16	0	7	0	0	0
P	5	22	20	1	16	0
S	9	6	56	3	17	3

表 2 2 分音符割合 (楽器)

	C1	C2	C3	C4	C5	C6
B	0	1	96	6	4	0
E	2	2	23	3	4	1
G	2	1	8	1	0	1
O	2	0	18	0	3	0
P	9	1	41	5	8	0
S	1	0	90	1	0	2

表 3 局所パターン (ジャンル)

	C1	C2	C3	C4	C5
US	0	12	4	14	5
EU	3	13	5	10	15
JA	1	0	17	22	0
OW	15	2	1	9	0
RU	5	0	3	2	11

表 4 主旋律音高平均 (ジャンル)

	C1	C2	C3	C4	C5
US	7	8	3	5	12
EU	2	13	4	12	15
JA	4	6	6	10	14
OW	3	8	2	7	7
RU	2	2	3	3	11

### 4.2.2 クラスタリングの性能評価

4.2.1 節での結果に対し、どれくらい正しく分類できているかを計るため、Purity 及び Entropy を用いて各クラスタリング結果の性能評価を行う。各特徴量のクラスタリング結果に対し、Purity による性能評価を行った結果を表 5, 7 に、Entropy による性能評価を行った結果を表 6, 8 に、楽曲属性毎に示す。

## 4.3 考察

クラスタリングの各性能評価の結果を見てみると、楽器に関しては、全音高平均、主旋律音高平均、全音高分散、上昇音高平均などの特徴量が上位に位置している。よって、これらの特徴量が、楽器毎の特徴を反映していると考えられる。また、ジャンルに関しては、局所パターン、全音高分散、全音高平均、テンポなどの特徴量が上位に位置しており、これらの特徴量がジャンル毎の特徴を反映していると考えられる。一方、音長に関する特徴量は、楽器、ジャンルのいずれにおいても概ね下位に位置しており、各楽曲

表 5 Purity による性能評価  
 (楽器)

特徴量	Purity の値
全音高平均	0.5684524
主旋律音高平均	0.5625
下降音高平均	0.4761905
全音高分散	0.4761905
上昇音高平均	0.4732143
主旋律音高分散	0.4613095
下降音高分散	0.452381
局所パターン	0.4471299
上昇音高分散	0.4404762
下降音割合	0.4375
水平音割合	0.4255952
16 分音符割合	0.4255952
テンポ	0.4166667
8 分音符割合	0.3988095
調性外楽音割合	0.3982301
4 分音符割合	0.3869048
上昇音割合	0.3809524
全音高ズレ分散	0.375
付点 4 分音符割合	0.3720238
2 分音符割合	0.3660714
全音符割合	0.3571429
付点 2 分音符割合	0.3511905

表 7 Purity による性能評価  
 (ジャンル)

特徴量	Purity の値
局所パターン	0.4852071
全音高分散	0.4792899
全音高平均	0.4260355
テンポ	0.408284
調性外楽音割合	0.4023669
下降音割合	0.3905325
下降音高平均	0.3786982
主旋律音高分散	0.3609467
上昇音高平均	0.3550296
付点 2 分音符割合	0.3550296
上昇音割合	0.3431953
上昇音高分散	0.3431953
全音符割合	0.3431953
水平音割合	0.3372781
下降音高分散	0.3313609
付点 4 分音符割合	0.3313609
8 分音符割合	0.3313609
4 分音符割合	0.3195266
主旋律音高平均	0.3136095
16 分音符割合	0.3017751
2 分音符割合	0.2840237

表 6 Entropy による性能評価  
 (楽器)

特徴量	Entropy の値
全音高平均	0.6220448
主旋律音高平均	0.6399332
全音高分散	0.7165634
主旋律音高分散	0.7320648
下降音高分散	0.7378158
テンポ	0.7449612
局所パターン	0.7521917
上昇音高平均	0.7554811
下降音高平均	0.758755
下降音割合	0.7652195
上昇音高分散	0.772679
16 分音符割合	0.7767722
水平音割合	0.7790158
全音高ズレ分散	0.7937278
8 分音符割合	0.7984945
付点 4 分音符割合	0.7999695
上昇音割合	0.805294
4 分音符割合	0.8097607
調性外楽音割合	0.8116282
2 分音符割合	0.8196802
全音符割合	0.8307155
付点 2 分音符割合	0.8330935

表 8 Entropy による性能評価  
 (ジャンル)

特徴量	Entropy の値
局所パターン	0.7296329
全音高分散	0.7859255
テンポ	0.8238935
全音高平均	0.8305142
調性外楽音割合	0.8385366
下降音割合	0.8509442
上昇音割合	0.8890649
上昇音高平均	0.8904839
下降音高平均	0.8987141
主旋律音高分散	0.9098266
上昇音高分散	0.9030486
水平音割合	0.9040377
下降音高分散	0.9091575
全音符割合	0.9102565
付点 2 分音符割合	0.9146482
8 分音符割合	0.9172375
付点 4 分音符割合	0.923797
16 分音符割合	0.9280687
4 分音符割合	0.9297045
2 分音符割合	0.9405231
主旋律音高平均	0.9499487

表 9 局所パターンの平均出現割合

	B	E	G	O	P	S
pattern1	0.08092	0.06248	0.07242	0.09847	0.05453	0.05071
pattern2	0.02117	0.02518	0.01531	0.03718	0.01473	0.01634
pattern3	0.03330	0.03270	0.05072	0.05266	0.04607	0.03743
pattern4	0.07087	0.01634	0.06206	0.08847	0.01718	0.04531
pattern5	0.03327	0.06443	0.07081	0.07480	0.10877	0.05182
pattern6	0.05193	0.08348	0.05310	0.08624	0.11211	0.06857
pattern7	0.05954	0.10966	0.14982	0.15392	0.13255	0.05622
pattern8	0.06967	0.03286	0.01544	0.03128	0.03864	0.04838
pattern9	0.20624	0.02040	0.02747	0.00126	0.00461	0.32601
pattern10	0.08567	0.03736	0.07570	0.03861	0.01182	0.04127
pattern11	0.03858	0.06085	0.08292	0.04501	0.09525	0.04565
pattern12	0.05128	0.07297	0.05581	0.03919	0.09693	0.06419
pattern13	0.07937	0.06936	0.13794	0.12507	0.12283	0.03918
pattern14	0.05344	0.06846	0.06021	0.04894	0.02322	0.02884
pattern15	0.02907	0.12058	0.06578	0.04215	0.07261	0.04330
pattern16	0.00823	0.00317	0	0.00568	0.00130	0.00880
pattern17	0.02737	0.11962	0.00439	0.03099	0.04677	0.02790

表 9 から、旋律の推移パターンについて、Bass や Synthesizer は水平パターンの出現確率が高いことがわかる。一方、Organ や Piano の場合は水平パターンの出現確率は低い。Guitar と Organ は、局所パターンにおける pattern7 や pattern13 の出現確率が高いなど、旋律推移が似ている。単調上昇パターンの出現確率に比べて単調下降パターンの出現確率が低い点は、各楽器共通している。

本研究で定義した特徴量には楽曲の主旋律を解析対象としているものがあるため、ループを対象に実験を行う際には、特徴抽出を行う前に主旋律推定を行った。しかし、ループの中には伴奏パートのみ演奏されているものもあるため、主旋律が必ず存在するとは限らない。本研究ではそのようなループに対しても主旋律推定を行い研究対象データとしたため、抽出結果には課題がある。主旋律推定時には、主旋律の有無を考慮する必要があると考える。

## 5. おわりに

本研究では、楽器やジャンルといった楽曲属性を反映する楽曲特徴量を、計算機上で動的に選出するモデルを提案した。具体的な流れとしては、まず楽曲特徴量を独自に定義し、楽曲属性毎に用意した楽曲から特徴量を抽出した。次に、特徴量毎に楽曲をクラスタリングし、クラスタリング結果の性能を評価することで、楽曲属性の特徴反映具合に関する楽曲特徴量のランキングを得た。

今後は、自動作曲への適用が有効と考えられる特徴量に対し、具体的にどのような値をとるのかを解析し、実際に自動作曲への適用を予定している。例えば、局所パターンに関しては、得られた出現確率を元に trigram の確率を求め、それによってメロディ生成を行うことで、自動作曲システムに反映できると考える。自動作曲システムとしては、Hevner の研究成果を用いて我々が実現した自動作曲システムを用いる。

属性の特徴を反映しているとは言いにくい。

したがって、例えば楽器の全音高平均の特徴量とジャンルの局所パターンの特徴量において、抽出した特徴量を解析して得られた値を自動作曲システムに用いる事により、特定の楽器及び特定のジャンルらしさを反映させることができ、各楽曲属性を組み合わせた作曲が可能になる、と考える。

特徴抽出結果の例として、楽曲属性を楽器、特徴量を局所パターンとして特徴抽出を行った結果を表 9 に示す。

## 参考文献

- [1] 中山達喜, 吉田真一: 音楽分類における特徴量の検討, 日本知能情報ファジィ学会ファジィシステムシンポジウム講演論文集 Vol.26, pp.290-, 2010.
- [2] 辻康博, 星守, 大森匡: 曲の局所パターン特徴量を用いた類似局検索・感性語による検索, 電子情報通信学会技術研究報告, SP96-124(1997-03), pp.17-24, 1997.
- [3] 吉野太智, 高木秀行, 清木康, 北川高嗣: 楽曲データを対象としたメタデータの自動生成とその意味的連想検索への適用, 情報処理学会研究報告, 1998-DBS-116(2), pp.109-116, 1998.
- [4] K.Hevner, "Expression in music: A discussion of experimental studies and theories," *Psychological Review*, Vol.42, pp.186-204, 1935.
- [5] K.Hevner, "Experimental studies of the elements of expression in music," *American Journal of Psychology*, Vol.48, pp.246-268, 1936.
- [6] K.Hevner, "The affective value of pitch and tempo in music," *American Journal of Psychology*, Vol.49, pp.621-630, 1937.
- [7] 芳村亮, 中西崇文, 北川高嗣: 任意の言葉を対象とした楽曲自動生成方式, 第17回データ工学ワークショップ論文集, 電子情報通信学会, 2006.