

# 任意の言葉の音相の印象に合致した楽曲の自動生成方式

岡田龍太郎<sup>†</sup> 芳村 亮<sup>†</sup> 本間 秀典<sup>†</sup> 北川 高嗣<sup>†</sup>

<sup>†</sup> 筑波大学大学院システム情報工学研究科

**あらまし** 本稿では、言葉の音声表現により人間が受ける印象を用いて、任意の日本語の印象に合致した楽曲自動生成方式を提案する。我々は、音相理論と呼ばれる理論を用いて、入力された任意の言葉から、その音声表現から受ける印象に合致した印象語をメタデータとして抽出する方式を実現している。また、Hevnerの研究を用いて、楽曲の印象を表す8つの印象語群と、6つの楽曲構造要素の相関を求めることにより、楽曲を自動生成する方式を実現している。本稿では、この印象語抽出方式と楽曲自動生成方式を意味の数学モデルを用いてメタレベルで連結することにより、任意の言葉の音相の印象に合致した楽曲自動生成方式を実現する。

**キーワード** 自動作曲, 情報生成, 楽曲メディアデータ, 音相, メタデータの抽出, 感性

## An Automatic Music Composition Method Corresponding to the Impression of the Sounds of Arbitrary Words

Ryotaro OKADA<sup>†</sup>, Ryo YOSHIMURA<sup>†</sup>, Hidenori HOMMA<sup>†</sup>, and Takashi KITAGAWA<sup>†</sup>

<sup>†</sup> Graduate School of Systems and Information Engineering, University of Tsukuba

**Abstract** This paper presents an automatic music composition method corresponding to the impression of arbitrary Japanese words by using the impressions received from the sound expression of the words. Founded on the “onso” theory, we have realized an automatic extraction method which extracts impression words as metadata corresponding to the impression by the sound expression of arbitrary words. And we have also realized a method of an automatic composition method of music, by searching for correlations between 8 groups of words about musical impressions and 6 musical structural elements by using the reserches done by Hevner. In this paper, we realize the method of automatic composition of music corresponding to the impression of the sounds of arbitrary words by connencting the impression words extraction method and the method of automatic composition of music on meta-level by using a mathmatical model of meanings.

**Key words** automatic music composition, generation of infomation, music media data, “Onso”, extraction of metadata, kansei

### 1. はじめに

コンピュータネットワークの広域化および高速化やコンピュータの廉価化に伴う計算機資源の増大により、多種多様かつ多量なメディアデータ群がネットワーク上に蓄積され利用されつつある。しかしながら、人間がコンピュータを介してこれらの情報群を利用する機会は増大する一方、膨大な情報群の中からその意図に合致した情報を獲得するための負担も大きくなっていく。現在、これらの情報群を対象としたシステムはメディアデータの種類ごとに独立して存在している。このような状況において、異種のメディアデータを自由に連結し、個々のシステムの連携や統合を可能とする方式の実現が重要となると考えられる。

我々はこれまでに、言葉と言葉の関係の計量による検索機構として、意味の数学モデルによる意味的連想検索を提案している。これは、単語群を文脈として解釈する機構により、言葉と言葉、あるいは言葉とメディアデータ、ドキュメント間の相関を文脈に応じて動的に計算することを可能とするモデルである。

また、これまでに、メディアデータから人間の心理的要素が反映されたメタデータを抽出する方式として、「音相理論 [5]」と呼ばれる理論を用いて、日本語の音の印象に合致したメタデータを抽出する方式 [4] を実現しており、また、印象を表すメタデータからメディアデータを生成する方式として、音楽心理学者 Hevner による研究 [1]~[3] を用いて、任意の印象語から楽曲を自動生成する方式 [9] を実現している。

これらの方式を意味の数学モデルに代表される意味的連想検

索を用いてメタレベルで連結することにより、独立に実装されたメディア情報の相互運用性を実現し、既存のデータベース群の利用価値を高め、新たな価値の創造が可能になると考えられる。

本稿では、音相と楽曲メディアデータに着目し、言葉の音声情報の持つ印象から音楽を生成する楽曲自動生成方式について述べる。本方式は、ユーザから与えられた任意の日本語から、任意の言葉を対象とした音の印象によるメタデータ自動抽出方式[4]と任意の印象語を対象とした楽曲自動生成方式[9]を意味の数学モデル[6],[7]を用いてメタレベルで連結することによって、任意の日本語をコンテキストとして楽曲を自動生成する。

本方式は印象語の抽出に言葉の音声表現を用いる方式であるので、利用者の意図する印象を持つ語の印象に合致した楽曲の生成が実現されることで、音楽を通じた情報の新しい価値を生み出すことが可能になる。さらに、未知語や擬音語など、普段我々がその言葉が持つ印象を意識しない言葉でさえ、その印象を表した楽曲を生成することができるので、人間の感性に新たな発見を見出す手助けになると考えられる。

## 2. 意味の数学モデルの概要

本節では、言葉と言葉の関係の計量を実現する意味の数学モデルの概要を示す。詳細は、文献[6],[7]に述べられている。

### (1) メタデータ空間 *MDS* の設定

検索対象となるメディアデータをベクトルで表現したデータにマッピングするための正規直交空間（以下、メタデータ空間 *MDS*）を設定する。

### (2) メディアデータのメタデータをメタデータ空間 *MDS* へ写像

設定されたメタデータ空間 *MDS* へ、メディアデータのメタデータをベクトル化し写像する。これにより、検索対象データのメタデータが同じメタデータ空間上に配置されることになり、検索対象データ間の意味的な関係を空間上での距離として計算することが可能となる。

(3) メタデータ空間 *MDS* の部分空間（意味空間）の選択  
検索者は与える文脈を複数の単語を用いて表現する。検索者が与える単語の集合をコンテキストと呼ぶ。このコンテキストを用いてメタデータ空間 *MDS* に各コンテキストに対応するベクトルを写像する。これらのベクトルは、メタデータ空間 *MDS* において合成され、意味重心を表すベクトルが生成される。意味重心から各軸への射影値を相関とし、閾値を超えた相関値（以下、重み）を持つ軸からなる部分空間（以下、意味空間）が選択される。

### (4) メタデータ空間 *MDS* の部分空間（意味空間）における相関の定量化

選択されたメタデータ空間 *MDS* において、メディアデータベクトルのノルムを検索語列との相関として計量する。これにより、与えられたコンテキストと各メディアデータとの相関の強さを定量化している。この意味空間における検索結果は、各メディアデータを相関の強さについてソートしたリストとして与えられる。

## 3. 任意の言葉の音相の印象に合致したメタデータ自動抽出方式の概要

本節では、日本語の音として表現可能な任意の言葉を対象として、その音声表現のみによって生じる印象を表す言葉をメタデータとして自動抽出する方式[4]について述べる。この方式は、音相理論[5]と呼ばれる研究に基づいて、それぞれの音の音素やそれを発音するための調音部位（調音点とも呼ばれる）や調音方法から、それらの組み合わせにより生じる印象を抽出する方式である。

### 3.1 音相理論

木通はそれぞれの言葉が持つ音の構造の違いによって生まれる表情を音相と呼んだ。音相理論では、音相を捉えるための単位として、それぞれの音の最小単位である音素と、それらの音素を発音するために使われる調音器官を表す調音点、および発音の方法にあたる調音法が挙げられている。さらに、各音相基がどのような印象を表すときに使われているかを調査することにより、その音素の動性（強さ）と輝性（明るさ）を明らかにした。音相理論において、音相の単位として挙げられたこれらの要素は音相基と呼ばれており、40の音相基が定義されている。そして単独の音相基、または2つの音相基の組み合わせと、音相理論では“表情”と呼ばれる、音相基から生じる印象の相関関係、および各表情とそれらに合致する印象語群についてまとめられている[5]。

### 3.2 音相の印象に合致したメタデータ自動抽出方式

本節では、音相理論に基づいて、言葉の音が持つ印象を表す言葉をメタデータとして自動抽出するための方式を示す。

#### (1) 音相基の抽出

この方式では、初期パラメータとして、入力語から40種の音相基を表現するのに必要十分な29の要素からなる初期ベクトル  $v_i$  を抽出し、音相基と表情の関係を用いて40種の音相基を表す40次元のベクトル  $v_0$  へと変換する。

#### (2) $v_0$ の拡張

それぞれの表情が有効であるかを判断する尺度として各音相基の標準使用率が示されているので、これを用いて  $v_0$  を補正する。さらに、2つの音相基の組み合わせによりさらに生じる38種の表情の判定を行い、 $v_0$  を78次元のベクトル  $v_0$  に拡張する。

#### (3) 印象語群とその重みの出力

それぞれの表情とそれらから抽出される印象語群の関係から変換行列  $T$  を作成し、以下により40語の印象語を表す40次元のベクトル  $w_0$  に変換する。

$$w_0 = T v_0.$$

ここでの40語の印象語は、2語を1組として1種の表現属性を表しており、20種の表現属性を表している。表情語と表現属性を表1に示す。この  $w_0$  が音相の印象に合致したメタデータ抽出方式により抽出される、言葉の音の印象を表す重み付き印象語群によるメタデータである。

表 1 音相を表す表情語と表情属性表.

表情語群	表情属性	訳語
A	シンプルな, 明白さ	plain, obvious
B	躍動感, 進歩的	vibrant, advance
C	新鮮さ, 新奇さ	fresh, unprecedented
D	動的, 活性的	dynamic, active
E	派手さ, 賑やかさ	florid, bustle
F	軽やかさ, 軽快感	light, trippingly
G	若さ, 澆刺さ	young, effervescent
H	現代的, 都会的	modern, urban
I	明るさ, 開放的	bright, open-minded
J	合理的, 現実的	reasonable, real
K	個性的, 特殊的	individual, special
L	強さ, 鋭さ	powerful, sharp
M	適応性, 庶民的	adaptable, popular
N	清らかさ, 爽やかさ	pure, brisk
O	健康的, 清潔感	healthy, clean
P	暖かさ, 安らぎ	warm, comfortable
Q	安定感, 信頼感	stable, confidence
R	高級感, 充実感	expensive, fulfil
S	高尚な, 優雅さ	profound, elegant
T	静的, 非活性的	static, inactive

		c6 bright cheerful gay happy joyous merry		c5 delicate fanciful graceful humorous light playful quaint sprightly whimsical		c4 calm leisurely lyrical quiet satisfying serene soothing tranquil
	c7 agitated dramatic exciting exhilarated impetuous passionate restless sensational soaring triumphant					
c8 emphatic exalting majestic martial ponderous robust vigorous						
	c1 awe-inspiring dignified lofty sacred serious sober solemn spiritual					
			c2 dark depressing doleful frustrated gloomy heavy melancholy mournful pathetic sad tragic		c3 dreamy longing plaintive sentimental tender yearning yielding	

図 1 Hevner による 8 つの印象語群

印象語群名	c1	c2	c3	c4	c5	c6	c7	c8
key	長調 4	短調 12	短調 20	長調 3	長調 21	長調 24	—	—
tempo	遅い 14	遅い 12	遅い 16	遅い 20	速い 6	速い 20	速い 21	速い 6
pitch	低い 10	低い 19	高い 6	高い 8	高い 16	高い 6	低い 9	低い 13
rhythm	固定 18	固定 3	流動 9	流動 2	固定 8	流動 10	固定 2	固定 10
harmony	単純 3	複雑 7	単純 4	単純 10	単純 12	単純 16	複雑 14	複雑 8
melody	上昇 4	—	—	上昇 3	下降 3	—	下降 7	下降 8

図 2 各印象語群に対する楽曲構造要素の相対重要性の表

	印象語群							
	c1	c2	c3	c4	c5	c6	c7	c8
key	0.0476	-0.1429	-0.2381	0.0357	0.2500	0.2857	0.0000	0.0000
tempo	-0.1217	-0.1043	-0.1391	-0.1739	0.0522	0.1739	0.1826	0.0522
pitch	-0.1149	-0.2184	0.0690	0.0920	0.1839	0.0690	-0.1034	-0.1494
rhythm	0.2903	0.0484	-0.1452	-0.0323	0.1290	-0.1613	0.0323	0.1613
harmony	0.0405	-0.0946	0.0541	0.1351	0.1622	0.2162	-0.1892	-0.1081
melody	0.1600	0.0000	0.0000	0.1200	-0.1200	0.0000	-0.2800	-0.3200

図 3 変換行列 T

#### 4. 任意の印象語を対象とした楽曲自動生成方式

本節では, 入力された言葉から印象のメタデータを抽出し, それに合致した楽曲を自動的に生成する, 楽曲自動生成方式について述べる. 詳細は文献 [9] に述べられている.

##### 4.1 Hevner の研究

Hevner の研究 [1]~[3] では, 楽曲構造要素として調性 (key)・テンポ (tempo)・音高 (pitch)・リズム (rhythm)・和声 (harmony)・旋律 (melody) の 6 つを挙げている. Hevner は, この 6 つの楽曲構造要素と 8 つの印象語群 (図 1) によって表現される印象との相関関係を調べ, 表にまとめた (図 2). 8 つの印象語群は, 印象語間で類似性があるものをまとめて一つの印象語群を作り, さらに印象語群間で類似性があるものを隣接するよう, 円形に配置している.

##### 4.2 任意の印象語を対象とした楽曲自動生成方式全体の概要

本方式は次のような手順により実現される.

- Step1: 楽曲印象語群ベクトルの出力

2. 節に示した意味の数学モデルを用いて, 入力に与えられた任意の言葉と, Hevner による 8 つの各印象語群との相関を求める. さらに, その各相関値を, 1) 相関値のうち最大のものを 1 に写像, 2) 相関値のうち最小のものを -1 に写像 することにより正規化し, 楽曲印象語群ベクトル  $c = (v_{c1}, v_{c2}, \dots, v_{c8})^T$ , ( $v_{ci}$  は語群  $c_i$  への重み) を生成する.

- Step2: 楽曲構造要素ベクトルの出力

図 2 で示した, Hevner による各印象語群に対する楽曲構造要素の相対重要性の表に正負の符号をつけ, 行ごとに正規化したものを変換行列  $T$  とする (図 3). 楽曲印象語群ベクトルを入力として, 変換行列  $T$  を作用させることにより,

6 つの楽曲構造要素への重みを表す楽曲構造要素ベクトル  $f = (key, tempo, pitch, rhythm, harmony, melody)^T$  を生成する.

$$f = Tc.$$

- Step3: 楽曲の特徴値の出力

出力された  $f$  の値をもとに, 生成する楽曲の特徴値  $kn, tem, mnap, unac, ac, tc, oc, um, dm, lm$  を決定する. 表 2 に楽曲の特徴値の意味を示す.

- Step4: 楽曲の生成・出力

Step3 で決定した特徴値をもとに, 4.3 節に示す楽曲の仕様にしたがって楽曲を生成する. 特徴値のうち, 和音による伴奏の生成には  $kn, mnap, unac, ac, tc, oc$  を使用し, 旋律の生成には  $kn, mnap, um, dm, lm$  を使用する.

#### 4.3 楽曲の仕様

ここでは, 本方式において生成する楽曲の仕様について示す.  
**楽曲の拍子・長さ** 本方式では本格的な曲を生成する必要はなく, 入力された言葉の持つ意味や感情が短いフレーズに反映さ

表 2 楽曲の特徴値

kn	楽曲の調性
tem	テンポ
mnap	旋律の音高の平均
umac	四分音符の和音の総演奏時間
ac	八分音符の和音の総演奏時間
tc	三和音の総演奏時間
oc	三和音以外の和音の総演奏時間
um	上昇する旋律音の総演奏時間
dm	下降する旋律音の総演奏時間
lm	水平な旋律音の総演奏時間

ればよいので、生成する楽曲は 4/4 拍子、長さは 4 小節とする。

**和音** 一つの楽曲に使用しうる和音は非常に多いので制限を与える。使用する和音は I 度から VII 度までの 7 つで、和音の構成音として使用する音はその調の音階に沿ったものとする。長調では I, II<sub>m</sub>, III, IV, V<sub>7</sub>, VI<sub>m</sub>, VII<sub>m</sub><sup>(-5)</sup>、短調では I<sub>m</sub>, II<sub>m</sub><sup>(-5)</sup>, III, IV<sub>m</sub>, V<sub>7</sub>, VI, VII である。また、和音の長さは四分音符と八分音符の 2 種類のみとする。

**旋律** 旋律に使用する音はその調の音階に沿ったものとする。また、Hevner の研究では旋律音の長さやリズムには言及していないので、ここでは旋律音の長さは四分音符と八分音符の 2 種類のみとする。

## 5. 任意の言葉の音相の印象に合致した楽曲の自動生成方式

本節では、任意の言葉の音声情報からその印象に合致した楽曲を自動生成する方式について示す。本方式の概要は以下の通りである。全体の概要図を図 4 に示す。

### (1) 言葉の印象の抽出

3.2 節に示した、対象となる語の音相に合致した印象語の抽出方式を用いて、ユーザから入力された任意の言葉の音相から、印象語ベクトル  $w_o$  を抽出する。

### (2) 印象語ベクトル $w_o$ の重みの補正

(1) により得られた印象語ベクトル  $w_o$  を、より入力語毎の差が出やすいように補正する。詳細は 5.1 節に示す。

### (3) 印象語ベクトルと各楽曲印象語群との相関の計量

2. 節に示した意味の数学モデルを用いて、(2) で得られた印象語ベクトル  $w_o$  と Hevner の指定する 8 つの各楽曲印象語群との相関を計量し、楽曲印象語群ベクトル  $c$  を構成する。

### (4) 楽曲の生成

(3) で得られた楽曲印象語群ベクトル  $c$  から、3.2 節に示した任意の印象語を対象とした楽曲自動生成方式を用いて楽曲を生成する。

### 5.1 印象語の重みの補正

音相を表すとして  $w_o$  に選ばれた語は、肯定的な意味を表す印象語のみで構成されているため、Hevner の指定する 8 つの印象語群との相関を計量した際に、特定の印象が強調されやすい傾向があり、出力される楽曲において入力語毎の印象の違いを正確に反映できていない可能性がある。

そこで、重みの小さい語をマイナスの印象として見ることに

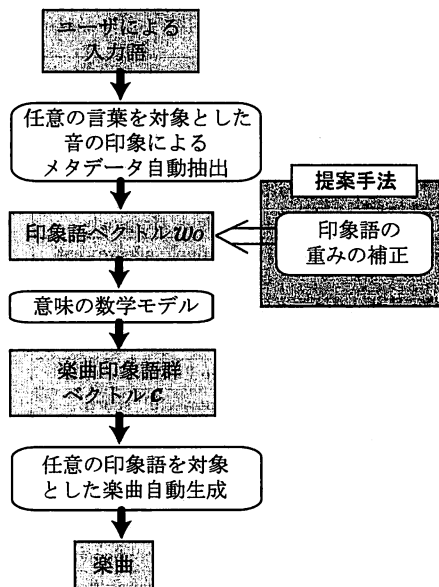


図 4 システム全体の概要図

より、印象語ベクトル  $w_o$  を Hevner の指定する 8 つの印象語群に変換した際に強調されにくい印象を反映することができると考えられる。

ここでは、予備実験に基づき、音相から選ばれた印象語の重みを、最小を -1、最大を 1 になるよう写像する。印象語ベクトル  $w_o$  の要素を  $w_{oi}$ 、写像後の要素を  $w'_{oi}$  としたとき、上記の写像を以下の式に示す。ただし、 $\max(w_o)$ 、 $\min(w_o)$  は、それぞれ  $w_o$  の要素のうち最大、最小のものを表す。

$$w'_{oi} = \frac{2w_{oi} - (\max(w_o) + \min(w_o))}{\max(w_o) - \min(w_o)}$$

## 6. 実験例

本節は、提案している異種メディアを対象とした連結方式の有効性を検証するため、本方式に基づく実験システムを構築した。

### 6.1 実験環境

本実験では、意味の数学モデルを用いて検索システムの構築を行った。意味の数学モデルの基本構成は 2. 節に示す通りである。メタデータ空間  $MDS$  の設定については、“Longman Dictionary of Contemporary English” [8] という英英辞典を利用した。同辞書は約 2,000 語の基本語を用いて約 56,000 語の見出し語を説明している。この基本語を特徴とみなし、見出し語の説明で肯定的に使われている語を “1”、否定的に使われている語を “-1”、説明に使われていない語を “0” として 2. 節 (1) におけるデータ行列  $M$  を作成した。これにより、約 2,000 次元の正規直交行列であるメタデータ空間  $MDS$  が生成された。 $MDS$  では、約  $2^{2000}$  通りの意味の様相を表現可能である。

### 6.2 実験システム

この実験システムを、C 言語および Perl 言語を用いて実装した。このシステムでは、WWW ブラウザをユーザインタフェースとし、任意の日本語を入力として生成された楽曲を呈示する。

表3 実験2, “ニコニコ”を入力した時の各値

楽曲印象語群ベクトル	楽曲構造要素ベクトル	楽曲特徴値
C1	-1.000000	Key 0.426002 kn E
C2	-0.091849	Tempo 0.407745 tem 145
C3	-0.968848	Pitch 0.181692 mnap 76.086006
C4	-0.528090	Rhythm -0.407078 unac 0.296461
C5	-0.148128	Harmony 0.187565 ac 0.703539
C6	1.000000	Melody 0.081811 tc 0.593782
C7	-0.507758	oc 0.406218
C8	-0.453857	um 0.415754
		dm 0.333943
		lm 0.250303

表4 実験2, “ガッカリ”を入力した時の各値

楽曲印象語群ベクトル	楽曲構造要素ベクトル	楽曲特徴値
C1	-0.565631	Key -0.025211 kn Am
C2	1.000000	Tempo -0.296871 tem 74
C3	-1.000000	Pitch -0.008186 mnap 68.607066
C4	0.872121	Rhythm -0.235652 unac 0.382174
C5	-0.517421	Harmony 0.160393 ac 0.617826
C6	0.016428	Melody 0.622389 tc 0.580196
C7	-0.990498	oc 0.419804
C8	-0.840017	um 0.657795
		dm 0.035406
		lm 0.306799

さらにその生成過程のデータとして、音相を表す印象語ベクトル、楽曲印象語群ベクトル、楽曲構造要素ベクトル、楽曲特徴値を呈示する。

### 6.3 出力例

入力として、文献[5]において“暖かさ、和やかさ”及び“明白さ、単純さ”に関する語として分類されている“ニコニコ”という言葉と、“不快、苦痛、哀歎”に関する語として分類されている“ガッカリ”という言葉を用いて楽曲の生成を行った。さらに、明確な印象の知られていない、コンピュータによってランダムで生成された“ヒンミパ”という言葉を用いて、楽曲の生成を行った。

楽曲自動生成を行った結果について、“ニコニコ”を入力した場合の、楽曲印象語群ベクトル、楽曲構造要素ベクトル、楽曲特徴値の各値の結果を表3に、“ガッカリ”を入力した場合の結果を表4に、“ヒンミパ”を入力した場合の結果を7に示す。また、実際に出力された楽曲の楽譜を、“ニコニコ”を図5、“ガッカリ”を図6、“ヒンミパ”を図7に示す。加えて、“ニコニコ”、“ガッカリ”、“ヒンミパ”の、音相から抽出された補正後の印象語ベクトル  $w_0$  を表6に示す。

### 6.4 実験の考察

“ニコニコ”、“ガッカリ”、及び“ヒンミパ”を入力したすべての場合において、楽曲が生成されていることが確認できた。

“ニコニコ”(表3)においては、印象語群 c6, すなわち、“bright”, “happy”などを表す印象が高くなっている。これは語の印象に一致していると考えられる。“ガッカリ”(表4)においては、印象語群 c2, すなわち、“dark”, “sad”などの印象

表5 実験2, “ヒンミパ”を入力した時の各値

楽曲印象語群ベクトル	楽曲構造要素ベクトル	楽曲特徴値
C1	-1.000000	Key 0.103564 kn Bb
C2	0.471624	Tempo 0.022286 tem 91
C3	-0.336413	Pitch 0.123486 mnap 73.815956
C4	1.000000	Rhythm -0.410771 unac 0.294615
C5	-0.191180	Harmony 0.094953 ac 0.705385
C6	0.526890	Melody 0.068167 tc 0.547477
C7	0.066480	oc 0.452523
C8	-0.324498	um 0.383960
		dm 0.315794
		lm 0.300246



図5 実験2, “ニコニコ”から生成された楽曲



図6 実験2, “ガッカリ”から生成された楽曲

が高くなっている。これは語の印象に一致していると考えられる。“ヒンミパ”(表7)においては印象語群 c4が高くなっている。楽曲特徴値においては、“ニコニコ”の調は“E”というメジャーキーであり、テンポも145と速く、旋律音の平均が高い楽曲が生成されている。また、“ガッカリ”の調は“Am”というマイナーキーであり、テンポも比較的遅く、旋律音の平均も低い楽曲が生成されている。“ヒンミパ”においては、調は“Bb”というメジャーキーであり、テンポも比較的速い楽曲が生成さ

表6 実験2, 補正後の印象語ベクトル  $w_o$ .

ニコニコ		ガッカリ		ヒンミバ	
印象語	重み	印象語	重み	印象語	重み
urban	1.000000	plain	1.000000	open-minded	1.000000
dynamic	1.000000	obvious	1.000000	bright	1.000000
active	1.000000	trippingly	0.909091	healthy	0.743589
modern	1.000000	light	0.909091	clean	0.743589
open-minded	0.652174	unprecedented	0.363636	plain	0.692307
bright	0.652174	reasonable	0.363636	obvious	0.692307
⋮	⋮	⋮	⋮	⋮	⋮
inactive	-0.627826	sharp	-0.909091	sharp	-0.846153
static	-0.627826	powerful	-0.909091	powerful	-0.846153
sharp	-1.000000	fulfil	-0.945455	inactive	-0.879598
plain	-1.000000	expensive	-0.945455	static	-0.879598
powerful	-1.000000	brisk	-1.000000	fulfil	-1.000000
obvious	-1.000000	pure	-1.000000	expensive	-1.000000



図7 実験2, “ヒンミバ” から生成された楽曲

れている。

これらにより、任意の言葉の音相の印象に合致した楽曲自動生成方式が実現されていること、及びに、印象語の重みの補正の有効性が確認できた。

## 7. おわりに

本稿では、音相と楽曲メディアデータに着目し、ユーザから与えられた任意の日本語の音声情報の持つ印象に合致した楽曲を自動生成する方式を実現した。

本方式により、音相と楽曲メディアデータという独立に実装されたメディア情報の相互運用性を実現し、既存のデータベース群の利用価値を高め、新たな価値の創造が可能になる。また、普段我々はその言葉が持つ印象を意識しない言葉の印象を表した楽曲を生成することは、人間の感性に新たな発見を見出す手助けになると考えられる。

また、本方式を利用して、商品名のイメージに合ったCMの楽曲を自動生成するシステムや、詩のイメージに合った楽曲を自動生成するシステムなど、人間の感性を増幅するシステムの開発が可能になると考えられる。

今後の課題としては、本方式の定量的な評価方式、および他

メディアを対象としたメディアデータ連携統合方式の実現が挙げられる。

## 文 献

- [1] Hevner, K., “Expression in music: A discussion of experimental studies and theories”, *Psychological Review*, vol.42, pp.186-204, 1935.
- [2] Hevner, K., “Experimental studies of the elements of expression in music”, *American J. Psychology*, vol.48, pp.246-268, 1936.
- [3] Hevner, K., “The affective value of pitch and tempo in music”, *American J. Psychology*, vol.49, pp.621-630, 1937.
- [4] 本間 秀典, 中西 崇文, 北川 高嗣: “任意の言葉を対象とした音の印象によるメタデータ自動抽出方式”, 電子情報通信学会技術研究報告, Vol.106, No.149, pp.7-12, 電子情報通信学会, 2006.
- [5] 木通隆行: “日本語の音相 — ことばのイメージを捉える技術, 表現する技術 —”, 小学館スクウェア, (2004).
- [6] Kitagawa, T. and Kiyoki, Y.: “The mathematical model of meaning and its application to multidatabase systems”, *Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems*, pp. 130-135(1993).
- [7] Kiyoki, Y., Kitagawa, T. and Hayama, T.: “A Metadata-based System for Semantic Image Search by a Mathematical Model of Meaning”, *Multimedia Data Management - using metadata to integrate and apply digital media -*, McGrawHill, A. Sheth and W. Klas(editors), Chapter 7 (1998).
- [8] *Longman Dictionary of Contemporary English*, Longman (1987).
- [9] 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.