

任意の言葉を対象とした音韻印象変換作用素の構成とその感性検索への適用

本 間 秀 典^{†1,*1} 中 西 崇 文^{†1,*1} 北 川 高 嗣^{†1}

本稿では、任意の言葉を対象とした音韻印象変換作用素の構成方式を示す。これは、与えられた任意の言葉からその言葉が持つ音の印象を自動抽出する作用素を実現する。この方式は、言葉の音素情報のみを用いてその印象の抽出を実現するため、未定義語や未知語にも柔軟に対応でき、明確に感性を表現しているにもかかわらず、辞書による定義が困難な言葉にも印象語による対応が可能であるという利点を持つ。これにより、ユーザの任意の言葉や未定義・未知の言葉による感性表現が可能となり、ユーザへの負荷が少ないインタフェースが実現されると考えられる。また、本稿では、本方式の適用例、および、実験システムとして、楽曲 MIDI データを対象とした本方式を有する感性検索を実現し、本方式の有効性を示した。

A Phoneme-Impression Transformation Operator Construction for an Arbitrary Word and Its Application to Kansei Retrieval

HIDENORI HOMMA,^{†1,*1} TAKAFUMI NAKANISHI^{†1,*1}
and TAKASHI KITAGAWA^{†1}

In this paper, we present a phoneme-impression transformation operator construction method for an arbitrary word. This method extracts impression that comes from sounds of arbitrary words, including undefined words, unknown words, and new words. It realizes interfaces between persons and computers utilizing user's feeling expression by undefined/unknown words. This paper presents a Kansei retrieval system by utilizing this method for music MIDI data as an experiment system and an application example. Finally we perform various experiments on the system in order to verify the effectiveness of this method.

1. はじめに

言葉は、意味を伝達するための記号としてだけでなく、発音そのものが話者の印象を伝達するメディアとしての機能も備えている。たとえば、日本語では、他の言語と比べると、オノマトペと呼ばれる擬態語・擬音語が多数存在し¹⁾、日常的なコミュニケーションにおいて、物事の状態を表すメディアとして、非常に重要な位置を占めている。

これまで、擬音語と喚起される印象の関係を示そうとする研究が行われている²⁾⁻⁴⁾。これらは、擬音語の音韻的特徴とイメージされる音の印象の関係を示そうとしている。また、文献 5) では、オノマトペメディアと名付けられた、Web 上からオノマトペの用例として適切な文章を自動的に収集し提示するシステムを実現している。このように、日本語において、言葉の音が表す印象を適切にとらえることが重要となる。

また、人、商品、企業、ブランドなどのネーミングにおいて“簡明性”、“視覚性”、“サウンド性”の重要性が指摘されている⁶⁾。簡明性とは、その対象を表現する言葉として分かりやすいかという意味についての観点である。視覚性とは、その対象を表現する言葉を字にしたときの視覚的な感覚である。サウンド性とは、その言葉自身が持つ音の響きである。特に、ネーミングはテレビ CM やラジオにおいて音として伝えられることが多く、サウンド性、つまり、その言葉が持つ意味情報を排除した音の印象が重要となってきている。このことから、我々が日常的に使用している意味情報を持つ言葉でさえも、意味情報を排除した音が、印象を伝えるコミュニケーションに非常に重要な役割を果たしている。

一方、文献 7), 8) では、“音相理論”と呼ばれる研究において日本語における言葉の音韻と音相の相関関係が体系的に示されている。音相理論によれば、一般に語や語句といった言葉は様々な印象を喚起するが、その言葉の印象はその語句の持つ意味からだけでなく音の特徴からも喚起されるものであり、これを“音相”と呼んでいる。音相は、我々の日常的なコミュニケーションにおいて、互いの感情を理解するための重要な要素の 1 つであると考えられる。これらのことから、言葉を介する人間とコンピュータ間のインタフェースにおいて、言葉の意味情報だけでなく、字面などの視覚的な印象やその言葉が持つ語感や音相から想起される印象を介したコミュニケーション機構の実現が重要であると考えられる。

^{†1} 筑波大学大学院システム情報工学研究科

Graduate School of Systems and Information Engineering, University of Tsukuba

*1 現在、情報通信研究機構

Presently with National Institute of Information and Communications Technology

本稿では、言葉が持つサウンド性に着目し、任意の言葉から、その言葉の音韻が持つ印象を自動抽出するための方式を示す。

これまで、我々は、心理学などの専門分野における研究成果に基づいてそのメディアデータから印象語を抽出するための枠組みとして Media-lexicon Transformation Operator \mathcal{ML} ⁹⁾ を提案してきた。 \mathcal{ML} を用いて様々なメディアデータを印象語で表すことにより、異種メディアデータの連結¹⁰⁾ が可能となる。ここで、言葉をメディアデータととらえて音相理論^{7),8)} を適用することにより、任意の言葉からその言葉が持つ音の印象を自動抽出する \mathcal{ML} が構成される。この作用素は、日本語で表記可能なすべての言葉やフレーズの音韻によって伝えられる印象を、言葉の音素情報のみを用いて抽出する。そのため、本方式は、未定義語や未知語にも柔軟に対応でき、明確に感性を表現しているにもかかわらず、新語や造語など辞書による網羅的な定義が困難な言葉にも対応が可能である。

これにより、任意の言葉や未定義・未知の言葉による感性表現が可能となり、ユーザへの負荷が少ないインターフェースが実現される。このように、印象や感情などの様々な情報を効果的に表すメディアを含む異種のメディアを統一的に扱うことが可能な方式を実現できれば、人間の感性や直感に合致した、マルチメディア環境での有効なコミュニケーションメディアの実現の第一歩となる。

本稿の構成は以下のとおりである。2章で Media-lexicon Transformation Operator \mathcal{ML} について概説する。3章で音相理論の概要と、それを用いた任意の言葉からその言葉の音の印象に合致した \mathcal{ML} の構成方式を示す。4章で本稿で実装した任意の言葉によるその言葉の音の印象に合致した楽曲を検索する感性検索システムの実現方式を示す。5章で4章で用いた実験により提案方式の有効性を示す。最後に、6章で本稿をまとめる。

2. Media-lexicon Transformation Operator \mathcal{ML}

本章では、Media-lexicon Transformation Operator \mathcal{ML} ⁹⁾ について述べる。 \mathcal{ML} とは、メディアデータからその分野の専門家による研究や評論、統計などにより、人間がそのメディアデータから受ける印象を表す単語（印象語）を導出する作用素である。これまで、我々は \mathcal{ML} の応用例として、楽曲メディアデータを対象としたメタデータ自動抽出方式^{9),11),12)}、画像メディアデータを対象としたメタデータ自動抽出方式¹³⁾ を実現してきた。

\mathcal{ML} は一般的に次のように表される。

$$\mathcal{ML}(Md) : Md \mapsto Ws. \quad (1)$$

(Md : メディアデータ, Ws : (重み付き) 印象語群)

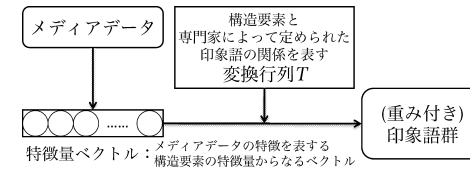


図1 Media-lexicon Transformation Operator \mathcal{ML} の構成図
Fig. 1 Configuration of Media-lexicon Transformation Operator \mathcal{ML} .

\mathcal{ML} により抽出された印象語をメタデータとして、検索者が発行する印象語から、その印象に合致したメディアデータの検索^{11)–13)} が可能となる。さらに、様々なメディアデータを人間の印象語を用いて表現することにより、画像メディアデータとそれらが持つ印象に合致した楽曲メディアデータを統合するような、異種メディアデータの連結が可能となる¹⁰⁾。

\mathcal{ML} の概要を図1に示す。 \mathcal{ML} の構成のポイントは変換行列 T にある。変換行列 T は専門家による研究や評論、統計による、メディアデータの m 個の構造要素とその特徴を表す n 個の印象語群との関係を示した $m \times n$ 行列として表される。構造要素とは、楽曲メディアデータではテンポなどの情報、画像メディアデータでは色彩情報などが該当する。メディアデータからその構造要素を表現する m 次元の特徴量ベクトル x を抽出し、変換行列 T によって n 次元の印象語ベクトル y を導出する。つまり、 \mathcal{ML} は次のように構成される。

$$y = Tx \quad (2)$$

印象語ベクトル y の値は、メディアデータに対する各印象語の相関とみることができる。これにより、 \mathcal{ML} はメディアデータに合致した重み付き印象語を導出する。

3. 任意の言葉を対象とした音韻印象変換作用素

本章では、提案方式である任意の言葉を対象とした音韻印象変換作用素の構成を示す。本方式では、2章で示した Media-lexicon Transformation Operator \mathcal{ML} の枠組みに専門家による研究成果として音相理論^{7),8)} を導入する。すなわち、式(1)の Md が任意の言葉、 Ws がその言葉の音の印象を表す重み付き印象語群となる。3.1節では、言葉の音とその印象の相関関係を示した研究の1つである音相理論を紹介する。3.2節では、提案方式である任意の言葉を対象とした音韻による \mathcal{ML} の構成方式について示す。

3.1 音相理論

文献7),8)では、日本語における音の構成要素とその印象の関係が音相理論としてまとめられている。音相とは、同じ言語を話すほとんどの話者が共通に持っている、言葉の音の

表 1 音相基の調音種とそれらが持つ輝性 (B), 勁性 (H) の関係 (子音の場合)

Table 1 The table of sound phase element for vowels.

		両唇音	前舌音		咽頭音
破裂音	無声音	p	t		k
		+B1.7 H1.6	+B1.4 H1.5		+B1.3 H1.3
	有声音	b	d		g
		-B2.2 H1.2	-B2.0 H1.0		-B2.0 H1.0
破擦音	無声音		ts	tʃ	
			+B1.7 H2.2	+2.2 H2.7	
	有声音		ɖ	ɟ	
			-B2.0 H1.0	-B1.5 H1.5	
摩擦音	無声音	f	s	ʃ	h
		B0.0 H0.0	+B0.3 H0.7	+B0.8 H1.2	B0.0 H0.0
弾き音 (流音)	有声音		r		
			+B0.4 H0.8		
鼻音	有声音	m	n		ŋ
		B0.0 H0.0	B0.0 H0.0		B0.0 H0.0
接近音	有声音	w	j		
		-B0.5 H0.0	+B0.5 H0.5		

- ・ ts: ツの子音
- ・ ɖ: ジを除くザ行の子音
- ・ f: フおよびファ行の子音
- ・ ɲ: ガ行鼻音
- ・ 促音 (Q) は後続子音に H1.3 を加える
- ・ 拗音は前子音に +B0.5, H0.5 をそれぞれ加える
- ・ 撥音 (N) は -B0.4 H0.0
- ・ tʃ: チおよびチャ行の子音
- ・ ɟ: ジ (ヂ) およびジャ行の子音
- ・ ʃ: シおよびシャ行の子音
- ・ j: ヤ行子音および拗音の第 2 子音

* 木通隆行著, 小学館スクウェア出版, p.107 表 17 より転載.

響きに対する感覚やイメージであるとされる。この理論によれば、我々は、その言葉の持つ意味の微妙な違いを、似た意味の言葉の中から、音相に基づいてよりそのイメージに近い音を持つ言葉を選んで使うことによっても伝えている。そのため音相は、我々が言葉によって伝えようとしている感情やイメージなどを直感的に表現していると考えられる。

音相理論では、音相基と呼ばれる音相を構成する要素を用いて対象となる言葉の印象を決定している。音相基の中でも特に、調音種と、それがどのような感情のときに用いられるか、という 2 つの観点で言葉の表情を生み出すとされている。前者は、調音点と呼ばれる、その音が発音される時に用いられる器官 (唇や舌など) と、調音法と呼ばれる、調音点の操作方法の 2 つからなる。後者は、その音の明るさを表す輝性 (Brightness, 以下 B) およ

表 2 音相基の調音種とそれらが持つ輝性 (B), 勁性 (H) の関係 (母音の場合)

Table 2 The table of sound phase element for vowels.

	有声音	B	H
a		B0.0	H0.0
i		+B1.0	H1.0
u		-B1.0	H0.0
e		B0.0	H0.0
o		-B0.7	H0.0

* 木通隆行著, 小学館スクウェア出版, p.107 表 17 より転載.

び強さを表す勁性 (Hardness, 以下 H) という要素を単位としている。それぞれの子音や母音の調音種とそれらが持つ輝性、勁性の関係は文献 8) において、表 1, 表 2 のように示されている。各表において、B の値は符号付きで表され、正であれば言葉の音の印象が明るく、負であれば暗いことを示し、絶対値の大きさはその度合いを示す。また、H の値は符号なしの数値で表され、数値が大きいほど言葉の音が強いインパクトを持つ。

文献 8) では、音相基から導出される表情を、音相基それぞれが持つ表情である 40 の甲類表情 (表 3) と、いくつかの音相基の組合せによって生じる 38 の乙類表情 (表 4) の 2 種類に分類し、音相による表情を表 5 に示すように 20 の表情語群に分類している。表 3 および表 4 が示すように、甲類表情と乙類表情はそれぞれ、表 5 に示す 20 の表情語群と対応関係が明示されている。表情語群は 40 の表情属性 (印象語に相当) によって構成されている。このように、言葉を構成する音素とその印象について詳細に調査されている。

3.2 任意の言葉を対象とした音韻による MLC の構成方式

本節では、任意の言葉を対象とした音韻の印象による作用素 MLC の構成方式を示す。MLC は、3.1 節の音相理論を専門家の研究として用いて構成する。この作用素により、任意の言葉や未定義・未知の言葉による感性表現による検索などにおけるユーザの入力インタフェースや、印象による商品名の検索などの実現が可能となる。

3.2.1 本方式の全体の流れ

提案方式の概要を図 2 に示す。この方式は、以下の手順により実現される。

前準備: MLC の変換行列 T_1, T_2 の生成 入力された任意の言葉から抽出された音相基を印象語群に変換するための変換行列 T_1, T_2 を構成する。基本的な MLC は 1 つの変換行列で構成されているが、本方式では、文献 7), 8) より、任意の言葉から抽出された音相基から表情を導出するための変換行列 T_1 と表情から表情語群ベクトルに変換するための T_2 を準備する。詳細は 3.2.2.1 で述べる。

Step1: 音相基ベクトルの抽出 入力された任意の言葉から音相基を抽出し、それぞれの値

表 3 甲類表情

Table 3 Individual properties.

表情 (音相基に相当)	抽出される表情語群
1. ア音多用	I, M, T, P
2. イ音多用	A, G, I, K, L, O, P
3. ウ, オ音多用	Q, S, T
4. エ音多用	I, K, M, P
5. 母音種多用	B, M, P, G, E, I, O
6. 母音種過少	C, K
7. 有声音多用	P, Q, S, T
8. 無声音多用	D, I, H, O, R
9. 有声破裂音系多用	Q, R, S, T
10. 無声破裂音系多用	D, F, G, I, J, N
11. 有声摩擦音多用	Q, T
12. 無声摩擦音系多用	I, N, O
13. 鼻音多用	F, M, P
14. 流音多用	D, F
15. 高調音種比	A, B, C, E, F, G, O
16. 低調音種比	K, Q
17. R (長音) 多用	Q, R, T
18. Q (促音) 多用	A, B, C, D, F, J, L, O
19. N (撥音) 多用	Q
20. 濁音多用	P, Q, R, S, T
21. 高動性	A, B, C, D, E, G, H, J, L
22. 低動性	P, Q, T
23. プラス輝性	C, D, F, G, H, I, N, O
24. プラス高輝性	G, H, J, K, M
25. マイナス輝性	Q, R, S
26. マイナス高輝性	Q, R, T
27. 高動輝拍多用	A, D, J, K, S
28. 順接拍多用	A, C, D, E, F, I, M, O
29. 逆接拍多用	P, Q, R, S, T
30. 多拍	P, R, S, T
31. 少拍	A, F, J, L
32. 無声化母音多用	B, C, F, G, H, I, N, O
33. 無声拗音多用	A, B, C, D, E, F, G, H, I, J
34. 子音拍多用	A, B, C, D, F, G, H, I, J, K, O, Q
35. 子音拍過少	R, S
36. 新子音多用	C, G, H
37. 高総合音価	B, C, D, E, G, H, I, M, P
38. 低総合音価	P, Q
39. 濁拗音多用	E, S, T
40. 摩擦音系多用	N, O, P, T

表 4 乙類表情

Table 4 Anaphoric properties.

表情 (甲類表情の組に相当)	抽出される表情語群
1. 甲 1×甲 17	P, T
2. 甲 2×甲 10	G, I
3. 甲 2×甲 27	A, J, K
4. 甲 3×甲 9	Q, S, T
5. 甲 5×甲 8	I, O, P
6. 甲 5×甲 28	E, I, M, O
7. 甲 6×甲 16	K
8. 甲 6×甲 27	K
9. 甲 6×甲 29	C, K
10. 甲 6×甲 37	C
11. 甲 7×甲 16	Q
12. 甲 7×甲 25	Q, S
13. 甲 7×甲 27	S
14. 甲 7×甲 29	P, Q, S, T
15. 甲 7×甲 30	P, S, T
16. 甲 8×甲 12	I, N, O
17. 甲 8×甲 37	D, H, I, P
18. 甲 10×甲 12	I, N
19. 甲 10×甲 18	F, J
20. 甲 10×甲 33	D, F, G, I, J
21. 甲 40×甲 13	P
22. 甲 12×甲 32	I, N, O
23. 甲 13×甲 14	F
24. 甲 15×甲 27	A
25. 甲 15×甲 28	A, B, C, E, F, O
26. 甲 15×甲 32	B, C, F, O
27. 甲 15×甲 37	B, C, E, G
28. 甲 16×甲 27	K
29. 甲 16×甲 29	Q
30. 甲 18×甲 28	A, C, F, O
31. 甲 25×甲 29	Q, R, S
32. 甲 26×甲 29	Q, R, T
33. 甲 27×甲 29	S
34. 甲 28×甲 32	B, C, F, I, O
35. 甲 29×甲 37	P
36. 甲 32×甲 34	B, C, F, G, H, I
37. 甲 32×甲 37	A, B, C, H, I
38. 甲 34×甲 37	B, C, D, G, I

表 5 表情語と表情属性表 (抜粋)

Table 5 The table of impression words and explanation attributes.

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

* 木通隆行著, 小学館スクウェア出版, p.118 表 21 をもとに著者が作成.

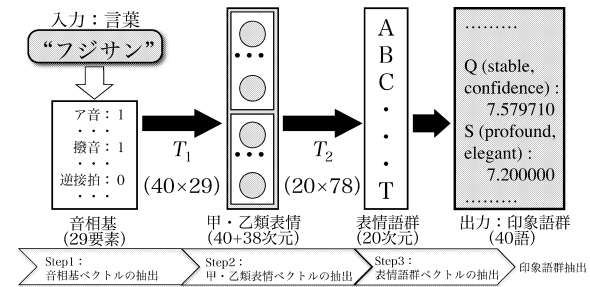


図 2 任意の言葉を対象とした音韻の印象による M_L の全体図

Fig. 2 An overview of operator M_L for arbitrary words by impression of phoneme.

を要素とする音相基ベクトルを抽出する. 詳細は 3.2.2.2 で述べる.

Step2: 甲・乙類表情ベクトルの抽出 音相基ベクトルに変換行列 T_1 を作用させることにより甲類表情ベクトルを導出し, さらにそれらの組合せで表現される乙類表情の各値を要素とする乙類表情ベクトルを導出する. さらに, 甲類表情ベクトルと乙類表情ベクトルを並べ合成した, 甲・乙類表情ベクトルを生成する. 詳細は 3.2.2.3 で述べる.

Step3: 表情語群ベクトルの抽出 甲・乙類表情ベクトルに変換行列 T_2 を作用させることにより, 表情語群ベクトルを抽出する. このベクトルの各要素の値は, 対応する表情語群との相関を表す. さらに表 5 に示す表情語群を印象語に相当する表情属性に展開することにより, 印象語とその重みを導出することができる. 詳細は 3.2.2.4 で述べる.

表 6 音相基群
Table 6 Sound phase elements.

ID	要素名	ID	要素名	ID	要素名
00	A音数	10	無声破擦音数	20	Hの総合音価
01	I音数	11	無声摩擦音数	21	Bの総合音価
02	ウ音数	12	鼻音数	22	高動輝拍数
03	工音数	13	流音数	23	逆接拍数
04	オ音数	14	接近音数	24	拍数
05	母音種数	15	調音種数	25	無声化母音数
06	有声音数	16	R数	26	無声拗音数
07	有声破裂音数	17	N数	27	新子音数
08	無声破裂音数	18	Q数	28	濁拗音数
09	有声破擦音数	19	濁音数		

3.2.2 実現方式

本項では、3.2.1節で示した各処理の詳細を述べる。

3.2.2.1 ML の変換行列 T_1, T_2 の生成(前準備の詳細)

本方式では、任意の言葉を対象とした音韻の印象による ML を実現するために、2つの変換行列 T_1, T_2 を準備する。

変換行列 T_1 は、入力された言葉から抽出される音相基ベクトルを甲類表情に対応した甲類表情ベクトルに変換する。今回使用する音相基は、表6に示す29要素からなる。また、甲類表情は表3に示す40要素からなる。文献8)により、 T_1 の各要素は、各甲類表情についてその音相基が関係する要素の値を1、それ以外の要素を0とすることによって特徴付けられ、 40×29 の行列として構成される。

変換行列 T_2 は、甲類表情と乙類表情からなる甲・乙類表情ベクトルを、印象を表す表情語群ベクトルに変換する。甲・乙類表情ベクトルは、表3で表す40の甲類表情と表4で表す38の乙類表情に対応する78次元のベクトルである。また、表情語群として、表5に示した20の語群が存在する。 T_2 は、列要素には78の表情を、行要素には20の表情語群を配置し、表3および表4から、各表情語群についてその表情が関係する要素を1、それ以外の要素を0とすることによって特徴付けられ、 20×78 の行列として構成される。

3.2.2.2 音相基ベクトル v_b の抽出(Step1の詳細)

入力された任意の言葉から表6に示す29の音相基を抽出し、それらを要素として構成される29次元の音相基ベクトル v_b を構成する。

$$v_b = (v_{b1}, v_{b2}, \dots, v_{b29})^T. \quad (3)$$

これは、言葉というメディアデータから抽出される、言葉の音という観点での特徴を表す構造要素、つまり音相基の特徴量からなるベクトルである。

3.2.2.3 甲・乙類表情ベクトルの抽出(Step2の詳細)

まず、音相基ベクトル v_b に変換行列 T_1 を作用させ甲類表情初期ベクトル v'_k を導出する。

$$v'_k = T_1 v_b = (v'_{k1}, v'_{k2}, \dots, v'_{k40})^T. \quad (4)$$

甲類表情初期ベクトル v'_k は、表3で表す40の甲類表情に対応する要素からなる。その値は、各甲類表情の出現数を表しているにすぎないため、これらの要素の値を、表3の甲類表情の項目に示されているとおり、「多用」、「過少」の度合いを示す値に変換する必要がある。変換された甲類表情の値 v_{ki} からなる甲類表情ベクトル v_k を次のように生成する。

$$v_k = (v_{k1}, v_{k2}, \dots, v_{k40})^T. \quad (5)$$

文献8)では、表7のような各音相基の標準使用率・標準値が示されている。これらの表から次のように場合分けし、各甲類表情の値 v_{ki} をそれぞれ計算する。

- 「多用」していることに着目する甲類表情

($i=1, 2, 3, 4, 7, \dots, 14, 17, \dots, 20, 27, \dots, 34, 36, 39, 40$ の場合)

文献8)によれば、音相基が「多用」していることの基準として、標準使用率の+15%であることがあげられている。そこで、本稿では、 N 拍の音を持つある言葉から抽出される v'_k の i 番目の要素を v'_{ki} 、および表7の音相基の標準使用率を avg_i とした場合の、その甲類表情の値 v_{ki} を以下の式により計算する。

$$v_{ki} = \begin{cases} 0 & \left(\frac{|v'_{ki}|}{N} - avg_i \leq 0 \right), \\ \left(\frac{|v'_{ki}|}{N} - avg_i \right) \times \frac{1}{0.15} & \left(0 < \frac{|v'_{ki}|}{N} - avg_i < 0.15 \right), \\ 1 & \left(\frac{|v'_{ki}|}{N} - avg_i \geq 0.15 \right). \end{cases} \quad (6)$$

ここでは、音相基の「多用」の基準がその音相基の使用率が標準使用率の+15%であることから、音相基の使用率が標準使用率と比べて0.15以上であるならば1、使用率が標準使用率以下の場合は0とし、それ以外の場合は0から1の間の値で特徴づけられる。ここで、使用率が15%を超えた場合の値を1に固定しているのは、予備実験と

表 7 各音相基の標準使用率と標準値

Table 7 Typical usage rates and typical values for sound phase elements.

音相基	標準値	音相基	標準値
ア音	32%	促音	1%(*)
イ音	20%	撥音	1%(*)
ウ・オ音	38%	濁音	10%(*)
工音	16%	高勁輝拍	0.8%
有声音	53%	順接拍	77%
無声音	47%	逆接拍	23%
有声破裂音系	10%(*)	無声化母音	12%
無声破裂音系	10%(*)	無声拗音	8%
有声摩擦音	10%(*)	子音拍	28%
無声摩擦音系	10%(*)	新子音	12%
鼻音	10%(*)	濁拗音	12%
流音	10%(*)	摩擦音系	26%
長音	1%(*)		

(*) 文献 8) では値が明示されていないため、本稿では例語から類推した値を用いている。

音相基	拍数	標準値
母音種	1~2	判定対象外
	3	3 種以上で「多用」、1 種以下で「過少」
	4	3 種以上で「多用」、1 種以下で「過少」
	5~7	4 種以上で「多用」、2 種以下で「過少」
	8~10	「多用」は判定対象外、3 種以下で「過少」
	11 以上	「多用」は判定対象外、4 種以下で「過少」
調音種比	1	判定対象外
	2	3 種以上で「高」、1 種以下で「低」
	3~4	4 種以上で「高」、2 種以下で「低」
	5~6	5 種以上で「高」、3 種以下で「低」
	7~9	6 種以上で「高」、4 種以下で「低」
	10~14	7 種以上で「高」、5 種以下で「低」
	15 以上	「高」は判定対象外、7 種以下で「低」
勁性	1	H = 1.0
	2	H = 1.5
	3	H = 2.0
	4 以上	H = (拍数) × 0.6 - 0.1
輝性	1	B = ±0.1
	2	B = ±0.4
	3	B = ±0.6
	4 以上	B = ±(拍数) × 0.6 - 1.5
多拍・少拍	8 拍以上は多拍、3 拍以下は少拍	

* 木通隆行著、小学館スクウェア出版、p.132-133 表 23 をもとに著者が作成。

して値の上限設定の有無で印象の抽出結果の比較を行った結果、値の上限を設定しない場合において、標準使用率の高さに依存して一部の音相基に対応する印象語群の重みのみが極端に高くなりやすくなる傾向が確認されたためである。

- 「過少」であることに着目する甲類表情 ($i = 35$ の場合)

文献 8) によれば、音相基が「過少」であることの基準として、標準使用率の -15% であることがあげられている。そこで、本稿では、「多用」の場合と同様、 N 拍の音を持つある言葉から抽出される v'_k の i 番目の要素を v'_{ki} 、および表 7 の音相基の標準使用率を avg_i とした場合の、その甲類表情の値 v_{ki} を以下の式により計算する。

$$v_{ki} = \begin{cases} 1 & \left(\frac{|v'_{ki}|}{N} - avg_i \leq -0.15 \right), \\ \left| \left(\frac{|v'_{ki}|}{N} - avg_i \right) \times \frac{1}{0.15} \right| & \left(-0.15 < \frac{|v'_{ki}|}{N} - avg_i < 0 \right), \\ 0 & \left(\frac{|v'_{ki}|}{N} - avg_i \geq 0.15 \right). \end{cases} \quad (7)$$

ここでは、音相基の「過少」の基準がその使用率が標準使用率の -15% であることから、音相基の使用率が標準使用率と比べて -0.15 以下であるならば 1、使用率が標準使用率以上の場合は 0 とし、それ以外の場合は 0 から 1 の間の値で特徴づけられる。

- 母音種、調音種比、勁性、プラス/マイナス高輝性、多拍/少拍

($i = 5, 6, 15, 16, 21, 22, 24, 26, 30, 31$ の場合)

これらの項目には、表 7 下部のような対応表が別途示されている。 N 拍の音を持つある言葉から抽出される v'_k の i 番目の要素を v'_{ki} とすると、拍数 N に該当している箇所と比較し、 v'_{ki} の値が条件を満たせば $v_{ki} = 1$ 、満たさない場合は $v_{ki} = 0$ と特徴づける。

- プラス/マイナス輝性 ($i = 23, 25$ の場合)

これらの項目については、文献 8) により、値の符号のみ着目する。プラス輝性の場合 ($i = 23$)、 v'_{ki} が正であれば $v_{ki} = 1$ 、それ以外は $v_{ki} = 0$ と特徴づける。マイナス輝性の場合 ($i = 25$)、 v'_{ki} が負であれば $v_{ki} = 1$ 、それ以外は $v_{ki} = 0$ と特徴づける。

- 高/低総合音価 ($i = 37, 38$ の場合)

文献 8) によれば、高総合音価とは勁性と輝性の両者の絶対値が大きいことを指し、低総合音価とは勁性と輝性の両者の絶対値が小さいことを指す。そのため、これらの値は、変換行列 T_1 からは導出されず、上記で導出した、高/低勁性、プラス/マイナス高

輝性, つまり v_{ki} ($i = 21, 22, 24, 26$) の値を用いて導出する必要がある. 本稿では, 上記の定義に従い, 予備実験として, 2つの要素の値の積を用いる方式と, 双方の値が正であれば一律で1を与える方式を用いて比較実験を行い, 値を1に固定した場合は2つの要素のいずれか一方の値のみの影響が大きくなってしまふことを確認した. そこで, 本稿では, 以下の式により高/低総合音価の値 v_{k37}, v_{k38} を導出する.

$$v_{k37} = \sqrt{\max(v_{k24}, v_{k26})} \times v_{k21}. \quad (8)$$

$$v_{k38} = \sqrt{\frac{1}{\min(v_{k24}, v_{k26})}} \times v_{k22}. \quad (9)$$

v_{k37} は, プラス/マイナス高輝性 v_{k24}, v_{k26} の大きい方どちらかと高動性 v_{k21} の積により, 動性と輝性の両者の絶対値の大きさを示している. v_{k38} は, プラス/マイナス高輝性 v_{k24}, v_{k26} の小さい方どちらかの逆数と低動性 v_{k22} の積により, 動性と輝性の両者の小ささを示している. 逆数をとる理由は, 低総合音価は上記の定義により, v_{k24}, v_{k26} の値の小さい方が小さければ小さいほど大きな値を付与するためである.

以上により, 甲類表情ベクトル v_k の各要素の値 v_{ki} が決定される.

また, 各乙類表情は甲類表情の組合せで構成されるため, 乙類表情ベクトル v_l は下記のように表される.

$$v_l = (v_{l1}, v_{l2}, \dots, v_{l38})^T. \quad (10)$$

$$v_{lk} = \begin{cases} v_{ki_1} v_{ki_2} & \text{if } v_{ki_1} v_{ki_2} \geq 1, \\ 0 & \text{if } v_{ki_1} v_{ki_2} < 1. \end{cases}$$

ここで, v_{ki_1}, v_{ki_2} は i 番目の乙類表情を構成する甲類表情の組のそれぞれの値を示している. たとえば $i = 1$ の場合, 表4から, 1番目の乙類表情は1番目の甲類表情と17番目の甲類表情の組合せとなる. すなわち, $i = 1$ の場合の $v_{ki_1} v_{ki_2}$ の値は, 式(6)で表される v_{k1} と v_{k17} の積で表される. また, 予備実験として, v_{lk} の各値に対し, 値が1未満であった場合にそのまま採用する方式(下限なし)と一律で0とする方式を用いて実験結果の比較を行ったところ, 下限なしの場合ではその乙類表情を構成する2つの甲類表情の値がともに小さい場合にも印象が付与され, これが検索実験におけるノイズとなることが確認された. 1未満を一律に0とするのは, このノイズを排除するためである.

最後に, 甲類表情ベクトル v_k と乙類表情ベクトル v_l を用いて, 78次元の甲・乙類表情ベクトル v_f を生成する.

$$v_f = (v_{k1}, v_{k2}, \dots, v_{k40}, v_{l1}, v_{l2}, \dots, v_{l38})^T. \quad (11)$$

3.2.2.4 表情語群ベクトルの抽出 (Step3の詳細)

前項の v_f に変換行列 T_2 を作用させることにより, 表情語群ベクトル v_o を導出する.

$$v_o = T_2 v_f \quad (12)$$

$$= (v_{o1}, v_{o2}, \dots, v_{o20})^T. \quad (13)$$

これにより, v_f が印象を表す表情語群ベクトル v_o に変換される. v_o の各要素は, 表5の表情語群 A~T との相関を表している. また, 表5を用いて表情語群から表情属性に展開することにより, 入力した言葉の音の印象を表す重み付き印象語を得ることができる.

4. 本作用素の感性検索への適用方式

本章では, 本作用素の適用例として, 感性検索への適用方式を示す. これにより, ユーザの任意の言葉や未定義・未知の言葉による感性表現が可能となり, ユーザへの負荷が少ないインタフェースが実現される. メディアデータに付与されたメタデータのマッチングだけでなく, オノマトベや未知語などニュアンスを表した言葉を用いた検索が実現されることにより, システムに対する感性表現の幅が広がり, ユーザの意図と合致した検索が可能になる.

本稿では, 実験システムとして, 本方式を有する楽曲 MIDI データを対象とした感性検索システムのプロトタイプを実現する. この全体図を図3に示す. 本システムは, 文献9), 11), 12) に示された, 楽曲メタデータ自動抽出方式を用いた楽曲検索に, 本提案方式である任意の言葉を対象とした音韻の印象による作用素 ML が加わったものとみることができる.

本システムは, 本提案方式, 楽曲メタデータ自動抽出方式, 印象語相関計算機構, 正規化機構からなる. 本提案方式では, ユーザによって入力された任意の言葉からその言葉の音によってイメージされる印象を印象語として抽出する. 楽曲メタデータ自動抽出方式は, 楽曲 MIDI データからその楽曲の印象を重み付き印象語で表現したメタデータを抽出する. これは文献9), 11), 12) で示している. ここでは, 音楽心理学者 Hevner による, 表8に示すような C1~C8 の8つの印象語群それぞれに -1以上1以下の大きさの重みを付与されたものがその楽曲の印象を表すメタデータとして付与される. これにより, 聴取者が楽曲に対していただく大まかな印象を表現することが可能である.

印象語相関計算機構は, 印象語で表現された楽曲のメタデータとユーザによって入力された言葉から抽出された印象語の関係を計算することで, その言葉の音から想起される印象に合致した楽曲 MIDI データの検索を実現する. ゆえに本機構では, 印象語どうしの関係を計算する必要がある. これを実現するために, 語どうしの関係を評価するための手法として

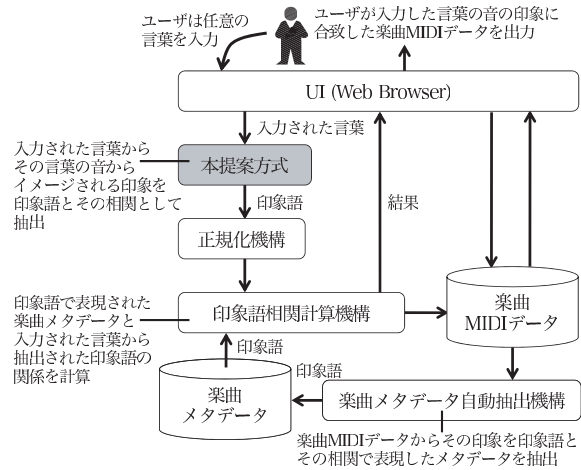


図 3 任意の言葉を対象とした音韻の印象による ML を用いた感性検索 (本方式の適用例)

Fig. 3 An overview of Kansei retrieval utilizing operator ML for arbitrary words by impression of phoneme (One of the examples of application for our method).

表 8 楽曲メタデータ自動抽出方式^{9),11),12)} の印象語群

Table 8 Categories of impression words for automatic music metadata extraction^{9),11),12)}.

	印象語群の表す性質	印象語
c1	真面目な, 神聖な	awe-inspiring, dignified, lofty, sacred, serious, sober, solemn, spiritual
c2	悲しい, 沈んだ	dark, depressing, doleful, frustrated, bloomy, heavy, melancholy, mournful, pathetic, sad, tragic
c3	感傷的な, 夢見るような	dreamy, longing, plaintive, sentimental, tender, yearning, yielding
c4	平穏な, 穏やか	calm, leisurely, lyrical, quiet, satisfying, serene, soothing, tranquil
c5	滑稽な, 優雅な	delicate, fanciful, graceful, humorous, light, playful, quaint, sprightly, whimsical
c6	楽しい, 幸せな	bright, cheerful, gay, happy, joyous, merry
c7	情熱的な, 鼓舞するよ うな	agitated, dramatic, exciting, exhilarated, impetuous, passionate, restless, sensational, soaring, triumphant
c8	堂々たる, 断固たる	emphatic, exalting, majestic, martial, ponderous, robust, vigorous

既存の任意の方式 (概念辞書, シソーラス, オントロジ, ベクトル空間モデルなど) を用いることができる. 本実装では, 意味の数学モデル^{14)–16)} を用いて実装を行った.

正規化機構は本提案方式から出力された印象語とその相関について, 相関の正規化を行う機構である. 本提案方式から導き出された印象語とその相関では, たとえば, “bright, open-minded” という印象は付与されるがその逆にあたる “dark, exclusive” といった印象はどんな音韻の入力に対しても付与されない, というようにポジティブな印象に正の重みが付与される一方, ネガティブな印象の表現が不足している. また, 音相理論において, たとえば表 4 “乙類表情” において印象語群 L と対応する表情が 1 つも存在しないなど, 抽出される印象の出現の傾向に偏りがあるため, 入力される語によっては非常に大きい値が付与されたり, 相関の小さい印象語にも正の相関が付与されて抽出されたりする傾向があり, これが, 検索のノイズとなる可能性がある. そこで本稿では, 音相から選ばれた印象語の重みを, 最小が -1 , 最大が 1 となるような正規化を行う. 印象語ベクトル v_o の要素を v_{o_i} としたときの, 正規化後の要素 v'_{o_i} を以下の式により定義する. ただし, $\max(v_o)$, $\min(v_o)$ は, それぞれ v_o の要素のうち最大, 最小のものを表す.

$$v'_{o_i} = \frac{2v_{o_i} - (\max(v_o) + \min(v_o))}{\max(v_o) - \min(v_o)}. \quad (14)$$

これにより, 中間の値が 0 となり, 小さな相関の印象語, もしくは関係のない印象語がマイナスの値になり, ノイズを緩和することが可能となる. 本実装では上記の式を用いたが, これ以外にも正規化方式は適用可能である.

本実験システムでは, オノマトペや未知語, 未定義語など言葉の音韻の印象が重要となる言葉が問合せとして利用可能となる. これにより, ユーザの任意の言葉や未定義・未知の言葉による感性表現が可能となり, ユーザへの負荷が少ないインタフェースが実現される.

5. 実験

4章で示した実験システムを構築し, 実験を行った. 5.2節では, 図3で示す“本提案方式”部のみを用いて, 実際にユーザが入力した言葉からどのような印象語が抽出されるかを定性的に評価することにより, 本方式の有効性を検証する. さらに, 5.3節では, 図3のシステム全体を用いて, 本方式を実際の楽曲 MIDI データを対象とした感性検索に適用することにより, 他のシステムへの適用における有効性について検証する.

5.1 実験環境

本実験では, 4章で示した図3のシステムの全体, または一部を用いて実験を行った.

表 9 実験 2 に用いる楽曲 MIDI データリスト
Table 9 A list of MIDI data for experiment-2.

<p>大学祝典序曲, おんまはみんな, アメージング・グレース, アニーローリー, 四季より「秋」, 茶色の小瓶, さらばジャマイカ, 今日の日はさようなら, 河は呼んでいる, カノン, 教会にて..., クラリネットを壊しちゃった, かわいいあの子, 黒い瞳, 枯葉, ダイアナ, はじめてのアドリブ, ホフマンの舟歌, 浜辺の歌, 庭の干し草, 四季の歌, グリーン・スリーブス, 聖しこの夜, ゴゼックのガボット, ユー・アー・マイ・サンシャイン, ウィリアム・テル序曲, はにゅうの宿, 幸せなら手を叩こう, 交響曲第 3 番 (ブラームス), 天国と地獄, 交響曲第 40 番 (モーツアルト), 新世界より「家路」, ハンガリア舞曲第 5 番, ジャンバラヤ, ラ・カンパネラ, ぼだい樹, ロンドン橋, ローレライ, 愛のロマンス, 太陽がいっぱい, 思い出, 漕げよマイケル, 月光, マイ・ボニー, プレリュード, サバの女王, 雨の日曜日, 追憶, 聖者の行進, ウォルシング・マチルダ</p>

それぞれの印象語は英語で抽出される。これは、本実装では、印象語相関計算機構において、意味の数学モデル^{14)–16)}を用いて実装しているが、印象語どうしの相関を計算するための元データとして“Longman Dictionary of Contemporary English”¹⁷⁾という英英辞書を使用しているためである。同辞書は約 2,000 語の基本語だけで約 56,000 語の見出し語を説明している。そのため、56,000 もの言葉を約 2,000 次元の正規直交空間で表現可能である。これにより、英語で表された様々な印象語についての相関計算が可能となる。

5.3 節で用いる検索対象となる楽曲 MIDI データとして、表 9 に示す、比較的有名で印象が分かりやすい 50 曲を用意した。各楽曲には、図 3 中の楽曲メタデータ自動抽出方式^{9),11),12)}によって、印象語群が抽出されている。

5.2 実験 1: 印象語の抽出実験

提案方式が任意の語において適用可能であることを示すため、感性検索において入力される語のパターンを想定し、それらに対して 4 章で示した図 3 のシステムの一部である“本提案方式”部を用いて、提案方式の定性的な検証を行った。

5.2.1 実験方法

提案方式により抽出される印象語の定性的な検証を行った。まず、5.2.2 項に示すように、感性検索において入力想定される語のパターンの分類および各パターンにおける提案方式の有効性に関する考察を行った。そして、各パターンごとに印象語の抽出実験を行った。

5.2.2 入力語に関する考察

感性検索において入力想定される語として、ユーザやキャラクタなどの人名、商品名、組織名といった名詞や、検索対象に対する印象や微妙なニュアンスなどを表現した形容詞、副詞、オノマトペがあげられる。そこで本稿では、入力語を、IPA 品詞体系を参考に下記のように分類し、この分類に基づいて考察を行った。

- 固有名詞 人名, 地名, 組織名, 商品名など, 命名対象の印象が明確であると考えられ, 提案方式による印象抽出が有効に働くと考えられる。
- 一般名詞 固有名詞と同様であるが, 方向, 数, 代名詞など, 意味や記号などの表現に用いられる一部の語については, 提案方式が必ずしも有効ではないと考えられる。
- 形容詞・動詞 状態や様子を表すことから明確な印象を持ち, 直感的に印象をイメージできるような音を持つ語も多い。これらの語は従来の方式においても, 語の辞書的な意味に基づく方式や被験者調査に基づく方式などにより, 感性検索に用いられている。
- 副詞・擬音語・擬態語 印象が明確である半面, ユーザに依存して多様な表現が存在し, 辞書などを用いた網羅的な利用が困難である。日常会話などにおいても微妙なニュアンスの表現に多く用いられ, 提案方式による印象抽出がきわめて有効であると考えられる。
- 未知語 新商品の名称やユーザが直感的に思いついた表現など, 既存の辞書や用語集などに収録されていないものに対しては, 感性検索での利用自体が困難である。このような場合, 提案方式により大まかな印象をとらえることはきわめて有効であると考えられる。
- 文章 文章や詩の印象に合致した曲や挿絵などを検索するような場合に入力が想定される。しかしながら, 提案方式を文章など拍数の多い入力語に適用する場合, 多くの印象語群に同程度の重みがついてしまったり, 接続詞や助詞などに意図と異なる印象が含まれてしまったりするなど, 印象を適切に表現するのが困難となることが想定される。

以上から、本提案方式は、ユーザが感性や印象を短いフレーズを用いて直感的に表現する場合に特に有効であると考えられる。一方、長いフレーズや文章を入力とする場合、特定の印象を抽出しやすい音韻を意図的に多用するなど、ユーザに対し提案方式に関する深い知識が要求されることとなり、自身の感性を反映させるのが困難となりうる。したがって、本稿では短いフレーズを入力として想定し、文章などへの適用に関しては今後の課題とする。

5.2.3 実験結果

前項での分類に基づき、一般名詞から“ゴージャス”、“うえ(上)”、“した(下)”, 固有名詞から商品名の例として“ルイピトン”および人名の例として“ジャッキーチェン”, 形容詞・動詞から“あかるい(明るい)”, 副詞・擬音語・擬態語から“さっぱり”および“ドロドロ”, そして未知語の例として、意味が分からない語である“ゴッゴル”および“ひんみば”を入力した場合の出力をそれぞれ示す。

入力した言葉が“ゴージャス”の場合の抽出された印象語とその相関が上位のものについて

表 10 “ゴージャス”での抽出された印象語とその相関
Table 10 Impression words with correlation of “gorgeous”.

印象語	相関		印象語	相関	
	正規化なし	正規化あり		正規化なし	正規化あり
elegant	14.066667	1.000000	stable	14.000000	0.394615
profound	14.066667	1.000000	inactive	12.866667	0.150877
confidence	14.000000	0.394615	static	12.866667	0.150877

表 11 “うえ”での抽出された印象語とその相関

Table 11 Impression words with correlation of “ue (up)”.

印象語	相関	
	正規化なし	正規化あり
stable, confidence	5.943396	1.000000
profound, elegant	4.943396	0.663492
plain, obvious	4.000000	0.346032
light, trippingly	4.000000	0.346032
static, inactive	3.943396	0.326984

表 12 “した”での抽出された印象語とその相関

Table 12 Impression words with correlation of “shita (down)”.

印象語	相関	
	正規化なし	正規化あり
bright, open-minded	14.000000	1.000000
plain, obvious	10.000000	0.428571
dynamic, active	10.000000	0.428571
young, effervescent	10.000000	0.428571
healthy, clean	9.000000	0.285714

て、表 10 に示す。この表では、本方式から計算された相関と式 (14) の計算を施した相関の両値を示している。文献 8) には、“ゴージャス”の音相の印象としては、「静的」、「高級感」、「安定感」の 3 つの語群が突出して高いポイントであるとして示されている。表 10 では、「静的、非活性的 (inactive, static)」、「高尚な、優雅さ (profound, elegant)」、「安定感、信頼感 (stable, confidence)」など、上記の分析結果に対応した印象語が上位に抽出されていることが分かる。他の文献 8) で分析されている語についても同様に、文献 8) の記述に合致した出力結果が確認できた。これにより、本方式によって、文献 7), 8) を専門家の研究として、言葉の音相に合致した印象語を抽出できたといえる。

入力した語が“うえ (上)”および“した (下)”の場合をそれぞれ表 11, 表 12 に示す。それぞれの語が持つ意味から「高級感」や「庶民的」といった印象語群によるニュアンスの表現が期待されるが、実験結果によれば、“うえ”からは「安定感、信頼感 (stable, confidence)」や「高尚な、優雅さ (profound, elegant)」など言葉の意味と近い表情語群が抽出された一方、“した”からは「明るさ、開放感 (bright, open-minded)」や「健康的、清潔感 (healthy, clean)」などが抽出され、特に“上”の対義語の組として見た場合、そのニュアンスを適切に表現できていない。このような傾向は、数値や方向などを表す語の一部においてみられ、

表 13 “ルイビトン”での抽出された印象語とその相関 (重み)

Table 13 Impression words with correlation (weight) of “Rui-Biton”.

印象語	相関	
	正規化なし	正規化あり
profound, elegant	12.567186	1.000000
stable, confidence	10.293173	0.625134
static, inactive	5.958491	-0.089429
expensive, fulfil	5.922581	-0.095349
warm, comfortable	4.800000	-0.280404

表 14 “ジャッキーチェン”での抽出された印象語とその相関 (重み)

Table 14 Impression words with correlation (weight) of “Jacky-Chen”.

印象語	相関	
	正規化なし	正規化あり
bright, open-minded	11.449134	1.000000
dynamic, active	10.224690	0.720833
young, effervescent	9.012085	0.444365
plain, obvious	7.853868	0.180297
reasonable, real	7.809700	0.170227

これらの語については、提案方式のみではなく、語の意味などを考慮した方式の適用が考えられる。

固有名詞の入力例として、入力が“ルイビトン”の場合の印象語とその相関が上位のものを表 13 に示す。この結果によれば、「高尚な、優雅さ (profound, elegant)」、「安定感、信頼感 (stable, confidence)」、「静的、非活性的 (inactive, static)」、「高級感、充実感 (expensive, fulfil)」などを表す印象語が抽出されていることが分かる。“ルイビトン”は高級ブランドであることから、「高級なもの」「優雅なもの」といったニュアンスを提案方式により表現することが可能であると考えられる。

人名の入力例として、入力を“ジャッキーチェン”とした場合の印象語とその相関が上位のものを表 14 に示す。「明るさ、開放感 (bright, open-minded)」、「動的、活性的 (dynamic, active)」、「若さ、湧刺さ (young, effervescent)」などを表す印象語が抽出されており、カンフー映画で有名な俳優「ジャッキー・チェン」のイメージと合致していると考えられる。一方、俳優の「ジャッキー・チェン」という芸名がそのイメージに合致していることから一般に受け入れられていると考えられ、このことから、命名対象が持つイメージをその名前が後付け的に表現していることを示している。この結果から、提案方式により、明るく動的な印象を入力して感性検索を行う際に、「アクションスターである“ジャッキー・チェン”のようなイメージ」をニュアンスとして入力する、といった利用が可能となると考えられる。

副詞の入力例として、入力を“サッパリ”、“ドロドロ”とした場合の出力をそれぞれ表 15, 表 16 に示す。“サッパリ”は「清潔で整っている」や「あっさりしている」といった様子を表すのに用いられる副詞であるが、実験結果によれば、「明るさ、開放感 (bright, open-minded)」、「動的、活性的 (dynamic, active)」、「健康的、清潔感 (healthy, clean)」および「シンプルな、明白さ (plain, obvious)」といった印象語群が抽出されており、“さっ

表 15 “サッパリ”での抽出された印象語とその相関
Table 15 Impression words with correlation of “Sappari”.

印象語	相関	
	正規化なし	正規化あり
bright, open-minded	12.010717	1.000000
dynamic, active	10.775694	0.769026
healthy, clean	10.518340	0.720895
plain, obvious	10.009967	0.625819
light, trippingly	9.581395	0.545668

表 16 “ドロドロ”での抽出された印象語とその相関
Table 16 Impression words with correlation of “Dorodoro”.

印象語	相関	
	正規化なし	正規化あり
profound, elegant	14.000000	1.000000
stable, confidence	12.000000	0.714286
...		
healthy, clean	0.543478	-0.922360
pure, brisk	0.000000	-1.000000

表 18 “ひんみば”での抽出された印象語とその相関
Table 18 Impression words with correlation of “himmiba”.

印象語	相関	
	正規化なし	正規化あり
plain	1.295348	1.000000
obvious	1.295348	1.000000
individual	1.193938	0.5985367
special	1.193938	0.5985367
reasonable	1.141632	0.494463
real	1.141632	0.494463

表 19 “ゴッゴル”での抽出された印象語とその相関
Table 19 Impression words with correlation of “goggoru”.

印象語	相関	
	正規化なし	正規化あり
elegant	12.333333	1.000000
profound	12.333333	1.000000
confidence	12.133333	0.306238
stable	12.133333	0.306238
inactive	8.133333	0.22199
static	8.133333	0.22199

表 17 “あかるい”での抽出された印象語とその相関
Table 17 Impression words with correlation of “Akarui (bright)”.

印象語	相関 (正規化なし)	
	相関 (正規化なし)	相関 (正規化あり)
bright, open-minded	7.865458	1.000000
healthy, clean	6.966390	0.748554
profound, elegant	6.428572	0.598140
young, effervescent	6.428572	0.598140
plain, obvious	6.059007	0.494782

ぱり”という語が持つ意味とも合致していることが分かる。一方、“ドロドロ”の場合、「高尚な、優雅さ (profound, elegant)」や「安定感、信頼感 (stable, confidence)」など、語の意味や用法とは必ずしも合致しない印象語群が抽出された。これは、音相理論において「重さ」や「鈍さ」のようなネガティブなイメージを持つ印象語群が規定されていないためである。ここで、下位の印象語群に着目してみると、「清らかさ、爽やかさ (pure, brisk)」や「健康的、清潔感 (healthy, clean)」といった印象が位置しており、“正規化機構”を用いてこれらの印象語群に負の重みを与えることにより、語の本来の印象にある程度近づけることが可能となると考えられる。

形容詞の入力例として、入力を“あかるい”とした場合の出力を表 17 に示す。「明るさ」「健康的」「高尚な」「若さ」「明白さ」といった印象の重みが大きくなっており、“明るい”という語の意味と合致した結果であるといえる。なお、ネガティブな印象を持つ形容詞に関しては、副詞の場合と同様の傾向が見られ、正規化を行い下位の印象語群に対して負の重みを与えることでより印象に近い出力が期待できる。

最後に、未知語の例として、意味が分からない語である“ひんみば”および“ゴッゴル”を入力とした場合について示す。

入力した言葉が“ひんみば”の場合の抽出された印象語とその相関が上位のものを表 18 に示す。「シンプルな、明白さ (plain, obvious)」、「個性的、特殊的 (individual, special)」、「合理的、現実的 (reasonable, real)」を表す印象語が抽出されている。特に、母音に着目すると、イ音が多く、表 3 の甲類表情から、「シンプルな、明白さ」や「個性的、特殊的」を表す印象語群 A, K の相関が大きくなったと考えられる。

さらに、入力した言葉が“ゴッゴル”の場合の抽出された印象語とその相関が上位のものを表 19 に示す。表 5 と見比べて、「高尚な、優雅さ (profound, elegant)」、「安定感、信頼感 (stable, confidence)」、「静的、非活性的 (inactive, static)」を表す印象語が抽出されている。特に“ゴッゴル”には“g”の音すなわち有声破裂音と、才音が多用されている。表 3 の甲類表情から、有声破裂音から「安定感、信頼感」、「静的、非活性的」を表す印象語群 Q, T が導かれ、才音から「高尚な、優雅さ」、「安定感、信頼感」、「静的、非活性的」を表す印象語群 Q, S, T が導かれ、これらの相関が大きくなったと考えられる。

5.2.4 考 察

これらの結果から、文献 7), 8) を専門家の研究として用いた Media-lexicon Transformation Operator が正しく動作していることが確かめられた。実験結果より、特に固有名詞、形容詞、未知語については適切に印象が抽出されており、感性検索などにおけるユーザの印象の入力のための手段として用いることが可能となると考えられる。

一方、文章に関しては、複数の音韻から複数の印象が抽出されてしまうため、複数の要素の選択や合成などといった処理が必要となるが、本提案方式は単語やフレーズなど短い語を対象としており、文章などの長い入力への対応は今後の課題とする。また、その他の品詞については、抽出される印象とその言葉の意味が必ずしも合致しないなど、提案方式のみでは十分ではない場合があることが確認された。このような場合、言葉の 3 要素⁶⁾、すなわち本

方式において扱う“サウンド性”に加え，“簡明性”，“視覚性”の3要素を組み合わせて印象をとらえる方式の実現が重要となる．本提案方式により，言葉の音素情報のみを用いてその印象の抽出を実現するため，未定義語や未知語にも柔軟に対応でき，辞書による定義が困難な言葉にもその言葉の印象から喚起される印象によるメタデータを扱うことが可能となる．

今後，これらの方式を検証するためには，エンドユーザによるアンケートなどの調査実験が必要になると考えられるが，このシステムのみでは，非常に難しいと考えられる．理由として，言葉の音のみの印象を分離して感じることは非常に難しいからである．ネーミングにおいて“簡明性”，“視覚性”，“サウンド性”が重要だといわれている⁶⁾が，これは言葉一般についてもいえると考えられる．本システムは，言葉の“サウンド性”を重視した作用素である．人間にとっての言葉の印象は，この3要素から決まると考えたとしても，そのうちの“サウンド性”のみをクリアに評価するのは難しいと予想される．言葉は，他のメディアに比べ，それぞれの要素がそれぞれに引っ張られやすいメディアであると考えられる．また，実験でも用いた，“ゴッゴル”は，2004年に開催された初回のSEOコンテスト¹⁸⁾という検索エンジン最適化を競うコンテストでテーマとしてつくられた，まったく意味を持たない語である．しかしながら，そのコンテスト後においても，Googleにおいて2009年7月現在，約57,000件ヒットすることから，当初，未知語として音の印象のみで受け取られていたものが，それ以外の印象が後付け的に広がることを意味している．

これらのことから，本方式を“サウンド性”に特化した基本機能として位置づけ，言葉の他の要素である，“簡明性”，“視覚性”についてそれぞれの基本機能を実現し，それを統合し，言葉の印象を総合的に抽出するシステムの実現が今後重要となる．そのシステムをアンケートなどの調査実験をすることにより，より有効性が検証できるものと考えられる．これらの基本機能実現，統合方式実現，その調査実験については今後の課題とする．

5.3 実験2：感性検索への適用による本方式の有効性の検証

5.3.1 実験方法

4章で示した図3のシステム全体を用いて，本方式が感性検索の入力インタフェースとして適用することについての有効性の検証を行った．

本実験において，入力する語とそれらに対応する正解楽曲を表20のように設定する．まず，入力される語として，文献8)において音相理論における各要素を決定するための調査に用いられた語として紹介されている8つの語群を用いることとした．これらは，表20に示すように，それぞれの言葉の用法に基づき，印象の明暗を基準に「暖かさ，和やかさ」「美しさ，爽やかさ」「明白さ，単純さ」「活性，行動」「不快，苦痛，哀歎」「不安，暗さ」「非

表20 入力として用いる語群とその正解楽曲データの対応
Table 20 8 groups of input words and answer music data.

性質	入力する言葉	正解楽曲
暖かさ，和やかさに関する語	“あっぱれ”，“あいらしい”など106語	印象語群 C4, C5, C6 の重みが最も大きい24曲（“庭の干し草”，“新世界より「家路」”，“ぼだい樹”，“ローレイ”，“カノン”，“ホフマンの舟歌”，“浜辺の歌”など）
美しさ，爽やかさに関する語	“うっとり”，“あでやか”など72語	印象語群 C5, C4, C6 の重みが最も大きい24曲（上記に同じ）
明白さ，単純さに関する語	“あからさま”，“あきらか”など167語	印象語群 C6 の重みが最も大きい17曲（“ウィリアム・テル序曲”，“アニー・ローリー”，“ゴゼックのガボット”，“はじめてのアドリブ”，“ダイアナ”など）
活性，行動に関する語	“いきいき”，“いさまし”など164語	印象語群 C7 および C8 の重みが最も大きい3曲（“枯葉”，“ハンガリア舞曲第5番”，“交響曲第40番（モーツァルト）”）
不快，苦痛，哀歎に関する語	“くよくよ”，“とぼとぼ”など138語	印象語群 C1 の重みが最も大きい13曲（“アメージング・グレース”，“今日の日はさようなら”，“河は呼んでいる”，“教会にて”，“ジャンバラヤ”，“ロンドン橋”など）
不安，暗さに関する語	“おどおど”，“どんより”など243語	印象語群 C2 の重みが最も大きい5曲（“黒い瞳”，“サバの女王”，“交響曲第3番”，“追憶”，“月光”）
非活性，非行動性に関する語	“まごまご”，“のろろ”など173語	印象語群 C3 の重みが最も大きい5曲（“グリーン・スリープス”，“ラ・カンパネラ”，“愛のロマンス”，“太陽がいっぱい”，“四季の歌”）
否定，抗争に関する語	“あてつけ”，“ぎすぎす”など225語	印象語群 C2 の重みが最も大きい5曲（“黒い瞳”，“サバの女王”，“交響曲第3番”，“追憶”，“月光”）

活性，非行動性」「否定，抗争」の8つに分類されている．そして，これら8つの語群と，表8に示された8つの楽曲印象語群を対応させることにより，各入力語に対する正解楽曲を表20のように設定した．本稿では，事前に提案方式および文献9)，11)，12)の楽曲メタデータ抽出方式に関する知識を持った被験者3名に語群と正解楽曲の対応付けを行ってもらい，全員の合意がとれたものを正解として設定した．

また，検索対象としては，一般に印象がよく知られている有名な楽曲50曲のMIDIデータを用いることとした．各楽曲は文献9)，11)，12)の楽曲メタデータ抽出方式を用いて特徴づけられており，対応する楽曲印象語群の相関（重み）が最も高い楽曲を，それぞれの楽曲印象語群に対する正解楽曲と設定した．

以上のように入力される語と正解楽曲を設定して検索実験を行い，再現率・適合率を求めた．また，4章の式(14)に示した正規化機構を用いた場合とそうでない場合でそれぞれ検索実験を行い，さらに正規化方式の比較対象として，重みの絶対値が小さな表情語群をノイ

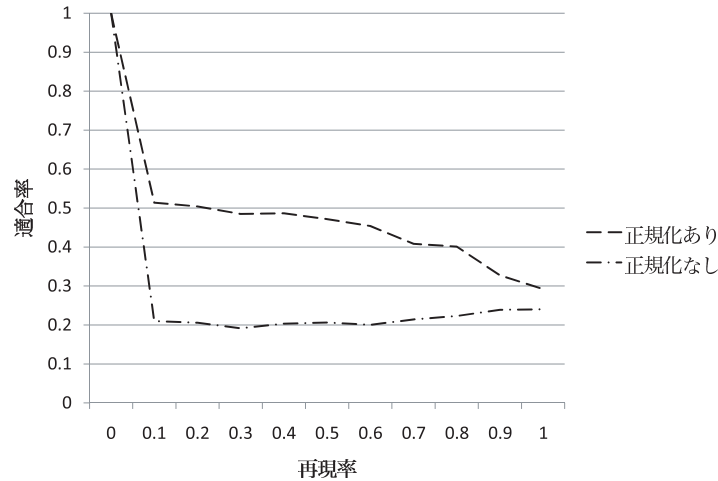


図 4 正規化の有無による再現率・適合率の比較

Fig. 4 Recall rates and precision rates comparison of normalization and non-normalization of correlation.

ズとして排除する方式に関しても同様に実験を行い、各結果の適合率・再現率を用いて有効性の検証を行った。

5.3.2 実験結果

まず、4章で示した正規化機構の有効性を示すため、“本提案方式”部の出力そのままの相関と表情語群の場合と式(14)によって正規化した相関と表情語群の場合とを検索結果によって比較する。さらに、重みの絶対値の小さな表情語群がノイズとなる可能性を検証するため、正規化を行った後に重みの絶対値が小さくなったいくつかの表情語群をカットして同様の実験を行って結果の比較を行った。正規化した場合としない場合の再現率・適合率による比較をしたグラフを図4に示す。同図において、“正規化あり”は正規化機構を用いて実験を行った場合を、“正規化なし”は正規化機構を用いず、“本提案方式”部による出力をそのまま用いて実験を行った場合をそれぞれ示している。また、各再現率における適合率は、それぞれの場合において表20に設定されたすべての入力に対する適合率の平均の値となっている。同図によれば、正規化機構を用いた場合の適合率の平均が、すべての再現率において正規化なしの場合を大きく上回っていることが分かる。この結果から、適切な正規化方式を適用することによって、本方式である任意の言葉からその言葉が持つ音韻の印象の自動抽

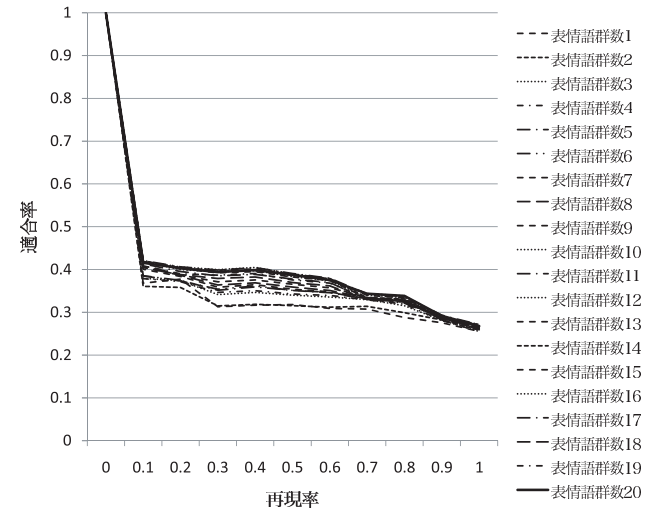


図 5 重みの小さな表情語群を除いた場合の各入力語群における適合率の平均

Fig. 5 Average precision rates of normalization in case of removing impressions having small weight.

出を実現する作用素が感性検索に有効に利用できることを示している。

次に、相関の絶対値の小さな表情語群を除去した場合の実験を行い、検索精度の比較を行った。本稿では、20の印象語群のうち相関の絶対値の小さい方から1, 2, 3, ..., 19個のそれぞれについて除去を行い、表情語群の除去を行わなかった場合と比較した。その結果は図5に示すとおりである。図5により正規化のみを行った場合と正規化後に表情語群の除去を行った場合の結果を比較したところ、表情語群数が1や2の場合に極端に検索精度が低下した点を除けば、各場合の結果に有意な差は見られなかった。この結果から、すべての語に対して一律で表情語群を除去する方式では、提案方式の精度改善は困難であるといえる。

以上の実験により、本稿においては、正規化機構として式(14)による正規化方式のみを用いることとする。より優れた正規化方式の検討は今後の課題とする。

一方、表4において、正規化機構を用いた場合であっても、再現率が上がるにつれて大きく適合率が低下している。そこで、正規化機構を用いた場合について、表20に示された8つの入力語群それぞれに対する平均適合率・再現率の比較を行った。この実験結果を図6

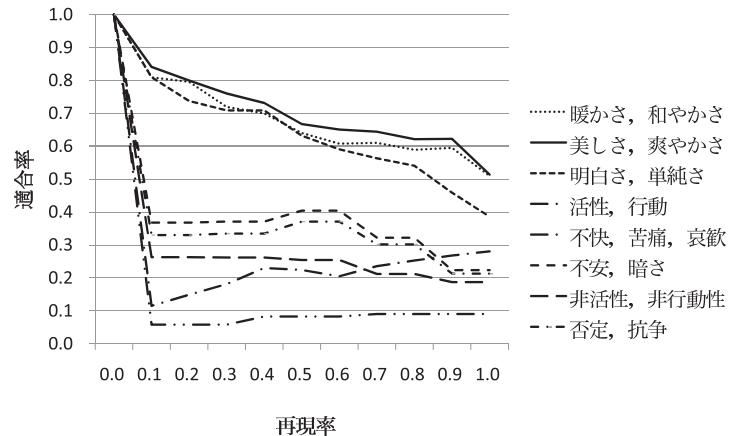


図 6 正規化ありの場合の各入力語群における適合率の平均
Fig. 6 Average precision rates of normalization.

に示す。これによれば、「暖かさ, 和やかさ」、「美しさ, 爽やかさ」および「明白さ, 単純さ」に関する語に対する検索精度が高い一方、「不快, 苦痛, 哀歎」などの暗い印象を持つ語に対してはどれも検索精度が低いという結果になった。

これは、音相理論において明確に規定が可能な表情語群の種類や、それらの出現頻度に偏りがあるためであると考えられる。正規化機構により明るい印象の表情語群にマイナスをつけることで暗い印象の表現をある程度可能としているが、それだけでは不十分であり、より多くの印象に対応可能な方式への改善などが必要になると考えられる。一方、「活性, 行動」に関する語に対する適合率が非常に低い結果となった。これは、入力される語として“いろめく”, “ふるいたつ”など様々な印象のものを含んでいたため正解設定が困難であったこと、および「活性, 行動」を明確に表す印象語群が楽曲のメタデータとして付与されていないことなどが理由である。

このように、一部の入力語群に関して検索精度が低下する原因として、この方式だけでは、本稿の正規化によって緩和はされているものの、印象としてあまり関係のない表情語群にも小さい相関(重み)が付与され、それがノイズになる可能性があることがあげられる。例として、「活性, 行動」に関する語として入力に用いられた“ピチピチ”の表情語抽出結果例を表 21 に示す。入力した言葉“ピチピチ”から、上位に「現代的, 都会的(modern, urban)」をはじめ、「若さ, 湧刺さ(young, effervescent)」、「明るさ, 開放的(bright, open-minded)」

表 21 “ピチピチ”から抽出された印象語
Table 21 Extracted impression words from “pichi pichi”.

印象語	相関		印象語	相関	
	正規化前	正規化後		正規化前	正規化後
modern	5.229970	1.000000	adaptable	2.412683	-0.077362
urban	5.229970	1.000000	popular	2.412683	-0.077362
vibrant	4.162457	0.591770	florid	2.032071	-0.222912
advance	4.162457	0.591770	bustle	2.032071	-0.222912
reasonable	3.866765	0.478694	powerful	1.779337	-0.319561
real	3.866765	0.478694	sharp	1.779337	-0.319561
young	3.839034	0.468090	healthy	1.694589	-0.351969
effervescent	3.839034	0.468090	clean	1.694589	-0.351969
plain	3.769567	0.441525	light	1.505773	-0.424175
obvious	3.769567	0.441525	trippingly	1.505773	-0.424175
bright	3.760254	0.437963	warm	1.113149	-0.574319
open-minded	3.760254	0.437963	comfortable	1.113149	-0.574319
individual	3.540967	0.354106	profound	0.338944	-0.870383
special	3.540967	0.354106	elegant	0.338944	-0.870383
dynamic	3.513548	0.343620	expensive	0.275192	-0.894763
active	3.513548	0.343620	fulfil	0.275192	-0.894763
fresh	3.403393	0.301496	stable	0.140409	-0.946306
unprecedented	3.403393	0.301496	confidence	0.140409	-0.946306
pure	2.722331	0.041050	static	0.117991	-0.954878
brisk	2.722331	0.041050	inactive	0.117991	-0.954878

など、直接楽曲メタデータの“c6”に属する語と関係した表情語群が抽出されている。また、小さい相関である下位の印象語として、「静的, 非活性的(static, inactive)」をはじめ、「安定感, 信頼感(stable, confidence)」など、“動きがない”, “静かな”といった表情語群が抽出されている。これらに対する相関は式(14)の正規化手法によって、マイナスの値になる。これにより、検索精度を上げているが、相関が小さくなったりマイナスになったりしたとはいえ、これほどの量の印象語が同時に入力されることは、文脈が定まらず、ノイズとなっている可能性がある。このことから、上記のような正規化手法のみでなく、不適切な表情語群のみ排除するなどの方式を導入する必要があると考えられる。

5.3.3 考 察

以上の結果から、本方式を感性検索の入力インタフェースとして適用した場合、「暖かさ, 和やかさ」などの印象を持つ語に対して有効であることが示された。一方、「不快, 苦痛, 哀歎」などの印象に対しては検索精度が高いとはいえないが、実験により、抽出される印象の相関の大きさを正規化することである程度の精度改善が見込めることが示された。これらの

結果から、本方式を感性検索に適用する場合、本方式の特性を考慮した適用方式が重要であることが分かる。実験1の結果も考慮すると、本方式は、任意の言葉すべてに対して適用するのではなく、人名や商品名といった固有名詞や、および副詞やオノマトペ（擬音語・擬態語など）、未定義・未知の語といった、ユーザの感性を明確に表現するが、既存の手法を用いて実現するのが困難な語に対して適用することによりその真価を発揮すると考えられる。

本実験では、楽曲 MIDI データを対象とした感性検索の入力インタフェースとして応用したが、本方式は、メディアデータを対象としたオノマトペや未定義語などによるタグ、メタデータの付与とそれらのデータの検索など様々な用途への適用が考えられる。

本実験では、式(14)で表されるような正規化方式を用いたが、正規化方式はこの方法に限らない。アプリケーションごとに正規化を変えることも可能であると考えられる。また、本実験から、正規化だけではなく、不要な表情語群を排除する機構も重要であると考えられる。本実験では重みの絶対値の小さな表情語群を一律で排除する方式を用いて結果の比較を試みたが、有意な検索精度の改善には至らなかった。

このような検索精度の悪化を引き起こした理由として、マルチメディア・コンテンツの多角性があげられる。言葉の印象に関わる要素として、言葉の3要素⁶⁾、すなわち“簡明性”、“視覚性”、および“サウンド性”の3つがあるが、提案方式では、入力される語の音韻、すなわち言葉の“サウンド性”のみに着目しているために抽出される印象に偏りが生じてしまうと考えられる。したがって、言葉の3要素すべてを扱う機構を実現し、より多様な印象の表現が可能な印象変換作用素の構築が重要である。このような複数の要素を複合した作用素の構築は今後の課題である。

さらに、本稿では、50曲の楽曲 MIDI データを対象として実験を行ったが、今後大量の対象データを用いての検証、および他のメディアデータやアプリケーションを対象とした検証を行う必要がある。その際に、上記の正規化方式、不要な印象語の排除方式の検証も同時に必要になると考えられる。また、楽曲検索において検索精度が低下した理由として、提案方式により言葉の音韻に付与される印象と楽曲 MIDI データに付与された楽曲印象語群との間に意味的なギャップがあるために、“印象語関連計量機構”において適切な計量が行えなかったことが考えられる。本稿では、楽曲からの印象抽出方式として、楽曲構造要素を用いる方式¹¹⁾を採用したが、楽曲の印象は、楽曲の構造だけでなく、音色や演奏など様々な要素から複合的に想起されると考えられる。そのため、楽曲に関しても言葉の音韻同様、メディアの多角性を考慮することが重要である。

今後の課題として、メディアの多角性に基づく印象変換作用素の実現が重要となる。メ

ディアの多角性を反映させることによってより多様な印象、微妙なニュアンスなどを適切に表現できるようになり、本提案方式の有効性はさらに増大すると考えられる。

6. おわりに

本稿では、任意の言葉からその言葉が持つ音の印象自動抽出を実現する作用素の構成方式について示した。この作用素の実現により、任意の言葉や未定義・未知の言葉による感性表現を用いた検索などのユーザの入力インタフェースや、印象による商品名の検索などに有効に用いることが可能となる。また、本方式の適用例、および、実験システムとして、楽曲 MIDI データを対象とした本方式を有する感性検索を実現し、本方式の有効性を示した。

我々がふだん使っている意味情報を持つ言葉でさえも、意味情報を排除した音が、印象を伝えるコミュニケーションに非常に重要な役割を果たしている。本方式は、言葉を介する人間とコンピュータ間のインタフェースとして、意味情報だけでなく、その言葉が持つ語感や音相から想起される印象を介したコミュニケーション機構であるといえる。これにより、ユーザの任意の言葉や未定義・未知の言葉による感性表現が可能となり、ユーザへの負荷が少ないインタフェースが実現される。音相に代表される、印象や感情などの様々な情報を効果的に表すメディアを含む異種のメディアを統一的に扱うことが可能な方式を実現できれば、人間の感性や直感に合致した、マルチメディア環境での有効的なコミュニケーションメディアの実現の第一歩となりうる。

今後の課題として、メディアの多角性を考慮した、他の印象を喚起させる要因を使った複合的な印象抽出方式の実現とその統合、本方式の被験者による検証実験、大規模データによる感性検索への応用、本方式を利用した異種メディア間連携・統合方式、さらには、提案方式の逆作用素を実装して、音楽などのメディア・コンテンツをその印象を最もよく表現する言葉に逆変換するなど、様々なアプリケーションへの応用が考えられる。

参 考 文 献

- 1) 田守育啓：日本語オノマトペ—多様な音と様態の表現，日本音響学会誌，Vol.54, No.3, pp.215-222 (1998).
- 2) 藤沢 望，尾畑文野，高田正幸，岩宮眞一郎：2 モーラの擬音語からイメージされる音の印象，日本音響学会誌，Vol.62, No.11, pp.774-782 (2006).
- 3) Takada, M., Tanaka, K. and Iwamiya, S.: Relationships between auditory impressions and onomatopoeic features for environmental sounds, *Acoustical Science and Technology*, Vol.27, No.2, pp.67-79 (2006).

- 4) Fujisawa, N., Iwamiya, S. and Takada, M.: Auditory Imagery Associated with Japanese Onomatopoeic Representation, *Journal of Physiological Anthropology and Applied Human Science*, Vol.23, No.6, pp.351-355 (2004).
- 5) Asaga, C., Mukarramah, Y. and Watanabe, C.: ONOMATOPEDIA: Onomatopoeia Online Example Dictionary System Extracted from Data on the Web, *Progress in WWW Research and Development: 10th Asia-Pacific Web Conference (APWeb2008)*, pp.601-612 (2008).
- 6) 岩永嘉弘: 絶対売れる! ネーミングの成功法則 コンセプトづくりから商標登録まで, PHP 研究所 (2002).
- 7) 木通隆行: 音相: 社名, 商品名から人名までヒット・ネーミングは“音”で決まる, プレジデント社 (1990).
- 8) 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).
- 9) Kitagawa, T. and Kiyoki, Y.: Fundamental framework for mediadata retrieval systems using medialexco transformation operator, *Information Modeling and Knowledge Bases*, IOS Press, Vol.12, pp.316-326 (2001).
- 10) Nakanishi, T. and Kitagawa, T.: An Implementation Method of a Heterogeneous Associative Media Data Search for Music Data and Image Data, *The IASTED International Conference on Databases and Applications*, pp.143-152 (2006).
- 11) 北川高嗣, 中西崇文, 清木 康: 楽曲メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的楽曲検索への適用, 電子情報通信学会論文誌, Vol.J85-D-I, No.6, pp.512-526 (2002).
- 12) 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).
- 13) 北川高嗣, 中西崇文, 清木 康: 静止画像メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的画像検索への適用, 情報処理学会論文: データベース, Vol.43, No.SIG12 (TOD16), pp.38-51 (2002).
- 14) Kitagawa, T. and Kiyoki, Y.: The Mathematical Model of Meaning and its Application to Multidatabase Systems, *The 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems*, pp.130-135 (1993).
- 15) Kiyoki, Y., Kitagawa, T. and Hayama, T.: A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning, *Multimedia Data Management - using metadata to integrate and apply digital media*, chapter7, Mc-GrawHill

(1998).

- 16) 清木 康, 金子昌史, 北川高嗣: 意味の数学モデルによる画像データベース探索方式とその学習機構, 電子情報通信学会論文誌, Vol.J79-D-II, No.4, pp.509-519 (1996).
- 17) Longman Dictionary of Contemporary English, Longman (1987).
- 18) SEO-Association. <http://www.seo-association.com/>

(平成 21 年 7 月 17 日受付)

(平成 22 年 2 月 5 日採録)



本間 秀典

独立行政法人情報通信研究機構 (NICT) 知識処理グループ研究員。2009 年筑波大学大学院システム情報工学研究科博士後期課程修了。博士 (工学)。電子情報通信学会, 日本データベース学会各会員。



中西 崇文 (正会員)

独立行政法人情報通信研究機構 (NICT) 知識処理グループ研究員。2006 年筑波大学大学院システム情報工学研究科修了。博士 (工学)。マルチメディアシステム, 知識処理システムに関する研究に従事。電子情報通信学会, 日本データベース学会各会員。



北川 高嗣 (正会員)

筑波大学大学院システム情報工学研究科教授。1978 年名古屋大学工学部卒業。1983 年同大学院工学研究科博士課程修了。スタンフォード大学計算機科学科客員研究員, 愛媛大学理学部数学科講師, 筑波大学電子・情報工学系助教授を経て現在に至る。数値解析, 逆問題, マルチメディア情報システムの研究に従事。日本応用数理学会会員。