

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

本 間 秀 典[†] 中 西 崇 文^{††} 北 川 高 嗣^{†††}

本稿では、言葉の音声表現に着目することにより、与えられた任意の言葉の音を持つ印象をメタデータとして抽出するための方式を提案する。本方式では、音相と呼ばれる概念を利用して、与えられた言葉の音から聞き手が受ける印象を印象語として抽出する。これにより、与えられた任意の言葉を、その印象を反映したメタデータにより構成されるベクトルとして表現することが可能となる。本方式を用いることにより、任意の語によるユーザの意図する感情に合致したメディアデータ検索機能が実現できる。さらに、本方式による任意の語と異種のメディアデータ間の連携が可能となり、ユーザの感性を反映した新たな機能を提供することが可能となると考えられる。

A Method of Automatic Extraction Corresponding to the Impression of the Sounds of Arbitrary Words

HIDENORI HOMMA,[†] TAKAFUMI NAKANISHI^{††} and TAKASHI KITAGAWA^{†††}

In this paper, we present a method of automatic metadata extraction corresponding to the impression by the sound expression of arbitrary words. This method can extract impression words corresponding to the "Onso" from words or phrases. By this method, an arbitrary word given by a user can be transformed to a vector by impression words corresponding to "Onso". This method realizes the search function for media data corresponding to user's impression by using arbitrary words. In addition, this method can realize the connections between arbitrary words and heterogeneous media data.

1. ま え が き

近年、コンピュータネットワークの広域化・高速化により、我々がコンピュータを介して様々な情報群と接触する機会が多くなってきている。このような環境において、多種多様なメディアデータがネットワーク上に散在しつつあり、これらを対象とした、高度な情報検索方式、及び知識の獲得方式が求められている。特に、これまでのようなコンピュータと人間の論理情報の伝達だけでは、操作によるユーザの負担が大きくなっていることから、人間の感性や直感に合致した、ユーザへの負担が少ないコミュニケーションメディアの実現が重要な課題となってきている。

一般に、我々のコミュニケーションにおいて、互いの感情を正確に理解することが重要である。メディアデータ群を対象としたシステム環境において、感性的

な情報が人間とシステムとの間のコミュニケーションメディアとして導入されれば、人間の感性や直感に合致した、ユーザへの負担が少ないインターフェースの実現が可能となると考えられる。

文献^{1),2)}によれば、一般に語や語句はさまざまな印象を喚起するが、語の印象はその語の持つ意味からだけでなく、語の音の特徴からも喚起されるとしている。そのため、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ひとつである。日本語における言葉の音とその印象の相関関係は、「音相理論」^{1),2)}と呼ばれる研究の中で示されている。この音相をメディアデータ群を対象としたシステム環境に適用することにより、操作によるユーザへの負担が少ないインターフェースが実現できると考えられる。

本稿では、言葉の音の印象に合致するメタデータを自動抽出することにより、任意の言葉をその印象に合致したベクトルに自動的に変換する方式について示す。本方式により、任意の言葉からそれらの音の印象に合致する印象語群を抽出し、それらを用いることにより、任意の言葉と他のメディアデータを統一的に扱い、言葉による、ユーザの意図した感情に合致した様々な機能を実現することが可能になると考えられる。

[†] 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

^{††} 独立行政法人 情報通信研究機構
National Institute of Information and Communications Technology

^{†††} 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

これまで、文献³⁾⁻⁵⁾で、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、言葉と検索対象のメディアデータ、ドキュメント間を文脈に応じて動的に計算することを可能とする。意味の数学モデルでは、検索対象をベクトル化し、メタデータ空間と呼ばれる空間に写像する。さらに、それらのベクトルをメタデータ空間の部分空間に射影して計量することにより、文脈に応じた連想検索を実現している。

本方式を、意味の数学モデル³⁾⁻⁵⁾などの連想検索方式に適用することにより、言葉の持つ印象に合致した連想検索が可能となる。また、本方式は、「音相理論」を適用し、言葉が持つ音から印象語群を抽出する方式であるため、未知語、擬音語など、これまでの検索システムにおいて利用することが困難であった語についても適用可能である。これにより、検索に対する負担を軽くするインターフェースとして有効な方式であると考えられる。

音相に代表される、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。さらに、これによりユーザへの負荷が少ないインターフェースの実現が可能となると考えられる。

2. 意味の数学モデルの基本構成

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³⁾⁻⁵⁾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。具体的には以下の通りメタデータ空間 MDS を設定する。

初めに、 m 個の基本データについて各々 n 個の特徴 (f_1, f_2, \dots, f_n) を列挙した特徴付ベクトル $d_i (i = 1, \dots, m)$ が与えられているものとし、そのベクトルを並べて構成する $m \times n$ 行列を M とおく(図1)。このとき、 M は、列ごとに2ノルムで正規化されている。

- (a) データ行列 M の相関行列 $M^T M$ を計算する。
- (b) $M^T M$ を固有値分解する。

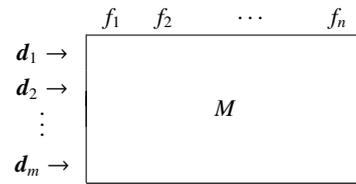


図1 データ行列 M によるメタデータの表現。

$$M^T M = Q \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_\nu & \\ & & & 0_{\nu_0} \end{pmatrix} Q^T, \quad (1)$$

$$0 \leq \nu \leq n.$$

ここで行列 Q は、

$$Q = (q_1, q_2, \dots, q_n) \quad (2)$$

である。この $q_i (i = 1, \dots, n)$ は、相関行列の正規化された固有ベクトルである。相関行列の対称性から、この固有値は全て実数であり、その固有ベクトルは互いに直交している。

- (c) メタデータ空間 MDS を以下で定義する。非ゼロ固有値に対応する固有ベクトルによって形成される正規直交空間をメタデータ空間 MDS と定義する。この空間の次元 ν は、データ行列 M のランクに一致する。この空間は、 ν 次元ユークリッド空間となる。

$$MDS := span(q_1, q_2, \dots, q_\nu). \quad (3)$$

$\{q_1, \dots, q_\nu\}$ は MDS の正規直交基底である。

(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。

具体的には、メディアデータのメタデータを以下のようにベクトル化する。

(a) メディアデータの特徴づけ

メディアデータ P を t 個の印象語(あるいは、 t 個のオブジェクト) o_1, o_2, \dots, o_t を用いて、次のように特徴づける。

$$P = \{o_1, o_2, \dots, o_t\}. \quad (4)$$

ここで、各印象語 o_i は、データ行列の特徴と同一の特徴を用いて表現される特徴付ベクトルである。

$$o_i = (o_{i1}, o_{i2}, \dots, o_{im}) \quad (5)$$

(b) **メディアデータ P のベクトル表現**

メディアデータ P を構成する t 個の印象語 o_1, o_2, \dots, o_t が、それぞれ n 次元のベクトルで定義されている。印象語 o_1, o_2, \dots, o_t は、合成することで n 次元ベクトル表現され、メディアデータベクトル p を形成する、さらにこのメディアデータベクトル p をメタデータ空間 MDS に写像する。これにより、同じ空間上に言葉とメディアデータが配置されることになり、言葉とメディアデータの間を空間上の距離として動的に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 の部分空間 (意味空間) 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. 任意の言葉を対象とした音の印象によるメタデータ自動抽出方式

本節では、任意の言葉を対象とした、言葉の音の印象に合致するメタデータの自動抽出方式について述べる。

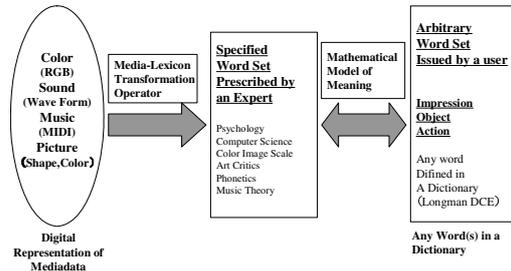


図 2 Media-lexicon Transformation Operator の概要。
Fig. 2 Media-lexicon Transformation Operator.

3.1 Media-Lexicon Transformation Operator の実現

本節では、Media-lexicon Transformation Operator⁶⁾ の概要を示す。

Media-lexicon Transformation Operator の概要図を図 2 に示す。Media-lexicon Transformation Operator ML は一般的に次のように表される。

$$ML(Md) : Md \mapsto Ws.$$

(Md : メディアデータ, Ws : (重み付き) 単語群) (6)

これらの演算子 ML により、任意の言葉をメディアデータ検索の間合わせとして利用し、意味の数学モデルによる、与えられた言葉の文脈に依存した連想検索の実現が可能となる。

3.2 言葉の音の印象に合致したメタデータ自動抽出方式

3.2.1 音相理論

文献^{1,2)}によれば、「音相」とは、言葉の音の響きを持つ印象であり、同じ言葉を話すほとんどの話者がその言葉に対して共通して持っているイメージであるといえる。音相理論によれば、我々は、その言葉のもつ意味の微妙な違いを、似た意味の語の中から、音相に基づいてよりそのイメージに近い音を持つ語を選んで使うことによっても伝えているとされる。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語の印象を決定している。音相基の中でも特に、それぞれの母音や子音が発音されるときに用いられる器官 (唇, 舌, 喉など) とそれらの操作の仕方にあたる「調音種」と、それぞれの音がどのような感情のときに用いられるかという 2 つの観点から音相による語の表情が作られるとされている。後者に関しては、その音の明るさを表す「輝性」(B, Brightness の略) および強さを表す「鋭性」(H, Hardness の略) という要素を単位としている。それぞれの

表1 音素と音価表 (子音の場合).
Table 1 The table of Onso for consonants.

		両唇音	前舌音		咽頭音
破裂音	無	p	t		k
		+B1.7 H1.6	+B1.4 H1.5		+B1.3 H1.3
	有	b	d		g
		-B2.2 H1.2	-B2.0 H1.0		-B2.0 H1.0
破擦音	無		ts	tʃ	
			+B1.7 H2.2	+2.2 H2.7	
	有		dʒ	ʧ	
			-B2.0 H1.0	-B1.5 H1.5	
摩擦音	無	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		r		
			+B0.4 H0.8		
鼻音	有	m	n		ŋ
		B0.0 H0.0	B0.0 H0.0		B0.0 H0.0
接近音	有	w	j		
		-B0.5 H0.0	+B0.5 H0.5		

- ・ 無：無声音
- ・ ts：ツの子音
- ・ dʒ：ジを除くザ行の子音
- ・ f：フおよびファ行の子音
- ・ ŋ：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
- ・ 有：有声音
- ・ tʃ：チおよびチャ行の子音
- ・ ʧ：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

表2 音素と音価表 (母音の場合).
Table 2 The table of Onso for vowels.

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

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²⁾において表1, 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、正であれば語の音の印象が明るく、負であれば印象が暗いことを示し、絶対値の大きさはその度合いを示している。また、Hの値は符号無しの数値で表され、数値が大きいほど語の音が強いインパクトを持っていることを表している。

文献²⁾では、音相による表情を表3に示すように20の群と40の表情属性(語)に分類している。さらに、音相を音相基それ自体が持つ表情である甲類表情と、幾つかの音相基の組み合わせによって生まれる乙類表情の2種類に分類している。甲乙それぞれの表情と表情語の関係は表4, 5のように示されている。

3.3 言葉の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、言葉の音相の印象に合致したメタデータ

表3 表情語と表情属性表 (抜粋).
Table 3 The table of impression words.

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, 澁刺さ	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, fulfill
S	高尚な, 優雅さ	profound, elegant
T	静的, 非活性的	static, inactive

(*) 本稿における実験で利用可能な英語

自動抽出方式について述べる。

本方式は以下の3ステップから成る。

Step1: 入力語の初期ベクトルへの変換

文献²⁾によれば、音相基とそれらの印象を表す言葉の関係として表4, 5に示される78種類が定義されている。これらのうち、甲類の40種類は、表6に示す29の音相基のいずれかの組み合わせにより判定可能である。そこで、表1, 2を用いて、入力語を29の音相基を要素とする29次元のベクトル w_{in} に変換する。

Step2: 表情の判定による初期ベクトルの拡張

文献²⁾において、表7のように各音相基ごとの使用数や拍数(その語の音の数)に対する割合の標準値が示されているので、これによって初期ベクトル w_{in} を、甲乙78種類の表情を反映した78次元のベクトル w_f に変換する。ここで、その語の拍数を N_H と表すこととすると、甲類表情におけるある判定対象値 x が「多用(もしくは過少)」であることに対応する要素 p は、標準値 avg 以下ならば0, および標準値 +15% 以上ならば1となるように、以下の式(7)で求められる。

$$p = \begin{cases} 0 & \text{if } x \leq avg \\ \left(\frac{|x|}{N_H} - avg \right) \times \frac{1}{0.15} & \text{if } avg < x < (avg + 0.15) \\ 1 & \text{if } x \geq avg \end{cases}$$

表 4 甲類表情.

Table 4 A part of individual properties.

音相基	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,I,L,O
19. N(撥音) 多用	Q
20. 濁音多用	P,Q,R,S,T
21. 高勁性	A,B,C,D,E,G,H,I,L
22. 低勁性	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,I,K,M
25. マイナス輝性	Q,R,S
26. マイナス高輝性	Q,R,T
27. 高勁輝拍多用	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 5 乙類表情.

Table 5 A part of anaphoric properties.

甲類表情の組	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, E, 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

また、乙類表情については、条件を満たす表情に対応する要素の値は全て 1 に設定した。

Step3: 表情語への変換行列の作成

表 3, 4, および 5 を用いることによって、Step2 で生成される 78 次元のベクトル w_f を表 3 に示されている 40 の表情語とその重みからなる 40 次元のベクトル w_c へ変換するための行列 T を作成することが可能である。

T は図 3 のように表現される。 T の行要素には 40 の表情語が配置され、列要素には 78 の表情が配置される。そして、ある表情語を抽出する表情に対応する要素の値を 1 とし、それ以外の要素を 0

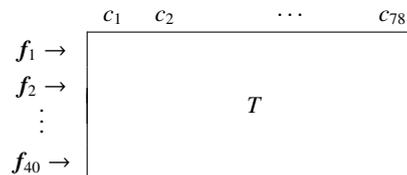


図 3 変換行列 T の表現.

Fig.3 Representation of the matrix T .

として行列の特徴付けを行う。図中の f は表情語を、 c は表情を表している。

Step4: 入力語の音相の印象と合致するメタデータの自動抽出

以下の演算により、入力語の音相の印象と合致す

表 6 初期ベクトルの構成に用いた要素 (音相基) 群.

Table 6 Elements for the first vector.

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

るメタデータの抽出と適切な重み付けが自動的に
行われる.

$$w_o = Tw_f$$

$$= (wf_1, wf_2, \dots, wf_{40})^T$$

ここで w_{f_n} は印象語 f_n の重みを表す.

4. 実 験

本方式の有効性を示すため、意味の数学モデルに基づくメディアデータ検索システムを構築し、検索実験を行った.

4.1 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”⁷⁾ 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.

4.2 実験例

本方式により抽出される印象語群をメタデータとすることにより、様々な商品名を検索対象に設定して意味的連想検索を行った場合の実験結果を表 8 に示す. ここで、検索対象としては、Web サイト「SMBC コンサルティング」⁸⁾ のコンテンツ「ヒット商品番付」などから、無作為に 30 語を選んだ.

表 7 音相基の標準値.

Table 7 Typical values for 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%(*)
エ音	16%	高勁脚拍	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%
有声摩擦音系	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²⁾ 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
母音種比	1~2	判定対象外
	3	3 以上で「多用」、1 以下で「過少」
	4	3 以上で「多用」、1 以下で「過少」
	5~7	4 以上で「多用」、2 以下で「過少」
	8~10	「多用」は判定対象外、3 以下で「過少」
	11 以上	「多用」は判定対象外、4 以下で「過少」
調音種比	1	判定対象外
	2	3 以上で「高」、1 以下で「低」
	3~4	4 以上で「高」、2 以下で「低」
	5~6	5 以上で「高」、3 以下で「低」
	7~9	6 以上で「高」、4 以下で「低」
	10~14	7 以上で「高」、5 以下で「低」
	15 以上	「高」は判定対象外、7 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	8 拍以上	
少拍	3 拍以下	

コンテキスト”quiet”, 及び”cool”とした場合の検索結果の例が表 8 に示されている. まず, ”quiet”は静かさや落ち着きを表すコンテキストである. この場合, “ラストサムライ”(映画のタイトル) や”アクオス”(液晶テレビの名前) といった, 落ち着きのある印象を持つ語が上位に現れていることが分かる. 次に, ”cool”は涼しさや心地よさを表すコンテキストである. この場合, “ニンテンドー DS”(携帯ゲーム機), “ショップ 99”(コンビニエンスストア) などのような, 新しさや新鮮さを連想する語が多く検索されている. これらの結果から, 提案方式により任意の言葉から抽出した印象語群が, ユーザの意図した感性や感情を反映していることが確認できた.

表 8 検索結果の例

context: quiet の場合		context: cool の場合	
word	correlation	word	correlation
ラストサムライ	0.252235	ニンテンドー DS	0.111455
アクオス	0.232417	クイーン	0.109399
みなとみらい線	0.223501	ショップ 99	0.109110
アテネオリンピック	0.219228	クラウン	0.108739
三越	0.218634	コエンザイム Q10	0.108613
郷ひろみ	0.218398	ハルウララ	0.108272
ヨン様	0.218351	iPod	0.107886

5. あとがき

本稿では、音相理論を用いることによって、任意の言葉の音が持つ印象に合致した印象語群をメタデータとして自動抽出する方式を提案した。また、検索実験により、本方式により任意の言葉をユーザの感性に合致したメディアデータとして利用することが可能であることを示した。本方式により、ユーザの感性や直感を反映したコンテキストの表現など、ユーザの負荷を軽減した新しいインターフェースの実現が可能となると考えられる。本方式は、ユーザの感性や直感に合致した、マルチメディア環境での効果的なコミュニケーションメディアの実現の第一歩になると考えられる。

今後の課題としては、音声入力などによる入力インターフェースの改善や、本方式と他のメディアデータの連携による新たな情報の生成及び利用方式の実現などが挙げられる。

参 考 文 献

- 1) 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).
- 2) 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).

- 3) Kitagawa, T. and Kiyoki, Y.: The mathematical model of meaning and its application to multi-database systems, Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems, pp. 130-135(1993).
- 4) Kiyoki, Y., Kitagawa, T. and Hayama, T.: “A Metadata 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).
- 5) 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,D-II, Vol.J79-D-II, No. 4, pp. 509-519 (1996).
- 6) T.Kitagawa, Y.Kiyoki: ”Fundamental framework for media data retrieval system using media lexico transformation operator”, *Information Modeling and Knowledge Bases*, vol.12, pp.316-326,(2001)
- 7) Longman Dictionary of Contemporary English, Longman (1987).
- 8) <http://www.smbc-consulting.co.jp/BizWatch/>