

任意の言葉の印象語抽出による連想検索のための問い合わせベクトル自動生成方式

本間 秀典[†] 中西 崇文[†] 北川 高嗣^{††}

[†] 筑波大学大学院 システム情報工学研究科 〒 305-8573 茨城県つくば市天王台 1-1-1 情報数理研究室

^{††} 筑波大学大学院 システム情報工学研究科 〒 305-8573 茨城県つくば市天王台 1-1-1

E-mail: [†]{homma,takafumi}@mma.cs.tsukuba.ac.jp, ^{††}takashi@cs.tsukuba.ac.jp

あらまし 本稿では、与えられた任意の言葉が持つ印象を印象語として抽出することにより、その印象と合致したメディアデータ連想検索の実現方式を提案する。本方式では、音相と呼ばれる概念を用いて、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印象語から成る問い合わせベクトルに変換し、メディアデータ連想検索のための問い合わせとして用いることが可能となる。本稿では、意味の数学モデルを用いてメディアデータを対象とした検索系を構築し、本方式により発行される印象語群を検索の問い合わせとすることによって任意の言葉を問い合わせとしたメディアデータ連想検索を実現し、検索結果についても示す。

キーワード 情報検索, メタデータの抽出, 意味的連想検索, 音相

An Automatic Creation Method of the Query Vector by a Metadata Extraction Corresponding to the Impression of Arbitrary Words

Hidenori HOMMA[†], Takafumi NAKANISHI[†], and Takashi KITAGAWA^{††}

[†] Graduate School of Systems and Information Engineering, University of Tsukuba

^{††} Graduate School of Systems and Information Engineering, University of Tsukuba

E-mail: [†]{homma,takafumi}@mma.cs.tsukuba.ac.jp, ^{††}takashi@cs.tsukuba.ac.jp

Abstract In this paper, we present an automatic creation method of the query vector by a metadata extraction corresponding to the impression of arbitrary words. This method realizes extraction of 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” of it. This method realizes that an arbitrary word can be used as a query vector for a semantic associative search. In addition, we realize a semantic associative search for media data by the mathematical model of meaning, and show the experimental results by using an arbitrary word as a query vector.

Key words information retrieval, extraction of metadata, semantic associative search, “Onso”

1. ま え が き

近年、コンピュータネットワーク上には多種多様なメディアデータ群が散在している。この中で、これまでのコンピュータと人間の論理情報の伝達だけでは、操作による人間の負荷が大きくなっており、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た。

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、感性的な情報を人間とシステムと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。

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、ユーザへの負荷が少ないインターフェースとして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

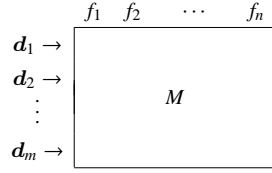


図1 データ行列 M によるメタデータの表現.

し、それを問合わせベクトルとした連想検索の実現方式について示す。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って、任意の言葉を検索者の印象に合致した連想検索のための問合わせとして利用することが可能となる。

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を抽出し、問合わせとして用いる方式であるため、未知語、擬音語など、これまで問合わせで使う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が統一的に扱うことが可能な方式の実現ができれば、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。

$$M^T M = Q \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_v & \\ & & & 0 \dots 0 \end{pmatrix} Q^T, \quad (1)$$

$$0 \leq v \leq n.$$

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 v は、データ行列 M のランクに一致する。この空間は、 v 次元ユークリッド空間となる。

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

$\{q_1, \dots, q_v\}$ 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。

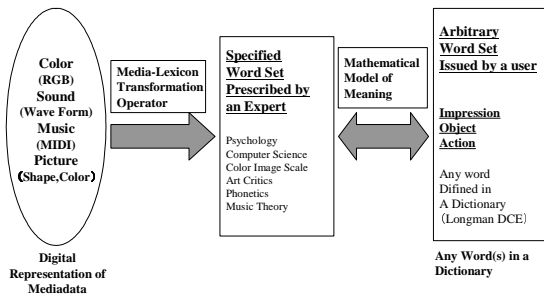


図2 Media-lexicon Transformation Operator の概要。
Fig.2 Media-lexicon Transformation Operator.

(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. 言葉の音の印象語抽出による問合わせベクトル自動生成方式

本節では、任意の言葉の音の印象に合致するメタデータの自動抽出方式、及び、それらを用いた連想検索のための問合わせベクトル自動生成方式について述べる。

3.1 Media-Lexicon Transformation Operator の実現

本節では、Media-lexicon Transformation Operator [6] 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。

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

$$(Md : \text{メディアデータ}, Ws : \text{(重み付き) 単語群}) \quad (6)$$

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問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 語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「勁性」(H, Hardness の略) 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 [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、数値が大きいく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、音相を音相基それぞれ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_m 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_m 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。このとき、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。

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		dʒ	dʒ	
			-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
- tʃ: チおよびチャ行の子音
- dʒ: ジを除くジャ行の子音
- dʒ: ジ(ヂ)およびジャ行の子音
- f: フおよびファ行の子音
- ʃ: シおよびシャ行の子音
- ŋ: ガ行鼻音
- j: ヤ行子音および拗音の第2子音
- 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

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

$$p = \begin{cases} 0 & \text{if } x \leq \text{avg} \\ \left(\frac{|x|}{\text{拍数}} - \text{avg} \right) \times \frac{1}{0.15} & \text{if } \text{avg} < x < (\text{avg} + 0.15) \\ 1 & \text{if } x \geq \text{avg} \end{cases} \quad (7)$$

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_o 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。

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, fulfil
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

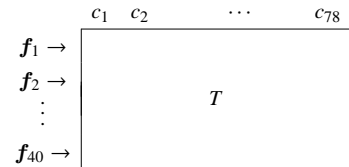


図3 変換行列 T の表現.

Fig. 3 Representation of the matrix T .

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。

$$w_o = T w_f = (w_{f1}, w_{f2}, \dots, w_{f40})^T$$

ここで w_{fn} は印象語 f_n の重みを表す。

3.4 問合わせベクトル自動生成方式

3.3節に示す方式を用いることにより、任意の語を40の印象語とそれらの重みからなるベクトル w_o に自動的に変換することが可能となる。適当な閾値を設定して抽出される語を限定することにより、このベクトルを意味的連想検索のための問合わせとして利用することが可能である。

4. 実験

提案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

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,J,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,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,I,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

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。

ここで、問合わせとして用いる印象語としては、上記の約

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, 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

2,000 語の基本語を用いて特徴付けが可能なすべての語を利用することができる。本実験において提案方式により抽出される全ての印象語は上記の約 56,000 語の見出し語のいずれかに対応しているため、変換後の問合わせベクトルを用いることにより、任意の語を問合わせとして用いることが可能である。

また、提案方式により抽出される印象語は、本実験で利用可能な英単語に翻訳して用いる。それらの訳語は、表 3 の第 3 列に示されている。

4.2 予備実験

本予備実験では、提案方式によって音相に基づいた語の印象語が適切に抽出されているかどうかを確認する。ここでは、文

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	高勁輝拍数
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

献 [2] で分析されている語を対象として, 提案方式により抽出された語の妥当性を示す.

語『ゴージャス』から抽出される印象語のうち, 重みが大きい 10 語を表 8 に示す. 文献 [2] によれば, 「非活性的, 静的」「充実感, 高級感」「信頼感, 安定感」の 3 つの語群が突出して高いポイントとなっているが, 表 8 によれば, それらに対応する訳語である “inactive”, “static”, “elegant”, “profound”, “confidence”, “stable” の重みが上位になっていることから, 音相に合致した印象語を抽出できていると考えられる.

4.3 実験方法

本実験では, 任意の語を問合わせとして用いて検索実験を行い, 検索結果を検証する. ここでは検索対象として 30 の画像を用いた. 各々の画像には表 9 に示すような印象語群がメタデータとして設定されており, これらを 2. 節 (2) の方法でメタデータ空間 *MDS* へマッピングした. なお, 検索対象の各画像に設定されている印象語群は人間の主観により決定した. さらに, 検索語としては, 提案方式により抽出される印象語群のうち, 重みが大きいものの上位 6 語を重み付きの問合わせベクトルとして用いることとした.

4.4 実験結果

表 10, 11 はそれぞれ, 検索語を『さんさん』とした場合の問合わせ及び検索結果である. 『さんさん』という言葉は「日の光がさんさんと降り注ぐ」などのように, 美しく光り輝くさまを表現する明るいイメージの擬態語である. 図 4 は検索結果で最上位となった画像であるが, 検索結果の上位には, “shine” や “delight”, “beautiful” など, 明るい印象のメタデータを設定された画像が検索されていることが分かる.

また, 表 12, 13 はそれぞれ, 検索語を『ゴッゴル』とした場合の問合わせ及び検索結果である. 『ゴッゴル』とは, “SEO コンテスト” [8] で検索語として利用された, 全く意味を持たない語である. 図 5 は検索結果の最上位となった画像であるが, “shine”, “calm” といったメタデータが設定されていた. その

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 [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 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

他の検索結果については, 『さんさん』の場合に比べ, “tender”, “dim” など, 暗いイメージに結びつくメタデータを持つ画像が検索される傾向が見られた.

4.5 考察

本方式により, 任意の語を音相に基づく印象語群からなるベクトルに変換し, これを用いて語のイメージに合致したメディアデータ連想検索を実現できることが分かった.

しかしながら, ポジティブな印象語が多く抽出される傾向があるため, 違う語で検索しても似通った検索結果を返すことがあることも分かった. そのため, ネガティブな印象語の取り扱い

表8 『ゴージャス』から抽出される印象語上位10件.

Table 8 Experimental results(gorgeous).

順位	印象語	重み
1	elegant	14.066667
	profound	
3	confidence	14.000000
	stable	
5	inactive	12.866667
	static	
7	expensive	8.200000
	fulfil	
9	comfotable	8.000000
	warm	

表9 検索対象に設定されたメタデータ.

Table 9 Metadata of each candidate mediadata.

画像名	印象語
chagall1	vivid quiet substance
chagall2	grief terrible sombre
chagall3	sober dynamic motion
chagall4	shine tender calm
corot1	beautiful grand calm
corot2	beautiful delicate calm
corot3	grief sombre sober
corot4	shine beautiful calm
gogh1	merry delight shine
gogh2	grief terrible somber
hiro1	twilight grand quiet
hiro2	cheer dim quiet
hiro3	beautiful quiet calm
hiro4	fine shine beautiful
hiro5	fine beautiful calm
hokusai1	dynamic strong motion
hokusai2	fight motion calm
hokusai3	delicate calm quiet
hokusai4	vivid motion speed
loirand1	shine grand calm
loirand2	delight shine calm
loirand3	delight grand calm
loirand4	quiet substance material
nelson1	grand dynamic motion
nelson2	twilight calm quiet
renoir1	dim tender quiet
renoir2	delight dim calm
renoir3	loud bustle crowd
renoir4	fine strong quiet
sarthou1	dynamic motion speed

いが必要であると考えられる.

5. あとがき

本稿では、語の音相に注目して、任意の語の音相の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、抽出されたメタデータを用いることにより、意味的連想検索の問合わせベクトルの自動生成方式についても示し

表10 コンテキスト『さんさん』の抽出語群.

Table 10 Impression words for "Sansan".

印象語	重み
healthy	7.200000
clean	
stable	5.289855
confidence	
open-minded	5.200000
bright	

表11 コンテキスト『さんさん』の検索結果上位10件.

Table 11 Experimental results(Sarada-kinembi).

検索結果	相関量
loirand1	0.285541
nelson2	0.274156
loirand2	0.272527
loirand3	0.268474
chagall4	0.265214
renoir2	0.262220
corot4	0.255645
hiro3	0.254157
corot1	0.251070
hokusai3	0.247163

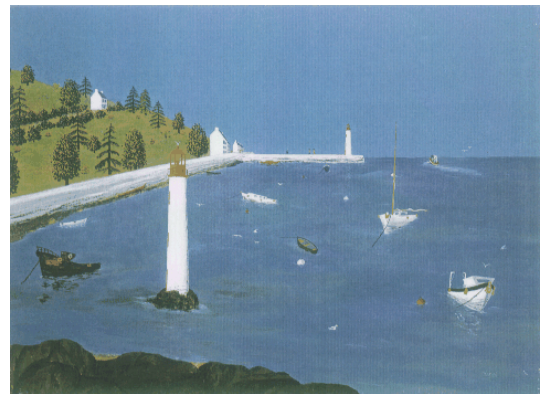


図4 コンテキスト『さんさん』の検索結果最上位(loirand1).

Fig.4 Experimental result ("loirand1").

表12 コンテキスト『ゴッゴル』の抽出語群.

Table 12 Impression words for "Goggoru".

印象語	重み
elegant	12.333333
profound	
confidence	12.133333
stable	
inactive	8.133333
static	

た。本方式を意味の数学モデルと組み合わせることにより、任意の語を意味的連想検索の問合わせとして利用することが可能となる。

しかしながら、本稿ではネガティブな印象を持つ語の抽出を行っていないため、検索実験の結果として明るい印象語をメタ

表 13 コンテキスト『ゴッゴル』の検索結果上位 10 件.

Table 13 Experimental results("Goggoru").

検索結果	相関量
chagall4	0.180767
corot1	0.174789
hiro3	0.170449
renoir1	0.170041
hiro5	0.166983
loirand1	0.166972
corot4	0.163333
loirand3	0.161957
hokusai2	0.160648
renoir4	0.157633



図 5 『ゴッゴル』の検索結果最上位 (chagall4).

Fig.5 Experimental result ("chagall4").

データに持つメディアデータが上位に挙がりやすい傾向が見られた。そのため、ネガティブな印象語の抽出は今後の課題として重要であると考えられる。また、音声入力などを用いた、よりユーザの負荷の小さい検索インターフェースの実現なども今後の課題である。

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 multidatabase 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 Metadatabase 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] Kitagawa, T. and Kiyoki, Y.: ”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] SEO-Association: <http://www.seo-association.com/>