

B1-2

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

中西 崇文[†] 北川 高嗣^{††} 清木 康^{†††}

本稿では、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ための感性作用素を画像メディアデータへ適用する方式について示す。メディアデータから抽出されるメタデータを対象として、人間の感性に反映した写像を実現する感性作用素を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に適用する。画像メディアデータ群を対象としたメタデータ自動抽出方式において、人間の感性や感覚を解釈する機構が導入されれば、人間の直感に合致した言葉によるメタデータの自動抽出が可能となる。さらに、意味的連想検索方式を用いることにより、画像メディアデータ群を対象とした人間の直感に合致する検索が可能となる。本稿では、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ための感性作用素を画像メディアデータに適用し、画像メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式について示す。

An Implementation Method of Automatic Metadata Extraction Method for Image Data and its Application to a Semantic Associative Search

NAKANISHI TAKAFUMI,[†] KITAGAWA TAKASHI^{††}
and KIYOKI YASUSHI^{†††}

This paper presents an implementation method of Kansei Operator for an automatic metadata extraction from image data. The Kansei Operator performs mapping with interpretation of human sensitivity for metadata extraction using logalismic function based on Fechner's law. The operator and a semantic associative search method enable semantic search according to the human Kansei for multimedia data. In this paper, we show an implementation method of automatic metadata extraction with Kansei Operator for image data. We clarify the effectiveness of our method by showing experimental results.

1. はじめに

現在、コンピュータネットワーク上に多種多様なメディアデータ群が散在している。また、それらを検索対象とするシステムの実現が行われつつある。メディアデータ群を対象とした情報獲得の機会の可能性が増大する一方、適切な情報獲得方式の実現が重要な課題となっている。

メディアデータを対象とした検索方式はメディアデータの特徴量を直接比較することによって検索を行う直接的な方法と、メディアデータをに付与された抽象データ（以下、メタデータ）を用いて検索する間接的な方法に大別できる。

我々は、メディアデータに対応するメタデータを言葉によって表現し、検索者の与える文脈に応じた意味的解釈を伴う間接的な検索方式として、メディアデータを対象とした意味的連想検

索方式^{4)~6)}を提案している。これにより、統計的に意味素を抽出して意味的解釈を実現する従来の研究¹⁾と比較して、言葉の意味を文脈に応じて解釈する機構より、言葉と言葉、あるいは、言葉とメディアデータ間の意味的な関係を与えられた文脈や状況に応じて動的に計算することが可能となる。現在の実現システムでは、文脈の様相の数は約 2^{2000} であり、ほぼ無限の文脈を表すことが可能である。

さらに我々は、文献^{3),5),7),8)}でメディアデータのメタデータを自動抽出するための実現方式について示している。メディアデータを対象としたメタデータ抽出において、人間の感性や感覚を解釈する機構が導入されれば、人間の感性や感覚に合致した言葉によるメタデータの自動抽出が可能となる。さらに、メディアデータを対象とした人間の直感に合致する検索が可能となる。これにより、メディアデータが人間に与える印象を抽出するメカニズムの解明の第1歩になると考えられる。

我々は、人間の感性や感覚を解釈する機構として、文献⁹⁾でメディアデータから抽出されるメタデータをメディアデータが人間に与える刺激の大きさや位置付け、人間の感覚により近い対数関数を反映することにより、人間の感性に合致した写像をする方式を実現する感性作用素を示している。また、文献⁹⁾では楽曲メディアデータを対象とした自動抽出方式^{3),7)}に感性作用素を適用し、楽曲メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式の実現と有効性を示している。

本稿では、カラーイメージスケール¹²⁾による画像メディア

[†] 筑波大学大学院システム情報工学研究科，つくば市
Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan
e-mail: takafumi@nalab.is.tsukuba.ac.jp

^{††} 筑波大学電子・情報工学系，つくば市
Institute of Information Sciences and Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan
e-mail: takashi@is.tsukuba.ac.jp

^{†††} 慶應義塾大学環境情報学部，藤沢市
Faculty of Environmental Information, Keio University, Fujisawa, Kanagawa 252-8520, Japan
e-mail: kiyoki@sfc.keio.ac.jp

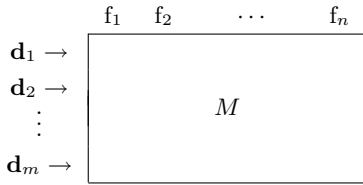


図1 データ行列 M によるメタデータの表現.

データを対象とした自動抽出方式⁸⁾に感性作用素⁹⁾を適用し、画像メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式の実現し、実験により有効性の検証する。

2. 意味の数学モデルと画像メタデータ自動抽出方式の概要

本節では、意味の数学モデル^{4)~6)}、および画像メタデータを対象としたメタデータ自動抽出方式⁸⁾の概要を示す。

2.1 意味の数学モデルの基本構成

人間が様々な印象を表す際に用いられる単語(以下、印象語)によって表現した問い合わせに対応したメディアデータを検索することを目的とした意味の数学モデルによるメディアデータ検索方式の概要を示す。詳細は、文献^{4)~6)}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。具体的な手順を以下に示す。

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。

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

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

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

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。具体的な手順を以下に示す。

(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) w_1, w_2, \dots, w_t を用いて、次のように特徴づける。

$$P = \{w_1, w_2, \dots, w_t\}. \quad (4)$$

ここで、各印象語 w_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。

$$w_i = (f_{i1}, f_{i2}, \dots, f_{in}) \quad (5)$$

(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 w_1, w_2, \dots, w_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 w_1, w_2, \dots, w_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 を形成し、メタデータ空間 MDS に写像される。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。印象語の合成方法の詳細は、3.1 節で述べる。

(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。

2.2 画像メタデータ自動抽出方式の概要

ここでは、デジタル化された色彩情報を対象として、人間が画像から受ける印象を抽出し、言葉によって表現されるメタデータを自動的に抽出する静止画像を対象としたメタデータ自動抽出方式⁸⁾について述べる。この方式は、マンセル表色系¹¹⁾で表される色彩情報と人間の印象の関係が示されているカラーイメージスケール¹²⁾を用いてマンセル表色系¹¹⁾で表される色彩情報を自動抽出し、それらの要素群と印象語の相関量を計算す

るためにカラーイメージスケール¹²⁾の示した相関関係を用いている。

2.2.1 カラーイメージスケール

カラーイメージスケール¹²⁾には、色彩と印象語の関係を表す統計データが示されている。

カラーイメージスケールの統計データは、マンセル表色系¹¹⁾における、10色相、12色調で表現される有彩色120色、および10階調で表される無彩色10色の計130色の基本色に対して180の印象語の関連性を5段階で表現されている。

マンセル表色系¹¹⁾とは、色彩の三属性の色相(Heu)、明度(Value)、及び彩度(Chroma)を用いる表現方法である。なお、3属性の概要を図2に示す。また、明度と彩度の相互の関連のもとに表現される明暗、濃淡など共通した色の状態(調子)を色相と呼ぶ。なお、明度、彩度と色調の関係を図3示す。

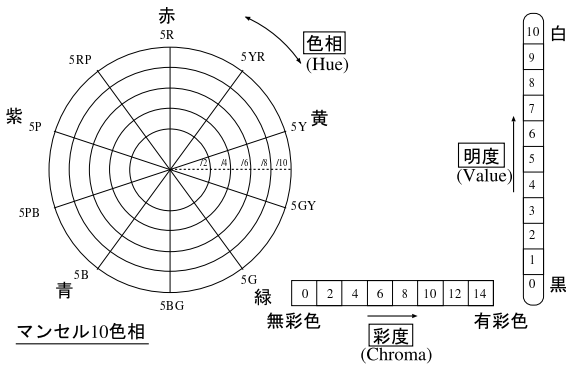


図2 マンセル表色系.

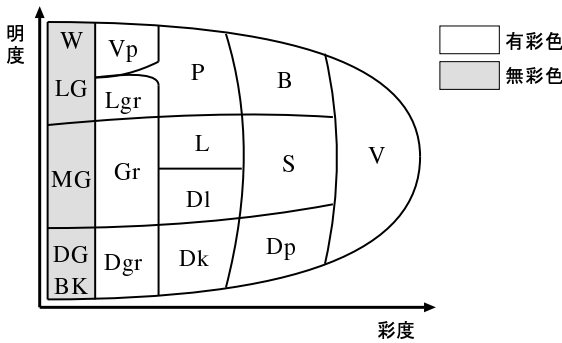


図3 明度・彩度と色調の関係.

2.2.2 画像メタデータ自動抽出方式のメタデータ抽出方法

ここでは、画像メタデータ自動抽出方式による静止画像メディアデータのメタデータ抽出方法⁸⁾を示す。また、メタデータ抽出方法の全体図を図4に示す。

(1) 色印象行列 C の作成

基本色 n 個の各々の色彩を、画像の定義に必要な十分な u 個の印象語 $w_{k\ell}$ ($\ell = 1, 2, \dots, u$) を特徴とした n 次元ベクトルとして表現する。このベクトルを、色印象ベクトルとして次のように表す。

$$c_k = (w_{k1}, w_{k2}, \dots, w_{ku})^T. \quad k = 1, 2, \dots, n \quad (6)$$

c_k の各要素は、色と印象語の関連の強さを示す数値データである。ここで色印象ベクトル c_k を行ベクトルとし (c_1, c_2, \dots, c_n) より構成される u 行 n 列の行列 C を、色

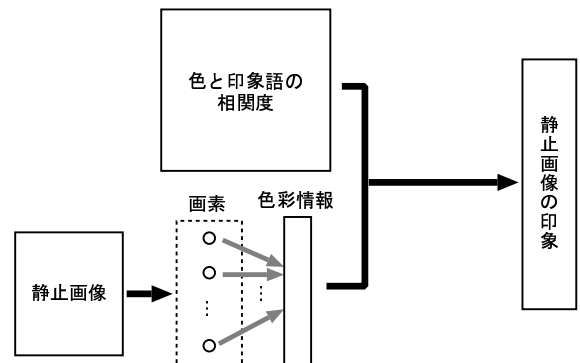


図4 画像メタデータ自動抽出方式の全体図.

印象行列 C とする (図5)。色印象行列 C は色と印象語との関連の強さを表す。

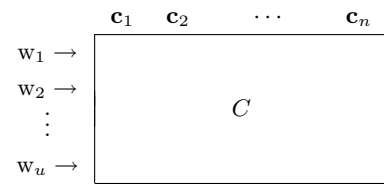


図5 色印象行列 C .

実装ではカラーイメージスケール¹²⁾で用いられている、マンセル表色系における有彩色120色、及び無彩色10色の基本色130色を用いる。

(2) 色彩情報の抽出

静止画像から色彩情報が抽出し、その色彩情報は画像全体における基本色 n 色の占める割合で構成される画像色彩ベクトル m によって表現される。画像色彩ベクトルを次に示す。

$$m = (m_1, m_2, \dots, m_n)^T. \quad (7)$$

但し、 m_k ($k = 1, 2, \dots, n$) は、画像における基本色 k の画像全体に占める割合に対応する。

(3) 画像のメタデータ抽出

色印象行列 C 、及び画像色彩ベクトル m を用いて、画像メタデータ I の抽出を行う。画像メタデータ I は、色印象ベクトル c_k に用いられる u 個の印象語と同一の印象語で特徴付けられるベクトルである。

画像メタデータ I を次のように表す。

$$I = Cm = \left(\sum_{k=1}^n w_{k1} m_k, \sum_{k=1}^n w_{k2} m_k, \dots, \sum_{k=1}^n w_{ku} m_k \right)^T. \quad (8)$$

3. 画像メタデータ自動抽出方式を対象とした感性作用素の適用

文献⁹⁾において、楽曲メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式の実現方法、及び有効性を検証してきた。本節では、文献⁹⁾で示した感性作用素の概要と、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⁸⁾に感性作用素を適用し、画像メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式の実現方法について述べる。

3.1 メディアデータを対象とした感性作用素

本節では、メディアデータから抽出されたメタデータをメタデータ空間 MDS へ人間の感性や直感に反映した写像を実現する感性作用素⁹⁾ について述べる。

メタデータ自動抽出方式などで、メディアデータから抽出されたメタデータは数個の印象語によって表されている。これらの印象語は n 次元ベクトルとして表される。 n 次元で表された印象語を合成することにより、メタデータはメタデータ空間 MDS に写像され、意味的連想検索が可能となる。

次に従来方式と提案する人間の感覚を反映した感性作用素として作用する方式(対数関数方式)の2種類の印象語方式について示す。

3.1.1 従来方式

従来方式は重みの大きさや特徴をメディアデータのメタデータにそのまま反映させる方式であり、以下のように表される。

メディアデータ P は、 t 個の印象語(あるいは、 t 個のオブジェクト) w_1, w_2, \dots, w_t から成る。ここで、 t は Hevner による印象語群のすべて語 ($t = 66$) を表している。

$$P = \{w_1, w_2, \dots, w_t\}. \quad (9)$$

各印象語 w_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。

$$w_i = (f_{i1}, f_{i2}, \dots, f_{in}). \quad (10)$$

従来方式を実現する演算子 \oplus_S を次のように定義する。

$$\oplus_S w_i := \left(\sum_{i=1}^t f_{i1}, \sum_{i=1}^t f_{i2}, \dots, \sum_{i=1}^t f_{in} \right). \quad (11)$$

3.1.2 対数関数方式

物理的な刺激と人間の感覚の関係を調べた研究で Fechner の法則¹³⁾ がある。本方式では、各特徴について印象語の重みを合成するときに、各特徴の総和を刺激の強さと位置付け、その刺激に対応する感覚の大きさを合成後の値として求めるために、Fechner の法則を用いる。

次に示す(1)では Fechner の法則について述べ、(2)~(4)で対数関数方式を定義する。

(1) Fechner の法則

E.H.Weber は重さの弁別の実験を行い、我々は物の間の差異そのものを知覚するのではなく、物の大きさに対するこの差異の比率を知覚するのであると述べた。

Fechner は Weber が出したこの事実を Weber の法則と名づけた。Weber の法則が一般的に成立すれば、刺激の大きさ(以下、刺激強度)と我々が感じる感覚の大きさについて、

$$d\gamma = k \frac{d\beta}{\beta}, \quad (12)$$

(k : 比例定数, β : 刺激強度, γ : 感覚の大きさ, $d\beta, d\gamma$: 刺激強度と感覚の大きさの微小の増分) が成立するとした。上記の式を積分して

$$\gamma = k(\log \beta - \log b). \quad (13)$$

(但し $\log b$ は積分定数)

よって、

$$\gamma = k \log \frac{\beta}{b}, \quad (14)$$

となる。Fechner はこれら式について「感覚の大きさ γ は、

メディアデータ $P = \{2.4w_1, 3.2w_2, \dots\}$

※ w_1, w_2, \dots : メディアデータの印象を表す印象語

印象語: w_1, w_2
 $w_1: f_1, f_2, f_4, f_6$
 $w_2: f_3, f_4, -f_5$
 ※ f_1, f_2, \dots, f_6 : 特徴

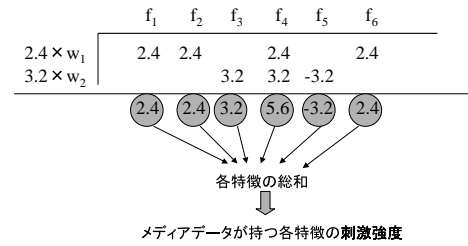


図6 刺激強度と印象語の特徴の関係。

刺激の絶対的大きさではなく刺激の大きさの対数で比例する。ここで刺激の大きさは閾値 b 、すなわち、感覚が生じかつ消失する刺激の単位としている。要するに、感覚の大きさは基本的刺激値の対数に比例する。」と説明している。これがその後 Fechner の法則と呼ばれている。

(2) 刺激強度と印象語の特徴の関係

本方式における、刺激強度と印象語の特徴の関係は次のように位置付けられる。

印象語 w_i に特徴づけられたひとつひとつの特徴をその特徴における刺激とみることができる。各特徴の総和を求めることにより、メディアデータがもっている各特徴における刺激の強さを求めていると考えられる。従って、印象語 w_i における各特徴の総和は各特徴の刺激強度として意味づけができる(図6)。

(3) 対数関数方式の導出

刺激強度(特徴の総和)が1のとき感覚の大きさは1であると定義する。これより式(13)に刺激強度 $\beta = 1$ 、感覚の大きさ $\gamma = 1$ を代入して、

$$1 = k(\log 1 - \log b). \quad (15)$$

$$\log b = -\frac{1}{k} \quad (k \neq 0). \quad (16)$$

これを式(13)に代入すると、

$$\gamma = k \log \beta + 1, \quad (17)$$

となる。刺激強度が0の場合、感覚の大きさは0と定義する。また刺激強度 $\beta < 0$ の場合以下のように定義する。

$$\gamma = -(k \log |\beta| + 1). \quad (18)$$

3.3.2(2)節で示した刺激強度と印象語の特徴の関係、式(17)、および式(18)より、メディアデータ P に式(9)のように付与する印象語 w_i が式(10)のように特徴づけられている場合、対数関数方式を実現する演算子 \oplus_F を次のように定義する。

$$\bigoplus_{i=1}^t F W_i := (\gamma_1, \gamma_2, \dots, \gamma_n).$$

$$\gamma_j = \begin{cases} k \log_{\alpha} |\beta_j| + 1 & (\beta_j > 0 \text{ のとき}) \\ 0 & (\beta_j = 0 \text{ のとき}) \\ -(k \log_{\alpha} |\beta_j| + 1) & (\beta_j < 0 \text{ のとき}) \end{cases}$$

$$\beta_j = \sum_{i=1}^t f_{ij}.$$

(19)

但し、 β が 0 に近い値の特徴はメディアデータの印象を的確に表していないため検索結果を悪化させる恐れがあることから、これらの特徴を排除する必要がある。

(4) 感性パラメータ

式 (19) において k を感覚ボリュウム係数とし、メディアデータの違いや検索者個々の感性の違いによって設定するパラメータとする。また式 (19) の対数の底の値 α についても同様の役割を果たすパラメータとする。これら二つのパラメータ (k, α) を感性パラメータとし感性のボリュウムとして設定できるパラメータとする。

3.2 画像メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式の実現方法

本方式の全体図を図 7 に示す。画像メディアデータを対象とした感性作用素を有するメタデータ自動抽出方式は次のような手順によって実現される。

- Step1: 重みつき印象語の出力 (従来の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)
文献^{5),8)}に示されている方式によりデジタル化された静止画像データからメディアデータの印象を抽出し重み (静止画像と印象語の相関) つき印象語が出力される。
- Step2: 不要な特徴の排除 (前処理)
画像メタデータ自動抽出方式により出力される重みつき印象語は特徴づけを行うことにより、 n 次元ベクトルとして表されている。重みが 0 に近い値の特徴は 3.1.2(3) 節より、メディアデータの印象を的確に表していないため検索結果を悪化させる恐れがある。これらの検索結果を悪化させる恐れがある特徴を排除する。
- Step3: 印象語合成 (感性作用素の適用)
特徴づけされた n 次元ベクトルで表された各印象語を合成することによって言葉とメディアデータの動的な関係を空間上での距離として計算できるメタデータ空間 MDS 上に写像される。

3.1 節より、メタデータ空間 MDS に写像する方式として、次の 2 方式が挙げられる。

- (1) 従来方式
人間の感覚を反映しない従来的方式
- (2) 対数関数方式
人間の感覚を反映した感性作用素として作用する方式

4. 実験

本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、検証実験を行った。

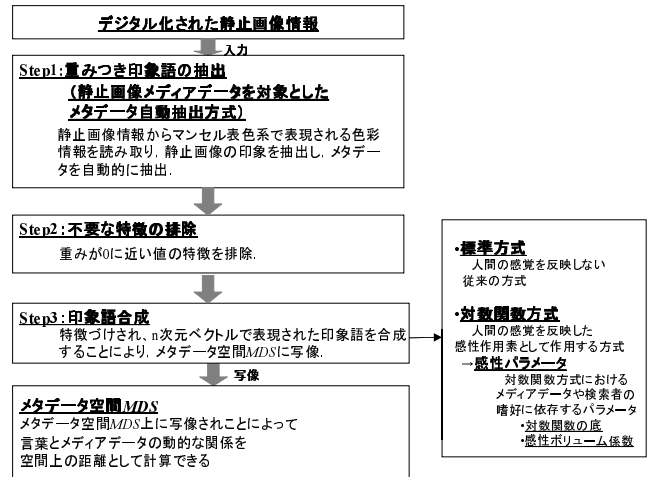


図 7 提案方式の全体図。

実験 1 では感性作用素として作用する対数関数方式の感性パラメータ (4.1.2(4) 節) の変化による意味的連想検索の検証を行い、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に適切な感性パラメータについて考察した。

実験 2 では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において、感性作用素として作用しない従来方式を用いた場合と、感性作用素として作用する対数関数方式を用いた場合との比較を仮想データを用いて行い、感性作用素の有効性の検証を行った。

実験 3 では実験 2 について行った実験を実際の画像データを用いて行い、感性作用素の有効性の検証を行った。

4.1 実験環境

メタデータ空間 MDS 作成については、"Longman Dictionary of Contemporary English"¹⁰⁾ という英英辞書を使用した。同辞書は、約 2000 語の基本語だけを用いて約 56000 語の見だし語を説明している。ここで基本語を特徴とみなし、各見出し語を説明する基本語が肯定の意味に用いられていた場合 "1"、否定の場合 "-1"、使用されていない場合 "0"、見出し語自身が基本語である場合その基本語の要素を "1" として、2.1(1) 節のデータ行列 M を作成した。これより、約 2000 次元の正規直交空間であるメタデータ空間 MDS を生成した。約 2000 次元のメタデータ空間 MDS では、約 2^{2000} 通りの意味の様相が表現可能である。

対象の画像メディアデータとして、風景写真 20 枚をデータセット-A として用意した。データセット-A の画像メディアデータを図 8 に示す。また、風景写真 35 枚をデータセット-B として用意した。データセット-B の画像メディアデータを図 9 に示す。さらに、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から出力されるメタデータを想定して手動で作成した仮想データを用意した。この仮想データは、「warm, active」を正解とする 100 データ、「quiet, fresh」を正解とする 100 データ、「dark, heavy」を正解とする 100 データを含む 1000 データからなる。

4.2 実験システム

実験システムの全体図を図 10 に示す。

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。



図 8 対象とした画像メディアデータセット-A.
Fig. 8 Image data set-A for experiments.

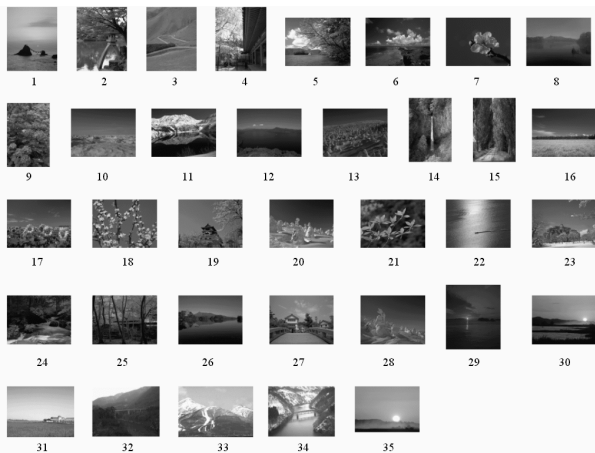


図 9 対象とした画像メディアデータセット-B.

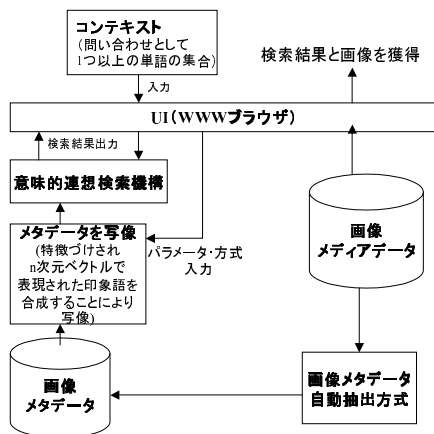


図 10 実験システム図.
Fig. 10 A figure of the system structure for experiments.

4.3 評価方法

本実験では図 8, 9 で示した画像メディアデータを対象として、意味的連想検索を行い検索結果を考察する。コンテキストとして、比較的検索者が使用すると想定される印象語を設定する。それに対する正解を画像の全体的な色彩と、カラーイメージスケール¹²⁾でそのコンテキストと印象が近い色彩を持つ画像に設定し

表 1 データセット-A のコンテキストと正解.

Table 1 Contexts and correct answer of data set-A.

コンテキスト	正解(画像 ID)
warm active	a, b, c, d, e
quiet fresh	f, g, h, i, j
dark heavy	k, l, m, n, o

表 2 データセット-B のコンテキストと正解.

Table 2 Contexts and correct answer of data set-B.

コンテキスト	正解(画像 ID)
warm active	1,16,22,29,30,31,35
quiet fresh	3,10,11,20,23,28,31,33,34
dark heavy	13,29,32

た。データセット-A のコンテキストと正解は表 1 に示す。また、データセット-B のコンテキストと正解は表 2 に示す。

また、本実験では検索結果を再現率、適合率という指標によって評価した。ここで再現率、適合率は以下のように表される。

$$\text{再現率} = \frac{\text{システムの検索結果に含まれる正解数}}{\text{本来の正解数 (正解として出力されるべきデータ数)}} \quad (20)$$

$$\text{適合率} = \frac{\text{システムの検索結果に含まれる正解数}}{\text{システムの検索結果出力数}} \quad (21)$$

4.4 実験 1

4.4.1 実験方法

画像メタデータ自動抽出方式から出力されたメタデータを対象として、感性作用素として作用する対数関数方式の感性パラメータ (4.1.2(4) 節) の変化による意味的連想検索を行い、検索結果の比較を行った。感覚ボリューム係数 $k = 1$ に固定し、底 α の変化による R 適合率の比較を行った。また、先の実験の R 適合率が一番高い値の底 α に固定し、感覚ボリューム係数 k の変化による R 適合率の比較を行った。実験対象データとして、データセット-A、および B をそれぞれ使用し、実験、比較を行った。

R 適合率とは、式 (20)、式 (21) のシステムの検索結果出力数を本来の正解数と同値にした場合の適合率を指し、再現率と適合率は同値になる。本実験ではシステムの検索結果出力数は表 1, 2 の正解数と同数とした。

さらに、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に適切な感性パラメータについて考察した。

4.4.2 実験結果

データセット-A を対象として、感覚ボリューム係数 $k = 1$ に固定し、底 α の変化による R 適合率の値の関係を図 11 に示す。但し、R 適合率の値は、表 1 で表す各コンテキストによる検索結果の R 適合率の平均値を表している。

これより、底 $\alpha = 8$ の前後で R 適合率が高くなっていることが分かる。

さらに、データセット-A を対象として、底 $\alpha = 8$ に固定し、感覚ボリューム係数 k の変化による R 適合率の値の関係を図 12 に示す。

以上から、データセット-A を対象とした場合、感覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ 前後で R 適合率が高くなることが分かる。

同様にデータセット-B を対象として、感覚ボリューム係数 $k = 1$

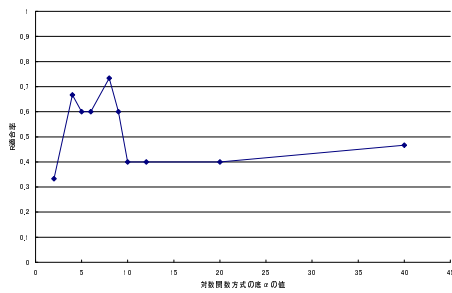


図 11 底 α の値の R 適合率の変化 (データセット-A).

Fig. 11 An alteration of R-Precision rates by the base α (data set-A).

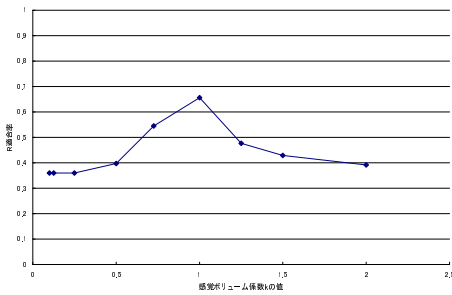


図 12 感覚ボリューム係数 k の R 適合率の変化 (データセット-A)(実験 1-1).

Fig. 12 An alteration of R-Precision rates by the sense volume value k (data set-A).

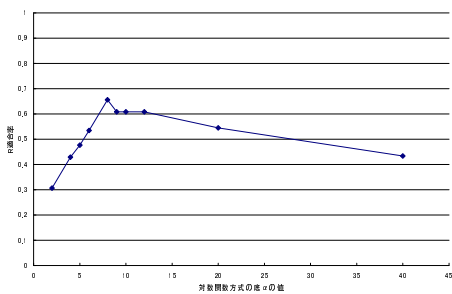


図 13 底 α の値の R 適合率の変化 (データセット-B)(実験 1-2).

Fig. 13 An alteration of R-Precision rates by the base α (data set-B).

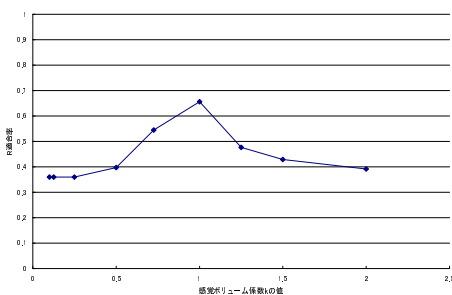


図 14 感覚ボリューム係数 k の R 適合率の変化 (データセット-B).

Fig. 14 An alteration of R-Precision rates by the sense volume value k (data set-B).

に固定し、底 α の変化による R 適合率の値の関係を図 13 に示す。

これより、底 $\alpha = 8$ の前後で R 適合率が高くなっていることが分かる。

さらに、データセット-B を対象として、底 $\alpha = 8$ に固定し、感覚ボリューム係数 k の変化による R 適合率の値の関係を図 14 に示す。

以上から、データセット-B を対象とした場合についても、感

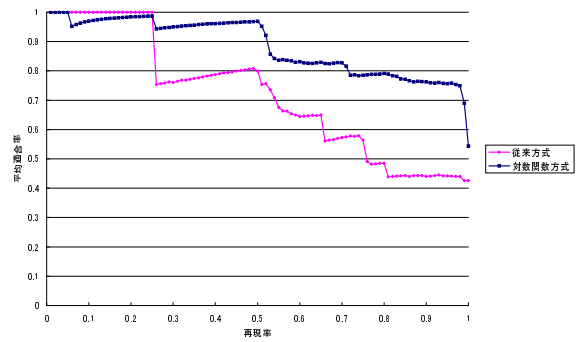


図 15 再現率・平均適合率 (実験 2).

Fig. 15 Recall rates and precision rates(experiment 2).

覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ 前後で R 適合率が高くなることが分かる。

4.4.3 考察

画像メディアデータの場合、感覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ 前後で設定するとよいことが分かった。さらに、データセット-A, B について同様の結果が出たことから、感性パラメータはメディアデータの種類の依存したパラメータであると考えられる。これにより、このパラメータはメディアデータの種類によって設定することで、検索結果を検索者の意図に近づけることが可能であることを示している。

4.5 実験 2

4.5.1 実験方法

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において、感性作用素として作用しない従来方式を用いてメタデータをメタデータ空間 MDS に写像した場合と、感性作用素として作用する対数関数方式を用いてメタデータを写像した場合とで、それぞれ意味的連想検索を行い比較した。実験対象データとして、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から出力されるデータを想定して手動で作成した仮想データ 1000 データを使用した。対数関数方式の感性パラメータについては、実験 1 の結果から、感覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ に設定する。

4.5.2 実験結果

感性作用素として作用しない従来方式を用いてメタデータをメタデータ空間 MDS に写像した場合と、感性作用素として作用する対数関数方式を用いてメタデータを写像した場合の再現率・平均適合率を表したグラフを図 15 に示す。平均適合率とは、各コンテキストにおける適合率の平均である。

全体的に、対数関数方式が従来方式に比べ、平均適合率が高くなっている。

4.5.3 考察

感性作用素として作用する対数関数方式では、従来方式に比べ、全体的に平均適合率高いことが分かった。

これより、人間の感性を反映した関数である、感性作用素を導入することにより、人間の感性に合致した検索が可能となることを示している。

4.6 実験 3

4.6.1 実験方法

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において、感性作用素として作用しない従来方式を用いてメタデータをメタデータ空間 MDS に写像した場合と、感性作用素とし

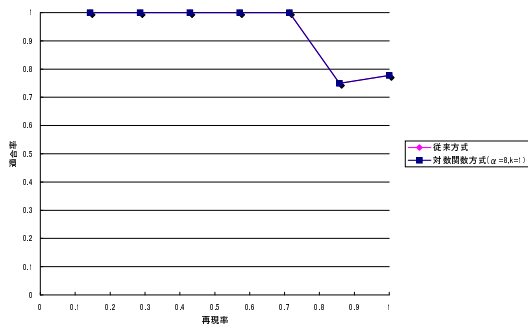


図 16 コンテキスト「warm active」の再現率・適合率 (実験 3-1).
Fig. 16 Recall rates and precision rates for the context "warm active" (experiment 3-1).

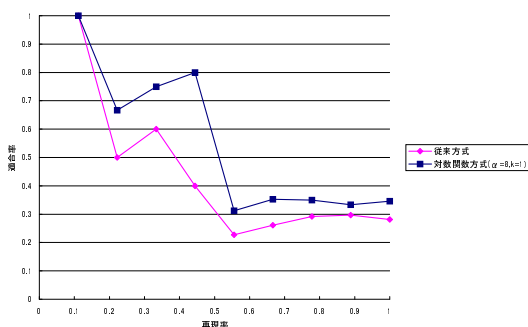


図 17 コンテキスト「quiet fresh」の再現率・適合率 (実験 3-2).
Fig. 17 Recall rates and precision rates for the context "quiet fresh" (experiment 3-2).

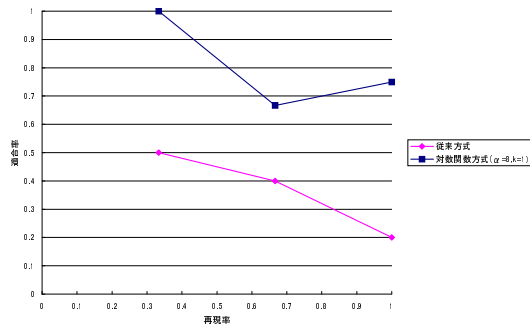


図 18 コンテキスト「dark heavy」の再現率・適合率 (実験 3-3).
Fig. 18 Recall rates and precision rates for the context "dark heavy" (experiment 3-3).

て作用する対数関数方式を用いてメタデータを写像した場合との比較を実際の画像メディアデータを用いて行った。実験対象データとして、データセット-Bを使用した。対数関数方式の感性パラメータについては、実験1の結果から、感覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ に設定する。

4.6.2 実験結果

感性作用素として作用しない従来方式を用いてメタデータをメタデータ空間 MDS に写像した場合と、感性作用素として作用する対数関数方式を用いてメタデータを写像した場合の再現率・適合率を表したグラフをコンテキスト「warm active」の場合、コンテキスト「quiet fresh」の場合、コンテキスト「dark heavy」の場合をそれぞれ、図 16、図 17、図 18 に示す。

コンテキスト「warm active」の場合では、従来方式と対数関数方式で各再現率において、適合率の値が変わらないが、コンテキスト「quiet fresh」, 「dark heavy」の場合では対数関数方式が従来方式に比べ、適合率が高くなっている。

4.6.3 考察

適切な感性パラメータを設定した感性作用素として作用する対数関数方式では、従来方式でも検索結果がよい場合はよいままで、従来方式であまり検索結果がよくない場合が大いに改善されることが分かった。

これより、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においても人間の感性を反映した関数である、感性作用素を導入することにより、人間の感性に合致した検索が可能となることを示している。

4.7 実験全体の考察

実験 1 では、画像メディアデータを対象とした感性作用素の

適切な感性パラメータの値として感覚ボリューム係数 $k = 1$ 、底 $\alpha = 8$ 前後で設定するとよいことを示した。

実験 2 では、画像メディアデータを対象とした感性パラメータを適切な値に設定した感性作用素として作用する対数関数方式を導入することで、より人間の直感に合致した検索が可能になることを示した。

さらに実験 3 では、感性パラメータを適切な値に設定した感性作用素として作用する対数関数方式が実際の画像データを対象としたメタデータ自動抽出方式への適用においても有効であることを示した。

この実験は、画像メディアデータを対象とした人間の感性を反映した関数である感性作用素を有するメタデータ自動抽出方式の有効性を示している。

5. おわりに

本稿では、人間の直感に応じた重みつき単語群によるメタデータの自動抽出を実現する感性作用素を、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に適用する方式を示した。本方式により、画像メディアデータの自動抽出方式により出力される重みつき単語群を対象として、人間の感性を反映した関数を導入することにより人間の直感に対応するメタデータを抽出可能となった。また、画像メディアデータを対象とした実験を行い、本方式の有効性を示した。

本方式により、画像メディアデータや楽曲メディアデータだけでなく、任意の種類メディアデータのメタデータ自動抽出方式から抽出されるメタデータを対象として、同様の方式によりメタデータ空間 MDS 上に写像でき、意味的連想検索をメディアデータの性質や検索者個々の嗜好の違いを吸収できると考えられる。

今後の課題として、検索対象データ量を増やし、大規模な画像データベースでの実験による本方式の詳しい検索特性の考察、本方式の各種メディアデータへの適用、メタデータ自動抽出方式への個人差の計量方式の導入、および、本方式による異種メディア間検索方式の実現が挙げられる。

参考文献

- 1) M.W.Berry, S.T.Dumains, G.W.O'Brien, "Using linear algebra for intelligent information retrieval", SIAM Review Vol. 37, No.4, pp.573-595, 1995.
- 2) "Retrieval Evaluation", Chapter 3 in R.Baeza-Yates

- and B.Ribeiro-Neto(editors), "Modern Information Retrieval", Addison-Wesley, 1999.
- 3) T.Kitagawa, Y.Kiyoki, "Fundamental framework for media data retrieval system using media lexco transformation operator", Information Modeling and Knowledge Bases, IOS Press, 2000.
 - 4) T.Kitagawa, Y.Kiyoki, "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, April 1993
 - 5) Y.Kiyoki, T.Kitagawa, H.Takanari, "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, Amit Sheth and Wolfgang Klas(editors), Chapter 7, 1998.
 - 6) 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.
 - 7) 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, 98-DBS-116(2), pp.109-116, 1998.
 - 8) 小谷拓矢, 清木康, 北川高嗣, "色彩情報による静止画像メタデータ生成方式と意味的画像検索への適用", 第9回データ工学ワークショップ (DEWS'98) 論文集, 電子情報通信学会, 1998.
 - 9) 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 (2002), accepted.
 - 10) "Longman Dictionary of Contemporary English", Longman, 1987.
 - 11) 千々岩英彰, "色彩学概説", 東京大学出版会, 2001.
 - 12) 小林重順, "カラーイメージスケール", 講談社, 1984.
 - 13) "新編 感覚・知覚心理学ハンドブック", 誠信書房, 1994.
-