

画像データベースを対象とした意味的連想検索の高速化アルゴリズム

宮川 明子[†] 清木 康^{††}
宮原 隆行^{†††} 北川 高嗣^{†††}

画像データなどのメディアデータを対象としたデータベースシステムの実現において、検索者が求めるメディアデータを適切かつ高速に抽出することは重要である。本論文では、意味の数学モデルを拡張した意味的画像連想検索を対象とした高速化アルゴリズムの実現方式を提案する。意味の数学モデルは、文脈あるいは状況に応じて動的に変化するデータ間の意味的な関係を計算するモデルである。本論文で提案するアルゴリズムは、指定された時間内の限られた計算回数で有効な検索結果を得ることを目的としたものである。このアルゴリズムによる意味的画像検索の実験を行い、実験の結果よりその有効性を明らかにした。

A Fast Semantic Associative Search Algorithm for Image Databases

AKIKO MIYAGAWA,[†] YASUSHI KIYOKI,^{††}
TAKAYUKI MIYAHARA^{†††} and TAKASHI KITAGAWA^{†††}

In the database system design for multimedia databases, it is important to develop a correct and fast retrieval method for media data(e.g. image, music). This paper proposes a fast algorithm and its implementation method for semantic associative image search based on our mathematical model of meaning. This model has been designed for computing semantic relation between data items dynamically according to context and situation.

The objective of this algorithm is to obtain the available and correct retrieval results within the given limited time in the semantic associative search. This paper also shows some experimental results of semantic image search to clarify the feasibility and effectiveness of the proposed algorithm.

1. はじめに

近年、計算機上において大量のマルチメディアデータを対象とした処理が行われるようになってきている。従来の文字・数値型データと比べ、メディアデータは、表現方法が無数にあり内包する情報量が非常に多いこと、扱うデータ量が膨大であることが特徴である。したがって、マルチメディアデータベースの分野では、膨大なメディアデータから適切な結果を効率的に獲得するための高速化の実現が重要である。

本論文では、メディアデータの一例として画像データを対象に、画像の印象や内容を表す抽象情報を利用した

意味的連想検索方式²⁾における高速化アルゴリズムを提案する。本検索方式は、文脈あるいは状況に応じて動的に変化するデータ間の意味的な関係を計算するモデルである意味の数学モデル⁴⁾を拡張したデータベース検索方式であり、検索者の印象を文脈として与えることにより画像などのメディアデータを対象とした意味的連想検索を実現する。

画像データベースを対象に、ユーザの印象にもとづいた検索を行う方式に関する研究の一つである文献⁶⁾の方法では、正準相関分析により、各ユーザの印象語と画像の特徴との相関関係を学習し、ユーザの印象による画像検索を可能としている。この方法は、画像の特徴を表わす静的に固定されたベクトルとユーザの印象語を表わすベクトルの相関から問い合わせ空間を生成し、その空間上での計算により画像データを抽出する方法である。また、文献⁷⁾の方法は、百科事典の全テキストの頻度多次元意味ベクトルを構築し、単語間の関連性をあらかじめ設定された特徴の分類項目によって判断しており、画像の説明文から作成した画像の意味ベクトルと検索語の意

[†] 慶應義塾大学 政策・メディア研究科

Graduate School of Media and Governance, Keio University

^{††} 慶應義塾大学 環境情報学部

Faculty of Environmental Information, Keio University

^{†††} 筑波大学 電子・情報工学系

Institute of Information Sciences and Electronics, University of Tsukuba

味ベクトルとの照合による画像検索を行っている。

一方、本意味的連想検索方式の特徴は、ユーザが与える印象に対応する特徴を直交空間からの部分空間選択という操作により抽出し、その特徴のみに着目した相関の強さから画像検索を行うことが可能な点にある。現在の実現システムでは、約 2^{800} 通りのほぼ無限に近い数の文脈 (印象) を表現することができる。

本論文で提案する高速化アルゴリズムの目的は、検索者が限られた時間内に検索結果を必要とする場合に、検索者の与えた文脈と相関の強いことが予想される画像を優先的に計算し、その時間内に適切な解を導くことである。

高速化手法としては、検索対象データがベクトル空間上の座標点として、固定的な位置に存在することを利用して、データのグループ化、索引付けを行う R-tree 構造¹⁾を用いるものが知られている。また、文献^{13),14)}において、ドキュメント群を検索対象とし、検索語群とドキュメントの相関の強さにしたがってドキュメント群をランキングするためのインデックスの実現方法が示されている。それら方法における相関量の計算では、相関量計算におけるベクトル (検索語群およびドキュメントに対応するベクトル) の特定要素に関する相関量を求めるのではなく、検索語群と各ドキュメント間の全体を考慮した相関量計算を行なっている。さらに、各要素に関する重みづけは静的に設定できるので、あらかじめ、全要素を対象としたインデックスを設定しておくことができ、その結果として高速化を実現している。

意味の数学モデルによる本意味的連想検索本式では、問い合わせによって与えられる特定の意味的要素に着目した相関量計算を行なう。検索者が文脈として与える検索対象データの印象に応じて動的に部分空間を選択し、その部分空間内に検索対象データ群を動的にマッピングすることにより、検索対象データ間の空間的な位置関係を決定するので、検索対象データ間の関係は文脈に応じて動的に変化する。動的に選択される部分空間、およびその部分空間中での軸の重みを考慮した相関量計算の高速化を実現するためには、インデックスを用意することなく、少ない相関量計算量で、問い合わせとの相関量が上位となる検索対象メディアデータを抽出する高速化アルゴリズムを考案することが必要である。したがって、本論文では、静的なインデックスによる検索の高速化ではなく、本画像検索方式に対応する動的な位置関係を前提としたアルゴリズムによる高速化方式を提案する。

本研究に関連する成果として、我々は、与えられた文脈にしたがって決定された部分空間上での単語間の距離を計算し、検索キーワードと意味的に関連する検索対象

データを高速に検索するための高速化アルゴリズム⁹⁾をすでに提案した。文献⁹⁾では、意味的な等価性、類似性に関する計算によりデータベース検索を行う意味的連想検索方式の高速化アルゴリズムを示している。このアルゴリズムは、検索キーワードおよびその文脈を説明する文脈語列を受けとることによって、その検索キーワードに意味的に近い検索対象データを高速に抽出するために用いられる。検索キーワードおよび文脈語列が与えられると、その文脈に対応する正規直交空間の部分空間を動的に選択する。この連想検索方式では、文脈理解を正規直交空間における部分空間選択によって実現し、与えられた文脈、すなわち、選ばれた部分空間において検索キーワードのベクトルと検索対象語ベクトルの位置により、その部分空間において、検索キーワードに距離が最も近い検索対象データを意味的に最も近いデータとして動的に抽出する。

文献⁹⁾の高速化アルゴリズムが、検索キーワードおよび文脈語列を受けとることによって、その文脈において検索キーワードに距離が近い検索対象データを高速に抽出するアルゴリズムであるのに対し、本論文で提案する高速化アルゴリズムは、検索時に問い合わせとして画像の印象や内容を表す文脈語列のみを受けとり、その文脈に対応する部分空間において、その部分空間上でのノルムが大きい検索対象データを、文脈との相関が強い検索対象データとして抽出するアルゴリズムである。したがって、前者が、部分空間上での2点間 (検索キーワードと検索対象データ) の距離を各検索対象データについて計算するのに対し、後者は、部分空間上での各検索対象データのノルムを計算することによって検索すべきデータを選択するので、それらにおける高速化の方法は異なるものとなる。

2. 意味の数学モデルを用いた意味的画像検索方式

2.1 概要

意味の数学モデルは、本来、言葉の意味を扱うためのモデルである。ここでは、印象や画像の内容を表現している文脈に対応した画像を選び出すために拡張した意味的画像検索方式について、概要を述べる。本方式については、文献^{2)~4)}において、詳しく述べている。

- (1) 前提： いくつかの単語を特徴づけたデータの集合が、 m 行 n 列の行列 (以下、“データ行列”と呼ぶ) の形で与えられているものとする。この行列において、 m 個のそれぞれの単語 (word) は、 n 個の特徴 (features) によって特徴づけられている。
- (2) イメージ空間 I の設定： データ行列から、特徴

づけに関する相関行列をつくる．そして，相関行列を固有値分解し，固有ベクトルを正規化する．相関行列の対称性から，このすべての固有値は実数であり，その固有ベクトルは互いに直交している．このとき，非ゼロ固有値に対応する固有ベクトル（以下，“意味素”と呼ぶ）の張る正規直交空間をイメージ空間 \mathcal{I} と定義する．この空間の次元 ν は，データ行列のランクに一致する．また，この空間は， ν 次元ユークリッド空間となる．

(3) 意味射影の集合 Π_ν の設定： イメージ空間 \mathcal{I} から固有（不変）部分空間への射影（以下，“意味射影”と呼ぶ）の集合 Π_ν を考える． i 次元の部分空間は， $\frac{\nu(\nu-1)\cdots(\nu-i+1)}{i!}$ ($i = 1, 2, \dots, \nu$) 個存在するので，射影の総数は 2^ν となる．つまり，このモデルは 2^ν 通りの意味の様相の表現能力を持つ．

(4) 意味解釈オペレータ S_p の構成： 文脈を決定する ℓ 個の文脈語列 s_ℓ としきい値 ϵ_s が与えられたとする．このとき，その文脈に応じた意味射影 $P_{\epsilon_s}(s_\ell)$ を決めるオペレータ（以下，“意味解釈オペレータ”と呼ぶ） S_p を次のように構成する．

(a) 文脈語列 s_ℓ を構成する ℓ 個の単語を各々イメージ空間 \mathcal{I} へ写像する．この写像では， ℓ 個の単語を各々イメージ空間 \mathcal{I} 内でフーリエ展開し，フーリエ係数を求める．ここでいうフーリエ展開とは，各単語に対応するベクトルと各意味素間での内積を計算することに対応し，各単語と各意味素の相関を求めることを意味する．

(b) 各意味素ごとに，フーリエ係数の総和を求める．これは，文脈語列 s_ℓ と各意味素との相関を求めることに相当する．このベクトルは， ν 個の意味素があるため， ν 次元ベクトルとなる．さらに，このベクトルを，無限大ノルムによって正規化する．（以下，文脈語列 s_ℓ の意味重心と呼ぶ．）

(c) このとき，文脈語列 s_ℓ の意味重心を構成する各要素において，しきい値 ϵ_s を超える要素に対応する意味素を，単語を射影する部分空間の構成に用いる．これにより，意味射影 $P_{\epsilon_s}(s_\ell)$ を決定する．部分空間を構成している各意味素の要素が，その意味素の重みとなる．

このオペレータは，文脈語列と相関の高い部分空間の自動的な選択を実現する．

(5) 部分空間における距離計算： 文脈語列 s_ℓ により，各意味素ごとに重みを定める．イメージ空間に写像された各画像のメタデータについて，部分

空間上におけるそのメタデータのノルムを文脈に対する画像の相関量とする．本方式では，部分空間においてノルムの大きい画像ほど，その文脈に強く関連しているものとする．

このモデルにより，文脈に応じた単語間の関係の解釈（部分空間の選択，およびその空間内における最良近似）が可能となる．

2.2 定式化

2.2.1 イメージ空間 \mathcal{I} の設定

各データ間の意味的な関係を計算するための基盤となる空間の生成を行う．

空間生成用メタデータとして， m 個の単語について各々 n 個の特徴 (f_1, f_2, \dots, f_n) を列挙した各単語に対する特徴付きベクトル \mathbf{w}_i ($i = 1, \dots, m$) を用意する．そのベクトルを並べた m 行 n 列のデータ行列を M とする（図 1）．

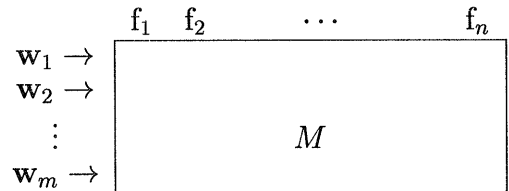


図 1 データ行列 M の構成
Fig. 1 Data Matrix M

- (1) データ行列 M の相関行列 $M^T M$ を作る．
- (2) $M^T M$ を固有値分解する．

$$M^T M = Q \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_\nu & & \\ & & & 0 \cdots 0 & \\ & & & & \ddots \end{pmatrix} Q^T,$$

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

ここで行列 Q は，

$$Q = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)^T$$

である．この \mathbf{q}_i は，相関行列の固有ベクトル，つまり意味素である．

- (3) このとき，イメージ空間 \mathcal{I} を以下のように定義する．

$$\mathcal{I} := \text{span}(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_\nu).$$

$(\mathbf{q}_1, \dots, \mathbf{q}_\nu)$ は \mathcal{I} の正規直交基底であり，この空間は， ν 次元ユークリッド空間となる．

2.2.2 画像データと検索語の n 次元ベクトル表現

2.2.1 節と同一の n 個の特徴を用いて，個々の画像，および，検索者の印象を表す文脈語列（印象語列）をそれぞれ n 次元ベクトルで表現し，画像のメタデータと印象

語のメタデータを作成する。これによって、各メタデータはイメージ空間 \mathcal{I} 上の一点に写像することが可能になる。

以下に、単語のメタデータ化を拡張した画像のメタデータ化の方法について概説する。

画像についての検索者の印象、および、画像に描かれている内容を合わせて、以下、オブジェクトと呼ぶ。

(1) 画像からオブジェクトの抽出

t 個のオブジェクト $\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t$ から成る画像 P を次のように定義する：

$$P = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t\}.$$

ここで、各オブジェクト \mathbf{o}_i は、データ行列 M の特徴と同一の特徴を用いて表現されている特徴付ベクトル $\mathbf{o}_i = (o_{i1}, o_{i2}, \dots, o_{in})$ である。

(2) 画像 P のベクトル表現

オブジェクト $\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t$ の和演算子 \oplus を次のように定義する：

$$\begin{aligned} \bigoplus_{i=1}^t \mathbf{o}_i &\equiv (\text{sign}(o_{\ell_1 1}) \max_{1 \leq i \leq t} |o_{i1}|, \\ &\text{sign}(o_{\ell_2 2}) \max_{1 \leq i \leq t} |o_{i2}|, \\ &\dots, \text{sign}(o_{\ell_n n}) \max_{1 \leq i \leq t} |o_{in}|) \end{aligned}$$

ただし、 $\text{sign}(a)$ は、“ a ”の符号（正または負）を表し、 $\ell_k (k = 1, \dots, n)$ は、 $|o_{ik}|$ の最大値を与える添字 $i (i = 1, \dots, t)$ を表している。

2.2.3 意味的画像検索

意味の数学モデルを適用した意味的画像検索は、以下の2つの機能にしたがって実現されている。

基本機能 1: 部分空間選択

イメージ空間から、画像の印象を表す文脈語列（印象語列） $s_\ell = (u_1, \dots, u_\ell)$ (u_1 から u_ℓ までの単語ベクトルは M 行列を作成したものと同一特徴で特徴付けられている）に対応した固有部分空間を選択する。

Step-1 : 文脈ベクトルを構成する単語ベクトルのフーリエ展開:

u_1 から u_ℓ までの各単語ベクトルについてフーリエ展開を行い、各々の意味素（固有ベクトル）との相関を計算する。

Step-2 : フーリエ展開によって得られた値を、各意味素（固有ベクトル）ごとに合算する:

文脈語と意味素の相関を計算するために、フーリエ展開によって得られた値を各々の意味素ごとに合算することにより意味重心ベクトルを形成する。

Step-3 : 部分空間の決定:

Step-2 で求めた意味重心ベクトルの各意味素上での値の絶対値が、与えられたしきい値 ϵ_s より大き

い意味素群を選択し、部分空間を構成する。

基本機能 2: 部分空間における意味的相関量の計算

イメージ空間に写像された画像のメタデータの部分空間上におけるノルムを、文脈に対する画像の相関量とする。本方式では、部分空間においてノルムの大きい画像ほど、その文脈に強く関連しているものとする。

画像のメタデータ $\mathbf{x} \in \mathcal{I}$ のノルム、以下のように定義する：

$$\eta(\mathbf{x}; s_\ell) = \frac{\sqrt{\sum_{j \in \Lambda_{\epsilon_s} \cap S} \{c_j(s_\ell) x_j\}^2}}{\|\mathbf{x}\|_2},$$

ただし、 $S = \{i | \text{sign}(c_i(s_\ell)) = \text{sign}(x_i)\}$ とする。

ここで、 $c_j(s_\ell)$ は、文脈を決定する言葉の列 s_ℓ に依存して決まる重みを表す²⁾。

上式では、単純なノルムによる計算式に正規化と符合の考慮という二つの機能が加わっている。これについては、文献^{5),8)}の実験の項でそれらの効果が実証されている。

一般的に画像を定義する特徴が多いほど意味素上の値が積み重なることから、部分空間における画像メタデータのベクトルのノルムは大きくなる傾向にある。そのため、本来意味的な関連が強いと思われる画像データよりも特徴の数が多い画像データの影響で、正しい解釈が行われなことがある。したがって、上記の計算式では、画像のメタデータを毎に2ノルムで正規化している。また、ある画像データのノルムを計算する際、データの軸上のベクトル値が、選択された部分空間を構成している軸と正負が逆の場合には、その値をノルムの計算に採用しないこととしている。

上式で各画像データのノルムを計算した結果から、意味的相関量の大きな画像データを出力解として導く。表1に、100画像を対象に、検索語に dynamic, power を指定しダイナミックで力強い印象を持つ画像を検索した結果を上位3画像について示す。2位に選ばれた画像は、定義した印象に検索語 dynamic, power の両単語とも含まれていないが、force, fight が定義されていることにより、意味的相関が強い画像として適切に選択されていることがわかる。

3. 意味的画像検索の高速化アルゴリズム

本画像検索方式では、検索者が与える文脈に応じて検索対象画像データをその文脈との相関の強い順に並べたリストを生成する。従来の方式では、すべての検索対象画像との意味的な相関量を計算した後、解の出力を行っ

表 1 dynamic, power な印象を持つ画像の検索結果

Table 1 A retrieval result of images associated with dynamic, power

	画像 ID	画像を定義した印象	印象との相関量
1	delacroix	motion dynamic force	0.308474
2	copley	vivid force fight	0.262297
3	schmidt	dynamic lively excite	0.243177

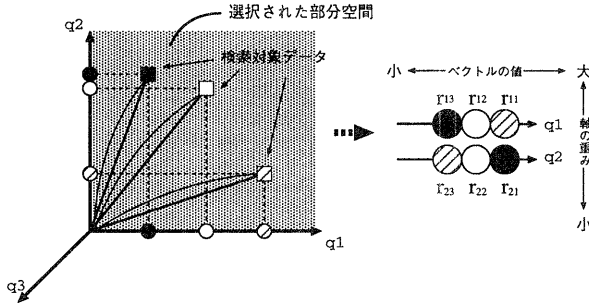


図 2 意味素 q_i 上の検索対象画像データ
Fig. 2 retrieval candidate image data on semantic element q_i

ていた。本稿で提案する 2 つの高速化アルゴリズムは、限られた時間内に解の出力を必要とする場合において、文脈との相関の強いと予想される画像を優先的に計算し、その時間内に適切な解を抽出することを目的としている。つまり、提案アルゴリズムを適用することによって、実時間性を備えた検索環境を提供することが可能になる。

以下に示す 2 つの提案アルゴリズムでは、次の 2 点の優先順位にしたがって計算順序を決定している。

- (1) 選択された部分空間を構成している軸の重み
- (2) 各軸上での画像メタデータのベクトルの値

提案アルゴリズムでは、あらかじめ、検索対象画像データは、各意味素の軸 $q_i (1 \leq i \leq \nu)$ ごとに、各検索対象画像データの軸上での値の大きさに応じて、ソートされていることを前提とする。このソート済みデータは、画像メタデータがイメージ空間に写像されたときに自動的に索引付けされる情報である。検索対象画像データが t 個存在し、ある印象を文脈として与えた時に決定される部分空間を構成する意味素(軸)の数を s 本とするとき、その中で j 番目に重みが大きい意味素 q_j 上で k 番目に大きい値を持つ検索対象画像データを $r_{j,k} (1 \leq j \leq s, 1 \leq k \leq t)$ と表す(図 2)。

与えられた文脈に対して、 i 番目に重みが大きい意味素上で最も大きい値を持つ検索対象画像データを r_{i1} で表し、 i 番目に重みが大きい意味素上で最も小さい値を持つ検索対象画像データを r_{it} で表す。

←小 意味素 q_i 上での値 大→

q_i :	ID 2	ID 5	ID 1	ID 4	ID 3
---------	------	------	------	------	------

図 3 意味素 q_i 上のソートされた画像データの例
Fig. 3 An example of retrieval candidate image data on semantic element q_i

←小 意味素上での値 大→

$r_1(s)$:	14	10	6	3	1	↑ 大 重み ↓ 小
$r_2(s)$:	17	13	9	5	2	
$r_3(s)$:	19	16	12	8	4	
$r_4(s)$:	20	18	15	11	7	

図 4 アルゴリズム A における相関の強さの計算順序例
Fig. 4 An Example of the execution process by the algorithm A applied to image data

図 3 は、検索対象画像データの意味素 q_i 上のソートされた画像データの例である。この例では、“ID 1” から “ID 5” までの 5 つの検索対象画像データが存在し、意味素 q_i 上で最も大きい値を持つ検索対象画像データ r_{i1} は “ID 3” であり、最も小さい値を持つ検索対象画像データ r_{i5} は “ID 2” であることを示している。

3.1 アルゴリズム A

2.2.3 節の相関の強さを求める式から、直観的には重みの大きい意味素上で大きい値を持つ検索対象画像データほど相関量が大きくなり、文脈との相関が強い画像データとして選ばれる可能性が高いと予想される。ゆえに、相関量の計算は、基本的に重みの大きい軸上で大きな値をもつ画像ベクトルから順に進められる。

この考え方を最も単純に実現したのがアルゴリズム A である¹⁰⁾。

- (1) $j_1 + k_1 > j_2 + k_2$ ならば $r_{j_1 k_1}$ よりも $r_{j_2 k_2}$ を先に計算する。
- (2) $j_1 + k_1 = j_2 + k_2$ の時、 $k_1 > k_2$ ならば $r_{j_1 k_1}$ よりも $r_{j_2 k_2}$ を先に計算する。

図 4 は、本アルゴリズムを適用した時の相関の強さの計算順序の例である。この例では、検索対象画像データが 5 つ存在し、文脈に対応した意味素群が 4 本選ばれた時の相関の強さの計算順序を表している。

3.2 アルゴリズム B

アルゴリズム B は、アルゴリズム A をさらに発展させたアルゴリズムである。本アルゴリズムの基本手順を次に示す。

まず、以下の初期設定を行う。

- 比較対象データ集合 Z 、計算済データ集合 D を

NULL とする.

- $j = 1, k = 1$ とする.

比較対象データ集合 Z とは, 次に相関量計算をすべき画像データを決定するために用いるデータ集合であり, 相関量計算が行われた画像データの画像 ID, 印象語との相関量, 計算が行われたときの j と k の値を Z に格納する. 次に相関量計算をすべき画像データは, Z の中で相関量が最大の画像を選択し, その画像の持つ j と k の値によって決定する. Z の中から最大の相関量を持つものとして選択された画像データは, Z から取り除く.

- (1) r_{jk} の画像と印象語との相関量を 2.2.3 節の式によって計算し, 計算の終了した画像データを D に加える. r_{jk} の画像がすでに相関量計算が終了している場合は, そのまま (3) に進む.
- (2) D の画像データを, 相関量の大きい順に並べる.
- (3) (1) の検索対象画像データ r_{jk} を Z に加える.
- (4) Z の中から相関量最大のデータ r_{xy} を選択する. 選ばれた最大データを Z から取り除き, M に加える.
- (5) (4) で選ばれた相関量最大データの位置によって, 前述の優先順位に従い, 次に計算するデータを決定する. このとき, すでに相関量計算が終了している画像は候補に含めない.
 - 選ばれたデータがその意味素の軸上で最も大きな値を持つとき (つまり $y=1$)
 r_{jk} について, $(j, k) = (x, (y+1))$ および $(j, k) = ((x+1), y)$ を計算対象とする.
 - 選ばれたデータが (a) 以外するとき (つまり, $y \neq 1$)
 r_{jk} について, $(j, k) = (x, (y+1))$ を計算対象とする.

以上の操作を, 制限時間まで, もしくは, 計算済データ集合 D に含まれる画像データが要求された出力個数となるまで繰り返し, ソーティングされた D のデータを結果として出力する.

図 5 は, 本アルゴリズムを適用した時の相関量の計算順序の例である. この例では, 検索対象画像データが 3 つ存在し, 文脈に対応した意味素群が 3 本選ばれた時の相関量の計算順序を表している.

以上のようなアルゴリズムに従って計算を進めたときの問題として, 比較対象データとなった画像ベクトルの中に極端に相関量の小さいものが含まれている場合に, 近くに相関量の大きな画像データがあるにもかかわらずその軸上では先に計算を進めることができないという現象に陥る可能性がある. このような現象を引き起こす相関量の値が極端に小さい画像ベクトルのことを, ここで

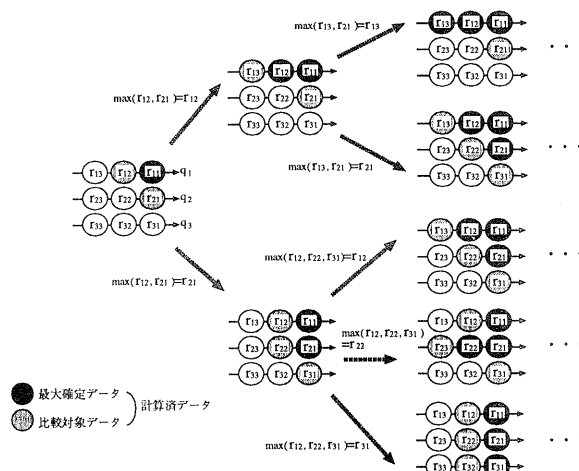


図 5 アルゴリズム B における検索対象画像データの計算順序例
 Fig. 5 An Example of the execution process by the algorithm B applied to image data

は, 極小データと呼ぶ.

そこで, 極小データの出現による計算経路の停滞を回避するため, 本アルゴリズム B を拡張したアルゴリズム B の拡張方式を提案する. 拡張方式では, 複数の検索対象データをまとめて一つのブロックを作り, ブロック単位に計算を進める. 基本方式では, 印象との相関を一つずつ計算し, 計算済みのデータはすべて比較対象データとして蓄えていたが, 本方式では, 1 ブロックに含まれる複数のデータの中から最も相関量の大きいものを比較対象データとして扱う. つまり, 次に計算する検索対象データのブロックを, 既に計算されたブロック内で最大の相関量となったデータ同士の相関量の比較により, 決定する.

また, 拡張方式の 1 ブロックで計算するデータの数を並列計算機のプロセッサの数に応じて設定し, ブロック内の検索対象データの相関量計算を並列に行うことができれば, さらなる高速化を期待できる.

4. 実 験

提案アルゴリズムの有効性を検証するために, 実験システムを構築し, 高速化アルゴリズム A, B の効果に関する実験を行なった.

4.1 実験環境

本実験では, 2 種類の英英辞書を参照した. Longman Dictionary of Contemporary English¹¹⁾ において基本語とされている 2328 単語を, The General Basic English Dictionary¹²⁾ を参照することにより定義した 2328 × 874 のデータ行列を作成し, イメージ空間を生

成した。なお、問い合わせに用いる印象語として 181 単語を用意し、検索対象画像データには 100 画像を印象で定義したものをを用いた。

4.2 実験方法

次の 3 項目に関する実験を行った。

実験 1: 従来の方式に対する高速化アルゴリズムの有効性の検証

実験 2: 極小データが出現する場合におけるアルゴリズム B 拡張方式の有効性の検証

実験 3: アルゴリズム B 拡張方式における並列化の有効性の検証

本実験システムでは、問い合わせにおいて複数の印象語を与えることによって、より検索者の要求に合致した画像を検索することが可能であるが、本実験の目的は、高速化アルゴリズムの有効性を検証することであるため簡略化し、1 回の問い合わせにおいて、検索語として 1 単語のみを使用している。

本実験では、アルゴリズム適用時の出力解の正確性を判断する指標として、正規順位を定義する。正規順位とは、全検索対象画像データの中での相関量の大きさによって決まる順位であり、すべての検索対象画像データについて相関量の計算を行った結果として得られる正解の順位である。ここでは、 R_i を、 i 番目に相関量計算を行った画像データの正規順位として用いる。(100 画像を検索対象としているので、 $i = 1 \dots 100$ である。) 本論文で提案する高速化アルゴリズムは、全画像データについて相関量計算が終了するまでは、必ずしも上位の順位にある正確な解から順に出力するわけではなく、上位となる可能性の高いデータから計算を進めるものである。本実験では、計算順位と正規順位が同じ場合を正解値とし、その正解値との比較において、提案アルゴリズムが正確性を保ちながら計算を進めているのか検証する。

従来方式では、すべての検索対象画像データの検索語との相関量を計算し、最後に相関量の多い順にソートした結果を生成していた。このときの検索画像データの計算順序は、軸の重みや軸上でのベクトルの値といった優先度は考慮せず、画像データ ID 順に計算される。したがって、従来方式は、相関量の観点から見ると、ランダムに画像を抽出し、その相関量を計算していることに対応する。本論文で提示した高速化アルゴリズムとの比較において、従来方式は、ランダムに画像を抽出し、相関量計算を行った場合の結果に相当する。

4.3 実験結果

4.3.1 実験 1

実験 1 では、3 つの観点から高速化アルゴリズムの有

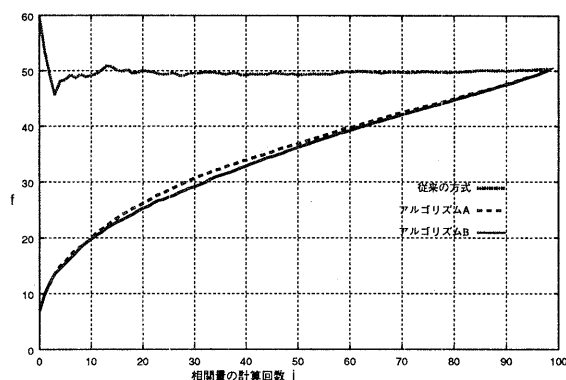


図 6 従来の方式との正規順位平均の比較

Fig. 6 Comparison of correctness between the conventional method and the proposed method

効性を検証している。

● 客観的な正確性の評価

相関量計算を i 番目の画像まで進めたときの正規順位の合計値は、 $\sum R_i$ で表され、1 回の計算における順位の期待値は $\frac{\sum R_i}{i}$ となる。100 番目、つまりすべての画像データについて相関量計算を終了した時に理論上期待値は 50 となるが、相関量計算の終了した画像データの少ない時点での期待値が小さいときほど、アルゴリズムの正確性が高いと言える。181 種類の印象語による問い合わせを行った結果、この正規順位の合計の期待値を 181 回の問い合わせについて平均したものを $f_i = \frac{\sum_{k=1}^{181} \sum_{181 \times i} R_{ki}}{181 \times i}$ (R_{ki} は、 k 回目の問い合わせのときの R_i) として、図 6 に示す。

従来方式では、計算の進行に伴い出力される正規順位の平均値はほぼ一様となっている。一方、アルゴリズム A とアルゴリズム B では、計算回数の少ない早い段階で、正規順位が上位のものが出力されていることがわかる。図 6 より、アルゴリズム A とアルゴリズム B の正確性について、大きな差は見られないものの、アルゴリズム B の方が正確性が高いことがわかる。

● 主観的な正確性の評価

ここでは、あらかじめ人手によって設定した正解画像に対し、各問い合わせにおける上位 5 画像に関して、従来方式と客観的な正確性の評価で最も精度の高かったアルゴリズム B の適合率と再現率を比較した。従来方式では、全 100 画像に対して相関量計算を行った結果を出力しているが、高速化アルゴリズムでは 20 画像のみを計算している。正解画像は、一つの問い合わせ(印象)につき 5 つ設定し、残りはどちらもいえない、当てはまらないに分類した。どちらもいえない画像は、正解画像の半分に対応するものとし、適合率と再現率を以下のよ

表2 主観的評価にもとづく適合率と再現率

Table 2 Examples of relevance ratio and reproduction ratio

検索 印象語	従来方式		アルゴリズム B	
	適合率	再現率	適合率	再現率
dynamic	1	1	1	1
funny	0.7	0.6	0.6	0.6
gentle	0.9	0.8	0.9	0.8
sad	1	1	1	1
silent	0.8	0.8	0.8	0.8
warm	0.9	0.8	0.9	0.8

表3 アルゴリズム間の相関量計算時間の比較

Table 3 Comparison of retrieval time among algorithms

アルゴリズム \ 時間 (msec)	100 画像	1000 画像
従来方式	16060	108334
アルゴリズム A	1992	4269
アルゴリズム B	1957	3617

うに定義した。(検索数 = 5, 設定した正解画像 = 5)
 適合率 = $\frac{\text{検索された (正解 + どちらともいえない) 画像数}}{\text{検索数}}$

再現率 = $\frac{\text{検索された正解画像数}}{\text{設定した正解画像}}$
 結果の一例を表2に示す。

高速化アルゴリズムでは、20 画像のみを計算した結果であるにもかかわらず、従来方式とほとんど同等の適合率および再現率で、検索者の意図する画像が検索できていることがわかる。

● 計算時間による評価

表3は、181 回の問い合わせに対して、正規順位1位のものが見つかるまでの平均時間(マイクロ秒)を示している。従来方式、アルゴリズム A に比べて、アルゴリズム B は最も少ない計算量で相関量の高いデータを検索できるため、実際の検索時間も最も短いことがわかる。

4.3.2 実験 2

アルゴリズム B において、極小データが現れる例として、印象語 brilliant による問い合わせを行ったときの結果を、表4、表5、および、図7、図8に示す。図7では、

$$h_i = \frac{i \text{ 番目までの正規順位合計の理想値}}{i \text{ 番目までの実際的な正規順位合計}} = \frac{\sum_{j=1}^i j}{\sum_{j=1}^i R_j}$$
 を、指標としている。ここでいう理想値とは、i 番目に計算された画像データの正規順位が i であることを指している。すなわち、この理想値を基準とした h_i が1に近付くほど、アルゴリズムの正確性は高いことが示される。図8の○の中の数字は正規順位、右肩の数字は計算順序を表している。拡張方式における1ブロックあたりの計算データ数は、 $4 = 2 \times 2$ とした。

基本方式では、相関量計算の4回目で正規順位が41位の画像データが登場し(図8)、極小データとなってい

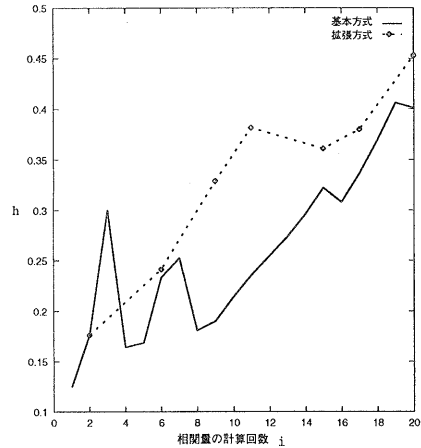


図7 基本方式と拡張方式の比較

Fig. 7 Comparison between the basic fast algorithm and the advanced algorithm

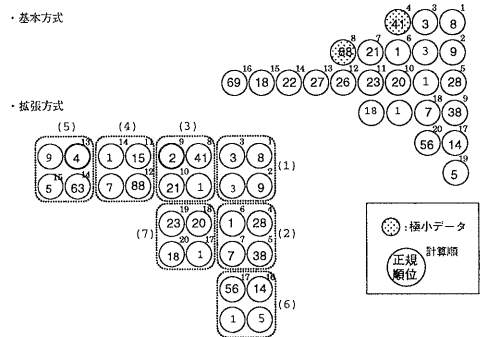


図8 計算順序の様子

Fig. 8 Execution process

る。この極小データは、相関量の大きい画像データが含まれる確率が高いと考えられる最も重みの大きい軸に存在しており、次に位置している画像データが正規順位が2位であるにもかかわらず、基本方式におけるここでの相関量の計算では、先に進むことができていない。一方、拡張方式では、極小データを吸収して相関量の計算が行えるため、出力結果(表4)や指標 h のグラフ(図7)を見ても基本方式よりも正確性が高い。

また、表5に示すように、極小データ画像として現れた41位の raeburn や 88位の salviati は、直観的にも、検索結果上位2位のデータに比べて、印象語の brilliant との関連が強くないことがわかる。

4.3.3 実験 3

図9に、181 回の問い合わせを行った結果の各アルゴリズムにおける指標 h の平均 $ave(h_i) = \frac{\sum_{k=1}^{181} h_k}{181 \times i}$ を表す。各方式の間にあまり差は見られないが、アルゴリズム B の基本方式、アルゴリズム A、アルゴリズム B の

表4 アルゴリズムBの基本方式と拡張方式の出力結果と計算順序

Table 4 Output results and execution path of basic fast process and advanced process

結果順位	基本方式		拡張方式	
	正規順位	計算順	正規順位	計算順
1	1	6	1	6
2	3	3	2	9
3	5	19	3	3
4	7	18	4	13
5	8	1	5	15

表5 上位出力画像と極小データ画像の例

Table 5 an example of retrieval results images in high order and local minimum images

正規順位	画像ID	相関量	画像を定義した印象
1	gogh	0.427176	dear brilliant huge
2	wadsworth	0.312555	shine summer sweet2
41	raeburn	0.211855	nimble freedom crisp
88	salviati	0.181144	fertile2 gravel luxury

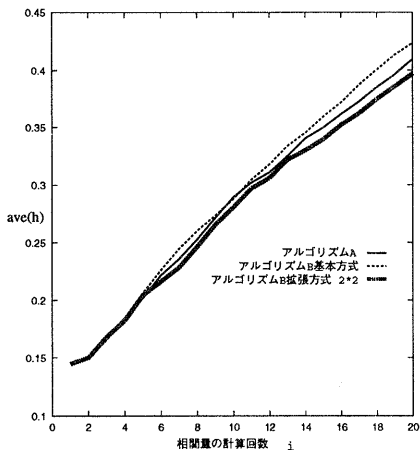


図9 アルゴリズム間の相関量計算回数の比較

Fig. 9 Comparison of number of computation data among algorithms

拡張方式の順に正確性の高いことがわかる。拡張方式よりも基本方式の方がきめ細かい判断によりアルゴリズムが進行するため、結果の正確性がやや高い。

しかし、1ブロックに計算するデータの数を増やしても、正確性の低下は少ないことが図10で確認できる。これは、拡張方式を利用し、並列処理が可能な計算機の各プロセッサに相関量計算を割り当て、並列計算を行うことによって、計算結果の正確性を保ちながら処理速度の向上が期待できることを示している。

4.4 考察

以上の実験結果から、実験に用いた3種類のアルゴリズム(アルゴリズムA, アルゴリズムBの基本方式・拡張方式)を総合的に比較すると、アルゴリズムBは、各実験結果より最も効果的なアルゴリズムであると考えられる。また、アルゴリズムBにおいて発生する可能性が

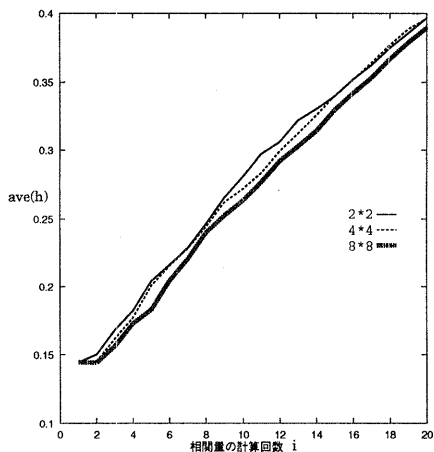


図10 拡張方式における1ブロックあたりの計算データ数の比較

Fig. 10 Comparison of number of computation data for the number of data per block in the advanced algorithm

ある極小データの出現は、拡張方式によって吸収することができ、拡張方式は、相関量計算の並列化による高速な処理へも適用可能である。

5. おわりに

本論文では、意味の数学モデルによる意味的画像検索方式を対象とした高速化アルゴリズムを提案した。これによって、時間制約を伴った大量のメディア情報からの検索において実時間処理を実現した。また、提案アルゴリズムを適用した実験システムを作成し、提案アルゴリズムの有効性を確認した。

本実験では、客観的、主観的尺度により提案アルゴリズムの有効性を示した。本アルゴリズムは、より大規模な画像データベースを対象とした場合、データ数に依存しない検索を可能とするアルゴリズムであるので、さらに大きな効果を期待できる。より大規模な画像データベースを対象とした場合の本アルゴリズムの実験の実現は、今後の検討課題とする。また、相関量計算の並列化によるさらなる高速化、異種メディアへ間の統一的な検索環境の実現などを旨とする予定である。

参考文献

- 1) A.Guttman: R-tree - A Dynamic Index Structure for Spacial Searching, in proceedings of ACM SIGMOD, pp.47-57 (1984).
- 2) Kiyoki, Y., Kitagawa, T. and Hayama, T.: A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning, ACM SIGMOD Record, Vol.23, No.4, pp.34-41 (1994).

- 3) 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(book), A. Sheth and W. Klas(editors), Chapter 7 (1998). (平成11年 6月20日受付)
(平成11年10月22日採録)
- 4) Kiyoki, Y., Kitagawa, T. and Hitomi, Y.: *A Fundamental Framework for Realizing Semantic Interoperability in a Multidatabase Environment*, Journal of Integrated Computer-Aided Engineering, Vol.2, No.1, pp.3-20, John Wiley & Sons, Jan (1995).
- 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) 栗田 多喜夫, 加藤 俊一, 福田 郁美, 坂倉 あゆみ: 印象語による画像データベースの検索, 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383 (1992).
- 7) 芥子 育雄, 池内 洋, 黒武者 健一: 百科辞典の知識に基づく画像の連想検索, 信会誌 (D-II), Vol.J97-D-II, No.4, pp.484-491 (1996).
- 8) 中村 恭子, 金子 昌史, 清木 康, 北川 高嗣: 意味の数学モデルによる意味的画像探索方式とその学習機構, 情処研報, DBS-104, pp.121-128 (1995).
- 9) 宮原 隆行, 清木 康, 北川 高嗣: 意味の数学モデルによる意味的連想検索の高速化アルゴリズムとその実現方式, 情報学論, Vol.38, No.7, pp.1399-1411 (1997).
- 10) 宮原 隆行, 清木 康, 北川 高嗣: 意味の数学モデルによる意味的画像探索のための高速化アルゴリズム, 情処研報, DBS-113, pp.245-250 (1997).
- 11) *Longman Dictionary of Contemporary English*, Longman, (1987).
- 12) Ogden, C.K.: *The General Basic English Dictionary*, Evans Brothers Limited (1940).
- 13) Persin, M.: Document filtering for fask ranking, Proc. of 17th ACM SIGIR Conf., pp.339-349 (1994).
- 14) W.Y.P.Wong.: Implementations of partial document ranking using inverted files, Information Processing and Management, 29(5), pp.647-669 (1993).

(担当編集委員 加藤 俊一)



宮川 明子 (正会員)

1973年生。1996年お茶の水女子大学理学部情報科学科卒業。1998年慶應義塾大学大学院修士課程政策・メディア研究科修了。現在、三菱電機 情報技術総合研究所勤務。データベースシステムの研究に興味をもつ。



清木 康 (正会員)

1978年 慶應義塾大学工学部電気工学科卒業。1983年 同大学院工学研究科博士課程了。工学博士。同年、日本電信電話公社武蔵野電気通信研究所入所。1984年～1995年筑波大学電子・情報工学系講師、助教授を経て、1996年慶應義塾大学環境情報学部助教授、1998年同学部教授。データベースシステム、知識ベースシステム、マルチメディアシステムの研究に従事。ACM, IEEE, 電子情報通信学会, 日本ソフトウェア科学会各会員。



宮原 隆行 (正会員)

1994年 筑波大学第三学群情報学類卒業。同大学大学院博士課程工学研究科電子・情報情報工学専攻を経て、現在、上武大学商学部講師。データベースシステムの研究に従事。



北川 高嗣 (正会員)

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。工学博士。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を経て1990年より筑波大学電子・情報工学系に勤務。現在同学系助教授。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。