

ウェーブレット変換を用いた対話的類似画像検索と 民俗資料データベースへの適用

小早川 倫広[†] 星 守[†]
大森 匡[†] 照井 武彦^{††}

本稿は、ウェーブレット変換と高次局所自己相関特徴量を用いた画像データ索引自動生成法と対話的類似画像検索システムを提案する。このシステムの3つの枠組みとして、1) ウェーブレット変換による多重解像度解析を用いる。2) 特徴量抽出に高次局所自己相関関数を用いる。3) 類似画像検索に近傍探索を用いる。上記の枠組みで、スケッチ画像検索、輪郭画像検索、例示画像検索や例示画像にスケッチを描き加えての検索などを繰り返して画像を検索できる対話的類似画像検索システムのプロトタイプシステムを構築し、評価実験を行った。このプロトタイプの画像データとして、国立歴史民俗博物館の民俗学資料カードの写真3315枚を用いた。このプロトタイプでは、検索者がスケッチ画、写真や写真に手を加えた画像を質問画像として、類似画像検索ができる。また、検索者は、検索結果の候補画像を見ながら、質問画像を作り直すことを繰り返す、対話的な検索ができる。評価実験から、対話的に類似画像検索を繰り返すことによってキーとなる最初の質問画像が曖昧な場合にも目的の画像を検索することができることが示された。

Interactive Image Retrieval Based on Wavelet Transform and Its Application to Japanese Historical Image Data

MICHIHIRO KOBAYAKAWA,[†] MAMORU HOSHI,[†] TADASHI OHMORI[†]
and TAKEHIKO TERUI^{††}

In this paper we propose a basic framework on which similarity-based retrieval system based on image contents can be uniformly realized for a variety of contents such as a sketch, a boundary, an example image and these combination. The framework consists of three parts: 1) Hierarchical decomposition of images using orthogonal discrete wavelet transform (DWT), 2) Feature extraction by 2nd-order local autocorrelation, 3) Similar image retrieval from the database by nearest neighbor search. Based on the framework above, we have developed a prototype system of an image database on a personal computer with a scanner. This system includes 3315 images of classical pictures, patterns, and instruments, which are Japanese historical heritages and/or legacy objects collected in the National Museum of Japanese History. We use a WWW browser and the standard graphical tools as user interface. The system allows similarity retrieval based on example images, user-constructed sketches and drawings, and interactive combinations of these. Furthermore, the system supports interactive retrieval by browsing which is needed when users are vague about their retrieval needs or unfamiliar with the information available in the database. Experiments showed a good performance of interactive similarity retrieval based on contents.

1. はじめに

博物館関係者の間では、長年にわたって蓄積されている画像・音楽などの収集データや、新しい収集データを計算機を使用して有効利用を図りたいという要求

が10年以上も前からあった。

さらに、近年急速にパーソナルコンピュータや周辺機器の高性能・低価格化やネットワーク化などが進み、個人ユースの計算機環境でも、大量のデータが取り扱えるようになってきたパーソナルコンピュータを用いて、収集した画像データを、素早く検索したり、閲覧したりすることのできるデータベースシステムの要望が高まった。しかし、現在実用になっている画像データベースシステムは、キーワード検索によるテキスト型のデータベースで、キーワード付加のコストや時間の

[†] 電気通信大学大学院情報システム学研究科
Graduate School of Information Systems, University of
Electro-Communications

^{††} 国立歴史民俗博物館
National Museum of Japanese History

問題、データベースを構築してもキーワード入力に専門分野の知識を必要とし、検索キーワードを知らなければ検索できないなどの問題あった。たとえば、本研究で利用した資料は、歴史民俗資料カードと呼ばれ、1972年から1974年の3年間をかけて収集・作成された、全国の歴史・考古・民俗の3分野の「もの」資料について作成されたカードである。これらカードには、保管場所、標準名、地方名や使用目的を記述したテキストデータと写真・スケッチの画像データが含まれている。この歴史民俗資料カードのデータベース化は、当初、「もの」の名前、ファイルIDや分類コードなどをキーワードとするテキスト型のデータベースシステムとして構築されたので、キーワードをまったく知らない人や初めてそのデータベースを利用する人には、適切なキーワードを指定することが難しく、データベースを検索することが難しい。

そこで、検索者が大まかな形を手描きした画像や写真や本などから抜き出した画像などを検索キーとして検索できる画像内容に基づいた類似画像検索システムが期待されている。同様に、博物館などの画像データだけではなく、一般の画像データの検索に対しても、画像内容に基づいた類似画像検索実現への期待は高い。

このように、画像内容に基づいた類似画像検索を可能にする検索方式の確立は重要な課題である。

画像内容に基づく類似画像検索システムは、

- スケッチ
- 輪郭
- テクスチャ
- 例示画像
- 画像の一部
- 色

などの画像内容から簡単に類似画像を検索する機能が必要となる。ここで我々は、画像をウェーブレット変換し、ウェーブレット係数空間上でそれぞれの画像内容をとらえ、さまざまな画像内容に基づく検索を統一的に扱う類似検索方式を提案する。本稿では、スケッチ画像検索・輪郭画像検索と例示画像検索を同一の枠組みの上で実現した類似画像検索システムのプロトタイプシステムを示す。ただし、本稿では検索機能の概要とその使用例を中心に述べた。ウェーブレット変換、高次局所自己相関特徴¹⁴⁾などの詳細は、文献3)を参照していただきたい。

2. 関連研究と提案システム

画像内容に基づく類似画像検索システムとしては、IBMのQBIC^{2),16)}、電総研の感性語による絵画検索

システム^{5),7)}、NTTのExsight¹⁾などがあげられるが、さまざまな画像内容による検索に対応しているのはQBICのみである。このQBICのシステムアーキテクチャは、個々の画像内容に対して、個別にそれぞれのアルゴリズムを用意してシステムを構成している。たとえば、Stollnitz¹³⁾の絵画検索手法などを取り込んで、色情報からの類似画像検索を行えるようにしたり、画像の輪郭抽出は、エッジオペレータを用いて行ったりと、いろいろな画像処理技術を駆使してシステムを構築している。個々の画像内容ごとに別々の特徴量抽出方法や検索方式を用いることは、検索システムの複雑化、巨大化を招くことになる。したがって、各検索機能ごとに異なった個別の検索方式を提供するだけでは、シンプルで効率的な検索システムを構築できない。

我々のシステムアーキテクチャの基本は、スケッチ画像検索、輪郭画像検索、例示画像検索、テクスチャ検索、部分画像検索、色検索、テクスチャ分割、テクスチャ合成、画像圧縮、ブラウジング用の画像生成を直交ウェーブレット係数空間を用いて、できるだけ統一的に行う。たとえば、本稿で提案しているスケッチ画像検索・輪郭画像検索や例示画像検索に対して、ウェーブレット変換による多重解像度空間⁹⁾の各解像度から特徴量を抽出し、索引空間を生成している¹⁰⁾。同様に、テクスチャ検索の索引に用いる特徴量もまた、ウェーブレット変換の多重解像度空間を用いて行っている¹¹⁾。また、テクスチャ領域分割などのテクスチャ解析もウェーブレット変換の多重解像度空間を用いている^{4),6)}。すなわち、画像内容に基づく類似画像検索を行うために個々の画像内容に対応した別々のアルゴリズムを組み合わせるのではなく、できるだけ統一的に検索機能を実現するために画像のウェーブレット係数空間を基本空間と考え、ウェーブレット変換による多重解像度空間から画像内容に対応する特徴量抽出を行い、画像の自動索引生成と対話的類似検索を行う。このような基本方針をとることにより類似画像検索システムがシンプルにまとまり、軽いシステムを構築することができる。

3. 類似画像検索システム

画像内容に基づく類似画像検索システムは、スケッチ、輪郭、テクスチャなどのさまざまな画像内容に対応するため、以下の検索機能が必要となる。

(1) スケッチ画像検索：ラフスケッチ画像を入力して類似画像を検索する。

例) 検索者が、図1(a)のスケッチ画像を質問

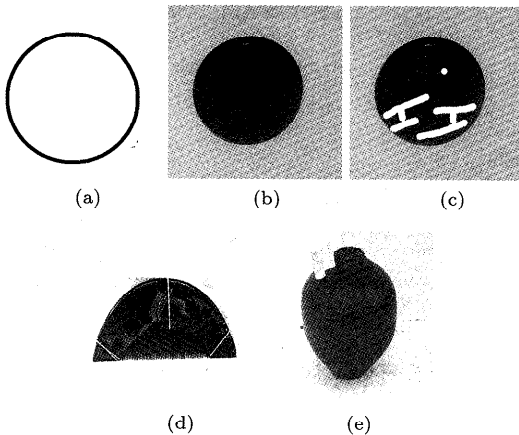


図1 質問画像として用いた画像

Fig. 1 Query by sketch (a), boundary image (b) (d), example image (e), and these combination (c).

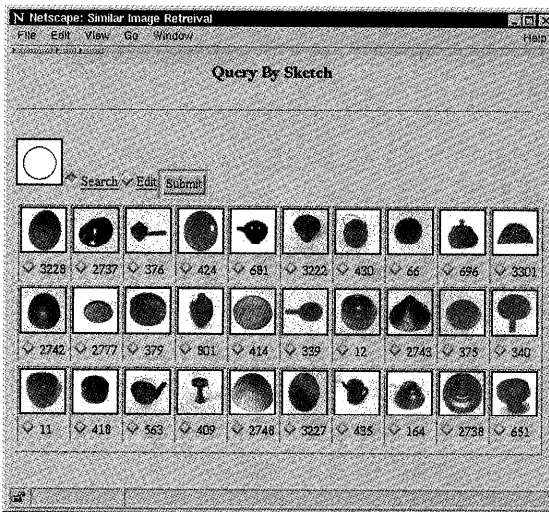


図2 スケッチ画からの検索

Fig. 2 Query by sketch.

画像として検索を行うと、システムは図2の候補画像群を表示する。

- (2) 輪郭画検索：画像を入力するとシステムがエッジ抽出を行い、質問画像の輪郭画像を自動作成し、この輪郭画像を用いて画像を検索する。
例) 検索者が、図1(b)の画像を質問画像とし検索を行うと、システムは図3の候補画像群を表示する。
- (3) 例示画像検索：写真などの画像を入力してそれと類似する画像を検索する。
例) 検索者が、図1(e)の画像を質問画像とし検索を行うと、システムは図6の候補画像群を表示する。

- (4) テクスチャ検索：入力画像と類似のテクスチャ画像を検索する。
- (5) 対話的検索：上記の検索の結果表示される候補画像群の中の1枚や候補画像群の中の1枚に手を加えながら繰り返し検索する。
例) 検索者が、図1(b)の画像に手を加えた図1(c)を質問画像とし、輪郭画像検索を行うと、システムは図4の候補画像群を表示する。
例) 検索者が、図2の候補画像から図1(d)を質問画像として選び輪郭画像検索を行うと、システムは図5の候補画像群を表示する。
例) 検索者が、図2の候補画像から図1(e)を選び、それを質問画像とし例示画像検索を行うと、システムは図6の候補画像群を表示する。
- (6) ブラウジング：画像をばらばらと見たり、データベースをざっと見る。

画像内容に基づく画像検索では、画像を解析し特徴量を計算する機能が必要となる。画像を解析する場合、画像の局所的な詳細な情報と画像の大域的な大まかな情報を抽出することが重要になる。このための一般的な手法としては、画像の多重解像度解析が用いられ、画像のピラミッド構造がよく利用される。多重解像度解析では、解像度が粗い場合、画像の大域的な構造を抽出することができ、解像度が細かい場合、局所的な詳細な情報を抽出することができる。また、画像をブラウジングする場合でも、いろいろな解像度の画像間の変換は簡便で可逆であることが必要である。

そこで我々は、直交ウェーブレット変換による多重解像度解析と高次局所自己相関関数¹⁴⁾を用いた画像特徴抽出方法、画像特徴量に基づいて索引と検索キーを自動的に生成する方法、検索結果を見ながら次の検索画像を入力する対話的類似画像検索法について提案し、これら方法に基づいた類似画像検索システムのプロトタイプについて述べる。

我々の対話的類似画像検索システム(図7)は、画像特徴抽出部、画像索引生成部、検索キー生成部、検索実行部、ユーザインタフェース部からなり、次のような機能を持つ。

画像特徴抽出部：画像特徴を抽出。

- (1) ウェーブレット変換による多重解像度分解。
- (2) 高次局所自己相関関数による特徴抽出。

画像索引生成部：特徴量空間から索引空間を生成。

- (1) 判別分析を用いて索引空間を生成。

検索キー生成部：質問画像から検索キーを生成。

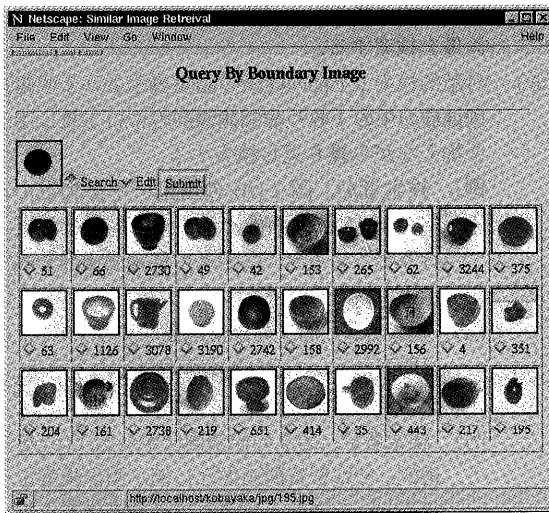


図 3 輪郭画像からの検索 (1)
Fig. 3 Query by boundary image (1).

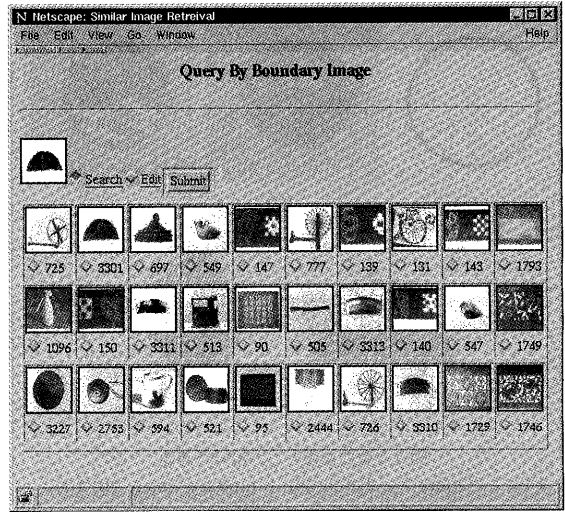


図 5 輪郭画像からの検索 (2)
Fig. 5 Query by boundary image (2).

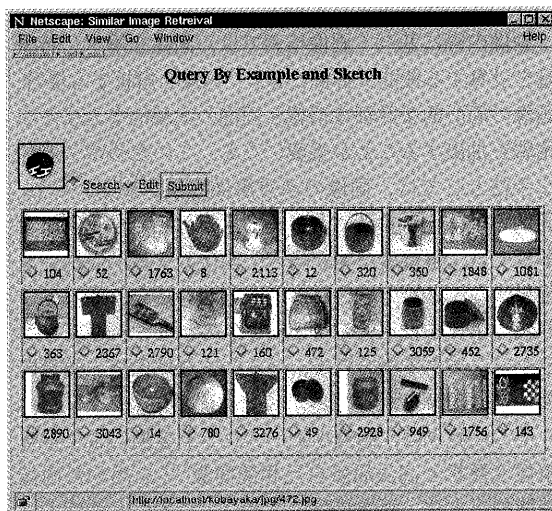


図 4 画像にスケッチを描き加えた画像による検索
Fig. 4 Query by combined an example image with a sketch image.

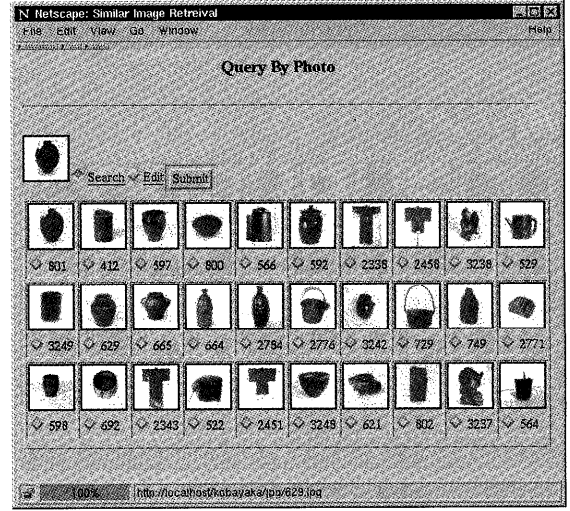


図 6 例示画像からの検索
Fig. 6 Query by example image.

- (1) 質問画像入刀.
- (2) 検索方式の指定.
- (3) 質問画像特徴の抽出 (画像特徴抽出部).

検索実行部: 類似画像検索を実行.

- (1) 質問画像の特徴量を索引空間に写像.
- (2) 近傍探索.

ユーザインタフェース部: WWW ブラウザ上で操作.

- 画像入力画面の表示.
- 画像検索方式の指定.
- 検索結果の表示.

3.1 特徴量抽出

従来、画像特徴抽出には、FFT、DCTなどの直交変換が多く用いられてきた。ここで我々は、変換基底の台が有限である直交ウェーブレット変換に着目し、その多重解像度解析と高次局所自己相関関数を画像特徴量抽出に用いる。ウェーブレット変換を用いた多重解像度解析は、MeyerやMallatにより発展し、多重解像度解析の画像処理への応用として、エッジ抽出⁸⁾、テキスト解析^{3),15)}への有効性が示されている。また、画像データベースへの応用としては、Stollnitzが、ウェーブレット分解空間そのものを特徴量空間として絵画データベースへ適用したものがある。我々は、ウェー

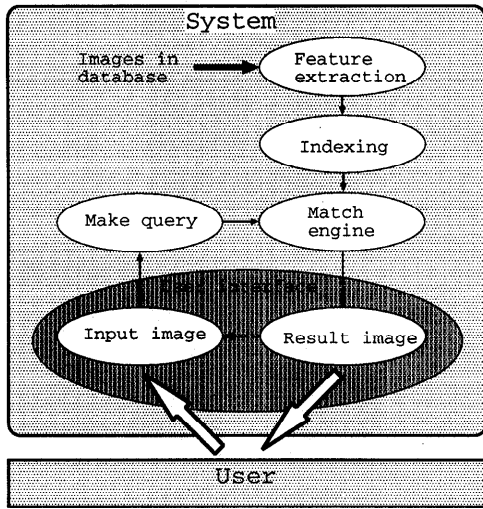


図7 類似画像検索システムの構造

Fig. 7 Similarity-based image retrieval system architecture.

レット変換による多重解像度空間の各解像度の関係に着目し、高次局所自己相関関数を用いて特徴量を抽出する。以下、3.2節、3.3節で、ウェーブレット変換と高次局所自己相関関数についての簡単に述べる。詳細な説明は、文献9)、12)を参照していただきたい。

3.2 ウェーブレット変換

2次元画像の場合を考える。スケーリング関数 $\phi(x, y)$ とマザーウェーブレット $\psi(x, y)$ が張る分解空間 S_j を、

$$S_j = (V_j, W_j^1, W_j^2, W_j^3). \tag{1}$$

と表す。ここで、サイズ $N \times N$ の原画像 A_0 ($A_0 \in V$) に対するウェーブレット変換を

$$S(A_0) = \left(\begin{array}{c|c} A_1 & D_1^1 \\ \hline D_1^2 & D_1^3 \end{array} \right), \tag{2}$$

$$P_{V_1}(A_0) = (a_{1,mn}) = A_1, \tag{3}$$

$m, n = 1, \dots, N/2$

$$P_{W_1^k}(A_0) = (d_{1,mn}^k) = D_1^k, \tag{4}$$

$1 \leq k \leq 3 \quad m, n = 1, \dots, N/2$

と表す。ただし P_{V_1} 、 $P_{W_1^k}$ は、空間 V_1 、 W_1^k への射影を表す。ここで、原画像 A_0 は、 A_0 の1/4のレベル1の縮小画像 A_1 とレベル1の差分画像 $\{D_1^k\}_{(1 \leq k \leq 3)}$ に分解される。再び A_1 に対してウェーブレット変換を行い、分解空間 S_2 を得る。このように A_j を再帰的にウェーブレット変換をすると、原画像 A_0 は、

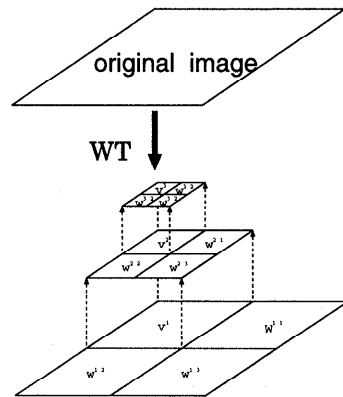


図8 ウェーブレット変換を用いた多重解像度解析

Fig. 8 Multiresolution analysis based on wavelet transform.

$$S_j(A_0) = \left(\begin{array}{c|c|c|c} A_j & D_j^1 & & \\ \hline D_j^2 & D_j^3 & \dots & D_2^1 \\ \hline \vdots & \ddots & & D_1^1 \\ \hline & D_2^2 & D_2^3 & \\ \hline & D_1^2 & & D_1^3 \end{array} \right), \tag{5}$$

のように分解される (図8)。

したがって、ウェーブレット変換を用いた多重解像度解析により、縮小画像 $\{A_j\}_{(1 \leq j \leq J)}$ のピラミッドと差分画像の $\{D_j^k\}_{(1 \leq k \leq 3)}_{(1 \leq j \leq J)}$ のピラミッドを一度に構成することができる。

本稿では、Haar 基底によるウェーブレット変換を用いているので、分解空間 D_j^1, D_j^2 は縮小画像の1階微分型のエッジフィルタをかけた空間となっている。

輪郭画像 B_j は、差分画像 $\{D_j^k\}_{(k=1,2)}$ を用いて作る。ここで、エッジ抽出オペレータ \mathcal{E} を、

$$\mathcal{E}\{D_{j,mn}^k (k=1,2)\} = B_j = (b_{j,mn}), \tag{6}$$

と定義する。ただし、 $b_{j,mn}$ (エッジ強度) は、

$$b_{j,mn} = \sqrt{(d_{j,mn}^1)^2 + (d_{j,mn}^2)^2}, \quad b_{j,mn} \in B_j. \tag{7}$$

この操作を行うことにより、原画像からの縮小画像ピラミッドと輪郭画像ピラミッドを、ウェーブレット変換から構成することができる。

3.3 高次局所自己相関特徴量

本研究では、ウェーブレット変換によって得られた分解画像群からの特徴量の抽出に高次局所自己相関関数を用いた。高次自己相関関数は次数や変位方向のとり方により無数に考えられる。ここでは簡単のため次

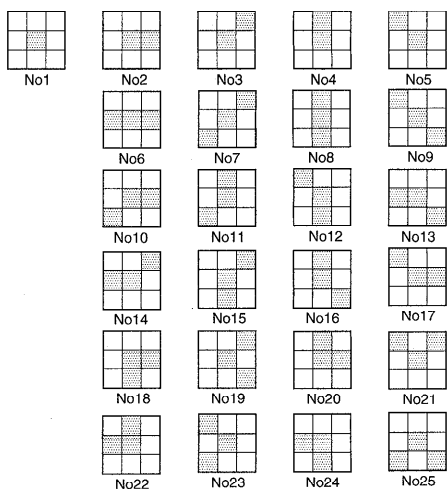


図9 高次局所自己相関特徴量の局所マスクパターン
Fig. 9 Local mask pattern of 2nd-order local auto correlation.

数をたかだか2までとし、また、画面に対する加法性の条件を満たすために、変位方向を参照点の周りの局所的な 3×3 画素の領域に限定した。平行移動に関して普遍な特徴を除くと、特徴の数は全部で25個になる。各特徴の計算は、局所パターンの対応する画素の値の積を全画素に対して足しあわせればよい。用いた局所パターンを図9に示す。No.1からNo.25までの各パターンに対応する特徴を f_1, \dots, f_{25} とすると、次のような25次元の特徴ベクトルを作ることができる。ここで、画像 I に対する高次局所自己相関特徴抽出 \mathcal{H} を

$$\mathcal{H}(I) = \mathbf{f} = (f_1, f_2, f_3, \dots, f_{24}, f_{25}), \quad (8)$$

と表す。このように計算された特徴は、明らかに対象の位置に関して普遍でしかも画面に関する加法性を満たしている。

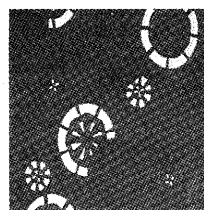
3.4 画像の特徴量の計算

実際に画像データベースに登録する画像に多重解像度解析を行い、画像特徴量を計算する。用いた画像のサイズは、 384×384 である。したがって、最大7レベルの解像度に分解できるが、特徴量算出のため6レベルまでを使用した。画像ピラミッドの例を図10に示す。特徴量算出の手順は次のようになる。

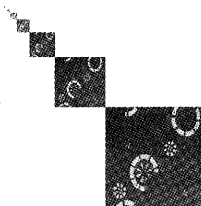
step1: ウェーブレット変換。

$$S_j(A_{j-1}) = \{A_j, \{D_j^k\}_{1 \leq k \leq 3}\}, \quad (j = 1, \dots, 6) \quad (9)$$

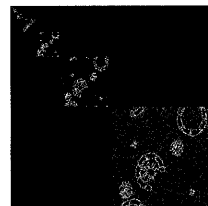
によって、画像はウェーブレット変換され、縮小画像 A_j ($j = 1, \dots, 6$)と差分画像 $\{D_j^k\}$ ($1 \leq k \leq 3$, $j = 1, \dots, 6$)に分解される。



(a)



(b)



(c)

図10 画像ピラミッドの構成：(a) 民俗学型紙画像、(b) 縮小画像ピラミッド、(c) 輪郭画像ピラミッド

Fig. 10 Pyramidal image: (a) original image, (b) approximated image pyramid, (c) boundary image pyramid.

step2: エッジ抽出オペレータ式(6)により輪郭画像 B_j を作成する。

step3 画像特徴量の計算。

縮小画像と輪郭画像 A_j, B_j ($j = 1, \dots, 6$)の高次局所自己相関特徴量抽出

$$\mathcal{H}(A_j) = \{f_{j,1}^a, \dots, f_{j,25}^a\} = \mathbf{f}_j^a \quad (10)$$

$$\mathcal{H}(B_j) = \{f_{j,1}^b, \dots, f_{j,25}^b\} = \mathbf{f}_j^b \quad (11)$$

を行う。ここで、 $\mathbf{f}_j^a, \mathbf{f}_j^b$ をレベル j の縮小画像特徴ベクトル、輪郭画像特徴ベクトルと呼ぶ。

縮小画像特徴ベクトルは、例示画像検索用の索引空間の生成に用いられ、輪郭画像特徴ベクトルは、スケッチ画像検索や輪郭画像検索用の索引生成に用いられる。また、縮小画像特徴ベクトルと輪郭画像特徴ベクトルは、ともに j が小さいほどより局所的で詳細な情報を表現し、 j が大きいほどより大域的で大まかな情報を表現していることに注意していただきたい。

3.5 画像索引生成部

画像データベースに登録する画像を N 枚とし、索引生成を行う。ここで、効率的な索引を生成するために判別分析を用いて、次のように索引空間を生成する¹²⁾。

n 番目の画像の特徴量ベクトル $\mathbf{f}_{n,1}, \dots, \mathbf{f}_{n,6}$ はクラス n に属するものとして、これらの特徴量ベクトルの集まりに対してクラス数を N として判別分析を行い、判別空間を作成する。この判別空間を画像データベースの索引空間とし、 n 番目の画像の代表ベクトル $\hat{\mathbf{i}}_n$ を索引とする。ただし、 $\hat{\mathbf{i}}_n$ は、判別空間に写像さ

れたクラス n の特徴量ベクトル群の重心である。

3.6 検索キー生成部

質問画像の画像特徴量を、検索者が指定した検索機能に応じてシステムに送る。

3.7 検索実行部

質問画像の特徴量を索引空間に写像する。写像された質問画像の代表ベクトルを \hat{i}_q とする。次に、索引空間中の索引 \hat{i}_j と質問画像 \hat{i}_q とのユークリッド距離

$$d_j = \|\hat{i}_j - \hat{i}_q\|, \quad (12)$$

を計算し、近いものから順番に一定数（利用者が指定する）を類似画像の候補として、ユーザインタフェース部へ送る。

3.8 ユーザインタフェース部

ユーザインタフェース部は、

- 質問画像を入力する。
 - 入力画像の確認と検索方式の指定をシステムに送信する。
 - 候補画像を表示する。
 - 候補画像を質問画像としシステムに入力する。
- 機能を持ち、すべて WWW ブラウザ上で実行される。

4. 対話的類似画像検索システムのプロトタイプ

質問画像を用いて検索する場合、1 回の検索で目的の画像を見つけ出すことは困難である。本稿では、検索者が質問画像を変えながら類似画像を検索する対話的類似画像検索システムを提案する（図 11）。

類似画像検索システムのプロトタイプシステムは、パーソナルコンピュータ上に構築した。プロトタイプシステムに登録した画像は、国立歴史民俗博物館保管の「もの」資料カードのモノクロ写真画像データ（サイズ 384×384 ）の 3315 枚である。また、手描き画像入力のためのスキャナと画像編集用ソフト GIMP（GNU Image Manipulation Program）を装備している。

4.1 検索機能

プロトタイプシステムは、画像データベース中の画像に対して、縮小画像特徴ベクトルを基に作成された索引空間（縮小画像索引空間）と輪郭画像特徴ベクトルを基に作成された索引空間（輪郭画像索引空間）の 2 つの索引を持っている。これら 2 つの索引を用いて各検索機能を実現する。

例示画検索： 写真の一部などを検索に用いる。

- 縮小画像索引空間において質問画像の縮小画像特徴量ベクトルとの距離を計算し、距離の近

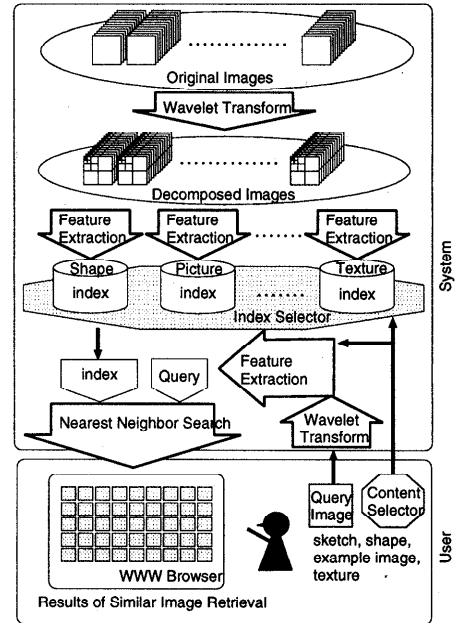


図 11 類似画像検索システムの概要

Fig. 11 Overview of the similarity-based image retrieval system.

いものから候補画像とする。

輪郭画検索： 写真などの一部からシステムが抽出した輪郭を検索に用いる。

- 輪郭画索引空間において質問画像の輪郭画像特徴量ベクトルとの距離を計算し、距離の近いものから候補画像とする。

スケッチ画検索： ユーザの描いたラフスケッチを検索に用いる。

- 輪郭画索引空間において質問画像の輪郭画像特徴量ベクトルとの距離を計算し、距離の近いものから候補画像とする。

対話的検索： 上記の検索の結果を見ながら以下のような検索を繰り返し行う。

- 検索方式を変更して検索を行う。
- 最初の質問画像に手を加え、その画像を質問画像として検索を行う。
- 出力された候補画像群の中から 1 枚を選び、選んだ画像を質問画像として検索を行う。
- 出力された候補画像群の中から 1 枚を選び、選んだ画像に手を加え、その画像を質問画像として検索を行う。

4.2 検索の例

ユーザが曖昧な画像の輪郭を頼りに、「着物」を検索する場合の検索手順について説明する。

(1) まず、ユーザは、「着物」の大まかな輪郭のス

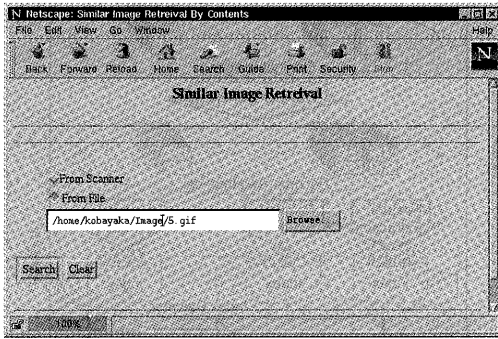


図 12 質問画像入力画面：スキャナ画像入力かファイルからの画像入力の選択

Fig. 12 User interface: input a query image: selector of an image from scanner or file.

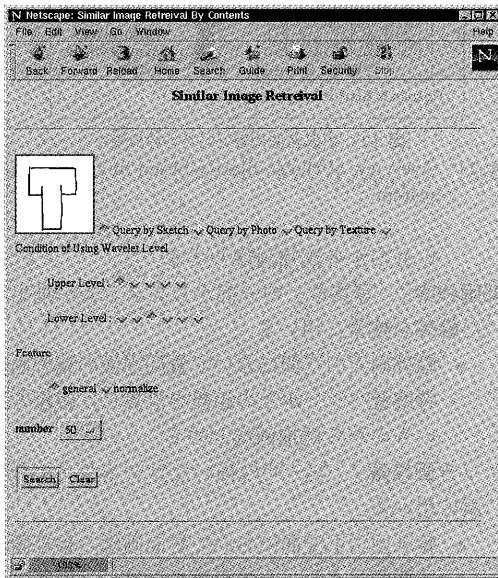


図 13 検索方式の指定画面

Fig. 13 User interface: selection of retrieval type.

ケッチ画をスキャナにセットし、類似画像検索システムに送る (図 12)。

- (2) システムは、入力された質問画像を表示するとともに検索方式について質問する (図 13) のので、ユーザは検索方式を指定する。この例では、スケッチ画検索を選択し、候補画像表示枚数を 50 枚に指定した。
- (3) システムは検索方式にあった方法で質問画像の特徴量を計算し、索引空間で近傍探索を行い、距離が近いものから順番に類似候補画像として表示する (図 14)。
- (4) 目的の画像が合った場合は、利用者がその画像をクリックし画像を拡大して表示する。目的の

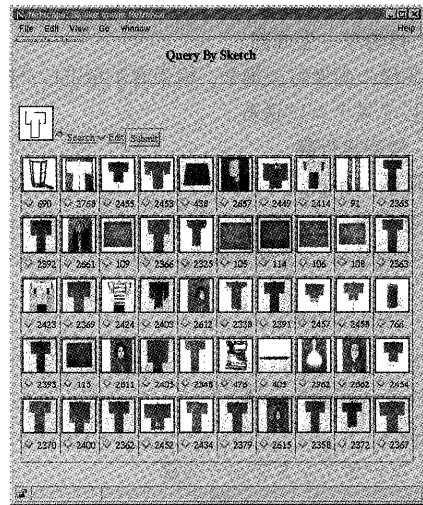


図 14 1 回目の検索結果

Fig. 14 Result of the first query.

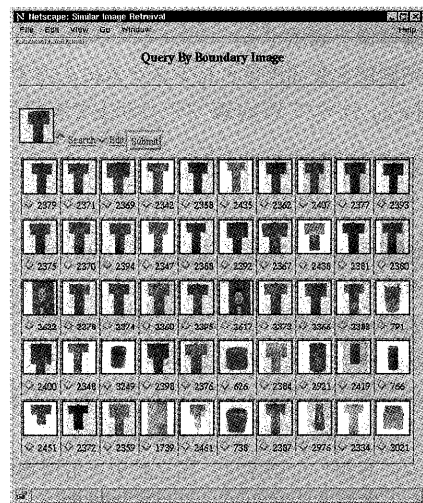


図 15 図 14 の候補画像の 1 枚を質問画像として検索した結果

Fig. 15 Result of the second query using one of the resulting images of the first query.

画像に近い画像が候補画像の中にあつた場合は、その 1 枚を選び、これを次の質問画像として (2) に戻る (図 15)。ユーザは候補画像の中に目的に近い画像がなければ、(1) に戻るか、(2) の検索方法の指定を変更し検索し直す。

5. 類似画像検索システムの評価

類似画像検索の評価は、人間の主観が関係する場合が多く、客観的な評価が難しい。従来キーワード検索で用いられているヒット率も、類似画像検索では、主観の影響を受ける。そのうえ、画像データベースが大

小さくなると実際上計算が困難であり、別の評価基準が必要である。類似画像検索では、目的に近い画像として出力された候補画像群の中に目的に近い画像がどれくらい含まれているかという基準と、目的に近い画像が候補画像のうち何番目に出力されているのかという基準は、重要な評価基準である。このような基準は、文献検索システムでもヒット率に代わる評価基準として用いられるようになってきている¹⁷⁾。

5.1 評価基準

まず、候補画像の総数中に対する類似画像の割合を出現率

$$r_1 = \frac{N_s}{N_m}, \quad (13)$$

とする。ここで、 N_s は候補画像として出力された画像の中でユーザが目的の画像に近いと思った画像の総数（類似画像）、 N_m は候補画像として出力された画像の総数（ユーザが指定する）、出現率 r_1 を、システムの大まかな性能評価として使用する。

もう1つの評価基準として、得点率を次のように定義する。まず、候補順位 n の画像には、次のような基準で得点

$$s_n = N - n \quad (14)$$

を与える。ただし N は、画像データベース中の画像総数。類似していると判断された画像の得点の総和 $\sum s_s$ と候補画像として出力された画像の総得点 $\sum_{i=1}^{N_m} s_i$ との比を、得点率

$$S = \frac{\sum s_s}{\sum_{i=1}^{N_m} s_i}, \quad (15)$$

と定義する。以下では、出現率と得点率を用いて、検索性能を評価する。出現率、得点率ともに高ければ、その利用者にとってシステムが有効であることが分かる。しかし、システムの絶対的な評価ではないことに注意が必要である。

5.2 評価実験

まずはじめに、検索者が「かま」、「着物」、「瓶」、「羽子板」、「はさみ（やっこ）」、「樽」、「櫛」、「ざる」をイメージして描いたスケッチ画像を質問画像（図16 (q1)～(q8)）として、スケッチ画像検索を行った。その検索結果を前記の2つの評価基準に関して、次のように調べた。ここでの評価は、検索者本人のほかに評価者5人で行った。まず検索者が、スケッチを手描きし、スケッチ画像検索を行う。その結果の候補画像の表示枚数を50枚として、次のように評価計算をする。まず、検索者は、どのような「もの」を検索しようとしたかを評価者5人に対して告げる。評価者は、検索結果の50枚の候補画像から、検索者が探している「もの」と

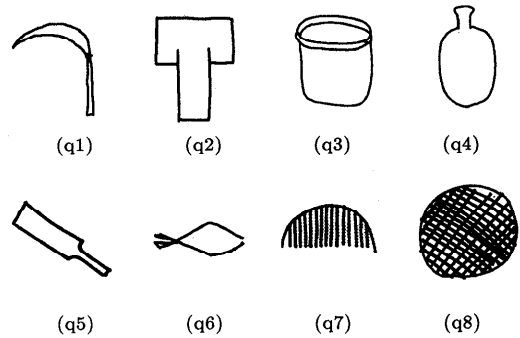


図16 評価実験に用いた手描きスケッチ画像

Fig. 16 Sketches using query.

表1 手描きスケッチを質問画像として質問したときの出現率
Table 1 Rate of appearance: results of query by sketch.

スケッチ画像番号	候補順位				
	10	20	30	40	50
q1	0.80	0.70	0.70	0.69	0.68
q2	0.63	0.52	0.61	0.56	0.63
q3	0.10	0.22	0.26	0.23	0.22
q4	0.57	0.58	0.45	0.43	0.40
q5	0.40	0.45	0.33	0.33	0.26
q6	0.50	0.43	0.31	0.28	0.23
q7	0.30	0.45	0.40	0.35	0.34
q8	0.10	0.05	0.07	0.07	0.07
平均	0.43	0.43	0.39	0.37	0.35

表2 手描きスケッチを質問画像として質問したときの得点率
Table 2 Rate of score: result of query by sketch.

スケッチ画像番号	候補順位				
	10	20	30	40	50
q1	0.80	0.70	0.70	0.69	0.68
q2	0.63	0.52	0.61	0.56	0.62
q3	0.10	0.22	0.27	0.22	0.22
q4	0.57	0.58	0.46	0.43	0.40
q5	0.40	0.45	0.33	0.33	0.26
q6	0.50	0.43	0.31	0.28	0.23
q7	0.30	0.45	0.40	0.35	0.34
q8	0.10	0.05	0.07	0.07	0.07
平均	0.43	0.43	0.39	0.37	0.35

類似していると判定した画像の数とそれらの画像の順位を調べる。たとえば、検索者が図16 (q1)を検索したときは、検索者は「かま」の画像を調べていることを、評価者に告げる。以上のようにして、検索者と評価者5人の計6人の平均出現率、平均得点率を求める。この結果を表1、表2に示す。

また、従来の評価方法であるヒット率（質問画像と同じカテゴリ（たとえば「かま」）として登録されている画像総数のうち検索の結果表示された画像の中で類似していると判断された画像数の割合）も求めた。ただし、どの画像を同じカテゴリに入れるかは、人の主観に大きく左右されたり、研究分野の違いにより分

表 3 画像データベース中の登録枚数

Table 3 The number of images of the same category.

分類名	登録画像数 (枚)
かま	264
羽子板	22
櫛	102
ざる	7

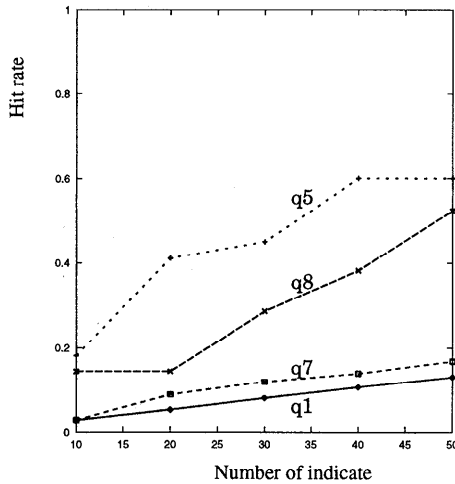


図 17 スケッチ画像検索におけるヒット率の推移

Fig. 17 Hit rate: query by sketch.

類区別が異なることもしばしばあるので、これもまたデータベース利用者の主観に左右される。たとえば、「着物」と「はんでん」は、大きくまとめれば、「着物」のカテゴリと見なすことができるが、「はんでん」を知っていれば違うカテゴリに属すると判別するであろう。ここでは、人による差異の少ない「かま」、「羽子板」、「櫛」、「ざる」についてのみヒット率を求めた。画像データベース中に登録されていた、「かま」、「羽子板」、「櫛」、「ざる」の数を表 3 に示す。また、ヒット率を図 17 に示す。

次に、さきほどのスケッチ画像検索の表示結果から検索者が目的に近いと思う画像 (図 18) を選び、輪郭画像検索を行った。ここで、スケッチ画像検索と同様に、出現率、得点率を計算した。その結果を表 4、表 5 に示す。また、質問画像図 18 (r1), (r5), (r7), (r8) に対するヒット率を図 19 に示す。

5.3 考 察

まず、最初に行った手描きスケッチを質問画像としたスケッチ画像検索の場合は、8 枚の質問に対して平均出現率 0.35、平均得点率 0.35 であった。また、「かま」、「羽子板」、「櫛」、「ざる」のヒット率の平均は 0.35 であった。ここで、「かま」と「ざる」のスケッチ画像検

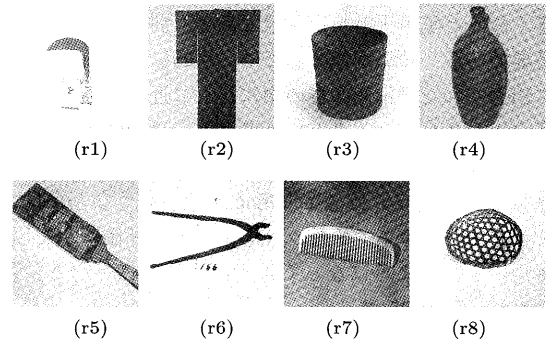


図 18 対話的検索で画像質問として用いた画像

Fig. 18 Query images used interactive retrieval.

表 4 スケッチ検索で表示された画像の 1 枚を質問画像として検索したときの出現率

Table 4 Rate of appearance: result of the second query using one of the resulting images of the first query.

画像番号	候補順位				
	10	20	20	40	50
r1	1.00	0.95	0.98	0.90	0.88
r2	1.00	1.00	0.90	0.83	0.72
r3	0.50	0.52	0.41	0.36	0.34
r4	0.50	0.45	0.33	0.30	0.26
r5	0.50	0.30	0.30	0.25	0.22
r6	0.53	0.33	0.26	0.22	0.17
r7	0.40	0.40	0.40	0.38	0.32
r8	0.10	0.10	0.12	0.14	0.13
平均	0.57	0.51	0.46	0.42	0.38

表 5 スケッチ検索で表示された画像の 1 枚を質問画像として検索したときの得点率

Table 5 Rate of Score: result of the second query using one of the resulting images of the first query.

画像番号	候補順位				
	10	20	30	40	50
r1	1.00	0.95	0.97	0.90	0.88
r2	1.00	1.00	0.90	0.83	0.78
r3	0.50	0.52	0.41	0.36	0.36
r4	0.50	0.45	0.33	0.30	0.26
r5	0.50	0.30	0.30	0.25	0.22
r6	0.53	0.33	0.26	0.22	0.17
r7	0.40	0.40	0.40	0.38	0.32
r8	0.10	0.10	0.12	0.14	0.13
平均	0.57	0.51	0.46	0.42	0.39

索について、出現率、得点率とヒット率について検討する。「かま」のスケッチ画像検索の出現率 0.68、得点率 0.68 は良い成績といえる。ヒット率は 0.13 で悪いように見える。しかし、データベース中には 264 枚の「かま」の画像が含まれており、候補画像の出力枚数を 50 枚としたときの、ヒット率の上限は $50/264=0.19$ であることを考慮すると、「かま」のスケッチ画像検索の結果は良かったといえる。

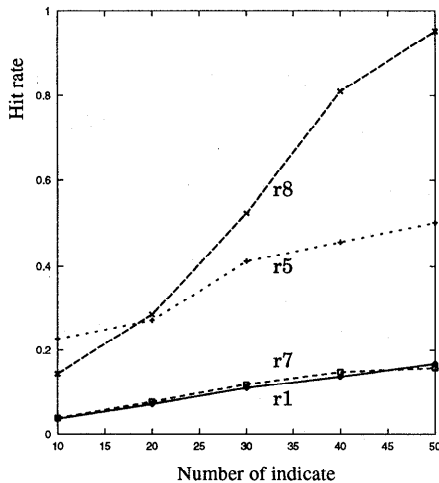


図 19 スケッチ画像検索の候補画像の 1 枚を用いたときのヒット率の推移

Fig. 19 Hit rate: query by sketch.

また、「ざる」のスケッチ画像検索では、出現率 0.07、得点率 0.07 でともに低い値であった。しかし、データベース中に含まれる「ざる」の画像は 7 枚しかないため、出現率の上限は、 $7/50=0.14$ であり、上記の出現率が必ずしも悪い値とはいえない。そして、ヒット率は 0.53 で 7 枚中約 3.7 枚と半数以上を選び出している。

以上の検討から、「かま」、「ざる」ともに該当する画像をよく選び出しているといえる。

次に、スケッチ画像検索の結果得られた候補画像の 1 枚を用いて、輪郭画像検索を行った場合、8 枚の質問に対して平均出現率 0.38、平均得点率 0.39 であった。また、「かま」、「羽子板」、「櫛」、「ざる」のヒット率の平均は 0.44 であった。「かま」に対しては、出現率 0.88、得点率 0.88 でスケッチ画像検索のときに比べて約 30% 高くなった。ヒット率も 0.17 と改善されている。この場合、ヒット率の上限 0.19 に近づいている。「かま」をよく選び出している。また、「ざる」に対しては、出現率 0.13、得点率 0.13 であるが、ヒット率は 0.95 で高いヒット率を実現している。以上のことから、手描きスケッチ画像を最初の質問画像としてスケッチ画像検索し、その検索結果の候補画像を用いて検索を行うことにより、出現率、得点率、ヒット率ともに上昇し、検索者の所望の画像が上位に候補としてあげられていることが分かる。このことから、対話的（繰返し）検索が有効であったといえる。

上述の評価実験結果でも分かるように、候補画像出力に対して、「かま」のように、同じカテゴリ内の画像総数が多い場合には、ヒット率は必然的に低くなる。

該当する画像をよく選び出していれば、出現率と得点率は高くなる。このような場合は、出現率、得点率での評価が妥当であろう。逆に、同じカテゴリ内の画像総数が少ない場合には、該当する画像をよく選び出しても、すなわちヒット率が高くても、出現率と得点率は低くなる。このような場合は、出現率と得点率だけを用いたのでは、正當に評価したとはいえない。このように同じカテゴリ内の画像総数が、評価に大きく影響を与える。このように、画像データを対象にした場合、評価基準の選定は、非常に難しい問題であることが分かる。今回の評価実験でも、このことが如実に現れている。

また、ここで評価基準として用いたヒット率、出現率と得点率は、すべて評価者の主観の影響を受ける評価基準であることに注意されたい。たとえば、ヒット率の計算では、画像データベースのすべての画像を何らかのカテゴリに分類し、次にそのカテゴリ内の画像総数に対する候補画像の出力枚数の比で表す。このような方法を用いる場合、画像データベースを、何個のカテゴリに分類すれば妥当であるかは、データベースの利用者の主観によって左右される。前述の「着物」と「はんでん」の例の場合のように、どういうカテゴリを作成し、どの画像をどこのカテゴリに分類するかは、検索者の検索意図、利用方法などの主観によるものが大きい。

同様に、出現率と得点率もまた検索者の検索意図によって大きく左右する。たとえば、検索者が「ざる」をイメージして検索し、「竹箆」が候補として出力された場合、「竹箆」は「ざる」に属するのか、その逆に「ざる」は「竹箆」に属するのかは、検索者の主観による。

以上述べたように、画像内容に基づく類似画像検索システムの評価は、検索者の主観が大きく関係し、客観的な評価が難しい。したがって、多くの方々はこのシステムを利用していただき、今後の評価を待ちたい。

なお、索引空間の評価の 1 つとして、データベース中の画像を質問画像として輪郭画像検索を行ったとき、質問画像と同じ画像がどの順位で候補として出力されるかを調べた。この結果を表 6 に示す。また、例示画像検索と縮小画像検索の評価については、文献 3) を参照していただきたい (画像数 112 枚の実験ではあるが、5 位以内に入るのが 98% であった)。

さらに、画像データベースの原画像の総容量と索引空間のサイズを表 7 に示す。このように、索引空間は原画像データの約 100 分の 1 のデータサイズで済み、提案システムは、大容量の画像データにも対応が可能

表6 スケッチ検索で表示された画像の1枚を質問画像として検索したときの質問画像が表示された順位

Table 6 The order of appearance.

画像番号	r1	r2	r3	r4	r5	r6	r7	r8
順位	1	28	3	4	1	1	1	1

表7 画像容量と索引空間のサイズの関係

Table 7 Size of image database, feature space, and index space.

データの種類	データサイズ	圧縮比 (%)
原画像データ	466.2 MB	-
特徴量空間	37.4 MB	8.0
索引空間	6.3 MB	1.3

である。

6. まとめ

本稿では、画像内容に基づく類似画像検索システムのための画像の自動索引生成方法、対話的検索方法を提案した。その枠組みは、以下のとおりである。

- (1) 画像解析方法として、ウェーブレット変換による多重解像度解析を用いる。
- (2) 多重解像度空間から画像内容に応じた画像特徴量を計算する。
- (3) 縮小画像特徴量と輪郭画像特徴量の算出に、高次局所自己相関関数を用いる。
- (4) 大域的な大きな特徴と局所的な詳細な特徴を持ち合わせた特徴空間を作成する。
- (5) 特徴空間を効率的に使用するために判別分析を用いて索引空間を作成する。
- (6) 類似測度は、ユークリッド距離とし、検索手法は、全探索で行う。
- (7) 候補の出力結果を検索者が見て、その画像（そのものあるいは加工したもの）を次の質問画像として検索する対話的検索を行う。

この枠組みに基づき、対話的類似画像システムのプロトタイプを構築した。

プロトタイプシステムは、以下のような特徴を持っている。

- ウェーブレット変換による多重解像度解析と高次局所自己相関関数による特徴量抽出のみによって、スケッチ画像検索、輪郭画像検索、例示画像検索を実現している。
- 手描きのスケッチ画などの曖昧な手がかりから類似画像を検索できる。
- 写真や写真に手を加えた画像から類似画像を検索できる。
- 繰返し検索（対話的検索）を行うことができる。

- 類似画像検索だけではなく、画像データベース中の特定の画像検索もできる。
- ウェーブレット変換による多重解像度解析を用いることにより、縮小画像特徴量と輪郭画像特徴量を同時に計算することができる。
- 縮小画像索引空間と輪郭画像索引空間を同じ枠組みで作成することができる。
- 国立歴史民俗博物館の「もの」資料カードの白黒写真を3315枚を画像データを用いている。この写真は、記録者が現地で撮影したものでノイズが非常に多い。
- 画像データに多くの雑音が含まれていても、画像のノイズ除去や鮮鋭化などの前処理を必要としない。

なお、より柔軟で確実な検索を実現するためにテキスト型データベースとの融合を試みている。

参考文献

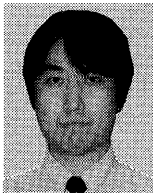
- 1) 赤間浩樹, 紺谷精一, 三井一能, 申間和彦: 画像内オブジェクトの自動抽出を使った画像検索システム Exsight—写真(PhotoDisk)への適用, 信学技報, DE97, No.31, pp.25-30 (1997).
- 2) Flickner, M., Sawhney, H., Niblack, W., Ashiley, J., Huang, Q., Com, B., Gorkani, M., Hafer, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P.: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol.28, No.9, pp.23-32 (1995).
- 3) 星守, 大森 匡: ウェーブレット変換を用いた類似・概略画像の検索方式と歴史民俗画像データへの適用, 平成7年度~平成8年度科学研究費補助金基礎研究(C)(2) 研究報告書 (1997).
- 4) 今村 寧, 中田雄一郎, 星守, 大森 匡: ウェーブレット変換を用いた画像のテキスト解析, 第51回情報処理学会全国大会, pp.197-198 (1996).
- 5) 加藤俊一, 栗田多喜夫: 画像の内容検索—電子美術館への応用, 情報処理, Vol.33, No.5, pp.466-477 (1992).
- 6) 小早川倫広: ウェーブレット変換を用いたテキスト領域分割と類似画像検索への応用, 修士論文, 電気通信大学大学院情報システム学研究科 (1998).
- 7) 栗田多喜夫, 下垣弘行, 加藤俊一: 主観的類似度に適応した画像検索, 情報処理学会論文誌, Vol.31, No.2, pp.227-237 (1990).
- 8) Mallat, S.: Wavelets for a Vision, *Proc. IEEE*, Vol.84, No.4, pp.604-614 (1996).
- 9) Mallet, S.: *A Wavelet Tour of Signal Processing*, Academic Press (1998).
- 10) 本谷佳代, 小早川倫広, 星守, 大森 匡: ウェーブレット変換を用いた画像データベースにおける

対話的類似画像検索方式, 信学技報, DE97, No.3, pp.13-18 (1997).

- 11) 中田雄一郎, 小早川倫広, 星 守, 大森 匡: ウェーブレット変換を用いたテクスチャ解析と類似画像検索への応用, 信学技報, PRMU69-166, pp.45-52 (1997).
- 12) 大津展之, 栗田多喜夫, 関田 巖: パターン認識—理論と応用, 朝倉書店 (1996).
- 13) Stollnitz, E.J., Derosc, T.D. and Salesin, D.H.: *Wavelets for Computer Graphics*, Morgan Kaufmann (1996).
- 14) Kurita, T., Otsu, N. and Sato, T.: A face recognition method using higher order local autocorrelation and multivariate analysis, *Proc. Int. Conf. on Pattern Recognition*, Hague, pp.213-216 (1992).
- 15) Unser, M.: Texture classification and segmentation using wavelet frames, *IEEE Trans. Image Processing*, Vol.4, No.11, pp.1549-1560 (1995).
- 16) Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C. and Taubin, G.: The QBIC Project: Querying Images By Content Using Color, Texture, and Shape, *SPIE Storage and Retrieval for Image and Video Database*, San Jose, pp.173-187 (1993).
- 17) Zobel, J. and Moffat, A.: Exporting the Similarity Space, *SIGIR*, Vol.32, No.1, pp.18-34 (1998).

(平成 10 年 8 月 31 日受付)

(平成 11 年 1 月 8 日採録)



小早川倫広

1966 年生。1998 年 3 月電気通信大学大学院情報システム学研究科博士前期課程修了, 同年 4 月電気通信大学大学院情報システム学研究科博士後期課程入学, 現在, 電気通信大学大学院情報システム学研究科博士後期課程在学中。データ工学, 多変量データ解析, データ圧縮に関心を持つ。電気学会, 電子情報通信学会各会員。



星 守 (正会員)

1970 年 3 月東京大学大学院工学系研究科修士課程修了, 同年 4 月電子技術総合研究所入所, その後千葉大学工学部をへて, 1992 年 4 月より電気通信大学大学院情報システム学研究科教授, 現在に至る。アルゴリズムとデータ構造 (特に探索法のための), 多変量データ解析, 数値解析に関心がある。IEEE Computer Society, IEEE Information Theory Society, Association for Computing Machinery, Intern. Association for Statistical Computing, 電子情報通信学会, 情報理論とその応用学会, 日本行動計量学会, 日本分類学会各会員。



大森 匡 (正会員)

1994 年 4 月から電気通信大学に助教として勤務。1990 年東京大学大学院情報工学専攻博士課程修了, 工学博士。データ工学, 特に並列データベース処理技術に関心を持つ。電子情報通信学会会員。



照井 武彦 (正会員)

1933 年生。1956 年東北大学工学部通信工学科卒業。同年より (株) 北辰電機。日本電気 (株) 勤務。1966 年より岩手大学工学部電子工学科, 情報工学科勤務。1980 年文化庁, 1981 年より文部省・国立歴史民俗博物館勤務, 現在, 同館情報資料研究部教授。日本の歴史・考古・民俗学を中心に人文学の情報処理研究に従事。電子情報通信学会, 情報通信学会, 情報考古学会, 各会員。