

# プリミティブ分解による多様な検索条件を扱うカラー画像検索

金原 史和<sup>†</sup> 佐藤 真一<sup>††</sup> 浜田 喬<sup>††</sup>

画像検索におけるひとつの重要な課題は、言葉では表現し難い画像メディア特有の検索要求への対応である。一般に、このような画像特徴に対する人間の検索要求は多分に曖昧でありかつ多様である。そこで本研究では、画像の特徴に関するユーザの多様な検索要求に応えることを目的とした画像検索プロトタイプシステムを構築した。本システムでは、領域分割処理を基に画像からプリミティブ領域を抽出し、それらの属性および領域間の関係を対象画像の特徴量とする。しかし、有意な領域分割を行うことは困難であるので、本研究では有意でない領域情報もプリミティブとして抽出して、検索のための特徴として用いる方針をとった。ユーザは例示画インタフェースにより所望の画像の特徴についての指示をすることにより検索を行う。このとき、着目する特徴や部分を明示的に“曖昧度”としてグラフィカルに指示できる。検索処理についてはシグネチャファイルを用いて効率化を図っている。これにより、従来とは異なる柔軟な例示画検索が行えるようになった。本論文では、本システムおよび提案する要素技術を通じて、ユーザの多様な検索要求を受容しうる画像検索の実現のためのひとつのアプローチについて論じる。

## Flexible Image Retrieval System Reflecting User's Diverse Query Requests Based on Primitive Extraction

FUMIKAZU KANEHARA,<sup>†</sup> SHIN'ICHI SATOH<sup>††</sup> and TAKASHI HAMADA<sup>††</sup>

We have developed a flexible image content retrieval prototype system for color images. In our process, primitives, which represent regions, are extracted from a color image by segmentation and merge processing, and a set of them are translated into two *signatures* as indices for image retrieving, which are respectively for attributes of each primitive and for relations between two of them. Since it is too difficult to extract only meaningful regions without the use of a priori knowledge, we have made a point of utilizing both meaningful and meaningless regions for retrieval. Queries are performed by the user's explicit instructions about attributes and relations with our developed image retrieval interface. Thus, by dealing with meaningless primitive, our system is capable of automatically extracting several features of the original image; thereby allowing a user to initiate queries representing diverse demands, intentions, and viewpoints in regard to local/global feature of the image.

### 1. はじめに

近年、計算機ハードウェアの発展にともない、ますます画像データベースへの需要が高まり、そのひとつの重要な要素である画像の類似検索に対する関心も強くなってきている。類似画像検索は、キーワードのような言葉では表現しにくい検索要求を満たすために必要と考えられるテーマである。

従来の類似画像検索の研究では例示画像をキーとす

る検索形態が多く、その例示画の与え方に差はあるものの、ユーザが思い描くイメージを例示画として表現する点においては共通している。しかしながら、ユーザのイメージや検索要求というのは一般的に曖昧かつ多様である。よって特定の目的の場合を除いて、柔軟な画像検索を提供するには、ユーザの検索の視点や意図をシステムに反映させることが必要になる。しかし、これは困難かつ漠然とした課題である。そこで本論文では、焦点を少し絞り、画像のどの部分のどの特徴にユーザが着目しているか、ということを検索に反映させることを考える。

このような観点から従来の研究をみた場合、画像や対象物体の大局的な類似度に重点を置いたものが多く<sup>1),2)</sup>、ユーザの視点の入り込む余地が少ない。一方、ユーザの着目部分に考慮したものとして、画像内のオ

<sup>†</sup> 東京大学大学院工学系研究科電子工学専攻  
Department of Electronic Engineering, Faculty of Engineering, University of Tokyo  
<sup>††</sup> 学術情報センター研究開発部  
Research and Development Department, National Center for Science Information Systems

プロジェクトの位置関係や<sup>3),4)</sup>, 部分画像を考慮した検索などの研究<sup>5),6)</sup>もあるが, 着目する特徴は反映されていない. 逆に検索条件として複数の特徴をユーザに提供している研究<sup>7),8)</sup>もあるが, どこに着目しているかは反映されていない.

そこで我々は以前, このようなユーザの視点を反映した画像検索を実現することを目的とし, シルエット画像を対象として, 形状特徴のみに焦点をあてた画像検索について検討した<sup>9),10)</sup>. 我々は, さらに改良を重ねるとともにカラー画像への拡張を行い, 画像検索プロトタイプシステムを構築した. 本論文では, 画像の特徴記述のための領域 (プリミティブ) 抽出方法および検索処理の改良について論じ, パターン画像の例である国旗画像と対象不特定の自然画像を対象例として, 本システムの有効性を検証するとともに, ユーザの多様な検索要求に応える画像検索の実現のためのひとつのアプローチを示す.

もちろん, 本研究と同じような視点に立ったと考えられる研究は存在する. Yihongの研究<sup>11)</sup>, PinPoint<sup>12)</sup>, QBIC<sup>13)</sup>などがそうである. しかし, 本システムでは, オブジェクト間の関係も特徴として用いている点, 対象画像に対して特別な知識やモデルを用いていない点, 原画像処理の際に人手を介さない点がこれらの研究と異なる. もちろん, 画像から有意な領域を抽出することは困難な技術であるので, 本研究では有意であるか否かにかかわらず領域を抽出して検索に用いるアプローチをとった.

以下, 2章でシステムの概要について述べ, 3章でプリミティブ抽出方法, 4章において検索処理について論じる. 5章では検索例を紹介するとともに評価を行い, 6章をもってまとめとする.

## 2. 検索処理の概要

本システムでは, 多角形オブジェクトの描画などでユーザが作成した問合せ画像中のオブジェクトが, ユーザが設定した検索条件のもとでデータ画像中にどれだけ含まれるかを測ることで検索を行う (図1).

このためには, データ画像から検索キーとなりうるオブジェクト領域をあらかじめ抽出しておかなければならない. しかし, これは対象を特定したり人手を介さない限り, 非常に困難な課題である. そこで, 本研究では有意であるか否かにかかわらず冗長に領域を抽出する方針をとる. この余計なものを含めた領域をプリミティブと呼ぶことにする.

一方, ユーザは, 描画した個々のオブジェクトごとに, 用意された特徴の中で着目する特徴の“ゆらぎ”

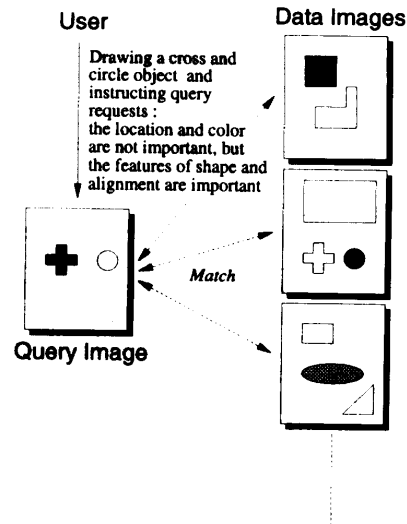


図1 検索処理の概念

Fig. 1 Concept of retrieval processing.

を指示してオブジェクトの集合である問合せ画像を作成する. この“ゆらぎ”はユーザの曖昧度また検索条件の厳しさを表すものである. すなわち, ゆらぎが大きく指示された場合は, 該当する特徴がユーザにとって曖昧であるか重要でないことを意味し, ゆらぎが小さい場合はその逆を表す. この際, 多様な検索要求を受容しうるために, オブジェクト間の関係も特徴として同様に扱う.

問合せ画像のオブジェクトの集合も, データ画像から抽出されたプリミティブの集合も, 後述するシグネチャに変換されて類似度が測られる.

## 3. プリミティブ抽出

### 3.1 抽出アルゴリズムの概要

まず, データ画像からのプリミティブの抽出手法から述べる. 本手法は領域分割処理に基づくが, 従来の多くの手法<sup>14)~17)</sup>のように分割結果の精度の向上を直接の目的とするのではなく, 有意な領域を少しでも多く抽出することを目指す.

図2に処理の概略を示す. まず前処理を施した対象画像からエッジ画像と初期分割画像の作成および各領域の併合処理を行う. この併合の過程においてプリミティブ領域を抽出し, 形状分解を施し, 最後に補正処理を行って出力する. 以下, 個々の処理について述べる.

### 3.2 プリミティブ抽出方法

#### 3.2.1 分割処理

まず, 原画像をガウスフィルタにより平滑化し, sobel オペレータを用いてエッジを抽出する. 得られた

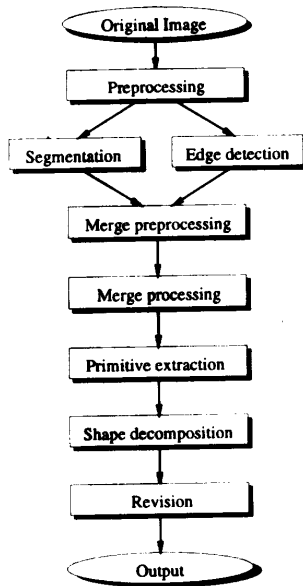


図2 分割処理のフローチャート  
Fig.2 The flow of algorithm.

エッジ輝度を  $I(x, y)$ , その最大エッジ輝度を  $I_{max}$ , 最小エッジ輝度を  $I_{min}$  として, 以下の式のように  $I(x, y)$  を正規化したものを  $f(x, y)$  とする.

$$f(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}} \quad (1)$$

一方, 色ヒストグラムを分割する再帰処理方法<sup>18)</sup>を用い, 平滑化後の画像を数百から数千の細かい領域に分割する. これを初期分割画像とする.

### 3.2.2 併合処理

初期分割画像は複数の領域集合  $\{R_i\}$  ( $i = 0, 1, \dots, N$ ) として表される.  $R_i$  は面積 ( $A_i$ ), RGB 色 ( $r_i, g_i, b_i$ ) の属性を持つ. 領域の色は該当する原画像の画素値の平均値とする. また任意の異なる2つの領域に含まれる画素どうしが4連結関係にある場合, その2つの領域は隣接関係にあると記す.

すべての隣接関係にある2つの領域  $R_i$  と  $R_j$  について, 併合した場合のコスト  $cost(R_i, R_j)$  が最小の組から併合する. すなわち階層的クラスタリングを行う.  $cost(R_i, R_j)$  を以下のように定義する.

$$cost(R_i, R_j) = \frac{E_i + E_j - E_{ij}}{B_{ij}}$$

$$E_i = (r_i^2 + g_i^2 + b_i^2)A_i$$

$$E_{ij} = (r_{ij}^2 + g_{ij}^2 + b_{ij}^2)(A_i + A_j) \quad (2)$$

$E_i$  は領域  $R_i$  のエネルギーを表しており,  $r_{ij}, g_{ij}, b_{ij}$  は領域  $R_i$  と  $R_j$  とを併合した場合の平均色である. すなわち,  $r_{ij} = (A_i r_i + A_j r_j) / (A_i + A_j)$  であり,  $g_{ij}, b_{ij}$  も同様である.

$B_{ij}$  は領域  $R_i$  と  $R_j$  との境界値である. ここで

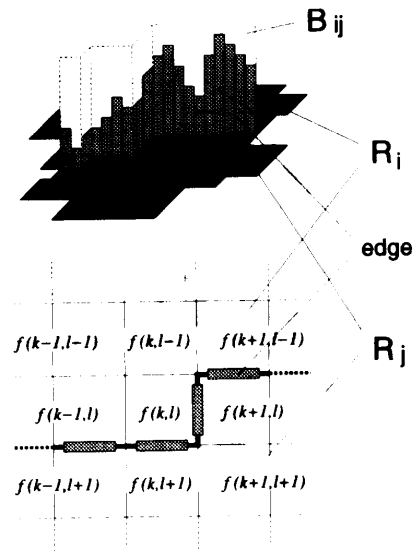


図3 エッジと併合領域  
Fig.3 An edge and regions to be merged.

いう境界値とは, 領域の併合の容易さを示すもので, エッジ情報を保存することと形状特徴を反映させることを目的としている. 具体的に  $B_{ij}$  は先に求めたエッジ輝度  $f(x, y)$  から計算される (図3). まず,  $R_i$  と  $R_j$  の隣接関係にある2つの画素に相当するエッジ輝度  $f(k, l), f(k', l')$  の平均を求めて画素間のエッジ輝度とする. そして, 最大エッジ輝度 (1.0) からこのエッジ輝度を差し引いた値についての境界すべての総和を  $B_{ij}$  とする.

領域  $R_i$  と  $R_j$  が併合された場合, その併合領域の属性は  $A_{ij} = A_i + A_j, r_{ij}, g_{ij}, b_{ij}$  になる. このように併合した場合の領域の色は2つの領域の平均色となる.

この  $cost(R_i, R_j)$  は, 小さな領域や画素値が近いものほど早期に併合され, しかも隣接境界エッジ輝度が強いものや隣接関係にある画素数が少ない領域どうしは併合されにくいという考えのもとに定義されている.

なお, 以上の手順では, すべての領域の中から最もコストの低い領域のペアを探さなければならないので効率が悪い. そこである一定以下の微小領域は, その領域と最小のコストで併合できる領域と強制的に併合させる (図2中の merge preprocessing に相当). つまり微小領域に関してはその併合順序に制限をなくすわけである. 具体的には, 原画像の総画素数に一定比率  $\epsilon_r$  を乗じた値よりも小さな画素数を持つ領域を微小領域とする.

### 3.2.3 プリミティブの決定

以上の処理を領域が1つになるまで行い, 併合の履歴をとる. 図4は併合の様子をノードとアークで示し

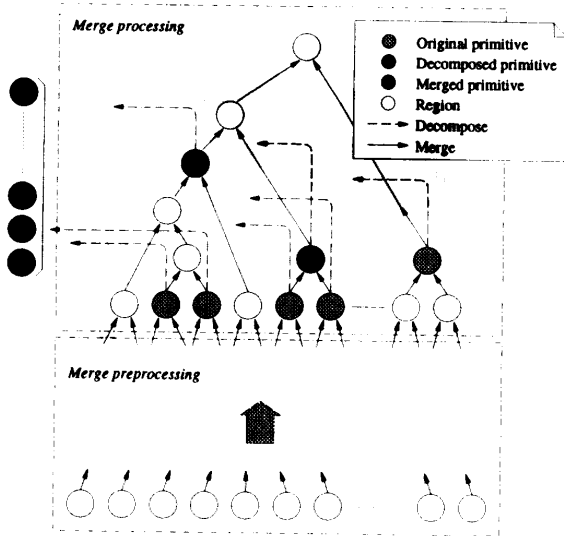


図4 プリミティブの抽出  
Fig. 4 Extraction of primitives.

たものである。ノードは領域に相当し、実線のアークは併合の過程を示し、全体として木構造になる。

さて、理想的には検索に用いることができる有意な領域をプリミティブとして抽出したいのだが、自動的に識別することは困難であるので、併合時のコスト  $cost(R_i, R_j)$  が大きい領域を有意である可能性があると考え、プリミティブとして抽出することにする。まず、領域が1つになるまでの全コスト  $cost_{all}$  を併合の履歴より計算し (merge preprocessing のコストも含む)、それに対し一定の割合 ( $\epsilon_c$ ) を上回るコストを要した領域をプリミティブとして出力する。つまり、 $cost(R_a, R_b) > cost_{all} \times \epsilon_c$  ならば  $R_a, R_b$  はプリミティブとして出力する。ここで、 $R_a$  と  $R_b$  がプリミティブと判定されたとして、もし  $R_a$  を根とする部分木内にすでに出力されたプリミティブがある場合、 $R_a$  を merged primitive と呼ぶことにする。また、そうでない場合は  $R_a$  を original primitive と呼ぶことにする。 $R_b$  も同様である。

なお、出力される際のプリミティブの色は平均色ではなくて代表色を用いる。ここで代表色とは、原画像中の該当領域内で最も出現頻度の高い色を指す。頻度は RGB 空間を  $16 \times 16 \times 16$  に量子化して計上している。平均色では色が濁ってしまい、検索の際に適当でないからである。

### 3.2.4 不適当なプリミティブの処理

このようにプリミティブを抽出するが、色とエッジ輝度を基にしたプリミティブを決定しているために、形状に関する制約が緩い。つまり実際には類似した形状にもかかわらずコストが大きいため無意味なプリ

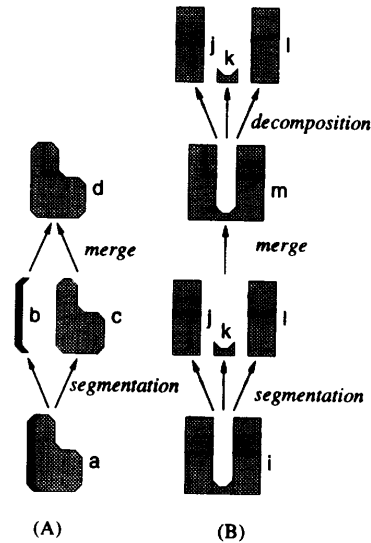


図5 不十分なプリミティブの判定  
Fig. 5 Extraction of inappropriate primitives.

ミティブを出力してしまう場合 (図5(A)) と、併合時に形状的に大きな変化があるにもかかわらずコストが小さいためにプリミティブとして出力されない場合 (図5(B)) が起きてしまう。

図5(A)では、 $b, c$  がプリミティブとして出力されたとして、 $d$  も後の併合過程でプリミティブとして出力されるおそれがある。明らかに  $c$  と  $d$  は類似しており、2つを出力するのは無駄である。そこで、この対策として周長と面積を考慮する。今、領域  $R$  がコストの条件を満たして、merged primitive として出力されうるとする。ここで  $R$  を根とした各葉の方向へのすべての経路の中で、最初に見つかるすでに primitive として出力済みの領域を  $R_i (i = 0, 1, 2, \dots)$  とし、その中で最も面積の大きい primitive を  $R_k$  とする。そして以下の式のどちらかの条件を満たした場合、 $R$  は merged primitive として出力されうるとする。これは形状の変化が小さければ、たとえコストが大きくとも出力しないための条件になる。ここで  $L_x, A_x$  は領域の周長と面積を表す。図5(A)では、 $d$  と  $c$  の周長と面積の差に変化がないので、 $d$  は出力されない。もちろん、この判定基準だけでは完全ではないが、ある程度の絞り込みはできる。

$$\begin{aligned} |L_R - L_{R_k}| / L_{R_k} &> \epsilon_p \\ |A_R - A_{R_k}| / A_{R_k} &> \epsilon_p \end{aligned} \quad (3)$$

次に、図5(B)では  $j, k, l$  が出力されないおそれがある ( $k$  はどちらにしろ不要であるが)。また、領域分割の段階ですでに  $i$  の状態以上に分割されないかもしれない。これに対しては図5(B)のように、以前我々が提案したシルエット画像に適用した形状分解手

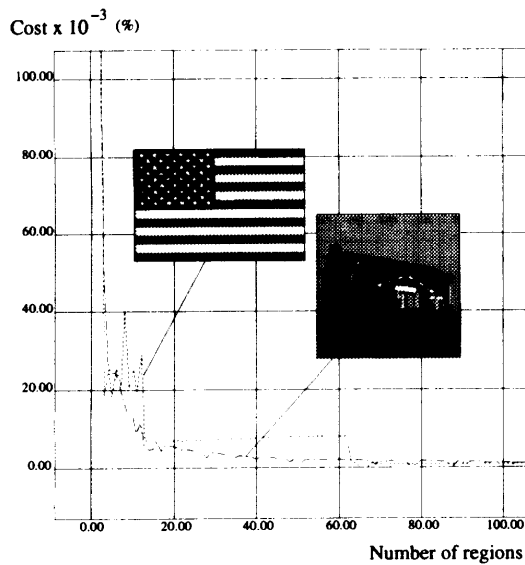


図6 領域数と併合コストの関係

Fig. 6 Graph regarding region count and cost.

法<sup>10),19)</sup>をここで用いて分割する。得られたプリミティブを decomposed primitive と呼ぶことにする (図4参照)。

なお、最後に後処理として、まったく重複するプリミティブと、形状分解によって生じた微小なプリミティブを削除する。

### 3.3 抽出例

さて、以上の処理において決定すべきパラメータの値について補足しておく。

$\epsilon_r$  は、ある程度小さい値にしておけば結果に及ぼす影響は小さく、0.001 以下が適当である。

次に  $\epsilon_c$  であるが、まず数枚の画像について、領域が1つになるまでの各併合処理で算出されるコストを測定した。その測定例を図6に示す。その結果、自然画では全体的に細かな増減を繰り返しながら滑らかにコストが上昇していくが、パターン画像の場合は色が比較的明瞭に区別されているために急激にコストが変化する。明らかに、これらの画像に一樣に最適な閾値  $\epsilon_c$  を設けるのは困難なことであるが、ここでは余計な領域を含んでも構わないので、プリミティブの数が多くなりすぎない程度に低い値にしておけばよい。複数のサンプル画像についての測定結果から、0.001~0.005 くらいの範囲が適当であることが分かった。

最後に  $\epsilon_p$  であるが、この値を大きくすれば merged primitive の判定基準を厳しくすることになる。人間の直観を考慮しなければならないが、元の周長や面積の半分程度の変化が起これば大きい変化であるとして0.5あたりを目安とする。

さて、以上の処理結果例を図7に示す。ここでは、

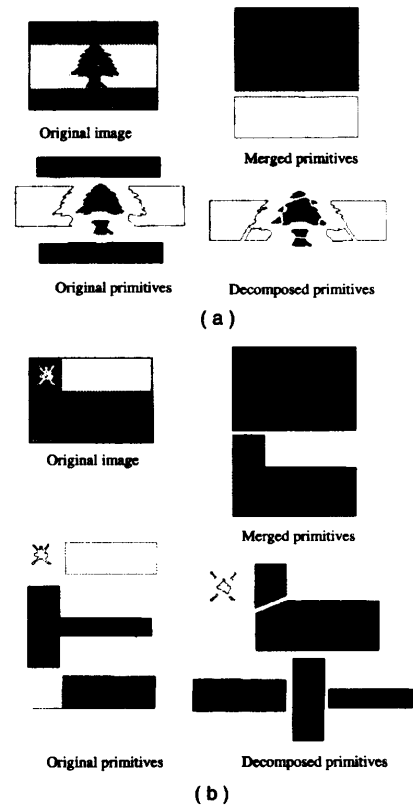


図7 国旗画像のプリミティブ抽出例

Fig. 7 Examples of the primitive extraction.

本手法の主張点が分かりやすくするためにパターン画像の例である国旗画像の処理例を示す。スキャナから取り込んだ  $450 \times 300$  の 24 bit の国旗画像である。パラメータは  $\epsilon_r = 0.0001$ ,  $\epsilon_c = 0.005$ ,  $\epsilon_p = 0.5$  とした。(a)では背景の白地が併合処理によって矩形として抽出できている。また、(b)では横向きのT字形と雑音の突起がある original primitive が分解されていることに注目されたい。

## 4. 検索処理

### 4.1 シグネチャ表現

前節までの処理で、データ画像から抽出されたプリミティブの集合は、連続したビット列であるシグネチャ表現<sup>20)</sup>に変換される。この内容に関してはすでに文献10)において議論したが、ここではもう一度簡単に述べるとともに、新たにマッチング処理の改良点について述べる。

各プリミティブはいくつの特徴量によって分類され、固定長のプリミティブシグネチャに表現される。各属性値は定められたビット数の範囲内で量子化され、該当する位置にビットが立てられ、画像全体はこのプリミティブシグネチャが連結したもので表現される(属性データシグネチャ)。また、2つのプリミティブ間の

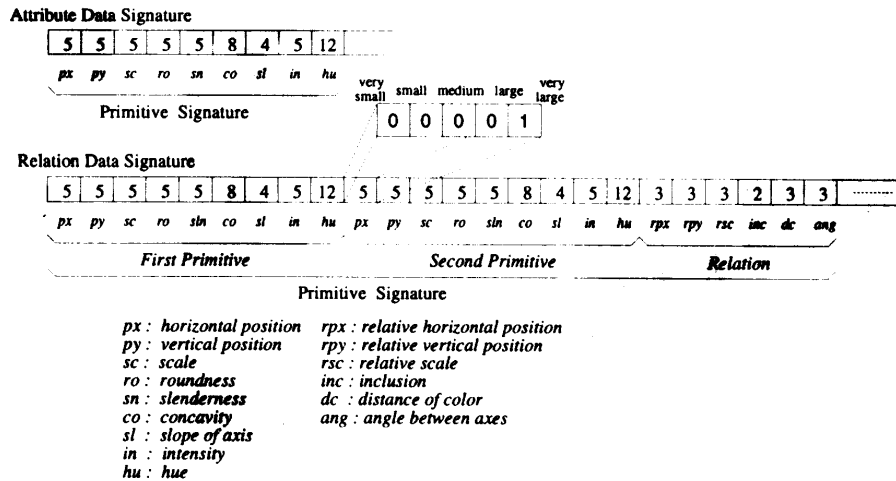


図8 シグネチャの形式

Fig. 8 Format of signature.

関係についてもその関係を表す値によって分類され、同様に関係データシグネチャが構成される。ただし関係のプリミティブシグネチャに関しては関係を結ぶ2つのプリミティブの属性のプリミティブシグネチャを先頭に付加する。この際、面積の大きいプリミティブを first primitive, 小さい方を second primitive とし、その順序に従って関係のプリミティブシグネチャの前に付加される。図8は国旗画像を対象とした場合のシグネチャの構成例である。個々のボックスはそれぞれ1つの属性または関係に相当し、その中の数字はビット数を表す。

本研究において具体的にどのような特徴を扱うかは、アプリケーションに依存するが、ユーザの検索要求を想定して多角的にプリミティブを分類できるような特徴を用いることが重要である。個々の特徴は単純かつ低レベルなものでも構わない。

このような視点に基づいて、以下に国旗画像を例とした場合に用いた属性と関係を記す。

#### 【属性】

- (1) 水平方向位置：プリミティブの中心の x 座標。5ビットを割り当てる。
- (2) 鉛直方向位置：プリミティブの中心の y 座標。5ビットを割り当てる。
- (3) 大きさ：画像全体の面積に対するプリミティブの面積の比率。5ビットを割り当てる。
- (4) 円形度：面積/(周長 × 周長) × 4π。5ビットを割り当てる。
- (5) 針状度：まずプリミティブの凸包を求め、その中心と主成分を主成分解析によって求める。次に求めた第一主成分を水平になるように凸包を回転させて、その状態での外接矩形を求める。

この外接矩形の縦の長さの横の長さに対する比を針状度と定義する。5ビットを割り当てる。

- (6) 凹凸度：プリミティブの輪郭を平滑化し、さらに多角形近似する。そのうえで各頂点の曲率を計算して凹凸のペアの数を計上する。8ビットを割り当てる。
- (7) 傾き：プリミティブの輪郭の主成分を求め、その第一主成分の傾きを軸の傾きとする。4ビットを割り当てる。
- (8) 明度：領域の RGB の代表色を HSI 変換して、その明度を用いる。5ビットを割り当てる。
- (9) 色相：同じく HSI 変換したもののうちの色相を用いる。12ビットを割り当てる。

以上計9属性、54ビットとなる。

#### 【関係】

面積の大きいプリミティブを第一プリミティブ (First Primitive), もう一方を第二プリミティブ (Second Primitive) として関係を結ぶ。

- (1) 相対水平方向位置：第一プリミティブから見た第二プリミティブの x 方向の位置関係。3ビットを割り当てる。
- (2) 相対鉛直方向位置：第一プリミティブから見た第二プリミティブの y 方向の位置関係。3ビットを割り当てる。
- (3) 相対的な大きさ：2つのプリミティブの面積比。3ビットを割り当てる。
- (4) 包含関係：第二プリミティブが第一プリミティブに包含されているかどうかを示す。2ビット(包含, 非包含)を割り当てる。
- (5) 色関係：色相間の距離。3ビットを割り当てる。
- (6) 角度：2つの軸がなす角度。3ビットを割り当

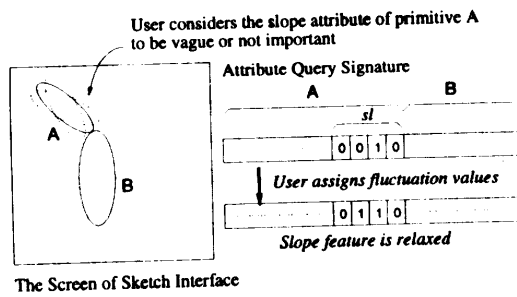


図9 ゆらぎを含む問合せシグネチャ  
Fig. 9 Query signature with the fluctuation.

てる。

計6関係17ビットで、1つの関係データシグネチャのプリミティブシグネチャは125ビットになる。相対水平方向位置関係は、第一プリミティブの占めるx方向の幅と第二プリミティブの中心のx座標との位置関係を示す。鉛直方向も同様である。色関係は双方のプリミティブの色相が定まっている場合のみ設定され、マンセルの色相環における2つの色相間の角度を距離とする。なお、各プリミティブの凸包の面積に比例した量だけそのプリミティブの凸包を膨張させて、交わったプリミティブどうしのみ関係を結ぶようにした。

4.2 問合せシグネチャ

マッチングは個々のデータシグネチャと問合せ画像から生成される問合せシグネチャとの間で行われる。すでに述べたように、データシグネチャは各属性と関係の特徴の存在をビット列によって表現しているが、問合せシグネチャでは、これに加えてゆらぎの情報が入り込む。前述したように、ユーザが注目しているまたは明確な特徴はゆらぎの値は小さく、注目していないまたは曖昧な特徴はその値が大きく指示される。この値は0.0から1.0までの値をとり、この値によって各属性・関係の値は幅を持ち、シグネチャ内の立つビットの数も変化する(図9)。

今、ある属性のとりうる値aの範囲が  $A_{min} \leq a \leq A_{max}$  として、ゆらぎの値が  $\sigma$  ( $0.0 \leq \sigma \leq 1.0$ ) であるとすると、属性値aは以下の式により最小値と最大値が与えられる。

$$a_{min} = a - (A_{max} - A_{min}) \times \sigma / 2.0$$

$$a_{max} = a + (A_{max} - A_{min}) \times \sigma / 2.0 \quad (4)$$

すなわち、該当するシグネチャは  $a_{min}$  に該当するビットと  $a_{max}$  に該当するビットの間に1がたったものになる。ただし、ゆらぎの値が1.0の場合は、その特徴に関してはまったく制約がないことを示し、すべてのビットに1をたてる。

4.3 マッチング処理

類似度は、上述のシグネチャを用いて、問合せ画像

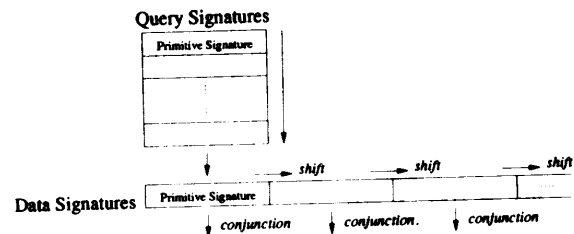


図10 マッチング処理(1)~プリミティブの対応~  
Fig. 10 Matching process.

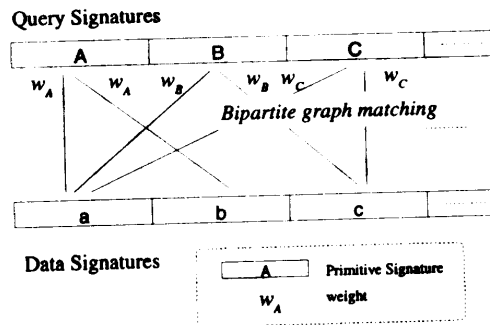


図11 マッチング処理(2)~2部グラフマッチング~  
Fig. 11 Matching process.

中のオブジェクトに相当するプリミティブがデータ画像中に含まれる(ここでは、この場合をマッチングと呼ぶ)割合を測ることで決定する。類似度は二段階の処理によって測られる。

第一段階は図10で示されるように、まず個々の問合せシグネチャとデータシグネチャとの間でシフト演算と論理積演算が行われる。もし問合せシグネチャ内のプリミティブシグネチャとデータシグネチャ内のプリミティブシグネチャとの論理積の結果が、そのデータシグネチャのプリミティブシグネチャと一致すれば、そのデータのプリミティブは検索要求に合致する可能性があることになる。このような可能性を持つプリミティブの対応の組合せをすべて見つける。この対応関係をノードとエッジで示したのが図11である。

第二段階の処理は、この可能性ある対応関係の中から最適の対応関係を求めることである。最適というのは対応関係が最も多くなる組合せのことをここでは指す。当然、これがユーザの検索要求に対して最適であるとは限らないが、少なくとも検索洩れを抑えることは期待できる。

この際、今まで1つのオブジェクト内の特徴間の重要度または曖昧度は考慮されていたが、オブジェクト間の重要度はすべて同じであった。たとえば、大きなオブジェクトがマッチングしても小さなオブジェクトがマッチングしても、類似度は同じになる。これは明らかに検索の精度を悪くする一因になる。しかし、こ

れまでもユーザに指示させるのは煩雑であるので、ここでは、以下のように単純に大きさに比例した重みを問合せ画像のプリミティブシグネチャに与えることにした。

$$w_k^A = A_k / \sum_{i=1}^N A_i \quad (5)$$

$$w_{k,i}^R = (w_k^A + w_i^A) / \sum_{(P_i, P_j) \in R} (w_i^A + w_j^A) \quad (6)$$

$A_i$  はオブジェクト  $i$  の面積であり、 $w_i^A$  と  $w_{i,j}^R$  は各々、属性と関係のプリミティブシグネチャの重みを表す。 $N$  は問合せ側のプリミティブの数である。 $w_{i,j}^R$  はプリミティブ  $i$  と  $j$  が結ぶ関係の重みであり、 $(P_i, P_j) \in R$  は関係が存在しうるプリミティブの組を示す。

結局、最適な対応関係を求めることは、2部グラフの重み付き最大マッチング問題を解くことになる。さらに、この場合は問合せ側の1つのプリミティブシグネチャから出る対応関係を示すエッジの重みはすべて同じ値を持つので(図11)、Hungaryの増大道を見つける方法<sup>21)</sup>で解くことができる。

最終的には類似度は以下の式によって定義される。

$$C_a = \sum_{i \in M} w_i^A \quad (7)$$

$$C_r = \sum_{(i,j) \in M} w_{i,j}^R \quad (8)$$

$$\text{Similarity} = \frac{C_a N_a + C_r N_r}{N_a + N_r} \quad (9)$$

$M$  は最大マッチングに成功したプリミティブの集合であり、 $N_a$  および  $N_r$  は問合せ属性シグネチャおよび関係シグネチャ中のプリミティブシグネチャの数であり、 $C_a$  と  $C_r$  はそれぞれ属性および関係シグネチャに関して最大2部グラフマッチングによって算出された類似度である。

## 5. 検索例と評価

### 5.1 国旗画像検索

国旗画像を対象としたのは、概念をとまなわないパターン画像の例として適切であり、本システムの実験に都合が良かったからである。よって国旗画像の検索システムの構築そのものが目的ではないことに注意されたい。

以下に検索例を示す。扱ったのは190枚の450×300の24bitの国旗画像である。プリミティブの数の平均は15.8であった。

まず図12の検索例での検索要求は「緑地に白の三

日月がある国旗画像」である。ユーザは図12(a)に示される検索インタフェースを用いて、長方形と三日月形の多角形オブジェクトを描画する。次に、各オブジェクトを選択して、ポップアップウィンドウを開き、属性または関係のゆらぎをスクロールバーを操作することで設定する(図12(a)および拡大図(c))。ゆらぎの値は初期設定では0.5であり、バーを右に移動させるとゆらぎの値は大きくなり、左に移動させるとゆらぎが小さくなる。このポップアップウィンドウに表示される属性は上段から順に、水平方向の位置、鉛直方向の位置、大きさ、円形度、針状度、凹凸度、傾き、輝度、色相を表す。関係についても同様に、上段から水平方向位置関係、鉛直方向位置関係、相対的な大きさ、包含関係、色関係、角度を表す。

この検索例では、三日月の輝度、凹凸度、円形度のゆらぎを小さくして色と形状に関する条件を厳しくした(図12(a)中の右側の属性ウィンドウ)。また緑地については、大きさや形状は不明であるが、だいたい長方形であることを想定して条件を設定した(図12(a)中の左側の属性ウィンドウ)。関係に関しても同様の操作で、包含関係と色関係のゆらぎを最小にして他は最大にした。検索結果は5件であり、余計なものが1件含まれていたが検索洩れはなかった(図12(b))。なお閾値は0.9である。

次に示す図13の検索例の検索要求は「赤白緑で全体をおおまかに横に区切った国旗画像」である。この場合、ユーザが描画した例示画像において、個々の長方形オブジェクトの鉛直方向に関する属性と関係のゆらぎを最大にし、その他はすべてゆらぎを最小にした。この設定のもとで、閾値0.9で検索数3件であり、ブラウジングした結果、図13(b)中の1, 7, 8番目の画像が検索された。ここで、もし、鉛直方向の位置関連の条件を厳しくしたならば、描画された色の並び方と同じ画像しか検索されないことに注意されたい。

次に、閾値0.9で位置と大きさに関する属性と関係のゆらぎを最大にした。この結果が図13(b)である。2, 5番目の画像は余計であり、図13(b)の枠内に示した2件が検索洩れである。ちなみに、形状に関する属性のゆらぎを大きくしたところ、この2件も検索結果として含まれた。

なお、上記の検索例では検索時間はいずれも約1秒ほどであった。

### 5.2 自然画像検索

次に、人物画やスポーツシーンや風景画などの対象不定の自然画像を扱った検索例を示す。扱う画像は、256×256の24bitの自然画像300枚である。プリミ



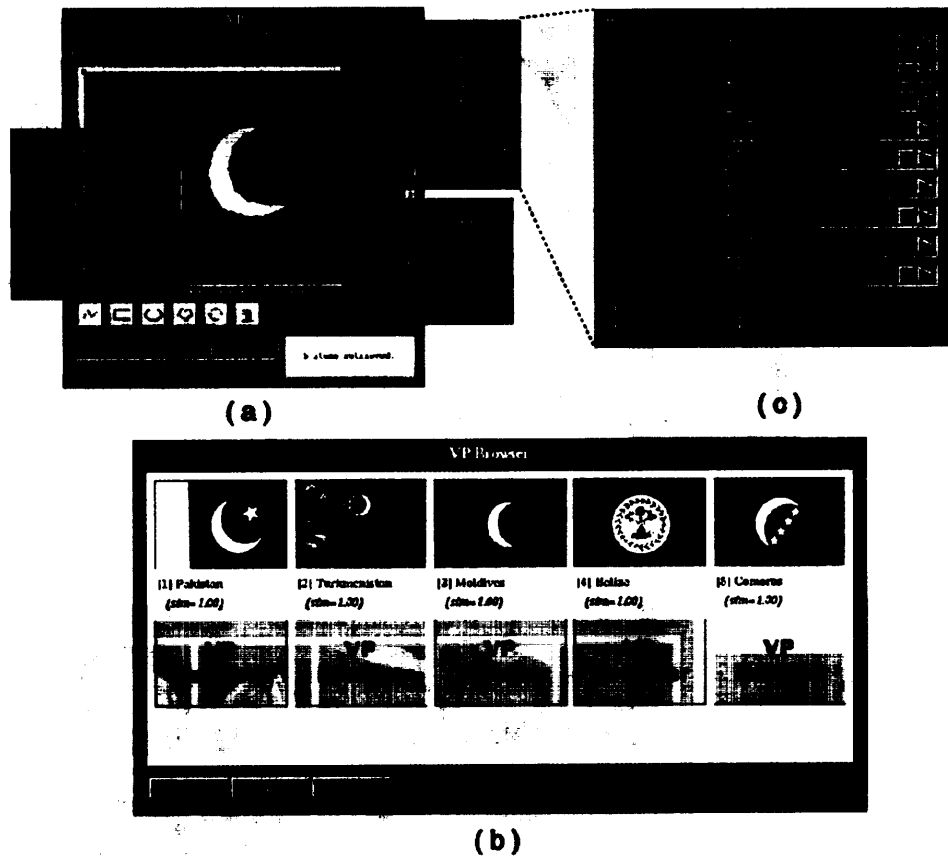


図 12 国旗画像検索例(1)

Fig. 12 Retrieval of national flags(1).

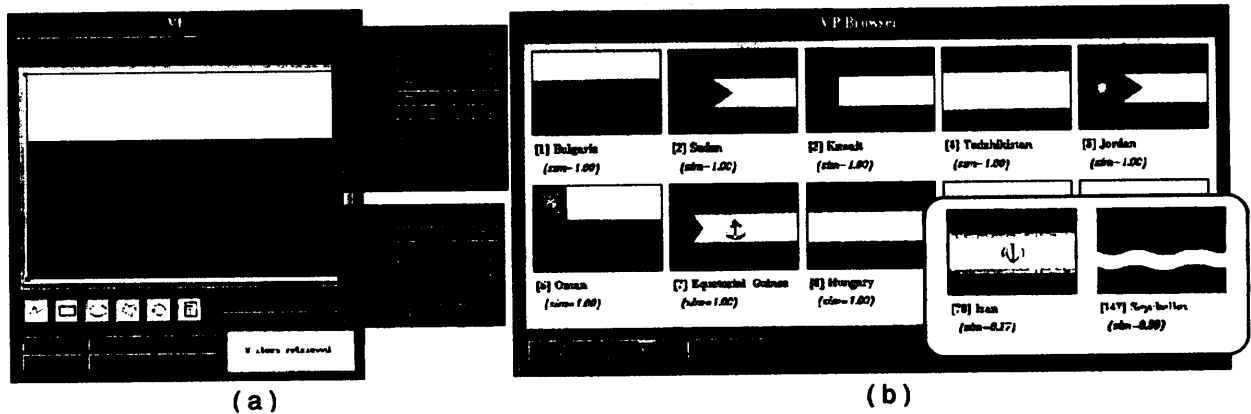


図 13 国旗画像検索例(2)

Fig. 13 Retrieval of national flags(2).

タイプを抽出する際には、処理時間の短縮のために128×128に縮小してから処理した。他はほぼ国旗画像同様に扱った。パラメータは $\epsilon_r = 0.001$ ,  $\epsilon_c = 0.003$ ,  $\epsilon_p = 0.75$ である。プリミティブの数は平均52.5であった。特徴に関しては、国旗画像の場合と同様であるが、自然画像ではオクルージョンや陰影などの影響で凹凸度の属性と角度関係の特徴の信頼性が極度に落ちるので省いた。よって用いる特徴は8属性46ピッ

ト、5関係14ビット(プリミティブシグネチャとしては106ビット)である。以下に検索例を示す。

図14に示す検索例における検索要求はヨットを念頭において「上部の背景が青色で船体の色が黄色っぽく三角に近い帆を持つ画像、ただし帆の色は不明」である。帆の色は不明なので、属性は分けて指示する。すなわち図14(a)中の右側上の属性ウィンドウが空と船体を示すオブジェクトのためのものであり、真中の

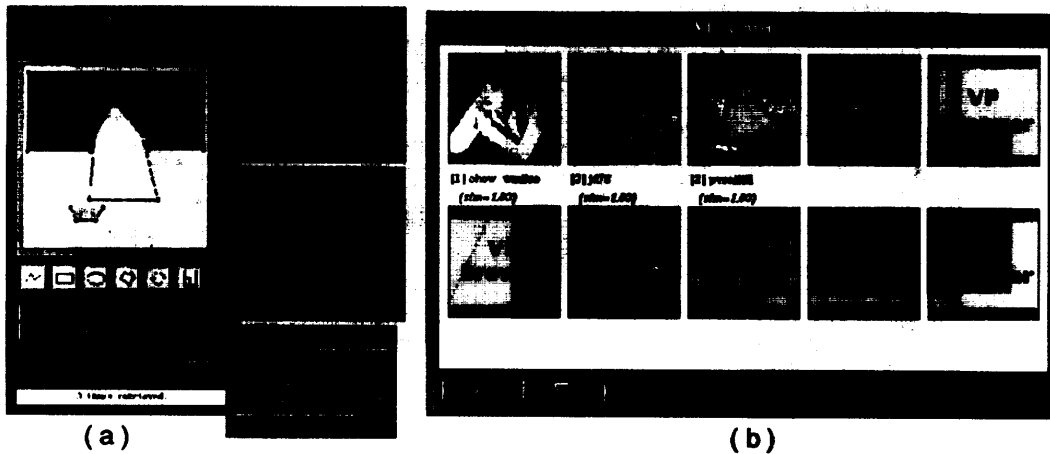


図 14 自然画像検索例(1)

Fig. 14 Retrieval of natural images(1).

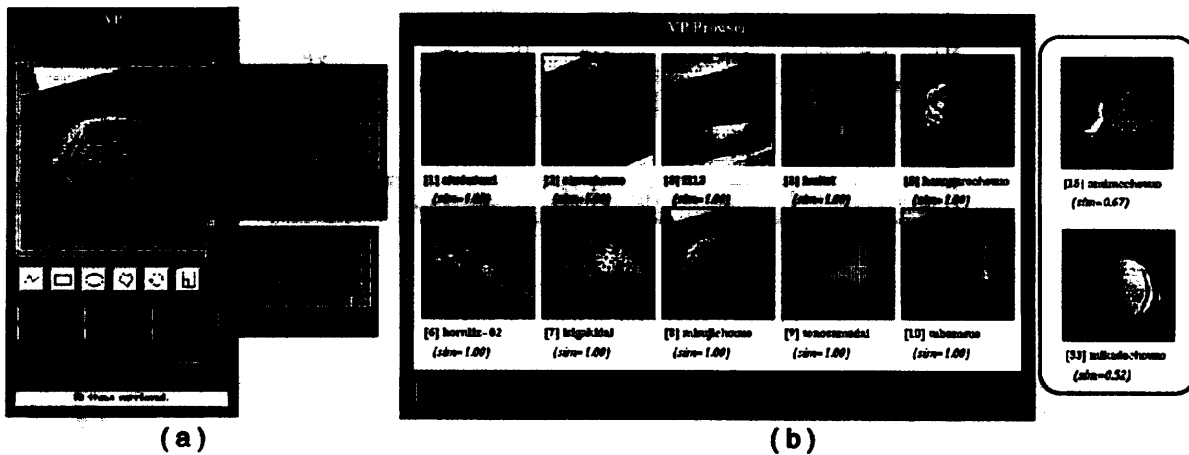


図 15 自然画像検索例(2)

Fig. 15 Retrieval of natural images(2).

属性ウィンドウが帆に相当する。関係は初期設定のままである。この結果、検索数は閾値 0.9 で 8 件であったが、閾値を 0.95 にあげたところ 3 件に絞れた。検索結果は図 14(b) のとおりである。所望の画像が含まれていることが分かる。

次に、実際の画像を起点とした検索例を図 15 に示す。まず、ユーザは別の検索の過程で見つけた魚の画像をインタフェース上にロードする。そして着目する箇所を多角形で指示する。この場合、検索要求は「同じ角度で撮影された魚の画像」である。背景を表す長方形オブジェクトと、中央付近に楕円オブジェクトを描画し、検索条件を設定した。この場合、これらのオブジェクトの色は表示された画像から抽出される。しかし、この場合、魚の色や背景の色には興味がないのでゆらぎを最大にしたが、背景はおそらく暗く魚はそれよりは明るい色であるだろうということで、輝度に

関してはゆらぎを小さくした。何度かの試行錯誤の結果、最終的に 10 件に絞り込んだ。結果は図 15(b) である。この際、余計な結果が 3 件含まれ、2 件が洩れた。

なお、ここでも検索時間はいずれも約 1 秒ほどであった。

### 5.3 評価

以上の検索例により、本システムが、個々のオブジェクトごとに属性と関係に関する検索条件を設定でき、従来とは異なる多様な検索条件を扱えることが確認できる。

次に本研究では適合率と呼出率を測定することで簡単な検索精度の評価を行った。まず、評価に用いる検索例を国旗画像について 15 例、自然画像について 20 例を想定した。この際、画像の特徴に注目した検索要求に限定した。また本システムでは、問合せとブラウジングを何度か繰り返して試行錯誤的に検索を行うの

表1 検索精度の評価  
Table 1 Retrieval effectiveness.

試行回数	1	2	3	4	5
国旗画像検索の呼出率 (%)	61	64	69	73	76
国旗画像検索の適合率 (%)	59	52	50	49	50
自然画像検索の呼出率 (%)	23	37	48	49	53
自然画像検索の適合率 (%)	29	28	30	30	32

で、10回を上限としてブラウジングごとに精度を測定した。また、ブラウジングは基本的に検索数が数件程度まで絞り込んだ時点で行うものとした。もちろん、そこから検索の幅を広げて、さらに追加のデータを検索することは許した。表1に各試行ごとの平均結果を示す。なお、検索者はシステムの操作に習熟していると仮定する。

また、ある検索要求に合致するデータの部分集合を  $R_q$ 、システムの検索結果の集合を  $E_q$ 、 $E_q$  のうち検索要求に合致する集合を  $E'_q$  とすると、適合率は  $|E'_q|/|E_q|$ 、呼出率は  $|E'_q|/|R_q|$  で定義される。

この結果より、試行錯誤的な検索により呼出率が向上していることが確認できる。これは、本システムの検索操作の柔軟性を示すものであり、最初の検索に失敗しても、ユーザが検索条件を操作することで試行錯誤的に所望の画像をさがしあてることが可能であることを示すものである。

この実験より、検索の精度が低下する原因は大きく分けて次の3つであることがいえる。

まず1つめはユーザの検索要求を適切にシステムに伝達できない場合である。現在のところ、本システムで提供しているのは低レベルの特徴のみであり、またユーザの考える特徴とシステムが扱う特徴の間には、どうしてもずれが生じてしまう。

2つめは、マッチング処理による余計な検索結果の混入である。2部グラフマッチングの導入により、マッチング処理が原因となる検索洩れについては改善できたが、本手法では、特徴を量子化したうえに、プリミティブ個々の属性と、その二項関係のみしか扱っていない。よって、実際にはまったく類似していないにもかかわらず高い類似度が観測される場合が起こりうる。

3つめは、原画像処理の問題である。国旗画像のように、比較的単純なパターン画像の場合は、有意なプリミティブを抽出できる確率は高いが、自然画像の場合だと依然困難である。本手法では、冗長な領域を抽出することで有意なプリミティブを拾うことを試みたが、まだまだ不十分である。有意なプリミティブが抽出できなかった場合、これが検索洩れの主たる原因になる。表1中でも、国旗画像検索に比べて自然画像検

索の精度が低いのはこれが原因である。

これらは今後の課題とするところである。

## 6. おわりに

本論文では、ユーザの視点を考慮した、多様な検索要求を柔軟に処理しうる画像検索について述べた。そして、実装したプロトタイプシステムを通じて、本手法が、ユーザに試行錯誤的な検索操作を提供し、従来とは異なった柔軟な検索が可能であることを確認した。また、本手法では、プリミティブの抽出を冗長に行うことで、画像処理技術の困難さを緩和するとともに、そのために増大したプリミティブの検索処理の効率化を図るために、シグネチャファイルを用いた。このような検索メカニズムは、特に一度、目にしたことのある画像を検索する場合や、断片的な情報を基にした画像検索には有効であると考えられる。

今後は、自然画像に対するプリミティブの抽出方法の改善、検索精度の向上および効率化を中心として、さらに検討を加えていく予定である。

## 参考文献

- 1) 加藤俊一, 下垣弘行, 藤村是明: 画像対話型商標・意匠データベース TRADEMARK, 電子情報通信学会論文誌, Vol.J72-D-II, No.4, pp.535-544 (1989).
- 2) Pentland, A., Picard, R.W. and Sclaroff, S.: Photobook: Tools for Content-Based Manipulation of Image Databases, *SPIE*, Vol.2368, pp.37-50 (1994).
- 3) Chang, S.K., Yan, C.W., Dimitroff, D.C. and Arndt, T.: An Intelligent Image Database System, *IEEE Trans. Softw. Eng.*, Vol.14, No.5, pp.681-688 (1988).
- 4) Bimbo, A.D., Campanai, M. and Nesi, P.: A Three-dimensional Iconic Environment for Image Database Querying, *IEEE Trans. Softw. Eng.*, Vol.19, No.10, pp.997-1011 (1993).
- 5) Swain, M.J.: Interactive Indexing into Image Database, *SPIE*, Vol.1908, pp.95-103 (1993).
- 6) 坂本弘章, 高畑 実, 上森 明: 複数のキー画像の部分的な特徴を使用する静止画像検索方式の提案—フレキシブル・モニタージュ検索—, 信学技報, Vol.PRU93-80, pp.31-36 (1993).
- 7) 美濃導彦, 岡崎 洋: 対象物の属性特徴による画像検索法—風景画像中の山を例として—, 情報処理学会論文誌, Vol.32, No.4, pp.513-522 (1991).
- 8) 平田恭二, 原 良憲: ハイパーメディアシステム「雅」におけるメディアナビゲーション方式, *Proc. Advanced Database System Symposium '93*, pp.17-26 (1993).

- 9) Kanehara, F., Satoh, S. and Hamada, T.: A Flexible Image Retrieval Using Explicit Visual Instruction, *Third International Conference on Document Analysis Recognition in Montreal, Canada*, pp.175-178 (1995).
- 10) 金原史和, 佐藤真一, 濱田 喬: 形状分解によるユーザの視点に基づいたシルエット画像検索, *情報処理学会論文誌*, Vol.36, No.12, pp.2800-2810 (1995).
- 11) Yihong, G., Zhang, H., Chuan, H.C. and Sakauchi, M.: An Image Database System with Content Capturing and Fast Image Indexing Abilities, *1994 International Conference on Multimedia Computing and Systems*, pp.121-130 (1994).
- 12) Bach, J.R., Fuller, C., Horowitz, B., Liu, K. and Fe Shu, C.: Managing Multimedia Data, *MISH'95* (1995).
- 13) Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C. and Taubin, C.: The QBIC Project: Querying Images By Content Using Color, Texture, and Shape, *SPIE*, Vol.1908, pp.173-187 (1993).
- 14) 高木幹男, 下田陽久 (編): 画像解析ハンドブック, 東京大学出版会 (1991).
- 15) 笠尾敦司: 分割 K-平均アルゴリズムを用いた大画像の領域分割, *電子情報通信学会論文誌*, Vol.J78-D-II, No.2, pp.221-228 (1995).
- 16) Panjwani, D.K. and Healey, G.: Markov Random Field Models for Unsupervised Segmentation of Textured Color Images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.17, No.10, pp.939-945 (1995).
- 17) Wu, Z. and Leahy, R.: An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.15, No.11, pp.1101-1113 (1993).
- 18) 大田友一, 金出武雄, 坂井利之: 領域分割処理によるカラー画像情報の構造化, *情報処理*, Vol.19, No.12, pp.1130-1136 (1978).
- 19) Kanehara, F., Satoh, S. and Hamada, T.: Shape Decomposition Based on Erosion Model, *Proc. the Workshop on Physics-Based Modeling in Computer Vision in Cambridge, USA*, pp.128-134 (1995).
- 20) Tavakoli, N. and Ray, A.: A New Signature Approach for Retrieval of Documents from Free-text Databases, *Information Processing and Management*, Vol.28, No.2, pp.153-163 (1992).
- 21) Hartsfield, N. and Ringel, G.: グラフ理論入門, サイエンス社 (1992).

(平成 8 年 2 月 23 日受付)

(平成 8 年 9 月 12 日採録)

#### 金原 史和 (正会員)

昭和 43 年生。平成 3 年東京大学工学部電子工学科卒業。平成 8 年同大学院工学系研究科電子工学専攻博士課程修了。画像データベースの研究に従事。



#### 佐藤 真一 (正会員)

昭和 62 年東京大学工学部電子工学科卒業。平成 4 年同大学院情報工学博士課程修了。同年学術情報センター助手。平成 7 年より、米国カーネギーメロン大客員研究員として Informedia 映像デジタルライブラリの研究を行っている。工学博士。図面・画像理解、画像データベース構築、映像データベース構築などの研究に従事。電子情報通信学会会員。



#### 濱田 喬 (正会員)

昭和 16 年生。昭和 39 年東京大学工学部電気工学科卒業。昭和 44 年同大学院工学系研究科電子工学専攻博士課程修了。工学博士。同年同大学生産技術研究所助教授。昭和 48 ~ 49 年カリフォルニア工科大学客員助教授。昭和 61 年学術情報センター教授。昭和 62 年東京大学工学部教授 (併任)。道路交通管制, 計算機言語, 分散処理, オブジェクト指向データベース等の研究に従事。著書「道路交通管制」, 「Road Traffic Control」。電気学会, 電子情報通信学会会員。

