

# 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—

美濃 導彦<sup>††</sup> 岡崎 洋<sup>†††</sup> 坂井 利之<sup>††††</sup>

画像データベースにおいて、各画像に検索のためのインデックスを自動的に付ける方法は、検索方法と関わる重要な問題である。我々は、人間に親和性のある検索方法を実現するためには画像を認識することが必要であるとの立場から、画像を認識し、その属性特徴をキーとして、画像を検索できるシステムを風景画像中の山を対象として作成した。まず、入力された画像に存在する山は、前処理、領域分割処理などの構造化処理を経て認識される。これにより、システムは山の領域を判別できるので、山の属性特徴としての、色、形、大きさ、頂上の位置の4つの属性特徴を計算し、これを画像検索のためのインデックスとする。検索方法は、システム主導型の検索方法である。すなわち、システムが山の属性特徴をメニュー形式で提示し、利用者はそれを対話的に選択していくことにより検索をすすめる。この方法は、他の方法（スケッチによる類似検索法など）よりも柔軟性にかける場合もあるが、記憶が曖昧なときでも画像を検索できるという大きな利点がある。実際にシステムを作成し、デモンストレーションにより評価した結果、対象物の属性特徴を用いた画像検索法の有効性が確認できた。

## 1. はじめに

画像情報の計算機による処理技術の発展に伴い、画像情報を含んだデータベース、すなわち画像データベースの研究が活発になってきた<sup>1)~3)</sup>。

画像データベースは、従来の数値・文字情報を中心としたデータベースの拡張版として位置付けることも可能ではある。例えば、従来のリレーショナルデータベースの1つのフィールドに画像を指すポインタを持つようなものも広い意味では、画像データベースである。この種の画像データベースの代表は文書ファインディングシステムで、光ディスクとともに実用化されている。検索キーは、画像の内容にかかわるものであるが、符号情報であるため、人間が画像を見て付加しなければならない。

画像に限らず、データベースに蓄えるデータは、大きく2つに分けられる。1つは、データの価値がかなり長い間、不変であるデータ、他方は価値が比較的短時間で消滅するようなデータである。どちらの種類のデータを扱うかでデータベースの設計思想が変わってくる。すなわち、前者のデータを扱うデータベースは

検索の効率が重要であり、データ入力時に手間をかけても問題はない。これに反して、後者のデータを扱う場合は、データ入力時の効率が重要であり、検索時に多少時間がかかっても支障はない。

本論文では、入力時の効率を重視する後者の画像データベースについて論じる。その特徴は、画像の入力時にシステムが画像内容に基づくインデックスを自動的に作成できる点である。柔軟な検索を実現するためには、人間の概念と対応のよい検索キーを抽出しなければならないので、画像の入力時に認識・理解処理が必要となる<sup>4)</sup>。

入力された画像はそのまま蓄積される。同時に、含まれている対象物が認識され、その色・形などの属性特徴がインデックスとして計算される。この属性特徴を利用して検索者の要求に柔軟に対応できる画像検索法を実現する。

このような画像検索システムの具体例として、山が存在する風景画像に対し、画像中の山を認識した後、山の属性特徴であるその形状、位置、大きさ、色を利用して検索を行うシステムを作成した。以後、2章で画像検索システムの考え方、特徴を明確にし、3章で画像検索システムの構築と検索処理について述べる。4章では、風景画像中の山を例としたシステムの構築と検索実験および考察を行う。

## 2. 対象物の属性特徴による画像データベースの検索

人間が画像を記憶するときは、画像を理解し、その

† A Retrieval Method of Image Database Using Attributive Features of Objects—In Case of a Mountain in the Scene Image as an Example—by MICHHIKO MINOH (Faculty of Engineering, Kyoto University), HIROSHI OKAZAKI (Information System Research Center, Canon Inc.) and TOSHIYUKI SAKAI (Faculty of Science and Technology, Ryukoku University).

†† 京都大学工学部

††† キヤノン(株)情報システム研究所

†††† 龍谷大学理工学部

内容を記憶する。すなわち、その画像のどこに何かどのような関係であったか、ということ記憶していると考えられる。人間に対して、柔軟な検索方法を提供するためには、このような記憶を自然な形で引き出す検索方法が望まれる。

従来までにも、様々な形で画像データベースに関する研究が行われている<sup>4)</sup>。それらは、大きく2つの流れに分類できる。1つは画像の内容を表現する意味的な検索キーを画像の入力・蓄積過程で人間が付加するタイプのもの<sup>5),6)</sup>、他方は入力蓄積時に画像から特徴量を抽出しておき、検索時にこの特徴量から類似度を計算して検索するものである<sup>7),10),11)</sup>。

ユーザから見た場合、前者は人間が求める検索キーを用いて検索できるので便利であるが、入力かわざらわしいのでデータの価値の高い画像データ（例えば地図や配管図など）の維持管理などに有効である。後者としては、商標・意匠データベース<sup>8)</sup>や民族学資料のデータベース<sup>9)</sup>などが実現されているが、画像の内容に踏み込んだ処理がなされていない。言い換えれば、画像から簡単に抽出できる特徴量を検索キーとしているだけなので、ユーザに対する柔軟な検索ができているとは言えない。名刺画像を対象としたデータベースシステム<sup>12)</sup>は文字の利点を生かして、人間に検索しやすい特徴（文字コード）と画像から抽出できる特徴（特徴パラメータ）間の対応をうまくとっているので、入力も容易で、検索も柔軟なものとなっている。

ここで提案する画像検索システムでは、画像を対象として、人間が求めている検索キーと画像から抽出できる検索キーとの間の対応を中心に考える。そのために、入力された画像から領域を単位とした記述（領域記述）を作成し、認識処理で対象物の認識を行い、対象物単位での記述（対象記述）を作成する。これらの結果から検索時に利用する属性特徴を抽出して蓄積する。属性特徴とは、具体的には認識された各対象物に関する形状・位置・大きさ・色などを意味する。

属性特徴による検索の手法として、計算機主導型の検索方式を提案する。検索システムは、認識された対象物の属性特徴ごとの類似度により、蓄積画像を分類したインデックスをもっている。検索時には分類された各クラスの代表パターンを属性特徴ごとに提示し、検索者は候補クラスを選択を行う。この分類・提示、候補選択を繰り返す検索と候補画像を確認する視察検索を併用した会話型検索により画像を検索する。

上記の画像検索システム全体における処理の流れを

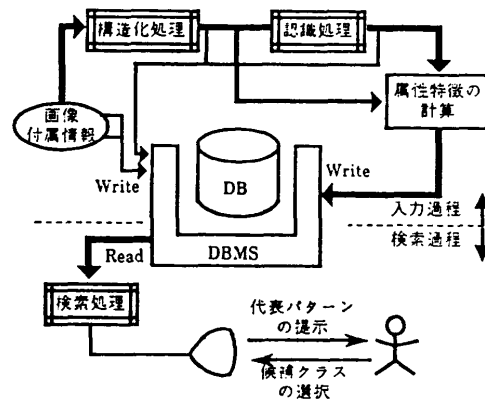


図1 処理の概要

Fig. 1 Overview of the processing.

図1に示す。システムの内部構成では入力過程における構造化処理・認識処理および検索過程における検索処理が中心となるので、次章で詳述する。

### 3. 属性特徴をインデックスとする画像データベースの作成

#### 3.1 画像の構造化

画像データベースに蓄積される画像は多種多様であるので、原画像を忠実に処理しなければならない。同時に、認識処理のための構造化処理は、認識側からの要求に対応できるような拡張性、高速性を持たなければならない。そこで、我々は原画に忠実に領域を分割する初期領域分割処理とその結果生成された基本領域を統合してゆく統合処理の2段階で構造化処理を行う。

初期領域分割のアルゴリズムとしては、領域成長法 (region growing) に基づく手法を適用する。この手法は、「濃度差が閾値 ( $Tr$ ) 以下の隣接画素の統合を繰り返すことにより領域を求める」ものである。本研究では、走査型のラベリングを利用して、これを実行する。

初期領域分割では、その結果が原画に忠実となるように、閾値を低く設定するので、非常に多くの小領域が出現する。特に、多く存在する孤立領域（面積が1の領域）は、あらかじめ、面積最大の隣接領域に統合する。

色情報を利用した領域分割の手法では、対象が風景画像などのようにテクスチャ領域を多く含んでいる場合には、多くの小領域がまとまって現れる。これらの小領域は色の情報を用いて統合することはできないの

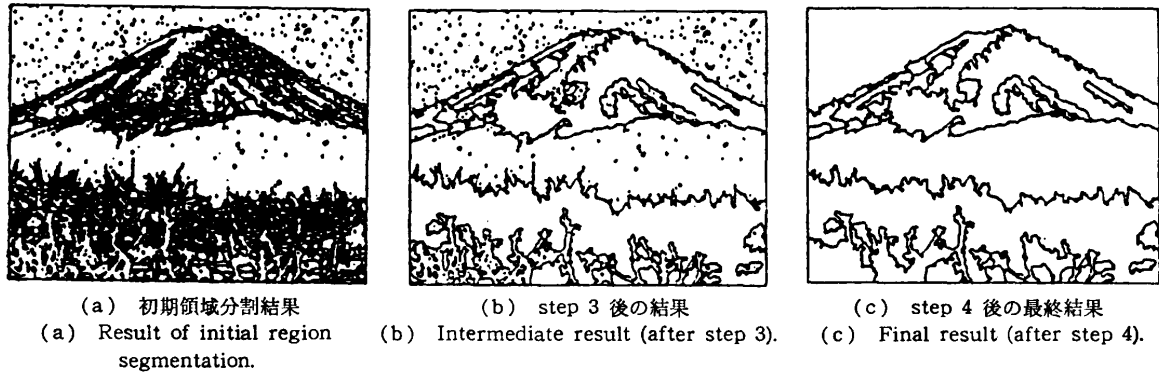


図 2 構造化処理結果  
Fig. 2 Results of region segmentation.

で、領域間の隣接関係を用いて統合する。ここでは、面積が閾値 ( $T_a$ ) 以下の小領域の統合を、初期領域分割で利用した、走査型のラベリングで実現する。

#### 【小領域統合処理】

Step 1: 初期領域分割結果の各領域に対して、面積が閾値  $T_a$  より小さい領域に属する画素に黒、その他の画素に白を与えた二値画像を作成する。

Step 2: その二値画像において黒連結成分のラベリングを行う。

Step 3: 二値画像の黒画素の部分においては、step 2 で行ったラベリングの結果を領域ラベルとする。白画素の部分には、もとの領域ラベルをそのまま保存する。

Step 4: step 2 で得られた領域の中で、面積が閾値  $T_a$  より小さい領域は、雑音、大きい領域間の境界上のエッジ部分、非常に小さい対象物などであるので、隣接する領域に適切な選択条件（面積最大、面積最小、共有する境界線最大、色差最小など）に基づいて統合する。 □

図 2 に以上の処理の流れを示す。(a) が初期領域分割後の領域境界線画像、(b)、(c) がそれぞれ step 3, step 4 で更新された領域分割結果の領域境界線画像である。ただし、図 2 の処理では、残った小領域を面積が最小の領域に統合している。

### 3.2 領域の形状記述

構造化結果は、画素に対応する各配列要素が領域番号を持つラベル画像であり、この形式はファイルとして蓄積するには好都合である。しかし、対象物を領域の形状に基づいて認識する処理においては、各領域の隣接関係などが簡単に抽出できるデータ構造にする必要がある。そこで、図 3 に示すような、領域、境界線、境界線分および境界点の 4 つの階層的な要素からなるデータ構造を利用し、領域間の隣接関係を境界線

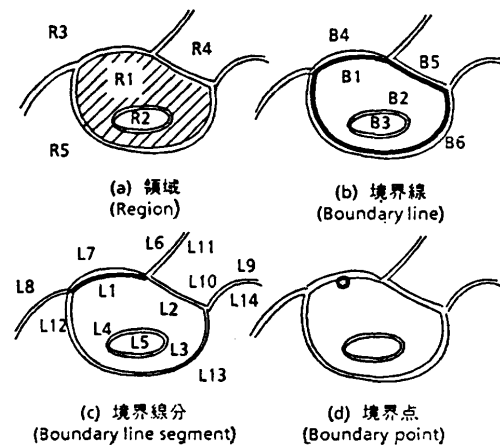


図 3 データ構造  
Fig. 3 Data structure.

分のペアで表現する。

#### (1) 領域 (Region) [図 3 (a) 参照]

ラベル画像で同一の領域番号を持つ連結成分であり、領域は 1 つ以上の境界線で表現できる。孔を持たない領域は 1 つの境界線（外周境界線）で表現でき（図 3 の R 2 は B 3 で表現）、孔を持つ領域は 1+（孔の数）の境界線（外周境界線と孔境界線）で表現できる。（図 3 の R 1 は B 1 と B 2 で表現。）

#### (2) 境界線 (Boundary line) [図 3 (b) 参照]

境界線は 1 つの閉曲線であり、隣接領域の違いによって 1 つ以上の境界線分で表現できる。（図 3 の B 1 は L 1, L 2, L 3 で表現でき、B 2 は L 4 で表現できる。）

#### (3) 境界線分 (Boundary line segment) [図 3 (c) 参照]

境界線分はその線分に対応する境界線上の複数画素で表現できる。境界線上の画素のことを境界点と呼ぶ

ことにする。

(4) 境界点 (Boundary point) [図3(d)参照]

境界点はその値として  $x$ ,  $y$  座標値を持つ。

さらに、領域間の隣接関係は境界線分のペアで表現する。図3の例では (L1, L7), (L2, L10), (L3, L13) などでの隣接関係を表現する。□

認識対象物の形状が複数の直線で表現できる場合を考える。まず、このデータ構造を用いて境界線を直線近似する。直線的な境界線を探す一般的な方法に、境界線上の各点において曲率を計算する方法があるが、この方法は境界線のマイクロなゆれに対処できない。そこで境界線ごとに、区分的直線近似法を適用し、すべての境界線分を直線近似して近似線分を得る。

近似線分をまとめて擬似的に直線と見なせる“擬似直線”を求めるために、ある近似線分から前後に近似線分を探索し、隣接する近似線分のなす角が、あらかじめ与えられた閾値  $THa$  より大きくなるまでまとめてゆく。求められた擬似直線の長さは、近似線分の長さの総和であり、傾きはそれぞれの近似線分の傾きを、それぞれの長さの重みとして加重平均した値とする。

次に、認識対象物の形状モデルを記述する。すなわち、 $N$ 本の擬似直線で形状を定義する場合、それらの擬似直線の長さ、傾きの条件と、連続する擬似直線の端点間の関係の条件を以下のように設定する。ただし、 $N$ 本の擬似直線は抽出したい対象物領域の境界線を右回りにたどった順になっている。

1) [ $i$ 番目の擬似直線に関する定義]

擬似直線の長さ  $length(i)$  と傾き  $slope(i)$  に対して条件を指定する。例えば

$$THL1(i) \leq length(i) \leq THL2(i)$$

$$THS1(i) \leq slope(i) \leq THS2(i)$$

2) [ $i$ 番目と  $i+1$ 番目の擬似直線の間における定義]

なんらかの条件により  $i$ 番目と  $i+1$ 番目の擬似直線の端点間の関係を定義する。例えば、端点間の距離  $dist\_end(i)$  が、

$$THD1(i) \leq dist\_end(i) \leq THD2(i)$$

または、端点間の直線数  $numl\_end(i)$  が、

$$THN1(i) \leq numl\_end(i) \leq THN2(i)$$

など。ただし、 $N$ 番目と1番目の擬似直線の間は条件設定しなくてもよい。□

### 3.3 画像の検索

“検索”とは検索者が検索キーとして示した属性特

徴に対して、それに近いクラスを取り出すことと捉えらると、その属性特徴の類似度に基づく“分類”と同質である。そこで、認識された対象物を属性特徴ごとの類似度によりいくつかのクラスに分類しておく。各クラスがどのような属性特徴をもつかをシステム側から提示し、検索者は自分の求める画像がどのクラスに含まれているかを判断する。このアプローチは検索者の自由度を多少制限するが、初心者や記憶が曖昧なときでも検索できるという特長を持つ。

この検索方式で得られた候補画像に対し、再び同様の検索方式を適用してさらに候補を絞る過程を繰り返すことにより実際に検索を行う。ある対象物に対するこの検索方式の手順は以下ようになる。

[検索手順]

全蓄積画像において、ある対象物の集合を  $R$  とする。そしてその対象物に関する属性特徴の数を  $Na$  とする。

Step 1: 集合  $R$  を属性  $i$  ( $i=1 \sim Na$ ) の類似度により  $Nc$  個のクラスに分類する。それにより  $Na$  種類の  $R$  の分割が求まる。すなわち、すべての  $i$  ( $i=1 \sim Na$ ) に対して

$$C_{i1} \cup C_{i2} \cup C_{i3} \cdots \cup C_{iNc} = R$$

$$C_{im} \cap C_{in} = \emptyset \quad \forall m, n \in \{1, 2, \dots, Nc\}$$

となる  $Na$  個の分割

$$\{C_{i1}, C_{i2}, C_{i3}, \dots, C_{iNc}\}$$

が求まる。

Step 2: 各クラス  $C_{ij}$  の代表パターンを検索システムが、検索者に提示する。

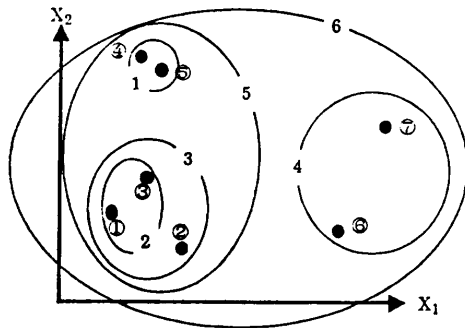
Step 3: 検索者は提示された代表パターンにより、求める画像が含まれていると推測するクラス  $C_{ij}$  を選択する。このとき、各属性特徴に対して選択するクラス数には制限はない。

Step 4: 属性特徴  $i$  に関する分割の各クラス中で選択された  $C_{ij}$  に関して和集合 (join) をとりそれを  $S_i$  とする。すべての  $i$  に対し、 $S_i$  の積集合 (intersection) をとったものを  $R'$  とする。この  $R'$  は  $R$  の部分集合となっている。すなわち、分割  $\{C_{i1}, C_{i2}, C_{i3}, \dots, C_{iNc}\}$  のクラス選択を  $\{1, 2, \dots, Nc\}$  の部分集合  $K_i$  で表すと、

$$S_i = \bigcup_{j \in K_i} C_{ij}, \quad R' = \bigcap_{i \in \{1 \sim Na\}} S_i$$

により  $R' (\subseteq R)$  が求まる。 $R'$  の各要素が属する画像の和集合が検索された候補画像となる。□

複数の対象物が認識されていれば、繰り返し検索の



(a) 与えられたデータ  
(a) Given data.



(b) 樹形図  
(b) Dendrogram.

図 4 階層的クラスタリング  
Fig. 4 Hierarchical clustering.

たびごとに異なった対象物を利用することにより、この検索法の枠組で対応できる。また、対象物による検索で求める画像が得られなかった場合は、領域記述の各領域の属性特徴を利用すればよい。これも蓄積画像中のすべての領域の集合を  $R$  とすることで、この検索法の枠組で対応できる。

この手順における領域  $R$  の分割には図 4 に示すような階層的クラスタリングを利用する。これは、1) 与えられた数のクラスを簡単に得られる、2) 初期クラスタリングにおける階層的構造からより細かいクラスタリングを近似的に求めることができる、などの理由による。

階層的クラスタリングにより得られた階層的構造(樹形図)を利用して、近似的に細分化されたクラスを得る方法を説明する。まず step 1 では集合  $R$  の各属性に対応する  $N_a$  個の階層的構造を得る。各階層的構造ごとに  $N_c$  個のクラスは簡単に求まる。検索者が検索を進めれば step 2 ~ step 4 によって新しい候補集合  $R'$  が求まる。繰り返し検索ではこの集合  $R'$  を再び各属性ごとに  $N_c$  個のクラスに分割するが、これらは集合  $R$  の対応する階層的構造から求める。すなわち、集合

$R$  のクラス数を階層的構造を併合過程とは逆にたどることにより  $N_c$  から順に 1 つずつ増やしていく。すべての要素が  $R'$  に含まれないクラスは除去し、残ったクラス数がちょうど  $N_c$  個になるかどうかを調べる。クラス数は 1 つずつ増えるので必ず  $N_c$  個になるときがある。その  $N_c$  個のクラスから、それぞれ  $R'$  に含まれない要素を除いたものを近似的に  $R'$  の分割に対応する  $N_c$  個のクラスとする。このようにすれば、繰り返しごとに実際にクラスタリングをする必要がないので、高速な応答性を持つ繰り返し検索が実現できる。

#### 4. 風景画像中の山を例とした実験

##### 4.1 システム構成

実現した画像検索システムのハードウェア構成を図 5 に示す。本システムは、大きく分けるとそれぞれのホストマシンを持つ 2 つのシステムが密結合した形になっている。1 つは IBM 3081 K を中心とする、大容量の外部記憶装置と高速な CPU を持つ汎用計算機システムで、高速なソフトウェア実行環境を実現している。

もう 1 つは、MicroVAX II を中心とする画像処理専用システムで、カラーのスクャナ・プリンタなどの入出力装置と、画像処理専用装置 IP 8500 が Micro VAX II に接続されている。IP 8500 にはフレームメモリをモニタするカラーモニタと、フレームメモリ間の高速画像転送が行える画像専用外部記憶装置 RTD が接続されている。

本実験システムでは、ほとんどの処理・蓄積を

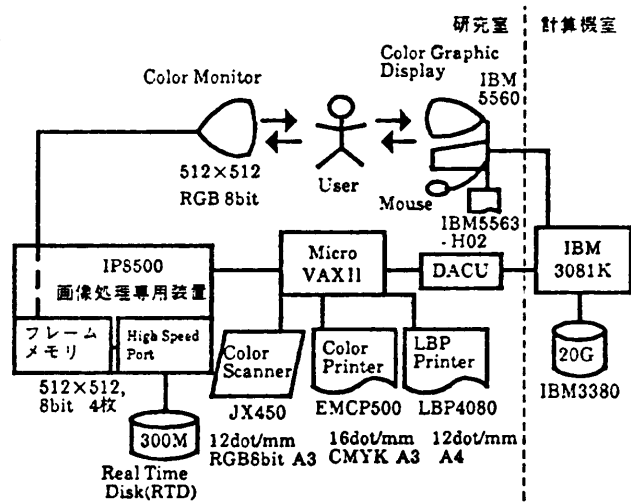


図 5 システム構成  
Fig. 5 System configuration.

IBM 3081 K で行い、画像の入出力および視察検索時に MicroVAXII に接続されている機器を利用する。そこで、この2つの計算機を含むシステム全体を IBM 3081 K 側の端末 (IBM 5560) のみから制御できるようにシステムを構築した。すなわち、MicroVAX II 上では、IBM 3081 K からの処理要求に応じるため、通信管理デーモンが起動され、IBM 3081 K 側の利用者プログラムから送られてくるファイル転送、入出力装置の制御、タスク生成などのコマンドを解釈実行する。このようなシステム構成は、検索処理における視察検索時に有用であり、IBM 3081 K 側の検索処理プログラムから、RTD に蓄積されている原画像を IP 8500 のフレームメモリへ転送するように、Micro VAXII に指示することができる。

#### 4.2 風景画像中の山の認識処理

入力画像は、カラースキャナより解像度 4 dot/mm で入力された  $512 \times 400$  dot、階調 8 bit の風景画像である。ここでは、山の形状モデルとして図 6 に示す簡単なものを利用しているので、入力対象とする風景画像は、山の左右の稜線が明確に示しているもの (富士山のような山を含むもの) でなければならない。この条

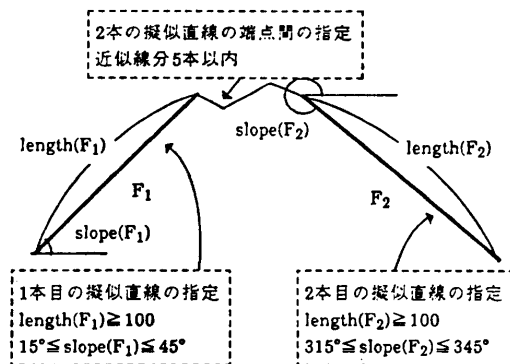


図 6 山の形状指定 (モデル)  
Fig. 6 Model of a mountain.

件を満たしていれば画像内に山が2つ以上存在してもよい。

まず、初期領域分割として、ラベリング型アルゴリズムを適用する。IBM 3081 K において閾値  $Tr=1$  とし、4 隣接での領域分割を行ったときの実行 CPU 時間は、画像に関係なく、約 1.8 秒であった。これは、従来までの領域成長型のアルゴリズム<sup>13)</sup> (CPU 時間が 40~600 秒 (平均 300 秒) 程度) と比較して、非常に高速であった。この手法は原理的には領域成長法と同一である。閾値  $Tr$  を大きくすれば領域内で一箇所でも濃度差の小さい部分があると隣の領域と統合されてしまうので、閾値  $Tr$  は 1 としている。

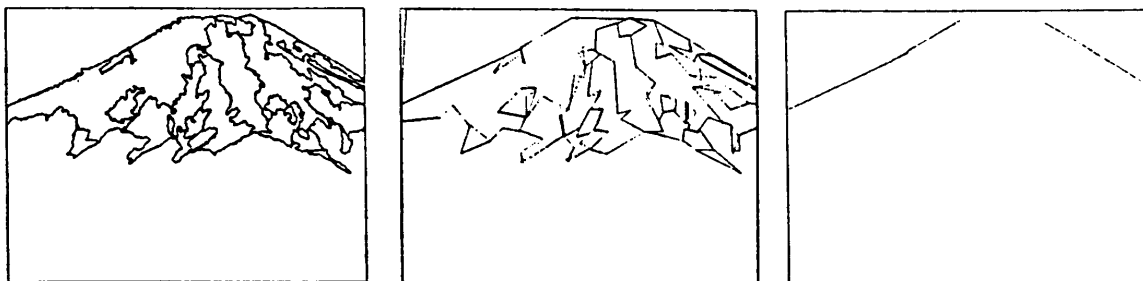
次に、閾値  $Ta$  を 300 として小領域の統合処理を行った。小領域の統合処理では異なるテクスチャの領域が隣接している場合は、1つの領域に統合されてしまうという問題点が明確になったが、それ以外は比較的うまく小領域が統合された。

その後、各領域の境界線を直線近似する。適用した区分近似法における誤差の閾値は、10 dot である。

山の認識処理では、まず、山の形状を2本の擬似直線で記述して、形状抽出を行い山の候補領域を得る。形状指定は図 6 に示すように、1本目の擬似直線が長さ 100 dot 以上、傾き  $15 \sim 45^\circ$ 、2本目の擬似直線が長さ 100 dot 以上、傾き  $315 \sim 345^\circ$ 、そしてその擬似直線間の関係は、端点間の近似線分の数が5本以内という設定である。

形状抽出した例を図 7 に示す。(a) が構造化結果の領域境界線画像、(b) が境界線ごとの直線近似結果、(c) が形状抽出で求めた2本の擬似直線である。求めた2本の擬似直線中に含まれる近似線分が属する領域の集合を山の候補領域とする。その例を図 8 (b) に示す。

形状抽出によって求められた2本の擬似直線は山の



(a) 構造化結果 (a) Result of region segmentation. (b) 直線近似結果 (b) Result of line approximation. (c) 形状抽出結果 (c) Result of contour extraction.

図 7 山の形状抽出  
Fig. 7 Contour extraction of a mountain.

両斜面に対応しているので、これらを含む最小の台形として山の面積が推定できる。山の候補領域の面積を Area 1, 推定による山の面積を Area 2 とする。

$P_{area} = \text{Area 1} / \text{Area 2}$  として,  $P_{area}$  の大きさにより, 山の候補領域を 3 つのタイプに分類する。図 8 にその例を示す。以下タイプ別処理について説明する。

(1)  $P_{area} > TH_{p1}$  (=2.5 程度) [タイプ 1]

推定された面積より山の候補領域の面積がかなり大きいので, 候補領域の分割処理を行う。

(2)  $P_{area} < TH_{p2}$  (=0.8 程度) [タイプ 2]

推定された面積よりも, 山の候補領域の面積が小さいので, 領域の統合を行う。この場合考えられる状況は, 山に対応する領域が多くの領域に過分割されているため, 山の候補領域が山の斜面付近 (2 本の擬似直線付近) の領域しか含んでいない状況である。

(3)  $TH_{p2} \leq P_{area} \leq TH_{p1}$  [タイプ 3]

推定された面積とほぼ一致する山の候補領域を得ているので, 何も処理は行わない。

図 8 で示した形状抽出結果に対して, タイプ別の処理を施した例を図 9 に示す。

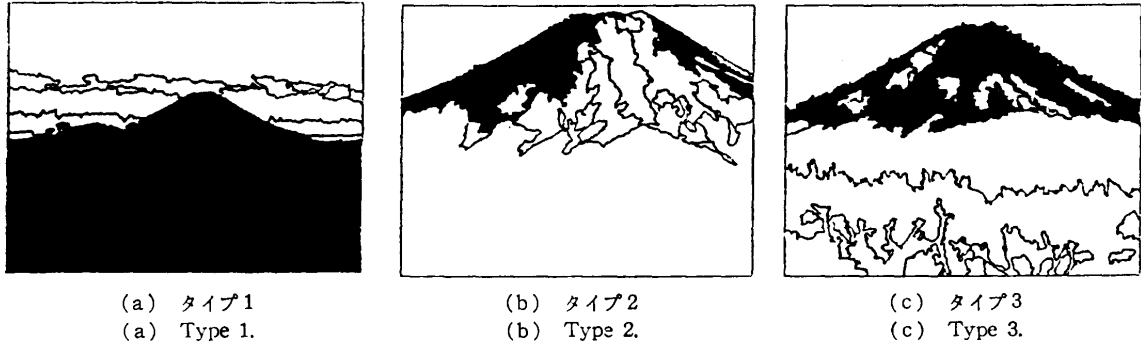


図 8 山の候補領域 (初期状態)  
Fig. 8 Candidates of mountain area (initial state).

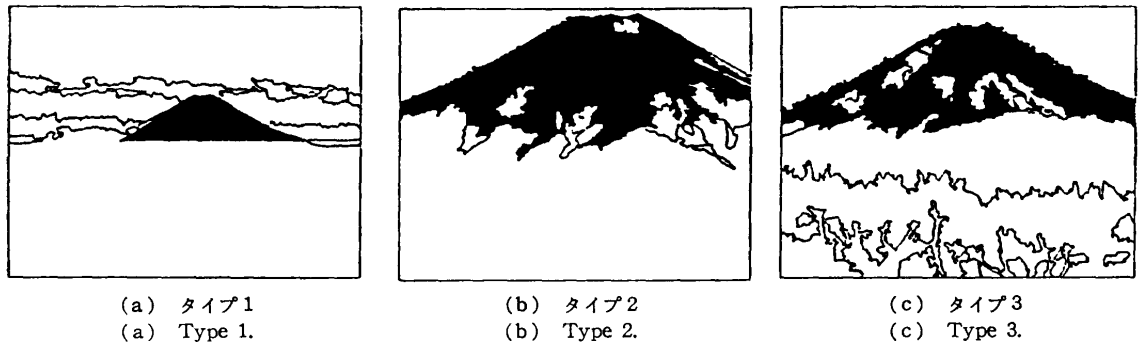


図 9 山の候補領域 (面積によるタイプ別処理結果)  
Fig. 9 Candidates of mountain area (different processings based on the area size).

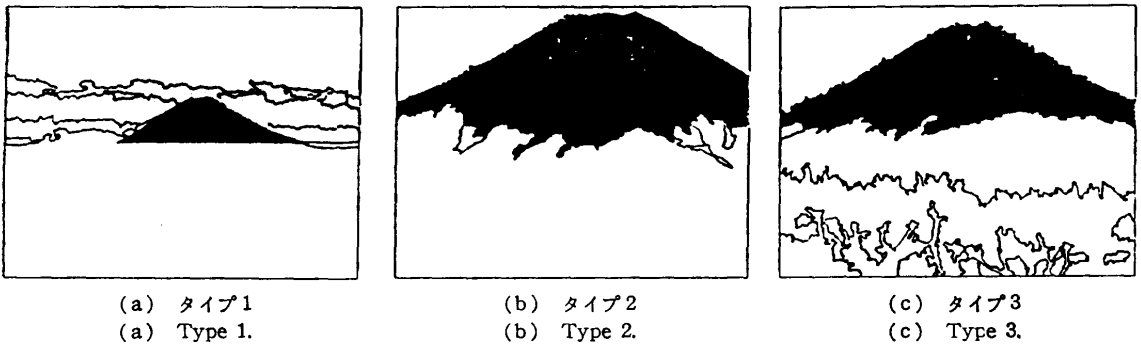


図 10 山の候補領域 (最終結果)  
Fig. 10 Candidates of mountain area (final result).

最後に隣接する小領域を統合して、山となる最終的な領域を求める。図9で示したタイプ別処理結果に対して、隣接小領域統合処理を行った例を図10に示す。これが山として認識された最終的な領域である。

実験では、富士山の写真36枚に対して認識処理を行った。その結果、認識できたものは27枚で、残りの9枚は認識できなかった。失敗例を分類すると、初期領域分割で山の領域がでていないもの4枚、小領域統合処理により山の領域が失われたもの2枚、山に雲がかかっていたため認識が失敗したもの3枚であった。認識された中で富士山以外の山も認識したものが11枚あった。

#### 4.3 対象を山とした検索実験

山の認識結果を用いて実際に検索システムを実現した。現在の認識対象は山だけであるので、属性特徴として形状、位置、大きさ、色を利用した( $N_a=4$ )。各属性特徴は、特徴ベクトルとして表現する。

(1) 形状：まず山の大きさをその外接長方形の横辺の長さが一定となるように正規化する。そして外接長方形の各辺の $Kr$ ( $=0.03$ )倍の距離だけ内側に入った点を通る直線を図11に示すように3本(上辺および左右の辺に対応する)引き、各直線と正規化した山の領域の境界との交点6つ( $P_1 \sim P_6$ )を求める。6つの交点のうち垂直線上にあるものは $y$ 座標、水平直線上にあるものは $x$ 座標を用いて形状を6次元の特徴ベクトルとして表現する。

(2) 位置：画像中の山の位置は、その重心位置より頂上の方が直感的で把握しやすいと考えられるので、図11で定義済みの6つの交点のうち、頂上に対応する2つの交点 $P_3, P_4$ の中点の $xy$ 座標(2次元の特徴ベクトル)で、山の位置を表現する。

(3) 大きさ：山が占める面積の平方根を1次元の特徴ベクトルとし、山の大きさを表現する。

(4) 色：山の色は山の領域に属する画素の $RGB$ 平均値を3次元の特徴ベクトルとして表現する。

クラスタリングによる分類のときに、類似度としてクラスタ重心間のユークリッド距離を用いた。分類し

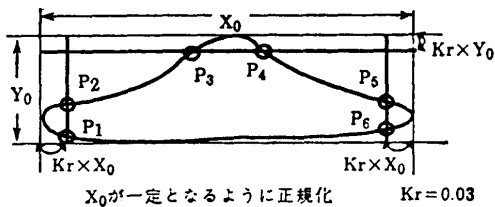


図11 形状特徴の正規化

Fig. 11 Normalization of the shape feature.

た各クラスの代表パターンは対応するクラスタ重心の特徴ベクトルとして、検索者に $N_c=4$ として次のように提示する。

- 1) 形状代表パターンの提示：クラスタ重心の特徴ベクトルの値に対応する6点の座標を求め、その各点を結んだ六角形を山の形状として表示する。
- 2) 位置代表パターンの提示：クラスタ重心の特徴ベクトルから頂上の位置( $xy$ 座標)を求め、その位置を中心とする一定半径の円で表示する。
- 3) 大きさ代表パターンの提示：クラスタ重心の特徴ベクトルの値を2乗することで面積を求める。左右対称でかつ山の平均的な形状を持つ基本パターンを定義しておき、求めた面積となるように相似変換を行い、それを提示する。ただし、相似変換したものが画像の横サイズより大きくなる場合は、山の底辺を下げることによって面積調整を行う。
- 4) 色代表パターンの提示：クラスタ重心の特徴ベクトルから各 $RGB$ 値を求め、擬似カラーで表示する。

以上の代表パターンの例として図12に検索端末の画面のハードコピーを示す。図12における数字はそのクラスに含まれている画像データの数を示している。検索は、曖昧な記憶でも実行できるように、各行ごとに複数個のクラスを選択できる。ある行で、すべてのクラスを選択すれば、その属性は検索には利用していないことになる。

検索実験で用いたデータ数は27で、数が少ないので検索率などを評価してもあまり意味がない。ただ、階層的クラスタリングに基づく段階的な検索を行っているので、データ数が多くなった場合でも検索のある段階では、実験で行ったような環境になる。ここで提案した検索法は、比較的類似した画像が何らかの手段で検索された後に、活用できると考える。

#### 4.4 考 察

画像の蓄積・検索システムは、データ数を増やして検索速度や検索率などで評価するのが普通である。し

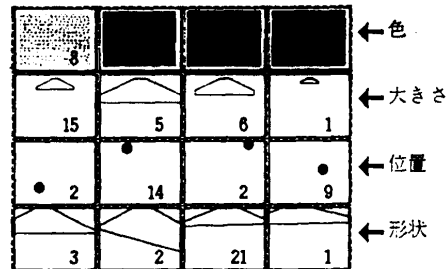


図12 探索端末の初期画面

Fig. 12 Initial display of a retrieval process.



しかし、ここで作成した実験システムが認識できる画像が少ないこと、ハードウェア上の種々制限からデータ量をあまり増やすことができないこと、および検索キーが曖昧で検索率などが明確に定義できないことなどの理由から、このような評価はできなかった。そこで、実験システムで2年間にわたりデモンストレーションを行い、多くの人々にご意見をうかがった。これらをもとにシステムを評価することにした。

クラス属性の数 ( $N_a$ ) は、これ以上は適当なものが考えられないが、クラス数 ( $N_c$ ) は、4である必要はない。ただ、あまりにも多いと人間が選択するときにはクラス間の違いが明確でなくなる。反対に少なすぎると持っている印象に合うクラスが存在しなくなる。今回の実験では、 $N_c=4$  としたが見学者の方々からは大きな異論はなかった。

見学に来られ、実際に検索をしていただいた約50名の方々の約半数は、一瞬見た画像をこのシステムとおしてうまく検索されていた。「検索できた」というのは、複数枚(2~5枚)の候補画像が検索されたときに、その中に求める画像が入っていたという意味である。

検索に利用した属性特徴は、検索者が欲しいと考えるであろう属性特徴の中でシステムが自動時に抽出できるものを選択している。しかし、検索者が直感的に捉える属性特徴とシステムが抽出できる属性特徴はまだ多くの点でずれている。これは、認識部分の能力不足と言うよりは属性特徴の提示方法の方に問題が多かったからだと考えられる。以下、各属性特徴について定性的に評価する。

形状はほとんど情報がなく、検索には有効に働かない。言い替えると、形状は画像によってあまり差がなく、特別な形をした例外的なもののみが初期検索で提示される(図12から27枚中21枚が同じクラスに属していることからわかる)。繰り返し検索を行っていくとクラス間の差が人間がみてはつきりしなくなってくる。これは、山を一定の形状によって認識しているためであり、対象物認識に用いた属性特徴は検索時に利用してもあまり有効でないことが判明した。

大きさは、画像によって比較的異なっているので検索には有効である。しかし、人間が山から受ける大きさの印象と実際にシステムが提示する山の大きさに大きなずれがある場合がある。これは、どこまでを山と考えるかの問題と提示がクラスの重心に対応するパターンであることの問題である。

頂上の位置の情報も比較的ばらつきがあり検索に有用であった。しかし山の位置だけを点で示しているため、直観的にわかり難いという問題点があり、形状、大きさ共に提示する方がよいという意見もあった。

色は色調が異なるものがうまくクラスに分かれていれば、その差でうまく検索できる。しかし、もともと色調が異なる色が混ざって1つのクラスになっていると、提示された色代表パターンからはどのような色が要素として含まれているかを推察し難い。山がいくつかの色の異なる領域からなっている場合も、表示される色はその平均値になるので、その山に対する人間の感覚とは異なってくる。また、色の表示も擬似カラーのため実際の色とは異なるという表示系の問題もある。

検索に用いた属性特徴の中で、どれが有効かの1つの尺度は、提示した4つのクラスのデータ数のエントロピーである。図12に従って計算すると、有効なものから、色、大きさ、頂上の位置、形状の順となる。この結果は、見学者らの評価とほぼ一致している。

この検索方式の利点・特長は次の3点に要約できる。

- 1) 検索者の記憶が曖昧なときも、その検索要求を簡単に与えることができるので、検索者自身の負担が少ない。
- 2) 消去法的な考えで候補クラスの選択を行い、繰り返し検索することで徐々に候補を絞っていけば、検索もれの少ない検索ができる。
- 3) 複数の認識された対象物や構造化結果の領域などによる検索においても本検索方式が利用できる。

この検索方式の最大の問題点は提示するパターンが各クラスの重心に位置するパターンであるという点で、平均をとられているために特に色、形状が重要な検索キーとなりえていない点である。また、階層的クラスリングを採用しているため提示される各クラスに含まれる画像の数に偏りがあることも問題点である。

もちろん、写真を一枚見せて、そのときの記憶をもとに検索してもらおうとすると、山以外のものが印象に残ることが多いので、画像全体の特徴的な属性(たとえば、赤いものがあつた画像など)も検索キーとして付加することも考えなければならない。

全般的には方向としては興味深いはまだ研究の余地が残っているというご意見が多かった。

## 5. おわりに

本論文では認識対象の属性特徴による検索が可能な

風景画像検索システムについて述べた。本システムではまず入力過程において風景画像中の対象物を認識する処理を行う。このとき、認識対象の各種属性特徴を算出・蓄積しておくことで、柔軟に検索者の要求に対応できる検索が可能となる。

本研究ではこの蓄積・検索方式を実際に風景画像中の山を例として実現した。すなわち、山を認識し、その形状・位置・色・大きさを属性特徴として算出、蓄積した。その後、これらの属性特徴ごとの代表パターンの提示と選択の繰り返しと視察検索の併用により実際に検索を行った。この検索実験により本画像検索システムが、検索者に柔軟かつ負担の少ない画像検索法を提供できることを確認した。

今後は、次のような研究を行う予定である。

- 1) 風景画像中の山に限らず他の対象物も認識できる汎用的な認識システムの構築。
- 2) 検索過程で検索者から得た情報を利用して、対象物の再認識が行えるシステムの実現。
- 3) 認識結果を利用した他の検索方式（例示検索、類似検索、部分検索など）の実現。

### 参 考 文 献

- 1) 坂内, 大沢: 画像データベース, 昭晃堂, 東京 (1987).
- 2) Chang, S. K.: Image Information Systems, *Proc. IEEE*, Vol. 73, No. 4, pp. 754-764 (1985).
- 3) 木戸出: 画像データベース, オーム社, 東京 (1986).
- 4) 坂内: 画像検索技術, 電子情報通信学会誌, Vol. 71, No. 9, pp. 911-914 (1988).
- 5) 島田, 江尻: 日本語インタフェイスを有する知識処理型マルチメディア地図情報システム GENTLE, 情報処理学会論文誌, Vol. 27, No. 12, pp. 1162-1173 (1986).
- 6) 高橋, 島, 岸野: 位置関係を利用した画像データベースシステム, 信学技報, Vol. 89, No. 336, pp. 23-28 (1989).
- 7) 栗田, 下垣, 加藤: 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) 加藤, 下垣, 藤村: 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. J 72-D-II, No. 4, pp. 535-544 (1989).
- 9) 佐藤, 橋原, 井岡, 黒川, 洪, 杉田, 久保, 山本: 民族学研究支援のための標本画像検索システム, 情報処理学会論文誌, Vol. 29, No. 12, pp. 1108-1118 (1988).
- 10) 岡田, 荒川, 坂井: 個人の色感覚に適応する織物パターンの蓄積・対話検索システム, 電子情報通信学会論文誌, Vol. J 70-D, No. 12, pp. 2563-

2574 (1987).

- 11) 加藤, 栗田, 坂倉: フルカラー絵画データベース ART MUSEUM—色彩感と略画による画像対話—, 信学技報, PRU 88-139, pp. 31-38 (1988).
- 12) 太田, 森, 坂井: 漢字パターン列の特徴パラメータによる検索—名刺画像における姓名の場合—, 電子情報通信学会論文誌, Vol. J 64-D, No. 11, pp. 997-1004 (1981).
- 13) 画像処理サブルーチンパッケージ SPIDER-II User's Manual, 協同システム開発(株), 東京 (1986).
- 14) 岡崎, 美濃, 坂井: 風景画像の認識とその認識対象属性による画像検索—風景画像中の山を例として—, 信学技報, IE 86-117 (1987).

(平成2年8月6日受付)

(平成3年1月11日採録)



美濃 導彦 (正会員)

昭和53年京都大学工学部情報工学科卒業。昭和58年同大学院博士課程修了。同年、工学部助手(情報工学科)。昭和62年3月より1年間米国UMASS客員研究員。平成元年助教授(工学部附属高度情報開発実験施設)。画像処理, 人工知能, 知的コミュニケーション関係の研究に従事。工学博士。電子情報通信学会, 画像電子学会, ロボット学会各会員。



岡崎 洋 (正会員)

昭和37年生。昭和60年京都大学工学部情報工学科卒業。昭和62年同大学院修士課程修了。同年、キャノン(株)入社。現在、同社情報システム研究所に勤務。画像処理ワークステーションのソフトウェアに関する研究に従事。平成元年第38回全国大会学術奨励賞受賞。



坂井 利之 (正会員)

昭和22年京都大学工学部電気卒業。同年同大学院特別研究生。昭和28年以降、同講師, 助教授を経て, 昭和35年京都大学教授。昭和63年京都大学名誉教授, 龍谷大学教授。平成元年, 龍谷大学理工学部部長。音声, 画像, 自然言語などのパターン情報処理とコンピュータネットワークの研究に従事。工学博士。著書「情報基礎学—通信と処理の基礎工学」など多数。電子情報通信学会, 電気学会, 音響学会など多数の学会会員。