

# 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

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

画像検索における一つの重要な課題は、言葉では表現し難い画像メディア特有な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、ユーザの多様な画像の特徴に関する検索要求に応えることを目的とし、今まであまり検索条件として詳細に取り上げられていない形状および構造情報に焦点をあてて、シルエット画像を対象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、新たに提案する手法によりシルエット画像の形状分解を行い、それらの部分領域の属性および領域間の関係を対象画像の特徴量とする。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。この時、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、本システムにおいて提案する検索処理方法、検索インタフェースおよび要素技術である形状分解手法について論じる。

## Flexible Image Retrieval Based on the Analysis of Shape and Structure

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

We have developed a flexible image content retrieval mechanism based on shape similarity for silhouette images. In our process, a silhouette image is decomposed into primitives by our novel method 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. Queries are performed by the user's explicit instructions about attributes and relations with our developed image retrieval interface. Thus, without the use of *a priori* knowledge, the proposed system is capable of automatically recognizing several features of each part composing the overall shape of the target image; thereby allowing a user to initiate queries representing detailed demands, intentions, and viewpoints in regard to local/global feature of the shape.

### 1. はじめに

近年、大容量記憶装置の発達、高解像度ディスプレイの出現などにより、計算機上で画像データを柔軟に扱えることが可能になってきたと共に画像データベースに対する需要も急速に増している。画像データベースを構築する際に解決すべき重要な問題の一つとして、その検索方法が挙げられる。現在、実用化レベルにある最も一般的な画像検索の方法は、人手により画像データに付与されたキーワードを用いて検索を行う

というものである。しかし、これには二つの大きな問題点がある。一つは大量な画像データに対する人手によるキーワード付与の労力の問題、もう一つは言葉では表現し難い画像メディア特有な検索要求への対応の問題である。前者についてはシーン解析や画像認識によるキーワードの自動抽出に関する研究<sup>1)</sup>などが行われており、後者に関してはスケッチなどによる例示画検索、ナビゲーション検索などの類似画検索の研究<sup>2)</sup>が主流である。将来、柔軟な画像データベースを構築するには、上記のようなキーワードなどの言葉による検索形態と例示画検索のような言葉に依らない検索形態が個々にまたは統合してユーザに提供されるべきであると考えられる。本稿ではこのうち後者に焦点をあてて議論する。

画像の特徴量を基にした検索に関する研究としては、画像内のオブジェクトの位置関係に着目したアイ

<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

コン検索<sup>3)</sup>や、色、形状、テキストチャなどの特徴を検索条件に用いたスケッチやナビゲーションによる検索システム<sup>4),5)</sup>、商標や配線図のような対象を特定したもの<sup>6),7)</sup>がある。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。しかるに、従来の多くの研究例においては、ユーザに許される検索要求は限られていたといえる。例えば、ナビゲーション検索では、選択された例示画の大局的な情報が重視されがちであり<sup>8)</sup>、局所的な特徴に着目した検索はできない。また、スケッチ検索においても、単に描かれた例示画をシステムがそのまま解釈するのではなく、ユーザがどの部分のどのような特徴に着目して描画したかというユーザの意図が反映できれば、より検索の幅はひろがる<sup>9)</sup>。特に、ユーザの漠然とした、または断片的な記憶を手がかりとした検索には望まれることである。

このような検索を実現するには、画像の局所的な特徴を個々に扱えるメカニズムが望まれ、そのためには画像を何らかの方法で構造化することが必要になる。また、上記のようなユーザの検索要求をシステムに伝達する仕組みも必要であり、さらに、それに伴う検索処理方法も重要な問題となる。このような点を意識したと考えられる研究としては、FMR<sup>10)</sup>や Jeffrey<sup>11)</sup>の研究などが挙げられる。FMRは複数の画像中からユーザが着目する領域を指示することによって、モニターのように画像を検索する手法を提案している。ただし、扱う特徴は主に色情報であり画像の構造化については考えられていない。Jeffreyのシステムでは顔画像を対象として、より細かい検索条件を受容できるメカニズムが提供されているが、検索形態は依然テキスト形式であり、スケッチ検索やビジュアルな指示を許容するナビゲーション検索などの例示画検索の形態に対応した仕組みについては論じられていない。また、QBIC<sup>5)</sup>や Yihongの研究<sup>4)</sup>では、検索条件として、位置、色、テキストチャ、形状などを用意しているが、対象の構造情報や細かい形状特徴までは立ち入っていない。

以上の背景を基に、本研究では、今まであまり検索条件として詳細に取り上げられていない形状および構造情報に焦点をあてて、シルエット画像を対象とした上記の課題を考慮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、シルエット画像の形状分解を行い、それらの部分領域の属性およびそれらの間の関係を対象画像の特徴量とする。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。この時、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、従来とは異なる多様で柔軟な例示画検索が行えるようになった。現在は、対象がシルエット画像の物体の輪郭形状に限定されているが、一般の多値画像から抽出された物体の形状に対しても適用可能である。本論文では、提案する検索処理方法、検索インタフェースおよび要素技術である形状分解手法について論じる。

以下、2章で検索処理の概要について述べ、3章で形状分解手法、4章で検索処理方法について提案する。5章ではシステムの検索例および評価を行い、6章をもってまとめとする。

## 2. 検索の概要

図1に本システムの検索処理の概要を示す。まず、オフライン処理において、シルエット画像は提案する形状分解手法により構造化される。分解された各領域(プリミティブ)はその骨格(スケルトン)を中心とした複数の属性特徴が抽出され、またそれらの属性値を基にした各プリミティブ間の2項関係が記述される。この関係を表す関係値も複数用意される。各属性はその属性値に従って後述するプリミティブシグネチャに変換され、それらの連結として属性データシグネチャが構成される。関係についても同様に関係データシグネチャが構成される。すなわち、この二つのデータシグネチャが一枚のシルエット画像のインデックスとなる。一方、オンライン処理におけるユーザからの検索は例示画インタフェースを用いて行う。ユーザは描画された例示画、または検索結果として提示される例示画に対して、各プリミティブに関する属性および関係について、曖昧度または重要度を明示的に指示する。ここではこの曖昧度をゆらぎと呼ぶ。ゆらぎを指示された各プリミティブはデータ画像同様に二つの属性間い合わせシグネチャと関係間い合わせシグネチャに変換される。間い合わせシグネチャはデータシグネチャと異なり、そのゆらぎの大きさによって立てられるビット数が変動する。マッチングは間い合わせシグネチャと各データシグネチャとの論理積演算とシフト演算によって行われ、これにより個々のマッチングに関しては高速化が図れる。次章以降では、各事項について詳細に述べる。

## 3. 形状分解と特徴抽出

### 3.1 形状分解

シルエット画像(図形)の形状分解は、古くから研究

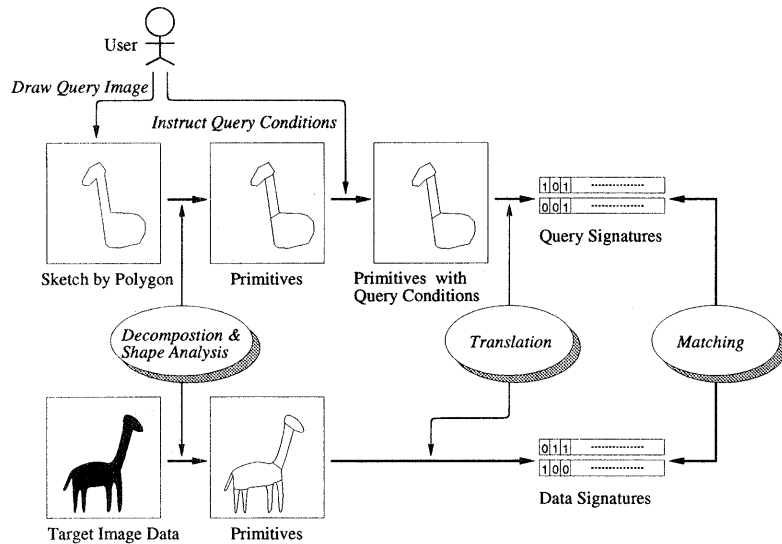


図1 検索の流れ

Fig. 1 Overview of retrieval.

されている画像処理分野の一つのテーマであり<sup>12)</sup>、輪郭線の曲率（凹凸）に着目する方法<sup>13),14)</sup>やモルフォロジー的に処理する方法<sup>15),16)</sup>などの手法が提案されてきた。本システムでは、ある程度高速で、かつ、できるだけ人間の直観に合った結果が得られる手法が望まれる。クラスタリングによる手法は計算時間に問題があり、モルフォロジー的手法は分割結果である各領域の形状が固定的であり拡張性に欠ける。一方、曲率に着目した方法は、雑音に敏感であるという短所はあるが、上記のような問題はない。しかし、従来の手法では完全に凸分解できないものや、単に曲率の変化だけに着目しただけのものが多く、必ずしも直観に合った結果が得られてはいない。

### 3.2 分割アルゴリズム

そこで本システムでは輪郭に着目した凸分解手法を提案する。本手法の基本的な考え方は輪郭の凹部に注目し、対象図形はそこから内部に浸食が生じて分割されるというものである。これは単に輪郭の凹凸だけに着目するのではなく、形状の大きさや厚みなど全体的な特徴が考慮されるのでより直観に即した結果が期待できる。以下にそのアルゴリズムの概要を示す。なお、このアルゴリズムは閉じた2次元の交差のない線図形を対象としている。

- (1) 前処理（輪郭線抽出、雑音除去および多角形近似）。
- (2) 分割操作が施されていないすべての領域（頂点列）に対して以下の操作を行う。
  - (a) 各頂点の曲率を求める。

- (b) 各頂点に対して浸食を表現するベクトル  $\vec{v}$  を求める。
  - (c)  $|\vec{v}|$  の極大値を持つ頂点を起点とした分割線を1本決定する。
  - (d) 領域を2分割する。
- (3) 分割により新たに生じた領域がある場合は2の操作を繰り返す。ない場合は終了。

以下、各手順について説明する。

#### 前処理

対象のシルエット画像から頂点追跡により輪郭頂点列を抽出し、それをフーリエ記述子表現を用いて高周波成分を遮断することにより平滑化を行う。その後復元した等間隔の頂点列を  $\{V_i\}$  とする。なお、分割処理前のこの最初の頂点列を輪郭線、分割後、新たに生成された頂点列を分割線と呼ぶ。

#### 曲率の設定

曲率は  $\overrightarrow{V_{i-1}V_i}$  と  $\overrightarrow{V_iV_{i+1}}$  とのなす角とする（反時計回りを正とする）。また、正の曲率（凸）をもつ  $V_i$  に  $C+$ 、0 または負の曲率（凹）をもつ  $V_i$  には  $C-$  のラベル付けをする。

#### $\vec{v}$ の計算

$V_i \in C-$  の各点における輪郭線内部を浸食する力を (1) 式で定義されるベクトル  $\vec{v}_i$  として表現する (図2 参照)。

$$\vec{v}_i = \phi_i^2 d_i \vec{n}_i, \quad \phi_i = \cos^{-1} \frac{\overrightarrow{V_p V_i} \cdot \overrightarrow{V_i V_n}}{|\overrightarrow{V_p V_i}| |\overrightarrow{V_i V_n}|}$$

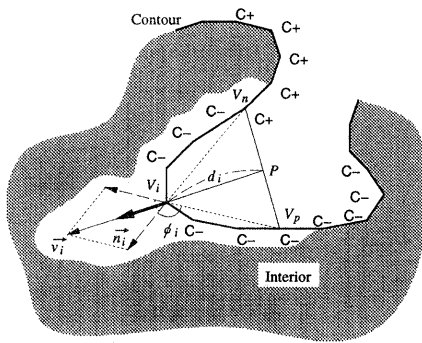


図2 浸食のモデル

Fig. 2 Computing the vector of erosion.

$$d_i = |\overrightarrow{PV}_i|, \vec{n}_i = \frac{\overrightarrow{PV}_i}{|\overrightarrow{PV}_i|} \quad (1)$$

$V_p, V_n$  は、 $V_i$  における角を表現する。ここで一般的注意としては、角は一見局所の特徴のように見えるが、実はそうではなく大域的特徴であるということである。よってここでは  $V_i$  における角をその隣接頂点によって測るのではなく、以下のようにして2頂点  $V_n, V_p$  を求めて大域的特徴としての角  $\angle V_p V_i V_n$  を定める(図3)。すなわち、 $V_i$  から輪郭沿いに頂点を調べていき、 $V_{i+k}, V_{i-k} (k=1, 2, \dots)$  の少なくとも一方が  $C+$  かつ輪郭線上にあるか、 $\cos(\angle V_i V_{i+k} V_{i+k+1}) \geq \epsilon_a$  または  $\cos(\angle V_i V_{i-k} V_{i-k-1}) \geq \epsilon_a$  であるならば、 $V_{i+k}, V_{i-k}$  を  $V_n, V_p$  とする。

$\vec{n}_i$  は輪郭内部に向かう単位ベクトルであり、求められた角  $\angle V_p V_i V_n$  を2等分する。

頂点  $P$  は  $V_p V_n$  と  $\angle V_p V_i V_n$  の2等分線との交点であり、 $\phi_i$  と  $d_i$  は、式からわかるように、 $V_i$  における大域的な曲率とその角の深さを表す。曲率を自乗しているのは、角の深さよりもその曲率の大きさの方が分割への寄与が大きいと推定されるからである。

なお、 $V_i \in C+$  の  $\vec{v}_i$  は  $\vec{0}$  である。

分割線の決定

本手法では、すべての凹部から形状内部への浸食が同時に起こり、速く形状を浸食し終える箇所から順に形状を分割していく方針をとっている。この際、角の深い凹部および曲率の大きい凹部では浸食する速度が速く、また、形状内部で厚みのある部分ほど浸食し終えるのに時間を要すると仮定する。この仮定は、例えば海岸などが削り取られていく自然現象を考慮したものであり、また人間の直観にも沿うことが期待される。

さて、以上の考え方に基づき、形状を最初に分割する分割線を求める。 $\vec{v}_i$  は上記の考え方から凹部にお

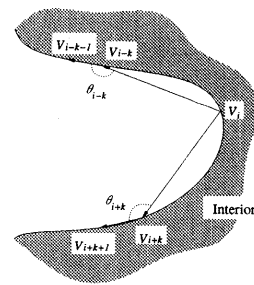


図3 角の決定

Fig. 3 Decision of angle at  $V_i$ .

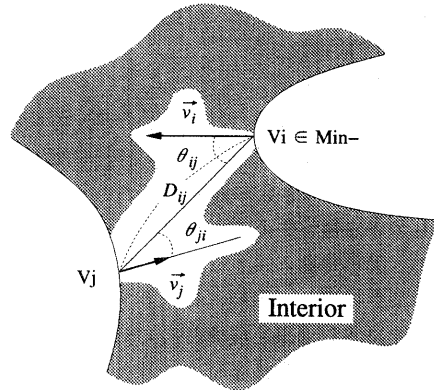


図4 浸食時間の決定

Fig. 4 Computation of erosion time.

ける浸食の速度または力を表現するものであるが、すべての  $\vec{v}_i$  についての振舞いを考慮するのは困難なので、ここでは、その大きさが極大値をとるものを代表として扱う。つまり、 $|v_i| > |v_{i-1}|, |v_i| > |v_{i+1}|$  を満たす  $V_i$  に  $Min-$  とラベル付けをし(負の曲率の極小値に相当)、 $V_i \in Min-$  における  $\vec{v}_i$  のみを考慮する。

次に、各  $V_i \in Min-$  について、 $V_i$  から他の点  $V_j$  への浸食し終えるまでの時間  $time$  を(2)式のように求める(図4参照)。ここで、 $V_i, V_j$  は共に輪郭線上でなければならない。すなわち分割線の端点が他の分割線上にあることを許さない。

$$time(V_i, V_j) = \begin{cases} D_{ij} / (|\vec{v}_i| \cos \theta_{ij} + |\vec{v}_j| \cos \theta_{ji}) & \text{for } (|\theta_{ij}| \leq \frac{\pi}{2}, |\theta_{ji}| \leq \frac{\pi}{2}) \\ D_{ij} / (|\vec{v}_i| \cos \theta_{ij}) & \text{for } (|\theta_{ij}| \leq \frac{\pi}{2}, |\theta_{ji}| > \frac{\pi}{2}) \\ \infty & \text{for } (|\theta_{ij}| > \frac{\pi}{2}) \end{cases} \quad (2)$$

ここで  $D_{ij}$  は  $V_i$  と  $V_j$  の距離であり  $\theta_{ij}$  は  $V_i$  におけるベクトル  $\vec{v}_i$  と  $\vec{V}_i V_j$  のなす角である。分割線は、

$time()$  の最小値を与える  $V_I, V_J$  を結ぶ線分となる。すなわち、最も速く浸食し終える箇所から分割を行うわけである。ただし、このままでは、どのような場合でも分割線は決定されてしまうので、閾値として、 $time()$  の最大値  $t_{max}$  を設けて、 $time(V_I, V_J) > t_{max}$  ならば分割線は設定されず、この領域はこれ以上分割されないとする。

### 分 割

分割線を境界として頂点列を2分割する。線分  $V_I V_J$  の間には各領域のために新たな頂点列を等間隔に補間する。

以上のように大きな凹部や分割線の長さが短くてすむ箇所から分割が生じる。

### 3.3 併合処理

前節で述べた停止条件だけでは過分割や分割不足が生じ、適切な分割結果は得られない。そこで、停止条件である  $t_{max}$  を大き目に設定しておき、一旦過分割を起こしてから併合する方針をとる。併合処理の手順は以下のとおりである。すなわち、まずすべての分割線をその長さに関して降順にソートする。次に、ソートされた順に各分割線を境界として持つ二つの領域について、併合の可否を検査して、可なら分割線に関して併合する。

併合の可否は以下のように決定する(図5)。まず対象となる二つの領域を併合した図形の凸包を求め、その凸包と分割線の延長との交点を端点とする線分の長さを  $l$  とする。分割線の長さを  $l_b$  とした時、 $l_b/l$  がある一定閾値  $\epsilon_m$  よりも大きい場合は併合可とする。これは、図形の輪郭の凹部は任意の二つの凸多角形が合わさって生じるという事実に基づいている。

この併合処理では大きな領域はほぼ適切に併合されるが、微小領域は併合されない場合がある。多くの場合、これは緩やかに曲率が変化するリボン状の形状に起こりがちである。ここでは、このような微小領域は、ヒューリスティックな知識をもとに隣接領域に併合する。

### 3.4 スケルトン抽出

以上の処理により対象シルエットは凸分解される。本検索では、各領域の特徴をその骨格(スケルトン)を主体として記述しており、スケルトンは、領域の輪郭線上の頂点列の第一主成分(図6(a))として求められる。ただし、対象シルエットを凸領域だけで記述するのは不十分である。例えば、関節で曲がっている動物の脚は二つの領域に分割してもかまわないが、それが動物の尾である場合は、むしろ分割するのは好ましくない。この判断にはあらかじめ対象に関する知識

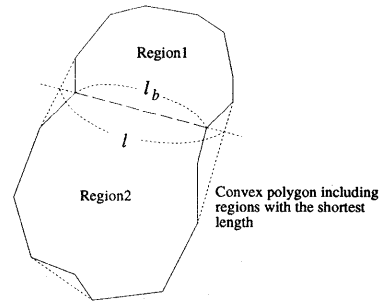
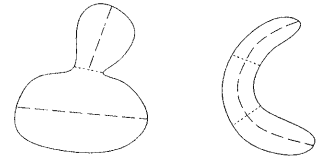


図5 併合処理  
Fig. 5 Merge processing.



(a) Skeletons computed by principal analysis (b) Skeletons computed by cross sections

図6 スケルトン抽出

Fig. 6 Extraction of skeletons.

が必要であるが、ここでは特に対象の知識を用いることはせず、特徴となりえる可能性のある特徴はすべて扱うことにする。すなわち、それが尾であるか脚であるかにかかわらず、分割したものと分割しないもの両方を扱う。

よって、次にすべき処理は今までの純凸分解から準凸分解への併合処理である。ここでは主にリボン状の形状領域を準凸多角形と定義する。併合処理は、リボン状の領域は分割線の一つまたは二つ持つ領域が連なっていることが多いというヒューリスティックな知識に基づいて行う。まず、分割線の一つまたは二つ有する隣接する二つの領域を探し、それらを分割する分割線の両端点を調べる。少なくとも一方の端点における元シルエットの輪郭線の曲率が一定値以上ならば併合する。つまり、その分割線が大きな二つの凹部で生じたのでなければ併合するわけである。

準凸多角形のスケルトンは主成分分析では求められないが、分割線が二つ以下の領域は、断面の始点終点が容易に決定できるので、併合時の断面の中点を結んだものとして求める(図6(b))。

### 3.5 分解例

図7に、512×512の大きさの画像データに対する本手法の分解例を示す。パラメータはフーリエ記述子の係数を  $\pm 50$  とし、復元時の参照点を300点、停止条件の  $t_{max}$  を200、 $\epsilon_a$  を  $-0.8$ 、 $\epsilon_m$  を  $0.8$  と設定した。例からわかるように、ほぼ直観に合った分解結果

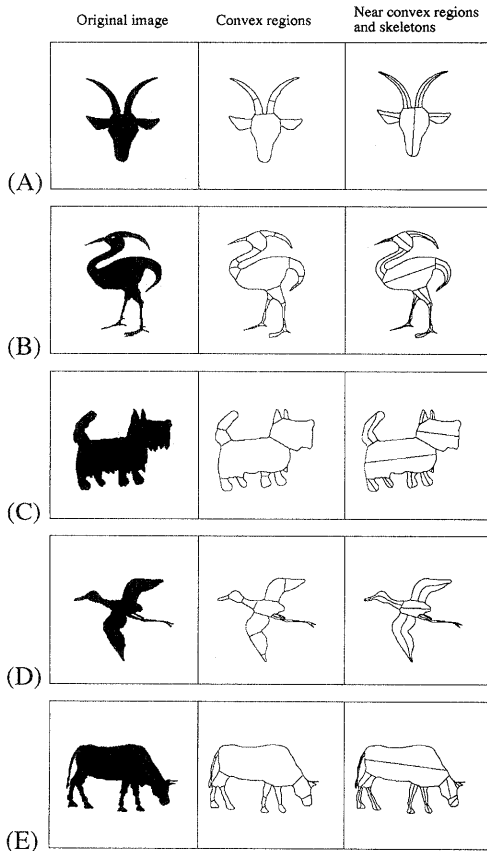


図7 分解例

Fig. 7 Examples of the decomposition results.

が得られている。また、処理時間であるが、この条件のもとで、Sun SparcStation20, SunOS 4.1.4 上で平均約 3.24 秒であり、高速に処理できている。このアルゴリズムは  $Min-$  の頂点の数を  $m$ 、輪郭線上の頂点数を  $n$  とすると  $O(mn)$  の計算量になる。

なお、現在は、輪郭線のみを対象としているので、(B) の目にあたるような“穴”は考慮していない。また、参照点を増やせばより好ましい結果が得られるが、図7のように参照点が 300 点程度でも良好な結果が得られる。

#### 4. 検索処理

##### 4.1 シグネチャ表現

シルエット画像データを、分割により得られた各プリミティブを基に検索用の表現に変換する。ここではシグネチャファイル<sup>17)</sup>を用い、対象データを連続したビット列であるシグネチャ構造に変換する。

シグネチャファイルは、本来全文データベースの索引付けなどに用いられるものであり、一つの単語に有

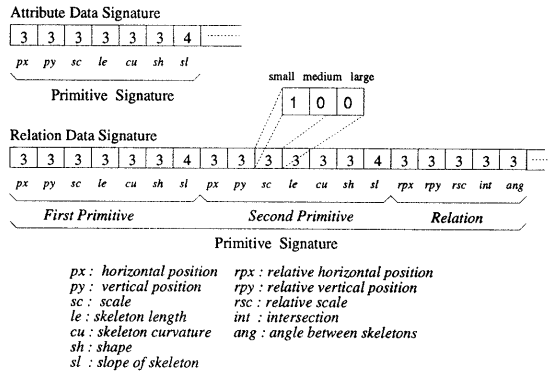


図8 シグネチャの形式  
Fig. 8 Format of signature.

限長のビットパターンを割り当て、ドキュメントをそのビットパターンの集合として表現するものである。シグネチャとはこのビットパターンまたはそれらの集合を指す。シグネチャファイルの代表的な形態として superimposed coding と word signature method がある<sup>18)</sup>。superimposed coding は、各単語に該当するビットパターンの論理和をドキュメントのシグネチャとするものであり、word signature method は単純に繋げたものをドキュメントのシグネチャとする。本研究においては、単語とドキュメントに相当するものは領域図形(プリミティブ)と画像データである。ここでは、このプリミティブの出現頻度の情報も重要となるので、頻度情報がそのままでは反映されない superimposed coding ではなく、word signature method の形態を応用した構成をとる。

各領域はいくつかの属性値(スカラー量)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。すなわち、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る。各属性は基本的にその値がすべて連続値をとるように設定され、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、二つの領域間の関係についてもその関係を表す関係値(スカラー量)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二つの領域の属性のプリミティブシグネチャを先頭に付加する。このように対象画像は基本的に一つずつの属性と関係のデータシグネチャにインデックス化される(図8)。例えば、現在用意されている属性の一つである scale 属性は、small, middle, large に相当させて3ビット用意されており、それぞれに境界値が設定されている。各領域の大きさを表す値は、この境界値に従って量子化

されて、例えば“100”(smallに相当)というビット列に変換される。他の属性も同様であり、各属性を表現するビット列が連なってプリミティブシグネチャが構成される。

このようにシグネチャを用いる理由は、マッチングをシフト演算と論理積演算で行うために、ユークリッド距離を測ったりグラフマッチングなどを行うより単純かつ高速であること、それにより局所的な類似度を測定するのに都合がよいからである。

以下、用いる属性および関係について記す。

[属性]

水平方向位置: 中心x座標. 3ビット (左, 中, 右).

鉛直方向位置: 中心y座標. 3ビット (下, 中, 上).

大きさ: 面積の平方. 3ビット (小, 中, 大).

スケルトン長: スケルトンの長さ. 3ビット (短, 中, 長).

形状: スケルトンの長さに対する領域の平均幅の比. 3ビット (細長い, 普通, ほぼ円).

スケルトンの曲率: スケルトンの曲率変化の絶対値の合計. 3ビット (直線, 湾曲, 極度に湾曲).

傾き: 軸の傾き. 4ビット (水平, 鉛直, 土斜め).

以上計7属性22ビットである。形状属性の値が大きいもの、すなわち円形に近いものは傾き属性は設定しない。ここで軸はプリミティブの輪郭頂点列の第一主成分であり、純凸領域の場合はスケルトンと軸は同一であるが、準凸領域の場合は異なることに注意されたい。

[関係]

スケルトンの長いプリミティブを基準プリミティブ(First Primitive)として、関係を結ぶ。

相対水平方向位置: x座標の位置関係. 3ビット (左, 中, 右)

相対鉛直方向位置: y座標の位置関係. 3ビット (上, 中, 下)

相対的な大きさ: スケルトンの長さの比. 3ビット (小, 同, 大)

交差位置: スケルトンの交差位置. 3ビット (中点寄り, 端点寄り, それ以外)

角度: 軸のなす角. 3ビット (平行, 斜め, 直角)

関係シグネチャには二つのプリミティブの属性を表すプリミティブシグネチャを付加するので、計5関係、59ビットになる。

相対水平方向位置関係は、基準プリミティブの占めるx方向の幅ともう一方のプリミティブの中心のx座標との位置関係を示す。位置関係が水平と鉛直方向に分けられているのは、本システムで扱う関係が

スカラ量であるためである。交差位置関係は二つのプリミティブの近接する点の、基準プリミティブのスケルトンに対する位置を表す。交差位置関係と角度関係は軸の傾き属性が設定されていなければならない。基準プリミティブがあいまいなもの、つまりスケルトンの長さが近いプリミティブの組合せは、二通りの関係を冗長に記述する。また、ここですべての領域間の関係を求めるわけにはいかないで、二つの領域(多角形)間の最短距離(二つの領域と交わる最小の長さの線分の長さ)が一定以内の距離(20)にあるもののみを扱う。

図7(A)の例では凸領域、準凸領域含めて、プリミティブが9、関係が8、属性データシグネチャが198ビット、関係データシグネチャが472ビットである。

4.2 マッチング処理

4.2.1 問い合わせシグネチャ

マッチングは個々のデータシグネチャと後述する検索インタフェースからユーザによって入力・指示および解析された問い合わせシグネチャとの間で行われる。データシグネチャは既に述べたように、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、問い合わせシグネチャでは、これに加えてゆらぎの概念が入り込む。ユーザが目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)。

4.2.2 マッチングアルゴリズム

マッチングは図10のように、極めて単純であり、

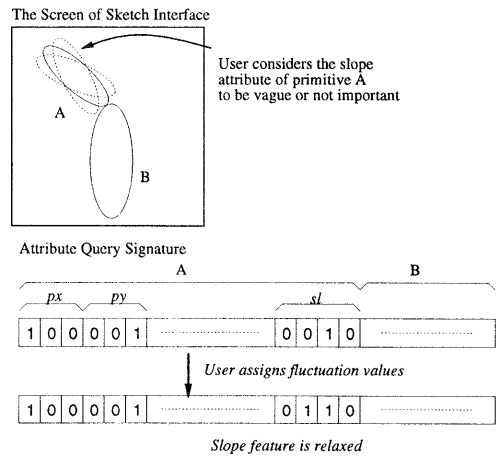


図9 ゆらぎを含む問い合わせシグネチャ  
Fig. 9 Query signature with the fluctuation.

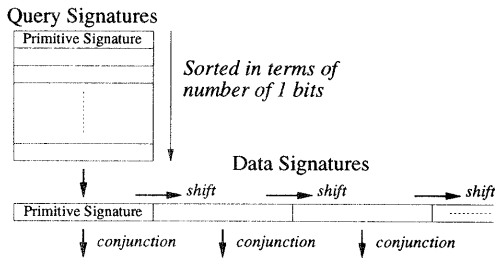


図 10 マッチング処理  
Fig. 10 Matching process.

個々の問い合わせシグネチャとデータシグネチャとのシフト演算と論理積演算で行われる。まず、問い合わせシグネチャは、そのプリミティブ単位にビット数に関して昇順にソートされ、ビット数の少ないプリミティブシグネチャから比較される。これはマッチングの精度を上げるためである。もし、データシグネチャ内のいずれかのプリミティブシグネチャとの論理積の結果がそのプリミティブシグネチャと一致すれば、類似度にカウントされ、一度一致したプリミティブシグネチャは 0 にリセットされ、重複して比較されないようにする。今、問い合わせシグネチャの属性および関係のワードシグネチャの数を  $N_a, N_r$  とし、一致したプリミティブシグネチャの数を  $C_a, C_r$  とすると、類似度は (3) 式で定義される。

$$\text{Similarity} = \frac{C_a + C_r}{N_a + N_r} \quad (3)$$

### 4.3 検索インタフェース

ユーザは検索インタフェースを用いてスケッチ画を作成してもよいし、ある画像データを起点として検索を行ってもよい (図 11)。描かれたスケッチ画はデータに施されたのと同様の処理により形状分解することが可能である。分割により再表示されたプリミティブ群に対して、ユーザは着目する特徴、曖昧な特徴などを次のようにして指示する。

ゆらぎの指示方法は、各属性・関係について直接ゆらぎの値を図 11 に示されるようなスクロールバーによって与える。バーにおいて最も左にカーソルを持っていけば、それはゆらぎ値が 0 であることを示す。ここで注意されたいのは、属性または関係の値を指示するのではないということである。例えばある領域の slope 属性のゆらぎ値を 0 にするとは、その領域の軸が水平であることを示すわけではなく、領域の slope 属性の値にかかわらず、キャンバス上で描かれている状態がユーザにとって明確であるということを示す。検索結果である画像データに対する処理も同様であり、ナビゲーションのように用いることも可能である。

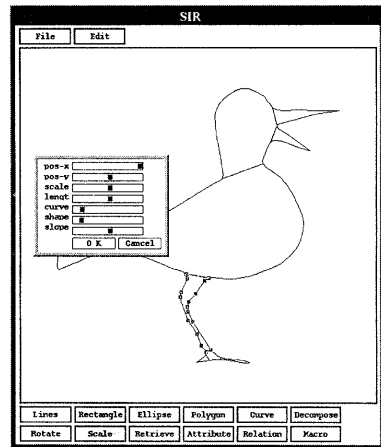


図 11 例示画像検索インタフェース  
Fig. 11 Query interface by visual example.

なお、このスクロールバーを含むポップアップウィンドウは属性および関係について個々に用意されており、図 11 は属性についての例である。また、複数の領域に対する一括指示も個々に対する指示も可能である。

## 5. 検索例と評価

本章では、いくつかの検索例を用いて本システムの評価を行う。本システムは Sun SparcStation 20 上に約 10,000 行の C++ で実装されている。現在扱うデータは約 100 枚の 512×512 の動物を中心としたシルエット画像である。

まず、図 12 は、楕円と折れ線の描画によるスケッチ検索の例である。ここでは、ユーザが一度目にしたことのある画像を検索対象にしていると仮定する。図 12 (a) では描かれた折れ線に対して属性を指示している。ここでは鉛直方向の位置とスケルトンの長さおよび曲率、形状が明確または重要であることが入力されている。楕円の方はデフォルトのゆらぎが与えられている (0.5)。図 12 (b) では同様に位置関係の制約が強められている。この検索結果のうち最も高い類似度を出した画像を図 12 (c) に示す。ほぼ検索要求は満たされている。次に、図 12 (d) のように鉛直方向の位置属性が重要でないと指示された場合には、図 12 (e) がさらに検索結果として含まれる。

図 13 は任意の多角形によるスケッチ検索例である。図 13 (a) のように描画された図は形状分解されて、図 13 (b) のようになる。ここでも同様に着目する特徴が指示される。すなわち図 13 (b) では、位置属性以外の属性のゆらぎは小さく指示されている。関係は



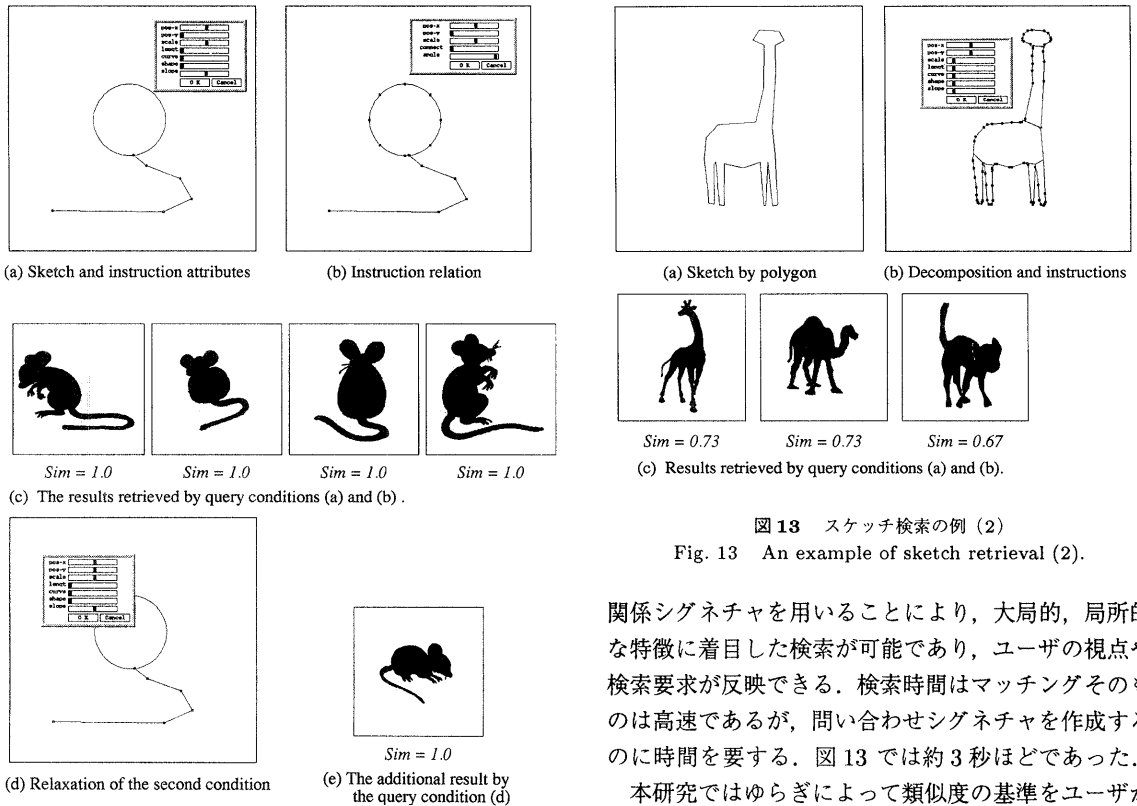


図 12 スケッチ検索の例 (1)

Fig. 12 An example of sketch retrieval (1).

デフォルトのままである。上位検索結果 3 枚の画像を載せておく。図 12 の検索がねずみの尾に着目した局所的な特徴による検索であるのに対し、図 13 は大局的な類似度を測っているといえる。

図 14, 図 15 は提示された画像を起点としたナビゲーション検索の例である。図 14 では図 13 同様に全体の形状特徴の類似度が測られ、図 15 では犬の胴体と脚だけに着目した検索が行われている。図 14 の結果は (j) 以外のらくだはほぼ良好に検索できていることがわかる。(c)(f)(i) のような全く余計なものが含まれるのは、これはシグネチャを用いているために生じる false drop と、2 項関係のみを扱っていることに起因する。

図 15 については脚が短く胴体が高い動物または鳥が検索されている。(b)(c)(g)(h) は適切な結果であるが、その他の三つは直観と合わない。特に (f) に関しては、脚の膝の関節を境に凸分割されているので、それが反応している。

これらの検索例を見てわかるように、本システムでは、プリミティブ分解およびゆらぎを混じえた属性・

図 13 スケッチ検索の例 (2)

Fig. 13 An example of sketch retrieval (2).

関係シグネチャを用いることにより、大局的、局所的な特徴に着目した検索が可能であり、ユーザの視点や検索要求が反映できる。検索時間はマッチングそのものは高速であるが、問い合わせシグネチャを作成するのに時間を要する。図 13 では約 3 秒ほどであった。

本研究ではゆらぎによって類似度の基準をユーザが調整できるので、類似度の値そのものはそれほど重要な意味を持たない。よって、検索条件を緩めていっても類似度の値が適切に増加しない場合が検索洩れの状態に相当することになる。この意味での検索洩れが生じるのは、主に原画像からのプリミティブ抽出が不適切である場合である。図 14 (j) は 1 本の脚が隠れており、さらに尾がないということで、(b)(d)(g)(h) よりも類似度が低くなっているが、それ以外に後ろの脚に関して良好な分割が得られなかったことが主な類似度の低下の原因の一つである。

しかし、今回のシルエット画像を対象とした場合では、この検索洩れよりは不要な検索結果の混入の方が全体の検索精度に与える影響が大きいことを確認している。この主たる原因は、属性および関係についてのシグネチャファイルを用いたことによって生じた false drop にある。false drop とは実際のデータは一致または類似しないにもかかわらず、シグネチャ上では一致してしまう状態のことをいう。よく言われる false drop は、superimposed coding の場合のものである。すなわち、複数の単語に相当するビットパターンの論理和をとって得られたドキュメントのシグネチャが、別の単語の組合せによって得られたものと同一である場合が起こりうるわけである。本手法では、word

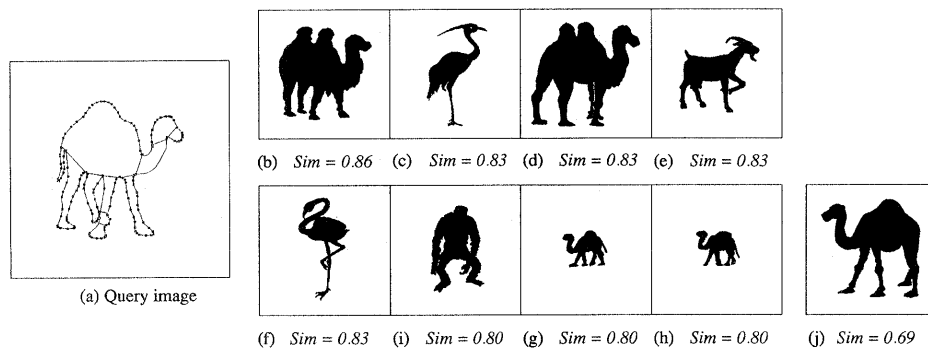


図 14 ナビゲーション検索の例 (1)

Fig. 14 An example of navigation retrieval (1).

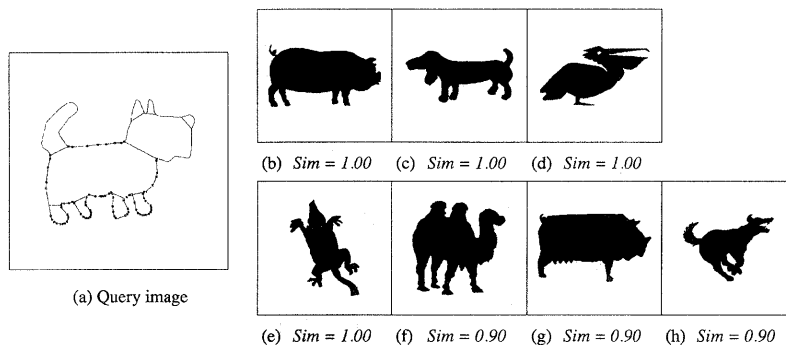


図 15 ナビゲーション検索の例 (2)

Fig. 15 An example of navigation retrieval (2).

signature method を応用した構成をとっているために論理和から生じる false drop の恐れはないが、プリミティブの属性や関係の連続値を量子化してビット列に変換しているために、あるカテゴリのプリミティブが一つのシグネチャで表現される。また属性および関係ともに局所的であるため、実際のデータ上で類似していなくとも、シグネチャ上では類似することがありうる。これらも一種の false drop である。

これは今後改善されるべき課題である。

## 6. おわりに

本論文では、ユーザの多様な検索要求を柔軟に処理しうるシルエット画像を対象にした検索システムについて述べた。また、その要素技術として形状分解手法およびシグネチャによる検索処理と、それらを基にした例示画検索による画像検索を提案した。また、本検索手法により、ユーザが検索要求をグラフィカルかつ明示的にシステムに伝達できることを示し、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その有効性を示した。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。

本検索手法の今後の課題としては領域分割による構造化の精度の向上と、検索の効率化および false drop の抑制が挙げられる。シグネチャの個々のマッチングは高速であるが、すべてのデータとマッチング処理を行うので、この点の効率化が必要であるし、前節で述べたように false drop を抑える処理も必要である。

現在は、シルエット画像のみを扱っているが、これに濃淡画像の領域分割技術や属性として色情報を加えることにより、より幅広い対象を扱いうことが期待され、今後の拡張の予定とするところである。

## 参考文献

- 1) 坂内正夫, 佐藤真一: 画像データベースにおけるモデル形成, 電子情報通信学会論文誌, Vol.J74-D-I, No.8, pp.455-466 (1991).
- 2) 加藤俊一, 栗田多喜夫: 画像の内容検索, 情報処理, Vol.4, No.5, pp.466-477 (1992).

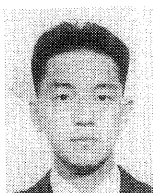
- 3) Chang, S.K., Yan, C.W., Dimitroff, D.C. and Arndt, T.: An Intelligent Image Database System, *IEEE Trans. Softw. Eng.*, Vol.14, No.5, pp.681-688 (1988).
- 4) Yihong, G., Zhang, H., Chuan, H.C. and Sakauchi, M.: An Image Database System with Content Capturing and Fast Image Indexing Abilities, *1994 International Conference on Multimedia Computing and Systems*, pp.121-130 (1994).
- 5) 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).
- 6) 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).
- 7) Wakimoto, K., Shima, M., Tanaka, S. and Maeda, A.: Content-based Retrieval Applied to Drawing Image Database, *SPIE*, Vol.1908, pp.74-84 (1993).
- 8) 尾田政臣: 人間のイメージ形成過程の特性を利用した画像検索システム, 情報処理学会論文誌, Vol.35, No.7, pp.1449-1456 (1994).
- 9) 金原史和, 佐藤真一, 濱田 喬: 図形間の幾何的および概念的関係を用いた作図支援, 情報処理学会論文誌, Vol.35, No.5, pp.897-907 (1994).
- 10) 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・モンタージュ検索—, 信学技報, Vol.PRU93-80, pp.31-36 (1993).
- 11) Jeffrey, R. and Santanu, P.: A Visual Information Management System for the Interactive Retrieval of Faces, *IEEE Transaction on Knowledge and Data Engineering*, Vol.5, No.4, pp.619-628 (1993).
- 12) Shapiro, L.G.: Recent Progress in Shape Decomposition and Analysis, *Progress in Pattern Recognition 2*, pp.113-123 (1985).
- 13) Rom, H. and Medioni, G.: Hierarchical Decomposition and Axial Shape Description, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.15, No.10, pp.973-981 (1993).
- 14) Bjorklund, C.M. and Pavlidis, T.: Gloval Shape Analysis by k-Syntactic Similarity, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.3, No.2, pp.144-155 (1981).
- 15) 原 潤一, 加藤博一, 井口征士: だ円膨張法を用い

た動物シルエット画像の記述と認識, 電子情報通信学会論文誌, Vol.J74-D-II, No.3, pp.366-375 (1991).

- 16) 木本伊彦, 浅井基博, 安田靖彦: だ円体によるシルエット図形の記述と表現, 電子情報通信学会論文誌, Vol.76-D-II, No.6, pp.1159-1167 (1993).
- 17) Tiberio, P. and Zezula, P.: Selecting Signature Files For Specific Application, *Inf.Process. Manage.*, Vol.29, No.4, pp.487-498 (1993).
- 18) Salton, G.: *Automatic Text Processing*, Addison Wesley (1989).

(平成7年 5月11日受付)

(平成7年 10月 5日採録)



金原 史和 (学生会員)

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5年同大学院工学系研究科電子工学専攻修士課程修了。現在、同大学院博士課程在学中。画像処理, 画像認識, 画像検索システム構築などの研究に従事。



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学客員研究員として *Inf. media* 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。同年東京大学生産技術研究所助教授。昭和48年~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, 電子情報通信学会各会員。