

輪郭線情報を用いた画像検索システム

加 藤 大志朗

- I はじめに
- II 輪郭線抽出
- III 短線分近似による粗グラフの抽出
- IV プリミティブグラフの抽出
- V 形状類似比較
- VI 結果の評価と今後の課題

I はじめに

近年、計算機の高速化、大容量化にともない、マルチメディア環境が身近なものになってきたのは、周知のとおりである。ユーザは大量のデータを管理しなくてはならず、そのデータ量にともない、データの検索のための労力も膨大なものを要求されつつある。本論文では、画像データベースにおける検索の一手法について論ずる。

画像データベースシステムにおいて、何が問題となるのかは、具体的な事例を考えて見れば分かり易い。例えば、カーテンや刺繍など、デザイン産業では、過去のデザインの膨大な蓄積がある。膨大な蓄積がある一方で、そのデザインはごく一部しか活用されていないのが実態である。1万件のデザインがあるとすると、顧客のニーズに応じて、そのなかから適当なデザインを選び出すことは、人間の能力をはるかに超えている。そこでデザインのデータベース化が自然と求められてくるわけであるが、既存のデザインを、ただ単に計算機上にデータとして保持するだけでは、問題は何も解決を見ない。顧客が「ピンク系統の、やさしい感じの、花柄のカーテン」を求めている場合、「ピンク」という色情報、「やさしい」という感性語、「花柄」という形状情報を検索キーとしてデータベー

スが検索できることが望まれるわけである。画像データベースを構築するにあたり、どのような検索キーを用いるかを決定することが重要である。

本論文では形状情報をもとに検索する画像データベースを考える。検索の過程は、図1に概略を示す。まずはじめに、データベースに登録する画像から、形状の情報を取り出す前処理を経て、その形状情報とともに、画像をデータベースに登録する。現在までの多くの画像データベースでは、登録時に人間が手作業で画像にラベルづけをし、そのラベルをもとに検索を行うような方法がとられている。計算機を用いて画像から人物を抽出するといったような、機械認識は種々の研究がなされているが、未だ実用には至っていない。したがって、手作業でラベルづけを行うようなデータベースの構築法は、現在のところ計算機を用いては行えないきめ細かな情報を画像に付加することは可能であるが、データベースに登録する画像の枚数の規模が増加するにしたがって、統一的な規則でラベルづけを行うのは非常に困難な作業である。本論文ではデータベースへの画像登録時に、計算機により自動的に形状に関する特徴量を抽出するシ

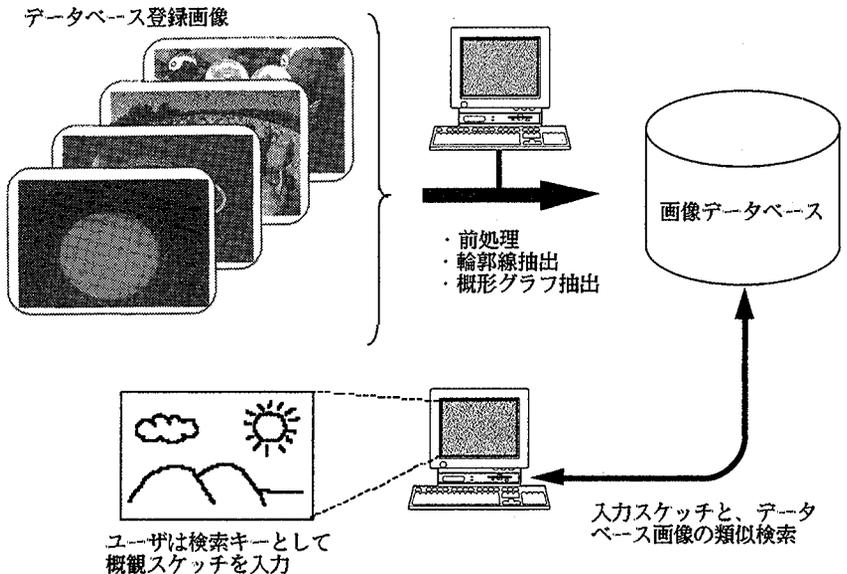


図1：画像データベースの作成・利用の流れ

システムを考えるが、機械認識の手法は一切用いていない。機械認識にかかわる部分は今後の課題であるが、ここで、本論文での特徴量の抽出について、どのような画像を、どのような意図を持った利用者が検索するか、状況設定を含めて説明したい。

まず、対象とする画像は、形状情報による検索を行うのであるから、輪郭の比較的はっきりとした画像である。イラスト画などを思い浮かべてもらえればよい。利用者は、データベースに登録されている画像を以前見たことがあり、「あの絵が引き出したい」ような場合、もしくは、「自転車」や「こんな形の家」のような、「これに近い絵はないか」というような意図を元にデータベースの検索を行う、いわゆる「あんなもの検索」[5]を想定している。実際の検索の段階では、利用者は、検索したいイメージを、輪郭線によるスケッチ(概観スケッチ)で計算機の画面上に描画して、そのスケッチをもとに、データベース画像を類似検索する。

概観スケッチを検索キーとした画像データベースでは、

- 1) 画像の形状情報の計算 (登録時の処理)
- 2) スケッチとの類似度計算 (検索時の処理)

の2つのフェーズに分けてそれぞれ処理を行う。本論文で、形状情報として、どのような特徴量を考えているかの説明も含めて、登録時の処理の流れを、大まかに追っていくことにしよう。登録時の処理の流れは、図2のとおりである。

まずはじめに、登録する画像から輪郭線を抽出する処理を行う。輪郭線の抽出には1次差分法によるエッジ抽出技法を用いた。このエッジ抽出には前処理としての画像平滑化の処理と後処理としての細線化の処理が必要となる。詳しくは2章で説明する。

次に輪郭線を用いた形状の特徴量を抽出するために、グラフ構造を抽出する。正確ではないが、一般にエッジ曲線は、画像に写っている被写体の輪郭か、色や濃度の異なる領域の境界線が得られるので、最終的に考えるグラフでは、閉曲線(ループ構造)または閉曲線を構成しない折れ線を、形状を把握するための最小単位(プリミティブ)として用いる。このプリミティブの抽出のため、

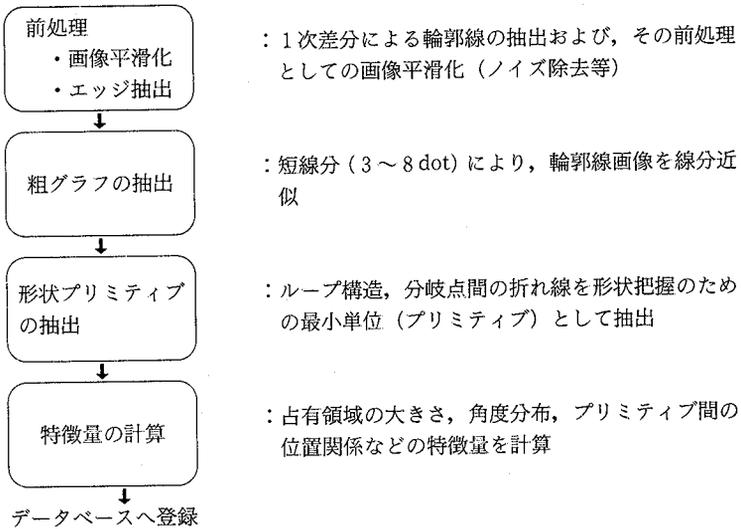


図2：データベース登録時の処理の流れ

また、抽出されたプリミティブの特徴量を計算するために、前処理の段階で抽出されたエッジ曲線から、長さ3～8の短線分により、エッジ曲線の折れ線近似をし、粗グラフを抽出する。粗グラフの抽出については、3章で説明する。

こうして得られた粗グラフから、プリミティブを抽出し、各プリミティブの属性値として、形状に関する情報を持たせることになる。プリミティブ抽出後は、プリミティブ（線分／ループ構造）がノード（node）で、プリミティブ間の連結関係がアロー（arrow）であるグラフが得られる（図3）。形状に関する情報としては、次の3種類の特徴量を用いている。

- ・位置情報
- ・占有領域の大きさ
- ・角度分布

形状に関する特徴量の抽出については、[3, 4, 6, 7, 9, 13] 等が研究されている。形状情報に限らず、特徴量の抽出とは、データ量の多い元データから、類似度計算等を行うための、比較的データ量が少ない、元データの特徴を表す

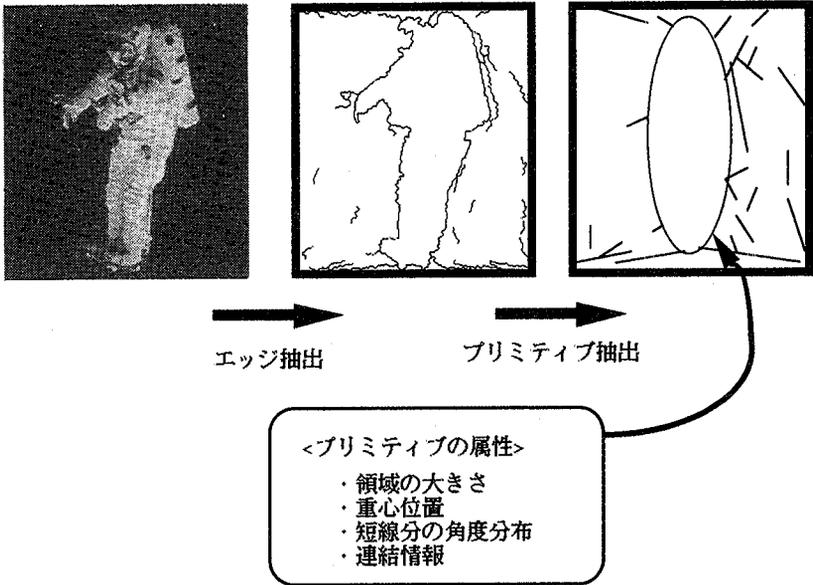


図3：特徴量抽出の流れ

情報を抽出する，情報圧縮技法と見ることができる。情報圧縮するわけであるから，技法によっては消失する情報も出てくる。形状比較のための特徴量として，Pフーリエ記述子を用いる方法，輪郭線の複素相関を用いる方法，音声認識等で用いられているメルケプストラム距離を用いる方法等が研究されている。本論文での方法は，角度分布を用いた，統計量による類似度検索と位置付けることができる。特徴量として統計量を採用した理由は，以下のとおりである。

まず，画像検索システムとして目指しているものは，特定の形状を抽出するものではなく，輪郭線を抽出できる画像全般を対象とした，一般的な画像検索システムである。検索システムの内部でどのような計算が行われているか，ユーザは詳細を知る必要もないし，一般ユーザにそのような知識を求めることは，システム的设计上避けるべきである。検索結果とユーザの入力スケッチをユーザが比較して，ユーザが納得できる，分かり易い結果を提示できることが重要

と考えた。形状の類似というのは、かなりの部分、主観に属するものであるから、検索結果は、とりあえずはユーザの望んだ結果ではないかも知れないが、分かり易い方がよい。第2に、上に述べたことと同様のことはあるが、システムを試作した段階で、その有効性の評価が行いやすいものと考えた。また、統計量を用いているが、形状の比較においては、データベース中の画像と、ユーザ入力のスケッチの間で、大きさの違い、回転による影響を吸収する方法がとられている。統計量以外の特徴量に関しては、6章で再び述べる。

II 輪郭線抽出

この章では、輪郭線の抽出までの処理を説明する。本論文では輪郭線抽出には、1次差分によるエッジ抽出技法を用いている。エッジ抽出技法には、2次微分(ラプラシアン; Laplacian)オペレータを用いた方法、他の1次微分に基づく方法(例えば Kirsch オペレータ等)、領域分割法を用いた方法等が知られているが[8, 10, 12], 今回は単純な1次差分法によるエッジ抽出を行った。ラプラシアンを使う方法では、抽出されるエッジ曲線は常にぼやけたものとなり、好ましい結果が得られなかった。Kirsh オペレータを用いた場合、1次微分法であるので、はっきりとしたエッジが検出され、また、オペレータ自身、局所的ノイズ除去の性質を持っているので、基本的には十分な性質を持ち合わせていると判断された。しかし、写真のような、ノイズを多く含み、また、コントラストが画面全体に均一でないような画像の場合、Kirsh オペレータを単独で用いても、期待されるエッジ曲線が得られない場合が非常に多かった。エッジ抽出においては、オペレータによって得られる画素の濃淡の変化の情報を、ある閾値を用いて、閾値以下の濃淡の変化はエッジ曲線として認めず、それ以上の変化をエッジ曲線として結果を出すという方法を取っているのに原因があると思われる。しかし、エッジ抽出における閾値法の採用は、不必要に大量のエッジ候補を抽出してしまったのでは、エッジ抽出後の処理に多大な負担をかけることになるので、避けられない。したがって、エッジ抽出の前処理として、ノイズ除去処理およびエッジ強調処理が必要となってくる。

そこで、本論文においては、[10, 11] のエッジを保ったスムージングと、単純な 8 方向 1 次差分法を組み合わせている。この組み合わせは、Kirsh 法を単独で用いた場合より、非常に良好な結果が得られている（図 4 の岩の表面写真参照）。

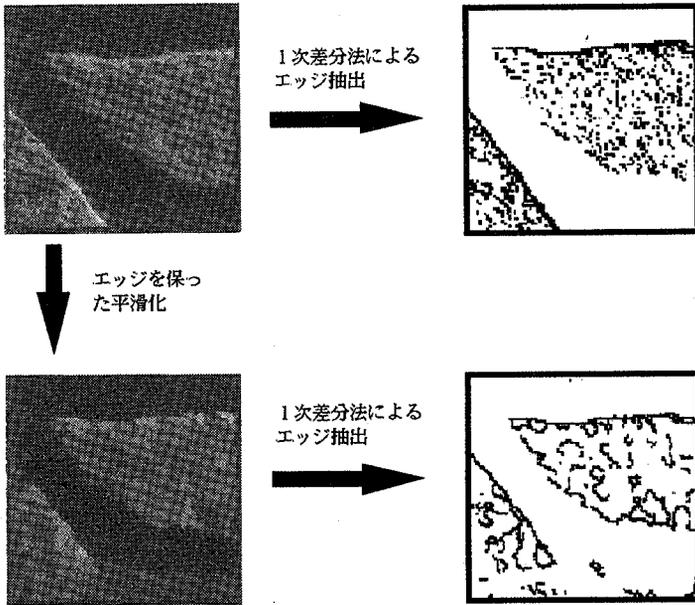


図 4：平滑化処理によるエッジ抽出の違い

また、差分法によるエッジ抽出では、一般に幅のある線分が得られる。幅のある線分では、後の処理である、グラフの抽出が的確に行われないので、幅のある線分を幅 1 の線分に細くする、細線化の処理が必要となる。

以上をまとめると、

- 1) ノイズ除去・エッジ強調 (長尾の方法)
- 2) 8 方向 1 次差分法によるエッジ抽出
- 3) 細線化処理 (Duff の方法)

の順に処理を行っていく。

デジタル画像において、ノイズ除去として、適当な $n \times n$ の正方形のメッ

シュ全体の濃度の平均値を，その中央の画素の新しい濃度とする方法が考えられている。しかし，この方法では，画像全体が暗くぼやけたものになってしまう。長尾らはエッジを保った平滑化を提案した。この方法は， 3×3 のメッシュをとり，中心点から8方向に伸ばした8つの，7点からなる矩型領域の内，濃度の分散が一番少ない矩型領域の濃度の平均を，中心点の新たな濃度とするもので，平滑化の結果はエッジも強調されている。

細線化の方法として，いくつかのものがいられているが，本論文のシステムではDuffの方法 [1, 8] を採用している。この方法は，8つの 3×3 のマスクを，条件が満たされなくなるまで，繰り返し画像に作用させて行くもので，ハードウェア化に適しているものである。

III 短線分近似による粗グラフの抽出

この粗グラフの抽出には，次のような必要性がある。1次差分法によって得られたエッジ曲線は，多くの場合，人間の直感に反して，海岸線のような複雑な折れ線を描くことになる。最終的には，形状の特徴量として，角度の分布，線分の長さを用いることを考えているのであるが，そのような複雑な折れ線では，得られる角度，線分の長さは無意味な量となる。人間が形状を認識するときは，そのような局所的な複雑さは無視し，大局的な量をもって形状を把握していると想像されるので，局所的なエッジ曲線の揺らぎは，大局的な量に吸収

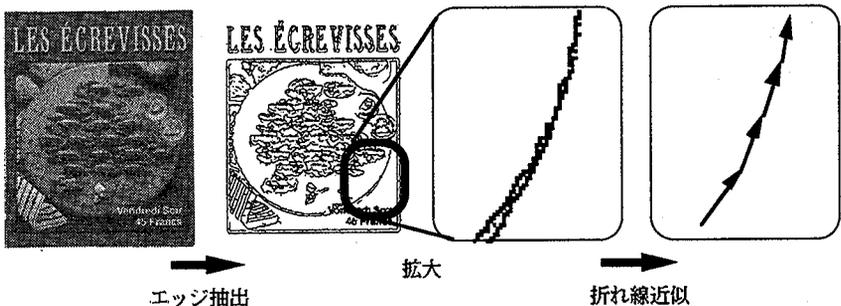


図5：粗グラフの抽出

されなければならない。図5にその例を示す。ザリガニが乗っている皿の右下部分を拡大した図を見ると、滑らかなはずの皿のエッジに細かい髭が生えている。元の画像に含まれているノイズ、エッジ抽出、細線化によって生じるノイズの多くは粗グラフの抽出で除去することができる。

粗グラフの抽出は、長さが3～8 dotの短線分により、エッジ曲線を近似することにより行われる。

IV プリミティブグラフの抽出

プリミティブの抽出は、図6のフローチャートに示されるように行われる。入力としては、3章までで得られた粗グラフである。

1次差分法による輪郭線抽出では、元画像から人間が受ける印象とは異なる、細かい輪郭線が抽出される。これは、人間の目では把握されない微妙な濃淡の変化も捉えてしまうためで、この問題を根本的に解決するためには、輪郭線抽出方式を別のものにしなければならないであろう。本論文のシステムで、この問題に対処する調整箇所は、輪郭線抽出段階で、エッジ曲線の濃度の閾値を上げるか、その後の段階での対処となる。しかし、エッジ曲線の濃度の閾値を上げることは、ほとんどの場合、輪郭線として抽出されるべきエッジ曲線が、閾値を下回る結果を招くことになる。したがって、輪郭線としては、妥当な輪郭線よりも多少多目のエッジ曲線を、輪郭線候補として抽出しておき、プリミティブ抽出段階で、小さな輪郭線候補からプリミティブを抽出することを抑制するようにしている。

また、これは本システムが、統計量による形状類似検索の有効性を確認するための試作的段階にあるためではあるが、現在のところ、輪郭線抽出以降は、領域の色情報や質感などのテクスチャ解析等による領域分割のための情報を使用していない。したがって、輪郭線からループ構造プリミティブを抽出する段階で、辺で連結したループ構造はすべてひとまとめにしてしまい、連結したループの最外周を1つのループ構造プリミティブとして抽出している。つまり、現段階では、形状情報は、部分的な形状の組み合わせとして形を認識するもので

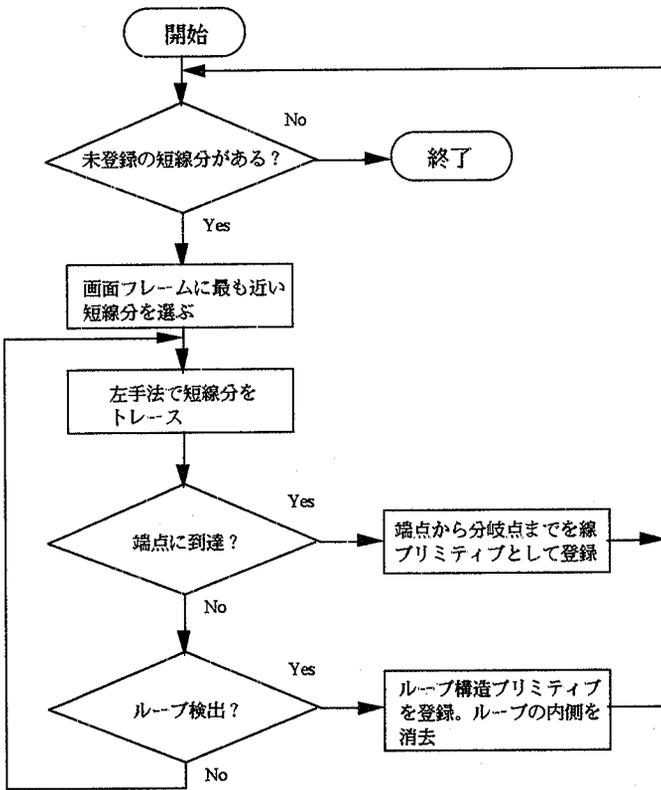


図 6：プリミティブ抽出処理

はなく、シルエットの外周を形状の認識単位としていることになる。これは、形状認識を行うための形状情報としてははなはだ不完全なものと捉えられる点であるが、本システムの目的は、ユーザが入力した概観スケッチから画像データベース中の画像を検索することである。あまりに敏感に形状を把握するものよりも、似たような画像候補を引き出してくれるものが期待されるので、現段階では、このような大ざっぱな形状情報にとどめておいた。この点は、6章の今後の課題のところで再び触れる。

V 形状類似比較

輪郭線による類似検索では、次の2種類の属性値を用いる。

- ・位置情報
- ・形状情報

形状の類似度計算も、これにともない大まかに捉えて2段階の計算で求めることになる。まずは若干の用語の定義を行う。

画像 G から得られたプリミティブグラフを G で表す。グラフ G の連結成分をコネクティブ (connective) と呼び、コネクティブの集合を C 、グラフのノードを構成するプリミティブの集合を P で表すこととする。コネクティブ、プリミティブそれぞれは属性値を持ち、それらの属性値の比較を行うことにより、画像間の類似度を計算する。

コネクティブの属性値として、

- ・コネクティブの重心位置
- ・占有領域の面積
- ・占有領域の固定軸方向の長さ

プリミティブの属性値として、

- ・プリミティブの重心位置
- ・プリミティブを構成する短線分の角度分布
- ・長さ (ループ構造の場合は外周長)
- ・占有領域の面積 (ループ構造のみ)

を用いた。コネクティブ $c \in C$ の重心位置の属性値を $c \cdot cent$ で表す。同様にコネクティブの他の属性 $a \in ATTRIB(C)$ の属性値は、 $c \cdot a$ で表す。この場合、 $ATTRIB(C) = \{cent, area, x-len, y-len\}$ である。プリミティブ $p \in P$ の属性値も同様に、属性 $a \in ATTRIB(P)$ に対し $p \cdot a$ で表す。 $ATTRIB(P) = \{cent, hist, len, area\}$ である。

ここで、1つ注意しておかなければならない。検索キーとしてのユーザのスケッチも、プリミティブグラフに変換され、類似検索は、プリミティブグラフ

間の類似度を計算することにより行われるのであるが、データベースに登録されている画像の大きさは、一般にユーザには未知である。検索キーとデータベースの画像の比較を行う場合、このフレームの大きさの違いを補正しなければならない。長さは対角線の比で補正を行い、位置に関しては、x 軸、y 軸それぞれの方向で、フレームの大きさととの比により補正を行っている。位置の属性値は、画像の左上端を原点とする座標値である。また、角度分布の属性値は補正を行わなかった。

2つのプリミティブグラフ G_1, G_2 が与えられたとき、本システムでは、この2つのグラフの類似度を以下の式で定義した。

$$\max_F \max_f \left(\prod_{c \in C_1} \prod_{a \in ATTRIB(C)} \Psi \cdot a(c \cdot a, F(c) \cdot a) \right) \times \left(\prod_{p \in P_1} \prod_{a \in ATTRIB(P)} \Phi \cdot a(p \cdot a, f(p) \cdot a) \right) \dots\dots\dots \text{式 1}$$

ここで、 F はグラフ G_1 のコネクティブと G_2 のコネクティブを1対1に対応付けする写像で、 f は G_1 のプリミティブと G_2 のプリミティブを1対1に対応付けする写像である。各プリミティブはいずれかのコネクティブに属しているので、写像 f は写像 F の選び方に影響を受ける。すなわち、 F が1つ定まれば、そのもとでの写像 f の選び方は F によるコネクティブの対応を越えてプリミティブを対応付けることは出来ない。

また、 $\Psi \cdot a$ は、コネクティブの属性 a に対応する、属性値の比較関数で、 $[0, 1]$ の範囲の実数値をとる。値1が完全な一致を表す。 $\Phi \cdot a$ は、プリミティブの属性 a に対応する、属性値の比較関数で、同様に $[0, 1]$ の範囲の実数値をとり、完全一致の時、値1をとる。例えば、コネクティブの位置の比較は、

$$\Psi \cdot cent(\alpha, \beta) = \left(1 - \frac{|\alpha_x - \beta_x|}{\max(\alpha_x, \beta_x)} \right) \times \left(1 - \frac{|\alpha_y - \beta_y|}{\max(\alpha_y, \beta_y)} \right) \dots\dots\dots \text{式 2}$$

となる。ここで、 α_x は一方のコネクティブの位置情報の x 座標を表す。長さ、面積に関する類似度計算関数も、式2と同様の形の式を使用している。

角度の分布の属性値は、 $0^\circ \sim 180^\circ$ までの角度を、適当な区間に区切り、その区間内の角度(x 軸との角度)を成す短線分の本数のヒストグラムである。この

ヒストグラムの類似度計算は、基本的にはヒストグラムを階段関数と見て、2つのヒストグラムの重なり部分の面積を、大きいほうのヒストグラムの面積で割って求める。

式1は、Fuzzy 論理の論理積として見てみれば、その意味合いは分かり易いであろう。また、式1は、本システムにおける形状マッチングの基本的性質を描き出している。つまり、画像の構成要素のどれとどれを類似形状の比較候補にすべきか、という組み合わせ問題が存在する。画像の構成要素の位置が、明確に判別できるものであれば、そのような問題は位置情報を考慮することにより回避できる問題であるが、検索のための入力、人間の曖昧なイメージを元にしたスケッチであるので、そのような解決方法は一般にとることが出来ない。そこで、計算量を軽減させるための工夫が必要となる。

組み合わせによる爆発を回避するための方策として、強欲算法による近似解法を用いた。また、このコネクティブまたはプリミティブの組み合わせの計算を行う上で、期待できない計算を途中で打ち切るために、SSDA法(残差逐次検定法; Sequential Similarity Detection Algorithm) [2, 8]における手法を採用した。式1の計算過程を考えてみよう。まず、コネクティブとプリミティブの対応画像 F および f を固定する。類似度の初期値は1で、計算の過程で、類似度の途中結果に次々に $\Psi \cdot a$ または $\Phi \cdot a$ の値を掛けて行く。 $\Psi \cdot a$ および $\Phi \cdot a$ の値は、0~1の実数値なので、類似度の途中結果は1から単調に減少していくことになる。例えば、式1の値を求める時は、はじめの閾値を0にしておき、閾値を下回らない間、 $\Pi \dots$ の乗算を繰り返し、閾値を下回れば計算の途中結果は放棄、最後まで進めば、それを新しい閾値にする、ということを繰り返し、計算回数を抑えながら最大値を求めることが出来る。

VI 結果の評価と今後の課題

本システムでの検索結果について述べる。

図7は、市販のイラスト集の画像で、全体で584画像ある内から、参考までに適当に選び出したものである。検索は、584全画像を対象に行った。データ

ベースに登録した画像のプリミティブグラフを検索キーとした時の検索結果を図8-1, 図8-2に示す。左上の画像が検索キーである。各画像の下の数値は、検索キーとの類似度である。図8-1が比較的良好な検索結果であり、図8-2は、不本意な検索結果である。以下、それぞれの場合について、考察を加える。

まず、図8-1、ノート型パソコンを検索キーとした場合。この場合、6位、8位、9位を除いて、全てノート型パソコンが検索結果として得られている。

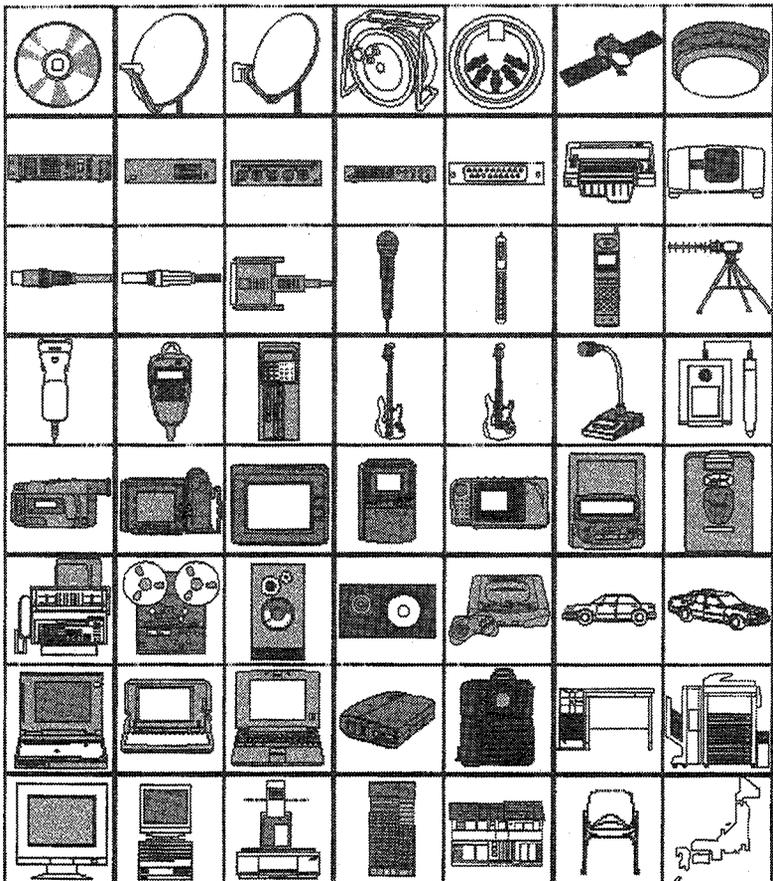


図7：テスト画像 (584画像中の56画像)

検索キーのノート型パソコンから得られるプリミティブグラフは、本体外側の輪郭線の四角形と、ディスプレイの輪郭線の四角形の2つのコネクティブをもち、一方が他方を包む形のものになっている。四角形の輪郭線から得られる短線分の角度分布は、4つの区間で鋭いピークを描き、他の区間では頻度がほぼ0のものとなる。したがって、角度のヒストグラムは、非常に特徴的なものとなり、同じ様な四角形から得られるヒストグラム以外との類似度は非常に小さくなる。また、1つのコネクティブが他方を包含する形は、2つのコネクティブ

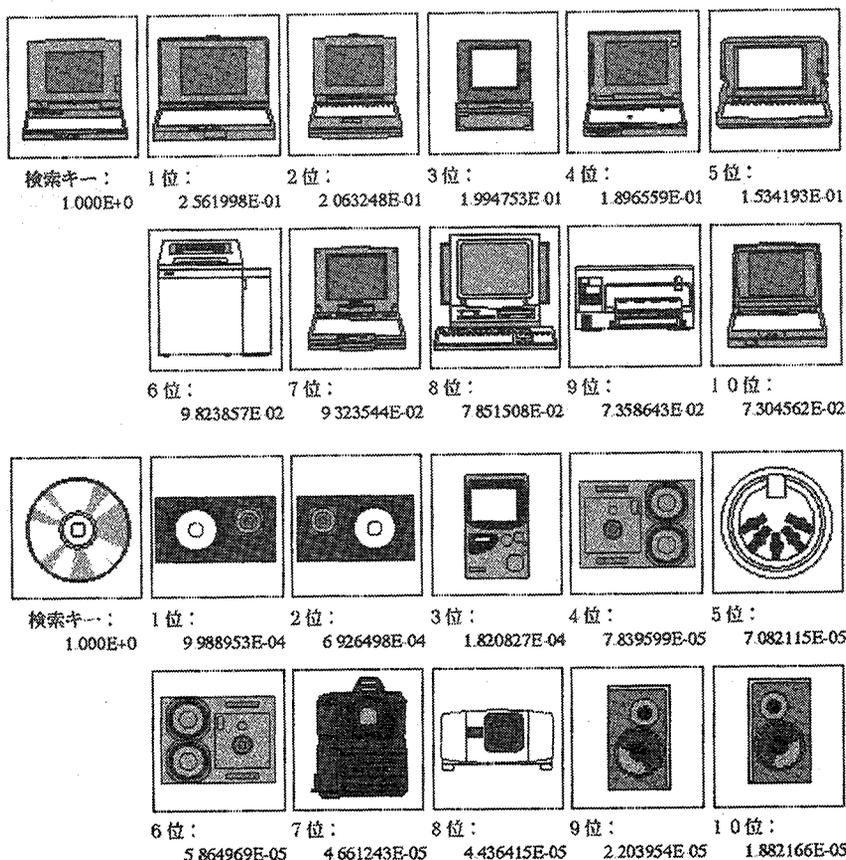


図8-1：検索結果その1

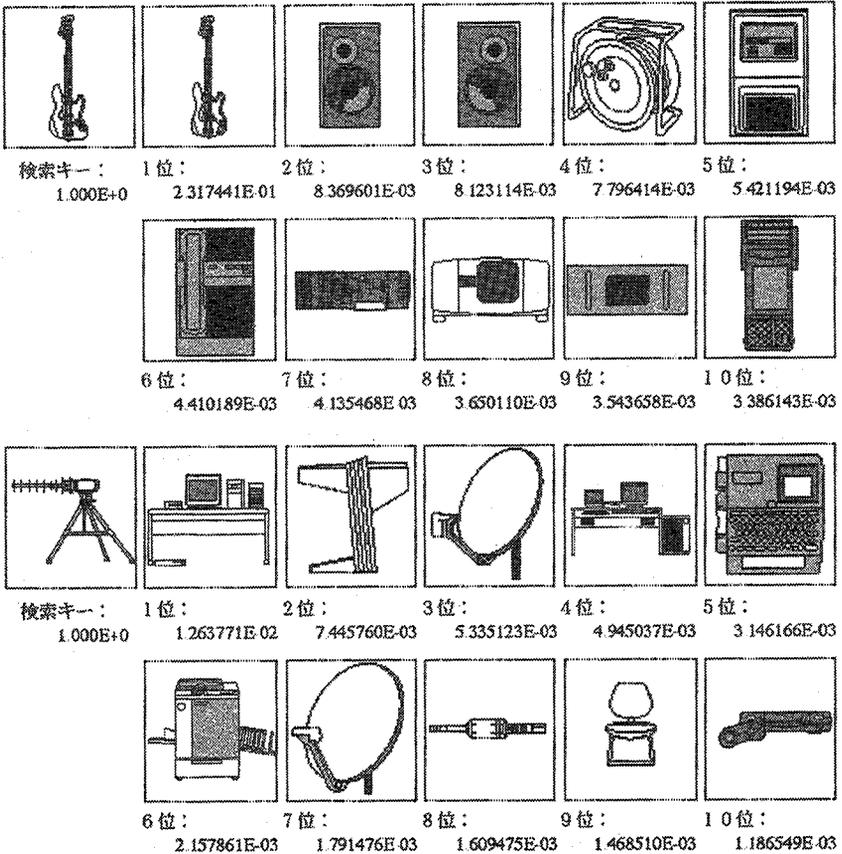


図8-2: 検索結果その2

ブが連結していないにもかかわらず、その重心位置が近接することになる。したがって、同心円状の輪郭線は、位置関係において、特徴的な配置関係となる。ノート型パソコンは、これら2つの特徴的の属性を構成するので、良好な検索結果が得られたものと推察される。

また、図8-1の、CDを検索キーとした場合は、CDの外周と中心の穴が同心円を形作っており、位置関係に関しては、上と同様である。角度のヒストグラムは、円の場合、各区間に均等な頻度が得られるので、円同士の類似度が非

常に高くなる。したがって、この場合も、角度のヒストグラム、位置関係の両方が特徴的な組み合わせなので、検索結果は比較的良好なものとなっている。

図8-2の上の例は、ギターを検索キーとして用いた場合の結果を示す。検索結果の1位はギターとなっており、検索キーそのものを含めて、上位2位がギターとなっているので、まだ救いはあるが、データベース中には、これら以外にあと2つのギターの画像があり、それが上位10位に入らなかった。ギター以外の検索結果は、(4位のコードリールも含めて)輪郭線の外周が四角形のものばかりである。この例の場合、グラフ抽出時の、短線分による折れ線近似により、ギターのフレームの曲線がつぶされてしまった結果、ギターの輪郭線が四角形として特徴量が抽出されてしまったのが原因のようである。特徴的な角度を保持することと、ノイズを除去することの調整の難点が表面に現れた例である。

図8-2の下例は、三脚付き八木アンテナを検索キーとした例である。この画像に類似する他の画像は、データベース中になく、検索結果に特定の傾向がみられないのはそのためであるが、ここで問題となるのは、類似の画像がないにもかかわらず、検索結果の類似度が比較的大きい値になっていることである。また、特に着目したいのは、1位、4位のシステムデスクと、3位、7位のパラボラアンテナである。この2つに加えて、ここでは登場していないが消防車の画像が、単一のコネクティブからなるような画像との検索において、結果の上位に良く現れる画像である。これは、それらの画像の角度のヒストグラムが、円のように均一ではないが、全ての区間に渡って、頻度が比較的高い値をもつようなものであるからで、角度のヒストグラムの類似度が高くなりやすい傾向にある。このような、多くの画像との間で、類似度が比較的高くなってしまいう画像が出てきてしまうのが、角度の分布に基づいた形状の類似検索の欠点である。

ここで、本システムにおける問題点と、それらの問題点に対し、解決のためのどのような方策を検討しているかについて述べる。以下、現在把握、検討している問題点を、輪郭線抽出部分に関する問題、位置情報の属性値の類似度計

算に関する問題、角度分布の属性値の類似度計算に関する問題、登録画像の数の規模に関する問題、それぞれに分けて説明する。

はじめに、輪郭線抽出部分の問題点。

形状情報を用いた画像の類似検索では、形状情報の情報源として、輪郭線を用いるのは極自然であり、この方向性は今後も維持されるものと思われる。輪郭線をもとに、各種特徴量を計算するので、はじめに抽出される輪郭線を、同じ画像から人間が把握する輪郭線にいかにか近づけるかが問題となる。輪郭線抽出部分では、ノイズ除去と、輪郭線の補完と言う、相反する方向の2つの種類の問題を解決しなければならない。

まず、ノイズ除去の観点から。本システムにおいては、輪郭線の抽出に、1次差分法を用いている。この1次差分法を用いて得られた輪郭線は、多くの場合、そのままでは人間の捉える輪郭線とは大きく異なる。図4に岩の表面の写真から抽出した輪郭線の例をあげた。人間が輪郭を把握する場合、1次差分法と同様に、濃度の不連続部分を物体のエッジ曲線として認識することも当然あるが、質感の差によってエッジ曲線を把握する場合も多々あると考えられる。図4はそのような事例に対応する例であると考えられる。質感の差は、画素のレベルで見ると大域的な情報なので、1次差分法のような局所的な情報しか用いないようなエッジ曲線の抽出技法ではおのずと限界がある。また、本システムでは、元画像の色情報は、輪郭線の抽出の段階でも、各種属性値の計算の段階でも使用されないのので、質感以前に色情報の利用を考えなければならない。

以上の考察から、輪郭線の抽出には、1次差分法によるエッジ曲線の抽出よりも、領域分割法によって得られる領域境界線を輪郭線として採用したほうが良好な結果が得られると予想される。しかし、現在のところ、領域分割法は、単純領域拡張法、反復型領域拡張法、分離・統合法など [8, 12] が知られているが、それらは、基本的に画素の濃淡または色による領域分割法で、質感といった、大域的な領域の性質を用いるものでもなく、質感の抽出に直接応用できるものでもない。本システムの研究を継続していく上で、次のテーマは、この質感を元にした領域分割法について取り組みたいと考えている。質感といった、

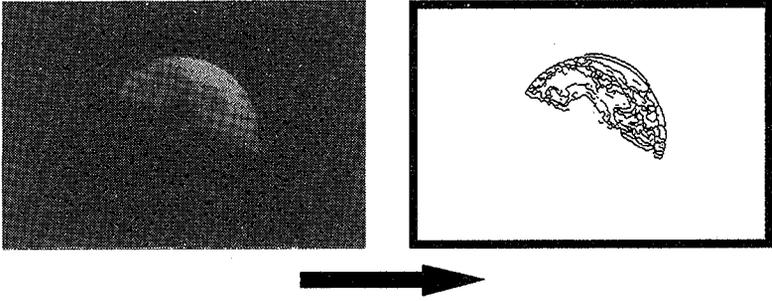
漠然とした特徴を、計算機の上でどのように数値化するかが問題となる。現在、Fractal が画像圧縮に応用されているが、Fractal による画像圧縮技法を、質感を定量化する特徴量として応用できないか、調査を行っているところである。

次に、輪郭線の補完について述べる。輪郭線の補完の必要性は、輪郭線抽出、プリミティブの抽出等、異なるレベルで必要とされる問題である。現在のシステムでは、輪郭線の補完の処理は含まれていない。まず、はじめに輪郭線の補完の必要性について、簡単に説明する。

まず、輪郭線抽出（エッジ曲線の抽出）の段階について。イラスト画など、コントラストが画像全般にわたって揺らぎが少ない様な画像については、輪郭線抽出段階でのエッジ曲線の補完の必要性が出てくることは少ない。1次差分法を用いる場合は、むしろ、輪郭線候補の濃度の閾値をいかにして決定付けるかが問題となる。この点も、今後の課題として解決して行かなければならない問題であるが、元画像から、1次差分もしくはラプラシアンを取って求めた濃度勾配から閾値を決定付ける方法が考えられ、現在のシステムにおいては、濃度勾配の上位 20% を閾値としている。この 20% という値は経験的なもので、今後、後の検索結果からのフィードバックによって、最適な閾値を求められないか、検討している。

議論を元に戻す。輪郭線抽出段階で、エッジ曲線が消失する現象の原因として、元画像において、コントラストが均一でないことが考えられる。図 9 にその例をあげる。図 9 は半月の写真であるが、月の昼の部分と夜の部分を分ける境界線は明確に抽出できず、得られた輪郭線の一部は、昼の部分の地形的な陰影が構成している。幸いなことに、このような現象はある程度、エッジを保った平滑化の処理で対応できるのであるが、それも限界がある。例えば、[10] のように、閾値よりも下回った濃度勾配についても、濃度勾配方向と直角方向にエッジ曲線を延ばす試みを行うものもある。しかし、この方法は試みたが、エッジ曲線の補完効果はほとんどなかった。輪郭線抽出の段階でエッジ曲線のある程度補完する必要性は認識しているが、残念ながらどのような解決法が望めるのか展望がない。これは、輪郭線抽出に領域分割法を用いた場合についても、

同様である。グラデーションのような、どこで区切りがあるのか分からない様な場合を考えて見ればよい。



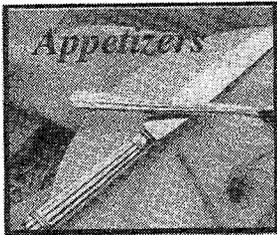
平滑化・エッジ抽出・細線化

図9：グラデーションによるエッジ曲線の消失

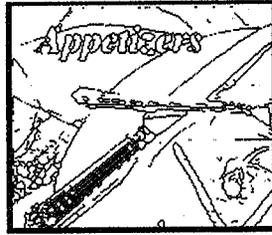
また、一度輪郭線候補を抽出した後、得られた輪郭線候補曲線の情報のみから、輪郭線の補完を行うことも必要である。輪郭線の欠如には、コントラスト等、画像の質が原因となる輪郭線の欠如ではなく、複数の被写体が重なることにより生じるものがある。画像検索においては、ユーザが「たしかこのような画像がデータベースの中に含まれていた」と、その画像のスケッチを入力するわけであるが、例えば図10の中の皿、ナイフ、フォークが重なっている部分を、重なっていないようにスケッチを書いて、検索キーとした場合を考えて見れば、隠蔽された輪郭線の補完の必要性が分かる。この手の輪郭線の補完を、ここで仮に大局的補完と名付けることにする。

大局的補完を実現するにあたっては、補完すべき輪郭線を描く対象物が何かということ、例えばそれは太陽のような円形のものなのか、フォークなのか、人間の足なのか、といった対象物の認識の過程を踏まえないと、補完すべき輪郭線の方向の予測がつかないのである。本システムでは、当初カルマンフィルター等を用いた、時系列予測を応用した大局的補完を考えていた。この場合、いくつかの形状に対応するフィルターをテンプレートとして持っておき、そのうちのどれかを補完部分に適用するようになりたいわけであるが、上で述べたように、テンプレート内のどれを適用すべきかを決定する部分で、機械認識を行

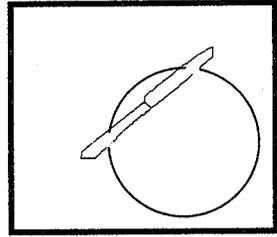
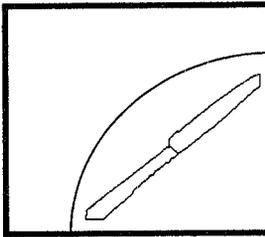
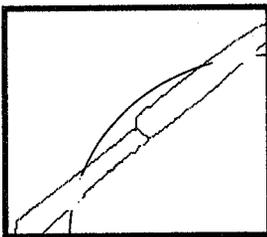
わなければならない。つまり、画像をデータベースに登録する段階で、輪郭線の補完をするために、機械認識の過程が入ることになる。画像データベースでは、将来的には機械認識を経てデータベースを構築することが必要とはなってくるが、現時点では認識の過程を含めることは考えていない。



元画像



輪郭線抽出画像



上の写真を思い浮かべて、ユーザはどのようなスケッチを描くのだろうか？

図10：元画像の重なりとユーザのイメージ

以上、2つのレベルの輪郭線補完の現状を考えると、輪郭線の欠如に関しては、補完という方向性で問題に対処するのではなく、別の方向性を考えたほうが良いように思われる。つまり、部分的な輪郭線から得られた形状情報による類似度計算の方法という方向性である。しかし、この方向性は、本質的な困難を持つと同時に、角度分布に基づく形状把握とは根本的に整合性が悪い。逆にいえば、角度分布に基づく形状把握の限界を、輪郭線の欠損の問題が如実に描き出している。ただし、このことは、形状認識に関しては、角度分布に基づく形状把握に限界があることを言っているのであって、画像検索に関してその有用性を否定しているわけではない。角度分布に基づく形状類似検索の有用性は、

1章で述べたとおりである。

次に、位置情報の属性値の類似度計算に関する問題点について述べる。5章の式1による類似度の計算式では、画像のどの位置に対象物があるのかという、絶対的な位置、もしくは画像のフレームとの相対位置により位置の類似度を算出している。しかし、画像フレームとの相対位置での類似度計算は問題がある。ユーザは、データベースから取り出したい画像を完全に把握しているわけではなく、印象に基づいて、検索キーとなるスケッチを書く。したがって、フレームとの相対的な大きさ、位置に関しては、忠実性がなく、描かれている対象物間の相対位置を、類似度計算における位置情報として用いるべきである。そのようにしなければ、例えば「自転車のある画像」を検索しようとして、自転車のみスケッチを検索キーとして描いても、位置情報の類似度が小さくなり、望ましい結果が得られない。5章では説明しなかったが、現時点でも、画像の対象物間の相対位置を元にした類似度計算を実現している。それは、式3に基づく類似度を用いている。式3においては、位置関係を画像 G_1 中の2つのコネクティブの相対位置と G_2 中の、対応する2つのコネクティブの相対位置の間で比較を行う。具体的な相対位置関係の類似度は式4で計算される。(ただし、式4中、分母の $\max \dots$ が0の場合は、分子も0なので、分母は1として計算。)

$$\begin{aligned} & \max_F \max_f \left(\prod_{\substack{c, c' \in C_1 \\ c \neq c'}} \Psi \cdot \text{cent}(c \cdot \text{cent}, c' \cdot \text{cent}, F(c) \cdot \text{cent}, F(c') \cdot \text{cent}) \right) \\ & \times \left(\prod_{c \in C_1} \prod_{a \in \text{ATTRIB}(C) \setminus \text{cent}} \Psi \cdot a(c \cdot a, F(c) \cdot a) \right) \\ & \times \left(\prod_{p \in P_1} \prod_{a \in \text{ATTRIB}(P)} \Phi \cdot a(p \cdot a, f(p) \cdot a) \right) \dots \dots \dots \text{式3} \end{aligned}$$

$$\begin{aligned} \Psi \cdot \text{cent}(a, \beta, a', \beta') &= \left(1 - \frac{|a_x - \beta_x - a'_x + \beta'_x|}{\max(|a_x - \beta_x|, |a'_x - \beta'_x|)} \right) \\ & \times \left(1 - \frac{|a_y - \beta_y - a'_y + \beta'_y|}{\max(|a_y - \beta_y|, |a'_y - \beta'_y|)} \right) \dots \dots \dots \text{式4} \end{aligned}$$

次に、角度分布の属性値の類似度計算に関する問題点について述べる。角度の分布の比較計算において、少なくとも2つの問題点がある。1つは、境界値問題。つまり、短線分の成す角度は、輪郭線抽出までの処理で生じるノイズも

含め、クリस्प値として見て意味のあるものではなく、ある程度の幅を考慮して考えなければならない。角度がヒストグラムの区間の中央値近辺の値であれば問題は生じないが、区間の境界値近くのものであれば、角度の値のブレによって、比較計算値に大きな影響を及ぼす結果となる。また、第2点として、ユーザの入力スケッチは、元画像の形状と、回転に関して一致しているわけではない点を考慮しなければならない。つまり、回転に関する普遍性をもっていない点ではない。今回のシステムでは、統計量を用いた形状比較方式の性質の確認が主な目的であるため、この2点は解決していない。

この2つの問題を解決法として、区間を Fuzzy 値で表現し、Fuzzy 値でのヒストグラムを構成することが考えられる。これにより、区間の境界値問題は解決される。回転補正に関する問題では、基本的には区間をずらして、ヒストグラムの類似度を求めることである。つまり、ある許容度の範囲内で、区間をずらし、そのうちで一番類似度の高い値を選ぶ。この方法では、回転のずれを、ヒストグラムの区間以下の値も許容する場合が問題となる。プリミティブの属性値としての、角度のヒストグラムは、プリミティブ抽出の前段階の短線分を元に算出される。短線分の x 軸と成す角度は、連続値をとるわけではなく、離散値となるので、ヒストグラムの区間の幅が、その離散値と比べて、比較的小さなものであれば、回転補正のための角度のずれは、ヒストグラムの区間単位で行えば十分で、区間幅以下の回転補正を考える必要はない。

最後に、データベースに登録する画像数の規模に関する問題点を述べる。データベースにおける検索では、データ数が多くなるに従い、検索時間が膨大となる。そのため、何らかの階層的な検索機構を設ける。本システムでは、現在のところ階層的な検索は実現していない。この点は、本システムの類似検索結果有効性の確認のためにも、画像データのクラスタリングを行い、階層化の実現可能性を確認していく予定である。

謝 辞

本システムを実現するにあたり、香川大学経済学部富永助教授に、ディスカッ

ションに参加し、有益な意見を述べていただくとともに、資料の調査を積極的に行っていただきました。また、京都大学数理解析研究所 泉田君、吉田君に、システムのコーディングを助けていただきました。OMRON 京都研究所の福島さん、山崎さん、藤井さん、栗林さん、京都ソフトウェアリサーチ 長沢さん、京都大学数理解析研究所 中島教授には、今回の研究のきっかけを与えていただき、また、器材面での援助をしていただきました。ここで、あわせて感謝いたします。

参 考 文 献

- [1] M. J. Duff, CLIP: A large scale integrated circuit array parallel processor, in *Proc. 3rd IJCP*, Cororado, CA, 1978, pp. 728-733
- [2] 尾上守夫, 前田紀彦, 斎藤優, 残差逐次検定法による画像の重ね合わせ, *情報処理*, Vol. 17, No. 7, pp. 634-640, 1976
- [3] 喜多伸之, 同心円特徴による類似形状検索, *信学技報 PRU* 90-121, pp. 79-86
- [4] 黒川雅人, 洪政国, 形状情報を用いた画像の類似検索システム, *情報処理学会論文誌* Vol. 32, No. 6, pp. 721-730, Jun 1991
- [5] 坂内正夫, 画像検索技術, *電子通信学会誌* Vol. 71, No. 9, pp. 911-914, 1988年9月
- [6] 佐藤真一, 大沢裕, 坂内正夫, 対象の多様性に対応しうる図面理解システムの一提案, *情報処理学会論文誌* Vol. 33, No. 9, pp. 1092-1102, Sep. 1992
- [7] 関田巖, 栗田多喜夫, 大津展之, 複素自己回帰モデルによる類似形状の検索, *信学技報* IE 89-7, pp. 17-24
- [8] 高木幹雄, 下田陽久, 画像解析ハンドブック, 東京大学出版会, 1991
- [9] 田邊勝義, 大谷淳, 形状類似画像検索における類似尺度の検討, *信学技報 PRU* 88-68, pp. 65-72
- [10] 長尾真, 画像と言語の認識工学, エレクトロニクス, モノグラフシリーズ2, コロナ社, 1989
- [11] 長尾真, 松山隆司, エッジを保ったスムージング, *電子通信学会研究会資料 PRL* 77-67, 1978年2月23日
- [12] 長尾智晴, 安居院猛, 画像の処理と認識, 昭晃堂, 1992
- [13] 新田祐介, 曾我真人, 今中武, 豊田順一, 柳田益造, 知覚に基づく類似図形パターン検索, *電子通信学会技術報告書データベースシステム* 82-4