

1. はじめに

近年の画像処理技術や光ディスク等の大容量記録媒体の普及に伴い、大量の画像データを蓄積・検索する画像データベースに対する要求が非常に高まってきている。それらの要求に答え、数多くの画像データベースシステムが構築されてきた。しかし従来の画像データベースシステムの取ってきたアプローチはキーワードを主としたもので画像の持つ情報の一部しか利用しておらず、画像の持つ情報を有効に利用した検索方法が望まれる。これに対して、画像処理の手法等を用いて索引となる情報を画像から直接抽出し、それらの値を用いて検索を行う画像の内容検索が一つの有効な手段としてその必要性を指摘されてきている[1]。

この種のシステムを構築する際には、

- (1) 計算機内部での記述形式：画像内の対象をどのようなモデルで記述するか。またモデル表現された2つの対象の類似度の判定方法をどのように定義するか。
- (2) マンマシンインタラクション：人間の検索したいものにたいする表象をどのような形でシステムに対して表現するか。及びその表現をシステム側でどのように解釈して(1)の記述形式にマッピングしていくか。の2つが大きな問題となる。

我々はこれらの問題点を解決して画像の持つ内容(情報)特に形状に基づいてより有効に画像を検索する方法(QB I / S : Query By Image contents/Shape)を研究している。[6]

最近になって、幾つかの研究報告がなされてきており[2][3][4][5]、代表的な例を挙げるとS. K. Chang等[2][3]は、画像をアプリアリに定義されたシンボルの集合とみなし、それらの間の空間的な関係を2-Dストリングという形式で記述した。画像検索はこれらのシンボルと関係を記述したストリングのマッチングのプロセスとなる。また、加藤等[5]は商標画像を対象に、画像の持つ空間周波数特徴や濃淡の特徴を局所的なパターン統計的な数値パラメータ列として表現した。画像検索はこれらの特徴ベクトルの距離計算によって行なわれる。前者は、空間的關係というシンボリックな関係を記述することが目的であり、それに対し後者は、画像パターンの持つ種々のアナログ的な特徴を数値パラメータで表現しようという試みである。この2つの記述方式は画像内の別の種類の情報を扱っていると考えられ、適切な画像の内容検索のためにはその両者を統合して扱うことが必要となる。

本稿では、シンボリックな関係として画像内対象物とその相互配置関係をグラフ構造で記述し、アナログ的な特徴として対象の輪郭形状及び相互配置関係の度合を数値的な属性表現で記述する。画像記述の方法としては属性付きグラフ構造を用いる。グラフ構造に対して類似の尺度を決め、この値によって類似検索を行う。属性付きグラフ構造を用いることの利点は、シンボリック、アナログの両者の特性を生かすだけではなく、位置関係や重なりという比較的あいまいさの介在する関係を属性値という数値情報で表現することにより比較可能な情報として扱うことができる点である。

また、マンマシンインタラクションに関しては、従来フリーハンドスケッチ入力[4]及びIconic-Index[2]による方法がある。フリーハンドスケッチの方法は表現の多様性があるが、逆にその多様性が特にあいまいな検索条件を持つユーザーに対しての負荷となる。ま

たIconic-Indexの方法は空間配置の表現を主な目的としているため、対象の形状等の詳細な表現は制限を受ける。ここでは、ユーザーに対して対象の部分画像を提示し、それらを組み上げて検索対象画像を作成するサンプル画像の選択と編集による方法[6]を提案する。

以下2章では本稿で用いる属性付グラフ構造の定義と原画像からの抽出手法について述べ、3章では検索のインターフェースの方法について述べる。4章では、ここで用いる属性付グラフ構造の類似度について述べ、5章では民族学の容器の画像を用いた実験システムについて述べる。

2. 属性付グラフ構造

2. 1. 属性付グラフ構造の定義

本稿では、画像の記述方式として属性付グラフ構造を用いる。Changらの定義[7]に従えば、属性付グラフ構造は、 V_n をノードラベル集合、 V_a をアークラベル集合とした時、以下の4つ組で表される構造である。

$$G = (N, E, u, e)$$

ここで

N : 有限のノード集合

E : ノード間の関係を表す順序対(アーク)集合

u : N と V_n を関係付ける関数

e : E と V_a を関係付ける関数

また、 V_n 及び V_a はどちらも (s, x) の形式をとり、

s : ノードシンボルまたはアークシンボル

x : $x = (x_1, x_2, \dots, x_n)$ の形を持つ属性ベクトルである。

各ノードはノードシンボルとノード属性を持つ。アークも同様にアークシンボル及びアーク属性をもつ。本稿では属性付グラフ構造を次のように定義した。ノードは画像から抽出された構成要素を意味し、アークは構成要素間の空間的な配置情報を示す。ノードシンボルは各構成要素の名前を示し、ノード属性としては、各構成要素の形状情報を数値パラメータで記述する。アークシンボルは2種類あり、構成要素間の空間的な関係として方向性及び重なりを示す。方向性に関しては、8方向のシンボルを用い、重なりに関しては、ないを示す2値のシンボルを用いる。(図1参照)

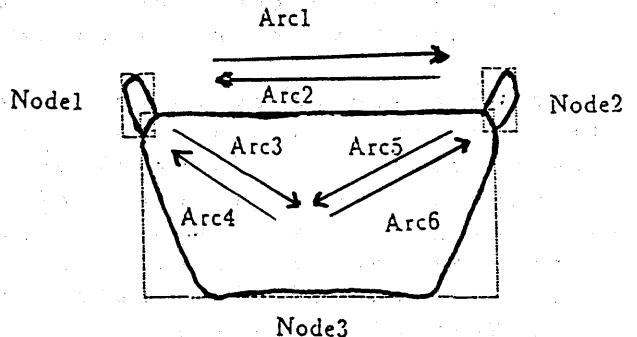


図1. 属性付グラフの例。

2. 2. 属性付グラフ構造の内容

ここでは、原画像から上記の属性付グラフ構造を作成するためのプロセスについて述べる。ただしここでは対象画像中の構成要素は、なんらかの画像処理手法で抽出されている事を前提とする。実際の実験システムでは濃淡8ビットの画像から構成要素を抽出するプロセスが存在するが、それについては第5章で具体的に述べる。

2. 2. 1. 形状情報の記述

蓄積・検索という立場から形状の記述を考える時、必要と思われる要素を挙げる。

(1) 位置・角度・スケールに関する不変性：
検索条件を例示画像で与える場合、回転・位置ずれ、拡大・縮小などは確実に起こりうる。このような入力条件を考えるとそれらの入力時の変形に安定な記述方法が望ましい。

(2) 記述の階層性：
検索のコストから考えると、まず粗くマッチングをとり、徐々に詳細なマッチングがとれるようなCoarse-to-Fineな構造を持ったものが望ましい。

以上の条件を満足する形状の表現方法として、ここではP-typeのフーリエ記述子を形状記述の方法として選択している。フーリエ記述子は、被展開関数によって幾つかの種類があり[8][9][10][11]、特に閉曲線の全曲率関数を用いたZahn[10]らのもの、閉曲線のある1点からの距離に関する関数とみたGranlund[9]らのもの、及びP表現を用いた上坂[8]のものが知られている。ここでは、適用範囲の広さを考え、閉曲線をも扱える上坂の記述子を用いる。以下では上坂のP表現について説明する。

曲線がX-Y平面上にあると仮定する。X座標を実数軸、Y座標を虚数軸と考えると曲線上の各点は複素数の形で表現できる。計算機上で曲線を実際に扱う場合は、有限個に離散化する必要がある。曲線Cを長さδのN個の点で離散化された多角形で近似して考えると各点は、

$$z(j) = x(j) + iy(j) \quad (j = 0, \dots, N-1)$$

の形で表現され、Cは複素関数Zで表現される。このZに対しP表現wを以下のように定める。

$$w(j) = (z(j+1) - z(j))/\delta$$

このP表現のフーリエ記述子をcとするとcはw(j)を離散的フーリエ変換して

$$\hat{c}(k) = \frac{1}{N} \sum_{j=-\frac{N}{2}+1}^{\frac{N}{2}} w(j) e^{2\pi i k j / N}$$

$$(k = -\frac{N}{2} + 1, \dots, \frac{N}{2})$$

の形で与えられる。

被展開関数としてP表現を用いた場合の妥当性については次のように考察される。スケールに関しては多角形近似の際のサンプリングの幅によって吸収できる。また移動及び回転に関してはP表現の開始点が曲線に対して一意に決定できれば吸収できる。また記述の階層性に関しては、フーリエ記述子の性質から低次の記

述子には大まかな曲線の特徴が記述され、高次の記述子にはより詳細な特徴が記述されていることが知られている。開始点の選択のためには、等辺多角形上の各辺の重心の分布を等辺多角形の重心を中心にして求め、これに対してKL展開[12]を施し、その主成分軸上に最も近い点を開始点として選択する。この正規化処理によって移動に関する影響は取り除かれる。各ノードに当たる構成要素にたいし上記のプロセスでフーリエ記述子を算出し、これをノード属性値として用いている。

2. 2. 2. 位置関係の記述

アークが表現する情報として2つの構成要素間の方向性と重なりを用いる。ここでは抽出された構成要素の重心と外接矩形を求め、方向性として重心間の方向ベクトル、重なりとして外接矩形の重なりを度合を数値化しアーク属性とする。方向性の場合、ベクトルの方向性を8方向に量子化してアークシンボルとし、また方向ベクトルの単位ベクトルをアーク属性として記述する。重なりの場合、外接矩形間に重なる部分がある場合、重なりのアークシンボルをつけ、重畳する部分の各外接矩形に対する位置をXおよびY方向に関して記述する。形状情報と同様に拡大・縮小等の影響を避けるため、方向性に関しては単位ベクトルを用い、重なりに関しては、各構成要素の外接矩形の辺を基準とした相対的な位置を記述する。(図2参照)

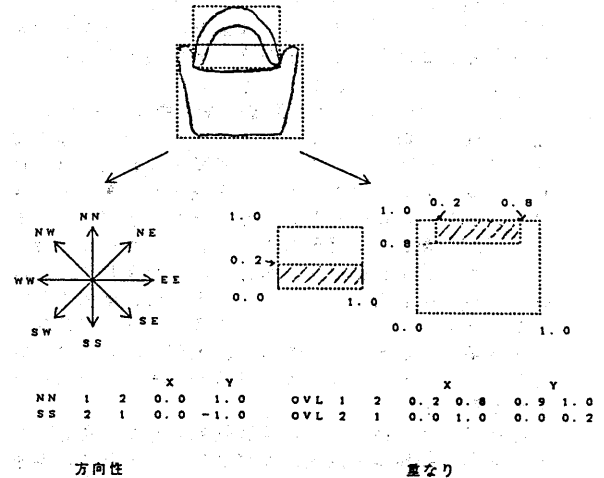


図2. アークラベル

3. 検索インタフェース

3. 1. ユーザーインタフェースの指針

計算機内部での対象データの記述と人間の持つ検索したいものに対する表象の間には大きなギャップがあるため、検索条件の表現方法は非常に重要な問題となる。特に、形状情報は画像パターンとして人間には容易に認識・記憶されるものだが、言葉等の2次情報で表現することは比較的難しい。このような情報を検索条件とする場合は本来画像として表現すべきである。

そのため従来は、フリーハンドによるスケッチ入力をもとにした検索条件の表現方法が報告されている[4][5]。しかし全く白紙の状態からスケッチを描かせることはユーザ側の負荷が大きく、スケッチ自身にも誤りが大きい事が予想される。またそれを解釈する計算機側から見るとラフに描かれたスケッチ画像は特徴抽出時に必要な情報の欠落が多く不正確な画像記述の要因となる。そこで本稿では計算機側で蓄積されている画像を構成要素毎に集めそれらを予めクラスタリングしておく。分類された各クラスタの代表パターン(サンプル画像と呼ぶ)をユーザに提示し、ユーザはそれらの中から自分の条件に合致したものを各構成要素毎に選択する。選択された画像を雛形としてそれらを組み合わせ、編集することによって自分の意図している形状を画像化していき最終的に検索用画像を作成する。

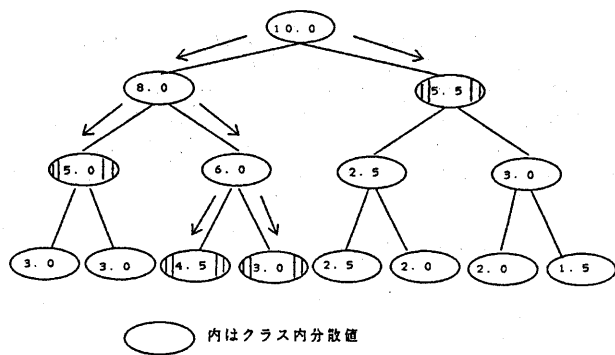


図3. サンプル画像の決定方法 (L=4の場合)

3. 2. ユーザーインターフェースの実現

ユーザーに対する提示は各構成要素毎に行なわれる。そのために各構成要素毎にデータを集めそれらの特徴パラメータ(ここではフーリエ記述子)を用いて階層的クラスタリングを行う。各クラスタの重心にユークリッド距離で最も近いものをそのクラスタの代表パターンとして用いる。次にクラスタ木を2分木に変換し、各クラスタに対して、インデックスとして2分木のコードを割当てる。インデックス、クラスタ内分散、及び代表パターンの番号をサンプル画像用データとして蓄積する。提示するサンプル画像は、クラスタ木の根から順にこれらのデータを見ながら決定していく。

決定のプロセスを、図3に示す。すなわち、L個のサンプル画像を集合中から決定する場合、最も分散の大きなクラスタから順次より小さなクラスタに分割していくという基準でクラスタ木を降りていって表示すべきクラスタを決定し、表示すべきクラスタがL個になったところでその代表パターンをサンプル画像として提示している。ユーザーは、提示されたサンプル画像の中から自分の持つ条件に最も合致した画像を選択する。また、提示された画像を指定してより下位のクラスタの画像を同じような表示形式で見ることが出来る。(この場合、上記のプロセスで根となるクラスタが指定された画像が代表しているクラスタと置き換わる。)

選択された画像に対する編集機能としては移動、拡大・縮小、輪郭線の消去、及びタブレット上での手書き入力がある。ユーザーはこれらの機能と選択したサンプル画像を用いて自分の要求に合った検索用画像を作成する。

4. 検索処理

本稿における検索処理は、基本的に検索用例示画像の属性付きグラフ G_i と蓄積されている画像の属性付きグラフ G_s のマッチング処理に相当する。マッチングのプロセスは検索用画像と蓄積画像のグラフ間の対応を取る構造マッチングと対応後の属性数値を用いた距離計算の2つのフェーズよりなる。

4. 1. 構造マッチング

ユーザーの作成した検索用画像は、本来検索すべき画像の部分画像であることが考えられる。そのため構造マッチングのプロセスは蓄積画像のグラフ(G_s)の中に検索用画像のグラフ(G_i)が部分グラフとして存在するか否かを判定するプロセスとなる。ここでは、以下の処理を各蓄積画像に対し行った。

- (1). G_s , G_i の各々について隣接行列を作成する。
- (2). G_i の各ノードに対しノードシンボルの等しい G_s 中のノードをリストとして与える。
- (3). (2)のリストを用い、構造マッチングを行う。

この問題は部分グラフの同型問題であり、 G_i と G_s の対応するノード(マッチングペア)によって作られる状態木の探索問題として扱うことができる。

実現する際には状態木の探索にはバックトラック付きの深さ優先探索を用い、検索用画像の全てのノードに対してマッチングペアが見つかった場合にはそのマッチングペアを候補として次の距離計算プロセスに渡すこととしている。対応の判定の際には、単にアークシンボルを比較するだけではなく、方向性の場合、 G_i の持つ方向ベクトルと G_s の持つ方向ベクトルの内積が一定のしきい値以内であれば他のシンボルのものでも対応すると見なしている。また、重なりの場合には、シンボルの判定後そのノードの持つ重畳部分の相対的な面積の比を取り、しきい値の範囲以内ならば対応がとれたものとしている。

4. 2. 距離計算

4. 1. の基準を満たした蓄積画像集合に対し、ノード属性を用いて距離を定義し、その値で順序付けを行なう。一般にフーリエ記述子の高次成分は、雑音等の影響を受けやすく信頼性が低い。ここでは16次元までの低次の記述子を用い、それらのユークリッド距離の和を距離尺度とする。上記の距離を算出し、値の小さなものから順に候補画像として出力する。

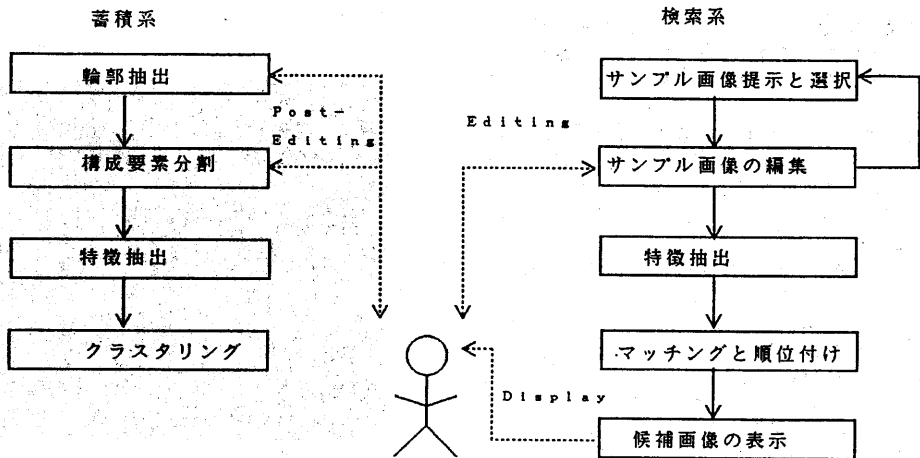


図4. 実験システムの概要.

5. 実験システムの概要

5. 1. システム構成

ここでは、前章迄で述べてきたアプローチを用いた試作システムについて述べる。処理対象としては民族学における容器の標本画像を用いた。これらはCCDカメラで正面方向から撮影されたもので、サイズは512X512画素、濃淡8ビットのデータである。対象画像に対し、本体、取手、足、首の4つの構成要素（ノードシンボル）を設定し、それらを用いてサンプル画像の作成及び検索を行うシステムを作成した。

5. 2. 蓄積系の処理

図4に処理の概要を示す。システムは蓄積系と検索系の2つに大別される。蓄積系の処理の内容は次の通りである。

(1)輪郭抽出：原画像中より背景を除去し対象となる標本を抽出する。対象となる濃淡画像の平滑化を行ない、確実に背景である部分（各画像の上下20ライン）の濃度分布を取り、そのピークをしきい値として2値化する。2値化された画像に対し膨張・収縮処理を各1度行ない、次に黒画素の連結成分のラベリングを行う。連結成分の面積及び外接矩形の辺の長さがしきい値よりも小さいものは雑音成分とみなして除去する。抽出された画像から8連結の意味での境界追跡を行って輪郭線を抽出する。抽出された輪郭線はユーザに提示され、不十分な部分を対話修正する。

(2)構成要素分割：修正された輪郭線画像に対し、ユーザ側からグラフィックカーソルを用いて構成要素の輪郭線部分を指示する。またその際に構成要素の名前を選択する。これが属性付グラフのノードシンボルとなる。

(3)特徴抽出：(2)までの処理でグラフのノード及びノードシンボルは決定される。アークシンボル及びアーク属性値の決定は、各構成要素に関して外接矩形及び重心を算出することによって求める。ノード属性であるフーリエ記述子の算出のためには、輪郭線をN個の等長の線セグメントに分割し、等辺多角形で近似する

。現在はN=256を用いている。作成された等辺多角形に対して2章で述べた正規化処理を施して開始点を決定する。その開始点から順にP表現を作成し、それに対して離散的フーリエ変換を施す。

5. 3. 検索系の処理

検索系の処理の内容は次の通りである。

(1)サンプル画像提示と選択：サンプル画像は図5の様にウィンドウ内に表形式で提示される。提示される代表パターンの選択には3章で述べている尺度を用い、代表パターン間の階層関係は表示されない。また、ユーザーはウィンドウ内の画像を指定することによって、分類木上でその画像の代表するクラスタの下位に当たる部分を参照することができる。この場合に提示する代表パターンは、指定したクラスタを根として同様に3章の尺度に従う。ユーザーは上記のプロセスを用いて各構成要素について自分の条件に合致した画像を選択する。

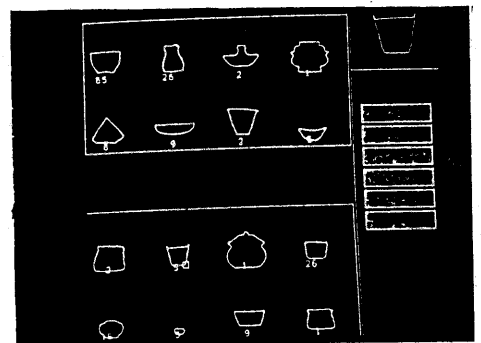


図5. サンプル画像表示の例（本体の部分）

(2) サンプル画像の編集：(1)のプロセスで選択された構成要素毎のサンプル画像を対話的に移動・編集することによって自分の表現したい画像を作成する。(図6参照)

(3) 特徴抽出：作成された検索用画像のグラフ構造を抽出する。ノードシンボルは選択された際の構成要素名を用いる。アークシンボル及びノード属性とアーク属性に関しては5.1.の(3)で述べた処理を用いて自動的に抽出する。

(4) マッチングと順位付け：検索用画像のグラフ構造と蓄積画像のグラフ構造間のマッチングを4章に述べた方法で行ない、距離の小さなものから順に並べて出力する。図7では左上が第1候補となっている。

システムは、FORTRANで記述され、IBM3081上のVM/CMSのもとで動作している。ユーザーインタフェース及び表示系はGKSを用い、装置としてはIBM5080グラフィックディスプレイ及び付属のロケータを用いている。

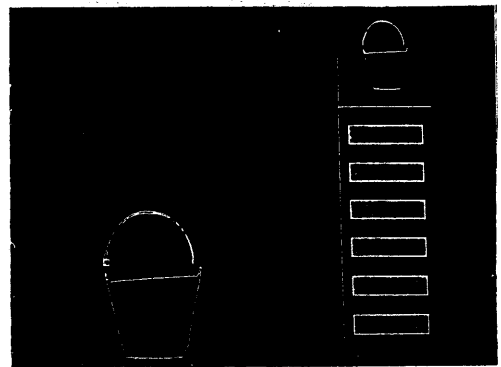


図6. サンプル画像編集の例。

5. 4. 実験結果

ここでは、上記のシステムの有効性を示すために行った予備実験及び実際の検索実験について述べる。まず予備実験として本稿で用いたフーリエ記述子の弁別能力及びサンプル画像の信頼性を確認する意味で、人間が見て明らかに同一のクラスとみなす画像群を用いて階層的クラスタリングを行った。実験には図8に示す15枚の画像(本体の部分)を用い、これらのデータは明らかに5つのクラスに分割される。これらの画像に対し実際にクラスタリングを施した結果が図8である。

次に実際の検索実験として108枚の画像を入力し、これらに対して上記の15枚を検索対象と考えて入力を行って検索した。各クラスの標本個数をM個とすると、検索対象画像が含まれているクラスの画像が上位M位以内に出現したものが89%であった。

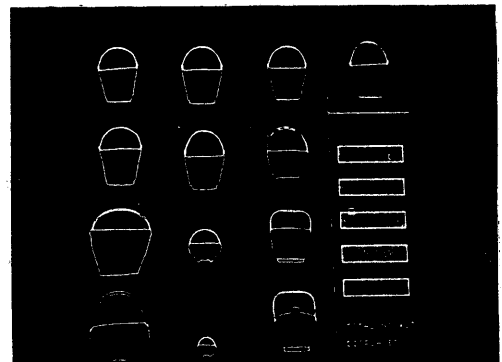


図7. 類似検索結果(右上の画像が検索用画像)。

6. まとめ

本稿では、画像検索における主要なテーマである画像の類似検索に関して属性付きグラフ構造を用いた形状情報及び空間配置情報の記述とその比較によって実現する方式について述べた。ここで用いている手法は、画像の持つシンボリックな特徴とアナログ的な特徴の両者の持つ情報を統合して扱っている点が特徴である。次に、この種のシステムの実用化を考えた場合に最も大きな問題となるマンマシンインタフェースに関しては、サンプル画像の選択と編集の方法を提案した。この方法はユーザーの持つ比較的あいまいな検索条件を、例示画像という自由度の高い表現形式で設定する際に有用な手段であると考えられる。また、ここで提案した手法をもとに、民族学の容器画像108件を対象に実験システムを作成した。実際に検索実験を行い、本手法による検索候補の限定及びその妥当性を示す事ができたと考える。

以下に問題点と今後の課題を挙げる。

(1) 構造マッチング：部分グラフの同型性判定の問題は、非常に計算コストがかかる問題であり、それに対して現状では、Blind-Searchに近い形で対応している。これは、ノード数やアーク数が増加した場合大きな問題となる。属性値等を用いた評価尺度によって、よ

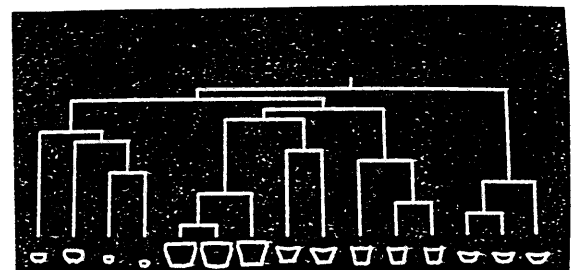


図8. 試験用データ及び分類結果。

り効率的なTree探索方法を組みこんでいく必要がある。
(2)形状類似尺度：現在は、フーリエ記述子による特徴空間を類似尺度の空間としているが、実際にはユーザーの持つ類似の感覚を何らかの空間として定義し、そこへのマッピングを定義することによってより効果的な類似検索が可能となろう。

謝辞

本研究は国立民族学博物館と日本アイ・ビー・エム(株)との間で行なわれた共同研究「民族学情報有効利用のためのコンピュータ応用手法についての基礎研究」(1986-1988年)の一環として実施されたものであり、ここで使用した画像の全ては国立民族学博物館より提供されたものです。本研究をすすめる上で協力頂いた東京基礎研究所の井岡幹博氏に感謝致します。

参考文献

- [1] T. Tamura and N. Yokoya: "Image Database Systems: A Survey", Pattern Recognition, Vol. 17, No. 1, pp. 29-43, 1984.
- [2] S. K. Chang et al: "An Intelligent Image Database System" IEEE Trans on Software, Vol. 14 No. 5, pp. 681-688, 1988.
- [3] S. Y. Lee et al: "Similarity Retrieval of Iconic Image Database", Pattern Recognition, Vol. 22 No. 6, pp. 675-682, 1989.
- [4] 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", 信学論(D), Vol. J72-D, No. 4, pp. 535-544, 1989.
- [5] 太田友一, 森良哉, 坂井利之: "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", 信学論(D), Vol. J64-D, No. 11, pp. 997-1004, 1981.
- [6] M. Kurokawa: "An Approach to Retrieving Images by Using Their Pictorial Features", Proc of ICIP89, pp. 588-592, 1989.
- [7] N. S. Chang, K. S. Fu: "A Relational Database System for Images": NATO Advanced Institutes Study Series Vol. 81, pp. 288-321, 1982.
- [8] 上坂吉則: "閉曲線にも適用できる新しいフーリエ記述子", 信学論(A), Vol. J67-A, No. 3, pp. 166-173, 1984.
- [9] G. H. Granlund: "Fourier Preprocessing for Hand-printed Character Recognition", IEEE Trans on Comput., Vol. C-21, pp. 195-201, 1972.
- [10] C. T. Zhan and R. S. Roskies: "Fourier Descriptors for Plane Closed Curves", IEEE Trans on Comput., Vol. C-21, pp. 269-281, 1972.

[11] T. P. Wallace and P. A. Wintz: "An Efficient Three-dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors", Comp, Vision, Graphics & Image Processing, Vol. 13, No. 2, pp. 99-126, 1980.

[12] 長尾真: "画像認識論", コロナ社, 情報工学講座 16, 1983.