

## 画像認識システムにおける例示による対象物モデルの獲得の一手法

有田 大作 鶴田 直之 谷口 倫一郎 雨宮 真人

{arita, tsuruta, rin, amamiya}@is.kyushu-u.ac.jp

九州大学総合理工学研究科

〒 8 1 6 春日市春日公園 6 - 1

あらまし

従来、画像認識システムでは対象物モデルは人手で与えており、モデル作成のため多くの手間がかかっていた。画像認識システムを様々な用途に用いるためには、いくつかのサンプルを与えて対象物モデルを自動的に作成する手法を確立することが、必要不可欠である。本稿で述べるモデル獲得の手法は、画像の2次元的特徴を基にモデルを作成するものであり、モデル作成の基になるのは、分割木と呼ばれる領域分割を階層的に表現した木である。分割木は根に近いノードほど、大局的な情報を担っている構造になっており、また、モデル自身も一種の分割木である。

和文キーワード 物体認識 モデルの自動獲得 領域分割

A Model Acquisition Method for Object Recognition  
by Pictorial Examples

Daisaku Arita, Naoyuki Tsuruta, Rin-ichiro Taniguchi and Makoto Amamiya

{arita, tsuruta, rin, amamiya}@is.kyushu-u.ac.jp

Graduate School of Engineering Sciences, Kyushu University

6-1, Kasuga-koen, Kasuga, Fukuoka 816 JAPAN

Abstract

In this paper, we describe a method to construct, from a series of images of objects, a model of an object class automatically. The model is described in a set of 2 dimensional features of segmented regions and relations among the segmented regions. To make models of objects, we introduce an idea of segmentation tree, which represents abstraction level of image segmentation, and introduce a model generation method from a series of the segmentation trees. Using this segmentation tree, we can cope with the problem of diversity of segmentation patterns and can make stable object models easily.

英文 key words Object Recognition Automatic Model Generation Region Segmentation

## 1. はじめに

画像認識において、画像中のどこに対象物があるかを探索したり、画像中の領域が何であるかを認識するためには、対象物に関する知識を予めシステムが持っている必要がある。これまで提案されている多くの画像認識システムでは、この知識をシステム開発者が構築している。しかし、画像認識の技術を様々な分野に応用しようとする場合、一つ一つの対象物に対しての知識を開発者が構築するのでは極めて困難があるので、システムが対象物に関する知識を自動的に獲得する方法が必要になってくる。

上述の観点から本研究では、ユーザが例題として「対象物の名前」およびその「画像上での領域」を入力することによって、システムがその対象物についての知識を獲得し、その知識を利用して対象物を探索・認識する方式を開発することを目指す。これによって、開発者は知識を構築する必要がなくなり、ユーザは対象物についての知識を容易にシステムに与えることができるようになる。

同様な知識獲得を目指した従来の研究としては、Connellらの研究[1]、秋山らの研究[2]、HarveyらのSPAM[3]、村瀬らの研究[4]等がある。Connell、秋山らの研究は、主に2次元の線画や輪郭線を対象としたものであり、実際の3次元物体を対象とした画像には適用し難い面がある。また、SPAMはユーザが対象物の構成部品の画像上での領域を指定することで、その部品に関する知識を自動的に獲得し、すべての部品についての知識を得ることで、対象物を認識できるようになるものである。しかし、この方法では知識獲得時にユーザが指定した部品が、認識時の領域分割において単一領域として抽出されなければならないが、構成部品の選択はユーザによって行われているので、この条件を満たす部品を選択することが難しいという問題がある。村瀬らの研究は、画像を固有ベクトル空間上で表現し、認識・学習しようとするものであり、単純な方法で比較良好な結果を得ている興味深いものである。しかし、認識時にも対象物が完全に切り出されている必要があるため、背景に様々なオブジェクトが存在するという一般的な状況では適用できないという問題がある。

本研究は、3次元物体が写っているカラー画像

を対象として、例として示される複数の画像（例示画像）から画像認識のためのモデルの獲得法を確立することを目的とする。モデルを獲得するための例示画像は、対象物のみが存在している（あるいは対象画像の中で、ユーザが対象物の全体領域を指定する）ことを仮定するが、実際に獲得されたモデルに基づいて認識を行う場合は、対象物のみが切り出されていることを仮定しない。本稿で述べる手法は、基本的には、対象画像の領域分割を基に、その領域の特徴や領域間の関係でモデルを記述しようとするものである。一般には、領域分割の結果は様々なパラメータによって影響を受けるため、安定した領域分割の結果が得られず、モデルを生成することが難しい。そこで、ここでは、領域の抽出され易さを基に、領域分割の結果を階層的に表現することで、パラメータによらない安定なモデルを作成することを試みる。本稿では、モデルの表現、モデルの作成法、本手法の有効性を示すための予備実験について報告する。

## 2. 対象モデルの表現

### 2.1 領域分割と分割度

先にも述べたように、画像の領域分割を行う場合には、どの程度の細かさで領域と見なすかということが問題になる。たとえば、エッジベースの方法では、どの程度の強さのエッジまでを領域の境界線と見なすかにより領域分割の結果が異なる。また、領域ベースの方法では、どの程度の類似性があったときに2つの領域を統合するのかにより結果が異なる。

ここで、この領域分割の細かさの程度（これは領域分割の結果を制御する領域間の類似度パラメータに対応する）を分割度と呼ぶことにする。ある画像に対して分割度を大きくしていくと、今まで一つだった領域が二つに分割されることになる。したがって、分割度を変えることで、分割領域には包含関係による階層構造ができることになり、階層の深さは分割度によって表されるといえる。以降では、分割度を大きくしていき、ある領域Aが領域BとCに分割されたときの分割度を領域Aの分割度と呼ぶことにする。

このような画像の領域分割における階層性を利用すると、領域分割手法の特性を反映した、対象

領域の「全体」と「部分」の関係をきわめて自然に抽出し、記述することができる。

## 2.2 対象物モデルの基本構造

前節で述べた領域の階層構造は、一般に二分木で表現することができる。領域間の包含関係による階層構造を表すために、図1に示すように、各領域を二分木の各ノードに対応付ける。このとき、木のルートノードは対象物の領域全体を表すノードであり、木の末端のノードはもうそれ以上分割されない領域を表す。以下、この木を分割木、ルートノードを初期ノード、末端のノードを終端ノードと呼ぶことにする。本手法における対象物モデルは、この分割木と特徴テーブルからなっている。特徴テーブルは、分割木の各ノードに対応する領域の特徴を保持するものであり、この特徴を基に領域間のマッチングを行う。

ここでは、領域分割の表現を木構造に限定したが、その利点としては以下のような点が挙げられる。

[1]木構造の場合、モデル更新は次節で述べるように枝刈を基本とした方法でできるが、木構造でない場合、このような単純な方法では実現でき

ず、計算量が多くなる。

[2]木構造の場合、子ノードどうしが隣接している場合は、その親ノードどうしも必ず隣接している。したがって、木構造の場合は、終端ノードどうしの隣接関係を記述しておけば、すべてのノードどうしの隣接関係がわかるという利点がある。

[3]木構造の方が、システムの作成が容易である。

## 2.3 特徴テーブル

特徴テーブルは分割木の各ノードに付加するものであり、そのノードの表す領域の形状および色の特徴を保持し、二つの分割木のノードどうしを比較する際に利用される。特徴としては、拡大縮小の影響を受けない不変性の高いものが望ましく、また、より正確な比較を行うために、なるべく多くの独立した特徴量を準備する必要がある。

ここでは、モデル生成時には大きさについての正規化がなされているという前提に立ち、現モデルが得られるまでの特徴量の分布をヒストグラムで保持する。ヒストグラムによって特徴量の分布を表現することによって、更新が容易で、しかも特徴量の分布からノードどうしに対応するかどうかを判断できる。尚、後に示す実験で用いた特徴

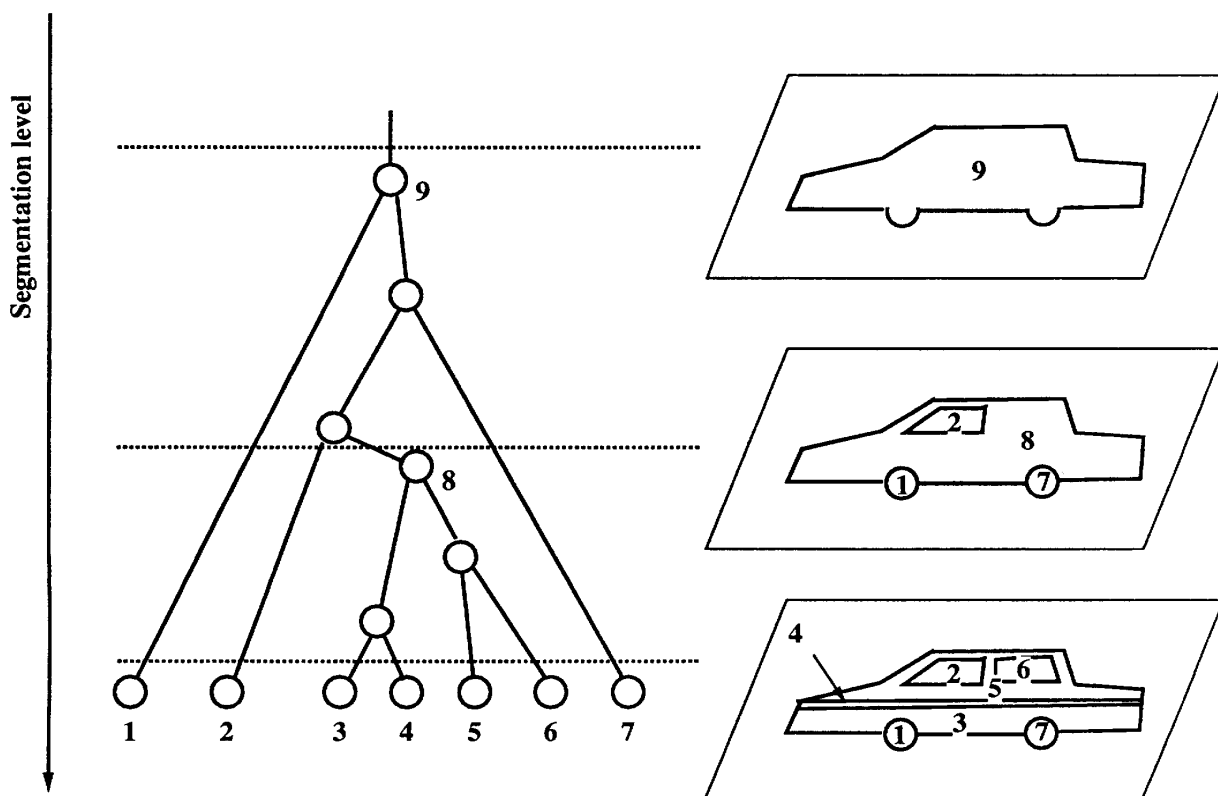


図1 分割木の概念

テーブルの内容は以下のようなものである。

面積，穴埋めした面積，周囲長，円形度，  
 主軸の方向，角の数，凸包の面積，  
 凸包の主軸の方向，凸包の角の数，  
 凸包の周囲長，RGB各成分の平均値，  
 重心位置，外接長方形の位置

尚，位置に関する特徴は，親ノードの外接長方形を1×1の正方形に正規化した領域を基準として表現する。これらの位置特徴を基に，領域の位置関係に関する制約を反映させることになる。

### 3. 認識システムの概要

前節で述べたモデルを用いることにより，画像認識システムの目指す以下の三つの機能を効率良く実現できる。

[対象物モデルの獲得] ユーザからの例示によって，対象物モデルを作成・更新する。

[対象物の探索] ユーザの指示する対象物を画像中から探す。

[対象物の認識] ユーザの指示する領域が何であることを認識する。

以下でそれぞれの機能がどのように実現されるかの概略について説明する。以下の記述では，対象物モデルのうち，データベース内に知識として蓄えられているものを知識モデル，入力画像から作成されたものを入力モデルと呼ぶことにする。

#### 3.1 対象物モデルの獲得

対象物モデル獲得処理の流れを図2に示す。まず，ユーザはシステムに対し画像を入力する。システムはこの画像を領域分割し，その結果をユーザに提示する。ユーザは分割された領域から対象物の名前とそれを構成する領域を選ぶ。システムは指定された領域から入力モデルを作成する。これをデータベースに格納してある知識モデルと比較し，新しい対象物モデルを作成する。これをデータベースに格納することによって，知識モデルを更新する。この手順を複数の画像に対して行うことによって，より正確な知識モデルが作成される。

モデルの比較更新の大まかな流れを以下に示

す。

[1]二つの木のノードの対応関係をトップダウン的に求める。このとき，対応のとれたノードについて，その子孫のノードどうしで，対応関係を求める（対応関係はノードの持つ特徴量から計算）。

[2]2つの木から対応関係がとれなかったノードを取り除く。それに応じてノード間の枝も付け直す。

[3]その結果，2つの木の構造は同じになる。

[4]各ノードの分割度，及び特徴量を更新する。

このようにして求めたモデルでは，対象物に不可欠で，しかも，利用している領域分割手法によって必ず抽出される領域を表すノードが残る。また，それぞれの特徴テーブルは，より多くの例題からヒストグラムを作成することになるので，そのノードのより一般的な特徴を表すことができるようになる。更新手続きについては，4.でさらに詳しく述べる。

#### 3.2 対象物の探索

対象物探索処理の流れを図3に示す。まず，ユーザはシステムに対し画像を入力する。システムはこの画像を領域分割し，画像全体を対象領域

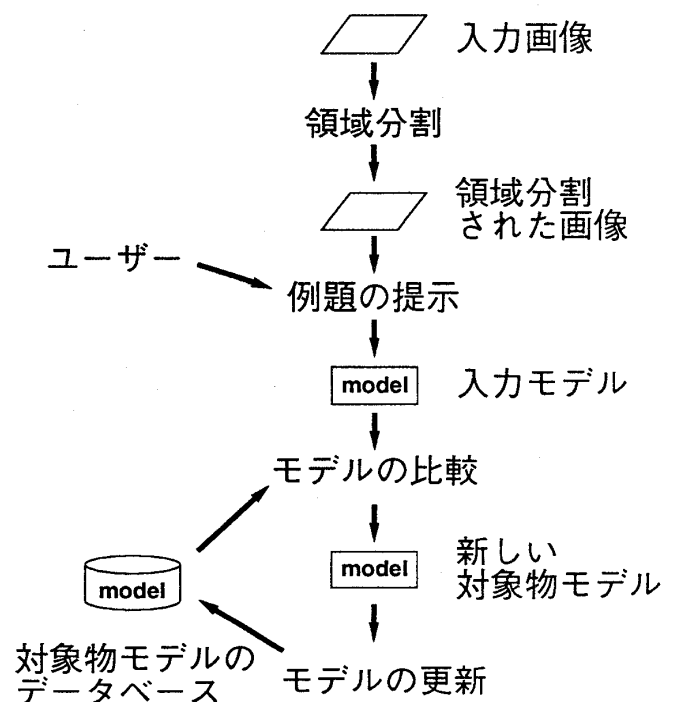


図2 モデル獲得の流れ

とする入力モデルを作成する。また、ユーザはシステムに対し探索したい対象物の名前を入力する。システムはこの名前の知識モデルをデータベースから検索する。この知識モデルを画像モデルの中から探索し、対象物の領域を出力する。

知識モデルの探索の大まかな流れは以下の通りである。

- [1]二つの木のノードの対応関係をトップダウン的に求める。このとき、対応のとれたノードについて、その子孫のノードどうしで、対応関係を求める（対応関係はノードの持つ特徴量から計算）。
- [2]すべての対応ノード対に対して、知識モデルの対応ノードの子ノードに対応するノードを入力モデルの対応ノードの子孫の中から探す。この処理を再帰的に行う。
- [3]対応する子孫のノードが見つからない対応ノード対は、不適切であるとして処理を終える。
- [4]知識モデルのすべてのノードに対応するノードが見つかったならば、その対応ノードが対象物である。

ここでは探索における分割木のノードどうしのマッチングの計算量が問題になる。しかし、本稿で提案した対象物モデルでは、知識モデルのノードの分割度が、その子ノードの領域が必ず存在する分割度を表しているのので、入力モデル側から対応するノードを探すときに、探索範囲を限定することができる（図4参照）。

### 3.3 対象物の認識

対象物認識処理における知識モデルと入力モデルの比較は前節の対象物探索のときとほぼ同様である。異なるのは、ユーザが分割された領域から認識させたい領域を選択し、システムは指定された領域から入力モデルを作成するという点である。これをデータベースに格納している全ての知識モデルと比較し、最も良く照合した対象物名を出力する。

## 4. 対象物モデルの更新

本章では対象物モデルの更新する方法について述べる。ここでは、知識モデルと入力モデルから作成される対象物モデルを更新モデルとよぶ。

更新モデルを作成するためには以下の処理が必要である。

- ・分割木のノード間の対応判定
- ・ノードの対応関係をもとにした分割木の更新
- ・それぞれのノードの特徴テーブルの更新

ここでは、これらについて詳しく述べる。

### 4.1 対応判定

ノード（換言すれば分割された領域）間の対応判定には知識モデルの特徴テーブルに格納されている特徴データを利用する。まず、特徴テーブルから各特徴についての評価関数を作成し、入力モデルのノードの持つ特徴量から確率を計算する。各特徴による確率からDempsterの結合則[5]により結合確率を求め、これからノードどうしの照合度を求める。各特徴から計算される基本確率は以下の通りである。

- ・肯定確率  $m(A)$

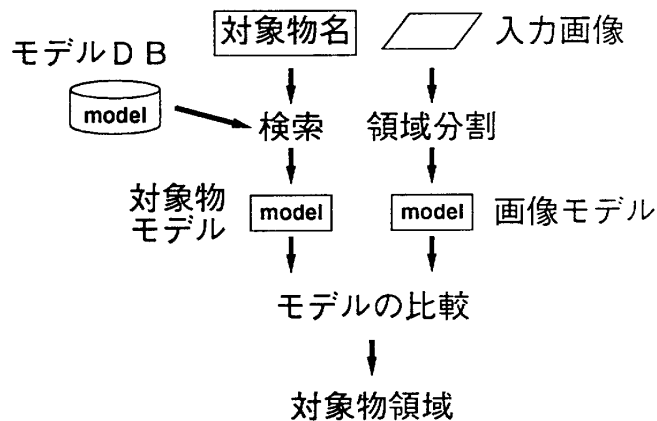


図3 対象物探索の流れ

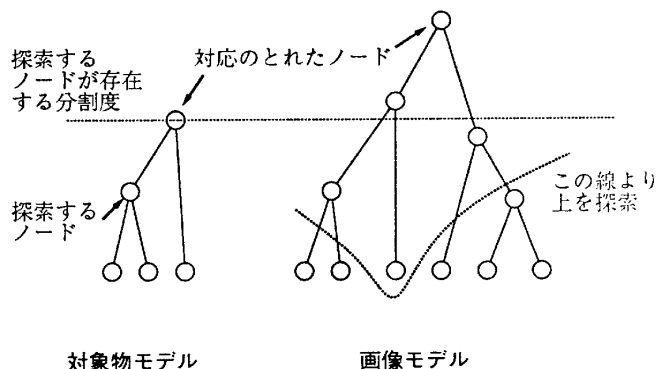


図4 探索範囲の限定

- ・ 否定確率  $m(\bar{A})$
- ・ 不定確率  $m(A \cup \bar{A})$
- ・  $m(A) + m(\bar{A}) + m(A \cup \bar{A}) = 1$

[ヒストグラムからの評価関数の作成]  
評価関数の条件として以下を与える。

- ・ 初期状態(ヒストグラムがすべて0)では  
 $m(A) = m(\bar{A}) = 0$
- ・ ヒストグラムのサンプル数が多い程、  
 $m(A \cup \bar{A})$ が小さくなる。

これらの条件を満たす関数として、以下を定義する。

$$m(A) = \frac{M\sqrt{H(x)}}{S}$$

$$m(\bar{A}) = \frac{\sqrt{N}}{S} - m(A)$$

$$m(A \cup \bar{A}) = 1 - m(A) - m(\bar{A})$$

$$= 1 - \frac{\sqrt{N}}{S}$$

ただし、各確率が、0以下になった場合は0、1以上になった場合は1とする。ここで、 $x$ はヒストグラムの区間、 $H(x)$ はヒストグラムの度数、 $N$ はサンプル数、 $M$ はヒストグラムの区間数、 $S$ はスケーリングのための定数である。

実際の照合度の計算は、各特徴から基本確率を求めた後、次式のDempsterの結合則によって、結合確率を求める。

[Dempsterの結合則]

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K}$$

ここで、

$$K = \sum_{B \cap C = \phi} m_1(B)m_2(C)$$

$$m_{1,2}(\phi) = 0$$

また、三つ以上の結合確率は、以下の式を帰納的に適用すれば求めることができる。

$$m_{1,2,\dots,k}(A) = m_{(1,2,\dots,k-1),k}(A)$$

ここで、 $2 \leq k \leq n$ 、 $n$ は特徴数である。

求められた結合確率から、ノード(領域)間の整合性を判定し、対応関係を定める。

#### 4.2 分割木の更新

分割木の更新の概略は3.で述べたとおりである。例えば図5(a)のような対応関係が取られた場合、図5(b)に示すノード以外のものが取り除かれ、最終的には図5(c)のような分割木が生成されることになる。

新しい分割木の生成に当たっては、対応ノードの二つの特徴テーブル統合して、新しい特徴テーブルを作成する。実際に行われる処理は以下の通りである。

- ・ 分割度の更新  
対応するノードの持つ分割度の大きい方を採用する。
- ・ ヒストグラムの更新  
各特徴に対して、入力モデルのノードの持つ特徴量をモデルのノードの持つヒストグラムに反映させる。
- ・ サンプル数の更新  
領域がマッチしたときだけ、サンプル数を1増やす。従って、ノードによりサンプル数は異なり、しかも例示画像の数とも異なっていることに注意。

#### 5. 予備実験

ここでは、本手法の妥当性を評価するために行った予備実験について簡単に述べる。まず、分割度を変化させることにより、画像の領域が全体から部分へと分割が進むことを示す。実験は、カラーの顔画像を対象として行った。対象画像のサイズは320×240画素である。

分割木の作成に用いた領域分割の手法は、画像の色空間上における類似性に基づいたものである[6]。実際に利用した類似性の尺度 $\Delta S_{pq}$ は以下のようである。

$$S_p = \sum_{j=1}^m \sum_{i=1}^{n_p} (x_{ji}^p - \bar{x}_j^p)^2$$

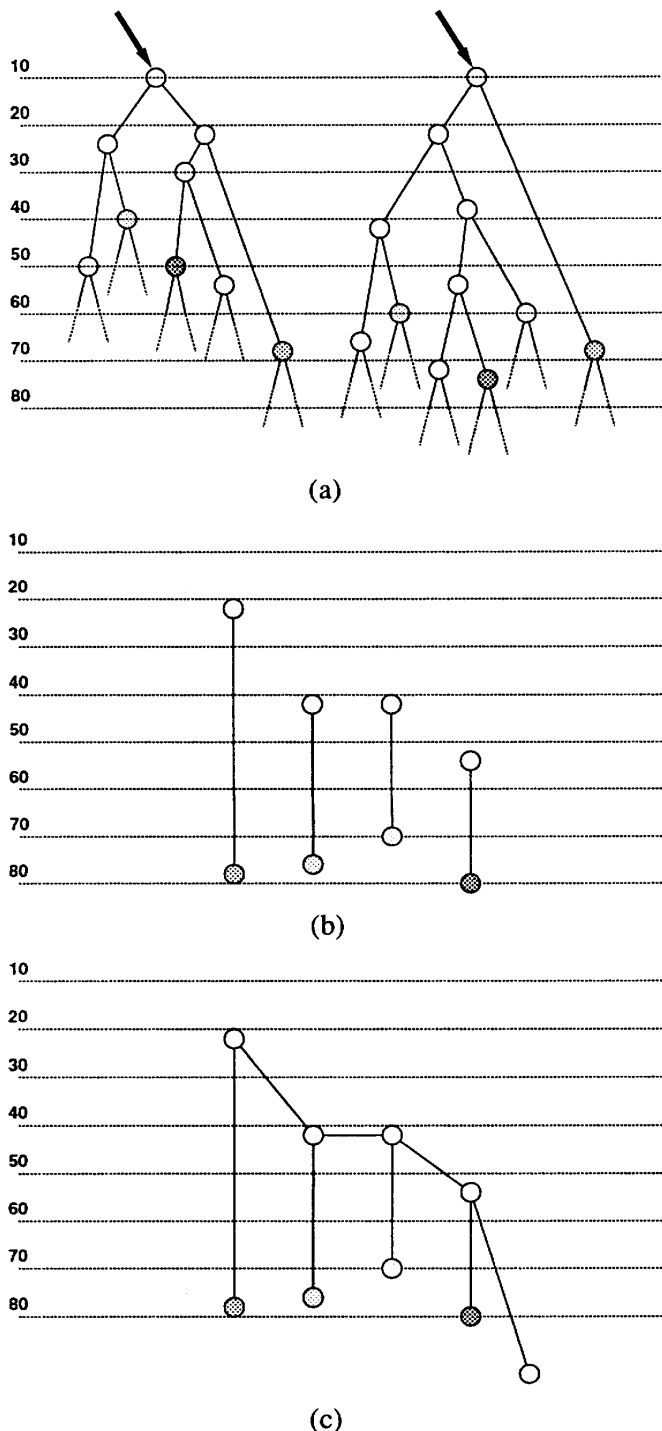


図5 分割木の更新

$$\Delta S_{pq} = \frac{n_p n_q}{n_p + n_q} \sum_{j=1}^m (\bar{x}_j^p - \bar{x}_j^q)^2$$

ここで、

$n_p, n_q$  はクラスタ (領域)  $p, q$  の画素数

$m$  は、画像のバンド数 (カラーの場合は3)

$j$  は、RGBの各画像に対応している

すなわち、クラスタ  $p, q$  の統合によるクラスタ

内偏差平方和の変化分であり、これが小さいものほど類似性が高いとした。

図6(a)が実験の対象画像であり、(b)がその分割木の上部である。この分割木全体には約3万個のノードが存在している (微小領域はほとんど意味がないため、ここでは64画素以上の領域しか特徴テーブルを持たせないようにしている)。 (b)の分割木中で横線で示す分割度で画像を領域分割した結果が、(d)、(e)、(f)である (分割度の小さい順に示している)。分割度を変化させることにより、画像の領域が全体から部分へと分割が進むことが分かる。

このような分割木を利用して、23個の顔画像を入力としてモデル獲得を行った結果、53個のノードを持つ分割木 (図(c)) が生成された。今回の実験では、10個程度の例示画像を与えた所でモデル木のノード削除が行われなくなり、安定してモデル作成が行われていることが判明した。

## 6. おわりに

本稿では、画像認識のための対象物モデル獲得の一手法について述べた。ここでは、ユーザが例題として対象物の名前およびその画像上での領域を入力することによって、システムがその対象物についての知識を獲得し、その知識を利用して対象物を探索・認識する方式を開発することを目指した。対象物モデルは領域分割をベースに記述されるが、領域の「全体」と「部分」を抽出するために、分割木概念を導入した。分割木は領域の分割され易さを木構造で表したものである。

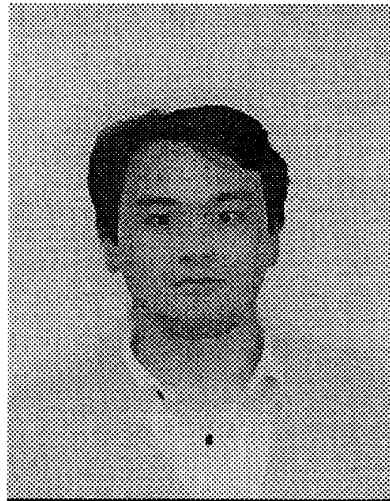
今回は、3次元物体を対象にしてはいるものの、その見え方はほぼ一定であるという制約を課している。従って、アスペクトを考慮して、様々な視点からの見え方を統合的に表現する方式、およびその場合のモデル獲得の手法を確立することが今後の課題である。

## 謝辞

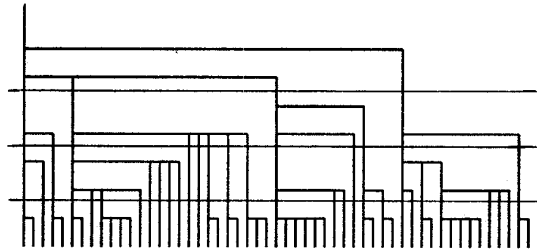
本研究は文部省科学研究費補助金 (一般研究(C): 05680300) 及び久留米鳥栖技術振興センターの援助を受けた。

## 参考文献

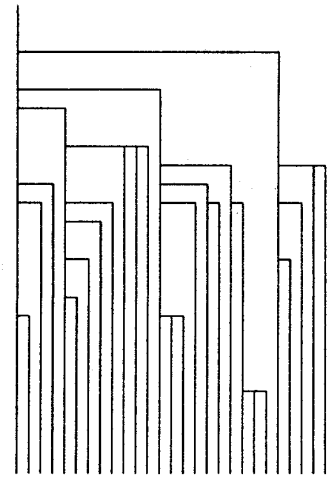
[1] J. H. Connell and M. Brady: Generating and Generalizing Models of Visual Objects, Artificial Intelli-



(a)



(b)



(c)



(d)



(e)



(f)

図6 実験結果

gence, Vol.31, pp.159-183, 1987.

[2]秋山, 荒井: 対象の幾何学的構造に関する知識を学習により記述するモデル駆動型画像理解システム, 電子情報通信学会技術研究報告, PRU93-29, pp.33-40, 1993.

[3]W. A. Harvey, M. Diamond, and D. M. McKeown, Jr.: Tools for acquiring spatial and functional knowledge in aerial image analysis, Proceedings of the DARPA Image Understanding Workshop, pp.857-873, 1992.

[4]村瀬, Nayar: 2次元照合による3次元物体の認識とその学習—パラメトリック固有空間表現—, 電子情報通信学会技術研究報告, PRU93-120, pp.31-38, 1994.

[5]G. Shafer: *A mathematical theory of evidence*, Princeton University Press, Princeton, 1976.

[6]後澤: 画像における階層的領域抽出, 第24回画像工学コンファレンス論文集, pp.391-3948, 1993.