

## 画像認識における階層的対象物モデルの獲得

有田, 大作  
九州大学システム情報科学研究院知能システム学部門

鶴田, 直之  
九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎  
九州大学システム情報科学研究院知能システム学部門

雨宮, 真人  
九州大学システム情報科学研究院知能システム学部門

<https://hdl.handle.net/2324/5671>

---

出版情報 : 映像情報メディア学会誌 : 映像情報メディア. 51 (8), pp.1240-1248, 1997-08-20. 映像情報メディア学会  
バージョン :  
権利関係 :



# 画像認識における階層的対象物モデルの獲得

有田 大作<sup>†</sup>, 鶴田 直之<sup>†</sup>,  
正会員 谷口 倫一郎<sup>†</sup>, 雨宮 真人<sup>†</sup>

## Generation of Hierarchical Object Models for Image Recognition

Daisaku Arita<sup>†</sup>, Naoyuki Tsuruta<sup>†</sup>,  
Rin-ichiro Taniguchi<sup>†</sup> and Makoto Amamiya<sup>†</sup>

**Abstract** An object recognition system is presented that generates object models from image examples. It also searches for and recognizes objects in images, using the object models. An object model is described by segmented regions and the hierarchical relations between them. The system generates one object model from one image example and compares the two to generate a more general object model. This paper describes a method for describing an object model and generating it from a series of image examples, and searching for an object in an image. Also shown are results of an experiment to verify proper system operation.

### 1. ま え が き

#### 1.1 背景

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あらかじ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、画像認識をさまざまな分野に利用しようとしたとき、ひとつひとつの対象物に関する知識を開発者が構築することは困難である。そこで、対象物に関する知識をシステムが自動的に獲得することが必要となる。

本稿では、この問題についての解決法を示し、知識の記述法、知識の獲得法、知識を利用した対象物探索

について述べる。

#### 1.2 目的

対象物に関する知識を自動的に獲得するシステムの実現の方法として、例題をシステムに与えることにより知識を獲得することが考えられる。そこで本研究では、ユーザが例題として「対象物の名前」および「画像上での対象物の位置」を入力することによって、システムがその対象物に関する知識を獲得し、その知識を利用して対象物の探索・認識を行う方式を開発することを目指す。

本稿では3次元物体が写っているカラー画像を対象物として、例として示される複数の画像（例示画像）から対象物に関する知識（対象物モデル）を獲得する。具体的には、1枚の例示画像から1つの対象物モデル

キーワード：画像認識、知識獲得、階層的対象物モデル

1997年1月21日受付、1997年4月14日再受付

†九州大学（〒816 春日市春日公園6-1, TEL 092-583-7618）

†Kyushu University (6-1, Kasuga-koen, Kasuga-shi, Fukuoka 816, Japan)

(これをデータモデルと呼ぶ)を作成し、複数のデータモデルを比較することでより一般的な対象物モデル(これを知識モデルと呼ぶ)を獲得する。つまり知識モデルは、対象物クラスに一般化されたものになる。また、この対象物モデルを利用して、カラー画像中から対象物を探索する。対象物クラスの対象物モデルを利用しているので、知識獲得時には与えていない個体についても探索可能である。例示画像は、対象物のみが存在している(実際は、画像中の対象物の領域をユーザが指定する)ことを仮定するが、探索時には対象物が切り出されていることを仮定しない。

### 1.3 他の手法との比較

他の手法との比較により本手法の特徴を明確にする。

#### (1) 領域分割を基に階層的な知識を獲得

Connellらの研究<sup>1)</sup>、秋山らの研究<sup>2)</sup>は、主に2次元の線画や輪郭線画像を対象としたものである。したがって、3次元物体の実画像に対しては、それらの線画を安定に抽出することが保証できず、適用し難い面がある。本研究は実画像を対象とし、領域分割における問題点を解決した階層的な対象物モデルを作成している。

#### (2) 例題画像から自動的に知識を獲得

HarveyらのSPAM<sup>3)</sup>の研究は、ユーザが対象物の構成部品の領域を画像上で教えることで、その構成部品に関する知識を獲得し、すべての構成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構成部品が、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、この条件を満たすような構成部品を選択することが非常に難しいという問題点がある。本研究では、ユーザの作業を最小限に抑え負担が少ないだけでなく、システムが構成部品を選択することによって、上記の問題がない構成部品を選ぶことが可能になる。

#### (3) 対象物クラスの知識を獲得

村瀬らの研究<sup>4)</sup>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、単純な方法で比較的良好な結果を得ている興味深いものである。しかし、これは個々の対象物の認識を目指すものであり、本研究のように対象物に関する知識を一般化し対象物クラスの認識を目指すものではない。

## 2. 対象物モデル

本研究における対象物モデルは、対象画像の領域分割によって得られる領域間の関係を表す分割木と、領域の特徴を表す特徴テーブルから構成されている。こ

こではそれらについて述べる。

### 2.1 分割木

一般に、領域分割を行うときには、領域内の画素をどの程度均一にするかを決めなければならない。このために、領域内の画素の均一性の閾値を設定することになる(この閾値を分割度と呼ぶことにする)。ある画像に対して領域分割を行うとき、分割度を大きく設定すると、領域内の画素の均一性が高い、つまり、画素値が非常に近いものから領域が成り立つことになり、小さな領域が生成される。反対に分割度を小さく設定すると、領域内の画素の均一性が低い、つまり、画素値がある程度離れたものまで1つの領域に含まれることになり、面積の大きな領域が生成される。このように、領域分割の結果は分割度の影響を受けてしまうため、対象物モデルを作成するのに適した分割度をどのように設定すれば良いかという難しい問題が生じる。

そこで、ある一定の分割度による領域分割結果を利用するのではなく、分割度による領域分割結果の変化を利用することを新たに提案する。図1の右のように、分割度を変化させたときの領域分割結果を比較してみると、分割度を徐々に小さくすることにより、それまで分かれていた領域が1つの領域に結合されていく(例えば領域2は領域4, 5, 6, 7が結合したものである)。これは、領域が複数のより小さな領域から成り立っていることを示す。つまり、領域にはこのよう

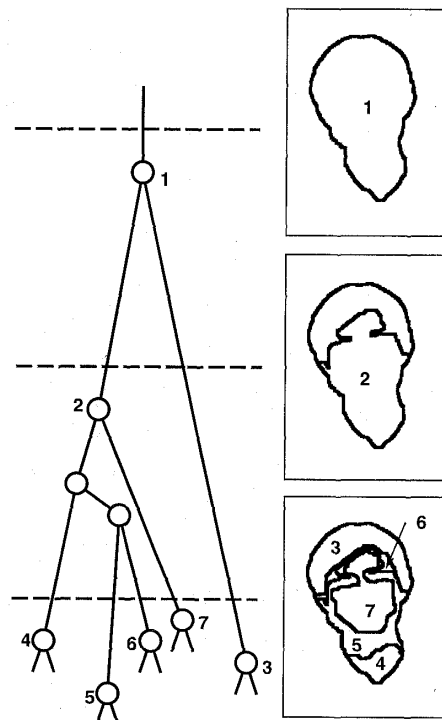


図1 分割木  
Segmentation tree.

な包含関係による階層構造があることがわかる。

本研究では、この領域の階層構造に、領域分割手法の特性を反映した対象物の本質的な情報が含まれていると考え、領域をノード、包含関係をリンクとする木(図1の左の木)によって、領域の階層構造を表現する。この木を分割木と呼び、これを利用して対象物モデルを構成する。

## 2.2 特徴テーブル

特徴テーブルは分割木の各ノードに付随し、そのノードに対応する領域の2次元的な特徴(色、大きさ、位置、形状に関する特徴)についてのデータを保持する。画像中の対象物の大きさは不定なので、拡大縮小の影響を受けない特徴である必要がある。

データモデルは1つの例題画像から作成されたものであり、その特徴テーブルは分割木の各ノードに対応するユニークな領域の特徴を表したものである。一方、知識モデルは複数の例題画像から得られたデータモデルを統合化したものであり、その特徴テーブルは、例題画像集合中で対応づけられた領域の持つ特徴をまとめたものであり、例題画像が与えられるたびに更新される。また、特徴テーブルは、2つの分割木のノードどうしを比較する際に利用される。したがって、各特徴データは比較および更新しやすい形式で保持される必要がある。

以下各特徴について説明する。

### (1) 色に関する特徴

色に関する特徴として、データモデルでは、領域内の画素についてのRGBそれぞれの平均値と分散値を用いる。一方、知識モデルでは、例題画像集合中で対応づけられた領域のもつ上記の特徴量の頻度分布(ヒストグラム)を特徴として用いる。ヒストグラムの階級数は、ここでは15にしている。

### (2) 位置・大きさに関する特徴

位置と大きさに関する特徴は、拡大縮小の影響を受けないようにするため、データモデルでは、対象物全体の領域の外接矩形に対する、その領域の外接矩形の相対的位置(上辺、下辺、左辺、右辺)と大きさ(幅、高さ)を用いる。一方、知識モデルでは、例題画像集合中で対応づけられた領域の位置・大きさの特徴データを、すべてリストにして保持している。色の場合と異なり、ヒストグラムを作成しないのは、対象物探索処理における対象物全体の領域の位置の推定のためである。これについては5章で説明する。また、対象物モデル獲得時には、これらの特徴はヒストグラム形式に変換され、色情報と同様に処理される。この変換は、対応づけられた領域の位置・大きさの情報がすべて

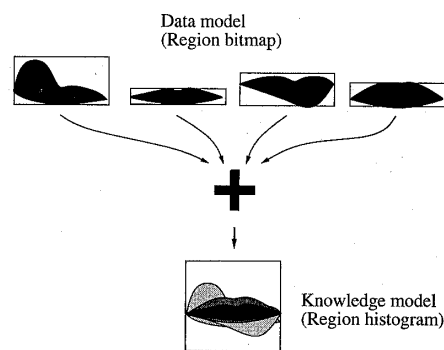


図2 領域ヒストグラム  
Region histogram.

ストに保持されているので、簡単に頻度分布が計算できる。

### (3) 形に関する特徴

形に関する特徴(形状特徴)は、データモデルでは、ノードに対応する領域内の画素を1、外の画素を0で示す2次元のビットパターン(領域ビットマップと呼ぶ)で表現している。ただし、画像全体の画素について領域内/外を示すわけではなく、領域の外接矩形中の画素についてのみ1/0のパターンで表現する。また、形状のみに注目するために、領域ビットマップの外接矩形の長辺が一定の長さになるように拡大縮小することによって、領域ビットマップの大きさを正規化している。

知識モデルにおける形状特徴は、図2に示すように、例題画像集合中で対応づけられた領域の領域ビットマップを重ねあわせた2次元の頻度分布として表す。これを領域ヒストグラムと呼ぶ。領域ヒストグラムは正方形であり、辺の長さは領域ビットマップの長辺と同じである。領域ビットマップと領域ヒストグラムの重ね合わせの位置は、4.2節のmatchの値が得られるときの位置、つまり、最もよく重なり合うと考えられる位置とする。

## 3. 対象物モデルの獲得

対象物モデル獲得処理は、ユーザがある対象物クラスについての例題画像を複数与えることにより、その対象物クラスの対象物モデルを獲得する処理である。このとき、複数の例題画像を同時に与える必要はなく、ユーザが例題画像を与えるたびに、対象物モデルは、対象物クラスのものより一般的なものに更新されていく。

対象物モデル獲得処理の手順を説明する。

- (1) システムはユーザから与えられた画像全体を領域分割し、画像全体についての分割木を作成する。

- (2) (1) で得られた分割木からユーザが対象物内部の領域を必要なだけ選択することで、対象物全体の領域を指定する。
- (3) 対象物内部の各領域に対応するノードの特徴テーブルを作成し、それ以外のノードを削除することで、対象物領域についてのデータモデルを作成する。
- (4) データモデルと知識モデルの間でノードマッチングを行い、2つの分割木のノード間での対応をとる。
- (5) 知識モデルを更新する。

この手順を複数の例示画像に対して行い、知識モデルを一般化する。ただし、最初の例題画像を与えたときには、知識モデルが存在しないので、手順(3)で作成したデータモデルがそのまま知識モデルとなる。以下、手順(1)(4)(5)について詳しく説明する。

### 3.1 領域分割

画像全体を画素結合法<sup>5)</sup>により領域分割する。画素結合法は、(1) 各画素を1つの領域と考え、(2) 隣接する領域のうち最も類似度の大きいものから順に結合し、(3) ある閾値よりも大きな類似度の隣接領域がなくなると処理を終了、という領域分割手法である。このとき、類似度の閾値をを非常に小さく(実際は0に)設定すると、処理終了時には領域が1つになる。この領域の結合の過程を木で表すことにより、分割木が得られる。また、結合した2つの領域の類似度が、生成された領域の分割度になる。

### 3.2 ノードマッチング

ノードマッチングとは、知識モデルの分割木の各ノードに対して、データモデルの分割木の中から対応するノードを探索することである。このとき、問題になるのは、「ノードの特徴の対応」と「分割木全体の整合性」を考慮して、どのように最適解を探すかということである。

#### ・ノードの特徴の対応

知識モデルの各ノードについて、特徴が最も合致しているノードをデータモデル中から選び、それらに対応させる。特徴が合致しているかどうかを判断するために、照合度を利用する(4章参照)。

#### ・分割木全体の整合性

ノードの対応が分割木における親子関係と整合しているかをチェックする。実際には、ノードの特徴の合致により対応したノード対について、分割木の構造と矛盾が生じている場合は、照合度のより大きいノード対を優先させ、そうでないノード対を破棄することで解決する。対応が破棄されたノードは次に照合度の大きい

ノードとの対応を試みる。

知識モデルの各ノードが「データモデルのノードと矛盾なく対応している」か「矛盾なく対応するデータモデルのノードがない」となったとき、ノードマッチングを終了する。

### 3.3 知識モデルの更新

知識モデルの対応がとれたノードについては特徴テーブルを更新し、対応がとれなかったノードについてはそのノードを削除し、上位と下位のノードを接続する。

このように、知識モデルの更新では、対応のとれたノードのみを残しているのので、知識モデルは、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\*から構成されることになる。また、特徴テーブルは、より多くの例題から作成されることになるので、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。

## 4. 照 合 度

照合度とは、知識モデルとデータモデルのノード対の特徴がどの程度合致しているかを表す度合である。照合度は、ノード対の特徴テーブルどうしを比較することによって求められる。まず、各特徴から2つのノードが対応するかどうかを表す基本確率を計算する。この基本確率は Dempster & Shafer の確率則<sup>6)</sup>における確率であり、以下の通りである。

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

無知確率は「その証拠からは  $A$  とも  $\bar{A}$  ともいえない」確率を表す。したがって、本システムにおける無知確率は「その特徴からはノード対が対応するかどうか判断できない」確率を表すことになる。つまり、ある特徴がそのノードを特徴付けるものである場合、言い替えるとそのノードであるかどうかを判断するために有効な特徴である場合に、無知確率は小さくなり、肯定か否定かをはっきりと判断することになる。

つぎに、すべての特徴の基本確率から Dempster の結合則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計算する。

以下、特徴の形式によって異なる2種類の基本確率の求めかたと、それからの照合度の求めかたについて説明する。

\* これらの領域は、ここで定めた尺度に基づいて選択されたものであり、人間の直観に合わないように見えることもある。

#### 4.1 ヒストグラムからの基本確率の計算

知識モデルの特徴テーブルから各特徴についての評価関数を作成し、データモデルのノードの各特徴データを代入することで基本確率を求める。

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  $m(A) = m(\bar{A}) = 0$ 。

・与えた例題画像数が多いほど、 $m(A \cup \bar{A})$  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。

$$m(A) = \frac{c}{Ms} h(x) \quad (1)$$

$$m(\bar{A}) = \begin{cases} \frac{cN}{Ms} & (h(x) = 0 \text{ のとき}) \\ 0 & (h(x) \neq 0 \text{ のとき}) \end{cases} \quad (2)$$

$$m(A \cup \bar{A}) = 1 - m(A) - m(\bar{A}) \quad (3)$$

ただし、各確率が、1以上になった場合は1とする。ここで、 $h(x)$  は階級  $x$  でのヒストグラムの度数、 $N$  は与えた例題画像数、 $M$  はヒストグラムの階級数 (今回の実験では  $M = 15$ )、 $c$  は  $h(x) = 0$  である階級の数、 $s$  はスケーリングのための定数 (今回の実験では  $s = 20$ ) である。

これらの式について説明する。 $c/M$  が大きくなると、 $m(A)$  と  $m(\bar{A})$  の値が大きくなり、 $m(A \cup \bar{A})$  の値は小さくなる。これは分布に偏りがあるときには、その特徴がノード対応判定に有効であると考えられるからである。また、それまでその特徴が階級  $x$  の値をとったことがない場合に、 $m(\bar{A})$  は正の値をとる。これは  $N$  の値が大きくなるほど意味が重くなるので、 $m(\bar{A})$  を  $N$  に比例させている。

#### 4.2 形状特徴からの基本確率の計算

領域ヒストグラムは、領域ビットマップを加算したものである。よって、形状が不変なノードでは加算される部分と、加算されない部分に分かれることになり、領域ヒストグラムの値は両極端に分かれることになる。したがって、領域ヒストグラムを画像とみなしたときの濃度ヒストグラムを作成すると、形状が不変なノードでは、濃度の分布が両端に偏っている。そこで、この濃度ヒストグラムの分散を基に、形状特徴の重要性を表す状態指数 (**Cond**) を (4) 式で求める。

$$\mathbf{Cond} = \frac{\mathit{Var}}{\mathit{Mean}(N - \mathit{Mean})} \quad (4)$$

ただし、それぞれ領域ヒストグラムの濃度分布の、 $\mathit{Var}$ : 分散値、 $\mathit{Mean}$ : 平均値であり、 $N$  は与えた例題画像数である。(4) 式の分母は、平均値が  $\mathit{Mean}$  のときの  $\mathit{Var}$  の最大値である。このため、**Cond** の値

は、領域ヒストグラムの濃度分布が平均値付近に集中しているときには0、逆に両端に偏っているときには1に近くなる。

これから、無知確率  $m(A \cup \bar{A})$  を (5) 式で求める。

$$m(A \cup \bar{A}) = 1 - \mathbf{Cond} \left(1 - \frac{1}{N^2 + 1}\right) \quad (5)$$

ここで、例題画像数が少ないと、知識モデルの情報も信用性が低いので、無知確率を上げるために、重要性を表す **Cond** を例題画像数によって調整している。

次に、 $m(\bar{A})$  と  $m(A)$  を求めるために、**match** という、領域ヒストグラムと領域ビットマップの一致度を表す値を、(6) 式によって計算する。領域ビットマップは領域ヒストグラムよりも小さいので、領域ビットマップを短辺方向にずらしながら **match** を求め、最大の値を採用する。

$$\mathbf{match} = \max_i \sum_x \sum_y W(x, y, i) \quad (6)$$

ただし、 $x$  方向にずらすときは、

$$W(x, y, i) = \begin{cases} 2 \times \mathit{Hist}(i + x, y) - N & (\mathit{Map}(x, y) = 1 \text{ のとき}) \\ -(2 \times \mathit{Hist}(i + x, y) - N) & (\mathit{Map}(x, y) = 0 \text{ のとき}) \end{cases}$$

ここで、 $\mathit{Map}$ : データモデルの領域ビットマップ、 $\mathit{Hist}$ : 知識モデルの領域ヒストグラムである。この式において、領域ヒストグラムの度数の高い部分と領域ビットマップの1の部分的重なるところ、および、度数の低い部分と0の部分的重なるところでは、 $W(x, y, i)$  は正になり、それ以外では負になる。

そして、(7)、(8) 式で、 $m(\bar{A})$  と  $m(A)$  を求める。

$$m(\bar{A}) = \frac{N \times \mathit{HistSize}^2 - \mathbf{match}}{2 \times N \times \mathit{HistSize}^2} \times (1 - m(A \cup \bar{A})) \quad (7)$$

$$m(A) = 1 - m(\bar{A}) - m(A \cup \bar{A}) \quad (8)$$

(7) 式の分母は分子のとり得る最大値であり、その分子の第1項は **match** のとり得る最大値である。したがって、領域ヒストグラムの **Cond** が大きくて、例題画像数も多く、そして、データモデルの領域画像と領域ヒストグラムが非常に似ていた場合に、 $m(\bar{A})$  は0に近づくことになる。

図3に示す形状特徴 (黒の部分が0を示す) に対して、基本確率を計算したものを表1に示す。領域ヒストグラム **H1**, **H2** はどちらも15枚の例題画像から作成したものであり、これらと、ある画像のデータモデ

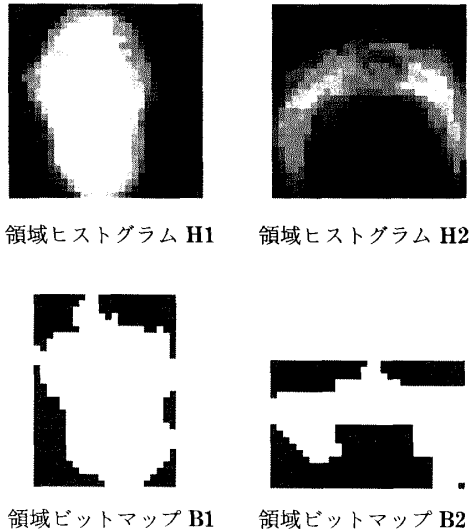


図3 形状特徴の例  
Examples of shape feature.

表1 形状特徴からの基本確率の例  
Examples of basic probability assignment  
calculated from shape feature.

	H1-B1	H1-B2	H2-B1	H2-B2
<i>Var</i>	38.3	38.3	19.1	19.1
<i>cond</i>	0.686	0.686	0.397	0.397
<i>match</i>	9646	1482	1604	6936
$m(A)$	0.584	0.378	0.221	0.299
$m(\bar{A})$	0.097	0.303	0.174	0.096
$m(A \cup \bar{A})$	0.319	0.319	0.605	0.605

ルの領域ビットマップ B1, B2 が対応づけられる基本確率を求めたものが表 1 である。H1 との組合せについては、無知確率が小さくなっており、H2 との組合せは、無知確率が大きくなっている。これは、H1 では、領域の形が一定であるのに対し、H2 では領域の形が一定ではないことによる。また、対応づけられるべき組合せである H1-B1 の肯定確率は大きくなっており、その他の組合せについては肯定確率が小さくなっていることもわかる。

#### 4.3 基本確率から照合度の計算

各特徴から確率を求めた後、Dempster の結合則によって、結合確率を求める。基本確率  $m_1$  と  $m_2$  の結合確率  $m_{1,2}$  を求める Dempster の結合則は (9) 式で与えられる。

$$m_{1,2}(z) = \frac{\sum_{x \cap y = z} m_1(x)m_2(y)}{1 - \sum_{x \cap y = \emptyset} m_1(x)m_2(y)} \quad (9)$$

ここで、 $x, y, z = \{A, \bar{A}, A \cup \bar{A}\}$ ,  $m_{1,2}(\emptyset) = 0$  である。

また、3 つ以上基本確率の結合確率は、(10) 式を帰納的に適用すれば求めることができる。

$$m_{1,\dots,i}(x) = m_{(1,\dots,i-1),i}(x) \quad (10)$$

ここで、 $2 \leq i \leq n$ ,  $n$  は特徴数である。

求められた結合確率から、照合度 *MATCH* を (11) 式を用いて計算する。

$$MATCH = m_{1,\dots,n}(A) + m_{1,\dots,n}(A \cup \bar{A}) \quad (11)$$

## 5. 対象物の探索

対象物探索処理は、ユーザが与えて画像からユーザが指示した対象物の領域を抜き出す処理であり、以下の手順で行われる。

- (1) ユーザはシステムに対し画像と対象物名を入力する。
- (2) システムは画像全体を領域分割し、画像全体についての分割木、および、それぞれのノードの特徴テーブルを作成する。これをデータモデルとする。
- (3) データモデルと知識モデルの間でノードマッチングを行い、2 つの分割木のノード間での対応をとる。
- (4) 探索された対象物の領域を出力する。

以下、手順 (3), (4) について説明する。

### 5.1 ノードマッチング

対象物モデル獲得処理と同様に、データモデルと知識モデルをノード間の照合度と分割木の整合性を基に対応させる。異なっているのは、対象物全体の領域が指定されていないために、照合度計算における位置と大きさに関する特徴が利用できないことであり、以下の手順によって照合度を求める。

- (1) 知識モデルとデータモデル間のノード対のすべての組合せについて、以下の処理を行う。
  - (a) 位置・大きさ特徴抜きでの照合度を計算する。
  - (b) 知識モデルのノードの持っている対象物全体の領域との相対位置情報 (2.2 節 (2)) と、データモデルのノードに対応する領域の位置から、探索画像中の対象物全体の領域の位置を推定する。この位置を推定全体領域位置と呼び、4 つのパラメータ (対象物全体の領域の外接矩形の上辺, 下辺, 左辺, 右辺の位置) によって表される\*1。

(c) 推定全体領域位置を4次元の空間に投票する。このとき、(1a)で求めたノード対の照合度を投票の重みとする。

この結果、全体領域の可能性の高い位置ほど多くの票を得ることになり、この得票数が、ある位置が領域全体の位置である可能性がどの程度あるかを示すことになる。

(2) 各ノード対の照合度を以下のように計算する。

- (a) そのノード対による推定全体領域位置での得票数<sup>\*2</sup>を調べる。これは、そのノード対が対応したと考えたときに推定される対象物全体の領域の位置・大きさの妥当性を示すものであり、したがって、そのノード対の位置・大きさ特徴の合致の程度を示すことになる。
- (b) (1a)で求めた照合度と、上で求めた得票数を積算することにより、最終的な照合度を求める。

## 5.2 探索された対象物の領域の出力

画像中に対象物が存在している場合には、対象物を表す分割木が、データモデルの分割木の部分木となっている。したがって、知識モデルに含まれているノードは必須のものだけであるから、それらのすべてがデータモデルの分割木の部分木のノードと対応がとれたときに、指定された対象物が探索されたことになる。このとき、この部分木のルートノードの表す領域が、対象物全体の領域として出力される。

## 6. 実 験

### 6.1 対象物モデル獲得実験

提案した方法による対象物モデル獲得実験を行った。実験では、例題画像として15枚の人間の頭部の写った実画像を用いた(図4, カラー, 120×160画素)。対象物モデルの分割木のノード数は、データモデルでは60から80個程度であり、知識モデルでは例題を与えるにつれて図5のように減少し、15枚の例題を与えた時点で6個になった。これは、この6個のノードがすべてのデータモデルに存在したことを意味し、これらに対応する領域が対象物に必須の領域であるということを示している。

このときの対象物モデルの分割木とその各ノードの領域ヒストグラムを図6に示す。そのノードの番号

\*1 知識モデルのノードの持つ相対位置情報は、知識獲得時与えられた例題の数だけ存在している。したがって、(1b) (1c)における推定全体領域位置はその数だけ存在することになる。

\*2 実際には複数の推定全体領域位置があるので、それぞれの得票数の平均をとる。

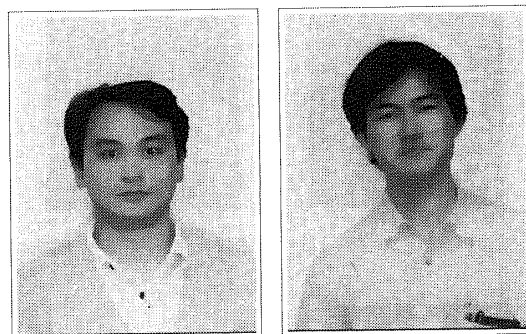


図4 例題画像  
Sample images.

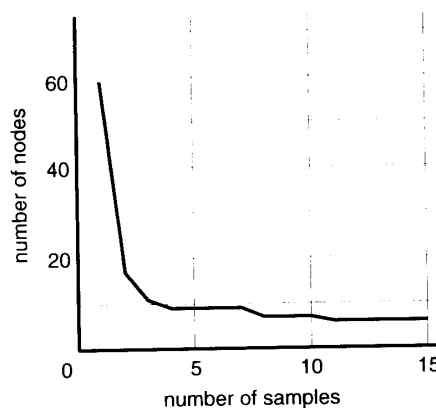


図5 分割木のノード数  
Number of nodes in segmentation tree.

は領域ヒストグラムの番号と対応している。ノード1は顔全体、ノード2は顔全体、ノード3は顔の中心部、ノード4は顔の周辺部、ノード5は髪、ノード6は髪の一部(照明によって光っていない部分)を表している。

### 6.2 対象物探索実験

前節で獲得した対象物モデルを利用し、探索画像から人間の頭部の領域を抜き出す実験を行った。実験では、図7を探索画像として与えた。これらの画像は対象物モデル獲得の例題としては与えていない。照合度計算のための対象物全体の領域の位置についての投票結果が図8である。濃い部分ほど多くの票を集めたことを示す。これを利用して照合度を求めノードマッチングを行った結果、抜き出された領域が図9である。

### 6.3 考 察

獲得した対象物モデルのノード数は、例題画像を与えるごとに減少していったが、11個目の例題画像で6個になってからは減少しなくなった。これより、人間の頭部に必須である6個の領域から成る対象物モデルが獲得されたと考えられる。

また、対象物探索の結果は良好であった。しかし、

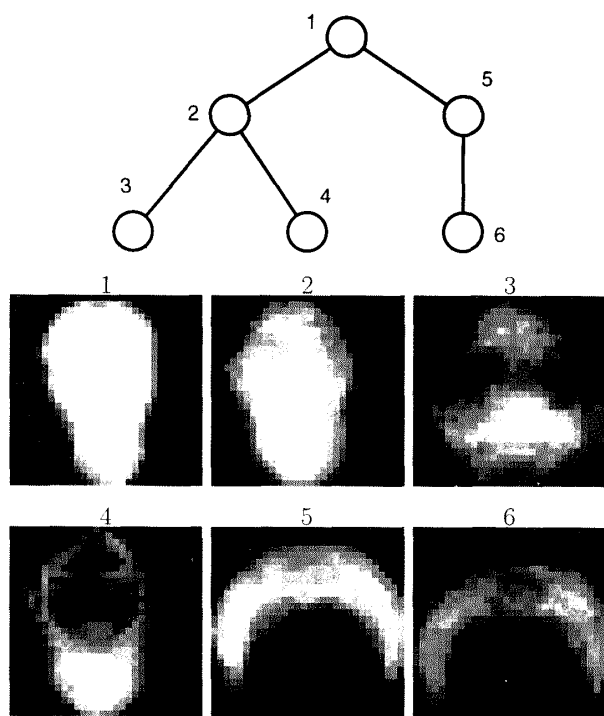


図6 獲得された分割木と領域ヒストグラム  
Generated segmentation tree and region histograms.

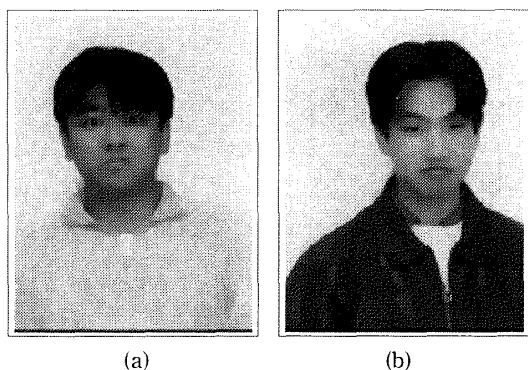
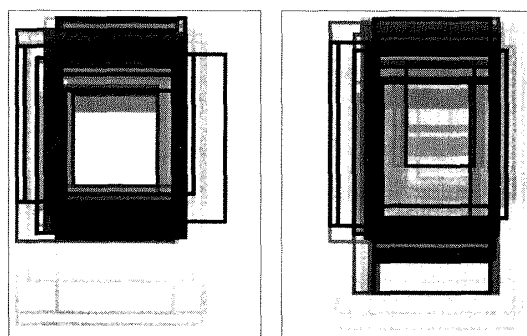


図7 探索画像  
Search image.

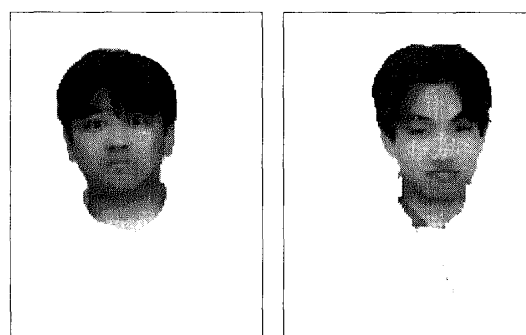
対象物の背景による領域分割誤りを考慮していないために、(b)の結果のように背景の一部を含んだ領域が出力されたり、探索がうまくいかない場合があった。

獲得した対象物モデル中に、目や口の領域がないのは、ある程度よりも小さな領域は、データモデルを作成する段階から削除しているからである。これは、小さい領域を扱うために生じる以下のような問題を避けるための処置である。

- ・デジタル画像において、小さすぎる領域は形状特徴が不正確になる。また、画像を領域分割し分割木を作成するとき、小さい領域ほど多く生成される。これにより、小さくて似たような（特徴テーブルの内容



(a) (b)  
図8 対象物全体の領域の位置  
Position of whole region of object.



(a) (b)  
図9 探索結果  
Search result.

が似ている)領域が多数存在することによる誤対応の問題が生じる。

- ・領域数の増大による処理量の増加の問題が生じる。目や口の領域を含む対象物モデルを獲得できるようにするためには、画像の解像度を上げ、それらの領域の面積を大きくするか、より小さな領域についても精度の高い特徴を利用することが必要である。

## 7. 本手法の有効性

ここでは、本手法の有効性について考える。本手法では、対象物の2次元的な「見た目」を基に対象物モデルを作成している。ここでいう見た目とは、色による領域分割の結果得られる分割木の構造と各ノードの持つ特徴量ということであり、1つの対象物クラスではこれらに大きな変化がないということが、本手法の前提となっている。本論文で示した実験では、簡単なRGBの強さの均一性を基にした領域分割の手法により分割木を作成したため、このままでは光源等の環境が大幅に変化したり、鏡面反射成分が強く現れている例題に対してはうまく動かない可能性がある。しかし、

それらの問題に対応した手法をベースに領域分割を行えば、基本的な枠組みはこのままで対応できるものと考えている。また、通常のカメラからの入力だけでなく、レンジセンサからの入力を用いた結果についても同じようなことが可能であり、その場合は、対象物の表面形状の類似性に基づいて知識モデルが作成されることになる。

この手法の一番の問題点は、3次元的な情報を基にしていないので、視点が大幅に異なれば、同じような分割木が構成されず、知識モデルが生成できないという点である。また、機能的には同じことを果たすが、見かけが大幅に異なるような抽象度の高い対象物クラス（例えば水道の蛇口）についてもユニークな知識モデルを構成することができない。これらの問題に対しては、1つの対象物クラスに対して複数の対象物モデルを保持するような仕掛けを構築することによって対処できると考えているが、その点は今後の課題といえよう。

## 8. む す び

本稿では、画像認識に必要な対象物モデルを例題画像から自動的に獲得する方法について述べた。また、獲得した対象物モデルに基づいて対象物探索が行えることを示した。

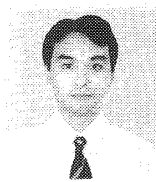
今後は、より正確で小さい領域にも対応できるように照合度計算法を改良し、また、背景の影響や隠蔽による領域分割誤りを考慮するように対象物探索を改良する予定である。

### [参 考 文 献]

- 1) J.H. Connell, M. Brady: "Generating and Generalizing Models of Visual Objects", Artificial Intelligence, vol.31,

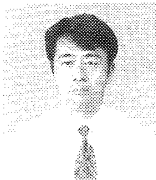
pp.159-183 (1987)

- 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", 信学技報, PRU93-29, pp.33-40 (1993)
- 3) W. Harvey, M. Diamond, D. Mckeown Jr.: "Tools for Acquiring Spatial and Functional Knowledge in Aerial Image Analysis", Proc. of the Darpa Image Understanding Workshop, pp.857-873 (1992)
- 4) 村瀬, S. 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—", 信学技報, PRU93-120, pp.31-38 (1994)
- 5) 高木, 下田編: "画像解析ハンドブック", pp.692-693, 東京大学出版会 (1991)
- 6) G. Shafer: "A Mathematical Theory of Evidence", Princeton University Press (1976)



ありた だいさく

**有田 大作** 1992年, 京都大学工学部情報工学科卒業。1994年, 九州大学大学院総合理工学研究科情報システム学専攻修士課程修了。現在, 同大学院システム情報科学研究科博士後期課程在学中。工学修士。



つるた なおゆき

**鶴田 直之** 1988年, 九州大学工学部情報工学科卒業。1990年, 同大学院総合理工学研究科修士課程修了。同年同助手。1996年, 同大学院システム情報科学研究科助手。工学修士。



たにぐちりんいちろう

**谷口 倫一郎** 1978年, 九州大学工学部情報工学科卒業。1980年, 同大学院工学研究科修士課程修了。同年, 同大学院総合理工学研究科助手。1988年, 同助教授。1996年, 同大学院システム情報科学研究科教授。工学博士。正会員。



あまみや まこと

**雨宮 真人** 1969年, 九州大学大学院工学研究科修士課程修了。同年, NTT電気通信研究所入所。現在, 九州大学大学院システム情報科学研究科教授。工学博士。