

2次元照合による3次元物体認識——パラメトリック固有空間法——

正員 村瀬 洋<sup>†</sup> 非会員 シュリー ナイヤー<sup>††</sup>

3D Object Recognition from Appearance——Parametric Eigenspace Method——

Hiroshi MURASE<sup>†</sup>, *Member* and Shree K. NAYAR<sup>††</sup>, *Nonmember*

あらまし 2次元画像から3次元物体を識別し、その物体の向きを検出する技術は、工業部品の分類など実用的な価値は高い。従来の代表的な手法としては、画像からエッジや表面形状などの3次元構造に着目する方法があるが、この枠組みでは3次元構造の抽出自体が困難であり、任意形状の物体に対して高い精度は得られていない。本論文では、2次元照合により3次元物体を認識する手法について述べる。2次元照合による手法は、3次元特徴の抽出が不要である等の特長はあるが、見る方向や光源の位置により複雑に変化する2次元画像をあらかじめ学習しておくことが記憶容量、計算量の点で困難であると考えられ、従来試みられていなかった。本手法では画像符号化を基本としたパラメトリック固有空間法の提案により、少ない記憶容量で3次元物体を2次元画像の集合体として学習することができるようになった。その結果、2次元画像例から物体を容易に学習すること、および困難な特徴抽出なしで3次元物体の認識とポーズ推定をすることが可能となった。本論文では、手法の提案と共に他の2次元照合的な手法との比較実験結果についても述べる。

キーワード 物体認識, ポーズ推定, 主成分分析, 固有ベクトル, 多様体, 視覚学習

1. まえがき

2次元画像から3次元物体を識別し、その物体の向きを検出する3次元物体認識技術は工業部品の分類や、一般環境内での移動物体の監視など幅広い応用があり、これまで多数の研究がなされてきた<sup>(1),(2)</sup>。従来の3次元物体認識は大別すると、モデルと入力特徴との照合に3次元構造を利用する手法と2次元照合を利用する手法に分類される。しかし、従来の研究の多くは3次元構造<sup>(2)</sup>に着目したものであり、2次元照合に着目した研究はこれまであまりなかった。その理由は、3次元物体は物体の向きや照明の方向によって見掛けの画像はさまざまに変形し、単純な2次元画像の照合では3次元物体はうまく認識できないと考えられてきたためである。1980年前後に提案されたMarrの3次元表現に関する魅力的なフレームワーク<sup>(3)</sup>は、更にそのアプローチに拍車をかけた。そのため、近年の物体認識手

法の多くは、2次元画像からいったん、影や動き、エッジなどの情報を利用してまず表面形状などの3次元構造を復元してから、それとあらかじめ記憶してある3次元モデルと照合をとる手法となっている。一方、2次元照合に基づく手法としてはエッジ位置等の特徴を利用するもの<sup>(4)</sup>、屈折点や端点の位置を利用するもの<sup>(5)</sup>はあるものの、これらは特徴点間の照合であり、2次元パターンを積極的に照合に利用するものではなかった。本論文では画像信号レベルの2次元照合により3次元物体を認識する手法について述べる。

ここでは、3次元物体の向きや光源の変化に対応して連続的に変動する2次元画像の変化を、画像の固有ベクトルから構成される部分空間(固有空間)上での多様体で表現するパラメトリック固有空間法を提案した。学習段階では物体の画像集合から、固有空間を計算し、その上で多様体を構成する。認識段階では入力画像をいったんこの固有空間上の点に投影し、その点に最も近い多様体上での位置を検出することにより、その物体の種類と物体の向きなどを検出する。本手法では、物体の例を与えるだけで自動的にその物体を学習することができる。また、認識段階では、入力画像中の3

<sup>†</sup> NTT 基礎研究所, 厚木市  
NTT Basic Research Laboratories, Atsugi-shi, 243-01 Japan  
<sup>††</sup> コロンビア大学, 米国  
Columbia University, New York, NY10027, USA

次元物体を認識すると同時に、その物体のポーズを検出することも可能である。

本手法はパターン認識における部分空間法<sup>(6),(7)</sup>と関係が深い。従来、画素値の固有ベクトルを認識へ応用した例としては、投影法や部分空間法による文字認識手法<sup>(8),(9)</sup>、あるいはアイゲンフェースによる顔画像認識手法<sup>(13),(14)</sup>などが挙げられる。しかし、これらはいずれもパターンの分類に主眼をおいたものであり、本手法のように物体の向きなどのパラメータを検出したり、3次元物体を表現しようとするものではなかった。

本論文ではまず、3次元構造を用いる手法(3次元照合法)と2次元照合を用いる手法(2次元照合法)との得失を考える。次に、2次元照合による物体認識のためのパラメトリック固有空間法について述べる。更に、本手法と他の2次元照合的な手法との比較を、認識実験により述べる。

## 2. 物体認識における3次元照合と2次元照合

3次元物体の画像認識には、前述したように3次元照合法と2次元照合法との2種類の異なったアプローチがある。ここではまず、それぞれの長所や短所を比較してみる。

3次元照合法の長所は、まず3次元的な完全性がある点である。つまりモデルの表現が3次元記述であるため、どの方向から見た物体も3次元の座標変換さえ施せば同じ表現になる。また、以下に述べる2次元照合法に比較してモデルの記述が一般的に小さくできる。この方法では、認識段階にいったん2次元画像から3次元形状あるいは3次元特徴を復元してから、3次元モデルと照合を行う。しかし、2次元画像から3次元情報を復元する手法にはいまだ完成された手法はなく、安定して3次元情報を復元できない。また、モデルを学習する段階でも、入力画像例から学習するためには同様の3次元情報復元の問題が生じる。そこで、一般的には人間の手作業によって入力されたCADモデル等を認識のモデルとして用いている。

2次元照合法は、基本的には、あらかじめ記憶してある2次元画像の集合と入力画像との2次元照合<sup>(11),(12)</sup>により3次元物体認識を行う。この手法は、3次元特徴の抽出が不要である等の特長がある。しかし、物体の見掛け画像は物体の向きや、光源の位置により大きく変化するため、この多様な2次元画像のバリエーションをあらかじめ学習しておくことは記憶容量、計算量

の点で困難であると考えられ、従来試みられていなかった。しかし、2次元照合法から記憶容量と照合の計算量に関する問題点を解決することができれば、3次元特徴抽出が困難な任意形状の対象に対しても汎用的な認識系を構築できることになる。

本手法では2次元照合法を採用する。ここでは画像符号化を基本としたパラメトリック固有空間法を提案したことにより、少ない記憶容量で3次元物体を2次元画像の集合体として記述することができる。その結果、2次元画像例から物体を容易に学習し、3次元物体を2次元照合により認識することが可能となった。

## 3. パラメトリック固有空間法による学習

3次元物体の見掛けの画像は、その物体の方向や照明の位置により大きく変動する。例えば、ある物体を1回転させただけで図1に示すような多様な画像が得られる。これをいかに記憶するかが、ここでの学習の問題となる。膨大な2次元の画像集合からその画像の情報の本質を抽出することは画像符号化の目的と同じである。ここではこの点に着目して符号化を基本とした画像の表現法を考える。ここではパラメトリック固有空間法という新しい画像表現法を提案する。

学習段階では、学習画像サンプルの集合からパラメトリック固有空間を構成する。この処理は2段階から構成される。1段階目は学習画像サンプルから固有ベクトルによる部分空間(固有空間)を構成する段階、2段階目は連続的に変化する学習画像の系列を固有空間上に投影し、部分空間上で多様体を構成することによ



図1 物体を回転させたときの見掛け画像の変化  
Fig. 1 A image set obtained by rotating an object.

りもとの画像系列を表現する段階である。物体の種類が複数の場合にはその数だけ多様体が構成される。認識段階では、まず入力画像を固有空間上の点に投影し、次にこの点と固有空間上の多様体との位置関係により認識および物体の方向を検出する。

### 3.1 画像の正規化

入力画像から、まず物体部分を切り出す。ここでは、仮にしきい値処理や背景との差分処理により切出しを行った。次に物体以外の部分に0の値を代入する。その後、物体を正方形に接するように、物体の縦横比を一定のまま大きさの正規化を行う。この画像をラスタースキャンし、その画素値を要素とするベクトル  $\hat{x}$ 、

$$\hat{x} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N]^T$$

でもとの画像を表現する。ここで  $N$  は画素数である。

次に、センサ感度の影響を除去するために、明るさの正規化を行う。正規化後の画像ベクトルを  $x$  とすると、ここではベクトル  $x$  の大きさが1となるように式、

$$x = \frac{\hat{x}}{\|\hat{x}\|}$$

により正規化する。

3次元物体の見掛け画像は、物体の向きと光源によって変化する。ここでは仮に物体の向きが1軸の周りを回転変化し、照明は1次元上を移動する点光源に背景光が重畳された場合を考える。これは工場内などの限られた環境の中では自然な設定である。また、物体の任意のポーズを扱うためには、パラメータ数を増やすことにより拡張できる。

ここで  $P$  種類の物体を学習する場合を考える。  $p$  番目の物体を1回転し、かつ光源の向きを変化させて収集した画像の集合を

$$\{x_{i1}^{(p)}, \dots, x_{k1}^{(p)}, x_{i2}^{(p)}, \dots, x_{kL}^{(p)}\}$$

で表現する。ここで  $R$  は回転方向の刻み数、  $L$  は光源の方向数を表す。これを第  $p$  物体の画像集合と呼ぶ。

またすべての物体に対する画像集合を、

$$\begin{aligned} & \{x_{i1}^{(1)}, \dots, x_{k1}^{(1)}, x_{i2}^{(1)}, \dots, x_{kL}^{(1)}, \\ & x_{i1}^{(2)}, \dots, x_{k1}^{(2)}, x_{i2}^{(2)}, \dots, x_{kL}^{(2)}, \\ & \vdots \\ & x_{i1}^{(P)}, \dots, x_{k1}^{(P)}, x_{i2}^{(P)}, \dots, x_{kL}^{(P)}\} \end{aligned}$$

で表現し、これを全物体の画像集合と呼ぶ。この集合内の各画像ベクトルを学習サンプルと呼ぶ。我々の実験ではこの学習サンプルの収集には、計算機制御で回転可能なターンテーブルと、光源方向の制御が可能なロボットアームを用いた。その様子を図2に示す。つまり学習サンプルは、対象となる物体をターンテー

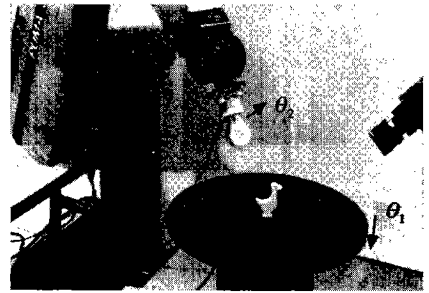


図2 3次元物体の学習画像収集システム  
Fig. 2 Setup used for automatic acquisition of object image sets.

ル上に置くことにより、すべて自動的に収集できる。

### 3.2 固有ベクトルの計算

図1の画像系列の例からもわかるように隣り合った二つの画像は極めて相関が高い。まず第1段階としてこの相関の性質を利用して、画像を圧縮する。ここでは画像集合に対して、2乗誤差の観点から最適に圧縮することが可能なカルーネン・レーベ展開を採用する。これは、画像集合の共分散行列の固有ベクトルが張る部分空間(固有空間)によりもとの画像を表現しようとする手法である。ここでは全物体に対する固有空間(以下ユニバーサル固有空間と呼ぶ)と物体  $p$  の固有空間の二つの固有空間を計算する。

まず、ユニバーサル固有空間を計算する。全学習サンプルの平均の画像  $c$ 、

$$c = \frac{1}{RLP} \sum_{p=1}^P \sum_{r=1}^R \sum_{l=1}^L x_r^{(p)l}$$

を計算し、次に各学習サンプルから平均画像を差し引き、行列  $X$ 、

$$X = [x_{i1}^{(1)} - c, \dots, x_{k1}^{(1)} - c, \dots, x_{iL}^{(P)} - c]$$

を作る。画像集合の共分散行列  $Q$  は、

$$Q = XX^T$$

により計算される。固有空間(例えば  $k$  次元)は次の固有方程式、

$$\lambda_i e_i = Q e_i$$

を解き、  $k$  個の大きい固有値 ( $\lambda_1 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_N$ ) に対応する固有ベクトル ( $e_1 \dots e_k$ ) を基底ベクトルとすることにより得られる。一般的に画像の共分散のように次元数(今回は16,384次元)の大きな行列の固有ベクトルの計算は困難である。しかし、画像数が少ない場合には、特異値分解やSTA法<sup>(10)</sup>などを利用することにより解くことが可能である。ユニバーサル固有空間は全物体の集合を表現するのに適した空間であり、物体の識別の際に利用する。

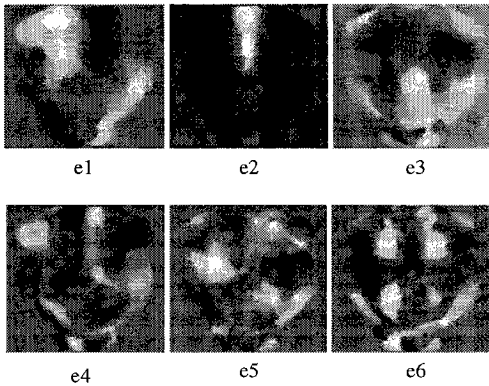


図3 図1に示した物体の固有ベクトルの例  
Fig. 3 Eigenvectors for the object shown in figure 1.

物体  $p$  の固有空間はその物体  $p$  の画像集合だけをを用いて計算する固有空間である。その共分散行列を  $Q^{(p)}$  とすると、第  $p$  物体の固有空間は、

$$\lambda_i^{(p)} e_i^{(p)} = Q^{(p)} e_i^{(p)}$$

を解き、その固有ベクトルを基底とすることにより得られる。物体  $p$  の固有空間はその物体を表現するのに適した空間であるため、物体の名前が識別された後に、その物体のポーズを推定する際に利用される。図3に図1で示した画像から作成した固有ベクトルの例を示す。

### 3.3 見掛け画像のパラメトリック固有空間表現

次に物体のポーズや、光源の位置の変化により連続的に変化する3次元物体を固有空間上の超平面により表現する。学習サンプルから平均画像を引いたベクトルを式

$$g_i^{(p)} = [e_1, e_2, \dots, e_k]^T (x_i^{(p)} - c)$$

により固有空間に投影すると、1枚の画像は点に対応する。例えば1回転分の学習サンプルを固有空間に投影するとそれは1次元の点の系列になる。なぜならば、一般的に物体のポーズの変化が少ない場合には画像の変化も少なく、相関が強い。また、強く相関をもった画像は固有空間上で近い位置に投影されるためである。例えば図1に示す図形の固有空間上での系列は、図4に示すようになる。実際には多次元空間での点列であるが、表示の都合上3次元で表示した。これらの点列は補間により連続的な変化として表現する。ここでは補間にはキュービックスプライン<sup>(17)</sup>を用いた。更に光源の位置を変化させた画像も同様に固有空間上に投影すると、物体のポーズと光源の位置の2パラメータにより表現される2次元多様体が固有空間上に構成され

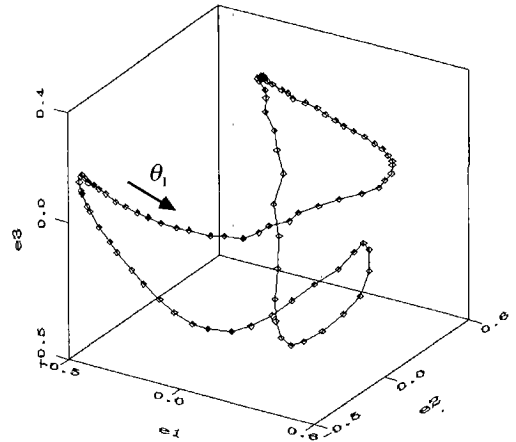


図4 パラメトリック固有空間の例(図1の物体の場合)  
Fig. 4 Parametric eigenspace representation of the object shown in Fig. 1.

る。この多様体を  $g^{(p)}(\theta_1, \theta_2)$  で表現する。  $\theta_1, \theta_2$  はそれぞれ回転と光源の位置のパラメータに対応する。この多様体は、学習サンプルに存在しない方向や光源の位置(中間の方向や光源の位置)に対する画像も、補間により表現していることになる。多様体は物体の種類の数だけユニバーサル固有空間上に構成される。

同様に物体  $p$  の固有空間に対して物体  $p$  の多様体を構成する。上記と同様に物体  $p$  の学習サンプルを物体  $p$  の固有空間に

$$f_i^{(p)} = [e_1^{(p)}, e_2^{(p)}, \dots, e_k^{(p)}]^T (x_i^{(p)} - c^{(p)})$$

により投影し、補間処理により多様体を構成する。ここで  $c^{(p)}$  は物体  $p$  の学習サンプルの平均である。そして、補間した表現を  $f^{(p)}(\theta_1, \theta_2)$  で表す。

## 4. 認 識

まず学習段階で用いたと同様の前処理を行う。つまり入力画像から、しきい値処理などにより物体領域を切り出す。切り出した後に大きさの正規化を行い、明るさの正規化を行う。正規化後の入力画像のベクトルを  $y$  とする。次にこのベクトルを次式によりユニバーサル固有空間上の点  $z$  に投影する。  $c$  は前述の平均画像である。

$$z = [e_1, e_2, \dots, e_k]^T (y - c)$$

認識はこの投影された点  $z$  が  $P$  種類ある多様体のどこに乗っているかを調べることになる。つまり、点  $z$  と多様体  $g^{(p)}(\theta_1, \theta_2)$  との距離

$$d_i^{(p)} = \min_{\theta_1, \theta_2} \| z - g^{(p)}(\theta_1, \theta_2) \|^2$$

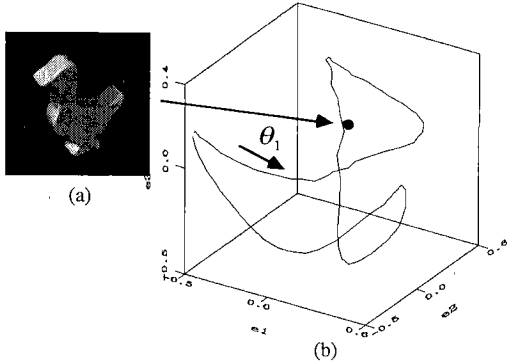


図5 (a)入力図形, (b)固有空間への投影(投影点の曲線上の位置が物体のポーズに対応)  
 Fig. 5 (a) An input image. (b) The input image is mapped to a point in the eigenspace. The location of the point on the curve determines the pose of the object.

を最小とする  $p$  を求めることにより実現できる。今回の実験では処理の高速化を図るために、この  $p$  を求める方法は、多様体上の代表点をすべて表に登録しておく、この表を探索することにより実現した。

物体名  $p$  が識別された後にその物体のポーズを推定する。まず、入力画像  $y$  を式

$$z^{(p)} = [e_1^{(p)}, e_2^{(p)}, \dots, e_k^{(p)}]^T (y - c^{(p)})$$

により物体  $p$  の固有空間に投影する。ポーズを検出するという事は、点  $z^{(p)}$  が多様体上のどこに位置しているかに対応している。そこで距離

$$d_2^{(p)} = \min_{\theta_1, \theta_2} \| z^{(p)} - f^{(p)}(\theta_1, \theta_2) \|$$

を最小とする  $\theta_1$  を見つける。図5に入力画像を固有空間上の点に投影し、その点に最小距離をもつ  $\theta_1$  を見つける様子を示す。この例では1次元多様体(曲線)の例で示すが、実際には2次元多様体(曲面)上の探索になる。

### 5. 認識実験

学習段階では物体を計算機制御の可能なターンテーブルの上に乗せ、学習画像サンプルを収集した(図1)。今回の実験では図6(a)に示す4種類の物体を用いた。学習サンプルとして4度おきに360度回転させたもの、つまり90種類のポーズの画像を収集した。光源位置は物体を取り巻くように見込み角を30度ずつ連続的に変化させた5箇所とした。つまり、全体で学習サンプル数は1,800(4×90×5)となる。今回の実験では画像は128×128の大きさに正規化した後に4バイト実数表現で格納したため、正規化後の学習サンプルの総データ

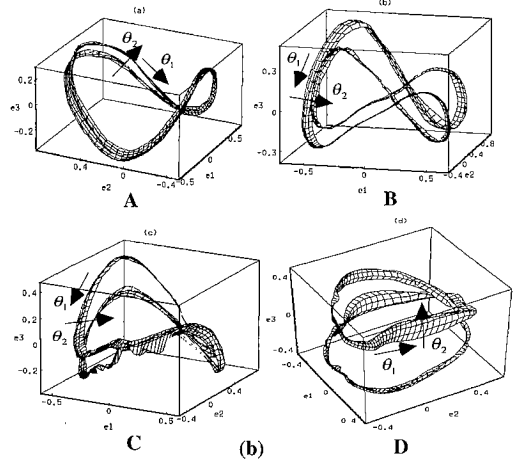
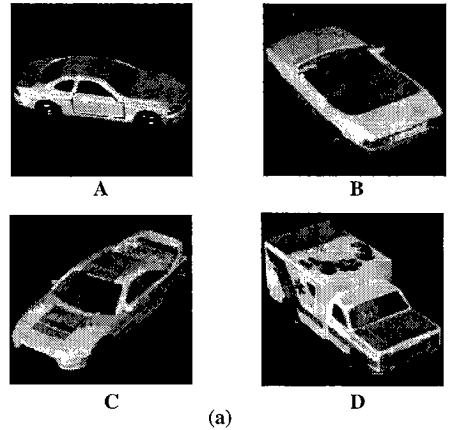


図6 (a)実験に使用した物体の例, (b)パラメトリック固有空間の例((a)の物体に対応)  
 Fig. 6 (a) The objects used in the experiments. (b) The parametric eigenspace for the objects in (a).

量は118 M バイトとなる。一方、認識実験用のデータとして、学習サンプルと2度位相の異なる90種類のポーズで光源位置を3箇所移動して収集した(合計1,080画像)。学習段階では3.で示した処理を施し、それぞれの物体についてパラメトリック固有空間を計算した。その例を図6(b)に示す。一つのパラメトリック固有空間のデータ量は、仮に8次元の固有空間上に一つの多様体が存在する場合には約640 K バイトとなる。その内訳は、固有ベクトル格納のために524 K バイト、多様体の表のために115 K バイトであった。

次に4.で示した認識処理を適用し、認識実験を行った。その実験結果を以下に示す。図7(a)は固有空間の次元に対する認識率を示す。8次元程度でほとんど認

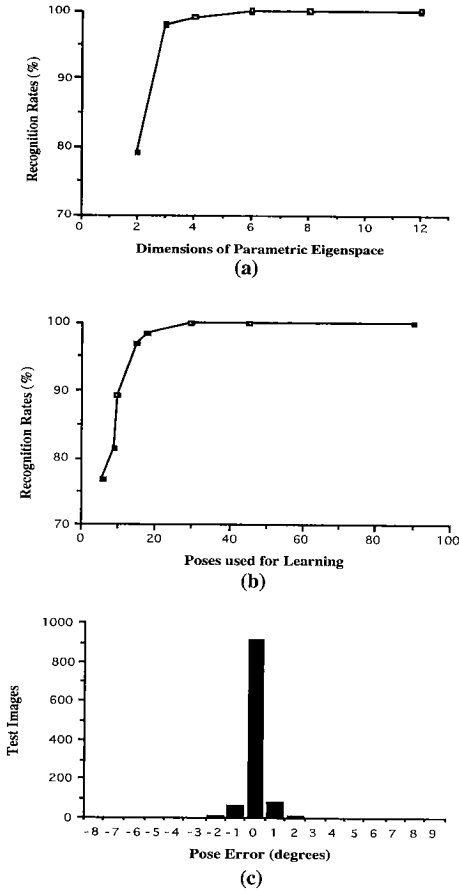


図7 (a)パラメトリック固有空間の次元と認識率との関係, (b)学習に用いたポーズ数と認識率との関係, (c)ポーズ推定誤差のヒストグラム

Fig. 7 (a) Recognition rate plotted as a function of the number of eigenspace dimensions. (b) Recognition rate plotted as a function of the number of poses of each learning object. (c) Histogram of the error in the pose estimation.

識率は飽和していることがわかる。本手法では学習サンプルに存在しないポーズの物体をもある程度、認識することが可能である。そこで限られたポーズ数でどの程度認識できるかを調べた。図7(b)に学習に利用するポーズを減らした際の認識率の変化を示す。今回示した程度の複雑さの物体では学習には15ポーズ程度で十分であることがわかる。図7(c)にはポーズ推定の精度を示した。平均値で1.2度のポーズ推定誤差であった。対象物体を背景から切り出す際に生じるエラーの認識精度に与える影響についても別に実験を行った。画像中での物体サイズの3%以下の切出し位置あるいは切出しサイズのエラーに対しても認識精度にはあまり影

響がなかった(文献(18)参照)。

## 6. 検 討

### 6.1 他手法との比較

本手法は基本的に2次元照合による3次元物体の認識である。従来、文字等の2次元図形認識用に開発された認識法が果たして3次元物体認識にどの程度適用可能かは興味深い。ここでは本手法と他の代表的な3手法と比較を行った。手法1は単純相関法、手法2はTurk<sup>(14)</sup>らが顔画像認識に用いたEigenface法、手法3は村瀬<sup>(9)</sup>らが文字認識で行った投影法である。手法の詳細はそれぞれの文献を参照されたい。

#### (1) 手法1：単純相関法

学習段階では物体 $p$ ごとに、学習用の画像集合の平均画像ベクトル $c^p$ を計算する。認識段階では入力画像 $y$ と平均画像ベクトルとの相関値 $d_s^p = y^T c^p$ を計算し、それが最大となる $p$ を認識結果とする。

#### (2) 手法2：Eigenface法(Turkほか<sup>(14)</sup>)

学習段階では画像集合の共分散行列の固有ベクトルを計算し固有空間(仮に8次元とした)をつくる。更に、学習サンプルをそこに投影し、その平均値 $c_s^p$ を物体 $p$ ごとに計算する。認識段階では、入力画像を固有空間に投影し、その点と平均値との距離 $d_t^p = \|z - c_s^p\|$ を計算し、それが最小となる $p$ を認識結果とする。

#### (3) 手法3：投影法(村瀬ほか<sup>(9)</sup>)

学習段階では物体 $p$ ごとに画像集合の相関行列の固有ベクトル $[e_1^p, e_2^p, \dots, e_k^p]$ を計算し、各物体 $p$ の固有空間をつくる。認識段階では、入力画像 $y$ を各物体の固有空間(仮に8次元とした)に投影し、その投影エネルギー $d_s^p = \sum_{k=1}^K (y^T e_k^p)^2$ が最大となる $p$ を認識結果とする。

この比較実験には図1で示した4種類の物体を含む8種類の物体を用いて行った。学習には各物体につき18の異なるポーズの画像(合計144枚)を用い、認識では90の異なるポーズの画像(合計720枚)を用いた。光源位置は固定とした。それぞれの認識率を表1に示す。手法1、手法2は変形の少ない2次元的な図形に対しては効果があるが、本手法のように見掛けの変化の大きい3次元物体にはあまり精度が得られないことがわかる。手法3は3次元物体に対しても比較的精度は得られるものの、物体のポーズ検出はできない。本提案の手法は他の3手法と比較して高い精度が得られることを確認した。また、同時に物体のポーズを検出できる点も優れている。

### 6.2 動画像への応用

本手法を動画像に適用した例を示す。この例では、入力画像(図8(a))から背景との差分により物体を切り出し(図8(b)),次に物体のポーズを検出している。図8(c)は最も近いポーズの学習サンプルを、図8(d)には検出した方向を示した。

表1 手法間の比較実験結果

	手法1	手法2	手法3	本手法
認識率	68.6%	66.8	98.7	99.8
ポーズ検出	不可能	不可能	不可能	可能

### 6.3 顔画像への応用

本手法は図9に示すような顔画像へもそのまま適用可能である。本手法によれば、任意の方向を向いた顔画像を認識すると同時に、その顔の方向(ポーズ)を自動的に検出することが可能となる。顔の方向を検出した実験結果(18方向の顔画像を学習に使用)の例を図9に示す。

### 6.4 人間の3次元物体認識法との対比

心理学の分野でも、人間が3次元物体を認識する際に、果たして2次元照合を利用しているか3次元照合を利用しているか興味もたれている。Edelmanらはメンタルローテーション<sup>(16)</sup>の心理実験により以下の知見<sup>(17)</sup>を示した。ある物体がその個人にとってあまり見

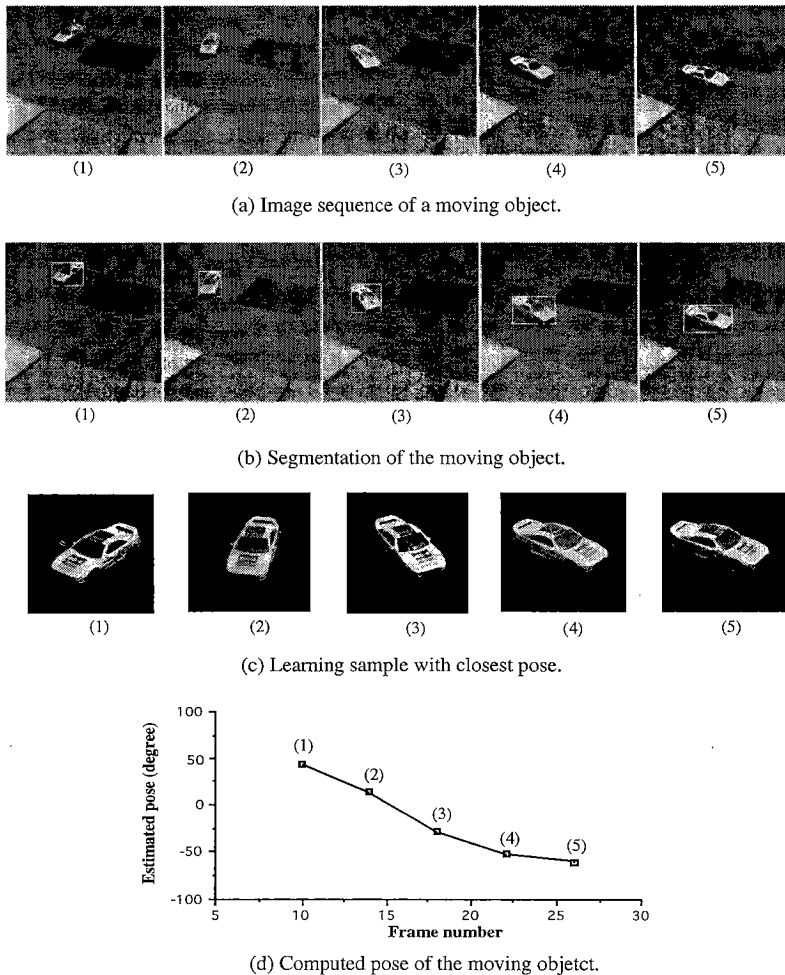


図8 パラメトリック固有空間法を動画像に適用した例  
Fig. 8 Parametric eigenspace method applied to the image sequence of a moving car.

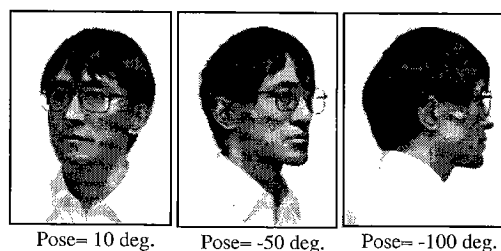


図9 顔画像認識への応用(正面を0度としたときの顔の方向の検出結果を示す)

Fig. 9 Pose estimation for the faces.

慣れていないような場合には、3次元構造を考慮しながらモデル物体との照合をとるが、よく見慣れた物体については2次元照合を行っている。これは、日常よく見る出現頻度の高い物体については、人間も処理の単純な2次元照合を利用していることを示している。

## 7. むすび

本論文では、任意の方向を向いた3次元物体を2次元照合により識別し、同時にその物体のポーズを検出する手法について述べた。ここで提案したパラメトリック固有空間法は、連続的に変化する画像系列を固有ベクトル空間上での多様体で表現する手法である。これにより少ない記憶容量で3次元物体を2次元画像の集合体として記憶することができるようになった。その結果、従来困難であったエッジや表現形状などの3次元構造を抽出することなく、2次元画像例から物体を学習し、2次元照合により3次元物体を認識し、ポーズを検出することが可能となった。今回の実験では、物体の1軸回転と光源の位置の2パラメータの場合を仮定したが、物体の任意のポーズ等を考えると更にパラメータ数が増える。今後はよりパラメータが増えた場合や、物体の種類が増えた場合などについて本手法の拡張性を検討していく予定である。

謝辞 貴重な意見を頂いたMITのT. Poggio教授、IBMワトソン研究所のD. Weinshallに感謝致します。日ごろ御指導頂く、木村NTT基礎研究所長、中津科学部長、内藤リーダに深謝します。

## 文 献

- (1) Chin R. T. and Dyer C. R. : "Model-Based Recognition in Robot Vision", ACM Computing Surveys, **18**, 1, pp. 67-108 (March 1986).
- (2) Besl P. J. and Jain R. C. : "Three-Dimensional Object Recognition", ACM Computing Surveys, **17**, 1, pp. 75-145 (1985).

- (3) Marr D. : "Vision", New York, W. H. Freeman (1982).
- (4) Ullman S. and Basri R. : "Recognition by Linear Combination of Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, **13**, 10, pp. 992-1006 (Oct. 1991).
- (5) Poggio T. and Edelman S. : "A networks that learns to recognize three-dimensional objects", Nature, **343**, pp. 263-266 (1990).
- (6) Fukunaga K. : "Introduction to Statistical Pattern Recognition", Academic Press, London (1990).
- (7) Oja E. : "Subspace Methods of Pattern Recognition", Research Studies Press, Hertfordshire (1983).
- (8) 飯島泰蔵 : "文字読み取り装置 ASPET/71", TV 学会誌, **27**, 3, pp. 157-164 (1973).
- (9) 村瀬 洋, 木村文隆, 吉村ミツ, 三宅康二 : "パターン整合法における特性核の改良とその手書き平仮名文字認識への応用", 信学論(D), **J64-D**, 3, pp. 276-283 (1981-03).
- (10) Murase H. and Lindenbaum M. : "Spatial Temporal Adaptive Method for Partial Eigenstructure Decomposition of Large Images", NTT Technical Report No. 6527 (March 1992) (IEEE Transaction on Image Processing, to appear).
- (11) Murase H. and Nayar S. K. : "Learning object models from appearance", AAAI-93, American Association for Artificial Intelligence, pp. 836-843 (July 1993).
- (12) Murase H. and Nayar S. K. : "Learning and recognition of 3D objects from appearance", IEEE Qualitative Vision Workshop, pp. 39-50, New York (June 1993).
- (13) Sirovich L. and Kirby M. : "Low dimensional procedure for the characterization of human faces", Journal of Optical Society of America, **4**, 3, pp. 519-524 (1987).
- (14) Turk M. A. and Pentland A. P. : "Face Recognition Using Eigenfaces", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591 (June 1991).
- (15) Press W., Flannery B. P., Teukolsky S. A. and Vetterling W. T. : "Numerical Recipes in C", Cambridge University Press, Cambridge (1988).
- (16) Tarr M. and Pinker S. : "Mental rotation and orientation-dependence in shape recognition", Cognitive Psychology, **21**, pp. 233-282 (1989).
- (17) Edelman S. and Weinshall D. : "A self-organizing multiple-view representation of 3D objects", Biological Cybernetics, **64**, pp. 209-219 (1991).
- (18) Murase H. and Nayar S. K. : "Illumination planning for object recognition in structured environments", Proc. of IEEE Conference on Computer Vision and Pattern Recognition (June 1994).

(平成5年10月25日受付, 6年4月20日再受付)





**村瀬 洋**

昭53名大・工・電子卒。昭55同大大学院修士課程了。同年日本電信電話社(NTT)入社。以来、文字・図形認識、コンピュータビジョンの研究に従事。平4から1年間米国コロンビア大学計算機学科に研究員として滞在。現在、NTT基礎研究所情報科学部主幹研究員と同時にNTT特別研究員。工博。昭60年度本会学術奨励賞受賞。1994年のCVPR会議(IEEE)で最優秀論文賞を受賞。IEEE、情報処理学会、AVIRG各会員。



**シュリー ナイヤー**

1990カーネギーメロン大学計算機学科にてph. Dを取得。1991からコロンビア大学計算機学科の助教授。1990のICCV会議でMarr賞を受賞。コンピュータビジョン、特に物理ベースビジョン、ロボットビジョンの研究に従事。