

# 複数カメラ入力によるアピアランスベース物体認識のための カメラ配置決定手法

高橋 友和<sup>†,††a)</sup> Lina<sup>†</sup> 松ヶ野修功<sup>†\*</sup> 井手 一郎<sup>†,†††</sup>  
目加田慶人<sup>††††</sup> 村瀬 洋<sup>†</sup>

Planning of Camera Positions for Appearance Based Object Recognition  
Using Multiple Cameras

Tomokazu TAKAHASHI<sup>†,††a)</sup>, Lina<sup>†</sup>, Osanori MATSUGANO<sup>†\*</sup>, Ichiro IDE<sup>†,†††</sup>,  
Yoshito MEKADA<sup>††††</sup>, and Hiroshi MURASE<sup>†</sup>

あらまし 複数のカメラを用いて物体認識を行う場合、認識率はカメラの配置に大きく影響を受ける。本論文では、複数のカメラを用いて精度良く物体認識を行うためのカメラ配置の決定手法を提案する。物体認識のフレームワークとしてはパラメトリック固有空間法を用いる。パラメトリック固有空間法では、学習段階において固有空間上に各物体に対応する多様体を形成する。提案手法は、これらの多様体の空間的な分布の形状特徴に基づいた評価関数を用いてカメラ配置を決定するものである。多様体間の距離に基づく評価関数を定義し、カメラを2台用いる場合、3台用いる場合についてカメラ配置決定に関する実験を行った。実験の結果から提案手法が認識性能の向上に有効であることを示す。

キーワード 物体認識, 複数カメラ, カメラ配置, パラメトリック固有空間法

## 1. ま え が き

三次元物体の認識は監視システムやヒューマンインタフェース、産業応用の分野で重要な技術である。物体認識を実現する手法では様々なアプローチが用いられるが、我々はノイズに対しロバストなアピアランスベースの物体認識に着目する。パラメトリック固有空間法 [1] はその中で広く用いられる手法の一つであり、物体の認識とその物体の姿勢、光源方向などのパラ

メータの推定を同時に行う手法である。この手法に関して様々な関連研究が報告されている [2] ~ [4]。

アピアランスベースの物体認識に関して、認識性能を低下させる主な要因として以下の二つが考えられる。一つは撮影状況に起因する入力画像の品質低下 (図 1) であり、もう一つは入力画像が撮影されたカメラの位置が物体の識別に適さない場合に生じる (図 2)。認識を行う際に複数の入力画像を用いることにより、これらの問題の解決を図る試みが報告されている。山口らは顔認識の入力画像として、ビデオシーケンス中の複数のフレーム画像を用いている [5]。アクティブリコ

<sup>†</sup> 名古屋大学大学院情報科学研究科, 名古屋市

Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi, 464-8603 Japan

<sup>††</sup> 日本学術振興会, 東京都

Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo, 102-8471 Japan

<sup>†††</sup> 国立情報学研究所, 東京都

National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

<sup>††††</sup> 中京大学生命システム工学部, 豊田市

Department of Life System Science and Technology, Chukyo University, Toyota-shi, 470-0393 Japan

\* 現在, トヨタ自動車株式会社

a) E-mail: ttakahashi@murase.m.is.agoya-u.ac.jp



(a) Low quality

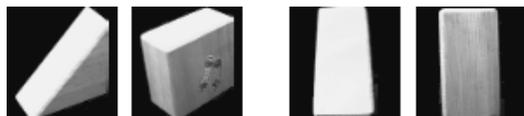
(b) High quality

図 1 撮影状況に依存する画像の品質の違い  
Fig. 1 Difference of image quality depending on capturing conditions.

グニション [4], [6] のフレームワークでは, 認識の過程において適応的にカメラ位置を移動させながら複数の入力画像を得ることにより, 認識性能の向上を図っている.

一方, 複数の入力画像を得る際により一般的に用いられる方法は, あらかじめ配置された複数のカメラから入力画像を得るものである. Selinger らは複雑な背景を含むシーン中の物体認識の問題に複数のカメラを用いている [7]. Shakhnarovich らは複数のカメラを用いて人物の顔と歩行姿勢の認識を行っている [8]. 実際の応用を考えた場合, 認識性能はカメラの配置と認識対象となる物体の形状特徴やテクスチャ特徴に依存すると考えられるため, あらかじめ設置するカメラの配置は慎重に検討しなければならない. カメラ配置決定の従来手法として次の二つの方法が考えられる. 認識対象に対する経験的な知識を利用して人手によりカメラ配置を決定する方法と, 様々なカメラ配置を用意して実際に認識実験を行った結果から決定する方法である. しかし, 前者には多種多様な認識対象への対応が困難であるという問題があり, 後者は手元にあるサンプルの集合を学習用とテスト用に分けることによって実現可能であるが, 学習用のサンプル集合が小さくなる問題と, 評価できるカメラ配置がテストサンプルを撮影したカメラ位置からなる配置に制限される問題がある. 学習用サンプル集合が小さくなる問題は, クロスバリデーション法を用いることによって軽減できるが, この場合には各カメラ配置について学習と認識を繰り返す必要があり, 手間と計算コストがかかる. また, 手元にあるサンプルの集合中に存在しないカメラ配置を評価するためには, そのカメラ配置からのテスト用サンプル集合を新たに取得する必要があり, この作業を各カメラ配置に対して繰り返し行うことは, 大変な手間を要し, 現実的ではない.

そこで, 我々は複数のカメラを用いて精度良く物体認識を行うためのカメラ配置の決定手法を提案する. 物体認識のフレームワークとしてパラメトリック固有



(a) Easy to recognize (b) Hard to recognize

図 2 カメラ位置の違いによる見え方の違い

Fig. 2 Difference of object appearances depending on camera positions.

空間法を用いる. 提案手法は, 学習段階において固有空間上に形成される多様体の情報のみを用いて, 与えられたカメラ配置を評価し, 実際の認識実験を行うことなく適切なカメラ配置を決定するものである. そのため, 多種多様な認識対象に容易に適用可能であり, サンプル集合全体を利用して学習した多様体を用いて, サンプルを撮影したカメラ配置以外の任意のカメラ配置を評価することが可能である. これにより, 上述のようなカメラ配置決定のための手間と計算コストの大幅な削減が期待できる. 以下, 2. で複数カメラ入力を用いる場合のパラメトリック固有空間法について述べ, 3. で提案手法の詳細を述べる. 4. で実験の方法を述べ, 5. で結果の考察を行い, 6. で本論文をまとめる.

## 2. 複数カメラ入力によるパラメトリック固有空間法

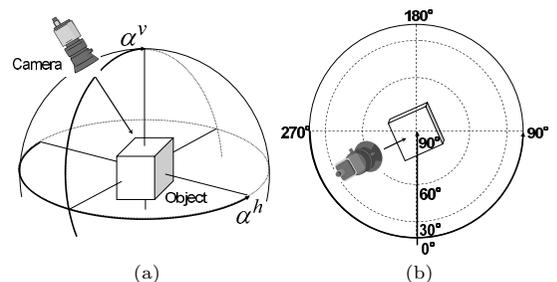
図 3 に本論文で用いるカメラの座標系を示す. 各カメラの位置  $\alpha$  は  $\alpha^h$  と  $\alpha^v$  の二つのパラメータで表される. これらはそれぞれ水平方向, 垂直方向のカメラの角度を表す. 以下, 複数カメラを用いたパラメトリック固有空間法について述べる.

### 2.1 学習段階

各学習画像に対し, 画素値を要素とした特徴ベクトル  $\hat{x}$  を  $x = \hat{x} / \|\hat{x}\|$  によって正規化したものから行列  $X$  を以下のように作成する.

$$X = (x_{1,1}^{(1)} - c, \dots, x_{h,v}^{(p)} - c, \dots, x_{H,V}^{(P)} - c) \quad (1)$$

ここで,  $h (= 1, 2, \dots, H)$ ,  $v (= 1, 2, \dots, V)$  はそれぞれ物体  $p (= 1, 2, \dots, P)$  の水平方向, 垂直方向の姿勢を表す識別子である.  $c$  はすべての  $h, v, p$  に関する  $x$  の平均ベクトルである. ユニバーサル固有空間を  $XX^T$  の上位  $k$  個の固有値に対応する固有ベクトル  $e_i (i = 1, 2, \dots, k)$  から形成する. 以下の式を用いて各特徴ベクトル  $x_{h,v}^{(p)}$  を固有空間上の点  $g_{h,v}^{(p)}$  に投影



(a) (b)

図 3 カメラ位置の座標系

Fig. 3 Coordinate system for camera positions.

する．

$$g_{h,v}^{(p)} = (e_1, e_2, \dots, e_k)^T (x_{h,v}^{(p)} - c) \quad (2)$$

$g_{h,v}^{(p)}$  から三次スプライン補間を用いて、各物体に対応する多様体  $G^{(p)}(\theta)$  を形成する．ここで、 $\theta = (\theta^h, \theta^v)$  は物体の姿勢パラメータベクトル、あるいは物体撮影時のカメラ位置を表す．

同様に各物体に対して、その物体に関する学習画像のみを用いてオブジェクト固有空間を形成し、オブジェクト固有空間上に多様体  $G^{(p)}(\theta)$  を形成する．これは、物体の姿勢パラメータ推定に用いる．

## 2.2 認識段階

任意の姿勢で入力される物体が、学習された物体セット中のどの物体であるかを認識する．それと同時に、その物体の姿勢パラメータを推定する．

### 2.2.1 単一カメラ入力の場合

カメラ位置を既知とする単一のカメラを入力に用いた物体認識について説明する．入力される物体の姿勢パラメータベクトル  $\tilde{\theta}$  が未知である場合を想定する．学習段階と同様に、カメラから入力される画像の特徴ベクトル  $y$  をユニバーサル固有空間上の点  $z$  に以下の式を用いて投影する．

$$z = (e_1, e_2, \dots, e_k)^T (y - c) \quad (3)$$

$z$  と  $G^{(p)}(\theta)$  の距離を以下のように定義する．

$$d^{(p)}(\theta) = \|z - G^{(p)}(\theta)\|^2 \quad (4)$$

各  $p$  に対して、 $d^{(p)}(\theta)$  の  $\theta$  に関する最小値を以下のように算出する．

$$d^{(p)} = \min_{\theta} d^{(p)}(\theta) \quad (5)$$

認識結果として、 $\tilde{p}$  を以下の式で求める．

$$\tilde{p} = \arg \min_p d^{(p)} \quad (6)$$

姿勢パラメータの推定に関して、入力された物体の姿勢パラメータベクトル  $\tilde{\theta}$  は以下の式を用いて推定する．

$$\tilde{\theta} = \arg \min_{\theta} d^{(p)}(\theta) \quad (7)$$

ここで、 $d^{(p)}(\theta)$  は物体  $\tilde{p}$  に関するオブジェクト固有空間上の投影点と多様体の間の距離を表すもので、式 (4) の  $G^{(p)}(\theta)$  を物体  $\tilde{p}$  に対するオブジェクト固有空間上の多様体  $G^{(p)}(\theta)$  に置き換えたものである．ここで得られる  $\tilde{\theta}$  は、カメラ位置パラメータに対する相対的な値である．

### 2.2.2 複数カメラ入力の場合

カメラ位置を既知とする  $M$  台のカメラを用いた物体認識について説明する．単一カメラ入力の場合と同様に、入力される物体の姿勢パラメータベクトル  $\tilde{\theta}$  が未知である場合を想定する．カメラ配置を  $A = \{\alpha_m | m = 1, 2, \dots, M\}$  で表し、 $m$  台目のカメラ位置を  $\alpha_m = (\alpha_m^h, \alpha_m^v)$  で表す． $m$  台目のカメラから入力される画像の特徴ベクトル  $y_m$  をユニバーサル固有空間上の点  $z_m$  に以下の式を用いて投影する．

$$z_m = (e_1, e_2, \dots, e_k)^T (y_m - c) \quad (8)$$

$z_m$  と  $G^{(p)}(\theta)$  の距離を以下のように定義する．

$$d^{(p)}(\theta) = \sum_{m=1}^M \|z_m - G^{(p)}(\theta + \alpha_m)\|^2 \quad (9)$$

認識結果として、 $\tilde{p}$  を式 (9) と式 (6) より求める．

物体の姿勢パラメータベクトル  $\tilde{\theta}$  は式 (7) を用いて推定する．ただし、式 (7) の  $d^{(p)}(\theta)$  は式 (9) の  $G^{(p)}(\theta)$  を  $G^{(p)}(\theta)$  に置き換えて算出する．

## 3. 物体認識に適切なカメラ配置の決定

この章では、はじめに複数カメラ入力の問題を単一カメラ入力の問題と等価なものとして扱うために固有空間の直積空間について説明し、次にカメラ配置の決定手法について説明する．

### 3.1 固有空間の直積空間

あるカメラ配置  $A$  が与えられたとき、固有空間のカメラ台数回の直積により形成される直積空間を考えると、複数のカメラを扱う問題をよりシンプルな単一カメラを扱う問題と等価なものとして考えることができる．固有空間の次元数が  $k$ 、カメラの台数が  $M$  であるとき、この直積空間の次元数は  $Mk$  次元となる．

2.2.2 で説明した複数カメラ入力による認識の問題を例に取って説明する． $M$  台のカメラからなるカメラ配置  $A = \{\alpha_m | m = 1, 2, \dots, M\}$  が与えられたとき、物体  $p$  に対応する直積空間上の多様体  $G_A^{(p)}(\theta)$  は、ユニバーサル固有空間上の多様体  $G^{(p)}(\theta)$  を用いて以下のように定義できる．

$$G_A^{(p)}(\theta) = (G^{(p)}(\theta + \alpha_1)^T, \dots, G^{(p)}(\theta + \alpha_M)^T)^T \quad (10)$$

ただし、この式から分かるように、直積空間上の多様

体  $G_A^{(p)}(\theta)$  はカメラ配置  $A$  に依存して位置や形状が変化する。

カメラ配置  $A$  の  $M$  台のカメラから入力され、ユニバーサル固有空間に投影された特徴ベクトル  $z_1, z_2, \dots, z_M$  は、直積空間上では以下のような一つの特徴ベクトル  $z_A$  として表現できる。

$$z_A = (z_1^T, z_2^T, \dots, z_M^T)^T \quad (11)$$

$G_A^{(p)}(\theta)$  と  $z_A$  の距離  $d_A^{(p)}(\theta)$  は以下のように定義でき、 $d_A^{(p)}(\theta)$  を最小とする  $p$  と  $\theta$  を選んだとき、その  $p$  が認識結果となる。

$$d_A^{(p)}(\theta) = \|z_A - G_A^{(p)}(\theta)\|^2 \quad (12)$$

この式は、2.2.1 に示した式 (4) と同様にシンプルな表現で定義されている。また、この式によって算出される値は 2.2.2 に示した式 (8) によって算出される値と等しい。

このように、固有空間の直積空間を考えることによって、複数のカメラを扱う問題と単一のカメラを扱う問題を等価なもののみなすことができる。

### 3.2 固有空間上でのカメラ配置の決定

複数カメラを用いた物体認識を精度良く行うためのカメラ配置の決定手法を提案する。提案手法は、学習段階において形成されたユニバーサル固有空間上の多様体の情報のみから、与えられたカメラ配置の認識に対する適切さを評価し、カメラ配置を決定するものである。

求めたいカメラ配置をカメラ位置の集合  $\tilde{A} = \{\tilde{\alpha}_m | m = 1, 2, \dots, M\}$  で表すとき、提案手法は  $\tilde{A}$  を以下に示す式を用いて決定する。

$$\tilde{A} = \arg \max_A F(A) \quad (13)$$

$F(A)$  はカメラ配置  $A$  の適切さを測る評価関数であり、ユニバーサル固有空間上の多様体  $G^{(p)}(\theta)$  を用いて構成する。

### 3.3 多様体間距離に基づいたカメラ配置の評価関数

ユニバーサル固有空間上には、複数の多様体同士が互いに交差、あるいは極端に接近している領域が存在する。認識の際、入力された特徴ベクトルがこれらの領域に投影された場合、一般に認識は困難となる場合が多い。このことから、カメラ配置に関する評価関数の一つとして、多様体間距離に基づく評価関数を定義

する。

$M$  台のカメラによる任意の配置  $A$  が与えられたとき、二つの物体  $p, q$  に対応する多様体間の距離を、直積空間上で以下のように定義する。

$$D_A^{(p,q)} = \min_{\theta_p, \theta_q} \|G_A^{(p)}(\theta_p) - G_A^{(q)}(\theta_q)\|^2 \quad (14)$$

この式から、 $P$  個の多様体を与えられたとき、各多様体間の距離の最小値を用いて、カメラ配置  $A$  に対するそれらの多様体同士の分散の度合を以下のように定義する。ただし、 $p = (1, 2, \dots, P)$ ,  $q = (1, 2, \dots, P)$ ,  $p \neq q$  とする。

$$D_A = \min_{p,q} D_A^{(p,q)} \quad (15)$$

例えば、あるカメラ配置  $A$  が与えられたとき、三つの多様体が直積空間上に図 4 のように分布している場合には、はじめに各多様体間の距離を算出し、その中で最小のものが  $D_A$  として決定される。 $D_A$  は多様体同士の分散の度合を測る尺度の一つであり、与えられたカメラ配置  $A$  に対する物体の識別の容易さを表すと考えられる。3.1 で述べたように、直積空間上の多様体  $G_A^{(p)}(\theta)$  は  $A$  に依存して位置や形状が変化するため、 $D_A$  を最大とする  $A$  は物体の識別が最も容易となるカメラ配置であると考えられる。 $D_A$  を  $A$  の関数  $D(A)$  として、直積空間を用いずにもとのユニバーサル固有空間上で表現すると以下ようになる。

$$D(A) = \min_{p,q} \min_{\theta_p, \theta_q} \sum_{m=1}^M \|G^{(p)}(\theta_p + \alpha_m) - G^{(q)}(\theta_q + \alpha_m)\|^2 \quad (16)$$

式 (13) と式 (16) より、 $F(A) = D(A)$  として、提

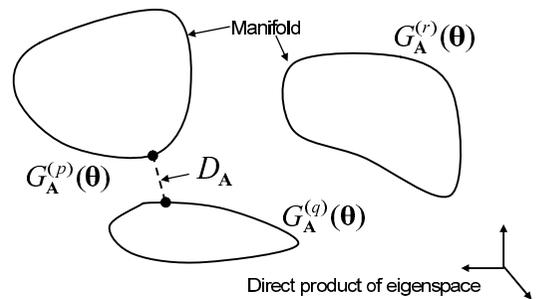


図 4 直積空間上でのカメラ配置  $A$  に対する  $D_A$  の算出  
Fig. 4 Calculation of  $D_A$  corresponding to camera arrangement  $A$  in a direct product of eigenspace.

案手法は求めたいカメラ配置  $\tilde{A}$  を以下のように決定する。

$$\tilde{A} = \arg \max_A D(A) \quad (17)$$

### 3.4 評価関数と認識率の関係

式 (16) に示した評価関数は、カメラ配置が与えられたとき、図 4 に示すように任意の二つの多様体の間の距離を計測し、それらの最小値をそのカメラ配置の評価値とするものである。つまり、これはそのカメラ配置において認識を行う際に、入力された特徴ベクトルが、多様体同士が最も接近している領域に投影される最悪の場合の識別の容易さを表しており、評価値が高いカメラ配置を選ぶことによってそのような最悪の場合の識別性能の向上が期待できる。しかしながら、これによってそれ以外の場合の識別性能が低下することも考えられるため、全体の認識率を向上させることは必ずしも直結しない。そこで、次章以降、様々なカメラ配置に関して、その評価値と実際の認識率の関係を実験的に調査することにより、提案手法の有効性を示す。

## 4. 実験方法

カメラが 2 台の場合と 3 台の場合について、提案手法によって決定されたカメラ配置を用いて認識実験を行うことにより、提案手法の有効性を調査した。カメラ配置の評価関数として式 (16) を用いた。72 種類の物体を以下に示す四つの物体セットに分割して用いた。

- FP (Football Players) 27 種のサッカー選手人形
- CC (Cartoon Characters) 15 種の漫画キャラクター人形
- WB (Wooden Blocks) 9 種の積み木
- FC (Faces) 21 人の顔画像

実験で用いた物体の一部を図 5 に示す。

学習段階では、各物体について、カメラ位置を水平方向 0~350 度を 10 間隔、垂直方向 0~90 度を 15 度

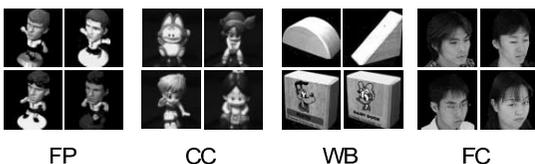


図 5 実験で用いた物体の例  
Fig. 5 Examples of objects used in the experiments.

間隔で変化させながら、合計 252 枚の画像を得た。各画像から物体が外接する最小の正方形領域を切り出し、画像サイズをそれぞれ 64 × 64 に伸縮させたものを学習画像とした。各物体セットに対してユニバーサル固有空間を形成し、各物体に対応する多様体を物体が属する物体セットの固有空間上に形成した。このとき、固有空間の次元数を 4 とした。

提案手法を用いてカメラ配置を決定する際には、各カメラの位置を水平方向 12 段階 (0~330 度, 30 度間隔)、垂直方向 4 段階 (0~90 度, 30 度間隔) で変化させ、それらの組合せ数分のカメラ配置を用意した。カメラを 2 台用いる場合と 3 台用いる場合について、用意したすべてのカメラ配置の中から提案手法を用いて適切なカメラ配置を決定した。この際に算出された各カメラ配置の評価値を記録した。多様体間距離を算出する際、任意の異なる多様体上の点同士の距離を算出する必要があるが、本実験では三次スプライン曲面で表された各多様体上に 360 (水平方向) × 90 (垂直方向) の点群をサンプリングし、それらの点同士の距離を近似的に用いた。

認識段階では、学習画像を撮影したカメラ位置を水平方向に 5 度ずらした位置から撮影した画像に対し、学習画像と同様に物体領域を正方形で切り出した。学習時と認識時の撮影環境の違いを想定して、それらの画像をサイズ 32 × 32, 16 × 16, 8 × 8 の 3 段階に伸縮させたものを入力画像とした (図 6)。入力画像を固有空間に投影する際には、線形補間を用いて画像のサイズを 64 × 64 に伸長したのから特徴ベクトルを形成し、それを投影した。各物体セット、各画像サイズに対して、提案手法によって決定されたカメラ配置と、それ以外のカメラ配置を用いた場合について、入力する物体の種類、並びにその物体の姿勢パラメータを変化させながら認識を繰り返し行い、実際の認識率を算出した。

ここで、カメラ配置決定、並びに認識段階に関して、

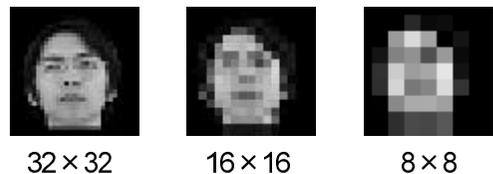


図 6 実験で用いた 3 段階の入力画像サイズ  
Fig. 6 Three levels of input image size used in the experiments.

3.1 において説明した直積空間を導入した．その理由はカメラ台数分の固有空間の直積空間上での距離計算を行うことにより，カメラの台数に依存しない同一のフレームワークを用いた処理が可能となるためである．ただし，結果として得られるカメラ配置や認識率，それを求めるために必要な計算量は直積空間を用いない場合と同じである．

### 5. 結果と考察

#### 5.1 物体認識の精度

表 1 は，物体セット CC，画像サイズ  $16 \times 16$  に対して，カメラを 2 台用いて行った実験に関して，また，表 2 は，物体セット FC，画像サイズ  $16 \times 16$  に対して，カメラを 3 台用いて行った実験に関して，そこで用いられたカメラ配置の例とそのときの認識率を示している．

表 1 2 台のカメラを用いた場合の認識率とカメラ配置の例 (物体セット CC，入力画像サイズ  $16 \times 16$ ): (c) は提案手法によって決定されたカメラ配置

Table 1 Two-camera recognition rates in CC,  $16 \times 16$  cases and their camera arrangements: (c) shows results of the proposed method.

カメラ配置	(a)	(b)	(c)
$(\alpha_1^h, \alpha_1^v)$	$(0^\circ, 0^\circ)$	$(0^\circ, 0^\circ)$	$(0^\circ, 30^\circ)$
$(\alpha_2^h, \alpha_2^v)$	$(0^\circ, 0^\circ)$	$(180^\circ, 0^\circ)$	$(0^\circ, 90^\circ)$
認識率	45.7%	77.4%	93.0%

表 2 3 台のカメラを用いた場合の認識率とカメラ配置の例 (物体セット FC，入力画像サイズ  $16 \times 16$ ): (c) は提案手法によって決定されたカメラ配置

Table 2 Three-camera recognition rates in FC,  $16 \times 16$  cases and their camera arrangements: (c) shows results of the proposed method.

カメラ配置	(a)	(b)	(c)
$(\alpha_1^h, \alpha_1^v)$	$(0^\circ, 90^\circ)$	$(0^\circ, 0^\circ)$	$(0^\circ, 30^\circ)$
$(\alpha_2^h, \alpha_2^v)$	$(0^\circ, 90^\circ)$	$(30^\circ, 0^\circ)$	$(120^\circ, 30^\circ)$
$(\alpha_3^h, \alpha_3^v)$	$(0^\circ, 90^\circ)$	$(330^\circ, 0^\circ)$	$(240^\circ, 30^\circ)$
認識率	34.3%	83.7%	100.0%

表 1, 表 2 のカメラ配置 (c) は提案手法によって決定されたものであり，他のカメラ配置 (a), (b) を用いた場合よりも高い認識率が得られた．図 7 は，物体セット CC，入力画像サイズ  $16 \times 16$  に対して，カメラを 2 台用いた実験において，また，図 8 は，物体セット FC，入力画像サイズ  $16 \times 16$  に対して，カメラを 3 台用いた実験において，用いられたすべてのカメラ配置に関して，式 (16) で算出されたカメラ配置の評価値 (Evaluation value of camera arrangement) とそのカメラ配置を用いて実際に調べた認識率 (Recognition rate) の関係を示している．図 7, 図 8 の (a), (b),

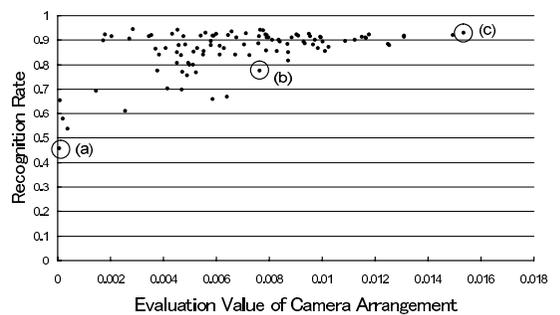


図 7 2 台のカメラを用いた場合の認識率とカメラ配置の評価値の関係 (物体セット CC，入力画像サイズ  $16 \times 16$ ): (a), (b), (c) は表 1 に示したカメラ配置にそれぞれ対応している

Fig. 7 Relation between two-camera recognition rates in CC,  $16 \times 16$  cases and evaluation values of camera arrangements: arrangements (a), (b), and (c) correspond to the arrangements in Table 1.

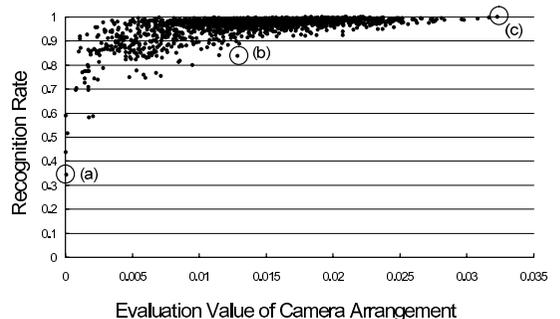


図 8 3 台のカメラを用いた場合の認識率とカメラ配置の評価値の関係 (物体セット FC，入力画像サイズ  $16 \times 16$ ): (a), (b), (c) は表 2 に示したカメラ配置にそれぞれ対応している

Fig. 8 Relation between three-camera recognition rates in FC,  $16 \times 16$  cases and evaluation values of camera arrangements: arrangements (a), (b), and (c) correspond to the arrangements in Table 2.

(c) はそれぞれ表 1, 表 2 のカメラ配置に対応している．これらの図から，評価値の高いカメラ配置からは高い認識率が得られ，認識率の低いカメラ配置の評価値は低いという関係が見られる．この関係は，カメラ台数や物体セット，入力画像サイズによらず，本実験のすべての場合において確認された．このことから，提案手法を用いて評価値からカメラ配置を決定することは，認識性能の向上に有効であるといえる．

表 3, 表 4 はそれぞれカメラ 2 台を用いた場合と 3 台を用いた場合について，提案手法によって決定されたカメラ配置（最大の評価値をもつカメラ配置）と不適切なカメラ配置（最小の評価値をもつカメラ配置）を用いたときの認識率を示している．適切でないカメラ配置を用いた場合と比較して，提案手法によって決定されたカメラ配置を用いることにより，入力画像サイズが小さく認識が困難な場合でも精度良く認識が行われていることが確認できる．

5.2 姿勢パラメータ推定の精度

物体認識が失敗した場合の姿勢パラメータ推定の結果は，一般的には意味をもたないと考えられるが，人間の顔のように姿勢の違いによる見えの変化に類似性がある場合には，認識が失敗した場合でも姿勢パラメータ推定に意味があると考え，物体セット FC に関して姿勢パラメータ推定実験を行った．図 9 は，物体セット FC，入力画像サイズ 16 × 16 に対して，カ

メラを 3 台用いた実験において，用いられたすべてのカメラ配置に関するカメラ配置の評価値（Evaluation value of camera arrangement）とそのカメラ配置を用いて実際に推定された姿勢パラメータの RMS 誤差（RMS Error of Pose Estimation）の関係を示している．この図から，上記の認識率に関する考察と同様に，評価値の高いカメラ配置からは高い姿勢パラメータ推定精度が得られ，姿勢パラメータ推定精度の低いカメ

表 4 提案手法により決定されたカメラ配置を用いることによる認識率の向上（カメラ 3 台）

Table 4 Improvement of recognition rates in three-camera cases by using camera arrangements planned by the proposed method.

画像サイズ	物体セット	認識率	
		評価値最小	評価値最大 (提案手法)
32 × 32	FP	71.5%	<b>99.4%</b>
	CC	74.6%	<b>98.1%</b>
	WB	86.7%	<b>97.5%</b>
	FC	90.4%	<b>100.0%</b>
16 × 16	FP	36.8%	<b>99.4%</b>
	CC	45.7%	<b>95.7%</b>
	WB	65.8%	<b>90.0%</b>
	FC	59.0%	<b>100.0%</b>
8 × 8	FP	20.3%	<b>89.8%</b>
	CC	18.5%	<b>76.1%</b>
	WB	49.4%	<b>64.2%</b>
	FC	28.0%	<b>59.3%</b>

表 3 提案手法により決定されたカメラ配置を用いることによる認識率の向上（カメラ 2 台）

Table 3 Improvement of recognition rates in two-camera cases by using camera arrangements planned by the proposed method.

画像サイズ	物体セット	認識率	
		評価値最小	評価値最大 (提案手法)
32 × 32	FP	71.5%	<b>99.4%</b>
	CC	74.6%	<b>97.0%</b>
	WB	86.7%	<b>96.9%</b>
	FC	90.4%	<b>99.9%</b>
16 × 16	FP	36.8%	<b>95.7%</b>
	CC	45.7%	<b>93.0%</b>
	WB	65.8%	<b>84.4%</b>
	FC	59.0%	<b>94.3%</b>
8 × 8	FP	20.3%	<b>61.4%</b>
	CC	18.5%	<b>60.7%</b>
	WB	49.4%	<b>75.0%</b>
	FC	28.0%	<b>48.8%</b>

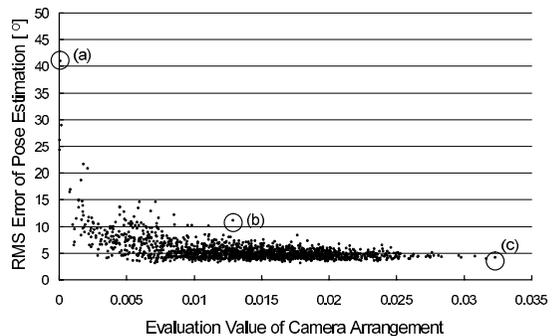


図 9 3 台のカメラを用いた場合の姿勢推定 RMS 誤差とカメラ配置の評価値の関係（物体セット FC，入力画像サイズ 16 × 16）: (a), (b), (c) は表 2 に示したカメラ配置にそれぞれ対応している

Fig. 9 Relation between three-camera pose estimation RMS errors in FC, 16 × 16 cases and evaluation values of camera arrangements: arrangements (a), (b), and (c) correspond to the arrangements in Table 2.

ラ配置の評価値は低いという関係が見られる．このことから，提案手法を用いて評価値からカメラ配置を決定することは，姿勢パラメータ推定精度の向上にも有効であることがうかがえる．

## 6. む す び

複数カメラを用いた物体認識に関して，精度良く認識を行うためのカメラ配置の決定方法を提案した．提案手法は，適切なカメラ配置を評価関数によって決定するものであり，そのような評価関数の一例として，多様体間距離に基づくシンプルな評価関数を定義した．その評価関数を用いて，カメラが2台の場合と3台の場合に対するカメラ配置の決定に関する実験を行い，実験の結果から，提案手法が認識性能並びに姿勢パラメータ推定精度の向上に有効であることを示した．

提案手法の意義は，認識率の高いカメラ配置を得るために多様体間距離の最小値を評価値として採用したこと，固有空間上の多様体を用いることによって任意のカメラ配置の評価を可能としたことにある．

今後は，カメラ配置決定のためのより正確で有用な評価関数の設計，学習サンプル数が変化した場合の影響の調査を検討している．また，与えられた物体セットに対して十分な認識性能に達するためのカメラ台数をカメラ配置の評価値を用いて決定する手法の検討も行いたい．

謝辞 日ごろより熱心に御討論頂く名古屋大学村瀬研究室諸氏に感謝する．本研究の一部は日本学術振興会科学研究費補助金，21世紀COEプログラム「社会情報基盤のための音声・映像の知的統合」による．

本研究では，画像処理にソフトウェアライブラリ MIST (<http://mist.suenaga.m.is.nagoya-u.ac.jp/>) を使用している．

## 文 献

[1] H. Murase and S.K. Nayar, "Visual learning and recognition of 3-D object recognition from appearance," *Int. J. Comput. Vis.*, vol.14, pp.5-24, 1995.

[2] H. Murase and S.K. Nayar, "Illumination planning for object recognition using parametric eigenspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.16, no.12, pp.1218-1227, 1994.

[3] H. Tanaka, I. Kitahara, H. Saito, H. Murase, K. Kogure, and N. Hagita, "Dynamic visual learning for people identification with sparsely distributed multiple surveillance cameras," *Proc. SCIA2005*, pp.130-140, 2005.

[4] H. Borotschnig, L. Paletta, M. Prantl, and A. Pinz,

"Active object recognition in parametric eigenspace," *Proc. BMVC'98*, vol.2, pp.629-638, 1998.

[5] O. Yamaguchi and K. Fukui, "'Smartface' — A robust face recognition system under varying facial pose and expression," *IEICE Trans. Inf. & Syst.*, vol.E86-D, no.1, pp.37-44, Jan. 2003.

[6] F.G. Callari and F.P. Ferrie, "Autonomous recognition: Driven by ambiguity," *Proc. CVPR'96*, pp.701-707, 1996.

[7] A. Selinger and R.C. Nelson, "Appearance-based object recognition using multiple views," *Proc. CVPR2001*, vol.1, pp.905-911, 2001.

[8] G. Shakhnarovich, L. Lee, and T. Darrell, "Integrated face and gait recognition from multiple views," *Proc. CVPR2001*, vol.1, pp.39-46, 2001.

## 付 録

### カメラ配置と画像の例

物体セット WB に対して3台のカメラを用いた場合に評価値が高かったカメラ配置と評価値が低かったカメラ配置の例を図 A・1 に示す．また，それぞれのカメラ配置から取得された画像の例を図 A・2 と図 A・3 に示す．

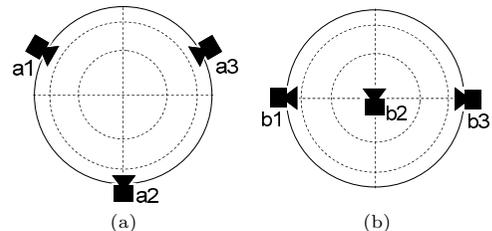


図 A・1 物体セット WB に対して3台のカメラを用いた場合の評価値が高かったカメラ配置 (a) と低かったカメラ配置 (b) の例

Fig. A・1 Examples of camera arrangements for WB with three cameras: (a) showed a high evaluation value and (b) showed a low evaluation value.

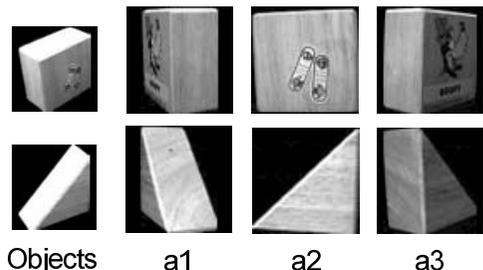


図 A・2 図 A・1 のカメラ配置 (a) から得られる画像の例

Fig. A・2 Examples of images captured from the camera arrangement (a) in Fig. A・1.

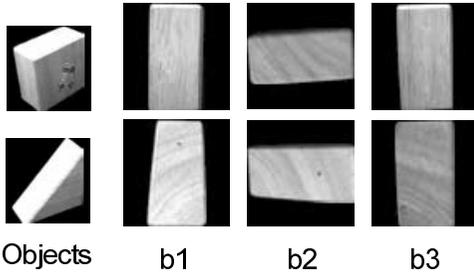


図 A-3 図 A-1 のカメラ配置 (b) から得られる画像の例  
Fig. A-3 Examples of images captured from the camera arrangement (b) in Fig. A-1.

(平成 18 年 10 月 6 日受付, 19 年 2 月 8 日再受付)



### 高橋 友和 (正員)

平 9 茨城大・工・情報卒。平 12 同大大学院理工学研究科博士前期課程了。平 15 同研究科博士後期課程了。博士(工学)。同年より 2 年間名古屋大学大学院情報科学研究科 COE 研究員。平 17 より日本学術振興会特別研究員, 現在に至る。三次元物体認識並びに三次元画像処理の基礎研究に興味をもつ。画像電子学会会員。



### Lina (学生員)

In 2001, bachelor degree from Department of Computer Science, Faculty of Engineering, Tarumanagara University, Indonesia. In 2004, master degree from Faculty of Computer Science, University of Indonesia. From 2006 until now, doctoral student at Graduate School of Information Science, Nagoya University, Japan. My field of study is 3D object recognition.



### 松ヶ野修功

平 15 名大・工・情報卒。平 17 同大大学院情報科学研究科メディア科学専攻修士課程了。同年トヨタ自動車(株)入社, 現在に至る。



### 井手 一郎 (正員)

平 6 東大・工・電子卒。平 8 同大大学院工学系研究科情報工学専攻修士課程了。平 12 同研究科電気工学専攻博士課程了。博士(工学)。同年国立情報学研究所助手。平 16 より名古屋大学大学院情報科学研究科助教授, 情報・システム研究機構国立情報学研究所客員助教授(連携)兼任。この間, 平 14~16 総合研究大学院大学数物科学研究科助手併任, 平 17 フランス情報学・統計システム研究所(IRISA)招聘教授。映像メディア処理全般に興味をもっている。平 7 第 51 回情報処理学会全国大会奨励賞受賞。人工知能学会, 情報処理学会, IEEE Computer Society, ACM 各会員。



### 目加田慶人 (正員)

平 3 名大・工・情報卒。平 8 同大大学院博士課程了。平 8 宇都宮大工学部情報工学科助手, 平 13 名大大学院工学研究科情報工学専攻助教授, 平 16 中京大学生命システム工学部教授, 平 18 より名大大学院工学研究科特任教授を兼任。画像処理・パターン認識とその医用応用に関する研究に従事。博士(工学)。平 10 日本医用画像工学会論文賞, 2002 FPGA/PLD Design Conference 審査委員特別賞, MIRU2004 優秀論文賞, 平 17 CADM 館野賞。IEEE, 日本医用画像工学会, コンピュータ支援画像診断学会各会員。



### 村瀬 洋 (正員)

昭 53 名大・工・電気卒。昭 55 同大大学院修士課程了。同年日本電信電話公社(現 NTT)入社。平 4 から 1 年間米国コロロンビア大客員研究員。平 15 から名古屋大学大学院情報科学研究科教授, 現在に至る。文字・図形認識, コンピュータビジョン, マルチメディア認識の研究に従事。工博。昭 60 本会学術奨励賞, 平 6 IEEE-CVPR 最優秀論文賞, 平 7 情報処理学会山下記念研究賞, 平 8 IEEE-ICRA 最優秀ビデオ賞, 平 13 高柳記念奨励賞, 平 13 本会ソサエティ論文賞, 平 14 本会業績賞, 平 15 文部科学大臣賞, 平 16 IEEE Trans. MM 論文賞, ほかに受賞。情報処理学会, IEEE 各会員。