

## 論 文

## 形状空間の幾何学的な関係に基づく三次元物体認識

五十嵐洋介<sup>†\*a)</sup> 福井 和広<sup>†b)</sup>

## 3D Object Recognition Based on Canonical Angles between Shape Spaces

Yosuke IGARASHI<sup>†\*a)</sup> and Kazuhiro FUKUI<sup>†b)</sup>

あらまし 時系列画像において抽出・追跡された特徴点集合の座標情報から因子分解法によって得られる形状空間は、特徴点集合の三次元的な幾何関係を表現しており、視点や対象物の動きに対して不変である。本論文ではこの形状空間の特性に着目し、形状空間同士の構造的な類似度に基づく物体認識のフレームワークを提案する。形状空間は特徴点数を次元とするベクトル空間中の三次元部分空間として得られるので、二つの形状空間の構造的な類似度は両者のなす正準角によって測ることができる。しかしながら、この類似度が意味をもって定義されるためには、二つの特徴点集合同士で特徴点に対応づけられている必要がある。この問題に対して、形状空間に付随して得られる直交射影行列を並べ換えることで対応付けを行う方法を提案する。提案する物体認識のフレームワーク及び対応付け法の有効性は、合成データを用いた性能評価、及びホクロなどの微小特徴に基づく顔認識実験により確認する。

キーワード 物体認識, 形状空間, 正準角, 因子分解法

## 1. ま え が き

本論文では、因子分解法によって得られる形状空間に基づいて、二つの特徴点集合間の幾何学的な類似度を測る方法を提案する。更にこの類似度を用いて、新しい三次元物体認識のフレームワークを構築する。提案する類似度に意味をもたせるためには、特徴点集合間で特徴点の対応付けが行われている必要がある。これに対応するために、形状空間に付随して得られる直交射影行列を並べ換えることで対応付けを行う方法を併せて提案する。

時系列画像から物体の三次元形状を復元する方法として、因子分解法 [1] がよく知られている。因子分解法では、まず対象物体の時系列画像から複数の特徴点の検出・追跡を行い、それらの座標を並べた計測行列  $\mathbf{W}'$  を生成する。次に  $\mathbf{W}'$  を特異値分解によって、カメラ運動を表現する行列  $\mathbf{M}$  と、物体の三次元形状を

表現する行列  $\mathbf{S}$  の積に分解する。

$\mathbf{S}$  の行ベクトルは、特徴点数を次元とするベクトル空間において、三次元部分空間である形状空間を張る。形状空間はアフィン投影モデルのもとで、カメラや物体の位置・姿勢の変化に不変である [2]。

これまで形状空間は、複数物体の運動分離 [2]~[4] や逐次型の因子分解法 [5] など、主に三次元形状復元の観点から利用されてきた。これに対して本論文では、形状空間が各特徴点の三次元的な相対位置関係（密に特徴点を取る場合には対象物体の三次元形状）を表し、更に部分空間である形状空間同士の構造的な類似度が両者のなす正準角 [6] により測れることに着目して、物体認識への適用を試みる。提案法の大きな特徴は陽に三次元復元を行わずに、二次元情報のみから三次元物体同士の形状類似度が得られる点にある<sup>(注1)</sup>。

図 1 に提案する特徴点集合間の構造的な類似度の概念を示す。まず各物体を撮影した多視点動画像から特徴点を追跡し、因子分解法によってそれぞれ形状空間を求める。次に二つの形状空間のなす正準角を求め、

<sup>†</sup> 筑波大学大学院システム情報工学研究科, つくば市  
Graduate School of Systems and Information Engineering,  
University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-  
8573 Japan

\* 現在, キヤノン株式会社

a) E-mail: igarashi@cvtlab.cs.tsukuba.ac.jp

b) E-mail: kfukui@cs.tsukuba.ac.jp

(注1): 文献 [7] では、与えられた二つの特徴点集合の三次元座標により定まる形状空間同士の距離に基づいて、特徴点集合間の三次元構造の一致度を測る方法が提案されているが、提案法は三次元座標を与える必要がない点で上記方法とは大きく異なる。

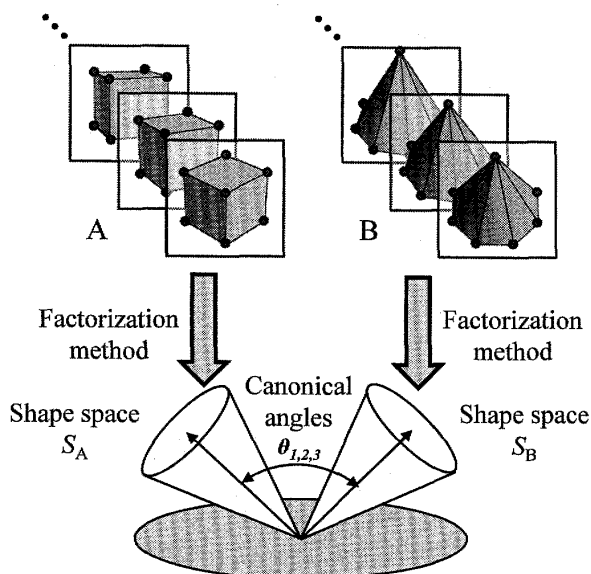


図1 提案する類似度の概念図：各物体上の特徴点集合の幾何情報を形状空間で表現し、両者のなす正準角を両者の類似度とする

Fig. 1 Concept of proposed similarity based on canonical angles between shape spaces.

これを用いて特徴点集合間の構造的な類似度を定義する。

このように正準角によって形状空間同士の類似度を測るためには、形状空間が計測行列における特徴点の並び順に依存して変化してしまう問題に対処する必要がある。本論文では、形状空間に付随して得られる直交射影行列の行と列を入れ換えることで、特徴点の対応付けを行う方法を提案する。

形状空間に基づく特徴点对応付けの従来法として、形状空間に対応する直交射影行列に対してピボット選択付きQR分解を適用し、得られるQ行列を並べ換える方法[8]が提案されている。この方法は、行列の並べ換えという点で提案法と類似したアプローチであるが、提案法は直交射影行列を直接並べ換えるのに対して、この方法ではQ行列を、形状空間を張る正規直交基底の組とみなして並べ換える。両者の特性を比較すると、特徴点の位置がノイズにより変動した際、Q行列は行列の分解操作を伴う分、直交射影行列に比べてより大きな影響を受ける可能性がある。これを検証するために、提案法とQR分解に基づく対応付け法を比較評価する。

提案法は汎用性の高い特徴点ベースの物体認識法として三次元体認識、ロボットビジョンにおける環境理解、車載画像処理、三次元物体データベースの類似形状検索などの様々な問題に適用可能と考えられるが、

本論文ではその中の一つとして、シミやホクロなどの微小特徴を用いた顔認識を取り上げ、提案法の有効性を示す。

以下、まず2.で形状空間の生成法とその特性について説明する。3.では、提案する特徴点集合間の類似度と特徴点の対応付け方法について述べる。4.では提案法の有効性を検証するために、合成データを用いた対応付けの評価と、顔画像を用いた認識実験を行う。5.はまとめである。

## 2. 形状空間の生成

本章では、因子分解法による形状空間の生成法と形状空間の特性について説明する。

TomasiとKanadeが提案した因子分解法は、時系列画像に対して特徴点群の追跡を行い、カメラ運動と形状情報(特徴点集合の三次元相対位置)を獲得する方法である[1]。いくつかのカメラモデルに対する方法が提案されているが、本論文ではアフィン投影モデルを仮定して議論を進める。

因子分解法では、まず視点を変えながら対象物体を撮影した $F$ 枚の連続画像に対して、 $P$ 個の特徴点を追跡する。特徴点 $p$ の第 $f$ 画像上での座標 $(x_{fp}, y_{fp})$ を用いて、行列 $\mathbf{W} \in \mathbb{R}^{2F \times P}$ を以下のように定義する。

$$\mathbf{W} = \begin{pmatrix} x_{11} & \dots & x_{1P} \\ y_{11} & \dots & y_{1P} \\ \vdots & \ddots & \vdots \\ x_{F1} & \dots & x_{FP} \\ y_{F1} & \dots & y_{FP} \end{pmatrix} \quad (1)$$

$\mathbf{W}$ の各列は各特徴点の運動軌跡を表す。

次に、 $\mathbf{W}$ の各行ごとの平均を求め、 $\mathbf{W}$ の各要素からこの行平均を引いた行列 $\mathbf{W}'$ を作る。 $\mathbf{W}'$ は計測行列と呼ばれ、以下のように運動の情報のみをもつ行列 $\mathbf{M}$ と形状の情報のみをもつ行列 $\mathbf{S}$ の積で表される[1]。

$$\mathbf{W}' = \mathbf{M}\mathbf{S} \quad (2)$$

ここで、 $\mathbf{S}$ は世界座標系における特徴点 $p$ の三次元座標 $\mathbf{r}_p = (x_p \ y_p \ z_p)^T$  ( $p = 1, \dots, P$ )を並べた行列である。ただし $\sum_{p=1}^P \mathbf{r}_p = \mathbf{0}$ である。 $\mathbf{M}$ は、 $\mathbf{r}_p$ を第 $f$ 画像に投影する $2 \times 3$ 行列 $\mathbf{M}_f$ を並べた行列である。

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1^T & \dots & \mathbf{M}_F^T \end{pmatrix}^T \in \mathbb{R}^{2F \times 3} \quad (3)$$

$$\mathbf{S} = \begin{pmatrix} \mathbf{r}_1 & \dots & \mathbf{r}_P \end{pmatrix} \in \mathbb{R}^{3 \times P} \quad (4)$$

したがって、 $\mathbf{W}'$  を式 (2) の形で分解できれば、運動と形状の情報を分離して取り出すことができる。

$\mathbf{W}'$  のランクが 3 であることを考慮すると、 $\mathbf{W}'$  は特異値分解により次のように分解できる。

$$\mathbf{W}' = \mathbf{U}_{2F \times 3} \mathbf{\Sigma}_{3 \times 3} \mathbf{V}_{P \times 3}^T \quad (5)$$

ここで  $\hat{\mathbf{M}} = \mathbf{U}\mathbf{\Sigma}$ ,  $\hat{\mathbf{S}} = \mathbf{V}^T$  とすると、 $\mathbf{W}'$  の一つの分解である  $\mathbf{W}' = \hat{\mathbf{M}}\hat{\mathbf{S}}$  が得られる。しかしながら、 $\mathbf{W}'$  の分解は一意ではない。なぜなら、任意の  $3 \times 3$  正則行列  $\mathbf{A}$  に対して  $\mathbf{M}^* = \mathbf{M}\mathbf{A}^{-1}$ ,  $\mathbf{S}^* = \mathbf{A}\mathbf{S}$  とすると、 $\mathbf{M}^*\mathbf{S}^* = \mathbf{M}\mathbf{A}^{-1}\mathbf{A}\mathbf{S} = \mathbf{W}'$  が成り立ち、 $\mathbf{M}^*$  と  $\mathbf{S}^*$  もまた、式 (2) を満たす  $\mathbf{W}'$  の分解となるからである。

$\mathbf{M}^*$  と  $\mathbf{S}^*$  はそれぞれ運動行列及び形状行列と呼ばれ、 $\mathbf{S}^*$  の行ベクトルが張る三次元部分空間は形状空間 (shape space) と呼ばれる [5]。形状空間は各特徴点の相対的な位置関係の情報を反映した部分空間となる。ここで注意すべきは、先に述べたように、与えられた特徴点集合に対して形状行列は任意性をもつが、形状空間は世界座標系のとり方に依存せず一意に決まる点である。

形状空間を張る正規直交基底は無数に存在するが、式 (5) の  $\mathbf{V}$  の列ベクトルはその一つである。そこで本論文では以降、 $\mathbf{V}$  を用いて議論する。

また提案フレームワークには適用できないが、CG データのように各特徴点の三次元座標が与えられている場合には、形状空間はこれらの座標を列ベクトルとして並べた行列  $\mathbf{S}'$  の行空間として得られ、形状空間を張る正規直交基底は行列  $\mathbf{S}'$  の行ベクトルに対して Gram-Schmidt の直交化法を適用することで求まる。この場合も世界座標系のとり方に依存して行列  $\mathbf{S}'$  が変化するのに対して、それから求まる形状空間は一意に決まる。なお、この生成法は、後で述べる直交射影行列を用いた特徴点の対応付け処理に限れば、適用可能である。

### 3. 提案法

本章では、まず直交射影行列を用いた特徴点の対応付け方法について提案する。次に特徴点に対応付けし得られる形状空間同士のなす正準角に基づいて物体認識を行う方法を提案する。

#### 3.1 直交射影行列に基づく特徴点対応付け

形状空間は、式 (5) によって得られる  $\mathbf{V}$  の列空間 (あるいは  $\mathbf{V}^T$  の行空間) として得られる。このとき、

$\mathbf{V}$  の行の並び順は、特徴点を並べる順番によって決まり、同一の特徴点集合から得られた二つの形状空間でも、特徴点の並び順が同一でない場合には一致しない。よって形状空間間の意味のある類似度を求めるためには、二つの特徴点集合間で特徴点に対応づけられている必要がある。この対応付けを行うために、 $\mathbf{V}$  から求まる直交射影行列  $\mathbf{V}\mathbf{V}^T$  の行と列を並べ換えることによって特徴点の対応付けを行う方法を提案する。

##### 3.1.1 直交射影行列の特性

二つの特徴点集合間で互いの特徴点を対応づける問題は、二つの形状行列において列同士を対応づけることと等価である。しかし形状行列は 2. で述べたように、同じ特徴点集合に対しても一意には定まらないため、形状行列の列同士を直接比較して対応付けを行うことはできない。

これに対して、形状空間は特徴点の順番が決まれば一意に定まり、形状空間に付随する直交射影行列もまた一意に定まる。したがって、同一物体から同じ特徴点集合が得られ、かつ各特徴点と同じ順番で並んでいれば、それらから得られる  $\mathbf{V}_A$ ,  $\mathbf{V}_B$  が  $\mathbf{V}_A \neq \mathbf{V}_B$  であっても、直交射影行列に関しては、 $\mathbf{Q} = \mathbf{V}_A\mathbf{V}_A^T = \mathbf{V}_B\mathbf{V}_B^T$  が成り立つ。この特性を利用して、形状行列同士を比較するのではなく、直交射影行列の行を比較することで特徴点対応付けを実現する。

##### 3.1.2 特徴点の入換えに伴う直交射影行列の変化

ここでは、特徴点の並び順を入れ換えた場合、形状行列と直交射影行列がそれぞれどのように変化するかを述べた上で、提案する特徴点対応付けの基本アイデアを説明する。

形状空間の正規直交基底として  $\mathbf{V}$  が与えられたとき、 $\mathbf{V}$  の  $i$  行と  $j$  行を入れ換わった行列を  $\mathbf{V}'$  とし、 $\mathbf{V}$  と  $\mathbf{V}'$  から得られる直交射影行列をそれぞれ  $\mathbf{Q}$ ,  $\mathbf{Q}'$  とすると、以下の式が成り立つ。ここで  $\mathbf{I}_p$  は、単位行列の  $i$  行と  $j$  行を入れ換えて生成される置換行列である。

$$\begin{aligned} \mathbf{V}' &= \mathbf{I}_p \mathbf{V} \\ \mathbf{Q}' &= \mathbf{V}'\mathbf{V}'^T = \mathbf{I}_p \mathbf{V}\mathbf{V}^T \mathbf{I}_p^T = \mathbf{I}_p \mathbf{Q} \mathbf{I}_p^T \end{aligned} \quad (6)$$

上式は、 $\mathbf{V}$  の行の入換え、すなわち形状行列の列を入れ換える変換が、直交射影行列において行と列を同時に入れ換える変換であることを意味する。そのため、直交射影行列を比較して対応付けを行う場合、二つの直交射影行列の行同士、または列同士のいずれかについて対応を求めればよい。

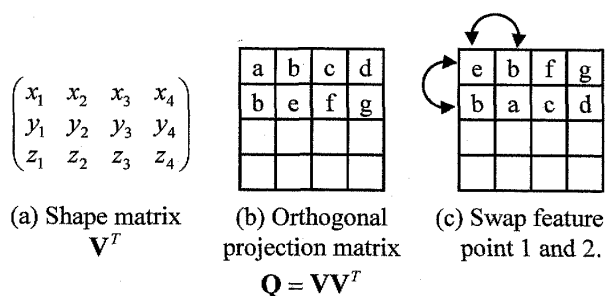


図2 特徴点の入換えに伴う、直交射影行列の行と列の入換わり ( $P = 4$ )  
 Fig.2 Example of swapping rows and columns by swapping feature points. ( $P = 4$ )

以下では図2に沿って、4点の特徴点を得られている場合を例に、特徴点の順番の入換えによって生じる直交射影行列の変化を示す。図2(a)に形状行列  $V^T$ , (b)に  $V$  から計算される直交射影行列  $Q$  を示す。

まず特徴点1と特徴点2が入れ換わると、 $V^T$  では1列と2列が入れ換わる。これに対応して直交射影行列  $Q$  では、1行と2行、また1列と2列が入れ換わり、(c)の行列が得られる。ここで(b)の1行と(c)の2行、(b)の2行と(c)の1行を比較すると、各行内で要素の順序は変わっているが、その行を構成する要素自体は変化していないことが分かる。この特性は、入れ換わる点の対が増えても同様に成立する。

この特性に基づけば、比較する二つの直交射影行列について各行の要素をソートすることで、行同士を比較することが可能となる。これにより、特徴点の対応付けの問題は、二つの直交射影行列において、ソートした行ベクトルの対で  $L1$  距離が最小になる組合せを見つける問題となる。

### 3.1.3 特徴点对応付けの流れ

提案する対応付けの流れを、図3に沿って説明する。図3では理解を容易にするために、直交射影行列ではなく対称行列を用いて模擬的に説明する。アルゴリズムの詳細については、付録を参照すること。

(a) 対応付けの対象となる二つの直交射影行列を、それぞれ行列  $Q_{A(1)}$ ,  $Q_{B(1)}$  とする。

(b)  $Q_{A(1)}$ ,  $Q_{B(1)}$  の各行の要素を昇順でソートした行列を  $Q'_{A(1)}$ ,  $Q'_{B(1)}$  として、行ベクトル間の  $L1$  距離が最も小さくなる組を1組目の対応とする。ここでは、 $Q'_{A(1)}$  の4行と  $Q'_{B(1)}$  の1行が対応づく。

(c) 対応づけられた4行と1行をそれぞれ削除し、4列と1列の列ベクトルを対応づける。この行列を次のステップの行列  $Q_{A(2)}$ ,  $Q_{B(2)}$  とする。

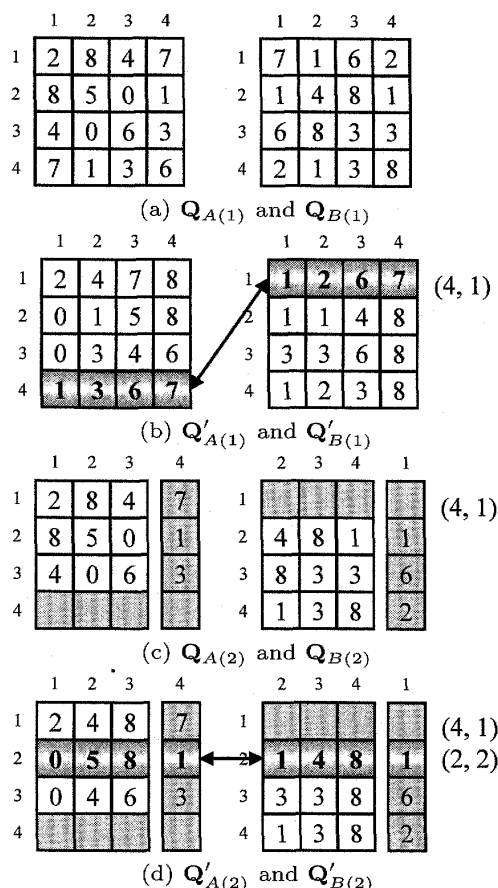


図3 提案法による対応付けの例  
 Fig.3 Example of the proposed matching process.

(d)  $Q_{A(2)}$ ,  $Q_{B(2)}$  の各行に対して、いまだ対応づけられていない要素のみをソートした後に、対応済みの列ベクトルの要素を含めた行ベクトル間距離が最小となる組を探索する。ここでは  $Q'_{A(2)}$  の2行と  $Q'_{B(2)}$  の2行が対応づけられる。以降(c), (d)をすべての行ベクトルが対応づけられるまで繰り返す。

### 3.2 正準角に基づく形状類似度

形状空間同士の類似度は、両者のなす角度である正準角[6]によって定量的に測ることができる。 $M$ 次元部分空間  $S_1$  と  $N$ 次元部分空間  $S_2$  ( $M \leq N$ ) に対して、 $M$ 個の正準角 ( $\theta_1 \leq \theta_2 \leq \dots \leq \theta_M$ ) が以下のように定義できる。

$$\cos^2 \theta_i = \max_{\substack{u_i \perp u_j, v_i \perp v_j \\ 1 \leq i, j \leq M, i \neq j}} \frac{(u_i \cdot v_i)^2}{\|u_i\|^2 \|v_i\|^2} \quad (7)$$

ここで  $(\cdot)$ ,  $\|\cdot\|$  はそれぞれ内積、ノルムを表し、 $u_i \in S_1$ ,  $v_i \in S_2$ ,  $\|u_i\| \neq 0$ ,  $\|v_i\| \neq 0$  である。 $S_1$ ,  $S_2$  を張る正規直交基底をそれぞれ並べた行列を  $\Phi = (\phi_1 \dots \phi_M)$ ,  $\Psi = (\psi_1 \dots \psi_N)$  とすると、第  $i$  正準角  $\theta_i$  の余弦  $\cos \theta_i$  は、 $\Phi^T \Psi$  あるいは  $\Psi^T \Phi$  の

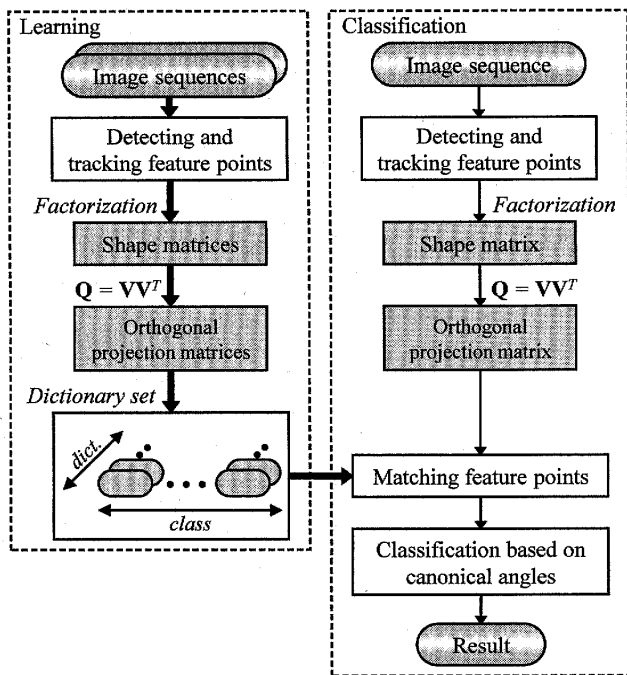


図4 提案する物体認識の流れ

Fig. 4 The flow of object recognition process based on proposed similarity.

$i$  番目に大きい特異値  $\sigma_i$  として得られる [9].

三次元部分空間である形状空間同士の関係は、3 個の正準角によって規定されるので、これらを用いて類似度  $Sim$  を以下のように定義する。

$$Sim = \frac{1}{3} \sum_{i=1}^3 \cos^2 \theta_i \quad (8)$$

類似度  $Sim$  は正規化量である。二つの形状空間が一致しているときは 1.0 となり、両者が完全に直交しているときに 0.0 となる。

### 3.3 提案法に基づく物体認識

図4に、提案法を用いた物体認識の流れを示す。提案するフレームワークは、大きく学習処理と識別処理に分かれる。

学習時には、次の手順で辞書を生成する。まず静止している対象物体を視点を変化させながら撮影し、多視点動画像を取得する。あるいは逆に静止カメラに対して物体を運動させてもよい。次に、この多視点動画像から特徴点を検出・追跡する。得られた特徴点集合から因子分解法によって形状行列を求め、これから直交射影行列を計算して辞書として保存する。

特徴点の検出・追跡は一般的に不安定なので、同じ物体を撮影した動画像であっても、照明や視点の変化によって得られる特徴点集合が変化すると考えられる。

そこで、このような変動を吸収するために、物体クラスごとに多様な照明条件・視点から撮影した複数の動画像を用意し、これらから複数の直交射影行列を求めておく。つまり各クラスごとに複数の直交射影行列が辞書として登録されることになる。

識別時には、学習時と同様に入力された多視点動画像から形状行列を求め、更に直交射影行列を求める。次に入力直交射影行列と各辞書の直交射影行列から、3. で述べた提案法を用いて特徴点の対応付けを行う。両者を対応づけた後に、入力形状空間と辞書形状空間の間の正準角から式(8)の類似度を求める。このとき、クラスごとに辞書が複数用意されているので、各クラスに対して複数の類似度が求まる。その中で最大の類似度を与える辞書データが属するクラスへ入力データを分類する。

提案フレームワークを物体認識に適用する際、アフィン変形により三次元配置が一致する特徴点集合同士は区別できないことに注意が必要である。つまり、特徴点集合はその物体固有であり、他の物体から得られた特徴点集合とアフィン変形により一致しない必要がある。評価実験では、このような例として、顔面上のシミ、ホクロなどの微小顔特徴点集合を用いた顔認識を取り上げる。

## 4. 実験

提案法の性能を評価するために、2種類の実験を行った。実験1では、合成データを用いて特徴点对応付けの精度を評価した。実験2では、顔画像から得られるホクロなどの特徴点集合に対して提案する物体認識のフレームワークを適用し、その有効性を検証した。特徴点の対応付け方法としては、提案法と1. で述べた直交射影行列のQR分解に基づく方法[8]を適用して、両者の性能を比較した。

### 4.1 実験1: 合成データによる特徴点对応付け

提案する対応付けの精度とノイズに対する頑健性について、合成データを用いて評価した。

まず、半径1の球表面に、 $P$ 点の特徴点をランダムに配置した特徴点集合を生成する。次に、この特徴点集合の各特徴点の座標に、標準偏差 $\sigma$ のガウス雑音を加えた特徴点集合を生成する。これら二つの特徴点集合について特徴点の対応付けを行った。評価では、 $P$ と $\sigma$ をそれぞれ変化させながら対応付けの正解率の変化を調べ、更に $P$ ごとに処理時間を測った。正解率は、各パラメータ当たり200回の試行の平均値とし

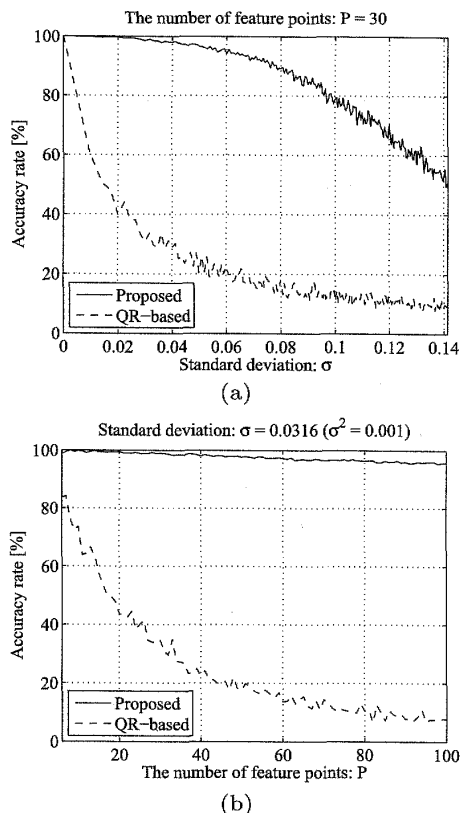


図5 合成データにおける特徴点对応付けの性能評価  
Fig. 5 Performances of proposed matching method.

た. 処理時間の計測は, Intel Xeon プロセッサ E5540 (4 core, 2.53 GHz), MATLAB 7.9.0 (R2009b) によって行った.

図5に, 対応付け結果のグラフを示す. (a)は $P=30$ として $\sigma$ を変化させたときの正解率の変化, (b)は $\sigma=0.0316$ として, $P$ を変化させたときの正解率の変化である. また図6に, $P=30$ , $\sigma=0.1$ における対応付け結果の一例を示す. 二つの特徴点集合をそれぞれ○と△で描画し, 正解の対応を直線で, 誤った対応を破線で示した. 表1に処理時間を示す.

図5から, 提案法が従来法に比べて, ノイズに対して極めて頑健であることが分かる. 図5(a)では, $\sigma=0.1$ という高ノイズに対しても80%の精度で対応付けに成功しており, 図5(b)では, $P=100$ 点で約95%の対応付けに成功している.

表1から, 提案法が従来法に比べて計算量が多いことが分かる. しかしながら図5が示すように, 従来法は特徴点数が多くなるとわずかなノイズによって正解率が急激に低下する. したがって, 総合的に評価すると, 提案法の従来法に対する優位性は大きいと考える. 以上により, 提案法により頑健かつ高精度な特徴点の

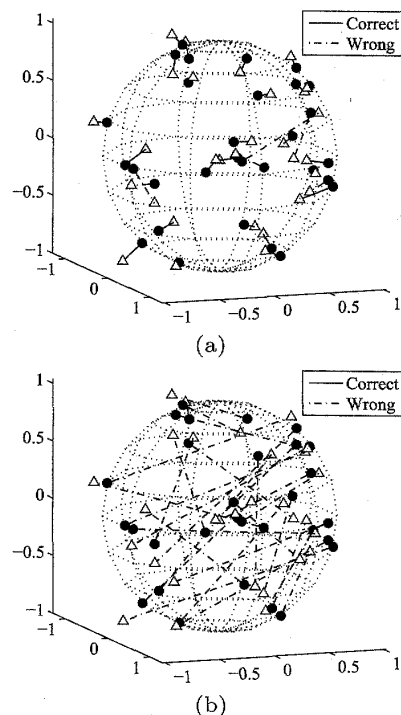


図6 合成データにおける対応付け結果の一例 ( $P=30$ ,  $\sigma=0.1$ ) (a) 提案法, (b) 従来法 [8]

Fig. 6 Example of sphere data matching ( $P=30$ ,  $\sigma=0.1$ ). (a) Proposed method and (b) QR-based method [8].

表1 特徴点对応付けの処理時間 [ms]

Table 1 Processing times of the matching methods.

Method	P			
	10	30	50	100
Proposed	1.8	40.1	145.6	765.7
QR-based	0.6	2.6	4.2	14.8

対応付けが可能であることが確認できた.

## 4.2 実験2: 顔画像認識

提案する物体認識のフレームワークを顔画像認識の問題に適用し, その有効性について検討した. 従来の代表的な顔認識手法 [10] は, 顔画像全体の情報を用いるアピアランス空間の方法であり, 既に実用化レベルにある. しかしながら, 親子や兄弟・双子など, 人が見ても間違ような人物同士を識別することは原理的に困難である. このような見た目が非常に類似している人物を識別する方法として, 近年, 顔表面のホクロやシミといった微小特徴の情報を認識に利用した方法の有効性が報告されている [11]~[13]. しかし, これらの方法は, 複数の特徴点の二次元的な位置関係に基づく識別であり, 特徴点集合の三次元的な幾何関係を考慮した識別にはなっていない. そこで本実験では, 顔画像から微小特徴を検出し, それらの三次元的な位

置関係に基づいた顔認識を試みた。正面顔に対する顔の奥行は5センチメートル程度であり、顔から50センチメートル程度離れた位置から撮影を行えば、アフィン投影モデルで十分に近似が可能である。

本実験が実験1と大きく異なるのは、評価において特徴点検出・追跡の不安定性を考慮する必要がある点である。これが問題となる状況として次の二つが考えられる。

[状況1] 特徴点を追跡する際には、必ず追跡失敗や誤追跡が生じる。これらの変動は実験1で評価したノイズとは特性が異なるため、改めて評価する必要がある。

[状況2] たとえ同一人物の顔画像からでも、照明変動や顔向きの変化により、同じ微小特徴を特徴点として検出することは難しい。この問題を完全に解決することは困難であり、同一人物に対しても対応付けの相手が存在しない特徴点の混入は避けられない。実際的な問題への適用を考えると、このような未対応点の混入によって、性能がどの程度劣化するかを評価しておく必要がある。

以下では状況1, 状況2を想定した評価実験を行った。

実験条件：被験者は28人とした。被験者はカメラから約0.5メートル離れた椅子に座り、ディスプレイ上で円運動するマークに顔向きを追従させた。このとき、正面顔で検出される特徴点の隠れが生じない範囲で顔の回転を行った。1人当たりマークが1周する間に50枚撮影し、合計マーク20周分1000枚の連続画像を撮影した。画像の解像度は $1024 \times 768$ ピクセル、撮影速度は15フレーム毎秒とした。

特徴抽出：まず、各画像に対して顔を検出[14]した後で、更に目鼻検出[15]を行った。この結果を用いて、人物間で共通の特徴である目鼻の近傍領域を特徴点の検出・追跡の対象から除去した上で、残りの領域に対して円形分離度フィルタ[15]を適用した。円形分離度フィルタは、円の内側と外側の輝度値の統計的な分離度を出力するフィルタであり、画像中のホクロのような円形状を有する微小特徴を強調した分離度マップを生成できる。図7に、入力画像から検出された顔・目鼻領域と、円形分離度フィルタにより得られた分離度マップの一例を示す。この分離度マップに対して、KLT法[16]を適用して特徴点の検出・追跡を行った。

#### 4.2.1 [状況1] 特徴点の追跡失敗に対する頑健性

前述した状況1を模擬するために、次の手順で特徴

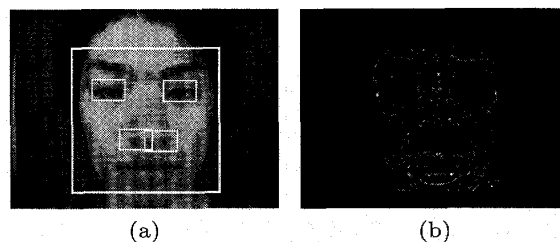


図7 特徴点検出・追跡のための前処理 (a) 入力画像に対する顔・目鼻検出, (b) 顔領域中の目鼻以外の領域に対して分離度フィルタを適用

Fig.7 Preprocessing for detecting and tracking feature points: (a) detected face and (b) separability degree map.

点を選択した。全連続画像1000枚の分離度マップに対してKLT法を適用し、特徴点を検出・追跡する。そして最後まで追跡された特徴点の中から、分離度が高い上位 $P$ 点を選択する。ここで注意すべきは、追跡は不完全であり、得られた特徴点集合には多くのノイズが含まれているということである。目視で確認したところ、追跡の途中で本来追跡すべき微小特徴から50ピクセルも離脱したり、誤った微小特徴を特徴点として追跡しているケースが多く見られた。

評価は以下の手順で行った。まず、人物ごとに特徴点の追跡結果を50枚ずつ20個に分割し、合計560個の画像セットを生成する。提案法の基本性能を確認するために、理想的な状況を想定して、顔の回転1周分のデータを1画像セットとする。次に、1個の画像セットを評価データとして取り出し、残りの画像セットをすべて辞書データとする。このとき、評価データが属するクラス(人物)は19個の辞書、それ以外のクラスでは20個の辞書をもつことになる。識別は式(8)で定義される類似度を用いた最近傍法により行う。これをすべての画像セットが評価データとして選ばれるまで繰り返した。

特徴点の対応付けには、提案法とQR分解に基づく方法[8]をそれぞれ適用した。評価尺度として、識別率とEER(Equal Error Rate)を用いた。EERは、FAR(他人受入率: False Acceptance Rate)とFRR(本人排他率: False Rejection Rate)が等しいときの誤識別率である。

図8に検出された30点の顔特徴点の例を示し、図9にある画像セットに含まれる顔画像の一部を示す。表2は識別に用いる特徴点数に対する識別性能を示している。識別率、EERの両方に関して、特徴点の対応付けに提案法を適用することで、従来法を適用した場合に



図 8 1000 枚の連続画像に対して追跡された特徴点の例 ( $P = 30$ )

Fig. 8 Examples of detected and tracked feature points through 1000 frames. ( $P = 30$ )

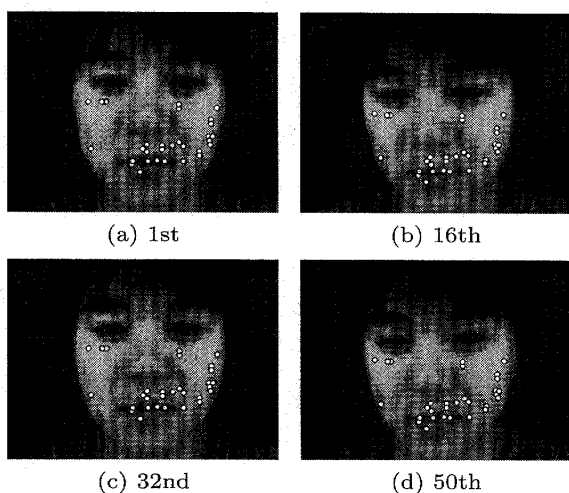


図 9 状況 1 における画像セットに含まれる顔画像の例  
Fig. 9 Examples included in a set of images in the case 1.

表 2 状況 1 における顔認識の結果

Table 2 Performances of face recognition in the case 1.

$P$	Recognition rate [%]		EER [%]	
	Proposed	QR-based	Proposed	QR-based
10	98.4	96.3	2.08	3.60
15	99.8	97.7	0.81	3.04
20	99.5	98.2	0.76	2.28
25	99.8	98.8	0.59	1.61
30	99.6	98.8	0.62	1.74
35	99.6	98.2	0.45	2.33

比べて性能が大幅に向上していることが分かる。また  $P = 35$  のときに識別率は 99.6%, EER は 0.45%と、良好な識別性能が得られていることから、提案した類似度が特徴点ベースの顔識別において有効に働くことが確認できた。

表 3 状況 2 における顔認識の結果 ( $P = 30$ )

Table 3 Performances of face recognition in the case 2. ( $P = 30$ )

$N$	RR [%]	EER [%]
0	99.6	0.62
5	99.2	0.99
10	98.1	1.80
15	97.2	2.68
20	96.6	2.68
25	96.2	2.90
30	96.1	2.94
35	95.9	3.13
40	96.9	2.90

表 4 辞書データ数の変化による識別性能の変化 ( $P = 30$ )

Table 4 Change of the performance with changing the number of dictionaries. ( $P = 30$ )

$N$	RR [%]			EER [%]		
	5	10	15	5	10	15
10	98.7	96.8	95.0	1.43	2.88	3.33
12	98.9	97.4	95.7	1.34	2.38	3.39
14	99.0	97.8	96.0	1.26	2.11	2.97
16	99.0	97.8	96.9	1.13	2.04	2.84
18	99.1	98.0	97.1	1.13	1.92	2.72
20	99.2	98.1	97.2	0.99	1.80	2.68

#### 4.2.2 [状況 2] 未対応点の混入に対する頑健性

状況 1 の実験で用いた  $P = 30$  点の特徴点に対して、対応点が存在しない特徴点 (画像セット内では全フレームで追跡が可能) を識別すべき特徴点集合に段階的に追加していくことで状況 2 を模擬する。なお、先に述べたように対応点が存在する特徴点も必ずしも安定した点ではなく、不完全な追跡による誤対応がかなり混入している。追加する特徴点はランダムに選択し、試行回数 5 回の平均を最終的な識別結果とした。追加する点数  $N$  を 0 点から 40 点まで変化させて、識別性能がどの程度劣化するかを調べた。特徴点の対応付けは提案法によって行い、識別は状況 1 と同様に、最近傍法を用いて行った。

表 3 に追加した特徴点数  $N$  と識別性能の関係を示す。  $N = 0$  のときが状況 1 の結果に対応する。追加される特徴点数  $N$  が増加するに伴い、識別性能は明らかに低下していく。しかしながら、  $P = 30$  点に対して  $N = 10$  点が追加された場合 (混入率:  $10/40 = 25\%$ ) でも識別率は 98.1%, EER は 1.80%と高い性能を維持している。

また表 4 に、  $P = 30$ ,  $N = 5, 10, 15$  において辞書データ数を変動させたときの識別性能の変化を示す。辞書数の増加に伴って性能が向上することが確認でき、



現実的な状況と考える  $N = 5$  (混入率 15%) においては、辞書数が 16 から 20 で性能が安定していることが分かる。

以上の結果から 30 点の特徴点に対して、混入率を 15% 程度に抑えれば、クラス当り 20 個の辞書でも実用的な性能が実現できることが分かった。この程度の混入率であれば、特徴点検出を工夫することで十分実現可能と考えられ、従来の顔画像全体を用いるアピアランス空間の方法とは異なるアプローチで、顔画像認識が実現できる可能性を示した意義は大きいと考える。

本実験データに対して、アピアランス空間の最も代表的な方法である固有顔法 [17] を適用した場合に、識別率は 99.8%, EER は 1.10% だった。これに対して、提案法の類似度と固有顔法の類似度の加算平均に基づいて識別を行うと、識別率が 99.8%, EER が 0.56% となり、識別性能の向上が確認できた。ただし識別時間については、固有顔法はクラス数に比例するのに対し、提案法は辞書データ数に比例し特徴点数の 3 乗に比例する。そのためリアルタイム性が必要となる場合には、固有顔法で誤識別を起こしやすい類似したクラス対のみに提案法を適用することになる。

## 5. むすび

本論文では、因子分解法によって得られる形状空間の特性を利用して、三次元空間における二つの特徴点集合同士の類似度を定義した。更にこの類似度に基づいて三次元物体認識を行うフレームワークを提案した。また提案法を実現するために必要な特徴点の対応付けを、形状空間から得られる直交射影行列を並べ換えることで解決する方法を提案した。

提案法の有効性は、二つの実験によって確認した。まず特徴点の対応付け方法について、ノイズを加えた合成データを用いて精度を評価した。その結果、従来法に比べて、提案法はノイズに対して非常に頑健であり、かつ高精度で対応付けが可能であることを確認した。また提案した物体認識のフレームワークについて、顔画像認識によって有効性を検証した。特徴点の追跡が極めて不安定な顔画像に対しても、提案する対応付けが有効であり、正準角による類似度によって良好な識別性能が得られることを確認した。更に同一人物において、検出された全特徴点数の 15% 程度の点で対応付けができなくても、実用的な識別性能を維持できる可能性があることが分かった。

今後は、提案法の識別性能を更に向上させるために、

識別に有効な特徴点を統計的に選択する方法について研究を進めていく。

## 文 献

- [1] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *Int. J. Comput. Vis.*, vol.9, no.2, pp.137-154, 1992.
- [2] J.P. Costeira and T. Kanade, "A multibody factorization method for independently moving objects," *Int. J. Comput. Vis.*, vol.29, no.3, pp.159-179, 1998.
- [3] N. Ichimura, "Motion segmentation based on factorization method and discriminant criterion," *International Conference on Computer Vision*, vol.1, pp.600-605, 1999.
- [4] K. Kanatani, "Motion segmentation by subspace separation and model selection," *International Conference on Computer Vision*, vol.2, pp.586-591, 2001.
- [5] T. Morita and T. Kanade, "A sequential factorization method for recovering shape and motion from image streams," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.19, no.8, pp.856-867, 1997.
- [6] C. Jordan, "Essai sur la géométrie à  $n$  dimensions," *Bulletin de la Société Mathématique de France*, vol.3, pp.103-174, 1875.
- [7] E. Begelfor and M. Werman, "Affine invariance revisited," *IEEE Conference on Computer Vision and Pattern Recognition*, vol.2, pp.2087-2094, 2006.
- [8] Z. Wang and H. Xiao, "Dimension-free affine shape matching through subspace invariance," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.2482-2487, 2009.
- [9] F. シャトラン (著), 伊理正夫, 伊理由美 (訳), 行列の固有値—最新の解法と応用, シュプリンガー・ジャパン, 2003.
- [10] K. Fukui and O. Yamaguchi, "Face recognition using multi-viewpoint patterns for robot vision," *International Symposium of Robotics Research*, pp.192-201, 2003.
- [11] 河原智一, 山口 修, 福井和広, "顔表面の微小特徴がなすグローバル構造を用いた人物認証," 第 5 回システムインテグレーション部門講演会, pp.619-620, 2004.
- [12] 浜田康志, 今岡 仁, "微小な構造特徴点の対応付けに基づく顔画像照合," 信学技報, PRMU2008-75, 2008.
- [13] J.S. Pierrard and T. Vetter, "Skin detail analysis for face recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1-8, 2007.
- [14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.511-518, 2001.
- [15] 福井和広, 山口 修, "形状抽出とパターン照合の組合せによる顔特徴点抽出," 信学論 (D-II), vol.J80-D-II, no.8, pp.2170-2177, Aug. 1997.
- [16] C. Tomasi and T. Kanade, "Detection and tracking of

point features,” Technical Report CMU-CS-91-132, Carnegie Mellon University, 1991.

- [17] M. Turk and A. Pentland, “Eigenfaces for recognition,” J. Cognitive Neuroscience, vol.3, no.1, pp.71-86, 1991.

## 付 録

### 特徴点对応付けのアルゴリズム

Algorithm 1 に対応付けのアルゴリズムを示す。  $N$  個の特徴点をもつ二つの集合から得られる形状空間の直交射影行列  $\mathbf{Q}_A$ ,  $\mathbf{Q}_B$  を入力すると、対応付けの結果である行列  $\mathbf{C}$  が出力される。  $\mathbf{C}$  の 1 列目に  $\mathbf{Q}_A$  に

対応する特徴点集合のインデックスが格納され、  $\mathbf{C}$  の 2 列目に  $\mathbf{Q}_B$  に対応する特徴点集合のインデックスが格納される。

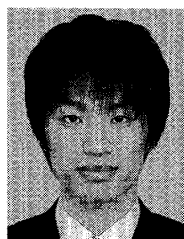
(平成 22 年 12 月 13 日受付, 23 年 3 月 4 日再受付)

### Algorithm 1 The proposed matching algorithm.

```

1: Input:  $N \times N$  Orthogonal projection matrices  $\mathbf{Q}_A, \mathbf{Q}_B$ 
2: Output:  $N \times 2$  Correspondence matrix  $\mathbf{C}$ 
3: // 対応付いていない行のインデックスを保持
4:  $\Phi = \{1, 2, \dots, N\}$ 
5:  $\Psi = \{1, 2, \dots, N\}$ 
6: for  $t = 1$  to  $N$  do
7:   //  $\mathbf{Q}_A, \mathbf{Q}_B$  の各行の要素を昇順ソート (図 3 (b), (d))
8:    $\mathbf{Q}'_A \leftarrow$  Sort the elements of  $\mathbf{Q}_A$  within each row
9:    $\mathbf{Q}'_B \leftarrow$  Sort the elements of  $\mathbf{Q}_B$  within each row
10:  // 未対応の行のみを対象に  $\mathbf{Q}'_A, \mathbf{Q}'_B$  の行ベクトル間
    距離を計算 (図 3 (b), (d))
11:  for  $i = 1$  to  $N - (t - 1)$  do
12:     $g \leftarrow \Phi_{[i]}$ 
13:     $\mathbf{u} \leftarrow$   $g$ th row of  $\mathbf{Q}'_A$ 
14:    for  $j = 1$  to  $N - (t - 1)$  do
15:       $h \leftarrow \Psi_{[j]}$ 
16:       $\mathbf{v} \leftarrow$   $h$ th row of  $\mathbf{Q}'_B$ 
17:      if  $t = 1$  then
18:         $d(i, j) = \sum_{k=1}^N |u_{[k]} - v_{[k]}|$ 
19:      else
20:        // 対応済みの列  $\mathbf{a}^{(\cdot)}, \mathbf{b}^{(\cdot)}$  の要素も含めて
         $L_1$  距離を計算 (図 3 (d))
21:         $d(i, j) = \sum_{k=1}^{N-(t-1)} |u_{[k]} - v_{[k]}|$ 
          +  $\sum_{k=1}^{t-1} |a_{[g]}^{(k)} - b_{[h]}^{(k)}|$ 
22:      end if
23:    end for
24:  end for
25:  // 行ベクトル間距離が最小になる組  $(i, j)$  を探索し, 保持
26:   $(r_A, r_B) = \underset{i, j}{\operatorname{argmin}} d(i, j)$ 
27:   $t$ th row of  $\mathbf{C} \leftarrow (\Phi_{[r_A]}, \Psi_{[r_B]})$ 
28:  //  $\mathbf{Q}_A, \mathbf{Q}_B$  の対応付いた列を  $\mathbf{a}^{(t)}, \mathbf{b}^{(t)}$  に移動
    (図 3 (c))
29:   $\mathbf{a}^{(t)} \leftarrow$   $r_A$ th column of  $\mathbf{Q}_A$ 
30:   $\mathbf{b}^{(t)} \leftarrow$   $r_B$ th column of  $\mathbf{Q}_B$ 
31:   $\mathbf{Q}_A \leftarrow$  Remove  $r_A$ th column of  $\mathbf{Q}_A$ 
32:   $\mathbf{Q}_B \leftarrow$  Remove  $r_B$ th column of  $\mathbf{Q}_B$ 
33:  //  $\Phi, \Psi$  から対応付いたインデックスを削除
34:   $\Phi \leftarrow$  Remove  $r_A$ th element of  $\Phi$ 
35:   $\Psi \leftarrow$  Remove  $r_B$ th element of  $\Psi$ 
36: end for
37: return  $\mathbf{C}$ 

```



五十嵐洋介 (正員)

2009 筑波大・第三学群・情報学類卒。2011 同大大学院システム情報工学研究科博士前期課程了。在学中はパターン認識・コンピュータビジョンの研究に従事。現在、キヤノン (株) に勤務。



福井 和広 (正員)

1986 九大・工・機械卒。1988 同大大学院工学研究科了。同年 (株) 東芝入社。同社研究開発センター主任研究員を経て、現在、筑波大学大学院システム情報工学研究科准教授。パターン認識・コンピュータビジョンの理論と応用に関する研究に従事。工博。2002 本会論文賞, 2006 本会 ISS 活動功労賞受賞。情報処理学会, IEEE 各会員。