

# ランダムサンプルコンセンサスを用いた 光学的アラインメントに基づく物体認識

岡部 孝弘<sup>†</sup> 佐藤 洋一<sup>†</sup>

3枚の基底画像を用いて任意方向の点光源下の拡散反射成分と attached shadow を表現する手法は、光学的アラインメント (photometric alignment) と呼ばれている。本論文では、照明変化をとともなう物体認識の問題に対して、光学的アラインメントに基づく認識手法を提案する。提案手法では、テスト画像に含まれるハイライトや影などの外れ値を考慮し、基底画像の推定においても利用されている RANSAC (RANDOM SAMPLE CONSENSUS) を用いて、テスト画像を安定に再現する。提案手法の有効性を示すために、顔画像データベース Yale Face Database B を用いた認識実験を行い、RANSAC が、基底画像の推定だけでなく認識にも有効であることを確認した。

## Object Recognition Based on Photometric Alignment Using Random Sample Consensus

TAKAHIRO OKABE<sup>†</sup> and YOICHI SATO<sup>†</sup>

Photometric alignment is a technique that represents both diffuse reflection components and attached shadows under an arbitrary point light source with three basis images. In this paper, we propose a method based on photometric alignment for object recognition under varying illumination. In order to synthesize a test image reliably in the face of outliers such as specular reflection components and shadows, our method utilizes RANSAC (RANDOM SAMPLE CONSENSUS) which has been used successfully for estimating the basis images. To demonstrate the effectiveness of the proposed method, we conducted experiments by using the Yale Face Database B and confirmed that RANSAC is effective not only for estimation of the basis images but also for object recognition under varying illumination.

### 1. ま え が き

物体認識には、物体姿勢、視点位置の変化と照明変化の問題があるが、前者が主に幾何学的特性に依存するのに対して、後者は主に光学的特性に依存するため、これらを分けて考えることができる。本論文では、照明変化の問題に着目し、物体姿勢および視点位置一定の条件下で照明のみが変化する場合の物体認識について議論する。

光源からの一次反射に限定すると、画像に含まれる成分は、拡散反射成分、ハイライト、attached shadow (陰)、cast shadow (影) の4つに分類される<sup>22),23)</sup>。照明変化をとともなう物体認識の問題に対して、Lambert モデルを仮定して、ある物体が任意照明下でとりうる拡散反射成分などを合成することで認識を行うアプローチは、生成的手法 (generative methods) と呼ばれている。

Lambert モデルを仮定すると、任意方向の点光源下の拡散反射成分は、3枚の基底画像の線形結合で表現される。また、線形結合で表現される拡散反射成分が負になる画素は attached shadow に対応していることから、負の画素値を0に置き換えることで、拡散反射成分だけでなく attached shadow も表現することができる。このようにして、任意方向の点光源下の拡散反射成分と attached shadow を3枚の基底画像を用いて表現する手法は、光学的アラインメント<sup>22),23)</sup> と呼ばれている。

光学的アラインメントに基づく物体認識には、基底画像の推定 (学習) とテスト画像の再現 (認識) の2つの要素がある。前者の学習に関しては、画像が拡散反射成分のみで構成される場合には、一次独立な方向の点光源に照らされた3枚の画像を基底画像とすればよい。しかしながら、学習画像には、一般に、ハイライト、attached shadow、および cast shadow が含まれているため、これらの成分を考慮して基底画像を推定する必要がある。また、後者の認識に関しても、テスト画像に含まれる拡散反射以外の成分を考慮して、

<sup>†</sup> 東京大学生産技術研究所

Institute of Industrial Science, The University of Tokyo

基底画像に対する線形結合係数を推定する必要がある。さらに、テスト画像を再現した後の具体的な識別法についても検討する必要がある。

上述の2つの要素のうち前者、すなわち学習過程は、3枚の基底画像が物体表面の法線と拡散反射率から決まるため、照度差ステレオ (photometric stereo)<sup>3,6)</sup> に関する問題である。拡散反射以外の成分を含む学習画像からの基底画像の推定法として、SVDMD (Singular Value Decomposition with Missing Data)<sup>24)</sup> に基づく手法<sup>8)</sup>、SVDMD と可積分条件に基づく手法<sup>7)</sup>、投票に基づく手法<sup>14)</sup>、および RANSAC<sup>6)</sup> を利用した手法<sup>11)</sup> が提案されている。また、運動物体からの基底画像の推定法も提案されている<sup>16)</sup>。

ところが、後者の認識過程については、これまで十分に検討されていなかった。そこで本論文では、光学的アラインメントに基づく物体認識における認識過程について議論する。特に、テスト画像を再現するための線形結合係数の推定法、および、再現後の識別法を提案する。具体的には、外れ値除去の一般的な手法であり、基底画像の推定<sup>11),16)</sup> においても利用されている RANSAC を用いて、結合係数を推定する。また、拡散反射成分と attached shadow からなる合成画像とテスト画像を比較して、再現できている画素数に基づいた識別を行う。

提案手法の有効性を示すために、顔画像データベース Yale Face Database B<sup>7)</sup> を用いた認識実験を行い、RANSAC が、基底画像の推定だけでなく認識にも有効であることを確認する。

本論文の構成は以下のとおりである。まず、2章で関連研究についてまとめ、光学的アラインメントの特長を述べる。次に、3章で光学的アラインメントについて詳しく述べ、物体認識に適用する際の問題点を明らかにする。4章では、RANSAC を利用した認識手法を提案する。5章で顔画像データベースを用いた実験結果を示してその考察を行い、6章でまとめと今後の研究課題を述べる。

## 2. 関連研究

従来、照明変化をともなう物体認識の問題に対しては、特徴に基づく手法 (feature-based methods) と見え方に基づく手法 (appearance-based methods) の2つのアプローチがとられてきた。

前者の特徴に基づく手法では、照明変化の影響を比較的受けにくいエッジやコーナーなどの特徴の抽出と照合により認識を行う。しかし、特徴がつねに安定に抽出できるとは限らないうえに、ローカルな情報を利

用するため、識別に有効な情報を損失する可能性がある<sup>5)</sup>。

これに対して、画像全体のグローバルな情報、つまり、全画素値を入力としてパターン認識の手法を適用する見え方に基づく手法は、照明や姿勢変化をともなう物体認識に有効であることが知られている<sup>3),15),18),21),25)</sup>。ところが、見え方に基づく手法には、一般に、学習時と同様の照明条件でなければ機能しないという欠点がある。

近年提案された生成的手法は、画像全体の情報を用いる点では見え方に基づく手法と同じであるが、Lambert モデルを仮定して、ある物体が任意照明下でとりうる拡散反射成分などを少数の学習画像から生成する点が特長である。したがって、生成的手法は、照明条件が学習画像撮影時と大きく異なる場合でも有効である。

画像が拡散反射成分のみで構成される場合には、Lambert モデルの仮定から、ある物体が任意照明下でとりうる画像は、画像空間内の3次元部分空間に拘束されることが知られている<sup>22),23)</sup>。この3次元部分空間とテスト画像との距離に基づく認識手法は、ハイライト、attached shadow、および cast shadow の寄与が小さい画像に対して有効であることが示されている<sup>7),8)</sup>。本論文では、この手法を照明部分空間法と呼ぶ。

一方、画像が拡散反射成分と attached shadow で構成される場合、つまり、凸物体かつ Lambert モデルを仮定した場合には、ある物体が任意照明下でとりうる画像集合は、画像空間において原点を頂点とする凸錐 (convex cone) を形成することが知られている。この凸錐は照明錐 (illumination cone) と呼ばれ、一般に、3枚の基底画像から生成される仮想照明画像の凸結合で表現される<sup>4)</sup>。照明錐とテスト画像との距離に基づく認識手法は、attached shadow の寄与が大きい画像に対しても有効であることが示されている<sup>7),8)</sup>。

照明錐モデルは、attached shadow を表現できる点で、照明部分空間法よりも優れている。しかしながら、照明錐を表現するには、一般に、物体表面の互いに異なる法線ベクトルの数の二乗のオーダーの仮想照明画像を必要とするため、実用上問題がある。これを受けて、拡散反射成分と attached shadow を、少数の (仮想照明) 画像を用いて効率良く表現する認識手法の研究が行われている。

その1つは、照明錐を少数の画像で近似的に表現しようとする試みである。任意照明下の画像は、実画像

文献 2), 3), 7), 8), 10), 12) では linear subspace method と呼ばれている。

を用いた実験から、少数の画像の線形結合で効率良く表現できることが知られており<sup>9)</sup>、照明錐モデルにおいても、主成分分析により照明錐を近似的に表現することで、効率化を図っていた<sup>7),8)</sup>。

さらに、最近になって、球面調和関数を用いた周波数空間での解析から、理論的にも、4枚から9枚の基底画像の線形結合で、照明錐を効率良く近似できることが示された<sup>1),20)</sup>。9枚の基底画像の張る部分空間とテスト画像との距離に基づく認識手法は、照明錐モデルと同程度の認識性能を持つことが示されている<sup>10),12)</sup>。しかしながら、これらの基底画像は低次の球面調和関数照明下の画像に対応しているため、基底画像を得るには、球面調和関数照明または特殊な方向に置かれた点光源<sup>10),12)</sup>の下で対象物体を撮影するか、物体の形状および拡散反射率を推定して合成しなければならないという問題がある。

もう1つの試みは、照明部分空間法を改良しようとするものである。凸物体の周りにランダムに分布する複数の点光源を考えると、法線ベクトルが類似の物体表面は、類似の点光源に照らされる。この点に着目し、画像を、対応する物体表面の法線ベクトルが類似の画素からなる小画像に分割し、各小画像について照明部分空間法を適用する、分割照明部分空間法(segmented linear subspace method)が提案されている<sup>2)</sup>。この手法においても照明錐モデルと同程度の認識性能が得られることが報告されているが、分割数の決定法については検討する必要がある。

本論文で議論する光学的アラインメントは、3枚の基底画像を用いて、任意方向の点光源下の拡散反射成分と attached shadow を表現する手法である。この手法は、照明部分空間法と比較して、attached shadow を表現できる点で優れている。また、拡散反射成分だけでなく attached shadow も表現できる照明錐モデルと比較して、大量の仮想照明画像を必要としないという特長がある。さらに、照明錐を低次元部分空間で近似する手法と比較して、特殊な照明を用いたり形状と拡散反射率を推定したりする必要がないという点で優れている。そのうえ、分割照明部分空間法と比較しても、画像を適切に分割する必要がないという特長がある。

一方、光学的アラインメントは、任意方向の点光源下の画像に対してのみ有効であり、複数の点光源や面光源などの複雑照明下の画像には適用できない。しかしながら、複雑照明環境においても、点光源と複雑照明の両方に照らされた画像と複雑照明のみに照らされた画像を取得して、その差分を入力とすることで、光学的アラインメントを適用することができる。

### 3. 光学的アラインメント

本章では、点光源下の画像を表現する光学的アラインメントについて、画像が拡散反射成分のみで構成されている場合と attached shadow を含む場合に分けて述べ、光学的アラインメントを物体認識に適用する際の問題点を明らかにする。

#### 3.1 拡散反射成分の線形表現

まず、画像が拡散反射成分のみで構成されている場合について考える。

Lambert モデルの仮定より、十分遠方の点光源(平行光線)に照らされた画像の  $i$  番目の画素値  $x_i^{(D)}$  は、

$$x_i^{(D)} = \rho_i \mathbf{n}_i^T \mathbf{s} \equiv \mathbf{b}_i^T \mathbf{s}, \quad (i = 1, 2, \dots, n) \quad (1)$$

で表される。ここで、 $\rho_i$ 、 $\mathbf{n}_i$  は  $i$  番目の画素に対応する物体表面の拡散反射率と単位法線ベクトル、 $\mathbf{s}$  は光源の向きを表す単位ベクトルにその強度を掛けたものである。したがって、拡散反射成分のみで構成される画像  $\mathbf{x}^{(D)}$  は、 $\mathbf{b}_i^T$  を  $n$  画素分縦に並べた  $n \times 3$  の行列を  $B$  として、

$$\mathbf{x}^{(D)} = B\mathbf{s} \quad (2)$$

で表現される。

式(2)は、任意方向の点光源下の画像が、 $n$  次元画像空間において、

$$\mathcal{L} = \{\mathbf{x} | \mathbf{x} = B\mathbf{s}, \forall \mathbf{s} \in \mathbb{R}^3\} \quad (3)$$

で定義される部分空間に拘束されることを意味している。この部分空間は照明部分空間(illumination subspace)と呼ばれ<sup>4)</sup>、物体形状が平面や円柱などの特殊な場合を除けば、一般に、その次元は3となる。ゆえに、任意方向の点光源下の画像は、一次独立な方向の点光源に照らされた3枚の画像  $e^{(j)}$  ( $j = 1, 2, 3$ ) の線形結合

$$\mathbf{x}^{(D)} = c_1 e^{(1)} + c_2 e^{(2)} + c_3 e^{(3)}. \quad (4)$$

で表現することができる。ここで  $c_j$  ( $j = 1, 2, 3$ ) は線形結合係数である。

本論文では、拡散反射成分を3枚の画像の線形結合で表現する手法を、拡散反射成分だけでなく attached shadow も表現する光学的アラインメントと厳密に区別して、照明部分空間法と呼ぶ。

#### 3.2 拡散反射成分と attached shadow の表現

次に、画像が拡散反射成分と attached shadow で構成されている場合について考える。

点光源の方向の変化に注目した議論を行うため、カメラの分光感度や光源の分光分布などへの依存性は考慮しない。

Attached shadow は物体表面の法線が光源方向を向いていないために生じる陰であるから，式 (1) の内積が負 ( $b_i^T s < 0$ ) となる画素  $i$  は attached shadow に対応している．したがって，拡散反射成分と attached shadow で構成される画像  $x^{(D+AS)}$  は，負の画素値を 0 に置き換えることで以下のように表現される．

$$x^{(D+AS)} = \max(Bs, \mathbf{0}). \quad (5)$$

ここで， $\max(z, \mathbf{0})$  は，ベクトル  $z$  の負の要素を 0 に置き換える操作である．

同様に，拡散反射成分のみで構成される 3 枚の基底画像を用いて，任意方向の点光源下の拡散反射成分と attached shadow を，以下のように表現することができる．

$$x^{(D+AS)} = \max\left(\sum_{j=1}^3 c_j e^{(j)}, \mathbf{0}\right). \quad (6)$$

このようにして，任意方向の点光源下の拡散反射成分と attached shadow を 3 枚の基底画像を用いて表現するのが，光学的アラインメントである．

### 3.3 物体認識への適用

光学的アラインメントは，各クラスについて 3 枚の基底画像を準備しておけば，任意方向の点光源下の拡散反射成分と attached shadow を合成できるため，照明変化をとともう物体認識に適用できることが指摘されている<sup>22),23)</sup>．

しかしながら，学習画像とテスト画像には，一般に，拡散反射成分だけでなく，ハイライト，attached shadow，および cast shadow が含まれている．そのため，たとえば，3 枚以上の学習画像から特異値分解により基底画像を求めても，得られた基底画像は拡散反射以外の成分の影響を受けてしまう．また，認識に関しても，たとえば，基底画像の張る 3 次元部分空間にテスト画像を射影，つまり，すべての画素を用いて最小二乗法により線形結合係数を求めても，正しい結合係数は得られない．したがって，学習と認識の両方の過程において，拡散反射以外の成分を適切に取り扱う必要がある．

さらに，光学的アラインメントに基づいてテスト画像を再現した後の具体的な識別法についても検討する必要がある．

## 4. 提案手法

3 枚の基底画像が物体表面の法線と拡散反射率から決まるため，基底画像の推定は照度差ステレオに関する問題であり，すでにいくつかの手法が提案されてい

る<sup>7),8),11),14),16),26)</sup>．そこで本論文では，光学的アラインメントに基づく物体認識の要素のうち，基底画像の推定以外の部分，すなわち，テスト画像を再現するための線形結合係数の推定法，および，再現後の具体的な識別法を提案する．

### 4.1 提案手法の概要

まず，基底画像のクラスとテスト画像のクラスが同一の場合について考える．一般に，テスト画像には，拡散反射成分だけでなく，ハイライト，attached shadow，および cast shadow が含まれている．これらの 4 つの成分のうち拡散反射成分に関しては，式 (4) のように，3 枚の基底画像の線形結合で表現される．したがって，テスト画像に含まれる拡散反射成分の線形関係から，線形結合係数を推定することができる．さらに，推定した結合係数と基底画像を用いて，式 (6) のようにして，拡散反射成分と attached shadow で構成される画像を合成することができる．このとき，テスト画像の拡散反射成分と attached shadow に対応する画素の値は，合成画像の画素値に等しい．特に，凸物体かつ Lambert モデルが成り立つ場合には，テスト画像と合成画像は完全に一致する．

一方，基底画像のクラスとテスト画像のクラスが異なる場合には，3 枚の基底画像を用いてテスト画像を再現するのは困難であることが予想される．

以上のことから，提案手法では，RANSAC を利用してテスト画像の拡散反射成分と attached shadow を最も多く再現するような線形結合係数を求め，再現できた画素数に基づいた識別を行う．

### 4.2 認識の手順

RANSAC を利用した提案手法の具体的な手順を示す．

各クラスについて，以下の手順で再現できた画素数を求める．

#### (1) 線形結合係数の算出

ランダムに選択した 3 画素が拡散反射成分に対応していると仮定して，画素値の線形関係から結合係数  $\hat{c}_j$  ( $j = 1, 2, 3$ ) を算出する．

#### (2) 画像合成とラベル付け

基底画像と手順 (1) で算出した結合係数から，以下のように拡散反射成分と attached shadow で構成される画像を合成する．

$$\hat{x}_i = \max\left(\sum_{j=1}^3 \hat{c}_j e_i^{(j)}, 0\right), \quad (i = 1, 2, \dots, n). \quad (7)$$

ここで  $e_i^{(j)}$  は  $j$  番目の基底画像の  $i$  番目の画

素値である．さらに，テスト画像と合成画像の画素値の誤差が閾値以下の画素は再現できていると見なして，各画素にラベル

$$\xi_i = \begin{cases} 1 & (|x_i^{(test)} - \hat{x}_i| < t : inlier) \\ 0 & (|x_i^{(test)} - \hat{x}_i| \geq t : outlier) \end{cases} \quad (8)$$

を付ける．ここで  $x_i^{(test)}$  はテスト画像の  $i$  番目の画素値である．閾値  $t$  は，拡散反射成分の Lambert モデルからの微小なずれ<sup>19)</sup> や画像の解像度などに依存する．

- (3) 再現できた画素の数え上げ  
再現できた画素 (inlier) の総数を

$$support = \sum_{i=1}^n \xi_i \quad (9)$$

と定義して，数え上げる．

- (4) 手順 (1) から (3) の反復  
手順 (1) から (3) までを繰り返し， $support$  を最大にする結合係数を求める．
- (5) 最小二乗法による結合係数の更新  
拡散反射成分に対応していると考えられる画素に対して最小二乗法を適用し，結合係数を更新する．具体的には，inlier のうち，誤差の閾値  $t$  の範囲で，attached shadow に対応しているもの，および，飽和しているものを除いて，重み

$$w_i = \begin{cases} 1 & (\xi_i = 1, t < \hat{x}_i < 1 - t) \\ 0 & (others) \end{cases} \quad (10)$$

を与え，評価関数

$$C = \sum_{i=1}^n w_i \left( x_i^{(test)} - \sum_{j=1}^3 \hat{c}_j e_i^{(j)} \right)^2 \quad (11)$$

を最小とする結合係数  $\hat{c}_j$  ( $j = 1, 2, 3$ ) を推定する．

- (6) 画像合成とラベル付け  
更新された結合係数に対して，手順 (2) の画像合成とラベル付けを行う．
- (7) 手順 (5) と (6) の反復  
手順 (5) と (6) を，ラベル  $\xi_i$ , ( $i = 1, 2, \dots, n$ ) が収束するまで繰り返す．
- (8) 再現できた画素数の数え上げ  
得られたラベルから  $support$  を求める．

最後に，各クラスの  $support$  を比較して， $support$

を最大とするクラスにテスト画像を分類する．

#### 4.3 従来手法との相違

RANSAC を利用した基底画像の推定手続きの一部として，すでに，RANSAC を用いた線形結合係数の推定法が提案されている<sup>11)</sup>．前節で述べた提案手法は，従来の推定法と比較して，以下の 2 つの特長がある．

従来手法では，テスト画像において画素値がある閾値以下の画素を，3 枚の基底画像の線形結合で表現したときの画素値が負の場合に attached shadow，正の場合に cast shadow に対応していると見なしたうえで，拡散反射成分，attached shadow，および cast shadow の 3 つの成分を inlier と定義している．したがって，画素値がある閾値以下の場合には，算出された線形結合係数の値にかかわらずつねに inlier と見なされるため，結合係数は，拡散反射成分のみから決定される．ゆえに，従来手法は，照明部分空間法に基づいているといえる．

一方，提案手法では，光学的アラインメントに基づいて，拡散反射成分と attached shadow を inlier と定義している．したがって，提案手法は，手順 (4) から分かるように，拡散反射成分だけでなく attached shadow の持つ情報も，結合係数の推定に利用している．

また，従来手法が  $support$  を最大とする線形結合係数を正解と見なしているのに対して，提案手法では， $support$  を最大とする結合係数を元に，手順 (5) 以降で，最小二乗法の枠組みで最適な結合係数を推定している．

## 5. 実験結果

提案手法の有効性を示すために，顔画像データベース Yale Face Database B<sup>7)</sup> を利用した認識実験を行った．顔画像については，照明変化をとまなう画像の変化が人物の違いによる画像の変化よりも大きくなる場合が多いために，照明変化をとまなう認識はきわめて困難な問題であることが知られている<sup>13)</sup>．

### 5.1 顔画像データベース

利用したデータベースは，10 人の顔を，9 つの姿勢で，64 方向の点光源および環境光の下で撮影した 5,850 枚の画像からなり，正面方向を向いている画像については，両目と口の 2 次元座標が与えられている．また，各画像は，点光源の方向とカメラの光軸のなす角度  $\theta$  に基づいて，5 つの部分集合に分類されている．

認識実験には，正面方向を向いている 650 枚の画像から顔付近を切り出した画像を用いた．その際，顔の位置や大きさの違いが認識結果に与える影響を避ける

画素値は，たとえば 8 bit であれば 255 で割ることで，0 から 1 の間の値をとるように正規化されているものとする．

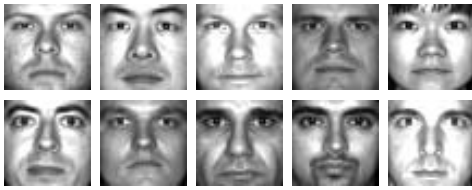


図 1 10人の顔画像

Fig. 1 Cropped images of ten individuals.



Subset1  $\theta \leq 12^\circ$     Subset2  $\theta \leq 25^\circ$     Subset3  $\theta \leq 50^\circ$     Subset4  $\theta \leq 77^\circ$     Subset5  $\theta > 77^\circ$

図 2 画像の光源方向依存性

Fig. 2 Example images in each subset: variability due to illumination.



図 3 基底画像

Fig. 3 Basis images.

ために、両目の座標が等しくなるように切り出して、 $40 \times 40$  の解像度の画像を作成した。また、環境光の影響を取り除くために、環境光の下で撮影した画像との差分をとった。図 1 に、実験に用いた 10 人の正面顔画像の例を示す。また、図 2 に示した各部分集合に属する画像の例から、同一人物であっても、光源の方向に依存して画像が大きく変化しているのを確認することができる。

## 5.2 学 習

各人物の学習画像として、部分集合 1 に属する 7 枚の画像を用いた。基底画像の推定には、画素値に基づいてハイライト, attached shadow, および cast shadow を除去したうえで特異値分解を適用する, SVDMD に基づく手法<sup>8)</sup> を利用した。図 3 に、推定した基底画像の一例を示す。

また、得られた基底画像から学習画像を合成して、式 (8) の閾値  $t$  を決定した。具体的には、基底画像の推定に用いた画素について、学習画像と合成画像の画素値の誤差を求め、その分布の標準偏差  $\sigma$  を閾値  $t$  の目安とした。

## 5.3 認 識

以下に示す 3 つの手法を用いた認識実験を行った。

### ● 照明部分空間法 (IS)

基底画像の張る 3 次元部分空間とテスト画像との距離を求め、距離が最小となるクラスに分類する。

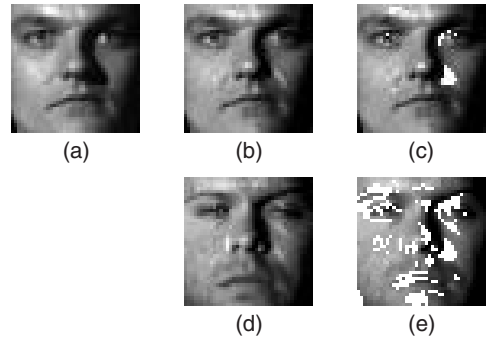


図 4 テスト画像と合成画像

Fig. 4 A test image and synthesized images.

- 射影を用いた光学的アラインメント (PA1)  
光学的アラインメントに基づいて合成した画像とテスト画像との距離を求め、距離が最小のクラスに分類する。ただし、線形結合係数は、基底画像の張る 3 次元部分空間にテスト画像を射影して求める。
- RANSAC を用いた光学的アラインメント (PA2)

4 章で提案したように、RANSAC を利用して結合係数を推定し、テスト画像を再現できた画素数に基づいて識別する。

図 4 に、提案手法 (PA2) の処理過程を示した。テスト画像 (a) に対して、同一人物の基底画像から合成した画像 (b) と、他の人物の基底画像から合成した画像 (d) を示した。各合成画像について、テスト画像との誤差が閾値  $t$  よりも大きい画素を白く表示したものが (c), (e) である。同一人物の基底画像から合成した画像が、鼻や目頭の cast shadow および額のハイライトを除いてほぼ再現できているのに対して、他の人物の基底画像から合成した画像は、目、鼻、口の周辺を中心として、比較的広い範囲で再現に失敗していることが分かる。

上記の 3 つの手法について、各部分集合に対する誤識別率を示した (表 1)。提案手法では、閾値  $t$  を  $4\sigma$  としたときの結果を示した。

実験結果から、拡散反射成分しか表現することのできない照明部分空間法 (IS) と比較して、拡散反射成分だけでなく attached shadow も生成できる光学的アラインメントに基づく手法 (PA1, PA2) が優れていることが分かる。特に、図 2 から分かるように、点光源の方向がカメラの光軸から外れるにつれて

閾値を  $3.5\sigma$  から  $5.5\sigma$  まで変化させても、誤識別率は、部分集合 4 に対して最大で 1% 程度、部分集合 5 に対して最大で 2% 程度しか変化しなかった。

表 1 誤識別率 (%)  
Table 1 Recognition error rates (%).

Method	Subset2	Subset3	Subset4	Subset5
IS	0	0	5.8	55.6
PA1	0	0	0.7	39.7
<b>PA2</b>	0	0	0	18.5
IC: attached <sup>7)</sup>	0	0	8.6	—
9PL <sup>12)</sup>	0	0	2.8	—
9PL <sup>10)</sup>	0	0.7	1.4	—
Segm. LS <sup>2)</sup>	0	0	0	—

attached shadow の寄与が大きくなるために、性能の差が大きくなって表れていると考えられる。また、同じ光学的アラインメントに基づく認識手法でも、線形結合係数の推定に RANSAC を用いた提案手法が、射影を用いた手法よりも優れているのを確認することができる。しかしながら、提案手法にも、ハイライトと cast shadow を生成することができないという問題がある。部分集合 5 に対する誤識別は、図 2 から、ハイライトや cast shadow が原因であると考えられる。

表 1 には、他の論文で報告されている結果も示した。なお、部分集合 5 に対する誤識別率は、他の論文では報告されていない。各実験で実際に使用している画像の切り出し方や解像度が若干異なるために、結果を単純に比較することはできないが、提案手法は、3 枚の基底画像しか用いていないにもかかわらず、大量の仮想照明画像を必要とする照明錐モデル<sup>7)</sup>よりも優れた性能を持つと考えられる。また、提案手法は、未知方向の点光源下で撮影された 7 枚の学習画像を用いているが、特殊な方向の点光源下の画像に対応する 9 枚の合成画像を用いた認識手法<sup>10),12)</sup>と比較しても、同程度の認識性能を示している。さらに、基底画像の適切な分割を必要とする分割照明部分空間法<sup>2)</sup>と比較しても、3 枚の基底画像をそのまま用いる提案手法は、同程度の認識性能を持つと考えられる。

## 6. む す び

照明変化をとまなう物体認識の問題に対して、任意方向の点光源下の拡散反射成分と attached shadow を 3 枚の基底画像を用いて表現する、光学的アラインメントに基づく認識手法を提案した。提案手法では、テスト画像に含まれる拡散反射以外の成分を考慮し、RANSAC を利用してテスト画像の拡散反射成分と attached shadow を最も多く再現するような線形結合係数を求め、再現できた画素数に基づいて認識を行った。また、顔画像データベースを用いた認識実験を行い、RANSAC が、基底画像の推定だけでなく認識にも有効であることを示した。

本論文で提案した光学的アラインメントに基づく認識手法には、cast shadow を表現することができないという問題がある。また、cast shadow を考慮した照明錐モデルにおいても<sup>7),8)</sup>、cast shadow を厳密に表現するのは困難であることが指摘されており<sup>17)</sup>、今後、生成的手法における cast shadow の取扱いについて検討したい。

謝辞 本研究では顔画像データベース Yale Face Database B<sup>7)</sup> を利用した。本研究の一部は、文部科学省科学研究費補助金特定領域研究(C)(2)「人間の意図・行動解析に基づく柔軟なヒューマン・マシン・インタラクションの実現」(課題番号:13224051)の助成により行われた。最後に、有益なご意見をいただいた査読者の方々に感謝いたします。

## 参 考 文 献

- 1) Basri, R. and Jacobs, D.: Lambertian reflectance and linear subspaces, *Proc. IEEE ICCV 2001*, pp.383-390 (2001).
- 2) Batur, A.U. and Hayes III, M.H.: Linear subspaces for illumination robust face recognition, *Proc. IEEE CVPR 2001*, 2, pp.296-301 (2001).
- 3) Belhumeur, P.N., Hespanha, J.P. and Kriegman, D.J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Trans. PAMI*, Vol.19, No.7, pp.711-720 (1997).
- 4) Belhumeur, P.N. and Kriegman, D.J.: What is the set of images of an object under all possible lighting conditions?, *Int'l. J. Computer Vision*, Vol.28, No.3, pp.245-260 (1998).
- 5) Brunelli, R. and Poggio, T.: Face recognition: features versus templates, *IEEE Trans. PAMI*, Vol.15, No.10, pp.1042-1052 (1993).
- 6) Fischler, M.A. and Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, Vol.24, No.6, pp.381-395 (1981).
- 7) Georghiades, A.S., Belhumeur, P.N. and Kriegman, D.J.: From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. PAMI*, Vol.23, No.6, pp.643-660 (2001).
- 8) Georghiades, A.S., Kriegman, D.J. and Belhumeur, P.N.: Illumination cones for recognition under variable lighting: faces, *Proc. IEEE CVPR '98*, pp.52-58 (1998).
- 9) Hallinan, P.W.: A low-dimensional representation of human faces for arbitrary lighting conditions, *Proc. IEEE CVPR '94*, pp.995-999

- (1994).
- 10) Ho, J., Lee, K.-C. and Kriegman, D.J.: On reducing the complexity of illumination cones for face recognition, *Proc. CVPR Workshop on Identifying Objects Across Variations in Lighting* (2001).
  - 11) 石井育規, 福井孝太郎, 向川康博, 尺長 健: 光学現象の分類基準を利用した画像の線形化, 画像の認識・理解シンポジウム (MIRU) 論文集 II, pp.167-176 (2002).
  - 12) Lee, K.-C., Ho, J. and Kriegman, D.J.: Nine points of light: acquiring subspaces for face recognition under variable lighting, *Proc. IEEE CVPR 2001*, 1, pp.519-526 (2001).
  - 13) Moses, Y., Adini, Y. and Ullman, S.: Face recognition: the problem of compensating for changes in illumination direction, *Proc. ECCV '94*, pp.286-296 (1994).
  - 14) Mukaigawa, Y., Miyaki, H., Mihashi, S. and Shakunaga, T.: Photometric image-based rendering for image generation in arbitrary illumination, *Proc. ICCV 2001*, pp.652-659 (2001).
  - 15) Murase, H. and Nayar, S.K.: Visual learning and recognition of 3-D objects from appearance, *Int'l. J. Computer Vision*, Vol.14, No.1, pp.5-24 (1995).
  - 16) Nakashima, A., Maki, A. and Fukui, K.: Constructing illumination image basis from object motion, *Proc. ECCV 2002* (LNCS 2352), pp.195-209 (2002).
  - 17) 岡部孝弘, 佐藤洋一: キャストシャドウ存在下の照明錐モデルに関する考察, 情報処理学会研究報告, CVIM 2002-131-20, pp.141-148 (2002).
  - 18) 岡部孝弘, 佐藤洋一: 照明変化を伴う物体認識へのサポートベクターマシンの適用, 画像の認識・理解シンポジウム (MIRU) 論文集 II, pp.177-186 (2002).
  - 19) Oren, M. and Nayar, S.K.: Generalization of the Lambertian model and implications for machine vision, *Int'l. J. Computer Vision*, Vol.14, No.3, pp.227-251 (1995).
  - 20) Ramamoorthi, R. and Hanrahan, P.: On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object, *J. Opt. Soc. Am. A*, Vol.18, No.10, pp.2448-2459 (2001).
  - 21) Shakunaga, T. and Shigenari, K.: Decomposed eigenface for face recognition under various lighting conditions, *Proc. IEEE CVPR 2001*, 1, pp.864-871 (2001).
  - 22) Sashua, A.: Geometry and photometry in 3D visual recognition, Ph.D. Thesis, MIT (1992).
  - 23) Sashua, A.: On photometric issues in 3D visual recognition from a single 2D image, *Int'l. J. Computer Vision*, Vol.21, No.1/2, pp.99-122 (1997).
  - 24) Shum, H.-Y., Ikeuchi, K. and Reddy, R.: Principal component analysis with missing data and its application to polyhedral object modeling, *IEEE Trans. PAMI*, Vol.17, No.9, pp.854-867 (1995).
  - 25) Turk, M.A. and Pentland, A.P.: Face recognition using eigenfaces, *Proc. IEEE CVPR '91*, pp.586-591 (1991).
  - 26) Woodham, R.: Photometric method for determining surface orientation from multiple images, *Optical Engineering*, Vol.19, No.1, pp.139-144 (1980).

(平成 14 年 9 月 6 日受付)

(平成 15 年 3 月 28 日採録)

(担当編集委員 向川 康博)



岡部 孝弘 (正会員)

1997 年東京大学理学部物理学卒業。1999 年同大学院理学系研究科物理学専攻修士課程修了。2000 年同大学院理学系研究科物理学専攻博士課程中退。2001 年より東京大学生産技術研究所技官。コンピュータビジョン, コンピュータグラフィックスに関する研究に従事。電子情報通信学会, IEEE 各会員。



佐藤 洋一 (正会員)

1990 年東京大学工学部機械工学科卒業。同大学院情報工学専攻を経て, 1991 年よりカーネギーメロン大学計算機科学部ロボティクス学科博士課程に在籍。1997 年 Ph.D. in Robotics 修了。同年より東京大学生産技術研究所研究機関研究員, 講師を経て, 現在同研究所助教授。コンピュータビジョン, ヒューマン・コンピュータ・インタラクション, およびコンピュータグラフィックスに関する研究に従事。平成 11 年山下記念研究賞, Int'l Conf. Shape Modeling and Applications '97 最優秀論文賞, MIRU2000 最優秀論文賞, 平成 11 年度日本バーチャルリアリティ学会論文誌論文賞, IEEE VR2001 Honorable Mention for the Outstanding Paper Award 受賞。電子情報通信学会, 日本バーチャルリアリティ学会, ACM, IEEE 各会員。