

## 画像線形化に基づく物体認識手法

岡部孝弘 佐藤洋一

東京大学生産技術研究所

〒153-8505 東京都目黒区駒場4-6-1

E-mail: {takahiro, ysato}@iis.u-tokyo.ac.jp

あらまし：照明変化を伴う物体認識に有効である照明錐モデルには、attached shadow を表現できるという利点がある一方で、膨大な数の仮想照明画像を生成しなくてはならないという欠点がある。本稿では、画像合成の分野で知られている画像線形化の枠組みに基づいて、3枚の基底画像を用いた認識手法を提案する。本手法は、3枚の基底画像を用いて認識を行う線形部分空間法が attached shadow を表現できないという欠点を、テスト画像撮影時の照明を推定することで解決するものである。顔画像データベースに対して提案手法を用いた認識実験を行い、画像線形化が、photometric image-based rendering だけでなく物体認識にも有効であることを示した。

キーワード：物体認識, 照明変化, 線形部分空間法, 画像線形化, Photometric image-based rendering

## Object Recognition Based on Image Linearization

Takahiro OKABE Yoichi SATO

Institute of Industrial Science, The University of Tokyo

Komaba 4-6-1, Meguro-ku, Tokyo, 153-8505 Japan

E-mail: {takahiro, ysato}@iis.u-tokyo.ac.jp

**Abstract** : It has been shown theoretically that the illumination cone model is effective for object recognition because it defines the set of images of an object under varying illumination, provided that the images consist of only diffuse components and attached shadows. It is, however, difficult to apply this model for object recognition in practice because a large number of extreme-images of an object are necessary for making the illumination cone of the object. To cope with this problem, we propose a method for object recognition under varying illumination; the method is based on the image linearization that has been shown useful in the context of photometric image-based rendering (PIBR). The proposed method is capable of treating both diffuse components and attached shadows even in the case where only as few as three images are given. We have tested our proposed method by using the Yale Face Database B and confirmed that the image linearization is effective not only for PIBR but also for object recognition under varying illumination.

**Keywords** : Object recognition, Varying illumination, Linear subspace method, Image linearization, Photometric image-based rendering

# 1 はじめに

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性質に依存するのに対して、後者は主に光学的性質に依存するため、これらを分けて考えることができる。本稿では、物体姿勢および視点位置一定の下での照明変化の問題について議論する。

従来、照明変化の問題に対しては、照明変化の影響を比較的受けにくいエッジ、コーナー等の特徴を利用した、特徴に基づく手法 (feature-based methods) が用いられてきた。しかし、特徴が常に安定して抽出できるとは限らないうえに、特徴というローカルな情報を利用するために、識別に有効な情報を損失している可能性がある [6]。

これに対して、画像全体のグローバルな情報を用いて画像と物体を直接対応づける、見え方に基づく手法 (appearance-based methods) が提案され、照明変化や物体姿勢変化を伴う物体認識においてその有効性が示されている [20][14][2][16]。ところが、これらの手法には、ある条件下で認識を行うにはあらかじめ同様の条件下で撮影した画像を用いて学習させなければならないという欠点がある。

近年提案された照明錐 (illumination cone) モデルは、照明のみが変化する場合に着目した画像認識手法の一つであり、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、反射モデルを仮定して、任意の照明下での物体の見え方を少数の学習画像から生成するという点で異なっており、見え方に基づく手法と区別して、生成的手法 (generative methods) と呼ばれる [4]。この手法は、照明が学習画像撮影時と大きく異なる場合にも有効であることが示されている [9][8]。

画像に含まれる成分は大きく分けて、拡散反射成分、ハイライト成分、attached shadow、cast shadow の4つに分類することができる。画像が拡散反射成分のみで構成されている場合には、光源方向が一次独立な3枚の画像の線形結合で、方向および個数が任意の平行光線下の画像を表現できることが知られている [19][15]。3枚の画像が画像空間で張る部分空間とテスト画像との距離に基づいた認識手法は線形部分空間法 (linear subspace method) と呼ばれ、ハイライトや影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 [9][8]。

一方、画像がハイライトや影を含む一般の場合には、重ね合わせの原理が成り立つことを仮定すると、任意の照明下で取り得る画像集合は、画像空間において凸錐 (convex cone) になることが示されており、これは照明錐と呼ばれている [4]。ここでいう重ね合わせの原理とは、二つの光源から同時に照らされた場合の画像が、個々の光源から照らされた場合の画像の和で表現されるとするものである<sup>1</sup>。凸物体かつ Lambert モデルを仮定した場合、つまり、画像が拡散反射成分と attached shadow で構成される場合の照明錐は、最少で3枚の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ことが示されており [4]、照明錐とテスト画像との距離に基づいた認識手法の有効性が示されている [9][8]。また、物体形状を復元 [5] することで、attached shadow だけでなく cast shadow も考慮した認識手法が提案されている [8]、この手法には原理的な問題があることが指摘されている [17]。

照明錐モデルは attached shadow を表現できる点で線形部分空間法よりも優れているが、一般に、画素数の二乗のオーダーの膨大な数の仮想照明画像を準備しなくてはならないという欠点がある。これを受けて、最近、拡散反射成分と attached shadow 成分を、少数の (仮想照明) 画像を用いて効率よく表現する認識手法の研究が行われている。

その一つは、照明錐を少数の画像で近似的に取り扱おうとする試みである。顔画像については実験から、任意の照明下での画像を少数の画像の線形結合で効率よく表現できることが知られていたが [10]、球面調和関数を用いた解析から理論的にも、顔画像に限らず、ある9枚の画像の線形結合で照明錐を効率よく近似できることが示された [18][3]。この9枚の画像の張る部分空間とテスト画像の距離に基づいた認識手法は、照明錐モデルと同程度の認識性能が得られることが示されている [3][11]、9枚の基底画像を得るには、球面調和関数光源下で対象物体を撮影するか、物体の形状を復元して球面調和関数光源下の画像を合成する必要がある。

もう一つの試みは、線形部分空間法を改良しようとするものである。凸物体のまわりにランダムに分布する単光源を考えたときに、法線ベクトルが類似の物体表面は類似の光源に照らされることに着目して、元の画像を、対応する物体表面の法線ベクトルが

<sup>1</sup> 但し、カメラの応答関数の非線形性が支配的になるような場合には、この仮定は成り立たない。

類似の画素からなる小画像に分割して、各小画像について線形部分空間法を適用する分割線形部分空間法 (segmented linear subspace method) が提案され、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 [1]. この手法の認識性能は画像の分割数に依存するが、その決定は自明ではない. 特に、認識対象物体の形状が大きく異なる場合の分割数の決定は容易でないと思われる.

本稿では、照明錐モデルが膨大な数の仮想照明画像を必要とするという問題点に対して、後者の線形部分空間法を改良する枠組みで、より簡便な認識手法を提案する. 提案手法は、3 枚の画像の線形結合で画像を表現する線形部分空間法の枠組みでありながら、照明錐モデルと同様に、attached shadow を生成できるという利点がある. 本手法は、画像合成の分野で知られている画像の線形化 [13] の枠組みを物体認識に適用したものであり、テスト画像撮影時の照明を推定することで、拡散反射成分だけでなく attached shadow も生成できることを利用している. 具体的には、まず、テスト画像の拡散反射成分を最もよく再現するように 3 枚の基底画像を組合わせて合成画像を作成し、合成画像の負の値を持つ画素値を 0 に置き換えることで attached shadow を表現する. 次に、各クラスごとに作成した合成画像とテスト画像を比較し、正しく再現できている画素数の最も多いクラスにテスト画像を分類する. 提案手法の有効性を示すために、10 名の顔を 64 方向の照明の下で撮影した顔画像データベース Yale Face Database B[7] を用いた認識実験を行い、画像の線形化が、photometric image-based rendering (PIBR) だけでなく、物体認識にも有効であることを示す.

本稿の構成は以下の通りである. まず、第 2 章で線形部分空間法および照明錐モデルについて述べる. 次に第 3 章で、画像の線形化に基づいた提案手法について述べる. 第 4 章では認識実験の結果を示してその考察を行い、第 5 章で結ぶ.

## 2 生成的手法

生成的手法は、内部反射の理想的なモデルである Lambert モデルを仮定して、任意の照明下の画像を生成することにより認識を行う手法である. 本節では、代表的な生成的手法である、線形部分空間法と照明錐モデルについて述べる.

### 2.1 線形部分空間法

Lambert モデルの仮定より、無限遠の点光源 (平行光線) に照らされた画像の  $i$  番目の画素値  $x_i$  は、

$$x_i = \rho_i n_i^T s \equiv b_i^T s, \quad (i = 1, 2, \dots, n) \quad (1)$$

で表される. ここで、 $\rho_i, n_i$  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、 $s$  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<sup>2</sup>. 凸物体を仮定すると、その画像は面のローカルな情報のみで決定され、 $b_i^T$  を  $n$  画素分縦に並べた  $n \times 3$  の行列を  $B$  として、

$$x = \max(Bs, 0) \quad (2)$$

で与えられる.  $b_i^T s < 0$  の成分は attached shadow に対応し、最大値をとる操作により 0 に射影される.

さらに、画像が attached shadow を含まずに拡散反射成分のみから構成される場合には、任意の照明下での画像集合は、

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

で表される [19][15].  $\mathcal{L}$  の次元は  $B$  のランクに等しく、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 3 次元であり、これは照明部分空間 (illumination subspace) と呼ばれる [4].

したがって、行列  $B$  は、最少で 3 枚の一次独立な方向の光源に照らされた画像を用いて推定することができる. すなわち、 $c$  枚の学習画像を各々  $x^{(j)}$  ( $j = 1 \dots c, c \geq 3$ ) としたときに、それらを並べた行列

$$X = (x^{(1)}, x^{(2)}, \dots, x^{(c)}) \quad (4)$$

の特異値分解により、行列  $B$  を推定することができる. しかし、画像が拡散反射成分のみから構成されることを仮定しているために、学習画像がハイライトや影を含む一般の場合には、単純に特異値分解して求めた行列  $B$  は式 (3) を満たさない. 3.1 節で述べる画像の線形化 [13] とは、ハイライトや影を含む複数枚の画像からランダムサンプリングと外れ値除去により拡散反射成分を求める手法である. また、

$$x = Bs = (BA)(A^{-1}s) = B^* s^* \quad (5)$$

より、光源方向が未知の場合の行列  $B$  の推定には、 $3 \times 3$  の正則な行列  $A$  の分だけ自由度が残る. この自由度については、3.1 節で述べる.

<sup>2</sup> 光源の方向と照射強度に注目した議論を行なうため、ここでは、カメラの分光応答関数や光源のパワースペクトル等への依存性を無視する.

## 2.2 照明錐モデル

照明錐は、無限遠方に存在する光源を仮定した場合に、任意の複数光源下において取り得る画像の集合である。ところが、重ね合わせの原理から、単光源に照らされた場合に取り得るすべての画像を求めれば、これらの線形和により容易に照明錐を構築できる。

凸物体かつ Lambert モデルを仮定した場合の照明錐は、行列  $B$  が既知の場合に、境界画像 (extreme-image) と呼ばれる仮想照明画像

$$e_{ij} = \max(B(b_i \times b_j), 0), \quad (i, j = 1, 2, \dots, n) \quad (6)$$

の凸結合で表現されることが示されている [4]。ここで  $b_i \times b_j$  方向からの平行光線は、境界光線 (extreme-ray) と呼ばれている [4]。照明錐モデルは、線形部分空間法と比較して、attached shadow を表現できるという利点があるが、式 (6) からわかるように、照明錐を表現するためには、互いに独立な法線ベクトルの数の二乗のオーダーの境界画像が必要になる。

## 3 提案手法

本稿の主題は、PIBR における画像の線形化を物体認識に適用することで、3 枚の画像の線形結合で画像を表現する線形部分空間法の枠組みでも、attached shadow を考慮した物体認識が可能であり、しかも有効であることを実験的に示すことである。本節では、ハイライトや影を含む複数枚の画像から拡散反射成分を求める画像の線形化 [13] について述べたのち、線形化の枠組みに基づく物体認識手法について述べる。

### 3.1 画像の線形化

式 (3) は、ある物体の任意の照明下の画像  $x$  が、3 枚の画像の線形結合で表現できることを意味している。線形化の手法は、拡散反射成分のみからなる 3 つの画素から、この 3 つの係数を決定することに着目した手法であり、この係数組を求めることは、式 (5) の自由度を残して、光源方向を推定していることに対応している。

ハイライトや影を含んだ  $c$  枚の画像を線形化する手続きは、以下の 3 つの過程からなる。まず、3 枚の画像  $x_1, x_2, x_3$  を選択してそれらを基底画像とし、残りの  $(c-3)$  枚の画像  $x_i (i = 4, 5, \dots, c)$  を基底画像の線

形和で表現する (基底画像による表現)。具体的には、 $x_i$  を基底画像の線形結合で

$$x_i = \alpha_i^{(1)} x_1 + \alpha_i^{(2)} x_2 + \alpha_i^{(3)} x_3 \quad (7)$$

のように表現し、 $x_i$  から任意に選択した 3 画素の値と、3 枚の基底画像中におけるそれらの画素値の関係から、係数組  $\alpha_i = (\alpha_i^{(1)}, \alpha_i^{(2)}, \alpha_i^{(3)})$  を求める。選択した 3 つの画素が拡散反射成分のみから成る場合には正しい係数組が得られる一方で、選択した 3 つの画素にハイライトや影が含まれている場合には正解から外れることから、ランダムに 3 画素を選択して係数組を算出することを繰り返して係数組の分布を求め、分布の重心を求めて外れ値を除去することを繰り返して最頻値となる係数組を  $\alpha_i$  とする。

次に、係数組  $\alpha_i (i = 4, 5, \dots, c)$  を用いて、基底画像を線形化する (基底画像の線形化)。具体的には、 $x_i$  から任意に選択した 3 枚の画像  $x_l, x_m, x_n$  とそれに対応した係数組  $\alpha_l, \alpha_m, \alpha_n$  を用いて、基底画像  $x_1, x_2, x_3$  を合成する。 $x_l, x_m, x_n$  のある画素が全て拡散反射成分のみから成る場合には合成画像の画素値も拡散反射成分と一致する一方で、ハイライトや影が含まれている場合には合成画像の画素値は拡散反射成分とは異なることから、ランダムに 3 枚の画像を選択して各画素の値を算出することを繰り返して画素値の分布を求め、分布の重心を求めて外れ値を除去することを繰り返して最頻値となる画素値で構成された基底画像を、 $x_1^L, x_2^L, x_3^L$  とする。

最後に、残りの  $(c-3)$  枚の画像を、係数組  $\alpha_i (i = 4, 5, \dots, c)$  を用いて  $x_1^L, x_2^L, x_3^L$  の線形結合で表現すれば線形化は完了する (全画像の線形化)。

このようにして行われる画像の線形化とは、ハイライトやキャストシャドウ成分を取り除くと同時に、式 (2) の非線形な操作を除去する手続きである。したがって、線形化された画像の各画素値は式 (1) で与えられ、 $b_i^T s < 0$  となる画素の値は負となる。

### 3.2 画像の線形化に基づく提案手法

提案手法は、与えられた画像から 3 枚の基底画像を求める学習の手続きと、それらを用いてテスト画像を分類する認識の手続きからなる。

学習に関しては、十分な数の学習画像が与えられている場合には、3.1 節で述べた線形化の手続きをそのまま適用できるが、学習画像の数が少なくなるにつれ

て外れ値除去による画素値の推定が不安定になる。4節の実験では、同じデータベースを用いた他の認識実験 [8][11][1] との比較を行うために、他の実験と同様に、各人物あたり 7 枚の学習画像を用いて認識を行うことから、画像の線形化を行わずに、正規化した画像を並べた行列の特異値分解により、各クラスあたり 3 枚の基底画像を求める。

こうして得られた各クラスあたり 3 枚の基底画像を用いてテスト画像を認識する手法は、以下の 3 つの手続きからなる。

まず、各クラスについてテスト画像の拡散反射成分を最もよく再現する 3 枚の基底画像の組合わせをランダムサンプリングと外れ値除去により求める。これは、画像線形化の最初の手続き (基底画像による表現) と同じである。

次に、各クラスごとに基底画像の組み合わせで表現される合成画像に対して、負の画素値を 0 と置き換えて、attached shadow を含む合成画像を作成する。

最後に、合成画像とテスト画像を比較して認識を行う。画像の線形化に基づく提案手法では、拡散反射成分と attached shadow を正しく合成することができるが、ハイライトと cast shadow を合成することはできない。ハイライトや cast shadow に対応する画素の値は拡散反射成分とは大きく異なるために、ある画素に注目したときに、ハイライトや cast shadow の影響で生じる誤差が、クラスが異なるために生じる誤差よりも大きくなる可能性がある。したがって、単純に合成画像とテスト画像の誤差の総和 (= 距離) に基づいて認識を行った場合に、誤認識を起こす原因となる可能性がある。ゆえに提案手法では、合成画像とテスト画像の誤差が閾値より大きい場合に距離 1、閾値以下の場合に距離 0 として、2 値化した距離を用いて、その距離が最少となるクラスに分類する。これは、テスト画像と合成画像の各画素値の差の絶対値が閾値以下の画素数をカウントし、カウント数の最も多いクラスにテスト画像を分類することと等価である。

## 4 実験結果

### 4.1 顔画像データベース

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[7] を利用した認識実験を行った。顔画像については、照明変化に伴う画像の変化が、

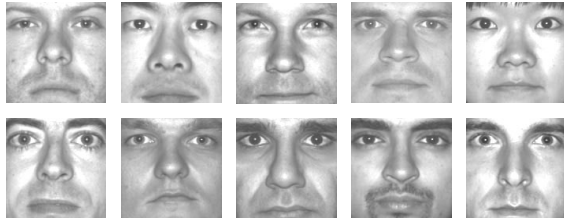


図 1: 10 名の顔画像

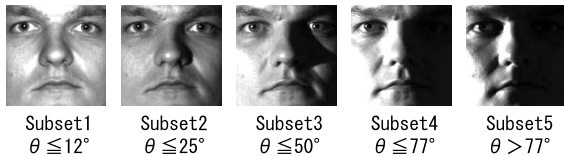


図 2: 画像の光源方向依存性

個人の違いによる画像の変化よりも大きくなることが多いために、照明変化を伴う認識は極めて困難な問題であることが知られている [12].

利用したデータベースは、10 人の顔を 9 つの姿勢で、64 方向の単光源および環境光の下で撮影した 5850 枚の画像からなる。認識実験には、正面を向いている 650 枚の画像から、顔付近を切り出した画像を用いた。その際に、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取った。このデータベースには、各画像について両目と口の 2 次元座標が与えられており、画像中の顔の位置や顔の大きさの違いが認識結果に与える影響を避けるために、全ての画像で両目の座標が等しくなるように平行移動、回転、縮小を行い、 $64 \times 64$  の画像を作成した (図 1)。また、単光源方向とカメラの光軸のなす角度  $\theta$  に基づいて、全ての画像は 5 つの部分集合のいずれかに分類されている。本実験では、各人物の学習画像として、 $\theta$  の最も小さい、部分集合 1 に属する 7 枚を用い、その他の部分集合に属する画像をテスト画像として用いた。図 2 に、各部分集合に属する画像の例を示す。

### 4.2 学習

3.2 節で述べたように、学習画像の数が十分でないことから、式 (4) の特異値分解を行い、特異値の大きい上位 3 つのベクトルを基底画像とした (図 3)。こうして得られた基底画像は、ハイライトや影のために、式



図 3: 基底画像

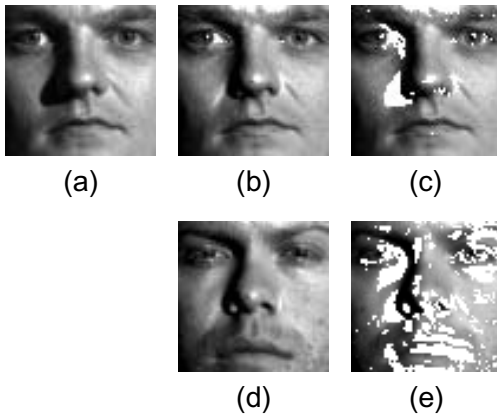


図 4: 処理過程

(3)を満たす保証はない。実際に、図3からも、額や鼻の先端でハイライトの影響を受けていると思われるが、特異値の累積寄与率が、上位3つで  $(96.0 \pm 0.3)\%$  ( $1\sigma$ ) となることから、式(3)からのずれは小さいものと考えられる。

### 4.3 認識

実際に認識を行うには、合成画像の画素値がテスト画像の画素値を正しく再現できているかどうかを判定するための閾値を設定する必要がある。本実験では、基底画像を用いて学習画像を再合成して画素値の誤差の分布の標準偏差  $\sigma$  を求め、閾値の目安とした。

図4に処理過程の画像を示した。テスト画像(a)に対して、同一人物の基底画像を用いて作成した合成画像(b)と、他の人物の基底画像を用いて作成した合成画像(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より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は、鼻や目頭の cast shadow 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、目や口の周辺を中心として比較的広い範囲

Error rate (%)				
Method	Subset 2	Subset 3	Subset 4	Subset 5
LS	0	0	6.5	36.5
LS (Linearization)	0	0	0	10.1

図 5: 提案手法と線形部分空間法の比較

で再現に失敗している。

閾値を  $4\sigma$  に設定したときの各部分集合に対する誤認識率を図5に示した<sup>3</sup>。拡散反射成分のみしか生成できない線形部分空間法と比較すると、拡散反射成分だけでなく attached shadow も生成可能な提案手法では、誤識別率が大幅に減少している。特に、図2からわかるように、光源の方向がカメラの光軸からはずれるにつれて影の寄与が大きくなるために、性能の差が大きくなって現れていると考えられる。しかし、提案手法は attached shadow しか生成できないために、テスト画像中で cast shadow の寄与が支配的になると認識に失敗する場合がある。実際に、部分集合5に対して誤認識が起こっているのは、cast shadow の影響であると考えられる。

次に、同じデータベースを用いた認識実験と比較するために、報告されている結果をまとめて図6に示した。9PL[11]では、主に部分集合4に属する方向の点光源に照らされた9枚の画像を学習画像としているが、それ以外の実験では、部分集合1に属する、各人物あたり7枚の画像を学習画像としている。また、提案手法以外の部分集合5に対する誤識別率は、報告されていない。各実験で実際に認識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関わらず、膨大な数の仮想照明画像を必要とする照明錐モデル[8]と同程度の性能を持つものと考えられる。また、本手法は、未知光源下で撮影された7枚の学習画像を用いているが、特殊な照明下で撮影された9枚の(仮想)画像を用いた認識手法[11]と比較しても、同程度の認識性能を示している。さらに、3枚の基底画像に対して画素のクラスタリングを行う分割線形部分空間法[1]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示していると言える。これらのことから、画像の線形化に基づく提案手法が有

<sup>3</sup> 閾値を  $2\sigma$  から  $6\sigma$  まで変化させても、結果は1%程度しか変わらなかった。



Method	Error rate (%)			
	Subset 2	Subset 3	Subset 4	Subset 5
IC (attached) [8]	0	0	8.6	---
IC (cast) [8]	0	0	0	---
9PL [11]	0	0	2.8	---
Segm. LS [1]	0	0	0	---
LS (Linearization)	0	0	0	10.1

図 6: Yale Face Database B を用いた認識実験結果

効であると同時に、画像の線形化の枠組みが、PIBR だけでなく物体認識に対しても有効であると言える。

## 5 結び

照明変化を伴う物体認識問題に対して、画像の線形化に基づく認識手法を提案し、顔画像データベースを利用した認識実験からその有効性を示した。提案手法は、学習画像の特異値分解で得られる 3 枚の基底画像をそのまま用いる単純な手法であるが、膨大な数の仮想照明画像を用いる照明錐モデル、特殊な照明下で撮影された 9 枚の (仮想) 画像を用いた認識手法、さらに、3 枚の基底画像に対して画素のクラスタリングを行う分割線形部分空間法と同程度の認識性能を持つと考えられる。これは、画像の線形化の枠組みが、画像合成だけでなく物体認識においても有効であることを示している。

本手法は、attached shadow を取り扱うことができないという線形部分空間法の問題点を、画像の線形化の枠組みにより、テスト画像撮影時の照明を推定することで解決するものである。本稿の実験で示したように、提案手法は、単光源下で撮影した画像に対しては有効である。しかし、複数光源や面光源下の画像に対しては、光源の推定が不安定になる可能性があるために、その有効性を検証する必要がある。Lambert モデルを仮定したときの照明錐が 9 次元部分空間で近似できることから [3][18]、基底画像の数を 3 から 9 に拡張することで複数光源や面光源に対しても安定な認識結果が得られることが期待される。さらに、9 枚の基底画像を用いることで任意の照明下の画像から拡散反射成分を取り出すことも可能であると思われるが、実画像を用いたこれらの検証については今後の課題としたい。

また、提案手法は、拡散反射成分と attached shadow を扱うことができるが、ハイライトや cast shadow を

扱えないために、それらが支配的な画像に適用することができない。これは、Lambert モデルが比較的良好に当てはまる物体しか認識できないことを意味する。これに対して、appearance-based methods の枠組みで、Lambert モデルを仮定せずに照明錐の性質を効率よく利用した手法が提案されている [16]。しかし、学習画像取得時と大きく異なる照明下での認識は困難であることが報告されていることから、生成的手法の枠組みでも、ハイライトや cast shadow を考慮した物体認識手法について検討する必要がある。

## 参考文献

- [1] A. U. Batur, and M. H. Hayes III, "Linear subspaces for illumination robust face recognition", In *IEEE CVPR*, pp.II-296-301, 2001.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection", *IEEE Trans. PAMI*, 19(7), pp.711-720, 1997.
- [3] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces", In *Proc. IEEE ICCV*, pp.383-390, 2001.
- [4] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible lighting conditions?", *Int. J. Computer Vision*, 28(3), pp.245-260, 1998.
- [5] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity", In *Proc. IEEE CVPR*, pp.1060-1066, 1997.
- [6] R. Brunelli, and T. Poggio, "Face recognition: features versus templates", *IEEE Trans. PAMI*, 15(10), pp.1042-1052, 1993.
- [7] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: generative models for recognition under variable pose and illumination", In *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.277-284, 2000.
- [8] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: illumination cone

- models for face recognition under variable lighting and pose”, *IEEE Trans. PAMI*, 23(6), pp.643–660, 2001.
- [9] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur, “Illumination cones for recognition under variable lighting: faces”, In *Proc. IEEE CVPR*, pp.52–58, 1998.
- [10] P. W. Hallinan, “A low-dimensional representation of human faces for arbitrary lighting conditions”, In *Proc. IEEE CVPR*, pp.995–999, 1994.
- [11] K. C. Lee, J. Ho, and D. Kriegman, “Nine points of light: acquiring subspaces for face recognition under variable lighting” In *Proc. IEEE CVPR*, pp.I-519–526, 2001.
- [12] Y. Moses, Y. Adini, and S. Ullman, “Face recognition: the problem of compensating for changes in illumination direction”, In *Proc. ECCV*, pp.286–296, 1994.
- [13] 向川康博, 宮木一, 三橋貞彦, 尺長健, “Photometric Image-Based Rendering による仮想照明画像の生成”, *情報処理学会 CVIM 論文誌*, 41(SIG 10), pp.19–30, 2000.
- [14] H. Murase, and S. K. Nayar, “Visual learning and recognition of 3-D objects from appearance”, *Int. J. Computer Vision*, 14, pp.5–24, 1995.
- [15] S. K. Nayar, and H. Murase, “Dimensionality of illumination manifolds in eigenspace”, CUCS-021-94, Technical Report, Department of Computer Science, Columbia University, New York, 1994.
- [16] 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”, *信学技報*, PRMU 2001-207, pp.9–16, 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”, *情処研報*, CVIM 2002-131-20, pp.141–148, 2002.
- [18] R. Ramamoorthi, and P. Hanrahan, “On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object”, *J. Opt. Soc. Am. A*, 18(10), pp.2448–2459, 2001.
- [19] A. Shashua, “Geometry and photometry in 3D visual recognition”, PhD thesis, MIT, 1992.
- [20] M. A. Turk, and A. P. Pentland, “Face recognition using eigenfaces”, In *Proc. IEEE CVPR*, pp.586–591, 1991.