

鏡面反射除去による姿勢・照明変動下の顔照合の高精度化

Specularity Removal for Enhancing Face Recognition under Variable Pose and Illumination

石山 壘[†]
Rui Ishiyama

塚田 正人[†]
Masato Tsukada

1. 概要

照明条件の変動に頑強な顔画像認識手法として、照明部分空間法が広く用いられている[1]。任意の照明条件での画像輝度値の拡散反射成分は低次元の線形部分空間で近似可能なことが示されており[2]、その基底画像は3D形状とアルベド(拡散反射率)から計算可能である[3]。3D形状は、3Dスキャナを用いれば手軽に高精度なデータを得ることができ[4]、また、1枚の画像から推定する手法も提案されている[5]。しかしアルベドは、肌の反射特性や現実環境の照明条件が複雑であるために、画像からの推定はもちろん、計測することも容易ではない。従来手法の多くは鏡面反射成分を無視してアルベドを推定し、照明基底を得ている[4,5,6]。このため、登録画像の鏡面反射成分がアルベドの推定誤差となり、照合性能に影響すると考えられる。

本論文では、照明条件と鏡面反射特性が未知である顔画像から、鏡面反射成分を除去する手法を提案する。提案手法は、一般的な顔画像の性質と3D形状を利用し、肌の反射特性に関して完全散乱等の仮定を用いないため、アルベド推定結果に元画像の鏡面反射成分が混入しない。実験により、提案手法を顔照合に適用することで、照明変動下における照合性能を改善できることを示す。

2. 鏡面反射成分の除去とアルベド推定

提案手法の流れを図1に示す。照明条件が不明である顔画像と、画素毎に対応付けされた3D形状を入力とし、画像から鏡面反射成分と陰影を除去し、アルベドを推定する。3Dスキャナを用いれば、テクスチャ画像とそれに対応付けられた3D形状データを容易に得ることができる。

画像の画素 x の輝度を $I(x)$ とし、アルベド、影を含んだ完全散乱成分、そして鏡面反射成分を $a(x)$ 、 $D(x)$ 、 $H(x)$ とすると、肌の反射モデルは次式で書ける。

$$I(x) = a(x)D(x) + H(x), \quad (1)$$

任意の照明条件での顔画像の拡散反射成分は、低次元の線形部分空間で近似可能である[2]。

$$a(x)D(x) = B(x)^T \lambda. \quad (2)$$

$B(x)$ と λ はそれぞれ、照明基底と照明パラメタである。

顔領域のうち、顔の特徴のエッジがなく、かつ、鏡面反射成分が無視できる領域、すなわち、 $a(x) \sim 1$ かつ $H(x) \sim 0$ である領域を α とする。領域 α では次式が成り立つ。

$$I(x) = D(x) = B(x)^T \lambda \quad \text{for } x \in \alpha. \quad (3)$$

ここで、 $B(x)$ はアルベドを1としたときの照明基底であり、従来研究[2,3]におけるアルベドを含む照明基底とは異なることに注意する。 $B(x)$ の計算に必要なのは3D形状のみであるから、アルベドが未知でも、次式の再構成誤差 E を最小化して、照明パラメタ λ を計算することができる。

$$E = \sum_{x \in \alpha} e(x)^2, \quad e(x) = \{ I(x) - B(x)^T \lambda \}. \quad (4)$$

ただし、この推定を行うためには領域 α を特定しなければならず、それは一般に容易でない。そこで、顔画像の一般的な性質として、以下の仮定を導入する。

(a) 顔全体は、領域 α 以外に、鏡面反射のある領域 β 、目や口などアルベドが変化する領域 γ 、の2つの領域に分けられ、 β と γ の面積は α より十分小さい。

(b) 目、口、鼻などの特徴によるアルベド変化は、肌色より暗い輝度を生じる変化である($a(x) < 1$)。

これらを仮定すると、以下に提案する手法により、鏡面反射成分を推定・除去できる。

式(4)の再現誤差 E ではなく、次式の顔全体に対する再現誤差 E' を最小化し、 λ の近似値 λ' を求める。

$$E' = \sum_{x \in \alpha, \beta, \gamma} |I(x) - B(x)^T \lambda'|^2 \\ = E(x) + \sum_{x \in \beta} e(x)^2 + \sum_{x \in \gamma} e(x)^2 \quad (5)$$

仮定(a)により、式(5)の第2、第3項は第1項にくらべ十分小さく、推定値 λ' は真値 λ の良い近似であるとみなせる。

定義に立ち返ると各領域の輝度値は次のように書ける。

$$I(x) = \begin{cases} D(x) & \text{for } x \in \alpha \\ D(x) + H(x) & \text{for } x \in \beta \\ a(x)D(x) & \text{for } x \in \gamma \end{cases} \quad (6)$$

再現画像 $D(x) = B(x)^T \lambda$ を元の画像 $I(x)$ から差し引くと、

$$I(x) - D(x) = \begin{cases} H(x) & \text{for } x \in \beta \\ -(1-a(x))D(x) & \text{for } x \in \gamma \end{cases} \quad (7)$$

仮定(b)より、 $(1-a(x))D(x) > 0$ が成り立ち、 $H(x) > 0$ であるから、式(7)の値が正となる画素を領域 β として特定でき、鏡面反射成分 $H(x)$ を次式で求めることができる。

$$H(x) = \max(I(x) - D(x), 0). \quad (8)$$

$H(x)$ を元画像から差し引くと鏡面反射成分を除去でき、陰影 $D(x)$ で正規化すればアルベドが得られる。

$$a(x) = \{ I(x) - H(x) \} / D(x) \quad (9)$$

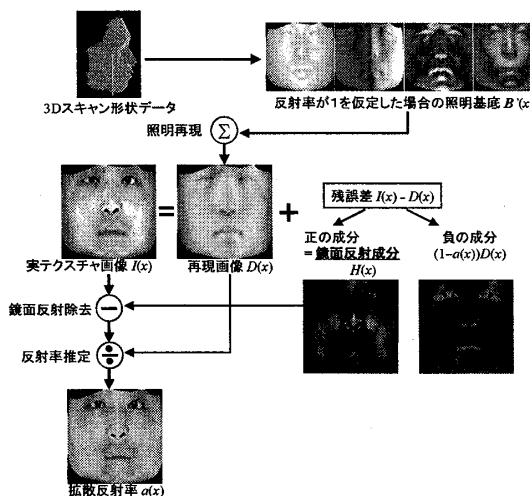


図1: 顔画像から鏡面反射成分を除去する提案手法

[†]NEC 共通基盤ソフトウェア研究所

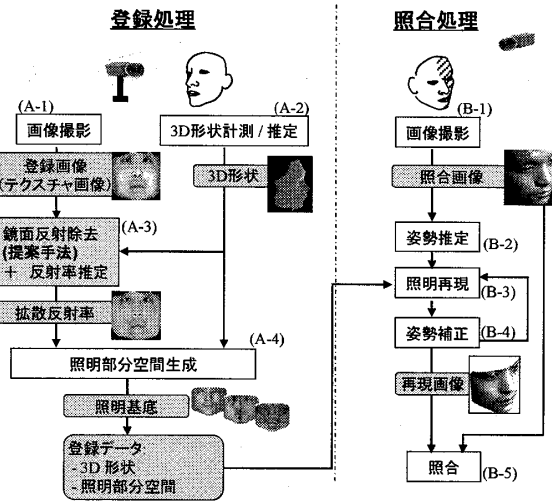


図 3：提案手法を組み込んだ顔照合手法の概要

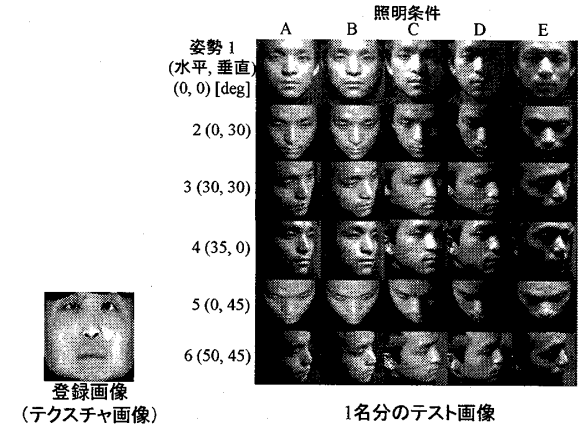


図 2：照合実験に用いた登録画像とテスト画像の姿勢・照明条件

3. 鏡面反射除去による顔照合性能の改善

3.1 顔照合手法への組み込み

2章の提案手法を[4]の顔照合手法に組み込み、鏡面反射の影響を除去し、照合性能を改善できることを示す。顔照合手法の概要を図3に示す。従来手法との違いは、登録画像から鏡面反射を除去しアルベドを推定する処理(A-3)を追加した点である。

登録処理は、3D スキャナによって各個人のテクスチャ画像(A-1)と 3D 形状データ(A-2)を計測し、測地照明基底を計算する(A-4)。従来手法では均一な照明条件で撮影された画像をアルベドに代用していた。提案手法では、画像から鏡面反射を除去し、アルベドを推定する(A-3)。

照合処理では、照合対象の2次元顔画像に対し、登録された各人物の3D形状および照明基底を用いて同一姿勢・照明条件の顔画像を再現し、その再現誤差によって人物を特定する。まず、姿勢推定の初期値を、あらかじめ入力した顔特徴点の画像上および3Dモデル上の位置を用いて計算する(B-2)。測地基底照明を該姿勢での画像に投影して照明基底画像を生成し、それらの線形和で再現画像を生成する(B-3)。そして、再現誤差が最小になるように姿勢推定値を更新する(B-4)。ステップ(B-3)と(B-4)を姿勢推定値が収束するまで繰り返し、最終的に得られた姿勢での再現誤差が最小となった人物を本人とする(B-5)。

3.2 照合実験

200名の人物について、それぞれ30種の大きく異なる姿勢・照明条件(図2右)で撮影した6,000枚の画像をテスト画像とし、顔照合実験を行った。テスト画像の姿勢は上下と左右の各方向の回転角度の合計で最大95度、照明方向は正面から横方向に最大90度まで変化させた。各個人について別途計測した3Dスキャンデータを登録データとした。登録画像であるテクスチャ画像は均一な照明を当て陰影を無くした状態で撮影されたものである(図2左)。

表1に、登録人数200名からの1位照合成功率を示す。従来手法でも、姿勢1~4かつ照明A~Cではほぼ100%の照合性能が得られていた。しかし、登録画像とテスト画像間の姿勢・照明の変動が極めて大きい条件(姿勢5~6かつ照明D~E)では性能が劣化していた。この原因として、影が多く照合に有効な特徴が少ない上に、姿勢・照明方向の違いによって鏡面反射成分が大きく異なっている(登録

表 1：顔照合実験結果 (1位照合成功率)

	Previous method					Avg	Our proposed method					Avg
	A	B	C	D	E		A	B	C	D	E	
1	100%	100%	99.5%	98.0%	92.0%	97.9%	100%	100%	99.5%	99.5%	97.5%	99.3%
2	100%	100%	99.5%	95.5%	90.0%	97.0%	100%	100%	99.5%	99.5%	98.0%	99.4%
3	100%	100%	99.5%	98.5%	90.5%	97.7%	100%	100%	99.5%	99.5%	98.0%	99.4%
4	100%	100%	99.5%	96.5%	93.0%	97.8%	100%	100%	99.5%	99.5%	99.0%	99.6%
5	98.0%	98.5%	96.5%	80.5%	81.0%	90.9%	98.5%	98.5%	98.5%	95.5%	97.0%	97.6%
6	98.0%	100%	100%	98.0%	83.0%	95.7%	99.5%	100%	99.5%	99.0%	95.0%	98.6%
Avg	99.3%	99.8%	99.0%	94.5%	88.3%	96.2%	99.7%	99.8%	99.3%	98.8%	97.4%	99.0%

画像では主に頬の部分に生じていたのに対し、テスト画像は鼻等の一部である)のために、照明再現における誤差が大きくなっていったことが一因と考えられた。

提案手法を用いた場合は、特にこれら厳しい条件での照合性能が大幅に改善された。平均1位照合率は従来法で96.2%であったが、提案手法によって照合エラーの73%が解消され、99.0%へ改善された。

4. まとめ

本論文では、照明条件が未知である顔画像から、鏡面反射成分を除去し、アルベドを推定する手法を提案した。また、提案手法を顔照合へ応用し、照合実験を行った。従来の研究で無視されていた、鏡面反射成分の顔照合性能への影響を明らかにするとともに、提案手法によって顔照合の性能を改善できることを示した。

参考文献

[1] 岩井儀雄, 勞世竈, 山口修, 平山高嗣, 画像処理による顔検出と顔認識, 情報処理学会研究報告, CVIM-149(37), 2005.
 [2] P. N. Belhumeur and D. J. Kriegman, "What Is the Set of Images of an Object Under All Possible Illumination Conditions?," *IJCV*, vol. 28, 245-260, 1998.
 [3] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," *IEEE Trans. PAMI*, vol. 25(2), pp. 218-233, 2003.
 [4] 石山壘, 濱中雅彦, 坂本静生, 顔の3次元見えモデルを用いた任意姿勢・照明変動下の顔画像認識, 電子情報通信学会論文誌, J88-D-II(10), 2069-2080, 2005.
 [5] L. Zhang and D. Samaras, "Face Recognition from a Single Training Image under Arbitrary Unknown Lighting Using Spherical Harmonics," *IEEE Trans. PAMI*, 28(3), 351-363, 2006.
 [6] S. Biswas, G. Aggarwal and R. Chellappa, "Robust Estimation of Albedo for Illumination-invariant Matching and Shape Recovery," *Proceedings of ICCV 2007*, 1-8.