

# 顔画像からの照明条件の推定と 照明ロボット制御への応用

岡田 修弥<sup>1,a)</sup> 栗山 繁<sup>1,b)</sup>

**概要：**芸術的な顔写真を撮影する際には、光源となる照明の位置や明るさは最も重要な要素となる。特に光源のセッティング（ライティング）が被写体の印象を大きく左右するため、プロの撮影現場では被写体の印象を考慮しながらアシスタントが照明器具やレフ板などを配置する。しかしながら、所望する印象や被写体の特徴に応じた細かな調整には芸術的な知識と経験が必要とされる。本研究では、ポートレート写真撮影における顔の「印象」が光源の位置や方向によって生じる陰影の状態と与えられるものと仮定し、撮影画像から生成した特徴量と撮影時の照明条件を対応付けて分類器に学習させることで、入力された所望の印象をもつ画像の照明条件を推定する手法を提案する。また、本手法の応用例として、推定された照明条件を照明ロボットを用いて実世界において再現する機構を提案し、顔画像撮影時の芸術的な印象操作に関わる照明制御の自動化を試みる。

## Estimation of lighting conditions from facial images for robotic lighting

OKADA HISAYA<sup>1,a)</sup> SHIGERU KURIYAMA<sup>1,b)</sup>

### 1. はじめに

写真撮影において光源となる照明の位置や明るさは最も重要な要素である。特に人の顔を被写体とするポートレート写真の撮影においては光源のセッティング（ライティング）が被写体の印象を大きく左右する。図1は照明を1灯のみ使用したライティングの代表例であるが、光源の位置を移動させただけでも被写体の鼻の高さや彫りの深さといった顔の各パーツの際立ちや印象は大きく変化する [1]。プロの撮影現場ではこのような印象を考慮しながらアシスタントが照明器具やレフ板などを配置するが、所望する印象や被写体の特徴に応じた細かな調整には専門的な知識と経験が必要とされる。

写真撮影におけるカメラマンの作業を軽減する研究例として、全身ポートレート写真撮影時に Kinect を用いてポジングの指示や提案をすることで構図の作成を支援する研

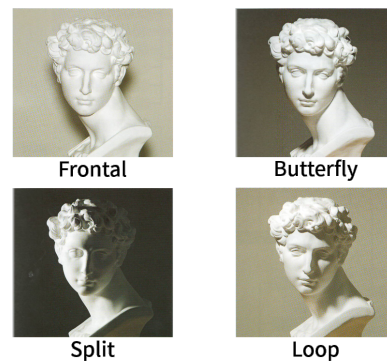


図1 ポートレート撮影における1灯ライティングの例 [1]

究 [2] や、移動体上にカメラを搭載してメイキング画像や同時に複数アングルからの画像を撮影するシステム [3] が提案されている。

撮影時の照明制御に関連する研究例としては、物体に対して異なる位置から照明光を当てた画像を複数枚撮影し、ユーザのスケッチに基づいて所望の陰影状態となるように画像を重畳するシステム [4] が提案されている。また、動的に移動する人物の追尾までも導入した研究例としては照

<sup>1</sup> 豊橋技術科学大学 大学院工学研究科 情報・知能工学専攻  
〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1

a) okada@val.cs.tut.ac.jp

b) sk@tut.jp

明器具を搭載したドローンを用いて所望の半逆光画像を撮影する手法 [5] や移動する複数の人物に対して複数の光源の光量を動的に制御する手法 [6] が提案されている。

画像処理によって顔画像の印象を操作する手法としては、3次元顔形状や静止画から抽出した2次元顔画像に対して陰影マスク画像を重畳し所望の陰影を描画する例 [7][8] が存在する。しかしながら、これらの手法では陰影の調整を誤ると物理的には生じ得ない陰影状態が生じることで不自然な印象を与える危険性がある。

本研究の究極的な目標はプロの照明技師の仕事をロボットが代替することによって撮影にかかる照明の配置や調整のコストを削減し、照明技法に関する芸術的または専門的な知識がなくとも所望の印象や特徴を指示するだけで簡単に印象的な写真を撮影可能にすることである。

本稿ではポートレート写真撮影における顔の「印象」を光源の位置や方向によって顔表面に生じる陰影の状態と定義し、照明を1灯のみ用いるシングルライティングの画像を対象として所望の印象をもつ任意の顔画像から照明条件を推定する手法を提案する。また、本手法の応用例として、推定された照明条件を照明ロボットを用いて再現することで印象の操作に関わる照明制御の自動化を試みる。

## 2. 画像特徴ベースの検索・照明条件の再現

図1のような1灯ライティングの例では、スプリット・ライティングの場合は顔に対して真横から、バタフライ・ライティングでは正面より斜め上方向からというように、印象と照明の位置などの制御変数が対応付けられている。本研究では、ある印象をもつ画像を撮影するためにはどのような制御変数を適用する必要があるのかを事前に把握し、目標とする陰影状態の画像から制御変数を推定することで、撮影者が実際の人物を撮影して得られる画像の印象を目標画像に近づけることを目的とする。

図2に示す概念図のように、照明ロボットを使用して撮影された画像について撮影時の制御変数を同時に記録しておく。ある制御変数を適用した時に得られる画像の特徴が類似しているとすれば、その特徴に基づいて制御変数を推定することができる。そこで、同一の制御変数で撮影された画像群が同じ分類クラスに属すると仮定し、その制御変数の下で得られた共通の画像特徴を学習させることで、制御変数のクラスを推定する分類器を構築する。

画像を用いた実験により、個人差による特徴の変化は照明条件の変化と比較すると影響の度合いが小さいことが判明した。したがって、最初に個人差を無視した一般的な特徴だけをを用いて照明条件を再現する。

## 3. 顔画像からの照明条件推定

本章では、照明条件を推定する上で用いる画像特徴量群を収集する方法について説明する。

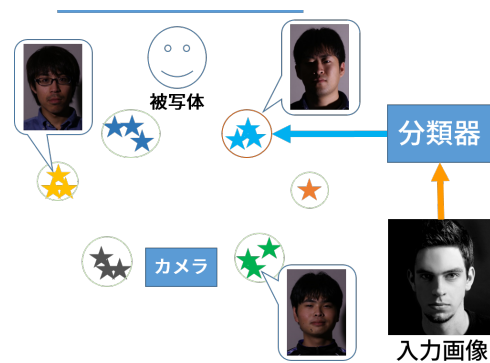


図2 撮影画像と印象の対応付けによる制御変数の推定

### 3.1 照明ロボット

照明ロボットの外観を図3(a)に示す。本ロボットは電源と任意の色・輝度に調光可能なLED電球 (Philips Hue) を3灯搭載したものであり、自由に位置や方向を制御するために iRobot Roomba をベースとする可動の本体から構成される。ロボット本体の制御は搭載された ZigBee モジュールを介してコンピュータとシリアル通信を行い、公開されている制御コマンドを送信する。

撮影時の制御変数としては、照明の設定値 (輝度、色温度) および照明ロボットの位置・方向を取得している。各照明は Philips Hue のブリッジに対して API を介して設定値を更新および取得するが、位置・方向については撮影する室内の天井に図3(b)のように Web カメラを設置し、照明ロボットに搭載された3個の色の異なる LED マーカを検出して算出する。これは Roomba 本体のオドメトリの精度が低いことに起因するものであり、車輪のスリップ等によってロボットの位置推定が困難になることから Web カメラによって得られる画像内のマーカ位置の座標および方向を記録している。

ロボットは図3(c)に示すアプリケーションからロボットの位置および各照明の設定値を制御できるようになっており、人が手動でセッティングを変更する作業が軽減される。また、カメラで撮影された画像が指定ディレクトリに保存されると、画像と撮影時点の制御変数をデータベースに記録する。これにより画像と撮影環境の対応関係を保持するとともに、アプリケーション下部に表示されている画像群から所望の印象の画像を選択すると、記録された撮影環境の制御変数を取り出して再現することもできる。

### 3.2 照明条件と人物の違いによる特徴の変化

実際に照明ロボットを使用し、左側・正面・右側から光を当てた2名の写真に対し、輝度および輝度勾配を疑似カラーで表現した画像を図4に示す。

照明位置が変化すると輝度の分布や輝度勾配の分布は大きく変化している。一方、鼻や頬といった顔の各パーツによって異なる陰影が生じるが、被写体の違いによって生じ

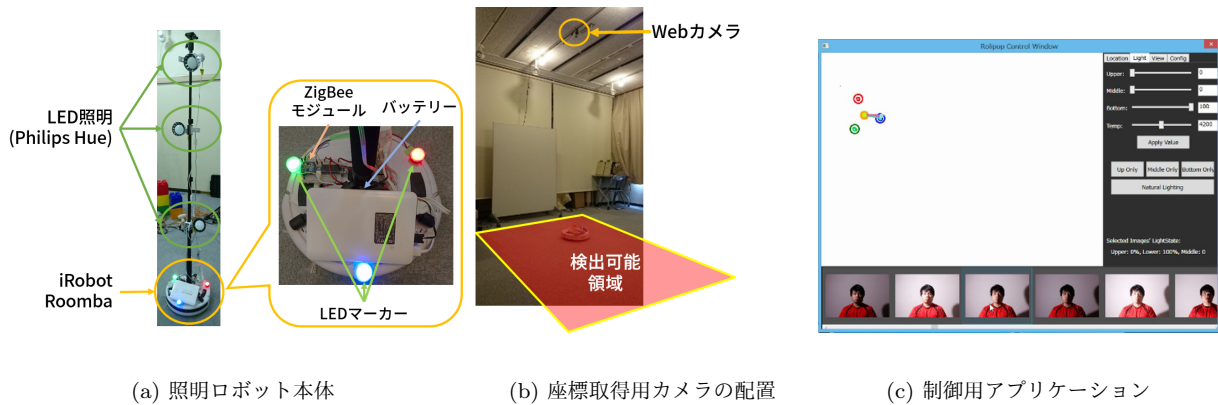


図 3 ロボット照明システム

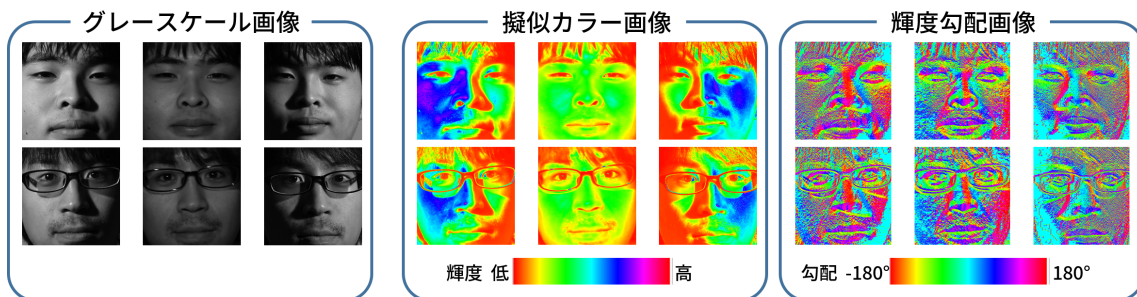


図 4 光源位置および個人差による輝度・輝度勾配の変化

る陰影の変化は照明位置に伴う変化よりも小さいことが確認できる。

以上の観察を踏まえ、被写体の特徴の違いによる差異は照明条件を推定したあとに調整するものとして、一枚の画像から代表的な陰影情報の抽出を試みる。

### 3.3 局所領域ごとの特徴量計算

顔の陰影が照明によってのみ変化することを前提として、それを識別・分類するために、各照明条件下で撮影された画像において顔部分における陰影の分布に着目する。すなわち、陰影の部分における明暗や勾配といった情報は照明条件によって大きく変化することを利用し、特徴として用いる。

このとき、画像全体から特徴を抽出するだけでは、各部位における特徴を捕捉できない。一方、顔の部位ごとに特徴を抽出しようとする個人差による影響を強く受け、スプリット・ライティングのような顔の半分以上が影となる画像では部位の検出自体に失敗する可能性が高くなる。

そこで、図 5 に示すように入力された顔画像を等間隔に分割し、それぞれの領域を局所領域として輝度および輝度勾配のヒストグラムを生成する。人の顔パーツの配置は個人ごとには大きく変化しないため、局所領域は大まかな顔の部位ごとの特徴を接続したベクトル表現となり、顔の向きが正面でない場合や顔領域の切り出し時に生じるズレに対してもある程度頑健となる。

これを (ヒストグラムのビン数) × (分割領域数) 次元の特徴量として扱い、同じ制御変数の代表値をもつすべての画像に対して SVM を用いて特徴量を学習し分類器を構築する。次章での実験では画像を縦横に 5 分割した合計 25 個の局所領域に対し、ビン数 20 の輝度および輝度勾配ヒストグラムを計算することで 1000 次元の特徴ベクトルを生成する。

## 4. 実験

前章で提案した特徴量を照明条件が既知のデータセットに対して計算し、特徴量から照明条件を推定する精度を検証した。学習器にはマルチクラス SVM を使用した。

### 4.1 検証用データセット

検証には 2 種類のデータセットを使用した。1 つ目は照明ロボットを使用して実際の環境で撮影したものであり、照明条件を上下 2 パターンおよび位置・角度 6 パターンの計 12 パターンの照明条件で 7 人の被写体を撮影した。被写体は眼鏡をかけた状態と外した状態で各状態 84 枚の画像を撮影し、顔識別器または手動で撮影画像から顔部分を切り出したものをそれぞれデータセットとして使用した。なお、照明色の制御には色温度を使用し、白色蛍光灯の色温度に相当する 4200[K] に設定した。画像の例を図 6 に示す。

2 つ目のデータセットは各種照明条件下における顔認

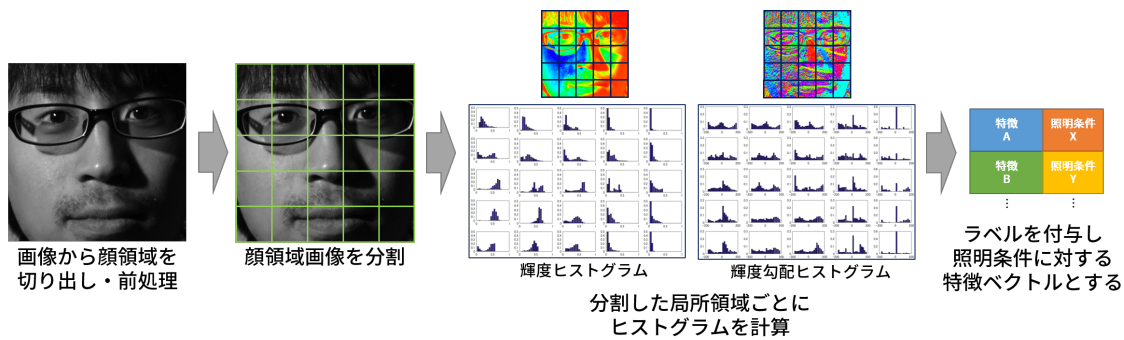


図 5 特徴量の生成

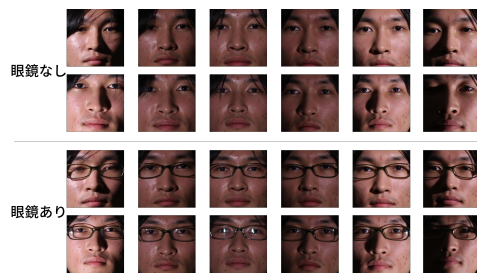


図 6 照明ロボットで撮影したデータセットの例

表 1 特徴量ごとの分類誤差 [%]

| 特徴量  |    | データセット        |               |        |
|------|----|---------------|---------------|--------|
|      |    | 照明ロボット (眼鏡有り) | 照明ロボット (眼鏡無し) | Yale B |
| 輝度   | 全体 | 50.0          | 59.9          | 69.8   |
|      | 局所 | 31.0          | 28.6          | 19.5   |
| 輝度勾配 | 全体 | 29.8          | 23.8          | 26.8   |
|      | 局所 | 32.1          | 25.0          | 16.4   |
| 併合   | 全体 | 20.2          | 25.0          | 23.1   |
|      | 局所 | 26.1          | 20.2          | 13.8   |

識精度を検証する目的 [9] で収集された Extended Yale B Database [10] (以降, Yale B) である. これは 1 灯ライティングで 64 パターンの照明条件下で撮影されたものであり, 眼鏡をかけていない 38 人の被写体に適用した画像が収録されている. ただし, 一部の画像は全体にノイズが載っているなどの悪条件のため対象から除外し, 2,414 枚を有効データとして用いた.

なお, 分類器の性能評価は Yale B は 10 分割交差検定, 照明ロボットで撮影したデータセットについては Leave-one-out 法を用いて検証した.

#### 4.2 特徴量による推定精度の比較

各データセットの画像に対して前章で述べた提案手法を用いて照明条件の推定精度を検証した. 特徴量は局所領域の輝度ヒストグラム, 輝度勾配ヒストグラム, および 2 つの特徴量のベクトルをそのまま接続した併合特徴量の 3 種類について, 画像全体から特徴を抽出した場合と局所領域に分割した場合の分類誤差を比較した. 画像全体からの特徴は局所領域の 1 領域分の生成方法と同じ手法で各特徴量につき 20 次元, 併合特徴量では計 40 次元ベクトル量となる. 各データセットに対する特徴量ごとの分類誤差を表 1 に示す.

画像全体から抽出した輝度ヒストグラム特徴のみを利用した場合は, 多くの画像で推定に失敗することが確認された. しかし, 局所領域に分割することで部位ごとの陰影特徴が表現されるため, 全体から特徴抽出した場合と比較すると 19~50%程度精度が向上することが判明した.

また, 輝度勾配ヒストグラムは光が顔の各部位に当たる

ことで生じる陰影の勾配傾向が各照明条件で類似していることを特徴として利用している. 画像全体から特徴抽出した場合と比較すると輝度ヒストグラムよりも分類精度が高いことが示されているが, 一方で局所領域に分割したことで Yale B データセットでは精度が向上しているにもかかわらず照明ロボットで撮影したデータセットでは精度が 5%前後低下している. これは Yale B データセットでは顔の切り出し範囲が一定であるが, 照明ロボットで撮影した画像は切り出し範囲にばらつきがあり, 局所領域における輝度勾配の傾向が画像ごとに異なることで生じたものと考えられる.

さらに, Yale B データセットおよび眼鏡無しのデータセットでは特徴量を併合して使用することで分類精度が向上し, 局所領域に分割した場合には分割しない場合と比較して 4~9%以上精度が向上した. このことから, 眼鏡をかけていない画像に対しては局所領域ごとに特徴量を計算し複数の特徴を利用して学習させることで照明条件を推定できることが示唆された.

一方で, 眼鏡有りのデータセットでは局所領域に分割することで推定精度が低下することが確認された. これは眼鏡のフレーム幅や生じる陰影量が変わることによって局所領域での特徴の分散が大きくなったことが原因として考えられ, 特徴量の重み付けなどによってこれらの影響を考慮する必要があることが示唆された.

#### 4.3 平均適合率に基づいた検証

全ての照明条件において正解した人物の画像と誤推定された照明条件の画像を比較すると, 図 7 のように同様な印

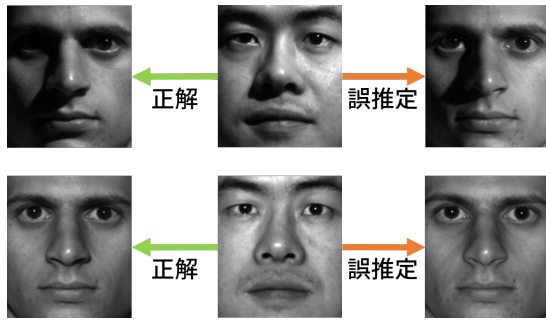


図 7 誤推定の例



図 8 主観評価用アンケートフォーム画面

象が得られていることが確認された。これを踏まえ、クラス（照明条件）の分類スコアをソートし、クラス全体の平均適合率（Mean Average Precision : mAP）および上位3件または5件以内の平均適合率を算出し、正解クラスが上位にランキングされることを検証した結果を表2に示す。

表 2 平均適合率

| データセット           | 指標 | 指標    |       |       |
|------------------|----|-------|-------|-------|
|                  |    | mAP@3 | mAP@5 | mAP   |
| Yale B           | 全体 | 0.846 | 0.855 | 0.858 |
|                  | 局所 | 0.911 | 0.915 | 0.916 |
| 照明ロボット<br>(眼鏡なし) | 全体 | 0.835 | 0.843 | 0.845 |
|                  | 局所 | 0.869 | 0.878 | 0.879 |
| 照明ロボット<br>(眼鏡あり) | 全体 | 0.887 | 0.892 | 0.892 |
|                  | 局所 | 0.849 | 0.855 | 0.857 |

照明条件の推定は最も分類スコアが高いものが選ばれるが、いずれのデータセットにおいても局所特徴では上位5件以内に0.85以上の平均適合率で正解とされる照明条件が含まれていることが確認された。また、眼鏡有りのデータセットに対する結果を見ると、正解数は低下したものの、他のデータセットの平均適合率に近い結果が得られた。

#### 4.4 主観評価による分類結果の検証

分類器による分類結果を人の目で観察した場合の類似性を検証するため、前節のYale Bデータセットで学習した分類器に対する分類スコアのランキング結果を用いて15人の被験者にアンケート形式の主観評価を実施した。

調査では図8のように1枚の対象画像に対して4枚の照

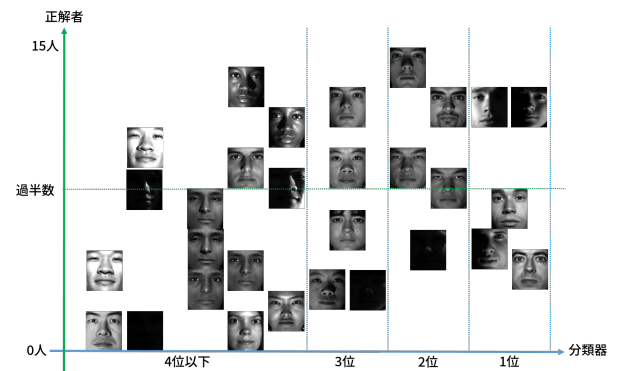


図 9 正解の照明条件を選択した被験者数と分類ランキングの相関

明条件の比較画像を提示し、類似していると思える照明条件の比較画像を全て選択させた。ただし、対象画像は正解の照明条件がランキングの1,2,3位に含まれている画像群から5枚ずつの計15枚の画像を計算に依る特徴識別の容易な例として抽出し、さらに4位以下の画像群から無作為に抽出した15枚の画像を特徴識別の困難な例として加えた計30枚の画像である。また、比較のために提示する画像は、異なる人物に対して対象画像と同じ照明条件を適用した画像とそれ以外の上位3位のランキングから選択した画像を用いた。なお、比較画像は全出題で同じ人物である。

主観評価の結果を基に同じ照明条件の比較画像を選択した被験者数とランキング順位の相関を図9に示す。

分類器では分類が困難であった4位以下の中から選択された対象画像の多くは顔表面の反射が強いものや、逆に反射が弱く陰影が顔表面に現れにくい人物であることが多い。

これらの画像の照明条件の多くは人間の目でも違いを見分けることが困難であるが、鼻筋などに特徴的な陰影が現れる画像においてはそれらの情報を元に人が類似している照明条件を判断している傾向が見られた。ゆえに、特徴量の局所領域ごとの重み付けや皮膚の反射成分と陰影成分を分離して学習することで、より照明条件による分類精度の向上が見込まれることが示唆された。

同様に、照明が正面から当たるような陰影の少ない画像についても、照明条件の判別が困難であることが傾向として読み取れる。これは、正面からのライティングによって陰影の特徴よりも皮膚のシミやテクスチャが目立ち、これらを陰影と解釈してしまうことが原因の一つとして考えられる。

一方で、分類器が正解と判断した照明条件に対して、人が必ずしも正解の照明条件を選択するとは限らないことが判明した。このような傾向は対象画像と比較画像の顔立ちが異なることで陰影形状によっては照明条件とは異なる基準で印象が決定される可能性があることを示唆しており、人間の主観により近い指標によって判断する必要があるといえる。

次に、被験者の過半数が類似していると選択した画像の

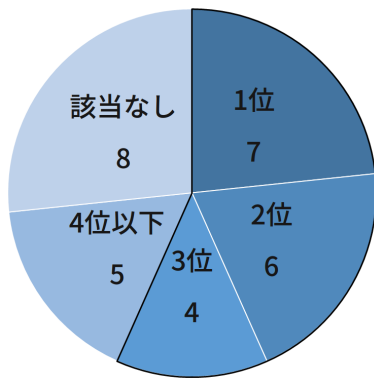


図 10 被験者の選択と分類ランキングの順位

分類ランキングにおける順位を図 10 に示す。

「該当なし」は過半数の被験者が類似画像が存在しないと回答した画像であり、この場合には顔のほぼ全面が陰影となり顔の一部または全てが見えない画像も多く含まれている。この結果から、被験者の選択は概ね分類器の提示する候補で充足されることが示唆されたが、一方で識別器の分類基準と人の視覚認知による分類基準との隔たりも確認された。

## 5. 照明ロボットによる印象の再現

応用例として、照明ロボットを使用して撮影したデータセットで学習した分類器に Yale B データセットの画像から抽出した特徴量を入力し、その推定結果の環境下で撮影した画像の印象を比較した。抽出に用いた画像と撮影した画像の比較例を図 11 に示す。

2つのデータセットは照明の種類や位置の対応関係は無く、推定対象の照明条件も 12 パターンと少ない。しかし、入力画像に近い条件のクラスへの推定がなされ、鼻筋の陰影や頬の光沢感が類似した照明条件が提示されていることが確認できた。

## 6. おわりに

本稿では、芸術的なポートレート写真撮影における照明制御のロボットによる自動化を目的とし、1枚の画像から照明条件を推定する手法を提案した。同じ照明条件で撮影された画像群を用いた実験結果より、類似の照明条件については実用上問題のない精度で推定できることが判明した。

ユーザが所望する印象をもつ入力画像の中には露光値が不明であったり、光量の異なる複数の光源やレフ板などを使用して撮影されている状況が起こり得る。このような場合において現在採用している特徴量の有効性は明らかにされておらず、今後の検証が必要である。

また、主観評価の結果より、個人の顔立ちや皮膚テクスチャによっても印象が変化する傾向が見られた。これを踏まえ、より人の主観に近い推定結果を提示することで所望の印象に近い画像を撮影できると考える。

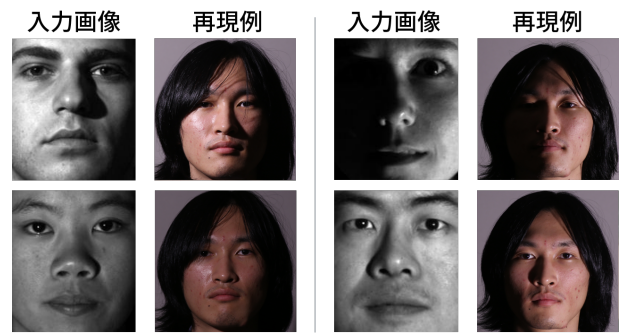


図 11 入力画像に対する印象の再現例

今後は照明条件の推定に関するロバスト性の向上を図るとともに、想定する照明ロボットの自律制御システムの開発にも並行して取り組んでいく。

## 参考文献

- [1] 河村民子: 基礎から始める, プロのためのポートレートライティング, 玄光社 (2013).
- [2] Hongbo Fu, Xiaoguang Han, Quoc Huy Phan.: *Data-driven Suggestions for Portrait Posing*, SIGGRAPH Asia 2013 Technical Briefs, 29, ACM, 2013.
- [3] Yuji Kokumai, Hideki Sasaki, Tomomi Takashina, Yutaka Iwasaki.: *A New Photography Style Using a Shooting Assistant Robot*, SIGGRAPH Asia 2013 Posters, 24, ACM, 2013.
- [4] Ankit Mohan, Jack Tumblin, Bobby Bodenheimer, Cindy Grimm, and Reynold Bailey.: *Table-top Computed Lighting for Practical Digital Photography*, ACM SIGGRAPH 2006 Courses, 2006.
- [5] Manohar Srikanth, Kavita Bala, and Frdo Durand.: *Computational Rim Illumination with Aerial Robots*, Proceedings of the Workshop on Computational Aesthetics, ACM, 2014.
- [6] 新宮淳, 亀田能成, 西口敏司, 美濃導彦: 複数人物に対する照明の動的制御, 信学技報 PRMU2002-135 pp.37-42, 2002.
- [7] 石垣智子, 高井勇志, 松山隆司: 陰影効果を利用した顔印象の編集 (顔とコミュニケーション), 電子情報通信学会技術研究報告.HCS, ヒューマンコミュニケーション基礎 (2008), 108 巻, 238 号, pp.21-25.
- [8] YiChang Shih, Sylvain Paris, Connelly Barnes, William T. Freeman, and Frdo Durand.: *Style Transfer for Head-shot Portraits*, ACM Trans. Graph. 33, 4, Article 148 (July 2014), 14 pages.
- [9] Kuang-Chih Lee, Jeffrey Ho, and David J. Kriegman.: *Acquiring Linear Subspaces for Face Recognition under Variable Lighting*, IEEE Trans. Pattern Anal. Mach. Intell. 27, 5 (May 2005), 684-698.
- [10] The Extended Yale Face Database B, 入手先 (<http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>), (2016.6.2).