

顔画像を用いた顔向きの変化に頑健な人物属性（年齢・性別）推定技術

山本 一真

増田 誠

顔画像から年齢や性別などの人物属性を推定する技術は多くの応用が考えられる。また、様々な顔向きの顔画像に対して年齢や性別を推定することができれば、カメラの設置条件を緩和することができ、応用範囲を広げることができる。そこで本稿では、OKIが開発している顔向きの変化に頑健な顔画像を用いた人物属性推定技術について紹介する。

はじめに

現在、町のいたるところにカメラが設置されており、今後その数はさらに増えることが予想される。これらのカメラから得られる画像を利用し、顔画像から年齢や性別などを推定することにより多くの応用が考えられる。たとえば、デジタルサイネージなどのディスプレイ上部にカメラを設置して広告視認者の年齢や性別を推定することで、視認者に応じた広告を表示したり広告の効果を測定したりすることができる。また、小売店のPOP（Point of purchase advertising）広告や立て看板などにカメラ設置して来店者の年齢や性別を推定することで、来店者が商品の購入を検討するまでの購買行動を分析することができる。また、商店街やショッピングセンターなどに設置されている監視カメラを利用して来店者の年齢や性別を推定することで、曜日や時間帯による客層の変化を把握したり販売促進の効果を測定したりすることができる。これらの応用を可能とする製品として、映像から人の姿や顔を検出し顔画像から年齢・性別を計測する広告効果測定支援システム「RESCAT®」*1) が販売されている¹⁾。

これらの用途で撮影される顔画像は正面を向いているとは限らない。そのため、年齢や性別などの人物属性推定の応用範囲を広げるためには、様々な方向を向いた顔画像へ対応する必要がある。たとえば監視カメラで撮影された画像では、図1のようにカメラに対して斜め下を向いた顔画像になる。また、小売店のPOPや立て看板にカメラを設置し購買行動を分析する場合には、来店者が自由に商品を見るためカメラに対して正面を

*1) RESCAT、FSE は沖電気工業株式会社の登録商標です。

向いていない顔画像になる。

近年、顔画像から年齢や性別などを推定する様々な技術が提案されている^{2)、3)}。しかし、多くの技術は正面を向いた顔画像を対象としており、応用範囲に限られる。そこで本稿では、OKIが開発している顔向きの変化に頑健な人物属性推定技術について紹介する。



図1 監視カメラ画像の例

顔画像からの人物属性推定

顔画像から年齢や性別などの人物属性を推定する処理の一般的な流れを図2に示す。まず、画像中から顔が存在する場所を探す顔検出処理を行う。次に、見つけた顔の中から目・鼻・口といった顔のパーツ（特徴点）が存在する場所を探す特徴点抽出処理を行う。最後に、顔画像と特徴点をもとにその顔の年齢や性別を推定する属性推定処理を行う。

顔検出処理と特徴点抽出処理については、OKIの顔認識ミドルウェア「FSE®」*1)などの既存方式を利用することができる。そのため、本稿では属性推定処理を対象とする。また、以降の手法の説明では属性として性別を利用する。

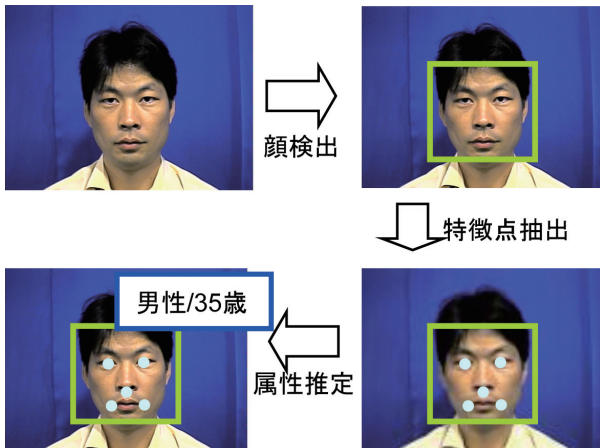


図2 顔画像からの人物属性推定の概要

機械学習による属性推定

属性推定処理については、一般に機械学習の手法を用いた方式が利用される。この方式は図3に示すように、属性の識別方法をあらかじめ学習しておく「学習」と、学習した識別方法をもとに属性を識別する「識別」とからなる。

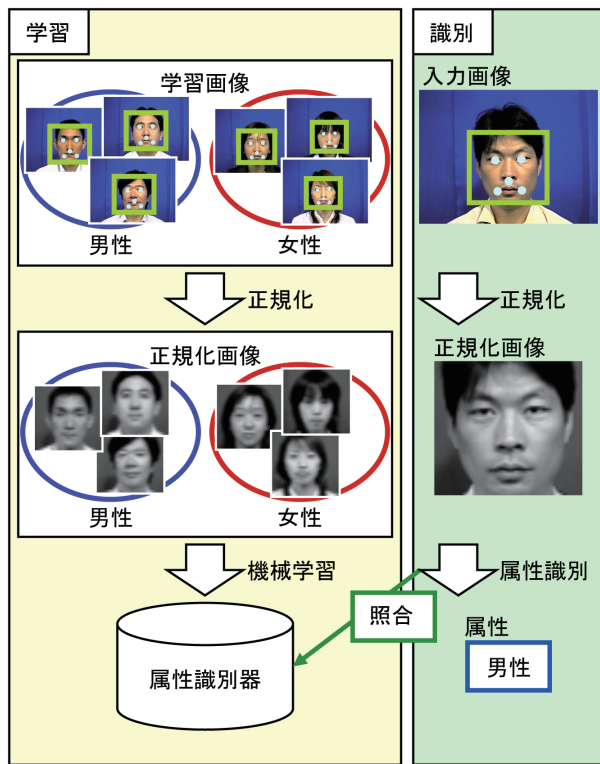


図3 機械学習による属性推定の概要

学習では、属性の分かっている大量の画像データを入力とする。まず、個人による特徴点の位置の違いや撮影状況による見え方の違いを吸収した正規化画像を作成する正規化処理を行う。次に、属性識別器を作成する機械学習を行う。機械学習では、正規化画像をもとに属性を識別するために有効な画像特徴を複数選択し、それを組合せる。これにより、高い精度で属性を識別することができる。男性と女性を分ける画像特徴として、たとえば口紅の有無により口周辺の明るさが異なるといった特徴や眉の形が異なるといった特徴が選択される。大量の画像データを利用することで、これらの画像特徴を自動的に選択することができる。

識別では、特徴点抽出された顔画像を入力とする。まず、学習時と同様に正規化画像を作成する正規化処理を行う。次に、学習した属性識別器と正規化画像を照らし合わせることで、顔画像の属性を識別する。

機械学習による属性推定では、入力となる画像を作成する正規化処理が重要となる。正規化を適切に行えない場合、属性の違いによる画像特徴の差より個人や撮影状況による画像特徴の差の方が大きくなる。そのため、どのようなアルゴリズムを利用しても属性を識別する画像特徴を適切に学習することができず、精度の高い属性推定を行うことができなくなる。

位置合わせによる正規化処理

属性推定における正規化処理では、一般に特徴点の位置合わせによる正規化が利用される。これは、特徴点の位置があらかじめ定義した座標になるよう2次元画像上での変換を行うことで顔画像を正規化する手法である。

様々な向きの顔画像に対して特徴点の位置合わせによる正規化を行った例を図4の位置合わせによる正規化画像に示す。この例を見ると、目・鼻・口といった特徴点の位置が正規化され、同じ位置になっていることが分かる。しかし、入力画像の顔向きが異なると正規化した画像の見え方が大きく異なる。この正規化画像を利用すると、同じ人物の顔でも顔向き毎に抽出できる画像特徴が大きく変わる。そのため、属性を識別する画像特徴を適切に学習することができない。

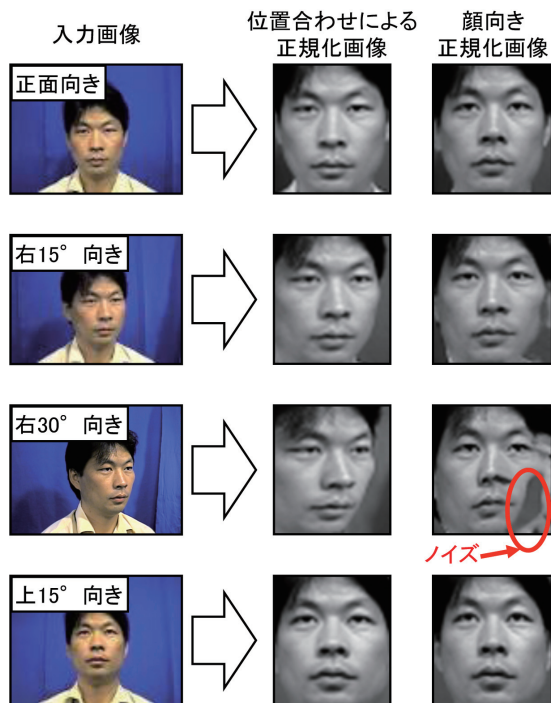


図4 位置合わせによる正規化画像と顔向き正規化による正規化画像の例

顔向きの変化に頑健な属性推定方法

OKIの属性推定手法は、顔向き正規化と二段階の属性識別とを組み合わせることで、顔向き変化への頑健性を向上している。

顔向きの違いによる画像特徴の差を小さくするために、顔の向きを合わせる正規化を行う。射影変換に基づいた3次元正規化法⁴⁾を利用することで、1枚の顔画像から3次元形状モデルを用いて顔の位置、大きさ、回転、向きを同時に正規化することができる。射影変換に基づいた3次元正規化法の概要を図5に示す。この正規化法では、2次元実画像における特徴点の2次元座標群 W と、平均的な顔形状の3次元モデルにおける特徴点の3次元座標群 S との対応関係を射影変換行列 M で表す。これら3つの行列の関係は $W=M \times S$ で表され、射影変換行列は、以下の式を計算することで求めることができる。

$$M = W \times \left\{ S^T (SS^T)^{-1} \right\}$$

つまり、2次元実画像で抽出した目・鼻・口といった特徴点と3次元モデル上の特徴点とから射影変換行列 M を求めることができる。この射影変換行列 M をもとに、2次元実画像が3次元モデルと同じ向きを向いた顔を正規化画像として作成することができる。

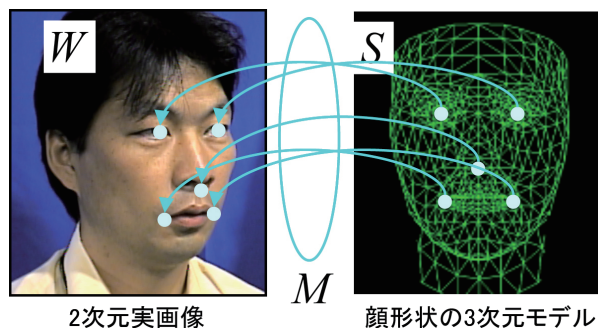


図5 射影変換に基づいた3次元正規化法の概要

様々な向きの顔画像に対して射影変換による顔向き正規化を行った例を図4の顔向き正規化画像に示す。位置合わせによる正規化画像と比較すると、画像の左側は入力画像の顔向きによらず同じ見え方となるのが分かる。これにより、顔向きによらず同じ画像特徴を抽出することができるようになる。そのため、属性を識別する画像特徴を適切に学習することができる。しかし、右30°向きのように大きく横を向いた顔の場合には、ノイズが現れる領域が大きくなるのが分かる。このノイズは、隠蔽の影響で2次元実画像上では見えない領域をもとに正規化画像を作成するため現れる。ノイズがある画像では正しく属性を識別することができない。

そこで、学習時に行う機械学習と識別時に行う属性識別とをそれぞれ二段階に分割することでノイズの影響を削減する。学習時に行う機械学習では、まず、入力となる正規化画像を顔の部分毎に分割し、それぞれの部分で属性識別器を学習する。次に、顔向き毎に回帰分析を行うことで部分毎の属性識別器の出力を重み付けして統合する方法を学習する。識別時に行う属性識別では、まず、正規化画像を顔の部分毎に分割し、それぞれの部分で属性識別を行う。次に、学習した顔向き毎に異なる重みを利用して部分毎の属性識別結果を統合し最終的な属性を決定する。右30°向きの顔に対し二段階の属性識別を行う例を図6に示す。まず正規化画像顔を左右に4分割し、それぞれの部分で属性を識別する。次に、右30°向きといった顔向きにあわせて学習した重みで部分毎の識別結果を統合することで男性0.8、女性0.2といったスコアを得る。男性のスコアが高いため、最終的に男性と識別している。二段階の属性識別を行うことで、顔向きにあわせてノイズのある部分の影響を除外することができる。そのため、顔向きが大きな場合にも正しく属性を識別することができる。

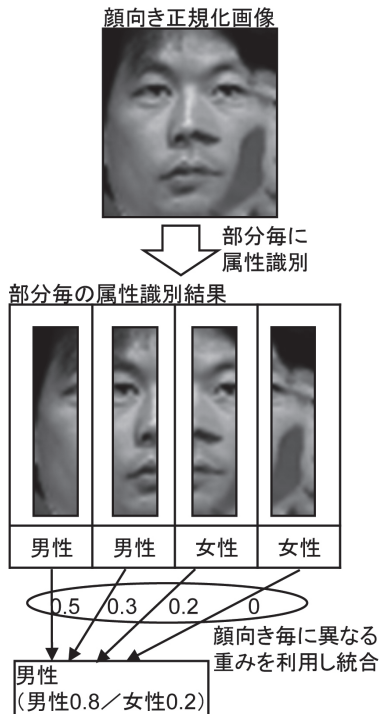


図6 二段階の属性識別の例

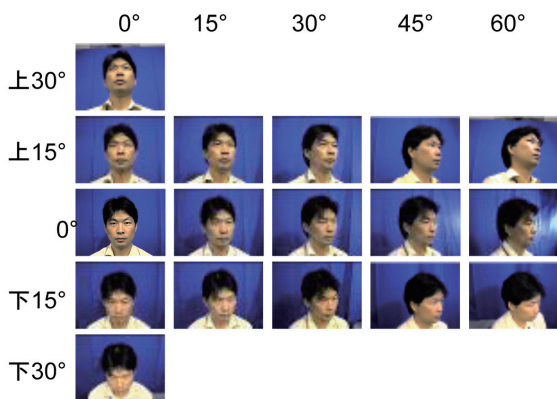


図7 顔向き毎の評価画像の例

精度評価

OKIの属性推定手法により顔向きの変化への耐性が向上したことを確認するために、精度評価を行った。画像特徴にHaar-like特徴を利用し、機械学習にBoostingを利用した²⁾。評価データとしては、図7に示す顔向きで撮影した顔画像を利用した。それぞれの向きで290人分の顔画像に対し、人手で入力した特徴点を利用して、年齢と性別の推定精度を評価した。年齢推定精度の評価指標は平均絶対誤差 (MAE) を利用した。これは、正解年齢から平均して何歳間違えるか

を表す指標であり、小さいほど精度がよい。性別推定精度の評価指標は、性別推定の正解率を利用した。位置合わせによる正規化を行い顔全体で一度に属性識別を行う比較手法と本手法との精度評価結果を表1に示す。比較手法と比べて本手法は、年齢・性別ともに全範囲で精度が向上した。このことから、顔向き正規化と二段階の属性識別により顔向き変化への耐性が向上したことが分かる。また、年齢・性別ともに正面でも精度が向上した。このことから、顔向き変化がない場合に精度劣化がないことが分かる。

表1 精度評価結果

	年齢推定MAE		性別推定正解率	
	正面	全範囲	正面	全範囲
比較手法	9.0	12.0	92%	74%
本手法	6.0	8.0	95%	88%

まとめと今後の課題

本稿では、顔向き正規化と二段階の属性識別を組み合わせることにより、顔向き変化に頑健な人物属性推定を行う手法を紹介した。今後の課題として、対応できる顔向きの範囲をさらに広げることや、マスクやサングラスといった装着物への対応を考えている。 ◆◆

参考文献

- 1) 塚本明利, 他: 映像認識による広告効果測定支援システム「RESCAT」, OKIテクニカルレビュー第218号, Vol.78 No.1, p.32-35, 2011.
- 2) Y Fu, et al.: "Age Synthesis and Estimation via Faces: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32 No.11, p.1955-1976, 2010.
- 3) E Makinen, R Raisamo: "An experimental comparison of gender classification methods", Pattern Recognition Letters, Vol.29 No.10, p.1544-1556, 2008.
- 4) 小坂谷達夫, 山口修: 顔認識のための射影変換に基づいた3次元正規化法, 電子情報通信学会技術研究報告 PRMU, Vo.105 No.375, p.49-54, 2005.

筆者紹介

山本一真: Kazuma Yamamoto. 研究開発センサセンシング技術研究開発部
 増田誠: Makoto Masuda. 研究開発センサセンシング技術研究開発部