

低解像度画像による性別推定

荻野洋夢¹ 吉田武史¹ 鷲見和彦¹ 波部齊² 満上育久³

概要：近年、至る場所で人物属性推定とデジタルサイネージを利用した広告が実用化されており、またカメラに写った人物の属性情報がマーケティング等に役立てられている。しかし多くの場合、カメラの正面に立った人物の詳細な顔情報が使用される。例えば、カメラに写っている詳細な顔情報を得られない歩行者の人物属性が推定できればデジタルサイネージや広告への誘導が可能になり、より効率的な情報発信が可能になると考えられる。また顔の詳細な情報を得ないのでプライバシーの問題にも配慮できる。そこで本研究では、詳細な顔情報が得られない歩行者の全身を対象とした性別推定を、物体検出や一般物体認識などで広く使われている特徴量を用いて行い、結果の比較・検討を行うことを目的とする。

キーワード：人物属性推定

1. はじめに

近年、至る場所で人物属性推定とデジタルサイネージを利用した広告が実用化されている。またカメラに写った人物の属性情報がマーケティング等に役立てられている。しかし多くの場合、カメラの正面に立った人物の詳細な顔情報が使用される。例えば図 1-1 のような、カメラに写っている詳細な顔情報が得られない歩行者の人物属性が推定できるようになれば、歩行者をデジタルサイネージや広告に誘導できるようになり、より効率的な情報発信ができるようになると思われる。また顔の詳細な情報を得ないのでプライバシーの問題にも配慮できる。

そこで本研究では、詳細な顔情報が得られない歩行者を対象とした性別推定を、物体検出や一般物体認識などで広く使われている特徴量を用いて行い、結果の比較・検討を行うことを目的とする。歩行者の全身画像を対象とし、まず特徴抽出を行い、性別を推定する情報を取得する。取得した情報を用いて機械学習を行い、男性か女性の 2 クラス分類問題を解き、性別の推定を行う。



図 1-1 撮影画像例

2. 関連研究

人物属性推定の代表的研究として、Moghaddam らは人物の顔画像を用いた性別推定を行っている[1]。この研究では、様々な識別器を用いて実験を行い、結果を比較している。Support Vector Machine (SVM)[2]を用いた推定において、実際の人の推定精度を上回る結果を残している。

Ueki らは顔のみではなく、顔と上半身画像を用いた性別推定を行っている[3]。顔、髪型、衣服の 3 種類の領域ごとに特徴抽出と機械学習を行い、最後に各領域の結果を統合している。特徴抽出に Principal Component Analysis (PCA) を、識別器に Gaussian Mixture Model (GMM)[4]を用いている。顔のみに比べ、統合した結果の方が高い精度を残している。

松濤らは天井に取り付けられた監視カメラを用いて、人物トラッキングを行い、その結果を用いて検出された人物の荷物の有無や性別の属性解析を行っている[5]。トラッキングは背景差分とブロップトラッキングを組み合わせで行っている。属性解析は、人物の位置や動きの変化を利用する手法と、人物の輪郭の時間変化を利用する手法の 2 種類の手法で実験を行っている。人物の位置や動きの変化を利用する手法では、人物領域の重心座標や面積、縦横比を特徴とし、SVM を用いて分類を行っている。輪郭の時間変化を利用する手法では、フーリエ記述子[6]を特徴量とし、動的計画法、GMM、k-means[7]を用いて分類を行っている。性別は人物の位置や動きの変化を利用する手法において、荷物の有無は輪郭の時間変化を利用する手法において良い結果が得られている。

Cao らは正面を向いた人物の全身を対象とした、PBGR 特徴量と Ada Boost[8]を改良した識別器を用いた性別推定を行っている[9]。

岩竹らは Bag Of Features[10]と Random Forests[11]を用いて、前向きと後向きの 2 種類の向きに着目した性別推定を行っている[12]。

しかし多くの場合歩行者は、詳細な顔情報を得ることは難しく、また必ずしも真正面を向いているわけではない。

1 青山学院大学
Aoyama Gakuin University
2 近畿大学
Kinki University
3 大阪大学
Osaka University

そこで本研究では詳細な顔情報が得られない歩行者画像を対象とし、性別推定を行う。

3. 検討内容

本研究では、低解像度の歩行者画像を対象とした性別推定を行う。実験1では、物体検出や一般物体認識などで広く使われている特徴量である、PCA, BOF, Histograms of Oriented Gradients (HOG)[13]の3種類の特徴量を共通の識別器を用いて性別推定を行い、結果を比較する。共通の識別器にはSVMを用いる。実験2では、画像のどの部分の特徴量が結果に影響しているか検証を行う。実験3では、画像中の背景が推定結果にどう影響を及ぼすか検証を行う。

3.1 特徴量

PCAとは、多変量で表されるデータの統計から一次結合で表現される新たな変量を構成し、互いに無相関な主成分に要約する手法である。画像に適用すると、その画像がどのような主成分で構成されているかを特徴量とする。

BOFとは、分類手法であるBag Of Words (BOW)を画像に適用した手法であり、BOWで文書を単語の集合とみなし、単語の語順を無視し、その頻度で文書の分類を行うのと同様に、画像を局所特徴量の集合とみなし、その位置情報を無視して画像をヒストグラムで表した特徴量である。今回は局所特徴量として、Scale Invariant Feature Transform (SIFT)[14]を用いる。SIFTとは、スケール空間を用いた照明変化や回転、拡大縮小に不変で頑強な特徴量である。

HOGとは、2005年にDalalらが発表した、局所領域の輝度の勾配方向をヒストグラムで表した特徴量である。幾何学的変換に強く、照明変動にも頑健である。まず輝度の勾配強度と方向の算出を行い、ヒストグラムを作成し、ブロック領域による正規化を行うことで算出される。輝度の勾配強度と勾配方向はそれぞれ式3.1, 3.2で算出される。

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad 3.1$$

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \quad 3.2$$

ヒストグラムは 0° から 180° までを 20° ずつ、計9方向に分割し作成される。ブロック領域の正規化は式3.3で算出される。ある n 番目のHOG特徴量について正規化を行う。分母は、1ブロック($q \times q$)に含まれるHOG特徴量の総和を表している。また N は勾配方向数である。

$$v(n) = \frac{v(n)}{\sqrt{(\sum_{k=1}^{q \times q \times N} v(k)^2) + 1}} \quad 3.3$$

3.2 識別器

SVMとは、1963年にVapnikが発表した、線形しきい素子を用いて2クラスのパターン識別器を構成する識別手法であり、1995年にカーネルトリックを用いることで非線形の識別手法へと拡張された識別手法である。カーネルトリックとは実際の空間での計算を避けて、カーネルの計算のみで最適な識別関数を構成する手法のことである。今回は、対象の画像から抽出された特徴量が線形分離できないデータであることを想定して、カーネルトリックを用いた非線形SVMを用いる。カーネルにはRadial Basis Function (RBF)を用いる。

4. 実験1

4.1 内容

本実験では、2013年度青山学院大学相模原祭の様子を撮影したデータを用いる。データから抽出された男女各50人、計100人を対象とする。データの例を図4-1に示す。画像サイズは 90×240 である。データ数の少なさを補うために交差検定法を用いてそれぞれの推定精度を比較し、平均した値を評価値とする。交差検定は、男女各50枚に対して、学習に各45枚、検定に各5枚を用いて、計10回の実験を行い、推定精度を算出する。

各特徴量の次元数は、PCAが90、BOFが50、HOGが2700である。PCAとBOFは手で設定した数値であり、HOGは、正規化回数と1ブロックのセル数と勾配方向数の乗算で求められるので、1セルを 15×15 ピクセル、1ブロックを 2×2 ブロックと設定し、正規化回数が75回、勾配方向数が9なので2700となる。



図 4-1 対象画像例

4.2 結果

実験 1 の結果を図 4-2 に示す。

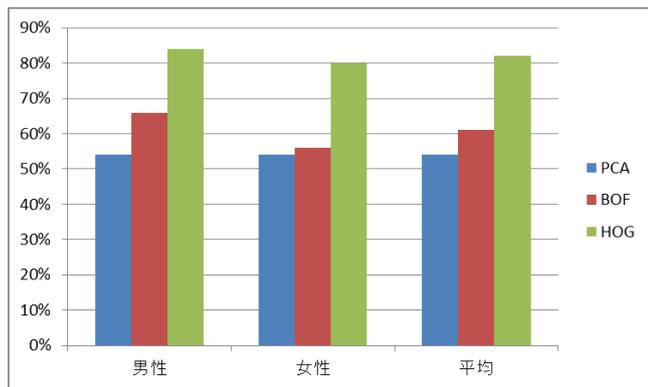


図 4-2 推定精度の比較



図 4-4 局所特徴抽出例 (女性)

4.3 考察

3 種類の特徴量の中で、HOG を用いた推定が最も良好な結果として得られた。PCA の精度が低くなってしまった原因として、ある主成分に対する特徴ベクトルの数値が、高ければ男性、低ければ女性、あるいは数値が高ければ女性、低ければ男性である、と明確に識別できる主成分が算出できていなかったためだと考えられる。BOF の精度が低くなってしまった原因として、図 4-3、図 4-4 のように、全身の服装や服の模様によって局所特徴量の抽出のされ方に差があり、性別ごとに分けられるような特徴ヒストグラムが算出されなかったためだと考えられる。



図 4-3 局所特徴抽出例 (男性)

HOG の結果が良好であった原因を検証するため SVM を使った時に生成される学習モデルの SV を 2 次元に次元数を削減することで可視化した。可視化画像を図 4-5 に示す。男性の SV は図の右、女性の SV は図の左に多く見られる。これらの SV が推定精度に影響したと考えられる。

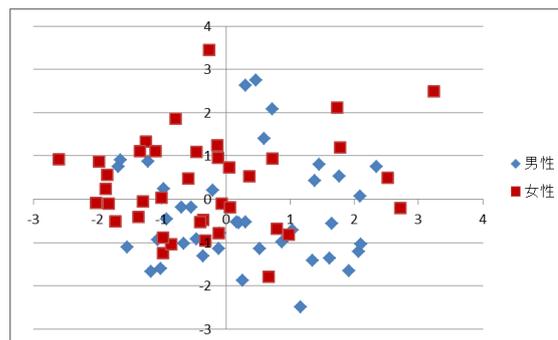


図 4-5 SV の可視化

5. 実験 2

5.1 内容

実験 1 より、HOG を用いた性別推定において良好な結果が得られた。そこで画像を上下 2 つに分割し、それぞれで性別推定を行い、どの部分の特徴量が結果に影響しているかを実験によって検証する。特徴量は HOG のみを用いる。画像や実験法は実験 1 と同様の条件で行う。

5.2 結果

実験 2 の結果を図 5-1 に示す。

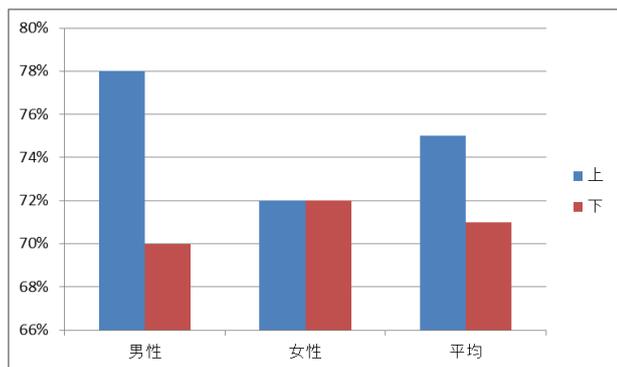


図 5-1 上下の比較

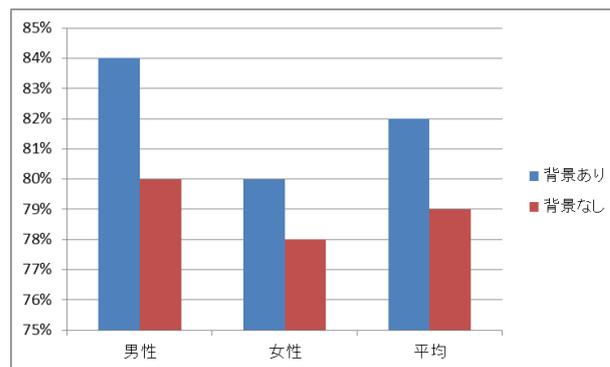


図 6-2 背景の有無による精度の比較

5.3 考察

下半分に比べ、上半分を対象とした時の結果の方が良好な結果が得られた。この結果から主に上半身から推定に用いられる特徴量が得られていると考えられる。

6. 実験 3

6.1 内容

本研究で用いている画像は同じ環境で得られたデータを用いているため、背景の影響が出てしまっている恐れがある。そこで背景を取り除いた画像を用意し、背景のある画像の結果と比較することで背景の影響を検証する。背景を取り除いた画像は手動で背景部分を取り除くことで準備する。画像例を図 6-1 に示す。また特徴量は HOG を用いる。



図 6-1 背景あり画像と背景なし画像

6.2 結果

実験 3 の結果を図 6-2 に示す。

6.3 考察

結果を比較したところ、精度に差はあまり見られなかった。この結果から HOG は背景の影響をあまり受けずに推定できていると考えられる。しかし背景を取り除くことで推定精度が下がってしまった。その原因として、背景の取り除き方に問題があったと考えられる。図 6-3 に背景を取り除くことで結果が変化してしまった画像例を示す。背景を取り除くことで人物の周りの勾配の方向が変化してしまい、得られる特徴量も変化してしまったと考えられる。



図 6-3 背景を取り除くことで失敗した画像例

7. 終わりに

本研究では、歩行者の全身を対象とした性別推定について、3 種類の実験を行い、結果に対して比較・検討を行った。実験では、PCA, BOF, HOG の 3 種類の特徴量と共通の識別器である SVM を用いた性別推定を行い、その結果を比較した。また、画像を上下に分割することで特徴量の結果への影響を検証し、主に上半分の特徴量が影響していることを確認した。さらに、背景のある画像と背景を取り除いた画像、それぞれで性別推定を行うことで背景の影響

を検証し、影響がほぼ無いことを確認した。

今回の実験を踏まえ今後は、今回用いた特徴量以外の性別をより顕著に表した特徴量の発見あるいは組み合わせの考案、また今回は識別器をSVMのみしか用いなかったが、複数の識別器を用いて結果を比較・検討、あるいはデータセットの充実、などを行うことで推定精度の向上を目指す。

謝辞 本研究の一部は、科学技術振興機構「JST」の戦略的創造研究推進事業「CREST」における研究領域「共生社会に向けた人間調和型情報技術の構築」の研究課題「歩容意図行動モデルに基づいた人物行動解析と心を写す情報環境の構築」の支援により行った。

参考文献

- [1] Baback Moghaddam, Ming-Hsuan Yang: "Gender Classification with Support Vector Machines", AFGR, pp.306-311, 2000.
- [2] Corinna Cortes, Vladimir Vapnik: "Support-Vector Networks Machine Learning", vol.20, pp.273-297, 1995.
- [3] Kazuya Ueki et al.: "A Method of Gender Classification by Integrating Facial, Hairstyle, and Clothing Images", ICPR, vol.4, pp.446-449, 2004.
- [4] J. H. Park, W. H. Cho, and S. Y. Park: "Color Image Segmentation Using a Gaussian Mixture Model and a Mean Field Annealing EM Algorithm", IEICE Trans Inf Syst, vol. E86-D, No. 10, pp.2240-2248, 2003.
- [5] 松濤智明, 山崎俊彦, 相澤清晴: "天井カメラ映像を用いた公共空間の人物属性解析", PRMU, pp. 135-140, 2011.
- [6] 沖田聡一郎, 藤尾光彦, 浅野晃: "フーリエ記述子の離散化方法の比較とテクスチャ解析への応用", PRMU, 102(708), pp. 67-72, 2003.
- [7] J. McQueen, "Some methods for classification and analysis of multivariate observations", In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp.281-297, 1967.
- [8] Yoav Freund, Robert E. Schapire: "A Short Introduction to Boosting", JSAI, 14(5):771-780, 1999.
- [9] L. Cao, M. Dikmen, Y. Fu, and T. S. Huang: "Gender recognition from body", ACM, pp.725-728, 2008.
- [10] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski and Cédric Bray: "Visual Categorization with Bags of Keypoints", ECCV, 2004.
- [11] L. Breiman: "Random Forests, Machine Learning", vol. 45, pp. 5-32, 2001.
- [12] 岩竹 隆志, 波部 斉: "多方向人物画像での性別推定のための効率的な学習", IEICE, ISS-SP-246, 2015.
- [13] Navneet Dalal and Bill Triggs: "Histograms of Oriented Gradients for Human Detection", CVPR, vol.1, pp.886-893, 2005.

- [14] David G. Lowe: "Object Recognition from Local Scale-Invariant Features", ICCV, vol. 2, pp1150-1157, 1999.