

パーツ間の関係性を考慮した歩行者属性分類

Pedestrian Attribute Classification Considering Relations between Body Parts

廣兼 優里^{*1} 長尾 智晴^{*2}
Yuri HIROKANE Tomoharu NAGAO

^{*1*2} 横浜国立大学 大学院環境情報学府

Graduate School of Environment and Information Science, Yokohama National University

Recently, driver assistances with vehicle camera have been actively conducted. As examples of the driving environment recognition by the vehicle camera, runway detection, pedestrian detection, and vehicle detection are included. In particular, many pedestrians have died in traffic accident. Therefore, the practical application of driver assistance technology using pedestrian detection and the further development are expected. However, at present, the characteristics of the detected pedestrian is not considered. Future, to show what attributes are in not only pedestrians, but also an object detected from the vehicle camera become increasingly necessary. In this report, we create a data set cut out from Google Maps image and perform attribute classification of person. At that time, we use the parts to correspond to different orientations and shielding of the person, and recognize by Random Forest.

1. はじめに

近年、車載カメラを用いた運転支援技術のための研究が盛んに行われている。走路検出、歩行者検出、車両検出などが車載カメラによる走行環境認識の例として挙げられる[石田 14]。特に、交通事故による死者の割合は歩行者が最も高いことから歩行者検出を用いた運転支援技術は今後実用化が進み、さらなる発展が見込まれる。しかし現状では、検出した歩行者に関する性質（以下、歩行者属性と定義）は考慮されていない場合が多い。今後、歩行者だけでなく車載カメラから検出した物体がどのような形状・性質などの属性をもつかを示すことがますます必要になると考えられる。

本論文では、車載カメラを想定して Google ストリートビュー [GoogleMap] から切り出した画像からデータセットを作成し、人物の属性分類を行う。その際、人物の様々な向きや遮蔽に対応するために、人物の頭、手、足といったパーツを用い、ノイズに強く高速で、パーツごとのカテゴリ分布を利用することができる Random Forest[Liaw 02] によって認識を行う。

従来の人物属性分類に関する研究は、顔画像を用いるものと全身画像を用いるものの2つに分けられる。顔画像を用いたものとして、1990年代から、顔の部分的な構造の特徴を用いた幾何学的特徴、顔画像全体の濃淡パターンを用いた見え方を利用してニューラルネットワークや Support Vector Machine(SVM)を用いて性別を分類する研究が行われている。また、Haar 特徴による AdaBoost 分類などによって性別や人種といった属性の分類も行われており [Brunelli 93]、近年では表情の認識 [Kumar 09] などが行われている。しかし顔画像を用いる方法は画像取得の点で限界があり、車載カメラで認識するには向き、大きさが多様であることから困難である。

そこで、全身画像を用いた属性分類として、歩容特徴を解析する手法 [Li 08]、人物のシルエット形状特徴による男女識別 [数藤 00] などが挙げられる。しかし、これらは撮影方向が限定的で、男女・年齢以外の属性の検討がなされていないことが課題として挙げられる。また、属性認識においては、オブジェクトを含む矩形を事前に検出すること、画像全体から特徴量を

抽出することが一般的であるが、人物属性認識の場合、人物画像全体が属性に応じた一定の形をしていないことから人物のポーズに対して適応的に特徴量を抽出することが必要となる。

人物属性認識にパーツを用いた従来研究として、Poselets[Bourdev 09] によって人物を含む矩形領域をパーツに分割し、段階的に SVM で人物の性別、T-シャツの有無、長髪かといった9つの2値分類を行うもの [Bourdev 11]、Deformable Parts Model[Felzenszwalb 10] とポーズ推定を利用して、SVM と Conditional Random Fields(CRF) によって26種類の服装・性別の属性を認識するもの [Chen 12] が挙げられる。これらは服装や性別以外の属性が検討されていないこと、属性が車載カメラでの認識に適していないこと、属性ごとのパーツ間の関係性とパーツごとの認識の統合方法が曖昧であることが課題として挙げられる。

そこで本論文では、車載カメラからの人物認識で必要となる『性別』『前方不注意』の属性をパーツを用いて分類する手法を提案する。

2. Random Forest

本研究で用いる機械学習の中の一手法である Random Forest について述べる。

Random Forest は決定木を用いたマルチクラス分類器である。木構造のフレームワークであるため高速な学習と分類が可能であること、ランダム学習であるため学習サンプルのノイズに対して強いことが特徴である。

2.1 分類器の学習

Random Forest は複数の決定木の集合で構成されている。各決定木には、入力されたサンプルが左右どちらの木に分類されるかを決める分岐ノードと、最終的にどのクラスのサンプルがどれだけの割合で着いたかをヒストグラムで表現するクラス分布をもつ終端ノードで構成される。

決定木の構築は入力データを分岐ノードによって分岐することで行う。分岐の評価には、左右に分岐した各クラスのサンプルとデータサンプルに付与されている教師データを用いて算出した情報エントロピーによる情報利得 ΔE を用いる。ランダムにパラメータを決定することを繰り返し、 ΔE が最大とな

連絡先: 廣兼 優里, 横浜国立大学大学院環境情報学府,
hirokane-yuri-yr@ynu.jp

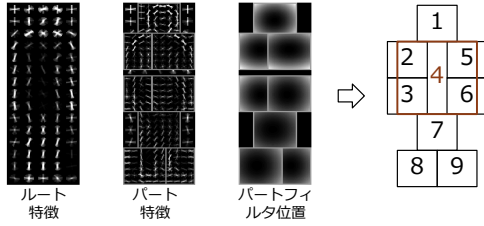


図 1: 人物画像のパーツの分割

る特徴量としきい値の組み合わせを分岐ノードのパラメータとして決定する。終端ノード l には分岐したデータ I_n の教師信号によってクラス分布が作成される。クラス分布は $|I|$ を全クラスのサンプル数、 $|I_{c_i}|$ をクラス c_i のサンプル数として式 (1) で示される。

$$p(c_i|l) = \frac{|I_{c_i}|}{|I|}. \quad (1)$$

2.2 分類

分類では、式 (2) で示される終端ノード l がもつクラス分布の平均 $P(c|L)$ を利用する。

$$P(c_i|L) = \frac{1}{T} \sum_{t=1}^T p_t(c_i|l_t), \quad (2)$$

T は決定木の数、 $p_t(c_i|l_t)$ は t 番目の決定木のクラス分布、 L は到達した終端ノード ($l_1, l_2, \dots, l_T$) である。そして式 (3) を用いてクラス推定を行う。 C_i^* は出力クラスを指す。

$$C_i^* = \arg \min_{c_i} P(c_i|L). \quad (3)$$

3. パーツ間の関係性を考慮した歩行者属性分類

3.1 パーツの分割

本手法では、Deformable Part Model[Felzenszwalb 10] を用いて図 1 のようにパーツごとに分割する。Deformable Part Model は、全身を捉える検出ウィンドウをルートフィルタとパーツを捉えるパーツフィルタを用いて、物体のモデルをパーツの集合として表現し、それぞれのパーツの妥当性と相対位置関係で評価を行う。具体的には、事前に定められたパーツフィルタとルートフィルタの見えに対するスコアと、パーツフィルタの位置に対するスコアの和が最大となるように学習する。このモデルは、姿勢の大きな変化に対応することができるため人物検出や物体認識で用いられている。

本手法では、Random Forest の学習および分類の前処理として入力画像を 9 つのパーツに分割する。頭部、両肘下、両肘上、腹部、足とパーツごとに検出することで属性の注目領域ごとに分類器を構築することが可能である。

3.2 Random Forest による歩行者属性分類

3.2.1 Random Forest の学習

Random Forest の学習の処理の流れを図 2 に示す。 T は決定木の個数、 s をパーツ番号として s_a はクラス a のサンプル、 s_b はクラス b のサンプルである。分割されたパーツごとにクラス分類器を構築する。

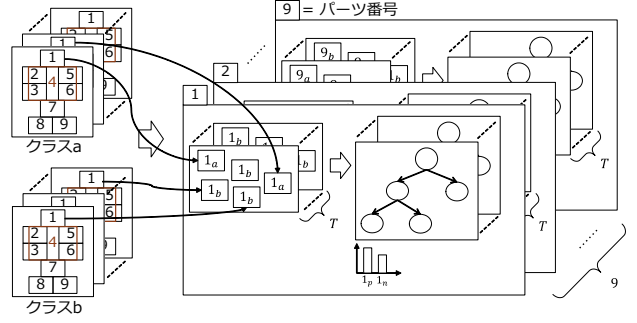


図 2: Random Forest の学習の処理の流れ

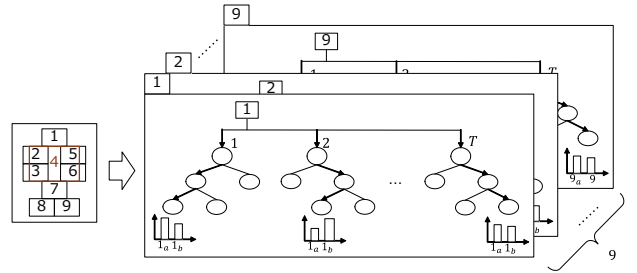


図 3: Random Forest による識別の処理の流れ

3.2.2 歩行者属性分類

分類の処理の流れを図 3 に示す。分割されたパーツごとに決定木に入力する。そして、各パーツの出力値は式 (4) で示すように到達した終端ノードのヒストグラムからクラスごとにカテゴリ分布を決定木の個数分足したものであり、出力クラス C_s は式 (5) である。 s はパーツ番号である。

$$\begin{cases} h_{s_a} = \sum_{t=1}^T p(c_a|l_a), \\ h_{s_b} = \sum_{t=1}^T p(c_b|l_b), \end{cases} \quad (4)$$

$$C_s = \begin{cases} a, & (h_{s_a} > h_{s_b}) \\ b, & (h_{s_a} \leq h_{s_b}) \end{cases} \quad (5)$$

全てのパーツの出力値を統合した最終出力クラス C は式 (6) で示される。

$$C = \begin{cases} a, & (H = \sum_{s=1}^9 (h_{s_a} - h_{s_b}) > 0) \\ b, & (H = \sum_{s=1}^9 (h_{s_a} - h_{s_b}) \leq 0) \end{cases} \quad (6)$$

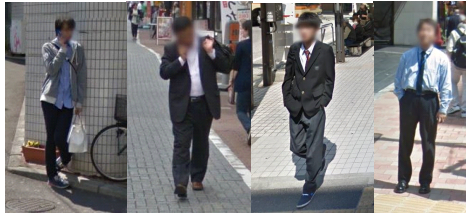
4. 歩行者属性分類実験と考察

4.1 人物属性データセット

車載のカメラや距離計測器などのセンサを用いて人物を検出したという前提の下、様々な向きや照明条件の人物全身画像を得ることができる Google マップから人物を切り出し、データセットを作成した。その際、人物の全身が入るよう矩形を切り出し、“性別 (男・女)”, “前方不注意 (有・無)”, “向き (表・裏・横)” の属性を付与した。図 4 に使用画像例を示す。

4.2 実験概要

本実験で分類する属性は『性別』『前方不注意』である。データセットの画像枚数を表 1 に示す。画像サイズは画像によって



性別	男性	男性	男性	男性
前方不注意	有り	有り	無し	無し
向き	表	表	表	表



性別	女性	女性	女性	女性
前方不注意	有り	有り	無し	無し
向き	横	横	表	表

図 4: データセットの画像例

異なる。また、使用した特徴量はパーツの矩形領域をさらに4×4分割した矩形領域ごとのエッジの傾きおよび強度、RGBヒストグラム、RGBごとの統計特徴量11種(平均階調値、最大階調値、最小階調値、標準偏差、最大・最小値幅、1 σ 内確率、3 σ 外確率、第1四分位数、第3四分位数、中央値、最頻値)である。Random Forestの学習パラメータを表2に示す。また、比較手法として各パーツごとにSVMを用いて属性分類を行い、全パーツの分類結果を多数決によって統合する手法を用いた。

表 1: 画像枚数

属性	学習 or テスト	属性値	枚数
性別	学習	男性	70 枚
		女性	71 枚
	テスト	男性	11 枚
		女性	11 枚
前方不注意	学習	あり	60 枚
		なし	57 枚
	テスト	あり	10 枚
		なし	10 枚

4.3 結果と考察

属性ごとの分類結果を表3、表4に示す。また、性別での分類例を図5、前方不注意での分類例を図6に示す。

比較手法と比べて提案手法では、統合後の分類率が良いことから、クラス分布を利用することで分類に有効なパーツとそうでないパーツを良好に考慮する事ができたと考えられる。また、提案手法ではパーツごとの学習画像の分類精度に差が少なく、そして未知画像の分類率が高いことが分かる。

パーツごとに見ると、性別では提案手法・比較手法ともにパーツ2～6の分類率が比較的高いことから、上半身の色情報

表 2: Random Forest のパラメータ

パラメータ	値
決定木の数	20
決定木の最大の深さ	20
特徴量のランダム選択回数	500
しきい値のランダム選択回数	10
各 Tree に与える教師信号のサブセットの割合	25%

表 3: 性別の分類結果

parts	SVM				Random Forest			
	統合前		統合後		統合前		統合後	
	学習	未知	学習	未知	学習	未知	学習	未知
1	0.83	0.35	0.82	0.5	0.82	0.50	0.80	0.63
2	1.0	0.40			0.82	0.59		
3	1.0	0.50			0.70	0.73		
4	0.72	0.65			0.80	0.68		
5	0.99	0.60			0.84	0.77		
6	0.72	0.70			0.78	0.64		
7	0.66	0.55			0.83	0.50		
8	1.0	0.45			0.82	0.68		
9	1.0	0.65			0.80	0.59		

表 4: 前方不注意の分類結果

parts	SVM				Random Forest			
	統合前		統合後		統合前		統合後	
	学習	未知	学習	未知	学習	未知	学習	未知
1	0.99	0.65	0.87	0.45	0.86	0.65	0.78	0.56
2	0.99	0.50			0.78	0.60		
3	0.72	0.45			0.72	0.60		
4	0.94	0.50			0.80	0.60		
5	0.75	0.65			0.80	0.50		
6	0.67	0.50			0.78	0.55		
7	0.75	0.35			0.78	0.55		
8	0.99	0.45			0.75	0.70		
9	0.78	0.60			0.79	0.60		

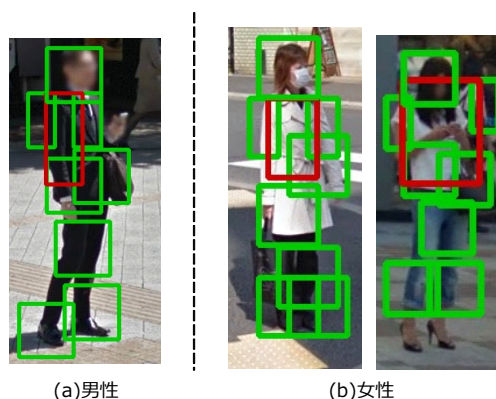


図 5: 性別での分類例

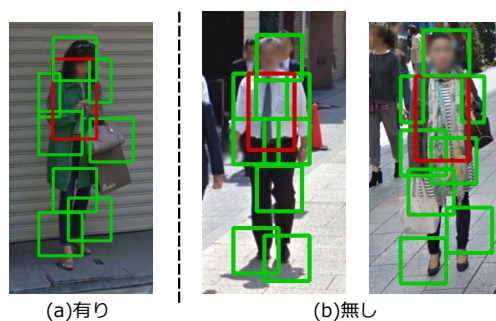


図 6: 前方不注意での分類例

が分類に重要であると言える。しかし、上半身の色情報が性別の分類全てにおいて有効ではないので、オブジェクトと背景を切り分け、髪の毛に相当する頭部の黒色の割合を特徴量に加えることや、体全体のエッジを利用することなどが必要となる。

一方、前方不注意では、図 6 からわかるように、無しの場合に頭部のエッジの傾きがあることが顕著に分かる。また、どちらかの腕を曲げている場合が多い。結果から、頭・腕・足のパーツでの認識率が高いことから、これらのパーツ内のエッジ強度を捉えて分類することができていると言える。しかし、前方不注意の画像データ中でも腕・足の特徴は多様であるため、腕・足のパーツを考慮した分類を行うといった、属性ごとに有効なパーツを組み合わせる工夫が今後必要になる。

5. まとめ

パーツ間の関係性を考慮するために、パーツごとのカテゴリ分布を用いた Random Forest を用いて『性別』『前方不注意』の歩行者属性分類手法を提案した。提案手法によって、パーツを用いた人物属性認識で一般的に用いられている SVM を用いた手法より、認識精度の向上、パーツごとのカテゴリ分布を用いた利点を確認することができた。今後は、Deformable Part Model の検出漏れを受容できるように画像枚数を増やすこと、そして識別器とパーツの組み合わせを最適化し、パーツを用いた人物・歩行者属性分類精度の向上を行うことを考えている。

参考文献

- [石田 14] 石田皓之: 車載カメラによる走行環境認識技術の現状, 2014 年電子情報通信学会総合大会 情報・システム講演論文集 2, DI-1-3, 2014.
- [GoogleMap] <https://www.google.co.jp/maps>, 2014.
- [Liaw 02] Liaw, A., and Wiener, M.: Classification and regression by randomForest, R news, Vol.2, No.3, pp.18-22.
- [Brunelli 93] Brunelli, R., and Poggio, T.: Face recognition: Features versus templates, IEEE transactions on pattern analysis and machine intelligence, Vol.15, No.10, pp.1042-1052, 1993.
- [Kumar 09] Kumar, N., Berg, A. C., Belhumeur, P. N., and Nayar, S. K.: Attribute and simile classifiers for face verification, Computer Vision, 2009 IEEE 12th International Conference on. IEEE, pp.365-372, 2009.
- [Li 08] Li, X., Maybank, S. J., Yan, S., Tao, D., and Xu, D.: Gait components and their application to gender recognition, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, Vol.38, No.2, pp.145-155, 2008.
- [数藤 00] 数藤恭子, 大和敦司, 伴野明, 石井健一郎: 入店客計数のためのシルエット・足音・足圧による男女識別法, 電子情報通信学会論文誌 D, Vol.83, No.8, pp.882-890, 2000.
- [Bourdev 09] Bourdev, L., and Malik, J.: Poselets: Body part detectors trained using 3d human pose annotations, Computer Vision, 2009 IEEE 12th International Conference on, pp.1365-1372, 2009.
- [Bourdev 11] Bourdev, L., Maji, S., and Malik, J.: Describing people: A poselet-based approach to attribute classification. Computer Vision, 2011 IEEE International Conference on, pp. 1543-1550, 2011.
- [Felzenszwalb 10] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D.: Object detection with discriminatively trained part-based models, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.32, No.9, pp.1627-1645, 2010.
- [Chen 12] Chen, H., Gallagher, A., and Girod, B.: Describing clothing by semantic attributes, Computer Vision - ECCV 2012, Springer Berlin Heidelberg, pp.609-623, 2012.