

歩行者の姿勢に注目した「歩きスマホ」認識に関する検討

○ 新村文郷†, 川西康友‡, 出口大輔††, 井手一郎‡, 村瀬洋‡, 藤吉弘亘‡‡

○ Fumito SHINMURA†, Yasutomo KAWANISHI‡, Daisuke DEGUCHI††,
Ichiro IDE‡, Hiroshi MURASE‡ and Hironobu FUJIYOSHI‡‡

†: 名古屋大学未来社会創造機構, shinmuraf@murase.m.is.nagoya-u.ac.jp

‡: 名古屋大学情報科学研究科, {kawanishi, ide, murase}@is.nagoya-u.ac.jp

††: 名古屋大学情報連携統括本部, ddeguchi@nagoya-u.ac.jp

‡‡: 中部大学工学部, hf@cs.chubu.ac.jp

<要約> 歩行者の「歩きスマホ」は周囲への注意が疎かになり、自動車の接近に気付にくくなることから、交通事故に遭う危険性が高い。そのため、車載カメラを用いた歩行者の「歩きスマホ」の認識は、安全運転支援に有効だと考えられる。そこで本発表では、歩行者の「歩きスマホ」を認識する手法を提案する。具体的には、「歩きスマホ」中に見られる特徴的な歩行者の姿勢に注目し、腕と頭の位置の事前知識を用いることで腕と頭の姿勢の共起性を特徴として利用する認識手法を提案する。そして、実験を通して提案手法の有効性を示す。

<キーワード> 歩きスマホ, ながら歩き, 人物属性認識, 車載カメラ画像

1. はじめに

近年、対歩行者事故の低減を目的とし、車載カメラを用いて自動車の周囲に存在する歩行者を認識する技術が広く研究されている。歩行者を認識して事故の危険性を予測することができれば、運転者への警告や自動的に回避や停止を行う支援が可能となり、より安全な運転の実現が可能だと考えられる。事故の危険を予測するためには、歩行者の状態を認識し、危険かどうかを推定することが重要となる。そのため、歩行者の細かな状態を表す属性を認識する技術が必要である。

認識すべき属性の一つに、歩行者が周囲に注意を払っている状態か否かが挙げられる。周囲に注意を払っていない状態では、歩行者は自動車の接近に気付かず、道路への飛び出し等の危険な行動を起こす可能性があると考えられる。近年のスマートフォンの普及により、歩行者がスマートフォンの画面に集中しながら歩く、いわゆる「歩きスマホ」が頻繁に見られるようになった。「歩きスマホ」中の歩行者は、視線の移動範囲が狭くなり、周囲への注意が疎かになる。そのため、自動車や歩行者が接近しても気付



図1 歩行者の「歩きスマホ」の例

かないことが多く、周囲へ注意を払っている歩行者に比べて事故に遭う危険性が高い。したがって、歩行者の「歩きスマホ」を認識することは、事故の危険予測に役立つと考えられる。

そこで本発表では、車載カメラ画像から検出された歩行者に対し、その歩行者が「歩きスマホ」をしているか否かを認識する技術の実現を目指す。

歩行者が「歩きスマホ」を行う場合、歩行者は手にスマートフォンを持ち、視線はスマートフォンの画面を向いた状態になる。図1に「歩きスマホ」をしている歩行者を示す。歩行者が肘を曲げ、顔を手

元に向けていることが確認できる。「歩きスマホ」中の歩行者の多くはこれと同様の姿勢になる。そのため、肘を曲げていることと、顔が手元を向いていることが同時に起こることが「歩きスマホ」の大きな特徴となる。そこで、歩行者の腕と頭の姿勢の関係性に注目し、両方の姿勢を同時に捉える特徴を用いることで「歩きスマホ」の認識を実現する。

2. 関連研究

これまでに、歩行者の「歩きスマホ」を認識する研究がいくつか報告されている。皆本ら[1]は、スマートフォンを持つ腕の形状に着目し、体の上半身から **Histograms of Oriented Gradients (HOG)** 特徴量[2]を求めて認識に利用している。我々のグループにおいても、腕や頭を含む体全体から求めた HOG 特徴量を「歩きスマホ」認識に用いる手法を提案している[3]。これらの手法は、画像から特徴を抽出し、統計的学習手法により識別器を構築することで、「歩きスマホ」の認識を行う手法である。

しかし、これらの手法は「歩きスマホ」の歩行者と似た姿勢をとる歩行者が誤認識されやすいという問題がある。皆本らの手法では上半身の形状を、我々のグループの手法では体全体の形状を特徴として利用しているため、「歩きスマホ」の歩行者と似たシルエットの歩行者では誤認識が発生する。例えば、肘を曲げただけの歩行者を「歩きスマホ」中の歩行者と誤認識する可能性がある。歩行者は「歩きスマホ」をしていなくても、荷物を持つ行為や腕を振る行為の際に肘を曲げた姿勢をとるため、誤認識が多く発生すると考えられる。

これらの誤認識を防止するためには、歩行者の腕と頭の姿勢に同時に注目する必要がある。腕と頭の両方の姿勢を同時に捉える特徴を用いることが有効と考えられる。すなわち、腕や頭の形状とそれらの共起を捉える特徴が「歩きスマホ」の認識に有効と考えられる。

特徴量の共起性を利用する手法として、尾崎らは複数の HOG 特徴量を組み合わせた **Joint HOG** 特徴量を提案している[4]。Joint HOG 特徴量は、画像全体から位置の異なる 2 個の局所領域を選択し、領域内の HOG 特徴量を組み合わせることで、単一の HOG 特徴量のみでは捉えられない形状の対称性や連続性といった特徴を捉えることが可能な手法である。

Joint HOG 特徴量を用いた認識では、組み合わせ



(a) 腕を曲げた歩行者 (b) 下を向いた歩行者
図 2 「歩きスマホ」と誤認識しやすい例

る 2 個の局所領域の位置は、AdaBoost により自動的に選択される。そのため、2 個の局所領域が必ずしも腕と頭の部分から選択されるとは限らない。例えば、腕の連続的な形状を捉えるために、腕の周囲の局所領域の組み合わせが選択されることもある。Joint HOG 特徴量の中から AdaBoost により認識に有効な特徴が選択されるため、認識に有効と考えられる腕と頭の部分の局所領域の組み合わせが自動的に選択されることが期待される。しかし、特徴の選択は学習によって行われるため、学習に使用したデータが選択に大きく影響する。このため、様々な歩行者の見えのバリエーションを含む大量のデータが学習に必要となる。そのような大量のデータの収集は非常に困難であるため、少ないデータからでも真に有効な特徴を選択する手法が必要である。

3. 提案手法

車載カメラ画像から検出された歩行者に対し、腕と頭の姿勢を捉える特徴の抽出を行い、歩行者が「歩きスマホ」をしているか否かを認識する識別器を構築する。図 3 に処理の流れを示す。

「歩きスマホ」中の歩行者の姿勢の特徴として、以下が挙げられる。

- ① 肘を曲げる
- ② 顔を手元に向ける

これらの姿勢の特徴に注目することで、「歩きスマホ」の認識を行う。

「歩きスマホ」と似た姿勢の歩行者として、肘を曲げただけの歩行者や、顔を下に向けただけの歩行者が挙げられる。図 2 にそれらの例を示す。これらの歩行者と「歩きスマホ」中の歩行者を識別するために、腕と頭の姿勢変化が同時に起きていることを捉える必要がある。そこで本手法は、「歩きスマホ」

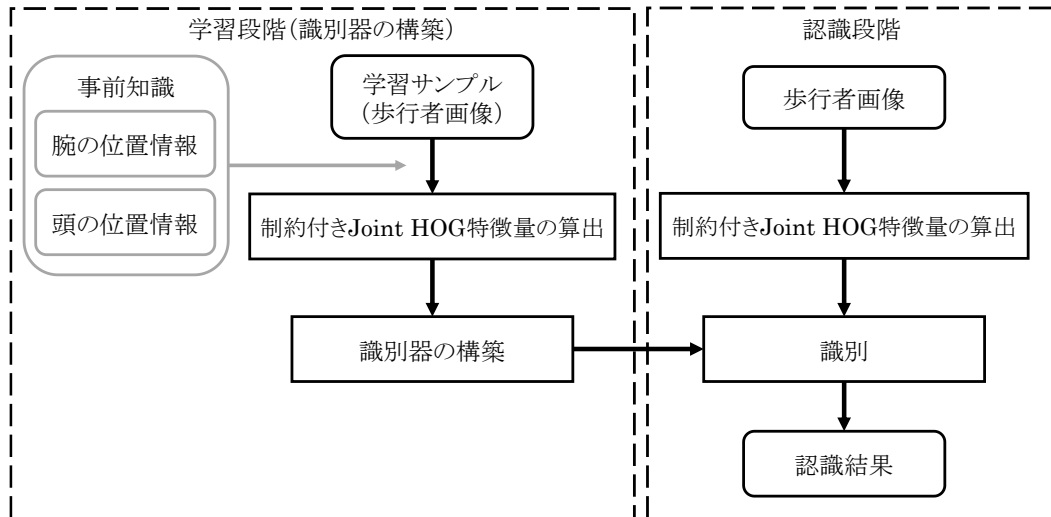


図3 「歩きスマホ」認識の処理の流れ

中の腕の姿勢と頭の姿勢の共起性を利用する。

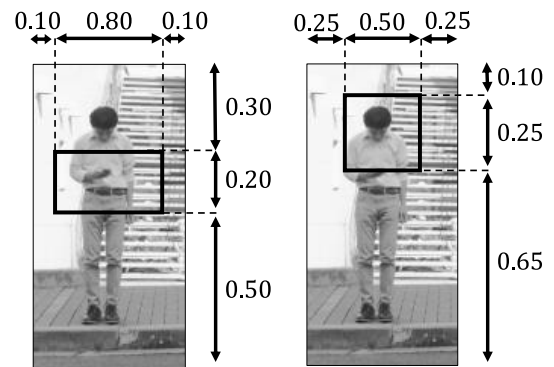
腕と頭の姿勢の共起性を捉えるため、Joint HOG 特徴量を用いた認識手法[4]を利用する。少量のデータを用いた学習では AdaBoost により認識に有効な特徴を選択することが難しいため、事前知識を用いて特徴を選択する位置に制約を付けた制約付き Joint HOG 特徴量を提案する。本手法では、事前に腕と頭の位置を与えることで、腕と頭の共起性を利用した認識を行う。腕の周辺と頭の周辺から局所領域を選択し、各領域内の HOG 特徴量を組み合わせた特徴を用いて認識を行う。これにより、腕と頭の形状とそれらの共起性を捉えた特徴を抽出する。これにより、腕や頭の片方の形状に依存した誤認識を抑えた「歩きスマホ」の認識を実現する。

3.1 HOG 特徴量の算出

腕や頭の形状を捉えるため、局所形状特徴量として Dalal らにより提案された HOG 特徴量[2]を利用する。HOG 特徴量は画像の局所領域における輝度勾配をヒストグラム化した特徴であり、物体の形状を表現することができるため、腕や頭の形状を捉える特徴量として適している。

まず、画像内の各局所領域から HOG 特徴量を算出する。本手法は、腕と頭の部分の形状特徴を利用するため、HOG 特徴量を算出する局所領域は、あらかじめ定めた腕の位置と頭の位置の周辺から行う。本手法で定めた腕周辺の局所領域と頭周辺の局所領域を図4に示す。

歩行者画像を 200×100 画素の大きさにリサイズした後、図3に示す各局所領域をセル (10×10 画素) に分割し、9方向の輝度の勾配方向ヒストグラムを作成する。これを局所領域全体で正規化し、



(a) 腕周辺の局所領域 (b) 頭周辺の局所領域
図4 各局所領域の定義

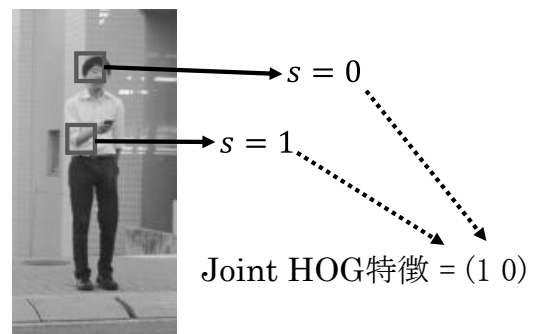


図5 Joint HOG 特徴

HOG 特徴量を算出する。

3.2 制約付き Joint HOG 特徴量の算出

腕部の HOG 特徴量と頭部の HOG 特徴量を用いて、Joint HOG 特徴量[4]を算出する。

まず、腕部の HOG 特徴量からセルを1つ選択する。選択されたセルの勾配ヒストグラムの中から、最も識別に有効な勾配方向を選択する。選択された

特徴量を利用して2値化符号を得る。次に、腕部と同様に、頭部の HOG 特徴量からセルを1つ選択し、勾配ヒストグラムの中から勾配方向を選択し、同様に2値化符号を得る。これにより得られた2個の2値化符号を組み合わせることで、2つのセルの組み合わせにおける Joint HOG 特徴量を算出する(図5)。

これをすべてのセルの組み合わせについて行うため、2つのセルの組み合わせの数の Joint HOG 特徴量が算出される。本手法では、腕周辺の局所領域内のセル数と頭周辺の局所領域内のセル数の積だけ識別器が構築される。

3.3 識別器の構築

算出した Joint HOG 特徴量の中から、識別に有効な特徴量を AdaBoost の学習により選択する。これにより、すべてのセルの組み合わせの中から、どの位置関係にあるセルの組み合わせが識別に有効かが学習されることになる。

このようにして構築した識別器では、「歩きスマホ」の識別に有効なセルの組み合わせが選択され、その Joint HOG 特徴量の中から認識に有効な特徴量を用いて識別が行われる。

4. 実験

提案手法の有効性を確認するため、歩行者画像に対し提案手法を適用し、「歩きスマホ」認識の性能評価を行った。

4.1 データセット

実験で使用するデータセットとして、市販のカメラ(PointGrey社製 Grasshopper3)を用いて歩行者を撮影し、歩行者画像を収集した。11人の歩行者それぞれについて、屋外の異なる場所にて、以下の4姿勢の画像を各90枚ずつ収集した。

- ① 「歩きスマホ」中の姿勢
- ② 前を向いた姿勢
- ③ 肘を曲げた姿勢
- ④ 下を向いた姿勢

これを、歩行者の正面と横(左右)の計3方向から撮影し、合計で11,880(=90枚×11人×4姿勢×3方向)枚の歩行者画像を用意した。なお、「歩きスマホ」中はすべての歩行者が右手にスマートフォンを持った。データセットの例を図6に示す。



(a) 「歩きスマホ」中の歩行者



(b) 前を向いた歩行者 (c) 肘を曲げた歩行者



(d) 下を向いた歩行者

図6 データセットの例

4.2 実験方法

用意したデータセットを使用して「歩きスマホ」認識を行い、精度を評価した。データセット中の11人の歩行者画像のうち、10人分のデータを用いて学習を行い、残り1人分のデータを用いて精度を評価した。これを11回繰り返し、11人のデータが1回ずつ評価に用いられるよう実験を行った。

精度評価では、最大F値を評価指標とし、11回の平均を評価に使用した。比較手法として、HOGとSVMを組み合わせた認識手法[3]、腕や頭といった領域の制限をせず Joint HOG 特徴量と2段階 AdaBoostにより学習を行う手法[4]を使用した。

4.3 実験結果

表1に評価実験の結果を示す。比較手法1は Joint HOG 特徴量と2段階 AdaBoost を用いた認識手法 (Joint HOG)、比較手法2は HOG と SVM を組み合わせた認識手法 (HOG+SVM) である。表の結果から確認できるように、2個の比較手法に比べて、提案手法の認識性能が優れていることが確認できる。

表 1 実験結果

	提案手法 (制約付き Joint HOG)	比較手法 1 (Joint HOG)	比較手法 2 (HOG+SVM)
最大 F 値	0.74	0.64	0.70

図 7 に提案手法や比較手法の精度を表した ROC 曲線を示す。横軸が誤検出率 (False Positive Rate), 縦軸が検出率 (Recall) を表している。なお, 図の ROC 曲線は, 11 回の評価実験のうちのある 1 回の結果である。

4.4 考察

比較手法 2 (HOG+SVM を用いた認識手法) よりも, 共起特徴を利用した比較手法 1 (Joint HOG 特徴量と AdaBoost を用いた認識手法) の精度が低い。これは, 学習に使用した歩行者画像が 10 人分と少ないため, 「歩きスマホ」の認識に有効な特徴を適切に選択できなかったためと考えられる。これに対し, 提案手法は事前知識によって共起特徴を抽出する位置を与えたことで, 認識に有効な特徴を利用することに成功し, 2 種類の比較手法よりも優れた認識性能が得られたと考えられる。

また, 図 7 の ROC 曲線を見ると, 誤検出率が低い部分において提案手法が高い性能を示している。これは, 比較手法では肘を曲げた歩行者や下を向いた歩行者を「歩きスマホ」歩行者と誤認識するため性能が低くなっているのに対し, 提案手法は腕と頭の姿勢の共起性を積極的に捉えた特徴量を利用することで, これらの歩行者を正しく認識できるようになったためと考えられる。

5. むすび

本発表では, 歩行者の姿勢を考慮した「歩きスマホ」認識手法を提案した。「歩きスマホ」の場合, 腕と頭の姿勢に大きな特徴が表れるため, これらの共起を捉える特徴を利用することで, 識別性能の向上を行った。

「歩きスマホ」中の歩行者は周囲への注意が不足するため, 衝突事故の危険がある。そのため, 本手法により危険のある歩行者を認識して運転者に注意を促すといった応用が可能となり, 運転時の安全性を高めることができると考えられる。

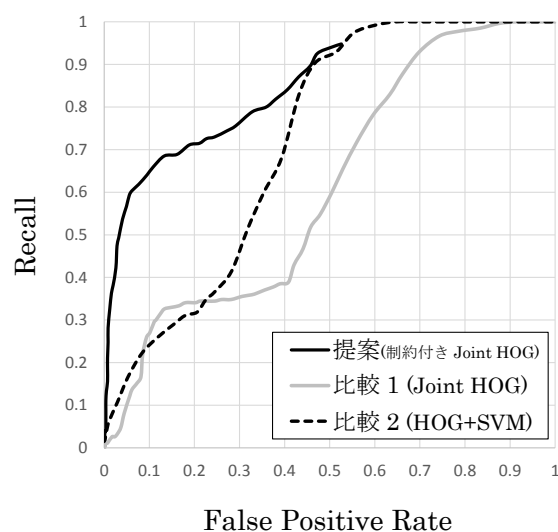


図 7 各手法の性能比較

今後の課題として, 歩行者の向きを考慮した認識手法の改善が挙げられる。歩行者の体の向きにより腕の位置や頭のシルエットが変化するため, 抽出される特徴量も変化し「歩きスマホ」を認識できない場合がある。この問題を解決するため, 歩行者の向きに合わせた特徴量の選択方法について検討を行う予定である。

謝辞

本研究の一部は国立研究開発法人科学技術振興機構 (JST) の, 研究成果展開事業「センター・オブ・イノベーションプログラム (名古屋 COI: 高齢者が元気になるモビリティ社会)」の支援及び, 科学研究費補助金の援助によって行われた。

参考文献

- [1] 皆本光, 佐野睦夫: “歩行者の不注意行動認識—歩きスマホ検出—”, 第 6 回視覚・聴覚支援システム (VHIS) 研究会予稿集 (2015)
- [2] N. Dalal, B. Triggs: “Histograms of Oriented Gradients for Human Detection”, In Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.886-893 (2005)
- [3] 新村文郷, 川西康友, 出口大輔, 井手一郎, 村瀬洋, 藤吉弘亘: “車載カメラ画像からの「スマホ歩き」認識に基づく歩行者の不注意度推定”, 信学技報, pp.83-88 (2015)

- [4] 尾崎貴洋, 山内悠嗣, 藤吉弘亘: “Joint HOG 特徴を用いた 2 段階 AdaBoost による車両検出”, DIA2008 講演論文集, I1-13 (2008)