

複数の候補領域を用いた歩行者検知システムの検討 A study on pedestrian detection method using multiple regions

高橋 圭[†] 藤谷 脩平[‡] 猿田和樹[†] 寺田裕樹[†] 武田和時[†]
Kei Takahashi Syuhei Fujiya Kazuki Saruta Yuki Terata Kazutoki Takeda

1. はじめに

近年、車載カメラ映像から歩行者を検知し、運転者の注意を喚起するシステムが実用段階となっている[1]。しかし、歩行者の背景の影響が大きいこと、歩行者の大きさや姿勢などが多種多様であること、歩行者の動きの予測が困難であることなど、他の特定物体の認識と比べた場合の課題が多いのが現状である。特に、歩行者候補領域を決定し、その候補領域に対して歩行者か否かの判定を行う場合、切り出される歩行者候補領域内における歩行者の位置ずれが検知精度に与える影響は大きい。本研究では、位置ずれが検知結果へ及ぼす影響を明らかにし、複数の歩行者候補領域を利用した検知手法により、検知精度の向上を目指す。

2. 歩行者検知システムの概要

本研究では車載カメラからの近赤外線映像に対する歩行者検知システムを用いる。車内に設置した近赤外線投光器により前方に近赤外光を照射し、反射光を近赤外線領域に対応したカメラで撮影する。

撮影したカメラ映像を、0.1秒毎に640×480ピクセルのフレーム画像として読み込み、歩行者の候補領域を抽出する。フレーム画像入力から歩行者の判別までの流れを図1に示す。まず、フレーム画像内で歩行者候補領域となる領域を高速物体検知手法により抽出し、候補領域に対し歩行者かどうかをSVMにより判別する[2]。SVMの特徴量には288次元のHOG特徴を用いる[3]。なお、SVMの実装にはIntel社の画像処理ライブラリOpenCVを用いる[4]。



図1 歩行者検知までの流れ

3. 歩行者検知実験

3.1 実験データ

学習用データには16×32ピクセルのグレースケール画像を用いる。ポジティブサンプルとして歩行者画像4940枚、ネガティブサンプルとして歩行者以外の画像（電柱、標識、背景のみ）20000枚を使用する。評価用データは動画画像から0.1秒毎に取り出した11401フレームの静止画（640×480ピクセル）である。また、歩行者候補領域の拡大縮小には、最近隣接補間法を用いる。

3.2 予備実験

はじめに、評価用データから切り出した歩行者候補領域に対し検知実験を行い、検知、未検知（歩行者を「歩行者以外」と判別）、誤検知（歩行者以外を「歩行者」と判別）、棄却の4通りの結果を求める。これらのデータを、画像内での歩行者の位置、輝度値から求めた勾配強度、重心の位置について分析して、歩行者領域の位置ずれがHOG特徴に及ぼす影響を明らかにする。予備実験の検知結果を表1に示す。

表1 予備実験の結果

	検知	未検知	誤検知	棄却
データ数	5863	3576	1735	27437

予備実験での検知画像の例、未検知画像の例を図2、図3にそれぞれ示す。図2、図3より検知の場合は歩行者が画像のほぼ中心に位置するが、未検知の場合は歩行者の位置が画像の中心から若干ずれていることが分かる。

次に、各画像の平均の勾配強度を求め、傾向を分析する。検知画像と勾配強度のグラフの対応を図4に示す。大まかな範囲として、1~36次元が歩行者の頭部、37~72次元が上半身上部、73~180次元が胴体部分、181~252次元が脚部、253~288次元が足元部分の勾配強度を表している。歩行者が正確に捉えられている検知画像の多くは、60次元付近と200次元付近、つまり肩付近と脚部付近でピーク値となっていることがわかった。また、検知データの多くは、歩行者の位置のずれの影響によりピーク値となる箇所もずれるケースがあり、歩行者としての特徴が失われ未検知と判別されている可能性があるといえる。

以上のことから、歩行者候補領域の切り出し位置をずらして歩行者を画像の中心に表示させることによって、HOG特徴の勾配強度及び検知結果が変化すると考えられる。次節では候補領域の切り出し位置をずらした場合について実験を行う。

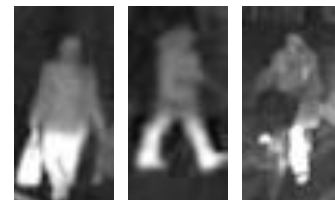


図2 検知画像の例

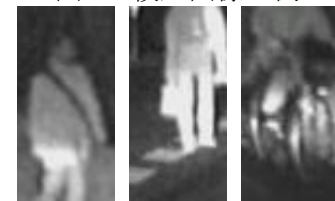


図3 未検知画像の例

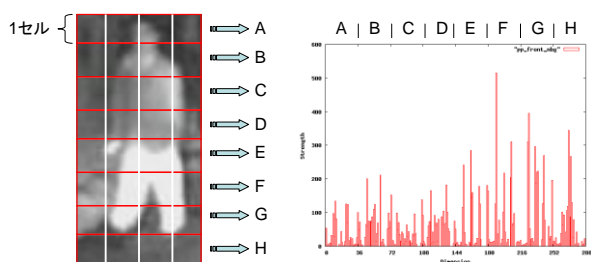


図4 画像と勾配強度グラフの対応

3.3 候補領域の位置ずらしに関する実験

3.3.1 実験概要

ここでは、フレーム画像からの歩行者候補領域の切り出し位置を、上下左右に1~2ピクセルずらして、それぞれの位置で検知結果の変化について考察する。図5に切り出し位置のずらし方向を示す。候補領域を①~⑧の方向に1ピクセルずらした後、歩行者かどうかを判定する。候補領域を1ピクセルずらす場合は元の位置も合わせると、1つの候補領域に対して9個の検知結果が出力され、2ピクセルずらす場合は25個の検知結果が出力される。

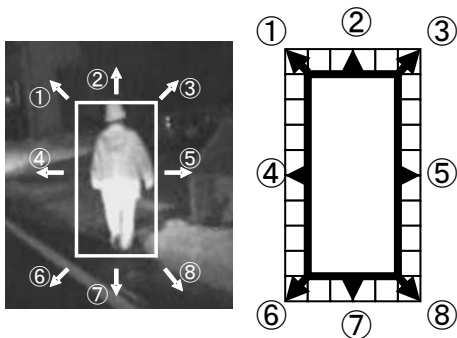


図5 切り出し位置のずらし方向

3.3.2 実験結果

表2に、候補領域を1ピクセルずらした場合の①~⑧各位置で検知できた件数を示す。例えば、②の方向に候補領域を1ピクセルずらした場合、判定結果が検知となった画像データが6146個ある、ということを表している。

表2 切り出し位置ずらしでの各位置の検知件数

①	②	③
5602	6146	6295
④	デフォルト位置	⑤
5197	5863	5993
⑥	⑦	⑧
4623	5294	5391

表2から、切り出し位置を1ピクセルずらすだけで検知件数が大きく異なることがわかる。③の位置で検知件数は最大となるが、誤検知数も大きく増加することが判明した。2ピクセルずらした場合でも同じような傾向が見られた。また、画像により検知できる位置はさまざまであり、検知できる位置は複数存在することが確認できた。よって複数の切り出し位置での判定結果を統合できれば、検知精度の向上が期待できる。

3.4 考察

ここでは、複数の候補領域の切り出し位置に対し、N箇所以上で歩行者と判断されれば、最終的な判別結果を“歩行者”とするというように、歩行者と判別された位置の総数がしきい値N以上であれば歩行者と判定する。切り出し位置を1ピクセルずらし、N=3~5に設定した場合の実験結果を表3に示す。ここで、表中の()内は、従来手法の結果と比較したときの増減数である。

表3 しきい値と検知・誤検知の件数

	N=3	N=4	N=5
検知	6634(+771)	6176(+313)	5691(-172)
誤検知	2199(+464)	1562(-173)	1126(-609)

表3から、N=3の場合は検知数が大きく増加するが、同時に誤検知数も増加しており、検知精度が向上したとはいえない。また、N=5の場合は誤検知の改善はできているが、検知数が減少しているため適していない。しかし、N=4では検知数が増加、なおかつ誤検知が減少しており、従来手法と比べると検知精度が向上していることが明らかとなった。歩行者と認識位置の総数をしきい値に設定する場合、1ピクセルずらした時はN=4が妥当といえる。同様に2ピクセルずらした時はN=9であることが判明した。

4. 結論

本研究により、画像内での歩行者の位置ずれによる検知結果への影響を明らかにできた。また、切り出し位置を複数設定し、歩行者と認識した位置の総数にもとづき判定することで、検知精度の向上が確認できた。

謝辞

本研究の一部は科学研究費補助金基盤研究(C)(22500161)の補助による。

参考文献

- [1]豊福邦彦, 勝野歳康, 萩里安雄, 石崎覚, 村野隆彦, 深町映夫, 澤田和久, 太場裕昌, “ナイトビュー (歩行者検知機能付)の開発”, TOYOTA Technical Review Vol.56 No.1(2008)
- [2]P.Viola and M.Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features”, Proc. of IEEE CVPR, vol.1, pp.511-518, 2001
- [3]藤吉弘亘, “Gradient ベースの特徴抽出-SIFT と HOG-”, 信学技報 PRMU2007-82, pp.211-223(2007)
- [4]奈良先端科学技術大学院大学 OpenCV プログラミングブック 製作チーム, “OpenCV プログラミングブック”, マイコミ(2007)

† 秋田県立大学大学院システム科学技術研究所
Graduate School of Systems Science and Technology, Akita Prefectural University
‡ クリエイティブソリューション株式会社
Creative Solutions Co.,Ltd.