

H-041

車載ステレオカメラシステムによる3次元情報を利用した歩行者検出の検討

A Study of Pedestrian Detection Using 3D Driving Space Data with In-Vehicle Stereo Camera System

服部 圭悟†
Keigo Hattori青木 孝文†
Takafumi Aoki

1. まえがき

近年、自動車の安全性や利便性の向上を目的としてカメラの車載化が進んでいる。交通事故発生の原因は約7割が認知遅れである。そこで、カメラを用いて事故要因を検出し、運転手に警告を与えるもしくは危機回避行動を行う高度運転支援システムに関する研究が注目されている。中でも歩行者検出は、事故削減において特に優先度の高い処理である。

従来の歩行者検出は、適当な大きさのウィンドウで画像全体を走査し、そのウィンドウ内に人がいるかどうかを学習アルゴリズムで判定する。現在までに、歩行者の判定方法については様々な手法が報告されているが[1]-[3]、歩行者の候補領域の生成についてはあまり議論されていない。ほとんどの場合、画像全体を様々な大きさのウィンドウで走査することで、距離や体格によって写り方が変わる歩行者に対応している。一方で、この処理による計算コストの高さと誤検出数の多さが問題となっている。これに対し、本論文では、走行空間の3次元情報を利用した歩行者候補領域の生成手法を提案する。3次元情報を利用することで、探索領域の制限と対象の大きさにあわせた候補領域の生成を行い、検出精度を向上させる。また、評価実験により提案手法の有効性を示す。

2. 走行空間の3次元計測

提案手法では、まず走行空間の3次元情報を取得するために、図1(a)に示すステレオカメラシステムを用いて走行空間を3次元復元する。使用する2台のカメラは平行となるように設置し、基線長は12cmに設定した。基線長を短くすることでステレオ画像間の対応付けを容易に行うことができるが、正確な3次元計測を行うことが困難であるという問題がある。この問題を解決するために、本論文では、位相限定相関法(Phase-Only Correlation: POC)に基づく高精度画像マッチングを用いることで、サブピクセル精度の対応付けを実現している[4]。これにより、従来は3次元計測の精度を向上させるために基線長を30cmから車両の横幅程度と広くする必要があったのに対し、提案手法ではコンパクトなステレオカメラでシステムを構成できる。

次に、走行空間の3次元情報に対して平面フィッティングを行い道路平面を検出する。3次元情報には先行車両や電柱などの道路以外の復元点や誤対応による外れ点が存在するため、RANSAC (RANdom SAmple Consensus) [5]を用いて道路平面を高精度に検出する(図1(b))。そして、検出した道路平面を基準とした座標系に3次元情報の座標系を変換し、道路より高い位置に存在する物体の領域を障害物領域として検出する。

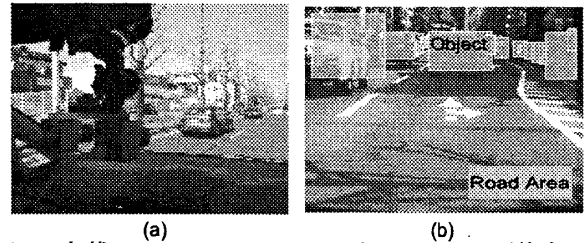


図1: 車載システム: (a) ステレオカメラ, (b) 道路と障害物の検出

3. 3次元情報を利用した歩行者検出

3.1 歩行者候補領域の生成

障害物領域の中から、3次元情報を利用して歩行者の候補領域を生成する。本論文では、歩行者として検出する対象を、成人男性の平均身長と平均肩幅を基に、高さ130cm~180cm、幅30cm~70cmと設定する。歩行者とカメラの相対的な距離に依存して画像上の歩行者のサイズが変わるため、前段の3次元計測の結果を利用して、対象の位置にあわせた適切なサイズの候補領域を生成する。以下に提案手法の処理手順を示す。

Step 1: 画像上の障害物領域から微小ブロックを切り出し、対応する3次元情報を取得する。そして、あらかじめ定義した歩行者の高さと横幅に合わせた大きさの候補領域を生成する。

Step 2: 候補領域を 2×2 の小領域に分割し、それぞれの領域の平均視差を算出する。この4つの平均視差が近い値を有しているか調べることで、候補領域が3次元的に対称性を持つかを判定し、対称性を持たない候補領域を破棄する。さらに、残った候補領域に対して、3次元的な対称性を持つ最大の領域となるように候補領域の大きさを調整する。

Step 3: 障害物領域全体を微小ブロックで走査しながら、Step 1~2の処理を行う。

以上の処理により、歩行者候補領域を生成する。提案手法では、障害物領域のみ探索をするため、従来手法に比べ計算コストと誤検出が少ない。また、Step 1の処理により、対象の位置にあわせた大きさの候補領域が生成できる。従来の1つのアプローチに、歩行者のエッジの対象性を利用する手法がある[1]。したがって、Step 2で候補領域の対称性を考慮することは、歩行者の特徴を捉える有効な手段である。図1(a)で取得した3次元情報は、キャリブレーション時に発生するカメラパラメータの推定誤差の影響を受ける。そのため、対称性に基づいて候補領域の大きさを調整することで、上記の誤差の影響に対応する。

3.2 歩行者の判定方法

歩行者の判定方法として Dalal らの手法 [2] と Shashua らの手法 [3] を用いる。Dalal らの手法は、特

†東北大学大学院情報科学研究科

表 1: 各手法の特徴

	Dalal の手法	Shashua の手法
特徴ベクトルの次元数	3240	117
検出精度	◎	○
姿勢変化	△	◎

特徴ベクトルとして HOG (Histograms of Oriented Gradients) を用いている。まず、候補領域の勾配強度と勾配方向を計算する。そして、候補領域をいくつかのブロックに分割し、ブロック毎に勾配強度で重み付けた勾配方向のヒストグラムを作成する。最後に、全てのヒストグラムを SVM (Support Vector Machine) で学習・判定する。それに対し Shashua らの手法は、まず体の部位毎に上記のヒストグラムを作成し、AdaBoost により部位毎の学習・判定を行う。そして、その判定結果を組み合わせ、SVM により学習・判定を行い、歩行者を検出する。各手法の特徴を表 1 に示す。Dalal らの手法は、候補領域全体を特徴空間に変換するため判定に使用する特徴ベクトルの数が多い、検出精度が高い。しかし、歩行者の姿勢と向きの変化に対応することができないという問題がある。それに対し、Shashua らの手法は、部位毎の判定により特徴ベクトルを減らすため検出精度はやや低下するが、部位毎の判定の組み合わせを用いることで体の向きや姿勢の変化にロバストな歩行者検出を実現できる。

4. 実験・考察

提案する歩行者検出システムを評価する。学習データの正解例は Dalal らの手法と Shashua らの手法で同じものを与える (約 4000 人)。不正解例は、構成した歩行者検出システムが誤検出したデータで与える。上記の処理を再帰的に行って、それぞれの学習スタイルに必要な不正解例を与える。評価に使うデータは実際に図 1(a) のシステムで撮影した 9 シーンを用いる。評価は、平均検出率 ($[\text{検出歩行者数}] / [\text{全歩行者数}]$)、平均誤検出数 ($[\text{誤検出歩行者数}] / [\text{全フレーム数}]$)、平均誤検出率 ($[\text{誤検出歩行者数}] / [\text{判定ウィンドウ数}]$) で与える。

実験結果を表 2、図 2 に示す。提案手法は、対象にあわせて大きさのウィンドウを作成できるため、歩行者の検出率が向上している。また、3 次元的な対称性を考慮することで、複数の物体を含む候補領域を除外でき、全探索に比べて誤検出数が削減している。誤検出率は全探索が提案手法よりも低い、これは全探索では参照するウィンドウ数が非常に多いためである。判定方法を比較すると、Shashua らの手法が歩行者の姿勢の違いに対してロバストに検出できるため、Dalal らの手法に比べ検出精度が高い結果となった。どちらの判定方法についても、提案手法は全探索より、検出率では約 10 ポイント、誤検出数では約 2.5 ポイント改善している。

5. むすび

本論文では、走行空間の 3 次元情報を利用した歩行者検出手法を提案した。実験を通して、検出率や誤検出数の点で従来の全探索よりも提案手法が優れていることを示した。また、提案手法は様々な判定アルゴリズムと組み合わせることが可能であり、幅広い応用が

表 2: 実験結果

	手法 1	手法 2	手法 3	手法 4
検出率 [%]	63.53	77.07	71.05	80.83
誤検出数 [人]	5.95	2.95	5.20	3.15
判定ウィンドウ数	972.63	132.60	1065.40	132.60
誤検出率 [%]	0.61	2.22	0.49	2.37

手法 1: 全探索+Dalal, 手法 2: 提案手法+Dalal
手法 3: 全探索+Shashua, 手法 4: 提案手法+Shashua

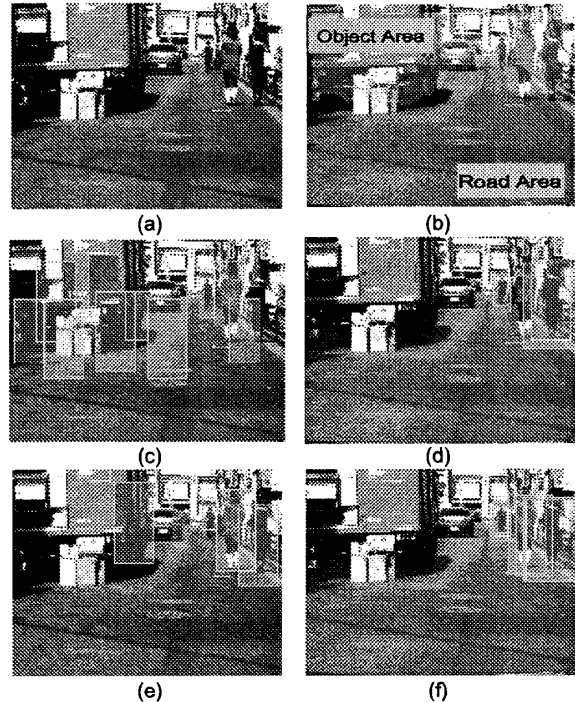


図 2: 実験結果: (a) 原画像, (b) 障害物領域の検出, (c) 全探索+Dalal, (d) 提案手法+Dalal, (e) 全探索+Shashua, (f) 提案手法+Shashua

期待できる。今後は、3 次元情報を利用したより高度な運転支援システムの検討や検出した歩行者のトラッキングを考えている。

参考文献

- [1] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-based pedestrian detection," *IEEE Intelligent Vehicles Symposium*, pp. 215-220, 2000.
- [2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 886-893, 2005.
- [3] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," *Intelligent Vehicles Symposium, IEEE*, pp. 1-6, 2004.
- [4] K. Takita, M.A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," *IEICE Trans. Fundamentals*, pp. 1913-1923, 2004.
- [5] Martin A. Fischler and Robert C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. of the ACM*, pp. 381-395, 1981.