

ステレオセンサを用いた歩行者認識

松島 宏典 胡 振程 内村 圭一

熊本大学大学院自然科学研究科

〒860-8555 熊本県熊本市黒髪 2 丁目 39 番 1 号

あらまし 歩行中の死傷者数を低減するため、車載カメラを利用した歩行者認識システムの要求が高まっている。本論文では、車両に設置した左右二台のカメラから得られる画像をもとに、三次元情報を考慮した歩行者認識手法について提案する。まず、左右の画像に対してステレオマッチングを行い視差画像を算出後、道路環境解析手法である U-V-disparity を適用することで歩行者候補領域を抽出する。それらの領域に対して、視差情報と形状判定により背景領域を削除し、歩行者候補の絞込みを行う。その後、Gabor Filter により歩行者の輪郭形状を抽出した後、Support Vector Machine(SVM) を用いて歩行者の認識を行う。これらの手法を道路撮影画像に適用し、有効性の確認と検証を行う。

キーワード ステレオ視, U-V-disparity, 歩行者認識, ITS

Pedestrian Recognition using Stereo Sensor

Kousuke Matsushima Zhencheng Hu Keiichi Uchimura

Graduate School of Science and Technology, Kumamoto University

2-39-1, Kurokami, Kumamoto-city, Kumamoto, 860-8555, Japan

Abstract Pedestrian-vehicle accidents account for about one third of the total automotive related fatality and injury. So the inclusion of pedestrian protection systems such as external airbags is being consider as a solution to preventing pedestrian fatality and injury. However, such systems require knowledge of perestrian presence for correct activation. In this work, it proposes the pedestrian recognition technique for considering three-dimensional information based on the image obtained from right and left cameras set up in the vehicle. These techniques are applied to the road scene images, and the confirmation and the verification of effectiveness are done.

Keywords Stereo Vision, U-V-disparity, Pedestrian Recognition, ITS

1. まえがき

近年、高度道路交通システム (ITS) に関する研究・開発が盛んに行われている。その中でも、車載センサを用いて道路や車両、歩行者などを分類する走行環境解析技術は、安全運転支援の重要な要素の一つとなっている。特に、歩行者の認識技術は、運転者の前方不注意などによる認知遅れを低減するため、実用化への要求が高まっており、様々な研究が行われている。

これまで、歩行者認識に関する研究は、歩行者の輪郭や顔部分といった特徴的な形状をモデルとして用意し、それらを撮影画像と照合させるモデルマッチングによる方法^{1)~4)}や、歩行リズムや歩幅など

を特徴要素として用い、移動物体が歩行者であるかどうかを識別する動作検出による方法^{5),6)}などが提案されている。しかし、前者は歩行者には様々な形状が存在するために、モデルを複数個用意しなければならず、かつ原画像上の全ての位置に対してマッチングを行う必要があるため、計算量が膨大となる問題がある。後者は、道路上を一定歩調で歩行していることが前提となっているため、停止している歩行者や歩行以外の動作を行っている歩行者には対応できない。また、赤外線カメラとレーザスキャナ、ジャイロセンサと車速センサをそれぞれ併せて用いるセンサフュージョンによる手法^{7)~10)}も提案されているが、これらはセンサが複数個必要となる



図1 一般道路における画像の例

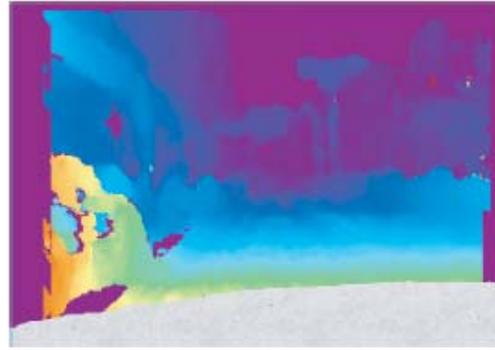


図3 一般道路における視差画像の例

ため、コスト増加が問題となる。

そこで、本研究では、一般的な車載カメラで構成されたステレオセンサによる歩行者認識手法を提案する。我々は、以前に走行環境解析手法である“U-V-disparity”を提案している¹¹⁾。よって、この方法を適用することで歩行者候補領域を抽出する。また、誤認識の要因となる背景情報に関しては、視差情報と形状判定を組み合わせた方法を用いて歩行者候補の絞込みを行う。その後、ガボールフィルタを適用して歩行者の輪郭形状を抽出し、SVMにより歩行者の認識を行う。以下、2.で本論文の提案手法の詳細について述べ、3.で実道路環境上での実験結果を示す。そして、最後に4.で結論をまとめる。

2. 提案手法

ここでは、我々が提案する歩行者認識手法について説明する。手法の概略を以下に示す。

(1) 左右のカメラから得られた2枚の画像に対してステレオマッチングを行い、視差画像を取得する。

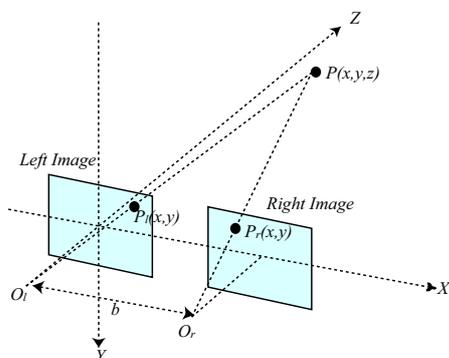


図2 ステレオ視におけるカメラ構成

- (2) 視差画像に U-V-disparity を適用し、歩行者候補領域を抽出する。
 - (3) 視差情報を形状判定を利用して背景部分の除去を行った後、歩行者領域にガボールフィルタを適用する。
 - (4) SVM を適用し、歩行者の認識を行う。
- 以降では、上記(1)~(4)について説明する。

2.1 ステレオ視における基本原理

図2にステレオ視における基本原理を示す。道路環境の撮影は、移動車両前方に設置された左カメラ、右カメラにより撮影される。従って、対応付けのためのマッチングは、キャリブレーション後、エピポーラ線上において左画像と右画像で行われる。なお、マッチングにはSSDA法に基づいた手法を適用している。

ここで、画像間の視差に関する基本的な関係を示す。左画像と右画像のそれぞれの対応点を $P_l(x_l, y_l)$, $P_r(x_r, y_r)$ とすると、左右画像間における視差 d_{lr} は(1)式で求められる。

$$d_{lr} = x_l - x_r = f \frac{b}{z} \quad (1)$$

このとき、 f は焦点距離、 b は基線長、 z は奥行きを示すパラメータである。図3に一般道路における左右画像間のマッチングにより得られた視差画像を示す。

2.2 U-V-disparity

2.2.1. U-V-disparity の基本概念

車両に取り付けたカメラで撮影した画像において、前方の車両以外に、路面・隣接車線や対向車線を走行中の車両およびそれ以外の障害物（歩行者、

自転車等)を検出し、クラスタリングすることを目的とする。

Labayrade は, (u, v) 画像座標系において垂直座標にあたる V を利用する“V-disparity”という概念を提案している¹²⁾。これは、視差画像の垂直方向の座標を V として縦軸に、画像の視差値を disparity として横軸にそれぞれ投影した 2 次元平面 (V-disparity 平面) を示す。この平面により、画像の垂直方向 V の視差変化に注目し路面から障害物・前方車両の後方を抽出するものである。しかし、実際の道路環境においては、路面上の障害物以外にも、車両横面やガードレール、路側帯、歩道橋など、V-disparity では検出できないものが多く存在する。そこで、本論文では三次元走行環境を、路面・道路上の障害物・路側帯にある建造物等に分類する“U-V-disparity”を利用する。

2.2.2 U-V-disparity 領域

ここでは、左右に設置されたカメラに対して説明する。両カメラは同じパラメータを持ち、道路平面に対して平行であるととし、そのときのピッチ角は θ であるとする。このとき、世界座標系を $WCS(X_w, Y_w, Z_w)$ 、左カメラ座標系を $LCCS(X_l, Y_l, Z_l)$ 、右カメラ座標系を $RCCS(X_r, Y_r, Z_r)$ としたときの図を図 4 に示す。カメラ座標系の原点は (u_0, v_0) で画像中央とし、座標は (u, v) で与えられるものとする。

図 4 に示すとおり、世界座標系からカメラ座標系への変換は、平行移動パラメータ $\pm \frac{b}{2}$ 、回転パラメータ $-\theta$ により表現することができる。従って、平行移動行列を $\mathbf{T}_l, \mathbf{T}_r$ 、回転行列を $\mathbf{R}_l, \mathbf{R}_r$ としたとき、世界座標系からカメラ座標系への変換行列 $\mathbf{D}_l, \mathbf{D}_r$ は以下のように表すことができる。

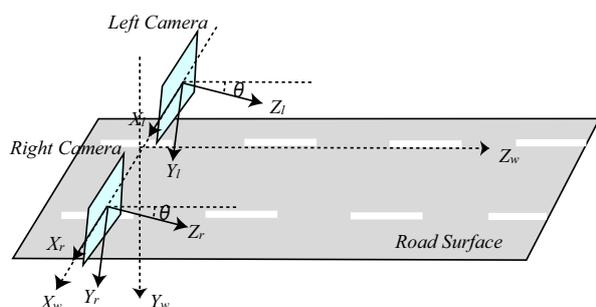


図 4 世界座標系とカメラ座標系との関係

$$\mathbf{D}_* = \mathbf{R}_* \mathbf{T}_* = \begin{pmatrix} 1 & 0 & 0 & \pm \frac{b}{2} \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

ここで、*は l または r とする (以降同様)。カメラのアスペクト比を 1.0 とし、左右カメラの焦点距離を f とすると、透視変換行列は以下のように表すことができる。

$$\mathbf{P} = \begin{bmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3)$$

従って、世界座標の同次行列 $(X_w, Y_w, Z_w, 1)^T$ と画像座標の同次行列 $(u, v, 1)^T$ の投影変換は以下のように表すことができる。

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P} \mathbf{D}_* \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \mathbf{M}_* \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (4)$$

ここで、 \mathbf{M}_* は変換行列を示しており、以下のように表すことができる。

$$\mathbf{M}_* = \begin{bmatrix} f & u_0 \sin \theta & u_0 \cos \theta & \pm \frac{b}{2} \\ 0 & v_0 \sin \theta + f \cos \theta & v_0 \cos \theta - f \sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \end{bmatrix} \quad (5)$$

式 (4)、(5) より、画像座標 (u_*, v) は簡単に算出することができる。

$$\begin{aligned} u_* &= u_0 + f \frac{X_w \pm b/2}{Y_w \sin \theta + Z_w \cos \theta} \\ v &= v_0 + f \frac{Y_w \cos \theta - Z_w \sin \theta}{Y_w \sin \theta + Z_w \cos \theta} \end{aligned} \quad (6)$$

式 (6) を簡略化するために、カメラの光軸に関して新たに画像座標 (U, V) を定義すると、式 (7) となる。

$$\begin{aligned} U_* &= u_* - u_0 = f \frac{X_w \pm b/2}{Y_w \sin \theta + Z_w \cos \theta} \\ V &= v - v_0 = f \frac{Y_w \cos \theta - Z_w \sin \theta}{Y_w \sin \theta + Z_w \cos \theta} \end{aligned} \quad (7)$$

また、視差 Δ は式 (6) から次のように表すことができる。

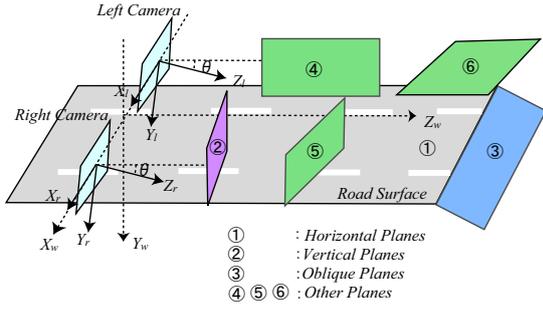


図5 世界座標系における水平面と垂直面

$$\Delta = u_l - u_r = f \frac{b}{Y_w \sin \theta + Z_w \cos \theta} \quad (8)$$

式 (7), (8) は、画像座標 (U, V) とステレオ視差 Δ の基本的な関係を示している。また、図5で示しているように、WCSに存在する物体は、水平面、垂直面、斜平面といった種々の平面にて定義づけを行っている。ここで、道路走行環境上の歩行者は垂直面に該当し、ある範囲の幅と高さを持つ物体とみなすことができる。よって、ROI(Region of Interest)により歩行者と関連性が小さい視差情報の削減を行った後、投影されたU-V-disparity領域に対して歩行者に該当する垂直面を抽出し、それらを歩行者候補領域とする。

2.3 3D-like Gabor Filter

本手法では、視差情報を考慮した画像に対して、ガボールフィルタを実行することで歩行者の輪郭特徴量を抽出する。まず、U-V-disparityにより求めた歩行者候補領域を、視差画像に対して一定範囲の視差値 D で分割したあと(図6参照)、2値化、ノイズ除去を実行する。次に、形状判定により背景領域等を削除し歩行者候補領域の絞込みを行うため、面積 (area)、最大長 (max_length)、周囲長 (boundary_length) を算出後、以下の式を適用する。

$$\text{Pedestrian Level} = \alpha \cdot RL + \beta \cdot CL \quad (9)$$

ここで、 RL , CL は丸さ度合い、凹凸度合いを表し、それぞれ以下の通りである。

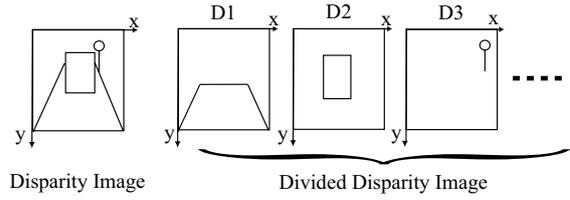


図6 分割された視差画像の例

$$RL = \frac{(\text{max_length})^2}{\text{area}} \times \frac{\pi}{4} \quad (10)$$

$$CL = \frac{(\text{boundary_length})^2}{\text{area}} \times \frac{1}{4\pi}$$

α, β は各度合いに対するパラメータである。式 (9) により、一定閾値以上の値を持つものを新たに歩行者候補領域とする。

その後、これらの領域に対して前処理後、ガボールフィルタ $g(x, y)$ を実行する。式は以下の通りである。

$$g(x, y) = \hat{g}(x, y) \cdot \exp\{j2\pi(u_0 t_x + v_0 t_y)\} \quad (11)$$

ここで、 $\hat{g}(x, y)$ はガウス関数であり、次のように表現される。

$$\hat{g}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp^{-\frac{t_x^2 + t_y^2}{2\sigma_x\sigma_y}} \quad (12)$$

$u_0, v_0, \sigma_x, \sigma_y$ はそれぞれ x, y 方向の周波数と窓幅を示しており、 t_x, t_y はカーネル中心からの距離である。

2.4 Support Vector Machine

Support Vector Machine(SVM) は、Vaptikらによって提案された2値分類のための学習アルゴリズムであり¹³⁾、 n 次元ユークリッド空間上に配置されたデータを2分する超平面(hyperplane)を見つけるためのアルゴリズムである。図7は線形の場合における分離超平面を示している。このとき、任意のデータ点 $\mathbf{x} = (x_1, \dots, x_d) \subseteq R^n$ (●: 正例, ○: 負例) の識別関数は次のように定義することができる。

$$f(x) = \sum_{j=1}^d w_j x_j + b \quad (13)$$

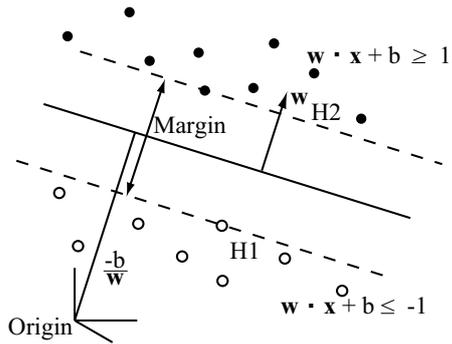


図7 線形の場合における分離超平面

ここで、 $w \in \mathbf{R}^n$ は重みベクトルであり、 $b \in \mathbf{R}$ はバイアス項である。図に示すように、近接する正例と負例のデータの間隔が大きいくほど、精度良くテストデータを分離することができる。なお、超平面 H1 あるいは H2 を超えて反対側に入り込むことを許すソフトマージンや、非線形に拡張するためのカーネルトリックも用いている。

3. 実験結果

図8は左カメラから得られた実道路環境画像と、その視差画像を示している。視差画像では、ROIを用いることによりカメラ遠方の視差情報が削除されているのが分かる。なお、ステレオマッチングの処理速度に関しては、PentiumIV 2.26GHzの汎用PCを用いた場合、最大視差32、画像サイズ640×480という条件で99.8msecという結果を得た。

また、U-V-disparityの結果を図9に示し、歩行者候補領域の抽出結果を図10に示す。歩行者の視差値の変化は、x軸方向、y軸方向でともに小さい。よって、U-disparity領域ではx軸方向にほぼ平行な線分として、V-disparity領域ではy軸方向にほぼ平行な線分としてそれぞれ抽出される。従って、それぞれの領域の線分を対応付けることによって、歩行者候補領域を抽出することができる。これにより、画像全体を探索せず歩行者候補領域のみを対象とすることができる。

図11に歩行者候補領域とその視差画像を、図12にガボールフィルタ結果と形状判定により背景領域等を削除したときのマスク画像を、図13に歩行者の特徴量抽出結果を示す。通常のフィルタ結果(図12左参照)では、背景領域の情報も抽出していることがわかる。従って、視差情報、形状判定を考慮す



図8 実験画像と視差画像

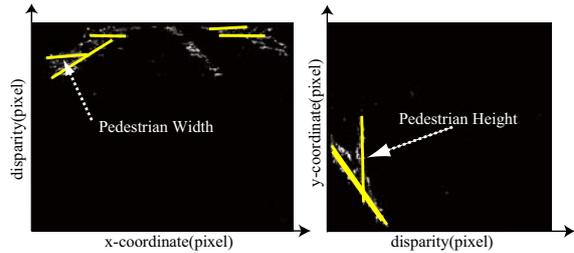


図9 U-V-disparity 領域



図10 歩行者候補領域の抽出結果



図11 歩行者候補領域と視差画像

ることにより、図13のように歩行者の輪郭部分のみを抽出することができる。また、SVMによる歩行者認識率は約88%であった。認識率については、ステレオマッチングやSVMのパラメータの最適化に関する検討、垂直エッジ処理を含めた背景領域の削除を行うことにより向上するものと思われる。



図 12 フィルタ結果とマスク画像



図 13 歩行者特徴量

4. まとめ

本論文では、ステレオセンサを用いた歩行者認識手法を提案した。歩行者候補領域の抽出には“U-V-disparity”を用いた。また、誤認識の要因となる背景情報に関しては、視差情報と形状判定を組み合わせた方法を用いることで、歩行者候補の絞込みを行い、SVMにより歩行者の認識を行った。今後は、各種センサとステレオセンサを融合したセンサフュージョンにより、更なる歩行者認識率の向上を行っていく予定である。

参考文献

- [1] D.M. Gavrila, J. Giebel and S. Munder, “Vision-Based Pedestrian Detection: The PROTECTOR System”, IEEE Intelligent Vehicle Symposium, June, 2004.
- [2] M. Mahlich, M. Oberlander, O. Lohlein, D. Gavrila and W. Ritter, “A Multiple Detector Approach to Low-resolution FIR Pedestrian Recognition”, IEEE Intelligent Vehicle Symposium, June, 2005.
- [3] M. Szarvas, A. Yoshizawa, M. Yamamoto and J. Ogata, “Pedestrian Detection with Convolutional Neural Networks”, IEEE Intelligent Vehicle Symposium, June, 2005.
- [4] A. Shashua, Y. Gdalyahu, G. Hayun, “Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System Level Performance”, IEEE Intelligent Vehicle Symposium, June, 2004.
- [5] A. Broggi, A. Fascioli, M. Carletti, T. Graf, M. Meinecke, “A Multi-resolution Approach for Infrared Vision-based Pedestrian Detection”, IEEE Intelligent Vehicle Symposium, June, 2004.
- [6] 安富 敏, 森 英雄, 清弘 智昭, “歩行のリズムに基づく歩行者検出の一手法”, 信学論, Vol.J78-D-, no.4, pp.608-617, 1995.
- [7] B. Fardi, U. Schuenert, G. Wanielik, “Shape and Motion-based Pedestrian Detection in Infrared Images: A Multi Sensor Approach”, IEEE Intelligent Vehicle Symposium, June, 2005.
- [8] 望月 大介, 矢野 良和, 橋山 智訓, 大熊 繁, “車載カメラによる歩行者検出のための背景除去及びアクティブ探索を用いたテンプレートマッチングの高速化”, 信学論, Vol.J87-D-, No.5, pp.1094-1103, 2004.
- [9] D.T. Linzmeier, D. Vogt, R. Prasanna, “Probabilistic Signal Interpretation Methods for a Thermopile Pedestrian Detection System”, IEEE Intelligent Vehicle Symposium, June, 2005.
- [10] T.Matthias, D.Reiner, M.Marc-Michael, Obojski, Marian Andrzej, “Radar Sensors and Sensor Platform used for Pedestrian Protection in the EC-funded project SAVE-U”, IEEE Intelligent Vehicle Symposium, June, 2004.
- [11] 松島 宏典, 松浦 博人, 木嶋 啓貴, 胡 振程, 内村 圭一, “三眼立体視による道路環境の認識”, 電気学会 電子・情報・システム部門大会, GS7-1, pp.841-846, 2005.
- [12] R. Labayrade, D. Aubert, and J. P. Tarel, “Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through “V-disparity” Representation”, IEEE Intelligent Vehicle Symposium, pp.646-651, 2002.
- [13] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.