

車載単眼カメラによる車両前方の障害物検出

山口 晃一郎 加藤 武男 二宮 芳樹

(株) 豊田中央研究所

あらまし: 本稿では, 予防安全のための運転支援システムの要素技術として, 車載単眼カメラを用いて自車両の運動を推定し, 車両前方の障害物を検出する手法を提案する. 走行環境において単眼画像から自車両の運動を推定する場合, 画像中に他の車両などの移動物体が含まれることや路面部分にパターンが少ないことから推定が困難となる課題がある. 提案手法では, 一時刻前での路面領域, 移動物体領域の検出結果を利用して, 運動推定に用いる特徴点を選択することで精度良く自車両の運動を推定する. そして, 推定された自車両の運動から走行環境の三次元形状を復元するとともに自車両によって生じる動きと異なる動きの領域を移動物体として検出する. 実走行画像を用いた実験によって, 前方の車両, 歩行者が検出可能であることを示す.

Obstacle Detection in Road Scene using Monocular Camera

Koichiro Yamaguchi Takeo Kato Yoshiki Ninomiya

Toyota Central R&D Labs., Inc.

Abstract: This paper proposes a method to estimate the ego-motion of the vehicle relative to the road and detect obstacles on the road by a monocular camera mounted on the vehicle. There are two problems in ego-motion estimation. First, a typical road scene contains moving objects such as other vehicles. Second, roads have few feature points. In our approach, ego-motion is estimated from correspondence points extracted in various regions except moving object regions. After estimating the ego-motion, three dimensional scene structure is reconstructed and moving obstacles are detected. In our experiments, it is shown that proposed method is able to detect moving obstacles such as vehicles and pedestrians.

1 はじめに

近年, 交通事故による死傷者の低減を目指して自動車の安全性を高める運転支援システムの開発が行われている. 運転支援システムの1つである衝突防止支援システムでは, 自車周辺の車両, 二輪車, 歩行者などの物体を検出することが要求され, レーザレーダ, ミリ波レーダ, 画像センサなどが利用される. それらのセンサの中で画像センサは, その高い空間解像度から物体の形状推定が可能となる点で期待されている. 単眼カメラによる障害物

検出手法としては, 水平エッジを利用して前方の車両を検出する手法 [1], パターン認識により歩行者を検出する手法 [2] などが提案されており, これらの方法により車両, 歩行者をそれぞれ検出することが可能である. 一方, 車載の単眼カメラから得られる時系列画像を用いて走行環境の三次元構造を推定し, 障害物の検出を行う方法も考えられ, その方法を用いると車両, 二輪車, 歩行者など様々な障害物を検出することができる. 時系列画像から走行環境の三次元構造を推定するにはカメラの運動, つまり自車両の運動を推定することが要求

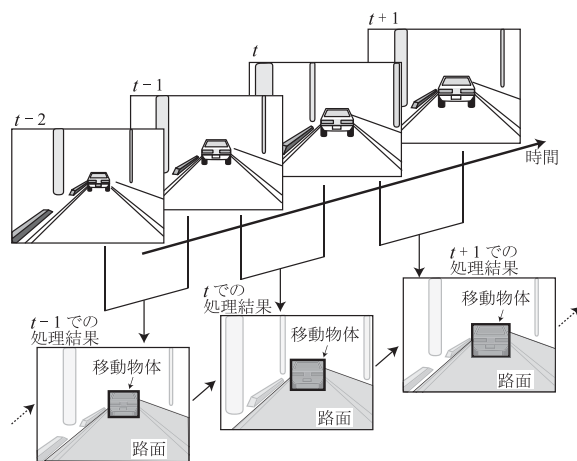
され、車速センサ、ヨーレートセンサからの情報を用いる方法が通常考えられるが、精度が十分でないことや車輪の滑りが発生した場合に正しく計測できないという問題がある。

単眼カメラにより、自車両の運動を推定する方法として、画像中での路面領域のオプティカルフローを利用する方法が提案されている [3]。しかし、路面領域にパターンが少ない場合には、正しいオプティカルフローを計算することは困難である。そこで、路面領域を小領域に分割し、ある運動に対して小領域ごとに計算される評価値を組み合わせた評価関数が最適となる運動を探索する Direct Method による方法が提案されている [4]。この方法では、オプティカルフローを計算する必要がないため、ロバスト性を向上させることができる。しかし、推定するパラメータが多い場合には安定した計算が困難となるため、回転、並進の 6 自由度のパラメータをすべて推定することは難しい。一方、画像間の対応点を利用してカメラ運動およびその三次元構造を推定する Structure from Motion(SFM) による手法 [5] を用いる場合、走行環境では他の車両など多数の移動物体が存在することが問題となる。

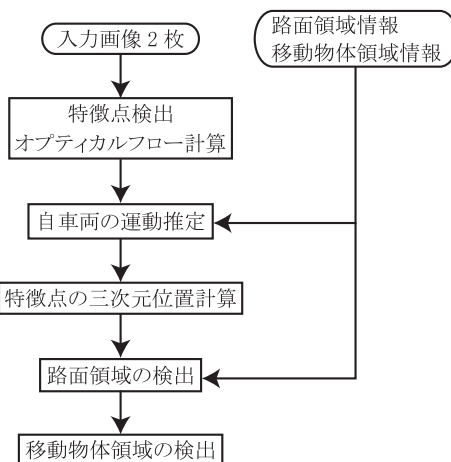
車載の単眼カメラから障害物を検出する手法としては、オプティカルフローから路面上にない点を判別し、障害物を検出する手法が提案されている [6]。しかし、単眼カメラから自車両の運動と走行環境の三次元構造を推定し、障害物を検出する実用可能な方法はこれまで十分検討されていない。

本研究では、車載単眼カメラにより得られる時系列画像から自車両の運動を推定し、車両、二輪車、歩行者などの移動物体を検出する手法を提案する。自車両の運動の推定では、一時刻前に検出した路面領域と移動物体領域の結果を利用し、運動推定に用いる画像中の特徴点を選択した上で SFM の方法を適用することにより移動物体の影響を除去し、推定誤差を低減する。また、特徴点を連続して追跡することにより自車両の運動により生じる動きとは異なる動きの特徴点を判別し、それらの特徴点から移動物体領域を検出する。

以下、まず 2 章において提案手法の概要を述べる。3 章では、自車両の運動推定、4 章では路面領域の検出、5 章では移動物体領域の検出の各処理について詳しく説明する。6 章では、実走行画像を用いた実験結果を示し、最後に 7 章でまとめと今後



(a) 時系列の処理の流れ



(b) 各時刻の処理

図 1: 障害物検出処理の流れ

の課題を述べる。

2 提案手法の概要

提案手法の処理の流れを図 1 に示す。なお本手法では、カメラは校正済みつまりカメラの内部パラメータは既知であると仮定し、レンズ歪みも補正済みであるとする。提案手法では、図 1(a) に示すように各時刻において、現時刻の画像と一時刻前の画像の 2 枚を用いて処理を行う。つまり、時刻 t においては、時刻 $t, t-1$ の 2 枚の画像を用いる。さらに、時刻 $t-1$ での処理結果も利用し、障害物検出処理を行うこととする。なお初期時刻では、移動物体領域はなし、路面領域はカメラの設置高さ、俯角を考慮してあらかじめ決めた領域とする。各時刻における処理の流れを図 1(b) に示す。各時刻

での入力とは時間的に連続する 2 枚の画像および一時刻前で検出した画像中の路面領域、移動物体領域の情報である。2 枚の画像が入力されると、まず画像中の特徴点を検出し、そのオプティカルフローを計算する。特徴点の検出には、Harris オペレータ [7] を使用する。すべての画素について Harris オペレータの特徴量を計算し、極大値かつしきい値以上となる点を特徴量の大きい順に指定した数だけ検出する。また、オプティカルフローの計算には、Lucas-Kanade 法 [8] を用いる。次に特徴点のオプティカルフローから画像間の運動パラメータを求める。このとき、一時刻前で路面領域、移動物体領域の検出結果を利用し、運動パラメータの推定に用いる特徴点を選択する。得られた運動パラメータから特徴点の三次元位置を求め、路面平面を推定して現時刻での路面領域を検出する。最後に、推定した自車両の運動とは異なる運動の特徴点を判別し、移動物体領域を決定する。

3 自車両の運動の推定

入力された 2 枚の画像間での特徴点の対応から、画像間の位置関係、つまり自車両の運動を推定する。推定する運動は、3 自由度の回転と 3 自由度の並進をあわせて 6 つのパラメータである。以下、運動推定に用いる特徴点の選択方法と運動の推定方法について述べる。

3.1 特徴点の選択

3.1.1 課題と選択手法の概要

静止環境の場合には正しい 8 点の対応点が取得できれば、画像間の運動を求めることができるが、走行環境では以下のような課題がある。

- (i) 走行環境では一般的に他の車両などの移動物体が存在し、画像中にはそれらの移動物体が含まれる。移動物体上の対応点は静止環境の仮定を満たさないため、自車両の運動を誤って推定する原因となる。
- (ii) 通常、路面領域にはパターンが少なく、建物や他の車両などの立体物を含む背景領域にはパターンが多いため、画像から単純に特徴点を抽出すると、図 2 に示すように画像の特定の

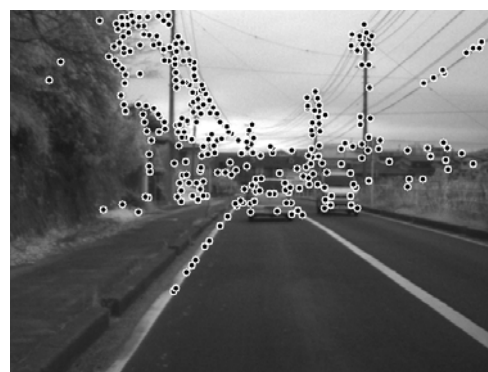


図 2: 走行環境で検出された特徴点

部分に特徴点が多く集中する傾向がある。遠方にある建物など一部に特徴点が集中し、画像内での特徴点の分布に偏りが生じると、推定した運動の誤差が大きくなる可能性がある。

これらの課題に対応するため、運動推定に用いる特徴点を以下のように選択する。

- (i) 一時刻前で検出された移動物体領域に含まれる特徴点を取り除くことにより、移動物体上の特徴点の数を低減する。
- (ii) 画像を路面領域、低い立体物の領域、高い立体物の領域の 3 つの領域に分割し、それぞれの領域から特徴点を選択することにより、画像中の様々な領域からバランスよく特徴点を抽出する。

以上のように特徴点を選択することにより、移動物体の影響を除去し、推定誤差を低減することが可能となる。

3.1.2 特徴点検出手法の説明

図 3 に示す前段の処理により検出された特徴点、および一時刻前で検出された路面領域、移動物体領域を用いて、運動推定のための特徴点の選択を行う。まず、図 4 に示すように画像を路面、低い立体物、高い立体物を多く含むような 3 つの領域となるよう水平に分割する。この画像の分割は以下の手順で行う。

1. 一時刻前で検出された路面領域のうち、ある割合以上の領域を含むように、画像の下部を水平に分割し、路面を多く含む領域を設定する。



(a) 特徴点



(b) 移動物体領域



(c) 路面領域

図 3: 特徴点と一時刻前での移動物体領域, 路面領域検出結果



図 4: 領域の分割

- 設定した路面を多く含む領域の上部の領域を水平に 2 分割し, 低い立体物領域と高い立体物領域とする.

以上のようにして分割された各領域ごとに, あらかじめ設定した割合の数だけ特徴点を選択する. 各領域では, Harris オペレータの特徴量が大きい特徴点から順に選択する. このとき, 一時刻前で検出された移動物体の領域内となる特徴点は選択しない.

以上の処理により, 図 5 に示すように, 移動物体上の特徴点が除去され, 路面領域, 背景領域の様々な部分から特徴点をバランスよく検出することができる. 一時刻前で正しく移動物体領域が検出されなかった場合には, 移動物体上の特徴点を除去することができないが, 画像を分割して特徴点を選



図 5: 提案手法により選択された特徴点

択することにより, 移動物体上に特徴点が集中することを防ぎ, 次節で述べるように自車両の運動の推定にロバスト推定法の RANSAC (Random Sample Consensus) [9] を用いることで移動物体の影響を除去することができる.

3.2 運動の推定

前節の方法により選択した特徴点とそのオプティカルフローから求まる画像間での対応点を用いて, 以下の式で表される基本行列 E を計算する.

$$E = [t]_{\times} R = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} R \quad (1)$$

ここで, R, t は画像間のカメラ運動を表し, R は回転行列, $t = (t_x, t_y, t_z)^{\top}$ は並進ベクトルである. 基本行列 E は以下の式を満たす.

$$\bar{p}'^{\top} E \bar{p} = 0 \quad (2)$$

$$\bar{p} = K^{-1} p, \quad \bar{p}' = K^{-1} p' \quad (3)$$

$$K = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

p は現時刻の画像の点で, p' は点 p に対応する一時刻前の画像の点である. また, \bar{p}, \bar{p}' はそれぞれ点 p, p' の正規化座標で, K はカメラのキャリブレーション行列を表す. カメラの内部パラメータは, 焦点距離 f_x, f_y , 画像中心 (c_x, c_y) , skew s で, 本研究では $s = 0$ とする. この基本行列 E は前節の方法で選択した特徴点の対応を用いて 8 点アルゴリズム [10] と RANSAC により計算する. RANSAC

を用いることにより，誤検出した対応点および特徴点選択時に除去できなかった移動物体の対応点の影響を取り除くことができる．

次に，得られた基本行列 E から回転行列 R ，並進ベクトル t を計算する．さらに，Levenberg-Marquardt 法を用いて運動パラメータの最適化を行う．

4 路面の検出

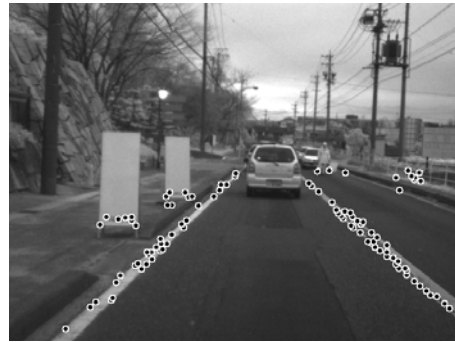
本章では，自車両の運動を推定し，特徴点の三次元位置を計算した後に，路面領域を検出する処理について説明する．まず，特徴点の三次元位置と一時刻前での路面領域の検出結果 (図 3(c)) を用いて，現時刻における三次元空間での路面平面の位置と傾きを推定する．そして，画像を小領域に分割し，小領域ごとに路面領域かどうかの判定を行う．また，単眼時系列画像から復元した三次元形状には定数倍の不定性が存在するため，推定した路面平面の位置と実際に設置したカメラの地上高からこの不定性を取り除く処理を行う．

4.1 路面平面の推定

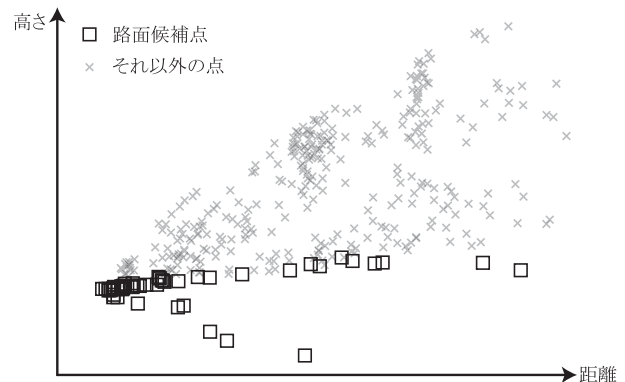
一時刻前での路面領域の検出結果を用いて，その領域内にある特徴点を抽出し，それらの特徴点を路面上の点の候補点とする．抽出した路面候補点から三次元空間での路面平面を推定する．図 6 に抽出された路面候補点とそれらの候補点の三次元位置を示す．現時刻でのカメラ座標系を (X, Y, Z) ，ただし原点はカメラの投影中心， X 軸は画像平面の水平方向， Y 軸は画像平面の垂直方向， Z 軸は光軸方向とすると，推定する平面の式は

$$aX + bY + cZ = 1 \quad (5)$$

と表すことができる．この平面のパラメータ a, b, c は最低 3 点の平面上の点の三次元位置があれば計算することが可能である．しかし，図 6 に示すように，抽出された特徴点の中には，一時刻前での路面領域の誤検出から路面上にない点やオプティカルフローの誤検出により三次元位置に誤りがある点が含まれる．そこで，LMedS (Least Median of Squares) 推定を用いて，それらの誤りの影響を低



(a) 路面候補点



(b) 三次元空間を横から見た図

図 6: 路面候補点とその三次元位置

減して路面平面の推定を行う．以下，路面平面推定の手順を示す．

1. 一時刻前での検出結果の路面領域にある特徴点を抽出し，路面候補点とする．
2. 路面候補点から 3 点をランダムに選択する．
3. 選択した 3 点から式 (5) のパラメータ a, b, c を求める．
4. 1. で抽出したすべての路面候補点について，推定した平面との距離を計算し，そのメディアン値を求める．
5. 2. ~ 4. を k 回繰り返す，推定平面との距離のメディアン値が最小となるパラメータの平面を路面平面とする．

繰り返し回数 k は，次のようにして決定する． k 回ランダムに 3 点を選択したとき，少なくとも 1 回は路面上の点のみとなる確率 P は，抽出したすべての路面候補点中の路面以外の点の割合を ϵ とすると，次式のように表せる．

$$P = 1 - \{1 - (1 - \epsilon)^3\}^k \quad (6)$$

よって、路面上でない点の割合 ϵ と確率 P を決めることにより必要な繰り返し回数 k を求めることができる。

また、単眼時系列画像から復元した三次元空間には定数倍の不定性があるため、推定した路面平面を用いてこの不定性を取り除く処理を行う。推定した路面平面が式 (5) のように表されるとき、カメラの投影中心から路面平面までの距離 d_0 は、以下の式で表される。

$$d_0 = \frac{1}{\sqrt{a^2 + b^2 + c^2}} \quad (7)$$

この距離 d_0 を実際のカメラの設置高さに合わせることで、三次元空間のスケールを決定する。

4.2 路面領域の検出

前節で推定された路面平面を用いて、画像中の路面領域を検出する処理について説明する。画像上の点 (x, y) が、平面上の点を投影した点であるとすると、そのオプティカルフロー (u, v) は、以下のような 8 つのパラメータの式で表すことができる [4]。

$$u(x, y) = \alpha_1 x + \alpha_2 y + \alpha_3 + \alpha_7 x^2 + \alpha_8 xy \quad (8)$$

$$v(x, y) = \alpha_4 x + \alpha_5 y + \alpha_6 + \alpha_7 xy + \alpha_8 y^2 \quad (9)$$

ここで、 $\alpha_i, i = 1..8$ は画像間の運動パラメータ、平面パラメータ、カメラ内部パラメータから計算される値であり、前段までの処理結果を用いて、推定した平面に対する式 (8),(9) の α_i を直接計算することができる。

次に、現時刻の画像を 10×10 画素の小領域 W_i に分割し、式 (8),(9) から一時刻前の画像で対応する小領域 $W'_i = \{(x', y') | x' = x + u(x, y), y' = y + v(x, y), (x, y) \in W_i\}$ を計算する。小領域内の点が推定した路面平面上の点であれば、 W_i と W'_i の相関は高くなり、そうでない場合には相関は低くなることから、小領域ごとに路面領域かどうかの判定を行う。しかし、推定した自車両の運動、路面平面の誤差により、実際には対応する小領域 W_i, W'_i には微少な位置ずれが生じる。一般的に路面では、白線部分で輝度が高く、それ以外の路面部分では輝度が低くなるため、白線とそれ以外の路面の両方を含む小領域で画素の輝度に基づいて相関を計算すると、

この微少なずれが大きく影響する可能性がある。そこで、領域内の輝度差が大きい場合には、領域内の輝度勾配が大きくなることから、小領域内で輝度勾配の平均を計算し、その値がしきい値以下なら輝度値に基づく SSD (Sum of Squared difference) による相関値 C_{SSD} 、しきい値以上なら微少なずれに対応するため、エッジ方向ヒストグラムを用いる SIFT (Scale Invariant Feature Transform) [11] の特徴による相関値 C_{SIFT} を計算し、相関値から小領域の路面領域判定を行う。 C_{SSD}, C_{SIFT} は以下の式の通り。

$$C_{SSD} = \sum_{(x,y) \in W_i} (I(x, y) - I'(x + u, y + v))^2 \quad (10)$$

$$C_{SIFT} = \sum_j (S_{W_i}(j) - S_{W'_i}(j))^2 \quad (11)$$

$I(x, y)$ は現時刻の画像の点 (x, y) の輝度値、 $I'(x', y')$ は一時刻前の画像の点 (x', y') の輝度値で、 $S_{W_i}(j)$ は小領域 W_i から計算される SIFT 特徴ベクトルの j 番目の要素を示す。

以上の方法により小領域ごとに路面領域判定を行った場合には、誤検出や未検出となる小領域も発生する。そこで、モルフォロジ演算のエロージョンとダイレーションを実行し、孤立した小領域の除去と抜けの小領域の補間を行う。

5 移動物体の検出

本章では、画像中の移動物体領域を検出する処理について説明する。まず、各時刻において推定した自車両の運動に従わない特徴点を検出する。しかし、各時刻で移動物体として検出された特徴点には、オプティカルフロー計算の誤った静止物体上の点も含まれる。またある時刻では、実際に移動物体上の点であっても、その動き方によって移動物体の点として検出されない場合も発生する。そこで、各特徴点を時間的に連続して追跡し、過去の時刻での判定結果も利用し、最終的な判定を行う。図 7(a) に最終的に移動物体の点と判定された特徴点とそのオプティカルフローを示す。

次に、移動物体の点を判定された特徴点を、その位置とオプティカルフローの大きさ、向きを基準としてグループ化し、図 7(b) に示すように各グ



(a) 移動物体点とそのフロー



(b) 移動物体領域の検出結果

図 7: 移動物体領域検出

ループに属する特徴点を含む矩形を移動物体領域として検出する。また、2点以上のグループとならなかった移動物体の特徴点は、誤検出として除去する。

移動物体上の特徴点は原理的に3次元位置を計算することができない。そこで、移動物体は路面上にあると仮定し、推定した路面平面を利用して移動物体領域の画像上での位置から距離を計算する。

6 実験

本章では、提案手法を実走行画像に適用した実験について述べる。実験では、車両に CCD カメラ SONY XC-55 を 1.28m の高さに俯角 -3 度で設置した。視野角は 32.6×24.8 度で、入力する画像の解像度は 640×480 画素、フレームレートは 10fps である。評価には、車両、二輪車、歩行者の移動物体を含む実走行画像 10 シーンを用いた。

図 8 に、先行車両、対向車両が存在するシーン 100 フレームにおいて、画像から検出した全特徴点

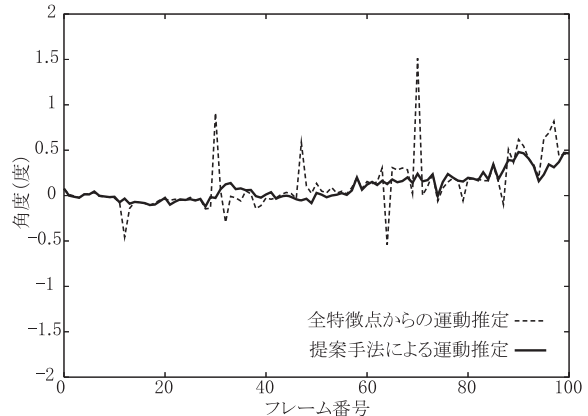


図 8: ヨー角の変化

を用いる運動推定法と提案手法により選択した特徴点を用いる運動推定法それぞれで推定したヨー角の変化を示す。全特徴点を用いる方法では、急激で大きな角度変化が頻繁に発生し、誤った推定となっていることが判断できるのに対し、提案手法による方法では、ほぼ安定して運動推定が行えていることが分かる。

図 9 に提案手法の検出結果例を示す。四角で囲んだ領域は移動物体領域、白く色づけした領域は路面領域の検出結果を示している。図 9(a) では先行車両が検出され、図 9(b) では、移動物体の誤検出も発生しているが、道路を横断する歩行者が検出できていることが確認できる。また、図 10 に復元した走行環境の 3 次元モデルの例を示す。図 10 から、走行環境の 3 次元構造が推定できており、また移動物体についても 3 次元位置が計算できていることが分かる。

7 まとめ

本稿では、車載単眼カメラにより自車両の運動を推定し、移動障害物を検出する方法を提案した。提案手法では、一時刻前の路面領域、移動物体領域の検出結果を利用し、計算に用いる特徴点を選択することで、走行環境において精度良く自車両の運動を推定でき、それにより路上を移動する障害物を検出可能である。実走行画像を用いた実験の結果、提案手法により単眼カメラから車両や歩行者などの物体が検出できることを確認した。今後の課題としては、静止障害物の検出や誤検出の低



(a) 車両検出例



(b) 歩行者検出例

図 9: 検出結果

減，領域検出の精度向上などが挙げられる．また，実時間処理のための処理量削減も課題である．

参考文献

- [1] R. Okada, Y. Taniguchi, K. Furukawa, and K. Onoguchi, "Obstacle detection using projective invariant and vanishing lines," Proc. Int. Conf. on Computer Vision, Vol. 1, pp. 330-337, 2003.
- [2] A. Shashua, Y. Gdalyahu, and G. Hayon, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," Proc. Intelligent Vehicle Symposium, pp. 1-6, 2004.
- [3] T. Suzuki and T. Kanade, "Measurement of vehicle motion and orientation using optical flow," Proc. Int. Conf. on ITS, pp. 25-30, 1999.
- [4] G.P. Stein, O. Mano, and A. Shashua, "A robust method for computing vehicle ego-motion," Proc. Intelligent Vehicles Symposium, 2000.



図 10: 復元した 3 次元モデル

- [5] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, "Visual modeling with a hand-held camera," Int. Journal of Computer Vision, Vol. 59, No. 3, pp. 207-232, 2004.
- [6] G. Lefaix, E. Marchand, and P. Bouthemy, "Motion-based obstacle detection and tracking for car driving assistance," Proc. Int. Conf. on Pattern Recognition, Vol. 4, pp. 74-77, 2002.
- [7] C. Harris and M. Stephens, "A combined corner and edge detector," Proc. Alvey Vision Conf., pp. 147-151, 1988.
- [8] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Int. Joint Conference on Artificial Intelligence, pp. 674-679, 1981.
- [9] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM, Vol. 24, No. 6, pp. 381-395, 1981.
- [10] R. Hartley, "In defense of the eight-point algorithm," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 6, pp. 580-593, 1997.
- [11] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.