

車載カメラ映像中の線分の 3 次元位置推定と走行可能領域検出への応用

内山 寛之[†] 出口 大輔[†] 高橋 友和^{††} 井手 一郎[†] 村瀬 洋[†]

y 名古屋大学大学院 情報科学研究科 〒 464-8601 愛知県名古屋市千種区不老町

yy 岐阜聖徳学園大学 経済情報学部 〒 500-8288 岐阜県岐阜市中鶯 1-38

E-mail: yfuchiyama,ddeguchi,ide,muraseg@murase.m.is.nagoya-u.ac.jp, yyttakahashi@gifu.shotoku.ac.jp

あらまし 走行可能領域検出は車両のナビゲーションや安全運転支援において重要である。市街地の構造物の多くはポールや壁面といった線分で構成されている。そこで本論文では、走行可能領域検出を目的とし、単眼カメラを用いて市街地の線分の 3 次元位置推定を行う手法を提案する。カメラの姿勢が既知の場合、線分の位置を推定するためには、画像間の線分の正確な対応付けが必要となる。そのため、少数の画像間で正しい線分の対応付けを行うために、多視点幾何拘束や市街地中の線分の特徴、車両の動きの事前知識を用いる。評価実験において、推定した線分の位置から走行可能領域の検出を行い、提案手法が走行可能領域検出に十分利用可能であることを確認した。

キーワード 走行可能領域検出, 3 次元復元, 車載カメラ

1. はじめに

走行可能領域検出 [1], [2] は、車両のナビゲーションや安全運転支援といった応用で重要である。本論文では、道路領域のような車両が通行可能な領域を走行可能領域と呼ぶ。衝突防止やレーンキープ等が走行可能領域検出により実現できる。従来、様々な走行可能領域検出手法が提案されている。多くの走行可能領域検出手法は、障害物がない領域の検出を行っている [2]。そのため、障害物を検出することは、走行可能領域検出において重要な要素技術となる。

環境中の物体の 3 次元位置を推定する手法として、レーザスキャナを用いたものがある [3]。しかし、現状では高価であり、一般車に搭載するには課題が多い。一方、車載カメラを利用することで環境中の物体の 3 次元位置推定が可能である [4], [5]。車載カメラは安価かつ小型であり、標識検出等、様々な用途に使用できるという利点がある。画像中の特徴点間の対応により、対象物体の 3 次元位置を推定する手法が提案されている [6], [7]。しかしながら、走行可能領域検出や衝突防止で重要となる細いポールのような物体は、テクスチャやコーナー等が乏しいため、安定した特徴点の検出が難しい。

一方、画像中の線分間の対応により、画像中の線分を 3 次元復元する手法が提案されている [8], [9]。市街地の例を図 1 に示す。線分は市街地の構造において主要な要素の 1 つであり、例えば、壁、ポール、ガードレール等は、多くが線分で構成されている。特に、鉛直の線分は、大部分の市街地の構造に含まれている。そのため、鉛直の線分を検出することは、走行可能領域検出において重要となる。市街地中の 3 次元線分を復元する手法として、Asai ら [10] はバンドル調整を用いたオフライン処理による手法を提案している。しかしながら、走行可能領域検出

のような走行中に行う処理を想定した場合には、多数の画像を利用するバンドル調整では問題がある。

本論文では、走行可能領域検出を目的とし、単眼カメラを用いて市街地映像中の線分の 3 次元位置推定を行う手法を提案する。鉛直な線分は、カメラの姿勢推定と画像間の線分の対応付けを基に 3 次元復元される。カメラの姿勢が既知の場合、正しく復元を行うためには、画像間の線分の対応付けが必要となる。また、走行可能領域検出のような走行中に行う処理を想定した場合には、少数の画像から線分の位置推定を行うことが望ましい。そのため、少数の画像間で正しい対応付けを行うために、多視点幾何拘束や市街地中の線分の特徴、車両の動きの事前知識を用いる。また、実験では、抽出した 3 次元線分が走行可能領域検出に十分利用可能であることを示す。

以降、2 節で線分の 3 次元位置推定手法について説明する。3 節で線分の 3 次元位置推定の実験と、検出した 3 次元線分を走行可能領域検出への適用した結果について述べる。最後に、4 節でまとめる。

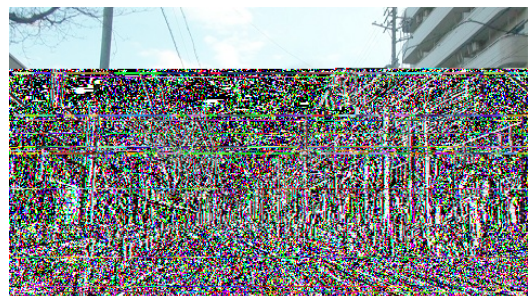


図 1 市街地の例。壁、ポール、ガードレールのように、市街地の構造の多くは線分を含む。

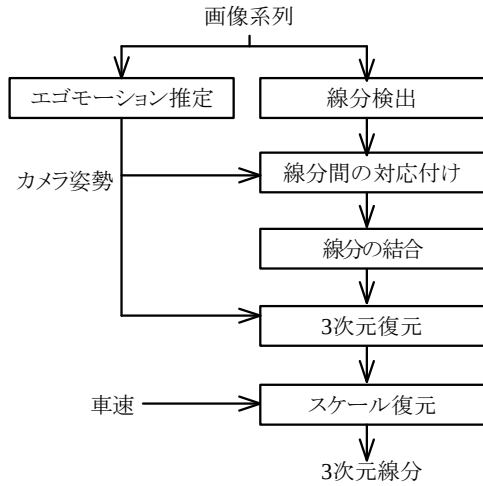


図2 線分の3次元位置推定手法の流れ.

2. 線分の3次元位置推定手法

2.1 手法の概要

本論文では、無限の長さをもつ細長い形状を直線と呼び、有限の長さで両端点をもつ形状を線分と呼ぶ。画像上の線分の幾何的な対応関係を求めるには、一般的に3視点以上が必要である。そのため提案手法では、線分の3次元位置推定は、車載カメラ映像から得た連続する3枚の画像により行う。正確な自車の位置が分かっているとすると、線分の3次元位置推定のためには、画像中の正確な線分の対応付けが問題となる。そこで、本研究では、線分間の対応付けに焦点を当てる。提案手法の方針は以下のとおりである。まず、線分の対応付けの候補を2視点幾何拘束を用いて探索する。ここで、2視点幾何拘束では対応が一意に定まらないという問題がある。そこで、さらに3視点幾何拘束を用いてさらに対応を絞り込む。より対応付け精度を高めるために、線分の色特徴やカメラの動きの事前知識を導入する。

提案手法は(1)エゴモーション推定、(2)線分検出、(3)線分間の対応付け、(4)線分の結合、(5)3次元復元の5つのステップで構成される。手法の流れを図2に示す。(1)のエゴモーション推定のステップでは、映像からカメラの姿勢を推定する。ここでは、画像間のSIFTキーポイント[13]の対応付けにより、各画像におけるカメラ行列と基礎行列を求める。(2)と(3)のステップでは、カメラの姿勢を用いて、画像中の線分の検出と対応付けを行う。ここで、背景と前景の状態によって、本来は1本の線分であるにもかかわらず、2本以上に分断されることがある。そこで、(4)では分断された線分の結合を行う。最後に(5)で、カメラの姿勢と線分の対応付けから、線分を3次元復元する。

本論文では、時刻 T_1, T_2, T_3 ($T_1 < T_2 < T_3$)で撮影された画像をそれぞれ画像1, 2, 3と表記する。また、図3のように、対応付いた線分の3つ組を $[s_1, s_2, s_3]$ と表

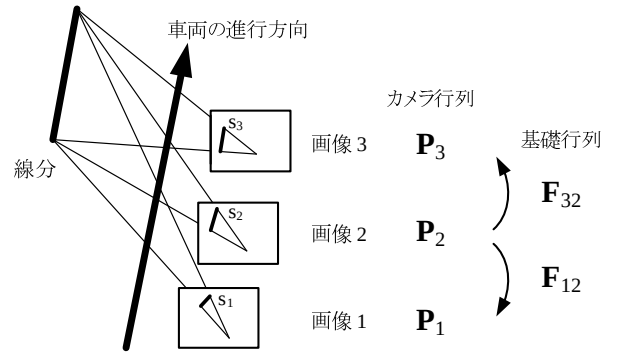


図3 視点間の幾何パラメータ.

記する。画像1, 2, 3のカメラ行列をそれぞれ P_1, P_2, P_3 とする。ここで、カメラ行列とは、カメラの姿勢を表す 3×4 の行列である。また、行列 F_{12}, F_{32} は、画像1と画像2, 3の間の基礎行列とする。

2.2 線分検出

画像から線分を検出するためにLSD (Line Segment Detector) [14]を用いる。LSDは輝度値の勾配を基に線分を検出する手法であり、線分の両端点の座標だけでなく、線分の幅も計測可能である。

2.3 2視点幾何拘束を用いた線分の対応付け

画像1, 2, 3中の対応する線分の3つ組 $[s_1, s_2, s_3]$ を得ることが、ステップ(2)の線分の対応付けの目的である。そのために、ある線分 s_2 に対応する線分 s_1, s_3 を探索する。ここではそれに先立ち、2視点幾何拘束とカメラの動きの事前知識により、 s_2 に対応する線分の組のおおまかな候補 $[s_1, s_2], [s_2, s_3]$ を探索する。

2.3.1 エピポーラ拘束による探索

線分 s_2 上の点 x と、基礎行列 F_{12} が与えられるとすると、画像1上の点はエピポーラ線 r' 上に存在する[11].

$$r' = F_{12}x \quad (1)$$

図4のように、対応する線分の候補はエピポーラ線と交差する。図4では、線分 s_2 上の任意の点 x に対応するエピポーラ線 r' を点線で、対応する線分の候補を太い実線で示している。線分 s_2 上のすべての点に対してエピポーラ線を計算し、交差する線分を見つけることで、対応付けの候補を探索する。画像1と同様に、画像3に対しても対応付けの候補を探索する。

2.3.2 カメラの動きの事前知識による探索

対応付け候補の誤りを削減するために、カメラの動きの事前知識を導入する。車両は前方に進むと仮定できる。さらに、画像上のエピポールはカメラの進行方向を示す。そのため、車両が進むにつれ、線分はエピポールから遠ざかる。そこで、エピポールと線分の重心の位置との距離が正しい順番となるような3つ組 $[s_1, s_2, s_3]$ を選択する。

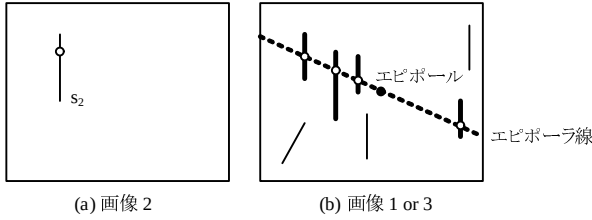


図 4 2 視点幾何拘束を用いた線分間の対応付け. 対応する線分の候補はエピポーラ線と交差する.

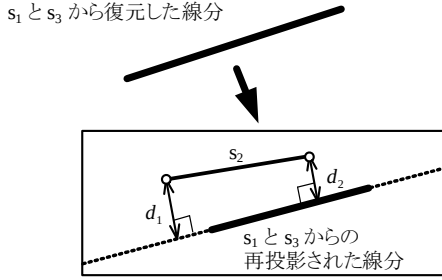


図 5 線分間の幾何的な距離の計算. 距離は再投影誤差により計算する.

2.4 3 視点幾何拘束を用いた線分間の対応付け

2 視点幾何拘束により対応付けの候補を探索した後, さらに 3 視点幾何拘束により絞り込みを行う. 次の 2 つの距離を対応付けに用いる.

- 幾何的な距離 d_g
- 色特徴による距離 d_c

市街地中には多くの鉛直の線分が存在するため, 誤った線分の 3 つ組が 3 視点幾何拘束を満たす場合がある. そのため, これらを除くために色特徴による距離を導入する. d_g , d_c がそれぞれしきい値 θ_g , θ_c 以下ならば, 3 つ組 $[s_1, s_2, s_3]$ を採択する.

2.4.1 幾何的な距離 d_g

線分の再投影誤差を幾何的な距離として用いる. まず, 3 次元空間中の線分を画像上の s_1 , s_3 を用いて復元する. 次にカメラ行列 $\mathbf{P}_2$ を用いて, 復元した線分を画像 2 に再投影する. このとき, 再投影した直線と画像 2 上の線分 s_2 との距離を, 次式で計算する.

$$d_g = d_1 + d_2, \quad (2)$$

ここで, d_1 , d_2 は再投影された直線と, 画像 2 上の線分 s_2 のそれぞれの端点との距離である (図 5).

π_1 , π_3 を画像 1, 2 上の線分とそれぞれのカメラ中心により構成される面とすると, 3 次元空間中の復元された線分は, π_1 と π_3 の交差する面として求められる.

$$\pi_1 = \mathbf{P}_1^T \mathbf{r}_1, \quad (3)$$

$$\pi_3 = \mathbf{P}_3^T \mathbf{r}_3, \quad (4)$$

ここで, $\mathbf{r}_1$, $\mathbf{r}_3$ は線分 s_1 , s_3 を示す.

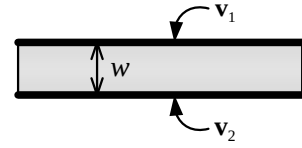


図 6 線分の色特徴. 色特徴ベクトル $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2]$ は, 線分領域の両側部分に沿った画素値を含む.

2.4.2 色特徴による距離 d_c

線分の色特徴を計算するために, 提案手法では線分の周囲の画素値を用いる. 線分は面のつなぎ目や背景と前景の境界に現れるが, 背景と前景の境界は視点や画像のぼけのため, 見た目が変化する. そのため, 色特徴を計算するために, これらの線分上の画素は除外する.

線分の画素を比較するために, F-guided matching [12] と以下で述べる線分の性質を用いる. F-guided matching は, 図 4 と同様に, 基礎行列から計算したエピポーラ線から, 線分上の画素間の対応関係を得る手法である. また, 線分は面のつなぎ目や背景と前景の境界に現れるため, 少なくとも線分のどちらかの側は, 視点の異なる画像中에서도同じ前景物体に含まれるという性質がある. そのため, 色特徴による距離を計算する際, 線分のどちらか一方の側のみを用いる.

色特徴による距離を計算するには, まず s_2 に関する色特徴ベクトル $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2]$ を作成する. ここで, 線分の幅 w は LSD により得られ, 図 6 のように, $\mathbf{v}$ は s_2 の両側の画素値を含む. F-guided matching を用い, 基礎行列 $\mathbf{F}_{12}$, $\mathbf{F}_{32}$ から s_1 , s_3 の色特徴ベクトル $\mathbf{v}' = [\mathbf{v}'_1, \mathbf{v}'_2]$, $\mathbf{v}'' = [\mathbf{v}''_1, \mathbf{v}''_2]$ を求める. $\mathbf{v}$, $\mathbf{v}'$, $\mathbf{v}''$ 間の色特徴による距離は次式により計算する.

$$d_c(\mathbf{v}, \mathbf{v}', \mathbf{v}'') = \min \{ k \mathbf{v}_1 - \mathbf{v}'_1 k_{L1} + k \mathbf{v}_1 - \mathbf{v}''_1 k_{L1}, \\ k \mathbf{v}_2 - \mathbf{v}'_2 k_{L1} + k \mathbf{v}_2 - \mathbf{v}''_2 k_{L1} \}, \quad (5)$$

ここで, $k \cdot k_{L1}$ はベクトルの L1 ノルムを表わす. また, d_c は $0 \leq d_c \leq 1$ の範囲に正規化する.

2.5 線分の結合

背景と前景の状態によって, 本来は 1 本の線分であるにもかかわらず, 2 本以上に分断されることがある. そこで, クラスタリングにより, 分断された線分の結合を行う. N 個の 3 つ組 $[s_1^{(1)}, s_2^{(1)}, s_3^{(1)}], \dots, [s_1^{(N)}, s_2^{(N)}, s_3^{(N)}]$ が線分間の対応付けにより得られているとする. 1 つの $s_1^{(n)}$ が複数の $s_2^{(n)}$ に対応づいている場合, これらの $s_2^{(n)}$ は同じ線分に属していると考えられることができる. そこで, ある 3 つ組の対の中に同じ線分を含むならば, その対は同じクラスタに統合する. 図 7 では, 3 つのクラスタに統合された例である. 最後に, クラスタリング結果を用い, 同じクラスタの線分を 1 つの線分に結合する.

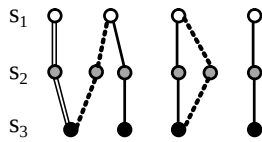


図 7 線分の結合のためのクラスタリング. ノードは画像中の線分を表し、エッジは対応する 3 つ組を示す. この例では、線分は 3 つのクラスに統合される.

表 1 提案手法と比較手法の条件

	比較手法 1	比較手法 2	比較手法 3	提案手法
幾何的な距離	なし	あり	あり	あり
色特徴による距離	線分周囲	なし	線分上	線分周囲

2.6 3 次元復元

対応する 3 つ組 $[s_1, s_2, s_3]$ から、線分を 3 次元復元する. 3 次元空間中の線分は、各画像のカメラ中心と画像上の線分のなす面の交わる線である. ここで、面同士のなす角が小さい場合、線分の位置推定精度は低下する. そのため、面 π_1, π_3 のなす角がしきい値 θ_a 以下の場合、それらの面により構成される線分を除外する. また、画像情報のみではスケールは復元できないため、3 次元復元した構造のスケールは車両の速度から取得する. 最後の、壁やポールなどの鉛直な線分は走行可能領域検出において特に重要なため、鉛直に近い線分のみを残し、他は除去する.

3. 実験と考察

提案手法の有効性を調査するために、市街地で撮影した映像を用いて実験を行った. まず、線分の 3 次元位置推定と線分の対応付けの精度の調査を行った. また、検出した 3 次元線分を用いて、走行可能領域検出を行った.

3.1 実験条件

線分の対応付け精度を評価するために、対応付けの失敗した割合を調査した. また、提案手法を次の 3 つの手法と比較した (表 1). 比較手法 1 は、幾何的な距離を使用しない. 比較手法 2 は、色特徴による距離を使用しない. また、比較手法 3 は、色特徴を線分上から抽出した (提案手法は色特徴を線分の周囲から抽出する). 比較手法 1, 2 のしきい値 θ_c, θ_g は、提案手法と同じとした. 比較手法 3 の色特徴のしきい値は、提案手法とほぼ同じ数の線分が検出される値に設定した.

エゴモーションは SIFT キーポイント [13] から推定した. SIFT キーポイントの画像間の対応から、基礎行列とカメラ行列を計算した.

映像は車両のフロントガラスに装着した民生用の HD カメラで撮影した. 元映像は $1,920 \times 1,080$ ピクセル、29.97 fps であるが、 960×540 ピクセル、6.0 fps に変換して使用した. 車両の速度は普及型 GPS で測定した. 10

表 2 10 シーン中の対応付いた線分の数

	比較手法 1	比較手法 2	比較手法 3	提案手法
対応数	503	705	354	379
誤り数	62	89	6	0
誤り率 [%]	12.3	12.6	1.7	0.0

個のシーンを評価に用いた. しきい値の値は $\theta_g = 3.0$ ピクセル, $\theta_c = 0.06$, $\theta_a = 0.3$ 度とした. カメラの内部パラメータは、あらかじめキャリブレーション [15] によって求め、映像の歪みは除去した.

3.2 線分の 3 次元位置推定の結果

線分の 3 次元位置推定の例を図 8 に示す. 図 8(a) は LSD の結果を示しており、画像 1, 2, 3 でそれぞれ 513, 505, 534 本の線分が検出された. 図 8(b) は対応付いた線分を示しており、39 本の線分が対応付いた. 図 8(c) は復元した 3 次元の線分を示している. 画像中の点は、推定されたカメラの位置を示しており、格子の間隔は 2.0 m である. その他の結果を図 9 に示す.

10 シーン中の対応付いた線分の数を表 2 に示す. 対応付け誤りは手で計数した. 提案手法は比較手法と比べ、最も対応付け誤りが少なかった.

復元した 3 次元空間中で図 8(b) 中の矢印間の長さを測定したところ、7.6 m であった. 実際の長さは 7.8 m であり、走行可能領域検出に用いるには十分な精度であると考えられる.

図 8(b) 中の遠くにある鉛直な線分が、図 8(c) で鉛直に復元されない事例が見られた. これは、映像を撮影する車両は直進するため、画像間の視差が小さいためである. これを解決する方法として、Manhattan world 仮説 [16] のような、市街地の構造の事前知識を導入する方法が挙げられる.

3.3 走行可能領域検出への応用

3.3.1 走行可能領域検出手法

前述のように、走行可能領域は道路領域のような車両が通行可能な領域である. 多くの手法は、走行可能領域は障害物がない領域であるという考えに基づいている [2]. そこで、本手法でもその方針に従い、検出した 3 次元線分を用いることで、単純に道路領域とその他の領域を区別する.

まず、路面に 3 次元線分を投影することで、障害物のマップを作成する (図 10(a)). このとき、頭上の標識などを検出しないために、車両の高さよりも高い場所にある線分は除去する. 次に、走行可能領域を求めるために、道路のそれぞれの側について、障害物の凸包を計算する (図 10(b)). 2 つの凸包の間の領域が走行可能領域である.

3.3.2 走行可能領域の検出結果

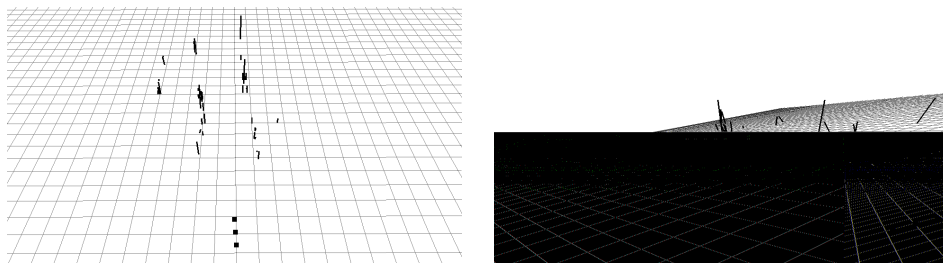
前回の実験と同じ映像に走行可能領域検出を適用した.



(a) 線分の検出結果 (画像 1, 2, 3)

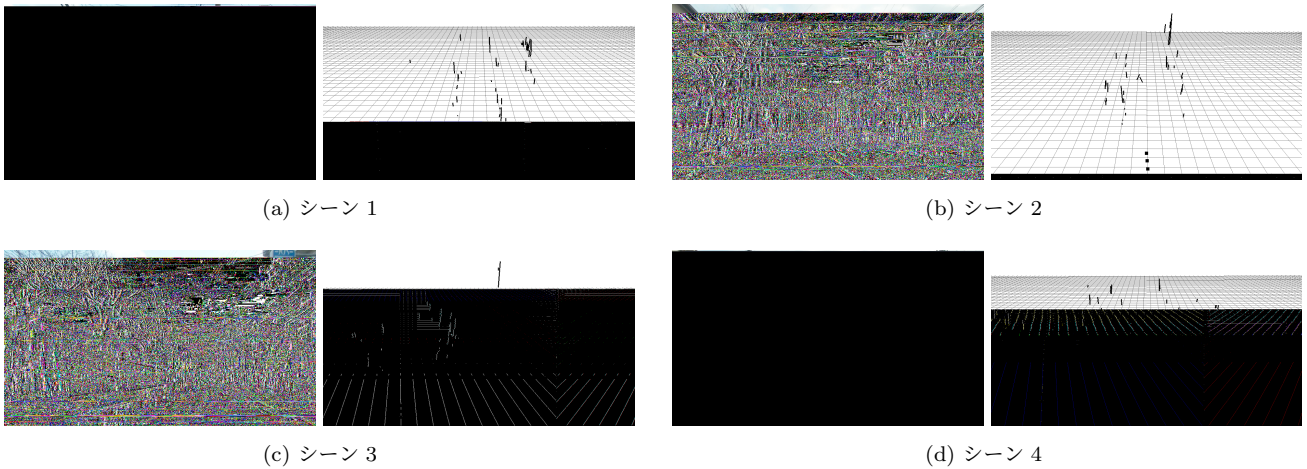


(b) 対応付け結果 (画像 1, 2, 3)



(c) 3次元復元結果

図 8 線分の 3 次元位置推定の例.



(a) シーン 1

(b) シーン 2

(c) シーン 3

(d) シーン 4

図 9 線分の対応付けと 3 次元位置推定の結果例.

実験結果の例を図 11 に示す. 左の図が障害物のマップであり, 右の図が凸包の一部分を入力画像に重畳表示したものである. 障害物のマップの格子の間隔は 5.0 m である. 右図の凸包は, 路側にあるポールや壁面などの物体により生成されており, 道路とその他の領域を分離していることがわかる.

4. ま と め

走行可能領域検出を目的とし, 単眼カメラを用いて市街地映像中の線分の 3 次元位置を推定する手法を提案し

た. 少数の画像間で正しい対応付けを行うために, 多視点幾何拘束や市街地中の線分の特徴, 車両の動きの事前知識を用いた. また, 実験では, 検出した 3 次元線分が走行可能領域検出に十分利用可能であることを示した.

今後の課題として, より精度を向上させるために, 市街地構造に関する事前知識の導入が挙げられる. また, より多様な環境で撮影した映像を用いて実験を行う予定である.

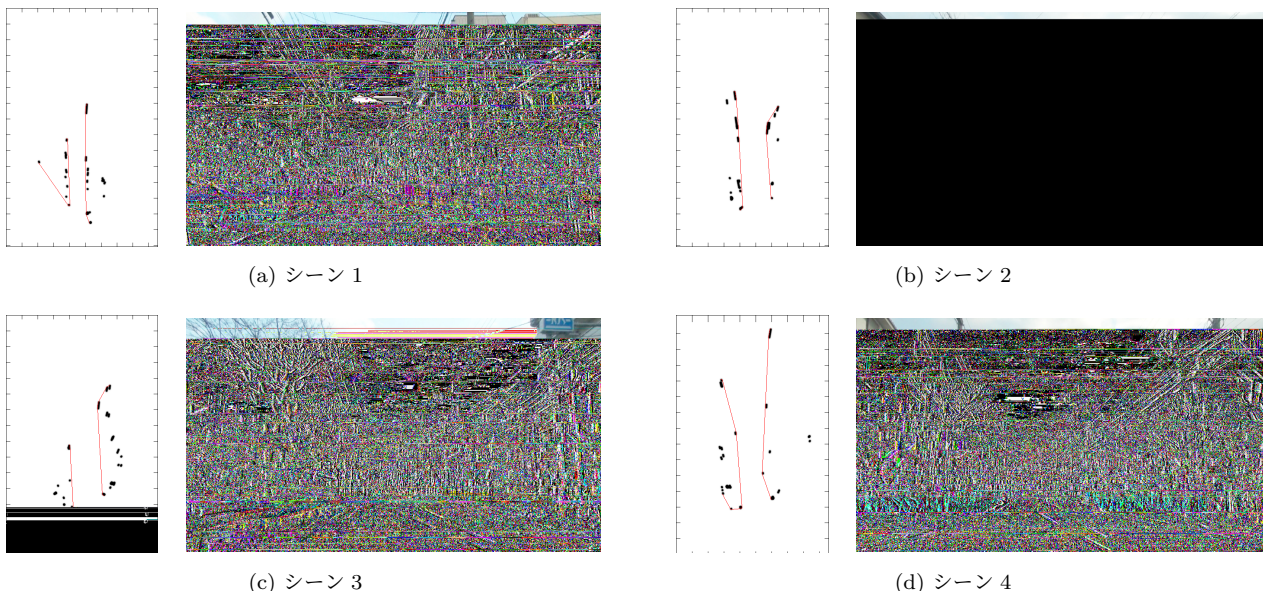


図 11 走行可能領域検出の適用例. 左の図が障害物のマップであり, 右の図が凸包の一部分を入力画像に重畳表示したものを示す. 障害物のマップの格子の間隔は 5.0 m である.

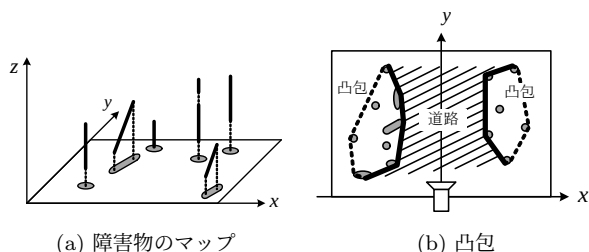


図 10 走行可能領域検出手法. (a) 路面に 3 次元線分を投影. (b) 凸包で道路とその他の領域を分離.

謝 辞

日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に深く感謝する. 本研究の一部は, JST 戦略的創造研究推進事業 CREST および科学研究費補助金による. また, 本研究では画像処理に MIST ライブラリ (<http://mist.murase.m.is.nagoya-u.ac.jp/>) を使用した.

文 献

- [1] H. Badino, U. Franke and R. Mester, “Free Space Computation using Stochastic Occupancy Grids and Dynamic Programming,” Proc. of ICCV2007 Dynamic Vision Workshop, xp.12, Oct. 2007.
- [2] Y. Matsumoto, K. Ikeda, M. Inaba and H. Inoue, “Exploration and Navigation in Corridor Environment Based on Omni-view Sequence,” Proc. of Int. Conf. on Intelligent Robots and Systems 2000, vol.2, pp.1505–1510, Nov. 2000.
- [3] H. Zhao and R. Shibasaki, “A Vehicle-borne Urban 3-D Acquisition System using Single-row Laser Range Scanners,” IEEE Trans. Systems, Man, and Cybernetics, vol.33, issue 4, pp.658–666, Aug. 2003.
- [4] W. Mark and D. M. Gavrila, “Real-time Dense Stereo for Intelligent Vehicles,” IEEE Trans. Intelligent Transportation Systems, vol.7, issue 1, pp.38–50, Mar. 2006.
- [5] K. Yamazaki, M. Tomono, T. Tsubouchi and S. Yuta, “Object Shape Reconstruction and Pose Estimation by a Camera Mounted on a Mobile Robot,” Proc. of Int. Conf. on Intelligent Robots and Systems 2004, vol.4 pp.4019–2025, Oct. 2004.
- [6] S. Ikeda, T. Sato, K. Yamaguchi and N. Yokoya, “Construction of Feature Landmark Database using Omnidirectional Videos and GPS Positions,” Proc. of Int. Conf. on 3-D Digital Imaging and Modeling 2007, pp.249–256, Aug. 2007.
- [7] I. Skrypnik and D. G. Lowe, “Scene Modelling, Recognition and Tracking with Invariant Image Features,” Proc. of Int. Symp. on Mixed and Augmented Reality 2004, pp.110–119, Nov. 2004.
- [8] C. J. Taylor and D. J. Kriegman, “Structure and Motion from Line Segments in Multiple Images,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.17, issue 11, pp.1021–1032, Nov. 1995.
- [9] D. Martinec and T. Pajdla, “Line Reconstruction from Many Perspective Images by Factorization,” Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition 2003, vol.1, pp.497–502, June. 2003.
- [10] T. Asai, K. Yamaguchi, Y. Kojima, T. Naito and Y. Ninomiya, “3D Line Reconstruction of a Road Environment using an In-vehicle Camera,” Proc. Int. Symp. on Visual Computing 2008, pp.897–904, Dec. 2008.
- [11] R. I. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision,” Cambridge University Press, ISBN: 0521623049, 2000.
- [12] C. Schmid and A. Zisserman, “Automatic Line Matching Across Views,” Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition 1997, pp.666–671, June 1997.
- [13] D. G. Lowe, “Distinctive Image Features from Scale-invariant Keypoints,” Int. J. of Computer Vision, vol.60, no.2, pp.91–110, Jan. 2004.
- [14] R. Gioi, J. Jakubowicz, J. Morel and G. Randall, “LSD: A Fast Line Segment Detector with a False De-

tection Control,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.32, issue 4, pp.722–732, Apr. 2010.

- [15] Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.22, issue 11, pp.1330–1334, Nov. 2000.
- [16] E. Delage, H. Lee and A. Y. Ng, “Automatic Single-image 3D Reconstructions of Indoor Manhattan World Scenes,” Robotics Research Results of the 12th Int. Symp. ISRR, Springer Tracts in Advanced Robotics, vol.28, pp.305–321, ISBN: 3540481109, 2007.