

## 都市環境における線分特徴の方向性分布を利用した SLAM

上 原 啓<sup>†</sup> 斎 藤 英 雄<sup>†</sup>(正会員) 原 孝 介<sup>††</sup><sup>†</sup>慶應義塾大学, <sup>††</sup>デンソーアイティラボラトリ

## Line-Based SLAM Considering Directional Distribution of Line Features in an Urban Environment

Kei UEHARA<sup>†</sup>, Hideo SAITO<sup>†</sup>(Member), Kosuke HARA<sup>††</sup><sup>†</sup>Keio University, <sup>††</sup>Denso IT Laboratory

〈あらまし〉 本論文では、線分特徴の事前分布を考慮した自動運転車のための SLAM (Simultaneous Localization and Mapping) を提案する。提案手法は、道路標示や建物など、線分を検出しやすい物体が存在する都市環境を走行する車両に搭載されたカメラにより撮影される動画像を入力とすることを前提とし、この動画像から検出される線分特徴の方向に関する制約を加えることにより SLAM の精度を向上させるものである。線分特徴の方向に関する制約を加えるため、本手法では道路標示中の線分が車両進行方向に対して平行、もしくは垂直であるものが多いことに着目し、それら 2 つの線分間の角度分布が混合ガウス分布に則ると仮定し、バンドル調整による最適化に導入する。これにより、カメラの位置姿勢と線分特徴の位置推定の精度を高めることが期待できる。本手法の有効性を確かめるため、カメラを 1 つ用いたシングルカメラシステムと、2 つ用いたマルチカメラシステムの両方に提案手法を適用し精度検証を行った。その結果マルチカメラによる従来手法の精度と、シングルカメラによる本手法の精度が同等となった。また提案手法により生成した路面地図は、線分の方向に関する制約を導入したことにより補正され、精度向上を確認した。

キーワード：線分 SLAM, マンタッハワールド仮説, 道路標示, ガウス分布, バンドル調整

<Summary> In this paper, we propose a line-based SLAM (Simultaneous Localization and Mapping) from an image sequence captured by a vehicle in consideration with the directional distribution of line features that detected in an urban environments. The proposed SLAM is based on line segments detected from objects in an urban environment, for example, road markings and buildings, that are too conspicuous to be detected. We use additional constraints regarding the line segments so that we can improve the accuracy of the SLAM. We assume that the angle of the vector of the line segments to the vehicle's direction of travel conform to four-component Gaussian mixture distribution. We define a new cost function considering the distribution and optimize the relative camera pose, position, and the 3D line segments by bundle adjustment. In addition, we make digital maps from the detected line segments. Our method increases the accuracy of localization and revises tilted lines in the digital maps. We implement our method to both the single-camera system and the multi-camera system. The accuracy of SLAM, which uses a single-camera system with our constraint, works just as well as a method that uses a multi-camera system without our constraint.

**Keywords:** line-based SLAM, manhattan world assumption, road markings, gaussian distribution, bundle adjustment

## 1. はじめに

今日、人の運転を必要としない完全自動運転に向けた研究が盛んに行われている。自動車が自律して運転を行うためには、車両自身が自己位置を推定し、地図上のどこを走ってい

るかを正確に把握する必要がある。Teramoto らの研究<sup>1)</sup>によると自動運転車の実用化に向けて、自己位置推定の誤差は数+センチメートル以内に抑える必要があるという。古くからある自己位置推定技術として GPS を利用したものがあり、RTK-GPS を用いた POSLV<sup>2)</sup>は誤差数十センチメートル以

下の精度を実現しているため、この要件を満たしている。しかし RTK-GPS はコストが高いことに加え、GPS 技術自体が電波の受信強度に大きく依存してしまうという問題がある。Google が開発する Google ドライバーレスカー<sup>3)</sup>はレーザースキャナによって様々な道路情報を認識し運転動作の最終決定をサポートしているが、レーザースキャナも同様にコストの高さが課題となっている。

これらの問題を解決する手段として、今日車載カメラを用いた Simultaneous Localization and Mapping (SLAM) 技術が盛んに研究されている。SLAM は画像のみをインプットとして、画像中の特徴量を用いて自己位置と周辺環境の推定を行う手法であり、特徴量が十分に検出される環境下において高い精度を誇る一方、特徴量が少ない環境下において精度が低下してしまうため、それらを補う様々な研究がなされている。Kawasaki ら<sup>4)</sup>は Pless によるカメラデザインとその精度の研究<sup>5)</sup>に基づいて非重複のマルチカメラを用いることによりステレオベースラインを長くとり SLAM の精度向上を実現している。また Koletschka ら<sup>6)</sup>は点と線をハイブリッドした SLAM によりテクスチャーの少ない環境下での SLAM を実現した。しかしながら、カメラ台数を増やすことはコストの増加を、点と線を用いることによる特徴量の増加は計算コストの増加を意味し、リアルタイム性が求められる車両のための SLAM には不向きである。

そこで本論文では、都市環境で検出された線分特徴量の方向性に関する事前分布を導入した線分 SLAM を提案する。道路標示を構成する線分の多くが車両進行方向に対して平行、もしくは垂直であることに着目し、車両進行方向ベクトルと線分ベクトルがなす角度の分布を混合ガウス分布で近似する。この分布を線分特徴地図の事前分布としてバンドル調整に導入し、最適化を行うことにより精度の向上を試みる。Zhou ら<sup>7)</sup>は、都心において建造物の面はお互いに平行または垂直、というマンハッタンワールド仮説<sup>8)</sup>に基づいた制約を加え、SLAM の向上を実現したが、都市環境を前提とした SLAM では、マンハッタンワールド仮説が成り立ちにくいカーブシーンに弱いという欠点があり誤差が多く含まれる。しかし線分特徴の方向に関する事前分布という新たな制約を加えることにより、マンハッタンワールド仮説が成立しない状況にも対応することができる。

また道路標示を用いた SLAM 研究が Wu ら<sup>9)</sup>によってなされている。Wu ら<sup>9)</sup>は道路標示を点の集合と捉え事前に学習することにより SLAM の向上を実現した。しかし筆者らの知る限り、道路標示の線分を検出し、その事前分布をガウス分布で近似し最適化を考慮した研究は未だ行われていない。そこで、本手法により従来よりも高精度の SLAM を実現できるかどうかを確認した。

本論文では、2 章において数式の表記を始めとする表記法の説明を行い、3 章で提案手法である線分のマッチング手法と、バンドル調整による車両運動の最適化について説明を行

う。4 章で実験結果を示し考察を行い、5 章でまとめと今後の課題を述べる。

## 2. 数式表記

本章では、本論文で用いる数式の表記法について説明する。マルチカメラシステムは 4 つの座標系  $W$ ,  $C_1^t$ ,  $C_2^t$ ,  $V^t$  によって成り立つ。 $W$  はワールド座標系、 $C_1^t$  は  $t$  フレーム目における後方カメラ座標系、 $C_2^t$  は  $t$  フレーム目における前方カメラ座標系、 $V^t$  は  $t$  フレーム目における車両の中心を原点とする車両座標系である。 $W$  から  $V^t$  への位置姿勢を  $\mathbf{R}_{vw}^t$ ,  $\mathbf{T}_{vw}^t$  と表記し、 $V^t$  から  $C_1^t$ ,  $C_2^t$  への位置姿勢をそれぞれ  $\mathbf{R}_{c_1v}$ ,  $\mathbf{T}_{c_1v}$ ,  $\mathbf{R}_{c_2v}$ ,  $\mathbf{T}_{c_2v}$  と表記する。なお後方カメラと前方カメラのキャリブレーションは事前に行っているものとする。これらの数式を用いると  $W$  の 1 点  $\mathbf{p} = (x, y, z)^T$  とそれに対応する  $C_k^t$  の 1 点  $\mathbf{q} = (q_x, q_y, q_z)^T$  の関係式は式 (1) のように表される。

$$\mathbf{q}_k = \begin{pmatrix} q_x \\ q_y \\ q_z \end{pmatrix} = \mathbf{R}_{c_k v} (\mathbf{R}_{vw}^t \mathbf{p} + \mathbf{T}_{vw}^t) + \mathbf{T}_{c_k v} \quad (1)$$

なお  $k$  は後方カメラ、もしくは前方カメラを表すものとする。また  $\mathbf{q}$  とそれに対応する画像上の 1 点  $\mathbf{u} = (u, v)^T$  の関係は式 (2) のように表される。

$$\mathbf{u}_k = \begin{pmatrix} u \\ v \end{pmatrix} = \pi \begin{pmatrix} q_x \\ q_y \\ q_z \end{pmatrix} = \begin{pmatrix} q_x/q_z \\ q_y/q_z \end{pmatrix} \quad (2)$$

続いて、 $W$  における線分  $\mathbf{L}_k$  は 6 次元ベクトル  $(\mathbf{p}_k^T, \mathbf{r}_k^T)^T$  で表され、それぞれ線分の中点  $\mathbf{p}_k$  と線分方向ベクトル  $\mathbf{r}_k$  を表す。また  $\mathbf{L}_k$  の画像上への投影線分  $\mathbf{l}_k$  を 4 次元ベクトル  $(\mathbf{u}_k^T, \mathbf{d}_k^T)^T$  と表す。 $\mathbf{u}$  と  $\mathbf{p}$  の関係は式 (1) と式 (2) と同様であり、 $\mathbf{r}_k$  とそれに対応する  $\mathbf{d}$  は式 (3) で表される。

$$\mathbf{D}_k = \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \mathbf{R}_{c_k v} \mathbf{R}_{vw}^t \mathbf{r}_k \quad (3)$$

$$\mathbf{d}_k = \begin{pmatrix} d_x \\ d_y \end{pmatrix} = \begin{pmatrix} q_z D_x - q_x D_z \\ q_z D_y - q_y D_z \end{pmatrix} \quad (4)$$

なお  $\mathbf{D}_k$  は  $\mathbf{r}_k$  に対応するカメラ座標系の 1 点となる。

最後に、本手法においてカメラの位置姿勢の初期値として用いられるオドメーターによる値は式 (5) のように表現される。

$$\mathbf{x}^{t+1} = \begin{pmatrix} x^{t+1} \\ z^{t+1} \\ \theta^{t+1} \end{pmatrix} = \mathbf{x}^t + \Delta \mathbf{x}^t + \varepsilon_{\mathbf{x}}^t \quad (5)$$

$\Delta \mathbf{x}^t$  は前フレームからの移動距離、 $\varepsilon_{\mathbf{x}}^t$  は  $\Delta \mathbf{x}^t$  のノイズを表す。

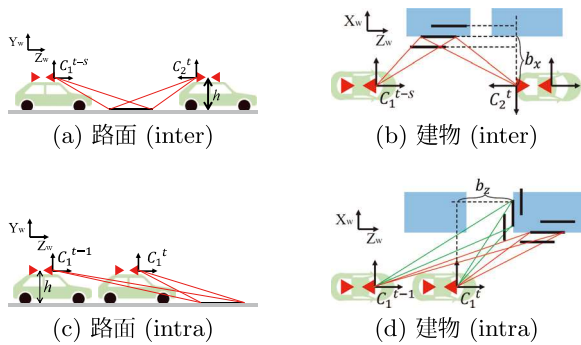


図1 別時刻画像マッチングと別カメラ画像マッチングの例  
Fig. 1 Examples of intra- and inter-camera correspondences

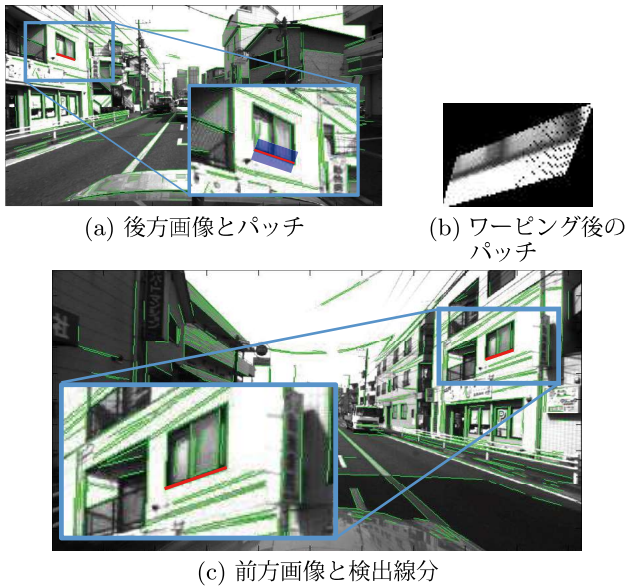


図2 ワーピングの結果  
Fig. 2 Result of Warping

### 3. 提案手法

本提案手法は大きく分けて、ワーピング処理を用いた線分マッチング部と、線分角度を考慮した最適化部の2つからなる。前者では、Line Segment Detector(LSD)を用いて画像群から検出した線分を、ワーピング処理を伴うマッチング手法により対応付け、マンハッタンワールド仮説に基づきそれら対応線分の3次元座標の初期値を推定する。また後者では、オドメーターから得られるカメラの位置姿勢を初期値として用いて、線分角度を考慮したバンドル調整によりそれら初期値の最適化を行う。以下、それぞれの処理に関して具体的な説明を行う。なお、マルチカメラシステムには前方方向と後方方向を向いた2台の非重複カメラを用いる Kawasaki ら<sup>4)</sup>による手法を採用した。

#### 3.1 線分マッチング

本節では、オドメトリを用いたワーピングによる線分マッチング手法を説明する。ワーピングとは、画像内の線分を含むパッチを画像変換し、対象画像の見え方に揃える処理を指す。

す。非重複カメラでは前方画像と後方画像に写された同一特徴量の見え方が大きく異なるため、ワーピングによる前処理を行うことによってマッチング精度が大幅に向上する。ワーピングは、マンハッタンワールド仮説に基づいた手法でパッチの3次元座標を算出し、オドメトリから取得した対象フレームのカメラ位置姿勢を用いて画像変換することにより可能となる。本手法では線分の法線方向に20ピクセルの幅を持つパッチを用いる。またマッチングは、前方画像、後方画像間のマッチングと、同一カメラ画像間のマッチングの2つに大別され、Lee ら<sup>10)</sup>の表記に基づき、それぞれを別カメラ画像マッチング (Inter-camera correspondence), 別時刻画像マッチング (Intra-camera correspondence) と呼ぶことにする。

##### 3.1.1 別カメラ画像マッチング

別カメラ画像マッチングは前方画像と後方画像間のマッチングを指す。本手法では後方画像中の線分を含んだパッチを前方画像中にワーピングし、見え方を揃える事を考える。図1(a)(b)より別カメラ画像マッチングでは線分が検出される場所が、路面上、もしくは建物の正面に限られる。それぞれの場合における線分の3次元座標の算出方法を次に示す。

路面上から線分が検出される場合、マンハッタンワールド仮説に基づき、全ての線分の  $y$  座標はカメラが設置される高さを表す  $h$  となる (図1(a))。ここでパッチの中心座標を  $\mathbf{u}_{m,2} = (u_{m,2}, v_{m,2})^T$  とするとパッチの画像座標は、中心座標からのユークリッド距離  $\alpha, \beta$  を用いて  $\mathbf{u}_{m,2} = (u_{m,2} + \alpha, v_{m,2} + \beta)^T$  と表せる。 $h$  は既知であるため、式2よりパッチ内の各点の後方カメラ座標系  $C_1^t$  における3次元点  $\mathbf{q}_r$  は式(6)で表現される。

$$\mathbf{q}_r = \begin{pmatrix} h(u_{m,2} + \alpha)/(v_{m,2} + \beta) \\ h \\ h/(v_{m,2} + \beta) \end{pmatrix} \quad (6)$$

次に建物の正面から線分が検出される場合を考える。同様にマンハッタンワールド仮説に基づいて考えると、線分の  $x$  座標は建物までの距離を表した  $b_x$  となる (図1(b))。よって式(2)よりパッチ内の各点の後方カメラ座標系  $C_1^t$  における3次元点  $\mathbf{q}_{bf}$  は式(7)で表現される。

$$\mathbf{q}_{bf} = \begin{pmatrix} b_x \\ b_x(v_{m,2} + \beta)/(u_{m,2} + \alpha) \\ b_x/(u_{m,2} + \alpha) \end{pmatrix} \quad (7)$$

以上の方法で算出したカメラ座標系の3次元点を、オドメトリから取得したカメラの位置姿勢  $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t$  を用いて世界座標系  $W$  の3次元点  $\mathbf{p}$  に変換し、ワーピング対象フレームのカメラの位置姿勢  $\mathbf{R}_{vw}^{t-s}, \mathbf{T}_{vw}^{t-s}$  を用いて投影先に色情報をマッピングすることにより、パッチを Warp することができる。本手法によるワーピングの結果を図2に示す。図2(b)に示された図2(a)内のパッチをワーピングしたものと、

図 2(c) に示されたワーピング対象の前方画像の見え方が一致しているのを見てとれる。

ワーピングを施した後方画像中のパッチと、前方画像中のパッチのマッチングは、Davison<sup>11)</sup>による拡張カルマンフィルタに基づく誤差楕円を用いて行い、異なるカメラ、フレーム間で撮影された線分の対応付けを行う。

しかしながら以上のマンハッタンワールド仮説に基づいた 3 次元座標算出法には 2 点の不確定要素がある。1 点目は画像中から検出された線分が建物から検出された線分か、路面上から検出された線分かが定かでない事、2 点目は建物までの距離  $b_x$  が定かでないことである。これらの不確定要素を解消するために、各線分に対して建物の正面から検出される場合、並びに路面から検出される場合の 2 つの場合に関してマッチングの計算を行う。その際  $b_x$  を 2m から 15m の探索範囲で 0.5m 間隔で変化させ、最も誤差楕円によるマッチングスコアの高かった条件をその線分における検出場所、 $b_x$  として採用する。なお探索範囲は線分検出が可能と思われる建物の範囲から決定した。

次に拡張カルマンフィルタに基づく誤差楕円<sup>11)</sup>を利用してワーピングを施した後方画像中のパッチと、前方画像中のパッチをマッチングさせる手法について説明する。誤差楕円内でワーピングしたパッチをラスタスキャンし、正規化相互相関によるマッチングスコアが一番高い座標を求める。このときの線分の座標と、誤差楕円に含まれる前方画像上の線分との比較を行い、ラスタスキャンしたパッチの線分の端点と前方画像上の線分との射影距離と角度の誤差の値が閾値より小さく、射影距離が一番小さいものを対応線分として採用する。

### 3.1.2 別時刻画像マッチング

別時刻画像マッチングは後方画像同士、或いは、前方画像同士の線分マッチングを指す。別時刻画像マッチングでは、路面、建物の正面に加えて建物の側面が線分の検出場所として考えられる (図 1(d))。建物の正面、路面に存在する場合は別カメラ画像マッチングと同様のため、本節では建物の側面にのみ焦点をあて説明を行う。図 1(d) より線分の  $z$  座標は  $b_z$  となるため、式 (2) よりパッチ内の各点の後方カメラ座標系  $C_1^t$  における 3 次元点  $\mathbf{q}_{bs}$  は次の式 (8) で表現される。

$$\mathbf{p}_{bs} = \begin{pmatrix} b_z(u_{m,1} + \alpha) \\ b_z(v_{m,1} + \beta) \\ b_z \end{pmatrix} \quad (8)$$

Inter-camera correspondence と同様に  $b_z$  を 2m から 15m の範囲で 0.5m 間隔で変化させ、各線分ごとの最適な  $b_z$  を決定し、誤差楕円によるマッチングを行う。

### 3.2 初期値推定

取得した線分の 3 次元位置の最適化を行うため、それらの初期値を推定する。前節で述べた通り、マンハッタンワールド仮説に基づき、検出線分が 3 つの平面にのみ存在することを利用すると、マッチングで得られた情報から、対応線分は

建物の正面・側面、路面のいずれかに存在していると分類される。よって式 (6)～式 (8) を用いて各線分の 3 次元座標を取得することができる。なお、ここで 1 つの対応につき、対応付けられた線分の数だけ 3 次元座標が得られるという問題が生じるため、最小メジアン法を用いてその中で最も正確な値を決定する。各 3 次元座標を対応線分の画像平面に再投影し、検出線分との再投影誤差の中央値が最も小さい線分の 3 次元座標を初期値として採用する。線分における再投影誤差とは、再投影した 2 次元直線の端点から検出線分への射影距離を指す。

またカメラの位置姿勢 ( $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t$ ) の初期値はオドメータによる値を用いる。

### 3.3 バンドル調整による最適化

カメラの位置姿勢 ( $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t$ ) と線分の 3 次元座標  $\mathbf{L}^j$  をバンドル調整で最適化することを考える。本手法では、最適化のためのコスト関数を、線分の再投影誤差項  $\mathbf{e}_l$  と角度誤差項  $\mathbf{e}_\theta$  からなる次の式 (9) で定義する。

$$E = \mathbf{e}_l^2 + \mathbf{e}_\theta^2 = \frac{1}{2} \sum_i \sum_{n=1}^2 d_\perp^2(\mathbf{g}_n^i, \mathbf{l}^i) + \frac{1}{2\sigma^2} \sum_m (\mathbf{e}_\theta^m)^2 \quad (9)$$

$d_\perp(\mathbf{a}, \mathbf{b})$  は点  $\mathbf{a}$  から線  $\mathbf{b}$  への射影距離を、 $\mathbf{g}_1^i, \mathbf{g}_2^i$  は画像中の線分の端点座標を、 $\mathbf{e}_\theta$  は車両進行方向と線分方向から算出される角度誤差を表している。再投影誤差項は  $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t, \mathbf{L}^j$  に依存し、角度誤差項は  $\mathbf{L}^j$  に依存するため、 $E$  を最小化することは、 $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t, \mathbf{L}^j$  を最適化することと等価となる。Madsen<sup>12)</sup>らによる非線形レーベンバーグ・マーカート法を用いた最適化法を各項に分けて説明する。

線分マッチングによって対応付けられた同一線分を表すセット  $\Omega$  を式 (10) のように表す。

$$\Omega = \{\omega_i = (t, k, j, p) | \\ t \in \{1, \dots, T\}, k \in \{1, 2\}, j \in \{1, \dots, J\}, p \in \{1, 2, 3\}\} \quad (10)$$

$i$  番目の線分はフレーム  $t$ 、場所  $p$  でカメラ  $k$  によって検出されたことを表す。

#### 3.3.1 再投影誤差項

線分の端点  $\mathbf{g}_n^i$  から再投影線分  $\mathbf{l}^i$  への射影距離  $d_\perp(\mathbf{g}_n^i, \mathbf{l}^i)$  を式 (11) より算出する。

$$\mathbf{e}_l^i = d_\perp(\mathbf{g}_n^i, \mathbf{l}^i) = \frac{d_y(g_x^i - u^i) - d_x(g_y^i - v^i)}{\sqrt{d_x^2 + d_y^2}} \quad (11)$$

なお、 $\mathbf{u}_k = (u, v)^T$  は  $\mathbf{l}^i$  の中点座標、 $\mathbf{d}_k = (d_x, d_y)^T$  は  $\mathbf{l}^i$  の方向ベクトルである。

式 (1)、式 (2)、式 (11) から再投影誤差項  $\mathbf{e}_l$  は  $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t, \mathbf{L}^j$  から成り立つことがわかる。バンドル調整による最適化を行うために、式 (11) を  $\mathbf{R}_{vw}^t, \mathbf{T}_{vw}^t, \mathbf{L}^j$  で微分し、ヤコビ行列を作成する。次に、微分のチェーンルールに基づいた微分式を示す。



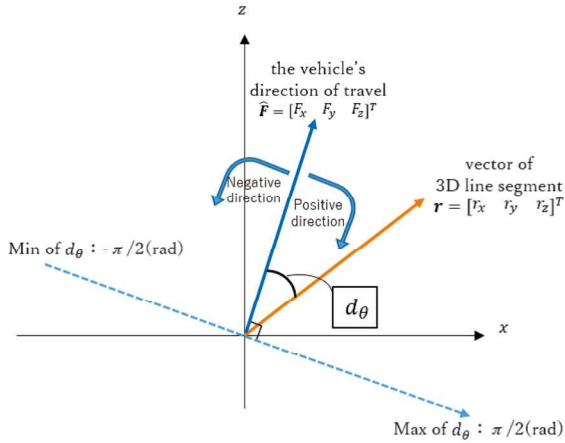


図3  $d_\theta$  の概略図  
Fig. 3 A schematic diagram of  $d_\theta$

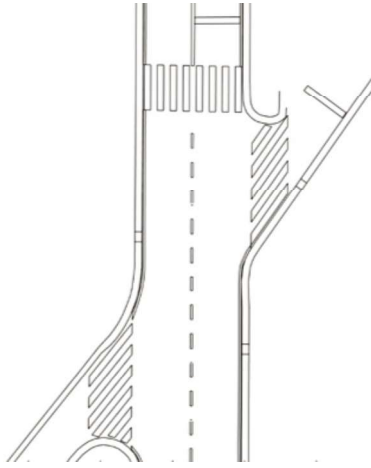


図4 新川崎の地図の例  
Fig. 4 Example of road maps in Shin-Kawasaki city

$$\frac{\partial e_l^i}{\partial \mathbf{R}_{vw}^t} = \frac{\partial e_l^i}{\partial \mathbf{q}^i} \frac{\partial \mathbf{q}^i}{\partial \mathbf{R}_{vw}^t} \quad (12)$$

$$\frac{\partial e_l^i}{\partial \mathbf{T}_{vw}^t} = \frac{\partial e_l^i}{\partial \mathbf{q}^i} \frac{\partial \mathbf{q}^i}{\partial \mathbf{T}_{vw}^t} \quad (13)$$

$$\frac{\partial e_l^i}{\partial \mathbf{p}^j} = \frac{\partial e_l^i}{\partial \mathbf{q}^i} \frac{\partial \mathbf{q}^i}{\partial \mathbf{p}^j} \quad (14)$$

$$\frac{\partial e_l^i}{\partial \mathbf{r}^j} = \frac{\partial e_l^i}{\partial \mathbf{D}^i} \frac{\partial \mathbf{D}^i}{\partial \mathbf{r}^j} \quad (15)$$

$\mathbf{q}$  と  $\mathbf{D}$  は、それぞれ式 (1) と式 (4) で定義したとおりである。

またマンハッタンワールド仮説に基づき、路面で検出された線分の  $y$  座標を全て  $h$  に固定する制約をコスト関数に施した。

### 3.3.2 角度誤差項

本手法における SLAM は都市環境を前提としているが、実際には正確にそれを満たしておらず推定に誤差が含まれてしまう。特にカーブシーンでは道路方向に対して建物が平行、または垂直に立っていないため、誤差量は顕著である。そこで最適化の際に、異なった制約を加えることにより、それらに起因する誤差を軽減することを考える。1 章で述べた通り

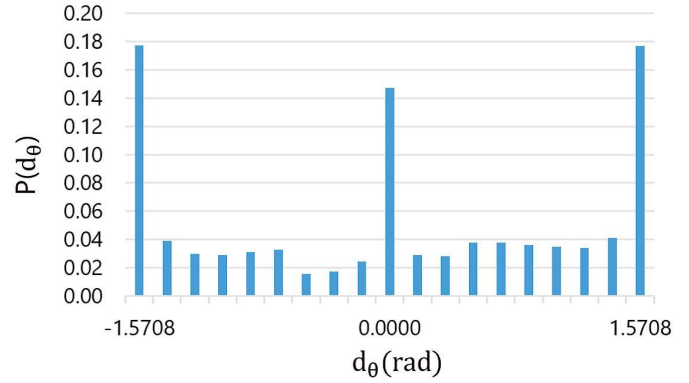


図5 新川崎の地図の  $d_\theta$  のヒストグラム  
Fig. 5 Histogram of  $d_\theta$  in Shin-Kawasaki

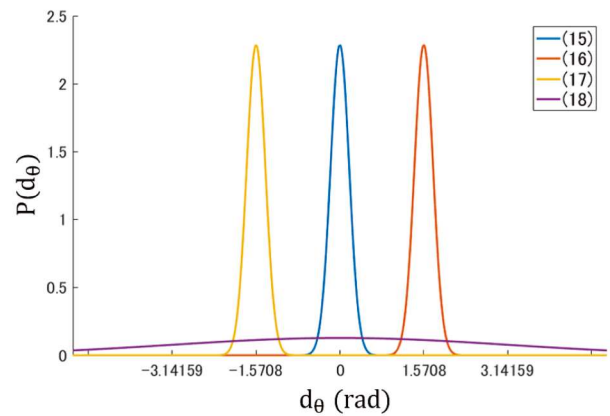


図6 4つからなる混合ガウス分布  
Fig. 6 A four-component Gaussian mixture distribution

本手法では、道路標示を構成する線分は車両の進行方向に対して平行、もしくは直交しているものが多くを占めていることに着目する。図3が示すように、車両進行方向ベクトルに対する路面上の線分の角度を  $d_\theta (-\frac{\pi}{2} \leq d_\theta \leq \frac{\pi}{2})$  とおき、新川崎の地図 (図4) を構成する各線分の  $d_\theta$  を算出し、図5のヒストグラムを作成した。この図からも道路標示を構成する線分が、車両進行方向に対して平行もしくは直交しているものが多いことがわかる。ここで本来、車両進行方向に対して平行、もしくは垂直である線分が含まうる誤差がガウス分布に則ると仮定すると、図5のヒストグラムは4つのガウス分布を用いて図6のような分布に近似できる。 $d_\theta = 0$  を中心とする突起したガウス分布は車両進行方向に平行な線分の存在確率が高いことを、 $d_\theta = \pm \frac{\pi}{2}$  を中心とするガウス分布は垂直な線分の存在確率が高いことを意味する。また平均を0とし、分散が広くとられたガウス分布は斜めの線分が存在することを表している。以下にそれぞれのガウス分布の式を示す。

$$P_1(d_\theta^m | \mu_1 = 0, \sigma_1 = \sigma_\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} (d_\theta^m)^2\right) \quad (16)$$

$$P_2(d_\theta^m | \mu_2 = \frac{\pi}{2}, \sigma_2 = \sigma_\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \left(d_\theta^m - \frac{\pi}{2}\right)^2\right) \quad (17)$$

$$P_3(d_\theta^m | \mu_3 = -\frac{\pi}{2}, \sigma_3 = \sigma_\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \left(d_\theta^m + \frac{\pi}{2}\right)^2\right) \quad (18)$$

$$P_4(d_\theta^m | \mu_4 = 0, \sigma_4 = \sigma_\beta) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left(-\frac{1}{2\sigma_\beta^2} (d_\theta^m)^2\right) \quad (19)$$

なお図6中の凡例番号がそれぞれの式番号に対応している。また車両進行方向ベクトルを  $\mathbf{F} = (F_x, F_y, F_z)$  とすると  $d_\theta$  は式(20)により算出される。

$$d_\theta = \arctan \frac{F_z}{F_x} - \arctan \frac{r_z}{r_x} \quad (20)$$

これらの式に基づき、角度誤差項におけるコスト関数を式(21)のように定義する。

$$e_\theta^m = \frac{d_\theta^m - \mu_n}{\sigma_n} \quad (21)$$

$n$  は4つのガウス分布のうちのいずれかを表し、各線分に対して  $P_n(d_\theta)$  の最大値をとる  $n$  を採用する。例えば、 $P_2(d_\theta)$  が他の  $P$  よりも大きいとき、 $\mu$  の値は  $\frac{\pi}{2}$  となる。つまり  $e_\theta^m$  は各線分の方角に関する誤差を表しており、このコスト関数を最小化することは、本来平行、もしくは垂直であるはずの線分が誤差により傾いたとき、それが正確に補正されることを意味している。また  $P_4(d_\theta)$  が他の  $P$  よりも大きいとき、つまり線分が本来の性質として傾いている場合、 $\sigma_4 = \sigma_\beta (> \sigma_\alpha)$  となり、式(21)よりコストが小さくなるため、車両進行方向に対して平行、もしくは直交に補正されにくくなる。

再投影誤差項と同様にコスト関数のヤコビ行列を作成することにより最適化を行う。次に角度誤差項による微分式を示す。

$$\frac{\partial e_\theta^j}{\partial \mathbf{r}} = \left( \frac{\partial e_\theta^j}{\partial r_x} \quad \frac{\partial e_\theta^j}{\partial r_y} \quad \frac{\partial e_\theta^j}{\partial r_z} \right) \quad (22)$$

$$\frac{\partial e_\theta^j}{\partial r_x} = \frac{r_z}{r_x^2 + r_z^2} \quad (23)$$

$$\frac{\partial e_\theta^j}{\partial r_y} = 0 \quad (24)$$

$$\frac{\partial e_\theta^j}{\partial r_z} = -\frac{r_x}{r_x^2 + r_z^2} \quad (25)$$

また  $\sigma_\alpha$  と  $\sigma_\beta$  の値を一定間隔で変化させていくことにより、各実験条件毎に最も適した  $\sigma_\alpha$  と  $\sigma_\beta$  を決定する。

## 4. 評価実験

本章では提案手法を評価するために行った実験について示す。本実験では自己位置推定と路面地図生成の精度をそれぞれ定量的に検証する。実験に用いたデータセットは、前方を向いた前方カメラと後方を向いた後方カメラ、また高い精度で自己位置を測定する RTK-GPS を実車に搭載し、新川崎周辺を走行することにより準備した。フレームレート 10fps で撮影した 72 フレームからなる直線シーン (dataset 1) と 200 フレームからなるカーブシーン (dataset 2) の2つに本手法を適用した。また本手法の有効性を検証するため、次の4つの実験条件を用意した。

- シングルカメラによる線のみを用いた SLAM
- マルチカメラによる線のみを用いた SLAM
- シングルカメラによる点・線を用いた SLAM
- マルチカメラによる点・線を用いた SLAM

上記の実験条件をそれぞれ line (S), line (M), point-line (S), point-line (M) と呼ぶことにする。

またプログラムは以下の環境で実行した。

- プロセッサ：Intel Core i7-6800K 3.40GHz
- RAM：64.0GB
- 実装環境：Matlab

### 4.1 マッチング結果

提案手法であるワーピングを用いたマッチング手法による結果と LEHF<sup>13)</sup>によるマッチング結果を比較したものを図7に示す。図7において、(a)～(d)は提案手法によるマッチング結果を、(g)～(j)は LEHF によるマッチング結果を、別カメラ画像マッチングの場合について示している。また (e), (f) は提案手法によるマッチング結果を、(k), (l) は LEHF によるマッチング結果を、別時刻画像マッチングの場合について示している。図中の赤線はインライアを、緑線はアウトライアを示している。別時刻画像マッチングにおいては、大きな見え方の違いがないため2つの手法間で有意な差はみられないが、別カメラ画像マッチングにおいては、提案手法による結果でインライアの割合が高いことが明らかであり、ワーピング手法の有効性を確認できる。

### 4.2 自己位置推定

2種類のデータセットに対して、それぞれ4つの軌跡を描画し比較を行った。1つ目は提案手法によって得られた運動推定結果、2つ目は従来手法(川崎らの方法<sup>4)</sup>を線分に拡張した方法)による運動推定結果、3つ目はオドメータによる軌跡、4つ目は RTK-GPS による軌跡である。今回、RTK-GPS によるデータを正解データとして用いた。

図8にそれぞれのデータセットにおける line(S) の運動推定結果を示す。また図9と図10にそれぞれのフレームにおける正解データとの距離誤差を示した。本手法における距離

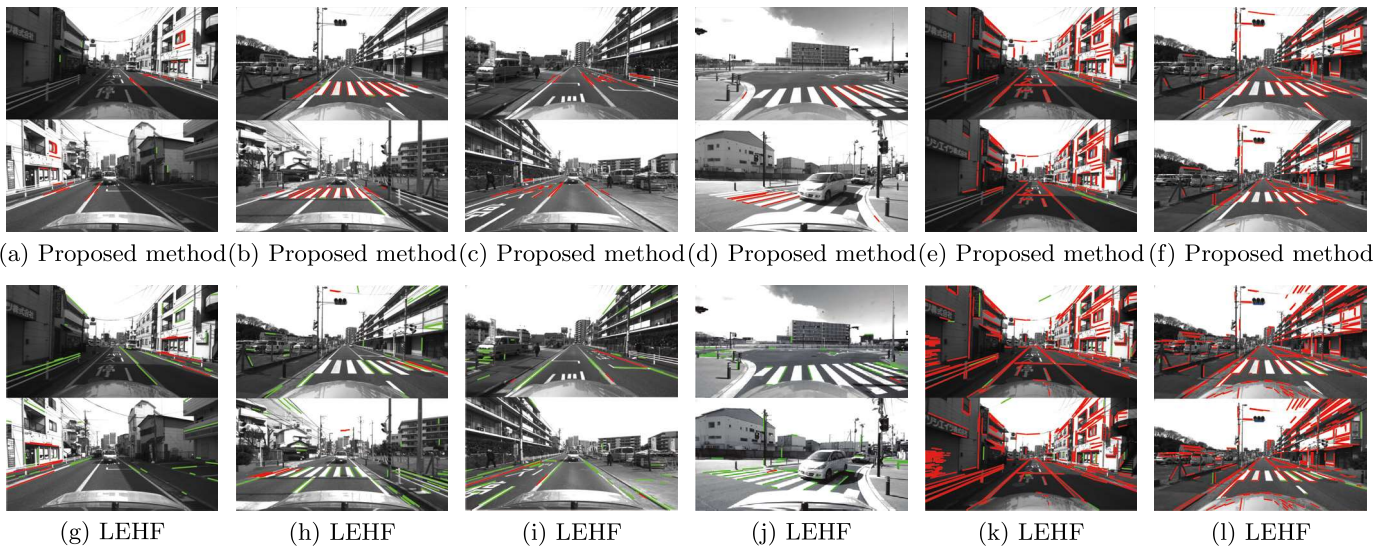


図7 提案手法とLEHFによるマッチング結果

Fig. 7 Matching result using proposed method and LEHF

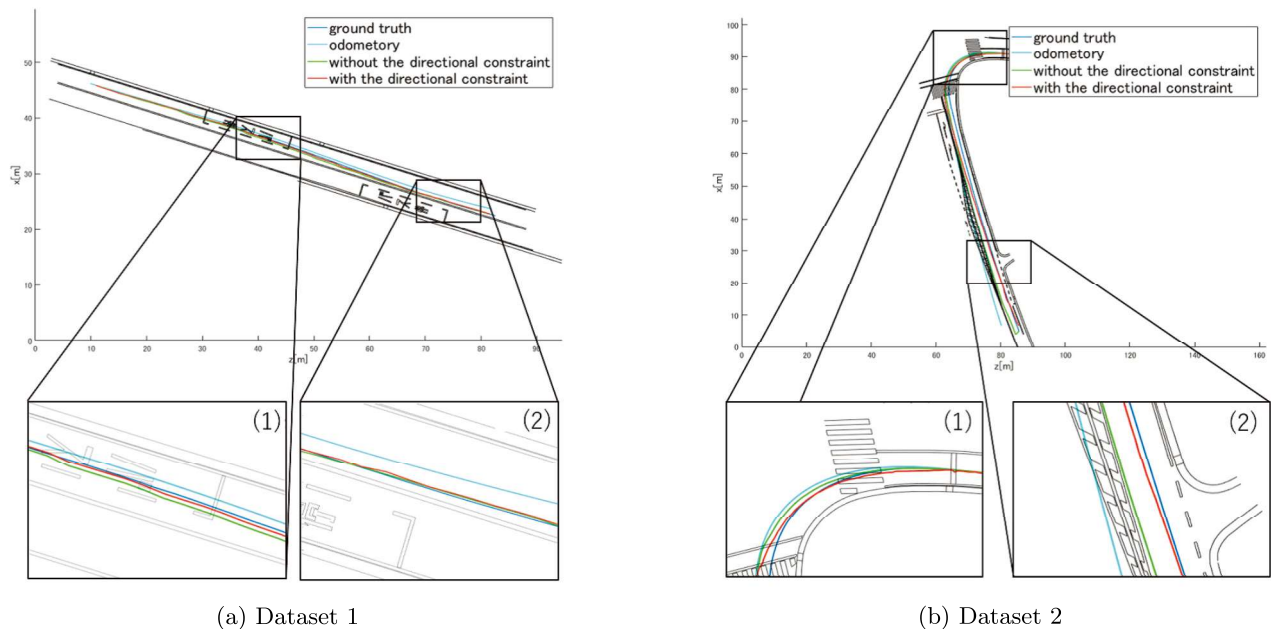


図8 運動軌跡推定の結果 (line (S))

Fig. 8 Results of trajectory estimation in the experiment of line (S)

誤差とは、各フレームにおける車両推定位置を正解データの軌跡に射影した距離を指す。これはGPSやオドメータの値を取得するタイミングとカメラの撮影するタイミングのわずかなずれを無視するためである。また、表1には全てのフレームの距離誤差を合計したものの比較を示した。

はじめに直線シーンに着目する。図8(a)の(1)を見ると、オドメトリや従来手法よりも提案手法の方が推定結果が正解データに近く、改善されているのが見て取れる。一方、(2)では従来手法よりも結果が悪くなっており、図9(a)にもそれが表れている。これは該当エリアにおいて本来斜めになっている線分が直線に補正されてしまったことによる誤差と考えられる。図9(c)にも同様の結果が表れており、(b)(d)に

は表れていない。(b)(d)はカメラを2台用いているため、対応線分が十分に検出され斜めの線が直線に補正されないという正確な結果が算出されたのであろう。表1より各実験条件における距離誤差の合計を見てみると、4つすべての条件において提案手法によって誤差が改善されていることがわかる。特にline (S)、line-point (S)において50%前後の改善が見られ、提案手法におけるline(S)の結果が、従来手法におけるline(M)の結果に、line-point(S)の結果が、従来手法におけるline-point(M)に近づいていることがわかる。

同様にカーブシーンに着目すると、図8、図10、表1から、全ての条件において誤差が大きく改善されているのが見て取れる。線分方向性分布を考慮しない従来法では、「マンハッ

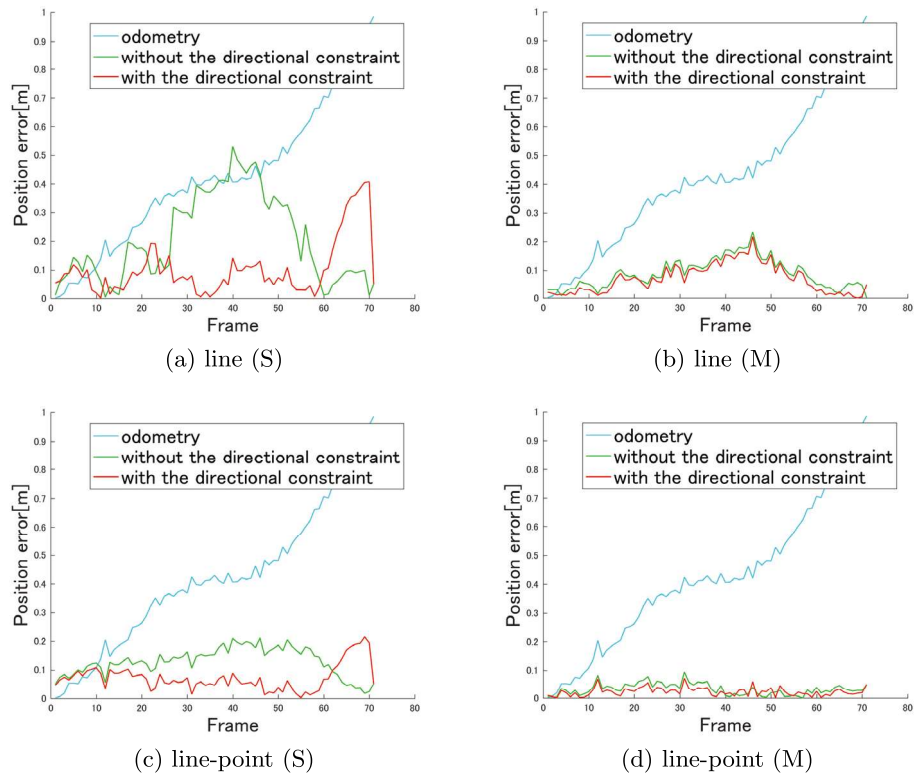


図 9 各フレームにおける距離誤差 (dataset 1)  
Fig. 9 Comparison of position error in each frame (dataset 1)

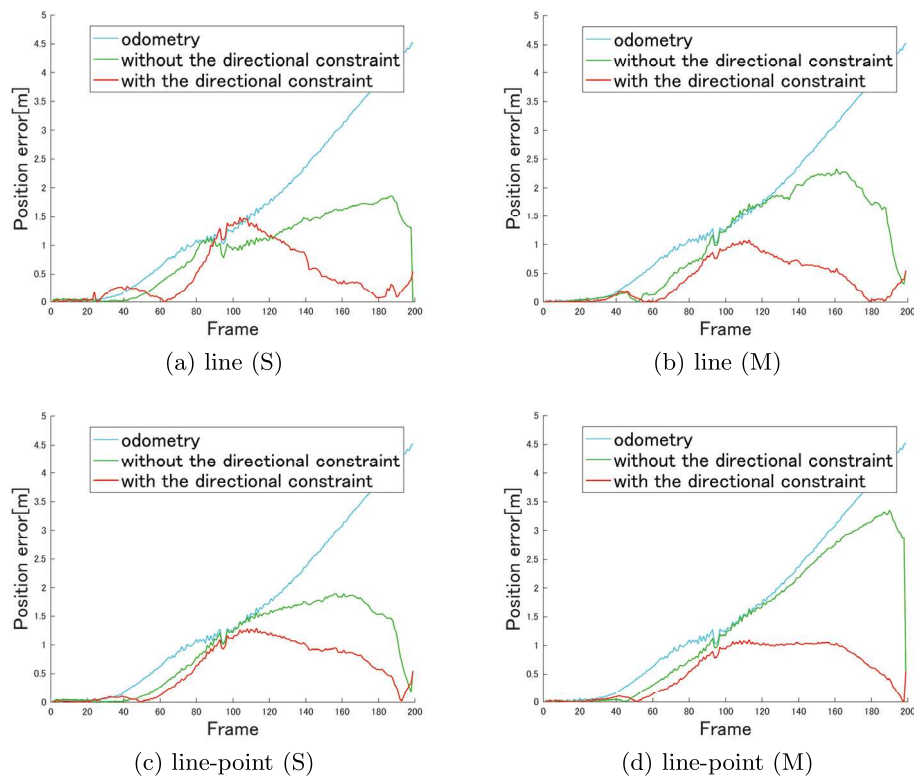


図 10 各フレームにおける距離誤差 (dataset 2)  
Fig. 10 Comparison of position error in each frame (dataset 2)

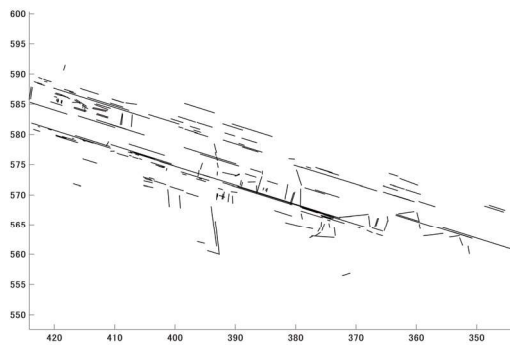
タンワールド仮説」, つまり, 都心における建造物の面は互いに平行, または垂直であるという前提に基づくので, それが成立しにくいカーブシーンではその前提による制約によ

て誤差を生んでしまう可能性があった。しかし, 提案手法による結果から, 道路標示中の線分特徴量の方向性に関する事前分布を導入し新たな制約を設けたことにより, この誤差を

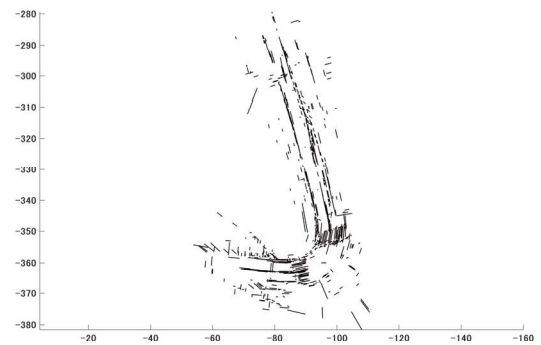


表 1 距離誤差の合計と、従来手法と比較した改善率  
Table 1 Sum of the position error and the improvement rate

	Dataset 1			Dataset 2		
	Without the constraint [m]	With the constraint [m]	Improvement rate [%]	Without the constraint [m]	With the constraint [m]	Improvement rate [%]
line (S)	13.57	6.52	52.0	176.11	96.31	45.3
line (M)	6.22	5.54	10.9	205.56	76.50	62.8
line-point (S)	8.45	4.32	48.9	186.18	114.05	38.7
line-point (M)	2.62	1.58	39.7	267.90	109.17	59.3



(a) Dataset 1



(b) Dataset 2

図 11 検出した線分から作成した路面地図  
Fig. 11 Digital maps made from detected line segments

表 2 インライアの割合

Table 2 The rate of inlier lines

	Dataset 1		Dataset 2	
	Without the constraint [%]	With the constraint [%]	Without the constraint [%]	With the constraint [%]
line (S)	28.5	47.5	40.1	43.7
line (M)	35.1	41.6	39.6	44.0
line-point (S)	39.1	49.7	41.5	43.8
line-point (M)	49.6	52.8	42.2	45.2

軽減することができていることを示している。

### 4.3 路面地図生成

図 11 に本手法で検出された線分から生成した路面地図を示す。定量的な評価を施すため、業者が作成した地図を正解データとして比較を行いインライアの割合を算出し、表 2 に示した。この値が高いほど、生成した地図中に誤検出した無駄な線分が少なく、正確な地図が復元できていることを示していると言える。なおインライアは検出線分の端点から、業者による地図中の線分への射影距離が 100mm 未満のものとした。表 2 を見てみると、全ての条件下でインライアの割合が増えており、特に dataset 1 が示す直線シーンにおいて高い改善が見られた。図 12 は、路面地図中の線分をインライアとアウトライアで色分けし描画したものであり、(a) は線分の方角性分布を考慮しない従来法による結果を、(b) は方角性分布を考慮した提案手法による結果を示す。図 12(a) 中の (1), (2) と図 12(b) 中の (1), (2) を比較すると、従来法中のアウトライアが提案手法によりインライアに変わっている

のがわかる。線分の方角に関する事前分布を考慮したことにより傾いた線分が補正されたためだと考えられる。Dataset 2 が示すカーブシーンでは、車両の進行方向と道路標示線分の関係性が仮定した分布からずれてしまうため、3%ほどの改善に留まったと考えられる。

### 4.4 実行時間

表 3 に Dataset1 における各実験条件のプログラム実行時間を示す。全ての実験条件において実行時間の減少に成功していることがわかる。これは最適化に新たな制約を加えたことにより、最適解に至るまでの反復回数が減少したためと考えられる。

また今回、3fps のカメラを用いたため 0.33sec/frame 以内の実行時間であれば、リアルタイム性を確保できているといえる。表 3 から、本手法により line (S) においてリアルタイム性の確保に成功し、line (M) においても同等の結果を得られていることがわかる。



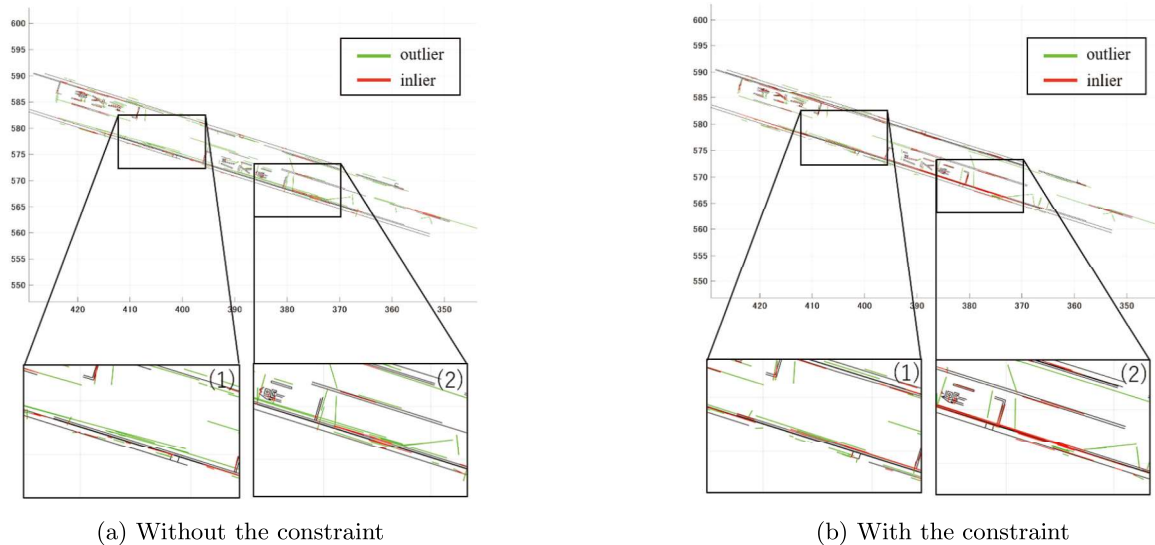


図 12 路面地図におけるインライア、アウトライアの描画結果 (line (S))  
**Fig. 12** Mapping results describing inlier and outlier lines in the experiment of line (S)

表 3 Dataset1 におけるプログラム実行時間

**Table 3** Execution time of the program of Dataset 1

	Without the constraint [sec/frame]	With the constraint [sec/frame]	Improvement rate[%]
line (S)	0.36	0.18	50.0
line (M)	0.50	0.35	30.0
line-point (S)	1.13	0.59	47.8
line-point (M)	1.53	1.26	17.6

## 5. 結 論

本論文では、従来手法である都市の3次元構造を利用した線分 SLAM に、道路標示線分の方角に関する事前分布を加える手法を提案した。車両の進行方向に対する道路標示線分の角度が混合ガウス分布に則ると仮定し、その分布に従うようバンドル調整に新たな制約を設けた。実車による評価実験によりシングルカメラ、マルチカメラ双方において本手法による線分 SLAM の有効性を実証した。直進シーンでは、従来手法でマルチカメラで実現していた運動推定精度を、提案手法によりシングルカメラで実現することに成功した。またカーブシーンにおいては、全実験条件下において自己位置推定誤差を改善させることに成功した。また検出した線分から路面地図を生成し、従来手法で傾いていた線分を、提案手法で方向を補正することによりインライアの割合を向上させることに成功した。さらには、最適化に制約を加えたことによりプログラム実行時間が低減でき、一部の実験条件においてリアルタイム性の確保に成功した。

今後の課題として、より精度の高い事前分布を考慮することがあげられる。今回、線分の方角分布をガウス分布に近似したが、業者による地図から学習し、より現実に近い事前分布を導入することができれば、より精度の高い SLAM が実

現できると考えられる。

## 参考文献

- 1) E. Teramoto, Y. Kojima, J. Meguro, N. Suzuki: "Development of The 'Precise' Automotive Integrated Positioning System and High-accuracy Digital Map Generation", R&D Review of Toyota CRDL, Vol. 43, No. 1, pp.13-23 (2012).
- 2) Applanix Corporation, Posion and Orientation Systems for Land Vehicles (POSLV), <https://www.applanix.com/downloads/products/brochures/LandBrochure.2017-web.pdf> (2017).
- 3) The New York Times, Google Cars Drive Themselves in Traffic, <http://www.nytimes.com/2010/10/10/science/10google.html?mcubz=0> (2010).
- 4) A. Kawasaki, H. Saito, K. Hara: "Motion Estimation for Non-Overlapping Cameras by Improvement of Feature Points Matching Based on Urban 3D Structure", Proc. of IEEE International Conference on Image Processing(ICIP), pp.1230-1234 (2015).
- 5) Robert Pless: "Using many Cameras as One", Proc. of IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Vol. 2, pp.587-593 (2003).
- 6) T. Koletschka, L. Puig, K. Daniilidis: "Mevo: Multi-Environment Stereo Visual Odometry Using Points and Lines", IEEE Intelligent Robots and Systems(IROS), pp.4981-4988 (2014).
- 7) H. Zhou, D. Zhou, L. Pei, R. Ying, P. Liu, W. Yu: "Structslam: Visual Slam with Building Structure Lines", IEEE Trans. on

- Vehicular Technology, pp.1364–1375 (2015).
- 8) James M Coughlan, Alan L Yuille: “Manhattan World: Compass Direction from a Single Image by Bayesian Inference.”, Proc. of the Seventh IEEE International Conference, Vol. 2, pp.941–947 (1999).
  - 9) T. Wu, A. Ranganathan: “Vehicle Localization Using Road Markings”, Proc. of IEEE Intelligent Vehicles Symposium(IV), pp.1185–1190 (2013).
  - 10) Gim Hee Lee, Friedrich Faundorfer, Marc Pollefeys: “Motion Estimation for Self-driving Cars with a Generalized Camera”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2746–2753 (2013).
  - 11) A. Davison, I. Reid, N. Molton, O. Stasse: “MonoSLAM: Real-Time Single Camera Slam”, IEEE Pattern Analysis and Machine Intelligence, pp.1052–1067 (2007).
  - 12) K. Madsen, H. Bruun, O. Tingleff: “Methods for Non-Linear Least Squares Problems”, Informatics and Mathematical Modelling, Technical University of Denmark (2004).
  - 13) K. Hirose, H. Saito: “Fast Line Description for Line-based SLAM”, Proc. of British Machine Vision Conference(BMVC), pp. 1–11 (2012).

(2017 年 2 月 24 日 受付)

(2017 年 6 月 19 日 再受付)



#### 上 原 啓

2016 年 慶應義塾大学情報工学科卒業。現在、同大学大学院理工学研究科前期博士課程に在学中。主にコンピュータビジョンに関する研究に従事。



#### 斎 藤 英 雄 (正会員)

1992 年 慶大院理工学研究科電気工学専攻博士課程修了。同年 慶大助手。その後専任講師、助教授を経て、2006 年 慶大理工情報工学科教授。この間、1997 年-1999 年まで、学術振興会海外特別研究員として、米国カーネギーメロン大学ロボティクス研究所に滞在し、主に Virtualized Reality の研究に従事。2000 年-2003 年 JST さきがけ研究「情報と知」領域研究員兼務。2006 年-2012 年 JST CREST 研究代表者。現在、主としてコンピュータビジョンとその VR 応用に関する研究等に従事。博士(工学)。電子情報通信学会、日本バーチャルリアリティ学会フェロー。情報処理学会、IEEE シニア会員。映像情報メディア学会、計測自動制御学会会員。



#### 原 孝 介

2002 年 東京農工大学工学研究科修了。同年 横河電機(株)入社。2004 年(株)デンソーアイティラボラトリ入社。2015 年 慶應義塾大学博士後期課程在学中。現在、主としてコンピュータビジョンに関する研究に従事。