

既走行経路における画像データベースを用いた 自車位置・姿勢推定

糸 秀行^{1,2,a)} Arne Suppé² 金出 武雄²

概要：駐車や通勤などの自動運転車の応用では、多くの場合、車両は走行毎に同じ経路を通ると考えられる。本稿では、事前に経路上を手動で走行することにより取得した動画画像からオフライン処理で作成されたデータベースを基に自動車の位置・姿勢を推定する手法を提案する。提案手法では、まず、Topometric Localization により、トポロジカルな情報とメトリックな情報を考慮しながら現在の画像に対応するデータベース画像を同定する。次に、オフライン処理において推定された特徴点の三次元位置と現在の画像上の特徴点の二次元位置の対応関係から自動車の位置・姿勢を推定する。実験では、屋内駐車場で撮影された動画画像を用いて、自動車の位置・姿勢を推定した結果を示す。

1. はじめに

近年、自動運転車の実現を目指した研究が多く成されており、公道での走行実験や市販車への機能搭載などが行われている。自動運転車には、通勤や旅行など様々なアプリケーションが考えられるが、中でも我々は、QOL (Quality of Life: 生活の質) の向上を目的とし、Valet Parking の自動化を目指す Virtual Valet プロジェクトを行なっている [1]。ここで、Valet Parking とは、レストランやホテルの入り口で車を乗り捨て、代わりに Valet と呼ばれる従業員に駐車場に駐車してもらい、北米では一般的なサービスである。Virtual Valet プロジェクトでは、これを自動化することにより、レストランなどの商業施設に加え、一般家庭においてもサービスを利用可能とすることで、車椅子利用者や重い荷物を持った人を助け、QOL の向上を図る。

自動運転車の実現には、パスプランニングや障害物検出など様々な要素技術が必要とされるが、本稿では、これらの中でも最も基本的な問題の一つである、自動車の現在の位置・姿勢を推定する Localization を扱う。現在までに様々な Localization 手法が提案されており、簡単かつ低コストで利用可能な GPS は多くの手法で用いられている。しかし、GPS には、ビルや樹木の近くや屋内など GPS 電波が遮蔽される環境では利用できないという問題

がある。また、レーザーレンジファインダなど三次元情報を計測可能な 3D センサを用いた手法が提案されており、DARPA (米国防高等研究計画局) 主催の自動運転車レース Grand Challenge や Urban Challenge で高い成果をあげている [2], [3]。しかし、3D センサは高価であり、また内部に高速に稼働する部品を持つため、耐久性が低いという問題がある。

本稿では、Valet Parking の自動化を実現するため、屋内でも使用可能かつ 3D センサと比較して安価なカメラを用いた自動車の Localization 手法を提案する。提案手法は、Valet Parking ではお店や家の入り口から駐車スペースまたはその逆というように、走行毎に同じ経路を通ることに着目し、事前に手動で経路上を走行することにより取得した画像データベースを基に Localization を行う。

2. 事前撮影された画像データベースに基づく 位置・姿勢推定

2.1 概要

本稿では、自動車に取り付けられたカメラを用い、事前に手動で経路上を走行することにより取得した画像データベースを基に Localization を行う手法を提案する。ただし、提案手法では、カメラの内部パラメータおよびカメラと自動車の間の相対位置・姿勢は事前のキャリブレーションにより既知とし、カメラの位置・姿勢を推定することで、自動車の位置・姿勢を推定する。図 1 に提案手法の処理とデータの流れを示す。図 1 に示すように、提案手法はオフライン処理とオンライン処理に分けられる。以下では、各処理の概要を述べ、それぞれの処理の詳細については、3

¹ 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute of
Science and Technology

² カーネギーメロン大学 ロボティクス研究所
The Robotics Institute, Carnegie Mellon University

a) hideyuki-k@is.naist.jp

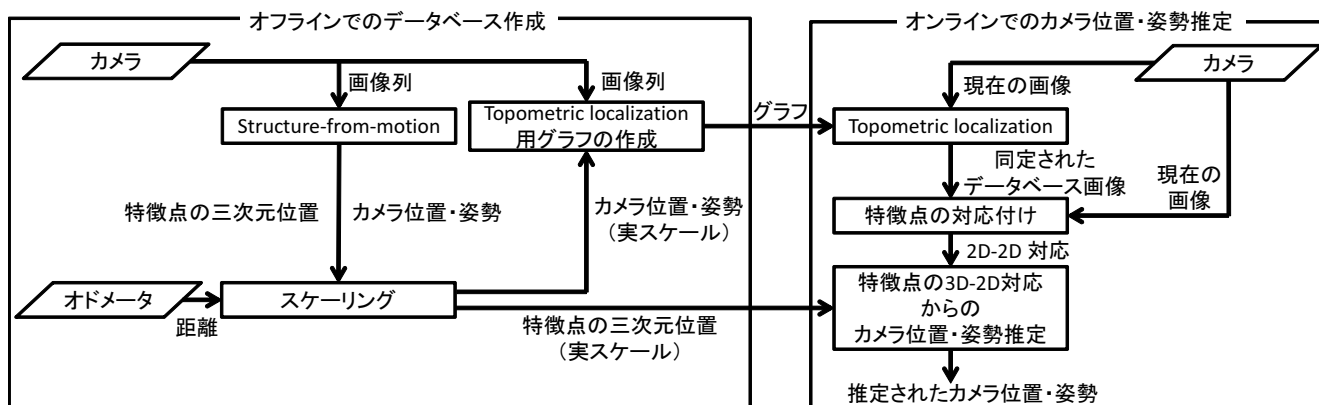


図 1 提案手法の処理とデータの流れ

節, 4 節で述べる.

2.1.1 オフラインでのデータベース作成

オフライン処理では, 事前に手動で経路上を走行することにより取得した動画を処理することで, オンラインでのカメラ位置・姿勢推定に利用するデータベースを作成する. まず, 取得された動画に対して Structure-from-Motion (SfM) 法 [4] を適用することで, カメラ位置・姿勢と特徴点の三次元位置を推定する. ここで, SfM 法ではカメラ位置・姿勢と特徴点の三次元位置のスケールを得ることができないため, オドメータから得られる走行距離を用いてこれらを実スケールに調整する.

次に, オンライン処理において現在のカメラ画像に対応するデータベース画像を同定するのに用いる Topometric Localization [5] 用のグラフを構築する. 具体的には, 推定されたカメラ位置を基に, 一定距離ごとに画像をサンプリングし, これらの位置をノード, ノード間の時系列の繋がりをエッジとする. また, サンプリングされた各画像から画像特徴を抽出し, 各ノードに対応付けて保存する.

2.1.2 オンラインでのカメラ位置・姿勢推定

オンライン処理では, オフラインで作成されたデータベースと現在のカメラ画像から現在のカメラ位置・姿勢を推定する. まず, Topometric Localization [5] により, オフラインで作成されたグラフのノードの中から, 現在カメラが位置する確率が最も高いものを選択することで, 現在の画像に対応するデータベース画像を同定する. ここで, Topometric Localization では, 入力画像とデータベース画像から抽出された画像特徴に加えて, トポロジカル (Topological) な情報としてエッジで表されるデータベース画像間の連結関係, メトリック (Metric) な情報としてデータベース画像間の距離を考慮して, 各ノードに対するカメラが位置する確率を計算する.

次に, 同定されたデータベース画像と現在の画像で特徴点の対応付けを行う. 最後に, 画像間の対応付けとオフライン処理で推定したデータベース画像中の特徴点の三次元位置から, 現在の画像における特徴点の三次元位置と二次

元位置の対応を取得し, これから現在のカメラ位置・姿勢を推定する.

2.2 関連研究

ここでは, 事前に撮影された画像を用いてカメラの位置・姿勢推定を行う従来手法を紹介し, 提案手法の位置付けを述べる. 従来手法は, SfM 法に基づく手法 [6], [7], [8], [9] とトポロジカルな位置・姿勢推定法に基づく手法 [10], [11] に分類できる.

2.2.1 SfM 法に基づく手法

SfM 法は画像間で特徴点を対応付けることにより, カメラ位置・姿勢と特徴点の三次元位置を推定する手法である. SfM 法に基づく手法では, 事前に撮影された画像群に対して SfM 法を適用することにより推定された特徴点の三次元位置と現在の画像上の特徴点の二次元位置の対応関係からカメラ位置・姿勢を推定する. この特徴点の三次元・二次元の対応関係からカメラ位置・姿勢を推定する問題は Perspective-n-Point (PnP) 問題として良く知られており, 多くの解法が存在する. 従って, SfM 法に基づく手法においては, どのように特徴点の対応関係を得るかが問題となる. 初期の手法として, Skrypnik ら [6] はデータベース中の全ての特徴点と現在の画像中の全ての特徴点との対応付けを, 回転・スケール変化に不変な SIFT 特徴量 [12] を用いて行なっている. しかし, この手法では, データベースに含まれる特徴点の数に比例して対応付けにかかる計算コストが大きくなるという問題がある. また, データベース中に見た目の類似した場所が含まれる場合には, 誤対応が多く発生するという問題がある.

これらの問題を解決するため, データベース中から対応付けに用いる特徴点を限定する手法が提案されている. Taketomi ら [7] は, オンライン処理において, 前のフレームで正しくデータベースと対応付けが成された特徴点を追跡し, 追跡された特徴点の三次元位置と現在の画像上の二次元位置の対応関係から暫定的なカメラ位置・姿勢を推定している. 暫定的なカメラ位置・姿勢の画角内におさまる

特徴点のみを対応付けの候補とすることで、計算コストの削減、対応付け精度の向上を実現しているが、急なカメラ位置・姿勢の変化や動物体による遮蔽などによりフレーム間の特徴点の追跡に失敗する場合がある。Irschara ら [8] や Arth ら [9] は、画像検索の手法を用い、現在の画像に最も類似している画像をデータベース中から同定し、同定された画像と現在の画像の間で特徴点を対応付ける手法を提案している。しかし、画像検索の手法では SIFT 特徴 [12] や SURF 特徴 [13] などの画像特徴のみを用いてデータベース中から画像を同定しており、データベース画像間のトポロジカル、メトリックな情報は用いられていない。

2.2.2 トポロジカルな位置・姿勢推定法に基づく手法

トポロジカルな位置・姿勢推定法では、オフラインで、事前に撮影された画像群から、各画像をノード、画像間の位置的な連結関係をエッジとするグラフを構築する。オンラインでは、入力画像とデータベース画像から抽出された画像特徴に加えて、トポロジカルな情報としてエッジで表されるデータベース画像間の連結関係を考慮しながら、現在の画像に対応するデータベース画像を同定する。この手法では、同定されたデータベース画像の位置・姿勢を現在のカメラの位置・姿勢として利用する。このような、大まかなカメラ位置・姿勢推定は、移動ロボットにおける部屋の認識などには有用であるが、自動運転を実現するためには、より高精度なカメラ位置・姿勢の推定が必要となる。

高精度な推定を実現するため、トポロジカルな位置・姿勢推定法に加えてエピポーラ幾何 [14] を利用する手法が提案されている。Kosecka ら [10] は二視点幾何を利用し、トポロジカルな位置・姿勢推定法によって同定されたデータベース画像と現在の画像間の相対的な位置・姿勢を推定している。しかし、二視点幾何では、相対的な位置・姿勢のスケールを求めることができない。これに対して、Murillo ら [11] は、トポロジカルな位置・姿勢推定法により、現在位置として最も確率が高い画像と二番目に確率が高い画像を同定し、これらの画像と現在の画像間の相対的な位置・姿勢を三視点幾何により推定している。この場合には、相対的な位置・姿勢のスケールは同定された二つのデータベース画像間の距離に対して決定される。

2.2.3 提案手法の位置づけ

提案手法は、SfM 法に基づく手法とトポロジカルな位置・姿勢推定法に基づく手法を組み合わせた手法である。提案手法では、Topometric Localization により、データベース中からトポロジカルな情報とメトリックな情報を考慮しつつ画像を同定することで、SfM 法に基づく手法において必要となる特徴点の探索範囲を限定する。また、SfM 法は二視点・三視点幾何より高精度な多視点幾何に基づいているため、トポロジカルな位置・姿勢推定法に加えてエピポーラ幾何を利用する手法より高精度な推定が期待される。

3. オフラインでのデータベース作成

本節では、事前撮影された動画から、オンラインでのカメラ位置・姿勢推定に利用するデータベースを作成する、オフライン処理について詳述する。

3.1 Structure-from-Motion

取得された動画に対して Structure-from-Motion (SfM) 法を適用することで、カメラ位置・姿勢と特徴点の三次元位置を推定する。ここでは、最新の SfM 実装の一つである、VisualSfM [4] を用いる。ただし、オンライン処理において、実スケールでカメラ位置を推定するためには、特徴点の三次元位置を実スケールで求める必要があるが、SfM 法のみではカメラ位置と特徴点の三次元位置のスケールを決定することができない。そこで、一般的な自動車には搭載されていると考えられる、オドメータから事前撮影時の自動車の走行距離を取得することで、SfM 法により得られたカメラ位置と特徴点の三次元位置を実スケールに調整する。

3.2 Topometric Localization 用グラフの作成

オンライン処理において現在のカメラ画像に対応するデータベース画像を同定するのに用いる Topometric Localization 用のグラフを構築する [5]。まず、SfM 法によって推定されたカメラ位置を基に、一定距離 ρ ごとに画像をサンプリングし、これらの位置をノード、ノード間の時系列の繋がりをエッジとする。ここで、一定距離 ρ ごとにサンプリングを行うのは、自動車の運動によるグラフ中での自動車の移動量の計算が容易になるためである。具体的には、自動車が単位時間 Δt の間、速度 u で動いた場合、グラフ中での自動車の移動量は $u\Delta t/\rho$ で計算される。

次に、オンライン処理において入力画像とデータベース画像間の相違度を計算するのに利用する画像特徴を各サンプリング画像から抽出し、グラフのノードと関連付けて保存する。ここで、入力画像とデータベース画像には撮影地点の違いがあることを考慮すると、画像特徴は回転やスケール変化に頑健である必要がある。ただし、一般的な SIFT 特徴 [12] や SURF 特徴 [13] などの局所特徴を利用すると、一つの画像につき数百点の特徴点が検出される。オンライン処理では、入力画像とデータベース中の複数の画像とで対応付けを行わなければならないため、これら数百点の特徴点同士での対応付けは計算コストが大きく、実時間処理が難しくなるという問題がある。そこで、Topometric Localization [5] では、局所特徴ではなく、画像全体から一つの特徴量を抽出する。具体的には、画像中心に画像全体を満たすスケールの特徴点を設定し、一つの SURF 特徴量を抽出する。これを Whole image SURF

(WI-SURF) 特徴量と呼び、以降ではベクトル $d \in R^{64}$ で表す。

4. オンラインでのカメラ位置・姿勢推定

本節では、オフラインで作成されたデータベースと現在のカメラ画像から現在のカメラ位置・姿勢を推定する、オンライン処理について詳述する。

4.1 Topometric Localization

Topometric Localization [5] は、オフラインで作成されたグラフのノードの中から、トポロジカルな情報とメトリックな情報を考慮しながら、現在のカメラが位置する確率が最も高いものを選択することで、現在の画像に対応するデータベース画像を同定する、

4.1.1 離散ベイズフィルタ

Topometric Localization [5] では各ノードにおける現在の自動車位置する確率を表す確率密度関数を離散ベイズフィルタにより計算する。離散ベイズフィルタのアルゴリズムを以下に示す。

Algorithm 1 離散ベイズフィルタ

Input: $\{p_{k,t-1}\}, u_t, \hat{d}_t$

Output: $\{p_{k,t}\}$

for all k do

 # Predict

$\hat{p}_{k,t} = \sum_{i=1}^N p(X_t = x_k | u_t, X_{t-1} = x_i) p_{i,t-1}$

 # Update

$p_{k,t} = \eta p(\hat{d}_t | X_t = x_k) \hat{p}_{k,t}$

end for

ここで、 X_t は時刻 t において自動車位置するノードを表し、 x_k は k 番目のノードを表す。自動車が時刻 t にノード x_k に位置する確率 $p(X_t = x_k)$ を $p_{k,t}$ と表すと確率密度関数は $p_{k,t}$ の全てのノードに対する集合 $\{p_{k,t}\}$ で表される。 u_t は時刻 t における自動車の速度、 \hat{d}_t は入力画像の WI-SURF 特徴量を表す。また N はグラフに含まれるノード数を表す。

離散ベイズフィルタでは予測と更新という二つの手続きにより、確率密度関数を計算する。予測手続きは速度 u_t の運動による確率密度関数の変化を状態遷移確率 $p(X_t = x_k | u_t, X_{t-1} = x_i)$ によって扱う。更新手続きは WI-SURF 特徴量 \hat{d}_t として得られる観測による確率密度関数の変化を観測確率 $p(\hat{d}_t | X_t = x_k)$ によって扱う。ここで、更新手続き中の η は更新後の確率密度関数の総和が 1 となるよう調整する正規化項である。以降、状態遷移確率および観測確率の定義について詳述する。

4.1.2 状態遷移確率

Topometric Localization [5] では、時刻 t における自動車の速度 u_t は平均 0、標準偏差 σ の正規分布に従うと仮

定し状態遷移確率を以下のように定義する。

$$p(X_t = x_k | u_t, X_{t-1} = x_i) = \frac{1}{\sqrt{2\pi\sigma'^2}} \exp\left(-\frac{(x_k - (x_i + u'_t))^2}{2\sigma'^2}\right) \quad (1)$$

ここで、 u'_t 、 σ' はそれぞれグラフ中での時刻 $t-1$ と t の間における自動車の移動量、標準偏差を表し、単位時間 Δt 、グラフのノード間の距離 ρ を用いて $u'_t = u_t \Delta t / \rho$ 、 $\sigma' = \sigma \Delta t / \rho$ と計算される。

4.1.3 観測確率

Topometric Localization [5] では、入力画像から得られる WI-SURF 特徴量 \hat{d}_t と k 番目のノードに対応するデータベース画像から得られる WI-SURF 特徴量 d_k の差のノルム $|\hat{d}_t - d_k|$ の関数として観測確率を定義する。ただし、WI-SURF 特徴量のノルムは 1 に正規化されているものとする。ここで、特徴量の差のノルムから確率への変換を行う関数は、真値付きのサンプルデータを用意し、実際の確率分布を計算することで求める。得られた確率分布はカイ二乗分布に近似しているため、パラメータを当てはめたカイ二乗分布を観測確率として用いる。

4.1.4 MAP 推定

時刻 t における自動車位置するノードは、離散ベイズフィルタによって得られた確率密度関数から最も確率が高いものを選択する (MAP 推定) によって

$$X_t = \underset{k}{\operatorname{argmax}} (p_{k,t}) \quad (2)$$

と決定される。

4.2 特徴点の対応付け

画像の回転、スケール変化に頑健な SIFT 特徴 [12] を用いて、Topometric Localization によって同定されたデータベース画像と現在の画像で特徴点の対応付けを行う。ここでは、計算時間の短縮のため、SIFT 特徴の GPU 実装である SiftGPU [15] を用いる。

4.3 特徴点の対応からのカメラ位置・姿勢推定

提案手法では、カメラ位置・姿勢を特徴点の三次元・二次元位置の対応から PnP 問題を解くことにより推定する。特徴点の三次元・二次元位置の対応は、オフライン処理において推定されたデータベース画像中の特徴点の三次元位置および Topometric Localization によって同定されたデータベース画像と現在の画像の間の特徴点の対応付けから得られる。ここで、PnP 問題の求解には、OpenCV に含まれる PnP 問題のソルバを用い、RANSAC により誤対応を排除した特徴点の対応関係よりカメラ位置・姿勢を推定する。このとき、インライア数が閾値より少ない場合には、推定失敗と扱う。以降の実験では、閾値は PnP 問題を線型に解く際の最小必要点数である 6 とした。

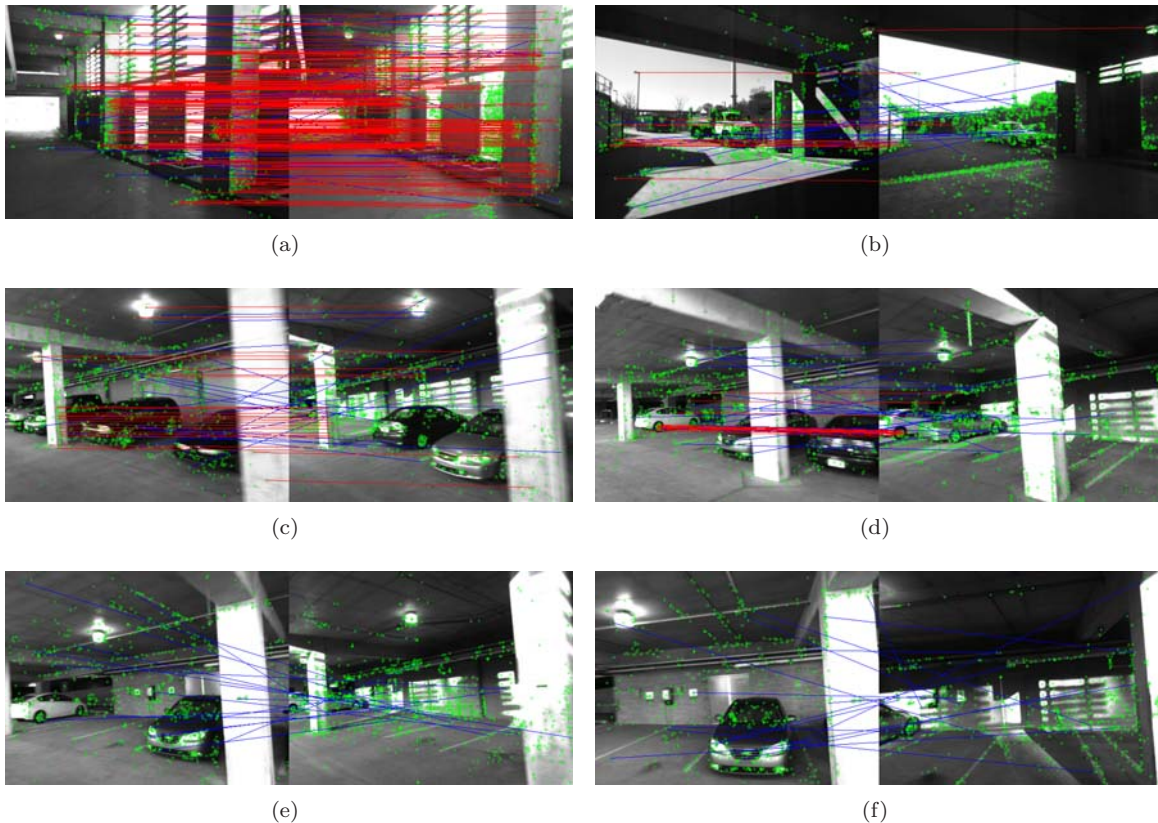


図 2 入力画像(右), Topometric Localization により同定されたデータベース画像(左), 特徴点の対応付け結果(赤線: インライア, 青線: アウトライア)の例

5. 実験

本節では, Valet Parking の自動化への応用を想定し, 屋内駐車場で撮影された動画を用い, 提案手法により自動車の位置・姿勢推定を行った結果を示す.

5.1 実験環境

カメラ(PointGrey Flea2), 走行距離を取得するオドメータを搭載した実験車両を用いて実験を行った. カメラは車両の進行方向に対して 45 度に取り付けられており, 動画(1024x768 pixel, 15fps)を取得する. オドメータは 3.1 節で述べたように, SfM 法の結果を実スケールに調整するために用いる.

データの取得は, Valet Parking の自動化への応用を想定し, 屋内駐車場で行った. 異なる日に, 駐車スペースから駐車場の入り口を通り, 駐車スペースに戻る経路上を手動で走行し, 二つの動画を取得した. これらの一つはオフライン処理におけるデータベースの作成に, 一つはオンライン処理における入力として用いた. 位置・姿勢推定精度の評価に用いる参照カメラ位置・姿勢は, SfM 法 [4] によって, 全ての画像間での特徴点の対応付けからオフラインで求めた.

5.2 位置・姿勢推定結果

図 2 に入力画像と Topometric Localization により同定されたデータベース画像, それらの画像間での特徴点の対応付け結果の例を示す. 図 3 にデータベース画像に対する車両位置・姿勢と特徴点の三次元位置, 提案手法により推定された車両位置・姿勢, 推定精度の評価に用いる参照車両位置・姿勢を示す. 提案手法では, 画像間で光源環境に差がある場合(図 2(b))やデータベース画像には存在しない自動車が入力画像に存在する場合(図 2(c))においても, 参照カメラ位置・姿勢と同様の位置・姿勢が推定された. しかし, 同じ車両が違う駐車スペースに停車している場合(図 2(d))には推定が不安定となった. また, Topometric Localization の誤り(図 2(e))や環境・光源の大きな違い(図 2(f))によって推定に失敗することもあった.

図 4 に参照車両位置・姿勢との比較による, 誤差のヒストグラムを示す. 表 1 に, PC (Intel Core i7-2600k 3.40 GHz CPU, NVIDIA GeForce GTX 580 GPU) を用いた場合における, 提案手法の平均計算時間を示す. これらの結果より, 提案手法は入力画像の約 70% において, 位置誤差 0.1 メートル, 姿勢誤差 0.3 度以内での推定を約 8 Hz で行えていることが確認できる.

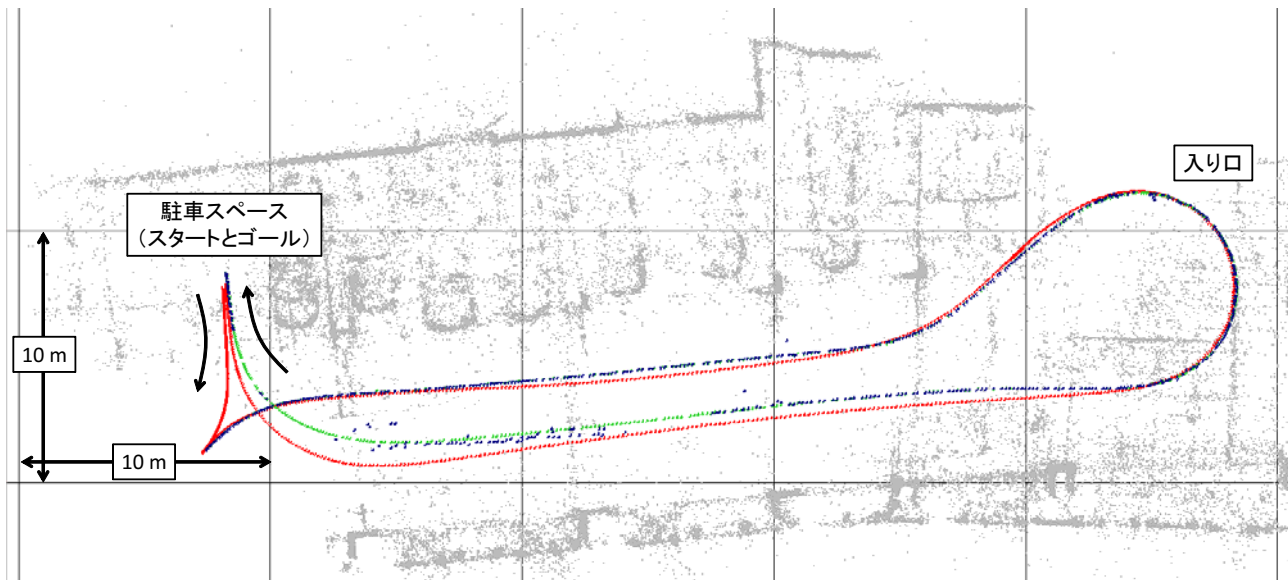


図 3 データベース画像に対する車両位置・姿勢 (赤) と特徴点の三次元位置 (灰), 提案手法により推定された車両位置・姿勢 (青), 推定精度の評価に用いる参照車両位置・姿勢 (緑)

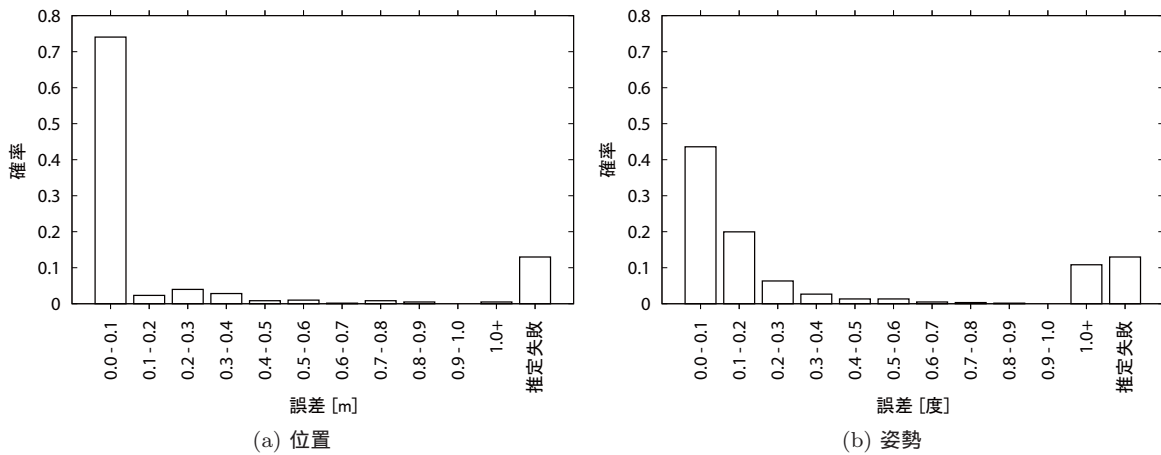


図 4 誤差のヒストグラム

表 1 平均処理時間 [msec]

Topometric localization	特徴点の対応付け	PnP 問題の求解	合計
8.8	47.8	67.6	124.2

6. まとめ

本稿では, Valet Parking の自動化を実現するため, 事前に経路上を手動で走行することにより取得した動画画像からオフライン処理で作成されたデータベースを基に, (1) Topometric Localization による現在の画像に対応するデータベース画像の同定, (2) 特徴点の三次元・二次元位置の対応関係からのカメラ位置・姿勢推定, を行うことで自動車の位置・姿勢を推定する手法を提案した. 実験では, 屋内駐車場で撮影された動画画像を用い, 提案手法による自動車の位置・姿勢推定を行い, 入力画像の約 70% において,

位置誤差 0.1 メートル, 姿勢誤差 0.3 度以内での推定を約 8 Hz で行えることを確認した. しかし, 環境や光源が変化している場合には, 不安定な推定や, 推定の失敗が起これることも確認した. 今後, 推定のロバスト性を向上させるため, オドメトリや安価な IMU など相対的な位置・姿勢を計測するセンサからの情報のカルマンフィルタによる融合, 前フレームにおいて対応付けに成功している特徴点を追跡することによる対応付け精度の向上, に取り組む.

謝辞 本研究は, National Science Foundation under Grant No. EEC-0540865 による. 糸 秀行は日本学術振興会「頭脳循環を活性化する若手研究者海外派遣プログラム」による支援を受けた.

参考文献

- [1] Suppé, A., Navarro-Serment, L. E. and Steinfeld, A.: Semi-autonomous virtual valet parking, *Proc. Int. Conf. on Automotive User Interfaces and Interactive Vehicu-*

- lar Applications*, pp. 139–145 (2010).
- [2] Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., Fong, P., Gale, J., Halpenny, M., Hoffmann, G., Lau, K., Oakley, C., Palatucci, M., Pratt, V., Stang, P., Strohband, S., Dupont, C., Jendrossek, L.-E., Koelen, C., Markey, C., Rummel, C., van Niekerk, J., Jensen, E., Alessandrini, P., Bradski, G., Davies, B., Ettinger, S., Kaehler, A., Nefian, A. and Mahoney, P.: Stanley: The robot that won the DARPA Grand Challenge, *J. of Field Robotics*, Vol. 23, No. 9, pp. 661–692 (2006).
- [3] Urmson, C., Anhalt, J., Bae, H., Bagnell, J. A. D., Baker, C. R., Bittner, R. E., Brown, T., Clark, M. N., Darms, M., Demitrish, D., Dolan, J. M., Duggins, D., Ferguson, D., Galatali, T., Geyer, C. M., Gittleman, M., Harbaugh, S., Hebert, M., Howard, T., Kolski, S., Likhachev, M., Litkouhi, B., Kelly, A., McNaughton, M., Miller, N., Nickolaou, J., Peterson, K., Pilnick, B., Rajkumar, R., Rybski, P., Sadekar, V., Salesky, B., Seo, Y.-W., Singh, S., Snider, J. M., Struble, J. C., Stentz, A. T., Taylor, M., Whittaker, W. R. L., Wolkowicki, Z., Zhang, W. and Ziegler, J.: Autonomous driving in urban environments: Boss and the Urban Challenge, *J. of Field Robotics*, Vol. 25, No. 8, pp. 425–466 (2008).
- [4] Wu, C.: VisualSFM: A visual structure from motion system, <http://www.cs.washington.edu/homes/ccwu/vsfin/> (2011).
- [5] Badino, H., Huber, D. and Kanade, T.: Real-time topometric localization, *Proc. Int. Conf. on Robotics and Automation* (2012).
- [6] Skrypnik, I. and Lowe, D. G.: Scene modelling, recognition and tracking with invariant image features, *Proc. Int. Symp. on Mixed and Augmented Reality*, pp. 110–119 (2004).
- [7] Taketomi, T., Sato, T. and Yokoya, N.: Real-time and accurate extrinsic camera parameter estimation using feature landmark database for augmented reality, *Int. J. of Computers and Graphics*, Vol. 35, No. 4, pp. 768–777 (2011).
- [8] Irschara, A., Zach, C., Frahm, J.-M. and Bischof, H.: From structure-from-motion point clouds to fast location recognition, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2599–2606 (2009).
- [9] Arth, C., Wagner, D., Klopschitz, M., Irschara, A. and Schmalstieg, D.: Wide area localization on mobile phones, *Proc. Int. Symp. on Mixed and Augmented Reality*, pp. 73–82 (2009).
- [10] Košecká, J., Li, F. and Yang, X.: Global localization and relative positioning based on scale-invariant keypoints, *Robotics and Autonomous Systems*, pp. 27–38 (2005).
- [11] Murillo, A., Guerrero, J. and Sagüés, C.: SURF features for efficient robot localization with omnidirectional images, *Proc. Int. Conf. on Robotics and Automation*, pp. 3901–3907 (2007).
- [12] Lowe, D. G.: Distinctive image features from scale-invariant keypoints, *Int. J. Computer Vision*, Vol. 60, No. 2, pp. 91–110 (2004).
- [13] Bay, H., Ess, A., Tuytelaars, T. and Gool, L. V.: Speeded-up robust features (SURF), *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346 – 359 (2008).
- [14] Hartley, R. I. and Zisserman, A.: *Multiple view geometry in computer vision*, Cambridge University Press, second edition (2004).
- [15] Wu, C.: SiftGPU: A GPU implementation of scale invariant feature transform (SIFT), <http://cs.unc.edu/~ccwu/siftgpu> (2007).