

深層学習を用いた自動運転向け自己位置推定精度の予測

松本 和人¹ Ehsan Javanmardi¹ 中里 仁¹ 塚田 学¹

概要: 近年、自動運転の社会実装に向けた研究開発や実証実験が盛んに行われている。自動運転を実現するにあたって、センサ情報から周囲の環境の情報を取得し、車体の位置を推定する必要がある。これを自己位置推定という。自己位置推定のセンサには 3DLiDAR がよく用いられる。3DLiDAR は測定精度が高く、周囲の明るさの影響を受けないため、高精度に自己位置推定を行えるが、周囲に特徴物が少ないところでは自己位置推定の精度が低下するという課題がある。本研究では、自己位置推定の精度を予測する手法を提案する。自己位置推定精度の予測を行い、精度が悪い場所に対して GNSS や IMU など 3DLiDAR 以外のセンサを用いたり、舗装マーキングを用いたりすることで、全体的な自己位置推定の精度を向上させることができる。オープンソース自動運転シミュレータを用いて自己位置推定精度予測のためのデータセットを作成した。実験では作成したデータセットに対して提案手法を行った。結果として、自己位置推定精度を高精度で予測できたことを報告する。

Predicting Localizability for Autonomous Driving Using Deep Learning

Matsumoto Kazuto¹ Ehsan Javanmardi¹ Jin Nakazato¹ Manabu Tsukada¹

1. はじめに

昨今、レベル4の自動運転の社会実装に向けた研究開発や実証実験が世界中で行われている。例えば、サンフランシスコでは、テック企業が無人商用ロボタクシーを公道で走行させた実証実験が行われている [1]。また、中国では Baidu が武漢市で一般人向けのロボタクシーを走行させ、2023年1月末までの累計乗車件数は140万件を超えており、今後200台の自動運転タクシーを追加する予定である [2]。

ここで、自動運転を実現するにあたって、センサ情報から周囲の環境の情報を取得し、車体の位置を推定する必要がある。これを自己位置推定という。安全な自動運転走行を行うためには、自己位置推定の誤差を数十センチメートル程度に抑えることが必要である。現在、自己位置推定で用いられるセンサには GNSS(Global Navigation Satellite System)、カメラ、IMU(Inertial Measurement Unit) や 3DLiDAR(Light Detection and Ranging sensor) などがあ

環境下では自己位置推定の誤差が大きくなるという問題がある [3]。一方でカメラは安価で使用できるが、周囲の明るさの影響を受けるという問題がある [4]。IMUを用いる自己位置推定は、車体の加速度と角速度の情報をもとに行う。IMUを用いて自己位置推定を行う場合、誤差が蓄積するという問題がある [4]。3DLiDARはレーザー光を照射して、その反射光の情報をもとに対象物までの距離や形などを測定する。そのため、3DLiDARは深度を測ることができ、周囲の明るさの影響を受けない。また、360°のデータを集めることができ、測定誤差が100mで2cm程度の誤差であり、精度がとても高い [5]。3DLiDARにはコストが高いという問題点があるが、近年安価になってきている [6]。そのため、最近では自己位置推定に用いるセンサとして 3DLiDAR がよく用いられており、主要な自動運転企業である Google, Toyota, Uber, Ford などは 3DLiDAR を採用している [7], [8], [9]。上記理由より本研究では、3DLiDAR を用いた自己位置推定に関して研究を行う。

現在、3DLiDARを用いた自己位置推定には大きく分けて二つの手法がある。一つ目が SLAM [10], [11], [12] で、二つ目がマップマッチングである。SLAMは自己位置推定と環境地図作成を同時に行う方法である。環境地図作成

¹ 東京大学

では、センサ情報から周辺環境を把握し地図を作成する。SLAMは作成した地図を自己位置推定に使い、自己位置推定の値を地図作成に用いることを繰り返すため、誤差が蓄積するという問題がある [13]。一方で、マップマッチングはセンサ情報と地図情報をマッチングすることにより自己位置推定を行う。この手法は事前にマップ情報を用意しておく必要があるが、最近では地図作成が低コストで実現できる。SLAMよりマップマッチングの方が高精度で自己位置推定できるため、本研究ではマップマッチングを用いて自己位置推定に焦点を当てる。

マップマッチングを用いた自己位置推定を行う手法として、スキャンマッチングがある。スキャンマッチングは異なる位置から撮影した三次元点群同士の位置関係を推定し、点群の合成を行うアルゴリズムである。スキャンマッチングにはICP(Iterative Closest Point) スキャンマッチング [14], [15], [16] とNDT(Normal Distribution Transform) スキャンマッチング [17], [18], [19] がある。ICP スキャンマッチングは二つの点群の点同士をマッチングすることでスキャンマッチングを実現する。NDT スキャンマッチングは点群単位でマッチングするのではなく、地図点群をボクセルに区切り、各ボクセル内で標準分布を作成し、点と分布でマッチングを行うことで、マッチングの精度を上げている。スキャンマッチングの手法は動的な物体の影響を受けやすいという特徴がある。動的な物体にロバストな方法として、特徴抽出を行う手法がある。

特徴抽出を行う手法は不変な情報を抽出するため、動的な物体の影響を受けにくいのでスキャンマッチングよりも高精度に自己位置推定を行うことができる。しかし特徴物が動的な物体に遮られた時に、マッチングができないという問題と、カーブや柱状のシンボルなどの特定の特徴物がないところでは自己位置推定の精度が落ちるという問題がある。点群データから高い精度で特徴抽出を行う手法として深層学習を用いて特徴抽出を行う手法がある。PointNetVLAD [20] はディープメトリックラーニング [21] を用いて三次元点群データの特徴を抽出した表現ベクトルを推測した。ディープメトリックラーニングとは、サンプル間の類似度に基づいて、入力データを特徴量空間への変換を学習させる手法である。こちらの手法は特定の特徴物がない場所でも特徴抽出ができるという利点がある。深層学習を用いることで、特徴が少ない場所でも特徴を抽出することができるようになったが、田舎道やトンネルなどの極度に特徴が少ない場所では依然として自己位置推定の精度は低いという課題がある。

マップマッチングでは特徴の無いトンネルや田舎道で精度よく自己位置推定することは難しい。GNSSやIMUなどの3DLiDARセンサを用いたり、舗装マーキングを用いたりすることで、自己位置推定が難しい区画の自己位置推定の精度を向上させる必要があり、そのためには自己位置推

定の精度が低い区画を特定する必要がある。Julianら [22] は3DLiDARセンサ情報から、自己位置推定の推測を行った。こちらの手法は環境に応じたパラメーターチューニングが必要であり、シミュレーション上で手作業でデータを収集する必要があるという問題がある。また、自己位置推定の誤差を算出する際に、近似値を用いているという課題もある。

本研究では深層学習を用いて、マップマッチングを用いた自己位置推定の精度を予測する手法を提案する。この手法を用いることで、各地点の3DLiDARセンサデータから、実際に車を走らせて自己位置推定することなく、自己位置推定誤差を予測することができる。具体的には、まず3DLiDARセンサから取得した点群データをMinkLoc3D-SI [23] を用いて特徴ベクトルに変換する。次に、自己位置推定を行い、誤差を導出し、特徴ベクトルと誤差の関係を深層学習のモデルで学習する。学習したモデルを用いることで、入力点群に対して自己位置推定の誤差を予測できる。

オープンソース自動運転シミュレータ「AWSIM [24]」を用いて自己位置推定精度予測のためのデータセットを作成する。シミュレータを用いると、自動運転車の実際の位置情報を得られるため、近似値ではない実際の自己位置推定の誤差を算出することができる。実験では作成したデータセットから得られる点群データから自己位置推定誤差の関係を予測を行う。

本論文の貢献は以下の通りである。

- 深層学習を用いた新しい自己位置推定精度予測手法の提案を行う。
- 自己位置推定値と実際の位置の情報が入ったデータセットの作成をして真の自己位置推定誤差を導出

本論文の構成は次のようになっている。2章では関連研究を述べたのち現状の課題を明らかにする。3章では、課題を解決するためのシステムアーキテクチャ及び実装方法について述べる。4章では、実験した結果とその評価について述べる。5章では、本研究の結論を述べる。

2. 関連研究

スキャンマッチングとは、異なる位置から撮影した三次元点群同士の位置関係を推定し、点群の合成を行うアルゴリズムである。3DLiDARを用いたスキャンマッチングの手法として、ICP(Iterative Closest Point) スキャンマッチング [14], [15], [16] とNDT(Normal Distribution Transform) スキャンマッチング [17], [18], [19] がある。ICP スキャンマッチングは二つの点群の点同士をマッチングすることでスキャンマッチングを実現する。NDT スキャンマッチングは点群単位でマッチングするのではなく、地図点群をボクセルに区切り、各ボクセル内で標準分布を作成し、点と分布でマッチングを行うことで、マッチングの精度を上げ

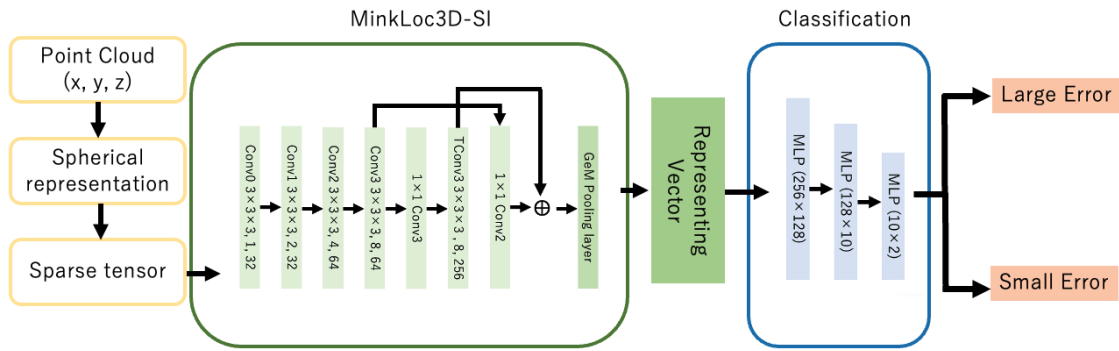


図 1 深層学習の構成図

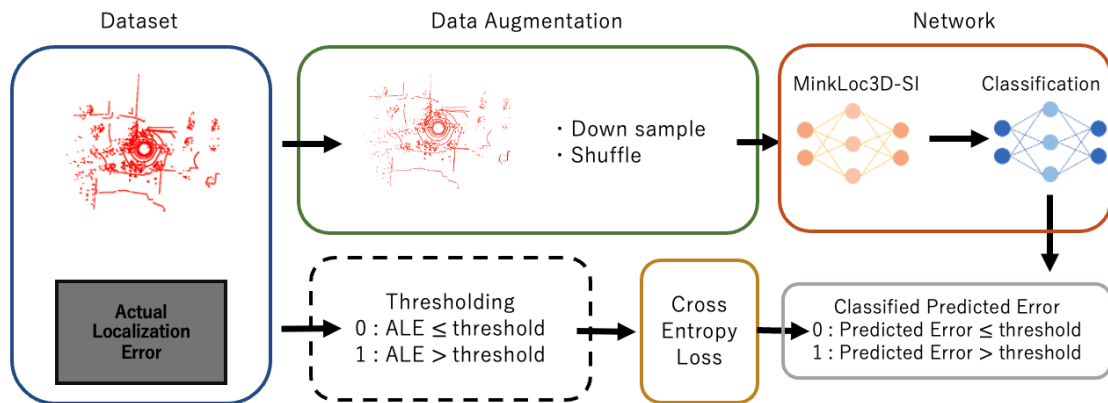


図 2 提案手法の全体図

ている。スキャンマッチングの手法は動的な物体の影響を受けやすいという欠点がある。

動的な物体にロバストな方法として、特徴抽出を行う手法がある。Yihuan ら [25] はセンサから得られた点群データから障害物を除き、道路縁石を抽出し、そのデータと地図データとのマッチングを行った。Yingfeng ら [26] はカーブと柱状の物体を抽出した二次元地図データを作成し、その二次元地図データとセンサデータとのマッチングを行った。Liang ら [27] はセンサデータに対し、ビームモデルを施し、正確なカーブの輪廓を抽出した。Mahdi ら [28] は高解像度航空監視データから道路特徴を自動で抽出した。特徴抽出を行う手法は不変な情報を抽出するため、動的な物体の影響を受けにくいのでスキャンマッチングよりも高精度に自己位置推定を行うことができる。しかし特徴物が動的な物体に遮られた時に、マッチングができないという問題と、カーブや柱状のシンボルなどの特定の特徴物がないところでは自己位置推定の精度が落ちるといった問題がある。

点群データから高い精度で特徴抽出を行う手法として深層学習を用いて特徴抽出を行う手法がある。PointNetVLAD [20] はディープメトリックラーニング [21] を用いて三次元点群データの特徴を抽出した表現ベクトルを推測した。ディープメトリックラーニングとは、サンプル間の類似度に基づいて、入力データを特徴量空間への変

換を学習させる手法である。SVT-Net [29] は局所的な特徴を抽出する深層学習のモデルである ASVT(Atom-based Sparse Voxel Transformer) と大域的な特徴を抽出する深層学習のモデルである CSVT(Cluster-based Sparse Voxel Transformer) を組み合わせることで幅広い粒度の特徴抽出を試みた。LPD-Net [30] はグラフベースの近傍集約モジュールを用いることで、局所構造を抽出しそれらを集約することで大域的な特徴を抽出した。

こちらの手法は特定の特徴物がない場所でも特徴抽出ができるという利点がある。深層学習を用いることで、特徴が少ないところでも特徴を抽出することができるようになったが、田舎道やトンネルなどの極度に特徴が少ない場所では依然として自己位置推定の精度は低いという課題がある。

マップマッチングでは特徴の無いトンネルや田舎道で精度よく自己位置推定することは難しい。GNSS や IMU などの 3DLiDAR センサを用いたり、舗装マーキングを用いたりすることで、自己位置推定が難しい区画の自己位置推定の精度を向上させる必要があり、そのためには自己位置推定の精度が低い区画を特定する必要がある。Ehsan ら [31] は事前に与えられる地図情報から、マップマッチングエラーの事前推定フレームワークを提案した。また、Julian ら [22] は 3DLiDAR センサ情報から、自己位置推

定の推測を行った。こちらの手法は環境に応じたパラメータチューニングが必要であり、シミュレーション上で手作業でデータを収集する必要があるという問題がある。また、自己位置推定の誤差を算出する際に、近似値を用いているという課題もある。

3. 提案手法

本論文では、3DLiDAR データから自己位置推定精度を予測する新しい手法を提案する。この手法は、3DLiDAR データを効果的に解析し、自己位置推定精度の予測精度を向上させることを目的としている。提案手法の流れは以下の通りである。まず、3DLiDAR 点群から特徴を抽出する。これには、MinkLoc3D-SI を使用する。MinkLoc3D-SI を用いることで、点群データから有用な特徴を効率的に抽出し、後続の処理に適した形式に変換することができる。次に抽出したデータと自己位置推定の誤差の関係を深層学習のモデルにより学習する。本提案手法では、姿勢データは扱わず、位置情報に焦点を当てる。これにより、位置推定タスクに特化した効率的な手法を実現することができる。提案手法の概要図を図1に示す。提案手法の概要としては、入力点群をダウンサンプリングと順序入れ替えを行うことでデータ拡張を行い、データを深層学習のモデルに入力する。自己位置推定のエラーを閾値処理し、正解ラベルとし、深層学習モデルを学習させる。深層学習の構成図を図2に示す。深層学習では、入力点群の特徴を MinkLoc3D-SI を用いて抽出し、抽出した特徴ベクトルから自己位置推定のエラーが閾値以上かどうかを判定する。

3.1 MinkLoc3D-SI

MinkLoc3D-SI は、3DLiDAR データから特徴を抽出し、特徴ベクトルに埋め込む手法である。この手法では、入力点群データをボクセル化し、ディープラーニングのネットワークである FPN アーキテクチャ [32] を適用することで、三次元点群データの局所的な特徴を抽出する。さらに、generalized-mean(GeM)pooling [33] を用いて局所的な特徴を集約し、大域的な特徴を表す特徴ベクトルを出力する。

FPN アーキテクチャを用いることで、さまざまなスケールでの特徴抽出が可能になり、大規模で複雑なシーンでの特徴抽出を実現した。また、ディープメトリックラーニングを用いることで、似た点群同士は距離が近く、似ていない点群同士は距離が遠くなるように点群データをベクトル空間に埋め込むことを実現した。

MinkLoc3D-SI は、高度な特徴抽出能力とディープメトリックラーニングによる効果的な埋め込み学習が組み合わさり、3D 空間における点群データの理解と分析に大きく貢献した。これにより、ロボットや自動運転車などのさまざまな応用が可能になり、現実世界での応用範囲が拡大していくことが期待される。

本研究では、一つの 3DLiDAR 点群データを 256 次元ベクトルに埋め込んだ。

3.2 深層学習モデル構成

MinkLoc3D-SI で抽出した 256 次元の特徴ベクトルから自己位置推定の誤差が閾値以上か以下かを分類するネットワークを考えた。三層のニューラルネットワークで構築した。ロス関数には Cross Entropy Loss [34] を用いた。Cross Entropy Loss は、分類タスクで一般的に使用されるロス関数であり、ネットワークが予測した確率分布と正解ラベルの確率分布の間の距離を測定する。このロス関数を用いることで、ネットワークは自己位置推定精度の予測精度を向上させるように学習が進められる。以上のネットワークアーキテクチャにより、256 次元の特徴ベクトルから自己位置推定精度に関係する情報を効果的に抽出し、自己位置推定精度を効果的に予測することができる。

3.3 データオーギュメンテーション

深層学習のモデルの汎化性能を上げるために学習データ拡張を行った。今回はダウンサンプリングと点群順序入れ替えを行った。

ダウンサンプリングは、点群データのサイズを縮小し、モデルが異なる解像度のデータに対応できるようにするために、ダウンサンプリングを行った。これにより、モデルはより効率的に学習できるだけでなく、さまざまな環境やセンサーのデータに対しても適応性を持たせることができる。今回はサンプリングレートを 0.01m から 0.1m まで 0.01m ずつ増やしてサンプリングした。

点群データ順序入れ替えは、点群データの順序をランダムに入れ替えることで、モデルが点群データの順序に依存しない特徴を学習するように促した。これにより、データの順序が変わってもモデルの性能が低下しないした。

これらのデータオーギュメンテーション手法により、学習データのバリエーションが増え、モデルはさまざまな状況に対応できるようになる。また、過学習を防ぐ効果もあり、モデルの汎化性能が向上する。

4. データセット

自己位置推定値と位置の真値が両方含まれるデータセットを作成した。これは、オープンソース自動運転シミュレータ「AWSIM」を用いて実現した。シミュレータを使用しない場合、位置の真値を正確に取得することが困難である。

このデータセットは、東京都の西新宿エリアを対象として作成されており、各ポジションにおいて位置の真値、自己位置推定値、3DLiDAR 点群データ、そしてダウンサンプリングした点群データが含まれている。自己位置推定の手法には NDT スキャンマッチングが用いられており、グリッ



図 3 西新宿エリアの地図

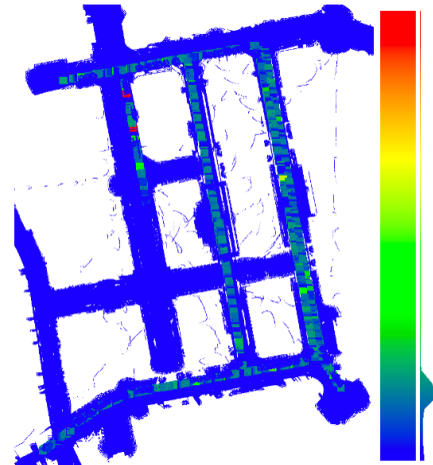


図 4 データセットの自己位置推定の誤差

| | 正答率 | 適合率 | 再現率 | F 値 |
|---------|------|-------|-------|-------|
| 閾値 0.3m | 0.73 | 0.831 | 0.841 | 0.836 |
| 閾値 0.5m | 0.66 | 0.842 | 0.744 | 0.790 |

ドサイズは 2 メートルとなっている。図 3 にデータセットを取得した場所を、図 4 に各位置の自己位置推定の誤差を示す。データに位置の真値と自己位置推定値があることで、自己位置推定の誤差を計算することができ、自己位置推定精度予測をより効果的に行える。

5. 実験

提案手法の精度を測るために実験を行った。シミュレータを用いて西新宿エリアで取得したデータセットを訓練データとテストデータに分割し、訓練データで深層学習モデルを訓練し、テストデータで評価した。自己位置推定の誤差の閾値を 0.3m と 0.5m に設定して実験を行った。図 3 の青線で塗られた PATH1 をテストデータとし、それ以外のパスで得られたデータを訓練データとした。

5.1 閾値が 0.3m のとき

実験結果を表 1 に示す。正答率は全体のデータの中で正しく分類できた割合である。適合率は自己位置推定の誤差が閾値より大きいと判定されたデータのうち、実際に大きいデータ数の割合である。再現率は実際に自己位置推定の誤差が閾値より大きいデータのうち、誤差が閾値より大きいと判定されたデータ数の割合である。F 値は適合率と再現率の調和平均であり、0 から 1 の範囲の値となり、1 に近づくほど良い分類モデルであると言える。本実験では、0.836 という高い F 値が得られた。提案手法の予測値と実際のエラーの関係を図 5 に、閾値処理した結果を図 6 に示

す。図 5 の青線は実際の自己位置推定の誤差を表し、点線は閾値である 0.3m を示している。橙色の点は今回の提案手法の予測値を表している。この図から、概ね精度良く予測できていることがわかる。図 6 の青い地点は自己位置推定誤差が 0.3m より小さく、赤い地点は自己位置推定誤差が 0.3m より大きいと予測されたことを示す。

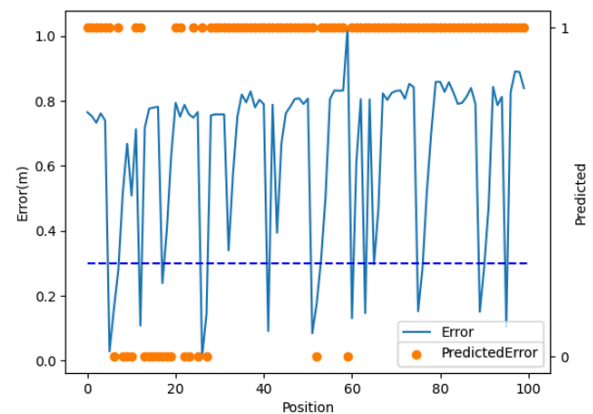
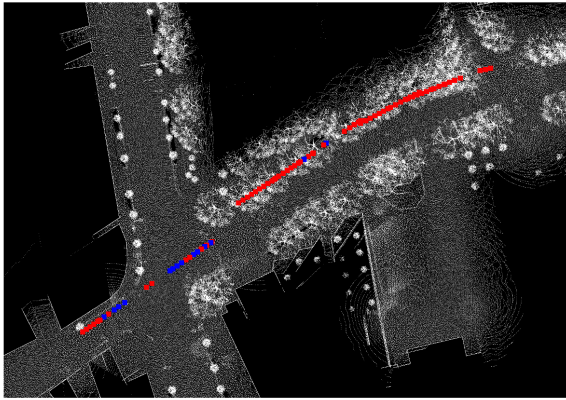


図 5 閾値が 0.3m の際のエラーと予測値の関係

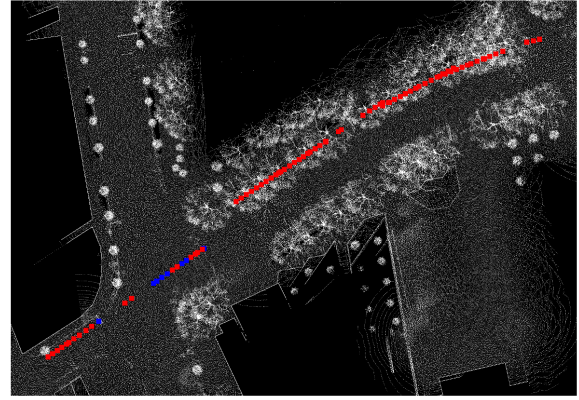
5.2 閾値 0.5m における結果及び議論

実験結果を表 1 に示す。閾値が 0.3m の場合に比べて F 値は若干低くなったが、それでも 0.79 という高い F 値が得られた。提案手法の予測値と実際のエラーの関係を図 7 に、閾値処理した結果を図 8 に示す。図 7 の青線は実際の自己位置推定の誤差を表し、青い点線は閾値である 0.5m を示している。橙色の点は今回の提案手法の予測値を表している。図より、閾値を上回っている地点は精度良く予測できているが、下回っている地点の予測精度が高くないこ



● 誤差が0.3mより小さい地点
● 誤差が0.3mより大きい地点

図 6 閾値を 0.3m とした時の結果



● 誤差が0.5mより小さい地点
● 誤差が0.5mより大きい地点

図 8 閾値を 0.5m とした時の結果

とがわかる。これは閾値を 0.5m とした時の方が 0.3m とした時より予測が難しいためであり、より大きな閾値で行うには、精度を上げていく必要があると考える。図 8 の青い地点は自己位置推定誤差が 0.3m より小さく、赤い地点は自己位置推定誤差が 0.3m より大きいと予測されたことを示す。

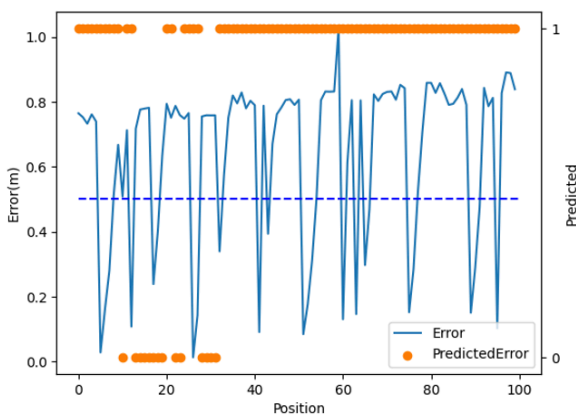


図 7 閾値が 0.5m の際のエラーと予測値の関係

5.3 所要時間

自己位置推定誤差を予測する際にかかる時間を表 2 に示す。点群データを特徴ベクトルに変換するのにかかった平均の時間は 0.0237 秒であり、特徴ベクトルを学習したモデルに入力し、予測値を得るのにかかった時間の平均は 0.0000102 秒であった。合計すると、予測にかかる時間の総計は 0.0237 秒であった。一秒あたり約 42 回予測ができるので、十分にリアルタイムに予測ができると考える。

表 2 所要時間

| | |
|------|---------------------------|
| 特徴抽出 | 2.37×10^{-2} (s) |
| 推論 | 1.02×10^{-5} (s) |
| 合計 | 2.37×10^{-2} (s) |

6. まとめ

本研究では、自己位置推定の誤差を予測する手法を提案した。実験の結果、誤差の閾値が 0.3m の場合には F 値が 0.836、0.5m の場合には 0.79 という高い精度で分類できることが示された。しかし、閾値が 0.3m の際は閾値より自己位置推定の誤差が大きい地点と小さい地点をよく分類できたが、閾値が 0.5m の際は閾値より誤差が小さい地点の予測精度が低かったため、今後より精度を上げていく必要があるとも考える。処理時間についてはリアルタイムに予測できることがわかった。

今後の展望としては、点群マップを入力とした手法を考慮し、更に精度の向上を目指すことが挙げられる。また、動的な障害物が存在する環境に対応した手法の検討も重要である。これらの改良により、さらに実用性の高い自己位置推定誤差予測手法の開発が期待される。

謝辞 本研究の一部は、NICT「Beyond 5G 研究開発促進事業（#01101）」及び科研費（#21H03423）の助成を受けて行われた。ここに記して感謝の意を表したい。

参考文献

- [1] Gil West. "we' re going commercial". cruise. <https://getcruise.com/news/blog/2022/were-going-commercial/>. 2023-03-22 参照.
- [2] Kevin Krolicki. "baidu, pony.ai start driverless robotaxi tests in beijing". REUTERS. <https://jp.reuters.com/article/baidu-autonomous-license-idCAKBN2TE05B>. 2023-03-31 参照.
- [3] Miguel Ángel de Miguel, Fernando García, and José María Armingol. Improved lidar probabilistic localization for autonomous vehicles using gnss. *Sensors*, Vol. 20, No. 11, p. 3145, 2020.
- [4] Athanasios Chalvatzaras, Ioannis Pratikakis, and Angelos A Amanatiadis. A survey on map-based localization techniques for autonomous vehicles. *IEEE Transactions on Intelligent Vehicles*, 2022.
- [5] Mahdi Elhousni and Xinming Huang. A survey on 3d li-

- dar localization for autonomous vehicles. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1879–1884. IEEE, 2020.
- [6] Inc Neuivition. "lidar price for cars". <https://www.neuivition.com/media/blog/lidar-price.html>. 2023-05-06 参照.
- [7] Mark Shand. "informing smarter lidar solutions for the future". <https://blog.waymo.com/2022/09/informing-smarter-lidar-solutions-.html>. 2023-05-17 参照.
- [8] Devin Coldewey. "here's how uber's self-driving cars are supposed to detect pedestrians". <https://techcrunch.com/2018/03/19/heres-how-ubers-self-driving-cars-are-supposed-to-detect-pedestrians/>. 2023-05-17 参照.
- [9] Kirsten Korosec. "ford, vw seeking buyer for argo ai's lidar unit". <https://techcrunch.com/2022/11/01/ford-vw-seeking-buyer-for-argo-ais-lidar-unit/>. 2023-05-17 参照.
- [10] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, Vol. 31, No. 5, pp. 1147–1163, 2015.
- [11] Tim Bailey and Hugh Durrant-Whyte. Simultaneous localization and mapping (slam): Part ii. *IEEE robotics & automation magazine*, Vol. 13, No. 3, pp. 108–117, 2006.
- [12] MWM Gamini Dissanayake, Paul Newman, Steve Clark, Hugh F Durrant-Whyte, and Michael Csorba. A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on robotics and automation*, Vol. 17, No. 3, pp. 229–241, 2001.
- [13] Josep Aulinas, Yvan Petillot, Joaquim Salvi, and Xavier Lladó. The slam problem: a survey. *Artificial Intelligence Research and Development*, pp. 363–371, 2008.
- [14] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, Vol. 1611, pp. 586–606. Spie, 1992.
- [15] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, Vol. 13, No. 2, pp. 119–152, 1994.
- [16] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the icp algorithm. In *Proceedings third international conference on 3-D digital imaging and modeling*, pp. 145–152. IEEE, 2001.
- [17] Martin Magnusson. *The three-dimensional normal-distributions transform: an efficient representation for registration, surface analysis, and loop detection*. PhD thesis, Örebro universitet, 2009.
- [18] Peter Biber and Wolfgang Straßer. The normal distributions transform: A new approach to laser scan matching. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, Vol. 3, pp. 2743–2748. IEEE, 2003.
- [19] Todor Stoyanov, Martin Magnusson, Henrik Andreasson, and Achim J Lilienthal. Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations. *The International Journal of Robotics Research*, Vol. 31, No. 12, pp. 1377–1393, 2012.
- [20] Mikaela Angelina Uy and Gim Hee Lee. Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4470–4479, 2018.
- [21] Sumit Chopra, Raia Hadsell, and Yann LeCun. Learning a similarity metric discriminatively, with application to face verification. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 539–546. IEEE, 2005.
- [22] Julian Nubert, Etienne Walther, Shehryar Khattak, and Marco Hutter. Learning-based localizability estimation for robust lidar localization. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 17–24, 2022.
- [23] Kamil Żywanowski, Adam Banaszczyk, Michał R Nowicki, and Jacek Komorowski. Minkloc3d-si: 3d lidar place recognition with sparse convolutions, spherical coordinates, and intensity. *IEEE Robotics and Automation Letters*, Vol. 7, No. 2, pp. 1079–1086, 2021.
- [24] TIER IV. "利用者 150 万人 2023 年も世界一を目指すバイドゥのロボタクシー". <https://tier4.github.io/AWSIM/>. (2023-03-22 参照).
- [25] Yihuan Zhang, Liang Wang, Xuhui Jiang, Yong Zeng, and Yifan Dai. An efficient lidar-based localization method for self-driving cars in dynamic environments. *Robotica*, Vol. 40, No. 1, pp. 38–55, 2022.
- [26] Yingfeng Cai, Ziheng Lu, Hai Wang, Long Chen, and Yicheng Li. A lightweight feature map creation method for intelligent vehicle localization in urban road environments. *IEEE Transactions on Instrumentation and Measurement*, Vol. 71, pp. 1–15, 2022.
- [27] Liang Wang, Yihuan Zhang, and Jun Wang. Map-based localization method for autonomous vehicles using 3d-lidar. *IFAC-PapersOnLine*, Vol. 50, No. 1, pp. 276–281, 2017.
- [28] Mahdi Javanmardi, Ehsan Javanmardi, Yanlei Gu, and Shunsuke Kamijo. Towards high-definition 3d urban mapping: Road feature-based registration of mobile mapping systems and aerial imagery. *Remote Sensing*, Vol. 9, No. 10, p. 975, 2017.
- [29] Zhaoxin Fan, Zhenbo Song, Hongyan Liu, Zhiwu Lu, Jun He, and Xiaoyong Du. Svt-net: Super light-weight sparse voxel transformer for large scale place recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36, pp. 551–560, 2022.
- [30] Zhe Liu, Shunbo Zhou, Chuanzhe Suo, Peng Yin, Wen Chen, Hesheng Wang, Haoang Li, and Yun-Hui Liu. Lpd-net: 3d point cloud learning for large-scale place recognition and environment analysis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2831–2840, 2019.
- [31] Ehsan Javanmardi, Mahdi Javanmardi, Yanlei Gu, and Shunsuke Kamijo. Pre-estimating self-localization error of ndt-based map-matching from map only. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 22, No. 12, pp. 7652–7666, 2020.
- [32] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117–2125, 2017.
- [33] Filip Radenović, Giorgos Toliás, and Ondřej Chum. Fine-tuning cnn image retrieval with no human annotation. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 41, No. 7, pp. 1655–1668, 2018.
- [34] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.