

キーフレーム間隔の動的制御による 高精度自己位置推定

北島克将¹ 藤田匠¹ 木内裕介¹

概要: V-SLAM (Visual Simultaneous Localization And Mapping) による地図作成では、地図容量削減のため、キーフレームを疎に保つことが一般的である。V-SLAM ではキーフレームとカメラ画像との照合により自己位置を推定するため、無人フォークリフトのように、安全を確保する上でパレットへの爪挿しが可能なレベルの誘導精度が必要な車両では、通常の V-SLAM ではキーフレームが過剰に疎となり位置推定精度が不足する場合があった。そこで本稿では、地図作成時にキーフレーム間隔を制御する手法を提案する。無人フォークリフトを用いた評価の結果、提案手法により位置推定誤差を低減することができた。

Visual Localization Based on Dynamically Controlled Keyframe Intervals

KATSUMASA KITAJIMA^{†1} TAKUMI FUJITA^{†1}
YUSUKE KINOUCHI^{†1}

Abstract: In the mapping process of V-SLAM (Visual Simultaneous Localization And Mapping), keyframes are often kept sparse to reduce the map size. Since a V-SLAM system performs localization by comparing the keyframes and input images, it may lead to insufficient accuracy for some applications including automated guided forklifts, which require accurate poses in order to insert the forks into pallets. In this paper we propose a method to dynamically control the keyframe density. Experimental results on an automated guided forklift show the proposed method reduces the localization error.

1. はじめに

自律移動ロボットを実現する上で最も重要な技術の一つが自己位置推定 (Localization) 技術である。ここで自己位置推定技術とは、環境の中でロボットがどこにいるのか推定する技術である。

GNSS (Global Navigation Satellite System) や環境中に設置した位置が既知のランドマークを利用すれば自己位置を得ることができる。しかし、GNSS は屋内や背の高い建物に囲まれた場所では利用が困難である。また、環境にランドマークを設置する手法は、設置時に工事コストが生じるうえ、屋内で用いる場合、棚等の構造物のレイアウト変更に制約が生じる恐れがある。そこで GNSS や環境中に設置したランドマークを利用せずに自己位置推定を行うことを考える。この場合、環境の情報を記録する地図作成 (Mapping) を事前に行っておけば、地図を用いて効率よく自己位置推定を行うことができる[1]。そこで、初めのうちは自己位置推定問題と地図作成問題を同時に解き (Simultaneous Localization And Mapping, SLAM) 地図を作成した上で、その後は作成済みの地図を用いて自己位置推定を行うことを考える。

SLAM に適用可能なセンサは複数存在するが、特によく用いられるのは LiDAR とカメラである。位置推定精度の観点では LiDAR を使用した SLAM (LiDAR SLAM) が有利である場合が多い。一方で、カメラを使用する SLAM (Visual SLAM, V-SLAM) は LiDAR SLAM よりもセンサ価格を低く抑えやすい利点がある。また、V-SLAM のうち単眼カメラを用いるものの特徴として、単眼カメラのみでは地図の縮尺を求めることが不可能であり、他の手段で縮尺を別途求める必要がある点が挙げられる。

LiDAR とカメラの両方を組み合わせることで、カメラによる自己位置推定を高精度化する取り組みも存在する。Sujiwo ら[2]は主に自動運転車向けを想定した技術として、地図作成時に LiDAR とカメラの両方を使用する手法を提案した。具体的には、V-SLAM で作成した地図に保存したフレーム (キーフレーム) に対し、LiDAR による高精度自己位置を紐づける。地図作成完了後はカメラのみを使用して自己位置推定を行い、その際キーフレームに紐づけられた情報を用いて自己位置推定結果を補正する。これにより、地図作成後はカメラのみで高精度かつスケールの正しい自己位置推定が可能となる。

一方で、特に高精度な自己位置情報を必要とする自律移動ロボットには Sujiwo らの手法をそのまま適用すること

¹ 三菱重工業株式会社
Mitsubishi Heavy Industries, Ltd.

が困難な場合がある。Sujiwo らの手法は自動運転車を主な適用先として想定しているが、例えば無人フォークリフト (Automated Guided Forklift, AGF) はパレットへの爪挿しを行うため、さらなる高精度化が求められる。

Sujiwo らの手法は、LiDAR に由来する高精度自己位置情報を用い、キーフレーム位置において V-SLAM の地図を校正しているとみなすことができる。したがって、キーフレーム近傍であるほど高精度となると考えられ、キーフレームを密にすることで精度向上につながる可能性がある。一方でキーフレームを過密にすると地図容量と計算負荷が増大するため、適切な間隔でキーフレームが配置されることが望ましい。

そこで本稿では、地図作成時にキーフレームの間隔を制御することで、キーフレームを不必要に増加させることなく自己位置推定精度を向上させる手法を提案する。提案手法を用いた位置推定システムを図 1 に示す。

2 章で、キーフレームに紐づいた高精度自己位置を用いてカメラによる高精度自己位置推定を行う手法 (図 1 の①③④⑤) について述べる。3 章で提案手法であるキーフレーム間隔制御手法 (図 1 の②) について述べる。4 章で評価結果について述べ、5 章で本稿の結論を述べる。

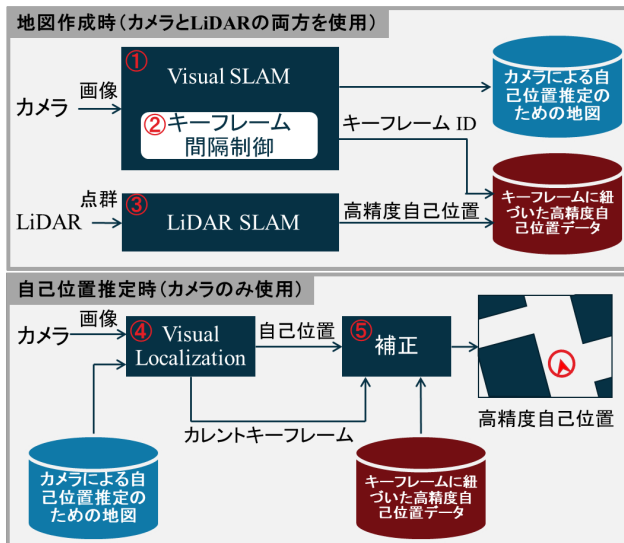


図 1 提案手法による位置推定システム。

2. カメラによる高精度自己位置推定

2.1 使用する SLAM アルゴリズム

地図作成時には、V-SLAM (図 1 ①) と LiDAR SLAM (図 1 ③) の両方を実行する。本稿では V-SLAM に ORB-SLAM2[3]を用い、LiDAR SLAM には NDT (Normal Distributions Transform) [4]による SLAM を利用する。ORB-SLAM2 は、V-SLAM に関連する主要な技術要素を網羅したオープンソースパッケージである[5]ORB-SLAM[6]

の拡張版であり、画像特徴に ORB 特徴[7]を用いる。NDT による SLAM には、Autoware[8]による実装を利用する。

地図作成後の自己位置推定時には、カメラによる自己位置推定 (図 1 ④) を行った後、[2]に基づく手法により自己位置を補正 (図 1 ⑤) する。この補正手法については 2.2 節で述べる。カメラによる自己位置推定には ORB-SLAM2 を利用する。

2.2 自己位置の補正

ある同一の点を座標系 $\{X\}$, $\{Y\}$ で表した位置ベクトル $\mathbf{p}_X$, $\mathbf{p}_Y (\in \mathbb{R}^4)$ (第 4 成分は常に 1 とする) について、 $\mathbf{p}_Y$ を $\mathbf{p}_X$ に変換する同次変換行列を $\mathbf{T}_{YX} (\in \mathbb{R}^{4 \times 4})$ と表記することにする。

$$\mathbf{p}_Y = \mathbf{T}_{YX} \mathbf{p}_X \quad (1)$$

同次変換行列について、任意の座標系 $\{X\}$, $\{Y\}$, $\{Z\}$ に対し、一般に以下の性質が成り立つ。

$$\mathbf{T}_{XZ} = \mathbf{T}_{XY} \mathbf{T}_{YZ} \quad (2)$$

また、 $\mathbf{T}_{YX}$ の回転行列部分を $\mathbf{R}_{YX} (\in \mathbb{R}^{3 \times 3})$ 、並進部分を $\mathbf{t}_{YX} (\in \mathbb{R}^3)$ と表すことにする。即ち

$$\mathbf{T}_{YX} = \begin{bmatrix} \mathbf{R}_{YX} & \mathbf{t}_{YX} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3)$$

本節で使用する座標系の種類と記号を表 1 にまとめる。また、各座標系と行列の関係を図 2 に示す。図中で赤の実線により示した変換 ($\mathbf{T}_{LW}$, $\mathbf{T}_{KW}$, $\mathbf{T}_{CW}$) は既知であり、赤の破線で示した変換 ($\mathbf{T}_{AW}$) は未知である。 $\mathbf{T}_{CW}$ が補正前、 $\mathbf{T}_{AW}$ が補正後の推定自己位置に対応する。 $\mathbf{T}_{LW}$, $\mathbf{T}_{KW}$ は地図作成後は既知である。緑の実線で示した 2 つの変換 ($\mathbf{T}_{CK}$, $\mathbf{T}_{AL}$) は、回転行列部分が等しく、並進部分は定数倍ずれている。

まず(2)式より、

$$\mathbf{T}_{CK} = \mathbf{T}_{CW} \mathbf{T}_{WK} \quad (4)$$

$$\mathbf{T}_{AL} = \mathbf{T}_{AW} \mathbf{T}_{WL} \quad (5)$$

が成り立つ。

$\mathbf{T}_{CK}$, $\mathbf{T}_{AL}$ は回転行列部分が等しいから、

$$\mathbf{R}_{CK} = \mathbf{R}_{AL} \quad (6)$$

となり、これと(4)(5)式より

$$\mathbf{R}_{CW} \mathbf{R}_{WK} = \mathbf{R}_{AW} \mathbf{R}_{WL} \quad (7)$$

となる。回転行列は直交行列であることに注意すると、 $\mathbf{R}_{AW}$ は以下のように求まる。

$$\mathbf{R}_{AW} = \mathbf{R}_{CW} \mathbf{R}_{KW}^T \mathbf{R}_{LW} \quad (8)$$

さらに、 $\mathbf{t}_{CK}$, $\mathbf{t}_{AL}$ は定数倍ずれた関係にあるから、スケールファクタを s とすると

$$\mathbf{t}_{AL} = s \mathbf{t}_{CK} \quad (9)$$

となる。本稿では s を以下の式で推定する。

$$s = \frac{\sum_i \|\hat{\mathbf{q}}_i - \hat{\mathbf{q}}_{center}\|}{\sum_i \|\mathbf{q}_i - \mathbf{q}_{center}\|} \quad (10)$$

ただし、 $\mathbf{q}_i (\in \mathbb{R}^3)$ は地図における i 番目のキーフレームの位置、 $\hat{\mathbf{q}}_i (\in \mathbb{R}^3)$ は i 番目のキーフレームに紐づいた高精

度位置, $\hat{q}_{center}$ と q_{center} はそれぞれ $\hat{q}_i, q_i (i = 1, 2, \dots)$ の重心である.

(4)(5)(9)式より

$$R_{AW}t_{WL} + t_{AW} = s(R_{CW}t_{WK} + t_{CW}) \quad (11)$$

が成立する. 一般に任意の座標系 $\{X\}, \{Y\}$ に対し

$$t_{YX} = -R_{XY}^T t_{XY} \quad (12)$$

となることに注意すると, t_{AW} は以下のように表せる.

$$t_{AW} = s(-R_{CW}R_{KW}^T t_{KW} + t_{CW}) + R_{AW}R_{LW}^T t_{LW} \quad (13)$$

以上(8)(13)式から T_{AW} が求まり, 補正後の自己位置推定結果を得る.

表 1 座標系一覧

座標系の種類		記号
世界座標系		$\{W\}$
キーフレーム座標系	ORB-SLAM2 による推定	$\{K\}$
	LiDAR SLAM による推定	$\{L\}$
カメラ座標系	補正前	$\{C\}$
	補正後	$\{A\}$

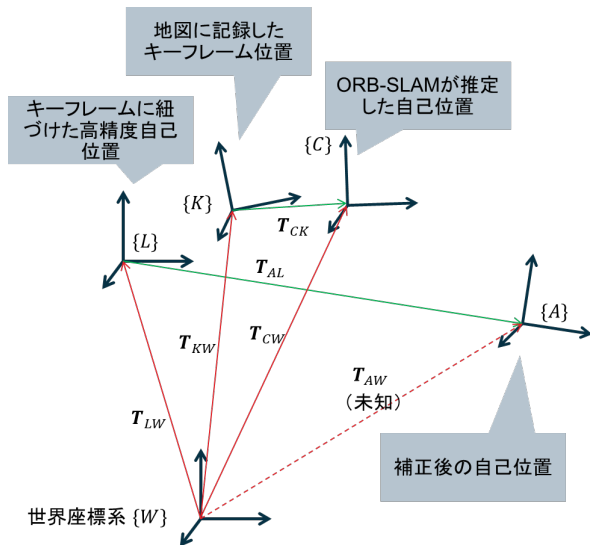


図 2 各座標系の関係

3. キーフレーム間隔制御手法

本章では, パラメータにより指定した間隔でキーフレームを挿入する手法 (図 1 ②) について述べる. キーフレーム間隔を制御するアルゴリズムのフローチャートを図 3 に示す. 図中の(1)(2)について以下で述べる.

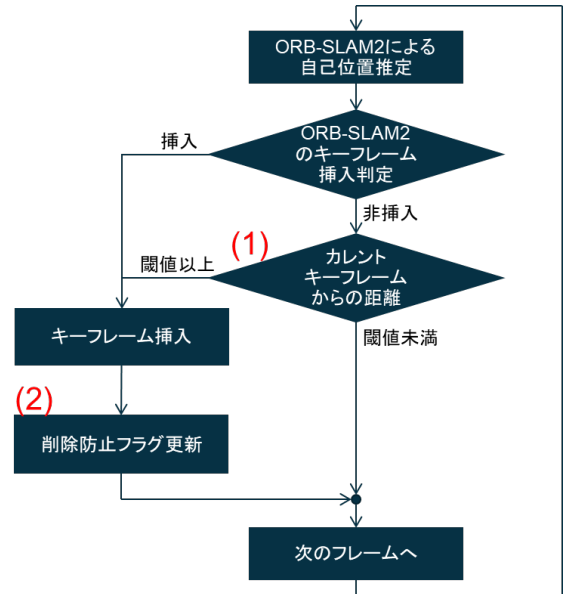


図 3 キーフレーム間隔制御のフローチャート

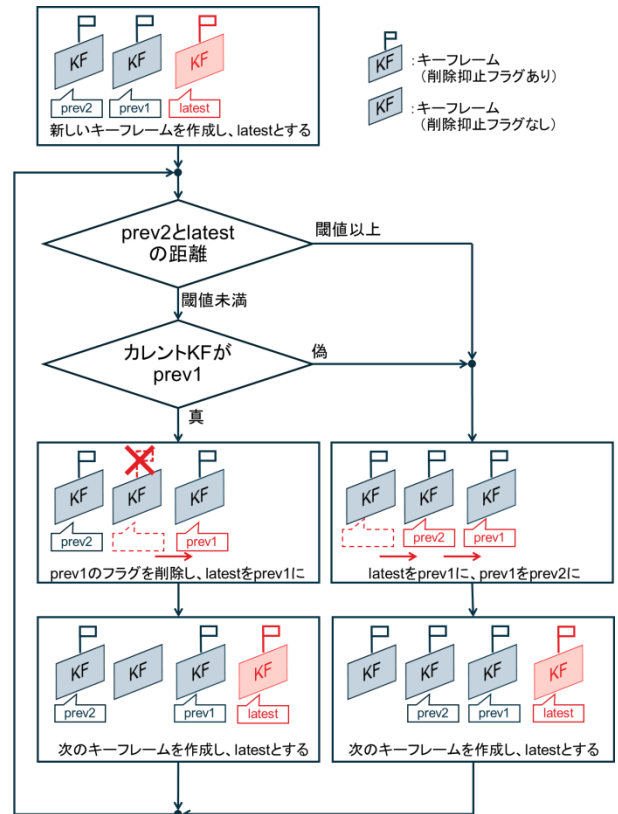


図 4 キーフレーム削除抑止フラグの管理方法

(1) 距離に応じたキーフレーム挿入

ORB-SLAM2 がキーフレームの挿入を不必要と判断しても, カレントキーフレームからの距離が閾値以上であればキーフレームを挿入する. ここでカレントキーフレームとは, 現在のカメラ画像と画像のパターンが最も近いキーフレームを, ORB-SLAM2 が地図中から選択したものである.

多くの場合、カレントキーフレームは現在の推定自己位置から最近傍のキーフレームである。これにより、ORB-SLAM2 のみではキーフレームが疎となっていた領域でも、パラメータで指定した間隔でキーフレームを生成できる。

(2) キーフレーム削除の抑止

(1)だけではキーフレームの間隔を保つうえで十分ではない。なぜなら ORB-SLAM2 は既知の特徴を多く含むキーフレームを自動的に削除するからである。

そこで、キーフレームに削除抑止フラグを割り当てる。このフラグが立っているキーフレームを削除処理から除外する。フラグの更新方法を図 4 に示す。新規に作成したキーフレームには予め削除抑止フラグを立てておき、削除しても間隔が閾値を上回らないと判断したキーフレームからフラグを除去する。これにより、キーフレームの削除によって間隔が大きくなることを抑止する。

4. 無人フォークリフトを用いた評価

4.1 評価方法

工場建屋内にて、AGF の実機にカメラ (STEREOLABS ZED), LiDAR (Velodyne VLP-16) を設置し、データ取得を行った。データ取得に使用したカメラはステレオカメラだが、評価には左眼画像のみを利用した。640 x 360 ピクセルの画像を 15 [fps] で取得した。

同じ経路を 2 回走行したデータのうち一方を用いて地図作成を行い、もう一方を用いて自己位置推定を行った。地図作成は、提案手法によりキーフレーム間隔を 200 [mm], 175 [mm], 150 [mm], 125 [mm], 100 [mm] とした場合と、キーフレーム間隔を制御しない場合の計 6 通りの条件で実施した。

自己位置推定精度を評価するため、レーザートラッカーを用いて AGF 軌跡を計測し、得られた軌跡を真値とした。レーザートラッカーを使用すれば、AGF に設置した専用マーカーを追尾することにより AGF 位置を高精度に取得することができる。ただし、AGF の走行中オクルージョンの発生によりレーザートラッカーの軌跡が一部欠損する。

そこで、レーザートラッカーの計測点のうち、欠損無く取得できたものを直線補完し、その上にカメラによる推定自己位置から垂線を下ろした (図 5)。垂線の長さの平均により自己位置推定誤差を評価した。

また、提案手法による地図の規模の変化について、作成した地図に含まれるキーフレームの枚数により評価した。

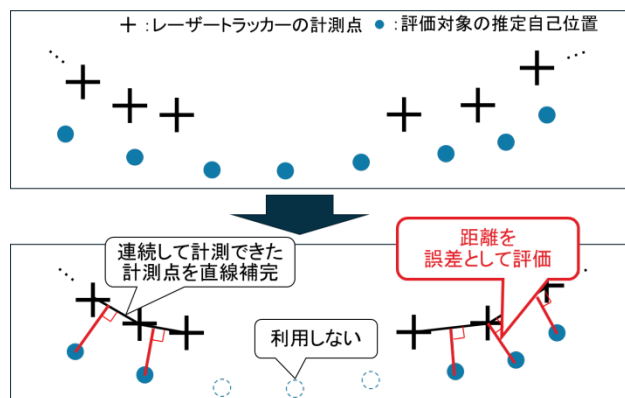


図 5 評価に使用する誤差

4.2 評価結果

試験の結果得た軌跡を図 6 に示す。本来は環状につながった軌跡であったが、レーザートラッカーで軌跡を得られなかった領域の軌跡を削除したため、3 つの軌跡に分割されている。

平均誤差の評価結果を図 7 に示す。キーフレーム間隔制御により、キーフレーム間隔制御を行わなかった場合と比較して平均誤差が最大約 21 [%] 減少した。平均誤差はキーフレーム間隔を 150 [mm] としたとき最小となり、約 22 [mm] であった。したがって、提案したキーフレーム間隔制御手法により位置推定精度を向上できることが分かった。

また、図 8 にキーフレーム個数の評価結果を示す。図 7 の誤差低減効果は 150 [mm] で向上が見込めなくなった一方で、キーフレーム間隔を小さくするに従いキーフレーム個数は単調増加した。このことから、キーフレームを増やせば増やすほど誤差が低減するわけではなく、ある値を超えて間隔を狭めても、誤差はそれ以上低減しなくなることが分かった。

以上の結果より、本稿の手法によって、キーフレームの間隔を適切な値に設定して地図作成を行うことで、不必要な地図の規模拡大を抑えつつ、自己位置推定精度を向上することが可能となることが分かった。

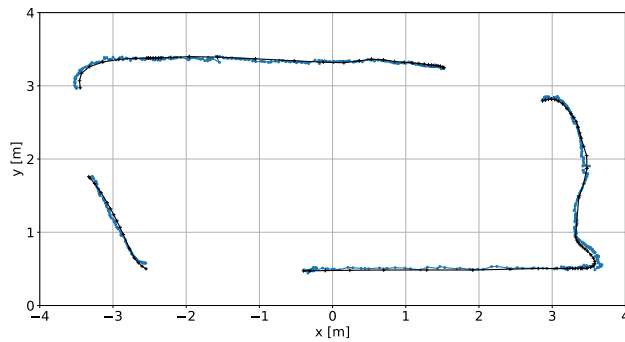


図 6 試験によって得られた軌跡。黒：直線補完後のレーザトラッカー計測点。青：評価対象の自己位置推定軌跡。

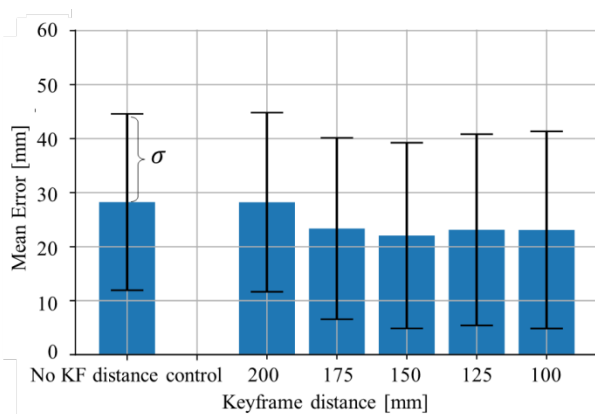


図 7 平均誤差評価結果

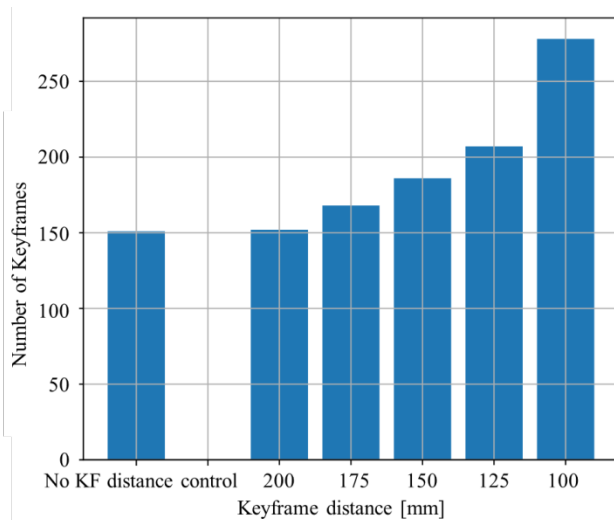


図 8 キーフレーム個数評価結果

位置推定精度を向上できることが確認できた。

参考文献

- [1] 友納正裕. SLAM 入門. オーム社, 2018, 3p.
- [2] Sujiwo, A. et al.. Localization Based on Multiple Visual-Metric Maps. IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 2017, p. 212-219.
- [3] Mur-Artal, R. et al.. ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. IEEE Transactions on Robotics, 2017, 33.5: 1255-1262.
- [4] Takeuchi, E. and Tsubouchi, T.. A 3-D Scan Matching using Improved 3-D Normal Distributions Transform for Mobile Robotic Mapping. 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, p. 3068-3073.
- [5] Taketomi, T. et al.. Visual SLAM algorithms: a survey from 2010 to 2016. IPSJ Transactions on Computer Vision and Applications, 2017, 9:16.
- [6] Mur-Artal, R. et al.. ORB-SLAM: a Versatile and Accurate Monocular SLAM System. IEEE transactions on robotics, 2015, 31.5, p. 1147-1163.
- [7] Rublee, E. et al.. ORB: An efficient alternative to SIFT or SURF. 2011 International conference on computer vision, 2011, p. 2564-2571.
- [8] Kato, S. et al.. AN OPEN APPROACH TO AUTONOMOUS VEHICLES. IEEE Micro, 2015, 35.6: 60-68.

5. おわりに

本稿では、地図作成時に LiDAR を援用することでカメラによる高精度自己位置推定を行う枠組みに、地図作成時のキーフレーム間隔を制御するアルゴリズムを組み合わせた自己位置推定手法を提案した。AGF を用いた評価の結果、提案手法によって不必要な地図規模拡大を抑止しつつ自己