

修士論文

題目

地上型レーザスキャナ点群と
ビデオ画像による Visual SLAM の精度評価

指導教員

松岡 真如 准教授

2023 年度

三重大学 工学研究科 情報工学専攻
データサイエンス研究室

新田 蓮 (422M521)

目次

1	はじめに	3
1.1	背景	3
1.2	目的	4
2	実験環境と使用機器・ソフトウェア	6
2.1	実験環境	6
2.2	基準点の配置	7
2.3	カメラの配置	7
2.4	カメラ	8
2.5	地上型レーザスキャナ	9
2.6	CloudCompare	9
2.7	ORB-SLAM3	10
3	研究手法	11
3.1	地上型レーザスキャナによるモデル作成	12
3.2	SLAM と 3 次元モデルの位置合わせ	13
3.2.1	SLAM	13
3.2.2	SLAM の位置合わせ	14
3.3	検証値作成の手法	15
3.3.1	SLAM のオフセット	15
3.3.2	カメラの同期	15
3.3.3	SLAM のトラッキング	16
3.3.4	カメラの姿勢推定	17
3.3.5	検証値のデータ作成	20
4	結果	21
4.1	カメラの内部評価	21
4.2	室内球の推定誤差	22
4.3	結合誤差と結合後マップ	23
4.4	SLAM の経路	23
4.5	SLAM の位置合わせ誤差	24
4.6	SLAM 球の推定位置	24
4.7	カメラの位置推定	25
4.8	SLAM の軌跡と検証値の比較	25
4.9	誤差のまとめ	26
5	考察	27
6	おわりに	29

参考文献	30
付録 A キャリブレーション	32

目 次

2.1	実験環境の屋内	6
2.2	配置された球	7
2.3	GoProHero8Black	8
2.4	SLAM 用のカメラ	8
2.5	FARO X330	9
2.6	ORB-SLAM の例	10
3.1	研究の流れ	11
3.2	地上型レーザスキャナによって得られた点群の一例	12
3.3	点群で取得された球	12
3.4	Fit Sphere 適用後	12
3.5	検証値の選択	14
3.6	オフセットの 3D モデル	15
3.7	カメラと画像によって求められるベクトル	16
3.8	外部標定によって求められた複数のベクトル	17
3.9	カメラの 2 次元位置を求めるイメージ	18
3.10	求められた基準点の位置と実際の位置のイメージ	18
4.1	再投影誤差の RMS の平均値と標準偏差	21
4.2	地上型レーザスキャナによって得られた結合後の点群	23
4.3	SLAM の経路と環境地図	23
4.4	求められた SLAM 球の位置	24
4.5	最適化後のカメラの視野	25
4.6	実際のカメラの視野	25
4.7	白線: 変換後の SLAM の経路 赤: 変換後の SLAM の点 緑: 検証値	25
4.8	赤: 変換後の SLAM の点	26
4.9	緑: 検証値	26

表 目 次

4.1	基準点の取得データ	22
4.2	各種変換の誤差	26

概要

近年、自己位置推定と環境地図生成を同時に行う Simultaneous Localization and Mapping (SLAM) が注目されている。映像のみを用いて行う Visual SLAM は、農場における作業履歴の記録などを低コストで行う事ができる。Visual SLAM を利用するには精度評価が必要である。

本研究では、地上型レーザスキャナとビデオ画像を用いて Visual SLAM の位置精度を評価した。研究室を実験環境とし、基準点として発泡スチロールの球 (以下、室内球) を 16ヶ所に配置した。また、地上型レーザスキャナを用いて室内の 3 次元点群モデルを作成した。SLAM のカメラの移動軌跡を記録するため、9 台のビデオカメラ (以下、室内カメラ) を配置した。この時、室内球が写るようにカメラを配置し、ビデオ画像上の室内球の位置から室内カメラの位置・姿勢を求めた。

Visual SLAM 用のカメラ (以下、SLAM カメラ) には、正四面体状に 4 つの球 (以下、SLAM 球) を取り付けた。SLAM カメラを台車に固定し、室内を移動して SLAM を行った。この時、室内カメラで SLAM カメラと 4 つの球を録画した。以下に示す手順で、3 次元モデル内の SLAM カメラの軌跡を決定した。はじめに、ビデオ画像に写った SLAM 球の中心の画素の位置を求めた。2 つの室内カメラで求めた SLAM 球の中心座標と室内カメラの 3 次元の位置・姿勢から、写真測量の手法を用いて SLAM 球の 3 次元位置を求めた。これを 4 つの SLAM 球に対して行うことで SLAM カメラの軌跡を求め、この座標を検証値とした。Visual SLAM で得られた軌跡を 3 次元モデルにおける位置に変換した。先述の検証値からほぼ等間隔で 3 点選び、同じ時刻の SLAM の自己位置と対応づけて変

変換行列を求めた。この変換行列を SLAM の軌跡に適用して、3 次元モデル上の位置へ変換した。

変換後の Visual SLAM の軌跡を求められた検証値と比較し、位置精度を評価した。精度評価には、位置の二乗平均平方根誤差 (RMSE) を用いた。Visual SLAM の位置の RMSE は 1.5 cm であった。地上型レーザスキャナでは、測距誤差が 5 mm であり、方位誤差は 0.015° である。SLAM カメラと室内カメラはキャリブレーションを行っており、すべてのカメラの再投影誤差の平均は 0.47 画素である。また、室内カメラの姿勢・位置推定後の室内球の位置と、求めたビデオ画像の室内球の位置の RMSE は 2.2 画素であった。また、ビデオ画像から手動で取得した球の中心は、手動調整により 2 画素のずれが生じると考えられる。以上の点を考慮すると、各変換・データ取得の検証値への影響は小さいものと考えられる。

レイトラッカーやモーションキャプチャカメラなどを用いた他の研究と比較して、本研究は汎用的なビデオ カメラを用いて SLAM の精度を評価できる。そのため、ビニルハウスなどの 3 次元モデルを作成し、定期的に SLAM の精度を評価しながら SLAM の手法を検討するといった利用法が考えられるであろう。

1 はじめに

1.1 背景

近年、Simultaneous Localization and Mapping (SLAM) が発展を遂げている。SLAM とは、カメラの映像やセンサの情報などから周辺の地図を生成すると共に自己位置や姿勢を推定する技術である [1]。Mur-Artal ら [2] は、大小、屋内外の環境でリアルタイムに動作する特徴点ベースの ORB-SLAM3 を発表した。ORB-SLAM は激しい動きや雑音に強く、一度通った箇所を再び通った時、全体の位置を修正するループクロージングと、自己位置推定できなくなった場合、同じ箇所を通った際に再び自己位置推定を始めるリローカリゼーションという機能を持つ。Long ら [3] は、動いている物体が存在する、特徴量の少ない屋内環境下に対する新しい高密度 RGB-D SLAM の手法を導入した。その SLAM は視野の広範囲が他の動いている物体で覆われる場合でも、SLAM を行いながら複数の物体の追跡を行う事ができる。位置推定、マッピング、動的セグメンテーション、物体の追跡において最先端の手法であり、カメラの蓄積誤差に対して頑健であることを実証している。

SLAM は、工事現場やロボットの自動走行などで利用されている。SLAM はマッピングと自己位置推定を可能にするが、正確なデータを取得するには姿勢を計測する慣性センサや周囲の形状を確認するレーザスキャナ、GPS など複数のセンサを必要とする。本研究ではカメラの画像のみから行う事ができる Visual SLAM に着目した。Visual SLAM を利用すれば、機器等のコストを減らす事ができ、SLAM の利用を普及する事ができる。例として、Krul ら [4] は、ビニールハウスや家畜の小屋にドローンベースの Visual SLAM

を用いて家畜や植物のモニタリングを行なっている。

これまで、筆者らは Visual SLAM による農地での作業位置の記録を目指し、3 次元モデル上に SLAM の移動経路を位置合わせした [5, 6]。見かけ上では SLAM の経路は概ね正しい経路を通っていたが、定性的な評価ができず、誤差の範囲が不明であることが課題となった。Visual SLAM の利用を促進するには、実用的な精度評価が必要となる。

実用的な精度評価を目指すには、実際に SLAM を利用する環境で精度評価を行うことができるデータセットが必要である。Helmberger ら [7] は、屋内外のさまざまな環境のデータセットである Hilti SLAM Dataset を発表した。これは高精度で信頼性の高い SLAM の開発に貢献している。検証値を測量機器であるトータルステーションを用いて取得している。また、Hilti SLAM Dataset で紹介され、かつ ORB-SLAM で使用されている、ASL Datasets [8] にも、レイトラッカーやモーションキャプチャカメラが利用されている。精度評価を行う際は、利用に合わせて、データセットを用意して検証することが必要であると考え。

本研究では、レイトラッカーやモーションキャプチャカメラなどを使わず、地上型レーザスキャナ点群とビデオ画像から SLAM の移動経路の精度評価を行うことを目的とした。

1.2 目的

本研究の目的は、ハンディカメラによる Visual SLAM の精度評価を地上型レーザスキャナによって得られた点群とビデオ画像を用いて行うことである。屋内で Visual SLAM を行い、それと同時にビデオ画像で SLAM のカメラを追跡して得られた座標を検証値として SLAM の経路と比較を行った。

物が散らばった環境下で行うことで実用的な精度評価を目指した。地上型レーザスキャナを用いた点群によって、正確な3次元モデルを生成し、基準点の位置や部屋の形状を緻密に取得することで、検証値を出来るかぎり正確に取得した。また、3次元モデル上に経路と検証値を表示することで可視性を得ることができた。Visual SLAM のデータセットを作成する際の機器などのコストが削減でき、各々の環境で精度を検証する際のデータセットの作成を容易にするものと考えられる。

2 実験環境と使用機器・ソフトウェア

2.1 実験環境

三重大学工学部情報工学コース、データサイエンス研究室を実験環境とした。やや広い空間で物が適度に配置されている。部屋の面積は約 13.5 m × 6.3 m であり、高さは約 2.6 m である。



図 2.1: 実験環境の屋内

2.2 基準点の配置

部屋には基準点として、16 個の発泡スチロールの球を配置した (以下、室内球)。直径約 15 cm と 20 cm のものを用いており、部屋の隅など全体的に配置した。これらは、後述の部屋に配置するカメラの位置合わせに使用した。



図 2.2: 配置された球

2.3 カメラの配置

部屋の隅などに部屋全体をカバーできるように 9 台のカメラを設置した (以下、室内カメラ)。図 2.1 は、カメラの内の一つの視点である。このカメラを用いて、SLAM を行っているカメラを追跡して検証値とした。

2.4 カメラ

カメラは、GoPro 社の GoProHERO8Black を用いた [9](図 2.3)。カメラは部屋に設置した 9 台と、SLAM 用の 1 台の計 10 台使用した。画素数は 1920×1080 画素、フレームレートは 60 フレーム/秒、手ぶれ補正を用いて撮影した。ピントが変わらないように焦点距離が固定の狭角モードを使用した。



図 2.3: GoProHero8Black

SLAM の経路を室内カメラによって追跡するために、SLAM 用のカメラ (以下、SLAM カメラ) に 4 つの球を取り付けた (以下、SLAM 球) (図 2.4)。球は 4 つとも半径約 9.5 cm で、SLAM カメラに正四面体状に取り付けた。水平方向の移動によるデータを検証するため、台車にカメラを固定して SLAM の撮影を行った。

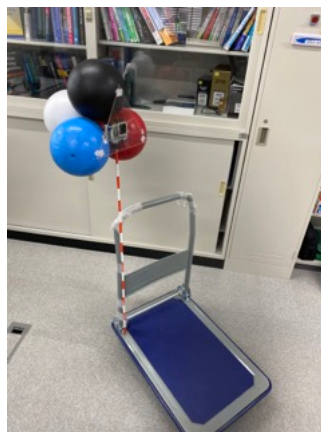


図 2.4: SLAM 用のカメラ

すべてのカメラはチェッカーボードを用いたキャリブレーションを行い、内部標定要素

を求めた。SLAM カメラには画像の中心位置、焦点距離、歪み補正值を利用した。室内カメラには、歪み補正值を利用しても効果が見られなかったため、画像の中心位置、焦点距離を利用した。

2.5 地上型レーザスキャナ

部屋の 3 次元モデルは、FARO X330[10] を用いて、地上型レーザスキャナによって作成した。距離の誤差は 5 mm であり、方位は 0.015° の誤差がある。

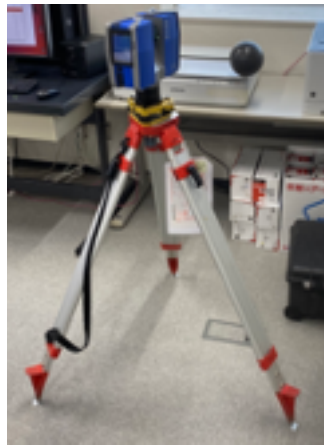


図 2.5: FARO X330

2.6 CloudCompare

CloudCompare は、点群データを編集可能なフリーソフトウェアである [11]。CloudCompare は、3 次元データの確認や室内球の 3 次元位置を得るために利用した。また、SLAM の座標系を 3 次元モデルと同じ座標系にする変換行列の出力に用いた。

2.7 ORB-SLAM3

SLAM のシステムとして、本研究では ORB-SLAM3 を利用した [12]。ORB-SLAM3 は、屋外や室内など様々な環境で扱える SLAM である。センサ付きカメラや複眼のカメラなどで動作し、本研究で使用する単眼カメラでも動作する。ORB-SLAM3 は、前後の画像の特徴点の動きから、カメラ座標を推定する。また、ORB-SLAM3 は、同じ箇所を通過することで、経路全体の累積誤差を修正できるループクロージング機能を持つ。ORB-SLAM3 では、撮影時のカメラの軌跡とカメラの向きを推定した。

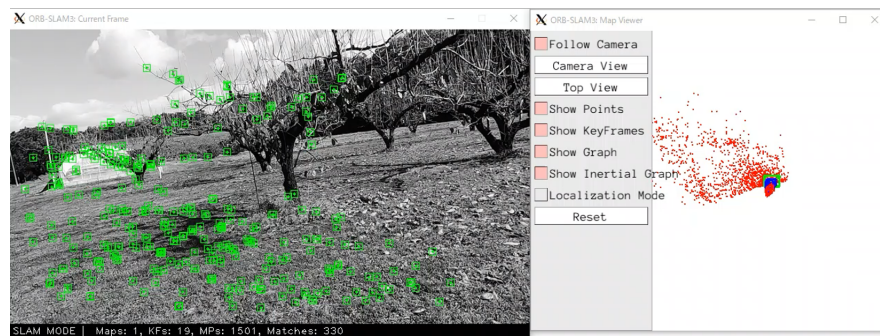


図 2.6: ORB-SLAM の例

3 研究手法

研究手法を図 3.1 に示す。

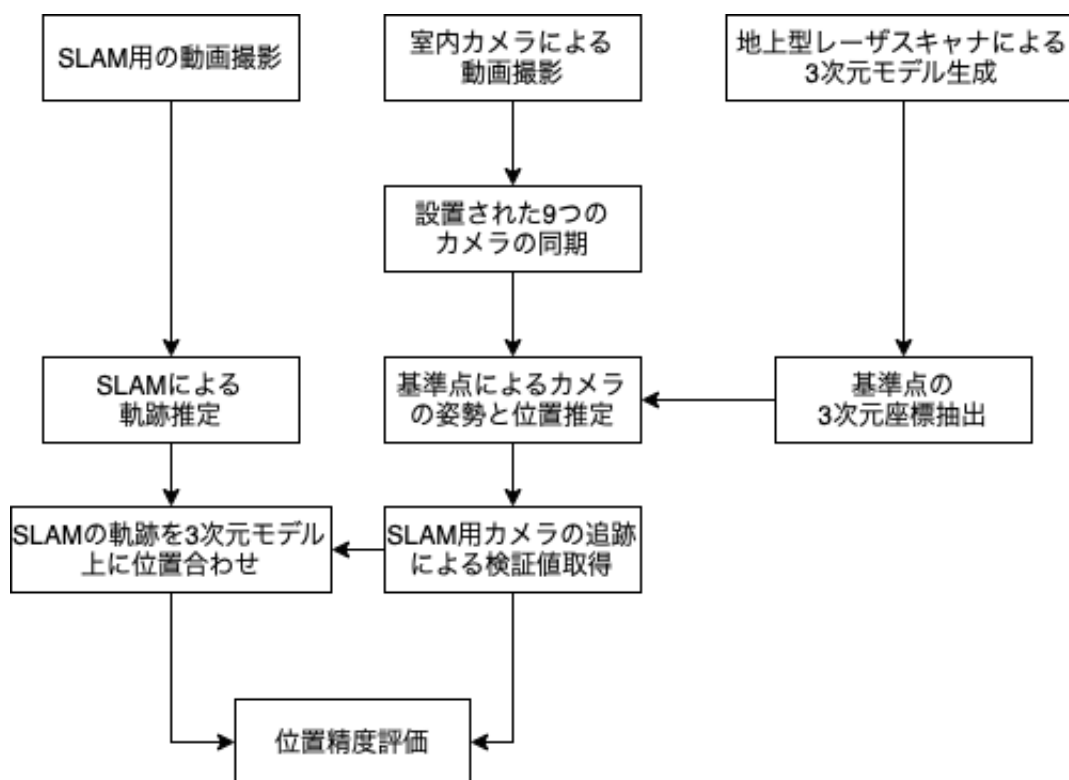


図 3.1: 研究の流れ

図 3.1 は研究全体の流れの図である。あらかじめ、地上型レーザスキャナで室内の 3 次元モデルを生成し、室内球の 3 次元座標を抽出した。SLAM 用の動画を撮影すると同時に室内カメラで撮影した。室内カメラをすべて同期させ、基準点を用いてカメラの位置・姿勢を外部標定を用いて最適化し、室内カメラで SLAM カメラを追跡できるようにした。室内カメラから SLAM 用のカメラを用いてそれを検証値とした。SLAM を行って、3 次元モデル上に位置合わせを行って、検証値と比較を行い精度評価した。

3.1 地上型レーザスキャナによるモデル作成

地上型レーザスキャナによって3次元モデルを生成した。部屋全体を3次元モデルにするために6つ地点で点群の取得を行った(図3.2)。

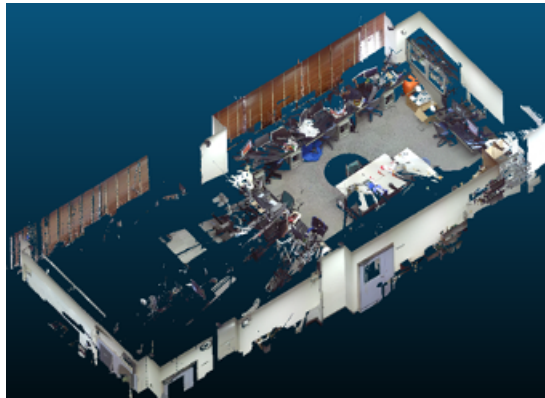


図 3.2: 地上型レーザスキャナによって得られた点群の一例

これらの点群の位置を合わせて合成するために室内球を使用した。6つの3次元モデルには室内球が点群として取得されている(図3.3)。CloudCompareの機能の1つであるFit Sphereを利用して室内球の3次元位置の中心座標を取得した(図3.4)。

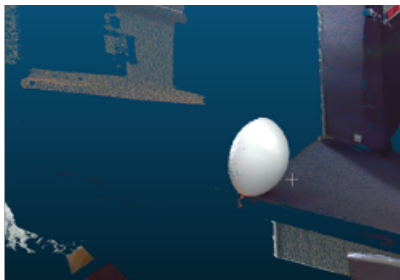


図 3.3: 点群で取得された球

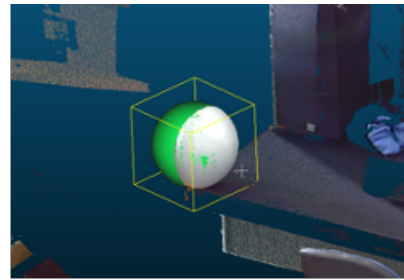


図 3.4: Fit Sphere 適用後

室内球の3次元位置を6つの3次元モデルで取得して、それらの位置の距離が最小になるように近づける変換を行って1つのモデルにした。

3.2 SLAM と 3 次元モデルの位置合わせ

3.2.1 SLAM

前処理として ORB-SLAM3 で実行する動画に対して変換を行った。ORB-SLAM3 で処理できるように動画を画像に分割した。60 フレーム/秒の動画で撮影を行っており、1 秒につき 60 枚の画像を生成した。ORB-SLAM3 に対応させるため画像サイズを 1920×1080 から、 720×423 に変換した。解析を容易にするために、ORB-SLAM3 のプログラムを編集した。ORB-SLAM3 は特徴点ベースである。特徴点とは画像の中の色の変化がはっきりしている点のことを指す。主に物体の角や境目、明るさの変化が大きい部分などが検出されやすく、平たい同色の色の部分は検出されにくい。特徴点ベースの SLAM とは、前後の画像の特徴点から共通する点を探り、その特徴点の移動からカメラの座標を求めるものである。その際、ORB-SLAM3 では、最初に取りべき特徴点の量を増やすことができ、その点を増やした。また、共通の点が少なくなると点を探す範囲が広がる。その範囲が広がる条件を緩めて、探索範囲を広くした。

ORB-SLAM には、カメラの内部評定要素を入力できる。チェッカーボードによるカメラのキャリブレーションの結果を用いて、画像の中心座標、歪み係数、焦点距離を設定した。

撮影された SLAM 用の動画に対して SLAM を行い、カメラの移動軌跡を取得した。今回は 47 秒の動画で 2834 フレームを用いた。

3.2.2 SLAM の位置合わせ

位置合わせには後述で求める検証値のうち3点を用いた。一定間隔になるように図 3.5 のように3点を選んだ。選び方は、検証値の偏りを考慮して、等間隔になるように 600、1300、1900 フレームを選んだ。CloudCompare の機能を使い、選んだ3点と対応する SLAM の3点が近づくような変換を行い、全体に適用した。これを検証値と比較して精度評価を行った。

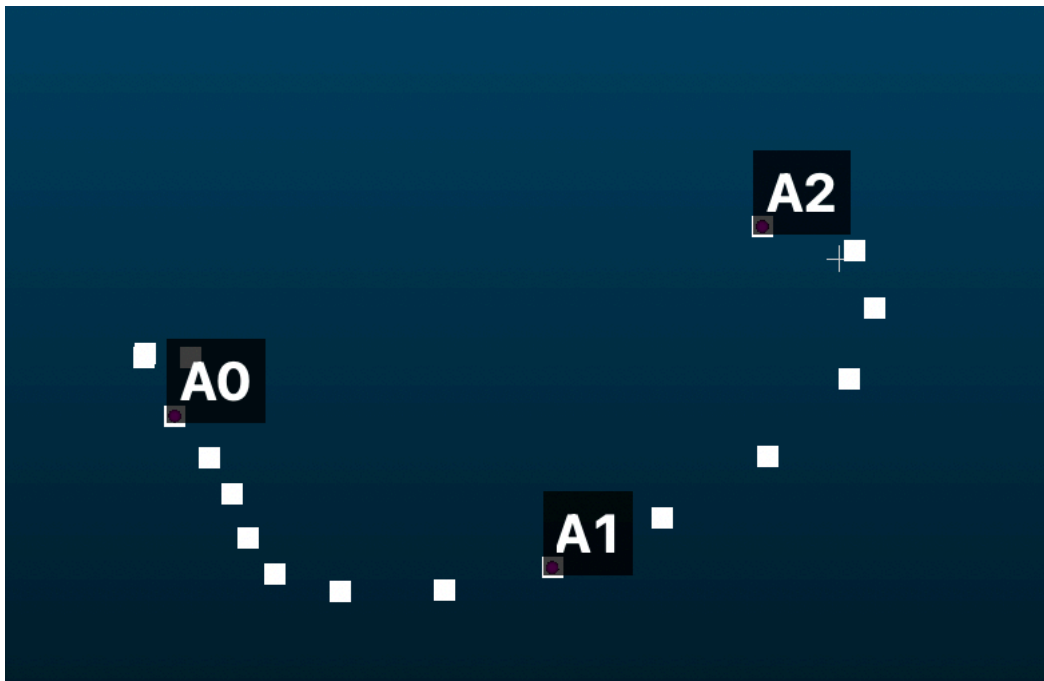


図 3.5: 検証値の選択

6

3.3 検証値作成の手法

3.3.1 SLAM のオフセット

SLAMで求められたのはカメラの位置である。比較を行うためには、SLAM 球の位置を求めて、球とカメラの位置関係を取得する必要がある。ipad に搭載されている LiDAR を用いて、Scaniverse[14] というアプリケーションによって、SLAM カメラと SLAM 球の 3 次元モデルを生成した (図 3.6)。室内球と同様に SLAM 球の中心を CloudCompare の Fit Sphere を用いて取得した。また、レンズの中心位置を手動で選択してカメラの中心として、球とカメラの位置関係を求めた。

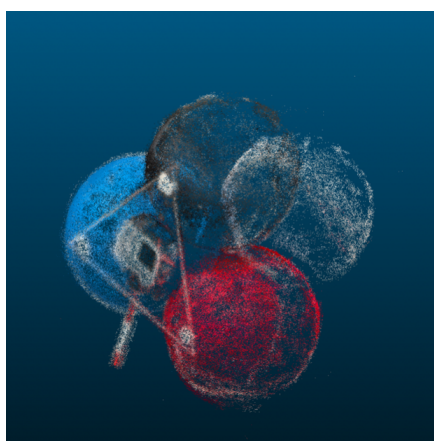


図 3.6: オフセットの 3D モデル

3.3.2 カメラの同期

9 台の室内カメラは SLAM を追跡するために同じ時間を映すように調整する必要があった。同一のストップウォッチを画面に映すことで同期を行った。ストップウォッチは 1/100 秒まで映すことができるものを利用した。ストップウォッチが写っているフレームをその時間として、時間が進む間隔が一定であることを確認するために、1 つのカメラにつき、

ストップウォッチの時間を動画全体の初め、中間、終わりの3回に分けて映した。カメラの時間とストップウォッチの時間を線形回帰して、全てのカメラの時間を同期させた。

3.3.3 SLAM のトラッキング

SLAM 用のカメラには、図 2.4 のように SLAM 球が取り付けられている。これを複数の室内カメラから 3 次元位置を求めることで追跡を行い、検証値とする。カメラの位置・姿勢とビデオ画像の 2 次元位置から、SLAM 球へのベクトルを求めた。図 3.7 のように、1 つのカメラとビデオ画像の SLAM 球の 2 次元位置より、球の 3 次元位置の中心へのベクトルが求められる。

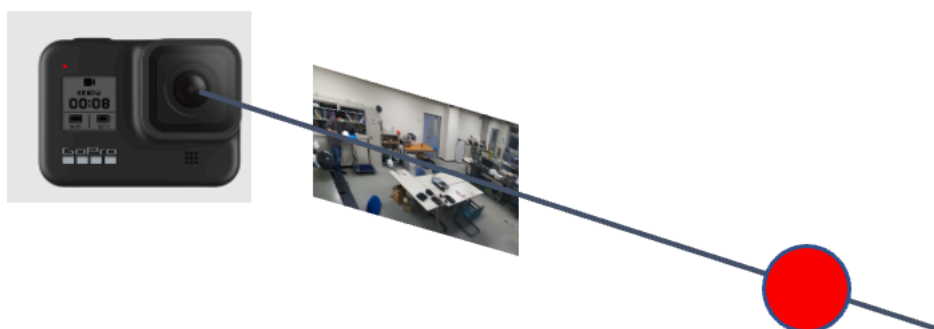


図 3.7: カメラと画像によって求められるベクトル

これを複数のカメラから求めることで、ベクトルはある1点に集中する(図3.8)。

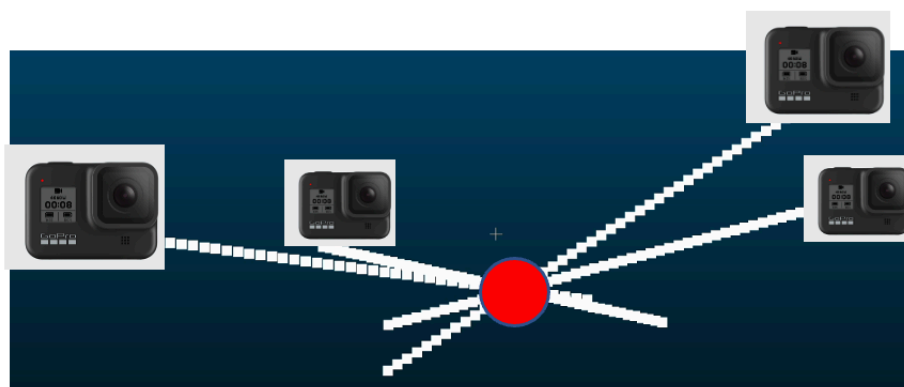


図 3.8: 外部標定によって求められた複数のベクトル

これを球の中心として、1つの画像につき、4つのSLAM球の3次元位置を求めた。今回は100フレームごとに球の位置を求めた。

3.3.4 カメラの姿勢推定

ベクトルを求めるにはカメラの位置と姿勢が必要である。9台の室内カメラの位置と姿勢を推定した。

3次元モデル上でのカメラの初期の位置と姿勢を大まかに与えた。解析写真測量[13]によると、カメラの姿勢と位置、室内球の3次元位置より、外部標定によってカメラの2次元での位置が求められる。図3.9のように3次元座標と姿勢を初期値として与えて、室内球の3次元座標によって、基準点の写真での2次元座標を求めた。

図3.10のように、カメラの位置と姿勢の初期値を大まかに与えたことによって、実際の球の位置と推定された球の位置は異なる。これらが近づくようにカメラの位置と姿勢の最適化を行う。

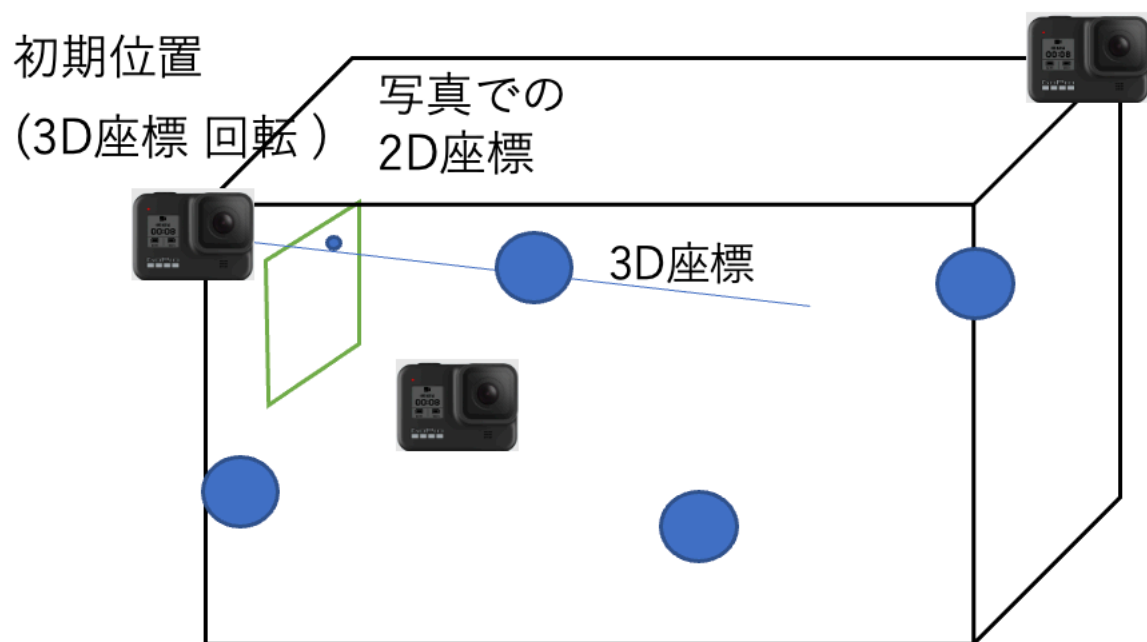


図 3.9: カメラの2次元位置を求めるイメージ

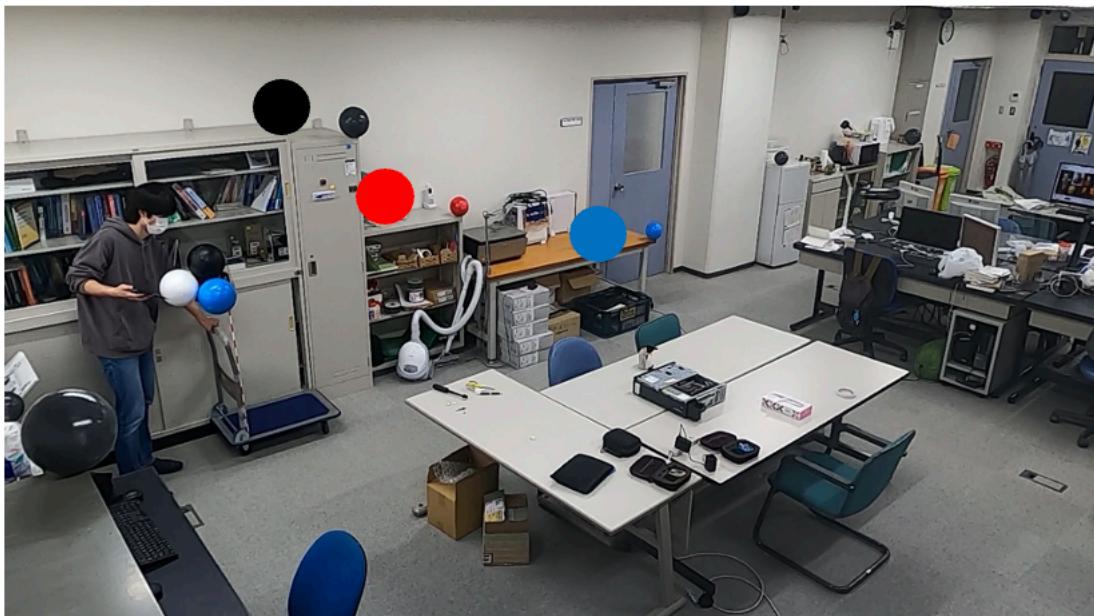


図 3.10: 求められた基準点の位置と実際の位置のイメージ

解析写真測量 [13] より、式 (3.1) から共線条件によって式 (3.2)、式 (3.3) が求められる。

式 (3.4)、式 (3.5) は、求められた値をデジタル画像上に変換する際の式である。

$$\begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{pmatrix} \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \kappa & -\sin \kappa & 0 \\ \sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X - X0 \\ Y - Y0 \\ Z - Z0 \end{pmatrix} \quad (3.1)$$

$$x' = -f_x \frac{x_p}{z_p} = -f_x \frac{a_{11}(X - X0) + a_{12}(Y - Y0) + a_{13}(Z - Z0)}{a_{31}(X - X0) + a_{32}(Y - Y0) + a_{33}(Z - Z0)} \quad (3.2)$$

$$y' = -f_y \frac{y_p}{z_p} = -f_y \frac{a_{21}(X - X0) + a_{22}(Y - Y0) + a_{23}(Z - Z0)}{a_{31}(X - X0) + a_{32}(Y - Y0) + a_{33}(Z - Z0)} \quad (3.3)$$

$$x = c_x - \frac{x'}{d_x} \quad (3.4)$$

$$y = c_y - \frac{y'}{d_y} \quad (3.5)$$

使用したカメラは、チェッカーボードを用いたキャリブレーションを行った。キャリブレーションによって求められた内部標定要素のうち、焦点距離と画像の中心の位置を使用した。 ω は x 軸の回転、 φ は y 軸の回転、 κ は z 軸周りの回転を表す。 (X, Y, Z) は基準点の 3 次元位置、 $(X0, Y0, Z0)$ はカメラの位置を表す。3 次元上の室内カメラの位置と 3 次元上の室内球と同じ直線状にビデオ画像上の室内球の中心は存在することを利用して、 f_x 、 f_y はカメラの焦点距離であり、 a は回転行列の各要素である。これにより、画像座標上の座標 x' 、 y' が求められる。 c_x 、 c_y はキャリブレーションによって求められた画像の中心の位置である。 d_x 、 d_y は 1 画素当たりのサイズ (実寸) である。これを用いて、2 次元画像上の位置 x 、 y に変換した。

この式により、室内球のビデオ画像上での位置を推定する。推定された複数の基準点の 2 次元位置と実際のビデオ画像の室内球の 2 次元位置と比較し、それらが近づくように式におけるカメラの位置と姿勢を Python の minimize 関数によって、室内カメラの位置と姿勢を求めた。

3.3.5 検証値のデータ作成

3.3.4 節で求められたカメラの姿勢・位置を用いて、3.3.3 節で SLAM 球の 3 次元位置を 100 フレームごとに求めた。SLAM 球の位置と 3.3.1 節のオフセットを用いてそれぞれの 4 つの球が重なるような変換を行った。同じ変換をカメラの中心にも適用し、カメラの位置として求めて、検証値とした。欠損箇所、SLAM の初期位置決定が終わるまでのデータを除き、18 地点の検証値を求め、SLAM と対応する位置を比較した。

4 結果

4.1 カメラの内部評定

チェッカーボードによるキャリブレーションを 90 回行った結果を図 4.1 に示す。評価指標として、再投影誤差を用いる。再投影誤差とは、3 次元空間上の点がカメラから投影された画像上の対応点と、実際の画像上の対応点との間の距離の誤差を表す指標である。すべてのカメラの再投影誤差の RMS の平均は 0.47 画素であった。

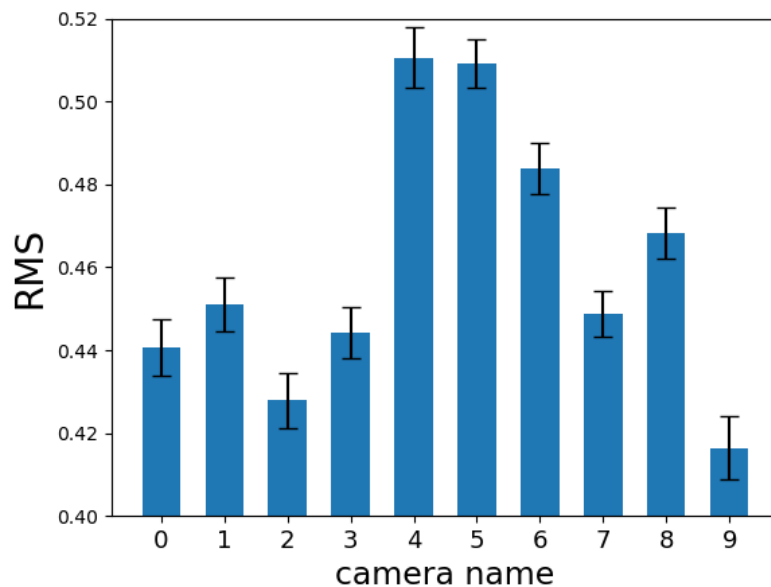


図 4.1: 再投影誤差の RMS の平均値と標準偏差

4.2 室内球の推定誤差

6つの3次元モデルで基準球の3次元位置をCloudCompareで取得した。表4.1は、それぞれのモデルで3次元座標を取得できた球の数と、3次元位置を取得する際に得られた球の3次元モデルと室内球の3次元モデルのRMSの平均である。

表 4.1: 基準点の取得データ

モデル番号	測定できた球の数	RMS [cm]
A	13	0.10
B	16	0.09
C	13	0.11
D	15	0.09
E	15	0.11
F	15	0.10

4.3 結合誤差と結合後マップ

モデル B で基準点がすべて測定できたので、B を基準として、基準点同士がモデル B の基準点に近づくように 5 つのモデル変換を行い、モデルの結合を行った。変換を行うと、基準点 1 つにつき、近づいた 6 つの基準点の中心が得られる。それらの距離の全組み合わせの RMSE は 5.4 mm であった。図 4.2 は、結合後の 3 次元モデルである。

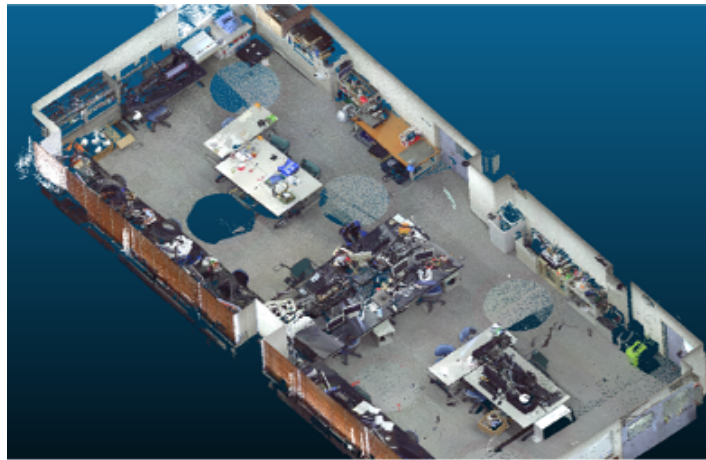


図 4.2: 地上型レーザスキャナによって得られた結合後の点群

4.4 SLAM の経路



図 4.3: SLAM の経路と環境地図

図 4.3 は、今回使用した動画から得られた SLAM の経路と環境地図である。

4.5 SLAM の位置合わせ誤差

SLAM の位置合わせは CloudCompare による 3 点の対応点を近づけるような変換で行った。変換行列は以下のように求められた。RMS は 5mm であった。

$$\begin{bmatrix} -4.276 & -0.216 & -4.171 & 5.81 \\ 4.17 & -0.120 & -4.275 & 4.86 \\ 0.070 & -5.97 & 0.236 & -25.20 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{bmatrix} \quad (4.1)$$

4.6 SLAM 球の推定位置

図 4.4 は、求められた 4 つの SLAM 球の 3 次元上での位置である。各色の球が各点の色と対応している。

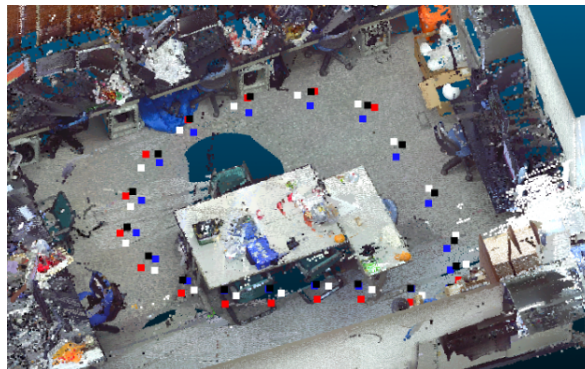


図 4.4: 求められた SLAM 球の位置

4.7 カメラの位置推定

図 4.8、図 4.9 はカメラの姿勢と位置を最適化した 3 次元モデル上の視野と実際のビデオ画像である。実際の基準点の位置とカメラの最適化後の基準点の位置の 2 次元位置の RMSE は 2.2 画素であった。

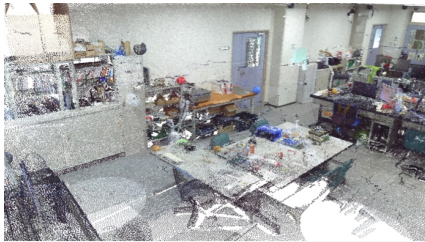


図 4.5: 最適化後のカメラの視野

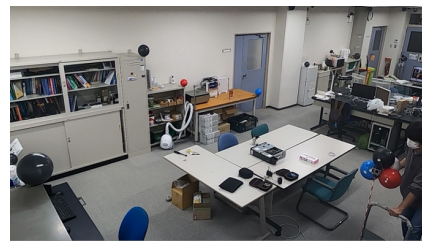


図 4.6: 実際のカメラの視野

4.8 SLAM の軌跡と検証値の比較

3 次元モデル上に表示された経路を図 4.7 に示す。

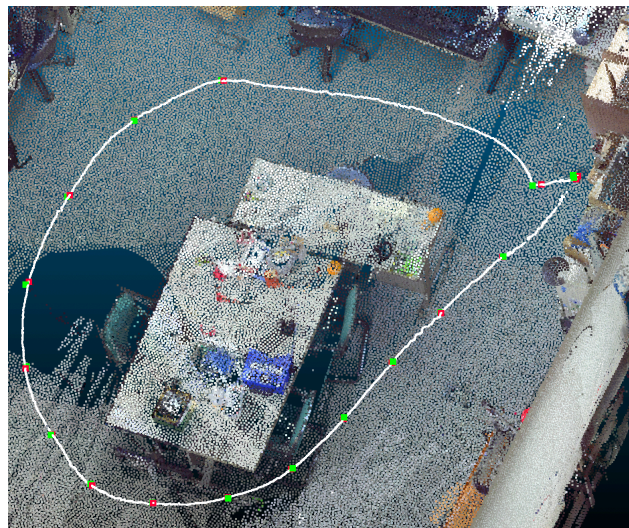


図 4.7: 白線: 変換後の SLAM の経路 赤: 変換後の SLAM の点 緑: 検証値

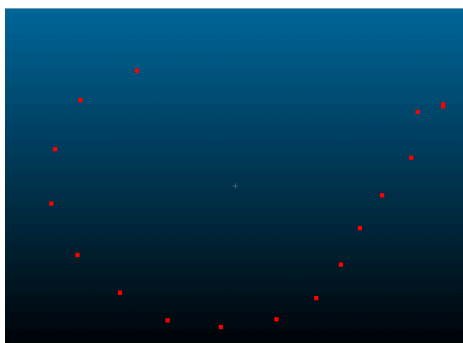


図 4.8: 赤: 変換後の SLAM の点

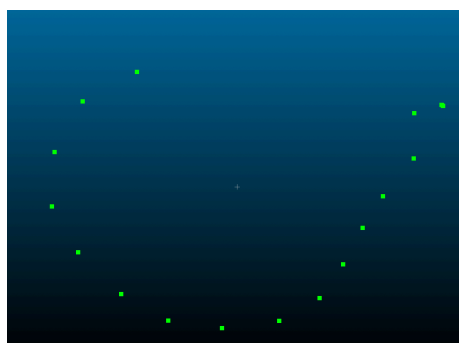


図 4.9: 緑: 検証値

白線は変換後の SLAM の経路を示している。また、緑の点はカメラから取得した検証値であり、赤は検証値と対応する変換後の SLAM の位置である。SLAM の経路は検証値の近くを通っており、経路も概ね実際に通ったものと同じである。位置合わせに使用した基準点を除いた 12 地点の SLAM と検証値との距離の RMSE は 1.5 cm である。

4.9 誤差のまとめ

各変換についての誤差を表 4.2 に示す。

表 4.2: 各種変換の誤差

球と経路の RMSE	1.5 cm
内部評定要素	RMS: 0.47 画素
地上型レーザスキャナの位置精度	測距:5 mm 方位:0.015°
基準点の 3 次元位置	RMS: 1 mm
モデルの結合	5.4 mm
SLAM の位置合わせ	5 mm
手動による 2 次元位置取得	2 画素
カメラ: 外部標定	(画像上の基準点) 2.2 画素

球と経路の RMSE が精度評価の結果である。他の項目は検証値を取得する際に生じる各変換で推定される誤差である。手動による球の 2 次元での中心座標取得は、画素値が球の中心になるように調整する際に 2 画素ほどのずれが生じる。

5 考察

ORB-SLAM による Visual SLAM の地上型レーザスキャナ点群とビデオ画像による精度評価を行った。検証値との誤差は RMSE 1.5cm と小さいことがわかった。検証値を求める際の各変換によって生じる誤差は、それぞれ小さい値を示していることから影響は少なく、想定される誤差は小さいものと考えられる。

室内カメラによる球の 2 次元位置取得は、画像の周辺の歪みの影響のより、ずれが生じていると考えられる。歪みの影響を調べる必要がある。SLAM 球の位置関係からカメラの位置を求めたが、実際のカメラの中心位置とは異なり、誤差に影響があるだろうと考えられる。SLAM がどこを自己位置であるカメラの位置としているかを考慮する必要がある。また、内部標定と外部標定を同時に求める方法として Zhang[15] のアルゴリズムがある。彼らのアルゴリズムは特徴点を利用するのに対して、我々は画像上と 3 次元モデルでの基準点の位置を用いている。特徴点が取得しにくい環境では本手法が利用できると考えられる。

SLAM カメラを台車に固定して行ったが、SLAM の球の位置がカメラに写っていれば検証点が得られるため、手持ちで行った際や速度に関わらず精度評価が可能であると考えられる。取得したデータの中には、台車に乗せた状態で今回の検証より長い距離を移動したものや、速度を変えたもの、台車を固定せず、手持ちで SLAM のカメラを動かして移動したものがある。それらの経路も検証し、今回と同等な精度評価が可能か検証する必要がある。

他の研究では、検証点を得る試手法として、レイトラッカーやモーションキャプチャカ

メラなどを用いている。それらを用いずにビデオカメラを用いることで、検証点の取得が容易になり、機器などのコストが削減できると考えられる。

カメラの位置と向きが分かれば、ビニルハウスなどの実験環境下で3次元モデルを作成することで、SLAMの精度を評価しながらSLAMの手法を検討するといった利用法が考えられる。実験規模によって、カメラの配置、室内球の数等を適切にすることで、効率的にSLAMの精度評価が行えるだろう。

6 おわりに

地上型レーザスキャナ点群とビデオ画像を用いて Visual SLAM の精度評価を行った。地上型レーザスキャナ点群によって、生成された 3 次元モデル上に SLAM の経路を位置合わせし、部屋に設置された基準点である球と室内カメラを元に SLAM を撮影している際の動きを追跡して検証値を求めた。SLAM の軌跡を 3 次元モデル上に位置合わせを行った経路と検証値を比較し、精度評価を行った。Visual SLAM の精度として、検証値との比較により RMSE を求めることができた。精度評価の結果について、各推定誤差をまとめ、本手法による誤差が小さいことがわかった。

今後は、様々な場所やパターンの軌跡を検証して、本手法の有用性を示すことが課題となる。ビニルハウスや農場、実際の利用者の動きを想定して精度評価を行い、実用性を示す必要がある。

参考文献

- [1] 友納 正裕, SLAM 入門: ロボットの自己位置推定と地図構築の技術, オーム社, pp.208, 2018.
- [2] Mur-Artal, R., Montiel, J., and Tardós, J., ORB-SLAM: A Versatile and Accurate Monocular SLAM System, IEEE Transactions on Robotics, 31(5), pp. 1147-1163. 2015, <https://doi.org/10.48550/arXiv.1502.00956>, (2024 年 3 月 6 日閲覧).
- [3] Long, R., Rauch, C., Zhang, T., Ivan, V., Lam, T., and Vijayakumar, S., RGB-D SLAM in Indoor Planar Environments with Multiple Large Dynamic Objects, IEEE Robotics and Automation Letters, 7(3), pp. 8209-8216, 2022, <https://doi.org/10.1109/LRA.2022.3186091>, (2024 年 3 月 6 日閲覧).
- [4] Krul, S., Pantos, C., Frangulea, M., and Valente, J., Visual SLAM for Indoor Livestock and Farming Using a Small Drone with a Monocular Camera: A Feasibility Study, Drones, 5(2), 41. 2021, <https://doi.org/10.3390/drones5020041>, (2024 年 3 月 6 日閲覧).
- [5] 新田 蓮, 松岡 真如, SfM と SLAM を用いた果樹園における移動履歴の記録.
- [6] 新田 蓮, 松岡 真如, SfM と SLAM を組み合わせた屋外における移動経路の記録, 日本写真測量学会, 令和 4 年度秋季学術講演会.
- [7] Helmberger, M., Morin, K., Berner, B., Nitish Kumar, Cioffi, G., and Scaramuzza, D., The Hilti SLAM Challenge Dataset, IEEE Robotics and Automation Letters,

- 7(3), pp. 7518-7525. 2022, <https://doi.org/10.1109/LRA.2022.3183759>, (2024 年 3 月 6 日閲覧).
- [8] ASL Datasets, <https://projects.asl.ethz.ch/datasets/doku.php?id=knavvisualinertialdatasets>, (2024 年 3 月 6 日閲覧).
- [9] GoPro, <https://gopro.com/ja/jp/shop/cameras/hero8-black/CHDHX-801-master.html>, (2024 年 3 月 6 日閲覧).
- [10] FARO X330 ユーザーマニュアル, <https://downloads.faro.com/index.php/s/4tpfLe8HeMGtiqM>, (2024 年 3 月 6 日閲覧).
- [11] CloudCompare, <https://www.danielgm.net/cc/>, (2024 年 3 月 6 日閲覧).
- [12] ORB-SLAM3, https://github.com/UZ-SLAMLab/ORB_SLAM3, (2024 年 3 月 6 日閲覧).
- [13] 長谷川博幸, 解析写真測量, 第 4 章 単写真標定, pp.46-56, 1995.
- [14] Scaniverse, <https://scaniverse.com>, (2024 年 3 月 6 日閲覧).
- [15] Zhengyou Zhang, A Flexible New Technique for Camera Calibration、IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pp.1330-1334, 2000. <https://doi.org/10.1109/34.888718>, (2024 年 3 月 6 日閲覧).

付録 A キャリブレーション

キャリブレーションを 90 回行った際の平均値を示す。なお、0 が SLAM に用いたカメラであり、他は設置されたカメラである。

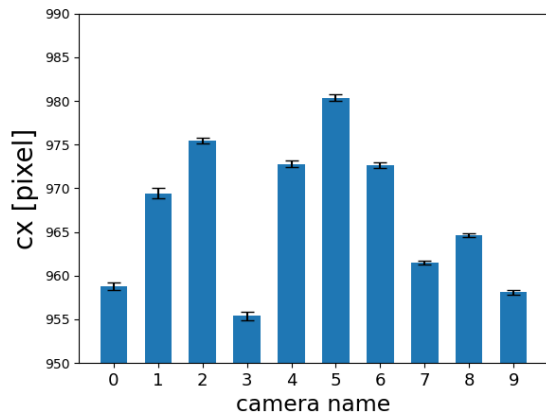


図 A- 1.1: 画像の中央位置:x

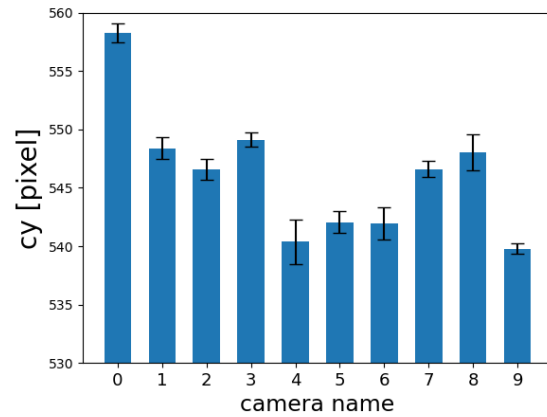


図 A- 1.2: 画像の中央位置:y

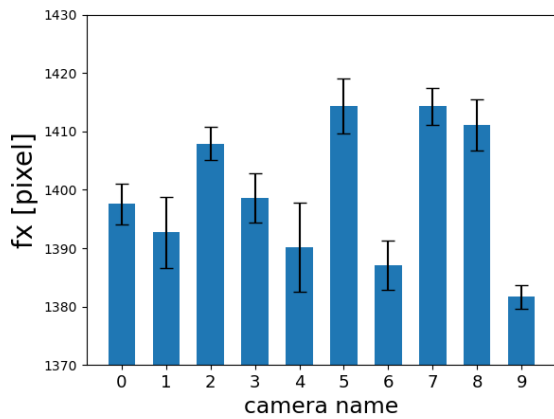


図 A- 1.3: 焦点距離:x

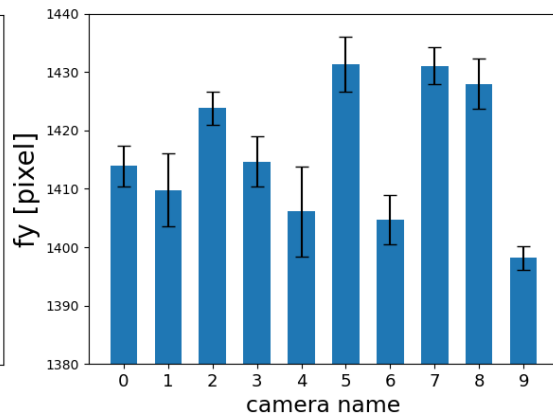


図 A- 1.4: 焦点距離:y

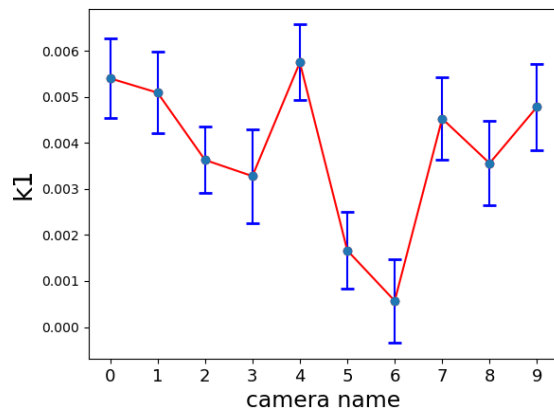


図 A- 1.5: 歪み補正值:k1

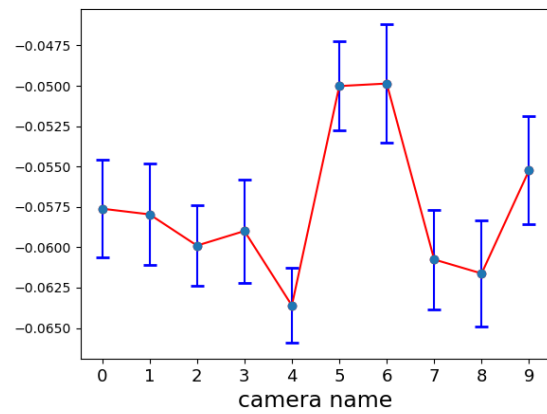


図 A- 1.6: 歪み補正值:k2

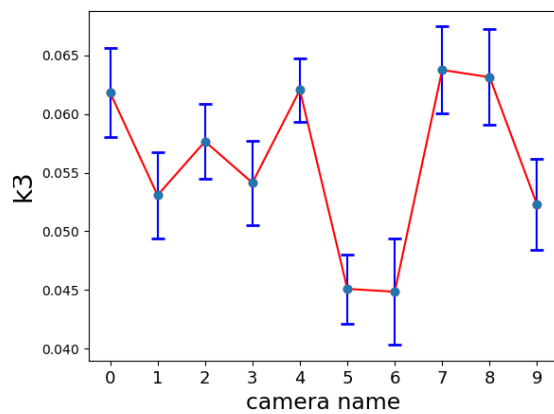


図 A- 1.7: 歪み補正值:k3

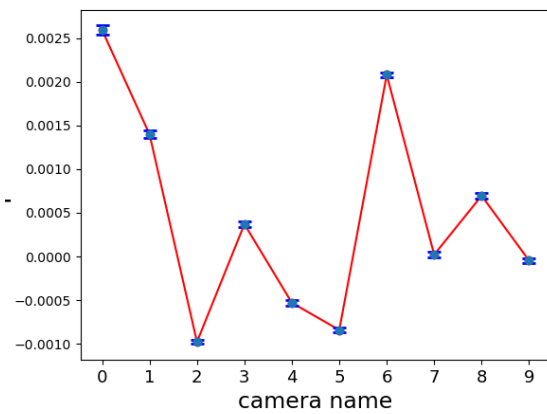


図 A- 1.8: 歪み補正值:p1

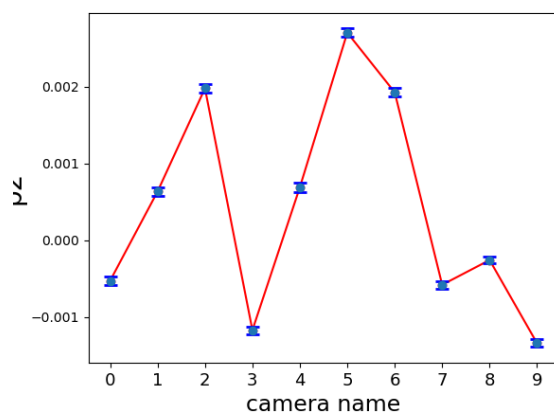


図 A- 1.9: 歪み補正值:p2

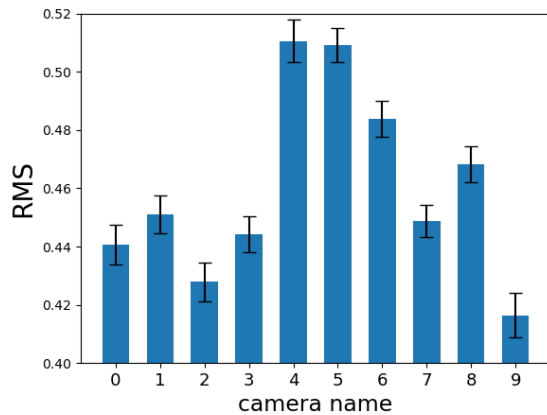


図 A- 1.10: 再投影誤差の RMS