

Title	特徴点マッチングにおけるモーション推定を改善するためのプレフィルタリング手法
Author(s)	宮浦, 啓太
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18331
Rights	
Description	Supervisor: 丁 洛榮, 先端科学技術研究科, 修士 (情報科学)

修士論文

特徴点マッチングにおけるモーション推定を改善するための
プレフィルタリング手法

2110165 宮浦啓太

主指導教員 丁洛榮

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和4年3月

Abstract

Most of the tasks accomplished by camera-equipped robot/mobile platforms are complicated and sophisticated (such as mapping, localization, navigation, and similar others). Most of these visual tasks are based on obtaining the motion of the camera/mobile platform. In the absence of any additional data apart from the visual data, motion estimation is done by using image matching (or image registration) methods. This method usually has been done through feature detection, description, and matching to establish correspondences between the images. But this procedure generates some mismatched points (e.g., due to noise), which are called outliers. There are some robust estimation methods (e.g., Random sample consensus (RANSAC)) to eliminate the outliers and estimate motion. The robust estimation methods occupy a large amount of computational cost whole image-matching pipeline due to their iterative procedure based on random sampling.

RANSAC is based on random sampling and threshold selection and it is one of the most widely known robust estimation methods in the image-matching framework since its easiness of implementation and high performance. But the computational cost of the method is high when the outlier ratio is high. Therefore, there have been proposed several methods to improve the RANSAC. Although they have improved their overall performance, their computational cost can still be high when the outlier ratio is large. In addition, there have been proposed several methods using geometric relationship filtering nowadays. For example, GR-RANSAC successfully removes outliers from the matched points under the assumption that the distances and angles between 2D features are similar for contiguously acquired images with small disparity, such as video images, leading to a reduction in the number of iterations. However, as noted by the authors, it is necessary to redetermine the angle and distance thresholds when using different datasets than the one used in the paper. Moreover, the authors also proposed GNN-RANSAC, which is a combination of GR-RANSAC and GNN. These methods are successfully reducing the number of iterations without modifying RANSAC. Therefore these methods are useful for other robust estimation methods in the image-matching framework.

We propose a novel pre-filtering step for eliminating the outliers obtained by established correspondences and reducing the number of iterations needed in the robust estimation method. Our pre-filtering step is motivated by the general procedure of visualizing the established matches among images. We tile images together and search for parallel lines connecting correspondences to eliminate the outliers.

In this research, we evaluate our pre-filtering step in extensive computational and comparative experiments using both simulated data involving along with real images and using a real dataset. Moreover, the GR_RANSAC algorithm is added

as a comparison baseline since it has a similar position in the image-matching pipeline to our proposed method. Our proposed method improves the motion estimation accuracy and the total number of iterations compared with conventional RANSAC. Moreover, GR_RANSAC's total number of iterations lower than our proposed method, but GR_RANSAC is mostly failed to estimate homography in our experiment. Therefore our pre-filtering step is considered more efficient and flexible compared with conventional RANSAC and GR_RANSAC.

目次

第1章	はじめに	1
1.1	背景	1
1.1.1	画像マッチング	1
1.1.2	Random sample consensus(RANSAC)	4
1.2	研究の目的	5
1.3	論文の構成	6
第2章	関連研究	7
2.1	RANSACの改良	7
2.2	フィルタリングを用いたRANSAC改良	7
2.3	本研究の位置付け	8
2.4	GR-RANSAC アルゴリズム	8
第3章	提案手法	14
3.1	モチベーション	14
3.2	アプローチ方法	14
第4章	実験・評価	19
4.1	計算機シミュレーションによる実験	19
4.1.1	実験環境	19
4.1.2	提案手法の評価方法	21
4.1.3	ホモグラフィ推定の成功率に関する結果	22
4.1.4	RANSAC 反復回数に関する結果	22
4.1.5	プレフィルタリングが失敗したケースの結果	23
4.1.6	プレフィルタリングがインライア比率を改善したケースの結果	25
4.2	実画像を用いた実験	26
第5章	おわりに	29

目 次

1.1	画像間のホモグラフィ	2
1.2	特徴ベース画像マッチングパイプライン	3
1.3	外れ値比率が増加した際の反復回数	6
2.1	フィルタリングステップ (GR_RANSAC, またはプレフィルタリング) を追加した, 特徴ベース画像マッチングパイプライン	9
2.2	(a) と (b) は論文で使用された画像 [20]. (c) と (d) はそれぞれの画像における特徴点 [8]. 幾何学的関係の違いは, 各画像における基準点と残りの全ての対応点の距離と角度を計測することで判断され, 事前に定義された閾値より大きい場合アウトライアとして識別される.	10
2.3	GR_RANSAC アルゴリズムフローチャート	11
2.4	(a) 画像 1 における仮の基準点 (b) 画像 2 における仮の基準点 (c) 基準点ペア. 基準点は, 各画像において仮の対応点と残りの対応点ペアからランダムに n 点抜き取り, その距離と角度を計測することで選ばれる. 計測された, 距離と角度が閾値より小さい場合, 基準点として決定づけられる. 我々は, n を 5 点で選択した.	12
3.1	画像をタイル状に並べた際の対応関係の可視化, 赤線: インライア, 緑線: アウトライア	14
3.2	(a) 従来手法で確立された対応関係. 14 個の対応関係の内, 9 個がインライア. (b) フィルタリングを追加して確立された対応関係. 10 個の対応関係の内, 9 個がインライアで, 1 個がアウトライア.	16
3.3	(a) 従来手法で確立された対応関係. 14 個の対応関係の内, 9 個がインライア. (b) フィルタリングを追加して確立された対応関係. 7 個の対応関係の内, 7 個がインライア.	17
3.4	(a) 従来手法で確立された対応関係. 14 個の対応関係の内, 9 個がインライア. (b) フィルタリングを追加して確立された対応関係. 9 個の対応関係の内, 9 個がインライア.	18
4.1	(a): 実験で使用した第 1 画像. (b) ~ (g): (a) に対して実験で使 用した変換を施したワープ画像例	20
4.2	誤差計算	21

4.3	プレフィルタリングステップを用いて，ホモグラフィ推定に失敗した際のインライア比率に関するヒストグラム．従来の RANSAC のみでホモグラフィを推定した際でも，この図で報告されている全事例の内 70.17% で失敗している．対応点が総数が 100 個の場合，インライア比率の違いによる失敗したケースが最も多くなっている．このことから，我々の提案手法は，インライア比率よりもインライアの総数が重要であると結論付けられる．	24
4.4	実験に使用した実画像	27
5.1	災害現場 [25]	30
5.2	開発中の Hexapod robot(日本文理大学伊藤研究室提供)	31

表 目 次

1.1	外れ値比率に対する反復回数	5
4.1	実験で使⽤したパラメータ	21
4.2	ホモグラフィ推定の成功率に関する結果のまとめ	23
4.3	RANSAC 反復回数に関する結果	25
4.4	平均計算時間の結果 (秒)	25
4.5	提案手法の成功と失敗したケースの結果	26
4.6	実画像を⽤いた実験による結果	28

第1章 はじめに

1.1 背景

1.1.1 画像マッチング

センシング技術の発達により、カメラを搭載したモバイルプラットフォームやロボットが身近になりつつある。日常生活における利用を向上させるためには、周囲の環境を認識・理解する能力が必要不可欠である。そのため、認識した情報の処理や、様々な視点から得られた情報の関係性や違いを理解することは、コンピュータビジョン、リモートセンシング、パターン認識等の様々な分野で重要なテーマとなっている。このような複雑なタスクにおける、重要かつ基本的な問題として画像マッチング（または画像登録）がある。

画像マッチングは、シーンがほぼ平面であると仮定して、重なり合う領域を持つ2枚の画像を重ね合わせる処理である。これまでに多くの手法が提案されてきたが、統一的なフレームワークを構築することは未だ困難な課題である [16]。そのため、提案されている手法の中でも、対象となる目的や画像データの特徴に応じて、適切に手法を選択する必要がある。画像を重ね合わせるには、画像の座標フレーム（左上角が原点）間の座標変換を計算する必要がある。この座標変換は、カメラや、モバイルプラットフォームの動きを表すこともできる。マッピング、ローカライゼーション、ナビゲーション、3D再構成等のロボット工学やコンピュータビジョンの分野における多くの高度なタスクは、モーション推定に依存しているためこの画像マッチング技術は重要である。また、他のデータソース（例えば、ナビゲーションセンサー、LIDAR等）がない場合、撮像された画像が唯一のソースとして残る。よって、それらのデータから、正確にモーション推定することは、これらの高度な方法において重要な役割を果たす。

画像座標フレーム間の座標変換は、シーンがほぼ平面である場合（図 1.1）や、カメラとシーンの距離がシーン内の3次元要素（建物等）に比べて大きい場合、平面射影変換（ホモグラフィ）で表現される。また、平面射影変換行列は、8自由度を持つパラメータで記述することが出来る。このことから、重なり合う画像間の変換を求める方法として、画素の輝度勾配から直接ホモグラフィを推定する特徴レス法と、画像から抽出した特徴点のマッチングによりホモグラフィを推定する特徴ベース法の2種類のフレームワークが良く用いられてきた [13]。

特徴レス法では、画素の輝度勾配を用いてホモグラフィを推定するため、特徴

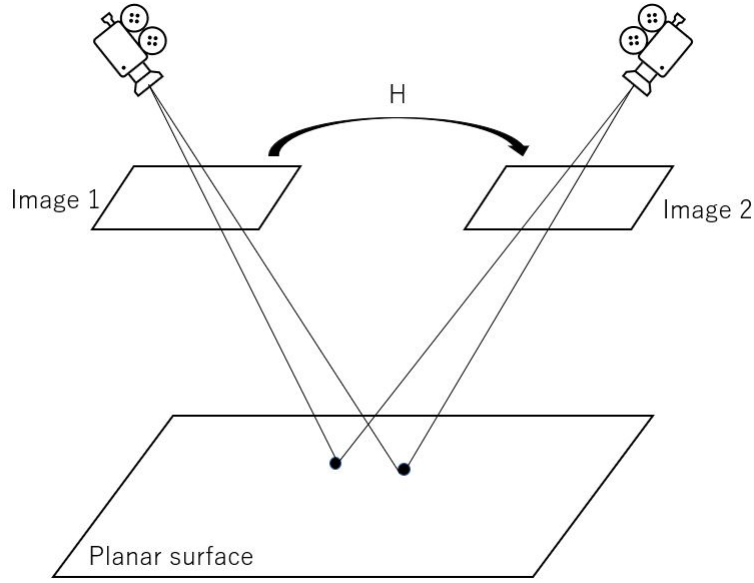


図 1.1: 画像間のホモグラフィ

ベース法のような前処理 (特徴点を用いた対応関係の確立) が不要である。しかし、この手法は、画像間のばらつき (視差) がほとんど存在しない場合を前提としている。更に、ホモグラフィのパラメータは、画像の輝度値と非線形な関係にあるため、複雑な最適化アルゴリズム (例えば、ガウス・ニュートン法) が必要になるといった問題がある。

特徴ベース法は、画像から特徴点を抽出し、2つの画像間に確立された対応点 (少なくとも4組の対応点) からホモグラフィを推定する。これは、(式 1.1) で表すことが出来る。Image1 (入力画像1), Image2 (入力画像2) をそれぞれ重複領域を持つ画像とし、Image1 上の点を (x, y) , Image2 上の点を (x', y') とし、 λ は任意の非ゼロのスケールファクターである。この方法の利点は、ホモグラフィパラメータを特徴点座標との線形関係でも表現することが出来るため、例えば疑似逆行列を用いることで容易にホモグラフィパラメータを求めることが出来る点である。また、画像間の視差が大きい場合にも有効な手法である。一方で欠点は、特徴点の抽出に依存し、対応点に多くの外れ値が含まれる場合、ホモグラフィパラメータの精度が低下しやすい点である。これらのことから、特徴ベース法は、コンピュータビジョン、ロボット工学、リモートセンシングの分野で、マッピング、ナビゲーション、ローカライゼーションの目的で広く利用されている。

$$\begin{pmatrix} \lambda x' \\ \lambda y' \\ \lambda \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1.1)$$

この 20 年間にわたり、Scale invariant feature transform (SIFT)[15] や Speeded up robust features (SURF)[2] の特徴点の検出と記述に関する開発により、例えば、様々なスケールや視点の変更等の極端な場合でも画像間の変換を計算することが出来るようになった。このことから、特徴ベース法は広く使われるようになった。また、特徴の検出と記述、マッチング [30, 16] のため、様々な深層学習ベースの手法も提案されており、その比較ベンチマークが [1] に示されている。[7] では局所特徴の検出と記述のための D2-Net フレームワークが提案されている。D2-Net は、画像マッチングでは限界があるが、ローカライゼーションでは他の手法を凌駕する性能を発揮する。これらの背景から、本研究では特徴ベース法に焦点を当てて進めていく。尚、特徴ベース法のパイプラインを (図 1.2) に示す。

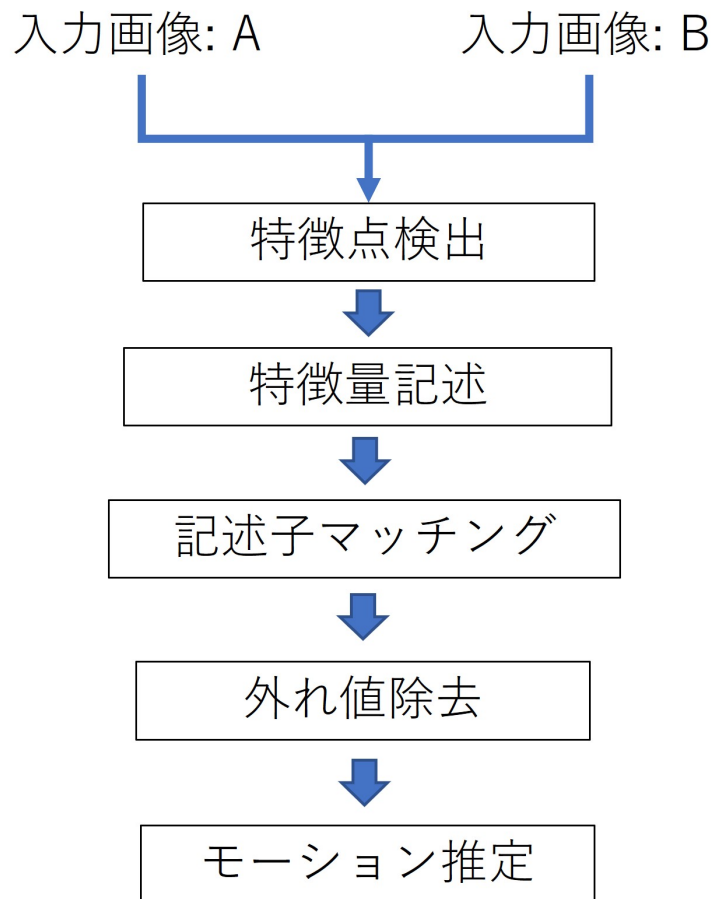


図 1.2: 特徴ベース画像マッチングパイプライン

まず、特徴に基づく画像マッチングは、各画像の特徴的な点（エッジ、コーナーなど）を検出する。次に、その特徴点の近傍の画素値を元に、特徴ベクトルを生成する。この特徴ベクトルのことを記述子（ディスクリプター）と呼び、この処理のことを特徴量記述と呼ぶ。そして、この記述子間のユークリッド距離を比較することで、マッチングを行い、画像間の対応関係を確立する。しかし、このマッチン

グ手順の結果は、例えば、ノイズや、視点変更の影響等により、いくつか誤った対応関係が含まれる。これは、確立された対応関係が画像間のカメラの動き（あるいは変換）に従わないものがあることを意味する。このようなミスマッチは、外れ値（アウト라이어）と呼ばれている。外れ値が存在する場合、ホモグラフィを正しく推定することが出来ない。そのため、外れ値を除去して、正しいホモグラフィを推定するために、Random sample consensus(RANSAC)[11] や Least median of squares regression (LMeds)[22] 等のロバスト推定法が利用されている。その中でも、実装が容易であること、性能が高いことから RANSAC が広く使われている。しかし、ロバスト推定法は反復手法であるため、マッチングパイプライン全体の計算コストの大半を占めていることが知られている。特に、視点が大きく変化する場合や、特徴点の検出やマッチングが困難な画像（テクスチャがない画像、ブレが含まれているような画像等）において問題となる可能性が高い。

そこで、本研究では、ロバスト推定法に必要な反復回数を削減することを目的とする。我々の提案手法を紹介する前に、以下に RANSAC の基本的な事項を紹介する。

1.1.2 Random sample consensus(RANSAC)

RANSAC は、与えられたデータから外れ値を取り除き、モデルパラメータを正確に推定することを目的としたロバスト推定法の 1 つである。以下に、オーバーラップする一対の画像間の対応関係から RANSAC を用いた、動き推定について述べる。

Step 1: マッチングにより、確立された対応点の中から、パラメータの推定に必要な最低限の数 n 個をランダムに抽出する。

Step 2: n を用いて、パラメータを推定する。

Step 3: Step1 で選んだ点以外のすべての点を、Step2 で得られたパラメータにあてはめ、そこから得られたデータと元々あったデータの誤差を比較する。

Step 4: 誤差が閾値 t 以下であれば、インライアとして数をカウントする。

Step 5: Step1 ~ 4 を繰り返し、インライアの数が多かった際のモデルパラメータ（ホモグラフィ行列）とその際のインライアを抽出する。

画像マッチングでは、外れ値を決定するために、特徴点位置間のユーグリッド距離に基づく閾値 t が用いられる。本研究では 2 次元で最も一般的な 8 自由度のパラメータを持つホモグラフィ推定に着目している。従って、サンプルサイズ n は 4 に等しい。RANSAC は反復手法であるため、上記のステップは最大反復回数 k ま

で実行される．この k は任意に設定することが出来る． k の理論的な限界は，(式 1.2) から計算できる．

$$k = \frac{\log(1-p)}{\log(1-\omega^n)} \quad (1.2)$$

$$\omega = \frac{\text{インライアの数}}{\text{マッチングにより得られた対応点の総数}}$$

p : n 個サンプリングされた中に少なくとも 1 つ外れ値が含まれていない確率

ω : インライア比率

外れ値比率が異なる場合の k の値を (表 1.1, 1.3) に示す．尚、ここでは、 $p = 0.99$, $n = 4$ で計算している。

外れ値比率 ($1-\omega$)	最大反復回数: k
0.1	4.3
0.2	8.7
0.3	16.8
0.4	33.2
0.5	71.4
0.6	177.6
0.7	566.2
0.8	2875.9
0.9	46049.4

表 1.1: 外れ値比率に対する反復回数

1.2 研究の目的

RANSAC [11] で必要とされる反復回数は、指数関数的に増加する．そこで、我々は、ロバスト推定法の性能向上を目的とした、新しいプレフィルタリングステップを提案する．プレフィルタリングステップは、外れ値の総数を減らすことを目的としているため、これは当然ながら、インライア比率の向上に繋がる．我々が提案するフィルタリングステップは、RANSAC だけでなく、画像マッチングの枠組みの中で他のロバスト推定法とともに併用することが出来る．提案手法は、重複領域のある画像をタイル状に並べた際に、インライアの多くは、通常平行線を形成することから動機づけられ、画像をタイル状に並べた際の対応関係を結ぶ線の角度ヒストグラムを用いて、インライアとアウトライアを識別するものであり、[18] の並進変換を斜め方向に拡張したものである．また、我々は、本アルゴリズムと従

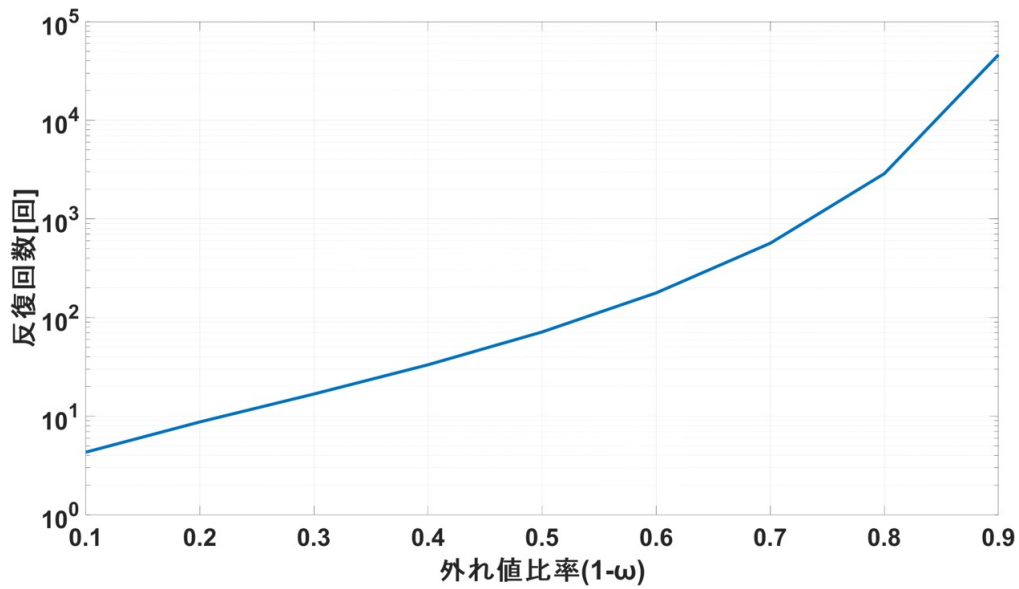


図 1.3: 外れ値比率が増加した際の反復回数

来 RANSAC を用いた場合と，GR_RANSAC アルゴリズム [8] を比較し，計算機シミュレーション実験，及び，実画像を用いた実験において本アルゴリズムの有用性を示す．GR_RANSAC [8] アルゴリズムの詳細については，第 2 章で説明する。

1.3 論文の構成

第 2 章 この章では，ロバスト推定法である RANSAC のパフォーマンスを向上させるために行われてきた研究について説明し，比較対象として選択した GR_RANSAC アルゴリズムの詳細に関して説明する。

第 3 章 この章では，提案した手法の詳細に関して説明する。

第 4 章 この章では，我々が行った実験と評価について，使用したデータセットも含めて説明する。

第 5 章 この章では，まとめと今後の課題について述べる。

第2章 関連研究

2.1 RANSACの改良

これまでに、ランダムサンプリングに基づき、閾値を用いてインライアとアウトライアを識別する RANSAC アルゴリズム性能を向上させるため、いくつかの拡張や派生アルゴリズムが提案されている。これは、研究の目的に応じて、パラメータ推定の精度、高速化、ロバストであること、に分類することが出来る [3]。計算時間の高速化を目的とした改良は、例えば、Progressive Sample Consensus (PROSAC) [4] があげられる。この手法は、特徴量記述子の類似度を用いてランク付けし、上位のランクスコアのものから順にサンプリングを行うことで、サンプリングの精度を高めることで高速化につなげている。同様にサンプリングを改善することによる計算時間の高速化は、N Adjacent Points Sample Consensus (NAPSAC [27]), Guided MLESAC [26] 等がある。データセットが変化した場合でもパラメータ推定の精度を保つには、閾値の自動選択が重要であり、これを目的とした改良は、例えば、いくつかの *automatic* RANSAC 方が提案されている [19, 21, 6]。これらの方法は、インライアとアウトライアを識別するための閾値を事前に固定する必要がなく、モデルパラメータとともに推定される。モデルパラメータ推定の精度向上を目的とした改良は、M-Estimator Sample Consensus (MSAC) [28], Maximum Likelihood Estimation Sample (MLESAC) [28] が提案されている。これらの手法は、全体的な性能は向上しているが、反復処理の性質上、外れ値の割合が大きい場合には、計算コストが高くなる。

2.2 フィルタリングを用いた RANSAC 改良

マッチングにより得られた対応関係集合のインライア比率は、RANSAC の計算時間に大きな影響を与える。そこで、マッチングのインライア比率を向上させるための研究がいくつか行われている。これは、RANSAC 実行前にマッチングによる外れ値を検出し、除去することを意味し、RANSAC 実行前のフィルタリングステップとして実装されている。例えば、画像間の対応点の空間的關係を利用した、Spatially consistent random sample consensus (SCRAMSAC [24]), spatial consistency on RANSAC (SC-RANSAC) [12] が提案されている。また、最近のいくつかの研究では、幾何学的關係を利用したフィルタリングを用いることで、事前

に外れ値を除去する方法もいくつか検討されている．[10] の論文では，幾何学的不変量を利用した，フィルタリングステップの有用性について示されている．しかし，画像から幾何学的不変量を安定的に抽出するには，計算コストがかかるといった課題が残されている．[8] の論文で提案されている GR_RANSAC は，ビデオ画像のような，連続して取得された隣接画像では，2次元特徴間の距離と角度は類似しているという仮定のもと，マッチングされた点からの外れ値除去に成功しており，反復回数の減少につなげている．しかし，著者らが指摘しているように，論文で使用されたデータセットと異なるものを使用する場合，角度と距離の閾値を再度決定する必要がある．GR_RANSAC アルゴリズムとグラフニューラルネットワークを組み合わせた (GNN-RANSAC) [9] も提案されている．これらのアルゴリズムは，RANSAC アルゴリズムを修正することなく，反復回数の減少に成功しているため，画像マッチングフレームワークにおける他のロバスト推定法にも有効である．他にも，ニューラルネットワークを用いてマッチングとフィルタリングを同時に実行する SuperGlue [23] も提案されているが，学習に時間がかかる．また，フィルタリングと RANSAC を最適化したものを組み合わせた，LP-RANSAC [29]，GH-COOSAC [5] 等も提案されている．

2.3 本研究の位置付け

我々の提案する新しいプレフィルタリングステップは，ロバスト推定法の実行前にフィルタリングを追加することで，ロバスト推定法の性能向上を目指している．そのため (図 2.1) の様に位置付けられる．また，GR_RANSAC アルゴリズムは，我々の提案手法と，画像マッチングパイプラインにおいて同様の位置付けにある (図 2.1)．そのため，本アルゴリズムとの比較に使用された．したがって，我々が提案するアルゴリズムの有用性を理解するには，GR_RANSAC がどのように働くのかを理解する必要がある．以下で，GR_RANSAC アルゴリズムの詳細に関して説明する．

尚，本研究で提案するアルゴリズムは，画像マッチングパイプラインにおける RANSAC 以外のロバスト推定法でも簡単に実装できるよう RANSAC への最適化は行わない．更に時間的制約や，多くのデータセットが必要になることからディープラーニングの手法も利用しない．そのため，GNN-RANSAC，SuperGlue や LP-RANSAC 等の学習や，RANSAC への最適化が行われているものは，本研究の対象外とする．

2.4 GR_RANSAC アルゴリズム

GR_RANSAC は，ビデオ画像のようなカメラの動きが小さく同一シーンを撮影している画像間では，検出された画像特徴間の距離と角度が類似していると仮定

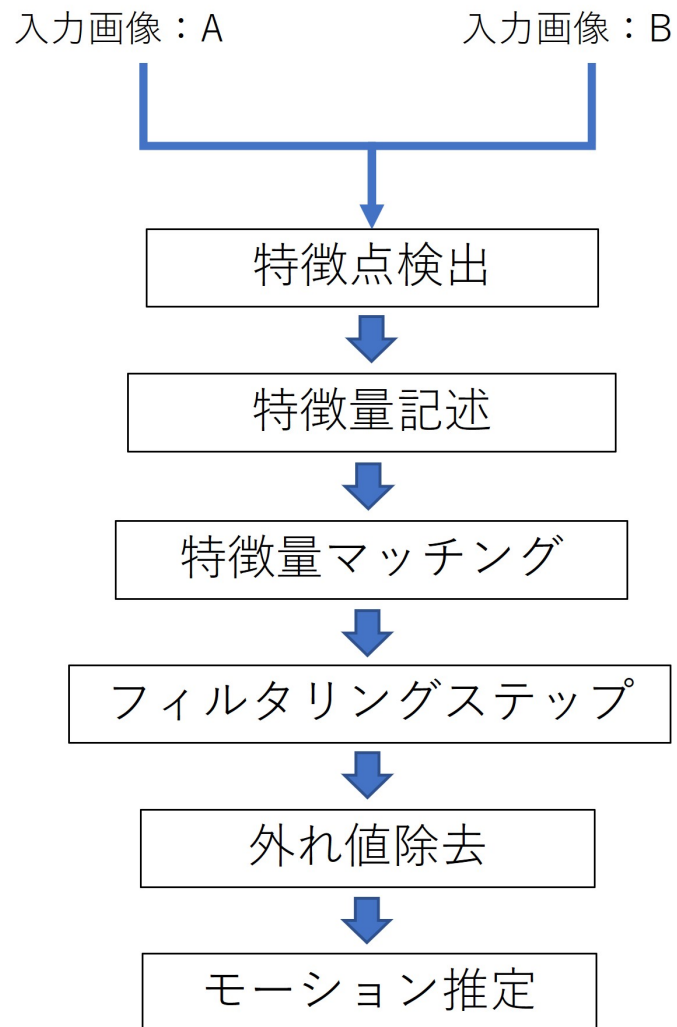


図 2.1: フィルタリングステップ (GR_RANSAC, またはプレフィルタリング) を追加した, 特徴ベース画像マッチングパイプライン

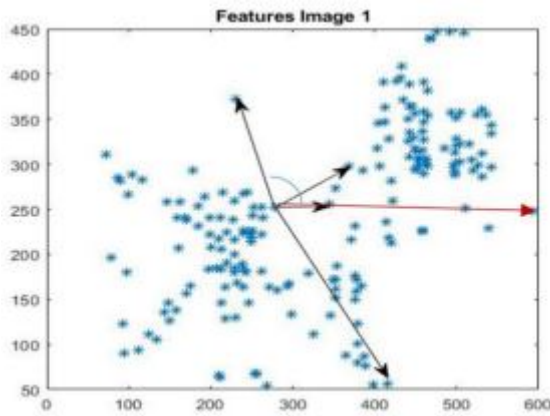
している．したがって，マッチングされた特徴点間の幾何学的関係の違い（閾値 t : 距離と角度）（図 2.2）を定義することで，外れ値を識別することが出来る．GR のフローチャートを（図 2.3）に示す．



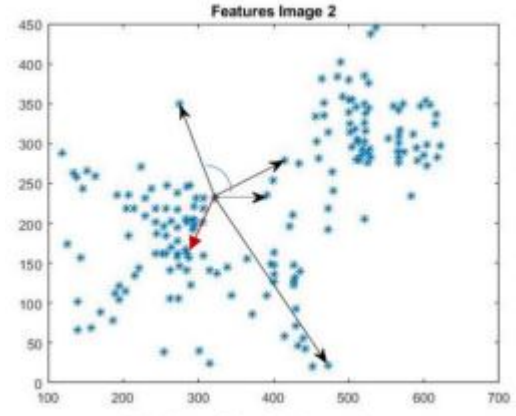
(a) 画像 1



(b) 画像 2



(c) 画像 1 における特徴点



(d) 画像 2 における特徴点

図 2.2: (a) と (b) は論文で使用された画像 [20]．(c) と (d) はそれぞれの画像における特徴点 [8]．幾何学的関係の違いは，各画像における基準点と残りの全ての対応点の距離と角度を計測することで判断され，事前に定義された閾値より大きい場合アウトライアとして識別される．

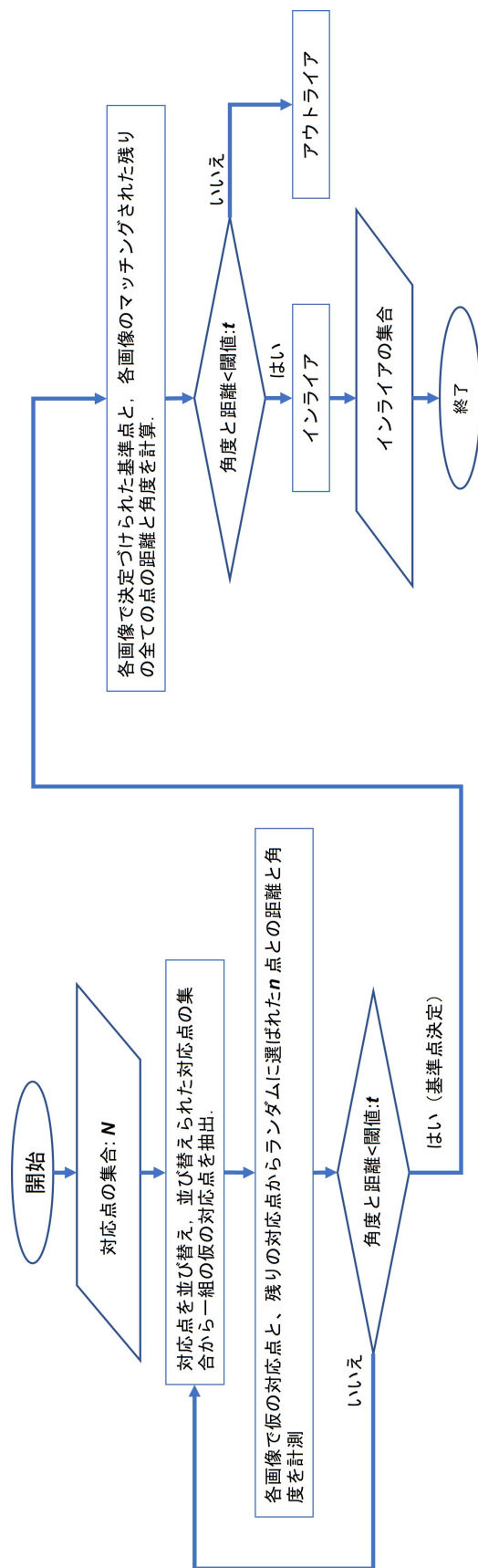
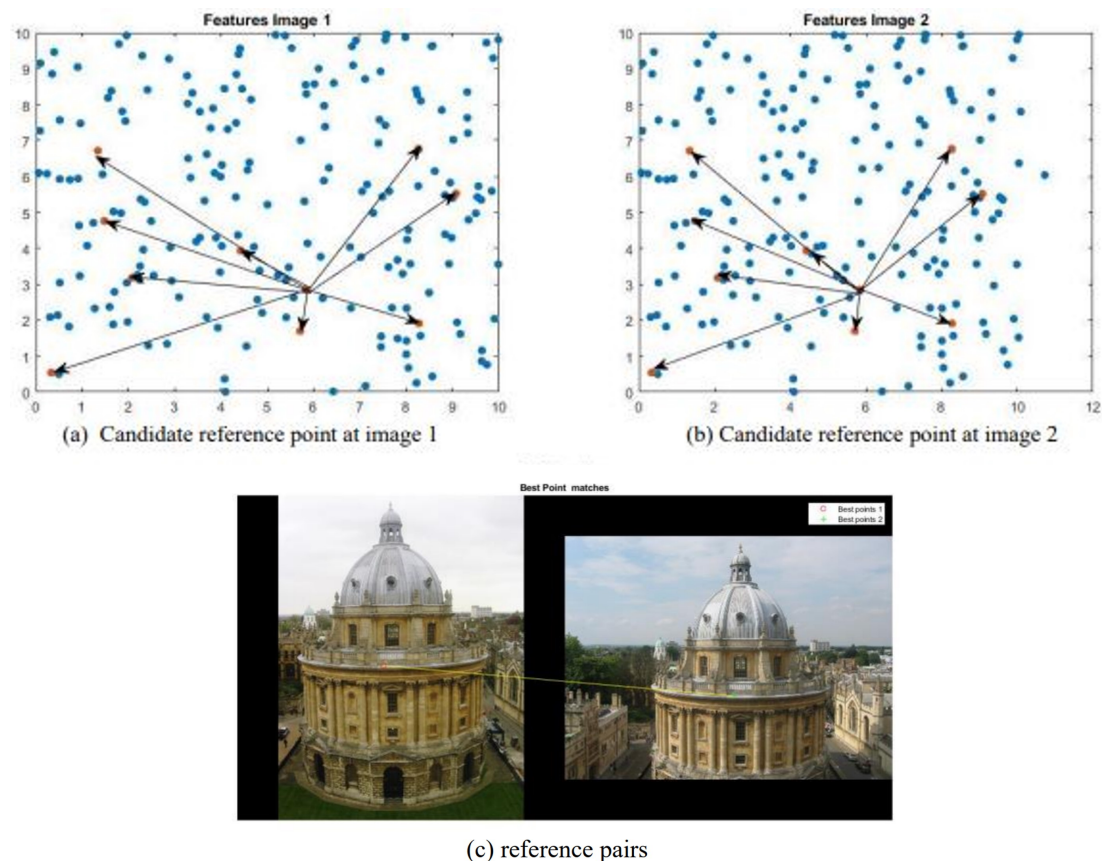


図 2.3: GR_RANSAC アルゴリズムフローチャート

まず、対応点は特徴量記述子のマッチングスコアによって並べ替えられる．並び替えられた対応点集合の中から、一組の仮の基準点を選ばれる．次に、各画像において、選ばれた仮の基準点と残りの対応点の中からランダムに選ばれた n 点の距離と角度を計算する．我々は、 $n = 5$ で選択した．計算結果を比較し、あらかじめ設定された閾値 (t : 距離と角度の値) より小さければ、基準点として決定する．そうでない場合は、並び替えられたリストの次の対応点を仮の基準点とし、基準点を決定できるまでこの処理を繰り返す (図 2.4)．



基準点の確立に関して[8]

図 2.4: (a) 画像 1 における仮の基準点 (b) 画像 2 における仮の基準点 (c) 基準点ペア．基準点は、各画像において仮の対応点と残りの対応点ペアからランダムに n 点抜き取り、その距離と角度を計測することで選ばれる．計測された、距離と角度が閾値より小さい場合、基準点として決定づけられる．我々は、 n を 5 点で選択した．

基準点が決定づけられた後、各画像の基準点と各画像のマッチングされた残りの全ての点の距離と角度が計算される． N 個の対応点の集合を仮定し、一つの基準点 $l_i = \{x_i, y_i\}$ と同じ画像中の残りの対応点 $l_j = \{x_j, y_j\}$ 間のユークリッド距離 L_p が計算される (ただし、 $i \neq j$)．更に、角度 $\theta_{i,j} = \arctan(x_i - x_j, y_i - y_j)$ も計算している．距離と角度が選択した閾値: t より小さい場合、インライアとし、そ

れ以外をアウトライアとして識別する.

インライア if $dL_p < td$ (距離閾値) and $|d\theta_{i,j}| < ta$ (角度閾値)

アウトライア otherwise

dL_p : 距離の差, $d\theta_{i,j}$: 角度の差

このアルゴリズムの閾値は, 距離と角度の値になっているため, 閾値の選択が困難である. そこで我々は, 閾値の値を直接用いるのではなく, 比率を用い, この比率値に対する閾値を用いる.

$$\begin{aligned}\text{距離比率} &: \frac{d1}{d2} \\ \text{角度比率} &: \frac{a1}{a2}\end{aligned}$$

$d1$ は決定された基準点から第 1 画像内の他の特徴点までの距離, 同様に $d2$ は第 2 画像内で算出された距離である. $a1$ は 1 枚目画像で計算された角度, $a2$ は 2 枚目画像で計算された角度である. 計算機シミュレーションにおいて, これらの比率の閾値を 1.2 とした.

第3章 提案手法

3.1 モチベーション

特徴ベース画像マッチングにおける，オーバーラップ画像間の対応関係の視覚化は，一般的にタイル状に並べられ，画像間の対応箇所を線で結ぶことで表される（図 3.1）．このような場合，ある種の変換に対する幾何学的不変量に基づいて，傾きの違いが保存されるため，インライアの多くは平行線のグループを形成する [14]．このようなモチベーションから，我々は，アウトライア除去とモーション推定のためのロバスト推定法を実行する前に，アウトライアを除去するためのシンプルかつ効率的な，新しいプレフィルタリングステップを提案する．



図 3.1: 画像をタイル状に並べた際の対応関係の可視化，赤線：インライア，緑線：アウトライア

3.2 アプローチ方法

本手法では，まず，2 枚目画像の特徴点位置に並進変換を施し（1 枚目の画像座標原点 (左上端) に対して 2 枚目画像の特徴点位置が相対的になるように），画像をタイル状に並べた際の平行線を得る．次に，対応関係が結ばれている直線の勾配

を計算する．勾配の計算後，直線の数が最大となる角度グループを見つけるため，角度ヒストグラムを生成する．これは，ヒストグラムの最頻値を見つけるのと同義である．この最大グループの直線を形成する特徴点は，インライアとみなされ，ロバスト推定法を用いてさらに処理するために保存される．我々の提案するアルゴリズムを Algorithm 1 に示す．

Algorithm 1: プレフィルタリングアルゴリズム

Input: 画像座標フレーム間において一致した特徴点位置

$\mathbf{P} = (x_i, y_i) \quad i = 1, 2, \dots, n$ and

$\mathbf{M} = (x_i, y_i) \quad i = 1, 2, \dots, n,$

画像サイズ： (u, v) ,

角度ヒストグラムを生成する際のビンの幅： w

Output: 保持された対応関係集合の配列インデックス, Ind

foreach 特徴量の対応関係 p and m **do**

A1 $\leftarrow$ 変換前の座標 $p = (x_p, y_p)$ と変換後の座標 $m = (x_m + u, y_m)$ を用いて，画像を水平方向にタイル状に並べた際の線の傾きを計算する

A2 $\leftarrow$ 変換前の座標 $p = (x_p, y_p)$ と変換後の座標 $m = (x_m, y_m + v)$ を用いて，画像を垂直方向にタイル状に並べた際の線の傾きを計算する

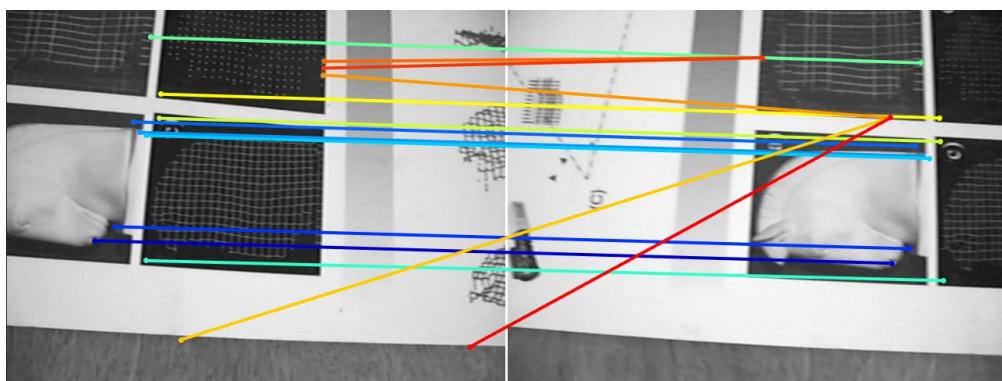
A3 $\leftarrow$ 変換前の座標 $p = (x_p, y_p)$ と変換後の座標 $m = (x_m + u, y_m + v)$ を用いて，画像を斜め方向にタイル状に並べた際の線の傾きを計算する

ビンの幅： w を用いて，**A1**，**A2** と **A3** の角度の値に対するヒストグラムを生成する．

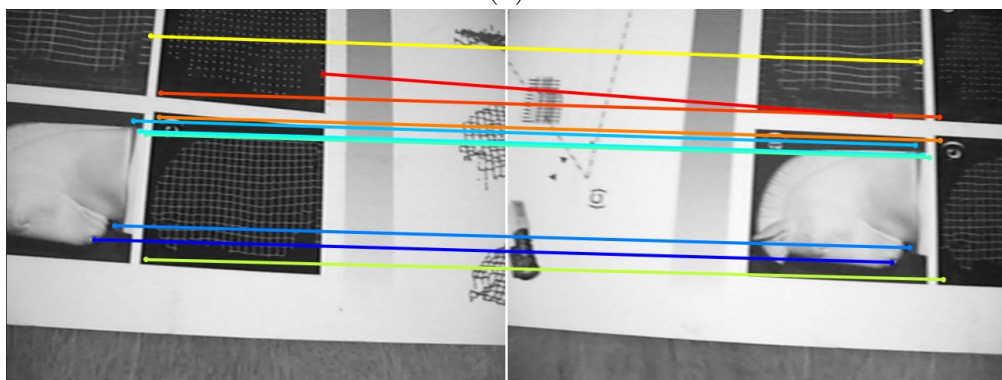
生成されたヒストグラムにおいて最大要素数となる，対応関係の配列インデックスを保持する．

$Ind \leftarrow$ **A1**，**A2**，**A3** の保持された配列インデックスを和集合で統一する．

提案したプレフィルタリングステップを用いて，水平方向，垂直方向，斜め方向にタイル状に並べた際の例を図 3.2, 3.3, 3.4 に示す．プレフィルタリングステップを追加する前に確立された対応関係の総数は，14 個で，その内 9 個がインライアであった．一方で、プレフィルタリングステップを追加した後の対応関係の総数は，10 個で、その内、9 個がインライア、1 個がアウトライアであった．



(a)



(b)

図 3.2: (a) 従来手法で確立された対応関係. 14 個の対応関係の内、9 個がインライア.
(b) フィルタリングを追加して確立された対応関係. 10 個の対応関係の内、9 個がインライアで、1 個がアウトライア.

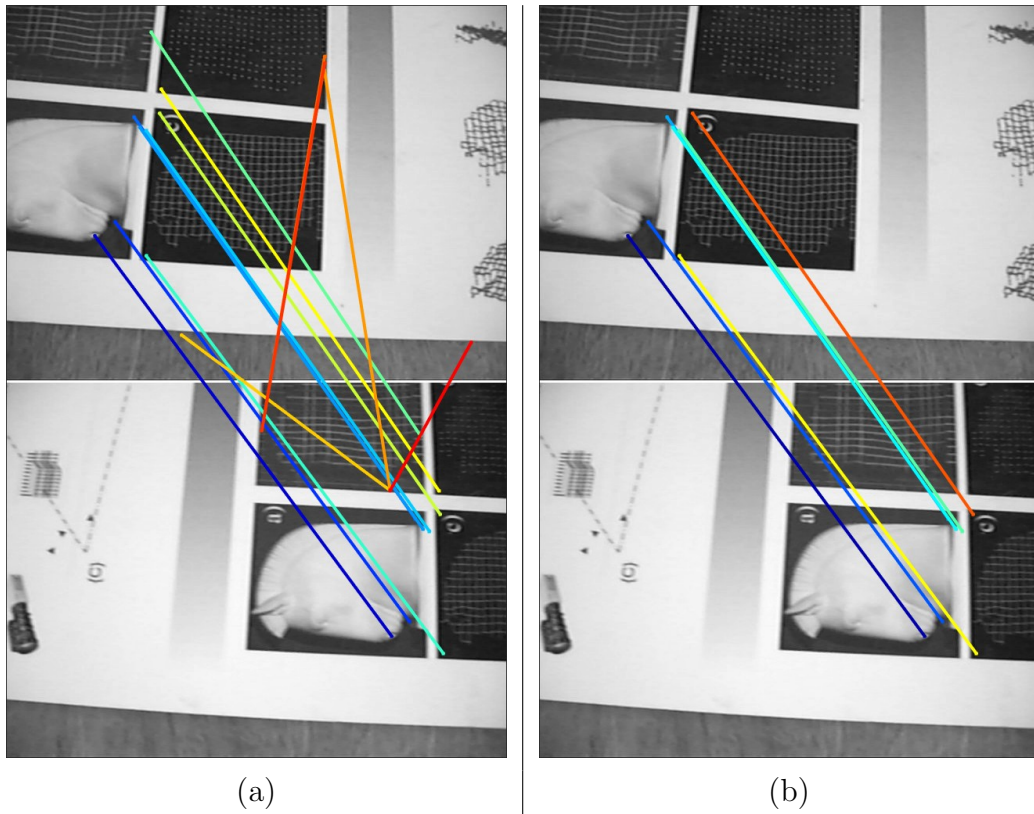
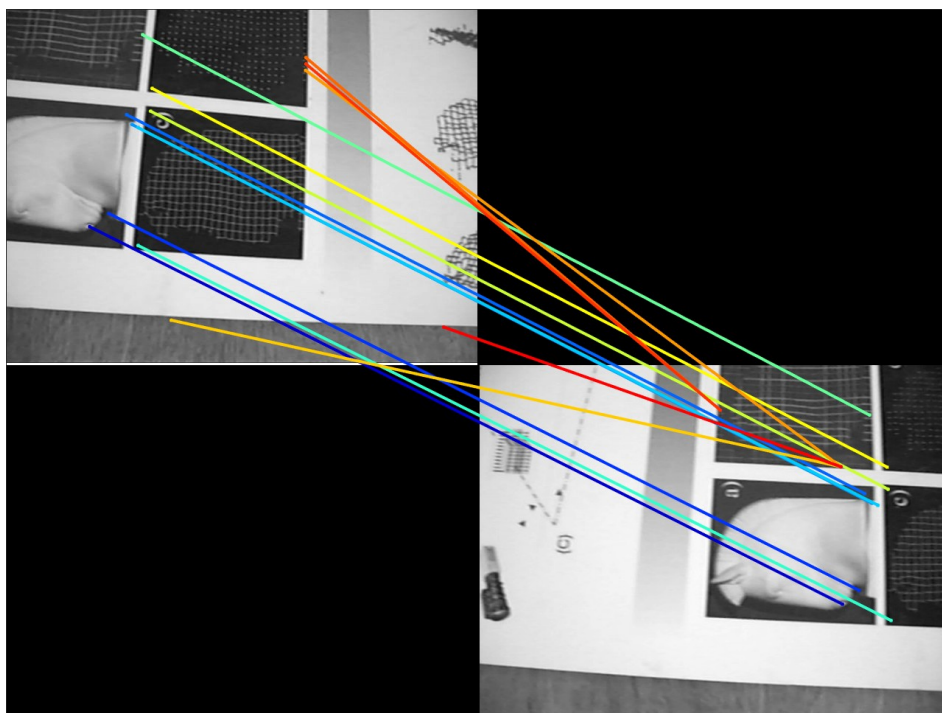
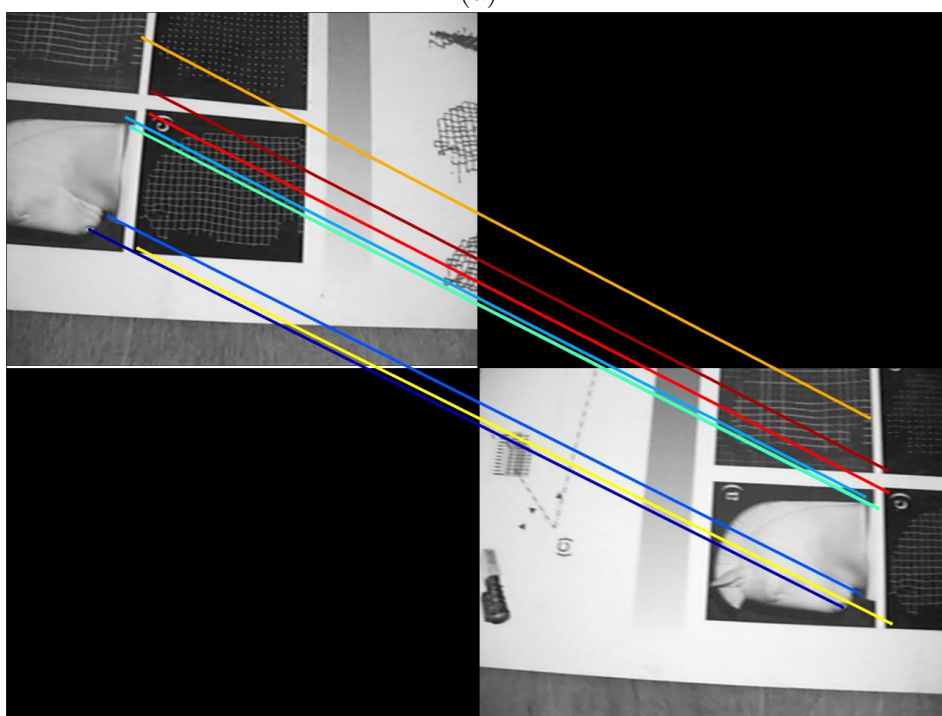


図 3.3: (a) 従来手法で確立された対応関係. 14 個の対応関係の内、9 個がインライア.
 (b) フィルタリングを追加して確立された対応関係. 7 個の対応関係の内、7 個がインライア.



(a)



(b)

図 3.4: (a) 従来手法で確立された対応関係. 14 個の対応関係の内、9 個がインライア.
(b) フィルタリングを追加して確立された対応関係. 9 個の対応関係の内、9 個がインライア.

第4章 実験・評価

我々は、提案したプレフィルタリングを、画像間の様々な視点変化があり、その際の動きを表したホモグラフィ行列を提供しているデータセット内の実データ [17] と、[17] で使われている異なる種類のホモグラフィを用いた計算機シミュレーションにより評価した。加えて、我々は、提案手法と画像マッチングパイプラインにおいて同様の位置付けにある GR_RANSAC を用いた比較、評価も行う。

4.1 計算機シミュレーションによる実験

本実験では、ノイズがあり、外れ値比率が高い場合において、我々の提案手法を従来の特徴点マッチングパイプラインに追加することで、ロバスト推定法の性能向上に対する有効性を確認する。

4.1.1 実験環境

シミュレーション実験のため、我々は、[17] で使用されている 22 種類のホモグラフィを利用し、ホモグラフィ集合を生成した。生成されたホモグラフィ集合を第 1 画像 (図 4.1(a)) に適用し、選択されたホモグラフィにより関連付けられた第 2 画像を生成した。生成された画像の例を (図 4.1(b) ~ (g)) に示す。その後、我々は SURF を用いて、第 1 画像と第 2 画像から特徴点の検出とマッチングを行った。マッチングにより対応関係が確立された後、RANSAC を実行することで、合成された画像であるにもかかわらず発生した外れ値を除去する。RANSAC により求められたホモグラフィと正解データホモグラフィが近い場合、インライアとし、この値を用いて表 4.1 にまとめた各パラメータに沿ってシミュレーションを実行した。我々は、外れ値比率に沿ったアウトライアとインライアの両方を含んだ対応関係を、パラメータにまとめてある対応関係の総数だけ、ランダムに生成した。その後、我々は、提案したプレフィルタリングステップと改良した GR_RANSAC [8] を用いた場合と、用いなかった場合で従来の RANSAC [11] を実行することにより、ホモグラフィを推定した。更に、我々是对応点の位置を破損させるため、ゼロ平均ノイズに標準偏差の異なるノイズを追加した。



(a) 第 1 画像



(b)



(c)



(d)



(e)



(f)



(g)

図 4.1: (a): 実験で使った第 1 画像. (b) ~ (g): (a) に対して実験で使った変換を施したワープ画像例

ホモグラフィ行列の数	22
対応関係の総数 (インライア + アウトライア)	[100, 150, 200]
外れ値比率	[0.5, 0.6, 0.7, 0.8, 0.9]
ノイズの標準偏差	[0.0, 0.5, 1.0, 1.5, 2.0]
RANSAC の最大反復回数	2, 500
距離閾値 (RANSAC 閾値, 誤差閾値)	5 ピクセル

表 4.1: 実験で使ったパラメータ

4.1.2 提案手法の評価方法

我々は、提案したプレフィルタリングステップを用いた場合と、用いなかった場合で従来の RANSAC [11] を実行することにより、ホモグラフィを推定し、正解データホモグラフィと比較することにより評価する (図 4.2). もし、式 (4.1) により計算された誤差が距離閾値 5 ピクセルより小さい場合、推定されたホモグラフィは正しく、それをホモグラフィ推定における成功としてカウントする. 誤差は、距離誤差の平均を求めることで得られる. これは、正解データのホモグラフィ (Ground-truth homography) と RANSAC により推定されたホモグラフィ (Estimated homography) を Image1 上の特徴点 X_i に適用し得られた、Image2 上の点 GX'_i と EX'_i 間の距離 d_i を求め、シミュレーションパラメータに沿って生成されたすべての対応点 (N) に対し計算し、その平均を求めることで得られる. その後、平均を求めることで得られる. 更に我々は、同様の手順で、改良した GR_RANSAC [8] を用いてホモグラフィを推定し、その結果を我々の提案手法と比較評価する.

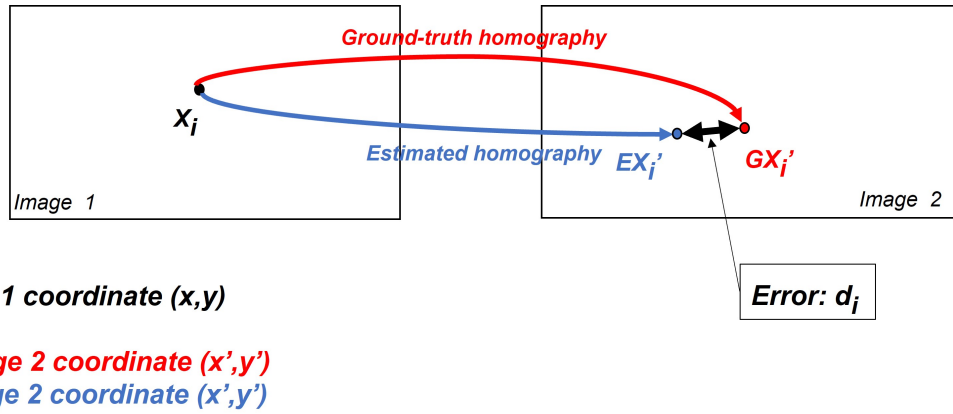


図 4.2: 誤差計算

$$\mu = \frac{\sum_{i=1}^N d_i}{N} \quad (4.1)$$

$$d_i = \|GX'_i - EX'_i\|$$

閾値の5ピクセルは、正解データの対応関係を崩すために使用した、ノイズを考慮した最大誤差として決定した。各ホモグラフィーにおいて、 $3 \times 5 \times 5$ のシミュレーションパラメータで試行し、これを1,000回繰り返し、合計1,650,000($22 \times 75 \times 1,000$)回の試行を行った。

4.1.3 ホモグラフィ推定の成功率に関する結果

ここでは、RANSACにプレフィルタリングステップまたは、GR_RANSACを追加し、ホモグラフィを推定した際の精度に関する結果を表4.2a, 4.2bに示す。表4.2aの合計より、RANSACにプレフィルタリングステップを追加することで、約14%結果の向上が確認できる。一方で、RANSACにGR_RANSACを追加した際の結果は、テストしたほとんどのケースで正確なパフォーマンスを発揮することが出来なかった。これは、GR_RANSACが、基準点の確立に失敗したとが原因であり、これはほとんどが画像間の動きによるものである。これらの結果から、GR_RANSACはほとんどの場合、3自由度や4自由度のようなより低い、自由度を持つ動作にほとんど適していると言える。以上のことから、我々の提案したプレフィルタリングステップはホモグラフィ推定の精度に関して、GR_RANSACと比較した際、より柔軟であるといえる。また、従来のRANSACに追加した場合にも、有用であるといえる。

4.1.4 RANSAC反復回数に関する結果

ここでは、プレフィルタリングステップや、改良したGR_RANSACを用いた場合と、従来RANSACのみでホモグラフィを推定した場合の反復回数に関する結果を統計的にまとめたものを表4.3に示す。また、フィルタリングを追加した際の平均計算時間の結果を表4.4に示す。表4.3より、我々が提案したプレフィルタリングステップや、GR_RANSACを追加することで、RANSACにおける必要な反復回数の削減に成功したことが確認できる。これは、主にフィルタリングステップ（プレフィルタリングステップ、GR_RANSAC）を追加することで、インライア比率を上げつつ、対応点数を減らすことができるためである。これはどちらも、反復回数を減らすのに有利である。また、プレフィルタリングステップと、GR_RANSACの反復回数を比較した場合、GR_RANSACが優れていることが確認できる。これは、GR_RANSACの方が平均インライア比率を向上させることが

表 4.2: ホモグラフィ推定の成功率に関する結果のまとめ

(a) 従来 RANSAC にプレフィルタリングステップを追加した際のホモグラフィ推定精度に関する結果

従来 RANSAC		失敗	成功	合計
提案手法	失敗	0.048 (78,887)	0.024 (38,781)	0.071 (117,668)
	成功	0.165 (272,169)	0.764 (1,260,163)	0.929 (1,532,332)
	合計	0.213 (351,056)	0.787 (1,298,944)	1.000 (1,650,000)

(b) 従来 RANSAC に GR_RANSAC を追加した際のホモグラフィ推定精度に関する結果

従来 RANSAC		失敗	成功	合計
GR_RANSAC	失敗	0.211 (348,127)	0.511 (842,765)	0.722 (1,190,892)
	成功	0.002 (2,929)	0.276 (456,179)	0.278 (459,108)
	合計	0.213 (351,056)	0.787 (1,298,944)	1.000 (1,650,000)

出来ているからである。しかし、従来 RANSAC と比較した場合、反復回数の削減に成功しているため有用であるといえる。また、表 4.4 より、従来の RANSAC にフィルタリングを追加した場合、平均計算時間の短縮に成功していることが確認できる。平均時間の計測は、RAM: 48GB, CPU: AMS Ryzen7 2700 Eight-core Processor を搭載したコンピュータで行った。

4.1.5 プレフィルタリングが失敗したケースの結果

我々のプレフィルタリングステップは、画像間の対応点をタイル状に並べた際の平行線の数に依存している。そのため、インライアの総数がインライア比率よりも重要である。つまり、インライア比率が 0.1 で、対応点の総数が 200 の場合、外れ値比率が同一で対応点の総数が 100 の場合よりも我々のプレフィルタリングステップは良好な性能を発揮する。我々のフィルタリングが失敗した全てのケース 117668 において使用したインライア比率に対する各対応点総数のヒストグラムを表したものが図 4.3 である。

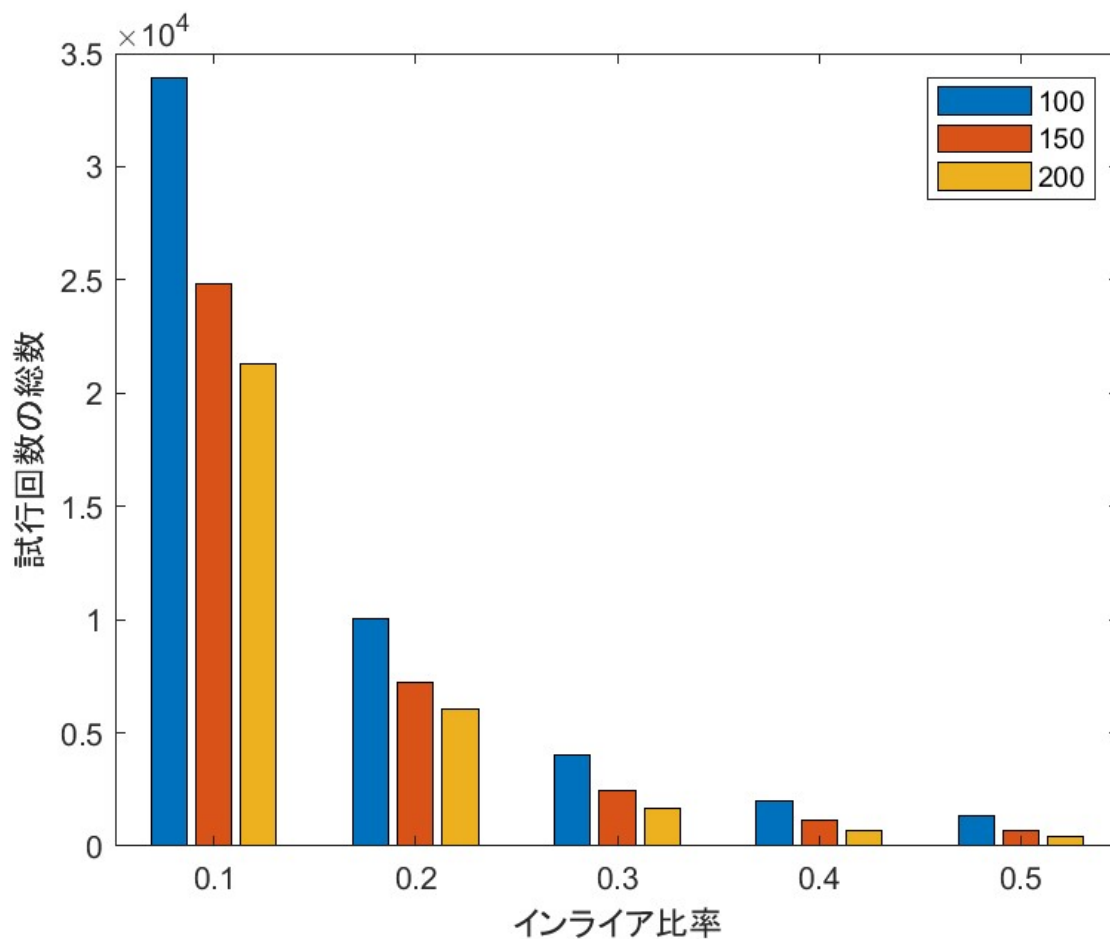


図 4.3: プレフィルタリングステップを用いて、ホモグラフィ推定に失敗した際のインライア比率に関するヒストグラム. 従来の RANSAC のみでホモグラフィを推定した際でも, この図で報告されている全事例の内 70.17% で失敗している. 対応点の総数が 100 個の場合, インライア比率の違いによる失敗したケースが最も多くなっている. このことから, 我々の提案手法は, インライア比率よりもインライアの総数が重要であると結論付けられる.

表 4.3: RANSAC 反復回数に関する結果

	最小	最大	平均	標準偏差	平均インライア比率
従来 RANSAC (全体)	58	2501	1300.4	1060.47	0.30
プレフィルタリング (全体)	4	2501	523.56	812.44	0.51
GR.RANSAC (全体)	2	2501	20.45	68.81	0.86
従来 RANSAC (成功)	58	2501	976.09	966.39	0.35
プレフィルタリング (成功)	4	2501	385.38	649.30	0.54
GR.RANSAC (成功)	2	2501	19.44	52.77	0.86
従来 RANSAC (失敗)	273	2501	2500.48	27.96	0.12
プレフィルタリング (失敗)	36	2501	2322.96	528.35	0.19
GR.RANSAC (失敗)	8	2501	758.64	944.16	0.52

表 4.4: 平均計算時間の結果 (秒)

	RANSAC	提案手法	提案手法+RANSAC	GR	GR+RASAC
全体	0.1762	2.5728E-04	0.0711	0.0026	0.0035
成功時	0.1332	2.5769E-04	0.0530	2.6433E-03	0.0034
失敗時	0.3352	2.5204E-04	0.3072	0.0059	0.0987

4.1.6 プレフィルタリングがインライア比率を改善したケースの結果

また、提案したプレフィルタリングにより、あらかじめ設定したインライア比率 (1- 外れ値比率) の向上に成功した際の総数と、インライア比率の平均を表 4.5 に示す. 表 4.5 の**大きい (>)** 列は、プレフィルタリングを追加することで、インライア比率が向上した総数を示した. 一方で、**小さい (<)**, **等しい (=)** の列は、インライア比率が減少、変化しなかった総数を示した. また、残りの 3 列は、インライア比率が向上したケース、向上しなかったケースとその両ケースにおけるインライア比率の平均を示した. 表 4.5 と表 4.2a で示されたプレフィルタリングが成功、失敗した際の合計には食い違いがある. これは、プレフィルタリングによりインライア比率が増加しているにも関わらず、ロバスト推定法である RANSAC で正しいモーションを得ることが出来なかったケースとその逆のケースがあるためである.

表 4.5: 提案手法の成功と失敗したケースの結果

プレフィルタリング						
inlier Ratio	試行回数の総数			平均比率		
	大きい (>)	小さい (<)	等しい (=)	大きい (>)	小さい (<)	両ケース
0.1	309916	20084	0	0.25	0.07	0.24
0.2	311416	18584	0	0.42	0.16	0.41
0.3	311591	18409	0	0.56	0.25	0.54
0.4	311910	17335	755	0.66	0.34	0.64
0.5	312121	16040	1839	0.74	0.44	0.73
Total	1556954	90452	2594			

4.2 実画像を用いた実験

本実験では、重複領域のある画像間に様々な視点変化があり、その際の動きを表したホモグラフィ行列を提供する *graffiti* データセット [17] 内の実画像を用いて、提案手法を評価した。また、GR_RANSACとも比較し、評価した。

我々は、標準的な SIFT ベースのパイプラインに従って、画像間の対応関係を求め、提案するプレフィルタリングを追加した場合と追加しない場合の両方でホモグラフィの推定を行った。同様の手順で GR_RANSAC によるホモグラフィ推定も行った。実験に使用した実画像を (図 4.4) に示す。RANSAC の誤差閾値は 7.5 ピクセル、最大サンプリング反復回数は 5000 回を使用した。得られた結果を表 (4.6) にまとめる。

データセットに含まれているホモグラフィを用いることで、画像間の正確な変換が得られる。そのため、我々は、それらのデータを用いることで、表の最後の 2 列に示してあるインライア数とインライア比率を求める。このインライア数とインライア比率は、RANSAC と同じ誤差閾値を用いて推定されたホモグラフィの正確さや、対応点のインライア比率、反復回数についての比較と考察を行うため使用した。

表より、我々の提案したプレフィルタリングステップは画像マッチングパイプラインを全体的に改善することができたことが確認できる。また、GR_RANSAC と比較した場合にも我々の提案した手法は優位性があることが示せた。画像ペア 1-4 のケースまでは、ホモグラフィの推定に成功している。特に、画像ペア 1-4 のケースでは、インライア比率を約 19% 向上させることに成功しており、従来 RANSAC と比較して約 92% 反復回数の削減に成功していることが確認できる。

画像ペア 1-5, 1-6 のケースでは、両ケースともホモグラフィを正しく推定できなかった。これは、記述子マッチングにより確立された対応関係に、ホモグラフィ推定に必要な、最低限の正しい対応点 (4) がないことが原因で、使用した特徴点抽出・マッチング手法の直接的な影響によるものと考えることが出来る。また、提案手法を追加した際、最低限の正しい対応点が無いのにも関わらず、反復回数が

少ないのは、我々の提案手法は、画像をタイル状に並べた際の平行線の数をカウントして外れ値を除去しているため、誤った対応点の平行線をインライアの集合としてしまい、ホモグラフィ推定が行われたのが原因であると考えられる。他にも、プレフィルタリングにより減らされた対応点の中から、RANSACがアウトライアをインライアと誤判断してしまったことが考えられ、これはRANSAC 閾値が大きいことを意味する。



図 4.4: 実験に使用した実画像

表 4.6: 実画像を用いた実験による結果

画像ペア	方法	対応点総数	RANSAC 反復回数	インライア数	インライア比率	正解データ	
						インライア数	インライア比率
1 2	RANSAC	484	5	415	0.857	415	0.857
	提案手法	411	6	379	0.922		
	GR	375	5	339	0.904		
1 3	RANSAC	106	51	45	0.425	45	0.425
	提案手法	65	50	32	0.492		
	GR	Failed	Failed	Failed	Failed		
1 4	RANSAC	72	949	13	0.181	13	0.181
	提案手法	32	73	12	0.375		
	GR	Failed	Failed	Failed	Failed		
1 5	RANSAC	51	5001	0	0.000	1	0.020
	提案手法	16	482	1	0.063		
	GR	Failed	Failed	Failed	Failed		
1 6	RANSAC	55	5001	0	0.000	0	0.000
	提案手法	22	1725	0	0.000		
	GR	Failed	Failed	Failed	Failed		

第5章 おわりに

画像マッチングは、多くの高度なコンピュータビジョンやロボティクスタスクにおいて、最も重要かつ基本的なステップの1つである。特徴点の検出、記述、マッチングステップにおいて、特にディープラーニングの手法を用いた研究により、目覚ましい進歩があった。しかしながら、マッチングステップでは、依然としていくつかミスマッチが発生してしまうため、ミスマッチを除去し、正確なモーション推定を行うためのロバスト推定法が必要である。しかし、ロバスト推定法は反復手法であるため、マッチングパイプライン全体における計算時間の大部分を占めるステップの1つとして知られている。そこで我々は、ロバスト推定法における反復回数を削減し計算時間を短縮するためのシンプルかつ効率的な新しいプレフィルタリングステップを提案する。我々のプレフィルタリングステップは、重複領域のある2枚の画像をタイル状に並べた際に得られる画像間の対応関係を結ぶ直線をグループ化することにより、マッチングにより得られた対応関係の集合のインライア比率を向上させること目的としている。我々は、従来RANSAC、提案手法+従来RANSAC、GR_RANSACで比較し、計算機シミュレーションによる実験と実画像を用いた実験により提案手法の有効性を示す。

計算機シミュレーションの結果より、提案手法+従来RANSACは、従来RANSACと比較して、約14% ホモグラフィ推定の成功率を向上させることが出来た。また、GR_RANSACは、テストしたほとんどのケースで正確なパフォーマンスを発揮することが出来なかった。更に、反復回数に関する結果より、プレフィルタリングとGR_RANSACは、どちらも反復回数の削減に成功した。プレフィルタリングとGR_RANSACを比較すると、GR_RANSACはプレフィルタリングよりも更に反復回数が少ないことが確認できた。しかしながら、GR_RANSACはテストしたほとんどのケースでホモグラフィ推定に失敗したため、我々の提案手法は、GR_RANSACや従来RANSACと比較して、より柔軟性があり効率的であることが確認できた。

実画像を用いた実験より、ホモグラフィ推定に必要な最低限の対応点数(4)がある場合でインライア比率が低い(外れ値比率が高い)ような状況においても、外れ値の総数を減らすことで、従来RANSACと比較した際、反復回数の削減に成功していることが確認できた。

以上のことから我々の提案した新しいプレフィルタリングステップは、画像マッチングの成功率を向上させるだけでなく、対応関係の総数を減らし、全体のインライア率を高めることができるため、ロバスト推定の際に行われる反復回数の削減に有効であると決定付けられる。

我々の実験では、画像間に複数の動きがあることは考慮していない．これは例えば、シーン内の物体が移動することによって発生する可能性がある．このような画像間に複数の動き（ホモグラフィ）が存在する場合は、プレフィルタリングは誤った平行線を検出してしまうことが考えられる．そのため、将来的には、そのようなケースに対応できるようプレフィルタリングステップを拡張することに注力する．

我々は現在、日本文理大学伊藤研究室と共同で、起伏のある地形（例えば、図 5.1）で走行可能な低コストな Hexapod robot（図 5.2）の開発を行っている．これは、6 脚ロボットは常に 3 点保持が可能であるということから、図 5.1 のような地形での探索（例えば、位置推定、マッピング等）に有効であると考えている．本学では、主にロボットの歩行制御部分と、カメラを用いたビジュアルアプリケーションのパートを担当する．本研究で提案した手法は、従来の画像マッチングパイプラインに対する追加・取り外しが容易にできるため、将来的には、そのロボットを用いた、画像マッチング技術がベースとなっているマッピングや位置推定等の高度なタスクを実行する際に組み込むことを考えている．



図 5.1: 災害現場 [25]

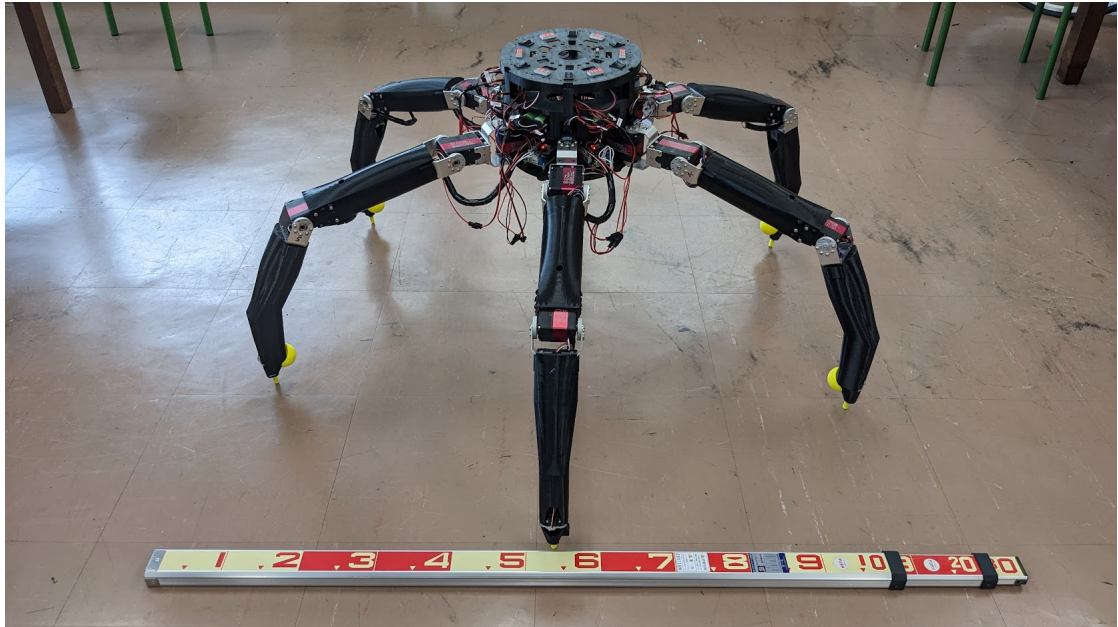


図 5.2: 開発中の Hexapod robot(日本文理大学伊藤研究室提供)

謝辞

本研究にあたり，論文のテーマ決定や，研究の進め方等，熱心に指導して頂いた丁洛榮教授，並びにエリボル・アルマガン准教授に心より感謝いたします。特に，エリボル・アルマガン准教授には，未知の分野であった画像処理に関して，丁寧に教えて頂き，研究に行き詰まる度にアドバイスを頂きました。心より感謝申し上げます。また，ロボットの開発に協力していただいた日本文理大学の伊藤順治教授，並びに研究室の学生の方にも深く感謝申し上げます。加えて，日頃からお世話になりました研究室メンバーの方々にも深く感謝申し上げます。

参考文献

- [1] Vassileios Balntas, Karel Lenc, Andrea Vedaldi, and Krystian Mikolajczyk. Hpatches: A benchmark and evaluation of handcrafted and learned local descriptors. In *CVPR*, 2017.
- [2] H. Bay, T. Tuytelaars, and L. J. Van Gool. SURF: Speeded up robust features. pp. 404–417, Graz, Austria, May 2006.
- [3] Sunglok Choi, Taemin Kim, and Wonpil Yu. Performance evaluation of ransac family. In *Proceedings of the British Machine Vision Conference*, pp. 81.1–81.12. BMVA Press, 2009. doi:10.5244/C.23.81.
- [4] Ondrej Chum and Jiri Matas. Matching with prosac-progressive sample consensus. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Vol. 1, pp. 220–226. IEEE, 2005.
- [5] Kuo-Liang Chung, Ya-Chi Tseng, and Hsuan-Ying Chen. A novel and effective cooperative ransac image matching method using geometry histogram-based constructed reduced correspondence set. *Remote Sensing*, Vol. 14, No. 14, p. 3256, 2022.
- [6] Andrea Cohen and Christopher Zach. The likelihood-ratio test and efficient robust estimation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [7] Mihai Dusmanu, Ignacio Rocco, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii, and Torsten Sattler. D2-Net: A Trainable CNN for Joint Detection and Description of Local Features. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
- [8] A. Elashry, B. Sluis, and C. Toth. Improving ransac feature matching based on geometric relation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLIII-B2-2021, pp. 321–327, 2021.

- [9] A. Elashry and C. Toth. Feature matching enhancement using the graph neural network (gnn-ransac). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLVI-M-2-2022, pp. 83–89, 2022.
- [10] Armagan Elibol and Nak Young Chong. Efficient image registration for underwater optical mapping using geometric invariants. *Journal of Marine Science and Engineering*, Vol. 7, No. 6, 2019.
- [11] M. A. Fischler and R. C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395, 1981.
- [12] Mehran Fotouhi, Hamid Hekmatian, Mohammad Amin Kashani-Nezhad, and Shohreh Kasaei. Sc-ransac: spatial consistency on ransac. *Multimedia Tools and Applications*, Vol. 78, No. 7, pp. 9429–9461, 2019.
- [13] Rui FC Guerreiro and Pedro MQ Aguiar. Global motion estimation: feature-based, featureless, or both?! In *International Conference Image Analysis and Recognition*, pp. 721–730. Springer, 2006.
- [14] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Harlow, UK, second edition, 2004.
- [15] D.G. Lowe. Distinctive image features from scale-invariant keypoints. Vol. 60, No. 2, pp. 91–110, 2004.
- [16] Jiayi Ma, Xingyu Jiang, Aoxiang Fan, Junjun Jiang, and Junchi Yan. Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision*, Vol. 129, No. 1, pp. 23–79, 2021.
- [17] Krystian Mikolajczyk, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and Luc Van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, Vol. 65, No. 1/2, pp. 43–72, 2005.
- [18] Keita Miyaura, Armagan Elibol, and Nak Young Chong. Performance enhancement step for motion estimation via feature-based image matching. In *2022 22nd International Conference on Control, Automation and Systems (ICCAS)*, pp. 1161–1166, 2022.
- [19] Lionel Moisan, Pierre Moulon, and Pascal Monasse. Automatic Homographic Registration of a Pair of Images, with A Contrario Elimination of Outliers. *Image Processing On Line*, Vol. 2, pp. 56–73, 2012.

- [20] James Philbin, Ondrej Chum, Michael Isard, Josef Sivic, and Andrew Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [21] Clément Riu, Vincent Nozick, and Pascal Monasse. Automatic RANSAC by Likelihood Maximization. *Image Processing On Line*, Vol. 12, pp. 27–49, 2022.
- [22] Peter J Rousseeuw. Least median of squares regression. *Journal of the American statistical association*, Vol. 79, No. 388, pp. 871–880, 1984.
- [23] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superglue: Learning feature matching with graph neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [24] Torsten Sattler, Bastian Leibe, and Leif Kobbelt. Scramsac: Improving ransac’s efficiency with a spatial consistency filter. In *2009 IEEE 12th International Conference on Computer Vision*, pp. 2090–2097, 2009.
- [25] SORABITO. 災害時のドローンの活用方法. <https://sora-bito.com/blog/column/123575>(accessed on 23 January 2023).
- [26] B.J. Tordoff and D.W. Murray. Guided-mlesac: faster image transform estimation by using matching priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 10, pp. 1523–1535, 2005.
- [27] Philip Hilaire Torr, Slawomir J Nasuto, and John Mark Bishop. Napsac: High noise, high dimensional robust estimation-it’s in the bag. In *British Machine Vision Conference (BMVC)*, Vol. 2, p. 3, 2002.
- [28] P.H.S. Torr and A. Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, Vol. 78, No. 1, pp. 138–156, 2000.
- [29] Gang Wang, Xiaoliang Sun, Yang Shang, Zi Wang, Zhongchen Shi, and Qifeng Yu. Two-view geometry estimation using ransac with locality preserving constraint. *IEEE Access*, Vol. 8, pp. 7267–7279, 2020.
- [30] X.Han, T. Leung, Y. Jia, R. Sukthankar, and A. C. Berg. Matchnet: Unifying feature and metric learning for patch-based matching. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3279–3286, 2015.