

画像マッチングを用いた全天球カメラの位置推定手法の一検討

宮崎 永^{†1} 阿部 亨^{†1,†2} 菅沼 拓夫^{†1,†2}

^{†1} 東北大学工学部電気情報理工学科 ^{†2} 東北大学サイバーサイエンスセンター

1 はじめに

近年、マップ機能やロボットによる自律ナビゲーションなど位置情報を利用したサービスの需要が高まっている [1]。この機能を実現するためにはサービスを実行する機器の自己位置推定が必要となる。屋外での自己位置推定には GPS を用いることが多いが、屋内では、GPS を用いた正確な自己位置推定は困難である。そこで、移動する機器に搭載したカメラで撮影した映像に基づき、屋内での自己位置を推定する手法が提案されている。この手法は、環境内の各地点で事前に撮影した画像をデータベースに保存し、新たに撮影した入力画像をそれらと比較することで現在の自己位置を推定するものである。しかし、この手法では、異なる地点で撮影したにも関わらず画像が類似する場合に、誤った地点を自己位置として推定してしまうエリアシング問題があった。

そこで本稿では、地理的に近い地点で撮影された複数の画像を比較に用いて、エリアシング問題に対処する自己位置推定手法について検討する。

2 関連研究

カメラの映像を用いた自己位置推定手法では、広い範囲を一度に撮影できる全天球カメラを用いた手法が多く提案されている。Payá ら [2] は、図 1 のようなパノラマ画像を環境内の各所で事前に撮影し、それらの類似度からクラスタリングを行い環境モデルを作成する手法を提案している。作成された環境モデルのイメージを図 2 に示す。この手法では、三段階のレベルで画像をグループ化しており、第一段階では、データベースの画像を部屋ごとにクラスタリングし、第二段階では、各部屋の画像をさらにクラスタリングしている。第三段階では、撮影地点の隣接関係とともに各画像を保存している。各クラスは一つの画像で代表されており、入力画像に対する自己位置推定は以下の手順で行われる。

1. 各クラスタの代表画像と入力画像を比較し、最も類似しているクラスタを選択
2. 選択されたクラスタ内のすべての画像と入力



図 1: 全天球カメラで撮影したパノラマ画像

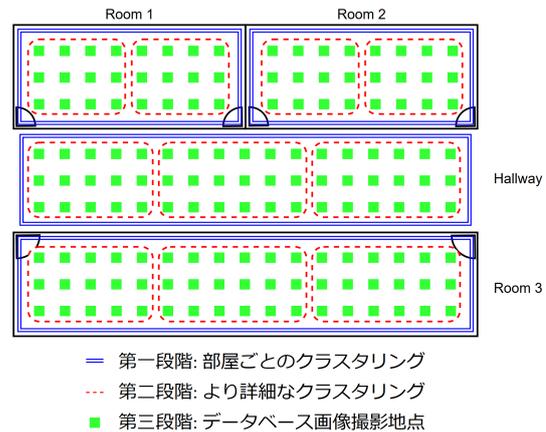


図 2: クラスタリングによる環境モデルの例

画像を比較し、最も類似している画像を選択

3. 選択された画像の撮影地点を自己位置として出力

しかし、この手法では異なる地点で撮影した画像が類似している場合、入力画像に対応したデータベース画像を適切に選択できず（エリアシング問題）、誤った地点を自己位置として推定する可能性がある。

3 提案

本稿では、カメラ映像を用いた自己位置推定におけるエリアシング問題へ対処するため、地理的に近い複数の画像との類似度を考慮し自己位置を推定する手法に関して検討を行う。

自己位置推定を行う際にエリアシング問題が発生する例を図 3 に示す。この例では直線上を移動しながら、各アルファベットの地点で環境を撮影した画像をデータベース画像とする。自己位置推定を行う場所として地点 C で撮影した画像を入力した

A Study on Localization of an Omnidirectional Camera by Image Matching

Haruka MIYAZAKI^{†1}, Toru ABE^{†1,†2}, and Takuo SUGANUMA^{†1,†2}

^{†1}Department of Electrical Information and Physics Engineering, School of Engineering, Tohoku University

^{†2}Cyberscience Center, Tohoku University

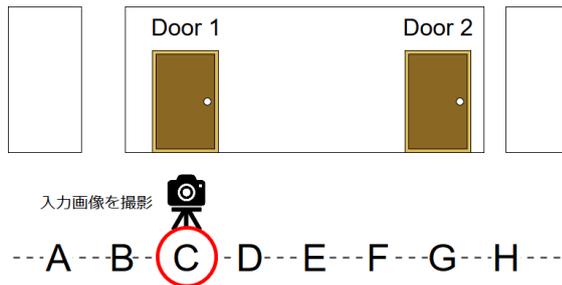


図 3: エリアシングが発生する想定環境

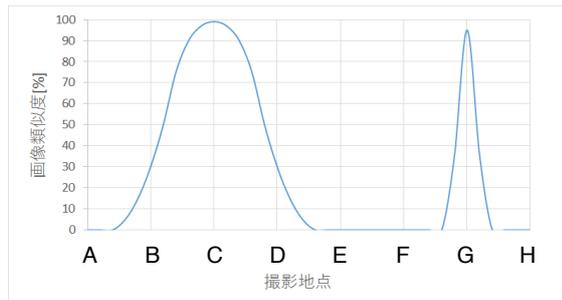


図 4: データベース内画像との類似度の予想グラフ

場合、扉の配置から、地点 C,G でのデータベース画像との類似度が両方とも高くなると予想される。しかし、周辺環境の違いにより、隣接した地点でのデータベース画像との類似度は変わることが予想される。地点 C の周辺である地点 B,D は、入力画像の撮影地点に近いので、図 4 に示すように、そのデータベース画像は入力画像との類似度が比較的高くなると考えられる。一方、地点 G では扉の影響で局所的に入力画像との類似度が高くなるが、その周辺（地点 F,H）の環境は異なるため、そのデータベース画像との類似度は低くなると考えられる。

そこで、地理的に近い画像を考慮するために、それぞれのデータベース画像と類似度を計算した後に、周辺地点での画像も含めて類似度が高い地点を選択することで、正しい自己位置推定を行うことができると考えられる。

これとは逆に、移動しながら連続して撮影した複数の入力画像と各データベース画像との類似を求めるときを考えると、各データベース画像について、複数の入力画像との類似度の平均値を求めると、正解地点では入力画像の撮影位置をずらしても類似度に大きな変化が出ないため、類似度の平均は高いまま維持されると予想される。一方で、正解とは異なる地点では、入力画像の撮影位置がずれば類似度が大きく低下することが予想されるため、平均値は低下すると考えられる。したがって、複数の入力画像との類似度を考慮することで正しい自己位置を推定できると考えられる。

4 設計

4.1 地理的に近い画像を考慮する手法の設計

以下のアルゴリズムにより自己位置推定を行う。

1. 入力画像とデータベース内画像の記述子を計算。文献 [2] を参考に、記述子の計算には Gist [3] を使用
2. 入力画像とデータベース内画像の類似度を計算。画像の記述子をベクトルで表現し、ユークリッド距離を類似度として出力
3. それぞれの画像に対し、一定の範囲の周囲の画像を含めた平均の類似度を計算
4. 平均の画像類似度が最も高い地点を自己位置として出力

各入力画像に対し、類似度を考慮する（類似度の平均を求める）データベース画像の範囲により、自己位置推定の精度が変化することが考えられる。考慮する範囲が狭すぎる場合、正解地点以外でも平均類似度が高くなり、エリアシング問題を改善できない可能性がある。また、考慮する範囲が広すぎる場合は、正解地点での平均類似度が低くなり、やはり推定精度は改善できないと考えられる。したがって、考慮するデータベース画像の範囲を変えて手法の評価を行う必要がある。

4.2 複数の入力画像を用いる手法の設計

複数の入力画像と各データベース画像との平均類似度を求める場合は、入力画像を撮影する間隔（移動距離）により自己位置の推定精度が変化することが考えられる。したがって、この場合も、入力画像を撮影する間隔を変えて手法の評価を行う必要がある。

5 おわりに

本稿では、カメラの映像を用いた自己位置推定において生じるエリアシング問題に対処するため、地理的に近い複数の画像の類似度を考慮する手法について検討した。一枚の入力画像と複数のデータベース画像との平均類似度、あるいは、複数の入力画像と一枚のデータベース画像との平均類似度を用いることで、自己位置推定の精度が向上可能であると考えられる。

今後は、本稿の提案手法に基づいて実装を行い、実環境において提案手法の有用性を評価する実験を行う予定である。

参考文献

- [1] 吉澤菜津子ほか: 屋内位置情報における推定技術の開発と新しいサービスの展開について, *Intec Technical Journal*, Vol. 13, pp. 44–51 (2013).
- [2] Payá, L. et al.: Modeling Environments Hierarchically with Omnidirectional Imaging and Global-Appearance Descriptors, *Remote Sensing*, Vol. 10, No. 4, p. 522 (2018).
- [3] Oliva, A. and Torralba, A.: Building the gist of a scene: The role of global image features in recognition, *Progress in brain research*, Vol. 155, pp. 23–36 (2006).