

H-005

画像認識に基づく現在地推定法の撮影シーン依存性の検証

Evaluating performance of a localization method based on image recognition in several scenes

沖谷卓哉*
Takuya Okitani

島田裕*
Yutaka Shimada

谷口行信*
Yukinobu Taniguchi

1 はじめに

近年、屋内施設の構造の複雑化に伴い歩行者ナビゲーションの必要性が高まっている。GPSの電波が届かない屋内で利用可能な現在地推定法として、Wi-FiやBluetoothなどにより、無線通信機器から発せられる信号の受信強度を用いる手法が挙げられる。しかし、これらの手法では屋内施設に多数の無線通信機器を設置する必要があるため、機器の設置や管理にコストがかかる。また、設置可能な場所が制限されるという問題点も存在する。そこで、本稿では画像認識技術により現在地を推定する手法に着目する。画像認識技術を用いた手法では、事前にデータベースに登録した位置情報付き参照画像と、歩行者が撮影したクエリ画像の特徴を比較して現在地を推定する。これにより、特別な機器を用いることなく歩行者ナビゲーションが可能となる。一方で、データベース内に類似した参照画像が複数存在することで誤認識が生じる問題点がある。本稿では、まず現在地の推定結果の正しさを検証する際に、その結果に対して信頼度を定義する。次に信頼度が閾値よりも低い場合、推定結果を保留することで誤認識を防ぐ。その際、閾値を適応的に変化させることで撮影シーンに対する依存性を低減する手法を提案する。

2 画像認識に基づく現在地推定法の流れ [1]

我々が文献 [1] で提案した画像認識による現在地推定の流れを図 1 に示す。また、図 1 中の各項目の詳細を以下に示す。

特徴抽出: データベース (以下 DB) 内の参照画像から SIFT 特徴量を抽出し、RootSIFT を用いて特徴を記述する [2]。

クエリ画像抽出: 動画からクエリ画像を抽出し、参照画像と同様に特徴を抽出する。

探索範囲制限: DB 内の全参照画像を探索対象とすると、類似した参照画像が複数存在する場合に誤認識が生じ、離れた地点にジャンプする問題がある。この問題を解決するため、歩行者の移動可能距離に応じて、参照画像の探索範囲を適応的に制限する。

ホモグラフィ行列の推定: 探索範囲の制限により限定した参照画像とクエリ画像の特徴点ペアを求めてホモグラフィ行列を推定する。推定したホモグラフィ行列により、誤った特徴点ペアを除いた正しい特徴点ペア (以下インライア) の数を求める。

信頼度判定: インライアの数上位 1 位と 2 位の参照画像のインライア数をそれぞれ I_1 , I_2 とする。こ

のとき、 I_1/I_2 を信頼度と定義し、推定結果が正しいかどうか判定をする。具体的には、信頼度 I_1/I_2 の値がある閾値 θ よりも小さい場合、すなわち DB 内に類似した参照画像が複数存在する場合は現在地推定の結果を保留し、信頼度 I_1/I_2 が閾値 θ よりも大きい場合は、インライア数が最大となる参照画像の撮影位置を現在地として更新する。

現在地出力: 信頼度判定により信頼度が高いと判定された場合、インライアの数が多い画像が撮影された位置を現在地として出力する。

我々は文献 [1] で、駅構内のデータを用いて信頼度判定の有効性を示したが、閾値 θ が固定であるため撮影シーンが変わると精度が低下する問題があった。本稿では、撮影シーンに依存しない適応的な閾値の決定手法を提案する。

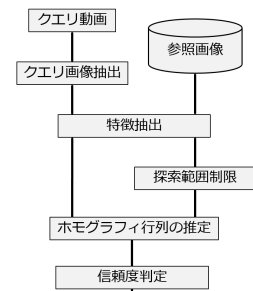


図 1: 画像認識技術を用いた現在地推定の流れ

3 提案手法

提案手法では、インライア数が上位 1 位と 2 位の画像を重ね合わせたときの重なり領域に着目して閾値を決定する。図 2(a) に示すように、参照画像 A, B の重なる領域が大きい場合は、信頼度 I_1/I_2 が小さくても現在地推定結果は正しい場合が多い。

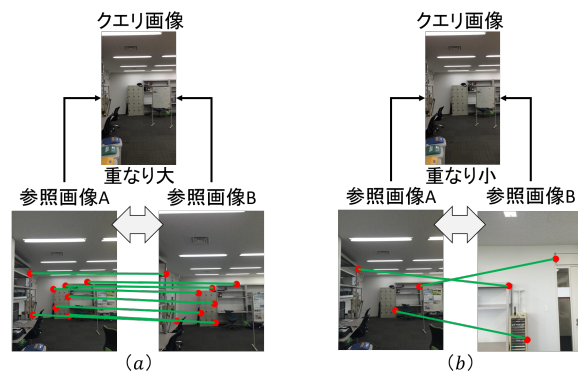


図 2: 参照画像 A, B に対する重なり領域の例

*東京理科大学工学研究科経営工学専攻

一方, 図 2(b) に示すように, 重なり領域が小さい場合は, 参照画像 A, B の間の信頼度 I_1/I_2 が偶然大きくなったとしても現在地推定結果は信頼できない. そこで, 閾値を決定するため DB 内の参照画像の組 A, B に対してホモグラフィ行列 H を求める. 画像 A の画素数を n_A , 画像 A をホモグラフィ行列 H によって画像 B に射影変換した後の重なり領域の画素数を $n_{A'}$ として, 面積比 $n_A/n_{A'}$ を算出する. この面積比を, ホモグラフィ行列 H が算出可能な全ての参照画像の組に対して求め, 推定結果が正しいかどうかの判定の際に閾値 θ として利用する. ただし, 一般に, 片方の画像をもう一方の画像に射影する行列は, 1つの参照画像の組につき 2つ得られる. 提案手法では, クエリ画像とのインライア数が上位 1 位の画像 A を, 上位 2 位の画像 B に射影した場合の面積比を求め, $I_1/I_2 > n_A/n_{A'}$ を満たしたとき, インライア数上位 1 位の画像を採用する. これは, 画像の重なり領域に含まれる特徴点の個数とその面積に比例するという仮定に基づいている. これを利用することで, 単にインライア数のみを用いるのではなく, 画像中の重なり領域を考慮した適応的な信頼度判定を行うことが可能となる.

4 実験条件

本手法の有効性を示すため, 東京都内の某駅構内(屋内)と研究室内で取得したデータを用いて実験を行った. 駅構内では図 3 に示す 140 地点, 研究室内では図 4 に示す 38 地点で事前に参照画像を撮影し, 位置情報を付加してデータベースに保存した. ただし, 各地点で 4 方向を撮影し参照画像を取得した. 駅構内の各撮影地点間隔は約 5m, 研究室内の撮影間隔は約 2.5m である. 撮影機器は iPhone5s, 画像サイズは 540×720 [pixel] とした. 駅構内は 2つ, 研究室内は 5つの異なるクエリ動画を用意し, 120fps の動画から 40 フレームごとに最もエッジ検出数の多い画像をクエリ画像として利用した. また, 各クエリ画像に対して事前に正解地点を与えておき, 正解範囲(各施設での参照画像の撮影間隔)内で現在地を推定できた割合を正答率として評価する. 本実験では, 図 1 において, (i) 固定の閾値を用いた信頼度判定の従来手法 1[1], (ii) 探索範囲の制限を行わない従来手法 2, と比較した.

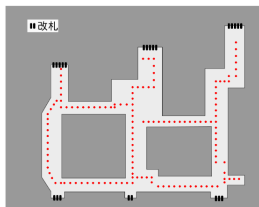


図 3: 駅構内の概略図 (赤丸は参照画像の撮影地点)

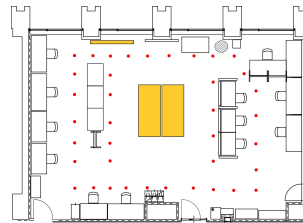


図 4: 研究室の概略図 (赤丸は参照画像の撮影地点)

5 実験結果

駅構内および研究室内で撮影した動画を用いて現在地推定を行った結果をそれぞれ図 5, 図 6 に示す. 駅構

内では, 従来手法 1 の固定閾値が上手く機能したため, 3つの手法の中で最も高い正答率を示しているが, 提案手法との大きな差はない. 研究室内では, 従来手法 1 の固定閾値が上手く機能せず大幅な精度の低下が生じた一方で, 提案手法は約 7 割の平均正答率で現在地を推定したことが確認できる. なお, 探索範囲を制限しない従来手法 2 との正答率の差は大きくないが, 探索範囲を制限することでホモグラフィ行列の推定のコストを大幅に削減でき効率的である. これより, 提案した閾値決定手法の有効性を確認した.

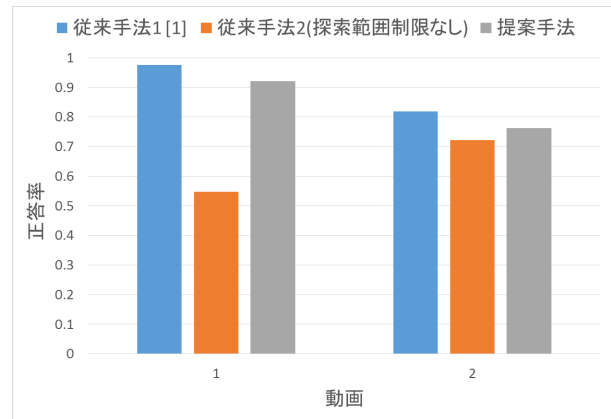


図 5: 駅構内での現在地推定実験の結果

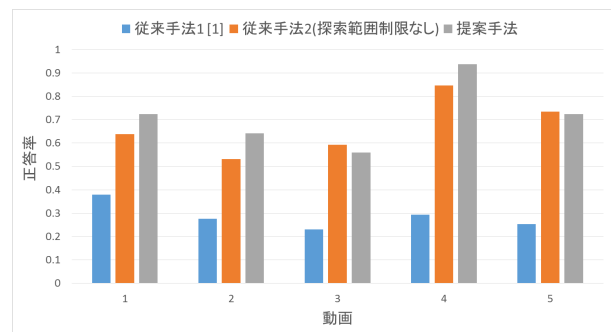


図 6: 研究室内での現在地推定実験の結果

6 おわりに

本稿では, 撮影シーンに依存しない頑健な現在地推定手法の開発を目的とし, 撮影シーンに依存した信頼度判定の閾値を自動で決定する手法を提案した. 具体的には, DB 内の参照画像の組の重なり領域を面積比とし, それに基づいて閾値を決定した. 駅構内と研究室内というシーンの異なる施設で実験を行ったことにより本手法の有効性を示した.

参考文献

- [1] 沖谷卓哉, 島田裕, 谷口行信, 「画像認識に基づく歩行者の移動軌跡推定」, 2016 年電子情報通信学会総合大会, D-12-77, 2016.
- [2] R. Arandjelovic and A. Zisserman, “Three things everyone should know to improve object retrieval,” in Proc. of CVPR, 2012.