

車載動画からの位置推定システム

松原一樹† 齋藤友彦†

湘南工科大学†

1. はじめに

近年、自動車やバイクにカメラを搭載し、撮影された車載動画が無料動画サイトなどに数多く投稿されている。旅行やドライブの参考になる一方、実際に撮影された場所へ行こうと考えた際、撮影された場所が分からないことがある。

本研究では、動画に映った 106 系の案内標識 (図 1) のみから撮影場所を推定するシステムを提案する。システムの概要は次の通りである。まず、画像認識の一つである YOLO (You Only Look Once) [1]を用いて案内標識を検出する。次に、Python 製 OCR (Optical Character Recognition) EasyOCR[2]を用いて標識に書かれている地名及び距離を認識する。認識した地名及び距離から、本研究で考案した位置推定法を用いて撮影場所を推定する。本研究では、実験により提案位置推定システムの精度を検証する。



図 1 案内標識

2. 関連研究

案内標識の位置を推定する文献[3]では、標識の色と形状の特徴を用いた標識領域の抽出及び、動画像中の抽出領域の変化と GPS (Global Positioning System) を用いた標識位置の推定を行っている。また、車載搭載センシング装置を用いた案内標識の位置、画像、記載された案内地名を抽出する文献[4]がある。位置情報の抽出は GNSS (Global Navigation Satellite System) による車両の走行軌跡を基にしている。画像の抽出はディープラーニングによる物体検出を用いている。案内地名の抽出は Google 社が提供する画像解析サービス Cloud Vision API を用いている。

3. 位置推定システム

3.1 画像認識及び文字認識

本システムでは、まず YOLO を用いて地名及び距離が記載された案内標識の認識を行う。案内標識の画像データを用意し、アノテーションソフトを用いて教師データを作成する。学習を行い、案内標識を検出した一例を図 2 に示す。次に、EasyOCR を用いて案内標識から文字認識を行う。文字を検出した一例を図 3 に示す。図 3 では、地名を緑色、距離を赤色の枠で囲っている。



図 2 案内標識の検出結果



図 3 文字の検出結果

3.2 地名及び距離からの位置推定法

文字認識の結果から、読み取った地名の座標を中心、距離を半径とした円を描く。そして、各円と円の交点を求め、交点の中心座標を計算する。なお、円と円が交わらないときは、半径の小さい方の円を 1km ずつ大きくし、交わった点の座標を求める。得られた座標の平均を推定位置の座標とする。

図 3 の例における、位置推定の過程を図 4 に示す。図 4 では、文字認識により得られた円を赤線、円と円の交点を青色マーク、交点の中心座標をオレンジ色マーク、推定した位置座標をピンク色マークで示している。なお、図 4 の緑色マークは案内標識が実際に設置されている位置を示す。

Location Estimation System from In-Vehicle Video

†KAZUKI MATSUBARA Shonan Institute of Technology

†TOMOHIKO SAITO Shonan Institute of Technology

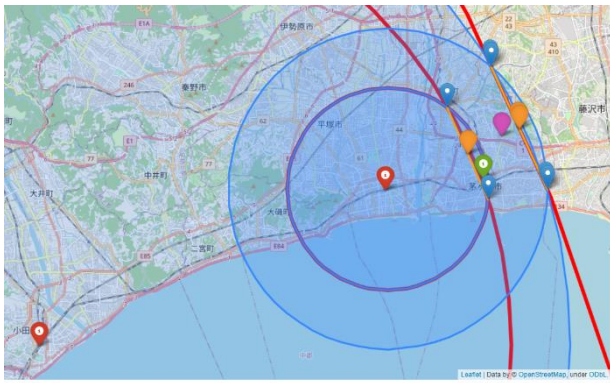


図4 円を描いた結果

3.3 位置推定システム全体

位置推定システム全体では、まず YOLO, EasyOCR を用いて案内標識から地名及び距離を認識する。認識の結果、地名及び距離のペアが揃った場合に限り、3.2 で述べた位置推定法を用いて位置座標を推定する。

4. 実験

4.1 画像認識及び文字認識の精度

案内標識が映ったテスト画像 50 枚を用いて実験を行う。また、テスト画像に含まれる地名及び距離は合計 270 個である。実験の結果、YOLO の認識率は 98% (50 個中 49 個認識), EasyOCR の認識率は 85% (270 個中 232 個認識) であった。

4.2 地名及び距離からの位置推定法の精度

ここでは、3.2 で述べた位置推定法の精度を検証するため、3.1 の画像認識及び文字認識は行わず、目視によって地名及び距離を入力し、位置推定を行う。4.1 と同様の画像を用いて実験を行った結果、平均推定誤差は約 3821m であった。また、階級の幅を 1km とし、位置推定誤差のヒストグラムを描くと図 5 のようになる。

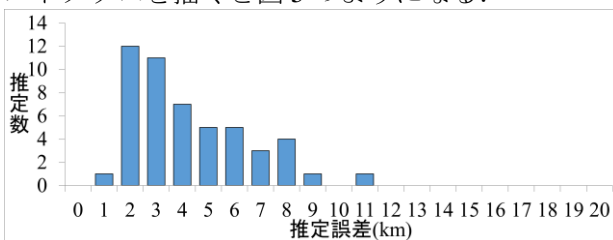


図5 4.2の位置推定結果

4.3 位置推定システム全体の精度

4.1 と同様の画像を用いて位置推定システム全体の精度を検証する。3.1 の画像認識及び文字認識を用いて地名及び距離のペアが揃った画像は 50 枚中 31 枚であった。これら 31 枚の画像に対し 3.2 の手法を適用した結果、平均推定誤差は約 5825m であった。階級の幅を 1km とし、位置推定誤差のヒストグラムを描くと図 6 のようになる。

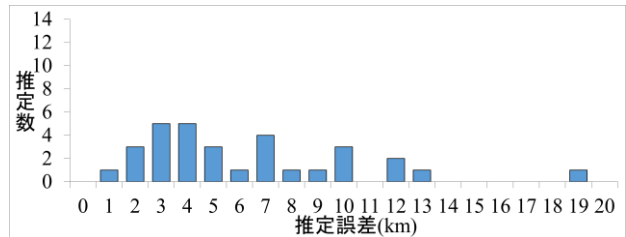


図6 4.3の位置推定結果

5. 考察

4.3 の結果より、位置推定システム全体の平均推定誤差は約 5825m であった。この距離を時速 40km で走行した場合、移動時間は約 9 分であり、撮影場所を推定するアルゴリズムとしては十分であると考えられる。また、4.2 の結果より、3.2 の位置推定法のみでの平均推定誤差は約 3821m であった。従って、画像認識精度及び文字認識精度を上げることで、更に位置推定誤差を小さくできると考えられる。

本研究の応用として、推定した位置座標を用いた近隣の観光や飲食店等を検索、表示するアプリの作成が考えられる。その際、10km 以内の誤差であれば時速 40km で走行した場合、約 15 分で到着でき、妥当な情報を表示できるため、本システムによる位置推定は有効だと考えられる。

6. おわりに

本研究では、車載動画に映った案内標識から位置を推定するシステムを提案した。実験により位置推定システムの精度を検証し、有効性を示した。

謝辞

本研究は JSPS 科研費 18K11585, 22K02811 の助成を受けたものです。

参考文献

- [1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You only look once : Unified, real-time object detection, ” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.779-788, 2016.
- [2] Jidda, “aidedAI/EasyOCR, ” <https://github.com/JaidedAI/EasyOCR>, (2023年1月11日閲覧).
- [3] 富永裕之, 内村圭一, 脇阪信治, “動画像解析による道路案内標識の自動位置推定, ” モバイルコンピューティングとワイヤレス通信高度通信システム, pp. 129-134, (2001).
- [4] 神波修一郎, 篠田浩太郎, “MMS で取得した点群及び画像データからの 道路案内標識データ抽出, ” 建設コンサルタント業務研究発表論文集, pp. 37-40, (2020).