

観光サイクリスト行動情報解析による停車地点の自動検出とその特徴分析

Preliminary analysis of stop points detected by density-based clustering of GPS data of sightseeing cyclists

加納 崇平
Sohei KANO

広島市立大学 言語音声メディア工学研究室
Language and Speech Research Laboratory, Hiroshima City University

概要 国内外から多くのサイクリストが集まるしまなみ海道を対象に、一般社団法人しまなみジャパンの協力を得て、GPS ロガーを使ってレンタサイクルを利用する観光サイクリストの行動情報を収集した。サイクリストの居住地や自転車の種類等の属性別に、サイクリストがよく留まる地点を検出するため、密度に基づくクラスタリングを行い、地図上に可視化できるシステムを構築した。また、停車地点の特徴を説明する文の自動生成を目指して、停車地点の特徴分析を行った。まだ予備的な段階ではあるが、その検討結果を報告する。

1 はじめに

地域の現状と課題を理解し、地域の課題解決に貢献する人材を育成するために、広島市を中心とする広域都市圏（広島市、呉市、竹原市、三原市、廿日市市、東広島市、山口県岩国市、山口県柳井市など 24 市町）と尾道市を対象に、観光関連データを収集し、教育研究活動に活用している。その一環として、国内外から多くのサイクリストが集まるしまなみ海道を対象に、一般社団法人しまなみジャパンの協力を得て、GPS ロガーを使ってレンタサイクルを利用する観光サイクリストの行動情報を収集した[1]。

現在、しまなみレンタサイクル事業では、年間約 72,000 台の実績を上げており、平均で 200 台/日となっている。今後の事業継続の安定化を考えた際、観光サイクリストの種別による行動パターンを把握することが非常に重要になっている。しまなみジャパンとしても課題の洗い出しや若者目線での対応策について検討する際のデータが不足している状況もあり、GPS ロガーで集めたデータの活用を大学と一緒に検討したいという考えであった。

サイクリストの居住地や自転車の種類等の属性別に、サイクリストがよく留まる地点を検出するため、密度に基づくクラスタリングを行い、地図上に可視化できるシステムを構築した[2]。属性別に結果を可視化するだけであっても、しまなみ海道における今後の観光サービスの検討に役立つと期待できるものである。

しかしながら、地図上に表示される情報が大量であるため、その特徴を把握することは簡単ではない。そこで、停車地点の特徴を説明する文の自動生成を目指して、停車地点の特徴分析を行った。まだ予備的な段階ではあるが、その検討結果を報告する。

2 観光サイクリストがよく留まる地点の検出

2.1 データの収集・解析

データの解析には、2018 年 2 月から 2018 年 12 月の期間に収集した、314 グループの属性データと位置情報データを用いる。属性データには、日付、電動アシストやクロスバイクなどの自転車の種類、ユーザの居住地、グループの人数などが含まれる。また、観光サイクリストの位置情報の取得には、GPS ロガーを自転車に装着することで実現した。本研究では、これらのデータに対して密度に基づくクラスタリングを行い、解析を行う。

密度に基づくクラスタリング手法には、DBSCAN (Density-based spatial clustering of applications with noise) [3] を用い、解析結果を地図上で見やすくするための改良を行った。DBSCAN とは 1 つのデータから一定の半径内に閾値以上のデータがあればクラスタを拡大していく手法である。よって、パラメータは半径と周りにあるデータ数の閾値である。半径内にあるデータの数がパラメータの条件を満たすとそのデータはコアとなり、コアと半径内にあるデータは同一クラスタとなる、半径内にコアが複数ある場合はそのコアとコアの半径内にあるデータはすべて同じクラスタに属する。つまり、データの密度が高いところをクラスタにする手法である。

GPS の軌跡データから速度 0 の留まっている地点を集め、DBSCAN を行った。GPS データに付与された属性は 4 つあり、居住地（県内、県外、海外）、性別（男、女）、自転車の種類（クロスバイク、電動アシスト、軽快車）、季節（春=3-5 月、夏=8-9 月、秋=10-11 月、冬=2 月と 12 月）である。DBSCAN のクラスタリング結果から多くのサイクリストが留まる地点がクラスタとして把握できる。

しかし、クラスタリングしただけでは、地図に表示する際に各クラスタに属するコアの平均を求めるため、視覚的にわかりづらくなる。よって、クラスタリング後に各クラスタに属するコアの平均を求め、その平均から一番離れたコアの距離を測る。それにより、地図で表示する際に各クラスタのコアの平均の 1 点のみのプロットと、その平均から一番離れたコアの距離からなる円を表示することができ、視覚的にわかりやすくなる。また、円を表示する際に各クラスタのコアの平均を中心にして、コアからの最大距離を半径にしているため、従来の DBSCAN よりも人が多いところだけを円で表すことができる。

2.2 可視化システム

DBSCAN のパラメータを半径 5 メートル、データ数の閾値を 5 個としたときの地図、半径 20 メートルでデータ数の閾値を 20 個としたときの地図を作成した。詳しい手順は以下のとおりである。

- ① 半径 5 メートル以内にデータ数が 5 つある地点に対してクラスタリングを行う。また、クラスタリングされた地点をコアとする。
- ② 同一半径内に複数のコアがあった場合、同一クラスタとしてクラスタリングを行う。
- ③ クラスタリング後、同一クラスタに属するコアの平均を求め、平均値から最も離れたコアとの距離を計算し、求めたコアの平均の値を 1 点のみプロットし、同時に距離から成る円を表示する。
- ④ 同様の作業を閾値 20 メートル・20 個でも行う。

可視化に用いる地図には OpenStreetMap を採用した。また、作成した地図のプロットした点に緯度、経度とそのクラスタ内にいる人数が分かるようにした。

3 説明文の自動生成を目指した基礎検討

佐藤等[4]は、レンタサイクル事業で用いられている自転車の GPS データを収集し、分析結果をもとに観光サイクルマップの開発を行っている。このシステムには、GPS データ、盛岡自転車会議の資料および Google Maps API を利用した「走行軌跡の表示」、「スポット情報の表示」および「ルート情報の表示」の 3 つの機能が備わっている。本研究では、説明文の自動生成を目指す。そのための第一歩として、実際にユーザの停車地点を Google Earth とストリートビュー機能を用いて、尾道市街地を対象に人手による分析を行う。

表 1 に停車地点の分析結果と説明文の生成の例を示す。注意の対象に注目すると、停車地点付近は眺めの良い海辺やお店、交差点が存在していることが確認された。また、説明文の例では、注意の対象に対応する説明文がある程度限られた種類で出現していることがわかった。

表 1: 停車地点の分析結果と説明文の生成の例

緯度	経度	人数	注意の対象	説明文の例
34.403508	133.193655	75 人	海辺, 向島渡船	景色, 自転車⇄船
34.403055	133.192953	57 人	海辺, 尾道駅前棧橋	景色
34.398781	133.177454	22 人	海辺	景色
34.406661	133.200864	14 人	尾道渡船フェリー乗り場	自転車⇄船
34.410607	133.201864	13 人	喫茶店	食事
34.410203	133.202726	9 人	商店街内	買い物
34.406917	133.200501	8 人	鮮魚店	買い物
34.410535	133.202287	8 人	土産店, 交差点	買い物, 信号
34.407895	133.200765	7 人	商店街入り口	自転車⇄徒歩
34.402963	133.192410	7 人	海辺, 尾道駅前棧橋	景色
34.402053	133.190504	6 人	海辺, 食事処, 自転車ショップ	景色, 食事, 買い物

4 まとめ

本報告では、尾道市街地を対象に基礎検討を行った。実際にストリートビュー機能を用いて、人手でユーザの停車地点を分析したところ、停車地点付近は眺めの良い海辺やお店、交差点が存在することが確認された。また、これらに対応する説明文の例では、限られた種類で表現できることがわかった。今後は、ユーザの注意の対象の自動判定すること、注意の対象に対応する説明文を生成することが挙げられる。注意の対象の自動判定には、まず、API などを利用することで位置情報周辺のデータを収集することを検討している。また、注意の対象に対応する説明文には、機あらかじめ用意した説明文のカテゴリに機械学習を用いて分類することや深層学習のテキスト生成技術を利用することを考えている。

5 謝辞

レンタサイクルのデータ収集にご協力いただいた一般社団法人しまなみジャパンの皆様へ心から感謝申し上げます。

「自主プロジェクト演習」でデータを解析してくれた広島市立大学大学院情報科学研究科の赤松知昭君に感謝する。

6 参考文献

- [1] 植松敏美, 難波英嗣, 田村慶一, 石野亜耶, 竹澤寿幸, 観光関連データベースの内容の拡充—観光サイクリストの行動情報収集と有用性の検討—, 観光情報学会第 18 回研究発表会, 2018 年 11 月.
- [2] 赤松知昭, 観光サイクリスト行動情報データの解析手法に関する研究, 広島市立大学大学院情報科学研究科「自主プロジェクト演習」研究成果発表会, 2020 年 2 月.
- [3] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases in Noise,” Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226-231, 1996.
- [4] 佐藤遼二, 工藤峻輔, 窪田諭, 市川尚, 阿部昭博, 自転車 GPS データによる観光サイクルマップの開発とそのまちづくり応用, 情報処理学会第 74 回全国大会, 1ZH-7, Vol. 4, pp. 875-876, 2012.

