

観光客の散策行動を考慮したエリア単位の行動分析

Tourist Behavior Analysis Considering Stroll in Sightseeing Area

武田直人 *1

Naoto TAKEDA

西村拓哉 *2

Takuya NISHIMURA

戸田浩之 *2

Hiroyuki TODA

関洋平 *3

Yohei SEKI

*1筑波大学大学院 図書館情報メディア研究科

Graduate School of Library Information and Media Studies, University of Tsukuba

*2日本電信電話株式会社 NTT サービスエボリューション研究所

NTT Service Evolution Laboratories, NTT Corporation

*3筑波大学 図書館情報メディア系

Faculty of Library, Information and Media Science, University of Tsukuba

In this paper, we propose a method clarifying tourists' behaviors in each sightseeing area from trajectories with focusing on the tourists' strolling activities in addition to staying activities. First, we extract staying segments and strolling segments from individual trajectory with a method for extracting arbitrary-shaped stay regions. Next, we adopt a density-based clustering method to centroids of staying segments and strolling segments. Consequently, the proposed method can extract sightseeing areas which contain more information of tourists' behaviors than existing methods. We evaluated our method based on real 1,500 tourists' trajectories around Fukuoka city. As a result, our method expanded the number of data to about 9 times in Dazaifu Tenman-gu Shrine area comparing to the existing methods.

1. はじめに

観光地における観光客の行動分析は、人気ルートの発見やそれに基づく情報推薦などを行う上で重要である。観光客の行動を捉える手法として、スマートフォンなどから得られる観光客の移動軌跡情報を利用した分析が挙げられる。移動軌跡とは特定の人から取得された測位点の集合であり、測位点とは GPS や Wi-Fi など取得された緯度・経度と、それが記録された時刻の 3 つの値の集合である。従来研究では、移動軌跡から、ユーザがある一点に一定時間以上留まった部分を抽出する滞留点抽出を行い、それに基づいた分析を行っている [Zheng 09][熊谷 15]。しかし、観光客の行動を分析する上では、一つの店舗への訪問などの、ほぼある一点に一定時間以上留まる行動だけでなく、有名な街並みをゆっくりと歩くなど、低速で移動する行動についても分析する必要がある。本研究では、このような行動を「散策」と呼ぶ。たとえば、福岡県にある人気の観光エリアである太宰府天満宮では、観光客は、広い境内を散策しながら観光を行う。従来の、一点への滞留に基づいた分析手法では、これらの観光エリアにおける観光客の散策を捉えることができない。

そこで、本研究では、単一の移動軌跡情報から、一点への滞留と低速で移動する散策の二つの行動をそれぞれ抽出することで、観光客の滞留だけでなく、散策も併せて抽出し、それらを用いてエリア単位での観光行動の分析を行う手法を提案する。まず、距離閾値と時間閾値を用いて単一の移動軌跡から滞留している測位点集合を抽出できる手法である D-Star[Nishida 15]を用いて、(1) 小さな距離閾値で適用することによる「滞留中の測位点」集合の抽出、(2) 残った移動軌跡に対して、(1) よりも大きな距離閾値で適用することによる「散策中の測位点」集合の抽出を行う。

連絡先: 西村拓哉, 日本電信電話株式会社 NTT サービスエボリューション研究所, 〒239-0847 神奈川県横須賀市光の丘 1-1, nishimura.takuya@lab.ntt.co.jp
この研究は第 1 著者が NTT サービスエボリューション研究所でのインターンシップ期間中に実習生として行ったものです。

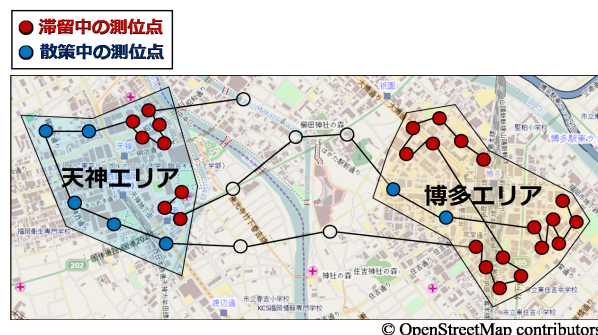


図 1: 本研究では、移動軌跡情報から滞留点と散策点を抽出し、それらをクラスタリングすることで人がよく訪れるエリアの抽出とその性質の分析に取り組む。図は 2 名の観光客の移動軌跡情報から滞留と散策を抽出し、それらをクラスタリングして二つのエリアを発見している様子を示す模式図である。

点」集合の抽出を行う。次に、このようにして得られた、単一の観光客の「滞留中の測位点」集合と「散策中の測位点」集合のそれぞれの重心を求める。最後に、求めた大勢の観光客の重心集合に対し、密度ベースのクラスタリング手法である OPTICS[Ankerst 99] を適用し、クラスタを形成する。これにより、観光客の移動軌跡情報から、従来のクラスタリング手法では捉えられなかった散策を含んだクラスタを抽出できる。また、観光客の滞留と散策の比率を分析でき、それぞれの観光エリアにおける、観光客の特徴を分析できる。

提案手法の概要を具体例を用いて、説明する。図 1 に、福岡県における 2 名の観光客の移動軌跡の例と、移動軌跡のクラスタリングの概要を示す。2 名の移動軌跡の各測位点は、D-Star を用いた 2 段階のクラスタリングによって、「滞留中の測位点」(赤点)と「散策中の測位点」(青点)がそれぞれ抽出されている。それぞれの段階で得られた測位点集合の重心を、OPTICS によってクラスタリングすることにより、天神エリアと博多エ

本研究の貢献は、(1) 単一の移動軌跡から滞留を抽出する手法を、距離閾値を変更することにより、散策も抽出できるように拡張した手法の提案、(2) 大勢の人が散策するような観光エリアを対象とした行動分析の実現、(3) 実データを用いた実験による、滞留と散策を考慮することの有効性の検証、の三つである。

本節では、移動軌跡を用いて人の行動や地域の特徴を分析する研究を紹介する。[Zheng 09] らは、大勢の観光客の移動軌跡から、人気の場所の発見と、人気の観光ルートを発見する手法を提案している。まず、大勢の移動軌跡から、滞留している測位点を抽出し、クラスタリングを行う。このようにして得られる、大勢が滞留するエリアにおいて、流入するユーザと流出するユーザについて、HITS アルゴリズムを適用することで、人気の観光エリアを発見している。また、[Ye 09] らは、GPS を用いて取得した移動軌跡から、個人の頻出する行動パターンの発見を行う手法を提案した。まず、実験対象者の移動軌跡から、滞留している測位点を抽出し、クラスタリングを行う。これにより、個人が頻繁に訪れるエリアを抽出することができる。さらに、エリア間の遷移パターンを分析し、実験対象者の生活における頻出行動パターンを発見している。また、[熊谷 15] らは、本研究でも用いるデータである、福岡の訪日観光客の移動軌跡から抽出された滞留点と、属性情報を組み合わせた分析手法を提案し、属性ごとの観光地における回遊パターンや、人気の観光エリアを明らかにした。

3. 提案手法

本研究では、単一の移動軌跡から滞留している測位点を抽出する手法である D-Star を用いて、観光客の移動軌跡から「滞留中の測位点」と「散策中の測位点」を抽出する。

示す．なお，測位点 5 は外れ値であり，測位点 6 から測位点 7 の間は，測位点が得られなかったこととする．まず，測位点 1 の移動窓 $q = 3$ 内であり，距離閾値 ϵ 内である，周辺の測位点集合を抽出する（測位点 1, 2, 3）．得られた周辺の測位点集合内の経過時間は， $6 - 1 = 5$ であり， $m_{time} = 15$ よりも小さいため，測位点 1 はコアポイントとはならない．このような操作を順に行うと，測位点 4（周辺の測位点集合：測位点 3, 4, 6, 7，経過時間：18）と測位点 7（周辺の測位点集合：測位点 4, 7, 8, 9, 10，経過時間：24）がコアポイントとなる．2 つのコアポイントの周辺の測位点集合には重なりがあるため，測位点 4 と測位点 7 の周辺の測位点集合は結合し，滞留点集合のクラスタとして採用する（滞留点集合のクラスタ：測位点 3, 4, 6, 7, 8, 9, 10）．

3.2 OPTICS によるエリア発見

本研究では、OPTICS を、大勢の観光客の、それぞれの移動軌跡における「滞留中の測位点」集合の重心と「散策中の測位点」集合の重心に適用することで、観光エリアを抽出する。



© OpenStreetMap contributors

図 3: 提案手法によって抽出された観光エリア。数字はエリア ID を示す (0: 博多駅周辺 41: 天神 149: 太宰府天満宮)。

3.1 節で、各測位点は、「滞留中の測位点」と「散策中の測位点」に区別したため、それぞれの比率を計算することで、抽出できた観光エリアにおける観光客の行動分析ができる。

4. 実験：福岡地域の訪日観光客の移動軌跡を用いた行動分析

4.1 実験方法

本節では、福岡地域の観光地における観光客の移動軌跡情報から、提案手法による観光エリアの抽出と、滞留のみを考慮した比較手法による観光エリアの抽出について、それぞれの測位点数を比較する。また、抽出できた観光エリアにおける観光客の行動分析を行うことで、提案手法の有効性を検証する。提案手法では、D-Star を 2 段階適用し、得られた重心集合から、OPTICS により観光エリアを抽出する。滞留のみを考慮する比較手法では、D-Star を 1 段階適用し、得られた重心集合から、OPTICS により観光エリアを抽出する。また、各観光エリアについて、滞留点のみを考慮する手法では得られない、滞留と散策別の観光客数や、滞在時間に基づく分析を行う。

D-Star の距離閾値は、1 段階目の滞留を抽出するための D-Star 適用では $\epsilon_1 = 15 (m)$ 、2 段階目の散策を抽出するための D-Star 適用では $\epsilon_2 = 200 (m)$ とした。なお、どちらの D-Star でも、移動窓 $q = 3$ 、 $m_{time} = 30 (min)$ とした。散策に関するパラメータ設定は、直進で時速 400 m 以下の低速移動を散策とみなしていることを示す。また、実験結果に示す OPTICS のパラメータは $Eps = 30 (m)$ 、 $MinPts = 25$ とした。

4.2 実験データ

実験データは、訪日外国人観光客を対象としたスマートフォンの観光アプリケーションによって収集した。福岡市および周辺観光地を訪れる外国人観光客に無料 Wi-Fi に簡単に接続できる機能や、観光情報を提供するアプリケーションを配布し、あらかじめ許諾を得た観光客から滞在中の移動履歴情報を収集している。実験に用いたユーザはランダムサンプリングで抽出した 1,500 名で、測位点の合計は 551,415 点である。

4.3 実験結果

観光エリアの抽出と測位点数の比較: 図 3 に、抽出できた観光エリアの一部を示す。図中のエリアは、得られたクラスターの外側の測位点を利用した凹包 [Edelsbrunner 83] で可視化している。エリア ID 0 は、博多駅周辺のエリア、エリア ID 41 は、繁華街である天神を含むエリア、エリア ID 149 は、太宰府天満宮を含むエリアである。

また、滞留と散策を考慮する提案手法と、滞留のみを考慮する比較手法の、それぞれのエリアにおける、測位点数を図 4 に

示す。図中の「全エリア」は、提案手法で抽出できた全てのエリアにおける測位点数を表す。なお、同エリア内で、同一の観光客が滞留と散策の両方を行っている場合は、それぞれにカウントしている。提案手法を用いることで、従来のクラスタリング手法では、捉えられなかった観光エリア内の観光客の散策を抽出でき、それぞれの観光エリアにおいて、分析対象となる測位点数の拡充ができた。特に、散策を行うことが多い太宰府天満宮では、提案手法を利用することで比較手法に比べて、約 9 倍という大幅な測位点数の拡充ができた。

それぞれの観光エリアにおける観光客の行動分析: 図 5 に、博多駅周辺エリア、天神エリア、太宰府天満宮エリアの時間帯別の観光客数（滞留と散策）を示す。なお、グラフは観光エリアに滞留ないし散策を開始した時刻について、ユーザ数をプロットしている。図 5 から、博多駅周辺の観光客は午前中と夜間に多いことがわかる。夜中に関しては、ホテルに滞在しているユーザの移動軌跡がその時間帯から取得開始されたと考えられる。また、観光客は、午前中に博多駅から他の観光エリアに移動し、夜に戻る、という可能性が示唆された。さらに、博多駅周辺では、多くの観光客が滞留することが明らかとなった。これは、電車やバスを待つ観光客や、駅周辺で食事をする観光客が多いためだと考えられる。天神では、すべての時間帯について、観光客数はほぼ一定であった。また、昼から夜にかけては、滞留する観光客と散策する観光客は、同程度であることが明らかとなった。太宰府天満宮では、午前中から昼にかけて、観光客数が多いという結果となった。また、ほとんどの観光客が滞留ではなく、散策をすることが明らかとなった。太宰府天満宮では、従来の滞留だけを考慮する行動分析では、測位点数が少なく、このような分析を行うことができない。散策の多い観光エリアでの、提案手法の有効性を確認できた。

4.4 考察

本研究では、従来の移動軌跡を利用した行動分析手法では考慮されていない観光客の散策を、滞留と区別して抽出した。実験データを用いた実験を通して、太宰府天満宮における散策する観光客は、滞留する観光客に比べて多いことなどの各観光エリアの特徴を明らかにできた。このような分析結果は、観光ルート推薦にとって有用な知見である。たとえば、太宰府天満宮では、散策する観光客が多いことから、次の観光エリアとして、休憩可能な観光エリアの推薦ができる。

また、観光客の散策を抽出することにより、各観光エリアでの測位点数を拡充できる。これにより、各観光エリアにおける観光客の行動をより多く捉えられるようになるため、測位点数を担保した上での分析ができる。たとえば、従来の滞留だけを考慮した分析では、太宰府天満宮のほとんどの観光客の行動を

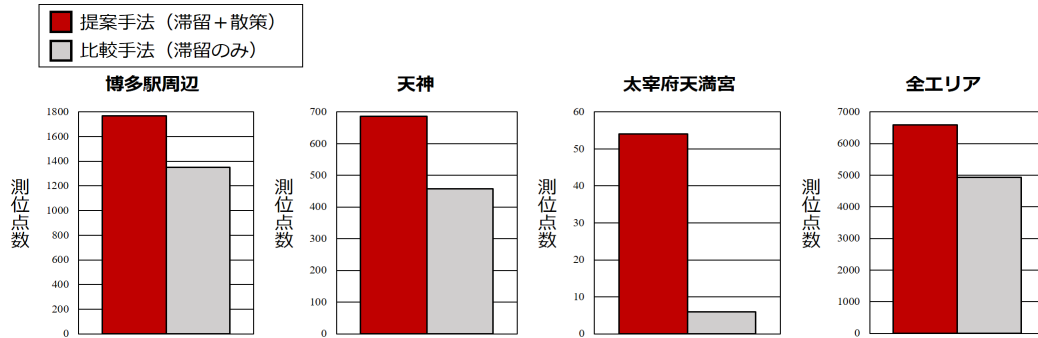


図 4: 各観光エリアにおける提案手法（滞留と散策を考慮）と比較手法（滞留のみを考慮）で抽出できた測位点数

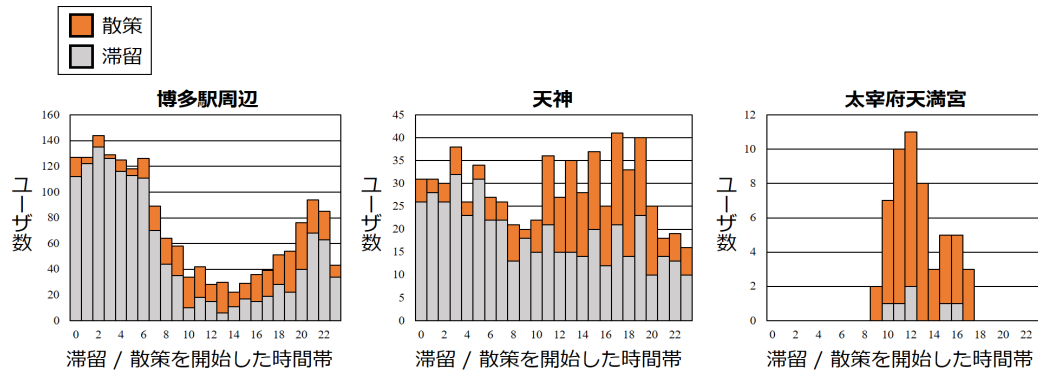


図 5: 各観光エリアにおける、その時間帯に滞留/散策を開始したユーザー数

捉えることができないが、提案手法を用いることで、観光客が滞留ないし散策を開始する時間帯を明らかにできた。これらのことから、提案手法は、観光エリアごとの観光客の行動を分析する上で、有効な手法といえる。

5. おわりに

本研究では、単一の観光客の移動軌跡情報に対して、2段階のクラスタリング手法を適用することで、滞留している測位点だけでなく、散策中の測位点を抽出できる手法を提案した。このようにして得られた「滞留中の測位点」と「散策中の測位点」の重心をクラスタリングすることで、観光エリアの抽出と、それぞれの観光エリアにおける観光客の行動を、滞留と散策に区別して分析できる。また、測位点数を確保した上で、各種分析が可能となる。

実データを用いた実験では、各観光エリアについて、測位点数が拡充でき、提案手法の有効性を確認した。また、複数の観光エリアにおける、観光客の行動が明らかとなった。たとえば、福岡の太宰府天満宮エリアでは、散策をする観光客が多いという知見を得た。

今後の課題としては、居住国や性別などの観光客の属性と組み合わせた分析が挙げられる。

参考文献

- [Ankerst 99] Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J.: OPTICS: Ordering Points to Identify the Clustering Structure, in *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data (SIGMOD 1999)*, pp. 49–60, Philadelphia, PA, USA (1999)
- [Edelsbrunner 83] Edelsbrunner, H., Kirkpatrick, D., and Seidel, R.: On the Shape of a Set of Points in the Plane, *IEEE Transactions on Information Theory*, Vol. 29, No. 4, pp. 551–559 (1983)
- [Nishida 15] Nishida, K., Toda, H., and Koike, Y.: Extracting Arbitrary-shaped Stay Regions from Geospatial Trajectories with Outliers and Missing Points, in *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science (IWCTS 2015)*, pp. 1–6, Seattle, WA, USA (2015)
- [Ye 09] Ye, Y., Zheng, Y., Chen, Y., Feng, J., and Xie, X.: Mining Individual Life Pattern Based on Location History, in *Proceedings of the 10th International Conference on Mobile Data Management (MDM 2009)*, pp. 1–10, Taipei, Taiwan (2009)
- [Zheng 09] Zheng, Y., Zhang, L., Xie, X., and Ma, W.-Y.: Mining Interesting Locations and Travel Sequences from GPS Trajectories, in *Proceedings of the 18th International Conference on World Wide Web (WWW 2009)*, pp. 791–800, Madrid, Spain (2009)
- [熊谷 15] 熊谷 雄介, 今井 良太, 松林 達史, 佐藤 吉秀, 堀岡 力: 非負値複合テンソル因子分解を用いた訪日外国人観光客の回遊行動分析, 信学技報 研究報告バイオ情報学 (BIO), Vol. 2015-BIO-42, No. 3, pp. 1–5 (2015)