

軌跡データのみを用いた観光スポット 遷移モデルの構築

Tourist Transition Model among Sightseeing Spot based on Trajectory Data

渡部 岳志 *1 笠原 秀一 *2 飯山 将晃 *2 美濃 導彦 *2
Takeshi Watabe Hidekazu Kasahara Masaaki Iiyama Michihiko Minoh

*1 京都大学工学部情報学科

Undergraduate School of Informatics and Mathematical Science, Kyoto University

*2 京都大学 学術情報メディアセンター

Academic Center for Computing and Media Studies, Kyoto University

In recent years, congestion of famous sightseeing spots and public transportation has become a problem in tourism areas. It is a way to solve it that local governments forecast congestion by a congestion simulation based on a tourist transition model and adjust the operation of public transportation. In this research, we propose a method to construct a tourist transition model among sightseeing spots based on trajectory data. We construct a tourist transition model using actual trajectory data of tourists who stayed around Kyoto city.

1. はじめに

近年観光都市では観光スポットや公共交通機関の混雑が問題となっている。それを解消する方法として行動モデルに基づいた混雑予測を元に交通機関の運行を変更するといった方法が考えられる。既存手法ではモデル構築に必要な情報を得るために Wikipedia など外部の情報を用いている。しかし、これらの手法では既知の観光スポットしかモデルに反映できない。そのため、実データに基づいて観光客が実際に訪問した観光スポット間の遷移を表す行動モデルが求められている。

本研究では観光客の軌跡データのみから構築できる観光スポット遷移モデルの構築手法を提案する。具体的には観光客の移動速度に着目して観光客が集まる点(観光客集中心点)を抽出する。その後、観光客集中心点を観光スポットと移動中に立ち寄る場所(経由地)に分類し、観光スポット間の遷移確率を求め、観光スポット遷移モデルを構築する。

2. 軌跡からの観光スポットの抽出

軌跡データから観光スポット遷移モデルを構築するには、測定誤差を除去した軌跡データから観光スポットを抽出する必要がある。観光スポットでは観光客は足を止めて観光を行うので、速度に着目すれば観光スポットを抽出できる。しかし、観光客はバス停や駅など経由地でも足を止めることがあるので、両者を区別しなければならない。観光スポットと経由地では観光客の移動方法の割合が異なり、前者では徒歩が多く、後者は非徒歩が多いと考えられるので、この特徴を用いて両者を分類する。以下、処理の詳細について述べる。

2.1 軌跡データの前処理

本研究で用いる軌跡データは、観光客が携帯する GPS 機器により計測した観光客の位置の時系列データである。軌跡データ内の各点は、軌跡 ID(id)、緯度(lat)、経度(lon)、時刻(t)を持つ。GPS の軌跡データは観光客が携帯する GPS 機器か

ら収集できる。そのため、容易に広い範囲のデータを収集することができる。

軌跡データは GPS の測定誤差を含んでいるので除去する。軌跡データが測定誤差を含んでいると訪問していない地点を観光スポットまたは経由地として推定してしまうからである。大きな誤差を含んだ軌跡データは移動速度が急に変化すると考えられる。観光客の移動速度は移動方法によって変化するが誤差による移動速度の変化はそれよりも大きい。このため、軌跡データの各点に前後の緯度 lat と経度 lon から求めた速度 v を付与し閾値 v_e より v が大きい点を異常値として除去する。

2.2 観光客集中心点の抽出

本節では、測定誤差を除去した軌跡データから、観光客が一定時間以上速度を低める点を観光集中心点として抽出する。観光客集中心点には観光スポットと経由地が含まれる。

観光客が一定時間以上速度を低めるパターンには、観光・乗り換えによる停止とそれらと関連のない一時的な立ち止まりの 2 種類がある。前者を停止点群として抽出する。一時的な立ち止まりは観光や乗り換えに比して短時間である。本研究では一時的な立ち止まりの最長時間を t_s 、観光・乗り換えや一時的な立ち止まりの中の移動速度の最大値を v_s とする。移動速度 v に対し、 t_s 以上の間連続して v が v_s を下回ったとき、その間の点を停止点群として抽出する。次に停止点群を mean-shift 法によってクラスタリングすることで観光客集中心点を得る。こうして得られた各クラスタの重心を観光客集中心点 c とする。

2.3 観光客集中心点の観光スポットと経由地への分類

本節では、前節で抽出した観光客集中心点を、観光スポットと経由地に分類する。観光客は観光スポットでは観光を行うため主に徒歩で移動し、経由地では乗り換えを行うため主に非徒歩で移動する傾向を持つ。この傾向を用いて観光客集中心点を分類する。

まず全ての観光客集中心点にボロノイ分割を用いてモデルの対象地域を分割した空間を集中心点空間 S とする。集中心点空間 S において徒歩で進入可能な領域を I_{walk} 、非徒歩で進入可能な領域を I_{trans}, I_{walk} と I_{trans} の和集合を I_{all} とする(図 1)。観光スポットと経由地の移動方法の違いから集中心点空間 S 内で領域 I_{all} のうち、領域 I_{trans} が占める面積の割合は観光ス

連絡先: 連絡先 〒 606-8501 京都市左京区吉田二本松町 京都大学学術情報メディアセンター南館 美濃研究室, 渡部 岳志,
E-mail:watabe@mm.media.kyoto-u.ac.jp

ポットでは小さく、経路地では大きいと仮定する。この仮定を用いて観光客集中点を観光スポットと経路地に分類する。

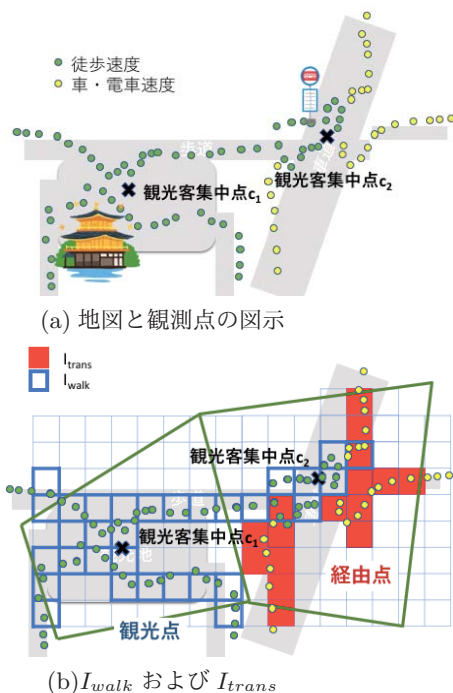


図 1: I_{trans} と I_{walk} の抽出例

I_{walk} と I_{trans} の面積を得るために I_{walk} と I_{trans} を画像として得る。まず、軌跡データから速度 v_w 以下の点を徒歩、それ以上の点を非徒歩で移動しているとする。集中点空間 S 内に存在する軌跡データの内、徒歩で移動している点からヒートマップを作成し、グレースケール画像を二値化することで二値画像 I_{walk}^c を作成する。これらの二値画像において値を持つ画素に対応する領域は観光客が徒歩で進入可能であるとし、値を持つ画素の数 $M(I_{walk}^c)$ をその面積とする。同様に軌跡データから非徒歩で移動している点から抽出し同様の二値画像 I_{trans}^c を作成し、 $M(I_{trans}^c)$ をその面積とする。 I_{trans}^c と I_{walk}^c の和集合を I_{all}^c とし、 $M(I_{all}^c)$ をその面積とする。 $M(I_{all}^c)$ と $M(I_{trans}^c)$ の面積比

$$R_c = \frac{M(I_{trans}^c)}{M(I_{all}^c)}$$

を画像の特徴量とする。特徴量 R_c が閾値 R_α より小さいものを観光スポット、大きいものを経路地として分類する。

3. 観光スポット遷移モデルの構築

3.1 観光スポットネットワークの構築

本節では、集中点ネットワークを定義し、このネットワークから得られた集中点間遷移確率を用いて観光スポットネットワークを構築する。

まず、集中点ネットワークとして、観光客集中点をノードとして持ち、観光客集中点間での遷移が軌跡データ内に存在するならばエッジを持つような有向重み付きネットワークを考える。このとき、エッジの重み $p_{i,j}$ は観光客集中点 i から行われた全遷移のうち、観光客集中点 j に遷移した割合とする。これがあるノード i から別のノード j へ直接遷移する確率とする。

次に観光スポットネットワークを構築する。遷移行列 $M = (p_{i,j})$ 、ノード i からノード j へのネットワーク上での最短距離を $d_{i,j}$ とする。このとき、観光客は最短経路のみを用いる訳ではない。そのためノード i とノード j の間を遷移する上で経路する可能性のあるノードの数の上限を $l_{i,j}$ とおくと求めたい $p_{i \rightarrow j}$ は以下のように求まる。

$$p_{i \rightarrow j} = \alpha \left(\sum_{r=d_{i,j}}^{l_{i,j}} M^r \right)_{ij} \quad (1)$$

本研究では $l_{i,j}$ は $d_{i,j}$ の関数であると考え、 $l_{i,j} = 2 \times d_{i,j}$ とした。 α は正規化項であり。

$$\alpha = \frac{1}{\sum_{c \in K} p_{i \rightarrow c}} \quad (2)$$

とする。 K は観光スポットからなる集合である。観光スポットをノードとして持ち $p_{i \rightarrow j}$ をエッジの重みとしてもつネットワークを観光スポットネットワークと呼ぶ。

3.2 観光スポット遷移モデルの構築

本研究では観光客の観光スポット遷移を単純マルコフモデルでモデル化する。つまり、観光客が次に遷移する観光スポットは現在訪問している観光スポットのみに依存する。観光スポットネットワーク上の対応するノード間のエッジの重みを、観光スポット間遷移確率とする。

4. 実験

4.1 軌跡データの前処理

本研究で使用したデータは、班別自由行動中の修学旅行生(班長が端末を携帯)及び旅行代理店スタッフの計 579 名の GPS 移動軌跡である [笠原 15]。実験データは、携帯端末を対象者に携帯したまま一日移動して貰い、その行動を携帯端末にインストールした GPS ロガーアプリケーションで収集した。GPS はおおよそ 1 秒毎に位置を計測している。対象期間は 2015 年 12 月である。実験地域としては修学旅行生が多数訪れている京都市周辺とするために緯度 34.80 度以上、35.15 度以下、経度 135.65 度以上、135.85 度以下の地域とした。また、閾値 v_e を 180(km/h) とし異常値を除去した。また、GPS 以外の A-GPS や wifi で測位された点も測定精度が低い場合全て除去した。実験データは 9,530,489 の観測点を含み、前処理によって実験地域外の観測点と異常値を削除した後で 5,108,676 の観測点を含んでいた。

4.2 観光客集中点の分類

本研究では実験に用いた軌跡データの特性をもとに、停止点群を抽出する際の定数として $v_s = 3.6$ (km/h)、 $t_s = 200$ (秒)を用いた。また、mean-shift 法のカーネル関数としてガウシアンカーネルを用い、バンド幅として 0.0010 を用いた。

提案手法によって観光客集中点が 354 箇所抽出された。抽出した観光客集中点に人の手によるタグ付けを行った結果は観光スポットが 170 箇所、経路地が 171 箇所であった。またその他に渋滞や交差点など観光客集中点の誤抽出が 13 箇所含まれていた。これに対し提案手法による観光客集中点の特徴量 R_c の分布が図 2、ROC 曲線が図 3 である。図 2 の特徴量 R_c の分布から観光スポットにおいて特徴量 R_c はほとんどの場合 0.4187 より小さく、経路地において R_c は同じく 0.4187 より大きいことがわかる。そのためモデルの構築には 0.4187 を閾

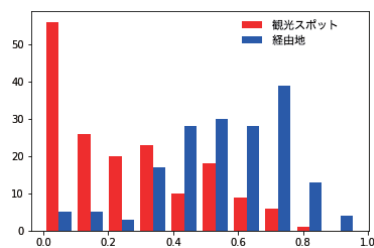
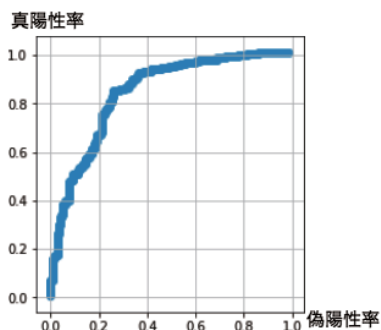
図 2: 提案手法における特徴量 R_c の分布

図 3: 提案手法の ROC 曲線

値として用いた。人の手によるタグ付けを正解としたとき、分類の正答率は 76.6% だった。

次に提案手法を用いて正しく分類できた例について観光スポットと経由地をそれぞれ図 4, 5 に示す。

図 4 は銀閣寺に対応する観光客集中心点において作成した二値画像である。この観光客集中心点の面積比 R_c は 0.040 と非常に小さい。下図より銀閣寺内の道が I_{all} として抽出されていることがわかる。また、上図より非徒歩で進入可能な領域 I_{trans} として抽出されている領域がほとんどないことがわかる。この領域内では車で移動できないが上図には非徒歩で進入した領域が存在している。これは GPS の測定誤差などによる影響で速度が正しく認識できていないためと考えられる。

図 5 は、山科駅上の観光客集中心点において作成した二値画像である。この観光客集中心点の特徴量 R_c は 0.855 と非常に大きい。下図より非徒歩で進入可能な領域として線路が抽出されていることがわかる。また、上図より進入可能な領域は非徒歩で移動できる領域とほとんど重なっていることがわかる。

次に分類が失敗した例を図 6 に示す。

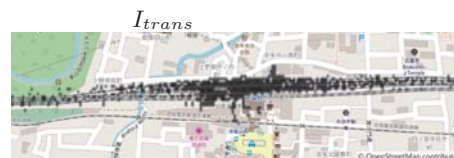
図 6 は東福寺に対応する観光客集中心点において作成した二値画像である。この観光客集中心点は東福寺の境内に存在している。地図から分かる通り集中心点空間内には東福寺と直接の関連を持たない道路や線路が西側に存在している。このため、この観光客集中心点の特徴量 R_c が 0.626 と大きく経由地に分類された。このように近隣に他の観光客集中心点がなく、集中心点空間が広く定まってしまう場合には空間内に本来観光スポットではない道路や線路が含まれてしまい、誤分類が発生することがある。このような誤分類を防ぐための方法としてボロノイ分割で空間を分割するのではなく停止点群の分布から集中心点空間を得る手法を用いることなどが考えられる。



非徒歩で進入可能な領域

進入可能な領域 I_{all} 図 4: 観光スポットへの分類成功例, 銀閣寺, $R_c = 0.040$
図左中央の太い道は参道であり徒歩以外では移動できない

非徒歩で進入可能な領域

進入可能な領域 I_{all} 図 5: 経由地への分類成功例, 山科駅, $R_c = 0.855$

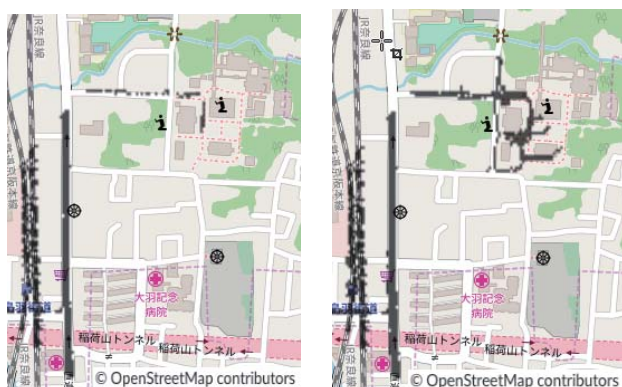
4.3 観光スポットネットワークの構築

実験データセットに対し提案手法を用い観光スポットネットワークを構築した。例として金閣寺の観光スポットからの遷移確率が高い観光スポット 10 箇所に関し、地図とともに図 7 に示す。

また、モデル全体の評価を行うために観光スポットをいくつかの地域に分け、それぞれの地域間の遷移確率を求めた。その結果を表 1 に示す。表 1 が示す通り、ほとんどの地域において、同一地域へ遷移する確率が最も高く、それに次いで清水寺または京都駅へ遷移する確率が高い。また、地理的に隣接している地域間では、遷移確率が高い傾向が見られる。この結果は、観光客は京都駅や清水寺のような人気のあるスポットには、距離にかかわらず遷移することがあるが、それ以外の場合は同一あるいは近隣の観光スポットに遷移することを示していると考えられる。

表 1: 地域間遷移確率

地域	目的地									
	京都駅	清水	河原	嵐山	岡崎	金閣	伏見	東山	銀閣	その他
京都駅	0.492	0.128	0.126	0.019	0.005	0.055	0.059	0.020	0.017	0.080
清水	0.101	0.591	0.078	0.013	0.016	0.029	0.017	0.006	0.064	0.080
河原町	0.072	0.223	0.453	0.036	0.011	0.027	0.022	0.010	0.031	0.111
嵐山	0.122	0.202	0.117	0.232	0.003	0.108	0.029	0.005	0.014	0.163
岡崎	0.101	0.237	0.068	0.013	0.224	0.029	0.012	0.004	0.167	0.140
金閣	0.108	0.098	0.090	0.057	0.015	0.347	0.013	0.004	0.098	0.164
伏見	0.229	0.132	0.061	0.015	0.008	0.050	0.360	0.023	0.019	0.099
東山	0.320	0.144	0.194	0.054	0.004	0.032	0.049	0.073	0.017	0.108
銀閣寺	0.018	0.268	0.118	0.006	0.093	0.044	0.013	0.001	0.281	0.151
その他	0.131	0.215	0.110	0.071	0.019	0.090	0.016	0.017	0.069	0.264

図 6: 観光スポットへの分類失敗例, 東福寺, $R_c = 0.626$

5. おわりに

本研究では、軌跡データのみから観光客集中点を抽出し、観光客集中点を観光スポットと経由スポットへ分類した。そして、得られた観光スポットと経由地を用いて観光客の観光スポット間遷移をモデル化する手法を提案した。この手法では経由地を考慮することで既存手法より観光客の移動に即したネットワークを構築できる。

今後の課題として、mean-shift 法のバンド幅や集中点ネットワーク上で観光客が取りうる経路、特徴量の閾値を自動的に決定する手法の考案などが挙げられる。

謝辞

本研究は JSPS 科研費 17K00438 の助成を受けたものです。

参考文献

- [笠原 15] 笠原秀一, 森幹彦, 椋木雅之, 美濃導彦.: 環境制約を用いた GPS 移動軌跡からの旅行者状態推定, 人工知能学会全国大会論文集 (CD-ROM), Vol.29, ROMBUNNO.2B3-NFC-02A-4, 2015.

遷移先	遷移確率
A 金閣寺舍利殿	0.136
B 銀閣寺	0.072
C 龍安寺前土産物屋	0.064
D 龍安寺	0.045
E 北野天満宮	0.045
F 銀閣寺前土産物屋街	0.045
G 白峯神社	0.037
H 京都駅	0.031
I 清水寺前土産物屋街	0.030
J 京都駅バスターミナル	0.023



図 7: 金閣寺からの遷移