

目的地予測に基づく移動軌跡からの逸脱行動の検出

渡部 岳志[†] 笠原 秀一^{††} 飯山 将晃^{††}

[†] 京都大学 大学院 情報学研究科 〒 606-8501 京都市左京区吉田本町

^{††} 京都大学 学術情報メディアセンター 〒 606-8501 京都市左京区吉田本町

E-mail: [†]{watabe,hidekazu.kasahara,iiyama}@mm.media.kyoto-u.ac.jp

あらまし 近年、観光客の増大による混雑が社会問題化している。その一因に、旅行者の道迷いによる観光地での滞留がある。道迷いを自動検知して旅行者に警告することで、こうした滞留を減らすことができる。本研究では、道迷いを含む遠回り移動を逸脱行動と定義し、その検出手法を提案する。逸脱行動の検出には目的地情報が必要であり、事前に目的地を知ることは難しいため、予測で代替する。だが、移動軌跡からのリアルタイム目的地予測の精度は低く、特に移動開始直後は難しい。そこで本研究では、混合ガウスモデルを用いた目的地推定と逸脱の度合いの推定を同時に行うことで、目的地予測の精度が低い場合でも逸脱行動を検出できる手法を提案する。道迷いの移動軌跡を収集し、提案手法の性能評価を行った。

キーワード 観光情報, 人の移動と行動, 機械学習

1 はじめに

世界的に観光客が増大する傾向が著しい。2018 年の訪日インバウンド観光客は前年比 8.7%増の 3.119 万人に達しており、その経済効果は大きい。反面、京都や鎌倉といった有名観光地の混雑など、オーバーツーリズムの弊害も生じており、その解消は世界的な課題である。そもそも居住地から離れて行動する観光客は、多くの場合観光地の地理に不慣れであるため、道に迷ったり、無駄な滞留を起こしやすい傾向がある。これが観光地の混雑の一因ともなっている。それゆえ、道迷いを無くしていくことで、オーバーツーリズムの弊害を減らすことができる。

GPS を搭載したスマートフォンは広く普及しており、GPS を用いたオンライン地図やナビを用いれば道迷いは防げるという見解もある。しかし、総務省の調査 [1] によれば、Google マップや地図ナビゲーションシステムの利用率は過半数に満たない。そのため、システムが自動的に道迷いを検知し、観光客あるいは団体旅行の引率者に警告するような仕組みが望まれる。例えば、団体旅行の典型である修学旅行において引率教員は管理責任を負っている。自由行動中の生徒が道に迷っていれば対処が必要だが、生徒の数は多く、たとえ GPS があっても手動での発見は難しい。それを自動化できれば、教員の負担を軽減して、生徒の安全を確保できるという効果も期待できる。

道迷いを警告するためには、まず何らかの方法で道迷い行動を検出する必要がある。モバイル GPS 機器の発達により観光客の現在地や移動軌跡をリアルタイムで観測・収集することが容易となった。そのため、GPS センシングデータによって道迷い行動を検出することが考えられる。道迷い行動に関する先行研究としては迷いにくくなる地図情報 [2] や AR を用いた経路情報の提示 [3] [4] などがあり、これらの研究ではルート推薦に対し、そのルートから外れた場合を道迷いと定義している。しかし、観光客は買い物などの意図的な寄り道や混雑回避のため

の迂回を行うことがある。こうした迂回・寄り道行動と道迷い行動は、観光客の主観によって区別され、センシングデータだけでは外形的に道迷いと区別がつかない。観光客の主観を用いて道迷い行動を定義することも考えられるが、観光客は自分自身が迷っていないと認識していても明らかな遠回りしていることがある一方、逆に自分が迷っていると認識しているにもかかわらず、出発地から目的地への最短経路を通ることがある。このように、外形的あるいは内面的に道迷い行動を狭く定義することは困難である。そこで本研究では、団体旅行の引率者の支援システムとして利用することを前提に引率管理者が判断できる道迷いを包括できるよう大きな遠回りを逸脱行動として定義し、逸脱行動の検出を行う。逸脱行動は、最短経路長と実際に移動した経路長を比較し、後者が十分長い場合と定義する。

徒歩で移動している観光客の GPS 軌跡から逸脱行動を検出する際の問題点として、必ず目的地を入力するカーナビとは異なり、観光客の目的地がシステムにとって未知の状態であることが挙げられる。これは様々な要因により目的地を途中で変更することがあり得る観光ドメインの特性である。そのため、最短経路を計算することができない。修学旅行などでは観光する予定の観光地を事前に申告させることがあるものの、混雑などが原因で観光する予定だった観光地に行かない場合や観光中に入手した情報によって予定になかった観光地に行くことがあり、確実ではない。また、観光中に次に行く目的地を逐次入力させることは観光客の負担が大きく、ユーザが能動的に入力する可能性も高くない。そのため、観光客の目的地は未知として予測を行い、予測に従って逸脱行動の検出を行った方が、より現実

に即している。GPS 軌跡から目的地を予測する研究はグリッドによってセルに分割された空間上での遷移を予測するもの [5] [6] [7] やスポット間の遷移を予測するもの [8]、地図上の 1 点を目的地として予測するもの [9] など様々な手法が提案されている。しかし、それらの手法で目的地を一意に求めるのは難しく、いくつかの候

補を挙げ、真の目的地がその中に含まれていれば予測できているというような評価や予測した目的地と真の目的地との距離を用いた評価を行なっている。そのため、既存手法をそのまま用いて逸脱行動を検出する場合、目的地予測を誤ると、迷うことなく移動を行っているのに逸脱行動しているものとして誤検出されてしまう恐れがある。そこで本研究では、陽に目的地を判定するのではなく、複数の目的地候補を推定した上でその推定の尤度を考慮して逸脱行動の検出を行う手法を提案する。すなわち、まず観光客の目的地予測を行う際に全ての目的地候補に対しそこが真の目的地である尤度を求める。その後、各目的地候補に対しその経路から逸脱している尤度を求め、それらを周辺化することで軌跡が逸脱行動である尤度を求める。この尤度が閾値を超えたとき、逸脱行動であるとして検出を行う。

2 関連研究

2.1 迷い行動

迷い行動に関する研究としてスマートフォンで取得したセンシングデータから歩容の特徴量を抽出し、サポートベクターマシンとランダムフォレストでリアルタイムに迷い行動を検出する手法が成本ら [10] [11] や高藤ら [12] によって提案されている。この手法では、心理学の知見である認知マップ [13] の概念を用いて迷い行動を定義し、加速度センサなどの歩容の特徴量に対して、秒単位で迷い行動を推定して高い検出制度を達成している。この手法では、対象者が迷ったと認識した結果、行動が変化するという仮定に基づいているため、本研究が想定しているような対象者が迷っていることに気づかない状況には適応できない。

また、そもそも迷い行動自体を起こさせないように情報提示する手法もいくつか提案されている。その代表的な事例として、迷いにくくなる地図情報 [2] や AR を用いた経路情報の定時やデバイスを用いて迷子を防止する手法 [4] などがある。梶田ら [2] は実際の地図とはランドマークや道路の位置や向きが実際とは異なるがわかりやすい経路情報の提示を提案している。Walk Navi [4] はスマートフォンなどのデバイス上に経路情報をアプリケーションや AR (拡張現実) を用いて提示することで迷子を防止する。これらの手法は基本的にプル型の情報提示を前提にしているが、観光客はそもそもスマートフォンなどのデバイスをそれほど頻繁に見ないため、効果が限定的である。これらに対して本研究では、プッシュ型の情報提示を能動的に行う。

2.2 目的地予測

GPS 軌跡から目的地を予測する場合、連続データである GPS 軌跡を離散化する必要がある。離散化手法としては、予測対象となる地域をグリッドで均等に分割したセル空間上での遷移を予測する手法 [5] [6] [7]、目的地となる観光スポットの所在を位置情報として与え、観光スポット間の遷移を予測する手法 [8]、地図上の 1 点を目的地として予測するもの [9] がある。

セル間遷移を予測する研究として Xue らの研究 [5] [6] や

Krumm らの研究 [7] がある。これらのモデルにおいてはセルの大きさが問題となる。Xue らはセルの一边が 2km の時にもっとも目的地の予測精度が高いとしている [5] [6]。また、Krumm らはセルの一边を 1km としている [7]。しかし、本研究の実験対象になった京都など、多くの観光地では目的地はこれより密集しており、これらの手法で良い結果を得た 1-2km 四方の範囲では、予測する目的地を一つに限定できないという問題がある。

スポット間遷移を予測する研究としては、Ashbrook ら [8] は軌跡から滞留点を用いて POI(Point Of Interest) を抽出し、それらの間の遷移をマルコフモデルによって予測をしている。

地図上の 1 点を目的地として予測する研究としてタクシーの降車地を予測する Bess [9] らの研究がある。Besse ら [9] はタクシーの軌跡をクラスタリングし、新しい軌跡が各クラスタに属する尤度とクラスタ内の軌跡の降車地の中心を用いて降車地の予測を行っている。

本研究では、予測すべき目的地の候補の数が比較的限定され、所在もすべて既知であるという状況を想定している。そのため、スポット間遷移の手法で目的地を予測する。そして、陽に目的地を判定するのではなく、複数の目的地候補を推定した上でその推定の尤度を考慮して逸脱行動の検出を行う手法を提案する。

3 問題設定

3.1 目的地候補

観光客が目的地とする可能性があるスポットの数はあまり多くなく、例えば京都市が観光地と考えているスポットは約 200 箇所しかない [16]。また、旅行口コミサイトであるトリップアドバイザーでは京都市全体で約 2000 箇所のスポットを観光地としている。そのため、本研究では観光客の目的地となりうる場所の数は有限で計算可能な数しかなく、また、全て既知であるとする。各目的地候補は多角形領域で定義され、重複は存在しないものとする。また、目的地候補の集合を \mathcal{D} とする。

3.2 GPS 軌跡

GPS 軌跡は一般的に時刻、緯度、経度を含む観測点で構成されている。このうち、本研究では時刻情報を観測点の順番を求めることにのみ使う。つまり、本研究では軌跡 T^i は

$$T^i = (p_0^i, p_1^i, \dots, p_{l^i}^i) \quad (1)$$

と定義する。ここで $p_k^i \in \mathbb{R}^2$ は軌跡 T^i の k 番目の観測点であり、緯度と経度からなる。また $l^i \in \mathbb{N}$ は軌跡 T^i の観測点の数である。

GPS 軌跡の最後の観測点が目的地候補 $d \in \mathcal{D}$ の多角形領域の内部にある場合、その GPS 軌跡の目的地は d であるとする。

3.3 データセット

本研究では逸脱行動の検出を行うための教師データとして過去の観光客の GPS 軌跡を用い、これを $\mathcal{T} = \{T^0, T^1, \dots\}$ と表記する。この \mathcal{T} に含まれる軌跡は全て目的地候補のうちどれか 1 つを目的地としており、全て既知とする。また、 \mathcal{T} 中の軌

跡のうち、目的地候補 $d \in \mathcal{D}$ を目的地とする軌跡を取り出した集合を \mathcal{T}_d とする。

3.4 地図情報

本研究で用いる地図情報は、交差点の位置、及び、それらの道路による接続関係を記述したデータである。その例を図1に示す。観光客は道路上および目的地候補の多角形領域内のみを通るものとする。

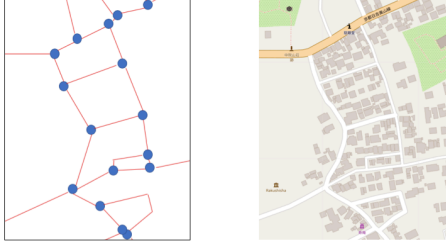


図1 本研究で用いる地図情報と同エリアの一般的な地図。地図情報は一般的な地図から交差点の位置、及び、それらの道路による接続関係のみを抽出したものである

3.5 移動軌跡からの逸脱行動検出

ある観光客の移動軌跡 T がリアルタイムに得られるとし、これを $T^c = (p_0^c, \dots, p_t^c)$ とする。 p^c は出発地点、 p_t^c は現在位置である。なお、この観光客の目的地は目的地集合 \mathcal{D} のいずれかではあるが未知であるとする。本研究では、与えられた T^c より現時点で観光客が逸脱行動の状態にあるかどうかの2クラス識別を行う。

4 提案手法

4.1 GPS 軌跡の前処理

観光客は通常道路に沿って移動を行う。しかし、GPSの観測誤差により道路以外の地点に観測点が存在することがある。また、道に沿って進んでいても、GPS軌跡が道に対して斜めに記録されることがある。そのため、各観測点間の距離を足し合わせたものは実際の移動距離とは異なることがある。これに対し、GPSの各観測点を地図情報によって補正するマップマッチング手法[17][18]が提案されている。本研究では2観測点間の移動距離を求める際、GPSの各観測点を最近傍の交差点に割り当て、それらの交差点間は最短経路を通ったものとして補間する交差点マッチングを行なう。

4.2 目的地予測

新しい軌跡 $T^c = (p_0^c, \dots, p_t^c)$ が与えられたとき各目的地候補 $d \in \mathcal{D}$ に対しそこが真の目的地である尤度 $P(d|p_0^c, \dots, p_t^c)$ を求める。

4.2.1 単一観測点の目的地予測

まず、ある観測点 p が与えられたとき、 p がどの目的地候補を目的地とする軌跡の観測点かの尤度を考える。目的地が同じ観光客は同じような道を通ると考えられる。そのため、 \mathcal{T}_d に含

まれる軌跡中の全ての観測点を考えたとき、点 p の近くに存在する観測点が含まれているなら点 p を通る軌跡の目的地が d である尤度は高く、逆に \mathcal{T}_d に含まれる軌跡が点 p の近くに観測点を持たないなら目的地が d である尤度は低くなると考えられる。よって、 \mathcal{T}_d 中の軌跡に含まれる全ての点の集合を \mathcal{P}_d としたとき、 p がどの \mathcal{P}_d に属するか尤度を求め、それを観測点 p が d を目的地とする軌跡の観測点である尤度とする。

\mathcal{P}_d から次式で与えられる2次元混合ガウスモデル Φ_n^d を作成し、それを用いて p が \mathcal{P}_d に属する尤度を求める。

$$\Phi_n^d(p) = \sum_{k=1}^n \pi_k \mathcal{N}_2(p|\mu_n, \Sigma_n) \quad (2)$$

ここで、 $\sum_{k=1}^n \pi_k = 1$, $\mu_n \in \mathcal{R}^2$, $\Sigma_n \in \mathcal{R}^{2 \times 2}$ で、 \mathcal{N}_2 は2次元正規分布である。

データセットとして \mathcal{P}_d を与えたとき、 Φ_n^d のパラメータ $\Theta_n^d = \{\pi_1^d, \mu_1^d, \Sigma_1^d, \dots, \pi_n^d, \mu_n^d, \Sigma_n^d\}$ はEMアルゴリズムを用いて求める。パラメータ Θ_n^d の尤度 $\mathcal{L}(\Theta_n^d|\mathcal{P}_d)$ は

$$\mathcal{L}(\Theta_n^d|\mathcal{P}_d) = \prod_{p \in \mathcal{P}_d} \Phi_n^d(p) \quad (3)$$

を最大化するように定める。また、これらの結果は n の値によって異なり、 n が小さすぎるとそのモデルでは観測点の分布を十分に表現できず、 n が大きすぎるとモデルが過適合を起こす可能性がある。そのため、適切な n を設定する必要がある。本研究ではベイズ情報量基準 $BIC(k) = -2\ln(\mathcal{L}(\Theta_k^d|\mathcal{P}_d)) + k\ln(|\mathcal{P}_d|)$ を用いて

$$n = \operatorname{argmin}_{k \in [1, \dots, 20]} BIC(k) \quad (4)$$

と定める。

新しい観測点はこうして得られた混合ガウスモデル Φ_n^d から生成されるとし、観測点 p が \mathcal{P}_d に属する尤度を

$$\phi^d(p) = \Phi_n^d(p|\Theta_n^d) \quad (5)$$

とする。

4.2.2 軌跡の目的地予測

次に新しい軌跡 $T^c = (p_0^c, \dots, p_t^c)$ について T^c の目的地が d である尤度 $s^d(T^c)$ を各観測点 p_n^c が \mathcal{P}_d に属する尤度 $\phi^d(p_n^c)$ を用いて以下のように定める。

$$s^d(T^c) = P(T^c|\phi^d) \quad (6)$$

$$= \prod_{i=0}^t \phi^d(p_i^c) \quad (7)$$

その後、この s^d を全ての目的地候補について求め、それらを正規化したものを T^c の目的地が d である尤度 $P(d|p_0^c, \dots, p_t^c)$ とする。つまり

$$P(d|p_0^c, \dots, p_t^c) = \frac{s^d(T^c)}{\sum_{d' \in \mathcal{D}} s^{d'}(T^c)} \quad (8)$$

とする。

4.3 逸脱行動の検出

逸脱行動とは観光客が本来目的地に向かう道から大きく外れている状態のことを指す。しかし、仮に目的地が分かっていたとしても観光客が用いる経路は一意に定まる訳ではない。そのため、本研究では観光客が迷わずに目的地に到着すると仮定した場合の経路長と、実際に観光客が通ってきた経路長、および、今後、観光客が目的地に向かう際に取りうる経路長に着目する。経路長を用いて逸脱行動かを判定する単純な指標として p_0 から d までの最短経路長と実際に用いると予想される経路長の比を用いる手法が考えられる。この場合、同じ距離だけ正しいルートから逸脱したとしても、逸脱度は異なってしまう。例えば、目的地に近い場合と遠い場合では、同じ距離の遠回りを行っても目的地に近い場合の方が逸脱度が高くなる。本研究では出発地から目的地までの距離による影響を除外するため、点 p_t^c から距離 α 遡った地点と点 p_t^c から目的地に距離 α 近づいた予想地点を考え、それらの間の最短経路長と実際の経路長との比を用いて、逸脱度を求める。逸脱度が閾値 ρ を超えた時、逸脱行動であるとして検出を行う。

このとき、 α の値によって同じ軌跡に対しても逸脱度は大きく異なると考えられる。例えば、出発地から目的地への最短経路を移動していた観光客が、ある地点あて最短経路から外れ、現在地点 B におり、また、B から目的地への最短経路上に A が存在する、つまり、地点 B で U ターンすることが現時点における最短経路となるような状況を考える。地点 A から現在地 B までの距離が α より大きいとすると、点 B から距離 α 移動する前の地点と点 B から目的地に距離 α 近づいた地点が一致するため、逸脱度は大きくなる。点 A から点 B までの距離が α より十分小さいとすると、点 B から距離 α 移動する前の地点と点 B から目的地に距離 α 近づいた地点との実際の経路長が最短距離に AB 間の距離の 2 倍を加えたものとなるため、逸脱度は小さくなる。そのため、 α が小さければ最短経路からの逸脱が少しでも逸脱度が大きくなり、 α が大きければ少し最短経路から逸脱しても逸脱度は大きくなる。ない。

以上の考察に基づき逸脱度 $D(d, p_0, \dots, p_t)$ を以下のように定める。

$$D(d, p_0, \dots, p_t) = \frac{\sum_{i=t-l(\alpha)}^{t-1} S(p_i, p_{i+1}) + L(p_t, p_{t+l(\alpha)})}{S(p_{t-l(\alpha)}, p_{t+l(\alpha)})} \quad (9)$$

逸脱度 $D(d, p_0, \dots, p_t)$ が閾値 ρ を超えた時、逸脱行動であるとして検出を行う。 $S(a, b)$ は点 a から点 b への道路ネットワークを考慮した最短距離、 $L(a, b)$ は点 a から点 b への予測移動距離である。 $p_{t+l(\alpha)}$ は p_t から d へ最短経路を用いて移動したとき、 p_t から距離 α だけ移動した地点である。 $p_{t-l(\alpha)}$ は p_t から距離 α だけ軌跡を遡った地点である。なお、本研究ではマップマッチングで GPS 軌跡の移動距離を交差点単位で補正しているため、 p_t から $p_{t+l(\alpha)}$ や $p_{t-l(\alpha)}$ までの距離は α と等しくならない。そのため、 $p_{t+l(\alpha)}$ と $p_{t-l(\alpha)}$ は p_t からの距離が α を超える最も p_t に近い点とする。このとき、観光客が逸脱行動を行っていないと仮定すると $L(p_t, p_{t+l(\alpha)})$ は

$S(p_t, p_{t+l(\alpha)})$ 以上、 $\rho \times S(p_t, p_{t+l(\alpha)})$ 以下である。本研究では $L(p_t, p_{t+l(\alpha)})$ は $S(p_t, p_{t+l(\alpha)})$ から $\rho \times S(p_t, p_{t+l(\alpha)})$ までの一様分布とする。なお、 p_0 から p_t までの移動距離が α 未満だった場合、もしくは、 p_t から d までの最短距離が α 未満だった場合、 $D(d, p_0, \dots, p_t) = 0$ とする。

これを用いて、逸脱度 $D(d, p_0, \dots, p_t)$ が ρ を超える確率を以下のように求める。

$$\begin{aligned} P(D(d, p_0, \dots, p_t) > \rho) \\ &= P\left(\sum_{i=t-l(\alpha)}^{t-1} S(p_i, p_{i+1}) + L(p_t, p_{t+l(\alpha)}) > \rho \times S(p_{t-l(\alpha)}, p_{t+l(\alpha)})\right) \\ &= P\left(L(p_t, p_{t+l(\alpha)}) > \rho \times S(p_{t-l(\alpha)}, p_{t+l(\alpha)}) - \sum_{i=t-l(\alpha)}^{t-1} S(p_i, p_{i+1})\right) \end{aligned} \quad (10)$$

$S(p_{t-l(\alpha)}, p_{t+l(\alpha)})$ は地図情報から、 $S(p_i, p_{i+1})$ はこれまでの軌跡データから得ることができる。

目的地が未知であるときの逸脱行動である確率 $P(D(p_0, \dots, p_t) > \rho)$ を式 (8),(10) から周辺化を用いて以下のように求める。

$$\begin{aligned} P(D(p_0, \dots, p_t) > \rho) \\ &= \sum_{d \in \mathcal{D}} P(D(d, p_0, \dots, p_t) > \rho) P(d | p_0, \dots, p_t) \end{aligned} \quad (12)$$

こうして得られる $P(D(p_0, \dots, p_t) > \rho)$ を観測点が追加されるごとに計算し、その値が閾値 θ を超えたときに逸脱行動として検出する。

5 実 験

5.1 データセット

5.1.1 GPS 軌跡の収集手法

提案手法の評価を行うために京都の嵯峨嵐山駅周辺で 2 日に分けて実験を行い、GPS 軌跡の収集を行なった。実験の対象となったエリアを図 2 に示す。参加者には GPS ロガーとしてスマートフォンと粒度の粗い地図のみを渡し、嵯峨嵐山駅から 9 つの目的地候補から指定した目的地までの徒歩での移動を GPS ロガーで記録した。

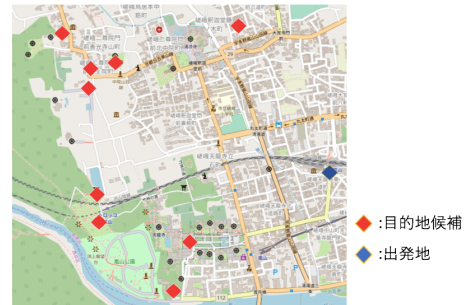


図 2 実験エリア周辺地図、図右の青い菱形が出発地、赤い菱形が目的地候補

5.1.2 逸脱行動のラベル付け

本研究は団体旅行の引率担当者の手による逸脱行動の検出の自動化を目的としている。そのため、提案手法の評価を行うための正解ラベルは観光客に主観に基づくのではなく、客観的な視点から人間が GPS 軌跡を見て逸脱しているかどうかの判定に基づく必要がある。また、提案手法では GPS 軌跡の各観測点ごとに逸脱行動かどうかの判定を行っており、それら全ての観測点に対し人の手でアノテーションを行うことは困難である。そのため、本研究では 1 本の GPS 軌跡を 20 の区間に分割し、それぞれの区間に対しアノテーションを行う。

GPS 軌跡収集に参加していない協力者 45 人が各区間それぞれに対して逸脱行動の判定を行った。人間が逸脱行動の判定を行った場合もその判定に大きな分散があるため、本研究では 45 人中過半数である 23 人以上の人が逸脱行動であると判定した区間を逸脱行動として正解ラベルを作成した。

5.2 目的地予測

提案手法を用いて目的地予測を行なった。その結果を図 3 に示す。横軸は軌跡の全長のうち入力に用いた割合、縦軸が正答率である。これらの軸に対し尤度がもっとも高かった目的地候補が正しい目的地だったときのみ正解とした場合の結果を青い線、上位 3 つに含まれていれば正解としたときの結果を橙色の線で示す。軌跡の序盤のみが入力されている間は正答率が高くないことがわかる。また、目的地に到着する寸前でも必ずしも目的地が予測できるとは限らないことがわかる。これは観光客が目的地の前にも間違えて通り過ぎてしまったり、通り過ぎて別の目的地に行く場合があるためである。

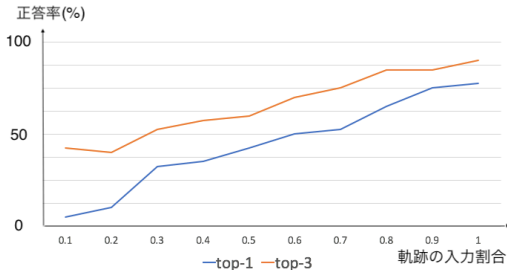


図 3 目的地予測

5.3 ρ の決定

本節では 4.3 で述べた ρ を教師データに基づいて決定する。まず、全ての軌跡について $D(p_0, \dots, p_t)$ を求めた。このとき、複数の ρ と α を用いて ROC 曲線を作成した結果を図 4 に示す。 ρ は、1.25, 1.5, 1.75, 2, 2.25, 2.5 を用いた。 α として 10m, 25m, 50m, 100m, 200m, 400m を用いた。図 4 より提案手法における ROC 曲線の形状は ρ の値によらず一定である。ROC 曲線の形状が類似しているならその分類性能も近い。そのため、提案手法における ρ が分類精度に与える影響は小さいと考えられる。本研究では ρ は 1.75 とした。

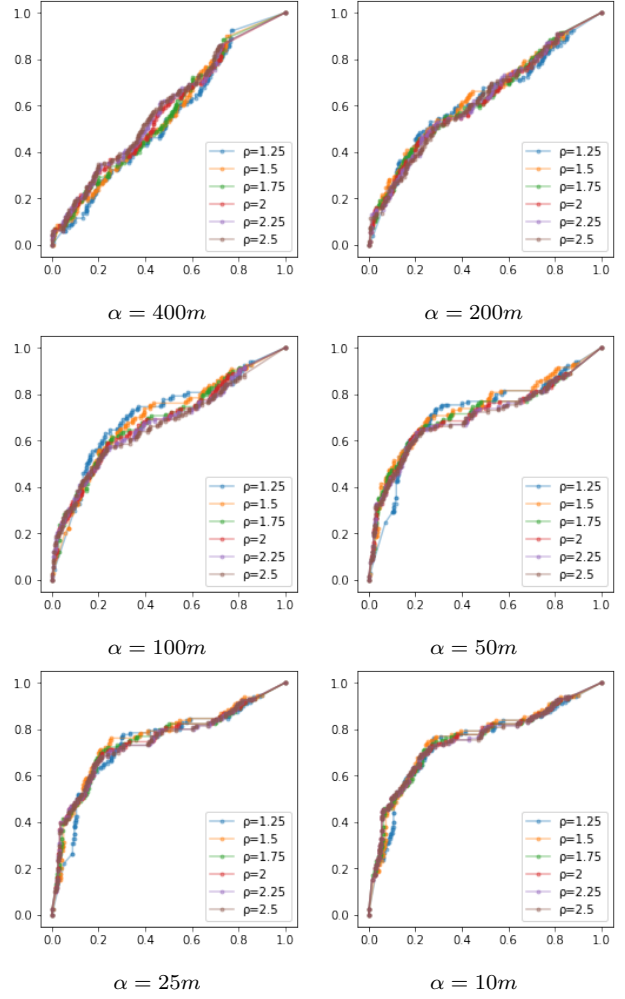


図 4 ρ による ROC 曲線の比較

5.4 比較手法

提案手法の評価のため、周辺化しない手法を比較手法として用いる。

5.4.1 比較手法 1: 目的地予測に基づく周辺化を行わない場合

比較手法 1 では目的地予測に基づく周辺化を行わず、もっとも尤度が高い目的地 d^{max} を真の目的地として逸脱行動の検出を行う。つまり、

$$D_1(p_0, \dots, p_t) = D(d^{max}, p_0, \dots, p_t) \quad (13)$$

とする。ここで d^{max} は

$$P(d^{max}|p_t, \dots, p_0) = \max\{P(d|p_t, \dots, p_0), d \in \mathcal{D}\} \quad (14)$$

より定義される。

5.4.2 比較手法 2: 真の目的地が既知である場合

比較手法 2 では事前に真の目的地 d^{true} が既知であるとして逸脱行動の検出を行う。つまり

$$D_2(p_0, \dots, p_t) = D(d^{true}, p_0, \dots, p_t) \quad (15)$$

とする。

5.5 検出された逸脱行動と人の手によるラベルの比較

提案手法により逸脱行動を検出するために全ての軌跡について $P(D(p_0, \dots, p_t) > \rho)$ を求めた。このとき、複数の α を用いて ROC 曲線を作成した結果を図 5 に示す。 α は 10m, 25m, 50m, 100m, 200m, 400m を用いた。同様に比較手法 1 と比較手法 2 で ROC 曲線を作成した結果を図 6 と図 7、各 α について提案手法と比較手法 1, 2 の ROC 曲線を比較したものを図 8 14 に示す。また、ROC 曲線から計算された AUC を表 1 に示す。

表 1 から提案手法においては α が小さくなるほど AUC が大きくなっている。また、比較手法 1 の場合、 $\alpha = 10m$ のときの AUC は $\alpha = 25m$ のときの AUC と近く、 $\alpha = 25m$ 以上の時は α が小さくなるほど AUC が大きくなっている。一方、比較手法 2 では $\alpha = 400m$ のときを除き、大きな差がない結果となった。

表 1 α 毎の AUC

α	10m	25m	50m	100m	200m	400m
提案手法	0.753	0.753	0.716	0.690	0.628	0.568
比較手法 1	0.710	0.714	0.687	0.661	0.602	0.522
比較手法 2	0.557	0.549	0.557	0.558	0.544	0.602

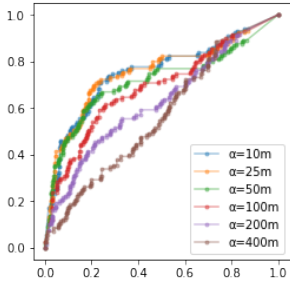


図 5 提案手法の ROC 曲線

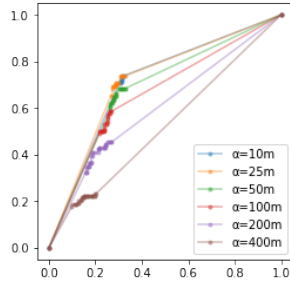


図 6 比較手法 1 の ROC 曲線

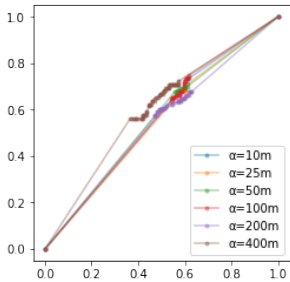


図 7 比較手法 2 の ROC 曲線

5.6 検出結果に関する考察

5.6.1 α の値による検出結果の違い

ある実験参加者の GPS 軌跡とその周辺地図を図 15 に示す。この参加者は図右下の青い菱形の地点から出発し、図上部の黄色の菱形の目的地へ向かった。人間によるラベル付けではこの参加者の GPS 軌跡は 8, 9, 16, 17, 18, 19 番目の区間が逸脱区間である。図 15 において GPS 軌跡を 20 分割し、逸脱区間を赤い線、そうでない区間を灰色の線で示している。また、こ

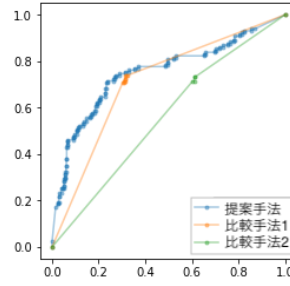


図 8 $\alpha = 10m$

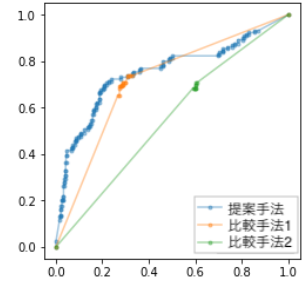


図 9 $\alpha = 25m$

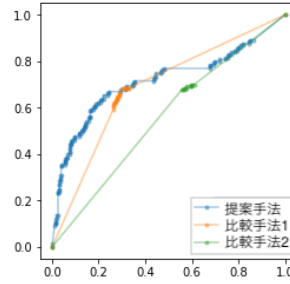


図 10 $\alpha = 50m$

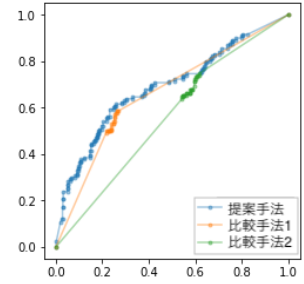


図 11 $\alpha = 100m$

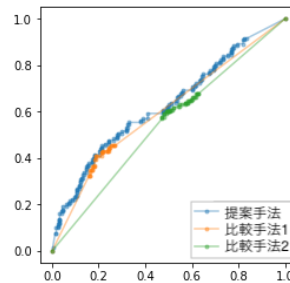


図 12 $\alpha = 200m$

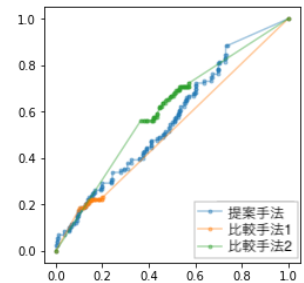


図 13 $\alpha = 400m$

図 14 提案手法と比較手法 1, 2 による ROC 曲線の比較

の軌跡を用いて複数の α で逸脱度を求めた結果を図 16 に示す。

$\alpha = 10m$ の時は 9, 10, 11, 16, 17, 18 番目の区間の逸脱度が 0.9 以上と非常に高い。一方、 $\alpha = 400m$ の時の逸脱度は軌跡全体を通して低く、最大でも 0.23 しかない。 $\alpha = 10m$ の時に検出できた逸脱行動が $\alpha = 400m$ の時に検出できなかった理由として 4.3 節で議論したように α が小さいほど小さな逸脱に対し逸脱度が大きくなりやすいことが考えられる。また、提案手法では目的地までの最短経路長が α 未満だった場合、逸脱度の計算が行えない。そのため、 $\alpha = 400mm$ の場合、16 から 19 番目の区間で逸脱行動が検出できていない。同じ理由により $\alpha = 100m$ 以下の場合でも 19 番目の区間の逸脱行動が検出できていない。

5.6.2 誤検出の原因

次に提案手法で検出できなかった逸脱区間について考察する。また逆に提案手法で検出されたが、ラベルがつかなかったケースの原因についても検討する。

まず、逸脱区間と検出できなかったケースを検討する。明らかな遠回りや U ターンを行ったもののすぐに最短経路に戻った場合が多く存在した。図 17 にその例を示す。図 17 右の丸で囲った部分や図 17 左の A 地点 \rightarrow B 地点 \rightarrow C 地点 \rightarrow B 地点 \rightarrow

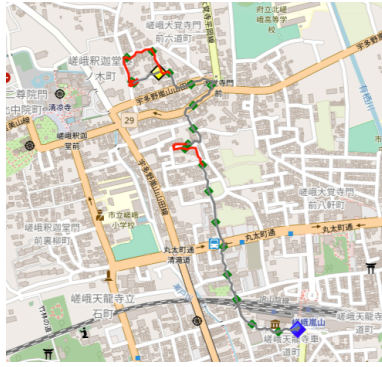


図 15 観光客の移動軌跡の例. GPS 軌跡を 20 分割したのち、逸脱行動である区間を赤色、そうでない区間を灰色で示した。

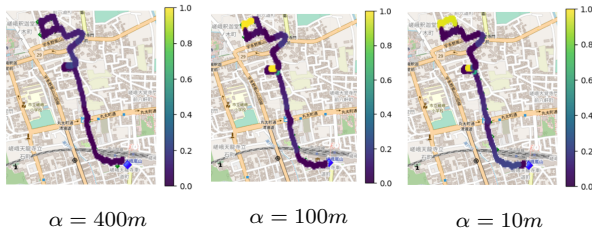


図 16 GPS 軌跡の各点における逸脱行動である確率

D 地点のように最短経路から外れるがすぐに U ターンし、最短経路に戻った場合が多く存在した。目的地と逆方向に進んだり U ターンを行なった場合、人の判断では明らかに逸脱と判断されやすい。しかし、本研究では移動距離を求める際、GPS の誤差の影響を除去するために交差点単位でマップマッチングを行っている。そのため、提案手法では明らかな遠回りを行なっても次の交差点に到着する前に引き返した場合、提案手法では検出することができない。これらの小さな観光客自身が気づいている逸脱行動を検出するためには、提案手法のような GPS を用いた客観的な手法ではなく、成本ら [10] [11] や高藤ら [12] のようなスマートフォン搭載センサを用いた観光客の行動意図の推定を行う手法がより適している。



図 17 FN の例. 逸脱区間を赤色、非逸脱区間を灰色で地図上に示した。U ターンや遠回りを行なっているが本研究では交差点マッチングを行っているため、これらの移動を考慮できていない。

次に、逸脱と検出されたが非逸脱区間とラベル付けされたケースについて考察する。このケースとしては、距離を計算すると遠回りを行なっているが地図を見てもそれが明らかにはわからない場合が多く見られた。観光客が地図を見て移動を行う

とき、大きな建物や線路の影響で方向は正しいのに遠回りな道を選択することがある。図 18 に例を示す。図 18 の右上の赤い丸で囲った交差点で左折を行った場合、目的地までの最短距離は 709m、直進した場合は 488m で約 1.5 倍の差があり、大きな遠回りとなる。しかし、この左折した区間に逸脱ラベルを付けた協力者は 45 人中 2 人だけである。そのような場合、GPS 軌跡とその周辺の地図を見てもラベル付けを行う人の目で遠回りだと判断することは難しい。また、遠回りを行なっているがその前後に目的地から遠ざかるようなより明確な逸脱行動がある場合、逸脱行動であると判断されづらい傾向がある。図 19 では目的地から遠ざかっているにもかかわらず、逸脱行動となっているのはその次の区間だけである。このように、人間の目では遠回りだと判断がつかないような場合でも、提案手法では遠回りであると判定できており、観光客の遠回りを防ぐという意味においては提案手法の有効性を示すことができている。

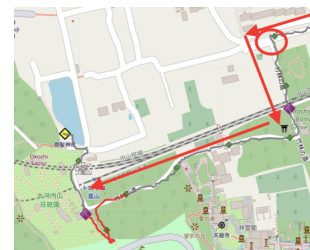


図 18 FP の例. 黄色の四角が目的地、紫の四角が逸脱行動が検出された地点である。観光客が図右上の赤い丸で囲った交差点で左折を行ったあと、逸脱行動が検出されている。



図 19 FP の例. 逸脱行動が検出された区間では目的地から遠ざかっており、遠回りを行なっているが、45 人中 7 人しか迷っているというラベル付けをしていない。



図 20 観光客が最初から方向を間違えた GPS 軌跡の例. 青い四角が出発地、黄色の四角が目的地、紫の四角が逸脱行動が検出された地点である。最初から目的地と逆向きに進んだため逸脱区間となっている。提案手法では移動開始から α 移動するまで逸脱度が定義できず、検出できていない。

提案手法は軌跡全体で目的地予測を行い、予測目的地との経



図 21 目的地周辺で迷った GPS 軌跡の例。黄色の四角が目的地である。目的地の目の前を通り過ぎており、逸脱行動とされている。提案手法では目的地に近すぎるため逸脱度が計算できず、検出できない。

路を用いて $p_{t-l(\alpha)}$ から p_t を用いて逸脱度の計算を行う。そのため、観光客が最初から間違えた方向に進んだ場合、目的地予測が適切に行えないため提案手法では正しく検出することができない。例として図 20 を示す。本実験では目的地候補は出発地の北方向から反時計回りに南西方向に分布している。そのため、出発直後に東方向に進んだ場合、どこが目的地であったとしても逸脱行動である。しかし、提案手法ではこの逸脱行動を検出できていない。また、現在地から目的地への最短経路長が α 未満の場合、提案手法では逸脱度の計算が行えないため、検出することができない。図 21 に例を示す。図 21 では観光客は目的地を通り過ぎており、人の手では逸脱行動と判定されているが、提案手法で逸脱行動は検出できていない。そのため、移動開始直前に支援を行うシステムや目的地周辺で支援を行う手法が必要である。

6 結 論

本研究では、陽に目的地を判定するのではなく、複数の目的地候補を推定した上でその推定の尤度を考慮して逸脱行動の検出を行う手法を提案した。また、目的地を明示的に判定する手法と比較することにより、提案手法の優位性を示した。

提案手法による逸脱行動の検出結果と人の手による逸脱行動のラベル付けの比較による評価を行った。比較手法として、予測目的地を単一に与える手法や真の目的地を事前に与える手法との比較を行った。その結果、提案手法を用いることでそれらの比較手法を用いるよりも検出精度が高くなることが分かった。また、提案手法で検出した逸脱行動と人の手による逸脱行動ラベルとが異なる場合の分類と考察を行った。

今後の課題として、本研究では最短経路を通るとしてその経路長を用いて逸脱度を算出しているが、観光客が実際に通るであろう経路を推定し、その経路長を用いることが考えられる。

謝 辞

本研究は JSPS 科研費 17K00438 の助成をうけたものです。

文 献

- [1] 総務省. 位置情報の利用に対する意識調査, 2014.
- [2] 梶田健史, 山守一徳, and 長谷川純一. デフォルメ地図自動生成システムの開発. 情報処理学会論文誌, 37(9):1736–1744, sep

- 1996.
- [3] Katashi Nagao. Agent augmented reality: A software agent meets the real world. In *Proceedings of the Second International Conference on Multi-Agent Systems (ICMAS'96)*, 1996.
- [4] J. Rekimoto, Y. Ayatsuka, and K. Hayashi. Augment-able reality: situated communication through physical and digital spaces. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, pages 68–75, Oct 1998.
- [5] Andy Yuan Xue, Rui Zhang, Yu Zheng, Xing Xie, Jin Huang, and Zhenghua Xu. Destination prediction by sub-trajectory synthesis and privacy protection against such prediction. In *2013 IEEE 29th international conference on data engineering (ICDE)*, pages 254–265. IEEE, 2013.
- [6] Andy Yuan Xue, Jianzhong Qi, Xing Xie, Rui Zhang, Jin Huang, and Yuan Li. Solving the data sparsity problem in destination prediction. *The VLDB Journal—The International Journal on Very Large Data Bases*, 24(2):219–243, 2015.
- [7] John Krumm and Eric Horvitz. Predestination: Inferring destinations from partial trajectories. In *International Conference on Ubiquitous Computing*, pages 243–260. Springer, 2006.
- [8] Daniel Ashbrook and Thad Starner. Using gps to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous computing*, 7(5):275–286, 2003.
- [9] Philippe C Besse, Brendan Guillouet, Jean-Michel Loubes, and François Royer. Destination prediction by trajectory distribution-based model. *IEEE Transactions on Intelligent Transportation Systems*, 19(8):2470–2481, 2017.
- [10] 成元 椋祐, 梶田宗吾, 山口弘純, 東野輝夫, et al. 認知マップの誤りによる迷い行動のスマートフォンを用いた検出手法. マルチメディア, 分散協調とモバイルシンポジウム 2017 論文集, 2017:1386–1395, 2017.
- [11] 成元 椋祐, 梶田宗吾, 山口弘純, 東野輝夫, et al. スマートフォンを用いた興味行動のセンシング. 研究報告ユビキタスコンピューティングシステム (UBI), 2018(49):1–7, 2018.
- [12] 高藤巧, 樋口雄大, 山口弘純, 東野輝夫, et al. スマートフォン利用者の迷い行動の検出に関する一検討. 研究報告ユビキタスコンピューティングシステム (UBI), 2016(1):1–6, 2016.
- [13] Reginald G Golledge et al. *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU press, 1999.
- [14] 瀧本祥章, 西田京介, 遠藤結城, 戸田浩之, 澤田宏, and 石川佳治. 時間帯を考慮したパーソナライズ目的地予測. 電子情報通信学会論文誌 D, 100(4):472–484, 2017.
- [15] Brian D Ziebart, Andrew L Maas, Anind K Dey, and J Andrew Bagnell. Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 322–331. ACM, 2008.
- [16] 笠原秀一, 森幹彦, 椋木雅之, 美濃導彦, et al. 遷移ネットワークを用いた大規模観光地の旅行者行動分析. *SIG-SPATIAL*, 17(3):1–8, 2013.
- [17] Christopher E White, David Bernstein, and Alain L Kornhauser. Some map matching algorithms for personal navigation assistants. *Transportation research part c: emerging technologies*, 8(1-6):91–108, 2000.
- [18] Sotiris Brakatsoulas, Dieter Pfoser, Randall Salas, and Carola Wenk. On map-matching vehicle tracking data. In *Proceedings of the 31st international conference on Very large data bases*, pages 853–864, 2005.