

対象者の状態を加味した
配列アライメントによる時空間行動分析手法の提案と分析
- 歩行速度と観覧対象との距離を活用した観覧行動類型化手法の提案 -
川瀬純也・伊藤史子

**Sequence Alignment Analysis for GPS logs
by Using Spatio-temporal Data of tourists' Activity**
Junya KAWASE and Fumiko Ito

Abstract: There are some previous studies applied sequence alignment on tourists' activities, which are suitable for typologies of overall view of tourists' activities as a whole in time and space. In this paper, we tried to expand the analysis by using tourists' walking speed and distance to the nearest exhibition as additional factors. As a result, we identified of the typologies characterized by time course of their activities.

Keywords: 配列アライメント手法 (sequence alignment method), 時空間行動 (spatio-temporal activity), GPS (global positioning system)

1. はじめに

近年, GPS などのセンサ類が搭載されたスマートフォンの普及により, 様々な分野・場面でこれを活用し, 人々の行動を時間と空間の両側面で把握し, 分析しようという試みがなされている。人々の時空間行動を定量的に把握して類型化を行う手法として, 文字列間の類似度を定量的に算出する配列アライメント手法を用いた試みがあるが, 時空間行動を文字列に置き換える時間解像度問題や, 配列アライメントの計算に用いる各種のパラメータ設定を的確に行う手法が確立されていないなど, 課題は多い。本研究においては, 筆者らがこれまで明らかとしてきた手法の課題点を踏まえ, 歩行移動者の時空間行動を対象とし

て, 対象者の状態を加味した文字列変換方法を検討・提案し, その手法による分析を試行する。

2. 配列アライメントによる時空間行動分析

配列アライメント (Sequence Alignment Method: SAM) とは, ある異なる文字列を同一の文字列にするために必要な編集操作 (挿入・削除・置換) の最小回数として定義される Levenshtein 距離の概念をもとに, 異なる文字列の違いを定量的に算出する手法である。Wilson et al. (1999) は, 時間地理学における日誌調査の分析に SAM を用いるため, 社会的な行動 (Sleep, Work, Shopping など) にそれぞれ文字を定め, 被験者の一日の行動を単位時間ごとに文字に置き換えることで「行動配列」と呼ぶ文字列を作り類似度を算出するアプリケーション”ClustalG”を提案している。また, Shoval and Isaacson (2007) は, 観光地を複数のエリアに分割して文字を定め, 観光者の GPS デー

川瀬純也 〒192-0397 東京都八王子市南大沢 1-1

首都大学東京大学院

都市環境科学研究科都市システム科学域

E-mail: kawase-junya1@ed.tmu.ac.jp

タから得られる観光地内での遷移を文字列に置き換えることで、時空間行動の類型化を行う手法を提案し、イスラエル・アッコを訪れた観光者の観光ルートの類型化を試みている。Wilson (2008) は、ClustalG の改良版として、それぞれの行動を行った地点の違いを「空間配列」と呼ぶ文字で表して行動配列に併記し、同じ行動でも地点が異なれば、類似度を相対的に低く見積もるアルゴリズムを採用した”ClustalTXY”を提案している。地点を表す文字列に座標を紐づけする際に、計算の負荷を軽減するために 5×5 のグリッドに置き換えてその代表点間のユークリッド距離で重みづけを行っている。また、矢部 (2010) は、多摩動物公園の来園者を対象に、園内を複数のエリアに分けて滞在エリアの遷移を表す文字列を用いるとともに、ClustalTXY における地点間の違いを重みづけする「空間配列」に対して、時間変化を表現するように値を入力して重みづけを行うことで、より明瞭な類型化が可能であることを明らかにしている。

3. 対象者の状態を加味した分析手法の提案

3.1 既存手法の課題点と提案手法の概略

対象空間を分割して文字を定め、対象者の移動をその文字列によって表現する場合、類型化の結果は空間の分割の仕方に大きく依存することとなる。これにより、例えば、ある時間だけそのエリア内に滞在していたとしても、そのエリア内を移動し続けていたのか、一か所で静止していたのかでは人々の時空間行動を理解する上で大きく意味が異なると考えられるが、それらは一種類の文字の連続として同様に扱われてしまうという問題が生じうる。ClustalTXY では空間的な違いによる重みづけが提案されているが、ある施設内の歩行移動など、移動距離が短く、さらに空間の接続関係が複雑な場合では、地点間のユークリッド距離による重みづけでは不十分であることが考えられる。また、GPS データなどから様々な具

体的な行動を推定することは難しく、日誌調査の分析を主たる目的とした ClustalTXY をそのまま歩行移動者の時空間行動分析に用いるには限界があると考えられる。

そこで、本研究では、既存の SAM による時空間行動の類型化手法と同様に、対象となる空間を複数のエリアに分割して文字を定め、対象者の空間的な遷移を表す文字列を用いるとともに、ClustalTXY で提案された重みづけグリッドとアルゴリズムを活用して、対象者の行動の状態を表して重みづけする手法の提案と分析を行う。筆者らがこれまで行ってきた SAM を用いた時空間行動分析 (川瀬・伊藤 2015) と同様に、2013 年 6 月 30 日 (日) 上野動物園において、一般来園者を対象に実施した GPS ロガー配布調査のデータを使用し、これまでの研究成果と新たな提案手法による分析結果との比較を行う。

動物園における観覧行動という特性を考慮した対象者の行動状態を表す指標として、対象者の「歩行速度」と「展示との距離」が挙げられる。川瀬ほか (2016) は、動物園来園者の観覧行動は、GPS ログから得られる歩行速度や展示との距離などを説明変数として用いることで、動物展示を観覧していたか否かが高い確率で推定可能であることを明らかにしている。そこで、本研究においても各対象者の単位時間窓内における「歩行速度」と「展示との距離」によって観覧状態を示す指標を作成し、空間遷移を表す文字列への重みづけを行うこととした。

3.2 被験者の状態を加味した文字列の作成

ClustalTXY では、配列 a と b の「行動配列」の類似度を $d(a_i, b_j)$ 、「空間配列」の類似度を $E(a_i, b_j)$ として、それぞれにウェイト u と v (ただし、 $u + v = 1$) をかけあわせることで、二つの配列要素の類似度 $q(a_i, b_j) = u * d(a_i, b_j) + v * E(a_i, b_j)$ として類似性を求める。

2013 年 6 月 30 日 (日) に上野動物園において、一般来園者を対象に実施した GPS ロガー配布調

査で回収したデータのうち、表門から入園し、東園から西園へと移動し退園した 46 組、東園から西園へ移動し再び東園へと戻った 54 組の GPS データを用いた。各施設の機能と空間的な接続とを考慮して、園内を 30 のエリアに分割してそれぞれ文字を割り当てた（図-1）。1 秒間隔で記録した GPS ログのうち、測位精度に問題があると判断したログを除外したのち、園内図上にプロットし、滞在エリアを表す文字を附加した。欠損したログは直前に有効であったログの滞在エリアに滞在していたとして扱うこととした。単位時間をこれまでの研究と同様に 1 分に設定し、1 分間隔ごとに GPS ログを抽出し、1 分間のうち最も割合の大きかったエリアをその時間窓内の滞在エリアとして、行動配列を作成した。

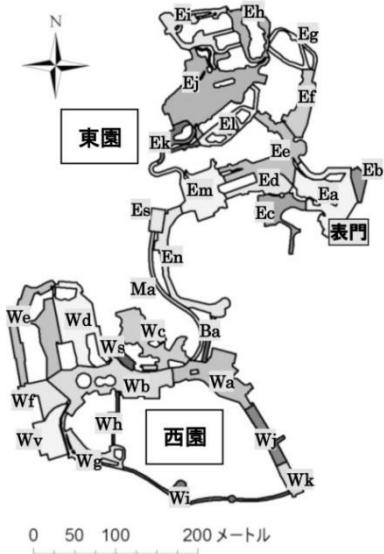


図-1 30 エリアに分割した上野動物園 園内図

次に、GPS ログを 1 分間隔ごとに抽出し、歩行速度と展示との距離の値をプロットして、それぞれカーネル密度推定により最高密度点を決定した。この最高密度点を用いて、歩行速度と動物展示の距離とをそれぞれ軸に取り、 5×5 の重み付けグリッドを作成する。今回は、外れ値を除外した後、2 軸とも均等な幅で区切ることとした。箱ひげ図などで用いられる、第 3 四分位点に IQR (interquartile range) の 1.5 倍の値を加えた値を上

限値とする手法 (Dekking, 2005) を用いて、上限値以上の値を外れ値とした。これを用いて、「空間配列」に代用するための観覧者の状態を加味する配列を作成した。

以上、二種類の配列を用いて、「東園→西園」と移動した 46 組、「東園→西園→東園」と移動した 54 組でそれぞれ配列アライメントによる樹上図を得て、類型化を行い、これまでの研究結果と比較した。なお、二種類の配列の類似度を足し合わせるウェイトは、 $u = 0.9$, $v = 0.1$ とした場合がこれまでの研究結果と最も比較しやすい結果を得ることができたため、その結果を用いる。

4. 結果と考察

それぞれの被験者群を対象として SAM を用いて得られた樹状図を作成し、各クラスターとしての特徴の解釈を行った。

まず、川瀬・伊藤（2015）の分析結果においては、園内のある箇所での長時間の滞在を特徴としてもつクラスターが複数得られていた。しかし、本研究で得られた樹状図では、それら長時間の滞在として特徴付けられたクラスターに属していたサンプルが各々異なるクラスターに属するという結果を得ることができた。これにより、同じエリアでの長時間の滞在でも、異なる時空間行動の特徴として扱われているということがうかがえる。

次に、各クラスターの GPS データから時空間パスを作成し、各クラスターの時空間行動の特徴を考察した。この時空間パスでは、表門での GPS 測位開始時刻から 1 分経過するごとに Z 軸方向に 1m だけ遷移するように設定し、平面園内図上に図示している。ここでは、川瀬・伊藤（2015）での結果とは異なり、特徴的な観覧ルートのまとまりをもつクラスターを確認することはできなかった。クラスターを作成する際に、空間的な遷移による類似度の影響が低く見積もられた結果であると考えられる。クラスター a の時空間パス（図

- 2) からは、被験者が園を訪れてから同じ程度の滞在時間が経過した後に東園から西園へと移動している様子がうかがえる。東園から西園への移動という空間的な遷移の類似に加え、多くの被験者が一定以上の歩行速度で移動していたという類似が反映された結果であると考察する。また、クラスターdとクラスターe(図-3)を比較すると、東園から西園へと空間的に遷移していく際のZ軸方向の傾きが大きく異なることがわかる。総滞在時間の影響もあるが、歩行速度の類似の影響が現れた類型の特徴であるとも考えられる。また、図中、時空間パスのうち赤い線(クラスターdはオレンジ)で描画された部分は、歩行速度・動物との距離がともに最小である文字が与えられた区間である。この赤い線がZ軸方向に縦に描画される区間は、長時間にわたり動物展示付近で観覧行動を行っていたと想定される区間であり、それぞれのクラスターを特徴付けている。

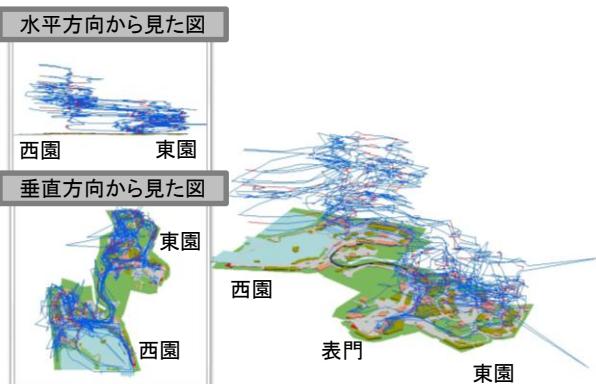


図-2 クラスターaの時空間パス

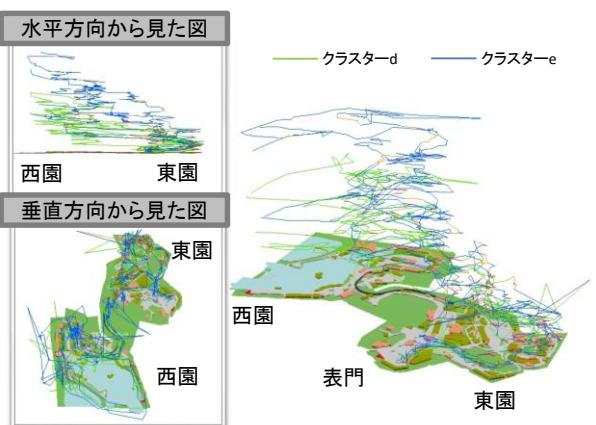


図-3 クラスターd及びeの時空間パス

5.まとめ

本研究では、既存のSAMによる時空間行動類型化手法の課題点を踏まえ、歩行速度と観覧対象物との距離を用いて対象者の状態を加味した文字列を作成し、空間的遷移の類似度に対する重みづけを行った。結果として、空間的な遷移の類似度により作成された類型よりも、歩行速度とそれによる時間的な変化が加味された類型を得ることができた。今後は、重みづけグリッドの作成方法を再検討するとともに、空間的な遷移と対象者の状態を反映した類型を得るための検討を進めていく。

参考文献

- 川瀬純也・伊藤史子(2015)：配列解析による上野動物園来園者の時空間行動類型化～閉じた空間内での回遊行動に着目して～, 地理情報システム学会講演論文集, 24, CD-ROM.
- 川瀬純也・伊藤史子・倉田陽平(2016)：時空間行動を表す複数の説明変数を用いた観覧行動推定 - 多摩動物公園におけるGPS調査実験をもとに-, GIS - 理論と応用, 24 (1), 1-11.
- 矢部直人(2010)：GPSデータに対する配列解析の援用, 地理情報システム学会講演論文集, 19, CD-ROM.
- Dekking, F. M., 2005. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media.
- Shoval, N., and Isaacson, M., 2007. *Sequence alignment as a method for human activity analysis in space and time*. *Annals of the Association of American geographers*, 97(2), 282-297.
- Wilson, C., 2008. *Activity patterns in space and time: calculating representative Hagerstrand trajectories*. *Transportation*, 35(4), 485-499.
- Wilson, C., Harvey, A., & Thompson, J., 1999. *ClustalG: Software for analysis of activities and sequential events*. *IATUR Conference Proceedings*.