

長期時系列におけるアノマリー分布の トラスビジュアリゼーションのインタラクティブ最適化

樋口 稔^{†,a} 熊野 雅仁^{†,b} 木村 昌弘^{†,c}

† 龍谷大学理工学部電子情報学科

a) *t130122@mail.ryukoku.ac.jp* b) *kumano@rins.ryukoku.ac.jp* c) *kimura@rins.ryukoku.ac.jp*

概要 近年, WEB 空間には人間生活に関連した膨大な時系列データが蓄積されている. 本研究では, 数年間の社会経済データや気象データのような人間生活と密接に関係する長期時系列において, 各日のアノマリーがどのように分布しているかをグローバルに分析するために, 週と曜日の二重周期性に着目したトラスビジュアリゼーション法を与える. さらに分析者の意向に応じて, アノマリー分布のトラスビジュアリゼーションを平面上にわかりやすく示すために, ビジュアリゼーションの視点を最適化するインタラクティブな手法を提案する. 気象観測データを用いた実験を行い, 提案法の有効性を検証する.

キーワード アノマリー分布, トラスビジュアリゼーション, 探索的データ解析, インタラクティブ最適化

1 はじめに

近年, WEB 空間に蓄積されるビッグデータの有益な利活用が模索されている. 我々は, ビッグデータに含まれる社会経済データや気象データのような人間生活と密接に関係する長期時系列データを可視化して, 視覚的および探索的にデータ解析を行うためのビジュアリゼーション手法に着目している. 本研究では図 1 のように, 3 次元トラス表面へ螺旋状に時系列情報を這わせた HGT (Helix glyph on torus) に異常性を可視化し, アノマリー分析を行う手法の研究を進めている [1].

HGT は, 図 1 のようにトラス表面の小円側に曜日を与えれば一次元の時系列データが 7 日で小円を一周し, 螺旋に沿って約 52 周することで年 (大円) を表現できる. このとき, HGT 上, 12 月末と 1 月初頭がつながることから数年間の傾向を HGT 上に表せば, 時系列の連続性を保持した数年の傾向を大円で視認することで人間生活に関係する曜日と週という二重周期性の観点からアノマリー分析が可能となる. ただし HGT には, 視点に対して裏側が視認しにくいというオクルージョンの問題がある. ビジュアリゼーションの研究では, この問題に対し, 視点選択 [2] や対話的に注視領域を透過処理する方法 [3] など, 様々な問題解決法が提案されている.

本研究では, トラスの小円・大円方向の任意の線に沿ってカットすることにより, 図 2 のように HGT をトラスと同相の長方形 (縦軸: 週周期 $i \equiv i + 52, i = 1, \dots, 52$, 横軸: 曜日周期 $j \equiv j + 7, j = 1, \dots, 7$) に平面展開し, HGT と長方形を相互に活用しながら対話的に分析者の意向に応じ最適な可視化を行うインタラクティブ最適化の方法を提案する. 人間生活と関連がある気象データを長期時系列とみなした実験で提案法の有効性を検証する.

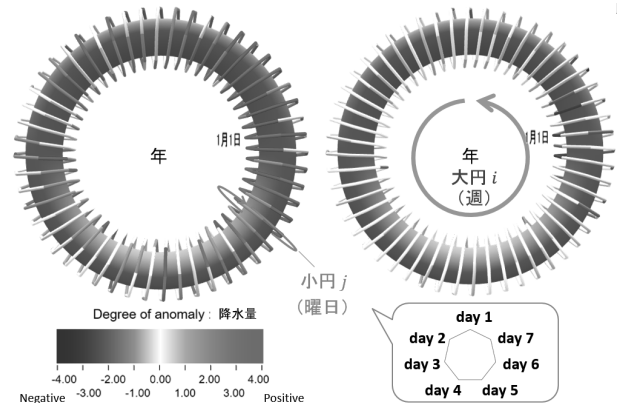
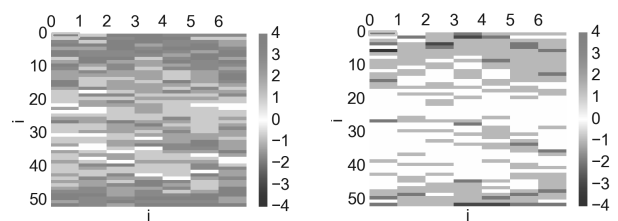


図 1: 二重周期性を考慮した HGT による異常可視化



(a) 基準 1 月 1 日 ('88-'90) (b) 基準 1 月 1 日 ('12-'14)

図 2: 図 1 の HGT を長方形に平面展開した例 (降水量の異常) .

2 長期時系列データのアノマリー分布のトラスビジュアリゼーション

2.1 長期時系列データのアノマリー

長期時系列データの分析とは, 連続する y 年間の気象データにおける平均的な 1 年間の挙動の分析のことを意味する. 各日の異常度を, 30 年間 (平年) のデータにおけるその日の平均と分散に基づいて定義する [1].

2.2 トーラスビジュアライゼーション

本研究では地域ごとに、図1のように、小円を曜日 $j \equiv j + 7, (j=1, \dots, 7)$, 大円を週 $i \equiv i + 52, (i=1, \dots, 52)$ と対応させ、各日の正の異常度に赤色、負の異常度に青色を割り当て、HGTとしてトーラス上に可視化する。

3 提案インタラクティブ最適化法

1月1日から始めて d 日目、($d=1, \dots, 364$)¹ を基準日として、HGTを図3のように展開する。ここに、 $A_{i,j}(d)$ は、第 $date(d, i, j) = \{7(i-1) + (j-1) + d\}$ 日目の異常度 [1] である。ただし、 $date(d, i, j) - 364 \equiv date(d, i, j)$ である。

ところで、この長方形はオクルージョンの問題が解消されるものの、カットによる非連続化により視認性に関する別の問題が生じる。図2は、 $d=1$ 日目(1月1日)を始点として小円・大円に沿ってカットした長方形の例であるが、図2aと図2bいずれも1月頃と12月頃に異常度の高い日が集中しており、HGTでは連続していた時期が平面展開すると分断されるときがある。これは分析者の着眼点により視認性が悪い状態となり得るが、 d を変更すれば視認性が高まる可能性がある。ただし、 d の変更による平面展開の仕方は364個存在し、多数の地域で可視化する各HGTの d を手動で決めることは非効率である。本研究では、用意した基準群から分析者の意向に応じて選択された基準の組み合わせについて基準日 d に基づいた最適化を行い、平面展開の候補から厳選された結果をユーザに提示することができるインタラクティブ最適化法を提案する。ここでは紙面の都合上、最もナイーブな二つの基準を取り上げる。

基準 1) 異常の少ない部分を平面の端に寄せるため、目的関数 $f_1(d) = |\bar{A}_1(d)| + |\bar{A}_{52}(d)|$ を d に基づき最小化。ただし、第 i 週の平均異常度 $\bar{A}_i(d) = 1/7 \sum_{j=1}^7 A_{i,j}(d)$ である。

基準 2) 異常の差が大きい部分を中心に寄せるため、 $\bar{A}_i(d)$ を用いて目的関数 $f_2(d) = |\bar{A}_c(d) - \bar{A}_{c+1}(d)|$ を d に基づき最大化。ただし $c=26$ 。

4 実験

4.1 実験データ

気象データ AMeDAS の一次細分区域 (142ヶ所) について1985年から2014年までのデータを扱うが、ここでは $y=3$ として降水量に関する結果のみを取り上げる。

4.2 実験結果

提案法の有効性を検証する上で、第 i 週に対して基準1の次に基準2、さらに第 j 曜日に対して基準1の次に

¹1年を364日で表現すれば平面展開図の列が同じ曜日に対応する。この特性を有効活用するため、閏年は12月29日以降、平年は12月30日以降の気象データを平均して統合し364日を1年とした。

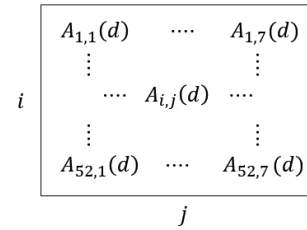


図3: HGTの平面展開における異常度 $A_{i,j}(d)$ の配置

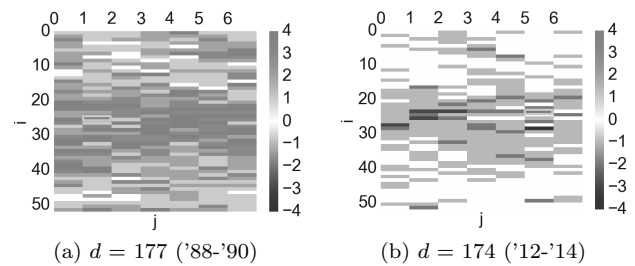


図4: インタラクティブ最適化の例 (鹿児島県奄美地方)。

基準2を適用した結果の一例として、図2aの最適化を図4a、図2bの最適化を図4bに示す。尚、図4aと図4b緑色の枠が1月1日の位置である。最適化の前(図2)に対し、図4aと図4bともに異常度の高い日々が中央に寄ることで12月と1月の異常が続く日々がかたまりとして視認性が高まったことがわかる。この観点において、提案インタラクティブ最適化法の有効性が示唆された。

5 まとめ

本研究では、長期時系列におけるアノマリー分布のトーラスビジュアライゼーションにおいて、オクルージョンの問題の解消法として分析者の意向に応じて選択された基準の組み合わせについて d に基づき HGT を最適に平面展開する手法を提案し、異常度の高い時期を長方形の中央に寄せる効果において有効性を示した。基準や複数の基準を組み合わせる方法は数多く存在し得るため、分析者の意向に沿う基準を新たに構築すれば、多様な最適化の結果を通じた分析が可能になる。今後は、多様かつ多種の時系列データを同時に分析可能なインタラクティブ最適化法および新基準の構築を目指す予定である。

参考文献

- [1] 西塚, 熊野, 木村: 気象観測時系列データのアノマリー可視化, 画像電子学会誌, Vol.45, No.1, 235, pp.128-129 (2016)
- [2] R. Miyazaki, T. Itoh: An Occlusion-Reduced 3D Hierarchical Data Visualization Technique. 13th International Conference on Information Visualisation, IV, pp.38-43 (2009).
- [3] X. Tong, C. Li, H. Shen: GlyphLens: View-dependent Occlusion Management in the Interactive Glyph Visualization, IEEE Scientific Visualization Conference (SciVis'16), TVCG.2016.2599049.