

## ゲーリッツ不変量による結び目データ解析

松葉 達也<sup>1</sup>

<sup>1</sup> 東京工業大学

e-mail : matsuba.t.aa@titech.ac.jp

### 1 はじめに

本研究では、ゲーリッツ不変量を用いた結び目データ解析の手法を提案する。位相的データ解析 (TDA: Topological Data Analysis) では、点群データが有する穴の構造に着目し、位相的特性を捉えるためにパーシステントホモロジー群を用いるが、点群データが潜在的に有する結び目構造を抽出するものではなかった [1, 2]。そこで、本研究では、結び目理論で提案されている位相不変量のひとつであるゲーリッツ不変量を活用して、点群データが潜在的に持つ結び目構造を解析する手法を提案する。

### 2 提案手法

本節では、位相不変量のひとつであるゲーリッツ不変量を用いた点群データ解析の手法について述べる。本研究で用いるゲーリッツ不変量とは、白黒彩色された結び目図式の交差点に符号を与えることでゲーリッツ行列を構成し、その行列を元に計算する不変量のことである。

なお、入力するデータは、ユークリッド空間  $R^3$  における有限個の点の集合  $P = \{p_1, p_2, \dots, p_n\}$  で与えられる点群データとする。

入力された各点群の半径を時間発展させることでマルチスケールの Čech 複体を構成する手順までは、先行研究であるパーシステントホモロジーを用いた位相的データ解析と同様である。しかし、本研究で提案するゲーリッツ不変量を用いた結び目データ解析では、点群の半径を離散的に時間発展させた際の各フィルトレーションステップにおいて接続されているエッジから結び目図式を抽出し、ゲーリッツ不変量のステップごとの変化を出力することにある。なお、閉路内で交差点を持つ単体的複体が構成された場合にゲーリッツ不変量を定義するものとする [3]。

以下では、ゲーリッツ不変量について概説する。

#### 2.1 結び目図式の白黒彩色と符号

結び目図式の白黒彩色とは、結び目図式の各領域を、隣接する領域が異なる色（たとえば白と黒）で塗られるように色分けすることである。

結び目の交差点に対して、以下のように符号を与える：

- 1) 右上から左下にかけて結ばれているエッジが、左上から右下にかけて結ばれているエッジの上を通る場合、その交差点に  $+1$  の符号を与える。
- 2) 上記の逆の場合、その交差点に  $-1$  の符号を与える。

#### 2.2 ゲーリッツ行列

ゲーリッツ行列  $G$  は、結び目図式で黒塗りされた各交差点における符号和を行列表示したもので

ある。ただし、行列の対角成分は、符号和に-1 を乗じたものとする。

## 2.3 ゲーリッツ不変量

ゲーリッツ不変量は、ゲーリッツ行列の行列式の絶対値で与えられる位相不変量である。

## 3 数値実験

本研究で提案した手法と、従来のパーシステントホモロジーによる位相的データ解析の出力結果の相違を確認するために数値実験を行った。

結果として、 $P_1 = \{(0, 10), (5, 5), (10, 0), (5, -5), (0, -10), (-5, -5), (-10, 0), (-5, 5)\}$ ,  $P_2 = \{(0, 10), (5, 5), (10, 0), (5, -5), (0, -10), (-5, -5), (-10, 0), (-5, 5), (2, 10), (7, 5), (12, 0), (7, -5), (2, -10), (-3, -5), (-8, 0), (-3, 5)\}$  の 2 つの点群データについて、パーシステントホモロジーによる位相的データ解析では同一の結果を出力したのに対し、ゲーリッツ不変量を用いた結び目データ解析では異なる出力結果を返した。

## 4 まとめ

本研究では、ゲーリッツ不変量を用いた点群データ解析手法を提案し、具体例の計算から、パーシステントホモロジーを用いた位相的データ解析では判別できない点群データを判別できることを確認した。

一方で、出力結果の頑健性の保証や膨大な点群データに適用する際の計算コスト削減などが今後の課題である。

## 参考文献

- [1] H. Edelsbrunner, D. Letscher, and A. Zomorodian, “Topological persistence and simplification,” *Discrete & Computational Geometry*, vol. 28, no. 4, pp. 511-533, 2002.
- [2] R. Ghrist, “Barcodes: The persistent topology of data,” *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61-75, 2008.
- [3] K. Inoue and Y. Yaguchi, “Goeritz matrices and polynomial invariants of links,” *Journal of Knot Theory and Its Ramifications*, vol. 16, no. 9, pp. 1135-1150, 2007.

# TDA に基づく胚細胞膜画像のセグメンテーションおよびトラッキング技術

織田 遥向<sup>1</sup>, 市川 尚文<sup>2</sup>, 井元 佑介<sup>2</sup>

<sup>1</sup> 東京大学, <sup>2</sup> 京都大学

e-mail : haruhisa-oda0722@g.ecc.u-tokyo.ac.jp

## 1 はじめに

細胞画像のセグメンテーション・トラッキングは生命科学画像解析の重要なテーマである。特に胚細胞の時系列データでは位置や体積といった 3 次元情報取得のため、3 次元セグメンテーション・トラッキング技術が有用である。本発表ではパーシステントホモロジー (PH) に基づく胚細胞膜画像のセグメンテーション技術を導入し、簡単なトラッキング手法と組み合わせて胚細胞の時系列データ解析に応用する。主に 8 細胞期の解析結果を取り上げ、発展的内容として多細胞期の解析についても述べる。本稿で省略する PH を適応する部分以外の画像処理は講演で紹介する。

## 2 2 次元画像処理の流れ

2 次元画像においては膜部分が周囲より高輝度であるため、sublevel filtration PH の  $H_0$  で検出を行う。得られたバーを HomCloud[1] の逆解析機能で元の画像内の volume に対応させる。検出された volume 部分を塗りつぶせば細胞部分が 1、外部が 0 の 2 値マスク画像が作成でき、volume の輪郭を用いれば膜部分の強調ができる。8 細胞期では前者を用いて、3 次元画像解析の入力を作成する。

## 3 3 次元画像処理の流れ

2 次元の 2 値マスク画像がスライスごとに作成された状態から解析を始める。 $z$  方向の解像度が  $xy$  方向より粗い場合は同じスライスを繰り返すことで対応し、3 次元全体で distance transform 後に PH を計算する。類似手法として [2] があるが、これは細胞輪郭を入力とするため  $H_2$  に注目したのに対し、今回は中実のデータを用いるため  $H_0$  に注目する。検出したバーの birth position をマーカーとして watershed セグメンテーションを行うことで 3 次元セグメンテーションが完了する。

## 4 8 細胞期時系列データの結果

手法の評価のため 8 細胞期胚細胞膜画像の時系列データ (全体で 34 枚の 3 次元画像) を解析した。まず、今回のデータでは適切なパラメータセットが存在して 34 枚すべてで 8 つの細胞が検出できることを確認した。さらに実用性の評価として、3 次元の PH 計算後のバーの選択のためのパーシステンスの閾値の調整について議論を行い、各時点における生成点が  $-10$  未満で  $k$  番目に長いバーの長さを  $P_k(t)$  と書くとき、 $P_8(t)$  と  $P_9(t)$  の差が平均して細胞サイズの  $1/2$  程度確保されていることを確認した。パラメータ調節の妥当性を裏付ける結果である (図 1)。さらに最小重み最大マッチングに基づくセグメンテーション結果のトラッキングを行い、PlantSeg[3] を用いた先行研究 [4] と比較した結果、共通して中心部分の細胞の体積が他より小さいという一貫性のある結果が得られた。

## 5 発展的内容: 多細胞期画像解析

多細胞期の画像では 8 細胞期のような明確な 3 次元構造がなく、目視でも正確なセグメンテーションが難しいため、個別のセグメンテーションを正確に行うことを目指すのは困難である。しかしながら形状が複雑な多細胞系ではマニュアルでの解析に労力を要するため、細胞数や平均体積といった集

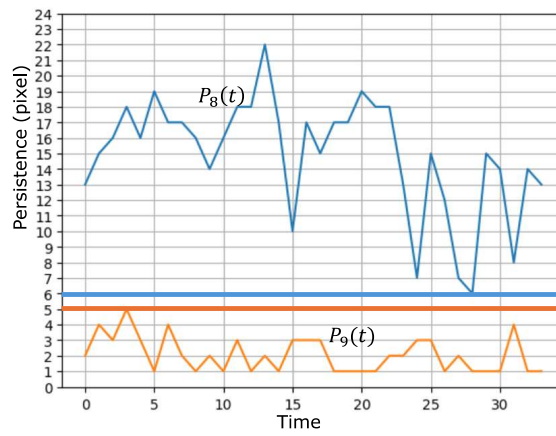


図 1. 生成点が  $-10$  未満の 8 番目および 9 番目に長いバーの長さ。今回のデータでは 34 枚すべてに対応できるパラメータセットが存在する。また 2 つのグラフの間には十分な差がある。

団の指標を計測することにも意義がある。今回導入した手法を改変して適応し、評価を行う。2 次元では検出された volume の輪郭を利用した膜の強調を行い、連結成分として細胞部分を検出する。3 次元ではまず、隣接するスライス間で重なった成分を利用してグラフを作成し、METIS[5] を用いた分割を行う。その後、分割が不十分な結果に対して 3 次元 PH を適用することで解析を完了する。結果として細胞数に関して先行研究 [6] とほぼ同様の結果を得たほか、平均体積についても同様の傾向が確認できた。

また、トラッキングに関して、8 細胞期と異なり、多細胞期では結果を目視で確認することが困難であるため、適切なトラック結果を抽出する方法が必要である。多細胞期には細胞体積が徐々に増加し、元の 2 倍程度になると分裂して半分になるという性質があるため、トラック結果の体積のグラフにこの性質が確認できることは良い指標になる。実際に先行研究のセグメンテーション結果に対し最小重み最大マッチングによるトラッキングを実施し、このパターンを確認した。

謝辞 8 細胞期の画像データは Dr. Dimitri Fabrèges からご提供いただいた。

## 参考文献

- [1] I. Obayashi, T. Nakamura and Y. Hiraoka, Persistent homology analysis for materials research and persistent homology software: HomCloud, Journal of the Physical Society of Japan, 91. (2022), 091013.
- [2] H. Oda, Persistent homological 3D cell detection, JSIAM Letters, 16. (2024), 25–28.
- [3] A. Wolny et al., Accurate and versatile 3D segmentation of plant tissues at cellular resolution, eLife, 9. (2020), e57613.
- [4] D. Fabrèges et al., Temporal variability and cell mechanics control robustness in mammalian embryogenesis, bioRxiv. (2023). <https://doi.org/10.1101/2023.01.24.525420>
- [5] G. Karypis and V. Kumar, A fast and highly quality multilevel scheme for partitioning irregular graphs, SIAM Journal on Scientific Computing, 20. (1999), 359–392.
- [6] T. Ichikawa et al., An *ex vivo* system to study cellular dynamics underlying mouse peri-implantation development. Developmental Cell, 57. (2022), 373–386.

## 位相的データ解析を用いた力学系の学習過程の評価

山田 泰輝<sup>1</sup>, 藤原 寛太郎<sup>1</sup>

<sup>1</sup> 東京大学大学院情報理工学系研究科

e-mail : yamada-taiki@g.ecc.u-tokyo.ac.jp

### 1 概要

パーシステントホモロジー (PH) は、位相空間の有限部分集合から元の位相空間のホモロジーを推定する手法であり、視覚的に判断できないデータの位相的性質を数値的に調査できる。本研究では、リカレントニューラルネットワークを用いて学習される高次元力学系に焦点を当て、力学系の漸近的振る舞いを理解する上で重要な不変集合の位相的性質を PH を用いて調査する。特に、学習に用いるデータ数の変化に伴って不変集合がどのように変化するかを分析し、力学系の学習の安定性について議論する。

### 2 背景

力学系は状態の発展規則を定め、様々な現象のダイナミクスを定式化する。観測データから力学系を学習することで、未知のダイナミクスの予測・制御に応用できる。特に、Echo State Network (ESN) を用いた手法 [1] は、その単純さから理論研究と応用研究の双方から注目されている。これまで、力学系の学習結果の収束性については研究されてきたが、有限データでの理論的保証は未解決の課題である [2]。ESN で学習した力学系は高次元であり、力学系の不変集合の全体像を視覚的に把握しにくいという課題をもつ。そこで本研究では、位相的データ解析を用いて高次元力学系の不変集合を定量化し、有限データでの ESN による学習結果を調査する。

位相的データ解析では、データがもつ形の情報を定量化する。パーシステントホモロジー (PH) [3] は、位相空間の有限部分集合から、元の位相空間における連結成分、穴、空洞といったホモロジーの情報を近似的に定量化する位相的データ解析手法の一つである。PH の出力結果をベクトル化することで、PCA による可視化や統計的処理に利用できる (図 1)。

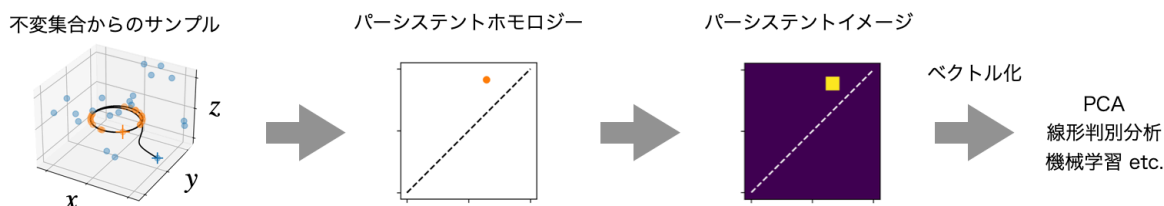


図 1. パーシステントホモロジーを用いた位相的データ解析の概要。

### 3 手法

真の力学系  $f_\circ : \mathbb{R}^{n_d} \rightarrow \mathbb{R}^{n_d}$  の軌道  $\{f_\circ^i(\mathbf{d})\}_{i \in \mathbb{N}}$  ( $\mathbf{d} \in \mathbb{R}^{n_d}$ ) を用いて、ESN の状態  $\mathbf{r} \in \mathbb{R}^{n_r}$  の発展規則を  $\mathbf{r}_{i+1} = \tanh[Af_\circ^i(\mathbf{d}) + B\mathbf{r}_i]$  で定めた。ここで、行列  $A \in \mathbb{R}^{n_r \times n_d}, B \in \mathbb{R}^{n_r \times n_r}$  は ESN のパラメータであり、各成分はそれぞれ区間  $[-1, 1], [-0.5, 0.5]$  上の一様分布から独立にサンプルした。 $\tanh$  は成分ごとに作用する。データを  $T$  個用いて ESN で学習した力学系  $f_T$  を  $f_T(\mathbf{r}) = \tanh[(AW_T + B)\mathbf{r}]$  で定めた。ここで、 $W_T = [f_\circ^{T_0}(\mathbf{d}) \dots f_\circ^{T_0+T-1}(\mathbf{d})][\mathbf{r}^{T_0} \dots \mathbf{r}^{T_0+T-1}]^+$

とした.  $X^+$  は行列  $X$  の擬似逆行列を表す.

力学系  $f_T$  の不変集合からの  $M$  個の点のサンプルの近似として,  $\mathcal{D}_s[f_T] = \{f_T^N(\mathbf{r}_i^{(s)})\}_{i=1}^M$  ( $N, M \in \mathbb{N}$ ) を用いた. ここで  $\mathbf{x}_i^{(s)} \in \mathbb{R}^d$  ( $i = 1, \dots, M$ ) は乱数シード  $s$  のもとで  $[-2, 2]^{n_r}$  上の一様分布から独立にサンプルした.  $\mathcal{D}_s[f_T]$  を入力として, 図 1 の流れでパーシステントホモロジーとパーシステントイメージを使用して算出したベクトルを  $\mathbf{v}(\mathcal{D}_s[f_T])$  で表す. 学習に用いるデータ数  $T$  と乱数シード  $s$  を変更して得られたベクトルの集合  $\{\mathbf{v}(\mathcal{D}_s[f_T])\}_{s,T}$  を用いて PCA を実行し, 各  $\mathbf{v}(\mathcal{D}_s[f_T])$  の第 1, 2 主成分をプロットした.

## 4 結果

ESN で学習した力学系  $f_T$  ( $T = 0, \dots, 9$ ) について, 定量化の結果と近似誤差との対応を図 2 に示す.

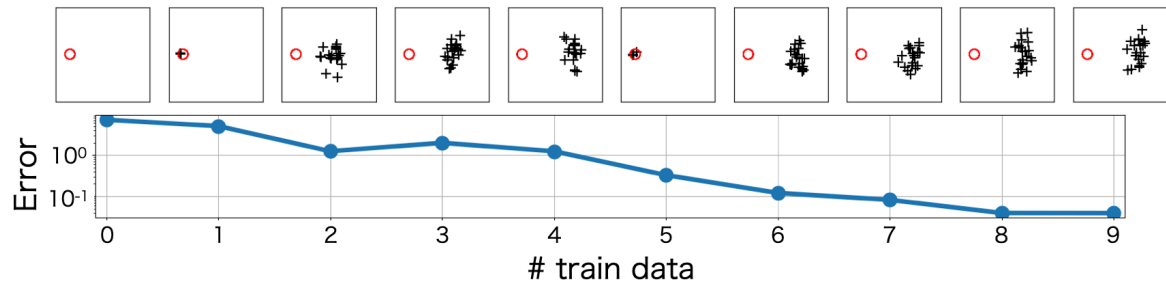


図 2.  $\{\mathbf{v}(\mathcal{D}_s[f_T])\}_{s,T}$  に対する PCA の結果と力学系の近似誤差との対応. 真の力学系  $f_0$  はリミットサイクルをもつ  $n_d = 2$  次元力学系とし,  $n_r = 5$  次元の ESN を用いた. また,  $N = 10^3$ ,  $M = 10^2$ ,  $s = 1, \dots, 10$  とした.

近似誤差は学習データの数の増加に伴って減少した (図 2-下). 一方, ESN で学習した力学系  $f_T$  の位相的性質 (図 2-上) は, 未学習の場合 ( $T = 0$ , 赤丸) から遠ざかってからその後一度戻った ( $T = 5$ ). この傾向は学習に利用した ESN のパラメータ  $A, B$  によらずに再現された. 学習した 5 次元力学系の不変集合の位相的性質を, ヒューリスティックな射影や縮約を行わずに評価している点に注意されたい. 本講演ではこの結果について, 力学系の学習の安定性の観点から考察する.

**謝辞** 本研究は, JSPS 科研費 JP20H00596, JP22K18419, JP24K15161, JST CREST JP-MJCR19K2, ムーンショット型研究開発事業 JPMJMS2021 の助成を受けた.

## 参考文献

- [1] Jaideep Pathak, Zhixin Lu, Brian R. Hunt, Michelle Girvan, and Edward Ott. Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(12):121102, December 2017. Number: 12 Publisher: American Institute of Physics.
- [2] Allen G. Hart, James L. Hook, and Jonathan H. P. Dawes. Echo State Networks trained by Tikhonov least squares are  $L_2(\mu)$  approximators of ergodic dynamical systems. *Physica D: Nonlinear Phenomena*, 421:132882, July 2021.
- [3] Vanessa Robins. Towards computing homology from finite approximations. In *Topology proceedings*, volume 24, pages 503–532, 1999. Issue: 1.