

平成 29 年 5 月 9 日現在

機関番号：11301

研究種目：挑戦的萌芽研究

研究期間：2014～2016

課題番号：26610042

研究課題名(和文) パーシステントポロジと逆問題：タンパク質の構造・機能解析における新手法の確立

研究課題名(英文) Inverse problem of persistent homology: A new method for structural analysis of proteins.

研究代表者

平岡 裕章 (Hiraoka, Yasuaki)

東北大学・原子分子材料科学高等研究機構・教授

研究者番号：10432709

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：本萌芽研究ではパーシステントホモロジーの逆問題解法を数学的に定式化し、実データへの適用が可能なソフトウェアの開発までを行った。まず順問題としてはポイントクラウドデータからパーシステント図を定める対応を可微分写像として定式化し、その局所構造を調べた。その後、陰関数定理を用いて逆対応を具体的に求めるアルゴリズムをニュートン法を用いて開発した。ここで開発された手法はタンパク質やガラスの構造解析に対して具体的に適用された。

研究成果の概要(英文)：In this research, we have formulated mathematical framework of inverse problems about persistent homology, and have developed a software which is applicable to real data. First, in direct problems, we studied the correspondence from point cloud data to persistence diagrams as differentiable maps. Then, we developed an algorithm to solve inverse solutions by using Newton method based on implicit function theorem. The developed software was applied to practical structural analysis of proteins and materials science.

研究分野：応用トポロジー

キーワード：パーシステント写像 特異点

1. 研究開始当初の背景

ソフトマター研究ではガラスやタンパク質といった個々の物質ごとに順問題と逆問題を解析することで構造-機能の相関関係を明らかにすることが求められている。ここで順問題(幾何→物性)では原子レベルの微視的幾何構造から巨視的な物性の特徴付けを行い、一方で逆問題(物性→幾何)では欲しい物性値を実現する幾何構造を決定し機能制御や新材料の開発を行う。申請者は本萌芽研究開始当初までに Protein Data Bank 内の原子配置情報から、幾何学的定量化としてパーシステントホモロジーを計算し、タンパク質圧縮率や H/D 交換率といった物性値を推定する順問題パートで成功をおさめた。しかしながら物質の機能制御や新材料の開発のためには、パーシステントホモロジーによって定量化された所望の物性値を実現する原子配置を逆問題として解く必要がある。

2. 研究の目的

本研究ではパーシステントホモロジーの逆問題に対して可微分構造を用いた解の追跡解法を提案し、さらにそのソフトウェア開発を行う。またその応用として、タンパク質構造解析や材料科学に現れる構造-機能相関に関する諸問題へ適用することを目指す。

具体的には以下の3つの課題を解決することが本研究の目的である。

課題1(順問題): 入力ポイントクラウドデータからパーシステントホモロジーを定める可微分写像の局所構造を解明する。

課題2(逆問題): パーシステントホモロジーの微小変形による入力データの追跡

課題3(ソフトウェア開発と応用): 追跡ソフトウェアの開発とタンパク質や材料科学に現れる構造-機能解析への応用。

現在パーシステントホモロジーは情報通信・材料科学・データ解析などの幅広い分野への応用が始まっている。ただしこれらの応用研究は全て順問題に属する内容であり、真に有効な応用数学手法として認められるのはその逆問題を解決する必要がある。可微分構造を用いた追跡理論を構築することで逆問題の解決を目指す本研究は、世界に先駆けた画期的な理論研究となる。

3. 研究の方法

課題1(順問題):

パーシステントホモロジーは加群としての単因子から一意に決定できる。そこで一般にユークリッド空間の有限点列から単因子を定める順問題の可微分写像としての構造を調べる。特に単因子の距離空間としての構造から有限点列の配置の局所対称性と微分の階数の関係を調べ、陰関数定理に基づく局所変形理論の構築をすすめる。

課題2(逆問題):

課題1で調べた可微分写像の局所的性質を用いて、解の追跡問題として逆問題を考察す

る。微分は一般に正方行列でないことに注意し、写像の正則点では一般化逆行列を用いたニュートン作用素の解の存在と一意性を明らかにする。また写像の臨界点では特異値と特異ベクトルを調べることで位置情報に対する分岐構造をすべて決定する。

課題3(ソフトウェア開発と応用):

課題2の結果を用いて、逆問題追跡ソフトウェアの開発を行う。アルゴリズムは一般化逆行列を用いたニュートン法と擬弧長法追跡により構成する。また本課題の成果をタンパク質やガラスといった材料構造解析へ応用する。

4. 研究成果

ユークリッド空間内の有限点列として与えられるポイントクラウドデータに対して、アルファ複体およびヴィートリスリップス複体を用いたフィルトレーションモデルに対して、そこから定まるパーシステント写像の well-defined 性とその可微分構造を明らかにした。ここでパーシステント写像はポイントクラウドデータからパーシステントホモロジーの単因子をあたえる対応から構成される。その意味でパーシステント写像の well-defined 性は単因子の一意性から示すことができる。一方でパーシステント写像の微分についての一意性は一般には成立せず、構成アルゴリズムの選択に微分写像は依存することになる。しかしながら、入力ポイントクラウドデータに適当な「一般の位置」の仮定を置くことで、微分写像の一意性についても示すことができた。また比較的単純な点列配置を用いて写像の特異点と一般化ニュートン法の非収束点の関係についても解明した。例えば3次元空間内の4点からなる例題では正四面体構造がパーシステント写像の境界を与え、そこから四面体を歪ませる方向に追跡が可能であることが確かめられた。その際に、徐々に対称性を落としながら歪んでいくことになるが、それぞれの対称性が壊れていく段階で、ニュートン法の収束が段階的に改善されることも確かめられた。さらに相似変形に対応する変形はパーシステント図内では対称性を壊す方向と横断的に交わる直線として与えられることも明らかにした。

さらにここで提案した追跡法をいくつかの点で改善を試みた。まず理論的側面としては、パーシステント図の生成元の個数が変化するより一般的な状況で理論的枠組みを構成した。アイデアはターゲット空間を無限次元ベクトル空間とし、そこへのパーシステント図の埋め込みを考えることで生成元の変化に対応させる写像の構成を行なった。ここでターゲット空間としては単射性が保証されているカーネル埋め込みを主に考察し、特に Persistence Weighted Gaussian Kernel と呼ばれる重み付きガウス型カーネルを用いた。このカーネルは別プロジェクトで位相

的統計理論の構築を目指して開発されたものであるが、データの摂動に対するリブシツツ安定性などの良い性質を有している。このような理論的な枠組みのもと、パーシステント図の微小変形が誘導するポイントクラウドの変化をヒルベルト空間上の勾配流として扱い、そこから定まる極値問題として追跡法を定式化しなおした。これにより既存の極値問題解法を本手法に適用することで、最適化理論ですでに開発されている高速解法が使用可能となった。

また最適化法との接点が見出されたことから、パーシステント図に対する LASSO 解析の可能性についても研究をおこなった。LASSO は近年スパースモデリングの分野で注目されている L1 最適化法であり、得られる最適解が疎構造をもつ。応用上は疎構造をもつことで現象の理解にもっとも影響を与える因子の特徴づけを可能とし、このことから現在幅広い分野で使われている。ここではパーシステント図を説明変数とするデータに対して線形回帰問題を考え LASSO を適用する理論枠組とアルゴリズムの構築を検討した。この際、パーシステント図のベクトル化にはパーシステントイメージとよばれる手法を用いた。これはパーシステント図を平面上の適当なグリッド分割上で定められるヒストグラムとみなし、各グリッド上での値を成分値とするベクトルである。これによりデータ解析に実用可能な LASSO を用いたスパース逆問題解法のプロトタイプを提案することに成功した。

ソフトウェアについては、萌芽研究とは独立に開発を進めている位相的データ解析ソフトウェア「HomCloud」に組み込むモジュールとしての開発を実施した。パーシステント写像の構成は PHAT および DIPHA をもちいた高速計算を実現することに成功した。その際、微分写像に相当する情報を計算アルゴリズムから抽出する必要があるが、この点についても問題なく対応できた。

実データへの応用としてはタンパク質構造解析およびガラスを含むソフトマターへの解析を実施した。特にシリカガラスの構造解析では、階層的なリング構造のいくつかの代表的な生成元に対して本研究課題で提案された追跡法を適用し、ガラスの脆性と原子配置幾何構造の関係解明についていくつかの結果を得ることができた。

また本萌芽研究の発展課題として、ポイントクラウドからなる一般の測度距離空間上での勾配流を扱う枠組みで逆問題を定式化する試みにも着手している。この定式化ではターゲットとするパーシステント図を実現する方向に勾配流を構成するが、解の時間大域解の存在や特異点の分類といった幾何解析的な興味深い問題と関わってくる。またそれに伴いパーシステント図上での最適輸送問題なども関わってくる。現在予備的な結果が得られつつあるが、本萌芽研究のさらなる

展開として極めて興味深い方向性であると思われる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 10 件)

- ① Marcio Gameiro, Yasuaki Hiraoka, Ippei Obayashi, Continuation of Point Clouds via Persistence Diagrams, *Physica D: Nonlinear Phenomena*, 査読有, Vol. 334, 2016, pp. 118-132, doi:10.1016/j.physd.2015.11.011.
- ② Yuko Fukuyama, Chihiro Nakajima, Shunsuke Izumi, Koichi Tanaka, Membrane Protein Analyses Using Alkylated Trihydroxyacetophenone (ATHAP) as a MALDI Matrix, *Analytical Chemistry*, 査読有, Vol. 88(3), 2016, pp. 1688-1695, doi : 10.1021/acs.analchem.5b03700.
- ③ Yasuaki Hiraoka and G. Kusano, Relative interleavings and applications to sensor networks, *Japan J. Indust. Appl. Math*, 査読有, Vol. 33, 2016, pp. 99-120, doi : 10.1007/s13160-016-0208-x.
- ④ Takenobu Nakamura, Yasuaki Hiraoka, Akihiko Hirata, Emerson. G. Escobar, and Yasumasa Nishiura, Persistent Homology and Many-Body Atomic Structure for Medium-Range Order in the Glass, *Nanotechnology*, 査読有, Vol. 26, 2015, pp. 304001, <http://dx.doi.org/10.1088/0957-4484/26/30/304001>.
- ⑤ Marcio Gameiro, Yasuaki Hiraoka, Shunsuke Izumi, Miroslav Kramar, Konstantin Mischaikow and Vidit Nanda, A topological measurement of protein compressibility, *Japan J. Indust. Appl. Math.*, 査読有, Vol. 32, 2015, pp. 1-17, doi:10.1007/s13160-014-0153-5.
- ⑥ Misa Takahashi, Jun Shigeto, Atsushi Sakamoto, Shunsuke Izumi, Kozi Asada, Hiromichi Morikawa, Dual selective nitration in Arabidopsis: almost exclusive nitration of PsbO and PsbP, and highly susceptible nitration of four non-PSII proteins, including peroxiredoxin II E, *Electrophoresis*, 査読有, Vol. 36(20), 2015, pp. 2569-2578, doi: 10.1002/elps.201500145.
- ⑦ Tomohiro Sasanami, Shunsuke Izumi, Naoki Sakurai, Toshifumi Hirata, Shusei Mizushima, Mei Matsuzaki, Gen Hiyama, Eriko Yorinaga, Takashi Yoshimura, Kazuyoshi Ukena, Kazuyoshi

Tsutsui, A unique mechanism of successful fertilization in a domestic bird, Scientific Reports, 査読有, Vol. 5, 2015, pp. 7700, doi:10.1038/srep07700.

- ⑧ Atsushi Sakamoto, Takashi Nishimura, Yoh-ichi Miyaki, Shunsuke Watanabe, Hiroshi Takagi, Shunsuke Izumi, Hiroshi Shimada, In vitro and in vivo evidence for oxalate oxidase activity of a germin-like protein from azalea, Biochemical and Biophysical Research Communications, 査読有, Vol. 458(3), 2015, pp. 536-542, doi: 10.1016/j.bbrc.2015.02.002.
- ⑨ Hiroyuki Nishi, Noriko Maeda, Shunsuke Izumi, Sayomi Higa-Nakamine, Seikichi Toku, Manabu Kakinohana, Kazuhiro Sugahara, Hideyuki Yamamoto, Differential regulation of epidermal growth factor receptor by hydrogen peroxide and flagellin in cultured lung alveolar epithelial cells, European Journal of Pharmacology, 査読有, Vol. 748, 2015, pp. 133-142, doi: 10.1016/j.ejphar.2014.12.016.
- ⑩ Satoshi Nakata, Ryoichi Tenno, Ayano Deguchi, Hiroya Yamamoto, Yoshikazu Hiraga, Shunsuke Izumi, Marangoni flow around a camphor disk regenerated by the interaction between camphor and sodium dodecyl sulfate molecules, Colloids and Surfaces, A: Physicochemical and Engineering Aspects, 査読有, Vol. 466, 2015, pp. 40-44, doi: 10.1016/j.colsurfa.2014.10.041.

[学会発表] (計 1 件)

- ① 平岡裕章, 32nd International Symposium on Computational Geometry (SoCG 2016), Topological Data Analysis on Materials Science: Statistical Characterization of Glass Transition, 平成 28 年 6 月 15 日-16 日, 「Boston(アメリカ)」.

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]

ホームページ等
http://www.wpi-aimr.tohoku.ac.jp/hiraoka_lab/

6. 研究組織

(1) 研究代表者

平岡 裕章 (HIRAOKA, Yasuaki)
東北大学・原子分子材料科学高等研究機構
・教授
研究者番号：10432709

(2) 研究分担者

泉 俊輔 (IZUMI Shunsuke)
広島大学・理学研究科・教授
研究者番号：90203116

(3) 連携研究者

()

研究者番号：

(4) 研究協力者

()