

X 線顕微分光イメージデータのパーシステントホモロジーによる TDA 解析

木村正雄¹, 城戸大貴¹, Mayrene UY¹, 丹羽尉博¹, Xichang Gao², 赤木和人², 大林一平³

¹ 高エネルギー加速器研究機構(KEK) 物質構造科学研究所

² 東北大学・材料科学高等研究所 (AIMR), ³ 岡山大学・AI・数理データサイエンスセンター

X 線顕微分光では、空間(3D)+エネルギー+時間=5D(次元)の多次元かつビッグデータが計測され、人間による解析だけでは、データに含まれる情報を十分に引き出すことが困難である。そこで応用数学(幾何学)を活用した位相的データ解析(TDA)による情報抽出の取り組みを紹介する。

放射光やクライオ電顕に代表されるように、材料や生体試料を **characterization** するための計測手法の発達はめざましい。例えば、XAFS による化学状態の計測では、従来は材料の $1\text{mm}^2 \times 10\mu\text{m}$ の領域の平均情報だったのが、数 10nm の空間分解能で三次元の顕微計測が可能になっている。その結果、空間(3D)+エネルギー+時間=5D(次元)の多次元データや複数のデータセットという、ビッグデータを手にすることになる(図1)。例えば、X 線顕微分光で三次元の化学状態解析を **operando** 計測すると、状況によっては TB のデータサイズとなり、データを眺めることすら大変である。

そうしたデータを解析して情報を抽出するには、人間が頑張ってデータを眺めるだけでなく、近年急速に広まっている情報科学や数理学、様々なデータベースを活用したアプローチが不可欠な時代に突入した(データの持ち腐れにならないためにも)。

X 線顕微分光では、材料やその反応の特徴が、多次元(ビッグ)データ空間での‘かたち’の‘うごき’として計測される。そこで、我々はその変化を指標化・定量化することにより、X 線顕微分光の多次元(ビッグ)データに含まれる情報を抽出することを着想した。具体的には、パーシステントホモロジーによる位相的データ解析(TDA)により、X 線顕微分光データから反応サイト (**trigger sites**) を予備知識無しに特定することに挑戦している[1,2]。

本アプローチを機械学習等の方法と組み合わせることにより、酸化物(焼結鈷)の還元反応に伴うき裂の発生サイトの予測[1]や、CFRP のナノスケールでのき裂発生メカニズム解明[3,4]を行うことに成功している。本アプローチは、特定された **trigger sites** を実際の材料空間で確認することができる **white-box** 的であるのが特徴で、従来蓄積されてきた材料科学的知見と照合してその妥当性を検証できる。これに多次元データの可視化技術[5]を組み合わせることにより、X 線顕微分光の多次元(ビッグ)データに含まれる情報を最大限活用することができる。

本研究を進めるにあたり JP19H00834, 22H05109, JPMJMI22C1 の支援を受けた。放射光実験は PF-PAC 2019S2-002, 2022S2-001 で実施した。

[1] M. Kimura, et al, *Sci. Rep.*, **8**, 3553 (2018), [2] I. Obayashi, Y. Hiraoka, M. Kimura, *J. App. Comput. Topo.*, **1**, 421 (2018), [3] M. Kimura, et al, *Comp. Sci. Tech.*, **230**, 109332 (2022), [4] M. Uy, et al., (submitted, 2024), [5] H. Igarashi et al., *J. Synchrotron Rad.*, (accepted, 2024).

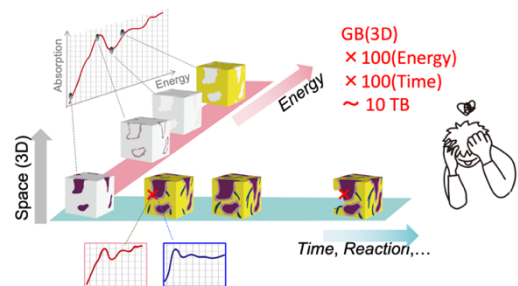


図 空間(3D)+エネルギー+時間=5D(次元)データからどのようにして情報を引き出すか?