

Graduate School of Advanced Science and Engineering
Waseda University

博士論文概要
Doctoral Dissertation Synopsis

論文題目
Dissertation Title

Development of Data-Driven Representation Methods for Microstructures of
Inorganic Thin Films with Two-Dimensional X-Ray Diffraction

2次元X線回折測定データを用いた無機半導体薄膜微細構造の特徴量抽出
手法の開発

申請者
(Applicant Name)
Akihiro YAMASHITA
山下 晶洸

Department of Advanced Science and Engineering, Research on Life Science and Medical
Bioscience

December, 2021

Society 5.0 をキーワードに高度に情報化されていく社会において、先端材料を要求するスピードも高速化している。この要求に応えるため、データ駆動による材料開発の高速化が期待されている。データ駆動型材料開発で特に期待されているのは、ビッグデータを活用し、所望の物性値から材料を予測する逆問題的アプローチである。このアプローチを支援するため、物質・材料研究機構の DICE や Materials Project のように公的なデータベースの構築が世界中で行われている。しかしながらこれらのデータベースと、個々の研究グループが所有しているデータとを連携してデータ駆動型材料開発を行うには、乗り越えるべきギャップがある。公的なデータベースに収録されているデータは、既存の論文から抽出したものやシミュレーション結果など、解析済みの、質の高いデータである。一方で個々の研究者が所有しているデータは、論文に用いた少量のデータの他は、質がより低いデータや、未解析測定データのような生データである。そのため外部のデータベースと連携して研究する際は、手持ちのデータの中でも質の良いデータを解析し、形式を揃える必要がある。これは非常に時間がかかり、結果として、手持ちのデータも活用してデータ駆動型材料開発を行うことが可能な研究グループは、外部にデータ整形を依頼する予算があるか、マンパワーが十分あるような大規模研究グループに限られる。このような背景の下、本研究では、小規模研究グループでも手持ちのデータによる独自性を持ったデータ駆動材料開発を行うことを可能とするための手法の探求を行った。

この目的を達成するため、測定データに教師なし学習の技術を適用することで、逆問題的アプローチで重要となる機械学習モデルに適用できる特徴量へ変換する方法について研究した。一般に機械学習モデルを適用する際にはデータをどう表現するかという特徴量の選択が重要である。この特徴量は機械学習モデルのパフォーマンスの主要な決定要因であるため、情報科学においては表現学習という名目の下、精力的に研究されてきた。それにも関わらず、データ駆動型材料開発の研究ではこの重要性は見落とされがちであった。なぜなら材料科学の立場からは、材料的新規性が乏しく、スコープの範囲外と見なされているからである。なお情報科学の立場からは、特徴量抽出はビッグデータを用いれば基本的に解決される問題であり、現在では異種データ間の連携など、深層学習を活用した、より高度な問題へ注目が移っているため、材料科学における問題点は興味をもたれていない。ビッグデータを準備するのが容易ではなく、また準備できても上述のような問題点がある材料科学においては、教師無し機械学習を用いた特徴量抽出の研究は、データ駆動型材料開発を進める上で未だに重要である。今回は数ある測定データの中でも、無機半導体薄膜の 2 次元 X 線回折 (2D-XRD) の測定データを対象とし、材料開発に適した特徴量となっているか評価するため、物性解析の観点から評価を行った。この研究により、多くの研究グループで、データベースと連携したデータ駆動型材料開発を可能となると期待される。さらに機械学習により測

定を高速・効率化する計測インフォマティクスのアプローチも行い、データ活用までの時間短縮を行った。

本論文は全 6 章で構成されている。以下で各章の内容を概説する。

第 1 章では序論として、研究の背景と目的について述べる。マテリアルズインフォマティクスとも呼称されているデータ駆動材料開発の概要、そしてこの分野における特徴量についてまず概説する。その後、本研究で着目している半導体薄膜、そして測定手法の背景を述べ、本研究の立ち位置を明確化する。

第 2 章では、2D-XRD 画像データの特徴量抽出手法として、非負値行列因子分解(Non-negative matrix factorization: NMF)の活用を提案する。準備できるデータ数の少なさ以外にも、以下の二つの理由のために、測定データをそのまま機械学習モデルの特徴量として用いるにはできない。①測定データは解像度の関係で基本的に高次元である。これはモデルのパラメータ数を増やし、結果として必要なデータ数を増やす。②材料科学の測定データは、情報科学の分野で用いられる画像データにくらべ、高次元である割にパターン数が少ない。加えてこの少ないパターンにおける微かな変化が重要であることが多い。この二つの問題点はさらに、情報科学で既に用いられている特徴量の活用も難しくしている。そのため NMF がこれらを解決する特徴抽出手法であるかを検討した。NMF は各データを、少数の非負な因子の線形結合で表現するため、他の特徴量抽出手法とは異なり解釈性が高いという利点がある。2D-XRD の画像データセットに適用したところ、回折角および χ 軸方向の広がりに基づいて次元削減を行い、特徴量を決定していることを確認した。これにより得られた特徴量には、結晶化具合が反映されており、材料科学の観点からの要求に応えられる程の分解能を有していることを確認した。また NMF が学習した特徴量抽出器は、成膜手法が異なる未学習データにも適用可能であることを確認した。

第 3 章では、第 2 章で述べた特徴量抽出手法によって、深層学習モデルの精度が向上するかの検討を行った。検討には逆問題的アプローチでの材料探索において重要な深層学習モデルの一つである、変分オートエンコーダを用いた。変分オートエンコーダは生成モデルと言われるモデルの一手法であり、入力データを連続確率変数に変換できると共に、新しいデータを生成できることが大きな特徴である。変分オートエンコーダが学習した潜在空間には、特徴量で示されている結晶化度を反映していることを確認した。これにより、潜在空間における試料間の関係を見ることで、成膜条件の検討に用いることが可能であることを報告する。実際にスパッタリング法は Pulsed Laser Deposition 法に比べて多様な膜を作製できると潜在空間で示唆されていることを確認した。このことは、生成モデルを用いることで、成膜手法や条件を視覚的に比較することにつながると期待できる。

第 4 章では、第 2 章、3 章で考察した手法を用い、In-Ga-O 半導体薄膜の 2D-XRD 画像データセットを、材料科学の観点から解析した。酸化インジウムに酸化ガリ

ウムを混ぜることでバンドギャップが広がることが知られており、新規パワー半導体材料として期待されている。しかしながら両者は結晶構造が大きく異なるため、成膜条件と結晶構造との対応関係では不明点が多い。第4章と第5章で提案した手法を用いることで、この関係性がより明らかになることを報告する。

第5章では、2D-XRD 画像データに含まれる、回折シグナルの密度情報が特徴量になりうるか可能性の検討を行った。データの密度を調べる手法である **Ordering points to identify the clustering structure (OPTICS)** を用いることで、回折シグナルの強度や広がり具合をグラフで表現できること確認した。さらにこの密度情報は、短時間測定でも十分に得ることができると分かったため、計測インフォマティクスの観点から、測定時間の短縮方法についても検討を行った。密度ベースクラスタリングの一手法である **Density-based spatial clustering of applications with noise (DBSCAN)** と組み合わせることで短時間測定でも高い S/N 比のデータが得られることを報告する。さらに 2D-XRD 画像中の密度関係をグラフで表現する方法についても検討を行った。グラフ表現では、NMF を用いた特徴量抽出とは異なり、回折シグナルの微妙な位置変化を捉えられると期待できる。2D-XRD 画像をグラフで表示したデータセットを、変分グラフオートエンコーダに学習させたところ、簡単な設定でも成膜手法の差異を学習していることを確認した。

第6章では結論として本研究の成果を総括し、実応用までの課題や展望、事業性について述べる。

List of research achievements for application of Doctor of Engineering, Waseda University

Full Name : 山下 晶洸

seal or signature

Date Submitted(yyyy/mm/dd): 2022/2/10

種類別 (By Type)	題名、発表・発行掲載誌名、 (theme, journal name, date & year of publication, name of authors inc. yourself)
論文	
○	1 <u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagyu, Toru Asahi, and Toyohiro Chikyow, "Direct feature extraction from two-dimensional X-ray diffraction images of semiconductor thin films for fabrication analysis" Science and Technology of Advanced Materials: Methods, in press
○	2 <u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagyu, Toru Asahi, and Toyohiro Chikyow, "Accelerating two-dimensional X-ray diffraction measurement and analysis with density-based clustering for thin films," Japanese Journal of Applied Physics, 60, SCCG04 (2020)
○	3 Takahiro Nagata, Takeshi Hoga, <u>Akihiro Yamashita</u> , Toru Asahi, Shinjiro Yagyu, and Toyohiro Chikyow "Valence Band Modification of a (GaxIn1-x)2O3 Solid Solution System Fabricated by Combinatorial Synthesis" ACS Combinatorial Science 2020, 22, 9, 433–439
講演	
1	<u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagyu, Toru Asahi, and Toyohiro Chikyow, "Quasi-Continuous Representation of Crystal Structure of Thin Films with Two-Dimensional X-Ray Diffraction and Non-Negative Matrix Factorization" 2021 MRS Fall Meeting & Exhibit, Boston & Online, Nov. 29 - Dec. 2 & Dec. 6-8, 2021
2	山下 晶洸、長田 貴弘、柳生 進二郎、朝日 透、知京 豊裕 「密度ベースクラスタリングを用いた2次元X線回折測定の高速化」 第68回応用物理学学会春季学術講演会、オンライン、2021年3月
3	<u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagyu, Toru Asahi, and Toyohiro Chikyow, "Automated Cluster Analysis of 2-Dimensional X-Ray Diffraction for Composition Spread Oxide Thin Film Fabricated by Combinatorial Synthesis, Aiming to Visual Information-Guided Material

List of research achievements for application of Doctor of Engineering, Waseda University

Full Name : 山下 晶洸

seal or signature

Date Submitted(yyyy/mm/dd): 2022/2/10

種類別 (By Type)	題名、発表・発行掲載誌名、 (theme, journal name, date & year of publication, name of authors inc. yourself)
	Discovery" 2020 Virtual MRS Fall Meeting & Exhibit, Online, Dec., 2020
4	<u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagy, Toru Asahi, and Toyohiro Chikyow, "Accelerating 2-Dimensional X-Ray Diffraction Measurement and Analysis with Density-Based Clustering for Thin Films" 33rd International Microprocesses and Nonotechnology Conference (MNC2020), Online, 11, 2020
5	<u>山下 晶洸</u> 、長田 貴弘、柳生 進二郎、朝日 透、知京 豊裕 「視覚情報による薄膜材料探索効率化に向けた、2次元X線回折手法の自動解析技術の開発」 第81回応用物理学会秋季学術講演会、オンライン、2021年9月
6	<u>山下 晶洸</u> 、長田 貴弘、柳生 進二郎、朝日 透、知京 豊裕 「計測インフォマティクスに向けた2次元X線回折解析プログラムの開発」 第67回応用物理学会春季学術講演会、東京、2020年3月
7	宝賀剛、 <u>山下 晶洸</u> 、朝日 透、知京豊裕、長田貴弘 「(GaxIn1-x)2O3固溶体薄膜における結晶構造および電気的特性の検討」 第67回応用物理学会春季学術講演会、東京、2020年3月
8	<u>Akihiro Yamashita</u> , Takahiro Nagata, Shinjiro Yagy, Toru Asahi, and Toyohiro Chikyow, "2D-X-Ray Diffraction Data Handling Flow for Time Efficient Measurement" 13th MANA International Symposium 2020, Tsukuba, 2, 2020