

先端材料開発に向けた、AI 先端計測技術の 多角的視点からの考察（第1回）

岡本和也*
青柳里果****

杉山昌章**
富谷茂隆*****

武藤俊介***

1. 序 論

社会基盤を支える先端材料研究において、“優れた研究”は“優れた評価・分析装置”に支えられているという一つの事実がある。各装置の高度化が進み原子レベルから量子レベルに迫る先端的な検査・計測・分析技術(以下、「先端計測技術」と略す)に関する国の事業の歴史は古く、2004年のJST(科学技術振興機構)の「先端計測分析技術・機器開発事業」にさかのぼる。ここでは、最先端の研究ニーズに応えるため、将来の創造的・独創的な研究開発に資する先端計測分析技術・機器およびその周辺システムの開発を推進され、多くの研究成果を経て2021年3月に、全課題の研究開発期間が終了しプログラムを終結した⁽¹⁾。

現在、研究開発の進め方そのものに変革を迫る第3次AI(Artificial Intelligence: 人工知能)の大きな流れが押し寄せてきており、この技術に関わる各種データからの情報抽出によるインフォマティクス研究が活況にある。特に、様々な機能を生み出す先端材料の研究開発を効率的に進めようとする領域では、マテリアルズ・インフォマティクスという手法が注目され、材料特性のデータベース化やその統計学的な有効活用が研究されている。また民間企業においても、様々な製造プロセスデータを各社とも独自の方法でデータベース化

し、効率的な製造方法や生産安定化のためのプロセス制御の高度化に関する研究が進められている。同様に、先端計測技術にもこのAI/ML(Machine Learning: 機械学習)技術の視点を取り込まれようとしている。しかもこれはデータ駆動という新しい概念を含んでおり、従来のような単なるハードウェア開発やソフトウェア開発を主軸においた領域だけではなく、この概念を巧みに活用することで分析装置間の複合化が進み、装置開発や原理に関わる研究においても新しい糸口になり得る可能性を秘めている。

先端計測技術について産業という視座でみると、高度な要求仕様に基づき付加価値が高い存在であるにもかかわらず、市場集中性(高位集中寡占の状況)にあり依然として各機器固有の機能強化に注目するReactiveな形態にある。将来に向けては、各機器から出力される時間軸も考慮した各種データ(画像・波形・信号)に内在する従来は捉えきれなかった意味を複合的にAI/ML活用により見出し、さらに様々な計測機器の機能を統合した「自律的に考えるシステム」としてのProactiveな形態が期待される。そのためには実践的な人材育成のあり方についても議論すべき時にあろう。

本解説は2回にわけて構成される。多岐に渡る先端計測技術の中で、材料研究によく使われ国内での研究者・技術者が多い電子顕微鏡分野、放射光やその他X線技術分野、二次イオン質量分析技術(SIMS: Secondary Ion Mass Spectro-

* 山口大学大学院；教授・副研究科長(〒755-8611 宇部市常盤台 2-16-1)，大阪大学エマージングサイエンスデザイン R³ センター；招聘教授

** 大阪大学大学院；特任教授

*** 名古屋大学未来材料・システム研究所；教授

**** 成蹊大学；教授

***** ソニーグループ株式会社 R&D センター；主幹研究員

Advanced Measurement and Analysis Systems Using AI/ML for Next Generation Materials Development —Multifaceted View 1—; Kazuya Okamoto*, Masaaki Sugiyama**, Shunsuke Muto***, Satoka Aoyagi****, Shigetaka Tomiya*****(*Graduate School of Innovation and Technology Management, Yamaguchi University, Ube. **Graduate School of Engineering, Osaka University, Suita. ***Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya. ****Faculty of Science and Technology, Seikei University, Tokyo. *****R&D Center, Sony Group Corporation, Atsugi)

Keywords: *measurement, metrology, electron microscopy, AI(artificial intelligence), machine learning, multimodal, operando, information entropy, innovation, intellectual property*

2022年 1 月18日受理[doi:10.2320/materia.61.470]

metry)に代表される質量分析分野における技術を中心に取り上げる。また、AI/ML 活用の様々な検査・計測・分析技術を「AI 先端計測技術」として提起し、考えておくべき論点について自然科学および社会科学の双方の広い視野まで拡張し考察する。

2. AI 先端計測の意義と最新動向

2.1 AI の基礎となる機械学習と歴史観

AI の基礎となる ML(機械学習)とは取得データから現象の傾向をとらえる手法の一つであり、入力データから予測出力を得るモデル(関数)の学習と捉えることができる。先端計測におけるデータ解析のフローを図 1 に示す。膨大な情報を有する実世界データ X を装置で計測(実験)すると、観測モデル $f(X)$ に従って変換された一部の情報がサンプリングされ、さらに計測器およびその制御系の物理的な状態の揺らぎに伴う雑音 N が付加されてデータ Y が取得される。データ解析とはデータ Y から逆問題を解きモデル $g(Y)$ を生成し、実世界における知りたい情報となる推定値 Z を得ることを目的とする。 Z を高精度に推定するためには、可能な限り雑音 N に惑わされることなく計測による実世界からのモデル $f(X)$ を人または機械が事前に理解しておくこと、もしくはデータから推察することが必要とされる。日本は質の良い膨大なデータを蓄積しており、かつその価値を理解している研究者・技術者も多いことから質の良いモデルを構築する機会が多く、AI 先端計測という分野(産業および技術)は将来に向けての格好の成長領域となりえる。加えて、大量の高次元計測データに普遍的に内在するスパース(疎)性を利用し、計算量が次元数に対し指数関数的に増大する状況においても、実際の時間でデータから最大限の情報を効率よく抽出できるスパースモデリング(SM: Sparse Modeling)⁽²⁾手法の研究も進んでいる。これは AI 先端計測技術にとって重要で検討すべき事案となっている。

次に、AI について簡単に歴史的な俯瞰をする。2022年現在、第3次 AI ブームの中心となっているのはニューラルネットワーク(Neural Network: NN)である。NN は生物の脳の機能をモデル化したものであり、1943年にマカロックと

ピッツによって提案された神経細胞(ニューロン)モデルから派生し、現在の深層学習(DL: Deep Learning)に至っており以下のように纏められる。

(1) 第1次(1940年代から1960年代: 推論・探索の時代): 神経細胞モデルの提案後、識別問題に適用可能なパーセプトロンが開発され注目を集めるものの、線形分離不可能な問題を解けない欠点が指摘された。

(2) 第2次(1980年代: 知識表現・エキスパートシステムの時代): 誤差逆伝播法⁽³⁾による学習アルゴリズムが開発され、NN により任意の関数が近似可能となった。しかしながら、当時はコンピュータ性能が劣り取得できるデータ数も限られ、さらには勾配消失問題といった NN 自体の問題もあった。

(3) 第3次(2000年代後半から現在: 知識学習・表現学習の時代): 2006年に上記、勾配消失問題を解決する NN の事前学習⁽⁴⁾が提案され、深い層を持つ NN 構築が可能になり、その後、画像認識に対して高い性能を有する畳み込み NN(Convolutional NN: CNN)が2012年の世界的な画像認識コンテスト ImageNet Large-Scale Visual Recognition Challenge(ILSVRC)において優れた性能を示したことで、深層学習(DL)がブームとなった。

現在、CNN は医用画像診断等で実際に適用されている。この AI の具現化の背景には、微細化や3次元高密度実装技術による革新的な半導体デバイスの進化⁽⁵⁾、GPU(Graphic Processing Unit)の拡張適用性(nvidia A100⁽⁶⁾など)や様々なアルゴリズム、ソフトウェア開発の劇的な進化などがある。また、学習モデルの構築においても、従来の専門家知識に基づいて設計する必要があった特徴量抽出を機械が自動で行えるようになったため、幅広い分野への適用が可能になったといえる。

2.2 AI 活用による先端計測技術

コンピュータの指数関数的な進化と個人レベルに浸透したネットワーク環境は、先述の第3次 AI 技術と融和し、産業構造に大きな変革を起こしている。事実、自動運転化技術やロボット化技術、画像および音声認識システムは人類の生活を変え、さらには産業界におけるモノづくりを変革させるだけでなく、デジタル技術を活用し顧客に付加価値を提供できる組織や文化を創り続ける DX(Digital Transformation)という概念生成に至っている。この大きな流れは Big Data が得られる領域を中心に展開されているが、データ自身を生み出しかつ日本が高い競争力を誇る先端計測技術分野に対しては、AI 技術との融合による研究開発そして価値創造に関する考察などは未だ萌芽期にあると考えられる。

先端計測技術を構成する代表的な3要素として、検査・計測・分析技術を考えてみる。検査技術では製造工程の欠陥検査や製品の品質検査、計測技術では環境計測や3次元形態計測、分析技術では物理分析装置や化学分析装置がある。現在、それぞれの技術から供給される膨大なデータに対し最適な AI 技術の開発が求められ、さらには AI 技術が自然に

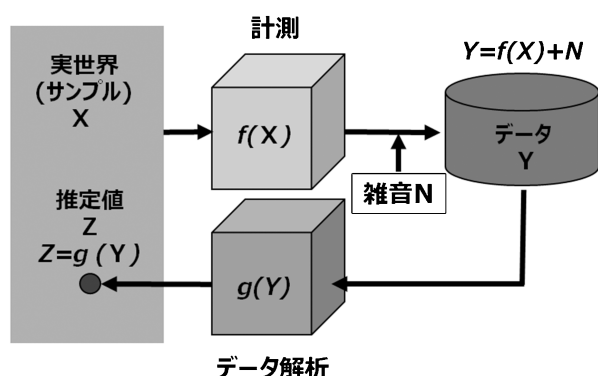


図1 先端計測におけるデータ解析のフロー。

先端計測装置の中に組み込まれた形で運用するための開発も重要となっている。その多くは解像度や検出感度の向上、微小試料など極限条件での計測・情報処理を主たる目的とし、複雑な計測や計算の原理を必要とする傾向が強い。そのため出力される測定情報は種々のノイズを含み、計測対象に関する不完全な情報しか含まないことも多い。さらに、機械学習は人間の学習とは根本的に異なるものであり、機械学習の場合にはコンピュータがどのように学習したのかが未知な部分が現状では含まれる。この場合、「見ると測るは科学の原点」などと言われていた従来の先端計測装置が持っている「正しさ」を、未知の部分を含む機械学習と組み合わせる良いのか、という根源的な問題が浮上してくる。その結果、現在の「AI 先端計測技術」の主たる技術論点は、得られた画像データ、スペクトルデータ、信号データなどをどのように機械学習させていくか、という事案が多く、先端計測装置自体が出してくるデータ部分までは、従来通りの原理原則を重視した正確なデータに基づくという立場を残している。このような前提をおくと、全く新しいアプローチが今後可能になりえる。すなわち、これらのデータを取り扱う時にその画像データが例えば、透過電子顕微鏡(TEM: Transmission Electron Microscope)、走査電子顕微鏡(SEM: Scanning Electron Microscope)、3次元アトムプローブ(3DAP: 3 Dimensional Atom Probe)のいずれかで得られた画像であっても、データのデジタル化処理(例えば前処理や特徴量抽出など)を施すことで異なる装置間であっても各種データのデジタル融合が容易となり、計測目的(対象)に対し単体装置では困難な新たな特性指標を得ることが可能となりえる。もちろん、分光光学という視点でみれば、TEMで得られるEELS(Electron Energy-Loss Spectroscopy)データと放射光で得られるXAFS(X-ray Absorption Fine Structure)データとは原理的にはほぼ等価な方法であり、このような観点でも新しいデータの融合が可能であり、これらを総称してデータ駆動と捉えて

いる。

一般にAI 先端計測というと、アルゴリズムに判断を委ねるような活用法が想定されるであろう。しかし実際に求められる重要な概念は、多角的な測定情報に対し異なる技術の組み合わせ・複合化によるデータの記述座標軸を増加、そして変数相互の相関関係を巧みに取り入れ、高効率かつ高精度に物理量を統合解析する「マルチモーダル」にあると言える。AI 技術は不完全情報から高い精度や信頼性を有する推定を行うことに長けている。すなわち、従来は個別装置の極限性能への向上が競争軸であったが、自動化や学習機能、さらには従来見えなかった領域のデータ抽出など、新しいユーザビリティ(ユーザの利便性)を主体としたソリューションとして新しい価値創造の世界で競争優位を發揮する時代が到来している。これはまさに、先端計測分野でのパラダイムシフトのトリガーを与えるものといえる。このためには当該産業の本質を見出す必要性もあり、AI 先端計測の本質に関わる社会科学的な研究も重要な因子となる。

この先端計測技術に関わる産業はどのようなものか。序論で述べたように計測機器分野の市場集中度は高く(高位集中寡占の状況にある)、多くの場合、産業分野ごとにその企業のセグメンテーションが明確であり新規参入は容易ではない。すなわち、現在のDX(およびIoT)社会のコアとなるのが先端計測技術であるにも関わらず、異なる視野での新しい産業分野としての創出は難しい。もちろん、GAFAM(Big Tech)のように独占企業や寡占企業こそがイノベーションを起こす場合もあるが、先端計測技術分野ではそのようなことは難しいであろう。

これまでの議論をもとにAI 先端計測技術の将来像を描くと、それは「未来型運用システム」の創成とその具体化にある。その方向性を3つの視点から解いたものを図2に示す。第1に、先端計測技術の重要性の拡大である。例えば、半導体製造プロセスにおける計測は過去には付加価値の

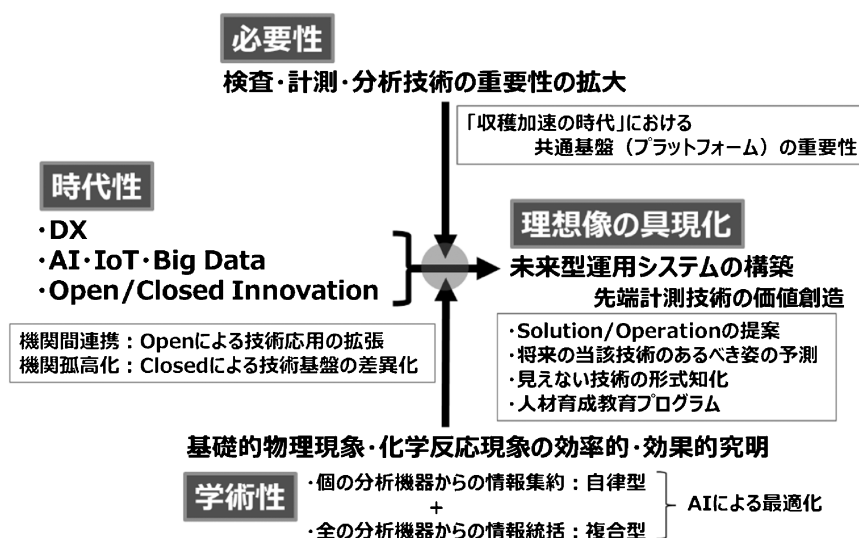


図2 AI 先端計測技術の将来像「未来型運用システム」。

ない工程とされていたが、現在では例えば Virtual Metrology という概念のもと製造プロセスに積極的に先端計測技術を適用し、高度な半導体 Chip コスト低減に至っている。とはいえ、これは半導体リソグラフィ分野であれば歴史的にはパターン計測からレチクルパターンを変更する光効果近接補正(OPC: Optical Proximity Correction)がその創発ともいえ、一つの技術的な進歩から次の進歩までの期間が指数関数的に短縮され、イノベーションが加速する収獲加速の時代⁽⁷⁾における共通基盤(プラットフォーム)の重要性という視点からも「必要性」が認知されていることである。第2が様々な分野へのAI活用という「時代性」である。DXは単なる対象やデータのデジタル化だけでなく産業や企業の事業構造をも変革する概念であり、先端計測はその基盤となり得る。イノベーションの視点からは様々なステークホルダーと連携して市場を創出する「Open Innovation」と個々の機関が独立して先行技術を創出する「Closed Innovation」の2つのイノベーションの両立による新しい世界観が重要であり、特にデータを生み出す先端計測技術においてはこの議論が重要である。なお、このイノベーションマネジメントのあり方もガイダンスとしてISO標準化されている⁽⁸⁾。第3が基礎現象の究明の「学術性」による差異化である。ここでの主題は、基礎的物理現象・化学反応現象の効率的・効果的究明にあり、「自律型」と「複合型」の2つの概念を考える。AI活用の視点から整理すると、自律型(Independent)とは「各機器にAIを活用し性能・ユーザビリティを向上させる」こと、複合型(Comprehensive)とは「単体もしくは繋がる複数の機器情報にAIを活用し新しい領域を発見すること」と本稿では定義する。以上、3つの方向性の融合により、未来型運用システムのあり方が明確にできる。それは、装置の枠を超えた課題解決と実行(Solution/Operation)の提案、将来の当該技術のあるべき姿の予測、見えない技術(暗黙知など)の見える化(形式知・デジタル化)ということにつながり、今後の展開に期待がかかる。

2.3 AI 先端計測の価値の議論

AI 先端計測において大切なことは、正しい答え・解釈、真実の提示である。例えば、電子顕微鏡技術などのデータ解析には多くの機械学習が適用され、その技術検討が行われている。しかし、検出原理である物理、また観察対象の中で起きている物理現象を正しく理解していない場合には誤った解釈や回答が生じている。この領域において、ものごとの真理を追究する学問に強い日本は最適なプレイヤーともいえ、AIによる自律的な新たな物理現象・科学的原理の発見と学理構築もあり得るだろう。他方、自動化に向かう過程で、技術的には画像や分析対象の三次元計測(3D)、時間軸でのダイナミック計測(4D)、さらには多数軸での計測が必要になる。それは空間的データと時間的データが揃うことで、我々が住んでいるリアルな3D/4Dの世界の現象と、分析・計測している対象とを完全に結びつけることができるからである。

このような背景のもと、日本学術振興会研究開発専門委員

会(2019年～2021年)⁽⁹⁾において、マルチモーダル計測(3.1節)、オペランド計測(3.2節)の視点が生み出された。そしてその先に、情報科学やシミュレーション技術と融合することにより、バーチャルな世界での分析・計測技術の必要性さえも議論として起きてくるかもしれない。もう一つの視点がデータの取扱いである。現在、インターネット、計算機、計測技術の急速な進化により膨大なデータ蓄積があること、そして、AI、データ科学の進化によりこれらのデータの潜在的価値抽出の必要性が増している。統計力学と情報科学に基づく量子情報科学分野はシグナル検出が関係する計測分野にとって大きな影響を及ぼす可能性を含み、その一つとして、本稿では情報エントロピー計測(3.3節)としての考察を試みた。

人材開発や科学技術の世界における日本の予算配分の様子をみると、先端計測技術は常に縁の下の力持ち的な立場であり、またその役割で良しとする文化がある。事実、JAIMA(日本分析機器工業会、2018)によると、国内製造業生産額124.3兆円に対してその産業を支える分析機器市場は2,162億円程度(～0.2%)に過ぎない。それは例えば、半導体製造工程においては、設計(論理からレイアウトまでの一連)が確実になされ、装置が安定に稼働し(工程能力指数 $C_p \geq 1.67$)、製造プロセスが安定に構築されていれば検査計測分析といった本質的に付加価値のつかないとされる工程はそもそも不要である、との風土が根付いていることにも起因しているのであろう(注：デバイスノード100 nm以降は状況が変化し、当該工程を積極的に取り入れた全体最適の視点から、半導体チップの製造歩留まりを上げるのが主流である。現在はDRAM ハーフピッチ17 nm(ノードレンジラベル3 nm)まで到達し、計測の重要性は自明である)。

しかしながらデータ駆動型社会において世界的に状況が変化している中、先端計測技術の開発データの取り扱い(所有権など)に関わるデータ契約(データの利用、加工、譲渡その他取扱いに関する契約)も重要な要素と認知され、日本でもそのガイドラインが構築され始めた⁽¹⁰⁾。さらにはそこからどのように融合し展開をすれば良いのか、という議論が欧米では盛んに行われている。そのうえで、知的財産権にも配慮しデータのオープン化を唱えている。データベース拠点の構築はそのような議論の上で成り立つものであり、先端計測という日本人の器用さとデータの信頼性を重視する文化に適した技術領域においては、日本は今後、世界のオープン化に対応し備えなければ孤立化していく可能性がある。このことから、AI 先端計測において重要となる技術として、(1)機能性暗号とプライバシー保護型機械学習のあり方、(2)プラットフォームにおけるブロックチェーンのあり方が重要となる。(1)として、機能性暗号技術が注目される。これはDX時代環境の制約・要件と新たに生ずるニーズ(データ保護と利活用の両立など)に応える機能を備えた暗号(プロキシ暗号、秘密計算など)である。これらは日本でも例えば、CRYPTEC(Cryptography Research and Evaluation Committees)主体で進められてきており、暗号化したままのBig Data分類や

プライバシー保護 DL といったシステムへ展開している (NICT・情報通信研究機構)⁽¹¹⁾。

次に、後者(2)ブロックチェーン (BC: Blockchain) は仮想通貨の基盤技術として広く知られているが、その優れた耐改ざん性、透明性等の特性を背景として広範なビジネス分野への応用が進められている。BC がビジネス変革の原動力とされるのは、従来の「中央集権的データベース」では実現困難であった様々な機関間でのデータ共有が可能となる非中央集権的なデータ構造にある。この BC を構成する主要な要素技術として、分散型台帳技術、P2P (Peer to Peer) ネットワーク、暗号化技術、合意形成アルゴリズム、スマートコントラクトの 5 つがあり、これらの技術が統合的に利用されることによって、非中央集権性、耐改ざん性、ゼロ・ダウンタイムといったブロックチェーンの特性が実現されている。そのデータ構造は極めて独特なものであり、過去のデータに新たなデータがチェーン状に連なることによって、完全なデータトレーサビリティと高度なセキュリティが実現される。事実、国内製薬業界では「Healthcare Blockchain Collaboration」が組織された。これにより、解決すべき課題である「疾患レジストリー」を非中央集権的に一元集約することによって、統合的で組織横断的な利活用を可能とされ、医薬品開発期間の短縮により医療の質向上に繋がることが期待される。この 2 つの観点の周辺技術は、AI 先端計測においても Data の取扱いの点で今後、学際的に重要となり得る。

3. AI 先端計測技術の具体例

3.1 マルチモーダル計測

3.1.1 マルチモーダル計測の意義と潜む課題

一般に科学計測は影絵に例えることができる。測定によって得られるデータとは計測対象物の全てではなく計測装置を通して特定の性質または応答を得たにすぎず、通常得られた情報に基づきその対象を特徴づける物理量を算出する。影絵も同様に 3 次元立体物にある方向から光を当てて特定の方向へ投影された形により、対象物が何であるかを観測者 (観客) は想定する。しかし、狐や犬に見える形は単なる手指の組み合わせを特定の方向から見た輪郭のみで模倣したものである。これはもちろん現代の科学計測に対しては単純すぎるものの、複数の違った角度 (多角的) から見ることで対象物を正しく理解する意義のメタファーとしては分かりやすい。これを AI 先端計測技術分野で考えてみる。一般に想定されがちなアルゴリズムだけに判断を委ねるような応用法でなく、先述の通り多角的な測定情報、例えば複数の異なる分光データの同時収集、時間軸追加 (次回 3.2 節のオペランド計測など)、データの S/N 比またはノイズモデルがそれぞれ異なる複数のテクニックの組み合わせなどを考える。これにより、これまで個別に測定・解析・解釈されていたデータを記述する座標軸を増やし、変数相互の相関関係をうまく取り入れ、効率よくかつより高精度に物理量を統合解析することが可能となる。よく言われている「マルチモーダル」というキーワ

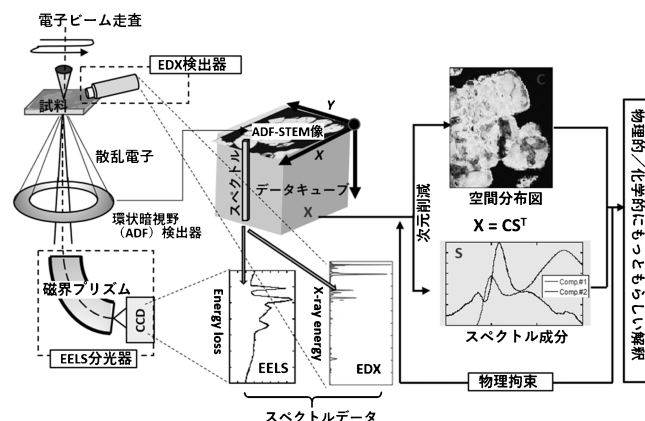


図3 走査透過型電子顕微鏡 (STEM) によるスペクトラムイメージデータ取得とそのデータ解析のための次元削減法の模式図。

ードに、これらの視点が含まれることの重要性を提唱する。

まずデータ解析によく使われる代表的な広義の ML (機械学習) テクニックを例にとり、計測という観点から十分に考えておかなければいけない課題について、以下に具体的な Case として記述する。

【Case 1】 ブラインド信号源分離、スペクトル分解などの次元削減 (DR: Dimension Reduction) 法: 図 3 に走査透過電子顕微鏡 (STEM: Scanning Transmission Electron Microscope) とそれに付随した代表的な分光法であるエネルギー分散 X 線分光 (EDS: Energy Dispersive X-ray Spectroscopy) と電子エネルギー損失分光 (EELS) におけるスペクトラムイメージデータから DR 法によって内在するスペクトル成分を分離し、それぞれの空間分布を可視化する「化学イメージング」の模式図を示す⁽¹²⁾⁻⁽¹⁴⁾。図に示すように、データキューブを展開した実験データ行列を X とし、 X を成分スペクトル行列 S とそれぞれの成分の空間濃度分布を表す行列 C の積に分解する⁽¹²⁾。ここで注意が必要な問題となる点について記述する。①アルゴリズムが排出した解答の正否を誰がどう保証するか: 分離された未知の相に対するスペクトル成分が解釈可能かどうかを判断することは一般に簡単ではない。②現実のデータに含まれるノイズをどのようにモデリングするか: 例えば、データのカウンタ数に依存しガウスノイズを仮定するか、ポアソンノイズの仮定によって結果が大きく変わりうる⁽¹⁵⁾。さらに非統計ノイズが含まれる場合はより複雑になる。③スパース (疎) でないデータにおけるクロストーク問題が発生する。すなわち、位置空間およびスペクトル空間でオーバーラップの大きいデータでは、成分間クロストークが生じることが多い。一例を図 4 に示す⁽¹²⁾⁽¹⁵⁾。半導体メモリ断面の暗視野 STEM 像である (a) の枠線部から取得したスペクトラムイメージデータを DR 法によって三成分分解した際の各成分の空間分布マップを (c) に示している。(d) の各スペクトル成分にみるように、↓の部分でスペクトル成分間のクロストークが生じている。これに対応して (c) の空間分布は (a) の相分布を正しく反映していない⁽¹¹⁾⁽¹²⁾。この

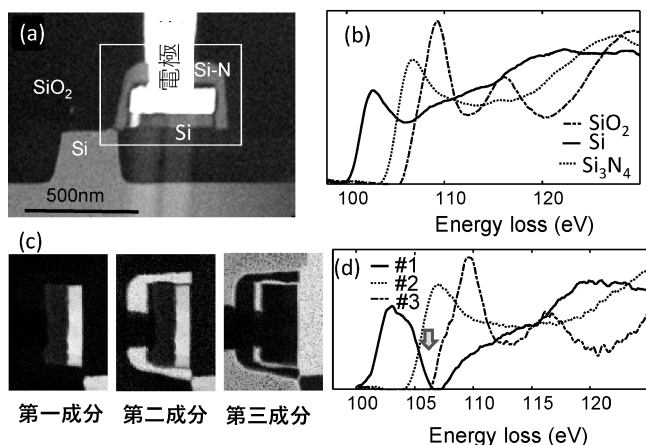


図4 (a) 半導体メモリ断面の暗視野STEM像. (b) (a) の各シリコン化合物部のSi-L_{2,3} 参照スペクトル. (c) (a) の枠線部から取得したスペクトラムイメージデータを次元削減法によって三成分分解した際の各成分空間分布マップ. (d) 各成分のスペクトル成分.

ように、行列分解法は成分スペクトル行列と重み行列が単純な行列積としてカップルしているため、一方のエラーは他方に伝搬する。事実このような誤った適用例が散見される。

【Case 2】 深層学習(DL)：①一般にニューラルネットワーク(NN)を使った解析には、まず訓練データとして大量(通常数十万件以上)の正解/不正解のタグのついた教師データを集める必要がある。この事実が正答率に深刻な制限を与える。また、実際の未知試料の結果において正答率がどの程度かは実は不明である。最近、Noise2Noise という「対象に対する正しい答えが不明な場合でも、同一対象に関する多数回の独立なノイズ付き計測結果が得られればDLによってノイズ低減器を得ることができる方法」が提案された⁽¹⁶⁾。しかし、実際の非統計ノイズが含まれる場合の有効性については不明である。②畳み込みNN(CNN)に代表される画像認識では、アルゴリズムが一体何をもって画像認識したか一般に知ることはできない。コンピュータは本当に画像データの本質を切り分けて判断したのか、あるいはまったく物理情報と関係のない些末な特徴を基に判断を下したかは全く闇の中である。もちろん、この説明可能AI(XAI: eXplainable Artificial Intelligence)実現への努力は続けられているが⁽¹⁷⁾、最終解答の選択肢が限られるTEM像分野におけるこのブラックボックスは後に大きな問題となりかねないと考えられる。

【Case 3】 ベイズ予測の適用：事前確率または事後確率モデルをどう取るべきかが論点となる。自明な場合を除き、多くの場合対象をどのような目で捉えるかという本質に関わり、予測結果はモデル依存であることは意外と見逃されている。モデルの選び方によっては、本質的に予測計算が収束しない場合もある。

【Case 4】 スパースモデリング(SM)および圧縮センシング(CS: Compressed Sensing)：次のような論点がある。実際のデータに特有のスパースな解を保証する適切なデータの変換をどのように選ぶか、選ばれたデータ変換空間において果

たして未知解は同様にスパース性を満たしているか。換言すると、圧縮処理後の逆変換によって元のデータの情報が保存されているか。さらには、正解が既知であるモデルデータで確認された場合においても、未知の構造が含まれた現実のデータではどのようにそれを保証するのが重要となる。

なお、昨今の急速なAIの活用事情を眺めてみると、上記Case 1 から4のデータ処理法に限らず、AIが質の悪いデータをあたかも魔法のように甦らせるかもしれないという大きな誤解を生んでいる場面がある。これは計測技術においてはさらに深刻な問題である。いかなる種類のデータであってもそこにある種の変換を施すことによって元のデータから不可逆的に情報が失われていくことは、数学的に証明されている⁽¹⁸⁾。元々そこに含まれていない、あるいは失われた情報はいかなる手段をもっても復活させることは不可能である。また、一つのスペクトルデータ(あるいは一枚の画像)から「ノイズのみ」を正しく切り分けて除去することは殆ど見込みがない。なぜなら統計ノイズは確率的に生成されるものであり、データ全体としての分布確率のみが与えられているため、ある特定の画素を取り出してその信号強度のうちノイズが占める割合を先験的に知ることはできない。実はこの問題は誤差論における精度と確度の関係に似ていて、このような課題を議論できる場が必要である。

3.1.2 広義のマルチモーダル計測の提案

計測とは、(おそらく多変数関数である)目的の物性の計測機器のもつ特定の物理変数への投影とみなすことができる。したがって、目的の物理量を陰に含んだ別の計測を組み合わせる(変数座標軸を増やす)ことはその物性本来の姿を浮かび上がらせることに他ならない。これらを踏まえたうえで、「計測現場から機械学習をどのように取り入れるか」という問いに対し「広義のマルチモーダル解析」を提案し、さらにこれを梃子としてこれまで進まなかった異分野融合を促進させる効果にも言及する。

「マルチモーダル」とは「複数の形式の、複数の手段による」などの意味に過ぎない。一方、「マルチスケール」解析ではマクロからミクロへ異なる階層構造へと順次分け入っていくイメージであるが、マルチモーダルはそれも包含しつつ同じ階層の中でも多角的に同時に観ること(データの多次元化)を目指している。以下、網羅的でなくSTEMスペクトラムイメージを主軸としていくつかの具体的な事例で、本解説で提案するマルチモーダル解析を緩やかに定義することにする。

(a) 複合分光(データ軸の多次元化)

図3のように同一場所(条件)から異なる種類の分光データ(この場合はEELSとEDSデータ)を取得する。この場合の最も簡単な取り扱い方は図5に示すようなConcatenation(連結)による次元削減である⁽¹⁹⁾。図5からわかるように、濃度行列Cを共通にとることで、分離された各相の化学状態(EELS)と元素組成(EDS)が一組のスペクトルとして行列Sの各行に格納される。連結処理における行列Cの共通化は強い拘束として働くため、信号雑音比(S/N比)の良い方

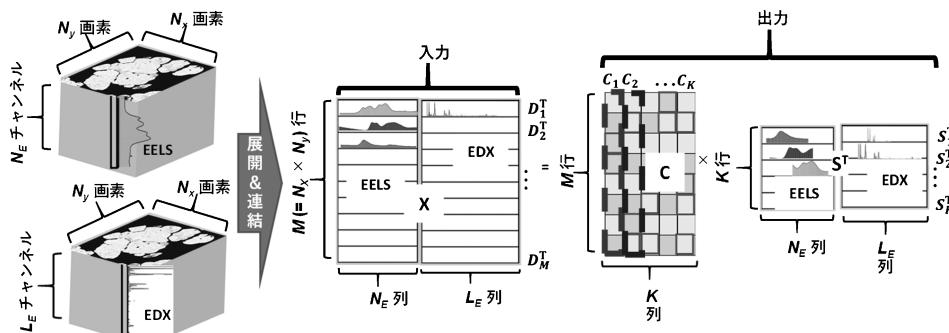


図5 二つの同時計測データを連結して次元削減する方法の模式図。

のスペクトルデータに重みをつけることにより，単独ではS/N比の良くないデータのスペクトル分離性能を向上させる効果も期待できる。

複合分光データのもう一つの有効な活用法として構造化データ融合(SDF: Structured Data Fusion)法が挙げられる⁽²⁰⁾。連結法と異なり，それぞれの分光データに対応する二つの行列方程式

$$\mathbf{X}_1 = \mathbf{C}_1 \mathbf{S}_1^T, \quad \mathbf{X}_2 = \mathbf{C}_2 \mathbf{S}_2^T \quad (\square^T: \text{行列転置}) \quad (1)$$

において行列 $\mathbf{C}_1, \mathbf{C}_2$ 間に $\mathbf{C}_1 = \mathbf{A} \mathbf{C}_2$ ($\mathbf{A}$ は任意の行列)で表される相関関係を設定することができる。このような一般的取り扱いの応用例は広く，高角度分解電子チャネリングX線/電子分光法(HARECXS/HARECES)として知られているサイト選択的要素分析/化学分析法に適用されている。図6に示すように特定のブラッグ反射列に対して試料への電子入射角度を連続的に変化させてEDS/EELSスペクトルを同時に取得することで，添加不純物の各ホスト元素サイトへの占有

率とサイト毎の不純物の化学状態(価数など)を定量的に求めることができる⁽¹²⁾。ここではMn, Niが二つのサイトにまたがって分布していることが判る。さらに理論計算とのフィッティングによって各サイトの元素占有率が求められる⁽²¹⁾。このスキームは上で次元削減法と称した行列分解法を一般次元に拡張したテンソル分解法と呼ばれる手法の一つである⁽²²⁾。二種の計測複合の例で取り扱うデータは3階テンソルであるが，テンソル化によってデータ種の数には制限がなくなり，任意の異なる指標間で個別に関係性を設定できる。したがって例えばナノビーム走査による収束電子回折収集データセットに代表される四次元(4D)-STEMデータ⁽²³⁾や，さらに他の分光データも同時に取り込んだ高次STEMデータへの有効利用が期待される。

(b) 時間軸の導入によるデータ軸次元拡張

最近特に注目されているオペランド計測(3.2節で議論)などの場合は，通常の計測変数軸に時間軸が加わる。データ構造としては前項と等価な形式になる。数学的な観点からは，単純にTwo-way解析の繰り返しと捉えることができ，上記連結データとはほぼ同じテクニックが適用できる。ある時刻のデータセットについて次元削減によって基底ベクトルを求めることを繰り返し，二次元座標-時間という三次元空間での各基底の時間発展をマッピングする，もしくは各エネルギーチャンネルの強度時間変動パターンとしての基底分解，という二種類の解析・可視化が考えられる。計測手段の複合化と併せることによりテンソル分解の手法がさらに柔軟な数学的取り扱いを与え得る。

(c) 異種計測情報結合，マルチスケールへの拡張

上記(a), (b)を更に拡張し，複数検出器で同時測定という制限も取り払い，各計測機器のデータをそれぞれAI活用によって数値化する。図7にAIを使った異種計測情報結合による物性解析のスキームを示す。同図で外側の円矢印上の各装置は共通した試料・物質測定で結ばれている。これらの中心にあるAIの様々なテクニックから各装置に伸びる矢印が各データの処理テクニックの適用を示している。こうして数値化されたデータは水平な右向き矢印によって示されるように統合され，物理的な拘束条件により複数の支配方程式に結び付けられている。すなわちこれにより多次元情報空間を構成し，そこにさらに別のAI技法による変数間の相関関係や拘

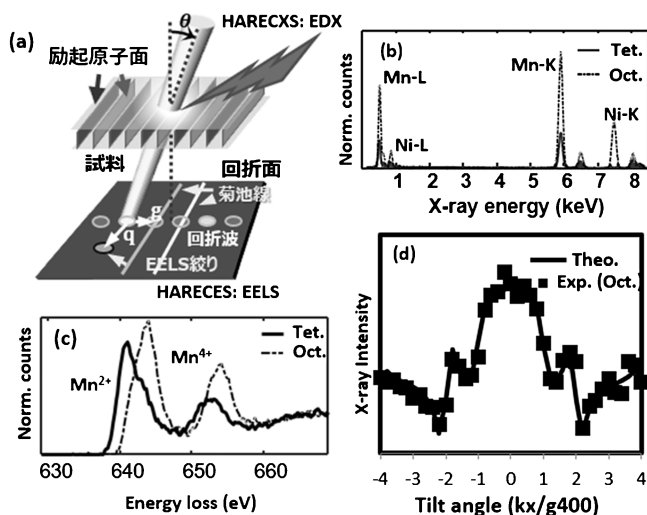


図6 (a) 高角度分解電子チャネリングX線/電子分光(HARECXS/HARECES)の模式図，(b) スピネル構造における酸素4配位四面体サイトと6配位八面体サイトの二つをMn, Niが占めている構造において，HARECXSデータから得られた二種類の原子サイトの元素ピーク群，(c) HARECXSデータから二つのサイトを占有するMnの価数状態を示すEELスペクトルの分離，(d) Mnスペクトル強度のビーム傾斜角度依存性。

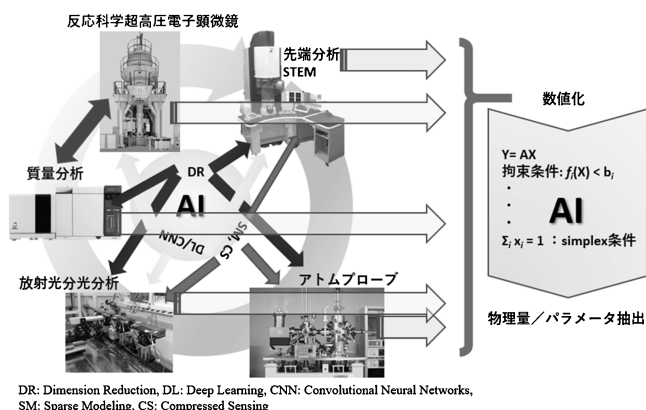


図7 AIを使った異種計測情報結合による物性解析のスキーム。

束条件を設定(あるいはそれらを最適化の解として導出)することにより特定の物性を立体的に浮かび上がらせるような試みをさらに提案する。この場合、標本点数、座標は必ずしも共有されていないため、大きな枠組みとしてのターゲット情報(求められる解)は設定されているものの、上記のような次元削減法をデータ全体に適用することはできない。そこで各データ(画像、スペクトル、数値表データ)あるいはその一部に対して、適切な情報整理・抽出処理を行った上でこれらを改めて融合することになる。最も簡便な方法は各情報を異なる軸の変数と捉えて、K-means法⁽²⁴⁾、サポートベクターマシン(SVM)⁽²⁵⁾などの良く知られている汎用クラスター解析を適用することであろうが、そこから意味を読み取るためには何らかのモデルシミュレーション、教師付きデータなどの経験によるタグ値が必要となろう。その点において結果の信頼性を担保しつつ自動化を図るためには、ハイスループット計測の成否がボトルネックとなることが容易に考えられる。

もう一つの方向性として、さらに進んだ手法としてマルチモーダル深層学習があげられる。このような試みはすでに医療分野で多くの事例が報告されており、解析のためのコードも公開されている⁽²⁶⁾。計測分野への応用として、例えばTEM/SEM像、スペクトラムイメージデータ、試料作製・熱処理条件などの情報を統合し機械特性や電気特性の経年劣化予測などが考えられる。ここでも十分な量の良質のデータが得られるかどうか、あるいは時間軸外挿のための適切な物理モデルを構築できるかどうかは鍵となる。

マルチスケール/マルチモーダルという観点では、データ解析を行う上での物理が共通である放射光分光/顕微鏡とTEM分光の連携が有効な事例の一つになると考えている。実はこのような戦略的・系統的な連携は世界的に見てもあまり進んでおらず、この点において、両分野でハードウェアリソースの豊富な日本が先端計測技術分野で世界をリードするチャンスがあると見ている。

3.1.3 マルチモーダル計測からみた新しい世界観

マルチモーダル科学計測へのAI応用の将来展望を探ってみると、以下のような可能性と問題点が考えられる。(1)測定と解析の自動化：すでに単体の計測機器では各メーカーで自

動化が進みつつある。複数機器間の直接連携のためにはデータフォーマットの統一化・共通化が必要なため、異なるメーカー間では進みにくい。(2)データ駆動による新たな科学法則発見：量子力学の誕生から100年以上経った現在、未だ基礎方程式の有効性を破る実験事実は見出されていない。マルチモーダルという観点では計測変数間の相関をデータ駆動で見出すという未開拓領域が残されている。歴史的に物性物理学における大きな発見は、従来の法則で外挿できない性質の発見から始まる。大量の実験データのAI/ML処理過程でデータ群から弾き出され通常捨てられるアウトライヤー(外れ値)が新たな発見の契機となる可能性を感じる。そこでは依然として深い洞察と経験によってシステムをコントロールするエキスパート(人間の)の存在が、実は本質的な要素となっていることが多い。

以上、これまで学際研究、異分野融合などのスローガンが叫ばれてから久しいが、現状では分野の縦割りやガラパゴス化は残念ながらさらに進んでいるのではないだろうか。そこで最近の計測DX化の動きは、ここで述べたような異分野間の垣根を取り払って必然的に連携を促す良い契機であると捉えたい。

第1回のまとめ

第1回はAI先端計測技術に関わる背景と概要を述べた。第3次AIの大きな流れにおいて、各種データからの情報抽出による様々なインフォマティクス研究が活況にあり、先端計測技術の領域にもこのAI/ML技術の視点が大きく取り込まれようとしている。ここにはデータ駆動という新しい概念を含んでおり、従来のような単なるハードウェア開発を主軸においた領域だけではなく、この思想を巧みに活用することで新しい装置開発や原理に関わる研究への新しい糸口さえも見え隠れしている。今回は多岐に渡る先端計測技術の中で、先端材料研究に多用される電子顕微鏡分野における「マルチモーダル計測」を取り上げた。これは、多角的な測定情報に対し異なる技術の組み合わせ・複合化によるデータの記述座標軸を増加、かつ変数相互の相関関係を巧みに取り入れ、高効率かつ高精度に物理量を統合解析する概念である。次回は、これを時間軸に展開した「オペランド計測」、スペクトルにおけるピークの多様性を数値化する「情報エントロピー計測」について述べる。さらに、先端計測分野の産業状況を知的財産情報等の視座から捉え、「AI先端計測技術」について包括的に解説する予定である。(つづく)

文 献

- (1) <https://www.jst.go.jp/sentan/gaiyou.html>
- (2) <http://sparse-modeling.jp/>
- (3) D. E. Rumelhart, G. E. Hinton and R. J. Williams: Nature, **323** (1986), 533-536.
- (4) G. E. Hinton and R. R. Salakhutdinov: Science, **313**(2006), 504-507.

- (5) 岡本和也, 佐藤了平: エレクトロニクス実装学会誌, **21** (2018), 531-541.
- (6) <https://www.nvidia.com/ja-jp/data-center/a100/>
- (7) レイ・カーツワイル: ポスト・ヒューマン誕生 コンピュータが人類の知性を超えるとき, NHK 出版, (2007).
- (8) <https://www.iso.org/standard/68221.html>
- (9) <https://jsps-ai-imat.mot.yamaguchi-u.ac.jp/>
- (10) 経済産業省: AI・データの利用に関する契約ガイドライン—データ編—(2018).
- (11) <https://www.cryptrec.go.jp/>
- (12) S. Muto and M. Shiga: Microscopy, **69** (2020), 110-122.
- (13) 武藤俊介, 志賀元紀: ふえらむ, **26** (2021), 32-40.
- (14) 武藤俊介, 志賀元紀: 表面科学, **37** (2016), 610-615.
- (15) 志賀元紀, 武藤俊介: 電気化学, **88** (2020), 42-46.
- (16) J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala and T. Aila: Proc. 35th Int'l. Conf. on Machine Learning, PMLR, **80** (2018), 2965-2974.
- (17) <https://human-centered.ai/explainable-ai/>
- (18) N. J. Beaudry: Quantum Information & Computation, **12** (2012), 432-441.
- (19) T. Thersleff, S. Budnyk, L. Drangai and A. Slabon: Ultramicrosc., **219** (2020), 113116.
- (20) L. Sorber, M. Van Barel and L. De Lathauwer: IEEE J. Sel. Top. Sig. Proc., **9** (2015), 586-600.
- (21) S. Muto, J. Spiegelberg, M. Shiga, M. Ohtsuka and J. Rusz: Proc. 10th Pacific Rim Int'l. Conf. Materials & Processing, PRICM10, (2019), 720-729.
- (22) A. Cichocki, D. P. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. F. Caiafa and A. H. Phan: IEEE Sig. Proc. Mag., **32** (2015), 145-163.
- (23) C. Ophus: Microsc. Microanal., **25** (2019), 563-582.
- (24) J. B. MacQueen: Proc. 5th Berkeley Symp. on Mathematical Statistics and Probability, (1967), 281-297.
- (25) V. Vapnik and A. Lerner: Aut. Remote Control, **24** (1963), 774-780.
- (26) B. Guan, F. Liu, A. Haj-Mirzaian, S. Demehri, A. Samsonov, T. Neogi, A. Guemazi and R. Kijowski: Osteoarthritis and Cartilage, **28** (2020), 428-437.



岡本和也



杉山昌章



武藤俊介



青柳里果



富谷茂隆

★★★★★★★★★★★★★★★★★★★★

山口大学大学院教授・大阪大学招聘教授
東京大学大学院工学系研究科修士。ハーバード大学経営大学院修了(PMD)。㈱ニコン・開発部門GMを経て、2016年山口大学大学院技術経営研究科教授。2017年より同副研究科長・現職。博士(工学)
専門分野: 半導体工学, 集積光学, 電子光学, 技術経営論

◎半導体デバイスの要素開発から製造・検査装置の開発に従事後, 新事業創成方法論の研究に従事。

★★★★★★★★★★★★★★★★★★★★

大阪大学大学院特任教授
1985年大阪大学大学院基礎工学研究科博士後期課程修了。同年, 新日本製鐵㈱第一技術研究所入社。新日鐵住金に統合後, 席主幹研究員を経て2015年大阪大学大学院特任教授。2019年より大阪大学日本製鉄材料基礎協働研究所 現職。博士(工学)
専門分野: 金属物性学, 鉄鋼材料

◎電子顕微鏡技術を中心に材料の微細組織評価と特性制御研究に従事。

★★★★★★★★★★★★★★★★★★★★

名古屋大学未来材料・システム研究所教授
1986年大阪大学大学院基礎工学研究科博士前期課程修了
1988年大阪大学教養部物理学教室助手, 1995年名古屋大学理工科学総合研究センター助教授, 2004年名古屋大学大学院工学研究科教授, 2013年4月より現職。博士(工学)
専門分野: ナノ電子分光, 数理データ解析

◎透過電子顕微鏡および関連分光法によるナノ物性可視化と新規分析手法開発に従事。本稿3.1節担当。

★★★★★★★★★★★★★★★★★★★★

成蹊大学教授
2002年早稲田大学大学院理工学研究科博士後期課程修了。2005年島根大学生物資源科学部助教授/准教授, 2014年成蹊大学理工学部准教授, 2017年4月から現職。博士(工学)
専門分野: 表面分析, データ解析(多変量解析・機械学習), 質量イメージング

◎生体試料・バイオ材料・有機物試料を主な対象とした飛行時間型二次イオン質量分析による表面分析とそのデータ解析に従事。本稿3.3節担当。

★★★★★★★★★★★★★★★★★★★★

ソニーグループ㈱主幹研究員
1988年慶應義塾大学大学院理工学研究科博士前期課程修了。同年, ソニー㈱入社, 2019年東京工業大学特任教授(兼任)。2019年12月より現職。応用物理学会フェロー。博士(工学)
専門分野: 半導体材料物性, 半導体デバイス, 材料解析, 半導体信頼性評価

◎材料解析(主として微細構造解析/光学評価)およびインフォマティクスを活用した半導体材料・デバイスの開発に従事。本稿3.2節担当。

★★★★★★★★★★★★★★★★★★★★