

SIP-MI プロジェクトにおける 特性空間分析システムの開発

井上純哉* 岡田真人** 長尾大道***
横田秀夫**** 足立吉隆*****

1. はじめに

2014年度にスタートした第1期 SIP「革新的構造材料」では、統合型の材料開発システムの礎として、材料開発の基本要素であるプロセス・組織・特性・性能の連関を計算機上でシームレスに実現し、材料の組織や性能の予測を可能にする「マテリアルズインテグレーション(以後、MI と記す)」の構築が目指された。本稿の主題である「特性空間分析システム」は、データ科学的手法の適用により、実験や数値計算に基づくデータベースを用いて物理モデルや現象論モデルのモデルパラメータを最適化し、MI 統合システムに導入された組織予測や性能予測といった各種モジュールの飛躍的な高精度化をサポートするためのシステムとして構想された。しかし、本プロジェクトでは単なる予測の高精度化だけにとどまらず、第2期 SIP の主題となる「逆問題」の効率的な解析も視野に入れ、現象を説明する有効な説明変数や物理モデル・現象論モデルの自動選択や、大規模な3次元材料組織データの効率的な管理手法や、3次元組織情報の有効な活用手法に関する研究開発も行った。

この様な特性空間分析システムは、大きく分け「データ駆動型予測システム」と「3D/4D 分析システム」の2つのシステムから構成される。以下ではそれぞれのシステムの内容を実際の適用事例を踏まえて簡単に紹介する。

2. データ駆動型予測システム

データ駆動型予測システムでは、個々のモジュールの高精度化に不可欠となるデータ同化手法と、説明変数やモデルの

選択を可能にするスパースモデリングの二つのデータ科学的手法に着目し、そのモジュール化と具体的な材料問題への適用を試みた。

(1) データ同化

データ同化は、気象学や海洋学の分野において盛んに研究されてきた手法である⁽¹⁾⁽²⁾。これら分野では、物理法則などに基づいて時空間モデルを構成し、大規模な数値解析を行なうことにより、実際の気象変動をすることが可能であるように思われる。しかしながら、この様な予測では、システムが持つ非線形性が原因となり、初期条件や境界条件に加えモデルに含まれるパラメータの与え方により、予測結果が現実の現象より大きく乖離した著しく不適切な結果となることが問題となる。そのため、数値解析における初期条件、境界条件ならびにモデルパラメータを、実際の観測に基づいた適切なものに設定することで、実現現象をより正確に再現した数値解析を可能にする手法が不可欠となっていた。それを可能にしたのがデータ同化である。構造材料を対象とした場合、予測の難しさは気象学や海洋学と同様に、構造材料が持つマルチスケールな不均一性や物理モデルの不完全性に起因する。データ同化ではこの様な不均一性や不完全性を踏まえた予測が可能になるだけでなく、その結果生じる数値解析結果の不確実性も定量的に評価できること(Uncertainty Quantification, UQ⁽³⁾)も、所謂一般的な機械学習ではなくデータ同化を用いることの最大のメリットと言える。データ同化を実現する手法には、時間を順方向に辿る逐次データ同化手法と、逆方向に辿る4次元変分法がある。

(a) 逐次データ同化手法

逐次データ同化手法は、対象とする物理現象をモデル化し

* 東京大学先端科学技術研究センター；准教授(〒153-8904 東京都目黒区駒場 4-6-1)

** 東京大学新領域創成科学研究科；教授

*** 東京大学地震研究所；准教授

**** 理化学研究所量子工学研究センター；チームリーダー

***** 名古屋大学工学部；教授

Development of Data-Driven System in Materials Integration; Junya Inoue*, Masato Okada**, Hiromichi Nagao***, Hideo Yokota**** and Yoshitaka Adachi***** (*Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo. **Graduate School of Frontier Science, The University of Tokyo, Kashiwa. ***Earthquake Research Institute, The University of Tokyo, Tokyo. ****Image Processing Research Team, RIKEN Center for Advanced Photonics, Wako. *****School of Engineering, Nagoya University, Nagoya)

Keywords: *materials integration, structural materials, integrated computational materials engineering, sparse modelling, data assimilation, database*

2019年6月17日受理[doi:10.2320/materia.58.503]

た数値シミュレーションモデルが既に手元にある場合に非常に有効な手段となる。逐次データ同化手法の概略を、代表的な手法であるアンサンブルカルマンフィルタ (Ensemble Kalman Filter, EnKF⁽⁴⁾) を用いて説明する。EnKF はモデルが非線形であっても適用できる上、並列化を含めた実装が比較的容易であるため、気象・海洋に限らず様々な分野で広く用いられている手法である。

逐次データ同化では、まず現象を説明する時間発展方程式 (システムモデル) が離散的に以下の形式で与えられるとする：

$$x_{t+1} = f(x_t, v_t)$$

ここで x_t は状態ベクトルと呼ばれ、システムモデルを記述するために必要な内部変数であり、例えば変位や圧力・温度などの状態量やモデルパラメータに対応する。また、 v_t はシステムモデルが現象を完全には表現できないことによる誤差である。また、一般に観測においては全ての状態ベクトル x_t を完全に計測することは不可能であることから、観測値 y_t は状態ベクトル x_t を入力とした観測演算子 $h(x_t)$ と観測ノイズ w_t により次式で与えられるとする：

$$y_t = h(x_t) + w_t$$

この様に考えると、例えば時刻 $t-1$ までの全ての観測 $y_{1:t-1}$ を用いて時刻 $t-1$ の状態ベクトル x_{t-1} の事前分布 $P(x_{t-1} | y_{1:t-1})$ が既に推定されていれば、時刻 t の状態ベクトル x_t の確率密度関数 $P(x_t | y_{1:t-1})$ はシステムモデルを用いて予測可能である。これを一期先予測と呼ぶ。この一期先予測を時刻 t に得られる観測 y_t と観測演算子 $h(x_t)$ を用いて、次式で与えられるベイズの定理により更新することで、時刻 t までの全ての観測 $y_{1:t}$ を用いて推定される状態ベクトル x_t の事後分布 $P(x_t | y_{1:t})$ が得られる：

$$P(x_t | y_{1:t}) = \frac{P(y_t | x_t) P(x_t | y_{1:t-1})}{P(y_t)}$$

以上を繰り返すことで、過去に観察された全ての観測値 $y_{1:t}$ を有効に活用し、時刻 t の状態ベクトル x_t の分布を得ようとするのが逐次データ同化手法である。

EnKF では状態ベクトル x_t の確率密度関数 $P(x_t | y_{1:t})$ を、離散的な状態ベクトル $x_t^{(i)}$ の集合として次式の様にアンサンブル近似する：

$$P(x_t | y_{1:t}) = \frac{1}{N} \sum_{i=1}^N \delta(x_t - x_t^{(i)})$$

これにより、時刻 t での離散的な一期先予測 $x_{t-1}^{(i)}$ は個々の状態ベクトル $x_{t-1}^{(i)}$ を既存の数値シミュレーションモデルに適用することで次式の様に簡単に求まる：

$$x_{t-1}^{(i)} = f(x_{t-1}^{(i)}, v_{t-1})$$

その結果、一期先予測 x_t の確率分布 $P(x_t | y_{1:t-1})$ は、以下のように近似的に推定できるのである：

$$P(x_t | y_{1:t-1}) = \frac{1}{N} \sum_{i=1}^N \delta(x_t - x_{t-1}^{(i)})$$

EnKF では更に観測演算子が次式の様な線形方程式で与えられるとし：

$$y_t = H_t x_t + w_t$$

更に観測ノイズが正規分布で与えられると仮定することで線形化されたベイズの定理を用いて、一期先予測 $x_{t-1}^{(i)}$ を次式で更新し、時刻 t の予測 $x_t^{(i)}$ を得ている：

$$x_t^{(i)} = x_{t-1}^{(i)} + \hat{K}_t (y_t + \hat{w}_t^{(i)} - H_t x_{t-1}^{(i)})$$

ここで、各項は全て単純な一次方程式で与えられる。MI 統合システムではこのような EnKF をモジュールとして導入しており、任意の数値シミュレーションモデルに対し逐次データ同化を適用することが可能になっている。

(b) 4次元変分法(アジョイント法)

4次元変分法⁽⁵⁾は EnKF とは異なり、既存の数値シミュレーションモデルをそのまま用いることはできない。しかし、少ない計算資源で効率的にデータ同化できるため、より大規模なシミュレーションモデルに適用できるという特徴がある。

4次元変分法では、まずシステムモデルが次式の時間発展方程式で与えられるとする：

$$\frac{\partial x}{\partial t} - F(x) = 0, \quad x(0) = x_0$$

また、逐次データ同化手法と同様に観測演算子 $h(x)$ を通して観測値 $y(t)$ が得られるものとする。解くべき問題は、全ての観測値 $y(t)$ を説明する最も尤もらしい状態ベクトル $x(t)$ の推定である。つまり、評価関数 J を次式で定義すると：

$$J = \int_0^T \mathcal{J}(x, y) dt$$

システムモデルを拘束条件として得られる次式のラグランジュ関数 \mathcal{L} の最適化問題を解くこととなる：

$$\mathcal{L} = \int_0^T \mathcal{J}(x, y) dt + \int_0^T \lambda^\dagger \left(\frac{\partial x}{\partial t} - F(x) \right) dt$$

ここで λ^\dagger は λ の転置を表す (以下でも同様)。最適解は \mathcal{L} の第一変分 $\delta \mathcal{L}$ がゼロとなる点として得られる：

$$\begin{aligned} \delta \mathcal{L} &= \int_0^T (\nabla_x \mathcal{J})^\dagger \delta x dt + \int_0^T \delta \lambda^\dagger \left(\frac{\partial x}{\partial t} - F(x) \right) dt \\ &\quad + \int_0^T \lambda^\dagger \left(\frac{\partial \delta x}{\partial t} - (\nabla_x F) \delta x \right) dt \\ &= \int_0^T \delta \lambda^\dagger \left(\frac{\partial x}{\partial t} - F(x) \right) dt + \lambda^\dagger(T) \delta x(T) - \lambda^\dagger(0) \delta x(0) \\ &\quad + \int_0^T \left((\nabla_x \mathcal{J})^\dagger - \lambda^\dagger (\nabla_x F) - \frac{\partial \lambda^\dagger}{\partial t} \right) \delta x dt \end{aligned}$$

ここで、最適解周辺では任意の $\delta \lambda$, δx に対して $\delta \mathcal{L} = 0$ となる必要があるため、最終的には最適解は以下の条件を満たす $x(t)$, $\lambda(t)$ となる事が分かる：

$$\frac{\partial x}{\partial t} - F(x) = 0$$

$$\nabla_x \mathcal{J}(x, y) - (\nabla_x F(x))^\dagger \lambda - \frac{\partial \lambda}{\partial t} = 0$$

$$\lambda(0) = \lambda(T) = 0$$

最初の式は $x(t)$ の時間発展方程式である。2番目の式は $\lambda(t)$ の時間発展方程式であり、アジョイント方程式と呼ばれる。

3番目は $\lambda(t)$ の初期条件と終端条件である。これらを同時に

満たす $x(t)$ と $\lambda(t)$ の探索方法として、4次元変分法では勾配法を用いる。 $x(t)$ の時間発展方程式、アジョイント方程式、更には終端条件 $\lambda(T) = 0$ が常に成り立つとすると、第一変分 $\delta\mathcal{L}$ は次式の様に変形できる：

$$\delta\mathcal{L} = \delta J = (\nabla_{x_0} J)^\dagger \delta x_0 = \lambda^\dagger(0) \delta x_0$$

つまり、 $\lambda(T) = 0$ を終端条件としてアジョイント方程式を時間方向に逆に積分することで得られる $\lambda(0)$ は、初期条件 x_0 に対する評価関数の勾配 $\nabla_{x_0} J(x, y)$ に対応することが分かる。従って、まず適当な初期条件 x_0 に対し $x(t)$ の時間発展方程式を解き、得られた $x(t)$ を用いてアジョイント方程式を時間方向に逆に積分することで勾配 $\nabla_{x_0} J(x, y) = \lambda(0)$ を求める。そして最後に勾配 $\nabla_{x_0} J(x, y)$ を用いて初期条件 x_0 を更新する。以上を勾配 $\nabla_{x_0} J(x, y)$ がゼロになるまで繰り返すことで、最適な初期条件 $x(0) = x_0$ 、つまり $x(t)$ と $\lambda(t)$ を得ることが可能になるのである。

以上のように、4次元変分法は勾配法を用いて少ない更新回数で精度良く最適解を得る手法となっている。そのため、アンサンブル近似を用いる EnKF より遥かに少ない計算資源で実行できるという利点を持つ。また、その定式化は有限要素法等の変分法に基づく数値解析手法と親和性が高いことが分かる。一方で、従来の4次元変分法は最尤推定であるため、最適解の一点だけが求まる。しかし、上述のように不確定要素が多い構造材料の解析においては、最適解に内在する不確実性の明示化(UQ)は推定結果の信頼性を議論する上で不可欠となる。そのため、MI プロジェクトでは推定に含まれる不確実性を明確にするための拡張も行った⁽⁶⁾⁽⁷⁾。紙面の都合上、詳細は参考文献(6),(7)をご参照頂きたい。

(c) データ同化の適用事例

データ同化の適用例として、Al合金の再結晶挙動をデータ同化により調査した例を以下に示す⁽⁸⁾。図1は実際に強加工されたアルミ合金を時効処理したときの再結晶挙動を、SEM-EBSDを用いて観察した結果である。積層欠陥エネルギーの大きなAl合金の再結晶挙動は、Humphreysによりサブグレイン成長モデルで良く説明できることが示されている⁽⁹⁾。そこで、ここではシステムモデルとしてMulti-Phase-Field(MPF)モデルを用いた。また、界面のモビリティと界面エネルギーに関しては、結晶方位依存性を考慮したモデルで良く説明できることが知られている⁽⁹⁾⁻⁽¹¹⁾。一方で、加工により導入されたひずみエネルギーは結晶方位毎に異なるが、従来はボトムアップ的なモデルを構築する事例が数多く報告されている。しかし、Humphreysも指摘している様に⁽¹²⁾、そのモデル化は模索段階と言える。そこでこ

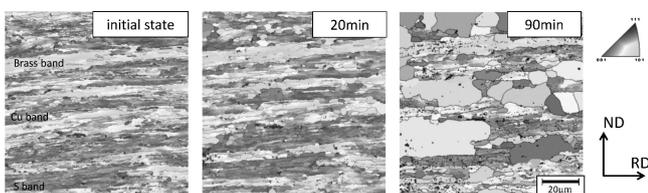


図1 Al合金の再結晶挙動(A1050, 573 K)。

では、逆に実験データを用いてひずみエネルギーをその時間推移を含めて推定できないかという観点で解析を行った。

図2にEnKFを用いて推定された結晶方位毎の体積率とひずみエネルギーの推移を示す。この解析では、観測値として結晶方位毎の体積率の変化を用いている。このような解析から、例えばS方位バンド内にはBrass方位バンド内よりも大きなひずみエネルギーが蓄えられていることや、Sバンド内に点在するCu方位粒には非常に小さなひずみエネルギーしか蓄えられていないことなどが読み取れる。また、Cu方位粒とS方位バンド内に蓄えられたひずみエネルギーの差は $10^{14} \sim 10^{15} / \text{m}^2$ 程度の転位密度に対応しており、強加工されたAl合金の転位密度と良く対応していることが分かる⁽¹³⁾。この例では集合組織の変化という限られた情報だけを用いているため、得られる情報も限られる。しかし、興味深い知見も得られており、将来的には表面構造写真の特微量抽出⁽¹⁴⁾やX線トモグラフィー等のより高度な測定と併用することで、より詳細な知見が得られると期待している。

(2) スパースモデリング

構造材料の特性データベースは、多くの場合でデータを得るためのコストが甚大なため、現在までに多大な労力が費やされてきたにも関わらず、蓄積された情報量には限りがある。そのため、多量のデータを前提とする機械学習手法の適用は現実的ではない場合が多い。これに対しスパースモデリングは、限られた情報量から現象を的確に表現しうる重要なパラメータを自動的に抽出する情報科学的モデリング技術であり、情報量が限られる構造材料とは極めて相性が良い手法と言える。また、データに含まれる本質を自動的に抽出するスパースモデリングは、未解明な現象が多く残されている構造材料の発展にも大きく寄与する可能性を秘めている。

スパースモデリングをベイズの定理に基づき定式化すると以下のようになる。ベイズの定理によると、観察 y を得た後のあるモデル M_i のモデルパラメータ x の事後分布 $P(x|y, M_i)$ は以下の式で与えられる：

$$P(x|y, M_i) \propto P(y|x, M_i)P(x|M_i)$$

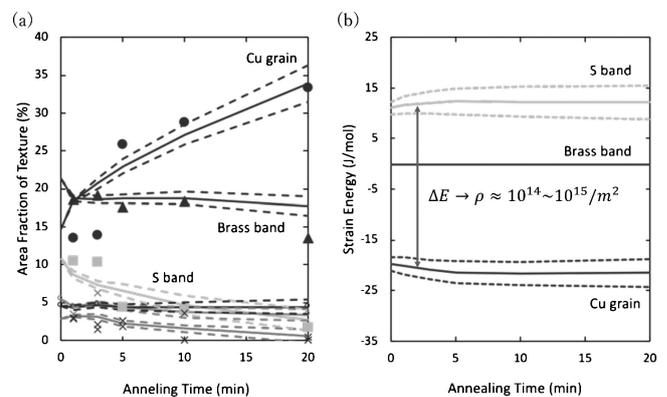


図2 EnKFによる同化結果。(a)体積率変化、(b)ひずみエネルギー。

さらに、現象を説明するモデル候補が有限個あるとき、個々のモデル M_i の事後分布は同様に次式で与えられる：

$$P(M_i|y) \propto P(y|M_i)P(M_i)$$

ここで、 $P(M_i)$ は個々のモデルの事前分布であり、事前知識としてモデルの優先度がない場合は全てのモデルで同じ値となる。つまり、一般にはモデルの事後分布 $P(M_i|y)$ は尤度 $P(y|M_i)$ に比例することが分かる。ここで、尤度 $P(y|M_i)$ は周辺尤度やエビデンスと呼ばれ、モデルパラメータ x で周辺化することで次式により得られる：

$$P(y|M_i) = \int P(y|x, M_i)P(x|M_i) dx$$

なお、周辺尤度の \ln を取って負にしたものをベイジック的自由エネルギー (Bayesian Free Energy, BFE) と呼ぶ。つまり、個々のモデル M_i に対し、観察 y の周辺尤度または BFE が計算できれば、どのモデルが現象を説明するにふさわしいモデルなのか議論できるのである。しかし、周辺尤度や BFE の計算は一般には難しく、ガウス分布を仮定することで解析的に導出する方法や、マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo, MCMC 法) を用いた数値積分が用いられる。

ここでは、重回帰分析における変数選択と複数の物理モデルからのモデル選択する場合の方法を紹介する。

(a) 重回帰分析の変数選択

得られるデータ量に制限がある中で、対象とする現象の説明変数の候補が多数あるとき、単純な重回帰分析は過学習を引き起こす。そのため、多数の説明変数の候補の中から有効な説明変数を選択する必要がある。変数選択問題として古くから様々な方法が提案されている。例えば、Pearson の相関係数 (Pearson's coefficient correlation)⁽¹⁵⁾ 等は古くから用いられている手法であるが、これは説明変数と出力の 2 つの変数間の直線関係の程度を表す指標に過ぎない。そのため、それだけでは最適な変数の組み合わせは教えてくれない。そのため、近年では LASSO (Least Absolute Shrinkage and Selection Operator) と呼ばれる L1 正則化手法⁽¹⁶⁾ が広く用いられる様になっている。LASSO は基本的には観測 y に対し評価関数 $J(x, y)$ が与えられたとき、説明変数の係数 x に対し $\sum |x| \leq t$ となる制約条件を付加することで、説明変数の次元を圧縮する手法である。このような LASSO により選ばれた説明変数の組み合わせは、確かにある超パラメータ λ に対する最適解ではある。しかし、超パラメータ λ によって最適となる組み合わせは異なり、また選ばれた最適解が他の組み合わせに対しどの程度の優位性があるのかという情報は示されない。多くの現象が複雑に絡み合う構造材料では、唯一絶対のモデルが存在し、それだけで現象が完全に表現できるという考えは受け入れにくい。そのため、MI 統合システムでは、線形重回帰分析に対しては LASSO に代わり ES-LiR (Exhaustive Search for Linear Regression)⁽¹⁷⁾⁽¹⁸⁾ と呼ばれる手法を利用したモジュールが導入されている。ES-LiR は線形重回帰分析に特化して BFE を網羅的に計算することで、最適解だけでなく最適解近傍の状態密度に関する情報も

明らかにする手法である。紙面の都合上、定式化の詳細は参考文献(17)(18)を参照されたい。

ES-LiR の適用事例として、鋼のマルテンサイト変態開始点 (Ms 点) の推定式を扱った例を示す⁽¹⁹⁾。鋼の Ms 点は炭素をはじめとする合金元素の濃度に強く依存することが知られ、古くから多くの推定式が提案されており⁽²⁰⁾⁻⁽²²⁾、その一例を以下に示す⁽²²⁾。

$$Ms = 512 - 453C + 15Cr - 16.9Ni - 9.5Mo + 217C^2 - 71.5CMn - 67.6CCr$$

このような推定式に含まれる合金元素の組み合わせは、個々の提案者の勘と経験に基づくものであり、多くの場合は諸々の事情にあわせて都合の良い推定式が用いられる。この問題に対し ES-LiR を適用し、BFE の全状態を明らかにした結果を図 3 に示す。なお、データには 1965 年以前の鋼が網羅的に調査された Andrews のデータを採用した⁽²²⁾。図 3 (a) から分かるように、Andrews の推定式は BFE の観点からは他の研究者の推定式より優秀ではあるものの、必ずしも最適な推定式ではない。一方で、図 3 (b) からは最適解自体も他の推定式に対し十分な優位性を示すわけではないことも分かる。このような場合、BFE を用いて全ての説明因子の重要度を評価した方が有用な情報が得られる (図 4)。図 4 (a) は上位 10 位までのモデルにおいて選択された説明因子を白色、選択されなかったものを黒色で示している。また、図 4 (b) は各説明因子に対し全モデルの確率を積算した値として重要度を示している。例えば、BFE を最適にする推定式だけからは、「W は必要だが Mo と C の交差項は不要」とな

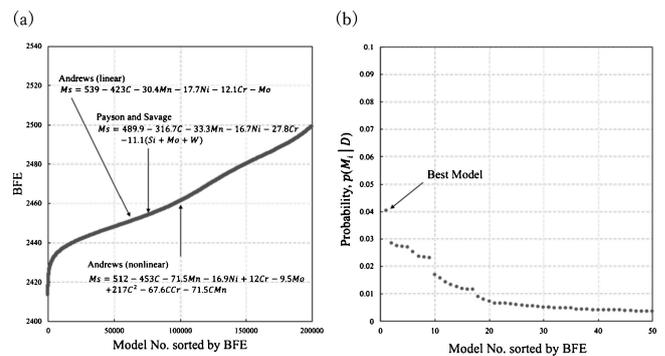


図 3 BFE の分布と推定式ならびに最適解。

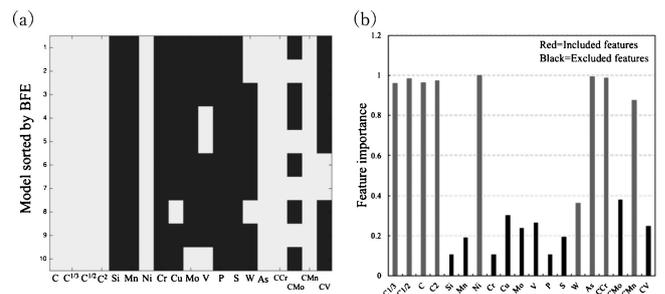


図 4 Top10 の変数選択と変数重要度。

る(図4(a)). しかし, 重要度の観点からは「W と CMo はほぼ同等の重要性がある」ことが分かる(図4(b)). 単なる予測だけに興味があるのであれば, 最適解だけを扱うことに異論はない. しかし, 選択された説明因子や背後の物理を真面目に考える上では, ES-LiR により得られる状態密度はより多くの示唆を与えるのである.

(b) 複数の提案モデルからのモデル選択

構造材料の問題においては, 一つの現象を説明するモデルとして複数のメカニズムが提案されることは珍しくない. また, いずれのメカニズムが実際の現象を支配しているのかといった判断は, 個々の研究者の経験と勘に基づいているといっても過言ではない. しかし, その様な判断が真っ当に出来るまでには長年に渡る経験が不可欠であり, 結局は個々の研究者の暗黙知のレベルで蓄積された知の深さが重要となる. しかし, その様な暗黙知の伝承は難しく, 経験と勘だけに頼らない方法論の確立が必要である. ここでは, 提案された複数のメカニズムに対し BFE を求め, 最も尤もらしいモデルの選択が可能か検証した例を示す.

一般に, BFE はモデルが非線形な場合は解析的に求めることは困難である. そのため, MCMC 法を用いた近似的な推定手法が広く用いられる(23). 中でも, 交換 MCMC 法は, 事後分布に多峰性があっても局所解に陥りにくいことや, 比較的安価に BFE の導出ができるため(24), MI システムで採用されている手法である. 交換 MCMC 法は, 逆温度 β_i を定義し, 個々の逆温度 β_i における事後分布を $P(y|x)^{\beta}P(x)$ とすることで, 異なる逆温度間での交換を従来の MCMC 法に加えて行う手法である. この場合, BFE は次式で与えられる:

$$BFE = -\ln \int P(y|x, M_i)P(x|M_i) dx$$

$$= -\sum_j \ln \langle \exp(-(\beta_{j+1} - \beta_j) \log P(y|x, M_i)) \rangle_{\beta_j}$$

ここで, $\langle \blacksquare \rangle_{\beta_j}$ は逆温度 β_j の事後分布における \blacksquare の平均を意味する. この様な導出は一見煩雑に見えるが, モンテカルロ法においては一般に平均値の収束は早く, 逆温度 $\beta=1$ の精緻な事後分布を得ることで BFE を求めるより, 高速に BFE を求めることができるのである.

交換 MCMC 法の適用事例として, 鋼の変態挙動を表す様々なモデルから実験データを用いて実際の変態挙動を最も尤もらしく説明するモデルを選択した例を示す(25). 図5(a)は鋼(Fe-0.15C-1.5Mn)を1400°Cから0.3°C/sで冷却した時の熱膨張測定により導出されたフェライト変態率の推移である. ここでの問題は, この様な変態率の推移だけからフェライト変態のキネティクスが議論できるのかということである. この問題の場合, 形成組織に関する事前知識が全く無い中で想定されるフェライト形態は3つある. 一つは非整合界面だけで構成されるレンズ上の粒界フェライト(LF), もう一つは整合界面と非整合界面で構成される Pill-box 型と呼ばれる粒界フェライト(PF), そして最後はフェライトサイドプレートと呼ばれる板状のフェライト(FSP)である. ま

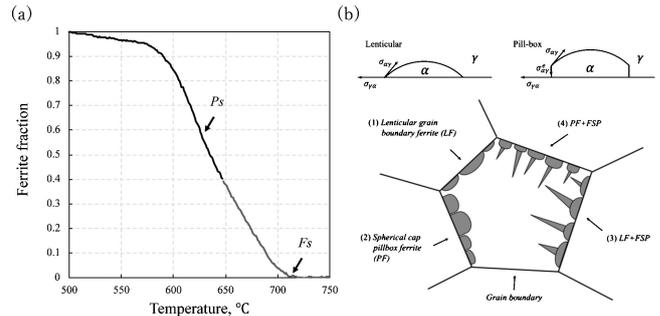


図5 フェライト変態率の推移(1400°C, 0.3°C/s)とフェライト変態モデル.

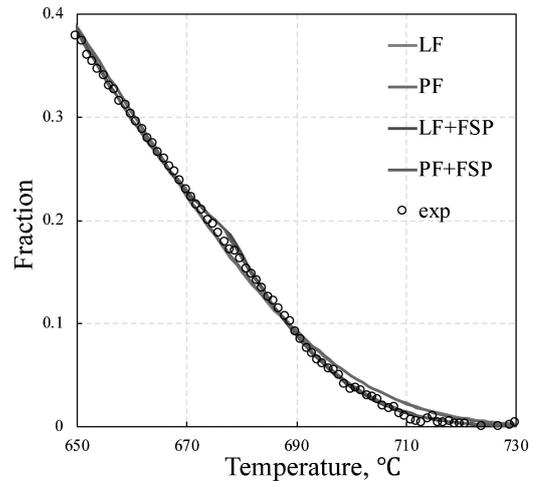


図6 フェライト変態率の最尤推定結果.

表1 フェライトモデルの推定誤差と BFE.

| Model class | Fitness | Complexity | BFE | Probability |
|-------------|---------|------------|-----|------------------------|
| LF | 112 | 18 | 130 | 4.09×10^{-35} |
| PF | 94 | 11 | 105 | 3.53×10^{-24} |
| LF + FSP | 42 | 24 | 66 | 3.05×10^{-7} |
| PF + FSP | 34 | 17 | 51 | 99.99 |

た, 観察される組み合わせとしては LF, PF がそれぞれ単独で形成される場合, LF の後に FSP が形成される場合, 更には PF の後に FSP が形成される場合の計4通りの組み合わせがある(図5(b)). なお, 個々のモデルに対する変態率の変化は Johnson-Mehr-Avrami-Kolmogorov (JMAK) 方程式を用いて与えられる(26)(27). 図6は, 全てのモデルのモデルパラメータを最尤推定により求め, 推定誤差が小さかった上位3つのモデルの推定結果と実験結果を比較した図である. いずれのモデルも推定誤差は十分に小さいことが分かり, これだけでは最適なモデル選択は難しい. しかし一方で, 表1に示す様に BFE からは, PF+FSP モデルが他のモデルよりも遥かに大きな尤度を示すことが分かる. 実際, 冶金学的観察からは Pill-box 型の粒界フェライトが形成された後に FSP が形成されることが観察される. また, その

遷移温度も PF + FSP モデルと一致しており、BFE を用いた推定結果の妥当性を支持する結果となっている⁽²⁵⁾。

3. 3D/4D 分析システム

3D/4D 分析システムでは、巨大な材料組織の 3 次元情報を得るための画像データを効率的に扱うクラウドシステムである Material Image Communication Cloud (MICC)⁽²⁸⁾と、材料組織の 3 次元構造を活用した特性予測を実際に体験出来る Materials Integration Phase Analyzer (MIPHA)⁽²⁹⁾の開発を行った。

(1) MICC

構造材料の特性を考える場合、材料組織の 3 次元構造の詳細を明らかにすることは重要と考えられている。しかし、そのためには 2 次元の断面組織画像を集約し、3 次元画像を構築する 3 次元画像処理環境が不可欠だが、その環境を手にするには高いコストや専門的な知識を必要とする。そのため、従来の材料組織の解析は主に 2 次元の断面組織画像が利用されているに過ぎなかった。また、たとえ材料組織の 3 次元画像が構築できたとしても、複雑な材料の幾何形状を抽出するには材料の組織形成に関する深い知識が必要であり、長年の経験を備えた専門家でも簡単ではない。そのため、知識や経験に頼らずに試行錯誤の中から適切な画像処理手法を組み合わせ見出すシステムが必要とされている。

MICC は、クラウド上で動作する 3 次元画像処理環境であり、ネットワークを通してアクセスすることで、あたかも手元の端末で大規模な画像処理を実行しているように見せることができるシステムになっている(図 7)。また、研究室や研究者ごとに画像データへのアクセス権限も管理でき、利用者間でのデータ共有の仕組みも備えている。そのため、複数の研究機関の連携や共同研究の効率化を支援できる構成となっている。さらに画像処理の履歴やその管理機能を有しており、過去に行ったすべての画像処理の履歴を記録することが可能である(図 8)。このような履歴を利用することで、将来的には多くの研究者による無数の試行錯誤の中から最適な画像処理方法を見出すことができると期待されている。

MICC の適用例として、鋼の溶接部への適用事例を紹介する。図 9(a)はシリアルセクションングにより取得した鋼組織の光学顕微鏡写真を再構成することで得られた材料の 3 次元組織を、様々な方向で切り出した断面画像を表示している様子を示している。この様に任意断面の材料組織を抽出することで、例えば旧オーステナイト粒界やマルテンサイト組織のブロックやパケットといった階層的な 3 次元構造を容易に抽出することが可能になるのである。また、図 9(b)はこの様にして抽出したブロックの 3 次元構造を示している。この例では専門家が様々な断面における組織形態から組

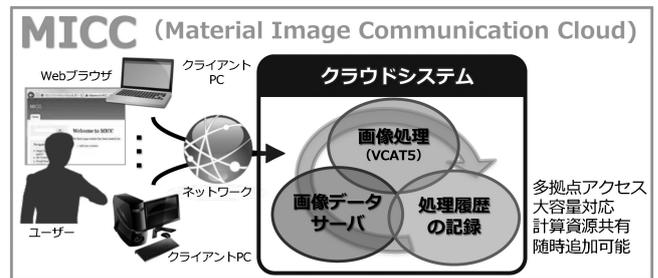


図 7 MICC の概略。

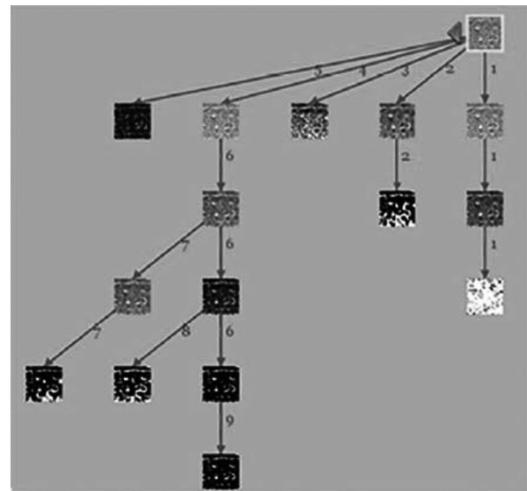


図 8 画像処理履歴のデータベース化。

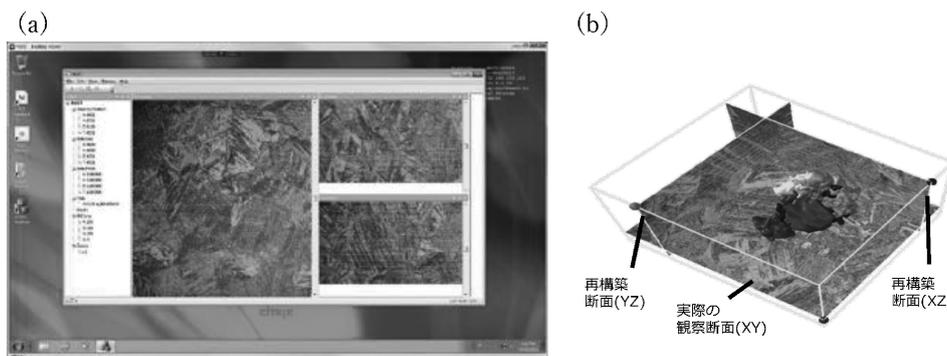


図 9 鋼組織の 3 次元組織再構成と組織抽出。

織境界の抽出を行っているが、この様な抽出をする過程の作業履歴を全て蓄積したデータを学習することで、シリアルセクションングにより得られた光学顕微鏡写真から自動的に組織判別や抽出が可能になると期待される。

(2) MIPHA

前述のように「材料組織の3次元幾何構造は特性の予測に重要である筈だ」と材料屋なら誰も想像している。しかし、実際に材料組織の3次元情報の集め方や利活用方法、更にはその効果となると必ずしも自明ではない。そこでMI統合システムが稼働する以前に、材料組織の3次元情報の扱い方や3次元の組織情報の効果をデモンストレーションする必要があり、そのプロトタイプとしてMIPHAを先行して開発を行った。

MIPHAはMI統合システムが目指した枠組みと同様に、材料組織の3次元情報を扱う上で不可欠な要素である、①組織識別、②2次元特徴量抽出、③3次元特徴量抽出、④特性推定、⑤逆解析を個々のモジュールとして実行できる統合型システムとなっている。特に画像の重要な特徴量としては、材料工学的に重要な特徴量(粒径、体積率など)と数学・画像工学的に重要な特徴量(自己相関関数、相互情報量、二点相関関数の距離、パーシステントホモロジーなど)があるが、MIPHAはワンクリックで前者の特徴量を画像から抽出するシステムとなっている。また、開発が続けられているshinyMIPHAでは後者の特徴量を抽出する。更に、MIPHAではこれらの特徴量を用いて、簡単に精度良く特性

の予測ができることが理解できるツールとなっている。

MIPHAの適用事例として、DP鋼の特性を推定した例を示す。図10はシリアルセクションングにより得られた光学顕微鏡写真から、機械学習によりフェライトやマルテンサイトを自動認識し、その情報を利用して3次元的な組織を再構成した例を示している。また、図11には応力-ひずみ曲線に影響する材料組織因子を示すが、ここに含まれる多くの幾何学的記述子をMIPHAでは抽出可能になっており、これらを様々な機械学習手法を適用することで、簡単に応力-ひずみ曲線が推定できることが体験できる⁽²⁹⁾。

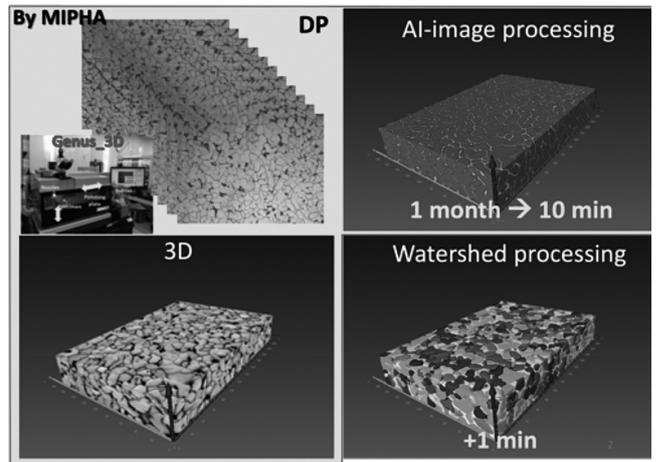


図10 DP鋼組織の3次元組織再構成と組織抽出。

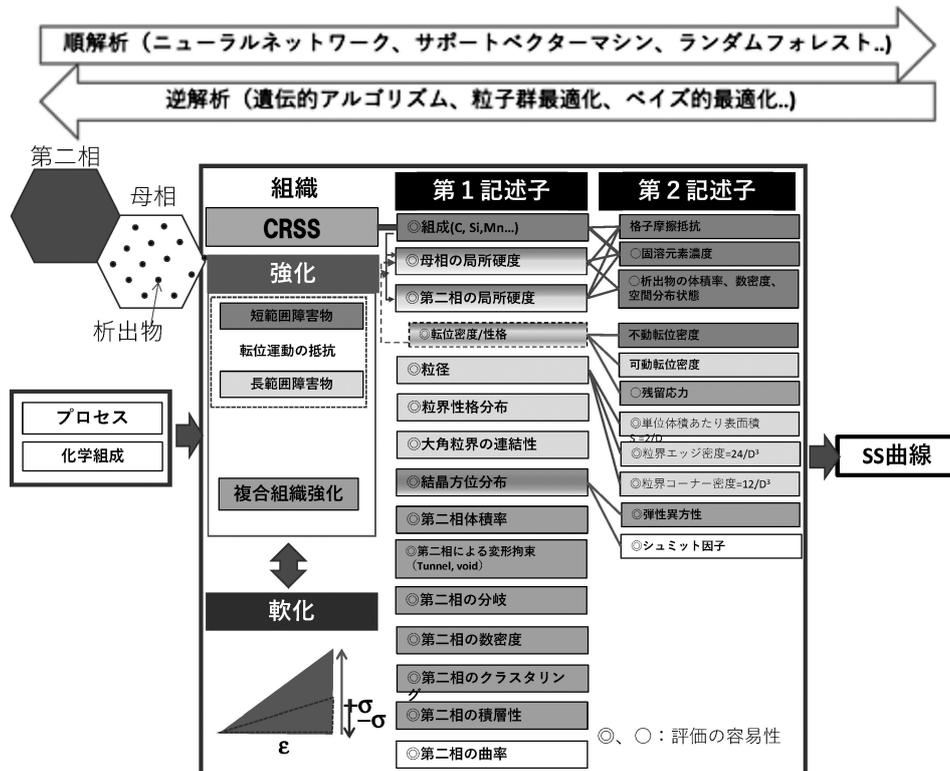


図11 MIPHAで利用可能な記述子と順解析/逆解析。

