

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年3月12日(12.03.2020)



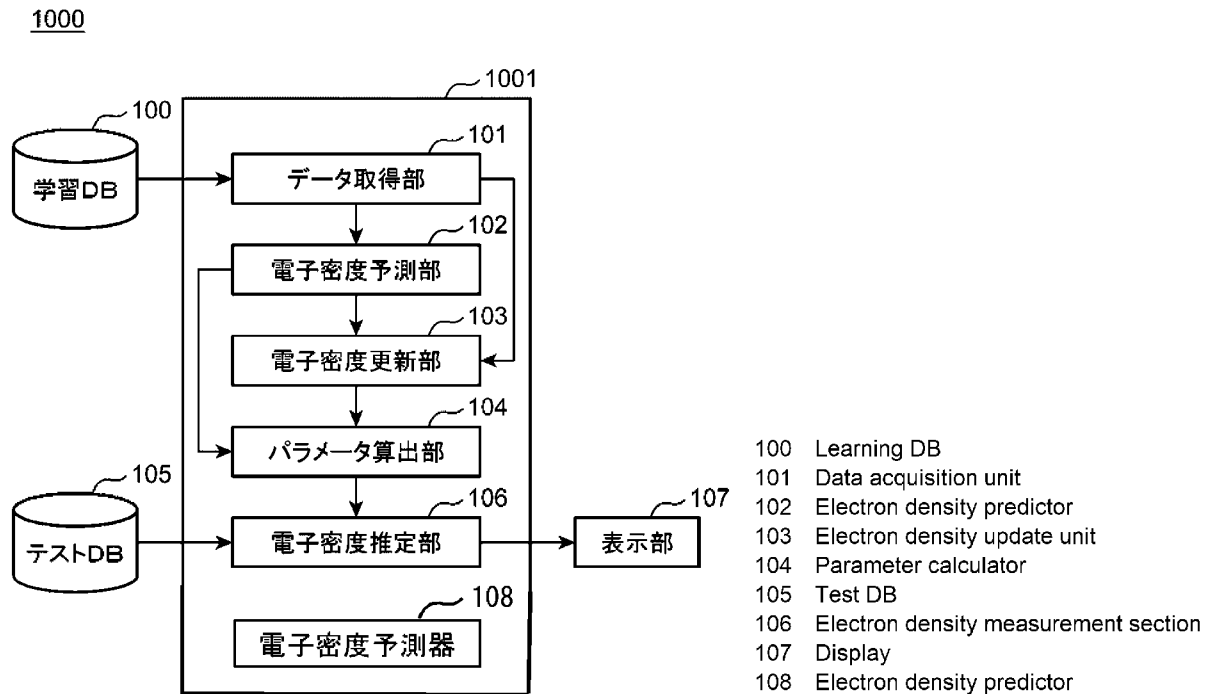
(10) 国際公開番号

WO 2020/049994 A1

- (51) 国際特許分類:
G16C 10/00 (2019.01) *G06N 20/00* (2019.01)
- (21) 国際出願番号: PCT/JP2019/032350
- (22) 国際出願日: 2019年8月20日(20.08.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-164752 2018年9月3日(03.09.2018) JP
- (71) 出願人: パナソニック IP マネジメント株式会社(PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) [JP/JP]; 〒5406207 大阪府大阪市中心区城見2丁目1番61号 Osaka (JP).
- (72) 発明者: 若杉 健介(WAKASUGI Kensuke). 森川 幸治(MORIKAWA Koji).
- (74) 代理人: 鎌田 健司, 外(KAMATA Kenji et al.); 〒5406207 大阪府大阪市中心区城見2丁目1番61号 パナソニック IP マネジメント株式会社内 Osaka (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,

(54) **Title:** ELECTRON DENSITY ESTIMATION METHOD, ELECTRON DENSITY ESTIMATION APPARATUS, AND ELECTRON DENSITY ESTIMATION PROGRAM

(54) 発明の名称: 電子密度推定方法、電子密度推定装置、及び電子密度推定プログラム



(57) **Abstract:** This electron density estimation method comprises: calculating a first electron density by inputting first input data to an electron density predictor; calculating a second electron density by executing a numerical simulation using the first input data and the first electron density, in which the numerical simulation is a process for setting the first input data and the first electron density to initial values and performing at least once a process of updating the electron

NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

一 国際調査報告（条約第21条(3)）

density using the density functional theory, and the second electron density is not a convergence value obtained by using the density functional method; training the electron density predictor by calculating parameters of the electron density predictor, which minimize a difference between the first electron density and the second electron density; acquiring second input data from a test database; inputting the second input data to the electron density predictor in which the calculated parameters are set; estimating a third electron density; and outputting the estimated third electron density.

(57) 要約：電子密度推定方法は、第1入力データを電子密度予測器に入力することで第1電子密度を算出し、第1入力データと第1電子密度とを用いた数値シミュレーションを実行して第2電子密度を算出し、数値シミュレーションは、第1入力データと第1電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を1回以上行う処理であり、前記第2電子密度は、前記密度汎関数法を用いて求めた収束値ではなく、第1電子密度と第2電子密度との差分を最小化する電子密度予測器のパラメータを算出することで前記電子密度予測器を学習させ、テストデータベースから第2入力データを取得し、算出されたパラメータが設定された電子密度予測器に前記第2入力データを入力し、第3電子密度を推定し、推定された前記第3電子密度を出力する。

明 細 書

発明の名称：

電子密度推定方法、電子密度推定装置、及び電子密度推定プログラム

技術分野

[0001] 本開示は、物質の電子密度を推定する技術に関するものである。

背景技術

[0002] 従来、物理法則に基づき、物質中の電子密度を更新する計算を反復し、電子密度の収束値を求める密度汎関数法と呼ばれる技術が開発されている（例えば、非特許文献１）。また、密度汎関数法を用いて触媒を開発する技術も知られている（例えば、特許文献１）。このように密度汎関数法は、新規な物質を開発するためのツールとして利用されている。

[0003] また、近年では、機械学習で得られた予測器を用いて電子密度を予測する技術も開発されている。例えば、非特許文献２は、物質中のポテンシャルから電子密度を推定する方法を開示する。具体的には、非特許文献２では、特定の物質について、複数の原子間距離及び角度を変えて密度汎関数法で電子密度の収束値を算出し、算出した電子密度の収束値を用いて電子密度を予測する予測器を学習させ、得られた予測器を用いて、あらゆる原子間距離及び角度に対する電子密度を予測する技術が開示されている。

[0004] 非特許文献３は、物質中のポテンシャルから電子のエネルギーを推定する方法を開示する。具体的には、非特許文献３では、二次元のランダムポテンシャル又は井戸型ポテンシャルから、シュレディンガー方程式を解いて得られる電子のエネルギーをディープニューラルネットワークを用いて推定する技術が開示されている。

先行技術文献

特許文献

[0005] 特許文献１：特許第６０１５３３６号公報

非特許文献

[0006] 非特許文献1: Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical review*, 140 (4A), A1133.

非特許文献2: Brockherde, F., Vogt, L., Li, L., Tuckerman, M. E., Burke, K., & Mueller, K. R. (2017). Bypassing the Kohn-Sham equations with machine learning. *Nature communications*, 8 (1), 872.

非特許文献3: Mills, K., Spanner, M., & Tamllyn, I. (2017). Deep learning and the Schroedinger equation. *Physical Review A*, 96 (4), 042113.

非特許文献4: Finger, L. W., & Hazen, R. M. (1978). Crystal structure and compression of ruby to 46 kbar. *Journal of Applied Physics*, 49 (12), 5823-5826.

非特許文献5: Cicek, oe., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016, October). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 424-432)

. Springer, Cham.

非特許文献6: Bergerhoff, G., Brown, I. D., & Allen, F. (1987). Crystallographic databases. International Union of Crystallography, Chester, 360, 77-95.

発明の概要

[0007] しかし、上記の従来技術は予測器の学習データとして電子密度の収束値が必要なため、更なる改善の必要がある。

[0008] 本開示は、上記の課題を解決するものであって、その目的は、電子密度の収束値を用いずに電子密度予測器を学習させる技術を提供することである。

[0009] 本開示の一態様に係る電子密度推定方法は、物質の組成及び構造に関する物質情報から前記物質の電子密度を推定する電子密度推定装置における電子密度推定方法であって、

前記電子密度推定装置のコンピュータが、

(a1) 前記物質情報を格納した学習データベースから第1入力データを取得し、

(a2) 前記第1入力データを電子密度予測器に入力することで第1電子密度を算出し、

(a3) 前記第1入力データと前記第1電子密度とを用いた数値シミュレーションを実行して第2電子密度を算出し、前記数値シミュレーションは、前記第1入力データと前記第1電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を1回以上行う処理であり、前記第2電子密度は、前記密度汎関数法を用いて求めた収束値ではなく、

(a4) 前記第1電子密度と前記第2電子密度との差分を最小化する前記電子密度予測器のパラメータを算出することで前記電子密度予測器を学習させ、

(a5) 前記物質情報を格納したテストデータベースから第2入力データ

を取得し、前記パラメータが設定された前記電子密度予測器に前記第2入力データを入力し、第3電子密度を推定し、

(a6) 推定された前記第3電子密度を出力する。

[0010] 本開示によれば、密度汎関数法で求めた電子密度の収束値を使わずに、電子密度予測器の出力値が密度汎関数法で求めた電子密度の収束値に近づくように電子密度予測器を学習させることができる。

図面の簡単な説明

[0011] [図1]電子密度の具体的な値の一例を示した図

[図2]密度汎関数法において、電子密度を更新する計算を反復することにより、電子密度が更新される様子を模式的に示した図

[図3]電子密度予測器を学習する際に算出される差分を示した図

[図4]電子密度予測器で得られた予測値を密度汎関数法における初期電子密度に設定し、電子密度を更新する計算を反復することにより、電子密度が更新される様子を模式的に示した図

[図5]予測値を用いて、密度汎関数法における電子密度を更新する計算を一回行って更新値を求め、そして、求めた更新値を教師データとして電子密度予測器を学習させることについて示した模式図

[図6]CIF10と呼ばれる自然画像データセットを用いたクラス分類問題に関して、学習データ数を変更した場合の学習の推移について示した図

[図7]本開示の実施の形態に係る電子密度推定システムの構成を示すブロック図

[図8]図7の電子密度更新部の詳細な構成を示すブロック図

[図9]CIFと呼ばれる形式で記載された第1入力データの一例を示す図

[図10]POSCARと呼ばれる形式で記載された第1入力データの一例を示す図

[図11]記述子の一例を示す図

[図12]初期電子密度、第1電子密度、及び電子密度予測器の一例を示す図

[図13]第1電子密度から第2電子密度が算出される処理の一例を示す図

[図14]第3 電子密度が推定される処理の一例を示す図

[図15]本開示の実施の形態1における電子密度推定装置の処理の全体的な流れを示すフローチャート

[図16]図15のステップS102における電子密度更新部の処理の詳細を説明するフローチャート

[図17]本開示の実施の形態2の処理を説明する図

[図18]本開示の実施の形態2の処理を説明する図

[図19]本開示の実施の形態2に係る処理の一例を示すフローチャート

[図20]A₁O₃に関する単位格子ベクトル、原子座標、原子番号、原子半径などを示す図

[図21]第1入力データと原子番号または原子半径とを組み合わせで表現したA₁結晶の記述子の一例などを示す図

発明を実施するための形態

[0012] (本開示に至った知見)

新規の材料開発を行う際、バンドギャップ及び形成エネルギー等の物質の特性値を予測する必要があるが、そのためには物質中の電子密度を計算する必要があるため、非特許文献1に示す密度汎関数法は、新規の材料開発の有効なツールとして従来から利用されている（例えば、特許文献1）。

[0013] 一方、近年、深層ニューラルネットワーク、カーネルリッジ回帰、及びサポートベクターマシンなどの機械学習手法によって、物質のポテンシャルから電子密度及び電子のエネルギーを予測する手法が提案されている。この手法では、密度汎関数法で求めた電子密度の収束値あるいは電子のエネルギーの収束値を用意し、収束値と予測値との誤差を最小化するように予測器のパラメータの更新を行ない、予測器を学習させる。

[0014] 例えば、非特許文献2には、特定の構造を持つ物質について、密度汎関数法で算出した電子密度の収束値をとして用いて電子密度予測器を学習させる技術が開示されている。

[0015] しかし、非特許文献2では、例えばペロブスカイト構造というような特定

の構造を持つ物質についてのみ予測器を学習させているため、学習した物質とは構造の異なる未知の物質については正確な電子密度を予測できないという課題がある。

[0016] また、密度汎関数法で収束値を得るには、現状の計算機の能力では1物質だけでも3週間程度はかかると言われており、これでは学習に必要な多数の物質の収束値を得るためには数年程度を要してしまう。そのため、密度汎関数法で収束値を得る非特許文献2の手法では、教師データの収集に膨大な時間がかかるという課題がある。

[0017] 非特許文献3では、シュレディンガー方程式を解いて得られる電子エネルギーをディープニューラルネットワークを用いて推定する技術が開示されている。

[0018] しかし、非特許文献3では、ディープニューラルネットワークの学習に大規模な学習データを必要とするため、学習データの用意に多大な時間を要するという課題がある。また、非特許文献3では密度汎関数法で求めた収束値を使わずにディープニューラルネットワークを学習させることについての開示はない。

[0019] なお、非特許文献1は、密度汎関数法の基本的な技術を開示する先行技術文献であり、機械学習で予測器を学習させることに関する開示はない。また、特許文献1は、900℃におけるアンカー粒子と貴金属粒子との比較結合エネルギーを算出する際に非特許文献1に示す密度汎関数法を利用することが開示されているに過ぎず、機械学習で予測器を学習させることに関する開示はない。更に、非特許文献4、5は、後述する技術を開示する先行技術文献である。

[0020] そこで、本発明者らは、密度汎関数法における電子密度の更新値を求める計算と機械学習による電子密度予測器のパラメータ更新を勘案し、電子密度の更新値を求める計算処理を電子密度予測器の学習処理に組み込むことで、密度汎関数法で求めた電子密度の収束値を使わずに、電子密度予測器の出力値が密度汎関数法で求めた電子密度の収束値に近づくように電子密度予測器

を学習させることができることを見出し、本開示を想到するに至った。

[0021] すなわち、本開示は、密度汎関数法で求めた電子密度の収束値を使わずに、電子密度予測器の出力値が密度汎関数法で求めた収束値に近づくように電子密度予測器を学習させることができる技術を提供することを目的とする。

[0022] 本開示の一態様に係る電子密度推定方法は、物質の組成及び構造に関する物質情報から前記物質の電子密度を推定する電子密度推定装置における電子密度推定方法であって、

前記電子密度推定装置のコンピュータが、

(a 1) 前記物質情報を格納した学習データベースから第1入力データを取得し、

(a 2) 前記第1入力データを電子密度予測器に入力することで第1電子密度を算出し、

(a 3) 前記第1入力データと前記第1電子密度とを用いた数値シミュレーションを実行して第2電子密度を算出し、前記数値シミュレーションは、前記第1入力データと前記第1電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を1回以上行う処理であり、前記第2電子密度は、前記密度汎関数法を用いて求めた収束値ではなく、

(a 4) 前記第1電子密度と前記第2電子密度との差分を最小化する前記電子密度予測器のパラメータを算出することで前記電子密度予測器を学習させ、

(a 5) 前記物質情報を格納したテストデータベースから第2入力データを取得し、前記パラメータが設定された前記電子密度予測器に前記第2入力データを入力し、第3電子密度を推定し、

(a 6) 推定された前記第3電子密度を出力する。

[0023] 本構成によれば、第1入力データと第1電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を1回以上行う数値シミュレーションが実行される。そして、この数値シミュレーションによって得られた第2電子密度と第1電子密度との差分が最小化されるように電子密度予測器

が学習されている。第2電子密度は、密度汎関数法を用いて求めた収束値ではない。このように、本構成は、数値シミュレーションにより得られた更新値である第2電子密度を用いて電子密度予測器が学習されている。そのため、本構成は、収束値を使わずに電子密度予測器を学習させることができる。その結果、膨大な時間及び労力をかけることなく学習データを準備することができ、電子密度予測器の生成に要するコストの低減を図ることができる。

[0024] 上記構成において、前記(a5)は、前記(a1)から前記(a4)までの処理が所定回数繰り返された後に実施されてもよい。

[0025] 本構成によれば、数値シミュレーションにより第2電子密度を更新しながら、電子密度予測器を学習させることができるため、数値シミュレーションと電子密度予測器の学習とを効率良く行うことができる。

[0026] 上記構成において、前記(a3)では、

(b1) 前記(a2)で更新された前記第1電子密度が所定の条件を満たす場合、前記数値シミュレーションを実施して前記第2電子密度を更新し、処理を前記(a4)に進め、

(b2) 前記(a2)で更新された前記第1電子密度が前記所定の条件を満たさない場合、前記第2電子密度を更新せずに、処理を前記(a4)に進めてもよい。

[0027] 本構成によれば、第1電子密度が更新される度に第2電子密度を更新させるのではなく、第1電子密度が第2電子密度にある程度近づいてから第2電子密度が更新されるので、処理コストのかかる数値シミュレーションの実行回数を抑制させて、効率良く電子密度予測器を学習させることができる。

[0028] 上記構成において、前記所定の条件は、前記(a2)で更新された前記第1電子密度と第4電子密度との第2差分が、前記(a2)で更新された前記第1電子密度と第5電子密度との第3差分よりも大きくなることであり、

前記第4電子密度は、前記(b1)で更新された前記第2電子密度の算出に用いられた前記第1電子密度であり、

前記第5電子密度は、前記(b1)で更新された前記第2電子密度であつ

てもよい。

[0029] 本構成によれば、更新された第1電子密度が、第2電子密度の更新に用いられた第1電子密度である第4電子密度よりも更新された第2電子密度である第5電子密度に近づいたことを条件に第2電子密度を更新させることができる。そのため、数値シミュレーションの実行回数を抑制させ、電子密度予測器を効率良く学習させることができる。

[0030] 上記構成において、前記物質情報は、結晶学共通データフォーマットを用いて記載された情報であってもよい。

[0031] 本構成によれば、既存のデータフォーマットを用いて物質情報が記載されているため、第1入力データのコンピュータへの実装が容易になる。

[0032] 上記構成において、前記(a4)では、前記第2電子密度が教師データとして用いられて前記電子密度予測器が学習されてもよい。

[0033] 本構成によれば、第1電子密度が第2電子密度に近づくように電子密度予測器を学習させることができる。

[0034] 上記構成において、前記(a6)では、前記第3電子密度を示す画像がディスプレイに表示されてもよい。

[0035] 本構成によれば、第3電子密度を視覚的にユーザに認識させることができる。

[0036] 上記の電子密度推定方法の作用効果は電子密度推定装置及び電子密度推定プログラムにおいても同様に実現される。

[0037] (実施の形態1)

まず、本開示の電子密度推定装置が密度汎関数法で求めた電子密度の収束値を使わずに電子密度予測器を学習させることが可能である理由について説明する。図1は、電子密度の具体的な値の一例を示した図である。密度汎関数法によって物質の単位格子中の電子密度を求める場合、単位格子が $24 \times 24 \times 24$ の3次元の複数のセルに区切られ、各セルに電子密度を示す実数値が保持される。ここで、単位格子は、例えば、体心立方格子または面心立方格子であり、これらの単位格子が $24 \times 24 \times 24$ の3次元のセルに区切

られる。空間中の電子密度を求める場合、空間の各点が1つのセルで表され、各セルが電荷を保持していると考え、各セルの電子密度の大きさが求められる。なお、単位格子を区切る際の各次元のセルの個数は24に限定されず、24以外の任意の値が採用されてもよい。また、単位格子を区切る際の各次元のセルの個数は、次元ごとに異なる値が採用されてもよい。

[0038] 図2は、密度汎関数法において、電子密度を更新する計算を反復することにより、電子密度が更新される様子を模式的に示した図である。図2において、縦軸はパラメータYを示し、横軸はパラメータXを示す。図2では、上述した $24 \times 24 \times 24$ のセルのうちある2個のセルにおける電子密度が、電子密度を更新する計算を反復することにより電子密度が更新される様子が示されている。パラメータXは2個のうち一方のセルの電子密度を示し、パラメータYは2個のセルのうち他方のセルの電子密度を示している。したがって、図2のグラフにおいて各点は、2個のセルの電子密度の集合で構成される。なお、図1の例では、セルの個数は $24 \times 24 \times 24$ 個であるため、この場合、図2のグラフにおいて各点は $24 \times 24 \times 24$ 個の電子密度の集合で構成される。これらのことは、図3、図4、図5でも同じである。

[0039] 初期値500と初期値501とのそれぞれは、密度汎関数法における初期電子密度を表す。収束値502は、密度汎関数法によって最終的に得られる電子密度を表す。すなわち、収束値502は、密度汎関数法における電子密度を更新する計算を反復して電子密度を収束させた値である。図2に示すように、密度汎関数法は、理論上、電子密度を更新する計算を反復する際の電子密度の初期値が異なっても、同じ収束値502に初期値500、501を収束させることができる。なお、図2の手法は、密度汎関数法において電子密度を更新する計算を反復して電子密度を収束させる、つまり、密度汎関数法で収束値502を算出する手法であるため、非特許文献1の手法に相当する。

[0040] 図3は、電子密度予測器を学習する際に算出される差分を示した図である。予測値600は電子密度予測器の出力を表す。差分601は収束値502

と予測値600との差分を表す。差分の計算には、例えばRMSE (Root Mean Square Error) が用いられる。学習時において、予測値600と収束値502との差分601を最小化するように、すなわち予測値600が収束値502に近づくように電子密度予測器のパラメータが更新される。なお、図3の手法は、密度汎関数法により電子密度予測器の収束値502が算出されており、非特許文献2、3の手法に相当しているため、本開示の手法とは異なる。

[0041] 図4は、電子密度予測器で得られた予測値600を密度汎関数法における初期電子密度に設定し、電子密度を更新する計算を反復することにより、電子密度が更新される様子を模式的に示した図である。更新値700は、予測値600を用いて、密度汎関数法における電子密度を更新する計算を一回行うことで得られた電子密度である。

[0042] 図2に示したように密度汎関数法では、異なる初期値500, 501を用いた場合でも、初期値500, 501を同じ収束値502に収束させることができる。よって、図4に示す予測値600を密度汎関数法における初期電子密度として用いても、密度汎関数法における電子密度を更新する計算を反復することで予測値600を収束値502に収束させることができる。図4の手法は、電子密度予測器で得られた予測値600を密度汎関数法における初期値電子密度として用い、密度汎関数法における電子密度を更新する計算を反復することで収束値502を得る手法である。したがって、図4の手法は密度汎関数法で収束値502を算出している点で本開示の手法とは異なる。

[0043] 図5は、予測値600を用いて、密度汎関数法における電子密度を更新する計算を一回行って更新値700を求め、そして、求めた更新値700を教師データとして電子密度予測器を学習させることについて示した模式図である。更新値700は、密度汎関数法において電子密度を更新する計算を一回行うことで得られた電子密度であり、密度汎関数法で求めた電子密度の収束値ではない。なお、図5は本開示の手法に相当する。差分800は更新値7

00と予測値600との差分である。差分800を最小化するように電子密度予測器のパラメータを更新することで、予測値600が更新値700に近づくように電子密度予測器を学習させることができる。

[0044] 更新値700は密度汎関数法によって予測値600を収束値502に向けて更新させる値である。予測値600が更新値700に近づくように電子密度予測器を学習させることで、電子密度予測器は、収束値502により近い予測値を出力できるようになる。

[0045] このように、図5の手法は、図2～図4の手法とは異なり、密度汎関数法が用いられているものの、密度汎関数法によって収束値502を算出しておらず、更新値700を算出するに留めている。そして、この更新値700を教師データとして電子密度予測器を学習させること、及び、密度汎関数法における電子密度を更新する計算を繰り返しながら、予測値600が徐々に収束値502に近づくように電子密度予測器の学習が行われる。そのため、収束値502を用いずに電子密度予測器を学習させることができる。その結果、密度汎関数法における電子密度を更新する計算を行う回数が少なくなるので、計算コストの削減を図りつつ、電子密度を正確に算出できる電子密度予測器を生成できる。

[0046] 次に、本開示の手法が、密度汎関数法によって得られた収束値を用いて電子密度予測器を学習させる比較例の手法に比べ、予測器の学習完了までの所要時間が短縮される理由を、式を用いて説明する。

[0047] <比較例の所要時間 T_{α} >

まず、比較例の所要時間 T_{α} について説明する。ここで、学習データ数を P 、学習時における1つの学習データの使用回数を Q 、密度汎関数法により収束値を求める場合の電子密度を更新する計算回数を R とおく。また、電子密度を更新する計算1回の所要時間を T 、1つの学習データあたりの電子密度予測器の学習に要する所要時間を L とおく。この場合、密度汎関数法で求めた収束値を学習データとして用意し、その後に電子密度予測器の学習を行なう場合の所要時間 T_{α} は以下のように計算される。すなわち、この比較例

は、P種類の物質のそれぞれについて密度汎関数法によるR回の電子密度の更新計算を行ってP個の学習データを取得し、P個の学習データのそれぞれについて電子密度予測器の学習をQ回行うものである。

[0048] 所要時間 $T_{\alpha} = T \times R \times P + L \times Q \times P$

右辺第1項の $T \times R \times P$ は、密度汎関数法に要する所要時間である。電子密度の更新計算1回の所要時間がT、電子密度の更新計算回数がRであるため、1つの学習データの計算に要する所要時間は $T \times R$ となる。よって、学習データ数Pの全ての計算に要する所要時間は $T \times R \times P$ となる。

[0049] 密度汎関数法では、下記に示す式(1)、(2)に示すコーンシャム方程式を解くことで電子密度が算出される。なお、式(1)、(2)は非特許文献1の式(2.8)、(2.9)に示されている。

[0050] [数1]

$$\left\{ -\frac{1}{2} \nabla^2 + [\varphi(\mathbf{r}) + \mu_{xc}(n(\mathbf{r}))] \right\} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (1)$$

$$n(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 \quad (2)$$

[0051] 計算対象となる物質の原子配置及び周期構造に基づき式(1)の左辺が定義され、式(1)が解かれることで波動関数 $\psi_i(\mathbf{r})$ が求められる。なお、式(1)の左辺の

[0052] [数11]

$$\varphi(\mathbf{r})$$

[0053] 及び $\mu_{xc}(n(\mathbf{r}))$ はそれぞれ、物質に応じた関数になる。また、式(1)の右辺の ϵ_i は波動関数 $\psi_i(\mathbf{r})$ に対応する軌道エネルギーであり、式(1)の方程式を解くことで得られる。

[0054] 電子密度 $n(\mathbf{r})$ は、式(2)により波動関数 $\psi_i(\mathbf{r})$ から算出される。しかしながら、式(1)左辺の定義には電子密度 $n(\mathbf{r})$ が必要であり、このままでは、式(1)を解くことができない。そこで、電子密度 $n(\mathbf{r})$ に適当な初期値が与えられ、与えられた電子密度の初期値を式(1)に代入

して、式（１）、（２）が解かれることで電子密度が更新される。そして、更新された電子密度を用いて再度式（１）の左辺が定義され、電子密度が更新される。このような電子密度を更新する計算を電子密度 $n(r)$ が収束するまで繰り返され、計算対象となる物質の原子配置及び周期構造に対応した電子密度が求められる。

[0055] 「密度汎関数法における電子密度を１回更新する計算」とは、「既知の電子密度と式（１）から $\phi_i(r)$ を求め、求めた $\phi_i(r)$ と式（２）とから新たな電子密度を求める」ことをいう。すなわち、「既知の電子密度を用いて、式（１）を１回、式（２）を１回使い、新たな電子密度を求める」ことである。

[0056] 所要時間 T_α の第２項の $L \times Q \times P$ は電子密度予測器の学習に要する所要時間である。学習において、１つの学習データの使用回数が Q 、１つの学習データの学習に要する所要時間が L なので、１つの学習データの学習の所要時間は $L \times Q$ となる。よって、 P 個の学習データの学習に要する所要時間は $L \times Q \times P$ となる。

[0057] <本開示の手法による所要時間 T_β >

以下の本開示の手法は、 P 種類の物質のそれぞれについて密度汎関数法における電子密度の更新計算を１回実施して P 個の学習データを取得する処理と、 P 個の学習データのそれぞれについて学習器を１回学習させる処理とからなる処理セットを、 Q 回繰り返すものである。すなわち、１個の学習データにつき、１回の電子密度の更新計算と１回の学習処理が行われるため、電子密度の更新計算回数 R と学習データの使用回数 Q とは等しくなる。

[0058] したがって、本開示の手法による所要時間 T_β は所要時間 T_α において $R = Q$ としたときに相当し、以下のように計算される。

[0059] 所要時間 $T_\beta = (T + L) \times Q \times P$

ここでは、１つの学習データについて電子密度予測器の学習を行なう際に、密度汎関数法による電子密度の更新計算が１回実施される。よって、１つの学習データについて電子密度予測器の学習に要する所要時間は、電子密度

の更新計算1回の所要時間 T と、1つの学習データの学習に要する所要時間 L との和($T+L$)となる。したがって、 P 個の学習データのそれぞれを Q 回学習させたときの所要時間は($T+L$) $\times Q \times P$ となる。

[0060] <所要時間 T_{α} と所要時間 T_{β} との比較>

所要時間 T_{α} から所要時間 T_{β} を引くと、下記の式が得られる。

[0061] $T \times (R - Q) \times P$

よって、密度汎関数法における電子密度の更新計算回数 R が電子密度予測器の学習データの使用回数 Q よりも大きい場合、 $T \times (R - Q) \times P$ が正となり、本開示の所要時間 T_{β} の方が比較例の所要時間 T_{α} よりも短くなる。

[0062] 図6は、C i f a r 1 0と呼ばれる自然画像データセットを用いたクラス分類問題に関して、学習データ数を変更した場合の学習の推移について示した図である。図6において、縦軸は正解値と予測値との差分を表し、この差分が収束した場合に電子密度予測器の学習が終了される。横軸は学習回数を表している。図6においては、5000、10000、20000、30000、40000、50000のデータサイズに対応する6つのグラフを用いて差分が収束していく様子が示されている。ここでは、データサイズは学習データの個数を示すものとする。

[0063] 図6に示されるように、学習におけるデータサイズを5000から50000まで変更しても収束までに必要な学習回数に大差がないことが分かる。例えば、データサイズが50000の学習において、学習回数が10000 $\times$ 100回で収束したとする。この場合、学習回数は、1つの学習データの使用回数 Q と学習データ数 P との積で表されるので、使用回数 Q は下記の式により20と算出される。

[0064] $Q = 10000 \times 100 / P = 20$

また、学習データ数 P が増えても収束に必要な学習回数が増えないことから、上式において、収束に必要な学習回数(10000 $\times$ 100)は一定とみなすことができる。したがって、上式において、学習データ数 P を増やした場合、収束に要する使用回数 Q は小さくできる。

[0065] 一方、電子密度の更新計算回数 R は、対象となる物質にもよるが、平均して 40 回程度の電子密度の更新計算を必要とする。また、電子密度の更新計算回数 R は、より高精度な計算をするほど増大させる必要がある。

[0066] 以上のことから、少なくとも 50000 個以上の学習データを用いる場合には、 $R = 40$ 、 $Q = 20$ により、 $T \times (R - Q) \times P$ が正になるため、本開示の手法の方が、比較例の手法に比べて所要時間を短くすることができる。

[0067] (実施の形態 1 : 構成の説明)

続いて、実施の形態 1 の構成について説明する。以下、本開示の実施の形態 1 に係る電子密度推定システム 1000 について、図面を用いて詳細に説明する。図 7 は、本開示の実施の形態に係る電子密度推定システム 1000 の構成を示すブロック図である。図 7 に示す電子密度推定システム 1000 は、電子密度推定装置 1001 と、学習データベース (DB) 100 と、テストデータベース (DB) 105 と、表示部 107 とを備える。

[0068] 電子密度推定装置 1001 は、データ取得部 101 と、電子密度予測部 102 と、電子密度更新部 103 と、パラメータ算出部 104 と、電子密度推定部 106 と、電子密度予測器 108 を備える。なお、図 7 において、電子密度推定装置 1001 は、CPU 等のプロセッサがコンピュータを電子密度推定装置 1001 として機能させる電子密度推定プログラムを実行することで実現される。また、図 7 において、学習データベース 100 及びテストデータベース 105 は、不揮発性のメモリで構成されている。

[0069] 以下、図 7 に示した各構成要素の詳細について述べる。

[0070] (学習データベース 100)

学習データベース 100 は、1 以上の物質のそれぞれの組成及び構造等の物質情報を含む第 1 入力データを予め記憶する。第 1 入力データは、例えば、結晶学共通データフォーマット (CIF) で記載された物質情報を含む。但し、物質情報の記載形式は、CIF のデータフォーマットに限定されず、組成式、結晶構造、空間群、及び格子ベクトルといった密度汎関数法におけ

る電子密度を更新する計算が実施可能な記載形式であれば、どのような記載形式であってもよい。

[0071] 図9は、CIFと呼ばれる形式で記載された第1入力データの一例を示す図である。CIFでは、組成式、単位格子ベクトルの長さ、原子同士が交わる角度、及び単位格子中の原子配置などで物質情報が記述される。図9では、物質「 Al_2O_3 」に関する物質情報が示されている。

[0072] 図10は、POSCARと呼ばれる形式で記載された第1入力データの一例を示す図である。POSCARでは、組成式、単位格子ベクトル、及び単位格子中の原子配置等で物質情報が記述される。なお、図9に示す記載内容は非特許文献4及び非特許文献6から引用した。

[0073] (データ取得部101)

図7に参照を戻す。データ取得部101は、学習データベース100から第1入力データを取得し、電子密度予測部102と電子密度更新部103とに出力する。

[0074] (電子密度予測部102)

電子密度予測部102は、データ取得部101から第1入力データを取得し、第1入力データから記述子を算出する。そして、電子密度予測部102は、算出した記述子を電子密度予測器108に送付する。電子密度予測器108は、電子密度予測器108の入力層に記述子を入力し、電子密度予測器108は第1電子密度を算出し、電子密度予測器108の出力層から第1電子密度を出力する。電子密度予測器108は、第1電子密度を電子密度予測部102に送付する。電子密度予測部102は第1電子密度を電子密度更新部103及びパラメータ算出部104に出力する。

[0075] 記述子は、第1入力データから算出可能なベクトル、或いは2次元または3次元行列で表現される。記述子は、構成元素の第1入力データと元素の物性値とを組み合わせで表現されてもよい。或いは、記述子は、非特許文献2のFig. 1、非特許文献3のFig. 1に示されているように、ポテンシャルで構成されていてもよい。或いは、記述子は、密度汎関数法における初

期電子密度で構成されてもよい。

[0076] 第1入力データと元素の物性値とを組み合わせる表現した記述子の一例を Al_2O_3 について以下に示す。図20は Al_2O_3 に関する単位格子ベクトル、原子座標、原子番号、原子半径などを示す図である。 Al_2O_3 の記述子は、 Al_2O_3 の単位格子ベクトル、Al原子の相対座標、O原子の相対座標、AlとOそれぞれの原子番号、原子半径を列挙した1次元ベクトルである。

[0077] 例えば、(Al_2O_3 記述子) = (4.805027 0 0 -2.40251 4.161275 0 0 0 13.11625 0 0 0.352096 8 146.2014 0 0 0.647904 8 146.2014 ...)である。当該ベクトルに含まる要素の並び順は任意に決定してよい。

[0078] 図21は、第1入力データと原子番号または原子半径とを組み合わせる表現したAl結晶の記述子の一例などを示す図である。例えば、Al結晶であれば、面心立方格子の原子座標を中心に電子密度がガウス分布状に存在すると考え、下記の式で示されるように、 $24 \times 24 \times 24$ 個のセルの値が計算される。

[0079] [数12]

$$\text{各セルの値} = \sum_{d < d'} \frac{\alpha}{\sqrt{2\pi}\sigma} \exp\left(-\frac{d^2}{2\sigma^2}\right).$$

ただし、 d = セルと原子の距離、 α = 原子番号、 $\sigma = \frac{\text{原子半径}}{3}$ 、 d' = 閾値

[0080] 各セルの中心点と各原子との距離に応じて減衰する値に、各原子の原子番号を乗じ、予め定められた半径以下に含まれる原子について和を取る。ガウス分布の σ には原子半径/3を用いればよい。予め定められた半径には、例えば、単位格子の対角線の内、最長の値を採用すればよい。なお、原子番号の代わりに、各原子の電気陰性度やイオン化エネルギーを用いても良い。

[0081] 図12は、初期電子密度、第1電子密度、及び電子密度予測器の一例を示す図である。

[0082] 図1では $24 \times 24 \times 24$ 個のセルの電子密度を表現している。図1は2

4枚の画像を示し、24枚の画像の各々は 24×24 の画素を示していると考えてもよい。そして、 24×24 の画素はそれぞれ画素値を有し、1つの画素の画素値が1セルの電子密度に対応すると考えてもよい。

[0083] 図12の電子密度予測器108の入力層のユニット数は $24 \times 24 \times 24$ であってもよい。図12の初期電子密度1401として示された24枚の画像の画素（個数= $24 \times 24 \times 24$ ）の画像値が電子密度予測器108の入力層のユニットに入力される。初期電子密度1401は電子密度予測部102が生成した記述子である。

[0084] 初期電子密度1401として示された24枚の画像の画素と電子密度予測器108の入力層のユニットは1対1に対応する。

[0085] 図12の電子密度予測器108の出力層のユニット数は $24 \times 24 \times 24$ であってもよい。図12の第1電子密度1403として示された24枚の画像の画素（個数= $24 \times 24 \times 24$ ）の画像値が電子密度予測器108の出力層のユニットから入力される。

[0086] 第1電子密度1403として示された24枚の画像の画素と電子密度予測器108の出力層のユニットは1対1に対応する。

[0087] 図12において、白い箇所は電子密度が高い箇所を示している。このことは、後述する図13、図14も同じである。初期電子密度1401は、電子を原子核の近傍に存在すると仮定した場合の電子密度である。ここでは、一例として、A1の単結晶について、 $24 \times 24 \times 24$ の各セルの電子密度を 24×24 の24枚の画像で表現したときの初期電子密度1401が示されている。

[0088] （電子密度予測器108）

電子密度予測器108は、初期電子密度1401（電子密度予測部102が生成した記述子）が入力され、第1電子密度1403を出力する。この例では、電子密度予測器108は、非特許文献5のFig. 2に示されるニューラルネットワークで構成されていてもよい。なお、電子密度予測器108は、 $24 \times 24 \times 24$ 次元の回帰問題を学習するランダムフォレスト又はサ

ポートベクターマシンで構成されてもよい。あるいは、図11の記述子と初期電子密度とのセットが電子密度予測器108に入力されてもよい。電子密度推定装置1001は電子密度予測器108を含まず、外部のサーバが電子密度予測器108を含んでもよい。この場合、電子密度推定装置1001と当該外部のサーバは有線及び／または無線でデータのやり取りを行ってもよい。

[0089] (電子密度更新部103)

図7に参照を戻す。電子密度更新部103は、データ取得部101から第1入力データを取得し、電子密度予測部102から第1電子密度を取得し、第1入力データと第1電子密度とを用いた数値シミュレーションを実行して第2電子密度を算出し、第2電子密度をパラメータ算出部104に出力する。

[0090] 数値シミュレーションは、第1入力データと第1電子密度とを用いて、密度汎関数法における電子密度を更新する計算を1回行う処理である。第2電子密度は、数値シミュレーションによって更新された電子密度である。

[0091] なお、当該数値シミュレーションは1回以上実施してもよい。つまり、まず、「第1入力データと第1電子密度とを用いて、密度汎関数法における電子密度を更新する計算を1回行い、第1の仮の2電子密度を求め」、次に「第1入力データと第1の仮の2電子密度とを用いて、密度汎関数法における電子密度を更新する計算を1回行い、第2の仮の2電子密度を求め」、～、最後に「第1入力データと第 $(n-1)$ の仮の2電子密度とを用いて、密度汎関数法における電子密度を更新する計算を1回行い、第 n の仮の2電子密度を求め」て、第 n の仮の2電子密度を、上述した第2電子密度としてもよい。電子密度の更新計算の回数は、密度汎関数法で電子密度の極限值を求める回数ではない。なお n は5以下であってもよい。

[0092] 当該数値シミュレーションを1回以上実施することで、数値シミュレーションで求める結果の安定性を向上させることが出来る。一般的な密度汎関数法では、電子密度の更新履歴を過去5回程度記録し、過去の電子密度と更新

後の電子密度とを適当な割合で足し合わせることで、数値シミュレーションの安定化を図っている。本開示の手法でも5回を上限に数値シミュレーションを1回以上行うことで、電子密度を更新する際の数値シミュレーションの安定化を図ることができる。

[0093] なお、当該数値シミュレーションを5回実施する際は、少なくとも250000個以上の学習データを用いる場合には、本開示の手法の方が、比較例の手法に比べて所要時間を短くすることができる。

[0094] 数値シミュレーションの回数は、数値シミュレーションで求める結果の安定性を図ることができる回数であり、かつ、数値シミュレーションで極限値を求める時間よりも短い時間で数値シミュレーションを完了できる回数であればよい。

[0095] 図13は、第1電子密度から第2電子密度が算出される処理の一例を示す図である。電子密度更新部103は、第1電子密度1403を式(1)に代入すると共に、第1入力データから得られる。

[0096] [数11]

$$\phi(r)$$

[0097] 及び $\mu_{xc}(n(r))$ を式(1)に代入することで、波動関数 $\phi_i(r)$ を求める。そして、電子密度更新部103は、求めた波動関数 $\phi_i(r)$ を式(2)に代入し、得られる電子密度 $n(r)$ を第2電子密度1501として算出する。式(1)には原子座標が含まれるため、第1電子密度1403のようなランダムな電子座標を初期値に設定しても、コーンシャム方程式を計算することで、原理的に、第2電子密度1501のように原子座標が反映された電子密度が得られる。

[0098] (パラメータ算出部104)

図7に参照を戻す。パラメータ算出部104は、電子密度予測部102から第1電子密度を取得し、電子密度更新部103から第2電子密度を取得する。そして、パラメータ算出部104は、第1電子密度と第2電子密度との

差分である第1差分を算出し、第1差分を最小とする電子密度予測器のパラメータを算出し、算出したパラメータを電子密度推定部106に出力する。

[0099] 第1差分としては、上述したRMSEが採用できる。但しこれは一例であり、MAE (Mean Absolute Error) 等の誤差を評価できる手法であればどのような手法を用いて第1差分が算出されてもよい。第1差分を最小とするパラメータの算出には例えばバックプロパゲーションが採用できる。バックプロパゲーションによるパラメータの更新式は、例えば、下記の式で表される。

$$[0100] \quad W_{t+1} = W_t - \gamma dD(W_t) / dW_t$$

W_t はパラメータの更新前の値を示し、 W_{t+1} はパラメータの更新後の値を示し、 $D(W_t)$ は第1差分を示し、 γ は学習係数を示す。学習係数の値としては、特に限定はないが、例えば0.01が採用できる。

[0101] 上記更新式によるパラメータ W_t の更新には、学習データベース100に格納された第1入力データの全てが用いられてもよい。あるいは、パラメータ算出部104は、学習データベース100から一部の第1入力データを取得し、取得した一部の第1入力データについて上記更新式を用いてパラメータ W_t を更新した後、再度、学習データベース100から別の一部の第1入力データを取得するというように、学習データベース100から取得する第1入力データを順次変更させながらパラメータ W_t の更新を繰り返してもよい。

[0102] パラメータ算出部104が、第1差分を最小とする電子密度予測器108のパラメータを算出することは、下記の処理(P1)～(P4)が実行されて、第kパラメータを特定することであると考えるてもよい。

[0103] (P1) 第1入力データに基づいて電子密度予測部102が生成した記述子が、第1パラメータが設定された電子密度予測器108に入力され、電子密度予測器108が第1電子密度 D_{11} を算出し、出力する。

[0104] (P2) 電子密度更新部103は、密度汎関数法を用いた電子密度を更新する計算を1回以上行い、第2電子密度を算出する。この1回以上の計算は

、第1入力データと第1電子密度 D_{11} に基づいた密度汎関数法を用いた計算を含む。第2電子密度は密度汎関数法を用いて求めた収束値ではない。

[0105] (P3) 上記生成された記述子が、第 i パラメータが設定された電子密度予測器108に入力され、電子密度予測器108が第1電子密度 D_{1i} を算出し、出力する。 $i = 2 \sim n$ であり、 i は自然数、 n は自然数、 n は2以上である。これにより、第1電子密度 $D_{12} \sim$ 第1電子密度 D_{1n} が得られる。

[0106] (P4) パラメータ算出部104は、第1電子密度 D_{1j} と2電子密度の差分である第1差分 DF_{1j} を算出する。 j は自然数である。これにより、第1差分 $DF_{11} \sim$ 第1差分 DF_{1n} が得られる。第1パラメータが第1差分 DF_{11} に対応し、 $\sim$ 第 n パラメータが第1差分 DF_{1n} に対応する。

[0107] パラメータ算出部104は、第1差分 $DF_{11} \sim$ 第1差分 DF_{1n} のうち、最も小さい第1差分 DF_{1k} を検出し、第1差分 DF_{1k} に対応する第 k パラメータを特定する。 $1 \leq k \leq n$ であり、 k は自然数である。以上で(P1) $\sim$ (P4)の説明を終える。

[0108] パラメータは以下に示すようなものであってもよい。

[0109] 電子密度予測器108は、入力層、1または複数の隠れ層、及び出力層を備えてもよい。入力層、1または複数の隠れ層、のそれぞれの層は、ニューロンに相当する複数のユニットを有する。出力層は1または複数のユニットを有する。

[0110] 隠れ層及び出力層おけるそれぞれのユニットに対して、重み値 $W_i = [w_{i1}, w_{i2}, \dots]$ が設定される。つまり、隠れ層及び出力層のユニット総数が z ならば、重み値 $W_1 = [w_{11}, w_{12}, \dots]$ 、 $\dots$ 、重み値 $W_z = [w_{z1}, w_{z2}, \dots]$ が存在する。入力層の各ユニットに入力データが入力されと、重み値 W_1 、 $\sim$ 、重み値 W_z を用いた演算がなされ、出力層の1または複数のユニットから出力データが出力される。

[0111] 上述した第1パラメータ、 $\sim$ 、第 n パラメータのそれぞれは重み値 W_1 、 $\sim$ 、重み値 W_z に対応する値を有する。つまり、第1パラメータは、重み値

$W1_1 = [w11_1, w12_1, \dots]$ 、 $\dots$ 、重み値 $Wz_1 = [wz1_1, wz2_1, \dots]$ 、 $\sim$ 、第 n パラメータは、重み値 $W1_n = [w11_n, w12_n, \dots]$ 、 $\dots$ 、重み値 $Wz_n = [wz1_n, wz2_n, \dots]$ である。

[0112] (テストデータベース 105)

テストデータベース 105 は、1 以上の物質のそれぞれの組成及び構造等の物質情報を含む第 2 入力データを予め記憶する。第 2 入力データは、例えば、結晶学共通データフォーマット (CIF) で物質情報が記載されている。但し、物質情報の記載形式は、CIF のデータフォーマットに限定されず、組成式、結晶構造、空間群、及び格子ベクトルといった密度汎関数法による計算が実施可能な記載形式であればどのような記載形式であってもよい。なお、第 2 入力データは、第 1 入力データと同様、図 9、図 10 に示すような内容で記載される。このように、第 2 入力データは第 1 入力データと記載形式は同一である。したがって、学習データベース 100 とテストデータベース 105 とは同一のデータベースで構築されてもよい。この場合、このデータベースに記憶された一部の入力データが第 1 入力データ、残りの入力データが第 2 入力データとして採用されてもよい。

[0113] (電子密度推定部 106)

電子密度推定部 106 は、パラメータ算出部 104 から更新後のパラメータを取得し、電子密度予測器 108 に取得したパラメータを設定する。そして、電子密度推定部 106 は、テストデータベース 105 から第 2 入力データを取得し、取得した第 2 入力データから記述子を算出する。そして、電子密度推定部 106 は、算出した記述子を電子密度予測器 108 に送付する。電子密度予測器 108 は、電子密度予測器 108 の入力層に記述子を入力し、電子密度予測器 108 は第 3 電子密度を算出し、電子密度予測器 108 の出力層から第 3 電子密度を出力する。電子密度予測器 108 は、第 3 電子密度を電子密度推定部 106 に送付する。電子密度推定部 106 は第 3 電子密度を表示部 107 に出力する。記述子と第 3 電子密度との算出方法は電子密

度予測部 102 における記述子と第 1 電子密度との算出方法と同じである。

第 3 電子密度の算出を、第 3 電子密度の推定と呼んでもよい。

[0114] 図 14 は、第 3 電子密度が推定される処理の一例を示す図である。初期電子密度 1601 は、第 2 入力データの記述子として得られた電子密度である。学習済みの電子密度予測器 108 に初期電子密度 1601 を入力することで、第 3 電子密度 1602 が得られる。第 3 電子密度 1602 は、初期電子密度 1601 とは異なり、各画像の中心および 4 隅に局在していた電子密度が減少し、代わりに中心および 4 隅の中間位置に非局在化した電子密度が表れており、正確に電子密度が推定されていることが分かる。

[0115] (表示部 107)

表示部 107 は、例えば、液晶パネル等の表示装置で構成され、電子密度推定部 106 が推定した第 3 電子密度を示す画像を表示する。表示部 107 は、第 3 電子密度を出力する出力部の一例である。第 3 電子密度を示す画像としては、例えば、図 14 に示す第 3 電子密度 1602 を表示する画像が採用されてもよいし、第 3 電子密度を示す数値群を表示する画像が採用されてもよい。

[0116] 図 8 は、図 7 の電子密度更新部 103 の詳細な構成を示すブロック図である。図 8 に示す電子密度更新部 103 は、シミュレーションパラメータ設定部 200 と、電子密度設定部 201 と、更新計算部 202 とを備える。

[0117] 以下、図 8 に示した各構成要素の詳細について述べる。

[0118] (シミュレーションパラメータ設定部 200)

シミュレーションパラメータ設定部 200 は、データ取得部 101 から第 1 入力データを取得し、数値シミュレーションの計算条件に設定する。これにより、式 (1) の例えば、 $\phi(r)$ 及び $\mu_{xc}(n(r))$ の関数が設定される。

[0119] (電子密度設定部 201)

電子密度設定部 201 は、電子密度予測部 102 から第 1 電子密度を取得し、数値シミュレーションの初期電子密度に設定する。これにより、式 (1

）における電子密度 $n(r)$ が設定される。

[0120] （更新計算部 202）

更新計算部 202 は、密度汎関数法における電子密度を更新する計算を 1 回以上行い、初期電子密度を更新し、第 2 電子密度を算出し、算出した第 2 電子密度をパラメータ算出部 104 に出力する。これにより、式（1）、式（2）を用いた 1 回以上の更新計算が行われ、図 13 に示すように、第 1 電子密度 1403 から第 2 電子密度 1501 が算出される。なお、更新計算の回数は、密度汎関数法で電子密度の極限值を求める回数ではない。

[0121] （実施の形態 1：動作の説明）

次に、電子密度推定装置 1001 の動作について説明する。

[0122] （フローチャート）

図 15 は、本開示の実施の形態 1 における電子密度推定装置 1001 の処理の全体的な流れを示すフローチャートである。

[0123] （ステップ S100）

データ取得部 101 は、学習データベース 100 から第 1 入力データを取得し、取得した第 1 入力データを電子密度予測部 102 と電子密度更新部 103 とに出力する。

[0124] （ステップ S101）

電子密度予測部 102 は、第 1 入力データから記述子を算出し、算出した記述子を電子密度予測器 108 に送付する。電子密度予測器 108 は第 1 電子密度を算出し、電子密度予測部 102 に送付する。密度を算出する。電子密度予測部 102 は、第 1 電子密度を電子密度更新部 103 に出力する。

[0125] （ステップ S102）

電子密度更新部 103 は、ステップ S100 で取得された第 1 入力データとステップ S101 で算出された第 1 電子密度とを用いて密度汎関数法による数値シミュレーションを行うことで第 2 電子密度を算出し、パラメータ算出部 104 に出力する。ここでは、電子密度更新部 103 は、ステップ S100～ステップ S104 の 1 つのループにおいて、数値シミュレーションの

計算を1回行うものとする。つまり、密度汎関数法における電子密度を更新する計算を1回行うものとする。この場合、数値シミュレーションの総回数は、ステップS100～ステップS104のループ数が k (k は1以上の整数)であれば、 k 回となる。更新計算の総回数 k は、密度汎関数法で電子密度の極限值を求める回数ではない。但し、これは一例であり、電子密度更新部103は、1つのループにおいて、数値シミュレーションの計算を m (m は複数)回実施してもよい。つまり、密度汎関数法における電子密度を更新する計算を m 回実施してもよい。この場合、数値シミュレーションの総回数は $k \times m$ となる。更新計算の総回数 $k \times m$ は、密度汎関数法で電子密度の極限值を求める回数ではない。

[0126] (ステップS103)

パラメータ算出部104は、第1電子密度と第2電子密度とから第1差分を算出し、第1差分が最小となるようにバックプロパゲーション等の手法を用いて電子密度予測器108のパラメータ W_t を更新する。

[0127] (ステップS104)

パラメータ算出部104は、パラメータの更新回数が所定の回数を超えた場合(ステップS104でYES)、更新されたパラメータを電子密度推定部106に出力し、処理をステップS105に進める。これにより電子密度予測器のパラメータ W_t を最小化する学習済みのパラメータが得られ、電子密度予測器の学習が終了される。

[0128] 一方、パラメータ算出部104は、パラメータの更新回数が所定の回数を超えていない場合(ステップS104でNO)、処理をステップS100に戻し、ステップS100からステップS103までの処理を繰り返す。ここで、所定の回数としては、図6で説明したように、電子密度予測器108の学習に用いられる第1入力データの数から第1差分が収束することが見込まれる予め定められた回数が採用できる。但し、これは一例であり、第1差分が所定の値以下に収束するまで、ステップS100からステップS103までの処理が繰り返されてもよい。

[0129] (ステップS105)

電子密度推定部106は、パラメータ算出部104から学習済みのパラメータを取得し、取得したパラメータを電子密度予測器108に設定する。そして、電子密度推定部106は、テストデータベース105から第2入力データを取得し、取得した第2入力データから記述子を算出し、算出した記述子を電子密度予測器108に送付する。電子密度予測器108は、第3電子密度を算出し、第3電子密度を電子密度推定部106に送付する。電子密度推定部106は第3電子密度を表示部107に出力する。

[0130] 図16は、図15のステップS102における電子密度更新部103の処理の詳細を説明するフローチャートである。

[0131] (ステップS200)

電子密度更新部103のシミュレーションパラメータ設定部200は、データ取得部101から取得した第1入力データを密度汎関数法の計算条件に設定する。ここで、計算条件としては、式(1)に示す $\phi(r)$ 及び μ_x 。 $(n(r))$ が挙げられ、これらの関数が、第1入力データから算出され、式(1)に設定される。

[0132] (ステップS201)

電子密度設定部201は、電子密度予測部102から取得した第1電子密度を密度汎関数法の初期電子密度に設定する。これにより、式(1)の電子密度 $n(r)$ の初期値として、電子密度予測器から得られた第1電子密度が設定され、式(1)の方程式を解くことが可能となり、波動関数 $\phi_i(r)$ 及び軌道エネルギー ε_i を算出することが可能になる。

[0133] (ステップS202)

更新計算部202は、密度汎関数法における電子密度を更新する計算を1回以上行い、初期電子密度に設定した第1電子密度を更新して第2電子密度を算出し、パラメータ算出部104に出力する。更新計算の回数は、密度汎関数法で電子密度の極限值を求める回数ではない。

[0134] 以上のように、電子密度更新部103はステップS200からステップS

202までの処理を行うことによって第2電子密度を算出する。

[0135] 以上説明したように、本実施の形態によれば、数値シミュレーションの1回以上の電子密度の更新計算により得られた更新値である第2電子密度を学習データとして用いて電子密度予測器を学習させることができる。更新計算の回数は、密度汎関数法で電子密度の極限值を求める回数ではない。そのため、本構成は、密度汎関数法で求めた電子密度の収束値を使わずに電子密度予測器を学習させることができる。その結果、膨大な時間及び労力をかけることなく学習データを準備することができ、電子密度予測器の生成に要するコストの低減を図ることができる。

[0136] (実施の形態2)

実施の形態2は、第1電子密度が第2電子密度にある程度近づくまで、数値シミュレーションをスキップさせ、第1電子密度が第2電子密度にある程度近づいた段階で第2電子密度を更新させるものである。なお、本実施の形態において、実施の形態1と同一の構成要素には同一の符号を付し、説明を省く。

[0137] 図17は、本開示の実施の形態2の処理を説明する図である。図17は、第2差分1703が第3差分1704よりも小さい状態での第1電子密度と第4電子密度と第5電子密度との距離関係を模式的に示している。なお、図17において、縦軸及び横軸は図2と同じである。

[0138] 第4電子密度1700は、最後に数値シミュレーションを実施した際に電子密度予測器が出力した第1電子密度である。第1電子密度1702は現在の電子密度予測器の出力である。第2差分1703は、第1電子密度1702と第4電子密度1700との差分であり、学習の進捗に対する電子密度予測器の出力の変動を表す。第3差分1704は第5電子密度と第1電子密度1702との差分であり、第2差分1703と同様、学習の進捗に対する電子密度予測器の出力の変動を表す。

[0139] 第5電子密度1701は、第4電子密度1700を初期電子密度とする数値シミュレーションによって更新された第2電子密度である。第5電子密度

1701は、第4電子密度1700から見て収束値1705側に位置している。そのため、第5電子密度1701は、第4電子密度1700から見たときの収束値1705のおよその方向を示す。よって、第5電子密度1701を教師データとして電子密度予測器を学習させることで、収束値1705に向けて電子密度予測器を学習させることができる。なお、収束値1702は、密度汎関数法において電子密度を更新する計算を繰り返して電子密度を収束させた値である。

[0140] 第2差分1703が第3差分1704より小さい状態では、第4電子密度1700から第5電子密度1701へ方向H1と、第1電子密度1702から第5電子密度1701へ方向H2とがおおよそ一致している。したがって、この状態では、第3差分1704を最小化するように電子密度予測器のパラメータを更新することで電子密度予測器の出力である第1電子密度1702を収束値1705に近づけることができる。このとき、本実施の形態では、数値シミュレーションがスキップされているため、計算時間を減らすことができる。

[0141] 図18は、本開示の実施の形態2の処理を説明する図である。図18は、図17において、第2差分1703が第3差分1704より大きい状態になっても数値シミュレーションを実施せずに電子密度予測器の学習を継続させた場合の第1電子密度と第4電子密度と第5電子密度との距離関係を模式的に示した図であり、本開示の比較例を表している。なお、図18において、縦軸及び横軸は図2と同じである。

[0142] 図18に示すように、第2差分1703が第3差分1704がより大きい状態になっても、数値シミュレーションにより第2電子密度を更新せずに、第5電子密度1701を教師データとする電子密度予測器の学習を継続すると、方向H2と方向H1とは一致するとはいえなくなる。

[0143] そこで、本実施の形態では、第2差分1703が第3差分1704より大きくなる、すなわち、第1電子密度1702が第4電子密度1700よりも第5電子密度1701側に近づくと、パラメータ算出部104は、数値シミュレーション

ュレーションを実施する。これにより、第2電子密度1706が更新され、更新後の第2電子密度が第5電子密度として設定され、電子密度予測器の学習が実施される。すなわち、方向H2と方向H1との向きが離れてくると、第2電子密度1706が更新されて方向H2を収束値1705の方向に向けて電子密度予測器を学習させることができる。その結果、効率の良い学習を実現できる。

[0144] 図19は、本開示の実施の形態2に係る処理の一例を示すフローチャートである。以下、図19を用いて、数値シミュレーションをスキップしながら、電子密度予測器を学習させる処理について説明する。

[0145] ステップS300からステップS301までの処理は、図15のステップS100からS101までの処理と同じであるため、説明を省く。

[0146] (ステップS302)

電子密度更新部103は、第1入力データについて、既に数値シミュレーションを実施したか否かを判定する。数値シミュレーションが実施されていない場合(ステップS302でNO)、処理はステップS303に進み、数値シミュレーションが実施されている場合(ステップS302でYES)、処理はステップS305に進む。ここで、ステップS302でNOと判定されるケースとしては、密度汎関数法による数値シミュレーションが1回も実施されていないケースが該当し、例えば、図19のフローの開始直後が該当する。

[0147] (ステップS303)

電子密度更新部103は、S301で算出された第1電子密度を密度汎関数法の初期電子密度に設定すると共に、第1電子密度を算出する際に電子密度予測器に入力された第1入力データを密度汎関数法の計算条件として設定して密度汎関数法による数値シミュレーションを実行し、第2電子密度を算出する。ここで、電子密度更新部103は、ステップS100～ステップS104の1つのループにおいて、数値シミュレーションの計算を1回行うものとする。つまり、密度汎関数法における電子密度を更新する計算を1回行

うものとする。但し、これは一例であり、この計算は1つのループにおいて複数実行されてもよい。

[0148] (ステップS304)

電子密度更新部103は、最後の数値シミュレーション、すなわち、ステップS303の処理で第2電子密度を算出する際に用いられた第1電子密度を第4電子密度としてメモリに記録する。また、電子密度更新部103は、最後の数値シミュレーション、すなわち、ステップS303の処理を実施した際に得られた第2電子密度を第5電子密度としてメモリに記録する。

[0149] (ステップS305)

電子密度更新部103は、第1電子密度と第4電子密度との差分である第2差分を算出する。また、パラメータ算出部104は、第1電子密度と第5電子密度との差分である第3差分を算出する。ここで、第1電子密度とは、ステップS301で算出された最新の第1電子密度を指す。ここでは、第2差分及び第3差分は、第1差分と同様、例えば、RMSEが採用できる。但し、これは一例であり、MAE (Mean Absolute Error) 等の誤差を評価できる手法であればどのような手法を用いて第2差分及び第3差分は算出されてもよい。

[0150] (ステップS306)

パラメータ算出部104は、第2差分が第3差分よりも大きいかな否かを判定する。第2差分が第3差分より大きければ(ステップS306でYES)、処理はステップS303に進み、第2差分が第3差分以下であれば(ステップS306でNO)、処理はステップS307に進む。ここで、第2差分が第3差分よりも大きいという条件は、所定の条件の一例に該当する。

[0151] (ステップS307)

パラメータ算出部104は、ステップS304で第2電子密度が更新された直後では、ステップS103と同様、その第2電子密度(第5電子密度)と、その第2電子密度の算出に用いられた第1電子密度(第4電子密度)との第1差分を算出し、第1差分が最小となるように電子密度予測器のパラメ

ータ W_t を更新する。

[0152] 一方、パラメータ算出部 104 は、第 2 電子密度の更新直後でない場合は、ステップ S305 で算出された第 3 差分が最小となるように電子密度予測器のパラメータ W_t を更新する。

[0153] ステップ S305 からステップ S306 までの処理は、図 15 のステップ S104 からステップ S105 までの処理と同じであるため、説明を省く。

[0154] 次に、図 19 のフローの概要について説明する。まず、フローが開始されると、数値シミュレーションが未実施のため、ステップ S302 で NO と判定され、ステップ S303 にて数値シミュレーションが実施されて第 2 電子密度が算出される。以後、第 2 差分が第 3 差分よりも大きくなるまで、S302 で YES → S305 → S306 で NO → S307 → S308 で NO のループが繰り返され、電子密度予測器の学習が実施される。これにより、第 1 電子密度が最後の数値シミュレーションで算出された第 2 電子密度、すなわち、第 5 電子密度に近づいていく。そして、第 2 差分が第 3 差分より大きくなると、ステップ S306 で YES と判定され、ステップ S304 の数値シミュレーションによって第 2 電子密度が更新される。以後、第 2 差分が第 3 差分より大きくなるまで、S302 で YES → S305 → S306 で NO → S307 → S308 で NO のループが繰り返される。

[0155] このように、本フローでは、第 2 差分が第 3 差分より小さい場合は数値シミュレーションがスキップされ、電子密度予測器の学習が進められ、第 2 差分が第 3 差分よりも大きくなった時点で、数値シミュレーションが実施されて第 2 電子密度が更新されている。そのため、本実施の形態では、第 1 電子密度が更新される度に数値シミュレーションを実行する態様を採用した場合に比べて、電子密度予測器の学習にかかる時間を短縮することができる。

[0156] (実施の形態 2 の変形例)

実施の形態 2 では、第 2 差分が第 3 差分より大きくなったときに第 2 電子密度が更新されたが、本開示はこれに限定されない。例えば、第 3 差分から所定値だけ減じた値より第 2 差分が大きくなったときに第 2 電子密度は更新

されてもよい。或いは、例えば、第3差分に所定値を加えた値より第2差分が大きくなったときに第2電子密度は更新されてもよい。

[0157] 以上、本開示に係る電子密度推定装置1001について、実施の形態に基づいて説明したが、本開示は、これらの実施の形態に限定されるものではない。本開示の趣旨を逸脱しない限り、当業者が思いつく各種変形を本実施の形態に施したものの、異なる実施の形態における構成要素を組み合わせて構築される形態なども、本開示の範囲内に含まれる。

[0158] (ハードウェア構成)

上記の電子密度推定装置1001は、具体的には、マイクロプロセッサ、ROM、RAM、ハードディスクドライブ、ディスプレイユニット、キーボード、及びマウスなどから構成されるコンピュータシステムにより構成されても良い。RAMまたはハードディスクドライブには、電子密度推定プログラムが記憶されている。マイクロプロセッサが、電子密度推定プログラムに従って動作することにより、電子密度推定装置1001は、その機能を達成する。ここで電子密度推定プログラムは、所定の機能を達成するために、コンピュータに対する指令を示す命令コードが複数個組み合わせられて構成されたものである。

[0159] さらに、上記の電子密度推定装置1001を構成する構成要素の一部または全部は、1個のシステムLSI (Large Scale Integration: 大規模集積回路) から構成されても良い。システムLSIは、複数の構成部を1個のチップ上に集積して製造された超多機能LSIであり、具体的には、マイクロプロセッサ、ROM、RAMなどを含んで構成されるコンピュータシステムである。RAMには、コンピュータプログラムが記憶されている。マイクロプロセッサが、コンピュータプログラムに従って動作することにより、システムLSIは、その機能を達成する。

[0160] さらにまた、上記の電子密度推定装置1001を構成する構成要素の一部または全部は、コンピュータに脱着可能なICカードまたは単体のモジュールから構成されても良い。ICカードまたはモジュールは、マイクロプロセ

ッサ、ROM、及びRAMなどから構成されるコンピュータシステムである。ICカードまたはモジュールは、上記の超多機能LSIを含んでも良い。マイクロプロセッサが、コンピュータプログラムに従って動作することにより、ICカードまたはモジュールは、その機能を達成する。このICカードまたはこのモジュールは、耐タンパ性を有しても良い。

[0161] また、本開示は、上記の電子密度推定装置1001により実行される電子密度推定方法であるとしても良い。また、この電子密度推定方法は、コンピュータが電子密度推定プログラムを実行することで実現されてもよいし、電子密度推定プログラムからなるデジタル信号で実現されても良い。

[0162] さらに、本開示は、電子密度推定プログラムまたは上記デジタル信号をコンピュータ読み取り可能な非一時的な記録媒体で構成されてもよい。記録媒体としては、例えば、フレキシブルディスク、ハードディスク、CD-ROM、MO、DVD、DVD-ROM、DVD-RAM、BD（Blu-ray（登録商標）Disc）、半導体メモリなどである。また、電子密度推定プログラムは、非一時的な記録媒体に記録されている上記デジタル信号で構成されてもよい。

[0163] また、本開示は、上記電子密度推定プログラムまたは上記デジタル信号を、電気通信回線、無線若しくは有線通信回線、インターネットを代表とするネットワーク、又はデータ放送等を経由して伝送することで構成されてもよい。

[0164] また、本開示は、マイクロプロセッサとメモリとを備えたコンピュータシステムであって、上記メモリは、電子密度推定プログラムを記憶しており、マイクロプロセッサは、電子密度推定プログラムに従って動作するとしても良い。

[0165] また、電子密度推定プログラムまたは上記デジタル信号を上記非一時的な記録媒体に記録して移送することにより、または電子密度推定プログラムまたは上記デジタル信号を、上記ネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施されてもよい。

[0166] また、本電子密度推定システムは、サーバと、サーバに対してネットワークを介して接続されたユーザが所持する端末とで構成されてもよい。この場合、図7において、電子密度推定装置1001、学習データベース100、及びテストデータベース105はサーバで構成し、表示部107は端末で構成する。電子密度推定部106は端末から所定の物質について電子密度の算出要求をネットワークを介して取得すると、該当する物質の第2入力データをテストデータベース105から取得して、その第2入力データを学習済みの電子密度予測器に入力することで該当する物質の電子密度を推定する。そして、サーバは推定した電子密度を通信装置を用いて端末に送信し、端末の表示部107に表示させればよい。

産業上の利用可能性

[0167] 本開示は、学習データがなくとも電子密度予測器の学習が可能であり、大規模な学習データが用意できない状況で、未知の材料の特性を予測する場合に有用である。

符号の説明

[0168] 100 学習データベース
101 データ取得部
102 電子密度予測部
103 電子密度更新部
104 パラメータ算出部
105 テストデータベース
106 電子密度推定部
108 電子密度予測器
1000 電子密度推定システム
1001 電子密度推定装置

請求の範囲

- [請求項1] 物質の組成及び構造に関する物質情報から前記物質の電子密度を推定する電子密度推定装置における電子密度推定方法であって、
- 前記電子密度推定装置のコンピュータが、
- (a 1) 前記物質情報を格納した学習データベースから第1入力データを取得し、
- (a 2) 前記第1入力データを電子密度予測器に入力することで第1電子密度を算出し、
- (a 3) 前記第1入力データと前記第1電子密度とを用いた数値シミュレーションを実行して第2電子密度を算出し、前記数値シミュレーションは、前記第1入力データと前記第1電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を1回以上行う処理であり、前記第2電子密度は、前記密度汎関数法を用いて求めた収束値ではなく、
- (a 4) 前記第1電子密度と前記第2電子密度との差分を最小化する前記電子密度予測器のパラメータを算出することで前記電子密度予測器を学習させ、
- (a 5) 前記物質情報を格納したテストデータベースから第2入力データを取得し、前記パラメータが設定された前記電子密度予測器に前記第2入力データを入力し、第3電子密度を推定し、
- (a 6) 推定された前記第3電子密度を出力する、
- 電子密度推定方法。
- [請求項2] 前記(a 5)は、前記(a 1)から前記(a 4)までの処理が所定回数繰り返された後に実施される、
- 請求項1に記載の電子密度推定方法。
- [請求項3] 前記(a 3)では、
- (b 1) 前記(a 2)で更新された前記第1電子密度が所定の条件を満たす場合、前記数値シミュレーションを実施して前記第2電子密

度を更新し、処理を前記（a 4）に進め、

（b 2）前記（a 2）で更新された前記第 1 電子密度が前記所定の条件を満たさない場合、前記第 2 電子密度を更新せずに、処理を前記（a 4）に進める、

請求項 2 記載の電子密度推定方法。

[請求項 4] 前記所定の条件は、前記（a 2）で更新された前記第 1 電子密度と第 4 電子密度との第 2 差分が、前記（a 2）で更新された前記第 1 電子密度と第 5 電子密度との第 3 差分よりも大きくなることであり、

前記第 4 電子密度は、前記（b 1）で更新された前記第 2 電子密度の算出に用いられた前記第 1 電子密度であり、

前記第 5 電子密度は、前記（b 1）で更新された前記第 2 電子密度である、

請求項 3 記載の電子密度推定方法。

[請求項 5] 前記物質情報は、結晶学共通データフォーマットを用いて記載された情報である、

請求項 1～4 のいずれかに記載の電子密度推定方法。

[請求項 6] 前記（a 4）では、前記第 2 電子密度が教師データとして用いられて前記電子密度予測器が学習される、

請求項 1～5 のいずれかに記載の電子密度推定方法。

[請求項 7] 前記（a 6）では、前記第 3 電子密度を示す画像がディスプレイに表示される、

請求項 1～6 のいずれかに記載の電子密度推定方法。

[請求項 8] 物質の組成及び構造に関する物質情報から前記物質の電子密度を推定する電子密度推定装置であって、

前記物質情報を格納した学習データベースから第 1 入力データを取得するデータ取得部と、

前記第 1 入力データを電子密度予測器に入力することで第 1 電子密度を算出する電子密度予測部と、

前記第 1 入力データと前記第 1 電子密度とを用いた数値シミュレーションを実行して第 2 電子密度を算出する電子密度更新部と、

前記数値シミュレーションは、前記第 1 入力データと前記第 1 電子密度とを初期値に設定し、密度汎関数法を用いた電子密度を更新する処理を 1 回以上行う処理であり、前記第 2 電子密度は、前記密度汎関数法を用いて求めた収束値ではなく、

前記第 1 電子密度と第 2 電子密度との差分を最小化する前記電子密度予測器のパラメータを算出することで前記電子密度予測器を学習させるパラメータ算出部と、

前記物質情報を格納したテストデータベースから第 2 入力データを取得し、前記パラメータが設定された前記電子密度予測器に前記第 2 入力データを入力し、第 3 電子密度を推定する電子密度推定部と、

推定された前記第 3 電子密度を出力する出力部とを備える、
電子密度推定装置。

[請求項 9] 請求項 1 ～ 7 のいずれかに記載の電子密度推定方法をコンピュータ
に実行させる、
電子密度推定プログラム。

[請求項 10] コンピュータによって行われる電子密度推定方法であって、
 (a 1) 物質の組成及び構造に関する物質情報及びテストデータベースを格納した学習データベースから第 1 入力データを取得し、前記物質情報は前記第 1 入力データを含み、前記テストデータベースは第 2 入力データを含み、
 (a 2) 電子密度予測器は、前記第 1 入力データと第 1 パラメータに基づいて、第 1 電子密度 D_{11} を算出し、
 (a 3) 密度汎関数法を用いて電子密度を更新する計算を 1 回以上行い、第 2 電子密度を算出し、前記 1 回以上の計算は、前記第 1 入力データと前記第 1 電子密度 D_{11} に基づいた前記密度汎関数法を用いた計算を含み、前記第 2 電子密度は、前記密度汎関数法を用いて求め

た収束値ではなく、

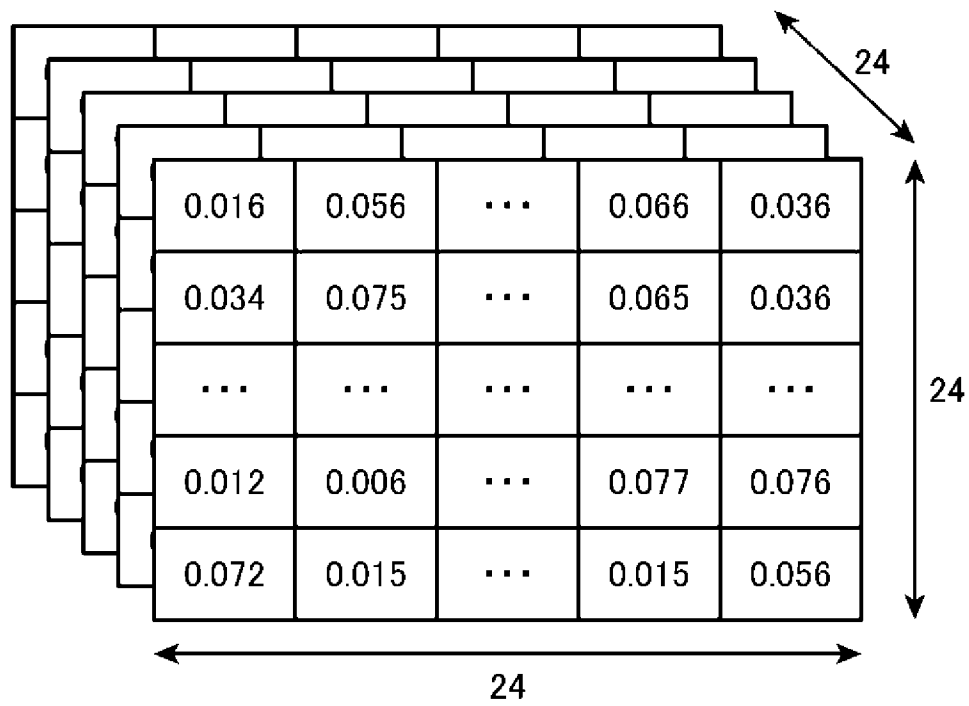
(a 4) 前記電子密度予測器は、前記第 1 入力データと第 i パラメータに基づいて第 1 電子密度 $D1_i$ を算出し ($i = 2 \sim n$)、

(a 5) 前記第 1 電子密度 $D1_j$ と前記 2 電子密度の差分である第 1 差分 $DF1_j$ を算出し ($j = 1 \sim n$)、前記第 1 差分 $DF1_1 \sim$ 前記第 1 差分 $DF1_n$ のち、最も小さい前記第 1 差分 $DF1_k$ を検出し ($1 \leq k \leq n$)、

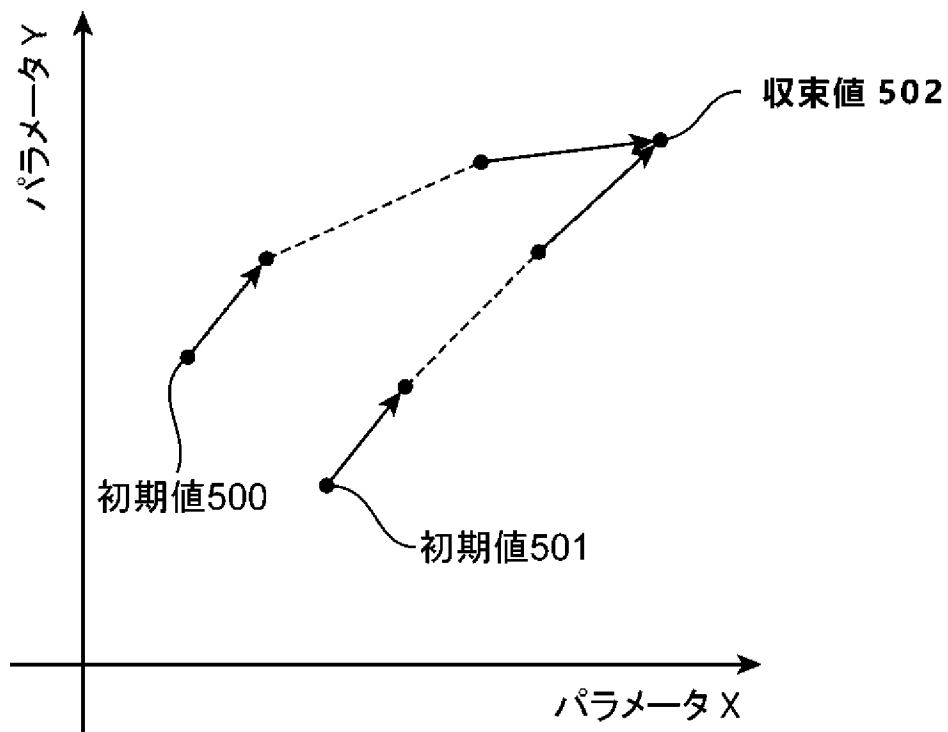
(a 6) 前記第 2 入力データを取得し、前記第 k パラメータが設定された前記電子密度予測器に前記第 2 入力データを入力し、第 3 電子密度を推定し、

(a 7) 推定された前記第 3 電子密度を出力する、
電子密度推定方法。

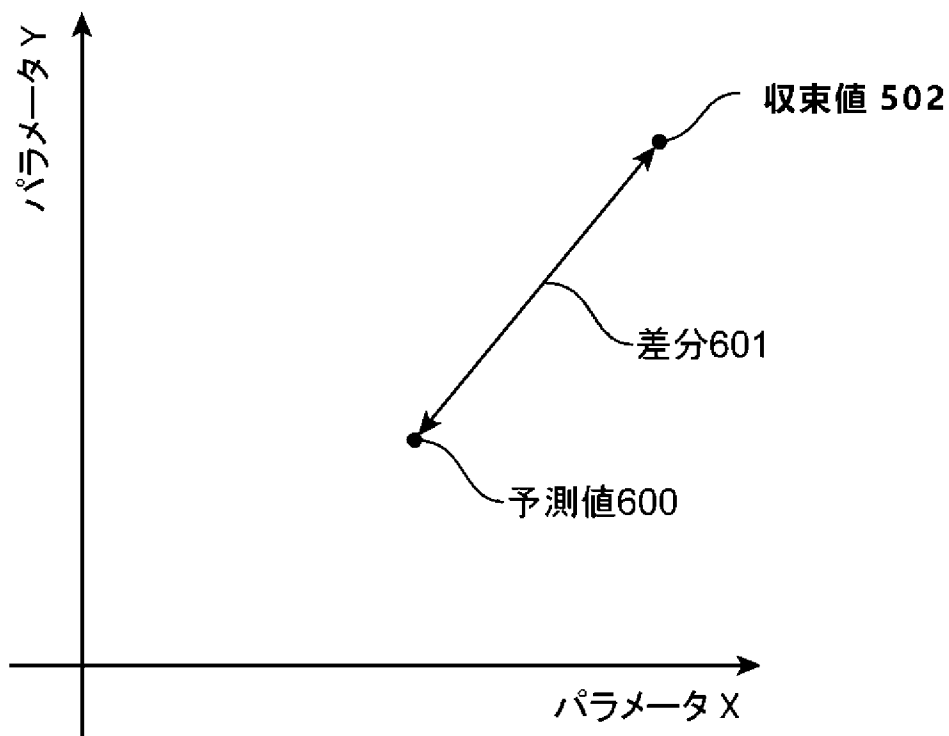
[図1]



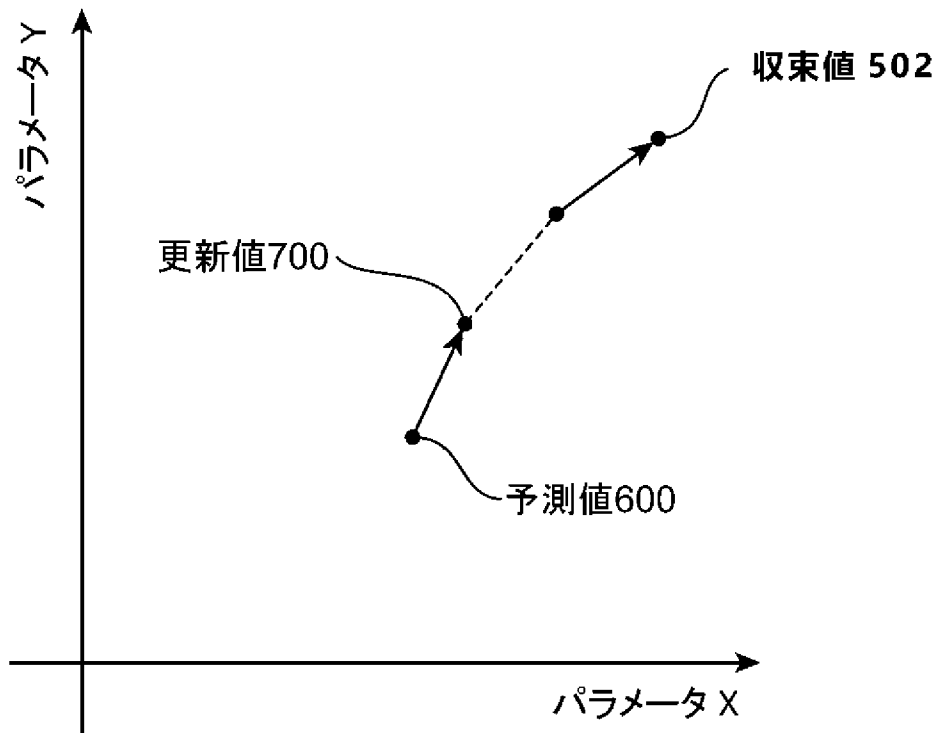
[図2]



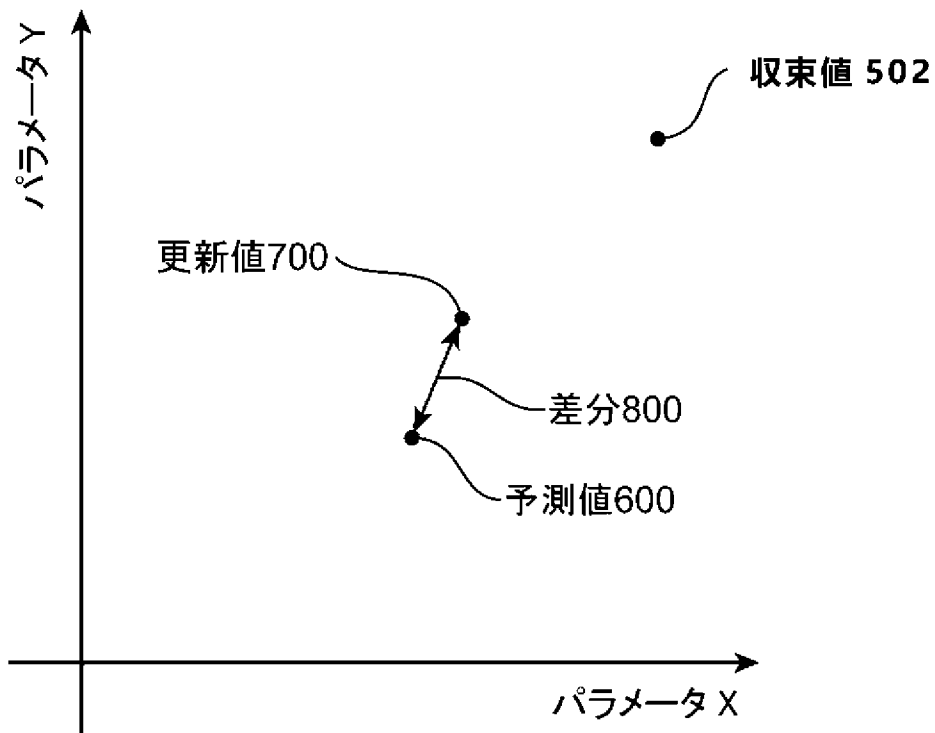
[図3]



[図4]

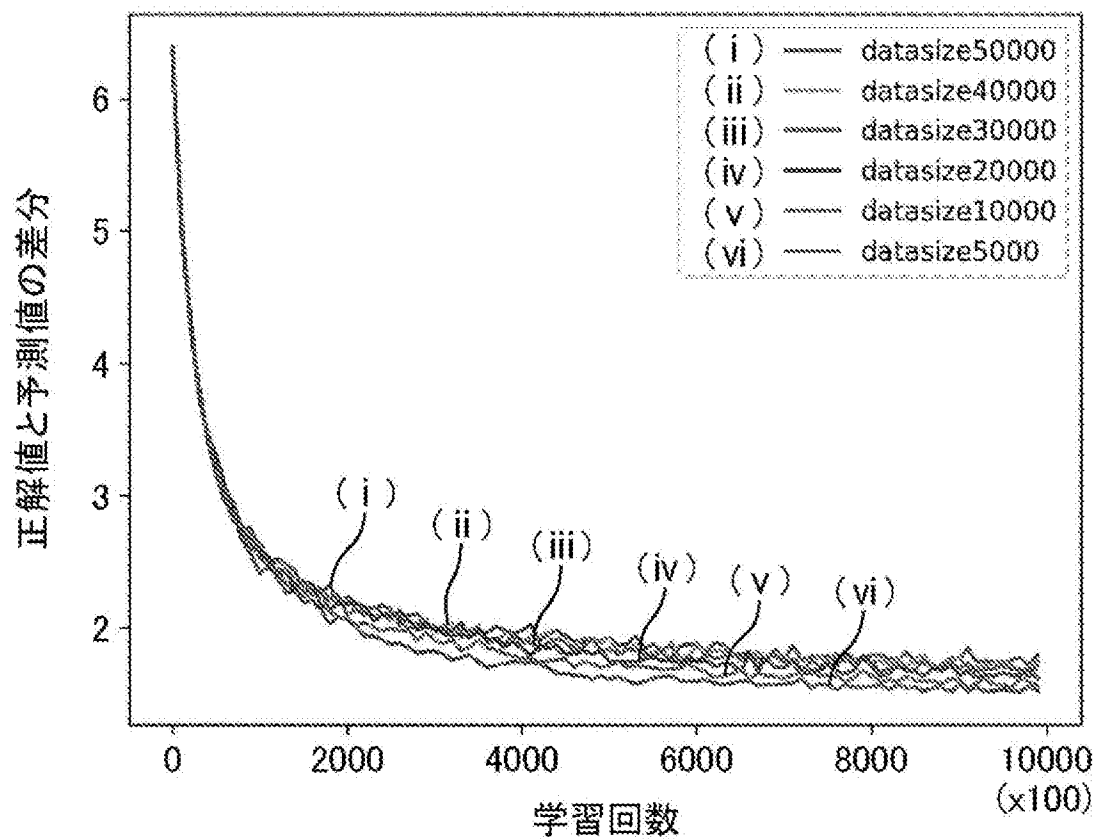


[図5]

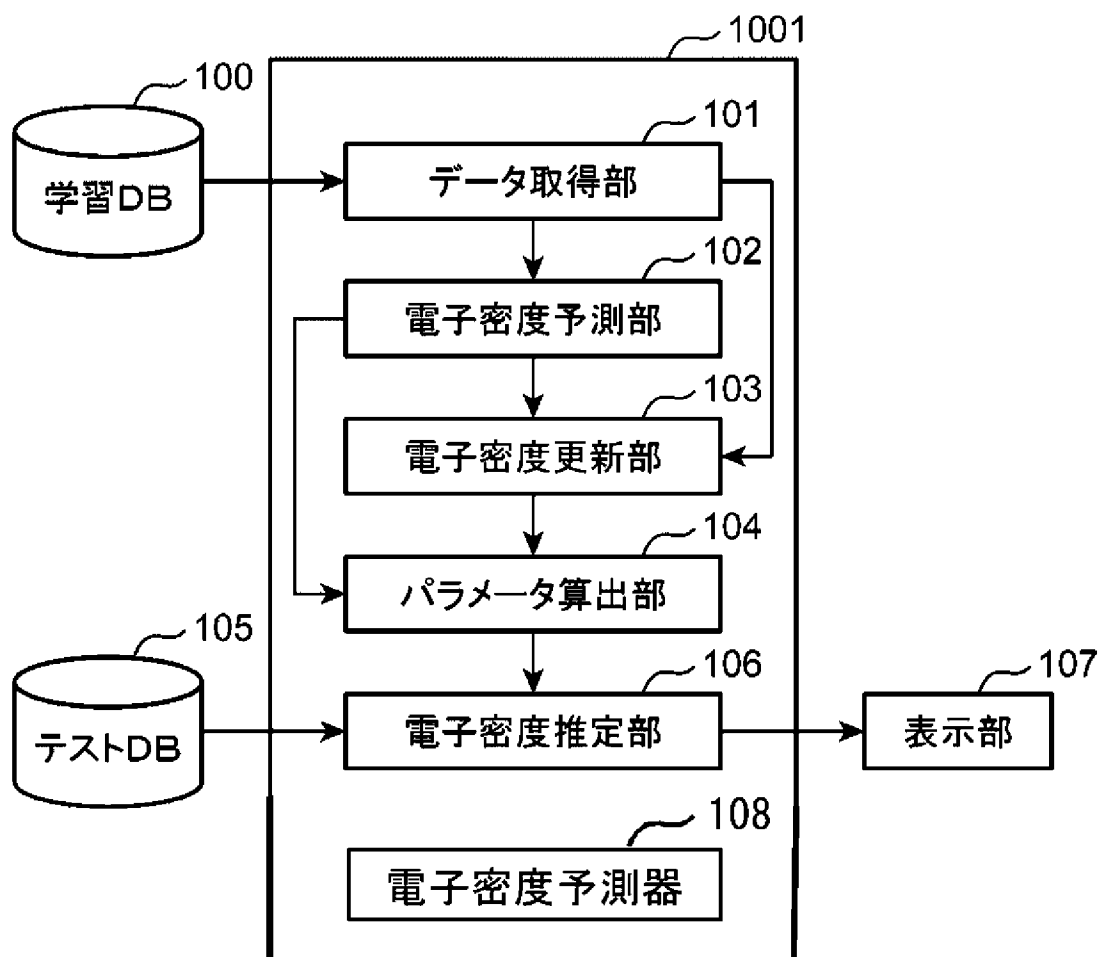


[図6]

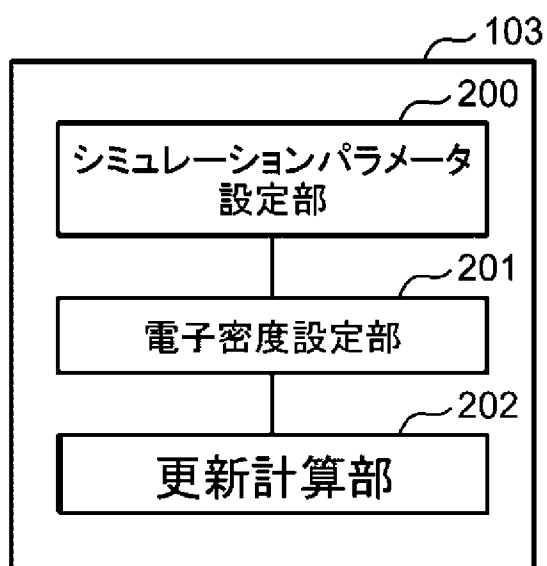
学習データ数毎の学習の推移



[図7]

1000

[図8]



[9]

```

_symmetry_space_group_name_H-M 'R -3 c H'
_symmetry_Int_Tables_number 167
_refine_ls_R_factor_all 0.03
loop_
_symmetry_equiv_pos_site_id
_symmetry_equiv_pos_as_xyz
1 'x-y, -y, -z+1/2'
2 '-x, -x+y, -z+1/2'
(中略)
35 '-y+1/3, x-y+2/3, z+2/3'
36 'x+1/3, y+2/3, z+2/3'
loop_
_atom_type_symbol
_atom_type_oxidation_number
Al3+ 3
O2- -2
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_symmetry_multiplicity
_atom_site_Wyckoff_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_B_iso_or_equiv
_atom_site_occupancy
_atom_site_attached_hydrogens
Al1 Al3+ 12 c 0 0.35216(8) 0.31(4) 1. 0
O1 O2- 18 e 0.3061(5) 0.025 0.33(5) 1. 0
#End of TTdata_9770-ICSD

```

Finger, L. W., & Hazen, R. M. (1978). Crystal structure and compression of ruby to 46 kbar. *Journal of Applied Physics*, 49(12), 5823–5826.

```

#(C) 2017 by FIZ Karlsruhe - Leibniz Institute for Information
Infrastructure. All rights reserved.
data_9770-ICSD
_database_code_ICSD 9770
_audit_creation_date 1984-09-04
_audit_update_record 2012-02-01
_chemical_name_systematic 'Aluminium oxide - Cr-doped'
_chemical_formula_structural 'Al2 O3'
_chemical_formula_sum 'Al2 O3'
_chemical_name_structure_type Corundum-Al2O3
_chemical_name_mineral Corundum
_exptl_crystal_density_diffn 3.98
_cell_measurement_pressure 1000.
_publ_section_title 'Crystal structure and compression of ruby
to 46 kbar'
loop_
_citation_id
_citation_journal_full
_citation_year
_citation_journal_volume
_citation_page_first
_citation_page_last
_citation_journal_id_ASTM
primary 'Journal of Applied Physics' 1978 49 5823 5826
JAPTAU
loop_
_publ_author_name
'Finger, L.W.'
'Hazen, R.M.'
_cell_length_a 4.7607(9)
_cell_length_b 4.7607(9)
_cell_length_c 12.9947(17)
_cell_angle_alpha 90.
_cell_angle_beta 90.
_cell_angle_gamma 120.
_cell_volume 255.06
_cell_formula_units_Z 6

```

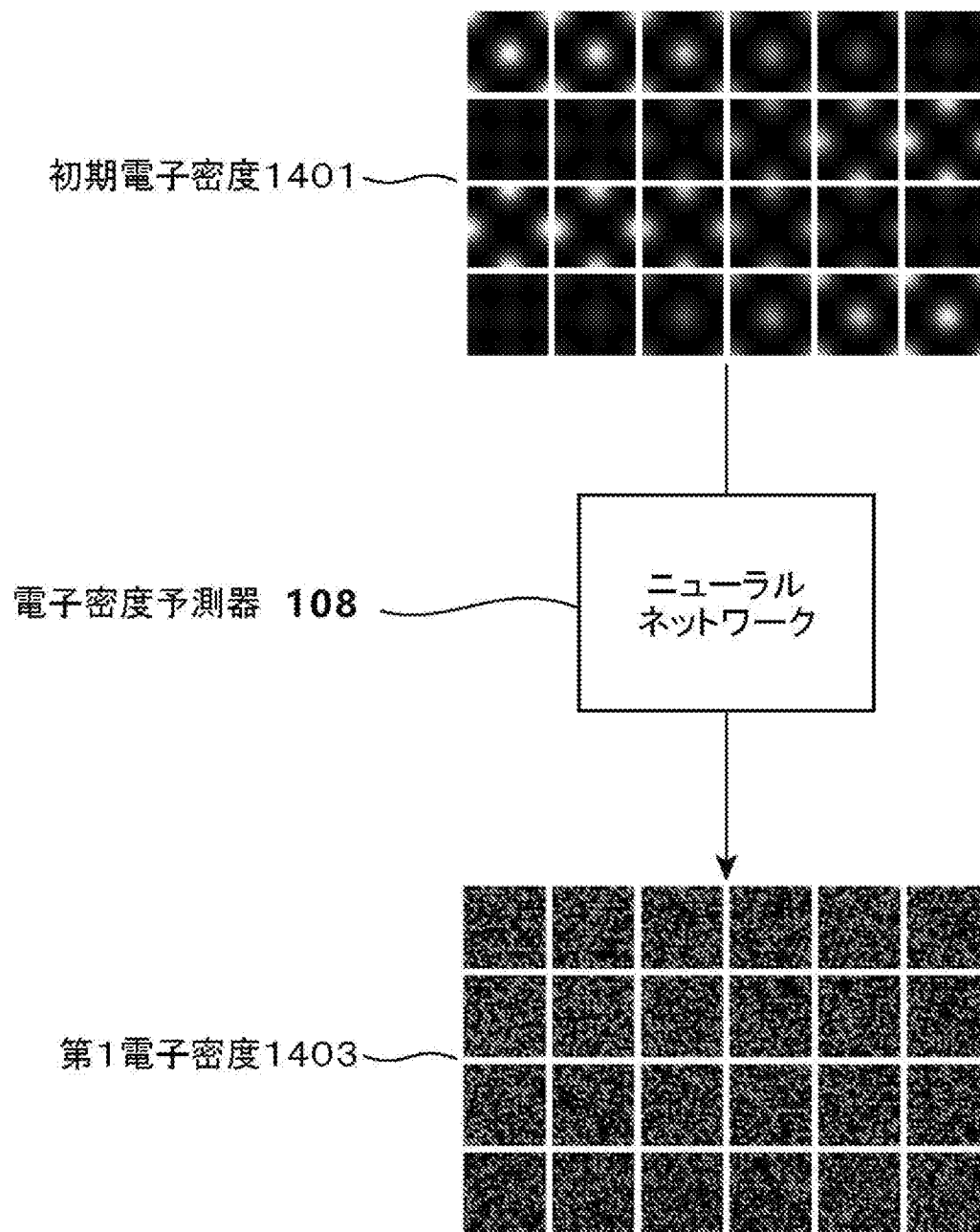
[図10]

```
Al4 O6
1.0
-2.380350 -4.122887 -0.000000
2.380350 -4.122887 -0.000000
-0.000000 -2.748591 4.331567
Al O
4 6
direct
0.852160 0.852160 0.443520 Al3+
0.352160 0.352160 0.943520 Al3+
0.147840 0.147840 0.556480 Al3+
0.647840 0.647840 0.056480 Al3+
0.443900 0.056100 0.750000 O2-
0.056100 0.750000 0.750000 O2-
0.750000 0.443900 0.750000 O2-
0.943900 0.250000 0.250000 O2-
0.250000 0.556100 0.250000 O2-
0.556100 0.943900 0.250000 O2-
```

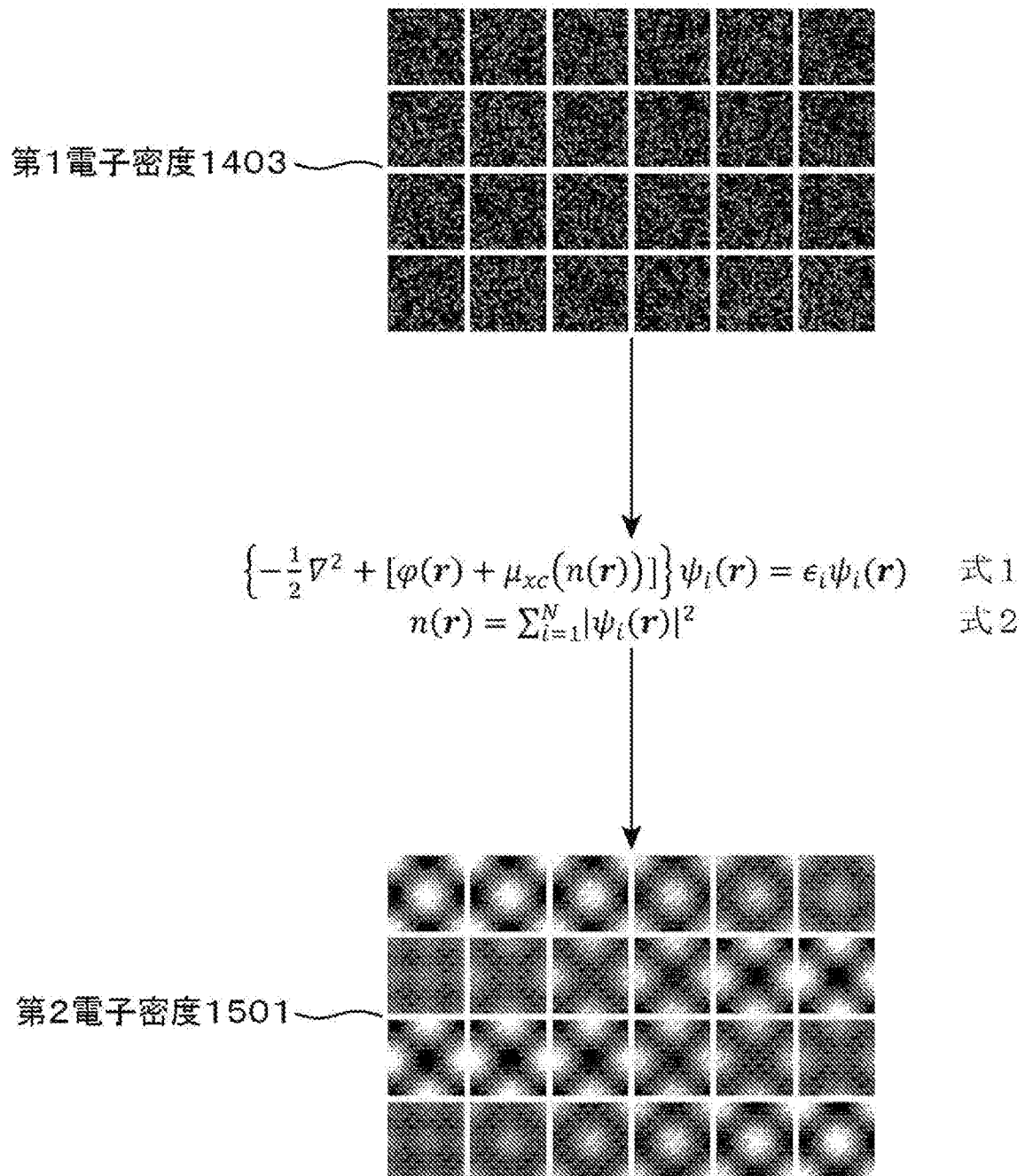

[図11]

組成式	原子番号の平均	原子半径の平均	体積の平均	...	原子番号の分散	原子半径の分散	体積の分散	...
S2.0Hf4.0	53.33	136.67	23.90	...	26.40	25.93	2.31	...
Nb8.0Co2.0Si2.0	34.17	137.50	17.17	...	10.37	12.83	2.86	...
P4.0Mo12.0	35.25	133.75	18.68	...	11.69	19.49	5.50	...
Tl1.0Te2.0Ag1.0	58.00	157.50	28.41	...	13.44	20.46	6.89	...
Ge1.0Ca7.0	21.50	173.13	40.40	...	3.97	18.19	6.71	...
Te4.0Ga4.0	41.50	135.00	26.78	...	10.50	5.00	7.17	...
Si2.0Co2.0Al1.0	19.00	123.00	15.72	...	6.54	11.22	4.06	...
Si1.0Sn1.0	33.00	122.50	27.07	...	17.00	22.50	0.10	...
S2.0Ta1.0	35.00	115.00	24.13	...	26.87	21.21	4.30	...
Si1.0Os1.0	45.00	120.00	17.00	...	31.00	10.00	3.02	...

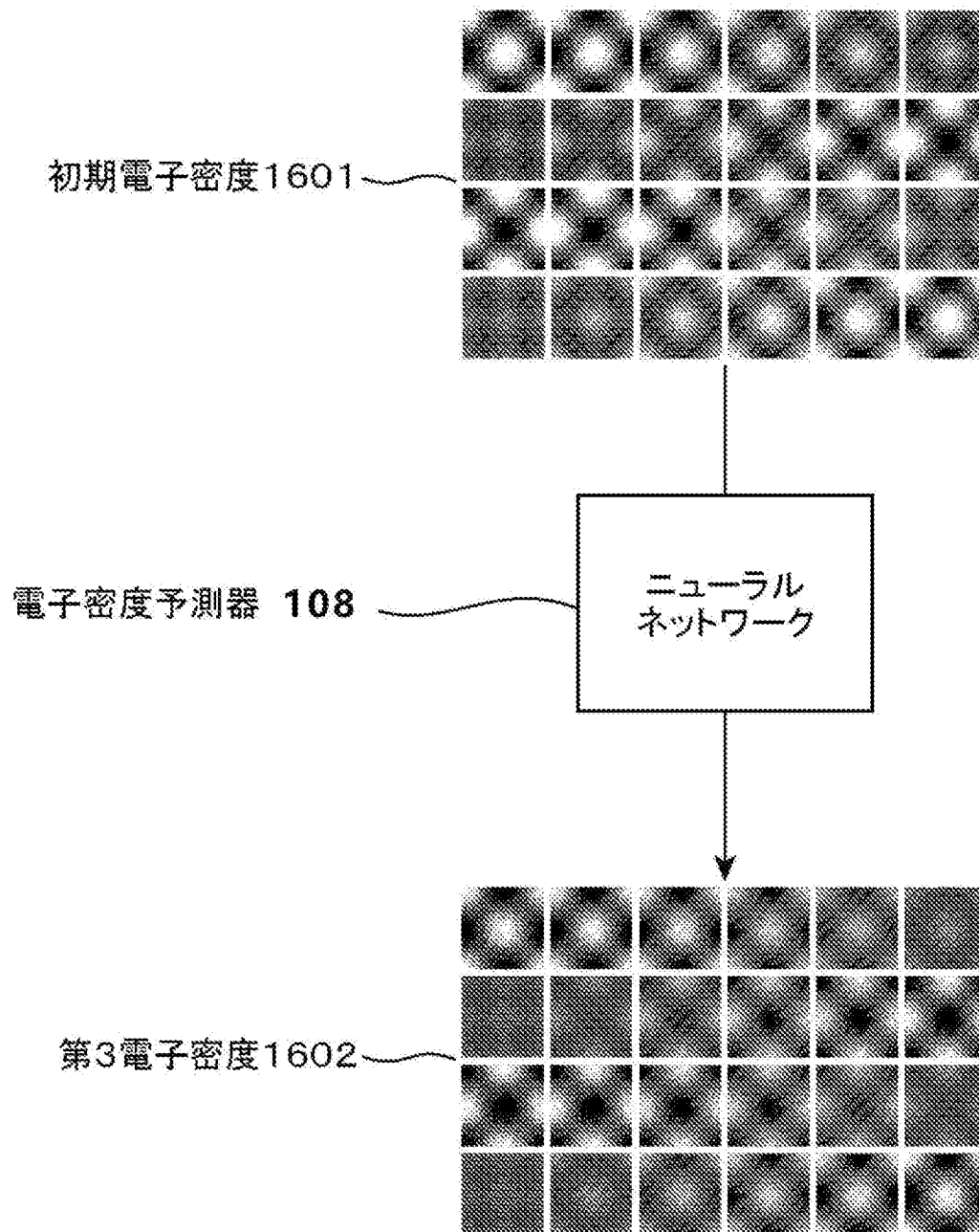
[図12]



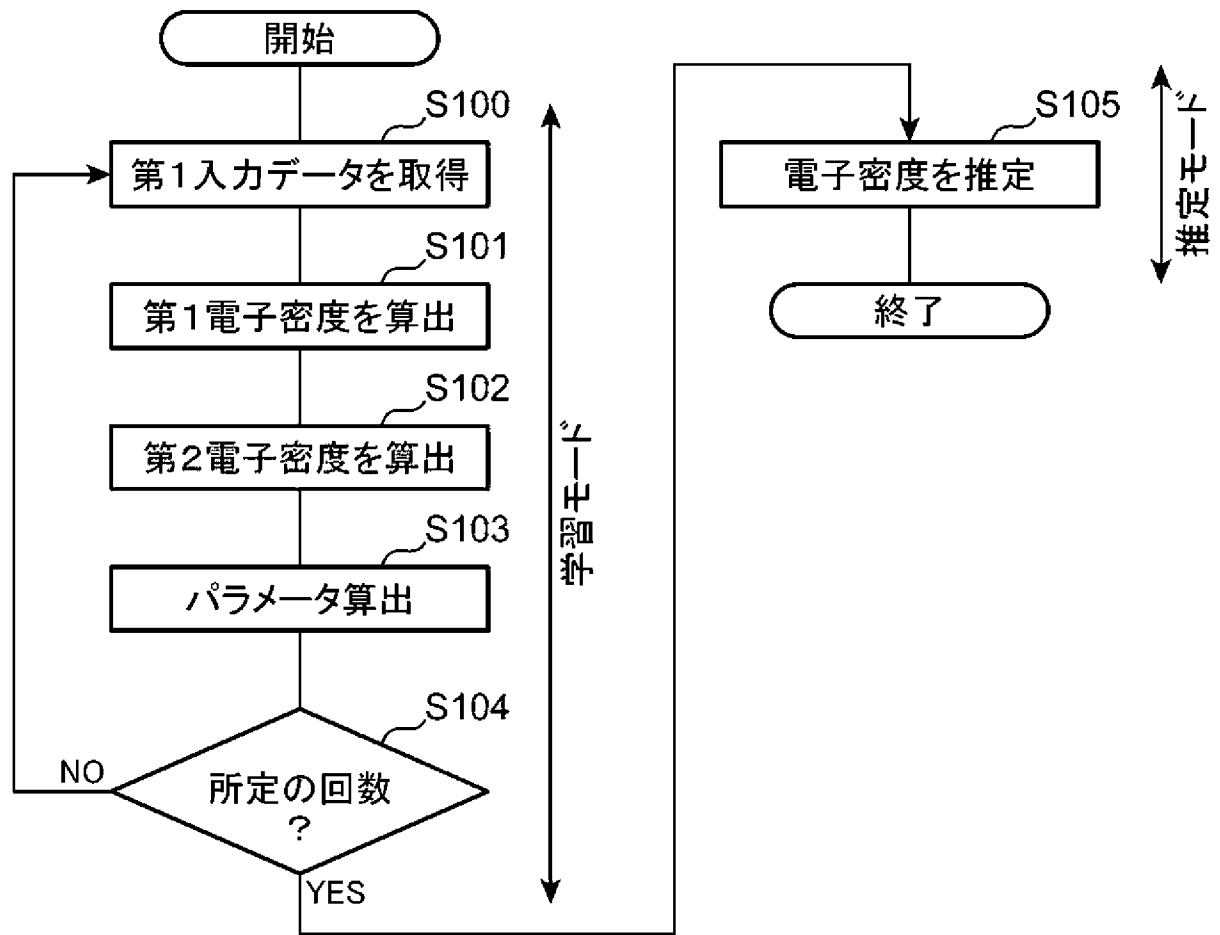
[図13]



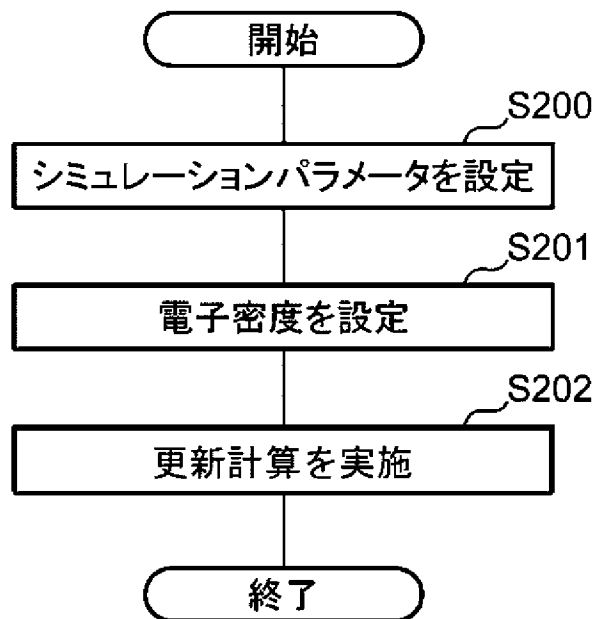
[図14]



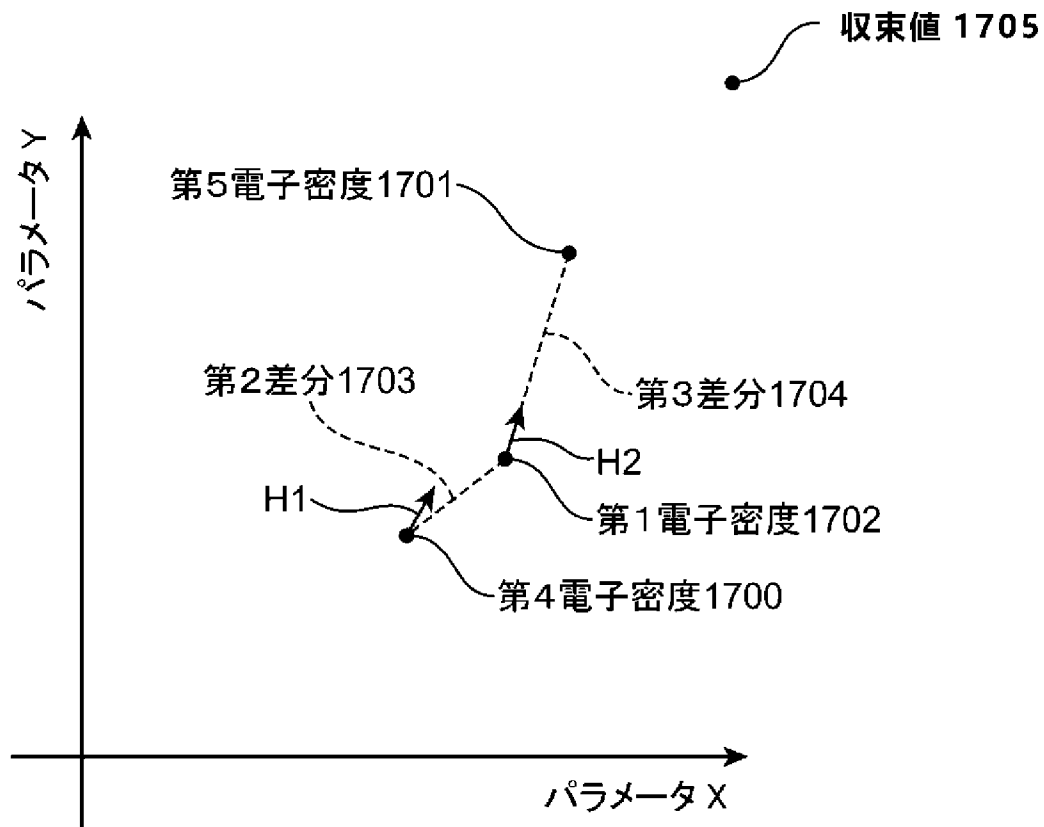
[図15]



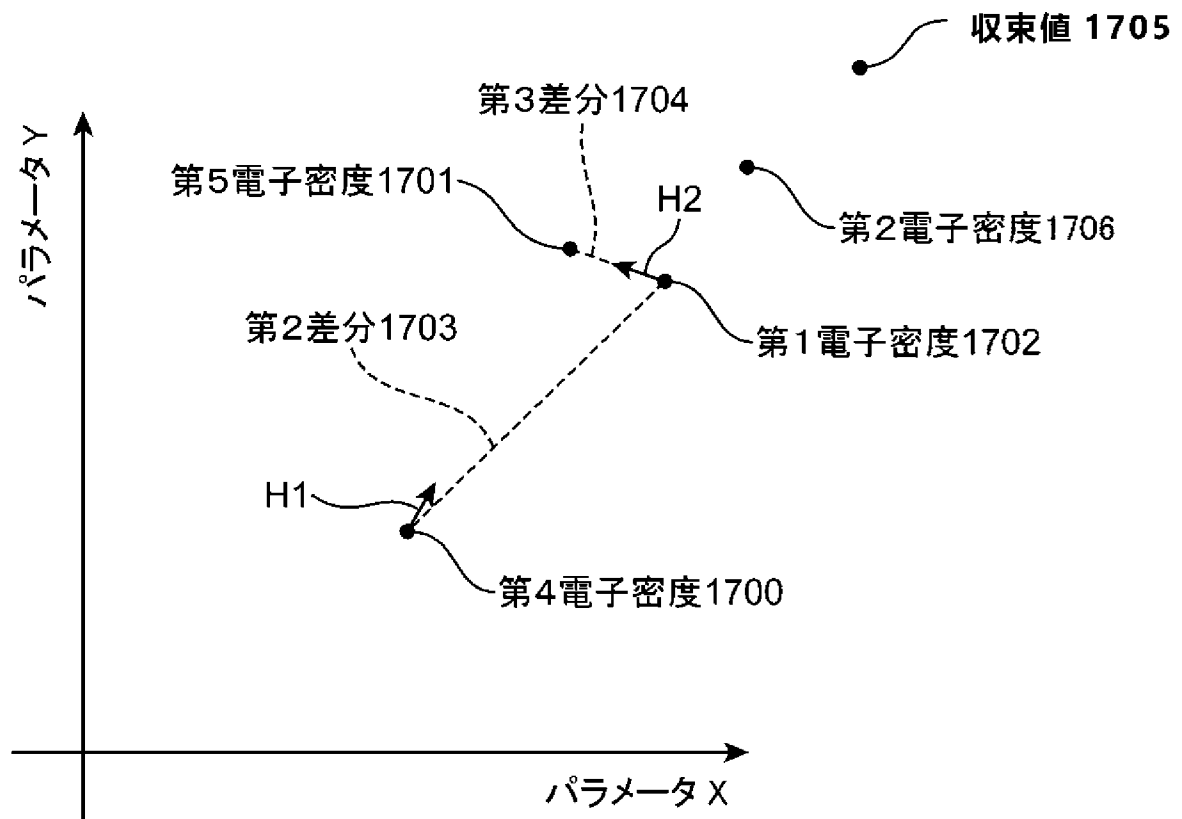
[図16]



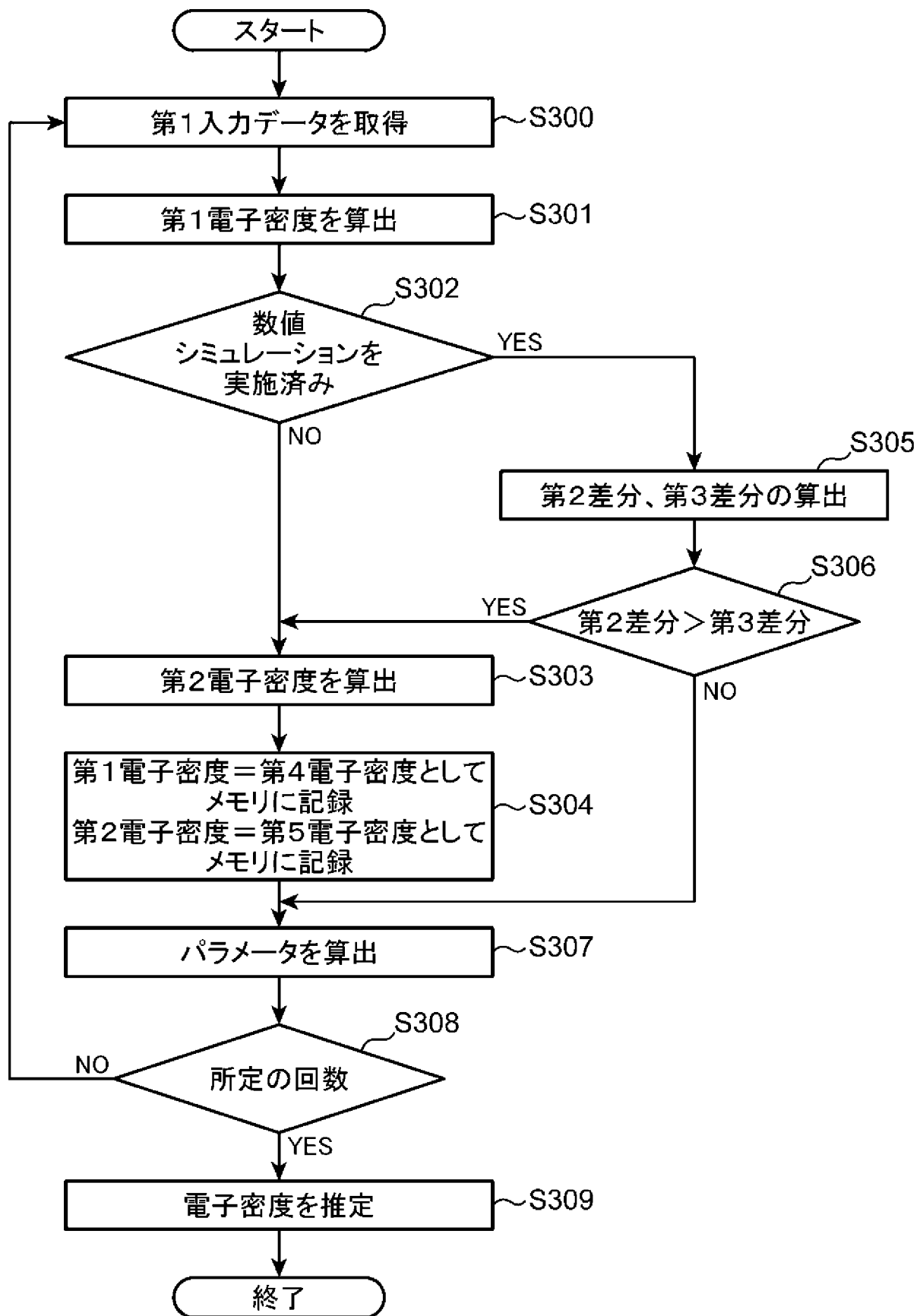
[図17]



[図18]



[図19]



[図20]

単位格子ベクトル

	x成分	y成分	z成分
単位格子ベクトルa	4.805027	0	0
単位格子ベクトルb	-2.40251	4.161275	0
単位格子ベクトルc	0	0	13.11625

原子座標, および, 元素の物性値

	相対x座標	相対y座標	相対z座標	原子番号	原子半径	...
原子1	0	0	0.352096	8	146.2014	...
原子2	0	0	0.647904	8	146.2014	...
原子3	0.333333	0.666667	0.814571	8	146.2014	...
(中略)			(中略)			...
原子28	0.972813	0.639479	0.583333	13	143	...
原子29	0.972813	0.333333	0.083333	13	143	...
原子30	0.360521	0.333333	0.583333	13	143	...

[図21]

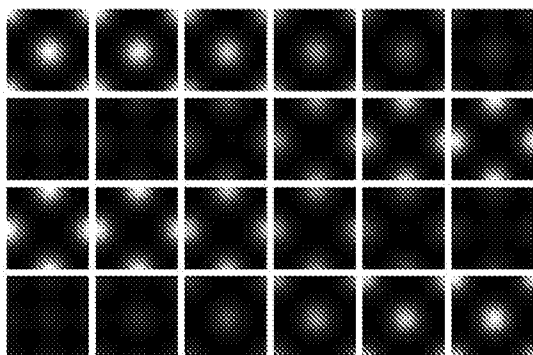
単位格子ベクトル

	x成分	y成分	z成分
単位格子ベクトルa	4.0679998398	0	0
単位格子ベクトルb	0	4.0679998398	0
単位格子ベクトルc	0	0	4.0679998398

原子座標, および, Al原子の物性値

	相対x座標	相対y座標	相対z座標	原子番号	原子半径	...
原子1	0	0	0	8	146.2014	...
原子2	0	0.5	0.5	8	146.2014	...
原子3	0.5	0.0	0.5	8	146.2014	...
原子4	0.5	0.5	0.0	8	146.2014	...

電子が原子座標を中心としたガウス関数状に
分布するとして算出した記述子



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/032350

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G16C10/00 (2019.01) i, G06N20/00 (2019.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. G16C10/00, G06N20/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus/JMEDPlus/JST7580 (JDreamIII), PubMed

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	BROCKHERDE, F. et al., Bypassing the Kohn-Shame equations with machine learning, Nature Communication, 11 October 2017, vol. 8, no. 1, pp. 1-10, DOI: 10.1038/s41467-017-00839-3	1, 2, 5-10 3, 4
Y A	佐藤文俊, 密度汎関数法によるタンパク質全電子シミュレーション, スーパーコンピューティングニュース, February 2009, vol. 11, 特集号 1, pp. 1-33, ISSN 1344-9567, (SATO, Fumitoshi, Supercomputing News, special issue 1), non-official translation (Protein all-electron simulation by density functional theory)	1, 2, 5-10 3, 4



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

07.10.2019

Date of mailing of the international search report

21.10.2019

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/032350

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2016-6617 A (YAHOO JAPAN CORPORATION) 14 January 2016, paragraphs [0015]-[0017] (Family: none)	10
Y	JP 2016-139336 A (CENTRAL RESEARCH INSTITUTE OF ELECTRIC POWER INDUSTRY) 04 August 2016, paragraphs [0058]-[0061] (Family: none)	10
Y	JP 1996-208136 A (FUJITEC CO., LTD.) 13 August 1996, paragraphs [0027], [0028] (Family: none)	10
A	US 2008/0104001 A1 (KIPP, J. E.) 01 March 2008, entire text, all drawings (Family: none)	1-10
A	US 6587845 B1 (BRAUNHEIM, B. B.) 01 July 2003, entire text, all drawings (Family: none)	1-10
A	WO 2012/028962 A2 (DELACOTTE, N. et al.) 08 March 2012, entire text, all drawings (Family: none)	1-10

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G16C10/00(2019.01)i, G06N20/00(2019.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. G16C10/00, G06N20/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlus/JMEDPlus/JST7580 (JDreamIII)
PubMed

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	BROCKHERDE, F. et al., Bypassing the Kohn-Sham equations with machine learning, Nature Communication, 2017. 10. 11, Vol.8 No.1, pp.1-10, DOI: 10.1038/s41467-017-00839-3	1, 2, 5-10 3, 4
Y A	佐藤文俊, 密度汎関数法によるタンパク質全電子シミュレーション, スーパーコンピューティングニュース, 2009. 02, Vol.11 特集号1, pp.1-33, ISSN 1344-9567	1, 2, 5-10 3, 4

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

07. 10. 2019

国際調査報告の発送日

21. 10. 2019

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

松野 広一

5 R

4806

電話番号 03-3581-1101 内線 3502

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2016-6617 A (ヤフー株式会社) 2016.01.14, 段落[0015]-[0017] (ファミリーなし)	10
Y	JP 2016-139336 A (一般財団法人電力中央研究所) 2016.08.04, 段落[0058]-[0061] (ファミリーなし)	10
Y	JP 1996-208136 A (フジテック株式会社) 1996.08.13, 段落[0027]-[0028] (ファミリーなし)	10
A	US 2008/0104001 A1 (KIPP, J. E.) 2008.03.01, 全文, 全図 (ファミリーなし)	1-10
A	US 6587845 B1 (BRAUNHEIM, B. B.) 2003.07.01, 全文, 全図 (ファミリーなし)	1-10
A	WO 2012/028962 A2 (DELACOTTE, Nicolas et al.) 2012.03.08, 全文, 全図 (ファミリーなし)	1-10