

明 細 書

発明の名称：材料特性予測システムおよび情報処理方法

技術分野

[0001] 本発明は、材料科学等の実験を支援する技術に関する。

背景技術

[0002] データ分析にかかる統計処理技術の発展に伴い、材料科学においてもデータ分析を行う需要が高まっている。特に、材料科学分野では、新素材の開発を効率的に行うために、次の実験の候補の選定を、既知のデータを元にして行うスクリーニングとよばれる方法が知られている。

[0003] スクリーニングの方法としては、各種実験のデータを情報システムに入力して機械学習を行って実験結果の予測のモデルを構築し、モデルの予測に基づいたスクリーニングを行う。この予測には、材料設計に関する様々なパラメータを引数にとり、材料の特性を返す関数を回帰分析によって求める方法がよく知られている。

先行技術文献

特許文献

[0004] 特許文献1：特開2004－086892号公報

特許文献2：特表2017－520868号公報

特許文献3：特開2008－081435号公報

発明の概要

発明が解決しようとする課題

[0005] 材料開発においては、材料特性の予測の精度を高めることで、よりの確に新材料の候補に対する有望性を見極めることができ、不要な実験を省略することで効率的な材料開発ができるようになると期待される。

[0006] 回帰分析では、関数の引数に相当する変数を説明変数、関数の返り値に相当する値を目的変数と呼ぶが、材料特性の予測では材料特性を目的変数とし、それを予測できるように材料の特徴を示す説明変数を選ぶ。この説明変数

の選択いかんによって予測の精度は上下するため、様々な材料特性の予測に対応できるようにバラエティに富んだ説明変数の生成手法をそろえることが肝要となる。

[0007] 特許文献 1 では、説明変数として成分配合比をもちいて材料特性を予測している。この方法は複数の物質が混合してできる材料の特性予測に用いることができる。しかし、単一の物質の特性の予測には適用できない。

[0008] 特許文献 2 には、分子の立体的な構造を、分子の周囲の空間を空間格子（ボクセル）にわけ、各ボクセル内の原子数によって表現し、それを説明変数に用いる方法が開示されている。この方法によれば、単一の分子に対してもその立体的な形状をもとにした物性予測が可能となる。

[0009] ところが、ボクセルを用いた方法では、座標系の決め方に自由度がある。すなわち、原点を分子のどこに置くのか、または、どの方向をx軸とするのか、などを決める手段がない。つまり、同じ物質でも多数のボクセルがありうることになる。

[0010] 特許文献 2 の発明では、原点や角度を変えたデータを多数生成することで、回帰分析にこの自由度を織り込もうとしているが、重複するデータを多数投入することになり計算時間等が大幅に増えてしまう。また、回帰分析の技法にもよるが、回帰分析のアルゴリズムがこの自由度を適切にモデルに織り込むことができるかは明確ではなく、予測精度をむしろ低下させる懸念が残る。さらに、仮にうまく予測できたとしても、逆計算が出来ないという課題もある。例えば、材料特性の予測値が最も高い条件を見つけたいという場合には、材料の特性を返す関数の最大値を探索すればよさそうだが、そのときのボクセルの説明変数を得たところで、対応する分子の構造は容易には推察できない。

[0011] 同様に分子の立体構造をもとにしたスクリーニングの方法としては、特許文献 3 に開示されているような既知の分子との類似性を評価する方法も知られている。この方法では他の分子を基準とするために分子単独の座標系の自由度の影響はすくないが、やはり逆計算が難しく、また、十分多数の分子が

既知でなければ適用できないという問題がある。

- [0012] そこで、座標系の選択に関する自由度をもたせることなく分子の空間構造を規定し、分子の立体構造に基づいた材料特性の予測をすることが望まれる。

課題を解決するための手段

- [0013] 本発明の好ましい一側面は、材料特性の予測を行うためのシステムであって、材料の構造式から分子を構成する原子の位置座標を算出する機能を備えた三次元分子構造算定部と、三次元分子構造算定部の算出した原子の位置座標をもとに、3つの原子を選び出して三角形を構築し、3つの原子と他の原子との距離を算出して空間構造特徴量とする機能を持つ空間構造特徴量算定部と、空間構造特徴量算定部の生成した空間構造特徴量を説明変数として材料特性を予測する材料特性予測部と、を持つことを特徴とする材料特性予測システムである。

- [0014] 本発明の好ましい他の一側面は、材料の構造式を入力とし、該材料の構造式から分子を構成する原子の位置を算出する三次元分子構造算定処理と、算出した原子の位置をもとに、3つの原子を選択して三角形を構築し、3つの原子と他の原子との距離を算出して空間構造特徴量を得る空間構造特徴量算定処理と、を実行する情報処理方法である。

発明の効果

- [0015] 座標系の選択に関する自由度をもたせることなく分子の空間構造を規定し、分子の立体構造に基づいた材料特性の予測が可能になる。

図面の簡単な説明

- [0016] [図1]実施例1の構成概略の例を示す機能ブロック図。
[図2]実施例1の物理的実装の構成の例を示すブロック図。
[図3]実施例1の使用手順の例を示す概念図。
[図4]実施例1の材料DB更新処理の例を示すフロー図。
[図5]実施例1の実験データ受付の画面表示の例を示すイメージ図。
[図6]実施例1の実験データの構成の例を示す表図。

[図7]実施例1の材料DBの実験データテーブルの例を示す表図。

[図8]実施例1の三次元分子構造データの1分子分の情報の例を示す表図。

[図9]実施例1で想定される三次元分子構造の模式図。

[図10]実施例1の空間構造特徴量計算処理の例を示すフロー図。

[図11]実施例1で想定される空間構造特徴量の計算過程の模式図。

[図12]実施例1の空間構造特徴量のテーブルの例を示す表図。

[図13]実施例1の材料特性予測処理の例を示すフロー図。

[図14]実施例1の材料特性予測表示の例を示すイメージ図。

[図15]実施例1の材料特性予測用データの構成の例を示す表図。

[図16]実施例2の構成概略の例を示す機能ブロック図。

[図17]実施例2の材料特性予測処理の例を示すフロー図。

[図18]実施例2の材料特性予測表示の例を示すイメージ図。

発明を実施するための形態

[0017] 実施の形態について、図面を用いて詳細に説明する。ただし、本発明は以下に示す実施の形態の記載内容に限定して解釈されるものではない。本発明の思想ないし趣旨から逸脱しない範囲で、その具体的構成を変更し得ることは当業者であれば容易に理解される。

[0018] 以下に説明する発明の構成において、同一部分又は同様な機能を有する部分には同一の符号を異なる図面間で共通して用い、重複する説明は省略することがある。

[0019] 同一あるいは同様な機能を有する要素が複数ある場合には、同一の符号に異なる添字を付して説明する場合がある。ただし、複数の要素を区別する必要がない場合には、添字を省略して説明する場合がある。

[0020] 本明細書等における「第1」、「第2」、「第3」などの表記は、構成要素を識別するために付するものであり、必ずしも、数、順序、もしくはその内容を限定するものではない。また、構成要素の識別のための番号は文脈毎に用いられ、一つの文脈で用いた番号が、他の文脈で必ずしも同一の構成を示すとは限らない。また、ある番号で識別された構成要素が、他の番号で識

別された構成要素の機能を兼ねることを妨げるものではない。

[0021] 図面等において示す各構成の位置、大きさ、形状、範囲などは、発明の理解を容易にするため、実際の位置、大きさ、形状、範囲などを表していない場合がある。このため、本発明は、必ずしも、図面等の開示された位置、大きさ、形状、範囲などに限定されない。

実施例 1

[0022] <1. システム構成>

図1に、実施例1の材料特性予測装置の例を示す。本実施例の材料特性予測装置(101)は、利用者(102)の操作を受け付ける装置であって、利用者から実験データを受け取る実験データ受付部(111)、材料の特徴とその特性が格納された材料データベース(DB: Data Base)(112)、材料の分子構造式の情報を受け取り、分子を構成する各原子の座標を推定する三次元分子構造算定部(113)、三次元分子構造算定部(113)の算定した原子位置を元に当該分子の空間構造特徴量を算定する空間構造特徴量算定部(114)、空間構造特徴量算定部(114)が算出した空間構造特徴量を、当該分子の識別子と対応付けて保存する空間構造特徴量DB(115)、材料DB(112)および空間構造特徴量DB(115)の情報を元に材料特性値が計測済みの分子の情報から未計測の分子の材料特性を予測する材料特性予測部(116)、材料特性予測部(116)の結果を利用者(102)に提示する材料特性予測提示部(118)、を含むシステムである。

[0023] 本実施例では材料特性予測装置(101)は、入力装置、出力装置、記憶装置、処理装置を含むサーバのような情報処理装置で構成することにした。計算や制御等の機能は、記憶装置に格納されたプログラムが処理装置によって実行されることで、定められた処理を他のハードウェアと協働して実現される。図1では、情報処理装置のハードウェア構成に代えて、機能ブロックを示している。各機能ブロックとして、計算機などが実行するプログラム、その機能、あるいはその機能を実現する手段を、「機能」、「手段」、「部」、「ユニット」、「モジュール」等と呼ぶ場合がある。

[0024] 図2に実施例1の物理的実装の構成の一例を示す。材料特性予測装置(101)

は、一般的なコンピュータを用いて実装できる。すなわち、演算性能を持ったプロセッサ(201)、高速に読み書きが可能な揮発性一時記憶領域であるDRAM (Dynamic Random Access Memory) (202)、HDD (磁気ディスク装置) やフラッシュメモリなどを利用した永続的な記憶領域である記憶装置(203)、操作を行うためのマウスやキーボード等の入力装置(204)、動作を使用者に示すためのモニタ(205)、外部と通信を行うためのシリアルポート等のインタフェース(206)を含む装置である。

[0025] 図1の実験データ受付部(111)、三次元分子構造算定部(113)、空間構造特徴量算定部(114)、材料特性予測部(116)、材料特性予測提示部(118)は、記憶装置(203)に記録されたプログラムをプロセッサ(201)が実行することによって実現できる。材料DB(112)、空間構造特徴量DB(115)は、データの蓄積を記憶装置(203)に行うようなプログラムをプロセッサ(201)が実行することで実装できる。

[0026] 図2の構成は、単体のコンピュータで構成してもよいし、あるいは、任意の部分が、ネットワークで接続された他のコンピュータで構成されてもよい。すなわち、複数のコンピュータによって、同様のシステムを構成してもよい。

[0027] 図3に実施例1のシステムの利用手順について模式的に示す。実施例1は、利用者が材料特性予測にまつわるデータを投入する材料データ投入(S310)と、材料特性予測の結果を確認する予測結果閲覧(S320)の2つの手順が実行できる。

[0028] 材料データ投入(S310)は、実験を行った材料のデータ、および、これから実験を行おうとしている材料のデータを格納したデータセットである実験データ(600)を材料特性予測装置(101)に投入する手順である。材料特性予測装置(101)はこれに応じて材料DB更新処理(S311)を実行することで材料DB(112)内部に保存されている情報を更新する。

[0029] 予測結果閲覧(S320)では、材料特性予測装置(101)は利用者(102)の求めに応じて材料特性予測提示処理(S321)を実行し、材料特性予測の結果を可視化

した画面である材料特性予測表示(322)を提示する。

[0030] <2. 材料データ投入処理>

図4に材料DB更新処理(S311)の処理手順の例を示す。材料DB更新処理(S311)では、最初に実験データ受付部(111)が利用者(102)から実験データ(600)を受け取って材料DB(112)を更新する(S401)。その後、三次元分子構造算定部(113)が材料DB(112)のデータに対応する三次元分子構造データ(800)を生成する(S402)。三次元分子構造データ(800)を用いて、空間構造特徴量算定部(114)が空間構造特徴量生成処理(S403)を実行して空間構造特徴量(1100)を算出、空間構造特徴量DB(115)に格納する。

[0031] 図5に材料DB更新処理(311)の最初のステップ(S401)で利用者(102)から実験データ(600)を受け取るために、モニタ(205)に表示する画面の一例を示す。実施例1では、利用者(102)が予め実験データをファイルに記憶しておき、そのファイルの位置をテキストボックス(501)に指定する形で実験データ(600)を受け渡す。受け渡されるファイルには、公知のCSV(Comma Separated Value)形式で表形式のデータが記述されており、それを解釈して表形式とした結果が表画面(502)に表示される。

[0032] 図5では、記述されている情報が実験の識別子である「ID」、実験時の温度を示す「Temp」、そのときの水溶性を示す「SOL」、材料の構造式を示す文字列「SMILES」が例示されている。この例では水溶性が予測を行いたい材料特性であり、SOL欄が空白のデータは未実験の条件ということを示している。なお、このデータの受け渡しは一例であり、表形式に変換可能な情報として、材料の構造式と材料特性を含む実験データを渡せる形態であれば他の方法でもよい。情報は表画面(502)に表示され、ボタン(503)により材料DB(112)に保存される。

[0033] 図6にこの実験データ(600)の1レコードの構成例を示す。本例では、実験データ(600)は、材料特性(601)、例えばSMILES形式などの材料の構造式を示せる情報である材料構造式(602)、温度や圧力など実験時の条件を示す実験条件(603)の情報を1レコードとする情報であり、実験データ(600)はこのレコー

ドを複数集めたものである。これらの情報は、図5の表画面(502)の各項目と対応しており、本実施例では各項目がどの要素に対応するか所定の項目名との対応をもって判定されている。この対応関係については画面から利用者(102)に入力してもらうなどしてもよい。また、材料特性(601)については、実験などによって判明している数値が格納され、未実験の場合は空白が格納される。材料特性(601)や実験条件(603)の定義や個数は任意である。

- [0034] 図4の材料DB更新処理(S311)の最初のステップ(S401)では、前記実験データ(600)を解釈・整形して材料DB(112)の実験データテーブルとして格納する。
- [0035] 図7に実験データテーブルの1レコードの情報を示す。このデータは、実験を一意に識別可能なように連番などで付番された実験ID(701)、実験データ(600)の材料特性(601)に由来する材料特性(702)、実験データ(600)の材料構造式(602)に由来する材料構造式(703)、実験条件(603)に由来する実験条件(704)を含む。これらは各由来となる情報に対して、単位や形式の変換を行い統一した表現に変換されるようにしてもよい。
- [0036] 図8に材料DB更新処理(S311)の第二のステップである処理(S402)で、三次元分子構造算定部(113)において算定される三次元分子構造データ(800)の一分子相当の構成を示す。このデータは、分子を構成する各原子の相対座標を記述した情報である。分子の種類によってその分子を構成する原子数は異なるため、ここではN個の原子からなっている分子を例にとっている。ここでいう相対座標は、材料構造式(703)にあらわれている原子とその結合の情報からディスタンス・ジオメトリー法など公知の方法で簡易的に算定できる座標値である。通常、原子間の距離がファン・デル・ワールス半径を下回っているような場合は原子間が結合している状態に対応する。
- [0037] 図9に三次元分子構造の例を模式的に示した。この図ではある一つの分子について3次元空間上での位置を示しており、球および元素記号が原子を、球の間の線が結合を示す形で描かれている（なお、水素原子は省略されている）。分子を構成する原子の種類と結合の情報は構造式に記載されているが、

通常、その原子の位置は記述されていない。例えば炭素原子(901)が存在すること、および結合する別の炭素原子については構造式に記載のある情報である。

[0038] しかし、例えば炭素原子(901)の座標値(1.0, 1.2, 5.0)や炭素原子(902)の座標値(7.0, 3.7, 5.0)は記載がないため算定が必要である。公知の方法の一例では、構造式にもとづきファン・デル・ワールス半径程度の位置に一度原子を配置し、結合の角度等について妥当な値になるようにその位置を最適化計算して配置を求めることができる。この計算方法については種々のものが公知であり、ある程度の精度が得られるようになっていれば、どのような方法を用いて求めてもよい。

[0039] この計算結果として得られる座標値については、あくまで相対座標であり、座標系は分子によって異なる。これについては、例えば分子の重心等を用いて何らかの統一的基準を作ろうとする方法があるが、本実施例ではこの座標系の基準はなくてもよく、任意のものでよいという利点がある。

[0040] 実験データテーブルの材料構造式(703)のそれぞれについて、原子の位置を算定して適当な順番で記載した結果として、三次元分子構造データ(800)の情報が得られる。この際、実験データ(600)との対応をとるために対応する実験ID(701)を対応付けておくことが求められる。

[0041] 材料DB更新処理(311)の第三のステップは、三次元分子構造データ(800)から特徴量を算出する空間構造特徴量生成処理(S403)である。

[0042] 図10にこの処理の詳細を示す。この処理では、最初に立体構造中の基準となる3つの原子を選択する(S1001)。本実施例では、分子を構成する炭素原子3つの組み合わせを調べ、最も面積の大きな三角形を構成する原子を選び出して用いるが、分子の特徴を表現するのに適した原子を別の基準によって選択することもできる。例えば、炭素原子以外でも分子量の大きい元素を候補に入れるなどすることで、分子内部の質量分布としての重点的位置に関連する数値が算定できる。

[0043] 本実施例では、三角形を構成する原子として炭素原子を優先している。こ

れは、材料が有機物の場合は構造の基幹が炭素原子であることが理由である。最も、炭素原子を選択することは必須ではなく、精度が高くなるものを適宜選択してもよい。実際は、予測したい物性によってどの原子を使うべきかが異なると推察されるので、ユーザが適宜設定できるようにすることが望ましい。

[0044] 立体構造中の基準となる三角形を構成する3つの原子は、後に図11、図12で説明するように、基準原子番号（整理番号）を付する等して巡回方向を定めておく。図11では、原子1101が基準原子番号1、原子1102が基準原子番号2、原子1103が基準原子番号3である。

[0045] 前述のように、どのような方法で三角形を構成する原子を選定するかは予測したい材料特性によって異なる可能性があるので、複数通り計算するようにして、最も精度のよいものを選択するなどしてもよい。

[0046] 次に、他の原子について、所定の基準にしたがって並べ替える(S1002)。ここで、この基準としては、前記基準原子の作る三角形の例えば重心からの最短直線距離を計算して、それが小さい順に並べるなどの方法がとれる。他にも、原子間の相対距離のみによって決定される順番に基づいて任意の並び順を用いることができる。この並び順に沿って、他の原子について各原子に識別番号が振られる。

[0047] 次に、前記の並び順に沿って、各原子と3つの基準原子との直線距離を算定し、これを特徴量とする(S1003)。前述のように、最も面積の大きな三角形を構成する原子を基準原子とすることにより、上記直線距離を大きくとることができる、誤差を圧縮することができる。

[0048] 図11に模式図を示す。図中には3つの基準原子(1101)(1102)(1103)がそれぞれ示されており、それ以外の原子(1104)との間に矢印(1105)(1106)(1107)が描かれている。これらの矢印の長さが d_1 、 d_2 、 d_3 として示されている。この3つの長さが定まると、原子(1104)の座標は、3つの基準原子が形成する三角形の面に対する裏表の逆を除いて一意に定まる。逆に座標が定まると3つの矢印(1105)(1106)(1107)の長さが定まる。すなわち、鏡像的な対称性を除

き1対1の関係にある。

[0049] ここで、対象の原子(1104)の方から見て3つの基準原子(1101)(1102)(1103)が作る三角形の巡回方向が時計回りの場合は正、逆の場合は負となるように(図11では反時計回りなので負値)、d1、d2、d3の符号を決定すると、原子の相対的な位置関係を一意に記述したことになる。

[0050] 本実施例ではこの値を一行に並べることによって分子の空間構造を記述する空間構造特徴量とする。この値は、分子内の原子の座標値に関して、座標系の向きや原点位置に対する依存性がなく、かつ、値が定まるとそこから分子構造を逆算可能であるという材料特性の予測に好ましい特徴を持っている。

[0051] 図12は、空間構造特徴量のテーブルを示す。算出された空間構造特徴量は図12に示す形式で空間構造特徴量DB(115)に格納される。空間構造特徴量DB(115)への格納にあたっては、材料DB(112)の実験データテーブルのどのレコードに由来するかを記録するため、実験ID(701)が付与されている。これについては並び順で対応を再現するなどの方法をとってもよい。そのほかには、基準原子の作る三角形の情報を記録するため、基準原子の原子番号(1201)(1202)(1203)、および基準原子間の距離(1204)(1205)(1206)を記録する。

[0052] その後、前記の空間構造特徴量として、各原子の基準原子との距離を記述する(1207)(1208)(1209)。この際、項目数は材料DB(112)に格納されている事例の中で原子の数が最も大きなものを基準として作成し、原子の数が最大でない分子について対応する原子が存在しない項目には0を付与する。なお、この際の距離は任意の単位でよいが、本例ではオングストロームを用いる。

[0053] 以上の処理によって、材料DB(112)に新しい実験データが追加できたことになる。すなわち、材料データ投入(S310)の手順が完了する。

[0054] <3. 予測結果閲覧処理>

図13を用いて、予測結果閲覧(S320)の際の材料特性予測提示処理(S321)について説明する。最初に、材料特定予測提示部(117)は、モニタ(205)により材料特性予測表示(322)を利用者(102)に提示して、予測対象となる実験デ

ータテーブルの指定を受ける(S1301)。この際、材料DB(112)に格納されている実験データテーブルの内容を用いてもよい。

[0055] 図14にその一例を示した。図中ドロップダウンボックス(1401)には実験データテーブルの例えばファイル名が候補として表示されている。予測値更新のボタン(1402)が押されると、材料特定予測提示部(117)は材料特性予測部(116)に当該実験データテーブル(図7)のレコードのうち材料特性(702)が空白のものについて予測値により補間を実行する指示を送り、その結果が画面(1403)に表示されるようになっている。図14中、下線を施した数値は、空白のデータを補間したことを示している。

[0056] 材料特性予測部(116)は材料特定予測提示部(117)の前記補間を実行する指示を受けると、指定された実験データテーブルのデータを材料DB(112)から取得し(S1302)、その実験ID(701)を用いて対応するレコードを空間構造特徴量テーブル(1200)から取得する(S1303)。これらに対応させることで、材料特性予測に用いるデータを生成する(S1304)。

[0057] 図15に材料特性予測用データの1レコード分の構成(1500)を示す。このデータは、材料DB(112)から実験ID(701)、材料特性(702)、実験条件(704)を、空間構造特徴量DB(115)からは基準原子1情報(1201)、基準原子2情報(1202)、基準原子3情報(1203)、基準原子1と2の距離(1204)、基準原子2と3の距離(1205)、基準原子1と3の距離(1206)、各原子の基準原子1との距離d1(1207)、各原子の基準原子2との距離(1208)、各原子の基準原子3との距離(1209)を、それぞれ取得し共通の実験ID(701)をもつレコードを結合したものである。

[0058] 材料特性予測部(116)は材料特性予測用データから材料特性(702)が未実測すなわち空白のものを除き、実験ID(701)と材料特性(702)を除く項目を説明変数、材料特性(702)を目的変数とにおいて、公知の回帰分析を実行し予測関数を得る(S1305)。この手順は、予測関数を $y = f(x)$ と書くとする、 y が目的変数、 x が説明変数であり、 x を定めると y が予測できるように f の関数形を定めるということを意味する。かくして回帰モデルを生成したのち、材料特性予測部(116)は、材料特性(702)が未実測すなわち空白のものを選び

、前記 $y = f(x)$ の予測関数を用いて、材料特性(702)の予測値を算出する(S1306)。

[0059] ここで、予測関数 f の構築に用いる方法としては、公知の多変量回帰分析法を用いることができ、たとえば回帰木、ランダムフォレスト、サポートベクター回帰、ガウシアンプロセス回帰、ニューラルネットワークなど、多変量を引数に取る回帰分析法であれば、公知の高精度な非線形回帰法を用いることができる。前述の通り、この予測結果は材料特定予測提示部(117)によって画面(1403)に反映されることとなる(S1307)。なお、本実施例では説明変数として空間構造特徴量と実験条件のみをもちいたが、実際には他の何らかの量（例えば、分子量、電荷）を導出してくわえて用いてもよい。また、公知のRecursive Neural Networkの方法のように、系列的な情報を用いて予測が可能な技法を用いると、各原子の基準原子1との距離 d_1 (1207)が0のデータを扱わずに予測することができ、より高い精度が得られることもありえる。

[0060] 以上の実施例により、実験計画のスクリーニングのために、分子の空間構造をなんら特別な後処理なく、予測に組み込んで評価できるようになる。これにより予測精度が向上すると期待される。

実施例 2

[0061] 実施例 2 は、材料特性が未計測の材料についての材料特性の予測値を算出するにとどまらず、予測が最適な材料特性となる条件を探索し、それを画面に表示することで、実験の計画策定の役に立てることを特徴とする。

[0062] 図 1 6 に実施例 2 の構成の例を示す。実施例 2 が実施例 1 と異なるのは、最適材料特性候補作成部(1601)である。最適材料特性候補作成部(1601)は、前記、予測が最適な材料特性となる条件を探索する機能を備えており、実施例 1 でいう材料特性予測提示処理(S321)において利用者に提示される材料特性予測表示(322)に最適値の候補となる仮想的な分子構造を立体的に表示することができる。

[0063] 図 1 7 に最適材料特性候補作成部(1601)を用いた実施例 2 の材料特性予測提示処理(S1700)を示す。この材料特性予測提示処理(S1700)は実施例 1 (図

13) の材料特性予測提示処理(S321)と異なり、材料特性予測部(116)が予測を行った後に、予測に用いた回帰分析の回帰モデルが最適材料特性候補作成部(1601)に渡され、それを用いて材料特性が最大のものや最小のものなど、所定の条件を満たし、材料DB(112)に含まれていない材料の候補を探索する処理(S1701)が行われる。この探索では、勾配降下法や共役勾配法、遺伝的アルゴリズムなど公知の関数の最適値探索方式によって、所定の条件を満たすような説明変数を調べることとなる。

この探索の結果は材料特性予測結果画面に表示される(S1702)。

[0064] 図18にこの表示結果の一例を示す。この画面が実施例1(図14)と異なるのは、下部に前記探索された仮想的な分子の構造が表示されている点である。この所定の条件はドロップダウンボックス(1801)に表示されており、利用者(102)はこれから一つを選択できる。すると、この選択された結果に対応する分子の立体構造が画面(1802)に表示される。この表示では、図11で示したとおりの論理によって、分子を構成する各原子の座標が算定されて用いられる。なお、原子間距離がファン・デル・ワールス半径を下回った場合には結合するものとして結合を表す線を描画している。この計算の前に、例えば炭素原子に結合している原子の数が4になるようにする、などの制約に関する評価を行うなどの加工をしてもよい。

[0065] 実施例2によれば、利用者(102)が与えた候補化合物以外にも候補を選ぶことができ、利用者も気づかなかった化合物を発見できる可能性が高まると期待される。

[0066] 以上説明した実施例によれば、分子の空間構造をもとにした予測を行うために、座標系の選択に関する自由度をもたせることなく、また、逆計算が可能のように、単独の分子の空間構造と1対1で対応するような特徴量を用いる。これにより、分子の立体構造を織り込んだ形での材料特性の予測ができ、ひいてはより適切なスクリーニングにつながるものである。

[0067] すなわち、実験計画のスクリーニングのための予測評価において、分子の持つ立体構造を予測に組み込んでより高精度な予測ができるようになる。ま

た、特定の予測値を持つような分子の立体構造の逆計算ができるため、望ましい性質を持った分子がどのような形状をしているかを推算できる。その結果、実験計画がたてやすくなり、ひいては少ない実験回数でよい材料を開発可能となる。

[0068] 以上の実施例で説明したように、発明者らは、材料特性予測の精度をあげるために分子の空間構造をもとにした特徴量を使用しようとする、分子内の座標系が一意に定まらず、また、特徴量から分子の形を逆計算できないため、最適解に対応する分子がわかりにくいという問題点に着目した。このため、実施例では、分子の立体構造を表現する特徴量として、分子内部の最も主要な3つの原子を選択し、その原子との直線距離をもって特徴量とするという手法を提案するものである。これにより、座標系の選択に関する自由度をもたせることなく分子の空間構造を規定し、分子の立体構造に基づいた材料特性の予測をすることが可能になる。

符号の説明

[0069] 101 材料特性予測装置
113 三次元分子構造算定部
114 空間構造特徴量算定部
115 空間構造特徴量DB
S311 材料DB更新処理
S321 材料特性予測提示処理
S403 空間構造特徴量生成処理
701 実験ID
1200 空間構造特徴量テーブル

請求の範囲

- [請求項1] 材料特性の予測を行うためのシステムであって、
材料の構造式から分子を構成する原子の位置座標を算出する機能を備えた三次元分子構造算定部と、
前記三次元分子構造算定部の算出した原子の位置座標をもとに、3つの原子を選び出して三角形を構築し、前記3つの原子と他の原子との距離を算出して空間構造特徴量とする機能を持つ空間構造特徴量算定部と、
前記空間構造特徴量算定部の生成した空間構造特徴量を説明変数として材料特性を予測する材料特性予測部と、
を持つことを特徴とする材料特性予測システム。
- [請求項2] 請求項1に記載のシステムであって、
前記材料特性予測部の予測する材料特性が所定の基準を満たすような空間構造特徴量の条件を探索し、当該条件を満たす原子の位置座標から仮想的な分子構造を構成する機能を持った最適材料特性候補作成部、
を持つことを特徴とする材料特性予測システム。
- [請求項3] 請求項1に記載のシステムであって、
前記空間構造特徴量算定部が、前記三角形を構成する3つの原子を選択する基準として、当該三角形の面積が最も大きくなる組み合わせを選択すること、
を特徴とする材料特性予測システム。
- [請求項4] 請求項1に記載のシステムであって、
前記空間構造特徴量算定部が、前記三角形を構成する3つの原子を選択する基準として、
炭素原子を優先的に選択すること、
を特徴とする材料特性予測システム。
- [請求項5] 請求項1に記載のシステムであって、

前記空間構造特徴量算定部が、前記3つの原子と他の原子との距離を算出する際に、前記三角形の形成する面に対する方向にもとづき、正負の符号を変化させて空間構造特徴量を算定すること、
を特徴とする材料特性予測システム。

[請求項6]

請求項1に記載のシステムであって、
前記空間構造特徴量算定部が、前記三角形を構成する3つの原子を選択する際に、前記3つの原子の巡回方向を規定すること、
を特徴とする材料特性予測システム。

[請求項7]

請求項1に記載のシステムであって、
前記空間構造特徴量算定部の生成した空間構造特徴量を空間構造特徴量データベースとして格納し、
前記空間構造特徴量データベースは、前記三角形を構成する3つの原子の順序を特定する情報と、前記三角形を構成する3つの原子相互の距離と、前記三角形を構成する3つの原子夫々と他の原子との距離を含むこと、
を特徴とする材料特性予測システム。

[請求項8]

材料の構造式を入力とし、該材料の構造式から分子を構成する原子の位置を算出する三次元分子構造算定処理と、
算出した原子の位置をもとに、3つの原子を選択して三角形を構築し、前記3つの原子と他の原子との距離を算出して空間構造特徴量を得る空間構造特徴量算定処理と、
を実行する情報処理方法。

[請求項9]

前記三角形を構成する3つの原子を選択する基準として、当該三角形の面積が最も大きくなる組み合わせを選択する、
請求項8記載の情報処理方法。

[請求項10]

前記三角形を構成する3つの原子を選択する基準として、炭素原子を優先的に選択する、
請求項8記載の情報処理方法。

- [請求項11] 前記三角形を構成する3つの原子を選択する際に、前記3つの原子の巡回方向を規定すること、
請求項8記載の情報処理方法。
- [請求項12] 前記3つの原子と他の原子との距離を算出する際に、前記3つの原子の巡回方向に基づいて、距離の正負の符号を変化させて空間構造特徴量を算定すること、
請求項11記載の情報処理方法。
- [請求項13] 前記空間構造特徴量算定処理で得た空間構造特徴量を空間構造特徴量データベースとして格納し、
前記空間構造特徴量データベースは、前記三角形を構成する3つの原子の順序を特定する情報と、前記三角形を構成する3つの原子相互の距離と、前記三角形を構成する3つの原子夫々と他の原子との距離を含む、
請求項8記載の情報処理方法。
- [請求項14] 前記空間構造特徴量算定処理で得た空間構造特徴量を説明変数として材料特性を予測する材料特性予測処理を実行する、
請求項8記載の情報処理方法。
- [請求項15] 前記材料特性予測処理は予測関数を用いて予測を行い、前記予測関数は、前記空間構造特徴量を説明変数、材料特性を目的変数とにおいて、回帰分析を実行して得たものである、
請求項14記載の情報処理方法。