

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号

特開2023-149303
(P2023-149303A)

(43)公開日 令和5年10月13日(2023. 10. 13)

(51)Int. Cl.
G 0 1 N 21/71 (2006. 01)

F I
G 0 1 N 21/71

テーマコード(参考)
2 G 0 4 3

審査請求 未請求 請求項の数 6 O L (全 17 頁)

(21)出願番号 特願2022-57807(P2022-57807)

(22)出願日 令和4年3月31日(2022. 3. 31)

特許法第30条第2項適用申請有り 第7回次世代イニシアティブ廃炉技術カンファレンス(NDEC-7) 2022年3月11日開催

(71)出願人 505089614

国立大学法人福島大学
福島県福島市金谷川1番地

(71)出願人 505374783

国立研究開発法人日本原子力研究開発機構
茨城県那珂郡東海村大字舟石川765番地1

(71)出願人 000140627

株式会社化研
茨城県水戸市堀町1044番地

(74)代理人 110001922

弁理士法人日峯国際特許事務所

(72)発明者 藤本 勝成

福島県福島市金谷川1番地 国立大学法人
福島大学内

最終頁に続く

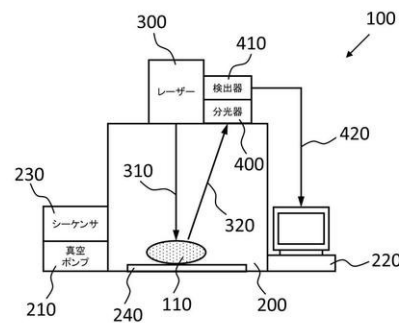
(54)【発明の名称】化合物同定方法及び化合物同定装置

(57)【要約】

【課題】未知試料に含まれる金属元素や化合物を分類・推定する方法を提供する。

【解決手段】本発明である化合物同定方法は、想定される化合物種の標準試料についてスペクトルデータを因子分解して因子及び前記因子の影響度を表す因子スコアを導出し、前記化合物種を決定木において前記因子スコアで分類するためのスコア閾値及び前記因子スコアを前記決定木にかけると因子適用順を予め学習・設定しておき、未知試料のスペクトルデータに対して前記標準試料の因子を用いて前記未知試料における因子スコアを算出し、前記因子適用順にしたがって前記決定木で前記未知試料の因子スコアを前記スコア閾値によって分類していくことにより、前記未知試料に含まれる化合物種を同定する、ことを特徴とする。

【選択図】図1



【特許請求の範囲】**【請求項 1】**

想定される化合物種の標準試料についてスペクトルデータを因子分解して因子及び前記因子の影響度を表す因子スコアを導出し、前記化合物種を決定木において前記因子スコアで分類するためのスコア閾値及び前記因子スコアを前記決定木に付ける因子適用順を予め学習・設定しておく、

未知試料のスペクトルデータに対して前記標準試料の因子を用いて前記未知試料における因子スコアを算出し、前記因子適用順にしたがって前記決定木で前記未知試料の因子スコアを前記スコア閾値によって分類していくことにより、前記未知試料に含まれる化合物種を同定する、

ことを特徴とする化合物同定方法。

【請求項 2】

前記スペクトルデータは、測定値が推定値を上回ったときより下回ったときに誤差が厳しく評価されるように罰則項付き非対称最小二乗法で近似させたベースラインに補正される、

ことを特徴とする請求項 1 に記載の化合物同定方法。

【請求項 3】

前記因子分解は、前記スペクトルデータの波長次元 K と試料数 N からなる K 行 N 列の行列を、指定された M 個の非負値基底ベクトルで分解して得られた前記因子の影響度を前記因子スコアとして導出する、

ことを特徴とする請求項 1 又は 2 に記載の化合物同定方法。

【請求項 4】

前記決定木は、前記化合物種を分類するとき、不純度を用いたコスト関数が最小値となるような因子とスコア閾値を導出しておく、

ことを特徴とする請求項 1 乃至 3 の何れか一に記載の化合物同定方法。

【請求項 5】

請求項 1 乃至 4 の何れか一に記載の化合物同定方法を実行する、

ことを特徴とする化合物同定装置。

【請求項 6】

前記試料にレーザーを照射する照射装置と、

前記レーザーが照射された前記試料から放出された光を分解する分光器と、

前記分光器で生成されたスペクトルデータを取得する検出器と、を有し、

前記検出器から入力されたスペクトルデータを用いて前記試料に含まれる化合物を同定する、

ことを特徴とする請求項 5 に記載の化合物同定装置。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、未知試料に含まれる金属元素や化合物を分類・推定する方法に関する。

【背景技術】**【0002】**

原子力発電所の廃炉に向けて、汚染水対策、使用済燃料プールからの燃料取出し、燃料デブリ取出し、廃棄物対策などが必要である。炉心溶融した原子炉では、溶解した燃料、制御棒などが混ざり合って燃料デブリとして固化していると考えられる。制御材に炭化ホウ素 (B_4C) を用いている原子炉では、酸化物の約 2 倍の硬度を持つホウ化物が生成されているため、金属、酸化物及びホウ化物を判別しながら燃料デブリを取り出すのが効率的である。なお、ホウ化物には、含有されているホウ素濃度と硬度に相関があることも確認されている。燃料デブリ取出方法の検討のため、未知試料の材料特定、材料硬度特定が必要である。また、未知試料内の材料分布、硬度分布の作成も検討されている。

【0003】

10

20

30

40

50

試料中の元素の定性・定量分析を行う方法として、誘導結合プラズマ発光分析（ICP-AES）や、レーザーアブレーション誘導結合プラズマ質量分析（LA-ICP-MS）がある。試料に外部から何らかのエネルギーを与えると、試料に含まれる元素は、元素固有の波長の光を放出するので、放出された光の波長と量を測定することで元素分析する。ICP-AESでは、干渉により正確に分析できないことがある。なお、特許文献1に記載されているように、試料に含まれる異物を正確かつ容易に解析することができる異物解析装置の発明も開示されている。

【0004】

また、固体、液体、気体等の状態に関わらず元素分析をする方法として、レーザー誘起ブレイクダウン分光法（LIBS）がある。LIBSは、高エネルギーのパルスレーザーを励起源とする原子発光分析法であり、複雑な前処理の工程を経ずに短時間で分析可能で、軽元素（ホウ素、酸素）から重元素（金属）まで同時計測が可能であることから、高レベル廃棄物の元素情報取得、硬度計測、ホウ素分布情報を基にした臨界管理等に応用が期待されている。なお、特許文献2に記載されているように、金属を主成分とする材料中の酸素やホウ素の組成や材料の硬度を算出する方法の発明も開示されている。

【0005】

LIBSでは、ナノ秒以下のパルス幅を持つピークパワーの高いレーザーを測定対象に照射すると数万のプラズマが瞬時形成される。この際、測定対象物はプラズマ化により原子化・励起され、プラズマ消失と共に励起された原子は基底状態へ緩和され、元素固有の蛍光発光を伴う。これを計測し、測定対象元素の同定と定量を行う。レーザー出力の変動や測定対象を構成する元素により生成されるプラズマが、時間・空間的に微小に変動を伴うため、結果として測定値にも変動が生じてしまう。

【0006】

LIBSであれば、プラズマ状態（時間・空間等）によるバックグラウンド、エネルギー弁別処理であれば、放射線と物質との相互作用により生じる散乱成分等が、計測値に大きく影響を与える。レーザー、X線や中性子等の放射線をプローブとして用いる分光計測においては、その各種要因や影響の大小はあるものの、それらをキャンセル又は補正する機能が、精度の高い計測をするために望まれている。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特許第6507757号公報

【特許文献2】特開2020-067337号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、未知の固体試料を分析する際に、多くの元素を含む合金や溶融物では、多くの発光波長を有するスペクトルデータが得られ、元素や化合物を判別することが困難な場合がある。例えば、鉄（Fe）やジルコニウム（Zr）の発光線は、非常に多く重なり合うと定量分析に支障をきたすことがある。

【0009】

また、レーザーのみならず、X線や中性子等の分光器を用いた計測では、分光器を閉止した条件で暗電流の計測を行い、その影響を除去している。測定環境（空気環境下や不活性環境下等）を除去するために、ブランクデータ（サンプルの無い状態での計測）を基に、その影響を除去しているものもある。しかし、これらの影響は装置機能に付加されていなかったり、類似したサンプルが存在しなかったりする場合などには適用できない。

【0010】

加えて、レーザーを用いた発光分析法（LIBSやLA-ICP-MS）や自己遮蔽を持つ中性子等では、試料固有の誤差を評価する必要があるが、中性子やプラズマ発光に関しては、照射エネルギーがコントロールできない場合が多い。

10

20

30

40

50

【 0 0 1 1 】

特に、LIBS測定では、雰囲気ガスによりスペクトルのベースライン（バックグラウンドと思われる点を繋いでいき、それを連続したラインとして示したもの）が変動することがある。例えば、アルゴン（Ar）雰囲気では発光感度が高く、ベースラインが広範囲の波長域でブロードになるホワイトノイズが発生する場合がある。また、190nmから900nmまでの波長の光を検出するため、複数の検出器を用いることになるが、光学系の特性で検出できる波長域が変わるため、検出器が切り替わる波長でベースラインにズレが生じてしまう。

【 0 0 1 2 】

このように、分光分析、放射線分析等のスペクトル分析におけるバックグラウンドノイズ成分（分光器の暗電流等の変動誤差）の軽減、各種分析における被測定対象物に対する光や放射線によるプローブ条件の適正化（被測定物に対するプローブによる損傷の緩和）、各種化合物（金属やセラミックス等を含む）の同定などが求められている。

10

【 0 0 1 3 】

そこで、本発明は、未知試料に含まれる金属元素や化合物を分類・推定する方法を提供することを目的とする。また、金属元素や化合物を分類・推定するに際して、試料固有の誤差を評価し、予め補正しておくことを別の目的とする。

【 課題を解決するための手段 】**【 0 0 1 4 】**

上記の課題を解決するために、本発明である化合物同定方法は、想定される化合物種の標準試料についてスペクトルデータを因子分解して因子及び前記因子の影響度を表す因子スコアを導出し、前記化合物種を決定木において前記因子スコアで分類するためのスコア閾値及び前記因子スコアを前記決定木にかける因子適用順を予め学習・設定しておき、未知試料のスペクトルデータに対して前記標準試料の因子を用いて前記未知試料における因子スコアを算出し、前記因子適用順にしたがって前記決定木で前記未知試料の因子スコアを前記スコア閾値によって分類していくことにより、前記未知試料に含まれる化合物種を同定する、ことを特徴とする。

20

【 0 0 1 5 】

前記化合物同定方法において、前記スペクトルデータは、測定値が推定値を上回ったときより下回ったときに誤差が厳しく評価されるように罰則項付き非対称最小二乗法で近似させたベースラインに補正される、ことを特徴とする。

30

【 0 0 1 6 】

前記化合物同定方法において、前記因子分解は、前記スペクトルデータの波長の標本数（波長次元） K と試料数 N からなる K 行 N 列の行列 Y を、指定された M 個の非負値基底ベクトルで分解して得られた前記因子の影響度を前記因子スコアとして導出する、ことを特徴とする。

【 0 0 1 7 】

前記化合物同定方法において、前記決定木は、前記化合物種を分類するときに、不純度を用いたコスト関数が最小値となるような因子とスコア閾値を導出しておく、ことを特徴とする。

40

【 0 0 1 8 】

前記化合物同定方法において、前記因子スコアをマッピングし、前記未知試料の成分や組成比を推定する、ことを特徴とする。

【 0 0 1 9 】

また、本発明である化合物同定装置は、前記試料にレーザーを照射する照射装置と、前記レーザーが照射された前記試料から放出された光を分解する分光器と、前記分光器で生成されたスペクトルデータを所定の範囲の波長ごとに分けて取得する複数の検出器と、を有し、前記検出器から入力されたスペクトルデータを用いて前記試料に含まれる化合物を同定する、ことを特徴とする。

【 発明の効果 】

50

【0020】

本発明によれば、未知試料に含まれる金属元素や化合物を分類・推定することができる。事前に既知物質のスペクトルデータを複数の因子に分解し、決定木により元素や化合物の分類方法を学習させておくことで、未知試料の成分の分類に利用することができる。

【0021】

また、試料固有の誤差を評価し、予め補正しておくことで、金属元素や化合物を精度良く分類・推定することができる。なお、後付けで補正することが可能で、ホワイトノイズも含め誤差を除去することができる。

【図面の簡単な説明】

【0022】

10

【図1】本発明である化合物同定方法における測定装置の概観図である。

【図2】本発明である化合物同定方法の前処理段階および学習段階の流れを示すフローチャートである。

【図3】本発明である化合物同定方法におけるスペクトルデータの補正を説明する図である。

【図4】本発明である化合物同定方法におけるスペクトルデータの補正を説明する図である。

【図5】本発明である化合物同定方法におけるスペクトルデータの因子分解を説明する図である。

【図6】本発明である化合物同定方法におけるスペクトルデータの因子分解を説明する図である。

20

【図7】本発明である化合物同定方法における決定木による分類を説明する図である。

【図8】本発明である化合物同定方法における決定木による分類を説明する図である。

【図9】本発明である化合物同定方法における決定木による分類を説明する図である。

【図10】本発明である化合物同定方法の運用段階の流れを示すフローチャート図である。

【図11】本発明である化合物同定方法で分類した結果を示す図である。

【発明を実施するための形態】

【0023】

以下に、本発明の実施形態について図面を参照して詳細に説明する。なお、同一機能を有するものは同一符号を付け、その繰り返しの説明は省略する場合がある。

30

【0024】

本発明は、可視光 - 赤外光やX線・ガンマ線・中性子等の放射線等を用いる分光分析の計測系に関して、計測系の安定性に起因するバックグラウンド変動の低減処理、計測対象のスペクトル同定処理の2つの処理を対象としたアルゴリズムに関するものである。本アルゴリズムは、計測データ（予測データを含む）を学習データとする機械学習方法およびその結果がもたらすデータ処理系の構成からなり、実施により計測精度やロバスト性の向上が期待される。また、ソフトウェア処理として実施可能なため、既存の分光分析装置に具備される解析プログラムへの実装性が高い。

【0025】

40

LIBS（レーザー誘起ブレークダウン分光法）による元素の同定及び定量化の分析精度を高度化するために、元素発光スペクトルに加えてプラズマ状態（時間・空間等）を特徴付けるパラメータ（参照パラメータ）を同時に計測し、双方の相関性を補正処理に用いる。具体的には、測定対象とする原子発光スペクトルの強度や形状と強く相関している参照パラメータを設定する。従来は、習熟経験が高い測定者が変動成分や特性を把握した上で設定していた。高い分析精度を得るためには、スペクトルの変動メカニズムを理解した評価が必要である。

【0026】

参照パラメータの設定には、機械学習を使う。従来手法では、測定対象のスペクトルを一定の幅を有するピークとしてサーチするのに対し、本発明では、人間では取扱いが困難

50

なビッグデータ（各種分析装置等から得られた波長 - カウント値（光強度値）がセットとなった全スペクトル成分）に機械学習手法（NMF：非負値行列因子分解）を適用し、参照パラメータを構築する。シグナルピークではなく、全スペクトル成分を用いることで、未知試料の同定の際に、スペクトルの強度や形状等も含む評価が可能である。

【0027】

本発明は、オーケストラの中からバイオリンやピオラ等の特定音を抽出する手法や、雑踏の中から特定の個人を抽出する手法に着想を得て、測定装置に展開したものである。本発明のアルゴリズムとしては、前処理段階（プレプロセッシングフェーズ）で加工した標準試料のスペクトルデータを、学習段階（トレーニングフェーズ）でNMF及び決定木を用いて学習させ、運用段階（運用フェーズ）で未知試料に適用している。

10

【実施例1】

【0028】

図1は、本発明である化合物同定方法における測定装置の概観図である。測定装置100は、チャンバー200、照射装置300、分光器400、検出器410、及び制御装置220などを備え、LIBS測定を行う。

【0029】

チャンバー200は、真空容器などであり、真空ポンプ210を用いて減圧された内部に試料110が載置される。試料110としては、主に、学習段階においては、標準試料110aなどの既知試料が使用され、運用段階においては、測定・分析対象の未知試料110bが使用される。標準試料110aは、金属元素や化合物など予め組成が判明しているもの、未知試料110bは、原子力発電所から運び込まれた燃料デブリなど組成が不明なものである。

20

【0030】

照射装置300は、レーザー発振器を備えたレーザーヘッド等であり、チャンバー200内の試料110にパルスレーザー等の照射光310を照射する。照射光310としては、YAG（イットリウム、アルミニウム、ガーネット）にNd（ネオジウム）を添加した結晶を媒体として用いた波長1064nmのNd：YAGレーザー等を使用する。

【0031】

照射光310を照射することにより試料110の表面をプラズマ化し、それが基底状態に戻る際に放出された放出光320を検出する。放出光320をコリメーター等で平行にした上で分光器400に入射させる。分光器400は、プリズムや回折格子などであり、光を波長ごとに分解し、強度分布を表したスペクトルデータ420を複数の検出器410を並置して所定の範囲の波長ごとに分けて取得させ、制御装置220に送る。

30

【0032】

制御装置220は、コンピュータ等であり、各装置を制御するとともに、取得したスペクトルデータ420から試料110の分類・分析を行う。その他、必要に応じて、シーケンサ230、電動のステージ240、冷却水循環設備などを備えても良い。

【0033】

例えば、試料110を載せたステージ240をシーケンサ230の制御で移動させて多地点を走査させる。ステージ240の位置情報（X-Y-Z軸の座標情報）からレーザーの照射位置を取得するとともに、変位計で試料110の形状情報（高さ情報）を取得し、制御装置220で座標をマッピングして立体的に形状と強度情報を可視化する。

40

【0034】

図2は、化合物同定方法の前処理段階および学習段階の流れを示すフローチャートである。前処理段階として、標準試料準備（標準試料の化合物種の設定）L1、スペクトルデータ取得（ベースライン推定）L2、学習段階として、NMF実行（標準試料の因子と因子スコアの取得）L3、及び決定木ルール生成（スコア閾値と因子適用順の設定）L4のステップを有する。

【0035】

前処理段階及び学習段階においては、想定される化合物種を設定して標準試料110a

50

を準備し、取得した標準試料 1 1 0 a のスペクトルデータ 4 2 0 を因子分解して標準試料 1 1 0 a の因子及び因子スコアを導出し、化合物種を決定木において因子スコアで分類するためのスコア閾値と、因子を決定木にかける因子適用順を、決定木のルールとして学習し、予め設定しておく。

【 0 0 3 6 】

前処理段階に係る工程は、図 2 の標準試料準備 L 1、スペクトルデータ取得 L 2 からなる。標準試料準備 L 1 ステップでは、測定対象の未知試料 1 1 0 b の成分として想定される金属元素や化合物などの化合物種を設定しておき、その金属元素や化合物を標準試料 1 1 0 a として準備する。標準試料 1 1 0 a は、未知試料 1 1 0 b に含まれる化合物種を同定する際に、どの因子がどの程度の影響を及ぼしているかを見るために予め用意しておくデータとなる。

10

【 0 0 3 7 】

スペクトルデータ取得 L 2 ステップでは、検出器 4 1 0 から送られてきたスペクトルデータ 4 2 0 (事前に記憶装置等に保存していたものを含む) を制御装置 2 2 0 に入力する。スペクトルデータ 4 2 0 にはノイズ等が含まれ、ベースライン 4 3 0 (基線) が乱れているので、制御装置 2 2 0 においてベースライン推定を行う。ベースライン推定では、ベースライン 4 3 0 が平滑となるようにスペクトルデータ 4 2 0 を補正する。

【 0 0 3 8 】

次に、上述した前処理段階に係る工程について、L I B S 測定を行って得た実例を図 3 から図 4 を参照して説明する。図 3 及び図 4 は、化合物同定方法におけるスペクトルデータの補正を説明する図である。

20

【 0 0 3 9 】

図 3 (a) は、L I B S 測定系において複数の測定波長範囲を分割し、各々に感度が高い分光器を用い広帯域化 (1 9 0 n m から 9 0 0 n m) を図り、取得したスペクトルである。図 3 (a) に示すように、検出器 4 1 0 の波長範囲の境界でベースライン 4 3 0 にズレを生じていることが分かる。ズレを生じることで、検出器 4 1 0 の感度や暗電流に起因するベースライン 4 3 0 に不連続性が発生することになる。

【 0 0 4 0 】

不連続性が発生した波長域では、計測結果に不確かな要因を多く含み、結果の信頼性が得られない。また、不連続部分にスペクトルがある対象の場合、正確な面積を算出することができず、場合によってはスペクトルの判別ができなくなる。

30

【 0 0 4 1 】

また、図 3 (b) は、図 3 (a) のスペクトルデータ 4 2 0 の一部を拡大したスペクトルである。図 3 (b) に示すように、ディテクタの分解能に対して最適なスムージングを適用しない限り、必要以上に調整されたベースラインになってしまう。

【 0 0 4 2 】

そこで、発明者らは、図 3 (c) に示す、非対称重みペナルティ最小二乗スムージング (罰則項付き非対称最小二乗法) により、複数の検出器 4 1 0 の波長範囲の境界におけるベースライン 4 3 0 のズレ (不連続性) を補正することを考案した。罰則項付き非対称最小二乗法は、測定値 y_i に対するベースライン 4 3 0 の推定値 z_i の好ましさ (良さ) を、測定値と推定値の差が最小とするだけでなく、平滑となるように評価関数を最小化することにより、好ましいベースライン 4 3 0 を導出するものである。

40

【 0 0 4 3 】

評価の観点としては、まず、二乗誤差 $(y_i - z_i)^2$ の観点から、 $y_i > z_i$ の際には誤差を非常に厳しくし、 $y_i < z_i$ の際には誤差を非常に甘くする非対称な評価を行い、さらに、 z_i の変動の観点から、変動の変化に対して厳しい評価を行う。 w_i は、非対称重みであり、 $y_i > z_i$ ならば p となり、 $y_i < z_i$ ならば $(1 - p)$ となる。 p ($p \in [0, 1]$: p は非常に小さい値) は、測定値が推定値を上回ったかどうかで w_i を変える役割を持つ。 μ は、正則化パラメータであり、推定値の変動 (非平滑度) $(\Delta^2 z_i)^2$ が罰則項となる。なお、 Δ^2 は 2 階微分演算子を表し、 w_i 、 μ は設計者が

50

経験的に定めるハイパーパラメータである。パラメータを変化させ、最適解を探っていけば良い。

【 0 0 4 4 】

これにより、スペクトルデータ 4 2 0 は、測定値が推定値を上回ったときよりも、測定値が推定値を下回ったときに誤差が厳しく評価されるように罰則項付き非対称最小二乗法で近似させたベースライン 4 3 0 に補正される。

【 0 0 4 5 】

図 4 は、図 3 (c) の補正により、図 3 のスペクトルデータ 4 2 0 を補正したスペクトルである。なお、図 4 (a) は、図 3 (a) のスペクトルデータ 4 2 0 を補正したスペクトルであり、図 4 (b) は、図 4 (a) のスペクトルデータ 4 2 0 の一部を拡大したスペクトルである。

10

【 0 0 4 6 】

図 4 (a) に示すように、補正後の平滑化されたベースライン 4 3 0 により、図 3 (a) のスペクトルの不連続性が解消され、バックグラウンド除去によりスペクトルデータ 4 2 0 の形状や強度を識別することが容易になることが分かる。これにより、元素同定及び定量性の観点で分析精度が向上し、今まで困難であったスペクトル評価が可能となる。なお、本実施例である L I B S 測定に限られず、分光器 4 0 0 を使う全ての測定装置に適用可能である。

【 0 0 4 7 】

また、非対称重みペナルティ最小二乗スムージングだけでなく、サンプルが無い条件でのベースライン 4 3 0 からパターン認識による除去等を適用することで、各種分析装置固有の性能向上やサンプル依存性の低減を図る。ラマン分光法や L A - I C P - M S 等のレーザー光源を用いた分析手法のみならず、分光器 4 0 0 を使う全ての測定装置 1 0 0 に適用可能である。

20

【 0 0 4 8 】

パターン認識による除去の例としては、試料が無い条件（空打ち）で測定環境のみを測定した際に得られたスペクトルをパターン認識させ、ベースラインで削除するものがある。一般的には、試料が無い条件を単純に差し引くことで行っているが、パターン認識を駆使すると測定毎に上下するベースラインに合わせた補正が可能である。

【 0 0 4 9 】

図 5 及び図 6 は、化合物同定方法におけるスペクトルデータの因子分解を説明する図である。NMF 実行 L 3 ステップでは、制御装置 2 2 0 において、補正後のスペクトルデータ 4 2 0 に対して、非負値行列因子分解（NMF）を行う。

30

【 0 0 5 0 】

NMF では、機械学習により重なり合った発光スペクトルデータ全体を指定した因子数の因子に分解し、因子スコアと物質名のペアで表現されているサンプルデータ群に対して、決定木でどの因子に注目してどんな閾値によって峻別していけば、正しく物質の分類ができるのかを探し出すアルゴリズムを使用し、関連する因子に分けることで未知試料 1 1 0 b から元素や化合物を分類する。

【 0 0 5 1 】

事前に既知物質のスペクトルデータ 4 2 0 を次元削減手法により、低次元の因子スコアベクトルにスペクトルを変換し、決定木により元素や化合物などを分類するアルゴリズムを学習させておくことで、未知試料 1 1 0 b の中の成分を推定する。

40

【 0 0 5 2 】

学習段階では、図 5 (a) に示すスペクトルデータ 4 2 0 (K 次元のベクトル群) を一纏めにし、NMF を使って M 個の因子 4 4 0 に分けて、決定木で扱いやすいように低次元化する。NMF は、スペクトルデータが 0 と正の値の行列表現ができれば、データの種類によらず適用可能である。

【 0 0 5 3 】

図 5 (b) に示すように、N 個の測定データ Y ($y_1 \sim y_N$) を近似的に因子 H と因子

50

スコア U に分解する。測定データ y_i は、 K 次元（例えば、 $K = 11562$ 種類の波長）のベクトルデータ（強度）であり、 $K \times N$ （ K 行 N 列）の行列で表される。因子 H は、 K 次元のベクトルデータに共通する何らかの因子を、 M 個（例えば、 $M = 3$ 個の因子）の基底ベクトルとして表したものであり、 $K \times M$ （ K 行 M 列）の行列で表される。因子スコア U は、 N 個の測定データ Y と M 個の因子 H に関する結合係数であり、 $M \times N$ （ M 行 N 列）の行列で表される。なお、 M はハイパーパラメータである。

【0054】

標準試料 110a の場合、スペクトルデータ 420 の波長の標本数（波長次元）と試料数からなる行列を非負値で因子分解し、得られた因子（例えば、第 1 因子～第 3 因子）をそれぞれ標準試料 110a の因子 440 とし、標準試料 110a における各因子 440 の影響度を表す係数がそれぞれ因子スコア 450 として導出される。因子スコア 450 は、測定データがどの因子 440 の影響を大きく受けているかを示すものである。

10

【0055】

LIBS を用いた例として、ホウ化物を中心に 4 種類の化合物種を設定し、標準試料 110a としてそれぞれ 37 個（ $N = 37 \times 4$ ）のスペクトルデータ 420 を抽出し、高次元（各スペクトルの波長次元数 $k = 11562$ ）のサンプルとした。それを NMF により低次元化（参照パラメータとして特徴量の異なる 3 つの因子 440 と因子スコア 450）した結果を示す。図 6 に示すように、各因子 440 に因子スコア 450 の影響を考慮した上で、3 つの因子の線形和により、測定したスペクトルデータ（図 5 (a)）を表現可能である。

20

【0056】

なお、因子 440 は、波長ごとの強度のスケールを揃えるために、平均値と標準偏差でスペクトルデータ 420 を標準化しても良い。標準化しない場合は、そのステップを省略しても良い。なお、レーザー照射によるサンプルの壊れやすさや放射線計測におけるサンプル特有の自己遮蔽いがあるので、標準化することでこれらを緩和することができる。

【0057】

図 7～図 9 は、化合物同定方法における決定木による分類を説明する図である。決定木ルール生成 L4 ステップでは、制御装置 220 において、標準試料 110a の因子スコア 450 を特徴量として決定木 500 を作成し、標準試料 110a の因子 440 を適用するとき因子スコア 450 によって分類するためのスコア閾値 460 と、標準試料 110a のどの因子 440 から適用して決定木 500 で分類していくか因子適用順を設定する。

30

【0058】

決定木 500 は、NMF で得た参照パラメータを元に「True」又は「False」を判断する分類方法であり、学習データに対して、着目した因子 440（参照パラメータ）とその判断に用いる因子スコア 450 のスコア閾値 460 を用いて、物質を正しく分類する探索アルゴリズムである。

【0059】

学習段階にある NMF と決定木 500 の組合せは、動的光散乱や誘導結合プラズマ質量分析装置（ICP-MS）等のスペクトルデータ 420 がある解析手法はもちろん、「0」と「正の値」の行列表現記載ができれば、データの種類によらず適用することができる。TEM（透過電子顕微鏡）や SEM（走査電子顕微鏡）及びそれに付随した EDX（エネルギー分散型 X 線分析）等の画像データ等にも展開が期待できる。

40

【0060】

図 7 (a) に示すように、決定木 500 で条件分岐させるために、標準試料 110a の因子スコア 450 について不純度を計算するコスト関数 J を使用し、コスト関数 J が最小となるようなスコア閾値 t_k ($1 \leq k \leq m$) を導出する。なお、コスト関数 J はジニ不純度などを用いて算出しており、 k は因子番号、 m は因子 440 の数、 m_{left} は一方に分ける因子（左サブセット）の数、 m_{right} は他方に分ける因子（右サブセット）の数、 G_{left} は左サブセットの不純度、 G_{right} は右サブセットの不純度である。

【0061】

50

なお、サブセット中に存在する化合物が1種類だけであれば、純粋なサブセット、複数種類の化合物が混合されていれば、不純なサブセットとなる。このとき、サブセットの不純度は、サブセット中に含まれる各化合物の比率に基づく、ジニ不純度や交差エントロピーによって表される。

【0062】

図7(b)に示すように、化合物種(雰囲気)がArとN₂の場合に、横軸側の因子440について因子スコア450のスコア閾値460(例えば、114.155)が導出され、決定木500において、因子スコア450がスコア閾値460より小さければ因子440は左サブセット(N₂と推定)に分類され、因子スコア450がスコア閾値460より大きければ因子440は右サブセット(Arと推定)に分類される。

10

【0063】

図8(a)は、LIBSで測定したホウ化物スペクトル(CrB系、FeB系、NiB系、ZrB系の4つの化合物種)に対して、因子440に分解し、特徴量(因子スコア450)を抽出して決定木500を作成する例を示す。因子440の因子適用順は、第2因子、第3因子、第1因子の順になったとする。

【0064】

決定木500のノードN1では、第2因子における因子スコア450のスコア閾値460が108.615と導出され、図8(b)に示すように、スコア閾値460より大きいものがCrとBを含む化合物として分類され、それ以外はノードN2で分類される。

【0065】

決定木500のノードN2では、第3因子における因子スコア450のスコア閾値460が65.837と導出され、図9(a)に示すように、スコア閾値460より大きいものがNiとBを含む化合物として分類され、それ以外はノードN3で分類される。

20

【0066】

決定木500のノードN3では、第1因子における因子スコア450のスコア閾値460が117.886と導出され、図9(b)に示すように、スコア閾値460より大きいものがZrとBを含む化合物として分類され、それ以外はFeとBを含む化合物として分類される。

【0067】

比較的スペクトルの多いジルコニウム(Zr)や鉄(Fe)よりもクロム(Cr)が決定木500の始めに分類されており、決定木500では人間の予想とは異なり、より正確に分類されている。これは、ZrやFeが干渉の影響を伴うピークを多数含むことも影響していると考えられる(従来手法のように、単純に干渉の無いシグナルピークによる分類だけだと正答率は低下する)。

30

【0068】

制御装置220において、標準試料110aから得られた因子440と因子スコア450、さらに決定木500のルールとして因子スコア450のスコア閾値と因子440の因子適用順をデータベース等の記憶装置に蓄積し、化合物同定指標を作成する。学習段階では、どのような標準試料110aからのデータか、因子スコア450がそれぞれの因子440でどのようにになっているかを把握し、運用段階において、未知試料110bを分類する際の化合物同定指標として利用する。

40

【0069】

図10は、化合物同定方法の運用段階の流れを示すフローチャート図である。運用段階として、未知試料測定S1、スペクトルデータ補正S2、因子分解処理S3、決定木分類処理S4のステップを有する。標準試料110aの各因子440とその因子スコア450が予め化合物同定指標として用意されており、未知試料110bの各因子440についての因子スコア450で振り分けることにより、未知試料110bをクラス分けする。

【0070】

運用段階においては、未知試料110bのスペクトルデータ420に対して、標準試料110aの因子440を用いて未知試料110bにおける因子スコア450を算出し、因

50

子440の因子適用順にしたがって決定木500で未知試料110bの因子スコア450をスコア閾値460によって分けていき、未知試料110bのクラスを推定することで、未知試料110bに含まれる化合物種を同定する。

【0071】

未知試料測定S1ステップでは、測定装置100において未知試料110bにレーザーを照射し、検出器410から送られてきたスペクトルデータ420を制御装置220に入力する。

【0072】

スペクトルデータ補正S2ステップでは、制御装置220において、未知試料110bのスペクトルデータ420に対しても、スペクトルデータ取得L2ステップと同様にベースライン補正を行う。

【0073】

因子分解処理S3ステップでは、制御装置220において、未知試料110bの補正済みのスペクトルデータ420に対して、標準試料110aの因子440を用いてNMFによる因子分解を行い、未知試料110bの因子スコア450を算出する。なお、必要に応じて標準化処理を施しても良い。

【0074】

決定木分類処理S4ステップでは、制御装置220において、未知試料110bの因子スコア450を決定木500のルール(学習させたアルゴリズム)に基づき因子440をクラス分けする。決定木500において、因子440の因子適用順にしたがって未知試料110bの因子スコア450をスコア閾値460によってクラス分けし、未知試料が属するクラスを推定していくことにより、未知試料110bに含まれる化合物種を同定する。

【0075】

決定木500に掛けることにより、因子スコア450と化合物種が対応付けられる。標準試料群の全てのスペクトルデータ420から、その共通する因子440(第1因子~第m因子)が導出され、各因子を通して各スペクトルデータ420は因子スコア450に変換される。そして、決定木500を通して化合物種にクラス分けされる。

【0076】

実際に、未知試料110bとして、ステンレス鋼(SUS)と炭化ホウ素(B₄C)の共晶化合物等を対象として測定したところ、テストデータ4種類×8個で正答率は100%の結果を得た。

【0077】

図11は、化合物同定方法で分類した結果を示す図である。因子スコア450をマッピングし、未知試料110bの成分や組成比を推定しても良い。図11(a)に示すように、模擬試料がある化合物に含まれる要素を含んでいるか成分を推定することもできる。また、図11(b)に示すように、CrB₂が多く含まれているか否か組成比を推定することもできる。

【0078】

X軸とY軸の2軸座標の制御と、変位計によりステージと変位計の間に載せた試料の距離を計測することで、高さZ軸の情報を取得し、試料形状を把握する。座標の制御は、電動ステージを用いると容易である。X-Y-Z軸を既知とした試料の位置をピンポイントでLIBS計測することにより、LIBS測定結果が得られる。LIBS測定結果から特徴的な元素の発光スペクトルの範囲を選択し、面積値や高さを算出する。二次元上元素のマッピング情報を表示するには、X座標とY座標情報を2軸に設定し、選択した元素発光スペクトルの強度をカラーで示すことで選択した元素をマッピングすれば良い。

【0079】

本発明によれば、未知試料に含まれる金属元素や化合物を分類・推定することができる。事前に既知物質のスペクトルデータを複数の因子に分解し、決定木により元素や化合物の分類方法を学習させておくことで、未知試料の成分の分類に利用することができる。

【0080】

10

20

30

40

50

また、試料固有の誤差を評価し、予め補正しておくことで、金属元素や化合物を精度良く分類・推定することができる。なお、後付けで補正することが可能で、ホワイトノイズも含め誤差を除去することができる。

【0081】

本発明では、プラズマ状態（時間・空間等）を特徴付ける参照パラメータにより測定対象とする原子発光スペクトルの採用・棄却を決める。また、参照パラメータの設定に機械学習を使っている。従来手法では、測定対象のスペクトルを一定の幅を有するピークとしてサーチするのに対し、本発明では、人間では取り扱うのが困難なビッグデータ（各種分析装置等から得られた全スペクトル成分）をNMFにより機械学習させ、参照パラメータを構築する。これにより、従来方法で行っていたシグナルピークによる処理ではなく、全スペクトル成分を対象に処理することで、評価に用いるデータ量が 10^3 まで増大し、その結果、今まで難しかったスペクトルの強度や形状等も含む評価ができるようになり、高い分析精度を確立する見通しも得られた。

10

【0082】

前処理で用いているベースライン補正は、その種類によって、非対称重みペナルティ最小二乗スムージング法のような関数や、パターン認識のような機械学習を行うことで、スペクトル毎に適用できる。学習データとして公的研究機関等のデータベースや既知試料の計測ができる環境があれば、後付けで既に所有している各種分析装置固有の性能向上（ラマン散乱やLA-ICP-MS等のレーザー光源を用いた分析手法のみならず、分光器を使う全ての測定器）やサンプル依存性の低減が可能である。

20

【0083】

学習データとして公的研究機関等のデータベース（分光や放射線スペクトル情報）や材料組成が既知である標準試料の計測をすれば良く、装置の解析ソフトウェアに実装するなどハードウェア改修することなく、簡便に各種分析装置固有の性能向上やサンプル依存性の低減が可能である。レーザー、X線や中性子等の光源をプローブに用いる分光計測のみならず、分光機能としてエネルギー分散型X線分析（EDX）が付加された電子顕微鏡（透過電子顕微鏡：TEM-EDX、走査電子顕微鏡：SEM-EDX）、電子プローブマイクロアナライザー（EPMA）等で分光データを面情報として画像化するものまで多種多様な応用が期待される。

【0084】

30

以上、本発明の実施例を述べたが、これらに限定されるものではない。例えば、NMFは、オーケストラの音からバイオリンやピオラ等の因子に分ける機械学習に使われる手法で、バイオリンやピオラの特徴音（参照パラメータ）を作成し分類することができる。このように、様々なビッグデータに対し、本発明を応用することができる。

【0085】

また、未知試料内の材料分布や硬度分布を作成しても良いし、多変量解析に応用しても良い。

【符号の説明】

【0086】

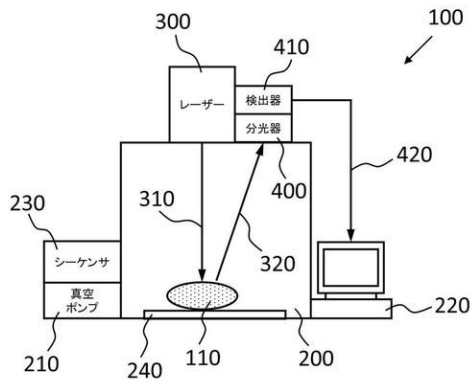
- 100：測定装置
- 110：試料
- 110a：標準試料
- 110b：未知試料
- 200：チャンバー
- 210：真空ポンプ
- 220：制御装置
- 230：シーケンサ
- 240：ステージ
- 300：照射装置
- 310：照射光

40

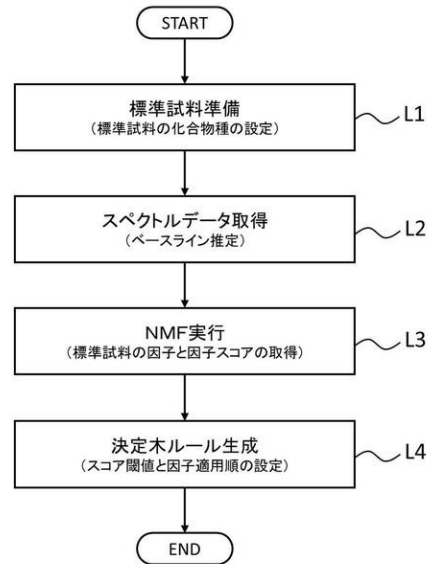
50

- 3 2 0 : 放出光
- 4 0 0 : 分光器
- 4 1 0 : 検出器
- 4 2 0 : スペクトルデータ
- 4 3 0 : ベースライン
- 4 4 0 : 因子
- 4 5 0 : 因子スコア
- 4 6 0 : スコア閾値
- 5 0 0 : 決定木

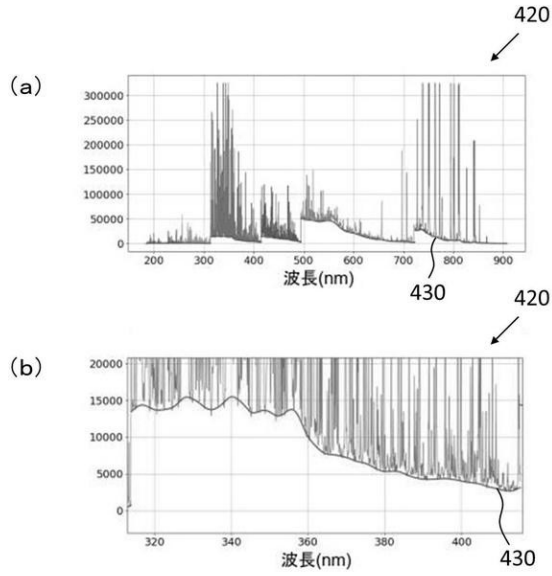
【 図 1 】



【 図 2 】



【 図 3 】

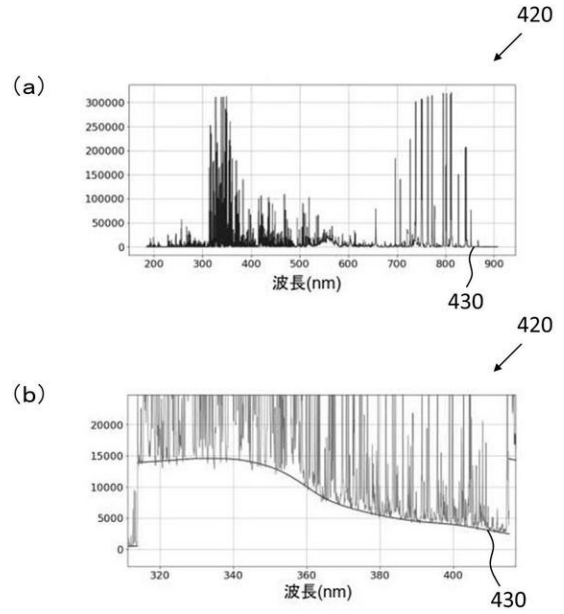


(c)

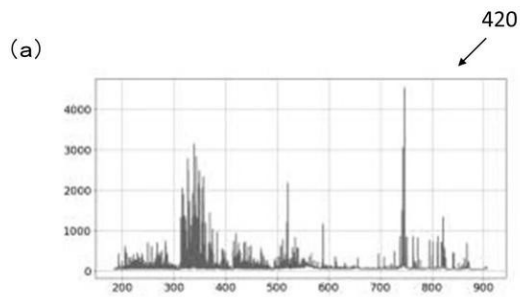
$$z = \arg \min_z \left\{ \sum_i w_i (y_i - z_i)^2 + \mu \sum_i (\Delta^2 z_i)^2 \right\}$$

$$w = p \text{ (if } y_i \geq z_i), \text{ and } 1 - p \text{ (if } y_i < z_i)$$

【 図 4 】



【 図 5 】



(b)

$$Y = [y_1, \dots, y_N] = (y_{k,n})_{K \times N}$$

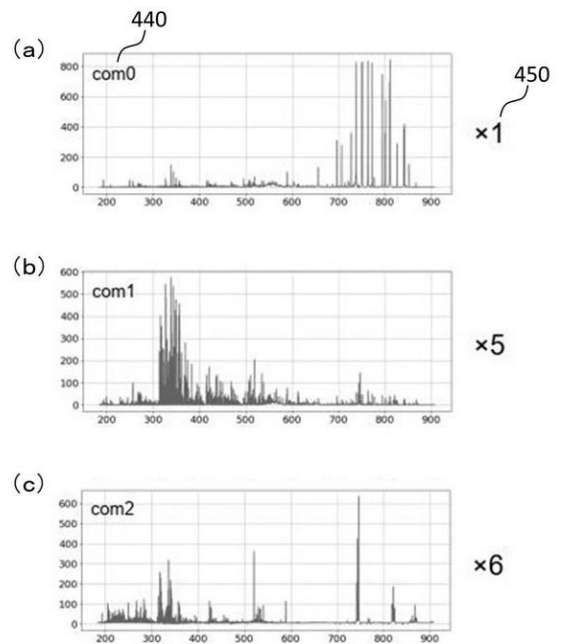
$$H = [h_1, \dots, h_M] = (h_{k,m})_{K \times M}$$

$$U = (u_{m,n})_{M \times N}$$

$$y_n \cong \sum_{m=1}^M h_m u_{m,n} (n = 1, \dots, N)$$

$$Y \cong HU$$

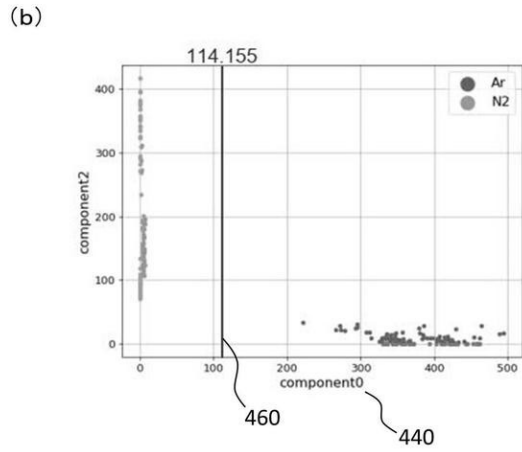
【 図 6 】



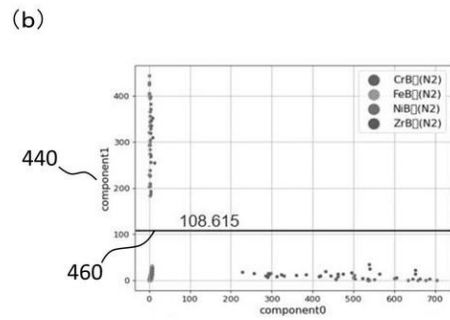
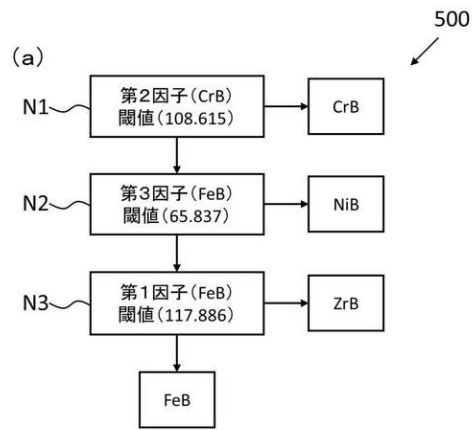
【 図 7 】

(a)

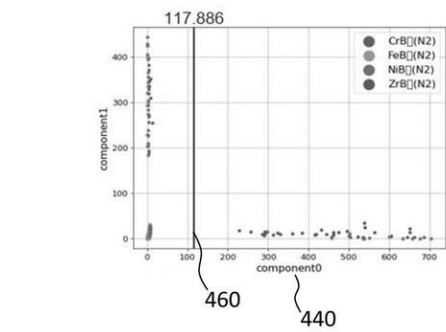
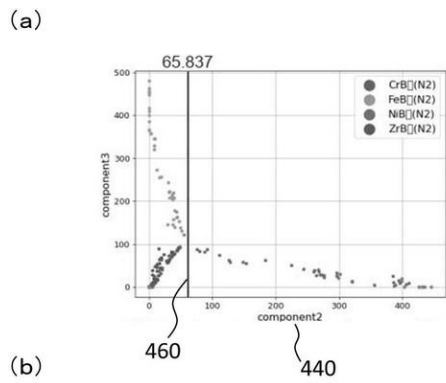
$$J(k, t_k) = \frac{m_{left}}{m} G_{left} + \frac{m_{right}}{m} G_{right}$$



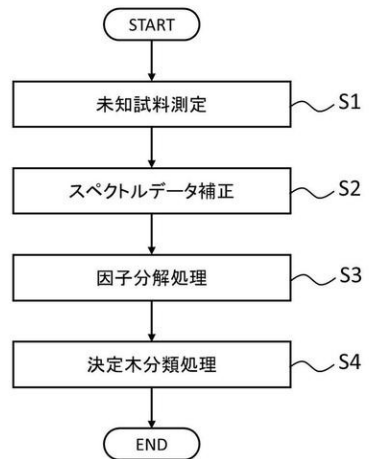
【 図 8 】



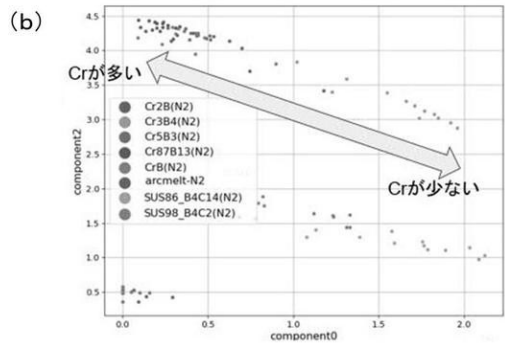
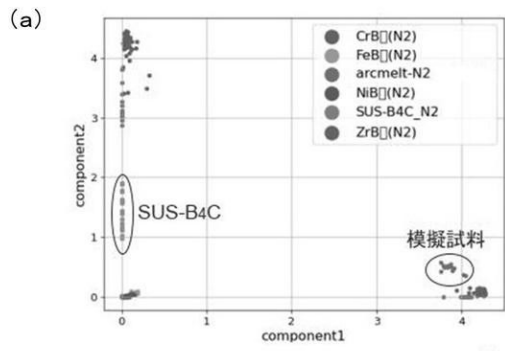
【 図 9 】



【 図 10 】



【 図 1 1 】



フロントページの続き

(72)発明者 横田 裕海

福島県福島市金谷川1番地 国立大学法人福島大学内

(72)発明者 阿部 雄太

茨城県東茨城郡大洗町成田町4002番地 国立研究開発法人日本原子力研究開発機構 大洗研究所内

(72)発明者 川上 智彦

茨城県水戸市堀町1044番地 株式会社化研内

Fターム(参考) 2G043 AA01 HA01 JA01 KA01 KA02 KA03 KA08 KA09