(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6630275号
(P6630275)
(45)発行日　令和2年1月15日(2020.1.15)　(24)登録日　令和1年12月13日(2019.12.13)

(51)Int.Cl.		F I		
G16B 40/10	(2019.01)	G16B	40/10	
G16B 50/00	(2019.01)	G16B	50/00	
G01N 27/62	(2006.01)	G01N	27/62	D
A61B 5/055	(2006.01)	A61B	5/055	380

請求項の数 14　(全 13 頁)

(21)出願番号　特願2016-545874 (P2016-545874)
(86)(22)出願日　平成27年1月9日 (2015.1.9)
(65)公表番号　特表2017-509945 (P2017-509945A)
(43)公表日　平成29年4月6日 (2017.4.6)
(86)国際出願番号　PCT/FR2015/050051
(87)国際公開番号　WO2015/104512
(87)国際公開日　平成27年7月16日 (2015.7.16)
審査請求日　平成29年12月15日 (2017.12.15)
(31)優先権主張番号　1450176
(32)優先日　平成26年1月10日 (2014.1.10)
(33)優先権主張国・地域又は機関
フランス(FR)

(73)特許権者　513237803
イマビオテク
ＩＭＡＢＩＯＴＥＣＨ
フランス国、エフ－５９１２０　ロス、アヴニュー・ウージェーン・アヴィネ　８８５
(74)代理人　110001508
特許業務法人　津国
(72)発明者　パムラール，ファビアン
フランス国、エフ－５９２８０　アルマンティエール、リュ・ベランジェ　３３、アパルトマン７

最終頁に続く

(54)【発明の名称】分子イメージングデータを処理する方法及び対応するデータサーバ

(57)【特許請求の範囲】
【請求項１】
特に分子イメージング方法によって利用される、複数のスペクトルデータセット（Ｊ１～Ｊｎ）を処理し、コンピュータによって実行される方法であり、それぞれのデータセット（Ｊ１～Ｊｎ）が、１セットの分子情報（Ｓ（Ｘｉ，Ｙｊ））を包含する少なくとも２つの次元の分子スペクトルがそれぞれ関連する１セットの空間位置（Ｘｉ，Ｙｊ）によって定められる方法であって、
－データセット（Ｊ１～Ｊｎ）ごとに、それぞれの位置（Ｘｉ，Ｙｊ）と関連した分子スペクトルを、それぞれが前記スペクトルの分子情報の一部を包含する、スペクトルのいくつかの区間（Ｔ１～Ｔｍ）へ分割する工程と、
－スペクトルの１セットの指標付き区間（Ｔ１～Ｔｍ）がそれぞれのデータセット（Ｊ１～Ｊｎ）のそれぞれの位置（Ｘｉ，Ｙｊ）と関連するように、それぞれのデータセット（Ｊ１～Ｊｎ）の位置（Ｘｉ，Ｙｊ）ごとに得られた区間（Ｔ１～Ｔｍ）をデータベース（ＢＤＤ）に書き込む工程と、
－目的分子情報に関する要求の後に、データベース（ＢＤＤ）において、目的分子情報を包含する区間（単数又は複数）（Ｔ１～Ｔｍ）を選択する工程と、
－それぞれの区間（Ｔ１～Ｔｍ）内で前記目的分子情報を選択する工程と、
－少なくとも１つの所定のピッチ（Ｐ’）によってすべての分子スペクトル（Ｓ（Ｘｉ，Ｙｊ））をスペクトルの区間（Ｔ１～Ｔｍ）に分割する工程と、
－いろいろな区間（Ｔ１～Ｔｍ）のポイント指標と、対応する分子情報との間の対応関

係を設定する基準軸（Ａｒｅｆ）を書き込む工程と、
－基準軸（Ａｒｅｆ）及び所定のピッチ（Ｐ’）に応じて、目的分子情報を包含するスペクトル区間（Ｔ１～Ｔｍ）を選択する工程と、
を含むことを特徴とする方法。
【請求項２】
分子イメージング方法の間に、複数のスペクトルデータセット（Ｊ１～Ｊｎ）を記録する方法であり、それぞれのデータセットが、１セットの分子情報（Ｓ（Ｘｉ，Ｙｊ））を包含する少なくとも２つの次元の分子スペクトルがそれぞれ関連する１セットの空間位置（Ｘｉ，Ｙｊ）によって定められ、コンピュータによって実行される方法であって、
－スペクトルデータの取得の間に、スペクトルデータセット（Ｊ１～Ｊｎ）ごとに、それぞれの位置（Ｘｉ，Ｙｊ）と関連した分子スペクトルを、それぞれが前記スペクトルの分子情報の一部を包含する、スペクトルのいくつかの区間（Ｔ１～Ｔｍ）へ分割する工程と、
－スペクトルの１セットの指標付き区間（Ｔ１～Ｔｍ）がそれぞれのデータセット（Ｊ１～Ｊｎ）のそれぞれの位置（Ｘｉ，Ｙｊ）と関連するように、それぞれのデータセット（Ｊ１～Ｊｎ）の位置（Ｘｉ，Ｙｊ）ごとに得られた区間（Ｔ１～Ｔｍ）をデータベース（ＢＤＤ）に書き込む工程と、任意に、
－目的分子情報に関する要求の後に、データベース（ＢＤＤ）において、目的分子情報を包含する区間（単数又は複数）を選択する工程と、
－それぞれの区間（Ｔ１～Ｔｍ）内で前記目的分子情報を選択する工程と、
－少なくとも１つの所定のピッチ（Ｐ’）によってすべての分子スペクトル（Ｓ（Ｘｉ，Ｙｊ））をスペクトルの区間（Ｔ１～Ｔｍ）に分割する工程と、
－いろいろな区間（Ｔ１～Ｔｍ）のポイント指標と、対応する分子情報との間の対応関係を設定する基準軸（Ａｒｅｆ）を書き込む工程と、
－基準軸（Ａｒｅｆ）及び所定のピッチ（Ｐ’）に応じて、目的分子情報を包含するスペクトル区間（Ｔ１～Ｔｍ）を選択する工程と、
を含むことを特徴とする方法。
【請求項３】
それぞれがデータセット（Ｊ１～Ｊｎ）に対応する、１セットのデータマップ（Ｃ１～Ｃｎ）の形で目的分子情報を視認する工程を含むことを特徴とする、請求項１又は２に記載の方法。
【請求項４】
データベース（ＢＤＤ）への区間（Ｔ１～Ｔｍ）の書き込みの前又は後に、前記方法が、スペクトル配列、及び／又はバックグラウンドノイズの減算、及び／又はイントラデータセット標準化（Ｊ１～Ｊｎ）からなる少なくとも１つの前処理工程を含むことを特徴とする、請求項１～３のうちのいずれか１項に記載の方法。
【請求項５】
最も微細な空間サイズを有するデータセット（Ｊ１～Ｊｎ）上に配列されるようにいろいろな空間サイズを有するデータセット（Ｊ１～Ｊｎ）の空間位置（Ｘｉ，Ｙｊ）をサイズ変更する工程を含むことを特徴とする、請求項１～４のうちのいずれか１項に記載の方法。
【請求項６】
最大強度（Ｐｋ１～Ｐｋｌ）、平均強度、ピーク下の面積、すべての分子情報のそれぞれの目的ピークのノイズ対信号比などの、データセットを抽出する工程を更に含むことを特徴とする、請求項１～５のうちのいずれか１項に記載の方法。
【請求項７】
抽出が、ノイズ対信号比及び／又はスペクトル分解能の品質基準によって定められるピーク選択基準に応じて、実行されることを特徴とする、請求項６に記載の方法。
【請求項８】
次いでいろいろなデータセット（Ｊ１～Ｊｎ）を互いに比較するために複数のデータセ

ット（Ｊ１～Ｊｎ）の分子スペクトルを互いに標準化する工程を含むことを特徴とする、請求項１～７のうちのいずれか１項に記載の方法。
【請求項９】
標準化工程を実行するために、
－データセット（Ｊ１～Ｊｎ）のすべてに共通な１つの分子又はいくつかの内因性若しくは外因性基準分子（Ｍｒｅｆ）を選択する工程と、
－データセット（Ｊ１～Ｊｎ）ごとに、これらの基準分子（単数又は複数）（Ｍｒｅｆ）の関数としての標準化要因（Ｆｎｏｒ）を計算する工程と、
－１セットの標準化されたデータセット（Ｊ１～Ｊｎ）を得るためにこの標準化要因によって目的分子情報を補正する工程と、を含むことを特徴とする、請求項８に記載の方法。
【請求項１０】
標準化要因（Ｆｎｏｒ）が、データセット（Ｊ１～Ｊｎ）に対応する１セットのサンプルに、又はデータセット（Ｊ１～Ｊｎ）に対応するサンプルの特定の基準ゾーン（Ｚ）に存在する少なくとも１つの基準分子（Ｍｒｅｆ）から定められることを特徴とする、請求項９に記載の方法。
【請求項１１】
標準化されたデータセット（Ｊ１～Ｊｎ）からそれぞれ生じる複数のデータマップ（Ｃ１’～Ｃｎ）の形で、マップのすべてに共通のカラースケールによって目的分子情報を視認、比較、及び分析する工程を含むことを特徴とする、請求項８～１０のうちのいずれか１項に記載の方法。
【請求項１２】
共通のカラースケールが、目的分子のためのすべてのデータセット（Ｊ１～Ｊｎ）の最小及び最大の強度に基づくことを特徴とする、請求項１１に記載の方法。
【請求項１３】
前記データセット（Ｊ１～Ｊｎ）が、質量分析方法によって得られることを特徴とする、請求項１～１２のうちのいずれか１項に記載の方法。
【請求項１４】
前記データセット（Ｊ１～Ｊｎ）が、Positron Emission Tomography（ＰＥＴ）又はMagnetic Resonance Imaging（ＭＲＩ）タイプのイメージング方法によって得られることを特徴とする、請求項１～１２のうちのいずれか１項に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
分野
本発明は、大容量データベース構造の使用、及び対応するデータのサーバをベースとした、分子イメージングデータを処理する方法に関する。分子イメージングは、可能な限り最小な侵襲手段、又は観察される生物学的若しくは生物物理学的システムを可能な限り崩壊させない手段によって、生体外又は生体内で器官及び／又は生物の分子機能を主に観察するすべてのイメージング技術を一般に包含する。
【０００２】
背景
分子イメージングは、数多くの技法の進化において、特に質量分析イメージング（ＭＳＩ）、磁気共鳴イメージング（ＭＲＩ）、又はラマン分光分析において、かなりの成長を続けている。これらの技法は、生物学的システムにおける役割を研究する、農薬、薬物、タンパク質又は脂質などの、標的内因性又は外因性化合物の生体内分布の研究を実行するために使用される。ソフトウェアの助力を得たこれらの技法の適用は、監視をせずに潜在的なバイオマーカを探索することを可能にする。
【０００３】
しかしながら、分子イメージングにおいて分析及び解釈されるデータの増加は、画像と

それぞれ関連したいくつかのデータセットに関する総合的な分析を制限する。データは、総合的な表現のために利用すること又は統計学的に分析することが困難になってきている。

【０００４】

実際に、データセットの多量の情報は、多くの場合、（スーパーコンピュータとは対照的に）従来のコンピュータのランダムアクセスメモリのサイズよりも大きい。例えば質量分析イメージングでは、５０，０００の位置の画像によって、すべての生情報がメモリに記憶できなくなり、したがってメモリへの影響を低減させこれらの画像の利用を可能にするために計算の実行が必要となる。計算によって、データが低減し、その結果、すべての利用可能な情報を考慮しないために、偏りが生じる。また、現在の分析ツールは、互いに関連したいくつかのデータセットを標準化する手段を有していない。

【０００５】

主にデータセットの取得ロボットの製造者に結びつけられる多くの記憶フォーマットが存在する。しかしながら、これらのフォーマットはいずれも、微細かつ迅速に問い合わせられるように適合及び最適化されない。分子イメージングの一部の分野は、ＭＲＩ（磁気共鳴イメージング）及び質量分析イメージングにおいて見いだされるＡｎａｌｙｚｅ７．５などの記憶フォーマットを開発している。質量分析イメージングはまたフォーマットｉＭＺＭＬを包むが、しかしながら、このフォーマットは、イメージングデータを扱う記憶システムを構成しない。

【０００６】

データの階層的な組織化に基づくタイプＨＤＦ５（“Hierarchical Data Format 5”）のフォーマットが、近年質量分析イメージングにおける大容量データの記憶に使用されている。これらのフォーマットは、インターネットブラウザを介した遠隔問い合わせインタフェースを構築する。それらは、非階層化されたファイルの場合よりも速い統計学的な計算を可能にする。しかしながら、これらのフォーマットは、いくつかのデータセットの研究において同時に比較及び検索可能な全体的なデータの分析をもたらさない。

【０００７】

したがって、分析の偏りがもたらされることを回避するために情報損失なく互いに関連したいくつかのデータセットを分析、比較、及び問い合わせる方法の必要性が存在する。

【０００８】

概要

本発明の目的は、特に分子イメージング方法によって利用される、複数のスペクトルデータセットを処理する方法であって、それぞれのデータセットが、１セットの分子情報を包含する少なくとも２つの次元の分子スペクトルがそれぞれ関連する１セットの空間位置によって定められる方法、を提案することであり、この必要性に効果的に応えることであるこの方法は、特に、

－データセットごとに、それぞれの位置と関連した分子スペクトルを、それぞれが低減された１セットの分子情報を包含する、分子情報のいくつかの区間へ、又はスペクトルの区間へ分割する工程と、

－スペクトルの１セットの指標付き区間がそれぞれのデータセットのそれぞれの位置と関連するように、それぞれのデータセットの位置ごとに得られた区間をデータベースに書き込む工程と、

－目的分子情報に関する要求の後に、データベースにおいて、目的分子情報を包含する区間（単数又は複数）を選択する工程と、

－それぞれの区間内で前記目的分子情報を選択する工程と、を含むことを特徴とする。

【０００９】

本発明はまた、分子イメージング方法の間に、複数のスペクトルデータセットを記録する方法であって、それぞれのデータセットが、１セットの分子情報を包含する少なくとも２つの次元を有する分子スペクトルがそれぞれ関連する１セットの空間位置によって定められる方法に関する。この方法は、

－スペクトルデータの取得の間に、スペクトルデータセットごとに、それぞれの位置と関連した分子スペクトルを、それぞれが前記スペクトルの分子情報の一部を包含する、スペクトルのいくつかの区間へ分割する工程と、
－スペクトルの１セットの指標付き区間がそれぞれのデータセットのそれぞれの位置と関連するように、それぞれのデータセットの位置ごとに得られた区間をデータベースに書き込む工程と、
任意に、
－目的分子情報に関する要求の後に、データベースにおいて、目的分子情報を包含する区間（単数又は複数）を選択する工程と、
－それぞれの区間内で前記目的分子情報を選択する工程と、を含むことを特徴とする。

【００１０】
より一般には、本発明は、前記データセットを構成する空間位置のセットの空間位置のそれぞれと関連した分子スペクトルを区間に分割することによって有利には自動的に取得される、データセットの処理を提案する。それぞれのスペクトルは、すべてが関連したスペクトルの分子情報のすべてを含む、スペクトルの区間の連続に対応する。本発明によれば、区間への分割工程は、スペクトルデータの取得の間に又はスペクトルデータの分析の間に、有利には自動的に行われ得る。

【００１１】
第１の実施形態では、関連スペクトルからの分子情報が損失されないか又は冗長でないように、連続する区間は重複せずに、単純に隣接する。別の実施形態では、一部の分子情報が同一のスペクトルの２つの連続するスペクトルの区間に存在することができるように、関連スペクトルの区間の一部が部分的に重複し得る。

【００１２】
本発明は、区間の選択によって、データベースの有用な情報を回復することができる程度で、分析されるデータの読み込み時間を実質的に低減させる。データの取得の間、データが記録されると同時に、スペクトルデータを分割することも可能であり、その結果、スペクトルデータの区間が、分子イメージング方法によるサンプルの分析の間に直接統合される。

【００１３】
スペクトルの低減なしで、信号の区間への分割により、本発明はまた、選別法（いくつかの薬物候補の分布比較）、定量化の方法、レセプタ占有率又は組織選別の研究、代謝若しくは３Ｄ再構成の研究の方法などの、完全に生理学的な事象を観察するための同一分析における薬物動態学的、薬力学的、及び内因性又は外因性分子の定量化の複合研究についての情報を低減させずに作用する。本発明はまた、統計学的処理に必要なメモリの使用及びいくつかのデータセット上の分子データイメージングの表現を最適化する。

【００１４】
一実施形態によれば、本発明のデータ処理又は記録の方法は、それぞれがデータセットに対応する、１セットのデータマップの形で目的分子情報を視認する工程を含む。

【００１５】
有利には、分子スペクトルは、同一か又はスペクトルごとで異なる可能性がある決定されたピッチによって、切り分けられる。

【００１６】
一実施形態によれば、本発明のデータ処理又は記録の方法は、
－少なくとも１つの所定のピッチによってすべての分子情報をスペクトルの区間に分割する工程と、
－スペクトルのいろいろな区間のポイント指標と、対応する分子情報との間の対応関係を設定する基準軸を書き込む工程と、
－基準軸及び所定のピッチに応じて、目的分子情報を包含するスペクトル区間を選択する工程と、を含む。

【００１７】

一実施形態によれば、データベースの区間の書き込みの前又は後に、前記方法は、スペクトル配列、及び／又はバックグラウンドノイズの減算、及び／又はイントラデータセット標準化からなる少なくとも１つの前処理工程を含む。

【００１８】

一実施形態によれば、本発明のデータ処理又は記録の方法は、最も微細な空間サイズを有するデータセット上に配列されるようにいろいろな空間サイズを有するデータセットの空間位置をサイズ変更する工程を含む。

【００１９】

一実施形態によれば、本発明のデータ処理又は記録の方法は、最大強度、平均強度、ピーク下の面積、すべての分子情報のそれぞれの目的ピークのノイズ対信号比などの、データセットを抽出する工程を更に含む。

【００２０】

一実施形態によれば、抽出は、ノイズ対信号比及び／又はスペクトル分解能の品質基準によって定められるピーク選択基準に応じて、達成される。

【００２１】

一実施形態によれば、本発明のデータ処理又は記録の方法は、次いでいろいろなデータセットを互いに比較するために複数のデータセットの分子情報のセットを互いに標準化する工程を含む。

【００２２】

一実施形態によれば、標準化工程に着手するために、前記方法は、

－データセットのすべてに共通な１つの基準分子又はいくつかの内因性若しくは外因性基準分子を選択する工程と、

－データセットごとに、これらの基準分子（単数又は複数）の関数としての標準化要因を計算する工程と、

－１セットの標準化されたデータセットを得るためにこの標準化要因によって目的分子情報を補正する工程と、を含むことができる。

【００２３】

一実施形態によれば、標準化要因は、データセットに対応する１セットのサンプルに、又はデータセットに対応するサンプルの特定の基準ゾーンに存在する少なくとも１つの基準分子から定められる。

【００２４】

一実施形態によれば、本発明のデータ処理又は記録の方法は、標準化されたデータセットからそれぞれ生じる複数のデータマップの形で、マップのすべてに共通のカラースケールによって目的分子情報を視認、比較、及び分析する工程を含む。

【００２５】

一実施形態によれば、共通のカラースケールは、目的分子のためのすべてのデータセットの最小及び最大の強度に基づく。

【００２６】

一実施形態によれば、前記データセットは、質量分析方法によって得られる。

【００２７】

一実施形態によれば、前記データセットは、Positron Emission Tomography（ＰＥＴ）又はMagnetic Resonance Imaging（ＭＲＩ）タイプのイメージング方法によって得られる。

【００２８】

有利には、本発明による記録及び／又は処理方法の工程のすべて又は一部は、自動化される。

【００２９】

本発明の別の目的は、本発明による複数のデータセットを処理する方法の少なくとも一部の工程を実行するメモリ記憶ソフトウェア命令を含むデータサーバである。

【００３０】

本発明の別の目的は、コンピュータによって実行可能でかつ本発明によるスペクトルデータを処理又は記録する方法の少なくとも１つの工程をコンピュータシステムが実行することができるように適合された命令を含む、コンピュータによって読み込み可能なデータ媒体である。

【００３１】

したがって、本発明は、プログラムがコンピュータ上で動作するときに上で説明された工程のすべて又は一部を実行するためのプログラムコード命令を含むコンピュータプログラムを提案する。

【００３２】

有利には、コンピュータプログラムは、少なくとも分子スペクトルをスペクトルの区間に分割する工程の実行のためのプログラムコード命令を含む。

【００３３】

本発明は、例示として与えられるが本発明を限定しない添付図面の以下の説明及び審査からより明確に理解されるであろう。

【図面の簡単な説明】

【００３４】

【図１】本発明による方法によって利用されるデータセットを得るための質量分析方法のいろいろな工程を概略的に示す。

【図２】本発明によるデータを処理する方法の実施のための統合データベース管理システムの概観図である。

【図３】データセットのそれぞれ位置と関連したスペクトル区間を得る工程の概略図である。

【図４】分子情報と対応するスペクトル区間との間の対応関係を設定する工程を示す。

【図５】いくつかの画像を正確に比較するためのインター画像標準化工程を概略的に示す。

【００３５】

同一、類似、又は同様の要素は、一方の図から他方の図まで同一の参照番号を維持する。

【００３６】

発明の詳細な説明

本発明による方法の実行の以下の例は、MALDIタイプの質量分析法のために説明される。当然、例えば、SIMS、DESI、LAESI、DIOS、ICP、Microscope MALDI、SNOM、SMALDI、LA-ICP、ESI（組織上での液体抽出）、MILDI、JEDI、ELDI、などのソースといった、MALDI以外のイメージング装置が、実質的に同一に使用され得る。本発明による方法はまた、例えばＰＥＴ（“Positron Emission Tomography”）又はＭＲＩ（Magnetic Resonance Imaging）などの、分子イメージング法によって利用可能なデータを生成する任意の他の方法に適用可能である。

【００３７】

図１は、MALDI質量分析法によって実行されるいろいろな工程を概略的に示す。第１の工程Ｅ１は、次いでスライド２上に配置される一片の組織１を、一般に低温切開によって、切り抜くことからなる。工程Ｅ２において、イオン化マトリックス３の微細均一層が、次いで、一片の組織１上に配置される。取得工程Ｅ３は、マトリックス３の分子をイオン化するために、所定のサイズを有した、例えば、取得ゾーン当たり約１００マイクロメートルの、自動レーザパルス４を生成することからなる。このようにしてイオン化された分子は、質量分析計によってそれ自体公知のように分析される。方法の実行に関するより詳細については文献FR2973112を参照することができる。

【００３８】

１セットの分子情報のそれぞれと関連した１セットの空間位置（Ｘｉ，Ｙｊ）によって定められるデータセットＪｋがもたらされる。空間位置は、軸Ｘ及びＹの基準座標系の関数として定められる。すべての分子情報は、この場合は特に分子量の関数としての強度を

示す２つの次元を有したスペクトルＳによって形成される。例えば、得られるデータセットＪｋは、２０，０００の空間位置Ｘｉ，Ｙｊ上にスペクトル当たり、５０，０００の強度（又は５０，０００ポイント）を有する。本発明による方法は、複数のデータセットＪ１～Ｊｎの処理に基づく。

【００３９】

図３に示されるように、画像の形式で示される可能性のあるそれぞれのデータセットＪｋのために、それぞれの位置（Ｘｉ，Ｙｊ）と関連した分子スペクトルＳ（Ｘｉ，Ｙｊ）は、スペクトルのいくつかの区間Ｔ１～Ｔｍに分割される。区間Ｔ１～Ｔｍは、所定のピッチＰ’によって切り分けられる。例えば、空間位置につき５０，０００ポイントのそれぞれのスペクトルのために、それぞれのスペクトルは、１０，０００ポイントのピッチＰ’によって切り分けられ得る。すなわち、スペクトルＳ（Ｘｉ，Ｙｊ）は、それぞれ１０，０００ポイントの５つの区間Ｔ１～Ｔ５に分割される。変数として、使用されるピッチＰ’を変化させることができる。

【００４０】

異なる区間Ｔ１～Ｔｍのポイント指標と対応する分子量との間の対応関係を設定する基準軸Ａｒｅｆがまた、図４に示されるように、定められる。

【００４１】

データベースＢＤＤへの書き込みの前に、方法は、好ましくは、モジュール１００を介してスペクトルＳ（Ｘｉ，Ｙｊ）の少なくとも１つの前処理工程を実行する。この前処理工程は、スペクトル配列からなることができる。この目的のために、基準ピークが考慮され、スペクトルは対応するピークの値をこれらの基準ピーク上に位置づけるようにオフセットされる。スペクトルの最小値とスペクトルの最小値が一致しなければならない基準値との間の距離に応じてスペクトルを再調整することによって、バックグラウンドノイズの減算を実行することも可能となる。

【００４２】

更に、いろいろなタイプの信号標準化（内部標準化）が、（内部標準によって、強度の総量を使用することによって）実行され得る。換言すれば、イントラデータセット標準化、すなわち、当業者にそれ自体公知の工程に対応する他のデータセットとは独立したデータセットごとの標準化が、ここで実行される。このイントラデータセット標準化は、データセットＪ１～Ｊｎを互いに正確かつ確実に比較する以下に詳述するデータセットＪ１～Ｊｎ間の次の標準化工程（インターデータセット標準化）に対する補完として、実行される。

【００４３】

また、方法は、スペクトルＳ（Ｘｉ，Ｙｊ）のそれぞれの目的ピークの、最大強度Ｐｋ１～Ｐｋｌ、平均強度、ピーク下の面積、ノイズ対信号比などの、データセットを抽出する工程を含む。「PeakPicking」と呼ばれる方法によってスペクトルＳ（Ｘｉ，Ｙｊ）のそれぞれのピークＰｋ１～Ｐｋｌの最大強度を抽出することが可能である。抽出は、ノイズ対信号比及び／又はスペクトル分解能の品質基準によって定められるピーク選択基準に応じて行われる。

【００４４】

それぞれのデータセットＪ１～Ｊｎの位置（Ｘｉ，Ｙｊ）ごとに得られる区間Ｔ１～Ｔｍ、並びに基準軸Ａｒｅｆ及び予め抽出されたピークＰｋ１～ＰＫｌは、図２に示されるように、データベースＢＤＤに書き込まれる。このようにして、１セットの指標付き区間Ｔ１～ＴＭ並びに１セットのピークＰｋ１～ＰＫｌ及び上述した補完的なデータ（平均強度、それぞれのピークの面積）は、それぞれのデータセットＪ１～Ｊｎのそれぞれの位置（Ｘｉ，Ｙｊ）と関連付けられる。

【００４５】

データベースＢＤＤは、書き込まれたデータの視認、管理、解釈、及び統計学的分析を可能にするＮｏＳＱＬ（“Not only Structured Query Language”）又はＳＱＬタイプの大容量データベースである。例えばデータベース管理システムは、以下の包括的ではない

システム、Hypertable、Haddoop、Cassandra、MongoDB、PolyBase、Azure上のHadoop、Hive、Pigから選択され得る。データベースシステムは、データのコヒーレンス、信頼性、及び適切さを確保する。複製及び保存の技術により、情報の信頼性を確保することができる。

【００４６】

データベースＢＤＤの適切な情報の選択工程が以下に説明される。目的分子情報と関連した要求の後、要求された分子情報を包含するデータセットＪ１～Ｊｎのセットの位置の区間（単数又は複数）が選択される。この目的のために、それぞれのデータセットＪ１～Ｊｎの位置Ｘｉ、Ｙｊごとに、要求された分子情報を包含する区間Ｔｐが、基準軸Ａｒｅｆ及び要求された分子情報の関数として選択される。

【００４７】

したがって、図４は、この工程のより簡単な理解のために、５ポイントのピッチＰ’による区間へのスペクトルの分割を示す。質量分析法によるスペクトルの場合によくあることだが、分子量のスケールは線形ではない。ユーザが例えば３ｍ／ｚの値の分子量を有する分子と関連した情報を必要とする場合には、基準軸Ａｒｅｆは分子量のこの値に対応するポイントの指標が９であると規定し、その結果、ピッチＰ’によって、目的分子量の強度を包含する区間Ｔｐが第２の区間Ｔ２であることが導出され得る。

【００４８】

あるいは、分子量の最小及び最大端並びにそれぞれの対応する区間Ｔ１～Ｔｍのポイントの指標化を記憶する表が使用され得る。この場合、要求された分子情報のために、区間Ｔ１～Ｔｍの識別子が、これらの最小及び最大端に応じて選択される。

【００４９】

すべてのデータセットＪ１～Ｊｎのポイントのそれぞれのために選択されたそれぞれの区間Ｔ１～Ｔｍ内で、要求された分子量の対応する強度が、次いで選択される。次いで、データセットＪ１～Ｊｎにそれぞれ対応する１セットのデータマップＣ１～ＣＮの形で、いろいろなマップＣ１～ＣＮ上に目的分子の存在の強度の分布を視認することができるようになる。色は、カラースケール上の分子情報の強度に左右される。代替的に、又は補完として、ピークＰｋ１～ＰＫｌの抽出が、目的分子の強度の分布の初期分析を提供する画像を迅速に生成する。

【００５０】

方法は、好ましくは、次いでいくつかのデータセットを互いに比較するためにデータセットＪ１～Ｊｎの分子情報を標準化する工程を含む（モジュール１０１参照）。この目的のために、すべてのデータセットＪ１～Ｊｎに共通の基準分子Ｍｒｅｆが選択される。代替的に、データセットＪ１～Ｊｎのすべてに共通のいくつかの内因性又は外因性分子が選択される。完全なサンプル中の、又はより詳細に対応するサンプルの基準ゾーンの、基準分子Ｍｒｅｆ（単数又は複数）の存在を考慮することが可能である。

【００５１】

より正確には、データセットＪ１～Ｊｎごとに、標準化要因Ｆｎｏｒが、これらの基準分子Ｍｒｅｆ（単数又は複数）の関数として計算される。目的分子情報は、次いで、標準化されたセットＪ１～Ｊｎのデータセットを生成するために、この標準化要因Ｆｎｏｒによって補正される。

【００５２】

例えば、図５に示されるマップＣ１～Ｃ３に対応する３つのデータセットのために、基準分子Ｍｒｅｆ＝５００ｍ／ｚが、標準化分子（標準化標準）として選択される。この分子Ｍｒｅｆは、マップＣ１～Ｃ３のゾーン Ｚに５、１０、及び１５の値をそれぞれ有する。マップＣ１～Ｃ３ごとに、標準化要因Ｆｎｏｒが、Ｍｒｅｆのこれらの値のそれぞれに基づいて計算される。Ｆｎｏｒ＝Ｃ１～Ｃ３のＭｒｅｆの値の平均／該マップのＭｒｅｆの値。図５に示される例では、マップＣ１、Ｃ２、及びＣ３のために２、１、及び１０／１５の標準化要因Ｆｎｏｒが、それぞれ得られる。この係数は、次いで、すべてのデータセットＪ１～Ｊｎのいろいろな位置と関連したスペクトルの強度値、特に位置Ｘｉ；Ｙｊ

の目的分子Ｍｉ＝３００ｍ／ｚの強度の値、のすべてに適用される。標準化されたマップＣ１’～Ｃ３’が得られ、互いに直接比較することが可能となる。

【００５３】

標準化は、サンプルの調製及びその分析のための試験条件の変化と結びつけられた強度を補正する。実際に、前の例では。マップＣ３に対応するサンプルの目的分子Ｍｉ＝３００ｍ／ｚの強力な存在の誤った結論に達する可能性があることが明白であるが、標準化要因Ｆｎｏｒの適用後に、研究は、位置Ｘｉ；Ｙｊで２００の強度を有する３つのサンプルの目的分子Ｍｉの均一な存在を明らかにする。

【００５４】

この標準化の後、標準化されたデータマップＣ１’～Ｃ３’の形で、マップＣ１’～Ｃ３’のすべてに共通のカラースケールで、分子情報を視認、比較、及び分析することが可能である（モジュール１０２参照）。得られるスケールは、目的分子Ｍｉのためのすべてのデータセットにわたって決定される最低及び最高強度に基づく。

【００５５】

当然、メモリにすべてを読み込むことなく、横断していくつかのデータセットＪ１～Ｊｎを統計学的に組み合わせることによって分析するために、３つ以上のデータセット及び対応するマップを生成することが可能である。イメージングデータは、次いで、潜在的なマーカを明らかにするために比較され得る。データを新たに解釈できるように共通の挙動を有する位置を組み合わせることも可能である。

【００５６】

換言すれば、本発明は、単一のツールで、定量化可能な品質の標準化された適切な情報を得るために、データを比較及び問い合わせる。いくつかのデータセットは、データを標準化するために画像のそれぞれに書き込まれた品質制御を読み込むことができる。

【００５７】

更に、方法は、必要に応じて、最も微細な空間サイズを有するデータセットに配列されるように、いろいろな空間サイズを有する同じ生物学的サンプルのデータセットＪ１及びＪ２の位置をサイズ変更することができる。ＰＥＴ（Positron Emission Tomography）又はＭＲＩ（Magnetic Resonance Imaging）タイプのシステムからのマップの処理と、質量分析計（イメージングMALDI）からのマップの処理を組み合わせることが可能である。

【００５８】

単一の画像に共分布のいくつかのデータセットＪ１～Ｊｎを表現することも可能である。「現況マップのフィールドｚ及びａのための（同じ空間サイズの）すべての位置ｘ及びｙのすべての値ｖ、ｗを選択する」タイプの要求を実現することが可能である。フィールドｚ及びａは、２つの異なるイメージング技法の値を表現する。

【００５９】

データベースＢＤＤへの書き込み要求によって加えられ得る一時的な統計学的処理を実行することも可能である。このようにして、本発明による方法は、多くの場合、分析方法の任意の時点で画像のそれぞれの位置に関する情報を加えることが可能な程度で、大きな柔軟性を提供する。

【００６０】

方法は、クライアント／サーバアーキテクチャに基づく。データベースＢＤＤは、単一のインタフェースによって問い合わせられ得る。データは、シングルユーザアーキテクチャ、又はローカルに若しくは（内部若しくは外部）ネットワークを介してアクセス可能なサーバのクラスタに、記憶及び問い合わせられる。クライアントインタフェースソフトウェアは、ワークステーションで又はウェブインタフェースバージョンでインストールされ得る。本発明の別の目的は、複数のデータセットＪ１～Ｊｎを処理する方法の少なくとも一部の工程を実行するメモリ記憶ソフトウェア命令を含むデータサーバ１０である。

【００６１】

また、データの処理数は容易に増大し、分析されるデータの量は分析に対する制限ではない。データの書き込みは、データベースＢＤＤのサイズに関係なく有効である。

【００６２】

当然、上記の説明は一例として与えられ、本発明の分野を限定せず、実施の詳細をすべての他の等価物と置き換えることはその範囲を越えるものではない。

【図１】

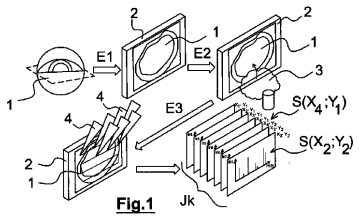
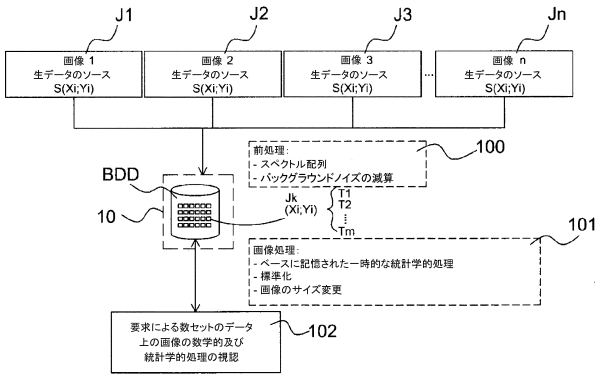
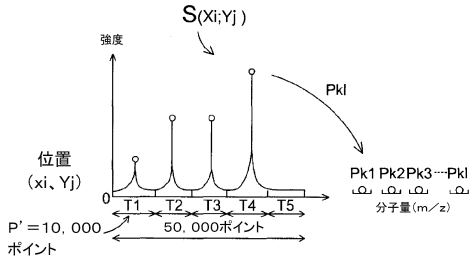


Fig.1

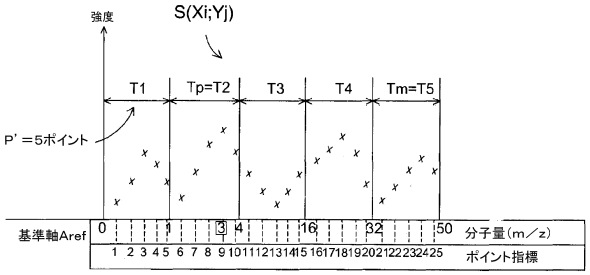
【図２】



【図３】



【図４】



【図５】

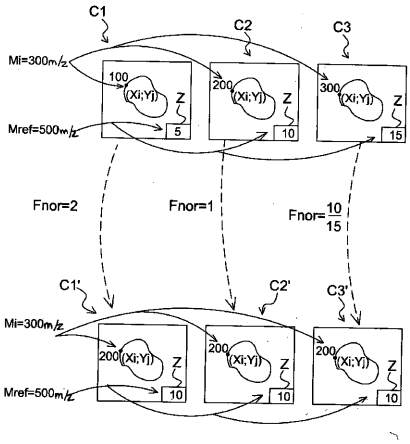

C1
C2
C3
Mi=300m/z
100
(Xi;Yj)
Z
5
200
(Xi;Yj)
Z
10
300
(Xi;Yj)
Z
15
Mref=500m/z
Fnor=2
Fnor=1
Fnor=10/15
C1'
C2'
C3'
Mi=300m/z
200
(Xi;Yj)
Z
10
200
(Xi;Yj)
Z
10
200
(Xi;Yj)
Z
10
Mref=500m/z

Fig.5

フロントページの続き

(72)発明者　スタウバー, ジョナサン・エム
フランス国、エフ－５９３２０　オーブールダン、リュ・デュ・マレシャル・ルクレール　１６

審査官　松野　広一

(56)参考文献　特表２００９－５４０３１９（ＪＰ, Ａ）
特開２０１２－２４７１９８（ＪＰ, Ａ）
特開平０７－１２８４５０（ＪＰ, Ａ）
特表２００４－５２６９５８（ＪＰ, Ａ）
特開２０１１－１９１２２２（ＪＰ, Ａ）
特表２００９－５００６１７（ＪＰ, Ａ）
国際公開第２０１３／０８０１７０（WO, Ａ１）
国際公開第０９／１５６８９７（WO, Ａ１）

(58)調査した分野(Int.Cl., ＤＢ名)
Ｇ１６Ｂ　５／００－９９／００
Ａ６１Ｂ　５／０５５
Ｇ０１Ｎ　２７／６２
Ｇ１６Ｈ　３０／００－３０／４０