

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4355281号  
(P4355281)

(45) 発行日 平成21年10月28日(2009.10.28)

(24) 登録日 平成21年8月7日(2009.8.7)

(51) Int.Cl.		F I		
GO 1 N 27/62	(2006.01)	GO 1 N 27/62		D
GO 1 N 30/86	(2006.01)	GO 1 N 30/86		M
GO 1 N 30/72	(2006.01)	GO 1 N 30/72		C

請求項の数 9 (全 21 頁)

(21) 出願番号	特願2004-359754 (P2004-359754)	(73) 特許権者	500309920
(22) 出願日	平成16年12月13日(2004.12.13)		株式会社インテックシステム研究所
(65) 公開番号	特開2006-170647 (P2006-170647A)		東京都江東区新砂1-3-3
(43) 公開日	平成18年6月29日(2006.6.29)	(74) 代理人	100091096
審査請求日	平成18年3月28日(2006.3.28)		弁理士 平木 祐輔
		(74) 代理人	100105463
			弁理士 関谷 三男
		(74) 代理人	100102576
			弁理士 渡辺 敏章
		(74) 代理人	100108394
			弁理士 今村 健一
		(72) 発明者	済木 育夫
			富山県婦負郡婦中町蛸川151-26

最終頁に続く

(54) 【発明の名称】 ピーク抽出方法およびピーク抽出装置

(57) 【特許請求の範囲】

【請求項1】

変数軸と強度軸とを有する複数のスペクトル特性における波形データからピークを抽出するピーク抽出方法であって、

それぞれの前記スペクトル特性の単一の波形データ中の点に対して、該点が前記変数軸上においてピークの近傍であるか否かを推定し、ピークの近傍であると推定された連続する点の集合をピーク領域とする処理を、パラメータであって該パラメータの値を複数通り変化させることによって前記ピーク領域が変化するパラメータに基づいて行うことにより値を複数通り変化させた前記パラメータのそれぞれに対して異なるピーク領域を求めるステップと、

前記パラメータにより求められた前記ピーク領域の包含関係を求め、該包含関係を有する数の多少に基づいてピーク領域の第1の信頼度を付与するステップと、

前記単一の波形データを複数用意し、該複数の波形データによる複数の前記第1の信頼度に基づいて前記変数軸上における第2の信頼度を求めるステップと、

該第2の信頼度に基づいてグループ化範囲を決定し、前記単一波形データにおいてピーク領域内における前記強度軸上の最大点として求められるピークの頂点の群のうち前記グループ化範囲内に存在する前記ピークの頂点の群を1つのグループとしてグループ化するステップと

を有することを特徴とするピーク抽出方法。

【請求項2】

前記ピークの近傍であるか否かを推定するステップは、波形データ中の点Aを画定し、該点Aに対して前記変数軸において探索方向に存在する点Bを、前記探索方向であって前記点Aの前記変数軸における変数の減少する減少方向に順次探索するステップと、該点Aの前記変数軸における変数の増加する増加方向に順次探索するステップと、を含み、

前記探索するステップは、前記点Aに対して探索方向の前記点Bが下記のいずれかに該当する場合に探索終了し、またそのケースを探索終了ケースとし、前記減少方向の探索と前記増加方向の探索とに関する該探索終了ケースの結果に基づいて前記点Aがピークの近傍であるか否かを推定することを特徴とする請求項1に記載のピーク抽出方法。

ケース1：前記点Bの前記強度軸座標が上方閾値を上回る。

ケース2：前記点Bの前記強度軸座標が下方閾値を下回る。

ケース3：上記のいずれのケースにも当てはまらず、波形データの端部まで到達する。

但し、前記上方閾値と前記下方閾値とは、前記パラメータに基づいて決められる値である。

#### 【請求項3】

前記第1の信頼度を付与するステップは、前記ピーク領域からそれを包含するピーク領域が存在する限り順次ピーク領域を辿り、該辿った回数を信頼度とすることを特徴とする請求項1又は2に記載のピーク抽出方法。

但し、辿る場合に複数の経路がある際には、辿った回数が最大の場合の値を信頼度とする。

#### 【請求項4】

前記第2の信頼度を求めるステップは、前記変数軸を複数の区間に区切り、該区間毎の前記第1の信頼度に基づくピーク密度曲線を生成するステップを含み、

前記グループ化範囲を決定するステップは、前記ピーク密度曲線とあるしきい値とに基づいて信頼度のレベルに応じたグループ化範囲を決定するステップを含むことを特徴とする請求項1から3までのいずれか1項に記載のピーク抽出方法。

#### 【請求項5】

前記スペクトル特性は、生体由来分子に起因するスペクトル特性であることを特徴とする請求項1から4までのいずれか1項に記載のピーク抽出方法。

#### 【請求項6】

コンピュータに請求項1から5までのいずれか1項に記載のピーク抽出方法を実行させるためのプログラム。

#### 【請求項7】

変数軸と強度軸とを有する複数のスペクトル特性における波形データを取得する波形データ取得部と、

単一波形データにおけるピークを該ピークの信頼度を付して抽出するピーク抽出部と、前記ピークの信頼度に基づいて前記ピークをグループ化するグループ化処理部とを有するピーク抽出装置において、

前記ピーク抽出部は、それぞれの前記スペクトル特性の単一の波形データ中の点に対して、該点が前記変数軸上においてピークの近傍であるか否かを推定し、ピークの近傍であると推定された連続する点の集合をピーク領域とする処理を、パラメータであって該パラメータの値を複数通り変化させることによって前記ピーク領域が変化するパラメータに基づいて行うことにより値を複数通り変化させた前記パラメータのそれぞれに対して異なるピーク領域を求め、前記パラメータにより求められた前記ピーク領域の包含関係を求め、該包含関係を有する数の多少に基づいてピーク領域の第1の信頼度を付与し、前記単一の波形データを複数用意し、該複数の波形データによる複数の前記第1の信頼度に基づいて前記変数軸上における第2の信頼度を求め、

前記グループ化処理部は、前記第2の信頼度に基づいてグループ化範囲を決定し、前記単一波形データにおいてピーク領域内における前記強度軸上の最大点として求められるピークの頂点の群のうち前記グループ化範囲内に存在する前記ピークの頂点の群を1つのグループとしてグループ化することを特徴とするピーク抽出装置。

10

20

30

40

50

## 【請求項 8】

さらに、前記グループ化されたピークに含まれるそれぞれのピークの強度の違い又は前記単一波形データそれぞれの前記グループ化範囲における強度の違いの少なくともいずれか一方に基づいて、一つ以上の前記単一波形データからなる波形データ群であって複数の異なる波形データ群の間の差異を検出する差異検出部を有することを特徴とする請求項 7 に記載のピーク抽出装置。

## 【請求項 9】

生体由来分子に起因するスペクトル特性のピークを抽出することを特徴とする請求項 7 又は 8 に記載のピーク抽出装置。

## 【発明の詳細な説明】

10

## 【技術分野】

## 【0001】

本発明は、電子計算機を用いた情報処理によって、スペクトル特性におけるピークを抽出する技術に関する。また、ピーク抽出技術を用いた診断装置に関する。飛行時間型質量分析装置(TOF-MS:time-of-flight Mass Spectrometry)、液体クロマトグラフなどの質量分析装置を用いて血漿、血清、尿、組織抽出液、髄液などの様々な生体試料を測定することにより得られる波形データ(以下「波形データ」あるいは「波形」と称する。)から、その試料に含まれるタンパク質等生体由来分子に対応する「ピーク」部分を、計算機を用いた情報処理によって抽出する手法および装置に関する。

## 【背景技術】

20

## 【0002】

スペクトル特性におけるピークの抽出技術としては種々の技術が提案されている。一般的なピーク抽出手法では、例えば、SN比が一定値以上であるなどの条件に基づいて単一波形からピークを抽出する方法を用いていた(この欄において第1のステップと称する)。

## 【0003】

例えば由来の異なる単一波形を比較する場合には、第2のステップとして、上記第1のステップで異なる単一波形から抽出したピークをグループ化して比較することが必要となる場合が多い。この際、第1のステップで抽出された全てのピークは、第2のステップであるピークのグループ化処理において全て同等に取り扱われる(特許文献1参照)。

30

## 【0004】

単一波形から得られたピークを、複数波形間で重ね合わせ、重ね合わせた状態における隣接する複数波形間のピーク同士がある距離よりも近い位置にある場合には、それらを同一のピークグループとみなし、同一由来のピークとして特定する処理が行われる。同一のピークグループに属する複数のピークについて、一部のピークが他のピークとその強度に関して十分な差が認められる場合には、そのピークを有する波形に関しては特異性がみられることになり、例えば、既知の物質と比較することによってその特異なピークに対応する物質を同定することやその特異性が特定の疾患患者に固有のものであればそれを診断に用いることなどが可能となる。

## 【0005】

40

【特許文献1】特開2003-315341号公報

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0006】

しかしながら、上記のピーク抽出方法を用いると、以下のような場合にピーク抽出精度が低下する可能性がある。

## (1) ノイズによる質低下

単一波形から得られたピークはピークらしいものもノイズのようなものも、次の処理であるピークグループ化処理においては全て同等に扱われる。図15は、上記の一般的な処理を行ってピーク抽出を行い、次いで、ピークのグループ化を行う手順の例を示す図であ

50

る。図15の左側中段の図に示すようにノイズが有る場合に、ピークのグループ化処理においてこのノイズに起因するピークも1つのピークグループとみなされる。その結果、ノイズのみ、或いは、大部分がノイズからなるピークグループが生成され易く、かかるノイズを多く含むピークグループとノイズを多くは含まないピークグループとで区別が付かないという問題点がある。

【0007】

(2) ピーク密集区間での質低下

もう1つの問題であるピーク密集区間でのピーク抽出の質低下について図16を参照しつつ説明を行う。図16は、上記の一般的な処理を行ってピーク抽出を行い、次いで、ピークのグループ化を行う手順の例を示す図である。

10

【0008】

図16に示すような細かなピークが密集している区間では、隣接ピークの間隔が狭まっている。そのため、一定の距離以内にあるピークをすべてグループ化する従来の単純なピークグループ化方法では、右図に示すように、結果として多くのピークが同一のグループに割り当てられてしまい、各波形から得られたピークの対応付けが正しくなされない。特にブロードなピーク、即ちピークトップが不明瞭あるいは隣接するピークが重なり合っている状態が多くみられる高分子領域においては、ピーク抽出によって得られた結果は図16に示すような「細かなピークが密集する」状態である場合が多く、この問題がより生じやすい。

【0009】

本発明は、上記の問題点に鑑みて、ピーク抽出とグループ化されたピークの対応付けとの精度を向上させることを目的とする。

20

【課題を解決するための手段】

【0010】

本発明は、個々のピークに対して信頼度を付与し、複数の波形において同じ位置に存在するピークの信頼度の合計値をピーク密度として求めこのピーク密度をもとにピークグループの信頼度を付与する方式を採用している。これにより、大部分がノイズのピークグループ、或いは、ごく一部の波形にしか見られない例外的なピークからなるグループの信頼度は低く設定される。これにより、ノイズによる質低下という上記(1)の問題に対しては、信頼度の低いピークグループを除去することにより対処できる。上記(2)の問題に対しては、近傍ピークを単純にグループ化するのではなく、微小区間ごとにピーク密度を計算し、ピーク密度の増減を考慮したグループ化を行うため、よりピークらしい部分のみをグループ化することが可能である(図17参照)。

30

【0011】

さらに、ピークグループに信頼度を付与するということは、ピークグループの序列化を行うことに対応する。従って、大量のピークグループが得られた場合に、信頼度の高いピークグループのみ、或いは、信頼度の高いピークグループから、後続の作業を行うことができ、作業の効率化が可能になる。

【発明の効果】

【0012】

本発明は、個々のピークに対して信頼度を付与し、複数の波形において同じ位置に存在するピークの信頼度の合計値をピーク密度として求め、このピーク密度をもとにピークグループの信頼度を付与する方式を採用している。これにより、大部分がノイズのピークグループ、或いは、ごく一部の波形にしか見られない例外的なピークからなるグループの信頼度は低く設定される。これにより、ノイズによる質低下という上記(1)の問題に対しては、信頼度の低いピークグループを除去することにより対処できる。上記(2)の問題に対しては、近傍ピークを単純にグループ化するのではなく、微小区間ごとにピーク密度を計算し、ピーク密度の増減を考慮したグループ化を行うため、よりピークらしい部分のみをグループ化することが可能である。

40

【0013】

50

さらに、ピークグループに信頼度を付与するという事は、ピークグループの序列化を行うことに対応する。従って、大量のピークグループが得られた場合に、信頼度の高いピークグループのみ、或いは、信頼度の高いピークグループから、後続の作業を行うことができ、作業の効率化が可能になる。

【発明を実施するための最良の形態】

【0014】

本明細書において、ピーク抽出処理とは、単一波形からピークを抽出し、抽出したピークを複数の波形でグループ化する処理を指す。

【0015】

以下、本発明の実施の形態によるピーク抽出技術について、図面を参照しつつ説明を行う。まず、具体的な実施の形態について説明する前に、本発明の概要についてSurface-Enhanced Laser Desorption/Ionization TOF MS (SELDI TOF - MS) を利用した質量分析を例にして、図1A、Bを参照しつつ説明を行う。

【0016】

図1Aは、プロテオーム解析装置の位置付けを示す図である。図1Aに示すように、SELDI TOF - MSは、例えば、金属製のチップと飛行時間(Time-of-flight)型の質量分析計を組み合わせたシステムである。生体内から得られた少量のサンプルを直接あるいは粗分画して用いることで、短時間で様々なタンパク質の解析を行うことができる。チップの表面にイオン交換性、IMAC(Ni、Cuなどの金属イオン)などの化学的な修飾が施されており、化学的性質を利用してチップ表面にタンパク質を捕捉し、TOF - MSで測定する。生体試料中のタンパク質の発現解析やタンパク質の同定、翻訳後修飾の解析などを行うことが可能である。このSELDI TOF - MSにより測定されたプロテオーム波形データは、横軸に分子量あるいはそれに準ずる量としての質量電荷比(m/z)を、縦軸に分子数あるいはそれに準ずる量としての強度をプロットしたスペクトルとして得られる。この場合には、各m/zに対応するピークがタンパク質等の特定分子に対応するピークである。さらに、このSELDI TOF - MSを用いて血漿等の生体試料を測定することにより得られる波形データから、その試料に含まれるタンパク質等生体由来分子に対応する「ピーク」部分を、計算機を用いた情報処理によって抽出するピーク抽出装置を備えている。SELDI TOF - MSを用いて複数の試料からそれぞれ得られた波形データに対し、同一分子由来と推測されるピークを複数波形データ間で対応づけし、ピークのグループ化を行う。ピークのグループ化の質は、その後のデータ解析処理の精度に大きな影響を与え、ひいては、そのデータ解析に基づく診断等の精度に大きな影響を与える。

【0017】

図1Bは、本実施の形態によるピーク抽出装置の概略構成例を示す図である。図1Bに示すように、本実施の形態によるピーク抽出装置は、波形データを取得する波形データ取得部1と、単一波形データにおけるピークを抽出するピーク抽出部3と、ピークのグループ化を行うグループ化処理部5と、データを出力するデータ出力部7と、を有している。さらに、グループ化されたピークに基づいて、単一波形データ間の差異を検出する差異検出部11を有する。それぞれの処理部は、実際には以下に説明する方法をコンピュータに実行させるためのプログラムを格納するROMと、ROM内のデータを展開するRAMと、RAMに展開されたプログラムを実行するCPUと、により処理を行うものである。

【0018】

本実施の形態によるピーク抽出装置は、単一の波形データからピークを抽出し、そのピークに信頼度を付与する。その後、複数波形にまたがるピークのグループ化と信頼度の付与を行う。

【0019】

すなわち、まず、波形データに関して、後述するピーク領域の幅を変化させることができるパラメータを設定し、このパラメータを任意に変えることにより、細かいピークから粗いピークまで様々なピークを「領域」(以下、「ピーク領域」と称する。)として抽出する。またそれらのピーク領域の包含関係からピーク領域の重なり度合い(以下、「包含

10

20

30

40

50

レベル」と称する。)を求める。この「包含レベル」は、「ピークらしさ」、すなわち抽出したピークの信頼度を表す指標として用いることができる。

#### 【0020】

この場合の複数波形データにまたがるピークの対応付けに際し、上述の「包含レベル」を用いてピークグループ自体の信頼度を数値化する。即ち包含レベルの高い(=信頼度の高い)ピークが多くの波形データで見られる位置では、高い信頼度のピークグループが得られ、一方、逆の場合には得られるピークグループの信頼度は低いと判断することができる。

#### 【0021】

以上に説明したように、ピークのグループ化の質を求めることにより、その後のデータ解析処理の精度、ひいては、最終的なデータ解析に基づく診断等の精度を知ることができる。

10

#### 【0022】

以下、単一波形データからのピーク抽出処理、複数波形データからのピークのグループ化処理の順に説明を行う。図2は、本実施の形態による単一波形データからのピーク抽出処理(ステップ1)及びピークのグループ化処理(ステップ2)の概要を示す図である。図2に示すように、複数の対象について測定してそれぞれ得られた単一の波形1から波形Nまでを、ステップ1において、波形1から波形Nまでのそれぞれの単一波形データからのピーク抽出を行い、次いで、ステップ2においてピークのグループ化を行う。以下に、ステップ1の単一波形データからのピーク抽出処理、ステップ2のピークのグループ化処理の順に説明を行う。

20

#### 【0023】

##### (1)ステップ1：単一波形データからのピーク抽出処理

本実施の形態によるピーク抽出装置では、ピークの存在する部分をピーク領域として検出する。このピーク領域として検出する処理において、内部で用いるパラメータを様々な値に設定することにより、様々な幅を持つピーク領域が得られる。換言すれば、様々な幅を持つピーク領域が得られるようにパラメータを選択してピーク領域を検出する処理を行う。このパラメータの値が大きい場合は、高いピークを含む幅の広い領域が得られる。逆にパラメータ値が小さい場合は、低いピークを含む幅の狭い領域が得られる。一般にパラメータ値が小さいと多くのピークが得られるがノイズも多く含まれるようになる。一方、パラメータ値が大きいとノイズが少ない一方でピークの検出漏れの確率が大きくなる。

30

#### 【0024】

本実施の形態によるピーク抽出装置では、パラメータ値を数通りに変更し、それぞれのパラメータ値毎に幅の異なるピーク領域を得る。これらのピーク領域の重なり具合からその包含関係を構成し、包含関係の深さを「包含レベル」として数値化する。この包含レベルを「ピークらしさ」の指標として取り扱うことにより、ピークに信頼度を付与することができ、ピークとしての信頼度を伴ったピーク領域を検出することができる。

#### 【0025】

図3は、単一波形データからピークを抽出し、包含レベルを求める過程を示す図である。図3に示す単一波形は、横軸が $m/z$ 、縦軸が強度である。但し、 $m$ は質量、 $z$ は電荷である。図3に示すように、右側から4つのピークが抽出されている(順に、ピーク1~4とする)。内部のパラメータを0.5、1、5、10とした場合のピーク領域を検出する。内部のパラメータを0.5とした場合に、ピーク1からピーク4までのいずれのピークにおいてもピーク領域が検出される。内部パラメータ=1でのピーク領域は、ピーク2及びピーク4においてピーク領域が検出される。内部パラメータ=5及び10でのピーク領域は、ピーク4においてピーク領域が検出される。

40

#### 【0026】

このように4種類の異なる内部パラメータ値におけるピーク領域を検出し、ピーク1~4のそれぞれに関して、検出されたピーク領域の構造に基づいて包含レベルを決定する。

#### 【0027】

50

以下に、ステップ1についてより詳細に説明する。ステップ1は、以下の2つのサブステップに分割される。以下、各サブステップについて詳細に述べる。

【0028】

(1-1) ピーク領域の検出

本実施の形態によるピーク領域検出ステップにおいては、図3に示すように、内部のパラメータPの値を数通りに変化させ、各値Pに対して単一波形からのピーク領域の検出を行う。変化させるパラメータPの具体的な値は予め与えられており、全ての波形データにおいて、その数、値ともに共通であるものとする。例えばN個のパラメータ値が事前に与えられていれば、単一の波形に対してN回のピーク領域の検出を行い、それぞれのパラメータ値に対して別個のピーク領域検出結果が得られる。

10

【0029】

ピーク領域検出処理は、波形中のすべての点に対して、その点がピークの近傍であるか否かを推定することによって行う。ピークの近傍であると推定された連続する点の集合が一つのピーク領域となる。各点がピークの近傍であるか否かの推定は、当該点の左方および右方を順次探索し、予め設定した閾値を超えて増加・減少しているか否かによって行う。当該点の推定の具体的手順について図4(a)から(d)までを参照しつつ具体的に説明する。図4(a)から図4(d)までは、探索を終了する4つのケースである。

【0030】

(1) まず、当該点を点Aとする。(2) 次に、点Aの左方、即ち、 $m/z$ 値が減少する方向を探索する。この際、(ア)BをAとする。(イ)Bの左隣の点を改めてBとする。(ウ)Bが下記探索終了ケースのいずれかに当てはまれば、下記(3)へ、それ以外は上記(イ)の処理に移る。(3) 点Aの右方探索を行う。すなわち上記(2)と同様の処理を右方即ち $m/z$ 値の増加する方向に対して行う。(4) 点Aの左方探索および右方探索の探索終了ケースをもとに、表1に基づいて、点Aがピークの近傍であるか否かを推定する。

20

【0031】

ここで、探索終了ケースは、以下の通りである。すなわち、図4(a)に示すように、探索終了ケース1として、点Bの強度が上方閾値 $T_1$ を上回った場合である。この場合には、左側により高いピークが存在することが推定されるため、点Aはピークの近傍ではないと推定できる。図4(b)に示すケース2では、点Bの強度が下方閾値を下回る。この場合には、自分(点A)よりも点Bが下がっているため、点Aはピークの近傍である可能性があるが、右方を探索するまではピークの近傍と推定することはできない。図4(c)に示すケース3では、ケース1、2のいずれかの状態になることなく、点Aから予め定められた距離Dまで到達した場合である。図4(d)に示すケース4は、上記のいずれにも合致せずに波形の端まで到達した場合である。ここで、図4(c)ケース3におけるDは、A又はBに依存する可変値であっても良い。

30

尚、図4(c)は省略することが可能である。

【0032】

【表 1】

表 1

ピークの近傍の推定表

		右方探索			
		ケース1	ケース2	ケース3	ケース4
左方探索	ケース1	×	×	×	×
	ケース2	×	○	左	○
	ケース3	×	左	左	左
	ケース4	×	○	左	○

○:ピークの近傍      ×:ピークの近傍ではない      左:Aの左隣点の推定結果と同じにする

10

【0033】

表1は、ピークの近傍を推定するための推定表である。「○」は、ピークの近傍であることを意味する。「×」は、ピークの近傍ではないことを意味する。「左」は、点Aの左隣の点の推定結果と同じにする（前の推定結果で代用する）ことを意味する。

【0034】

左方と右方とを探索し、それぞれの探索結果がケース1～4までのいずれに当てはまるかを判定することにより、ピークの近傍であるか否かを推定することができる。この推定結果に基づいて、ピーク領域を決めることができる。尚、表1及び以下の計算手法は、例示であり、変更又は変形が可能であることはいうまでもない。

20

【0035】

尚、上記探索終了ケースにおける「上方閾値 T1」および「下方閾値 T2」は、パラメータPに依存する値である。閾値の決め方には様々な方法があるが、以下の表2に代表的な3つの方法を示す。

【0036】

【表 2】

30

表 2

方法1	$T1=Y(A)+P$ $T2=Y(A)-P$	方法2	$T1=Y(A)+P$ $T2=\max\{Y(A),\dots,Y(B)\}-P$
方法3	$T1=Y(A)+n(A)\times 2^P$ $T2=\max\{Y(A),\dots,Y(B)\}-n(A)\times 2^P$		

40

Y(A):Aの強度       $\max\{Y(A),\dots,Y(B)\}$ :AからBに至るすべての点の強度の最大値

n(A):Aの位置におけるノイズの強さ

【0037】

ここで、 $Y(A)$ は、AのY座標即ち強度であり、 $\max\{Y(A),\dots,Y(B)\}$ は、AからBに至るすべての点の強度の最大値であり、 $n(A)$ はAの位置におけるノイズの強さを示す。表2に示されるように、T1、T2は、A又はBに依存する可変値であっても良い。

【0038】

50

表 2 に示すように、方法 1 では、T 1 が点 A の強度よりもパラメータ P だけ上がった値、T 2 が点 A の強度よりもパラメータ P だけ下がった値として定義される。方法 2 では、T 1 が点 A の強度よりもパラメータ P だけ上がった値、T 2 が位置 A から B までの間における強度の最大値からパラメータ P だけ下がった値として定義される。方法 3 では、T 1 は点 A の強度よりもノイズの強さとパラメータ P とで表される関数  $n(A) \times 2^P$  だけ大きい値であり、T 2 は位置 A から B までの間における強度の最大値からノイズの強さとパラメータ P とで表される関数  $n(A) \times 2^P$  だけ小さい値である。ここで、関数  $n(A) \times 2^P$  は、 $P = 1$  であれば、ノイズの強さ  $n(A)$  の 2 倍が上記の関数となる。以下では、方法 3 を採用した場合について説明する。

以上のような方法により、単一波形におけるピーク領域を検出することができる。

10

#### 【 0 0 3 9 】

##### ( 1 - 2 ) ピーク領域の選別と包含レベルの設定

サブステップ ( 1 - 2 ) では、上述のサブステップ ( 1 - 1 ) において得られた数通りのパラメータ値によるピーク領域群に対して、それらの包含関係をもとにピーク領域の選別および包含レベルの設定を行う。図 5 は、サブステップ ( 1 - 1 ) で得られたピーク領域の例を示す図である。以下、図 5 を参照しつつサブステップ ( 1 - 2 ) の処理について説明する。パラメータ P を様々に変化させて得られたピーク領域、すなわち、パラメータ P が  $P_1$ 、 $P_2$ 、 $P_3$ 、 $P_4$  の場合について得られた図 5 において太線で示される全てのピーク領域に関して、そのうちの任意の 2 組に対する包含関係を検査する。図 6 は、包含関係について検査した様子を示す図である。図 6 に示すように、包含関係を、包含される側 (例えば符号 A で示す。) から包含する側 (例えば符号 B で示す。) へ向けての矢印により示している。尚、包含関係を検査する際に、一方が他方を完全に含まなくともほぼ含んでいれば包含関係が成立すると判断されるように、後述する「包含関係成立条件」を設定する。これにより、若干のはみ出しを許容する包含関係の構築が可能となる。なお、包含関係は必ずしもパラメータ P の大小とは一致しない。

20

#### 【 0 0 4 0 】

図 7 は、包含関係成立条件のうち不完全な包含関係ではあるが包含していると思なすことができる例を示す図である。下記の式 ( 1 ) を満たすような完全な包含関係が成立しなくても、式 ( 2 )、( 3 ) に示す条件が全て満たされる場合には、ピーク領域 A はピーク領域 B に包含されるものとする。

30

#### 【 0 0 4 1 】

完全な包含関係は以下の関係である。

$$c < 0 \quad ( 1 )$$

但し、ピーク領域の左端、右端ともに満たされることが条件である。

#### 【 0 0 4 2 】

不完全ではあるが、本実施の形態において包含関係になるとみなす関係は以下の関係である。

$$c < k a \quad ( 2 )$$

#### 【 0 0 4 3 】

但し、ピーク領域の左端、右端ともに満たされることが条件である。

40

$$p b \quad a < q b \quad ( 3 )$$

#### 【 0 0 4 4 】

但し、a : ピーク領域 A の幅    b : ピーク領域 B の幅    c : ピーク領域 A のはみ出しの長さ、k , p , q : 予め決められら定数 (例 :  $k = 0.2$  ,  $p = 0.2$  ,  $q = 0.95$  )。k は、0.2 から 0.3 の間、q は 1 に近い 1 以下の値が適当である。状況によって調整することで、包含関係を正確に把握することができる。

#### 【 0 0 4 5 】

次に、図 8 に示すように、各ピーク領域の包含レベル ( CL ) を設定する。図 8 に示すように、各ピーク領域からそれを包含するピーク領域が存在する限りそれらを順次辿る。すなわち、矢印の方向に包含関係がとれずに矢印を付すことができなくなるまで辿る。そ

50

して、その辿った回数を包含レベル  $CL$  と定義する。但し、複数の経路がある場合は、辿った回数の最大値を包含レベル  $CL$  とする。

【0046】

次に、図9に示すように、他のいずれかのピーク領域を包含しているピーク領域、および、他のピーク領域との包含関係が存在しないピーク領域を取り除く（矢印の指示される側に相当する領域を除去する）。これにより、他のピーク領域に包含されており、かつ、他のピーク領域を包含していないピーク領域のみを残すことができる。

【0047】

最後に、図10に示すように、2つのピーク領域間で重なりがある場合は、いずれか一方のみを取り除く。どちらを取り除くかについては様々な決め方があるが、「パラメータ  $P$  の小さいほうを取り除く」と規定した場合の例を図10に示している。

10

以上の手順により、ピーク領域を精度良く抽出することができる。さらに、抽出されたピーク領域には、包含レベル、すなわち、ピーク領域の信頼度が付与されている。尚、ここでは、ピーク領域内において最も強度の大きい点をそのピーク領域の頂点とする。

【0048】

(2) ステップ2：複数波形データからのピークのグループ化

ステップ2においては、複数の波形データから得られたピークのグループ化処理を行う。複数の波形データそれぞれにおいてピーク抽出を行い、それぞれのピークを比較する場合は、ある波形からのピークが他の波形におけるどのピークと対応するのかを決める必要がある。

20

【0049】

しかしながら、対応する複数ピークの  $m/z$  値が完全に一致することは稀であり、 $m/z$  値の完全一致をもってピークの対応をとることは困難である。そこで、 $m/z$  値が多少ずれていても、比較的近いピークに関しては同じグループに属するものとして適切にグループ化する必要がある。本実施の形態によるピーク抽出装置では、ステップ1において得られた各ピーク領域の包含レベルを利用することにより、信頼度を付加したピークのグループ化を実現する。ピークのグループ化処理は以下のサブステップからなる。

(2-1) 複数の波形から得られたピーク領域の頂点座標を重ね合わせる。

(2-2)  $m/z$  軸を微小区間に分割する。

(2-3) 区間ごとに包含レベルの合計を計算し、全区間にわたるピーク密度曲線を得る

30

(2-4) ピーク密度曲線をもとに、ピークのグループ化範囲を決定する。

(2-5) 同一のグループ化範囲に属するピークをグループ化する。

【0050】

以下では、上記各サブステップについて説明を行うが、その前にステップ2の流れの概略について図11を参照しつつ説明を行う。まず、個々の波形データからのピークであって、ステップ1において抽出されたピーク領域とピークの頂点座標とを有する波形データを準備する。これらの個々の波形におけるピークの頂点座標を重ねて表示し、この状態において  $m/z$  軸を微小区間に分割する。次いで、この微小区間毎に包含レベルの合計を計算し、曲線化する。これをピーク密度曲線と称する。このピーク密度曲線に対して、あるしきい値  $L$  を基準にしてこれを越える値をグループ化範囲とする。これにより、グループ化範囲に対して信頼度を付与することができる。

40

以下に、より詳細に各サブステップについて説明する。

【0051】

(2-1)：ピーク頂点の重ね合わせステップ

まず、各単一波形から求めたピークの頂点を、 $m/z$  軸、強度軸双方が揃うように重ね合わせる。

【0052】

(2-2)： $m/z$  軸の分割

$m/z$  軸を微小区間に分割する。区間の分割方法としては、a) 区間幅を固定にして分

50

割する方法、b) 区間幅を変動させて分割する方法、c) 隣り合う区間同士の重なりを許容する分割方法、d) 許容しない分割方法など、がある。以下に表3と図12とを参照しつつ代表的な分割方法を挙げて説明する。図12は、分割方法1～4までの具体的な分割例を示す図である。

【0053】

【表3】

表 3

分割方法1	波形の左端から等間隔に分割(区間の重なりなし)
分割方法2	波形の左端から、 $m/z$ の値に応じて幅が変動する区間に分割(区間の重なりなし)
分割方法3	小数点以下が0である位置を中心に、幅が固定(例えば1.2)の区間に分割(区間の重なりあり)
分割方法4	小数点以下が0である位置を中心に、 $m/z$ の値に応じて変動する幅を持つ区間に分割(区間の重なりあり)

10

20

【0054】

表3に示す方法のうち、分割方法1は、波形の左端から等間隔に分割(区間の重なりなし)する方法であり、 $m/z$ に関して等間隔に分割する基本的な方法である。これに、対して、分割方法2は、波形の左端から、 $m/z$ の値に応じて幅が変動する区間に分割(区間の重なりなし)する方法である。例えば、SELDI TOF-MSでは、 $m/z$ が大きくなると、解像度(分解能)が粗くなる(低下する)ため、このような方法により $m/z$ 軸を実効的に等間隔に近い方向にすることができる。分割方法3は、少数点以下が0である位置を中心に、幅が固定(例えば1.2)の区間に分割(区間の重なりあり)する方法であり、分割範囲の中心位置同士が等間隔になる。分割方法4の方法は、少数点以下が0である位置を中心に、 $m/z$ の値に応じて変動する幅を持つ区間に分割(区間の重なりあり)する方法であり、分割方法2と分割方法3とを加味した方法であり、本実施の形態では、SELDI TOF-MSに適用することを考慮して適正化された分割方法4を採用している。

30

【0055】

(2-3) ピーク密度曲線の導出

$m/z$ 軸の分割に関するサブステップ(2-2)で得られた区間のそれぞれについて、当該区間に頂点が含まれるピーク領域を全て列挙し、それらの包含レベルの合計値を当該区間に対するピーク密度とみなす処理を行う。包含レベルが高いほど、また、ピークが多く含まれる区間ほど、ピーク密度が高くなる。それぞれの区間毎にピーク密度を計算しこれを全ての区間に対してプロットすることにより、このプロットに沿ったピーク密度曲線が得られる。尚、ピーク密度計算方法には、上述の包含レベルの合計を行う基本的な方法(基本方式)の他に、様々な代替方法が存在する。表4にピーク密度計算方法の代表的な例を示す。

40

【0056】

【表 4】

## 表 4

ピーク密度 計算方法1	<p>ある単調増加関数F(x)に対して、 ピーク密度 = <math>\sum_{i=1, n} F(CLi)</math>とする。 (P1~Pn : 当該区間に属するピーク) (iCLi : Piの包含レベル)</p> <ul style="list-style-type: none"> <li>・F(x)=x →基本方式と同じ</li> <li>・F(x)=x・x →包含レベルの2乗の合計</li> <li>・F(x) = <math>\begin{cases} 0(x &lt; 3) \\ 1(x \geq 3) \end{cases}</math> →包含レベルが3以上のピークの個数</li> </ul>
ピーク密度 計算方法2	<p>当該区間を中心として含み幅が当該区間よりS倍広い区間を参照区間とする。当該区間および参照区間について、前項の方法「ピーク密度計算方法1」でピーク密度を計算してそれぞれD1、D2とし、数値S・D1/D2を改めて当該区間のピーク密度とする。当該区間の周辺に比べた前項方法の数値の相対的な高さを表す。(図 13参照)</p>

10

20

## 【0057】

表4に示すピーク密度計算方法1は、ある単調増加関数F(x)に対して、ピーク密度を以下の式で求める。

$$\text{ピーク密度} = \sum_{i=1, n} F(CLi) \quad (4)$$

## 【0058】

但し、式(4)において、P1からPnまでは、当該区間に属するピークであり、CLiはPiの包含レベルである。単調増加関数F(x)を具体的に決めることによって、様々なピーク密度計算方法を実現できる。以下にいくつかの例を示す。

$$F(x) = x \quad (5)$$

## 【0059】

単調増加関数F(x)が式(5)で表される場合は、ピーク密度計算方法1は基本方式と一致する。即ちピーク密度は包含レベルの合計に相当する。

他には以下の式を用いることができる。

$$F(x) = x \cdot x \quad (6)$$

## 【0060】

F(x)として式(6)を用いると、ピーク密度は包含レベルの2乗の合計に相当する。式(6)は、包含レベルが高いほどピーク密度に与える影響が大きいことを考慮して重み付けしたものである。

$$F(x) = \begin{cases} 0(x < 3) \\ 1(x \geq 3) \end{cases} \quad (7)$$

40

式(7)は、包含レベルの低いピークはノイズに起因するピークと推定して無視し、包含レベルが3以上のもののみを考慮する方法である。

## 【0061】

次に、ピーク密度の計算方法2について説明する。ピーク密度計算方法2では、当該区間を中心として含み、幅が当該区間よりS倍広い区間を参照区間と規定する。当該区間および参照区間について、上記ピーク密度計算方法1のいずれかの方法によりピーク密度を計算し、それぞれD1、D2とし、数値S・D1/D2を改めて実際のピーク密度とする。図13は、この方法によりピーク密度を求めるピーク密度の計算方法の概念図である。当該区間の周辺に比べた前項方法の数値の相対的な高さを表す。この方法によれば、前後の広い範囲を参照して前後と比較することで、より精度の良いピーク密度を求めること

50

ができる。尚、本実施の形態では計算方法 2 による方法を用いている。また計算方法 2 で用いる下位のピーク密度計算方法には単調増加関数 $F(x)$ として式(6)を用いている。

【0062】

(2-4) グループ化範囲の決定

サブステップ(2-3)で得られたピーク密度曲線をもとに、グループ化範囲を以下の手順により決定する。図14は、グループ化範囲決定方法の概略を示す図である。まず、第1に、ピーク密度曲線において閾値 $L_0$ よりも大きいピーク密度を持つ区間が存在しなければ終了する。第2に、ピーク密度曲線において閾値 $L_0$ よりも大きいピーク密度を持つ区間が存在すれば、最も高いピーク密度を持つ区間をグループ化の中心 $C$ とする。またそのピーク密度を $L$ とする(a)。第3に、後述するような予め指定された方法により、中心 $C$ を含む一定の範囲をグループ化範囲とする(a')。またそのグループ化範囲の信頼度を $L$ とする(a)。第4に、グループ化範囲内の全区間のピーク密度を「0」にする(b)。また、グループ化範囲の左方および右方を順次検査し、ピーク密度が減少する限りにおいて、その区間のピーク密度を「0」にする(b)。ここで第1に戻り、残るピークに関して同様の操作を行う(この際、都度 $L$ が変更される、(c, d, e))。閾値 $L_0$ よりも大きいピーク密度を持つ区間が存在しなくなれば、処理を終了し最終的なグループ化範囲を確定する(e)。上述の $C$ を含む一定の範囲をグループ化範囲とする方法としては種々の方法を用いることができる。代表的なグループ化範囲を決定する方法を表5に示す。

10

【0063】

【表5】

20

表 5

グループ化範囲決定方法1	Cを中心に±wの区間をグループ化範囲とする。 wは固定値でも良いし、Cの $m/z$ の値に依存して変動する値でもよい。
グループ化範囲決定方法2	Cの左方および右方を順次検査し、ピーク密度が $L'$ まで下がったところまでをグループ化範囲とする。 ここで、 $L'$ は、Cのピーク密度 $L$ を用いて、例えば以下のような式により決められる値である。 $\cdot L' = L \cdot \alpha \quad (\alpha \text{は } 0 \leq \alpha < 1 \text{ なる定数 ; } 0.1 \text{ など})$ $\cdot L' = L - \beta \quad (\beta \text{は } 0 < \beta \leq L \text{ なる定数})$
グループ化範囲決定方法3	上記を併用する。

30

【0064】

表5におけるグループ化範囲決定方法1は、点Cを中心に±wの区間をグループ化範囲とする方法である。wは固定値でも良いし、点Cの $m/z$ の値に依存して変動する値でもよい。グループ化範囲決定方法2は、点Cの左方および右方を順次検査し、ピーク密度が $L'$ まで下がったところまでをグループ化範囲とする方法である。ここで、 $L'$ は、点Cのピーク密度 $L$ を用いて、例えば以下のような式により決められる値である。式としては、第1に $L' = L \cdot \alpha$ である。但し、 $\alpha$ は $0 < \alpha < 1$ なる定数であり、例えば0.1程度の値である。或いは、 $L' = L - \beta$ に基づいて決定することも可能である。但し、 $\beta$ は $0 < \beta \leq L$ となる定数である。グループ化範囲決定方法3は、上記1と2とを併用する方法である。本実施の形態では、グループ化範囲決定方法3を用いている。

40

【0065】

上記の方法1ではピーク密度の変化を考慮せず予め定められた範囲をグループ化範囲とするため、細かなピークが密集している部分では、隣り合う別ピークが同じグループ化範

50

囲に割り当てられるという問題がある。一方、方法 2 では得られるグループ化範囲の幅に制限を設けてないため、実際には有り得ない非常に幅の広いグループ化範囲が得られる場合がある。方法 3 では両方法を併用することで、これらの問題が生じるのを回避できる。

【0066】

(2-5) ピークのグループ化

元に戻り、各ピーク頂点を見て、ピーク頂点がサブステップ(2-4)において得られたいずれのグループ化範囲に所属するかについて検査する。そして、同じグループ化範囲に属するピークをグループ化する。いずれのグループ化範囲にも属さないピークはグループ化されず、即ちノイズであったとみなす。

【0067】

以上のように、本発明の実施の形態においては、個々のピークに対して信頼度を付与し、複数の波形において同じ位置に存在するピークの信頼度の合計値をピーク密度として求め、このピーク密度をもとにピークグループの信頼度を付与する。これにより、大部分がノイズのピークグループ、或いは、ごく一部の波形にしか見られない例外的なピークからなるグループの信頼度は低く設定される。従って、ノイズによる質低下を防止するとともに、近傍ピークを単純にグループ化するのではなく、微小区間ごとにピーク密度を計算し、ピーク密度の増減を考慮したグループ化を行うため、よりピークらしい部分のみをグループ化することが可能である。

【0068】

さらに、ピークグループに信頼度を付与することで、大量のピークグループが得られた場合に、信頼度の高いピークグループのみ、或いは、信頼度の高いピークグループから、後続の作業を行うことができ、作業の効率化が可能になるという利点がある。

【実施例】

【0069】

以下に、40匹のラットの血漿を用いた例について記載する。図18においては、用いる40匹のラットはその性質により「正常なラット」、「外科的処置を施したラット」などのいくつかの群に分けられるものとする。図18は、本実施例による処理の流れを示すフローチャート図である。図18に示すように、ステップS1において、SELDI TOF-MSにより測定を行う。採取されたラット40匹分の血漿に対し、SELDI TOF-MSを用いてそれぞれ2回測定を行い、測定結果を計算機に取り込める形式の波形データとして出力する。1回の測定に対しそれぞれ波形データが得られる(ステップS1)。

【0070】

次に、計算機に取り込まれた各波形データに対して、後述する計算機処理(ステップS2)を行う。ステップS2において、まず、波形データを準備し、ベースラインの除去(ステップS2-1)と平滑化処理を行い(ステップS2-2)、ステップS2-3でピーク抽出処理を行う。ステップS2-3のピーク抽出処理においては、ステップS2-3-1の単一波形からのピーク抽出(上記ステップ1)と、ステップS2-3-2の複数波形からのピークのグループ化処理(上記ステップ2)を行う。次いで、ステップS2-4において、ステップS2-1の結果またはステップS2-2の結果と、ステップS2-3の結果とに基づいて、グループ間の差異を検出する。

【0071】

ベースラインの除去(ステップS2-1)および平滑化の処理(ステップS2-2)によって得られた波形データ(図中(a)点に相当)のサンプル10例を図19に示す。図19に示すサンプル1から10までの各波形データをピーク抽出装置に入力すると、それぞれ個別に単一波形からのピーク抽出の処理が施される。その結果、各波形データに対するピークと各ピークに対する包含レベルとが得られる(図18(b)点に相当)。図19のサンプル1に対するピーク抽出結果を図20に示す。図20中の一部を拡大したデータを図21に示す。これらの図中の波形は平滑化処理後のデータである。また図中の点は、ピーク頂点を示す。

10

20

30

40

50

## 【 0 0 7 2 】

各波形データに対してピーク抽出および包含レベルの算出を行った後、「ステップ2 複数波形からのピークのグループ化」を行う。グループ化の結果を図 2 2 に示す。また、その一部を拡大したデータを図 2 3 に示す。図中の上段においては、サンプル 1 から 1 0 までの波形データと抽出されたピーク頂点とが示されている。図中の中段は、得られたピーク密度曲線が描かれている。図下段では、ピーク密度曲線をもとにグループ化されたピークが示されている。同じマークで示されたピーク頂点が同一グループに属するものである。

## 【 0 0 7 3 】

次にピーク抽出装置から得られたピークグループ及びそのグループ化範囲に基づき、正常なラット群(以下A群)と外科的処置を施したラット群(以下B群)の差異を検出する。図 2 4 は A 群と B 群とで差がみられる代表的なグループ化範囲周辺において、A 群と B 群とを上下にわけて波形データを示した図である。図 2 5 は同じグループ化範囲周辺におけるピーク頂点あるいは後述の代表点を示した図である。以下に、図 2 4、図 2 5 を参照しつつ差異検出について説明する。まず、あるピークグループに含まれるすべてのピーク頂点について、それぞれのピーク頂点の強度を得る。得られた強度の中で A 群(図 2 5 の黒丸)または B 群(図 2 5 の×印)のラットに関するものを取り出し(強度取得処理)、Mann-Whitney の U 検定などの適切な統計的検定手法によって A 群と B 群の強度の差に対する有意確率を算出する。ピーク抽出装置から得られた全ピークグループに対して同様の有意確率を算出し、十分小さい有意確率が得られたピークグループを A 群と B 群の間で差異が認められるピークグループとして検出する。尚、上記「強度取得処理」の留意点について、以下に説明する。

## 【 0 0 7 4 】

(i) A 群または B 群のラットに関するある同一波形データ上のピークが、当該ピークグループに複数含まれる場合は、それらの中から一つのピークを選んで代表点とし、代表点に対する強度のみを取り出す。代表点の選び方としては、例えば強度が最大のものを代表点とする方法などがある。

## 【 0 0 7 5 】

(ii) A 群または B 群のラットに関するある同一波形データ上のピークが、当該ピークグループに一つも含まれない場合は、当該ピークグループのグループ化範囲で前述の波形データ上の代表点を一つ選び、その強度を取り出す。代表点の選び方としては、例えば最大の強度を持つ点を代表点とする方法などがある。

## 【 0 0 7 6 】

(iii) (i) および (ii) において代表点を選択する際、対象とする波形データは平滑化前と平滑化後(それぞれ図 1 8 中の(e)、(f))のどちらでも良い。ただし、いずれを対象とするかは、統計的検定の対象となるデータ全体を通して統一しておく必要がある。

・ 必要に応じて強度の正規化を行っても良い。

## 【 0 0 7 7 】

尚、本実施例ではラットの血漿を用いたが、他の動物あるいはヒトの血漿、血清、尿、組織抽出液、髄液などの様々な生体試料へも同様に適用が可能である。また比較する群を「健常人」、「疾患 S の患者」などとした場合は、健常人と疾患患者とを区別できるグループ化範囲を見つけることが可能となる。更には、別途取得した血漿等の生体試料に対して SELDITOF - MS による測定を行って波形データを得ておき、差異のあるグループ化範囲におけるその波形データの強度を調べることで、同生体試料を提供したヒトに関する疾患の有無の推定、即ち診断を行うことが可能となる。

## 【 0 0 7 8 】

また、本実施例ではラット群間の差異の検出に統計的検定手法を用いたが、差異を検出可能な方法であれば別の方法を用いてもよい。また、本実施例では、単独のピークグループごとにラット群間の差異の検出を行ったが、複数のピークグループの組み合わせによって差異を検出することも可能である。さらに、本実施例では A 群と B 群との 2 群間の差異

10

20

30

40

50

を検出したが、Kruskal-Wallisの検定などの適切な手法を用いることで、3つ以上の群間の差異を検出することも可能である。

【0079】

尚、本実施の形態による手法は、SELDI TOF-MSに留まらず、液体クロマトグラフや質量分析など出力が波形で得られる装置全般に適用可能である。特にピークがブロードな場合、即ちピークトップが不明瞭であったり、隣接するピークが重なり合ったりするような場合に対して有効である。

【産業上の利用可能性】

【0080】

飛行時間型質量分析装置(TOF-MS:time-of-flight Mass Spectrometry)、液体クロマトグラフなどの質量分析装置を用いて血漿、血清、尿、組織抽出液、髄液などの様々な生体試料を測定することにより得られるスペクトル特性の波形データから、生体試料に含まれるタンパク質等の生体由来分子に対応する波形のピークを、電子計算機を用いた情報処理によって抽出する技術に関する。

【図面の簡単な説明】

【0081】

【図1A】本発明の一実施の形態によるプロテオーム解析装置の位置付けを示す図である。

。

【図1B】本実施の形態によるピーク抽出装置の概略構成例を示す図である。

【図2】本実施の形態による単一波形データからのピーク抽出処理(ステップ1)及びピークのグループ化処理(ステップ2)の概要を示す図である。

【図3】本実施の形態による単一波形データからピークを抽出し、包含レベルを求める課程を示す図である。

【図4】図4(a)から図4(d)までは、ピークの探索を終了する4つのケースである。

。

【図5】サブステップ(1-1)において得られた数通りのパラメータ値によるピーク領域群を示す図である。

【図6】パラメータPが $P_1$ 、 $P_2$ 、 $P_3$ 、 $P_4$ の場合について包含関係を検査した様子を示す図である。

【図7】包含関係成立条件のうち不完全な包含関係ではあるが包含していると見なすことができる例を示す図である。

【図8】各ピーク領域の包含レベル(CL)を設定する様子を示す図である。

【図9】他のいずれかのピーク領域を包含しているピーク領域、および、他のピーク領域との包含関係が存在しないピーク領域を取り除く手順を示す図である。

【図10】2つのピーク領域間で重なりがある場合にいずれか一方のみを取り除く手順を示す図である。

【図11】複数の波形データから得られたピークのグループ化処理を示す図である。

【図12】区間の分割方法方法1~4までの具体的な分割例を示す図である。

【図13】ピーク密度の計算方法2であって、当該区間を中心として含み、幅が当該区間よりS倍広い区間を参照区間と規定する方法によりピーク密度を求めるピーク密度の計算方法の概念図である。

【図14】グループ化範囲決定方法の概略を示す図である。

【図15】単一波形データからピーク抽出を行い、次いで、ピークのグループ化を行う手順の例を示す図である。

【図16】上記の一般的な処理を行ってピーク抽出を行い、次いで、ピークのグループ化を行う手順の例を示す図である。

【図17】微小区間ごとにピーク密度を計算し、ピーク密度の増減を考慮したグループ化を行う方法を示す図である。

【図18】本発明の実施例による処理の流れを示すフローチャート図である。

【図19】ベースラインの除去および平滑化の処理によって得られた波形データ(図18

10

20

30

40

50

中(a)点に相当)のサンプル10例を示す図である。

【図20】図19のサンプル1に対するピーク抽出結果を示す図である。

【図21】図20中の一部を拡大したデータを示す図である。

【図22】各波形データに対してピーク抽出および包含レベルの算出を行った後、複数波形からのピークのグループ化を行った結果を示す図である。

【図23】図22の一部を拡大したデータを示す図である。

【図24】正常なラット群(以下A群)と外科的処置を施したラット群(以下B群)とで差がみられる代表的なグループ化範囲周辺において、A群とB群とに上下にわけて波形データを示した図である。

【図25】正常なラット群(以下A群)と外科的処置を施したラット群(以下B群)とで差がみられる代表的なグループ化範囲周辺において、各グループ化範囲内での各波形の代表点を示した図である。

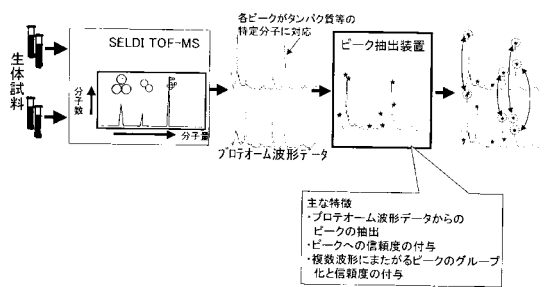
10

【符号の説明】

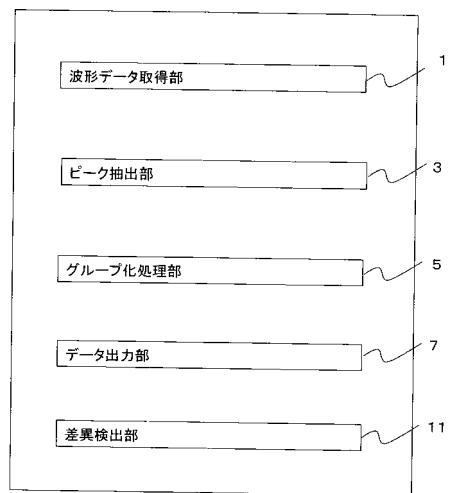
【0082】

1 ... 波形データ取得部、 3 ... ピーク抽出部、 5 ... グループ化処理部、 7 ... データ出力部、  
11 ... 差異検出部。

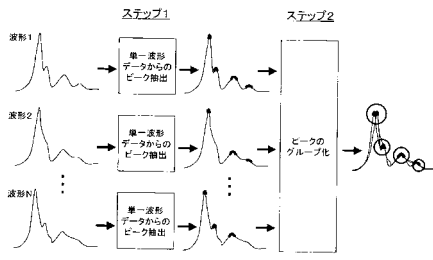
【図1A】



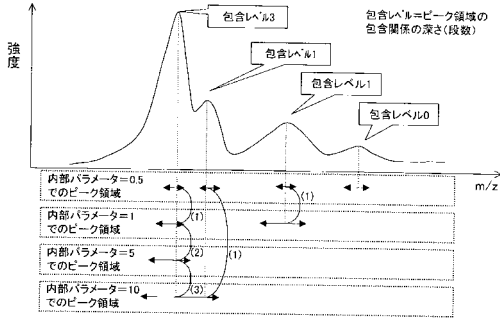
【図1B】



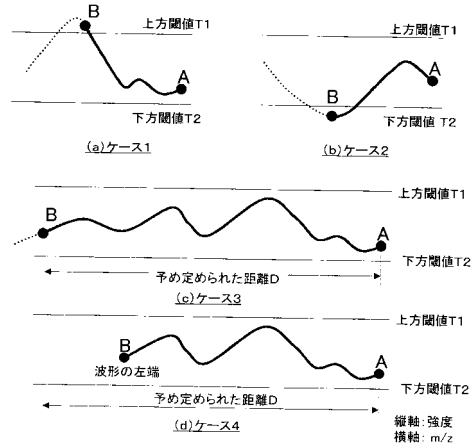
【図2】



【図3】



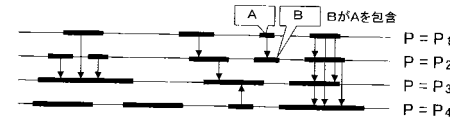
【図4】



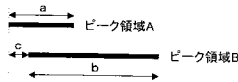
【図5】



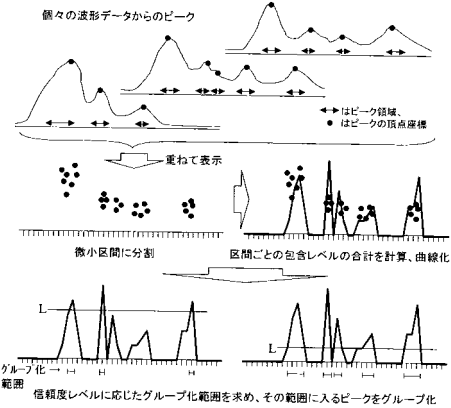
【図6】



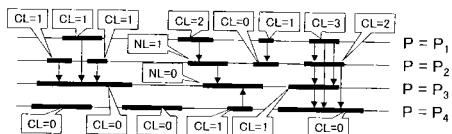
【図7】



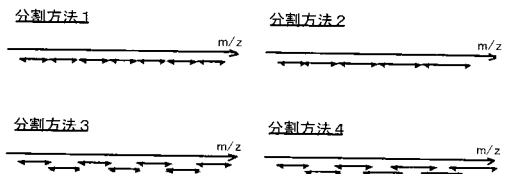
【図11】



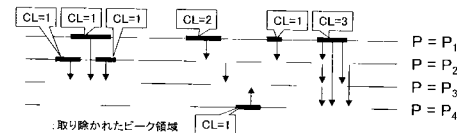
【図8】



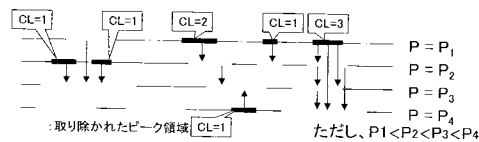
【図12】



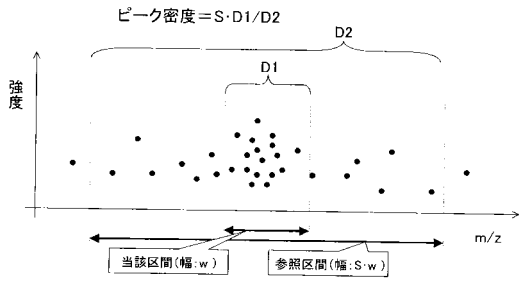
【図9】



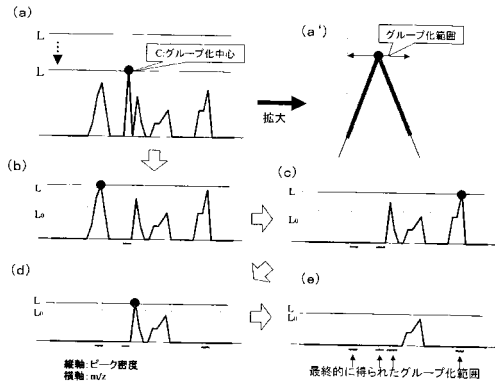
【図10】



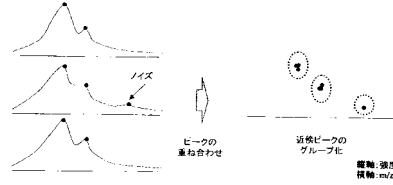
【図13】



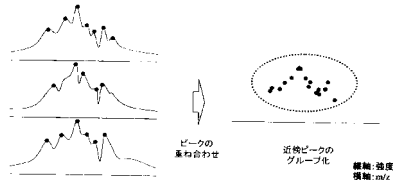
【図14】



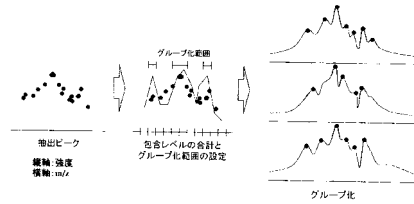
【図15】



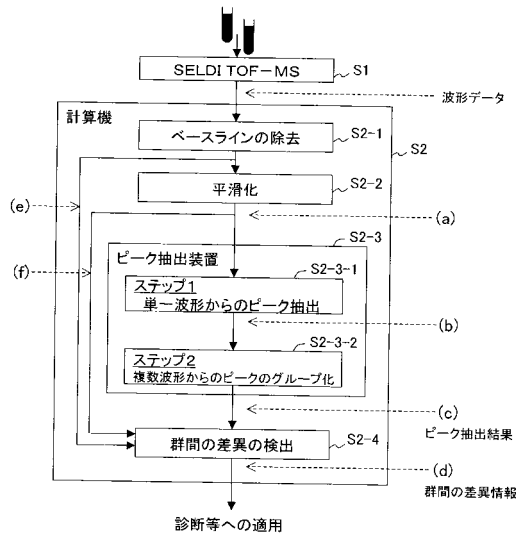
【図16】



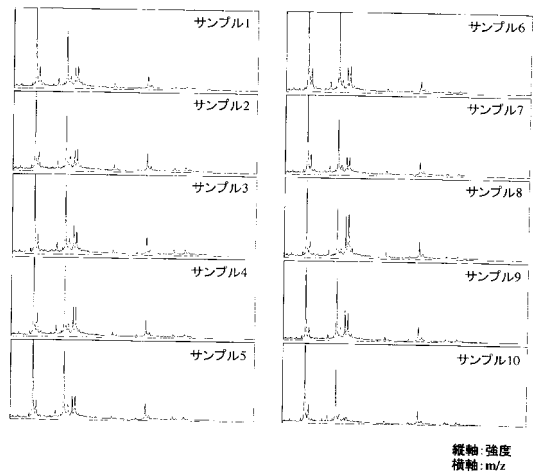
【図17】



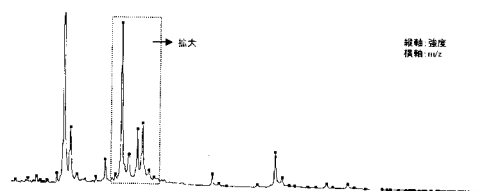
【図18】



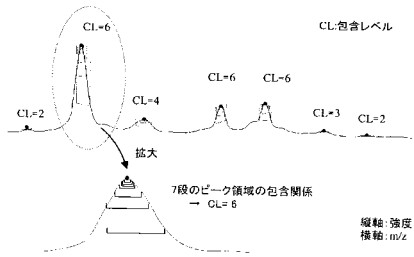
【図19】



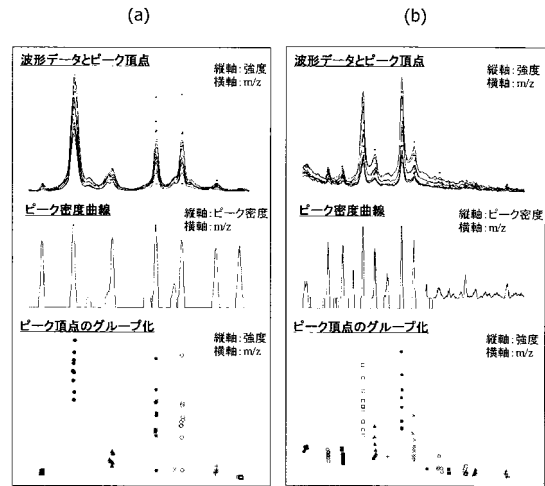
【図20】



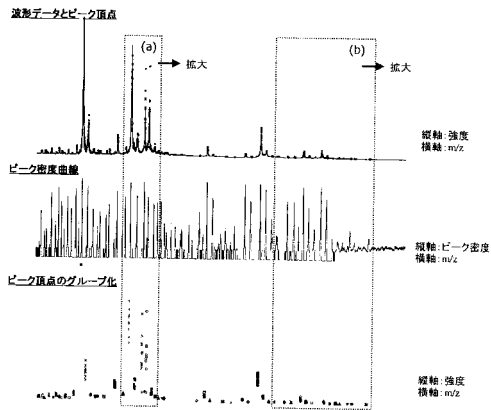
【図 2 1】



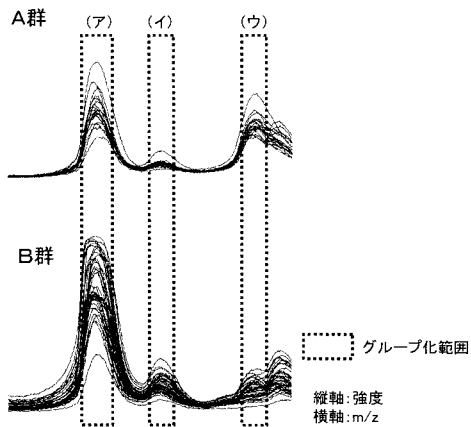
【図 2 3】



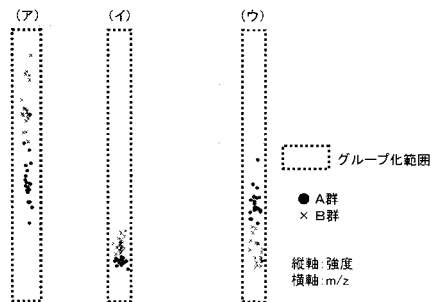
【図 2 2】



【図 2 4】



【図 2 5】



---

フロントページの続き

- (72)発明者 尾山 卓也  
富山県富山市下新町3番23号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式  
会社内
- (72)発明者 柴垣 ゆかり  
富山県富山市下新町3番23号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式  
会社内
- (72)発明者 小川 和生  
茨城県稲敷郡阿見町吉原3586 株式会社ツムラ内

審査官 高 場 正光

- (56)参考文献 特開2004-257922(JP,A)  
E.T.Fung, 他, "ProteinChip(R) Clinical Proteomics: Computational Challenges and Solutions", BioTechniques, 2002年 3月, Vol.32 Supplement, pp.34-38,40-41, [Computational Proteomics]

- (58)調査した分野(Int.Cl., DB名)  
G01N27/62-27/70  
H01J49/00-49/48  
G01N30/00-30/96  
JSTPlus(JDreamII)  
JMedPlus(JDreamII)