

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第5964983号
(P5964983)

(45) 発行日 平成28年8月3日 (2016. 8. 3)

(24) 登録日 平成28年7月8日 (2016. 7. 8)

(51) Int. Cl.

F I

GO 1 N 27/62 (2006. 01)

GO 1 N 27/64 (2006. 01)

GO 1 N 27/62 V

GO 1 N 27/62 D

GO 1 N 27/64 B

請求項の数 7 (全 10 頁)

(21) 出願番号	特願2014-544031 (P2014-544031)	(73) 特許権者	502073946
(86) (22) 出願日	平成24年11月30日 (2012. 11. 30)		バイオメリュー・インコーポレイテッド
(65) 公表番号	特表2015-500466 (P2015-500466A)		アメリカ合衆国2 7 7 1 2 ノースカロライ
(43) 公表日	平成27年1月5日 (2015. 1. 5)		ナ州ダーラム、ロドルフ・ストリート1 0
(86) 国際出願番号	PCT/IB2012/056860		〇番
(87) 国際公開番号	W02013/080170	(74) 代理人	100108453
(87) 国際公開日	平成25年6月6日 (2013. 6. 6)		弁理士 村山 靖彦
審査請求日	平成27年6月22日 (2015. 6. 22)	(74) 代理人	100064908
(31) 優先権主張番号	11306610. 4		弁理士 志賀 正武
(32) 優先日	平成23年12月2日 (2011. 12. 2)	(74) 代理人	100089037
(33) 優先権主張国	欧州特許庁 (EP)		弁理士 渡邊 隆
(31) 優先権主張番号	61/566, 025	(74) 代理人	100110364
(32) 優先日	平成23年12月2日 (2011. 12. 2)		弁理士 実広 信哉
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 質量分析法により微生物を特定するための方法

(57) 【特許請求の範囲】

【請求項 1】

質量分析法によって微生物を特定するための方法であって、
前記微生物の少なくとも1つの質量スペクトルを取得するステップと、
取得した質量スペクトルごとに、
前記質量スペクトルのピークを所定の質量範囲で検出するステップと、
質量対電荷比の範囲における各所定の再分割区間において高々1つのピークを特定する、ピークのリストを生成するステップであって、前記再分割区間の幅は、関係

【数 1】

$$L(b) = \exp\left(\frac{b - \beta}{\alpha}\right) \times \left(\exp\left(\frac{1}{\alpha}\right) - 1\right)$$
$$\alpha = \frac{b_{\min} - (b_{\max} + 1)}{\ln m_{\min} - \ln m_{\max}}$$
$$\beta = \frac{(b_{\max} + 1) \times \ln m_{\min} - b_{\min} \times \ln m_{\max}}{\ln m_{\min} - \ln m_{\max}}$$

10

に従って前記質量対電荷比に沿って増大し、前記再分割区間は、前記範囲における最小の

20

質量対電荷比に対して整数 b_{min} から前記範囲における高質量対電荷比に対して整数 b_{max} までの 1 より大きい整数で参照され、 $L(b)$ は整数 b で参照される区間の幅であり、 m_{min} は質量対電荷比の前記範囲の下界であり、 m_{max} は質量対電荷比の前記範囲の上界である、ステップと、

過去に特定した微生物および / または微生物の種類の知識ベースに従って取得したピークのリスト (複数可) を分析するステップと、

を含む、方法。

【請求項 2】

質量対電荷比の所定の範囲は 3、000 トムソンから 17、000 トムソンの範囲内にある、請求項 1 に記載の方法。

10

【請求項 3】

900 から 1500 個の区間が存在する、請求項 1 または 2 に記載の方法。

【請求項 4】

1200 から 1400 個の区間が存在する、請求項 3 に記載の方法。

【請求項 5】

前記再分割区間に保持された前記ピークは最大ピークである、請求項 1 乃至 4 の何れか 1 項に記載の方法。

【請求項 6】

前記質量分析法は MALDI - TOF 分光分析である、請求項 1 乃至 5 の何れか 1 項に記載の方法。

20

【請求項 7】

特定すべき微生物の質量スペクトルを生成可能な質量分光計と、

請求項 1 乃至 6 の何れか 1 項の方法を実施することによって、前記質量分光計により生成された前記質量スペクトルに関連付けられた前記微生物を特定可能な計算ユニットと、
を備える、質量分析法により微生物を特定するための装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は微生物、特にバクテリアを質量分析法により特定することに関する。

【背景技術】

30

【0002】

質量分析法を用いて微生物、より具体的にはバクテリアを特定することが知られている。微生物のサンプルが用意され、その後、当該サンプルの質量スペクトルを取得し前処理、即ち、スペクトル・デノイズ (ノイズ除去)、(検出器に起因する) バックグラウンド・ノイズのフィルタを行う。次に、前処理されたスペクトルの有意なピークを検出し、このようにして得られたピークのリストを、特定した微生物または 1 群の微生物の典型的なピークのリストから構築した知識ベースのデータ (種、綱、科など) と「分析」し「比較する」。

【先行技術文献】

【非特許文献】

40

【0003】

【非特許文献 1】 Jackson O. Lay ' s document, "Maldi-tof spectrometry of bacteria", Mass Spectrometry Reviews, 2001, 20, 172-194

【発明の概要】

【発明が解決しようとする課題】

【0004】

この原理は一見単純に見えるが、その実装は繊細である。実際、先ず、質量スペクトルに含まれる情報の量、特にピークの数是非常に多く、堅牢な知識ベースを生成し分類、比較、決定アルゴリズムを実装するための、非常に強力な計算ツールが必要である。

【0005】

50

次に、特にスペクトルにおけるピークの位置に関して、高い測定の不確実性がある。実際、同一の分光計での或る測定から他の測定にかけて、ならびに、或る分光計から他の分光計にかけて、所与の分子を表すピークが測定スペクトル内の固定位置を有さないこと、または、少なくとも当該ピークが或る範囲内に含まれないことが観測されることがある。したがって、取得スペクトルおよび所与のタンパク質分子に対応するピークを、分類アルゴリズムによって、当該タンパク質分子に対応するものとして特定することはできない。最後に、当該不確実性は質量対電荷比の範囲にわたって一定ではなく、この比が増大するにつれて増大する。

【課題を解決するための手段】

【0006】

10

本発明は、分析すべき情報量の減少と質量スペクトル・ピークの位置に関する精度不足の影響の減少とに起因する質量分析法による微生物の堅牢な特定を可能とする方法を提供することを目的とする。

【0007】

この目的のため、本発明の目的は、質量分析法によって微生物を特定するための方法であって、

当該微生物の少なくとも1つの質量スペクトルを取得するステップと、

取得した質量スペクトルごとに、

当該スペクトルのピークを所定の質量範囲で検出するステップと、

質量対電荷比の範囲における各所定の再分割区間において高々1つのピークを特定する、ピークのリストを生成するステップであって、当該再分割区間の幅は、関係

20

【0008】

【数1】

$$L(b) = \exp\left(\frac{b - \beta}{\alpha}\right) \times \left(\exp\left(\frac{1}{\alpha}\right) - 1\right)$$

$$\alpha = \frac{b_{\min} - (b_{\max} + 1)}{\ln m_{\min} - \ln m_{\max}}$$

30

$$\beta = \frac{(b_{\max} + 1) \times \ln m_{\min} - b_{\min} \times \ln m_{\max}}{\ln m_{\min} - \ln m_{\max}}$$

【0009】

に従って当該質量対電荷比に沿って増大し、当該再分割区間は、当該範囲における最小の質量対電荷比に対して整数 $b_{\min}$ から当該範囲における高質量対電荷比に対して整数 $b_{\max}$ までの1より大きい整数で参照され、 $L(b)$ は整数 b で参照される当該区間の幅であり、 $m_{\min}$ は質量対電荷比の当該範囲の下界であり、 $m_{\max}$ は質量対電荷比の当該範囲の上界である、ステップと、

40

過去に特定した微生物および/または微生物の種類の知識ベースに従って取得したピークのリスト(複数可)を分析するステップと、

を含む方法を提供することである。

【0010】

換言すれば、質量対電荷比の連続空間、即ちトムソン空間が、対数的に量子化され、幾つかのピークが当該区間に存在する場合に単一のピークが各量子化区間に保持される。これにより、処理すべきデータ量を大幅に減らすことができる。さらに、ピークの正確な位置が、当該ピークが属する区間への参照で置き換えられる。これにより、正確な位置を知識ベースと比較する必要がないので、ピークの位置に関する測定の不確実性が低下する。当該ピークが或る区間に属するかどうかを判定する。最後に、区間幅を対数的に増大させ

50

ることにより、機器が一定の相対精度を有するという事実、即ち、

$$p = \Delta \mu / m = \text{定数}$$

に適合させることができる。

【 0 0 1 1 】

1 実施形態によれば、所定のトムソン範囲は3、000トムソンから17、000トムソンの間である。発明者は実際に、当該範囲が大抵のバクテリアおよびイースト/モールドの特定に十分であることを発見した。特に、3、000トムソンでの局所ピークが多数の微生物に一般的でありしたがって特異的でないことが分かる。

【 0 0 1 2 】

1 実施形態によれば、900から1、500個の区間、特に、1、200から1、400個の区間が存在する。発明者は、これらの区間が、トムソン空間の量子化により導入される情報の損失と正確なピーク位置を区間で置き換えることで得られる精度との間の最適な妥協点をもたらすことを発見した。

【 0 0 1 3 】

1 実施形態によれば、再分割区間に保持されたピークは最大強度を有するピークである。しかし、他の選択も可能である。例えば、区間に存在するピークの強度の平均値または中央値を選択することが可能である。1 実施形態によれば、当該質量分析法はMALDI-TOF分光分析である。

【 0 0 1 4 】

本発明の別の目的は、質量分光分析により微生物を特定する装置であって、特定すべき微生物の質量スペクトルを生成可能な質量分光計と、上記方法を実施することによって、当該分光計により生成された当該質量スペクトルに関連付けられた当該微生物を特定可能な計算ユニットとを備える、質量分析法により微生物を特定するための装置を提供することである。

【 0 0 1 5 】

本発明は、例としてのみ与えた以下の説明を添付図面と関連して読むことでより良く理解される。

【図面の簡単な説明】

【 0 0 1 6 】

【図1】本発明に従う方法の流れ図である。

【図2】本発明に従う量子化の区間数に従って質量スペクトルから排除されたピークの数プロットを示す図である。

【発明を実施するための形態】

【 0 0 1 7 】

図1を参照して、本発明に従う、MALDI-TOF(「matrix-assisted laser desorption/ionization time of flight」の頭字語)タイプの質量分析法を用いたバクテリア特定方法を説明する。

【 0 0 1 8 】

当該方法は、ステップ10で、特定すべきバクテリアのサンプルを準備することで開始し、次いでステップ12でMALDI-TOFタイプの質量分析法により準備したサンプルの1つまたは複数の質量スペクトルを取得する。MALDI-TOF 質量分析法は当業界で公知であり、以降ではさらに詳細に説明することはしない。例えば、非特許文献1を参照されたい。

【 0 0 1 9 】

当該方法は続いて、ステップ14で、取得したスペクトルを前処理して、特に当該スペクトルを雑音除去し平滑化する。より具体的には、分光計のバックグラウンド・ノイズを表す当該スペクトルのベース・ラインを除去する。

【 0 0 2 0 】

次に、ステップ16で、取得したスペクトルに存在するピークの特定を、例えばピーク検出アルゴリズムにより、例えば極大点の検出に基づいて実施する。このようにして、ス

10

20

30

40

50

ベクトルのピークの位置と強度を含むスペクトルごとのピークのリストが生成される。

【 0 0 2 1 】

有利なことに、当該ピークは限定されたトムソン・レンジ $[m_{min}; m_{max}]$ 、好ましくはトムソン・レンジ $[m_{min}; m_{max}] = [3, 000; 17, 000]$ で特定される。実際、微生物を特定するのに十分な情報はこの質量対電荷比の範囲に含まれていることが分かっており、したがってより広い範囲を考慮する必要はないことが分かっている。

【 0 0 2 2 】

【数 2】

当該方法は続いて、ステップ 18 で量子化または「ビンニング」ステップを行う。これを
実現するために、範囲 $[m_{min}; m_{max}]$ を、関係

$$L(b) = \exp\left(\frac{b-\beta}{\alpha}\right) \times \left(\exp\left(\frac{1}{\alpha}\right) - 1\right) \quad (1)$$

$$\alpha = \frac{b_{min} - (b_{max} + 1)}{\ln m_{min} - \ln m_{max}} \quad (2)$$

$$\beta = \frac{(b_{max} + 1) \times \ln m_{min} - b_{min} \times \ln m_{max}}{\ln m_{min} - \ln m_{max}} \quad (3)$$

に従って、トムソンに沿って対数的に増大する幅を有する区間に分割する。ここで、再分割区間は、例えば 1 に等しい整数 b_{min} から整数 b_{max} までの 1 より大きい整数で参照され、 $L(b)$ は整数 b で参照される区間の幅である。整数 b_{min} は範囲 $[m_{min}; m_{max}]$ における最小の質量対電荷比の区間に対応し、整数 b_{max} は範囲 $[m_{min}; m_{max}]$ をにおける最大の質量対電荷比の区間に対応する。そうすると、トムソン軸が関係

$$b(m) = \lfloor \alpha \ln m + \beta \rfloor$$

に従って量子化される。ここで、 $\lfloor \cdot \rfloor$ は次に最小の整数値への丸めを表す。

【 0 0 2 3 】

複数のピークを含む区間ごとに、単一のピークが保持され、有利には当該ピークが最大強度を有する。このように、測定したスペクトルごとにベクトルが生成される。当該ベクトルの各成分は、量子化区間に対応し、当該区間に対して保持されたピークの強度を値として有する。値「0」はピークが当該区間で検出されなかったことを意味する。

【 0 0 2 4 】

例えば、図 1 のステップ 18 で、特定したピークの 3 つのリスト、即ち、「リスト 1」、「リスト 2」、および「リスト 3」が示されている。各々、測定した質量スペクトルに対応する。トムソン空間を、「ピン 1」から「ピン 8」へと、幅が対数的に増大する 8 個の区間に分割する。最大強度を有するピークのみが各区間に保持される。したがって、1 番目のリスト「リスト 1」の区間「ピン 6」に対して、1 つのピークが削除されている。リスト「リスト 1」、「リスト 2」、および「リスト 3」に対しては、例えば次の行列が得られる。各行はリストに対応する。

【 0 0 2 5 】

10

20

30

40

【数 3】

$$\begin{pmatrix} 980 & 0 & 98 & 0 & 1,300 & 1,556 & 400 & 2,000 \\ 505 & 700 & 200 & 0 & 500 & 200 & 345 & 256 \\ 700 & 0 & 0 & 100 & 2,340 & 1,786 & 0 & 2,507 \end{pmatrix}$$

【0026】

したがって、上述のような量子化により、質量の増大に伴うピーク位置に関する不確実性の増大が考慮されることが分かる。特に、本発明に従うトムソン軸の再分割により

10

$$p = \Delta \mu / m \quad (4)$$

のタイプの不確実性を考慮することができる。ここで、 p はピーク位置の精度であり、 $\Delta \mu$ は分光計ピーク位置の測定の不確実性であり、 m はピークの実際の位置である。したがって、当該量子化は、質量分光計の測定誤差を考慮した適合的な量子化である。

【0027】

測定したピーク位置を、当該ピークが属する区間への参照で置き換えることは、当該区間の中央にピーク位置を描えることと等価である。本発明に従う対数再分割により関係(4)に従う不確実性を低下できることを証明できる。実際、

【0028】

【数 4】

20

$$\frac{L(b)}{m_{bar}(b)} = \frac{2 \left(\exp \left\{ \frac{1}{\alpha} \right\} - 1 \right)}{\left(\exp \left\{ \frac{1}{\alpha} \right\} + 1 \right)} = ct$$

【0029】

である。ここで、 $m_{bar}(b)$ は、参照 b により参照される区間の中央である。

【0030】

ピークの強度は、或るスペクトルから別のスペクトルおよび／または或る分光計から別の分光計へと非常に可変的である。この可変性のため、生の強度値を考慮することは非常に困難である。有利なことに、任意ではあるが、当該方法は続いて強度離散化ステップを行う。当該ステップが例えば、単純な「二値化」(存在／不存在)を含んでもよい。

30

【0031】

したがって、行列の各行が「二値化」されて正規化され、当該行列は取得されたスペクトルごとに区間内のピークの存在または不存在を特定する。例えば、上述の行列は行列

【0032】

【数 5】

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

40

【0033】

に二値化される。

【0034】

バクテリアの特定に関連する情報がピークの不存在および／または存在に原則として含まれること、および、強度情報はその高い可変性のためあまり関係がないことに留意されたい。したがって、例えば、ロジスティック回帰、判別分析、分類木、LASSO法、SVM(「Support Vector Machine」の頭字語)タイプのアルゴリ

50

ズムのような通常の分類ツールによりこの種のリストに基づいてバクテリアを特定することができる。したがって、このように二値化した行列を全ての既知の分類ツールにおいて使用してもよい。

【 0 0 3 5 】

当該方法は続いてステップ 20 で、以前のステップで取得した行列を分析する。より具体的には、分類決定アルゴリズム 22 が、過去に特定した微生物のピークおよび / または微生物のタイプのリストに従って構築された知識ベース 24 に従って実装される。1 つもしくは複数の候補または微生物のタイプ (科、菌、種、亜種) を、分析サンプル用に特定する。

【 0 0 3 6 】

このように、本発明に従う方法により、可変サイズかつ 2 軸 (m/z 、強度) に沿った連続値のピークのリストを合理的な固定サイズのベクトルに削減することができる。

【 0 0 3 7 】

上述のように生成した、過去に特定した微生物および / または微生物のタイプに関連するピークのリストから知識ベース 24 が構築される。本発明は、任意の種類の分類アルゴリズムと知識ベースに適用されることが分かる。本発明に従う量子化により特に、データ量を減らすことができ、ピーク位置精度の問題を排除することができ、したがって、より堅牢な知識ベースを単純な方法で構築することができる。実装は (例えば許容距離) の計算よりかなり単純であり、知識ベースをほぼ全自動で構築することができる。

【 0 0 3 8 】

有利なことに、微生物を特定するために、区間数が 900 から 1、500 の間、好ましくは 1、200 から 1、400 の間で選択される。発明者は、これらの区間が、トムソン空間の量子化により導入された情報の損失と正確なピーク位置を区間で置き換えることで得られる精度との間の最適な妥協点をもたらすことを発見した。発明者は、図 2 に示すように、試験を行い、区間数に従って量子化により排除されたピークの数モデル化した。幾つかの区間を超えて、データ量の減少は無視でき、所定数のもとでは、排除されるピークの数指数関数的に増大することに特に留意されたい。

【 0 0 3 9 】

本発明の対数量子化と定数量子化について比較試験を実行した。定数量子化では、全ての区間で幅が同一であり、質量分光計、分類決定アルゴリズム、および知識ベース構築が同一である。これらの試験を以下の表に示す。誤差は微生物特定誤差に対応する。

【 0 0 4 0 】

10

20

30

【表 1】

量子化	区間数	誤差 (%)	占有メモリ空間 (Mo)
対数	300	10.33	120
対数	600	6.25	240
対数	800	5.3	320
対数	1,000	5.4	400
対数	1,300	5.0	520
対数	1,700	6.9	680
対数	2,300	8.52	920
対数	4,700	12.2	1,880
定数	300	12.4	120
定数	600	8.75	240
定数	800	7.2	320
定数	1,000	6.6	400
定数	1,300	6.2	520
定数	1,700	5.9	680
定数	2,300	7.22	920
定数	4,700	11.0	1,880

【 0 0 4 1 】

一定幅の 1、000 個の区間を選択することによって、区間の幅は、17、000 トムソンに等しい質量対電荷比に対する試験に使用される質量分光計の解像度に等しくなる。一定幅の 4、700 個の区間を選択することによって、区間の幅は、3、000 トムソンに等しい質量対電荷比に対する質量分光計の解像度に等しくなる。

【 0 0 4 2 】

$b_{min} = 1$ として関係 (1) 乃至 (3) に従って 1、700 個の対数区間を選択することによって、各区間の幅は、当該区間の中央に等しい質量対電荷比に対する分光計の精度に等しくなる。しかし、平均して、1、300 個の区間が最低誤差の特定率と最小の占有メモリ空間をもたらすことが分かる。特に、上記表に示されているように、1、700 個の区間と比較すると、これは一見して良好に適合されているように見えるが、2 の誤差ポイント (- 28 % の誤差) が得られる一方、占有メモリ空間は減る。したがって、1、300 が本発明を実装するのに好適である。

【 0 0 4 3 】

また、本発明に従う量子化は、定数量子化と比べて少なくとも 1 ポイント (- 15 % の誤差) だけ少ない最大誤差率と、少ないメモリ空間 (- 25 %) をもたらすことに留意されたい。少数の区間に対して、本発明に従う量子化は定数量子化よりも良好な結果をもたらす。したがって、これにより、質量分光計の解像度または保持されるトムソン範囲 $[m_{min}; m_{max}]$ が増大しても少数の区間を保つことができる。したがって、同じ誤差率、例えば、ほぼ 6 % に対して、本発明に従う量子化では 700 個の区間しか必要でないが定数量子化では 1、700 個の区間が必要であることが分かる。

【符号の説明】

【 0 0 4 4 】

10

20

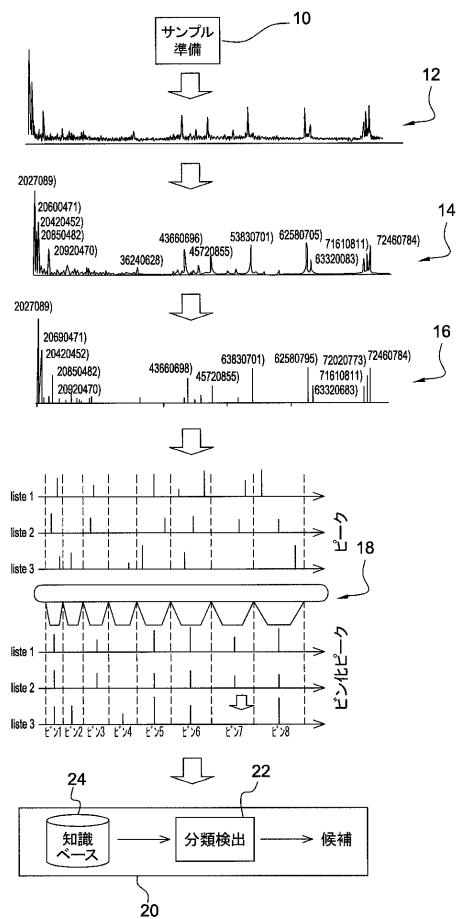
30

40

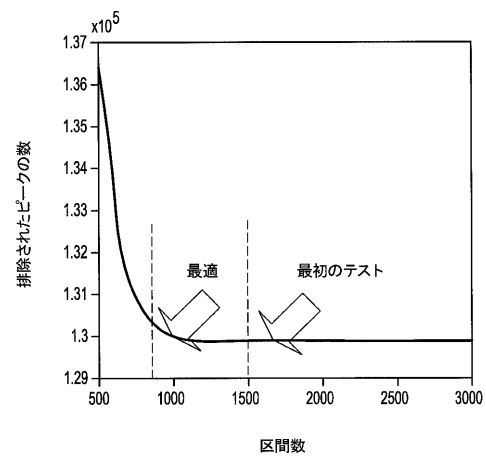
50

- 1 0 サンプル準備
- 2 2 分類検出
- 2 4 知識ベース

【図 1】



【図 2】



フロントページの続き

- (72)発明者 グレゴリー・ストリューベル
フランス・04100・マノスク・モンテ・ドゥ・サント・ルスタニュ・349
- (72)発明者 モード・アルサック
フランス・42400・サン・シャモン・リュ・ドゥ・ランゴナン・13
- (72)発明者 ドゥニ・デスレ
フランス・01120・モンリュエル・リュ・デ・エコルシャ・359
- (72)発明者 ピエール・ジャン・コト・パタ
フランス・01150・ラニュー・リュ・デュ・クルランドン・105

審査官 藤田 都志行

- (56)参考文献 特表2006-522340(JP,A)
特表2010-515915(JP,A)
米国特許出願公開第2011/0202282(US,A1)
特開2007-316063(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G01N 27/62
G01N 27/64