

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2005年3月10日 (10.03.2005)

PCT

(10) 国際公開番号
WO 2005/022166 A1

- (51) 国際特許分類: G01N 33/68
- (21) 国際出願番号: PCT/JP2004/008827
- (22) 国際出願日: 2004年6月17日 (17.06.2004)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2003-172898 2003年6月18日 (18.06.2003) JP
- (71) 出願人 (米国を除く全ての指定国について): 独立行政法人理化学研究所 (RIKEN) [JP/JP]; 〒3510198 埼玉県和光市広沢 2 番 1 号 Saitama (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 川端 邦明 (KAWA-BATA, Kuniaki) [JP/JP]; 〒3510198 埼玉県和光市広沢

2 番 1 号 独立行政法人理化学研究所内 Saitama (JP). 浅間 一 (ASAMA, Hajime) [JP/JP]; 〒3510198 埼玉県和光市広沢 2 番 1 号 独立行政法人理化学研究所内 Saitama (JP). 齊藤 佳奈子 (SAITO, Kanako) [JP/JP]; 〒3510198 埼玉県和光市広沢 2 番 1 号 独立行政法人理化学研究所内 Saitama (JP). 國光 智 (KUNIMITSU, Satoshi) [JP/JP]; 〒3510198 埼玉県和光市広沢 2 番 1 号 独立行政法人理化学研究所内 Saitama (JP). 三島 健稔 (MISHIMA, Taketoshi) [JP/JP]; 〒3510198 埼玉県和光市広沢 2 番 1 号 独立行政法人理化学研究所内 Saitama (JP).

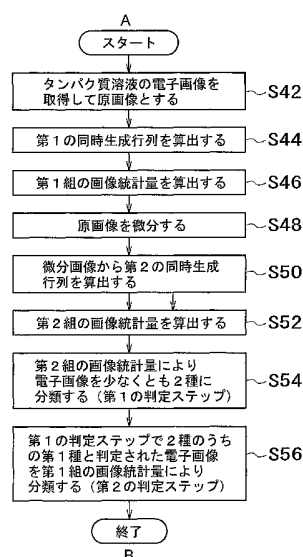
(74) 代理人: 奥山 尚一, 外(OKUYAMA, Shoichi et al.); 〒1070052 東京都港区赤坂 3 丁目 2 番 1 2 号 赤坂ノアビル 8 階 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,

[続葉有]

(54) Title: PROTEIN CRYSTALLINE STATE JUDGING METHOD AND ITS SYSTEM

(54) 発明の名称: タンパク質結晶化状態判定方法およびそのシステム



A... START
 S42... CAPTURE ELECTRONIC IMAGE OF PROTEIN SOLUTION AND USE IT AS ORIGINAL IMAGE
 S44... CALCULATE FIRST SIMULTANEOUSLY GENERATED MATRIX
 S46... CALCULATE IMAGE STATISTICAL AMOUNT OF FIRST SET
 S48... DIFFERENTIATE ORIGINAL IMAGE
 S50... CALCULATE SECOND SIMULTANEOUSLY GENERATED MATRIX FROM DIFFERENTIATED IMAGE
 S52... CALCULATE IMAGE STATISTICAL AMOUNT OF SECOND SET
 S54... SORT ELECTRONIC IMAGES INTO AT LEAST TWO TYPES ACCORDING TO IMAGE STATISTICAL AMOUNT OF SECOND SET (FIRST JUDGING STEP)
 S56... SORT ELECTRONIC IMAGES JUDGED AS TYPE 1 OUT OF TWO TYPES AT FIRST JUDGING STEP ACCORDING TO IMAGE STATISTICAL AMOUNT OF FIRST SET (SECOND JUDGING STEP)
 B... END

(57) Abstract: The crystalline state of a protein solution is judged. A method for judging the crystalline state of a protein solution through analysis using an observed electronic image about the protein solution is disclosed. The method makes a judgment on a sample of a protein solution by applying an original image capturing step S42, an original image matrix calculating step S44, an original image statistical amount calculating step S46, an image differentiating step S48, a differentiated image matrix calculating step S50, a differentiated image statistical amount calculating step S52, a first judging step S54, and a second judging step S56.

(57) 要約: 本発明は、タンパク質溶液の結晶化状態を判定することを課題とする。タンパク質溶液についての観察された電子画像を用いて解析するタンパク質溶液の結晶化状態の判定方法。タンパク質溶液のサンプルについて、原画像取得ステップS42と、原画像行列算出ステップS44と、原画像統計量算出ステップS46と、画像微分ステップS48と、微分画像行列算出ステップS50と、微分画像統計量算出ステップS52と、第1の判定ステップS54と、第2の判定ステップS56を適用して判定する。



BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,

IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告書
- 請求の範囲の補正の期限前の公開であり、補正書受領の際には再公開される。

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明 細 書

タンパク質結晶化状態判定方法およびそのシステム

技術分野

本発明は、タンパク質溶液におけるタンパク質結晶生成を自動化するためのタンパク質結晶状態を判定する方法およびそのシステムに関する。

背景技術

従来より、タンパク質の結晶構造を解析するためにX線結晶構造解析法が用いられている。かかる構造解析法においては、構造解析対象のサンプルとなるタンパク質の良質な結晶を得るために、構造解析サンプル作成の準備として、タンパク質の溶液（水溶液）から水分を蒸発させて結晶を生成する方法（結晶化法）が用いられる。この際、構造や性質が未知であるタンパク質溶液から結晶が得られるかどうかを選別するスクリーニング作業が行なわれている。

従来のこのスクリーニング作業では、結晶状態を顕微鏡画像から人の目で逐次検討し、そのサンプルの結晶生成の様子を判定している。この判定においては、サンプルのタンパク質溶液が、溶液のままであるか、沈殿を起こしているか、微結晶が得られているか、あるいは、結晶が得られるかが判定される。表1に、このスクリーニングにおいて用いられる判定基準を示す。

（表1）

スコア	状態名	代表的な観察結果	計算の分類
0	透明	透明	A
1	沈殿 (i)	粒が無く、一様に黒ずんでいる	B
2	沈殿 (ii)	少し点々がある	C
3	沈殿 (iii)	アモルファス	D
4	沈殿 (iv)	大きくなったアモルファス	E
5	微結晶	50 μ m以下の結晶 (角がある)	
6	結晶 (i)	針状	
7	結晶 (ii)	板状	
8	結晶 (iii)	結晶が重なり合っているもの	
9	結晶 (iv)	良質な結晶	

F : いずれにも属さない

表1において、「透明」とは、タンパク質溶液が何らの結晶も生成しない様子を表わしている。また、「沈殿」は、4つに分類される。即ち、粒が見られず黒ずんできたり褐色を示すもの（沈殿 (i)）、点状の組織が観察されるもので白色にみえ、少し点々があるもの（沈殿 (ii)）、アモルファス様の組織で透明に見えることがあるもの（沈殿 (iii)）、大きなアモルファス様の組織で粒子が見えることがあるもの（沈殿 (iv)）に分類される。また、結晶が生成されているものは、5つに分類される。即ち、50 μ m程度以下の結晶で頂点が観察されるもの（微結晶）、針状結晶が観察されるもの（結晶 (i)）、板状結晶が観察されるもの（結晶 (ii)）、重なり合った結晶が観察されるもの（結晶 (iii)）、良質な結晶が観察されるもの（結晶 (iv)）というように分類される。便宜的に、各分類に対してスコアとなる数字が割り当てられて分類されることもある。ただし、タンパク質の結晶の成長過程に従ってこのスコアが順次増してゆくように変化するものとは限らない。このように分類される画像の模式図を、スコアごとに図7に示す。特に、未知タンパク質溶液の結晶化実験においてこれらに分類されるサンプルの数が多数であるため、大変有用な手段となる。

従来の人の観察を基礎とするスクリーニング作業では、観察者はトレーニングによって養成されて、経験的にこのような分類を行なうことができるようになる。養成された観察者が多数のサンプルについて得た顕微鏡画像を観察して選別を行なう。このときに判定された分類は、それ自体も、X線構造解析向けのサンプルについての有用な情報であり、タンパク質の構造解析や性質の決定に際して重要な指標となっている。

この結晶を得るための手法においては、大別すると、タンパク質溶液をカバーガラスの下面に下垂させて蒸発を行なう手法（ハンギングドロップ蒸気拡散法）と、タンパク質溶液を蒸発容器の上に載置させて蒸発を行なう手法（シッティングドロップ蒸気拡散法）とが用いられる。多数のタンパク質溶液の顕微鏡画像を得るには、自動的に顕微鏡画像を記録することが有効であり、機械化の容易なシッティングドロップ蒸気拡散法が望ましい。

特に、表1に示した透明、沈殿（i）、沈殿（i i）、沈殿（i i i）に未知タンパク質溶液のサンプルが分類されることが多いが、これ等に分類されるサンプルは、いまだ結晶が生成されておらず、観察者がこれらを全て観察するのは、効率が悪い。

また、従来のテクスチャ解析手法の例が、R. M. Haralick, K. Shanmugam および I. Dinstein による "Texture features for image classification" (IEEE Trans. Syst., Man, Cybern., vol. SMC-3, no6, pp.610-621, 1973) に記載されている。

発明の開示

上記従来のタンパク質溶液の結晶化状態判定方法では、人の目による分類が欠かせず、スクリーニング作業で処理できるサンプルの数に限界がある。また、その分類基準についても、観察者ごとに判定が異なる場合があり、さらに、同じ観察者であっても繰り返し同じ判定が行なえない場合がある。

- 5 多数のサンプルの判定を行なうためには、シッティングドロップ蒸気拡散法による機械化が望ましいが、結晶化状態の判定は、ハンギングドロップ蒸気拡散法の顕微鏡画像に比べて判定が難しい傾向がある。

本発明は、かかる従来のタンパク質溶液の結晶化状態判定方法を、顕微鏡画像に対してテクスチャ解析の手法を用いて機械化することを課題とする。また、機械化されて自動撮影される顕微鏡画像においても良好に判定が行なえる判定方法
10 を提供し、撮影から判定までを機械化することができるような効率の高い安定したタンパク質の結晶化状態の判定を可能にする方法やシステムを提供することを課題とする。さらに、表1に示した透明、沈殿 (i)、沈殿 (i i)、沈殿 (i i i) を適切に分類する方法を確立することも課題とする。

- 15 本発明では、タンパク質溶液の電子画像を取得して原画像とする原画像取得ステップと、該原画像の少なくとも一部分について第1の同時生成行列を算出する原画像行列算出ステップと、該第1の同時生成行列から第1組の画像統計量を算出する原画像統計量算出ステップと、該原画像に微分処理を行なって微分画像を得る画像微分ステップと、該微分画像の少なくとも一部分について第2の同時生成
20 成行列を算出する微分画像行列算出ステップと、該第2の同時生成行列から第2組の画像統計量を算出する微分画像統計量算出ステップと、該第2組の画像統計量を用いて、タンパク質溶液の電子画像を少なくとも2種に分類する第1の判定

ステップと、前記第1の判定ステップで分類された結果、前記2種のうちの第1種に属すると判定されたタンパク質溶液の電子画像をさらに前記第1組の画像統計量を用いて少なくとも2種に分類する第2の判定ステップとを含む、タンパク質溶液の結晶化状態の判定方法が提供される。

- 5 画像の解析は、一般にテクスチャ解析と呼ばれる手法を用いることができる。この解析対象に、タンパク質溶液の原画像（電子的に取得して明るさや領域の切り出し等の予備的な調整をしたのみの画像で、微分していない画像）と、その原画像を微分した微分画像を共に用いることにより、未知タンパク質の溶液における結晶化の状態を良好に判定できる。

- 10 また、判定を2段階に行なうことにより、タンパク質溶液の結晶化状態を判定する作業の機械化が容易となる。本方法によれば、表1に示した透明と沈殿（i）とを他のものから適切に分類することが可能となる。

- また、本発明においては、前記第1の判定ステップで分類された結果、前記第1種以外の種に分類されたタンパク質溶液の電子画像を、前記第2組の画像統計
15 量を用いてさらに少なくとも2種に分類する第3の判定ステップをさらに含む方法が提供される。

第3の判定ステップにより、透明と沈殿（i）との分類に加えて、沈殿（i i）と沈殿（i i i）とを他のものから分類することができる。

- さらに、本発明においては、上記方法において、予め判定されているタンパク
20 質の結晶状態の電子画像を用いて、前記第1の判定ステップにおいて用いる第1の判定基準と第2の判定基準とを、前記第2組の画像統計量に基づく線形判別法により確立するステップと、前記予め判定されているタンパク質の結晶状態の電

子画像を用いて、前記第2の判定ステップにおいて用いる第3の判定基準を、前記第1組の画像統計量に基づく線形判別法により確立するステップと、前記予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第3の判定ステップにおいて用いる第4の判定基準を、前記第2組の画像統計量に基づく線形判別法により確立するステップとをさらに含む方法が提供される。ここで、前記第1の判定ステップは、前記第1の判定基準と前記第2の判定基準とにより、前記第2組の画像統計量を用いて電子画像を3種に分類するものであり、前記第2の判定ステップは、前記第3の判定基準により、該第1組の画像統計量を用いて前記3種のうちの第1種に属する電子画像をさらに2種に分類するものであり、前記第3の判定ステップは、前記第4の判定基準により、該第2組の画像統計量を用いて前記第1種以外のいずれかの種に属する電子画像をさらに2種に分類するものである。

判定ステップを適切に選択された4つの基準を適切な順序で用いるものとして構築すれば、判定の際のエラーが低減され、人による判定を十分に再現できる良好な判定が人による分類を経ずとも行なえる。

本発明においては、上記方法において、透明状態、第1～第4の沈殿状態、微結晶状態、および結晶状態を要素として構成される複数の状態のいずれかにタンパク質溶液の電子画像を分類する方法が提供される。ここで、前記第1の判定基準は、タンパク質溶液が、透明状態と第1～第3の沈殿状態とからなる状態のいずれか、あるいは、第4の沈殿状態と微結晶状態と結晶状態とからなる状態のいずれかのどちらにあるかを判定するものであり、前記第2の判定基準は、タンパク質溶液が、透明状態と第1の沈殿状態とからなる状態のいずれか、あるいは、

第2～第4の沈殿状態と微結晶状態と結晶状態とからなる状態のいずれかのどちらにあるかを判定するものであり、前記第3の判定基準は、タンパク質溶液が、透明状態、あるいは、第1の沈殿状態のどちらにあるかを判定するものであり、前記第4の判定基準は、タンパク質溶液が、第2の沈殿状態、あるいは、第3の沈殿状態のどちらにあるかを判定するものであり、前記第1の判定ステップは、前記第1の判定基準と前記第2の判定基準とにより、タンパク質溶液が、透明状態と第1の沈殿状態からなる状態のいずれか、第2～第3の沈殿状態からなる状態のいずれか、あるいは、第4の沈殿状態と微結晶状態と結晶状態とからなる状態のいずれかのいずれにあるかを判定するステップであり、前記第2の判定ステップは、前記第3の判定基準により、透明状態と第1の沈殿状態からなる状態のいずれかにあるタンパク質溶液が、透明状態、あるいは、第1の沈殿状態のどちらにあるかを判定するステップであり、前記第3の判定ステップは、前記第4の判定基準により、第2～第3の沈殿状態からなる状態のいずれかにあるタンパク質溶液が、第2の沈殿状態、あるいは、第3の沈殿状態のどちらにあるかを判定するステップである。

テクスチャ解析においては、判定基準は一般には様々な画像に基づいて確立することができるが、本発明においては、原画像と微分画像を用い、それぞれの基準が判定する特徴に適した画像の種類をこのように選択することが出来る。これにより、判定のエラーが低減され、人による判定を十分に再現可能な良好な判定の機械化が可能となる。

本発明では、さらに、タンパク質溶液について観察された画像を解析することによって該タンパク質溶液における結晶化状態を判定する結晶化状態判定シス

- テムが提供される。このシステムは、タンパク質溶液の電子画像を取得して原画像とする原画像取得手段と、該原画像の少なくとも一部分について第1の同時生成行列を算出する原画像行列算出手段と、該第1の同時生成行列から第1組の画像統計量を算出する原画像統計量算出手段と、該原画像に微分処理を行ない、微分画像を得る画像微分手段と、該微分画像の少なくとも一部分について第2の同時生成行列を算出する微分画像行列算出手段と、該第2の同時生成行列から第2組の画像統計量を算出する微分画像統計量算出手段と、該第2組の画像統計量を用いて、タンパク質溶液の電子画像を少なくとも2種に分類する第1の判定手段と、前記第1の判定手段で分類された結果、前記2種のうちの第1種に属すると判定されたタンパク質溶液の電子画像をさらに前記第1組の画像統計量を用いて少なくとも2種に分類する第2の判定手段とを含んで構成される。

本発明の解析システムは、タンパク質溶液の結晶化を判定するサンプルについて電子画像を用いて結晶化状態の判定を行なうことが出来る。これにより、従来は観察者に頼っていた結晶化状態の判定作業の機械化が可能となる。

- さらに本発明の解析システムに、タンパク質の溶液の結晶化サンプルの自動観察手段を組み合わせれば、未知タンパク質の観察とその未知タンパク質の結晶化状態の判定とを連続して自動で行なうタンパク質結晶化状態観察解析システムが構築できる。

図面の簡単な説明

図1は、本発明の実施の形態におけるサンプルの構成を示すサンプル容器の断面図である。

図 2 は、本発明の実施の形態における結晶化状態判定システムの構成サンプルの構成を示す構成図である。

図 3 は、本発明の実施の形態におけるテクスチャ解析の手法を説明する説明図である。

図 4 は、本発明の実施の形態におけるテクスチャ解析における微分方法を説明する説明図である。

図 5 は、本発明の実施の形態における判定基準の確立の方法を示すフローチャートである。

図 6 は、本発明の実施の形態における判定方法を示すフローチャートである。

図 7 は、タンパク質溶液の結晶状態の画像による分類を模式的に示す説明図である。

図 8 は、本発明により提供されるタンパク質溶液の結晶状態の判定方法を表すフローチャートである。

発明を実施するための最良の形態

本発明は、コンピュータ装置を用いて実施される。このコンピュータには、適当な画像入力手段が接続されている。その画像入力手段は、例えば顕微鏡に備えられた CCD（電荷結合デバイス）などの撮像装置、A/Dコンバータ、適当な
5 メモリなどによって電子的な画像を取得し、コンピュータに入力する。また、コンピュータには、演算装置、記憶装置、表示装置、通信装置などの通常の装置が備えられる。本発明の各ステップや各手段は、演算装置に実現された機能ブロックとして処理が行なわれ、機能が実現される。例えば、判定ステップや判定手段

は、演算装置において、記憶装置から読み込まれたプログラムが必要なレジスタやメモリによって構成される記憶動作と演算処理動作によって判定を行なう機能として実現される。本発明は、このようなコンピュータを用いて、画像処理および画像の分類を用いて、タンパク質溶液の結晶化状態を判定するための方法や

5 手段を提供するものである。

以下、図面を参照して、本発明の実施の形態について説明する。

図1は、本発明において観察対象の溶液においてタンパク質の結晶化（析出）をさせるために製作されるサンプルの準備された様子を表わす断面図である。本発明の観察対象は、タンパク質溶液1を観察する光学顕微鏡の観察像である。タンパク質溶液1は、水分を含んだ溶液である。このタンパク質溶液1は、容器3

10 内に配置された滴下ステージ5に滴下して準備される。容器3の開口部は、カバースリップ2によって覆われ、容器3とカバースリップ2の接触部はシールされて、以後気密が保たれる。容器3には、予め沈殿材4が配置されており、容器内の水分を吸収してゆく。

15 このようなサンプルでは、数日から数年の時間をかけて、溶液中の水分が減少してゆき、タンパク質の飽和溶解度に達すると、徐々にタンパク質の結晶が成長してゆくことがある。タンパク質の種類によっては、全く結晶を生成しない場合や、沈殿を生じる場合もある。

図2は、本実施形態の結晶化状態判定システム100の構成を説明する構成図

20 である。本システムには、原画像取得手段12、原画像行列算出手段14、原画像統計量算出手段20、画像微分手段22、微分画像行列算出手段24、微分画像統計量算出手段26、判定基準設定手段28、判定基準格納手段32、判定手

段 3 4 が備えられている。なお、判定手段 3 4 には、図示しないが、第 1 ～第 3 の判定手段が含まれているものとしてもよい。また、本システムにおいては、判定基準を確立するために、予め結晶化状態が分類されたタンパク質の画像について、外部にある撮像装置 1 6 からの電子画像のデータ 1 8 を取得し、結晶状態データ 3 0 を取得するものであってもよい。本システムは、未知タンパク質溶液の画像について、外部にある撮像装置 1 6 からの電子画像のデータ 1 8 を取得し、その未知タンパク質溶液の結晶化状態を判定して、その結晶状態のデータ 3 6 を出力する。

原画像取得手段 1 2 は、例えば顕微鏡に備えられた撮像装置 1 6 によって撮影され、デジタル化された電子画像のデータ 1 8 を取得して、その電子画像を調整して原画像のデータ 1 2 2 とする。調整は、デジタル化されたデータの明度等を適切な範囲に調整したり、計算に十分な範囲で階調の削減を行ったり、必要な領域を切り出す処理などの一般的な処理である。

原画像行列算出手段 1 4 は、その調整された原画像のデータ 1 2 2 の少なくとも一部分の各画素を走査して、該原画像のデータ 1 2 2 に基づいて第 1 の同時生成行列 1 4 2 を算出する。

ここで、同時生成行列とは、画像の濃度（画素値） i の点から一定の変位 $\delta = (r, \theta)$ （図 3 a 参照）だけ離れた点の濃度が j となる確率を要素として持つような行列である。例えば、図 3 b のような画像については、 (r, θ) 毎に次式のような同時生成行列が計算される。

$$\begin{array}{cc}
 \begin{array}{c} \rightarrow j \\ 0 \quad 1 \quad 2 \quad 3 \\ P_{\theta} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \downarrow i \end{array} \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} & \begin{array}{c} \rightarrow j \\ 0 \quad 1 \quad 2 \quad 3 \\ P_{\theta} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \downarrow i \end{array} \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \\
 (b) \ (r, \theta) = (1, 0^\circ) & (c) \ (r, \theta) = (1, 90^\circ)
 \end{array}$$

$$\begin{array}{cc}
 \begin{array}{c} \rightarrow j \\ 0 \quad 1 \quad 2 \quad 3 \\ P_{\theta} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \downarrow i \end{array} \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} & \begin{array}{c} \rightarrow j \\ 0 \quad 1 \quad 2 \quad 3 \\ P_{\theta} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \downarrow i \end{array} \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\
 (d) \ (r, \theta) = (1, 135^\circ) & (e) \ (r, \theta) = (1, 45^\circ)
 \end{array}$$

ここでは、処理対象の範囲の画素が走査されて、同時生成行列の行列要素に加算されている。

- この各行列に現われる要素は、その (r, θ) と (i, j) の組合わせごとの
- 5 画素の出現頻度を表わしている。この画素の出現頻度を確率に換算するために正規化処理を行なう。この確率を要素とする同時生成行列の行列要素を以後 $E(i, j)$ とする。なお、 (r, θ) 毎にこの要素が計算されているが、記載を省略する。原画像行列算出手段 14 はこのような確率を求める機能を有しており、適
- 10 当なメモリーを用いて頻度を計算し、その頻度を正規化して確率を要素とする同時生成行列を出力する機能を有している。

このようにして定められる同時生成行列 $E(i, j)$ に基づいて、原画像統計量算出手段 20 は、第 1 組の画像統計量 202 を算出する。この画像統計量の組は、次式で表わされる。

$$\text{角 2 次モーメント : } \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{E(i, j)\}^2 \quad \dots (1)$$

$$\text{コントラスト : } \sum_{k=0}^{N-1} k^2 E_{x-y}(k) \quad \dots (2)$$

$$\text{相関 : } \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} ij E(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad \dots (3)$$

$$\text{分散二乗和 : } \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu_x)^2 E(i, j) \quad \dots (4)$$

$$\text{逆差分モーメント : } \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i - j)^2} E(i, j) \quad \dots (5)$$

$$\text{平均和 : } \sum_{k=0}^{2N-2} k E_{x+y}(k) \quad \dots (6)$$

$$\text{分散和 : } \sum_{k=0}^{2N-2} (k - \sum_{l=0}^{2N-2} l E_{x+y}(l))^2 E_{x+y}(k) \quad \dots (7)$$

$$\text{エントロピー和 : } - \sum_{k=0}^{2N-2} E_{x+y}(k) \log \{E_{x+y}(k)\} \quad \dots (8)$$

$$\text{エントロピー : } - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E(i, j) \log \{E(i, j)\} \quad \dots (9)$$

$$\text{分散差 : } \sum_{k=0}^{N-1} \left\{ k - \sum_{k=0}^{N-1} k E_{x-y}(k) \right\}^2 E_{x-y}(k) \quad \dots (10)$$

$$\text{エントロピー差 : } - \sum_{k=0}^{N-1} E_{x-y}(k) \log \{E_{x-y}(k)\} \quad \dots (11)$$

$$\text{相関計測情報 1 : } \frac{HXY - HXY1}{\max\{HX, HY\}} \quad \dots (12)$$

$$\text{相関計測情報 2 : } [1 - \exp\{-2.0(HXY2 - HXY)\}]^{1/2} \quad \dots (13)$$

$$\text{最大相関係数 : } (Q \text{ の 2 番目に大きい固有値 })^{1/2} \quad \dots (14)$$

ここで、この各画像統計量の計算には、次の関係式が用いられている。

$$\begin{aligned}
E_x(i) &= \sum_{j=0}^{N-1} E(i, j), \quad i = 0, 1, \dots, N-1 \\
E_y(j) &= \sum_{i=0}^{N-1} E(i, j), \quad j = 0, 1, \dots, N-1 \\
E_{x+y}(k) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E(i, j), \quad i+j=k, k=0, 1, \dots, 2N-2 \\
E_{x-y}(k) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E(i, j), \quad |i-j|=k, k=0, 1, \dots, N-1 \\
\mu_x &= \sum_{i=0}^{N-1} i E_x(i), \quad \mu_y = \sum_{j=0}^{N-1} j E_y(j) \\
\sigma_x^2 &= \sum_{i=0}^{N-1} (i - \mu_x)^2 E_x(i) \\
\sigma_y^2 &= \sum_{j=0}^{N-1} (j - \mu_y)^2 E_y(j) \\
HXY &= - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E(i, j) \log\{E(i, j)\} \\
HX &= - \sum_{i=0}^{N-1} E_x(i) \log\{E_x(i)\} \\
HY &= - \sum_{j=0}^{N-1} E_y(j) \log\{E_y(j)\} \\
HXY_1 &= - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E(i, j) \log\{E_x(i) E_y(j)\} \\
HXY_2 &= - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E_x(i) E_y(j) \log\{E_x(i) E_y(j)\} \\
Q(i, j) &= \sum_{k=0}^{N-1} \frac{E(i, k) E(k, j)}{E_x(i) E_y(j)}
\end{aligned}$$

なお、これらの画像統計量は、非特許文献1によって上記と同様に導入されている。

- 画像微分手段22は、原画像取得手段12の出力する原画像122に微分処理
 5 を行なう。この微分処理は、本実施の形態においては、一次微分フィルタとして
 知られるSobelフィルタ（水平微分フィルタ：図4a参照、垂直微分フィル

タ：図4 b 参照）を適用する。この際、図4では、注目画素である中央の画素に対して、周囲の画素を図示する各重みを考慮して加算する処理が、注目画素を順次走査して行なわれる。これにより、画像微分手段22によって微分画像222が得られる。

- 5 微分画像行列算出手段24は、原画像行列算出手段14が原画像122に行なった処理と同様の処理により、微分画像222に基づいて第2の同時生成行列242を算出する。また、微分画像統計量算出手段26は、原画像統計量算出手段20と同様の処理により、第2の同時生成行列242から第2組の画像統計量262を算出する。
- 10 判定基準設定手段28は、原画像122が表1の結晶状態のいずれのものあるかを結晶状態データ30として受け付ける。この結晶状態データ30は、表1のスコア0～9の10段階のいずれであることを示すデータである。また、判定基準設定手段28は、前記第1組の画像統計量202と、前記第2組の画像統計量262とを受け付ける。
- 15 判定基準設定手段28は、これらの受け付けたデータから、4種の判定基準を線形判別法により確立する。線形判別法は、統計標本が持ついくつかの特性から、統計標本がどの群（クラス）に属するかを識別する手法である。このとき、統計標本としては、結晶状態データ30を用い、その識別が行なわれる空間（特徴空間）として、微分を用いない原画像に基づいている第1の画像統計量の14次元空間、あるいは、微分画像に基づいている第2の画像統計量の14次元空間が
- 20 選択される。

判定基準は、線形識別関数という関数として表現され、一つの線形識別関数で

判定対象を2つに分類するような基準である。この線形識別関数は、14次元空間内の超平面（14次元空間を2つの半空間に仕切る13次元の自由度を有する面）として決定される。つまり、線形識別関数は、14次元空間に散布されたデータ（各々が、予め判定されているタンパク質の結晶状態のいずれかの情報によ

5 って分類可能であり、第1の画像統計量または第2の画像統計量のデータを座標値としてプロットされている点）を2つに仕切る際の超平面として定義される。

線形識別関数を決定するための指標は、直感的には、予め分類されている各データを出来るだけ分離させるような仕切り方となる超平面（線形識別関数）を定めることである。具体的には、第1の画像統計量または第2組の画像統計量を要素とするデータ $u = (u_1, \dots, u_{14})^T$ と、係数行列（係数ベクトル） A と

10 の内積をとって、スカラー量 z を算出する（式（15））。なお、ここでは、係数行列 A は、後にフィッティングされる未知パラメータを要素としている。また、データ u は、多数のデータが用いられるが、各々が、予め結晶状態が判定されていて結晶状態データ30が得られ、第1の画像統計量および第2画像統計量が算

15 出される複数のデータである。

$$z = A^T u \quad \dots\dots\dots (15)$$

次に、クラス内共分散行列 Σ_Z 、クラス間共分散行列 Σ_B が、式（16）、（17）により求められる。

$$\Sigma_W = \sum_{i=1,2} P(\omega_i) \Sigma_i \quad \dots\dots\dots (16)$$

$$\Sigma_B = \sum_{i=1,2} P(\omega_i) (m_i - m)(m_i - m)^T \quad \dots\dots\dots (17)$$

20 ここで、 $P(\omega_i)$ は2つに分類された分類の各クラス ω_i の事前確率、 Σ_i はク

ラス ω_i の共分散行列、 m_i はクラス ω_i の特徴データ（14次元の座標値）の単純平均によって作られるベクトル、 m は全パターンにおける同様の平均のベクトルを表わす。

データ u を係数行列 A で変換したことに対応させて、スカラー量 z のクラス内

- 5 およびクラス間の分散 $\Sigma \sim Z$ 、 $\Sigma \sim B$ を求めると、

$$\tilde{\Sigma}_W = A^T \Sigma_W A \quad \dots\dots\dots (18)$$

$$\tilde{\Sigma}_B = A^T \Sigma_B A \quad \dots\dots\dots (19)$$

のようになる。これを用いて、評価関数 $J \Sigma (A)$ を式 (20) のように定める

。

$$J_{\Sigma}(A) = \frac{\tilde{\Sigma}_B}{\tilde{\Sigma}_W} = \frac{A^T \Sigma_B A}{A^T \Sigma_W A} \quad \dots\dots\dots (20)$$

- 10 この評価関数は、分類されるクラス間の分離度を表わしている。従って、最適となる係数行列 A は、この評価関数 $J \Sigma$ の最大値を与えるものである。

評価関数 $J \Sigma$ の最大値を与える A を算出するには、式 (21) によって与えられる行列の最大の固有値を与える固有ベクトルを求めればよい。

$$\Sigma_W^{-1} \Sigma_B \quad \dots\dots\dots (21)$$

- 15 このようにして求められた最適な A は、式 (22) に定義される識別関数の超平面の法線ベクトルを与えるものである。

$$\begin{aligned} g(u) &= A^T u + u_0 \\ &= a_1 u_1 + \dots + a_{14} u_{14} + u_0 \end{aligned} \quad \dots\dots\dots (22)$$

なお、式 22 の u_0 は、14元のベクトル u に m を代入した場合に $g(m) = 0$

が成立するように定められる。

以上のような処理を、判定基準設定手段 28 が実行する。このために、判定基準設定手段 28 には、図示はしないが、行列演算を適切に行なうための手段が備えられる。具体的には、内積計算（式 15、17 等）、積和演算（式 16、17 5）、固有値演算（固有値の算出、固有ベクトルの決定）が処理できる演算機能を有する。上記の式の説明では、式変形の順を追って説明しているが、判定基準設定手段 28 は式 18～20 については特段計算する必要がない。

本実施形態の結晶化状態判定システム 100 の機能としての判定基準設定手段 28 の具体的な処理を、図 5 を用いて説明する。判定基準設定手段 28 は、まず、予め結晶状態が判定されている観察画像（教示用の観察画像）について、結晶状態データ 30 の入力（図 5、S2）と、その画像について計算された第 1 組の画像統計量および第 2 組の画像統計量の入力（S4）を受け付ける。次に、確立する判定基準を選択する（S6）。そして、処理に用いる画像統計量の組が判定基準に依存するので、確立する判定基準を選択し、予め判定された結晶状態のデータによって各画像を分類する（S8）。この分類は、表 1 に示した各分類を要素とし、その要素の組合わせによってそれぞれが構成された 2 つの分類である。この分類は、以下に説明するように、そのとき確立しようとする判定基準によって定まる。

本実施の形態では、判定基準 g1～g4 を確立し、これを判定に用いる。g1 は、タンパク質溶液が、透明状態と第 1～第 3 の沈殿状態とからなる状態にあるか、または、第 4 の沈殿状態と微結晶状態と結晶状態とからなる状態にあるかを判定する（第 1 の判定基準）。g2 は、タンパク質溶液が、透明状態と第 1 の沈

殿状態とからなる状態にあるか、または、第2～第4の沈殿状態と微結晶状態と結晶状態とからなる状態にあるかを判定する（第2の判定基準）。g3は、タンパク質溶液が、透明状態にあるか、または、第1の沈殿状態にあるかを判定する（第3の判定基準）。g4は、タンパク質溶液が、第2の沈殿状態にあるか、または、第3の沈殿状態にあるかを判定する（第4の判定基準）。なお、各判定基準は式22の多項式の係数を定めるものであり、判定されるべき画像についての画像統計量（14個の統計量）uによって各判定基準の多項式の値を計算し、その正負によって分類がなされる。判定基準g1～g4の正負をどのように定めるかについては、上述した最大固有値の固有ベクトルが、反転されても同様に同じ固有値の固有ベクトルとなる任意性を利用して、任意に定め得る。本実施の形態では、表1のスコアが小さい側の分類が、各判定基準の多項式の評価値が正となるように定める。

例えば判定基準g1による判定においては、透明状態と第1の沈殿状態とからなる状態（第1の状態）と、第2～第4の沈殿状態と微結晶状態と結晶状態とからなる状態（第2の状態）とのいずれにあるかを、分類することに相当する。このように、各判定基準によって異なる分け方で分類が行なわれる。

そして、ステップS8において判定基準に合わせて分類された分類を用いて、クラス内共分散行列とクラス間共分散行列を、式16および17に従って算出する（S10）。そして、式21の行列を算出して（S12）、その行列の最大固有値に対応する固有ベクトルを算出する（S14）。この計算には、行列の固有値や固有ベクトルを数値的に算出する任意の計算方法を用いることが出来る。固有ベクトルが求まると、その固有ベクトルを係数とする多項式（式22）が、ベク

トル u に m (m は全パターンにおける特徴データの平均のベクトル) を代入した場合に $g(m) = 0$ が成立するように、 u_0 を定める (S 16)。

これにより、式 22 の全ての係数が算出されて、その判定基準が確立する。

判定基準 $g_1 \sim g_4$ の全てについてこのような過程を繰り返すことにより (S 18)、判定基準が全て確立することとなる。なお、 g_3 、 g_4 は、既に g_1 、 g_2 によって分類されたあと判定にのみ用いるため、ステップ S 8 における判定基準に合わせた分類は、確立した g_1 、確立した g_2 によって分類されたものについて判定を行なえばよい。以上のようにして、判定基準設定手段 28 が判定基準を確立する。

10 判定基準設定手段 28 は、判定基準 $g_1 \sim g_4$ を判定基準格納手段 32 に格納する。このような判定基準 $g_1 \sim g_4$ によって表 1 の右端の列に示す計算による分類 A～E が分類される。また、いずれの分類にも入らないものが F として分類される。

本実施の形態では、実際にスコア 0～9 に分類される画像を用いて判定基準 $g_1 \sim g_4$ を確立している。判定基準の確立に用いた画像は、スコア 0～9 の各分類について、予め熟練した観察者によって分類されている 20 枚から 60 枚程度の画像である。

続けて、本実施の形態のシステム 100 によって未知タンパク質溶液のサンプルについて判定を行なう方法について説明する。この判定に際しては、判定手段 34 が判定基準格納手段 32 から判定基準 $g_1 \sim g_4$ を読み込むことにより行なう。

未知タンパク質溶液についても、図 1 に示したようにサンプルが準備される。

そして、光学像として観察画像を取得し、判定基準を確立する際と同一の手段（原画像取得手段、原画像行列算出手段、原画像統計量算出手段、画像微分手段、微分画像行列算出手段、微分画像統計量算出手段）によって同一の処理がなされる。これにより、該未知タンパク質溶液のサンプルについての第1組の画像統計量と第2組の画像統計量が得られる。

第1組の画像統計量と第2組の画像統計量は、各々14次元の数値データからなるベクトルであるので、これらに、判定基準設定手段28によって確立された4種の判定基準 $g_1 \sim g_4$ を適用して判定を行なう。つまり、判定基準を示す多項式にこの未知タンパク質溶液から得られた画像統計量のベクトルを代入して、

10 値の正負を評価することにより、判定を行なう。判定結果のデータ36は、判定基準設定手段28から出力され、適当な表示手段に表示されたり、記憶手段に格納される。

次に、図6に従って、第1～第4の判定基準（ $g_1 \sim g_4$ ）を適用する具体的なステップについて説明する。

15 判定手段は、まず、第1の判定基準 g_1 と第2の判定基準 g_2 とを用いて、判定を行なう（図6、S20）。ここで、 g_1 が正であることは、透明状態と第1～第3の沈殿状態とからなる状態にあることに対応し、 g_1 が負であることは、第4の沈殿状態と微結晶状態と結晶状態とからなる状態にあることに対応する。同様に、 g_2 の正と負は、それぞれ、透明状態と第1の沈殿状態とからなる状態

20 にあること、第2～第4の沈殿状態と微結晶状態と結晶状態とからなる状態にあることに対応する。したがって、 g_1 と g_2 が共に正のときは、未知タンパク質溶液が透明状態と第1の沈殿状態とからなる状態（図6、分枝202；表1、Aま

たはB)にある。また、 g_1 が正、 g_2 が負のときは、第2～第3の沈殿状態からなる状態(分枝204;表1、CまたはD)にあり、 g_1 と g_2 が共に負のときは、第4の沈殿状態と微結晶状態と結晶状態とからなる状態(分枝206;表1、E)にある。なお、判定基準において、 g_1 が負、 g_2 が正という組合わせについて、実際のタンパク質溶液の結晶化状態を示す画像がこのようになることはほとんどなく、この場合はエラーとして図示しない処理がなされ、表1のFに分類される。

次に、画像が透明状態と第1の沈殿状態からなる状態(分枝202;表1、AまたはB)に判定されると、第3の判定基準 g_3 を用いて判定を行なう(S22)。 g_3 の正負は、それぞれ、未知タンパク質溶液が、透明状態(分枝222;表1、A)にあるか、第1の沈殿状態(分枝224;表1、B)にあるかに対応する。

また、画像が、第2～第3の沈殿状態からなる状態(分枝204;表1、CまたはD)に判定されると、第4の判定基準 g_4 を用いて判定を行なう(S24)。 g_4 の正負は、それぞれ、未知タンパク質溶液が、第2の沈殿状態(分枝242;表1、C)にあるか、第3の沈殿状態(分枝244;表1、D)にあるかに対応する。

以上のようにして、表1のA～Eのいずれであるかが判定された後、各判定結果に応じて、判定結果を保持する変数Xにその結果を格納する(S26a～S26e)。そして、その判定結果を出力する(S28)。

表2に、本実施の形態において得られた未知タンパク質の画像の各分類の結果を示す。表の各行には、同じ画像を人により分類した結晶化状態であり、各列に

は、本実施の形態のシステムによって処理をして求められた分類である。各データは、人による分類と計算による分類とに当てはまる画像の実際の枚数である。

「結果」の列には、人により分類されている画像が、計算によって適切に分類されている割合を示している。

5 (表 2)

		計算による分類						結果
		A	B	C	D	E	F	
人による分類	0 : 透明	50	0	0	0	0	0	100%
	1 : 沈殿 (i)	2	41	5	1	0	0	84%
	2 : 沈殿 (i i)	3	16	27	3	0	0	55%
	3 : 沈殿 (i i i)	0	0	2	30	18	0	60%
	4 : 沈殿 (i v)	0	0	0	2	48	0	96%
	5 : 微結晶	0	1	1	7	41	0	82%
	6 : 結晶 (i)	0	0	0	2	23	0	92%
	7 : 結晶 (i i)	0	0	1	1	22	1	92%
	8 : 結晶 (i i i)	0	0	0	2	48	0	96%
	9 : 結晶 (i v)	0	0	1	1	46	1	93%

このように、人によって分類されていた分類のうち、透明 (表 1、A)、沈殿 (i) (B)、沈殿 (i i) (C)、沈殿 (i i i) (D) が、これらの相互からも、そして、残りの分類 (E) から良好に分類されている。

- 10 本実施の形態の構成から、本発明によって提供されるタンパク質溶液の結晶化状態の判定方法をまとめると図 8 のようになる。

- 即ち、タンパク質溶液の電子画像を取得して原画像とする (原画像取得ステップ、S 4 2)。そして、該原画像の少なくとも一部分について第 1 の同時生成行列を算出する (原画像行列算出ステップ、S 4 4)。さらに、該第 1 の同時生成
- 15 行列から第 1 組の画像統計量を算出する (原画像統計量算出ステップ、S 4 6)。

原画像には、微分処理を行なって微分画像が求められる (画像微分ステップ、S 4 8)。微分画像の少なくとも一部分について第 2 の同時生成行列が算出され

る（微分画像行列算出ステップ、S 5 0）。この第 2 の同時生成行列から第 2 組の画像統計量が算出される（微分画像統計量算出ステップ、S 5 2）。第 2 組の画像統計量は、タンパク質溶液の電子画像を少なくとも 2 種に分類するのに用いられる（第 1 の判定ステップ、S 5 4）。第 1 の判定ステップで分類された結果、

5 前記 2 種のうちの第 1 種に属すると判定されたタンパク質溶液の電子画像は、さらに前記第 1 組の画像統計量を用いて少なくとも 2 種に分類される（第 2 の判定ステップ、S 5 6）。ここで、第 1 の判定ステップは、図 6 においてステップ S 2 0 として表わされ、第 2 の判定ステップは、図 6 においてステップ S 2 2 として表わされている。この処理は、第 1 の判定ステップおよび第 2 の判定ステップ

10 に本実施の形態のようにして確立した判定基準を用いてコンピュータの動作により実行される。

本発明においては、タンパク質溶液の電子画像の原画像と微分画像を共に用いることにより、タンパク質の溶液における結晶化の状態を良好に判定できる。また、判定ステップを 2 段階とすることにより、タンパク質溶液の結晶化状態を判定

15 する作業の機械化が容易となり、判定のエラーが低減される。また、透明、沈殿（i）、沈殿（i i）、沈殿（i i i）を適切に分類することが可能となる。判定基準とその順序を適切に構築することにより、判定の際のエラーが低減され、人による判定が良好に再現されて機械化が可能となる。さらに本発明の解析システムに、タンパク質の溶液の結晶化サンプルの自動観察手段を組み合わせれば、未知タンパク質の観察とその未知タンパク質の結晶化状態の判定とを連続して自動で行なうタンパク質結晶化状態観察解析システムが構築できる。

20

以上のように、いくつかの例示的な実施形態が本発明に関して詳細に説明され

た。当業者には、本発明の新規な開示および効果から実質的に離れることなくこれら実施形態を様々に改良することが可能であることが分るであろう。したがって、そのような改良も全て本発明の範囲に含められるべきことが意図されている。

請 求 の 範 囲

1. タンパク質溶液の電子画像を取得して原画像とする原画像取得ステップと、
、
該原画像の少なくとも一部分について第1の同時生成行列を算出する原画像
5 行列算出ステップと、
該第1の同時生成行列から第1組の画像統計量を算出する原画像統計量算出
ステップと、
該原画像に微分処理を行なって微分画像を得る画像微分ステップと、
該微分画像の少なくとも一部分について第2の同時生成行列を算出する微分
10 画像行列算出ステップと、
該第2の同時生成行列から第2組の画像統計量を算出する微分画像統計量算
出ステップと、
該第2組の画像統計量を用いて、タンパク質溶液の電子画像を少なくとも2種
に分類する第1の判定ステップと、
15 前記第1の判定ステップで分類された結果、前記2種のうちの第1種に属する
と判定されたタンパク質溶液の電子画像をさらに前記第1組の画像統計量を用
いて少なくとも2種に分類する第2の判定ステップと
を含む、タンパク質溶液の結晶化状態の判定方法。
- 20 2. 前記第1の判定ステップで分類された結果、前記第1種以外の種に分類さ
れたタンパク質溶液の電子画像を、前記第2組の画像統計量を用いてさらに少な
くとも2種に分類する第3の判定ステップ

をさらに含む、請求項 1 に記載の方法。

3. 予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第 1 の判定ステップにおいて用いる第 1 の判定基準と第 2 の判定基準とを、前記第 2 組の画像統計量に基づく線形判別法により確立するステップと、

前記予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第 2 の判定ステップにおいて用いる第 3 の判定基準を、前記第 1 組の画像統計量に基づく線形判別法により確立するステップと、

- 10 前記予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第 3 の判定ステップにおいて用いる第 4 の判定基準を、前記第 2 組の画像統計量に基づく線形判別法により確立するステップと

をさらに含む、

前記第 1 の判定ステップは、前記第 1 の判定基準と前記第 2 の判定基準とにより、前記第 2 組の画像統計量を用いて電子画像を 3 種に分類するものであり、

- 15 前記第 2 の判定ステップは、前記第 3 の判定基準により、該第 1 組の画像統計量を用いて前記 3 種のうちの第 1 種に属する電子画像をさらに 2 種に分類するものであり、

- 前記第 3 の判定ステップは、前記第 4 の判定基準により、該第 2 組の画像統計量を用いて前記第 1 種以外のいずれかの種に属する電子画像をさらに 2 種に分類するものである、請求項 2 に記載の方法。

4. 請求項 3 に記載の方法において、透明状態、第 1 ～第 4 の沈殿状態、微結

晶状態、および結晶状態を要素として構成される複数の状態のいずれかにタンパク質溶液の電子画像を分類する方法であって、

前記第 1 の判定基準は、タンパク質溶液が、透明状態と第 1 ～第 3 の沈殿状態とからなる状態のいずれか、あるいは、第 4 の沈殿状態と微結晶状態と結晶状態
5 とからなる状態のいずれかのどちらにあるかを判定するものであり、

前記第 2 の判定基準は、タンパク質溶液が、透明状態と第 1 の沈殿状態とからなる状態のいずれか、あるいは、第 2 ～第 4 の沈殿状態と微結晶状態と結晶状態とからなる状態のいずれかのどちらにあるかを判定するものであり、

前記第 3 の判定基準は、タンパク質溶液が、透明状態、あるいは、第 1 の沈殿
10 状態のどちらにあるかを判定するものであり、

前記第 4 の判定基準は、タンパク質溶液が、第 2 の沈殿状態、あるいは、第 3 の沈殿状態のどちらにあるかを判定するものであり、

前記第 1 の判定ステップは、前記第 1 の判定基準と前記第 2 の判定基準とにより、タンパク質溶液が、透明状態と第 1 の沈殿状態からなる状態のいずれか、第
15 2 ～第 3 の沈殿状態からなる状態のいずれか、あるいは、第 4 の沈殿状態と微結晶状態と結晶状態とからなる状態のいずれかのいずれにあるかを判定するステップであり、

前記第 2 の判定ステップは、前記第 3 の判定基準により、透明状態と第 1 の沈殿状態からなる状態のいずれかにあるタンパク質溶液が、透明状態、あるいは、
20 第 1 の沈殿状態のどちらにあるかを判定するステップであり、

前記第 3 の判定ステップは、前記第 4 の判定基準により、第 2 ～第 3 の沈殿状態からなる状態のいずれかにあるタンパク質溶液が、第 2 の沈殿状態、あるいは

、第3の沈殿状態のどちらにあるかを判定するステップである、方法。

5. タンパク質溶液の電子画像を取得して原画像とする原画像取得手段と、
該原画像の少なくとも一部分について第1の同時生成行列を算出する原画像

5 行列算出手段と、

該第1の同時生成行列から第1組の画像統計量を算出する原画像統計量算出
手段と、

該原画像に微分処理を行ない、微分画像を得る画像微分手段と、

該微分画像の少なくとも一部分について第2の同時生成行列を算出する微分

10 画像行列算出手段と、

該第2の同時生成行列から第2組の画像統計量を算出する微分画像統計量算
出手段と、

該第2組の画像統計量を用いて、タンパク質溶液の電子画像を少なくとも2種
に分類する第1の判定手段と、

15 前記第1の判定手段で分類された結果、前記2種のうちの第1種に属すると判
定されたタンパク質溶液の電子画像をさらに前記第1組の画像統計量を用いて
少なくとも2種に分類する第2の判定手段と

を含む、タンパク質溶液の結晶化状態の判定システム。

20 6. 前記第1の判定手段で分類された結果、前記第1種以外の種に分類された
タンパク質溶液の電子画像を、前記第2組の画像統計量を用いてさらに少なくと
も2種に分類する第3の判定手段

をさらに含む、請求項5に記載のシステム。

7. 判定基準設定手段をさらに備え、

該判定基準設定手段は、

- 5 予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第1の判定手段において用いる第1の判定基準と第2の判定基準とを、前記第2組の画像統計量に基づく線形判別法により確立し、

前記予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第2の判定手段において用いる第3の判定基準を、前記第1組の画像統計量に基づく

- 10 線形判別法により確立し、

前記予め判定されているタンパク質の結晶状態の電子画像を用いて、前記第3の判定手段において用いる第4の判定基準を、前記第2組の画像統計量に基づく線形判別法により確立するものであり、

前記第1の判定手段は、前記第1の判定基準と前記第2の判定基準とにより、

- 15 前記第2組の画像統計量を用いて電子画像を3種に分類するものであり、

前記第2の判定手段は、前記第3の判定基準により、該第1組の画像統計量を用いて前記3種のうちの第1種に属する電子画像をさらに2種に分類するものであり、

前記第3の判定手段は、前記第4の判定基準により、該第2組の画像統計量を

- 20 用いて前記第1種以外のいずれかの種に属する電子画像をさらに2種に分類するものである、請求項6に記載のシステム。

8. 請求項7に記載のシステムにおいて、透明状態、第1～第4の沈殿状態、微結晶状態、および結晶状態を要素として構成される複数の状態のいずれかにタンパク質溶液の電子画像を分類するシステムであって、

前記第1の判定基準は、タンパク質溶液が、透明状態と第1～第3の沈殿状態
5 とからなる状態のいずれか、あるいは、第4の沈殿状態と微結晶状態と結晶状態
とからなる状態のいずれかのどちらにあるかを判定するものであり、

前記第2の判定基準は、タンパク質溶液が、透明状態と第1の沈殿状態とから
なる状態のいずれか、あるいは、第2～第4の沈殿状態と微結晶状態と結晶状態
とからなる状態のいずれかのどちらにあるかを判定するものであり、

10 前記第3の判定基準は、タンパク質溶液が、透明状態、あるいは、第1の沈殿
状態のどちらにあるかを判定するものであり、

前記第4の判定基準は、タンパク質溶液が、第2の沈殿状態、あるいは、第3
の沈殿状態のどちらにあるかを判定するものであり、

前記第1の判定手段は、前記第1の判定基準と前記第2の判定基準とにより、
15 タンパク質溶液が、透明状態と第1の沈殿状態からなる状態のいずれか、第2～
第3の沈殿状態からなる状態のいずれか、あるいは、第4の沈殿状態と微結晶状
態と結晶状態とからなる状態のいずれかのいずれにあるかを判定するものであ
り、

前記第2の判定手段は、前記第3の判定基準により、透明状態と第1の沈殿状
20 態からなる状態のいずれかにあるタンパク質溶液が、透明状態、あるいは、第1
の沈殿状態のどちらにあるかを判定するものであり、

前記第3の判定手段は、前記第4の判定基準により、第2～第3の沈殿状態か

らなる状態のいずれかにあるタンパク質溶液が、第2の沈殿状態、あるいは、第3の沈殿状態のどちらにあるかを判定するものである、システム。

FIG.1

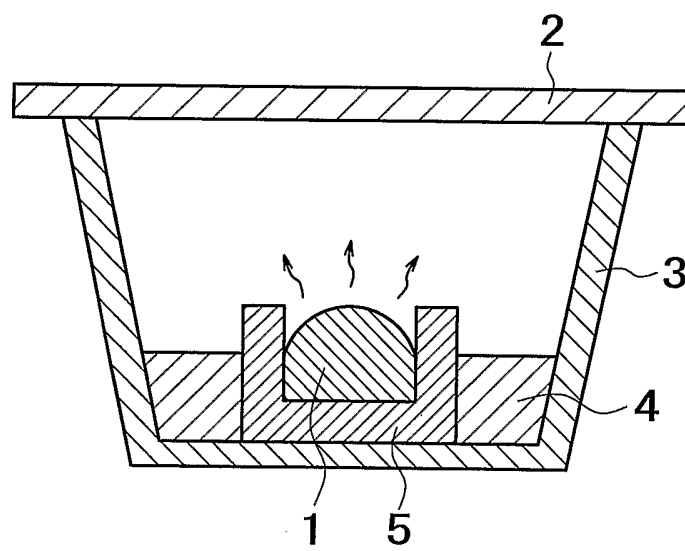


FIG.2

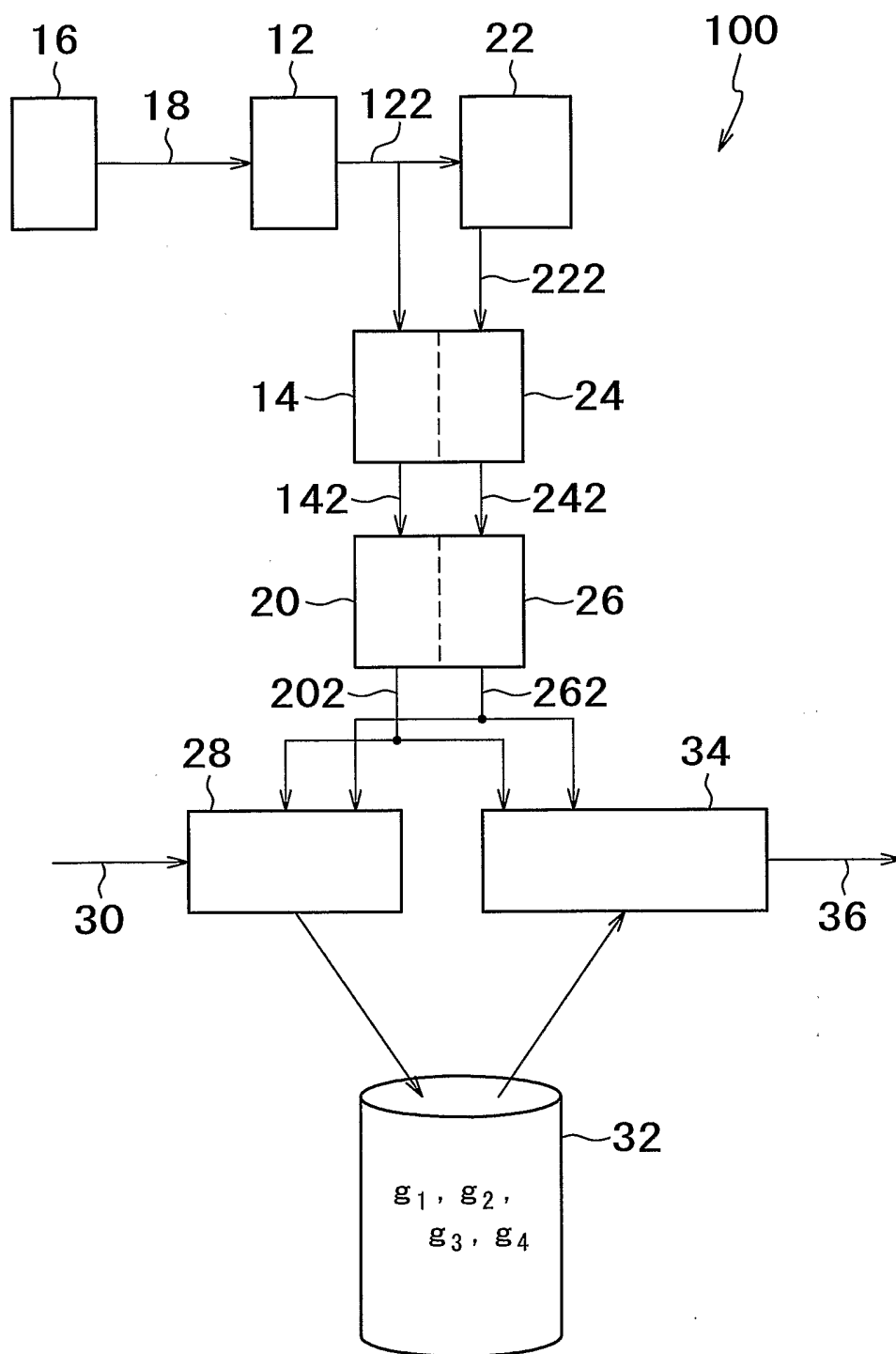
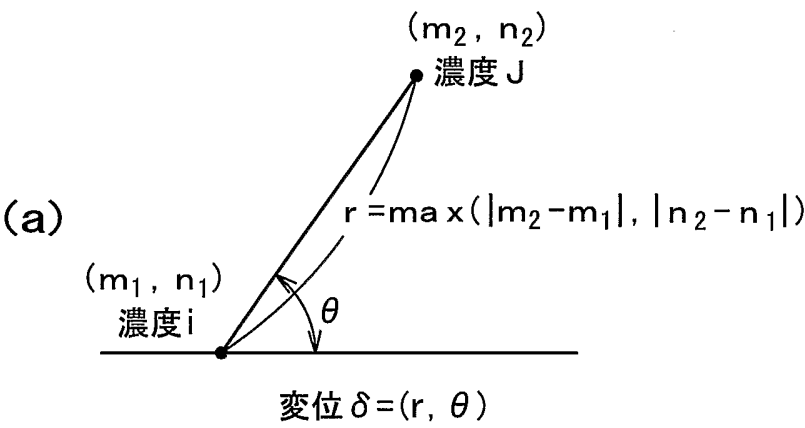


FIG.3



(b)

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

FIG.4

(a)

- 1	0	1
- 2	0	2
- 1	0	1

(b)

- 1	- 2	- 1
0	0	0
1	2	1

5/8

FIG.5

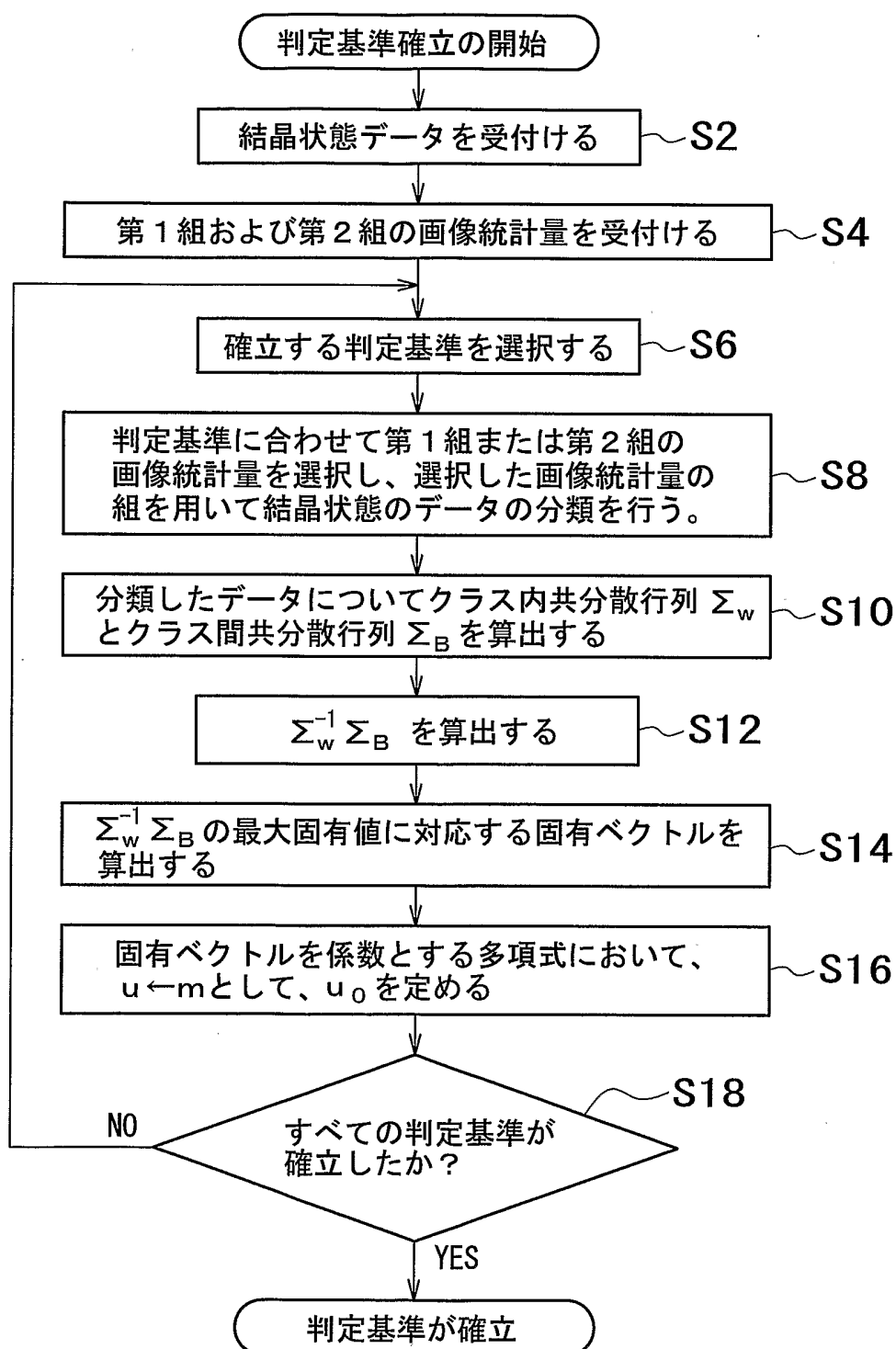


FIG.6

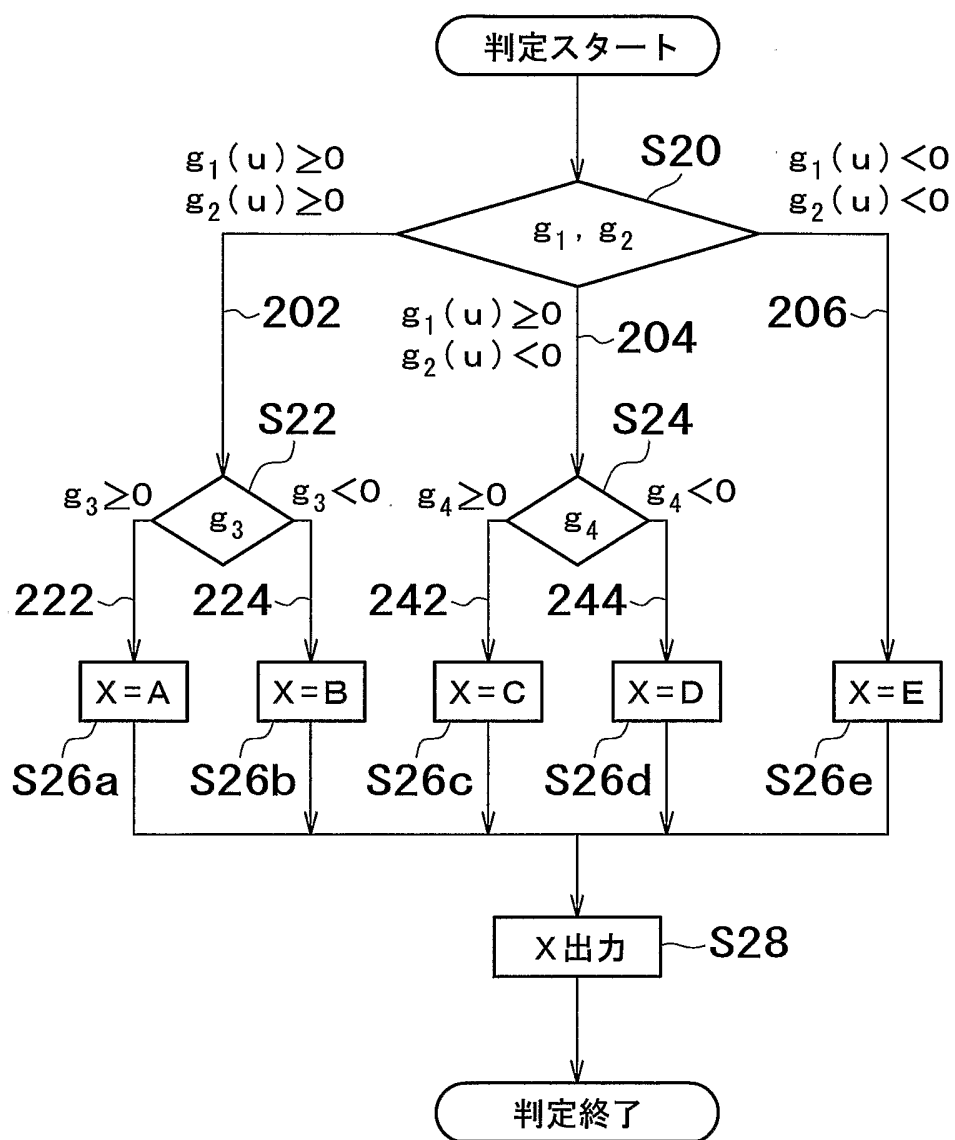


FIG.7

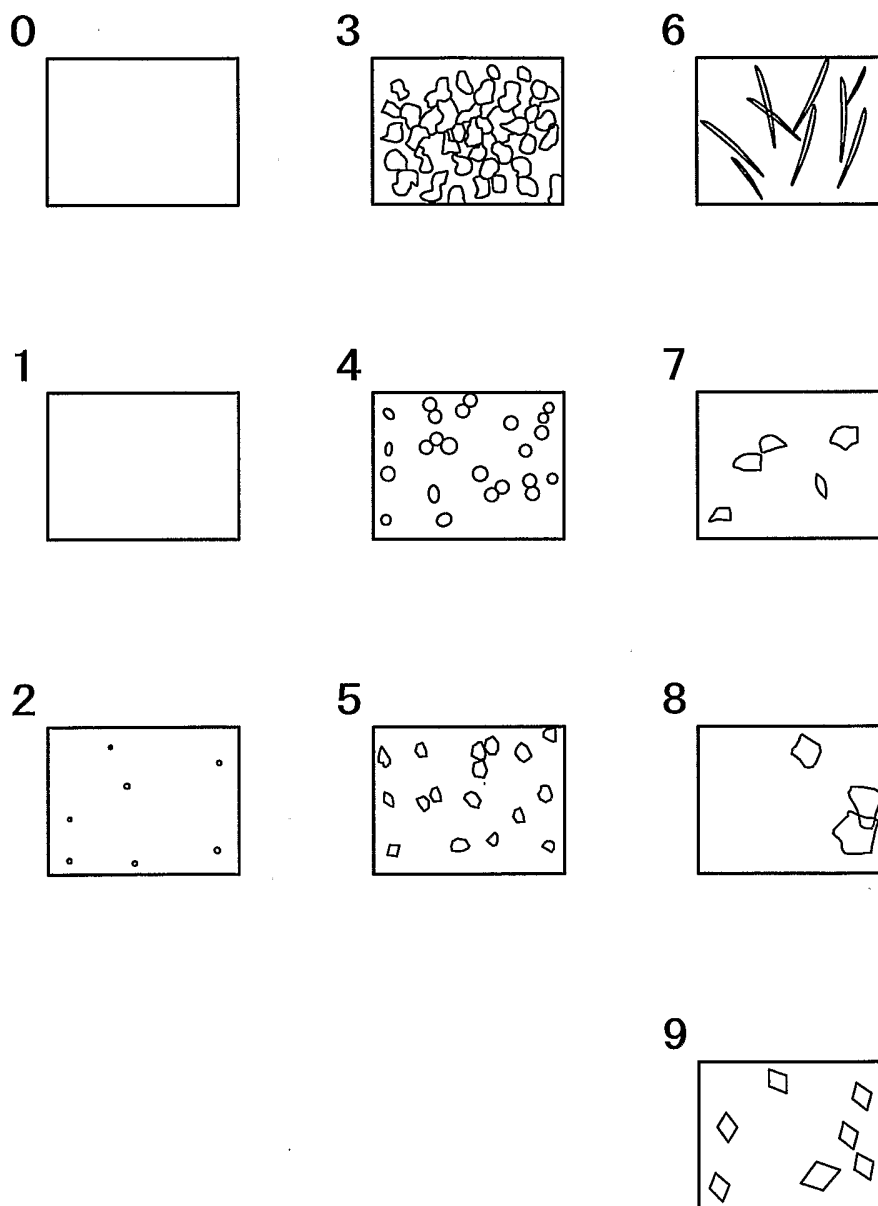
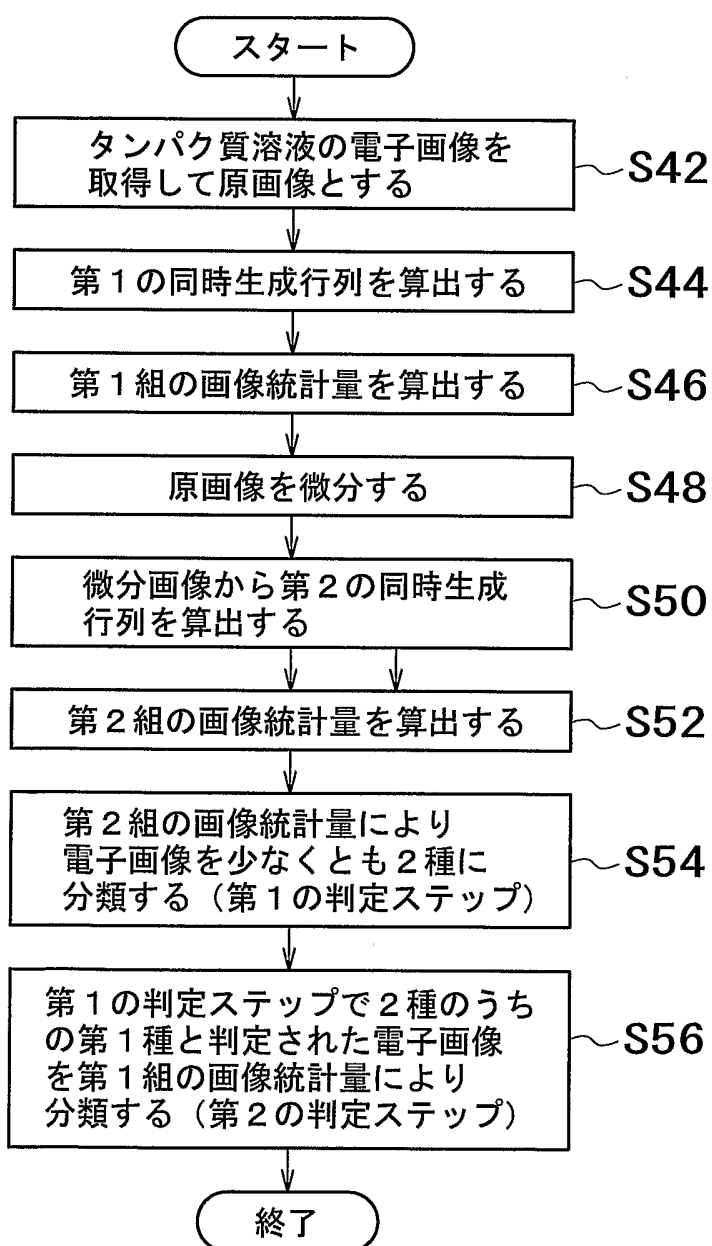


FIG.8



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/008827

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G01N33/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 7-35687 A (Fujitsu Ltd.), 07 February, 1995 (07.02.95), (Family: none)	1-8
A	JP 2002-541455 A (The UAP Research Foundation), 03 December, 2002 (03.12.02), & WO 00/60345 A & EP 1181548 A & CA 2369935 A & AU 4197300 A & US 2002-164812 A	1-8
A	ROBERT M. et al., "Textural Features for Image Classification", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, Vol.SMC-3, No.6, pages 610 to 621, (1973)	1-8

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
24 December, 2004 (24.12.04)

Date of mailing of the international search report
25 January, 2005 (25.01.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G01N33/68

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G01N33/68

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2004年
 日本国登録実用新案公報 1994-2004年
 日本国実用新案登録公報 1996-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 7-35687 A (富士通株式会社) 1995. 02. 07 (ファミリーなし)	1-8
A	J P 2002-541455 A (ザ ユーエーピー リサーチ ファウンデーション) 2002. 12, 03 & WO 00/60345 A & EP 1181548 A & CA 2369935 A & AU 4197300 A & US 2002-164812 A	1-8

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

24. 12. 2004

国際調査報告の発送日

25. 1. 2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

竹中 靖典

2 J

9507

電話番号 03-3581-1101 内線 3251

C (続き) . 関連すると認められる文献		
引用文献の カテゴリ*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	ROBERT M. et.al, "Textural Features for Image Classification", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, Vol.S MC-3, No.6, P610-621, (1973)	1 - 8