

明 細 書

発明の名称：

生体組織解析装置、生体組織解析プログラムおよび生体組織解析方法 技術分野

[0001] 本発明は、ハイパースペクトルデータを用いて生体組織を解析する技術に関し、特に、細胞核のガン化状態等のように、生体組織の変質状態を識別するのに好適な生体組織解析装置、生体組織解析プログラムおよび生体組織解析方法に関するものである。

背景技術

[0002] 従来、細胞核のガン化状態等のように、生体組織の変質状態を診断するに際しては、病理専門医による顕微鏡画像を利用した形態学的診断が行われている。しかしながら、そのような形態学的診断は、病理専門医の知識や経験によるところが大きく、一人前になるまでには数年に渡るトレーニングが必要とされるほど困難なものとされている。

[0003] そこで、近年、分光スペクトルを用いて生体組織を光学的に分析する方法が検討されている。例えば、特開 2017-203637 号公報には、試料に含まれる細胞の分光スペクトルに基づいて、統計的手法、機械学習又はパターン認識により、当該細胞が腫瘍細胞であるか否かを判定する分析工程を有する腫瘍細胞検出方法が提案されている（特許文献 1）。

先行技術文献

特許文献

[0004] 特許文献 1：特開 2017-203637 号公報

発明の概要

発明が解決しようとする課題

[0005] しかしながら、特許文献 1 に記載された発明においては、分析対象の細胞が腫瘍細胞であるか否かを判定するものに過ぎない。このため、例えば、発症初期の大腸癌等における 4 段階のガン化状態（非ガン細胞、軽度異形成（

L G D : Low Grade Dysplasia)、高度異形成 (H G D : High Grade Dysplasia)、ガン細胞) 等のように、生体組織の変質状態を細かく識別することはできないという問題がある。

[0006] 本発明は、このような問題点を解決するためになされたものであって、ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別することができる生体組織解析装置、生体組織解析プログラムおよび生体組織解析方法を提供することを目的としている。

課題を解決するための手段

[0007] 本発明に係る生体組織解析装置は、ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別するという課題を解決するために、二次元画像を構成する各ピクセルごとにスペクトル情報に対応付けてなるハイパースペクトルデータを用いて生体組織を解析する生体組織解析装置であって、下記 (i) , (ii) を有するとともに、下記 (iii) および／または (iv) を有する、前記生体組織解析装置；

(i) ハイパースペクトルカメラから生体組織の前記ハイパースペクトルデータを取得するハイパースペクトルデータ取得部；

(ii) 前記ハイパースペクトルデータに基づいて、前記生体組織の二次元画像から解析対象部位に相当するピクセルを画像処理によって抽出する解析対象部位抽出部；

(iii) 前記解析対象部位に相当するピクセルのそれぞれを教師なし学習によって分類することにより、前記生体組織の変質状態を大別する変質状態大別部；

(iv) 前記解析対象部位に相当するピクセルのそれぞれを教師あり学習によって分類することにより、前記生体組織の変質状態を識別する変質状態識別部。

[0008] また、本発明の一態様として、生体組織の変質状態を予め大別することで、解析精度を向上するという課題を解決するために、前記変質状態識別部によって識別する前に、前記解析対象部位に相当するピクセルのそれぞれを教

師なし学習によって分類する変質状態大別部を有していてもよい。

[0009] さらに、本発明の一態様として、解析用データに対する光源の影響を除去し、解析精度を向上するという課題を解決するために、前記ハイパースペクトルデータの取得時に用いた光源のスペクトル情報によって、前記解析対象部位に相当するピクセルの前記スペクトル情報を除算する光源影響除去部を有していてもよい。

[0010] また、本発明の一態様として、解析用データに対する染色の影響を除去し、解析精度を向上するという課題を解決するために、前記解析対象部位に相当するピクセルの前記スペクトル情報に基づいて、前記解析対象部位の染色濃度によって変化するスペクトル成分を除去する染色影響除去部を有していてもよい。

[0011] さらに、本発明の一態様として、解析対象部位をピクセル単位で抽出するという課題を解決するために、前記解析対象部位抽出部は、前記スペクトル情報の類似度に基づいて各ピクセルを解析対象部位とその他の部位とにラベリングし、前記解析対象部位としてラベリングされたピクセルをフィルタリングしてもよい。

[0012] また、本発明の一態様として、解析対象部位を細胞核単位で抽出するという課題を解決するために、前記解析対象部位抽出部は、前記生体組織の二次元画像を画像認識処理によって腺管を特定するとともに、前記腺管上における細胞核を解析対象部位として画像認識処理によって特定してもよい。

[0013] さらに、本発明の一態様として、潰瘍やポリープにおける細胞核のガン化状態を病理診断に従って識別するという課題を解決するために、前記変質状態大別部によって特定される前記変質状態は、潰瘍における、非ガン細胞および軽度異形成（L G D : Low Grade Dysplasia）の2ステージであり、前記変質状態識別部によって識別される前記変質状態は、潰瘍における、非ガン細胞、軽度異形成（L G D : Low Grade Dysplasia）、高度異形成（H G D : High Grade Dysplasia）およびガン細胞の4ステージ、またはポリープにおける、軽度異型腺腫（L G A : Low Grade Adenoma）、高度異型腺腫（H G A

: High Grade Adenoma) およびガン細胞の３ステージによって特定されてもよい。

[0014] また、本発明の一態様として、生体組織の変質状態を解析するのに適したハイパースペクトルデータを取得するという課題を解決するために、前記ハイパースペクトルデータの波長帯域に可視光が含まれていてもよい。

[0015] また、本発明に係る生体組織解析プログラムは、ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別するという課題を解決するために、二次元画像を構成する各ピクセルごとにスペクトル情報を対応付けてなるハイパースペクトルデータを用いて生体組織を解析する生体組織解析プログラムであって、下記 (i) , (ii) としてコンピュータを機能させるとともに、下記 (iii) および／または (iv) としてコンピュータを機能させる、前記生体組織解析プログラム；

(i) ハイパースペクトルカメラから生体組織の前記ハイパースペクトルデータを取得するハイパースペクトルデータ取得部；

(ii) 前記ハイパースペクトルデータに基づいて、前記生体組織の二次元画像から解析対象部位に相当するピクセルを画像処理によって抽出する解析対象部位抽出部；

(iii) 前記解析対象部位に相当するピクセルのそれぞれを教師なし学習によって分類することにより、前記生体組織の変質状態を大別する変質状態大別部；

(iv) 前記解析対象部位に相当するピクセルのそれぞれを教師あり学習によって分類することにより、前記生体組織の変質状態を識別する変質状態識別部。

[0016] さらに、本発明に係る生体組織解析方法は、ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別するという課題を解決するために、二次元画像を構成する各ピクセルごとにスペクトル情報を対応付けてなるハイパースペクトルデータを用いて生体組織を解析する生体組織解析方法であって、下記 (i) , (ii) を有するとともに、下記 (iii) および／または

(iv) を有する、前記生体組織解析方法；

(i) ハイパースペクトルカメラから生体組織の前記ハイパースペクトルデータを取得するハイパースペクトルデータ取得ステップ；

(ii) 前記ハイパースペクトルデータに基づいて、前記生体組織の二次元画像から解析対象部位に相当するピクセルを画像処理によって抽出する解析対象部位抽出ステップ；

(iii) 前記解析対象部位に相当するピクセルのそれぞれを教師なし学習によって分類することにより、前記生体組織の変質状態を大別する変質状態大別ステップ；

(iv) 前記解析対象部位に相当するピクセルのそれぞれを教師あり学習によって分類することにより、前記生体組織の変質状態を識別する変質状態識別ステップ。

発明の効果

[0017] 本発明によれば、ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別することができる。

図面の簡単な説明

[0018] [図1]本発明に係る生体組織解析装置の一実施形態を示すブロック図である。

[図2]ハイパースペクトルデータの一例を示す図である。

[図3]本実施形態の解析対象部位抽出部によって実行される、(a)スペクトル情報抽出処理、(b)ラベリング処理および(c)フィルタリング処理を示す図である。

[図4]本実施形態の生体組織解析装置および生体組織解析プログラムによって実行される生体組織解析方法を示すフローチャートである。

[図5]実施例1において、各ステージにおける全特徴量の組み合わせについての識別率を示す表である。

[図6]実施例2において、主成分分析によって特定された各主成分の寄与率を示すグラフである。

[図7]実施例3において、ロジスティック回帰による予測結果の混同行列を示

す表である。

[図8]実施例4において、(a) サポートベクターマシンのみによる解析結果、および(b) クラスタ分析とサポートベクターマシンによる解析結果を示す表である。

[図9]実施例5において、ハイパースペクトルデータの独立成分分析結果を示すグラフである。

発明を実施するための形態

[0019] 以下、本発明に係る生体組織解析装置、生体組織解析プログラムおよび生体組織解析方法の実施形態について図面を用いて説明する。

[0020] 本実施形態の生体組織解析システムは、図1に示すように、生体組織のハイパースペクトルデータを取得するハイパースペクトルカメラ10と、このハイパースペクトルカメラ10によって取得されたハイパースペクトルデータに基づいて、生体組織の変質状態を解析する生体組織解析装置1とから構成されている。以下、各構成について説明する。

[0021] なお、本発明において、生体組織とは、人間や動物等を含む生体において、複数の細胞が所定のパターンで集合してなる全ての組織を含む概念である。また、本発明において、変質状態とは、ガンの進行度やガンの悪性度等のように、生体組織がどの程度変質しているかを示す全ての状態を含む概念である。

[0022] ハイパースペクトルカメラ10は、二次元の空間情報と、複数の波長におけるスペクトル情報（ハイパースペクトル情報）とを同時に取得するものである。具体的には、図2に示すように、二次元画像を構成する各ピクセルごとにスペクトル情報を対応付けてなるハイパースペクトルデータを取得するものである。なお、スペクトル情報は、分光スペクトルとも呼ばれ、各バンド（波長）ごとの光強度の分布を示すものである。

[0023] 本実施形態において、上記二次元画像を構成する単位ピクセルは、生体組織における解析対象部位の内部に含まれる大きさに設定されており、当該解析対象部位について複数のピクセルを有するハイパースペクトルデータを取

得するようになっている。また、本実施形態では、潰瘍やポリープにおけるガン化状態（ガンの進行度）を解析する場合、細胞核が解析対象部位とされる。しかしながら、解析対象部位は、リンパ球、粘液、組織球（白血球）等のように、解析しようとする生体組織の変質状態に応じて適宜選択され得る。

[0024] また、本実施形態において、生体組織は、ホルマリン等で固定してからパラフィン等で包埋して切片を作成し、ヘマトキシリン・エオシン染色等を施した上でプレパラートに挟まれている。そして、当該プレパラートを顕微鏡にセットした状態で、可視光源によって照射して得られる透過光を拡大し、当該透過光を接眼レンズに当接させたハイパースペクトルカメラ10によって計測するようになっている。

[0025] なお、一般的な医療機器に使用される赤外線や紫外線は、顕微鏡レンズの透過性（光学コーティング性能）が可視光域に比べて低く、収差の影響も受けやすい。また、顕微鏡レンズに使用されるフッ化カルシウムは、近赤外光の励起で蛍光を発するため、弱いラマン信号の測定を妨害することが知られている。一方、可視光はX線に比べて光子エネルギーが小さく、人体への影響が少ない。また、可視光の分光情報は、細胞核内のクロマチン濃度の影響で変化することが知られており、ラマン分光と同様にクロマチンだけでなく多くの細胞組織の情報を潜在的に含んでいる可能性が高い。そこで、本実施形態では、ハイパースペクトルデータの波長帯域に可視光が含まれるように可視光源を使用している。

[0026] つぎに、生体組織解析装置1は、パーソナルコンピュータ等のコンピュータによって構成されており、図1に示すように、主として、表示手段2と、入力手段3と、記憶手段4と、演算処理手段5とを有している。以下、各構成手段について詳細に説明する。

[0027] 表示手段2は、液晶ディスプレイ等で構成されており、生体組織解析装置1による解析結果等を表示するものである。入力手段3は、キーボードやマウス等で構成されており、ユーザによる指示や選択を入力するものである。

なお、本実施形態では、表示機能のみを備えた表示手段2と入力機能のみを備えた入力手段3とを別々に使用しているが、この構成に限定されるものではなく、表示機能および入力機能を兼ね備えたタッチパネル等の表示入力手段を使用してもよい。

[0028] 記憶手段4は、各種のデータを記憶するとともに、演算処理手段5が演算処理を行う際のワーキングエリアとして機能するものである。本実施形態において、記憶手段4は、ハードディスク、ROM (Read Only Memory)、RAM (Random Access Memory)、フラッシュメモリ等で構成されており、図1に示すように、プログラム記憶部41と、教師データ記憶部42とを有している。以下、各構成部についてより詳細に説明する。

[0029] プログラム記憶部41には、本実施形態の生体組織解析装置1を制御するための生体組織解析プログラム1aがインストールされている。そして、演算処理手段5が、当該生体組織解析プログラム1aを実行することにより、生体組織解析装置1としてのコンピュータを後述する各構成部として機能させるようになっている。

[0030] なお、生体組織解析プログラム1aの利用形態は、上記構成に限られるものではない。例えば、CD-ROMやDVD-ROM等のように、コンピュータで読み取り可能な記録媒体に生体組織解析プログラム1aを記憶させておき、当該記録媒体から直接読み出して実行してもよい。また、外部サーバ等からクラウドコンピューティング方式やASP (Application Service Provider) 方式で利用してもよい。

[0031] 教師データ記憶部42は、生体組織に関するスペクトル情報としての教師データを記憶するものである。本実施形態では、潰瘍における細胞核を解析対象部位としているため、潰瘍における、非ガン細胞、軽度異形成 (LGD : Low Grade Dysplasia)、高度異形成 (HGD : High Grade Dysplasia) およびガン細胞という4ステージによって特定される変質状態のそれぞれについて、スペクトル情報が教師データとして登録されている。

[0032] なお、教師データは、上記に限定されるものではなく、解析しようとする

生体組織の変質状態に応じて適宜選択される。例えば、ポリープにおける細胞核のガン化状態を解析する場合には、ポリープにおける、軽度異型腺腫（L G A : Low Grade Adenoma）、高度異型腺腫（H G A : High Grade Adenoma）およびガン細胞という３ステージによって特定される変質状態のそれぞれについて、スペクトル情報が登録されることとなる。

[0033] 演算処理手段５は、ハイパースペクトルカメラ１０からハイパースペクトルデータを取得し、生体組織の解析処理を実行するものである。本実施形態において、演算処理手段５は、ＣＰＵ（Central Processing Unit）等で構成されており、記憶手段４にインストールされた生体組織解析プログラム１ａを実行することにより、図１に示すように、ハイパースペクトルデータ取得部５１と、解析対象部位抽出部５２と、光源影響除去部５３と、染色影響除去部５４と、データ数削減部５５と、変質状態大別部５６と、変質状態識別部５７として機能するようになっている。以下、各構成部についてより詳細に説明する。

[0034] ハイパースペクトルデータ取得部５１は、ハイパースペクトルカメラ１０から生体組織のハイパースペクトルデータを取得するものである。本実施形態において、ハイパースペクトルデータ取得部５１は、ハイパースペクトルカメラ１０に制御信号を出力して生体組織をスキャンさせ、撮像素子から出力された画像信号を１ラインごとにハイパースペクトルデータとして取得するようになっている。

[0035] 解析対象部位抽出部５２は、ハイパースペクトルデータに基づいて、生体組織の二次元画像から解析対象部位に相当するピクセルを画像処理によって抽出するものである。本実施形態において、解析対象部位抽出部５２は、まず、図３（ａ）に示すように、二次元画像を構成する各ピクセルのスペクトル情報を抽出する。つぎに、図３（ｂ）に示すように、抽出したスペクトル情報の類似度に基づいて、各ピクセルを解析対象部位とその他の部位（リンパ球、組織球、粘液等）とにラベリングする。そして、図３（ｃ）に示すように、解析対象部位としてラベリングされたピクセルをフィルタリングする

。このフィルタリングされた各ピクセルのスペクトル情報（各バンドごとの光強度）が解析用データとなる。

[0036] なお、本実施形態において、解析対象部位抽出部52は、ベクトル間距離やマハラノビス距離等によってスペクトル情報の類似度を算出している。また、本実施形態では、解析対象部位およびその他の各部位ごとに類似度の閾値が設定されており、いずれかの閾値以上の類似度を有するピクセルが、当該閾値に対応する部位としてラベリングされるようになっている。

[0037] また、本実施形態において、解析対象部位抽出部52は、スペクトル情報の類似度に基づいて解析対象部位を抽出しているが、この構成に限定されるものではなく、他の抽出方法を採用してもよい。具体的には、潰瘍やポリープは腺管構造を有しており、当該腺管上に解析対象部位となる細胞核（腺上皮細胞）が存在している。このため、本実施形態のように、潰瘍やポリープにおけるガン化状態を解析する場合、解析対象部位抽出部52は、ハイパースペクトルデータから生体組織の二次元画像（RGB画像）を合成し、当該二次元画像を画像認識処理することによって腺管を特定するとともに、当該腺管上における細胞核を解析対象部位として画像認識処理によって特定してもよい。

[0038] 光源影響除去部53は、光源が解析用データに与える影響を除去するものである。本実施形態において、光源影響除去部53は、ハイパースペクトルデータの取得時に用いた光源のスペクトル情報によって、解析対象部位に相当するピクセルのスペクトル情報を除算することにより、光源の影響を除去するようになっている。

[0039] 具体的には、光源影響除去部53は、解析用データの各バンドごとに下記式（1）によって分光反射率 R_i を算出する。

$$R_i = I_i / I_0 \quad \cdots \text{式（1）}$$

ただし、各符号は以下を表す。

I_i : i 番目の細胞核のスペクトル情報

1₀: 光源のスペクトル情報

- [0040] なお、本実施形態では、上記のとおり、解析対象部位抽出部52が、解析対象部位に相当するピクセルを抽出しているため、ピクセルごとに分光反射率 R_i を算出している。しかしながら、上記のとおり、解析対象部位抽出部52が、画像認識処理によって解析対象部位を抽出する場合、当該解析対象部位ごとに分光反射率 R_i を算出する。また、ハイパースペクトルデータの取得時における光源の条件が一定であれば、光源影響除去部53を機能させる必要はなく、スペクトル情報がそのまま解析用データとなる。
- [0041] 染色影響除去部54は、解析対象部位に対する染色が解析用データに与える影響を除去するものである。本実施形態において、染色影響除去部54は、解析対象部位に相当するピクセルのスペクトル情報に基づいて、解析対象部位の染色濃度によって変化するスペクトル成分を除去することにより、染色による影響を除去するようになっている。
- [0042] なお、解析対象部位の染色濃度によって変化するスペクトル成分は、実施例5で後述するとおり、独立成分分析等によって事前に特定される。また、解析対象部位の染色濃度が一定であれば、染色影響除去部54を機能させる必要はなく、上記スペクトル成分を除去していないスペクトル情報がそのまま解析用データとなる。
- [0043] データ数削減部55は、解析用データのデータ数を削減するものである。本実施形態において、データ数削減部55は、主成分分析によって解析用データの次元を削減することにより、解析速度を向上するようになっている。なお、データ数の削減方法は、主成分分析に限定されるものではなく、解析用データに含まれる任意のバンドのみを選択することにより、データ数を削減してもよい。また、解析速度よりも解析精度を優先する場合は、データ数削減部55を機能させずに、全ての解析用データについて解析すればよい。
- [0044] 変質状態大別部56は、解析対象部位に相当するピクセルのそれぞれを教師なし学習によって分類することにより、前記生体組織の変質状態を大別するものである。本実施形態において、変質状態大別部56は、教師なし学習

として、異なる性質のものが混ざり合った集団から、互いに似た性質を持つものを集め、クラスターを作るクラスター分析（ウォード法）を採用している。このため、本実施形態のように、潰瘍におけるガン化状態を解析する場合、実施例1で後述するとおり、上述した4ステージのうち、識別しやすい非ガン細胞および軽度異形成（LGD）の2ステージが高精度に大別される。

[0045] なお、教師なし学習としては、クラスター分析に限定されるものではなく、生体組織の変質状態を大別しうるものであれば適宜採用することができる。また、変質状態大別部56による分類は、変質状態識別部57による識別の前に行うことによって識別対象を絞り込み、識別精度および汎用性を向上するものであるが、必ずしも実行する必要はない。さらに、ポリープにおけるガン化状態を解析する場合は、非ガン細胞というステージはないため、変質状態大別部56による分類は不要である。

[0046] 変質状態識別部57は、解析対象部位に相当するピクセルのそれぞれを教師あり学習によって分類することにより、生体組織の変質状態を識別するものである。本実施形態において、変質状態識別部57は、教師あり学習として、高い分類機能を有するロジスティック回帰を採用している。このため、実施例3で後述するとおり、変質状態大別部56では識別が困難であった高度異形成（HGD）およびガン細胞の2ステージについても高精度に識別される。

[0047] なお、教師あり学習としては、ロジスティック回帰に限定されるものではなく、生体組織の変質状態を識別しうるものであれば、サポートベクターマシン（SVM: Support Vector Machine）、ランダムフォレスト、ディープラーニングなどのニューラルネットワーク関連手法および判別分析等を適宜採用することができる。また、変質状態大別部56によって非ガン細胞および軽度異形成（LGD）を予め大別しない場合には、変質状態識別部57によって直接、上記4ステージのそれぞれを識別してもよい。さらに、変質状態大別部56による分類のみで十分な場合は、変質状態識別部57による識

別を実行する必要はない。

- [0048] つぎに、本実施形態の生体組織解析装置 1、生体組織解析プログラム 1 a および生体組織解析方法による作用について説明する。
- [0049] 本実施形態の生体組織解析装置 1、生体組織解析プログラム 1 a および生体組織解析方法を用いて生体組織の変質状態を解析する場合、図 4 に示すように、まず、ハイパースペクトルデータ取得部 5 1 が、ハイパースペクトルカメラ 1 0 から生体組織のハイパースペクトルデータを取得する（ステップ S 1：ハイパースペクトルデータ取得ステップ）。
- [0050] このとき、本実施形態では、赤外線や紫外線と比較して顕微鏡レンズの透過性が高く、収差の影響を受けにくい可視光源を使用する。また、可視光は、X線と比較して人体への影響が少なく、その分光スペクトルには多くの細胞組織に関する情報が潜在的に含まれている。よって、生体組織の変質状態を解析するのに適したハイパースペクトルデータが取得される。
- [0051] つぎに、解析対象部位抽出部 5 2 が、ハイパースペクトルデータ取得部 5 1 によって取得されたハイパースペクトルデータに基づいて、生体組織の二次元画像から解析対象部位に相当するピクセルを抽出する（ステップ S 2：解析対象部位抽出ステップ）。これにより、当該ピクセルのスペクトル情報が解析用データとして出力されるため、当該解析用データに対して、解析モードに応じた処理が実行される（ステップ S 3）。
- [0052] 具体的には、高精度モードに設定されている場合（ステップ S 3：YES）、光源影響除去部 5 3 が、解析対象部位に相当する各ピクセルのスペクトル情報を光源のスペクトル情報で除算する（ステップ S 4：光源影響除去ステップ）。これにより、解析用データに対する光源の影響が除去されるため、解析精度が向上する。
- [0053] つづいて、染色影響除去部 5 4 が、解析対象部位の染色濃度によって変化するスペクトル成分を除去する（ステップ S 5：染色影響除去ステップ）。これにより、解析用データに対する染色の影響が除去されるため、解析精度が向上する。

- [0054] 一方、高精度モードに設定されていない場合（ステップS 3：NO）、データ数削減部5 5が、解析用データのデータ数を削減する（ステップS 6：データ数削減ステップ）。これにより、実施例2で後述するとおり、一定の解析精度が維持されたまま、解析速度が向上する。
- [0055] 解析モードに応じた処理が実行された後、変質状態大別部5 6が、解析対象部位に相当するピクセルのそれぞれを教師なし学習によって分類する（ステップS 7：変質状態大別ステップ）。これにより、生体組織の変質状態が大別され、変質状態識別部5 7によって識別すべき対象が絞り込まれるため、解析精度が向上する。
- [0056] つづいて、変質状態識別部5 7が、解析対象部位に相当するピクセルのそれぞれを教師あり学習によって分類する（ステップS 8：変質状態識別ステップ）。これにより、生体組織の変質状態が高精度に識別される。なお、本実施形態において、解析結果は、全ピクセル数に対する、変質状態のそれぞれに分類されたピクセル数の割合によって表示手段2等に表示される（ステップS 9：解析結果表示ステップ）。
- [0057] なお、本発明において、上述したステップS 1およびステップS 2は、生体組織解析装置1とは別体の解析装置（プログラム）に実行させてもよい。すなわち、別体の解析装置によって、ハイパースペクトルカメラ10から生体組織のハイパースペクトルデータを取得させ、当該ハイパースペクトルデータに基づいて、解析対象部位に相当するピクセルを抽出させる。そして、当該ピクセルを生体組織解析装置1が取得して解析してもよい。
- [0058] また、上述したステップS 3からステップS 6およびステップS 9については必須の処理ではなく、必要に応じてまたはユーザの選択に応じて適宜実行される処理である。さらに、上述したステップS 7およびステップS 8は、要求される解析精度に応じていずれか一方または双方が実行されるものであり、ステップS 7のみ、ステップS 8のみ、ステップS 7およびステップS 8の双方の順に解析精度が向上する。
- [0059] 以上のような本実施形態の生体組織解析装置1、生体組織解析プログラム

1 a および生体組織解析方法によれば、以下のような効果を奏する。

1. ハイパースペクトルデータを用いて生体組織の変質状態を高精度に識別することができる。
2. 解析用データに対する光源の影響を除去し、解析精度を向上することができる。
3. 解析用データに対する染色の影響を除去し、解析精度を向上することができる。
4. 解析対象部位をピクセル単位で抽出することができる。
5. 解析対象部位を細胞核単位で抽出することができる。
6. 潰瘍やポリープにおける細胞核のガン化状態を病理診断に従って識別することができる。
7. 波長帯域に可視光を含ませることにより、生体組織の変質状態を解析するのに適したハイパースペクトルデータを取得することができる。

[0060] つぎに、本発明に係る生体組織解析装置 1、生体組織解析プログラム 1 a および生体組織解析方法の具体的な実施例について説明する。

実施例 1

[0061] 本実施例 1 では、上述した本実施形態において、変質状態大別部 5 6 による教師なし学習として採用されたクラスター分析の識別性能を確認する実験を行った。

[0062] 具体的には、まず、生体組織として潰瘍性大腸炎に関連する検体を用意し、予め病理医によって、非ガン細胞、軽度異形成（L G D）、高度異形成（H G D）およびガン細胞の 4 ステージを識別させた。つぎに、識別された 4 ステージのそれぞれについて、細胞核のハイパースペクトルデータを取得した。

[0063] 本実施例 1 において、ハイパースペクトルデータは、ハイパースペクトルカメラ 1 0（北海道衛星株式会社：H S C 1 7 0 2）を正立顕微鏡（オリンパス株式会社）の接眼レンズに直接当てて取得した。このとき、顕微鏡に内蔵された白色 L E D を光源とし、顕微鏡の倍率は 4 0 0 倍（接眼 1 0 倍、対

物40倍)であった。また、計測波長域は350nmから1050nmの可視光と近赤外域であり、波長サンプリング間隔を5nmとして計141バンドのハイパースペクトルデータを取得した。

[0064] つぎに、取得したハイパースペクトルデータについてウォード法によるクラスター分析を行い、3～4つの特徴量をクラスターとして抽出した。そして、各ステージについて抽出した全特徴量の識別率(類似度)を主成分分析と線形判別分析とによって算出し、各ステージの指標となる特徴量がないかどうかを検討した。その結果を図5に示す。

[0065] 図5に示すように、非ガン細胞の各特徴量である非ガン1～4は、他の特徴量に対する識別率がほぼ100%であった。よって、非ガン細胞は、クラスター分析によって容易に識別しうることが示された。また、軽度異形成(LGD)の特徴量の一つであるLGD1は、他の特徴量に対する識別がいずれも83.3%以上であり、平均すると約95.5%であった。よって、軽度異形成(LGD)は、ガン細胞や高度異形成(HGD)が持たない特有の特徴量を有しており、この特徴量を用いることにより軽度異形成(LGD)は、クラスター分析によって他のステージから識別できることが示された。

[0066] 以上の本実施例1によれば、教師なし学習としてのクラスター分析は、潰瘍のガン化状態における4ステージのうち、非ガン細胞および軽度異形成(LGD)を高精度に識別できることが示された。

実施例 2

[0067] 本実施例2では、上述した本実施形態において、データ数削減部55によるデータ数の削減方法として採用された主成分分析の性能を確認する実験を行った。

[0068] 具体的には、まず、クラスター分析では識別が困難であったガン細胞と高度異形成(HGD)のサンプルを用意し、実施例1と同様の方法によりハイパースペクトルデータを取得した。つぎに、主成分分析によってハイパースペクトルデータのデータ数を二次元データに削減し、寄与率を算出した。その結果を図6に示す。図6に示すように、第一主成分から第二主成分までの

累積寄与率が90%を超えているため、第三主成分までの次元を削減しても情報量の損失がきわめて少ないことが示された。

- [0069] 以上の本実施例2によれば、主成分分析は、ハイパースペクトルデータの情報量をほとんど損失させることなくデータ数を削減でき、解析速度を向上しうることが示された。

実施例 3

- [0070] 本実施例3では、上述した本実施形態において、変質状態識別部57による教師あり学習として採用されたロジスティック回帰の識別性能を確認する実験を行った。

- [0071] 具体的には、まず、実施例2でデータ数を削減したガン細胞と高度異形成（HGD）のハイパースペクトルデータを用意した。つぎに、数値解析ソフト（Python）の拡張モジュール（scikit-learn）中にあるLogisticRegression関数を使用し、ロジスティック回帰によって解析を行った。このとき、高度異形成（HGD）のサンプル数（490個）と、ガン細胞のサンプル数（320個）の比率を保ったまま、7：3の割合でトレーニングデータとテストデータに分割し、学習と予測（識別）を行ったところ、その識別率は95.9%という高い値であった。

- [0072] また、予測結果について、真陽性、偽陽性、偽陰性および真陰性の個数を表す混合行列を図7に示す。図7の混合行列に基づいて、予測結果の評価尺度の一つであるF値（精度と再現率の調和平均）を算出したところ、96.9%という高いスコアであった。なお、行列内の各数値は、以下の値を示すものである。

左上（真陽性：TP）：ガンであり、ガンと識別された細胞

右上（偽陽性：FP）：HGDであるが、ガンと識別された細胞

左下（偽陰性：FN）：ガンであるが、HGDと識別された細胞

右下（真陰性：TN）：HGDであり、HGDと識別された細胞

- [0073] 以上の本実施例3によれば、教師あり学習としてのロジスティック回帰は、潰瘍のガン化状態における4ステージのうち、ガン細胞および高度異形成

(HGD)を高精度に識別できることが示された。

実施例 4

[0074] 本実施例4では、教師あり学習の一つであるサポートベクターマシンによって解析した場合の解析精度と、教師なし学習の一つであるクラスター分析によって予め解析対象を絞り込んでからサポートベクターマシンによって解析した場合の解析精度とを比較する実験を行った。

[0075] 具体的には、潰瘍性大腸炎に関連する検体を用意し、サポートベクターマシンのみによって、潰瘍における、非ガン細胞、軽度異形成(LGD)、高度異形成(HGD)およびガン細胞の4ステージを識別した。また、同様の検体について、クラスター分析によって予め非ガン細胞を識別した後、サポートベクターマシンによって、残りの軽度異形成(LGD)、高度異形成(HGD)およびガン細胞の3ステージを識別した。その結果を図8に示す。

[0076] 図8(a)に示すように、サポートベクターマシンのみによって解析した場合、ガン細胞、軽度異形成(LGD)、高度異形成(HGD)および非ガン細胞の4ステージについての識別率は、それぞれ92.2%、78.8%、55.6%および100.0%であった。一方、図8(b)に示すように、クラスター分析とサポートベクターマシンとを組み合わせる解析した場合、ガン細胞、軽度異形成(LGD)および高度異形成(HGD)の4ステージについての識別率は、それぞれ90.0%、71.0%および86.6%であった。

[0077] すなわち、クラスター分析とサポートベクターマシンとを組み合わせる解析した場合、軽度異形成(LGD)の識別率は、7.8%向上し、高度異形成(HGD)の識別率は15.4%向上した。

[0078] 以上の本実施例4によれば、潰瘍のガン化状態の識別は、変質状態識別部57による教師あり学習のみによって行うよりも、変質状態大別部56による教師なし学習と組み合わせる方が、識別精度が向上することが示された。

実施例 5

[0079] 本実施例5では、上述した本実施形態において、染色影響除去部54によ

って除去すべき、解析対象部位の染色濃度によって変化するスペクトル成分を特定する実験を行った。

[0080] 具体的には、まず、ガンの表層上皮性腫瘍でホルマリン固定したものからパラフィン包埋切片を作製し、ヘマトキシリン・エオジン染色を施した病理標本を用意した。なお、ヘマトキシリン・エオジン染色では、核内のクロマチンが濃青紫色に染色される。このため、細胞核内のクロマチン濃度によって光の透過率が変化し、染色液の吸収スペクトルに依存したスペクトル変化がみられることとなる。

[0081] つぎに、実施例1と同様のハイパースペクトルカメラ10を使用し、倍率を200倍に固定して1つの標本から平均3枚のハイパースペクトルイメージを撮影した。なお、ガン細胞では、細胞核が肥大化し細胞質を確認することが困難になるため、1枚から10個の細胞核について分析し、1つの細胞核ごとの平均スペクトルを使用した。

[0082] つづいて、ハイパースペクトルイメージにおける染色濃度を目視で3段階に分け、染色濃度が低いものから、染色レベル1、染色レベル2、染色レベル3とした。そして、独立成分の数を3つとして、ハイパースペクトルデータの独立成分分析を行った。その結果を図9に示す。なお、分析に際しては、445nm～745nmの波長間を5nmのバンドで区切って計測された反射強度を使用した。

[0083] 図9に示すように、3つの独立成分 (Signal1, Signal2, Signal3) のうち、特にSignal1で示す独立成分は、540nm～700nmの波長範囲において段階的に変化していることが確認された。一方、その他の独立成分 (Signal2, Signal3) は、全ての染色レベルにおいてほぼ共通していることが確認された。よって、Signal1で示す独立成分は、染色濃度によって変化するスペクトル成分といえる。

[0084] 以上の本実施例5によれば、ハイパースペクトルデータを独立成分分析することにより、解析対象部位の染色濃度によって変化するスペクトル成分を抽出できることが示された。

[0085] なお、本発明に係る生体組織解析装置 1、生体組織解析プログラム 1 a および生体組織解析方法は、前述した実施形態に限定されるものではなく、適宜変更することができる。

[0086] 例えば、上述した本実施形態では、変質状態大別部 5 6 による教師なし学習として、クラスター分析が採用されているが、これに限定されるものではない。例えば、教師なし学習として主成分分析を採用し、当該主成分分析によって次元を削減したハイパースペクトルデータを線形判別分析によって解析してもよい。

[0087] また、上述した本実施形態では、生体組織の変質状態として、潰瘍やポリープにおける細胞核のガン化状態を識別しているが、具体的には、以下に示すような生体組織の変質状態を識別することが可能である。

(1) 膵管上皮細胞における、正常細胞、膵腺房細胞癌 (A C C : Acinar Cell Carcinoma of pancreatic body)、神経内分泌腫瘍 (N E T : Neuroendocrine Tumor)、膵嚢胞性腫瘍 (S P N : Solid Pseudopapillary Neoplasm)

、膵ガン細胞

(2) 前立腺細胞における、萎縮細胞、前立腺ガン細胞、前立腺肥大細胞

(3) 子宮細胞における、軽度異形成 (C I N 1)、中等度異形成 (C I N 2)、高度異形成・上皮内ガン (C I N 3)

(4) 肺細胞における、肺扁平上皮ガン、肺腺ガン

(5) 乳管細胞における、浸潤性乳管ガンの 3 グレード、非浸潤性乳管ガンの 3 グレード

[0088] さらに、上述した本実施形態では、データ数削減部 5 5 が主成分分析によって解析用データの次元を削減しているが、この手法に限定されるものではなく、ニューラルネットワークを使用した次元圧縮のためのアルゴリズムであるオートエンコーダ (自己符号化器) を使用してもよい。あるいは、教師データや解析対象データの前処理として、教師なし学習の 1 つであるクラスター分析を用いて複数のクラスターを形成し、その一部をノイズクラスターとして除去してもよい。

[0089] また、上述した本実施形態では、生体組織を計測する際に、ヘマトキシリン・エオシン染色を施しているが、これに限られるものではなく、細胞レベルでの識別可能性を有するものであれば、抗原抗体反応（免疫反応）を可視化するために用いられる免疫染色等を用いてもよい。

[0090] さらに、上述した本実施形態では、顕微鏡の接眼レンズにハイパースペクトルカメラ 10 を当接させて生体組織を計測しているが、この構成に限定されるものではなく、三眼顕微鏡の接眼レンズにハイパースペクトルカメラ 10 を当接させて生体組織を計測してもよい。

符号の説明

- [0091]
- 1 生体組織解析装置
 - 1 a 生体組織解析プログラム
 - 2 表示手段
 - 3 入力手段
 - 4 記憶手段
 - 5 演算処理手段
 - 10 ハイパースペクトルカメラ
 - 41 プログラム記憶部
 - 42 教師データ記憶部
 - 51 ハイパースペクトルデータ取得部
 - 52 解析対象部位抽出部
 - 53 光源影響除去部
 - 54 染色影響除去部
 - 55 データ数削減部
 - 56 変質状態大別部
 - 57 変質状態識別部