

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6910068号
(P6910068)

(45) 発行日 令和3年7月28日(2021.7.28)

(24) 登録日 令和3年7月8日(2021.7.8)

(51) Int.Cl.	F I
GO6T 7/00 (2017.01)	GO6T 7/00 612
GO1N 33/48 (2006.01)	GO1N 33/48 P
GO1N 33/483 (2006.01)	GO1N 33/48 M
	GO1N 33/483 C

請求項の数 24 (全 28 頁)

(21) 出願番号	特願2017-563619 (P2017-563619)	(73) 特許権者	506147560
(86) (22) 出願日	平成28年6月10日 (2016. 6. 10)		ユニバーシティ オブ ピッツバーグーオブザ コモンウェルス システム オブ ハイヤー エデュケーション
(65) 公表番号	特表2018-525707 (P2018-525707A)		UNIVERSITY OF PITTS BURGH OF THE COMMON WEALTH SYSTEM OF HI GHER EDUCATION
(43) 公表日	平成30年9月6日 (2018. 9. 6)		アメリカ合衆国 15260 ペンシルベ ニア, ピッツバーグ サッカレー アベニ ュー 130, ガードナー スティール カンファレンス センター ファーストフ ロア
(86) 国際出願番号	PCT/US2016/036825		
(87) 国際公開番号	W02016/201186	(74) 代理人	110000523
(87) 国際公開日	平成28年12月15日 (2016. 12. 15)		アクシス国際特許業務法人
審査請求日	令和1年6月10日 (2019. 6. 10)		最終頁に続く
(31) 優先権主張番号	62/174, 187		
(32) 優先日	平成27年6月11日 (2015. 6. 11)		
(33) 優先権主張国・地域又は機関	米国 (US)		

(54) 【発明の名称】ヘマトキシリン・エオシン (H&E) 染色組織画像における関心領域を調べて、多重化/高多重化蛍光組織画像で腫瘍内細胞空間的不均一性を定量化するシステム及び方法

(57) 【特許請求の範囲】

【請求項 1】

染色組織画像において関心領域を特定する方法において、

前記染色組織画像を表す画像データを受け取る工程であって、前記画像データが、染色された組織画像を表すRGB空間内のRGBデータを反対色空間に変換して反対色空間データを生成して、前記反対色空間データを正規化することにより得られた、前記反対色空間における正規化画像データである、工程と、

受け取った画像データに基づいて、前記画像データにおける隣接ピクセル間の統計的関連性を示す相互情報量データを決定することによって前記染色組織画像について局所的空間的統計を定量化する工程と、

定量化された局所的空間的統計に基づいて、決定された相互情報量データ及びグラフに基づくスペクトル・セグメント化アルゴリズムを使用して、前記染色組織画像内の組織学的構造を特定し、境界を検出する工程と、

前記染色された組織画像を表す、受け取った画像データと、特定した前記組織学的構造の前記検出された境界とを使用して、特定した前記組織学的構造がセグメント化された、セグメント化染色組織画像を生成する工程と、を含む方法。

【請求項 2】

前記染色された組織画像は、ヘマトキシリン・エオシン (H & E) 染色組織画像であり、

前記RGBデータを反対色空間に変換することは、紫及びピンクのピクセルの色相値が複素色面において最大限に分離されるようにピンク及び紫の色が反対に配置される、選択されたピンク及び紫のピクセルの集まりに基づく直交射影行列を利用することを含み、請求項1に記載の方法。

【請求項3】

前記染色された組織画像は、ヘマトキシリン・エオシン(H&E)染色組織画像であり、

前記RGBデータを反対色空間に変換することは、前記RGBデータをH&E-色相チャンネル、H&E-彩度チャンネル、及びH&E-明度チャンネルに変換することを含み、

前記反対色空間データを正規化することは、前記反対色空間データにおけるピンク、紫、及び白のクラスのピクセルを特定し、前記ピンク、紫、及び白のクラスのピクセルの各々について統計量をマッチングすることを含み、前記ピンク、紫、及び白のクラスのピクセルを特定することは、一変量フォン・ミーゼス分布の混合を用いたH&E-色相チャンネルを使用して色相値をモデル化することを含み、請求項1に記載の方法。

【請求項4】

前記画像データは、前記反対色空間における正規化色相データを含み、前記相互情報量データを決定することは、前記正規化色相データにおける隣接ピクセル間の色相角の同時分布を推定すること、及び前記同時分布の自己相互情報量(PMI)を計算することを含み、前記PMIは前記相互情報量データである、請求項1に記載の方法。

【請求項5】

特定する工程は、前記PMIから親和性関数を作ることと、スペクトラルクラスタリングを使用して、親和性関数に基づいて前記境界を検出することとを含む、請求項4に記載の方法。

【請求項6】

前記染色された組織画像は、ヘマトキシリン・エオシン(H&E)染色組織画像であり、前記セグメント化染色組織画像は、セグメント化H&E染色組織画像である、請求項1、4又は5のいずれかに記載の方法。

【請求項7】

命令を含む1以上のプログラムを格納する非一時的なコンピュータ可読媒体であって、コンピュータによって実行されると、コンピュータに請求項1に記載の方法を実行させる、非一時的なコンピュータ可読媒体。

【請求項8】

染色組織画像において関心領域を特定するコンピュータシステムであって、
処理装置を備えており、当該処理装置は、

前記染色組織画像を表す受け取った画像データに基づいて、前記染色組織画像について局所的空間的統計を定量化する工程であって、前記画像データが、染色された組織画像を表すRGB空間内のRGBデータを反対色空間に変換して反対色空間データを生成して、前記反対色空間データを正規化することにより得られた、前記反対色空間における正規化画像データである工程を行うように構成された定量化コンポーネントであって、前記画像データにおける隣接ピクセル間の統計的関連性を示す相互情報量データを決定することによって前記局所的空間的統計を定量化するように構成された定量化コンポーネントと、

定量化された局所的空間的統計に基づいて、決定された相互情報量データ及びグラフに基づくスペクトル・セグメント化アルゴリズムを使用して、前記染色組織画像内で組織学的構造を特定し、境界を検出する工程を行うように構成された特定コンポーネントと、

前記染色された組織画像を表す、受け取った画像データと、特定した前記組織学的構造の前記検出された境界とを使用して、特定した前記組織学的構造がセグメント化された、セグメント化染色組織画像を生成する工程を行うように構成されたセグメント化組織画像生成コンポーネントと、
を含む、システム。

【請求項9】

10

20

30

40

50

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり

、
前記RGBデータを反対色空間に変換することは、紫及びピンクのピクセルの色相値が複素色面において最大限に分離されるようにピンク及び紫の色が反対に配置される、選択されたピンク及び紫のピクセルの集まりに基づく直交射影行列を利用することを含む、請求項8に記載のシステム。

【請求項10】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり

、
前記RGBデータを反対色空間に変換することは、前記RGBデータをH & E - 色相チャンネル、H & E - 彩度チャンネル、及びH & E - 明度チャンネルに変換することを含み、

前記反対色空間データを正規化することは、前記反対色空間データにおけるピンク、紫、及び白のクラスのピクセルを特定し、前記ピンク、紫、及び白のクラスのピクセルの各々について統計量をマッチングすることを含み、前記ピンク、紫、及び白のクラスのピクセルを特定することは、一変量フォン・ミーゼス分布の混合を用いたH & E - 色相チャンネルを使用して色相値をモデル化することを含む、請求項8に記載のシステム。

【請求項11】

前記画像データは、前記反対色空間における正規化色相データを含み、前記相互情報量データを決定することは、前記正規化色相データにおける隣接ピクセル間の色相角の同時分布を推定すること、及び前記同時分布の自己相互情報量（PMI）を計算する工程を含み、前記PMIは前記相互情報量データである、請求項8に記載のシステム。

【請求項12】

特定する工程は、前記PMIから親和性関数を作ることと、スペクトラルクラスタリングを使用して、親和性関数に基づいて前記境界を検出することとを含む、請求項11に記載のシステム。

【請求項13】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、前記セグメント化染色組織画像は、セグメント化H & E染色組織画像である、請求項8、11又は12のいずれかに記載のシステム。

【請求項14】

染色組織画像において関心領域を特定する方法において、

前記染色組織画像を表す画像データを受け取る工程であって、前記画像データが正規化RGBデータであり、前記正規化RGBデータが、染色された組織画像を表すRGB空間内のRGBデータを反対色空間に変換して反対色空間データを生成して、前記反対色空間データを正規化して、正規化反対色空間データを生成して、前記正規化反対色空間データをRGB空間に変換することにより得られた、正規化データである、工程と、

前記正規化RGBデータから決定された核間距離分布に基づいて、前記染色組織画像について局所的空間的統計を定量化する工程と、

定量化された局所的空間的統計に基づいて、前記染色組織画像内の組織学的構造を特定し、境界を検出する工程と、

前記染色された組織画像を表す、受け取った画像データと、特定した前記組織学的構造の前記検出された境界を使用して、特定した前記組織学的構造がセグメント化された、セグメント化染色組織画像を生成する工程と、
を含む方法。

【請求項15】

定量化する工程は、前記正規化RGBデータから推定核位置をスーパーピクセル形式で特定することと、各スーパーピクセルとそれに最も近い幾つかの隣接したものとの間の点別距離に基づいてスーパーピクセル・グラフを作ることと、前記スーパーピクセルを標識セグメントへとクラスター化することとを含む、

特定する工程は、前記標識セグメントを前記組織学的構造にマージすることを含む、請

10

20

30

40

50

求項 1 4 に記載の方法。

【請求項 1 6】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、前記セグメント化染色組織画像は、セグメント化 H & E 染色組織画像である、請求項 1 4 又は 1 5 に記載の方法。

【請求項 1 7】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、
前記 RGB データを反対色空間に変換することは、紫及びピンクのピクセルの色相値が複素色面において最大限に分離されるようにピンク及び紫の色が反対に配置される、選択されたピンク及び紫のピクセルの集まりに基づく直交射影行列を利用することを含み、請求項 1 4 に記載の方法。

【請求項 1 8】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、
前記 RGB データを反対色空間に変換することは、前記 RGB データを H & E - 色相チャンネル、H & E - 彩度チャンネル、及び H & E - 明度チャンネルに変換することを含み、
前記反対色空間データを正規化することは、前記反対色空間データにおけるピンク、紫、及び白のクラスのピクセルを特定し、前記ピンク、紫、及び白のクラスのピクセルの各々について統計量をマッチングすることを含み、前記ピンク、紫、及び白のクラスのピクセルを特定することは、一変量フォン・ミーゼス分布の混合を用いた H & E - 色相チャンネルを使用して色相値をモデル化することを含み、請求項 1 4 に記載の方法。

【請求項 1 9】

命令を含む 1 以上のプログラムを格納する非一時的なコンピュータ可読媒体であって、コンピュータによって実行されると、コンピュータに請求項 1 4 に記載の方法を実行させる、非一時的なコンピュータ可読媒体。

【請求項 2 0】

染色組織画像において関心領域を特定するコンピュータシステムであって、
処理装置を備えており、当該処理装置は、
前記染色組織画像を表す受け取った画像データに基づいて、前記染色組織画像について局所的空間的統計を定量化する工程であって、前記画像データが正規化 RGB データであり、前記正規化 RGB データが、染色された組織画像を表す RGB 空間内の RGB データを反対色空間に変換して反対色空間データを生成して、前記反対色空間データを正規化して、正規化反対色空間データを生成して、前記正規化反対色空間データを RGB 空間に変換することにより得られた、正規化データである工程を行うように構成された定量化コンポーネントであって、前記正規化 RGB データから決定された核間距離分布に基づいて、前記局所的空間的統計を定量化するように構成された定量化コンポーネントと、

定量化された局所的空間的統計に基づいて、前記染色組織画像内の組織学的構造を特定し、境界を検出する工程を行うように構成された特定コンポーネントと、

前記染色された組織画像を表す、受け取った画像データと、特定した前記組織学的構造の前記検出された境界を使用して、特定した前記組織学的構造がセグメント化された、セグメント化染色組織画像を生成する工程を行うように構成されたセグメント化組織画像生成コンポーネントと、
を含む、システム。

【請求項 2 1】

定量化する工程は、前記正規化 RGB データから推定核位置をスーパーピクセル形式で特定することと、各スーパーピクセルとそれに最も近い幾つかの隣接したものととの間の点別距離に基づいてスーパーピクセル・グラフを作ることと、前記スーパーピクセルを標識セグメントへとクラスター化することとを含み、

前記特定コンポーネントは、前記標識セグメントを前記組織学的構造にマージすること

10

20

30

40

50

により特定するように構成されている、請求項 20 に記載のシステム。

【請求項 22】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、前記セグメント化染色組織画像は、セグメント化 H & E 染色組織画像である、請求項 20 又は 21 に記載のシステム。

【請求項 23】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、
前記 RGB データを反対色空間に変換することは、紫及びピンクのピクセルの色相値が複素色面において最大限に分離されるようにピンク及び紫の色が反対に配置される、選択されたピンク及び紫のピクセルの集まりに基づく直交射影行列を利用することを含み、請求項 20 に記載のシステム。

【請求項 24】

前記染色された組織画像は、ヘマトキシリン・エオシン（H & E）染色組織画像であり、
前記 RGB データを反対色空間に変換することは、前記 RGB データを H & E - 色相チャンネル、H & E - 彩度チャンネル、及び H & E - 明度チャンネルに変換することを含み、
前記反対色空間データを正規化することは、前記反対色空間データにおけるピンク、紫、及び白のクラスのピクセルを特定し、前記ピンク、紫、及び白のクラスのピクセルの各々について統計量をマッチングすることを含み、前記ピンク、紫、及び白のクラスのピクセルを特定することは、一変量フォン・ミーゼス分布の混合を用いた H & E - 色相チャンネルを使用して色相値をモデル化することを含み、請求項 20 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

[関連出願の相互参照]

本願は、米国特許法第 119 条 e 項の下、2015 年 6 月 11 日に出願された、発明の名称が「A Common Framework for Finding Regions of Interest in Hematoxylin and Eosin (H&E) Stained Tissue Images and Quantifying Intratumor Cellular Spatial Heterogeneity in Multiplexed/Hyperplexed Fluorescence Tissue Images」である米国仮出願第 62/174,187 号に基づく優先権を主張しており、その内容は引用を以て本明細書に組込まれる。

【0002】

本発明は、デジタルパソロジー（digital pathology）に関しており、特に、ヘマトキシリン・エオシン（H & E）で染色された組織画像中の関心領域を調べるための、そして、多重化 / 高多重化（multiplexed/hyperplexed）蛍光組織画像における腫瘍内細胞空間的不均一性を定量化するための共通のフレームワークに関する。

【背景技術】

【0003】

デジタルパソロジーとは、組織学的に染色された組織試料の取得、保存及び表示を言い、最初に、セカンドオピニオン遠隔病理診断、免疫染色判断や術中遠隔病理診断のようなニッチな用途で注目を集めつつある。通常、3 ~ 50 個のスライドで構成される大量の患者データが生検標本から作られて、顕微鏡下で病理医により目視で評価されるが、デジタル技術を用いると高解像度モニターで検分される。手作業が含まれるために、現在のワークフローの実務は、時間がかかり、間違いやすく、主観的である。

【0004】

癌は不均一性疾患（heterogeneous disease）である。ヘマトキシリン・エオシン（H & E）で染色された組織画像は、不均一性は、上皮内癌、浸潤癌、脂肪組織、血管及び正

常な管のような種々の組織学的構造の存在で特徴付けられている。腫瘍内不均一性及び腫瘍間不均一性の両方に対するプレジジョン・メディシンアプローチの1つは、生検組織の配列を決定し、そして、具体的には各患者について疾患関連ゲノム・シグネチャーのパネルと単一の患者の腫瘍の明確な領域とを特定することである。しかしながら、ゲノム/エピゲノム・プロファイリングには、組織の粉碎又は腫瘍から複数のコア試料の取出しの何れかをする必要があり、時間がかかり、不均一状態を理解するには解像度が低い。種々の組織学的構造間の空間的相互関係は、疾患の兆候を示す。例えば、(浸潤)血管に成長する腫瘍巣は、転移のリスクが高まっていることを示唆する。従って、組織学的構造の正確なセグメント化は、空間的相互関係マップの構築を助けることで、ディープ・モレキュラー・プロファイリングのためにタンパク質、DNA及びRNAバイオマーカーを組み合わせたプレジジョン・メディシン調査を容易にする。更に、この空間的相互関係マップは、病理医には探索ツールとして、又は、更なる分子プロファイリングには顕微解剖へのガイドとして役に立ち得る。

【0005】

正常な管や上皮内癌のような構造は明確な境界を有するが、浸潤癌やストロマのような他の多くのものは明確な境界を持たないために、組織学的構造セグメント化はかなり困難な仕事となる。また、構造形態は、組織源(例えば、胸対肺)と、組織標本手法及び染色手法とに大きく依存する。歴史的には、核は全てのより高いレベルの組織構造に関してブロックを構築しているため、生物学的画像分析の研究報告は、核をセグメント化することに注目してきた。より最近の方法は、核-内腔関連性(nuclei-lumen association)、領域成長、マルコフ確率場と組み合わせた領域ベース動的輪郭、及び深層学習に基づくアプローチを用いた、前立腺及び胸部の組織画像における腺のような他の組織学的構造のセグメント化にまで広がっている。他の幾つかのアプローチは、症状特異エクストラクタ(extractor)及び臓器特異エクストラクタを設計して、MITOS(有糸分裂像)及びGlaS(腺)のような一般に利用可能なデータセットの分析を促進する。例えば、代表的な腺のセグメント化の方法は、最初に内腔を特定する工程と、次いで細胞の周囲上皮層をサーチする工程とを含んでいる。しかしながら、この方法は、導管腔が腫瘍細胞により完全に満たされている非浸潤性乳癌の場合には使えない。基本的な数学的基礎理論は、H&E画像の管理されたセグメント化を進展させてきたが、その基礎理論は、2,3以上の例を超えて試験されてこなかった。

【0006】

さらに、多くの悪性腫瘍においては、分子細胞不均一性は、異なる患者の腫瘍の間で、1人の患者の腫瘍の異なる部位間で、及び単一の腫瘍内において、際立った特徴となる。腫瘍内不均一性は、腫瘍微小環境(TME)を含む、表現型的に異なる癌細胞クローン亜集団と他の細胞型とに影響を及ぼす。これらの癌細胞クローン亜集団及び他の細胞型には、局所的な骨髄由来ストロマ的幹細胞及び前駆細胞と、腫瘍促進又は腫瘍致死の何れかである免疫炎症性細胞のサブクラスと、癌関連線維芽細胞と、内皮細胞と、周皮細胞とがある。TMEは、発展する生態系と見ることができ、ここでは、癌細胞がこれら他の細胞型との異型相互関係にかかわり、利用可能な資源を使用して増殖し、生き延びる。この考えに則って、TME内の細胞型間の空間的相互関係(即ち、空間的不均一性)は、疾患進行及び治療抵抗の主な駆動機構の1つと思われる。従って、TME内の空間的不均一性を明らかにして、特定の疾患サブタイプを正しく決定し、個々の患者の最適な治療コースを特定することは是非とも必要なことである。

【0007】

今日まで、腫瘍内不均一性は、3つの主要なアプローチを使用して調査されてきた。第1のアプローチは、腫瘍の特定領域からコア試料を採取し、集団平均を測定する。試料の不均一性は、全エクソーム解析、エピジェネティクス、プロテオミクス、及びメタボロミクスなどの複数の技術を使用して、腫瘍内の複数のコアを分析することにより測定される。第2のアプローチは、組織からの細胞の分離後に、上記の方法、RNAseq、画像化又はフローサイトメトリーを使用する「単一細胞分析」を含んでいる。第3のアプローチ

10

20

30

40

50

は、光学顕微鏡画像化の空間分解能を利用して空間コンテキストを維持し、分子特異的レベルと組み合わせられて生体内の細胞のバイオマーカーを測定するものである。

【 0 0 0 8 】

光学顕微鏡画像化を利用する空間分析によれば、細胞レベル及び細胞下レベルで、大領域の組織セクション及び / 又は複数の腫瘍マイクロアレイセクションの分析が容易になる。例えば、細胞下レベルの分解能により、特定のバイオマーカーの活性化状態（例えば、転写因子の核への転移）の特定が可能になる。加えて、質量分析画像化における近年の発展によれば、多くの細胞構成要素の組織セクションに亘る測定を光学顕微鏡よりも低い分解能で行うことができる。

【 0 0 0 9 】

透過光及び蛍光発光を含む腫瘍内の細胞バイオマーカー発現レベルを特徴付けるために、数種の光学顕微鏡画像化プラットフォームが開発されている。DNA、RNA及びタンパク質バイオマーカー、通常同じ試料内に1～7個までの蛍光標識された（多重化蛍光発光として知られる）バイオマーカーに基づく大領域組織セクション及び組織マイクロアレイ（TMA）の画像から、蛍光発光に基づく多変量情報が取得されている。現在、幾つかの市販のプラットフォームを使用して、取得し、処理し、セグメント化し、組織試料のバイオマーカーの信号レベルの幾つかの基本的な分析を行うことができる。近年、標識化、画像化及び蛍光発光のクエンチの反復サイクルにおいて、60種までの蛍光標識された抗体と、数種のDNA又はRNA交雑プローブとを取得可能なプラットフォームが実証されている。また、現在、組織セクション及びTMAにおける細胞外構成要素だけでなく、特異的細胞型の位置、細胞活性化の状態、細胞バイオマーカー発現レベル及び部位を「マップする」ことも可能である。

【 0 0 1 0 】

主たる課題は、バイオマーカーのパネルに基づいてTME内の重要な空間的關係（相互關係及びその欠落）を定量化できるアルゴリズムを開発することである。組織セクションの不均一性の測定における初期の試みは、シャノンエントロピーやラオの二次エントロピーのような生態学的研究からのダイバーシティメトリック（diversity metrics）を適用した。しかしながら、これらの方法は、多重化（7種までのバイオマーカー）又は高多重化（7種を超えるバイオマーカー）免疫蛍光（IF）データには適切でなかった。高次元データを説明する他の方法には、各バイオマーカーを「オン」又は「オフ」のみにする精巧な細胞表現型方法がないかもしれない。さらに、これらの方法が、不均一性スコアにおいてバイオマーカー・パターン間に空間的關係を組み入れることはほとんどない。実際、TMEの空間組織は、癌細胞及び非癌細胞から選択されたバイオマーカーの発現レベルに加えて、重要な診断バイオマーカーであると仮定されてきた。

【 0 0 1 1 】

他の不均一性特徴化方法は、（i）ネットワークベースのアプローチの使用や多重化の利点を得ることなく、関心領域のサンプリングにより空間情報を組み入れており、（ii）非線形関連性や空間情報を考慮せずに、多重化 / 高多重化IFデータでバイオマーカー間の線形關係を分析しており、（iii）腫瘍内の空間構成の特徴付けをすることなく、多重化細胞表現型の関連性を特徴付けてきた。さらに、他の方法の大半は、腫瘍内不均一性を単一のスコアとして報告しており、従って、TMEの2つの空間的に異なる構成を誤って同じスコアにマッピングすることもあり得る。

【 0 0 1 2 】

故に、多重化 / 高多重化蛍光組織画像における、H & E画像のセグメント化及び腫瘍内細胞空間的不均一性の定量化の分野には、改良の余地がある。

【 発明の概要 】

【 0 0 1 3 】

ある実施形態では、H & E染色組織画像において関心領域を特定する方法が提供される。その方法は、H & E染色組織画像を表す画像データを受け取る工程と、受け取った画像データに基づいて、H & E染色組織画像について局所的空間的統計を定量化する工程と、

10

20

30

40

50

局所的空間的統計に基づいて、H & E 染色組織画像内の組織学的構造を特定する工程と、受け取った画像データ及び特定した組織学的構造を使用して、セグメント化 H & E 染色組織画像を生成する工程とを含んでいる。

【0014】

別の実施形態では、コンピュータによって実行された場合、コンピュータに上記の方法を行わせる命令を含む 1 以上のプログラムを格納する非一時的なコンピュータ可読媒体が提供される。

【0015】

さらに別の実施形態では、H & E 染色組織画像において関心領域を特定するコンピュータシステムが提供される。そのシステムは処理装置を含んでおり、当該処理装置は、(i) H & E 染色組織画像を表す受け取った画像データに基づいて、H & E 染色組織画像に関する局所的空間的統計を定量化するように構成された定量化コンポーネントと、(i i) 局所的空間的統計に基づいて、H & E 染色組織画像の組織学的構造を特定するように構成された特定コンポーネントと、(i i i) 受け取った画像データ及び特定した組織学的構造を使用して、セグメント化 H & E 染色組織画像を生成するように構成されたセグメント化組織画像生成コンポーネントとを含む。

10

【0016】

さらに別の実施形態では、蛍光組織画像にて腫瘍内細胞空間的不均一性を定量化する方法が提供される。その方法は、幾つかの蛍光組織画像を表す画像データを受け取る工程と、受け取った画像データで細胞セグメント化を行い、幾つかの蛍光組織画像にて幾つかの細胞を特定する工程と、各細胞を複数の所定のバイオマーカー強度パターンの 1 つに割り当てる工程と、割り当てられた所定のバイオマーカー強度パターンに基づいて、幾つかの蛍光組織画像の空間的統計を定量化する工程と、定量化された空間的統計の視覚表示を作成する工程とを含む。

20

【0017】

別の実施形態では、コンピュータによって実行された場合、当該コンピュータに上記の方法を行わせる命令を含む 1 以上のプログラムを格納する非一時的なコンピュータ可読媒体が提供される。

【0018】

さらに別の実施形態では、蛍光組織画像にて腫瘍内細胞空間的不均一性を定量化するコンピュータシステムが提供される。そのシステムは処理装置を含んでおり、当該処理装置は、(i) 幾つかの蛍光組織画像を表す画像データで細胞セグメント化を行い、幾つかの蛍光組織画像における幾つかの細胞を特定するように構成された細胞セグメント化コンポーネントと、(i i) 各細胞を複数の所定のバイオマーカー強度パターンの 1 つに割り当てるように構成された割当てコンポーネントと、(i i i) 割り当てられた所定のバイオマーカー強度パターンに基づいて、前記幾つかの蛍光組織画像の空間的統計を定量化する定量化コンポーネントと、(i v) 定量化された空間的統計の視覚表示を作成する視覚表示作成コンポーネントとを含む。

30

【図面の簡単な説明】

【0019】

【図 1】図 1 は、試料の H & E 染色画像である。

【図 2】図 2 は、H & E 色相空間に変換した図 1 の H & E 画像であって、ヒートマップ (左) 及び角度ヒストグラム (右) として示されている。

【図 3】図 3 は、本開示概念の一態様による第 1 のセグメント化方法を示すフローチャートである。

【図 4】図 4 は、本開示概念の一態様による第 2 のセグメント化方法を示すフローチャートである。

【図 5】図 5 は、本開示概念の一態様による一組の優性バイオマーカー強度パターンを構築する方法を示すフローチャートである。

【図 6】図 6 は、バイオマーカー強度分布グラフを示しており、縦軸は、本明細書で記載

40

50

される方法で使用される例示的なバイオマーカーについて2つの異なるレジームL1及びL2を示している。

【図7】図7は、図6に示す2つのバイオマーカー強度レジームL1及びL2について、それぞれ得られたパターン辞書を示す。

【図8】図8は、L1レジームの細胞集団からのバイオマーカー・データの3D表示を示しており、各細胞は、辞書中の唯一のパターンに属するようにフェノタイプ化されており（phenotyped）、別々の色で示されている。

【図9】図9は、バイオマーカー強度レジームL1及びL2のパターン辞書の最適寸法の決定を示す。

【図10】図10は、レジームL1及びL2のパターン辞書の統合を示す。

10

【図11】図11は、例示的な実施形態による多重化/高多重化蛍光組織画像の空間的不均一性を定量化する方法のステップを示すフローチャートである。

【図12】図12は、例示的な実施形態により使用される所定の優性バイオマーカー強度パターンの模式的表示を示す。

【図13】図13は、例示的な実施形態による細胞空間依存画像を示す。

【図14】図14は、例示的な実施形態によるPMIマップを示す。

【図15】図15は、代表的な細胞空間依存画像及びPMIマップを示す。

【図16】図16は、代表的な細胞空間依存画像及びPMIマップを示す。

【図17】図17は、代表的な細胞空間依存画像及びPMIマップを示す。

【図18】図18は、本明細書に記載したH&E画像をセグメント化する手順を実施するシステムの概略図である。

20

【図19】図19は、本明細書に記載する腫瘍内空間的不均一性を定量化する手順を実施するシステムの概略図である。

【発明を実施するための形態】

【0020】

本明細書で使用されるように、単数形「1つ」及び「その」は、別段の明確な記載がない限り、複数形態も述べている。

【0021】

本明細書で使用されるように、2つ以上の部分又は要素を「結合する」という記載は、1つ以上の中間部分又は要素を介して、それらの部分が、直接又は間接的に、繋げられる又は共に動作することを意味する。

30

【0022】

本明細書で使用されるように、用語「幾つか」は、1又は2以上の整数（即ち、複数）を意味する。

【0023】

本明細書で使用されるように、用語「コンポーネント」及び用語「システム」は、ハードウェア、ハードウェアとソフトウェアの組合せ、ソフトウェア、又は実行されているソフトウェアのようなコンピュータに関連したエンティティを意味する。例えば、コンポーネントは、プロセッサ上で作動するプロセス、プロセッサ、オブジェクト、実行可能なファイル、実行の脈絡、プログラム及び/又はコンピュータであり得るが、これらに限定されない。例としては、サーバで作動するアプリケーション及びサーバはどちらもコンポーネントであり得る。1つ以上のコンポーネントが、実行プロセス及び/又はスレッド内に存在してよく、1つのコンポーネントが、1つのコンピュータに局在してよく及び/又は2つ以上のコンピュータに配置されてよい。本明細書では、ユーザに情報を表示する幾つかの方法は、幾つかの図又はグラフについてスクリーン又はスクリーンショットとして示されて、説明されているが、当業者は、種々の他の代替法を使用できることを理解するだろう。スクリーン又はスクリーンショットは、表示記述として、グラフィカル・ユーザ・インターフェースとして、又はスクリーン（例えば、パーソナル・コンピュータ、PDA、携帯電話、又は他の適切なデバイスのどれか）に情報を描写するその他の方法によって保存及び/又は転送され、ここで、ページに表示されるレイアウト及び情報又はコンテン

40

50

ツは、メモリ、データベース、又は別の格納場所に保存される。また、スクリーン又はスクリーンショットは、必要に応じて、プリントされてもよい。

【 0 0 2 4 】

本明細書で使用されるように、用語「スーパーピクセル」は、同様な画像統計を有するピクセルのコヒーレントパッチ又はグループを意味する。

【 0 0 2 5 】

本明細書で使用する方向を示す表現は、例えば、上、底、左、右、上部、下部、前、後、及びこれらの派生語であってこれらに限定されないが、図に示された要素の方向を示しており、明示的に述べられていない限り、特許請求の範囲を限定するものではない。

【 0 0 2 6 】

< A . H & E 染色組織画像のセグメント化 >

開示されている概念の第 1 態様は、特に、浸潤癌のような胸部組織、上皮内癌、非定型及び正常管、脂肪組織及び / 又はリンパ球のような組織の H & E 染色画像における組織学的構造のセグメント化でのデジタル・パソロジー・システムの機能及び動作の改良（例えば、処理能力の改良）に焦点を当てている。本発明者らは、空間画像統計は、組織学的構造の広範なクラスをセグメント化するための判別可能なフィンガープリントを与えると仮定した。本明細書においてより詳しく記載する本開示概念のこの態様は、2つのグラフ理論的セグメント化方法を提供しており、各々は、局所的空間的統計の特徴付けに依存している。

【 0 0 2 7 】

第 1 の方法では、グラフの各ノードは画像のピクセルに対応しており、エッジは 2 つのノードが同じグループに属する強さに対応している。エッジの強さは、H & E 画像におけるピンクと紫の染色の間の分離を強調するように作られた反対色空間において、二変量フォン・ミーゼス混合分布の形で、ペアをなすピクセル統計を測定することによって決定される。スペクトル法がグラフを区分化するのに使用される。第 1 の方法は、はっきりした境界を有する構造（例えば、脂肪組織及び血管）のセグメント化においてより成功することが期待される。

【 0 0 2 8 】

第 2 の方法は、無定型な空間範囲を有する組織学的構造（例えば、腫瘍巣）を抽出するのに好適に設計されている。この策では、仮想の核中心が、H & E 画像における核の空間分布を取得するために作成されるグラフのノードとなる。核間空間距離にデータ駆動閾値（data-driven thresholds）を適用することにより、ネットワークは、均質な画像パッチに区分される。

【 0 0 2 9 】

本明細書に記載する 2 つのセグメント化方法は、2つの共通エレメント、即ち、反対色表示及び外見正規化（appearance normalization）を有しており、これらはそれぞれ、以下に詳細に記載される。セグメント化方法は、これらが画像統計をどのようにして取得し、及び方法を区画するグラフにこれらをどのようにしてはめ込むかで相違する。その方法のこれらの態様は、本明細書で個別に記載される。

【 0 0 3 0 】

公知の典型的な反対色（反対色軸として赤 - 緑、黄 - 青を有する）色相 - 彩度 - 明度（H S V）変換が、H & E からの赤 - 緑 - 青（R G B）画像に適用される場合、ピンクと紫の範囲は、色環の青 - 赤四分円に限定される。本開示概念のこの態様の目的は、下流の空間分析パイプラインがより強固になるように、ピンクと紫の色の間の分離を強調することである。このために、色空間の構築が最適化されて、ピンク及び紫の色が反対に配置される。具体的には、例示的な実施例では、専門家は、ピンク及び紫のピクセルの集まりを選択することを許された。次いで、特異値分解がデータのこの集まりで行われて、サイズ 3 × 3 の直交射影行列が得られた。本開示概念のこの態様は、反対空間 H S V のように、投影座標に特定の解釈をもたらす。具体的には、第 1 の特異ベクトル（非負の値を有するように強制されている）上への投影は、H & E - 輝度値 b を得る。残り 2 つの投影座標 c 2

10

20

30

40

50

及び c_3 は、H & E - 彩度 $s = \sqrt{c_2^2 + c_3^2}$ と、H & E - 色相 $\theta = \tan^{-1}(c_3 / c_2)$ である複素面を形成する。この構成から、紫及びピンクのピクセルの色相値は、複素色面において最大限に分離されると予測される。例として、ここで、試料 H & E 染色画像である図 1 のピンク及び紫のピクセルの平均色相値における角度差は、図 2 に示すように、1 : 7 ラジアンであって、図 1 の H & E 画像を H & E 色相空間に変換したものであって、図 2 は、ヒートマップ (左) 及び角度ヒストグラム (右) を示している。この広がり、標準 H S V 反対色空間から得られる約 0 : 4 ラジアンの値を超えている。彩度が低い場合、色相値は不安定である。これは、複素面の基点にマップされたピクセルに関して正しい ($c_2 > 0, c_3 = 0$)。典型的な H S V 表示においては、全ての白のピクセルは、低い彩度値を持ち、従って不安定な色相角を有するだろう。ここで留意すべきは、脂肪組織、内腔、組織の裂け目、及び収縮のため、白のピクセルは H & E 画像の重要な部分を形成できることである。本開示概念のこの態様の反対色表示では、専門家が選択したピンク / 紫のピクセルの集まりからの回転行列を知得することで、白のピクセルに、より高い彩度値及びより安定な色相角を与えることができる。しかしながら、複素面の基点にマップする、低い彩度値 (例えば、 $< 0 : 0.05$) を有するピクセルの集団があるだろう。この集団は、使用されたサイズ $2K \times 2K$ の H & E 画像については約 0 : 3 % と経験的に推定される。

10

【 0 0 3 1 】

加えて、区分化、染色及び画像化におけるばらつきは、H & E 画像のカラーアピランスの変化となる。従って、例示的な実施形態では、データは正規化される。従来の正規化方法は、非負値行列因子分解のような染色ベクター評価法 (stain vector estimation methods) を利用していた。これらの方法は、幾つか画像の色分布がほとんど紫又はほとんどピンクにひどく傾くので、本開示概念のこの態様については効果がないことが分かった。本発明者らは、2つの画像のカラーアピランスは、これらの色統計量がマッチするならば、類似すると仮定した。しかしながら、ソース画像及びターゲット画像における全ピクセル集団の統計量をマッチングさせることは、意図しないアーチファクトをもたらす可能性がある。例えば、もしソース画像が主としてピンクのピクセル (ストロマ) を有しており、ターゲット画像が主として紫のピクセル (浸潤癌) を有するならば、ソース画像統計量のターゲット画像統計量へのマッチングは、ソース画像中の多くのピンクのピクセルを紫に変えてしまい、ストロマから核へとこれらのピクセルの細胞要素の素性を間違っ

て変化させるだろう。この問題に対処するために、ピンク (エオシン)、紫 (ヘマトキシリン)、及び白 (例えば、脂肪、収縮) である 3つのクラスのピクセルが最初に特定されて、統計量は、これらのクラスの各々について、別々にマッチングされる。3つのクラスを特定するため、H & E 画像は、説明したように、H & E - 色相チャンネル、H & E - 彩度チャンネル、及び H & E - 明度チャンネルに変換される。H & E - 色相空間は角度があり、この空間のピンク、紫及び白のピクセルのクラウドの間に分離が与えられると、色相値は、一変量フォン・ミーゼス分布の混合を用いてモデル化される。角統計の一変量フォン・ミーゼス分布は、線形統計の一変量正規分布の等価相対物である。フォン・ミーゼス分布は、2つのパラメータ、つまり、平均 μ と、濃度パラメータ $\kappa > 0$ とを特徴とし、 $f(x) = \frac{1}{2\pi\kappa} \{ I_0(\kappa \cos(x - \mu)) \}^{-1} \exp(-\kappa \cos(x - \mu))$ (式中、 $I_0(\cdot)$ は、0 次の第 1 種変形ベッセル関数である) で表される。K-変量フォン・ミーゼス分布の混合は、下式で表される。

20

30

40

【 数 1 】

$$\sum_{k=1}^K m_k f_k(x | \mu_k, \kappa_k)$$

ここで、 m_k は、事前確率であり、 μ_k 及び κ_k は、平均パラメータ及び濃度パラメータである。低い彩度値及び不安定な色相角を有するピクセルを明示的に構成するために、均

50

一な角ノイズが、事前確率が約0.3%である追加の混合要素として加えられる。一変量フォン・ミーゼス混合のパラメータを、期待値最大化(EM)アルゴリズムを用いて求めることができる。分布の統計は、モーメントの無限集合を特徴とし得る。しかしながら、例示的な実施形態においては、分析の便宜のため、モーメントは、4次(平均、標準偏差、歪度、尖度)までのみをコンピュータで計算される。各チャンネルにおいて、ソース画像からの各ピクセルクラスのモーメントが、ターゲット画像にマッチされる。例えば、ソース画像における紫のピクセルのモーメントは、全3チャンネル中のターゲット画像における紫のピクセルのモーメントにマッチされる。H & E 反対色空間中の統計量を正規化した後、得られたピクセル値は、上記の回転行列の逆関数を使用して、(正規化RGBデータを生成するために)RGB空間に変換される。

10

【0032】

2つのセグメント化方法の2つの共通エレメント、即ち、反対色表示及びアピアランスの正規化を説明した。次に、各セグメント化方法の残りを詳細に記載する。各セグメント化方法においては、正規化画像データが入力として使われる。具体的には、正規化H & E - 色相データは、第1の方法において入力として使用され、正規化RGBデータは、第2の方法において入力として使用される。

【0033】

第1の方法に関して、正常な胸部組織は、それぞれが細胞の集合体である小領域の管を取り囲む、ピンクで染色された大領域の結合組織(CT)を有する。これらの細胞の核は暗い紫に染色されるであろうが、核からの紫の染色が細胞質に波及し得るので、核を取り巻く細胞質は、ピンク及び紫の混合を示す。統計学的に言えば、これらの核の何れかにあるならば、核を表す紫のピクセルと細胞質を表すピンク - 紫のピクセルとに囲まれると予測されるであろう。これらの細胞が管構造物を構成するならば、各細胞の特定の近隣において、類似の特性を発揮する他の細胞が見つかるはずである。一方、通常は結合組織中に散乱するのが見られるものが線維芽細胞核にあるならば、その近隣で、ほとんどピンクのピクセルを見つけるであろう。管のような構造内の統計的関連性はその境界のところより高いと仮定すれば、結合組織のあちこちに散乱した線維芽細胞を無視して、管をセグメント化できるはずである。

20

【0034】

一変量フォン・ミーゼス分布の混合を使用することで、画像ピクセルを、ピンク、紫及び白のクラスに分離することができるが、腺/管のような組織学的構造の輪郭を描くには、これらの構造は3つのクラスの全てからのピクセルを含むので、不十分である。本開示概念のこの態様では、これらの構造をセグメント化するために、管のような構造内における統計的関連性はその境界にわたるよりも高いと仮定され、この統計的関連性が、二変量フォン・ミーゼス分布の混合を使用して、本開示概念のこの態様によってモデル化される。H & E - 色相は角変数なので、2つの隣接ピクセルからの色相値の同時分布 $P(A, B)$ は、トーラス上にある。この同時密度は、二変量フォン・ミーゼス分布の混合としてモデル化される。H & E - 色相空間におけるピクセルA及びBの値を、それぞれ θ_1 及び θ_2 とする。2つの角変数 $\theta_1 - \theta_2 < \pi$ 及び $\theta_2 - \theta_1 < \pi$ の二変量分布は、 $f_c(\theta_1, \theta_2) = C_c \exp[\mu_1 \cos(\theta_1 - \mu) + \mu_2 \cos(\theta_2 - \mu) - \mu_3 \cos(\theta_1 - \mu - \theta_2 + \mu)]$ となる。ここで、 μ は平均であり、 $\mu_1, \mu_2 > 0$ は、それぞれ、 θ_1 及び θ_2 の濃度であり、 μ_3 は相関係数であり、 C_c は正規化定数である。完全な二変量フォン・ミーゼスモデルは8つのパラメータを有するが、例示的な実施形態では、正の相互関係を有する縮小された5パラメータ・コサイン・モデルを使用する。周辺密度は、 $f_c(\theta) = C_c^2 I_0(\mu_3) \exp\{\mu_2 \cos(\theta - \mu)\}$ である。 μ_3 の値は、分布が単峰性であるか二峰性であるかを示す。具体的には、 $\mu_1 > \mu_3 > 0$ 及び $\mu_2 > \mu_3 > 0$ である場合において、 $\mu_3 < \mu_1 \mu_2 / (\mu_1 + \mu_2)$ であれば、同時密度は単峰性であり、 $\mu_3 > \mu_1 \mu_2 / (\mu_1 + \mu_2)$ であれば二峰性である。

30

40

【0035】

H & E - 色相空間におけるH & E画像の隣接ピクセルの値を考慮する場合、トーラス上

50

塊には最大で6つの可能性、即ち、紫 - 紫、ピンク - ピンク、白 - 白、及び3つの異なる対相互関係がある。この同時分布をモデル化するために、6個の単峰性二変量フォン・ミーゼス分布の混合を使用する。K個の二変量フォン・ミーゼス分布の混合モデルは、下式によってパラメータで表すことができる。

【数2】

$$f_c(\varphi, \psi) = \sum_{i=1}^K m_i f_i(\varphi, \psi | \mu_i, \nu_i, \kappa_{1i}, \kappa_{2i}, \kappa_{3i})$$

10

μ_i 、 ν_i 、 κ_{1i} 及び κ_{2i} の初期値は、画像の全ピクセルに関する一変量フォン・ミーゼスの混合から作成される。濃度パラメータ κ_{1i} 及び κ_{2i} と、相関パラメータ κ_{3i} とは、 f_i に関する単峰性条件を満たす。 κ_{3i} は、楕円形パターン（試料化されたデータで観察される）の歪みを回避するために、-1と1の間の値を有するよう拘束される。上記の制約と合わせて、混合のパラメータは、EMアルゴリズムによって推定される。先に論じた様に、混合モデルの多くても6つの要素が存在するので、明確なモデル選択のステップは、混合モデルに関して保障されない。H & E 画像が3つの基本色、つまり、紫、ピンク及び白の何れか1つを欠く場合、その色に関連した事前確率又はクラスターの混合比は、0に近くなるであろう。

【0036】

20

H & E 反対色空間における隣接ピクセルの色相角間の統計的依存性をモデル化することを考える。同時確率を統計的関連性の尺度として使用する場合、結合組織におけるピンク - ピンクのピクセルペアは、管内部の紫 - 紫のピクセルペア又はCT - 管の境界にわたったピンク - 紫のピクセルペアよりも高い確率を有することが分かるかも知れない。しかしながら、幾つかのH & E 画像ではピンクが過剰であるために、CT - 管境界にわたったピンク - 紫のピクセルペアの組合せは、管内部の紫 - 紫のピクセルペアと同じ又はそれより高い確率を有することもある。ピンク - ピンクのペアは最も高い同時確率を持ち、紫 - 紫ペアは紫 - ピンクペアと同様な同時確率を有することもある。換言すれば、同時確率は正しい境界を感知するには十分ではないかもしれない。これは、相互情報量(MI)を使用して、相対的存在量を修正することで改善できる。MIをコンピュータで計算するために、下式の特徴

【数3】

30

$$\vec{f}_A \text{ 及び } \vec{f}_B$$

(例えば、H & E - 色相角) を有する複数のピクセルペア(A, B)を、閾値未満の距離内にある画像の全ての場所から無作為に選択する。距離dだけ離れた特徴A及びBの同時確率は、 $P(A, B; d)$ で表される。全体的な同時確率は、下式で規定される。

【数4】

40

$$P(A, B) = \frac{1}{Z} \sum_{d=d_0}^{\infty} w(d) p(A, B; d)$$

dの値は、パラメータに依存しており、具体的には、 $d = 2 + 2|r|$ であり、ここで、 $r \sim (0,)$ である。核は、10倍の倍率で直径が約15ピクセルである。セグメント化アルゴリズムは核の集合体をターゲットとするので、サンプルのピクセルペアの間の距離は、少なくとも核の直径をカバーすべきである。故に、は3に設定される。自己

50

相互情報量 (pointwise mutual information) (PMI) は、二変量フォン・ミーゼス分布の混合によってモデル化された同時確率と、一変量フォン・ミーゼス分布の混合によってモデル化された周辺確率 $P(A)$ 及び $P(B)$ とによって計算される。具体的には、下式となる。

【数 5】

$$\text{PMI}_\rho(A, B) = \log \frac{P(A, B)^\rho}{P(A)P(B)}$$

10

例示的な実施形態では、 $\log [P(A, B) / \{P(A)P(B)\}]$ の上限について正規化するために、 $\rho = 2$ とされる。

【0037】

さらに、親和性関数が、PMI から規定されて、2つのピクセルを同じ組織学的構造にグループ化する尤度を示す。エレメント $w_{i, j}$ を有する親和性行列 W は、ピクセル i 及び j の間の類似性を表す。ここで、

【数 6】

$$w_{i, j} = e^{\text{PMI}_\rho(\vec{f}_i, \vec{f}_j)}$$

20

である。親和性関数は、標準スペクトルグラフセグメント化方法に対して入力として使用される。当該方法は、自然画像をセグメント化する現状技術であって、例えば、Arbelaez, P. らの「Contour Detection and Hierarchical Image Segmentation」、IEEE TPAMI、33(5)、898-916(2012)に記載されている。親和性行列 W から、一般化システムの固有対

【数 7】

30

$$\vec{v}, \lambda$$

は、下式とわかる。

【数 8】

$$(D - W)\vec{v} = \lambda D\vec{v}$$

40

主固有ベクトルマップ (小さい固有値) は、潜在的な組織学的構造の境界位置を示す。周知のように、単一固有ベクトルは、複素画像における全ての可能性のある境界を取得できないだろう。従って、通常の実施では、多数の主固有ベクトルの方向空間微分からエッジ強さマップが計算される。後処理ステップが、擬似境界ピクセルを削除するために使用される。

【0038】

図 3 は、詳細に説明した第 1 セグメント化方法をまとめたフローチャートである。その方法は、対象スライドに関する正規化 H & E - 色相データを受け取るステップ 5 で始まる。次に、ステップ 10 で、H & E - 色相データにおける複数の隣接ピクセル間の正規化 H

50

& E - 色相値の同時分布が推定される。次いで、ステップ 15 で、正規化 H & E - 色相データに関する PMI が、推定された同時分布に基づいて計算される。ステップ 20 にて、親和性関数が、計算された PMI から規定される。最後に、ステップ 25 で、対象スライドにおける組織が、親和性関数及びスペクトルグラフセグメント化方法（スペクトラルクラスタリングとしても知られている）を使用してセグメント化される。

【 0 0 3 9 】

第 2 セグメント化方法に関して、局所的空間的統計は、胸部組織における種々の組織学的構造間で変化する。例えば、腺管上皮内癌の細胞の凝集は、それらの境界が互いに極めて近接するように集まる傾向にある。これは、生体内腫瘍が増殖するが、管内部に留まるからである。一方、浸潤癌における上皮細胞は、空間的に遠く離れている。それらはまた、増殖するが、胸部ストロマに自由に侵入できて、管内に制限されない。正常な管の局所統計はより規則正しく、特に正常な上皮（内部）及び筋上皮細胞（外側）は、空洞（内腔）を取り囲む 2 つの層を形成する。

10

【 0 0 4 0 】

脂肪組織では、核は小さくて細胞の片側にある。脂肪組織の大半は、脂肪滴で構成される。本発明の発明者は、異なる組織学的構造は、核間距離（局所統計量）の異なる分布を有すると仮定した。以下に説明するように、本開示概念のこの態様の第 2 セグメント化方法は、この仮説に基づいている。

【 0 0 4 1 】

組織病理学的及び細胞病理学的画像における核のセグメント化は、詳しく研究されている課題である。しかしながら、乳癌における上皮細胞の近接と、有糸核分裂像（分裂細胞）の広まりとは、核の境界を正確に検出することを困難にし、これはヒトの眼でも困難である。この問題を避けるため、第 2 セグメント化方法では、推定核位置をスーパーピクセル形式で特定することで、核を近似的に示し、スーパーピクセルを接続するグラフを構築することで、各スーパーピクセル対に関する隣接・距離情報が得られる。より具体的に言うと、例示的な実施形態では、H & E 画像からスーパーピクセルを作成するために、先ず、ピクセル色が先に記載したように正規化される。次いで、Tosun, A. B. 及び Gunduz - Demir, C. 「Graph Run-length Matrices for Histopathological Image Segmentation」IEEE TMI, 30 (3), 721 - 732 (2011) で提案されているアルゴリズムが実行されて、円状にスーパーピクセルが合わせられる。簡単に言うと、このアルゴリズムは、先ず、k - 平均アルゴリズムを用いて強度に基づいて 3 つのクラスにピクセルをクラスター化する。クラスター中心は、主要素分析を使用して、無作為に選択されたトレーニング画像にわたって決定される。これらの 3 つのクラスは、紫領域、ピンク領域及び白領域を示し、これらは、それぞれ、核領域、ストロマ領域及び内腔 / 白領域に相当する。次に、このアルゴリズムは、円状のスーパーピクセルを、核、ストロマ及び内腔 / 白要素に関するクラスター化ピクセルに合わせる。スーパーピクセル分解後、スーパーピクセルの中心座標に基づいてデローニの三角分割が形成され、各スーパーピクセルの近隣が決定される。各スーパーピクセル・ペアに関する距離情報があれば、グリーディ法でこのグラフを区画し、特定のタイプのセグメントについてマージ規則（merging rules）を適用することによって、組織学的構造の最終的なセグメント化が達成される。これは以下の部分で詳述される。提案方法は、核間距離分布を動機とするが、紫及び白のピクセルクラスの両方からのスーパーピクセルは、管、血管及び脂肪組織のような複合体組織学的構造を占めると考えられる。例えば、正常な管は、白の内腔領域を囲む 2 つの細胞層を形成する紫の核を有する。一方、ストロマ（ピンク）クラスは、バックグラウンドであると考えられ、グラフ分割ステップに含まれない。

20

30

40

【 0 0 4 2 】

より具体的には、各スーパーピクセルは、グラフにおけるノードと考えられ、グラフの接続性は、距離閾値により決定される。各クラスについて、スーパーピクセルの中心とそれに最も近い 15 個の隣りのもの（デローニの三角分割により特定される）の間の対距離

50

(pairwise distance) を計算する。距離閾値 は、距離分布の中央値 () に比例するとして設定される。比例定数は、全データベースに関するアルゴリズムのパフォーマンスを極大化するように設定される。スーパーピクセル・グラフを作った後、貪欲接続要素分析アルゴリズム (greedy connected component analysis) を使用して、スーパーピクセルを標識セグメントへとクラスター化する。例示的な実施形態では、組織領域に関する最も大きな 15 個のセグメントが選択される。例示的な実施形態における組織画像のサイズは $2K \times 2K$ なので、管、腫瘍巣、脂肪滴の一握りだけが、与えられた任意の画像にて予測される。この点で、標識セグメントの 2 つのセットは、紫及び白のスーパーピクセルから得られている。

【 0 0 4 3 】

紫のセグメントと白のセグメントを最終的な組織学的構造へとマージするため、数種の簡単なルールに従い、核クラスターによって形成された重要な構造が見過ごされていないことを確かめる。白のセグメントが、紫のセグメントによって完全にカバーされている場合、紫の全領域は、紫のセグメントのラベルを得る。白のセグメントが紫のセグメントと重なっている場合、重なった領域に関わらず、重なった部分は紫のセグメントのラベルを得、重なっていない部分は白のセグメントのラベルを得る。紫のセグメントが白のセグメントにより完全にカバーされている場合、紫の領域は、紫のセグメントのラベルを取り、残りの白の領域は、白のセグメントのラベルを保持する。これは、血管内に存在する核凝集を見過ごさないためである。紫及び白のセグメントがマージされた後、残りのラベル化されていない領域は、バックグラウンド又はストロマと考えられる。

【 0 0 4 4 】

図 4 は、詳細に記載した第 2 セグメント化方法をまとめたフローチャートである。その方法は、対象スライドに関する正規化 RGB データを受け取るステップ 30 で始まる。次に、ステップ 35 で、RGB データから、推定核位置をスーパーピクセル形式で特定する。次いで、ステップ 40 で、各スーパーピクセルとその最も近い隣接したものの幾つかとの間の対距離に基づいて、スーパーピクセル・グラフを作る。次に、ステップ 45 で、スーパーピクセル・グラフを使用して、スーパーピクセルを標識セグメントにクラスター化又はグループ分けする。最後に、ステップ 50 で、標識セグメントを最終的な組織学的構造へとマージする。次いで、決定された最終的な組織学的構造を、対象スライドをセグメント化 (即ち、対象スライドがセグメント化されている画像を作る) ために使用する。

【 0 0 4 5 】

< B . 腫瘍内空間的不均一性の定量化 >

本明細書でより詳しく説明されるように、本開示概念の別の態様は、デジタル・パソロジー・システムの機能及び動作における改善 (例えば、処理の改善) を提供する。特に、この態様は、腫瘍内空間的不均一性を定量化する方法を提供し、これは、単一のバイオマーカー、多重化、又は高多重化免疫蛍光 (IF) データを処理する。その方法は、アプローチが全体論的であり、TMA における全腫瘍組織セクション及び / 又はスポットの発現及び空間情報の両方を使用して、空間的関連性を特徴付ける。本明細書で詳細に記載する例示的な実施形態では、本方法は、主要な及びマイナーな亜集団の両方の空間的関連性を明示的に解明する 2 次元不均一性マップを生成する。腫瘍内空間的不均一性の特徴化は、癌の進行、増殖及び治療に対する反応について重要な診断バイオマーカーとなり、従って、本開示概念のこの態様の方法及びシステムは、価値のある診断及び治療ツールとなると信じられる。

【 0 0 4 6 】

本開示概念のこの態様によれば、特定のバイオマーカーの所定のセットを使用して、多重化 / 高多重化蛍光組織画像における空間的不均一性が定量化される。目的の明確化のため、本開示概念のこの態様が、限定でない例示的な実施形態において実施される。当該実施形態では、3 種の乳癌バイオマーカー (エストロゲン受容体 (ER)、ヒト上皮成長因子 2 (HER2) 及びプロゲステロン受容体 (PR)) が、核、細胞膜、細胞質及び上皮細胞を含むセグメントについてバイオマーカーと組み合わせて使用されて、空間的不均一

10

20

30

40

50

性が定量化される。しかしながら、本開示概念のこの態様において、異なる及び/又は追加のバイオマーカーが使用されてよいことは理解されるだろう。また、空間的腫瘍内不均一性を定量化するために自己相互情報量 (PMI) を用いる本開示概念のこの態様の効果は、本明細書に記載する特定の例示的な実施形態を超えて広がりうることも理解されるだろう。例えば、本開示概念のこの態様は、数多くの癌及びストロマバイオマーカーで標識化された全体スライド (whole-slide) IF 画像の分析に広げてもよく、これに限定されない。

【0047】

さらに、本開示概念のこの態様は、本明細書では表現型とも言う、優性のバイオマーカー強度パターン (使用される特定のバイオマーカーの所定のセットに基づいている) の所定のセットを使用して、細胞空間的不均一性を測定して定量化する。従って、図5を参照しながら、最初に、優性バイオマーカー強度パターンを構築する、限定ではない代表的な方法を以下に説明する。その後、優性バイオマーカー強度パターンを使用して空間的不均一性を定量化することを説明する。

【0048】

図5を参照すると、まず、ステップ105にて、既に細胞セグメント化されたデジタル・バイオマーカー画像を各コントロール・スライドが含むような一組のデジタル・コントロール・スライドを得る (即ち、ここでは細胞セグメント化方法が行われる)。例示を目的として、本明細書に記載する例示的な実施形態では、少なくとも前記3種のバイオマーカー (ER、PR及びHER2) を使用して、コントロール・スライドのバイオマーカー画像が作成される。次に、ステップ110で、コントロール・スライドについて、免疫蛍光 (IF) データが作成される。具体的には、ステップ110では、各バイオマーカー画像について、バイオマーカー画像における各セグメント化細胞の所定のバイオマーカーの強度レベルを得ることによって、IFデータが作成される。従って、本明細書の記載から理解できるように、IFデータは、コントロール・スライドの各バイオマーカー画像の各セグメント化細胞に関するバイオマーカー強度レベル・データを含むだろう。

【0049】

次に、ステップ115で、信号強度が真のバイオマーカー発現を示すと仮定して、各バイオマーカーの信号強度の分布に基づいて、IFデータからの細胞が、(以下に記載する閾値を用いて) 2つの区画に分離される。図6は、代表的なバイオマーカー (ER、PR、HER2) のそれぞれについて、バイオマーカー強度分布グラフを示している。図6で示すLog-発生分布は、それぞれ、2つ以上の線形方程式によってモデル化できる。これらの2つの異なるモデルが出くわすであろうノッチが、特定のバイオマーカー・チャンネルにおいて閾値として設定され、図6のバイオマーカー強度分布グラフでは、垂直線で描かれている。任意の所与の細胞について、そのバイオマーカー強度の1又は複数が閾値を超える場合、その細胞はレベル1 (L1) に属する。任意の所与の細胞のバイオマーカー強度の全てが、それらに対応するバイオマーカー・チャンネルにおける閾値未満であれば、その細胞はレベル2 (L2) に属する。これらの2つの区画は、それらの信号雑音比で解釈することができ、比較すると、L1はより高い信号雑音比を有し、L2はより低い信号雑音比を有する。細胞の各区画を使用して、バイオマーカー強度パターンの固有のセットがわかる。L1及びL2データに関するパターン係数の分布は、一般的に異なるガウ斯的性質を有することを考えると、このアプローチは、特に賢明である。図6で示すように、調べたバイオマーカー強度は、長い尾の形をした分布を有しているので、数的に安定したパターン認識アルゴリズムを導くためにLog強度表示が選択される。

【0050】

次に、図5のステップ120で、本明細書では表現型とも称する優性バイオマーカー強度パターンのセットを、以下のようにして区画IFデータから知る。まず、図7に示すように、IFデータの各区画、L1及びL2について、スパース信号表示を導き出す。より具体的には、図7を参照すると、列はIFデータにおける各細胞を示し、行は各細胞のログ・バイオマーカー強度を示している (それぞれ、上から下にER、HER2、PR) 所

10

20

30

40

50

与のデータ行列 X は、行列 D 及び W の積により近似できる。 D は、データセット X における細胞の集合から得られる潜在的なバイオマーカー強度パターンの辞書を示しており、各列は、データから得られるパターンの 1 つを示し、各行は、各パターンの個々のバイオマーカー強度を示す。 W はスパース行列であり、 X における各細胞を、特定のスケール係数で、 D における特異的パターンにフェノタイプ化する。従って、各細胞（ W における列）は、スパース・コードがあるバイオマーカー・パターン（ D における列）に相当するただ 1 つの細胞表現型により示される。各行列の色スペクトルは、 1 つの色、例えば、青（低強度）から他の色、例えば黄（高強度）まで変化する。行列 DW は、実際のデータ行列とその復元との間の類似性を表すために示されている。行列 X 及び DW は、それらが最も一致する辞書エレメントによってカラムソートされており、列である行列 X 及び DW を視察することによって、バイオマーカー強度パターンの各々がデータに存在していることが観察できる。データのこの復元の利益は、少数の解釈可能なバイオマーカー強度パターンで、大きなアレイの細胞レベル・データを示して、図 8 で示すようなデータベースに固有の高度にクラスター化されたクラウドを記述できる能力である。 $3DL \log$ ・バイオマーカー強度空間の各細胞は、その表現型でコード化されている色である。辞書 D 及び辞書係数行列 W への所定のデータセット X の線形表示の復元エラーは、 D の次元に、即ち、データセット X を記述するために使用されるパターンの数に非常に依存する。

10

【 0 0 5 1 】

D の理想的な次元を選ぶために、図 9 に示すように、データ復元の 10 倍クロス検証が実行される。これらの分析において典型的であるように、次元が増加するにつれて、エラー分散が次元とともに増加し始めるある点まで、復元エラー及びエラーの分散は減少する。例示的な実施形態では、 11 個のパターンの辞書サイズが、データ区画 $L1$ 及び $L2$ の両方について、復元エラー及びエラーの分散の両方を最適化することが分かっている。データ $L1$ 及び $L2$ の重なり合わない区画の各々について 11 個パターンのセットを知ること、 2 つの辞書を、全データセットを記述できる 1 つの大きなバイオマーカー強度パターンの辞書にマージすることができた。しかしながら、これらのパターンは、同じ実験条件下で取得された同じデータセットに由来する区画から別々に得られたので、 $L1$ データから得た辞書と $L2$ データから得た辞書との間に幾つかの重複があることが分かった。故に、例示的な実施形態では、 k - 平均クラスター化を使用して、大きな 22 パターン辞書（各区画からの 11 個のパターンを有する）を、本明細書に記載するアプローチから発見された特有のパターンのみを含むようなより小さな最終辞書に集約する。図 10 は、 $L1$ から得た 11 個のパターンと、 $L2$ から得た 11 個のパターンとを示している。各バイオマーカー・パターンは、その ER 、 $HER2$ 及び PR 強度のステムプロットとしてそれぞれ示されている便宜上、ステムプロットの強度パターンは、高、中、低として記述されるであろう。例えば、 $L1$ 辞書（左に示している）のパターン 8 は、 ER 高、 $HER2$ 中、及び PR 低と記述できる。

20

30

【 0 0 5 2 】

図 10 において右側に示された k - 平均クラスター化の結果は、 8 個のバイオマーカー強度パターンの最終辞書次元となる。例示的な実施形態では、最終次元は、クラスター評価についてのシルエット基準の結果に基づいて選ばれた。なお、 1 個のパターンは、区画 $L2$ に特有であり、最終パターン・セットのパターン 7 が、 ER 低発現、 $HER2$ 中発現、及び PR 高発現であった。これは、データを 2 つのグループ、 $L1$ 及び $L2$ に区分することの価値を実証しており、 1 つの区画では優勢であるが他の区画では優勢ではないパターンが解明され得る。

40

【 0 0 5 3 】

一組の優性バイオマーカー強度パターンを得る代表的な手順を説明してきたが、議論は、次に、この優性バイオマーカー強度パターンを使用して空間的不均一性を定量化する方法に移る。具体的には、図 11 は、例示的な実施形態による、多重化 / 高多重化蛍光組織画像の空間的不均一性を定量化する方法のステップを示すフローチャートである。

【 0 0 5 4 】

50

図 1 1 を参照すると、その方法は、分析すべき複数のデジタル多重化蛍光スライドを得るステップ 1 2 5 から始まる。ステップ 1 2 5 で得られるスライドの数は、1 個のスライドでも、1 人の患者からの複数のスライドでも、又は全患者群の複数のスライドでもよい。本明細書から理解できるように、ステップ 1 2 5 で得られたスライドの各々は、関心腫瘍のセクションを含んでおり、そのセクションのバイオマーカー画像となる。本明細書の他の部分で述べたように、例示的な実施形態で使用されるバイオマーカーは、E R、P R 及び H E R 2 である。次に、ステップ 1 3 0 で、細胞セグメント化が、対象スライド（複数を含む）のデジタル・スライド・データ上で実行される。ステップ 1 3 0 では、数多くの公知の又は後で開発される適切な細胞セグメント化アルゴリズムの何れかが使用されてよい。次いで、ステップ 1 3 5 で、対象スライド（複数を含む）における各細胞について空間位置及びバイオマーカー強度データが得られる。次に、ステップ 1 4 0 で、対象スライド（複数を含む）の各細胞を、細胞のバイオマーカー強度構成に基づいて、所定の優性バイオマーカー強度パターン（即ち表現型の 1 つ）の 1 つに割り当てる。図 1 2 は、1 ~ 8 と付された所定の優性バイオマーカー強度パターンの各々の模式的表示 1 6 0 を示す。例示的な実施形態では、各模式的表示 1 6 0 は、その模式的表示を他のものから容易に区別できるように、固有の色（複数を含む）で表されている。次に、ステップ 1 4 5 で、細胞割当てと図 1 2 に示す模式的表示とを利用して、対象組織試料（複数を含む）の不均一性を目に見えるように表示する細胞空間依存画像（cell spatial dependency image）が生成される。図 1 3 は、本開示概念の 1 つの特定の例示的な実施形態による細胞空間依存画像 1 6 5 を示す。図 1 3 でわかるように、細胞空間依存画像 1 6 5 は、模式的表示 1 6 0 を使用して、対象スライド（複数を含む）の細胞の間の空間依存性を示す。例示的な実施形態では、細胞空間依存画像 1 6 5 は、以下ケースの確率を記録する：（i）免疫細胞が癌細胞の近くで発生する場合、（ii）免疫細胞及び癌細胞が互いを抑制する場合、及び（iii）免疫細胞及び癌細胞は互いに非依存性である場合。細胞空間依存画像 1 6 5 は、特定の組織構造を示してはいない。

【 0 0 5 5 】

次に、ステップ 1 5 0 で、空間ネットワークを構築して、対象スライド（複数を含む）における優性バイオマーカー強度パターンの構成を記述する。次いで、ステップ 1 5 5 で、対象スライド（複数を含む）の不均一性を、本明細書に記載したように、スライド（複数を含む）の P M I マップを生成することによって、定量化する。例示的な実施形態では、ステップ 1 5 0 及び 1 5 5 は、以下に述べるようにして実行される。

【 0 0 5 6 】

対象スライドのバイオマーカー画像（即ち、組織 / 腫瘍試料）におけるバイオマーカー・パターンの空間構成を示すために、ネットワークを、対象スライドについて構築する。腫瘍試料に関する空間ネットワークの構築により、細胞バイオマーカー強度データ（ネットワークのノードで）は、空間データ（ネットワークのエッジで）に固有に結合される。ネットワーク構築における仮定は、細胞は、ある限界、例えば $250 \mu\text{m}$ まで、近くの細胞とのコミュニケーション機能を有しており、その限界内での細胞のコミュニケーション機能は、細胞の距離に依存するということである。従って、例示的な実施形態の確率分布は、対象スライドにおける細胞とそれに最も近い 1 0 個の隣接物との間の距離についてコンピュータで計算される。ハード限界は、（標準偏差を検出するため）この分布の中央値の 1 . 5 倍で選択されて、ネットワーク内の細胞は、この限界内でのみ接続された。次いで、ネットワークにおける細胞間のエッジは、隣り合う細胞間の距離により重み付けられる。

【 0 0 5 7 】

次に、自己相互情報量（P M I）を使用して、対象スライドについて、辞書におけるバイオマーカー・パターン、つまり異なる細胞表現型の各ペアの間の関連性を測る。この測定手法は、線形及び非線形の両方の一般的統計的関連性を取得し、従来の研究はスピアマンのロー係数のような線形計量法を使用していた。バイオマーカー・パターンの各ペアについて、P M がコンピュータで計算されると、対象スライドのデータにおける全ての関連

10

20

30

40

50

性の測定値がPMIマップに表示される。典型的なPMIマップ170を図14に示す。

【0058】

PMIマップ170は、対象スライドの微小環境内における異なる細胞表現型間の関係を記述する。特に、PMIマップ170におけるエントリ172は、全ての表現型にわたるランダム（又はバックグラウンド）分布により予測される相互関係と比べた場合に、データセットにおいて、2つの表現型（行及び列の番号で表される）間の特定の空間的相互関係がどの程度の頻度で起こるかを示す。赤のような第1色におけるエントリは、表現型間の強い空間的関連性を表しており、黒のような第2色におけるエントリは、共局在化がないこと（表現型間の弱い空間的関連性）を表す。他の色は、他の関連性を表すために使用されてもよい。例えば、緑のような第3色に色付けされたPMIエントリ172は、全データセットにわたる細胞表現型のランダム分布にすぎないという関連性を表す。また、PMIマップ170は、エントリ172が青のような第4色で示されると、反関連性を表してよい（例えば、表現型1がほとんど現れなければ、空間的に近い表現型3）

10

【0059】

従って、PMIマップ170は、強い対角エントリ及び弱い非対角エントリを有しており、全体的には不均質であるが局所的には均質な腫瘍を記述する。そのようなPMIマップ170Aの例を図15に示す。この例では、表現型2、4及び8に関する対角エントリにおける関連性が強い。このことは、図15で示す腫瘍試料画像の個々の微小ドメインの構成によって示されるように、これらの表現型が、同じ表現型の細胞と空間的に関連していることを示唆している。逆に、強い非対角エントリを有するPMIマップ170Bは、局所的に不均質な腫瘍を記述することができる。そのようなPMIマップ170Bの例を図16に示す。この例では、細胞表現型1及び6、細胞表現型2及び4、ならびに細胞表現型3及び8の間の関連性が空間的に局在化している。さらに、PMIマップ170Bは、表現型7の細胞とそれ自身とのただ1つの関連性を示している。図17で示す例示的なPMIマップ170Cは、腫瘍画像における全ての表現型間の関連性を示し、従って、PMIマップ170Cは、色が完全に混在している。既存の測定にわたったPMIマップの利益は、そのマップが表現型間の空間的関係を再現することである。それらは、細胞構成の概要だけでなく、腫瘍トポロジーの近似も提供する。簡単にするため、より複雑なPMIマップの例は含んでいないが、全てのPMIマップ170は、これらの単純な相互関係で成り立っていることは理解されるだろう。

20

30

【0060】

例示的な実施形態では、対象スライドのPMIは、以下のように計算される。過完備辞書D（Dの各列は別個のパターン d_i である）へのIFデータセットX（Xの各列は、細胞 x_k である）の線形脱構築（linear deconstruction）と、各細胞をただ1つの単一バイオマーカー強度パターンに割り当てるスパースコーディング行列Wとが与えられると、本明細書で説明したように（ステップ140）、各細胞は、表現型 f_i （ i は、Wの列 w_k におけるゼロ以外のインデックスである）を有するように割り当てられる。アルゴリズムの潜在的な落とし穴は、信号強度が高い及び低い細胞が、同じ細胞表現型に割り当てられる可能性があることである。所与のネットワーク又はネットワークセット s についての1

40

【数9】

$$PMI_s(f_i, f_j) = \log \frac{P(f_{i_s}, f_{j_s})}{P(f_{i_t})P(f_{j_t})}$$

ここで、 $P(f_{i_s})$ は、表現型 f_i がネットワークセット s で起こる確率であり、 $P(f_{i_t})$ は、表現型 f_i がネットワークの完全集合から得られるバックグラウンド確率分布である。個々のネットワークを全体として組織スライドの分布と比較するために、パッ

50

クグラウンド分布は、全データセットに基づいていることに留意のこと。この構成は、DNA又はタンパク質配列について位置特異的スコア行列（PSSM）に類似しており、バックグラウンド分布は、任意の所与の位置について、配列のデータセットにわたる任意の特定のヌクレオチド又はアミノ酸を見出す確率を示す。PMIマップは、所与のネットワークセットsについて、単語集のパターンの可能なペア毎にPMIスコアで構成される。不均一性を十分に理解するために、二次元PMIマップの解釈を唱道してきたが、文献にある他の一次元スコアとの比較に関心のある読者の便宜のため、PMIマップから一次元不均一性スコア値も導き出すことにする。情報不十分な一次元不均一性スコアは、下式で規定される。

【数10】

$$HET_{PMI_s} = \sum_{i,j} \left| \log \frac{P(f_{i_s}, f_{j_s})}{P(f_{i_t})P(f_{j_t})} \right|,$$

ここで、より高いスコアは、バックグラウンド分布からのより大きな差異を表す。一次元スコアは、それらのPMIマップによって示されるように、TMEの2つの空間的に異なる構成を同じスコアに間違っマップする可能性がある。

【0061】

対象スライドについてPMIマップ170をコンピュータで計算し、有意な相互関係又は相互関係モチーフを特定した後で、この有意な関連性に寄与した細胞を調べる必要がある。有意な相互関係は、PMI値が±1に近い場合に考慮される。1に近いPMI値は、バイオマーカー・パターンのこの特定の空間的相互関係が、バックグラウンド分布において観察されるよりも頻繁に起こることを意味する。-1に近いPMI値は、1つのパターンがネットワークにおいて観察される場合、他のパターンが、バックグラウンド分布から予測されるよりも数少なく観察されることが見られることを意味する。0に近いPMI値は、バックグラウンド分布によって適切に記述され得る相互関係を意味する。

【0062】

< C . システムの実施 >

図18は、本明細書で記載するH&E染色組織画像セグメント化する手順を実施可能な例示的なシステム・デジタル・パソロジー（pathology）200の概略図である。図18に示すように、システム200は、本明細書で説明されているように、H&E染色組織画像を表すデジタル画像データを受け取って、これらの画像を処理するように構成されたコンピューティング・デバイスである。システム200は、例えば、PC、ラップトップ・コンピュータ、タブレット・コンピュータ、スマートフォン、又は本明細書に記載した機能を行うように構成された任意の他の適切なデバイスであってよく、これらに限定されない。システム200は、入力装置202（例えば、キーボード）、ディスプレイ204（例えば、LCD）、及び処理装置206を含む。ユーザは入力装置202を使用して処理装置206に入力することができ、処理装置206は、出力信号をディスプレイ204に提供し、本明細書で詳細に記載するように、ディスプレイ204にユーザに対して情報（例えば、セグメント化組織画像）を表示するようにする。処理装置206は、プロセッサとメモリとを含む。プロセッサは、例えば、マイクロプロセッサ（μP）、マイクロコントローラ、又はメモリとインターフェイスで接続する他の適切なプロセッシング・デバイスでもよく、これらに限定されない。メモリは、限定ではないが、コンピュータの内部記憶領域の形をしたデータ記憶のためのRAM、ROM、EPROM（複数を含む）、EEPROM（複数を含む）、FLASH、記憶レジスタをもたらずもそのような、種々のタイプの内部及び/又は外部記憶媒体の任意の1種又は複数であってよく、例えば、機械読み取り可能媒体であり、揮発性メモリ又は不揮発性メモリであってよい。メモリは、本明細書に記載された開示概念を実施するためのルーチンを含む、プロセッサによって実行可能な数多くのルーチンをそこに格納している。特に、処理装置206は、H&E染色組織

10

20

30

40

50

画像を表す、受け取った画像データに基づいて、本明細書に記載されているように、H & E 染色組織画像に関する局所的空間的統計を定量化するように構成された定量化コンポーネント 208 と、本明細書に記載されているように、局所的空間的統計に基づいて、H & E 染色組織画像内で組織学的構造を特定するように構成された特定コンポーネント 210 と、受け取った画像データ及び特定した組織学的構造を使用して、セグメント化 H & E 染色組織画像を生成するように構成されており、次いで、画像をディスプレイ 204 に提供してもよいセグメント化組織画像生成コンポーネント 212 とを含んでいる。定量化コンポーネント 208 は、H & E 画像データにおける隣接ピクセル間の統計的関連性を示す相互情報量データの決定によって局所的空間的統計を定量化するように構成された 1 つ以上のコンポーネントを含んでよく、特定コンポーネント 210 は、本明細書に記載された、相互情報量データ及びグラフに基づくスペクトル・セグメント化アルゴリズムを使用して組織学的構造を特定するように構成された 1 つ以上のコンポーネントを含んでもよい。或いは、定量化コンポーネント 208 は、RGB データから推定核位置をスーパーピクセル形式で特定し、各スーパーピクセルと複数のそれに最も近い隣接ものとの間の点別距離 (pointwise distance) に基づいてスーパーピクセル・グラフを作成して、スーパーピクセルを標識セグメントにクラスター化する 1 つ以上のコンポーネントを含んでもよく、特定コンポーネント 210 は、本明細書に記載されているように、標識セグメントを組織学的構造にマージして特定するように構成された 1 つ以上のコンポーネントを含んでもよい。

【0063】

図 19 は、本明細書に記載する空間的腫瘍内不均一性を定量化する手順を実施してもよい代表的なデジタル・パソロジー・システム 300 の概略図である。図 19 で示すように、システム 300 は、本明細書に記載されているように、蛍光組織画像を表すデジタル画像データを受け取り、これらの画像を処理するように構成されたコンピューティング・デバイスである。システム 300 は、例えば、PC、ラップトップ・コンピュータ、タブレット・コンピュータ、スマートフォン、又は任意の他の本明細書に記載する機能性を行うように構成された適切なデバイスであってもよく、これらに限定されない。システム 300 は、入力装置 302 (例えば、キーボード)、ディスプレイ 304 (例えば、LCD)、及び処理装置 306 を含む。ユーザは、入力装置 302 を使用して処理装置 106 に入力することができ、処理装置 306 は、出力信号をディスプレイ 304 に提供し、本明細書で詳細に記載されているように、ディスプレイ 304 にユーザに対して情報 (例えば、空間依存画像及び PMI マップ) を表示するようにする。処理装置 306 はプロセッサとメモリとを含む。プロセッサは、例えば、マイクロプロセッサ (μP)、マイクロコントローラ、又はメモリとインターフェイスで接続する他の適切なプロセッシング・デバイスでもよく、これらに限定されない。メモリは、限定ではないが、コンピュータの内部記憶領域の形をしたデータ記憶のための RAM、ROM、EPROM (複数を含む)、EEPROM (複数を含む)、FLASH、記憶レジスタをもたらずものような、種々のタイプの内部及び/又は外部記憶媒体の任意の 1 種又は複数であってもよく、例えば、機械読み取り可能媒体であり、揮発性メモリ又は不揮発性メモリであってもよい。メモリは、本明細書に記載された開示概念を実施するためのルーチンを含む、プロセッサによって実行可能な数多くのルーチンをそこに格納している。具体的には、処理装置 306 は、幾つかの蛍光組織画像を表す画像データで細胞セグメント化を行い、該幾つかの蛍光組織画像の幾つかの細胞を特定するように構成された細胞セグメント化コンポーネント 308 と、細胞の各々を複数の所定のバイオマーカー強度パターンの 1 つに割り当てるように構成された割当てコンポーネント 310 と、割り当てられた所定のバイオマーカー強度パターンに基づいて、蛍光組織画像の空間的統計を定量化する定量化コンポーネント 312 と、細胞空間依存画像 165 又は PMI マップ 170 のような、定量化された空間的統計の視覚表示を作成する視覚表示作成コンポーネント 314 とを含む。定量化コンポーネント 312 は、空間ネットワークを構築して空間的統計を定量化し、幾つかの蛍光組織画像における所定のバイオマーカー強度パターンの構成を記述し、所定のバイオマーカー強度パターンのペア毎に自己相互情報量をコンピュータで計算して幾つかの蛍光組織画像の不均一性を定量

10

20

30

40

50

化するように構成された1つ以上のコンポーネントを含んでもよい。

【0064】

請求項において、括弧の間に置かれている符号は、請求項を限定するように解釈されるべきではない。用語「備える」又は「含む」は、特許請求の範囲に挙げているもの以外の要素や工程の存在を排除しない。幾つかの手段を列挙する装置の請求項において、これら手段の幾つかは、ハードウェアの1つ及び同じアイテムによって具体化されてよい。要素の前にある用語「1つ」は、そのような要素が複数あることを排除しない。幾つかの手段を列挙する装置の任意の請求項において、これら手段の幾つかは、ハードウェアの1つ及び同じアイテムによって具体化されてよい。ある要素が互いに異なる従属請求項において列挙されるという単なる事実は、これらの要素を組み合わせ使用できないことを示してはいない。

10

【0065】

説明の目的で、現在、最も実用的で好ましい実施形態であると考えられることに基づいて本発明を詳細に記載してきたが、そのような詳細は、単に、その目的のためだけであって、本発明は、開示した実施形態に限定されることなく、逆に、添付の請求項の精神及び範囲内にある変更及び等価な構成を含むものと理解される。例えば、本発明は、可能な範囲で、任意の実施形態の1つ以上の特徴を任意の別の実施形態の1つ以上の特徴と組み合わせることができることを意図しているものと理解される。

【図1】

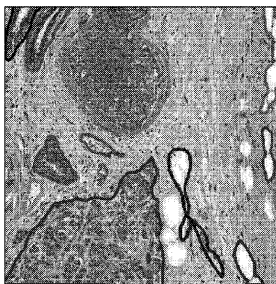


FIG.1

【図2】

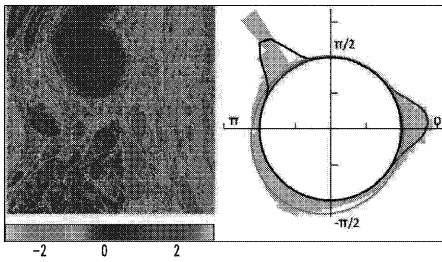


FIG.2

【図3】

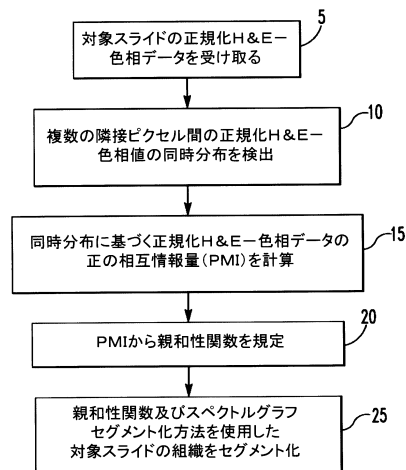


FIG.3

【 図 4 】

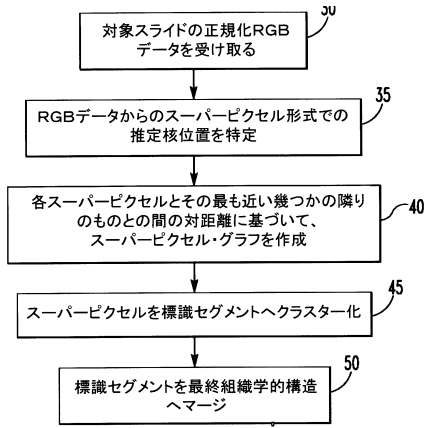


FIG. 4

【 図 5 】

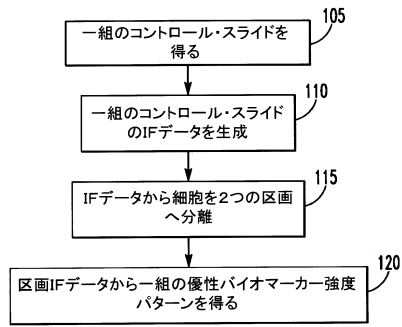


FIG. 5

【 図 7 】

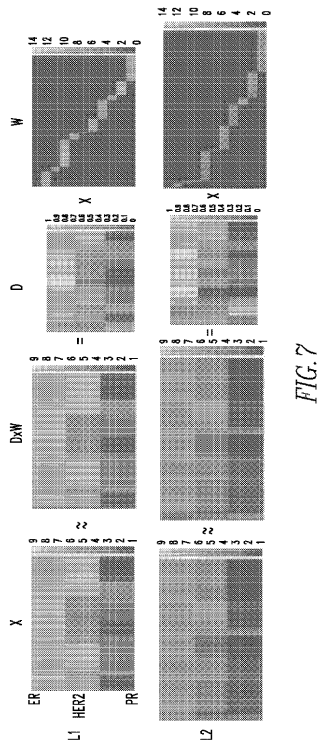


FIG. 7

【 図 6 】

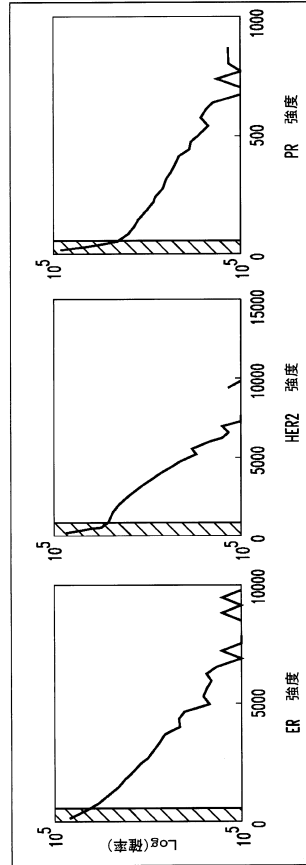


FIG. 6

【 図 8 】

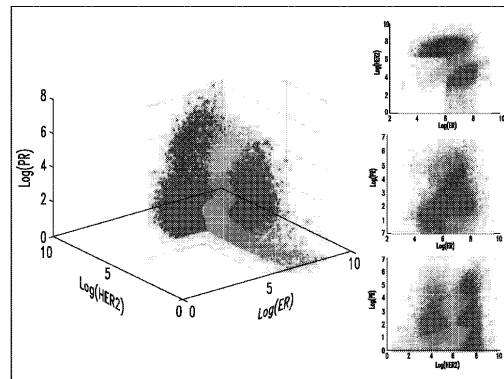


FIG. 8

【図9】

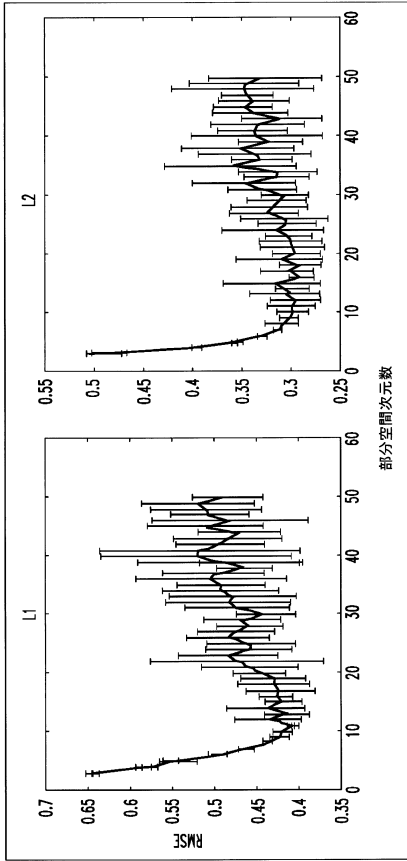


FIG.9

【図10】

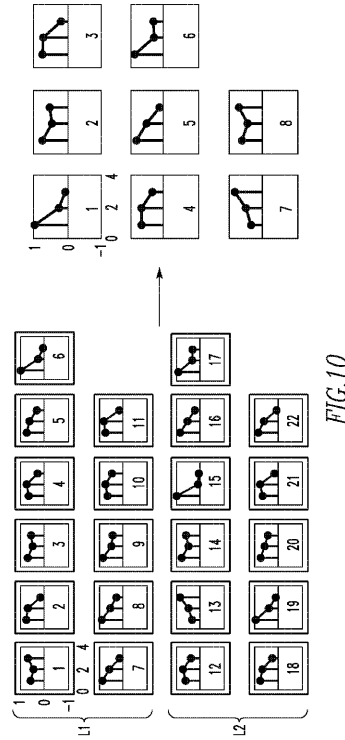


FIG.10

【図11】

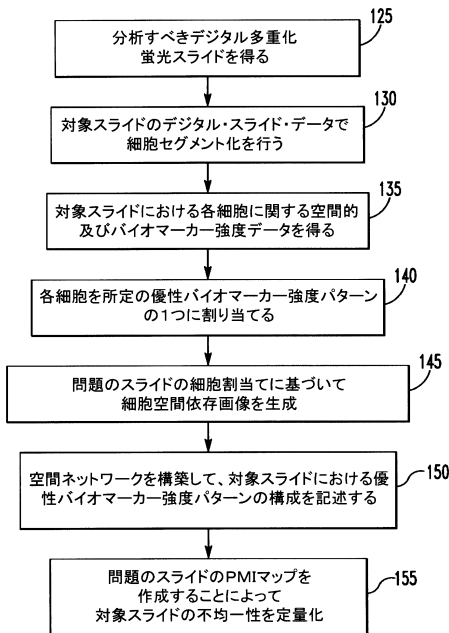


FIG.11

【図12】

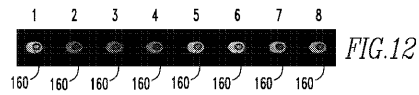


FIG.12

【図13】

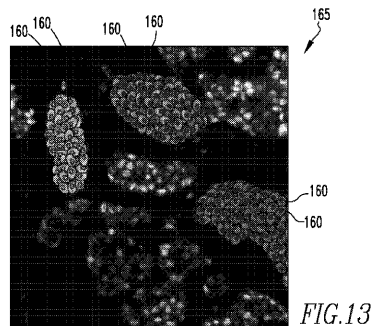


FIG.13

【 図 1 4 】

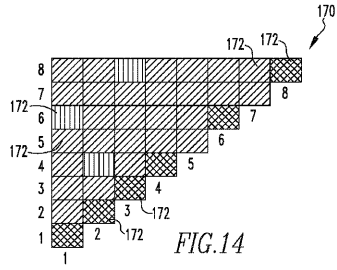


FIG.14

【 図 1 5 】

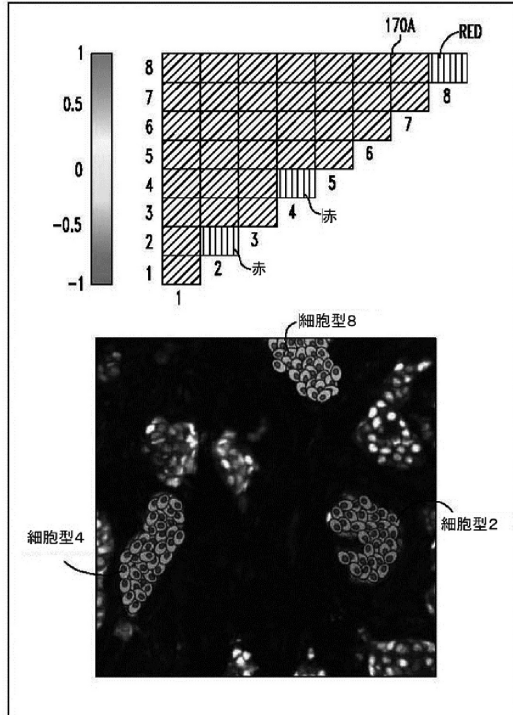


FIG.15

【 図 1 6 】

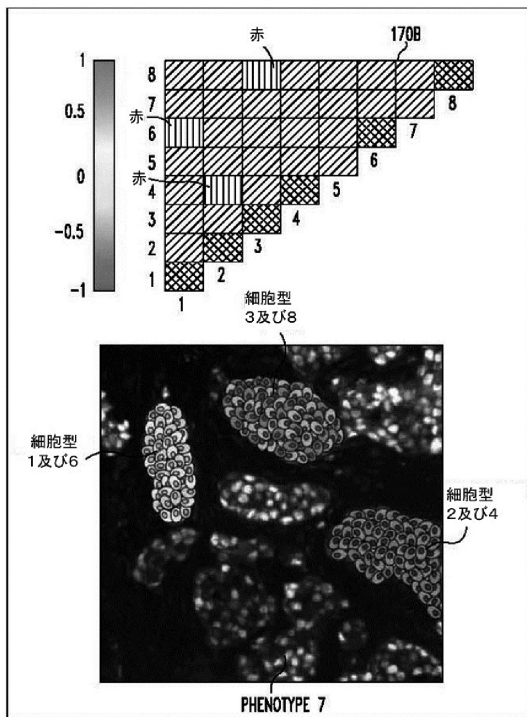


FIG.16

【 図 1 7 】

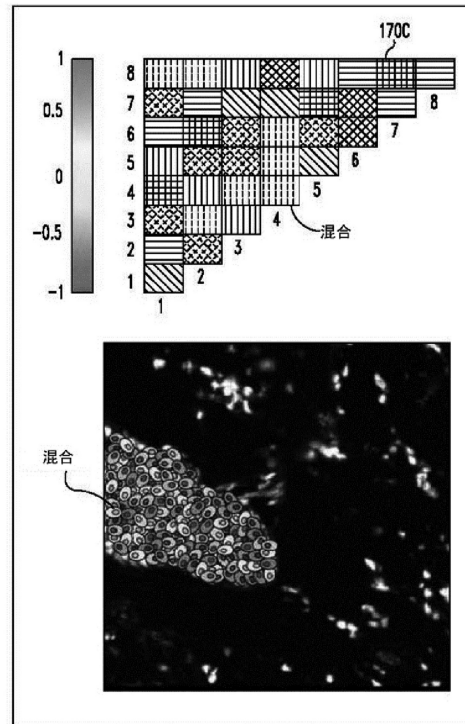


FIG.17

【図18】

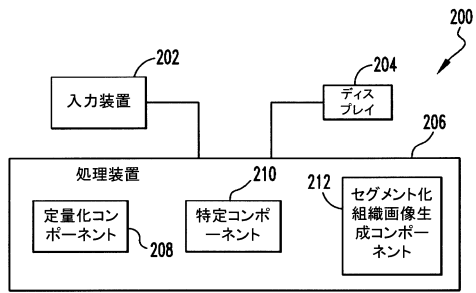


FIG.18

【図19】

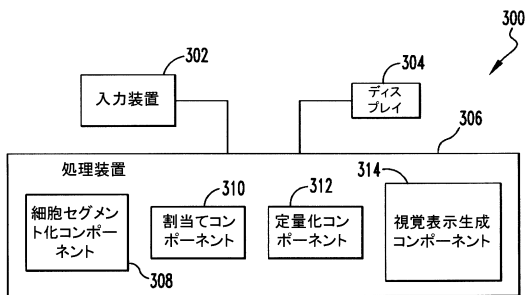


FIG.19

フロントページの続き

- (72)発明者 チェンヌボトラ, エス. チャクラ
アメリカ合衆国 1 5 2 1 7 ペンシルベニア, ピッツバーグ, エプディー ストリート 6 3 3
4
- (72)発明者 テイラー, ディー. ランシング
アメリカ合衆国 1 5 2 1 5 ペンシルベニア, ピッツバーグ, ノーター デイム プレイス 9
1 0

審査官 藤原 敬利

- (56)参考文献 米国特許出願公開第2011/0286654 (US, A1)
特開2015-087167 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 0 1 N	3 3 / 4 8	-	3 3 / 9 8
G 0 6 T	1 / 0 0	-	1 / 4 0
G 0 6 T	3 / 0 0	-	9 / 4 0