

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3686439号
(P3686439)

(45) 発行日 平成17年8月24日(2005.8.24)

(24) 登録日 平成17年6月10日(2005.6.10)

(51) Int. Cl.⁷

H04N 1/40

F I

H04N 1/40

F

請求項の数 10 (全 29 頁)

(21) 出願番号	特願平6-318778	(73) 特許権者	398038580
(22) 出願日	平成6年12月21日(1994.12.21)		ヒューレット・パカード・カンパニー
(65) 公開番号	特開平7-231388		HEWLETT-PACKARD COMPANY
(43) 公開日	平成7年8月29日(1995.8.29)		アメリカ合衆国カリフォルニア州パロアルト
審査請求日	平成13年12月19日(2001.12.19)		ト ハノーバー・ストリート 3000
(31) 優先権主張番号	171551	(74) 代理人	100063897
(32) 優先日	平成5年12月21日(1993.12.21)		弁理士 古谷 馨
(33) 優先権主張国	米国 (US)	(74) 代理人	100076680
			弁理士 溝部 孝彦
		(74) 代理人	100087642
			弁理士 古谷 聡

最終頁に続く

(54) 【発明の名称】 デジタル・イメージのフォト領域検出システム、及び方法

(57) 【特許請求の範囲】

【請求項1】

デジタルイメージの複数のセルからなる領域を、フォト領域又は非フォト領域のどちらか一方として分類する方法であって、

前記複数のセルの各々について、前記セルがフォト・セルである可能性を示す1つ以上の特徴を抽出するステップ(a)と、

前記ステップ(a)で抽出された前記1つ以上の特徴に基づいて、前記セルの各々がフォト・セルになり得るか否かを判定するステップ(b)と、

前記ステップ(b)の結果に基づいて前記セルの各々を分類し、フォト・セル、非フォト・セルまたは分類されないセルのうち1つとしてラベルを付けるステップ(c)と、

前記セルを用いて、前記領域をフォト領域または非フォト領域のどちらか一方として成長させるステップ(d)と、

からなり、

さらに、前記ステップ(a)は、前記セルの共起行列の主対角線に関する慣性モーメントの量を判定するステップ(i)を含み、

さらに、前記ステップ(i)は、

前記セルの各々に対して、第1のグレイレベルを有するピクセルが第2のグレイレベルを有するピクセルに隣接する回数を表す要素を有する共起行列を構築するステップ(1)と

、前記共起行列のグレイレベル遷移の量を表す慣性モーメントを判定するステップ(2)とを含む、

10

20

ように構成される前記方法。

【請求項 2】

前記ステップ (a) ~ (d) の後に、前記領域の各々の前記分類に基づいてピクセル当たりのビット数を各々の領域へ割り当てるステップ (e)、
をさらに含む請求項 1 の方法。

【請求項 3】

前記ステップ (e) の後に、前記非フォト領域の各々に対して閾値レベルを定めるステップ (f) と、
前記ステップ (f) の後に、前記非フォト領域の各々をピクセル当たり 1 ビットの形式へ変換するステップ (g) と、
をさらに含む請求項 2 の方法。

10

【請求項 4】

前記ピクセルの各々が複数のグレイレベルのうちの 1 つで表現され、
前記特徴を抽出するステップ (a) が、
前記複数のセルの各々において各々のグレイレベルにある前記ピクセルの数を示すグレイレベルヒストグラムを構築するステップ (ii)、
をさらに含む、請求項 1 の方法。

【請求項 5】

前記特徴を抽出するステップ (a) が、
ステップ (a) (ii) の前に、前記デジタルイメージを量子化し、ピクセルの各々を表現
するのに用いられる前記ビット数へ変換するステップ (iii)、
をさらに含む、請求項 4 の方法。

20

【請求項 6】

前記ステップ (a) (iii) が、人間の目の応答に適應させるため、不均一量子化を用いて実施される、請求項 5 の方法。

【請求項 7】

前記 1 つ以上の特徴を抽出するステップ (a) の間に抽出される前記 1 つ以上の特徴が、
前記セルに対するグレイレベル分布度であり、前記ステップ (a) がさらに、前記複数のセルの各々に対するグレイレベル分布度を判定するステップ (iii) を含み、
前記ステップ (iii) が、
前記グレイレベルヒストグラムにおいて、どのグレイレベルがピークであるか及びどのグレイレベルが谷であるかを判定するステップ (1) と、
前記グレイレベルヒストグラムの各々のピークに対して、部分母集団を判定するステップ (2) と、
前記部分母集団の各々の確率を判定するステップ (3) と、
前記部分母集団の各々の分布度を判定するステップ (4) と、
前記部分母集団の各々の重複確率を判定するステップ (5) と、
前記部分母集団の各々の確率、前記部分母集団の各々の分布度および前記部分母集団の各々の重複確率を用いて、前記グレイレベルヒストグラムの前記グレイレベル分布度を判定するステップ (6) と、
を含む、請求項 4 の方法。

30

40

【請求項 8】

非フォト・セルとしてラベルが付けられた前記セルの各々に対し、前記非フォト領域を成長させるステップ (d) (i) がさらに、
前記非フォト・セルの前記グレイレベルヒストグラムの明るさのピークのグレイレベルと、隣接する分類されないセルの前記グレイレベルヒストグラムの明るさのピークのグレイレベルとの差が、1 グレイレベルより大きいか否かを判定するステップ (1) と、
前記差が 1 グレイレベル以下の場合、前記隣接するセルを非フォト・セルとして分類するステップ (2) と、
前記非フォト・セルと前記隣接するセルの組み合わせを、非フォト領域として分類するス

50

テップ(3)と、
からなる、請求項1の方法。

【請求項9】

デジタルイメージ内のある領域を、フォト領域又は非フォト領域のどちらか一方として分類する方法であって、前記領域の各々は複数のセルからなり、前記複数のセルの各々はピクセルからなり、前記ピクセルの各々は複数のグレイレベルのうち1つを表し、

前記方法は、

前記複数のセルの各々について、前記セルがフォト・セルである可能性を示し且つ前記セルに対するグレイレベル分布度を含む1つ以上の特徴を抽出するステップ(a)と、

前記ステップ(a)で抽出された前記1つ以上の特徴に基づいて、前記セルの各々がフォト・セルになり得るか否かを判定するステップ(b)と、

前記ステップ(b)の結果に基づいて前記セルの各々を分類し、フォト・セル、非フォト・セルまたは分類されないセルのうち1つとしてラベルを付けるステップ(c)と、

前記セルを用いて、前記領域をフォト領域または非フォト領域のどちらか一方として成長させるステップ(d)と、

からなり、

さらに、前記ステップ(a)は、前記複数のセルの各々において各々のグレイレベルにある前記ピクセルの数を示すグレイレベルヒストグラムを構築するステップ(i)と、前記複数のセルの各々に対してグレイレベル分布度を判定するステップ(ii)とを含み、

さらに、前記ステップ(ii)は、

前記グレイレベルヒストグラムにおいて、どのグレイレベルがピークを表すか及びどのグレイレベルが谷を表すかを判定するステップ(1)と、

前記グレイレベルヒストグラムの各々のピークに対して、部分母集団を判定するステップ(2)と、

前記部分母集団の各々の確率を判定するステップ(3)と、

前記部分母集団の各々の分布度を判定するステップ(4)と、

前記部分母集団の各々の重複確率を判定するステップ(5)と、

前記部分母集団の各々の確率、前記部分母集団の各々の分布度および前記部分母集団の各々の重複確率を用いて、前記グレイレベルヒストグラムの前記グレイレベル分布度を判定するステップ(6)と、

を含むように構成される前記方法。

【請求項10】

前記特徴を抽出するステップ(a)が、さらに、

前記ステップ(a)(ii)の前に、人間の目の応答に適應させるため不均一量子化を用いて前記デジタルイメージを量子化し、前記ピクセルの各々を表すのに用いられるビット数へ変換するステップ(iii)を含む、請求項9の方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

本発明は、一般に、イメージ・スキャナの技術に関するものであり、より詳しくは、デジタル・イメージ内のフォト領域を検出するためのシステム、及び方法に関する。

【0002】

【従来の技術】

デジタル・イメージ・スキャナは、コンピュータによって扱うことができる様に、文書をデジタル化するために用いられることが多い。例えば、スキャナを利用してテキストをデジタル化し、エディタがその文書を書き直したり、修正したりすることが可能な様に、ワード処理プログラムによって認識されたフォーマットに変換する。

【0003】

プロセッサ速度、及びスキャナ技術の進歩によって、印刷されたテキストだけでなく、絵画イメージのデジタル化も都合良く行える様になってきた。スキャナの性能を高めようと

10

20

30

40

50

、開発者はデジタル・イメージを表すために用いられるビット数の最適化方法を探し求める努力をしてきた。イメージを表すために用いられるビット数を単純に減らすと、イメージの品質を低下させることになる。ビット数を増やすと、イメージのデジタル化、及び後続のイメージ処理に必要なプロセッサ時間を増大させることになる。

【0004】

イメージを表すために用いられるビット数を最適化するため、開発者はイメージを十分に表すことができるグレイレベルの数（即ち、グレイレベルの数を表現可能なビット数）に基づいて、スキャンされたイメージを複数領域に分割しようとしてきた。例えば絵画イメージを含む領域の様に、領域の中にはイメージを十分に表すのに、多数のグレイレベルを必要とするものもある。これらの領域は、「フォト領域」と呼ばれる。ブレイク・テキスト、又はライン・アートを含む領域の様な他の領域は、イメージを効果的にデジタル化するのに、それほど多くのグレイレベルを必要としない。これらの領域は「非フォト領域」と呼ばれる。

10

【0005】

イメージの解像度を保つために、フォト領域はピクセル毎に複数（例えば、4つ以上）のビットで表される。一方、非フォト領域は、一般にピクセル毎に1つのビットで表わされるが、それでも重要な情報を保持することができる。非フォト領域の場合、ピクセルを表すために利用される1ビットは、走査されたグレイレベルが、特定のグレイレベル閾値より大きいか、小さいかに基づいて、セットされる。換言すれば、閾値グレイレベルが選択され、閾値グレイレベルより暗いピクセルは黒にセットされ、閾値グレイレベルより明るいピクセルは白にセットされる。レーザ・プリンタの様な2レベル出力装置でイメージを印刷すべき場合、その品質を保つために、フォト領域ではハーフ・トーン、又は拡散までも必要とされるが、非フォト領域は、一般に上述の様な閾値処理がされうる。

20

【0006】

OCR（光学文字認識）製品の「ページ分解」又は「ページ分析」コンポーネントに利用される様な、伝統的な解決策の大部分は、閾値イメージ（ピクセルあたり1ビット）を使用しており、グラフィックスとテキストを区別することが可能である。しかし、こうした従来のシステムは、グラフィックの性質に関する情報（即ち、写真、ライン・アート）を正確に示すことはできない。

【0007】

【発明が解決しようとする課題】

本発明の目的は、デジタル方式で文書を表す方法を最適化するためのシステム、及び方法を提供することにある。

30

【0008】

【課題を解決するための手段】

本発明によれば、デジタル化された文書は複数領域に分割される。その領域はフォト領域、又は非フォト領域として分類される。非フォト領域には、テキスト、又は単純な線画の様な簡単な内容が含まれる。一方、フォト領域は絵の様な、より複雑なイメージを含んでいる。フォト領域に含まれる情報の複雑さのために、これらの領域は、ピクセル毎に複数のビットを用いることによって、より適切に表される。しかし、非フォト領域は、通常ピクセル毎に1ビットで表されうる。従って、デジタル化された文書を表すために用いられるビット数は、フォト領域に対してはピクセル毎に複数ビットを割り当て、非フォト領域に対してはピクセル毎に1ビットを割り当てることによって、最適化されうる。

40

【0009】

文書内の各領域がフォト領域、又は非フォト領域として特徴付けられる。これは、デジタル・イメージを、それぞれがピクセルのグループを含んでいる複数のセルに分割することによって実現される。複数のセルから選択されたセルが個別に検査され、各セルがフォト・セル（即ち、フォト領域の一部）と、非フォト・セル（非フォト領域の一部）のどちらの可能性が高いかが判定される。その可能性の判定が終わると、セルはフォト・セル、又は非フォト・セルのどちらかとして分類される。次に、フォト領域、又は非フォト領域の

50

どちらかを成長させるために、各セルが利用される。

【0010】

セルがフォト・セルの可能性があるかどうかの判定は、セルから特徴を抽出することによって開始される。セルがフォト・セルである可能性があるかどうかを示す、多くの特徴を抽出することができる。抽出される最も重要な特徴の1つは、グレイレベル分布度である。グレイレベル分布度を決定するため、グレイレベル・ヒストグラムと呼ばれる1次元ヒストグラムが、セルから作成される。グレイレベル・ヒストグラムは、セル内の各グレイレベル毎にピクセル数を示す。例えば、イメージが16個のグレイレベルで表される場合、グレイレベル・ヒストグラムは、16個のグレイレベルのそれぞれにおいて、そのセル内にピクセルがいくつ含まれているかを表示する。

10

【0011】

グレイレベル・ヒストグラムが作成されると、それが検査され、どのグレイレベルがヒストグラムのピークを表し、どのグレイレベルがヒストグラムの谷を表すのかが決定される。ピークは、局部的に最大数のピクセルが存在するグレイレベルによって定義される。例えば、白の背景に黒のテキストが印刷されたセルは、2つのピーク、即ち黒、又は黒に近いテキストに関するグレイレベルの低い第1のピーク、及び白、又は白に近い背景に関するグレイレベルの高い第2のピークを示すことになる。

【0012】

ピーク、及び谷が分かると、部分母集団、即ちモードが決定される。モードはピークと、ピークの左右の2つの谷によって定義される。各モード毎に、そのモードの確率、及び分布度が決定される。更に、各モードに関連して重複確率が決定される。

20

【0013】

ピーク、谷、及びモード情報を利用して、ヒストグラムのグレイレベル分布度が決定される。グレイレベル分布度は、他の特徴と共に、セルがフォト・セルと非フォト・セルのどちらの可能性が高いかを示す。ピーク同士がより接近し、かつ/又はピークの幅がより広がる場合、グレイレベル分布度が増大することになる。このグレイレベル分布度の増大は、セルがフォト・セルである可能性がより高いということを示している。対照的に、よりピークの間隔が広く、ピークの幅が狭いピークを示す、より小さいグレイレベル分布度は、非フォト・セルの可能性がより高いことを表している。これは明るい背景に暗いテキストが現れる、上記で用いられた例と一致する。この例の場合、2つのピーク、即ち、白の背景に関するピークと、黒のテキストに関するピークが存在する。2つの幅の狭いピークは、グレイレベル・スペクトルのほぼ全域の幅だけ離れているので、このヒストグラムに関するグレイレベル分布度は小さい。

30

【0014】

可能性が判定されると、セルはその可能性に基づいて、フォト・セル、又は非フォト・セルのどちらかに分類される。分類されたセルは、次に、フォト領域、又は非フォト領域のどちらかに成長するための種子として用いられる。非フォト領域は、非フォト・セルから成長する。非フォト領域は、非フォト・セルである可能性の高いセルの位置を突き止め、同様の特性を備えた、まだ分類されていない隣接するセルを結合することによって成長する。これらの結合されたセルは、次に、非フォト領域として分類される。同様に、フォト領域は、フォト・セルである可能性の高いセルから始まり、まだ分類されていないが、同様の特性を有する隣接セルを結合することによって成長する。

40

【0015】

グレイレベル分布度に加えて、他の特徴を利用することによって、セルがフォト・セルであるか、非フォト・セルであるかを判定することができる。利用可能な他の特徴には、セルの共起行列(co-occurrence matrix)の慣性モーメントがある。共起行列は、別のグレイレベルのピクセルに隣接したピクセルの各グレイレベル毎のピクセル数を示す2次元行列である。2進領域は、一般に、大きい黒・白遷移を有する。これらの領域では、慣性モーメントが大きくなる傾向がある。一方、フォト領域は、グレイレベルの遷移がより平滑になる傾向がある。これらの領域では、慣性モーメントが小さくなる傾向がある。

50

【0016】

文書内の領域がフォト領域、又は非フォト領域に分類されると、各領域毎に、イメージを表すのに用いられるビット数が選択されうる。非フォト領域の場合、単純な2進表現で十分である。換言すれば、各ピクセル毎に、1ビットが用いられる。例えば、白の背景に黒のテキストを有する領域の場合、各ピクセル毎に1ビットを用い、そのピクセルが黒か白かを表示することが可能である。フォト領域の場合、もっと多くのビットを利用することによって、黒と白の間のグレイレベルを追加することが可能になる。明らかに、フォト領域内の各ピクセルを表すために利用するビットが増えると、フォト領域内のグレイレベル(又は、色)を表す上での柔軟性が大きくなる。従って、デジタル文書を表すために必要なビット数は、各データ・タイプを十分に表すのに必要とされるビット数のみが提供 10

【0017】

本発明のこれ以外の特徴、及び利点、並びに本発明の各種実施例の構成、及び作用については、添付の図面に関連して詳細に後述する。

【0018】

本発明が添付の図面に関連して記述される。これらの図面中では、同じ参照番号は同一のものか、又は機能的に同様の要素を示している。

【0019】

【実施例】

1.0 発明の概要、及び解説 20

本発明は、デジタル・スキャナにスキャンされ、4ビット・グレイスケールで記憶されるページ内のフォト領域の検出を目的とするものである。フォト領域の検出を行う目的は、これらの領域と非フォト領域を別個に処理できるようにすることにある。ほとんどの場合、フォト領域は非フォト領域と別個に取り扱われる必要があり、その中に含まれる情報を保持する。例えば、ライン・アート、又はテキストの様な非フォト領域は、適合する固定、又は可変局部閾値でグレイスケールの閾値処理を行うことによって、正確に表されうる。一方、フォト領域は、正確に表現するためにハーフ・トーン処理が行われる。同様に、ライン・アート、又はテキスト・データに関して有効なデータ圧縮技法は、フォト・データに利用するには適しておらず、特殊な技法を用いなければならない。

【0020】 30

本書に開示される方法は、OCR(ページ分解の一部として)、イメージング(例えば、ユーザ選択、及び編集に関するページ内からの「フォト」タイプの自動検出)、及び文書管理(必要な場合に、グレイスケールを保持し、圧縮することによって、品質、及び有効な記憶域の処理を行う)といった分野に適用可能である。

【0021】

好適実施例において、本発明の方法は3つの主要ステップに分割することができる：

ステップ104-特徴抽出；

ステップ108-初期段階でフォト・セル、及び非フォト・セルの可能性の高いセルの判定；

ステップ112-類似性測定、及び局部コンテキストに基づく領域成長。

【0022】 40

これら3つのステップが図1に示されている。各ステップについては、後続のセクションにおいて詳述する。明確には、ステップ104はセクション2.0で解説し、ステップ108はセクション3.0で解説し、ステップ112はセクション4.0で解説する。

【0023】

これらのステップはデジタル・イメージに適用される。デジタル・イメージは、テキスト、ライン・グラフィックス、及び写真又は写真と同様のものの領域のうち任意の領域、又は全領域から構成可能な文書(例えば、ページ)のイメージである。このイメージは、デジタル・スキャナを利用したスキャン、又はデスクトップ・パブリッシング・システムを利用した生成の結果として、デジタル・フォーマットとすることが可能である。

【0024】 50

2.0 特徴抽出

好適実施例において、デジタル・イメージがセルに分割される（後述）。全てのセルについて特徴が抽出され、各セルがフォト・セルである、即ちフォト領域の特性を備えている可能性が判定される。

【0025】

セルは全て、フォト・セル、非フォト・セルのラベルが付けられるか、又は分類されない。分類されたセル（即ちフォト・セル、及び非フォト・セル）は、非分類セルが「成長」する元となる種子として用いられる。

【0026】

次に、特徴抽出のステップ104について記述する。ステップ104において、スキャンされた文書の個々の領域がフォト領域である可能性を示すことができる特徴が、スキャンされた文書から抽出される。この特徴には、グレイレベル分布度、共起行列の慣性モーメント、行プロファイル、及び列プロファイルを含めることができるが、これらに限定されるわけではない。

10

【0027】

図6は、好適実施例において特徴の抽出を行う方法を示した動作流れ図である。図6を参照すると、この実施例においてスキャンされた文書の全てのセルに関して特徴が抽出される。300dpiのスキャン・ページの場合、64x64ピクセルのセル・エリアを選択することができる。これは約5.08mmx5.08mmに等しい。このフォーマットは、ヒューレット・パカード社のAccupageTMスキャナの自動強度(autointensity)アルゴリズムにおいて用いられているセルのサイズに一致する様に選択されている。当該分野の技術者であれば、本明細書に解説の発明が、他のフォーマットによっても実現可能なことが明らかである。

20

【0028】

好適実施例では、4ビット・グレイスケールのデータが用いられる。これは、スキャナからホスト・コンピュータへのデータ転送時間を最小にし、同時にフォト領域を検出し、更に、適度な品質でそれを保存するのに十分な情報を保持するために行われる。8ビット・トーン・マッピング（256色、又はグレイ・シェード）を使用するスキャナに関しては、8ビット・フォーマットから4ビット・フォーマットへの変換を行う必要がある。

【0029】

ステップ504では、元の8ビット・データに不均一な量子化を施すことによって、4ビット・データが得られる。人間の目の応答に最も効果的に適応させるため、この変換には、明るいレベルが暗いレベルに比べてより広くなる様に間隔をあける、8ビットから4ビットへのトーン・マッピングが含まれる。これは、本質的に8ビット反射率データから4ビット「明度」データ L^* （ L^* abカラー・スペースで用いられる）への変換である。 L^* abは、当業界において既知の多くの規格の1つであるカラー・スペースである。 L^* abは、8ビット・グレイスケール・データを4ビット・グレイスケール・データに量子化する場合に用いられるものであるため、通常使用されない。下記の表1には、このトーン・マッピングが示されている。

30

【0030】

【表1】

40

8ビット反射率	4ビットL*
0 - 63 (白)	15 (白)
64 - 92	14
93 - 119	13
120 - 142	12
143 - 163	11
164 - 180	10
181 - 196	9
197 - 209	8
210 - 220	7
221 - 229	6
230 - 236	5
237 - 242	4
243 - 246	3
247 - 249	2
250 - 251	1
252 - 255 (黒)	0 (黒)

10

【 0 0 3 1 】

ステップ508において、スキャンされたセルに関して、1次元ヒストグラムが作成される。ヒストグラムは、各グレイレベル内のセルにおけるピクセル数を示している。例えば、セル内の238個のピクセルが、グレイレベル12である場合、ヒストグラムの12番目の要素に値238が割り当てられる。ヒストグラムは、各ピクセルを個々に調べ、そのグレイレベルに対応するヒストグラムの要素をインクリメントすることによって、簡単に作成することが可能である。ステップ508で作成される1次元ヒストグラムは、各グレイレベルにおけるセル内のピクセル数を示しているため、「グレイレベル・ヒストグラム」と呼ばれる。

20

【 0 0 3 2 】

2.1 フォト領域の判定に用いられる特徴

2.1.1 グレイレベル分布度

非フォト領域は、通常間隔の広い、狭く集中した2つのグレイレベルを示す。例えば、テキストは、2つのピーク、即ち前景のピークと背景のピークとしてヒストグラムに示される。大部分のデータは、各ピークのまわりに狭い範囲で分布している。白、又はカラー・ページの様な極めて均一な背景は、まわりに狭い範囲でグレイが分布した単一のピークとして示される。同様に、暗く、狭い幅で定義されたテキストは、同様に単一のピークとして示される。

30

【 0 0 3 3 】

一方、フォト領域は、1つ以上のピークのまわりに広く分布したグレイレベルとして示される。従って、グレイレベル分布度は、セルがフォト領域に含まれるか、非フォト領域に含まれるかを示す有効な指針である。グレイレベル分布度はステップ512において決定される。

40

【 0 0 3 4 】

図7は、好適実施例に基づいて、グレイレベル分布度を決定する方法を示した動作流れ図である。ここで図7を参照すると、ステップ604において、ヒストグラムのピークが決定される。セル内の複数グレイレベルから成るヒストグラムは、以下の様に定義することができる：

$histo[K], K=0, \dots, 15$

ピークは、以下の様に定義することができる：

以下の2つの条件が満たされる場合、Kはピークである ($0 \leq K \leq 15$) :

(i) $histo[K-1] < histo[K] > histo[K+1]$;かつ

(ii) $2histo[K] - histo[K-1] - histo[K+1] > histo[K]/10$

50

あるいは、以下の2つの条件が満たされる場合、Kはピークである：

(i) $histo[K-1] < histo[K] > histo[K+2]$ (ここで、 $histo[K] = histo[K+1]$);かつ

(ii) $2histo[K] - histo[K-1] - histo[K+2] > histo[K]/10$

この定義は、ヒストグラムにおける幅 = 2ピン(グレイレベル2つ分)の「フラットな」ピークの存在を考慮したものである。ヒストグラムの端における効果を考慮すると、ヒストグラムの両端の値は、次の様に定義される：

$histo[-1] = 0$;及び

$histo[16] = histo[17] = 0$ 。

【0035】

ステップ608において、ヒストグラムの谷が決定される。谷は2つのピークの間で最小グレイレベル値として定義される。換言すれば、グレイレベル P_1 、及び P_2 が連続したピークとすると、以下の条件を満たせば、 $P_1 < V_1 < P_2$ の場合に V_1 が谷であるということができる：

$histo[V_1] \leq histo[i]$ ($P_1 < i < P_2$ の全ての*i*に対して)。

【0036】

ステップ612において、ヒストグラムの部分母集団、即ちモードが決定される。モードは、ピーク、及びその左右の2つの谷のそれぞれによって定義される。ピークは、 $peak[K]$ と書くことができる。2つの谷は、左の谷、及び右の谷に関してそれぞれ $val[K]$ 、及び $val[K+1]$ と書くことができる。この用語を利用すると、モードKは、ヒストグラムにおける $val[K]$ から $val[K+1]$ までの領域として定義される。

【0037】

図8は、例示のヒストグラムに関して、モードを求める方法を示す図である。ここで図8を参照すると、この例では、4つのモード704、708、712、及び716が存在する。それぞれのモードがピークと、そのまわりの2つの谷によって定義される。ヒストグラムの最初と最後の谷 $val[0]$ 、及び $val[4]$ が、それぞれ次の様に定義されるという点に留意されたい：
 $histo[val[0]] = histo[val[4]] = 0$ 。

【0038】

ステップ616において、各モードの確率が決定される。各モードの確率は以下の式より得られる：

【0039】

【数1】

$$p(K) = \frac{0.5histo[val[K]] + \sum_{val[K]+1}^{val[K+1]-1} histo[i] + 0.5histo[val[K+1]]}{\sum_0^{15} histo[i]}$$

【0040】

ステップ620において、各モードの分布度が決定される。分布度を定義するため、ピークから、ピークの値の1/4に当たるポイントまで外側に、モードの測定が実施される。図9は、モードの分布度の例を示す図である。ここで図9を参照すると、モードは、ピークの左右に、それぞれ2つの値 $X1$ 、及び $X2$ を有している。

【0041】

$X1$ は、モードの分布度の左の境界を定義し、 $X2$ は、モードの分布度の右の境界を定義している。左の谷とピークの間位置する $X1$ (即ち、 $val[2] \leq X1 \leq peak[2]$)は、以下の様に表される：

(i) $histo[X] \geq histo[peak[2]]/4$ ($X1 \leq X \leq peak[2]$);及び(ii) $histo[X] < histo[peak[2]]/4$ ($val[2] \leq X < X1$)

10

20

30

40

50

histo[val[2]] \geq histo[peak[2]]/4のため、(ii)を満たすことができない場合、X1は、val[2]に等しくなる様にセットされる。

【0042】

同様に、ピークと右の谷の間に位置するX2(即ち、peak[2] \leq X2 \leq val[3])は、以下の様に表される：

(i) histo[X] \geq histo[peak[2]]/4 (peak[2] \leq X \leq X2)(ii)histo[X] $<$ histo[peak[2]]/4 (X2 $<$ X $<$ val[3]) 同様に、histo[val[3]] \geq histo[peak[2]]/4のため、(ii)を満たすことができない場合、X2はval[3]に等しくなる様にセットされる。

【0043】

これらの端点を利用すると、peak[2]に対応するモードの分布度は、以下の様に定義される：

sd[2] = X2 - X1 + 1。

【0044】

ステップ624において、各モードに関連した「重複確率」が定義される。図10には、モードの重複確率が示されている。ここで図10を参照すると、peak[i]に関連した重複確率pov[i]は、以下の様に定義される：

【0045】

重複確率によって、部分母集団間、即ちモード間の分離の尺度が得られるという点に留意されたい。図11、及び12には重複確率によってどの様にこの分離が示されているかを表している。ここで図11、及び12を参照すると、図11に例示のヒストグラムにおける各モードの重複確率は、ゼロのヒストグラム値を有している。

【0046】

重複確率によって、部分母集団間、即ちモード間の間隔の測定値が得られるという点に留意されたい。図11、及び12には重複確率によってどの様にその間隔が示されているかを表している。ここで図11、及び12を参照すると、図11に例示のヒストグラムにおける各モードの重複確率は、ゼロのヒストグラム値を有している。

【0047】

一方、図12に示すヒストグラムにおける各モードの重複確率はゼロではない。図12の点線は、モードがどの様に重複するかを示している。

【0048】

ステップ628において、ヒストグラムのグレイレベル分布度が決定される。グレイレベルG_sは、以下の様な式で得られる：

【0049】

【数3】

$$G_s = \sum_i sd[i](p[i] + pov[i])$$

【0050】

ここで、sd[i]はモードiの分布度であり、p[i]はモードiの確率であり、pov[i]はモードiの重複確率である。

【0051】

ヒストグラムに関するグレイレベル分布度を表すため、2、3の例を示すことにする。示される例は、グレイレベル分布度に対して重複確率が及ぼす効果を示すために選択されたものである。

【0052】

例1

図13には、ヒストグラムに関するグレイレベル分布度を決定する1つの例が示されている。ここで図13を参照し、G_sに関する上記式に値を代入すると、次の様になる：

10

20

30

40

50

$$G_s = 3(0.5 + 0) + 3(0.5 + 0) = 3。$$

【 0 0 5 3 】

例 2

図14には、ヒストグラムに関するグレイレベル分布度を決定する第2の例が示されている。ここで図14を参照し、 G_s に関する上記式に値を代入すると、次の様になる：

$$G_s = 3(0.5 + 0.25) + 3(0.5 + 0.25) = 4.5$$

重複確率の項がなければ、例1、及び例2はグレイレベル分布度が全く同じになる。しかし、ヒストグラムのレベル2のピークは、より接近していて、かつ/又は幅が広いため、それらはより大きなグレイレベル分布度を有することになる。例2の、より大きなグレイレベル分布度（即ち、より「写真」の様な）は、このデータと同様のデータを有するものと視覚的認識が一致する。

10

【 0 0 5 4 】

丁度前述したグレイレベル分布度の特徴は、セル・ヒストグラムに基づいている。従って、グレイレベル分布度は、文書内のスキャンされた各セル毎に定義されうる。抽出されることも可能な、スキャンされ、デジタル化された文書の追加的特徴もある。これらの中には、セル・ヒストグラムに基づくものもある。セル・ヒストグラムに基づく特徴のうち、グレイレベル分布度がおそらく最も重要である。グレイレベル分布度は、2進（2色 - 例えばテキスト、及び線画）イメージ、及びコントーン（即ち、写真状の）イメージとの識別に用いられる、主要な特徴である。

【 0 0 5 5 】

20

2.1.2 代替、又は補足の特徴抽出

上述のグレイレベル分布度の特徴は、セル・ヒストグラムに基づいている。セル・ヒストグラムに基づかない追加的な特徴も多数存在する。ここで、これらのいくつかについて記述する。これらの特徴は、グレイスケール分布度の代わりに、又はそれに加えて抽出することが可能である。

【 0 0 5 6 】

2.1.2.1 共起行列の慣性モーメント

セル・ヒストグラムに基づかない第1の特徴は、対角線の周りのセルの共起行列の慣性モーメントである。ステップ516では、慣性モーメントの特徴を抽出するオプションのステップが行われる。共起行列の慣性モーメントを決定するため、ページの各セル毎に共起行列が作成される。上述の様に、セルはスキャンされた文書の、事前に定義された部分である。セル・サイズは変更可能であり、性能/品質のトレード・オフを行うことにより選択される。セル・サイズが、より小さく指定されれば（つまり、ページ毎により大きいセルの数を有する）、各ページ毎により多くの閾値決定が行われうる。この指定は、ページのコントラストがページ全体にわたって均一でない時に有益な場合がある。

30

【 0 0 5 7 】

次に、共起行列、及びその作成方法について、より詳細な説明を行うことにする。図2は、共起行列のフレームワークを示す図である。図3は、共起行列の作成に使用するための、各ピクセル毎のグレイレベル値を対にする方法を示した図である。図4は、共起行列の作成が含まれるステップを示す動作流れ図である。

40

【 0 0 5 8 】

ここで図4を参照すると、ステップ304において共起行列が作成される（又はメモリから検索される）。次に図2を参照すると、共起行列200はグレイレベルの数に等しい数の行、及び列を有する行列である。従って、ピクセル毎に4ビットを使用するシステムの場合、共起行列200は16x16の対称行列である。各行は行番号208によって識別され、各列は列番号212によって識別される。

【 0 0 5 9 】

各行、及び列の交点は、共起行列200の要素204である。例えば、図2では、3つの要素204が強調表示されている：即ち、行7と列9の交点には要素204Aが、行9と列10の交点には要素204Cが、行10と列11の交点には要素204Bが示されている。これらの要素は、例えば

50

それぞれ(7,9)、(9,10)、及び(10,11)の様に、その座標によって参照することも可能である。

【0060】

このステップ304で作成される共起行列200は、各要素204における初期値(例えばゼロ(0))を有する空行列である。

【0061】

ステップ308において、ピクセル・グレイレベル対が決定される。次に図3を参照すると、文書がスキャンされた時、結果として得られるピクセル・パターンはピクセル・グレイレベル対220と呼ばれる、隣接するピクセル・グレイレベル値224の対にすることが可能である。この対は、第1のグレイレベル値224のピクセルが、第2のグレイレベル値224を有するピクセルに隣接していることを示している。図3に示す例には、2つのピクセル・グレイレベル対が示されている。第1の対220は、グレイレベル224のピクセル7が、グレイレベル224のピクセル9に隣接することを示している。第2の対は、グレイレベル224のピクセル10(16進数でA)が、グレイレベル224のピクセル11(16進数でB)に隣接することを示している。

10

【0062】

図3に示す様な、ピクセルのグレイレベル224を対にするグルーピング操作は、本発明の可能な1つの実施例を記述したに過ぎない。追加、及び/又は代替グルーピング操作を実施することが可能である。例えば追加の対220は、グレイレベル224のピクセル9とグレイレベル224のピクセル10(16進数でA)を使用して定義することが可能である。発明者の判断によれば、多くのアプリケーションでは、この追加の対を含むことによって得られるスキャンされたものの品質の上限によって、対220の数を増加させるといったことを取り扱う後続ステップで、必要な追加処理時間を正当化できない可能性がある。

20

【0063】

ステップ312において、共起行列200はステップ308において決定されたピクセル・グレイレベル対220を使用して更新される。次に、この実施方法について記述する。上述の様に、各ピクセル・グレイレベル対220は、グレイレベル値224のピクセルがもう1つのグレイレベル値224のピクセルに隣接していることを示している。対220を形成するこれら2つのグレイレベル224が同じ、即ち隣接ピクセルが同じグレイレベル224を有する可能性がある点に留意されたい。図3に示す例の場合、1つの対220は、グレイレベル224のピクセル7がグレイレベル224のピクセル9に隣接することを示している。この7、及び9による対220に関して、共起行列200の要素204A(即ち要素(7,9))がインクリメントされる。同様に、各グレイレベル対220(例えばX、及びY)毎に、行X、列Yの交点に対応する要素204(例えば要素(X,Y))がインクリメントされる。

30

【0064】

従って、各X、Yの交点に関する各要素204は、グレイレベル224のピクセルXが、グレイレベル224のピクセルYに隣接する回数を示す数が含まれている。

【0065】

対220は、隣接するグレイレベルを示しているので、(X,Y)によって示される対は、対(Y,X)と同じである。例えば、グレイレベル224の9に隣接したグレイレベル224の7は、グレイレベル224の7に隣接したグレイレベル224の9と同じである。従って、ステップ110(再び図1参照のこと)において、共起行列200は、共起行列の主対角線(点線232で表示)上、及びその上方の要素204だけがインクリメントされるように折り畳まれる。例えば、対(7,9)及び(9,7)の場合、要素204Aはインクリメントされるが、要素214はインクリメントされない。従って、主対角線上、及びその上方の要素204におけるデータだけしか処理する必要はない。

40

【0066】

図5には、共起行列に挿入された、セルからのデータを有する共起行列200の例が示されている。ここで図5を参照すると、この例ではグレイレベル3のピクセルがグレイレベル4のピクセルに隣接することが7回あり、グレイレベル3のピクセルがグレイレベル5の

50

ピクセルに隣接することが2回あり、グレイレベル4のピクセルがグレイレベル4のピクセルに隣接することが103回あり、グレイレベル5のピクセルがグレイレベル4のピクセルに隣接することが95回ある等々。図5に示す様に、データが入力されるのは行列200の主対角線上、及びその上方の要素204だけである。

【0067】

要するに、共起行列の各要素204はグレイ値*i*がグレイ値*j*に隣接することになる回数として定義される。共起行列200の対角線に関する慣性モーメント M_I は、次の様に定義される：

ここで、 C_{ij} は、共起行列の行*i*、及び列*j*の要素である。

【0068】

【数4】

$$M_I = \frac{\sum_i \sum_j (i-j)^2 C_{ij}}{\sum_i \sum_j C_{ij}}$$

10

【0069】

計算上同様の代替定義は、以下に示す通りである：

【0070】

【数5】

$$M_I^* = \frac{\sum_i \sum_j |i-j| C_{ij}}{\sum_i \sum_j C_{ij}}$$

20

【0071】

2進領域は、通常黒・白遷移が大きく、従ってこれらの領域では、 M_I が大きくなりやすい。一方、コントーン領域は、グレイレベルの遷移がより平滑になりやすく、従ってこれらの領域では、 M_I が小さくなる傾向がある。 M_I が、共起行列の対角線に関する分布度の測定値であることに留意されたい。この測定値は、更に次の様に単純化することが可能である：

【0072】

【数6】

$$\max_{i,j} |i-j|$$

30

40

【0073】

$C_{ij} > 0$ である全ての*i*、*j*について。

【0074】

2.1.3 追加セル・ヒストグラムに基づく特徴

セルを記述するには、本セクションにおいて前述のグレイレベル分布度の特徴に加え、ヒストグラムに基づくいくつかの他の特徴も有効である。セルのヒストグラムに基づくこれらの特徴は、セルがフォト領域に含まれる可能性を判定するのに有効である。これらの特徴は、ステップ524において決定されるが、以下でその説明を行う。

【0075】

50

2.1.3.1 範囲

範囲は、次の様に定義される：

グレイレベルの最大値 - グレイレベルの最小値 + 1

ここで、グレイレベルの最大値、及びグレイレベルの最小値は、ヒストグラムから簡単に得られる。グレイレベルの最大値は、ヒストグラムにおける非ゼロ要素を備えたものの中で最も高いグレイレベルである。逆に、グレイレベルの最小値は、非ゼロ要素を備えたものの中で最も低いグレイレベルである。

【0076】

2.1.3.2 暗さ率

暗さ率は、以下の様に定義される：

黒のピクセル数 / 全ピクセル数 × 100

黒のピクセル数は、閾値を用いた決定の後で、ヒストグラムから求めることができる。閾値は、2レベル・ヒストグラムに関して、閾値を超えるグレイレベルの全ピクセルが黒になり、閾値未満の全ピクセルが白になる様に定義することが可能である。

2レベル・ヒストグラムに関して、この閾値はピークの間の谷とすることができる。

【0077】

2.1.3.3 ヒストグラム形状

ヒストグラム形状は、セルがフォト領域に含まれる可能性を検出するために利用し得る有効な特徴である。図20は、ヒストグラム形状の特徴が抽出される方法を示す動作流れ図である。ここで図20を参照すると、ヒストグラムは、1レベル、2レベル、及び3（又はより高い）レベルに分類することができる。更に、1レベル、及び2レベル・ヒストグラムは弱、又は強に分類することも可能である。これらの属性のおのおのについては、別個に記述する。

【0078】

ステップ1704において、ヒストグラムのレベルが決定される。図21の動作流れ図には、ヒストグラムのレベルを決定する方法が示されている。ここで図21を参照すると、ヒストグラムが正確に1つのピークを備えている場合（決定ブロック1804）、ブロック1808に示す様に、そのヒストグラムは1レベルのヒストグラムである。

【0079】

正確に2つのピークがある場合（判定ブロック1812）、ブロック1816に示す様に、そのヒストグラムは2レベルのヒストグラムである。暗い方のグレイレベルのピークは「DARK-PEAK」と呼ばれ、明るい方のグレイレベルのピークは「BRIGHT-PEAK」と呼ばれる。

【0080】

3つ以上のピークがある（判定ブロック1812に対する答が否定である）場合、ピークの特性に従って、ヒストグラムは2レベル、又はそれ以上となる可能性がある。この場合、動作は「重要な」ピークの数を決めることによって続けられる。

【0081】

ステップ1820において、3つ以上のピークがあると、最初に最も高い3つのピークが検査される。これら3つの最も高いピークのうち、最も暗いピーク（最低のグレイレベルに相当する）は、DARK-PEAKと呼ばれ、3つのピークのうちの最も明るいピークは、BRIGHT-PEAKと呼ばれ、中間のピークは、MID-PEAKと呼ばれる。中間のピークの高さが最高のピークの高さの25%未満の場合（判定ブロック1824）、ステップ1816に示す様に、ヒストグラムは2レベルであると宣言される。それ以外であれば、ヒストグラムはステップ1828に示す様に3レベル、又はそれ以上のレベルであると宣言される。

【0082】

再び図20を参照すると、ステップ1708で、セルに対して形状ラベルが割り当てられる。1レベル・ヒストグラムの場合、グレイレベル分布度が所定の最小値（以下で定義される）以上の場合、そのセルに対して形状ラベル「MONOLEVEL PHOTO」が割り当てられる。範囲が6以下の場合、そのセルに対してラベル「MONOLEVEL STRONG」が割り当てられる。範囲が6を超えると、形状ラベルは「MONOLEVEL WEAK」になる。

10

20

30

40

50

【 0 0 8 3 】

2レベル・ヒストグラムに関して、グレイレベル分布度が所定の最小値（以下で定義される）以上の場合、そのセルには形状ラベルBILEVEL PHOTOが割り当てられる。それ以外であれば、セルに対して3つのラベル、即ちSTRONG-BILEVEL、MID-STRONG-BILEVEL、又はWEAK-BILEVELのうち1つが割り当てられる。この割り当てを行う方法については、いくつかの重要な用語の紹介に関連して後述する。

【 0 0 8 4 】

用語PEAK-DISTは、BRIGHT-PEAKとDARK-PEAKとの距離として定義される。これは、次の様に書くことができる：

$$\text{PEAK-DIST} = \text{BRIGHT-PEAK} - \text{DARK-PEAK}.$$

10

【 0 0 8 5 】

用語DARK-DISTは、ヒストグラムにおけるDARK-PEAKと最低のグレイレベルとの距離として定義される。これは、次の様に表現される：

$$\text{DARK-DIST} = \text{DARK-PEAK} - \text{GRAY-MIN}.$$

【 0 0 8 6 】

用語BRIGHT-DISTは、ヒストグラムにおけるBRIGHT-PEAKと最低のグレイレベルとの距離として定義される。これは、次の様に表現される：

$$\text{BRIGHT-DIST} = \text{BRIGHT-PEAK} - \text{GRAY-MIN}.$$

【 0 0 8 7 】

値PEAK-DIST-MINはBRIGHT-PEAKの関数であり、以下の表に示す様に割り当てられうる：

20

【 0 0 8 8 】

【表 2】

BRIGHT-PEAK	0-8	9-10	11-12	13-15
PEAK-DIST-MIN	3	4	5	6

【 0 0 8 9 】

最後に、H-DARKはDARK-PEAKの高さであり、H-BRIGHTはBRIGHT-PEAKの高さである。

【 0 0 9 0 】

4つの条件が以下に定義される。これらの条件は、セルに対して、STRONG-BILEVEL、MID-STRONG-BILEVEL、又はWEAK-BILEVELセルのいずれかのラベル付けを行うために使用される。4つの特性、又は条件は次の通りである：

30

(i) $\text{PEAK-DIST} \geq \text{PEAK-DIST-MIN}$

(ii) $\text{PEAK-DIST} \geq 2\text{DARK-DIST}$

(iii) $\text{H-DARK} < \text{H-BRIGHT}/25$

(iv) $\text{BRIGHT-DIST} \geq \text{PEAK-DIST-MIN}.$

【 0 0 9 1 】

条件(i)及び(ii)が満たされると、セルには形状ラベルSTRONG-BILEVELが割り当てられる。条件(i)及び(ii)が満たされなければ、セルには2つの形状ラベルMID-STRONG-BILEVEL、又はWEAK-BILEVELの一方が割り当てられる。

40

【 0 0 9 2 】

条件(iii)及び(iv)が満たされると、セルには形状ラベルMID-STRONG-BILEVELが割り当てられる。これらの条件が満たされなければ、セルにはラベルWEAK-BILEVELが割り当てられる。

【 0 0 9 3 】

図16、及び17にはSTRONG-BILEVELヒストグラム形状とWEAK-BILEVEL形状の違いが示されている。図16は、条件(i)及び(ii)を満たしており、従ってラベルSTRONG-BILEVELが割り当てられる。図17は、条件(i)を満たしておらず、従ってSTRONG-BILEVELは割り当てられない。白地に黒のテキスト、及びライン・アートのヒストグラムは、通常、図16に示すヒス

50

トグラムに似ている。これらのヒストグラムは、比較的大きいPEAK-DIST(1404)だけBRIGHT-PEAKから離れた、DARK-PEAKの狭い分布（即ち、小さなDARK-DIST(1402)）によって特徴付けられる。これは、暗いグレイレベルから中間のグレイレベル、更に明るいグレイレベルへの遷移を表しているので、BRIGHT-PEAKに向かうDARK-PEAKの分布がより広がるものと予測される。

【0094】

MID-STRONG-BILEVELに関する条件(iii)は、セル内において、暗さのレベルの数が明るさのレベルの数に比べて少ない場合に有効である。換言すれば、それは細い線で構成された小さいフォントのことである。

【0095】

1レベル、又は2レベルの宣言がなされないセルは、3レベル以上の形状のヒストグラムとして宣言される。セルが、ある値以上のグレイレベル分布度を有している場合、そのセルに対して形状ラベルTRI-LEVEL PHOTOが割り当てられる。それ以外であれば、ラベルTRI-LEVELが割り当てられる。

【0096】

2.1.4 他の特徴

ヒストグラムのグレイレベル分布度測定値と共起行列の慣性モーメントの両方によって、セル内のグレイレベルの統計的特性に関する情報が得られる。しかし、これらでは、セルの構造、幾何学形状、又は位置特性の特徴に関連した情報はあまり得られない。セルの行、及び列プロファイルに基づく特徴によって、こうした情報が得られる。グレイレベル分布度の特徴に加えて、これらの特徴の計算を行うことも可能である。これらの特徴はステップ520において計算される。

【0097】

$N \times N$ のセルGの行プロファイルRは $N \times 1$ のベクトルとして以下の様に定義される。

【0098】

【数7】

$$R_i = \sum_j G_{ij}$$

【0099】

ここで、 G_{ij} は、セルにおける (i, j) 番目のピクセルのグレイ値である。列プロファイルは $1 \times N$ のベクトルであり、以下の様になる。

【0100】

【数8】

$$C_j = \sum_i G_{ij}$$

【0101】

行、及び列プロファイルは、従ってセル内の2次元情報に関する1次元プロファイルである。これらのプロファイルによって、水平、又は垂直罫線（線）、テキストの行等の様な幾何学的特徴の有無に関する手がかりが得られる。これらのプロファイルを用いて、セル内の水平、又は垂直フォト境界の始め、又は終わりが検出される。

【0102】

64×64 セルの場合、プロファイル・ベクトルは 64×1 と 1×64 である。データ量を更に減少させ、ノイズの低減を実現するために、もとのプロファイルから2つの連続した要素全てをそれぞれ加算して単一要素とし、縮小されたプロファイルを生成することによって、こ

10

20

30

40

50

これらのベクトルは32×1と1×32に縮小される。プロファイル値を最も重要なバイトだけに置き換えることによって、更に単純化することもできる。

【0103】

これらのプロファイルの利用について説明するため、例えばある領域内の情報が、テキストに相当する可能性があるか否かを判定する方法について考察することにする。図15には、行プロファイル内のデータがプロットされるとどの様に見えるかが示されている。行プロファイル値1302は縦座標に沿って、行インデックス1303は横座標に沿ってプロットされる。白地に黒テキストのページの場合、小さい行プロファイル値1302はテキスト行1304を表し、大きい行プロファイル値1302は背景1308を表している。

【0104】

更に、この情報がテキストである可能性を確かめる方法は、現在のセルの行プロファイルと、その左右の隣接セルの行プロファイルを比較することである。ほとんどの場合、セルのテキスト行と背景行の間の境界線は、ほぼ一直線となる。計算によってこれを実施する簡単な方法は、 R_i が8ビット値であると仮定して、以下の様に閾値を選択することである：

【0105】

【数9】

$$\text{閾値} = \max_i \{R_i\} = 16$$

【0106】

現在のセルの上方、及び下方に位置するセルを調べることによって、所定のエリアにおけるテキスト行の幅、及び間隔を決定することが可能である。

【0107】

3.0 フォト領域の可能性が高いセルの判定

フォト検出システムにおける第2のステップ108では、どのセルがフォト領域の一部をなす可能性が極めて高く、どのセルが非フォト領域の一部をなす可能性が高いかを判定する。可能性の判定が済むと、セルはステップ110において、フォト・セル、非フォト・セル、又は非分類セルとして分類、又はラベル付けされる。分類は前記可能性に基づいて行われる。フォト・セル、又は非フォト・セルとして分類されたセルはステップ112で領域成長プロセスを開始するための「種子」として使用される。

【0108】

3.1 フォト・セルの可能性が高いセル

グレイレベル分布度が指定の最小値以上であるセルは、フォト・セルの可能性が高いセルとして指定される。指定の最小値は、「背景」ピークとみなされるピークのグレイレベルの関数であり、ヒストグラムが1レベルであるか否かの関数でもある。

【0109】

ヒストグラムが1レベルの場合、所定のグレイレベル分布度の最小値は、背景ピークの関数である。1レベル・ヒストグラムの場合、背景ピークが唯一のピークである。好適実施例による1レベル・ヒストグラムに関するグレイレベル分布度が、表3に示されている：

【0110】

【表3】

背景ピーク	0-4	5-11	12-14	15
グレイレベル分布度の最小値	2	3	4	3

【0111】

従って、例えば1レベルのセルが、グレイレベルが4の背景ピークを有している場合、グレイレベル分布度(ステップ108において、上述の様に判定される)が2以上であれば、フォト領域である可能性が高い。

【0112】

ヒストグラムが2レベル、又はそれ以上の場合、ある実施例では、背景ピークが、BRIGHT-PEAK、又はDARK-PEAKになる。DARK-PEAKの高さが、BRIGHT-PEAKの高さの3倍を超えると、背景ピークは、DARK-PEAKとみなされる。それ以外であれば、背景ピークは、BRIGHT-PEAKになる。表4には、好適実施例に基づく複数レベルのヒストグラムにおける種々の背景ピークのグレイレベルに関して、どの様にグレイレベル分布度の最小値を指定するかが示されている。

10

【0113】

【表4】

背景ピーク	0-4	5-11	12-15
グレイレベル分布度の最小値	3	4	4

【0114】

従って、例えば2レベル・セルが、グレイレベルが4の背景ピークを有している場合、グレイレベル分布度(ステップ108において、上述の様に判定される)が3以上であれば、

20

フォト領域である可能性が高い。

【0115】

3.2 非フォト領域の可能性が高いセル

非フォト領域の可能性が高い3タイプが、種子、即ち背景の種子、2レベルの種子、及びテキストの種子として使用される。これらの種子については、それぞれセクション3.2.1、3.2.2、及び3.2.3において説明する。

【0116】

3.2.1 背景の種子

背景の種子は、通常均質な背景のエリア(ページ余白)が存在するページの上部から領域成長プロセスを開始することが有効であり、必要である。図18は、あるセルが背景の種子か否かを判定するために使用されるセル領域を表した図である。ここで図18を参照すると、領域1508は複数のセル1504から構成される。この領域は $n \times m$ の領域であり、セル1504は座標 (i, j) によって識別することが可能である。好適実施例では、領域サイズは1セル \times 4セルである。

30

【0117】

特定の (i, j) セルは、以下の2つの条件が満たされる場合、背景の種子である：

(i) 矩形領域1508内における (i, j) 、及びそれに隣接する全てのセル1504が形状ラベルが「MONOLEVEL STRONG」を有している。

(ii) 矩形領域内の任意の2つのセル間におけるピークのグレイレベルの最大差が1である。

40

【0118】

3.2.2 2レベルの種子

2レベルの種子は、ページの任意の場所におけるテキスト、又はライン・アートに関する領域成長を開始するために使用される。再び図18の $n \times m$ のセルから成る矩形領域内のセル (i, j) 1504を参照する。問題となるセル (i, j) 1504は、以下の条件が満たされる場合、2レベルの種子である：

(i) 矩形領域内の (i, j) 、及びそれに隣接する全てのセル(即ち、そのセルに「すぐ隣接するセル」)が、形状ラベル「STRONG-BILEVEL」、「MID-STRONG-BILEVEL」、又は「MONOLEVEL STRONG」を有している。

(ii) $n \times m / 4$ 以上の「STRONG-BILEVEL」が存在する。

50

(iii) 矩形領域の特徴の暗さ率が、全体として30%未満である。

(iv) 矩形領域の任意の2つのセル間におけるピークのグレイレベルの最大差が1である。
好適実施例では、この領域のサイズは3セル×4セルである。

【0119】

3.2.3 テキストの種子

テキストの種子は、フォト領域に分類された領域に極めて近接したセルが、2レベルの種子の条件を満たすが、実際にはフォト領域の一部である場合に生じる、いくつかの曖昧な状況を解決するために使用される。この場合、追加テストが実施される。このテストは、セルがテキストの種子に関する条件を満たすか否かを判定するために実施される。テストで、条件を満たすことが分かれば、セルは2レベル領域を成長させるための種子として使用される。一方、セルがテキストの種子でなければ、2レベル領域を成長させるための種子としては使用されない。セルをテキストの種子として分類するのに必要な条件は行、及び列プロファイルの特徴について上で説明したものと同一である。

10

【0120】

4.0 領域成長

最終ステップであるステップ112では、セルから抽出された特徴に基づいて領域を成長させる。その基本戦略は最初に非フォト領域を成長させ、次にフォト領域を成長させるというものである。フォト領域の成長は、フォト・セル可能性の高いセルから始めて、分類せずに放置されたセルに成長させていくやり方で実施される。

【0121】

20

4.1 非フォト領域の成長

非フォト領域の成長は、非フォト・セルの可能性の高いセルから実施される。この領域の成長は、非フォト・セルの可能性の高いセルの位置を探し、該セルと「同様の」特性（即ち、同様の特徴）を備えた近傍セルとを組み合わせることによって実施される。この成長プロセスは水平、及び垂直方向に実施される。水平方向の成長は、左から右、及び右から左に実施することが可能である。好適実施例では、垂直方向の成長は、上から下に（スキヤンの方向に）行われる。

【0122】

4.1.1 水平方向の左から右

次に、左から右へのスキヤンによってセルを分類する方法について説明する。図19は左から右へのスキヤンを使用した分類のプロセスを示す動作流れ図である。ここで図19を参照すると、ステップ1604において第1のセルを調べ、そのセルが分類されているか否かの判定を行う。セルが分類されていない場合（判定ブロック1608）、ステップ1612においてそのすぐ左のセルが調べられる。このセルが非フォト・セルとして分類され（判定ブロック1616）、両方のセルのBRIGHT-PEAKの差が1以下であれば（判定ブロック1618）、ステップ1620において、そのセルは非フォト・セルとして分類される。これら最後の2つの条件のいずれかが満たされない場合、そのセルは、ブロック1624に示す様に分類されないことになる。領域成長プロセスの開始時では、分類されたセルだけがフォト、及び非フォト・セルの可能性の高いセルであるという点に留意すべきである。

30

【0123】

40

4.1.2 水平方向の右から左

右から左へのスキヤンによってセルを分類する方法は、セルが右から左にスキヤンされるという点を除けば、左から右へのスキヤンに関する上述のものと同じであり、ステップ1612において左の近傍セルではなく、右の近傍セルが調べられる。

【0124】

4.1.3 垂直方向の上から下

上から下へのスキヤンによってセルを分類する方法は、セルが上から下にスキヤンされるという点を除けば、左から右へのスキヤンに関する上述のものと同じであり、ステップ1612において左の近傍セルではなく、上の近傍セルが調べられる。

【0125】

50

4.1.4 領域成長をガイドする境界

フォト領域が形成されるにつれて、前述の行、及び列プロファイルの特徴を使用することによって、セル内における領域の境界が正確に位置付けられ得る。水平方向の境界は、セルの行プロファイルを利用して明らかにされる。垂直方向の境界は、セルの列プロファイルを利用して明らかにされる。イメージ内の既に処理済みの行のセルの中で、分類されたフォト・セルの垂直方向の境界が、水平方向における左から右、及び右から左への非フォト領域の成長をガイドし、あるいは制限するために利用可能である。これは以下の様に実施される。成長のための検査を受けている現在のセルのすぐ上に位置するセルが、西側の境界を通過しているフォト・セルである場合、西側の境界と現在のセルの上に位置するセルが一直線になっている垂直方向の境界、又は端があるかを探すために、現在のセルの列プロファイルが検査される。上記境界、及び端があれば、現在のセルに関して、左から右への非フォト領域の成長が抑止される。同様に、現在のセルの上に位置するセル上の東側の境界を利用して、現在のセルにおける右から左への非フォト領域の成長を抑止することが可能である。

【0126】

5.0 典型的なアーキテクチャ

次に、デジタル・イメージにおけるフォト領域を検出するためのシステムに関する典型的なアーキテクチャについて提示し、説明することにする。このアーキテクチャは、デジタル方式によるイメージ表現の効率を高めるための、本発明の実施方法の一例に関して提示される。この例によれば、デジタル・イメージは、ヒューレット・パッカード社のSCANJETTM スキャナの様なデジタル・スキャナを利用して、文書のページをスキャンすることによって生成される。図22は、デジタル・スキャナとインターフェイスをとった本発明を示すブロック図である。

【0127】

ここで図22を参照すると、デジタル・スキャナ1904を利用して、ページのスキャンを行うことにより、ページのテキスト、及び/又はグラフィック・データがデジタル・イメージ1932に変換される。デジタル・イメージ1932はページを表すピクセルの配列から構成される。各ピクセルは、いくつかのビットを利用して、ピクセルに関するカラー、又はグレイレベルを表現する。この例の場合、デジタル・イメージ1932は、ピクセルあたり8ビットのフォーマットによるものである。

【0128】

オプションの量子化器1908は量子化を実施して、デジタル・イメージ1932のフォーマットを変換する。例えば、1つの実施例では、量子化器1908はピクセルあたり8ビットのフォーマットのデジタル・イメージ1932からピクセルあたり4ビットのフォーマットの量子化デジタル・イメージ1936に変換する。この量子化は、上記表1に示した様に実施される。

【0129】

好適実施例では、デジタル・イメージ1932が、文書がスキャンされるにつれて、連続して量子化器1908に供給される。この実施例では、量子化器1908は、各セルがスキャンされるにつれて、デジタル・イメージ1932を受信する。

【0130】

量子化デジタル・イメージ1936は、特徴決定モジュール1912に入力される。特徴判決モジュール1912は、本明細書のセクション2において記述した様に、デジタル・イメージ1936のセルの1つ以上の特徴1940を決定するために使用される。特徴決定モジュール1912は、特徴の決定に必要なとされる場合、グレイレベル・ヒストグラム、及び共起行列200を生成することが可能である。

【0131】

各セルの特徴1940が、セル分類モジュール1916に供給される。セル分類モジュール1916は、セルがフォト・セル、又は非フォト・セルである可能性を判定し、この可能性の情報1944を領域成長モジュール1920に供給する。

【0132】

10

20

30

40

50

領域成長モジュール1920は、領域をフォト領域、又は非フォト領域として成長させるために使用される。領域成長モジュール1920は量子化デジタル・イメージ1936のセルを受信する。受信したセルが分類されていない場合、領域成長モジュールは、本明細書のセクション4に記述された手順に従って、近傍セルを検査し、分類された領域を成長させる。

【0133】

分類を利用するため、分類された領域1948が閾値モジュール1924に送られる。閾値モジュール1924は非フォト領域に関する閾値を決定し、これらの領域のピクセルあたり複数ビットの表現をピクセルあたり単一ビットのフォーマットに変換する。例えば、閾値モジュール1924は、グレイレベルが11以上のピクセルが白で、グレイレベルが11未満のピクセルが黒であると決定することができる。閾値モジュール1924は、フォト領域に関してピクセルあたり複数ビットのフォーマットを使用し、非フォト領域に関してピクセルあたり単一ビットのフォーマットを使用した、最適化デジタル・イメージ1952を出力する。

【0134】

6.0 結論

本発明の種々の実施例について説明してきたが、これらは単に例示を目的としたものであり、この実施例に制限されるものではないということを理解すべきである。従って、本発明の外延、及び範囲は、上述の代表的な実施例によって制限されるものではなく、請求の範囲、及びその同等物のみに基づいて規定されるべきである。

【0135】

以下に本発明の実施態様を列挙する。

【0136】

1. デジタル・イメージ内の領域をフォト領域、又は非フォト領域として分類する方法であって、それぞれの領域が複数のセルを備え、

(a)前記各セルが、フォト・セルである可能性があるか否かを判定するステップ、

(b)前記セルのそれぞれに対して、フォト・セル、非フォト・セル、又は分類されないセルとしてラベル付けを行うステップであって、前記分類が前記ステップ(a)の結果に基づいて行われ、

(c)前記セルを使用して、フォト領域、又は非フォト領域として領域を成長させるステップから構成される分類方法。

【0137】

2. 前記ステップ(a)の前に、前記セルに関して1つ以上の特徴を抽出するステップが更に含まれ、前記特徴が、前記セルがフォト・セルである可能性を示すものであることを特徴とする、項番1に記載の方法。

【0138】

3. 各ピクセルが、複数のグレイレベルの1つによって表され、前記特徴抽出ステップが、各グレイレベルに含まれる、前記セル内のピクセル数を表したグレイレベル・ヒストグラムを生成するステップから成ることを特徴とする、項番2に記載の方法。

【0139】

4. デジタル・イメージを量子化して、各ピクセルを表すために使用されるビット数を変換するステップを更に含むことを特徴とする、項番3に記載の方法。

【0140】

5. 前記量子化ステップが、人間の目の応答に適應させるため、不均一量子化を使用して実施されることを特徴とする、項番4に記載の方法。

【0141】

6. 前記特徴抽出ステップが、前記セルに関するグレイレベル分布度を決定するステップを更に含み、前記セルに関するグレイレベル分布度を決定するステップが、

(i)前記グレイレベル・ヒストグラムのどのグレイレベルがピークを表しているか、どのグレイレベルが谷を表しているかを決定するステップと、

(ii)前記グレイレベル・ヒストグラムの各ピーク毎に部分母集団を決定するステップと、

(iii)各部分母集団の確率を求めるステップと、

10

20

30

40

50

(iv)各部分母集団の分布度を求めるステップと、
 (v)各部分母集団の重複確率を求めるステップと、
 (vi)ステップ(iii)、(iv)、及び(v)における決定結果を使用して、前記グレイレベル・ヒストグラムのグレイレベル分布度を決定するステップから構成されることを特徴とする項番3に記載の方法。

【0142】

7. 前記グレイレベル・ヒストグラムが、 $histo[K]$ ($K = 0, \dots, M$)によって定義される $M+1$ 個の要素から構成されることと、前記ピークを決定する前記ステップ(i)が、
 $histo[K-1] < histo[K] > histo[K+1]$ 、かつ
 $2histo[K] - histo[K-1] - histo[K+1] > histo[K]/10$
 であるかどうかを判定するステップからなることを特徴とする、項番6に記載の方法。

10

【0143】

8. 前記グレイレベル・ヒストグラムが、 $histo[K]$ ($K = 0, \dots, M$)によって定義される $M+1$ 個の要素から構成されることと、前記ピークを判定する前記ステップ(i)が、
 $histo[K-1] < histo[K] > histo[K+2]$ (ここで、 $histo[K] = histo[K+1]$)、かつ
 $2histo[K] - histo[K-1] - histo[K+2] > histo[K]/10$
 であるかどうかを判定するステップからなることを特徴とする、項番6に記載の方法。

【0144】

9. 前記グレイレベル・ヒストグラムが、 $histo[K]$ ($K = 0, \dots, M$)によって定義される $M+1$ 個の要素から構成されることと、グレイレベル K における前記ピークが $peak[K]$ によって定義されることと、 $peak[K]$ に隣接した2つの谷が $val[K]$ 、及び $val[K+1]$ によって定義されることと、各モードの確率を求めるための前記ステップ(iii)が、

20

【0145】

【数10】

$$p(K) = \frac{0.5histo[val[K]] + \sum_{val[K]+1}^{val[K+1]-1} histo[i] + 0.5histo[val[K+1]]}{\sum_0^M histo[i]}$$

30

【0146】

によって得られることを特徴とする、項番6に記載の方法。

【0147】

10. 非フォト領域、又はフォト領域として領域を成長させる前記ステップ(c)が、最初に、非フォト領域を成長させるステップと、次に他の全ての領域をフォト領域として分類するステップから構成されることを特徴とする、項番3に記載の方法。

【0148】

11. 非フォト領域を成長させる前記ステップが、
 最初のセルを調べて、分類されているか否かを判定するステップと、
 前記最初のセルが分類されていないならば、隣接セルを調べて、前記隣接セルが分類されているか否かを判定するステップと、
 前記隣接セルが非フォト・セルであれば、前記隣接セルの前記グレイレベル・ヒストグラムにおけるBRIGHT-PEAKのグレイレベルと、前記第1のセルの前記グレイレベル・ヒストグラムにおけるBRIGHT-PEAKのグレイレベルとの差が、1グレイレベルを超えるか否かを判定するステップと、
 前記差が1グレイレベル以下であれば、前記第1のセルを非フォト・セルとして分類するステップから構成されることを特徴とする項番10に記載の方法。

40

【0149】

50

12. 前記特徴抽出ステップが、
前記各セル毎に、共起行列を作成するステップと、
前記共起行列の慣性モーメントを決定するステップから構成され、
前記慣性モーメントが、前記セルの黒と白の遷移が大きいか、あるいはグレイレベル遷移が平滑かを示すということの特徴とする項番2に記載の方法。

【0150】

13. (a)デジタル・イメージを、それぞれ複数のピクセルを含むセルに分割するステップと、
(b)各セルがフォト・セルである可能性を判定するステップと、
(c)各セルがフォト・セルである可能性に基づいて、フォト・セル、非フォト・セル、又は分類されないセルとして、各セルにラベル付けを行うステップと、
(d)それぞれ、フォト領域、又は非フォト領域として分類され、複数のセルから構成される領域を成長させるステップと、
(e)各領域の分類に基づいて、各領域に対してピクセル毎にいくつかのビットを割り当てるステップから構成されるデジタル・イメージを表すために使用されるビット数を最適化するための方法。

10

【0151】

14. 前記ステップ(a)の前に、前記デジタル・イメージの量子化を行い、各ピクセルを表すのに用いられるビット数を変更するステップが更に含まれることを特徴とする、項番13に記載の最適化方法。

20

【0152】

15. 前記量子化ステップが、人間の目の応答に有効に適用する様に、前記デジタル・イメージを不均一に量子化するステップであることを特徴とする、項番14に記載の最適化方法。

【0153】

16. 更に、前記ステップ(e)の後で、前記各非フォト領域の閾値レベルを決定し、前記非フォト領域をピクセルあたり1ビットのフォーマットに変換するステップが含まれることを特徴とする、項番13に記載の最適化方法。

【0154】

17. デジタル・イメージを受信し、デジタル・イメージのセルの少なくとも1つの特徴を決定する様に構成された特徴決定モジュールと、
前記特徴決定モジュールに接続され、前記少なくとも1つの特徴に基づいて、前記セルがフォト領域である確率を決定し、前記確率に基づいて、前記セルをフォト領域、又は非フォト領域として分類する様に構成されたセル分類モジュールと、
前記セル分類モジュールに接続され、いくつかのセルから成る領域をフォト領域、又は非フォト領域として分類する様に構成された、領域成長モジュールから構成される、デジタル・イメージ内のフォト領域を検出するためのシステム。

30

【0155】

18. 前記特徴決定モジュールに接続されて、デジタル・イメージのフォーマットを変換する様に構成された量子化器を更に備えていることを特徴とする、項番17に記載の検出システム。

40

【0156】

19. 前記量子化器に接続され、文書をスキャンして、そのデジタル・イメージを生成するデジタル・スキャナを更に備えていることと、複数のピクセルによって前記デジタル・イメージが表され、各ピクセルが複数のビットで表されることを特徴とする、項番18に記載のシステム。

【0157】

20. 更に、前記領域成長モジュールに接続され、非フォト領域の閾値レベルを計算し、前記閾値レベルに基づいて、前記非フォト領域をピクセルあたり1ビットのフォーマットに変換する様に構成された閾値モジュールを更に備えていることを特徴とする、項番17に

50

記載のシステム。

【0158】

【発明の効果】

本発明により、デジタル・イメージを表すために用いられるビット数を、フォト領域、及び非フォト領域の両方について、イメージの品質を低下させることなく最適化することが可能である。

【図面の簡単な説明】

【図1】文書の領域をフォト領域、又は非フォト領域として分類する方法を示した、高レベルの動作流れ図である。

【図2】共起行列の構造を示す図である。

10

【図3】ピクセル・グレイレベル対を示す図である。

【図4】共起行列を作成する方法を示す動作流れ図である。

【図5】共起行列の例を示す図である。

【図6】セルに関して特徴を抽出する方法を示す動作流れ図である。

【図7】セルのグレイレベル分布度を決定する方法を示す動作流れ図である。

【図8】グレイレベル・ヒストグラムの部分母集団、即ちモードを示す図である。

【図9】部分母集団、即ちモード分布度を示す図である。

【図10】部分母集団、即ちモードの重複確率を示す図である。

【図11】各モードの重複確率がゼロである例を示す図である。

【図12】モードが重複する例を示す図である。

20

【図13】グレイレベル分布度が、2つの別個のグレイレベルのピークを示す例を示す図である。

【図14】グレイレベル分布度が、ピークが互いに接近し、より幅広であることを表している例を示した図である。

【図15】行プロファイル値を利用して、テキスト、及び背景の行を示すことが可能な方法を示す図である。

【図16】強2レベル・ヒストグラム形状を示す図である。

【図17】弱2レベル・ヒストグラム形状を示す図である。

【図18】背景の種子を示す図である。

【図19】非フォト領域が成長する方法を示す動作流れ図である。

30

【図20】グレイレベル・ヒストグラムの形状ラベルが決定される方法を示す高レベルの動作流れ図である。

【図21】グレイレベル・ヒストグラムが、1レベル、2レベル、又はそれ以上のレベルのヒストグラムとして分類される方法を示す動作流れ図である。

【図22】デジタル・イメージ内のフォト領域を検出するためのシステムに関する代表的アーキテクチャを示すブロック図である。

【符号の説明】

1904 デジタル・スキャナ

1908 量子化器

1912 特徴決定モジュール

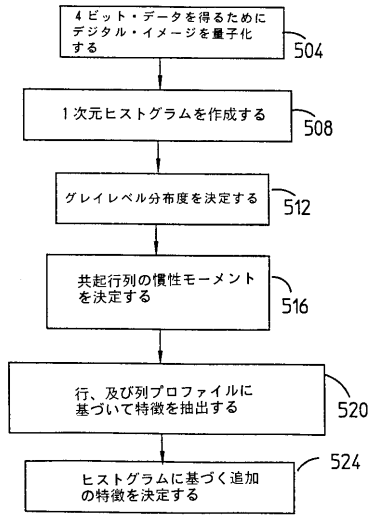
1916 セル分類モジュール

1920 領域成長モジュール

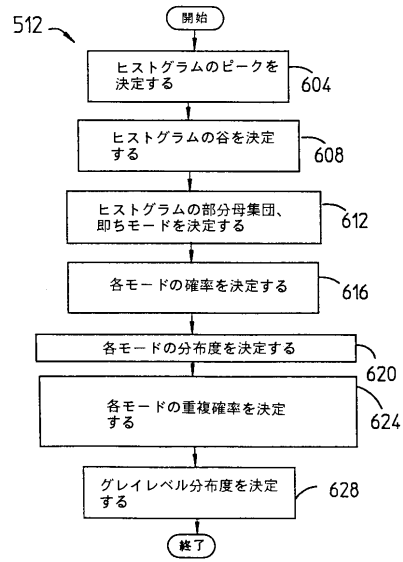
1924 閾値モジュール

40

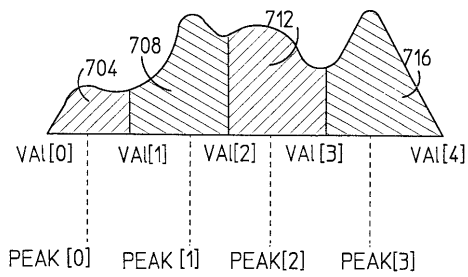
【 図 6 】



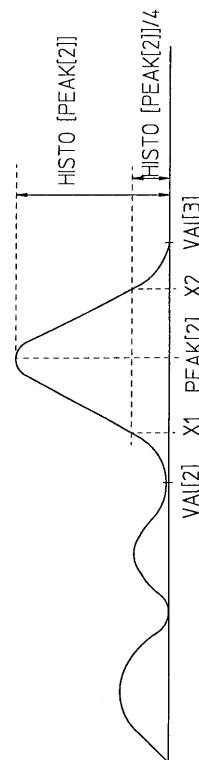
【 図 7 】



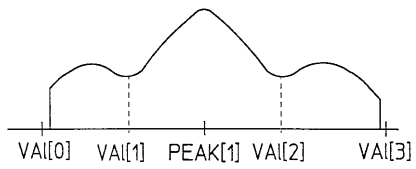
【 図 8 】



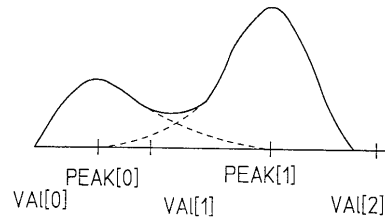
【 図 9 】



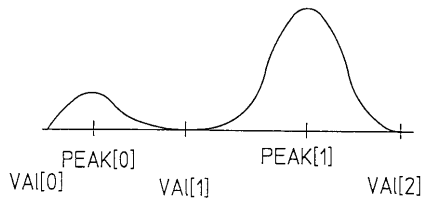
【図10】



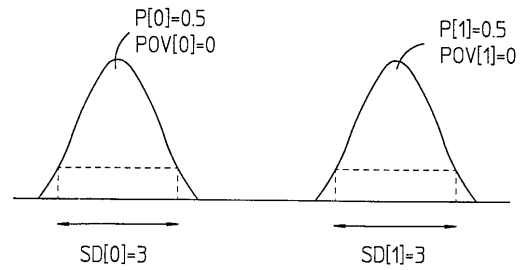
【図12】



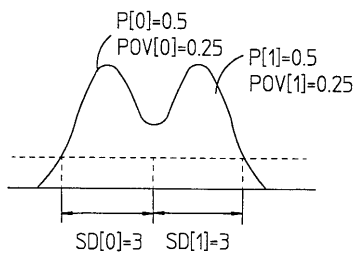
【図11】



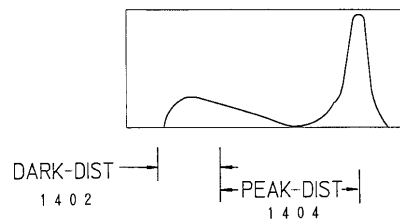
【図13】



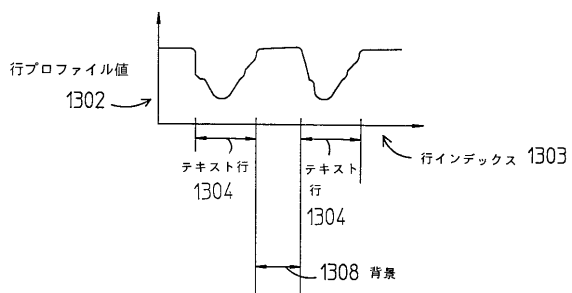
【図14】



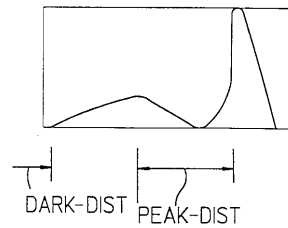
【図16】



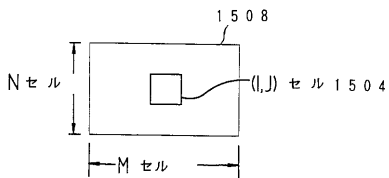
【図15】



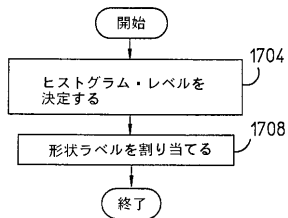
【図17】



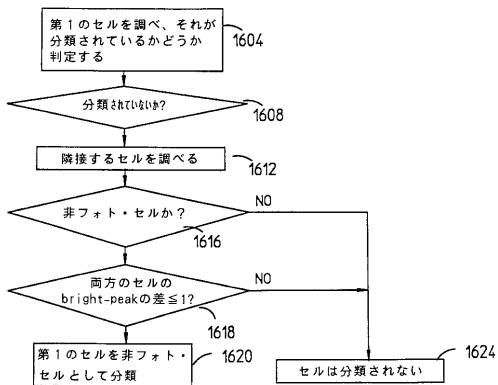
【図18】



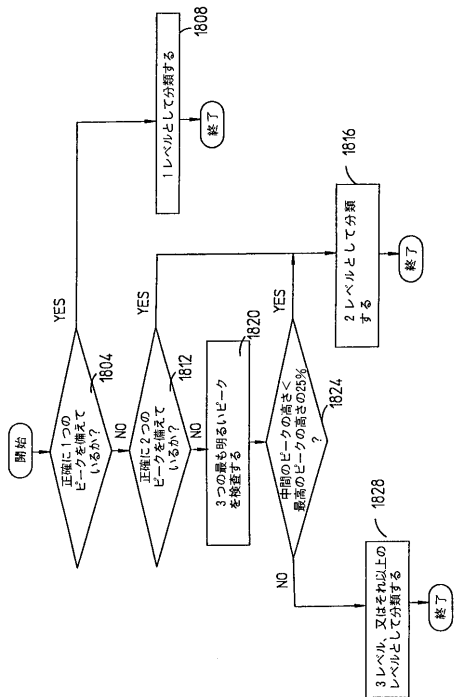
【図20】



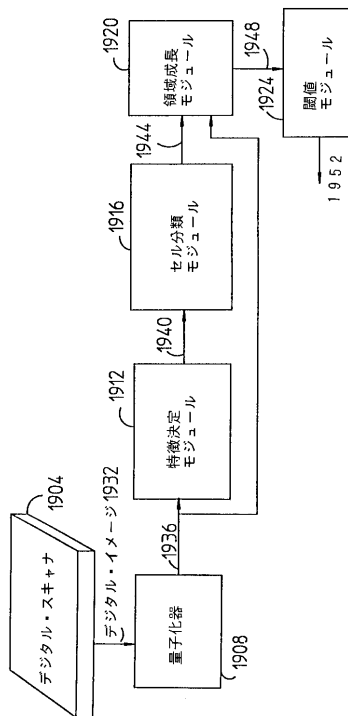
【図19】



【図21】



【図22】



フロントページの続き

(72)発明者 オスカー・エイ・ズニガ
アメリカ合衆国コロラド州80526フォート・コリンズ, ベッドフォード・サークル・1728

審査官 松永 稔

(56)参考文献 特開平05-083545(JP,A)
特開平05-146433(JP,A)
特開平06-319133(JP,A)
特開昭61-225975(JP,A)
特開平07-231387(JP,A)

(58)調査した分野(Int.Cl.⁷, DB名)
H04N 1/40