

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-236800

(P2005-236800A)

(43) 公開日 平成17年9月2日(2005.9.2)

(51) Int. Cl.⁷
H04N 1/40

F I
H04N 1/40

テーマコード (参考)
5C077

審査請求 未請求 請求項の数 8 O L (全 17 頁)

(21) 出願番号 特願2004-45227 (P2004-45227)
(22) 出願日 平成16年2月20日 (2004.2.20)

(71) 出願人 303000372
コニカミノルタビジネステクノロジーズ株式会社
東京都千代田区丸の内一丁目6番1号
(74) 代理人 100090446
弁理士 中島 司朗
(72) 発明者 吉田 宏樹
東京都日野市さくら町1番地 コニカミノルタテクノロジーセンター株式会社内
Fターム(参考) 5C077 MP05 MP06 PP27 PP43 PQ19

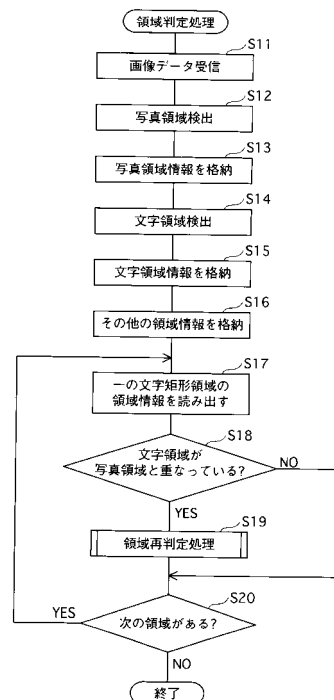
(54) 【発明の名称】 画像領域判定方法、画像処理装置およびプログラム

(57) 【要約】

【課題】 制御手段への負担が少なく済み、処理時間の短縮を図りながら、画像領域の誤判定をできるだけ少なくすることが可能な画像領域判定方法を提供すること。

【解決手段】 受信した画像データに基づいて被判定領域について写真領域を検出する(ステップS12)。次に、当該被判定領域について文字領域を検出する(ステップS14)。検出された文字領域が、検出された写真領域と重なっている場合には(ステップS18で「YES」)、当該重なっている部分について、ステップS14で文字領域を検出するために用いた方法よりも文字判定精度が高い第2の方法を用いて、本当に文字領域であるか否かを再判定する領域再判定処理を実行する(ステップS19)。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

被判定領域内の文字領域と写真領域を判定する画像領域判定方法であって、
前記被判定領域内に存在する写真領域を検出する第 1 のステップと、
前記被判定領域内に存在する文字領域を検出する第 2 のステップと、
検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分
を検出する第 3 のステップと、
重複部分について、前記第 2 のステップで文字領域を検出するために用いた第 1 の方法
とは異なる第 2 の方法を用いて文字領域であるか否かを判定する第 4 のステップと、
を含むことを特徴とする画像領域判定方法。

10

【請求項 2】

前記第 2 の方法は、
前記第 1 の方法よりも文字領域判定の精度が高い方法であることを特徴とする請求項 1
に記載の画像領域判定方法。

【請求項 3】

前記第 4 のステップは、
重複領域に属する各画素データから画像の明度分布を示すヒストグラムを作成し、作成
されたヒストグラムのピーク数、幅、度数の増減量の少なくとも一つを検出して、その検
出結果に基づいて文字領域の判定を行うことを特徴とする請求項 1 または 2 に記載の画像
領域判定方法。

20

【請求項 4】

前記第 4 のステップは、
重複領域を所定量だけ拡張する第 5 のステップを含み、
拡張された領域に属する各画素データに基づいて、文字領域であるか否かの判定を行う
ことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の画像領域判定方法。

【請求項 5】

前記第 5 のステップは、
拡張後の領域の一部が前記検出された写真領域内からはみ出る場合には、そのはみ出る
部分については拡張を行わないことを特徴とする請求項 4 に記載の画像領域判定方法。

【請求項 6】

前記第 5 のステップは、
拡張後の領域の一部が前記検出された写真領域内からはみ出る場合には、拡張すべき領
域を当該写真領域内に収まる方向にシフトさせて所定量だけ拡張させることを特徴とする
請求項 4 に記載の画像領域判定方法。

30

【請求項 7】

被判定領域内の文字領域と写真領域を判定する画像処理装置であって、
前記被判定領域内に存在する写真領域を検出する第 1 の検出手段と、
前記被判定領域内に存在する文字領域を検出する第 2 の検出手段と、
検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分
を検出する第 3 の検出手段と、
重複部分について、前記第 2 の手段が文字領域を検出するために用いた第 1 の方法とは
異なる第 2 の方法を用いて文字領域であるか否かを判定する判定手段と、
を備えることを特徴とする画像処理装置。

40

【請求項 8】

被判定領域内の文字領域と写真領域を判定する画像領域判定処理を実行するコンピュ
ータで動作するプログラムであって、
前記被判定領域内に存在する写真領域を検出する第 1 の処理と、
前記被判定領域内に存在する文字領域を検出する第 2 の処理と、
検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分
を検出する第 3 の処理と、

50

重複部分について、前記第2の処理で文字領域を検出するために用いた第1の方法とは異なる第2の方法を用いて文字領域であるか否かを判定する第4の処理と、

を含む処理をコンピュータに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、文字領域と写真領域を判定する画像領域判定方法、画像処理装置およびプログラムに関する。

【背景技術】

【0002】

近年、スキャナの普及に伴い、スキャナを用いてカラー原稿を読み取り、読み取って得られた画像データを圧縮して保存することが頻繁に行われている。この圧縮方式としては、例えばJPEGが知られており、写真画像等を圧縮するのに圧縮率が高く、画質も良好である。ところが、文字画像のような高周波部分については圧縮率が低下し、画質も劣化することがある。そこで、1頁分の原稿画像について文字、写真等の領域分割を行い、文字領域の画像データをMMR圧縮、写真領域の画像データをJPEG圧縮等することにより、画質をできるだけ劣化させることなく高圧縮率を得る方法が提案されている。

【特許文献1】特開2003-338934号公報

【特許文献2】特開平5-342408号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、従来の方法では、例えば写真領域内に小さな文字画像が存在している場合、その文字が判別されずに単なる写真領域と誤判定されたり、その逆に写真画像の一部なのに文字と誤判定されてしまうといった問題がある。特に、実際には写真画像の一部であるのに文字であると誤判定された部分のデータがMMRにより2値化圧縮されてしまうと、復元画像の画質が劣化、例えば隣接する画素間の微妙な色変化を再現できないといった問題が生じる。

【0004】

文字の誤判定を少なくするには、文字の判定精度ができるだけ高くなる方法を用いれば良い。しかしながら、文字の判定精度が高いということは、その分、文字であるか否かの判定のための処理が増大することになるから、例えば大量の原稿について、1頁ごとに原稿画像の全画素を対象に当該方法を用いて領域判別を行うとすると、その処理のためにCPU等の制御手段の負担が膨大になり、また処理時間も長くなるという問題がある。

【0005】

本発明は、上記の問題点に鑑みてなされたものであって、制御手段への負担が少なく済み、処理時間の短縮を図りながら、画像領域の誤判定をできるだけ少なくすることが可能な画像処理装置、画像領域判定方法およびそのプログラム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】

【0006】

上記目的を達成するため、本発明に係る画像領域判定方法は、被判定領域内の文字領域と写真領域を判定する画像領域判定方法であって、前記被判定領域内に存在する写真領域を検出する第1のステップと、前記被判定領域内に存在する文字領域を検出する第2のステップと、検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分を検出する第3のステップと、重複部分について、前記第2のステップで文字領域を検出するために用いた第1の方法とは異なる第2の方法を用いて文字領域であるか否かを判定する第4のステップと、を含むことを特徴とする。

【0007】

また、前記第2の方法は、前記第1の方法よりも文字領域判定の精度が高い方法である

10

20

30

40

50

ことを特徴とする。

さらに、前記第4のステップは、重複領域に属する各画素データから画像の明度分布を示すヒストグラムを作成し、作成されたヒストグラムのピーク数、幅、度数の増減量の少なくとも一つを検出して、その検出結果に基づいて文字領域の判定を行うことを特徴とする。

【0008】

また、前記第4のステップは、重複領域を所定量だけ拡張する第5のステップを含み、拡張された領域に属する各画素データに基づいて、文字領域であるか否かの判定を行うことを特徴とする。

さらに、前記第5のステップは、拡張後の領域の一部が前記検出された写真領域内からはみ出る場合には、そのはみ出る部分については拡張を行わないことを特徴とする。

10

【0009】

また、前記第5のステップは、拡張後の領域の一部が前記検出された写真領域内からはみ出る場合には、拡張すべき領域を当該写真領域内に収まる方向にシフトさせて所定量だけ拡張させることを特徴とする。

本発明に係る画像処理装置は、被判定領域内の文字領域と写真領域を判定する画像処理装置であって、前記被判定領域内に存在する写真領域を検出する第1の検出手段と、前記被判定領域内に存在する文字領域を検出する第2の検出手段と、検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分を検出する第3の検出手段と、重複部分について、前記第2の手段が文字領域を検出するために用いた第1の方法とは異なる第2の方法を用いて文字領域であるか否かを判定する判定手段と、を備えることを特徴とする。

20

【0010】

本発明に係るプログラムは、被判定領域内の文字領域と写真領域を判定する画像領域判定処理を実行するコンピュータで動作するプログラムであって、前記被判定領域内に存在する写真領域を検出する第1の処理と、前記被判定領域内に存在する文字領域を検出する第2の処理と、検出された文字領域と写真領域の位置情報に基づいて、文字領域と写真領域の重複部分を検出する第3の処理と、重複部分について、前記第2の処理で文字領域を検出するために用いた第1の方法とは異なる第2の方法を用いて文字領域であるか否かを判定する第4の処理と、を含む処理をコンピュータに実行させることを特徴とする。

30

【0011】

ここで、上記写真領域とは、写真だけでなく絵画、図表等、階調性を有する画像領域を示す意味で用いられる。

【発明の効果】

【0012】

これより、例えば第2の方法を第1の方法よりも判定精度が高い(その分CPU等の制御手段への負担が大きくなる)方法とすれば、まず第1の方法を用いて文字領域を判定し、次に誤判定され易い文字と写真の重複部分についてのみ判定精度が高い第2の方法を用いて判定することが可能になる。したがって、原稿1頁分の画像の全画素について第2の方法を用いる構成に比べて、制御手段への処理負担の軽減、処理時間の低減が可能になり、かつ第1の方法だけを用いて判定する構成に比べて、判定精度を向上させることができるという効果を奏する。

40

【発明を実施するための最良の形態】

【0013】

以下、本発明の実施の形態について図面を参照しながら説明する。

(1) 画像処理システムの構成

図1は、本発明に係る画像処理装置を含む画像処理システム(以下、「システム」と略する。)の全体構成を示す図である。

同図に示すように、本システムは、画像処理装置1と、スキャナ2と、ファイルサーバ3とが、ネットワーク、ここではLAN(Local Area Network)4を介して接続されてな

50

り、互いに各種データのやりとりが可能になっている。

【0014】

スキャナ2は、原稿画像を1頁単位で読み取って画像データを得る公知の画像読取装置であり、読み取って得られた画像データを画像処理装置1に送る。

画像処理装置1は、スキャナ2から画像データを受信すると、原稿1頁ごとにその原稿に含まれる文字領域、写真領域、その他の領域をそれぞれ判定する領域判定処理と、文字画像の各画素のデータをMMR (Modified Modified READ)を用いて圧縮し、写真画像およびその他の領域の画像の各画素のデータをJPEG (Joint Photographic Experts Group)を用いて圧縮する画像圧縮処理と、圧縮データに個別のファイル名を付けて、LAN4を介してファイルサーバ3 10

【0015】

ファイルサーバ3は、ハードディスクドライブを有し、画像処理装置1から送信されて来るデータを受信すると、それらをハードディスクドライブに格納する。また、画像処理装置1からのファイルの送信要求があると、それらを読み出して送信する。

本実施の形態の画像処理装置1は、上記領域判定処理において、領域判定結果を再判定する処理を実行するようにしており、この点に特徴がある。以下、画像処理装置1の構成および領域判定処理と画像圧縮処理の内容を詳述する。

【0016】

図2は、画像処理装置1の機能構成を示すブロック図である。 20

同図に示すように、画像処理装置1は、パーソナルコンピュータからなり、主な構成要素として本体部10、ディスプレイ11、キーボード12およびマウス13を備えている。

本体部10は、インターフェース(I/F)部101、制御部102および記憶部103を備えており、これらはバス104を介して接続されている。

【0017】

I/F部101は、LANカード、LANボードといったLAN4に接続するためのインターフェースである。

記憶部103は、ハードディスクドライブ等からなり、OS(オペレーティングシステム)、文書作成等のための各種アプリケーションソフト等が格納されている。 30

制御部102は、主な構成要素としてCPU105、ROM106、RAM107を備えている。

【0018】

CPU105は、記憶部103からOS、アプリケーションソフト等を読み出して、その機能をユーザに提供する等、画像処理装置1の全体の動作を統一的に制御する。また、ROM106に格納されている領域判定処理、画像圧縮処理および送信処理のためのプログラムを読み出して、これらを実行する。

RAM107には、画像データ格納部108、領域情報格納部109および圧縮データ格納部110が設けられ、順に、画像データ、領域情報、圧縮データが格納される。

【0019】

図3は、制御部102による領域判定処理の内容を示すフローチャートである。 40

同図に示すように、制御部102は、スキャナ2から原稿の画像データを受信すると(ステップS11)、受信した画像データを一旦RAM107内の画像データ格納部108に格納させた後、その画像データに基づいて原稿画像に含まれる写真領域を検出する処理を行う(ステップS12)。

【0020】

この写真領域の検出は、公知の方法により行うことができる。例えば、カラーの画像データの場合、原稿1頁分の画像(被判定領域)の各画素について、そのr(赤)、g(緑)、b(青)色の画素データ(256階調)を公知の変換式を用いて明度データに変換し2値化を行う。 50

2 値化画像に対し 4 連結ラベリングを行い、ラベリングされた画像それぞれについて、その大きさが所定の大きさ（文字相当程度の大きさ）よりも大きいものについては、その領域を写真領域とする。一方、小さいものについては、文字または文字の一部とみなして矩形領域の設定を解除する。

【 0 0 2 1 】

例えば、図 4 に示すような入力画像に対し、2 値化処理、写真領域の設定が行われると、写真の全体が写真領域として設定されるが、文字列の部分は写真領域としては設定されないことになる。なお、ここで写真領域として判定される領域には、写真の他、絵画、図表など階調性を有する画像の領域が含まれるものである。

検出された写真領域の座標、ここでは対角の 2 点の X - Y 座標、およびラベリングされた画像を構成する各画素の座標（ラベリングされたラベル値を含む。以下「写真マップ」という。）のデータを写真領域情報として RAM 107 内の領域情報格納部 109 に格納させる（ステップ S 13）。

【 0 0 2 2 】

図 4 の例では、写真領域の座標として A 点（X 1、Y 1）、B 点（X 2、Y 2）のデータが保存され、写真マップとして写真画像の各画素の座標データが保存されることになる。

図 3 に戻って、ステップ S 14 では、文字領域を検出する処理を行う。

この文字領域の検出は、公知の方法により行うことができる。ここでは、原稿 1 頁分の画像（被判定領域）の各画素について、その画素の明度データに公知のエッジフィルタをかけ 2 値のエッジ画像を生成する。生成された 2 値のエッジ画像について、当該エッジ画像を取り囲む矩形領域を設定する。そして、各矩形領域について、その領域内のエッジ画像の局所的な形状をそのエッジ画像の特徴量として抽出し、抽出した特徴量からそのエッジ画像が文字によるものであるか否かを判定する。具体的には、画像の局所的な形状として、カーブ量や傾斜方向、閉ループ数、十字交差数、T 字交差数等を抽出し、抽出された特徴点が、予め保持している文字判定のためのパターンの特徴点と一致している数が所定値（閾値）以上であれば、文字として判定し、所定値よりも少なければ文字ではないと判定する（第 1 の判定方法）。

【 0 0 2 3 】

文字と判定された矩形領域についてのみ、その領域（文字矩形領域）の座標、ここでは対角の 2 点の X - Y 座標、およびその領域内において実際に文字画像（エッジ画像により現される画像）を構成する各画素の座標（以下、「文字マップ」という。）のデータを文字領域情報として領域情報格納部 109 に格納させる（ステップ S 15）。

例えば、図 5 において領域 P が文字矩形領域として検出された場合には、領域 P の座標として C 点（X 3、Y 3）、D 点（X 4、Y 4）のデータと領域 P 内の文字画像の各画素の座標を示すデータが保存される。

【 0 0 2 4 】

また、領域 Q が文字矩形領域として検出された場合（実際に図 4 の写真画像内に文字が入っている場合と、生成された 2 値のエッジ画像に山の稜線等が複雑な細線になって現れそれを文字と誤判定した場合のいずれかになる。）には、E 点（X 5、Y 5）、F 点（X 6、Y 6）のデータと領域 Q 内の文字画像の各画素の座標を示すデータが保存されることになる。なお、写真、文字領域の検出方法については、上記のものに限られることはなく、他の公知の方法を用いることができる。

【 0 0 2 5 】

そして、写真および文字領域以外のその他の領域（下地部分等）の画素の座標のデータをその他の領域情報として領域情報格納部 109 に格納させる（ステップ S 16）。

次に、一の文字矩形領域の座標を領域情報格納部 109 から読み出す（ステップ S 17）。そして、当該一の文字矩形領域が、いずれかの写真領域と重なっているか否かを判断する（ステップ S 18）。この判断は、次のようにして行われる。すなわち、領域情報格納部 109 から各写真領域の座標を読み出し、読み出した写真領域の座標位置のデータと

10

20

30

40

50

、当該一の文字矩形領域との座標位置のデータ（位置情報）に基づいて、両者間の重複部分を検出するものである。例えば、図5の領域Qの場合であれば、写真領域Rと重複していることが検出されることになる。

【0026】

当該一の文字矩形領域がいずれかの写真領域と重なっていることを判断すると（ステップS18で「YES」）、領域再判定処理を行う（ステップS19）。

図6は、領域再判定処理のサブルーチンの内容を示すフローチャートである。

この領域再判定処理は、写真領域に存在していると判定された文字画像が本当に文字画像であるのかを再判定する処理である。すなわち、本実施の形態では、上記したように文字判定の方法として、エッジ画像からカーブ量等を画像の特徴量として抽出し、抽出された特徴量に基づいて文字であるか否かを判定する方法をとっている。この方法によれば比較的CPU105の負担が少なく済ませられるのであるが、例えば実際には写真の一部であるがエッジ画像で見るとたまたま文字の特徴量と一致したため文字と誤判定されることが生じる場合がある。そこで、文字と誤判定され易い写真領域内に文字が重なっていると判断された部分について、第1の判定方法よりも判定精度が高い第2の方法を用いて文字か否かを再判定するものである。

10

【0027】

同図に示すように、写真領域と重なっている文字矩形領域の大きさを所定の範囲だけ拡張する文字拡張領域の設定を行う（ステップS21）。

図7は、図5に示す文字矩形領域Qの大きさを所定の範囲（斜線部）だけ拡張した場合の例を示す図である。図7（a）の例は、文字矩形領域Qを、その主、副走査（図では上下左右）方向に、ここでは10画素分の範囲だけ拡張したものになっている。文字矩形領域と拡張部とを合わせた領域が文字拡張領域になる。

20

【0028】

このように拡張部を設定するのは、次の理由による。すなわち、本実施の形態では、再判定の方法として、文字拡張領域を構成する各画素についての明度ヒストグラムを作成し、その明度ヒストグラムの形状、ピーク数等の特徴量から、一度文字であると判定された画像が本当に文字であるのか、逆に写真画像の一部であるのかを再判定する方法（第2の判定方法）をとっており、この方法の場合、文字矩形領域だけでなく、写真画像の一部分を加えた領域（文字拡張領域）について明度ヒストグラムを作成したほうが、その形状等の特徴量から、文字と写真をより区別し易いからである。

30

【0029】

なお、領域を拡張しようとした際に、その一部が写真領域からはみ出る場合には、図7（b）のように、はみ出る方向への拡張を行わないようにして、文字拡張領域が写真領域内に含まれるように拡張を行う。

そして、設定された文字拡張領域の各画素の明度データから明度ヒストグラムを作成する（ステップS22）。ここでは、256階調の明度データを16階調に変換し、変換後のデータを用いて明度ヒストグラムを作成する。

【0030】

図8（a）は、ある文字拡張領域の16階調明度データに基づく明度の度数分布表の例を示す図であり、図8（b）は、当該明度の度数分布表を基に作成したヒストグラム（ここでは、折れ線グラフで表している。）の例を示す図である。なお、図8（a）に示す度数分布表の「勾配」欄の数値は、ヒストグラム（折れ線グラフ）の勾配の様子を表すための値であり、領域再判定に用いるものである。具体的には、階調数が増えたときに、度数が増加していると「1」としている。例えば階調数が6の場合、5の場合よりも度数が増えているから「1」になる。また、度数が減少していると「-1」、勾配が正から負へ転じる場合に「2」、負から正へ転じる場合に「-2」としている。

40

【0031】

図6に戻って、ステップS23では、ヒストグラム解析処理を実行する。このヒストグラム解析処理は、ステップS22で作成されたヒストグラムに基づいて文字の再判定を行

50

う処理である。

図 9 は、ヒストグラム解析処理のサブルーチンの内容を示すフローチャートである。

同図に示すように、制御部 102 は、まず文字フラグ A と写真フラグ B を共に「0」に設定する（ステップ S31）。

【0032】

そして、作成された明度ヒストグラムからピーク数を計数して、その数が「1」または「6」以上であるか否かを判断する（ステップ S32）。このピーク数の判断は、図 8（a）の勾配欄を参照し、その数値が「2」を示す階調数がいくつあるのかを検出することにより行われる。図 8（a）の例では、ピーク数が「3」ということになる。

ピーク数が「1」または「6」以上であることを判断すると（ステップ S32 で「YES」）、文字として判定された領域は、実際には写真である蓋然性が高いとして、写真フラグ B の値を「1」に設定し（ステップ S33）、その逆に「2」～「5」の範囲内であることを判断すると（ステップ S32 で「NO」）、判定通り文字である蓋然性が高いとして、文字フラグ A の値を「1」に設定する（ステップ S34）。

10

【0033】

このように処理するのは、文字として判定された部分が実際には写真画像の一部である場合には、文字拡張領域の明度ヒストグラムをとると、階調がなだらかであれば、その形状が一つの大きな山型になったり（ピーク数が「1」）、逆に階調変化が多ければピークが多数、ここでは 6 以上になる蓋然性が高くなる。その一方で、実際に文字である場合には、文字自体は 1 色で構成されることがほとんどであり、文字矩形領域のほとんどが当該文字を構成する画素で占められると考えられ、そうすると文字と写真の各部分についてピーク数が 2～5 の範囲内に収まる蓋然性が高くなるからと考えられるからである。

20

【0034】

再判定に用いられる閾値（ここでは、1 と 6）は、予め実験等により写真に文字が含まれる各種画像、含まれない各種画像それぞれについて作成された明度ヒストグラムの形状が実際に文字が含まれる場合と含まれない場合でどのような傾向になるのかが参照され、その結果から写真である蓋然性が高くなると想定されるピーク数として求められたものであり、そのデータは ROM 106 に格納されており、当該処理時に読み出される。

【0035】

次に、ヒストグラムの幅条件解析処理を実行する（ステップ S35）。

30

図 10 は、ヒストグラムの幅条件解析処理のサブルーチンの内容を示すフローチャートである。

同図に示すように、まず閾値を文字拡張領域の全画素数の 0.3（%）にした場合のヒストグラムの幅 W_1 が所定値 TW_1 よりも大きいかが否かを判断する（ステップ S51）。

【0036】

ここで、閾値を全画素数の 0.3（%）にした場合とは、例えば全画素数が 10000 の場合であれば、閾値が 30 ということになり、閾値を 30 にした場合のヒストグラムの幅 W_1 とは、ヒストグラムにその縦軸上の 30 の位置を直交する水平線を引いたときに切断されるヒストグラムの水平方向の幅を示すものである。図 8（b）の例では、 TH_1 が 0.3（%）の場合の閾値を示し、 W_1 がヒストグラムの幅を示している。

40

【0037】

ヒストグラムの幅 W_1 が広いということは、それだけ階調に幅があることを意味し、その画像は写真である蓋然性が高いといえる。そこで、本実施の形態では、ヒストグラムをどの位置（上記閾値に相当）で切断したときの幅 W_1 がどの値（所定値 TW_1 ）を越えれば写真とできる確度が高くなるのかを各種画像を用いた実験から求め、そのデータを ROM 106 に予め格納しておき、本解析処理において読み出して、幅 $W_1 >$ 所定値 TW_1 の場合には（ステップ S51 で「YES」）、実際には写真である蓋然性が高いとして、現在の写真フラグ B の値に「1」をインクリメントし（ステップ S52）、逆に $W_1 < TW_1$ の場合には（ステップ S51 で「NO」）、判定通り文字である蓋然性が高いとして、現在の文字フラグ A の値に「1」をインクリメントして（ステップ S53）、ステップ S

50

54に移る。

【0038】

次に、閾値を文字拡張領域の全画素数の1.6(%)にした場合のヒストグラムの幅W2が所定値TW2よりも大きいか否かを判断する(ステップS54)。この幅W2と所定値TW2は、上記W1とTW1と同様に写真である確からしさを判断するために予め求められてROM106に格納されているものである。図8(b)の例では、TH2が1.6(%)の場合の閾値を示し、W2がヒストグラムの幅を示している。

【0039】

制御部102は、幅W2>所定値TW2の場合には(ステップS54で「YES」)、写真の蓋然性が高いとして、現在の写真フラグBの値に「1」をインクリメントし(ステップS55)、W2<TW2の場合には(ステップS54で「NO」)、文字である蓋然性が高いとして、現在の文字フラグAの値に「1」をインクリメントして(ステップS56)、ステップS57に移る。

10

【0040】

同様に、閾値を文字拡張領域の全画素数の2.1(%)にした場合のヒストグラムの幅W3が所定値TW3よりも大きいか否かを判断する(ステップS57)。この幅W3と所定値TW3も、上記W1とTW1等と同様に写真画像である確からしさを判断するために予め求められてROM106に格納されているものである。図8(b)の例では、TH3が2.1(%)の場合の閾値を示し、W3がヒストグラムの幅を示している。

【0041】

幅W3>所定値TW3の場合には(ステップS57で「YES」)、写真の蓋然性が高いとして、現在の写真フラグBの値に「1」をインクリメントし(ステップS58)、W3<TW3の場合には(ステップS57で「NO」)、文字である蓋然性が高いとして、現在の文字フラグAの値に「1」をインクリメントして(ステップS59)、ステップS60に移る。

20

【0042】

さらに、閾値を文字拡張領域の全画素数の6.5(%)にした場合のヒストグラムの幅W4が所定値TW4よりも大きいか否かを判断する(ステップS60)。この幅W4と所定値TW4も、上記W1とTW1等と同様に写真画像である確からしさを判断するために予め求められてROM106に格納されているものである。図8(b)の例では、TH4

30

【0043】

幅W4>所定値TW4の場合には(ステップS60で「YES」)、写真の蓋然性が高いとして、現在の写真フラグBの値に「1」をインクリメントし(ステップS61)、W4<TW4の場合には(ステップS60で「NO」)、文字である蓋然性が高いとして、現在の文字フラグAの値に「1」をインクリメントして(ステップS62)、ヒストグラム解析処理のサブルーチンにリターンする。

【0044】

なお、閾値を全画素数の何パーセントにし、そのときのヒストグラムの幅がどの値よりも小さければ写真と判定すべきであるのかについては、上記した各条件に限られることはなく、最適と考えられる条件を適用することが可能である。

40

図9に戻って、ステップS36では、ヒストグラムの各ピーク(勾配が正から負に転じる部分)の内、その値(度数)が最大のもとの、2番目のものとの間隔が閾値TH5よりも小さいか否かを判断する。図8(b)の例では、W5が当該間隔を示している。

【0045】

このようにするのは、判定通り文字である場合、通常、文字は1色で構成され、かつ見易いように下地との明度の差(コントラスト)が大きいことがほとんどであり、そうであればヒストグラムに出現するピークの内、文字画像に対応するピークと、背景画像に対応するピークが1、2番目に大きくなり、かつこれらピークは、水平方向に離れた位置に現れることが多いと考えられ、幅W5の大きさを写真であるか否かの再判定の一条件として

50

用いることができるからである。

【0046】

このTH5の値は、文字である蓋然性が高くなると想定されるピーク間隔として予め実験等から求められたものであり、そのデータはROM106に格納されており、当該処理時に読み出されることになる。

幅 $W5 < TH5$ であることを判断すると(ステップS36で「YES」)、写真である蓋然性が高いとして、現在の写真フラグBの値に「1」をインクリメントし(ステップS37)、その逆に幅 $W5 \geq TH5$ と判断すると(ステップS36で「NO」)、判定通り文字である蓋然性が高いとして、現在の文字フラグAの値に「1」をインクリメントして(ステップS38)、ステップS39に移る。

10

【0047】

ステップS39では、ヒストグラムの勾配数が閾値TH6よりも大きいかなんかを判断する。このヒストグラムの勾配数とは、図8(a)に示す「勾配」欄の「1」と「-1」の数を計数した値のことである。図8(a)の例では、「6」になる。

勾配数が多いということは、階調変化が大きい画像であることを意味するから、勾配数の大きさも写真であるか否かを再判定する一条件とすることができる。

【0048】

このTH6の値は、写真である蓋然性が高くなると想定される勾配数として予め実験等から求められたものであり、そのデータはROM106に格納されており、当該処理時に読み出される。

20

勾配数 $> TH6$ であることを判断すると(ステップS39で「YES」)、写真である蓋然性が高いとして、現在の写真フラグBの値に「1」をインクリメントし(ステップS40)、その逆に勾配数 $\leq TH6$ と判断すると(ステップS39で「NO」)、判定通り文字である蓋然性が高いとして、現在の文字フラグの値に「1」をインクリメントして(ステップS41)、ステップS42に移る。

【0049】

図11(a)は、写真領域に実際に存在する文字画像について文字拡張領域を設定した場合の明度ヒストグラムの例を示す図であり、図11(b)は、写真領域に文字が存在すると誤判定された画像(実際には写真の一部)について文字拡張領域を設定した場合の明度ヒストグラムの例を示す図である。

30

図11(a)を見ると、ピークが3つあり、最大と2番目のピークの間隔が広く、かつ勾配数も少なく(中間部分がほぼ平坦になっているため)、ヒストグラムの幅がどの閾値をとっても大変狭くなっていることが解る。したがって、上記した各判定条件(S32、S36、S39、S51、S54、S57、S60)に当てはめるとそのほとんどにおいて文字の蓋然性が高いと判断され、その結果、文字フラグの数値が増えることになる。

【0050】

一方で、図11(b)を見ると、ピークが1つであり、勾配数が多く(ピークが少ない分、勾配数が多くなっている)、ヒストグラムの幅条件における閾値を小さくした場合(0.3%、1.6%、2.1%等)にヒストグラムの幅が大変広がっていることが解る。したがって、各判定条件に当てはめるとそのほとんどにおいて写真の蓋然性が高いと判断され、写真フラグの数値が増えることになる。

40

【0051】

図9に戻って、ステップS42では、フラグAとBの値の大小関係を判断する。

そして、 $B > A$ の場合には(ステップS42で「YES」)、各判定条件の内、写真と判断された数が多いということになるから、実際には写真画像の蓋然性が極めて高く、文字の判定は誤判定であると判断し(ステップS43)、メインルーチンにリターンする。

一方、 $B \leq A$ の場合には(ステップS42で「NO」)、逆に文字と判断された数が多いことになるから、文字の蓋然性が依然高い、すなわち誤判定ではないと判断し(ステップS44)、領域再判定処理のサブルーチンにリターンする。

【0052】

50

図6に戻って、ステップS24では、ステップS43またはS44における判断結果を参照する。

誤判定であると判断すると(ステップS24で「YES」)、ステップS17で読み出した当該文字矩形領域は、1度、文字として判定されたが、実際には写真画像の一部である蓋然性が極めて高いと考えられるから、当該文字矩形領域を写真領域に変更して(ステップS25)、メインルーチンにリターンする。写真領域への変更とは、具体的には、領域情報格納部109に格納されている当該文字矩形領域に対する文字領域情報(ステップS15において保存されたもの。)を領域情報格納部109から削除するものである。

【0053】

一方、誤判定でないと判断すると(ステップS24で「NO」)、文字領域である蓋然性が極めて高いと考えられるが、本実施の形態ではもう1回だけ再判定(2回目の再判定)を行う(ステップS26)。この2回目の再判定の方法(第3の判定方法)としては、第2の判定方法よりもさらに判定精度が高い方法であれば良く、例えばステップS19の領域再判定処理における判定条件に、さらに別の判定条件を加えて再判定する方法や、ステップS14の文字判定と同じ方法ではあるが、その判定基準を高くする、すなわち抽出すべき画像の特徴量の数を増やし、特徴点の一致数がステップS14の処理における閾値よりも大きな値を越える場合に文字として判定する方法も考えられる。さらに、別の方法、例えば公知の文字認識の手法(文字として認識するための辞書データを予め格納しておき、被判定画像の形状がその辞書に格納されているパターンの形状と完全または所定部位が一致した場合にのみ文字と認識する方法)を用いて、文字判定を行うこともできる。

【0054】

文字と判定されなければ(ステップS27で「NO」)、ステップS25に移り、文字と判定されると(ステップS27で「YES」)、そのままメインルーチンにリターンする。この場合、当該文字矩形領域は、最初の判定通り、文字領域であると最終決定されることになる。

図3に戻って、ステップS20では、領域再判定処理をまだ行っていない文字矩形領域があるか否かを判断する。

【0055】

領域再判定処理を実行すべき文字矩形領域がある場合には(ステップS20で「YES」)、ステップS17に戻って、ステップS17～S20の処理を行う。領域再判定処理を実行すべき文字矩形領域がないと判断されるまで、ステップS17～S20の処理を繰り返し実行し、領域再判定処理を実行すべき文字矩形領域がないと判断すると(ステップS20で「NO」)、当該領域判定処理を終了する。

【0056】

図12は、画像圧縮処理の内容を示すフローチャートである。

同図に示すように、制御部102は、文字領域、写真領域、その他の領域について、各領域の文字画像、写真画像、その他の領域の画像の各画素データを読み出す(ステップS71)。

具体的には、文字領域については、領域情報格納部109から、文字領域情報として現に格納されている文字マップ(文字画像のみを構成する画素の座標)を読み出し、読み出した座標位置に位置する各画素のデータ(文字画像データ)を、画像データ格納部108に格納されている画像データから読み出すことにより行う。

【0057】

写真領域については、写真領域情報として領域情報格納部109に格納されている写真マップ(文字画像を構成する画素の座標)を読み出して、読み出した座標位置に位置する各画素のデータ(写真画像データ)を、画像データ格納部108に格納されている画像データから読み出す。また、その他の領域については、その他の領域情報として領域情報格納部109に格納されている画素の座標を読み出して、読み出した座標位置に位置する各画素のデータ(その他の領域の画像データ)を当該画像データから読み出すものである。

【0058】

10

20

30

40

50

そして、文字画像データをMMR圧縮し、写真画像データおよびその他の領域の画像データをJPEG圧縮して(ステップS72)、これら圧縮データを別ファイルとして圧縮データ格納部110に格納し(ステップS73)、当該画像圧縮処理を終了する。格納された圧縮データは、上記の送信処理においてサーバ3に送信されることになる。

以上説明したように、本実施の形態では、文字領域と写真領域の領域判別を行った後、写真領域に文字領域が重なっている部分について、その文字領域と判定された領域が本当に文字であるのかを、領域判別の際に用いた第1の方法とは異なる第2の方法を用いて再判定するようにしている。

【0059】

通常、文字と写真領域の判別では、文字と写真が重なっていない部分については文字領域と写真領域を判別し易い(すなわち、判定精度があまり高くない方法を用いても誤判定され難い)が、文字と写真の重複部分については誤判定され易い傾向がある。

これより、例えば第2の方法を第1の方法よりも判定精度が高い(その分CPUへの負担が大きくなる)方法とすれば、まず第1の方法を用いて文字領域を判定し、次に誤判定され易い文字と写真の重複部分についてのみ判定精度が高い第2の方法を用いて判定することができる。したがって、例えばスキャナ2で大量の原稿を1頁ごとに順次読み取ってその画像データを圧縮する処理を実行する場合であっても、原稿1頁分の画像の全画素について第2の方法を用いる構成に比べて、CPUへの処理負担の軽減および処理時間の低減が可能になり、かつ第1の方法だけを用いて判定する構成に比べて判定精度を向上させることができるという効果を奏する。

【0060】

なお、上記第1～第3の判定方法は、上記したものに限られることはいうまでもなく、判定精度が第1、第2、第3の順に高くなるものであれば良い。また、第1、第2の方法としては、判定精度がほぼ同じものでも異なる方法であれば一定の効果を得ることも出来る。

また、文字拡張領域を生成する際の拡張部の大きさも上記10画素分に限られることはなく、明度ヒストグラムに基づいて文字、写真の判定を行う場合の判定精度がより高くなると想定される値が実験等により予め決められることになる。また、拡張部の大きさを固定とする必要もなく、例えば文字領域の大きさに応じて可変させる構成にすること等が考えられる。

【0061】

また、本実施の形態では、文字画像を取り囲む文字矩形領域を文字領域として設定し、設定された文字矩形領域と写真領域との重複部分を検出するとしたが、矩形領域を設定せずに文字画像の部分だけを文字領域として設定し、設定された文字領域と写真領域との重複部分(写真領域内に入っている部分)を検出する方法をとることも可能である。

さらに、文字矩形領域全体が写真領域内に入っている場合だけでなく、その一部が写真領域と重なっている場合に、当該重なっている部分について領域再判定を行うとしても良い。

【0062】

また、原稿1頁分の画像の範囲を被判定領域としたが、例えば原稿を複数のブロックに分けてブロック単位で被判定領域を設定し、ブロックごとに領域判定処理を実行するとしても良い。

なお、本発明は、画像処理装置に限られず、上記した文字、写真領域を判定する画像領域判定方法にも適用することができる。また、その方法をコンピュータが実行するプログラムであるとしてもよい。

【0063】

また、本発明に係るプログラムは、例えば磁気テープ、フレキシブルディスク等の磁気ディスク、DVD、CD-ROM、CD-R、MO、PDなどの光記録媒体、Smart Media(登録商標)などのフラッシュメモリ系記録媒体等、コンピュータ読み取り可能な各種記録媒体に記録することが可能であり、当該記録媒体の形態で生産、譲渡等がなされる場

10

20

30

40

50

合もあるし、プログラムの形態でインターネットを含む有線、無線の各種ネットワーク、放送、電気通信回線、衛星通信等を介して伝送、供給される場合もある。

【0064】

(変形例)

以上、本発明を実施の形態に基づいて説明してきたが、本発明は、上述の実施の形態に限定されないのは勿論であり、以下のような変形例が考えられる。

(1) 上記実施の形態では、文字拡張領域を設定しようとする際に、拡張領域の一部が写真領域をはみ出る場合には、そのはみ出る方向への拡張をしないようにしたが、そのようにすると拡張すべき領域の面積が本来の拡張面積よりも減少することになる。そこで、例えば図13に示すようにはみ出る部分の拡張領域(破線部分)を写真領域内に収まる方向にシフトさせて(はみ出る部分を網点部分で示す範囲に振り分けて)、拡張領域を設定することで、いずれの場合でも拡張領域の面積をほぼ一定にして同一条件で再判定を行う構成をとることもできる。

10

【0065】

(2) 上記実施の形態では、第2の方法において文字拡張領域を設定するとしたが、例えば文字矩形領域だけから明度ヒストグラムを作成した場合に、第1の方法よりも判定精度を高くできる場合には必ずしも拡張することを要せず、その分処理負担を軽減することも可能である。

(3) 上記実施の形態では、画像処理装置としてパーソナルコンピュータを用いる場合の例を説明したが、本発明は、上記した領域判定処理を実行できるものであれば、例えば複写機、スキャナ、プリンタ、FAX、MFP(Multiple Function Peripheral)等の画像処理装置に適用できる。

20

【産業上の利用可能性】

【0066】

本発明に係る画像領域判定方法は、文字領域と写真領域をCPU等にかかる処理負担を軽減しながらより早く精度良く判定する方法に利用できる。

【図面の簡単な説明】

【0067】

【図1】本発明に係る画像処理装置を含む画像処理システムの全体構成を示す図である。

【図2】画像処理装置1の機能構成を示すブロック図である。

30

【図3】領域判定処理の内容を示すフローチャートである。

【図4】被判定領域における入力画像の例を示す図である。

【図5】文字矩形領域の設定例を示す図である。

【図6】領域再判定処理のサブルーチンの内容を示すフローチャートである。

【図7】(a)(b)は、図5に示す文字矩形領域Qの大きさを所定の範囲だけ拡張した場合の例を示す図である。

【図8】(a)は、ある文字拡張領域の16階調明度データに基づく明度の度数分布表の例を示す図であり、(b)は、当該明度の度数分布表を基に作成したヒストグラムの例を示す図である。

【図9】ヒストグラム解析処理のサブルーチンの内容を示すフローチャートである。

40

【図10】ヒストグラムの幅条件解析処理のサブルーチンの内容を示すフローチャートである。

【図11】(a)は、写真領域に実際に存在する文字画像について文字拡張領域を設定した場合の明度ヒストグラムの例を示す図であり、(b)は、写真領域に文字が存在すると誤判定された画像(実際には写真の一部)について文字拡張領域を設定した場合の明度ヒストグラムの例を示す図である。

【図12】画像圧縮処理の内容を示すフローチャートである。

【図13】文字矩形領域の大きさを拡張した場合の別の例を示す図である。

【符号の説明】

【0068】

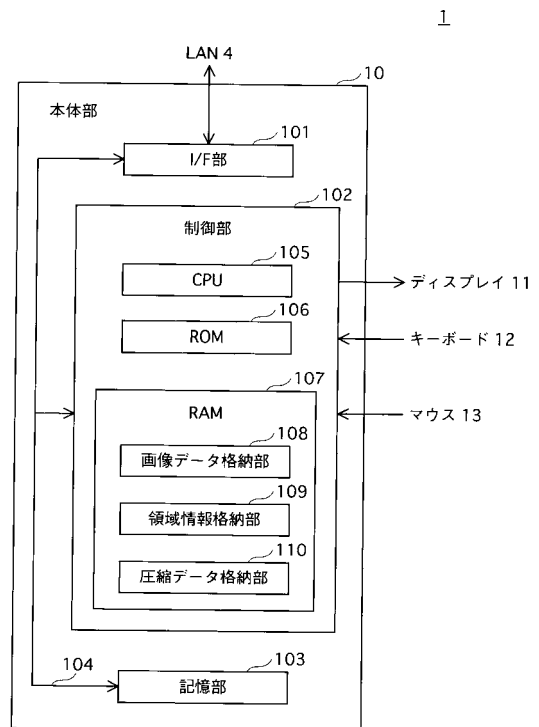
50

- 1 画像処理装置
- 102 制御部
- 105 CPU
- 106 ROM
- 107 RAM
- 109 領域情報格納部

【図1】

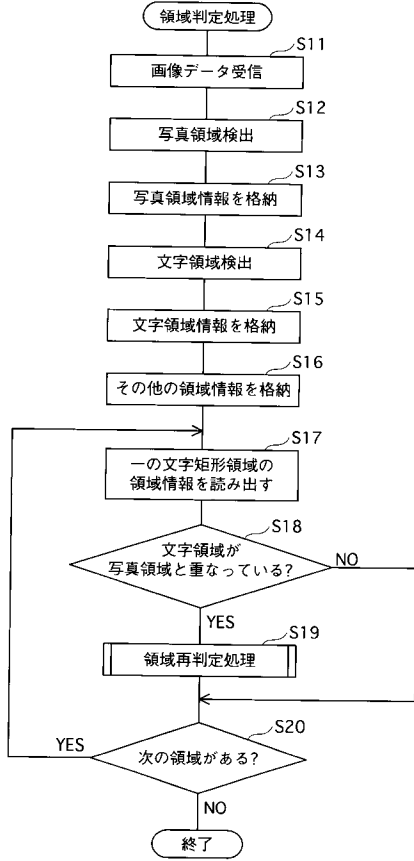


【図2】

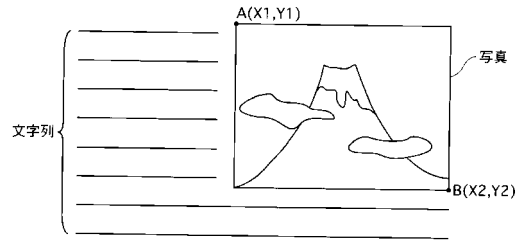


1

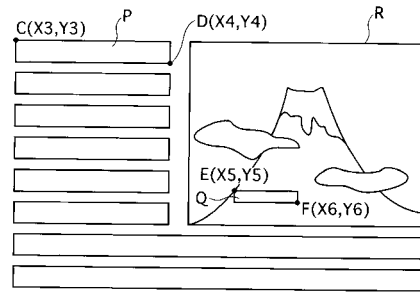
【 図 3 】



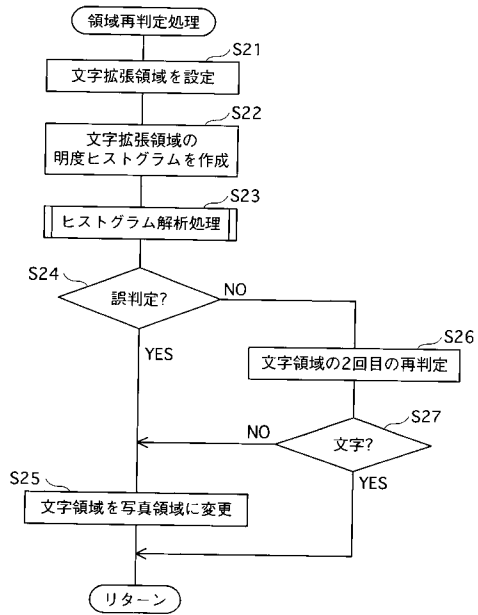
【 図 4 】



【 図 5 】

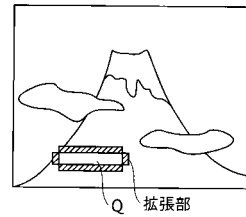


【 図 6 】

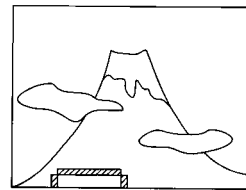


【 図 7 】

(a)



(b)

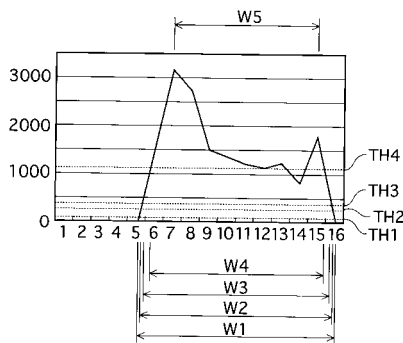


【 図 8 】

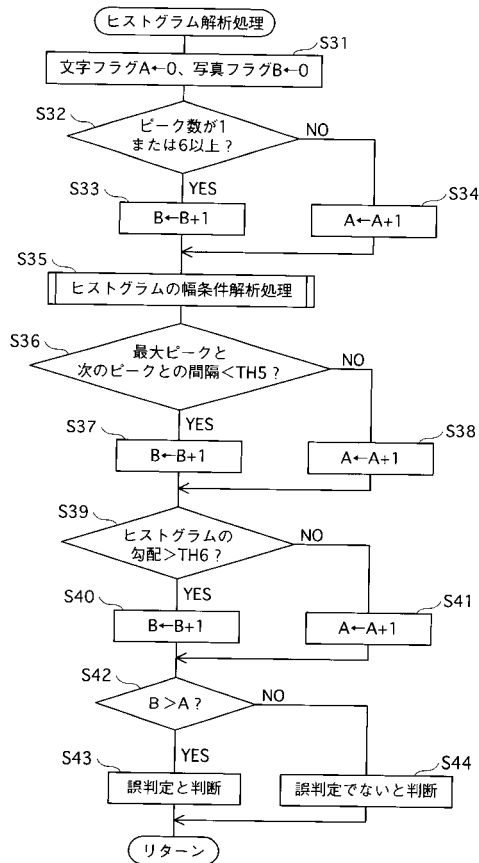
(a)

階調数	度数	勾配
1	0	0
2	0	0
3	0	0
4	0	0
5	0	1
6	1659	1
7	3163	2
8	2730	-1
9	1504	-1
10	1334	-1
11	1191	-1
12	1111	-2
13	1200	2
14	808	-2
15	1766	2
16	0	-

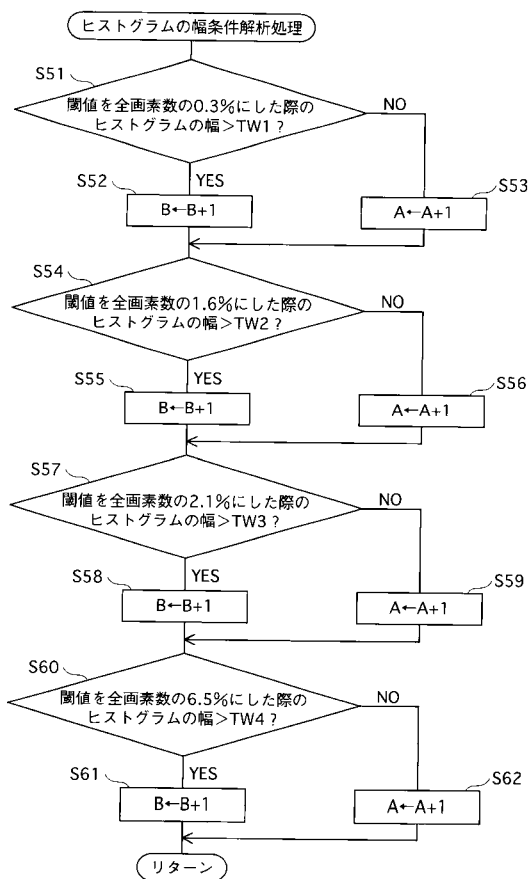
(b)



【 図 9 】

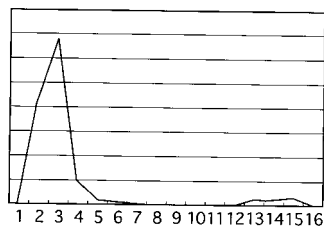


【 図 10 】



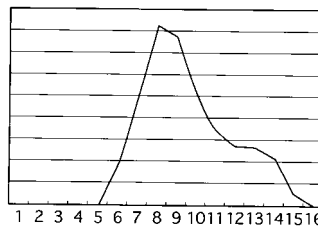
【 図 11 】

(a)



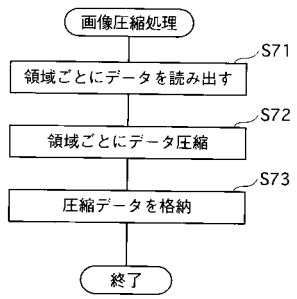
写真領域に文字が存在する場合の例

(b)



写真領域に文字が存在しない場合の例

【 図 1 2 】



【 図 1 3 】

