

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7292984号
(P7292984)

(45)発行日 令和5年6月19日(2023.6.19)

(24)登録日 令和5年6月9日(2023.6.9)

(51)国際特許分類 F I
G 0 6 V 30/148 (2022.01) G 0 6 V 30/148
G 0 6 V 30/14 (2022.01) G 0 6 V 30/14 3 4 0 K

請求項の数 20 (全22頁)

(21)出願番号	特願2019-110205(P2019-110205)	(73)特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	令和1年6月13日(2019.6.13)	(74)代理人	110001243 弁理士法人谷・阿部特許事務所
(65)公開番号	特開2020-201864(P2020-201864 A)	(72)発明者	山中 理 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(43)公開日	令和2年12月17日(2020.12.17)	審査官	淀川 滉也
審査請求日	令和4年6月10日(2022.6.10)		

最終頁に続く

(54)【発明の名称】 行分割装置および方法、ならびにプログラム

(57)【特許請求の範囲】

【請求項1】

原稿を光学的に読み取ったスキャン画像を二値化して得た二値画像内における行単位の大きさや位置を示す行情報と、前記二値画像内の行における文字単位の大きさや位置を示す文字情報とを取得する取得手段と、

前記文字情報から文字の高さを示す文字高さ情報を抽出する抽出手段と、

前記文字高さ情報に基づき、文字が高さ方向で他の文字と接触するかどうかを判定する第1の判定手段と、

前記行情報と前記第1の判定手段による判定結果に基づいて、前記原稿における行のうち注目する行が高さ方向で他の行と接触するかどうかを判定する第2の判定手段と、

前記第2の判定手段により他の行と接触すると判定された注目する行において、当該注目する行における幅方向にて、黒画素の数をカウントしたカウント値を導出する導出手段と、

前記第2の判定手段により他の行と接触すると判定された前記注目する行に対し、前記行情報で示される行の位置と前記文字高さ情報とから導出される基準点の周辺にて、前記カウント値において前記注目する行の高さ方向で最少となる第1の位置を行分割位置として行分割を実行する行分割手段と、

を有する、ことを特徴とする行分割装置。

【請求項2】

前記行分割手段は、前記カウント値の最少値が前記カウント値の合計値から導出した平

均値より小さい、かつ、前記第 1 の位置で分割されときの前記第 1 の位置の下側または上側の行の高さが、前記文字高さ情報から導出された文字の高さより大きいときに、前記第 1 の位置を前記行分割位置として行分割を実行することを特徴とする請求項 1 に記載の行分割装置。

【請求項 3】

前記行分割手段は、前記カウント値の最少値が前記カウント値の合計値から導出した平均値以上であるときに、前記基準点および前記文字高さ情報から導出された第 2 の基準点の周りにて、前記カウント値が最少となる第 2 の位置を前記行分割位置として行分割を実行することを特徴とする請求項 2 に記載の行分割装置。

10

【請求項 4】

前記行分割手段は、前記カウント値の最少値が前記カウント値の合計値から導出した平均値以上であり、かつ、前記基準点および前記文字高さ情報から導出された第 2 の基準点の周りにて、前記カウント値が最少となる第 2 の位置のカウント値が前記平均値より小さいときに、前記第 2 の位置を前記行分割位置として行分割を実行することを特徴とする請求項 2 に記載の行分割装置。

【請求項 5】

前記行分割手段は、前記カウント値の最少値が前記カウント値の計値から導出した平均値以上であり、前記基準点および前記文字高さ情報から導出された第 2 の基準点の周りにて、前記カウント値が最少となる第 2 の位置のカウント値が前記平均値より小さい、かつ、前記第 2 の位置で分割されときの当該第 2 の位置の下側または上側の行の高さが前記文字高さ情報から導出された文字の高さより大きいときに、前記第 2 の位置を前記行分割位置として行分割を実行することを特徴とする請求項 2 に記載の行分割装置。

20

【請求項 6】

前記抽出手段は、前記文字高さ情報から文字高さの最頻値を抽出し、
前記第 1 の判定手段は、前記文字情報で示される各文字の高さが前記最頻値と同じであるときに、前記文字が通常文字であると判定し、
前記第 2 の判定手段は、前記第 1 の判定手段により前記通常文字と判定された文字が前記注目する行に存在しないときに、前記注目する行が行接触行であると判定し、
前記行分割手段は、前記注目する行に対し、前記第 1 の位置を前記行分割位置として行分割を実行することを特徴とする請求項 2 から 5 の何れか一項に記載の行分割装置。

30

【請求項 7】

前記第 2 の判定手段は、前記第 1 の判定手段により前記通常文字と判定された文字が前記注目する行に存在しない、かつ、前記注目する行の高さが前記最頻値と異なるときに、前記注目する行が行接触行であると判定することを特徴とする請求項 6 に記載の行分割装置。

【請求項 8】

前記第 1 の判定手段は、前記文字情報で示される各文字の高さが前記最頻値と異なるときに、前記文字が行接触文字であると判定し、
前記第 2 の判定手段は、前記第 1 の判定手段により前記通常文字と判定された文字が前記注目する行に存在しない、前記注目する行の高さが前記最頻値と異なる、かつ、前記第 1 の判定手段により前記行接触文字と判定された文字が前記注目する行に存在するときに、前記注目する行が行接触行であると判定することを特徴とする請求項 7 に記載の行分割装置。

40

【請求項 9】

前記第 1 の判定手段は、前記文字情報で示される各文字の高さが前記最頻値と異なる、かつ、前記文字情報で示される各文字の信頼度が所定の閾値未満であるときに、前記文字が行接触文字であると判定する

50

ことを特徴とする請求項 8 に記載の行分割装置。

【請求項 10】

前記第 1 の判定手段は、前記文字情報で示される各文字の高さが前記最頻値と異なる、前記文字情報で示される各文字の信頼度が所定の閾値未満である、かつ、前記文字情報で示される各文字の高さが前記最頻値の 2 以上の整数倍と同じであるときに、前記文字が行接触文字であると判定する

ことを特徴とする請求項 8 または 9 に記載の行分割装置。

【請求項 11】

前記第 1 の判定手段は、前記文字情報で示される各文字の高さが前記最頻値と異なる、前記文字情報で示される各文字の信頼度が所定の閾値未満である、前記文字情報で示される各文字の高さが前記最頻値の 2 以上の整数倍と同じである、かつ、前記文字のアスペクト比が 2 であるときに、前記文字が縦倍角文字であると判定し、

前記行分割手段は、前記第 1 の判定手段により前記縦倍角文字と判定された文字が前記注目する行に存在するときに、前記縦倍角文字の最下位置を前記行分割位置として行分割を実行する

ことを特徴とする請求項 6 から 10 の何れか一項に記載の行分割装置。

【請求項 12】

前記行分割手段は、前記行情報で示される行上端位置から前記文字高さ情報に基づき導出された前記最頻値だけ下へ進めた位置を前記基準点としたときの前記行分割位置にて行分割を実行する

ことを特徴とする請求項 6 から 11 の何れか一項に記載の行分割装置。

【請求項 13】

前記行分割手段は、前記行情報で示される行下端位置から前記文字高さ情報に基づき導出された前記最頻値だけ上へ進めた位置を前記基準点としたときの前記行分割位置にて行分割を実行する

ことを特徴とする請求項 6 から 11 の何れか一項に記載の行分割装置。

【請求項 14】

前記行分割手段は、

前記行情報で示される行上端位置から前記文字高さ情報に基づき導出された前記最頻値だけ下へ進めた位置を前記基準点としたときの前記行分割位置にて行分割を実行する第一の行分割手段と、

前記行情報で示される行下端位置から前記文字高さ情報に基づき導出された前記最頻値だけ上へ進めた位置を前記基準点としたときの前記行分割位置にて行分割を実行する第二の行分割手段と、

を有し、

前記第一の行分割手段による前記行分割位置と、前記第二の行分割手段による前記行分割位置とに基づき、前記第一の行分割手段による行分割結果を採用するか、または前記第二の行分割手段による行分割結果を採用するかを決定する決定手段をさらに有する

ことを特徴とする請求項 6 から 11 の何れか一項に記載の行分割装置。

【請求項 15】

前記決定手段は、前記第一の行分割手段による前記行分割位置と、前記第二の行分割手段による前記行分割位置が同じであるときに、前記第一の行分割手段による行分割結果、または前記第二の行分割手段による行分割結果のどちらかを採用すると決定する

ことを特徴とする請求項 14 に記載の行分割装置。

【請求項 16】

前記決定手段は、前記第一の行分割手段による前記行分割位置と、前記第二の行分割手段による前記行分割位置が異なるときに、前記第一の行分割手段および前記第二の行分割手段により得られた行分割結果に対して OCR 処理を行い、前記 OCR 処理により認識した文字の信頼度に基づき採用する行分割結果を決定する

ことを特徴とする請求項 14 または 15 に記載の行分割装置。

10

20

30

40

50

【請求項 17】

前記二値画像から罫線を除去する第一の除去手段を有し、

前記取得手段は、前記第一の除去手段により前記罫線を除去した前記二値画像から前記行情報および前記文字情報を取得する

ことを特徴とする請求項 1 から 16 の何れか一項に記載の行分割装置。

【請求項 18】

前記二値画像からノイズを除去する第二の除去手段を有し、

前記取得手段は、前記第二の除去手段により前記ノイズを除去した前記二値画像から前記行情報および前記文字情報を取得する

ことを特徴とする請求項 17 に記載の行分割装置。

10

【請求項 19】

原稿を光学的に読み取ったスキャン画像を二値化して得た二値画像内における行単位の大きさと位置を示す行情報と、前記二値画像内の行における文字単位の大きさと位置を示す文字情報を取得する取得ステップと、

前記文字情報から文字の高さを示す文字高さ情報を抽出する抽出ステップと、

前記文字高さ情報に基づき、文字の高さ方向で他の文字と接触するかどうかを判定する第 1 の判定ステップと、

前記行情報と前記第 1 の判定ステップによる判定結果に基づいて、前記原稿における行のうち注目する行が高さ方向で他の行と接触するかどうかを判定する第 2 の判定ステップと、

20

前記第 2 の判定ステップにて他の行と接触すると判定された前記注目する行において、当該注目する行における幅方向にて、黒画素の数をカウントしたカウント値を導出する導出ステップと、

前記第 2 の判定ステップにて他の行と接触すると判定された前記注目する行に対し、前記行情報で示される行の位置と前記文字高さ情報とから導出される基準点の周辺にて、前記カウント値において前記注目する行の高さ方向で最少となる第 1 の位置を行分割位置として行分割を実行する行分割ステップと、

を有する、ことを特徴とする行分割方法。

【請求項 20】

コンピュータに、請求項 1 から 18 の何れか一項に記載の行分割装置の各手段として機能させるためのプログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、スキャン画像の文字領域から行画像を分割する技術に関する。

【背景技術】

【0002】

従来、カメラやスキャナなどにより光学的に原稿を読み取って得たスキャン画像内の文字を認識する文字認識技術（OCR 技術）がある。OCR 技術では、スキャン画像の文字領域を行単位に分割したり（以下、行分割と称す）、文字単位に分割したりする（以下、文字分割と称す）ことで、各文字の位置や大きさを正確に把握する必要がある。原稿の読み取り時にゴミなどによってスキャン画像内の文字が行間でつながると、行分割を適切に行うことができず、OCR 精度を低下させてしまう。

40

【0003】

特許文献 1 には、画像を二値化して行の幅方向にて黒画素の数をカウントしてカウント値を取得し、取得した黒画素のカウント値から導出した平均文字高さを基準として、文字が行間でつながる接触行であるかどうかを判定する技術が記載されている。

【先行技術文献】

【特許文献】

【0004】

50

【文献】特開平 6 - 2 2 3 2 2 4 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、行間が極めて狭く複数箇所で行同士が接触する行接触を生じて、文字が行間でつながる接触行がある帳票では、特許文献 1 の技術を用いてもスキャン画像にて行分割を適切に行うことができない場合があった。例えば、印刷に使用するロール紙を節約するレシート印刷機では、文字領域の行間が極めて狭く、複数箇所で行接触を生じているレシートが生成される場合がある。

【0006】

特許文献 1 の技術では、黒画素のカウント値から導出した平均文字高さを基準として接触行であるかどうかを判定しており、接触行が複数存在すると、接触行を含めた平均の文字高さを実際の行高さとのずれが大きくなる。これにより、接触行であるかどうかを正確に判定できず、行分割を適切に行うことができない可能性があった。また、行数が少ない原稿のスキャン画像では、数箇所でも接触行があると平均文字高さが実際の文字高さから大きくずれてしまい、同様に、接触行であるかどうかを正確に判定できず、行分割を適切に行うことができない可能性があった。

【0007】

本発明は、原稿を光学的に読み取ったスキャン画像に対して行分割を適切に行うことを目的とする。

【課題を解決するための手段】

【0008】

本発明の一態様に係る行分割装置は、原稿を光学的に読み取ったスキャン画像を二値化して得た二値画像内における行単位の大きさと位置を示す行情報と、前記二値画像内の行における文字単位の大きさと位置を示す文字情報とを取得する取得手段と、前記文字情報から前記文字の高さを示す文字高さ情報を抽出する抽出手段と、前記文字高さ情報に基づき、前記文字が高さ方向で他の文字と接触するかどうかを判定する第 1 の判定手段と、前記行情報と第 1 の判定手段による判定結果に基づいて、前記原稿における行のうち注目する行が高さ方向で他の行と接触するかどうかを判定する第 2 の判定手段と、前記第 2 の判定手段により他の行と接触すると判定された前記注目する行において、当該注目する行における幅方向にて、黒画素の数をカウントしたカウント値を導出する導出手段と、前記第 2 の判定手段により他の行と接触すると判定された前記注目する行に対し、前記行情報で示される行の位置と前記文字高さ情報とから導出される基準点の周辺にて、前記カウント値において前記注目する行の高さ方向で最少となる第 1 の位置を行分割位置として行分割を実行する行分割手段と、を有する、ことを特徴とする。

【発明の効果】

【0009】

本発明によれば、原稿を光学的に読み取ったスキャン画像に対して行分割を適切に行うことができる。

【図面の簡単な説明】

【0010】

【図 1】情報処理システムの概略を示す図である。

【図 2】行分割処理の手順例を示すフローチャートである。

【図 3】スキャン画像例を示す図である。

【図 4】文字属性判定処理の手順例を示すフローチャートである。

【図 5】行属性判定処理の手順例を示すフローチャートである。

【図 6】第二の行分割処理の手順例を示すフローチャートである。

【図 7】スキャン画像の二値画像例を示す図である。

【図 8】行接触行を含む領域の二値画像例およびその射影画像例を示す図である。

【図 9】文字属性判定処理の手順例を示すフローチャートである。

10

20

30

40

50

【図10】第二の行分割処理の手順例を示すフローチャートである。

【図11】第二の行分割処理の手順例を示すフローチャートである。

【図12】行分割処理の手順例を示すフローチャートである。

【図13】行分割結果決定処理の手順例を示すフローチャートである。

【図14】情報処理装置における機能構成例を示すブロック図である。

【発明を実施するための形態】

【0011】

以下、本発明を実施するための形態について図面を用いて説明する。ただし、この実施形態に記載されている構成要素はあくまで例示であり、本発明の範囲をそれらに限定する趣旨のものではない。また、実施形態で説明されている構成要素の組み合わせのすべてが、課題を解決するための手段に必須のものとは限らず、種々の変形及び変更が可能である。

10

【0012】

(実施形態1)

図1は、本実施形態に係る情報処理システムの概略構成例を示す図である。情報処理システムは、複写機100と、情報処理装置110とを有する。複写機100は、スキャナ101と、プリンタ102と、複写機側通信部103とを有する。スキャナ101は、原稿を光学的に読み取るスキャンを行って得たスキャン画像を生成する。プリンタ102は、紙などの記録媒体に画像を形成する。複写機側通信部103は、ネットワークを介して外部装置と通信処理を行う。複写機側通信部103は、例えば、後述の情報処理装置110の通信部117との通信処理を行う。

20

【0013】

情報処理装置110は、CPU111と、ROM112と、RAM113と、HDD114と、表示部115と、入力部116と、情報処理装置側通信部117とを有する。CPU111は、ROM112に記憶された制御プログラムを読み出して各種処理を実行する。RAM113は、CPU111の主メモリ、ワークエリア等の一時記憶領域として用いられる。HDD114は、各種データや各種プログラム等を記憶する。なお、後述する情報処理装置110の機能や処理は、CPU111がROM112又はHDD114に格納されているプログラムを読み出し、このプログラムを実行することにより実現されるものである。

【0014】

情報処理装置側通信部117は、ネットワークを介して外部装置と通信処理を行う。情報処理装置側通信部117は、例えば、複写機100の通信部103との通信処理を行う。表示部115は、各種情報を表示する。入力部116は、キーボードやマウスなどの入力装置を有し、ユーザによる各種操作を受け付ける。なお、表示部115と入力部116は、タッチパネルのように一体に設けられていてもよい。また、表示部115は、プロジェクタによる投影を行うものであってもよく、入力部116は、投影された画像に対する指先の位置を、カメラで認識するものであってもよい。

30

【0015】

本実施形態においては、複写機100のスキャナ101がレシートなどの原稿をスキャンし、入力画像(スキャン画像301)を生成する。そして、入力画像は、複写機側通信部103により情報処理装置110に送信される。情報処理装置110においては、情報処理装置側通信部117が入力画像を受信し、これをHDD114などの記憶部に記憶する。

40

【0016】

図2は、本実施形態に係る行分割処理の手順例を示すフローチャートである。なお、レシートなどの処理対象の原稿は、スキャナ101でスキャン可能に複写機100の所定の位置に載置されているとする。

【0017】

S201において、複写機100のスキャナ101により原稿(文書)を光学的にスキャンしてスキャン画像を取得する。複写機100が取得したスキャン画像は、通信部10

50

3による情報処理装置110の通信部117との通信処理により、情報処理装置110に送られる。情報処理装置110のCPU111が複写機100から送られたスキャン画像をHDD114に格納する。図3は、原稿のスキャン画像例を示す図である。スキャン画像301は、レシートをスキャンした画像である。本実施形態では、レシートのスキャン画像を用いたが、レシート以外の、行間が狭く、複数箇所で行同士が接触する行接触が発生している原稿や、スキャン時のゴミやノイズによって文字が行間でつながって行接触した原稿のスキャン画像でもよい。原稿の罫線に対応する領域であるスキャン画像301の罫線領域302よりも上方側にあつては、文字列の行間が広く、各文字列の文字を認識可能な状態となっている。スキャン画像301の罫線領域302よりも下方側にあつては、文字列の行間が狭く、各文字列の文字を認識しづらい状態となっている。スキャン画像301の詳細は後述する。

10

【0018】

図2のフローの説明に戻る。S202において、CPU111は、スキャン画像301に対して二値化処理を行い、CPU111が二値化処理で取得した二値画像をRAM113に格納する。二値化処理とは、画像を白と黒の2階調に変換する処理のことであり、閾値より濃い色の画素は黒画素、その閾値より薄い色の画素は白画素となる。本実施形態では、スキャン画像301全体のヒストグラムから閾値を決定して二値画像を作成する手法を用いたが、その後の文字認識が可能な精度で白と黒の二値に変換されれば具体的な手法は問わない。

【0019】

20

S203において、CPU111は、S202の処理で取得した二値画像に対して罫線除去(罫線領域除去)を行う。罫線除去とは、スキャン画像301の罫線領域302に対応する二値画像内の罫線領域を検出し、検出した罫線領域を二値画像から除去する処理である。除去対象の罫線領域としては、罫線領域302のような破線だけでなく、例えば実線や縦方向の罫線などの罫線が存在する原稿のスキャン画像であれば、同様に、スキャン画像に対する二値処理により取得した二値画像に対して罫線除去を実行する。

【0020】

S204において、CPU111は、S203の処理を行った二値画像に対してモルフォロジー変換処理などのノイズ除去を行う。スキャン時のゴミの混入などによってスキャン画像301に微小な黒画素がノイズとして混入することがあり、そのノイズを除去する。本実施形態では、二値画像に対してモルフォロジー変換処理を行うことによりノイズ除去を行っているが、その後のOCRが可能な精度でノイズが除去されれば具体的な手法は問わない。

30

【0021】

S205において、CPU111は、S204の処理を行った二値画像に対して第一の行分割を行い、CPU111が分割した行の情報をRAM113に格納する。本実施形態では、二値画像全体を行の幅方向に射影をとり、行の高さ方向にて黒画素数が所定の閾値以上の領域を行と判定する。二値画像全体を行の幅方向に射影をとるとは、行の高さ方向のそれぞれの位置にて黒画素の数を行の幅方向でカウントして黒画素数を取得することである。例えば、行の幅方向にて黒画素の出現頻度を表す行方向の黒画素の射影ヒストグラムを作成し、作成した黒画素の射影ヒストグラムに基づき黒画素数を取得してもよい。CPU111は、行と判定した領域の位置、高さ、幅の情報を行情報としてRAM113に格納する。

40

【0022】

S206において、CPU111は、S205の処理で取得した行の情報を基に、S204の処理を行った二値画像に対して文字分割を行い、CPU111が文字分割で取得した、分割した文字の情報をRAM113に格納する。本実施形態では、分割した各行に対して行と垂直な方向に射影をとり、黒画素数が所定の閾値以上の領域を文字と判定する。CPU111は、文字と判定した領域の位置、高さ、幅の情報を文字情報として取得し、行情報と関連付けてRAM113に格納する。S205の行分割、S206の文字分割に

50

については、行接触した文字を除き、原稿の二値画像における文字領域を行単位、文字単位に分割できれば本実施形態以外の手法を用いてもよい。

【0023】

S207において、CPU111は、S206で取得した文字情報を基に、S204の処理を行った二値画像に対してOCR処理を行い、CPU111がOCR処理結果を文字情報と関連付けてRAM113に格納する。OCR処理結果には、各文字情報に複数の認識結果候補が格納されており、各認識結果候補には信頼度が関連付けられている。本実施形態において、信頼度とは、認識対象文字の特徴量と、OCR処理結果の文字の特徴量の一致度を数値化(0~255)したものである。信頼度の数値が低いほどOCR処理結果が認識対象文字とOCR処理結果の文字との一致率が高いことを表し、より信頼できる結果であることを示す。また、信頼度は、各文字のOCR結果の確からしさが客観的に比較できれば、その表現方法は問わない。

10

【0024】

S208において、CPU111は、S205で取得した各文字の高さ情報から文字高さ最頻値Hを導出し、導出結果である文字高さ最頻値Hを文字情報と関連付けてRAM113に格納する。本実施形態では、文字高さ情報を用いて文字高さの頻度グラフを作成し、最も頻度の高い中央値を文字高さ最頻値Hと定義する。

【0025】

S209において、CPU111は、文字属性判定処理を行い、S207の処理で取得した各文字(認識対象文字)に対して文字属性を付与し、文字情報と関連付けてRAM113に格納する。

20

【0026】

図4は、本実施形態に係る文字属性判定処理の手順例を示すフローチャートである。S401において、CPU111は、原稿の文字領域内の各文字の文字情報をRAM113から取得する。

【0027】

S402において、CPU111は、S401で取得した原稿の文字領域内の文字情報に対して文字属性未判定の文字を検索し、未判定の文字が存在するか存在しないかを判定する。CPU111は、未判定の文字が存在するとの判定結果を得た場合(S402のYES)、処理をS403へ移行する。CPU111は、未判定文字が存在しないとの判定結果を得た場合(S402のNO)、文字属性判定処理を終了する。

30

【0028】

S403において、CPU111は、S402の処理で未判定の文字と判定された当該未判定文字の文字高さおよび文字信頼度を含む文字情報を取得する。S403に続くS404、S406、S407において、未判定文字の文字属性を判定する。

【0029】

S404において、CPU111は、S403で取得した文字情報に基づき、処理対象の文字の文字高さがS208で取得した文字高さ最頻値Hと同程度であるか否かを判定する。CPU111は、処理対象の文字の文字高さが文字高さ最頻値Hと同程度であるとの判定結果を得た場合(S404のYES)、処理をS405へ移行する。CPU111は、処理対象の文字の文字高さが文字高さ最頻値Hと同程度ではないとの判定結果を得た場合(S404のNO)、処理をS406へ移行する。本実施形態において、文字高さ最頻値Hの±30%以内の文字高さの文字を、文字高さ最頻値Hと同程度とした。なお、文字高さ最頻値Hと同程度は文字高さ最頻値Hの±30%以内に限定されない。文字高さ最頻値Hと同程度は、文字高さ最頻値Hの±25%以内などのよりも狭い範囲や文字高さ最頻値Hの±35%以内などのよりも広い範囲としてもよい。

40

【0030】

S405において、CPU111は、処理対象の文字の文字属性を通常文字と判定する。CPU111は、文字属性の判定結果である通常文字をS403で取得した文字情報と関連付けてRAM113に格納する。CPU111は、文字属性の判定結果をRAM11

50

3へ格納した後、処理をS402へ移行する。

【0031】

S406において、CPU111は、S403で取得した文字情報に基づき、処理対象の文字の信頼度が閾値以上か否かを判定する。CPU111は、処理対象の文字の信頼度が閾値以上であるとの判定結果を得た場合(S406のYES)、処理をS405へ移行する。そして、上述の通り、文字属性を通常文字と判定し、通常文字をS403で取得した文字情報と関連付けてRAM113に格納する。CPU111は、処理対象の文字の信頼度が閾値未満であって、処理対象の文字の信頼度が閾値以上ではないとの判定結果を得た場合(S406のNO)、処理をS407へ移行する。本実施形態では、閾値を150とした。なお、文字の信頼度判定用の閾値は150に限定されない。文字の信頼度判定用の閾値は150より小さい数値や150より大きい数値としてもよい。

10

【0032】

S407において、CPU111は、取得した文字高さ(処理対象の文字の文字高さ)が文字高さ最頻値Hの2以上の整数倍と同程度であるか否かを判定する。本実施形態において、文字高さ最頻値Hの2以上の整数倍の±30%以内の文字高さの文字を文字高さ最頻値Hの2以上の整数倍と同程度とした。なお、文字高さ最頻値Hの2以上の整数倍と同程度はこの範囲に限定されない。文字高さ最頻値Hの2以上の整数倍と同程度は、文字高さ最頻値Hの2以上の整数倍のマイナス25%から文字高さ最頻値Hの2以上の整数倍のプラス25%範囲内などのより狭い範囲としてもよい。文字高さ最頻値Hの2以上の整数倍と同程度は、文字高さ最頻値Hの2以上の整数倍のマイナス35%から文字高さ最頻値Hの2以上の整数倍のプラス35%範囲内などのより広い範囲としてもよい。CPU111は、処理対象の文字の文字高さが文字高さ最頻値Hの2以上の整数倍と同程度であるとの判定結果を得た場合(S407のYES)、処理をS408へ移行する。CPU111は、処理対象の文字の文字高さが文字高さ最頻値Hの2以上の整数倍と同程度ではないとの判定結果を得た場合(S407のNO)、処理をS402へ移行する。

20

【0033】

S408において、CPU111は、処理対象の文字の文字属性を行接触文字と判定する。CPU111は、文字属性の判定結果である行接触文字をS403で取得した文字情報と関連付けてRAM113に格納する。CPU111は、文字属性の判定結果をRAM113へ格納した後、処理をS402へ移行する。

30

【0034】

上述したS402からS408の処理を行うことにより、文字属性判定処理を行っていない原稿の文字領域内の全ての文字に対して文字属性が判定されることになる。

【0035】

図2のフローの説明に戻る。S210において、CPU111は、行属性判定処理を行い、各行に対して行属性を付与し、行情報と関連付けてRAM113に格納する。

【0036】

図5は、本実施形態に係る行属性判定処理の手順例を示すフローチャートである。S501において、CPU111は、原稿の文字領域内の、各行の行情報およびその行に含まれる文字の文字属性情報をRAM113より取得する。

40

【0037】

S502において、CPU111は、S501で取得した文字の文字属性情報に対して行属性未判定の行を検索し、未判定の行が存在するか存在しないかを判定する。CPU111は、未判定の行が存在するとの判定結果を得た場合(S502のYES)、処理をS503へ移行する。CPU111は、未判定の行が存在しないとの判定結果を得た場合(S502のNO)、行属性判定処理を終了する。

【0038】

S503において、CPU111は、S502の処理で未判定行と判定された当該未判定行に含まれる文字の文字属性および行高さ情報を取得する。S503に続くS504、S506、S507において、未判定行の行属性を判定する。

50

【 0 0 3 9 】

S 5 0 4において、C P U 1 1 1は、行内の文字の文字属性が全て通常文字であるか否かを判定する。C P U 1 1 1は、行内の文字の文字属性が全て通常文字である（通常文字のみである）との判定結果を得た場合（S 5 0 4のY E S）、処理をS 5 0 5へ移行する。C P U 1 1 1は、行内の文字の文字属性が通常文字ではない文字を含んでおり、通常文字のみではないとの判定結果を得た場合（S 5 0 4のN O）、処理をS 5 0 6へ移行する。

【 0 0 4 0 】

S 5 0 5において、C P U 1 1 1は、行属性を通常行と判定する。C P U 1 1 1は、行属性の判定結果である通常行をS 5 0 3で取得した情報と共に行情報に関連付けてR A M 1 1 3に格納する。C P U 1 1 1は、行属性の判定結果をR A M 1 1 3へ格納した後、処理をS 5 0 2へ移行する。

10

【 0 0 4 1 】

S 5 0 6において、C P U 1 1 1は、取得した行高さがS 2 0 8で取得した文字高さ最頻値Hと同程度であるか否かを判定する。本実施形態において、文字高さ最頻値Hの±30%以内の行高さの行を文字高さ最頻値Hと同程度とした。C P U 1 1 1は、取得した行高さが文字高さ最頻値Hと同程度であるとの判定結果を得た場合（S 5 0 6のY E S）、処理をS 5 0 5へ移行する。C P U 1 1 1は、取得した行高さが文字高さ最頻値Hと同程度ではないとの判定結果を得た場合（S 5 0 6のN O）、処理をS 5 0 7へ移行する。

【 0 0 4 2 】

S 5 0 7において、C P U 1 1 1は、処理対象の文字の文字属性が行接触文字である文字が行内に1つ以上存在するか否かを判定する。C P U 1 1 1は、処理対象の文字の文字属性が行接触文字である文字が行内に1つ以上存在するとの判定結果を得た場合（S 5 0 7のY E S）、処理をS 5 0 8へ移行する。C P U 1 1 1は、処理対象の文字の文字属性が行接触文字である文字が行内に1つも存在しないとの判定結果を得た場合（S 5 0 7のN O）、処理をS 5 0 2へ移行する。

20

【 0 0 4 3 】

S 5 0 8において、C P U 1 1 1は、行属性を行接触行と判定する。C P U 1 1 1は、文字属性の判定結果である行接触文字をS 5 0 3で取得した情報と共に行情報と関連付けてR A M 1 1 3に格納する。C P U 1 1 1は、行属性の判定結果をR A M 1 1 3へ格納した後、処理をS 5 0 2へ移行する。

30

【 0 0 4 4 】

上述したS 5 0 2からS 5 0 8の処理を行うことにより、行属性判定処理を行っていない原稿の文字領域内の全ての行に対して行属性が判定されることになる。

【 0 0 4 5 】

図7は、図3に示すスキャン画像301に対しS 2 0 2の二値化処理からS 2 1 0の行属性判定処理を実施した後の二値画像700を示す図である。二値画像700内の破線矩形領域701は、行属性が通常行である行を示している。二値画像700内の実線矩形領域702は、行属性が行接触行である行を示している。実線矩形領域702において、1行目「小計」の「言（ごんべん）」と、2行目「（）」とが行接触している。同様に、実線矩形領域702において、1行目「0」と2行目「）」と、2行目「内」と3行目「計」の「言（ごんべん）」、2行目「）」と3行目「0」とが行接触している。すなわち、二値画像700の実線矩形領域702においては、3行にわたり4箇所で行接触が発生している。

40

【 0 0 4 6 】

図2のフローの説明に戻る。S 2 1 1において、C P U 1 1 1は、二値画像における行であって行属性が行接触行である行に対して第二の行分割を行い、C P U 1 1 1が分割した行の情報をR A M 1 1 3に格納する。

【 0 0 4 7 】

図6は、本実施形態に係る第二の行分割処理の手順例を示すフローチャートである。S 6 0 1において、C P U 1 1 1は、S 2 1 0の判定処理で得た各行の行属性情報をR A M

50

113より取得する。そして、CPU111は、取得した各行の行属性情報を基に、スキャン画像（二値画像）の文字領域内に、処理対象の行の行属性が行接触行である行であって、第二の行分割処理が行われていない行が存在するか否かを判定する。CPU111は、スキャン画像（二値画像）の文字領域内に、行属性が行接触行である行であって、第二の行分割処理が行われていない行が存在するとの判定結果を得た場合（S601のYES）、処理をS602へ移行する。CPU111は、スキャン画像（二値画像）の文字領域内に、行属性が行接触行である行であって、第二の行分割処理が行われていない行が存在しないとの判定結果を得た場合（S601のNO）、第二の行分割処理を終了する。

【0048】

S602において、CPU111は、行属性が行接触行である行であって、第二の行分割が行われていない行の位置、高さ、幅を示す行情報をRAM113より取得する。すなわち、CPU111は、注目する行の、行上端位置や行下端位置などの位置や高さなどを示す行情報を取得する。

10

【0049】

S603において、CPU111は、スキャン画像301の二値画像における、行接触行に対応する箇所に対して行方向に射影を作成し、処理対象の行接触行に対応する箇所を含む矩形領域の射影情報を取得する。本実施形態では、CPU111は、二値画像700の実線矩形領域702の射影情報を取得する。

【0050】

S604において、CPU111は、S603で取得した射影情報を基に、黒画素数の平均値Aを導出する。例えば、CPU111は、黒画素数（カウント値）の合計値を導出し、導出した黒画素数の合計値を処理対象の行接触行の高さ方向の画素数（カウント数）で除算することによって、黒画素数の合計値の平均値A（カウント値の平均値）を導出する。

20

【0051】

S605において、CPU111は、二値画像700の実線矩形領域702の射影画像において、行分割位置を特定するための基準点を決定する。

【0052】

図8は、二値画像700の実線矩形領域702に対応する二値画像800およびその射影画像801を示す図である。CPU111は、S602で取得した行情報に含まれる行接触行の行上端位置（先頭位置）803より、S208で取得した文字高さ最頻値Hだけ下へ進めた位置を基準点として決定する。

30

【0053】

S606において、CPU111は、S605の処理で決定した基準点の周辺のうち、黒画素数が最少値Sとなる位置P（第1の位置）を特定する。本実施形態において、基準点から、行と垂直な方向にて±2画素分の黒画素数を比較して黒画素数が最少値Sとなる位置Pを分割位置804として特定した。なお、基準点の周辺の大きさは、基準点から±2画素分に限定されない。基準点の周辺の大きさは、基準点から±1画素としたり、基準点から±2画素より大きくしたりしてもよい。

【0054】

S607において、CPU111は、黒画素数最少値S（図8の分割位置804）が黒画素数の平均値A（図8の位置802）より小さいか否かを判定する。CPU111は、黒画素数最少値Sが黒画素数の平均値Aより小さい（黒画素数の平均値A未満である）との判定結果を得た場合（S607のYES）、処理をS608へ移行する。CPU111は、黒画素数最少値Sが黒画素数の平均値A以上（カウント値の平均値以上）であるとの判定結果を得た場合（S607のNO）、処理をS601へ移行する。

40

【0055】

S608において、CPU111は、分割位置804で行分割を行った場合の、行分割後の残りの領域の行高さ（以下、行分割後の行高さともいう）を導出する。すなわち、CPU111は、分割位置804で行分割を行った場合の、二値画像800にて分割位置8

50

04の下側の領域の高さであって、分割位置804から行接触行の最下位置806までの距離を導出する。そして、CPU111は、導出した行分割後の行高さがS208で導出した文字高さ最頻値Hよりも大きいかなかを判定する。CPU111は、導出した行分割後の行高さが文字高さ最頻値Hよりも大きいとの判定結果を得た場合(S608のYES)、処理をS609へ移行する。CPU111は、導出した行分割後の行高さが文字高さ最頻値H以下であるとの判定結果を得た場合(S608のNO)、処理をS601へ移行する。

【0056】

S609において、CPU111は、処理対象である二値画像800に対してS606で特定した分割位置804にて行分割を実行する。この行接触行に対する行分割が第二の行分割処理に対応する。CPU111は、第二の行分割処理を実施した後、処理をS603へ移行する。

10

【0057】

二値画像800にあつては、CPU111は、分割位置804で行分割を実行する場合と同様、分割位置804の下側の領域に対して、S601からS609の第二の行分割処理を実行する。これにより、分割位置805を特定し、特定した分割位置805にて行分割が実行される。

【0058】

よって、二値画像800にあつては、行接触行は、3つの行に分割される。

【0059】

20

以上、説明したとおり、本実施形態によれば、行間が極めて狭い文書や複数箇所で行接触した原稿のスキャン画像や、行接触した行の割合が多い原稿のスキャン画像においても、より正確な行分割を実現することができる。その結果、より高精度なOCR処理を実現することができる。

【0060】**(実施形態2)**

本実施形態については、実施形態1との差分のみの説明とし、特に明記しない部分については実施形態1と同じ構成および手順となる。図2に示す行分割処理のS201からS208、S210の処理は、実施形態1の場合と同じであり、その説明を省略する。

【0061】

30

実施形態1では、通常文字および行接触文字の2種類の文字属性に対応する文字を含む原稿の行分割処理について説明した。3種類以上の文字属性に対応する文字を含む原稿の行分割処理としてもよい。

【0062】

例えば、レシートには、文字幅に対する文字高さの比(アスペクト比)が2である文字(以下、縦倍角文字と称す)がしばしば用いられる。本実施形態では、文字属性の通常文字よりも文字高さが大きい縦倍角文字を文字属性の判定に追加した文字属性判定処理について説明する。

【0063】

図9は、本実施形態に係る文字属性判定処理の手順例を示すフローチャートである。S901からS908の各処理は、図4のS401からS408の各処理と同じであり、その説明を省略する。すなわち、図9は、図2に示すフローにS909およびS910を追加したフローチャートとなっている。

40

【0064】

S907において、CPU111は、処理対象の文字の文字高さが文字高さ最頻値Hの2以上の整数倍と同程度でないとの判定結果を得た場合(S907のNO)、処理をS909へ移行する。

【0065】

S909において、CPU111は、通常文字または行接触文字と判定された文字以外の文字のアスペクト比が2であるかなかを判定する。CPU111は、処理対象の文字の

50

アスペクト比が2であるとの判定結果を得た場合（S909のYES）、処理をS910へ移行する。CPU111は、処理対象の文字のアスペクト比が2ではなく、2以外であるとの判定結果を得た場合（S909のNO）、処理をS902へ移行する。

【0066】

S910において、CPU111は、処理対象の文字の文字属性を縦倍角文字と判定する。CPU111は、文字属性の判定結果である縦倍角文字をS903で取得した文字情報と関連付けてRAM113に格納する。CPU111は、文字属性の判定結果をRAM113へ格納した後、処理をS902へ移行する。

【0067】

上述したS902からS910の処理を行うことにより、文字属性判定処理を行っていない原稿の文字領域内の全ての文字に対して文字属性が判定されることになる。

10

【0068】

図10は、本実施形態に係る第二の行分割処理の手順例を示すフローチャートである。S1001からS1006、S1009からS1011の各処理は、図6のS601からS606、S607からS609の各処理と同じであり、その説明を省略する。すなわち、図10は、図6に示すフローにS1007およびS1008を追加したフローチャートとなっている。

【0069】

CPU111は、黒画素数が最少値Sとなる位置Pを特定するS1006の処理に続くS1007において、位置Pに縦倍角文字が存在するか存在しないかを判定する。CPU111は、位置Pに縦倍角文字が存在するとの判定結果を得た場合（S1007のYES）、処理をS1008へ移行する。CPU111は、位置Pに縦倍角文字が存在しないとの判定結果を得た場合（S1007のNO）、処理をS1009へ移行する。

20

【0070】

S1008において、CPU111は、分割位置をS1006で特定した位置Pから変更し縦倍角文字の最下位置に設定し、分割位置として設定した縦倍角文字の最下位置にて行分割を実行する。CPU111は、行分割処理を実行した後、処理をS1003へ移行する。このように、文書に縦倍角文字を含んでいる場合でも、適切に設定した分割位置にて行分割を実行することができる。

【0071】

以上、説明したとおり、本実施形態によれば、文字サイズの異なる文字間で行接触している原稿のスキャン画像、および行接触した行の割合が多い原稿のスキャン画像においても、より正確な行分割を実現することができる。その結果、より正確なOCR精度を実現することができる。

30

【0072】

（実施形態3）

本実施形態については、実施形態1、2との差分のみの説明とし、特に明記しない部分については実施形態1、2と同じ構成および手順となる。図2に示す行分割処理のS201からS207、S210の処理は、実施形態1の場合と同じであり、その説明を省略する。

40

【0073】

実施形態2では、縦倍角文字を追加した3種類の文字属性に対応する文字を含む原稿をスキャンして得られたスキャン画像に対する行分割処理について説明した。文字属性の縦倍角文字の代わりに文字属性の通常文字とは異なる特徴（大きさ）の文字の文字属性に対応する文字を含む原稿をスキャンして得られたスキャン画像に対する行分割処理としてもよい。

【0074】

例えば、図2のS208において、CPU111は、文字高さの最頻値Hを導出する際に、文字高さ情報を用いて文字高さの頻度グラフを作成する。CPU111は、作成した頻度グラフにピークが複数存在する場合、例えば、文字属性をピーク数に応じて第二最頻

50

文字、第三最頻文字などと定義ためのピーク値を導出する。縦倍角文字の判定処理 S 9 0 9 を第二最頻文字、第三最頻文字を判定するための特殊文字の判定処理とする。S 9 0 9 の判定結果の処理を、判定結果に応じて文字属性を第二最頻文字または第三最頻文字に設定するか、文字属性を設定しない処理とする。

【 0 0 7 5 】

そして、位置 P での縦倍角文字の存在判定処理 S 1 0 0 7 を、位置 P での第二最頻文字、第三最頻文字の、文字の存在判定処理とする。位置 P に第二最頻文字、第三最頻文字の文字が存在すると判定された場合には、分割位置を S 1 0 0 7 で判定した判定結果に応じて第二最頻文字、第三最頻文字の特徴文字の最下位置にて行分割を実行する。特徴文字が存在しないと判定された場合には、処理を S 1 0 0 9 へ移行する。S 1 0 0 8 にて、処理を S 1 0 0 3 へ移行する。このように、原稿に特徴文字を含んでいる場合でも、適切に設定した分割位置にて行分割を実行することができる。

10

【 0 0 7 6 】

以上、説明したとおり、本実施形態によれば、文字サイズの異なる文字が複数存在する原稿、および行接触した行の割合が多い原稿において、異なるサイズの文字間で行接触している場合でも、より正確な行分割処理を実現することができる。その結果、より正確な OCR 精度を実現することができる。

【 0 0 7 7 】

(実施形態 4)

本実施形態については、実施形態 1 との差分のみの説明とし、特に明記しない部分については実施形態 1 と同じ構成および手順となる。

20

【 0 0 7 8 】

実施形態 2 では、文字属性として縦倍角文字を追加することで、通常文字と縦倍角文字との行接触でも正確に行分割できる行分割処理について説明した。第二の行分割処理条件を変えることでより正確な行分割処理を実現できる。

【 0 0 7 9 】

図 1 1 は、本実施形態に係る第二の行分割処理の手順例を示すフローチャートである。S 1 1 0 1 から S 1 1 0 9 の各処理は、図 6 の S 6 0 1 から S 6 0 9 の各処理と同じであり、その説明を省略する。すなわち、図 1 1 は、図 6 に示すフローに S 1 1 1 0 から S 1 1 1 4 を追加したフローチャートとなっている。

30

【 0 0 8 0 】

S 1 1 0 7 において、CPU 1 1 1 は、黒画素数最少値（カウント値の最少値）S が黒画素数平均値（カウント値の平均値）A 以上であるとの判定結果を得た場合（S 1 1 0 7 の NO）、処理を S 1 1 1 0 へ移行する。CPU 1 1 1 は、黒画素数最少値 S が黒画素数平均値 A 未満であるとの判定結果を得た場合（S 1 1 0 7 の YES）、処理を S 1 1 0 8 へ移行し、実施形態 1 と同様、S 1 1 0 8 にて、位置 P で行分割処理を実行するかの判定を実行する。

【 0 0 8 1 】

S 1 1 1 0 において、CPU 1 1 1 は、S 1 1 0 5 で特定した基準点から S 2 0 8 で取得した文字高さ最頻値 H だけ下へ進んだ位置を新たな基準点 2 を特定する。このとき、CPU 1 1 1 は、基準点 2 の周辺の黒画素数を取得する。本実施形態において、上述の実施形態 1 と同様、基準点の周辺の大きさは、基準点から ± 2 画素分に限定されない。基準点の周辺の大きさは、基準点から ± 1 画素としたり、基準点から ± 2 画素より大きくしたりしてもよい。

40

【 0 0 8 2 】

S 1 1 1 1 において、CPU 1 1 1 は、S 1 1 1 0 の処理で特定した新たな基準点 2 から、行と垂直な方向にて ± 2 画素分の黒画素数を比較して黒画素数が最少値 T となる位置 Q（第 2 の位置）を分割位置 8 0 4 として特定する。なお、基準点の周辺の大きさは、基準点から ± 2 画素分に限定されない。基準点の周辺の大きさは、基準点から ± 1 画素としたり、基準点から ± 2 画素より大きくしたりしてもよい。

50

【 0 0 8 3 】

S 1 1 1 2において、C P U 1 1 1は、S 1 1 1 1にて位置Qの特定に用いた黒画素数最少値TがS 1 1 0 4で導出した黒画素数平均値Aより小さいか否かを判定する。C P U 1 1 1は、黒画素数最少値Tが黒画素数平均値Aよりも小さい（黒画素数平均値A未満である）との判定結果を得た場合（S 1 1 1 2のY E S）、処理をS 1 1 1 3へ移行する。C P U 1 1 1は、黒画素数最少値Tが黒画素数平均値A以上であるとの判定結果を得た場合（S 1 1 1 2のN O）、処理をS 1 1 0 1へ移行する。

【 0 0 8 4 】

S 1 1 1 3において、C P U 1 1 1は、S 6 0 8の処理と同様、S 1 1 1 1で特定した位置Qで行分割処理を行った場合の、行分割後の行高さを導出する。すなわち、C P U 1 1 1は、S 1 1 1 1で特定した位置Qを分割位置として行分割を行った場合の、二値画像にて分割位置の下側の領域の高さであって、分割位置から行接触行の最下位置までの距離を導出する。そして、C P U 1 1 1は、導出した行分割後の行高さがS 2 0 8で導出した文字高さ最頻値Hよりも大きいと判定する。C P U 1 1 1は、導出した行分割後の行高さが文字高さ最頻値Hよりも大きいとの判定結果を得た場合（S 1 1 1 3のY E S）、処理をS 1 1 1 4へ移行する。C P U 1 1 1は、導出した行分割後の行高さが文字高さ最頻値H以下であるとの判定結果を得た場合（S 1 1 1 3のN O）、処理をS 1 1 0 1へ移行する。

10

【 0 0 8 5 】

S 1 1 1 4において、C P U 1 1 1は、処理対象である二値画像8 0 0に対してS 1 1 1 1で特定した位置Qにて行分割処理を実行する。この行接触行に対する行分割処理が第二の行分割処理に対応する。C P U 1 1 1は、第二の行分割処理を実施した後、処理をS 1 1 0 3へ移行する。

20

【 0 0 8 6 】

以上、説明したとおり、本実施形態によれば、文字サイズの異なる文字間で行接触している原稿や、および行接触した行の割合が多い原稿においても、文字属性を追加することなく、より正確な行分割処理を実現することができる。その結果、より正確なO C R精度を実現することができる。

【 0 0 8 7 】

（実施形態5）

本実施形態については、実施形態1との差分のみの説明とし、特に明記しない部分については実施形態1と同じ構成および手順となる。

30

【 0 0 8 8 】

実施形態1では、3行にわたり行接触が発生した行に対応する文字を含む文書の行分割処理について説明した。4行以上にわたって行接触が発生する行に対応する文字を含む文書の行分割処理としてもよい。

【 0 0 8 9 】

図1 2は、本実施形態に係る行分割処理の手順例を示すフローチャートである。S 1 2 0 1からS 1 2 1 1の各処理は、実施形態1におけるS 2 0 1からS 2 1 1の各処理と同じであり、その説明を省略する。すなわち、図1 2は、図2に示すフローにS 1 2 1 2およびS 1 2 1 3を追加したフローチャートとなっている。

40

【 0 0 9 0 】

C P U 1 1 1は、第二の行分割処理を行うS 1 2 1 1に続くS 1 2 1 2において、第三の行分割処理を実行する。第三の行分割処理は、第二の行分割処理と同じ手順である。第二の行分割処理では、行接触行の上端から基準点、および黒画素数が最少値となる位置を特定する。これに対し、第三の行分割処理では、行接触行の後尾位置（行下端位置）から基準点、および黒画素数が最少値となる位置を特定する。

【 0 0 9 1 】

S 1 2 1 3において、C P U 1 1 1は、第二の行分割処理結果と、第三の行分割処理結果を比較し、比較結果に応じて行分割結果を決定する処理を実行する。図1 3は、本実施

50

形態に係る行分割結果決定処理の手順例を示すフローチャートである。

【0092】

S1301において、CPU111は、S1211で特定した第二の行分割位置と、S1212で特定した第三の行分割位置を行ごとに比較する。

【0093】

S1302において、CPU111は、第二の行分割位置と第三の行分割位置が同じであるか否かを判定する。CPU111は、第二の行分割位置と第三の行分割位置が同じであるとの判定結果を得た場合(S1302のYES)、処理をS1303へ移行する。CPU111は、分割位置にずれがあり、第二の行分割位置と第三の行分割位置が同じではないとの判定結果を得た場合(S1302のNO)、処理をS1304へ移行する。

10

【0094】

S1303において、CPU111は、第二の行分割結果を最終的な行分割結果として採用する。これにより、行接触行に対する行分割処理は、第二の行分割処理で実行される行分割位置に確定される。

【0095】

S1304において、CPU111は、第二の行分割結果と第三の行分割結果のそれぞれに対して文字分割処理を実行し、続いてOCR処理を実行する。

【0096】

更に、S1305において、CPU111は、S1304で導出された各OCR処理結果を用いて平均文字信頼度を導出し、導出した各平均文字信頼度を比較する。平均文字信頼度は、各行分割結果に対応する文字信頼度の平均値である。CPU111は、第二の行分割結果に対応するOCR処理結果を用いて平均文字信頼度を導出する。CPU111は、第三の行分割処理結果に対応するOCR処理結果を用いて平均文字信頼度を導出する。CPU111は、導出した第二の行分割処理結果に対応する平均文字信頼度と、導出した第三の行分割処理結果に対応する平均文字信頼度とを比較する。

20

【0097】

S1306において、OCR処理結果が認識対象文字とOCR処理結果の文字との一致率が高いことを示す、平均文字信頼度の数値が低い行分割処理結果を最終的な行分割処理結果として採用する。これにより、行接触行に対する行分割処理は、平均文字信頼度に応じて採用された、第二の行分割処理または第三の行分割処理で実行される行分割位置に確定される。本実施形態では、行接触行全体の平均文字信頼度を比較した。行接触行内の各行の平均文字信頼度を導出して行ごとに比較し、より信頼度の良い方を採用しても良い。また同様に、文字ごとに文字信頼度を比較して、より信頼度の良い方を採用しても良い。

30

【0098】

以上、説明したとおり、本実施形態によれば、行間が狭く、かつ、非常に多くの行にわたり行接触が発生している原稿、および行接触した行の割合が多い原稿においても、より正確な行分割処理を実現することができる。その結果、より正確なOCR精度を実現することができる。

【0099】

図14は、CPU111がROM112に格納されたプログラムを実行する際の、上述した実施形態1、2、3、4、5の情報処理装置110の機能ブロック構成図である。図示の各構成要素1401から1406は、プログラムを実行したCPU111が行うものであるが、これらの一部がハードウェアにより実現しても構わない。情報処理装置110は、行・文字情報取得部1401、文字高さ情報抽出部1402、文字接触判定部1403、行接触判定部1404、カウント値導出部1405、行分割部1406を有する。

40

【0100】

行・文字情報取得部1401は、原稿を光学的に読み取ったスキャン画像を二値化して得た二値画像内における行単位の大きさや位置を示す行情報を取得する。また、行・文字情報取得部1401は、二値画像内における文字単位の大きさや位置を示す行情報を取得する。文字高さ情報抽出部1402は、文字情報から文字高さ情報を抽出して文字高さ情

50

報を取得する。

【 0 1 0 1 】

文字接触判定部 1 4 0 3 は、文字高さ情報に基づき、文字が高さ方向で他の文字と接触するかどうかを判定する。行接触判定部 1 4 0 4 は、行情報と文字接触判定部 1 4 0 3 による判定結果に基づいて、注目する行が高さ方向で他の行と接触するかどうかを判定する。

【 0 1 0 2 】

カウント値導出部 1 4 0 5 は、行接触判定部 1 4 0 4 により他の行と接触すると判定された注目する行において、当該注目する行における幅方向にて、黒画素の数をカウントしてカウント値を導出する。

【 0 1 0 3 】

行分割部 1 4 0 6 は、行接触判定部 1 4 0 4 により他の行と接触すると判定された注目する行に対し、カウント値において注目する行の高さ方向で最少となる位置を行分割位置として行分割を実行する。

【 0 1 0 4 】

[その他の実施形態]

上記では、行接触行の行上端位置より文字高さ最頻値だけ下へ進めた位置を基準点と決定し、基準点の周辺で行分割したときの分割位置の下側の領域の高さなどにて所定の条件を満たすときに基準点の周辺から基準に分割位置を特定する手法について説明した。分割位置の特定手法は、これに限定されない。行接触行の行下端位置より文字高さ最頻値だけ上へ進めた位置を基準点と決定し、基準点の周辺で行分割したときの分割位置の上側の領域の高さなどにて所定の条件を満たすときに基準点の周辺から基準に分割位置を特定する手法としてもよい。この場合、実施形態 1、2、4 の説明の上側を下側に変更するなど、上下方向を逆に読み替えて設定すればよい。

【 0 1 0 5 】

上記では、スキャナ 1 0 1 で光学的に原稿を読み取る複写機 1 0 0 を例に説明したが、このような機器に限定されない。例えば、光学的に原稿を読み取ることが可能なカメラと、情報処理装置 1 1 0 と通信可能な通信部とを備えた、スマートフォンやタブレットなどの携帯型の端末としてもよい。

【 0 1 0 6 】

本発明は、上述の実施形態の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1 以上の機能を実現する回路（例えば、ASIC）によっても実現可能である。

【 符号の説明 】

【 0 1 0 7 】

- 1 0 0 …… 複写機
- 1 1 0 …… 情報処理装置
- 8 0 1 …… 射影画像（行方向の射影結果）
- 8 0 2 …… 平均黒画素数

10

20

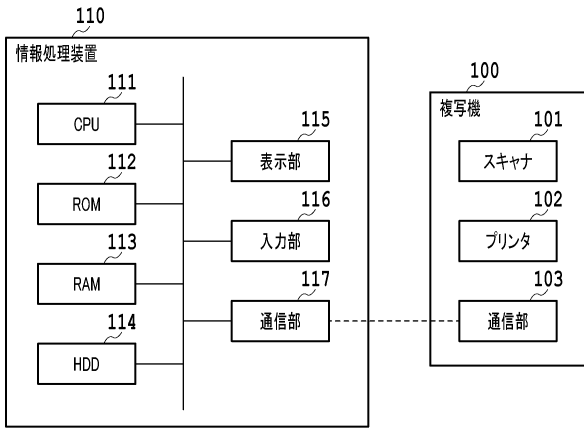
30

40

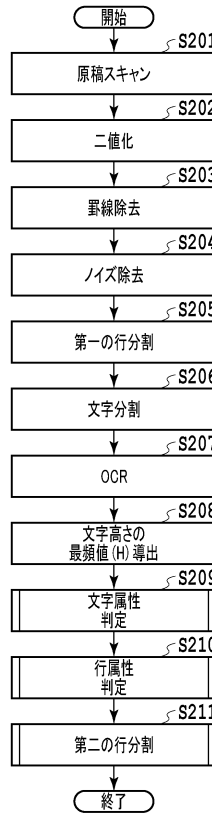
50

【 図面 】

【 図 1 】



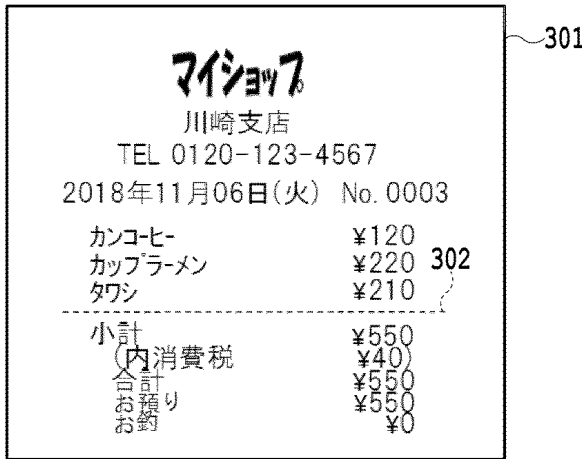
【 図 2 】



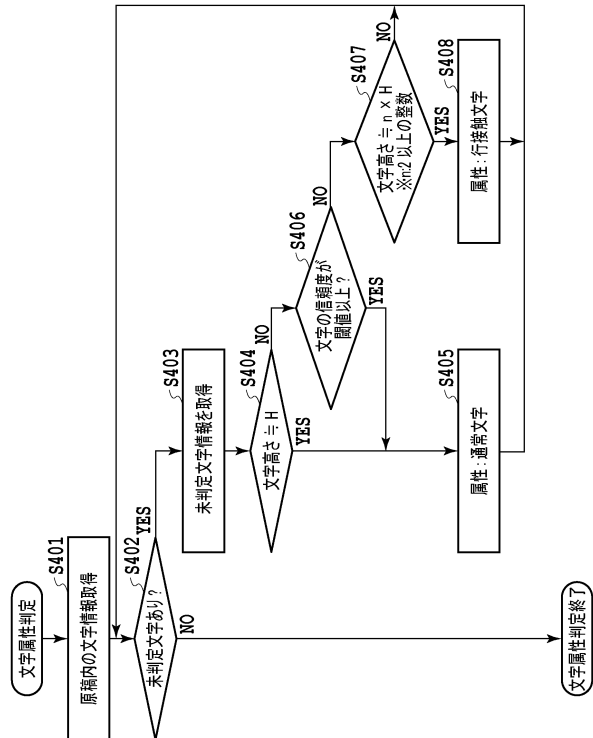
10

20

【 図 3 】



【 図 4 】

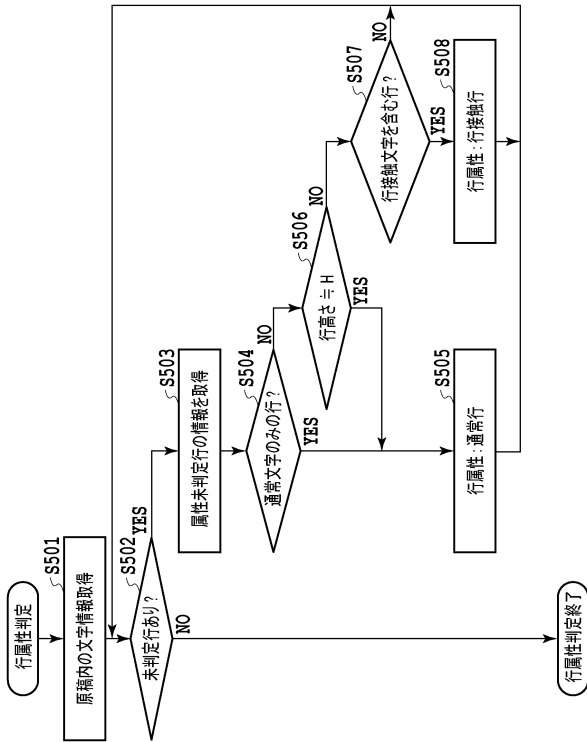


30

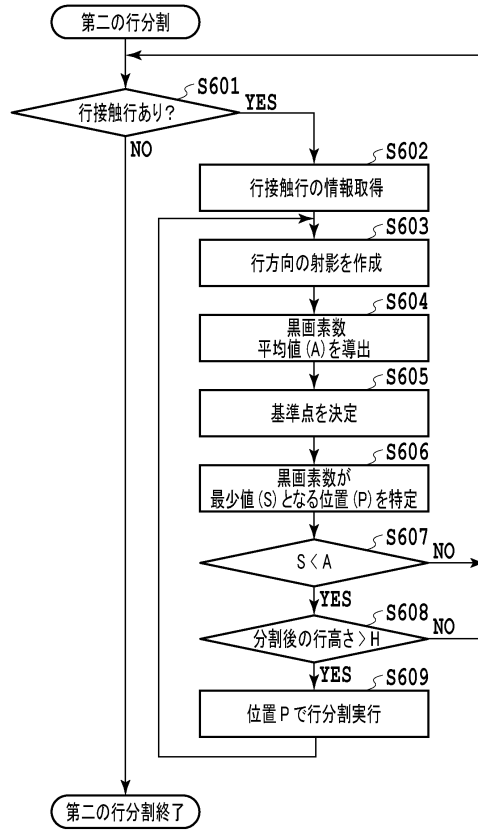
40

50

【図5】



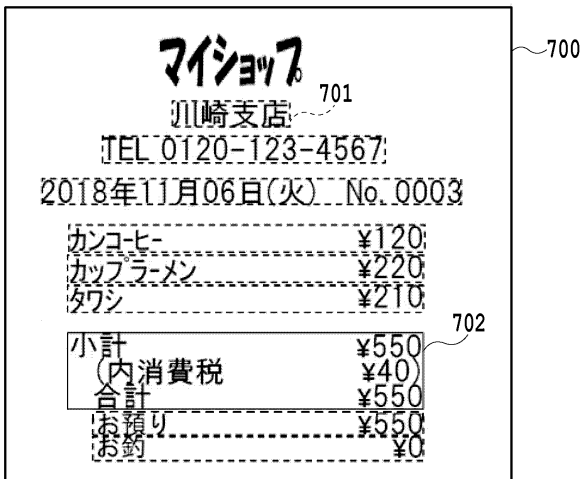
【図6】



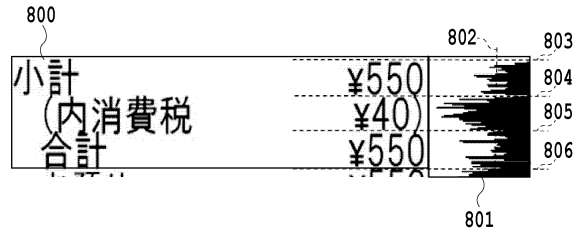
10

20

【図7】



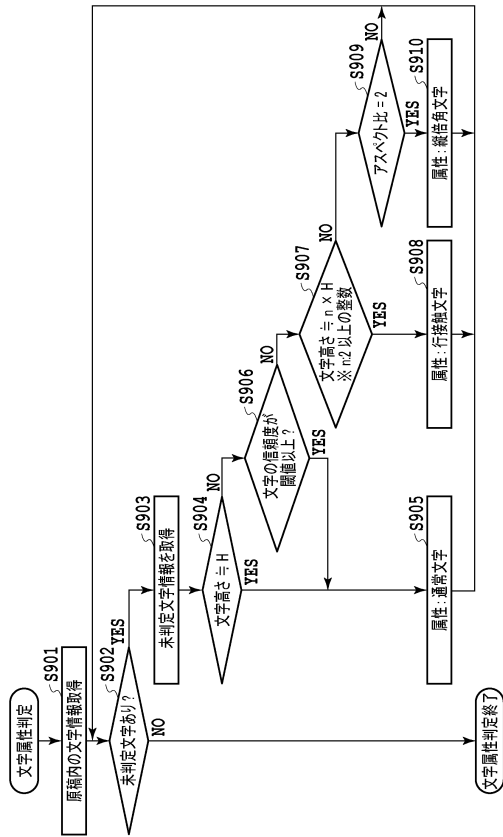
【図8】



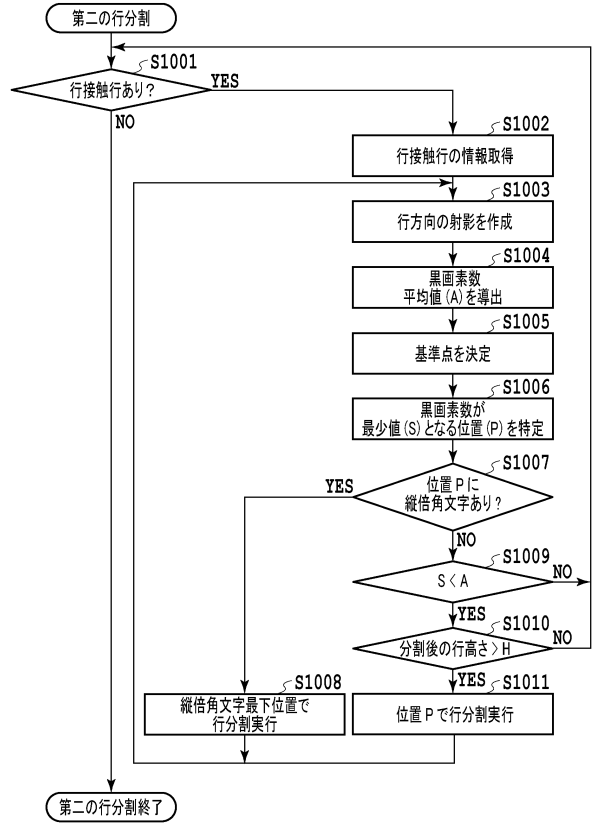
30

40

【図 9】



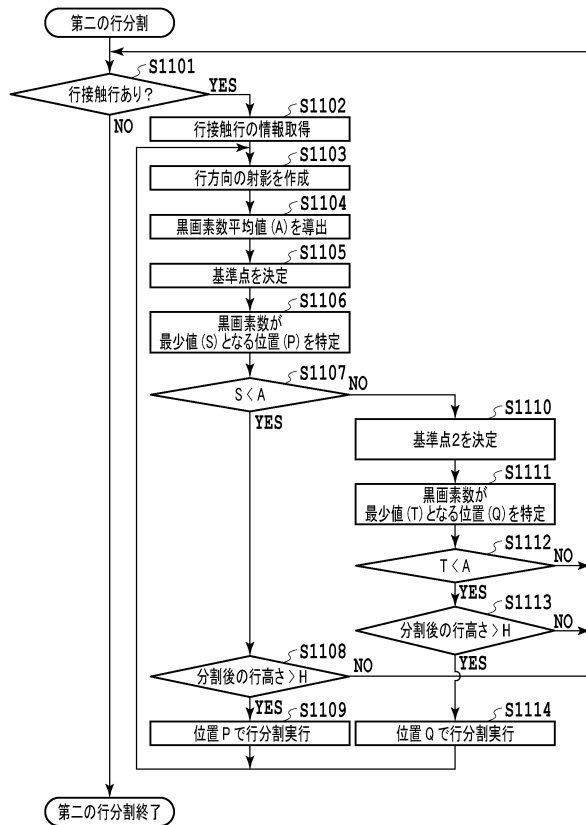
【図 10】



10

20

【図 11】



【図 12】

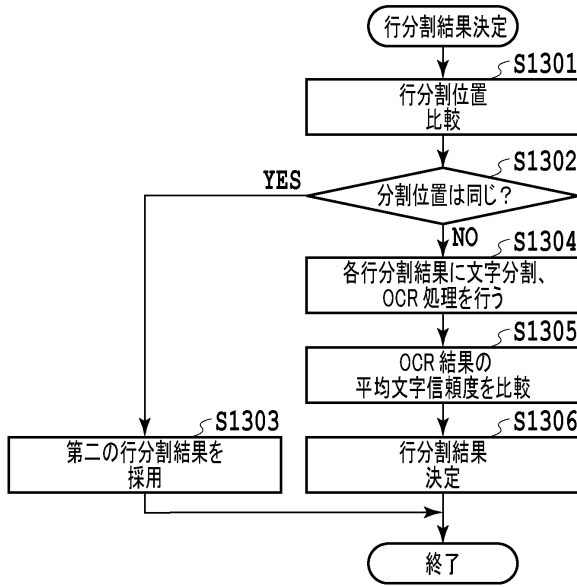


30

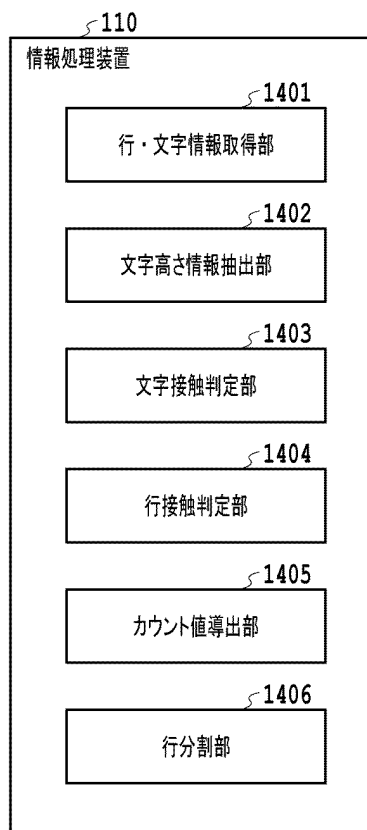
40

50

【 図 1 3 】



【 図 1 4 】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開平 8 - 3 0 5 7 9 3 (J P , A)
特開平 5 - 1 8 2 0 2 4 (J P , A)
特開平 7 - 0 1 3 9 9 4 (J P , A)
- (58)調査した分野 (Int.Cl. , D B 名)
G 0 6 V 3 0 / 1 4 8
G 0 6 V 3 0 / 1 4