

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7689439号
(P7689439)

(45)発行日 令和7年6月6日(2025.6.6)

(24)登録日 令和7年5月29日(2025.5.29)

(51)国際特許分類

F I

G 0 6 V 30/146(2022.01)

G 0 6 V 30/16 (2022.01)

G 0 6 V 30/146

G 0 6 V 30/16

請求項の数 7 (全13頁)

(21)出願番号	特願2021-67356(P2021-67356)	(73)特許権者	000001007
(22)出願日	令和3年4月12日(2021.4.12)		キヤノン株式会社
(65)公開番号	特開2022-162474(P2022-162474 A)	(74)代理人	100114775
(43)公開日	令和4年10月24日(2022.10.24)		弁理士 高岡 亮一
審査請求日	令和6年4月11日(2024.4.11)	(74)代理人	100121511
			弁理士 小田 直
		(74)代理人	100208580
			弁理士 三好 玲奈
		(72)発明者	鳴海 悠貴
			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
		(72)発明者	大石 剛
			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

最終頁に続く

(54)【発明の名称】 画像処理システム、画像処理方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

文書画像に対して、傾き補正を実施する画像処理システムにおいて、
手書き文字と活字とが混在する文書画像を取得する文書画像取得部と、
手書き文字を含む画像により手書き文字の特徴を学習させたニューラルネットワークを用いて、前記文書画像内に含まれる手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する分離部と、
前記生成された手書き文字以外の画像を用いて、傾き角度を推定する傾き角度推定部と、
前記推定された傾き角度に基づいて、前記手書き文字と前記活字とが混在する前記文書画像の傾きを補正する傾き補正部と、を有する
ことを特徴とする画像処理システム。

【請求項2】

前記分離部は、前記文書画像の傾き角度の範囲に対応する手法を適用して、前記文書画像から手書き文字を除外することにより、前記手書き文字以外の画像を生成することを特徴とする請求項1に記載の画像処理システム。

【請求項3】

前記分離部は、
前記文書画像の傾き角度の範囲が規定値以上であれば、第1のニューラルネットワークを用いて前記文書画像内に含まれる手書き文字の画素を判断し、当該判断された手書き文

字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する一方、

前記文書画像の傾き角度の範囲が前記規定値より小さければ、第2のニューラルネットワークを用いて前記文書画像内に含まれる手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成するものであり、

前記第2のニューラルネットワークの学習に用いた文書画像の傾きは、前記第1のニューラルネットワークの学習に用いた文書画像の傾きより小さい、
ことを特徴とする請求項1に記載の画像処理システム。

【請求項4】

前記分離部は、

傾き角度が異なる複数通りの手書き文字の画像により手書き文字の特徴を学習させた前記ニューラルネットワークを用いて、前記手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する

ことを特徴とする請求項1に記載の画像処理システム。

【請求項5】

前記分離部は、

傾き角度が異なる複数通りの手書き文字を含む文書画像により手書き文字の特徴を学習させた前記ニューラルネットワークを用いて、前記手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する

ことを特徴とする請求項1に記載の画像処理システム。

【請求項6】

文書画像に対して、傾き補正を実施する画像処理方法において、

手書き文字と活字とが混在する文書画像を取得する文書画像取得工程と、
手書き文字を含む画像により手書き文字の特徴を学習させたニューラルネットワークを用いて、前記文書画像内に含まれる手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する分離工程と、

前記生成された手書き文字以外の画像を用いて、傾き角度を推定する傾き角度推定工程と、

前記推定された傾き角度に基づいて、前記手書き文字と前記活字とが混在する前記文書画像の傾きを補正する傾き補正工程と、を有する

ことを特徴とする画像処理方法。

【請求項7】

コンピュータを、請求項1乃至5のいずれか1項に記載の画像処理装置として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、手書き文字と活字が混在する文書画像に対して、傾き補正を実施する技術に関する画像処理システム、画像処理方法及びプログラムに関する。

【背景技術】

【0002】

従来、画像形成装置でスキャンした文書画像データに光学文字認識処理（以下、OCR処理とする）を行うことで、画像データ内の文字列をデジタルデータとして抽出する技術が存在する。しかしながら、スキャンした文書画像が傾いている場合、正しいOCR処理が実施できない場合がある。そこで、OCR処理の前処理として、文書画像の傾き角度を推定し、正しい角度になるよう補正を行う技術（以下、傾き補正）が存在する（例えば、

10

20

30

40

50

特許文献 1、特許文献 2、特許文献 3)。

【 0 0 0 3 】

特許文献 1 に記載の技術では、文書画像の回転角の関数として画素の分散を測定し、分散が最大である文書回転角度（傾き角度）で、傾き補正を実施している。また、特許文献 2 に記載の技術では、表領域を検出後、罫線の傾きに基づいて入力画像の傾き補正を実施する。また、特許文献 3 に記載の技術では、文書画像のエッジを検出することで、画像の中身を確認することなく、傾き補正を実施する。

【先行技術文献】

【特許文献】

【 0 0 0 4 】

【文献】特開平 3 - 2 6 8 1 8 9 号公報

【文献】特開平 8 - 4 4 8 2 2 号公報

【文献】特開 2 0 2 0 - 5 3 9 3 1 号公報

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 5 】

しかしながら、特許文献 1 では、行間隔やピッチ、角度にバラつきのある手書き文字が多く混在する場合、正確な傾き角度を推定することが困難である。特に、活字に対して、手書き文字数の割合が多い場合、または、手書き文字の濃度が濃い（活字に対して輝度差が大きい）場合に、正確な傾き角度を推定できないおそれがある。また、特許文献 2 では、原稿画像に罫線情報がない場合、傾き補正が実施できないおそれがある。また、特許文献 3 では、エッジの検出ができない場合や、文書原稿が四角くない場合（やぶれ等）、正確な傾き補正が実施できないおそれがある。

【 0 0 0 6 】

本発明は、上記事情に鑑みてなされたものであり、手書き文字と活字が混在する文書画像に対する、傾き補正の精度を向上させた画像処理システムを提供することを目的とする。

【課題を解決するための手段】

【 0 0 0 7 】

上記の目的を達成するために、本発明における画像処理システムは、文書画像に対して、傾き補正を実施する画像処理システムにおいて、手書き文字と活字とが混在する文書画像を取得する文書画像取得部と、手書き文字を含む画像により手書き文字の特徴を学習させたニューラルネットワークを用いて、前記文書画像内に含まれる手書き文字の画素を判断し、当該判断された手書き文字の画素を前記文書画像から除去することにより、前記手書き文字以外の画像を生成する分離部と、前記生成された手書き文字以外の画像を用いて、傾き角度を推定する傾き角度推定部と、前記推定された傾き角度に基づいて、前記手書き文字と前記活字とが混在する前記文書画像の傾きを補正する傾き補正部と、を有することを特徴とする。

【発明の効果】

【 0 0 0 8 】

本発明によれば、手書き文字と活字が混在する文書画像に対する、傾き補正の精度を向上させることができる。

【図面の簡単な説明】

【 0 0 0 9 】

【図 1】本発明の実施形態における画像処理システムの一例を示すブロック図。

【図 2】第 1 実施形態における画像処理の手順を示すフローチャート。

【図 3】第 1 実施形態における文書画像と処理結果の一例を示す図。

【図 4】第 2 実施形態における画像処理の手順を示すフローチャート。

【図 5】第 2 実施形態における手書き文字分離のための学習データの一例を示す図。

【発明を実施するための形態】

【 0 0 1 0 】

10

20

30

40

50

以下、添付図面を参照して本発明の実施形態を詳しく説明する。なお、以下の実施形態は特許請求の範囲に係る本発明を限定するものでなく、また本実施形態で説明されている特徴の組み合わせの全てが本発明の解決手段に必須のものとは限らない。なお、同一の構成要素には同一の参照番号を付して、説明を省略する。

【0011】

< 画像形成システム概要 >

図1は、本発明の実施形態における画像処理システムの一例を示すブロック図である。図1に示すように、画像処理システムは、画像形成装置100と、ホストコンピュータ170と、サーバー191（クラウドサーバーであってもよい）を含む。

【0012】

本実施形態では、画像形成装置100として、印刷機能、読取機能、FAX機能等、複数の機能が一体化された複合機（MFP：Multi Function Printer）が用いられるものとして説明する。また、サーバー191は、文書管理機能を有するものとして説明する。画像形成装置100とホストコンピュータ170とサーバー191は、LAN（Local Area Network）190等のネットワークを相互に通信可能に接続されている。画像形成装置100とホストコンピュータ170とサーバー191は、複数台接続されていてもよく、他の装置が接続されていてもよい。また、ネットワークは、本実施形態ではLAN190を例示して説明するが、有線ネットワークや無線ネットワーク、もしくは、それらが組み合わされた構成の場合がある。

【0013】

画像形成装置100は、制御装置110、リーダー装置120、プリンタ装置130、操作部140、記憶装置150を含む。制御装置110は、リーダー装置120、プリンタ装置130、操作部140、記憶装置150のそれぞれと接続される。

【0014】

制御装置110は、画像形成装置100を統括的に制御する制御基板（コントローラ）である。制御装置110は、CPU111、ROM112、RAM113、画像処理部114を含む。

【0015】

CPU111は、システムバス（不図示）を介して、制御装置110内の各ブロックを制御する。例えば、CPU111は、ROM112やRAM113、記憶装置150、又は、他の記憶媒体に記憶されたプログラムを読み出して実行することにより、画像形成装置100の機能を実行する。

【0016】

ROM112は、例えば、制御プログラムや、画像形成装置100の機能を実行する上で必要なテーブルや設定データ等を記憶する。RAM113は、例えば、CPU111のワークメモリとして用いられる。

【0017】

画像処理部114は、リーダー装置120によって生成された読取画像データや、外部から受信した画像データに対して、変換、補正、編集、圧縮／解凍など、種々の画像処理を実行する。画像処理部114は、ハードウェアで構成される場合があれば、ソフトウェアで実現される場合もある。

【0018】

リーダー装置120は、スキャナエンジンの構成を有し、原稿を光学的に読み取る原稿スキャン処理をし、光学的に読み取った原稿から読取画像データ（文書画像）を生成する。原稿スキャン処理は、原稿台にセットされた原稿を光学的に読み取る方法でもよいし、自動原稿給送装置（ADF：Automatic Document Feeder）から給送された原稿を光学的に読み取る方法でもよい。

【0019】

プリンタ装置130は、インクジェット記録方式や電子写真方式等、各種の記録方式に対応したプリンタエンジンの構成を有する。これにより、プリンタ装置130は、記録媒

10

20

30

40

50

体上に画像を形成する。

【 0 0 2 0 】

操作部 1 4 0 は、ユーザーの操作を受付けるための操作キー、及び、各種設定やユーザーインターフェース画面の表示などを行う液晶パネルを備える。操作部 1 4 0 は、ユーザー操作等によって受け付けた情報を制御装置 1 1 0 へ出力する。

【 0 0 2 1 】

記憶装置 1 5 0 は、ユーザー情報を記憶する。ユーザー情報としては、例えば、画像データや、モードやライセンスなどの機器情報や、アドレス帳やカスタマイズなどがある。

【 0 0 2 2 】

なお、画像形成装置 1 0 0 は、図 1 に示す構成に限られず、画像形成装置 1 0 0 の実行可能な機能に応じて他の構成を含む。例えば、FAX 機能の実行に必要な構成や、近距離無線通信を可能とする構成を含む場合もある。

10

【 0 0 2 3 】

サーバー 1 9 1 は、制御装置 1 9 8、操作部 1 9 5、記憶装置 1 9 6、表示部 1 9 7 を含む。制御装置 1 9 8 は、操作部 1 9 5、記憶装置 1 9 6、表示部 1 9 7 のそれぞれと接続される。

【 0 0 2 4 】

制御装置 1 9 8 は、サーバー 1 9 1 を統括的に制御する制御基板（コントローラ）である。制御装置 1 9 8 は、CPU 1 9 2、ROM 1 9 3、RAM 1 9 4 を含む。

【 0 0 2 5 】

CPU 1 9 2 は、システムバス（不図示）を介して、制御装置 1 9 8 内の各ブロックを制御する。例えば、CPU 1 9 2 は、ROM 1 9 3 や RAM 1 9 4、記憶装置 1 9 6、又は、他の記憶媒体に記憶されたプログラムを読み出して実行することにより、サーバー 1 9 1 の機能を実行する。

20

【 0 0 2 6 】

ROM 1 9 3 は、例えば、オペレーティングシステムプログラム（OS）等の各種制御プログラムや、サーバー 1 9 1 の機能を実行する上で必要なテーブルや設定データ等を記憶する。RAM 1 9 4 は、例えば、CPU 1 9 2 のワークメモリとして用いられる。

【 0 0 2 7 】

操作部 1 9 5 は、ユーザーの操作を受付けるためのキーボードやポインティングデバイス等を備え、受け付けたユーザー操作等の情報を制御装置 1 9 8 へ出力する。記憶装置 1 9 6 は、例えば、各種アプリケーションプログラムや、データ、ユーザー情報、機器情報などを記憶する。表示部 1 9 7 は、例えば液晶ディスプレイであり、各種ユーザーインターフェース画面や情報の表示を行う。

30

【 0 0 2 8 】

ホストコンピュータ 1 7 0 は、LAN 1 9 0 を介して、画像形成装置 1 0 0 やサーバー 1 9 1 のそれぞれと接続される。この構成により、ホストコンピュータ 1 7 0 による操作や指示に基づいて、画像形成装置 1 0 0 やサーバー 1 9 1 を操作することもできる。

【 0 0 2 9 】

上述のような構成の画像処理システムを例に挙げ、具体的な実施形態を以下に述べる。なお、下記の実施形態で用いる「手書き文字」とは、人の手によって手書き入力された文字を指す。

40

【 0 0 3 0 】

[第 1 実施形態]

手書き文字と活字が混在する文書の場合、文字間隔やピッチが不均一な手書き文字の影響により、従来の傾き補正が失敗する可能性がある。本実施形態では、手書き文字がある場合、手書き文字の影響を除外した傾き角度推定用の画像を生成して、傾き角度推定を実施する。

【 0 0 3 1 】

図 2 は、第 1 実施形態における画像処理の手順を示すフローチャートである。図 3 は、

50

第 1 実施形態における入力文書画像と処理結果の一例を示す図である。なお、以下の説明で、「傾き」とは、傾きの基準となる左右方向に延びる基準線 L（図 3 参照）に対する入力文書画像の傾き角度をいう。

【 0 0 3 2 】

以下、図 2 を用いて画像処理の手順の説明を進めるが、必要に応じて、図 3 を参照して、画像処理の手順を説明する。図 2 の処理は、例えば、CPU 1 1 1 が、ROM 1 1 2 に記憶されたプログラムを RAM 1 1 3 に読みだして実行することで実現される。

【 0 0 3 3 】

ステップ S 2 0 1 において、入力文書画像の取得処理を行う。入力文書画像とは、画像処理システムに入力される文書画像 3 0 0 のことである。入力文書画像の取得処理において、CPU 1 1 1 は、操作部 1 4 0 を介してユーザーからの原稿スキャン処理の指示を受け付けると、リーダー装置 1 2 0 にスキャン指示を行い、スキャンを実施する。これにより、原稿に対応した読取画像データ（文書画像）を取得する。図 3 に示す文書画像 3 0 0 は、ステップ S 2 0 1 の文書画像取得処理で取得した入力文書画像の一例である。原稿が原稿台に斜めにセットされる場合には、リーダー装置 1 2 0 により取得された文書画像 3 0 0 にも傾きが生じる。また、ADF を用いて原稿を読み込んだ場合であっても、原稿のセットの仕方や搬送用モータの左右の速度差等により、取得された文書画像に傾きが生じる場合がある。そのため、取得した文書画像の傾きを識別し、補正することが必要である。このように、CPU 1 1 1 は、画像処理システムにおいて、文書画像を取得する文書画像取得部として機能する。

【 0 0 3 4 】

ステップ S 2 0 2 において、手書き文字分離処理を行う。手書き文字の分離処理において、CPU 1 1 1 は、ステップ S 2 0 1 で生成した読取画像データに対して、手書き文字の記載個所を分離する処理を行う。これにより、手書き文字の画像と手書き文字以外の画像を生成する。ここで生成した手書き文字以外の画像を、本分離処理の後、傾き角度推定のために用いる。このように、CPU 1 1 1 は、画像処理システムにおいて、手書き文字と判断された手書き文字の画像と、手書き文字と判断されない手書き文字以外の画像とに分離する手書き文字分離部として機能する。

【 0 0 3 5 】

本実施形態の手書き文字分離手法では、まず、画像内の手書き文字領域とそれ以外の背景領域を、ニューラルネットワーク（NN）に学習させる。次に、ニューラルネットワークの学習に基づいて、画素ごとに手書きか手書きでないかを判断する。これにより、手書き文字の画像特徴と一致する場合に手書き文字と判断し、画素を抽出することが可能である。例えば、図 3 の文書画像 3 0 0 の読取画像データに対して本処理を行うことで、画素 3 1 0 ~ 3 1 2 に示すような画素を手書き文字と判断する。次に、手書き文字と判断された画素 3 1 0 ~ 3 1 2 を除去することで、手書き文字以外の文書画像 3 0 1 が得られる。なお、本実施形態の手書き文字分離手法は一例であって、手書き文字を分離する手法を本実施形態の手法に限定するものではない。

【 0 0 3 6 】

従来の手書き文字分離は、文字種別に特化した OCR 処理への投入を目的としていた。なお、OCR 処理とは、光学的文字認識（OCR：Optical Character Recognition）による文字データの抽出処理のことをいう。

【 0 0 3 7 】

これに対し、本実施形態では、傾き認識の妨げにならないよう手書き文字を分離する。つまり、本分離処理により手書き文字を除外して、以後の傾き角度推定に用いる画像を生成している。手書き文字以外の文書画像 3 0 1 のような、手書き文字を除外した画像を用いることで、傾き角度の推定精度を向上させることが期待できる。

【 0 0 3 8 】

ステップ S 2 1 1 で、CPU 1 1 1 は、読み取った文書画像 3 0 0 に手書き文字が混在するかを判断する。ステップ S 2 0 2 において、手書き文字として抽出できた画素が一定

10

20

30

40

50

量以上の場合、手書き文字がある (Y e s) と判断し、ステップ S 2 0 2 の処理へ進む。一方、ステップ S 2 0 2 において、手書き文字として抽出できた画素が一定量を下回る場合、手書き文字がない (N o) と判断し、ステップ S 2 1 2 の処理へ進む。

【 0 0 3 9 】

本実施形態では、ステップ S 2 1 1 において手書き文字があるかないかの判断において、ステップ S 2 0 2 で画像分離した手書き文字の画像の画素と手書き文字以外の画像の画素の割合に基づいて、手書き文字の有無を判断した。手書き文字として抽出した画素が一定量を下回る場合、画像ノイズ等の可能性が高く、傾き角度推定に及ぼす影響はほとんど無い。あるいは、真に手書き文字を抽出した場合であっても、対する活字など手書き文字以外の画素の方が一定割合以上あれば、傾き角度推定に影響を与えることはほとんど無い。従って、手書き文字の有無を判断する際には、ステップ S 2 0 2 で画像分離した手書き文字の画像の画素が、手書き文字以外の画像の画素よりも一定割合以上多い場合に、手書き文字あり (Y e s) という判断を行った。

10

【 0 0 4 0 】

ステップ S 2 0 3 において、傾き角度推定処理を行う。傾き角度推定処理では、C P U 1 1 1 は、ステップ S 2 0 2 で生成した手書き文字以外の文書画像 3 0 1 を用いて傾き角度推定を行う。行間隔やピッチ、角度にバラつきのある手書き文字を除外し、手書き文字以外の文書画像 3 0 1 を用いて傾き角度推定を行うことで、傾き角度推定の精度が向上する。このように、C P U 1 1 1 は、画像処理システムにおいて、手書き文字以外の画像の傾き角度を推定する傾き角度推定部として機能する。

20

【 0 0 4 1 】

本実施形態で用いた傾き角度 (回転角度) を推定する手法は、文書画像内の文字列や線が印字前のデータ上では水平方向に並んでいることを利用する。例えば、様々な方向への射影ヒストグラムを取り、ヒストグラムのピークとボトムが短い周期で大きく振動するヒストグラムに対応する角度を選択することにより、傾き角度の推定が可能である。これは、正しい方向への射影であれば、同一ラインの文字列や、同一方向の罫線等の水平の線が、ヒストグラム上の同一のビンに投票され、行間にあたる部分には何も投票されないため、大きな振幅が文字間の周期で生じるためである。

【 0 0 4 2 】

ここまでの手法で推定した角度は、文字の向きが考慮されておらず、180度の不確実性が存在する。文字の向きについては、簡易的な文字認識処理を行った際の文字の確からしさ情報を用いて、判定することが可能である。これにより、文字の向きも考慮した角度情報を算出することが可能になる。この傾き角度推定手法は、行間隔が均一で行間隙間が所定の隙間以上に空いており、横ストロークが水平となる活字や罫線主体の文書において有効な手法である。従って、手書き文字以外の文書画像 3 0 1 のような、活字ベースの文書においては、傾き角度を正確に判断することができる。なお、上述の水平方向とは、図3における基準線 L と平行な方向であることをいう。基準線 L は、紙面における左右方向に延びる線であり、紙面における上下方向と直交する線である。

30

【 0 0 4 3 】

本実施形態で、手書き文字以外の文書画像 3 0 1 に対して、傾き角度推定処理を実施することで、基準線 L に対する傾き角度 を得ることができる。ただし、画像の傾き角度を識別する手法については特定の手法に限定するものではない。

40

【 0 0 4 4 】

ステップ S 2 1 2 において、C P U 1 1 1 は、ステップ S 2 1 1 で手書き文字がないと判断された文書画像について、傾き角度推定処理を実施する。傾き角度推定処理は、ステップ S 2 0 3 で行った処理と同様である。

【 0 0 4 5 】

ステップ S 2 1 3 において、ステップ S 2 0 3 およびステップ S 2 1 2 で推定した傾き角度に基づいて、文書画像に傾きがあるかを判断する。傾き角度が一定角度以上ある場合、傾きがある (Y e s) と判断し、ステップ S 2 0 4 の傾き補正処理へ進む。一方、傾き

50

角度が一定角度に満たない場合、傾きがない（No）と判断し、傾き補正をスキップし、ステップS205のOCR処理へ進む。

【0046】

ステップS204で、CPU111は、ステップS203またはステップS212で推定した傾き角度を用いて、S201で取得した文書画像に対して傾き補正処理を実施する。本実施形態における傾き補正は、ステップS203およびステップS212で推定した傾き角度による回転座標変換を行った。なお、補正手段はこれに限定するものではない。本実施形態では、図3に示す傾き角度を用いて、手書き文字以外の文書画像301に対して本傾き補正処理を行うことで、傾き補正後の補正画像302を得ることができる。傾き補正処理の後、ステップS205のOCR処理へ進む。このように、CPU111は、画像処理システムにおいて、傾き角度に基づいて文書画像300を補正する傾き補正部として機能する。

10

【0047】

ステップS205で、CPU111は、ステップS204で補正した補正画像302に対してOCR処理を行う。本実施形態においては、ステップS202で分離した手書き文字と、手書き文字以外の文書画像301について、それぞれ、手書き用、活字用に特化したOCR処理を実施する。その後、手書き文字のOCR結果と、手書き文字以外の文書画像301のOCR結果とを、マージする処理を行った。

【0048】

また本実施形態においては、OCR処理の前に文字列領域を判定し、文字列領域となった各領域についてOCR処理を実施し、文字列領域における文字列の文字コードを取得する。この領域判定により、文字列領域以外の領域を処理しなくて済む。この結果、処理負荷の軽減や文字認識の精度を向上することができる。なお、OCR処理については、様々な手法が考案されており、本実施形態の手法に限定するものではない。

20

【0049】

ステップS206で、CPU111は、文書画像300や補正画像302にステップS205で得たテキスト情報を付加したものを、記憶装置150にデータを登録して、本処理を終了する。データを登録する際に、文書画像を画像データのままでなく、OCR処理結果を用いて、全文検索可能なPDFなどの文書フォーマットに変換してもよい。

【0050】

本実施形態では、画像形成装置100上で全ての処理を行ったが、これに限るものではない。例えば、処理負荷を分散するために、LAN190を介して、ステップS201で生成した読取画像データを、サーバー191に送信し、サーバー191でユーザーからの操作受付以外の処理を行ってもよい。

30

【0051】

[第2実施形態]

本実施形態では、第1実施形態における手書き文字を分離する処理（図2のステップS202の処理）において、傾きの大きい文書画像に対しても、手書き文字の分離精度を高く維持できる手法について説明する。傾きの大きい文書画像に対しては、手書き文字の分離精度を維持する処理を実行し、傾きの小さい文書画像に対しては、十分な精度が発揮できるだけのシンプルな処理を実行する。図4は、第2実施形態における画像処理の手順を示すフローチャートである。以下、第2実施形態について、第1実施形態と異なる点を主に説明する。

40

【0052】

ステップS410で、CPU111は、ステップS201で取得した文書画像の傾きの範囲を判断する。傾きの範囲とは、どの程度傾いた画像が入力される可能性があるか、その取り得る傾きの程度の幅を指す。例えば、ADFで取得した文書画像より、原稿台にセットして取得した文書画像の方が置き方の自由度が高いため、傾きの範囲は大きいと言える。また、ADFであっても、原稿のセットの仕方や搬送用モータの左右の速度差等により文書画像に傾きが生じる可能性がある。特に、マルチサイズの前稿に対応可能なADF

50

を利用する場合は、特定原稿サイズ対応の A D F を利用する場合と比べて、傾きの範囲が大きくなる。このように、C P U 1 1 1 は、画像処理システムにおいて、傾き角度が取り得る範囲を判断する傾き角度範囲判断部として機能する。

【 0 0 5 3 】

例えば、本実施形態で用いる画像形成装置 1 0 0 における A D F は、小サイズ（はがき、領収書など）などの原稿サイズを自動で検知する。この検知した原稿サイズが、給送可能な最大原稿サイズより小さい場合、セット位置のずれなどで原稿が傾き易い。この場合は、傾きの範囲が規定以上である（ Y e s ）と判断する。また、原稿台から原稿を読み取って取得した場合も、傾きの範囲が規定以上である（ Y e s ）と判断する。対して、A D F で検知した原稿サイズが、給送可能な最大原稿サイズである場合は、傾きの範囲は小さく、規定を下回る（ N o ）と判断する。このように、ステップ S 4 1 0 においては、文書画像が傾く可能性が大きく、傾きの範囲が規定以上の場合（ Y e s ）、ステップ S 4 0 1 に進む。一方、傾きの範囲が規定を下回る場合（ N o ）、ステップ S 4 0 2 に進む。

10

【 0 0 5 4 】

本実施形態で用いた手書き文字の分離手法は、画像内の手書き文字領域とそれ以外の背景領域をニューラルネットワーク（ N N ）に学習させて、画素ごとに手書きかそうでないかを判断する、といった手法である。以下、入力文書画像の傾き角度の範囲が規定以上であるか否かによって、場合分けをして説明する。

【 0 0 5 5 】

ステップ S 4 0 1 において、ステップ S 4 1 0 で入力文書画像の傾き角度の範囲が規定以上と判断されているため、C P U 1 1 1 は、その範囲に対応可能な、傾き角度の大きい文書用の手書き文字分離処理を行う。ステップ S 4 0 1 の処理で用いるニューラルネットワークには、手書き文字の画像データとして、手書き文字の傾き角度を異ならせた複数通りのパターン画像を学習させる。図 5 は、第 2 実施形態における手書き文字分離のための学習データの一例を示す図である。図 5 に示すように、様々な角度で手書き文字を学習させることで、傾き角度の大きい入力文書画像に対しても、手書き文字の抽出精度を維持することが可能である。

20

【 0 0 5 6 】

なお、本実施形態では、手書き文字の傾き角度を異ならせることとしたが、これに限るものではない。例えば、ニューラルネットワークの学習画像の傾き角度の範囲を制限して、本処理において、入力文書画像の傾き角度が異なるようにしながら手書き文字の抽出を行い、取り得る傾き角度を網羅することとしてもよい。

30

【 0 0 5 7 】

ステップ S 4 0 2 において、ステップ S 4 1 0 で入力文書画像の傾きの範囲が規定を下回ると判断されているため、C P U 1 1 1 は、その範囲に対応可能な、傾きが小さい文書用の手書き文字分離処理を行う。ステップ S 4 0 2 の処理で用いるニューラルネットワークには、手書き文字分離のための画像データとして、ステップ S 4 0 1 の処理で用いるニューラルネットワークよりも少ない数の回転パターンの画像を学習させる。学習パターンの少ないシンプルなネットワーク構造のニューラルネットワークを利用した方が、同一精度を目指した場合に、推論のコストを低減することが可能である。そのため、傾き範囲が小さいと思われる場合には、十分な精度が見込めるシンプルなネットワーク構造を用いる。本処理で用いるニューラルネットワークには、図 5 の画像 5 0 1 に示すような、正常の向きの画像のみを学習に使用する。

40

【 0 0 5 8 】

このように、本実施形態においては、傾きの範囲が大きい文書画像に対して、回転した複数パターンの手書き文字を学習したニューラルネットワークを用いて手書き文字の分離処理を行う。このため、手書き文字の分離精度を高く維持することが可能である。また、傾きの範囲が小さい文書画像に対しては、学習パターンを減らしたニューラルネットワークを用いて、手書き文字の分離処理を行う。このため、シンプルな処理で十分な精度を発揮することが可能である。

50

【 0 0 5 9 】

[その他の実施形態]

以上、本発明をその好適な実施形態に基づいて詳述したが、本発明はこれら特定の実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の様々な形態も本発明に含まれる。また、上述の実施形態の一部を適宜組み合わせてもよい。特に、上述の実施形態においては、画像処理をするCPUとして、画像形成装置100のCPU111を例示したが、サーバー191のCPU192を用いてもよい。

【 0 0 6 0 】

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つ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 0 6 1 】

100 ... 画像形成装置
110 ... 制御装置
120 ... リーダー装置
191 ... サーバー
198 ... 制御装置

10

20

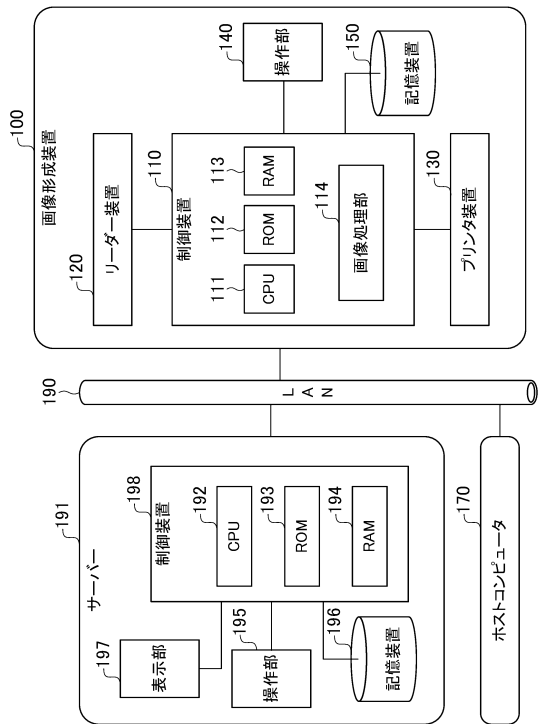
30

40

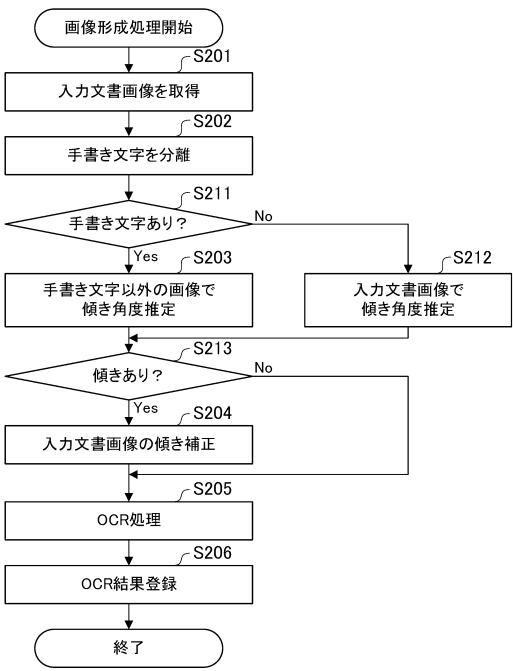
50

【図面】

【図 1】



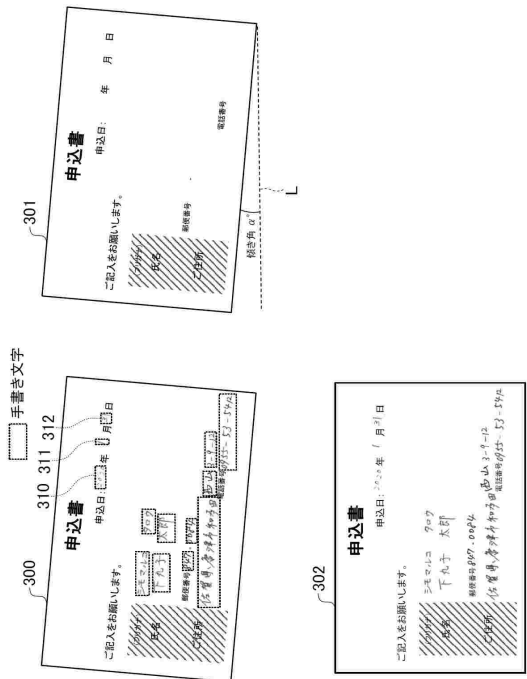
【図 2】



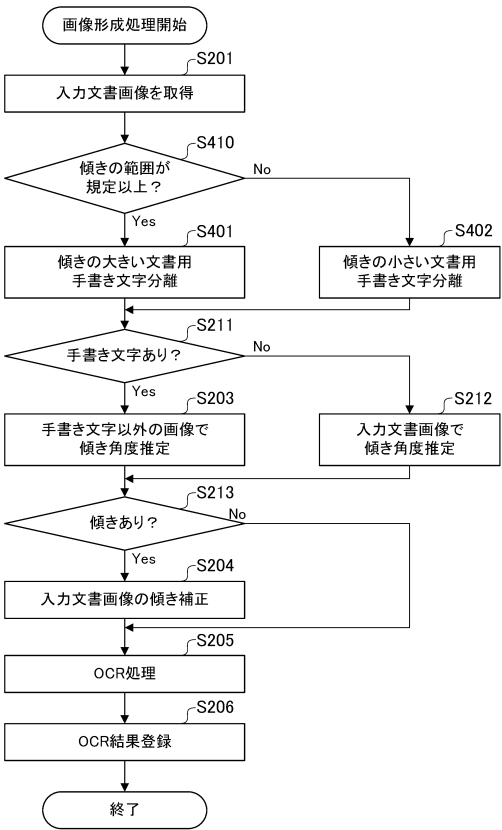
10

20

【図 3】



【図 4】

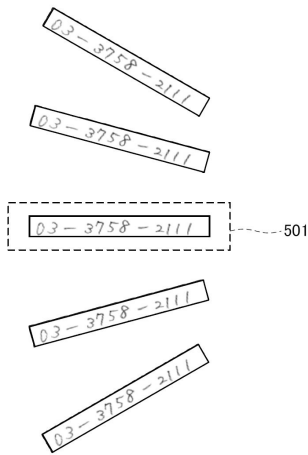


30

40

50

【 図 5 】



10

20

30

40

50

フロントページの続き

審査官 岡本 俊威

(56)参考文献 米国特許出願公開第 2 0 0 3 / 0 2 1 5 1 3 6 (U S , A 1)

(58)調査した分野 (Int.Cl., D B 名)

G 0 6 V 3 0 / 0 0 - 3 0 / 1 8 6