

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7516170号
(P7516170)

(45)発行日 令和6年7月16日(2024.7.16)

(24)登録日 令和6年7月5日(2024.7.5)

(51)国際特許分類		F I	
G 0 6 V	30/412 (2022.01)	G 0 6 V	30/412
G 0 6 V	30/14 (2022.01)	G 0 6 V	30/14 3 4 0 K
H 0 4 N	1/387(2006.01)	H 0 4 N	1/387 7 0 0
		H 0 4 N	1/387 8 0 0

請求項の数 19 (全30頁)

(21)出願番号	特願2020-148383(P2020-148383)	(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2年9月3日(2020.9.3)	(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2021-144673(P2021-144673 A)	(72)発明者	宮内 崇 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3年9月24日(2021.9.24)	審査官	小池 正彦
審査請求日	令和5年9月1日(2023.9.1)		
(31)優先権主張番号	特願2020-43075(P2020-43075)		
(32)優先日	令和2年3月12日(2020.3.12)		
(33)優先権主張国・地域又は機関	日本国(JP)		

最終頁に続く

(54)【発明の名称】 画像処理装置、画像処理方法、およびプログラム

(57)【特許請求の範囲】

【請求項1】

入力画像におけるテキストブロックを検出する検出手段と、
複数の登録文書の中から、前記入力画像に対応する登録文書を特定する特定手段と、
前記特定された登録文書において規定されている、処理対象の項目に対応する第1のテキストブロックと前記第1のテキストブロックの近傍に存在する少なくとも1つの第2のテキストブロックとを含む部分レイアウトに基づき、前記入力画像における前記処理対象の項目に対応するテキストブロックの決定をする決定手段と、
前記決定されたテキストブロックに対して文字認識処理を行うことにより、前記入力画像における前記処理対象の項目に対応する文字列を取得する取得手段と、
を有することを特徴とする画像処理装置。

【請求項2】

前記決定手段は、
前記入力画像における探索範囲の何れかの位置に前記部分レイアウトを重畳させて、前記部分レイアウトに含まれるテキストブロックと前記入力画像におけるテキストブロックとが重なる領域の大きさに基づく一致度を導出することにより、前記決定を行う
ことを特徴とする請求項1に記載の画像処理装置。

【請求項3】

前記決定手段は、
前記特定された登録文書における前記第1のテキストブロックの位置に対応する前記入

力画像における位置を含む所定の領域を前記探索範囲として、前記一致度の導出を行うことを特徴とする請求項 2 に記載の画像処理装置。

【請求項 4】

前記決定手段は、

前記部分レイアウトに含まれるテキストブロックと前記探索範囲内のテキストブロックとの縦方向の位置の差に基づき、前記決定のための縦方向の位置群を導出する

ことを特徴とする請求項 2 または 3 に記載の画像処理装置。

【請求項 5】

前記決定手段は、

前記探索範囲内の前記縦方向の位置群の横方向に、前記部分レイアウトを重畳した場合の、夫々の位置における前記一致度を導出する

ことを特徴とする請求項 4 に記載の画像処理装置。

10

【請求項 6】

前記決定手段は、

前記一致度が閾値以上であり、かつ、前記一致度が最大となる前記入力画像における位置を、前記決定を行うための前記入力画像における位置として導出する

ことを特徴とする請求項 2 から 5 のいずれか 1 項に記載の画像処理装置。

【請求項 7】

前記決定手段は、

前記一致度が閾値以上となる前記入力画像における候補位置を導出し、

前記候補位置の数が 1 つの場合、前記候補位置を、前記決定を行うための前記入力画像における位置として導出し、

前記候補位置の数が 2 つ以上の場合、前記部分レイアウトに含まれるテキストブロックを前記特定された登録文書内の何れかの位置に重畳させた場合に前記一致度を導出する方法と同一の方法で導出された一致度が閾値以上となる前記特定された登録文書内の位置を類似位置として取得し、前記候補位置と、前記類似位置とを対応付けることにより、前記決定を行うための前記入力画像における位置を導出する

ことを特徴とする請求項 5 に記載の画像処理装置。

20

【請求項 8】

前記決定手段は、

前記候補位置の数が 2 つ以上の場合で、かつ、前記候補位置の数と前記類似位置の数とが同じ場合、

同一条件で並べられた前記候補位置と前記類似位置とを一方の側から順に対応付けを行った結果、前記第 1 のテキストブロックに対応する前記類似位置に対応付けられた前記候補位置を、前記決定を行うための前記入力画像における位置として導出する

ことを特徴とする請求項 7 に記載の画像処理装置。

30

【請求項 9】

前記候補位置の数が 2 つ以上の場合で、かつ、前記候補位置と前記類似位置の数が異なる場合、

同一条件で並べられた前記候補位置と前記類似位置とをそれぞれ一方の側から順に対応付けを行った結果、前記第 1 のテキストブロックに対応する前記類似位置に対応付けられた前記候補位置が示す第 1 の位置と、

前記同一条件で並べられた前記候補位置と前記類似位置とをそれぞれ他方の側から順に対応付けを行った結果、前記第 1 のテキストブロックに対応する前記類似位置に対応付けられた前記候補位置が示す第 2 の位置と、を求め、

前記第 1 の位置と前記第 2 の位置とのうち、所定の条件を満たす方の位置を、前記決定を行うための前記入力画像における位置として導出する

ことを特徴とする請求項 7 または 8 に記載の画像処理装置。

40

【請求項 10】

前記決定手段は、

50

前記一致度に基づき導出された前記入力画像における前記位置に、前記特定された登録文書における前記第 1 のテキストブロックを配置して、前記配置されたテキストブロックと重なり合う前記入力画像におけるテキストブロックが所定の条件を満たす場合、前記重なり合うテキストブロックを前記入力画像における前記処理対象の項目に対応するテキストブロックと決定する

ことを特徴とする請求項 6 から 9 のいずれか 1 項に記載の画像処理装置。

【請求項 1 1】

前記所定の条件は、

前記配置されたテキストブロックと前記重なり合うテキストブロックとの重なり度合いが所定値以上であり頂点の距離が一定の範囲内の場合である

10

ことを特徴とする請求項 1 0 に記載の画像処理装置。

【請求項 1 2】

前記特定された登録文書における前記第 1 のテキストブロックを基準とした所定の範囲が設定されており、

前記一致度は、

前記部分レイアウトを前記入力画像に重畳させた場合、前記入力画像における前記所定の範囲に含まれるテキストブロックのうち、前記部分レイアウトに含まれるテキストブロックと重ならないテキストブロックの面積が大きいほど、前記一致度が下がるように調整される

ことを特徴とする請求項 2 から 1 1 のいずれか 1 項に記載の画像処理装置。

20

【請求項 1 3】

前記部分レイアウトには、前記特定された登録文書における前記第 1 のテキストブロックと、前記特定された登録文書における前記第 1 のテキストブロックを基準とした所定の範囲に含まれる前記第 2 のテキストブロックと、が含まれる

ことを特徴とする請求項 1 から 1 2 のいずれか 1 項に記載の画像処理装置。

【請求項 1 4】

前記所定の範囲は、

前記特定された登録文書において、前記第 1 のテキストブロックを基準とした領域であって、テキストブロックが所定の数以上が含まれる領域に基づく領域である

ことを特徴とする請求項 1 2 または 1 3 に記載の画像処理装置。

30

【請求項 1 5】

前記特定手段は、

前記検出された前記入力画像におけるテキストブロックのレイアウトと前記複数の登録文書それぞれにおけるテキストブロックのレイアウトとの類似度に基づいて、前記入力画像に対応する登録文書を特定する

ことを特徴とする請求項 1 から 1 4 のいずれか 1 項に記載の画像処理装置。

【請求項 1 6】

前記特定手段が前記入力画像に対応する登録文書を特定できない場合、前記入力画像を新たな文書として登録する登録手段をさらに有する

ことを特徴とする請求項 1 から 1 5 のいずれか 1 項に記載の画像処理装置。

40

【請求項 1 7】

前記取得手段によって取得された前記入力画像における前記処理対象の項目に対応する文字列に基づき、前記入力画像のプロパティを設定する設定手段をさらに有する

ことを特徴とする請求項 1 から 1 6 のいずれか 1 項に記載の画像処理装置。

【請求項 1 8】

入力画像におけるテキストブロックを検出する検出ステップと、

複数の登録文書の中から、前記入力画像に対応する登録文書を特定する特定ステップと、前記特定された登録文書において規定されている、処理対象の項目に対応する第 1 のテキストブロックと前記第 1 のテキストブロックの近傍に存在する少なくとも 1 つの第 2 のテキストブロックとを含む部分レイアウトに基づき、前記入力画像における前記処理対象

50

の項目に対応するテキストブロックの決定をする決定ステップと、

前記決定されたテキストブロックに対して文字認識処理を行うことにより、前記入力画像における前記処理対象の項目に対応する文字列を取得する取得ステップと、

を有することを特徴とする画像処理方法。

【請求項 19】

コンピュータを、請求項 1 から 17 のいずれか 1 項に記載の画像処理装置の各手段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】

本開示は、画像に含まれるインデックスを抽出する技術に関する。

【背景技術】

【0002】

帳票等の紙文書を画像読み取り装置でスキャンすることにより得られたスキャン画像に含まれる所望の項目の文字列（以下、インデックスという）を抽出する方法がある。文書の内容からインデックスを抽出するには、OCR 処理が必要となる。しかし、スキャン画像全体に対して OCR 処理を実行すると処理負荷が増し、ユーザの待ち時間の増加することがある。

【0003】

特許文献 1 には、文書の種類ごとにインデックスが含まれる領域の情報を予め登録し、登録されているインデックスの領域に対して部分的に OCR 処理を行い、スキャン画像からインデックスを抽出する方法が開示されている。

【先行技術文献】

【特許文献】

【0004】

【文献】特開 2019 - 128715 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、同じ種類の文書であっても、記載される内容によってインデックスが含まれる文字列領域（以下、テキストブロックという）の位置がずれていることがある。このため、登録されているインデックスの領域に対して部分的に OCR 処理を行っても、インデックスの抽出に失敗してしまうことがある。

【0006】

本開示の技術は、スキャン画像のテキストブロックの位置が、登録されている位置とずれている場合であっても、抽出対象のインデックスを抽出することを目的とする。

【課題を解決するための手段】

【0007】

本開示の画像処理装置は、入力画像におけるテキストブロックを検出する検出手段と、複数の登録文書の中から、前記入力画像に対応する登録文書を特定する特定手段と、前記特定された登録文書において規定されている、処理対象の項目に対応する第 1 のテキストブロックと前記第 1 のテキストブロックの近傍に存在する少なくとも 1 つの第 2 のテキストブロックとを含む部分レイアウトに基づき、前記入力画像における前記処理対象の項目に対応するテキストブロックの決定をする決定手段と、前記決定されたテキストブロックに対して文字認識処理を行うことにより、前記入力画像における前記処理対象の項目に対応する文字列を取得する取得手段と、を有することを特徴とする。

【発明の効果】

【0008】

本開示の技術によれば、スキャン画像のテキストブロックの位置が登録されている文書と異なる場合であっても、抽出対象のインデックスを抽出することができる。

10

20

30

40

50

【図面の簡単な説明】

【0009】

【図1】システムの構成例を示す図である。

【図2】画像形成装置のハードウェア構成例を示す図である。

【図3】画像形成装置の機能構成を示す図である。

【図4】スキャン画像のファイル生成処理のフローチャートである。

【図5】インデックス抽出処理のフローチャートである。

【図6】ブロックセレクション処理の例を示す図である。

【図7】インデックス抽出ルールの例を示す図である。

【図8】インデックスブロック推定処理のフローチャートである。

10

【図9】ペアブロックの決定方法を説明する図である。

【図10】部分パターンの例を示す図である。

【図11】Y候補位置の決定処理を説明する図である。

【図12】Y方向のシフト量のヒストグラムの例を示す図である。

【図13】部分パターンの一致度の算出を説明する図である。

【図14】部分パターンの一致度の算出を説明する図である。

【図15】部分パターン範囲の決定方法を説明する図である。

【図16】インデックスブロック推定処理のフローチャートである。

【図17】部分パターンの例を示す図である。

【図18】XY候補位置群の例を示す図である。

20

【図19】類似位置群の例を示す図である。

【図20】類似位置群とXY候補位置群の対応付けを説明する図である。

【発明を実施するための形態】

【0010】

以下、添付図面を参照して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開示の技術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開示の技術の解決手段に必須のものとは限らない。

【0011】

<実施形態1>

本実施形態の画像形成装置は、文書原稿をスキャンして、得られたスキャン画像の先頭ページの画像に含まれる所定の項目の文字列を組み合わせることでファイル名を生成する。そして生成したファイル名をそのスキャン画像のファイル名としてユーザにレコメンドする。しかしながら、スキャン画像から所定の項目の文字列を抽出するには処理負荷が増すことがある。

30

【0012】

このため、文書の種類ごとに所定の項目のテキストブロックの位置情報を登録しておく。そしてスキャン画像の文書の種類を特定して、特定された文書における登録されたテキストブロックの位置に基づき、スキャン画像から所定の項目の文字列を抽出することが考えられる。しかしながらこの場合も、同じ文書の種類であっても、記載内容の変更等によりスキャンされた画像におけるテキストブロックの位置は登録されている位置と異なってしまうことがある。

40

【0013】

例えば、図11(a)の文書が登録されており、テキストブロック1003の位置を示す情報が発行元会社名を示す文字列が含まれる領域の情報として登録されているものとする。一方、図11(b)は、図11(a)と同じ種類の文書のスキャンして得られたスキャン画像であるが、表構造内の項目行数が増えており、抽出されるべき発行元会社名のテキストブロック1101が、図11(a)と比較して下方向にシフトしている。このため図11(b)のスキャン画像を得るためにスキャンされた文書が図11(a)と同じ種類であると特定できても、図11(b)の画像の発行元会社名を示す文字列の抽出に失敗することがある。なお、図11(c)の説明については後述する。

50

【 0 0 1 4 】

このため実施形態では、スキャン画像に含まれる項目のテキストブロックを抽出するために、スキャンされた文書原稿と同じ種類の文書における項目を示すテキストブロックと、それ以外の少なくとも1つのテキストブロックとのレイアウトを用いる。本実施形態では、そのレイアウトとの一致度が高い領域をスキャン画像から探索して、探索された結果に基づきスキャン画像に含まれる項目のテキストブロックを推定する方法を説明する。

【 0 0 1 5 】

なお、本実施形態では、画像内の座標は例えば、原点が左上で、縦方向がY方向、文字列が連続する横方向がX方向に延びる座標系が用いられる。テキストブロックの位置は、例えば、左上座標値が夫々の位置として保持される。

10

【 0 0 1 6 】

〔システム構成〕

図1は、本実施形態を適用可能なシステムの全体構成を示す図である。本実施形態のシステム105は、画像形成装置100および端末101を有する。図1に示すように、画像形成装置100はLAN102に接続され、Internet103等を介してPCなどの端末101等と通信可能になっている。なお、本実施形態においては、端末101は無くてもよく、画像形成装置100のみの構成だけでもよい。

【 0 0 1 7 】

画像形成装置100は、表示・操作部123（図2参照）、スキャナ部122（図2参照）及び、プリンタ部121（図2参照）等を有する複合機（MFP）である。画像形成装置100は、スキャナ部122を用いて文書原稿をスキャンするスキャン端末として利用することが可能である。また、タッチパネルやハードボタンなどの表示・操作部123を有し、ファイル名や格納先のレコメンド結果を表示したり、ユーザからの指示を受け付けたりするためのユーザインタフェースの表示を行う。

20

【 0 0 1 8 】

〔画像形成装置のハードウェア構成〕

図2は、画像形成装置100のハードウェア構成を示すブロック図である。本実施形態の画像形成装置100は、表示・操作部123、スキャナ部122、プリンタ部121、及び制御部110を有する。

【 0 0 1 9 】

制御部110は、CPU111、記憶装置112（ROM118、RAM119、HDD120）、プリンタI/F部113、ネットワークI/F部114、スキャナI/F部115、表示・操作I/F部116を有する。また、制御部110ではこの各部がシステムバス117を介して互いに通信可能に接続されている。制御部110は、画像形成装置100全体の動作を制御する。

30

【 0 0 2 0 】

CPU111は、記憶装置112に記憶された制御プログラムを読み出し実行することにより、後述のフローチャートにおける読取制御や画像処理、表示制御などの各処理を実行する手段として機能する。

【 0 0 2 1 】

記憶装置112は、制御プログラム、画像データ、メタデータ、設定データ及び、処理結果データ等を格納し保持する。記憶装置112には、不揮発性メモリであるROM118、揮発性メモリであるRAM119及び、大容量記憶領域であるHDD120などがある。ROM118は、制御プログラムなどを保持する不揮発性メモリであり、CPU111はその制御プログラムを読み出し制御を行う。RAM119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揮発性メモリである。

40

【 0 0 2 2 】

ネットワークI/F部114は、制御部110（画像形成装置100）を、システムバス117を介してLAN102に接続する。ネットワークI/F部114は、LAN102上の外部装置に画像データを送信したり、LAN102上の外部装置から各種情報を受

50

信したりする。

【 0 0 2 3 】

スキャナ I / F 部 1 1 5 は、スキャナ部 1 2 2 と制御部 1 1 0 とを、システムバス 1 1 7 を介して接続する。スキャナ部 1 2 2 は、文書原稿を読み取ってスキャン画像データを生成し、スキャナ I / F 部 1 1 5 を介してスキャン画像データを制御部 1 1 0 に入力する。なお、スキャナ部 1 2 2 は、原稿フィードを備え、トレイに置かれた複数の原稿を 1 枚ずつフィードして、連続的に読み取ることが可能とする。

【 0 0 2 4 】

表示・操作 I / F 部 1 1 6 は、表示・操作部 1 2 3 と制御部 1 1 0 とを、システムバス 1 1 7 を介して接続する。表示・操作部 1 2 3 には、タッチパネル機能を有する液晶表示部、ハードボタンなどが備えられている。

10

【 0 0 2 5 】

プリンタ I / F 部 1 1 3 は、プリンタ部 1 2 1 と制御部 1 1 0 とを、システムバス 1 1 7 を介して接続する。プリンタ部 1 2 1 は、CPU 1 1 1 で生成された画像データをプリンタ I / F 部 1 1 3 を介して受信し、当該受信した画像データを用いて記録紙へのプリント処理が行われる。以上のように、本実施形態に係る画像形成装置 1 0 0 では、上記のハードウェア構成によって、画像処理機能を提供することが可能である。

【 0 0 2 6 】

[画像形成装置の機能構成]

図 3 は、画像形成装置 1 0 0 の機能構成を示すブロック図である。なお、図 3 では画像形成装置 1 0 0 が有する諸機能のうち、文書原稿をスキャンして電子化（ファイル化）し、保存を行うまでの処理に関わる機能に絞った機能を示す。

20

【 0 0 2 7 】

表示制御部 3 0 1 は、表示・操作部 1 2 3 のタッチパネルに、各種のユーザ操作を受け付けるためのユーザインタフェース画面（UI 画面）を表示する。各種のユーザ操作には、例えば、スキャン設定、スキャンの開始指示、ファイル名設定、ファイルの保存指示などがある。

【 0 0 2 8 】

スキャン制御部 3 0 2 は、UI 画面でなされたユーザ操作（例えば「スキャン開始」ボタンの押下）に応じて、スキャン設定の情報と共にスキャン実行部 3 0 3 に対しスキャン処理の実行を指示する。スキャン実行部 3 0 3 は、スキャン制御部 3 0 2 からのスキャン処理の実行指示に従い、スキャナ I / F 部 1 1 5 を介してスキャナ部 1 2 2 に文書原稿の読み取り動作を実行させ、スキャン画像データを生成する。生成したスキャン画像データは、スキャン画像管理部 3 0 4 によって HDD 1 2 0 に保存される。

30

【 0 0 2 9 】

画像処理部 3 0 5 は、スキャン画像データに対して、テキストブロックの検出処理、OCR 処理（文字認識処理）、類似文書の判定処理といった画像解析処理の他、回転や傾き補正といった画像加工処理を行う。画像処理部 3 0 5 によって、画像形成装置 1 0 0 は画像処理装置としても機能する。スキャン画像から検出される文字列領域は「テキストブロック」とも呼ばれる。なお画像処理の詳細については後述する。

40

【 0 0 3 0 】

図 3 の各部の機能は、画像形成装置 1 0 0 の CPU が ROM に記憶されているプログラムコードを RAM に展開し実行することにより実現される。または、図 3 の各部の一部または全部の機能を ASIC や電子回路等のハードウェアで実現してもよい。

【 0 0 3 1 】

[スキャン画像のファイル生成処理のフローチャート]

画像形成装置 1 0 0 が文書原稿を読み取り、文書原稿の先頭ページのスキャン画像に対して画像処理を行い、スキャン画像に含まれる文字列を利用してファイル名を生成し、表示・操作部 1 2 3 を通じてユーザにレコメンドする処理の全体について説明する。

【 0 0 3 2 】

50

図4のフローチャートで示される一連の処理は、画像形成装置100のCPUがROMに記憶されているプログラムコードをRAMに展開し実行することにより行われる。また、図4におけるステップの一部または全部の機能をASICや電子回路等のハードウェアで実現してもよい。なお、各処理の説明における記号「S」は、当該フローチャートにおけるステップであることを意味し、以後のフローチャートにおいても同様とする。

【0033】

S400においてスキャン制御部302は、表示・操作部123を介してユーザのスキャン指示を受け付けると、スキャン実行部303に、スキャナ部122の原稿フィーダのトレイから複数の文書原稿を1枚ずつ読み取り（スキャン）を実行させる。そして、スキャン制御部302は、スキャンの結果得られた画像（スキャン画像とよぶ）の画像データを取得する。

10

【0034】

S401において画像処理部305は、S400で取得した画像データを解析し、スキャン画像に含まれるインデックスを抽出する処理（インデックス抽出処理）を行う。「インデックス」とは、文書のタイトル、管理ナンバー、会社名などの所定の項目の文字列である。本実施形態ではインデックスは、スキャン画像を保存する際のファイル名またはメタデータとして使用される。本ステップのインデックス抽出処理の詳細については、図5を用いて後述する。

【0035】

インデックスの使用方法はファイル名の生成またはメタデータの抽出に限られない。フォルダパスなどの他のプロパティ情報を設定するために用いられてもよい。つまり、ファイル名およびメタデータは、スキャン画像データに関するプロパティ（属性）として設定される情報の一種である。

20

【0036】

S402において表示制御部301は、S401で抽出されたインデックスを用いてファイル名を生成し、生成されたファイル名およびメタデータを、表示・操作部123に表示させてユーザに提示（レコメンド）する。また、表示制御部301は、ユーザによる確認または提示したファイル名の修正を受け付ける。表示制御部301は表示・操作部123を介してユーザから確認または修正を受け付けると、提示したファイル名または修正された場合は修正後のファイル名がスキャン画像のファイル名として決定される。ユーザが表示・操作部123を介して修正した場合は、インデックス抽出ルールが更新される。インデックス抽出ルールについては後述する。

30

【0037】

S403において画像処理部305は、S400で取得した画像データからファイルを作成し、S402で決定されたファイル名を設定する。本実施形態では、一例として、ファイル形式としてPDF（Portable Document Format）化してスキャン画像を保存するものとして説明する。PDFの場合には、画像データをページに分け保存することが可能であり、S400において複数の文書原稿をスキャンした場合には、各文書原稿に対応する画像データを別々のページとして1つのファイルに保存される。

【0038】

S404においてスキャン画像管理部304は、S403で作成したファイルを、LAN102を通じて所定の送信先に送信する。

40

【0039】

[インデックス抽出処理（S401）について]

図5は、S401のインデックス抽出処理の詳細を示すフローチャートである。インデックス抽出処理の詳細について図5を用いて説明する。インデックス抽出処理では、画像データの1ページに対して、向きの補正を行い、文書の種類を特定し、文書の種類に応じたインデックス抽出を行う処理を行う。

【0040】

S500において画像処理部305は、画像データからスキャン画像の傾きの角度を検

50

出し、検出した傾きだけ逆方向に画像を回転することでスキャン画像の傾きを補正する。傾き補正の対象となる傾きは、例えば、文書原稿のスキャン時にスキャナ部 1 2 2 の原稿フィーダ内のローラの摩耗などが原因でまっすぐに文書原稿が読み取られないことで発生する。または、スキャンされた文書原稿が印刷時にまっすぐに印刷されなかったために発生する。

【 0 0 4 1 】

傾きの角度の検出方法として、まず、画像データ内に含まれるオブジェクトを検出し、水平方向あるいは鉛直方向に隣り合うオブジェクト群を連結する。そして、連結されたオブジェクト群の中心位置を結んだ角度が、水平方向または鉛直方向からどれだけ傾いているかを導出して傾きを求める。なお、傾きの検出方法はこの方法に限られない。他にも例えば、画像データ内に含まれるオブジェクトの中心座標を取得し、0 . 1 度単位で中心座標群を回転させて、中心座標群が水平方向あるいは垂直方向に並ぶ割合がもっとも高い角度をスキャン画像の傾きとして求める方法でもよい。スキャン画像の傾きを補正することによって、以降に行われる、回転補正、ブロックセレクション処理、およびOCR処理のそれぞれの処理精度を上げることができる。

10

【 0 0 4 2 】

S 5 0 1 において画像処理部 3 0 5 は、S 5 0 0 の処理の結果得られた傾き補正後のスキャン画像に対して、画像内の文字が正立する向きになるように、9 0 度単位で画像を回転補正する。回転補正の方法は、例えば、傾き補正後のスキャン画像を基準画像として、基準画像と、基準画像を 9 0 度回転した画像と、基準画像を 1 8 0 度回転した画像と、基準画像を 2 7 0 度回転した画像と、の 4 枚の画像を用意する。そして、それぞれの画像に対し、高速処理可能な簡易的なOCR処理を実行して、一定値以上の確信度で認識された文字の数が最も多い画像を回転補正後の画像とする方法がある。ただし、回転補正の方法はこの方法に限るものではない。なお以降のスキャン画像とは、特に断りが無い限り S 5 0 0 および S 5 0 1 で補正されたスキャン画像のことを指すものとする。

20

【 0 0 4 3 】

S 5 0 2 において画像処理部 3 0 5 は、スキャン画像に対しブロックセレクション処理を実行する。ブロックセレクション処理とは、画像を前景領域と背景領域に分類した上で、前景領域をテキストブロックとそれ以外のブロックに分割して、テキストブロックを検出する処理である。

30

【 0 0 4 4 】

具体的には、白黒に二値化されたスキャン画像に対し輪郭線追跡を行って、黒画素輪郭で囲まれる画素の塊を抽出する。そして、面積が所定の大きさよりも大きい黒画素の塊については、内部にある白画素に対しても輪郭線追跡を行い白画素の塊を抽出し、さらに一定の大きさ以上の面積の白画素の塊の内部から再帰的に黒画素の塊を抽出する。こうして得られた黒画素の塊を前景領域と決定する。決定された前景領域は、大きさ及び形状で分類し異なる属性を持つ領域に分類する。例えば、縦横比が 1 に近く大きさが一定の範囲の前景領域を文字相当の画素塊とし、さらに近接する文字が整列良くグループ化され得る領域は文字列の領域 (T E X T) と決定する。扁平な画素塊は線領域 (L I N E) と決定する。一定大きさ以上でかつ矩形の白画素塊を整列よく内包する黒画素塊の占める範囲を表領域 (T A B L E) と決定する。不定形の画素塊が散在している領域を写真領域 (P H O T O) と決定する。そして、それ以外の形状の画素塊を図画領域 (P I C T U R E) と決定する。こうしてオブジェクトの属性毎に領域分割されたものの中から、文字属性を持つと決定された前景領域 (T E X T) がテキストブロックとして検出される。

40

【 0 0 4 5 】

図 6 は、ブロックセレクション処理の結果の一例を示す図である。図 6 (a) は回転補正後のスキャン画像を示す。図 6 (b) は図 6 (a) のスキャン画像に対するブロックセレクション処理の結果を示しており、点線で示した矩形が前景領域を表している。なお、図 6 (b) では、全ての前景領域の属性が決定されているが、属性については一部の前景領域に対してのみ表示している。本ステップで検出された各テキストブロックの情報 (属

50

性と各ブロックの位置およびサイズを示す情報)は、後続処理である、OCR処理および類似度計算等で用いられる。

【0046】

本ステップのブロックセレクション処理ではテキストブロックだけを検出する。その理由は、文字列の位置はスキャン画像の構造をよく表現し、インデックス情報と密接に関連するためである。したがって、写真領域や表領域等の他の属性を持つと判定されたブロックの情報を後続の処理で利用することを排除するものではない。

【0047】

S503において画像処理部305は、HDD120からインデックス抽出ルールを取得しRAM119に展開する。

【0048】

図7は、インデックス抽出ルール(以下単に、抽出ルールとよぶ)の一部を示す図である。図7は、抽出ルールに含まれる帳票IDとして「0001」が付与され登録されている抽出ルールのレコードを示している。抽出ルールでは、登録されている文書1つについて、「文書ID」と、「サムネイル」と、「文書識別情報」と、「インデックス情報」との各データが、レコード単位で対応付けられている。抽出ルールは登録済み文書の数だけこれらの組み合わせ(レコード)を保持する。文書IDは、文書の種類を表すユニークなIDである。

【0049】

文書識別情報は、登録されている文書のスキャン画像に対してブロックセレクション処理を実行した結果得られるテキストブロックの位置およびサイズの情報である。文書識別情報は、文書の種類を特定するための情報であり後述する文書マッチングで使用される。

【0050】

インデックス情報は、スキャン画像に含まれるインデックスを抽出するための情報である。インデックスは、ファイルに付与するファイル名またはメタデータを決定するために使用される。インデックス情報は、具体的には、登録されている文書内における、それぞれの項目の文字列(インデックス)が含まれるテキストブロックの座標およびサイズの情報が含まれる。図7の「インデックス情報」の画像701はそれぞれの項目が含まれるテキストブロックの位置およびサイズを画像上の座標に配置して図示したものである。また、インデックス情報にはファイル名を生成するために用いられるインデックスとその順番を示す情報、メタデータとして付与するための情報が含まれる。

【0051】

インデックス情報の「ファイル名ルール」には、タイトル(title)、発行元会社名(sender)、帳票番号(number)の項目のインデックスを、セパレータであるアンダースコアでつなげてファイル名を生成することが示されている。また、「メタデータ」には合計金額(total_price)の項目のインデックスをメタデータとして利用することが示されている。つまり、所定の項目のインデックスを抽出することで、ユーザにレコメンドするファイル名の生成、およびメタデータの抽出をすることができる。

【0052】

なお、本実施形態では、抽出されたインデックスをファイル名またはメタデータとして利用する例を示しているが、他のプロパティ情報であるファイルの送信先のフォルダ情報を決定するためのルールを保持してもよい。その場合も、インデックスを用いて生成されたプロパティ情報がS402でユーザにレコメンドされて、S403でプロパティ情報がスキャン画像のファイルに設定される。

【0053】

また、登録されている文書の抽出ルールとして、図7の「サムネイル」に示したように、登録された文書に対応するスキャン画像のサムネイルと一緒に保持してもよい。

【0054】

S504において画像処理部305は、スキャン画像に対して文書マッチングを実行する。文書マッチングでは、スキャン画像を得るためにスキャンされた文書(入力文書)と

10

20

30

40

50

同じ種類の文書が、抽出ルールに登録されている文書群にあるかどうかを判定する。そして、入力文書と同じ種類の文書が登録されていると判定された場合、その種類を特定する処理である。

【 0 0 5 5 】

本実施形態では、まず、スキャン画像と、抽出ルールに登録されている夫々の文書と、を1対1で比較し、含まれるテキストブロックの形状および配置がどれだけ類似しているかを表す類似度の算出を行う。類似度の算出の方法として、例えば、スキャン画像のテキストブロック全体と、登録されている文書のテキストブロック全体で位置合わせを行う。そして、スキャン画像の各テキストブロックと登録されている文書の各テキストブロックとが重なる面積の総和の二乗（値Aとする）を求める。さらにスキャン画像のテキストブロックの面積の総和と登録されている文書のテキストブロックの面積の総和との積（値Bとする）を求める。そして、値Aを値Bで割った値を類似度とする方法がある。この類似度の算出を、スキャン画像と抽出ルールに登録されている全ての文書との間で行う。

10

【 0 0 5 6 】

そして、所定値以上の類似度であり、かつ、最も類似度が高い、抽出ルールに登録されている文書が、スキャンされた入力文書と同じ種類の文書と特定される。また、抽出ルールに、類似度が所定値以上の文書が無かった場合は、入力文書と同じ種類の文書は、抽出ルールには登録されていないと判定される。

【 0 0 5 7 】

S 5 0 5 において画像処理部 3 0 5 は、S 5 0 4 で実行した文書マッチングの結果、入力文書と同じ種類の文書が抽出ルールに登録されていたかを判定する。入力文書が登録済み文書でなかった場合（S 5 0 5 が N O）、本フローチャートの処理を終了する。登録済み文書でなかった場合は、前述したように新たにIDが付されて、S 5 0 2 で検出したテキストブロックのレイアウト情報等が抽出ルールに登録される。この場合、S 4 0 2 ではファイル名およびメタデータのユーザにレコメンドはされずに、表示制御部 3 0 1 は、ユーザによるファイル名の入力を受け付ける。表示制御部 3 0 1 は表示・操作部 1 2 3 を介してユーザから入力を受け付けると、入力されたファイル名がスキャン画像のファイル名として決定される。

20

【 0 0 5 8 】

入力文書と同じ種類の文書が登録されている場合（S 5 0 5 が Y E S）、S 5 0 6 において画像処理部 3 0 5 は、S 5 0 4 で入力文書と同じ種類と特定された抽出ルールの文書と同じ文書IDを、スキャン画像に付与する。

30

【 0 0 5 9 】

S 5 0 7 において画像処理部 3 0 5 は、S 5 0 6 で付与された文書IDに紐づいた抽出ルールに基づいて、スキャン画像内における抽出対象（処理対象）の項目のインデックスのテキストブロックを推定するインデックスブロック推定処理を実行する。タイトル、発行元会社名、帳票番号等の項目を示す文字列（インデックス）が含まれるテキストブロックをインデックスブロックと呼ぶことがある。インデックスブロック推定処理の詳細については、後述する。

【 0 0 6 0 】

S 5 0 8 において画像処理部 3 0 5 は、S 5 0 7 で推定された夫々の項目のインデックスブロック群に対して、部分的なOCRを実行し、各インデックスブロックに対応する文字列をインデックスとして抽出する。

40

【 0 0 6 1 】

[インデックスブロック推定処理（S 5 0 7）について]

図8は、S 5 0 7 のインデックスブロック推定処理のフローチャートである。インデックスブロック推定処理の詳細について図8を用いて説明する。なお、以下、登録文書とは、S 5 0 3 で取得した抽出ルールにおいて登録されている文書のうち、S 5 0 6 でスキャン画像に付与された文書IDに対応する文書のことをいう。本フローチャートの説明では、登録文書は図7の文書ID「0 0 0 1」の文書であるものとして説明する。

50

【 0 0 6 2 】

S 8 0 0 において画像処理部 3 0 5 は、抽出ルールから、S 5 0 6 で付与された文書 I D に紐づいた文書識別情報を取得する。そして、画像処理部 3 0 5 は、スキャン画像内の全体のテキストブロックと、登録文書の全体のテキストブロックとで全体の位置合わせを行う。

【 0 0 6 3 】

S 4 0 0 で取得されたスキャン画像の入力文書は、登録文書と同じ種類の文書であり、夫々の項目は登録文書の項目と同じ座標に印刷される。しかし、印刷およびスキャンのタイミングまたは印刷時の機器による違い等により、スキャン画像上のテキストブロックの位置と登録文書のテキストブロックの位置とにズレが生じてしまうことがある。そこで、
10
本ステップではそのズレの影響を軽減して以降の処理の精度を向上させるため、全体の位置合わせを行う。なお、本実施形態では、図 5 の S 5 0 0 で傾き補正を行っているため、本ステップの全体の位置合わせでは、スキャン画像上のテキストブロック全体をシフト（平行移動）する補正のみを行う例について説明する。

【 0 0 6 4 】

全体の位置合わせでは、登録文書のテキストブロックに対してどれだけスキャン画像のテキストブロックがシフトしているかというシフト量を算出して、シフト量だけスキャン画像の各テキストブロックがシフトするように座標の修正を行う。

【 0 0 6 5 】

図 9 は、スキャン画像のテキストブロックと登録文書のテキストブロックとを同じ座標系に描画した画像の一部分を切り出した図である。図 9 を用いて全体の位置合わせのためのシフト量の算出の具体的な手順を説明する。図 9 において、実線の矩形はスキャン画像内のテキストブロック群のうちから選択された 1 つのテキストブロック 9 0 0 を示し、破線の矩形は、テキストブロック 9 0 0 の周囲にある登録文書のテキストブロック 9 0 1 ~ 9 0 3 を示している。また、図 9 において、一点鎖線の円 9 0 4 は、スキャン画像のテキストブロック 9 0 0 の左上頂点を中心に一定距離を半径とした範囲を示している。
20

【 0 0 6 6 】

シフト量の算出のために、スキャン画像の各テキストブロックと対応する候補となる登録文書のテキストブロック（ペアブロックとよぶ）を決定する。ここでスキャン画像のテキストブロックのペアブロックの決定について説明する。
30

【 0 0 6 7 】

初めに、登録文書のテキストブロック 9 0 1 ~ 9 0 3 のうち、スキャン画像内のテキストブロック群から選択された 1 つのテキストブロック 9 0 0 の左上頂点を中心とする円 9 0 4 の中に、左上頂点が入るテキストブロックを探す。図 9 では、テキストブロック 9 0 1、9 0 2 が該当することになる。次に、スキャン画像のテキストブロック 9 0 0 と、登録文書のテキストブロック 9 0 1、9 0 2 それぞれとのオーバーラップ率を求める。オーバーラップ率は、スキャン画像のテキストブロックと登録画像のテキストブロックとの左上頂点同士を合わせて、両テキストブロックの共通部分の面積を算出する。そして、（共通部分の面積） / （両テキストブロックのうち大きい方の面積）によって値を求めてオーバーラップ率とする。
40

【 0 0 6 8 】

オーバーラップ率が、所定の条件を満たす登録文書のテキストブロックを、ペアブロックとする。所定の条件は、例えば、スキャン画像のテキストブロックとのオーバーラップ率が、最大オーバーラップ率に係数 を乗算した値以上であり、かつ、所定の閾値以上であることである。この場合において、係数 は最大オーバーラップ率と近いオーバーラップ率を持つ組合せを選択するためのもので、例えば 0 . 5 ~ 0 . 8 のような 1 . 0 未満の値とする。また、所定の閾値は最低ラインを規定するものであり、例えば 0 . 3 ~ 0 . 7 のような 1 . 0 未満の値とする。

【 0 0 6 9 】

図 9 では、登録文書のテキストブロック 9 0 1、9 0 2 のうち、スキャン画像のテキス
50

トブロック 900 と形状の近い、テキストブロック 901 のみがペアブロックとして選択される。所定の条件を満たすテキストブロックが他にもあればペアブロックは複数選択されることもある。このように、スキャン画像内から選択された 1 つのテキストブロックに対応するペアブロック群のそれぞれに対して、スキャン画像内から選択されたテキストブロックとの左上頂点の X 方向および Y 方向の差分量（シフト量）を算出する。そして、差分量をシフト量ヒストグラムに投票する。この場合のヒストグラムのピンの範囲は任意でよい。

【0070】

図 9 の場合、テキストブロック 900 については、登録文書のテキストブロック 901 との左上頂点の X 方向および Y 方向の差分量（シフト量）が算出されて、シフト量がシフト量ヒストグラムに投票される。

10

【0071】

スキャン画像内のテキストブロックに対応するペアブロック群を決定し、シフト量ヒストグラムに投票するまでの処理を、スキャン画像の全てテキストブロックに対してそれぞれ行う。そして、最終的に得られたシフト量ヒストグラムにおける最大のピーク点となる位置を決定する。決定された位置が示すシフト量を全体の位置合わせのシフト量とする。

【0072】

なお、ノイズの影響が懸念される場合は、生成したシフト量ヒストグラムに対してスムージングを掛けてもよい。また、最大となるピーク点以外の局所的なピーク点についても、シフト量の候補として選び、その候補の中から全体の位置合わせに用いるシフト量を選んでよい。例えば、シフト量の各候補について、スキャン画像のテキストブロックの座標をシフトさせて、図 5 の S504 の文書マッチングと同様の類似度算出を行い、最も類似度が高くなる候補を、最終的なシフト量として決定してもよい。

20

【0073】

上記の手順で決定されたシフト量だけ、スキャン画像の各テキストブロックの座標をシフトすることで、位置合わせされたスキャン画像のテキストブロック群を得ることができる。なお、テキストブロックの位置合わせの方法は上記の方法に限るものではない。スキャン画像全体のシフト（平行移動）に関する補正のみを行う例について説明したが、印刷およびスキャンのズレとして、倍率に関するズレが想定される場合には、シフト量だけでなく、倍率のズレも考慮した位置合わせを行ってもよい。

30

【0074】

なお以下のステップにおけるスキャン画像またはスキャン画像のテキストブロック群は、この全体の位置合わせされたスキャン画像またはテキストブロック群を指すものとする。

【0075】

次に、S506 で付与された文書 ID に紐づいた登録文書のインデックス情報を取得する。そして S801 でインデックス情報に含まれるインデックスの項目のいずれかを処理対象に選んで S801 ~ S810 を繰り返す。そして、スキャン画像のテキストブロック群から、処理対象の項目のテキストブロックを推定する処理を行う。処理対象の項目に対する処理が終了すると、再度、未処理の項目の中から処理対象の項目が選択される。

【0076】

40

S801 において画像処理部 305 は、登録文書のインデックス情報に登録されている項目のうち未処理のインデックスの項目を 1 つ選択して処理対象の項目とする。本実施形態では、図 7 のインデックス情報に保持されている、タイトル (title)、発行元会社名 (sender)、帳票番号 (number)、合計金額 (total_price) の項目の何れかが処理対象として選択される。

【0077】

S802 において画像処理部 305 は、処理対象の項目の「部分パターン」を取得する。部分パターンには、登録文書に含まれるテキストブロックの一部のレイアウト（部分レイアウト）の情報と、部分レイアウトを含む範囲（部分パターン範囲）の情報と、が含まれる。

50

【 0 0 7 8 】

図 1 0 (a) は、図 7 で文書 I D 「 0 0 0 1 」として登録されている登録文書における、それぞれの項目のインデックスブロックの位置およびサイズを図示したものである。図 1 0 (a) の破線の矩形は、タイトル、帳票番号、合計金額、発行元会社名のそれぞれの項目のインデックスブロック 1 0 0 0 ~ 1 0 0 3 を表している。

【 0 0 7 9 】

図 1 0 (b) は、「発行元会社名 (sender)」の項目の部分パターンを示す図である。図 1 0 (b) の一点鎖線の矩形で表される範囲は、「発行元会社名 (sender)」の項目の部分パターン範囲 1 0 0 6 を示す。部分パターン範囲 1 0 0 6 は、「発行元会社名 (sender)」の項目のテキストブロックであるインデックスブロック 1 0 0 3 を基準として予め設定された値を使って決定される。

10

【 0 0 8 0 】

テキストブロック 1 0 0 4、1 0 0 5 は、登録文書における、部分パターン範囲 1 0 0 6 に少なくとも一部が含まれるテキストブロックを表している。このテキストブロック 1 0 0 4、1 0 0 5 と、インデックスブロック 1 0 0 3 で表される登録文書内の部分的なレイアウトが、発行元会社名の項目の部分レイアウトである。部分レイアウトは、処理対象の項目のテキストブロックと、処理対象の項目のテキストブロック以外の少なくとも 1 つのテキストブロックとで表される。レイアウトとは、夫々のテキストブロックの位置情報と、夫々のテキストブロックのサイズと、を表す情報である。

【 0 0 8 1 】

発行元会社名の項目の部分パターンに含まれる情報として、部分パターン範囲 1 0 0 6 と、インデックスブロック 1 0 0 3 とテキストブロック 1 0 0 4 および 1 0 0 5 とからなる部分レイアウトと、が決定される。このように、登録文書の夫々の項目に対応する部分パターンが決定されて記憶されている。

20

【 0 0 8 2 】

詳細は後述するが、本実施形態では、部分レイアウトと配置が類似または一致しているスキャン画像内の位置を探索して、スキャン画像内における処理対象の項目のテキストブロックを推定する。

【 0 0 8 3 】

図 1 0 (c) は、「タイトル (title)」の項目の部分パターンを示す図である。タイトルについても同様に、部分パターン範囲 1 0 0 7 と、タイトルのインデックスブロック 1 0 0 0 と部分パターン範囲 1 0 0 7 に含まれるテキストブロック 1 0 0 1、1 0 0 8 ~ 1 0 1 3 とからなる部分レイアウトと、が部分パターンとして決定されている。

30

【 0 0 8 4 】

なお、部分パターン範囲 1 0 0 7 のサイズは、図 1 0 (b) の部分パターン範囲 1 0 0 6 と比べてサイズが異なる。このように項目の性質に応じて部分パターンサイズは異ならせてもよい。または、部分パターン範囲のサイズは、全ての項目で共通のサイズが用いられてもよい。部分パターン範囲のサイズの決定方法については実施形態 2 で説明する。

【 0 0 8 5 】

なお、部分パターンは、文書原稿をスキャンした後に行われるインデックス抽出処理の実行が行われるごとに決定される必要はない。例えば、文書の登録時において、項目ごとに部分パターンを決定し、図 7 で示した抽出ルールの一部として予め記憶させてもよい。つまり、S 8 0 2 では、記憶されている処理対象の項目の部分パターンが取得されればよい。

40

【 0 0 8 6 】

次の S 8 0 3 および S 8 0 4 では、処理対象の項目の部分レイアウトとの一致度が高い領域のある、スキャン画像内の位置 (X Y 候補位置) を決定する。X Y 候補位置の決定方法としては、例えば、テンプレートマッチングのようにスキャン画像内の探索範囲に対して部分パターンを走査して一致度を算出することで候補位置を推定してもよい。本実施形態では計算量を抑制させるため、探索範囲における Y 方向の候補となる位置を決定して Y

50

方向の位置（Y位置）を絞り込む。その上で、Y位置の候補（Y候補位置）群それぞれにおいて、X方向に部分パターンを走査してXY候補位置を決定することで、計算量を抑える方法を説明する。

【0087】

S803において画像処理部305は、スキャン画像のテキストブロック群から、登録文書における処理対象の項目の部分パターンのテキストブロックに類似するY候補位置群を決定する。

【0088】

図11は、Y候補位置群の決定処理を説明するための図である。処理対象の項目が発行元会社名（sender）であるものとして説明を行う。

【0089】

図11(a)は、登録文書における発行元会社名（sender）の部分パターンを示す図であり図10(b)と同様の図である。図11(b)は、スキャン画像であり破線の矩形は、位置合わせがされたテキストブロック群を表している。また、図11(b)で示したスキャン画像が示す文書は、登録文書「0001」と同じ種類の文書として判定された文書であるが、図7の登録文書に比べて表構造内の項目行数が増えている例を示している。よって、スキャン画像における推定されるべき発行元会社名（sender）のインデックスブロック1101が、登録文書における発行元会社名（sender）のインデックスブロック1002の位置と比較して下方向にシフトしてしまっている。

【0090】

図11(c)は、発行元会社名の部分パターンに含まれる部分レイアウトを表すテキストブロック1003～1005のうちの1つのテキストブロック1003を、スキャン画像のテキストブロック群と同じ座標系に重畳させた図である。Y候補位置群の決定について、部分パターン内のテキストブロック1003に注目して図11(c)を用いて説明する。

【0091】

図11(c)の、一点鎖線の矩形で表される探索範囲1100は、処理対象の項目のY候補位置群を決定するために探索する範囲を表している。破線の矩形で表されるテキストブロック1101～1109は、図11(b)に示すスキャン画像のテキストブロックのうち、矩形の中心が探索範囲1100の中にあるテキストブロックである。

【0092】

Y候補位置群の決定には、はじめに、部分レイアウトに含まれる1つのテキストブロック（図11(c)ではテキストブロック1003）が選択される。そして選択されたテキストブロックをスキャン画像のテキストブロック群と同じ座標系に重畳し、探索範囲内のスキャン画像のテキストブロック（図11(c)ではテキストブロック1101～1109）との矩形の中心のY位置の差分量をそれぞれ算出する。そして、算出された差分量がY方向のシフト量ヒストグラムに投票される。シフト量ヒストグラムのビンの範囲は任意でよい。

【0093】

図12は、Y方向のシフト量ヒストグラムの例を示す図である。図12(a)は、図11(c)における部分パターンのテキストブロック1003と、スキャン画像のテキストブロック1102とのY位置の差分量を投票した後のシフト量ヒストグラムである。hは基準からのY方向の探索範囲の絶対値の上限を示している。テキストブロック1003とテキストブロック1102とのY方向の差分量に従い、位置1200に投票が行われている。同様に、部分パターンに含まれる1つのテキストブロックと、スキャン画像の探索範囲内の全てのテキストブロックとのY中心の差分量に応じた投票が行われる。この投票を、部分パターン内の全テキストブロックに対して行う。つまり、部分パターンのテキストブロック1004、1005についても、探索範囲内のテキストブロック1101～1109とのY中心の差分量が算出されてシフト量ヒストグラムに投票される。そして、Y方向のシフト量ヒストグラムを完成させる。なお、ノイズの影響が懸念される場合は、Y方

10

20

30

40

50

向の生成したシフト量ヒストグラムに対してスムージングを掛けてもよい。

【0094】

図12(b)は最終的に生成されるY方向のシフト量ヒストグラムである。シフト量ヒストグラムの生成が完了した後、ヒストグラム内の位置1201~1206に示すようなピーク点を決定し、各ピーク点のピンに応じたY方向のシフト量に基づきY候補位置群を決定する。

【0095】

なお、図11(c)のY候補位置群を決定するための探索範囲1100は、部分パターンのインデックスブロックの位置を基準に、あらかじめ設定された値で自動決定される。なお、探索範囲のサイズについては、全ての項目で共通の範囲を使用してもよいし、処理対象の項目の属性に応じて決定してもよい。例えば、タイトルのインデックスブロックは文書内で固定の位置にあることが多い。よって、処理対象の項目がタイトルの場合、探索範囲を狭くしても探索範囲から推定されるべきインデックスブロックが外れる可能性は低いため、探索範囲を狭く設定してもよい。探索範囲を狭くすることで、計算量を抑えつつ、余計な候補位置が決定されることを防ぐことができる。一方、項目が合計金額のインデックスブロックは、文書内の表構造の項目行数の変化に応じて、位置が上下に変化することがある。このため、処理対象の項目が合計金額の場合は他の項目よりも探索範囲を上下に広く設定してもよい。

10

【0096】

S804において画像処理部305は、S803で決定された夫々のY候補位置を基準に、部分パターンの部分レイアウトとスキャン画像のテキストブロック群との一致度を導出する。

20

【0097】

図13は、スキャン画像内のある位置に処理対象の項目の部分レイアウトを重ねて置いた場合の、部分レイアウトとスキャン画像のテキストブロックのレイアウトとの重なり状態を示した図である。図13を用いて、部分レイアウトとスキャン画像のテキストブロック群の一致度の導出方法について説明する。

【0098】

図13において、実線の矩形は、処理対象の項目の部分レイアウトを構成するテキストブロック1003~1005である。一点鎖線の矩形は、部分パターン範囲1006を表している。破線の矩形は、スキャン画像のテキストブロック1101、1104~1106、1109を表す。斜線塗りつぶし領域1309、1310は、部分レイアウトのテキストブロック1003~1005とスキャン画像のテキストブロックの重なっている領域を表している。

30

【0099】

部分レイアウトとスキャン画像のテキストブロックとの一致度Scoreは、以下の式(1)で導出する。

【0100】

【数1】

$$Score = \frac{1}{N_R} \sum_{r \in R} Correlation(r) \quad \dots \text{式(1)}$$

40

【0101】

上記式(1)において、Rは部分レイアウトを構成する全テキストブロックを表しており、またNRは部分レイアウトを構成するテキストブロックの総数を表す。図13において、Rは、テキストブロック1003~1005であり、NRは3である。

【0102】

Correlation(r)は、部分レイアウトを構成する一つのテキストブロックrの個別一致度である。テキストブロックrの個別一致度Correlation(r)は、式(2)によって導出する。

50

【 0 1 0 3 】

【 数 2 】

$$Correlation(r) = \begin{cases} \frac{\sum_{q \in \text{Overlapping}Q} \left(\frac{\text{OverlapArea}(r,q)}{\text{Area}_r} \right) \times \left(\frac{\text{OverlapArea}(r,q)}{\text{Area}_q} \right)}{N_{\text{Overlapping}Q}}, & N_{\text{Overlapping}Q} > 0 \\ 0, & N_{\text{Overlapping}Q} = 0 \end{cases} \quad \dots \text{式 (2)}$$

【 0 1 0 4 】

OverlappingQは、テキストブロック r と重なりのあるスキャン画像のテキストブロックの集合である。OverlapArea(r,q)は、テキストブロック r とOverlappingQのテキストブロックうちの1つのテキストブロック q との重なり領域の面積である。また N_{OverlappingQ}はOverlappingQの総数を表す。

10

【 0 1 0 5 】

図 1 3 において、r をテキストブロック 1 0 0 3 とした場合、OverlappingQ はテキストブロック 1 1 0 5 のみでありOverlapArea(r,q)は領域 1 3 0 9 である。r をテキストブロック 1 0 0 5 とした場合、OverlappingQ は、テキストブロック 1 1 0 4 のみでありOverlapArea(r,q)は領域 1 3 1 0 が該当する。r をテキストブロック 1 0 0 4 とした場合、該当するOverlappingQは無いためN_{OverlappingQ}は0であることから、Correlation(r)は0である。

20

【 0 1 0 6 】

Area_rはテキストブロック r の面積であり、Area_qはテキストブロック q の面積である。

【 0 1 0 7 】

なお、式 (1) による一致度の導出では、スキャン画像のテキストブロックの数が多く、またテキストブロックの面積が大きいほど、個別一致度Collrelation(r)の値は大きく導出されてしまうことがある。そこで、一致度Scoreは、以下の式 (1) ' に示すようにペナルティ項PenaltyTermを追加してもよい。

【 0 1 0 8 】

【 数 3 】

$$Score = PenaltyTerm \times \frac{1}{N_R} \sum_{r \in R} Correlation(r) \quad \dots \text{式 (1) '}$$

30

【 0 1 0 9 】

式 (1) ' におけるペナルティ項PenaltyTermは、式 (3) によって導出する。

【 0 1 1 0 】

【 数 4 】

$$PenaltyTerm = \frac{TotalArea_R}{TotalArea_R + TotalArea_NonOverlappingQ} \quad \dots \text{式 (3)}$$

40

【 0 1 1 1 】

TotalArea_Rは、部分レイアウトを構成する全テキストブロックの総面積である。図 1 3 ではテキストブロック 1 0 0 3 ~ 1 0 0 5 の総面積である。

【 0 1 1 2 】

TotalArea_{NonOverlappingQ}は、部分パターン範囲内に存在するスキャン画像のテキストブロックのうち、部分レイアウトを構成するテキストブロックの何れとも重ならないテキストブロック群の面積の総和である。図 1 3 の場合、部分パターン範囲 1 0 0 6 内のテキストブロック 1 1 0 1、1 1 0 4、1 1 0 5、1 1 0 6、1 1 0 9 のうちテキストブロック 1 0 0 3 ~ 1 0 0 5 と重ならないテキストブロック 1 1 0 1、1 1 0 6、1 1 0 9 の面積の総和である。

50

【 0 1 1 3 】

ペナルティ項を設けることによって、部分パターン範囲 1 0 0 6 内の部分レイアウトを構成するテキストブロックが存在しなかった範囲に、スキャン画像内のテキストブロックが存在する場合に一致度を減点するように調整することができる。よって、部分レイアウトを構成するテキストブロックが少ない場合であっても、部分パターン範囲内の部分レイアウトを構成するテキストブロックが存在しない領域の情報を活用して一致度を導出することができる。なお、一致度の導出方法は、上記の式による導出に限るものではなく、部分レイアウトとの一致度が決定できればよい。

【 0 1 1 4 】

S 8 0 4 において画像処理部 3 0 5 は、S 8 0 3 で決定した Y 候補位置群のうちのいずれかの Y 候補位置に、インデックスブロックが位置するように部分パターン（部分レイアウトおよび部分パターン範囲）を置く。そして、画像処理部 3 0 5 は、部分パターンを X 方向に走査して、各位置における一致度を導出する。画像処理部 3 0 5 は、これを全ての Y 候補位置群に対して行う。

10

【 0 1 1 5 】

図 1 4 は、S 8 0 3 で決定した Y 候補位置群のうちの一つの Y 候補位置における本ステップの処理を表した図である。図 1 4 (a) において、実線の矩形は、部分レイアウトを構成するテキストブロック 1 0 0 3 ~ 1 0 0 5 であり、一点鎖線の矩形は部分パターン範囲 1 0 0 6 を表している。また破線の矩形は、スキャン画像のテキストブロック 1 1 0 1 、 1 1 0 5 、 1 1 0 6 を表し、斜線の領域は、部分レイアウトのテキストブロックとスキャン画像のテキストブロックとの重なっている領域を表している。また、図 1 4 では、本ステップにおける処理が図 1 4 (a) ~ (e) から順に処理が進むように示されており、探索範囲内で部分パターンを X 方向に（左から右に）走査しながら、それぞれの位置における一致度を導出する様子を示している。同様の処理が夫々の Y 候補位置において行われる。

20

【 0 1 1 6 】

S 8 0 5 において画像処理部 3 0 5 は、S 8 0 4 で導出した一致度が最大となる位置を X Y 候補位置と決定する。例えば、図 1 4 の場合、部分パターン（部分レイアウト）が、図 1 4 (c) に示す位置で一致度が最大となる。このため、図 1 4 (c) における部分レイアウトに含まれるインデックスブロックを示すテキストブロック 1 0 0 3 の位置が、X Y 候補位置として決定される。

30

【 0 1 1 7 】

S 8 0 6 において画像処理部 3 0 5 は、S 8 0 5 で決定した X Y 候補位置における一致度が所定の閾値以上かどうかを判定する。

【 0 1 1 8 】

一致度が閾値以上の場合（S 8 0 6 が Y E S ）、S 8 0 7 において画像処理部 3 0 5 は、S 8 0 5 で決定したスキャン画像上の X Y 候補位置を処理対象の項目のテキストブロック（インデックスブロック）のある位置と推定する。画像処理部 3 0 5 は、推定した位置に基づき、スキャン画像内の処理対象の項目のインデックスブロックを推定する処理を行う。

40

【 0 1 1 9 】

例えば、登録文書における処理対象の項目のインデックスブロックをスキャン画像内の X Y 候補位置にシフトさせた場合に、重なり合うスキャン画像内のテキストブロックが、所定の条件を満たすかが判定される。所定の条件とは、例えば、登録文書における処理対象のインデックスブロックとの重なり度合いを示す重なり率が所定の値以上、かつ、登録文書における処理対象のインデックスブロックとの左上座標の距離が一定の範囲内に入っているかという条件である。

【 0 1 2 0 】

所定の条件を満たすテキストブロックがあると判定された場合（S 8 0 7 が Y E S ）、S 8 0 8 に進む。S 8 0 8 において画像処理部 3 0 5 は、S 8 0 7 で所定の条件を満たす

50

と判定されたスキャン画像のテキストブロックを、S 8 0 1 で選択した処理対象の項目を示す文字列を含むテキストブロック（インデックスブロック）と推定する。

【 0 1 2 1 】

一致度が閾値未満の場合（S 8 0 6 が NO）または該当のテキストブロックがないと判定された場合（S 8 0 7 が NO）、S 8 0 9 に進む。S 8 0 9 において画像処理部 3 0 5 は、S 8 0 1 で選択した処理対象の項目に対応するテキストブロックはスキャン画像内には無いと決定する。例えば、スキャン画像において処理対象の項目に対応する文字列が所定の領域に記載されていない場合、あるいは、S 8 0 4 で誤って位置を推定してしまった場合、S 8 0 9 において決定が行われる。

【 0 1 2 2 】

S 8 1 0 において画像処理部 3 0 5 は、登録文書のインデックス情報に登録されている全ての項目について、インデックスブロックを推定する処理を完了したかを判定する。未処理の項目があれば S 8 0 1 に戻る。

【 0 1 2 3 】

全ての項目について処理が完了していれば本フローチャートの処理を終え S 5 0 8 に進む。S 5 0 8 において画像処理部 3 0 5 は、推定された夫々の項目のインデックスブロックに OCR 処理を実行し、それぞれの項目に対応する文字列をインデックスとして抽出する。

【 0 1 2 4 】

以上説明したように本実施形態では、テキストブロックのレイアウトの一部を利用してスキャン画像に含まれるインデックスの抽出を行う。このため、本実施形態によれば、入力文書おける記載内容の変化等によって、スキャン画像に含まれるインデックスブロックの位置が登録文書と異なる場合であっても、インデックスを抽出することができる。また、本実施形態では、文書マッチングによって入力文書の種類を特定して、文書の種類に紐づいた抽出ルールを利用する。このため、テキストブロックの部分的なレイアウトによるインデックスブロックを推定する処理であっても、インデックスの誤抽出を抑制することができる。また、文書マッチングおよびインデックスブロック推定処理では、OCR 処理の前処理の結果として得られる前景領域のうちテキストブロックのみを使用する。このため、余計な計算コストをかけることなく、インデックス抽出処理を行うことができる。

【 0 1 2 5 】

< 実施形態 2 >

実施形態 1 では、部分パターン範囲は、予め設定された値に基づき決定する方法について説明した。しかしながら、部分パターン範囲を広く設定しすぎると、インデックスブロックの周囲のみレイアウトが変わっているような場合、適切にインデックスブロックの位置を推定することができない。一方、部分パターン範囲が小さくなると部分レイアウトを構成するテキストブロックの数が少なく決定されることがあり、スキャン画像内の一致度の高い領域を探索するのが難しくなる。このため本実施形態では、部分パターン範囲を適切なサイズに決定する方法を説明する。なお、本実施形態については、実施形態 1 からの差分を中心に説明する。特に明記しない部分については実施形態 1 と同じ構成および処理である。

【 0 1 2 6 】

文書の種類に応じてインデックスブロックの周囲に存在するテキストブロックの数、レイアウトは変わる。このため、本実施形態では、部分パターン範囲のサイズを決定するために、段階的に対象の項目のインデックスブロックを含む領域を広げながら、その領域と重なるテキストブロックの数をカウントする。そして重なるテキストブロックの数が一定数以上になったときの領域を、その項目の部分パターン範囲として決定する。

【 0 1 2 7 】

図 1 5 は、本実施形態における部分パターン範囲の決定方法を説明するための図である。図 1 5 (a) における、実線の矩形はタイトルのインデックスブロック 1 0 0 0 であり、一点鎖線の矩形は、タイトルの部分パターン範囲を決定するための領域である。領域は

10

20

30

40

50

、それぞれ、初期領域 1 5 0 0、2 段階目の領域 1 5 0 1、最大領域 1 5 0 2 を示している。図 1 5 (a) では、タイトルの項目における部分パターン範囲を決定するための領域が段階的に変更される様子を示している。初期領域から最大領域まで段階的に領域を広げながら、その領域と重なるインデックスブロックを除くテキストブロックをカウントする。そして、カウントされたテキストブロックが所定の数以上になったときの一点鎖線の矩形で示す領域を、その項目の部分パターン範囲として決定する。なお、所定の数は、1 個以上であることが好ましい。本実施形態では、所定の数が 5 であるものとして説明する。

【 0 1 2 8 】

本実施形態の部分パターン範囲の決定方法について具体的に説明する。はじめに、初期領域 1 5 0 0 と少しでも重なっているテキストブロックの数をカウントする。この場合、インデックスブロック 1 0 0 0 以外のテキストブロックが存在しないため、次の段階へ進む。

【 0 1 2 9 】

次に、領域を広げて、2 段階目の領域 1 5 0 1 と少しでも重なっているテキストブロックをカウントする。図 1 5 (b) は、部分パターン範囲を決定するための領域を 2 段階目の領域 1 5 0 1 とした場合の図である。図 1 5 (b) に示すように 2 段階目の領域 1 5 0 1 とは、テキストブロック 1 0 0 1、1 0 0 8 ~ 1 0 1 3 が重なる。このため 2 段階目の領域 1 5 0 1 と重なるテキストブロックは 7 個とカウントされる。そして重なるテキストブロックの数が所定の数である 5 以上であると判定される。このため、タイトルの部分パターン範囲については 2 段階目の領域 1 5 0 1 が示す位置およびサイズに決定される。このため部分パターン範囲に少なくとも一部が含まれるテキストブロック 1 0 0 1、1 0 0 8 ~ 1 0 1 3 と、インデックスブロック 1 0 0 0 とからなるレイアウトが、タイトルの部分レイアウトとして決定される。

【 0 1 3 0 】

または、項目によって、周囲のテキストブロックの数は異なり、記載内容によるテキストブロックのレイアウトの変化が少ない領域は異なる。このため、例えば、項目の属性に応じて部分パターン範囲のサイズを異ならせてもよい。つまり、項目の属性に応じた部分パターンのサイズを予め設定してもよい。

【 0 1 3 1 】

項目がタイトルの場合、タイトルのテキストブロックの近傍にはテキストブロックが存在しないことが多いという特徴がある。また、タイトルは、文書の記載内容の変化によるテキストブロックのレイアウトの変化が少ない文書の上部に存在するという特徴がある。このため、図 1 0 (c) の部分パターン範囲 1 0 0 7 に示すように、項目が文書のタイトルであれば、X 方向は画像幅全体が収まり、Y 方向も画像の約 4 分の 1 が収まるような領域が部分パターン範囲として決定されてもよい。

【 0 1 3 2 】

以上説明したように本実施形態では、文書に応じて部分パターン範囲が決定される。このため、文書に応じて適切な部分パターン範囲によって、インデックスブロック推定処理の精度を向上させることができる。

【 0 1 3 3 】

< 実施形態 3 >

実施形態 1 では、部分パターンを利用して導出された一致度が最大となる位置を X Y 候補位置として決定し、X Y 候補位置の一致度が所定の閾値以上であれば、X Y 候補位置に基づき処理対象の項目のインデックスブロックのある位置を推定する方法を説明した。

【 0 1 3 4 】

しかしながら、入力文書には、登録文書の部分レイアウトと配置が類似したテキストブロックを含む領域が複数存在することがある。入力文書内に部分レイアウトと類似する領域が複数存在する場合、実施形態 1 の方法では、入力文書内における処理対象の項目のインデックスブロックの推定に失敗してしまうことがある。

【 0 1 3 5 】

そこで本実施形態では、処理対象の項目の部分レイアウトに類似した領域が入力文書内に複数存在する場合であっても、入力文書内のインデックスブロックの位置を適切に推定する方法について説明する。なお、本実施形態については、実施形態1からの差分を中心に説明する。特に明記しない部分については実施形態1と同じ構成および処理である。

【0136】

図16は、本実施形態におけるS507のインデックスブロック推定処理を説明するためのフローチャートである。本実施形態におけるインデックスブロック推定処理の詳細について、図16のフローチャートに従い説明する。S1600～S1604はS800～S804と同一であるため説明を省略する。

【0137】

S1605において画像処理部305は、S1604で導出した一致度が所定の閾値以上となるスキャン画像内のXY位置を決定する。本ステップの結果、複数のXY位置が決定されない場合もあるが、便宜的に本ステップによって決定されるXY位置をXY候補位置群と呼ぶ。

【0138】

図17は、インデックスブロックとその周囲のブロックからなる部分レイアウトと類似する領域が複数存在する登録文書の例を示す図である。図17(a)は、登録文書の一例を示す図である。図17(b)は、図17(a)の登録文書における「見積日付(Quotation Date)」の項目に対応する文字列を含むテキストブロック1705をインデックスブロックとした場合の部分パターンを示す図である。図17(b)において、一点鎖線の矩形は、「見積日付」の項目の部分パターン範囲1700を示し、実線の矩形で表されるテキストブロック1701～1706は、「見積日付」の項目の部分レイアウトを構成するテキストブロックを示している。図16のフローチャートの説明では、「見積日付」を処理対象の項目とした場合の処理について説明する。

【0139】

図18は、入力文書を説明するための図である。図18(a)は、入力文書を示す図であり、本フローチャートの説明では、この入力文書がスキャンされた結果得られたスキャン画像に対して、インデックスブロック推定処理が行われるものとして説明する。また、S504の文書マッチングにより、図18(a)の入力文書に類似する文書は、図17の登録文書が特定されたものとして説明する。

【0140】

図18(b)～(e)は、それぞれ、図18(a)の入力文書のスキャン画像に対してブロックセレクション処理を行った結果検出されたテキストブロックを表す画像に、図17(b)の「見積日付」の部分パターンを重畳した図である。図18(b)～(e)の夫々の図における矩形は、部分パターンを示す。即ち、実線の矩形は、部分レイアウトを構成するテキストブロックであり、一点鎖線の矩形は部分パターン範囲である。

【0141】

図18(b)～(e)で示す、部分パターンの位置は、S1604で導出した一致度が所定の閾値以上となったときの位置である。このため部分レイアウトを構成する実線の矩形で表したテキストブロックのうち、インデックスブロックのXY位置1801～1804が、本ステップの処理の結果、XY候補位置群として決定されている。

【0142】

図18(a)に示す入力文書のように、単純なテキストブロックの配置が繰り返し存在する文書において、その繰り返して配置されているテキストブロックの中にインデックスブロックが存在される場合には、一致度が閾値以上となるXY位置が複数決定される。このため、図18(a)に示す入力文書に対して、本ステップの処理がされた結果決定されるXY候補位置群の数は2以上となる。

【0143】

S1606において画像処理部305は、S1605で決定したXY候補位置群の数に応じて処理を切り替える。XY候補位置群の数が1個であれば、S1610に進み、XY

10

20

30

40

50

候補位置群の数が0個であれば、S 1 6 1 2に進む。S 1 6 1 2の処理はS 8 0 9と同一であるため説明を省略する。

【0 1 4 4】

X Y候補位置群の数が2個以上である場合はS 1 6 0 7に進む。S 1 6 0 7において画像処理部3 0 5は、登録文書内の位置であって、処理対象の項目の部分レイアウトとの一致度が所定の閾値以上となる位置である類似位置(群)を取得する。

【0 1 4 5】

登録文書内の位置に、処理対象の項目の部分パターンに含まれる部分レイアウトを重畳させてテキストブロックの一致度の導出を行い、一致度が所定の閾値以上となる登録文書内のX Y位置が「類似位置」として決定される。登録文書内のテキストブロックと部分レイアウトのテキストブロックとの一致度の算出方法は、S 1 6 0 2～S 1 6 0 4と同様の方法で導出されればよい。即ち、入力文書を対象としていたところを、登録文書を対象として同様の手順で一致度を導出すればよい。

【0 1 4 6】

図19は、登録文書内の類似位置を説明するための図である。図19(a)は、図17(a)と同一の登録文書を示す図である。図19(b)～(e)は、それぞれ、図19(a)の登録文書のスキャン画像に対してブロックセレクション処理を行った結果検出されたテキストブロックを表す画像に、図17(b)の「見積日付」の部分パターンを重畳した図である。図19(b)～(e)の夫々の図における矩形は、部分パターンを示す。即ち、実線の矩形は、部分レイアウトを構成するテキストブロックであり、一点鎖線の矩形は部分パターン範囲である。

【0 1 4 7】

図19(b)～(e)の、部分パターンの位置は、導出された一致度が所定の閾値以上となったときの、それぞれの位置である。このため部分レイアウトを構成するテキストブロックのうちインデックスブロックのX Y位置が、類似位置群1 9 0 1～1 9 0 4として決定されている。本ステップでは、処理対象の項目の類似位置群の位置情報が取得される。類似位置群1 9 0 1～1 9 0 4には、類似位置1 9 0 2のように、図17(b)で示した登録時のインデックスブロック1 7 0 5のX Y位置も含まれる。

【0 1 4 8】

なお、S 1 6 0 7で登録文書内の類似位置を決定する処理が行われる必要はない。例えば、文書の登録時において、項目ごとに部分パターンを決定した後に類似位置群を決定し、類似位置群の情報を図7で示した抽出ルールの一部として予め記憶させてもよい。つまり、S 1 6 0 7では、記憶されている処理対象の項目の抽出ルールの1つとして類似位置群が取得されればよい。

【0 1 4 9】

S 1 6 0 8において画像処理部3 0 5は、S 1 6 0 7で取得した登録文書の類似位置群と、S 1 6 0 5で決定した入力文書におけるX Y候補位置群との対応付けを行う。具体的には、Y位置でソートされた類似位置群に対して、類似位置群と同一条件でソートされたX Y候補位置群を、Y位置の一方の側から順番で対応付けを行い、さらにY位置の他方の側からの順番で対応付けを行う。

【0 1 5 0】

図20は、本ステップの処理を説明するための図である。表中の数値は、図18または図19で示した文書内の位置を示す符号を示す数値である。

【0 1 5 1】

図20(a)は、図18および図19で示したように、類似位置群とX Y候補位置群の数が一致している場合の対応付けを示す図である。列2 0 0 1はY位置でソートされた類似位置群である。列2 0 0 2はY位置でソートされたX Y候補位置群であり、列2 0 0 1の類似位置群に対してY位置の上から順番で対応付けられたX Y候補位置群である。列2 0 0 3はY位置でソートされたX Y候補位置群であり、列2 0 0 1の類似位置群に対してY位置の下から順番で対応付けられたX Y候補位置群である。図20(a)では、列2 0

02のXY候補位置群も列2003のXY候補位置群も、それぞれ同じ類似位置と対応付けられる。

【0152】

図20(b)は、XY位置群の数に対して、類似位置群の数が少ない場合の本ステップの対応付けの方法を説明するための図である。例えば、図19(e)に示す登録文書の位置に部分パターンを重畳させた場合の登録文書との一致度が閾値未満であり、S1607では類似位置群1901~1903のみが取得された場合の、対応付けを表した図が図20(b)である。列2011はY位置でソートされた類似位置群である。列2012は、列2011の類似位置群に対してY位置の上から順番で対応付けられたXY候補位置群である。列2013は、列2011の類似位置群に対してY位置の下から順番で対応付けられたXY候補位置群である。図20(b)では、上からの対応付けと下からの対応付けでは、類似位置群に対応するXY候補位置群が異なる結果となっている。

10

【0153】

図20(c)は、XY候補位置群の数に対して、類似位置群の数が多の場合の本ステップの対応付けの方法を説明するための図である。図18(e)に示す入力文書の位置に部分パターンを重畳させた場合の入力文書との一致度が閾値未満であり、S1605ではXY位置1801~1803のみがXY候補位置群として決定された場合の、対応付けを表した図が図20(c)である。列2021はY位置でソートされた類似位置群である。列2022は、列2021の類似位置群に対してY位置の上から順番で対応付けられたXY候補位置群である。列2023は、列2021の類似位置群に対してY位置の下から順番で対応付けられたXY候補位置群である。上からの対応付けと下からの対応付けとは異なる結果となり、上からの対応付けでは類似位置1904に対応するXY候補位置群は見つからず、下からの対応付けでは類似位置1901に対応するXY候補位置群は見つからない結果となる。

20

【0154】

S1609において画像処理部305は、S1608で行った対応付けの結果に基づき、S1605で決定されたXY候補位置群から1つのXY候補位置を決定する。

【0155】

S1608で行われた対応付けの結果が、図20(a)に示したように、上からの対応付けと下からの対応付けの結果が一致する場合がある。この場合は、XY候補位置群のうち、登録時のインデックスブロックの位置を示す類似位置に対応付けられたXY位置を、1つのXY候補位置として決定する。図20(a)の例では、インデックスブロックの位置を示す類似位置1902に対応付けられたXY位置1802が、1つのXY候補位置として決定される。

30

【0156】

一方、S1608で行われた対応付けの結果が、図20(b)および(c)で示したように、上からの対応付けと下からの対応付けの結果が一致しない場合がある。この場合ははじめに、上からの対応付けを行った場合の、インデックスブロックの位置を示す類似位置に対応付けられた入力文書のXY位置を決定する。さらに、下からの対応付けを行った場合の、インデックスブロックの位置を示す類似位置に対応付けられた入力文書のXY位置を決定する。

40

【0157】

図20(b)の例では、インデックスブロックの位置を示す類似位置1902に対応付けられた、XY位置1802とXY位置1803とが決定される。図20(c)の例では、類似位置1902に対応付けられた、XY位置1802とXY位置1801とが決定される。そして、決定された2つのXY位置のうち、S1604で導出した一致度が高い方を、XY候補位置群のうちの1つのXY候補位置として決定する。なお、一致度をいなしで、2つのXY位置から1つの中から1つのXY位置を選択してもよい。例えば、2つのXY位置を表示させてユーザからの指示を受け付け、上からの対応付けと下からの対応付けのどちらを利用するかを項目ごとに覚えておいて利用してもよい。

50

【 0 1 5 8 】

X Y 候補位置群から 1 つの X Y 候補位置を決定されると S 1 6 1 0 に進む。S 1 6 1 0 では、S 8 0 7 の処理と同様に、X Y 候補位置を処理対象のインデックスブロックのある位置として推定して、スキャン画像のテキストブロックから、処理対象の項目のインデックスブロックを推定する処理を行う。S 1 6 1 1 は S 8 0 8 と、S 1 6 1 3 は S 8 1 0 とそれぞれ同一であるため説明を省略する。

【 0 1 5 9 】

以上説明したように本実施形態では、入力文書において一致度が閾値以上となる X Y 候補位置が複数存在した場合に、部分パターンとの一致度が閾値以上となる登録文書の類似位置群との対応付けを行った上で 1 つの X Y 候補位置を決定する。このため、インデックスブロックとその周囲のテキストブロックからなる部分レイアウトに類似した領域が文書内に複数存在する場合でも、インデックスブロック推定処理の精度を向上させることができる。

10

【 0 1 6 0 】

< その他の実施形態 >

上述の実施形態では、画像形成装置 1 0 0 が単体で図 4 のフローチャートの各ステップの処理を行う例を説明した。他にも、これらの処理の全部または一部を図 3 の機能を有するシステム 1 0 5 上の他の画像処理装置で行う形態でもよい。

【 0 1 6 1 】

例えば、スキャン処理を画像形成装置 1 0 0 で実行して、スキャン画像を端末 1 0 1 にネットワークを介して送信する。端末 1 0 1 が画像処理部 3 0 5 と同様の機能を有しており、端末 1 0 1 においてインデックス抽出処理を実行してもよい。この場合、端末 1 0 1 はインデックス抽出結果を画像形成装置 1 0 0 に返信して、画像形成装置 1 0 0 は取得したインデックス抽出結果に基づきファイル生成およびファイル送信をする。

20

【 0 1 6 2 】

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、A S I C）によっても実現可能である。

【 符号の説明 】

30

【 0 1 6 3 】

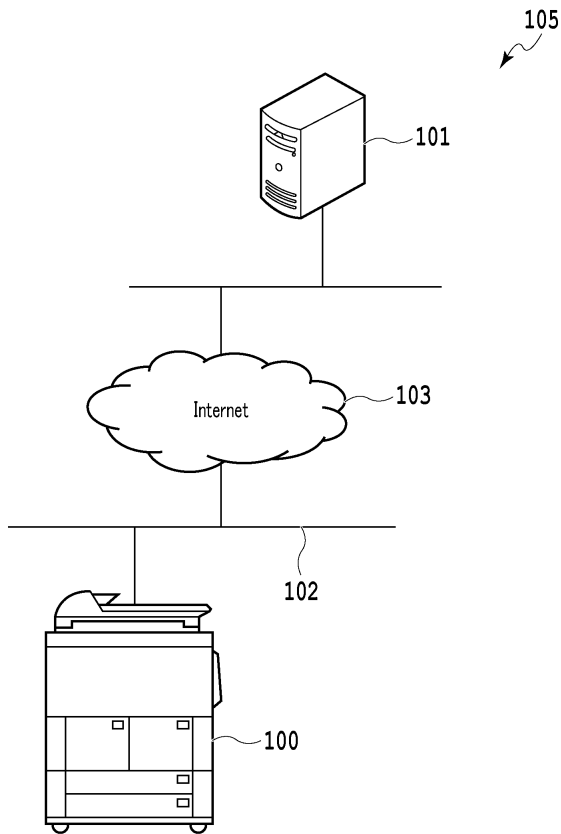
1 0 0 画像形成装置
3 0 5 画像処理部
1 1 1 C P U

40

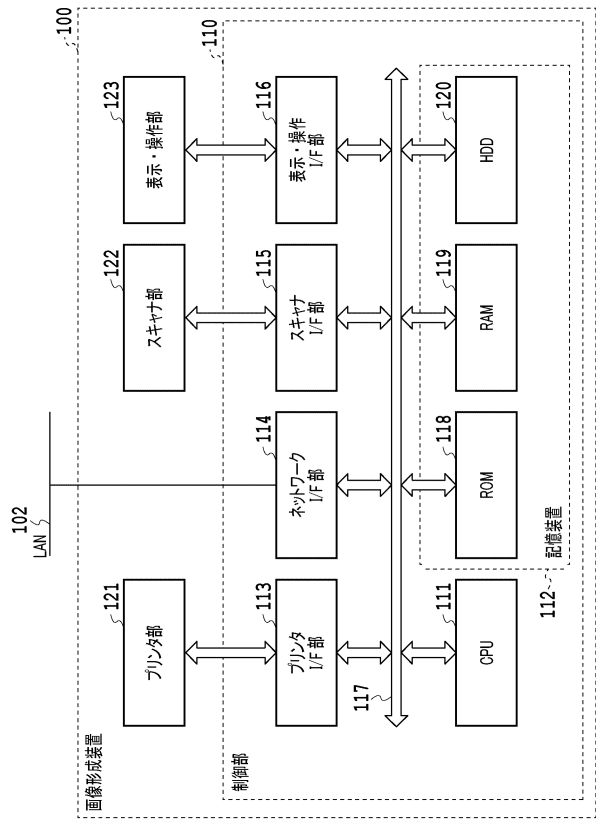
50

【 図面 】

【 図 1 】



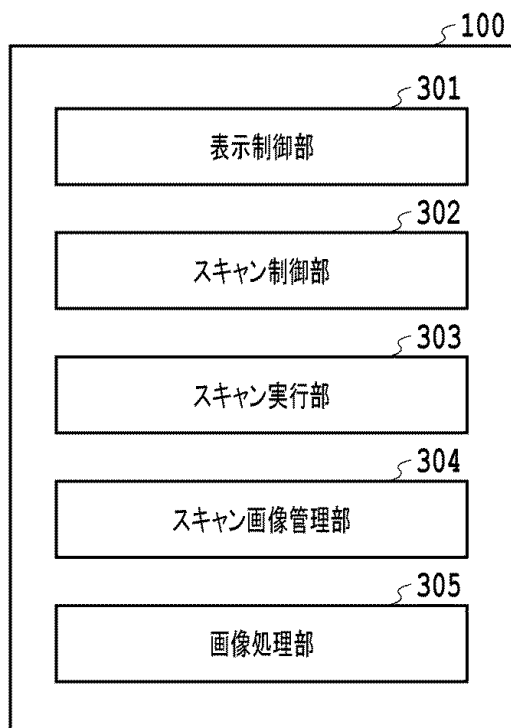
【 図 2 】



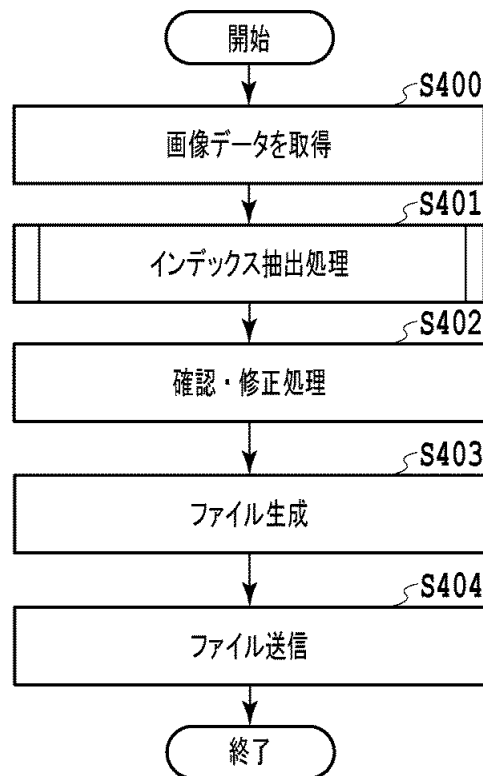
10

20

【 図 3 】



【 図 4 】

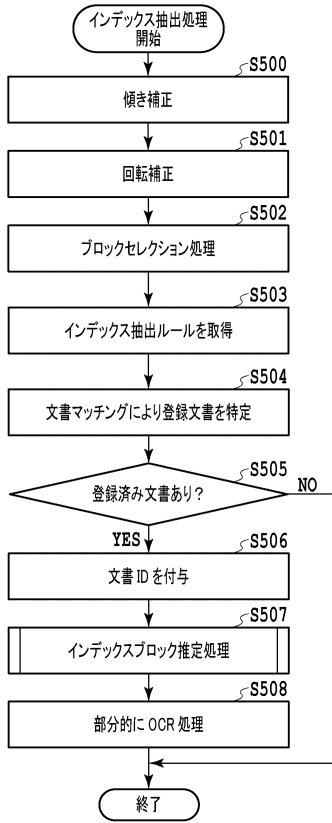


30

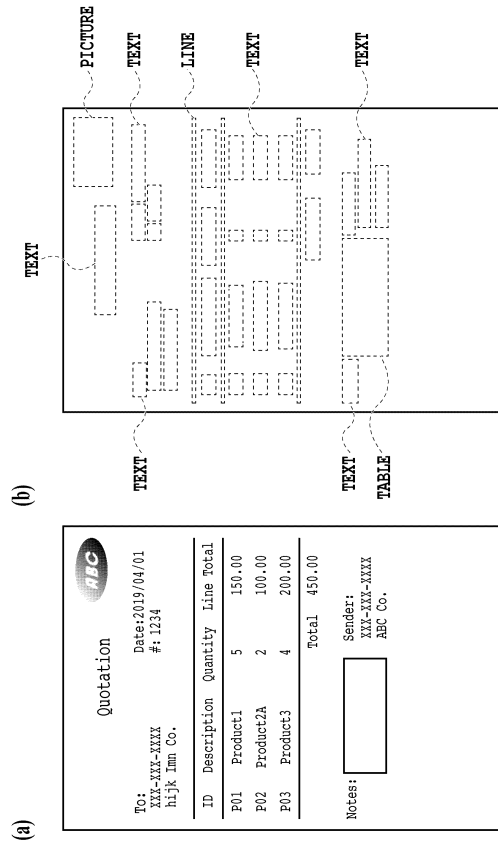
40

50

【図 5】



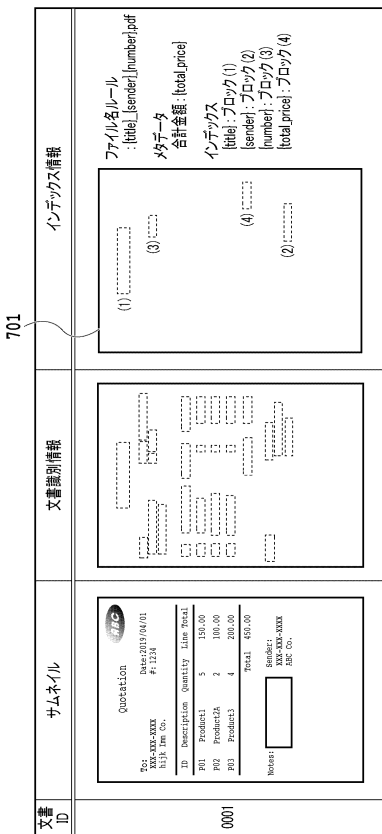
【図 6】



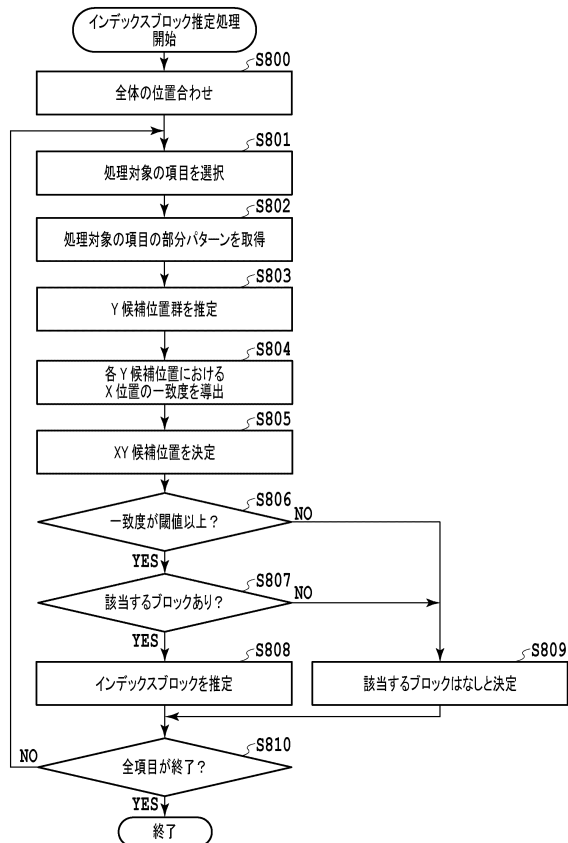
10

20

【図 7】



【図 8】

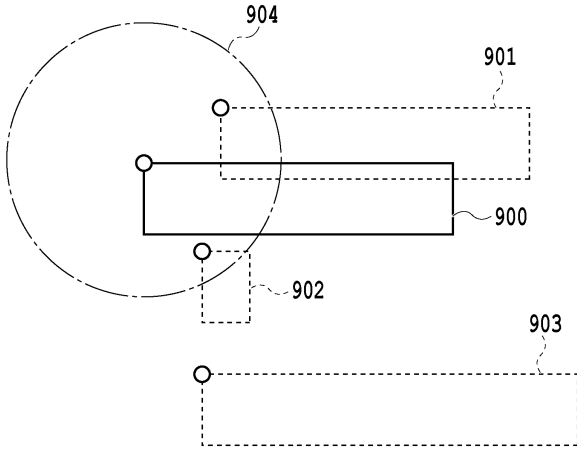


30

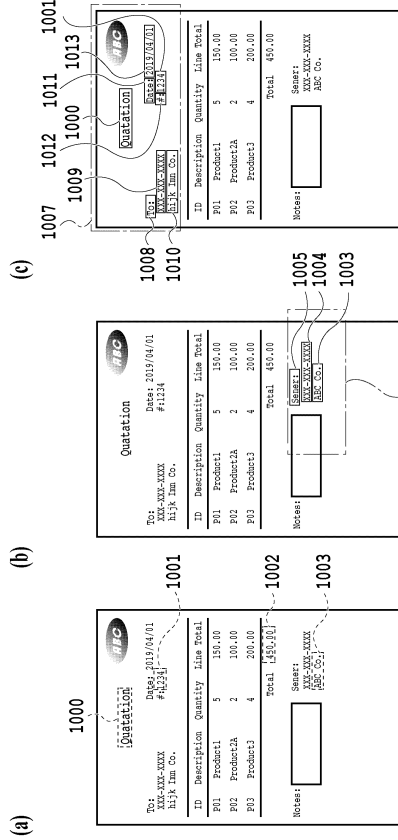
40

50

【 9 】



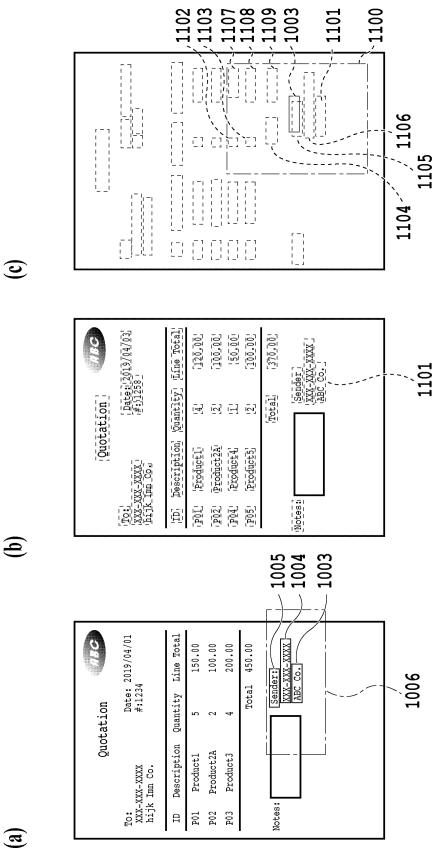
【 10 】



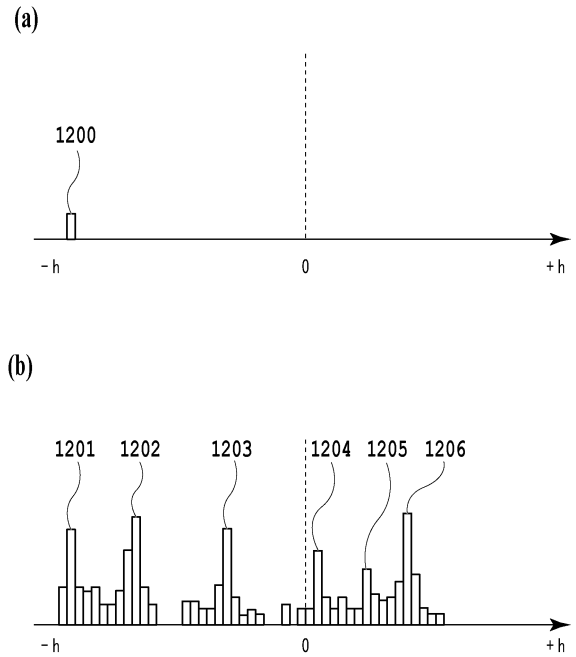
10

20

【 11 】



【 12 】

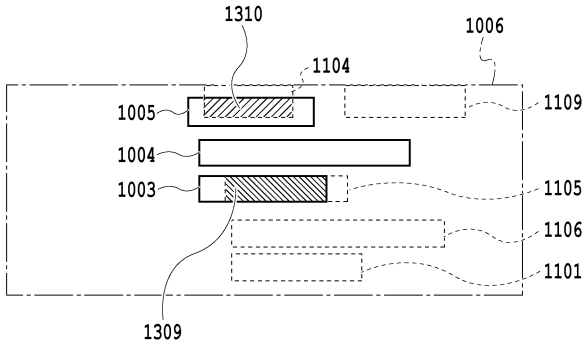


30

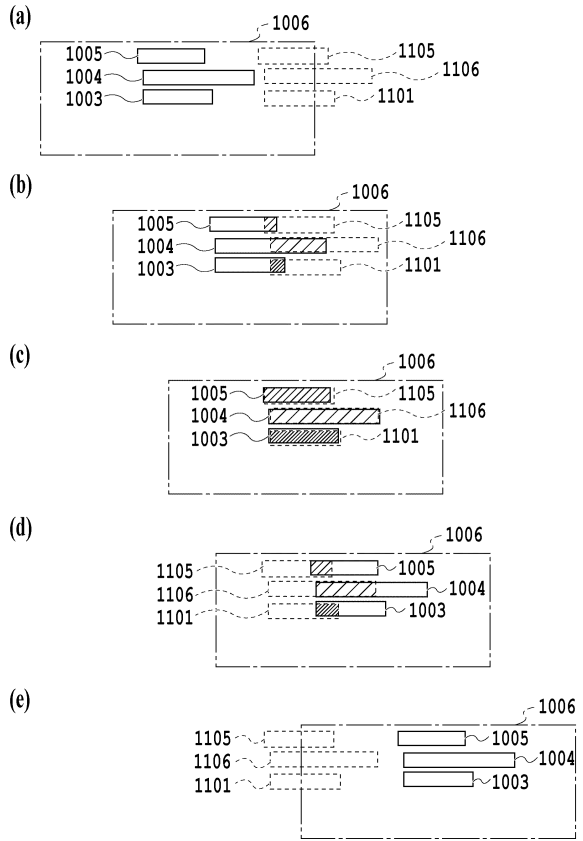
40

50

【図13】



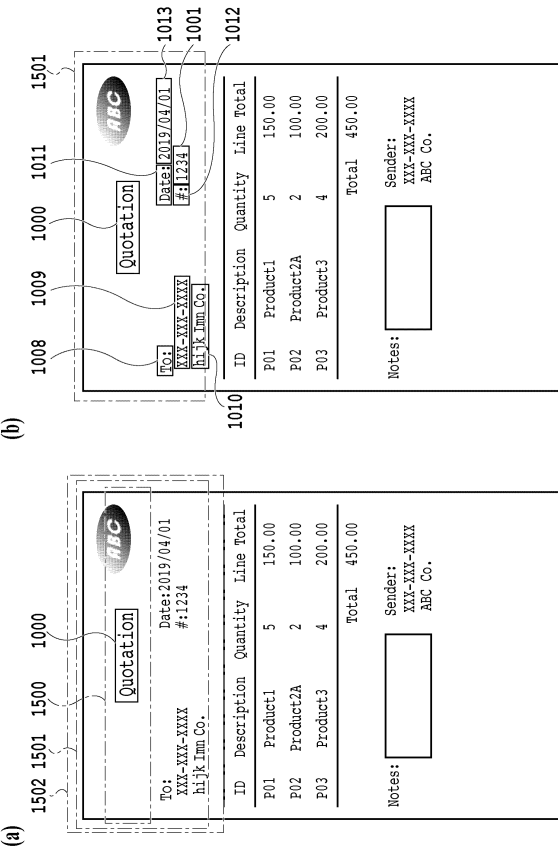
【図14】



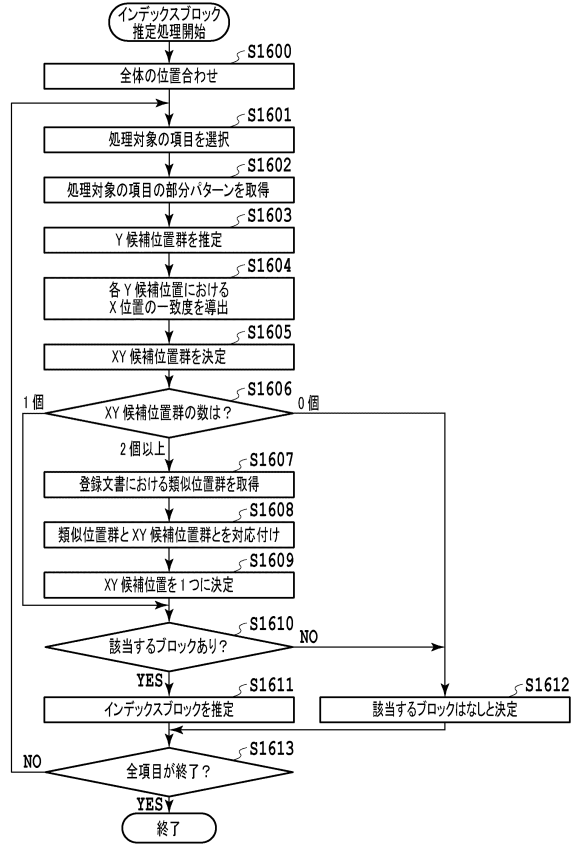
10

20

【図15】



【図16】



30

40

50

【 図 1 7 】

(a) 登録文書

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

(b)

Customer Number	AAA1234	1700
Quotation Number	Q3456789	1704
Quotation Date	4/30/2019	1705
Sales Person	ABC DEFGH	1706
Payment Terms	Credit Card	
Expiration Date	5/29/2019	

【 図 1 8 】

(a) 入力文書

Customer Number	AAA1234
Quotation Number	Q3456820
Quotation Date	5/17/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	6/16/2019

(b) 1801

Customer Number	AAA1234
Quotation Number	Q3456820
Quotation Date	5/17/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	6/16/2019

(c)

1802

Customer Number	AAA1234
Quotation Number	Q3456820
Quotation Date	5/17/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	6/16/2019

(d)

1803

Customer Number	AAA1234
Quotation Number	Q3456820
Quotation Date	5/17/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	6/16/2019

(e)

1804

Customer Number	AAA1234
Quotation Number	Q3456820
Quotation Date	5/17/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	6/16/2019

10

20

【 図 1 9 】

(a) 登録文書

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

(b) 1901

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

(c)

1902

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

(d)

1903

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

(e)

1904

Customer Number	AAA1234
Quotation Number	Q3456789
Quotation Date	4/30/2019
Sales Person	ABC DEFGH
Payment Terms	Credit Card
Expiration Date	5/29/2019

【 図 2 0 】

(a)

	2001	2002	2003
類似位置群	上からの対応付け	下からの対応付け	
1901	1801	1801	
1902	1802	1802	
1903	1803	1803	
1904	1804	1804	

類似位置群とXY候補位置群の数が同じ

(b)

	2011	2012	2013
類似位置群	上からの対応付け	下からの対応付け	
1901	1801	1802	
1902	1802	1803	
1903	1803	1804	

類似位置群の方が少ない (1904が無い場合)

(c)

	2021	2022	2023
類似位置群	上からの対応付け	下からの対応付け	
1901	1801	無し	
1902	1802	1801	
1903	1803	1802	
1904	無し	1803	

XY候補位置群の方が少ない (例 1804が無い場合)

30

40

50

フロントページの続き

- (56)参考文献 特開 2019 - 057173 (JP, A)
特開 2016 - 191989 (JP, A)
特開 2008 - 181460 (JP, A)
特開 2019 - 128715 (JP, A)

- (58)調査した分野 (Int.Cl., DB名)
G06V 30/412
H04N 1/387
G06V 30/14