(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4227508号
(P4227508)
(45)発行日　平成21年2月18日(2009.2.18)　(24)登録日　平成20年12月5日(2008.12.5)

(51)Int.Cl.			F I		
G06T	1/00	(2006.01)	G06T	1/00	200D
G06T	7/60	(2006.01)	G06T	7/60	150P
H04N	1/21	(2006.01)	H04N	1/21	

請求項の数 12　(全 23 頁)

(21)出願番号　特願2003-412214(P2003-412214)
(22)出願日　平成15年12月10日(2003.12.10)
(65)公開番号　特開2005-173922(P2005-173922A)
(43)公開日　平成17年6月30日(2005.6.30)
審査請求日　平成17年11月15日(2005.11.15)

(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
(74)代理人　100076428
弁理士　大塚　康徳
(74)代理人　100112508
弁理士　高柳　司郎
(74)代理人　100115071
弁理士　大塚　康弘
(74)代理人　100116894
弁理士　木村　秀二
(72)発明者　齋藤　和之
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】情報処理装置及び情報処理方法

(57)【特許請求の範囲】
【請求項１】
画像入力手段を介して入力されたページ単位の画像データを、章を含む階層構造を有する統合文書として登録することが可能な情報処理装置であって、
前記画像入力手段を介して入力された各画像データの方向を検出する方向検出手段と、
前記方向検出手段により検出された前記各画像データの方向が前ページの画像データの原稿方向と同一であるか否かを検出する連続性検出手段と、
前記連続性検出手段による検出結果に基づき、画像データの方向が前ページと同一であるときに当該画像データを前のページと同じ章のページとして登録し、画像データの方向が前ページと同一でないときに、当該画像データを新規の章のページとして登録する登録手段と、
を備えることを特徴とする情報処理装置。
【請求項２】
方向検出手段と、連続性検出手段と、登録手段と、画像入力手段とを有し、前記画像入力手段を介して入力されたページ単位の画像データを、章を含む階層構造を有する統合文書として登録する情報処理装置における情報処理方法であって、
前記方向検出手段が、前記画像入力手段を介して入力された各画像データの方向を検出する方向検出工程と、
前記連続性検出手段が、前記方向検出工程により検出された前記各画像データの方向が前ページの画像データの原稿方向と同一であるか否かを検出する連続性検出工程と、

前記登録手段が、前記連続性検出工程による検出結果に基づき、画像データの方向が前ページと同一であるときに当該画像データを前のページと同じ章のページとして登録し、画像データの方向が前ページと同一でないときに、当該画像データを新規の章のページとして登録する登録工程と、

を備えることを特徴とする情報処理方法。

【請求項３】

前記方向検出工程により検出された各画像データの方向が予め定められた方向に該当しない場合、前記連続性検出工程は、当該各画像データを、前のページに連続する画像データとすることを特徴とする請求項２に記載の情報処理方法。

【請求項４】

前記方向検出工程により、前記画像データが白紙として検出された場合、前記連続性検出工程は、当該白紙の画像データを、前のページに連続する画像データとすることを特徴とする請求項２に記載の情報処理方法。

【請求項５】

前記方向検出工程は、前記各画像データの方向を検出するために、異なる複数の方向に関して、確からしさを示す方向検出確度を前記各画像データについてそれぞれ演算し、当該演算された複数の方向検出確度を比較して、最大の方向検出確度に対応する方向を、当該各画像データの方向として検出することを特徴とする請求項２に記載の情報処理方法。

【請求項６】

前記連続性検出工程は、前後のページに対応する画像データの方向に基づき、孤立する画像データを検出し、当該孤立する画像データが連続する数に従って、前記方向検出工程により検出された画像データの方向を修正することを特徴とする請求項２に記載の情報処理方法。

【請求項７】

前記孤立する画像データが１ページの場合は、前記連続性検出工程は、当該孤立する画像データの方向を、前のページに連続する画像データとして修正することを特徴とする請求項６に記載の情報処理方法。

【請求項８】

前記孤立する画像データが２ページ連続の場合は、前記連続性検出工程は、当該孤立する各画像データの方向検出確度の大小関係に従って、いずれか一方の画像データの方向を修正することを特徴とする請求項６に記載の情報処理方法。

【請求項９】

前記孤立する画像データが３ページ以上連続する場合、前記連続性検出工程は、当該孤立する各画像データの方向を修正しないことを特徴とする請求項６に記載の情報処理方法。

【請求項１０】

前記登録工程により登録された、前記画像データと前記検出結果に基づいて、表示手段に当該画像データをプレビュー表示する表示制御工程を備えることを特徴とする請求項２乃至９のいずれか１項に記載の情報処理方法。

【請求項１１】

前記表示制御工程は、前記プレビュー表示において、前記検出結果に基づき、方向が修正されている画像データを識別するための識別表示を当該画像データのプレビュー表示に付加することを特徴とする請求項１０に記載の情報処理方法。

【請求項１２】

請求項２乃至１１のいずれか１項に記載の情報処理方法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、様々なデータを一つの電子化文書としてまとめ、その文書に対する印刷条件

や編集機能の設定を行なう文書処理技術に関するものである。特に、複数の種類の原稿を一括して電子化し、この電子化された原稿を再利用（印刷設定等）する場合に、入力した原稿の種類毎に区分けをし、この区分けに従って印刷条件等を設定して文書処理を行う情報処理装置及び情報処理方法に関するものである。

【背景技術】

【０００２】

複数種類の原稿を一まとめに電子化する際、入力した原稿を種類毎に区分けする方法としては、原稿と原稿の間に白紙を挿入する方法によるもの（例えば、特許文献１を参照）や、原稿と原稿の間に所定の紙（仕切り用紙）を挿入する方法によるもの（例えば、特許文献２を参照）などがある。

【０００３】

また、白紙や仕切り用紙等を使用せずに区分けをする方法としては、入力する原稿の向き（縦横）を検出して、検出した原稿の向きが直前の原稿の向きと異なるときに、向きの変化を入力原稿の区切りとして認識し、原稿の向きが変わる前と、向きが変わった後の原稿とを別文書として外部記憶装置に登録させるもの（例えば、特許文献３を参照）や、原稿に記録されている文字情報を認識の確からしさを評価する文字認識確度をパラメータとして帳票の向きを判定し、直前の帳票の向きと異なった向きを検出した場合には、その変化の前後を帳票の境界として認識し、その境界の前後を単位として帳票画像データをファイルするもの（例えば、特許文献４を参照）がある。

【特許文献１】特開平０８－７９４９９号公報

【特許文献２】特開平０３－２９２５６８号公報

【特許文献３】特開平０７－１９２１０９号公報

【特許文献４】特開平１１－２１３１７１号公報

【発明の開示】

【発明が解決しようとする課題】

【０００４】

しかしながら、上述した従来の技術では、原稿を読み取りこれを一つの文書としてまとめようとすると、以下のような問題が生じる。

【０００５】

（１）白紙を挿入する方法の場合、例えば、両面印刷された原稿Ａと片面印刷された原稿Ｂとを、一括して自動給紙装置（ＡＤＦ）で連続的に入力する場合、片面印刷された裏側の白紙を原稿の区切り（「章の区切り」）と誤認識してしまう問題が生じる。また、両面印刷の原稿を連続して入力する場合でも原稿に記録されている印刷データが極端に少ない場合は、白紙と誤認識してしまうという問題もある。

【０００６】

（２）また、白紙や所定の紙（仕切り用紙）を挿入する場合には、一括して電子化しようとする原稿の区分けに必要な数量だけ、白紙や仕切り用紙を予め用意し、原稿間に白紙や仕切り用紙を挿入しなければならないという人為的な作業が必要となり、印刷に付帯する作業負担が大きくなるとともに、大量の原稿を処理するという観点からすると印刷処理に関する効率が低下するという問題も生じる。

【０００７】

（３）また、特許文献３や特許文献４のように、様々な原稿を一括して電子化する際に、区切り毎に各々異なる文書として保存や登録を行うものは、いずれも別文書あるいは別ファイルとして保存されてしまうこととなり、複数の原稿を一括して電子化し、一冊の印刷物の章として区分けをして編集し、印刷設定を行うことができないという問題もある。この場合、オペレータは、一旦保存した各文書を、再度読み込んで、個別の原稿に対して種々の条件設定し直すなどの手間が必要とされるという問題も生じる。

【０００８】

（４）更に、文字認識を利用した方向検出方法の場合、ほとんど空白に近いページや文字が無く絵や写真等が含まれている書類は、そのページの方向を正しく検出することが困

難であり、原稿の方向によって区切りを付ける際に誤判定してしまうという問題も生じる。

【０００９】

本発明は、上記背景技術に鑑みてなされたもので、複数種類の原稿を入力する場合、電子化文書を原稿方向の連続性に基づいて区分けした階層構造としてまとめ、一冊の原稿として印刷、編集設定することを可能にする文書処理装置及び文書処理方法を提供することを目的とする。

【００１０】

あるいは、本発明は、階層構造にまとめる際、誤認識の可能性のある原稿を特定するための識別表示をすることで、誤認識を修正するための操作、編集処理を容易にする情報処理装置及び情報処理方法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】

【００１１】

上記課題を解決するべく、本発明にかかる情報処理装置等は、主として以下の構成よりなることを特徴とする。

【００１２】

すなわち、画像入力手段を介して入力されたページ単位の画像データを、章を含む階層構造を有する統合文書として登録することが可能な情報処理装置は、

前記画像入力手段を介して入力された各画像データの方向を検出する方向検出手段と、

前記方向検出手段により検出された前記各画像データの方向が前ページの画像データの原稿方向と同一であるか否かを検出する連続性検出手段と、

前記連続性検出手段による検出結果に基づき、画像データの方向が前ページと同一であるときに当該画像データを前のページと同じ章のページとして登録し、画像データの方向が前ページと同一でないときに、当該画像データを新規の章のページとして登録する登録手段と、

を備えることを特徴とする。

【００１６】

あるいは、方向検出手段と、連続性検出手段と、登録手段と、画像入力手段とを有し、前記画像入力手段を介して入力されたページ単位の画像データを、章を含む階層構造を有する統合文書として登録する情報処理装置における情報処理方法は、

前記方向検出手段が、前記画像入力手段を介して入力された各画像データの方向を検出する方向検出工程と、

前記連続性検出手段が、前記方向検出工程により検出された前記各画像データの方向が前ページの画像データの原稿方向と同一であるか否かを検出する連続性検出工程と、

前記登録手段が、前記連続性検出工程による検出結果に基づき、画像データの方向が前ページと同一であるときに当該画像データを前のページと同じ章のページとして登録し、画像データの方向が前ページと同一でないときに、当該画像データを新規の章のページとして登録する登録工程と、

を備えることを特徴とする。

【発明の効果】

【００１７】

本発明は、複数種類の原稿を入力する場合、電子化文書を原稿方向の連続性に基づいて区分け（章分割）した階層構造としてまとめ、一冊の原稿として印刷、編集設定することが可能になる。

【００１８】

あるいは、階層構造にまとめる際、誤認識の可能性のある原稿を特定するための識別表示をすることで、誤認識を修正するための操作を容易にすることが可能になる。

【発明を実施するための最良の形態】

【００１９】

図１は本発明が提供する文書処理技術を適用した装置の概略的な構成を説明するブロッ

ク図である。以下、図面を参照しつつその構成を説明する。

【００２０】

図１において、１０１は原稿画像を入力し、入力した原稿の方向を識別して原稿の連続性を検出するための原稿入力処理部であり、同処理部１０１には原稿をスキャン画像として入力する画像入力部１０４、スキャン画像における原稿方向を検出する原稿方向検出部１０５、原稿方向検出部１０５により検出された原稿方向の連続性を検出する原稿方向連続性検出部１０６、スキャン画像や原稿方向や原稿方向連続性に関する情報を保管するデータ保管部１０７、そして、各処理部１０４～１０６、データ保管部１０７とのデータのやり取りや、保管データに関する管理テーブル自体の管理を行う保管データ管理部１０８が設けられている。

【００２１】

１０２は統合文書を作成し印刷設定編集を行う情報処理装置（印刷設定編集部とも呼ぶ）である。この印刷設定編集部１０２内における各機能ブロックは実際にはハードウェアのみで実現する機能、或いは、ハードウェアとソフトウェアとの協働で実現できる機能を示す。同編集部１０２には、原稿入力処理部１０１の保管データ管理部１０８と接続する統合文書作成部１０９が設けられている。この統合文書作成部１０９は画像入力部１０４によって入力されたスキャン画像をページを表すデータとし、原稿方向の連続性の情報に基づいてスキャン画像の種類毎に階層化し、１つの文書として保持する統合文書を作成する。例えば、原稿入力処理部１０１を、プリンタ機能及びコピー機能を備えたデジタル複合機とし、印刷設定編集部１０２をユーザが利用するホストコンピュータに相当させた場合には、保管データ管理部１０８、統合文書作成部１０９には、ネットワークインタフェース機能が備えられており、保管データ管理部１０８と、統合文書作成部１０９とは、イーサーネット（登録商標）ケーブル等により接続される。また、別の形態としては、原稿入力処理部１０１、印刷設定編集部１０２とをデジタル複合機内に含めるようにしても良い。

【００２２】

更に、情報処理装置１０２は、ページレイアウトを編集するページレイアウト編集部１１１、印刷データ出力部への印刷指示を編集する印刷指示編集部１１２、印刷指示並びにページレイアウトの情報を印刷設定情報としてジョブチケットファイルに保存する印刷設定情報保存部１１３、印刷設定情報とページデータを１つの印刷データとする印刷データ変換部１１４を有し、これらの編集情報や印刷データへの変換情報等に基づいて、印刷設定制御部１１６は統合文書作成部１０９や各編集部（１１１、１１２）、印刷設定情報保存部１１３、印刷データ変換部１１４、表示制御部１１７を制御する。また、表示制御部１１７は編集処理結果や作成した統合文書のページ情報等の内容に基づいて表示部１１０上に処理内容をプレビュー等の表示するための制御を行う。印刷設定制御部１１６で処理されたデータは、表示制御部１１７に送られ、登録されているスキャンページデータのレイアウト表示が表示部１１０上に表示される。

【００２３】

また、印刷データ変換部１１４の処理に基づき、印刷設定情報とページデータが１つの印刷データとして変換された印刷データは印刷データ送信部１１５に送られ、印刷データを印刷する印刷データ出力部１０３から出力することができる。印刷データ出力部１０３は、原稿入力処理部１０１をデジタル複合機とした場合には、原稿入力処理部１０１と同一の装置に含めるようにしても良いし、原稿入力処理部１０１のデジタル複合機とは別の画像形成装置とするようにしても良い。また、原稿入力処理部１０１、印刷設定編集部１０２とをデジタル複合機内に含める場合には、印刷データ出力部を同じデジタル複合機内に含めるようにしても良い。

【００２４】

図２は図１のシステム（文書処理装置とも呼ぶ）における処理の流れを説明するためのフローチャートである。まず、ステップＳ２０１において、原稿をスキャン画像として入力するための原稿入力処理を行う。この処理の具体的な流れは図３のフローチャートに従

う。まず、ユーザは、入力の対象となる原稿を自動給紙装置（ＡＤＦ）にセットし（図３のステップＳ３０１）、原稿がＡＤＦにセットされると画像入力部１０４が原稿の入力処理を実行する（図３のステップＳ３０２）。この際、入力される原稿が複数種類ある場合、ユーザは種類毎に文書の方向を変えてＡＤＦ上にセットする。例えば、図４に例示するように、原稿Ａ５０１と原稿Ｂ５０２の２種類の原稿を入力する場合、原稿Ａを正立方向（図６の方向１）にセットして入力し、原稿Ｂを原稿Ａの正立方向とは逆向きの倒立方向（図６の方向３）にセットして入力する。尚、図４に例示した原稿は２種類の場合を示しているが、本実施形態にかかる文処理は、これに限定されるものでなく、更に、原稿の種類が増えた場合であっても入力処理をすることは可能であり、この場合は図４に示した原稿Ａ、原稿Ｂのように、正立方向と倒立方向のように交互に向きを変えてＡＤＦにセットして入力すればよい。

【００２５】

画像入力部１０４により読み取られた原稿はスキャン画像としてデータ保管部１０７に保管され（Ｓ３０３）、処理をステップＳ３０４に進める。ステップＳ３０４では、全ての原稿の入力が終了したか否かを判断し、原稿の入力が全て終了していない場合には（Ｓ３０４－Ｎｏ）、処理をステップＳ３０２に戻し、再度、画像の入力処理（Ｓ３０２）以降の処理を実行し、ステップＳ３０４の判断で、全ての原稿の入力が終了した場合（Ｓ３０４－Ｙｅｓ）には、ステップＳ２０１の原稿入力処理は終了する。

【００２６】

説明を図２のフローチャートに戻し、次に、処理をステップＳ２０２の原稿方向連続性検出処理に進め、先のステップＳ２０１で入力した原稿に関して原稿の向き（方向）が同じ方向で入力されているか否か（原稿方向の連続性）を検出するための処理を、原稿方向検出処理部１０５及び原稿方向連続性検出処理部１０６が実行する。この連続性を検出するための具体的な処理の流れは、図５のフローチャートに従う。

【００２７】

まず、ステップＳ５０１で、データ保管部１０７に一時保管されているスキャン画像を最初のページから読み出し、ステップＳ５０２で原稿の向き（原稿方向）を検出するための検出処理を原稿方向検出処理部１０５が実行する。原稿方向は図６に示すように「方向１～４」に類別され、方向１を正立方向とすると方向３は逆向きの倒立方向となる。方向２及び方向４に関しても同様であり、原稿を構成する各ページ単位に、方向を検出するためフラグが定められる。原稿が白紙の場合は、原稿方向の検出処理を行わずに予め定めた例外処理を行うために、フラグは強制的に「０」と設定される。

【００２８】

ここで、原稿方向検出処理部１０５は、原稿の向きが方向１～４の何れに対応するかを検出するために、本実施形態では原稿中に含まれる１つの文字に該当する画像を切り出し、その画像とデータ保管部１０７に保管されている認識辞書の文字とのマッチングの度合いを評価し、最も高い類似度の認識結果を与える方向（１～４）を原稿の方向として特定する。

【００２９】

切り出した画像と認識辞書データとのマッチング処理は、具体的には以下の演算処理に従う。ここで、切り出した文字に該当する画像数=ｃ、各方向毎の類似度の和をΣｎ（ｎ＝１～４)、各方向毎の平均類似度をｋ（ｎ＝１～４）とすると、

ｋ（ｎ）　＝　Σ（ｎ）　／　ｃ　　・・・（１）

（１）式により平均類似度ｋ(n)を各方向毎に求め、最大の平均類似度ｋ（ｎ）を与える方向が原稿の方向として検出される。

【００３０】

ここで検出された原稿方向は、一時保管データ管理テーブル（図７）の「検出方向」のデータとして記録され、スキャン画像と同一のデータ保管部１０７に保管される。一時保管データ管理テーブルに記録した原稿方向（「検出方向」）から、原稿を構成する各ページの連続性を検出するために、原稿方向連続性検出処理部１０６は連続性開始フラグを設

定するための処理を行う。

【００３１】

説明を図５のフローチャートに戻し、ステップＳ５０３において、原稿方向連続性検出処理部１０６は、最初の原稿のページか否かを判断する。原稿の最初のページである場合（Ｓ５０３－Ｙｅｓ）は、ステップＳ５１０に処理を進め「連続性開始フラグ」をｏｎとして登録する（図７の７０１を参照）。

【００３２】

一方、ステップＳ５０３の判断で、原稿の最初のページでないと判断された場合（Ｓ５０３－Ｎｏ）は、処理をステップＳ５０４に進める。

【００３３】

ステップＳ５０４では、ステップＳ５０２の処理により検出された原稿方向を示すフラグに基づき原稿方向の変化の有無を判断する。原稿方向を示すフラグが、前ページにおけるスキャン画像の原稿方向としてフラグに設定されているデータと同一である場合（Ｓ５０４－Ｎｏ）は連続性に変化は無いとして、処理をステップＳ５０５に進め一時保管データ管理テーブル（図７）の「連続性開始フラグ」をｏｆｆとして記録する（Ｓ５０５）。

【００３４】

一方、ステップＳ５０４の判断で、前ページにおけるスキャン画像の原稿方向としてフラグに設定されているデータと異なる場合（Ｓ５０４－Ｙｅｓ）は、処理をステップＳ５０６に進める。ステップＳ５０６では、原稿方向としてフラグに設定されているデータが方向＝０、すなわち白紙か否かを判定し、白紙（方向＝０）である場合（Ｓ５０６－Ｙｅｓ）は、処理をステップＳ５０８に進め、強制的に前ページの原稿方向と連続しているとして、前ページの原稿方向と同じ方向を示すデータ（この場合は方向１）を一時保管データ管理テーブル（図７）の「最終的方向」に記録する（図７の７０３を参照）。

【００３５】

ステップＳ５０９では、当初の検出方向（白紙の場合はフラグの値を「０」として設定された結果による）と異なる方向が最終的に設定された履歴として「方向修正フラグ」をｏｎと記録する（図７の７０５を参照）。

【００３６】

一方、ステップＳ５０６の判断で、白紙ではない場合（Ｓ５０６－Ｎｏ）は、処理をステップＳ５０７に進め、予め定めた方向以外の方向が検出されているか否かが判断される。本実施形態では、予めＡＤＦに対して原稿を正立方向＝１若しくは倒立方向＝３にセットする場合が想定されているので、例えば、正立方向＝１もしくは倒立方向＝３以外の方向（方向２もしくは方向４）が検出された場合は、ステップＳ５０２における原稿方向の検出処理に誤りがあると判断する。

【００３７】

例えば、図７の一時保管データ管理テーブルにおける「ページ３」のように、方向２が検出された場合は、原稿方向の検出処理に誤りがあると判断し、強制的に前ページの原稿方向と同一の方向を示すデータが一時保管データ管理テーブル（図７）の「最終的方向」に記録され（Ｓ５０８、図７の７０２を参照）、当初の検出方向（＝２）と異なる方向が最終的に設定されたこと示す履歴として「方向修正フラグ」をｏｎと記録する（Ｓ５０９、図７の７０６を参照）。

【００３８】

そして、処理をステップＳ５０５に進め、一時保管データ管理テーブル（図７）の「連続性開始フラグ」をｏｆｆとし、前ページとの間では連続性が維持された状態として検出結果が記録される。

【００３９】

一方、ステップＳ５０７の判断で、検出された方向が予め定めた方向である場合（方向２もしくは４ではない場合）（Ｓ５０７－Ｎｏ）は、処理をステップＳ５１０に進め、前ページとは不連続となる連続性変化開始ページであるとして、一時保管データ管理テーブルの「連続性開始フラグ」をｏｎとして記録する（図７の７０４を参照）。この処理では

、例えば、前ページまで連続した方向とは逆の方向（倒立方向）のページが検出されることになる。

【００４０】

ステップＳ５１１において、保管してある全てのスキャン画像の読み出しが終了していない場合（Ｓ５１１－Ｎｏ）、処理を再び、ステップＳ５０１に戻し、方向の連続性判断の処理を継続し、スキャン画像の読み出しが終了した場合（Ｓ５１１－Ｙｅｓ）は、データ保管部１０７に全ページのスキャン画像とともに一時保管データ管理テーブルとして、原稿方向の連続性に関する情報が一時保管されることになる。そして、データ保管部１０７の保管されたデータが保管データ管理部１０８を介して統合文書作成部１０９に転送される。

【００４１】

図５のフローチャートに従い、原稿方向連続性検出処理（図２のステップＳ２０２）が終了すると、統合文書作成部１０９により、次のステップＳ２０３の統合文書作成処理が実行される。統合文書は、「ページを表すデータ」と「ページに関連する情報」を階層的に一つの文書として保持する文書であり、画像入力部１０４によって入力されたスキャン画像を「ページを表すデータ」とし、原稿方向連続性の情報に基づくページに関連する情報とスキャン画像とを種類毎に階層化した１つの文書として保持する統合文書が作成される。本実施形態において、階層の区分としては原稿の連続性の区切りとなる「章」とし、章を構成する「ページ」の帰属に関する処理は図８のフローチャートに従い統合文書作成部１０９により処理される。原稿方向の違いは、「章」として区別された階層の違いとして処理される。

【００４２】

以下、ステップＳ２０３の統合文書作成処理の流れを図８のフローチャートに従い説明する。まず、ステップＳ８０１において、階層の区分である「章」の番号Ｎを０で初期化する。次に、一時保管してあるスキャン画像を、データ保管部１０７から最初のページから読み出す。原稿入力処理部１０１と印刷設定編集部１０２とがネットワークを介して接続される場合には、ステップＳ８０２の処理はネットワークを介して原稿入力処理部１０１から転送されてくるデータ保管部に格納されたデータを受信し読み込む処理に相当する。そして読み出したページに関連する情報として、一時保管データ管理テーブル（図７）を参照し（Ｓ８０３）、連続性開始フラグがｏｎであるか否かを判定する（Ｓ８０４）。ステップＳ８０４の判断で、読み出したページの連続性開始フラグがｏｎであれば（Ｓ８０４－Ｙｅｓ）、処理をステップＳ８０５に進め、そのページは新しい章の始まりであると判断し、章番号Ｎに１を加算する。

【００４３】

次に、処理をステップＳ８０６に進め、そのページを表すスキャン画像の原稿方向が正立か否か、すなわち方向＝１（図６）であるか否かを判定する。ステップＳ８０６の判断で原稿方向が正立の場合（Ｓ８０６－Ｙｅｓ）は処理をステップＳ８０８に進める。一方、スキャン画像が正立以外の場合は（Ｓ８０６－Ｎｏ）、正立になるような角度の回転処理を行うことによりスキャン画像を正立に変換する（Ｓ８０７）。

【００４４】

ステップＳ８０８では、統合文書の第Ｎ章の階層を構成するページデータとしてスキャン画像を登録する。第Ｎ章の階層に係る情報は例えば統合文書の構造を示すＸＭＬにより実現される。ここで、図９はステップＳ８０２において読み取られたスキャン画像に基づき作成された統合文書９０１としてのドキュメント情報の構成を示す図であり、第Ｎ章の階層９０２にはページ情報の階層９０３が含まれ、このページ情報の階層９０３には、ページ情報に対応するスキャン画像がスキャンページデータ９０４として登録される。そしてこの登録処理において登録された統合文書９０１が後述する図１１の表示に反映される。

【００４５】

説明を図８に戻し、全てのスキャン画像が読み出され、統合文書へのページ情報とスキ

ャン画像の登録が終了したら（Ｓ８０９－Ｙｅｓ）、処理を終了し、ステップＳ８０９の判断でページデータの登録が終了していない場合（Ｓ８０９－Ｎｏ）は、処理を再び、ステップＳ８０２に戻し、データの読み出し、統合文書の作成処理を続行する。図８のフローチャートに従ったステップＳ２０３の統合文書作成処理が終了すると、次に、処理をステップＳ２０４の印刷設定処理に進め、印刷設定制御部１１６が印刷設定編集を図１０のフローチャートに基づいて行う。

【００４６】

説明を図１０に移し、まず、ステップＳ１００１において、印刷設定制御部１１６は表示部１１０上に入力されたスキャン画像をプレビュー表示する（図１１を参照）。このステップの段階では、印刷設定情報であるジョブチケットはデフォルト設定を取っている。

【００４７】

ここで、図１１は、登録されているスキャン画像に基づく統合文書９０１をレイアウトして表示する部分１１０１と、章に帰属するページの区分け（章分割）表示部分１１０２とを組み合せた表示部１１０上のプレビュー表示画面を例示する図であり、オペレータはこの表示画面を参照しながら、章分割された結果を確認することができ、この表示を参照しながらオペレータは、以下に説明するページレイアウトの編集等を行うことができる。

【００４８】

次に、ステップＳ１００２において、オペレータはページレイアウト編集部１１１において、ページレイアウトの編集（例えば、章の登録されているページの順番を入れ替えたり、アノテーションやページ番号等を設定したり、ページが登録されている章を変更する等）を行ない、ステップＳ１００３において、オペレータは印刷指示編集部１１２にて、印刷デバイスへの指示の編集（例えば、印刷部数を設定したり、ステープルやパンチ等の処理を設定したりするなど）を行うことができる（図１２）。ここで、図１２は、印刷設定編集時の画面表示の例を示す図であり、スキャン画像のレイアウト表示１２０１と区分け（章分割）された章に帰属するページの階層構造の表示１２０２に加え、スキャン画像を印刷するための詳細設定を指示するための入力画面１２０３が表示される。

【００４９】

ステップＳ１００３の印刷指示編集処理が終了すると、処理を次のステップＳ１００４に進め、印刷設定情報保存部１１３にて、印刷設定情報（ページレイアウト編集の情報や、印刷指示編集の情報）をジョブチケットとして保存し、図１０のフローチャートの処理を終了する。ここで、「ジョブチケット」とは、電子化原稿を構成するページを最小単位とする構造を有する中間データである。

【００５０】

説明を図２に戻し、ステップＳ２０４の印刷設定処理が終了すると、処理をステップＳ２０５に進め、データ変換部１１４は印刷データ変換処理を実行する。尚、ステップＳ２０４の印刷設定を含む、ステップＳ２０３において作成された統合文書９０１を所定の記憶部に保持し、別のタイミングでステップＳ２０５の印刷データ変換処理を行なわせるようにしても良い。印刷データ変換部１１４は、ジョブチケットの印刷設定情報と各ページデータとを印刷データ出力部１０３が出力可能な印刷データに変換し、印刷データ送信部１１５は、変換された印刷データを印刷データ出力部１０３に送信し（Ｓ２０６）、印刷データ出力部１０３は送信された印刷データに基づき印刷する（Ｓ２０７）。

【００５１】

以上説明したとおり、本実施形態によれば、複数種類の原稿を入力する場合、電子化文書を原稿方向の連続性に基づいて区分け（章分割）した階層構造としてまとめ、一冊の原稿として印刷、編集設定することが可能になる。

【００５２】

＜第２実施形態＞

第１実施形態における原稿方向の検出のように、予め定めた原稿方向以外は原稿方向検出に誤りがあると判断する代わりに、原稿方向検出の確度（確からしさ）によって原稿方向が連続か非連続かを判定することも可能である。

【００５３】
また、統合文書作成（図２のステップＳ２０３）の際に、「方向検出確度」や、「方向修正フラグ」等の情報を、更に統合文書内のテーブルに保持してもよい。これにより、原稿方向の連続性検出処理の結果、原稿が連続か非連続か不確実である場合に、オペレータの判断により入力原稿におけるページ検出の結果を修正するための情報をオペレータに可視化して報知することが可能になる。以下、この具体的な内容を図１３のフローチャートに従い説明する。尚、本実施形態では、既に原稿入力処理（図２のステップＳ２０１）により、データ保管部１０７に原稿の全てのページを読み取ったスキャン画像が一時保管されているものとする。

【００５４】
まず、ステップＳ１３０１において、一時保管してあるスキャン画像を最初のページから読み出す。原稿入力処理部１０１と印刷設定編集部１０２とがネットワークを介して接続される場合には、ステップＳ８０２の処理はネットワークを介して原稿入力処理部１０１から転送されてくるデータ保管部に格納されたデータを受信し読み込む処理に相当する。そして、次の原稿方向検出処理においてページの原稿方向（図６）を検出する（Ｓ１３０２）。ステップＳ１３０２の処理は、第１実施形態で説明した（１）式に従い、最も高い平均類似度ｋ(n)を求め、最大の平均類似度ｋ(n)をもつ方向を原稿の方向として検出し、検出した原稿方向を一時保管データ管理テーブル（図１４を参照）の「検出方向」に記録し、スキャン画像と同じデータ保管部１０７に保管する。原稿方向の検出の際に得られる文字認識処理結果により、原稿方向連続性検出処理部１０６は、各方向（図６の１～４）の類似度から原稿方向の確からしさを確率的に評価するパラメータ（方向検出確度）を算出し、これを一時保管データ管理テーブル（図１４）の「方向検出確度」（１４０５）に記録し、スキャン画像と同じデータ保管部１０７に保管する（Ｓ１３０３）。

【００５５】
原稿方向に関する方向検出確度の算出方法として、最も高い平均類似度 ｋ(a)と２番目に高い平均類似度ｋ(b)をそれぞれデータ保管部１０７に一時的に保管しておき、それらの差をδとし、

方向検出確度　K=ａδ　（ａは係数）　　・・・（２）

（２）式に従って方向検出確度を算出する。

【００５６】
次に、一時保管データ管理テーブルに記録した原稿方向（「検出方向」）から、原稿を構成する各ページの連続性を検出するために、原稿方向連続性検出処理部１０６は連続性開始フラグを設定するための処理を行う。ステップＳ１３０４において、最初の原稿のページか否かを判断する。原稿の最初のページである場合（Ｓ１３０４－Ｙｅｓ）は、ステップＳ１３１１に処理を進め「連続性開始フラグ」をｏｎとして登録する（図１４の１４０１を参照）。

【００５７】
一方、ステップＳ１３０４の判断で、原稿の最初のページでない場合（Ｓ１３０４－Ｎｏ）は、処理をステップＳ１３０５に進める。ステップＳ１３０５では、ステップＳ１３０２の処理により検出された原稿方向としてフラグに設定されているデータに基づき原稿方向の変化の有無を判断する。原稿方向としてフラグに設定されているデータが、前ページにおけるスキャン画像の原稿方向としてフラグに設定されているデータと同一である場合（Ｓ１３０５－Ｎｏ）は連続性に変化は無いとして、処理をステップＳ１３０６に進め、一時保管データ管理テーブル（図１４）の「連続性開始フラグ」をｏｆｆとして記録する（Ｓ１３０６）。

【００５８】
一方、ステップＳ１３０５の判断で、前ページにおけるスキャン画像の原稿方向としてフラグに設定されているデータと異なる場合（Ｓ１３０５－Ｙｅｓ）は、処理をステップＳ１３０７に進める。ステップＳ１３０７では、原稿方向としてフラグに設定されているデータが方向＝０、すなわち白紙か否かを判定し、白紙（方向＝０）である場合（Ｓ１３

０７－Ｙｅｓ）は、処理をステップＳ１３０９に進め、強制的に前ページの原稿方向と連続しているとして、前データの原稿方向と同じ方向を示すデータを一時保管データ管理テーブル（図１４）の「最終的方向」に記録する（Ｓ１３０９、図１４の１４０３を参照）。

【００５９】

ステップＳ１３１０では、当初の検出方向（白紙の場合はフラグの値を「０」として設定された結果による）と異なる方向が最終的に設定された履歴として「方向修正フラグ」をｏｎと記録する（図１４の１４０５を参照）。

【００６０】

ステップＳ１３０７の判断で、白紙でない場合（Ｓ１３０７－Ｎｏ）は、処理をステップＳ１３０８に進め、方向検出確度が閾値ｍ未満か否かにより方向検出処理結果の確実性を判定する。例えば、以下に示す例では、原稿方向連続性検出処理部１０６は、閾値ｍを６０として方向検出確度と比較をするが、本発明の趣旨は、閾値の値が６０に限定されるものでなく、閾値は任意に設定することができるパラメータであることは言うもでもない。図１４の「ページ３」のように、方向検出確度が６０未満のもの（１４０７）が検出されている場合、原稿方向連続性検出処理部１０６は、原稿方向の検出処理の結果が不確実であると判断し、強制的に前ページの原稿データと同一の方向を示すデータ（この場合は方向１）を一時保管データ管理テーブル（図１４）の「最終的方向」に記録し（Ｓ１３０９、図１４の１４０２を参照）、当初の検出方向（＝２）と異なる方向が最終的に設定されたことを示す履歴として「方向修正フラグ」をｏｎとして記録する（Ｓ１３１０）。そして、ステップＳ１３０６において、一時保管データ管理テーブル（図１４）の「連続性開始フラグ」をｏｆｆとし、前ページとの間では連続性が維持された状態として検出結果が記録される。

【００６１】

また、ステップＳ１３０８において、方向検出確度が閾値ｍ以上である場合（Ｓ１３０８－Ｎｏ）、処理をステップＳ１３１１に進め、前ページとは不連続となる連続性変化開始ページであるとして、一時保管データ管理テーブルの「連続性開始フラグ」をｏｎとして記録する（図１４の１４０４を参照）。この処理では、例えば、前ページまで連続した方向とは逆の方向（倒立方向）のページが検出されることになる。

【００６２】

ステップＳ１３１２において、保管してある全てのスキャン画像の読み出しが終了していない場合（Ｓ１３１２－Ｎｏ）、処理を再び、ステップＳ１３０１に戻し、方向の連続性判断の処理を継続し、スキャン画像の読み出しが終了した場合（Ｓ１３１２－Ｙｅｓ）は、データ保管部１０７に全ページのスキャン画像とともに一時保管データ管理テーブルとして、原稿方向の連続性に関する情報が一時保管されることになる。図１３のフローチャートに従い、原稿方向連続性検出処理（図２のステップＳ２０２）が終了すると、統合文書作成部１０９により、ステップＳ２０３の統合文書作成処理が、図１５のフローチャートに従い実行される。以下、図１５を参照しつつ統合文書作成処理を説明するが、図１５のステップＳ８０１～ステップＳ８０７の処理に関しては、図８において対応する各ステップと処理内容が重複するのでここでは説明を省略する。

【００６３】

ステップＳ８０６の判断で、スキャン画像の原稿方向が正立の場合、または、ステップＳ８０７の処理により正立に変換されたスキャン画像に対し、ステップＳ１５０８では、統合文書の第Ｎ章の階層を構成するページデータとしてスキャン画像を登録する。ここで、図１６は第２実施形態において生成される統合文書１６０１としてのドキュメント情報の構成を示す図であり、第Ｎ章の階層１６０２にはページ情報の階層１６０３が含まれ、このページ情報の階層１６０３には、ページ情報に対応するスキャン画像がスキャンページデータ１６０４として登録される。この時、ページ情報の階層１６０３にはスキャンページデータ１６０４に対応する原稿方向関連情報１６０５として、「方向修正フラグ」ならびに「方向検出確度」が登録される（Ｓ１５０９）。更に、白紙のページが検出されて

いる場合は、この情報も原稿方向関連情報１６０５として登録してもよい。

【００６４】

以上のようにして、保管してある全てのスキャン画像の読み出しと統合文書へのページ情報、スキャン画像の登録が終了したら（Ｓ１５１０－Ｙｅｓ）、統合文書作成処理を終了し、ステップＳ１５１０の判断で、ページデータの登録が終了していない場合（Ｓ１５１０－Ｎｏ）は、処理を再びステップＳ８０２に戻し、データの読み出し、統合文書の作成処理を続行する。

【００６５】

図１５のフローチャートに従った統合文書作成処理が終了すると、印刷設定制御部１１６は印刷設定編集を図１０のフローチャートに基づいて行う（具体的な処理の流れは、第1実施形態における説明と重複するので、ここでは説明を省略する）。この際、ステップＳ１００１のプレビュー表示処理において、表示部１１０上には、章に帰属するページの区分け（章分割）表示部分１７０２とともに表示されるスキャン画像のレイアウト表示１７０１(図１７)に対して、入力されたスキャン画像の章分割における区分け（章分割）の結果をオペレータが視覚的に把握できるような識別情報１７０３が表示制御部１１７により付加される。表示制御部１１７により付加される識別情報１７０３の表示には原稿方向関連情報１６０５として統合文書１６０１に登録されている、方向修正フラグまたは方向検出確度が用いられる。

【００６６】

例えば、印刷設定制御部１１６は、検出された方向を強制的に修正したことを示す方向修正フラグが「on」になっているページに対しては該当ページの外枠を黄色の太枠で囲んで示したり、方向検出確度が閾値ｍ未満であることを示すために、該当ページの外枠を赤の太枠により表示したりすることで、オペレータが原稿方向の検出結果を容易に識別できるように表示制御する。あるいは、表示制御部１１７は、オペレータが認識しやすいように警告を報知するためのダイアログメッセージを表示部１１０上に表示制御することも可能である。

【００６７】

表示制御部１１７は、原稿方向関連情報１６０５に含まれる白紙の情報に関しても同様に、ページの方向の検出処理に関して、オペレータの注意を喚起するようにページの外枠の色を変えるようにすることもできる。

【００６８】

オペレータはこの表示に基づいて検出結果が妥当か否かを判断することができ、修正が必要な場合は、変更対象となる画面のレイアウト表示を不図示の入力手段を介して指定して、ページレイアウトの編集（例えば、章に登録されているページの順番を入れ替えたり、スキャン画像がページとして登録されている章を変更する等）や、対応するページ情報（９０３、１６０３等）を更新することができる。また、ユーザは修正が必要ないと判断した場合、あるいは、ページレイアウトの編集により誤った検出結果が解消されたと判断する場合は、入力手段を介してその内容を確定させることができる。

【００６９】

以上説明したとおり、本実施形態によれば、複数種類の原稿を入力する場合、電子化文書を原稿方向の連続性に基づいて区分け（章分割）した階層構造としてまとめ、一冊の原稿として印刷、編集設定することが可能になる。あるいは、階層構造にまとめる際、誤認識の可能性のある原稿を特定するための識別表示をすることで、誤認識を修正するための操作を容易にすることが可能になる。

【００７０】

＜第３実施形態＞

上述の第１実施形態においては、原稿方向の連続性を検出する方法として、予め定められた方向に原稿方向が該当するか否を判断し、予め想定された方向に該当しない場合は、強制的に検出方向を前ページの原稿方向と同じ方向に設定していた。また、第２実施形態においては、方向毎の方向検出確度をパラメータとして、所定の閾値ｍ未満の場合は、強

制的に前ページと同じ方向に原稿方向を設定していた。

【００７１】

本実施形態では、原稿方向の連続性を判断する処理として、原稿方向の検出対象となるＸ（Ｐ）番目のページに関して、その前後のページ（Ｘ（Ｐ－１）ページ、Ｘ（Ｐ＋１）ページ）の原稿方向と、前後のページに関する方向検出確度とを考慮した原稿方向の連続性の検出処理を説明する。尚、本実施形態において原稿方向連続性検出処理の前提となる第１実施形態で説明した原稿方向入力処理（図２のステップＳ２０１）、平均類似度の算出、第２実施形態で説明した方向検出確度の算出等に関しては説明が重複するので詳細な説明は省略する。

【００７２】

図１８Ａ、Ｂは、第３実施形態の処理の流れを説明するフローチャートである。まず、ステップＳ１８０１で、データ保管部１０７に一時保管されているスキャン画像を最初のページから一旦全て読み出す。原稿入力処理部１０１と印刷設定編集部１０２とがネットワークを介して接続される場合には、ステップＳ１８０１の処理はネットワークを介して原稿入力処理部１０１から転送されてくるデータ保管部に格納されたデータを受信し読み込む処理に相当する。そして、ステップＳ１８０２で原稿の向き（原稿方向）を検出するための検出処理を原稿方向検出処理部１０５が実行する。この原稿方向の検出は、第１実施形態でも説明したように、原稿を構成する各ページ単位に方向が定められる（図６を参照）。原稿方向検出処理部１０５は、原稿の向きが方向１～４の何れに対応するかを検出するために、原稿中に含まれる１つの文字に該当する画像を切り出し、その画像とデータ保管部１０７に保管されている認識辞書の文字とのマッチングの度合いを評価し、最も高い類似度の認識結果を与える方向（１～４）を原稿の方向として特定する。そして、ステップＳ１８０１で読み出した保管データの全てに関して原稿方向検知処理を終えたら処理をステップＳ１８０３に移行する。

【００７３】

ステップＳ１８０３において、最初の原稿のページか否かを判断する。原稿の最初のページである場合（Ｓ１８０３－Ｙｅｓ）は、ステップＳ１８１１に処理を進め全ての保管データに対する処理が終了したか否かを判断する。ステップＳ１８１１の処理で全保管データの処理が終了した場合（Ｓ１８１１－Ｙｅｓ）は処理を図１８Ｂのステップ１８２１に進める。ステップＳ１８１１の判断で全保管データに対する処理が終了していないと判断した場合（Ｓ１８１１－Ｎｏ）、処理を再び、ステップＳ１８０１に戻し、一時保管データの読み取り処理から開始して、ページの方向の連続性に関する判定処理を続行する。ステップＳ１８０３の判断で、原稿の最初のページでないと判断された場合（Ｓ１８０３－Ｎｏ）は、処理をステップＳ１８０４に進める。

【００７４】

ステップＳ１８０４では、ステップＳ１８０２の処理により検出された原稿方向としてフラグに設定されたデータに基づき原稿方向の変化の有無を判断する。ここで、原稿方向の検出対象のページをＸ（Ｐ）ページとすると、その前後のページＸ（Ｐ－１）、Ｘ（Ｐ＋１）に関する原稿方向として設定されているフラグの値（1～４のいずれか）をデータ保管部１０７から読み出し、前後のページ原稿方向を比較する。原稿方向として設定されているフラグのデータが、前後のページにおけるスキャン画像の原稿方向としてフラグに設定されているデータと同一である場合（Ｓ１８０４－Ｙｅｓ）は、ページの方向の連続性に変化は無いとして、処理をステップＳ１８０５に進める。全ての保管データに対してページの方向の連続の判定処理が終了した場合（Ｓ１８１１－Ｙｅｓ）は処理を図１８Ｂのステップ１８２１に進める。ステップＳ１８１１の判断で読み込んだ全保管データに対する処理が終了していないと判断した場合（Ｓ１８１１－Ｎｏ）、処理を再び、ステップＳ１８０３に戻し、孤立フラグの判定処理を読み込んだ全ての保管データに対して行なうまでステップＳ１８０３から１８０８の処理を繰り返す。

【００７５】

一方、ステップＳ１８０４の判断で、前後ページにおけるスキャン画像の原稿方向とし

てフラグに設定されているデータと異なる場合（Ｓ１８０４－Ｎｏ）は、処理をステップＳ１８０６に進める。

【００７６】

この場合、例えば、Ｘ（Ｐ－１）ページ＝方向１
Ｘ（Ｐ）ページ＝方向３
Ｘ（Ｐ＋１）ページ＝方向１

のように、検出対象となるＸ（Ｐ）ページの前後のページが同じ原稿方向（＝１）であり、検出対象となっているＸ（Ｐ）ページ目は倒立方向（＝３）と認識されている場合（前後のページの原稿方向が同一方向で、対象となるＸ（Ｐ）ページ目の原稿方向が前後のページの原稿方向と異なる状態を「孤立」という）、原稿方向連続性検出処理部１０６は、Ｘ（Ｐ）ページ目が誤認識されている可能性があると判断し、一時保管データ管理テーブル（図１９）の、孤立ページを特定するための「孤立フラグ」（１９１０）の設定をｏｎとして記録する（Ｓ１８０８、図１９の１９１１を参照）。

【００７７】

また、検出対象となるＸ（Ｐ）ページ目が前後のページに対して孤立していない場合（前後のページＸ（Ｐ－１）、Ｘ（Ｐ＋１）の原稿方向が異なる場合）（Ｓ１８０６－Ｎｏ）、処理をステップＳ１８０７に進める。

【００７８】

例えば、Ｘ（Ｐ－１）ページ＝方向１
Ｘ（Ｐ）ページ＝方向３
Ｘ（Ｐ＋１）ページ＝方向３

のような場合、Ｘ（Ｐ）ページは前後のページに対して孤立していないので、原稿方向連続性検出処理部１０６は孤立フラグをｏｆｆとして登録する（Ｓ１８０７）。ステップＳ１８１１の判断で、全ページに対して孤立フラグのｏｎ/ｏｆｆ設定が終了した場合（Ｓ１８１１－Ｙｅｓ）、図１８Ａの処理を終了する。この場合、すくなくとも図１９の一時保管データ管理テーブルを保持したまま、図１８ＢのステップＳ１８１２に処理を進め、孤立フラグがｏｎとして登録されているページが連続するか否かを検査する。

【００７９】

例えば、各ページの方向が、
Ｘ（Ｐ－１）＝方向１
Ｘ（Ｐ）＝方向３（孤立フラグｏｎ）
Ｘ（Ｐ＋１）＝方向１

という設定の場合、孤立フラグがｏｎになっているのは、Ｘ（Ｐ）ページ目だけであり、この場合は孤立フラグの記録は不連続として判断され、処理はステップＳ１８１３に進められる。

【００８０】

ステップＳ１８１３において、不連続な孤立フラグが記録されているページ（Ｘ（Ｐ））の原稿方向を、強制的に前後ページの原稿方向と連続していると判断し、一時保管データ管理テーブル（図１９）の「連続性開始フラグ」をｏｆｆとして修正して記録する。その際、前後ページの原稿方向と同じ方向を示すデータ（この場合は方向１）を一時保管データ管理テーブルの「最終的方向」に記録する（図１９の１９０２を参照）。

【００８１】

ステップＳ１８１４では、当初の検出方向（＝３）が修正され、これと異なる方向（＝１）が最終的に設定された履歴として「方向修正フラグ」をｏｎとして記録する（図１９の１９０６を参照）。

【００８２】

一方、ステップＳ１８１２において、孤立フラグがｏｎとして記録されているページが連続する場合（Ｓ１８１２－Ｙｅｓ）、例えば、各ページが、
Ｘ（Ｐ－１）＝方向１
Ｘ（Ｐ）＝方向３（孤立フラグｏｎ）

Ｘ（Ｐ＋１）＝方向１（孤立フラグｏｎ）
Ｘ（Ｐ＋２）＝方向３
という設定の場合、孤立フラグがｏｎになっているのは、Ｘ（Ｐ）ページとＸ（Ｐ＋１）ページであり、この場合は孤立フラグがｏｎとして記録されているページは２連続として判断され、処理はステップＳ１８１７に進められる。

【００８３】

ステップＳ１８１７では、孤立フラグがｏｎとして記録されているページの連続が３以上か否かを判断し、連続するページが３以上である場合（Ｓ１８１７－Ｙｅｓ）は処理をステップＳ１８１５に進める。一方、連続するページが３以上でない場合（Ｓ１８１７－Ｎｏ）は処理をステップＳ１８１８に進める。ステップＳ１８１８では、原稿方向連続性検出処理部１０６は、孤立フラグが連続する各原稿ページの方向検出確度１９０５を、一時保管データ管理テーブル（図１９）から読み出して比較をし、ステップＳ１８１９において、ステップＳ１８１８の処理に基づく比較結果に従い、方向検出確度の小さい方のページの原稿方向を、その前後のページと同一原稿方向に強制的に修正する。その際、原稿方向連続性検出処理部１０６は、前後ページの原稿方向と同じ方向を示すデータを一時保管データ管理テーブルの「最終的方向」に記録する。この場合、Ｘ（Ｐ）ページの方向検出確度をｋ１、Ｘ（Ｐ＋１）ページの方向検出確度をｋ２とし、両者の関係が、

ｋ１＜ｋ２　　・・・（３）

の場合は、方向検出確度の小さい方向ページＸ（Ｐ）の原稿方向が方向１として修正される。更にステップＳ１８１９では、当初の検出方向（＝３）が修正され、これと異なる方向（＝１）を最終的に設定した履歴として「方向修正フラグ」をｏｎとして記録する。

【００８４】

一方、ステップＳ１８１７の判断で、連続するページが３以上の場合、例えば、Ｘ（Ｐ－１）＝方向１
Ｘ（Ｐ）＝方向３（孤立フラグｏｎ）
Ｘ（Ｐ＋１）＝方向１（孤立フラグｏｎ）
Ｘ（Ｐ＋２）＝方向３（孤立フラグｏｎ）
Ｘ（Ｐ＋３）＝方向１（孤立フラグｏｎ）
Ｘ（Ｐ＋４）＝方向３
という設定の場合のように、原稿方向が１ページごとに異なるような場合は、ユーザが意図的に１ページ毎に不連続となるようにＡＤＦに原稿をセットしたものとして判断し、原稿方向の修正は行なわない。この場合は、各ページが不連続となる連続性変化開始ページであるとして、一時保管データ管理テーブルの各ページの「連続性開始フラグ」をｏｎとして記録し、各ページは各々異なる章に分割される。

【００８５】

ステップＳ１８１５において、全保管データに関する処理が終了しているか否かを判断し、終了していない場合（Ｓ１８１５－Ｎｏ）は処理をステップＳ１８２３に戻して、検査ページを次ページにセットして、孤立フラグの連続性を判断するステップＳ１８１２以降の処理を続行する。全保管データに対する処理が終了した場合（Ｓ１８１５－Ｙｅｓ）は、処理をステップＳ１８２２に進める。ステップＳ１８２２の処理は、図５で説明した原稿方向の連続性を検出処理であり、本ステップでは、図５のステップＳ５０７、Ｓ５０９を除く、各ステップが実行され、各ページに対する「連続性開始フラグ」のｏｎ／ｏｆｆが設定され、処理を終了する。尚、図５のフローチャートに対する各ステップに対する処理は既に第１実施形態において説明しているので、ここでは説明を省略する。

【００８６】

図１８Ａ、Ｂのフローチャートに従い、原稿方向連続性検出処理（図２のステップＳ２０２）が終了すると、統合文書作成部１０９により、次のステップＳ２０３の統合文書作成処理が実行される。以降の処理に関しては第１実施形態と内容が重複するので説明は省略する。

【００８７】

以上説明したとおり、本実施形態によれば、複数種類の原稿を入力する場合、電子化文書を原稿方向の連続性に基づいて区分け（章分割）した階層構造としてまとめ、一冊の原稿として印刷、編集設定することが可能になる。あるいは、本実施形態によれば、上述の実施形態では検出できない、倒立方向の原稿を正立方向に修正する処理を可能にし、より精度良く原稿の種類ごとに電子化文書を区分け（章分割）することが可能になる。

【００８８】

＜第４実施形態＞

上述各実施形態において説明したように、原稿方向の検出処理、原稿方向の連続性に基づいて区分け（章分割）した階層構造としてまとめ、一冊の原稿として印刷、編集設定をするための構成は、図１に示す原稿入力処理部１０１と印刷設定編集部１０２とに分割した構成に限定されるものではなく、例えば、第1実施形態で説明した原稿入力処理部１０１に設けられている原稿方向検出処理部１０５、原稿方向連続性検出処理部１０６、データ保管部１０７、保管データ管理部１０８の構成を印刷設定処理部１０２側に設けて、画像入力部１０４で取り込んだ原稿を構成するページ単位の原稿データに対して、印刷設定編集部１０２が、原稿の章を構成するページとして、原稿方向の検出処理、原稿方向の連続性に基づく区分け（章分割）編集処理等を実行するようにしてもよい。

【００８９】

尚、上述の各実施形態で示したフローチャートにおいて、その前後は処理が未完成にならない限り入れ替えることが可能である。また、本発明は、複数の機器(例えばホストコンピュータ、インタフェース機器、リーダ、プリンタなど)から構成されるシステムに適用しても、一つの機器からなる装置（複写機、プリンタ、ファクシミリ装置など）に適用してもよい。

【００９０】

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェア（図２、図３、図５、図８、図１０、図１３、図１５、図１８の手順）のプログラムコードを記憶した記憶媒体を、システムあるいは装置のコンピュータ（またはＣＰＵやＭＰＵ）が記憶媒体に格納されたプログラムコードを読出し実行することによっても、達成される。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【００９１】

プログラムコードを供給するための記憶媒体としては、例えば、フロッピー（登録商標）ディスク、ハードディスク、光ディスク、光磁気ディスク、ＣＤ－ＲＯＭ、ＣＤ－Ｒ、磁気テープ、不揮発性のメモリカード、ＲＯＭなどを用いることができる。

【００９２】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼動しているＯＳ（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれる。

【００９３】

さらに、記憶媒体から読出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるＣＰＵなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれる。

【図面の簡単な説明】

【００９４】

【図１】文書処理装置の概略的な構成を説明するブロック図である。

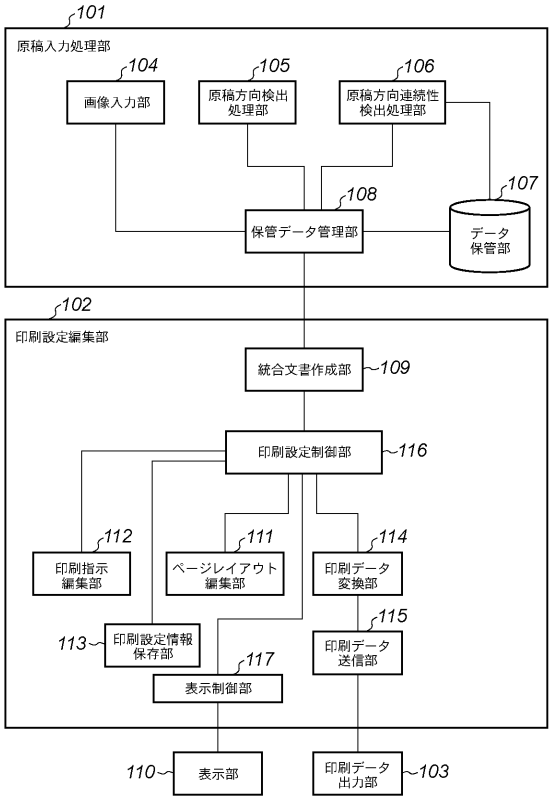
【図２】第１実施形態にかかる文書処理の流れを説明するためのフローチャートである。

【図３】原稿入力処理の具体的な流れを説明するフローチャートである。

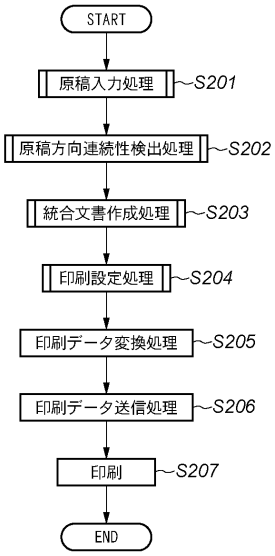
【図４】原稿Ａ５０１と原稿Ｂ５０２の２種類の原稿の入力を例示的に説明する図である

。
【図５】第１実施形態にかかる原稿方向の連続性を検出するための処理の流れを説明するフローチャートである。
【図６】原稿方向の類別を説明するための図である。
【図７】第１実施形態にかかる一時保管データ管理テーブルの内容を示す図である。
【図８】第１実施形態における統合文書作成処理の流れを説明するフローチャートである。
【図９】第１実施形態における統合文書の構成例を示す図である。
【図１０】印刷設定処理の流れを説明するフローチャートである。
【図１１】第１実施形態において作成された統合文書のプレビュー表示例を示す図である。
【図１２】印刷設定編集時の画面表示の例を示す図である。
【図１３】第２実施形態にかかる原稿方向の連続性を検出するための処理の流れを説明するフローチャートである。
【図１４】第２実施形態にかかる一時保管データ管理テーブルの内容を示す図である。
【図１５】第２実施形態にかかる統合文書作成処理の流れを説明するフローチャートである。
【図１６】第２実施形態において生成される統合文書の構成例を示す図である。
【図１７】第２実施形態において作成された統合文書のプレビュー表示例を示す図である。
【図１８Ａ】第２実施形態にかかる原稿方向の連続性を検出するための処理の流れを説明するフローチャートである。
【図１８Ｂ】第２実施形態にかかる原稿方向の連続性を検出するための処理の流れを説明するフローチャートである。
【図１９】第３実施形態にかかる一時保管データ管理テーブルの内容を示す図である。

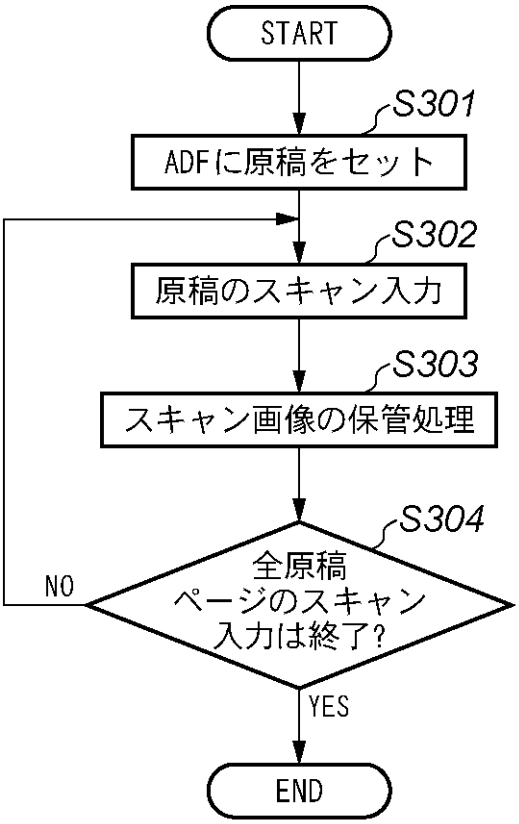
【図１】



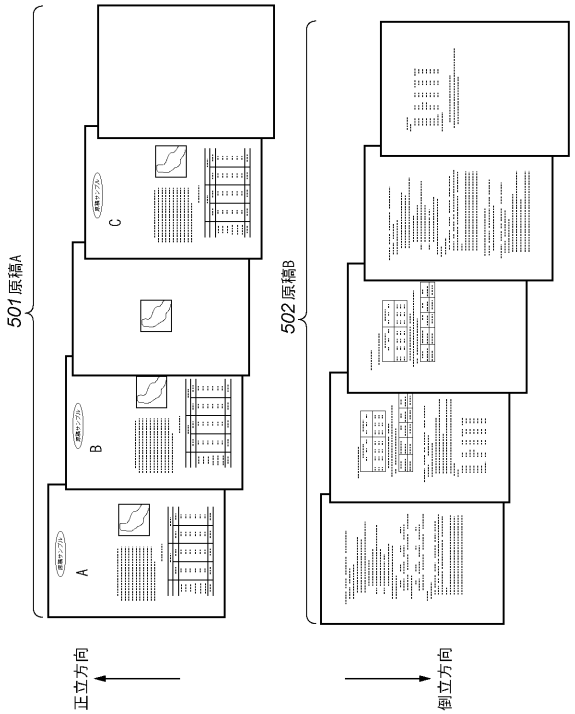
【図２】



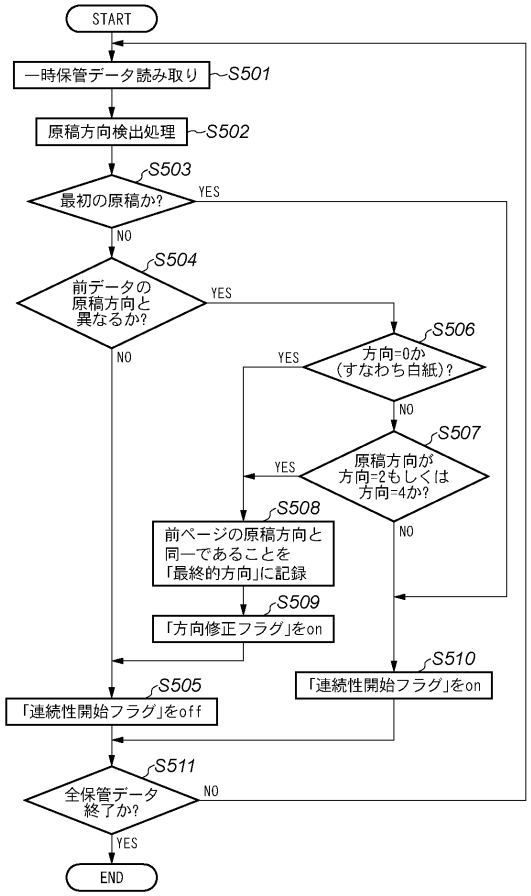
【図３】


START
S301
ADFに原稿をセット
S302
原稿のスキャン入力
S303
スキャン画像の保管処理
S304
全原稿
ページのスキャン
入力は終了?
NO
YES
END

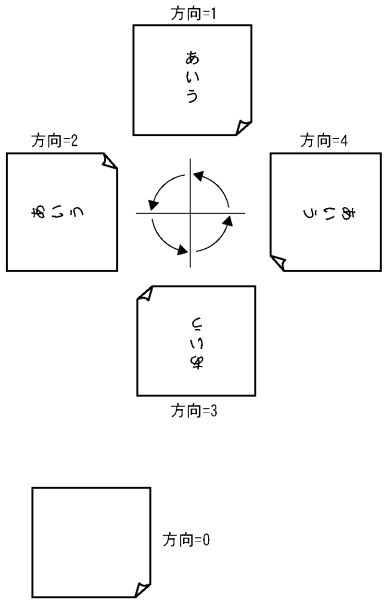
【図４】


501 原稿A
502 原稿B
正立方向
倒立方向

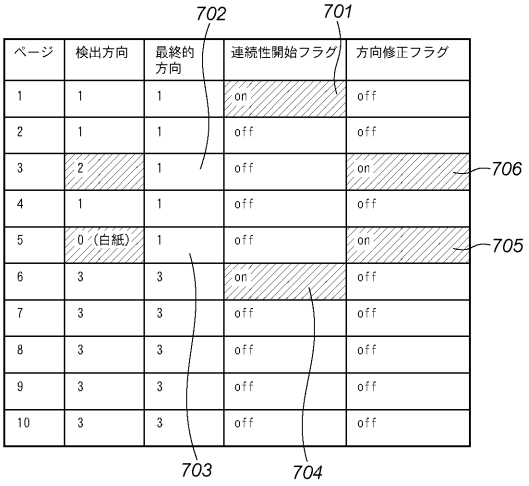
【図５】


START
一時保管データ読み取り
S501
原稿方向検出処理
S502
S503
最初の原稿か?
YES
NO
S504
前データの
原稿方向と
異なるか?
YES
NO
S506
方向=0か
(すなわち白紙)?
YES
NO
S507
原稿方向が
方向=2もしくは
方向=4か?
YES
NO
S508
前ページの原稿方向と
同一であることを
「最終的方向」に記録
S509
「方向修正フラグ」をon
S510
「連続性開始フラグ」をon
S505
「連続性開始フラグ」をoff
S511
全保管データ
終了か?
NO
YES
END

【図６】

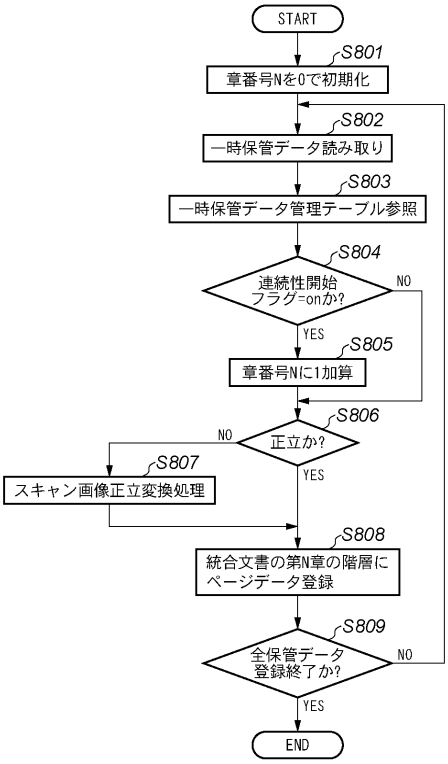

方向=1
あ
い
う
方向=2
方向=4
方向=3
方向=0

【図７】

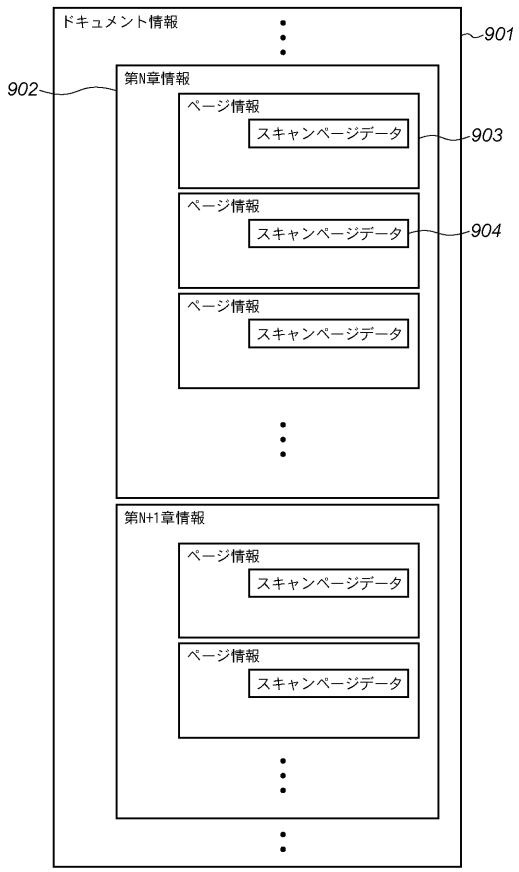


ページ	検出方向	最終的方向	連続性開始フラグ	方向修正フラグ
1	1	1	on	off
2	1	1	off	off
3	2	1	off	on
4	1	1	off	off
5	0（白紙）	1	off	on
6	3	3	on	off
7	3	3	off	off
8	3	3	off	off
9	3	3	off	off
10	3	3	off	off

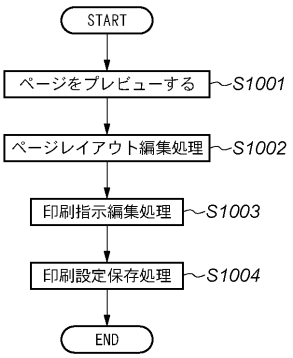
【図８】



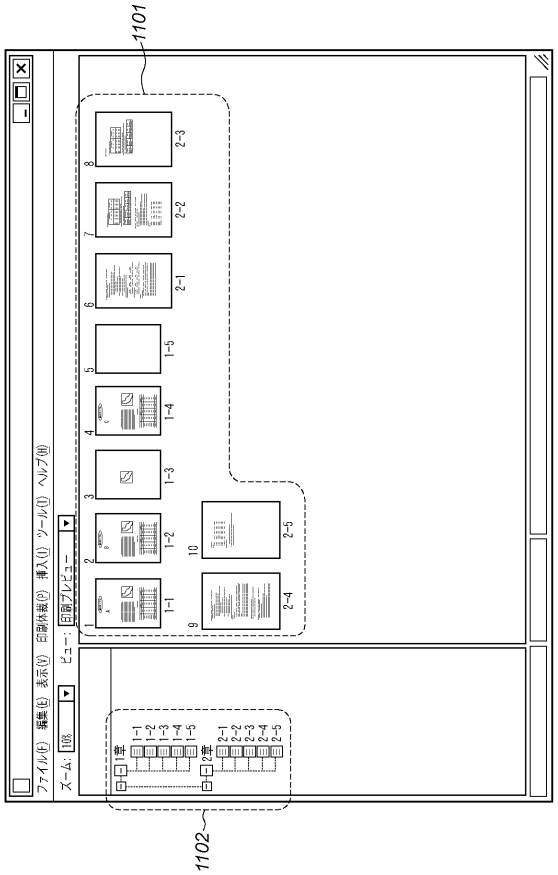
【図９】



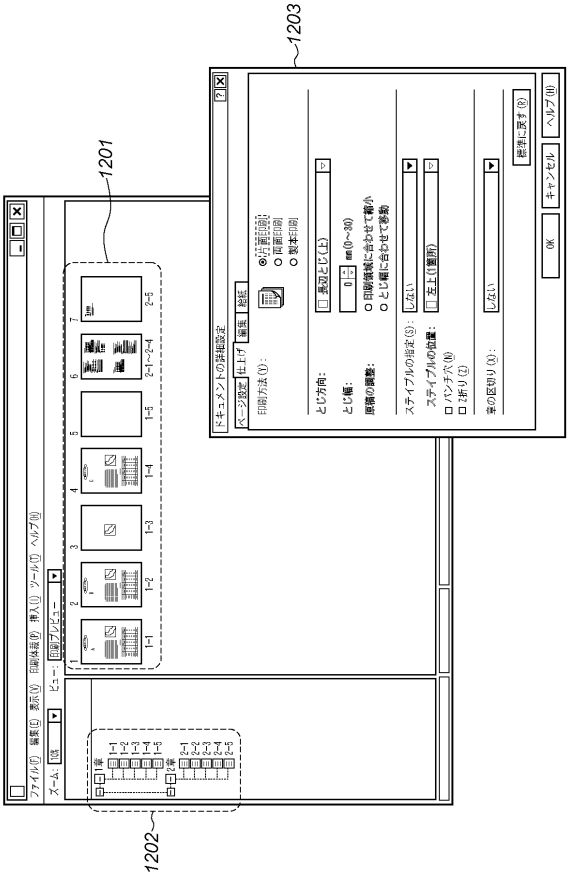
【図１０】



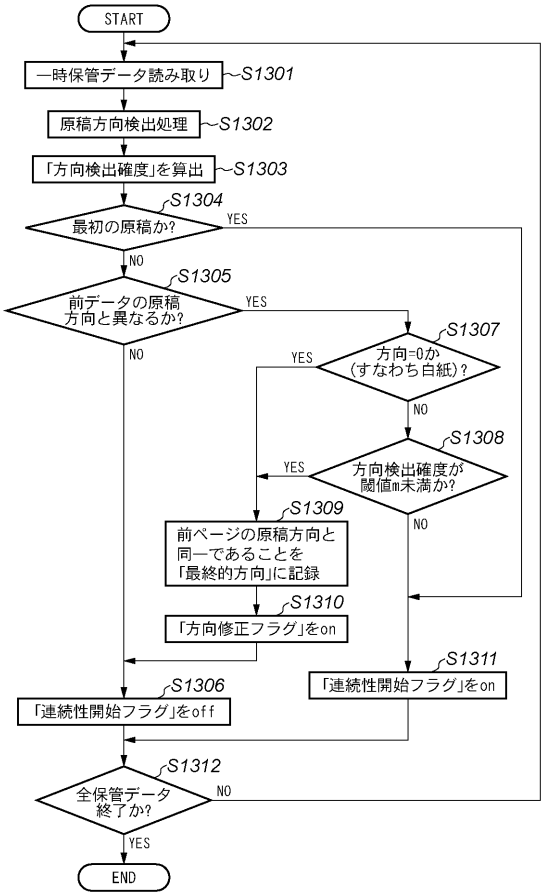
【図１１】



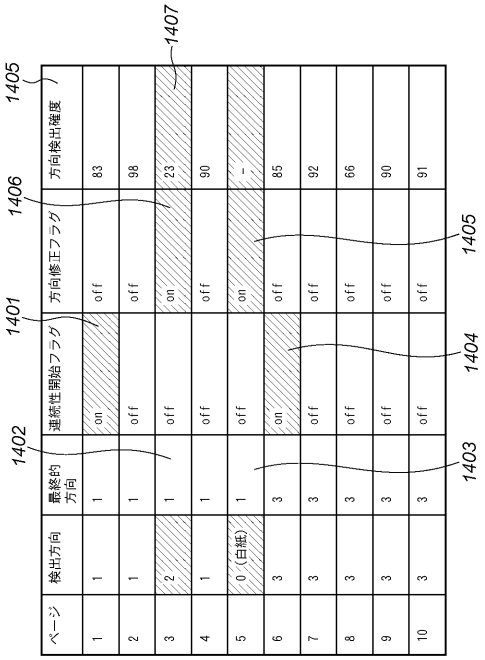
【図１２】



【図１３】

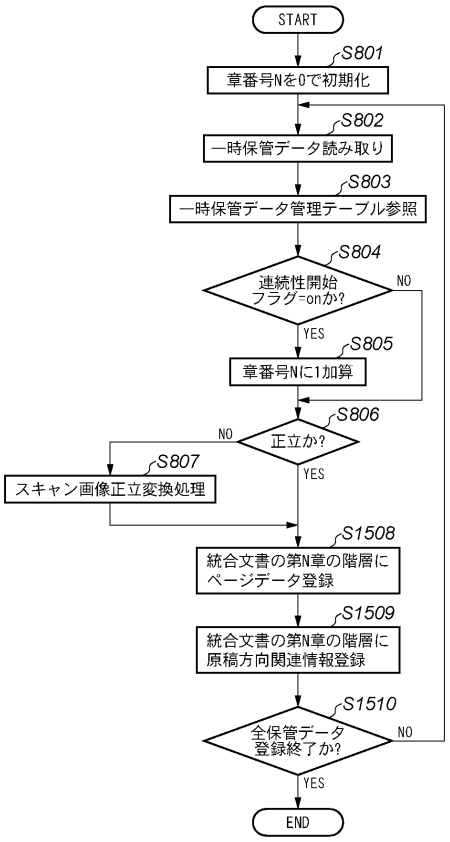


【図１４】

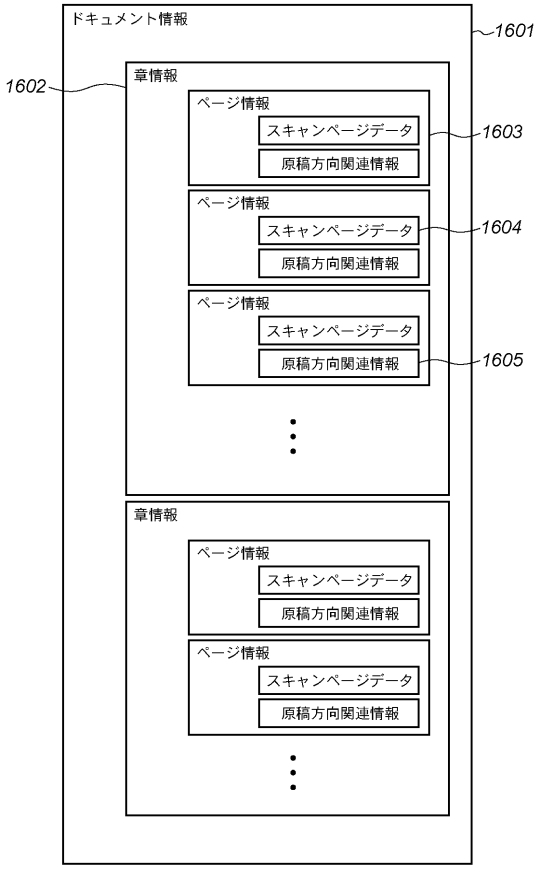


ページ	検出方向	最終的方向	連続性開始フラグ	方向修正フラグ	方向検出確度
1	1	1	on	off	83
2	1	1	off	off	98
3	2	1	off	on	23
4	1	1	off	off	90
5	0(白紙)	1	off	on	-
6	3	3	on	off	85
7	3	3	off	off	92
8	3	3	off	off	66
9	3	3	off	off	90
10	3	3	off	off	91

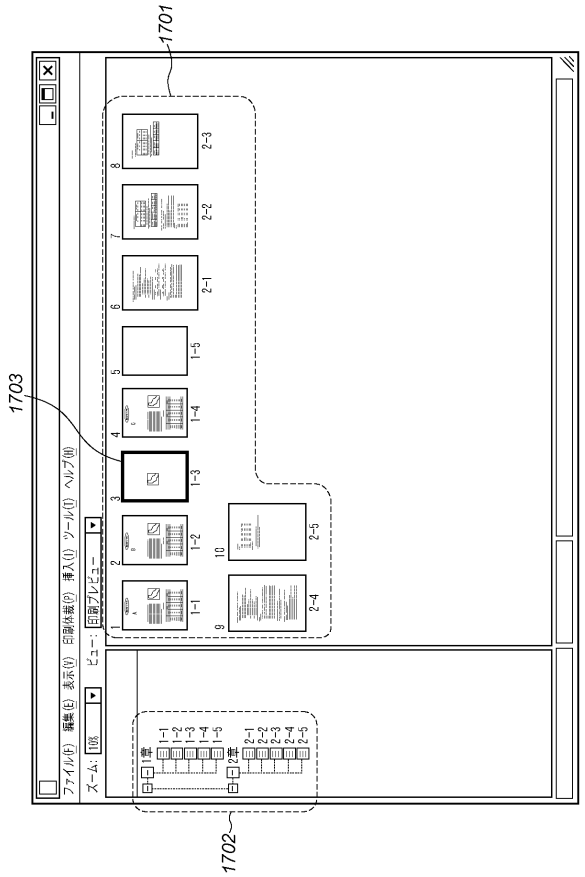
【図１５】



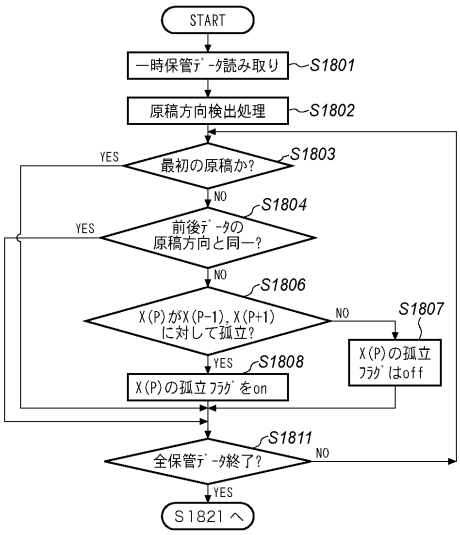
【図１６】



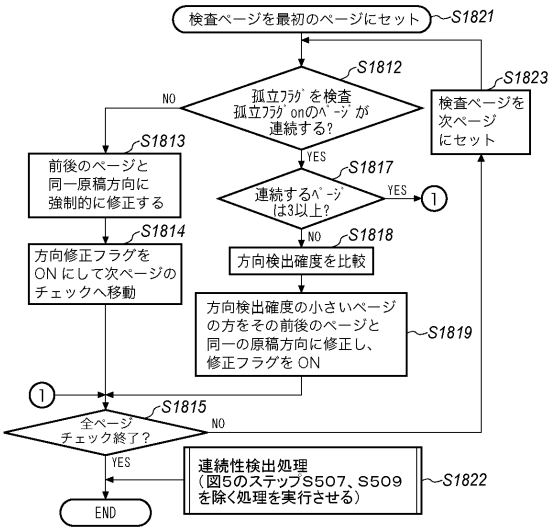
【図１７】



【図１８Ａ】



【図１８Ｂ】



【図１９】

1911

	1910 孤立フラグ	1905 方向検出確度	1906 方向修正フラグ	連続性開始フラグ	1902 最終的方向	検出方向	ページ
	off	83	off	on	1	1	1
	off	98	off	off	1	1	2
	on	23	on	off	1	3	3
	off	90	off	off	1	1	4
	off	-	on	off	1	0（白紙）	5
	off	85	off	on	3	3	6
	off	92	off	off	3	3	7
	off	66	off	off	3	3	8
	off	90	off	off	3	3	9
	off	91	off	off	3	3	10

フロントページの続き

審査官　千葉　久博

(56)参考文献　特開平１１－２１３１７１（ＪＰ，Ａ）
特開２００２－０４４３１８（ＪＰ，Ａ）
特開平１１－１２０３２１（ＪＰ，Ａ）
特開２００３－３０８２００（ＪＰ，Ａ）
特開２００２－０３２７６５（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｔ　　１／００
Ｇ０６Ｔ　　７／００，７／２０－７／６０
Ｈ０４Ｎ　　１／２１
Ｇ０６Ｆ　１７／３０