

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4492701号
(P4492701)

(45) 発行日 平成22年6月30日 (2010. 6. 30)

(24) 登録日 平成22年4月16日 (2010. 4. 16)

(51) Int. Cl. F I
G O 6 T 1 / 0 0 (2006. 01) G O 6 T 1 / 0 0 2 0 0 D

請求項の数 4 (全 15 頁)

<p>(21) 出願番号 特願2008-42 (P2008-42) (22) 出願日 平成20年1月4日 (2008. 1. 4) (65) 公開番号 特開2009-163475 (P2009-163475A) (43) 公開日 平成21年7月23日 (2009. 7. 23) 審査請求日 平成21年9月16日 (2009. 9. 16)</p>	<p>(73) 特許権者 000005496 富士ゼロックス株式会社 東京都港区赤坂九丁目7番3号 (74) 代理人 100075258 弁理士 吉田 研二 (74) 代理人 100096976 弁理士 石田 純 (72) 発明者 鈴木 昭宏 神奈川県川崎市高津区坂戸3丁目2番1号 K S P R & D ビジネスパークビル 富士ゼロックス株式会社内 審査官 千葉 久博</p>
--	---

最終頁に続く

(54) 【発明の名称】 情報処理装置、情報処理システム、および情報処理プログラム

(57) 【特許請求の範囲】

【請求項1】

ユーザの識別情報を取得する識別情報取得手段と、
ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段と、

1 又は複数の種類から成る複数ページの証憑が画像読取手段により読み取られて生成された複数ページの画像情報を取得する第1の画像情報取得手段と、

前記複数ページの証憑から選択された少なくとも1ページの証憑が画像読取手段により読み取られて生成された少なくとも1ページの画像情報を取得する第2の画像情報取得手段と、

前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて出力する出力手段と、

を有することを特徴とする情報処理装置。

【請求項2】

ユーザの識別情報を取得する識別情報取得手段と、
ユーザの識別情報と証憑情報とを関連付けて記憶する証憑情報記憶手段と、
前記証憑情報記憶手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段と、

10

20

1 又は複数の種類から成る複数ページの証憑を読み取って複数ページの画像情報を生成する第1の画像読取手段と、

前記複数ページの証憑から選択された少なくとも1ページの証憑を読み取って少なくとも1ページの画像情報を生成する第2の画像読取手段と、

前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて記憶する記憶手段と、

を有することを特徴とする情報処理システム。

【請求項3】

コンピュータを、

ユーザの識別情報を取得する識別情報取得手段、

ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段、

1 又は複数の種類から成る複数ページの証憑が画像読取手段により読み取られて生成された複数ページの画像情報を取得する第1の画像情報取得手段、

前記複数ページの証憑から選択された少なくとも1ページの証憑が画像読取手段により読み取られて生成された少なくとも1ページの画像情報を取得する第2の画像情報取得手段、

前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて出力する出力手段、

として機能させるための情報処理プログラム。

【請求項4】

コンピュータを、

ユーザの識別情報を取得する識別情報取得手段、

ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段、

1 又は複数の種類から成る複数ページの証憑を読み取って複数ページの画像情報を生成する第1の画像読取手段、

前記複数ページの証憑から選択された少なくとも1ページの証憑を読み取って少なくとも1ページの画像情報を生成する第2の画像読取手段、

前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて記憶手段に記憶させる手段、

として機能させるための情報処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置、情報処理システム、および情報処理プログラムに関する。

【背景技術】

【0002】

複数の文書原稿を読み取る前に文書間に予め特定の文書区切り用原稿を挿入しておくことで、イメージスキャナで連続的に原稿を読み取った場合でも、この文書区切り用原稿を認識し、原稿の画像を文書単位に登録する装置がある（例えば、特許文献1の「従来の技術」の欄を参照）。

【0003】

特許文献1には、イメージスキャナにより複数の文書の原稿を読み取るときに、文書区切り用原稿を用いることなく、各文書の区切りを判定して原稿画像を取得する装置が記載

10

20

30

40

50

されている。具体的には、複数の文書の原稿を読み取って画像データを生成し、すべての画像データから特徴量を算出し、当該特徴量に基づき文書区切りを判定する装置が記載されている。

【0004】

また、特許文献2には、複数の帳票画像を入力し、入力された複数の帳票画像を順番に登録帳票と比較していき、入力された帳票画像に類似する登録帳票の種別を識別した時点で、当該識別した登録帳票の構成枚数と同じ枚数分の上記入力された帳票画像を1組の帳票画像としてファイリングする装置が記載されている。

【0005】

【特許文献1】特開2002-24258号公報

【特許文献2】特開2005-115572号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

本発明は、文書区切り用原稿を使用せずに、複数ページの原稿から読み取られた画像情報を所望の箇所分割することが可能な情報処理装置、情報処理システム、または情報処理プログラ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】

【0007】

本発明に係る情報処理装置は、ユーザの識別情報を取得する識別情報取得手段と、ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段と、1又は複数の種類から成る複数ページの証憑が画像読取手段により読み取られて生成された複数ページの画像情報を取得する第1の画像情報取得手段と、前記複数ページの証憑から選択された少なくとも1ページの証憑が画像読取手段により読み取られて生成された少なくとも1ページの画像情報を取得する第2の画像情報取得手段と、前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて出力する出力手段と、を有することを特徴とする。

【0008】

本発明に係る情報処理システムは、ユーザの識別情報を取得する識別情報取得手段と、ユーザの識別情報と証憑情報とを関連付けて記憶する証憑情報記憶手段と、前記証憑情報記憶手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段と、1又は複数の種類から成る複数ページの証憑を読み取って複数ページの画像情報を生成する第1の画像読取手段と、前記複数ページの証憑から選択された少なくとも1ページの証憑を読み取って少なくとも1ページの画像情報を生成する第2の画像読取手段と、前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて記憶する記憶手段と、を有することを特徴とする。

【0009】

本発明に係る情報処理プログラムは、コンピュータを、ユーザの識別情報を取得する識別情報取得手段、ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段、1又は複数の種類から成る複数ページの証憑が画像読取手段により読み取られて生成された複数ページの画像情報を取得する第1の画像情報取得手段、前記複数ページの証憑から選択された少なくとも1ページの証憑が画像読取手段により読み取られて生成された少なくとも1ページの画像情報を取得する第2の画像情報取得手段、前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画

10

20

30

40

50

像情報を、当該証憑情報に関連付けて出力する出力手段、として機能させるためのものである。

【 0 0 1 0 】

本発明に係る情報処理プログラムは、コンピュータを、ユーザの識別情報を取得する識別情報取得手段、ユーザの識別情報と証憑情報とを関連付けて記憶する手段から、前記識別情報取得手段により取得されたユーザの識別情報に関連付けられた証憑の種類を示す証憑情報の一覧を取得する証憑情報取得手段、1又は複数の種類から成る複数ページの証憑を読み取って複数ページの画像情報を生成する第1の画像読取手段、前記複数ページの証憑から選択された少なくとも1ページの証憑を読み取って少なくとも1ページの画像情報を生成する第2の画像読取手段、前記複数ページの画像情報のうち、前記少なくとも1ページの画像情報が生成された前記少なくとも1ページの証憑の種類毎に分割された前記複数ページの画像情報を、当該証憑情報に関連付けて記憶手段に記憶させる手段、として機能させるためのものである。

10

【発明の効果】

【 0 0 1 1 】

請求項1に記載の発明によれば、文書区切り用原稿を使用せずに、複数ページの原稿から読み取られた画像情報を所望の箇所で分割することが可能となる。

【 0 0 1 2 】

請求項2に記載の発明によれば、文書区切り用原稿を使用せずに、複数ページの原稿から読み取られた画像情報を所望の箇所で分割することが可能となる。

20

【 0 0 1 3 】

請求項3に記載の発明によれば、文書区切り用原稿を使用せずに、複数ページの原稿から読み取られた画像情報を所望の箇所で分割することが可能となる。

【 0 0 1 4 】

請求項4に記載の発明によれば、文書区切り用原稿を使用せずに、複数ページの原稿から読み取られた画像情報を所望の箇所で分割することが可能となる。

【発明を実施するための最良の形態】

【 0 0 1 5 】

以下、本発明の実施の形態を図面に従って説明する。

【 0 0 1 6 】

[情報処理システムの構成]

図1は、本実施の形態に係る情報処理システム1の構成の一例を示すブロック図である。この情報処理システム1は、複数ページの原稿を読み取って複数ページの画像情報を生成し、当該複数ページの画像情報を分割して所定の記憶装置に記憶させるものである。情報処理システム1は、物理的に1つの装置により実現されてもよいし、複数の装置により実現されてもよい。一つの態様では、情報処理システム1は、スキャン機能、プリント機能、およびコピー機能などを有する複合機である。

30

【 0 0 1 7 】

図1において、情報処理システム1は、第1の画像読取部10、第2の画像読取部20、および情報処理装置30を有する。

40

【 0 0 1 8 】

第1の画像読取部10は、複数ページの原稿を読み取って複数ページの画像情報を生成する。第1の画像読取部10は、例えば、自動原稿送り装置により複数ページの原稿を順番に光学的に読み取って電子画像情報を生成するスキャナである。

【 0 0 1 9 】

上記複数ページの原稿は、例えば、複数の紙文書である。紙文書としては、例えば、証憑類（見積書、発注書、納品書、領収書、契約書、申込書など）が挙げられるが、特に限定されない。

【 0 0 2 0 】

第2の画像読取部20は、上記複数ページの原稿から選択された少なくとも1ページの

50

原稿を読み取って少なくとも1ページの画像情報を生成する。第2の画像読取部20は、例えば、自動原稿送り装置により少なくとも1ページの原稿を順番に光学的に読み取って電子画像情報を生成するスキャナである。なお、この他にも書画カメラやデジタルカメラなど、原稿情報を取り込むものであればスキャナに限定しない。

【0021】

上記少なくとも1ページの原稿は、複数ページの画像情報を分割する際の分割位置の基準となるものである。

【0022】

上記第1の画像読取部10および第2の画像読取部20は、単一のスキャナにより実現されてもよい。

10

【0023】

情報処理装置30は、第1の画像読取部10により生成された複数ページの画像情報を取得し、当該複数ページの画像情報を分割して出力する装置である。情報処理装置30は、情報処理システム1全体の制御を行ってもよい。情報処理装置30は、一つの態様では、ハードウェア資源とソフトウェアとの協働により実現され、例えばコンピュータである。具体的には、情報処理装置30の機能は、記録媒体に記録された情報処理プログラムがメインメモリに読み出されてCPU(Central Processing Unit)により実行されることによって実現される。上記情報処理プログラムは、CD-ROM等のコンピュータ読み取り可能な記録媒体に記録されて提供されることも可能であるし、データ信号として通信により提供されることも可能である。ただし、情報処理装置30の機能は、ハードウェアのみにより実現されてもよい。

20

【0024】

情報処理装置30は、第1の画像情報取得部31、第2の画像情報取得部32、特定部33、および出力部34を有する。

【0025】

第1の画像情報取得部31は、複数ページの原稿が第1の画像読取部10により読み取られて生成された複数ページの画像情報を取得する。

【0026】

第2の画像情報取得部32は、複数ページの原稿から選択された少なくとも1ページの原稿が第2の画像読取部20により読み取られて生成された少なくとも1ページの画像情報を取得する。

30

【0027】

特定部33は、第1の画像情報取得部31により取得された複数ページの画像情報と、第2の画像情報取得部32により取得された少なくとも1ページの画像情報とを比較して、上記複数ページのうち選択されたページを特定する。具体的には、特定部33は、上記複数ページの画像情報のうち、上記選択されたページの画像情報と実質的に一致するページを特定する。画像情報同士が一致するか否かの判断において、特定部33は、ビットマップ形式の画像情報同士を照合してもよいし、各画像情報から抽出されるテキスト情報同士を照合してもよい。

【0028】

出力部34は、特定部33により特定されたページに基づいて、第1の画像情報取得部31により取得された複数ページの画像情報を分割して出力する。具体的には、出力部34は、上記特定されたページ(すなわち選択されたページ)を分割位置の基準として、複数ページの画像情報を分割して出力する。すなわち、出力部34は、上記特定されたページに基づいて特定される分割位置で、複数ページの画像情報を分割して出力する。

40

【0029】

一つの態様では、出力部34は、複数ページの画像情報を分割して所定の記憶装置に記憶させる。例えば、出力部34は、複数ページの画像情報を分割して複数の電子文書を生成し、生成された各電子文書を、当該電子文書を識別するための文書識別情報と関連付けて、所定の記憶装置に記憶させる。上記所定の記憶装置は、図1の例では、情報処理装置

50

30とLAN(Local Area Network)等の通信路を介して接続された文書管理システム3である。上記文書識別情報としては、例えば、各電子文書から抽出されるテキスト情報や、各電子文書に対して割り当てられるシリアル番号などが挙げられる。

【0030】

ただし、出力部34は、複数ページの画像情報をソーター等により分割して印刷出力させてもよいし、複数ページの画像情報を複数のウィンドウ等により分割して画面上に表示させてもよい。

【0031】

なお、情報処理装置30は、ページ数指定受付部36および出力部37をさらに有してもよい。

10

【0032】

ページ数指定受付部36は、ユーザからのページ数の指定を受け付ける。

【0033】

出力部37は、第1の画像情報取得部31により取得された複数ページの画像情報を、上記指定されたページ数毎に分割して出力する。出力部37は、画像情報を、所定の記憶装置に記憶させてもよいし、印刷出力させてもよいし、表示させてもよい。

【0034】

[出力部34による分割処理の具体例]

以下、出力部34による分割処理の具体例を示す。

【0035】

20

(第1の例)

本例では、出力部34は、特定されたページと当該ページの1つ前のページとの間を分割位置として、複数ページの画像情報を分割する。

【0036】

図2は、出力部34による分割処理の第1の例を説明するための概念図である。

【0037】

図2では、7ページの原稿が読み取られて、7ページの画像情報201が生成されている。また、上記7ページの原稿から選択された1ページ目および5ページ目の原稿が読み取られて画像情報202が生成されている。この場合、特定部33により、7ページのうち1ページ目および5ページ目が、選択されたページとして特定される。

30

【0038】

出力部34は、特定された1ページ目とその1つ前の0ページ目(実際には存在しない)との間、および、特定された5ページ目とその1つ前の4ページ目との間を分割位置として、7ページの画像情報を分割して出力する。すなわち、出力部34は、1~4ページ目の画像情報203と、5~7ページ目の画像情報204とをそれぞれ一纏まりの文書情報として、2つの文書情報を出力する。

【0039】

本例における使用の一態様では、ユーザは、4ページの第1の紙文書と3ページの第2の紙文書とを順に配列した合計7ページの原稿を情報処理システム1に読み取らせる。ついで、7ページの原稿のうち、第1の紙文書の先頭ページ(1ページ目)と、第2の紙文書の先頭ページ(5ページ目)とを抜き出して、情報処理システム1に読み取らせる。すると、情報処理システム1により、第1の紙文書に対応する画像情報と、第2の紙文書に対応する画像情報が、互いに分けて出力される。

40

【0040】

なお、本例において、1ページ目の原稿の選択は省略されてもよい。

【0041】

(第2の例)

本例では、出力部34は、特定されたページと当該ページの1つ後のページとの間を分割位置として、複数ページの画像情報を分割する。

【0042】

50

図3は、出力部34による分割処理の第2の例を説明するための概念図である。

【0043】

図3では、7ページの原稿が読み取られて、7ページの画像情報301が生成されている。また、上記7ページの原稿から選択された4ページ目および7ページ目の原稿が読み取られて画像情報302が生成されている。この場合、特定部33により、7ページのうち4ページ目および7ページ目が、選択されたページとして特定される。

【0044】

出力部34は、特定された4ページ目とその1つ後の5ページ目との間、および、特定された7ページ目とその1つ後の8ページ目（実際には存在しない）との間を分割位置として、7ページの画像情報を分割して出力する。すなわち、出力部34は、1～4ページ目の画像情報303と、5～7ページ目の画像情報304とをそれぞれ一纏まりの文書情報として、2つの文書情報を出力する。

10

【0045】

本例における使用の一態様では、ユーザは、4ページの第1の紙文書と3ページの第2の紙文書とを順に配列した合計7ページの原稿を情報処理システム1に読み取らせる。ついで、7ページの原稿のうち、第1の紙文書の最終ページ（4ページ目）と、第2の紙文書の最終ページ（7ページ目）とを抜き出して、情報処理システム1に読み取らせる。すると、情報処理システム1により、第1の紙文書に対応する画像情報と、第2の紙文書に対応する画像情報とが、互いに分けて出力される。

【0046】

なお、本例において、7ページ目の原稿の選択は省略されてもよい。

20

【0047】

[出力部37による分割処理の具体例]

図4は、出力部37による分割処理の一例を説明するための概念図である。以下、図4を参照して、出力部37による分割処理の一例を説明する。

【0048】

図4では、8ページの原稿が読み取られて、8ページの画像情報401が生成されている。また、ページ数指定受付部36は、ページ数の指定として、2ページの指定を受け付けている。

【0049】

出力部37は、8ページの画像情報401を、2ページ毎に分割して出力する。すなわち、出力部37は、1～2ページ目の画像情報402、3～4ページ目の画像情報403、5～6ページ目の画像情報404、および7～8ページ目の画像情報405を、それぞれ一纏まりの文書情報として、4つの文書情報を出力する。

30

【0050】

本例における使用の一態様では、ユーザは、それぞれ2ページの第1～第4の紙文書を順に配列した合計8ページの原稿を情報処理システム1に読み取らせる。また、ユーザは、情報処理システム1に対し、ページ数「2」を指定する。すると、情報処理システム1により、第1～第4の紙文書に対応する4つの画像情報が、互いに分けて出力される。

【0051】

[情報処理システムの動作]

図5は、本実施の形態に係る情報処理システム1の動作手順の一例を示すフローチャートである。以下、図5を参照して、情報処理システム1の動作の一例を説明する。

40

【0052】

なお、本例では、情報処理システム1には、LAN等の通信路を介して、業務情報管理システム2および文書管理システム3が接続されている。

【0053】

業務情報管理システム2は、ユーザを識別するためのユーザ識別情報（以下、「ユーザID」と称す）と、当該ユーザの業務情報とを関連付けて記憶する。業務情報は、1つ以上の証憑情報を含む。この証憑情報は、証憑の種類（見積書や発注書等）を示す情報と、

50

当該証憑のシリアル番号とを含む。

【 0 0 5 4 】

文書管理システム 3 は、電子文書と、当該電子文書を識別するための文書識別情報と関連付けて記憶する。

【 0 0 5 5 】

図 5 を参照すると、情報処理システム 1 (例えば複合機) は、まずログイン処理を行う (S 1)。具体的には、情報処理システム 1 は、ユーザインタフェース (例えば複合機に設けられたタッチパネル式の液晶ディスプレイ) を介して、ユーザから当該ユーザのユーザ ID を含むユーザ認証情報を受け付ける。そして、情報処理システム 1 は、当該ユーザ認証情報に基づいてユーザ認証を行う。

10

【 0 0 5 6 】

上記ユーザ認証に成功すると、情報処理システム 1 は、業務情報管理システム 2 から、上記ユーザ ID に関連付けられている証憑情報の一覧を取得し、当該証憑情報の一覧をユーザインタフェースに表示する (S 2)。

【 0 0 5 7 】

また、情報処理システム 1 は、ユーザインタフェースを介して、ユーザから登録モードの選択を受け付ける (S 3)。本例では、登録モードには、可変ページモードと、固定ページモードとがある。

【 0 0 5 8 】

情報処理システム 1 は、可変ページモードが選択された場合には、後述する可変ページ処理を実行し (S 4)、固定ページモードが選択された場合には、後述する固定ページ処理を実行する (S 5)。

20

【 0 0 5 9 】

可変ページ処理または固定ページ処理が終了すると、情報処理システム 1 は、ユーザインタフェースを介して、ユーザから登録処理を終了するか否かの指示を受け付け (S 6)、登録処理を終了する旨が指示された場合には、処理を終了させ、そうでない場合には、処理をステップ S 2 に戻す。

【 0 0 6 0 】

(可変ページモード)

図 6 は、可変ページモードの処理の一例を示すフローチャートである。以下、図 6 を参照して、可変ページモードの処理の一例を説明する。

30

【 0 0 6 1 】

ユーザは、複数の証憑が順に積層された複数ページの原稿を情報処理システム 1 の自動原稿送り装置にセットして、情報処理システム 1 に原稿のスキャンを指示する。なお、証憑の各々は、1 または複数ページの紙原稿である。

【 0 0 6 2 】

情報処理システム 1 は、上記セットされた複数ページの原稿を自動原稿送り装置により順番に読み取って、複数ページの画像情報を生成する (S 1 1)。

【 0 0 6 3 】

そして、情報処理システム 1 は、生成された全ページの画像情報を対象に OCR (光学的手法による文字認識 : Optical Character Recognition) 処理を実施して各ページからテキストデータを抽出し、各ページのページ番号と、各ページから抽出されたテキストデータとが互いに関連付けられて記録された全ページテーブルを作成する (S 1 2)。

40

【 0 0 6 4 】

ついで、ユーザは、上記複数ページの原稿から各証憑の先頭ページを抜き出し、各証憑の先頭ページが順に積層された少なくとも 1 ページの原稿を情報処理システム 1 の自動原稿送り装置にセットして、情報処理システム 1 に原稿のスキャンを指示する。

【 0 0 6 5 】

すると、情報処理システム 1 は、セットされた少なくとも 1 ページの原稿を自動原稿送り装置により読み取って、少なくとも 1 ページの画像情報を生成する (S 1 3)。

50

【0066】

そして、情報処理システム1は、生成された少なくとも1ページの画像情報を対象にOCR処理を実施して各ページからテキストデータを抽出し、各ページから抽出されたテキストデータ（以下、「先頭ページのテキストデータ」と称す）が記録された先頭ページテーブルを作成する（S14）。

【0067】

ついで、情報処理システム1は、図7に示されるファイル作成処理を実行する（S15）。

【0068】

図7を参照すると、情報処理システム1は、先頭ページテーブルに記録されている先頭ページのテキストデータのうち、未だ対象テキストデータとされていないものを対象テキストデータとする（S151）。

10

【0069】

そして、情報処理システム1は、対象テキストデータと、全ページテーブルに記録されている各ページのテキストデータとを比較し、全ページテーブルの中から対象テキストデータと一致するテキストデータを検索する（S152）。

【0070】

そして、情報処理システム1は、全ページテーブルにおいて、上記ステップS152で検索されたテキストデータに対応するページ番号に先頭ページフラグを付与する（S153）。

20

【0071】

ついで、情報処理システム1は、先頭ページテーブルに記録されている全てのテキストデータについて検索が終了したか否かを判断し（S154）、終了していないと判断された場合には（S154:NO）、処理をステップS151に戻し、終了したと判断された場合には（S154:YES）、処理をステップS155に進める。

【0072】

ステップS155では、情報処理システム1は、全ページテーブルにおいて、先頭ページフラグが付与されているページ番号のうち、最も番号が小さいページ番号と、二番目に番号が小さいページ番号とを特定する。

【0073】

そして、情報処理システム1は、上記特定された2つのページ番号に基づき、全ページの画像情報から所定範囲の画像情報を抽出し、当該画像情報を1つの文書ファイルとする（S156）。ここで、上記所定範囲は、上記特定された最も小さいページ番号に対応するページから、上記特定された二番目に小さいページ番号に対応するページの1つ前のページまでの範囲である。

30

【0074】

ついで、情報処理システム1は、全ページテーブルにおいて、上記特定された最も小さいページ番号に付与されている先頭ページフラグを消去する（S157）。

【0075】

ついで、情報処理システム1は、全ページテーブルにおいて、先頭ページフラグの個数が1個であるか否かを判断し（S158）、1個でないと判断された場合には（S158:NO）、処理をステップS155に戻し、1個であると判断された場合には（S158:YES）、処理をステップS159に進める。

40

【0076】

ステップS159では、情報処理システム1は、全ページテーブルにおいて、先頭ページフラグが付与されているページ番号を特定する。

【0077】

そして、情報処理システム1は、上記特定された1つのページ番号に基づき、全ページの画像情報から所定範囲の画像情報を抽出し、当該画像情報を1つの文書ファイルとし（S160）、ファイル作成処理を終了させる。ここで、上記所定範囲は、上記特定された

50

ページ番号に対応するページから、最終ページまでの範囲である。

【0078】

再び図6を参照すると、上記ファイル作成処理が終了した後、情報処理システム1は、図8に示されるファイル登録処理を実行する(S16)。

【0079】

図8を参照すると、情報処理システム1は、ステップS12で取得された証憑情報の一覧のうち、未だ対象証憑情報とされていないものを対象証憑情報とする(S161)。

【0080】

ついで、情報処理システム1は、上記ファイル作成処理で作成された文書ファイルの中から、対象証憑情報と一致するテキストを含む文書ファイルを検索する(S162)。このとき、情報処理システム1は、全ページテーブルを参照してもよい。

10

【0081】

そして、情報処理システム1は、文書ファイルが見つかったか否かを判断する(S163)。

【0082】

文書ファイルが見つかった場合(S163: YES)、情報処理システム1は、当該文書ファイルを対象証憑情報と関連付けて文書管理システム3に登録し(S164)、当該文書ファイルの登録先を対象証憑情報と関連付けて業務情報管理システム2に通知し(S165)、処理をステップS166に進める。

【0083】

20

一方、文書ファイルが見つからなかった場合(S163: NO)、情報処理システム1は、処理をステップS166に進める。

【0084】

ステップS166では、全ての証憑情報について検索が終了したか否かを判断し、終了していないと判断された場合には(S166: NO)、処理をステップS161に戻し、終了したと判断された場合には(S166: YES)、処理を終了させる。

【0085】

なお、上記の例では、全ページの読み取りが行われた後に先頭ページの読み取りが行われているが、先頭ページの読み取りが行われた後に全ページの読み取りが行われてもよい。

30

【0086】

(固定ページモード)

図9は、固定ページモードの処理の一例を示すフローチャートである。以下、図9を参照して、固定ページモードの処理の一例を説明する。

【0087】

情報処理システム1は、ユーザインタフェースを介して、ユーザからページ数(1以上の整数値)の指定を受け付ける(S21)。

【0088】

ついで、ユーザは、複数の証憑が順に積層された複数ページの原稿を情報処理システム1の自動原稿送り装置にセットして、情報処理システム1に原稿のスキンを指示する。ここで、各証憑のページ数は、互いに同じである。例えば、複数ページの原稿は、各々2ページの4つの証憑が順に積層された合計8ページの原稿である。

40

【0089】

情報処理システム1は、セットされた複数ページの原稿を自動原稿送り装置により順番に読み取って、複数ページの画像情報を生成する(S22)。

【0090】

ついで、情報処理システム1は、上記複数ページの画像情報を、上記指定されたページ数毎に分割し、分割された各画像情報を1つの文書ファイルとする(S23)。

【0091】

ついで、情報処理システム1は、上記作成された全文書ファイルを対象にOCR処理を

50

実施して各文書ファイルからテキストデータを抽出する（S 2 4）。

【0 0 9 2】

ついで、情報処理システム 1 は、ステップ S 1 2 で取得された証憑情報の一覧のうち、未だ対象証憑情報とされていないものを対象証憑情報とする（S 2 5）。

【0 0 9 3】

ついで、情報処理システム 1 は、上記抽出された各文書ファイルのテキストデータに基づき、上記作成された文書ファイルの中から、対象証憑情報と一致するテキストを含む文書ファイルを検索する（S 2 6）。

【0 0 9 4】

そして、情報処理システム 1 は、文書ファイルが見つかったか否かを判断する（S 2 7）。

10

【0 0 9 5】

文書ファイルが見つかった場合（S 2 7：YES）、情報処理システム 1 は、当該文書ファイルを対象証憑情報と関連付けて文書管理システム 3 に登録し（S 2 8）、当該文書ファイルの登録先を対象証憑情報と関連付けて業務情報管理システム 2 に通知し（S 2 9）、処理をステップ S 3 0 に進める。

【0 0 9 6】

一方、文書ファイルが見つからなかった場合（S 2 7：NO）、情報処理システム 1 は、処理をステップ S 3 0 に進める。

【0 0 9 7】

20

ステップ S 3 0 では、全ての証憑情報について検索が終了したか否かを判断し、終了していないと判断された場合には（S 3 0：NO）、処理をステップ S 2 5 に戻し、終了したと判断された場合には（S 3 0：YES）、処理を終了させる。

【0 0 9 8】

なお、本発明は、上記実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲内で種々変更することができる。

【図面の簡単な説明】

【0 0 9 9】

【図 1】実施の形態に係る情報処理システムの構成の一例を示すブロック図である。

【図 2】出力部 3 4 による分割処理の第 1 の例を説明するための概念図である。

30

【図 3】出力部 3 4 による分割処理の第 2 の例を説明するための概念図である。

【図 4】出力部 3 7 による分割処理の一例を説明するための概念図である。

【図 5】実施の形態に係る情報処理システムの動作手順の一例を示すフローチャートである。

【図 6】可変ページモードの処理の一例を示すフローチャートである。

【図 7】ファイル作成処理を示すフローチャートである。

【図 8】ファイル登録処理を示すフローチャートである。

【図 9】固定ページモードの処理の一例を示すフローチャートである。

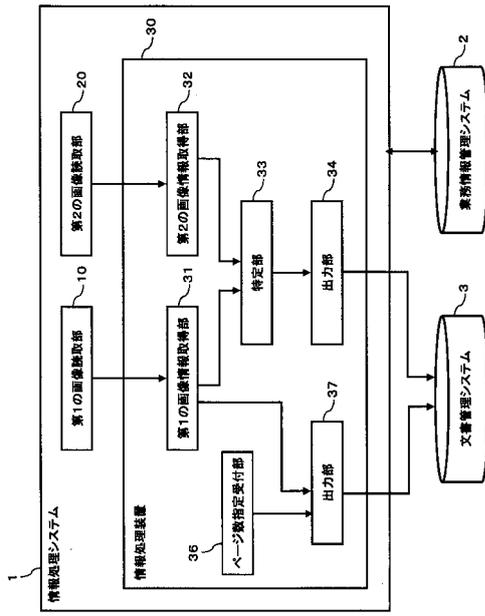
【符号の説明】

【0 1 0 0】

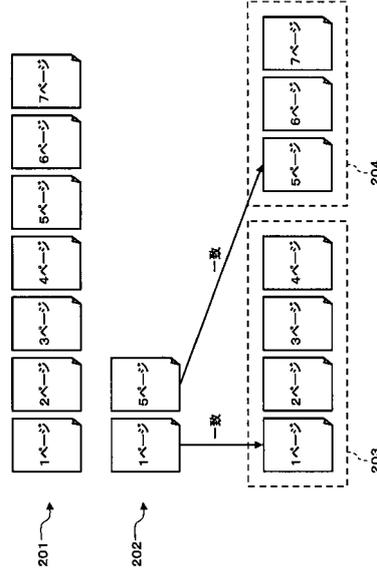
40

1 情報処理システム、1 0 第 1 の画像読取部、2 0 第 2 の画像読取部、3 0 情報処理装置、3 1 第 1 の画像情報取得部、3 2 第 2 の画像情報取得部、3 3 特定部、3 4 出力部、3 6 ページ数指定受付部、3 7 出力部。

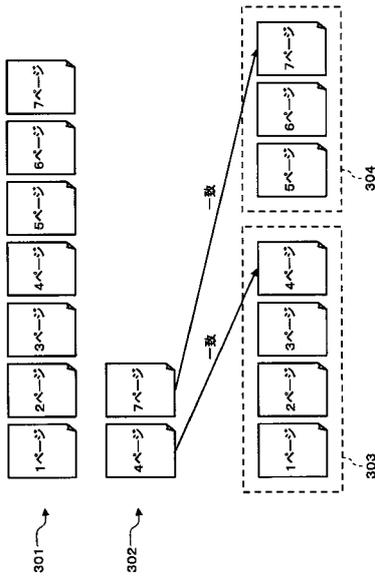
【図 1】



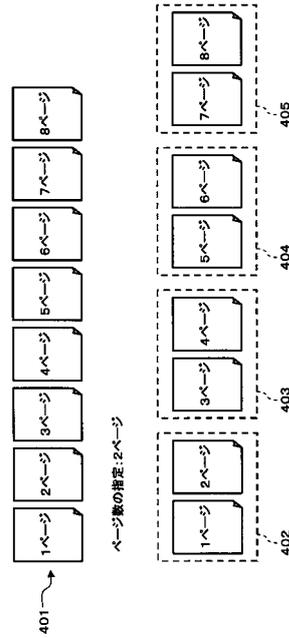
【図 2】



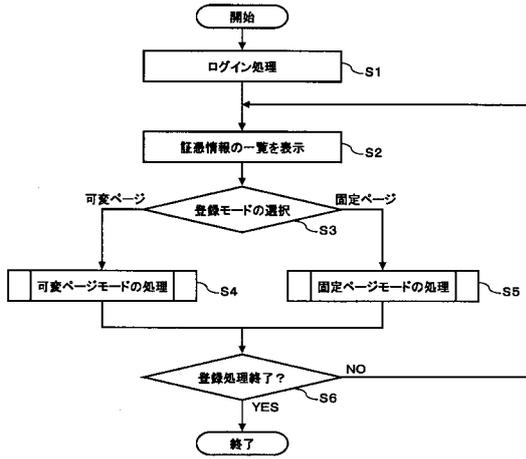
【図 3】



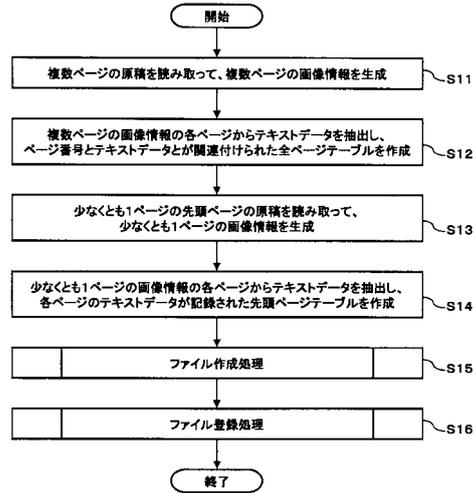
【図 4】



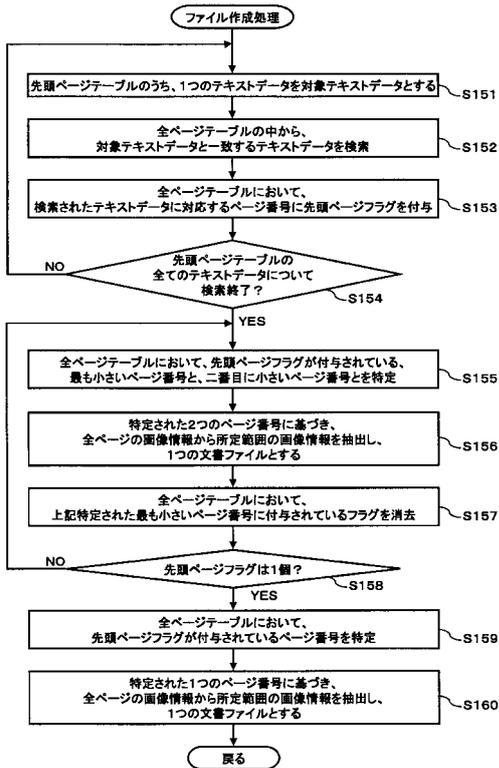
【図5】



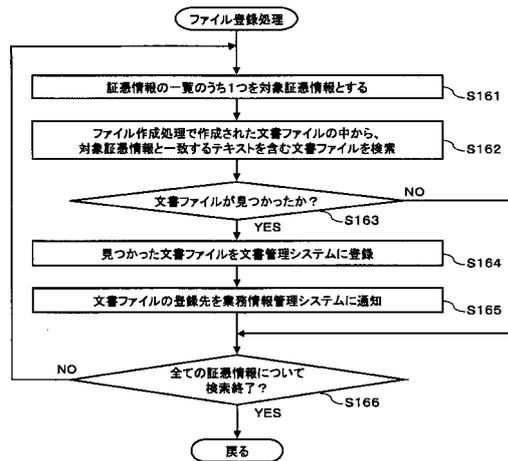
【図6】



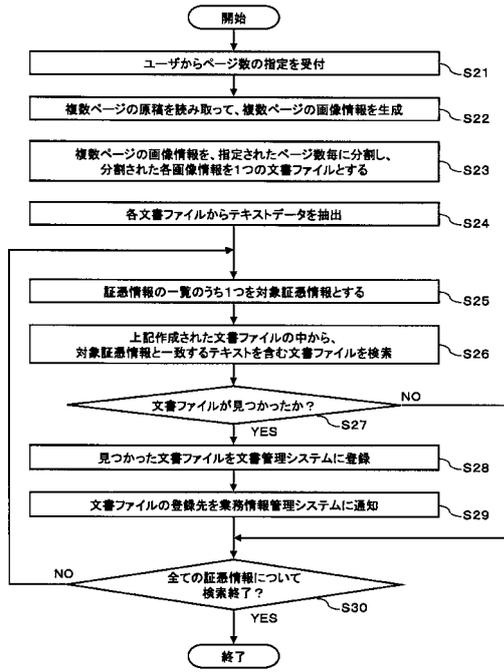
【図7】



【図8】



【図9】



フロントページの続き

- (56)参考文献 特開2006-338188(JP,A)
特開2005-115572(JP,A)
特開2004-254050(JP,A)
特開2002-024258(JP,A)
特開平06-314303(JP,A)
国際公開第2004/017240(WO,A1)

- (58)調査した分野(Int.Cl., DB名)
G06T 1/00