(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4785655号
(P4785655)

(45)発行日　平成23年10月5日(2011.10.5)　(24)登録日　平成23年7月22日(2011.7.22)

(51)Int.Cl.			F I		
H04N	1/387	(2006.01)	H04N	1/387	
G06T	11/60	(2006.01)	G06T	11/60	100A
G06T	3/00	(2006.01)	G06T	3/00	400J

請求項の数 14　(全 26 頁)

(21)出願番号　特願2006-190826(P2006-190826)
(22)出願日　平成18年7月11日(2006.7.11)
(65)公開番号　特開2008-22159(P2008-22159A)
(43)公開日　平成20年1月31日(2008.1.31)
審査請求日　平成21年7月8日(2009.7.8)

(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　高田　智美
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】文書処理装置及び文書処理方法

(57)【特許請求の範囲】
【請求項１】
文書処理装置であって、
文書画像から複数の物理ページを抽出する第１抽出手段と、
前記第１抽出手段によって抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出手段と、
前記第２抽出手段によって抽出された夫々のオブジェクトの有するテキストの特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定手段と、
前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合手段と、
を有することを特徴とする文書処理装置。
【請求項２】
文書処理装置であって、
文書画像から複数の物理ページを抽出する第１抽出手段と、
前記第１抽出手段によって抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出手段と、
前記第２抽出手段によって抽出された夫々のオブジェクトの有する表の特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定手段と、

前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合手段と、
を有することを特徴とする文書処理装置。
【請求項３】
文書処理装置であって、
文書画像から複数の物理ページを抽出する第１抽出手段と、
前記第１抽出手段によって抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出手段と、
前記第２抽出手段によって抽出された夫々のオブジェクトの色や形状の特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定手段と、
前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合手段と、
を有することを特徴とする文書処理装置。
【請求項４】
文書処理装置であって、
文書画像から複数の物理ページを抽出する第１抽出手段と、
前記第１抽出手段によって抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出手段と、
前記第２抽出手段によって抽出された夫々のオブジェクトの位置関係を解析し、当該位置関係に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定手段と、
前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合手段と、
を有することを特徴とする文書処理装置。
【請求項５】
前記結合手段は、前記複数の物理ページの位置又は倍率に基づいて、前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合することを特徴とする請求項１乃至４の何れか１項に記載の文書処理装置。
【請求項６】
前記結合手段は、前記オブジェクトの有するテキストのサイズと位置座標とに応じて、前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合することを特徴とする請求項１に記載の文書処理装置。
【請求項７】
前記第２抽出手段は夫々のオブジェクトに関するメタデータを抽出し、
前記オブジェクトと前記抽出されたメタデータとを関連付けて格納する格納手段を更に有することを特徴とする請求項１乃至６の何れか一項に記載の文書処理装置。
【請求項８】
前記第１抽出手段によって抽出された前記複数の物理ページの夫々のレイアウトを解析するレイアウト解析手段と、
前記レイアウト解析手段によって解析されたレイアウトに基づいて前記文書画像の論理構造を解析する論理構造解析手段とを更に有し、
前記第２抽出手段は、前記論理構造解析手段によって解析された論理構造とページ構成に基づいてメタデータを抽出することを特徴とする請求項７に記載の文書処理装置。
【請求項９】
オブジェクトを検索するための検索条件を入力するための検索条件入力手段と、
前記検索条件入力手段によって入力された検索条件に基づいてオブジェクトに関連付けられたメタデータを検索する検索手段とを更に有することを特徴とする請求項８に記載の文書処理装置。
【請求項１０】

文書処理装置の文書処理方法であって、
文書画像から複数の物理ページを抽出する第１抽出工程と、
前記第１抽出工程において抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出工程と、
前記第２抽出工程において抽出された夫々のオブジェクトの有するテキストの特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定工程と、
前記判定工程において前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合工程と、
を有することを特徴とする文書処理方法。
【請求項１１】
文書処理装置の文書処理方法であって、
文書画像から複数の物理ページを抽出する第１抽出工程と、
前記第１抽出工程において抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出工程と、
前記第２抽出工程において抽出された夫々のオブジェクトの有する表の特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定工程と、
前記判定工程において前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合工程と、
を有することを特徴とする文書処理方法。
【請求項１２】
文書処理装置の文書処理方法であって、
文書画像から複数の物理ページを抽出する第１抽出工程と、
前記第１抽出工程において抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出工程と、
前記第２抽出工程において抽出された夫々のオブジェクトの色や形状の特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定工程と、
前記判定工程において前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合工程と、
を有することを特徴とする文書処理方法。
【請求項１３】
文書処理装置の文書処理方法であって、
文書画像から複数の物理ページを抽出する第１抽出工程と、
前記第１抽出工程において抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出工程と、
前記第２抽出工程において抽出された夫々のオブジェクトの位置関係を解析し、当該位置関係に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定工程と、
前記判定工程において前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合工程と、
を有することを特徴とする文書処理方法。
【請求項１４】
コンピュータを、請求項１乃至９の何れか１項に記載の文書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】
本発明は、文書処理装置及び文書処理方法に関する。

【背景技術】
【０００２】
近年、電子文書の普及に伴い、それらを有効活用したいという需要が高まっている。
【０００３】
図１は、電子文書に対する処理の流れの一例を示す図である。図１に示すように、電子文書に対する操作は、一度作成・利用した後に蓄積・保存し、更にこれを編集・加工することによって新しい文書を作成するなど、文書作成のコスト削減のために再利用するのが一般的である。一方、印刷文書もコンピュータに取り込み、その内容を再利用したい、という要求がある。
【０００４】
印刷文書や電子文書を効率的に再利用するためには、大量の文書の中から必要な情報を探し出すための検索技術が重要となる。文書の中には、オブジェクトデータとして、文字情報だけでなく、図、表、写真等の画像情報も含まれており、特に利用頻度が高いと考えられる。文書に含まれる文字情報の場合は、指定された検索語と文字情報のマッチングを行うことで容易に検索することができる。しかし、画像情報等の場合は、それ自体は文字情報をもたないため、画像情報等に検索のためのメタデータを付加する技術が提案されている。
【０００５】
文書画像を複数の領域に分割し、各領域の特徴量によりテキストや画像等の種類を識別する技術（例えば、特許文献１参照）が提案されている。
【特許文献１】特開２０００－２９３６７１号公報
【発明の開示】
【発明が解決しようとする課題】
【０００６】
ところで、文書では、情報量の多い図表や画像を１ページに記述すると小さくて見難いため、複数のページにまたがって記述することがある。また、雑誌等の書籍では、向かい合った左右の２ページに１つの内容を記載する見開きを使用することがよくあり、これらは物理的には２ページ、論理的には１ページとみなすことができる。
【０００７】
このような物理的に複数のページが論理的な１ページである文書画像から、画像やテキスト等のオブジェクトを抽出する場合、次のような問題があった。
【０００８】
文書を物理ページ毎に読み込むと、複数の物理ページにまたがって記述されている１つのオブジェクトが分割されて抽出される。
【０００９】
また、ページ画像を読み込む際にはページ画像や領域毎に色や濃度を最適化し、ページ画像毎に倍率を調整するため、分割された画像毎に異なった画像処理が施される。また、ページ画像に歪みが発生することもあり、分割された画像を単に結合するだけでは、必ずしも元の画像が得られない。
【００１０】
また、論理ページを構成する各物理ページを別々に読み込み、ページ画像として合成した場合、左右の物理ページの間に空白が入ったり、位置が上下にずれたりすることがある。このような文書画像から抽出される画像やテキスト等のオブジェクトは、複数の物理ページにまたがって記述されている１つのオブジェクトが分割されて抽出されたものである。
【００１１】
また、ページ画像から分割して抽出された全てのオブジェクトについて、分割されたオブジェクトの各領域の違いを解析し、色情報や位置・倍率等の全ての要素を正確に補正して結合することは文書処理装置のＣＰＵに大変負荷のかかる処理である。また、補正して結合した各オブジェクトを文書処理装置の内部に保持しておくと、ディスクの負荷が増大

する。
【００１２】
本発明は、文書中の見開きのようなページに含まれる分割されたオブジェクトデータを１つのオブジェクトデータとして有効に利用することを目的とする。
【課題を解決するための手段】
【００１３】
本発明は、文書処理装置であって、文書画像から複数の物理ページを抽出する第１抽出手段と、前記第１抽出手段によって抽出された前記複数の物理ページの夫々からオブジェクトを抽出する第２抽出手段と、前記第２抽出手段によって抽出された夫々のオブジェクトの有するテキストの特徴を解析し、当該特徴に基づいて少なくとも１つのオブジェクトが前記複数の物理ページにまたがっているか否かを判定する判定手段と、前記判定手段によって前記複数の物理ページにまたがっていると判定されたオブジェクト同士を結合する結合手段と、を有することを特徴とする。
【発明の効果】
【００１４】
本発明によれば、文書中の見開きのようなページに含まれる分割されたオブジェクトデータを１つのオブジェクトデータとして有効に利用することができる。
【発明を実施するための最良の形態】
【００１５】
以下、図面を参照しながら発明を実施するための最良の形態について詳細に説明する。
【００１６】
［第１の実施形態］
第１の実施形態では、
図２は、本発明の一実施形態に係る文書処理システムが構築されるコンピュータ装置の基本構成を示すブロック図である。
【００１７】
図２において、２０１はＣＰＵであり、後述するＲＯＭやＲＡＭのプログラムに従って第１の実施形態の文書処理装置における各種制御を実行する。また、ＣＰＵ２０１自身の機能や計算機プログラムの機構により、複数の計算機プログラムを並列に動作させることができる。２０２はＲＯＭであり、ＣＰＵ２０１の制御手順を記憶する計算機プログラムや制御データが格納されている。２０３はＲＡＭであり、ＣＰＵ２０１が処理するための制御プログラムを格納すると共にＣＰＵ２０１が各種制御を実行する際の作業領域を提供する。
【００１８】
２０４はアルファベット、ひらがな、カタカナ、句点等を入力する文字記号入力キーや、カーソル移動を指示するカーソル移動キーのような各種機能キーを備えたキーボードであり、ユーザによる各種入力操作環境を提供する。また、マウスのようなポインティングデバイス、タッチパネル、スタイラスペンを含むこともできる。２０５はシステムバス（アドレスバス、データバスなど）であり、各構成を接続する。１０６は様々なデータなどを記憶するための外部記憶装置であり、ハードディスク、光ディスク、磁気ディスク、光磁気ディスク、不揮発性のメモリカード等の記録媒体と、記憶媒体を駆動し、情報を記録するドライブなどで構成される。保管された計算機プログラムやデータはキーボードなどの指示や各種計算機プログラムの指示により、必要な時にＲＡＭ上に完全或いは部分的に呼び出される。
【００１９】
２０７は表示器であり、ディスプレイなどで構成され、各種入力操作の状態をユーザに対して表示する。２０８は他の通信装置等と通信を行うためのネットワークコントロールユニット（ＮＣＵ）である。ネットワーク（ＬＡＮ）などを介して不図示の遠隔地に存在する装置と通信し、プログラムやデータを共有することが可能になる。２０９は画像を読み取るためのイメージスキャナであり、セットされた紙原稿を１枚ずつ光学的に読み取り

、イメージ信号をデジタル信号列に変換する。読み取られた画像データは、外部記憶装置やＲＡＭ等に格納される。

【００２０】

尚、通信手段としては、有線通信や無線通信など、何でも良く、またアダプタ装置などと接続され、通信を行っても良い。有線通信としては、ＲＳ２３２ＣやＵＳＢ、IEEE1394、Ｐ１２８４、ＳＣＳＩ、モデム、イーサネット（登録商標）などである。また無線通信としては、Bluetooth（登録商標）、赤外線通信、IEEE802.11xなどである。

【００２１】

また、画像データは、イメージスキャナ２０９だけでなく、ＮＣＵ２０８に接続されたネットワークスキャナやコピー装置等の入力機器を介して入力されても良い。読み取られた画像データも、外部記憶装置やＲＡＭなどではなく、ネットワークに接続されたサーバやコピー機等の外部記憶装置等に格納しても良い。

【００２２】

以上説明した構成は、第１の実施形態における一例であり、特にこれに限定されるものでない。

【００２３】

図３は、見開きのページ画像に対して領域抽出処理を行った結果を示す図である。このページ画像は、見開きを構成する各物理ページを別々に読み込んだ後、ページ画像として合成したため、左右の物理ページの間に空白があり、また位置が上下にずれている。そのため、異なる物理ページにまたがって記述されている画像等のオブジェクトが分割されて抽出されている。

【００２４】

図３において、３００は見開きを構成する左右の物理ページを一度にスキャンしたページ画像である。３１６及び３１７は各々抽出された物理ページの領域である。３０１～３１４は抽出されたオブジェクトデータを示す領域である。３０１、３１３及び３１４は、３０３～３１２の本文を構成する領域とは空間的に離れているため、それぞれ独立した文字領域又は画像領域として抽出される。

【００２５】

３０２～３０７と３０９～３１０は文字領域である。本実施形態では、文字列の方向が同じで、文字サイズと文字間値・行間値がほぼ均一であり、更に行方向の配置（字下げ、センタリング、揃えなど）が同じ部分が一つの文字領域として抽出される。３０８～３１０は画像領域であり、図として識別されている。３１５は後述する処理で抽出された物理ページの分割位置である。

【００２６】

尚、詳細は後述するが、３０９及び３１０、３１１及び３１２は、それぞれ一つのオブジェクトを構成するが、分割されて抽出されている。また、図３は、第１の実施形態における領域抽出結果の一例を示す図であるが、画像と文字の領域が抽出できれば、他の領域抽出結果でも構わない。

【００２７】

図４は、見開きページを物理ページ毎に読み込んだページ画像に対して領域抽出処理を行った結果を示す図である。図４において、４００及び４０１はスキャンしたページ画像である。４００は見開きの左側のページ画像であり、４０１は見開きの右側のページ画像であり、この例では位置が上下にずれている。４０２～４１２は抽出されたオブジェクトデータを示す領域である。４０２、４０３、４１２は、４０４～４１１の本文を構成する領域とは空間的に離れているため、それぞれ独立した文字又は画像領域として抽出される。

【００２８】

４０４、４０５、４０７、４１０、４１１は文字領域である。これらの文字領域は、文字列の方向が同じで、文字サイズと文字間値・行間値がほぼ均一であり、更に行方向の配置（字下げ、センタリング、揃えなど）が同じ部分が一つの文字領域として抽出される。

４０６、４０８、４０９は画像領域であり、図として識別されている。４１３及び４１４は、抽出された物理ページの領域である。

【００２９】

尚、詳細は後述するが、４０８及び４０９、４１０及び４１１は、一つのオブジェクトであるのに分割されて抽出されている。また、図４は、第１の実施形態における領域抽出結果の一例を示す図であるが、画像と文字の領域が抽出できれば、他の領域抽出結果でも構わない。

【００３０】

また、図３、図４では、画像とテキストが混在した文書画像を例に挙げたが、必ずしも複数の種類のオブジェクトが混在する必要はなく、例えば画像のみで構成された文書画像であっても構わない。

【００３１】

図５は、第１の実施形態における文書入力時の処理の一例を示すフローチャートである。この処理を示すプログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。

【００３２】

尚、図５に示す処理の説明では、一例として、イメージスキャナ２０９などの入力機器で読み取られた紙文書を対象として説明を行う。しかし、紙文書だけでなく、ワードプロセッサや編集ソフトで作成した文書、ＨＴＭＬなどで記述された文書、ＰＤＦなどの形式の電子文書でも構わない。

【００３３】

但し、電子文書の場合、ステップＳ５０１の入力処理において、フォーマット変換などの処理が必要となる。また、文章を文字コードで保持している文書の場合は、ステップＳ５０３の文字認識処理は不要となる。

【００３４】

まず、ステップＳ５０１において、ＣＰＵ２０１は、イメージスキャナ２０９やネットワークに接続されたコピー機などの入力機器を用いて文書を読み取り、電子化されたページ単位の文書画像を得る。入力機器によって入力される文書画像には、２値画像、カラー画像などがある。ページ画像を読み込む際に、ページ画像や領域毎に色情報等を最適化したり、またページ毎に位置や向き等が異なったりすることがある。

【００３５】

尚、電子化された文書画像を得た後、各ページ画像について、ノイズ除去処理や向きと傾きの補正処理を行っても良い。ページ画像の向きと傾きを判定し、修正する方法としては、公知のどのような方法を用いても構わない。

【００３６】

次に、ステップＳ５０２において、ＣＰＵ２０１は、ステップＳ５０１で読み取った文書の各ページ画像について領域分割を行う。そして、文字、図、表、写真などの画像を内包する矩形領域をその矩形の種類とサイズ、ページ内での位置座標等の物理的な情報と共に抽出する。

【００３７】

尚、文字領域については、ＣＰＵ２０１は、縦書き・横書きなどの文字列の読み方向と文字サイズを検出し、検出結果に基づいて文字列行と文字を抽出する。ここでは、文字列の方向が同じで、文字サイズと文字間値と行間値がほぼ均一である領域を一まとまりの文字領域として抽出する。尚、文字領域内の行方向の配置（字下げ、センタリング、揃えなど）を検出し、検出結果に基づいて文字領域を行方向に分割することで、更に、行方向の配置が同じ領域を一まとまりとしても良い。

【００３８】

また、非文字領域については、写真、表、枠や線などを検出し、領域として抽出する。入力された文書画像がカラー画像などの多値の場合は、２値に変換することで同様に領域分割処理を行うことができる。この領域分割方法としては、公知のどのような方法でも構

わない。

【００３９】

次に、ステップＳ５０３において、ＣＰＵ２０１は、全ての文字領域に対して文字認識処理を行い、その処理結果を全てＲＡＭ２０３や外部記憶装置２０６などの記憶媒体に格納する。そして、ステップＳ５０４において、文書の各ページ画像から物理ページを抽出する処理を行う。この処理は、自動又は手動で行う。自動で行う場合は、各ページ画像に対する物理ページの構成を判別する。そして、１枚のページ画像が複数の物理ページで構成されていれば、各ページ画像を物理ページ単位に分割する。物理ページの構成の判別は、ページ画像の縦横比率やステップＳ５０２で抽出された領域を利用する。例えば、横長のページ画像において最上部・最下部にヘッダやページ番号と思われる左右（上下）対象の領域が存在するかによって判別される。この物理ページ構成の判別方法は一例であり、他にもいろいろな方法が考えられる。また、文書入力時にユーザが指定しても良い。

【００４０】

次に、ステップＳ５０５において、ＣＰＵ２０１は、文書の各物理ページにおけるレイアウトを抽出し、テキストや画像などのコンテンツの種類毎に矩形領域で分割する。そして、得られた矩形領域の物理的な情報に従って各物理ページ画像における各矩形領域の空間的な関係を抽出する。例えば、物理ページ画像内の２つの領域に対する空間的な関係を各矩形領域の位置座標やサイズを用いて解析し、判定する。空間的な関係としては、互いの領域が存在する上下左右の方向や、２つの領域が重なっている、接している、含まれているなどの状態、２つの領域の大小関係などである。また、２つの領域が接していない場合には、隣接する各領域間の物理ページ画像全体における距離の比較から遠い又は近いなどを判定する。また、文字領域については、物理ページ画像内の他の文字領域との位置を比較することにより、行方向の配置を抽出しても良い。

【００４１】

以上の解析結果は、物理ページ毎に木構造やネットワーク構造で表現することができる。ここで挙げた各矩形領域間の関係及びその表現方法は、第１の実施形態における一例であり、他の関係が抽出されても良いし、また解析結果を他の方法で表現しても構わない。例えば、レイアウトとして、各矩形領域の物理ページ全体に対する相対的な位置やサイズなどを抽出しても良い。

【００４２】

図６は、ある物理ページ画像における各領域の空間的な関係を抽出した結果の一例を示す図である。図６では、ページ画像内の２つの領域に対する空間的な関係、更に、２つの領域が接していない場合には、隣接する２つの領域間の相対的な距離をネットワーク構造で表現している。例えば、領域１と領域２の空間的な関係は、領域５が領域４の下にあり、接していないが近い距離にあることを示している。

【００４３】

図５に戻り、ステップＳ５０６において、文書の全ての物理ページに対して、連続する複数の物理ページが論理的な１ページを構成しているか、或いは物理的な１ページが論理的な１ページであるかを自動又は手動で判別する。複数の物理ページから成る論理ページの判別を自動で行う場合は、文書の方向やステップＳ５０５で抽出したレイアウト、即ち「左のページ」「右のページ」のような見開きページ内で使用される言語表現などを利用する。

【００４４】

尚、ページ番号などを利用して物理ページの連続性を判定し、連続する物理ページについてのみ論理ページの組を判別する。そして、不連続な物理ページについては判別を行わないようにすると効率良く判別できる。更に、文書の種類が折り込みページのない書籍の場合は、向かい合うページと背中合わせのページが必ず交互に並ぶことを考慮して論理ページ構成を判別しても良い。論理ページの判別方法はこれに限るものではなく、他にもいろいろなものが考えられる。

【００４５】

次に、ステップＳ５０７において、ＣＰＵ２０１は、ステップＳ５０６の判別結果に基づき論理ページを取得する。そして、ステップＳ５０８において、ステップＳ５０７で取得した論理ページが見開きのように、複数の物理ページから成る論理ページであるか否かを判定する。複数の物理ページから成ると判定した場合はステップＳ５０９へ進み、１物理ページから成ると判定した場合はステップＳ５１２へ進む。

【００４６】

このステップＳ５０９では、ＣＰＵ２０１は、論理ページの組になる各物理ページ内の矩形領域についてサイズと位置を合わせる処理を行う。例えば、各ページの背景画像や飾り、抽出された矩形領域のレイアウトの規則性、物理ページの結合位置付近にある矩形領域の位置関係や位置座標・サイズなどを利用し、組になるページサイズの比率と、位置のずれを求めて調整する。矩形の範囲には誤差があるので、矩形の位置やサイズを調整しても、内部の画像やテキストが合致するとは限らないし、矩形のサイズや位置座標等の情報にも誤差があるので、ページ内の全ての矩形領域を完全に合致させるのは難しい。従って、完全に合わせる必要はなく、ある程度の誤差の範囲内で調整できれば良い。ページの倍率と位置を合わせる方法としてはこれに限るものではなく、他にもいろいろな方法が考えられる。

【００４７】

次に、ステップＳ５１０において、ＣＰＵ２０１は、論理ページの組になる各物理ページの結合位置付近にある２つの領域が物理ページによって分割された１つのオブジェクトか否かを判別する。この判別処理の詳細については、更に後述する。

【００４８】

次に、ステップＳ５１１において、ＣＰＵ２０１は、ステップＳ５０９、Ｓ５１０の結果に従って、ステップＳ５０５のレイアウト抽出結果に対する補正を行う。即ち、見開きなどの論理ページを対象としたレイアウト抽出処理を行い、ステップＳ５０５のレイアウト抽出結果に対して、論理ページに対するレイアウト情報を追加する。論理ページ上でのレイアウト情報として、第１の実施形態では、各領域の位置とサイズから各領域が物理ページのどちら側に属するか、或いは両方に属しているかなどの情報を追加する。補正方法や補正する情報はこれに限るものではなく、他にもいろいろなものが考えられる。

【００４９】

次に、ステップＳ５１２において、全ての論理ページに対して、ステップＳ５０７からステップＳ５１１までの処理が終了したか否かを判定する。ここで、未処理の論理ページがある場合は、次の論理ページに対してステップＳ５０７からステップＳ５１１の処理を行う。

【００５０】

以上、図５を用いて説明した文書の入力処理は、処理の一例であり、他にもいろいろなものが考えられる。これは、文書入力処理の一例であり、処理の順や処理内容は、これに限定されるものではない。

【００５１】

また、第１の実施形態では、文書入力時に、各オブジェクトについての分割判別処理を行っているが、分割判別処理のタイミングとしては文書入力時に限定されるものではなく、他のタイミングで行うようにしても良い。

【００５２】

図７は、ある文書におけるページ画像や各ページ画像から抽出された領域に関する各種物理的な情報の一例を示す図である。この例では、ページ画像に対して、ページサイズや読み込み時の解像度、電子化されたページ画像データの格納位置などの物理的な情報が付与されている。また、各ページ画像から抽出した物理ページについて、位置やサイズなどの情報と、同じ論理ページを構成している物理ページを示す情報が付与されている。

【００５３】

また、抽出された各矩形領域に対して、文字領域、画像領域などの領域種別、矩形領域のサイズ、ページ内での位置座標等の物理的な情報とが付与されている。更に、１つのオ

ブジェクトが分割されている矩形領域の場合には分割された他のオブジェクトを示す情報が付与されている。更に、文字領域については、文字サイズ、文字認識した結果である文字列が付与され、画像領域については、写真、表などの画像種別が付与されている。

【００５４】

例えば、ページ画像１は、幅が２９０ｍｍ、高さが２１０ｍｍで、処理解像度が３００ｄｐｉであり、領域１と領域２はページ画像１から抽出された物理ページで見開きページである。また、領域６は、Ｘ座標２０ｍｍ、Ｙ座標５０ｍｍの位置にある、幅５５ｍｍ、高さ５０ｍｍの文字領域で、文字サイズ９ポイントで記述されている文字列である。また、領域７及び領域９、領域８及び領域１０は、異なる物理ページに分割された１つのオブジェクトである。

【００５５】

図７は、領域の物理的な情報の一例を示しているが、物理的な情報とはこれに限るものではなく、次のステップにおいて、レイアウト抽出ができれば、他の情報が抽出されても良い。例えば、図７では、矩形領域のサイズと位置座標情報を抽出しているが、矩形領域の左上の位置座標と右下の位置座標を抽出するようにしても良い。

【００５６】

図８は、図５に示すステップＳ５１０における判別処理の詳細を示すフローチャートである。この処理は、２つの物理ページそれぞれの結合位置付近にある２つの領域が、物理ページによって分割された１つのオブジェクトか否かを判別する処理である。この処理のプログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。

【００５７】

まず、ステップＳ８０１において、ＣＰＵ２０１は、位置情報に基づいて、論理ページの組となる２つの物理ページの結合位置付近にある２つの領域を取得する。そして、ステップＳ８０２において、領域に含まれるオブジェクトの種類が同じか否かを判定する。判定の結果、オブジェクトの種類が同じと判定した場合はステップＳ８０３へ進み、ステップＳ５０９で調整した領域の矩形のサイズと位置、ページ内のレイアウトなどを利用して２領域が１つのオブジェクトである可能性を判定する。

【００５８】

例えば、図３では、物理ページを水平方向に結合するので、領域３０９と領域３１０の調整後の高さとＹ座標がほぼ同じであれば、１つのオブジェクトの可能性がある。また、例えば右側の物理ページの主な領域の左上Ｘ座標位置よりも、領域３１０は分割位置３１５に近く、また左側の物理ページの主な領域の右上Ｘ座標位置よりも、領域３０９は分割位置３１５に近い。これにより、領域３０９及び領域３１０は一つのオブジェクトである可能性が高いと言える。また、領域３０９及び領域４１０の距離は左右の物理ページ間の距離とほぼ一致することからも、一つのオブジェクトである可能性が高いと言える。

【００５９】

尚、矩形の範囲やサイズ・位置座標等には誤差があるので、サイズや位置の比較を行う場合には、誤差とみなせる程度の違いであれば完全に一致していなくても良い。矩形領域のサイズと位置関係を利用して判定する方法としては、これに限るものではなく、他にもいろいろな方法が考えられる。

【００６０】

次に、ステップＳ８０３において、ＣＰＵ２０１は、２つの領域が１つのオブジェクトであると判定した場合はステップＳ８０４の処理へ進む。そして、領域に含まれるオブジェクトの種類毎にその特徴を利用して２領域が１つのオブジェクトである可能性を判定する。矩形の範囲には誤差があるので、矩形の位置やサイズを調整しても、内部の画像やテキストの位置やサイズが合致するとは限らない。そこで、領域内に記述されている各オブジェクトを解析することによって判定を行う。

【００６１】

テキスト領域については、テキスト領域の文字の特徴や文字認識した文字列を利用する。例えば、領域内の文字サイズやスタイル、飾りなどが文書内の標準文字のそれと異なり

、かつ一致している場合は、１つのオブジェクトの可能性が高いと言える。また、例えば２つのテキスト領域を分割した場合と結合した場合の領域内の各テキスト文字列について、辞書とのマッチングや形態素解析を行い、解析の結果得られる評価値が大きい方が１つのオブジェクトである可能性が高い。また、見出しやキャプション等と思われる領域については、文字列の特徴を利用することによって判定できる。

【００６２】

例えば、図３に示す領域３１１、領域３１２、領域３０７は、画像領域と接しており、それぞれ領域３０９、領域３１０、領域４０８のキャプションと識別できる。また、領域３１１及び領域３０７のテキスト文字列は「“図”＋英数字＋“：”」から始まっているのに対して領域３１２はこのパターンに当てはまらない。これにより、領域３１２は領域３１１と組になると考えられる。文字サイズや文字コードなどは、正確に認識できないこともあるので、完全に一致しなくても誤差とみなせる程度の違いであれば良い。

【００６３】

表領域については、罫線の位置座標やセルのサイズ、マトリクス構造、セル内のテキストや画像などの領域を利用して判定する。これらの情報は、正確に認識できないこともあるので、完全に一致しなくても誤差とみなせる程度の違いであれば良い。

【００６４】

写真等については、結合部分の色や形状等の情報を利用して判定する。画像の色等は、スキャン時にページ画像毎にチューニングされている場合があるので、完全に一致していなくても、誤差の範囲内で判定すれば良い。

【００６５】

これらは、オブジェクトの特徴を利用して判定する方法の一例であり、他にもいろいろな方法が考えられる。

【００６６】

次に、ステップＳ８０４において、１つのオブジェクトと判定された場合はステップＳ８０５へ処理が進む。そして、ＣＰＵ２０１は、テキスト以外の領域についてテキスト領域との関係を利用して１つのオブジェクトである可能性を判定する。例えば、図３に示す領域３１１と領域３１２は、それぞれ画像領域３０９と画像領域３１０のキャプションであり、１つのオブジェクトと判定する（ステップＳ８０４）。そこで、領域３０９と領域３１０も、１つのオブジェクトと判定することができる。これは、テキスト領域との関係を利用して判定する方法の一例であり、他にもいろいろな方法が考えられる。

【００６７】

次に、ステップＳ８０５において、２つの領域が１つのオブジェクトと判定された場合はステップＳ８０６へ処理が進む。ＣＰＵ２０１は、ステップＳ８０１で取得した２つの領域の領域抽出結果に対して、一つの領域であることを示す情報を追加する。

【００６８】

次に、ステップＳ８０７において、ＣＰＵ２０１は、２つの物理ページの結合位置付近にある全ての領域に対して、ステップＳ８０１からステップＳ８０６の処理が終了したか否かを判定する。判定の結果、未処理の領域がある場合は、ステップＳ８０１に戻り、全領域について処理を終了するまで、上述の処理を繰り返す。

【００６９】

以上、図８に示す処理は、ステップＳ５１０の処理の一例であり、他にも様々なものが考えられる。例えば、ステップＳ８０２～Ｓ８０５の全ての判定処理を行う必要はなく、処理内容や処理順序はこの通りでなくても良い。

【００７０】

また、この例では、ステップＳ８０２～Ｓ８０５の何れかで可能性がないと判定された場合、それらは別の領域であると判定した。しかし、例えばステップＳ８０２～Ｓ８０５の何れかで可能性があると判定された場合、その確信度等によるポイントを加算し、全てのステップでの判断による総合ポイントによって判定を行っても良い。

【００７１】

次に、
図９は、第１の実施形態における見開きページの補正・結合処理を示すフローチャートである。この処理プログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。
【００７２】
まず、ステップＳ９０１において、ＣＰＵ２０１は、処理対象となるオブジェクトを取得する。利用するオブジェクトは、利用する目的やアプリケーションなどによって異なる。また、オブジェクトではなく論理ページ画像を取得しても良い。
【００７３】
次に、ステップＳ９０２において、ＣＰＵ２０１は、ステップＳ９０１で取得した利用対象が分割されているか否かを判定する。即ち、利用対象がオブジェクトの場合は、１つのオブジェクトを含む領域が複数に分割されているか否かを判定する。また、利用対象が論理ページ画像の場合は、その論理ページ内に含まれるオブジェクトを含む領域が物理ページによって分割されているか否かにより判定する。分割されていると判定した場合はステップＳ９０３へ処理を進め、分割されていないと判定した場合は、この処理を終了する。
【００７４】
図５を用いて説明したように、文書入力時に各オブジェクトについての分割判別処理を行っているので、ここではその情報を利用する。しかし、分割判別処理のタイミングとしては文書入力時に限るものではなく、ここで行うようにしても良い。
【００７５】
次に、ステップＳ９０３において、ＣＰＵ２０１は、分割されたオブジェクトを含む領域或いはページについて、色や濃度等を補正するか否かを判定する。補正すると判定した場合はステップＳ９０４へ進み、補正しないと判定した場合はステップＳ９０５へ進む。補正するか否かは、オブジェクトの種類や利用目的によって異なる。例えば、背景やページ飾りなどは、再利用性が低いので補正しなくても良い。
【００７６】
ステップＳ９０４では、ＣＰＵ２０１は、分割された領域又はページを色、濃度、倍率、或いは位置などについて正確に補正する処理を行う。色、濃度、倍率、位置の全てについて補正してもよいし、またこの中の何れか一つについて補正しても良いし、また、この中の組み合わせを補正しても良い。どのように補正するかは、オブジェクトの種類や利用目的によって異なる。
【００７７】
色の補正は、分割されたオブジェクトを含む画像データについて、例えば各画像領域の彩度、明度、色調の分布を利用することで行える。分割された画像オブジェクトの位置や倍率の補正は、例えば各画像領域の境界部分から複数の対応点を抽出し、対応点のずれを利用して画像領域間の変換式を算出することで行える。テキストオブジェクトの位置や倍率の補正は、各テキスト領域部分の画像データについて、領域中の各テキスト行のサイズと位置を利用して補正することができる。尚、領域中の各テキスト行のサイズは、例えば図４や図５に示すように物理ページを結合するのであれば、高さである。
【００７８】
表オブジェクトの補正は、オブジェクトを含む領域の画像データを補正しても良いし、表の罫線の位置座標やセルのサイズ、マトリクス構造などの情報を利用して補正しても良い。位置と倍率については正確に補正しない場合でも、図５のステップＳ５０９で矩形の位置・サイズを調整した際の情報を利用して、大体の位置と倍率を補正しても良い。また、論理ページ画像の場合は、ページ内に含まれるオブジェクトを含む領域を利用して補正する。
【００７９】
上述した補正方法はこれに限るものではなく、他にもいろいろなものが考えられる。
【００８０】

次に、ステップＳ９０５において、ＣＰＵ２０１は、分割された領域又はページを結合する処理を行う。この結合処理は、分割された領域又はページについて、結合した画像データを生成するが、表オブジェクトの場合は、画像データを生成するのではなく、結合した表データを抽出しても良い。また、テキストオブジェクトの場合は、各テキスト領域部分を結合した画像データを生成し、再度文字認識処理を行って文字サイズや文字コード等の文字情報を抽出する。結合したデータは、利用後は破棄して構わない。結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。

【００８１】

図９に示す処理は、第１の実施例における利用時の補正・結合処理の一例であり、処理の順や処理内容は、この通りでなくても良い。

【００８２】

第１の実施形態によれば、１つのオブジェクトデータが、複数のページに分割して記述されていることを判定できるようにすることにより、複数のページに含まれている分割されたオブジェクトデータを１つの領域として有効に利用することができる。

【００８３】

また、分割されたオブジェクトデータを含む複数のページを１つのページとして有効に利用することができる。

【００８４】

［第２の実施形態］

次に、図面を参照しながら本発明に係る第２の実施形態について詳細に説明する。第２の実施形態では、見開きのようなページに含まれている分割されたオブジェクトデータを１つのオブジェクトデータとして、必要に応じた精度で表示し、有効に利用する場合を説明する。

【００８５】

尚、第２の実施形態における文書処理システムの構成は、第１の実施形態の構成と同様であり、その説明は省略する。

【００８６】

図１０は、第２の実施形態におけるオブジェクト表示時の処理を示すフローチャートである。この処理のプログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。

【００８７】

第２の実施形態では、検索アプリケーションで検索を行った結果の一覧表示、検索結果を確認するために一覧の中から選択して拡大表示、一覧の中から選択したものを編集して再利用するための表示を想定している。しかし、検索結果の一覧表示だけでなく、例えば特定のフォルダやディレクトリ内に格納されているもの、又は何らかの方法でグループ化されたものを表示しても構わない。

【００８８】

まず、ステップＳ１００１において、ＣＰＵ２０１は、表示対象データを取得する。通常、表示対象は、表示を行うアプリケーションなどによって異なるので、各表示プログラムに応じた適切なものを取得する。例えば、画像検索結果を表示する場合は、検索結果の自然画像や写真等の画像オブジェクトを含む領域のデータを取得し、表検索結果を表示する場合は検索結果として得られた表オブジェクトを含む領域のデータを取得する。尚、画像や表以外のオブジェクトを表示対象としても良いし、論理ページ画像を表示対象としても良い。

【００８９】

次に、ステップＳ１００２において、ＣＰＵ２０１は、ステップＳ１００１で取得した表示対象が分割されているか否かを判定する。即ち、表示対象がオブジェクトの場合は、１つのオブジェクトを含む領域が複数に分割されているか否かを判定する。また、表示対象が論理ページ画像の場合は、その論理ページ内に含まれるオブジェクトを含む領域が物理ページによって分割されているか否かにより判定する。分割されていると判定した場合

はステップＳ１００３へ処理を進め、分割されていないと判定した場合はステップＳ１０１０へ処理を進める。

【００９０】

図５を用いて説明したように、文書入力時に各オブジェクトについての分割判別処理を行っているので、ここではその情報を利用する。しかし、分割判別処理のタイミングとしては文書入力時に限るものではなく、ここで行うようにしても良い。

【００９１】

次に、ステップＳ１００３～Ｓ１００５において、表示目的及び表示方法を判定する。また、ステップＳＳ１００３～Ｓ１００５に記述されたもの以外にも様々な表示目的及び表示方法がある。

【００９２】

このステップＳ１００３では、多くの表示対象を同時に表示する一覧表示か否かを判定する。ここで一覧表示と判定された場合にはステップＳ１００６へ進み、上述のステップＳ１００１で取得した表示対象を結合し、一覧表示のためのサムネイル画像データを生成する。一覧表示の場合、各表示対象は小さい画像であり、大体どのようなものかが分かればよいので補正する必要はない。より厳密に処理する場合は、位置や倍率については図５に示すステップＳ５０９で矩形の位置・サイズを調整した際の情報を利用して補正しても良い。

【００９３】

次に、ステップＳ１００４では、ＣＰＵ２０１は、ユーザに選択された特定のオブジェクトや論理ページを確認するための拡大表示か否かを判定する。ここで拡大表示と判定された場合にはステップＳ１００７へ進み、ステップＳ１００１で取得した表示対象を結合し、拡大表示するための画像データを生成する。その際、表示するデータの種類と表示の目的に応じて補正を行う。例えば、検索結果を確認する場合は、データの種類と検索アルゴリズムに応じて確認したい要素を補正して結合する。

【００９４】

例えば、色特徴量による画像検索結果の場合は、ユーザは画像の色情報を確認したいと想定できる。よって、分割された各画像領域部分の色の違いを正確に補正して結合した画像データを生成する。色の補正は、例えば、分割された各画像領域の彩度、明度、色調の分布を利用することで行うことができる。また、形状特徴量による画像検索結果の場合は、ユーザは画像の形状情報を確認したいと思われるので、分割された各画像領域部分の画像データの倍率と位置を正確に補正して結合した画像データを生成する。位置や倍率の補正は、例えば分割された各画像領域の境界部分から複数の対応点を抽出し、対応点のずれを利用して画像領域間の変換式を算出することで行うことができる。補正には、画像特徴量抽出時の補正情報を利用しても良い。

【００９５】

また、表の場合は、分割された各表領域部分の画像データの位置や倍率などを補正して結合した画像データを生成しても良い。更に抽出された表の情報、即ち表に関する罫線の位置座標やセルのサイズ、マトリクス構造などを利用して、結合した表示用データを生成しても良い。

【００９６】

また、論理ページ画像の場合は、ある程度のページ内容が分かればよいと思われるので、論理ページ内に含まれるオブジェクトを含む領域を利用して、物理ページ画像の位置や倍率をある程度補正して結合した画像データを生成する。

【００９７】

拡大表示は、検索結果の確認の場合だけとは限らないので、拡大表示する目的に応じて補正する内容は異なる。補正方法と結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。結合した画像データや表示用データは、表示を行った後は破棄して構わない。

【００９８】

次に、ステップＳ１００５では、ＣＰＵ２０１は、ユーザに選択された特定のオブジェクトや論理ページを編集するための表示か否かを判定する。ここで編集のための表示と判定された場合にはステップＳ１００８へ進み、ステップＳ１００１で取得した表示対象を編集のために補正して結合する処理を行う。例えば、画像オブジェクトの場合、分割された各画像領域部分の画像データの色と倍率と位置を正確に補正して結合した画像データを生成する。また、表オブジェクトの場合、分割された各表領域の位置や倍率等を正確に補正して結合し、表の情報、即ち表に関する罫線の位置座標やセルのサイズ、マトリクス構造等を抽出して、表示用データを生成する。

【００９９】

また、論理ページ画像の場合、論理ページ内に含まれる各オブジェクトをそれぞれ上述した方法で補正して結合した画像データ・表示用データを合成し、論理ページ画像データを生成する。その際、分割されているテキスト領域は、各テキスト領域部分の画像データについて、領域中の各テキスト行のサイズ（例えば、図３や図４に示すように物理ページを結合するのであれば、高さ）と位置が合うように補正して結合した画像データを生成する。そして、再度文字認識処理を行って文字サイズや文字コード等の文字情報を抽出し、結合した表示用データを生成する。また、背景やページ飾りなどは、再利用性が低いので補正しなくても良い。補正方法と結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。

【０１００】

次に、ステップＳ１００３～Ｓ１００５の何れにも該当しない場合はステップＳ１００９の処理へ進む。そして、ＣＰＵ２０１は、分割されている表示対象を表示対象の種類と表示目的及び表示方法に応じて、色、濃度、位置、サイズ等を補正・結合する処理を行う。

【０１０１】

次に、ステップＳ１０１０において、ＣＰＵ２０１は、ステップＳ１００６～Ｓ１００９の何れかで結合された表示対象、又は分割されていない表示対象を各画面に表示する処理を行う。そして、ステップＳ１０１１において、全ての表示対象に対して、処理を終了したか否かを判定する。未処理の表示対象がある場合は、ステップＳ１００１に戻り、表示対象がなくなるまで上述の処理を繰り返す。

【０１０２】

尚、図１０に示す処理は、第２の実施形態における表示処理の一例であり、処理の順や処理内容は、この通りでなくても良い。

【０１０３】

図１１は、検索結果や特定のフォルダ内に格納されているオブジェクト及び論理ページを一覧表示した画面例を示す図である。図１１は、専用のアプリケーションでウィンドウシステムを利用した場合の画面の例であるが、Ｗｅｂブラウザなどによって同様の機能が提供されるのでも構わない。

【０１０４】

図１１において、１１０１はタイトルバーと呼ばれるもので、このウィンドウのタイトル表示と、例えば移動や大きさの変更など全体の操作を行う部分である。１１０２、１１０３はこのウィンドウに関する機能を提供するボタンで、ヘルプの表示やこのウィンドウを閉じる操作などを指示するためのものである。

【０１０５】

次に、矩形領域１１０４、１１０５は、オブジェクトや論理ページのサムネイル画像を表示する領域である。矩形領域１１０４に表示されているオブジェクト及び論理ページは分割されており、矩形領域１１０５は分割されていないことを示している。また矩形領域１１０４が太枠となっているのは、この領域がユーザによって選択されていることを示しており、１１０５は選択されていない領域を示している。

【０１０６】

１１０６は、この一覧表示画面に表示することができないオブジェクトや論理ページの

表示を指示する部分である。「前画面」ボタンが選択されたことを検出すると、この画面に表示された一覧の前の一覧を表示し、「次画面」ボタンが押下されると、次の一覧を表示する。

【０１０７】

１１０７は選択された領域内に表示されたオブジェクトや論理ページを拡大表示することを指示するためのボタンであり、このボタンが選択されたことを検出すると、拡大表示のための画面へ移行する。

【０１０８】

１１０８は選択された領域内に表示されたオブジェクトや論理ページを編集することを指示するためのボタンであり、このボタンが選択されたことを検出すると、編集のための画面へ移行する。そして、１１０９の「終了」ボタンが選択されたことを検出すると、一覧表示画面を終了する。

【０１０９】

図１２は、第２の実施形態において、あるオブジェクト及び論理ページを拡大表示した画面例を示す図である。図１２は、専用のアプリケーションでウィンドウシステムを利用した場合の画面の例であるが、Ｗｅｂブラウザなどによって同様の機能が提供されるのでも構わない。

【０１１０】

図１２において、１２０１はタイトルバーと呼ばれるもので、このウィンドウのタイトル表示と、例えば移動や大きさの変更など全体の操作を行う部分である。１２０２、１２０３はこのウィンドウに関する機能を提供するボタンで、ヘルプの表示やこのウィンドウを閉じる操作などを指示するためのものである。

【０１１１】

次に、矩形領域１２０４は、図１１の一覧表示画面等を利用して選択されたオブジェクト及び論理ページを表示する領域であり、ここでは分割されたオブジェクト及び論理ページが表示されている。そして、１２０５の「終了」ボタンが押下されると、拡大表示画面を終了する。

【０１１２】

図１３は、第２の実施形態において、あるオブジェクト及び論理ページを編集する画面例を示す図である。これは、専用のアプリケーションでウィンドウシステムを利用した場合の画面の例であるが、Ｗｅｂブラウザなどによって同様の機能が提供されるのでも構わない。

【０１１３】

図１３において、１３０１はタイトルバーと呼ばれるもので、このウィンドウのタイトル表示と、例えば移動や大きさの変更など全体の操作を行う部分である。１３０２、１３０３はこのウィンドウに関する機能を提供するボタンで、ヘルプの表示やこのウィンドウを閉じる操作などを指示するためのものである。

【０１１４】

次に、矩形領域１３０４は、図１１の一覧表示画面等を利用して選択されたオブジェクト及び論理ページを表示する領域であり、分割されたオブジェクト及び論理ページができるだけ正確に補正された状態で表示されている。１３０５は編集を行うためのメニューを表示する部分である。ここでは、例として、「コピー」「切り取り」「貼付」等の項目が表示されているが、編集のメニュー項目としては、これに限るものではなく、他にもいろいろなものが考えられる。

【０１１５】

１３０６は編集された結果を保存することを指示するためのボタンであり、このボタンが選択されたことを検出すると、編集されたオブジェクト及び論理ページを保存するための画面へ移行する。そして、１３０７の「終了」ボタンが選択されたことを検出すると、編集画面を終了する。

【０１１６】

第２の実施形態によれば、複数ページに含まれている分割されたオブジェクトデータを、必要に応じて色、濃度、位置座標、倍率の少なくとも何れか一つ又はこれらの組み合わせを補正して結合するか、或いは補正しないで結合する。これにより、文書処理装置のＣＰＵとメモリに負担をかけずに、必要に応じた精度で、１つのオブジェクトデータとして表示することができる。

【０１１７】

また、分割されたオブジェクトデータを含む領域が存在する複数のページを、必要に応じて色、濃度、位置座標、倍率の少なくとも何れか一つ又はこれらの組み合わせを補正して結合するか、或いは補正しないで結合する。これにより、文書処理装置のＣＰＵとメモリに負担をかけずに必要に応じた精度で、１つのページとして表示することができる。従って、文書中のオブジェクトデータ又はページを有効に再利用することができる。

【０１１８】

［第３の実施形態］

次に、図面を参照しながら本発明に係る第３の実施形態について詳細に説明する。第３の実施形態では、見開きのようなページに含まれている分割されたオブジェクトデータから、文書に関するメタデータを精度良く抽出する場合を説明する。

【０１１９】

尚、第３の実施形態における文書処理システムの構成は、第１の実施形態の構成と同様であり、その説明は省略する。

【０１２０】

図１４は、第３の実施形態における検索用メタデータ抽出時の処理を示すフローチャートである。この処理のプログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。この処理は、ある一つの検索エンジンのための検索メタデータを袖出する際の処理であり、例えばこのシステムに複数の検索エンジンが実装されている場合には、この処理が複数回実行される。

【０１２１】

まず、ステップＳ１４０１において、ＣＰＵ２０１は、検索対象となる写真、図、表などのオブジェクトデータが含まれる領域情報を取得する。検索対象となるオブジェクトデータの種類は、検索の種類によって異なるので、各検索エンジンに応じた適切なオブジェクトデータを取得する。例えば、画像検索の場合は検索対象として自然画像や写真などの画像オブジェクトを取得し、表検索の場合は検索対象として表オブジェクトを取得する。尚、画像や表以外のオブジェクト領域を検索対象として取得しても良い。

【０１２２】

次に、ステップＳ１４０２～Ｓ１４０４において、ＣＰＵ２０１は、この後抽出されるメタデータを利用する検索エンジンの種類を判定する。尚、ステップＳ１４０２～Ｓ１４０４に記述されたもの以外にも様々な検索方法がある。

【０１２３】

ステップＳ１４０２では、ＣＰＵ２０１は、言語情報による検索か否かを判定する。ここで、言語情報による検索と判定された場合にはステップＳ１４０５へ進む。

【０１２４】

ステップＳ１４０３では、ＣＰＵ２０１は、色特徴量による画像検索か否かを判定する。ここで、色特徴量による画像検索と判定された場合にはステップＳ１４０７へ進む。

【０１２５】

ステップＳ１４０４では、ＣＰＵ２０１は、形状特徴量による画像検索か否かを判定する。ここで、形状特徴量による画像検索と判定された場合にはステップＳ１４０８へ進む。

【０１２６】

ステップＳ１４０２からステップＳ１４０４のいずれにも該当しない場合は、ステップＳ１４０９へ進む。

【０１２７】

ステップＳ１４０５では、ＣＰＵ２０１は、文書中から言語メタデータの抽出対象となる全てのテキスト領域を取得する。ここでは、検索対象となる写真、図、表などの画像オブジェクトに関連付けられているテキスト領域を取得するが、他のテキスト領域をメタデータの抽出対象として取得しても良い。画像オブジェクトとテキスト領域の関連付けは、後述する図１５に示すステップＳ１５０２で行われる。

【０１２８】

次に、ステップＳ１４０６において、ＣＰＵ２０１は、ステップＳ１４０５で取得したテキスト領域のうち、分割されているテキスト領域を結合し、結合した領域からテキストを取り出す。その際、各テキスト領域部分の画像データについて、領域中の各テキスト行のサイズと位置が合うように補正して結合した画像データを生成し、再度文字認識処理を行って文字情報を抽出する。領域中の各テキスト行のサイズは、例えば図３や図４に示すように物理ページを結合するのであれば、高さである。テキスト領域の補正方法と結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。また、文字情報を抽出した後は、結合した画像データは破棄して構わない。

【０１２９】

また、ステップＳ１４０１で取得した検索対象のオブジェクトについては、分割されていても１つのオブジェクトであることと他の領域との位置関係が分かれば良いので、補正も結合もする必要はない。

【０１３０】

次に、ステップＳ１４１０において、ＣＰＵ２０１は、テキスト領域のテキスト情報から検索対象となる写真、図、表などの画像オブジェクトに関連する言語メタデータを抽出する。その際、後述する論理構造解析を利用して画像オブジェクトについて説明している文字列を言語メタデータとして抽出しても良い。例えば、キャプションと思われるテキスト領域の文字情報から画像番号（「図１」）と画像名（「システム構成図」）を抽出し、画像名をメタデータとする。また、段落と思われるテキスト領域の文字情報から画像番号を含む文を抽出し、メタデータする。また、例えば「上（の）」のような画像の方向を示す語と画像を示す語を含む文をメタデータとして抽出し、その語が示す画像の方向とステップＳ５１１で抽出された論理ページ内での各領域の空間的な関係を照合し、画像と言語メタデータを関連付けても良い。以上は、言語メタデータを抽出する処理方法の一例であり、他にも様々な方法が考えられる。

【０１３１】

ステップＳ１４０７では、ＣＰＵ２０１は、ステップＳ１４０１で取得した検索対象オブジェクトが分割されていれば画像領域を結合する。その際、画像オブジェクトの色特徴量が正確に抽出できるように分割された各画像領域部分の色の違いを正確に補正して結合した画像データを生成する。色の補正は、例えば分割された各画像領域の彩度、明度、色調の分布などを利用することで行うことができる。位置や倍率については、検索アルゴリズムが精度をそれほど要求しない場合は、ある程度調整してあれば正確でなくても良いので、図５に示すステップＳ５０９で矩形の位置・サイズを調整した際の情報を利用して補正すれば良い。画像領域の補正方法と結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。

【０１３２】

次に、ステップＳ１４１１において、ＣＰＵ２０１は、検索対象となる画像オブジェクトの画像特徴を解析し、色特徴量を抽出する。色特徴量としては、例えば画像全体や画像を格子状に分割したブロックにおける色分布のヒストグラムや平均色などがある。画像オブジェクトが分割されていた場合に、結合した画像データは、色特徴量を抽出した後は破棄して構わない。また、検索結果表示時に利用できるよう、補正のための情報を保持しておくようにしても良い。

【０１３３】

ステップＳ１４０８では、ＣＰＵ２０１は、ステップＳ１４０１で取得した検索対象オブジェクトが分割されていれば画像領域を結合する処理を行う。その際、画像オブジェク

トの形状特徴量が正確に抽出できるように、分割された各画像領域部分の画像データの倍率と位置を正確に補正して結合した画像データを生成する。位置や倍率の補正は、例えば各画像領域の境界部分から複数の対応点を抽出し、対応点のずれを利用して画像領域間の変換式を算出することで行うことができる。色や濃度については、検索アルゴリズムが精度をそれほど要求しない場合は、補正しなくても良い。画像領域の補正方法と結合方法はこれに限るものではなく、他にもいろいろなものが考えられる。

【０１３４】

次に、ステップＳ１４１２において、ＣＰＵ２０１は、検索対象となる画像オブジェクトの画像特徴を解析し、形状特徴量を抽出する。形状特徴量としては、例えば画像全体や画像を格子状に分割したブロックにおける輝度勾配方向の離散化された強度分布などがある。画像オブジェクトが分割されていた場合に、結合した画像データは形状特徴量を抽出した後は破棄して構わない。また、検索結果表示時に利用できるように、補正のための情報を保持しておくようにしても良い。

【０１３５】

ステップＳ１４０９では、検索エンジンの種類がステップＳ１４０２～Ｓ１４０４の何れにも該当しない場合に、分割されているオブジェクトを検索エンジンに応じて、色、濃度、位置、サイズなどを補正・結合する。例えば、表を検索する検索エンジンで使用するメタデータを抽出する場合は、分割された表を含む領域について、位置座標やサイズなどを補正して結合する。

【０１３６】

図５を用いて説明したように、文書入力時に各オブジェクトについての分割判別処理を行っており、ステップＳ１４０７～Ｓ１４０９では、その結果を利用している。しかし、分割判別処理のタイミングとしては文書入力時に限るものではなく、ステップＳ１４０７～Ｓ１４０９の前に行うようにしても良い。

【０１３７】

次に、ステップＳ１４１３において、各検索エンジンに応じた方法で検索用メタデータを抽出する。例えば、表検索エンジンの場合は、分割された表に関する罫線の位置座標やセルのサイズ、マトリクス構造をメタデータとして抽出する。オブジェクトが分割されていた場合に結合した画像データは、検索用メタデータを抽出した後は破棄して構わない。また、検索結果表示時に利用できるように、補正のための情報を保持しておくようにしても良い。

【０１３８】

そして、ステップＳ１４１４において、各検索エンジンに応じた全ての検索対象オブジェクトに対して、ステップＳ１４０１～Ｓ１４１３の処理が終了したか否かを判定する。未処理の検索対象オブジェクトがあると判定した場合はステップＳ１４０１に戻り、次の検索対象オブジェクトに対してステップＳ１４０１～Ｓ１４１３の処理を行う。

【０１３９】

図１４に示す処理は、第３の実施形態におけるメタデータ抽出処理の一例であり、処理の順や処理内容は、この通りでなくても良い。

【０１４０】

次に、図１５を用いて、第３の実施形態における文書登録時の動作について詳細に説明する。図１５は、第３の実施形態における文書登録処理を示すフローチャートである。この処理のプログラムは、ＲＯＭ２０２に格納されており、ＣＰＵ２０１によって実行される。

【０１４１】

まず、ステップＳ１５０１において、ＣＰＵ２０１は、画像及び文字情報が混在した１ページ以上で構成される文書画像を入力し、その文書画像を解析し、次の論理構造抽出処理のための前処理を行う。ステップＳ１５０１の処理については、図５を用いて説明した通りである。

【０１４２】

次に、ステップＳ１５０２において、ＣＰＵ２０１は、各領域に関する各種情報、レイアウト抽出結果、及び文字領域に含まれる文字情報の特徴などに基づき、論理構造解析規則に従って解析を行い、文書の論理構造を抽出する。論理構造とは、図７に示すように、ステップＳ１５０１で抽出された領域やページに対して論理的な意味属性を抽出して付与したもの、及びそれらの論理的な関係を推定し構造化したものである。論理構造解析規則には、上述の論理ページを処理対象とする規則と物理ページを処理対象とする規則がある。

【０１４３】

次に、ステップＳ１５０３において、検索用メタデータの抽出処理を行う。ステップＳ１５０３の処理については、図１４を用いて説明した通りである。

【０１４４】

そして、ステップＳ１５０４において、ステップＳ１５０３で抽出された画像とメタデータを関連付けてＤＢに格納する。

【０１４５】

次に、第３の実施形態において、抽出されたメタデータを利用して文書に含まれる写真、図、表などのオブジェクトを検索する時の動作について説明する。

【０１４６】

第３の実施形態では、写真、図、表などのオブジェクトデータに関連付けられているメタデータを利用して検索を行う。検索は、まずユーザが指定した検索キーワードやキーワードのリストなどの検索条件と各オブジェクトデータに関連付けられたメタデータを対比する。そして、その検索条件と適合するメタデータが付与されているオブジェクトデータをピックアップして検索結果として表示する。

【０１４７】

検索条件と各オブジェクトデータに関連付けられたメタデータを対比する方法は、各検索エンジンによって異なる。また、検索時に、検索条件とピックアップした各オブジェクトデータのメタデータとの類似度を計算して求めても良い。ここで言う類似度とは、ユーザが入力した検索条件が、各オブジェクトデータに付与されたメタデータとの関係を示す表現としてどの程度適切であるかを示すものである。これは、検索方法の例であり、検索方法としてはこれに限るものではなく、どのような方法でも構わない。

【０１４８】

また、メタデータを利用することにより、文書及び文書中のオブジェクトデータを蓄積する時に、効率的に分類・整理・管理することができるようになる。例えば、メタデータとして付与されている語を分析し、関連するカテゴリでオブジェクトデータを分類することができ、分類するカテゴリはユーザが与えても良いし、クラスタリング等の統計的手法によって自動的に分類するようにしても良い。また、分類時に、カテゴリと各オブジェクトデータのメタデータの類似度を計算して求め、分類に利用しても良い。これは、分類方法、文書管理方法の一例であり、文書管理方法としてはこれに限るものではなく、どのような方法でも構わない。

【０１４９】

第３の実施形態によれば、複数ページに含まれている分割されたオブジェクトデータから、文書中に含まれるオブジェクトデータに関するメタデータを抽出する時に、色、濃度、位置座標、倍率の少なくとも何れか一つ又はこれらの組み合わせを補正する。そして、分割されたオブジェクトデータを結合することにより、文書処理装置のＣＰＵとメモリに負担をかけずに、メタデータを精度良く抽出することができる。

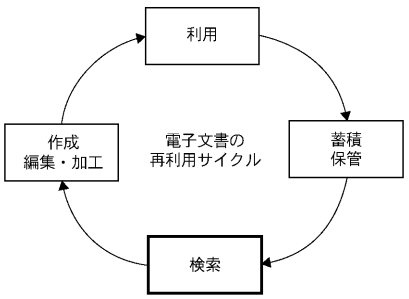
【０１５０】

また、複数ページに含まれている分割されたオブジェクトデータから、文書に関するメタデータを抽出する時に、色、濃度、位置座標、倍率の少なくとも何れか一つ又はこれらの組み合わせを補正する。そして、分割されたオブジェクトデータを結合することにより、文書処理装置のＣＰＵとメモリに負担をかけずにメタデータを精度良く抽出することができる。

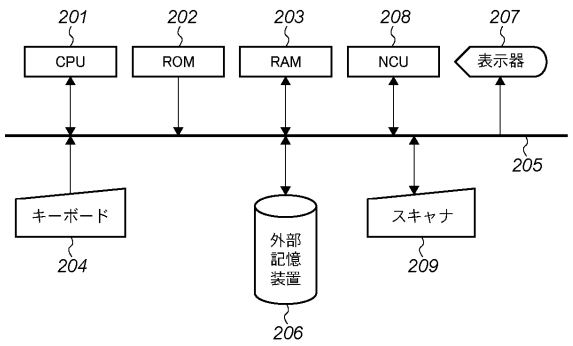
【０１５１】
従って、文書中のオブジェクトデータを効率的に再利用できる。また、メタデータを利用することにより、文書及び文書中のオブジェクトデータを蓄積する時に、効率的に分類・整理・管理することができる。
【０１５２】
尚、本発明は複数の機器（例えば、ホストコンピュータ，インターフェース機器，リーダ，プリンタなど）から構成されるシステムに適用しても、１つの機器からなる装置（例えば、複写機，ファクシミリ装置など）に適用しても良い。
【０１５３】
また、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ（ＣＰＵ若しくはＭＰＵ）が記録媒体に格納されたプログラムコードを読出し実行する。これによっても、本発明の目的が達成されることは言うまでもない。
【０１５４】
この場合、記録媒体から読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録媒体は本発明を構成することになる。
【０１５５】
このプログラムコードを供給するための記録媒体として、例えばフレキシブルディスク，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。
【０１５６】
また、コンピュータが読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、次の場合も含まれることは言うまでもない。即ち、プログラムコードの指示に基づき、コンピュータ上で稼働しているＯＳ（オペレーティングシステム）などが実際の処理の一部又は全部を行い、その処理により前述した実施形態の機能が実現される場合である。
【０１５７】
更に、記録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込む。その後、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又は全部を行い、その処理により前述した実施形態の機能が実現される場合も含まれることは言うまでもない。
【図面の簡単な説明】
【０１５８】
【図１】電子文書に対する処理の流れの一例を示す図である。
【図２】本発明の一実施形態に係る文書処理システムが構築されるコンピュータ装置の基本構成を示すブロック図である。
【図３】見開きのページ画像に対して領域抽出処理を行った結果を示す図である。
【図４】見開きページを物理ページ毎に読み込んだページ画像に対して領域抽出処理を行った結果を示す図である。
【図５】第１の実施形態における文書入力時の処理の一例を示すフローチャートである。
【図６】ある物理ページ画像における各領域の空間的な関係を抽出した結果の一例を示す図である。
【図７】ある文書におけるページ画像や各ページ画像から抽出された領域に関する各種物理的な情報の一例を示す図である。
【図８】図５に示すステップＳ５１０における判別処理の詳細を示すフローチャートである。
【図９】第１の実施形態における利用時の補正・結合処理を示すフローチャートである。
【図１０】第２の実施形態におけるオブジェクト表示時の処理を示すフローチャートであ

る。
【図１１】検索結果や特定のフォルダ内に格納されているオブジェクト及び論理ページを一覧表示した画面例を示す図である。
【図１２】第２の実施形態において、あるオブジェクト及び論理ページを拡大表示した画面例を示す図である。
【図１３】第２の実施形態において、あるオブジェクト及び論理ページを編集する画面例を示す図である。
【図１４】第３の実施形態における検索用メタデータ抽出時の処理を示すフローチャートである。
【図１５】第３の実施形態における文書登録処理を示すフローチャートである。
【符号の説明】
【０１５９】
２０１　ＣＰＵ
２０２　ＲＯＭ
２０３　ＲＡＭ
２０４　キーボード
２０５　システムバス
２０６　外部記憶装置
２０７　表示器
２０８　ＮＣＵ
２０９　スキャナ

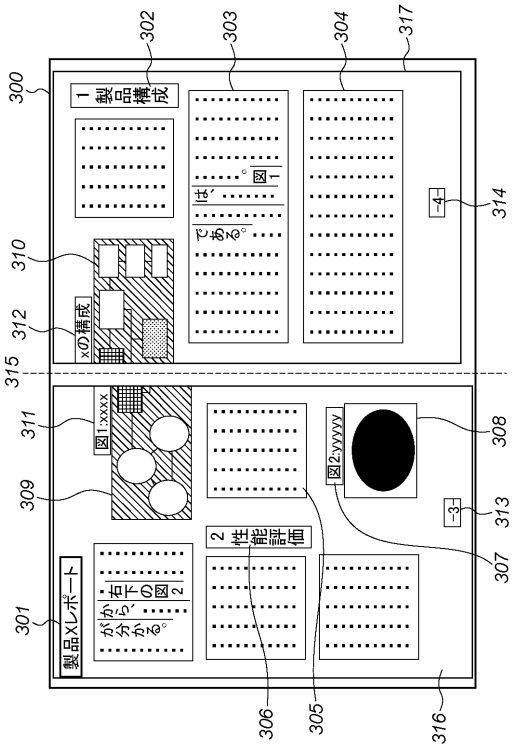
【図１】



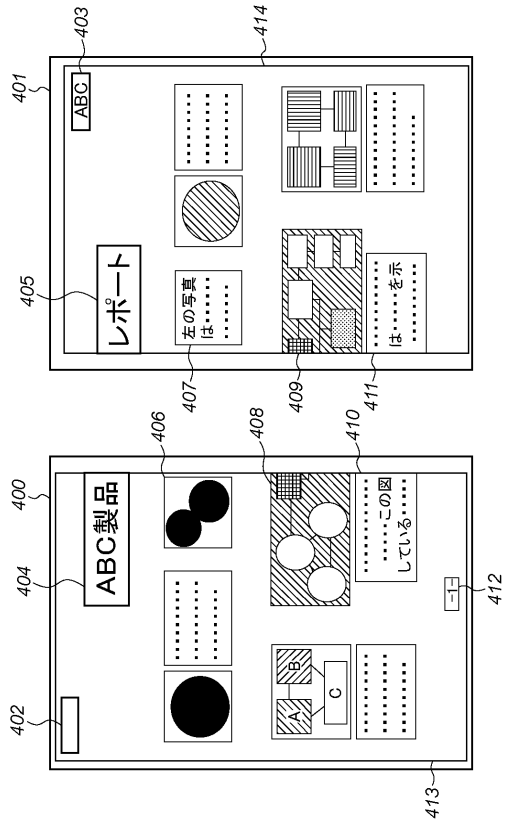
【図２】



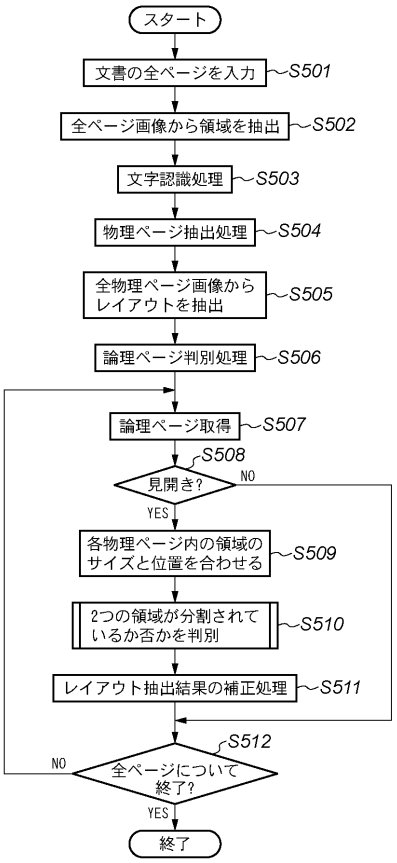
【図３】



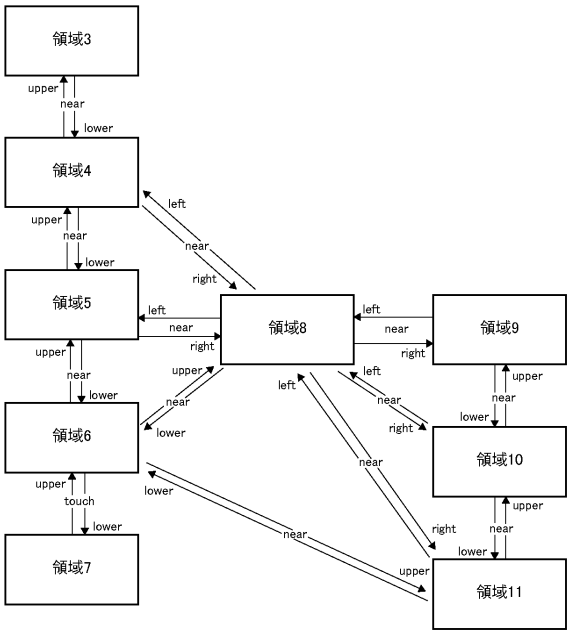
【図４】


403
414
401
ABC
レポート
405
左の写真は……
……を示は……
407
409
411
406
408
410
400
ABC製品
404
……この図している……
412
-1-
402
A
B
C
413

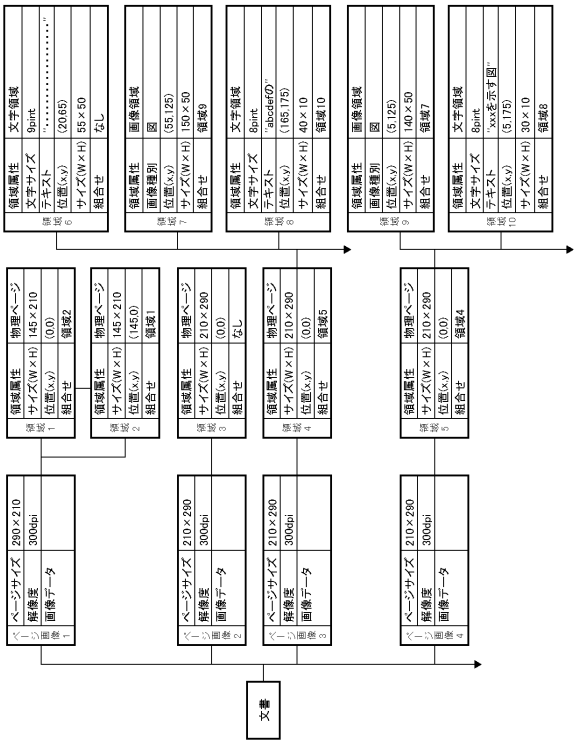
【図５】


スタート
文書の全ページを入力
S501
全ページ画像から領域を抽出
S502
文字認識処理
S503
物理ページ抽出処理
S504
全物理ページ画像から
レイアウトを抽出
S505
論理ページ判別処理
S506
論理ページ取得
S507
S508
見開き?
NO
YES
各物理ページ内の領域の
サイズと位置を合わせる
S509
2つの領域が分割されて
いるか否かを判別
S510
レイアウト抽出結果の補正処理
S511
S512
NO
全ページについて
終了?
YES
終了

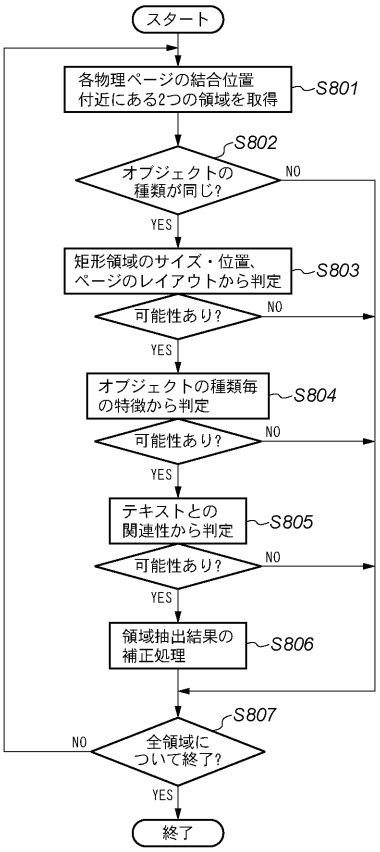
【図６】


領域3
upper
near
lower
領域4
upper
near
lower
領域5
upper
near
lower
領域6
upper
touch
lower
領域7
left
near
right
領域8
left
near
right
領域9
upper
near
lower
領域10
upper
near
lower
領域11

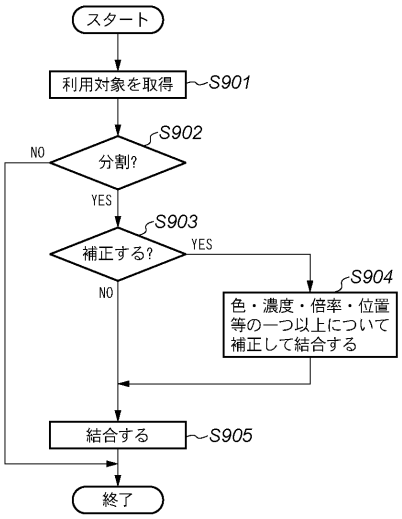
【図７】


文書

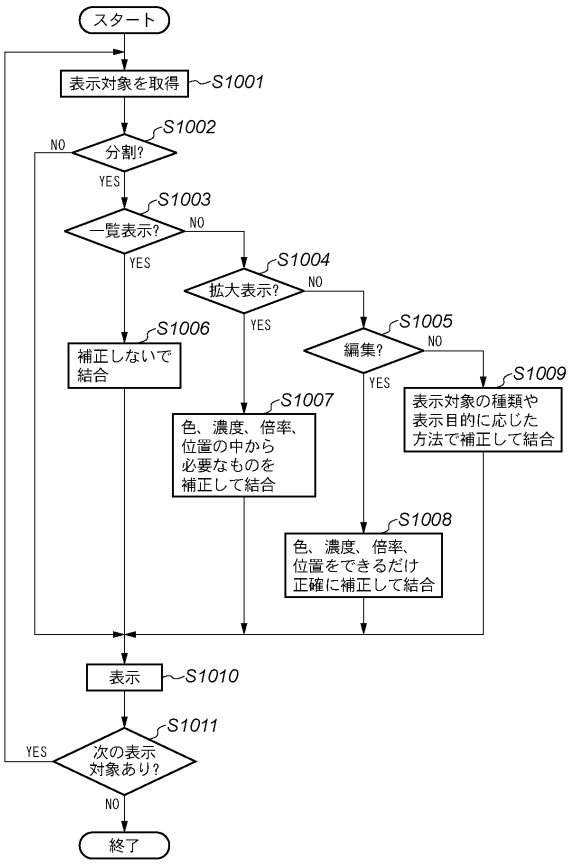
【図８】


スタート
各物理ページの結合位置付近にある2つの領域を取得
S801
S802
オブジェクトの種類が同じ?
NO
YES
矩形領域のサイズ・位置、ページのレイアウトから判定
S803
可能性あり?
NO
YES
オブジェクトの種類毎の特徴から判定
S804
可能性あり?
NO
YES
テキストとの関連性から判定
S805
可能性あり?
NO
YES
領域抽出結果の補正処理
S806
S807
全領域について終了?
NO
YES
終了

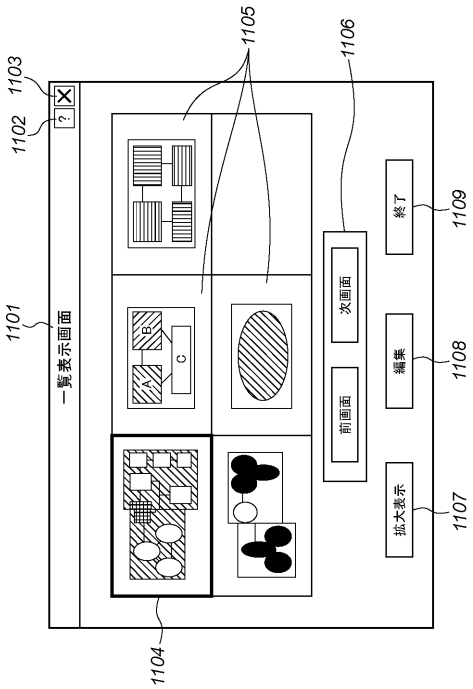
【図９】


スタート
利用対象を取得
S901
S902
分割?
NO
YES
S903
補正する?
YES
NO
S904
色・濃度・倍率・位置等の一つ以上について補正して結合する
結合する
S905
終了

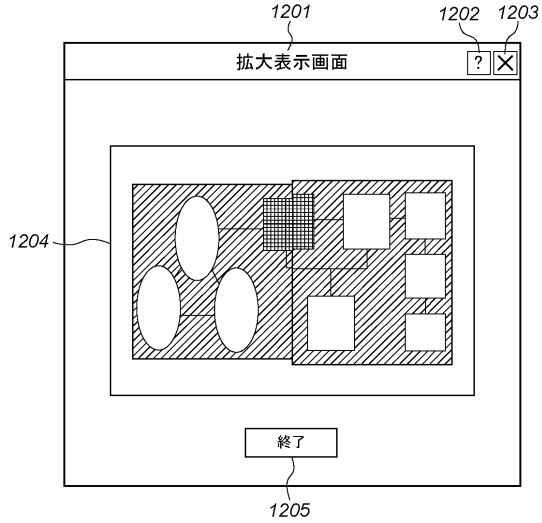
【図１０】


スタート
表示対象を取得
S1001
S1002
分割?
NO
YES
S1003
一覧表示?
NO
YES
S1004
拡大表示?
NO
YES
S1005
編集?
NO
YES
S1006
補正しないで結合
S1007
色、濃度、倍率、位置の中から必要なものを補正して結合
S1008
色、濃度、倍率、位置をできるだけ正確に補正して結合
S1009
表示対象の種類や表示目的に応じた方法で補正して結合
表示
S1010
S1011
次の表示対象あり?
YES
NO
終了

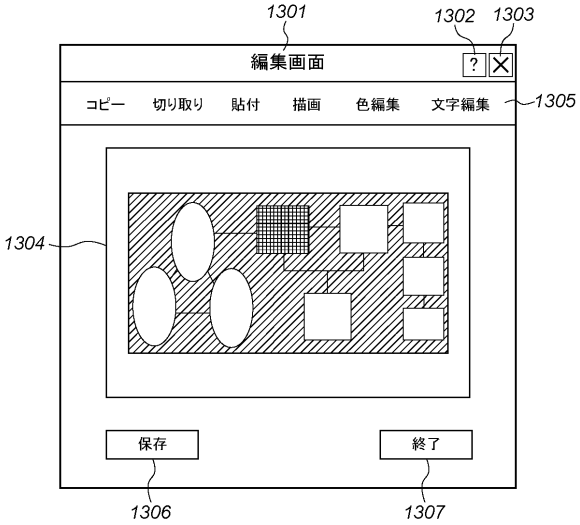
【図１１】


1105
1106
1103
1102
1101
一覧表示画面
次画面
前画面
終了
編集
拡大表示
1109
1108
1107
1104

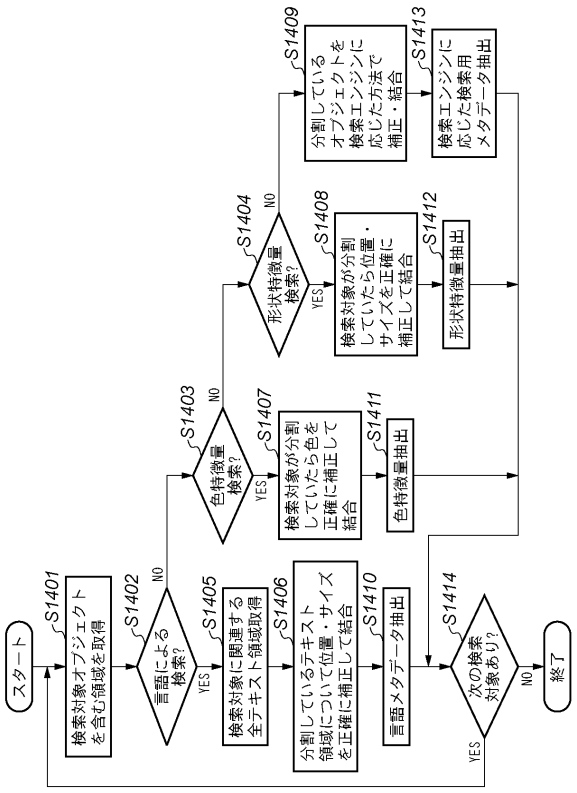
【図１２】



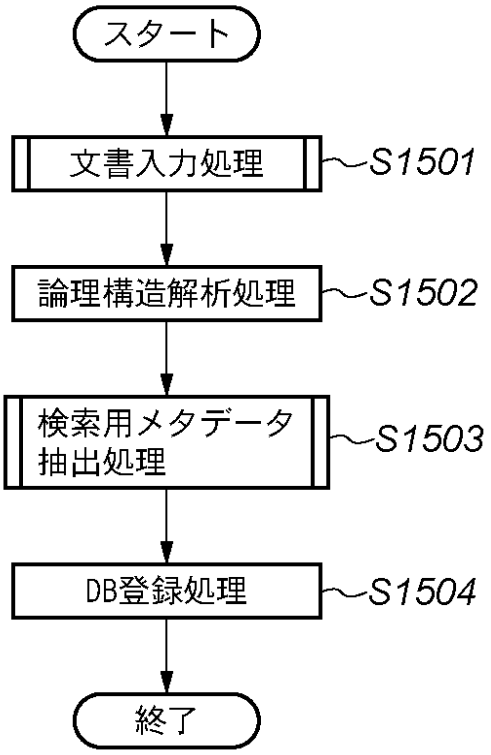
【図１３】



【図１４】



【図１５】



フロントページの続き

審査官　秦野　孝一郎

(56)参考文献　特開２０００－２９３６７１（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｎ　　１／３８７
Ｇ０６Ｔ　　３／００
Ｇ０６Ｔ　１１／６０