

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4480109号
(P4480109)

(45) 発行日 平成22年6月16日(2010.6.16)

(24) 登録日 平成22年3月26日(2010.3.26)

(51) Int.Cl.

F I

G 0 6 T 1/00 (2006.01)

G 0 6 T 1/00 2 0 0 C

請求項の数 17 (全 15 頁)

(21) 出願番号	特願2000-174017 (P2000-174017)	(73) 特許権者	000001007
(22) 出願日	平成12年6月9日(2000.6.9)		キヤノン株式会社
(65) 公開番号	特開2001-351089 (P2001-351089A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成13年12月21日(2001.12.21)	(74) 代理人	100076428
審査請求日	平成19年6月11日(2007.6.11)		弁理士 大塚 康德
		(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
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 画像管理装置および画像管理方法

(57) 【特許請求の範囲】

【請求項 1】

画像を管理するための画像管理装置であって、

所定の分類毎に前記画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第1生成手段と、

前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第2生成手段と、

前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第3生成手段と、

生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存手段と、

選択指示された出力形式に必要なデータを前記保存手段が保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換手段と、

を備えたことを特徴とする画像管理装置。

【請求項 2】

前記保存手段は、

更に、前記画像の全体を示す全体画像データを、前記部分画像データと前記領域データ

と前記配置データとに関連付けて保存することを特徴とする請求項 1 に記載の画像管理装置。

【請求項 3】

前記全体画像データを圧縮する手段を備え、

前記保存手段は、圧縮された前記全体画像データを保存することを特徴とする請求項 2 に記載の画像管理装置。

【請求項 4】

前記保存手段は、前記部分画像データを、前記識別された各領域の分類にしたがって適応的に圧縮を行って保存することを特徴とする請求項 1 に記載の画像管理装置。

【請求項 5】

前記保存手段は、前記識別された領域の分類が多値のピクチャ領域の場合は多値画像用圧縮を行った部分画像データを保存し、前記識別された領域の分類がテキスト領域の場合は 2 値画像にして 2 値画像用圧縮を行った部分画像データを保存することを特徴とする請求項 4 に記載の画像管理装置。

【請求項 6】

前記分類には、テキスト、表、枠、線画、若しくは、ピクチャが含まれることを特徴とする請求項 1 に記載の画像管理装置。

【請求項 7】

前記第 2 生成手段は、前記識別された領域の分類がテキストである場合、当該領域の画像に含まれる文字に対して前記文字認識処理を実行し、当該文字認識処理で得られるテキストデータを前記領域データとして生成することを特徴とする請求項 1 乃至 6 のいずれか 1 項に記載の画像管理装置。

【請求項 8】

前記第 2 生成手段は、前記識別された領域の分類が表である場合、当該領域の画像に前記表解析処理を実行し、当該表解析処理で得られる表の構造に関する情報及び当該表に含まれる文字を示すテキストデータを、前記領域データとして生成することを特徴とする請求項 1 乃至 7 のいずれか 1 項に記載の画像管理装置。

【請求項 9】

前記第 2 生成手段は、前記識別された領域の分類が枠である場合、当該領域の画像に前記ベクトル変換処理を実行し、当該ベクトル変換処理で得られる枠線のベクトルデータを前記領域データとして生成することを特徴とする請求項 1 乃至 8 のいずれか 1 項に記載の画像管理装置。

【請求項 10】

前記第 2 生成手段は、前記識別された領域の分類が線画である場合、当該領域の画像に前記ベクトル変換処理を実行し、当該ベクトル変換処理で得られる線画のベクトルデータを前記領域データとして生成することを特徴とする請求項 1 乃至 9 のいずれか 1 項に記載の画像管理装置。

【請求項 11】

前記保存された領域データから、指示された分類の領域データを読み出す手段を更に備えたことを特徴とする請求項 1 に記載の画像管理装置。

【請求項 12】

前記形式変換手段は、前記選択指示された出力形式がテキスト中心の出力形式であった場合、前記識別された分類がテキストである領域については、当該領域の前記領域データとして前記保存手段が保存した、前記文字認識処理で得られたテキストデータを抽出し、前記識別された分類がピクチャである領域については、前記保存手段が保存した前記部分画像データを抽出し、前記識別された分類が表である領域については、当該領域の領域データとして前記保存手段が保存した、前記表解析処理で得られる表の構造に関する情報と当該表に含まれる文字のテキストデータとを抽出し、当該抽出したデータを用いて当該選択指示されたテキスト中心の出力形式のデータに変換することを特徴とする請求項 1 乃至 11 のいずれか 1 項に記載の画像管理装置。

【請求項 1 3】

前記形式変換手段は、前記選択指示された出力形式が画像中心の出力形式であった場合、前記保存手段が保存した前記各領域の前記部分画像データを抽出し、当該抽出したデータを用いて当該選択指示された画像中心の出力形式のデータに変換することを特徴とする請求項 1 乃至 1 2 のいずれか 1 項に記載の画像管理装置。

【請求項 1 4】

前記形式変換手段は、前記選択指示された出力形式が表中心の出力形式であった場合、前記保存手段が前記領域データとして保存した、前記表解析処理で得られる表の構造に関する情報と当該表に含まれる文字のテキストデータとを抽出し、当該抽出したデータを用いて当該選択指示された表中心の出力形式のデータに変換することを特徴とする請求項 1 乃至 1 3 のいずれか 1 項に記載の画像管理装置。

10

【請求項 1 5】

前記形式変換手段は、前記選択指示された出力形式が要約の出力形式であった場合、前記保存手段が保存した、前記文字認識処理で得られたテキストデータを抽出し、当該抽出したテキストデータの要約処理を行うことにより要約文章を作成し、当該作成された要約文章に前記画像のサムネイル画像を付加したレポート文書に変換することを特徴とする請求項 1 乃至 1 4 のいずれか 1 項に記載の画像管理装置。

【請求項 1 6】

画像を管理するための画像管理方法であって、

第 1 生成手段が、所定の分類毎に前記画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第 1 生成工程と、

20

第 2 生成手段が、前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第 2 生成工程と、

第 3 生成手段が、前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第 3 生成工程と、

保存手段が、生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存工程と、

形式変換手段が、選択指示された出力形式に必要なデータを前記保存工程で保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換工程と、

30

を含むことを特徴とする画像管理方法。

【請求項 1 7】

コンピュータを、

所定の分類毎に画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第 1 生成手段、

前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第 2 生成手段、

40

前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第 3 生成手段、

生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存手段、

選択指示された出力形式に必要なデータを前記保存手段が保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換手段、として機能させるプログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

50

【発明の属する技術分野】

本発明は、文書等を示す画像の管理に関する。

【0002】**【従来の技術】**

従来、紙文書を電子化して、文書画像データとして保存する場合は、ドキュメントスキャナ等でスキャンして画像データを作成して保存する方法が採用されている。この場合、文書画像データはデータ容量が大きい場合が多いため、必要に応じて圧縮処理を施して保存することも行われている。

【0003】

ここで、ドキュメントスキャナの1度のスキャンで作成される文書画像データを1ファイルとしたときに、1つの紙文書から作成されるファイルは、通常、複数にわたり、データベース管理されるのが一般的であった。いわゆる文書データベースである。

10

【0004】

データベースを用いると文書画像データのファイル1つ1つにキーワードを付加することができ、検索に利用することができる。しかしながら、画像を丸ごと保管し、キーワードを付加する作業は、かなり面倒な作業であった。そこで、文書画像データを文字認識処理し、全文検索可能な形式にし、これと文書画像データを共に管理保管する方法も提案されている。

【0005】

図2は、文書画像データをデータベースに登録保存する従来の処理の流れを示した図である。

20

【0006】

図2(a)は、文書画像データに、キーワードを付加してデータベースに登録する手法を採用した従来の処理の流れを示した図である。紙文書201は、202で、スキャナによって文書画像データ化される。文書画像データ化されたファイルは、203でキーワードが付加され、204で、データベースに登録および保管される。データベース205の中には、キーワードを保存するインデックス(206)と文書画像データの実体(207)とが保管される。

【0007】

一方、上述した処理では、キーワード入力(203)の作業が面倒なため、図2(b)に示すような文字認識を利用したデータベースへの登録も提案されている。この文字認識を利用する方法では、文字コードがそのままキーワードと同じような役割を果たし、全文検索による検索も可能となる。

30

【0008】

図2(b)の210および211は、図2(a)の201および202に示すものと同様である。212では、文字認識処理が行われ、文書画像データ中のテキスト(TEXT)コードが抽出され、文字コードインデックス(216)が作成される。その後、213で、文字コードインデックス(216)および文書画像データ(217)が、データベースに登録・保管される。

【0009】

40

キーワードインデックス215を利用する場合と同様に、データベース214中に保存された文字コードインデックス216を利用することにより、文書画像データ217を特定することが可能であるため、利用者は、キーワードインデックス215の入力なしでも、文字コードインデックス216により検索が可能となった。

【0010】

また、図2(a)または(b)に示す方法によらず、通常のファイル保管方法でフォルダに保存する場合は、キーワードや文字コードのインデックスを付したファイルの保管とはならないことから、保管するファイルについてTEXT形式のものも作成し、双方のファイルに実質的に同じ名前のファイル名を付し、同じ場所に保存するなどの方法も提案されている。

50

【 0 0 1 1 】

このように、従来の文書画像の管理手法では、画像データのみならず、その検索等のために利用する種々の情報を画像データと共に管理されている。

【 0 0 1 2 】

【発明が解決しようとする課題】

しかしながら、従来の文書画像データをデータベースに保存する方法では、登録の際に、わざわざキーワードや文字認識によるコード情報を付加しても、所望の実体の検索や取出後に、これらの付加情報を文書画像データの実体に添付することができないという問題があった。そのため、付加情報を別ファイル中に取り込み、文書画像データにその付加情報を含む別ファイルを添付して渡す方法がとられ、管理が煩雑になっていた。

10

【 0 0 1 3 】

また、単にフォルダに保存した場合も、文書の実体と付加情報が別々に管理され、検索や所望の実体の取出後に、文書画像データが引き渡されるのみで、せっかく付加した情報は捨てられてしまうという問題があった。

【 0 0 1 4 】

要するに、従来の手法では、画像データの管理のために作成された種々の情報を有効活用できないという問題があった。

【 0 0 1 5 】

従って、本発明の目的は、管理される画像に関連する種々の情報を有効に活用し得る画像管理装置、画像管理方法及び記録媒体を提供することである。

20

【 0 0 1 6 】

【課題を解決するための手段】

本発明によれば、画像を管理するための画像管理装置であって、所定の分類毎に前記画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第1生成手段と、前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第2生成手段と、前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第3生成手段と、生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存手段と、選択指示された出力形式に必要なデータを前記保存手段が保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換手段と、を備えたことを特徴とする画像管理装置が提供される。

30

【 0 0 1 7 】

また、本発明によれば、画像を管理するための画像管理方法であって、第1生成手段が、所定の分類毎に前記画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第1生成工程と、第2生成手段が、前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第2生成工程と、第3生成手段が、前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第3生成工程と、保存手段が、生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存工程と、形式変換手段が、選択指示された出力形式に必要なデータを前記保存工程で保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換工程と、を含むことを特徴とする画像管理方法が提供される。

40

【 0 0 1 8 】

また、本発明によれば、コンピュータを、所定の分類毎に画像に含まれる各領域を識別し、前記画像中における当該識別された各領域の配置を示す配置データを生成する第1生成手段、前記識別された各領域に対して、文字認識処理と表解析処理とベクトル変換処理

50

との中から当該各領域の分類に応じて選択される処理を実行することにより、前記各領域の領域データを生成する第2生成手段、前記識別された全ての領域について、当該領域それぞれの画像を示す部分画像データを生成する第3生成手段、生成された前記部分画像データと前記領域データと前記配置データとを関連付けて保存する保存手段、選択指示された出力形式に必要なデータを前記保存手段が保存した各領域の前記部分画像データと前記領域データとの中から前記識別された各領域の分類に応じて抽出し、当該抽出したデータを用いて当該選択指示された出力形式のデータに変換する形式変換手段、として機能させるプログラムを記録した記録媒体が提供される。

【0019】

【発明の実施の形態】

以下、本発明の好適な実施の形態について、添付図面を参照して説明する。

【0020】

図1は、本発明の一実施形態に係るドキュメント管理装置の一部を示すブロック図である。この装置において、保存管理の対象である文書画像データは、画像解析がなされ、解析で得られた情報がレイヤ形式の管理情報として付加される。以下、更に詳細に説明する。

【0021】

101は、紙文書等を読み取り文書画像データを作成するスキャナである。この文書画像データは、読み取られ紙文書の1または複数ページからなるデータを含むことができる。

【0022】

文書画像データは、圧縮処理部104において、画像圧縮される。画像圧縮の方法には、例えば、多値画像の場合は、JPG圧縮を用いることができ、また、2値画像の場合は、MMR圧縮を用いることができるが、他の方法でも問題はない。圧縮された文書画像データは、メモリ110に保存される。

【0023】

ここで、本実施の形態においては、画像データ及びこれに関連するデータを特定の形式(以下、PAF(Page Analysis Format)ともいう。)で保存する。そして、全体を示す文書画像データは、メモリ110上に構成された全体画像保存レイヤに保存される。

【0024】

なお、本実施形態においては、文書画像データを圧縮部104において圧縮して保存することとしたが、圧縮せずに原画像データのままで保存してもよい。この際、データ量が多い場合は、実体データを別のメモリに保存し、メモリ110には、その実体データの場所を特定するリンク情報のみを保存するようにしてもよい。

【0025】

2値化処理部102では、文書画像のデータが多値画像データの場合に2値化処理を行う。これは、引きつづき処理されるレイアウト解析、文字認識処理のために行われるものである。この2値化処理の目的は、画質を向上させることを直接の目的とするのではなく、文書画像に含まれるテキスト部分やピクチャ部分を際立たせることである。2値化の処理として、例えば、多値画像の濃度の平均値を算出し、最適なレベルで2値化を行う方式を用いる。

【0026】

レイアウト解析部103では、2値化処理部102から出力された文書画像データを解析し、所定の分類毎に文書画像に含まれる各領域を識別し、各領域の画像データについての情報を作成する。本実施形態では、例えば、テキスト領域、表領域、ピクチャー(絵、図形、写真等を含む)領域、及び、枠(フレーム)領域を識別し、各領域の画像データについての情報を作成する。領域の分類の識別はこれに限られず、例えば、文書画像中のタイトルが表示された領域を、タイトル領域として識別することもできる。

【0027】

また、このレイアウト解析により得られた情報(各領域の配置情報を含む。以下、レイアウト解析情報という。)を作成し、これを、メモリ110のレイアウト情報レイヤに格納する。レイアウト解析部103での処理には、公知の輪郭線追跡手法などによる処理が適

10

20

30

40

50

用でき、分類・領域識別が可能である。

【 0 0 2 8 】

文字認識処理部 1 0 5 では、テキスト領域と識別された領域の画像に含まれる文字を認識するための文字認識処理を行う。認識されたテキストデータは、メモリ 1 1 0 のテキスト & ベクトル保存レイヤに保存される。

【 0 0 2 9 】

表解析処理部 1 0 6 では、表領域と判別された領域に対して表解析を行う。表解析処理とは、表領域と識別された領域の画像に含まれる表を解析する処理であって、例えば、(1) 表構造の解析 (例えば、行数 X、列数 Y からなる表の構成) (2) 表の枠線の認識、(3) 表内セルの解析などの処理を行い、表解析情報を出力する処理である。この表解析情報のうち、表の構造に関する情報は、メモリ 1 1 0 のレイアウト情報レイヤに、表中の文字に関する情報はメモリ 1 1 0 のテキスト & ベクトル保存レイヤに保存される。ここで、レイアウト情報レイヤとテキスト & ベクトルレイヤは、別のレイヤではあるが、リンクで密接に結びつけられ、関連付けられている。

10

【 0 0 3 0 】

線画像ベクトル変換部 1 0 7 では、レイアウト解析部 1 0 3 で、フレームと呼ばれる大きな枠を含む領域が認識された場合、その枠を表わすデータを所定の形式のベクトルデータ化する処理 (ベクトル変換ともいう。)を行う。

【 0 0 3 1 】

例えば、枠が四角形の場合は、四角形の 4 隅の位置座標を 4 つのベクトルに変換する処理を行う。ただし、角が丸みをおびているなどの所定の場合には、4 つ以上のベクトル情報に変換することもある。この点を図 4 を用いて説明する。

20

【 0 0 3 2 】

図 4 の 4 0 2 は、直線の枠で囲まれた線画である。この枠をベクトル変換すると 4 0 4 に示したように 4 辺で囲まれた枠は、4 つのベクトルの情報に変換される。また、4 0 3 に示すような 4 隅が内側に曲がっている枠の場合は、4 0 5 に示すように 8 つのベクトル情報に変換される。この処理 (上述の 8 つのベクトル情報に変換する処理) は、実際には、画像の枠領域に対して輪郭線追跡処理を行い、その情報に基づき、4 隅のベクトルを追加して持つか否かを判断し、追加するとした場合にこの処理が行われる。その他、線画像ベクトル変換部は、文書画像データ中の一般の線画もベクトル変換される場合、これら線画像のベクトル情報も、フレームの情報同様に、領域の配置に関する情報はレイアウト情報レイヤ、ベクトル情報はテキスト & ベクトル保存レイヤに保存される。

30

【 0 0 3 3 】

部分画像切り出し部 1 0 8 では、レイアウト解析処理部 1 0 3 での処理において認識された各領域について文書画像データの切り出しが行われ、その領域の分類にマッチした画像圧縮がなされる。

【 0 0 3 4 】

例えば、ピクチャ領域の画像データであって、オリジナルの画像データ 1 0 1 が多値の場合、多値画像圧縮である J P E G が選択され、テキスト領域の場合、多値よりも 2 値の画像データとして保持した方が好都合の場合が多く、2 値画像データとし、M M R 圧縮を行い、保存することができる。このように各分類別に適応的に圧縮を行う。最も圧縮の方法は限定されない。

40

【 0 0 3 5 】

圧縮の結果得たデータは、メモリ 1 1 0 の部分画像保存レイヤに保持する。

【 0 0 3 6 】

上述の各部での処理により、メモリ 1 1 0 上にレイヤ構成をもつ文書画像データに関する情報が一つのファイルとして保存される。

【 0 0 3 7 】

この情報の作成において、所定の場合には、特定のレイヤについての情報を削除することができる。例えば、全体画像のデータと部分画像のデータとは、実質的に同じ情報を重複

50

して持つこととなるため、文書画像がテキストや表を含むものではなく、複数のピクチャから構成される場合は、全体画像を削除して、部分画像のみを保存するのも良い。一方、全体画像が複数のテキスト領域から構成される場合は、テキストデータも保存するため、情報の形態が異なり、重複して持っても左程無駄ではない。

【 0 0 3 8 】

以上述べた本実施の形態に係るレイヤ構成をとることにより、所定のレイヤ毎に必要な情報を保持することが可能となった。これは、本レイヤ構成をとる場合の特徴である。図 1 の 1 1 1 は、そのレイヤ構成が視覚的に分かるように、レイヤ構成を模式的に示した図である。

【 0 0 3 9 】

次に、図 5 を用いて、保存される情報について、さらに詳細に説明する。

【 0 0 4 0 】

文書画像データ 5 0 1 の全体に対してレイアウト解析を行うと、タイトル (TITLE)、テキスト (TEXT)、ピクチャ (PICTURE) という分類毎の領域を識別する。それぞれの領域のデータは、5 0 2 に示された木構造でデータ保存される。各領域のデータは、リージョン (REGION) と呼ばれる矩形情報を持ち、その矩形の配置情報 (不図示) も保持する。

【 0 0 4 1 】

図 5 の 5 0 1 では 3 つのリージョン (REGION) が保持され、タイトル (TITLE)、テキスト (TEXT)、ピクチャ (PICTURE) 等の分類の情報が保存されている。

【 0 0 4 2 】

タイトルリージョンに関しては、その位置関係、文字の大きさ等の情報に基づき、テキスト分類のサブ分類として Title という情報を付加している。

【 0 0 4 3 】

5 0 2 のテキストリージョンの木構造は、イメージリンク (image_link)、イメージインフォ (image_info)、コード (code) といった子の木を有する構造である。これには、テキスト領域の部分画像と、その配置情報、及び、文字認識処理により得られるテキストデータが付加される。実際どのようにリンクされるかは、後述する。

【 0 0 4 4 】

同様に、ピクチャリージョンについては、イメージリンク (image_link)、イメージインフォ (image_info) が付加される。こちらには、コード (code) という子はない。

【 0 0 4 5 】

図 6 に、図 5 で説明した形態の情報保存を、実際の文書画像データに当てはめた例を示す。図 7 には、さらに詳細な記述方法を示した。本実施の形態における情報の作成は、XML (eXtensible Markup Language) を用いて行われる。図 7 の 7 0 1 に、処理対象の文書画像データを示す。

【 0 0 4 6 】

本文書画像データは、レイアウト解析処理部での処理の結果、タイトル (TITLE)、テキスト (TEXT)、ピクチャ (PICTURE) の領域に認識される。各分類の領域のデータは、さらに必要な情報処理が施され、情報が付加され、その結果、7 0 2 に示すタグ付きの文書フォーマットの画像情報が生成される。

【 0 0 4 7 】

各タグには、それぞれ意味付けがされている。ここで意味付けとは、木構成において、ある特定のタグには、決められた意味内容を有するタグが配置されるということである。

【 0 0 4 8 】

例えば、<PAF_REG> のタグは、終了タグ </PAF_REG> の間に、テキストを保持する <TXT> タグや画像を保持する <PICTURE> タグを配置することができる。

【 0 0 4 9 】

例えば、

10

20

30

40

<PAF_REG>

<TXT>This is a TITLE</TXT>

<PICTURE>..¥Meeting¥document1.mmr</PICTURE>

</PAF_REG>

のようにである。すなわち、<PAF_REG> タグは、子の木として、<TXT> と <PICTURE> が配置可能という事になる。また、<TXT> には、<BLK_IMG> タグと <BLK_OCR> タグを保持することが可能である。

【 0 0 5 0 】

例えば、

<PAF_REG>

<TXT>This is a TITLE

<BLK_IMG href=..¥Meeting¥配布資料.mmr/>

<BLK_OCR href=..¥Meeting¥OCRファイル />

</TXT>

</PAF_REG>

である。

【 0 0 5 1 】

以上では、保存管理の対象である文書画像データをそのまま保存するのではなく、文書画像データについての所定の解析情報も付加して保存することが可能な構成について説明してきた。

【 0 0 5 2 】

次に図 3 を用いて、作成・保存した情報のうち、所定の情報を取出す方法について説明する。

【 0 0 5 3 】

図 3 のメモリ 3 0 1 は、図 1 に示すメモリ 1 1 0 と同じであり、上述した手法で各情報が保存されている。

【 0 0 5 4 】

抽出処理部 3 0 2 は、出力形式選択指示 3 0 3 に従って、所望のレイヤから必要な情報を抽出する。この抽出は、従来 1 ページ単位でしか扱えなかったのに対して、文書画像データの中の表のみを取出すといった文書内コンテンツレベルでの抽出も可能である。これは、先に説明したように、本実施形態では、文書内のコンテンツにそれぞれタグを設定しているため、そのタグを頼りに抽出することが可能となったことによる。例えば、図 6 で示した文書画像データの "TITLE" の部分画像を抽出したいと指示すれば、その部分のみ持ってくる事ができる。この文書内コンテンツは、1 ページのコンテンツのみならず複数ページに渡るコンテンツを持ってくることも可能である。

【 0 0 5 5 】

出力形式変換部 3 0 4 は、文書全体あるいは、文書内コンテンツを抽出したもののに対して、所望のアプリケーションで読み込める形式に変換する処理部である。これは、引き続き次の処理部とともに説明を行う。

【 0 0 5 6 】

図 3 の 3 0 5 は、出力形式変換部 3 0 4 で変換された、テキスト (Text) 中心文書である。これは、ワープロ文書のように文章主体の書類の中にピクチャが存在するような書類である。このような出力形式は、ワープロソフト等のアプリケーションソフトのために設けられる。例えば、マイクロソフト社のワード (商標) のなどのソフトウェア向けの形式である。テキスト (Text) 中心文書 3 0 5 を、GUI を用いて表示したときの例を 3 0 9 に

10

20

30

40

50

示した。この場合、テキスト領域の画像データとしてはテキストデータとして保存されているデータを用い、ピクチャ（PICTURE）領域についてはそのまま部分画像のデータを用いる。表領域については、その表構造を表解析情報に基づき作成し、その表内セルのデータを文字情報に基づき作成する。

【 0 0 5 7 】

画像中心文書 3 0 6 は、3 0 5 と同様に、マイクロソフト社のパワーポイント（PowerPoint（商標））のようなソフトウェア向けの出力形式である。1 つ 1 つの文書内コンテンツの表示には、基本的に部分画像を用いる。部分画像の集まりにより、文書を構成する。しかしながら、利用者の指示により、ある部分はテキストデータを渡してもよい。画像中心文書 3 0 6 を、GUI を用いて表示したときの例を 3 1 0 に示した。

10

【 0 0 5 8 】

表中心文書 3 0 7 は、3 0 5 等と同様にマイクロソフト社のエクセル（商標）のようなソフトウェア向けの出力形式である。表ソフトへ挿入するデータは、上述の表解析処理結果が表形式情報に変換されてから渡される。表中心文書 3 0 7 を、GUI を用いて表示したときの例を 3 1 1 に示した。

【 0 0 5 9 】

また、要約文書 3 0 8 は、管理情報付き画像データ 3 0 1 のテキスト & ベクトル保存レイヤから抽出されたテキスト情報に対し、要約処理がなされた結果得られる、簡略化した文書をいう。要約処理には、文章についてなされる形態素解析などの公知の手法が適用でき、文章相互のつながりや段落の切れなどの情報に基づいて行う、いくつかの手法が知られている。そのような処理を行う、要約処理エンジンにより、要約が作成される。3 1 2 に、要約文章に画像のサムネイル画像を付加して作成されたレポート文書を示す。

20

【 0 0 6 0 】

以上説明したように、本実施形態では、利用者が望む所定の文書内コンテンツを抽出することが可能となると共に、画像と文字コードが混在する場合でも表示することが可能である。さらに、文書画像データを通常のファイル管理におけるファイルの状態で保存し、電子文書について用いるのと同様な全文検索ツールにより検索が可能となる。

【 0 0 6 1 】

次に、図 8 は、文書画像データを上述した手順で複数ページ保存し、その文書画像データと同列で電子文書も統合管理する電子バインダを表す図である。

30

【 0 0 6 2 】

紙文書 8 0 1 は、スキャナ等を用いて読み取られて文書画像データが得られる。その際、上述の処理により、メモリ 1 1 0 のレイヤ構成のフォーマットで、各レイヤ別、矩形領域の情報が保存される。

【 0 0 6 3 】

電子文書 8 0 2 は、ワープロ等により電子情報として保持されていた文書である。この電子情報は、本ドキュメント管理に適したフォーマットに変換され、Clear File 8 0 3 が得られる。この処理は、XML 保存形式で保存することにより実現される。

【 0 0 6 4 】

一般に、電子文書には、文字情報、ベクトル情報、PICTURE、表といった属性別に情報が保存されている。しかしながら、電子文書の形式は、作成されたアプリケーションに依存するため、他のアプリケーションによって作成された電子文書の保存形式は不明であった。そのため、他のアプリケーションによって作成された電子文書の情報を取出すことは困難であった。しかしながら、XML 形式は、公開性やドキュメントを人間が読むことを保証する可読性のあり、独自性を排除することになった。また前記属性は、共通化された表現で記述されつつある。8 0 3 を Clear File と名称付けたのは、文書画像データや独自形式電子文書が共通の土台で扱えるようになる意味合いで付けたものであり、Clear File 8 0 3 は、複数ページのドキュメントフォーマットで作成される。

40

【 0 0 6 5 】

8 0 4 に、これらを束ねた電子バインダを示した。従来、電子文書 8 0 2 は、形式が異なる

50

るため同列で扱うことはできなかった。しかしながら、両者とも、文書を管理する前記タグを用いることにより公開性があることおよびXML記述という同一の表現を用いていることから、統合が可能でとなった。

【0066】

図9に、電子バインダの木構造の例を示した。901内のSHEETとは、1ページの文書画像データである。この中に解析情報であるレイアウト記述、及び文字認識結果が格納される。

【0067】

次に、下のDocumentというタグは、電子文書からの情報の保存形式である。この電子文書には、レイアウト記述とその中身である各要素別情報が保管されている。例えば、表について、枠線はVector記述で保持され、表の中の文字はCode情報で保持されている。木の構造は異なるが、文書内の要素別に記述されているため、文書内コンテンツ管理が可能となる。

10

【0068】

図8において、電子バインダを開いて表示した例を805に示した。805の右側に各保存されている文書のサムネイル画像を示し、左側にバインダ内から所望のコンテンツを抽出して表示した例を示す。ここで、文書内の各コンテンツの各属性が管理されているため、例えば、表だけを抽出したりすることも可能となる。

【0069】

なお、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（例えば、複写機、ファクシミリ装置など）に適用してもよい。

20

【0070】

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のプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、その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み出し実行することによっても、達成されることは言うまでもない。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。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働しているオペレーティングシステム（OS）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ことは言うまでもない。

30

【0071】

さらに、記憶媒体から読み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カ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、その機能拡張カ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ことは言うまでもない。

40

【0072】

本発明を上述記憶媒体に適用する場合、その記憶媒体には、先に説明したフローチャートに対応するプログラムコードが格納されることになる。

【0073】

【発明の効果】

以上説明したように、本発明によれば、管理される画像に関連する種々の情報を有効に活用することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係るドキュメント管理装置の一部を示すブロック図である。

50

【図 2】(a) 及び (b) は、文書画像データをデータベースに登録する従来の処理の流れを示した図である。

【図 3】本発明の一実施形態に係るドキュメント管理装置の一部を示すブロック図である。

【図 4】線画像ベクトル変換部 107 で、フレームをベクトル変換する方法についての例を示す図である。

【図 5】文書画像データの中の各分類別に情報を保存する際の保存形態についての例を示す図である。

【図 6】実際の文書画像と文書画像データの保存形態を照らし合わせた例を示す図である。

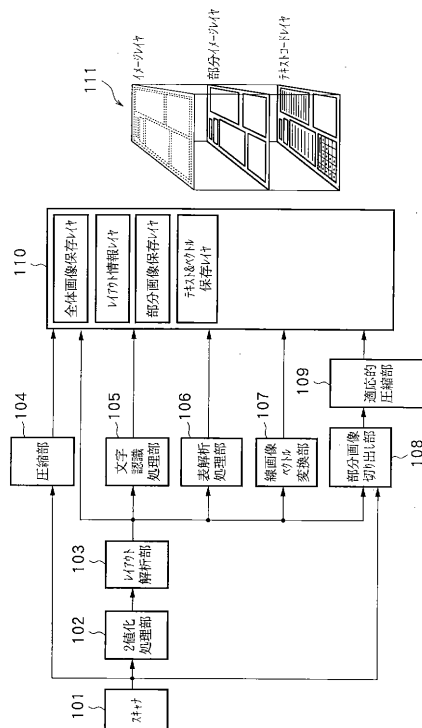
【図 7】実際のドキュメントフォーマットの例を示す図である。

【図 8】電子バインダの例を示す図である。

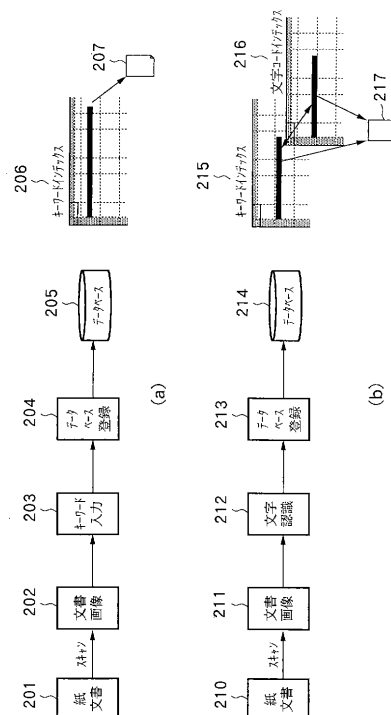
【図 9】電子バインダの木構成例を示す図である。

10

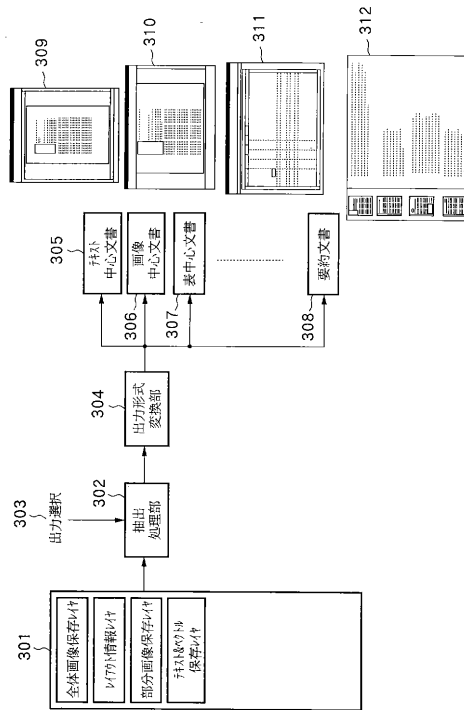
【図 1】



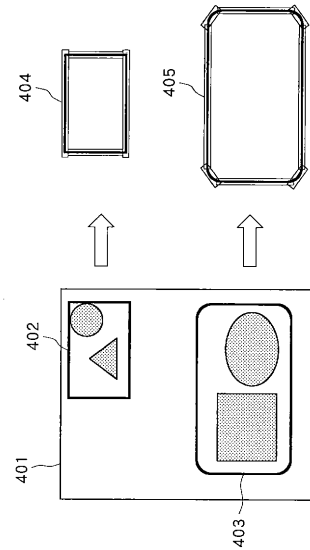
【図 2】



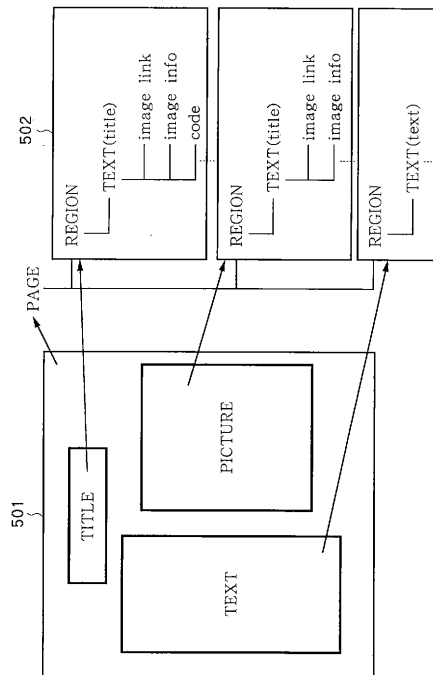
【図 3】



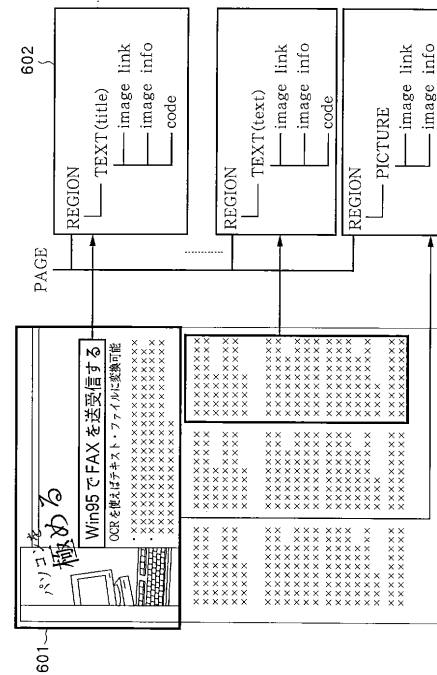
【図 4】



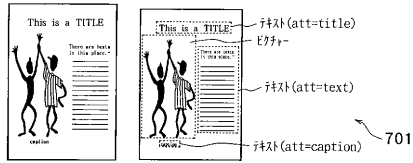
【図 5】



【図 6】



【 図 7 】



simple full complete example (except embedded part)

```
<?XML version="0.0">
<!-- The following is FAMIL example for Fig 4.8 -->
<PAF version="2.0" type="pdf" app=" "
  <!-- "F_INFO" is META data for all of a document (or image) -->
  <PAF_INFO>
    <CREATOR>.../</CREATOR>

  </PAF_INFO>

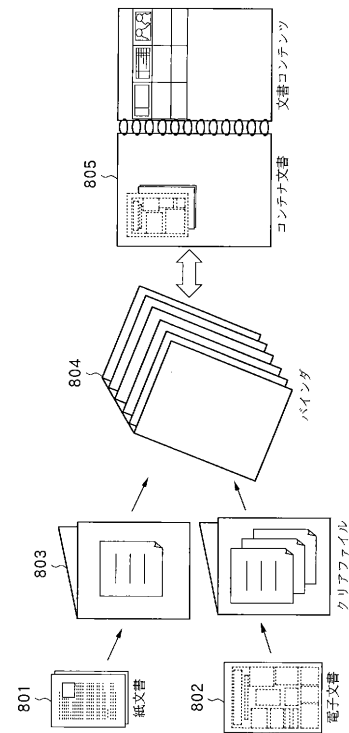
  <!-- Content of "GE" is information in only one page -->
  <PAGE_page="error"/>
  <!-- GE_REGION is META data for this page -->
  <PAGE_INFO n_block="5" width="2000" height="3000" doi="400"
    n_bwimg="1" ..._gimpr">

    <!-- Next "F_REGION" represent info in Fig 4.9 -->
    <PAF_REG id="ft" op="0" opid="200" eight="00" order="1"
      <TXT att="info" action="region" _horizontal">
        <BLK_IMG href="#foot/compress="mm/mml:left","mmblyWt="..."
          imgId="..." mgltype="..." mgDPI="..."
        <BLK_OCR href="..." orId="..."
          This is a TITLE
        </TXT>
      </PAF_REG>

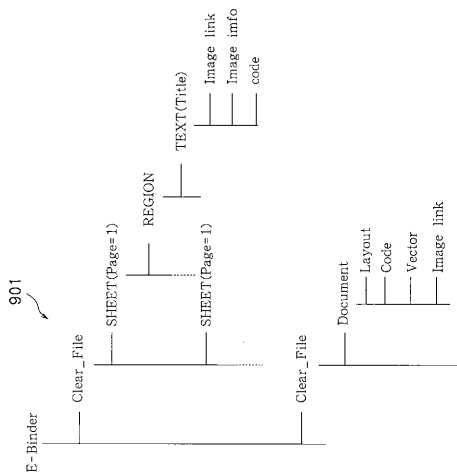
      <!-- Next "GE_REGION" is 'ummy frame' in Fig 4.10 -->
      <PAF_REG id="efts" op="0" opid="0" eight="400" order="1"
        <VRT n_child="1"
          <!-- "F_REGION" is 'cture' in Fig 4.9 surrounded by virtual frame -->
          <PAF_REG id="rect" op="0" opid="0" eight="200" order="1"
            <PICTURE overlap="0" child="1"
              <BLK_IMG href="#foot/compress="png/mgl:left","mmblyWt="..."
                imgId="..." mgltype="..." mgDPI="..."
            </PICTURE>
          </PAF_REG>
          <!-- "F_REGION" is 'action' in Fig 4.9 surrounded by virtual frame -->
          <PAF_REG id="efts" op="0" opid="700" eight="00" eight="0" order="1"
            <TXT att="action" action="region" _horizontal">
              <BLK_IMG href="#foot/compress="mm/mml:left","mmblyWt="..."
                imgId="..." mgltype="..." mgDPI="..."
            <BLK_OCR href="..." orId="..."
              caption
            </TXT>
          </PAF_REG>
        </VRT>
      </PAF_REG>

    <!-- Next "GE_REGION" is 'ext' in Fig 4.9 -->
    <PAF_REG id="efts" op="0" opid="0" eight="400" order="1"
      <TXT att="ext" action="region" _horizontal">
        <BLK_IMG href="#foot/compress="mm/mml:left","mmblyWt="..."
          imgId="..." mgltype="..." mgDPI="..."
        <BLK_OCR href="..." orId="..."
          There are texts in this place. ....
        </TXT>
      </PAF_REG>
    </PAGE>
  </PAF>
```

【 図 8 】



【 図 9 】



フロントページの続き

(72)発明者 高岡 真琴

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

審査官 村松 貴士

(56)参考文献 特開平10-320412(JP,A)

特開平11-120185(JP,A)

特開昭60-140471(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00