(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3711636号
(P3711636)
(45)発行日　平成17年11月2日(2005.11.2)　(24)登録日　平成17年8月26日(2005.8.26)

(51)Int.Cl.[7]　F I
G06F 17/30　G06F 17/30　１７０Ｂ
G06F 17/30　３２０Ａ

請求項の数 2　(全 15 頁)

(21)出願番号　特願平8-177282
(22)出願日　平成8年6月18日(1996.6.18)
(65)公開番号　特開平10-3483
(43)公開日　平成10年1月6日(1998.1.6)
審査請求日　平成14年4月18日(2002.4.18)

(73)特許権者　000005496
富士ゼロックス株式会社
東京都港区赤坂二丁目１７番２２号
(74)代理人　100085899
弁理士　岩上　昇一
(74)代理人　100093182
弁理士　南野　貞男
(72)発明者　糸乗　勝彦
神奈川県足柄上郡中井町境４３０　グリーンテクなかい　富士ゼロックス株式会社内

審査官　深津　始

最終頁に続く

(54)【発明の名称】情報検索装置および方法

(57)【特許請求の範囲】

【請求項１】

文書画像内の一部領域を検索用パターンとして指定する指定手段と、

前記指定手段によって指定された検索用パターンにより前記文書画像内を検索し、前記指定された検索用パターンと一致する画像パターンが検出されたとき、その画像パターンを含む領域を関連領域として抽出する関連領域検索手段と、

その関連領域検索手段により抽出された領域に対する文字認識を行い、検索用情報を得る文字認識手段と、

複数のテキストあるいは画像を蓄積し、前記検索用情報を用いて蓄積情報を検索する情報蓄積検索手段とを備えたことを特徴とする情報検索装置。

【請求項２】

文書画像内の一部領域を検索用パターンとして指定する指定手段と、前記指定手段によって指定された検索用パターンにより前記文書画像内を検索し、前記指定された検索用パターンと一致する画像パターンが検出されたとき、その画像パターンを含む領域を関連領域として抽出する関連領域検索手段と、その関連領域検索手段により抽出された領域に対する文字認識を行い、検索用情報を得る文字認識手段と、複数のテキストあるいは画像を蓄積し、前記検索用情報を用いて蓄積情報を検索する情報蓄積検索手段とを備えた情報検索装置による情報検索方法であって、

前記指定手段が、文書画像内の一部領域を検索用パターンとして指定するステップと、

関連領域検索手段が、前記指定手段によって指定された検索用パターンにより前記文書

画像内を検索し、前記指定された検索用パターンと一致する画像パターンが検出されたとき、その画像パターンを含む領域を関連領域として抽出するステップと、
　前記文字認識手段が、前記抽出された領域に対する文字認識を行い、検索用情報を得るステップと、
　複数のテキストあるいは画像を蓄積した情報蓄積検索手段が、蓄積情報を前記検索用情報を用いて検索するステップと
を備えたことを特徴とする情報検索方法。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、データベースに登録されている情報や同一文書内に記載されている関連ある情報を検索する装置および方法に関する。本発明は、具体的には、紙面上に印刷された文書を画像として読み込み、簡便な方法で指定した項目に関連した情報を検索、あるいは検索・表示する情報検索装置および方法に関する。
【０００２】
【従来の技術】
従来、データベースの検索は適当なキーワードを機械に指定することにより行われている。また、関連ある情報間にリンクを張り、必要な部分を指定すると自動的に詳細な情報、あるいは関連性の強い情報を提示する仕組みが提案されている。これらの従来技術は、あらかじめ用意されたコード情報（キーワードであればテキスト、情報間のリンクであればリンク先のアドレス等）を解析することによって実現されている。
【０００３】
このような仕組みは、テキストのようなコードデータを対象として提案され、画像データに付いても同様な仕組みを提供するためにいくつかの方法が提案されている。その１つとして、画像データを登録する時に、画像データと共にその画像の内容を表すテキストデータを同時に登録し、検索時にはこのテキストデータを使用するものである。この画像が文書であれば、画像の内容を解析して、文字認識を行い、その認識結果を登録し、全文検索を行うことになる（特開平７－８５０７４号公報、特開平６－２８２５８２号公報）。また、このテキストデータを利用することにより、同一文書中の単語の検索が容易となるので、文書中の特定の単語を検索して、関連する情報が記載されている部分を容易に得ることができる。
【０００４】
また、このように画像全体の内容をテキストデータとして変換するのではなく、利用者が適当な部分を選択し、その選択した一部分のみを文字認識し、その結果を用いて必要なデータをデータベース中から検索するという方法も提案されている（特開平７－２１０５７６号公報）。
【０００５】
【発明が解決しようとする課題】
しかしながら、データベースなど大量に蓄積されている情報は、コード化されているものばかりではなく、画像として蓄積されているものも多く、また、利用者が出版物など個人的に登録するような場合も、画像で入力することの方がコード化して登録する（文書入力装置で入力し直す、あるいは文字認識装置で入力する）よりは簡単に作業をすることができる。
【０００６】
もともと文書は１文書中にいろいろな方法で関連情報を記載している。例えば、論文などでは参考文献として、関連する情報を詳細に掲載している文献を紹介し、しかも本文中には関連性のある適当な部分で参照すべき参考文献の番号を載せている。読者はこの番号を頼りに巻末に記載されている参考文献集より、適当な文献を捜すことができる。
【０００７】
また、関連情報として、書中の注釈や図／表番号などは、本文中に記載されている内容の

詳細を知らせる重要な情報である。例えば、『図ＸＸに示しているように、～は～』のような記述では、図を見ながら読みすすめると理解が深まる。また、注釈においても関連性のある内容を数字（１、２、・・・・）あるいは記号（＄、§、‡、・・・・）で表し、頁の下部にけい線で区切った脚注欄に記載したり、章末／巻末に記載したりといった方法が採られている。

【０００８】

このような文書中しかも画像中の関連情報をたどるには、画像を解析して文字認識を行い、コード情報に置き換えた上で従来技術を用いて、検索を行うといった手段があるが、この方法では画像とコード情報との関連性が失われるため、画像上を直接指定して必要な情報を検索することができず、あらたにキーワードを入力して情報を検索する必要があった。

【０００９】

本発明は、文書画像上の領域を直接指定することにより、関連情報を求め、この関連情報から検索用情報を得て、その検索用情報によりデータベース等の検索を行う情報検索装置及び方法を実現することを課題とする。

【００１３】

【課題を解決するための手段】

本発明（請求項１）の情報検索装置は、文書画像を表示するための画像表示手段（１５１）と、文書画像内の一部領域を検索用パターンとして指定する指定手段（１５２）と、前記指定手段によって指定された検索用パターンにより前記文書画像内を検索し、前記指定された検索用パターンと一致する画像パターンが検出されたとき、その画像パターンを含む領域を関連領域として抽出する関連領域検索手段（１５３）と、その検索手段により抽出された領域に対する文字認識を行い、検索用情報を得る文字認識手段（１５４）と、複数のテキストあるいは画像を蓄積し、検索用情報を用いて検索する情報蓄積検索手段（１５５）を備えたことを特徴とする。この情報検索装置の作用においては、画像入力手段により入力した画像中の一部領域を、検索用パターンとして指定手段により指定する。関連領域検索手段により、その指定した検索用パターンと一致する画像パターンを、文書画像中から検出し、その検出した画像パターンに関連付けられた領域を抽出する。文字認識手段は、その抽出された領域に対して文字認識を行い、認識結果の文字列から検索用情報（例えば、認識結果の文字列をキーワードとする検索式）を生成し、情報蓄積検索手段の検索に用いる。本発明によれば、キーワードなどの検索用情報の入力などの操作を省略することができ、操作性のよい情報検索装置を実現できる。また、本発明によれば、検索した結果が画像情報であろうとテキストデータであろうと、その操作方法は同じなので、検索結果の情報の形態を意識することのない、操作性のよい情報検索装置を提供できる。

【００１５】

【発明の実施の形態１】

本発明の一実施の形態の情報検索装置について説明する。図１は、その情報検索装置の基本構成図である。この装置は、２値画像、特に文書画像を入力する画像入力部１０と、入力した文書画像を表示する画像表示部１１と、入力した文書画像を構成する領域を解析する文書領域解析部１２と、入力した文書画像の書式を指定する書式指定部１３と、入力した文書画像の一部領域を指定する領域指定部１４と、指定した領域内の記号と同一の記号を文書中の別の領域、あるいは書式指定部１３によって指定された領域から検出する記号検出部１５と、前記記号検出部１５で検出された記号を含む領域を、入力した文書画像から切り出す文書領域切り出し部１６と、画像入力部１０、画像表示部１１、文書領域解析部１２、書式指定部１３、領域指定部１４、記号検出部１５、及び文書領域切り出し部１６の動作を制御する中央制御装置１７、その中央制御装置１７が各部を制御するためのプログラムなどを格納する記憶装置１８を備えている。

【００１６】

以下、図１を用いて各部の詳しい説明を行う。

まず、画像入力部１０により画像を入力する。これは、紙面に印刷された画像をディジタ

ルデータとして入力するために用いるイメージスキャナのような装置であっても、外部のデータベースに登録されている画像を入力する構成であってもよい。

【００１７】

入力された画像は、画像表示部１１に表示される。利用者は、この画像表示部１１を通して入力された画像や、作業の状況を知ることができる。

【００１８】

また、文書領域解析部１２により、入力された文書画像を構成する複数の領域を抽出する。例えば、図２に示すように、検出される領域は文字領域２１、図形領域２２、けい線領域２３などである。この文書領域解析部１２においては、まず画像中の画素連結成分を抽出することから解析が始まる。

【００１９】

画素連結成分は、一般にラベル付け処理（例えば、通産省工業技術院監修、協同システム開発（株）発行“画像処理サブルーチン・パッケージＳＰＩＤＥＲ　ＵＳＥＲ’Ｓ　ＭＡＮＵＡＬ”，ＰＰＩＩＩ－３４～ＩＩＩ－３５参照）として知られる処理によって抽出することができる。ラベル付け処理は、画素値が１である画素（黒画素）について、図３に示すように画像中の注目点（ｉ、ｊ）の画素の周辺（ｉ－１、ｊ－１）、（ｉ、ｊ－１）、（ｉ＋１、ｊ－１）、（ｉ－１、ｊ）の画素の画素値及びラベル値を調べ、周辺の画素に１以上のラベル値が与えられていれば、画素（ｉ、ｊ）にそのラベル値と同じラベル値を与え、周辺の画素の画素値がすべて０（白画素）であれば、新しいラベル値を画素（ｉ、ｊ）に与える処理である。もし、周辺の画素に複数のラベル値が与えられている時は、その中の最小値をラベル値として画素（ｉ、ｊ）に与えて、周辺画素に与えられているすべてのラベルが同一のラベルであることを画素連結テーブル４１に記録しておく。例えば、周辺の画素にラベル値２とラベル値３のラベルがあったときには、注目画素のラベルをラベル値２に設定し、ラベル値２とラベル値３が連結した画素のラベルであることを画素連結テーブル４１に記録する。最後にこの画素連結テーブル４１を参照して、すべてのラベルを調整し直して、連結している領域ごとに異なるラベルを与える。ラベル付けの様子を図４（ａ）及び（ｂ）、図５に示す。図４（ａ）では、画像の走査によりラベルを各画素に与えた結果を示し、同図（ｂ）はその過程で生成される画素連結テーブル４１を示す。図４（ｂ）の画素連結テーブル４１によると、ラベル値２とラベル値３を持つ画素は連結していることが分かる。したがって、この情報を元に画素値を更新して図５のように、ラベル値３を持つ画素をすべてラベル値２に更新する。

【００２０】

このように生成されたラベル画像に対して、各ラベル毎にその外接矩形６１、６２を求めて、この外接矩形によって画素連結成分を表す。その結果、図６（ａ）に示す文書画像は図６（ｂ）のように外接矩形６１、６２によって表現される。次に、この外接矩形の面積、縦横比によって各外接矩形の属する領域を判定する。例えば、文書画像の大半を占める文字の外接矩形６１より図形を表現する画像の外接矩形６２の方が面積が大きい。また、けい線は細長い線分なので、外接矩形の幅を高さで割った値は極端に大きな値を取る。このような処理は適当な閾値を設けることで簡単に行うことができる。図形領域として判定するための閾値Ｇ（面積）は、本文中で用いられている最大文字サイズが２４ポイント、最小文字サイズが６ポイントであれば、Ｇ＝７０．５（平方ｍｍ）、またけい線は３文字以上の長さを持つと仮定すると、けい線領域を判定する閾値Ｒ（縦横比）はＲ＝１２と設定することができる。したがって、閾値Ｇより大きい面積を持つ外接矩形を図形領域、閾値Ｒより大きい縦横比を持つ外接矩形をけい線領域とする。

【００２１】

更に、ここで求めた図形領域とけい線領域に対して処理を行う。図形領域に関しては、図を説明するキャプション部分も図形領域に含まれるように領域を拡張する。図形領域の拡張は、閾値Ｈを設定して図形領域に近い文字矩形を含むように図形領域を拡張する。この閾値Ｈは、文書中で使用される最大文字サイズを２４ポイントとすると、その１／３の値Ｈ＝２．８ｍｍとして設定することができる。したがって、文書画像中のすべての外接矩

形と図形領域の間の距離を計算して、閾値Ｈより近接している矩形を図形領域として統合する。この処理により、図形領域内に、そのキャプションを取り込むことができる。

【００２２】

次に、けい線領域に関しては、けい線の下側には脚注が存在すると判断して、けい線領域２３より下側にある外接矩形をすべて統合して脚注領域７１（図７）とする。ただし、けい線領域は文書画像中のあらゆる部分（文書の最上部、本文中など）に記入されていると考えられるので、ここでは脚注として可能性の高い、文書画像の下側１／２に含まれるけい線領域を対象として、その下側に存在する外接矩形を統合して１つの領域７１として生成する（図７）。

【００２３】

図２では、これらの外接矩形をさらに統合を進めて、文字だけの領域２１を検出しているが、本発明では文書全体に対するこれ以上の統合処理は必要ないので、外接矩形の統合処理は文書領域解析部１２では行わない。文書領域解析部１２で得られたすべての外接矩形の位置と大きさ、それとけい線領域、脚注領域と図形領域の位置を記憶装置１８に記憶する。

【００２４】

次に、書式指定部１３によって入力した文書画像の注釈位置を指定する。注釈の書かれている場所としては、脚注あるいは節末／章末があるので、入力された文書画像ではどちらに書かれているかを、利用者に指定してもらう。このとき、脚注が指定されると、記憶装置１８に記憶されている脚注領域を注釈領域として記憶し直しておく。

【００２５】

また、節末／章末が指定された時は、さらに注釈領域を利用者に指定をしてもらう。例えば、図８のようにマウスカーソル８１を用いて注釈の開始位置８２と終了位置８３を指定することができる。まず、マウスカーソル８１を注釈の開始位置８２に移動してから、マウスボタンを押し下げることで注釈の開始位置を指定し、次にマウスカーソル８１を注釈の終了位置８３に移動して、同様にマウスボタンを押し下げることで注釈の終了位置を指定する。ここで指定された注釈開始位置８２と注釈終了位置８３の間に存在する、文書領域解析部１２で抽出した外接矩形を統合し、注釈領域として記憶装置１８に記憶する。

【００２６】

以上のようにして記憶装置１８に記憶されている入力画像の解析結果を用いて、文書画像内の注釈番号あるいは図番号と関連する画像を検索して表示を行う。以下に、注釈番号あるいは図番号を指定して、関連する画像を表示するまでの手順を説明する。

【００２７】

まず、利用者が領域指定部１４により画像表示部１１上に表示されている画像の一部領域９１を指定する。指定の方法は、例えば図９に示すようにマウスカーソル８１を使用して、画像上の一部分を囲むようにする。更に詳しくは、指定したい領域の左上の角にマウスカーソル８１を移動してマウスボタンを押し下げ、そのまま指定したい領域の右下までマウスカーソル８１を移動しマウスボタンを離すことで領域９１を指定することができる。

【００２８】

領域指定部１４により指定された領域内の画像は切り出され、記憶装置１８に記憶される。記号検出部１５では、領域指定部１４で指定された領域内の画像パターンの表す記号に一致する記号を領域解析部１２で抽出された図形領域内の画像と注釈領域内の画像から検出する。このとき、領域指定部１４で指定された領域内に、参照を示す記号が存在する保証はないので、領域指定部１４では、簡単な文字認識を行い、参照を示す記号以外であれば警告して利用者に適当な領域を指定するように促すことも可能である。ここで、簡単な文字認識とは、参照を表す記号が数種類の文字（数字、”図”、”表”、”＊”、”＄”、等の文字）に限定されるため、これらの文字だけを対象とした小さな認識辞書による文字認識である。

【００２９】

具体的に記号の検出は、記憶装置１８に記憶された領域指定部１４により指定された領域

の画像を、検索対象となっている文書領域解析部１２によって抽出された図形領域又は注釈領域の指定の領域内の画像に重ね合わせ、画素値が一致する画素数をカウントする。記憶装置１８に記憶されている記号画像の全画素数に対する画素値が一致した画素数の割合を画像の一致度として、図形領域あるいは注釈領域全体に対して順次重ね合わせて、一致度の最も高い部分を検出することで記憶装置１８に記憶された記号と同一の記号を検出する。また、文書領域解析部１２で抽出した画素連結成分を利用して、図形領域あるいは注釈領域内の画素連結成分ごとに前記の検出方法を行うことで同様に記号位置を検出することもできる。このとき、図形領域内の図形として判断される大きさ（面積）を持つ画素連結成分の外接矩形を、検索対象から除外することで処理時間を短縮することができる。

【００３０】

しかし、領域指定部１４で指定された領域内の文字サイズと、検索対象となる図形領域あるいは注釈領域内の文字サイズが異なる場合には、上記の記号検出部１５ではうまく記号を検出できない場合がある。そのため、記憶装置１８に記憶された図形領域あるいは注釈領域内の画素連結成分の外接矩形のサイズと一致するように、記憶装置１８に記憶された指定領域の画像を変形することで、この問題を解決することができる。

【００３１】

上記のようにして検出された記号は、その位置と検出された位置を含む領域を記憶装置１８に記憶する。文書領域切り出し部１６では、記憶装置１８に記憶された、記号検出部１５で検出された記号位置を含む図形領域あるいは注釈領域を文書画像２０から切り出し、さらに記号検出位置に記号検出位置マーク１０１を付加した参照位置画像１０２を生成する。この文書領域切り出し部１６で生成した参照位置画像１０２を指定領域の近傍に、画像表示部１１によって表示を行う。この様子を図１０では図番号が指定された場合、図１１では注釈記号が指定された場合を示す。

【００３２】

以上の操作により、入力した文書画像中に記載されている関連情報を取り出し、表示することができる。また、表示される関連情報は画像なので、関連情報内にさらに参照する情報が存在すれば、同様の処理により表示することも可能である。

【００３３】

本実施形態では、簡単に説明するためにあたかも１画像について操作しているように説明しているが、実際には複数の画像で１つの文書を構成している場合が多い。そのため、複数の画像を１つの情報の単位として扱う仕組みが必要である。このような仕組みに関しては、１つの画像ファイルの先頭に幾つの画像が格納されているかの情報を記述して、１つのファイルに複数の画像を格納する方法がある。

【００３４】

また、図１２のように１つの文書がどの画像から構成されているかを示すために、文書名、頁数、画像ＩＤ等を有する文書画像インデックス情報１２１を作成し、この文書画像インデックス情報１２１によって画像の表示、検索を行うこともできる。なお、この方法によれば、注釈の記載されている頁番号を文書画像インデックス情報１２１に登録することで、書式指定部１３を省略することができる。

【００３５】

このように複数の画像を１つの文書として扱い可能とすることで、指定した図番号あるいは注釈記号の参照場所が、別の頁に存在していても、必要な参照情報を即座に表示することが可能となる。図１３、図１４に別頁に記載されている参照情報を表示している例を示す。図１３は図番号の領域９１を指定し、別頁の対応する図面の画像１０２を表示している様子を示している。図１４は注釈の領域９１を指定し、別頁の対応する注釈の画像１０２を表示している様子を示している。

以上の操作により、複数画像によって構成され、同一画像中にない関連情報を取り出し、表示することができる。

【００３６】

以上に説明した本発明の第１の実施の形態によれば、画像中の一部を指定して、指定した

領域に記述されている記号に関連する情報を記述している領域を切り出して、指定した領域の近傍に表示するようにしたので、テキストデータと同様の操作性で文書中の情報を検索することができると共に、関連情報を分かり易い位置に自動的に表示できる操作性のよい情報検索装置を提供できる。

【００３７】

更に、本発明の第１の実施の形態によれば、複数の画像を１つの情報単位として扱うことができるようにしたので、実際の紙の文書を読むのに近い感覚で扱うことのできる操作性のよい情報検索装置を提供できる。

【００３８】

【発明の実施の形態２】

図１５は、本発明の第二の実施の形態による情報検索装置の基本構成図である。この情報検索装置は、２値画像を入力する画像入力部１５０と、入力した画像を表示する画像表示部１５１と、入力した画像の一部領域を指定する領域指定部１５２と、指定した領域内の記号と同一の記号を画像中のほかの部分から検出し、検出した記号に対応する関連領域を求める関連領域検索部１５３と、検出した記号に対応する関連領域内の文字を認識する文字認識部１５４と、文字認識の結果をキーワードとして情報を検索する情報蓄積検索部１５５と、画像入力部１５０、画像表示部１５１、領域指定部１５２、関連領域検索部１５３、文字認識部１５４、及び情報蓄積検索部１５５等の動作を制御する中央制御装置１５６と、中央制御装置１５６が各部を制御するためのプログラム等を格納する記憶装置１５７を備えた構成を有している。

【００３９】

なお、本実施形態では論文を対象としてその参考文献を検索するように記述されているが、論文ばかりではなく一般の文書で、記号を使って他文書の情報との対応を取って関連づけているものについても同様の方法で他文書の情報の検索が可能である。

【００４０】

以下、図１５を用いて各部の詳しい説明を行う。まず、画像入力部１５０により画像を入力する。これは、紙面に印刷された画像をディジタルデータとして入力するために用いるイメージスキャナのような装置であっても、情報蓄積検索部１５５のように予め蓄積されている画像データを入力する装置であってもよい。

【００４１】

入力された画像は、画像表示部１５１に表示される。利用者は、この画像表示装置を通して入力された画像や、作業の状況を知ることができる。

【００４２】

次に、領域指定部１５２により画像表示部１５１上に表示されている画像の一部領域１６２（図１６）を指定する。指定の方法は、例えば図１６に示すようにマウスカーソル１６１を使用して、画像上の一部分を囲むようにする。さらに詳しくは、指定したい領域の左上の角にマウスカーソルを移動してマウスボタンを押し下げ、そのまま指定したい領域の右下までマウスカーソルを移動しマウスボタンを離すことで領域１６２を指定することができる。

【００４３】

領域指定部１５２により指定された領域内の画像は検索用画像パターンとして切り出され、記憶装置１５７に一時的に記憶される。その検索用画像パターンの内容は例えば参照を示す記号（数字（１），（２）や記号＄，§…など）である。関連領域検索部１５３では、領域指定部１５２で指定された検索用画像パターン（記号画像）に対応する項目を文書中の別の部分から検出し、記号が示す参照の参照先が記述してある部分を抽出する。一般に、このような参照を示す記号に対応する項目は、この記号が記述された部分より文書の後ろ側に出現することが多い。したがって、関連領域検索部１５３は、入力画像の後ろ側から切り出した記号画像とのマッチングを行い、一致した部分を検出する。この場合、利用者により指定される領域内に適当な記号が存在する保証はないので、指定領域内の記号を文字認識部１５４により文字認識を行い、記号のコード以外である場合には警告をして

利用者に適当な領域を指定するように促すこともできる。また、記号に対応する項目は、図１７のように記号を文頭に置き、対応する項目が記号に続いて書かれるため、画像の左端を優先的に検索して選択した記号にマッチングする項目を捜して、検出の時間を短縮することもできる。

【００４４】

具体的に記号の検出は、記憶装置１５７に記憶された検索用画像パターンを検索対象となっている文書画像に重ね合わせ、画素値が一致する画素数をカウントする。検索用画像パターンの全画素数に対する画素値が一致した画素数の割合を画像の一致度として、検索対象画像全体に対して記号画像を順次重ね合わせて、一致度の最も高い部分を検出することで検索用画像パターンの表す記号と同一の記号を表す画像部分を検出する。また、検索対象画像の黒画素で構成される連結領域を抽出して、この連結領域毎に検索用画像パターンとの一致度を調べることにより検出することもできる。

【００４５】

しかし、領域指定部１５２で指定された領域内の記号が肩文字などの本文の文字サイズ（参考文献を記述している文字サイズ）と異なるサイズで記述されている場合には、上記の関連領域検索部１５３ではうまく記号を検出できない場合がある。そのため、記憶装置１５７に記憶されている画像の大きさを本文文字のサイズと一致するように変形させることで、この問題を解決する。具体的には、まず、検索対象の画像全体の画素の横方向への周辺分布を作成する。この周辺分布において、画素の存在しない部分に挟まれる領域が文字行に相当するので、この領域の高さに合わせるように、記憶装置１５７に記憶されている画像の大きさを変形することで、関連領域検索部１５３の精度を向上させることができる。

【００４６】

次に、関連領域検索部１５３は先に検出した項目の周辺を解析して、記号に対応する文字列を切り出す。この文字列の切り出し方法は、検出された項目の周辺において、画像の横方向への周辺分布１７２を作成し、画素の存在しない部分に挟まれる領域が文字行に相当するので、記号検出位置に対応する文字行の部分１８１を切り出す。なお、関連領域検索部１５３で、記憶装置１５７に記憶されている画像を周辺分布を用いて変形している場合には、関連領域検索部１５３で用いた周辺分布をそのまま使用することができる。

【００４７】

この切り出された文字行の部分１８１（図１８）には、例えば論文であれば参考文献の著者名、タイトル等が書かれている。この切り出した文字行の部分を文字認識部１５４により認識し、著者名、タイトル、所属等を表すコード情報を得る。ただし、文献の記載が１行で終わる保証はなく、記号が検出された部分の周辺を解析して、文頭のインデントなどの情報によって、参考文献の記載が複数行にわたっているか調べる必要がある。

【００４８】

例えば、記号に対応する項目が記述されている行を切り出した後、この行の先頭位置（Ｘ座標）Ｒ１を測定して記憶装置１５７に記憶する。次に今検出した行の次の行を周辺分布の情報から切り出し、その行の先頭位置を測定してＲ２とする。今、切り出した２つの行が同じ位置から始まっていれば、次の行は別の項目を示していると判断できるが、下側の行の方が右側（Ｒ１＜Ｒ２）から始まっているなら、下側の行は前の行の続きであると判断できる。したがって、Ｒ１＜Ｒ２である場合、記号に対応する項目が記述されている行とその次の行を切り出す。また、２行以上であっても同様に行の連結性を検査することができるので、３行以上に渡って行が連結している場合でも、同様に切り出すことができる（図１８参照）。

【００４９】

関連領域検索部１５３で切り出した領域は、文字認識部１５４で認識され文字コードに変換される。文字認識の方法は、一般的な方法でよく、特別な手法は必要でない。具体的には、予め文字種類全てに対して、画像的な特徴量を計算しておき辞書を作成しておく。入力となる文字についても同じ特徴量を計算して、求めた特徴量と辞書に記憶された特徴量

とを比較して、最も近い文字を出力する。

【００５０】

このようにして求めた文字コードには、数％程度の認識誤りが含まれている。この誤りをそのままにしておくと、情報が検索できない場合があるので、認識誤りを補う必要がある。最近の文字認識手法では、認識結果の正しさの程度を数値で表せるので、この情報を用いて正しさが低い文字については、その文字をワイルドカード（どの文字を当てはめてもよい特別な記号）を表す記号と置き換えて、情報蓄積検索部１５５に渡して、検索を行うことで、文字認識の認識誤りを補うことができる。例えば、文字認識の結果『あした』という文字列を得て、このうち『し』の文字の認識の正しさが低いとする。ここで、ワイルドカードを表す記号が『＊』であるとすると、この文字列は『あ＊た』という文字列に置き換えられることになる。この置き換えにより、この文字列と一致する文字列は『あした』のほかに『あなた』、『あきた』などがある。この方法は、数文字程度の単語では一致する単語が複数出現するので検索には有効な方法ではないが、論文のタイトルなど、ある程度長い文字列に対しては、一致する文字列が少なくなるので、簡単で有効な手段となる。

【００５１】

文字認識部１５４によって得られた文字列は、著者名、タイトルなどの属性の異なる情報が盛り込まれているので、これらを分離してから情報蓄積検索部１５５に入力する必要がある。そのため、それぞれの属性を分離している『　：　』や『　、』『　；　』などの文字で分割して、それぞれの文字をＡＮＤ条件で情報蓄積検索部１５５の入力とする。例えば、文字認識部１５４の結果、図１９に示すように、『山田太郎：ＸＸに基づく検索装置：Ｋ学会』の文字列１９１が得られたとすると、『　：　』や『　、』のセパレータで区切って、『”山田太郎”ＡＮＤ”ＸＸに基づく検索装置”ＡＮＤ”Ｋ学会”』という検索式１９５に変換する。

【００５２】

また、書誌情報の並びは各文書によって一意に決定できる。例えば、ある文書では『タイトル：著者名：所属』の並びで記載されてることもあるし、別の文献では『著者名：タイトル：出版社』の並びで記載されている。この情報は、一文書内では変わることがないので、事前にこの情報を知ることができれば、より正確な検索を情報蓄積検索部１５５において実施することができる。例えば、図１９で例示している文字列１９１の場合、その書誌情報は『著者名：タイトル：出版社』の順番で記載されているので、著者名として山田太郎、文献のタイトルとして”ＸＸに基づく検索装置”、出版社としてＫ学会を指定して、その検索式は『（著者名：”山田太郎”）　ＡＮＤ　（タイトル：”ＸＸに基づく検索装置”）ＡＮＤ（出版社：”Ｋ学会”）』とすることができる。このような書誌情報は、画像を入力した時点で、利用者が指定することもできるし、情報蓄積検索部１５５に既に蓄積されている画像であれば、その画像情報（画像の幅や高さや生成された日付などの情報）に追加しておくことも可能である。

【００５３】

この検索式を用いて情報蓄積検索部１５５に蓄積されている情報を検索し、必要な情報を出力する。ここで、情報蓄積検索部１５５はシステム内に存在する必要はなく、ネットワークで接続されて遠隔地に存在する利用可能なデータベースを利用してもよい。

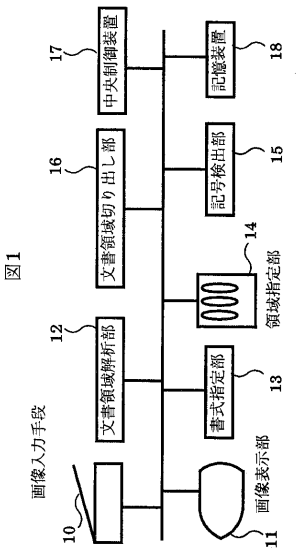
【００５４】

以上の操作により、入力した文書画像中に存在する関連情報項目を取り出し、それにより関連情報を検索して表示することができる。また、検索した結果がテキストデータであっても、画像データであっても同じ操作手順により、さらに情報の検索を行うことができる。すなわち、検索した結果が画像情報であれば、上記の方法でさらに検索を進めることができる。検索した結果がテキストデータであれば、参考文献番号をマウスで指定して、テキスト内から一致する文字を検索する。これは、一般的な文書作成装置などで用いられている文字探索方法を使用することができる。また、テキストデータであれば検索した行の文字列を取り出し、文字認識をする必要もなく検索式を作ることができる。

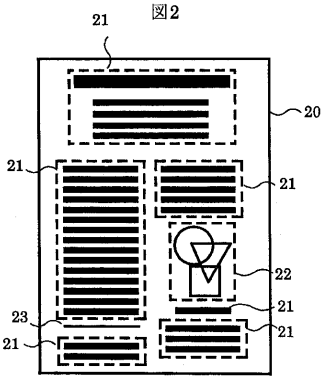
【００５５】
また、文書を表すようなデータは、ほとんどが複数ページにわたって記述されているのが普通である。本実施形態では、簡単に説明するためにあたかも１画像について操作しているように記述しているが、実際には複数の画像を１つの文書として扱い、複数の画像を１つの情報の単位として扱う仕組みが必要である。このような仕組みに関しては、１つの画像ファイルの先頭に幾つの画像が格納されているかの情報を記述して、１つのファイルに複数の画像を格納する方法がある。したがって、本実施例の関連領域検索部１５３は、最終頁を表す画像からマッチングを始めることになる。
【００５６】
また、第１の実施の形態において図１２により説明したのと同様に、１つの文書がどの画像から構成されているかを示す文書画像インデックス情報１２１を作成し、この文書画像インデックス情報１２１によって画像の表示、検索を行うこともできる。なお、この方法によれば、参考文献集の記載されている頁番号を文書画像インデックス情報１２１に登録することで、関連領域検索部１５３は指定されている画像についてのみ処理を行えばいいので、処理時間を短縮することができる。
【００５７】
以上説明したように、本発明の上記実施の形態によれば、画像中の一部を指定して、指定した領域に記述されている内容に関連する項目を検出し、検出された項目に対してのみ文字認識を行い、文字認識した結果から検索式を生成して情報を検索するようにしたので、キーワードの入力などの操作を省略でき、操作性のよい情報検索装置を実現できる。
【００５８】
また、本実施の形態によれば、検索した結果が画像情報であろうとテキストデータであろうと、その操作方法は同じなので、検索結果の情報の形態を意識することのない、操作性のよい情報検索装置を提供できる。
【００５９】
さらに、本実施の形態によれば、複数の画像を１つの情報単位として扱うことができるようにしたので、実際の紙の文書を読むのに近い感覚で扱うことのできる操作性のよい情報検索装置を提供できる。
【００６４】
【発明の効果】
本発明（請求項１）によれば、検索用情報の入力のための操作を省略することができ、操作性のよい情報検索装置を実現できる。また、本発明によれば、検索した結果が画像情報であろうとテキストデータであろうと、その操作方法は同じなので、検索結果の情報の形態を意識することのない、操作性のよい情報検索装置を提供することができる。
【図面の簡単な説明】
【図１】本発明の実施の形態１の基本構成図
【図２】文書を構成する領域を説明するための図
【図３】画像のラベル付けのための走査方法を説明するための図
【図４】（ａ）及び（ｂ）は画像のラベル付けを説明するための図
【図５】画像のラベル付けを説明するための図
【図６】（ａ）及び（ｂ）はラベルの外接矩形によって表した画像を説明するための図
【図７】文書領域解析部によって抽出される領域を説明するための図
【図８】節末／章末における注釈領域の指定方法を説明するための図
【図９】領域指定部による領域指定を説明するための図
【図１０】図番号を指定した場合の参照位置画像の表示を説明するための図
【図１１】注釈記号を指定した場合の参照位置画像の表示を説明するための図
【図１２】複数の画像によって構成される文書の管理方法を説明するための図
【図１３】複数画像の文書において図番号を指定した場合の参照位置画像の表示を説明するための図
【図１４】複数画像の文書において注釈記号を指定した場合の参照位置画像の表示を説明

するための図
【図１５】本発明の実施の形態２の基本構成図
【図１６】領域指定部による領域指定を説明するための図
【図１７】指定された記号に対応する項目及び記号検出位置周辺における周辺分布を示す図
【図１８】複数行で記載されている文献項目を切り出す方法を説明するための図
【図１９】認識結果を検索式に変換する方法を説明するための図
【符号の説明】
１０…画像入力部、１１…画像表示部、１３…書式指定部、１４…領域指定部、１５…記号検出部、１６…文書領域切り出し部、１７…中央制御装置、１８…記憶装置、２０…文書画像、２１…文字領域、２２…図形領域、２３…けい線領域、４１…画素連結テーブル、６１…外接矩形、７１…脚注領域、８１…マウスカーソル、８２…注釈開始位置指定マーク、８３…注釈終了位置指定マーク、９１…指定領域、１０１…記号検出位置マーク、１２１…文書画像インデックス情報、１５０…画像入力部、１５１…画像表示部、１５２…領域指定部、１５３…関連領域検索部、１５４…文字認識部、１５５…情報蓄積検索部、１５６…中央制御装置、１５７…記憶装置、１６０…画像、１６１…マウスカーソル、１６２…指定領域、１７１…記号検出位置、１８１…１つの文献を記載した領域、１９１…文字認識部１４による認識結果、１９２…著者名に対応する文字列、１９３…文献名に対応する文字列、１９４…出版社に対応する文字列、１９５…１９２、１９３、１９４の文字列を組合せた検索式。

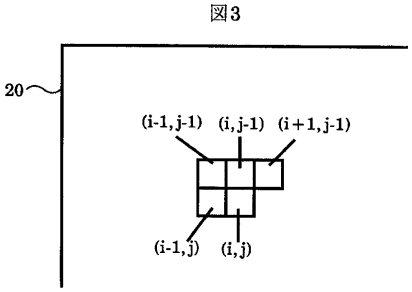
【図１】



【図２】



【図３】



【図４】

図4

(a)

20

1	1	1	1	1	1	1	1
1	0	0	0	0	0	0	1
1	0	2	0	3	3	0	1
1	0	2	0	3	3	0	1
1	0	2	2	2	2	0	1
1	0	0	2	2	2	0	1
1	0	0	0	0	0	0	1
1	1	1	1	1	1	1	1

(b)

41

画素ラベル値	連結画素ラベル値
1	1
2	2
3	2

【図５】

図5

20

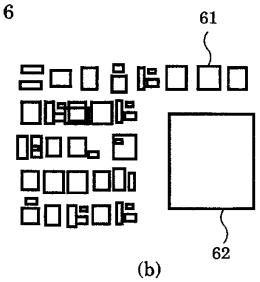
1	1	1	1	1	1	1	1
1	0	0	0	0	0	0	1
1	0	2	0	2	2	0	1
1	0	2	0	2	2	0	1
1	0	2	2	2	2	0	1
1	0	0	2	2	2	0	1
1	0	0	0	0	0	0	1
1	1	1	1	1	1	1	1

【図６】

図6

このように生成された画像に対して、次の処理を加えるために

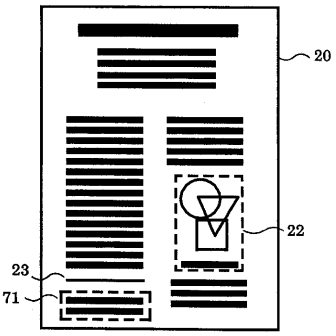
(a)



(b)

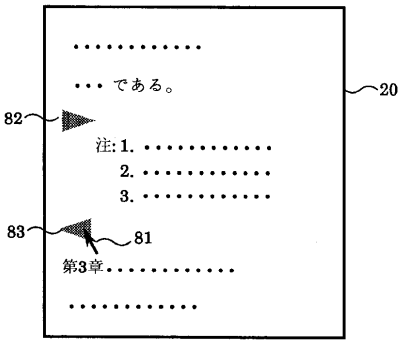
【図７】

図7



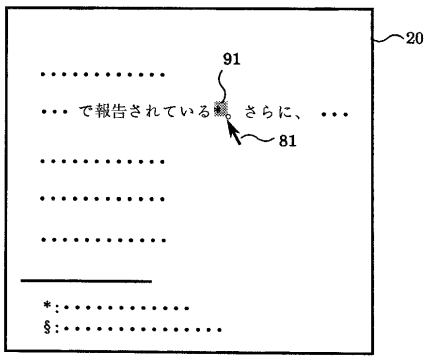
【図８】

図8



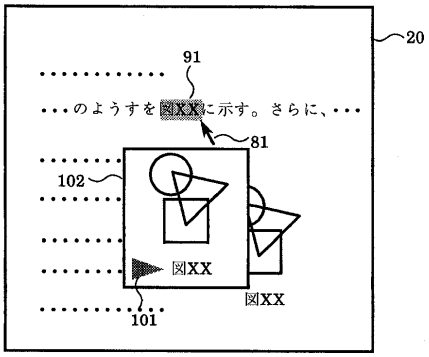
【図９】

図9



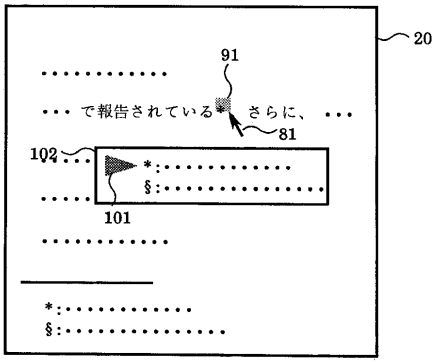
【図１０】

図10



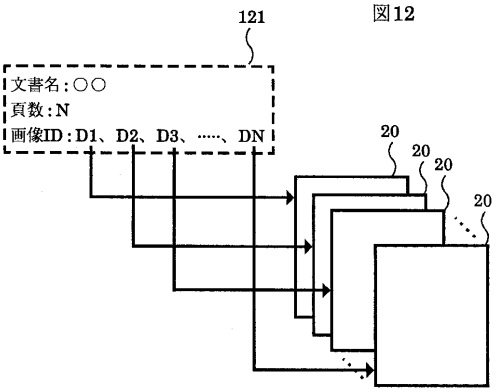
【図１１】

図11



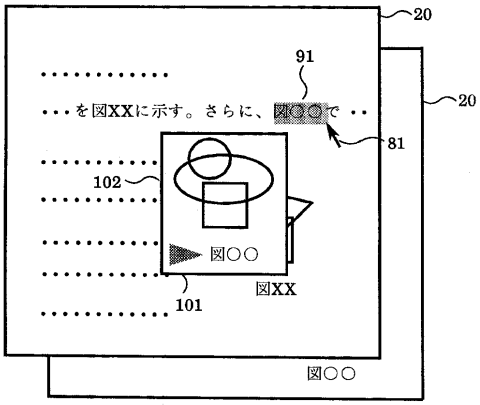
【図１２】

図12



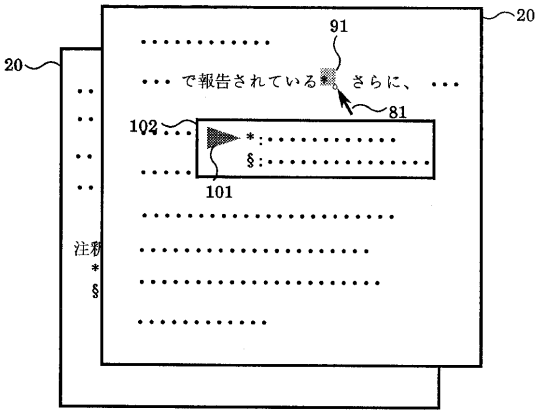
【図１３】

図13



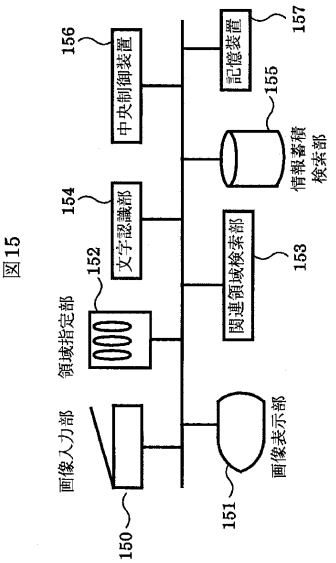
【図１４】

図14



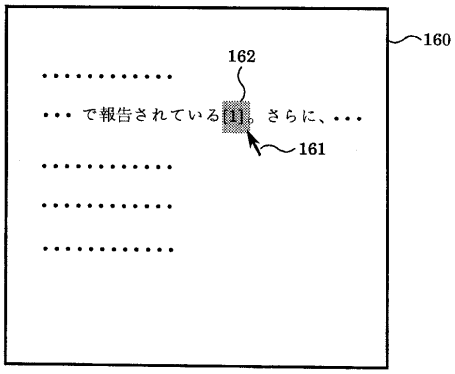
【図１５】

図15



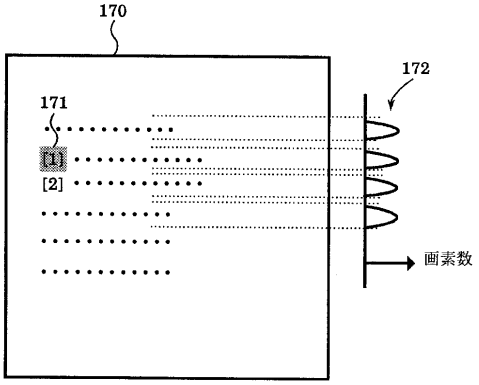
【図１６】

図16



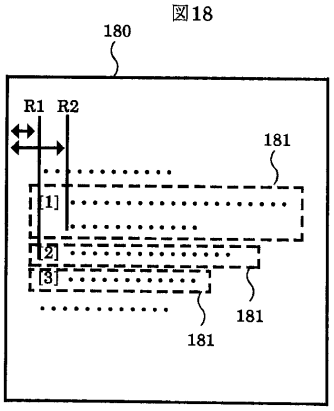
【図１７】

図17



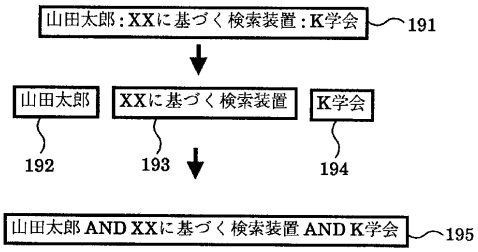
【図１８】

図18



【図１９】

図19



フロントページの続き

(56)参考文献 特開昭６３－０６２０３３（ＪＰ，Ａ）
特開平０４－１８８３６５（ＪＰ，Ａ）
特開平０１－１５９７５９（ＪＰ，Ａ）
坂本弘章、高畑実、上森明，複数のキー画像の部分的な特徴を使用する静止画像検索方式の提案　フレキシブル・モンタージュ検索，電子情報通信学会技術研究報告，日本，１９９３年１１月１２日，９３巻、３１７号，３１－３６頁、ＡＩ９３－６３

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G06F 17/30 - G06F 17/30 419
JICSTファイル(JOIS)