

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-257543

(P2008-257543A)

(43) 公開日 平成20年10月23日(2008.10.23)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 200D	5B009
G06F 17/21 (2006.01)	G06F 17/21 530A	5B050
G06F 17/30 (2006.01)	G06F 17/30 170B	5B075
	G06F 17/30 350C	5B109
	G06T 1/00 200C	

審査請求 未請求 請求項の数 3 O L (全 15 頁)

(21) 出願番号 特願2007-100154 (P2007-100154)
 (22) 出願日 平成19年4月6日(2007.4.6)

(71) 出願人 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 加藤 雅弘
 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
 Fターム(参考) 5B009 NA01 QA06 VA09
 5B050 BA10 BA16 BA20 GA08
 5B075 ND07 NK21 PR06 QM08
 5B109 NA01 QA06 VA09

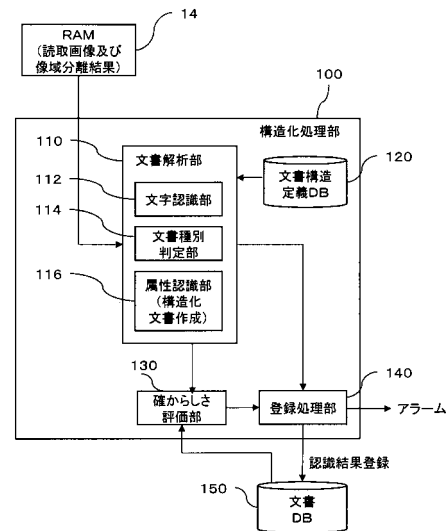
(54) 【発明の名称】 画像処理システム及びプログラム

(57) 【要約】

【課題】 入力された画像から認識した属性の値が、当該属性の値として妥当か否かを判定できるシステムを提供する。

【解決手段】 文書解析部110は、画像読取装置が出力した読取画像を解析し、その画像に含まれる各属性の値を認識する。確からしさ評価部130は、文書解析部110が認識した属性の値を、文書DB150に登録された当該読取画像の文書種類と同一種類の過去の文書内での同じ属性の値と比較することで、認識した属性の値の確からしさの評価値を求める。その評価値が閾値より低い場合、登録処理部140は、認識した属性の値の確認をユーザに求める。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

過去に入力された画像情報ごとに、その画像情報の文書種別と、その画像情報から認識された各属性の値と、を記憶する記憶手段と、

入力された第 1 の画像情報に対応する文書種別を判別する判別手段と、

前記第 1 の画像情報から、前記判別手段が判別した文書種別に対応した各属性の値を認識する認識手段と、

前記認識手段が前記第 1 の画像情報から認識した各属性の値を、それぞれ、前記記憶手段に記憶された前記第 1 の画像情報と同じ文書種別の過去の画像情報群から認識された同じ属性の値と比較することにより、前記認識手段が認識した各属性の値がそれぞれ当該属性の値であることの確からしさを示す属性評価値を計算する属性評価値計算手段と、

前記評価値計算手段が計算した各属性の属性評価値がすべて所定の閾値以上である場合は、前記認識手段が認識した各属性の値を前記判別手段が判別した文書種別に対応づけて前記記憶手段に記憶させ、そうでない場合は例外処理を実行する処理手段と、

を備える画像処理システム。

【請求項 2】

前記入力された第 1 の画像情報の画質を示す画質評価値を計算する画質評価値計算手段を更に備え、

前記属性評価値計算手段は、前記画質評価値計算手段が評価した画質評価値を反映した属性評価値を計算する、

ことを特徴とする請求項 1 記載の画像処理システム。

【請求項 3】

コンピュータを、

過去に入力された画像情報ごとに、その画像情報の文書種別と、その画像情報から認識された各属性の値と、を記憶する記憶手段、

入力された第 1 の画像情報に対応する文書種別を判別する判別手段、

前記第 1 の画像情報から、前記判別手段が判別した文書種別に対応した各属性の値を認識する認識手段、

前記認識手段が前記第 1 の画像情報から認識した各属性の値を、それぞれ、前記記憶手段に記憶された前記第 1 の画像情報と同じ文書種別の過去の画像情報群から認識された同じ属性の値と比較することにより、前記認識手段が認識した各属性の値がそれぞれ当該属性の値であることの確からしさを示す属性評価値を計算する属性評価値計算手段、

前記評価値計算手段が計算した各属性の属性評価値がすべて所定の閾値以上である場合は、前記認識手段が認識した各属性の値を前記判別手段が判別した文書種別に対応づけて前記記憶手段に記憶させ、そうでない場合は例外処理を実行する処理手段、

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、画像処理システム及びプログラムに関する。

【背景技術】**【0002】**

請求書等の紙文書を光学的に読み取ってそこに記載されている各種の情報を認識し、認識した情報をデータベースに登録するシステムが利用されている。近年では、紙文書に示された各項目の情報を認識し、その認識結果を X M L (eXtensible Markup Language) などの構造化記述言語で記述して保存するシステムが普及しつつある。

【0003】

特許文献 1 の装置は、光学的に読み取った文書画像から文字領域及び画像領域をそれぞれ区別して抽出し、各領域のレイアウトを求め、文字領域には文字認識を行い、レイアウトと文字認識結果を用いて構造化文書を作成する。この装置では、この一連の処理を自動

10

20

30

40

50

処理で行っている。

【0004】

特許文献2の装置は、読み取った文書の画像を文字領域、表領域、イメージ領域に分離し、各領域の矩形を表示する。その矩形をユーザが移動させることによって、レイアウト編集が行われ、編集されたレイアウトにしたがったハイパーテキストが作成される。

【0005】

【特許文献1】特開平11-066196号公報

【特許文献2】特開平11-312231号公報

【発明の開示】

【発明が解決しようとする課題】

10

【0006】

本発明は、入力された画像から認識した属性の値が、当該属性の値として妥当か否かを判定できるシステ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】

請求項1に係る発明は、過去に入力された画像情報ごとに、その画像情報の文書種別と、その画像情報から認識された各属性の値と、を記憶する記憶手段と、入力された第1の画像情報に対応する文書種別を判別する判別手段と、前記第1の画像情報から、前記判別手段が判別した文書種別に対応した各属性の値を認識する認識手段と、前記認識手段が前記第1の画像情報から認識した各属性の値を、それぞれ、前記記憶手段に記憶された前記第1の画像情報と同じ文書種別の過去の画像情報群から認識された同じ属性の値と比較することにより、前記認識手段が認識した各属性の値がそれぞれ当該属性の値であることの確からしさを示す属性評価値を計算する属性評価値計算手段と、前記評価値計算手段が計算した各属性の属性評価値がすべて所定の閾値以上である場合は、前記認識手段が認識した各属性の値を前記判別手段が判別した文書種別に対応づけて前記記憶手段に記憶させ、そうでない場合は例外処理を実行する処理手段と、を備える画像処理システムである。

20

【0008】

請求項2に係る発明は、請求項1に係る発明において、前記入力された第1の画像情報の画質を示す画質評価値を計算する画質評価値計算手段を更に備え、前記属性評価値計算手段は、前記画質評価値計算手段が評価した画質評価値を反映した属性評価値を計算する

30

【0009】

請求項3に係る発明は、コンピュータを、過去に入力された画像情報ごとに、その画像情報の文書種別と、その画像情報から認識された各属性の値と、を記憶する記憶手段、入力された第1の画像情報に対応する文書種別を判別する判別手段、前記第1の画像情報から、前記判別手段が判別した文書種別に対応した各属性の値を認識する認識手段、前記認識手段が前記第1の画像情報から認識した各属性の値を、それぞれ、前記記憶手段に記憶された前記第1の画像情報と同じ文書種別の過去の画像情報群から認識された同じ属性の値と比較することにより、前記認識手段が認識した各属性の値がそれぞれ当該属性の値であることの確からしさを示す属性評価値を計算する属性評価値計算手段、前記評価値計算手段が計算した各属性の属性評価値がすべて所定の閾値以上である場合は、前記認識手段が認識した各属性の値を前記判別手段が判別した文書種別に対応づけて前記記憶手段に記憶させ、そうでない場合は例外処理を実行する処理手段、として機能させるためのプログラムである。

40

【発明の効果】

【0010】

請求項1に係る発明によれば、入力された画像情報から認識した属性の値が当該属性の値として妥当か否かを判定できる。

【0011】

請求項2に係る発明によれば、入力された画像情報から認識した属性の値の妥当性を、

50

その画像情報の画質を考慮して判定することができる。

【0012】

請求項3に係る発明によれば、入力された画像情報から認識した属性の値が当該属性の値として妥当か否かを判定できる。

【発明を実施するための最良の形態】

【0013】

以下、図面を参照して本発明の好適な実施の形態を説明する。図面において、同様の構成要素又はステップには同一符号を付して、重複した説明は省略する。

【0014】

本実施形態では、一例として、見積書等の定型文書の画像を解析してその定型文書に含まれる各文書要素の値を求め、それら各文書要素の値を含んだ構造化文書の文書データ（以下単に「構造化文書」と呼ぶ）を生成するためのシステムを説明する。文書要素は、「属性」ともいう。また、以下では、一例として、スキャナ、プリンタ、複写機などの機能をあわせ持つ画像形成装置（このような画像形成装置は、デジタル複合機とも呼ばれる）をベースにしたシステムを例にとって説明する。すなわち、以下では、紙等の媒体に印刷された文書を画像形成装置にて読み取り、読取結果の画像を解析して構造化文書を作成する処理を例にとって説明する。このシステムは、例えば、紙の見積書を読み取り、その読み取り画像から、発行者や宛先、個々の商品の単価や個数、合計金額などの要素を認識し、それら要素からなる構造化文書を生成するのに用いられる。

10

【0015】

本実施形態の画像形成装置の制御機構の主要部を図1に示す。図1においてROM（リード・オンリ・メモリ）12には、画像形成装置の動作制御のための制御プログラムなどのデジタル情報が格納されている。CPU（中央処理装置）10がこのROM12内の制御プログラムを実行することにより、いかに説明する画像形成装置の各部の制御が実現される。

20

【0016】

RAM（ランダム・アクセス・メモリ）14は、画像形成装置の主記憶装置であり、制御プログラムの実行の際にワークメモリとしても用いられる。RAM14は、例えば、画像出力装置24に供給する1ページ分の画像データを蓄えるページバッファとして用いることもできる。

30

【0017】

大容量記憶装置16は、各種のデータを保存するための補助記憶装置であり、例えばハードディスクやEEPROM（Electrically Erasable Programmable Read-Only Memory）などの不揮発性の記憶装置である。

【0018】

操作パネル18は、この画像形成装置のユーザインタフェースのための表示や、ユーザからの各種指示の入力受付などのためのユーザインタフェース手段である。操作パネル18は、例えば、スタートボタンなどの機械的な操作ボタンと、GUI（グラフィカル・ユーザ・インタフェース）のための液晶タッチパネルと、を備える。液晶タッチパネルは、CPU10で実行される制御プログラムが生成したGUI画面を表示し、そのディスプレイに対するユーザのタッチ位置を検出して制御プログラムに渡す。制御プログラムは、そのタッチ位置の情報からユーザの入力内容を解釈する。

40

【0019】

通信インタフェース20は、LAN（ローカル・エリア・ネットワーク）等のネットワークを介して他の装置とデータ通信を行う際の制御を担う装置である。リモートホストからのプリント指示等は、この通信インタフェース20を介して画像形成装置内に入力される。また、ネットワーク上にある言語翻訳サーバなどといった外部サーバにアクセスする際にも、通信インタフェース20は用いられる。

【0020】

画像読取装置22は、原稿を光学的に読み取って電子的な画像データを生成する装置で

50

ある。例えば、ラインセンサやエリアセンサを用いたスキャナがその一例である。自動原稿送り装置（ADF）（図示省略）にセットされた原稿は、ADFの機能により1枚ずつ画像読取装置に送られ、光学的に読み取られる。

【0021】

画像出力装置24は、CPU10の制御により供給される画像データを用紙に画像形成（印刷）する装置であり、プリントエンジンとも呼ばれる。

【0022】

ファクシミリモジュール26は、ファクシミリデータの送受信を行うモジュールである。

【0023】

像域分離回路28は、画像読取装置22により得られた読取画像の像域分離を行う回路である。よく知られるように、像域分離では、画像のエッジ強度や空間周波数などの各種の画像特徴を用いて文字や連続調画像などの像域（以下紛れのない場合に、像域を「領域」と呼ぶこともある）を判別する。像域分離回路28は、読取画像を同じ画像特徴を有する連続した領域ごとに区切る。その区切りの結果得られた各連続領域がそれぞれ像域と呼ばれる。複写機等の画像形成装置は、印刷画質向上のために、読み取った画像中の文字領域と連続調画像領域とを分離する像域分離の機能を備えるものが一般的になっている。また、高速な印刷処理のため、この像域分離の機能の多くの部分はASIC(Application Specific Integrated Circuit)やDSP(Digital Signal Processor)等のハードウェア回路として実装されている。像域分離回路28としては、そのような従来から画像形成装置が備える回路を用いることができる。従来は、像域分離回路は、もっぱら印刷などといった画像出力のためだけに用いられていたが、この像域分離回路28は、画像出力だけでなく、文書画像の解析処理にも利用される。

【0024】

画像処理回路30は、文書画像に対し、印刷やファクシミリ送信等の用途に応じた画像処理を施す回路である。例えば、印刷を行う場合、像域分離回路28で分離された文字と画像の像域に対し、それぞれ文字用の画像処理（エッジ強調など）と連続調画像用の画像処理（階調補正など）を施す。

【0025】

次に、この画像形成装置の構造化文書作成処理に関する機能部分の構成及びその処理内容を、図2及び図3を参照して説明する。この画像形成装置は、構造化文書作成処理のための機能として、図2に示す構造化処理部100を備える。構造化処理部100は、以下に説明する機能又は手順を記述したプログラムをCPU10に実行させることにより実現できる。このプログラムは、ROM12又は大容量記憶装置16に記憶されており、CPU10により実行される。このプログラム又はその一部は、ディスク等の可搬型媒体又はネットワーク経由のダウンロードにより大容量記憶装置16にインストールしてもよい。

【0026】

構造化処理部100は、画像読取装置22が読み取った原稿の画像を解析して、その原稿の画像から各属性（文書要素）の値を認識し、その認識結果に従ってその原稿に表された情報の内容を示した構造化文書を生成する。画像形成装置は、操作メニューの一項目として構造化文書作成機能を備える。ユーザが操作パネル18に表示される操作メニューからこの機能を選び、原稿をプラテン又はADFにセットして処理の実行を指示すると、画像読取装置22がその原稿を読み取り（S11）、構造化処理部100がその読取結果の画像に対する処理を実行する。画像読取装置22が読み取った画像はRAM14に蓄積される。また像域分離回路28は、その画像に対して像域分離処理を施し、その処理結果のデータをRAM14に蓄積する。構造化処理部100は、RAM14に蓄積された読取結果の画像（以下「読取画像」と呼ぶ）と像域分離結果のデータとに基づき、構造化文書を作成する。

【0027】

構造化処理部100は、機能モジュールとして文書解析部110、文書構造定義DB（

10

20

30

40

50

データベース) 120, 確からしさ評価部 130, 及び登録処理部 140を備える。文書解析部 110は、読取画像及び像域分離結果を解析する手段であり、文字認識部 112、文書種別判定部 114及び属性認識部 116を備える。

【0028】

文字認識部 112は、読取画像中の文字領域の画像に対して文字認識を行う。文字認識部 112は、周知のOCR (Optical Character Recognition: 光学文字認識) 技術を用いたものでよい。文字認識部 112は、漢字や仮名などの日本語の文字、アルファベット、数字、記号などの各種の文字のうちの1以上を認識する。

【0029】

文書種別判定部 114は、読取画像の文書種別を判定する。「文書種別」は、例えば見積書、請求書、領収書など、使用目的から見た文書の種類のことである。文書種別の判定は、文書構造定義DB 120に蓄積された各文書種別の文書構造定義に基づき行う。また、「AAAAA社」宛の発注書などのように、使用目的による分類と当該文書の宛先による分類との組合せにより、文書種別を規定してもよい。

10

【0030】

ここで、文書構造定義について説明する。文書構造定義は、文書の論理構造を定義するデータである。例えば構造化文書をXMLで記述する場合には、文書構造定義は、例えばDTD (Document Type Definition) やW3C XML schema、RELAX NG (Regular Language description for XML New Generation) などの、XML用のスキーマ言語により記述される。文書構造定義は、文書種別ごとに作成される。文書構造定義は、文書を構成する複数の要素の論理構造を表す。また、文書構造定義には、それら各要素の要素名、データ型等の情報を含めることもできる。データ型には、文字列型、整数型、日付型などといった様々な型がある。また、文書構造定義には、各要素のとり得る値の範囲を示す情報 (例えば下限値と上限値の組) が含まれていてもよい。またそれら単純な型を組み合わせた複合型なども定義できる。また文書構造定義には、文書中での要素の出現回数や要素の配列順序などの情報を含めることもできる。

20

【0031】

文書構造定義DB 120には、このような文書構造定義が登録される。また、文書構造定義DB 120には、図4に示すように、文書ごとに、スタイル情報204又は検索キー情報206又はその両方を、上述のような文書構造定義202と対応づけて登録してもよい。

30

【0032】

スタイル情報204は、文書を構成する各要素の表示 (可視化) の仕方を規定する情報である。ここで、例えば像域分離により求められ個々の像域が、それぞれ文書の要素に対応する。スタイル情報204は、例えば、それら各要素に対応する像域の読取画像内での位置、形状の情報を含む。また、スタイル情報204には、それら各要素の種類の情報が含まれていてもよい。要素の種類は、例えば、文字列、連続調画像、グラフィック図形などがある。文字列を、更に詳細に例えば数字列と日本語文字列とに分類してもよい。これら要素の種類は、像域分離及び文字認識の処理において特定することができる。また、文字領域に対応する要素の場合、その領域に含まれる文字のサイズ、フォント、字詰めなどは、その要素に対するスタイル情報204の項目の一例である。スタイル情報204は、例えばCSS (Cascading Style Sheets) やXSL (eXtensible Stylesheet Language) などのスタイルシート言語で記述されたものでよい。

40

【0033】

検索キー情報206は、文書構造定義を検索する際の検索条件の情報である。例えば、会社名の文字列 (例えば「株式会社AAA」)、会社のロゴマークや社印の画像情報、又は文書の表題の文字列 (例えば「見積書」、「領収書」) などが、検索キー情報206の一例である。1つの文書構造定義に対応する検索キー情報206は、会社名、ロゴマーク、表題などといった情報のうちのいずれかを含んでいればよい。また、文書構造定義DB 120に登録された文書構造定義202のすべてについて検索キー情報206が登録さ

50

れている必要は必ずしもない。

【 0 0 3 4 】

文書種別判定部 1 1 4 は、各文書種類の文書構造定義 2 0 2 (及びそれに対応するスタイル情報 2 0 4 , 検索キー情報 2 0 6)の中から、読取画像に最もよく適合するものを選ぶことで、その読取画像の文書種別を特定する。

【 0 0 3 5 】

この判定では、例えば、文書の読取画像中に含まれる文書の見出し(表題)を抽出し、抽出した表題により文書構造定義の探索範囲を絞り込む。例えば、読取画像中の見出しの存在する可能性の高い範囲(例えば文書の上部)にある文字領域に対して文字認識を行い、この結果得られた文字列が文書の見出しとなり得る文字列であれば、その文書の見出しの文字列と判定する。文書の見出しとなり得る文字列のリストは、あらかじめ構造化処理部 1 0 0 に登録しておけばよい。そして、文書構造定義 DB 1 2 0 に登録された文書構造定義 2 0 2 のうち、対応する検索キー情報 2 0 6 がその見出しの文字列(又はその類義語)を含む文書構造定義 2 0 2 を、読取画像の文書構造定義の候補として抽出する。

10

【 0 0 3 6 】

以上では、文書の見出しによる絞り込みの例を説明したが、見積書等の定型文書中にはその文書を発行した発行者の社名、住所、又はロゴマークが含まれる場合も多いので、それらのうちいずれかを用いて候補の絞込を行うこともできる。この場合、社名、住所等も検索キー情報 2 0 6 の項目として文書構造定義 DB 1 2 0 に登録しておけばよい、また、見出し、社名、住所、ロゴマークのうち2以上の組合せにより候補を絞り込んでもよい。なお、このような見出し、社名等の文字列は、文字認識部 1 1 2 により求められる。

20

【 0 0 3 7 】

以上のような絞込を行う代わりに、読取画像から抽出した見出しや社名等の各種検索キーに応じて、文書構造定義 DB 1 2 0 中の各文書構造定義について、読取画像に対する適合の度合いを示すスコアを求めてもよい。例えば、文書見出しについてのスコアは、読取画像から抽出した部署の見出しの文字列と意味的に近い見出し文字列を検索キー情報 2 0 6 に含んだ文書構造定義 2 0 2 ほど高い値となるようにつけばよい。また例えば、見出し、社名などといった項目ごとにこのようなスコアを求め、それらを総合した値(例えばそれらの平均値)を総合的なスコアとし、その総合的なスコアが最も高い文書構造定義を、読取画像に対応する文書構造定義と判定してもよい。

30

【 0 0 3 8 】

また、文書種別判定部 1 1 4 は、そのような検索キー情報 2 0 6 を用いた探索の代わりに、或いはそのような探索に加えて、文書要素群のレイアウトに基づく探索を行ってもよい。レイアウトに基づく探索は、例えば、像域分離回路 2 8 の像域分離結果を利用することができる。例えば、像域分離により求められた文字領域、連続調画像領域等のレイアウトと、文書構造定義 DB 1 2 0 内の各文書構造定義 2 0 2 に対応するスタイル情報 2 0 4 が示すレイアウトとを比較することで、読取画像とそれら各文書構造定義 2 0 2 とのレイアウトに関する類似度合いを求めることができる。類似度合いは、例えば、読取画像のレイアウトにおける各領域を、スタイル情報 2 0 4 が示すレイアウトにおいてその領域と同じ種類(文字、連続調画像、又はグラフィック図形など)の領域のうち最も近い領域と重複なく対応付け、それら対応付けられた領域同士の位置やサイズの差を総合した値(例えばそれら差の自乗和など)から類似度合いのスコアを求めればよい。例えば、対応づけられた領域同士の差が小さいほど類似度合いのスコアは高くなるような関数を用いればよい。

40

【 0 0 3 9 】

また、レイアウトの類似度合いを求める際に、文字領域、連続調画像領域などといった画像の種類だけでなく、各領域のデータ型の適合の度合いを考慮してもよい。例えば、読取領域の各文字領域内の文字列のデータ型は、文字認識部 1 1 2 による文字認識により判明する。文字認識部 1 1 2 は、例えば、単純な文字認識により得られた文字列を、金額特有の表記パターン、日付の文字列パターンや、住所特有の表記パターンなどと照合するこ

50

とにより、その文字列が金額、日付、住所などといったデータ型のいずれに該当するかを判定することができる。一方、文書構造定義202には各要素のデータ型が記述されている。したがって、例えば、読取画像の領域ごとに、その領域のデータ型と、文書構造定義202内の要素のうちその領域に対応する要素のデータ型とを求め、データ型が適合する領域又は要素の数、又はその数の文書構造定義の全要素数に対する割合に基づき、レイアウトの類似度合いのスコアを求めることができる。読取画像の領域に対応する文書構造定義202の要素は、例えば、その文書構造定義202に対応するスタイル情報204が示すレイアウトにおいて、その領域に最も近い要素である。

【0040】

また、レイアウトの類似度合いを求める際に、読取画像の各領域に含まれる見出し文字列を考慮してもよい。例えば発注書の場合、「商品名」や「数量」などといった見出し文字列が含まれる。文書種別判定部114は、読取画像のある領域の文字認識結果の中にそのような見出しに該当する文字列があれば、その文字列と、文書構造定義202中の要素のうちその領域に対応する要素に含まれる見出し文字列との類似度合いを求め、その類似度合いに基づき、レイアウトの類似度合いのスコアを求めることができる。

10

【0041】

以上、レイアウトの類似度合いのスコアの求め方をいくつか例示した。このような手法で求めたスコアが最も高い文書構造定義202を、読取画像に対する適合度が最も高い文書構造定義202と判定してもよい。

【0042】

なお、レイアウトの類似度合いを求めるにあたり、以上に例示した領域同士の位置やサイズの類似度、データ型の適合度、及び見出し文字列の類似度のうちの2以上を総合してもよい。

20

【0043】

また、文書の見出しの類似度合いと、レイアウトの類似度合いとを総合したスコアを求め、そのスコアが最も高い文書構造定義202を、読取画像に対する適合度が最も高い文書構造定義202と判定してもよい。また、文書の見出しでもレイアウトでもない他の観点の類似度合いを用いてもよい。

【0044】

なお、以上の処理において、適合度に対して閾値を設定しておき、文書構造定義202群のうちの最高の適合がその閾値以上の場合にのみ、その最高値に対応する文書構造定義を最適な文書構造定義として採用するようにしてもよい。

30

【0045】

文書種別判定部114は、以上のようにして求めた読取画像に対する適合度が最も高い文書構造定義に対応する文書種別を、読取画像の文書種別と判定する(S12)。

【0046】

このようにして読取画像に対する適合度が最も高い文書構造定義が求められると、次に属性認識部116が、読取画像から抽出した各領域の値をその文書構造定義に規定される各要素(属性)の値に当てはめる(S13)。適合度が最も高い文書構造定義を求める過程で、読取画像の各領域がその文書構造定義のどの要素に対応するかは判明しているので、ここでは各領域の値(例えば文字認識の結果)をそれらに対応する各要素に対応づければよい。この対応付け処理により、読取画像の情報内容を表す構造化文書が作成される。

40

【0047】

登録処理部140は、文書解析部110が生成した構造化文書を文書DB150に登録する処理を行う。ここで、構造化文書と対応づけて、その元になった読取画像のデータを文書DB150に登録してもよい。文書DB150は、構造化文書を蓄積するデータベースである。文書DB150は、画像形成装置内に設けられていてもよいし、画像形成装置からアクセス可能なネットワーク上に設けられていてもよい。

【0048】

ここで、以上のように紙文書を読み取った画像を解析して構造化文書を作成する場合、

50

解析が正確に行われないと、誤った属性値を含んだ構造化文書が作成されることになる。例えば、紙文書に印刷された数量を表す値のいずれかの数字が、その値を取り囲む罫線と接触した結果、文字認識の過程で数字と認識されない可能性がある。また、印刷の汚れや人手による書き込みなどが文字認識において小数点と認識されることにより、文書中の数値が本来と異なる値として認識される場合がある。また、小数点と、カンマ等の位取り記号とを誤って認識することで、文書中の数値が本来と異なる値として認識される場合がある。このような数値の認識誤りは往々にして起こり得る。例えば企業活動で作成された記録の紙文書からそのような構造化文書を作成して蓄積する場合、誤った属性値が含まれていたのでは、データとしての価値が低くなってしまふ。

【 0 0 4 9 】

誤った属性値を含んだ構造化文書を、そうでない構造化文書と同列に文書DB150に蓄積したのでは、文書DB150の信頼性が低下する。そこで、属性値に誤りがあるか否かを判定し、誤りがあると判定される場合には、その旨をユーザに知らせるなどの対策をとることが考えられる。

【 0 0 5 0 】

このために、例えば、文字認識で求められた各属性の値が、その属性のデータ型に適合しているか否か、その属性の値としてとり得る範囲に属しているか否かを自動的に判定することも考えられる。しかし、このようなデータ型や値のとり得る範囲によるチェックでは、認識された属性値がその範囲に収まってしまう場合には、誤りを見落としがちである。

【 0 0 5 1 】

そこで、以下では、データ型及びとり得る値の範囲のチェックのみでは見つけられない属性値の誤りを自動判定し、その判定結果に応じて登録処理部140の処理内容を制御するための構成の一例を説明する。

【 0 0 5 2 】

本実施形態の画像形成装置は、このような制御のために、確からしさ評価部130を備えている。確からしさ評価部130は、その構造化文書の各属性(要素)の値の確からしさの評価値を計算する(S14)。ここで言う「確からしさの評価値」は、その値が、当該属性の値であることの確からしさの程度、を示す評価値である。本実施形態では、この確からしさの評価値は、評価対象の属性値と、過去の同種の文書における当該属性の値との関係に基づき求める。すなわち、評価対象の属性の値が、過去の同種の文書における当該属性の値からかけ離れているほど、評価対象の値が該属性の値である確からしさは低くなる、と判断する。過去の同種の文書における属性の値は、文書DB150に蓄積された構造化文書群から求めることができる。

【 0 0 5 3 】

一例として、例えば今回の読取画像から認識した当該属性の値が、直近の同種の文書の同一属性の値がどの程度変化したかという変化量に基づき、確からしさの評価値を求めることができる。例えば、定期的発行する部品の発注書のように時間的に順に作成される同種の文書を本実施形態の画像形成装置により読み取って文書DB150に登録していく場合を考える。これら発注書の中から、同一の部品「CCCC」の発注数量の属性値を認識し、それら各発注書に対応する構造化文書にそれぞれ組み込んで文書DB150に蓄積しているとする。ここで、過去の各時点での部品「CCCC」の発注数量の推移が図5に示す棒グラフで与えられたとする。この発注数量の推移から、所定回数(図示例では3回)の発注ごとの発注数量の移動平均値を求めたものが図5の実線のグラフである。そして、今回読み取った発注書から認識した発注数量の、直前の移動平均値に対する変化量を求める。そして、その変化量に基づき、確からしさの評価値を計算する。この評価値Vの計算式の例を以下に示す。

【 0 0 5 4 】

$$V = \max(0, 100 - |\text{変化量}| / \text{平均})$$

【 0 0 5 5 】

10

20

30

40

50

この計算式では、0と $(100 - |\text{変化量}| / \text{平均})$ とのうち大きい方が評価値Vの値となる。最小値は0で最大値は100である。分母の「平均」は、変化量の計算の際に用いた「直前の移動平均値」である。ただし、「直前の移動平均値」の代わりに、過去の発注数量の全平均を用いてもよい。この計算式で求められた評価値Vは、平均に比して変化量が大いほど、小さな値となる。なお、この計算式はあくまで一例に過ぎない。

【0056】

このように求めた評価値Vを、あらかじめ構造化処理部100に登録されている変化量についての閾値と比較し、評価値Vがその閾値以上であれば、読取画像から認識した属性「発注数量」の値は十分に「確からしい」と判定される。これに対し、評価値Vが閾値より小さければ、読取画像から認識した値が「発注数量」の値であることの確からしさは十分でないということである。

10

【0057】

図5の例では、波線のグラフが、直前の移動平均値に対する今回読み取った属性値の変化量を表している。この例では、前回までの変化量は、例えば20以下で安定していたが、今回の変化量はそれらに比べて突然大きくなっている。このようなケースでは、今回の評価値はそれまでに比べて大幅に低くなり閾値以下となる。

【0058】

確からしさ評価部130は、読取画像中から認識した各属性(要素)の値について上述のように評価値を求め(S14)、それら評価値を、読取画像から作成した構造化文書と共に登録処理部140に渡す。登録処理部140は、それら各属性の評価値をそれぞれ閾値と比較する(S15)。適切な閾値は、あらかじめ実験やシミュレーション等を行って求めておき、構造化処理部100に登録しておく。この比較の結果、それらすべての属性の評価値がすべて閾値以上であれば、文書解析部110の認識結果(すなわち作成した構造化文書の各要素の値)は十分に確からしいということなので、登録処理部140は、その構造化文書を文書DB150に登録する(S16)。一方、それら各属性の評価値の中に1つでも閾値未満のものがあれば、例外処理を実行する(S17)。

20

【0059】

例外処理では、例えば、評価値が閾値未満である属性の値をユーザに通知し、ユーザに確認を求める。評価値が閾値未満である属性が複数ある場合は、それら各属性について確認を求める。このとき、読取画像から作成した構造化文書はまだ文書DB150に登録しないままで保留する。その通知は、例えば操作パネル18に対し、その属性の値と名称との組を、その値の確認を促すメッセージ等と共に表示することで行えばよい。例えば「以下の属性の読取値は正しいかご確認ください。」等のメッセージと共に「商品CCCCの発注数量：5」などという属性値の情報を表示する。この表示画面には、例えば確認ボタンと、属性値を変更するためのユーザインタフェースとが表示される。ユーザは、例えば読み取った紙原稿と表示された属性値とを見比べ、表示された属性値が正しいか否かを判断する。ここで、読み取った属性値と、読取画像とを同時に操作パネル18に表示することで、ユーザの確認の便宜を図ってもよい。ユーザは、表示された属性値が正しければ確認ボタンを押下し、正しくなければ表示された属性値に対応する正しい値を入力する。ユーザが確認ボタンを押下した場合は、登録処理部140は、保留していた構造化文書を文書DB150に登録する。また、ユーザが属性値を修正した場合は、登録処理部140は、保留していた構造化文書中の当該属性に対応する要素の値を修正し、この修正を反映した構造化文書を文書DB150に登録する。

30

40

【0060】

以上の例では、評価値が閾値未満の属性値があった場合、すぐにユーザに通知して確認を求めたが、例外処理はこのようなものに限らない。例えば、作成した構造化文書に、当該構造化文書の各要素(属性)のうち評価値が閾値未満のものを特定する付属情報に対応づけて文書DB150に蓄積してもよい。その後で、管理者等が例えば定期的にそのような付属情報に対応づけられた構造化文書が蓄積されているか否かを確認し、そのような構造化文書があれば、それを評価値が閾値未満である属性値が正しいかどうかの確認作業を

50

行えばよい。この確認作業のためのシステムは、例えば、付属情報を元に構造化文書中で確認すべき属性の値を表示し、その構造化文書と対応づけて蓄積された読取画像をあわせて表示すればよい。管理者は、その表示をみて、その属性の値を確認すればよい。

【 0 0 6 1 】

以上の例では、新たに認識した属性値が、同一種類の過去の文書におけるその属性値の直近所定数回の平均値からみてどれだけ変化したかという変化量に基づき「確からしさ」を評価したが、これは一例に過ぎない。例えば直近所定数回の平均値からの「変化量」の代わりに、同一種類の過去のすべての文書における当該属性値の平均値からの偏差（新たに認識した値と平均値との差）に基づき評価を行ってもよい。差の絶対値が大きいほど、評価値は低くなる。差そのものの代わりに、偏差値など、平均からの偏差を表す他の統計的指標値を用いてもよい。

10

【 0 0 6 2 】

また、定期的に作成される文書を構造化して蓄積する場合には、直近所定数回の平均値からの変化量を用いる代わりに、前年の同じ時期（例えば前年同月）の文書における値からの変化量や、過去数年の同じ時期の文書における値の平均値からの変化量を用いて評価してもよい。

【 0 0 6 3 】

また、更に、読取画像の画質を推定し、前述の確からしさの評価値を、推定した画質に応じて補正してもよい。すなわち、読取画像の画質が悪いと、その読取画像から認識した属性値の確からしさも低くなると考えられるので、画質を確からしさの評価値に反映させるのである。

20

【 0 0 6 4 】

すなわち、この例では、構造化処理部 1 0 0 は、図 6 に示すように画質評価部 1 6 0 を備える。画質評価部 1 6 0 は、読取画像の画質評価値を求める。確からしさ評価部 1 3 0 は、画質評価部 1 6 0 が求めた画質評価値により、確からしさの評価値を補正する。

【 0 0 6 5 】

この補正には、例えば、図 7 に例示するように、画質の評価値が低くなるほど小さくなる係数を用いればよい。このような係数を、過去の同一属性の値との関係から求めた確からしさの評価値に乗じることで、補正された確からしさの評価値を求める。この補正された評価値をステップ S 1 5 で閾値と比較してもよい。

30

【 0 0 6 6 】

ここで、画質の評価値は、例えば、読取画像中での微小な連結成分の数に基づき求めることができる。ここでいう連結成分は、黒画素の連結成分であり、これを求めるには、読取画像を二値化する。文字認識などのためには読取画像を二値化することが一般的なもので、そのような二値化の結果を利用すればよい。「微小」というのは、読取ノイズ等のノイズに対応する大きさである。すなわち、請求書などの定型文書では、文字部分以外の「地」の部分は白等の単一の色と考えられるので、微小な連結成分の数が多い場合、それはノイズが多いということであり、読取画像の画質が悪いと判定できる。読取画像中の連結成分が「微小」か否かの判定では、その成分に含まれる黒画素の数を、あらかじめ定めた閾値（例えば数十画素）と比較すればよい。黒画素数が閾値以下である場合に、その成分が「微小」と判定できる。このように、微小であると判定された連結成分の単位面積あたりの数（即ち密度）に基づき、画質評価値を求める。概略的に言えば、微小な連結成分の密度が大きくなるほど、画質評価値は小さく（0 に近く）なるようにすればよい。

40

【 0 0 6 7 】

なお、画像読取装置 2 2 が読み取った画像には、画像処理回路 3 0 等によりノイズ除去の処理が行われる場合がある。画像処理回路 3 0 の設定値によっては、ノイズ除去処理が強行われすぎ、その結果本来の画像までノイズとして除去されてしまう可能性も考えられる。このような可能性を考慮して、微小な連結成分の数が少なすぎる場合に画質評価値を低くしてもよい。このような例を図 8 に示す。

【 0 0 6 8 】

50

以上、実施形態を説明した。以上の実施形態では、紙の文書を読み取ってその読み取り結果の画像から構造化文書を作成したが、本システムは、外部から入力された文書画像に対しても同様の処理を行うことができる。

【0069】

また、以上の例では、読み取った画像の解析に、像域分離回路28の出力を利用したが、これも必須のことではない。読み取った画像を領域分けする技術は他にも知られており、そのような他の技術を用いてもよい。

【0070】

また、上記実施形態のシステムは、画像形成装置をベースとしたものであったが、画像形成装置の代わりに、情報処理部100の機能を表すプログラムを汎用のコンピュータにインストールしても、同様の処理が行える。この場合、プログラムは、可搬型の記録媒体又はネットワーク経由でコンピュータにインストールすることができる。

10

【図面の簡単な説明】

【0071】

【図1】実施形態の画像形成装置のハードウェア構成の一例を示す図である。

【図2】構造化処理部の構成の一例を示す図である。

【図3】構造化処理部の処理手順の例を示す図である。

【図4】文書構造定義DBに登録されるデータの構造の例を示す図である。

【図5】属性値の変化量に基づく確からしさの判定を説明するための図である。

【図6】確からしさの評価値を読取画像の画質によって補正する機能を備えた構造化処理部の例を示す図である。

20

【図7】画質評価値と補正係数との関係の一例を示す図である。

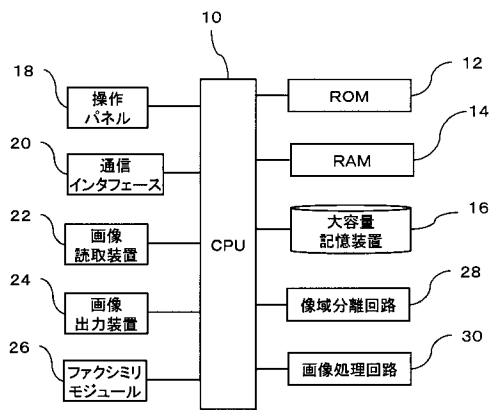
【図8】画質評価値の求め方の一例を説明するための図である。

【符号の説明】

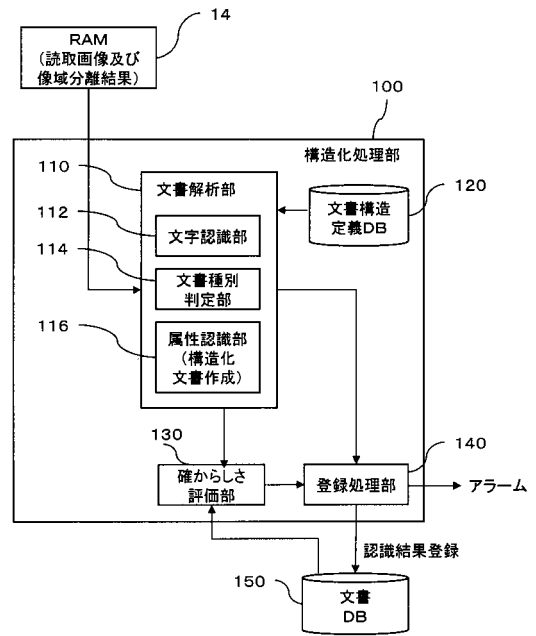
【0072】

14 RAM、100 構造化処理部、110 文書解析部、112 文字認識部、114 文書種別判定部、116 属性認識部、120 文書構造定義DB、130 確からしさ評価部、140 登録処理部、150 文書DB。

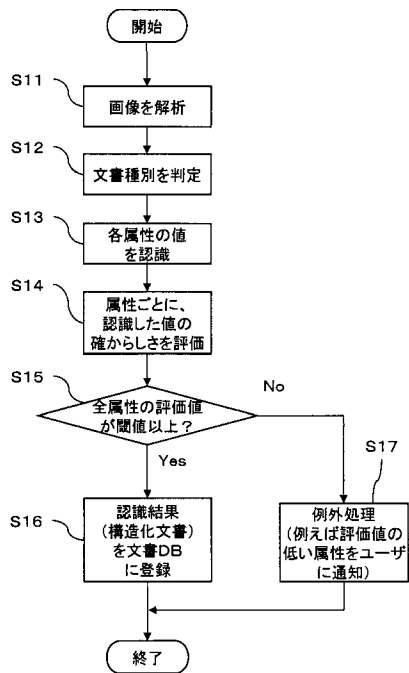
【図1】



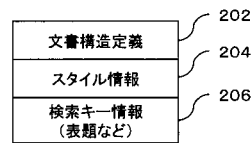
【図2】



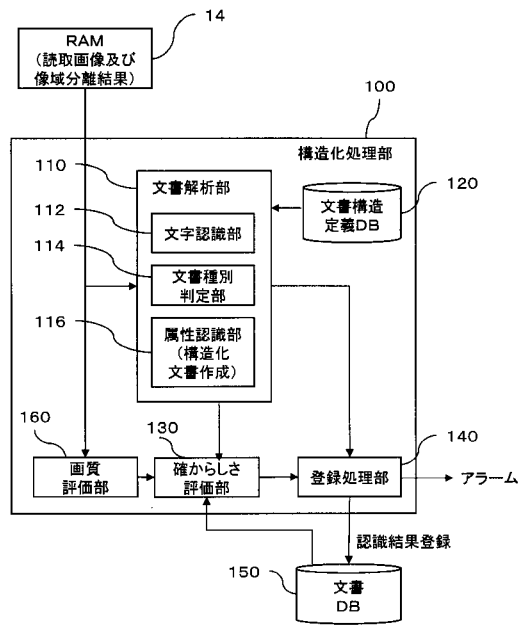
【図3】



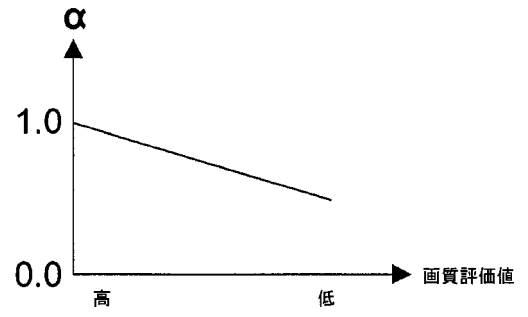
【図4】



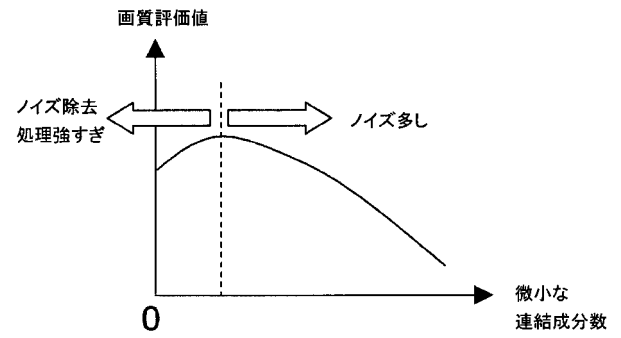
【図6】



【図7】



【図8】



【 図 5 】

