

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4613397号
(P4613397)

(45) 発行日 平成23年1月19日(2011.1.19)

(24) 登録日 平成22年10月29日(2010.10.29)

(51) Int.Cl.

G06K 9/20 (2006.01)

F 1

G06K 9/20 340 L

請求項の数 7 (全 17 頁)

(21) 出願番号 特願2000-195146 (P2000-195146)
 (22) 出願日 平成12年6月28日 (2000.6.28)
 (65) 公開番号 特開2002-15280 (P2002-15280A)
 (43) 公開日 平成14年1月18日 (2002.1.18)
 審査請求日 平成19年5月21日 (2007.5.21)

(73) 特許権者 303000372
 コニカミノルタビジネステクノロジーズ株
 式会社
 東京都千代田区丸の内一丁目6番1号
 (74) 代理人 100072349
 弁理士 八田 幹雄
 (74) 代理人 100110995
 弁理士 奈良 泰男
 (74) 代理人 100114649
 弁理士 宇谷 勝幸
 (74) 代理人 100129126
 弁理士 藤田 健
 (74) 代理人 100130971
 弁理士 都築 正則

最終頁に続く

(54) 【発明の名称】 画像認識装置、画像認識方法および画像認識プログラムを記録したコンピュータ読取可能な記録媒体

(57) 【特許請求の範囲】

【請求項 1】

画像認識装置において、

入力された画像データ中の文字画像から文字コードを認識する文字認識部と、

文字コードが認識された文字画像についての、近隣の文字画像との連続性の度合いである文字連続性度を検出する文字連続性度検出部と、

文字コードが認識された文字画像についての文字連続性度の、第1の閾値及びそれよりも小さい第2の閾値との大小関係を判定する出力形式判定部と、

を有し、

文字連続性度が第1の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連続性が第2の閾値よりも大きく第1の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成することを特徴とする画像認識装置。

【請求項 2】

前記文字連続性度検出部は、文字コードが認識された文字画像と該文字画像の近隣の文字画像との距離、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントサイズの差、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントのタイプの相違、文字コードが認識された文字画像が含まれる連続した文字画像の列の長さ、および文字コードが認識された文字画像と該文字画像の近隣の文字画像との色の差のうち、少なくともいずれか一つ以上を検出することを特徴とする請求項1に記載

の画像認識装置。

【請求項 3】

電子ファイルを作成するファイル作成部、をさらに有することを特徴とする請求項 1 または 2 に記載の画像認識装置。

【請求項 4】

原稿を読み取ることによって画像データを得るスキャナ部、をさらに有することを特徴とする請求項 1 ~ 3 のいずれかに記載の画像認識装置。

【請求項 5】

画像データを受信する画像データ受信部と、データを用紙に印刷するプリント部と、をさらに有することを特徴とする請求項 1 ~ 4 のいずれかに記載の画像認識装置。 10

【請求項 6】

画像認識方法において、

入力された画像データ中の文字画像から文字コードを認識する過程と、

文字コードが認識された文字画像についての、近隣の文字画像との連続性の度合いである文字連続性度を検出する過程と、

文字コードが認識された文字画像についての文字連続性度の、第 1 の閾値及びそれよりも小さい第 2 の閾値との大小関係を判定する過程と、

文字連続性度が第 1 の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連続性が第 2 の閾値よりも大きく第 1 の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成する過程と、 20

を有することを特徴とする画像認識方法。

【請求項 7】

画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、

入力された画像データ中の文字画像から文字コードを認識する過程と、

文字コードが認識された文字画像についての、近隣の文字画像との連続性の度合いである文字連続性度を検出する過程と、

文字コードが認識された文字画像についての文字連続性度の、第 1 の閾値及びそれよりも小さい第 2 の閾値との大小関係を判定する過程と、

文字連続性度が第 1 の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連続性が第 2 の閾値よりも大きく第 1 の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成する過程と、 30

をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、画像認識技術に関し、特に、入力された画像データ中の文字画像を、文字コードデータに変換するか否かを判定する処理についての画像認識技術に関する。 40

【0002】

【従来の技術】

最近の画像認識装置は、読み取りに関する条件の良い原稿（例えば同一フォントの文字のみで構成された原稿）であれば、極めて高い精度で文字画像を文字コードとして認識可能である。しかしながら、原稿上の文字の品質が悪かったり、原稿上の文字等のレイアウトが複雑であると、文字画像を文字コードとして認識する認識精度が大きく低下し、誤認識される文字画像が増加してしまう。

【0003】

このような問題に対し、原稿を読み取って得られた画像データのすべてをバックアップ用に保存するシステムが知られている。しかし、バックアップ用の画像データを含んで保存 50

されたファイルの容量は、当然ながら入力された画像データよりも大幅に増えてしまう。

【0004】

そこで、誤認識されている可能性が高い文字画像のみを、文字コードデータに変換することなく、文字画像データ（例えばビットマップ形式の画像データ）のまま出力する画像認識装置が提案されている。ここで、誤認識の可能性の判断には、例えばあらかじめ記憶されている標準パターンとの類似度などの結果から検出された、文字コードを認識する確からしさの情報が使用される。

【0005】

【発明が解決しようとする課題】

しかしながら、文字コードを認識する確からしさの情報だけでは、画像データ中の文字画像のうち、どの文字画像が正確に文字コードとして認識されていて、どの文字画像が誤認識されているかという判別の精度はあまり高いものではなく、結果的に誤認識される文字画像を十分に除去することが困難である。

10

【0006】

特に、最近では、読み取られる原稿は多様化しており、カラーの原稿や複雑なレイアウトの原稿などが増えている。これに伴って、画像データから文字画像が存在する文字領域を抽出することが難しくなってきている。したがって、例えば、図11に示すような、非文字である図形が文字部の中に埋め込まれた原稿を読み取る場合、図12に示すように、得られた画像データ中の非文字画像である図形画像から誤って文字コードを認識して、文字コードデータで出力される事態が生じ得る。なお、図中の符号E1は、図形画像から文字コードが誤認識されて得られた文字コードデータ、符号E2は、文字画像から文字コードが誤認識されて得られた文字コードデータを示す。しかも、図形画像から文字コードデータE1に変換される際の文字コードを認識する確からしさが、本来の文字画像とあまり変わらない場合がある。このため、文字コードを認識する確からしさの情報だけで文字コードデータに変換するか否かの判定を行うと、図形画像から誤認識されて得られた文字コードデータを除去することができない。

20

【0007】

一方、特開平8-55185号公報には、画像データ中の文字画像が存在する文字領域を抽出する段階で、近隣の文字画像との位置関係を調べながら文字領域を抽出するという技術が提案されている。また、特開平8-161421号公報には、文字領域を抽出する段階で、近隣の文字画像との位置関係を調べて文字領域の候補となる文字候補領域をまず抽出し、この文字候補領域内の画像データのうち文字コードを認識することができる文字画像によって文字領域を再構成するという技術が提案されている。しかしながら、上記公報に開示されたいずれの技術も、画像データ中から文字領域を抽出するための技術に過ぎない。

30

【0008】

本発明は、上記従来技術の問題点に鑑みてなされたものであり、本発明の目的は、入力された画像データ中の文字画像が文字コードデータに変換される割合を高くしつつ、誤認識の文字コードデータが出力されることを有效地に防止することにある。

【0009】

40

【課題を解決するための手段】

本発明の目的は、下記する手段により達成される。

【0010】

(1) 画像認識装置において、入力された画像データ中の文字画像から文字コードを認識する文字認識部と、文字コードが認識された文字画像についての、近隣の文字画像との連続性の度合いである文字連続性度を検出する文字連続性度検出部と、文字コードが認識された文字画像についての文字連続性度の、第1の閾値及びそれよりも小さい第2の閾値との大小関係を判定する出力形式判定部と、を有し、文字連続性度が第1の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連続性が第2の閾値よりも大きく第1の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コー

50

データに重なることなく連続して配置される文字画像データを作成することを特徴とする画像認識装置。

【0011】

(2) 前記文字連續性度検出部は、文字コードが認識された文字画像と該文字画像の近隣の文字画像との距離、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントサイズの差、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントのタイプの相違、文字コードが認識された文字画像が含まれる連続した文字画像の列の長さ、および文字コードが認識された文字画像と該文字画像の近隣の文字画像との色の差のうち、少なくともいずれか一つ以上を検出することを特徴とする上記(1)に記載の画像認識装置。 10

【0014】

(3) 電子ファイルを作成するファイル作成部、をさらに有することを特徴とする上記(1)または(2)に記載の画像認識装置。

【0015】

(4) 原稿を読み取ることによって画像データを得るスキャナ部、をさらに有することを特徴とする上記(1)~(3)のいずれかに記載の画像認識装置。 20

【0016】

(5) 画像データを受信する画像データ受信部と、データを用紙に印刷するプリント部と、をさらに有することを特徴とする上記(1)~(4)のいずれかに記載の画像認識装置。 20

【0017】

(6) 画像認識方法において、入力された画像データ中の文字画像から文字コードを認識する過程と、文字コードが認識された文字画像についての、近隣の文字画像との連續性の度合いである文字連續性度を検出する過程と、文字コードが認識された文字画像についての文字連續性度の、第1の閾値及びそれよりも小さい第2の閾値との大小関係を判定する過程と、文字連續性度が第1の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連續性度が第2の閾値よりも大きく第1の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成する過程と、を有することを特徴とする画像認識方法。 30

【0018】

(7) 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、入力された画像データ中の文字画像から文字コードを認識する過程と、文字コードが認識された文字画像についての、近隣の文字画像との連續性の度合いである文字連續性度を検出する過程と、文字コードが認識された文字画像についての文字連續性度の、第1の閾値及びそれよりも小さい第2の閾値との大小関係を判定する過程と、文字連續性度が第1の閾値よりも大きいと判定された文字画像を文字コードデータに変換し、文字連續性度が第2の閾値よりも大きく第1の閾値よりも小さいと判定された文字画像を切り出して、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成する過程と、をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。 40

【0019】

【発明の実施の形態】

以下、添付した図面を参照して、本発明の実施形態を説明する。

【0020】

図1は、本発明の第1実施形態に係る画像認識装置を含む情報機器システムの構成図である。

【0021】

情報機器システム100は、画像認識装置としての機能を有するコンピュータ10、画像入力装置であるスキャナ50、および、画像出力装置であるプリンタ60等の情報機器がネットワークN上に接続されて構成されている。 50

【 0 0 2 2 】

コンピュータ10、スキャナ50、および、プリンタ60は、ネットワークNを経由して互いにデータの授受を行うことができる。

【 0 0 2 3 】

図2は、コンピュータ10の概略構成を示すブロック図である。

【 0 0 2 4 】

コンピュータ10は、各情報機器との間でデータの授受を行うためのインターフェース(I/F)18と、スキャナ50を制御するためのスキャナドライバ12と、プリンタ60を制御するためのプリンタドライバ14と、所定のプログラムやデータを記憶するROM15と、一時的にデータを記憶するRAM16と、各種情報の表示やユーザからの指示入力が可能な操作パネル17と、文字認識等の処理を行う画像処理部13(詳細後述)と、上記各部を制御する主制御部11とを有している。なお、スキャナドライバ12、プリンタドライバ14は、ソフトウェアであり、図示しないハードディスク等の記憶装置に記憶されている。10

【 0 0 2 5 】

図1に示す情報機器システムの動作指示は、コンピュータ10の操作パネル17から行うことができる。本実施形態において操作パネル17は、コンピュータ10のディスプレイ上に表示される架空のパネルであり、コンピュータ10を操作するためのキーボードやマウス等の機器により、動作指示を行う。

【 0 0 2 6 】

例えば、操作パネル17上にはスキャン&ファイルボタンとスキャン&プリントボタンとが表示されており、スキャン&ファイルボタンで動作指示を行った場合には、スキャナ50は原稿読取動作を開始し、読み取りによって得られた画像データはコンピュータ10に送信され、画像処理部13にて文字認識等の画像処理がなされた後、所定の形式のファイルとしてコンピュータ10内の記憶装置に保存される。20

【 0 0 2 7 】

また、スキャン&プリントボタンで動作指示を行った場合には、スキャナ50の読み取りによって得られた画像データは、画像処理部13にて文字認識等の画像処理がなされた後、プリンタ60に送信され用紙上に印刷される。

【 0 0 2 8 】

次に、画像処理部13について詳細に説明する。

【 0 0 2 9 】

画像処理部13は、文字認識部31、文字色検出部32、出力形式判定部33、文字画像データ作成部34、文字消去部35、およびファイル作成部36を備えている。

【 0 0 3 0 】

文字認識部31は、個々の文字画像から文字コードを認識して文字コードデータを得るほか、文字コードを認識する確からしさの度合いである文字認識確度を検出する。文字画像から文字コードを認識する方法は、例えば、各文字画像の特徴量とあらかじめ記憶されている辞書パターンとの一致の度合いに基づいて行われる。文字認識確度は、例えば10段階で与えられ、値が大きいと文字認識確度が高いことを意味する。また、文字認識部31は、文字画像の位置情報を認識する。個々の文字画像の位置情報は、例えば図3に示すように、画像データの左上を原点として、各文字画像の外接矩形の左上および右下の座標として得られる。40

さらに、文字認識部31は、文字画像から、フォントのタイプやフォントサイズ、さらには太字・斜体などのスタイル、アンダーライン等の文字属性を認識する。

【 0 0 3 1 】

フォントのタイプは、例えばセリフ系とサンセリフ系などに分類されて判別される。また、フォントサイズは、個々の文字画像の外接矩形の大きさと、文字コードが認識されて得られた文字コードデータを一般的なフォントで表した場合の大きさとを比較することにより求められる。所定の代表的なフォントの文字高さや文字幅などの情報、すなわちフォン50

トメトリックデータは、あらかじめデータとして記憶されている。

【0032】

文字色検出部32は、文字コードが認識された個々の文字画像の色を検出する。

【0033】

出力形式判定部33は、文字コードが認識された文字画像を、文字コードデータに変換するか否かの判定を行う。本実施形態では、出力形式判定部33は、文字コードが認識された文字画像を調べて、文字コードデータに変換するか、文字コードデータに変換せずに、
10
入力された画像データ中の文字画像が切り出されて形成され、隣接する文字コードデータに重なることなく連続して配置される文字画像データを作成するか、あるいは入力された画像データのまま元の位置に残すか、の判定を行う。そして、文字画像データ作成部34は、文字コードが認識された文字画像のうち、文字画像データを作成する旨の判定が行われた文字画像を、入力された画像データから切り出して文字画像データを作成する。

【0034】

文字消去部35は、元の画像データ、すなわち入力されたカラーの画像データから、文字コードデータまたは文字画像データとされた後の元の文字画像を消去する。

【0035】

ファイル作成部36は、入力された画像データ中の文字画像から得られた文字コードデータや文字画像データなどを用いて、所定の形式のファイルを作成する。

【0036】

次に、画像処理部13の動作を図4および図5のフローチャートを用いて説明する。

20

【0037】

まず、スキャナ50で原稿を読み取ることによって得られた画像データに対し、RGB表色系からLab表色系へと色空間の変換を行う(S11)。Lab表色系は、スキャナあるいはプリンタ等の個々の情報機器に依存しないので、ネットワークN上で共通の色情報として扱うことができるものである。

【0038】

次に、文字認識部31に入力する画像データを作成するために、OCR前処理が行われる(S12)。後に文字認識部31がモノクロ2値画像データから特徴量を抽出して文字コードの認識を行うことから、OCR前処理では、まず、Lab表色系で表示されたカラー画像データに対して2値化処理を行うことにより、文字画像が黒で、その背景画像が白のモノクロ2値画像データを作成する。本実施形態では、文字画像およびその背景画像のL、a、bのレベルを検出することにより、例えば白地に書かれた色文字や、色地に書かれた黒文字等が原稿上に表されていても、原稿を読み取ることによって入力された画像データ中の文字画像から十分に文字コードを認識することができるよう2値化処理を行う。この他、OCR前処理では、入力された画像データ中の孤立点等のノイズの除去、原稿が傾いて読み取られた場合の補正処理、つぶれたりかすれたりしている文字画像の補正なども行われる。

30

【0039】

OCR前処理により得られたモノクロ2値画像データは、文字認識部31に入力され、文字認識処理が行われる(S13)。文字認識部31は、モノクロ2値画像データから文字画像が存在する文字領域の抽出を行った後、文字領域内の個々の文字画像から文字コードを認識する。また、文字認識部31は、文字コードを認識する確からしさの度合いである文字認識確度を検出することができ、さらに、文字画像の位置情報のほか、フォントのタイプ、フォントサイズ、太字・斜体などのスタイル、アンダーライン等の文字属性を認識する。

40

【0040】

次いで、文字色検出処理が行われる(S14)。すなわち、文字コードが認識された個々の文字画像の色は、文字色検出部32により検出される。ここでは、Lab表色系に色変換された画像データが使用される。文字画像の色の検出は、例えば、画像データ中の文字画像が存在する文字領域のLabの値を読み取り、文字画像を構成する全画素について、

50

L、a、bそれぞれの平均値を求ることにより行う。文字画像とその背景画像との画素の区別は、OCR前処理(S12)で得られたモノクロ2値画像データを用いて行う。つまり、個々の文字画像の外接矩形内において、モノクロ2値画像データにおける黒画素を文字画像の画素、白画素を背景画像の画素と区別することができる。

【0041】

文字画像を文字コードデータに変換して出力する場合、文字色検出部32により検出された個々の文字画像の色情報に基づいて、色を指定して文字コードデータが出力される。これにより、例えば原稿上の文字の色を再現することができる。

【0042】

そして、文字領域内の個々の文字画像について出力形式の判定処理が行われる(S15)。
すなわち、出力形式判定部33により、文字コードが認識された文字画像を、文字コードデータに変換するか否かの判定が行われる。

10

【0043】

出力形式判定部33は、大きく分けて2つの評価量に基づいて、出力形式の判定を行う。出力形式の判定を行うための第1の評価量は、文字コードを認識する確からしさの度合いである文字認識確度である。出力形式の判定を行うための第2の評価量は、文字コードが認識された文字画像についての、近隣の文字画像との連続性の度合いである文字連続性度である。この文字連続性度は、文字画像から認識される文字属性の連続性の面から、文字らしさを評価した値であり、その詳細は後述する。

【0044】

20

そして、出力形式判定部33は、文字コードが認識された文字画像のうち、文字コードデータに変換しないと判定した文字画像を、さらに2つの出力形式に分類して判定する。すなわち、出力形式判定部33は、文字コードが認識された文字画像が文字画像には違いないものの文字認識確度が小さい場合、隣接する文字コードデータに重なることなく連続して配置するために、入力された画像データ中の文字画像を切り出して文字画像データを作成すると判定する。また、出力形式判定部33は、文字コードが認識された文字画像が非文字画像である図形画像であるにもかかわらず誤って文字コードが認識されたと考えられる場合、当該文字画像を入力された画像データのまま元の位置に残すと判定する。

【0045】

次に、文字コードが認識された文字画像についての出力形式の判定処理についてさらに詳しく説明する。

30

【0046】

文字認識部31からは、個々の文字画像に対して、文字コードデータのほか、セリフ系・サンセリフ系などのフォントのタイプ、フォントサイズ、太字・斜体などのスタイル、アンダーライン等の文字属性、文字認識確度、文字コードが認識された各文字画像ごとの位置情報、などの情報が出力される。そして、これら文字認識部31からの情報をもとに、文字画像の位置や文字属性についての連続性を検出して文字らしさを評価する。前述したように、文字認識処理では、ある程度周辺の文字画像との位置関係を調べて、より文字らしい部分を文字領域として画像データから抽出するが、抽出された文字領域内に非文字画像である図形画像が文字画像と間違えられて含まれている場合がある。本実施形態では、特に、これらの孤立した文字画像を誤って文字コードデータに変換することのないように、すべての文字画像から文字コードを認識した後に、当該文字画像を文字コードデータに変換するか否かの判定を行なうようにしている。

40

【0047】

文字連続性度、すなわち文字らしさの評価量は、文字コードが認識された文字画像と該文字画像の近隣の文字画像との距離、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントサイズの差、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントのタイプの相違、文字コードが認識された文字画像が含まれる連続した文字画像の列の長さ、および文字コードが認識された文字画像と該文字画像の近隣の文字画像との色の差、の5項目について順に調べ、各項目において文字画像の連続性

50

に着目した評価ポイントを算出することにより検出される。

【0048】

具体的には、図5に示すように、まず、注目する個々の文字画像の順番を示す変数nを初期化する(S21)。

【0049】

そして、文字コードが認識された文字画像と該文字画像の近隣の文字画像との距離に基づいて、第1の評価ポイントを計算する(S22)。ここでは、注目文字画像の前後±k個の文字画像の位置情報(図3参照)が調べられ、次式により第1の評価ポイントP1nが計算される。なお、式中の添字は、個々の文字画像の順番を示す(以下の式において同様)。

10

【0050】

【数1】

$$P1_n = M1 - \min \left[\sum_{i=0}^k \{(X1_{n+i+1} - X2_{n+i}) + (Y1_{n+i+1} - Y2_{n+i})\}, \right.$$

$$\left. \sum_{i=-k}^0 \{(X1_{n+i+1} - X2_{n+i}) + (Y1_{n+i+1} - Y2_{n+i})\} \right]$$

【0051】

20

但し、 $P1n < 0$ のときは、 $P1n = 0$

M1：第1の評価ポイントの最大値

X1：文字コードが認識された文字画像の外接矩形の左X座標

X2：文字コードが認識された文字画像の外接矩形の右X座標

Y1：文字コードが認識された文字画像の外接矩形の上Y座標

Y2：文字コードが認識された文字画像の外接矩形の下Y座標

である。

【0052】

上記式に示すように、近隣文字画像との距離が大きい程、第1の評価ポイントは小さくなる。すなわち、注目文字画像が近隣文字画像から離れている程、連続性が小さく、文字画像である可能性が低いということを表している。

30

【0053】

なお、上記式からわかるように、例えば複数の文字画像が横に並んで形成された行の右端や左端の文字画像を考慮するため、注目文字画像の右側の近隣文字画像との距離と左側の近隣文字画像との距離とを調べて、両者のうち小さい方が評価ポイントの計算に使用される。

【0054】

次に、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントサイズの差に基づいて、第2の評価ポイントを計算する(S23)。ここでは、注目文字画像の前後±k個の文字画像のフォントサイズが調べられ、次式により第2の評価ポイントP2nが計算される。

40

【0055】

【数2】

$$P2_n = M2 - \frac{\sum_{i=-k}^k |S_{n+i} - S_n|}{(\sum_{i=-k}^k S_{n+i})/(2k+1)}$$

50

【0056】

但し、 $P_{2n} < 0$ のときは、 $P_{2n} = 0$

M_1 : 第2の評価ポイントの最大値

S : フォントサイズ

である。

【0057】

上記式の分母は、近隣文字画像の平均フォントサイズを表す。上記式に示すように、近隣文字画像とのフォントサイズの差が大きい程、連続性が小さく、第2の評価ポイントは小さくなる。

【0058】

次に、文字コードが認識された文字画像と該文字画像の近隣の文字画像とのフォントのタイプの相違に基づいて、第3の評価ポイントを計算する（S24）。

ここでは、注目文字画像の前後 $\pm k$ 個の文字画像のフォントのタイプが調べられ、注目文字画像とフォントのタイプが相違している文字画像の数を N_f として、次式により第3の評価ポイント P_{3n} が計算される。

【0059】

$$P_{3n} = M_3 - N_f$$

但し、 $P_{3n} < 0$ のときは、 $P_{3n} = 0$

M_3 : 第3の評価ポイントの最大値

である。

10

【0060】

上記式に示すように、近隣文字画像とのフォントのタイプの相違が多い程、連続性が小さく、第3の評価ポイントは小さくなる。

【0061】

次に、文字コードが認識された文字画像が含まれる連続した文字画像の列の長さに基づいて、第4の評価ポイントを計算する（S25）。ここでは、注目文字画像の前後に、順に個々の文字画像の外接矩形の間隔を調べ、当該間隔が所定の閾値 T_H 以上であれば、文字画像列の区切りであると判断する。なお、閾値 T_H は、例えば近隣文字画像の平均サイズによって変化させるのが望ましい。そして、注目文字画像を n 番目の文字画像として、注目文字画像の左側に存在する i_1 番目の文字画像から、注目文字画像の右側に存在する i_2 番目の文字画像までが、一つの文字画像列であると判断されると、次式により第4の評価ポイント P_{4n} が計算される。

20

【0062】

$$P_{4n} = i_2 - i_1 + 1$$

但し、 $P_{4n} > M_4$ のときは、 $P_{4n} = M_4$

M_4 : 第4の評価ポイントの最大値

である。

30

【0063】

上記式に示すように、文字画像列が長い程、連続性が大きいとされ、第4の評価ポイントは大きくなる。

40

【0064】

最後に、文字コードが認識された文字画像と該文字画像の近隣の文字画像との色に基づいて、第5の評価ポイントを計算する（S26）。ここでは、注目文字画像の前後 $\pm k$ 個の文字画像の色が調べられ、次式により第5の評価ポイント P_{5n} が計算される。

【0065】

【数3】

$$P5_n = M5 - \left\{ \sum_{i=-k}^k |L_{n+i} - L_n| + \sum_{i=-k}^k |a_{n+i} - a_n| + \sum_{i=-k}^k |b_{n+i} - b_n| \right\}$$

【0066】

但し、 $P5_n < 0$ のときは、 $P5_n = 0$

M5：第5の評価ポイントの最大値

L, a, b ：文字画像の色を L a b 表色系で表した値
である。

10

【0067】

上記式に示すように、近隣文字画像との色の差が大きい程、連続性が小さく、第5の評価ポイントは小さくなる。

【0068】

以上のようにして求められた第1～5の評価ポイントから、総合評価ポイントとしての文字連続性度 P_n が、次式により計算される（S27）。

【0069】

$P_n = W1 \times P1_n + W2 \times P2_n + W3 \times P3_n + W4 \times P4_n + W5 \times P5_n$

但し、 $W1 \sim W5$ は、それぞれの評価ポイントにたいするウェイトであり、適宜設定することができる。

20

【0070】

一方、この文字連続性度 P_n とは別に、文字コードを認識する確からしさの度合いである文字認識確度 C_n が、各文字画像ごとに得られる。

【0071】

そして、出力形式判定部33は、文字認識確度 C_n のほか、近隣の文字画像との連続性の度合いである文字連続性度 P_n を用いて、既に文字コードが認識された各文字画像について、出力形式の判定を行う。

【0072】

図6は、出力形式の判定基準を説明するための図である。

30

【0073】

まず、文字認識確度 C_n の閾値として、例えば小さい方の閾値 $CTh1$ と大きい方の閾値 $CTh2$ とが設定され、文字連続性度 P_n の閾値として、例えば小さい方の閾値 $PTh1$ と大きい方の閾値 $PTh2$ とが設定される。

【0074】

出力形式判定部33は、まず、文字認識確度 C_n が大きい方の閾値 $CTh2$ より大きい文字画像を、文字コードデータに変換する旨の判定を行う（S28でYES、S33）。出力形式判定部33は、また、文字認識確度 C_n が小さい方の閾値 $CTh1$ より大きく、且つ文字連続性度 P_n が小さい方の閾値 $PTh1$ より大きい文字画像も、文字コードを正しく認識していると判断して、文字コードデータに変換する旨の判定を行う（S29でYES、S33）。すなわち、文字認識確度 C_n が非常に大きい文字画像 ($C_n > CTh2$) は、文字コードデータに変換すると判定され、文字認識確度 C_n が中ぐらいの文字画像 ($CTh1 < C_n < CTh2$) は、文字連続性度 P_n が閾値 $PTh1$ より大きい場合に限って文字コードデータに変換すると判定される。

40

【0075】

上記の2つのステップS28およびS29で文字コードデータに変換しないと判定された文字画像は、文字連続性度 P_n により、さらに2つの出力形式に分類されて判定される。すなわち、文字コードデータに変換しないと判定された文字画像のうち、文字連続性度 P_n が閾値 $PTh2$ より大きい文字画像は、入力された画像データ中の文字画像を切り出して文字画像データを作成すると判定される。

50

一方、文字コードデータに変換しないと判定された文字画像のうち、文字連續性度 P_n が閾値 P_{Th2} 以下の文字画像は、入力された画像データのまま元の位置に残すと判定される。

【0076】

このように、文字認識部 31 により文字コードが認識された文字画像は、出力形式判定部 33 により、最終的に 3 つに分類される。すなわち、個々の文字画像は、確実に文字コードが認識されたと判断されて、文字コードデータで出力する旨の判定が行われたもの（図 6 中の A1 で示す領域）と、文字画像には違いないが正しい文字コードが認識できたか否かが疑わしい文字画像と判断されて、隣接する文字コードデータに重なることなく連続して配置される文字画像データで出力する旨の判定が行われたもの（図 6 中の A2 で示す領域）と、近隣文字画像との連續性が少なく、非文字画像である図形画像であるにもかかわらず誤って文字コードが認識された文字画像であると判断されて、入力された画像データのまま元の位置に残す旨の判定が行われたもの（図 6 中の A3 で示す領域）と、に分類される。

10

【0077】

そして、注目する文字画像の順番を示す変数 n を繰り上げ（S34）、画像データ内の全文字画像について、上記の処理を繰り返し行う（S35）。

【0078】

このようにして、画像データ中の個々の文字画像について出力形式の判定処理が終了すると、図 4 に示すメインフローチャートに戻り、文字画像データの作成が行われる（S16）。すなわち、文字画像データ作成部 34 は、文字認識部 31 から出力される文字画像の位置情報にしたがって、画像データ中の文字画像を切り出して文字画像データを作成する。このとき、個々の文字画像ごとに外接矩形領域の画像データが切り出される。

20

【0079】

ここで、文字画像データとして出力する文字画像を、原稿を読み取って得られた画像データとして元の位置に残さず、個々の文字画像単位で切り出すのは、文字コードデータと文字画像データとが重ならないようにするためである。つまり、一般に、入力フォントと出力フォントとが全く同じである場合は少ないので、例えば文字画像から変換された文字コードデータが配置される位置と、入力時の画像データ中の文字画像の位置とは、多少ずれてしまう。したがって、文字画像を一文字画像ずつ切り出さずに元の画像データとして残した場合、そのまま文字コードデータを配置すると、文字コードデータと原稿を読み取って得られた元の画像データ中の文字画像とが重なるといった問題が生じ得る。そこで、本実施形態では、一文字画像ずつ切り出しておいて、文字画像データを作成し、例えば文字コードデータを出力フォントで配置した位置に続けて、切り出しておいた文字画像データを配置する。

30

【0080】

また、文字画像データ作成部 34 は、文字画像データを作成する際に、文字画像の外接矩形内に対して、エッジ強調、または当該文字画像に適した 変換を行う。これにより、例えばビットマップで表された文字画像データでも、用紙に印刷して鮮明な文字を再現することができる。

40

【0081】

次に、文字消去処理が行われる（S17）。すなわち、文字画像から文字コードデータまたは文字画像データを得た後、文字消去部 35 により、原稿を読み取ることによって入力されたカラーの画像データから、文字コードデータまたは文字画像データとされた後の元の文字画像が消去される。

【0082】

図 7 は、文字消去処理を説明するための図である。図 7 に示すように、個々の文字画像の外接矩形内における画像データ（図 7（A））から、文字画像に当たる部分を一旦消去し（図 7（B））、周囲の画像データで補完する（図 7（C））。消去する部分は、当該文字画像の外接矩形内におけるモノクロ 2 値画像データから黒を 2 ~ 3 画素膨張させる処理

50

を行うことによって求める。なお、原稿を読み取ることによって入力された元の画像データから文字画像に当たる部分を除去した後の画像データには、図7（B）に示すように、文字画像以外の画像データ、すなわち、写真や線画に対応する画像、背景画像などが残っている。

【0083】

以上の処理を実施することにより、例えば図8に示すように、画像データ中の文字画像から、文字コードデータD1と文字画像データD2とが、また、非文字画像である図形画像をそのまま残した画像データD3が得られる。なお、図9は、文字コードデータ、文字画像データ、および画像データの内容を別々に示す図である。図8および図9に示されるように、非文字画像である図形画像であるにもかかわらず誤って文字コードが認識された孤立して存在している文字画像は、文字コードデータに変換されないので、図12のように図形画像の部分に誤認識の文字コードデータが出力されることはない。10

【0084】

そして、これらの文字コードデータ、文字画像データ、および画像データは、所定のファイル形式でページ内に配置されて保存される（S18）。なお、スキャン＆プリントボタンが選択されている場合は、所定のデータがプリンタ60に送信されて用紙に印刷される。。

【0085】

このように、本実施形態によれば、文字画像から文字コードを認識した後に、当該文字画像を文字コードデータに変換するか否かの判定を、文字連続性度を用いて行うようにしたので、非文字画像である図形画像であるにもかかわらず誤って文字コードが認識されて、文字コードデータとして出力される事態を有效地に防止することができる。20

【0086】

しかも、文字連続性度を用いることにより、文字コードを認識する確からしさに高目の厳しい閾値を設けなくても誤認識の文字コードデータを除去することが可能となるため、入力された画像データ中の文字画像が文字コードデータに変換される割合を高く維持することができる。

【0087】

本発明は、上記した実施形態のみに限定されるものではなく、特許請求の範囲内において、種々改変することができる。30

【0088】

例えば、上記実施形態では、入力された画像データの色空間を、RGB表色系からLab表色系に変換して画像処理するようにしたが、スキャナ50で読み取って得られた画像データのRGB表色系等、他の色空間を使用して画像処理を行うことも可能である。

【0089】

また、上述した文字画像の連続性に着目した第1～第5の評価ポイントを算出する方法は、上記実施形態で示した式等に限定されるものではない。例えば、次式により第2の評価ポイントP2nを計算することも可能である。

【0090】

【数4】

$$P2_n = M2 - \sum_{i=-k}^k |S_{n+i+1} - S_{n+i}|$$

【0091】

また、上記実施形態では、文字連続性度Pnは、第1～5の評価ポイントに基づいて総合的に求められるが、これに限定されるものではない。文字連続性度Pnは、例えば、第1～5の評価ポイントのうちの一つ、あるいは二つ以上を任意に組み合わせたものに基づいて求められるようにすることも可能である。50

【0092】

また、上記実施形態では、出力形式判定部33は、文字認識確度Cnと文字連續性度Pnとに基づいて、図6に示す方法で、各文字画像についての出力形式の判定を行っているが、これに限定されるものではない。出力形式判定部33は、例えば、文字認識確度Cnの閾値CThを文字連續性度Pnに応じて連続的に変化させる図10に示す方法で、各文字画像についての出力形式の判定を行うことも可能である。

【0093】

また、簡易な出力形式の判定処理を行う場合には、文字画像から文字コードを認識する確からしさの度合いを定量的に（例えば10段階で）検出することなく、文字画像から文字コードを認識できたか否かのみを検出し、出力形式判定部33が、文字コードが認識された文字画像を文字コードデータに変換するか否かを、文字連續性度を用いて判定するようにしてよい。

10

【0094】

さらに、上記実施形態では、画像認識装置としてコンピュータを例に挙げて説明したが、本発明の画像認識装置はこれに限定されるものではなく、例えば同様の処理をスキャナで行ってコンピュータやプリンタ等に送信するシステム、スキャナから直接画像データを受信してプリンタで同様の処理を行うシステム、原稿の読み取りから用紙への印刷までの処理をカラーのデジタルコピー機すべて行うシステム、等にも適用することが可能である。

20

【0095】

なお、上記した実施形態において、画像認識の制御は、上記した処理手順（図4、図5参照）を記述した所定のプログラムを主制御部11が実行することによって行われるものであり、この所定のプログラムは、コンピュータ読取可能な記録媒体（例えば、フロッピーディスクやCD-ROM等）によって提供されることもできる。また、この所定のプログラムは、上記各処理を実行するアプリケーションソフトウェアとして提供されてもよいし、コンピュータ、デジタルコピー機、スキャナ、プリンタ等の各情報機器や管理サーバの一機能として各情報機器や管理サーバのソフトウェアに組み込んでもよい。

【0096】**【発明の効果】**

以上説明したように、本発明によれば、文字画像から文字コードを認識した後に、当該文字画像を文字コードデータに変換するか否かの判定を、文字連續性度を用いて行うようにしたので、非文字画像である図形画像であるにもかかわらず誤って文字コードが認識されて、文字コードデータとして出力される事態を有効に防止することができる。

30

【0097】

しかも、文字連續性度を用いることにより、文字コードを認識する確からしさに高目の厳しい閾値を設けなくても誤認識の文字コードデータを除去することが可能となるため、入力された画像データ中の文字画像が文字コードデータに変換される割合を高く維持することができる。

【0098】

また、文字コードデータに変換しないと判定された文字画像のうち文字連續性度が所定値以上の文字画像を切り出して文字画像データを作成し、隣接する文字コードデータに重なることなく連続して配置するようにしたので、文字コードデータと文字画像データとを重なることなく出力することが可能となる。

40

【図面の簡単な説明】

【図1】 本発明の実施形態に係る画像認識装置を含む情報機器システムの構成を示す図である。

【図2】 図1に示されるコンピュータの概略構成を示すブロック図である。

【図3】 各文字画像の位置情報を説明するための図である。

【図4】 画像認識処理を示すフローチャートである。

【図5】 図4に示される出力形式判定処理を示すサブルーチンのフローチャートである

50

。

【図6】 出力形式の判定基準を説明するための図である。

【図7】 文字消去処理を説明するための図である。

【図8】 文字コードデータ、文字画像データ、および非文字画像である図形画像をそのまま残した画像データをあわせて配置した一例を示す図である。

【図9】 文字コードデータ、文字画像データ、および画像データの内容を別々に示す図である。

【図10】 出力形式の他の判定基準を説明するための図である。

【図11】 非文字である図形が文字部の中に埋め込まれた原稿の一例を示す図である。

【図12】 従来の画像認識装置により、図11の原稿を読み取って得られた画像データを画像処理して出力した一例を示す図である。 10

【符号の説明】

10 ... コンピュータ(画像認識装置)、

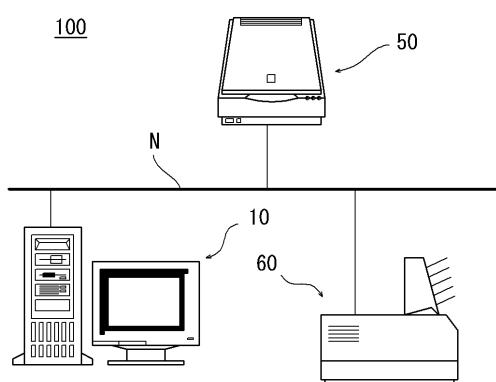
31 ... 文字認識部、

33 ... 出力形式判定部、

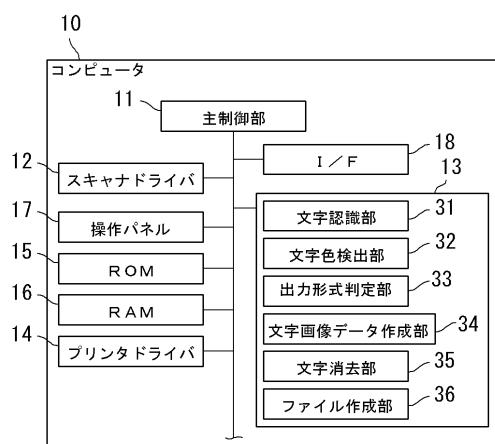
34 ... 文字画像データ作成部、

36 ... ファイル作成部。

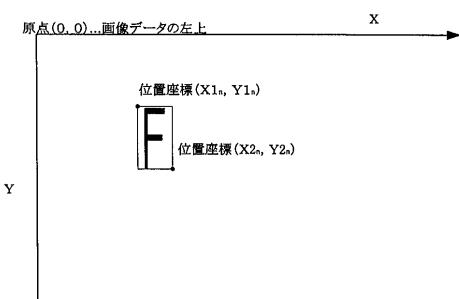
【図1】



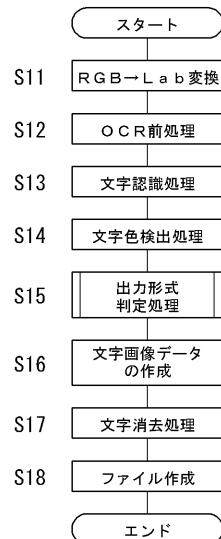
【図2】



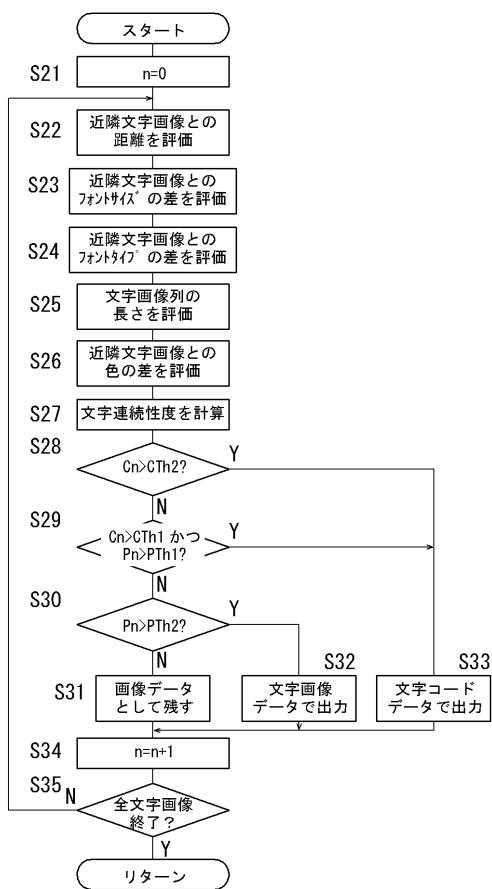
【図3】

Font type

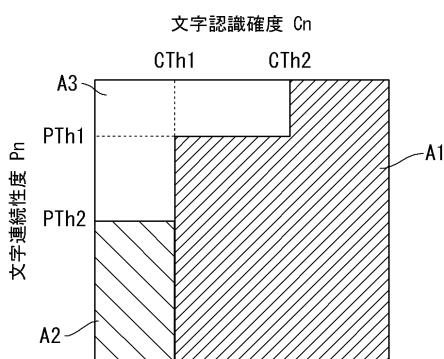
【図4】



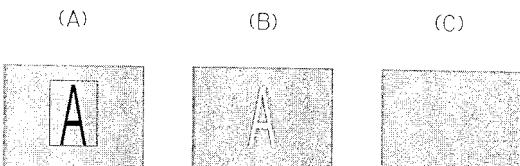
【図5】



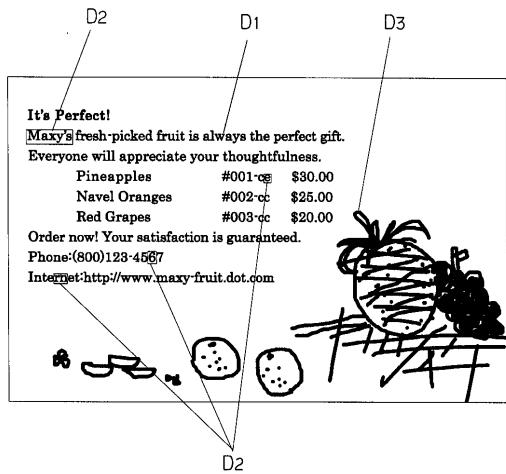
【図6】



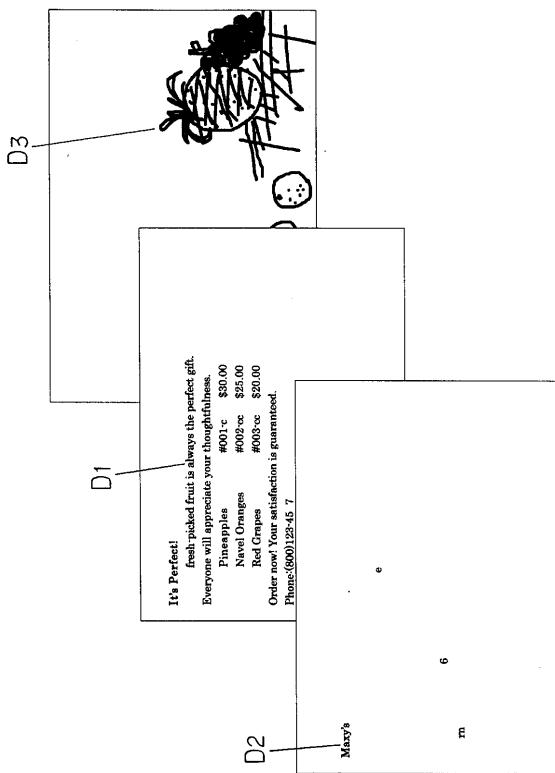
【図7】



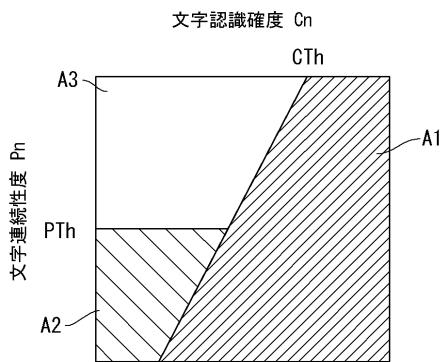
【図8】



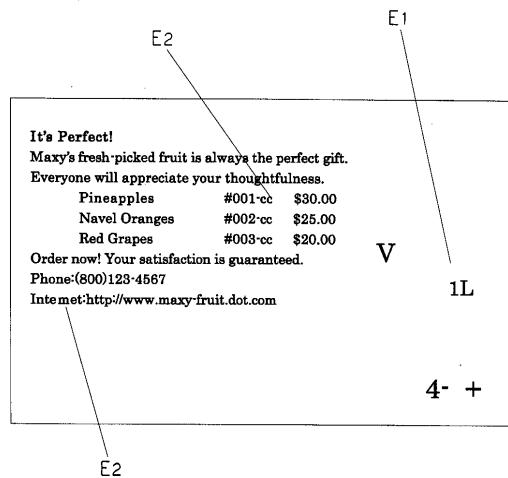
【図9】



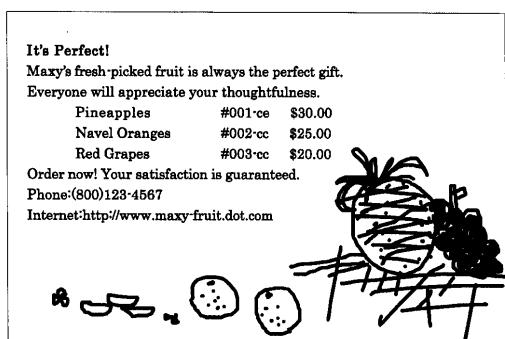
【図10】



【図12】



【図11】



フロントページの続き

(74)代理人 100134348

弁理士 長谷川 俊弘

(72)発明者 藤原 葉子

大阪市中央区安土町二丁目3番13号 大阪国際ビル ミノルタ株式会社内

審査官 宮里 達也

(56)参考文献 特開平01-191992(JP,A)

特開平02-090384(JP,A)

特開平04-293185(JP,A)

特開平05-108872(JP,A)

特開平11-085899(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06K 9/20