(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3792759号
(P3792759)
(45)発行日　平成18年7月5日(2006.7.5)　(24)登録日　平成18年4月14日(2006.4.14)

(51)Int.Cl.　F I
G06K 9/03 (2006.01)　G06K 9/03 C

請求項の数 11　(全 16 頁)

(21)出願番号　特願平7-235334
(22)出願日　平成7年9月13日(1995.9.13)
(65)公開番号　特開平9-81689
(43)公開日　平成9年3月28日(1997.3.28)
審査請求日　平成14年9月11日(2002.9.11)

(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　100093908
弁理士　松本　研一
(72)発明者　松林　一弘
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
(72)発明者　砂川　伸一
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　新井　則和

最終頁に続く

(54)【発明の名称】文字認識方法とその装置

(57)【特許請求の範囲】

【請求項１】

複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得工程と、

前記複数の異なる画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピング工程と、

前記生成された文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定工程と、

前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定工程と

を備えることを特徴とする文字認識方法。

【請求項２】

前記複数の異なる画像が第１の画像と第２の画像であるとき、

前記文字グループ判定工程では、前記第１の画像において生成された第１の文字グループに含まれる各文字パターンの候補文字コードと前記第２の画像において生成された第２の文字グループに含まれる各文字パターンの候補文字コードとを比較し、対応する全ての

文字パターンの候補文字コードの中に同一の候補文字コードがある場合、前記第１の文字グループと前記第２の文字グループとが同一の文字グループであると判定することを特徴とする請求項１に記載の文字認識方法。

【請求項３】

前記複数の異なる画像が動画像における異なるシーンの画像であり、且つ、前記抽出される文字パターンが当該動画像における動かないテロップであるという条件が加えられた場合、

前記文字グループ判定工程では、前記生成された複数の文字グループに含まれる文字パターンの、文字数と座標と大きさとに基づいて、前記複数の異なる画像間で同一の文字グループを判定することを特徴とする請求項１に記載の文字認識方法。

【請求項４】

前記文字候補獲得工程では、前記抽出された文字パターンから所定の特徴抽出を行って得られた特徴と、所定の辞書に格納されている各文字コードに対応する特徴とのマッチングに基づいて、前記候補文字コードとその確信度を獲得することを特徴とする請求項１に記載の文字認識方法。

【請求項５】

前記グルーピング工程では、前記複数の異なる画像の各々において、前記抽出した文字パターン間の距離が近い文字パターンを同じグループとしてグループ化することにより、複数の文字グループを生成することを特徴とする請求項１に記載の文字認識方法。

【請求項６】

前記グルーピング工程では、前記複数の異なる画像の各々において、前記抽出した文字パターン間の距離が近く、さらに、サイズが概同じ文字パターンを同じグループとしてグループ化することにより、複数の文字グループを生成することを特徴とする請求項５に記載の文字認識方法。

【請求項７】

前記文字コード決定工程では、前記同一の文字グループに含まれる文字パターンに対応する候補文字コードごとに確信度を合計した値を、該候補文字コードの前記集計された確信度とすることを特徴とする請求項１に記載の文字認識方法。

【請求項８】

前記文字コード決定工程では、前記同一の文字グループに含まれる文字パターンに対応する候補文字コードごとに求めた確信度の最大値を、該候補文字コードの前記集計された確信度とすることを特徴とする請求項１に記載の文字認識方法。

【請求項９】

前記複数の異なる画像は、動画像における複数の異なるシーンの画像であることを特徴とする請求項１に記載の文字認識方法。

【請求項１０】

複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得手段と、

前記複数の異なる画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピング手段と、

前記生成された複数の文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定手段と、

前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定手段と

を備えることを特徴とする文字認識装置。

【請求項１１】

メモリに記憶されているコンピュータ実行可能なプログラムを読み出してＣＰＵにおいて実行することにより、画像内の文字パターンに対応する文字コードを決定する装置であって、

前記プログラムは、複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得ステップと、前記異なる複数の画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピングステップと、前記生成された複数の文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定ステップと、前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定ステップとを、コンピュータに実行させるためのプログラムコードを含むことを特徴とする装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、画像から抽出した文字を認識する文字認識方法とその装置に関する。

【０００２】

【従来技術】

従来、画像から抽出した文字を認識して対応する文字コードを出力する装置が知られている。図１３は、従来の文字認識手順を示すフローチャートであり、簡単に説明する。

【０００３】

まず、Ｓ１０００で画像データ入力を行い、Ｓ１００２でその画像データ記憶させる。Ｓ１００３では、画像データから文字パターンを含む領域を抽出する。Ｓ１００４では、文字認識辞書１００５を探索して、文字認識処理を行ない、Ｓ１０１０では認識結果を記憶する。

【０００４】

【発明が解決しようとする課題】

しかしながら、前記従来の文字認識方法では、充分な認識の確度が得られないという問題があった。

本発明の目的は、上記従来例に鑑みてなされたもので、より高い確信度で文字認識が可能な文字認識方法とその装置を提供することを目的とする。

【０００５】

【課題を解決するための手段】

上記目的を達成するため、本発明の文字認識方法は、以下の構成を備える。即ち、複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得工程と、前記複数の異なる画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピング工程と、前記生成された文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定工程と、前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定工程とを備える。

【０００６】

上記文字認識方法において、前記複数の異なる画像が第１の画像と第２の画像であるとき、前記文字グループ判定工程では、前記第１の画像において生成された第１の文字グル

ープに含まれる各文字パターンの候補文字コードと前記第２の画像において生成された第２の文字グループに含まれる各文字パターンの候補文字コードとを比較し、対応する全ての文字パターンの候補文字コードの中に同一の候補文字コードがある場合、前記第１の文字グループと前記第２の文字グループとが同一の文字グループであると判定することを特徴とする。

【０００７】

上記文字認識方法において、前記複数の異なる画像が動画像における異なるシーンの画像であり、且つ、前記抽出される文字パターンが当該動画像における動かないテロップであるという条件が加えられた場合、前記文字グループ判定工程では、前記生成された複数の文字グループに含まれる文字パターンの、文字数と座標と大きさとに基づいて、前記複数の異なる画像間で同一の文字グループを判定することを特徴とする。

【０００８】

上記文字認識方法において、前記文字候補獲得工程では、前記抽出された文字パターンから所定の特徴抽出を行って得られた特徴と、所定の辞書に格納されている各文字コードに対応する特徴とのマッチングに基づいて、前記候補文字コードとその確信度を獲得することを特徴とする。

【０００９】

上記目的を達成するため、本発明の文字認識装置は、以下の構成を備える。即ち、複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得手段と、前記複数の異なる画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピング手段と、前記生成された複数の文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定手段と、前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定手段とを備える。

【００１０】

上記目的を達成するため、本発明の装置は、以下の構成を備える。即ち、メモリに記憶されているコンピュータ実行可能なプログラムを読み出してＣＰＵにおいて実行することにより、画像内の文字パターンに対応する文字コードを決定する装置であって、前記プログラムは、複数の異なる画像それぞれから複数の文字パターンを抽出して文字認識処理を行うことにより、当該抽出した文字パターンそれぞれに対応する複数の候補文字コードとその確信度を獲得する文字候補獲得ステップと、前記異なる複数の画像の各々において、前記抽出した各文字パターンの位置に基づいて前記文字パターンをグループ化することにより、複数の文字グループを生成するグルーピングステップと、前記生成された複数の文字グループそれぞれに含まれる各文字パターンに対応する複数の候補文字コードに基づいて、前記複数の異なる画像間で同一の文字グループを判定する文字グループ判定ステップと、前記判定された同一の文字グループそれぞれに含まれる文字パターンに対応する候補文字コードの確信度を集計し、前記集計された確信度に基づき、前記同一の文字グループに含まれる文字パターンに対応する文字コードを決定する文字コード決定ステップとを、コンピュータに実行させるためのプログラムコードを含む。

【００１１】

【発明の実施の形態】

はじめに、本発明に係る一実施の形態のポイントを要約した後に、その詳細な説明に入る。

本発明に係る文字認識方法の一実施の形態では、複数の画像から文字パターンを抽出し、抽出された文字パターンの認識処理を行って、文字パターンに対応する文字コードの候補とその確信度（得点）を求める。また、各文字パターンの位置関係で、互いに近いもの同

士を１つのグループとしてグルーピングする。そして、上述の複数の画像間で、そのグループ単位で、選ばれた文字コードの候補に基づいて、文字パターンの対応をとり、グループ単位でのマッチングをとる。

【００１２】

次に、マッチングが取られたグループに含まれる文字パターンについて、文字コード候補ごとの得点を集計する。そして、その得点集計結果に基づいて、最終的に、文字パターンに対する文字コードを確定する。

この様に、単独の画像単位で認識するのではなく、複数の画像間での文字グループのマッチングをとることに基づいて文字認識を行うことにより、より確度の高い文字認識が可能になった。

【００１３】

即ち、複数の画像間で対応する文字が異なった認識結果が得られた場合でも、上述の得点集計結果に基づいて、単独の画像内での認識結果を修正できるため、より信頼性の高い文字認識を可能とした。

次に、本実施の形態に係る文字認識装置の詳細な説明を行う。

（第１の実施の形態）

図１は、本実施の形態の文字認識装置のハードウェア構成図である。

【００１４】

２１は画像入力装置であり、ビデオカメラやイメージスキャナから画像を取り込んだり、ハードディスクやＣＤ－ＲＯＭから画像データを読み込んだり、通信機器から画像データを受信したりする。

２２は画像メモリであり、入力された画像データを一時的に記憶する。

２３はＣＰＵであり、後述する各文字認識処理のフローチャートに対応するプログラムを格納しているプログラム記憶メモリ２４から、プログラムを逐次読み出し、解釈し、実行する。

【００１５】

２５は辞書パターン記憶メモリであり、文字認識で用いる辞書パターンを記憶する。

２６はデータメモリであり、文字認識処理の途中結果や最終結果のデータを記憶したり、ＣＰＵによる文字認識処理の実行に必要な作業領域として使われる。２７は指示装置であり、キーボード、マウス、ペンなどを備え、本文字認識装置での各種動作の指示やデータ入力を行なう。

【００１６】

２８は表示装置であり、ＣＲＴやＬＣＤなどを備え、処理対象の画像や各種データを画面に表示する。

次に、図２は、図１のハードウェア上で処理される文字認識処理手順を示すフローチャートである。

尚、以下に示すフローチャートに対応するプログラムは、プログラム記憶メモリ２４に格納されており、ＣＰＵ２３によって実行される。

【００１７】

ステップＳ１では、画像入力装置２１から複数の静止画像または動画像を入力する。ステップＳ２では、入力された複数の画像データを画像メモリ２２に記憶する。ステップＳ３では、画像データから各文字パターンを含む領域を抽出する処理を行なう。

【００１８】

ステップＳ４では、辞書パターン記憶メモリ２５の辞書パターンを参照して、文字認識処理を行なう。これは、通常行われている所定の文字特徴に関して、辞書パターンの特徴と認識対象の文字の特徴との距離によって行えばよい。この距離の近い順に認識候補の文字コードが辞書から抽出され、各抽出された認識候補の文字コードにはその距離、即ち、確信度（例えば、距離の小さいものには、高い確信度、即ち、より大きな数字を対応させる）の情報がアタッチされる。

【００１９】

尚、後述する本ステップのさらに詳細な処理手順の一例を図１５に示す。
ステップＳ５では、ステップＳ４の認識結果の認識候補の文字コードと確信度をデータメモリ２６に格納する。また、各文字パターンの画像でのレイアウト位置の近く、また、同じサイズの文字パターンを同じグループとするグルーピングを行う。
【００２０】
ステップＳ６では、複数の入力画像間でグループ単位で、グループ内の文字コードが一致するものを探索する。
ステップＳ７では、ステップＳ６で一致が検出されたグループ単位で、そのグループに属する各文字パターンの認識候補文字コードの確信度の集計を行う。例えば、同じ認識候補文字コードごとの合計を取る。
【００２１】
ステップＳ８では、ステップＳ７での集計結果に基づいて、確信度の最も高い最適な候補を選択し、対応する文字パターンの認識結果として決定する。そして、その認識結果の文字を表示装置２８に出力する。
ステップＳ９では、決定した認識結果をデータメモリ２６に記憶する。
以上説明したように、単独の画像単位で認識するのではなく、複数の画像間での文字グループのマッチングをとることに基づいて文字認識を行うことにより、より確度の高い文字認識が可能になった。
（第２の実施の形態）
以上、第１の実施の形態の文字認識手順の概略を説明した。次に、第１の実施の形態での基本的処理コンセプトは、同じであるが、別の処理手順による実現方法をより詳細に説明する。
【００２２】
まず、図３のフローチャートを用いて、本発明の第２の実施の形態について説明する。
ステップＳ３１において、１枚分の画像データを画像メモリに入力する。例えば図４（ａ）のような地図の画像を入力する。
ステップＳ３２において、画像データからグループごとの文字群を抽出し、辞書２５を参照して、認識候補の文字コードとその確信度を得る。この結果を、データメモリ２６に格納する。
【００２３】
そして、例えば、その抽出され、認識された文字パターンに関して、図５のようなデータ構造を生成する。文字群のグループ分けは、各文字パターンが互いに近く配置されおり、サイズが同じものを集めことによっておこなう。
図５のデータ構造は、大きく３カテゴリに分類できる。第１のカテゴリは、５００、５０１に示すような、各グループごとに有するグループテーブルである。第２のカテゴリは、そのグループテーブル（５００、５０１）からアドレスポイントされ、各グループに含まれる各文字パターン（画像中の）のサイズに関する情報（ｘ，ｙ座標、幅、高さ、等）と文字パターンに対応する文字コードを格納している認識候補テーブル（５０３、５０４）へのアドレスポインタを有する文字パターン情報テーブル（５０２）である。最後の第３のカテゴリは、文字パターン情報テーブル（５０２）からアドレスポイントされ、各文字パターンに対応する認識結果である各認識候補の文字コードとその得点（即ち、確信度）を格納している認識候補テーブル（５０３、５０４）である。
【００２４】
本処理によって、図４（ａ）の画像に関して処理した場合において、生成されたデータ構造が図５であり、この内容に関して以下説明する。
図５を参照して、グループテーブル（５００）は、図４（ａ）の「宮沢町」に関するもので、他方、グループテーブル（５０１）は、図４（ａ）の「東区」に関するものである。
グループテーブル（５００、５０１）の左欄は、テーブル自体内での、各文字パターンに関する情報が格納されている位置のアドレスポインタ（右欄）を格納しているアドレスを示す。

【００２５】
尚、各文字の位置や大きさの情報からこれらの５文字のうち“宮”、“沢”、“町”の３文字は“宮沢町”という語を構成するとみなし、１つの文字群としてグループ単位でグループテーブル（５００）で管理している。同様に“東”、“区”の２文字の“東区”という語を構成する１つのグループとして、グループテーブル（５０１）で管理している。
【００２６】
文字パターン情報テーブル（５０２）のアドレス２０００Ｈから順に、“宮”、“沢”、“東”、“区”“町”の各文字の「ｘ座標、ｙ座標、幅、高さ」、と文字認識候補の情報が格納されている位置を示すアドレスポインタがそれぞれ格納される。尚、“東”、“区”、“町”の各パターンに対しては、記述を簡単にするため、アドレスポインタの行き先を示す矢印を省略している。
【００２７】
認識候補テーブル（５０３）のアドレス３０００Ｈからは、“宮”という文字パターンの認識候補として“宮”、“官”、“宜”の文字コードと得点が格納される。同様に、認識候補テーブル（５０４）のアドレス３０１０Ｈからは、“沢”という文字パターンの認識候補として“沢”、“況”という文字コードと得点が格納される。
【００２８】
尚、図示していないが、“東”、“区”“町”の各文字についても、同様に候補の文字コードと得点が格納される。
次に、ステップＳ３３において、上のステップで処理した画像が、１枚目の入力画像であるかチェックし、１枚目であれば、ステップＳ３５へ進む。そして、ステップＳ３５では、全画像について処理を終了したかチェックし、まだ、残りの画像があれば、ステップＳ３１に戻り、次の画像の入力を行う。２枚目の入力画像の例を図４（ｂ）に示す。以下、この画像に関して、上述のステップＳ３２、ステップＳ３３、ステップＳ３４の処理を実行してゆく。
【００２９】
そして、ステップＳ３５の判定で、全画像について処理を終了すれば、ステップＳ３６へ進み、全画像の全文字グループについて認識候補の中から、集計の確信度の高いものを選択して、文字コードを確定し、その結果を表示装置に表示する。尚、ステップＳ３６の処理は、図９を参照して後述する。
上述のステップＳ３３で、処理した画像が１枚目でなければ、ステップＳ３４に進み、ステップＳ３４では、１枚目の画像と２枚目の画像で、同一のグループ化された語があるかどうか判定する。本ステップでの詳細な処理は、図７を参照して後述するが、結果として、１枚目の画像で“宮沢町”と認識された文字グループと、２枚目の画像で“宮沢町”と認識された文字グループは、同一であると判定される。他方、１枚目の画像で“東区”と認識された文字グループと２枚目の画像で“西区”と認識された文字グループは同一でないと判定される。
【００３０】
尚、図４（ａ），（ｂ）の２枚の画像を入力して認識する例を示したが、２枚以上であってもよいことは言うまでもない。
次に、図６のフローチャートを用いて、図３のステップＳ３２での画像データから文字群を抽出する詳細な処理手順を説明する。
ステップＳ４１において、画像から文字を含む領域を抽出する。これには文字と背景の輝度や色の違いによって抽出する方法や、周波数成分の特徴の違いによって抽出する方法などを用いればよい。
【００３１】
ステップＳ４２において、文字を１文字ずつのパターンに分離する。これには特定の色の画素を連結したパターンについて形や大きさが所定の条件を満たすかどうかで判定する。
ステップＳ４３において、分離されたパターンに対して、辞書パターン記憶メモリ２５に格納されている辞書パターンを検索することで文字認識を行ない、候補の文字コード及び

得点（認識の確からしさ：確信度）を得る。また、文字のｘ座標、ｙ座標、幅、高さの情報も得る。

【００３２】

ステップＳ４４において、各文字の座標から所定の範囲内の距離にある文字を統合し、文字群としてグループ単位で管理する。例えば、注目する文字に対して上下左右１文字分の範囲内に他の文字があればその文字を同じ文字群に統合する。これにより、縦書き、横書き、斜め書きにかかわらず、１つの語を構成する文字をまとめる。

【００３３】

次に、図７のフローチャートを用いて、図３のステップＳ３４での処理、即ち、複数の画像間で、同一のグループ語があるかどうか判定する動作について説明する。

ステップＳ５１において、ステップＳ３２で選択された、認識文字コードの候補とその確信度を含む文字グループをデータメモリから取り込む。

【００３４】

ステップＳ５２において、ステップＳ５１で取り込んだ文字グループと一致する文字グループが他の画像の文字グループにあるかをチェックするために、含まれる文字コード単位で比較する。そして、同一の文字でなければステップＳ５６へ進み別の語と判定し、ステップＳ５５へ進む。

逆に、同一の文字があれば、ステップＳ５３へ進む。

【００３５】

ステップＳ５３では、文字グループ内の全文字について、比較が終了したかどうかをチェックする。そして、文字グループ内の全文字が一致すれば、ステップＳ５４において同一文字グループとし、その結果をデータメモリ２６に格納する。

ステップＳ５５では、全文字グループの組み合わせにおいて、比較が終了したかどうかチェックし、終了していなければ、ステップＳ５１に戻り、同様の処理を繰り返す。全文字グループの全組み合わせでの比較が終了すれば、本処理を終了する。

【００３６】

尚、上述のステップＳ５４やステップＳ５６での判定は、図８に示す様なグループ一致判定テーブルを、データメモリ２６上に作成して、一致マーク（図８では丸印）を格納しておくことで実現される。図８の表示で示されるように、画像１の文字グループ1と画像２の文字グループ２とは、文字１と文字１、文字２と文字２、文字３と文字３とがそれぞれ同じ候補を含むので同一の語とみなすことができる。また画像１の文字グループ２と画像２の文字グループ１とは、文字２と文字２は同じ候補を含むが文字１と文字１は同じ候補を含まないので別の語とみなす。

【００３７】

次に、図９のフローチャートを用いて、図３のステップ３６の文字コードの確定処理を説明する。

まず、ステップＳ６１において、同一の文字グループの組をデータメモリ２６から取り込む。

ステップＳ６２において、ステップＳ６１で取り込んだ文字グループ内の対応する同じ候補文字コードの得点を集計する。集計方法としては、各候補ごとの全画像の得点を加算してもよいし、また、各候補ごとの全画像中の最高点を取り出して得点としてもよい。

【００３８】

ステップＳ６３において、集計点が高い候補を選び出し、対応する文字の認識結果として確定し、表示装置に表示する。

図１０（ａ）は、集計処理の一例を説明するための図であり、ここでは、５枚の入力画像での文字パターン「宮」に対する、“宮”、“官”、“宜”の候補文字の得点の例として加算する方法を用いた集計結果を示す。この結果、得点合計は、それぞれ２５０点、２００点、４０点となり、“宮”が最も集計点が高い。したがって１から５までのすべての画像において、この文字を“宮”として確定する。すなわち２枚目の画像と５枚目の画像の誤認識が修正されることになる。

【００３９】
本実施の形態では、複数の画像の中に共通に含まれる語について共通の認識結果を出力するので、複数の画像のうち、ある画像の文字が誤認識しても、他の画像の認識結果によって自動的に修正することができ、結果として認識の確度を上げることができるという効果がある。
（第３の実施例）
本発明の第３の実施の形態においては、図１１のように動画像内のテロップの認識方法について説明する。
【００４０】
テロップは、背景画像の影響などで正しく認識されない場合がある。例えば、図１０（ｂ）の画像２のように、正しい文字“宮”の得点が０で、文字候補から漏れてしまう場合もあり得る。
そこで、背景画像が図１１（ａ）から図１１（ｂ）のように変わってもテロップは動かないという条件を加えることで、文字グループの文字数及び各文字の座標、大きさが同じであれば同一の語とみなすことができる。
【００４１】
図１２のフローチャートを用いて、同一の語があるかどうか判定する動作について説明する。
ステップＳ１２１において、２つの画像から文字グループを１つずつ入力する。
ステップＳ１２２において、文字グループの文字数が等しいかどうか判定し、等しくなければステップＳ１２８へ進み、別の語と判定する。
【００４２】
他方、文字数が等しいかどうかチェックし、等しければステップＳ１２３へ進む。
ステップＳ１２３では、文字の座標（ｘ座標とｙ座標の双方）が等しいかどうか判定し、等しくなければステップＳ１２８へ進み別の語と判定する。等しければステップＳ１２４へ進む。
【００４３】
ステップＳ１２４では、文字の大きさ（幅と高さの双方）が等しいかどうか判定し、等しくなければステップＳ１２８へ進み別の語と判定する。等しければステップＳ１２５へ進む。
ステップＳ１２５では、文字グループ内の全文字に関して処理が終了したかチェックし、終了していなければ、ステップＳ１２３に戻り、同様の処理を繰り返す。終了すれば、ステップＳ１２６に進み、文字グループ内の全文字について座標と大きさが等しい文字の組み合わせができたとして、同一文字グループと判定する。
【００４４】
ステップＳ１２７では、全文字グループの組み合わせについて、上述の処理がが終了したかチェックし、終了していなければ、ステップＳ１２１に戻り、終了するまで繰り返す。
上述の処理を、同一文字グループの判定にさらに加えることで、例えば、図１０（ｂ）に示すように、“宮”の合計点が最も高い結果が得られ、１から５までのすべての画像において、この文字を“宮”として確定できる。
【００４５】
尚、本実施の形態では、テロップのように文字が静止しているという条件をさらに加えることで、複数の画像のうち、ある画像が誤認識で候補から漏れても、他の画像の認識結果によって自動的に修正することができ、認識確度を上げることができるという効果がある。
尚、上述の説明では、複数の画像間での文字グループのマッチングに基づいて、文字パターンに対する文字コードの決定を行っている例を示したが、これは、複数の画像間ではなく、同じ画像内での文字グループのマッチングに基づいて、文字パターンに対する文字コードの決定を行ってもよいことは言うまでもない。
【００４６】

次に、図１４は、プログラム記憶メモリ２４にアサインされた上述したフローチャートの各処理に対応する各プログラムのレイアウトの一例を示す。尚、このプログラムは、フロッピーディスクなどの可搬可能な媒体に格納され、実行時に、メモリ２０２にロードされて、ＣＰＵ２３によって実行されてもよいことは言うまでもない。

【００４７】

１４０には、ステップＳ１－Ｓ２の画像入力／格納処理を行うプログラムが格納されている。

１４２には、ステップＳ３の文字パターン抽出処理を行うプログラムが格納されている。

１４３には、ステップＳ３で抽出された文字パターンに対応する候補文字パターンを辞書２５から抽出し、確信度を求める処理を行う文字認識プログラムが格納されている。

【００４８】

１４４には、ステップＳ３で抽出された文字パターンのグルーピング処理を行うプログラムが格納されている。

１４５には、ステップＳ７の得点（確信度）集計処理を行うプログラムが格納されている。

【００４９】

１４６には、ステップＳ８での文字コード決定処理を行うプログラムが格納されている。

１４７には、ステップＳ８での文字コード決定処理を行った後、表示装置にその結果を表示するプログラムが格納されている。

尚、図１４のレイアウトの順に、特別の意味はなく任意の順でよい。

【００５０】

次に、図２のステップＳ４の文字認識処理の詳細な処理の一例を、図１５を参照して説明する。

まず、図１５を参照して、ステップＳ１５１では、ステップＳ３で切り出された各文字パターンを入力する。

ステップＳ１５２では、入力した文字パターンの幅と高さをそれぞれ拡大、または、縮小して、所定の大きさに正規化する。

【００５１】

ステップＳ１５３では、正規化パターンを所定の閾値で２値化する。

ステップＳ１５４では、従来から知られている細線化の方法で２値化されたパターンを細線化処理する。

ステップＳ１５５では、細線化処理されたパターン、または、その所定の特徴量と文字認識辞書５とのマッチングを行う。

【００５２】

ステップＳ１５６では、マッチング結果として、マッチング距離の近い、即ち、確信度の高い文字候補を選択する。

以上、ステップＳ４での詳細な処理の一例を示した。

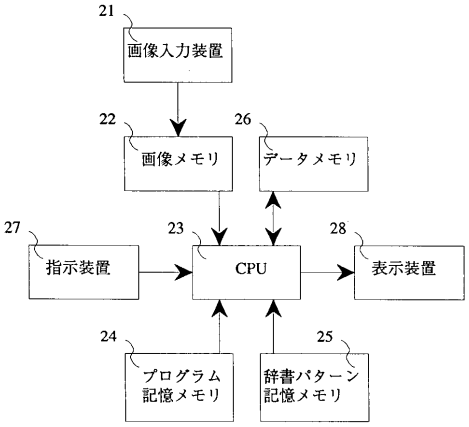
尚、本発明は、ホストコンピュータ、インタフェース、表示装置等の複数の機器から構成されるシステムに適用しても、１つの機器からなる装置に適用しても良い。また、本発明はシステム或は装置にプログラムを供給することによって達成される場合にも適用できることはいうまでもない。この場合、本発明を達成するためのソフトウェアによって表されるプログラムを格納した記憶媒体から、該プログラムを該システム或は装置に読み出すことによって、そのシステム或は装置が、本発明の効果を享受することが可能となる。

【００５３】

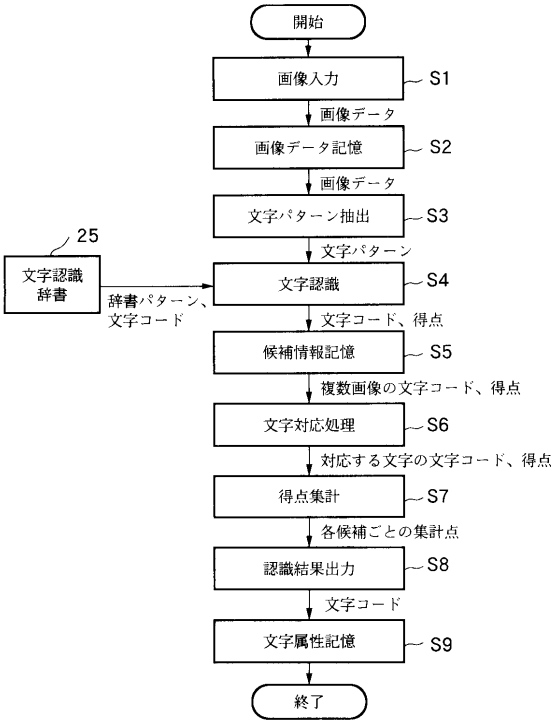
以上述べたように、画像から文字パターンを抽出する手段と、抽出された文字パターンを認識し文字コードの候補及び各候補の得点を出力する手段と、複数の画像間でグループ単位で文字パターンを対応させる手段と、対応するグループの文字パターンについて文字コードの候補ごとの得点を集計する手段と、得点集計結果に基づいて前記文字コードの候補及び各候補の得点を修正する手段とを設けたことにより、より確度の高い文字認識を行うことができる。

【００５４】
【発明の効果】
以上説明したように本発明によれば、より高い確信度で文字認識が可能な文字認識をおこなうことができる。
【図面の簡単な説明】
【図１】本発明の文字認識装置のブロック図である。
【図２】本発明の文字認識処理手順を示すフローチャートである。
【図３】本発明の第２の実施の形態の文字認識処理手順を示すフローチャートである。
【図４】入力画像の一例を示す図である。
【図５】画像データから文字グループを抽出した結果得られるデータの例を示した図である。
【図６】画像データから文字グループを抽出する動作を示したフローチャートである。
【図７】同一グループ語を判定する動作を示したフローチャートである。
【図８】複数の画像間での文字の対応の例を示したグループ一致判定テーブルの図である。
【図９】判定認識候補の中から文字コードを確定する処理手順を示したフローチャートである。
【図１０】複数の入力画像に対する各候補の得点の例を示す図である。
【図１１】入力画像の例を示す図である。
【図１２】同一語を判定する処理手順を示したフローチャートである。
【図１３】従来の文字認識装置のフローチャートである。
【図１４】本実施の形態に係る文字認識処理の各プログラムのレイアウトを示す図である。
【図１５】図２のステップＳ４での文字認識処理の詳細な処理例を示すフローチャートである。
【符号の説明】
２１　画像入力装置
２２　画像メモリ
２３　ＣＰＵ
２４　プログラム記憶メモリ
２５　辞書パターン記憶メモリ

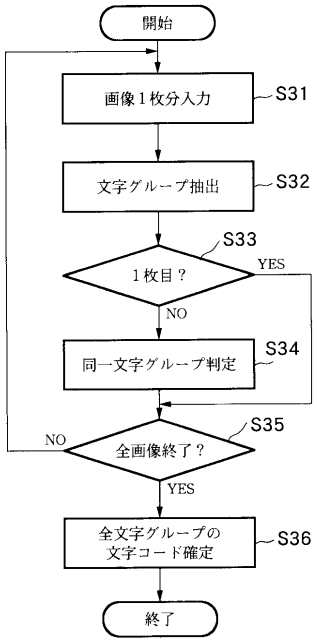
【図１】


21
画像入力装置
22
画像メモリ
26
データメモリ
27
指示装置
23
CPU
28
表示装置
24
プログラム記憶メモリ
25
辞書パターン記憶メモリ

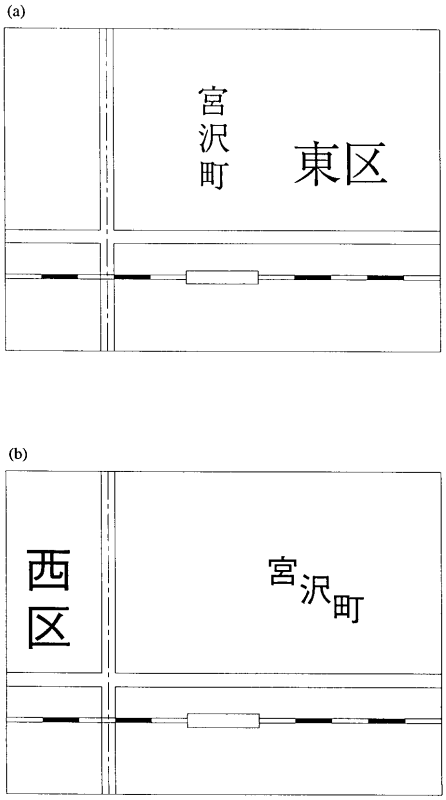
【図２】


開始
画像入力
S1
画像データ
画像データ記憶
S2
画像データ
文字パターン抽出
S3
文字パターン
25
文字認識辞書
辞書パターン、文字コード
文字認識
S4
文字コード、得点
候補情報記憶
S5
複数画像の文字コード、得点
文字対応処理
S6
対応する文字の文字コード、得点
得点集計
S7
各候補ごとの集計点
認識結果出力
S8
文字コード
文字属性記憶
S9
終了

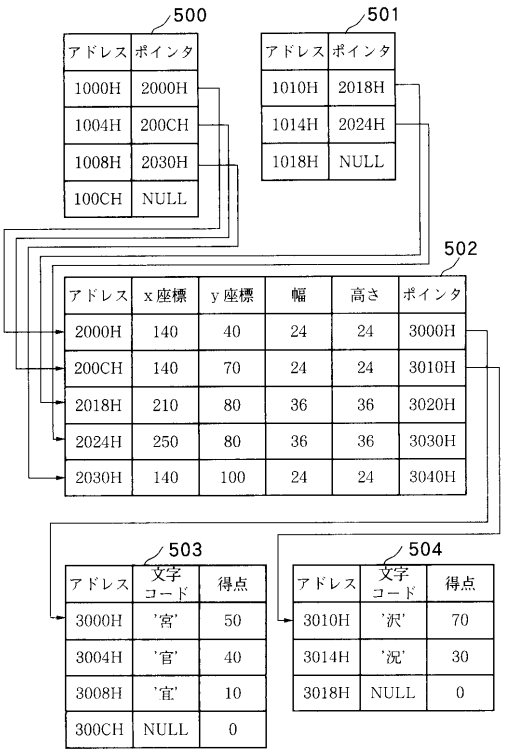
【図３】


開始
画像１枚分入力
S31
文字グループ抽出
S32
S33
１枚目？
YES
NO
同一文字グループ判定
S34
S35
NO
全画像終了？
YES
全文字グループの文字コード確定
S36
終了

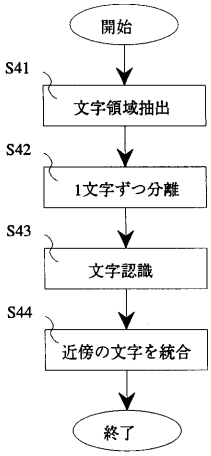
【図４】


(a)
宮沢町
東区
(b)
西区
宮沢町

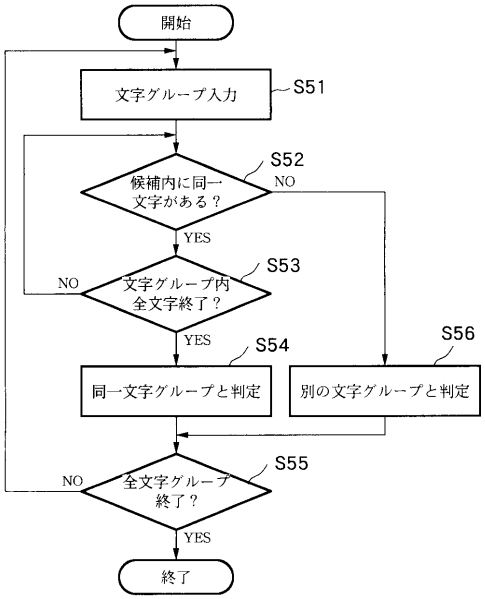
【図５】



【図６】



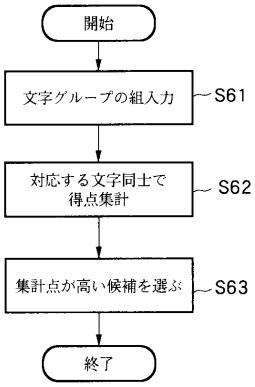
【図７】



【図８】

					画像2									
					文字グループ1				文字グループ2					
					文字1		文字2		文字1		文字2			文字3	
					候補1	候補2	候補1	候補2	候補1	候補2	候補1	候補2	候補3	候補1	候補2
					西	酉	区	凶	官	宮	沢	決	況	町	司
画像1	文字グループ1	文字1	候補1	宮						○					
			候補2	官					○						
			候補3	宜											
		文字2	候補1	沢							○				
			候補2	況									○		
		文字3	候補1	町										○	
			候補2	司											○
	文字グループ2	文字1	候補1	東											
			候補2	車											
			候補3	束											
		文字2	候補1	区			○								
			候補2	匹											

【図９】



【図１０】

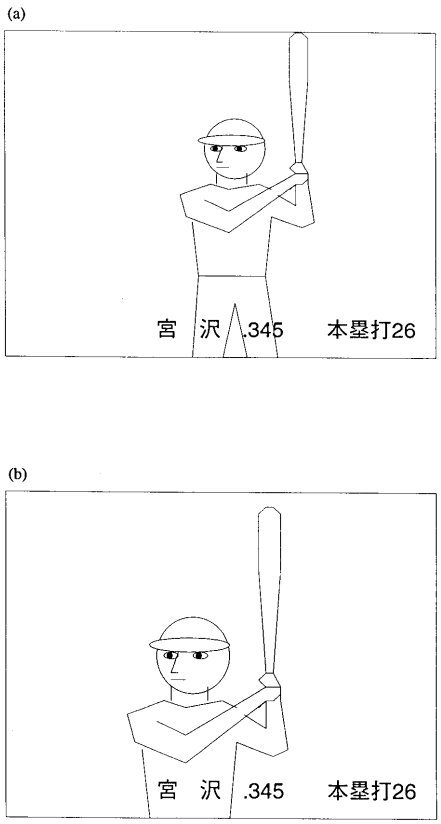
(a)

候補文字＼画像	1	2	3	4	5	計
宮	50	40	50	70	40	250
官	40	60	30	20	50	200
宜	10	0	20	10	0	40

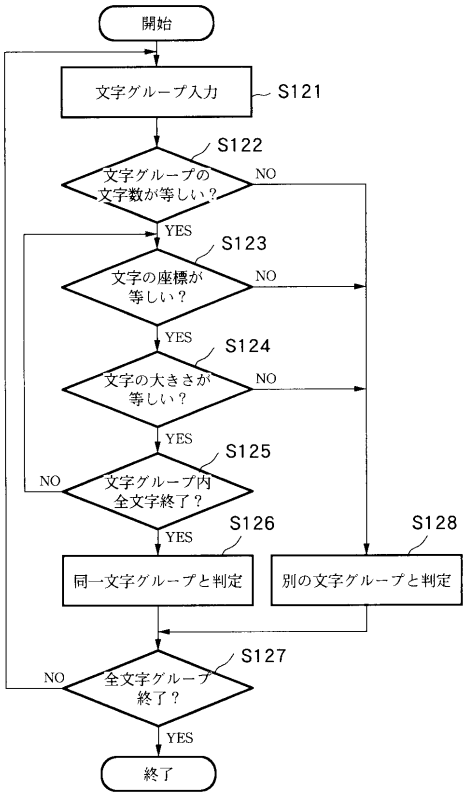
(b)

候補文字＼画像	1	2	3	4	5	計
宮	50	0	50	70	40	250
官	40	0	30	20	50	200
宜	10	0	20	10	0	40
客	0	30	0	0	0	30
容	0	10	0	0	10	20

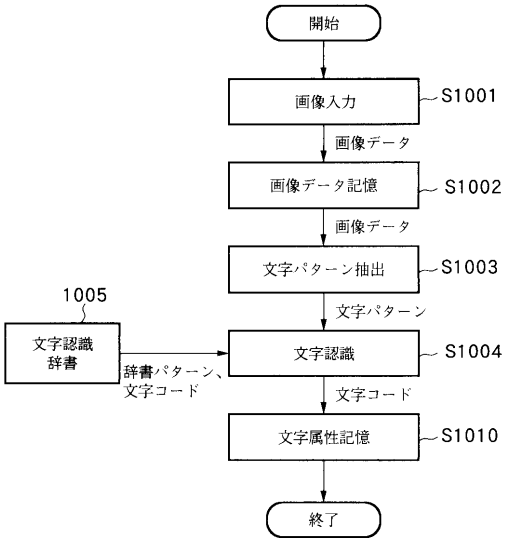
【図１１】



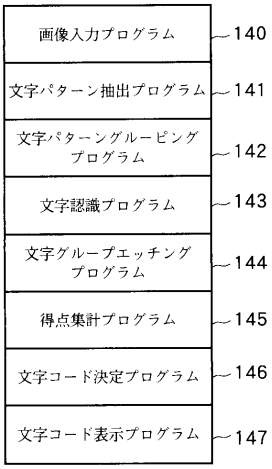
【図１２】



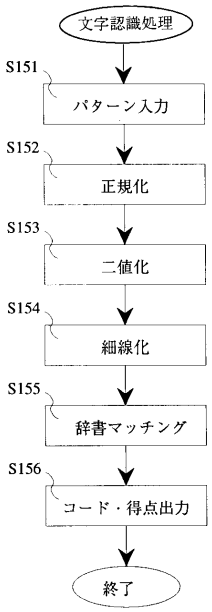
【図１３】


開始
画像入力
S1001
画像データ
画像データ記憶
S1002
画像データ
文字パターン抽出
S1003
1005
文字認識辞書
文字パターン
文字認識
S1004
辞書パターン、文字コード
文字コード
文字属性記憶
S1010
終了

【図１４】


画像入力プログラム
140
文字パターン抽出プログラム
141
文字パターングルーピングプログラム
142
文字認識プログラム
143
文字グループエッチングプログラム
144
得点集計プログラム
145
文字コード決定プログラム
146
文字コード表示プログラム
147

【図１５】


文字認識処理
S151
パターン入力
S152
正規化
S153
二値化
S154
細線化
S155
辞書マッチング
S156
コード・得点出力
終了

フロントページの続き

(56)参考文献 特開平０５－２５８１００（ＪＰ，Ａ）
特開平０５－１８９６０６（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
G06K 9/00～9/76