

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4594765号
(P4594765)

(45) 発行日 平成22年12月8日(2010.12.8)

(24) 登録日 平成22年9月24日(2010.9.24)

(51) Int. Cl. F I
G06K 9/03 (2006.01) G06K 9/03 C
G06K 9/62 (2006.01) G06K 9/62 630C

請求項の数 3 (全 16 頁)

(21) 出願番号	特願2005-64386 (P2005-64386)	(73) 特許権者	000004226
(22) 出願日	平成17年3月8日(2005.3.8)		日本電信電話株式会社
(65) 公開番号	特開2006-251920 (P2006-251920A)		東京都千代田区大手町二丁目3番1号
(43) 公開日	平成18年9月21日(2006.9.21)	(74) 代理人	100096459
審査請求日	平成19年2月21日(2007.2.21)		弁理士 橋本 剛
		(74) 代理人	100086232
			弁理士 小林 博通
		(74) 代理人	100104938
			弁理士 鶴澤 英久
		(74) 代理人	100140361
			弁理士 山口 幸二
		(72) 発明者	草地 良規
			東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 文字認識装置、文字認識方法、及び文字認識プログラムの記録媒体

(57) 【特許請求の範囲】

【請求項1】

画像中の文字を認識する文字認識装置であって、

入力した画像を走査しながら局所画像を切り出して特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度に基づいて文字候補を決定する第一段文字認識手段と、

前記決定した文字候補について、既に特徴抽出に用いられた特徴とは異なる特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度から前記文字候補の絞込みを行う一以上の異特徴文字認識手段と、を有し、

前記第一段文字認識手段と前記一以上の異特徴文字認識手段のうちのいずれか1つの手段が、

入力された画像について、複数の方向に線の線幅を検出する線幅検出手段と、

前記各方向の線の線幅を特徴値とした特徴ベクトルを出力する出力手段と、を有し、

前記線幅検出手段は、

線幅を検出する方向および線の線幅に応じた複数のフィルタを生成するフィルタ生成手段と、

前記画像と前記複数のフィルタとの一致度を算出する一致度算出手段と、

前記算出した一致度において最大の一致度を示すフィルタの線の線幅の値と前記一致度を乗算して得た値を線の線幅とする線幅計算手段と、

を有することを特徴とする文字認識装置。

【請求項2】

10

20

画像中の文字を認識する文字認識装置における文字認識方法であって、

第一段文字認識手段が、入力した画像を走査しながら局所画像を切り出して特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度に基づいて文字候補を決定する第一段文字認識ステップと、

一以上の異特徴文字認識手段が、前記決定した文字候補について、既に特徴抽出に用いられた特徴とは異なる特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度から前記文字候補の絞込みを行う異特徴文字認識ステップと、を有し、

前記第一段文字認識ステップと前記異特徴文字認識ステップのいずれかのステップにおいて、

線幅検出手段が、入力された画像について、複数の方向に線の幅を検出する線幅検出ステップと、

出力手段が、前記各方向の線の線幅を特徴値とした特徴ベクトルを出力する出力ステップと、を有し、

前記線幅検出ステップにおいて、

フィルタ生成手段が、線幅を検出する方向および線の線幅に応じた複数のフィルタを生成するフィルタ生成ステップと、

一致度算出手段が、前記画像と前記複数のフィルタとの一致度を算出する一致度算出ステップと、

線幅計算手段が、前記算出した一致度において最大の一致度を示すフィルタの線の線幅の値と前記一致度とを乗算して得た値を線の線幅とする線幅計算ステップと、

を有することを特徴とする文字認識方法。

【請求項 3】

上記請求項 1 または 2 いずれかに記載の文字認識装置又は文字認識方法を、コンピュータで実行可能に記載したプログラムを記録したことを特徴とする記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、文字列を画像として撮影し、この文字列を識別する画像識別技術に関するものである。この具体的な産業応用システムとして、例えば看板の日本語翻訳システムなどが挙げられる。

【背景技術】

【0002】

景観に存在する文字列識別は、一般的には文字列位置特定、文字領域特定、2値化、及び文字識別という4ステップを経る。しかし、このような技術では、照明変動や複雑な背景などの影響によって、文字列位置特定、文字領域特定、2値化がうまくいかず、文字列識別精度が低いという問題があった。

【0003】

このような問題を解決するために、景観に存在する文字列識別技術として、全画面探索、文字候補絞込み、及び言語モデルによる文字列推定という3ステップを経る方法がある(例えば、非特許文献1参照。)

【非特許文献1】草地良規、伊藤直己、鈴木章、荒川賢一、「画像インデクシングを目的としたテキスト領域不要の景観中文字認識」、電子情報通信学会 信学技報 PRMU2004-89、(2004-10)、p.37-42

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、上記方法では、特徴の定義及び識別アルゴリズムに限界があり、背景と文字との区別が完全にはつかず、背景部分に文字候補が多数存在してしまい、言語モデルによる文字列推定がうまく働かないという課題があった。

【0005】

10

20

30

40

50

本発明は、かかる事情に鑑みてなされたものであり、その目的は、上記課題を解決した文字認識技術を提供することにある。

【課題を解決するための手段】

【0006】

そこで、上記課題を解決するために、請求項1に記載の発明は、画像中の文字を認識する文字認識装置であって、入力した画像を走査しながら局所画像を切り出して特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度に基づいて文字候補を決定する第一段文字認識手段と、前記決定した文字候補について、既に特徴抽出に用いられた特徴とは異なる特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度から前記文字候補の絞込みを行う一以上の異特徴文字認識手段と、を有し、前記第一段文字認識手段と前記一以上の異特徴文字認識手段のうちのいずれか1つの手段が、入力された画像について、複数の方向に線の線幅を検出する線幅検出手段と、前記各方向の線の線幅を特徴値とした特徴ベクトルを出力する出力手段と、を有し、前記線幅検出手段は、線幅を検出する方向および線の線幅に応じた複数のフィルタを生成するフィルタ生成手段と、前記画像と前記複数のフィルタとの一致度を算出する一致度算出手段と、前記算出した一致度において最大の一致度を示すフィルタの線の線幅の値と前記一致度とを乗算して得た値を線の線幅とする線幅計算手段と、を有することを特徴とする。

10

【0010】

また、請求項2に記載の発明は、画像中の文字を認識する文字認識装置における文字認識方法であって、第一段文字認識手段が、入力した画像を走査しながら局所画像を切り出して特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度に基づいて文字候補を決定する第一段文字認識ステップと、一以上の異特徴文字認識手段が、前記決定した文字候補について、既に特徴抽出に用いられた特徴とは異なる特徴を抽出し、当該特徴と辞書に登録された文字の特徴との類似度から前記文字候補の絞込みを行う異特徴文字認識ステップと、を有し、前記第一段文字認識ステップと前記異特徴文字認識ステップのいずれかのステップにおいて、線幅検出手段が、入力された画像について、複数の方向に線の幅を検出する線幅検出ステップと、出力手段が、前記各方向の線の線幅を特徴値とした特徴ベクトルを出力する出力ステップと、を有し、前記線幅検出ステップにおいて、フィルタ生成手段が、線幅を検出する方向および線の線幅に応じた複数のフィルタを生成するフィルタ生成ステップと、一致度算出手段が、前記画像と前記複数のフィルタとの一致度を算出する一致度算出ステップと、線幅計算手段が、前記算出した一致度において最大の一致度を示すフィルタの線の線幅の値と前記一致度とを乗算して得た値を線の線幅とする線幅計算ステップと、を有することを特徴とする。

20

30

【0014】

また、請求項3に記載の発明は、上記請求項1または2いずれかに記載の文字認識装置又は文字認識方法を、コンピュータで実行可能に記載したプログラムを記録したことを特徴とする。

【発明の効果】

【0015】

請求項1～3に記載の発明によれば、第一段文字認識手段で決定した文字候補について、異特徴文字認識手段がさらに絞込みを行うので、背景部分の文字候補を削減することができる。

40

【0017】

また、第一段文字認識手段と異特徴文字認識手段とで異なる特徴を組み合わせて絞り込みを行うことで、さらに文字候補を削減することができる。

【発明を実施するための最良の形態】

【0018】

以下、本発明の実施形態について図面を用いて説明する。

【0019】

50

(文字列翻訳システム)

まず図1を用いて、文字認識装置を文字列翻訳システムに適用した例を説明する。文字列翻訳システムは、カメラ付きPDAなどの携帯端末11、文字認識装置12、文字列推定装置13、及び翻訳装置14から構成される。

【0020】

この文字列翻訳システムにおいて、ユーザは携帯端末11にて画像を撮影して文字認識装置12に送付する。文字認識装置12は画像から文字候補を抽出して文字列推定装置13に送付する。文字列推定装置13は文字候補から文字列を推定して翻訳装置14に送付する。翻訳装置14は文字列を翻訳して翻訳結果を携帯端末11に送付し、ユーザはこの翻訳結果を閲覧することができる。このように文字列翻訳システムにより、ユーザは撮影した文字列の画像を基に、この文字列の翻訳結果を見ることが可能となる。

10

【0021】

(文字認識装置)

ここで図2を用いて、文字認識装置12の構成を説明する。図2に示すように文字認識装置12は、第一段文字認識手段21、及びN個の異特徴文字認識手段22-1~22-Nから構成される。

【0022】

(第一段文字認識手段)

第一段文字認識手段21は、画像から文字候補を抽出する手段であり、例えば、「画像インデクシングを目的としたテキスト領域不要の景観中文字認識」、電子情報通信学会信学技報PRMU2004-89(2004-10)、p.37-42、に記載の方法により実現できる。

20

【0023】

ここで、第一段文字認識手段21について、例を挙げて説明する。この第一段文字認識手段21では、大きさの異なる文字に対応するために、複数解像度画像を生成し、位置をずらしながら、定められた大きさの画像を切り出し、粗密検索を行う。この結果をインデクスとして利用する。画像検索では、キーワードが入力されると、インデクスから該当する文字のみを抽出して規則性を判定し、規則的と判定された画像を結果として出力する。

【0024】

ここで、粗密探索はパターン学習、及びパターン識別より構成される。以下に、パターン学習、及びパターン識別について説明する。

30

【0025】

[パターン学習] パターン学習は、特徴抽出、カテゴリの階層構造の作成、幾何学変形によるパターン生成、及び辞書生成の4つの段階から構成される。

【0026】

特徴抽出では、文字を正面から撮影した原パターン($w \times w$ とする)を用意し、特徴を抽出する。特徴は、加重方向指数ヒストグラム特徴(WDCH)を利用する。WDCHはOCRに用いられ、2値画像を対象としていたが、グレー画像に容易に拡張可能である。以下にアルゴリズムの概要を示す。ただし、M、Nは正定数である。

- 1: 原パターンから、ソーベルオペレータを用いて微分の値及び方向を求める。
- 2: 微分方向をM方向に量子化する。
- 3: 原パターンを $N \times N$ のグリッドに分割する。
- 4: 各グリッドの各M方向で、微分値の大きさを加算する。
- 5: $N \times N \times M$ の特徴ベクトルと考え、ノルムを正規化する。

40

【0027】

WDCHは、微分値をベースとしているため、明るさ変動の影響を受けにくい。また、グリッド内の微分値の和を用いることにより、フォントによる変形などの形状の微小変動を吸収することができる。

【0028】

カテゴリの階層構造の作成では、特徴ベクトルの類似性から各カテゴリをクラスタリン

50

グし、階層構造を作成する。各ノードは、複数のカテゴリが含まれる。最下層のノードでは、単体のカテゴリのみが含まれる。

【0029】

幾何学変形によるパターン生成では、各カテゴリに対し、視点の変動に伴う文字の変形パターンを生成する。原パターンを回転、垂直スキュー、水平スキュー、縦横比、及び拡大縮小の5パラメータのアフィン変換により幾何学変形する。生成されたパターンの大きさは、原パターンよりも大きくなる場合があるが、現パターンの窓サイズ内の部分パターンのみを用いて特徴を抽出し、この特徴ベクトルを辞書生成に用いる。

【0030】

辞書生成では、以下の手順で各ノードの辞書を作成する。

10

【0031】

第1段階として、特徴圧縮を行う。まず、各階層において、特徴を圧縮する。幾何変形パターンを含む全特徴ベクトルを主成分分析し、上位の固有値を有する固有ベクトルを用いて圧縮する。この圧縮特徴ベクトルを $f(c, r, p)$ と表す。ただし、 c はカテゴリ、 r は圧縮率、 p は変形パラメータである。

【0032】

第2段階として、各ノードでの辞書生成を行う。各ノードのカテゴリ集合を C とすると、 $f(C, r, p)$ のベクトルを主成分分析し、部分空間 $E_d(C, r)$ を求める。ただし、 d は部分空間の次元数であり、寄与率により求めるが、システムにより定められる整数である。

20

【0033】

各階層の圧縮率は、下層に向かうに従い低く設定することで粗密探索を実現する。上層では、精度は低いが、高速な識別を行い、下層では、低速であるが高精度な識別を実行する。

【0034】

[パターン識別] 複数解像度画像全面に位置を動かしながら、大きさ $W \times W$ の小領域を切り出し、パターン識別を行う。パターン識別では、階層構造において複数のルートを辿りながら、粗密探索を実行する。以下にアルゴリズムの概要を示す。

1. 特徴抽出：各解像度画像全面に対し位置を変化させつつ、領域を切り出して特徴を抽出する。すべての切り出し領域の特徴をあらかじめ算出しておく。

30

2. 初期化：木構造のルートノードを出発点とする。

3. 候補ノードの設定：すべての切り出し領域に、候補ノードとして第一階層のノードをセットする。各切り出し領域に対して、4～6を繰り返す。

4. 圧縮：下層の圧縮率を用いて切り出し領域の特徴を圧縮する。これを $I'(r)$ と表す。

5. 投影距離計算：以下の式に従い、候補ノード C の部分空間を用いて投影距離 $L(C)$ を求める。

【0035】

【数1】

$$L(C) = |I'(r)|^2 - \sum_{d=1}^D \{I'(r) \cdot E_d(C, r)\}^2$$

40

【0036】

ただし、 D は部分空間次元である。

6. スクリーニング：上記の距離値から、各候補ノードの順位を計算する。この距離及び順位の閾値から、候補ノードを更新する。

7. ピーク検出：すべての切り出し領域の各候補ノードについて、3次元(縦、横、解像度)の空間的な連結性を算出し、セグメントを求める。各セグメント内の距離値の最小ピークを有する候補ノードのみを残す。その他の候補ノードは削除する。

50

8．局所領域でのスクリーニング：同一解像度のピークの集合各々に対して以下の処理を行う。まず空間をブロック分割し、各ブロック内に含まれるピークを距離値によってソートして、上位から一定個数以内だけ採用する。その後、ブロック分割の位置を水平/垂直に半ブロックずらして同じ処理を行う。

9．同一候補文字のピーク統合：同一候補文字を持つ2つのピークを取り出して中心座標と解像度が互いに近ければ距離値の小さい方に統合する処理を、統合するピークの対が存在しなくなるまで繰り返す。

10．候補ノードの更新：候補ノードの下層に接続されたノードを新しい候補ノードとして登録する。

11．終了判定：最下層に辿り着いていれば残った候補ノードをインデクスとして出力して終了、その他であれば上記4に戻る。インデクスのフォーマットは(カテゴリ名、位置、大きさ、類似度)である。

【0037】

上記7の空間的な連結性は3次元だけでなく2次元(縦、横)等も考えられる。また、上記7～9は、処理量削減のための処理であり、精度及び処理量のトレードオフとなる。すべての階層で行う必要はなく、定められた階層のみでおこなえばよい。

【0038】

[画像検索] 画像検索では、キーワードが文字列として入力されるとパターン識別で得たインデクスの中からパターンが空間的に規則的に配置された個所を探索し、そのような個所が存在するインデクスを有する画像を検索結果として出力する。パターンの空間的な配置の規則として、ここでは、

(1)パターンの大きさがほぼそろっていること

(2)ピッチがほぼ一定であること

(3)ピッチの大きさが個別のパターンの大きさに対して相対的に一定の範囲内に収まっていること

(4)パターンの並ぶ順序が入力された文字列の順序と一致し、かつパターンの並ぶ方向と水平方向又は垂直方向との角度の差が一定の範囲内であることを用いる。

【0039】

この場合の探索アルゴリズムでは、入力文字列を構成する順方向の任意の2個の文字の組み合わせがインデクスの中で存在する個所をすべて探し、これらの個所で仮想的な入力文字列の開始位置の2次元座標、及び文字送りを表す2次元のベクトルの値を算出し、これらのパラメータで構成される投票空間に投票を行う。ただし、投票の際には、その組み合わせが上記(1)、(3)、(4)の規則に反しないか否かをチェックし、反すると判定した場合には投票を行わない。そして最後に、投票空間の中からスコアが閾値以上の個所の有無を探索する。

【0040】

このアルゴリズムでは、投票の際に処理対象となる候補文字は2つのカテゴリだけに限定するため偽の候補文字を多く含むインデクスに対しても高速な処理が可能となり、かつ投票処理の特性により部分的な正解の欠落に対してロバスト性を有することになる。

【0041】

このようにして第一段文字認識手段21は画像から文字候補(候補カテゴリ、位置、大きさ、類似度)を決定する。

【0042】

(異特徴文字認識手段)

異特徴文字認識手段22-1～22-Nは、文字候補(候補カテゴリ、位置、大きさ、類似度)から文字を認識する手段であり、画像から特徴を抽出した後、部分空間法等により文字識別を行う。なお、部分空間法は、以下の文献に詳細が記載されている。“E. Oja. Subspace Methods of Pattern Recognitions. Research Studies Press, 1983.”。

【0043】

10

20

30

40

50

例えば、第一段文字認識手段 2 1 は撮影画像から特徴を抽出し、この特徴と辞書に登録された文字の特徴との類似度に基づいて、図 3 に示すような文字候補の出力をするものとする。これに対し、異特徴文字認識手段 2 2 - A は、第一段文字認識手段 2 1 から出力された文字候補情報の位置及び大きさに応じた画像について、第一段文字認識手段 2 1 が用いた特徴とは異なる特徴を抽出し、文字カテゴリの識別を行う。そして中間データとして図 4 に示すような文字候補のデータを出力する。この文字候補データに対して、さらに異なる異特徴文字認識手段 2 2 - B が、異特徴文字認識手段 2 2 - A とは異なる特徴を抽出して、文字候補の絞込みを行う。最後には、各文字候補の距離閾値（例えば、1000 以内）又は順位閾値（例えば、5 位以内）等によって、図 5 のような出力を得る。すなわち、複数の異なる特徴に着目して文字候補の絞込みを行う異特徴文字認識手段 2 2 を用いて処理を行うことにより、文字候補を削減することができる。

10

【0044】

（第一段文字認識手段と異特徴文字認識手段とが特徴抽出手段を共有）

ここで図 6 を用いて、第一段文字認識手段 2 1 と異特徴文字認識手段 2 2 とが特徴抽出手段 6 1 を共有する場合について説明する。図 6 に示すように、第一段文字認識手段 2 1 と異特徴文字認識手段 2 2 とが特徴抽出手段 6 1 をシェアしており、そのパラメータのみが異なるものとする。

【0045】

例えば、図 7 ~ 図 9 を用いて、特徴抽出手段 6 1 がパラメータとしてブロック数を入力可能な拡張加重方向指数ヒストグラムを用いる場合について説明する。

20

【0046】

図 7 は、拡張加重方向指数ヒストグラムの処理の流れを示した図である。図 7 に示すように、

エッジの方向 / 大きさを求める (S 7 1)。

【0047】

エッジの方向を量子化する (S 7 2)。

【0048】

ブロック分割をする (S 7 3)。

【0049】

各ブロック / 各量子化方向でエッジの大きさの和を計算する (S 7 4)。

30

【0050】

平滑化を行う (S 7 5)。

【0051】

各値を 1 つの特徴ベクトルの要素とみなしてノルムの大きさを正規化する (S 7 6)。

【0052】

上記の処理を経て拡張加重方向指数ヒストグラムの処理を行う。

【0053】

なお、上記 S 7 2、S 7 4、S 7 5 については、加重方向指数ヒストグラムの詳細を記載した以下の文献に詳細が記載される。「T. Wakabayashi, S. Tsuruoka, F. Kimura, Y. Miyake, "Accuracy Improvement through Increased Feature Size in Handwritten Numeral Recognition", IEICE, Vol. J77-D-II, No. 10, pp. 2046-2053, 1994 (in Japanese)」。

40

【0054】

S 7 1 では、図 8 に示すように、エッジの大きさ及び方向を求める。具体的な例としては、縦と横のソーベルオペレータを用いて算出することができる。

【0055】

S 7 3 では、図 9 に示すように、ブロック数が 8 である場合は、画像を 8 × 8 のブロックに分割する。また、ブロック数が 4 である場合は、画像を 4 × 4 のブロックに分割する

50

【 0 0 5 6 】

S 7 6 では、一般的な技術を用いて、各値を 1 つの特徴ベクトルの要素とみなしてノルムの大きさを正規化することができる。

【 0 0 5 7 】

このように、第一段文字認識手段 2 1 で用いる特徴と異特徴文字認識手段 2 2 で用いる特徴とがパラメータのみが異なるものを用いることにより、識別率を低下させずに文字候補を削減することが可能となる。

【 0 0 5 8 】

(第一段文字認識手段と異特徴文字認識手段とが異なる特徴を用いる)

ここで、図 1 0 を用いて、第一段文字認識手段 2 1 が加重方向指数ヒストグラム特徴抽出手段 1 0 1 を有し、異特徴文字認識手段 2 2 が線幅特徴抽出手段 1 0 2 を有する場合について説明する。なお、第一段文字認識手段 2 1 が線幅特徴抽出手段 1 0 2 を有し、異特徴文字認識手段 2 2 が加重方向指数ヒストグラム特徴抽出手段 1 0 1 を有してもよい。

【 0 0 5 9 】

(線幅特徴抽出手段)

図 1 1 を用いて線幅特徴抽出手段 1 0 2 について説明する。図 1 1 は、線幅特徴抽出手段 1 0 2 の構成図である。図 1 1 に示すように線幅特徴抽出手段 1 0 2 は、文字パターン群又は画像(画像パターンと呼ぶ)を入力する画像入力手段 1 1 1、任意の画素において複数の方向の線幅を検出する線幅検出手段 1 1 2、及び各画素の各方向の線幅を特徴値とした特徴ベクトルとして出力する出力手段 1 1 3 から構成される。

【 0 0 6 0 】

図 1 2 は、線幅検出の計算例であって、「ア」の画像に対して 0 度、4 5 度、9 0 度、及び 1 3 5 度の方向の線幅を算出する。例えば、局所的に 2 値化を行い、各方向の連結の長さを測定する。図 1 2 に示すように各画素で 4 つの方向の線幅値が得られ、これを特徴ベクトルとみなす(ただし、数値の記載がない画素の線幅の値は 0 である)。特徴ベクトルの次元は、方向数×画素数である。

【 0 0 6 1 】

ここで図 1 3 を用いて線幅ヒストグラム計算手段 1 3 3 を有する線幅特徴抽出手段 1 0 2 について説明する。図 1 3 に示すように線幅ヒストグラム計算手段 1 3 3 を有する線幅特徴抽出手段 1 0 2 は、文字パターン群又は画像(画像パターンと呼ぶ)を入力する画像入力手段 1 3 1、任意の画素において複数の方向の線幅を検出する線幅検出手段 1 3 2、定められた局所領域内において方向別に線幅のヒストグラムを計算する線幅ヒストグラム計算手段 1 3 3、及び各局所領域の各方向の線幅のヒストグラムを特徴値とした特徴ベクトルとして出力する出力手段 1 3 4 から構成される。

【 0 0 6 2 】

図 1 4 は、線幅ヒストグラムの計算例である。図 1 4 に示すように、図 1 3 の線幅検出手段 1 3 2 にて算出された線幅に対し、画像に対して局所領域を設定し、局所領域内で角度方向別に、線幅の大きさを加算する。ここでは、0 度の中央付近の局所領域では、各画素の線幅が 1、1、0.5 であるため、これらを加算して 2.5 となる。図 1 4 に示すように、各領域で 4 つの方向の線幅ヒストグラムが得られ、これを特徴ベクトルとみなす。特徴ベクトルの次元は、方向数×局所領域数である。なお、局所領域は、例えば画像を格子状に単純に分割することにより求める。

【 0 0 6 3 】

(線幅検出手段)

ここで図 1 5 を用いて線幅検出手段 1 1 2、1 3 2 について説明する。図 1 5 に示すように、線幅検出手段 1 1 2、1 3 2 は、任意の方向および線幅に応じた複数のフィルタを算出するフィルタ生成手段 1 5 1、着目画素を中心とした領域と複数のフィルタとの一致度を算出する一致度算出手段 1 5 2、及び最大の一致度を示すフィルタの線幅の大きさと一致度とを乗算して線幅とする線幅計算手段 1 5 3 から構成される。

【 0 0 6 4 】

図 1 6 は、フィルタ生成手段 1 5 1 により生成されるフィルタの例である。図 1 6 に示すように、4 方向、かつ 3 種類の線幅を検出するフィルタ例である。この場合の線幅は 2、3、4 である。

【 0 0 6 5 】

(一致度計算手段)

一致度計算手段 1 5 2 では、生成した各フィルタと着目画素を中心とした領域との一致度を計算する。ここで、図 1 7 に一致度計算手段 1 5 2 の構成について説明する。図 1 7 に示すように、一致度計算手段 1 5 2 は正規化相関手段 1 7 1 を有している。この正規化相関手段 1 7 1 は、フィルタ生成手段 1 5 1 が生成したフィルタ並びに文字パターン群若しくは画像を入力して、フィルタと文字パターン群又は画像との正規化相関値を算出し、この正規化相関値を一致度として出力する。この正規化相関値の算出においては、例えば、内積値や距離などを用いることができる。

10

【 0 0 6 6 】

このように、第一段文字認識手段 2 1 が加重方向指数ヒストグラム特徴抽出手段 1 0 1 を有し、異特徴文字認識手段 2 2 が線幅特徴抽出手段 1 0 2 を有する、すなわち、第一段文字認識手段 2 1 と異特徴文字認識手段 2 2 とがまったく原理の異なる特徴を組み合わせることにより、若干の識別率の低下を伴うが劇的に文字候補を削減することが可能となる。

【 0 0 6 7 】

(プログラム等)

なお、上記実施形態において、文字認識装置は、例えば、文字認識装置を構成するコンピュータ装置が有する CPU によって実現され、必要とする第一段文字認識処理、異特徴文字認識処理、特徴抽出処理、拡張加重方向指数ヒストグラム特徴抽出処理、線幅特徴抽出処理、正規化相関計算処理などをアプリケーションプログラムとして搭載することができる。

20

【 0 0 6 8 】

また、第一段文字認識処理、異特徴文字認識処理、特徴抽出処理、拡張加重方向指数ヒストグラム特徴抽出処理、線幅特徴抽出処理、正規化相関計算処理などで行った処理結果や計算結果等のデータを内部メモリや外部記憶装置等に書き込み・読み出しができるようにしてもよい。

30

【 0 0 6 9 】

また、本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録媒体を、システム又は装置に供給し、そのシステム又は装置の CPU (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も可能である。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上記実施形態の機能を実現することになり、このプログラムコードを記憶した記憶媒体としては、例えば、CD-ROM、DVD-ROM、CD-R、CD-RW、MO、HDD 等がある。

【 図面の簡単な説明 】

【 0 0 7 0 】

【 図 1 】 文字列翻訳システムの構成図。

【 図 2 】 文字認識装置の構成図。

【 図 3 】 第一段文字認識手段の出力例。

【 図 4 】 異特徴文字認識手段の出力例。

【 図 5 】 異特徴文字認識手段の出力例。

【 図 6 】 第一段文字認識手段と異特徴文字認識手段との特徴抽出手段の共有例。

【 図 7 】 拡張加重方向指数ヒストグラムの処理フロー図。

【 図 8 】 エッジの大きさ及び方向を求める例を示す図。

【 図 9 】 画像をブロックに分割する例を示す図。

【 図 1 0 】 第一段文字認識手段と異特徴文字認識手段との構成図。

40

50

【図 1 1】線幅特徴抽出手段の構成図。

【図 1 2】線幅検出の計算例を示す図。

【図 1 3】線幅特徴抽出手段の構成図。

【図 1 4】線幅ヒストグラムの計算例を示す図。

【図 1 5】線幅検出手段の構成図。

【図 1 6】フィルタ生成手段により生成されるフィルタの例を示す図。

【図 1 7】一致度算出手段の構成図。

【符号の説明】

【 0 0 7 1 】

1 1 ... 携帯端末

1 2 ... 文字認識装置

1 3 ... 文字列推定装置

1 4 ... 翻訳装置

2 1 ... 第一段文字認識手段

2 2 ... 異特徴文字認識手段

6 1 ... 特徴抽出手段

1 0 1 ... 拡張加重方向指数ヒストグラム特徴抽出手段

1 0 2 ... 線幅特徴抽出手段

1 1 1 ... 画像入力手段

1 1 2 ... 線幅検出手段

1 1 3 ... 出力手段

1 3 1 ... 画像入力手段

1 3 2 ... 線幅検出手段

1 3 3 ... 線幅ヒストグラム計算手段

1 3 4 ... 出力手段

1 5 1 ... フィルタ生成手段

1 5 2 ... 一致度計算手段

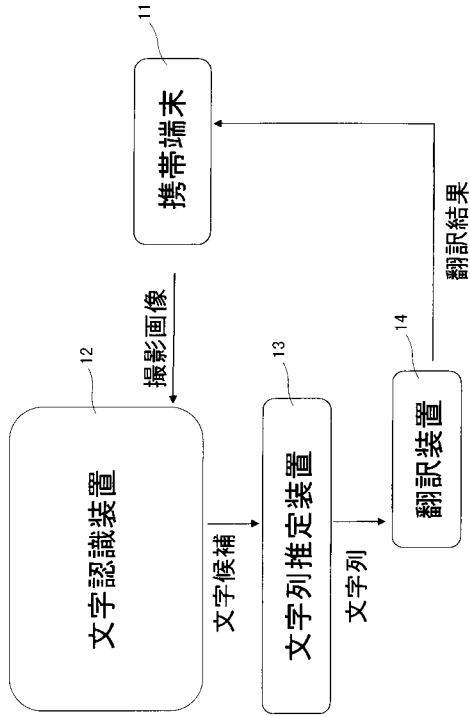
1 5 3 ... 線幅計算手段

1 7 1 ... 正規化相関手段

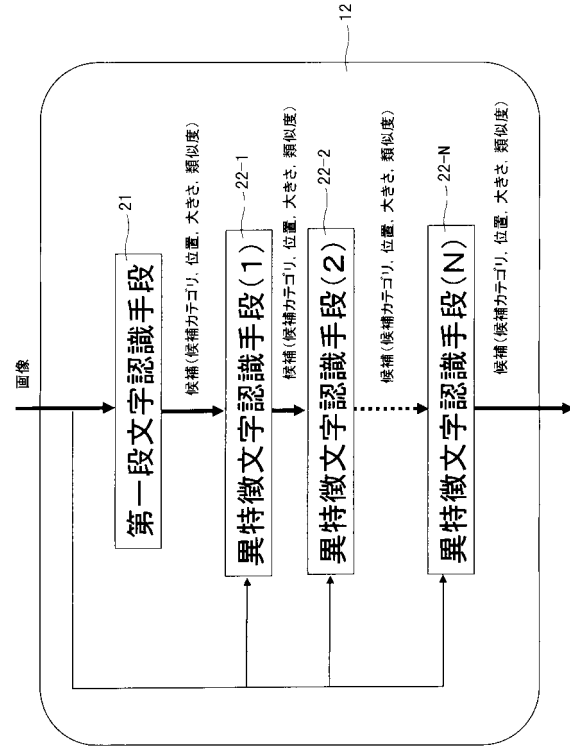
10

20

【図 1】



【図 2】



【図 3】

候補番号	横座標	縦座標	サイズ	カテゴリ名	順位	距離
1	40	40	50	あ	1	500
2	40	40	50	い	2	600
3	70	40	50	え	1	300
4	70	40	50	あ	2	700

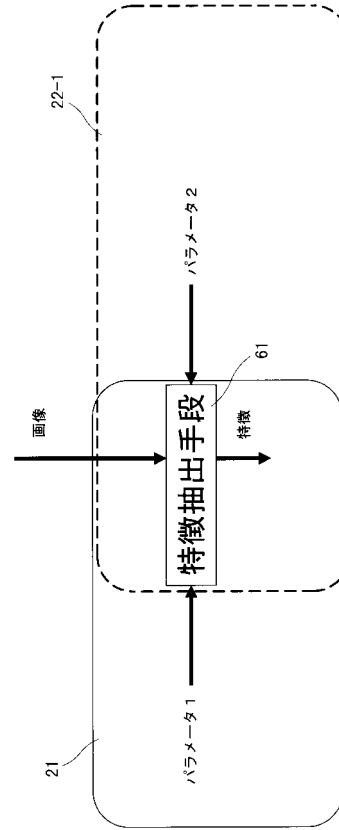
【図 4】

候補番号	横座標	縦座標	サイズ	カテゴリ名	順位	距離
1	40	40	50	あ	6	1500
2	40	40	50	い	1	600
3	70	40	50	え	2	300
4	70	40	50	あ	9	400

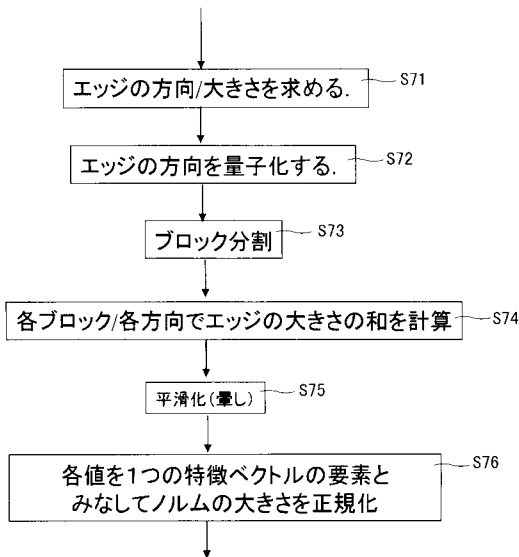
【図5】

候補番号	横座標	縦座標	サイズ	カテゴリ名	順位	距離
2	40	40	50	い	1	600
3	70	40	50	え	2	300

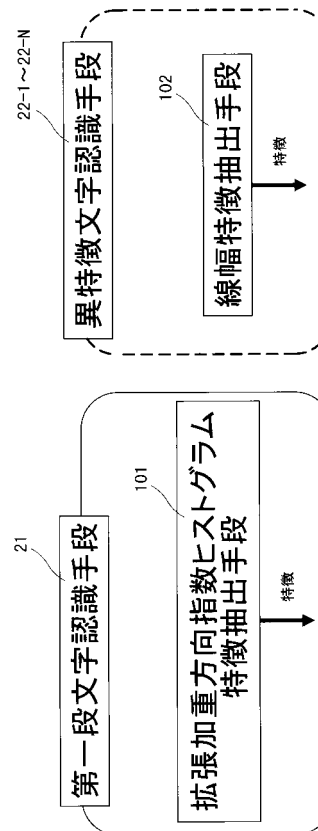
【図6】



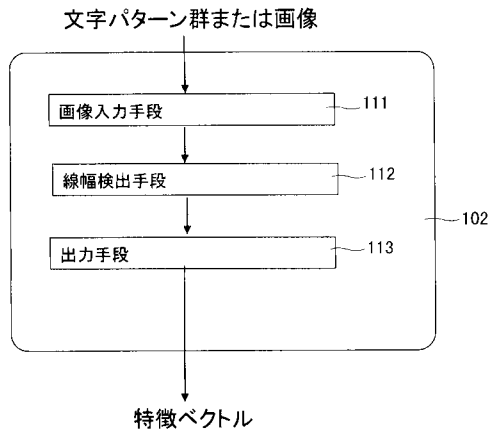
【図7】



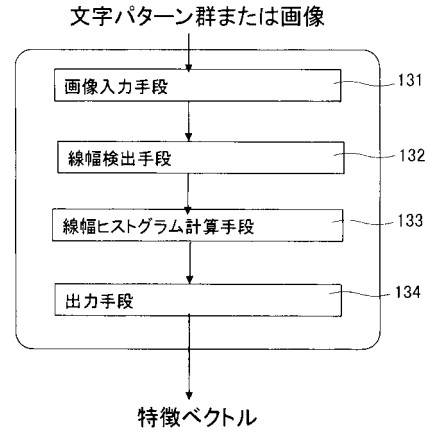
【図10】



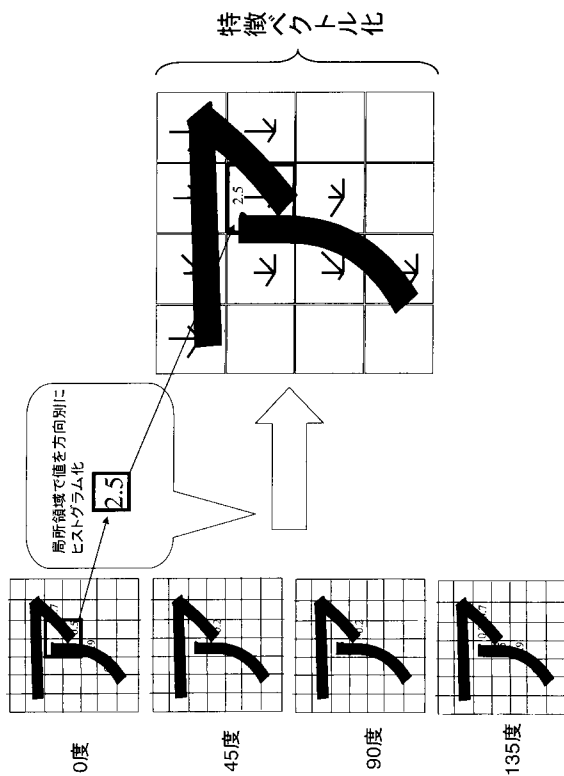
【図 1 1】



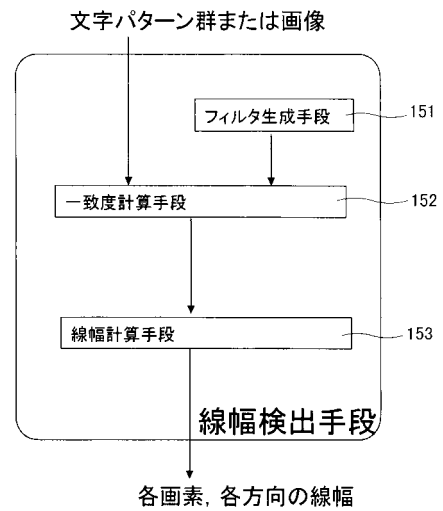
【図 1 3】



【図 1 4】



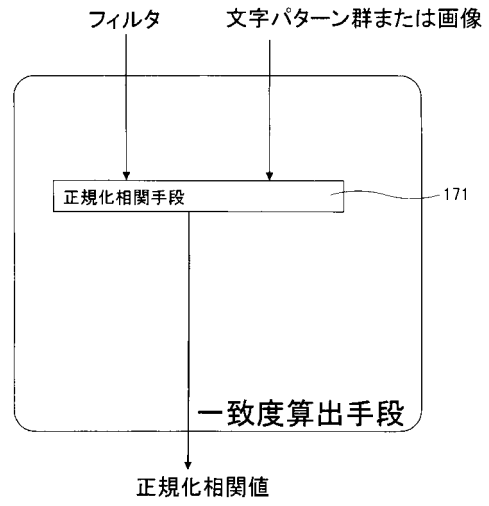
【図 1 5】



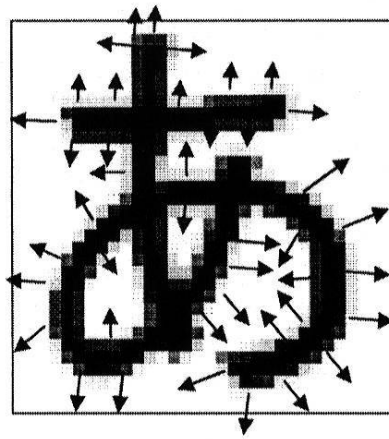
【図 16】

線幅	4				
	3				
	2				
	方向	0度	45度	90度	135度

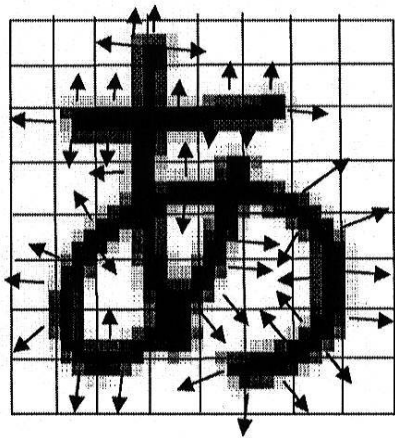
【図 17】



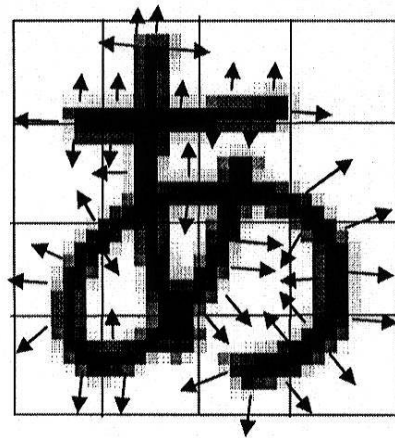
【図 8】



【図9】

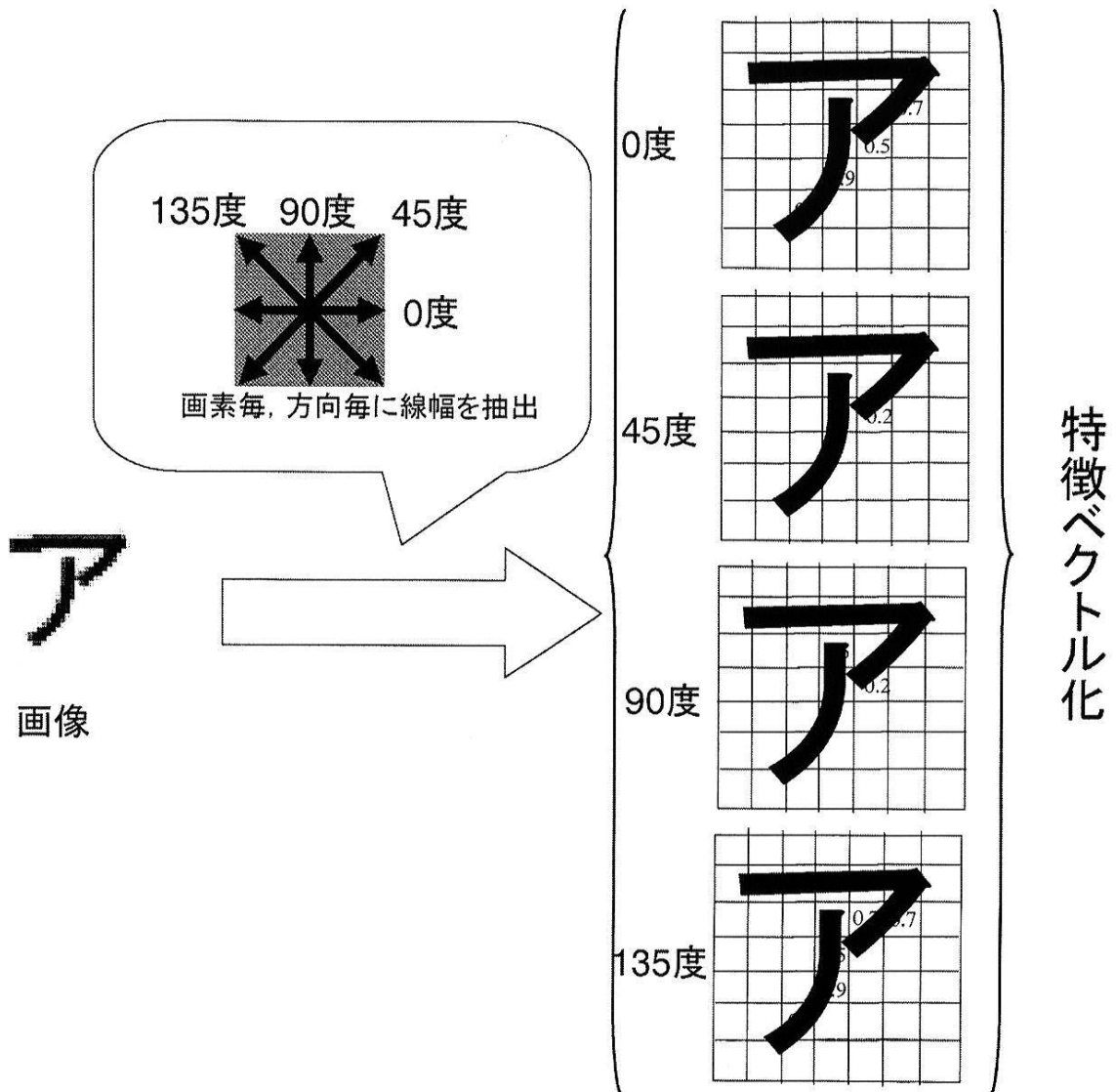


ブロック化数=8



ブロック化数=4

【図12】



フロントページの続き

- (72)発明者 鈴木 章
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 荒川 賢一
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 安藤 慎吾
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 宮里 達也

- (56)参考文献 特開昭48-017637(JP,A)
特開昭62-125485(JP,A)
特開平06-162262(JP,A)
特開平08-180138(JP,A)
特開2000-155803(JP,A)
草地良規,伊藤直己,鈴木章,荒川賢一,画像インデクシングを目的としたテキスト領域抽出不要の景観中文字認識,電子情報通信学会技術研究報告 PRMU,日本,社団法人 電子情報通信学会,2004年10月14日,第104巻,P.37-42

- (58)調査した分野(Int.Cl.,DB名)
G06K 9/03
G06K 9/62