

(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2010-15555

(P2010-15555A)

(43) 公開日 平成22年1月21日(2010.1.21)

(51) Int.Cl.	F I	テーマコード (参考)
G 0 6 K 9/20 (2006.01)	G 0 6 K 9/20 3 2 0 N	5 B 0 2 9
G 0 6 K 9/62 (2006.01)	G 0 6 K 9/62 6 1 0 D	5 B 0 6 4

審査請求 有 請求項の数 27 O L (全 46 頁)

(21) 出願番号	特願2009-139107 (P2009-139107)	(71) 出願人	000005049
(22) 出願日	平成21年6月10日 (2009. 6. 10)		シャープ株式会社
(31) 優先権主張番号	12/164, 985		大阪府大阪市阿倍野区長池町2 2 番 2 2 号
(32) 優先日	平成20年6月30日 (2008. 6. 30)	(74) 代理人	110000338
(33) 優先権主張国	米国 (US)		特許業務法人原謙三国際特許事務所
		(72) 発明者	リチャード ジョン キャンベル
			アメリカ合衆国 ワシントン州 9 8 6 0
			7, カマス, ノースウェスト パシフィッ
			ク リム ブールバード 5 7 5 0 シャ
			ープ ラボラトリーズ オブ アメリカ,
			インコーポレイテッド内

最終頁に続く

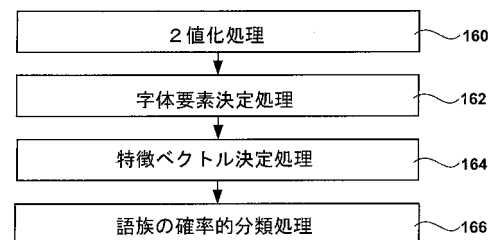
(54) 【発明の名称】 デジタル画像の特徴判定方法およびシステム

(57) 【要約】

【課題】字体のマッチングのためのパターンデータを格納するメモリサイズを削減できる、電子文書のページ方向、言語または語族等を特定する方法およびシステムを実現する。

【解決手段】本発明に係るデジタル画像の特徴判定方法およびシステムは、字体要素に散らばる特徴の分布を各象限毎にカウントし測定することにより、画像データ中のテキスト方向、言語、または語族を特定する方法およびシステムからなる。よって、トレーニング・データ・セットから見本特徴ベクトルおよび共分散行列を生成しておき、確率モデルに適用することで、電子文書の言語または語族を特定することができる。したがって、必要とするデータを格納するメモリサイズを大きく削減することができる。

【選択図】図 1 2



【特許請求の範囲】**【請求項 1】**

デジタル画像における画像特徴判定方法であって、

a) デジタル画像において、第 1 字体要素を特定するステップと、

b) 上記第 1 字体要素の第 1 参照位置を決定するステップと、

c) 上記第 1 字体要素において、第 1 字体特徴に関連する第 1 特徴配置を決定するステップと、

d) 上記第 1 特徴配置の、上記第 1 参照位置に対する第 1 相対位置を決定するステップと、

e) 上記第 1 字体特徴および上記第 1 相対位置に関して、その出現回数を第 1 度数分布として積算するステップと、 10

f) 上記第 1 度数分布に基づいて、分布特性を決定するステップと、

g) 第 1 確率モデルは、上記第 1 字体特徴と第 1 画像特徴とに関連付けられており、上記分布特性と当該第 1 確率モデルとに基づいて、上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップと、

を有することを特徴とする画像特徴判定方法。

【請求項 2】

上記第 1 確率モデルに関連付けられた第 1 の複数のパラメータを受け取るステップをさらに有することを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 3】

20

上記第 1 字体要素を特定するステップは、上記デジタル画像から 2 値化された画像を生成するステップを含むことを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 4】

上記分布特性と第 2 確率モデルとに基づいて、第 2 画像特徴の評価を上記デジタル画像に関連付けるステップをさらに有し、上記第 2 確率モデルは、上記第 1 字体特徴と上記第 2 画像特徴とに関連付けられていることを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 5】

a) 上記第 1 画像特徴は、方向の特徴であり、

b) 上記第 2 画像特徴は、言語の特徴であることを特徴とする請求項 4 に記載の画像特徴判定方法。 30

【請求項 6】

a) 上記第 1 確率モデルと上記第 2 確率モデルとは、共同の確率モデルであり、

b) 上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップと、上記第 2 画像特徴の評価を上記デジタル画像に関連付けるステップとは、上記共同の確率モデルによって併せて行われることを特徴とする請求項 4 に記載の画像特徴判定方法。

【請求項 7】

上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップは、第 1 方向に対応して上記分布特性を並び替えるステップを含んでおり、上記第 1 字体要素は、上記第 1 方向を回転させたものに対応する第 2 方向に関するものであることを特徴とする請求項 1 に記載の画像特徴判定方法。 40

【請求項 8】

上記確率モデルは、正規確率密度関数を含むことを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 9】

上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップは、最大事後確率 (MAP) 分類法、最尤 (ML) 分類法、および最小リスク分類法のいずれか 1 つの分類処理を含むことを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 10】

上記第 1 参照位置は、上記第 1 字体要素の重心であることを特徴とする請求項 1 に記載 50

の画像特徴判定方法。

【請求項 1 1】

上記第 1 参照位置を決定するステップは、

- a) 上記第 1 字体要素の外接矩形を決定するステップと、
- b) 上記外接矩形の中心を決定するステップと、

を含むことを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 1 2】

上記第 1 相対位置を決定するステップは、上記第 1 参照位置に基づいて、上記第 1 特徴配置を象限に分解するステップを含むことを特徴とする請求項 1 に記載の画像特徴判定方法。

10

【請求項 1 3】

上記第 1 画像特徴が、方向の特徴または言語の特徴であることを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 1 4】

上記第 1 画像特徴の評価に対する信頼度を決定するステップをさらに有することを特徴とする請求項 1 に記載の画像特徴判定方法。

【請求項 1 5】

a) 上記第 1 字体要素において、第 2 字体特徴に関連する第 2 特徴配置を決定するステップと、

b) 上記第 2 特徴配置の、上記第 1 参照位置に対する第 2 相対位置を決定するステップと、

c) 上記第 2 字体特徴および上記第 2 相対位置に関して、その出現回数を上記第 1 度数分布に積算するステップと、

をさらに有することを特徴とする請求項 1 に記載の画像特徴判定方法。

20

【請求項 1 6】

a) 上記デジタル画像において、第 2 字体要素を決定するステップと、

b) 上記第 2 字体要素の第 2 参照位置を決定するステップと、

c) 上記第 2 字体要素において、上記第 1 字体特徴に関連する第 2 字体特徴配置を決定するステップと、

d) 上記第 2 特徴配置の、上記第 2 参照位置に対する第 2 相対位置を決定するステップと、

e) 上記第 1 字体特徴および上記第 2 相対位置に関して、その出現回数を第 2 度数分布に積算するステップと、

f) 上記第 2 度数分布と上記第 1 度数分布とを結合させて、上記分布特性を決定するステップと、

をさらに有する、請求項 1 に記載の画像特徴判定方法。

30

【請求項 1 7】

デジタル画像における画像特徴判定方法であって、

a) デジタル画像において第 1 の複数の字体要素を特定するステップと、

b) 上記第 1 の複数の字体要素における各字体要素に対して、字体特徴配置度数分布を決定し、第 1 の複数の度数分布を生成するステップと、

c) 上記第 1 の複数の度数分布に基づいて、分布特性を決定するステップと、

d) 第 1 確率モデルは、複数の字体特徴と、第 1 画像特徴とに関連付けられており、上記分布特性および当該第 1 確率モデルに基づいて、上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップと、

を有することを特徴とする画像特徴判定方法。

40

【請求項 1 8】

上記第 1 画像特徴の評価を上記デジタル画像に関連付けるステップは、最大事後確率 (MAP) 分類法、最尤 (ML) 分類法、および最小リスク分類法のいずれか 1 つの分類処理を含むことを特徴とする請求項 1 7 に記載の画像特徴判定方法。

50

【請求項 19】

上記第1画像特徴は、方向の特徴または言語の特徴であることを特徴とする請求項17に記載の画像特徴判定方法。

【請求項 20】

上記第1の複数の字体要素を特定するステップは、第1の複数の連結した要素からテキスト字体要素を判定するステップを含むことを特徴とする請求項17に記載の画像特徴判定方法。

【請求項 21】

上記第1の複数の字体要素を特定するステップは、ノイズを排除するために、第2の複数の字体要素をフィルタリングするステップを含むことを特徴とする請求項17に記載の画像特徴判定方法。

10

【請求項 22】

デジタル画像における画像特徴判定方法であって、

a) デジタル画像において、第1字体要素を特定するステップと、

b) 第1パターンは、上記第1字体要素内の第1字体画素と当該第1字体画素の近傍にある第1の複数の画素とに基づいており、上記第1字体画素に関連付けられる上記第1パターンを特定するステップと、

c) 上記第1パターンに関する出現回数を、第1度数分布に積算するステップと、

d) 上記第1度数分布に基づいて、分布特性を決定するステップと、

e) 第1確率モデルは、上記第1パターンと第1画像特徴とに関連付けられており、上記分布特性と当該第1確率モデルとに基づいて、上記第1画像特徴の評価を上記デジタル画像に関連付けるステップと、
を有することを特徴とする画像特徴判定方法。

20

【請求項 23】

上記第1画像特徴は、方向の特徴または言語の特徴であることを特徴とする請求項22に記載の画像特徴判定方法。

【請求項 24】

デジタル画像のページ方向を判定する方法であって、

a) デジタル画像の中の字体要素を特定する第1ステップと、

b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する第2ステップと、

30

c) 上記参照位置を原点として上記字体要素を四象限に分割する第3ステップと、

d) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する第4ステップと、

e) 上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する第5ステップと、

f) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算する第6ステップと、

g) 上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する第7ステップと、

40

h) 複数の所定のページ方向のトレーニングデジタル画像の各々に対して、上記第1から第7ステップと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する第8ステップと、

i) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定のページ方向の各々における上記対象特徴ベクトルの尤度を求める第9ステップと、

j) 上記複数の所定のページ方向についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像のページ方向を判定する第10ステップと、

50

を有することを特徴とするデジタル画像のページ方向を判定する方法。

【請求項 25】

デジタル画像の言語または語族を判定する方法であって、

- a) デジタル画像の中の字体要素を特定する第 1 ステップと、
 - b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する第 2 ステップと、
 - c) 上記参照位置を原点として上記字体要素を四象限に分割する第 3 ステップと、
 - d) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する第 4 ステップと、
 - e) 上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する第 5 ステップと、
 - f) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算する第 6 ステップと、
 - g) 上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する第 7 ステップと、
 - h) 複数の所定の言語または語族のトレーニングデジタル画像の各々に対して、上記第 1 から第 7 ステップと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する第 8 ステップと、
 - i) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定の言語または語族の各々における上記対象特徴ベクトルの尤度を求める第 9 ステップと、
 - j) 上記複数の言語または語族についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像の言語または語族を判定する第 10 ステップと、
- を有することを特徴とするデジタル画像の言語または語族を判定する方法。

【請求項 26】

デジタル画像のページ方向を判定するシステムであって、

- a) デジタル画像の中の字体要素を特定する字体要素決定部と、
 - b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する参照位置決定部と、
 - c) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する特徴配置決定部と、
 - d) 上記参照位置を原点として上記字体要素を四象限に分割し、上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する相対位置決定部と、
 - e) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算し、上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する分布特性決定部と、
 - f) 複数の所定のページ方向のトレーニングデジタル画像の各々に対して、上記対象特徴ベクトルを求めるのと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する見本分布特性取得部と、
 - g) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定のページ方向の各々における上記対象特徴ベクトルの尤度を求め、上記複数の所定のページ方向についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像のページ方向を判定する画像特徴判定部と、
- を有することを特徴とするデジタル画像のページ方向を判定するシステム。

【請求項 27】

デジタル画像の言語または語族を判定するシステムであって、

- a) デジタル画像の中の字体要素を特定する字体要素決定部と、
 - b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する参照位置決定部と、
 - c) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する特徴配置決定部と、
 - d) 上記参照位置を原点として上記字体要素を四象限に分割し、上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する相対位置決定部と、
 - e) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算し、上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する分布特性決定部と、
 - f) 複数の所定の言語または語族のトレーニングデジタル画像の各々に対して、上記対象特徴ベクトルを求めるのと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する見本分布特性取得部と、
 - g) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定の言語または語族の各々における上記対象特徴ベクトルの尤度を求め、上記複数の所定の言語または語族についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像の言語または語族を判定する画像特徴判定部と、
- を有することを特徴とするデジタル画像の言語または語族を判定するシステム。

10

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、デジタル画像の特徴を判定する方法およびシステムに関するものであり、特に、文書の方法および語族を判定する方法およびシステムに関するものである。

【背景技術】

【0002】

電子文書におけるページ方向は、原稿がスキャンされた方向、撮像装置のブラテンに対する原稿の方向、またはその他の要因によって、便宜的に決められた名目上のページ方向である「原稿のページ方向」に対応していないことがある。電子文書におけるページ方向と名目上のページ方向とが一致しないと、電子文書进行处理する際に、好ましくない、予期しない、最適でないなどといった不十分な処理結果となりうる。例えば、上記のように方向が異なると、電子文書を印刷したものに最終処理を施す際に、好ましくない結果をもたらす。最終処理としては、例えば、紐で綴じる、ステープルで綴じるといった処理が挙げられる。さらに、許容できるレベルの正確さで、例えば、光学文字認識 (optical character recognition: OCR) といった画像処理操作を行うためには、具体的な方向についてのデータを入力する必要がある。また、名目上のページ方向に対する電子文書のページ方向が不明だと、コンピューターのモニターや携帯用ディスプレイ、またはその他の表示装置などにおいて、適切な表示方向が得られない。

30

【0003】

40

特許文献1にはOCR技術を用い文字認識を行い、読み取った原稿のページ方向を特定する方法が記載されている。これによると、原稿内の一文字一文字を切り出し、切り出した文字をパターン化する。切り出した文字パターンの特徴とデータベース化された文字パターン情報とを比較する。マッチングの方法としては、データベース化された文字パターンに切り出した文字パターンを重ね合わせ、画素毎の白黒を比較し、全てが合致した時のデータベース化された文字パターンをその文字であると判別する。全てが合致する文字パターンがない場合、マッチングする画素が最も多い文字パターンの文字であると判別する。なお、所定のマッチング割合に達しなければ判別不能と判断する。上記の処理を、切り出された文字パターンを90°、180°、270°回転させたものについて同様に行う。文字パターンの回転角毎に判別可能な文字数の比較を行い、判別可能な文字数が最も多

50

い回転角を文字の方向とすることにより、原稿のページ方向を特定することができる。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開平6-189083（平成6年7月8日公開）

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、上記従来の方法では、文字パターン情報を記録するデータベースが大きくなるという問題を生じる。切り出した文字をマッチングするためには文字数分のパターンデータが必要となり、パターンデータ用のメモリサイズが大きくなってしまふ。またこのデータベースは文書に書かれた文字の言語が不明であるときはさらに増大する。

10

【0006】

本発明は、上記の問題点に鑑みてなされたものであり、その目的は、字体のマッチングのためのパターンデータを格納するメモリサイズを削減できる、電子文書のページ方向、言語、または語族等を特定する方法およびシステムを実現することにある。

【課題を解決するための手段】

【0007】

本発明に係るデジタル画像の特徴判定方法は、a) デジタル画像において、第1字体要素を特定するステップと、b) 上記第1字体要素の第1参照位置を決定するステップと、c) 上記第1字体要素において、第1字体特徴に関連する第1特徴配置を決定するステップと、d) 上記第1特徴配置の、上記第1参照位置に対する第1相対位置を決定するステップと、e) 上記第1字体特徴および上記第1相対位置に関して、その出現回数を第1度数分布として積算するステップと、f) 上記第1度数分布に基づいて、分布特性を決定するステップと、g) 第1確率モデルは、上記第1字体特徴と第1画像特徴とに関連付けられており、上記分布特性と当該第1確率モデルとに基づいて、上記第1画像特徴の評価を上記デジタル画像に関連付けるステップと、を有することを特徴としている。

20

【0008】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、画像の特徴を特定することができる。

30

【0009】

上記第1確率モデルに関連付けられた第1の複数のパラメータを受け取るステップをさらに有する構成としてもよい。

【0010】

上記第1字体要素を特定するステップは、上記デジタル画像から2値化された画像を生成するステップを含んでもよい。

【0011】

上記分布特性と第2確率モデルとに基づいて、第2画像特徴の評価を上記デジタル画像に関連付けるステップをさらに有し、上記第2確率モデルは、上記第1字体特徴と上記第2画像特徴とに関連付けられている構成としてもよい。

40

【0012】

上記第1画像特徴は、方向の特徴であり、上記第2画像特徴は、言語の特徴であってもよい。

【0013】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、電子文書の言語およびページ方向を特定することができる。

【0014】

上記第1確率モデルと上記第2確率モデルとは、共同の確率モデルであり、上記第1画

50

像特徴の評価を上記デジタル画像に関連付けるステップと、上記第2画像特徴の評価を上記デジタル画像に関連付けるステップとは、上記共同の確率モデルによって併せて行われる構成としてもよい。

【0015】

上記第1画像特徴の評価を上記デジタル画像に関連付けるステップは、第1方向に対応して上記分布特性を並び替えるステップを含んでおり、上記第1字体要素は、上記第1方向を回転させたものに対応する第2方向に関するものであってもよい。

【0016】

これにより、1つの分布特性から、デジタル画像を回転させた複数の分布特性を得られるので、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズをさらに削減することができる。

10

【0017】

上記確率モデルは、正規確率密度関数を含んでもよい。

【0018】

上記第1画像特徴の評価を上記デジタル画像に関連付けるステップは、最大事後確率(MAP)分類法、最尤(ML)分類法、および最小リスク分類法のいずれか1つの分類処理を含んでもよい。

【0019】

上記第1参照位置は、上記第1字体要素の重心であってもよい。

【0020】

20

上記第1参照位置を決定するステップは、上記第1字体要素の外接矩形を決定するステップと、上記外接矩形の中心を決定するステップと、を含んでもよい。

【0021】

上記第1相対位置を決定するステップは、上記第1参照位置に基づいて、上記第1特徴配置を象限に分解するステップを含んでもよい。

【0022】

上記第1画像特徴が、方向の特徴または言語の特徴であってもよい。

【0023】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、電子文書の言語およびページ方向を特定することができる。

30

【0024】

上記第1画像特徴の評価に対する信頼度を決定するステップをさらに有する構成としてもよい。

【0025】

これにより、特定された画像の特徴の信頼性を高めることができる。

【0026】

上記第1字体要素において、第2字体特徴に関連する第2特徴配置を決定するステップと、上記第2特徴配置の、上記第1参照位置に対する第2相対位置を決定するステップと、上記第2字体特徴および上記第2相対位置に関して、その出現回数を上記第1度数分布に積算するステップと、をさらに有する構成としてもよい。

40

【0027】

上記デジタル画像において、第2字体要素を決定するステップと、上記第2字体要素の第2参照位置を決定するステップと、上記第2字体要素において、上記第1字体特徴に関連する第2字体特徴配置を決定するステップと、上記第2特徴配置の、上記第2参照位置に対する第2相対位置を決定するステップと、上記第1字体特徴および上記第2相対位置に関して、その出現回数を第2度数分布に積算するステップと、上記第2度数分布と上記第1度数分布とを結合させて、上記分布特性を決定するステップと、をさらに有する構成としてもよい。

【0028】

50

複数の字体要素について度数分布を積算し、結合することにより、より正確に画像の特徴を特定することができる。

【0029】

本発明に係るデジタル画像の特徴判定方法は、a) デジタル画像において第1の複数の字体要素を特定するステップと、b) 上記第1の複数の字体要素における各字体要素に対して、字体特徴配置度数分布を決定し、第1の複数の度数分布を生成するステップと、c) 上記第1の複数の度数分布に基づいて、分布特性を決定するステップと、d) 第1確率モデルは、複数の字体特徴と、第1画像特徴とに関連付けられており、上記分布特性および当該第1確率モデルに基づいて、上記第1画像特徴の評価を上記デジタル画像に関連付けるステップと、を有することを特徴としている。

10

【0030】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、画像の特徴を特定することができる。

【0031】

上記第1画像特徴の評価を上記デジタル画像に関連付けるステップは、最大事後確率(MAP)分類法、最尤(ML)分類法、および最小リスク分類法のいずれか1つの分類処理を含んでもよい。

【0032】

上記第1画像特徴は、方向の特徴または言語の特徴であってもよい。

20

【0033】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、電子文書の言語およびページ方向を特定することができる。

【0034】

上記第1の複数の字体要素を特定するステップは、第1の複数の連結した要素からテキスト字体要素を判定するステップを含んでもよい。

【0035】

上記第1の複数の字体要素を特定するステップは、ノイズを排除するために、第2の複数の字体要素をフィルタリングするステップを含んでもよい。

30

【0036】

これにより、実際には字体ではない字体要素を排除することができるので、より正確に画像の特徴を特定することができる。

【0037】

本発明に係るデジタル画像の特徴判定方法は、a) デジタル画像において、第1字体要素を特定するステップと、b) 第1パターンは、上記第1字体要素内の第1字体画素と当該第1字体画素の近傍にある第1の複数の画素とに基づいており、上記第1字体画素に関連付けられる上記第1パターンを特定するステップと、c) 上記第1パターンに関する出現回数を、第1度数分布に積算するステップと、d) 上記第1度数分布に基づいて、分布特性を決定するステップと、e) 第1確率モデルは、上記第1パターンと第1画像特徴とに関連付けられており、上記分布特性と当該第1確率モデルとに基づいて、上記第1画像特徴の評価を上記デジタル画像に関連付けるステップと、を有することを特徴としている。

40

【0038】

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、画像の特徴を特定することができる。

【0039】

上記第1画像特徴は、方向の特徴または言語の特徴であってもよい。

【0040】

50

これにより、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを削減しながら、電子文書の言語およびページ方向を特定することができる。

【0041】

本発明に係るデジタル画像のページ方向を判定する方法は、a) デジタル画像の中の字体要素を特定する第1ステップと、b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する第2ステップと、c) 上記参照位置を原点として上記字体要素を四象限に分割する第3ステップと、d) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する第4ステップと、e) 上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する第5ステップと、f) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算する第6ステップと、g) 上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する第7ステップと、h) 複数の所定のページ方向のトレーニングデジタル画像の各々に対して、上記第1から第7ステップと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する第8ステップと、i) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定のページ方向の各々における上記対象特徴ベクトルの尤度を求める第9ステップと、j) 上記複数の所定のページ方向についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像のページ方向を判定する第10ステップと、を有することを特徴としている。

【0042】

本発明に係るデジタル画像のページ方向を判定するシステムは、a) デジタル画像の中の字体要素を特定する字体要素決定部と、b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する参照位置決定部と、c) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する特徴配置決定部と、d) 上記参照位置を原点として上記字体要素を四象限に分割し、上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する相対位置決定部と、e) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算し、上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する分布特性決定部と、f) 複数の所定のページ方向のトレーニングデジタル画像の各々に対して、上記対象特徴ベクトルを求めるのと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する見本分布特性取得部と、g) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定のページ方向の各々における上記対象特徴ベクトルの尤度を求め、上記複数の所定のページ方向についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像のページ方向を判定する画像特徴判定部と、を有することを特徴としている。

【0043】

これにより、字体のマッチングのために必要となるパターンデータ数は、字体の特徴の種類の数ですみ、また各言語に対するそれぞれの見本特徴ベクトルとして「字体の特徴の種類の数×象限の数」のデータが必要であるのみであり、各言語に対するそれぞれの共分散行列として「(字体の特徴の種類の数×象限の数)²」のデータが必要であるのみである。そのため、必要とするデータを格納するメモリサイズを削減しながら、電子文書のページ方向を特定することができる。また、確率モデルを用いることにより、より正確に電子文書のページ方向を特定することができる。

【0044】

本発明に係るデジタル画像の言語または語族を判定する方法は、a) デジタル画像の中

の字体要素を特定する第 1 ステップと、b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する第 2 ステップと、c) 上記参照位置を原点として上記字体要素を四象限に分割する第 3 ステップと、d) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する第 4 ステップと、

e) 上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する第 5 ステップと、f) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算する第 6 ステップと、g) 上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する第 7 ステップと、h) 複数の所定の言語または語族のトレーニングデジタル画像の各々に対して、上記第 1 から第 7 ステップと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する第 8 ステップと、i) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定の言語または語族の各々における上記対象特徴ベクトルの尤度を求める第 9 ステップと、j) 上記複数の言語または語族についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像の言語または語族を判定する第 10 ステップと、を有することを特徴としている。

10

【0045】

本発明に係るデジタル画像の言語または語族を判定するシステムであって、a) デジタル画像の中の字体要素を特定する字体要素決定部と、b) 上記字体要素の重心または上記字体要素の外接矩形の中心を、参照位置として決定する参照位置決定部と、c) 複数の字体の特徴の各々を示す所定のパターンにマッチする、上記字体要素の特徴点の位置を決定する特徴配置決定部と、d) 上記参照位置を原点として上記字体要素を四象限に分割し、上記特徴点の位置が、上記四象限のいずれに位置するかを示す相対位置を決定する相対位置決定部と、e) 上記デジタル画像の所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算し、上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する分布特性決定部と、f) 複数の所定の言語または語族のトレーニングデジタル画像の各々に対して、上記対象特徴ベクトルを求めるのと同じ手順を用いて得られる見本特徴ベクトルと、当該見本特徴ベクトルに関する共分散行列とを取得する見本分布特性取得部と、g) 上記対象特徴ベクトルと、上記見本特徴ベクトルと、上記共分散行列とを正規分布の確率モデルに適用し、上記複数の所定の言語または語族の各々における上記対象特徴ベクトルの尤度を求め、上記複数の所定の言語または語族についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像の言語または語族を判定する画像特徴判定部と、を有することを特徴としている。

20

30

【0046】

これにより、字体のマッチングのために必要となるパターンデータ数は、字体の特徴の種類の数ですみ、また各言語に対するそれぞれの見本特徴ベクトルとして「字体の特徴の種類の数×象限の数」のデータが必要であるのみであり、各言語に対するそれぞれの共分散行列として「(字体の特徴の種類の数×象限の数)²」のデータが必要であるのみである。そのため、必要とするデータを格納するメモリサイズを削減しながら、電子文書の言語または語族を特定することができる。また、確率モデルを用いることにより、より正確に電子文書の言語または語族を特定することができる。

40

【発明の効果】

【0047】

本発明に係るデジタル画像の特徴判定方法およびシステムによれば、字体のマッチングのためのパターンデータおよびデジタル画像の特徴の判定のための確率モデルに用いるデータを格納するメモリサイズを大きく削減しながら、画像の特徴を特定することができる。また、確率モデルを用いることにより、より正確に画像の特徴を特定することができる。

50

。

【図面の簡単な説明】

【0048】

【図1】(A)は「可読方向」にあるテキスト文書の例を示す図であり、(B)は(A)を時計回りに90°回転させたテキスト文書の例を示す図であり、(C)は(A)を時計回りに180°回転させたテキスト文書の例を示す図であり、(D)は(A)を反時計回りに90°回転させたテキスト文書の例を示す図である。

【図2】文書のページ方向の修正処理を含む、本発明の実施形態を示すフロー図である。

【図3】画像の2値化処理を行った後に行われる非テキストブロック排除処理を含む、本発明の実施形態を示すフロー図である。

10

【図4】要素フィルタリング処理を含む、本発明の実施形態を示すフロー図である。

【図5】(A)は字体要素を囲む外接矩形の中心を字体要素の参照点とする字体要素の例を示す図であり、(B)は字体要素の重心を字体要素の参照点とする字体要素の例を示す図である。

【図6】(A)は「上向き」といわれる字体特徴に対応する画素パターン例を示す図であり、(B)は「右向き」といわれる字体特徴に対応する画素パターン例を示す図であり、(C)は「下向き」といわれる字体特徴に対応する画素パターン例を示す図であり、(D)は「左向き」といわれる字体特徴に対応する画素パターン例を示す図であり、(E)は「右上開口」といわれる字体特徴に対応する画素パターン例を示す図であり、(F)は「右下開口」といわれる字体特徴に対応する画素パターン例を示す図であり、(G)は「左下開口」といわれる字体特徴に対応する画素パターン例を示す図であり、(H)は「左上開口」といわれる字体特徴に対応する画素パターン例を示す図である。

20

【図7】字体要素の例を示す図である。

【図8】字体要素の特徴数の積算処理を含む、本発明の実施形態を示すフロー図である。

【図9】文書のページ方向検出を行うための教師付トレーニング処理を含む、本発明の実施形態を示すフロー図である。

【図10】言語検出を行うための教師付トレーニング処理を含む、本発明の実施形態を示すフロー図である。

【図11】特徴/見本比較処理を含む、本発明の実施形態を示すフロー図である。

【図12】語族の確率的分類処理を含む、本発明の実施形態を示すフロー図である。

30

【図13】要素のフィルタリング処理を含む、本発明の実施形態を示すフロー図である。

【図14】非テキストブロック排除処理を含む、本発明の実施形態を示すフロー図である。

。

【図15】確率的分類法を用いて、語族とページ方向とを同時に分類する分類処理を含む、本発明の実施形態を示すフロー図である。

【図16】字体画素を中心に置くウィンドウに基づく、パターン特徴を含む、本発明の実施形態を示すフロー図である。

【図17】非テキストブロック排除処理を含む、本発明の実施形態を示す図である。

【図18】要素のフィルタリング処理を含む、本発明の実施形態を示す図である。

【図19】本発明に係る画像特徴特定システムを示すブロック図である。

40

【発明を実施するための形態】

【0049】

本発明の実施形態は、字体特徴の分布を測定することによって、デジタル画像におけるページ方向を判定する方法およびシステムに関するものである。また、本発明の他の実施形態は、字体特徴の分布を測定することによって、電子文書におけるテキストの言語または語族を判定する方法およびシステムに関するものである。本発明の別の実施形態は、字体の特徴分布を測定することによって、ページ方向と共に、言語あるいは語族とを判定する方法およびシステムに関するものである。本発明のさらに別の実施形態は、一連のトレーニング・データ・サンプルを基に、ページ全体あるいはテキスト領域全体における字体特徴の見本分布を構築するために、教師付トレーニング方法を用いる方法およびシステム

50

に関するものである。本発明のさらに別の実施形態は、統計的手法を用いる方法およびシステムに関するものであり、字体特徴の分布値のセットに基づいて、入力ページあるいは領域の語族または方向を特定する方法およびシステムに関するものである。

【0050】

上記した本発明の目的、およびその他の目的、特徴、優れた点は、以下に示す本発明の詳細な説明および添付図面によって十分分かるであろう。

【0051】

本発明の実施の形態は、添付の図面を参照することにより、最もよく理解することができる。本実施形態全体を通して、同様の要素には同じ参照符号が付されている。上記した各図は、この発明を実施するための形態の一部として明示的に組み込まれている。

10

【0052】

本発明における構成要素は、図中全般に記述および記載されているように、様々な異なる形態に変更および設計することができる。それゆえ、以下に本発明における方法およびシステムの実施形態を詳細に記述することは、本発明の範囲を限定するものではなく、単に本発明の好ましい実施形態を記述するためのものである。

【0053】

本発明の実施形態における要素はハードウェア、ファームウェア、および/またはソフトウェアにより具現化できる。以下ではこうした形態のうちの一実施形態のみ記載しているが、当業者は本発明の請求項に示される範囲内において、任意の形態で実施しうる。

【0054】

テキストまたは画像等の向きに依存する電子文書（文書画像、テキスト画像、または画像）におけるページ方向は、「原稿のページ方向」（名目上のページ方向と呼ぶ）と一致していないことがある。名目上のページ方向は、テキストまたは画像等の向きとは無関係に仮に決められている。これはスキャン方向やスキャナのプラテン上の原稿方向、その他の要因が原因となっている。このように電子文書においてページ方向と名目上のページ方向が違うと、電子文書进行处理する際に、好ましくない、予期しない、最適でない等の不十分な結果をまねく。例えば、上記のように方向が異なると、電子文書がプリントされたものに対して最終処理を行う際に好ましくない結果を引き起こす。この最終処理としては、たとえば、紐で綴じる、ステープルで綴じる、などが挙げられる。さらに、許容できるレベルの正確さで画像処理操作、たとえば、光学文字認識（optical character recognition : OCR）を行うためには、具体的な方向データを入力する必要がある。また、名目上のページ方向に対する電子文書のページ方向が不明だった場合、コンピューターのモニター、携帯用ディスプレイなどの表示装置において適切な表示方向を得ることができない。

20

30

【0055】

本発明に係るある実施形態は、電子文書における最有力のテキスト方向の自動検知に関するものである。テキスト方向は名目上のページ方向と関連付けることができる。

【0056】

スキャン、コピーなどのデジタル画像処理を行う際、文書は撮像装置のプラテンあるいはドキュメント・フィーダ上に標準方向ではない向きで置かれることがある。その場合、スキャンなどによって取り込まれた文書の電子コピーが、通常の可読方向に対して逆さまあるいは回転したものになってしまう。

40

【0057】

図1(A)～(D)は、「上」方向を可読方向とした場合の可読方向に対する文書方向を示すものである。図1(A)は、可読方向に置かれたテキスト文書2を示している。つまりテキスト文書2は可読方向に対して0度の向きに置かれている。図1(B)は、可読方向に対して90度回転したテキスト文書4を示している。つまり、テキスト文書4は、可読方向に対して時計回りに90度、反時計回りに270度、右回りに90度、左回りに270度回転した向きに置かれている。図1(C)は、可読方向に対して180度回転したテキスト文書6を示している。つまり、テキスト文書6は、可読方向と逆さまの向きに置かれている。図1(D)は可読方向に対して270度回転したテキスト文書8を示して

50

いる。つまり、テキスト文書 8 は、時計回りに 270 度、反時計回りに 90 度、右回りに 270 度、左回りに 90 度回転した向きに置かれている。

【0058】

電子ページが回転している場合、ユーザーが手動で方向を修正する必要がある。たとえば、コピーの場合は、原稿を物理的に回転させる必要があり、スキャンの場合は、文書ビューアまたは画像・文書処理アプリケーションにおけるページの向きをデジタル処理によって逆向きにしたり回転させたりする必要がある。

【0059】

本発明に係る実施形態では、電子文書画像が可読方向であるかどうか、あるいは電子文書画像が上記可読方向に対して 90 度、180 度、270 度回転しているかどうかを判断することができる。このような方向を基本 4 方向と呼ぶ。また、本発明における実施形態では、電子文書画像あるいは画像の一部におけるテキストの言語または語族を判断することができる。

10

【0060】

本発明に係るまた別の実施形態では、字体（グリフ）特徴の分布を測定することによって、電子文書画像のページ方向を判断することができる。本発明におけるさらに別の実施形態では、字体特徴の分布を測定することによって、電子文書画像または画像の一部におけるテキストの言語または語族を判断することができる。

【0061】

本発明に係るまた別の実施形態では、一連のトレーニング・データ・サンプルを基に、あるページ全体もしくはある文字領域全体における字体特徴の見本分布を構築するために、教師付トレーニング方法を用いている。各トレーニング・サンプルには、そのトレーニング・サンプルにおける文字の方向とその文字の言語あるいは語族が、与えられている。トレーニング・データは、見本ページ（もしくは他の領域）の字体特徴の分布のライブラリを形成する。

20

【0062】

本発明に係るさらに別の実施形態では、あるデータのページ（または他の領域）の分布をマッチングするために上記見本ページ（もしくは他の領域）の字体特徴分布ライブラリを用いる。

【0063】

30

本発明におけるさらにまた別の実施形態では、ページ方向と語族を判断するために、語族字体特徴分布の多変量の統計モデルを用いる。こうした実施形態の中には、最尤（Maximum Likelihood: ML）分類法を用いるものや、最大事後確率（Maximum a priori Probability: MAP）分類法を用いるものがある。さらには、最小リスク（Minimum Risk）分類法を用いるものもある。

【0064】

また、本発明に係る別の実施形態では、文字内における二次元パターンの出現頻度と上記パターンの空間的な関係に対応する字体特徴分布（分布特性）を備える。上記のような実施形態では、字体特徴分布が、90 度の回転特性を有しており、その場合、回転したモデルまたは観測データを判定するために、モデルまたは観測データを操作してもよい。こうした実施形態においては、トレーニング処理には、単一方向についてのモデル化または観測が必要となる。

40

【0065】

さらに、本発明に係る他の実施形態は、字体の特徴量を用いて、デジタル文書におけるページ方向と語族を検知するものである。「字体」とは、個々の文字を指すものであり、書き言葉において使用される記号である。本発明に係るまた別の実施形態は、文書方向を判断するために、記載された文字体系において、言語と関係している字体の非対称性を用いる。また、本発明に係るさらに別の実施形態は、文書の言語あるいは語族を判断するために、記載された文字体系において、言語と関係している字体の非対称性を用いる。本発明に係る他の実施形態では、検知結果に対する信頼性を判断する。語族とは、方向に対し

50

て類似した特徴を有する書き言葉をひとまとめにしたものである。

【 0 0 6 6 】

また、本発明に係る別の実施形態では、2 値化されたコーナー・パターンを用いており、文字の非対称性をコード化するために、字体の参照位置に対する象限（四象限）の位置をコード化する。本発明に係るまた別の実施形態では、テキストデータの言語あるいは語族を検知するために、言語分類処理に、字体の2 値化されたコーナー・パターン分布の数をを用いる。また、本発明に係る他の実施形態では、分類処理部をトレーニングし、方向の見本を構築するために、2 値化されたコーナー・パターンを用いる。本発明に係るさらに別の実施形態では、多変量の統計モデルを基にした分類処理部を構築するために、2 値化されたコーナー・パターンを用いる。

10

【 0 0 6 7 】

本発明に係る実施形態について、図2を参照して説明する。まず、電子文書画像に対して2 値化処理が行われる（処理10）。2 値化処理の方法としては、例えば、局所的適応2 値化法やグローバル2 値化法がある。局所的2 値化法としては、Bernsenによる局所法、ChowおよびKanekoによる局所法、Eikvil他による局所法、MardiaおよびHainsworthによる局所法、Niblackの局所法、Taxt他による局所法、YanowitzおよびBrucksteinによる局所法、Parkerによる局所法、WhiteとRohrerの動的閾値アルゴリズム、WhiteとRohrerの積分関数アルゴリズム、などの局所法が挙げられる。また、グローバル2 値化法としては、Abutalebのグローバル法、Kapur他によるグローバル法、KittlerおよびIllingworthのグローバル法、Otsuによるグローバル法、などが挙げられる。本発明の実施形態においては、 32×32 の画素ブロックを用いる、ブロックベースOtsu（大津）2 値化法によって2 値化処理が行われる。画像データの2 値化は、画像データの背景画素とそれ以外（字体要素）とを区別できるようにするために行う。

20

【 0 0 6 8 】

2 値化処理10の後、字体要素判定処理12が行われる。字体（グリフ）とは、記載された言語に用いられた個々の文字あるいはシンボルである。字体要素判定処理12では、2 値化された結果が連結領域アルゴリズムによって処理され、字体要素が特定、つまりラベル化される。本実施形態においては、字体要素判定処理12が、四方連結領域解析、あるいは八方連結領域解析を含んでいてもよい。さらに、字体要素判定処理12が、字体要素特性の算出処理を含んでいてもよい。字体要素特性としては、例えば、字体要素を取り囲む外接矩形、字体要素面積、字体要素重心、字体要素を取り囲む外接矩形面積、字体要素を取り囲む外接矩形の中心、などの特性が挙げられる。ここで字体要素の外接矩形とは字体要素に外接するような矩形である。ひとつの字体要素はひとつの字体文字からなる。画像ノイズといった要因によって、ひとつの字体要素が複数の文字または非字体要素を含んだり、あるいはひとつの字体要素文字が複数の要素に分割されてしまったりすることがある。

30

【 0 0 6 9 】

字体要素と字体要素特性によって、ページ方向の判定処理14が行われる。そして、検知されたページ方向に従って、可読方向あるいは所望の方向を向くよう、電子文書画像の修正処理16が行われる。

40

【 0 0 7 0 】

次に、図3を参照して、本発明の別の実施形態について説明する。本実施形態では、2 値化処理10が行われた後、非テキストブロックを排除（除去）するブロックレベル・フィルタリング処理11が行われ、続いて字体要素判定処理12が行われる。本発明の別の実施形態としては、ブロックレベル・フィルタリング処理11が、ブロック内のコントラスト基準を満たさないブロック、あるいはクラスの分離度が基準を満たさないブロックを排除する処理を含んでいてもよい。

【 0 0 7 1 】

図4を参照して、以下に本発明の別の実施形態について説明する。字体要素判定処理12が行われた後、ノイズとなりうる字体要素を排除するために、要素フィルタリング処理

50

13が行われ、その後、ページ方向判定処理14が行われる。本発明に係る実施形態では、要素フィルタリング処理13は、サイズ基準を満たさない字体要素を排除する処理を含んでいてもよい。サイズ基準としては、例えば、平均的な要素面積に基づいたサイズ下限値、またはサイズ上限値、あるいは字体幅の下限値などのサイズ基準が挙げられる。

【0072】

本発明に係る他の実施形態においては、ブロックレベル・フィルタリング処理とともに要素フィルタリング処理を含んでいてもよい。

【0073】

本発明に係るさらに別の実施形態において、ページ方向判定処理14は、字体要素の参照点（参照位置）に対する、字体要素の2値化された特徴点分布（特徴配置）に基づいて行われる。以下、図5（A）に示す本発明の実施形態について説明する。字体要素30の参照点は、字体要素を囲む外接矩形34の中心32である。該字体要素の外接矩形34の中心32（ c_x 、 c_y ）は、下記の式によって決定される。

10

【0074】

【数1】

$$c_x = \frac{\max\{glyph\}_x + \min\{glyph\}_x}{2}, c_y = \frac{\max\{glyph\}_y + \min\{glyph\}_y}{2}$$

【0075】

20

ここで、式中の{glyph}は、上記字体要素を含む画素の組の座標を示している。

【0076】

図5（B）に示す本発明の実施形態を、以下に説明する。字体要素40の参照点は、字体要素40の重心42である。該字体要素40の重心42（ c_x 、 c_y ）は下記の式によって決定される。

【0077】

【数2】

$$c_x = \frac{\sum\{glyph\}_x}{area}, c_y = \frac{\sum\{glyph\}_y}{area}$$

30

【0078】

ここで式中のareaとは上記字体要素40の面積を指す。本発明に係る実施形態では、上記字体要素40の面積は、字体画素の数によって測定されてもよい。

【0079】

本発明に係る実施形態において、字体要素における特徴点の相対的な位置は、該字体要素の参照点と関連付けて記載されていてもよい。また、本発明に係る別の実施形態としては、字体要素の特徴点の位置は、該特徴点が位置する象限によって記載されていてもよい。本発明の実施例において、字体要素の参照点は、四象限に分解した原点であってもよい。

40

【0080】

図5（A）に示される本発明の実施形態では、参照点は、字体要素の外接矩形34の中心32である。例えば、ラスタスキャン座標系では、画像の左上に原点をとり、画像の水平位置を正のx座標、画像の垂直位置を正のy座標で記述しており、（ f_x 、 f_y ）で示される字体の特徴点の象限は下式によって決定される。

【0081】

【数 3】

$$(fx - c_x) \geq 0 \quad \text{かつ} \quad (fy - c_y) \leq 0 \quad \text{右上象限 3 5}$$

$$(fx - c_x) \geq 0 \quad \text{かつ} \quad (fy - c_y) \geq 0 \quad \text{右下象限 3 6}$$

$$(fx - c_x) \leq 0 \quad \text{かつ} \quad (fy - c_y) \geq 0 \quad \text{左下象限 3 7}$$

$$(fx - c_x) \leq 0 \quad \text{かつ} \quad (fy - c_y) \leq 0 \quad \text{左上象限 3 8}$$

10

【0082】

本発明の実施形態において、特徴点が、象限を分割する線 3 1、3 3 上に位置する場合、該特徴点は 2 つの象限に位置するとしてもよい。あるいは、特徴点が、象限を分割する線 3 1、3 3 上に位置する場合、該特徴点は、分割線に対して定められたデフォルトの象限に位置するとしてもよい。

【0083】

また、本発明の実施形態において、特徴点が、上記字体要素の参照点 3 2 上に位置する場合、該特徴点は 4 つすべての象限に位置するとしてもよい。あるいは、特徴点が、字体要素参照点 3 2 上に位置する場合、該特徴点は、デフォルトの象限に位置するとしてもよい。

20

【0084】

図 5 (B) で示される本発明の実施形態において、参照点は字体要素 4 0 の重心 4 2 である。たとえば、ラスタスキャン座標系では、画像の左上に原点を取り、画像の水平位置を正の x 座標、画像の垂直位置を正の y 座標で記述しており、(f x、f y) で示される字体の特徴点の象限は、下式によって決定される。

【0085】

【数 4】

$$(fx - c_x) \geq 0 \quad \text{かつ} \quad (fy - c_y) \leq 0 \quad \text{右上象限 4 4}$$

$$(fx - c_x) \geq 0 \quad \text{かつ} \quad (fy - c_y) \geq 0 \quad \text{右下象限 4 5}$$

$$(fx - c_x) \leq 0 \quad \text{かつ} \quad (fy - c_y) \geq 0 \quad \text{左下象限 4 6}$$

$$(fx - c_x) \leq 0 \quad \text{かつ} \quad (fy - c_y) \leq 0 \quad \text{左上象限 4 7}$$

30

【0086】

本発明の実施形態において、特徴点が、象限を分割する線 4 1、4 3 上に位置する場合、該特徴点は 2 つの象限に位置するとしてもよい。あるいは、特徴点が、象限を分割する線 4 1、4 3 上に位置する場合、該特徴点は、分割線に対して定められたデフォルトの象限に位置するとしてもよい。

40

【0087】

また、本発明の実施形態において、特徴点が、上記字体要素の参照点 4 2 上に位置する場合、該特徴点は 4 つすべての象限に位置するともみなしてもよい。あるいは、特徴点が、上記字体要素の参照点 4 2 上に位置する場合、該特徴点は、デフォルトの象限に位置するとしてもよい。

【0088】

本発明の実施形態においては、右上象限を象限 1 あるいは第一象限、右下象限を象限 2 あるいは第二象限、左下象限を象限 3 あるいは第三象限、そして、左上象限を象限 4 あるいは第四象限としているが、各象限は違う番号が付与されていてもよい。

50

【 0 0 8 9 】

さらに、本発明に係る実施形態においては、字体の特徴点の位置が、該字体の一つの領域に関連付けられて表されている。あるいは、字体の特徴点の位置が、該字体の外接矩形におけるサブ領域と関連付けられて表されてもよい。こうした実施形態においては、上記領域が互いに重なっていてもよく、重なっていなくてもよい。つまり、本発明の実施形態においては、外接矩形が9つのサブ領域に区切られていてもよく、その場合、上記サブ領域はそれぞれ同じ大きさで、 3×3 となるよう区切られており、各サブ領域は互いに重ならない構成となる。あるいは、上記外接矩形は2つのサブ領域に区切られていてもよく、その場合、各サブ領域は、該外接矩形の半分に相当する。上記のような2つ区切りとしては、例えば、上下に区切られるものと、左右に区切られるものとがある。

10

【 0 0 9 0 】

さらに、本発明の実施形態においては、画素パターンに基づく特徴量が、ページ方向判定処理に用いられてもよい。図6(A)~(H)は、8種類のパターン特徴の例に関連させて、 3×3 の画素パターン例を12個示したものである。図6(A)に示されるパターン60、62は、「上向き」と呼ばれる字体特徴を示すものである。図6(B)に示されるパターン64、66は、「右向き」と呼ばれる字体特徴を示すものである。図6(C)に示されるパターン68、70は、「下向き」と呼ばれる字体特徴を示すものである。図6(D)に示されるパターン72、74は、「左向き」と呼ばれる字体特徴を示すものである。図6(E)に示されるパターン76は、「右上開口」と呼ばれる字体特徴を示すものである。図6(F)に示されるパターン78は、「右下開口」と呼ばれる字体特徴を示すものである。図6(G)に示されるパターン80は、「左下開口」と呼ばれる字体特徴を示すものである。図6(H)に示されるパターン82は、「左上開口」と呼ばれる字体特徴を示すものである。

20

【 0 0 9 1 】

字体要素におけるある画素とそれを取り囲む近傍の8画素とが、特定の字体特徴に対応するいずれかの画素パターンとマッチした場合、当該画素はその特定の字体特徴に対する字体特徴点とみなされる。そして、この画素は、該当する特徴パターンにおける中心画素とみなされる。ここで、字体特徴を有するパターンは、字体要素において非字体画素に相当する画素と、該字体要素において字体画素に相当する画素と、字体要素において非字体画素にも字体画素にも相当する画素、つまり「どちらでもよい」画素とで表される。図6(A)~(H)において示されるパターン例において、上記非字体画素は白の四角で示される画素(例えば、符号55)である。黒の四角で示される画素(例えば、符号56)は、字体画素に相当するものであり、さらに、斜線の網掛けで示される画素(例えば、符号57)は、「どちらでもよい」画素に相当する。

30

【 0 0 9 2 】

ここで、字体の特徴を示すパターンは、字体の角部(鋭角または鈍角の角部)または字体の頂点を示すパターンとしてもよい。この場合、字体の角部または頂点が向いている方向の種類によって複数のパターンにわけてもよい。より具体的には、字体の鋭角の角部または頂点に位置する字体画素を基準画素としたとき、当該基準画素に隣接する背景画素の当該基準画素に対する方向、および、当該基準画素に隣接する字体画素の当該基準画素に対する方向の組み合わせに応じたパターンであってもよい。例えば、図6(A)に示すように、基準画素に隣接する背景画素の当該基準画素に対する方向が上方向及び左方向、または、上方向及び右方向であり、基準画素に隣接する字体画素の当該基準画素に対する方向が下方向である組み合わせを、字体の特徴を示す一つのパターンとすればよい。また字体の鈍角の角部を示すパターンは、字体の鈍角の角部に位置する背景画素を基準画素とし、当該基準画素に隣接する背景画素の当該基準画素に対する方向、及び、当該基準画素に隣接する字体画素の当該基準画素に対する方向の組み合わせに応じたパターンであってもよい。例えば、図6(E)に示すように、基準画素に隣接する背景画素の当該基準画素に対する方向が上方向及び右方向であり、基準画素に隣接する字体要素の当該基準画素に対する方向が下方向及び左方向である組み合わせを、字体の特徴を示す一つのパターンとす

40

50

ればよい。

【 0 0 9 3 】

上記した字体の特徴点について、図 7 に示す字体要素 9 0 に関連付けて説明する。非字体画素は、符号 9 1 で示されるような白の四角で示されており、字体画素は、符号 9 2 で示されるようなクロスハッチの網掛けで示されている。画素 9 3 は、字体の特徴点であり、「上向き」特徴と「左向き」特徴との両方を有する字体特徴点である。画素 9 4 は、「上向き」特徴と「右向き」特徴との両方を有する字体特徴点である。画素 9 5 は、「下向き」特徴と「左向き」特徴との両方を有する字体特徴点である。画素 9 6 は、「下向き」特徴と「右向き」特徴との両方を有する字体特徴点である。画素 9 7 は、「左下開口」特徴を有する字体特徴点である。画素 9 8 は、「右下開口」特徴を有する字体特徴点である。画素 9 9 は、「下向き」特徴と「右向き」特徴との両方を有する字体特徴点である。そして、画素 1 0 0 は、「下向き」特徴と「左向き」特徴との両方を有する字体特徴点である。

10

【 0 0 9 4 】

本発明に係る実施形態において、領域あるいはページごとの統計を形成するために、ある字体要素に対する、画素パターンに基づく特徴の出現度数およびそれらの相対位置の出現度数とが、他の字体要素における画素パターンに基づく特徴の出現度数およびそれらの相対位置の出現度数とに組み合わせられていてもよい。領域あるいはページ毎の統計によって、ページ方向をより確かに判定することができる。

【 0 0 9 5 】

20

本発明に係る実施形態において、画素パターンに基づく特徴の出現度数およびその相対位置の出現度数は、分布マトリクス（度数分布）に積算される。該分布マトリクスでは、マトリクスの各行が象限位置を示し、各列が画素パターンに基づく特徴を示している。つまり、字体要素を g で示し、その字体要素に対する分布マトリクスを dm_g で示すと、字体要素 g 中におけるそれぞれの字体特徴の出現に対して、それに相当する要素が分布マトリクス dm_g 内に積算される。

【 0 0 9 6 】

図 6 (A) ~ (H) で示した 8 個の画素パターン特徴に対する分布マトリクス dm_g は下記のような形式となる。

【 0 0 9 7 】

30

【 数 5 】

$$dm_g = \begin{pmatrix} e_{1,1} & e_{1,2} & e_{1,3} & e_{1,4} & e_{1,5} & e_{1,6} & e_{1,7} & e_{1,8} \\ e_{2,1} & e_{2,2} & e_{2,3} & e_{2,4} & e_{2,5} & e_{2,6} & e_{2,7} & e_{2,8} \\ e_{3,1} & e_{3,2} & e_{3,3} & e_{3,4} & e_{3,5} & e_{3,6} & e_{3,7} & e_{3,8} \\ e_{4,1} & e_{4,2} & e_{4,3} & e_{4,4} & e_{4,5} & e_{4,6} & e_{4,7} & e_{4,8} \end{pmatrix}$$

【 0 0 9 8 】

40

ここで式中の dm_g の要素 $e_{i,j}$ （エレメント： $e_{i,j}(g)$ とも表される）は、画素パターンに基づく特徴 j が字体要素 g の象限 i において出現した回数を表し、各 j は、上記 8 個の画素パターンに基づく特徴例のいずれかひとつに対応する。 k 番目の領域に対して積算された分布マトリクスを DM_k と示すと、当該総和分布マトリクス DM_k は、領域 R_k 内すべての字体要素 g の分布マトリクス dm_g において各要素（エレメント）に関して和をとったものである。これは下記の式のように表される。

【 0 0 9 9 】

【数 6】

$$\mathbf{DM}_k = \sum_{\forall g \subseteq R^k} \mathbf{dm}_g$$

$$e_{i,j}^k = \sum_{\forall g \subseteq R^k} e_{i,j}(g)$$

【0 1 0 0】

ここで式中の $\mathbf{DM}_k$ における要素（エレメント）は $e_{i,j}^k$ で示される。

【0 1 0 1】

10

字体要素の分布マトリクス $\mathbf{dm}_g$ と該字体要素を回転させた字体要素の分布マトリクスとの間には、ある関係がある。図 6 (A) ~ (H) に基づいて説明した上記 8 個の画素パターンの特徴例と図 5 (A) ~ (B) に図示された四象限に分解された例において、字体要素の分布マトリクスにおける要素（エレメント）は、字体要素を右回りに 90 度回転させた場合、下記のように並び替えて表される。

【0 1 0 2】

【数 7】

$$\begin{pmatrix} e_{4,4} & e_{4,1} & e_{4,2} & e_{4,3} & e_{4,8} & e_{4,5} & e_{4,6} & e_{4,7} \\ e_{1,4} & e_{1,1} & e_{1,2} & e_{1,3} & e_{1,8} & e_{1,5} & e_{1,6} & e_{1,7} \\ e_{2,4} & e_{2,1} & e_{2,2} & e_{2,3} & e_{2,8} & e_{2,5} & e_{2,6} & e_{2,7} \\ e_{3,4} & e_{3,1} & e_{3,2} & e_{3,3} & e_{3,8} & e_{3,5} & e_{3,6} & e_{3,7} \end{pmatrix}$$

20

【0 1 0 3】

また、上記字体要素を右回りに 180 度回転させた場合、下記のように並び替えて表される。

【0 1 0 4】

【数 8】

$$\begin{pmatrix} e_{3,3} & e_{3,4} & e_{3,1} & e_{3,2} & e_{3,7} & e_{3,8} & e_{3,5} & e_{3,6} \\ e_{4,3} & e_{4,4} & e_{4,1} & e_{4,2} & e_{4,7} & e_{4,8} & e_{4,5} & e_{4,6} \\ e_{1,3} & e_{1,4} & e_{1,1} & e_{1,2} & e_{1,7} & e_{1,8} & e_{1,5} & e_{1,6} \\ e_{2,3} & e_{2,4} & e_{2,1} & e_{2,2} & e_{2,7} & e_{2,8} & e_{2,5} & e_{2,6} \end{pmatrix}$$

30

【0 1 0 5】

さらに、上記字体要素を右回りに 270 度回転させた場合、下記のように並び替えて表される。

40

【0 1 0 6】

【数 9】

$$\begin{pmatrix} e_{2,2} & e_{2,3} & e_{2,4} & e_{2,1} & e_{2,6} & e_{2,7} & e_{2,8} & e_{2,5} \\ e_{3,2} & e_{3,3} & e_{3,4} & e_{3,1} & e_{3,6} & e_{3,7} & e_{3,8} & e_{3,5} \\ e_{4,2} & e_{4,3} & e_{4,4} & e_{4,1} & e_{4,6} & e_{4,7} & e_{4,8} & e_{4,5} \\ e_{1,2} & e_{1,3} & e_{1,4} & e_{1,1} & e_{1,6} & e_{1,7} & e_{1,8} & e_{1,5} \end{pmatrix}$$

【0 1 0 7】

10

上記式中において、特徴 1 は「上向き」特徴、特徴 2 は「右向き」特徴、特徴 3 は「下向き」特徴、特徴 4 は「左向き」特徴、特徴 5 は「右上開口」特徴、特徴 6 は「右下開口」特徴、特徴 7 は「左下開口」特徴、そして特徴 8 は「左上開口」特徴にそれぞれ相当する。

【0 1 0 8】

90 度回転は、下式で表される並び替え操作として要約できる。

【0 1 0 9】

【数 1 0】

20

$$\mathbf{dm}_g(90^\circ) = \text{permute}(\mathbf{dm}_g) = P_{Row} \mathbf{dm}_g P_{Col} \cdots$$

【0 1 1 0】

ここで、式中の P_{Row} および P_{Col} は以下の通りである。

【0 1 1 1】

【数 1 1】

$$P_{Row} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } P_{Col} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

30

【0 1 1 2】

180 度回転は、 P_{Row} および P_{Col} が上記のような場合、下式で表される並び替え操作として要約できる。

40

【0 1 1 3】

【数 1 2】

$$\mathbf{dm}_g(180^\circ) = \text{permute}(\mathbf{dm}_g(90^\circ)) = P_{Row} \mathbf{dm}_g(90^\circ) P_{Col} = P_{Row} P_{Row} \mathbf{dm}_g P_{Col} P_{Col}$$

【0 1 1 4】

さらに、270 度回転は、 P_{Row} および P_{Col} が上記のような場合、下式で表される並び替え操作として要約できる。

【0 1 1 5】

【数 1 3】

$$\mathbf{dm}_g(270^\circ) = \text{permute}(\mathbf{dm}_g(180^\circ)) = P_{Row} \mathbf{dm}_g(180^\circ) P_{Col} = P_{Row} P_{Row} P_{Row} \mathbf{dm}_g P_{Col} P_{Col} P_{Col}$$

【0 1 1 6】

本発明に係る実施形態において、分布マトリクス間には上記のような関係があるため、トレーニング法によってひとつの文書方向についてトレーニングすることで、4つの基本方向すべてに対して分布マトリクスを生成することができる。

【0 1 1 7】

また、本発明の実施形態において、異なる方向における字体特徴は上記したように並び替える関係にあるため、第一の方向に対応する方向データが格納されれば、必要に応じて当該方向データを並び替えることで、他の方向における識別ベクトルを生成することができる。あるいは、必要な方向すべてに対して、方向データを格納してもよい。

【0 1 1 8】

本発明に係る実施形態においては、画素パターンに基づく特徴量が、言語あるいは語族の識別に用いられてもよい。この場合、対象の領域またはページ全体における字体要素について算出された画素パターンに基づく特徴の数の分布が、言語あるいは語族判定において有効な統計となる。各種語族に対してそれぞれ異なる文字言語（書き言葉）用の字体要素群を用いることで、各語族について、字体要素ごとの特徴数の分布をそれぞれ形成することができる。例えば、ラテン語系の文字（アルファベット）は、アジア系の言語よりも、文字ごとの画素パターンに基づく特徴数が少ない傾向にある。

【0 1 1 9】

図8に示される本発明の実施形態では、字体要素ごとに特徴数の分布が決定される。2値化処理110が行われた後、字体要素判定処理112、そして各字体要素について特徴判定処理114が行われる。続いて、字体要素ごとに特徴数集計処理116が行われ、集計された上記字体要素あたりの特徴数について積算する積算処理118が行われることで、該当する領域あるいはページの分布特徴が生成される。上記で積算された配列は、ヒストグラム、つまり分布ベクトルDVとみなされる。ヒストグラムにおける各ビン（値域）、あるいはベクトルにおける各要素（エレメント）は、画素パターンに基づく特徴数が該ビンに対応する数である字体要素の数を、対象領域またはページについて積算したものである。分布ベクトルは以下のように決定される。

【0 1 2 0】

【数 1 4】

1. $\mathbf{DV} = \mathbf{0}$ /*分布ベクトルをゼロベクトルで初期化*/

2. 対象となる領域 R^k に含まれる全ての字体要素について、

$$\text{num_features}_g = \sum_i \sum_j e_{i,j} \quad /*字体について特徴の数を合計する*/$$

$$DV(\text{num_features}_g) = DV(\text{num_features}_g) + 1$$

/*字体の特徴の数に対応するビンについて回数を積算する*/

【0 1 2 1】

ここで上式における = は代入演算子であり、/* */に挟まれているのはコメントである。

【0 1 2 2】

本発明に係る実施形態では、上記分布特徴を用いて、電子文書において記載されたテキストの言語あるいは語族を判定する。

【0 1 2 3】

図9に示す本発明の実施形態においては、教師付トレーニング（supervised-training

）法に基づいて、テキスト方向の判定を行う。本実施形態では、まず、トレーニング・データ・セットについて、特徴の測定処理 1 2 0 が行われる。上記トレーニング・データ・セットと、あらかじめ分かっているトレーニング・データの方角に関するデータ（方向ラベル）とから作られた測定結果を用いて、分類処理部に対するトレーニング処理 1 2 2 を行う。その後、文書画像に対して特徴の測定処理 1 2 4 が行われ、上記でトレーニングされた分類処理部および上記で測定された文書画像の特徴を用いて、文書方向の判定処理 1 2 6 が行われる。

【 0 1 2 4 】

また、図 1 0 に示す本発明の実施形態においては、教師付トレーニング法に基づいて、テキスト言語の判定を行う。本実施形態では、まず、トレーニング・データ・セットに対して、特徴の測定処理 1 3 0 が行われる。上記トレーニング・データ・セットと、あらかじめ分かっているトレーニング・データの言語に関するデータ（言語ラベルおよび方向ラベル）とから作られた測定結果を用いて、分類処理部に対してトレーニング処理 1 3 2 を行う。その後、文書画像に対して特徴の測定処理 1 3 4 が行われ、上記でトレーニングされた分類処理部および文書画像の特徴を用いて、当該文書画像における言語の判定処理 1 3 6 が行われる。

10

【 0 1 2 5 】

また、本発明に係る実施形態においては、各語族に対する見本分布マトリクスおよび見本分布ベクトルを用いて分類処理部のトレーニングを行ってもよい。これらの見本分布（見本分布マトリクスまたは見本分布ベクトル）は、各語族の典型的な文書（見本文書）に基づく教師付きトレーニングによって得られる。

20

【 0 1 2 6 】

見本分布は、各言語のトレーニング・データ・セットにおける分布を正規化し、その正規化された分布を平均化することによって得られるものであってもよい。本発明に係る実施形態においては、トレーニング・データに対する方向ラベルおよび言語ラベルは、当該トレーニング・データの 2 値化された画像データ部分と関連付けられたメタデータとして含まれる。本発明に係る実施形態においては、上記トレーニング法によって、参照方向（例えば通常上方向とみなされる可読方向）となるように、トレーニング・データを正規化してもよい。参照方向における n 番目のトレーニング・サンプルに対するこの分布マトリクスは、 $DM_{n,D=U}$ と表される。ここで、トレーニング・サンプルは、文書画像のある領域またはページを含んでいる。言語に対する各トレーニング分布マトリクスを正規化することによって、下記の式によって示される、正規化された分布マトリクスを得ることができる。ここで正規化された分布マトリクスを $NDM_{n,D=U}$ で示す。 $DM_{n,D=U}$ と $NDM_{n,D=U}$ の関係は以下の式で示される。

30

【 0 1 2 7 】

【 数 1 5 】

$$NDM_{n,D=U} = \frac{DM_{n,D=U}}{\sum_i \sum_j e_{i,j}^n}$$

40

【 0 1 2 8 】

ここで $e_{i,j}^n$ の上付添字 n は、 n 番目のトレーニング・サンプルについての $e_{i,j}$ であることを示す。また添字 $D=U$ は、ページ方向（direction）が上向き（up）のものであることを示す。

【 0 1 2 9 】

あるいは、下記の式のように、正規化された分布マトリクスは、各要素の和が 1（unit y）ではない値をとってもよい。

【 0 1 3 0 】

【数 1 6】

$$\mathbf{NDM}_{n,D=U} = \frac{C}{\sum_i \sum_j e_{i,j}^n} \mathbf{DM}_{n,D=U}$$

【0 1 3 1】

上記式において、C は所定の定数である。

【0 1 3 2】

見本分布マトリクスは、言語または語族（F で示す）についての複数のトレーニング・サンプル全体に渡って、上記の正規化された分布マトリクスの各要素について下記の式に従って平均化することによって得られる。

【0 1 3 3】

【数 1 7】

$$\mathbf{NDM}_{F,D=U} = \frac{1}{n_T} \sum_{n \in F} \mathbf{NDM}_{n,F,D=U}$$

【0 1 3 4】

ここで添字 n は、参照方向の語族 F についての n 番目のトレーニング・サンプルであることを示す。また n_T は平均をとるトレーニング・サンプルの数である。

【0 1 3 5】

他の方向に対する分布マトリクスは、参照方向に対応する分布マトリクスを並び替えることによって得られる。本発明の実施形態においては、参照方向とは「可読」方向のことであるが、他の実施形態では、参照方向が「可読」方向以外の方向であってもよい。また上記の実施形態においては、各言語に対するそれぞれの見本分布マトリクスとして 3 2 個のデータ（8 種類の字体の特徴 × 4 象限）が必要であるのみであり、必要とするデータベースを小さくすることができる。したがって計算に必要なメモリサイズを大きく削減することが可能となる。

【0 1 3 6】

本発明に係る実施形態においては、各語族の見本分布ベクトル $\mathbf{NDV}_F$ は、下記の式にしたがって、複数のトレーニング・サンプルに基づいて決定される。

【0 1 3 7】

【数 1 8】

$$\mathbf{NDV}_n = \frac{\mathbf{DV}_n}{\sum_i \sum_j e_{i,j}^n}$$

$$\mathbf{NDV}_F = \frac{1}{n_T} \sum_{n \in F} \mathbf{NDV}_n$$

【0 1 3 8】

ここで式中の n は、語族 F についての n 番目のトレーニング・サンプルであることを示し、F は、複数のトレーニング・サンプル全体を平均化して得られた語族 F の分布であることを示す。また $\mathbf{DV}_n$ は n 番目のトレーニング・サンプルについての分布ベクトルを示す。また n_T は平均をとるサンプルの数である。見本分布ベクトル $\mathbf{NDV}_F$ は多数のトレーニング・サンプルに渡って平均をとることによって得られる。分布ベクトルは、ページ方向とは無関係である。

【0 1 3 9】

10

20

30

40

50

本発明に係る実施形態において、上記分類処理部は、最適なマッチング結果を検知するマッチング検知器であり得る。ページ方向検知モジュールは、トレーニングによって算出された上記見本分布（見本分布マトリクスまたは見本分布ベクトル）を用いて、ページ方向、言語あるいは語族、および信頼度を特定する。

【 0 1 4 0 】

本発明に係る実施形態では、信頼度は、評価されたページ方向結果の信頼性が低い場合にそのページ方向結果を排除するために利用される。あるいは、ページ方向のマッチング段階で算出された情報を、ページ方向に関するその他の情報源から得た付加情報とあわせて用いてもよい。ページ方向のマッチング段階で算出される情報とは、例えば、マッチしそうなものすべての信頼度を測定したものなどが挙げられる。

10

【 0 1 4 1 】

図 1 1 を参照にして、本発明のさらに別の実施形態について説明する。文書画像に対して、2 値化処理 1 4 0 が行われ、字体要素候補の判定処理 1 4 2 が行われる。2 値化処理を経て、字体要素候補判定処理が行われた結果は、次に、要素フィルタリング処理 1 4 4 にかかけられ、期待する字体要素構造に適合しない要素は排除される。その結果、字体要素からなる 2 値化字体マップが得られる。次に、字体マップから特徴を抽出する処理 1 4 6 が行われ、それによって、正規化された分布 N D M および N D V が生成される。文書画像の特徴を示すこれらの分布（分布特徴）は、あらかじめ一連のトレーニング手順で算出された複数の言語およびページ方向に関する見本分布と比較される（処理 1 4 8）。ここで N D M および N D V は、字体要素の特徴の出現数（ $\sum_i \sum_j e_{i,j}$ ）でそれぞれ D M および D V を割り正規化したものである。

20

【 0 1 4 2 】

本発明に係る実施形態では、上記のように比較することで、測定された分布 N D M と見本分布 N D M_{F, D} との積和演算 N D M ・ N D M_{F, D} に応じたマッチング値を生成する。この積和演算は、2 つのマトリクスの相関関係を示す、すなわち 2 つのマトリクスの要素同士の積の和をとったものであり、次式で定義される。

【 0 1 4 3 】

【 数 1 9 】

$$\text{NDM} \bullet \text{NDM}_{F,D} = \sum_i \sum_j e_{i,j} \cdot e_{i,j}^{F,D}$$

30

【 0 1 4 4 】

単位レベル（ユニット）で正規化されたマトリクスでは、マッチング値は 0 から 1 までの値をとる。このとき、0 はマッチング度が最も低いことを表し、1 はマッチング度が最も高いことを表している。すなわち、これは、文書画像から得られた分布マトリクスと見本分布から得られた分布マトリクスとの特性が近ければ値が大きくなる、信頼度というべきものである。積和演算によってページ方向と、言語または語族との全ての組み合わせについて上記のマッチング値を算出する。そしてその中から最もふさわしい結果（ページ方向と、言語または語族との最もふさわしい組み合わせ）が次式の選択条件により決定される。

40

【 0 1 4 5 】

【 数 2 0 】

$$(d, f) = \max_{D, F} \left(\sum_i \sum_j e_{i,j} \cdot e_{i,j}^{F,D} \right)$$

【 0 1 4 6 】

ここで式中の D はページ方向、F は語族を示す。このようにして得られた（d, f）によって、測定されたデータに対して最も高い相関値（マッチング度）を示すページ方向 d、および、言語または語族 f の見本分布が特定される。

50

【 0 1 4 7 】

また、本発明に係る別の実施形態においては、マッチング値は、下記の式に示されるように、2つのマトリクス間の差分の絶対値の総和にしたがって生成される。

【 0 1 4 8 】

【数 2 1】

$$\sum_i \sum_j |e_{i,j} - e_{i,j}^{F,D}|$$

【 0 1 4 9 】

本実施形態では、マッチング値 0 はマッチング度が最も高いことを示し、マッチング値 2 はマッチング度が最も低いことを示す。すなわち、これは、文書画像から得られた分布マトリクスと見本分布から得られた分布マトリクスとの特性が近ければ値が小さくなる、信頼度というべきものである。上記演算によってページ方向と、言語または語族との全ての組み合わせについて上記のマッチング値を算出する。そしてその中から最もふさわしい結果が次式の選択条件により決定される。

【 0 1 5 0 】

【数 2 2】

$$(d, f) = \min_{D, F} \left(\sum_i \sum_j |e_{i,j} - e_{i,j}^{F,D}| \right)$$

【 0 1 5 1 】

ここで式中の D はページ方向、F は語族を示す。このようにして得られた (d, f) によって、測定されたデータに対して差分が最も小さい (マッチング度が最も高い) ページ方向 d、および、言語または語族 f の見本分布を特定する。

【 0 1 5 2 】

本発明のいくつかの実施形態では、見本分布をあらかじめメモリ等の記録媒体に保存しておき、そこから取得するようにしてもよい。また、見本分布をインターネット等のネットワークまたは電気信号回線を経由して取得してもよい。

【 0 1 5 3 】

本発明のある実施形態においては、分布ベクトルを利用して言語または語族を特定するためのマッチング値は、積和演算 $NDV \cdot NDV_F$ によって決定してもよい。上記の演算はベクトルの内積である。他の実施形態では、マッチング値は分布ベクトルの各要素同士の差分の絶対値の和 $\sum_i |(NDV - NDV_F)_i|$ によって与えられる。ここで ()_i は、ベクトルの i 番目の要素を示し、| | は、ベクトルの要素 (成分) の絶対値を示し、 $\sum_i$ はベクトルの全ての要素 (成分) の和をとることを示す。最もふさわしい言語を特定するための選択条件はそれぞれ、積和演算については $(f) = \max_F (NDV \cdot NDV_F)$ であり、差分の絶対値の和については $(f) = \min_F (\sum_i |(NDV - NDV_F)_i|)$ である。

【 0 1 5 4 】

本発明の実施形態において、言語分類処理は、ページ方向検出処理の前に行われる。本実施形態では、言語分類処理を先に行うことで、ページ方向検出処理に使用される語族の範囲を限定させることができる。

【 0 1 5 5 】

あるいは、言語判定処理とページ方向検出処理とを並行して行ってもよい。そして分布マトリクスのマッチング結果と、そのマッチング結果から最適なものを選択するための選択条件とは、ページ方向の特定および語族の特定の両方に役立つ情報を返す。この情報は限定された語族のマッチング結果と組み合わせられる。すなわち、分布マトリクスは複数の語族の見本分布マトリクスと比較され、そのマッチング結果の中から、選択条件により最もマッチング結果がよいものを選び出し、語族およびページ方向が特定される。本発明

10

20

30

40

50

に係るある実施形態では、最もマッチング結果がよいものを選択する際に、各マッチング結果に重みをつけて評価する。あるいは、マッチング値を分類処理部に入力してもよい。

【 0 1 5 6 】

本発明の実施形態を、最適マッチングの検出処理部に関するものとして上に記述した。本発明の他の実施形態は、パターン認識法に基づく分類処理部を含んでもよい。

【 0 1 5 7 】

本発明に係るある実施形態においては、語族とページ方向を示すのに用いられる見本分布は、正規化された分布ベクトルの平均に基づいて決定される。本実施形態では、ページ、あるいはページの一部分が、最も分布に近い見本語族およびページ方向に属すよう分類される。別の手段としては、統計的手法を用いてもよく、その場合は、字体特徴の分布値の集合が与えられ、入力ページあるいは領域の語族またはページ方向が特定される。

10

【 0 1 5 8 】

本発明の実施形態は、各語族 L_i についての特徴ベクトル X の尤度 $p(X | L_i)$ を記述する特徴分布の確率モデルを含んでいてもよい。本実施形態においては、尤度と、各語族の相対的な出現度数についての上述の情報とをベイズの公式に当てはめて、事後確率を演算する。特徴ベクトル X が与えられた場合の i 番目の語族の事後確率は $P(L_i | X)$ で表される。この事後確率は、下記の式によって決定される。

【 0 1 5 9 】

【数 2 3】

$$P(L_i | X) = \frac{p(X | L_i)P(L_i)}{p(X)}$$

20

【 0 1 6 0 】

ここで上式において、 $p(X)$ は特徴ベクトル X の尤度、 $P(L_i)$ は i 番目の語族の事前確率である。

【 0 1 6 1 】

本発明の実施形態では、下記の式によって最大事後確率を有する語族を決定することによって、特徴ベクトルを与える語族を決定する。

【 0 1 6 2 】

【数 2 4】

$$MAP(L_i) = \max_i (P(L_i | X)) = \max_i (p(X | L_i)P(L_i))$$

30

【 0 1 6 3 】

上記のように、本実施形態において、分類処理部は、観測されたデータに対して最も高い確率を有する語族として、該データに関連する語族を選択する。この方法を、MAP 分類法 (MAP classifier) と呼ぶ。

【 0 1 6 4 】

別の方法として、 N を考慮している言語の分類数としたとき、すべての語族は一樣であるとみなしてもよい。すなわち、 $P(L_1) = \dots = P(L_N)$ と仮定する。この場合、上記 MAP 分類法は、下記の式で表される最尤分類法に等しくなる。

40

【 0 1 6 5 】

【数 2 5】

$$ML(L_i) = \max_i (p(X | L_i))$$

【 0 1 6 6 】

さらに、本発明に係る実施形態においては、各語族 L_i についての特徴ベクトル X の尤度 $p(X | L_i)$ を算出する処理を含んでいてもよい。ある実施形態では、密度分布とも

50

見なせる尤度関数 $p(X | L_i)$ を、言語候補である各言語の代表的なサンプルのトレーニング・データに基づいてモデル化する。このとき、各語族の上記密度分布を、パラメータを用いて特徴づけてもよい。あるいは、各語族の上記密度分布を、尤度の離散近似によって特徴付けてもよい。本発明の実施形態では、メモリ制限、速度要件、トレーニング・データ・セットのサイズなどの要因は、どの方法を選択するかに影響を及ぼす。

【0167】

本発明の実施形態では、密度分布は、正規確率密度関数 $N(\mu, \Sigma)$ としてモデル化される。ここで、密度平均ベクトル μ_{L_i} および共分散行列 Σ_{L_i} は、各語族の密度分布を特徴付けている。本発明の実施形態では、パラメータ μ_{L_i} および Σ_{L_i} が決定され、特徴ベクトルの尤度は、下記の式によって決定する。

10

【0168】

【数26】

$$p(\mathbf{X}, L_i) = (2\pi)^{-\frac{n}{2}} |\Sigma_{L_i}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{X} - \mu_{L_i})^T \Sigma_{L_i}^{-1} (\mathbf{X} - \mu_{L_i}) \right\}$$

【0169】

上式において、 n は特徴ベクトルにおける特徴の数（ベクトルの要素の数）、 $|\quad|$ は行列式、そして T は行列の転地を示す。

【0170】

20

ここで、特徴ベクトル（対象特徴ベクトル：分布特性） X は、判定対象としているデジタル画像の字体要素の特徴の分布を示すものであり、正規化された分布マトリクス NDM 等と同等の情報をもつベクトルであり得る。ここでの字体要素の特徴とは、字体要素において、例えば図6(A)～図6(H)に示す画像パターンに対応する点と該点が位置する象限との組み合わせを示すものである。特徴ベクトル X は、デジタル画像から、例えばパターンマッチング等を用いて正規化された分布マトリクス NDM と同様の方法で求めることができる。密度平均ベクトル（見本特徴ベクトル） μ_{L_i} は、語族 L_i のトレーニング・データ・セットから特徴ベクトル X と同様の方法によって求められる。また、共分散行列 Σ_{L_i} は、語族 L_i のトレーニング・データ・セットと密度平均ベクトル μ_{L_i} とから、その共分散を求めることで得ることができる。この確率的モデルを、上述した32個の要素（8種類の特徴×4象限）について字体要素の特徴を抽出する例に適用する場合、特徴ベクトルおよび密度平均ベクトルは対応する32個の要素（成分）を有する列ベクトルとして記述でき、共分散行列 Σ_{L_i} は対応する32個の要素の共分散を示す 32×32 の行列として記述できる。

30

【0171】

本発明に係る実施形態において、すべての語族は、該当する可能性が同等であると考えられる場合がある。そのような場合、以下の式に示す尤度の対数を用いることによって、同等の分類処理を行うことができる。

【0172】

【数27】

40

$$\log(p(\mathbf{X}, L_i)) = -\frac{1}{2} (\mathbf{X} - \mu_{L_i})^T \Sigma_{L_i}^{-1} (\mathbf{X} - \mu_{L_i}) + \log(k_{L_i})$$

【0173】

ここで上式において、 k_{L_i} は、語族 L_i に対する規格化定数項を表す。

【0174】

本発明に係る実施形態において、パラメータ μ_{L_i} と Σ_{L_i} とは、語族 L_i に対して、当該語族の見本となるトレーニング・データを用いて決定される。もしくは、このパラメータ μ_{L_i} と Σ_{L_i} とは、語族 L_i に対して、その言語の文字体系からなる記号と、その

50

言語が記載されたテキストにおける当該記号の出現頻度とを分析することによって決定されてもよい。

【0175】

ここで、図12を参照しながら、本発明に係る実施形態について説明する。まず、文書画像に対して2値化処理160を行う。2値化処理法としては、例えば、局所的適応2値化法やグローバル2値化法がある。局所的2値化法としては、Bernsenによる局所法、ChowおよびKanekoによる局所法、Eikvil他による局所法、MardiaおよびHainsworthによる局所法、Niblackの局所法、Taxt他による局所法、YanowitzおよびBrucksteinによる局所法、Parkerによる局所法、WhiteとRohrerの動的閾値アルゴリズム、WhiteとRohrerの積分関数アルゴリズム、などの局所法が挙げられる。また、グローバル2値化法としては、Abutalebのグローバル法、Kapur他によるグローバル法、KittlerおよびIllingworthのグローバル法、Otsuによるグローバル法、などが挙げられる。本発明の実施形態においては、 32×32 の画素ブロックを用いる、ブロックベースのOtsu2値化法によって2値化処理が行われる。画像データの2値化は、画像データの背景画素とそれ以外（字体要素）とを区別できるようにするために行う。

【0176】

上記2値化処理160が行われた後、字体要素判定処理162が行われる。字体とは、記載された言語に用いられた個々の文字あるいは記号である。字体要素判定処理162では、2値化された結果が連結領域アルゴリズムによって処理され、字体要素が特定、つまりラベル化される。本実施形態においては、字体要素判定処理162が、四方連結領域解析、あるいは八方連結領域解析を含んでいてもよい。さらには、字体要素判定処理162が、字体要素特性の算出処理を含んでいてもよい。字体要素特性としては、例えば、字体要素を取り囲む外接矩形、字体要素面積、字体要素重心、字体要素を取り囲む外接矩形面積、字体要素を取り囲む外接矩形の中心、などの特性が挙げられる。ひとつの字体要素はひとつの字体文字からなる。画像ノイズといった要因によって、ひとつの字体要素が複数の文字または非字体要素を含んだり、あるいはひとつの字体要素文字が複数の要素に分割されてしまったりすることがある。

【0177】

次に、特徴ベクトルXの決定処理164が行われ、決定された特徴ベクトルに基づいて、確率的分類処理166が行われ、入力された文書画像に関連する語族が判定される。

【0178】

さらに、図13を参照して、本発明に係る別の実施形態について説明する。まず、2値化処理170、そして字体要素判定処理172が行われる。続いて、要素フィルタリング処理174が行われ、ノイズとなりうる字体要素を排除する。本発明に係る実施形態においては、要素フィルタリング処理174において、サイズ基準を満たさない字体要素を排除する処理を含んでいてもよい。サイズ基準としては、例えば、平均的な要素面積に基づいたサイズ下限値、またはサイズ上限値、あるいは字体幅の下限値などのサイズ基準が挙げられる。その後、要素フィルタリング処理174によって排除されなかった字体要素を用いて、入力データに対応する特徴ベクトルの決定処理176が行われる。ここで決定された特徴ベクトルは、確率的分類処理178に用いられ、当該入力データの言語あるいは語族が判定される。本発明の実施形態においては、上記確率的分類処理178は、MAP分類処理でもよく、あるいは、ML分類処理でもよい。上記実施形態では、上記画像のページ方向は既知である。つまり、画像データのページ方向が見本のページ方向と一致すれば、測定（決定）された特徴ベクトルXを用いて上記分類処理178が直ちに行われる。一方、画像データのページ方向が見本のページ方向と一致しない場合は、見本を回転させたものに対する特徴ベクトルを求め、それから上記分類処理178を行う。

【0179】

次に、図14に示す本発明の実施形態を説明する。まず、2値化処理180を行った後、字体要素判定処理184を行う前に、非テキストブロックを排除するため非テキストブロック排除処理182が行われる。本発明の実施形態においては、ブロックレベル・フィ

ルタリング処理（非テキストブロック排除処理）１８２が、ブロック内のコントラスト基準を満たさないブロック、あるいはクラスの分離度が基準を満たないブロックを排除する処理を含んでいてもよい。上記非テキストブロック排除処理１８２によって排除されなかったブロックは、字体要素判定処理１８４に用いられ、字体要素が特定される。そして、特定された字体要素を用いて、入力データに対応する特徴ベクトルの決定処理１８６が行われる。上記決定された特徴ベクトルを用いて、確率的分類処理１８８が行われ、入力データの言語あるいは語族が判定される。本発明における実施形態においては、上記確率的分類処理１８８は、MAP分類処理でもよく、あるいは、ML分類処理でもよい。上記実施形態では、上記画像のページ方向は既知である。つまり、画像データのページ方向が見本のページ方向と一致すれば、測定（決定）された特徴ベクトル X を用いて上記分類処理１８８が直ちに行われる。一方、画像データのページ方向が見本のページ方向と一致しない場合は、見本を回転させたものに対する特徴ベクトルを求め、それから上記分類処理１８８を行う。

10

【０１８０】

本発明に係る実施形態は、ブロックレベル・フィルタリング処理と字体要素フィルタリング処理の両方含んでいてもよい。

【０１８１】

本発明に係る実施形態では、上記特徴ベクトル X は、字体要素に対する総和分布マトリクス DM_k から得られる（処理１６４、１７６、１８６）。ここで、 DM_k は、下記の式によって与えられる。

20

【０１８２】

【数２８】

$$DM_k = \sum_{\forall g \subseteq R^k} dm_g$$

【０１８３】

ここで上式の g は、対象となっている領域 R_k における字体要素候補である。

【０１８４】

本発明に係る実施形態において、上記特徴ベクトルは、総和分布マトリクスの行または列をベクトル形式に統合することによって得たベクトルからなる。

30

【０１８５】

例えば、下記に示すような総和分布マトリクス DM_k の場合、この総和分布マトリクス DM_k に対応する特徴ベクトル X は以下ようになる。すなわち、分布マトリクス DM_k が、

【０１８６】

【数２９】

$$DM_k = \begin{pmatrix} e_{1,1}^k & e_{1,2}^k & e_{1,3}^k & e_{1,4}^k & e_{1,5}^k & e_{1,6}^k & e_{1,7}^k & e_{1,8}^k \\ e_{2,1}^k & e_{2,2}^k & e_{2,3}^k & e_{2,4}^k & e_{2,5}^k & e_{2,6}^k & e_{2,7}^k & e_{2,8}^k \\ e_{3,1}^k & e_{3,2}^k & e_{3,3}^k & e_{3,4}^k & e_{3,5}^k & e_{3,6}^k & e_{3,7}^k & e_{3,8}^k \\ e_{4,1}^k & e_{4,2}^k & e_{4,3}^k & e_{4,4}^k & e_{4,5}^k & e_{4,6}^k & e_{4,7}^k & e_{4,8}^k \end{pmatrix}$$

40

【０１８７】

である場合、特徴ベクトル X は、

【０１８８】

【数 3 0】

$$X = \begin{bmatrix} e_{1,1}^k & e_{1,2}^k & e_{1,3}^k & e_{1,4}^k & e_{1,5}^k & e_{1,6}^k & e_{1,7}^k & e_{1,8}^k & e_{2,1}^k & e_{2,2}^k & e_{2,3}^k & e_{2,4}^k & e_{2,5}^k & e_{2,6}^k & e_{2,7}^k & e_{2,8}^k & e_{3,1}^k & e_{3,2}^k & e_{3,3}^k & e_{3,4}^k & e_{3,5}^k & e_{3,6}^k & e_{3,7}^k & e_{3,8}^k & e_{4,1}^k & e_{4,2}^k & e_{4,3}^k & e_{4,4}^k & e_{4,5}^k & e_{4,6}^k & e_{4,7}^k & e_{4,8}^k \end{bmatrix}^T$$

【0 1 8 9】

となる。この特徴ベクトルの場合、上記分布マトリクスの転置された各行を辞書式順序に連結させて列ベクトルを形成することによって、上記ベクトルが得られる。

【0 1 9 0】

特徴ベクトルを、別のページ方向における特徴分布を得るために、並び替えてもよい。本発明に係る実施形態においては、画像データの測定された特徴ベクトルが、見本とは違うページ方向を向いている場合に、見本と同じページ方向の特徴ベクトルを得るためにこのような並び替えが必要となる。

10

【0 1 9 1】

上述した並び替え可能な関係性を利用して、上記例に挙げた特徴ベクトルに対して並び替えを行う。

【0 1 9 2】

上記の例の入力データの画像のページ方向を90度回転させると、並び替え操作の式、 $X^{(90^\circ)} = \text{permute}(X^{(0^\circ)})$ で表される。上記並び替え操作は、下記の式によるマッピング(mapping)によって定義される。

【0 1 9 3】

20

【数 3 1】

$$\text{mapping} = (28, 25, 26, 27, 32, 29, 30, 31, 4, 1, 2, 3, 8, 5, 6, 7, 12, 9, 10, 11, 16, 13, 14, 15, 20, 17, 18, 19, 24, 21, 22, 23),$$

【0 1 9 4】

ここで、 $X^{(90^\circ)}(i) = X^{(0^\circ)}(\text{mapping}(i))$ であり、 $X^{(\text{orientation})}(i)$ は、上記回転角度(orientation)の場合における特徴ベクトルXのi番目の要素(成分)を示す。

【0 1 9 5】

30

また、上記の例の入力データの画像のページ方向を180度回転させると、並び替え操作の式、 $X^{(180^\circ)} = \text{permute}(X^{(90^\circ)})$ で表され、このとき、上記並び替え操作は、上記のマッピングによって定義される。

【0 1 9 6】

さらに、上記の例の入力データの画像のページ方向を270度回転させると、並び替え操作の式、 $X^{(270^\circ)} = \text{permute}(X^{(180^\circ)})$ で表され、このとき、上記並び替え操作は、上記のマッピングによって定義される。

【0 1 9 7】

本発明に係る実施形態においては、上記特徴ベクトルXは、正規化された総和分布マトリクスから得たものであってもよい。この場合、正規化定数は、領域 R^k における字体要素候補の数であってもよい。また、マトリクスの形を経ずとも、ベクトルの各要素に字体要素の特徴および象限の出現回数を積算し、領域 R^k における字体要素の数で正規化を行うことにより特徴ベクトルXを求めてもよい。

40

【0 1 9 8】

本発明に係る実施形態では、可読方向(0度)におけるトレーニング・データを用いて、複数の語族に対する多変量の正規分布パラメータを決定する。このパラメータは、 L_i をi番目の語族、0を可読方向としたとき、 $N(\mu_{L_i, 0}, \Sigma_{L_i, 0})$ で表される。

【0 1 9 9】

図15に示す本発明に係る実施形態は、ページ方向および語族の尤度を同時に算出する処理を含む。まず、電子文書画像に対して2値化処理190を行う。本発明に係る実施形

50

態においては、上記２値化処理１９０を行った後、字体要素判定処理１９４を行う前に、非テキストブロックを排除するために非テキストブロック排除処理１９２が行われる。あるいは、非テキストブロック排除処理１９２は行わなくてもよい。その場合、２値化処理１９０および字体要素判定処理１９４を行った後に、要素フィルタリング処理を行い、ノイズとなる字体要素を排除する。さらに別の方法としては、非テキストブロック排除処理および要素フィルタリング処理の両方を行ってもよい。

【０２００】

上記字体要素判定処理１９４の後、特徴分布マトリクス（総和分布マトリクス）の判定処理１９６が行われ、上記判定された特徴分布マトリクスに基づいて、特徴ベクトルの生成処理２００が行われる。この特徴ベクトルに基づいて、並び替えられた特徴ベクトルの生成処理２０２が行われる。そして、確率的分類処理２０４に基づいて、語族とページ方向とを併せて判定する。本発明に係る実施形態においては、上記確率的分類処理２０４は、MAP分類処理でもよく、ML分類処理でもよい。

10

【０２０１】

本発明に係る実施形態においては、MAP分類処理を行ってもよい。この場合、語族とページ方向は下記の式によって選択される。

【０２０２】

【数３２】

$$MAP(L_i, k) = \max_{i,k} \left(P(L_i | \mathbf{X}^{(k)}) \right) = \max_{i,k} \left(p(\mathbf{X}^{(k)} | L_i) P(L_i) \right)$$

20

【０２０３】

ここで、 L_i は語族を示し、 k はページ方向を示す。本発明に係る実施形態においては、確率密度関数が正規確率密度関数としてモデル化されており、語族とページ方向は、下記の式によって選択される。

【０２０４】

【数３３】

$$MAP(L_i, k) = \max_{i,k} \left(\left| \Sigma_{L_i} \right|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{X}^{(k)} - \mu_{L_i})^T \Sigma_{L_i}^{-1} (\mathbf{X}^{(k)} - \mu_{L_i}) \right\} P(L_i) \right)$$

30

【０２０５】

あるいは、対数尤度を用いて、上記に相当するMAP分類処理を下記の式のように行ってもよい。

【０２０６】

【数３４】

$$MAP(L_i, k) = \max_{i,k} \left(\left\{ -\frac{1}{2} (\mathbf{X}^{(k)} - \mu_{L_i})^T \Sigma_{L_i}^{-1} (\mathbf{X}^{(k)} - \mu_{L_i}) \right\} - \frac{1}{2} \log \left(\left| \Sigma_{L_i} \right| \right) + \log \left(P(L_i) \right) \right)$$

40

【０２０７】

本発明に係るある実施形態においては、ML分類処理を行ってもよい。この場合、語族とページ方向は、下記の式によって選択される。

【０２０８】

【数３５】

$$ML(L_i, k) = \max_{i,k} \left(p(\mathbf{X}^{(k)} | L_i) \right)$$

【０２０９】

ここで、 L_i は語族を示し、 k はページ方向を示す。本発明に係る実施形態においては、

50

確率密度関数が正規確率密度関数としてモデル化されており、語族とページ方向は、下記の式によって選択される。

【 0 2 1 0 】

【 数 3 6 】

$$ML(L_i, k) = \max_{i,k} \left(\left| \Sigma_{L_i} \right|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left(\mathbf{X}^{(k)} - \mu_{L_i} \right)^T \Sigma_{L_i}^{-1} \left(\mathbf{X}^{(k)} - \mu_{L_i} \right) \right\} \right)$$

【 0 2 1 1 】

あるいは、対数尤度を用いて、上記に相当する M L 分類処理を下記の式のように行ってもよい。

10

【 0 2 1 2 】

【 数 3 7 】

$$ML(L_i, k) = \max_{i,k} \left(\left\{ -\frac{1}{2} \left(\mathbf{X}^{(k)} - \mu_{L_i} \right)^T \Sigma_{L_i}^{-1} \left(\mathbf{X}^{(k)} - \mu_{L_i} \right) \right\} - \frac{1}{2} \log \left(\left| \Sigma_{L_i} \right| \right) \right)$$

【 0 2 1 3 】

本発明に係る実施形態は、判定された語族およびページ方向に対する信頼度を判定する処理を含んでいてもよい。

20

【 0 2 1 4 】

M A P 分類処理を含む実施形態において、信頼度としては、例えば、最大尤度値、最大対数尤度値、尤度の最小値と最大値との差、対数尤度値の最小値と最大値との差、最大尤度値と 2 番目に大きい尤度値との差、最大対数尤度値と 2 番目に大きい対数尤度値との差、および、現時点での最善の判定とその代わりとなる判定との間の広がりなどの他の測定量が挙げられる。

【 0 2 1 5 】

M L 分類処理を含む実施形態において、信頼度としては、例えば、最大尤度値、最大対数尤度値、尤度の最小値と最大値との差、対数尤度値の最小値と最大値との差、最大尤度値と 2 番目に大きい尤度値との差、最大対数尤度値と 2 番目に大きい対数尤度値との差、および、現時点での最善の判定とその代わりとなる判定との間の広がりなどの他の測定量が挙げられる。

30

【 0 2 1 6 】

本発明に係る実施形態は、確率的分類処理部を含んでいてもよく、この確率的分類処理部は、不適当な判定をするリスクまたは不適当な判定をした際の損失に関連する因子を、判定基準に組み込んでいる。こうした実施形態では、与えられたデータの或る状態における尤度だけでなく、最もありうる状態からかけ離れた実際の状態を与える判定を受け入れるリスクをも因子として含むよう、上記判定基準の特性を最適化してもよい。

【 0 2 1 7 】

40

条件付き損失 (conditional loss) を導入することによって、判定を行うことの損失が判定基準に加わる。本実施形態においては、これによって、どの判定の尤度も低い、または、その判定を下すリスクが高い場合、上記確率分類処理部は、より安全な、または、より損失の少ない判定を行うようになる。特徴ベクトル X が与えられた場合、判定結果 O_i であると判定する条件付き損失 (すなわちリスク) $Risk(O_i | X)$ は、下記のように定式化できる。

【 0 2 1 8 】

【数 3 8】

$$Risk(O_i | \mathbf{X}) = \sum_{O_j = \{R0, J0, R90, J90, \dots\}} \lambda(O_i | O_j) P(O_j | \mathbf{X})$$

【0 2 1 9】

ここで、語族とページ方向とをひとつの指標で表記するよう書き換えている。指標の先頭についているアルファベットが同じであれば、同じ語族であることを示す（例えば、指標 R 0、R 9 0、R 1 8 0、R 2 7 0 がある語族を示す場合、J が先頭についている指標 J 0、J 9 0、J 1 8 0、J 2 7 0 は、別の語族を示す）。指標 O_j は、語族およびページ方向の値が結合した集合の要素である。上記の条件付き損失の定式は、以下のように解釈できる。つまり、測定結果（特徴ベクトル） $\mathbf{X}$ が与えられた場合、ページ方向および言語 O_i を選択することの条件付き損失または条件付きリスクは、実際の系の状態が O_j であるときにページ方向および言語 O_i を選択することの損失 $\lambda(O_i | O_j)$ に、上記測定結果 $\mathbf{X}$ が与えられたときに実際の状態が O_j である確率を乗じたものを、状態 O_j について和をとったものに等しくなる。

10

【0 2 2 0】

各ページ方向および語族が一様にありえそうな場合、条件付き損失の定式は、さらに簡略化される。このとき、上記条件付き損失は、以下の通りである。

【0 2 2 1】

20

【数 3 9】

$$Risk(O_i | \mathbf{X}) \propto \sum_{O_j = \{R0, J0, R90, J90, \dots\}} \lambda(O_i | O_j) p(\mathbf{X} | O_j)$$

【0 2 2 2】

本発明に係る実施形態において、上記判定基準は、全体の損失を最小化する状態を選択するように、以下のように定式化される。

【0 2 2 3】

【数 4 0】

$$MinLoss(O_i) = \min_{i = \{R0, J0, R90, J90, \dots\}} (Risk(O_i | \mathbf{X}))$$

30

【0 2 2 4】

本発明に係る実施形態において、2つの語族が同等に有望である場合、条件付き損失算出のマトリクス式は、以下の通りである。

【0 2 2 5】

【数 4 1】

$$\begin{pmatrix} Risk(O_{R0} | \mathbf{X}) \\ Risk(O_{J0} | \mathbf{X}) \\ Risk(O_{R90} | \mathbf{X}) \\ Risk(O_{J90} | \mathbf{X}) \\ Risk(O_{R180} | \mathbf{X}) \\ Risk(O_{J180} | \mathbf{X}) \\ Risk(O_{R270} | \mathbf{X}) \\ Risk(O_{J270} | \mathbf{X}) \end{pmatrix} = \begin{pmatrix} \lambda(O_{R0} | O_{R0}) & \lambda(O_{R0} | O_{J0}) & \dots & \lambda(O_{R0} | O_{J270}) \\ \lambda(O_{J0} | O_{R0}) & \lambda(O_{J0} | O_{J0}) & & \\ \vdots & & \ddots & \vdots \\ \lambda(O_{J270} | O_{R0}) & & & \lambda(O_{J270} | O_{J270}) \end{pmatrix} \begin{pmatrix} p(\mathbf{X} | O_{R0}) \\ p(\mathbf{X} | O_{J0}) \\ p(\mathbf{X} | O_{R90}) \\ p(\mathbf{X} | O_{J90}) \\ p(\mathbf{X} | O_{R180}) \\ p(\mathbf{X} | O_{J180}) \\ p(\mathbf{X} | O_{R270}) \\ p(\mathbf{X} | O_{J270}) \end{pmatrix}$$

40

【0 2 2 6】

50

また、本発明に係る実施形態において、特徴ベクトルは回転特性（例えば、上述したような特性）を有しているので、条件付き損失は、以下のように記述することもできる。

【 0 2 2 7 】

【 数 4 2 】

$$\begin{pmatrix} Risk(O_{R0} | \mathbf{X}) \\ Risk(O_{J0} | \mathbf{X}) \\ Risk(O_{R90} | \mathbf{X}) \\ Risk(O_{J90} | \mathbf{X}) \\ Risk(O_{R180} | \mathbf{X}) \\ Risk(O_{J180} | \mathbf{X}) \\ Risk(O_{R270} | \mathbf{X}) \\ Risk(O_{J270} | \mathbf{X}) \end{pmatrix} = \begin{pmatrix} \lambda(O_{R0} | O_{R0}) & \lambda(O_{R0} | O_{J0}) & \cdots & \lambda(O_{R0} | O_{J270}) \\ \lambda(O_{J0} | O_{R0}) & \lambda(O_{J0} | O_{J0}) & & \\ \vdots & & \ddots & \vdots \\ \lambda(O_{J270} | O_{R0}) & & & \lambda(O_{J270} | O_{J270}) \end{pmatrix} \begin{pmatrix} p(\mathbf{X}^{(0)} | O_R) \\ p(\mathbf{X}^{(0)} | O_J) \\ p(\mathbf{X}^{(90)} | O_R) \\ p(\mathbf{X}^{(90)} | O_J) \\ p(\mathbf{X}^{(180)} | O_R) \\ p(\mathbf{X}^{(180)} | O_J) \\ p(\mathbf{X}^{(270)} | O_R) \\ p(\mathbf{X}^{(270)} | O_J) \end{pmatrix} \quad \begin{matrix} 10 \\ 20 \end{matrix}$$

【 0 2 2 8 】

上記の場合、各方向および各語族の組み合わせに対する別々の見本を使用する代わりに、上記測定値（特徴ベクトルの要素）が並び替えられる。

【 0 2 2 9 】

上述した実施形態は、ページ方向および語族を同時に（並列に）判定する処理を含む分類処理部を含むものである。以下に説明する実施形態では、ページ方向および語族の判定はそれぞれ別々に行われてもよい。

【 0 2 3 0 】

別の実施形態においては、語族は、N個の可能性のある語族 1 1 ~ 1 N から、リスク（条件付き損失）を最小にすることによってページ方向とは独立に判定される。この場合、上記リスクは、以下のように決定される。

【 0 2 3 1 】

【 数 4 3 】

$$\begin{pmatrix} Risk(L_{I1} | \mathbf{X}) \\ \vdots \\ Risk(L_{IN} | \mathbf{X}) \end{pmatrix} = \begin{pmatrix} \lambda(L_{I1} | O_{I10}) & \cdots & \lambda(L_{I1} | O_{IN0}) & \cdots & \lambda(L_{I1} | O_{I1270}) & \cdots & \lambda(L_{I1} | O_{IN270}) \\ \lambda(L_{I2} | O_{I10}) & \cdots & \lambda(L_{I2} | O_{IN0}) & \cdots & \lambda(L_{I2} | O_{I1270}) & \cdots & \lambda(L_{I2} | O_{IN270}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ \lambda(L_{IN} | O_{I10}) & \cdots & \lambda(L_{IN} | O_{IN0}) & \cdots & \lambda(L_{IN} | O_{I1270}) & \cdots & \lambda(L_{IN} | O_{IN270}) \end{pmatrix} \mathbf{p}$$

【 0 2 3 2 】

ここで、p は尤度値のベクトルであり、以下の式によって決定される。

【 0 2 3 3 】

【数 4 4】

$$\mathbf{p} = \begin{pmatrix} p(\mathbf{X} | O_{I0}) \\ \vdots \\ p(\mathbf{X} | O_{IN0}) \\ p(\mathbf{X} | O_{II90}) \\ \vdots \\ p(\mathbf{X} | O_{IN90}) \\ p(\mathbf{X} | O_{II180}) \\ \vdots \\ p(\mathbf{X} | O_{IN180}) \\ p(\mathbf{X} | O_{II270}) \\ \vdots \\ p(\mathbf{X} | O_{IN270}) \end{pmatrix} = \begin{pmatrix} p(\mathbf{X}_0 | L_{II}) \\ \vdots \\ p(\mathbf{X}_0 | L_{IN}) \\ p(\mathbf{X}_{90} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{90} | L_{IN}) \\ p(\mathbf{X}_{180} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{180} | L_{IN}) \\ p(\mathbf{X}_{270} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{270} | L_{IN}) \end{pmatrix}$$

10

20

【0 2 3 4】

ここで、特徴ベクトルは回転特性（例えば、上述したような特性）を有している。

【0 2 3 5】

また、本実施形態においては、ページ方向は、リスクを最小にすることによって語族とは独立に判定される。このとき、上記リスクは、以下のように決定される。

【0 2 3 6】

【数 4 5】

$$\begin{pmatrix} Risk(O_0 | \mathbf{X}) \\ Risk(O_{90} | \mathbf{X}) \\ Risk(O_{180} | \mathbf{X}) \\ Risk(O_{270} | \mathbf{X}) \end{pmatrix} = \begin{pmatrix} \lambda(O_0 | O_{I0}) & \cdots & \lambda(O_0 | O_{IN0}) & \cdots & \lambda(O_0 | O_{II270}) & \cdots & \lambda(O_0 | O_{IN270}) \\ \lambda(O_{90} | O_{I0}) & \cdots & \lambda(O_{90} | O_{IN0}) & \cdots & \lambda(O_{90} | O_{II270}) & \cdots & \lambda(O_{90} | O_{IN270}) \\ \lambda(O_{180} | O_{I0}) & \cdots & \lambda(O_{180} | O_{IN0}) & \cdots & \lambda(O_{180} | O_{II270}) & \cdots & \lambda(O_{180} | O_{IN270}) \\ \lambda(O_{270} | O_{I0}) & \cdots & \lambda(O_{270} | O_{IN0}) & \cdots & \lambda(O_{270} | O_{II270}) & \cdots & \lambda(O_{270} | O_{IN270}) \end{pmatrix} \mathbf{p}$$

30

【0 2 3 7】

ここで、 $\mathbf{p}$ は尤度値のベクトルであり、以下の式によって決定される。

【0 2 3 8】

【数 4 6】

$$p = \begin{pmatrix} p(\mathbf{X} | O_{I0}) \\ \vdots \\ p(\mathbf{X} | O_{IN0}) \\ p(\mathbf{X} | O_{I90}) \\ \vdots \\ p(\mathbf{X} | O_{IN90}) \\ p(\mathbf{X} | O_{I180}) \\ \vdots \\ p(\mathbf{X} | O_{IN180}) \\ p(\mathbf{X} | O_{I270}) \\ \vdots \\ p(\mathbf{X} | O_{IN270}) \end{pmatrix} = \begin{pmatrix} p(\mathbf{X}_0 | L_{II}) \\ \vdots \\ p(\mathbf{X}_0 | L_{IN}) \\ p(\mathbf{X}_{90} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{90} | L_{IN}) \\ p(\mathbf{X}_{180} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{180} | L_{IN}) \\ p(\mathbf{X}_{270} | L_{II}) \\ \vdots \\ p(\mathbf{X}_{270} | L_{IN}) \end{pmatrix}$$

10

20

【0 2 3 9】

ここで、特徴ベクトルは回転特性（例えば、上述したような特性）を有している。

【0 2 4 0】

本発明に係る別の実施形態は、言語または語族を判定する、第一最小損失分類処理部を含む。また、本実施形態は、上記判定された言語または語族 L を与えられ、ページ方向を以下の式によって判定する第二最小損失分類処理部を含む。

【0 2 4 1】

【数 4 7】

$$\begin{pmatrix} Risk(O_0 | \mathbf{X}) \\ Risk(O_{90} | \mathbf{X}) \\ Risk(O_{180} | \mathbf{X}) \\ Risk(O_{270} | \mathbf{X}) \end{pmatrix} = \begin{pmatrix} \lambda(O_0 | O_{L0}) & \lambda(O_0 | O_{L90}) & \lambda(O_0 | O_{L180}) & \lambda(O_0 | O_{L270}) \\ \lambda(O_{90} | O_{L0}) & \lambda(O_{90} | O_{L90}) & \lambda(O_{90} | O_{L180}) & \lambda(O_{90} | O_{L270}) \\ \lambda(O_{180} | O_{L0}) & \lambda(O_{180} | O_{L90}) & \lambda(O_{180} | O_{L180}) & \lambda(O_{180} | O_{L270}) \\ \lambda(O_{270} | O_{L0}) & \lambda(O_{270} | O_{L90}) & \lambda(O_{270} | O_{L180}) & \lambda(O_{270} | O_{L270}) \end{pmatrix} \mathbf{p}_L$$

30

【0 2 4 2】

ここで、 $\mathbf{p}_L$ は尤度値のベクトルであり、以下の式によって決定される。

【0 2 4 3】

【数 4 8】

$$\mathbf{p}_L = \begin{pmatrix} p(\mathbf{X} | O_{L0}) \\ p(\mathbf{X} | O_{L90}) \\ p(\mathbf{X} | O_{L180}) \\ p(\mathbf{X} | O_{L270}) \end{pmatrix} = \begin{pmatrix} p(\mathbf{X}_0 | L_L) \\ p(\mathbf{X}_{90} | L_L) \\ p(\mathbf{X}_{180} | L_L) \\ p(\mathbf{X}_{270} | L_L) \end{pmatrix}$$

40

【0 2 4 4】

ここで、特徴ベクトルは回転特性（例えば、上述したような特性）を有している。

【0 2 4 5】

50

最小損失分類処理を含む本発明の実施形態において、信頼度は、最小損失分類処理の判定結果に基づいて決定される。ここで、上記信頼度としては、例えば、最小損失値、最小対数損失値、損失値の最小値と最大値との差、対数損失値の最小値と最大値との差、最小損失値と２番目に小さい損失値との差、最小対数損失値と２番目に小さい対数損失値との差、その他、現時点での最善の判定とその代わりとなる判定との間の広がりなどの測定量が挙げられる。

【０２４６】

本発明に係る実施形態では、語族が不明である場合に正しいページ方向を選択する損失が、語族が正しく特定されている場合に間違ったページ方向を選択する損失よりも小さくなるような方法が用いられる。

【０２４７】

また、本発明に係る実施形態では、実際のページ方向が回転した方向である場合にページ方向として可読方向を選択する損失が、実際のページ方向が可読方向である場合にページ方向として回転した方向を選択する損失よりも小さくなるような方法が用いられる。

【０２４８】

本発明に係る実施形態においては、上述の２値化したコーナー・パターンに基づいた特徴分布空間ではなく、別の特徴分布空間を、上記した確率的モデルに用いてもよい。この場合、各語族およびページ方向のトレーニング・データ・セットを用いて、各語族に対し、一つ以上の判別に最適な特徴（判別に用いるのに適した特徴）を決定する。つまり、各語族に対して、図１６に示す流れにしたがって、判別に最適な特徴を決定する。トレーニング・データ・セットに入力された各入力画像に対して、２値化処理２１０を行い、字体の位置を特定する処理（字体要素判定処理）２１２を行う。字体要素判定処理２１２に続いて、上記で特定した字体の各字体画素を中心に、 $m \times m$ の大きさのウィンドウを形成して配置する（処理２１４）。上記中心にある字体画素に対応する $m \times m$ の画素パターンに対するカウンター数を増やしていく（処理２１６）。すなわち、トレーニング・データ・セットの字体要素について上記 $m \times m$ の画素パターンの出現度数を積算していく。本発明に係る実施形態においては、上記ウィンドウは、 5×5 の大きさであってもよい。そして、上記の画素パターンに対して蓄積されたデータに基づき、判別に最適な特徴を決定する（処理２１８）。画素パターンは字体要素の中の特定の特徴を表すものであり、処理２１８においては、各画素パターンの判別性能（判別に用いるのに適しているか否か）を判定する。

【０２４９】

上記実施形態において、ある $m \times m$ の画素パターンの出現度数が、出現度数閾値よりも大きかった場合、その画素パターンが判別に最適な特徴であるとみなしてもよい。また、本発明に係るある実施形態では、 $m \times m$ の画素パターンを p_i とすると、下記の式を満たすとき、当該画素パターン p_i は判別に最適な特徴であるとみなしてもよい。

【０２５０】

【数４９】

$$\frac{\text{count}(p_i)}{\text{count}(\text{glyphs})} > T_{\text{count}}$$

【０２５１】

また、本発明に係る実施形態においては、判別に適しているか否かの選択基準（例えば上式）を満たす上位 M 個の画素パターンのみを考慮してもよい。あるいは、判別に適しているか否かの選択基準を満たす画素パターンのいずれをも考慮するようにしてもよい。

【０２５２】

図１７に示す本発明に係る実施形態では、２値化処理２２０を行った後、字体要素判定処理２２４を行う前に、非テキストブロック排除処理２２２を行い、非テキストブロックを排除する。そして、字体要素判定処理２２４を行った後、ウィンドウ形成処理２２６、

10

20

30

40

50

パターン出現度数積算処理 2 2 8、そして判別性能判定処理 2 3 0 を行う。

【 0 2 5 3 】

さらに、図 1 8 に示す本発明に係る実施形態では、2 値化処理 2 4 0 および字体要素判定処理 2 4 2 を行った後、要素フィルタリング処理 2 4 4 を行い、ノイズとなる字体要素を排除する。要素フィルタリング処理 2 4 4 を行った後、ウィンドウ形成処理 2 4 6、パターン出現度数積算処理 2 4 8、そして、判別性能判定処理 2 5 0 を行う。

【 0 2 5 4 】

本発明に係るさらに別の実施形態においては、非テキストブロック排除処理と要素フィルタリング処理とを両方行ってもよい。

【 0 2 5 5 】

本発明に係る実施形態では、語族の判別に最適な特徴を組み合わせたセットを用いて、MAP 分類法の定式にしたがって、確率的分類処理部を形成してもよく、当該分類処理部は、各語族を識別するよう設計されている。本発明に係る別の実施形態では、語族の判別に最適な特徴を組み合わせたセットを用いて、ML 分類法の定式にしたがって、確率的分類処理部を形成してもよく、当該分類処理部は各語族を識別するよう設計されている。さらに別の実施形態では、語族の判別に最適な特徴を組み合わせたセットを用いて、最小リスクの定式にしたがって、確率的分類処理部を形成してもよく、当該分類処理部は各語族を識別するように設計されている。本発明に係る実施形態においては、特徴を組み合わせたセット（例えば図 6（A）～図 6（H）に示す画素パターン）に対する特徴値を、入力画像に対してそれぞれ測定し、それによって特徴ベクトルを形成しており、上記確率的分類処理部によって、語族が判定される。さらに別の実施形態では、ページ方向が判定される。さらにまた別の実施形態では、語族とページ方向とが判定される。

【 0 2 5 6 】

本発明に係るある実施形態は、上述の各実施形態を実現するシステムとして提供される。図 1 9 を参照しながら本実施形態の画像特徴判定システム 3 0 0 に関して以下に説明する。画像特徴判定システム 3 0 0 に、スキャナ等により読み取られたデジタル画像の画像データが入力される。入力された画像データは 2 値化画像生成部 3 0 1 に入力され、2 値化された画像データが生成される。画像データの 2 値化は、画像データの背景画素とそれ以外（字体要素）とを区別できるようにするために行う。上記 2 値化された画像データは字体要素決定部 3 0 2 に入力され、2 値化された画像データの中の字体要素が決定される。

【 0 2 5 7 】

上記字体要素を含む 2 値化された画像データは参照位置決定部 3 0 3 に入力され、字体要素の参照位置が決定される。上記参照位置は上記字体要素の重心、または上記字体要素の外接矩形の中心、としてもよい。

【 0 2 5 8 】

また上記字体要素を含む 2 値化された画像データは特徴配置決定部 3 0 4 に入力され、字体要素の中の、複数の字体の特徴の各々を示す所定のパターンにマッチする、特徴点の位置（特徴配置）が決定される。特徴点の位置は字体要素の、パターンにマッチする部分の中心としてもよい。

【 0 2 5 9 】

ここで、複数の字体の特徴の各々を示す所定のパターンとは、図 6 に示されるように、例えば、角部（鋭角または鈍角の角部）を示すものであり、当該角部が向いている方向の種類毎に予め定められたパターンである。なお、角部とは、字体要素のエッジを構成する複数の辺の交点部分であり、角部が向いている方向とは、当該複数の辺の両方または何れか一方において、当該交点が上下左右のいずれの端に位置するかを示すものである。例えば、図 6（A）では、上下に延びる辺の上端に交点が存在するパターンを示している、と言える。

【 0 2 6 0 】

字体要素の上記参照位置と上記特徴点の位置とを示すデータは相対位置決定部 3 0 5 に

10

20

30

40

50

入力され、参照位置に対する特徴点の相対的な位置（相対位置）が決定される。上記相対的な位置は、上記参照位置を原点として字体要素を四象限に分け、上記特徴点の位置が位置する象限を示すデータとしてもよい。

【0261】

相対位置を示すデータは分布特性決定部306に入力される。分布特性決定部306は、上記画像データの所定の領域に含まれる複数の上記字体要素について、上記字体の特徴および上記相対位置毎に、上記字体の特徴の出現回数を度数分布として積算し、上記度数分布を上記字体要素あたりの分布として正規化することにより得られる対象特徴ベクトルを算出する。

【0262】

対象特徴ベクトルは画像特徴特定部（画像特徴判定部）307に入力される。画像特徴特定部307は、上記対象特徴ベクトルに基づき画像の特徴（ページ方向、言語または語族等）を特定する。画像特徴特定部307が行う処理として以下に具体例を記す。画像特徴特定部307は、上記対象特徴ベクトルと、見本特徴ベクトル（密度平均ベクトル）および共分散行列とに対して正規分布の確率モデルを適用し、所定の言語または語族（またはページ方向）を仮定した場合の上記対象特徴ベクトルの尤度を求める。画像特徴特定部307は、複数の言語または語族（またはページ方向）についての各尤度に基づき、最大事後確率分類法、最尤分類法、または最小リスク分類法を用いて上記デジタル画像の言語または語族（またはページ方向）を特定する。

【0263】

なお、上記見本特徴ベクトルは、所定のページ方向、言語または語族のトレーニングデジタル画像に対して、対象特徴ベクトルと同様の手順を用いることで算出することができる。また、上記共分散行列は、上記所定のページ方向、言語または語族における上記見本特徴ベクトルに関する共分散行列であり、上記トレーニングデジタル画像と上記見本特徴ベクトルとから求めることができる。各言語、語族、またはページ方向の見本特徴ベクトルおよび共分散行列は、あらかじめメモリ等の記録媒体に保存しておき、図示しない見本分布特性取得部がそこから取得し、画像特徴特定部307に与えるようにしてもよい。また、上記見本特徴ベクトルおよび共分散行列をインターネット等のネットワークまたは電気信号回線を経由して取得してもよい。

【0264】

ニューラルネットワーク（Neural Networks）、サポートベクターマシン（Support Vector Machines：SVM）、隠れマルコフモデル（Hidden Markov Model：HMM）、およびその他の方法を含むトレーニング方法を、画素パターンに基づく特徴、および、ページ方向と言語とを決定するための上述した演算方法、と連携させて使用してもよい。

【0265】

上の明細書中で使用した用語および表現は、説明するために使用したものであり、範囲を限定するものではない。また図示および記述された特徴と同等のものを排除する用語および表現の使用を意図するものでもない。本発明の範囲は、特許請求の範囲に示した請求項でのみ定義および限定されるべきものである。

【0266】

最後に、画像特徴特定システム300の各ブロックは、ハードウェアロジックによって構成してもよいし、次のようにCPUを用いてソフトウェアによって実現してもよい。

【0267】

すなわち、画像特徴特定システム300は、各機能を実現する制御プログラムの命令を実行するCPU（central processing unit）、上記プログラムを格納したROM（read only memory）、上記プログラムを展開するRAM（random access memory）、上記プログラムおよび各種データを格納するメモリ等の記憶装置（記録媒体）などを備えている。そして、本発明の目的は、上述した機能を実現するソフトウェアである画像特徴特定システム300の制御プログラムのプログラムコード（実行形式プログラム、中間コードプログラム、ソースプログラム）をコンピュータで読み取り可能に記録した記録媒体を、上記

10

20

30

40

50

画像特徴特定システム 300 に供給し、そのコンピュータ（または CPU や MPU）が記録媒体に記録されているプログラムコードを読み出し実行することによっても、達成可能である。

【0268】

上記記録媒体としては、例えば、磁気テープやカセットテープ等のテープ系、フロッピー（登録商標）ディスク／ハードディスク等の磁気ディスクや CD-ROM / MO / MD / DVD / CD-R 等の光ディスクを含むディスク系、IC カード（メモリカードを含む）／光カード等のカード系、あるいはマスク ROM / EPROM / EEPROM / フラッシュ ROM 等の半導体メモリ系などを用いることができる。

【0269】

また、画像特徴特定システム 300 を通信ネットワークと接続可能に構成し、上記プログラムコードを通信ネットワークを介して供給してもよい。この通信ネットワークとしては、特に限定されず、例えば、インターネット、イントラネット、エキストラネット、LAN、ISDN、VLAN、CATV 通信網、仮想専用網（virtual private network）、電話回線網、移動体通信網、衛星通信網等が利用可能である。また、通信ネットワークを構成する伝送媒体としては、特に限定されず、例えば、IEEE 1394、USB、電力線搬送、ケーブル TV 回線、電話線、ADSL 回線等の有線でも、IrDA やリモコンのような赤外線、Bluetooth（登録商標）、802.11 無線、HDMI、携帯電話網、衛星回線、地上波デジタル網等の無線でも利用可能である。なお、本発明は、上記プログラムコードが電子的な伝送で具現化された、搬送波に埋め込まれたコンピュータデータ信号の形態でも実現され得る。

【0270】

上述した明細書において用いた用語や表現は、説明するために用いたものであり、これに限定されるものではない。また、上記で使用された用語や表現によって、上記で示されたあるいは記載された特徴またはその一部分と同等のものを排除するものではない。本発明の範囲は、特許請求の範囲においてのみ限定されるものである。

【産業上の利用可能性】

【0271】

スキャン、コピー、およびデジタル画像処理における電子文書の最有力のテキスト方向またはテキスト言語の自動判定に利用できる。

【符号の説明】

【0272】

- 2, 4, 6, 8 テキスト文書
- 30 字体要素
- 32 中心（参照点）
- 34 字体要素外接矩形（外接矩形）
- 35, 36, 37, 38 象限
- 40 字体要素
- 42 重心（参照点）
- 44, 45, 46, 47 象限
- 55 背景画素
- 56 字体画素
- 57 無関係画素
- 90 字体要素
- 160 2 値化処理（2 値化画像生成）
- 162 字体要素判定処理（字体要素特定）
- 164 特徴ベクトル決定処理（対象特徴ベクトル算出）
- 166 確率的分類処理
- 300 画像特徴特定システム
- 301 2 値化画像生成部

10

20

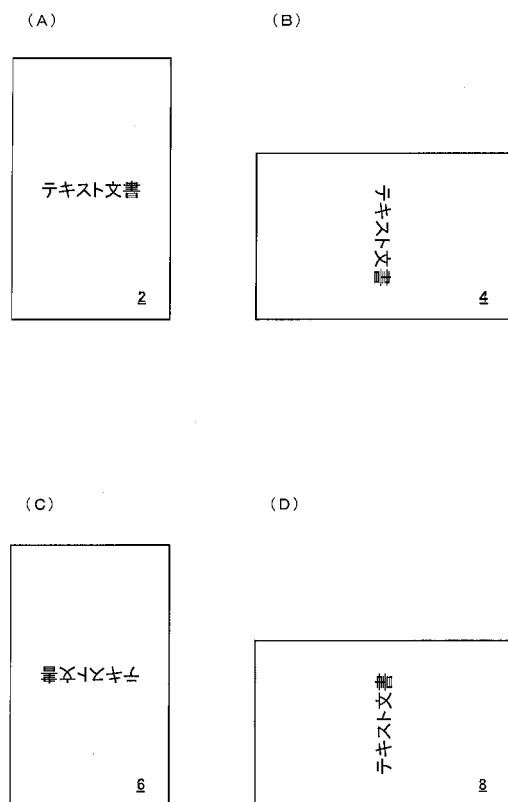
30

40

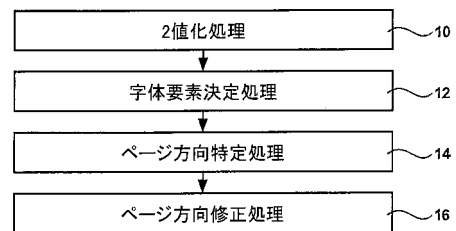
50

- 3 0 2 字体要素決定部
- 3 0 3 参照位置決定部
- 3 0 4 特徴配置決定部
- 3 0 5 相対位置決定部
- 3 0 6 分布特性決定部
- 3 0 7 画像特徴特定部 (画像特徴判定部)

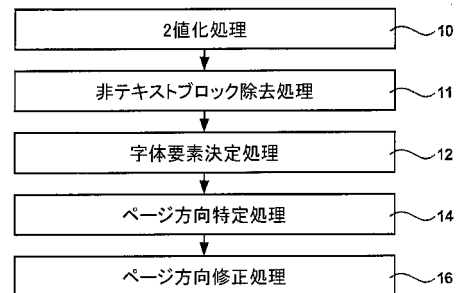
【図 1】



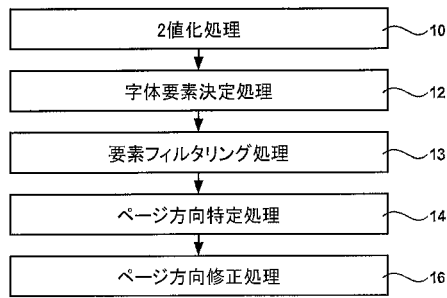
【図 2】



【図 3】

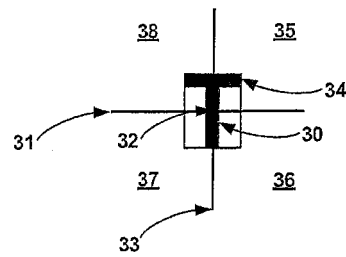


【図 4】

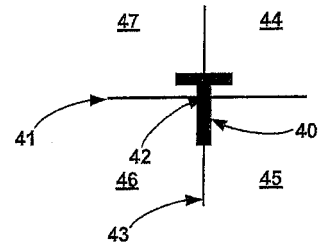


【図 5】

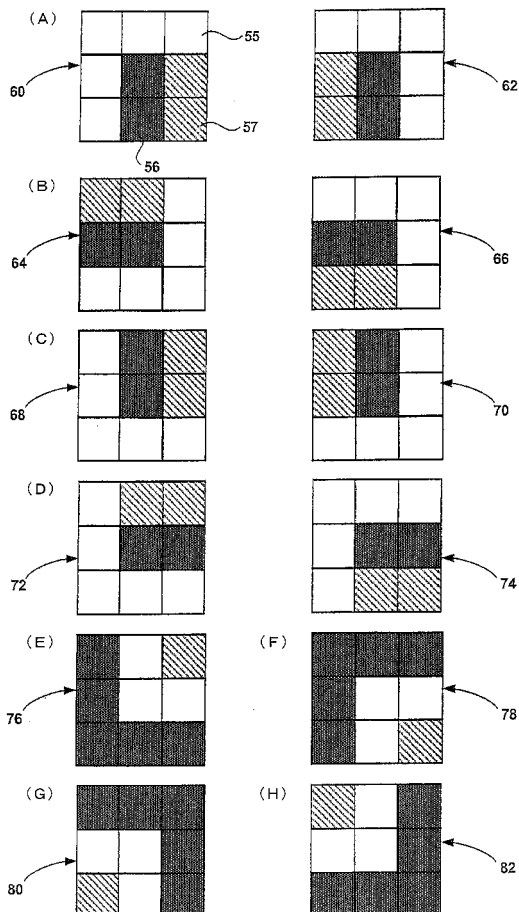
(A)



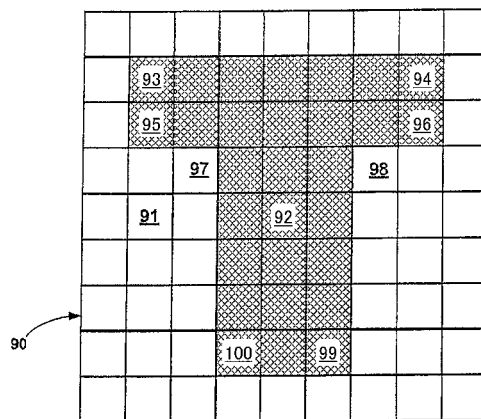
(B)



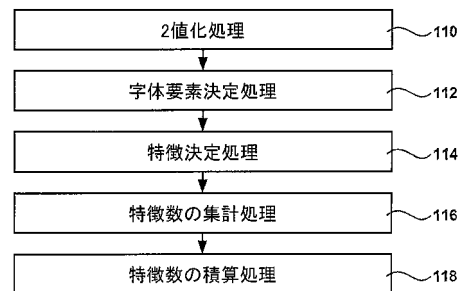
【図 6】



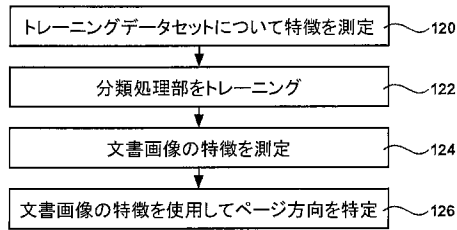
【図 7】



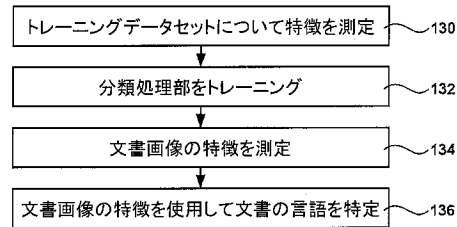
【図 8】



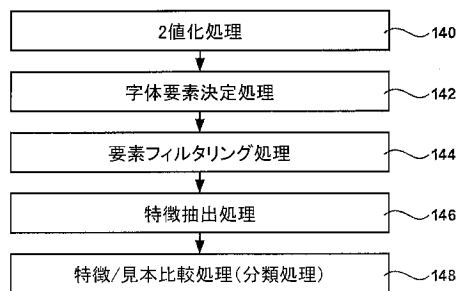
【図 9】



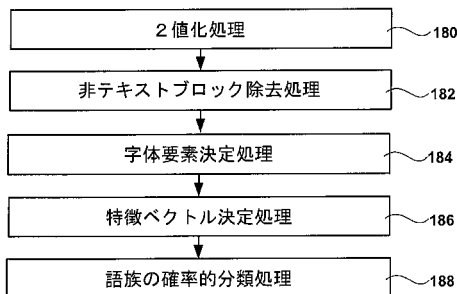
【図 10】



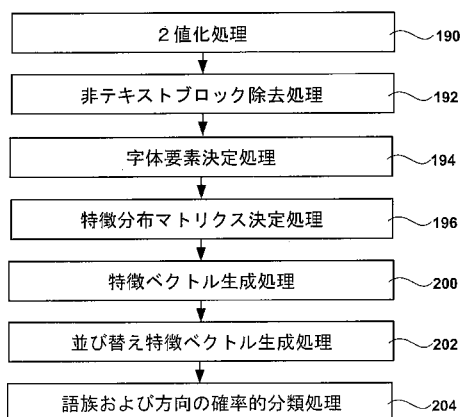
【図 11】



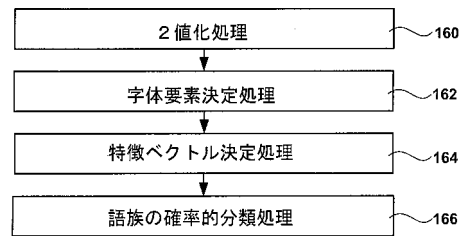
【図 14】



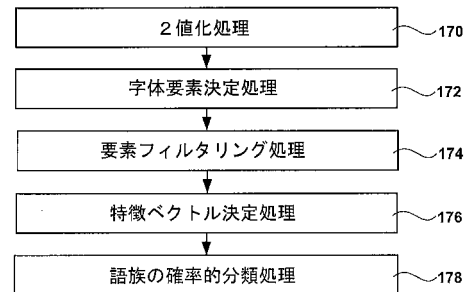
【図 15】



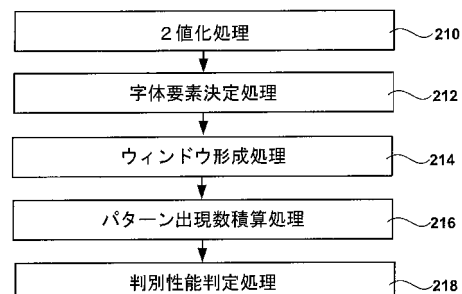
【図 12】



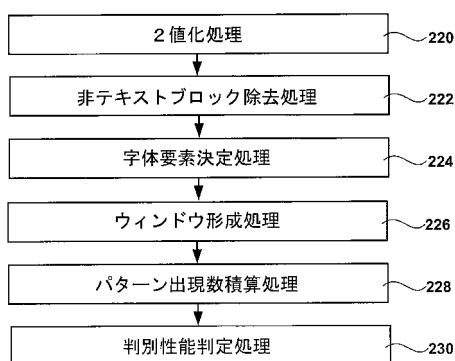
【図 13】



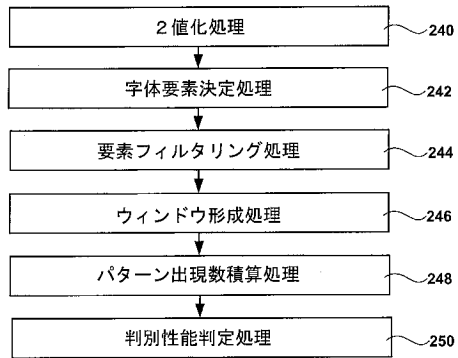
【図 16】



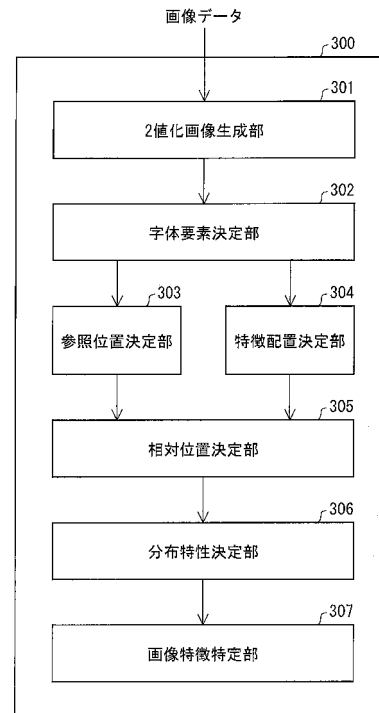
【図 17】



【図 18】



【図 19】



フロントページの続き

(72)発明者 ジョン マシュー スピーグル

アメリカ合衆国 ワシントン州 98607, カマス, ノースウェスト パシフィック リム ブ
ールバード 5750 シャープ ラボラトリーズ オブ アメリカ, インコーポレイテッド内

Fターム(参考) 5B029 AA01 BB02 CC17

5B064 AA01 AB02 AB16 BA01