

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5004082号
(P5004082)

(45) 発行日 平成24年8月22日(2012.8.22)

(24) 登録日 平成24年6月1日(2012.6.1)

(51) Int.Cl.		F I			
G06T	1/00	(2006.01)	G06T	1/00	200C
G06F	17/30	(2006.01)	G06F	17/30	170B
G06T	7/00	(2006.01)	G06T	7/00	300F

請求項の数 13 (全 20 頁)

(21) 出願番号	特願2007-196574 (P2007-196574)	(73) 特許権者	505127721
(22) 出願日	平成19年7月27日 (2007.7.27)		公立大学法人大阪府立大学
(65) 公開番号	特開2009-32109 (P2009-32109A)		大阪府堺市中央区学園町1番1号
(43) 公開日	平成21年2月12日 (2009.2.12)	(74) 代理人	100065248
審査請求日	平成22年7月14日 (2010.7.14)		弁理士 野河 信太郎
		(74) 代理人	100130672
			弁理士 伊藤 寛之
		(72) 発明者	中居 友弘
			大阪府堺市中央区学園町1-1 公立大学法
			人大阪府立大学内
		(72) 発明者	黄瀬 浩一
			大阪府堺市中央区学園町1-1 公立大学法
			人大阪府立大学内

最終頁に続く

(54) 【発明の名称】 文書画像検索方法、文書画像登録方法、そのプログラムおよび装置

(57) 【特許請求の範囲】

【請求項1】

撮像されあるいは読取られた文書および/または画像(検索質問文書画像)の特徴点から計算される特徴量とデータベース中に登録された複数の文書および/または画像(登録文書画像)の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する方法であって、

検索質問文書画像から複数の連結成分を抽出し、

抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、

幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、

第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、

各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定し、

前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とする文書画像検索方法。

【請求項2】

前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べたベクトルとして得られる請求項1記載の文書画像検索方

法。

【請求項 3】

検索質問文書画像または登録文書画像は、分かち書きされない言語の文書を含む請求項 1 または 2 に記載の文書画像検索方法。

【請求項 4】

前記第 1 不変量は、前記組み合わせに係る各特徴点を結ぶ複数の多角形の面積比として得られる請求項 1 記載の文書画像検索方法。

【請求項 5】

前記第 2 不変量は、前記組み合わせに係る各連結成分の面積を正規化して離散化したものを要素とするベクトルである請求項 1 記載の文書画像検索方法。

10

【請求項 6】

前記第 2 不変量は、前記組み合わせに係る各連結成分の面積比を要素とするベクトルである請求項 1 記載の文書画像検索方法。

【請求項 7】

前記第 2 不変量は、前記組み合わせに係る各連結成分の面積と前記連結成分の凸包の面積との比を要素とするベクトルである請求項 1 記載の文書画像検索方法。

【請求項 8】

前記第 2 不変量は、前記組み合わせに係る各連結成分の周長と前記連結成分の凸包の周長との比を要素とするベクトルである請求項 1 記載の文書画像検索方法。

【請求項 9】

20

撮像されあるいは読取られた文書および/または画像(文書画像)の特徴点から計算される特徴量を前記文書画像と対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する登録方法であって、

登録すべき文書画像から複数の連結成分を抽出し、

抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、

幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、

30

第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、

各特徴量を前記文書画像に対応付けてデータベースに登録し、

前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とする文書画像登録方法。

【請求項 10】

撮像されあるいは読取られた文書および/または画像(検索質問文書画像)の特徴点から計算される特徴量とデータベース中に登録された複数の文書および/または画像(登録文書画像)の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する処理をコンピュータに実行させるためのプログラムであって、

40

検索質問文書画像から複数の連結成分を抽出し、

抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、

幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、

第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、

各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定する処理を実行させ、

前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とする文書画像検

50

索プログラム。

【請求項 1 1】

撮像されあるいは読取られた文書および／または画像（文書画像）の特徴点から計算される特徴量を前記文書画像に対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する処理をコンピュータに実行させるためのプログラムであって、登録すべき文書画像から複数の連結成分を抽出し、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、各特徴量を前記文書画像に対応付けてデータベースに登録する処理を実行させ、前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とする文書画像登録プログラム。

10

【請求項 1 2】

撮像されあるいは読取られた文書および／または画像（検索質問文書画像）の特徴点から計算される特徴量とデータベース中に登録された複数の文書および／または画像（登録文書画像）の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する装置であって、検索質問文書画像から複数の連結成分を抽出する連結成分抽出部と、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とする特徴点決定部と、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求める不変量算出部と、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出する特徴量算出部と、各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定する検索部とを備え、前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とすることを特徴とする文書画像検索装置。

20

30

【請求項 1 3】

撮像されあるいは読取られた文書および／または画像（文書画像）の特徴点から計算される特徴量を前記文書画像に対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する装置であって、登録すべき文書画像から複数の連結成分を抽出する連結成分抽出部と、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とする特徴点決定部と、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求める不変量算出部と、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出する特徴量算出部と、各特徴量を前記文書画像に対応付けてデータベースに登録する登録部とを備え、前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べた面積の順位に係るものであることを特徴とすることを特徴

40

50

とする文書画像登録装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、データベースに登録された文書および/または画像の中から、検索質問としての文書および/または画像に対応するものを検索する処理に係る文書画像検索方法、文書画像登録方法、そのプログラムおよび装置に関する。

【背景技術】

【0002】

発明者らは、Webカメラを用いたリアルタイム文書画像検索を以前に提案している（例えば、非特許文献1参照）。これは、Webカメラで撮影された印刷文書の画像を検索質問としてリアルタイムで検索処理を行い、データベースから対応する文書画像（文書および/または画像）を見つけて提示するものである。

【0003】

このようなカメラに基づく文書画像検索の目的は、印刷文書からの情報のアクセス手段を実現することである。すなわち、カメラ付き携帯電話などのデジタルカメラとディスプレイが一体となった機器を用いて対象を撮影し、検索を通じた認識に基づいて撮影対象に関連するサービスを起動することを目指している。また、リアルタイム化することで利用者の自発的な要求に基づかずにサービスの提供を行うことができるようになる。つまり、カメラに写るものを常に検索することで、サービスの関連付けられた対象が撮影された際に自動的に利用者に提示するといった、プッシュ型のサービスを実現できる。

【0004】

我々の提案しているリアルタイム文書画像検索は、LLAH(Locally Likely Arrangement Hashing)と呼ばれる画像認識手法（例えば、非特許文献2、3および特許文献1参照）に基づくものである。LLAHは、画像から抽出された特徴点の配置に基づいて画像を表現する特徴量を計算し、その検索を行うものである。LLAHの特徴として、高速かつ高精度に画像の認識が可能であり、隠れや紙面の湾曲などの外乱に強い（ロバストである）という点が挙げられる。また、特徴点の座標のみから特徴量を計算するために、SIFTなどの複雑な処理を要する特徴計算を必要としない点もリアルタイム処理を実現する上で大きな利点となっている。

【特許文献1】国際公開第2006/092957号パンフレット

【非特許文献1】中居，黄瀬，岩村：「特徴点の局所的配置に基づくリアルタイム文書画像検索とその拡張現実への応用」，電子情報通信学会技術研究報告，PRMU2006-66，pp.41-48（2006）。

【非特許文献2】中居，黄瀬，岩村：「特徴点の局所的配置に基づくデジタルカメラを用いた高速文書画像検索」，電子情報通信学会論文誌（D），J89-D，9，pp.2045-2054（2006）。

【非特許文献3】中居，黄瀬，岩村：<デジタルカメラを用いた高速文書画像検索におけるアフィン不変量および相似不変量の利用=5 信学技報，Vol. 105，No. 614，PRMU2005-188，pp.25-30（2006）。

【発明の開示】

【発明が解決しようとする課題】

【0005】

その一方で、LLAHの検索性能は特徴点の配置に大きく依存するため、柔軟性に欠けるといふ側面もある。即ち、特徴点の配置のみから特徴量を計算するため、特徴点が規則的な並びとなる場合などは特徴量の識別性が低くなるため画像を識別することが困難になる。また、登録される画像と検索質問の画像から得られる特徴点が大きく異なる場合、特徴量の安定性が低くなるため対応する画像を見つけることは難しい。以上の理由から、LLAHの適用対象は単語の重心という安定かつ識別性の高い特徴点を与えるような、単語が分かち書きされた文書（例えば英文文書のように、単語と単語の間にスペースが置かれる文書）

10

20

30

40

50

に限られるのが実情であった。

【0006】

この発明は、以上のような事情を考慮してなされたものであって、LLAHの特徴量計算処理を改良し、これまで困難であった単語が分かち書きされない文書（たとえば、日本語文書）についても検索可能な文書画像検索手法を提供するものである。即ち、LLAHの特徴量計算方法を改良し、識別性や安定性に問題のある特徴点からでも高精度な検索を可能にした画像認識法を提供する。この発明により、従来困難であった日本語文書等単語が分かち書きされない文書の高精度な検索が実現できる。

【課題を解決するための手段】

【0007】

この発明は、撮像されあるいは読取られた文書および/または画像（検索質問文書画像）の特徴点から計算される特徴量とデータベース中に登録された複数の文書および/または画像（登録文書画像）の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する方法であって、検索質問文書画像から複数の連結成分を抽出し、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定する文書画像検索方法を提供する。

【発明の効果】

【0008】

この発明の文書画像検索方法は、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出するので、各特徴点とその近傍の特徴点との配置関係だけでは十分な識別性が得られないような文書画像であっても、各連結成分の属性を加えることによって高い精度の検索結果を得ることができる。

【0009】

即ち、この発明は、特徴点の配置に加えて連結成分の面積を特徴量計算に用いることで実現される。実験により、日本語文書を対象としたリアルタイム文書画像検索が実現されたことが確認された。

また、前記文書画像検索方法に対応するものとして、この発明は、撮像されあるいは読取られた文書および/または画像（文書画像）の特徴点から計算される特徴量を前記文書画像と対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する登録方法であって、登録すべき文書画像から複数の連結成分を抽出し、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、各特徴量を前記文書画像に対応付けてデータベースに登録する文書画像登録方法を提供する。

【0010】

また、異なる観点から、この発明は、撮像されあるいは読取られた文書および/または画像（検索質問文書画像）の特徴点から計算される特徴量とデータベース中に登録された複数の文書および/または画像（登録文書画像）の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する処理をコンピュータに実行させるためのプログラムであって、検索質問文書画像から複数の連結成分を抽出し、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、幾何学的歪みに対して実質的

10

20

30

40

50

に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定する処理を実行させるための文書画像検索プログラムを提供する。

【0011】

また、前記文書画像検索プログラムに対応するものとして、この発明は、撮像されあるいは読取られた文書および/または画像(文書画像)の特徴点から計算される特徴量を前記文書画像に対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する処理をコンピュータに実行させるためのプログラムであって、登録すべき文書画像から複数の連結成分を抽出し、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とし、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求め、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出し、各特徴量を前記文書画像に対応付けてデータベースに登録する処理を実行させるための文書画像登録プログラムを提供する。

10

【0012】

さらに、異なる観点から、この発明は、撮像されあるいは読取られた文書および/または画像(検索質問文書画像)の特徴点から計算される特徴量とデータベース中に登録された複数の文書および/または画像(登録文書画像)の特徴点から得られる特徴量とを比較して検索質問文書画像に対応する登録文書画像を検索する装置であって、検索質問文書画像から複数の連結成分を抽出する連結成分抽出部と、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とする特徴点決定部と、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求める不変量算出部と、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出する特徴量算出部と、各特徴量の比較結果を統計的に処理して検索質問文書画像に対応する登録文書画像を特定する検索部とを備えることを特徴とする文書画像検索装置を提供する。

20

30

【0013】

また、前記文書画像検索装置に対応するものとして、この発明は、撮像されあるいは読取られた文書および/または画像(文書画像)の特徴点から計算される特徴量を前記文書画像に対応付けて予め登録しておき、検索質問として撮像されあるいは読取られた文書画像の特徴点から得られる特徴量を登録された特徴量と比較して検索質問に対応する文書画像を検索するために用いられるデータベースに前記文書画像を登録する装置であって、登録すべき文書画像から複数の連結成分を抽出する連結成分抽出部と、抽出された連結成分の重心を求めて各連結成分に対応する特徴点とする特徴点決定部と、幾何学的歪みに対して実質的に不変な第1および第2不変量であって、各特徴点とその近傍の特徴点との組み合わせから得られる第1不変量および前記組み合わせに係る各連結成分の属性から得られる第2不変量をそれぞれ求める不変量算出部と、第1不変量と第2不変量とを組み合わせる各特徴点に対応する特徴量を算出する特徴量算出部と、各特徴量を前記文書画像に対応付けてデータベースに登録する登録部とを備えることを特徴とする文書画像登録装置を提供する。

40

【0014】

以下、この発明の好ましい態様について説明する。

前記第2不変量は、前記組み合わせに係る各特徴点を、それに対応する連結成分の面積の大きい順または小さい順に並べたベクトルとして得られるものであってもよい。

【0015】

また、検索質問文書画像または登録文書画像は、分かち書きされない言語の文書を含ん

50

でいてもよい。

【0016】

さらにまた、前記第1不変量は、前記組合せに係る各特徴点を結ぶ複数の多角形の面積比として得られてもよい。

【0017】

あるいは、前記第2不変量は、前記組合せに係る各連結成分の面積を正規化して離散化したものを要素とするベクトルであってもよい。

【0018】

あるいはまた、前記第2不変量は、前記組合せに係る各連結成分の面積比を要素とするベクトルであってもよい。

10

【0019】

また、前記第2不変量は、前記組合せに係る各連結成分の面積と前記連結成分の凸包の面積との比を要素とするベクトルであってもよい。

【0020】

あるいはまた、前記第2不変量は、前記組合せに係る各連結成分の周長と前記連結成分の凸包の周長との比を要素とするベクトルであってもよい。

【0021】

ここで示した種々の好ましい態様は、それら複数を組み合わせることもできる。

【発明を実施するための最良の形態】

【0022】

20

この実施形態では、単語が分かち書きされない文書の代表例として日本語文書を対象としたリアルタイム文書画像検索法の一例を説明する。これは、非特許文献2で提案されたようなリアルタイム文書画像検索法、即ち、単語が分かち書きされた文書に好適な文書画像検索法の検索対象を日本語文書等単語が分かち書きされない文書へ拡張するものである。

【0023】

日本語文書の検索を実現する上では、前記文書画像検索の処理手順のうち特徴点抽出処理が問題となる。分かち書きのされない日本語文書では、非特許文献1で英文文書に対して行ったように単純な画像処理で単語の抽出を行うことは困難である。そこで、提案手法では連結成分の重心を用いる。

30

以下、図面を用いてこの発明をさらに詳述する。なお、以下の説明は、すべての点で例示であって、この発明を限定するものと解されるべきではない。

【0024】

従来のLLAHを適用した文書画像検索処理 - 参考例

はじめに、従来のLLAHを適用した文書画像検索の処理手順の概要を説明する。これによって、本願発明の位置付けがより明確に理解されるであろう。

【0025】

従来のLLAHを適用した文書画像検索システムの構成を図6に示す。図6で、登録文書は、検索対象の文書画像データベースに登録される文書画像である。登録文書および検索質問は、特徴点抽出により点の集合に変換された後、登録の場合は登録処理へ、検索の場合は検索処理へ入力される。登録処理および検索処理では、抽出された特徴点から不変量を用いた特徴量をそれぞれ求めるが、この特徴量計算は同じ処理を用いる。

40

【0026】

登録処理では、特徴点から得られた特徴量をハッシュ表のインデックスに変換し、それを用いて文書画像データベースへ登録する。一方、検索処理では、特徴量から同様にインデックスを計算し、投票することで所望の文書画像を検索する。以下、各々について述べる。

【0027】

特徴点抽出

まず、登録文書または検索質問を特徴点の集合として表現すべく、特徴点抽出処理が行

50

われる。文書画像の特徴を点（特徴点）の集合で表し、各特徴点に係る特徴量を用いて識別できるようにするためである。

【 0 0 2 8 】

特徴点抽出で重要なことは、特徴点の再現性、すなわち射影変換・ノイズ・低解像度の影響下でも同一の特徴点を得られることである。英文文書における単語の重心は、この条件を満たす特徴点の1つである。これは、英文文書では単語と単語の間に空白があり、分離が比較的容易なためである。

【 0 0 2 9 】

例を用いて手順の概略を説明する。入力画像(図7)は、まず適応2値化により2値画像(図8)に変換される。次に、2値画像から以下のように単語領域を得る。まず、ガウシアンフィルタを適用して2値画像をにじませる。このときガウシアンフィルタのパラメータは文字サイズの推定値（連結成分の面積の最頻値の平方根）に基づいて適応的に定める。そして、にじませた画像に対して、再度適応2値化を適用し、2値画像(図9)を得る。この画像の連結成分を単語領域とみなし、その重心を特徴点とする。図9に対しては図10が得られる。

【 0 0 3 0 】

幾何的歪に対する不変量の特徴量計算への適用

幾何的歪に対する安定性を得るために、各特徴点に係る特徴量として、幾何的歪に対する不変量を用いる。この実施形態では、幾何的歪みの一種であるアフィン歪みに対する不変量（アフィン不変量）を用いる。アフィン不変量は同一平面上の4点から計算され、アフィン変換における不変量である。アフィン変換とは、直線の平行性が維持される幾何変換であり、デジタルカメラで撮影された画像で生じる射影変換よりも自由度の低い変換である。射影変換を受けた平面においても、多くの場合、局所領域ではその変換がアフィン変換に近似されるため、アフィン不変量を利用することが可能となる。アフィン不変量は同一平面上の4点ABCDの座標から $P(A,C,D)/P(A,B,C)$ で計算される。即ち、同一平面上の4点の組み合わせから得られる2つの三角形の面積比として求められる。アフィン不変量を用いて求めた各特徴点についての特徴量を離散化し、各特徴点のインデックスとして用いる。なお、他の種類の不変量の特徴量計算に適用することも可能である。例えば、射影歪みに対する不変量として複比が知られている。この複比を特徴量計算に用いてもよい。あるいは、相似歪みに対して、直線間の角度、距離の比、面積の比、距離の2乗と面積の比などの相似不変量を用いてもよい。

【 0 0 3 1 】

特徴量に求められる性質

特徴量とは、文書画像の各特徴点を表現する量である。文書画像検索は、検索質問および登録文書のそれぞれについて、特徴点から得られる特徴量を計算し、それらの値を比較することで検索質問と登録文書が対応しているか否かを判断する。検索質問に対応する登録文書を正確に、かつ高速に検索できる特徴量が優れた特徴量であるといえる。

【 0 0 3 2 】

即ち、優れた特徴量の第一の条件は、各種の幾何的歪み（射影歪み、アフィン歪みなど）の影響を受けても同じ文書の同じ点から同じ特徴量が得られることである（特徴量の安定性）。もし登録文書と検索質問から異なる特徴量が得られれば、検索によって正しく対応する特徴点を見つけることはできない。優れた特徴量の第二の条件は、異なる点からは異なる特徴量が得られることである（特徴量の識別性）。もし異なる文書から同じ特徴量が得られれば、検索の際に正しく対応する特徴点だけでなく対応しない特徴点まで見つかることになる。また、いうまでもなく、安定性や識別性の高い特徴量であっても、計算量が膨大であれば利用は困難となる。従って、計算量が小さいことも特徴量の満たすべき条件である。

【 0 0 3 3 】

第一の条件（安定性）を満たすために、幾何的歪に対する不変量を用いることは前述したが、安定性をより高めるため、LLAHでは各特徴点の近傍点から複数の点の組み合わせを

10

20

30

40

50

作成し、そこから複数の特徴量を計算する。これにより、着目した特徴点を複数の特徴量の組み合わせ（特徴ベクトル）で表す。これは、図11で、射影歪みの影響があってもある程度広い範囲の近傍 n 点（図11では8点）のうち m 点（図11では7点）までは同じ点が含まれるという仮定に基づいている。図11(a)と(b)は、同じ原稿が異なる射影歪みを受けたものである。近傍 n 点のうち m 点が同一であるならば、図12のように n 点からすべての m 点の組み合わせ $P_m(0), P_m(1), \dots, P_m(\binom{n}{m}-1)$ を作成し、それぞれ特徴量を計算することで、少なくとも1つは同じ特徴量が得られると考えられるからである。

【0034】

また、第二の条件（識別性）をより高めるため、各特徴点の特徴量計算に用いる近傍特徴点の数として十分な識別性が得られるような数を選択する。 m が適度に大きければ、計算される不変数の数が多くなるため、同じ特徴量が偶然に現れる可能性は低くなる。 m が大きすぎると不変数の数が増加するが、誤差の影響で異なる不変数が計算される可能性がかえって高くなる。

10

【0035】

登録・検索

以上で説明した特徴量を用いた文書画像の登録および検索方法について述べる。ここでは概略のみを説明する。詳細については、例えば、非特許文献1または3を参照されたい。まず、登録について述べる。ここまでに述べたように、この手法では、各特徴点の近傍 n 点から m 点を取り出し、 $\binom{n}{m}$ 次元のアフィン不変量のベクトルによって表現される m 点の配置を特徴量としている。この特徴量を、以下に示すハッシュ関数によってハッシュ表のインデックス H_{index} に変換する。

20

【0036】

【数1】

$$H_{index} = \left(\sum_{i=0}^{\binom{n}{m}-1} r_{(i)} k^i \right) \bmod H_{size} \quad (1)$$

【0037】

ここで、 k は不変量の量子化レベル、 H_{size} はハッシュ表のサイズである。ただし、式(1)のハッシュ関数は一例であって、これに限定されるものではない。

得られたインデックスを用いて図13に示されるようなハッシュ表へ、登録文書の識別番号である文書IDと点の識別番号である点ID、不変量 $r_{(i)}$ ($i=0, 1, \dots, \binom{n}{m}-1$)を登録する。登録時に衝突が生じた場合、データは図13のようにリスト構造で付け加えられる。

30

【0038】

次に検索について述べる。登録時と同様に、特徴点の局所的配置から特徴量を求め、前式(1)を用いてハッシュ表のインデックスを求める。インデックスを用いて登録処理で作成されたハッシュ表にアクセスし、登録されている文書IDの文書に対して投票する。このような処理をすべての点について繰り返し、最終的に最も多くの得票数を得た文書を検索結果とする。

【0039】

日本語文書への拡張

40

次に、上述の従来手法を日本語文書へ適用できるようにしたこの発明の手法について、従来手法との差異点である特徴点抽出および特徴量計算について説明する。

【0040】

特徴点として日本語文書における連結成分の重心を用いる場合、その識別性が問題となる。日本語文書では連結成分の重心は多くの場合で文字の重心であり、さらに文字はほぼ等間隔に配置されているため、得られる特徴点は規則的な配置をもつ。その結果、特徴点の配置のみから得られる特徴量は、異なる文書からも同じものが得られるため、識別性に欠ける。

【0041】

図1は、従来のLLAHを適用した文書画像検索手法による検索質問 Q と登録画像 D の各特

50

徴点の対応関係を示す説明図である。即ち、日本語文書のPDFファイルを変換して得た登録画像と、登録画像を印刷したものをカメラで撮影して得た検索質問画像の特徴点の対応関係を従来のLLAHで求めたものである。右側の長方形はデータベース中の文書画像（登録文書）Dを表し、左側の四角形はカメラで撮影した画像（検索質問）Qを表す。登録文書Dおよび検索質問Q内の点は、それぞれの画像から抽出された特徴点を示す。また、両者の間に引かれた無数の線は、特徴点同士の対応関係を表すものである。従来法では、特徴量の識別性が不足しているために、図1に示されるように、正しく対応する特徴点だけでなく誤ったものにおいても対応が生じている。

【0042】

特徴量の識別性が不足する場合に、従来のLLAHで用いられる対策は、より多くの特徴点を用いて特徴量を計算することである。特徴量計算に用いられる点の数が多ければ多いほど、異なる文書が同じ点の配置をもつ可能性は低くなるため、その識別性は高くなる。ただし、そのようなアプローチは特徴点の安定性が高いことを前提としている。特徴点の安定性が低いと、特徴量計算に用いる特徴点の数が多いときに同じ文書からでも同じ特徴点を得ることが困難になり、特徴量の安定性が低くなる。そして、日本語文書における連結成分の重心は、その配置に識別性が欠けるだけでなく、安定性についても問題がある。そのため、より多くの特徴点を用いるという方策は特徴量の安定性を低下させてしまう。

【0043】

この実施形態では、特徴量計算に用いる点の数を増やすことなく特徴量の識別性を向上させるため、連結成分の面積を用いる。まず、あらかじめ特徴点抽出処理で連結成分の重心だけでなく面積も計算しておく。そして、特徴量計算処理において特徴点の元となった連結成分の面積の順位を求め、それを従来の特徴点の配置に基づく特徴量に加えることでより識別性の高い特徴量とする。

【0044】

特徴量計算の処理手順

図2は、この実施形態に係る特徴量計算の処理手順を示すフローチャートである。以下では、この手順に沿って詳しく説明する。

【0045】

まず、図2で、非特許文献1に記載のものと同様の手法によって、画像から連結成分を抽出する（Step1）。英文文書を対象とした場合には、単語領域が連結成分となるように画像処理のパラメータを調整している。ところが、日本語文書には分かち書きされた単語が存在しないため、この処理をそのまま適用することはできない。そこで、本手法では、文字の一部あるいは全部が連結成分となるようにパラメータを調整し、画像処理を施す。他の分かち書きされていない言語による文書でも同様のアプローチが可能である。連結成分が求められると、その重心を求めて特徴点とする（Step3）。この処理は非特許文献1に記載のものと同じである。

【0046】

次に、具体的に特徴量を抽出していく。このとき、面積比という量と面積順位という量の2つを用いる。前者の面積比は図2のStep5で求められるものである。面積比は、特徴点の組み合わせから得られる幾何学的不変量（具体的にはアフィン不変量）であり、非特許文献1で用いたものと同じである。もう一方の面積順位はStep7で求められる。こちらは、従来の特徴点ではなく、連結成分から得られる特徴量であり、不変量あるいはそれに類する性質を持つものであればよい。面積の順位は射影変換のもとでの不変量ではないが、本特徴抽出処理のように局所領域を見る場合においては、不変量に準じるもの、すなわち通常の場合、概ね不変として扱うことができるものである。他には、連結成分の面積を正規化した上で離散化したものを用いることができるほか、アフィン不変量である「連結成分の面積比」や「連結成分の凸包の面積と連結成分の面積の比」などの面積に関する特徴量、連結成分やその凸包の周囲長の比などの長さに関するものなど、連結成分から得られる特徴量であって、不変量あるいはそれに準ずるものであれば、この処理に用いることができる。なお、Step5とStep7とは、いずれを先に処理してもよい。図2は、この点を明

10

20

30

40

50

示すべくStep5とStep7とを並列的に記載している。

【0047】

以上のように、非特許文献1の手法では捨て去っていた連結成分を見直し、新しい特徴量を導入することによって、特徴点により正確に区別できるようになり、認識精度の向上が期待できる。最後に、これらの2つの特徴量を合わせて、特徴ベクトルを作成する(Step9)。

前述の説明からわかるように、この発明の特徴は、主として図2のStep5,7,9の処理にある。そこで、これらの処理についてさらに詳しく述べる。

【0048】

不変量(面積比)と面積順位を用いた特徴量計算の具体例

10

図3は、この実施形態に係る文書画像検索処理において、不変量(面積比)と面積順位を用いた特徴量計算の具体例を示す説明図である。ここで、右図中央にある白抜きの小さい円は注目する特徴点Pを表し、その他の小さい円は周囲の特徴点 $i_1 \sim i_6$ を表す。また、小さい円を含んで、様々な形をした大きな図形は、特徴点の元となった連結成分を表す。

【0049】

非特許文献1の場合と同様、本手法では、注目する各特徴点の周囲に存在する特徴点を用いて、面積比を計算する。具体的には、以下の3ステップである。まず、注目する特徴点の周囲 n 点(ここでは、 $n=7$ とする)から、特徴量計算の対象となる m 点(ここでは、 $m=6$ とする)を選ぶ。選択の可能性は nC_m 通り存在し、その各々について特徴ベクトルが計算されるので、各特徴点は、 nC_m 個の特徴ベクトルによって索引付けされることになる。

20

【0050】

今、図3の6点選ばれたとしよう。次に行う処理は、これらの6点から4点を選ぶ組み合わせをすべて求め、各組み合わせから面積比を計算することである。4点を用いれば、3角形を2つ作ることができるので、その面積の比を不変量(面積比)として登録する。6点から4点を選ぶ組み合わせは、 ${}_6C_4=15$ 通りあるので、面積比は15個得られる。これを面積比のベクトル (s_1, \dots, s_{15}) とする。以上は、非特許文献1で用いている面積比の特徴量と同じである。これは、特徴点の組合せから得られる特徴量である。

【0051】

もう一方の「面積順位」の特徴量は以下のように求める。注目する特徴点の周囲 m 点について、それを生成する元となった連結成分の面積を調べる。そして、大きいものから順に、(1), ..., (6)のように順位を定める。各特徴点は、識別番号 i_k を持っているので、1位から順に識別番号を並べたリストを作成する。図3では、 $i_5, i_2, i_4, i_6, i_1, i_3$ の順であるので、リストは、 $(i_5, i_2, i_4, i_6, i_1, i_3)$ となる。これを面積順位の特徴ベクトルとする。

30

【0052】

最後に面積比のベクトルと面積順位のベクトルをあわせて、より高次元で識別性の高い特徴量ベクトル $(s_1, \dots, s_{15}, i_5, i_2, i_4, i_6, i_1, i_3)$ を得る。この特徴量ベクトルを用いた後続の処理は、非特許文献1と同じである。

【0053】

図4は、この実施形態に係る文書画像検索手法による検索質問Qと登録画像Dの各特徴点の対応関係を示す説明図である。これは、図1の例と同じものを上記の新しい特徴量ベクトルによって検索した処理の例である。図1と比べて対応する特徴点の数は大幅に減ったものの、正しい対応関係が求められていることがわかる。

40

【0054】

図5は、図4の検索質問Qと登録画像Dの画像同士で対応する部分を示した説明図である。図5に示すように、この実施形態の手法によれば、検索質問Qと登録画像Dの各特徴点が正しく対応づけられていることが分かる。

【0055】

実験例

提案手法の有効性を検証するため、非特許文献1の手法と提案手法を用いて日本語文書

50

検索実験を行った。文書画像データベースに収めた文書は、各種学会誌、論文誌より収集した日本語文書1万ページである。また、検索質問としては、1万ページの中から50ページを選び、角度60度（正面を90度としたときの角度）から撮影した画像50枚を用いた。画像サイズは1280万画素である。また、これを60%の大きさに縮小した画像も、検索質問として用いた。使用計算機は、CPUがAMD Opteron 2.4GHz、メモリが16GBのものである。

結果を表1に示す。以下、表に沿って考察する。

【0056】

【表1】

実験結果

		画像サイズ	
		100% (1280万画素)	60% (768万画素)
非特許文献1の手法	検索精度	92%	74%
	処理時間	512ms	421ms
提案手法	検索精度	98%	100%
	処理時間	205ms	175ms

【0057】

まず検索精度について述べる。非特許文献1の手法は、画像のサイズが十分大きいとき（100%）については、92%という検索精度を得ていたが、画像サイズが小さくなると精度が急激に悪化した。これは、小さい画像に対しては、特徴点だけでは識別性の十分高い特徴量が得られないことを示している。一方、提案手法は、両方の画像サイズにおいて、98%、100%という値を得ており、画像サイズが少なくとも800万画素程度までであれば、極めて高い精度を得られることがわかった。

【0058】

次に処理時間について述べる。ここで、処理時間とは、検索質問の画像1枚を処理するのに必要であった平均時間である。非特許文献1の手法が400～500ミリ秒程度必要であるのに対して、提案手法は半分以下の200ミリ秒程度で処理が完了している。この理由は、識別性の高い特徴量ベクトルを用いることによって、ハッシュにおける衝突が回避され、結果として処理の手間が省けたことによる。

以上を総合すると、提案手法は、日本語などの分かち書きされていない文書を高い精度で高速に検索する手法として、従来法に比べて優れたものであるといえる。

【0059】

実験例でも述べたように、この発明に係る文書画像検索方法は、CPUがプログラムに従って所定の処理手順を実行することにより実現することができる。また、例えば、画像データを体系的に管理して格納する機能を備えたデジタル複合機など、組み込み型の機器あるいはシステムとして実現され得る。そのような観点から、この発明の文書画像検索方法は、プログラムあるいは装置としての側面から捉えることも可能である。

【0060】

前述した実施の形態の他にも、この発明について種々の変形例があり得る。それらの変形例は、この発明の範囲に属しないと解されるべきものではない。この発明には、請求の範囲と均等の意味および前記範囲内でのすべての変形とが含まれるべきである。

【図面の簡単な説明】

【0061】

【図1】従来のLLAHを適用した文書画像検索手法による検索質問Qと登録画像Dの各特徴点の対応関係を示す説明図である。

【図2】この実施形態に係る特徴量計算の処理手順を示すフローチャートである。

【図3】この実施形態に係る文書画像検索処理において、不変量（面積比）と面積順位を用いた特徴量計算の具体例を示す説明図である。

【図4】この実施形態に係る文書画像検索手法による検索質問Qと登録画像Dの各特徴点

10

20

30

40

50

の対応関係を示す説明図である。

【図5】図4の検索質問Qと登録画像Dの画像同士で対応する部分を示した説明図である。

【図6】従来の文書画像検索システムを示すブロック図である。

【図7】従来の文書画像検索システムへの入力画像の一例を示す説明図である。

【図8】図7の入力画像が適応2値化処理されたものを示す説明図である。

【図9】図8の適応2値化画像をさらににじませた画像を示す説明図である。

【図10】図9の画像の連結成分の重心を特徴点とし、入力画像を特徴点の集合で示した説明図である。

【図11】従来の文書画像検索システムにおいて、特徴点pとその近傍の8つの特徴点が射影歪みを受けたときの影響を示す説明図である。

10

【図12】従来の文書画像検索システムにおいて、特徴点pの近傍n点からすべてのm点の組み合わせを作成する様子を示す説明図である。

【図13】従来の文書画像検索システムで登録文書の各特徴の特徴量がハッシュ表に登録される様子を示す説明図である。

【符号の説明】

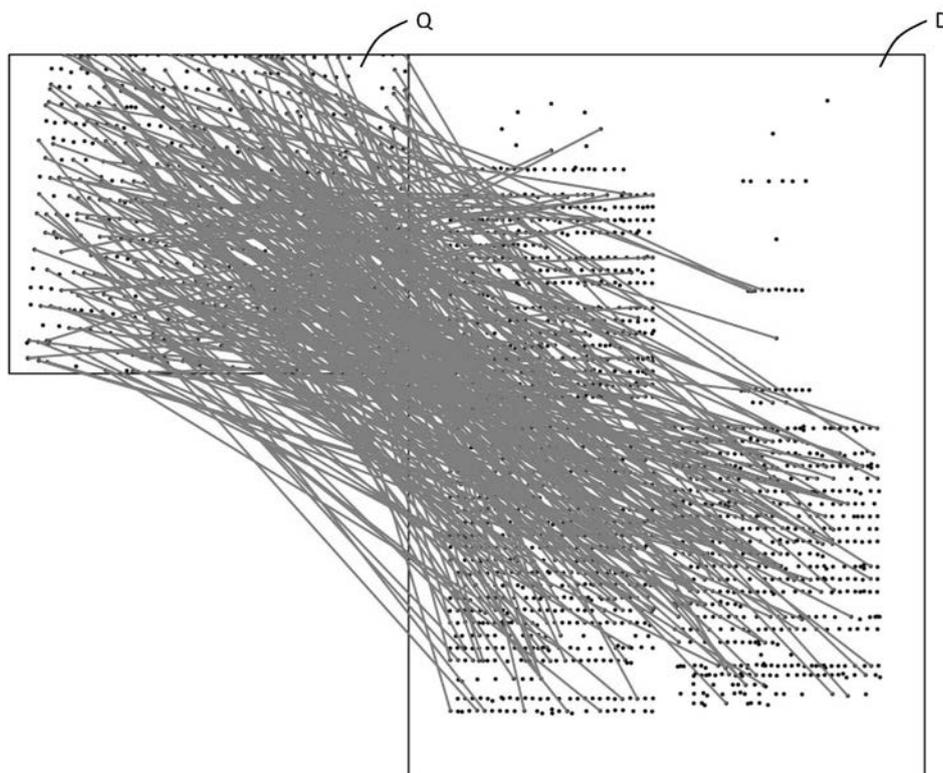
【0062】

Q：検索質問

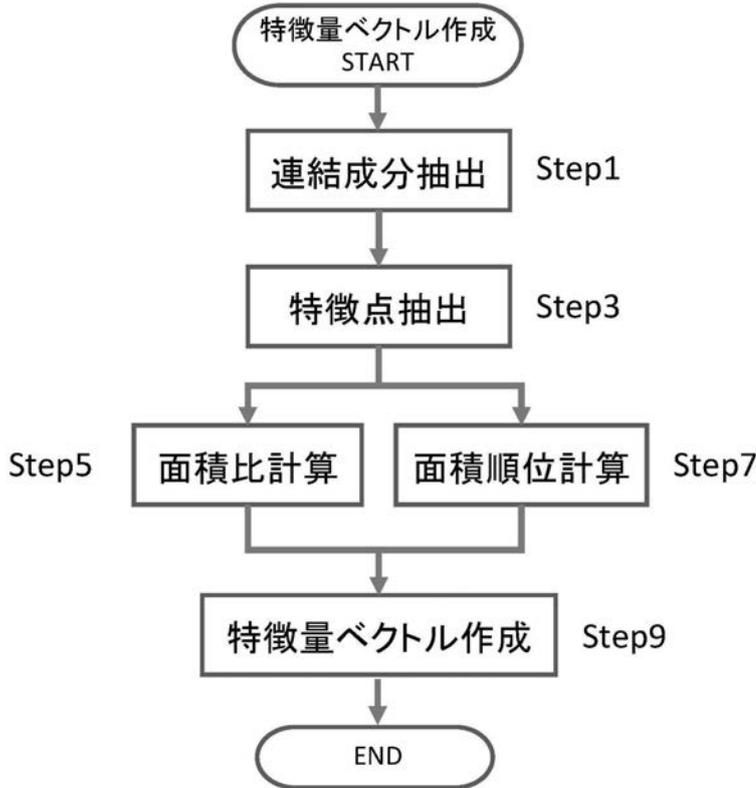
D：登録文書

【図1】

特徴点の配置のみを用いた場合の特徴点对応
(右側の点が登録文書の特徴点, 左側の点が検索質問の対応点)

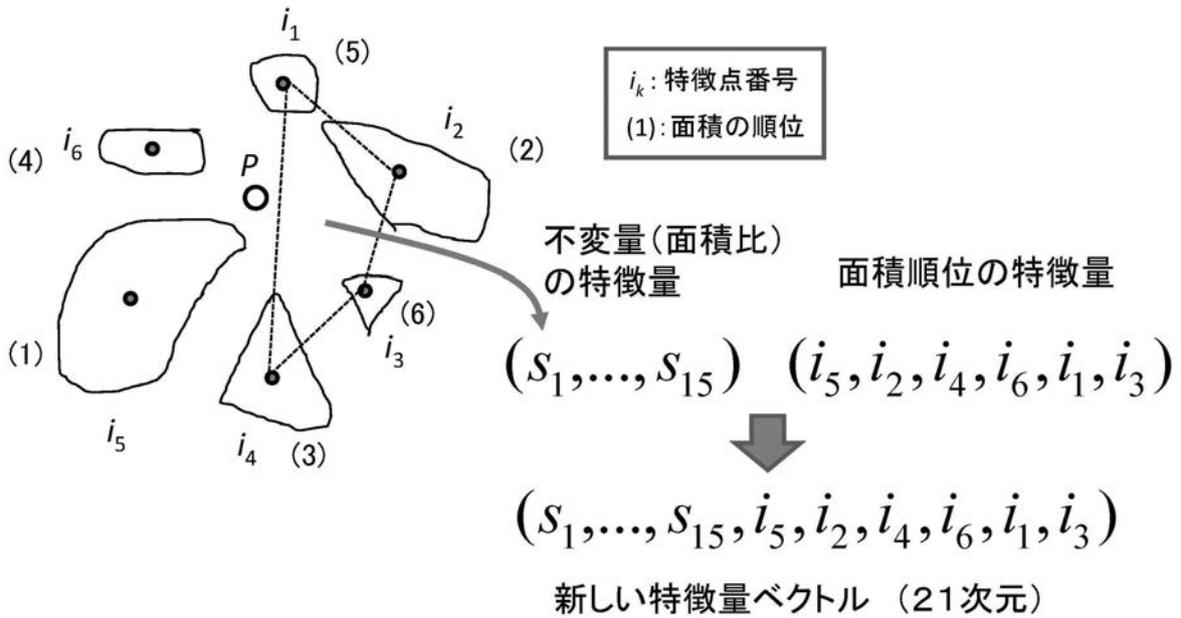


【 図 2 】



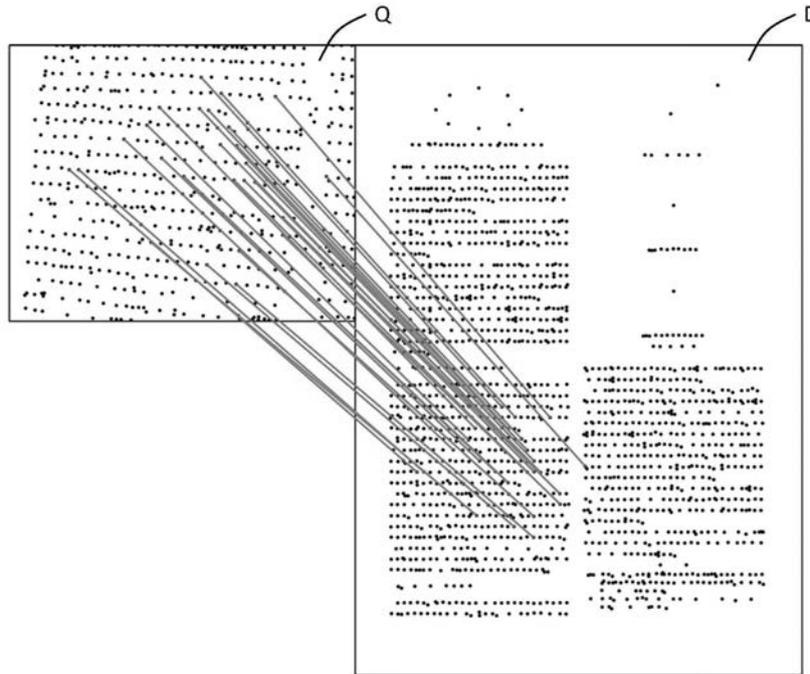
【 図 3 】

面積比と面積順位を用いた特微量ベクトル



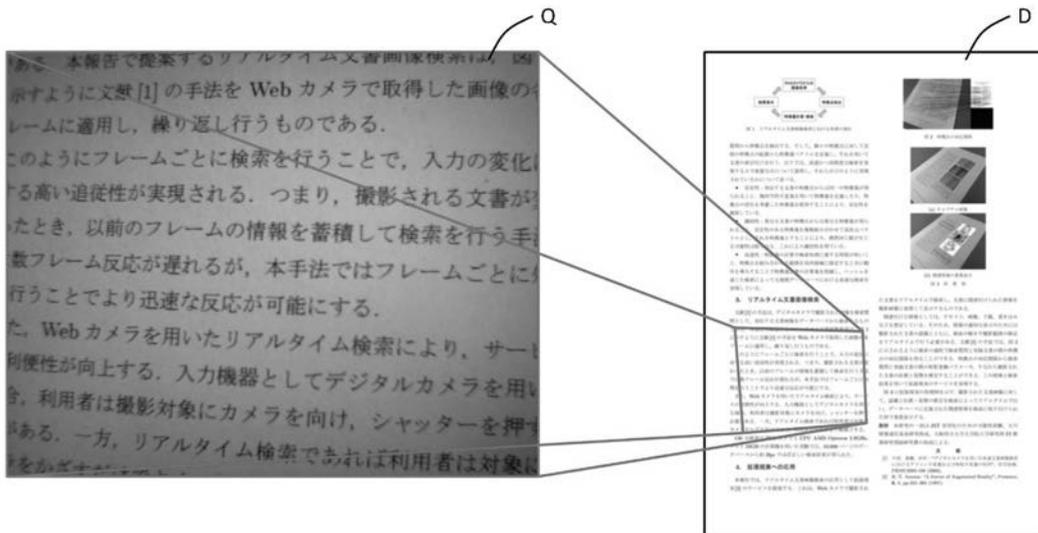
【図4】

連結成分の面積順位を加えた場合の特徴点对应
(右側の点が登録文書の特徴点, 左側の点が検索質問の対応点)

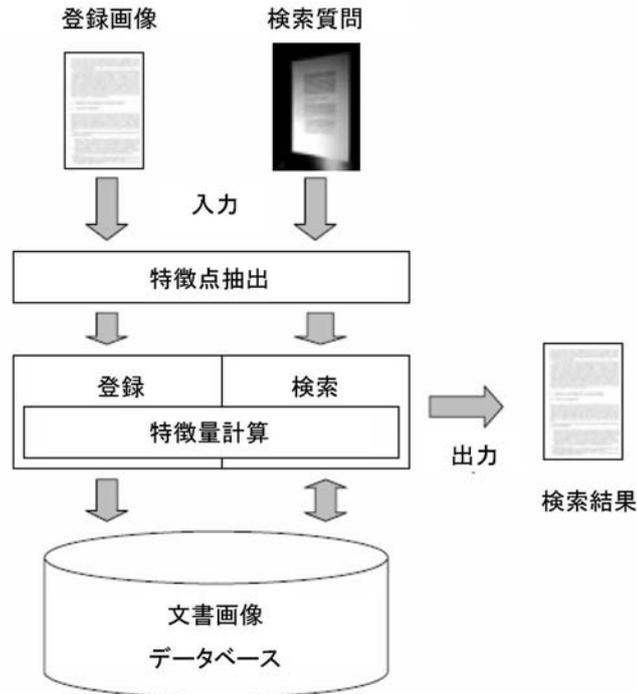


【図5】

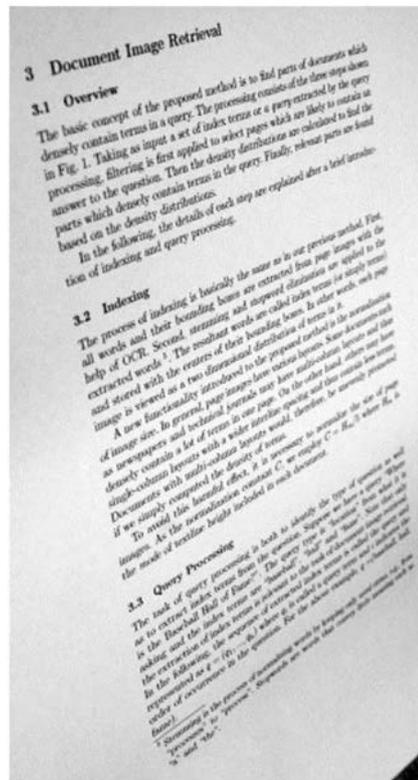
日本語文書の検索例



【 図 6 】



【 図 7 】



【 8 】

3 Document Image Retrieval

3.1 Overview

The basic concept of the proposed method is to find parts of documents which densely contain terms in a query. The processing consists of the three steps shown in Fig. 1. Taking as input a set of index terms or a query extracted by the query processing, filtering is first applied to select pages which are likely to contain answers to the question. Then the density distributions are calculated to find the parts which densely contain terms in the query. Finally, relevant parts are found based on the density distributions.

In the following, the details of each step are explained after a brief introduction of indexing and query processing.

3.2 Indexing

The process of indexing is basically the same as in our previous method. First, all words and their bounding boxes are extracted from page images with the help of OCR. Second, remaining and unpaired characters are applied to the extracted words. The remaining words are called *index terms* for simplicity and stored with a two-dimensional distribution of terms in a document. A new functionality introduced in the proposed method is the combination of newspapers and technical journals into one page. On the other hand, others are not stored as a list of terms in our page. The other hand, others are not stored as a list of terms in our page. The other hand, others are not stored as a list of terms in our page.

To avoid this harmful effect, it is necessary to normalize the size of page images. As the normalization constant C , we take $C = \frac{1}{\sqrt{A}}$, where A is the area of the image. As the normalization constant C , we take $C = \frac{1}{\sqrt{A}}$, where A is the area of the image.

3.3 Query Processing

The task of query processing is both to identify the type of queries and to extract index terms from the queries. Suppose that a query is given in the form of a list of index terms. The query type is identified by the number of index terms. The number of index terms is used to identify the type of queries. The number of index terms is used to identify the type of queries.

【 9 】

3 Document Image Retrieval

3.1 Overview

The basic concept of the proposed method is to find parts of documents which densely contain terms in a query. The processing consists of the three steps shown in Fig. 1. Taking as input a set of index terms or a query extracted by the query processing, filtering is first applied to select pages which are likely to contain answers to the question. Then the density distributions are calculated to find the parts which densely contain terms in the query. Finally, relevant parts are found based on the density distributions.

In the following, the details of each step are explained after a brief introduction of indexing and query processing.

3.2 Indexing

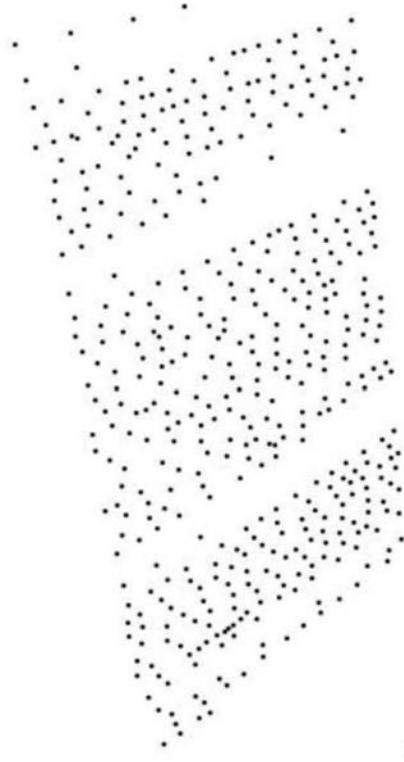
The process of indexing is basically the same as in our previous method. First, all words and their bounding boxes are extracted from page images with the help of OCR. Second, remaining and unpaired characters are applied to the extracted words. The remaining words are called *index terms* for simplicity and stored with a two-dimensional distribution of terms in a document. A new functionality introduced in the proposed method is the combination of newspapers and technical journals into one page. On the other hand, others are not stored as a list of terms in our page. The other hand, others are not stored as a list of terms in our page.

To avoid this harmful effect, it is necessary to normalize the size of page images. As the normalization constant C , we take $C = \frac{1}{\sqrt{A}}$, where A is the area of the image. As the normalization constant C , we take $C = \frac{1}{\sqrt{A}}$, where A is the area of the image.

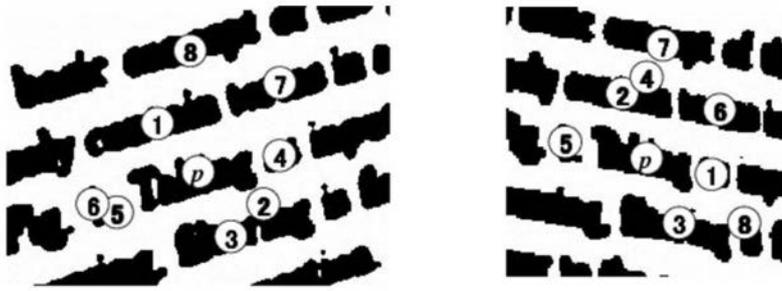
3.3 Query Processing

The task of query processing is both to identify the type of queries and to extract index terms from the queries. Suppose that a query is given in the form of a list of index terms. The query type is identified by the number of index terms. The number of index terms is used to identify the type of queries. The number of index terms is used to identify the type of queries.

【 図 1 0 】



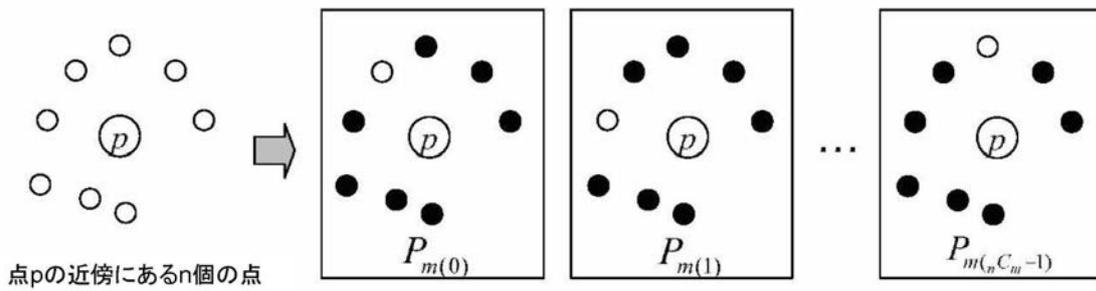
【 図 1 1 】



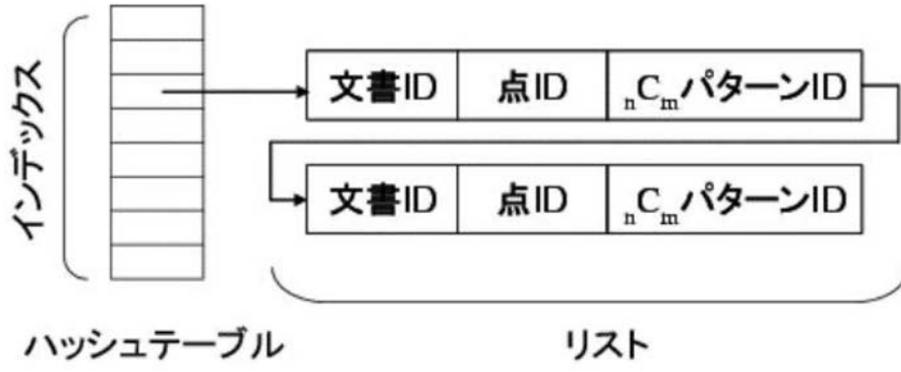
(a)

(b)

【 図 1 2 】



【図13】



フロントページの続き

(72)発明者 岩村 雅一
大阪府堺市中区学園町1 - 1 公立大学法人大阪府立大学内

審査官 岡本 俊威

(56)参考文献 特開2008 - 257670 (JP, A)
国際公開第2006 / 092957 (WO, A1)

(58)調査した分野(Int.Cl., DB名)

G06T 1 / 00
G06T 7 / 00 - 7 / 60
G06F 17 / 30