

LLAH によるリアルタイム文書画像検索の日本語文書への拡張

中居 友弘[†] 黄瀬 浩一[†] 岩村 雅一[†]

[†] 大阪府立大学工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

E-mail: †nakai@m.cs.osakafu-u.ac.jp, ††{kise,masa}@cs.osakafu-u.ac.jp

あらまし 著者らは既に、英文文書を対象としたリアルタイム文書画像検索法を提案している。この手法は、LLAH と呼ばれる高速かつ高精度な画像認識手法を用いることで、大規模データベースからの文書画像検索をリアルタイムで実現したものである。しかし、LLAH は特徴点の配置のみを手掛かりとして検索を行うため、その精度が特徴点の識別性・安定性に大きく依存するという問題がある。そのため、この手法の対象は単語の重心という特徴的かつ安定な特徴点只得られる英文文書に限られていた。本稿では、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案する。提案手法により、従来困難であった日本語文書の検索が実現できる。発表では、日本語文書を対象としたリアルタイム文書画像検索のデモを行う。

キーワード 文書画像検索, リアルタイム処理, Web カメラ

Extension of Real-Time Document Image Retrieval with LLAH for Japanese Documents

Tomohiro NAKAI[†], Koichi KISE[†], and Masakazu IWAMURA[†]

[†] Graduate School of Engineering, Osaka Prefecture University

1-1 Gakuencho, Naka, Sakai, Osaka, 599-8531, Japan

E-mail: †nakai@m.cs.osakafu-u.ac.jp, ††{kise,masa}@cs.osakafu-u.ac.jp

Abstract We have already proposed a real-time retrieval method for English documents. It is realized with a fast and accurate image recognition method called LLAH. Since LLAH exploits only arrangements of feature points, its accuracy depends heavily on discrimination power and stability of feature points extracted from images. Therefore the method has been applicable to English documents which give centroids of word regions as discriminative and stable feature points. In this paper, we propose an improved LLAH which enables to retrieve images with less discriminative and less stable feature points such as those extracted from Japanese documents. In the presentation, we show a demonstration of real-time retrieval of Japanese documents.

Key words Document image retrieval, Real-time processing, Web camera

1. はじめに

我々は Web カメラを用いたリアルタイム文書画像検索を提案している [1]。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を見つけて提示するものである。

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。

我々の提案しているリアルタイム文書画像検索は、LLAH と呼ばれる画像認識手法 [2] 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ロバストであるという点が挙げられる。また、特

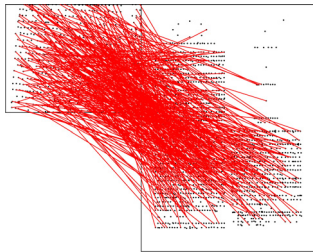


図 1 特徴点の配置のみを用いた場合の特徴点对応（右側の点が登録文書の特徴点，左側の点が検索質問の対応点）

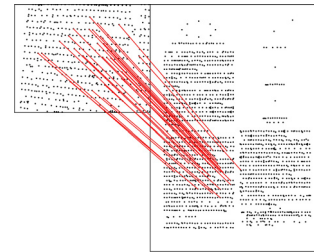


図 2 連結成分の面積を加えた場合の特徴点对応（右側の点が登録文書の特徴点，左側の点が検索質問の対応点）

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。

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問題もある。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 の適用対象はこれまでのところ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英文文書に限られてきた。

本発表では、LLAH の特徴量計算処理を改良し、これまで困難であった日本語文書についても検索可能な手法を提案する。これは、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を確認した。

2. 提案手法

本稿では、日本語文書を対象としたリアルタイム文書画像検索法を提案する。これは、[1] で提案された英語文書のリアルタイム文書画像検索法を日本語文書へ拡張するものである。

日本語文書の検索を実現する上では、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、[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。

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。図 1 は、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 Web 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。このように、特徴量の識別性が不足しているため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。

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、

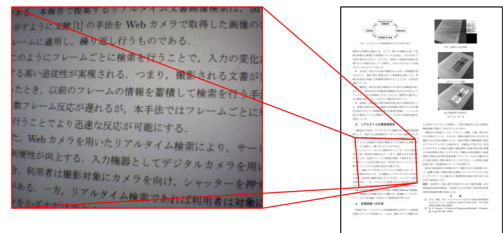


図 3 日本語文書のリアルタイム検索

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。

提案手法では、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。具体的には、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。図 2 は、新しい特徴量を用いた場合の対応関係である。このように、識別性が向上したために誤った対応が除かれて正しい対応関係のみが得られている。また、Web カメラを用いたリアルタイム検索の実験を行った結果、図 3 に示すように日本語文書の検索が可能であることが示された。

なお、本稿で述べたシステムのサンプルプログラムは [3] で公開されている。

謝辞 本研究の一部は日本学術振興会科学研究費補助金(基盤研究(B)19300062, 特別研究員奨励費 19-7621)の補助による。

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”, 信学技報, PRMU2006-66, pp.41-48 (2006).
- [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”, 信学論 (D), J89-D, 9, pp.2045-2054 (2006).
- [3] <http://imlab.jp/LLAH/> .