

Web カメラを用いたリアルタイム文書画像検索とその拡張現実への応用

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 本報告で我々はリアルタイム文書画像検索の手法を提案する。これは、Web カメラで撮影された文書に対応する文書画像をデータベースからリアルタイムで検索する手法である。本手法の特長は、大規模データベースにおけるリアルタイム検索が可能で、高い追従性である。これらの特長は我々のすでに提案した高速文書画像検索法を検索処理に用いることで実現されている。実験では、10,000 ページのデータベースから約 3fps で検索が可能であることを示した。また、リアルタイム検索の応用として、文書に対する拡張現実のサービスについても提案する。これは、撮影された文書へ関連情報を、文書の位置・姿勢に適合した形で、リアルタイムで重畳表示するものである。

キーワード 文書画像検索, リアルタイム処理, Web カメラ, 拡張現実

Real-Time Document Image Retrieval using a Web Camera and Its Application to Augmented Reality

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 In this report, we propose a method of real-time document image retrieval that is a task of retrieving a document image corresponding to a query image captured using a web camera. This method enables real-time retrieval in a large scale database with high following capability realized by the use of our fast document image retrieval method. From experimental results we have shown that the method can retrieve a document from a database with 10,000 pages at about 3 fps. We also propose an implementation of augmented reality of documents in which relevant information is overlaid onto the captured document image.

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

1. はじめに

文書画像検索とは、データベースから利用者の与えた検索質問に対応する文書画像を検索する処理である。我々はすでに、デジタルカメラで撮影した文書画像を検索質問として、データベースから対応する文書を検索する文書画像検索法を提案している [1]。この手法の特長として、取り扱いの容易なデジタルカメラの利用を可能にした点が挙げられる。従来手法の多くは画像の幾何歪みに対するロバスト性をもたなかったため、画像の取り込みの手間の大きいスキャナなどを入力機器としてきた。我々の提案した手法では、幾何学的不変量を利用することで画像の幾何歪みに対処している。また、この手法は大規模データベースにおいても極めて高速な検索が可能（画像処理を除いた検索時間は、10,000 ページのデータベースで検索質問 1 枚あた

り約 0.1 秒）である。これは、ハッシュを利用した検索によって実現されている。

本報告では、文献 [1] の手法の高速性を利用して、Web カメラを用いたリアルタイム検索を提案する。これは、Web カメラから得られるフレームを検索質問として、対応する文書画像を検索するという処理をフレームごとに繰り返すものである。

また、リアルタイム検索を用いたサービスとして、文書への拡張現実を提案する。これは、検索の過程で得られる検索質問と登録文書の特徴点の対応関係を用いて撮影された文書の位置と姿勢を推定し、関連情報を重畳表示するものである。

2. 特徴点の局所的配置に基づく高速文書画像検索法

文献 [1] の手法について説明する。この手法では、まず検索

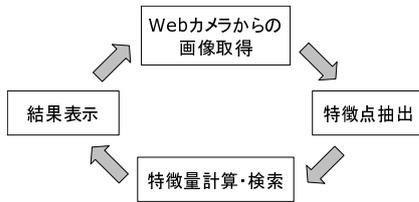


図 1 リアルタイム文書画像検索における処理の流れ

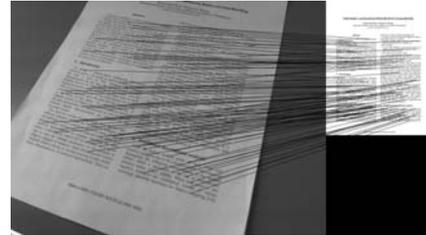


図 2 特徴点の対応関係

質問から特徴点を抽出する．そして、個々の特徴点に対して近傍の特徴点の配置から特徴量ベクトルを定義し、それを用いて文書の索引付けを行う．以下では、高速かつ高精度な検索を実現する上で重要な点について説明し、それらがどのように実現されているかについて述べる．

- 安定性：対応する文書の特徴点からは同一の特徴量が得られること．幾何学的不変量を用いて特徴量を定義したり、特徴点の消失を考慮した特徴量を採用することにより、安定性を確保している．
- 識別性：異なる文書の特徴点からは異なる特徴量が得られること．安定性のある特徴量を複数組み合わせ高次元ベクトルとし、それを特徴量とすることにより、偶然同じ値が生じる可能性は低くなる．これにより識別性を得ている．
- 高速性：特徴量の計算や検索処理に要する時間が短いこと．特徴点を組み合わせる範囲を局所領域に限定すると共に順序を導入することで特徴量計算の計算量を削減し、ハッシュを通じた検索によって大規模データベースにおける高速な検索を実現している．

3. リアルタイム文書画像検索

文献[1]の手法は、デジタルカメラで撮影された画像を検索質問として、対応する文書画像をデータベースから検索するものである．本報告で提案するリアルタイム文書画像検索は、図1に示すように文献[1]の手法をWebカメラで取得した画像の各フレームに適用し、繰り返し行うものである．

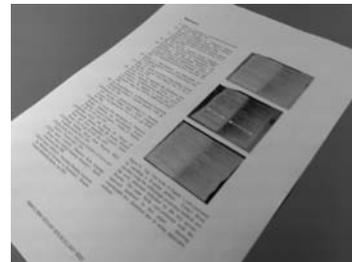
このようにフレームごとに検索を行うことで、入力の変化に対する高い追従性を実現される．つまり、撮影される文書が変わったとき、以前のフレームの情報を蓄積して検索を行う手法では数フレーム反応が遅れるが、本手法ではフレームごとに処理を行うことでより迅速な反応が可能にする．

また、Webカメラを用いたリアルタイム検索により、サービスの利便性が向上する．入力機器としてデジタルカメラを用いる場合、利用者は撮影対象にカメラを向け、シャッターを押す必要がある．一方、リアルタイム検索であれば利用者は対象にカメラをかざすだけでよく、利用者の負荷をより軽減できる．

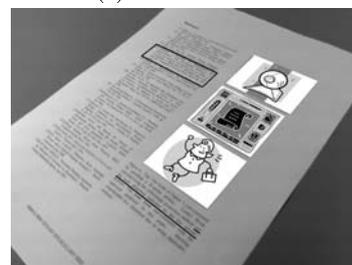
130万画素のWebカメラとCPU AMD Opteron 2.8GHz、メモリ16GBの計算機を用いた実験では、10,000ページのデータベースから約3fpsでほぼ正しい検索結果が得られた．

4. 拡張現実への応用

本報告では、リアルタイム文書画像検索の応用として拡張現実[2]のサービスを提案する．これは、Webカメラで撮影され



(a) キャプチャ画像



(b) 関連情報の重畳表示

図 3 処理例

た文書をリアルタイムで検索し、文書に関連付けられた情報を撮影画像に重畳して表示するものである．

関連付ける情報としては、テキスト、画像、下線、書き込みなどを想定している．そのため、情報の適切な表示のためには撮影された文書の認識とともに、紙面の傾きや撮影範囲の推定をリアルタイムで行う必要がある．文献[1]の手法では、図2に示されるように検索の過程で検索質問と登録文書間の特徴点の対応関係を得ることができる．特徴点の対応関係から検索質問と登録文書間の射影変換パラメータ、すなわち撮影された文書の位置と姿勢を推定することができる．この情報と検索結果を用いて拡張現実のサービスを実現する．

図3に拡張現実の処理例を示す．撮影された文書画像に対して、認識と位置・姿勢の推定を検索によってリアルタイムで行い、データベースに定義された関連情報を紙面に貼り付けられた形で重畳表示する．

謝辞 本研究の一部はJST実用化のための可能性試験、大川情報通信基金研究助成、大阪府立大学大学院工学研究科FI推進研究奨励研究費の助成による．

文 献

- [1] 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”, 信学技報, PRMU2005-188 (2006).
- [2] R. T. Azuma: “A Surfer of Augmented Reality”, Presence, 6, 4, pp.355-385 (1997).