

# 多数の特徴点によって記述されるパターンの高速認識法

A Method of Fast Recognition of Patterns Represented by Many Feature Points

中居友弘 †, 黄瀬浩一 ‡, 岩村雅一 ‡

Tomohiro NAKAI†, Koichi KISE‡ and Masakazu IWAMURA‡

大阪府立大学大学院工学研究科

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

**概要** パターン認識の手法として、パターンから特徴点を抽出し、特徴点の配置に基づいて認識するものが提案されている。しかし、従来手法には、パターンに含まれる特徴点の数が多いと処理時間が膨大になるという問題がある。我々はすでに、特徴点の局所的配置を利用することで多数の特徴点で記述されるパターンを高速に認識する手法を提案している。また、この手法の高速性を利用した拡張として、英文文書を対象としたリアルタイム文書画像検索法についても提案している。本稿では、このパターン認識法の実用性を確認するために、様々な条件で Web カメラを用いたリアルタイム文書画像検索の実験を行った。その結果、認識対象の一部からでもほぼ 100% の検索が可能であり、低い解像度や紙面の湾曲に対するロバスト性をもつことが示された。

## 1 はじめに

看板、ポスター、文書、商品の箱など、我々の周囲にある多くの物体には文字などのパターンが記され、我々に情報を提供している。このようなパターンを誰でも簡単にデータベースに登録でき、バーコード認識のような手軽さで認識できれば、様々なサービスに利用可能であると考えられる。

パターン認識は、パターンから特徴量を抽出してその比較を行うことで実現される。画像の局所領域から、輝度変化に基づいて特徴量を求める手法として、SIFT [1] や PCA-SIFT [2] が提案されている。これらの手法で得られる特徴量は、画像の拡大縮小や回転、視点のずれにロバストであり、これらの特徴量を用いた認識法 [3] も提案されている。しかし、文書画像のように同じ部分パターンが繰り返し現れる対象では、同じような特徴量ばかりが得られる [4] ため、認識に用いることは容易ではない。

そこで、パターンから特徴点を抽出し、特徴点を組み合わせて特徴量を求める手法が考えられる。Geometric Hashing (GH) [5] は、特徴点を組み合わせて幾何学的変換に不変な特徴量を求め、認識を行う手法である。しかし、GH では射影歪みに不変なもの

するためには検索対象の特徴点数  $N$  に対して  $O(N^5)$  の計算量が必要となるため、処理速度の面で大きな問題がある [6]。

我々の提案した特徴点の局所的配置に基づくパターン認識法 [7] では、特徴点の局所的な配置に基づいて幾何学的不変量を利用した特徴量を求め、これをインデックスとして特徴点をハッシュ表に格納することで高速検索を実現する。この手法の特長として、特徴量計算の計算量が  $O(N)$  であるため、文書画像のように多くの特徴点をもつ画像についても高速に特徴量計算ができる点が挙げられる。

また、文献 [7] の手法に基づくリアルタイム文書画像検索法についても提案している [8]。この手法の特長は、(1) 解像度の低い Web カメラを入力機器とし、(2) 文書の撮影範囲が限定されていたり紙面の湾曲がある、という厳しい条件の下でも、大規模データベースからリアルタイムで精度の高い検索が実現できる点にある。この特長は、文献 [7] の手法が厳しい条件の下でも能力を失わない点に加え、リアルタイム検索のために新たに導入したパイプライン処理の効果によるものである。さらに、検索の副作用として特徴点の対応関係が得られるという特長もあ

る．これによって，ユーザは撮影された箇所を知ることができるほか，拡張現実などの応用処理をすることも可能となる．

本稿では，文献 [8] の手法の実用性を明らかにするためにさまざまな条件での検索実験を行った結果を報告する．実験により，対象文書の 1/4 程度の範囲が撮影されていればほぼ 100% での検索が可能であり，カメラの解像度が低い場合や紙面が湾曲した場合でも検索が可能であることが示された．

## 2 特徴点の局所的配置に基づくパターン認識法

文献 [7] で提案したパターン認識法について簡単に述べる．認識処理は，与えられたパターンから特徴量を求め，あらかじめ計算しておいたデータベースの特徴量と比較して対応するものを見つけることで行われる．特徴量は，パターンから得られる特徴点ごとに，その近傍点の配置を幾何学的不変量で記述することで求められる．つまり，特徴点一つひとつについて特徴量が得られる．多数の特徴量それぞれについて対応する特徴量を求め，対応する特徴量の数を得票数として，最も多くの得票数を得たパターンを認識結果とする．

この手法では，特徴点ごとに近傍点のみを用いて特徴量を求めるため，計算量は特徴点の数に比例したものとなり，特徴点の数が多くなっても計算量が爆発的に増大することはない．また，特徴量の比較にはハッシュ表を用いることにより，衝突のない場合は  $O(1)$  で比較が可能である．そのため，特徴量の比較の計算量は登録されているパターンの数に影響されにくい．

一方，特徴量の計算に局所的な点の配置を用いているため，与えられたパターンとデータベースに登録されているパターンの同一部分から同じ特徴点を得られないと，正しく検索を行うことができない．文献 [7] では対象を英文文書とし，ノイズ等の影響を受けても変化しにくい単語の重心を特徴点として用いている．また，幾何学的不変量の計算の際に対象が平面であることを仮定しているため，湾曲などの非射影的な変化があると得られる特徴量が変化する．ただし，局所的に平面に近似可能な程度の変化であれば対処できる．

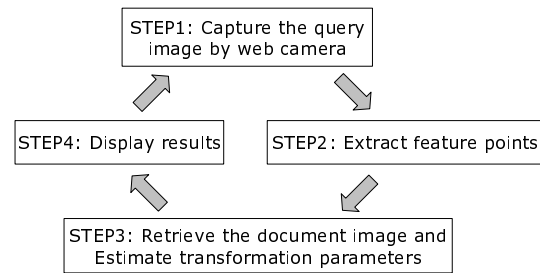


図 1: リアルタイム検索の処理の流れ



図 2: キャプチャ画像

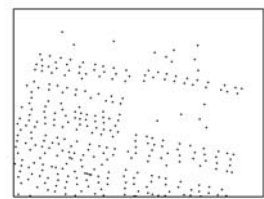


図 3: 特徴点抽出結果

## 3 リアルタイム文書画像検索

文献 [7] 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．文献 [8] で提案したリアルタイム文書画像検索は，文献 [7] 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，繰り返す行うものである．

以下では図 1 に示す処理の概要を述べる．まず Web カメラによって検索質問となる画像を取得し (STEP1)，特徴点を抽出する (STEP2)．Web カメラによるキャプチャ画像を図 2 に，特徴点抽出結果を図 3 に示す．特徴点の抽出は，キャプチャ画像を適応二値化し，ガウシアンフィルタを適用して再度二値化することで単語ごとに結合された画像を作成して，各連結成分の重心を求めることで行われる．

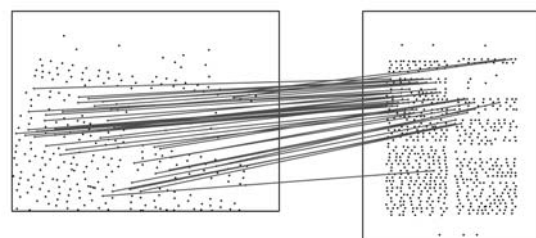


図 4: 検索結果と特徴点の対応関係

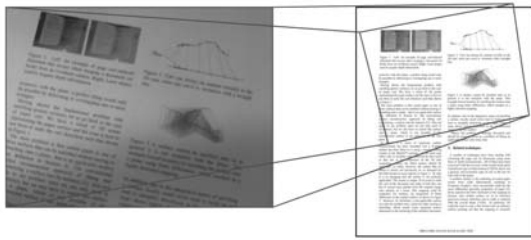


図 5: 射影変換パラメータの推定による撮影範囲の特定

次に、特徴点に基づいて文書画像検索を行う。これは、各特徴点について近傍点の配置から特徴量を計算し、あらかじめ計算しておいたデータベースの登録文書の特徴量から対応するものを見つけることで実現される。なお、特徴量の計算には文献 [9] のようにアフィン不変量を用いる。また、この処理では対応する登録文書の特定だけでなく、図 4 のような特徴点単位の対応関係の取得も同時に行う。そして、対応関係をもとに、RANSAC [10] を用いて撮影画像の射影変換パラメータを推定する (STEP3)。これにより、図 5 のように撮影範囲を特定することができる。最後にこれらの結果を表示する (STEP4)。

一連の処理が完了すると、再び Web カメラによる画像の取得に戻り、処理を繰り返す。

また、フレームレートを向上させるため、クライアントサーバシステムによるパイプライン処理を行う。STEP1, STEP2, STEP3 はクライアントが担当し、STEP4 はサーバが担当する。クライアントとサーバが並行して処理を行うことにより、単位時間当たりの処理フレーム数を増加させることができる。

## 4 実験

文献 [8] の手法の有効性を確認するため、Web カメラで英文文書を撮影した動画を入力として様々な条件で検索処理を行い、性能を調べた。

まず最初に、この手法を適用するためにはどの程度の性能の Web カメラが必要かを調べるため、カメラの解像度を変化させて性能を調べた。この手法では、英文文書における単語領域の重心を特徴点として用いるため、カメラの解像度が低いと、特徴点抽出処理で単語の結合に失敗し、精度が低下すると考えられる。



図 6: モーションブラーを受けたフレームの例

低解像度のカメラで特徴点の抽出を行うためには、撮影範囲を狭くし、拡大した状態で撮影する必要がある。この手法は特徴点の局所的配置を特徴量として用いているため、文書全体が撮影されていなくても検索が可能である。しかし、撮影範囲が小さくなると、画像から得られる特徴点の数が減少するため、ノイズ等の影響で正しい特徴量が得られないとき、精度に影響する可能性がある。そこで、検索を行うためにはどの程度の範囲が撮影される必要があるかを調べるため、撮影範囲を変化させて性能に対する影響を調べた。

また、書籍となっている文書を対象とする場合、書籍の厚みによる湾曲が問題となる。この手法では、紙面が平面であることを仮定して特徴量を計算するため、湾曲のある場合は正しく特徴量が求められない場合がある。そこで、紙面が湾曲している場合の性能についても調べた。

性能の評価のために、特徴点抽出および検索に要する処理時間と、検索および撮影範囲推定の精度を調べた。精度については、フレームを単位とし、撮影された文書や範囲が正しく検索された割合を用いた。なお、撮影範囲が正しく得られたかどうかについては、目視で判定した。また、精度については、すべてのフレームを対象とした精度と、強いモーションブラーのかかったフレームを除いた精度を求めた。これは、図 6 のような強いモーションブラーがかかった画像の場合、本手法では原理的に特徴点の抽出が不可能となり、検索に失敗するため、検索処理の性能をより正確に測るためにはこのようなフレームを除く必要があるためである。なお、実験に用いた動画は [11] で公開している。

実験では、クライアントとして CPU Intel Pentium4 3.0GHz、メモリ 1.5GB の計算機、サーバとし

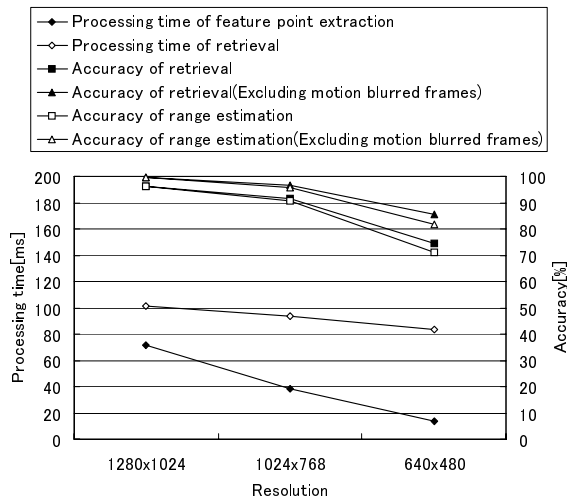


図 7: カメラの解像度と処理時間・精度の関係

て CPU Opteron 2.8GHz, メモリ 16GB の計算機を用いた。動画の撮影にはサンワサプライ社製 130 万画素 Web カメラ CMS-V20SETSV を用いた。データベースの登録ページ数は 10,000 とした。文献 [7] の検索手法におけるパラメータは  $n = 7, m = 6, k = 15, H_{size} = 2^{27} - 1$  とした。

#### 4.1 カメラの解像度と性能の関係

リアルタイム検索を実現するにはどの程度の解像度のカメラが必要かを調べるため、Web カメラの解像度を变化させたときの処理時間と精度を測定した。入力としては、文書のほぼ全体が撮影された、7fps、約 400 フレームの動画を用いた。

結果を図 7 に示す。まず、精度について述べる。解像度が低くなるに従って、精度は低下した。これは、解像度が低いと正確に単語の重心を求めることが困難になり、特徴点抽出の精度が低下するためと考えられる。ただし、この実験では撮影範囲を文書全体に固定していたため、撮影範囲を狭くすれば低解像度のカメラでも高い精度を実現できる可能性はある。

次に、処理時間について述べる。特徴点抽出および検索ともに解像度の低下に伴って処理時間は減少した。これは、特徴点抽出については処理する画像のサイズが小さくなったためであり、検索については得られる特徴点の数が減少したためであると考えられる。

表 1: 各撮影範囲の動画のフレーム数と時間

撮影範囲	フレーム数	時間 [s]
1/1	420	60
1/2	382	54
1/4	406	58
1/8	388	55
1/16	384	55

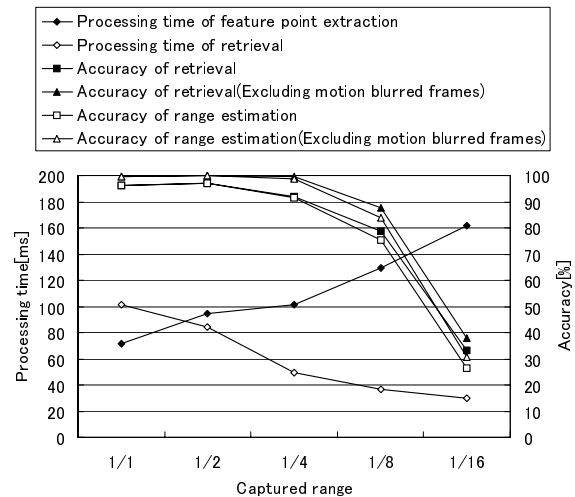


図 9: 撮影範囲と処理時間・精度の関係

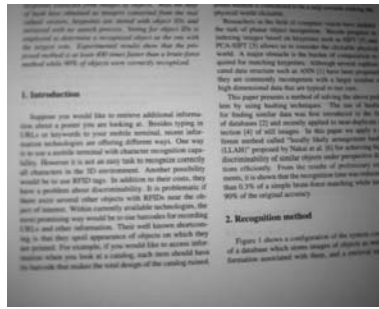
#### 4.2 撮影範囲のと性能の関係

この実験では、撮影範囲を 1/1 (文書全体), 1/2, 1/4, 1/8, 1/16 と変化させ、撮影範囲と処理時間および精度の関係を調べた。入力としては、130 万画素の Web カメラを用いて 7fps で撮影した動画を用いた。各動画のフレームの例を図 8 に示す。また、各動画のフレーム数と時間を表 1 に示す。なお、特徴点抽出の際に用いるガウシアンフィルタのマスクの大きさは、撮影範囲ごとに手動で適切に設定した。撮影範囲が狭いほど、マスクは大きいものとした。

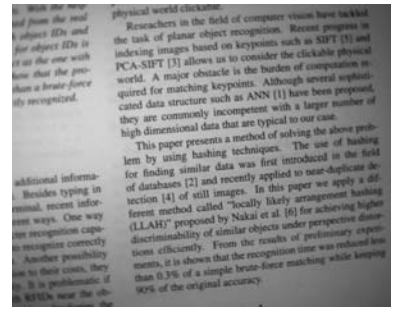
実験結果を図 9 に示す。まず、精度について述べる。撮影範囲が狭くなるに従って、検索および撮影範囲の推定精度は低下した。特に、撮影範囲が 1/4 の場合までは、モーションブラーを受けたフレームを除いたときの精度はほぼ 100% を維持していたが、撮影範囲が 1/8 になると急激に精度が低下した。これは、前述したように撮影範囲が狭いとフレーム画像に含まれる特徴点の数が小さくなるため、ノイズ



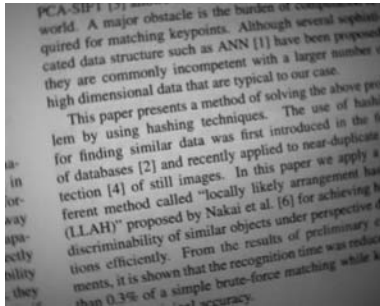
(a) 1/1



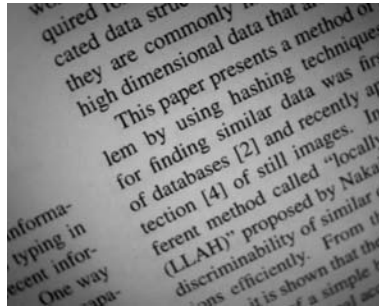
(b) 1/2



(c) 1/4



(d) 1/8



(e) 1/16

図 8: 各撮影範囲のフレームの例

等の影響で特徴量が正しく得られなかったときの影響が大きくなるためと考えられる。従って、高精度な検索を実現するためには 1/4 から 1/8 の範囲が撮影されている必要があるといえる。また、撮影範囲が 1/4 までは検索の精度と撮影範囲推定の精度はほぼ等しく、検索に成功したものについてはほぼすべてが撮影範囲の推定にも成功した。しかし、撮影範囲の狭いものでは検索の精度と撮影範囲推定の精度に差が現れた。これは、検索の過程で得られる点対応の数が小さくなるためと考えられる。

次に、処理時間について述べる。撮影範囲が狭くなるに従って、特徴点抽出の処理時間は増大した。これは、特徴点抽出処理において用られるガウシアンフィルタのマスクが大きくなるためである。ただし、この問題は、撮影範囲が狭い場合には画像をサンプリングして縮小するなどの処理を加えることで回避できる可能性がある。一方、検索の処理時間は撮影範囲が狭くなるに従って減少した。これは、フレーム画像に含まれる特徴点の数が少なくなったためである。検索質問画像から得られる各特徴点についてハッシュ表からの検索を行うため、検索の処理時間は検索質問画像に含まれる特徴点の数に比例したものになる。



図 10: 湾曲した紙面のフレームの例

### 4.3 紙面の湾曲に対する耐性

厚みのある書籍の紙面をカメラで撮影する場合、湾曲による歪みが生じる。この実験では、図 10 に示すような湾曲した紙面を撮影した場合の性能を調べた。入力に用いた動画は、解像度 130 万画素、345 フレーム、7fps で文書のほぼ全体が写るように撮影したものである。

実験結果を 4.2 の湾曲がなく文書全体が撮影された場合のものと合わせて表 2 に示す。湾曲がない場合と比較すると、精度はやや低下した。これは、提案手法の特徴量が紙面の平面性を前提としているため、

表 2: 紙面の湾曲の有無と性能 (括弧内はモーションブラーを受けたフレームを除いた場合)

	精度 [%]		処理時間 [ms]	
	検索	撮影範囲の推定	特徴点抽出	検索
湾曲なし	96.2(99.8)	96.2(99.8)	71.2	101.4
湾曲あり	91.0(94.6)	88.7(92.2)	72.1	94.8

紙面が湾曲すると正しい特徴量を得ることが難しくなることが原因であると考えられる。ただし、モーションブラーを受けたフレームを除いた場合では、湾曲する紙面であっても、約 95% のフレームで検索に成功し、精度の低下は約 5% にとどまった。

## 5 まとめ

本稿では、特徴点の局所的配置を用いることで多数の特徴点をもつパターンを高速に認識できる手法と、この手法を利用したリアルタイム文書画像検索法について説明した。様々な条件での実験を行い、リアルタイム文書画像検索法は 1/4 程度の範囲が撮影されていればほぼ 100% の検索が可能であり、カメラの解像度が低い場合や紙面が湾曲した場合でも検索が可能であることが示された。

今後の課題としては、リアルタイム検索によって得られる撮影された文書と位置の情報を利用したサービスの考案、様々な条件下での口バラスト性の向上などが挙げられる。写真やポスターなどの、他の対象への拡張も重要な課題である。

## 参考文献

- [1] D. G. Lowe: Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, vol. 60, no. 2, pp.91-110, 2004.
- [2] Y. Ke and R. Sukthankar: Pca-sift: A more distinctive representation for local image descriptors, *CVPR2004*, vol. 2, pp.506-513, 2004.
- [3] 野口和人, 中居友弘, 黄瀬浩一, 岩村雅一: 特徴ベクトルの近傍探索と物体認識の効率に関する実験的検討, *信学技報*, PRMU2006-68, pp.57-64, 2006.
- [4] J. Liang, D. DeMenthon and D. Doermann: Camera-based document image mosaicing, *Proc. The 18th International Conference on Pattern Recognition*, vol. 2, pp.476-479, 2006.

- [5] H. J. Wolfson and I. Rigoutsos: Geometric hashing: an overview, *IEEE Computational Science & Engineering*, vol. 4, no. 4, pp.10-21, 1997.
- [6] 岩村雅一, 中居友弘, 黄瀬浩一: 特徴点の局所的配置に基づく物体認識の精度・速度に関する一考察, *信学技報*, PRMU2006-67, pp.49-56, 2006.
- [7] 中居友弘, 黄瀬浩一, 岩村雅一: 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)*, vol. J89-D, no. 9, pp.2045-2054, 2006.
- [8] 中居友弘, 黄瀬浩一, 岩村雅一: 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.
- [9] 中居友弘, 黄瀬浩一, 岩村雅一: 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, pp.25-30, 2006.
- [10] M. A. Fischler and R. C. Bolles: Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography, *Comm. ACM*, vol. 24, no. 6, pp.381-395, 1981.
- [11] <http://www.m.cs.osakafu-u.ac.jp/~nakai/demo2/>

中居友弘: 平 16 阪府大・工・情報卒。平 18 同大大学院修士課程了。現在, 同大学院博士課程に在籍。文書画像処理に関する研究に従事。FIT2005 ヤングリサーチャー賞, MIRU2006 デモセッション賞受賞。電子情報通信学会学生会員。  
URL: <http://www.m.cs.osakafu-u.ac.jp/~nakai/>  
黄瀬浩一: 昭 61 阪大・工・通信卒。昭 63 同大大学院博士前期課程了。同年同大大学院博士後期課程入学。平 2 阪府大・工・電気助手。現在, 同大大学院工学研究科教授。博士(工学)。平 12 年~13 ドイツ人工知能研究センター客員教授。文書画像解析, 情報検索などの研究に従事。情報処理学会, 人工知能学会, 電気学会, IEEE, ACM などの会員。  
岩村雅一: 平 10 東北大・工・通信卒。平 15 同大大学院博士課程了。同年同大大学院工学研究科助手。平 16 阪府大大学院工学研究科助手。現在に至る。博士(工学)。パターン認識, コンピュータビジョンに関する研究に従事。IEEE, 情報処理学会, Pattern Recognition Society 各会員。