

# 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

中居 友弘<sup>†</sup> 黄瀬 浩一<sup>†</sup> 岩村 雅一<sup>†</sup>

<sup>†</sup> 大阪府立大学大学院工学研究科

〒599-8531 大阪府堺市学園町 1-1

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

あらまし 我々はすでにデジタルカメラを用いた高速文書画像検索の一手法として特徴点の局所的配置に基づく手法を提案している．この手法では射影不変量である複比を用いているが，複比は特徴点の位置の変動に対して不安定であることが問題となっている．そこで，本稿ではより安定な性質をもつアフィン不変量および相似不変量の利用を考える．実験により，アフィン不変量を用いることで画像の歪みが大きい場合でも高精度な検索が可能であることが示された．また，相似不変量は，画像の歪みに対してロバストでないため，大きな射影歪みの生じる場合には適さないことが示された．

キーワード 文書画像検索，デジタルカメラ，アフィン不変量，相似不変量，ハッシュ，投票

## Use of Affine Invariants and Similarity Invariants in Fast Camera-based Document Image Retrieval

Tomohiro NAKAI<sup>†</sup>, Koichi KISE<sup>†</sup>, and Masakazu IWAMURA<sup>†</sup>

<sup>†</sup> Graduate School of Engineering, Osaka Prefecture University

1-1 Gakuen-cho, 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 method of camera-based document image retrieval which utilizes perspective invariants called cross-ratios. However, the cross-ratios have defects that they are unstable against changes of positions of feature points. In this paper, we evaluate the use of affine invariants and similarity invariants which are considered to be more stable. Experimental results have shown that higher accuracy is obtained with affine invariants even under perspective distortion. It is also shown that similarity invariants are not suitable for retrieval of largely distorted images since they are not robust to perspective distortion.

**Key words** Document image retrieval, Digital camera, Affine invariant, Similarity invariant, Hashing, Voting

### 1. はじめに

文書画像検索とは，利用者の与えた検索質問に対応する文書画像をデータベースから検索する処理である．本稿では特に，デジタルカメラで撮影された文書画像を検索質問として，データベースから同一の文書を検索する場合を考える．

この問題について，我々はすでに特徴点の局所的配置に基づく文書画像検索法[1], [2]を提案している．提案手法は，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．

提案手法の特長として，検索対象の特徴点数  $N$  に対して

$O(N)$  の計算量であるため，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検索できる点が挙げられる．Geometric Hashing (GH) [3] のような従来手法では，射影歪みに不変なものにするためには  $O(N^5)$  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．提案手法はこの問題を解決するものである．

しかし，複比には特徴点の位置が射影変換以外の要因（ノイズ等）で変動した場合，安定性に欠ける問題がある．複比の値は，同一平面上の5点から計算される面積の比の比で計算される．そのため，特徴点の位置の変動が微小であっても複比の値が大きく変動し，検索精度が低下するという問題が生じる．

そこで本稿では，提案手法へのアフィン不変量および相似不

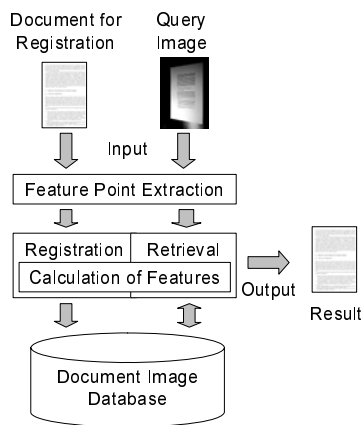


図 1 文書画像検索システム

変量の導入を検討する。これらの不変量は、特徴点の位置の変動に対して複比よりも安定であるため、精度の改善に有効であると考えられる。これらの不変量は射影変換に対して不変ではないが、特徴点の局所的配置に基づく提案手法では、画像全体が射影変換を受けていても局所的にはアフィン変換および相似変換に近似できる場合があるため、有効であると考えられる。実験により、アフィン不変量を用いることにより画像の歪みが大きい場合も高精度で検索できることが確認された。

## 2. 提案手法

### 2.1 処理の流れ

提案する文書画像検索システムの構成を図 1 に示す。登録文書 (document for registration) および検索質問 (query image) は、特徴点抽出 (feature point extraction) により点の集合に変換された後、登録の場合は登録処理 (registration) へ、検索の場合は検索処理 (retrieval) に入力される。これらの処理では、特徴点から不変量を用いた特徴量を求める特徴量計算 (calculation of features) に同じ処理を用いる。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 (document image database) 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。

### 2.2 特徴点抽出

本稿では、[1], [2] と同様に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。単語領域の重心は射影歪みやノイズの影響下でも比較的安定に抽出できるため、安定な特徴量の計算に有効である。

### 2.3 幾何学的不変量

提案手法では、同一平面上の  $f$  点から計算される幾何学的不変量を用いて特徴量を計算する。幾何学的不変量とは、幾何変換を通じて変化しない値であり、幾何変換の種類に応じて様々な幾何学的不変量が存在する。幾何学的不変量の計算に要する点の数  $f$  は不変量の種類によって異なる。

(1) 複比 … 複比は、射影変換の不変量であり、同一平面上の 5 点  $ABCDE$  ( $f = 5$ ) の座標から

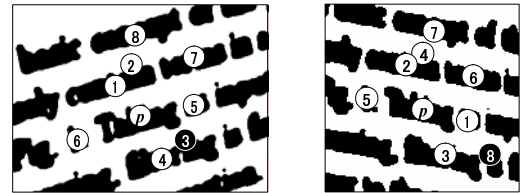


図 2 射影歪みによる特徴点  $p$  の近傍点の変化 (数字は点  $p$  との距離の順位)

$\{P(A, B, C)P(A, D, E)\} / \{P(A, B, D)P(A, C, E)\}$  で求められる [4]。ここで、 $P(A, B, C)$  は頂点  $A, B, C$  が成す三角形の面積である。複比は射影変換の不変量であるため、点  $ABCDE$  の座標が射影歪みによって変化しても、その値は同一に保たれる。

(2) アフィン不変量 … アフィン不変量は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、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、局所領域ではその変換がアフィン変換に近似されることがあるため、点の局所配置に基づく提案手法では複比の代わりにアフィン不変量を利用することが考えられる。アフィン不変量は同一平面上の 4 点  $ABCD$  ( $f = 4$ ) の座標から  $P(A, C, D) / P(A, B, C)$  で計算される [5]。

(3) 相似不変量 … 相似変換はアフィン変換よりもさらに自由度の低い変換であり、拡大・縮小、回転、並進のみからなる変換である。相似変換においては直線間の角度、距離の比、面積の比、距離の 2 乗と面積の比などが不変となる [5]。本稿では、3 点  $ABC$  ( $f = 3$ ) から  $AC / AB$  で計算される距離の比を用いる。

さて、画像上の特徴点から得られる不変量の値は連続値だが、インデックスに用いる際には離散化する必要がある。本手法では、予備実験により特徴点から得られる不変量のヒストグラムを作成し、各離散値における不変量の頻度が等しくなるように離散値を割り当て、 $k$  階調に量子化する。

### 2.4 特徴量計算

#### 2.4.1 特徴量の満たすべき条件

提案手法における特徴量とは、文書画像の特徴点を表現する量である。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。

特徴量の評価基準としては、検索精度や検索に要する計算量が挙げられる。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。

本稿では、特徴量の精度に関する条件として以下の 2 条件を定義する。第一の条件は、各種歪みの影響を受けても同じ文書の同じ点からは同じ特徴量が得られなければならないということである。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ないことはできない。この条件を「特徴量の安定性」と呼ぶ。第二の条件とは、異なる点からは異なる特徴量が得られなければならないということである。もし異なる文書から同じ特徴量が得られれば、

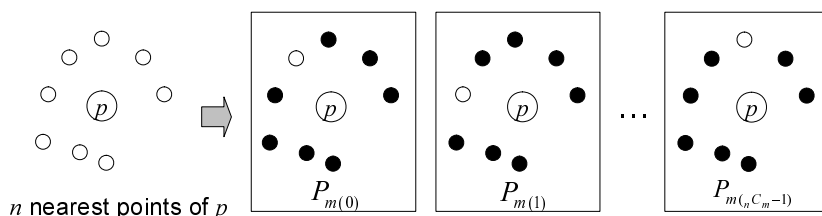


図3  $n(=8)$  点から  $m(=7)$  点を選ぶすべての組み合わせ

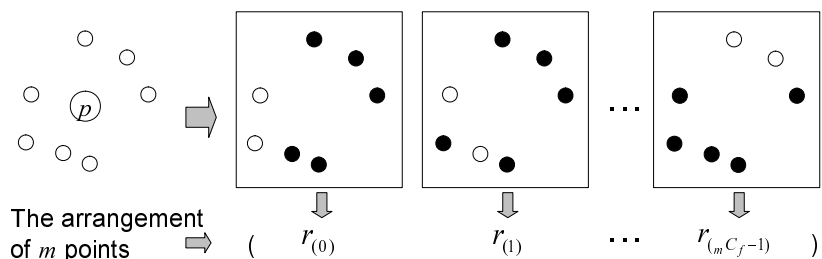


図4 すべての  $f$  点の組み合わせから得られる不変量で表現される  $m(=7)$  点の配置

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．この条件を「特徴量の識別性」と呼ぶ．

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、特徴量は精度に関する上記の2条件を満たすだけでなく、計算量が小さい必要もある．

これら3つの条件が高速かつ高精度な文書画像検索における特徴量の満たすべき条件である．

#### 2.4.2 特徴量の安定性

まず特徴量の安定性について述べる．前述したように、提案手法では各特徴点の近傍点から不変量を用いて特徴量を計算する．そのため、特徴量が安定であるためには、射影歪みによって近傍点の座標が変化しても特徴量の計算に用いる点に同じ点が必要である．図2に示されるように、近傍点は射影歪みの影響で変化するため、特徴点  $p$  の近傍  $f$  点から計算される不変量を特徴量とすると、同じ特徴点  $p$  から同じ特徴量を得ることはできない．

そこで、提案手法ではより広い範囲の近傍点から複数の点の組み合わせを作成し、そこから複数の特徴量を計算する．これは、図2のように、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2では8点）のうち  $m$  点（図2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．近傍  $n$  のうち  $m$  点在同一であるならば、図3のように  $n$  点からすべての  $m$  点の組み合わせ  $P_{m(0)}, P_{m(1)}, \dots, P_{m(C_m-1)}$  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．

#### 2.4.3 特徴量の識別性

次に、特徴量の識別性について述べる．提案手法では、1つの特徴量の計算に用いる特徴点の数  $m$  を増やすことで識別性を高める． $m$  点の配置の表現方法としては、図4のように  $m$  点から得られるすべての  $f$  点の組み合わせから計算される不変量の列  $r_{(0)}, r_{(1)}, \dots, r_{(mC_f-1)}$  を用いる． $m$  が大きければ大き

いほど、計算される不変量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．

ただし、 $m$  が大きすぎると特徴量の安定性が低下する．なぜなら、特徴量が一致するためには特徴量の不変量すべてが一致する必要があるが、 $m$  が大きいと不変量の数  $mC_f$  が増加するので、誤差の影響で異なる不変量が計算される可能性が高くなるためである．

#### 2.4.4 計算量と記憶容量

以上に述べたように、 $n$  を大きくすると広い範囲から複数の特徴量を計算するようになり、特徴量の安定性が向上する．また、 $m$  を大きくすると1つの特徴量の計算に用いる点の数が増加するため、特徴量の識別性が向上する．しかし、これらのパラメータを大きくしすぎると計算量の面で問題が生じる． $n$  および  $m$  が大きすぎると、不変量の計算回数が増加する．そのため、登録および検索に要する処理時間が大きくなる．また、登録時には計算される特徴量を記憶しておく必要があるため、大きな記憶容量が必要となる．

#### 2.4.5 不変量の量子化レベル

安定性および識別性に影響するパラメータは  $n, m$  だけではない．不変量の量子化レベル  $k$  もこれらに影響する． $k$  の値が大きい（不変量が細かく離散化される）と、誤差の影響で同じ  $f$  点から計算される不変量が異なる値に離散化される可能性が高くなり、安定性が低下する． $k$  の値が小さい（不変量が粗く離散化される）と、異なる  $f$  点から計算された不変量が同じ値に離散化される可能性が高くなり、識別性が低下する．

以上のことから、高い検索精度と同時に高速な処理や小さい記憶容量を実現するには、 $n, m$  および  $k$  を適切に設定する必要があるといえる．

### 2.5 登録

以上の準備に基づき、図5に示す登録アルゴリズムについて説明する．登録処理では、登録文書の各特徴点について、その近傍  $n$  点からすべての  $m$  点を選ぶ組み合わせを生成し、そこ

```

1: for each  $p \in \{ \text{特徴点} \}$  do
2:    $P_n \leftarrow p$  の近傍  $n$  点の集合
3:   for each  $P_m \in \{ P_n \text{ から } m \text{ 点を選ぶすべての組み合わせ} \}$  do
4:      $P_m$  から適当に 1 点  $p_0$  を選び,  $P_m$  の要素を  $p$  を中心として  $p_0$  を先頭に時計回りに並べて特徴点の列  $L_m = (p_0, \dots, p_{m-1})$  を作る. 以後,  $p_i$  の順序とは添字の若い順のことをいう
5:      $(L_f(0), \dots, L_f(i), \dots, L_f(mC_f - 1)) \leftarrow L_m$  から  $f$  点を選んではいる部分列  $L_f$  をすべて求めて辞書式順序に並べたリスト
6:     for  $i = 0$  to  $mC_f - 1$  do
7:        $r_{(i)} \leftarrow L_f(i)$  から計算される不変量の離散値
8:     end for
9:      $H_{\text{index}} \leftarrow$  式 (1) で計算されるハッシュ表のインデックス
10:     $H_{\text{index}}$  を用いて文書 ID, 点 ID, 不変量  $r_{(0)}, \dots, r_{(mC_f - 1)}$  をハッシュ表に登録
11:  end for
12: end for

```

図 5 登録アルゴリズム

から不変量を用いたインデックスを求めて図 6 に示すハッシュ表に登録する.

以下, アルゴリズムに沿って説明する. 1 行目で特徴点の集合から 1 点を取り出して  $p$  とし, 2 行目で  $p$  の近傍  $n$  点を取り出して  $P_n$  とする. 3 行目では  $P_n$  から  $m$  点を取り出して  $P_m$  とし, 4 行目で  $P_m$  の要素に対して  $p$  を中心とした時計回りの順序を与えて特徴点の列  $L_m$  を定める. 次に, 5 行目で  $L_m$  の要素から順序を保ったまま  $f$  点を選択してできる特徴点の列  $L_f$  をすべて求め, 辞書式順序に並べる. 例えば,  $m = 7, f = 5$  のときには,  $(L_f(0), \dots, L_f(\tau C_5 - 1))$  として  $((p_0, p_1, p_2, p_3, p_4), (p_0, p_1, p_2, p_3, p_5), \dots, (p_2, p_3, p_4, p_5, p_6))$  が得られる. さらに 7 行目において,  $L_f(i)$  の点を前から順に  $A, B, C, \dots$  として 2.3 で示した不変量の計算式に代入して不変量を求め, 離散化して  $r_{(i)}$  とする. このようにして得られた  $r_{(i)}$  より, 9 行目で以下の式 (1) に示されるハッシュ関数を用いてハッシュ表のインデックス  $H_{\text{index}}$  を求め, 10 行目で  $H_{\text{index}}$  を用いて登録文書の識別番号である文書 ID (Document ID) と点の識別番号である点 ID (Point ID), 不変量 (Invariants)  $r_{(i)} (i = 0, 1, \dots, mC_f - 1)$  をハッシュ表に登録する.

本稿で用いるハッシュ関数を以下に示す.

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

ここで,  $k$  は不変量の量子化レベル,  $H_{\text{size}}$  はハッシュ表のサイズである. また, 登録時に衝突が生じた場合, データは図 6 のようにリスト構造で付け加えられる.

このような処理をすべての  $p$  について施し, 文書を登録する.

## 2.6 検 索

次に検索について述べる. 図 7 に検索アルゴリズムを示す. 登録処理と同様に, 1~3 行目で,  $p, P_n, P_m$  を得る. 4~5 行目では, 登録処理とは異なり,  $P_m$  のすべての点を開始点  $p_0$  として  $L_m$  を作成する. これは, 図 5 の登録アルゴリズムの 4 行

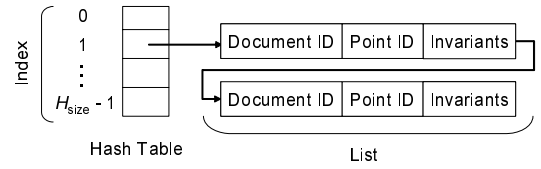


図 6 ハッシュ表の構造

```

1: for each  $p \in \{ \text{特徴点} \}$  do
2:    $P_n \leftarrow p$  の近傍  $n$  点の集合
3:   for each  $P_m \in \{ P_n \text{ から } m \text{ 点を選ぶすべての組み合わせ} \}$  do
4:     for each  $p_0 \in P_m$  do
5:        $P_m$  の要素を  $p$  を中心として  $p_0$  を先頭に時計回りに並べて特徴点の列  $L_m = (p_0, \dots, p_{m-1})$  を作る
6:        $(L_f(0), \dots, L_f(i), \dots, L_f(mC_f - 1)) \leftarrow L_m$  から  $f$  点を選んではいる部分列  $L_f$  をすべて求めて辞書式順序に並べたリスト
7:       for  $i = 0$  to  $mC_f - 1$  do
8:          $r_{(i)} \leftarrow L_f(i)$  から計算される不変量の離散値
9:       end for
10:       $H_{\text{index}} \leftarrow$  式 (1) で計算されるハッシュ表のインデックス
11:       $H_{\text{index}}$  を用いてハッシュ表にアクセスし, リストを得る
12:      for each リストの各要素について do
13:        if 誤対応を防止するための条件 (文献 [1]) が成立する then
14:          文書 ID に投票
15:        end if
16:      end for
17:    end for
18:  end for
19: end for
20: 得票数から式 (2) を用いて各文書のスコアを得る
21: 最大のスコアをもつ文書を検索結果とする

```

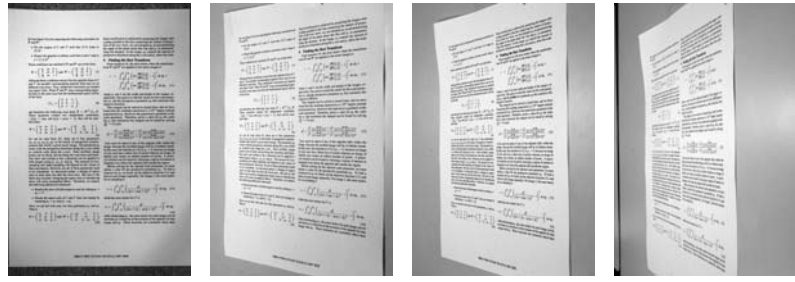
図 7 検索アルゴリズム

目において, 画像の回転を考慮せずに,  $P_m$  から得られる列として  $L_m$  を 1 つだけ登録しているために必要となる. 点  $p$  を中心とした特徴点の時計回りの順序は, 射影変換を受けた後でも, 開始点の任意性を残して一定である. すなわち,  $L_m$  の巡回置換を考えると, その中に必ず 1 つは登録時に用いた順序のものが存在するといえる. 6~10 行目で登録時と同様の手法によりハッシュ表のインデックスを求め, 11 行目でハッシュ表を読んで図 6 にあるようなリストを得る. 12~14 行目でリストの各要素について, 登録されている文書 ID のものに投票する. ここで, 誤対応による投票を制限するため, 文献 [1] と同じ 3 条件を用いる. 簡単には, (1) 不変量の列が一致する, (2) 検索質問の 1 点 が 1 つの登録文書の複数点に対応しない, (3) 登録文書の 1 点は検索質問の複数点と対応しない, である.

このような処理をすべての点について行くと, 各登録文書の得票数が得られる. しかし, このようにして得られた得票数は依然として誤対応によるものを含んでいる. 誤った投票は登録文書に含まれる特徴点数に概ね比例するため, 多くの特徴点をもつ登録文書は不当に多くの得票数を得ることになる. このような誤った投票を補正するため, 以下の式のように文書  $d_i$  に



図 8 データベース中の登録文書の例



(a) 撮影角度 90° (b) 撮影角度 60° (c) 撮影角度 45° (d) 撮影角度 30°

図 9 検索質問の例

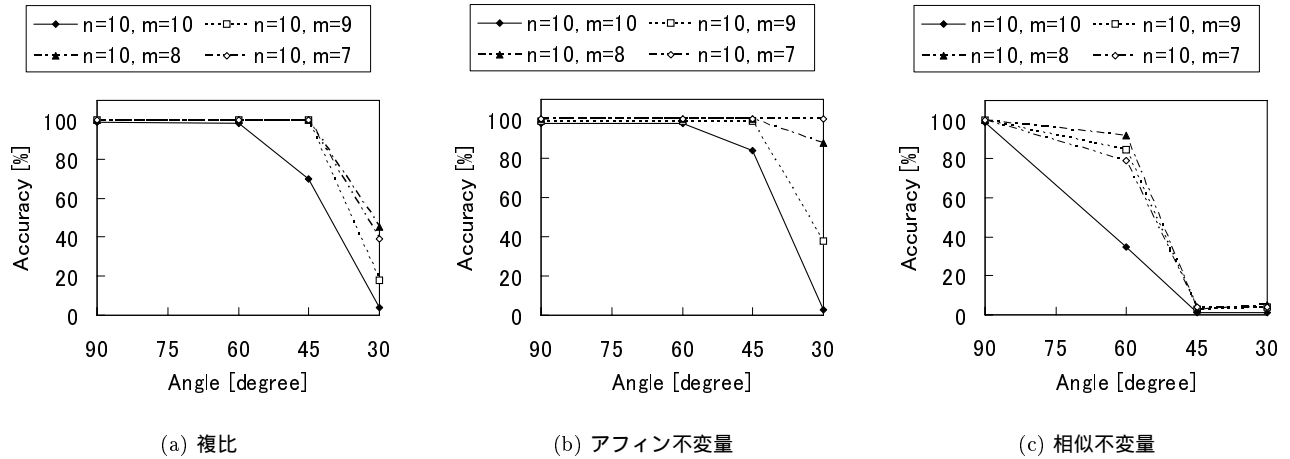


図 10 さまざまなパラメータでの撮影角度と検索精度の関係

対してスコア  $S(d_i)$  を定義する .

$$S(d_i) = V(d_i) - cN(d_i) \quad (2)$$

ここで,  $V(d_i)$  は  $d_i$  の得票数,  $N(d_i)$  は  $d_i$  に含まれる特徴点の数,  $c$  は予備実験で定められる特徴点数と誤投票の比例定数である . 最終的に, 最大のスコアを得た文書を検索結果とする .

### 3. 実験

#### 3.1 実験概要

提案手法におけるアフィン不変量および相似不変量の有効性を検証するため, 検索精度および処理時間の比較実験を行った . 実験では, 図 8 に示すような英語論文の電子文書を変換して得られた文書画像のデータベースと, 印刷文書をデジタルカメラで撮影した検索質問を用いた . 文書画像データベースとしては, それぞれ 100, 1,000, 10,000 ページの文書画像からなるデータベース A, B, C を用いた . データベース C は B を, B は A をその一部として含んでいる . また, 電子文書としては主に CVPR, ICPR, ICCV などの類似したレイアウトをもつ国際会議の予稿集のものを用いた . 検索質問としては, データベース A を印刷したものを紙面に対してそれぞれ 90°, 60°, 45°, 30° 程度の角度から撮影したものをを用いた . 画像の数はそれぞれ 100 枚である . 検索質問の例を図 9 に示す . 検索質問の作成には, CANON EOS Kiss Digital (630 万画素) と付属のレンズ EF-S 18-55mm USM を使用した . 実験に用いたハッシュ表のサイズは  $H_{size} = 2^{27} - 1$  とした . また, CPU AMD Opteron 2.8GHz, メモリ 16GB の計算機を用いた .

#### 3.2 実験 1: 撮影角度と検索精度の関係

まず, 特徴量の計算に複比, アフィン不変量, 相似不変量を用いたときの検索質問の撮影角度と検索精度の関係を調べた . 提案手法では, 特徴量計算時の組み合わせ数を定めるパラメータ  $n$  および  $m$  や, 不変量の量子化レベル数  $k$  の値によって性能が大きく変化する . この実験では,  $n = 10$  とし,  $m = 10, 9, 8, 7$  のそれぞれについて最も高い精度を与えた  $k$  を用いた . また, 撮影角度 90° ~ 30° の検索質問と 100 ページのデータベース A を用いた . 図 10 に (a) 複比, (b) アフィン不変量, (c) 相似不変量のそれぞれについて各  $n, m$  の組み合わせにおける撮影角度と検索精度の関係を示す .

図 10(a)~(c) より, 角度が小さくなるにつれて全体的に精度が低下することがわかる . これは, 歪みによって特徴点の近傍構造が変化し, 近傍  $n$  点のうち  $m$  点の一致という条件が満たされなくなることが原因であると考えられる . ただし,  $n = 10, m = 7$  のように  $n$  と  $m$  の差が大きいときには精度の低下が比較的小さくなっている .

図 10(b) に示すアフィン不変量は角度の減少に対する精度の低下が最も小さく, 次が (a) の複比, 精度の低下の最も大きいものが (c) の相似不変量という結果になった . これは, 以下に述べる不変量の性質が原因であると考えられる .

まず, 複比は射影不変量であるため, 特徴点の位置の射影歪みによる変動に対しては不変である . しかし, 射影歪みでない変動に対しては不安定な性質をもっている . 提案手法では特徴点として単語領域の重心を用いており, これは画像が歪んだ場

表 1 各不変量における撮影角度ごとの登録ページ数と検索精度・処理時間の関係  
( $n = 8, m = 7$ )

登録ページ数	検索精度 [%]												処理時間 [msec]		
	複比				アフィン不変量				相似不変量				複比	アフィン不変量	相似不変量
	90°	60°	45°	30°	90°	60°	45°	30°	90°	60°	45°	30°			
100	99	99	98	12	99	99	99	46	100	73	3	3	80.8	109.3	118.3
1,000	99	99	97	3	99	99	99	32	99	48	1	2	83.0	109.5	136.6
10,000	98	98	72	0	98	98	98	8	98	28	0	0	97.4	119.6	332.3

合、一定の変動を受ける。そのため、大きな射影歪みを受けた場合、特徴点の座標の変動により特徴点から計算される複比の値が変化し、不安定となると考えられる。

アフィン不変量は、特徴点の位置の射影歪みによる変動に対しては不変ではない。しかし、画像全体が射影歪みを受けていても特徴量の計算の行われる局所領域における歪みはアフィン変換とみなすことができるため、アフィン不変量による特徴量は安定となる。また、アフィン不変量は複比に比べて特徴点の座標の変動に対して比較的安定であるため、アフィン変換に近似できる範囲の射影歪みであれば、複比よりも安定な特徴量を得ることができると考えられる。

相似不変量の場合もアフィン不変量と同様に、局所領域の歪みが相似変換に近似できる範囲では安定となる。しかし、相似変換は変換としての制約が強過ぎるため、射影歪みが大きいと局所領域でも相似変換に近似できなくなる。そのため、射影歪みが大きいと特徴量としての安定性が低くなる。

以上の不変量の性質より、射影歪みおよび特徴点の変動の双方に対して安定なアフィン不変量での精度が最も高いものになったと考えられる。

### 3.3 実験 2：登録ページ数と検索精度の関係

次に、不変量ごとに登録ページ数と検索精度の関係を調べた。 $n = 8, m = 7$  とし、 $k$  は登録ページ数 100 のときに最も高い精度を与えたものを用いた。表 1 に結果を示す。実験 1 と同様に、撮影角度の減少に伴って検索精度が低下した。また、登録ページ数が増加するに従って精度は低下した。これは、登録ページ数の増加に伴って類似した点の配置をもつ文書の登録される確率が増加することが原因と考えられる。実験 1 と同様に、アフィン不変量のときの精度が最も高く、次に複比、最も精度の低いものが相似不変量という結果になった。

### 3.4 実験 3：登録ページ数と処理時間の関係

各不変量での登録ページ数と処理時間の関係についても調べた。ここでの処理時間とは、1 枚の検索質問について図 7 に示される検索処理のみに要した時間であり、その前段階の特徴点抽出処理は含まない。パラメータは実験 2 と同じものを用いた。例として、検索質問に撮影角度 60° のものを用いたときの結果を表 1 に示す。なお、検索質問の撮影角度が変化しても処理時間はほぼ一定であった。

全体として、登録ページ数の増加に伴って処理時間は増大した。これは、ハッシュに登録されるデータが増加するため、衝突率が高くなることが原因であると考えられる。

処理時間は複比のときに最も短く、次にアフィン不変量、最

も長い処理時間を要したものが相似不変量という結果になった。これは、特徴量を求める際の不変量の計算回数  $mC_f$  の相違に起因すると考えられる。 $m = 7$  の場合、 $mC_f$  の値は  $f$  が小さいほど大きくなる。そのため、 $f = 5$  である複比に比べて、 $f = 4$  のアフィン不変量、 $f = 3$  の相似不変量の方が処理時間が長くなっていると考えられる。

また、相似不変量では登録ページ数 10,000 のときに急激に処理時間が増大している。これは、相似不変量はその計算に用いる点の数が少ないため、同じ離散値をとりやすく、ハッシュでの衝突が急激に増加したためと考えられる。

## 4. ま と め

本稿では、特徴点の局所的配置に基づく文書画像検索法に対して、アフィン不変量および相似不変量を導入し、その有効性を検証した。これは、射影歪みを受けた文書画像であっても局所的にはその変換がアフィン変換や相似変換に近似できる場合があるという考えに基づくものである。実験により、アフィン不変量を用いることで複比の場合よりも射影歪みにロバストな検索が可能になることが示された。これは、アフィン不変量のもつ特徴点の変動に対する安定性によるものと考えられる。相似不変量については、歪みに対する安定性が低く、大きな射影歪みの生じる場合には適さないことが示された。また、処理時間についても、アフィン不変量は複比よりもやや長い処理時間を要するものの、それほど大きな差はみられなかった。以上のことから、提案手法においては複比よりもアフィン不変量の方が有効であると考えられる。

今後の課題としては、精度の改善、日本語文書への対応、文書に限らないシーン中のオブジェクトの検索手法の考案が挙げられる。

## 文 献

- [1] T. Nakai, K. Kise and M. Iwamura: "Hashing with Local Combinations of Feature Points and Its Application to Camera-Based Document Image Retrieval —Retrieval in 0.14 Second from 10,000 Pages—", *Proc. CBDAR'05*, pp.87–94 (2005).
- [2] 中居, 黄瀬, 岩村: "デジタルカメラによる文書画像検索—1 万ページから 0.1 秒で検索する—", *情報科学技術レターズ*, vol. 4, LI-001, pp.133–136 (2005).
- [3] H. J. Wolfson and I. Rigoutsos: "Geometric hashing: an overview", *IEEE Computational Science & Engineering*, vol. 4, no. 4, pp.10–21 (1997).
- [4] T. Suk and J. Flusser: "Point-based projective invariants", *Pattern Recognition*, vol.33, no.2, pp.251–261 (2000).
- [5] 佐藤: コンピュータビジョン—視覚の幾何学—, コロナ社 (1999).