

射影不変量とホモグラフィ行列を用いた 文字列画像検索方式の提案

An Image Retrieval Method for Character Strings Using Projective Invariants and Homography Matrices

秋山 達勇†
Tatsuo Akiyama

西脇 大輔†
Daisuke Nishiwaki

1. はじめに

現実世界には、案内板、看板など、文字列情報を提示することを目的とするものがある（以下、これらを総称して看板と呼ぶ）。看板に記すことのできる情報量は限られているため、補足的な情報を提供できれば、看板の効果を高められると考えられる。このようなことを実現するための一つの方法として、看板の画像データとその補足的な情報とをネットワーク上に登録しておき、カメラデバイスを使って看板を撮影し、この撮影画像をキーとして、ネットワーク上に登録した情報を検索することを考える。具体的な利用場面として、カメラ撮影された看板画像への拡張現実的な情報提示、看板に記載される文字列情報の翻訳表現や補足説明の提示等が想定できる。

このような画像を用いた情報検索を行う場合、カメラで撮影した画像（検索画像）の視点は一定しないため、ネットワーク上に登録した、固定された視点から撮影した画像（登録画像）との照合を行うには、射影ひずみによる変形を考慮する必要がある。本報告では、画像間で見かけ上生じる看板の変形を射影変換でモデル化し、検索画像と登録画像とを照合する方式を提案する。まず、各画像上の特徴点の局所的な配置から計算される射影不変量の列を用いて両画像間の局所的特徴点配置の対応関係を求める。その後、対応する局所的特徴点配置の対から推定される射影変換に関するパラメータ（ホモグラフィ行列）を用いて検索画像と登録画像の位置合わせ（画像補正）を行い、こう配特徴を用いて照合する。

2. 従来技術

射影変換や2次元アフィン変換といった幾何学的変換を画像間の変形モデルとする従来方法に、中居らの文書画像検索方法[1]がある。この方法は重心特徴点の局所的配置を特徴点群から計算される幾何学的不変量[2]の列として特徴付けることができるため、画像間でモデルに含まれる変形が起きても原理的に検索可能である。しかし、重心の位置情報だけでは、重心位置が（ほぼ）一致しているのに形状が異なる連結領域を見分けられない。中居らは、英文文書のように連結領域が単語単位であるものを想定しているが、漢字やひらがな等、連結領域が文字毎になるものを想定すると、検索精度を確保する上で工夫が必要である。筆者らは、中居の方式を、これらの点で改良し、2次元アフィン変換を変形モデルとした画像同定（検索）方法を提案し、スキャナ撮像された郵便住所等の文字列画像を対象に実験を行った[7]。一方、画像の局所的なこう配情報に基づく局所特徴量 SIFT[3]、

† NEC システムIPコア研究所

SURF[4]は本来回転・スケール不変であるが、微小な3次元傾き（アフィン変化）に対してある程度頑健であること[5]や利用例[6]が報告されている。本報告では[7]における画像変形モデルを射影変換に置き換えた方式を提案し、SURFとの比較を行う。

3. 提案方式

3.1 概要

本報告の提案方式は、（特徴点である重心等が射影変換変形の影響で微妙に位置ずれすることを除けば、）射影変換で変形された画像対に関して連結領域形状を反映した照合が可能である。方式の概略は次の通りである。

- (1) 登録画像と検索画像の双方から、画像中の特徴点（重心等、詳細は[8]参照）の局所的配置を特徴付ける特徴量（射影不変量）の列を計算する。
- (2) 特徴量の列を利用して対応する登録画像の特徴点配置と検索画像の特徴点配置との対応関係を求める。
- (3) 看板の平面性を仮定し、対応する特徴点配置対から推定されるホモグラフィ行列を用いて画像補正する。
- (4) 連結領域毎に局所こう配特徴を利用した照合を行い、各照合結果を統合して、最終的な同定判定を行う。

3.2 射影不変量

射影不変量とは、照合対象の2画像間に射影変換で記述される変形があっても不変となる特徴を持つ量である。例えば、3次元空間上のある平面を撮影したとき、その平面上の5点に関する画像中の点を A, B, C, D, E とし、 $P(X, Y, Z)$ を三角形 XYZ の面積とすると、複比

$$\frac{P(A, B, C)P(A, D, E)}{P(A, B, D)P(A, C, E)} \quad (1)$$

が射影不変量となる[2]。本提案では、[8]の手法で計算される特徴点の局所的配置に対して、中居らの画像検索方法[1]と同様の方法で複比の列を計算する。計算された複比列をベクトルとみなし、複比列対から計算されるベクトル間距離が一定値より小さい場合、その複比列対が計算された基である特徴点配置対が一致したと判定する。

3.3 ホモグラフィ行列

ホモグラフィ行列 H とは、2つの視点から3次元空間上のある平面上の任意の点 i について、2つの視点の画像中の点 i の座標 (x_i, y_i) と (x_i', y_i') を関係付ける行列である。具体的には、

$$\alpha_i \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = H \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2)$$

なる関係を持つ[9]。(ここで α_i は点*i*に応じて定まるスカラー量である。)本発表では、この行列を推定するために、一致したと判定された特徴点配置を構成する点($m=7$ 個)の座標を利用する。この H は3.1節(3)の画像補正を行うために利用される。

4. 実験

提案方式に関して、基礎的な評価実験を行った。評価に用いた画像は、図1の上下段2種類の画像である。また、2種類それぞれの正対画像を人工的に変形させて評価画像を作成した。この変形は、 $z=d$ の位置にある画像中心(0,0,d)のオブジェクトに対して、画像中心に視点を向けたまま平面 $y=0$ 上で角度 θ だけカメラを回転させてオブジェクトを撮影した場合に得られる変形に相当する。

本実験では、 $d=60$ [cm]、 θ は -80 [deg]から 80 [deg]の範囲を 10 [deg]刻みで変動させた。2種類の正対画像を登録画像として、変形した画像を検索画像として用いた。

実験結果を図2に示す。図2に示すグラフの横軸は回転角 θ 、縦軸は類似度である。提案手法の類似度は、局所的な領域の照合に関して、(一致した局所領域数)/(ホモグラフィ行列により投影先の有効領域内に画像を補正できた領域数)で定義した。比較のため、図2にはOpenCV2.0を利用して作成したSURF特徴による照合についての評価結果も記した。提案手法との比較のため、SURF特徴を用いて照合した場合の類似度を(一致したSURF特徴点数)/(登録画像から計算されたSURF特徴点数)で定義した。

提案手法はSURFに対して、全体的に検索画像と登録画像とが異なる種類の画像の場合に類似度が小さい傾向があった(図2)。また、検索画像と登録画像が同じ種類の画像の場合には、相対的な傾きが 20 [deg]~ 60 [deg]の範囲において、提案手法の類似度が高い傾向があった。

検索画像と登録画像が同じ種類の画像の場合の類似度をさらに向上させるためには、ホモグラフィ行列の推定精度、射影変換歪みに起因する画像解像度低下などに関する考察が必要である。図2には、理想的な H の値を用いて3.1節(3)の画像補正を行った場合の類似度も示した。ホモグラフィ行列の推定精度向上によって、類似度がある程度向上することが可能と推測される。

5. まとめ

本発表では、カメラ等で傾いて撮影された看板等の画像を検索することを目的として、射影不変量およびホモグラフィ行列を用いる検索方式を提案した。人工的に作成したテスト画像を用いた基礎的な実験の結果、画像検索においてよく利用されるSURF特徴量を利用した検索方法と比較して同程度以上の精度が得られた。今後、実際の情景画像中の看板画像についても評価する予定である。

参考文献

[1] 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, Vol. J89-D, No.9, pp. 2045-2054, 2006.
 [2] T. Suk, J. Flusser, “Point-based projective invariants”, Pattern Recognition, Vol. 33, No. 2, pp. 251-261, 2000.
 [3] D. Lowe, “Distinctive Image Features from Scale-Invariant

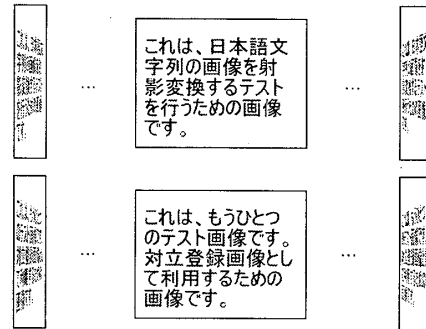
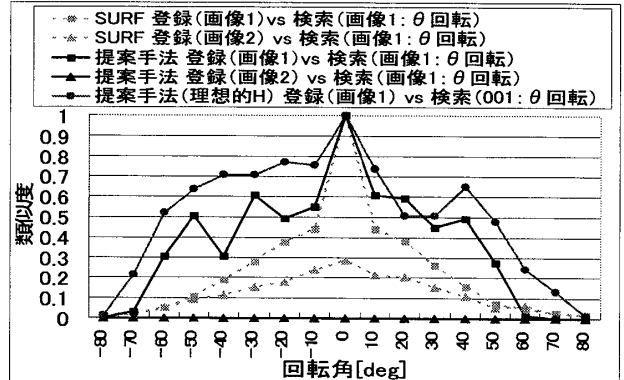
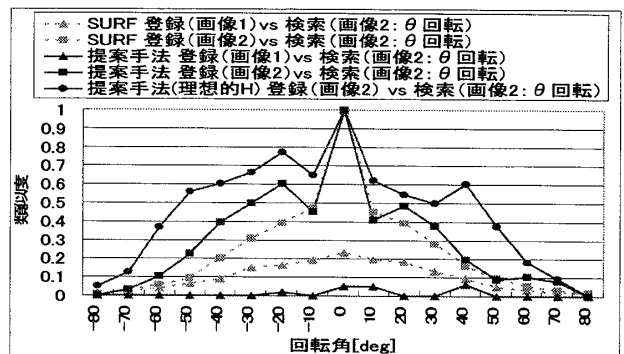


図1. 評価用テスト画像 (上段: 画像1, 下段: 画像2)



(a) 検索画像が画像1の場合



(b) 検索画像が画像2の場合

図2. 回転角 θ と類似度との関係

Keypoints”, International Journal of Computer Vision 60(2), pp. 91-110, 2004.

[4] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “Speeded-Up Robust Features (SURF)”, Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359, 2008.

[5] 藤吉弘亘, “Gradientベースの特徴抽出 - SIFTとHOG -”, 情報処理学会研究報告 CVIM 160, pp. 211-224, 2007.

[6] 内田, 伊東, 岩村, 大町, 黄瀬, “ペン先画像からの手書き復元”, 信学論 D, Vol. J93-D, No.1, pp. 64-67, 2010.

[7] 秋山, 高橋, 大網, 西脇, “画像領域全体の粗い色特徴配置と配達先住所領域内の局所的なこう配特徴を利用した発送物の同定”, ViEW2009講演論文集, pp. 170-175, 2009.

[8] 秋山, 近藤, 西脇, “アフィン不変量を用いた欧文印刷住所画像の同定方式に関する提案”, 信学技報, Vol. 107, No. 491, pp. 7-12, PRMU2007-215, 2008.

[9] 出口光一郎, “ロボットビジョンの基礎”, コロナ社, 2000.