

# 背表紙画像を用いた書籍の位置特定のための 特徴点マッチングにおける特徴量の比較評価

山内 翔平<sup>†</sup> 梅澤 猛<sup>‡</sup> 大澤 範高<sup>‡</sup>

千葉大学工学部<sup>†</sup> 千葉大学大学院融合科学研究科<sup>‡</sup>

## 1 はじめに

大量の書籍が収納された書棚から、画像処理技術を使って書籍の位置を特定することができれば、配架が乱れている場合でも短時間で効率よく目的のものをを見つけるのに役立つ。高橋らは、携帯端末で書棚の動画を撮影し、目的の書籍の背表紙画像に対して SURF 特徴量を用いた特徴点マッチングを行うことで配置場所をリアルタイムで特定する手法を提案した[1]が、特定精度は 65%程度と実用化には課題が残る。

SURFは拡大縮小・回転・照明変化に頑健であるとされるが、ほかに SIFT, AKAZE, ORB, BRISK の各特徴量も同様の性質があるとされる。このうち ORB と BRISK は FAST と呼ばれる高速な特徴点検出手法と特徴量記述にバイナリコードを使用することで SURF よりも高速なマッチングを可能にしている。AKAZE は非線形スケールスペースと M-LDB ディスクリプタを使用している。このため SURF, SIFT よりも高速で、ORB, BRISK よりも特徴点抽出と特徴量記述において高い性能を期待できる。

そこで本研究では、SURF, SIFT, AKAZE, ORB, BRISK の 5 つの各特徴量が書籍の位置特定精度に与える影響を比較検討することで精度の向上の方策を探る。

## 2 検索方法

### 2.1 検索手順

本研究では研究[1]の検索手順を踏襲し、以下の 5 つの手順で検索を行う。

- 1) 目的の書籍の選択
- 2) 書棚の撮影
- 3) 特徴点マッチング
- 4) ホモグラフィ行列の計算
- 5) 検索結果の表示

まず 1) では目的の書籍の背表紙画像を選択する。今回は、予め携帯端末で撮影した背表紙画像を使用する。2) では携帯端末で書棚を撮影し、3) では 1), 2) で得られた画像を元に特徴点マッ

Comparative evaluation of features point matching for finding book using spine image

<sup>†</sup>Shohei Yamauchi, Faculty of Engineering, Chiba University

<sup>‡</sup>Takeshi Umezawa, Noritaka Osawa, Graduate School of Advanced Integration Science, Chiba University

ングを行う。

特徴点マッチングより得られた対応点から、4) でホモグラフィ行列を求め、背表紙画像の頂点座標を書棚画像の対応する座標に変換する。その際に、対応点に外れ値が存在する場合、RANSAC アルゴリズム[2]を用いて削除する。そして 5) で検索結果を表示する。求めたホモグラフィ行列を用いて、背表紙画像の角の 4 点の座標を書棚画像の座標に変換し、書籍の外周を四角形で囲み、検索結果として表示する。

### 2.2 特徴点マッチング

特徴点マッチングでは、背表紙画像と書棚画像から各特徴量の算出アルゴリズムにより特徴点を抽出し、特徴点それぞれに対し特徴量記述を行う。得られた特徴量ベクトルに対して 2 近傍探索を行い、第 1 候補距離を第 2 候補距離で割った値が閾値以下となれば第 1 候補を対応点として結びつける。

### 2.3 ホモグラフィ行列

ホモグラフィとは、ある平面を別の平面に射影することができる変換である。特徴点マッチングで求めた対応点である背表紙画像と書棚画像の座標を  $(x_i, y_i)$ ,  $(x'_i, y'_i)$ 、ホモグラフィ行列を  $H$  とすると変換式は以下のように表される。

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad \left( H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \right)$$

そこから、以下の逆投影誤差を最小化するようなホモグラフィ行列を求める。

$$\sum_i \left( x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left( y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2$$

ホモグラフィ行列はスケールに依存するため、 $h_{33} = 1$ として正規化する。

### 2.4 マッチング例

書棚画像から目的の書籍の位置を特定するプログラムを実装し、それを用いて実験を行った。図 1 に特徴点マッチングによって書籍の位置を特定したときの処理結果の一例を示す。(a)は書棚から目的とする書籍の背表紙と一致する書籍の位置を特定した様子である。左端が背表紙画像で、その右側が書棚画像である。抽出した特徴点を丸くプロットし、対応する点同士を直線で

結んでいる。また、書棚画像中の位置を特定した書籍の外周を線で囲み、検索結果として表示している。(b)はこのときの背表紙画像の一部を拡大した図、(c)は書棚画像中の該当書籍を拡大した図である。



(a) 背表紙および書棚の画像



(b) 背表紙画像の拡大図



(c) 書棚画像の拡大図

図1 特徴点マッチングによる書籍位置の特定

### 3 比較実験

先述した検索手順に従い比較実験を行う。書棚から50cmの地点で撮影した書棚画像を用いて、それぞれの特徴量において、閾値を0.5~0.9まで0.1ずつ変化させ、5冊分の背表紙画像と特徴点マッチングを行い比較する。なお、検出精度の評価には、書籍の検索結果と書棚における正解領域との適合率、再現率及び適合率と再現率の調和平均であるF値を用いる。

検索結果の領域をA、書棚の正解領域をBとすると、適合率は $(A \cap B)/A$ 、再現率は $(A \cap B)/B$ で表わされ、それぞれ検索の正確性と網羅性を示している。

### 4 結果とまとめ

図2に実験により得られた各特徴量の閾値を変化させた際の適合率、再現率及びF値を表わしたグラフを示す。エラーバーは標準偏差を示している。

図2より、ほとんどの閾値で適合率、再現率及びF値が0.9を超えており、どの特徴量でも書籍の検索精度が高いことがうかがえる。また、AKAZEとORBの閾値が0.5の場合は閾値が低く、十分な対応点がとれずに書籍の検索ができなかったため、他よりも低い値になっている。再現率は、AKAZEで閾値が0.6以上の場合に一番高い値を示している。そのためAKAZEが最も書籍の正解領域を検出していることがわかる。

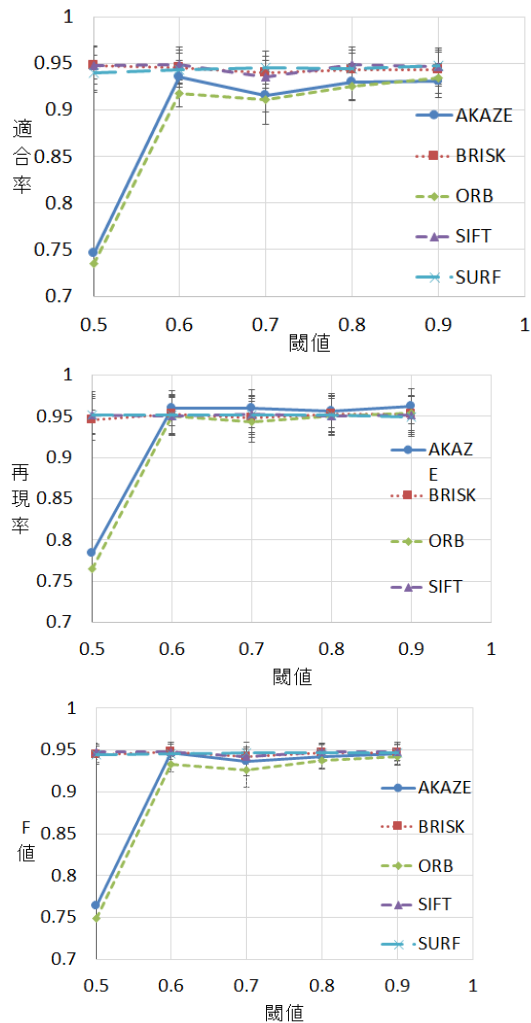


図2 各特徴量における適合率、再現率及びF値

F値は、BRISK、SURF、SIFTが安定した値をとっている。

しかし、リアルタイムで動作するよう実装を目指すには処理時間も大きく関わってくるため、今後は処理時間をはじめ、書棚からの距離や角度、書棚画像の画素数なども比較条件として評価し、検索精度の向上を目指す。

### 参考文献

- [1] 高橋 政樹, 富澤 勇介, 高井 昌彰, “拡張現実技術を用いた背表紙画像による書棚の図書検索システム”, 電子情報通信学会技術研究報告, MVE, マルチメディア・仮想環境基礎, vol.111, no.235, pp.11-15, 2011.
- [2] M. A. Fischer and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography”, Commun. ACM, no.24, vol.6, pp.381-395, June 1981.