

# 特定物体認識における SIFT 特徴点マッチング手法の改善

中西 雅哉<sup>†</sup> 大野 将樹<sup>‡</sup> 獅々堀 正幹<sup>‡</sup>

徳島大学大学院 先端技術科学教育部<sup>†</sup> 徳島大学大学院 理工学研究部<sup>‡</sup>

## 1. はじめに

特定物体認識とは、特定の目標物が写っている画像を検索するものである。特定物体認識では、SIFT 特徴量[1]などの局所特徴量を用いて認識をするのが一般的である。しかし、対象となる物体以外の特徴点とマッチングすることが問題となる。そこで、本研究では SIFT 特徴量の類似性・出現位置を考慮してノイズの除去を行い、誤対応を減らすことで検索精度を向上させることを目的とする。

## 2. 関連研究

SIFT 特徴量を用いた特定物体認識に関する研究として、SIFT 特徴量を用いた交通道路標識認識[2]や機械用基板における物体認識[3]など、幅広く研究されている。

しかし、従来手法による特定物体認識では SIFT 特徴量を用いてクエリ画像と検索画像の特徴点をマッチングした際に、クエリ画像の特徴点が目標の物体以外の特徴点にマッチングしてしまい、誤対応が生まれてしまうという問題があった。図 1 に示すように、左側のクエリ画像（ロゴマーク）が、右側の検索画像のロゴマーク以外の特徴点にマッチングし、誤対応が生じてしまっている。

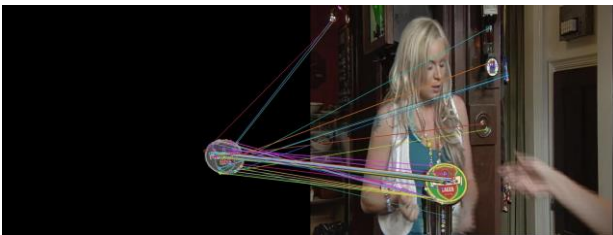


図 1：従来手法による SIFT 特徴点の誤対応の例

## 3. 提案手法

従来手法により生じた誤対応を減らすために、SIFT 特徴量の類似性・出現位置を考慮してノイズを除去する手法を提案する。提案するシステムの概要を図 2 に示す。

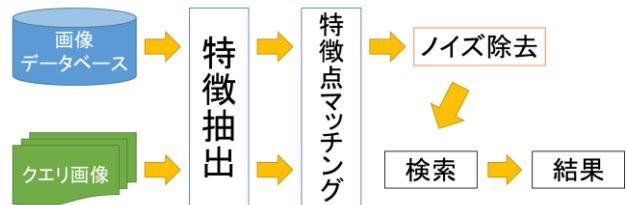


図 2：システム概要

検索画像とクエリ画像から抽出した SIFT 特徴量を比較し、特徴点マッチングを行う。そこで生まれるノイズを以下の方法で除去する。

### 3.1 相対的類似度によるノイズ除去

SIFT 特徴量間の類似性に着目する。毛髪や柱などの特徴量は画像中にありふれており、特徴量間の類似性が高い。一方、クエリ内の特徴量はクエリ以外に類似した特徴量が少ない。そこで、SIFT 特徴量の類似度上位 1 件目を 2 件目で割ることで上位 2 件の比を求める。これを相対的類似度とする。閾値を定め、上位 2 件の類似度が高いものはノイズとして除去する。

### 3.2 相対的距離によるノイズ除去

SIFT 特徴量の類似度上位 2 件の出現位置に着目する。図 3 の示すように、クエリ内とクエリ以外の特徴点の類似度上位 2 件の座標間の距離をそれぞれ調べると、クエリ内の特徴点は座標間の距離が 300pixel 以内の点が多いという結果を得た。そこで、この座標間の距離を相対的距離とし、閾値を定め、上位 2 件の距離が遠いものはノイズとして除去する。



図 3：SIFT 特徴点の出現位置の違い

### 3.3 双方向の類似性を考慮したノイズ除去

クエリ画像と検索画像の SIFT 特徴点を比較し、互いに最近傍の点だったとき、その点を特徴的な点とする。そうでない点はノイズとして除去する。

## 4. 実験

### 4.1 データベース作成用画像

データベース画像として、映像コンテンツの

Improvement of matching between SIFT feature points in specific object search

Masaya Nakanishi<sup>†</sup>, Masaki Oono<sup>‡</sup> and Masami Shishibori<sup>‡</sup>

<sup>†</sup>Department of Information Science and Intelligent Systems, University of Tokushima

<sup>‡</sup>Faculty of Science and Technology, Tokushima University

解析や検索の高度化を目的とした競争型プロジェクトである TRECVID[4]のインスタンスサーチタスクで用いられる動画をショット出現時間ごとに分割した画像を用いる。

#### 4.2 実験内容

提案手法の有効性を示すため、以下の 3 つの実験を行った。

- 従来手法に対する実験（実験 1）  
最近傍の SIFT 特徴量間のユークリッド距離をすべて採用する。
- 相対的類似度及び距離を用いた実験（実験 2）  
相対的類似度が 0.8 未満，相対的距離が 300pixel 未満を満たす点のみを採用する。
- 双方向の類似性を考慮した実験（実験 3）  
クエリ画像と検索画像のキーポイントが互いに最近傍の点のみを採用する。

また、検索するクエリ画像は 8 枚使用する。検索画像は 2,000 枚のデータベースを用意し、そのうち正解画像を 100 枚含める。検索結果の上位 300 件中にどれだけ検索対象が含まれているかを評価対象とし、適合率と再現率で評価する。

#### 4.3 実験結果

全クエリ画像に対する適合率・再現率の平均を図 4 に示す。適合率、再現率のグラフは右上にあるほど高い精度であるといえるので、実験 1 よりもノイズを除去した実験 2，実験 3 のほうが高い精度で検索できることがわかる。これより、本手法は従来手法で発生していた特徴点の誤対応を減らすことができたと考えられる。

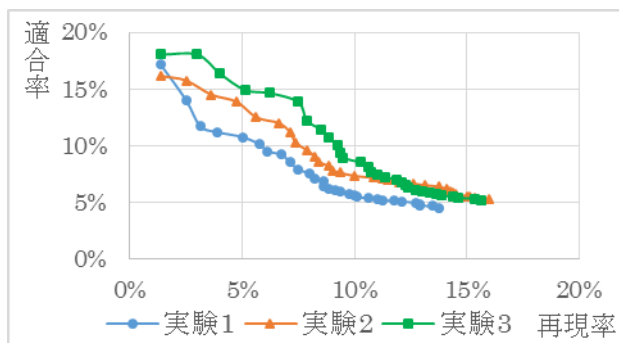


図 4：実験結果

#### 5. 考察

実験結果より全体平均で見ると、本手法は従来手法よりも精度を向上させることができたことがわかる。しかし、個々のクエリに対する結果を見ると、腕時計のようにサイズの小さいクエリは、従来手法よりも提案手法の方が精度が低下していた。原因としては、腕時計から取れるキーポイントの数が少なく、ノイズの除去によって対応点がほとんどなくなってしまうこと

である。図 5，図 6 に示すように、ノイズの除去前と除去後と比較すると、誤対応を大きく減らしてはいるが、正しい対応もほとんどなくなってしまうことがわかる。これより、クエリによって、ノイズ除去の閾値の値を正しく設定する必要がある。

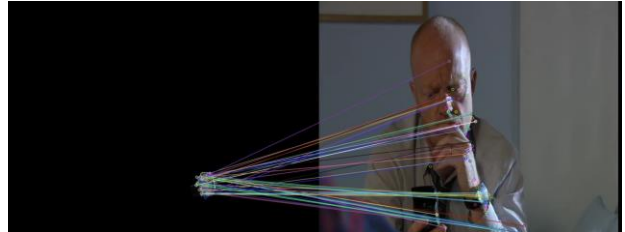


図 5：ノイズ除去前のマッチング画像

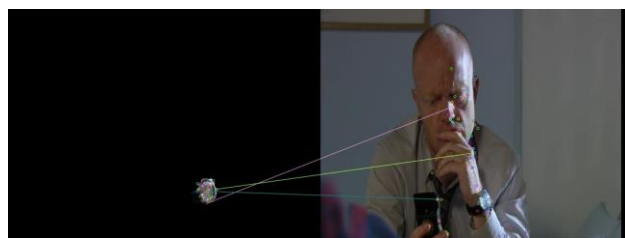


図 6：ノイズ除去後のマッチング画像

#### 6. まとめ

特定物体認識手法として、SIFT 特徴点のマッチングをする際に生じる誤対応を減らすために、局所特徴量の類似性・出現位置に着目して、ノイズ除去を行う手法を提案した。従来手法との比較をおこなった所、精度の向上が見られた。しかし、クエリによっては低い精度を示すものもあった。

今後の課題としては、ノイズの除去がうまくいかなかったクエリに対し、正対応・誤対応をうまく判別するための SVM による機械学習を導入し、更なる精度の向上を目指したい。

#### 参考文献

- [1] D. G. Lowe, “Object recognition from local scale-invariant features”, Proc. of IEEE International Conference on Computer Vision (ICCV), pp.1150-1157 1999.
- [2] 高木雅成, 藤吉弘亘, “SIFT 特徴量を用いた交通道路標識認識”, 第 13 回画像センシングシンポジウム SSII07, LD2-06, 2007.
- [3] 三吉建尊, 越野亮, 木村春彦, “HOG と Bag of Keypoints を用いた工作機械用制御盤内における物体認識に関する研究”, 第 27 回ファジィシステムシンポジウム講演論文集, pp.1183-1188, 2011.
- [4] TRECVID, <http://trecvid.nist.gov>