

Web 画像を用いたカテゴリ別 Visual Words による一般物体認識

田中 雄翔[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学工学部情報知能工学科

^{††} 神戸大学自然科学系先端融合研究環

1. はじめに

近年、画像データの大容量化やロボット産業の発展に伴って、計算機による一般物体認識の必要性が高まっている。局所特徴量による一般物体認識において、有効性を示しているのが、

BoF(Bag of Features)を用いたアプローチ[1]である。これは、画像から SIFT(Scale-Invariant Feature Transform)特徴量を抽出し、その集合を用いて画像全体を特徴付ける手法であるが、新たなカテゴリを認識するためには新たにそのカテゴリの画像を加えて学習する必要があり、認識するカテゴリが増えるほど、認識率は低下するという問題点がある。そこで、本稿では、上記の問題を解決するために、Web 画像を学習に用い、カテゴリ別に Visual Words を作成しカテゴリ照合を行う手法を提案する。

2. 提案手法

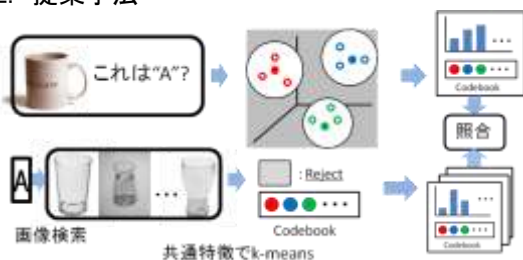


図1. 提案手法の流れ

提案手法の流れを図1に示す。ある画像について、「その画像がAであるか」というタスクが与えられると、A というワードで画像検索を行い、画像を収集する。集まった画像に共通する特徴[2]で、BoFを構築する。Web 画像を利用することで、未知の物体認識にも対応できる。このとき、A に関係ない特徴点は排除したい。そこで k-means を実行する際に、Reject 領域という特徴点棄却域を作ることで、不必要な点を棄却する。

3. 実験

まず、認識問題として、従来の BoF との比較を行い、その後、照合実験を行った。用いた画像クラスは、Caltech-101 に存在する 10 クラスであり、各クラス 30 枚を画像検索で収集して学習し、Caltech-101 に含まれている各クラス 20 枚をテストとして、10 クラス分類を行った。分類にはともに k-NN 法を用いた。実験結果を表 1 に示す。

表 1. 実験結果(%)

	従来手法 (BoF)	提案手法
認識率	32.5	42.5

BoF を用いた他の研究に比べ認識率が低いのは、学習に用いたデータが実際に Web で収集した画像であるためノイズが非常に多いからであるが、従来手法よりは精度は高くなった。

次に、照合実験では、上記と同じ 10 クラスに関して照合を行った。precision-recall の最大 F 値における平均結果を示す。

表 2. 実験結果

Precision	Recall	F
0.3831	0.675	0.4883

4. まとめ

本稿では、カテゴリ別 Visual Words を用いて一般物体認識を行う手法を提案し、その有効性を示した。

参考文献

- [1] P.Duygulu,K.Barnard,N.d.Freitas and D.Forsyth,"Visual categorization with bags of keypoints," European Conference on Computer Vision, pp.IV:97-112,2002.
- [2] Mohamed Aly, Mario Munich, Pietro Perona, "Using More Visual Words in Bag of Words Large Scale Image Search", April 2011.