

特徴点クラスタリングを用いた詳細画像識別に関する一検討

田中健太[†] 青木輝勝^{†‡}

[†] 東北大学大学院情報科学研究科

[‡] 東北大学未来科学技術共同研究センター

1.はじめに

詳細画像識別は、視覚的に類似する物体を分類するタスク(例:車カテゴリを車種別クラスに分類)であり、近年注目を集めている。従来研究[1][2]では、一般的な BoF フレームワーク[3]を利用しているため、Visual words 生成のために局所特徴量だけが使用されている。しかし、物体認識を行う上で、大域的な情報もきわめて重要である。本稿では、大域特徴量および局所特徴量の両方の性質を併せ持ち、局所特徴量として扱うことができる特徴量記述方式を提案する。

2.提案手法

2.1.局所特徴量と大域特徴量の融合

画像特徴量は、大域特徴量と局所特徴量に大別することができる。SIFTをはじめとする局所特徴量はロバスト性が高く、画像から多数抽出することにより BoF フレームワークに用いることができる。しかし局所特徴量は局所領域から抽出するため、局所領域の情報しか有していない。そのため、局所特徴量のみを抽出した場合、それが何の情報であるか決定するには限界がある。一方で、大域特徴量を与えた上で局所特徴量に注目したとき、周辺情報と比較することが可能になるため、より高い精度で何の情報であるか決定することができるようになる(図1)。

このことから明らかなように大域特徴量は、物体情報を得る上で極めて重要である。しかし大域特徴量は画像1枚に対して1つしか得られないため、BoF フレームワークを用いることができない。そこで、大域特徴量と局所特徴量を融合することにより、両方の性質を持つ新たな特徴量記述を行う。本稿で提案する特徴量は図2の赤斜線部に位置するものとなる。

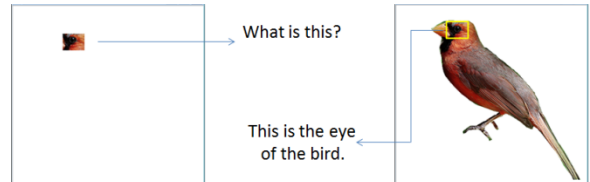


図1 局所特徴量 VS 大域特徴量中の局所特徴量

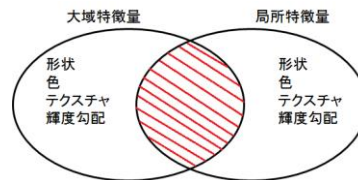


図2 提案特徴量の位置付け

2.2.提案特徴量の記述方法

提案特徴量の概要を図3に示す。局所特徴量と大域特徴量を融合した提案特徴量の具体的な記述方法を以下に示す。

1. GrabCut を用いて予め物体領域を算出し、物体領域中の画素を用いて大域特徴量を算出する。
2. Dense sampling を用いて特徴点の検出を行い、各特徴点およびその周辺画素を用いて、局所特徴量を抽出する。
3. 次式に従い、局所特徴量と大域特徴量を融合する。

$$v_{(n,i)} = \begin{cases} v_{(l,i)} & (\text{if } v_{(l,i)} \neq 0) \\ -v_{(g,i)} & (\text{if } v_{(l,i)} = 0) \end{cases} \quad (1)$$

A Novel Semantic Keypoint Clustering Method for Fine-grained Visual Categorization

Kenta Tanaka[†] Terumasa Aoki^{†‡}

[†] Graduate School of Information Sciences, TOHOKU University

[‡] New Industry Creation hatchery Center, TOHOKU University

ここで、 P は特徴量次元、 $V_n = (v_{(n,1)}, \dots, v_{(n,P)})$ は提案特徴量、 $V_g = (v_{(g,1)}, \dots, v_{(g,P)})$ は大域特徴量、 $V_l = (v_{(l,1)}, \dots, v_{(l,P)})$ は局所特徴量である。

提案方式では、局所特徴量はそのまま保存し、局所特徴量に存在しない情報は、大域特徴量に置換しマイナスの値で表現する。これにより、大域特徴と局所特徴の性質を併せ持ち、局所情報をより顕著に表すことができる局所特徴量となる。提案特徴量の記述方法を図 4 に示す。

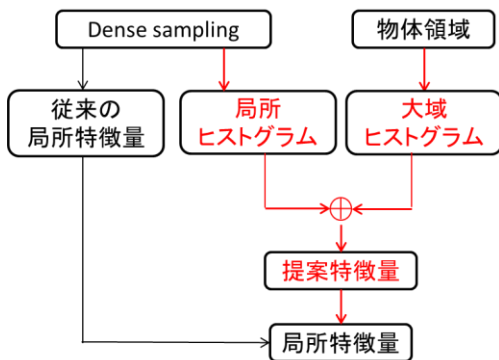


図 3 提案特徴量の概要

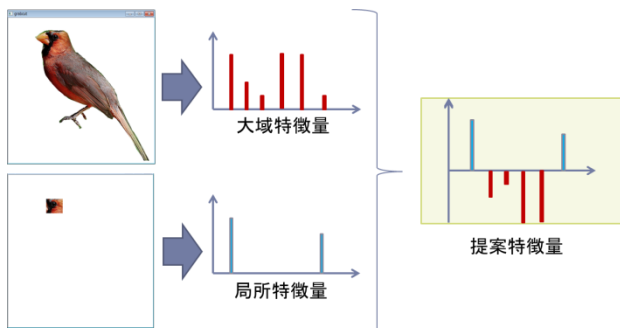


図 4 提案特徴量の記述方法

3. 評価実験

提案特徴量の有効性を検証するため、Caltech-UCSD Bird-200-2011 データセット [4] から 10 クラス 597 枚の画像をピックアップし、実験を行った。実験では、提案特徴量と従来の特徴量を統合し、BoF フレームワークおよび SVM により認識を行った。また従来の局所特徴量には SURF 特徴量、提案特徴量には色情報として Hue Histogram を用いた。

提案特徴量を用いて精度比較を行った結果を図 5 に示す。提案方式において、従来の特徴量や単純な統合による特徴量よりも優れた認識精度が得られた。

		特徴量	認識精度(%)
従来の特徴量	局所情報のみ	SURF	53.8
		V_l +SURF	65.5
提案特徴量	大域情報+局所情報	V_g +SURF	61.4
		提案方式+SURF	69.0

図 5 実験結果

4. まとめ

本稿では、詳細画像識別の高精度化を目的として、大域特徴量と局所特徴量の両方の性質を併せ持つ新しい特徴量の記述方法を提案した。また実験により、従来の局所特徴量よりも高精度に識別できることを示した。今後の課題としては、大域特徴量の抽出方法、提案した特徴量の記述方式について更なる検討を行う。

文献

[1] Nakayama, H., "Augmenting descriptors descriptors for fine-grained visual categorization using polynomial embedding". ICME, 2013.

[2] P-H. Gosselin, N. Murray, H. Jegou, F. Perronnin, Inria+Xerox@FGcomp. "Boosting the Fisher Vector for Fine-Grained Classification". INRIA Tech, 2013.

[3] C. Dance, J. Willamowski, L. Fan, C. Bray, and G. Csurka. "Visual Categorization with bags of keypoints". ECCV 2004.

[4] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. "The Caltech-UCSD Birds-200-2011 Dataset". CNS-TR-2011-201.