

PRMU アルコン 2015：靴のロゴ検出

48-156212 小林 勇也

1. 概要

「3. コンテストにチャレンジしてみた」を行った。挑戦したコンテストは 2015 年度 PRMU アルゴリズム コンテスト ～あの靴、どこ?～である。与えられた靴画像 (図 1, 図 2 参照) からメーカーのロゴ検出をするプログラムが求められる。作業はロゴ認識とロゴ領域検出の 2 種類の要素からなるが、今回はロゴ認識のみを対象にした。



図 1. 学習画像例



図 2. テスト画像例

2. 適応手法

試した手法は以下の 2 通りである。

1. SIFT 特徴量同士が最近となるものに投票する方法
2. SIFT+BoF+SVM を用いた方法

2.1 手法 1

学習画像群から得られた SIFT 特徴量を用いて、クエリに対する最近点対の効率的な計算が可能な kd 木を構成する。テスト画像からの SIFT 特徴量各々に最近となる kd 木内の SIFT 特徴量を求めて、その特徴量に対応する学習画像のロゴラベルに投票を行う。最大の得票となるロゴラベルをそのテスト画像のロゴラベルとする。

2.2 手法 2

SIFT, BoF, SVM を用いる手法は物体認識で広く使われている手法である [1][2]。SIFT 特徴量を用いて、BoF (Bag of Features) と呼ばれるデータ群を作成する。BoF から得られる画像と 1 対 1 対応するベクトルを入力、ロゴラベルを出力とする多クラス SVM を学習し、予測を行う。

Bag of Features

学習画像群から得られた SIFT 特徴量を Kmeans 法によりクラスタリングする。この結果、学習画像に対応するクラスタリングラベルのヒストグラムを得る。このヒストグラムを正規化したものを画像を表現する特徴量とする方法が BoF である。[図 3 を参照] また、テスト画像に関しては SIFT 特徴量に最も近い Kmeans 法のクラスタリング中心を与えるクラスタリングラベルを対応づけることで、ヒストグラムを得る。

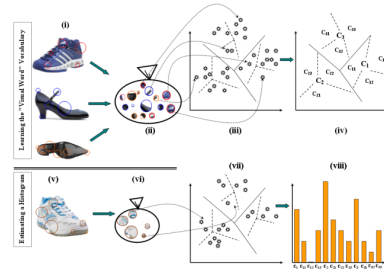


図 3. BoF の図解, [2] より転載

3. 実験

<https://sites.google.com/site/alcon2015prmu/home> で配布されている画像を実験に用いた。画像は認識の困難さに応じて、Lv1-3 までレベル分けされている。

$$\text{precision} = \frac{\text{正解データ数}}{\text{正解矩形数}}$$

$$\text{recall} = \frac{\text{正解データ数}}{\text{出力矩形数}}$$

上記の precision, recall により性能を評価した。

手法 1	Lv1	Lv2	Lv3
precision[%]	67	57	42
recall[%]	62	54	38

表 1. 手法 1 の認識精度

手法 2	Lv1	Lv2	Lv3
precision[%]	30	26	20
recall[%]	31	24	19

表 2. 手法 2 の認識精度

4. 追記

本発表のソースコード及び、発表スライドを <https://github.com/PRMU-algo-2015/yuyak> に載せる。

参考文献

- [1] T. Vo, D. Tran and W. Ma, Tensor Decomposition of Dense SIFT Descriptors in Object Recognition, In Proceedings of European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2014.
- [2] <http://www.codeproject.com/Articles/619039/Bag-of-Features-Descriptor-on-SIFT-Features-with-O>, last access 2015/7/6.