
Objects Recognition Based on SOG with SVM

Final Report for CIS 519

Shangyi Cheng
Yao Chu
Chenyang Zhao

SHANGYI@SEAS.UPENN.EDU
CHUYAO@SEAS.UPENN.EDU
CHZHAO@SEAS.UPENN.EDU

Abstract

We studied the question of object recognition using histograms of oriented gradients (HOG) with support vector machine (SVM) with Gaussian Kernel on ball detection as a test case. After reviewing several existing methods for feature extraction, our project verified that HOG performs well in ball recognition. We show the whole process from choosing regions of interest manually, extracting features from these regions, using principal component analysis (PCA) to reduce dimension of the dataset and then applying SVM with Gaussian kernel to do the classification.



Figure 1. Analysis of Available Data

1. Introduction

The target for this project is to detect the object of our interest and give the label for given images. The dataset we used are images of four kinds of balls from Caltech 256(Griffin, G. Holub, AD. Perona, P.), which contain 98 images of golf in the folder “088.golf-ball”, 174 images of soccer ball in “193.soccer-ball”, 104 images of bowling-ball in “017.bowling-ball” and 98 images of tennis ball in “216.tennis-ball”. As shown in Figure 1, among images of the same kind(take soccer-ball as example), some of the images have a whole soccer in the middle and fill the image, while in other images(such as a photo of a soccer game), the soccer may cover a small portion in the corner, only part of a soccer or several soccers are contained. We decide to use the image with a single target occupying most of the image as the training set. So the first thing we need to do before training is to create such standard images for training from the raw dataset.

2. Model Details

2.1. Overview of the Method

Inspired by the paper presented(Dalal & Triggs, 2005) by Navneet Dalal and Bill Triggs, we will try the method described in the paper. The interested region will be labeled manually for the training data. HOG features can be extracted from the interested region. Then SVM or other classifier will be applied to train the model.

2.2. Region of Interest (ROI) Selection

Before we use the dataset directly as the training data, we need to select the region of interest which contains an object we would like our model to recognize. We tried several ways to select such region.

At the beginning, we considered using corner detection directly on the original image to select ROI and at the same time extract features. However, as shown in Figure (?), extracting corner features based on the Harris Corner Detection and Adaptive Non-maximal Suppression method citebrown2005multi is not so helpful for following reasons: firstly, different pattern can be found in a single kind of balls so that geometric corner feature doesn't provide enough information for classification; secondly,

the complicated background may generate large amount of features which are irrelevant to the ball.

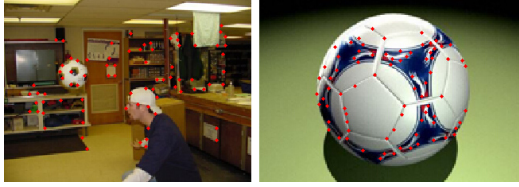


Figure 2. Corner Detection Results

Then we tried several methods for circle detection including (Atherton T.J., 1999), (Victor A., 2006), (D’Orazio T., 2004) to pick up the circular region and build the training set. One of the robustest approach is Circle Hough Transform (CHT), which takes the image M and desired radius range $[r_{min}, r_{max}]$ as inputs and return the positions of circle centers. As shown in Figure (?), this method performs good for the standard images with a single object occupying the whole image but has a poor performance when there are other ball-like distractions such as heads and the tree holes on a bowling ball, or only partial of the object, rather than a whole circle, appear on the image. It’s also possible for CHT to return some position with high confidence for some circular patterns inside the ball. We found it was really hard to pick up the ROI just by the shape and without any referring to the features inside the region.



Figure 3. CHT Detection Results

Thus, we finally decided to pick up ROI for each image in dataset manually. To increase the number of instances in our training data, for each image, let the user chooses the ROI with a movable, resizable rectangle and position it interactively using the mouse. And then shift the area a little in one direction, thus a raw image will create at most nine images and at least one image for training. Finally, the picked area is resized to a 40×40 square. The algorithm is presented in Algorithm 1.

Algorithm 1 Select ROI Manually

Input: image I
 Initialize accumulator matrix $M = 0$.
for each edge point (i, j) in image I **do**
 for each point (m, n) in image I **do**
 if $r_{min}^2 < (m - i)^2 + (n - j)^2 < r_{max}^2$ **then**
 $M(m, n) = M(m, n) + \phi_{m,n}$
 end if
 end for
end for
Return: positions of local maximum points in accumulator matrix M

2.3. PCA

The number of features we get from last step using HOG is 900. Since the number of instances we use is about 4000 which is not large for 900 features, the training data is more likely to cause overfitting. On the other hand, to improve efficiency and decrease the time for training, it’s necessary to reduce the dimension first. On the training dataset, we compute covariance matrix Σ and data mean (average over all rows of X). Next, we choose the d most important PCA basis vectors as the new training dataset. Then we should use Σ and X got previously on training dataset on the test data.

Acknowledgments

This project is supported by Upenn CIS 519. Thanks Eric Eaton and Xiaoxiang Hu for their instruction. Thanks MATLAB for providing so many useful toolbox.

References

- Atherton T.J., Kerbyson D.J. Size invariant circle detection. pp. 795–803, 1999.
- Brown, Matthew, Szeliski, Richard, and Winder, Simon. Multi-image matching using multi-scale oriented patches. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pp. 510–517. IEEE, 2005.
- Dalal, Navneet and Triggs, Bill. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pp. 886–893. IEEE, 2005.
- D’Orazio T., Guaragnella C., Leo M. Distant A. A new algorithm for ball recognition using circle hough transform and neural classifier. pp. 393–408, 2004.

Victor A., Carlos H.G., Arturo P. Raul E.S. Circle detection on images using genetic algorithms. pp. 652–657, 2006.