Pedestrian Detection

Víctor E. Ríos Martínez

Pedestrian Detection

Implementation on the HOG descriptor with a support vector machine to detect people in a determined area.

Contents

Introduction

Detecting humans is a task that has a great importance and value in the field of computer vision.

In the last decade were published some of the most important papers in this field by Dalal and Triggs.

The HOG descriptor is one of the easiest techniques to implement.

Objective

The objective of this project is to get a basic understanding of the techniques used to detect people and to get and overview of the state of the art.

Pedestrian detection will help to detect people in streets and other places, such banks, stores, airports, etc.

Justification

Pedestrian detection is very useful in video surveillance systems, it could be used in banks, stores and other places.

Also it can be used to detect if there is people in the street.

Development

This work is based in the HOG descriptor and SVMs.

The HOG (Histogram of Oriented Gradients) descriptor is capable to describe the features of an object by computing the gradient directions of each pixels, and grouping them into blocks, which are used to normalize and group the histograms.

The SVM is a supervised learning algorithm used to classify. It constructs a hyperplane to separate objects by its features.

HOG descriptor

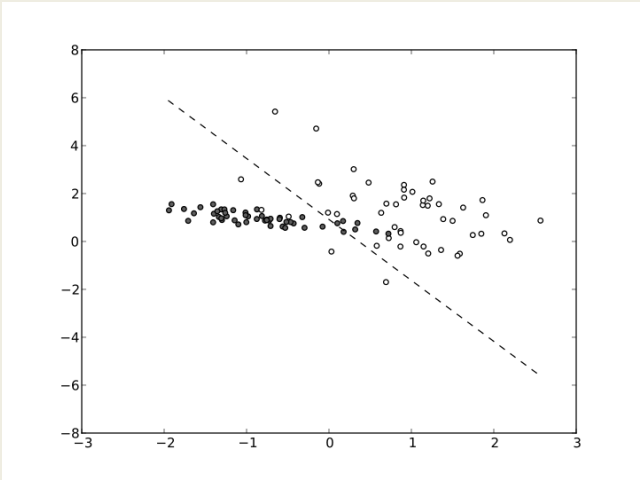
The procedure is the following:

1. The image is divided in cells.
2. Compute gradient orientation of each pixel.
3. The cells are discretized taking into account the gradient orientation.
4. Each cell's pixel votes for the orientation based on the gradient orientation.
5. The cells are grouped in blocks.
6. Histograms are normalized for each block.
7. The descriptor is the set of normalized histograms.

SVM

The Support Vector Machine needs to be trained to be able to classify objects.

The OpenCV library has a method that returns the coefficients for a trained SVM, which will determine whether an object with certain HOG features is a human or not.



Implementing HOG descriptor in python

This is the code that computes the HOG descriptor with OpenCV in python.

```
# Create a hog descriptor object.
hog = cv2.HOGDescriptor()

# Set the coefficients for the SVM
hog.setSVMDetector(cv2.HOGDescriptor_getDefaultPeopleDetector())

# Detect people in the image
hogParams = {'winStride': (0, 0), 'padding': (0, 0), 'scale': 1.05}
found = hog.detectMultiScale(bigger_img, **hogParams);
```

Region of Interest

Sometimes is required only a region of a video or image to be analyzed, the region of interest. I can select that region in a window and make the analysis only in this area.

```
def selectROI(event, x, y, flags, params):
    global roi
    if event == cv2.EVENT_LBUTTONUP:
        roi.append((x, y))
        if len(roi) == 2:
            cv2.rectangle(img, roi[0], roi[1], (255, 0, 0), 2)
```

Gaussian filter and background subtraction

These techniques were rejected because they do not seem to affect positively the results of the analysis, the result gets worse. However this is part of the code used for it.

```
blur = cv2.GaussianBlur(img, (3,3), 0)
fgmask = fgbg.apply(img, None, 0.5)
fgmask = cv2.morphologyEx(fgmask, cv2.MORPH_OPEN, kernel)
```

Showing results

Drawing the bounding boxes.

```
# A subroutine to get different colors
col = colorsfile.color(noofpersons)

# Bounding box
for i in xrange(0, noofpersons):
    person = found[0][i].tolist()
    x=person[0]
    y=person[1]
    xf=person[0]+person[2]
    yf=person[1]+person[3]
    cv2.rectangle(img, (x, y), (xf, yf), col[i], 1)

cv2.putText(img, "Number of persons: " + str(noofpersons), (0, 30), cv2.FONT_HERSHEY_SIMPLEX, 1, (255, 255, 255), 2)
```

Results

The code receives the input video as the first argument, if it is not passed then the camera is used.

The first frame is used to select the region of interest in the video. Before starting the analysis it is possible to restart the selection with the key r.

The analysis will be shown in another window.

Detecting people

The region of interest is the area enclosed in a blue rectangle.

Detecting people

Another region of interest.

Detecting people

A woman being detected.

Future work

There are some limitations to the HOG descriptor and the parameters used by the getDefaultPeopleDetector method.
The SVM can be trained to detect people in a better way handling with occlusion, accessories and clothes.
Tracking the people would be useful to identify them and other applications.
The user interface for this implementation should be improved to get a better user experience.
Normalization, filters, and background subtraction should be tested to see if they are really beneficial in pedestrian detection.

Conclusions

HOG descriptor along with the SVMs works pretty well detecting people but there are still some problems that could be enhanced as mentioned before.

Using background subtraction and techniques for noise reduction didn’t yield better results.

There is still work to do to detect people with simple and quick methods.

References

Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE.

Histogram of Oriented Gradients (HOG) Descriptor. (n.d.). Retrieved May 23, 2015, from https://software.intel.com/en-us/node/529070

Rosebrock, A. (2015, March 9). Capturing mouse click events with Python and OpenCV - PyImageSearch. Retrieved May 25, 2015, from http://www.pyimagesearch.com/2015/03/09/capturing-mouse-click-events-with-python-and-opencv/