Histogram of oriented gradients (HOG)

In this text, we introduce a method for creating feature vector that can be used for object classification. So far we have introduced features computed from the object area, length of object border, squarness, or circumference. When we use such features, we needed the segmented objects, and we ignore the changes of the brightness within the object areas. Today, we will use a method that does not require the segmented areas but it uses the pixel values directly. As it can be derived from its name, we will use gradient and histograms - the values from the histograms are used as the features. You know the gradients from the exercise about edge detection, the histograms were presented in Hough transform. The illustration of the HOG method is in Fig. 1. Originally, this method was proposed for human detection [1]. The algorithm can be divided in two steps.

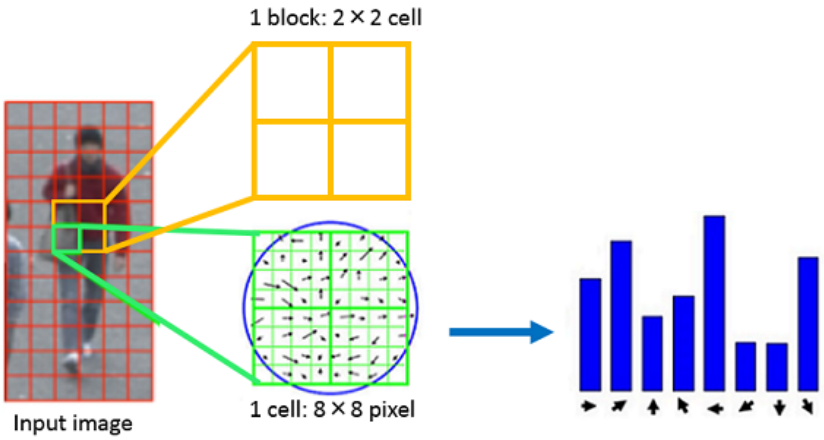

Figure 1: HOG Illustration [http://www.lsi-contest.com/2017/shiyou_3-1e.html]

Step 1: Compute the Gradient Image

For each image pixel (x, y), compute the orientation of gradient as

$$\varphi(x,y) = \arctan\left(\frac{f_y(x,y)}{f_x(x,y)}\right), \tag{1}$$

and the size of gradient as

$$e(x,y) = \sqrt{f_x^2(x,y) + f_y^2(x,y)}\,, \tag{2}$$

where $f_x(x,y)$ and $f_y(x,y)$ are the differences of brightness in the x and y direction, respectively. For their computing, we can use the following equations

$$f_x(x,y) = f(x+1,y) - f(x,y)\,, \tag{3}$$
$$f_y(x,y) = f(x,y+1) - f(x,y)\,. \tag{4}$$

Step 2: Create the Histograms

Let the image be split into blocks of size $B_x \times B_y$, and let the blocks be split into cells of size $C_x \times C_y$ (see 1). Create a histogram of gradient orientations in each cell. Divide the range of orientations ($0-180°$ or $0-360°$, when to use the first and when the second one?) into N bins (for example, in Fig. 1, we have $N = 8$ and we use the interval $0-360°$, each bin in histogram represents the interval of angles with the step $45°$, i.e. $\langle 0-45), \langle 45-90), \ldots$). For each image pixel (x, y), add its gradient size $e(x, y)$ into the bin representing the angle interval in which the gradient orientation $\varphi(x, y)$ of the pixel belongs. Create such a histogram for each cell, and normalize the histograms within the blocks. The normalization is done to have the method less sensitive to the light variation in the image.

As a result of this method, we get a feature vector with the histogram values. The size of the feature vector depends on the number of bins and on the number of cells in the image. For example, let us have an image of the size 80×80, the size of cell is 8×8 pixels, the size of block is 16×16 pixels (or 2×2 cells), the number of bins is $N = 8$. Each cell

provides 8 values from the histograms, it is 32 values from the block. According to the size of block, we have 5 verical and 5 horizontal blocks, which leads to the feature vector of the size $32 \cdot 5 \cdot 5 = 800$.

YOUR TASK

Implement the HOG method.

References

[1] Dalal, Navneet and Triggs, Bill: Histograms of Oriented Gradients for Human Detection, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1, pp. 886–893,(2005)