Programming Homework 4

CS540

June 22, 2019

1 Instruction

Please submit your output files and code on Canvas → Assignments → P4. Please do not put code into zip files and do not submit data files. The homework can be submitted within 2 weeks after the due date on Canvas without penalty (50 percent penalty after that).
Please add a file named "comments.txt", and in the file, you must include the instructions on how to generate the output, for example:

- Data files required: train.csv, test.csv. Run: main.jar.
- Data folder required: data/train1.png ... data/train100.png . Compile and Run: main.java.

2 Details

All the requirements are listed on the course website. The following is only an example workflow to solve the problem.

1. Download the complete webpage, not individual images.
2. Read an image, compute $\frac{1}{3}R + \frac{1}{3}G + \frac{1}{3}B$ for each pixel.
3. Convolve the image with the derivative filter or the Sobel filter (below) to get the gradient and compute the magnitude G and direction Θ for each pixel.

$$W_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, W_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\nabla_x = I * W_x, \nabla_y = I * W_y$$

$$G = \sqrt{\nabla_x^2 + \nabla_y^2}$$

$$\Theta = \arctan\left(\frac{\nabla_y}{\nabla_x}\right), \text{ in most programming languages } atan2\left(\nabla_y, \nabla_x\right)$$

Also try $atan2\left(\nabla_x, \nabla_y\right)$. The function is different for different languages.
Important: use the atan2 function, not the arctan function to find the angle. The usual arctan function

in most programming languages is not the inverse tangent function.

G and Θ are 150×150 matrices.

When you implement convolution with image matrix indexed from 0 to 149 and filter matrix indexed from 0 to 2, the formula is slightly changed for $I * W$.

$$(I * W)_{jj'} = \sum_{t=-1}^{1} \sum_{t'=-1}^{1} W_{t+k,t'+k} X_{j-t,j'-t'}$$

The k is the number so that the filter size is $(2k+1) \times (2k+1)$, here for Sobel filters, $k = 1$.

4. Divide the image into 10×10 regions (225 of them), in each 10×10 region (call it R), sum the magnitude of the gradients with similar directions. Similar direction means they belong to the same bin according to the following binning (for 8 bins):

$$B = \left[0, \frac{\pi}{8}\right), \left[\frac{\pi}{8}, \frac{\pi}{4}\right), \left[\frac{\pi}{4}, \frac{3\pi}{8}\right), \left[\frac{3\pi}{8}, \frac{\pi}{2}\right), \left[\frac{\pi}{2}, \frac{5\pi}{8}\right), \left[\frac{5\pi}{8}, \frac{3\pi}{4}\right), \left[\frac{3\pi}{4}, \frac{7\pi}{8}\right), \left[\frac{7\pi}{8}, \pi\right]$$

There is a formula to compute this directly:

$$\Theta'_{i,i'} = \begin{cases} \Theta_{i,i'} & \text{if } \Theta_{i,i'} \geqslant 0 \\ \pi + \Theta_{i,i'} & \text{if } \Theta_{i,i'} < 0 \end{cases}$$

$$\text{bin} = \text{floor}\left(\frac{\Theta'_{i,i'}}{\frac{\pi}{8}}\right) \bmod 8$$

The histogram for bin $i = 0, 1, 2, ..., 7$ is:

$$h_j = \sum_{(i,i') \in R} G_{ii'} \mathbb{1}_{\left\{\Theta'_{i,i'} \in B_j\right\}}$$

5. Output the flattened histograms (one image on each line).

Note: In the previous version of the file, the way bins are defined is incorrect. The directions 0 and π and $-\pi$, or $\frac{\pi}{4}$ and $-\frac{3\pi}{4}$, or $\frac{\pi}{2}$ and $-\frac{\pi}{2}$ etc should be considered the same direction. In the previous incorrect version, the bins are:

$$B = \left(-\frac{\pi}{8}, \frac{\pi}{8}\right], \left(\frac{\pi}{8}, \frac{3\pi}{8}\right], \left(\frac{3\pi}{8}, \frac{5\pi}{8}\right], \left(\frac{5\pi}{8}, \frac{7\pi}{8}\right],$$
$$\left(\frac{7\pi}{8}, -\frac{7\pi}{8}\right], \left(-\frac{7\pi}{8}, -\frac{5\pi}{8}\right], \left(-\frac{5\pi}{8}, -\frac{3\pi}{8}\right], \left(-\frac{3\pi}{8}, -\frac{\pi}{8}\right]$$

If you completed the homework with this method of binning, it's effectively only 4 bins instead of 8, but it is okay to submit this version of the output. I apologize for the mistake.