Computer Vision 1: Homework 4

Deadline 22.11. 12:15

Important: Submit your programming solutions through Moodle. The deadline for submitting your work is always on Thursday, at 12.15, the week after handing out the homework.
For other, non-programming homework, bring your solution with you to the exercise class. For each homework problem, one student will be chosen at random to present their solution.

Programming tasks.

Some computational attention systems use the concept of difference-of-Gaussians (DoG) pyramids to calculate contrasts. As noted in [1], "to distinguish bright objects on dark background from dark objects on bright background, the contrast computation is separated into on-off and off-on contrasts, corresponding to cells in the human visual system that respond only to one type of contrast respectively". On-off and Off-on contrasts may be calculated using DoG pyramids as follows. First we create two Gaussian pyramids P_1 and P_2 using σ_1 and σ_2, respectively. Usually, $\sigma_1 < \sigma_2$, and then P_1 is called the "center pyramid" and P_2 the "surround pyramid", due to the respective sizes of the receptive fields of the filters. Then, each of the layers of each pyramid is subtracted from the other. The ith layer of the on-off contrast pyramid is given by

$$\texttt{On-Off}(i) = P_1(i) - P_2(i),$$

and the ith layer of the off-on pyramid is

$$\texttt{Off-On}(i) = P_2(i) - P_1(i).$$

Figure 1: A test image.

- Use the test image shown in Figure 1 from Moodle (`test-image-MSRA-3_97_97769.jpg`). Convert the test into an intensity image I by 1) converting it to floating point data type, 2) calculating $I = (R + G + B)/3$.
- Calculate the on-off and off-on DoG pyramids for I using $L = 4$ layers, and visualize the results. Set $\sigma_1 = 3$, $\sigma_2 = 9$. Use `skimage` and `matplotlib` for displaying images. Do you observe bright objects on dark background highlighted in the on-off pyramid images, and vice versa for off-on? What is the meaning of different layers?

Other tasks.

1. Consider an image X of arbitrary size, where $x_{i,j}$ denotes the pixel at coordinates (i, j). We are given two kernels of the same size, $K, L \in \mathbb{R}^{(2m+1)\times(2m+1)}$ where $m \in \mathbb{N}$.

Recall from the lecture notes that the cross correlation $G = K \otimes X$ can be written in a sum form, such that the (i, j)th element of G is

$$(K \otimes X)_{i,j} = \sum_{u=-m}^{m} \sum_{v=-m}^{m} k_{u,v} \cdot x_{i+u,j+v}.$$

Using the definition, prove that cross-correlation satisfies

(a) homogeneity, i.e., that for $\alpha \in \mathbb{R}$, $(\alpha K \otimes X) = \alpha(K \otimes X)$, and

(b) additivity, i.e., that $((K + L) \otimes X) = K \otimes X + L \otimes X$.

Show the equalities hold for an arbitrary element (i, j) of the result. You do not need to consider padding. Based on the result, what can you say about the linearity of convolution?

2. We are given a 3-by-3 image X, and a 2-by-2 kernel K, represented as matrices

$$X = \begin{bmatrix} x_0 & x_3 & x_6 \\ x_1 & x_4 & x_7 \\ x_2 & x_5 & x_8 \end{bmatrix}, \quad K = \begin{bmatrix} k_0 & k_2 \\ k_1 & k_3 \end{bmatrix}.$$

Without using any padding of X, the result of the cross correlation will be a 2-by-2 matrix

$$K \otimes X = \begin{bmatrix} r_0 & r_2 \\ r_1 & r_3 \end{bmatrix}.$$

Show that any element r_i can be represented as a matrix multiplication $r_i = h(i)^T \cdot v$. Here, $h(i)$ is some 9-by-1 vector to be determined that contains elements of K and zeros. Furthermore, v is the 9-by-1 vector obtained when "flattening" the values of the image as $v = \begin{bmatrix} x_0 & x_1 & \dots & x_8 \end{bmatrix}^T$.

Hints:

- Begin by writing an expression for r_i as a sum of products of elements of K and X.
- Convert the sum into a matrix multiplication of v from the left by some vector $h(i)^T$; and write out an expression for the vector $h(i)$.

References

[1] Simone Frintrop, Thomas Werner, and German Martin Garcia. Traditional saliency reloaded: A good old model in new shape. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.