Computer Vision

James L. Crowley

M2R MoSIG option GVR — Fall Semester

9 nov 2017

Lesson 4 — Exercises

1) Discrete Scale Space:
Let *P(x, y)* be an image array of size *1024x1024* pixels, where (x, y) are integers. A discrete scale space is computed as:

$$P(x,y,k) = P(x,y) * G(x,y,\sigma_k)$$

where

$\sigma_k = 2^k$ For k=0 to 10

a) Write the formula for the 2D sampled Gaussian filter, including window and normalization factor.
b) Determine the minimum window size of the Gaussian filter for each value of k from 0 to 10.
c) Write the formula for the convolution of the 2D Gaussian with the image.
d) What is the computational cost in terms of multiplications and additions for computing the space using 2D convolutions for each value of k?
e) Write the formula for the convolution using a succession of 1-D Gaussian filters with the image.
f) What is the computational cost in terms of multiplications and additions for computing the space using 1D convolutions for each value of k?

2) Computing a multi-scale image pyramid.

The images in the discrete scale space can be re-sampled using

$$P(i, j, k) = S_{\Delta xk}\{P(x,y,k)\} = P(i\Delta x_k, j\Delta y_k, k)$$

using a sample size of $\Delta x_k = 2^k$

However, there was no need to compute the convolution at image positions that are removed by resampling. Thus we can optimize the convolution by skipping these samples during the convolution. This can be written as:

$$P(i,j,k) = \sum_{u=-R_k}^{R_k} \sum_{v=-R_k}^{R_k} P(u - i\Delta x_k, v - j\Delta x_k) G(u,v,\sigma_k)$$

for integer i, j from (0,0) to (2^{10-k} , 2^{10-k}) and k from 0 to 10 and where R_k is the half size of the Gaussian window at level k. ($N_k=2R_k+1$)

a) What is the size (number of rows and columns) of each resampled image for k=0 to 9?
b) What is the computational cost of this operation using 2D convolutions?
c) What is the computational cost of this operation using a sequence of 1D convolutions?

3) Note that using $\Delta x_k = \sigma_k$ in the above algorithms keeps the algebra simple, but results in substantial noise from aliasing. This can be remedied by using $\sigma_k = 2^{k+1}$ and $\Delta x_k = 2^k$. What does aliasing do to the images in the pyramid?