Computer Vision

James L. Crowley and Nachwa Aboubakr

MoSIG M2 Fall Semester
Project 3 22 October 2020

Face Detection with a sliding window detector at multiple scales.

The objective for this exercise is to use your best MLP face detector constructed last week to detect faces at multiple scales using windows from a Gaussian Pyramid. You will first construct a sliding window face detector using your best MLP, and then optimize this detector using a full octave Gaussian Pyramid.

This exercise is composed of four parts.

1) Write a program to construct a scale-invariant Gaussian pyramid, using the algorithm shown in section 3.2 of the course notes. Demonstrate the impulse response of your pyramid by creating a 512 x 512 image with a single non-zero pixel at the center position (256x256). Display the contents of central 13 columns (cols 250 to 262) from row 256 from each channel of each level of your pyramid. Do this for $\sigma_0=1$ and $\sigma_0=\sqrt{2}$ and compare the results.
2) Write a program to extract and flatten a sliding window from an image over a range of sizes from 10 x 10 to 40 x 40 using a scale factor of 1.2. Each window must be transformed to the standard size of input vector for your MLP face detector from last week. Use your best MLP to label each window as face, or not face. Report precision, recall and computing time for evaluation with the images in folds 9 and 10 of FDDB.
3) Adapt your sliding window detector to extract and flatten windows of sizes from 10 x 10 to 40 x 40 from each level of your scale invariant pyramid, using a scale factor of 1.2. Use this program to detect faces from all images in your pyramid.
4) Compare precision, recall and computing time for the face detection from an image and from a pyramid using the images in folds 9 and 10 of FDDB

Document your work in the Jupyter Notebook by commenting it and send the .ipynb file to: James.Crowley@inria.fr, Nachwa.Aboubakr@inria.fr, Yangtao.Wang@inria.fr. Results are due before class on Thursday 5 nov.