Exercises, Lectures 6, Intelligent Multimedia Systems

Master AI, 2012

Exercise 1. Integral Images and Haar Features

The integral image is used as an effective way of calculating the sum of pixel values in a given image. The integral image is used by the Viola-Jones Object Detection Framework to efficiently calculate the Haar features. Consider the following image patches:

5	2	2
6	4	5
2	3	6

Figure 1 Intensity values of image patch M.

1	1	1	1
1	1	1	1
0	0	0	0
0	0	0	0

Figure 2 Intensity values of image patch N.

a) Compute the integral image s(x,y) using s(x,y) = i(x,y) + s(x-1,y) + s(x,y-1) – s(x-1,y-1) for patch M, where i(x,y) is the pixel value at location (x,y).
b) Calculating the sum of pixels in (sub)rectangle of the original image patch can be done by computing i(x',y') = s(A) + s(D) – s(B) – s(C). Calculate the sum of pixels for the lower, right 2x2 sub-rectangle for patch M.
c) What is the time complexity?
d) Compute the integral image for patch N.
e) For which of two-rectangle Haar features, an edge is found?
f) Compute the difference of the sum of pixels for the upper and lower (2x4) rectangles.

Exercise 2. Sliding Window Approach

Sliding window approach is a popular way to detect objects in images.

a) What is the basic pipeline for window-based object detection?
b) What are the advantages of a sliding-window approach?
c) Given an image of 256x256, how many windows are required to detect objects for 8 different orientations and 6 scales?
d) Assume that for a strong classifier (e.g. non-linear SVM), the time for window classification is about 0.01 seconds. How many hours does it take to detect objects in 10,000 images?
e) How can one reduce the number of bounding boxes for detection? Now, how long does it take to detect objects in 10,000 images?

Exercise 3. Facial Expressions

Facial expressions may indicate certain intentions of humans and their corresponding emotions. Facial expressions can be described by action units.

a) What is an action unit?
b) What are the six basic emotions of humans?
c) What are the facial characteristics of a happy person?

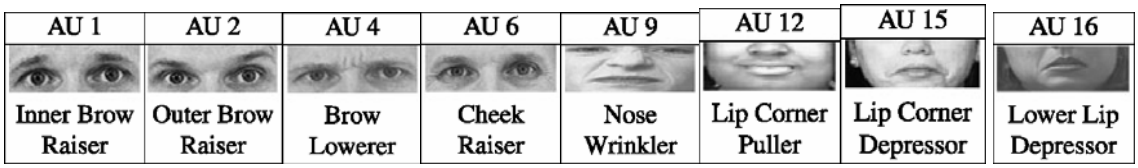

Figure 3: A subset of action units.

d) Considering Figure 3, which AU's indicate positive emotions/expressions?
e) Which ones of them indicate negative emotions/expressions?
f) Which of the action unit or action unit combination can describe sadness expression?
g) Which of the action unit or action unit combination can describe a smile?
h) Which features are important to recognize a real enjoyment smile?