

## Exercises, Lectures 4-5, Intelligent Multimedia Systems

Master AI, 2012

### Exercise 1. Object Tracking

After initialization (the initial image location and size of the object is provided) objects can be *tracked* over time by computing their locations in subsequent frames.

- a) What is template matching and how can this technique be used for tracking?
- b) Could you define a pixel-wise similarity measure for template matching?
- c) What are the possible image transformations between the template and possible candidates? What is the search area?
- d) What are the pros and cons of template matching for object tracking?
- e) What is the difference between the similarity measure of template matching and mean-shift?

### Exercise 2. Edges and Corners

Consider the following image patches:

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

Figure 1 Intensity values of image patch A.

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

Figure 2 Intensity values of image patch B.

- a) What are the interest points of image patch A?

- b) Compute the gradient and the Harris corner response of image patch A (using a simple derivative filter e.g.  $[-1\ 1]$ ).
- c) What are the interest points of image patch B?
- d) Compute the gradient and the Harris corner response of image patch B (using a simple edge filter e.g.  $[-1\ 1]$ ).
- e) Compute the eigenvalues of  $M$  for patch B where  $M$  is the  $2 \times 2$  matrix computed from the image derivatives i.e. second moment matrix (autocorrelation matrix).

### **Exercise 3. Object Descriptors**

Object recognition is important in computer vision. Objects can be *recognized* by considering their presence and locations in unknown images. The standard approach is to use image descriptors in a bag-of-features approach.

- a) What is an image descriptor?
- b) What are the advantages of using histograms as image descriptors? What about quantization (# of bins)?
- c) What kind of image structures are descriptors made of?
- d) Compute the histogram of oriented gradients and pixel values for patch A given in Figure 1.
- e) Compute the histogram of oriented gradients and pixel values for patch B given in Figure 2. Which histogram is more discriminative?
- f) What is the SIFT descriptor?
- g) Is the SIFT descriptor invariant under a change in (in-plane) rotation of the object? Please explain.
- h) What is a color SIFT descriptor? Give an example of robust color SIFT descriptor.

### **Exercise 4. Back-of-Features**

The bag-of-features approach is an easy way to represent images for image classification.

- a) What is the difference between dense and point sampling?
- b) What are the basic steps of the bag-of-feature approach?
- c) What are visual words and how is the visual vocabulary computed?
- d) What are spatial pyramids and why are they useful?
- e) Using the back-of-features approach with SVM for object recognition, do you expect that certain objects may be confused during recognition? Give examples.
- f) Do you think that context is important for object recognition? Can you give an example of certain objects?