Technische Universität München
Fakultät für Informatik
PD Dr. Slobodan Ilic

WS 16/17
Exercise Sheet 3
November 10, 2016

Exercises in Tracking & Detection

The objective of this exercise is object detection using feature descriptors. For this purpose SIFT and HOG descriptors will be used. You are given four scene images *test_shell1.jpg*, *test_shell2.jpg*, *test_pot1.jpg* and *test_pot2.jpg* and two object images *shell.jpg* and *pot.jpg* that can be downloaded from the TDCV website.

The task is to develop matching algorithms using SIFT and HOG features. Input to the algorithms is one test image and the object image and the output is the bounding box around the detected object in the test image. For extracting and matching SIFT and HOG descriptors use VLFeat library (http://www.vlfeat.org/).

Exercise 1 Matching with SIFT

In this exercise SIFT will be used for finding the object in the scene. Given the scene image and the image of the object to be detected, the following steps in your algorithm have to be performed:

- Extraction of SIFT features from the scene and the object image and visualization (position, scale and orientation of the feature points have to be shown).
- Matching SIFT descriptors and visualization of putative matches (visualization should draw lines between matched features).
- Removing outliers (use function estimateGeometricTransform from Computer Vision System Toolbox with assumed affine transformation between images). Visualize pruned matches.
- Draw bounding box around detected object.

The method has to be tested on all images and conclusions drawn based on the obtained results. The answers to the following questions are expected:

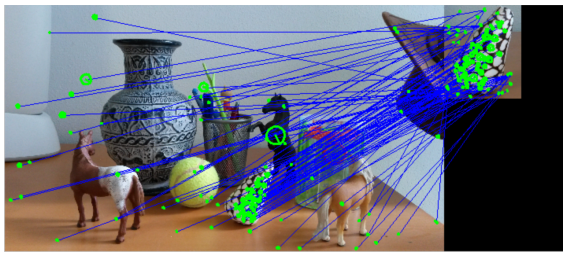
- What is the main difference between two given objects and in which images they can be detected using SIFT features?
- Do the occlusions obstruct detection?

The examples of the possible outputs of steps 2 and 3 are shown in Fig. 1a and in Fig. 1b.

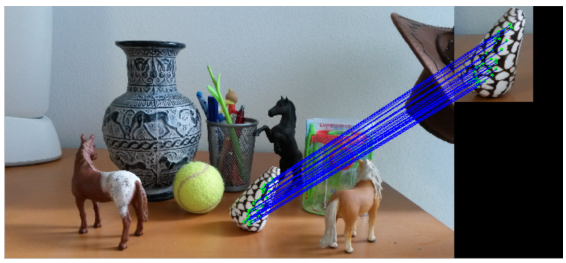
Exercise 2 Matching with HOG

In this exercise you HOG will be used for finding the object in the scene. Given the scene image and the image of the object to be detected, the following steps in your algorithm have to be performed:

- Extraction of HOG features from the scene and the object image.



(a) Putative matches of SIFT descriptors.



(b) Pruned matches of SIFT descriptors.

Figure 1: Matching with SIFT

Figure 2: Matching with HOG

- Match HOG descriptor of the object to the scene using sliding window at multiple scales.
- Draw bounding box around detected object with the maximal matching score.

For convolution of the HOG descriptor of the object image with the HOG image of the scene use *scores = vl_ nnconv(hog, w, [])*, which is part of *http://www.vlfeat.org/matconvnet* library. The examples of the possible outputs of the detection is shown in Fig. 2.

The answers to the following questions are expected:

- What is the main difference between two given objects and in which images they can be detected using HOG features?
- Do the occlusions obstruct detection?
- How the change of the number of scales influences the result?

Note. For the HOG exercise resize scene image to 1/4 and object image to 1/8 for the optimal speed and performance.