Computer Vision

Exercises of Lab 4

Exercise 4.1: Camera Calibration

The goal of this lab exercise is to calibrate a camera using multiple correspondences of 2D image coordinates and 3D world coordinates of a known object.

First, download the Camera Calibration Toolbox for MATLAB from [vision.caltech.edu](). Also, download [Peter Kovesis' Computer Vision Matlab functions]() – direct link [here](). Open Exercise4.1.m and change the variables VLFEATROOT, CALIBROOT, and MATLABFNSROOT to the unpacked paths at your hard drive.

Run the code and point out the four corners (clockwise from the upper left corner) in the first frame. Understand how the 2D homogenous coordinates are stored in the variable p.

Your task is to use the SIFT feature descriptor and feature matching to automatically find the corresponding corners in the subsequent frames. For frame f=2:F, this can be done with the following steps:

1. Find matches between SIFT features in the current frame f and a previous frame. Use the function vl_ubcmatch(f_old,f_new) to return these matches.
2. Estimate a homography based on the matching points, describing the transformation from the previous frame to frame f. Use the function ransacfithomography(p_old, p_new, 0.1) to find the best fitting 2D homography. For this problem, 0.1 is a suitable distance threshold for the RANSAC algorithm.
3. Apply the transformation of the previous corner coordinates to the current frame f and store them in p(f,:,:).

You can visualize the results with the code provided in line 60-68.

Try to apply the above method by matching frame f and f-1, and then try always matching frame f to frame 1. Which method works best, and why?

Once your corner matches look convincing enough (Figure 1), run the camera calibration (from line 73), and inspect Figure 2 (3D) carefully. Click the button "Switch to world-centered view". When would you use a camera-centered view and when would you use a world-centered view?

Plot a projection of an (x,y,z)-axis on top of the images. Use the estimated intrinsic and extrinsic parameters from the camera calibration. Slide 8 in Lecture4.1-2_StereoEpipolarAndImageAlignment.pdf describe the order of transformations required for transforming from 3D world coordinates to 2D image coordinates. The parameters returned from the camera calibration toolbox are described here.

NOTE: The convention of the extrinsic parameters in the toolbox differs from the convention in slide 13-14. The toolbox uses the following convention:

$$P_{camera} = R \times P_{world} + T$$

where R is a rotation matrix, T is a translation vector, and P_{world} and P_{camera} are non-homogeneous 3D coordinate vectors. How does this influence the formulas in slide 8?