2 Instructions for Second Assignment

For this assignment you must solve at least 3 exercises. You can choose any 3 of the proposed exercises below. Submit a report of no more than 5 pages (IEEE style) with the general description of your solutions without including the source code. You can use any function provided by OpenCV or any other vision library. In particular for the *Segmentation of Disparity Maps* algorithms, remember that the idea is to detail the strengths and weaknesses of each confidence measure.

3 Exercises

3.1 Problem 1

Map Surface of Cylindrical Shape into a Plane. Provide a solution for the following task: Given is a set of images showing one cylindrical object such as a can or a bottle (see Figure 1). The occluding contour of the cylindrical object may be generated interactively rather than by an automated process.

Figure 1: Left: Example of an input image of a cylindrical shape with glued-on label. Middle: Approximate location of occluding contours of cylinder, generated by manual interaction. Right: Extracted flat label from multiple input images.

Use multiple images of the cylindrical object showing an "interesting" part of the surface (such as the label in Figure 1) and calculate an image showing this surface part as being mapped into the plane (also using image stitching for merging segments from multiple taken images).

3.2 Problem 2

By combining calibrated camera with an estimation of homographies between planes it is possible to compute the camera's pose (rotation and translation) if the image contains a planar marker object. This marker object can be almost any flat object. Augmented reality (AR) is a collective term for placing objects and information on top of image data. The classic example is placing a 3D computer graphics model so that it looks like it belongs in the scene, and moves naturally with the camera motion in the case of video. Given an image with a marker plane as in the section above, we can compute the camera's position and

pose and use that to place computer graphics models so that they are rendered correctly.

Figure 2: Example of computing the projection matrix for a new view using a planar object as marker. Matching image features to an aligned marker gives a homography that can be used to compute the pose of the camera. Template image with a gray square (top left); an image taken from an unknown viewpoint with the same square transformed with the estimated homography (top right); a cube transformed using the estimated camera matrix (bottom).

3.3   Problem 3

Segmentation of Disparity Maps and Generation of Mean Distances. Use as input stereo data recorded from a car (see, e.g., data provided on KITTI, HCI, or EISATS). Apply a segmentation algorithm on disparity maps, calculated by a selected stereo matcher (e.g. in OpenCV), with the goal to identify segments of objects "meaningful" for the traffic context. Consider the use of temporal consistency for segments. Calculate mean distances to your object segments and visualize those in a generated video, which summarizes your results. See Figure 3 for an example.

3.4   Problem 4

Select or program two different stereo matchers and apply those on challenging HCI stereo sequences. See Figure 4 for examples. Select a confidence measure for calculated disparities and visualize confidence measure results as a graphical overlay on the input sequence for (say) the left camera. Discuss the "visible correlation" between low confidence values and shown situations in the input sequence.

Figure 3: Mean distances to segmented objects. Left: Distances to objects within the lane area only. Right: Distances to all the detected objects.
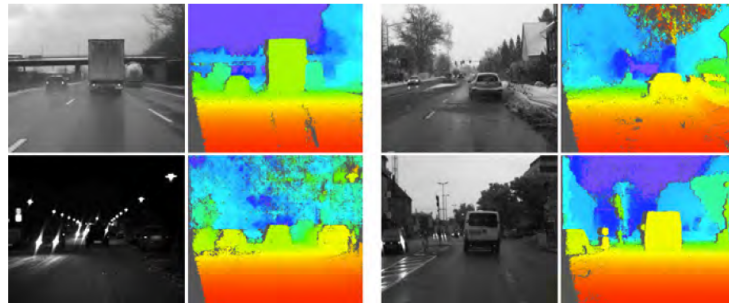

Figure 4: Challenging HCI stereo sequences and examples of calculated disparity maps (using iSGM).

3