Computer Vision - 3D Modeling

All printed documents are allowed. The different sections below are independent. Answers must be justified and should be synthetic.

1 Reconstruction of 3D Points (8 points)

1. **Precision;** Consider the stereo system sketched in figure 1 (top view). The configuration depends only on the angle α. The task considered here is the 3D reconstruction of points seen in both images. Because of the noise in the image point positions, the coordinates of the reconstructed points will not be perfectly precise. In particular, the errors in the reconstructed 3D coordinates may be differently large along the different coordinate axes. For the configuration shown in the figure, which coordinate will be most affected (the one corresponding to the upright axis or the one corresponding to the horizontal axis)?

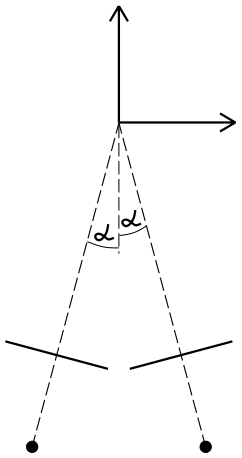

Figure 1: Figure related to question (1).

2. **Undetermined 3D reconstruction.** 3D points can be reconstructed from image points determined in the two image planes, for example by intersecting the lines of sight corresponding to these image points. However, there exist 3D points which cannot be reconstructed in this manner, i.e. the intersection of the lines of sight is not unique or not defined. For which 3D points is this the case (describe their locus, with respect to the optical centers of the two cameras)?

3. **Camera setups for 3D reconstruction.** Consider now the two scenarios shown in figure 2. The one on the left corresponds to a camera that takes two images while moving forward. The right one corresponds to a lateral camera motion. Which one of the two scenarios seems more appropriate for carrying out a 3D reconstruction, and why? (The answer may rely on that to question (c)).

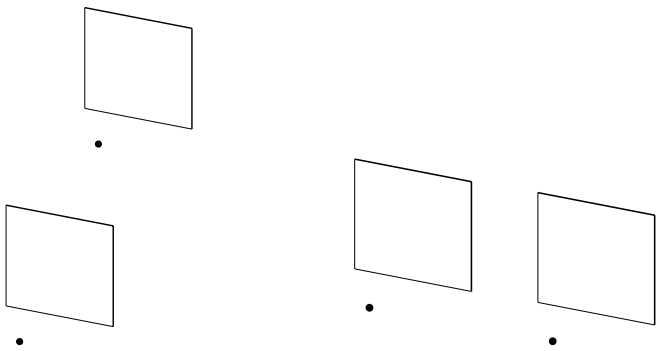

Figure 2: Figure related to question (3).

4. **3D reconstruction using several images.** The goal here is to reconstruct 3D points from their image points, determined in several images, instead of only two. In practice, it may happen that one or more of the image points are incorrect, i.e. are outliers, for instance due to a mistake in image matching. If such outliers are used for 3D reconstruction, the result may be bad.

   Describe a *robust estimation* procedure (see lecture notes) that could solve this problem. Briefly sketch (with words, no formulae required) the steps of the procedure, especially how to compute candidate estimates for a 3D point and how to compute a score for the quality of these candidate estimates.

2 Projective Geometry (6 points)

1. What are the differences between Euclidean, affine and projective geometries ?
2. Why is parallelism an affine notion ?
3. How many correspondences between pairs of points are required to estimate a homography in P^n ?
4. Write the equations given by a pair of corresponding points in P^2 and explain then how a homography can be determined in this space.

3 Human Pose Recovery (6points)

Recent visual acquisition systems allow real time human motion recovery. Among such systems, the Microsoft Kinect approach [1] consists in recovering the pose of humans by extracting, from a learned dataset, the most likely pose with respect to the camera observations.

1. The strategy followed is to use a learned dataset. Another strategy is to track a known parametric model of humans. Shortly explain these two strategies and emphasize there differences.
2. In the Kinect approach [1], studied during a lecture, what are the input and the output of the algorithm ?
3. What is the contribution of the approach with respect to the state-of-the-art ?
4. Would the photometric information be useful here ?
5. What are the limitations of such approach ?

References

[1] Jamie Shotton, Andrew FItzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, Andrew Blake Real-time Human Pose Recognition in Parts from Single Depth Images, *IEEE Conference on Computer Vision and Pattern Recognition*, 2011.