1 Short explanations

a) What does it mean that computer vision is an "inverse problem" (1 point)

b) What is perspective distortion and what causes it? How can an image of a sphere be affected by this type of distortion. (2 points)

c) What is a "homogeneous least squares" problem? How did we solve in class? (2 points)

d) In object detection and recognition histograms often used for describing objects. Please name two robust measures for comparing histograms and describe why they are more robust than the Euclidean distance, for instance. (2 points)

2 Image formation

a) What is the goal of camera calibration? Describe the main steps of the calibration algorithm that you learned in class. Which input data is required? (3 points)

b) Briefly describe the Bayer pattern and why many cameras use it. Why or why not is there an equal distribution of the colors? How can missing colors be obtained? (3 points)

3 Image processing

a) Name two assumptions justifying the use of Gaussian filters to reduce the noise in images. (2 points)

b) Describe how the Gaussian and Laplacian Pyramid are related computationally. Based on that, describe why the Laplacian pyramid provides a frequency decomposition of an image. (4 points)

4 PCA

a) Name two advantages of using a low-dimensional model obtained by PCA (1 point)

b) In practice people usually perform a singular value decomposition on the data instead of a direct eigendecomposition of the data covariance. What are the main reasons for this? How does one obtain the eigenvectors and eigenvalues of PCA through SVD? Please derive the equations. Assume the SVD of the normalized (mean-subtracted) data matrix $\mathrm{X} \in R^{M \times N}$ (M is the dimension and N is the number of data points) is given as $\mathrm{X} = \mathrm{U}\mathrm{S}V^T$. (3 points)

c) List at least two cases for which simple appearance-based matching methods for detection are likely to fail. (2 points)

5 Interest points

a) How can we characterize interest points and why are they userful? (2 points)

b) Name one undesirable property of the Harris interest point detector that can be compensated for by using the Harris-Laplace detector. How does the Harris-Laplace achieve this? (3 points)

c) Name one advantage and one disadvantage of using the color histogram of a patch as a local descriptor? (2 points)

6 Single-view geometry

a) How many degrees of freedom does a homography have? Describe intuitively where this number comes from. (2 points)

b) What is the criterion that RANSAC tries to optimize? (1 point)

c) In image stitching problems, all images should be transformed onto the same image plane,

which can be done by image warping. Here an inverse warping scheme is more appropriate than forward warping. What is the reason for this? What is the difference between these two warping schemes? (3 points)

7 Two-view geometry

a) In the epipolar geometry of two views, what are epipoles and epipolar lines? (2 points) Why are the epipolar lines useful? (1 point) Sketch the configuration of the epipoles and epipolar lines for stereo images. (1 point)

b) How many point correspondences are needed for estimating the fundamental matrix of two given views? (1 point) Algorithms for estimating the fundamental matrix usually do not operate on pixel coordinates as inputs. Briefly explain why and how one computes the inputs from the pixel coordinates. (1 point)

c) What is a main problem of window-based matching for stereo estimation? How can it be addressed? (2 points)

8 Motion

a) What is the difference between the problems of binocular stereo matching and optical flow estimation. In which circumstance does the optical flow problem correspond to stereo matching? (2 points)

b) Write down the optical flow constraint equation (OFCE). (1 point) What are the two assumptions behind it? (2 points)

c) What is the basic idea behind the Lucas-Kanade method for estimating a dense optical flow field with large (fast) motion? (2 points)

9 Bag-of-Words model and categorization

a) Describe briefly how you learn a bag-of-words feature representation. (3 points)

b) Suppose we want to categorize n images into c classes. How many codeword dictionaries do we need to build at least? (1 point)

c) What is the idea behind Bayesian decision theory respectively an optimal Bayes classifier? (2 points)

10 Deep Learning

a) You want to build a deep neural network for predicting homographies from given pairs of input images. Please explain how you would acquire appropriate ground truth data for this problem. (2 points) Which loss function would you choose for training and why? (2 points)

b) Imagine that an $\mathrm{m} \times \mathrm{m}$ filter is iteratively applied n times. What is the size of the receptive field in which surrounding pixels affect the output at a certain pixel? Specify a formula depending on m and n. (2 points)

c) Explain what architectural component distinguishes residual networks. (1 point) What is the benefit of residual networks? (1 point)