1.Short explanations

a).Briefly explain the terms “principal point” and “principle axis”.

princle points: camera centre at image plane

princle axis

b).How do you need to adjust the aperture to get a smaller depth-of-field(such that the range in focus is smaller)?

enlarge the size of the aperture to get a smaller depth-of-field.

The image will be brighter if we enlarge the size of the aperture, because more light comes in.

c).Pyramid representations such as the Gaussian image pyramid are often used to accelerate computer vision algorithms. Name at least one additional reason for using an image pyramid.

each level/subband contains only image structure of a particular range of spatial frequencies.

d). Name one blob detector and one corner detector

corner detector:Harris Detector

blob detector : SIFT

2.Image formation

a).when photographing with a pinhole camera, the image will appear blurry when the hole is either too small or too big. What are the reasons in both cases?

b).What are world, camera, and image coordinate systems and how are they related? Write down the matrices relating the different coordinate systems.

3. Image processing

a). Name one linear and one non-linear filter for image filtering.

Gaussian filter. Median filter.

b).To compute the output of a 3x3 filter at a certain pixel, we need to look at the 3x3 neighborhood of that pixel. Imagine that a m x m filter is incratively 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.

(m-1)*n+1 （我唯一看到的标准答案

c).How do we detect edges using 1st derivatives? What is a possible problem with this approach?

smooth with Gaussian; calculate derivative; find its local optimal.

This approach may cause Hysteresis.

d).Usually, we first smooth an image before calculating its derivatives. How can we simplify this process?

4. PCA

a).what is the advantage of using Singular Value Decomposition(SVD) instead of Eigendecomposition to solve the PCA objective? Briefly explain your answer.

we never have to explicit build and store the convaraince matrix.

explain: define

b).How are the eigenvalues of the covariance matrix related to the singular values of the mean-centered matrix? What do these values represent in terms of data statistics?

c).Explain why the principal components are computed as the eigenvectors of the largest eigenvalues of the covariance matrix.

5.Interest Points

a). Assume that we have an image and we want ro compute interest points. The Harris values for each pixel have already been computed in the grid below. Now we want to apply non-maximum suppression to Harris values. Thin out the values by applying non-maximum suppression with a window size of 3 x 3, while only suppressing strictly smaller values (use "<" as the comparison operator and not "<=").

Draw a circle around each remaining Harris value in the grid below.

1	2	3	2	1	1
1	1	2	3	3	2
2	3	3	2	1	2
5	7	3	5	2	1
6	3	6	6	2	3
4	2	1	3	5	4

b). Suppose we would like to find interest points in 3D images(e.g. in CT scans), for which there is an intensity value for each (x,y,z) voxel.Describe a generalization of the Harris point detector by describing the main steps of the method, including a test to decide when a voxel is an interest point.

6. Single-view geometry

a).A scene is captured by two images I1 and I2 such that we can relate the 2D image coordinates of both images by 3X3 homography H. How does the homography constraint our camera setup? Given a homography from image I1 to I2, how can we transform the pixel coordinates in image I1 into pixel coordinates of I2? Write down the constraint equation(pay attention to the right order).

only rotation

HX1 * X2=>X=K[I | 0]x=Kx(就那个转换公式)

b). When estimating a homography matrix, numerical stability typically becomes an issue. Explain why this is the case and give a solution to avoid this issue.

7.Two-view geometry

a). Why do we use rectified images for stereo estimation ?

b). How many correspondence pairs are needed at least for calculating the essential matrix and why?

c).Outline an algorithm that computes a dense depth map from two input images of two calibrated cameras(i.e.their intrinsic parameters are known). Which subproblems do you need to solve?

8. Motion

a).Window-based matching is a simple optical flow estimation method. What objective does this method optimize? How does the window size affect the estimated optical flow?

1.ground truth

2.too big: discontinuities at motion boundary

too small: not enough image information

b).Briefly describe the aperture problem in optical flow estimation.

the motion of on2-dimensional spatial structure can not be determined

unambiguous if it is viewed through a small apeture such that the ends of the

stimulus are not visible.

c).Name two weaknesses of the Lucas-Kanade method for optical flow estimation. How can we tackle these weaknesses?

9.Bags-of-words model and categorization

a). Why do we have to normalize the final histogram in the bag-of-words model?

b).Briefly describe how k-means clustering works. How is k-means clustering used i the context of bag-of-words?

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

10.Deep Learning

a). What is the key idea behind deep learning for computer vision? What is the major

advantage in comparison to using SIFT or HOG features?

助教说重点在 features

b).What is back propagation? How can you check if back propagation works correctly in a neural network ?

c). What is max pooling? Why is it used in many classification networks?

d). What loss function is typically used in image classification? Write it down and explain the terms.