THE AUSTRALIAN NATIONAL UNIVERSITY

Mid-semester Quiz

COMPUTER VISION

2020

ENGN6528

Writing period: 120 minutes.

Reading period: nil.

Please note that this is practice only

Attempt to answer as many questions as you can.
Answer your questions clearly and briefly, without the use of excessive sentences.
The clarity of your answers will substantially affect the marks.

All questions are to be completed on the exam paper.

Marks are as shown. Total is 100.

Marks: To be self-marked

Q1: (10 marks) [basic concepts]

Answer the following questions concisely. Each of the questions must be answered in no more than 5 lines of text. Longer answers will be penalized.

(1) Consider the HSV colour space. What does H, S and V stand for ?
Based on the HSV colour representation, what is the main difference between a "red" colour and a "pink" colour ?

(2) What does SIFT stand for ?

Using the PCA technique, any image can be represented as a linear combination of some so-called “eigenfaces”, plus a noise image. The eigenfaces can be computed by using eigen-value decomposition of certain covariance matrix A.

Do the eigenfaces correspond to Eigen-values of the matrix A, or Eigen-vectors of the matrix A ?

Q2: (15 marks) [basic concepts]

Answer the following questions concisely. Each of the questions must be answered in at most 5 lines of text. Longer answers will be penalized.

(1) What is the size of the camera matrix P for a standard perspective pinhole camera ?

One can decompose the camera matrix P into three components as P= K[R|t]. Please explain the meanings of K, R, t. Use a diagram if you like.

(2) Consider a pinhole camera observing a static scene.

What is the size of the matrix K ?
Name the calibration parameters of K.

How many degrees of freedom does K have ?

How many point correspondences are minimally needed to compute a combined matrix P of intrinsic and extrinsic coordinates using the linear DLT method ?

(3) Compare the following two operations: (1) you first apply a 3x3 Sobel filter (S) to an image I then followed by a Gaussian filter (G); and (2) you first apply a Gaussian filter G to the image then followed by a Sobel filter (S).

Mathematically speaking, these two different operations should give you equivalent results, i.e., (S*G)*I = (G*S)*I.

However, in practice one operation often produces better result than the other. Please explain why, and which one is your preferred edge detector?

(4) The vertical and horizontal Sobel filters are given below

$$L_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * L \quad \text{and} \quad L_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * L.$$

Please explain the effects of these two filters.

How would you use these filters to obtain the

orientation of the gradient?

Q3: (15 marks) [basic calculations]

Consider the 5×5 image below. The pixel grey values are indicated by the values in the cells.

3	2	1	2	4
2	1	3	200	3
6	7	8	7	9
8	100	6	6	7
7	9	6	8	8

Apply a 3×3 median filter to the image. Note that to avoid problems at the edges of the image you only need to calculate the filtered values for the central 3×3 region.

Apply the vertical edge filter used by the Sobel edge detector to the image above. Again, you only need to calculate the results for the central 3×3 region.

Note: in solving this problem, please use a "convolution" rather than "correlation" operator to implement the image filtering.

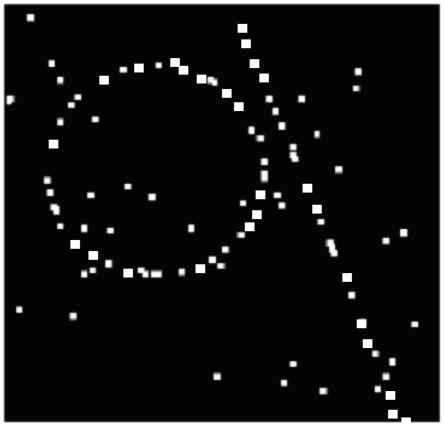
Q4: (15 marks) [Basic algorithms]

(1) Please outline the steps of a classic K-means algorithm briefly. In doing this task, you may use plain English, or use Matlab-like or Python-like pseudo-code (with necessary comments) if that is easier.

(2) Modify the above classic K-means algorithm in a way such that you can control the maximal size of each cluster. In other words, you can specify a parameter M such that the number of elements in any cluster is no greater than M. Outline the steps of your modified algorithm.

Q5: (15 marks) [basic design problem]

Given below is an input binary image. You are asked to design an algorithm that can extract the circle in it. You need to estimate all the parameters of the circle.



(i) Please outline your algorithm briefly.

(ii) Briefly explain how the circle is extracted by your algorithm, and how the line is ignored by your algorithm.

Q6: (15 marks)

Suppose the following four pairs of corresponding feature points from two image planes are related by an unknown homography matrix H. (These points are given in standard Euclidean coordinates).

(0; 0) --> (2; 3)
(1; 0) --> (3; 4)
(0; 1) --> (0; 4)
(1; 1) --> (2; 5)

(a) (2 marks) What is the size of the above homography matrix ?

How many independent variables (i.e. DoFs) does a homography have?

(b) (3 marks) Can you use the above four pairs of corresponding points to compute the homography matrix linearly ?

If your answer is "Yes", you do not need to compute the H matrix; all you need to do is to list and explain your main algorithm briefly.

If your answer is "No", then please further answer: How many pairs of points (correspondences) are minimally needed in order to compute the homography matrix linearly?

(b) (5 marks) Without actually computing H, is it still possible to tell where the point (0.5; 0.5) in the first image maps to, in the second image ? i.e., (0.5,0.5) --> (? , ?). If so, state the corresponding point's coordinates in the second image; Otherwise, explain why.

Q7: (15 marks) [basic design problem]

Design a computer vision algorithm that can be used to detect the chair in the image. Briefly explain the key steps of your algorithm.

======= END of ALL QUESTIONS in the QUIZ ==========