NATIONAL UNIVERSITY OF SINGAPORE

EE4212 – COMPUTER VISION

(Semester 2 : AY2019/2020)

Time Allowed : 2 Hours

INSTRUCTIONS TO STUDENTS

1. Please write only your Student Number. Do not write your name.
2. This assessment paper contains **EIGHT** essay questions and **EIGHT** multiple choice questions. It comprises **EIGHT** printed pages.
3. Students are required to answer **ALL** questions.
4. Students should write the answers for each question on a new page.
5. This is a OPEN BOOK assessment.
6. Non-programmable calculators are allowed.
7. Total Marks is **ONE HUNDRED (100)**.

Q1

From a few point correspondences, you have used the classical 8-point algorithm to obtain an estimate of the fundamental matrix **F**. The singular value decomposition (SVD) of **F** is given as below, where the central matrix is a diagonal matrix, and $\mathbf{u}_i$ and $\mathbf{v}_i$ represent column vectors.

$$\mathbf{F} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0.0002 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}^{\mathrm{T}}$$

Note that the **F** above does not have the requisite rank property. What would happen to the intersection point(s) of the epipolar lines obtained from this fundamental matrix **F**? Describe in words what should be modified in the above expression so that the improved **F** possesses the desired rank property. Write down the left epipole in pixels in terms of elements of the above expression.

(9 marks)

Q2

Given the intrinsic matrices, you have upgraded the fundamental matrix **F** to the Essential matrix **E**. If the null space of this matrix **E** is $(a,b,c)^{\mathrm{T}}$, explain how you would

(a) find the left epipole (x_0, y_0) in mm, given that you know the focal length for the left camera is f_0 in mm,

(b) Write down the two possible solutions for the translation vector **t**, given that you know the magnitude of **t** is k, and that the magnitude of $(a,b,c)^{\mathrm{T}}$ can be represented by M.

(c) Write down the null space of **F** if the translation is pointing in the Y-direction, and there is no rotation between the two cameras.

(9 marks)

Q3

Figure A depicts the view from inside a car driven by an autonomous driving system. Suppose the autonomous driving system has estimated the optical flow vectors of this scene in view using the Horn and Schunck's algorithm. Discuss in details whether the following regions/issues would pose problems for the optical flow estimation:

(a) The traffic lights;

(b) The surfaces of the cars in the scene;

(c) The thin elongated structures in the scene (e.g. the scaffolding on the left);

(d) The fact that these cars are moving independently of the camera's egomotion (assuming that the magnitudes of these independent motions are not too large).

(12 marks)

Figure A

Q4

Figure A depicts the view from inside a car driven by an autonomous driving system. Assume that perfect optical flow for this scene has been obtained at every pixel location.

(a) Explain whether you could obtain good motion parallax across the depth discontinuities of the distant objects, in particular, those that exist between the trees and the building.

(b) Explain why you should not directly apply the approximate motion parallax algorithm to the optical flow to estimate the Focus of Expansion (FOE) of the camera's egomotion.

(c) Explain the main advantage of using a camera with hemispherical field of view (i.e. 180° field of view) in egomotion estimation.

(10 marks)

Q5

Textures are one of basic visual patterns of images. Gray level co-occurrence matrix (GLCM) is usually used as features for describing texture patterns. Given an image patch as shown in **Figure B**, please calculate and write down its GLCM features at four different angles: 0, 45, 90 and 135 degrees and two displacement distances d=1 and 2. (Note: there are in total 8 GLCM features you need to compute).

(10 marks)

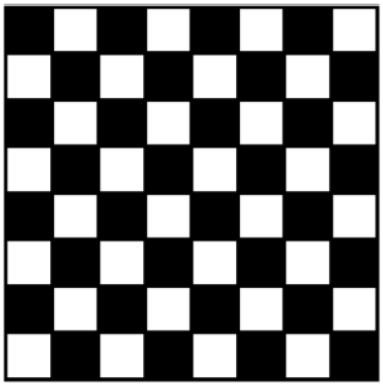

Figure B

Q6

Interactive image segmentation is a common function in many image editing software. Given several inputs of the users (in the form of clicks or curves) to indicate some foreground and background pixels, the interactive image segmentation system can automatically segment the foreground out from the background by following the user-provided annotations. **Figure C** shows one example (left: input image with user annotation; right: segmentation result). Please use the Markov Random Field (MRF) model to design such an image segmentation system.

Explain the formulation of the MRF objective function for interactive image segmentation and explain the optimization algorithm in text.

(10 marks)

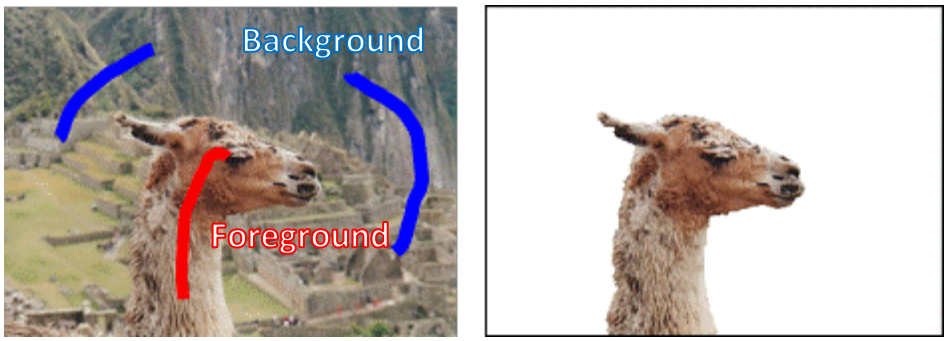

Figure C

Q7

SIFT is a widely used image descriptor for many applications including image matching, image stitching and panorama generation. Here, given two images, we are aiming to find their pixel-to-pixel correspondence. Please describe the SIFT feature extraction procedure from the input image and how to obtain the image matching step by step.

(10 marks)

Q8

Explain why SIFT features are rotation and scale invariant.

(10 marks)

Multiple Choice Questions (2.5 marks each)

A. Which of the following is a valid homogeneous representation for some point in a two-dimensional image?

(a) (0,1,0)
(b) (0,0,0)
(c) (1,1)
(d) None of the above

B. The vector given in **Figure MCQ-1** is the optical flow measured at pixel ***x***, for a camera that is both rotating and translating. Knowing that the translation *Z*-component $T_Z < 0$, which of the following is a possible location for the Focus of Expansion?

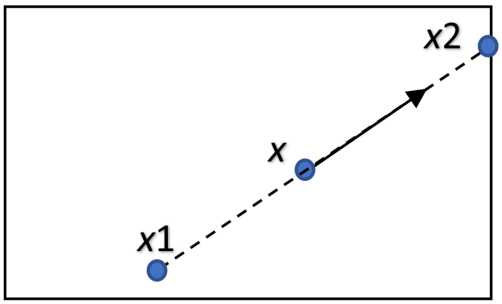

Figure MCQ-1

(a) x1
(b) x2
(c) both x1 and x2 are possible
(d) Cannot be ascertained, due to the presence of rotation

C. **Figure MCQ-2** shows the 2D motion flow pattern arising from a moving camera (which has a relatively small field of view) in front of a static scene. Which of the following is the possible camera motion that gives rise to the flow? (In the figure, the flow is along the line in the direction away from the dot.)

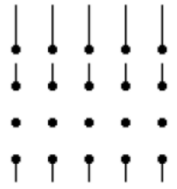

Figure MCQ-2

(a) Pure translation along the camera's vertical axis
(b) Pure rotation about the camera's horizontal axis
(c) A combination of (a) and (b)
(d) None of the above

D. Which of the following is true?

(a) A straight line in an image is straight in 3D.
(b) If the tips of two lines coincide in an image, then they are coincident in 3D.
(c) Lines collinear in an image are collinear in 3D.
(d) All of the above
(e) None of the above

E. You have learnt that the motion parallax should pass through the FOE (focus of expansion). Is it possible to have a 3D scene such that some of the motion parallaxes in the image point towards the FOE whereas the remaining motion parallaxes point away from the FOE?

(a) Yes
(b) No
(c) Yes, but the scene must contain independent motions for this to happen
(d) Yes, but there must be relative rotation between the camera and the scene for this to happen

F. Which of the following descriptions on the relation between the focal length and field of view (FOV), and the relation between the aperture size and the depth of field (DOF) is correct?

(a) If the focal length is larger, FOV is larger; if the aperture size is larger, DOF is larger
(b) If the focal length is larger, FOV is larger; if the aperture size is larger, DOF is smaller
(c) If the focal length is larger, FOV is smaller; if the aperture size is larger, DOF is larger
(d) If the focal length is larger, FOV is smaller; if the aperture size is larger, DOF is smaller

G. Radial distortion bends straight lines into circular arcs and would cause shape distortion of captured photo. Suppose we have two fixed parameters $k_1 \geq 0$ and $k_2 \geq 0$. Let $r^2 = x_d^2 + y_d^2$. Which of the following solutions can correct the radial distortion?

(a) $x = x_d(1 + k_1 r^2 + k_2 r^4), y = y_d(1 + k_1 r^2 + k_2 r^4)$
(b) $x = x_d(1 - k_1 r^2 + k_2 r^4), y = y_d(1 - k_1 r^2 + k_2 r^4)$
(c) $x = x_d(1 + k_1 r^2 - k_2 r^4), y = y_d(1 + k_1 r^2 - k_2 r^4)$
(d) $x = x_d(1 - k_1 r^2 - k_2 r^4), y = y_d(1 - k_1 r^2 - k_2 r^4)$

H. Which of the following descriptions about K-means for clustering is correct?

(a) The clustering results from K-means are not sensitive to the initial centers.

(b) The clustering results from K-means are sensitive to the initial centers.

(c) Using different features for the data to cluster, K-means will give the same clustering results.

(d) Using different similarity metric for the data to cluster, K-means will give different clustering results.

END OF PAPER