

NANYANG TECHNOLOGICAL UNIVERSITY**SEMESTER 2 EXAMINATION 2023-2024****EE6222 – MACHINE VISION**

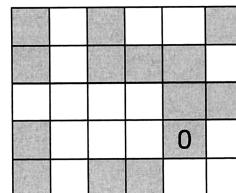
May 2024

Time Allowed: 3 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 4 pages.
 2. Answer all 4 questions.
 3. All questions carry equal marks.
 4. This is a closed-book examination.
 5. Unless specifically stated, all symbols have their usual meanings.
-

1. A binary image is represented by set A .
 - (a) Design an algorithm to extract an 8-connected component from a binary image represented by set A . (7 Marks)
 - (b) Design an algorithm to extract a 4-connected component from a binary image represented by set A . (4 Marks)
 - (c) Let set A contain the shaded pixels in Figure 1. The pixel indicated by “0” in Figure 1 is selected to extract an 8-connected component. Write the value n in the proper pixel cells/locations to indicate the sequence of which the pixels are extracted in the n^{th} iteration of your algorithm. (7 Marks)

**Figure 1**

Note: Question No. 1 continues on page 2.

- (d) Show the result for part (c) for extracting a 4-connected component instead of 8-connected component. (7 Marks)
2. You are in an office on the school campus. A student of height 1.67m knocks on the door. You need to judge whether the student is male or female with the minimum probability of a wrong decision. You also need to show that your decision minimizes the probability of a wrong decision with specific assumptions.
- (a) What is your decision, and what is the probability that your decision is wrong, if you only know that there are 300 male students and 700 female students in the school? (5 Marks)
- (b) What is your decision if you know that there are an equal number of male and female students in the school, and the average height of male students is 1.7m and that of female students is 1.62m? (8 Marks)
- (c) Suppose there are 300 male and 700 female students in the school. What is your decision if you know that the average height and standard deviation of male students are 1.7m and 0.2m, respectively, and those of female students are 1.62m and 0.3m, respectively? (12 Marks)
3. A gray-level digital image X of size 100×100 is cropped into small images of size 3×3 by sliding the center of a 3×3 window to every pixel of X . If the center of the window is on the boundary of the image X , zero padding is used to obtain images of size 3×3 so that we get 10000 small images of size 3×3 . Each 3×3 image is flattened into a column vector and is expressed as $\mathbf{z}^k = (z_1^k, \dots, z_9^k, 1)^T$, $k = 1, \dots, 10000$, and is input to a typical fully-connected layer of a multilayer perceptron (MLP) with a linear activation function to generate the output vectors $\mathbf{y}^k = (y_1^k, \dots, y_6^k)^T$. The network parameters of this layer are denoted by w_{ij} and b_j , $0 < i < 10$, $0 < j < 7$.
- (a) Express the outputs y_j^k in terms of the inputs z_i^k . (5 Marks)
- (b) Construct the matrix \mathbf{W} that contains all network parameters, and express the output vector \mathbf{y}^k in terms of the input vector \mathbf{z}^k . (5 Marks)

Note: Question No. 3 continues on page 3.

- (c) Six output images, \mathbf{Y}_j , $j = 1, \dots, 6$, of the same size with the input image \mathbf{X} , are constructed by the 10000 output vectors \mathbf{y}^k , $k = 1, \dots, 10000$. Express the output images \mathbf{Y}_j in terms of the input image \mathbf{X} . (10 Marks)
- (d) Suppose this network is trained by 100 images of size 100×100 . What is the number of training samples used to train the network parameters \mathbf{W} or w_{ij} and b_j ? (5 Marks)
4. (a) To measure the actual distance between two points in the 3D space, one useful strategy is to leverage on the projection invariant properties so that the distance can be measured with the corresponding projected points of the image. One such property is the cross-ratio.
- (i) Suppose there are four points in the real world denoted as points a, b, c and d , strictly in sequence and are situated on a line. Express the cross-ratio between these four points.
 - (ii) They are then projected to the image as points A, B, C and D , respectively, and strictly in sequence situated also on a line. Prove that the cross-ratio is projective invariant.
 - (iii) Suppose we first measure the distances among those image points and the outcomes are $|AB| = 20\text{px}$, $|BC| = 40\text{px}$, and $|CD| = 60\text{px}$. Correspondingly, we also know the distances among the points in the 3D space are $|ab| = 6\text{m}$ and $|cd| = 8\text{m}$. What is the distance between the points b and c in space, denoted as $|bc|$? Observe the computation and state briefly the advantages and disadvantages of leveraging cross-ratio for measuring points in space. (15 Marks)
- (b) The depth of an object can be estimated by the disparity of the object in a stereo pair via stereo geometry. Suppose we have an ideal stereo camera, whose baseline is of length B . We further assume that the focal length is f and the disparity of the object of interest is d .
- (i) Formulate how to obtain the depth of the object of interest, Z , with the disparity equation.

Note: Question No. 4 continues on page 4.

- (ii) Suppose we know that there is a disparity error Δd and have previously estimated the depth Z . However, we do not know the actual disparity. With the existence of a disparity error, there must be an error in the estimated depth, denoted as ΔZ . Derive in detail the estimated depth error ΔZ in terms of the disparity error Δd and the estimated depth Z . What can be observed from the derived formula suggesting the estimated depth error?

Note: The estimated depth error should be in its simplest form via Taylor series approximation.

(8 Marks)

- (c) Large Language-Vision Models (LLVMs) is currently state-of-the-art for video / image recognition. State, in brief, the key difference in training previous models (i.e., CNNs and variants or ViT) opposed to LLVMs.

(2 Marks)

END OF PAPER

EE6222 MACHINE VISION

Please read the following instructions carefully:

- 1. Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**
2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.
3. Please write your Matriculation Number on the front of the answer book.
4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.