NATIONAL UNIVERSITY OF SINGAPORE

ME5405 – MACHINE VISION

(Semester 1 : AY2016/17)

Time Allowed : 2.5 Hours

<u>INSTRUCTIONS TO STUDENTS</u>

1. Please write your Student Number only. Do not write your name.
2. This assessment paper contains **FOUR(4)** questions and comprises **EIGHT(8)** 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 an **OPEN BOOK** assessment.
6. Programmable calculators are **NOT** allowed for this examination.

Question 1

A research engineer mounts a camera on a ground robot. She drove the robot on a flat planar ground while the camera captures two images I and I' with overlapping field-of-view at two different times. Let the camera reference frames where the images are taken be $(\mathrm{x},\mathrm{y},\mathrm{z})$ and $(\mathrm{x}',\mathrm{y}',\mathrm{z}')$. As shown in Fig. 1, the relative motion between these two frames is denoted by $(t_x, t_z, \theta_y = 0)$, where (t_x, t_z) is the translation vector on the xz-plane and θ_y is the rotation angle around the y-axis.

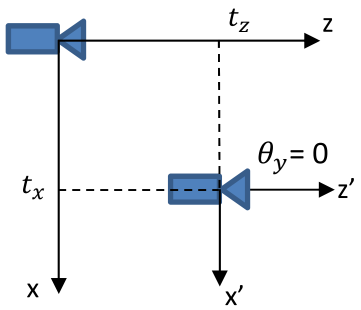

Fig. 1

(a) Write the expression for the essential matrix E between I and I' in terms of (t_x, t_z, θ_y), where $\theta_y = 0$.

(*5 marks*)

(b) The camera focal length and principle point are given by (f_x, f_y) and $(C_x = 0, C_y = 0)$ respectively. A 3D point P in the scene appears as $p = (p_x, p_y)$ in image I, write the expression for the corresponding epipolar line L' in image I' in terms of (f_x, f_y), (C_x, C_y), (p_x, p_y) and $(t_x, t_z, \theta_y = 0)$. Show your working clearly.

(*6 marks*)

(c) Given a pair of point correspondence $p \leftrightarrow p'$ i.e. $(p_x, p_y) \leftrightarrow (p'_x, p'_y)$ from I and I'.

(i) Show that the relationship between p and p' is given by

$$p_y(f_x t_x - p'_x t_z) - p'_y(f_x t_x - p_x t_z) = 0.$$

(*6 marks*)

(ii) What is the minimum number of image correspondences needed to solve for the relative motion $(t_x, t_z, \theta_y = 0)$?

(*2 marks*)

(d) Suppose four 3D points lying on a plane in the scene are concurrently seen by I and I'. If the plane is parallel to both the image planes and is located at d distance away from the reference frame of the first camera $(\mathrm{x},\mathrm{y},\mathrm{z})$, find the homography in terms of $(t_x, t_z, \theta_y = 0)$ and d that relates I, I' and the 3D plane.

(*6 marks*)

Question 2

(a) $G(x, y, \sigma)$ is a 2D image filter defined as

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^3 + x^2 + xy^2 + y^2}{2\sigma^2 x + 2\sigma^2}\right).$$

(i) Show that $G(x, y, \sigma)$ is a separable filter.

(*5 marks*)

(ii) Starting from the definitions of correlation and convolution, show that the outputs from correlation and convolution are the same for the filter $G(x, y, \sigma)$ on an image $I(x, y)$.

(*5 marks*)

(iii) Name the filter and describe its effect on an image.

(*1 mark*)

(b) The image gradients I_x and I_y of a 3x3 patch in the x and y directions are given by

$$I_x = \begin{bmatrix} 0 & -1 & -1 \\ 0 & -2 & -2 \\ 0 & -3 & -3 \end{bmatrix}, \quad I_y = \begin{bmatrix} 3 & 2 & 1 \\ 3 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Assuming a constant weight of 1 on each entry of the 3x3 patch, show that the 3x3 patch is likely to be a corner in the image.

(*8 marks*)

(c) A quantitative analyst made a discovery that the value of the Straits Times Index (STI) over time t can be modelled by a cubic polynomial function:

$$y = a_0 + a_1 t + a_2 t^2 + a_3 t^3,$$

where (a_0, a_1, a_2, a_3) are coefficients of the polynomial function that can be found from a least-squares fit of the polynomial over n observations (t_i, y_i), for all $i = 1 \ldots n$.

(i) Unfortunately the n observations (t_i, y_i), $i = 1 \ldots n$ are corrupted with outliers which severely undermined the accuracy of the least-squares fit. Determine the minimum number of observations needed to robustly estimate the coefficients using RANSAC. Justify your answer.

(*2 marks*)

(ii) Find the probability that any selected point is an inlier if a minimum number of 72 RANSAC iterations is needed for robust estimation of the coefficients (a_0, a_1, a_2, a_3). Assume that the probability of at least one of the RANSAC iterations is outlier free is 0.99.

(*4 marks*)

Question 3

(a) The histogram of a given image has two principal brightness regions. The background and the object regions are represented, respectively, by the two Rayleigh distributions given below:

$\frac{2z}{\beta_1}\exp\left(-\frac{z^2}{\beta_1}\right)$ and $\frac{2z}{\beta_2}\exp\left(-\frac{z^2}{\beta_2}\right)$; where z is the gray level and β_1 and β_2 are constants. The respective *à priori* probabilities are P_1 and P_2.

(i) Based on minimum error criterion, determine the optimal threshold of the image.

(ii) From the expression of the optimal threshold that you have determined in Part (i) above, deduce the conditions under which this optimal threshold is valid. Is Rayleigh distribution useful in this case?

[*Hint: restrictions posed by the relationships between P_1 and P_2, and between β_1 and β_2.*]

Note: The mean and variance of a Rayleigh distribution $\frac{2z}{\alpha}\exp\left(-\frac{z^2}{\alpha}\right)$ are given by $\sqrt{\frac{\pi\alpha}{4}}$ and $\frac{\alpha(4-\pi)}{4}$, respectively, and α is a constant.

(15 marks)

(b) Fig. 3(a) shows a 500 by 500 pixel size image. Each of the vertical bars (enclosed in the box) are 5 pixels wide and 100 pixels high. Their separation is 20 pixels exactly. Fig. 3(b) shows the profile of part of a horizontal scan line through the vertical bars.

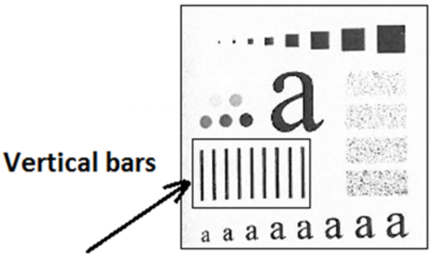

Fig. 3(a)

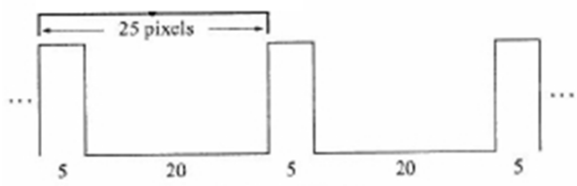

Fig. 3(b)

Fig. 3(a) is now smoothed with square averaging filter masks of size 23, 25 and 45. The resulting images are shown in Fig. 3(c), (d) and (e), respectively.

Fig. 3(c)

Fig. 3(d)

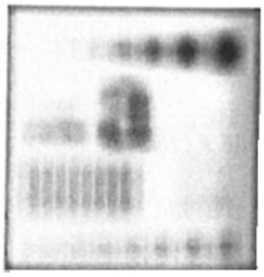

Fig. 3(e)

The vertical bars in Fig. 3(c) and Fig. 3(e) are blurred, but a clear separation exists between them. However, the vertical bars in Fig. 3(d) has merged into one single region. Explain the reason for this observation.

(10 marks)

Question 4

(a)

(i) Fig. 4(a) shows a binary image A.

(1) Design a structure element and the associated morphological operation(s) to remove the "leg" of the figure as shown in Fig. 4(b).

(2) To obtain Fig. 4 (c) from Fig. 4(a), design the structure elements and the necessary morphological operation(s) to achieve the objective [*Hint: You will need two morphological operations with different structure elements*].

Note that the pixel marked "+" is the origin of the image.

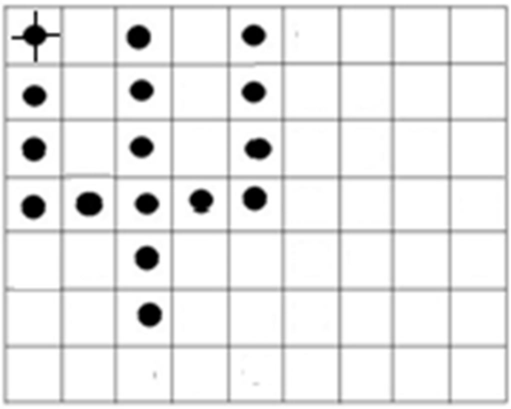

Fig. 4(a) Image A

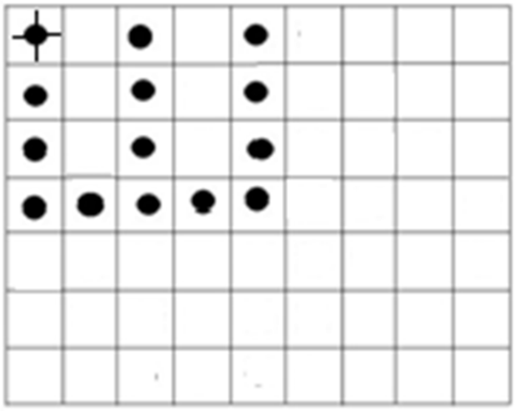

Fig. 4(b) Image B

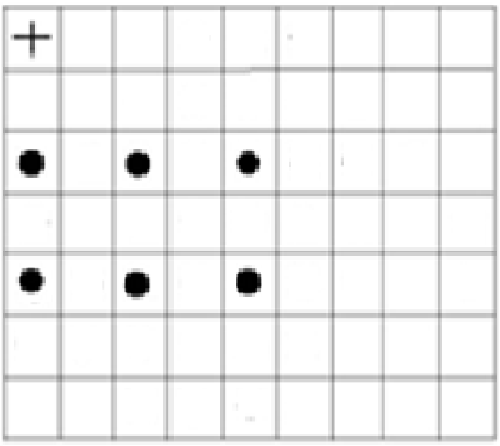

Fig. 4(c) Image C

(ii) Fig. 4(d) shows a binary image D. Figures Fig 4(e) and 4(f) are two different structure elements E and F, respectively. Determine the result of the following morphological operations. Note that the pixel marked "+" is the origin of the image.

(1) $G = D \ominus E$
(2) $G \oplus F$

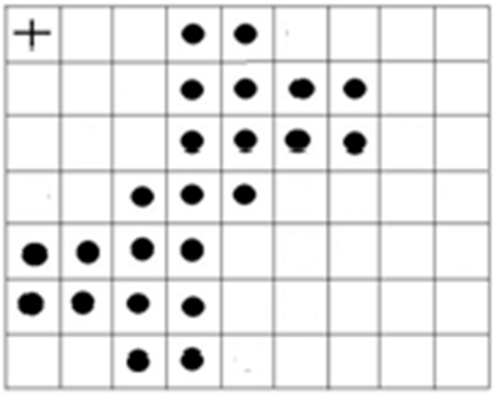

Fig 4(d) Image D

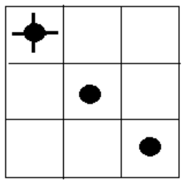

Fig 4(e) Structure Element E

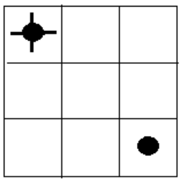

Fig 4(f) Structure Element F

(13 marks)

(b)

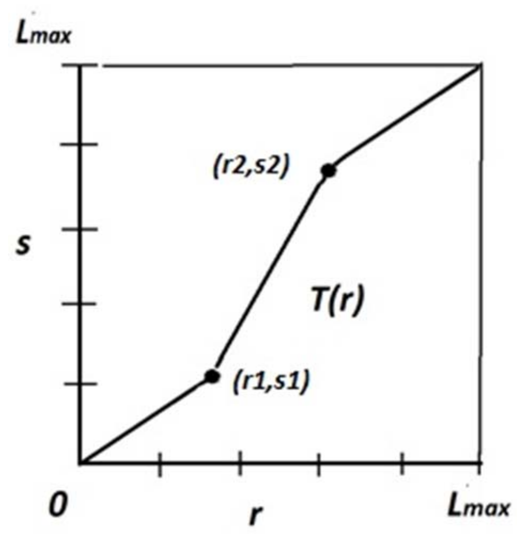

Fig. 4(g)

In the global enhancement of an image by point transformation, a Transfer Function *T(r)* is designed to transform the gray level of the original image (*r*) to new gray level (*s*) of the transformed image with the desired enhancement effects. *T(r)* usually consists of line segments joined at the point where transformation of a certain effect begins. Fig 4(g) shows an example of a $s = T(r)$ with points (r_1, s_1) and (r_2, s_2). In the figure *Lmax* is the maximum gray level value available.

(i) Given a gray-level image with gray level *r* lying in the interval [0,240]. We wish to compress the gray level intervals [0,80] and [160,240] by a factor of 2, and expand the gray level in the interval [80,160] by a factor of 2. Find the transfer function *T(r)* required to achieve the transformation. Sketch the transfer function and label the points (r_1, s_1) and (r_2, s_2) .

(ii) Suppose now we have an image with gray level *r* ranges in the interval [0,30]. Determine the transfer function *T(r)* for the following transformation:

- Stretches the gray level *r* in interval [0,10] to *s* in the interval [0,15];
- Shift the gray level *r* in the interval [10,20] to *s* in the interval [15,25]; and
- Compresses the gray level *r* in the interval [20,30] to *s* in the interval [25,30].

Sketch the transfer function and label the points (r_1, s_1) and (r_2, s_2).

(12 marks)

-- END OF PAPER --