RKMVERI

MSc. Big Data Analytics, Year 2024
Mid-term Examination
Computer Vision
Maximum Marks 60

Time: 120 minutes Date: ____/03/2024

a. Short questions (any ten, 2 marks each):

1. What are the percentage coverage of red, green and blue filters in Bayer filter of digital colour camera?
   Red 25%, Green 50% and Blue 25%

2. State two differences between rod and cone cells in human retina.

Any two:	
Rod cells	**Cone cells**
Sensitive to low intensity light.	Sensitive to high intensity light and colour.
Smaller is size.	Larger in size.
More in number at periphery (away from fovea).	More in number around fovea.

3. List the intrinsic parameters of a camera. Considering homogeneous coordinate system, write down the intrinsic parameter matrix to map world point to image pixel.
   Intrinsic parameters: focal length f, photoreceptor density ϕ_x and ϕ_y, half of image plane size $d_x$and d_y.
   Intrinsic parameter matrix:

$$\begin{bmatrix} \phi_x & 0 & d_x/f & 0 \\ 0 & \phi_y & d_y/f & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

4. Why is Gamma correction required?
   Gamma correction is required to linearize the non-linear response of photosensors to incident light intensity.

5. Given two cameras what is meant by epipolar constraint?
   For any point on the image plane of Camera-1, corresponding point lies on its epipolar line on the image plane of Camera-2 and *vice versa*. This is *epipolar constraint*.

6. Suppose there are two cameras. Which point is called the epipole of the image plane in camera-2? Whose image is it?

All epipolar lines in image plane of Camera-2, due to image points in Camera-1, meet at a point. This meeting points is called *epipole*.
Epipole is the image of optical centre of Camera-1 on the image plane of Camera-2, and *vice versa*.

7. What is the main objective of rectification technique?
   In dense stereo reconstruction the objective of rectification technique is to geometrically transform the image(s) in such a way that for a point in one image the corresponding point lies on the same horizontal scanline in the other image. This reduces search complexity.

8. Suppose minimum and maximum gray levels in an image are 'a' and 'b', whereas that of available gray range are 'L' and 'M' respectively. Write the linear stretching transformation from input graylevel 'x' to output gray level 'y'.
   $$y = \frac{M - L}{b - a}(x - L) + M$$

9. State one advantage and one disadvantage of median filter over mean filter for image noise smoothing.
   Median filter preserves the edge sharpness better than mean filter, but it is computationally more expensive.

10. How can smoothing technique be used to sharpen an image?
    Smoothing technique removes fine details including edge sharpness from image. So, smoothened image subtracted from the original image gives fine details removed by smooting. Hence, adding some multiple or fraction of this detail to original image enhances its sharpness.
    $$I_{sharp} = I_{org} + \alpha(I_{org} - I_{smooth})$$

11. What are two different approaches to image segmentation?
    a) Region extraction (based on some measure of homogeneity).
    b) Edge detection (based on abrupt change in some feature).

12. Define Laplacian for an image. How does it help in detecting edge point (pixel) in the image?
    Laplacian (in continuous domain): $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$
    Laplacian (in continuous domain): $\nabla^2 f = f * \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
    Edge pixel may be detected as zero-crossing in $\nabla^2 f$.

b. Regular questions (any four, 5 marks each):

13. Suppose the tip of a tree forms an image on the image plane of a digital camera having focal length 1.2 cm and image plane of size 8 cm x 4.5 cm. If the tip of the tree be at (-2 m., 5 m. 10 m.) with respect to the optical centre of the camera, find the image coordinate of the corresponding image point. Given that the photoreceptor density is 1000/cm in both horizontal and vertical direction.

Given focal length f, photoreceptor density $\phi_x = \phi_y = 1000$,

Displacement from centre to corner:

$d_x = \frac{s_x \times \phi_x}{2} = \frac{8 \times 1000}{2} = 4000$ and $d_y = \frac{s_y \times \phi_y}{2} = \frac{4.5 \times 1000}{2} = 2250.$

$$x = f\frac{u\phi_x}{w} + d_x = 1.2\frac{(-2)1000}{10} + 4000 = 3760$$

$$y = f\frac{v\phi_y}{w} + d_y = 1.2\frac{(5)1000}{10} + 4000 = 4600$$

Image point: $(3760, 4000)$

14. Two cameras having same intrinsic parameters and same orientation. However, one camera is shifted by a translation vector $[t_x, t_y, 0]^T$ with respect to the other. If an object point is imaged at (x_1,y_1) and (x_2,y_2) respectively in two cameras, prove that

$$\frac{t_x}{x_1 - x_2} = \frac{t_y}{y_1 - y_2}$$

Let with respect to first camera the 3D object point is (u, v, w).

Then with respect to second camera the same 3D point is $(u - t_x, v - t_y, w)$.

Thus $x_1 = f\frac{u}{w}$, $y_1 = f\frac{v}{w}$, $x_2 = f\frac{u - t_x}{w}$ and $y_2 = f\frac{v - t_y}{w}$

$x_1 - x_2 = f\frac{t_x}{w}$ and $y_1 - y_2 = f\frac{t_y}{w}$, then $\frac{w}{f} = \frac{t_x}{x_1 - x_2}$ and $\frac{w}{f} = \frac{t_y}{y_1 - y_2}$

$$Hence, \quad \frac{t_x}{x_1 - x_2} = \frac{t_y}{y_1 - y_2}$$

15. Suppose the fundamental matrix of Camera-2 with respect to Camera-1 is given by $\begin{bmatrix} 0.005 & 0 & -2 \\ 0.001 & 0.002 & 0 \\ 0.08 & -0.06 & 5 \end{bmatrix}$, find the equation of epipolar polar line in the image plane of Camera-2 due to point $(1000,\ 1500,\ 1)$ in the image plane of Camera-1.

Equation of epipolar polar line in the image plane of Camera-2 due to point $p_h^{I_1}$:

$$p_h^{I_1} F^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

Hence, $(1000\ \ 1500\ \ 1)\begin{bmatrix} 0.005 & 0 & -2 \\ 0.001 & 0.002 & 0 \\ 0.08 & -0.06 & 5 \end{bmatrix}^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$

$$(1000\ \ 1500\ \ 1)\begin{bmatrix} 0.005 & 0.001 & 0.08 \\ 0 & 0.002 & -0.06 \\ -2 & 0 & 5 \end{bmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

$(3\ \ 4\ \ -5)\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$ or $3x + 4y - 5 = 0$

16. Define image segmentation in mathematical terms.

Segmentation is a technique that divides image domain D into, say, n regions denoted by $R_1,\ R_2,\ \ldots,\ R_n$ such that

$$\cup_{i=1}^{n} R_i = D,\ \ R_i \cap R_j = \emptyset,\ \ Prop.(R_i) = True \ \text{ for all } \ i, \text{ and}$$
$$Prop.(R_i \cup R_j) = False, \quad \text{if } R_i \text{ and } R_j \text{ are adjacent.}$$

17. Graylevel frequency count of an 8-level image of size 100x100 is given by the following table.

Graylevel	0	1	2	3	4	5	6	7
Frequency count	0	400	1300	700	3500	2500	1600	0

If its contrast is enhanced by histogram equalization technique, what will be the graylevel frequency count of the output image?

r_i	0	1	2	3	4	5	6	7
n_i	0	400	1300	700	3500	2500	1600	0
p_i	0	0.04	0.13	0.07	0.35	0.25	0.16	0
c_i	0	0.04	0.17	0.24	0.59	0.84	1.00	1.00
$c_i \times 7$	0	0.28	1.19	1.68	4.13	5.88	7.00	7.00
$[c_i \times 7]$	0	0	1	2	4	6	7	7
s_i	0	1	2	3	4	5	6	7
N_i	400	1300	700		3500		2500	1600

c. Long questions (any two, 10 marks each):

18. Suppose a camera is aligned with world co-ordinate system and its optical centre is at origin (0,0,0). Second camera is rotated and translated with respect to the first camera. The said rotation and translation is represented by a 3x3 matrix R and 3x1 vector $\boldsymbol{t}$. Let the intrinsic parameter matrices of the first and second cameras are K_1 and K_2. Express the *fundamental matrix* F in terms of intrinsic and extrinsic parameter matrices. Hence, prove that

$\left(p_h^{I_2}\right)^T F p_h^{I_1} = 0$, where $p_h^{I_i}$ denotes the (3x1) homogeneous coordinate of cooresponding image points in 1st and 2nd cameras. Superscript *T* denote transposition.

19. Suppose in an image there are two types of regions. The graylevel of the pixels in the first type of regions follows Gaussian distribution with mean μ_1 and standard deviation σ_1, while that in second also follows Gaussian distribution with mean μ_2 and standard deviation σ_2. Suppose *apriori* probability of the pixels to belong to first type of regions is P_1 and that for second type of regions is P_2. Derive the optimum threshold *t* for image segmentation by graylevel thresholding.
20. State three criteria of a good edge detector. Hence, describe Canny's edge detector that is designed to satisfy these criteria.

--- o ---