AVD624/ESG667

Indian Institute of Space Science and Technology Trivandrum

I SEMESTER , 2024

ExamType: Quiz 1

DEPARTMENT OF AVIONICS

computer vision/ computer vision and advanced image processing

(Time allowed: ONE hours)

NOTE: Read all questions first. **There are questions worth 30 marks.** If something is missing in a problem description, clearly mention your assumptions with your solution. If required, use sketches to illustrate your findings.

1. What is Homography? What is the use of tomography in computer vision? Explain the notion of the algebraic distance and the geometric distance in the context of estimating a homography. (5 marks)

2. write the camera matrix. What do the different rows and the different columns of the camera matrix stand for? Provide simple algebraic proofs for your answer. (5 marks)

3. Explain the working of Zhang's camera calibration scheme. Write the necessary steps required to obtain the camera calibrations. (Derivation is not required) (5 marks)

4. Parallel Lines under Perspective Transforms (5 marks)

   (a) The two boxes in Figure 1 represent the same 3D shape rendered using two projective techniques, explain their different appearance and the types of projections used to map the objects to the image plane.

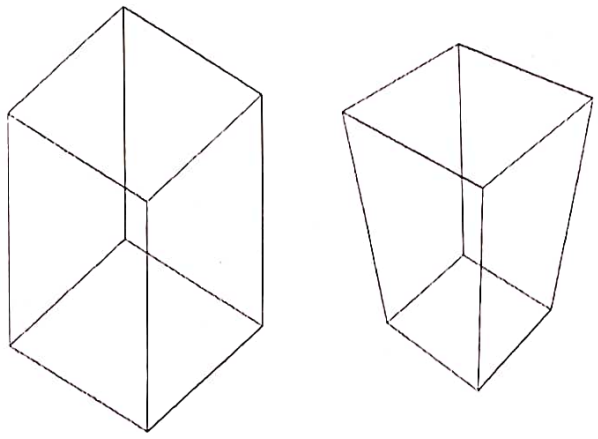

Figure 1: Boxes rendered using different projections

(b) For each projection, if the edges of the cubes were to be extended to infinity, how many intersection points would there be?

(c) What is the maximum number of vanishing points that are possible for an arbitrary image?

(d) How would you arrange parallel lines so that they do not appear to have a vanishing point?

5. Using RANSAC to find circles: Suppose we would like to use RANSAC to find circles in R^2. Let $D = ((x_i, y_i)_{i=1}^n$ be our data, and let I be the random seed group of points used in RANSAC. (5 marks)

(a) The next step of RANSAC is to fit a circle to the points in I. Formulate this as an optimization problem. That is, represent fitting a circle to the points as a problem of the form minimize

$$minimize \sum_{i \in I} L(x_i; y_i; c_x; c_y; r)$$

where L is a function for you to determine which gives the distance from $(x_i; y_i)$ to the circle with center $(c_x; c_y)$ and radius r

(b) What might go wrong in solving the problem you came up with in when number of points in I is too small

(c) The next step in our RANSAC procedure is to determine what the inliers are, given the circle $(c_x; c_y; r)$. Using these inliers we refit the circle and determine new inliers in an iterative fashion. Define mathematically what an inlier is for this problem. Mention any free variables.

6. Compare and contrast the following (write the transformation equations and highlight what remains invariant and what gets modified due to these transformations) (5 marks)

(a) Translation Transformation

(b) Euclidean transformation

(c) Similarity Transformation

(d) Affine Transformation

(e) Projective Transformation

Indian Institute of Space Science and Technology Trivandrum

I SEMESTER , 2024

ExamType: Quiz 1

DEPARTMENT OF AVIONICS

computer vision/ computer vision and advanced image processing

(Time allowed: ONE hours)

NOTE: Read all questions first. **There are questions worth 30 marks.** If something is missing in a problem description, clearly mention your assumptions with your solution. If required, use sketches to illustrate your findings.

1. We have discussed corner detection for 2D images. First, briefly describe the Harris corner detection method for 2D images. Now, we want a method for corner detection for use with 3D images, i.e., there is an intensity value for each (x,y,z) voxel. Describe a generalization of either the Harris corner detector by giving the main steps of an algorithm, including a test to decide when a voxel is a corner point (5 marks)

2. Write Important steps of SIFT algorithm. Explain how we obtain a 128-dimensional feature descriptor. (5 marks)

3. The following operator is often applied to an image I(x, y) in computer vision algorithms, to generate a related function h(x, y): (10 marks)

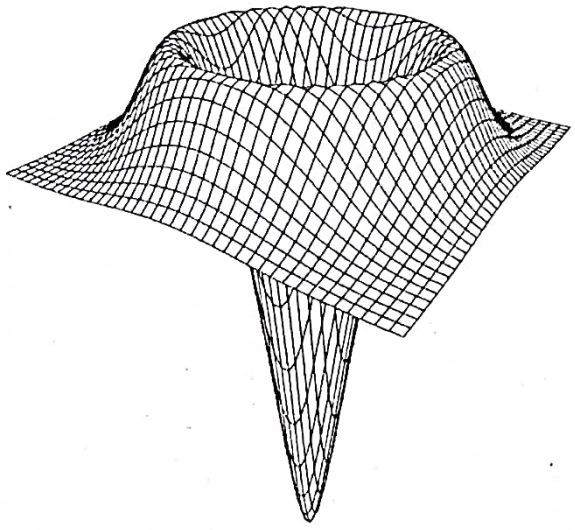

Figure 1: Filtering operator

$$h(x,y) = \int_{\alpha}\int_{\beta} \nabla^2 e^{-((x-\alpha)^2+(y-\beta)^2)/\sigma^2} I(\alpha,\beta) d\beta d\alpha$$

where

$$\nabla^2 = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$$

(a) Give the general name for the type of mathematical operation that computes $h(x, y)$, and the chief purpose that it serves in computer vision.

(b) What image properties should correspond to the zero-crossings of the equation, i.e. those isolated points (x, y) in the image $I(x, y)$ where the above result $h(x, y) = 0$?

(c) What is the significance of the parameter σ? If you increased its value, would there be more or fewer points (x, y) at which $h(x, y) = 0$?

(d) Describe the effect of the above operator in terms of the two-dimensional Fourier domain. What is the Fourier terminology for this image-domain operator? What are its general effects as a function of frequency and as a function of orientation

(e) If the computation of $h(x, y)$ above were implemented entirely by Fourier methods, would the complexity of this computation be greater or less than the image-domain operation expressed above, and when? What would be the tradeoffs involved?

(f) If the image $I(x, y)$ has 2D Fourier Transform $F(u, v)$, provide an expression for $H(u, v)$, the 2D Fourier Transform of the desired result $h(x, y)$ in terms of only the Fourier plane variables (u, v), the image transform $F(u, v)$, and the parameter .

4. BoVW/HoG/Fourier image processing (5 marks)

(a) Fourier analysis decomposes images according to a basis set. What is that basis set? Is the Fourier transform a linear operation? Why or why not?

(b) Explain Histogram of Oriented Gradients(HoG) Feature extraction scheme?

(c) What is the aperture problem? When considering image features, what characteristic(s) should we look for in an effort to avoid this?

5. Machine Learning Basics

(a) What is a supervised classifier? Briefly explain how the Convolutional neural network works. (2 marks)

(b) Given n linearly independent feature vectors in n dimensions, show that for any assignment to the binary labels, you can always construct a linear classifier with weight vector w, which separates the points. Assume that the classifier has the form $sign(w^T x)$. Note that a square matrix composed of linearly independent rows is invertible. (1 mark)

(c) Explain the difference between linear and logistic regression. Write the appropriate mathematical equation in support of your answer. (2 marks)

Indian Institute of Space Science and Technology Trivandrum

II SEMESTER , 2024

ExamType: End Semester

DEPARTMENT OF AVIONICS

computer vision/ computer vision and advanced image processing

(Time allowed: THREE hours)

NOTE: Read all questions first. Attempt all Questions.

1. What is Bi-linear Interpolation? Highlight its applications in computer vision.
   let (x,y) be a point with real coordinates such that $a \leq x < a+1$ and $b \leq y < b+1$ for integers a and b. Also, let $\Delta x = x - a$ and $\Delta y = y - b$ Bilinear interpolation yields an image value at (x, y) that has the following format

   $$I(x,y) = I(a,b) - - - - + I(a+1,b) - - - + I(a,b+1) - - - - - + I(a+1,b+1) - - - -$$

   Fill in the missing parts above. (5 marks)

2. Consider following linear systems (10 marks)

   $$3x + 4y = 2$$
   $$3x + 4y = 3$$

   (a) What are A and b if we write the system in this problem in the following form? $Ax = b$
   (b) What is the rank of A?
   (c) Give a unit column vector r that spans the row space of A.
   (d) Give a unit column vector that spans the null space of A?
   (e) Write the matrix V in the SVD $A = U\Sigma V^T$ of A.
   (f) Compute the matrices U and Σ in the SVD of A.
   (g) Compute Pseudo inverse of A,
   (h) Find the minimum-norm solution x^* of the system $Ax = b$

3. Explain Affine transform. write the necessary equations to explain that. how many DoF are there in Affine transform? (2 marks)

4. Draw and Explain the working principles of the UNeT architecture. How this algorithm can be used for image segmentation. (5 marks)

5. Suppose we are viewing an object 2 meters in front of a camera with a focal length of 50 mm. How far behind the lens will the image of this point be brought into focus? (3 marks)

6. Let us consider a scenario where a 3D point Q is observed by two cameras. Let the 2 camera matrices be given by

$$K_1 = K_2 = \begin{bmatrix} 200 & 0 & 320 \\ 0 & 200 & 240 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation matrices R_1=R_2=I Translation matrices $\mathbf{t}_1 = [0,0,0]^T$,$\mathbf{t}_2 = [100,0,0]^T$. The corresponding 2D points on the images are given by:

$$q_1 = \begin{bmatrix} 520 \\ 440 \\ 1 \end{bmatrix}$$

$$q_2 = \begin{bmatrix} 320 \\ 440 \\ 1 \end{bmatrix}$$

Compute the 3D point Q. (5 marks)

7. What is the use of Loss function in machine learning? Write binary cross entropy loss function and explain each term of this loss function. Write the loss function that can be used for the regression type of problems. (3 marks)

8. What is a Bilateral Filter? Write the mathematical expression in support of your answers. How bilateral filter overcomes the limitations of Gaussian Filters. (2 marks)

9. What is Object Detection? write salient differences between classification, localization, and detection. Draw the architecture of RCNN and FastRCNN. Explain the key aspects of these architectures. (5 marks)

10. Derive Lucas and Kanades optical flow equations from the first principle. What is the main difference between (Horn and Schunck) and (Lucas and Kanades). When Lucas and Kanade method will fail. (4 marks)

11. What is Epipolar Geometry what is its use in stereo vision? Derive the Essential matrix expression using the stereo geometry idea we discussed in class. (3 marks)

12. What is the size of the fundamental matrix, the number of degrees of freedom in it, and the minimum number of point correspondences to compute it using linear algebra, assuming sufficiently many images? Under what conditions is the essential matrix identical to the fundamental matrix? (3 marks)