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 require, use sketches to illustrate your findings.

1. In class, we discussed the concept of homogeneous coordinates. In this example, we will confine ourselves to the real 2D plane. A point $(x, y)^T$ on the real 2D plane can be represented in homogeneous coordinates by a 3-vector $(wx, wy, w)^T$, where $w \neq 0$ is any real number. All values of $w \neq 0$ represent the same 2D point. Dividing out the third coordinate of a homogeneous point (x, y, z) converts it back to its 2D equivalent: $(x/z, y/z)^T$.

Consider a line in the 2D plane, whose equation is given by $ax+by+c = 0$. This can equivalently be written as $l^T x = 0$, where $\mathbf{l} = (a, b, c)^T$ and $\mathbf{x} = (x, y, 1)^T$. Noticing that $\mathbf{x}$ is a homogeneous representation of $(x, y)^T$, we define $\mathbf{l}$ as a homogeneous representation of the line $ax+by+c = 0$. Note that the line $(ka)x + (kb)y + (kc) = 0$ for $k \neq 0$ is the same as the line $ax + by + c = 0$, so the homogeneous representation of the line $ax + by + c = 0$ can be equivalently given by $(a, b, c)^T$ or $(ka, kb, kc)^T$ for any $k \neq 0$. All points $(x, y)^T$ that lie on the line $ax + by + c = 0$ satisfy the equation $\mathbf{l}^T\mathbf{x} = 0$, thus, we can say that a condition for a homogeneous point x to lie on the homogeneous line l is that their dot product is zero, that is, $\mathbf{l}^T\mathbf{x} = 0$. We note this down as a fact:

A point x in homogeneous coordinates lies on the homogeneous line l if and only if $\mathbf{x}^T\mathbf{l} = \mathbf{l}^T\mathbf{x} = 0$

Now let us solve a few simple examples:

(a) Give at least two homogeneous representations for the point $(3, 5)^T$ on the 2D plane, one with $w > 0$ and one with $w < 0$. (1 mark)

(b) What is the equation of the line passing through the points $(1, 1)^T$ and $(-1, 3)^T$ [in the usual Cartesian coordinates]. Now write down a 3-vector that is a homogeneous representation for this line. (1 mark)

(c) Consider the two lines $x + y - 5 = 0$ and $4x - 5y + 7 = 0$. find their intersection in homogeneous coordinates. Convert this homogeneous point back to standard Cartesian coordinates. Is your answer compatible with what basic coordinate geometry suggests? (2 marks)

(d) Consider the two lines $x + 2y + 1 = 0$ and $3x + 6y - 2 = 0$. What is the special relationship between these two lines in the Euclidean plane? What is the interpretation of their intersection in standard Cartesian coordinates? (2 marks)

(e) Write the homogeneous representations of the above two lines and compute their point of intersection in homogeneous coordinates. What is this point of intersection called in computer vision parlance? (2 marks)

(f) Give (with justification) an expression for the homogeneous representation of the line passing through two homogeneous points x1 and x2. (2 marks)

2. Find a line passing through the following points (2 marks)
$u1 = [101, 203]^T$, $u2 = [49, 102]^T$,$u3 = [27, 51]^T$ $u4 = [201, 403]^T$, $u5 = [74, 151]^T$.
We can use SVD here and you can leave the output in terms of SVD.

3. Find a plane passing through the following points (2 marks)
$x1 = [9.99, 101, 203]^T$, $x2 = [5.1, 49, 102]^T$, $x3 = [2.5, 27, 51]^T$ $x4 = [21, 201, 403]^T$, $x5 = [7.6, 74, 151]^T$.
We can use SVD here and you can leave the output in terms of SVD.

4. Find the point of intersection of the 3D line $x = \lambda d + x0$ with the 3D plane $p^T x = 0$. The parameters have the following (2 marks)
$d = [1, 2, 0]^T$,$x0 = [3, 4, 5]^T$, $p = [1, 2, 0, 7]^T$

5. Which of the following statements describes an affine camera but not a general perspective camera?

(a) Relative sizes of visible objects in a scene can be determined without prior knowledge.
(b) Can be used to determine the distance from a object of a known height.
(c) Approximates the human visual system.
(d) An infinitely long plane can be viewed as a line from the right angle

6. (write all the options that apply) Which of the following could affect the intrinsic parameters of a camera?

(a) A crooked lens system.
(b) Diamond/Rhombus shaped pixels with non right angles.
(c) The aperture configuration and construction.
(d) Any offset of the image sensor from the lenss optical center.

7. Which of the following factor does not aect the intrinsic parameters of camera model?

(a) Focal length
(b) Offset of optical center
(c) Exposure
(d) Image resolution

8. Assuming the camera coordinate system is the same as the world coordinate system, the intrinsic and extrinsic parameters of the camera can map any point in homogenous world coordinates to a unique point in the image plane.(True or false, give a suitable reasoning)

9. Consider a world coordinate system W, centered at the origin $(0, 0, 0)$, with axes given by unit vectors $\hat{\mathbf{i}} = (1, 0, 0)^T$, $\hat{\mathbf{j}} = (0, 1, 0)^T$ and $\hat{\mathbf{k}} = (0, 0, 1)^T$. Recall our notation where boldfaces stand for a vector and a hat above a boldface letter stands for a unit vector.

(a) Consider another coordinate system, with unit vectors along two of the orthogonal axes given by $\hat{\mathbf{i}}\text{'} = (0.9, 0.4, 0.1\sqrt{3})$ and $\hat{\mathbf{j}}\text{'} = (-0.41833, 0.90427, 0.08539)$ Find the unit vector, $\hat{\mathbf{k}}\text{'}$, along the third axis orthogonal to both $\hat{\mathbf{i}}\text{'}$ and $\hat{\mathbf{j}}\text{'}$. Is there a unique unit vector orthogonal to both $\hat{\mathbf{i}}\text{'}$ and $\hat{\mathbf{j}}\text{'}$? If not, choose the one that makes an acute angle with $\hat{\mathbf{k}}\text{'}$. (3 marks)

(b) Find the rotation matrix that rotates any vector in the $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ coordinate system to the $\hat{\mathbf{i}}\text{'}$, $\hat{\mathbf{j}}\text{'}$ and $\hat{\mathbf{k}}\text{'}$ coordinate system. (3 marks)

(c) What is the extrinsic parameter matrix for a camera that is located at a displacement $(-1, -2, -3)$ from the origin of W and oriented such that its principal axis coincides with $\hat{\mathbf{k}}\text{'}$, the x-axis of its image plane coincides with $\hat{\mathbf{i}}\text{'}$ and the y-axis of the image plane coincides with $\hat{\mathbf{j}}\text{'}$? (3 marks)

(d) What is the intrinsic parameter matrix for this camera, if its focal length in the x-direction is 1050 pixels, aspect ratio(Aspect ratio is primarily dictated by the size of your camera's sensor, taken from the width and height of an image (W:H). For instance, if your camera sensor is 36mm wide and 24mm high, its aspect ratio would be 3:2.) is 1.0606, pixels deviate from rectangular by 0.573 degrees and principal point is offset from the center $(0, 0)^T$ of the image plane to the location $(10, -5)^T$. (3 marks)

(e) Write down the projection matrix for the camera described by the configuration in parts (c) and (d). (3 marks)

10. Consider set of corresponding points x'_1 and x'_2 in pixel coordinates in two views, which are related by pure rotation R. Assume that all the parameters of the calibration matrix K are known except the focal length f. Describe in detail following steps of an algorithm for recovering R and the focal length of the camera f

(a) The projection points in two views are related by an unknown 3×3 matrix H. Write down the parametrization of matrix H in terms of rotation matrix entries r_{ij} and the focal length f.

(b) Describe a linear least squares algorithm for estimation of matrix H. What is the minimal number of corresponding points needed in order to solve for H ?

(c) Given the parametrization of H derived in a) describe a method for estimating the actual rotation and the focal length of the camera.

11. Explain what are the key ideas that are used in edge detection. What all things can cause an edge in an image. Name a few methods to detect edges in the image. How Sobel edge detection method works. (5 marks)

12. The image below is an image of a 3 pixel thick vertical line. (5 marks)

(a) Show the resulting image obtained after convolution of the original with the following approximation of the derivative filter $[-1;0;1]$ in the horizontal direction. How many local maxima of the filter response do you obtain ?

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(b) Suggest a filter which when convolved with the same image would yield a single maximum in the middle of the line. Demonstrate the result of the convolution on the original image.

(c) What is a difference between Gaussian smoothing and median filtering? How would you decide to use one vs. another ?

13. Given two lines in the image denoted by their projective coordinates l_1 and l_2, how would you compute an intersection point of these two lines ? (1 mark)

14. Show that the line through two 2D points x and x' is $l = x \times x'$. (1 mark)

15. Under homography, how can we write the transformation of points in 3D from camera 1 to camera 2? Explain (mention all the steps needed) how do we estimate the homography. (3 marks)

16. What does it mean for the 2D convolution kernel (filter) to be separable ? (1 mark)

17. Which of the following statements are true of a pinhole camera? Which are false? (1 mark)

(a) (A) Images in a pinhole camera are upside down.

(b) (B) A pinhole camera has a fixed focal length,f.

(c) (C) Images in a pinhole camera are a perspective projection

18. (write all that apply) Which of the following could affect the intrinsic parameters of a camera? (1 mark)

(a) A crooked lens system.

(b) Diamond/Rhombus shaped pixels with non right angles.

(c) The aperture configuration and construction.

(d) Any offset of the image sensor from the lenss optical center.

19. state true or false (with a proper reasoning)

(a) Pincushion distortion modifies the image only along the vertical direction and barrel distortion modifies the image only along the horizontal direction. (1 mark)

(b) The vanishing point associated with a line in 3D space (when viewed through a pinhole camera) can never be a point at infinity (1 mark)

(c) The result of applying a scale, rotation, and translation (in that order) to a vector **v** is equal to the result of applying the same rotation, translation, and scale (in that order) to the same vector **v**. (1 mark)

(d) It is impossible to estimate the intrinsic parameters of the camera given a single image even with prior information about the scene (1 mark)

20. Which of the following always hold(s) under affine transformations? (1 mark)

(a) Parallel lines remain parallel

(b) Ratio of lengths of parallel line segments remain the same

(c) Ratio of areas remain the same

(d) Perpendicular lines remain perpendicular

(e) Angles between two line segments remain the same

21. The figure (1) below shows the outputs of applying one of transformations (projective, affine, similarity, and isometric) to a square with vertices at (1,1), (1,-1), (-1,-1), (-1,1) tell which is the

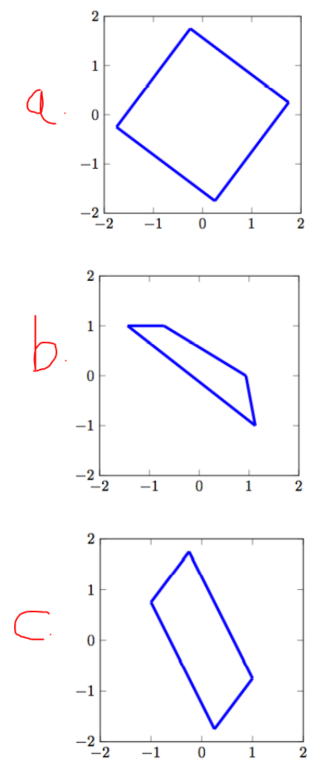

Figure 1: figure for understanding the various transfrom (a) (b) and (c) top-to-bottom respectively)

most specific transformation used to generate each output. if possible write the transformation matrix and write the degrees of freedom. (2 marks)

22. Suppose we want to solve for the camera matrix K and that we know that the world coordinate system is the same as the camera coordinate system. Assume the matrix K has the structure outlined below. Note that K_{33} is an unknown. Assume that we are given n correspondences. Each correspondence consists of a world point (x_i, y_i, z_i) and its projection (u_i, v_i) for $i = 1, ..., n$. (5 marks)

$$\begin{pmatrix} K_{11} & K_{12} & K_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{pmatrix}$$

(a) What is the minimum number of correspondences needed to solve for the unknowns in the matrix K?

(b) Set up an equation of the form $Ax = 0$ to solve for the unknowns in K (where A is a matrix, and x and 0 are vectors). Be specific about what the matrix A and vector x are.

(c) Explain how to solve for the unknowns in the camera matrix K. Make sure $K33 = 1$.