

Exam 1: Classical Vision

● Graded

Student

Vudit Dharmendra Pokharna

Total Points

52.833 / 100 pts

Question 1

Filters & Features

20.5 / 25 pts

1 (a)

+ 1 pt Correct Option

✓ + 0 pts Incorrect Option

+ 2 pts Correct Explanation

✓ + 0 pts Incorrect Explanation

1 b)

✓ + 2 pts i) Correct

+ 0 pts i) Incorrect

✓ + 2 pts ii) Correct

+ 0 pts ii) Incorrect

✓ + 2 pts iii) Correct

+ 0 pts iii) Incorrect

1 (c)

✓ + 1 pt (i) Correct Option

+ 0 pts (i) Incorrect Option

✓ + 1 pt (i) Correct Explanation

+ 0 pts (i) Incorrect Explanation

✓ + 1 pt (ii) Correct Option

+ 0 pts (ii) Incorrect Option

✓ + 1 pt (ii) Correct Explanation

+ 0 pts (ii) Incorrect Explanation

1 (d)

✓ + 2 pts Correct R1 value

+ 0 pts Incorrect R1 value

✓ + 2 pts Correct R2 value

+ 0 pts Incorrect R2 value

+ 0.5 pts Correct Window 1

+ 0 pts Incorrect Window 1

+ 0.5 pts Correct Window 2

+ 0 pts Incorrect Window 2

+ 0.5 pts Correct Scaling Property

+ 0 pts Incorrect Scaling Property

+ 0.5 pts Correct Rotation Property

+ 0 pts Incorrect Rotation Property

1 (e)

+ 0.75 pts (i) Correct Choice

+ 0 pts (i) Incorrect Choice

+ 1 pt (i) Correct Explanation

+ 0 pts (i) Incorrect Explanation

+ 0.75 pts (ii) Correct Choice

+ 0 pts (ii) Incorrect Choice

+ 1 pt (ii) Correct Explanation

+ 0 pts (ii) Incorrect Explanation

+ 0.75 pts (iii) Correct Chhoice

+ 0 pts (iii) Incorrect Choice

+ 0.75 pts (iv) Correct Choice

+ 0 pts (iv) Incorrect Choice

+ 1 pt (iv) Correct Explanation

+ 0 pts (iv) Incorrect Explanation

Question 2

Transformations & Fitting

11 / 24 pts

+ 3 pts 2 (a)

+ 3 pts 2 (b)

✓ + 3 pts 2 (c)

+ 2 pts 2 (d)

✓ + 2 pts 2 (e)

+ 2 pts 2 (f)

+ 2 pts 2 (g)

✓ + 1 pt 2 (h)

+ 1 pt 2 (i)

✓ + 1 pt 2 (j)

✓ + 2 pts 2 (k)

✓ + 2 pts 2 (l)

Question 3

Hough Transform

5.333 / 15 pts

(a)

+ 0 pts (a) Incorrect

+ 3 pts (a) All correct

(b)

+ 0 pts (b) i incorrect

+ 1 pt (b) i correct - theta

+ 1 pt (b) i correct - d

+ 0 pts (b) ii incorrect

+ 1 pt (b) ii correct - theta

+ 1 pt (b) ii correct - d

+ 0 pts (b) iii incorrect

+ 1 pt (b) iii correct - theta

+ 1 pt (b) iii correct - d

(c)

+ 0 pts (c) All incorrect

+ 0.333 pts (c) One correct

+ 0.666 pts (c) Two correct

+ 1 pt (c) Three correct

+ 1.333 pts (c) Four correct

+ 1.666 pts (c) Five correct

+ 2 pts (c) Six correct

+ 2.333 pts (c) Seven correct

+ 2.666 pts (c) Eight correct

+ 3 pts (c) Nine correct

+ 3.333 pts (c) Ten correct

+ 3.666 pts (c) Eleven correct

+ 4 pts (c) Twelve correct

(d)

✓ + 0 pts (d) Incorrect

+ 1 pt (d) Angle correct

+ 1 pt (d) Distance correct

Question 4

4.a

 - 0 pts Correct

- 2 pts Incorrect

4.b

- 0 pts Correct

 - 1 pt Incorrect

4.c

 - 0 pts Correct

- 1 pt Incorrect

4.d

 - 0 pts Correct

- 1 pt Incorrect

4.e

 - 0 pts Correct

- 0.5 pts Partially correct

- 1 pt Incorrect

4.f

- 0 pts q1 Correct

- 0.5 pts q1 Correct choice, wrong explanation

 - 1 pt q1 Incorrect - 0 pts q2 Correct

- 0.5 pts q2 Correct choice, wrong explanation

- 1 pt q2 Incorrect

4.g

- 0 pts Correct

 - 3 pts Incorrect

4.h

- 0 pts Correct

 - 3 pts Incorrect

Question 5

Epipolar Geometry

10 / 22 pts

a

+ 0 pts Click here to replace this description.

✓ + 2 pts Click here to replace this description.

b

✓ + 0 pts Click here to replace this description.

+ 2 pts Click here to replace this description.

c

+ 0 pts Click here to replace this description.

✓ + 2 pts Click here to replace this description.

d

+ 0 pts Click here to replace this description.

✓ + 2 pts Click here to replace this description.

e

+ 0 pts Click here to replace this description.

✓ + 4 pts Click here to replace this description.

f

✓ + 0 pts Click here to replace this description.

+ 3 pts Click here to replace this description.

g

✓ + 0 pts Click here to replace this description.

+ 4 pts Click here to replace this description.

h

✓ + 0 pts Click here to replace this description.

+ 3 pts Click here to replace this description.



Name:

Vidit Pokharna

GTID:

903772087

This exam contains 14 pages (including this cover page) and 5 (multi-part) questions.

There are 100 total points.

Time Limit: 60 Minutes

Important: Make sure to write your name and GTID on this page **and** your GTID in the box at the top right corner of each following page.

For multiple choice questions fill in the square: → ■

Grade Table (for staff use only)

Question	Points	Score
Filters & Features	25	
Transformations & Fitting	24	
Hough Transform	15	
Projective Geometry & Camera Calibration	14	
Epipolar Geometry	22	
Total:	100	



1: Filters & Features (25 points)

- (a) (3 points) Which filter would be the most effective in achieving image smoothing for the given image?



- Gaussian Filter
- Mean Filter
- Median Filter
- Box Filter

Briefly explain the reasoning behind your answer:

The box filter averages out image, and it turn smoothes the image with the filter matrix

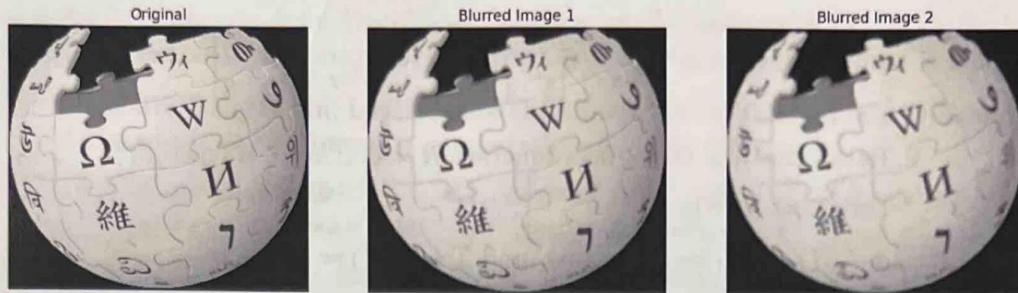
- (b) (6 points) Provide the appropriate mathematical expression for the following properties of convolution (denoted by *), given filters f , g , h .

i. Commutative $f * g = g * f$

ii. Associative $(f * g) * h = f * (g * h)$

iii. Distributive (over +) $f * (g + h) = (f * g) + (f * h)$

- (c) (4 points) In the following question, analyze the effect of hyperparameters in the given image operations.



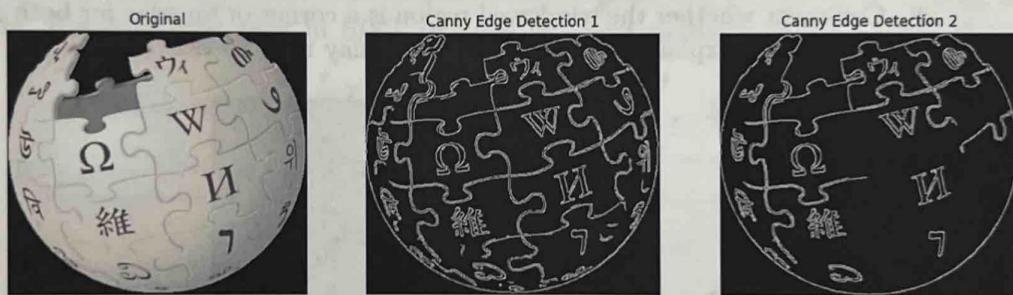
Which image above is the result of using a higher value of σ (standard deviation) while applying Gaussian Blur?

Blurred Image 1

Blurred Image 2

Briefly explain your choice:

Blurred image 2 represents more enhanced blur from the original than blurred image 1, depicting a larger σ used for it.



Which image above is the result of using a high threshold while performing Canny Edge Detection?

Canny Edge Detection 1

Canny Edge Detection 2

Briefly explain your choice:

image 2 does not contain all features from the original image compared to image 1, depicting a higher threshold used leading to less features.

- (d) (6 points) Harris corner detection algorithm computes a 2×2 matrix M by summing over operations on partial gradients in a windowed region. We provide two such matrices from different 3×3 windows of an image:

$$M_1 = \begin{bmatrix} 2 & 5 \\ 5 & 4 \end{bmatrix} \quad M_2 = \begin{bmatrix} 4 & 2 \\ 2 & 8 \end{bmatrix}$$

- i. Recall that the cornerness function $R = \det(M) - \alpha \text{Trace}(M)^2$.

For a 2×2 matrix $\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$,

$$\det(M) = m_{11}m_{22} - m_{12}m_{21} \text{ and } \text{Trace}(M) = m_{11} + m_{22}.$$

Compute the cornerness function R , for both M_1 and M_2 defined above.

Use $\alpha = 1/12$.

Use space below as scratch and put final answer in box.

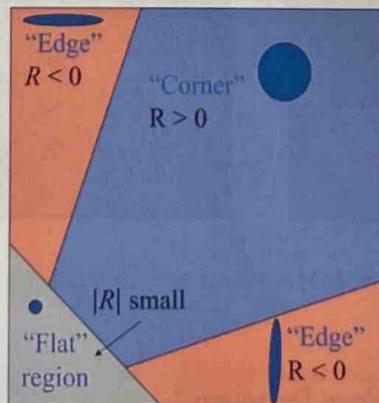
$$\begin{aligned} R_1 &= \det(M_1) - \alpha \text{trace}(M_1)^2 \\ &= (8-25) - \frac{1}{12}(6)^2 = -17 - 3 \\ &= -20 \end{aligned}$$

$$\begin{aligned} R_2 &= \det(M_2) - \alpha \text{trace}(M_2)^2 \\ &= (32-4) - \frac{1}{12}(12)^2 \\ &= 28 - 12 = 16 \end{aligned}$$

$$R_1 = -20$$

$$R_2 = 16$$

- ii. Comment whether the windowed region is a corner or an edge for both M_1 and M_2 . Briefly explain your reasoning. You may refer to the figure below.



M_1 is an edge because $R_1 = -20 < 0$

M_2 is a corner because $R_2 = 16 > 0$

- iii. Is the Harris Corner detector invariant to (i) Scaling and (ii) Rotation?

(i) invariant to scaling

(ii) invariant to rotation



- (e) (6 points) Below are four True/False question on the properties of SIFT (Scale-Invariant Feature Transform). Write a brief explanation if the statement is False.

- i. In the SIFT algorithm, Laplacian of Gaussian filter is preferred over Difference of Gaussians as the former is computationally more efficient.

False. Decomposing 2D convolution into 2 1D convolutions takes advantage of separability of Gaussian filters and lowers time complexity.

- ii. SIFT descriptors are invariant to changes in scale, but are sensitive to changes in rotation.

False. SIFT is not rotation & scale invariant because we use gradients in X and Y directions, which is not rotation/scale invariant.

- iii. Normalizing the SIFT descriptors imparts some degree of illumination invariance to the descriptors.

True. SIFT built on gradient information, capturing underlying edge and texture information in a way inherently more stable to changes in illumination and contrast.

- iv. Keypoint matching in SIFT requires an exact match between all descriptor elements

False. It is based on threshold, not exact match

homework 10: coordinate transformation and convolution

Given two vectors $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

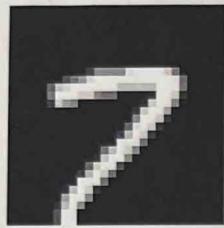
Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

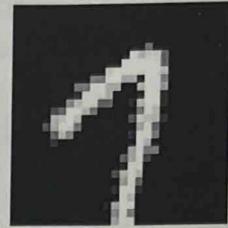
Given $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$, the dot product is given by $\mathbf{v}^T \mathbf{w}$.

2: Transformations & Fitting (23 points)

- (a) (3 points) Provided the original and transformed image, what is the transformation matrix applied to the image?



(a) Original image



(b) Transformed image

$\begin{bmatrix} \cos(30) & -\sin(30) \\ \sin(30) & \cos(30) \end{bmatrix}$

$\begin{bmatrix} \cos(30) & \sin(30) \\ 2 * \sin(30) & \cos(30) \end{bmatrix}$

$\begin{bmatrix} \cos(30) & 1.5 * \sin(30) \\ \sin(30) & \cos(30) \end{bmatrix}$

$\begin{bmatrix} \cos(90) & -\sin(90) \\ \sin(90) & \cos(90) \end{bmatrix}$

- (b) (3 points) In the matrix selected in the previous question, what were the transformations that were applied? (Select all that apply.)

Rotation
 Shear

Rotation then Shear
 Shear then Rotation

- (c) (3 points) Given a 2D vector A : $\begin{bmatrix} x \\ y \end{bmatrix}$ select all transformations that are performed by the following expression on A for **any** x and y . (Select all that apply.)

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

where $a_1, a_2, a_3, a_4, b_1, b_2$ are not zero.

Rotation and translation
 Shear and translation

Only shear
 Only rotation

- (d) (2 points) Affine transformations consist of which of the following transformation types: (select all that apply)

- 1 Projective transformations
 Rotation transformations Translation transformations
 Scaling transformations



Here is the RANSAC Algorithm:

Repeat N times.

- Draw s points uniformly at random
- Fit line to these s points
- Find inliers to this line among the remaining points.
 - Inlier = points whose distance from the line is less than t
- If there are d or more inliers, accept the line and refit using all inliers

(e) (2 points) We define convergence as finding the best fit of the data. What is the effect of N on the convergence of the procedure?

- Increasing N guarantees convergence.
- Decreasing N guarantees convergence.
- Increasing N increases chance of convergence, but does not guarantee it.
- Decreasing N increases chance of convergence, but does not guarantee it.

(f) (2 points) What is the effect of s on the outcome?

- Decreasing s will make the algorithm less sensitive to the outliers.
- s has no effect on the outcome.
- Increasing s will make the algorithm less sensitive to the outliers.

(g) (2 points) What is the effect of d on the outcome? (Select all that apply.)

- In most cases, decreasing d might make the algorithm complete the procedure faster (but not necessarily find the *best fit*).
- Increasing d too much will make the algorithm prone to missing the best fit.
- Decreasing d too much will make the algorithm prone to producing a false best fit.
- d does not have any effect on the convergence of the algorithm.

Given (x, y) co-ordinates, find the resultant co-ordinates (x', y') by performing the corresponding transformation.

Note: You don't need to compute the actual numerical values of your answers. For example, your final answer can be $\sqrt{3}/2$, no need to simplify it further.

- (h) (1 point) Co-ordinates: (2, 3), transformation: rotate by 90° in anticlockwise direction about (0, 0).

$$x' = 2 \cos(90^\circ) - 3 \sin(90^\circ) = -3$$

$$y' = 2 \sin(90^\circ) + 3 \cos(90^\circ) = 2$$

(-3, 2)

- (i) (1 point) Co-ordinates: (1, 1), transformation: rotate by 30° in clockwise direction about (0, 0).

$$\begin{aligned} x' &= 1 \cos(-30) - 3 \sin(-30) = \frac{\sqrt{3}}{2} + \frac{3}{2} \\ y' &= 1 \sin(-30) + 3 \cos(-30) = -\frac{1}{2} + \frac{3\sqrt{3}}{2} \end{aligned}$$

\left(\frac{\sqrt{3}}{2} + \frac{3}{2}, -\frac{1}{2} + \frac{3\sqrt{3}}{2}\right)

- (j) (1 point) Co-ordinates: (3, 4), transformation: shear by a factor of 4 in X-axis.

$$\begin{aligned} x' &= 3 + 4 \cdot 4 = 19 \\ y' &= 0 \cdot 3 + 4 = 4 \end{aligned}$$

(19, 4)

- (k) (2 points) Co-ordinates: (5, 6), transformation: translate by (2, 2) and then rotate by 45° in clockwise direction about (0, 0).

$$\begin{aligned} x' &= 7 \cos(-45) - 8 \sin(-45) = \frac{7\sqrt{2}}{2} + \frac{8\sqrt{2}}{2} = \frac{15\sqrt{2}}{2} \\ y' &= 7 \sin(-45) + 8 \cos(-45) = -\frac{7\sqrt{2}}{2} + \frac{8\sqrt{2}}{2} = \frac{\sqrt{2}}{2} \end{aligned}$$

\left(\frac{15\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)

- (l) (2 points) Co-ordinates: (10, 10), transformation: reflect across $y = -x$ and rotate by 45° in clockwise direction about (0, 0)

$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 10 \end{bmatrix} = \begin{bmatrix} -10 \\ -10 \end{bmatrix}$$

$$\begin{aligned} x' &= -10 \cos(-45) + 10 \sin(-45) = -10\sqrt{2} \\ y' &= -10 \sin(-45) - 10 \cos(-45) = 0 \end{aligned}$$

(-10\sqrt{2}, 0)

3: Hough Transform (15 points)

(Use π directly in the result for this part)

Given a clip of the edges (E) and gradient angles (A) of an image, we want to use the Hough Transform to find lines in the image. Specifically, we are looking for lines with angles of $\theta \in [-\pi/4, 0, \pi/4, \pi/2]$. Recall that when using Hough Transform, we consider lines in polar coordinates of the form:

$$d = x \cos(\theta) + y \sin(\theta)$$

Here's a brief idea about Hough Transform line detection algorithm. After achieving the gradient angles for each valid image pixel (given below), we will transform the pixels into polar coordinates. Then, the vote is increased by one if there's a pixel agree with the equation given above. The coordinates with the highest vote are considered as the detected line.

E: Edge Image			A: Gradient Angle			Reference Table					
	x	y		x	y	f	θ	$-\pi/4$	0	$\pi/4$	$\pi/2$
0	0	0	0	0	0	0	$-\pi/4$				
-1	0	1	-1	0	-1	0	$-\pi/4$	0			
-2	1	1	-2	-2	-1	-2	0	0			
	0	1		0	1		0	1	2	1	0

- (a) (3 points) Which of the following image pixel(s) (i.e., (x, y) coordinates) will contribute votes in the Hough space? (Select all that apply)

- $\circ =$ $(x, y) = (0, 0)$ $(x, y) = (1, 0)$ $(x, y) = (2, 0)$
 $\partial = -1 \sin(0)$ $(x, y) = (0, -1)$ $(x, y) = (1, -1)$ $(x, y) = (2, -1)$
 $l = -2 \sin(-\frac{\pi}{4})$ $(x, y) = (0, -2)$ $(x, y) = (1, -2)$ $(x, y) = (2, -2)$

- (b) (6 points) Specify θ and d for the following (x, y) coordinates considering the gradient angles give above:

- i. $(0, 0)$: $\theta =$ 0 $d =$ 0
ii. $(2, 0)$: $\theta =$ $-\pi/4$ $d =$ 1
iii. $(2, -2)$: $\theta =$ 0 $d =$ 1

- (c) (4 points) Fill in the Hough accumulated votes for the appropriate angles, θ , and quantized d values below. You can assume the implementation only casts votes to bins that match the gradient angles given and the array is initialized with zeros.

Hint: start by writing θ and calculating d per remaining edge point.

Hough Array

	$-\pi/4$	1	1	1
θ	0	1	0	12
	$\pi/4$	0	0	1
	$\pi/2$	1	1	0

$\sqrt{2}$ 1 2 d

- (d) (2 points) Based off your Hough Array, what parameters d and θ correspond to the best fit line in your image?

$$\theta = \underline{\quad 0 \quad}$$

$$d = \underline{\quad 1 \quad}$$

4: Projective Geometry & Camera Calibration (14 points)

(a) (2 points) Where do the projections of parallel 3D lines intersect?

- Optical center
 A vanishing point

- A perpendicular bisector
 Infinity

(b) (1 point) How many intersecting points does the perspective projection of the edges of a cube define?

- 1
 3

- 6
 12

(c) (1 point) How many minimum pairs of corresponding points do we need for estimating the Homography matrix?

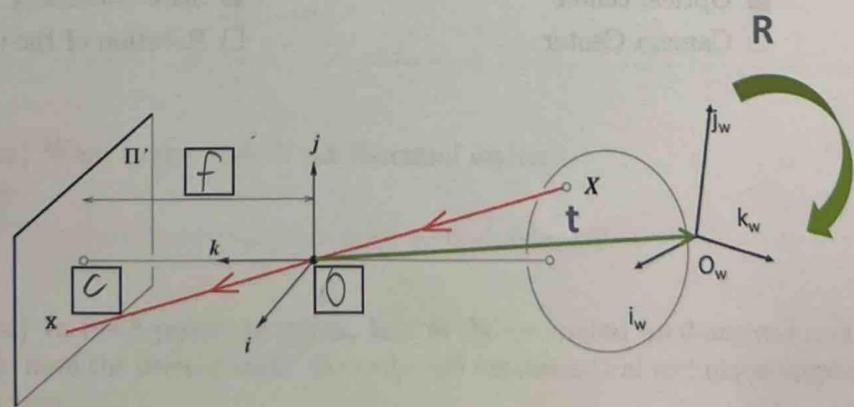
- 4
 6

- 8
 12

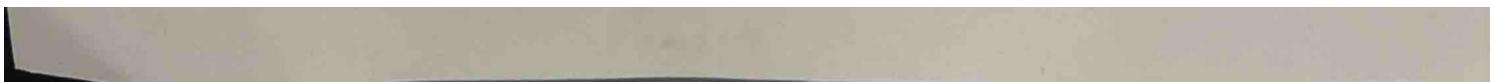
(d) (1 point) Which of the following are invariant (doesn't change) to Projective Transformations?

- Circle
 Parallel Lines

- Square
 Straight Lines

(e) (1 point) Identify the following elements in the camera projection diagram below. Write the letters f , C , and O in the boxes next to the correct location in the diagram. $f \rightarrow$ Focal Length, $C \rightarrow$ Optical Centre, $O \rightarrow$ Camera Centre

(f) (2 points) Below are two True/False questions related to Perspective Projection and Camera Calibration. Write a brief explanation if the statement is False.



- i. If two lines intersect in an image, they must be intersecting lines in the 3D world.

True

- ii. In projective geometry, the lengths of objects are preserved when projected from 3D to 2D.

False: due to perspective projection, objects may be distorted in the process and may not preserve length.

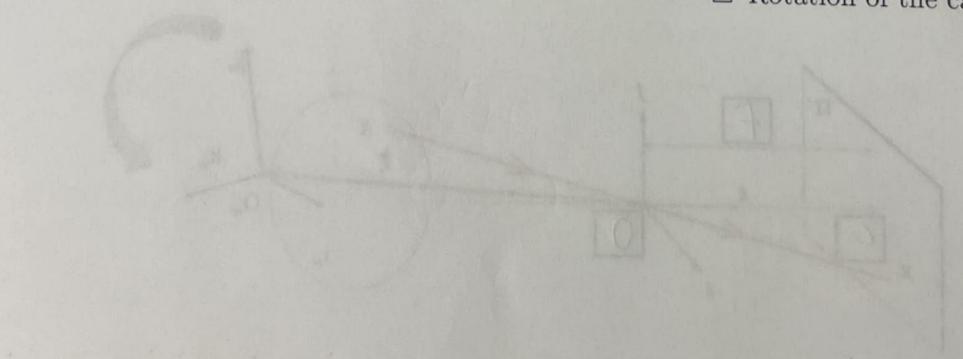
Recall that the camera projection equation is given by: $p = K [R \mid t]P$

- (g) (3 points) Which of the following claims are correct for the camera calibration equation below? (Select all that apply).

- K (intrinsic matrix) contains 6 parameters, R and t contain 5 extrinsic parameters.
- K (intrinsic matrix) contains 5 parameters, R and t contain 6 extrinsic parameters.
- R and t represent transformations of the world coordinate system defined in the camera coordinate system.
- R and t represent transformations of the camera coordinate system defined in the world coordinate system.

- (h) (3 points) Which of the following are components of the intrinsic matrix K ? (Select all that apply).

- | | |
|--|--|
| <input checked="" type="checkbox"/> Focal length | <input type="checkbox"/> Aspect ratio |
| <input checked="" type="checkbox"/> Optical center | <input checked="" type="checkbox"/> Skew coefficient |
| <input type="checkbox"/> Camera Center | <input type="checkbox"/> Rotation of the camera |





5: Epipolar Geometry (22 points)

- (a) (2 points) Which matrix/matrices are used to describe the relationship between corresponding points in stereo images, and is defined in the space of the original image coordinates?
select all that apply.

Projection matrix
 Fundamental matrix

Homography matrix
 Essential matrix

- (b) (2 points) Which of the following statements is/are true regarding the 8-point algorithm?
select all that apply

It only works with calibrated cameras.
 It can be applied to uncalibrated stereo images.

It requires knowledge of the camera's intrinsic parameters.
 It is limited to estimating the essential matrix. \times

- (c) (2 points) How many degrees of freedom (DOF) are there in the Essential matrix?

5

- (d) (2 points) What is the rank of the Essential matrix?

2

- (e) (4 points) In the 8-point algorithm, how is the estimated fundamental matrix enforced to have the correct rank? Describe the mathematical technique employed for this purpose.

By enforce the rank 2 constraint by performing SVD on F and setting a singular value to zero

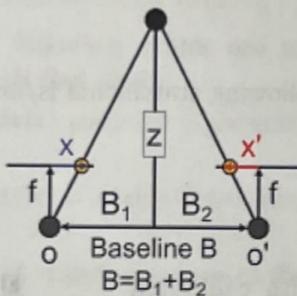
$$F = V \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T$$

↑ rank 2

- (f) (3 points) In images captured through the forward motion of a camera, comment on the epipole's position across both images and on the appearance of the corresponding epipolar lines in this scenario.

Epipole is projected horizontally due to forward motion and epipolar lines are perpendicular to image plane.

Consider the stereo setup below for two identical cameras with no relative rotation, each with focal length f . We would like to recover the depth of the 3D world point that projects onto the corresponding image points, x and x' in camera o and o' , respectively.



- (g) (4 points) Using similar triangles, write the two ratio equations needed to solve for depth, Z , with triangulation:

Hint: The right pixel, x' , is left of the optical center (i.e., negative coordinate value)

$$\text{Ratio 1: } \frac{f}{x} = \frac{z}{B_1} \quad \text{Ratio 2: } \frac{f}{x-x'} = \frac{z}{B_1+B_2}$$

- (h) (3 points) Solve for depth, Z , by combining the ratios above. Show your derivation in the space below and put your final answer in the box. Your answer should be of the form, $Z = \dots$

$$Z = \frac{(x-x')B_1}{x(B_1+B_2)}$$

$$\begin{aligned} \frac{fx}{B_1} &= \frac{f(x-x')}{B_1+B_2} \\ \cancel{fx} &= \cancel{f(x-x')} \frac{B_1}{B_1+B_2} \\ \cancel{f} &= \cancel{x} \frac{x'}{B_1} \\ \cancel{f} &= \cancel{x} \frac{x'}{B_2} \\ \frac{x}{B_1} &= \frac{x'}{B_2} \\ \frac{x}{B_1} &= \frac{x'}{B_2} \\ \frac{B_1}{B_1+B_2} &= \frac{x}{x+x'} \end{aligned}$$