Name:

GTID:

This exam contains 12 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) (6 points) Identify the filters below and their functions:

a) $\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ c) $\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

(b) (4 points) Write the time complexity of a Gaussian filter, and a Separable Gaussian filter, in big O notation.

i. Guassian filter ______

ii. Seperable Guassian filter ______

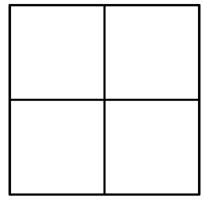
Which one of them exhibits a lower time complexity and why?

(c) (5 points) For an image, I, consider the following responses after filtering with an x and y derivative (difference) filter.

$\frac{dI}{dx} = \begin{array}{|c|c|} \hline 1 & 4 \\ \hline 4 & 2 \\ \hline \end{array}$ $\frac{dI}{dy} = \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & 2 \\ \hline \end{array}$

i. Compute the corresponding gradient magnitude image:
(Answers may include square roots.)

ii. Write the corresponding binary edge image for a threshold of 3.



(d) (6 points) Below are three True/False question related to Harris Corner Detection. λ_1 and λ_2 are eigenvalues of the matrix M for a region. Write a brief explanation if the statement is False.

i. The region is an edge if $\lambda_1 >> \lambda_2$

ii. The region is a corner if λ_1 and λ_2 are small values.

iii. The region is flat if $\lambda_2 >> \lambda_1$

(e) (4 points) You are given an interest point detector that can choose regions of interest at different scales and rotations. Now you need to select an algorithm to describe these regions. You have both RGB pixel values and per-pixel gradient angles, re-oriented by the direction of max gradient. You consider the following:

- **A**: RGB pixel values concatenated into a vector
- **B**: A normalized histogram of RGB pixel values
- **C**: Re-oriented gradient angles concatenated into a vector
- **D**: A normalized histogram of re-oriented gradient angles

To help you make your choice, select all features that satisfy the following properties.

i. Invariant to scaling

☐ A ☐ B ☐ C ☐ D

ii. Invariant to rotations

☐ A ☐ B ☐ C ☐ D

iii. Retains local geometric information

☐ A ☐ B ☐ C ☐ D

iv. Invariant to photometric variations

☐ A ☐ B ☐ C ☐ D

2: Transformations & Fitting (25 points)

(a) (3 points) Which of the following are true for RANSAC? *(select all that apply)*

- ☐ It employs an iterative algorithm.
- ☐ Hyerparameter tuning is necessary.
- ☐ It is extremely sensitive to outliers.
- ☐ It will converge even when given a uniform random set of points.

Recall the following equations that describe a transformation of (x, y) into (x', y').

2D rotation around (0,0)	2D shear
x' = x $\cos\theta - y\sin\theta$	x' = x + αy
y' = x $\sin\theta + y\cos\theta$	y' = $\beta x + y$

(b) (3 points) Construct a transformation matrix that results in a 2D rotation by $3\pi/4$ radians around (0,0). *Note:* $cos(\pi/4) = 1/\sqrt{2}$, $sin(\pi/4) = 1/\sqrt{2}$

(c) (3 points) Construct a transformation matrix that results in a 2D shear by $\alpha = 6$ in the horizontal direction and by $\beta = 3$ in the vertical direction.

(d) (2 points) Which of these 2D affine transformations can be represented by a single 2x2 matrix multiplication to the input vector: *(select all that apply)*

- ☐ Shearing transformations
- ☐ Rotation transformations
- ☐ Translation transformations
- ☐ Scaling transformations

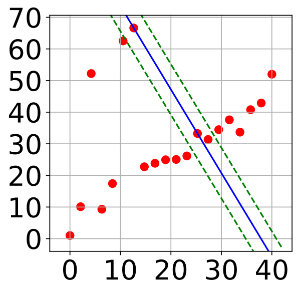
(e) (3 points) The following matrix represents a transformation. What is the transformation?

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

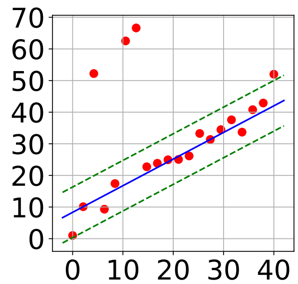
- ☐ Projection on $y = x$.
- ☐ Reflection across $y = -x$.
- ☐ Rotation across $y = -x$.

☐ Rotation $y = x$.

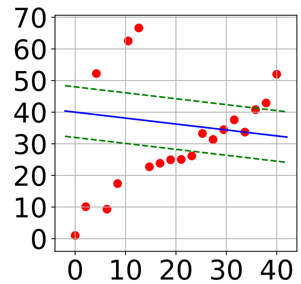
Assume that we run RANSAC for 3 iterations as pictured below. The blue solid line represents the line we have fit, while green dotted lines represent the threshold for inliers (any points within or on the green dotted lines are inliers).



(a) iteration1



(b) iteration2



(c) iteration3

(f) (6 points) For each run of the algorithm, compute the number of inliers and indicate whether the model is stored or not.
You can assume that the hyperparameter for the minimum number of inliers is $d = 5$

Iteration	**Number of Inliers**	**Model Stored as Current Best? *(yes/no)***
1		
2		
3		

(g) (1 point) Which of the three iterations will result in the model selected by RANSAC?

☐ Iteration 1 ☐ Iteration 2 ☐ Iteration 3

(h) (2 points) In which situation would you prefer to use RANSAC over K-means? *(Select one)*

☐ The data has very few outliers.
☐ The data has a lot of outliers.

(i) (1 point) After an affine transformation *(Select one)*

☐ Parallel lines become intersecting lines.
☐ Parallel lines remain parallel.

3: Hough Transform (15 points)

Given the following (x, y) coordinates, what are their possible (d, θ) coordinates after being transformed into Hough Space? Recall that when using the Hough Transform, we consider lines in polar coordinates of the form:

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

(a) (2 points) $x = 0, y = 0$

☐ $(d, \theta) = (0{,}0)$ ☐ $(d, \theta) = (0, \frac{-8\pi}{33})$ ☐ $(d, \theta) = (2{,}0)$

(b) (2 points) $x = 2, y = 1$

☐ $(d, \theta) = (2{,}1)$ ☐ $(d, \theta) = (\frac{\sqrt{2}}{2}, \frac{\pi}{4})$ ☐ $(d, \theta) = (0, \pi/6)$

(c) (2 points) $x = -1, y = 2$

☐ $(d, \theta) = (\frac{\sqrt{2}}{2}, \frac{-\pi}{4})$ ☐ $(d, \theta) = (-1{,}0)$ ☐ $(d, \theta) = (-\frac{1}{2}, \frac{\pi}{2})$

Choose all the correct answers for the following questions.

(d) (2 points) What are the main differences between using Hough transform for detecting lines and Hough transform for detecting circles?

- ☐ The Hough Transform for lines uses a 2D parameter space, while the Hough Transform for circles uses a 3D parameter space.
- ☐ The Hough Transform for circles can detect lines, but the Hough Transform for lines cannot detect circles.
- ☐ There is no difference; both use the same parameter space and techniques.

(e) (2 points) What are the benefits of using Hough Transform for line detection?

- ☐ Robustness to noise
- ☐ Tolerance to disconnected segments
- ☐ High detection speed and low memory use
- ☐ All of the above

(f) (2 points) What does a point in the accumulator space of a Hough Transform represent?

- ☐ The intensity of a pixel in the original image
- ☐ A potential line in the image space
- ☐ The color of a detected line
- ☐ The gradient of an edge in the image

(g) (3 points) In a 300x300 pixel image, the Hough Transform is applied to detect lines. After processing the image, the highest value in the accumulator matrix is found at the coordinates (90, 106). What do these coordinates represent in the context of the Hough Transform?

- ☐ The slope and intercept of the detected line in the image space.
- ☐ The angle and distance from the origin of the detected line in the image space.
- ☐ The number of lines and their average length in the image space.
- ☐ The center and radius of a detected circle in the image space.

4: Projective Geometry & Camera Calibration (14 points)

(a) (2 points) In the context of camera calibration, what does the extrinsic matrix account for?

- ☐ Properties of the camera lens
- ☐ Rotation and translation of the camera with respect to the world
- ☐ The light absorption of the camera sensor
- ☐ The resolution of the camera sensor

(b) (2 points) How many vanishing points can be found in an image where parallel lines are converging?

- ☐ One for each set of parallel lines
- ☐ Only one regardless of the number of parallel lines
- ☐ Two for each set of parallel lines
- ☐ It varies depending on the angle of the camera

(c) (3 points) Below are three True/False question related to Perspective Projection and Camera Calibration. Write a brief explanation if the statement is False.

i. Perspective projection always maintains the parallelism of lines from the 3D world in the 2D image plane.

ii. When calibrating a camera, the intrinsic matrix is used to correct for lens distortion.

iii. A square in the real world always projects as a square in the image captured by a camera.

Recall that the camera projection matrix is given by: $M = K\,[R \mid t]$

(d) (3 points) Briefly describe what the camera projection matrix is used for.

(e) (4 points) How many degrees of freedom are in each of the following matrices/vectors?

i. 3D Rotation matrix, R ____________

ii. 3D translation vector, t ____________

iii. Camera Intrinsic matrix, K ____________

iv. Camera Projection matrix, M ____________

5: Epipolar Geometry (22 points)

(a) (2 points) Which of the following quantities are required to be known in order to **solve for** (not define) the Essential matrix using the epipolar constraint?
select all that apply.

- ☐ R (rotation)
- ☐ t (translation)
- ☐ K (instrincs)
- ☐ 3D to 2D correspondences
- ☐ 2D correspondences

(b) (2 points) Which of the quantities needed to solve for the Essential matrix (i.e., selected above) are unknown when solving for the Fundamental matrix?

(c) (3 points) What is the dimension of the Essential / Fundamental matrices?

(d) (3 points) What is the rank of the Essential / Fundamental matrices?

(e) (4 points) When estimating the fundamental matrix F using the 8-point algorithm, you compute the SVD (Singular value decomposition) of the estimated F matrix, as shown below.

$$F = U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} V^T$$

Does this estimated F give the correct result?

☐ Yes ☐ No

If not, what steps would you take to fix it?

(f) (3 points) Why would one choose to use image rectification before searching for correspondences across stereo images?

(g) (3 points) Write the steps in basic stereo matching algorithm.

(h) (2 points) In stereo vision, what does the term "baseline" refer to?
select all that apply.

- ☐ The distance between the principal point and the image plane.
- ☐ The distance between the optical centers of two cameras.
- ☐ The distance between the camera and the object being viewed.
- ☐ The distance between the lenses and the image sensor.
- ☐ The distance between the camera lens and the camera body.