

Midterm 1

Computer Vision I (CSCI 3240U)

October 8, 2026 (70 min.)

Family name: _____

Given names: _____

Student number: _____

Question	Marks		
1	_____	/	20
2	_____	/	10
3	_____	/	10
4	_____	/	10
5	_____	/	20
Total	_____	/	70

Instructions

- You may consult one double-sided letter-size cheatsheet. Calculators are allowed. No other aids are permitted.
- Write in pen.
- Show your working. Partial marks are awarded for a correct method with a small arithmetic error.
- State any assumption you make (for example, the rotation convention, the boundary handling, or the sign of a derivative kernel).
- Write “**I do not know**” to receive 10% marks in lieu of an answer.
- Total pages: 16.

Question 1 [20 Marks]

Consider a camera whose rotation matrix and translation are

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \mathbf{t} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix},$$

respectively.

Assume the principal axis is along the camera's z -direction, the focal length is 4, the principal point is at the origin, and pixels are square with unit aspect ratio.

Part A [10%]

Write down the *intrinsic matrix* $\mathbf{K}$ for this camera.

Part B [25%]

Write down the *extrinsic matrix* $[\mathbf{R} \mid \mathbf{t}]$ for this camera.

Part C [25%]

Write down the camera projection matrix $\mathbf{P} = \mathbf{K}[\mathbf{R} \mid \mathbf{t}]$.

Part D [25%]

Now use projection matrix $\mathbf{P}$ to project the world point $\mathbf{p}_{\text{world}}^T = (2, 4, 1)$ to the image plane.

Part E [15%]

Check your answer by (a) expressing $\mathbf{p}_{\text{world}}^T$ in the camera coordinate system and (b) applying the pinhole camera model to compute the projection of the transformed point in the image plane.

This page is intentionally left blank.

Question 2 [10 Marks]

Consider a signal

$$x = [2 \quad -1 \quad 1]$$

and a kernel

$$h = [1 \quad 2 \quad -1 \quad 0 \quad 3 \quad -1 \quad 2].$$

Compute $x * h$. Assume “reflect” boundary conditions.

This page is intentionally left blank.

Question 3 [10 Marks]

The bi-variate Gaussian is

$$G(\mathbf{x}; \mu, \Sigma) \propto \exp \left(-\frac{1}{2} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right).$$

Evaluate the argument to the exponent for $\mathbf{x} = \begin{pmatrix} 2 & 1 \end{pmatrix}^T$ when $\mu = \begin{pmatrix} -1 & 3 \end{pmatrix}^T$ and $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Part A [15%]

Compute $(\mathbf{x} - \mu)^T$.

Part B [25%]

Compute Σ^{-1} .

Part C [30%]

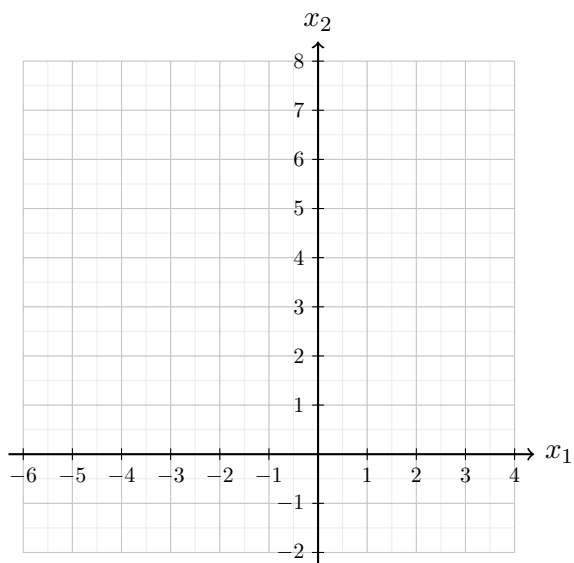
Compute the argument to the exponent.

Part D [10%]

Write down the argument to the exponent at the mean.

Part E [20%]

On the axes below, sketch a representative iso-contour of this Gaussian. Mark the mean μ , and state whether the contour is (i) a circle, (ii) an axis-aligned ellipse, or (iii) a tilted ellipse. Justify your choice in one sentence.



Question 4 [10 Marks]

Part A [40%]

Answer True or False.

1. [**T** / **F**] Under a pinhole camera, distant objects appear smaller.
2. [**T** / **F**] A Gaussian pyramid can be inverted exactly to recover the original image.
3. [**T** / **F**] Convolution with a Gaussian is a high-pass operation.
4. [**T** / **F**] Applying a Gaussian derivative filter to an image gives the same result as first smoothing the image with a Gaussian and then differentiating.

Part B [60%]

The 3×3 neighbourhood below has been extracted from an image, centred on a pixel at position (m, n) :

$$I = \begin{pmatrix} 40 & 45 & 50 \\ 30 & 35 & 40 \\ 20 & 25 & 30 \end{pmatrix}.$$

Rows increase downward (y increases downward), columns increase to the right. Apply the Sobel kernels

$$S_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}, \quad S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix},$$

treated as *convolution* kernels, to compute I_x and I_y at the centre pixel. Then report the gradient magnitude $\|\nabla I\|$ and the gradient direction θ (in degrees).

This page is intentionally left blank.

Question 5 [20 Marks]

We are asked to convolve a 512×512 image with a 2D kernel of half-width 15. In this context, answer the following:

Part A [33%]

Write down the number of operations (additions, multiplications) needed to perform this convolution directly.

Part B [33%]

Now assume that the kernel is separable. Write down the number of operations needed if we take advantage of separable convolution.

Part C [34%]

Now assume that the kernel is a **box filter**, i.e., all $(2k + 1)^2$ entries are equal to $1/(2k + 1)^2$. Using an *integral image* (summed-area table) of the input, write down the number of operations needed for:

- (i) precomputing the integral image, and
- (ii) computing the box-filtered output for the whole image using the integral image.

End of exam.