Question 1 ♣ Let W be a 3×3 window centered about the pixel p. Let the gradient vectors in this window (top to bottom, left to right) be $\{(0,0), (0,1), (0,1), (0,0), (-1,0), (0,0), (0,0), (-1,0), (0,0)\}$. Compute the correlation matrix C in this window without normalizing it then compute a as the sum of of entries in this matrix. Let Q be the matrix $\sum_i \nabla I(p_i) \nabla I(p_i)^T x_i$ where $\nabla I(p_i)$ are gradients at locations in W and x_i are the x-coordinates of corresponding locations p_i in W. Assume that the first row of Q is given by $[1 \;\; 2]$ and that the second row of Q is given by $[2 \;\; 3]$. Using the matrix Q and the matrix C you computed before, compute the location t of the corner in this window. Compute b as the sum of the x- and y-coordinates of the location t. Mark **all** the statements that are **correct**.

A b=6	D b=3	G b=7	J a=3	M a=2	P a=0
B b=0	E a=-1	H a=4	K a=5	N a=7	Q a=1
C b=8	F b=5	I b=1	L b=2	O b=4	R a=6

Question 2 ♣ Let v the projected motion field obtained by projecting an actual 3D motion field V. Let $\hat{v}$ be the optical flow field observed in the images. Let I_x and I_y be the x- and y-derivatives of the image at a given time, respectively. Let I_t be the time derivative at a given location. Mark **all** the statements that are **correct**.

Aperture Problem: When observing motion through a small aperture or in a local neighborhood, we can only detect motion perpendicular to the edge or feature within that neighborhood. This leads to ambiguity in determining the full motion vector.

It introduces a global smoothness constraint across the entire image.In contrast to local methods like Lucas-Kanade, which solve several small problems independently, Horn-Schunck solves one large global problem

radial or linear motion field depending on motion type

- [ ] A The projected motion field v is always a radial motion field.
- [ ] B The Horn-Schunck algorithm for optical flow estimation considers the estimation of each optical flow vector at each location independent of all other locations.
- [x] C The projected motion vectors v of a pure translation in which there is a translation in Z (i.e. $T_z \neq 0$) form a radial motion field.
- [x] D The relative motion field of instantaneously coincident points is always a radial motion field.
- [ ] E When decomposing the projected motion field according to the translational and rotational components of the 3D motion V, the translational component does not depend on the depth Z.

the rotational comp does not depend on z

- [ ] F Assuming that the motion is constant in a neighborhood allows finding the optical flow vector in the center of the neighborhood regardless of the content of the neighborhood.

This assumption alone is insufficient due to the aperture problem

- [x] G The Horn-Schunck algorithm for optical flow estimation assumes that the image brightness at each object point is constant.

book

- [ ] H The Optical Flow Constraint Equation (OFCE) at a given point allows computing the optical flow vector at that location without any ambiguity.

It provides one equation for two unknowns, leading to the aperture problem

- [x] I The Optical Flow Constraint Equation (OFCE) assumes that the image brightness at each object point is constant.

Question 3 ♣ Let $P = (1, 2, 3)$ be a world point and $p = (1, 1)$ be a corresponding image point. Let M be the unknown perspective projection matrix. Let m_1^T, m_2^T, m_3^T be the rows of the matrix M. Let x be a vector composed of the three rows of M. Let $Ax = \mathbf{0}$ be the system of 2 equations that is formed using the points P and p to solve for x. Write the elements of the 12×2 matrix A using the corresponding points P and p. Compute a as the sum of the elements in this matrix. Let the the first, second, and third rows of M be $[2,\ 0,\ ,1\ ,0]$, $[1,\ 2,\ ,0\ ,1]$, $[0,\ 1,\ ,0\ ,2]$, respectively. Let ρ be the unknown scale of the estimated matrix M. Compute b as the absolute value of ρ. Mark **all** the statements that are **correct**.

$Am = 0$

$$\underbrace{\begin{bmatrix} P_i^T & 0^T & -x_i P_i^T \\ 0^T & P_i^T & -y_i P_i^T \end{bmatrix}}_{2n \times 12} \underbrace{\begin{bmatrix} m_{11} \\ \vdots \\ m_{33} \end{bmatrix}}_{12 \times 1} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{2n \times 1}$$

$n = 1$ here

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 & 0 & 0 & -1 & -2 & -3 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 & 1 & -1 & -2 & -3 & -1 \end{bmatrix} \Rightarrow a = 0$$

$$m_{33} = 2 \;;\; \rho = \frac{1}{2} = 0.5 = b$$

per normalization

A a=-2	D a=6	G b=$\sqrt{3}$	J b=1	M a=-6	P a=0
B b=$\sqrt{2}$	E a=7	H b=3	K a=12	N b=0	Q a=-12
C b=2	F b=$\sqrt{5}$	I a=-4	L b=4	O a=14	R b=5

Question 4 ♣ Let R be a 4×4 3D rotation matrix representing rotation by 90 degrees about the Z-axis in homogeneous coordinates. Compute a as the sum of elements in R. Let T be a 4×4 3D translation matrix representing translation by $[1, 2, 3]$ in homogeneous coordinates. Compute b as the sum of elements in T. Mark **all** the statements that are **correct**.

$$R_z(\theta) = \begin{bmatrix} \cos 90 & -\sin 90 & 0 & 0 \\ \sin 90 & \cos 90 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$a = 2$

$$T = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$b = 10$

- [A] b=10
- [B] b=9
- [C] b=8
- [D] b=4
- [E] a=4
- [F] b=5
- [G] a=2
- [H] b=12
- [I] a=3
- [J] a=1
- [K] a=0
- [L] b=7
- [M] a=3.5
- [N] a=1.5
- [O] a=0.5
- [P] a=2.5
- [Q] b=11
- [R] b=6

Question 5 ♣ Let S be a set of data points that contain several outliers and/or Gaussian noise. Let l be a model that is estimated based on the set S. Let $E(l)$ be the error computed for the model l using the set S. Mark **all** the statements that are **correct**.

- A Using the RANSAC algorithm is a good approach for trying to deal with outliers.
- B The RANSAC algorithm continuously updates the distance threshold used to detect inliers.
- C The best model l is the one that maximizes $E(l)$.
- D The RANSAC algorithm continuously updates the number of trials that has to be conducted.
- E The gradient of $E(l)$ is commonly used in trying to find the optimal l.
- F Outliers are important for an accurate model and should be included when computing the model parameters.
- G Using the RANSAC algorithm is a good approach for trying to deal with Gaussian noise.
- H The RANSAC algorithm continuously updates the number of points that has to be drawn at each evaluation.
- I The Geman-McClure function used in M-estimators gives higher weight to large errors.

Question 6 ♣ Let K_l^* and K_r^* be the intrinsic camera parameters of a stereo system. Let R and T be the rotation and translation of the right image with respect to the left image. Let p_l and p_r be a pair of corresponding points on the left and right images, respectively, given in image coordinates (i.e. pixels). Mark **all** the statements that are **correct**.

depending on the camera configuration.
it can be outside the images

x [A] The epipoles must always be inside the images.

✓ [B] Given the point p_l in the left image the right epipolar line in the right image corresponding to p_l is given by Fp_l. F-->as pl in img coord

x [C] The matrices K_l^*, K_r^*, R, and T, are all necessary for computing the essential matrix E.

x [D] The matrix F must always be a full rank matrix.
rank 2

✓ [E] Normalizing the set of corresponding points helps in obtaining a more stable estimate of the fundamental matrix.

F is 3x3-->9-1=8 equations so 8 points needed

x [F] The fundamental matrix may be estimated based on a set of 6 corresponding points in the left and right images.

x [G] Given the point p_l in the left image the right epipolar line in the right image corresponding to p_l is given by $p_r^T F$.

✓ [H] The matrices K_l^*, K_r^*, R, and T, are all necessary for computing the fundamental matrix F. can be found without these but with 8 pt algorithm

✓ [I] The epipoles may be computed by applying SVD to the fundamental matrix F.

The epipoles correspond to the left and right null spaces of F.

Question 7 ♣ Assume an axis-aligned stereo system where the distance between the optical centers is $T = 4$, the focal length of both cameras is $f = 2$. Let $p_l = (1, 2)$ and $p_r = (2, 2)$ be a set of corresponding points in the left and right images respectively (in camera coordinates). Compute a as the z-coordinate of the 3D point which produced p_l and p_r (in camera coordinates). Assume that the intrinsic camera parameters matrix is given by K^* where the first, second, and third rows of this matrix are given by $[2,\ \ ,0\ \ ,0\ \ ]$, $[0,\ \ ,2\ \ ,0\ \ ]$, $[0,\ \ ,0\ \ ,1\ \ ]$, respectively. Let $\bar{p}_l$ be the image coordinates of the point p_l. Use the matrix K^* to compute $\bar{p}_l$. Compute b as the sum of the x- and y-coordinates of $\bar{p}_l$. Mark **all** the statements that are **correct**.

$$z = f\frac{T}{d} = \frac{fT}{(x_r - x_l)} = \frac{2 \times 4}{2-1} = 8$$

$$\bar{p}_l = K_l\, p_l = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} = (2, 4) \Rightarrow b = 6$$

A b=3	D b=8	G a=6	J b=6	M b=2	P b=1
B a=8	E a=-2	H a=-6	K b=0	N b=5	Q b=4
C a=4	F a=-4	I b=7	L a=2	O a=-8	R a=0

Question 8 ♣ Let $S = \{(1,1),(2,2),(3,3)\}$ be a set of points. Let $l = (-1,\ \ 1,\ \ 0)$ be the parameters of a line that are computed based on S using a line fitting algorithm. Compute the error produced when using l to model the line containing the points S and set a as this error. Compute the normal N of this line, then compute b as the sum of the elements of the normalized vector N. Mark **all** the statements that are **correct**.

A a=4	D a=5	G a=3	J a=14	M a=-1	P b=-1
B b=0	E b=2	H b=4	K a=1	N b=1	Q b=5
C a=2	F a=6	I b=14	L b=6	O a=0	R b=3

Question 9 ♣ Let G_1 be a 1D discrete Gaussian filter with $\sigma = 4$ which is represented by a 1×5 array. Let G_2 be a 2D discrete Gaussian filter with $\sigma = 4$ which is represented by a 5×5 array. Let P be a Gaussian image pyramid that is produced by applying G_1 or G_2 to the image I. Mark **all** the statements that are **correct**.

- A It is not possible to correctly use G_1 for producing a Gaussian image pyramid because it is a 1D filter.
- B It is possible to produce the Gaussian pyramid P by convolving the image with a Laplacian filter instead of a Gaussian filter.
- C The Gaussian filter G_2 enhances edges in the image and so is a good way to perform edge detection.
- D The Gaussian pyramid P enables analyzing the image at multiple scales.
- E The size of the array that is used to store G_2 is not sufficient for accurately representing G_2.
- F The size of the array that is used to store G_1 is sufficient for accurately representing G_1.
- G The Gaussian pyramid P consumes less memory for storing an image compared with storing the image in a single array.
- H Applying G_1 to the image twice can be used instead of applying G_2 once, but this is less efficient.
- I The Gaussian filter G_2 is an effective filter for removing impulsive (salt & pepper) noise from the image.

Question 10 ♣ Let M be a 3×4 projection matrix of a camera with a focal length of $f = 2$ with no scale, skew, translation, or rotation. Assuming that the $(3, 3)$ element of M is 1, compute a as the sum of elements in M. Let $P = (1, 2, 4)$ be a point in 3D. Compute the coordinates of the 2D point $p = (x, y)$ that is produced by projecting P using M, then compute b as the sum of elements in p (i.e. $x + y$). Mark **all** the statements that are **correct**.

$$M = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix} \Longrightarrow \begin{array}{l} x = 1/2 \\ y = 1 \end{array}$$

A b=1	D b=4	G a=8	J b=0	M a=1	P a=0
B a=3	E a=6	H b=6	K a=4	N a=5	Q b=1.5
C a=2	F b=0.5	I b=10	L b=2	O a=7	R b=8