Winter 2017 - CSE/ECE 344/544
Computer Vision
Mid Sem - Feb. 21, 2017
(Solution)

Maximum score: 70 | **Time: 75 mins** |

1. (*20 points*) State whether the following statements are **true** or **false** with appropriate justification.

a) Matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, both have rank r. $\mathbf{A} + \mathbf{B}$ will always have rank r.

Answer. False. Take an example, $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ are rnk 2 matrices. However, $A + B$ is a rank 1 matrix.

b) The effect of radial distortion reduces as you radially move away from the principal point.

Answer. False. The effect of radial distortion increases as you radially move away from the principal point. Refer to the figure in the slides.

c) If the image has arbitrarily shaped color regions, mean shift segmentation should be the preferred method over k-means based segmentation.

Answer. True. K-means based segmentation detects spherical segments.

d) In the Harris corner detection approach, an edge is detected if the second moment matrix $\mathbf{H}$ is rank-deficient and has large $||\mathbf{H}||_2$.

Answer. True. Since second moment of matrix $\mathbf{H}$ is rank deficient, which means one of the eigenvalue is zero and a large $||\mathbf{H}||_2$ indicates that the other eigenvale is large. Since we have one eigenvalue large and other eigenvalue zero, it is either vertical or horizontal edge.

e) For line detection using Hough transforms, a point in the image space corresponds to a circle in the Hough parameter space.

Answer. False. If you are using slope and intercept for hough parameter space, a point in the image space corresponds to a line in the hough parameter space.

OR

If you are using radius and angle for hough parameter space, a point in the image space corresponds to a sinusoid in the hough parameter space.

f) Harris corner detection is computationally inefficient because the SSD error has to be explicitly computed over a window around every pixel for small translations.

Answer. False. Harris corner isn't computationally inefficient as we don't have to compute SSD error explicitly for harris corner detection. We can use the image derivative information at that point for corner detection.

g) The Sobel operator is a good choice of filter kernel for performing blob detection.

Answer. False. Sobel can be used for edge detection, but for blob detection Gaussian kerenl with differnt sigma parameter will be a better option.

h) Harris corner detection is invariant to 2D rotations of an image.

Answer. True. The 2D rotations doesn't effect the eigenvalues of the image derivative matrix and just rotates direction of it's eigenvectors.

i) The skew parameter in the intrinsic camera matrix primarily depends on the length to width ratio of the pixel on the sensor.
Answer. False. Skewness parameter depends on skewness angle which can be same for different length to width ratio.

j) In Canny edge detection, non-maxima suppression is performed by retaining the maximum gradient magnitude in a $k \times k$ window around a pixel while setting all other pixels to zero.
Answer. False. In canny edge detection, non-maxima suppression is performed along the gradient direction not a $k \times k$ window around the pixel.

2. (*20 points*) President T. has ordered the construction of a new airport coming up in Idiotville, Oregon. He wants a surveillance camera along with a low-cost single board computer installed along a lane, such that the image plane is parallel to the cars moving on the road. His administration has chosen you to work on this surveillance system at the Idiotville International airport, with the first task being classification of passing scooters vs. buses using their tires. However, the hardware given to you can not support sophisticated feature extraction and classification, i.e., you can not extract SIFT like features or run a deep net or even classifiers like SVMs. Your choices are restricted to some basic image processing and perhaps some additional computation. President T. has said that unless you solve this problem at Idiotville, all Indian student visa applications will be suspended indefinitely. How would you humor President T.?
{*Hint*: You may assume that size of the tires of buses and scooters are known.}
{More about Idiotville, Oregon - `https://en.wikipedia.org/wiki/Idiotville,_Oregon`}
Answer.
• According to the specification provided, hough transform is the best solution. Let us assume that, all the scooters and buses have tire size of r_1 and r_2 respectively.
• Since we know the radius of the tires we will need very limited computation to detect circles using hough transform.
• If the circle with radius r_1 is detected we would classify that the image has a scooter and viceversa.

OR

I am not sure how accurate this method will be
• We can also use the blob detection by assuming that all the scooters and buses have tire size of r_1 and r_2 respectively.
• Since we know the radius of the tires we will compute the difference of gaussians, for differnt σs. For each radius value, we will get differnt σ values for which the tire gives local extrema.
• If the set of σs corresponding to radius r_2 gives better extrema then we would classify that the image has a bus and vicevers.

3. (*30 points*) Consider a vector $(2, 5, 1)^\top$, which is rotated by $\pi/2$ about the Y-axis, followed by a rotation about X-axis by $-\pi/2$ and finally translated by $(-1, 3, 2)^\top$.

a) (*6 points*) What is the coordinate transformation matrix in this case?

b) (*4 points*) Find the new coordinates of this vector. Where does the origin of the initial frame of reference get mapped to?

c) (*10 points*) What is the direction of the axis of the combined rotation in the original frame of reference and what is the angle of rotation about this axis?

d) (*10 points*) Using Rodrigues formula, show that you achieve the same rotation matrix as you get by sequentially applying the two rotations.

Answer.

a)

$$\mathbf{R}_y = \begin{bmatrix} \cos\frac{\pi}{2} & 0 & \sin\frac{\pi}{2} \\ 0 & 1 & 0 \\ -\sin\frac{\pi}{2} & 0 & \cos\frac{\pi}{2} \end{bmatrix}$$

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\frac{-\pi}{2} & -\sin\frac{-\pi}{2} \\ 0 & \sin\frac{-\pi}{2} & \cos\frac{-\pi}{2} \end{bmatrix}$$

$$\mathbf{t} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$$

The transformation to be applied is:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R}_x\mathbf{R}_y & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 3 \\ 0 & -1 & 0 & 2 \end{bmatrix}$$

b) The new coordinates are,

- Origin $[-1, 3, 2]^\top$.
- End point of vector $[0, 1, -3]^\top$.
- Hence the vector will be $[1, -2, -5]^\top$.

c) Here, final rotation matrix is,

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \tag{1}$$

The direction $\mathbf{n}$ and θ angle both can be computed using rodrigues' formula,

$$\theta = \cos^{-1}\left(\frac{trace(\mathbf{R}) - 1}{2}\right) = 120^\circ$$

$$\mathbf{n} = \frac{1}{2\sin\theta}\begin{bmatrix} R_{32} - R_{23} \\ R_{13} - R_{31} \\ R_{21} - R_{12} \end{bmatrix} = \begin{bmatrix} -0.5769 \\ 0.5769 \\ -0.5769 \end{bmatrix}$$

d) If we use the equation $\mathbf{n}$ and θ calculated above using the equation,

$$\mathbf{R} = \mathbf{I} + \sin\theta\ \mathbf{N} + (1 - cos\theta)\ \mathbf{N}^2$$

where,

$$\mathbf{N} = \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}$$

We will get Rotation matrix as,

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

4. (*15 points*) {*Extra Credit*}: The image formation process can be summarized in the equation $\mathbf{x} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X}$, where $\mathbf{K}$ is the intrinsic parameter matrix, $[\mathbf{R}|\mathbf{t}]$ are the extrinsic parameters, $\mathbf{X}$ is the 3D point and $\mathbf{x}$ is the image point respectively, both in homogeneous coordinates. Consider a scenario where there are two cameras (C_1 and C_2) with intrinsic matrices $\mathbf{K}_1$ and $\mathbf{K}_2$ and *corresponding* image points $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively. Assume that the first camera frame of reference is known and is used as the world coordinate frame. The second camera orientation (pose) is obtained by a *pure* 3D rotation $\mathbf{R}$ applied to the first camera's orientation. Show that the homogeneous coordinate representation of image points $\mathbf{x}_1$ and $\mathbf{x}_2$ of C_1 and C_2 respectively, are related by an equation $\mathbf{x}_1 = \mathbf{H}\mathbf{x}_2$, where $\mathbf{H}$ is an invertible 3×3 matrix. Find the matrix $\mathbf{H}$ in terms of $\mathbf{K}_1, \mathbf{K}_2$ and $\mathbf{R}$.

Answer.

We have point $\mathbf{X}_W$ in world coordinate system in homogeneous form, $\mathbf{X}_W = [\widetilde{\mathbf{X}}_W, 1]$ where, $\widetilde{\mathbf{X}}_W = [x, y, z]$. We also hae transformation from $\mathbf{X}_W$ to image given as,

$$\mathbf{x}_1 = \mathbf{K}_1[\mathbf{R}_1|\mathbf{t}_1]_{3\times4}\mathbf{X}_W \tag{1}$$

Since, the camera-1 coordinate system is also world coordinate system, we have extrinsic parameter as Identity,i.e., $\mathbf{R}_1 = \mathbf{I}$ and the translation is $\mathbf{t}_1 = [0, 0, 0]^\top$ in non-homogeneous co-ordinates, which changes the transformation as,

$$\mathbf{x}_1 = \mathbf{K}_1\widetilde{\mathbf{X}}_W \tag{2}$$

Now, for camera-2 coordinate system, we have extrinsic parameter as pure rotation ,i.e., $\mathbf{t}_1 = [0, 0, 0]^\top$ in non-homogeneous coordinates, which makes the transformation for the world point X_W as,

$$\mathbf{x}_2 = \mathbf{K}_2\mathbf{R}_2\widetilde{\mathbf{X}}_W \tag{3}$$

Since, K_1 is an invertible matrix we can write,

$$\widetilde{\mathbf{X}}_W = \mathbf{K}_1^{-1}\mathbf{x}_1 \tag{4}$$

Hence, using equation 3 and equation 4 we can write that,

$$\mathbf{x}_2 = \mathbf{K}_2\mathbf{R}_2\mathbf{K}_1^{-1}\mathbf{x}_1 \tag{5}$$

Since, all matrices $\mathbf{K}_2$, $\mathbf{R}_2$ and $\mathbf{K}1$ are invertible, we have

$$\mathbf{x}_1 = \mathbf{K}_1\mathbf{R}_2^{-1}\mathbf{K}_2^{-1}\mathbf{x}_1 \tag{6}$$

From equation 6 we have,

$$\mathbf{H} = \mathbf{K}_1 \mathbf{R}_2^{-1} \mathbf{K}_2^{-1}$$