Winter 2018 - CSE/ECE 344/544 Computer Vision
Mid Sem - Feb. 18, 2018

Maximum score: 70 **Time: 60 mins**

Instructions:

- Try to attempt all **four** questions. You can score up to 20 extra credit points.
- True/False questions without justification will not be awarded any points. Keep in mind that a statement is true only if it is *always* true.
- Do not copy. Institute plagiarism policy is strictly enforced.
- In the unusual case that a question is not clear, even after discussing with the invigilating TAs, please state your assumptions *clearly* and solve the question. Reasonable assumptions will be accounted for while grading.

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 rank 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) For a pinhole camera, size of the image of a world object decreases as the focal length is increased.
Answer. False. It will increase because of the relationship $x = \frac{fX}{Z}$. Since the image coordinates x and y are both directly proportional to f, they will increase with f, thereby increasing the size of the image of a world object.

d) A similarity transformation in 3D has seven degrees of freedom.
Answer. True. 3 from rotation, 3 from translation and 1 from scaling.

e) Given that the cross-ratio is preserved under the projective transformation, it is preserved under affine transformation as well.
Answer. True. Projective transformations are more general than affine, and all invariants of projective transformations are preserved in its special cases, i.e., affinities, similarities and isometries (Euclidean transformations).

f) The intrinsic matrix $\mathbf{K}$ can be rank-deficient for some cameras.
Answer. False. Intrinsic matrix is an upper triangular matrix, which is always invertible. Hence, $\mathbf{K}$ cannot be rank-deficient.

g) Points at infinity can be used to find the slope of a set of mutually parallel lines.
Answer. True. The points at infinity encode the direct of the line in a plane, therefore the slope can be computed.

h) The line at infinity can be estimated from a set of four mutually parallel lines.
Answer. False. In order to find line at infinity we need two pairs of parallel lines in different directions.

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) Let $\ell'_\infty = \mathbf{H}_P^{-\top}\ell_\infty$, where $\ell_\infty = (0,0,1)^\top$ and $\mathbf{H}_P$ is any projective transformation, specifically with the last row $(\mathbf{v}, b)$ such that $\mathbf{v} \neq 0$. Let $\mathbf{H}_A$, with last row $(0,0,1)$, be an affine transformation applied to ℓ'_∞, i.e., $\ell''_\infty = \mathbf{H}_A^{-\top}\ell'_\infty$. The position of ℓ'_∞ will be preserved under the affine transformation $\mathbf{H}_A$.
Answer. False. Since $\mathbf{H}_P$ is a projective transformation, the line at infinity ℓ_∞ is mapped to a finite line $\ell'_\infty \neq (0,0,1)^\top$. On this finite line ℓ'_∞, $\mathbf{H}_A$ applies an affine transformation, modifying it. Therefore, $\ell'_\infty \neq \ell''_\infty$.

2. (*20 points*) Let $\mathbf{H}$ be a 2×2 transformation matrix of a 1-D projective space $\mathcal{P}^1$ such that $\mathbf{u} = \mathbf{Hv}$, where $\mathbf{u}$ and $\mathbf{v}$ are in $\mathcal{P}^1$ and represented using homogeneous co-ordinates. How many degrees of freedom does $\mathbf{H}$ have? How many point correspondences do you need to solve for $\mathbf{H}$? Can you re-write the equation $\mathbf{u} = \mathbf{Hv}$ in the form $\mathbf{Ah} = \mathbf{0}$, where $\mathbf{h} = \text{vec}(\mathbf{H})$. Write the structure of $\mathbf{A}$ in terms of u_i and v_j.
Answer: Points u and v in $\mathcal{P}^1$ are related as:

$$\mathbf{Hv} = \begin{bmatrix} \mathbf{h}^{1T}\mathbf{v} \\ \mathbf{h}^{2T}\mathbf{v} \end{bmatrix}$$

where $\mathbf{h}^{1T}$ and $\mathbf{h}^{2T}$ are the rows of $\mathbf{H}$.

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}^{1T}\mathbf{v} \\ \mathbf{h}^{2T}\mathbf{v} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\mathbf{u}_1}{\mathbf{u}_2} \\ \frac{\mathbf{u}_2}{\mathbf{u}_2} \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{h}^{1T}\mathbf{v}}{\mathbf{h}^{2T}\mathbf{v}} \\ \frac{\mathbf{h}^{2T}\mathbf{v}}{\mathbf{h}^{2T}\mathbf{v}} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\mathbf{u}_1}{\mathbf{u}_2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{h}^{1T}\mathbf{v}}{\mathbf{h}^{2T}\mathbf{v}} \\ 1 \end{bmatrix}$$

$$\frac{\mathbf{u}_1}{\mathbf{u}_2} = \frac{\mathbf{h}^{1T}\mathbf{v}}{\mathbf{h}^{2T}\mathbf{v}}$$

On cross multiplying we get:

$$\mathbf{u}_1\mathbf{h}^{2T}\mathbf{v} = \mathbf{u}_2\mathbf{h}^{1T}\mathbf{v}$$

This gives:

$$\begin{bmatrix} \mathbf{u}_1\mathbf{v}_1 & \mathbf{u}_1\mathbf{v}_2 & -\mathbf{u}_2\mathbf{v}_1 & -\mathbf{u}_2\mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \end{bmatrix}$$

Since each point correspondence gives us one equation and there are three degrees of freedom in $\mathbf{H}$, we need three equations to solve for H, hence three point correspondences are required.

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

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

b) (*4 points*) Find the new coordinates of this vector $\mathbf{y}$. 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)

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

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

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

The transformation to be applied is:

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

b) The new coordinates are,
- Origin $[-2, -1, 1]^\top$.
- End point of vector $[-3, -3, 6]^\top$.
- 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,

$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{trace(\mathbf{R})-1}{2}\right)\\ &= 120^\circ\\ \mathbf{n} &= \frac{1}{2\sin\theta}\begin{bmatrix} R_{32}-R_{23}\\ R_{13}-R_{31}\\ R_{21}-R_{12}\end{bmatrix}\\ &= \begin{bmatrix}0.5774\\ -0.5774\\ -0.5774\end{bmatrix}\end{aligned}$$

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. (*20 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. Let $\mathbf{X}$ be an arbitrary 3D point in the scene and $\mathbf{x}_1$ and $\mathbf{x}_2$ be its corresponding projections on to two image planes of two corresponding cameras ($\{\mathbf{K}_1, \mathbf{R}_1, \mathbf{t}_1\}$ and $\{\mathbf{K}_2, \mathbf{R}_2, \mathbf{t}_2\}$). If we assume that the two camera centers coincide (but the coordinate axes may not be aligned), then there exists a homography $\mathbf{H}$, such that $\mathbf{x}_1 = \mathbf{H}\mathbf{x}_2$, where $\mathbf{H}$ is an invertible 3×3 matrix. Find the matrix $\mathbf{H}$ in terms of $\mathbf{K}_1, \mathbf{K}_2, \mathbf{R}_1$ and $\mathbf{R}_2$. Can such a homography exist if the camera centers do not coincide? Explain why or why not. {*Hint*: You can assume that one of the cameras is the world reference frame.}

Answer.

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

$$\mathbf{x}_1 = \mathbf{K}_1[\mathbf{R}_1|\mathbf{t}_1]_{3\times 4}\mathbf{X} \tag{1}$$

Assuming, 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,

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

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 as,

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

Since, K_1 is an invertible matrix we can write,

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

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,

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

- Yes, homography will exist, in that case we will also consider the translation between the camera centers in equation number 3.