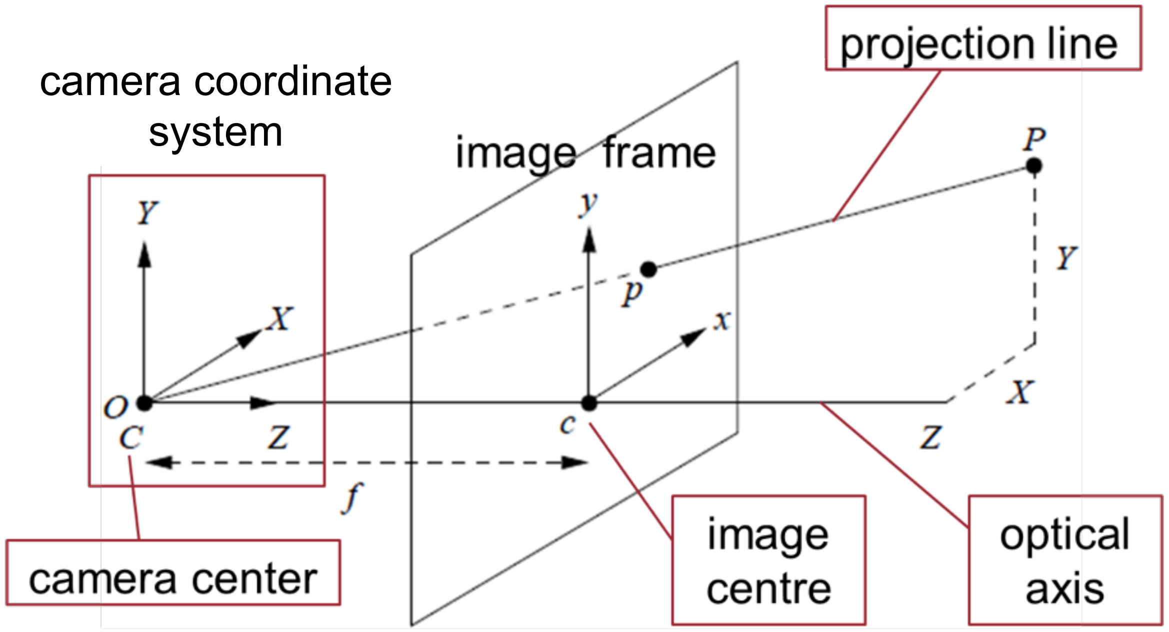
- **Review the lectures**
- **Assessment**

Camera models

- Pinhole camera model



$$x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}$$

Camera models

- Perspective projection model
  - Camera sensor's pixels not exactly square

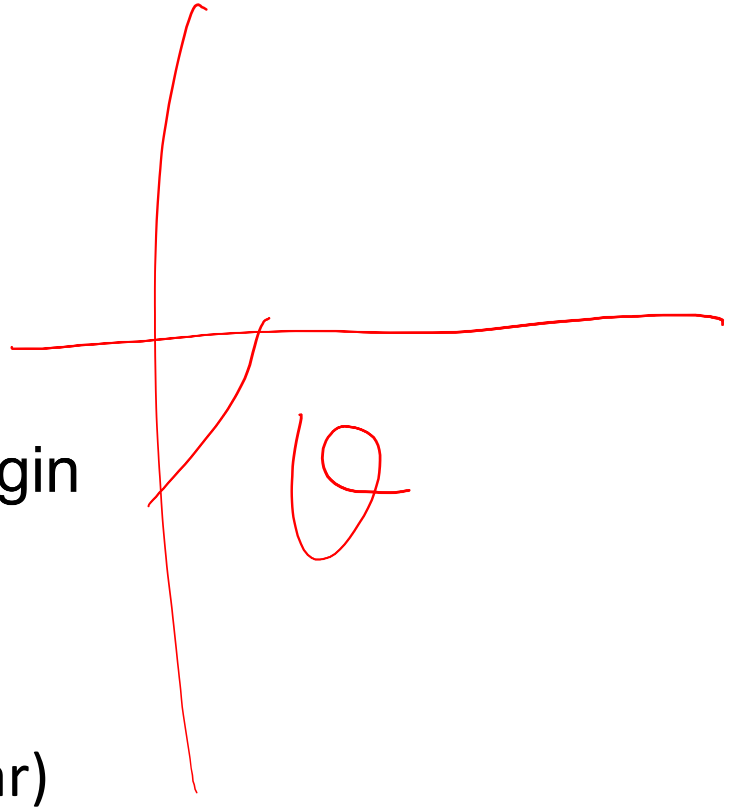
$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

  - Image center or principal point c may not be at origin

$$x = f_x\frac{X}{Z} + c_x, \quad y = f_y\frac{Y}{Z} + c_y$$

  - Distortion (image frame may not be exactly rectangular)

$$x = f_x\frac{X}{Z} - f_x \cot\theta\frac{Y}{Z} + c_x, \quad y = \frac{f_y}{\sin\theta}\frac{Y}{Z} + c_y$$



Camera models

- Perspective projection model
  - Intrinsic parameters

$$\tilde{\mathbf{x}} = \frac{1}{Z}\mathbf{K}\mathbf{X}, \quad \mathbf{K} = \begin{bmatrix} f_x & -f_x \cot\theta & c_x \\ 0 & \dfrac{f_y}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \quad \mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Intrinsic parameter matrix

Camera models

- Perspective projection model
  - Extrinsic parameters
    - Camera coordinate system is not aligned with world coordinate system
    - Camera can move and rotate

$$ {}^{C}\mathbf{X} = {}^{C}_{W}\mathbf{R}\,{}^{W}\mathbf{X} + {}^{C}_{W}\mathbf{T} $$

Camera models

- Perspective projection model

$$\mathbf{P}' = \mathcal{M}\mathbf{P}_w$$
$$= \mathcal{K} \begin{bmatrix} \mathcal{R} & \mathbf{T} \end{bmatrix} \mathbf{P}_w$$

Internal (intrinsic) parameters

External (extrinsic) parameters

Camera Calibration

- Recovering K
- Recovering R and T

$$\mathbf{P}' = \mathcal{M}\mathbf{P}_w$$

$$= \mathcal{K}\begin{bmatrix}\mathcal{R} & \mathbf{T}\end{bmatrix}\mathbf{P}_w$$

Internal (intrinsic) parameters

External (extrinsic) parameters

Camera Calibration

- How many parameters to recover?
  - 5 intrinsic parameters
    - 2 for focal length
    - 2 for offset
    - 1 for skewness
  - 6 extrinsic parameters
    - 3 for rotation
    - 3 for translation

$$\mathbf{P}' = \mathcal{M}\mathbf{P}_w$$
$$= \mathcal{K}\begin{bmatrix}\mathcal{R} & \mathbf{T}\end{bmatrix}\mathbf{P}_w$$

$$K = \begin{bmatrix} \alpha & -\alpha\cot\theta & u_o \\ 0 & \dfrac{\beta}{\sin\theta} & v_o \\ 0 & 0 & 1 \end{bmatrix} \quad R = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{bmatrix} \quad T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

Camera Calibration

- 11 parameters to recover
- Corresponding 3D-2D point pairs
  - Each 3D-2D point pair -> 2 constraints
  - 11 unknown -> 6 point correspondence
  - Use more to handle noisy data

$$\mathbf{p}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = M\mathbf{P}_i = \begin{bmatrix} \frac{\mathbf{m}_1\mathbf{P}_i}{\mathbf{m}_3\mathbf{P}_i} \\ \frac{\mathbf{m}_2\mathbf{P}_i}{\mathbf{m}_3\mathbf{P}_i} \end{bmatrix}$$

$$u_i(\mathbf{m}_3\mathbf{P}_i) - \mathbf{m}_1\mathbf{P}_i = 0$$
$$v_i(\mathbf{m}_3\mathbf{P}_i) - \mathbf{m}_2\mathbf{P}_i = 0$$

Camera Calibration

- Solved using SVD

$$P\mathbf{m} = 0$$

SVD decomposition of P

$$U_{2n\times 12}\, D_{12\times 12}\, V^{T}_{12\times 12}$$

Epipolar Geometry

- Essential matrix
  - Canonical camera

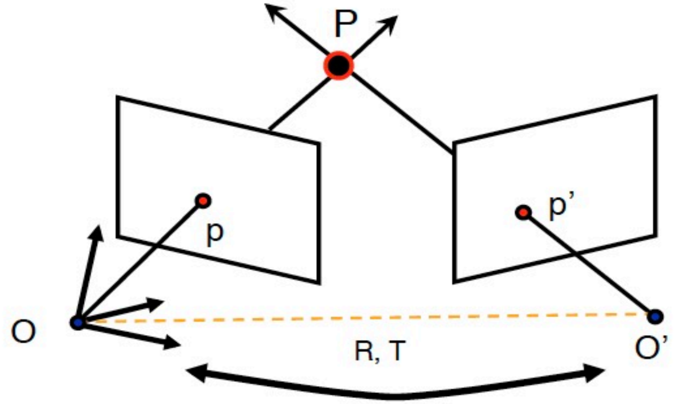
$$p'^{T} E\, p = 0,\ E = [T_{X}]R$$

$$K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Fundamental matrix

$$p'^{T} F\, p = 0,\ F = K'^{-T}[T_{\times}]RK^{-1}$$

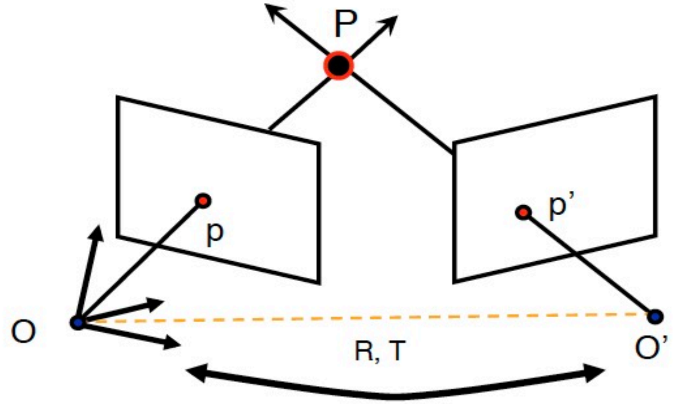
  - Relate matching points of different views
    - No need 3D point location
    - No need intrinsic parameters
    - No need extrinsic parameters



Epipolar Geometry

- Fundamental matrix
  - 3 by 3
  - homogeneous matrix
  - 7 degrees of freedom
    - 9 elements
    - scale doesn’t matter (homogeneous matrix)
    - determinant(F) = 0

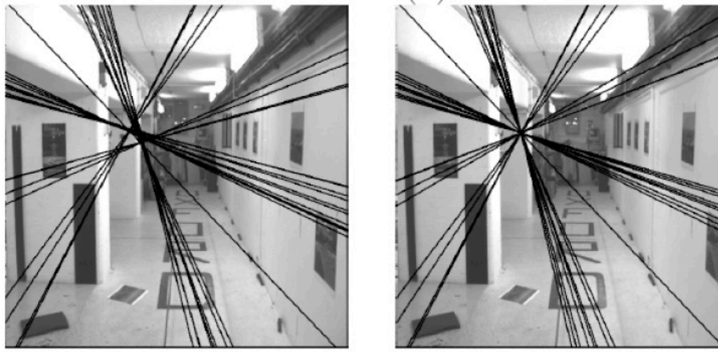
$$p'^T F\, p = 0, \qquad F = K'^{-T}[T_{\times}]RK^{-1}$$



Epipolar Geometry

- Fundamental matrix
  - 3 by 3
  - homogeneous matrix
  - 7 degrees of freedom
    - 9 elements
    - scale doesn't matter (homogeneous matrix))
    - determinant(F) = 0
      - rank(F) = 2

Fundamental matrix has rank 2 : $\det(\mathtt{F}) = 0$.



Left : Uncorrected $\mathtt{F}$ – epipolar lines are not coincident.

Right : Epipolar lines from corrected $\mathtt{F}$.

Recover F from Matched Image Points

- 8-point algorithm (>= 8 point pairs)
- Normalized 8-point algorithm
  - Idea: normalize image points before constructing the equations
    - Translation: the centroid of the image points is at origin
    - Scaling: average distance of points from origin is $\sqrt{2}$

$$q_i = T p_i \qquad q'_i = T' p'_i$$

Recover F from Matched Image Points

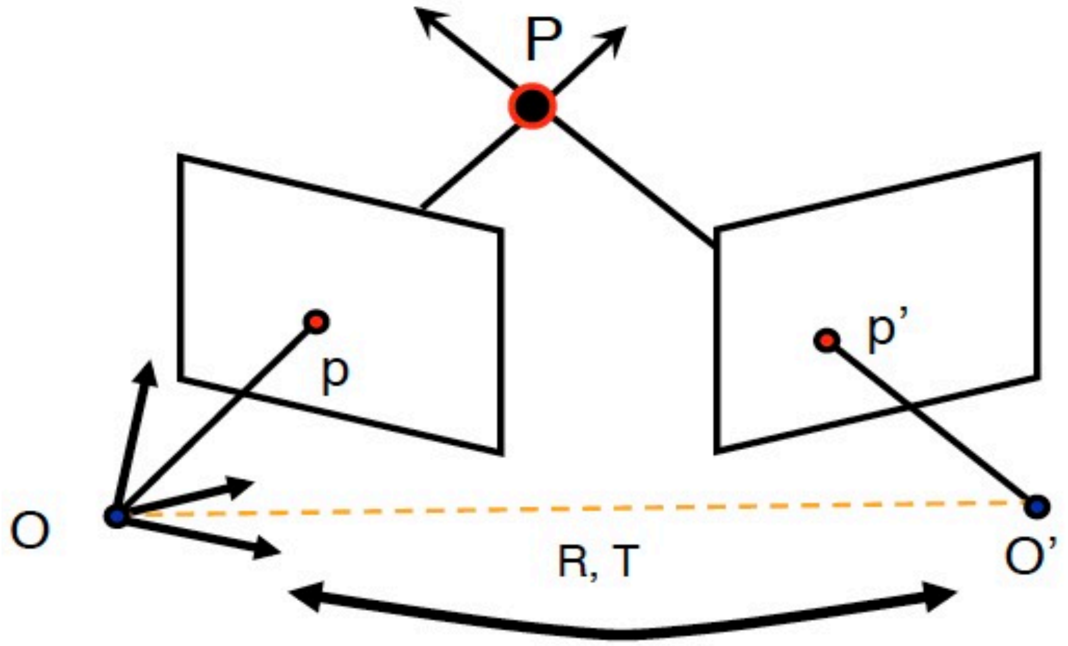
- 8-point algorithm (>= 8 point pairs)
- Normalized 8-point algorithm
  - Idea: normalize image points before constructing the equations
  - Construct linear system using the normalized points
  - Solve using SVD
  - Constraint enforcement
  - De-normalization

Relative Pose from Fundamental Matrix

- Essential matrix from fundamental matrix
  - Known intrinsic parameters
    - Estimation
    - Calibration

$$F = K'^{-T}[t_\times]RK^{-1}$$

$$E = [t_\times]R = K'^{T}FK$$



Relative Pose from Fundamental Matrix

- Essential matrix from fundamental matrix
- Relative pose from essential matrix
  - SVD of E
  - determinant(R) > 0
    - Two potential values
  - T up to a sign
    - Two potential values

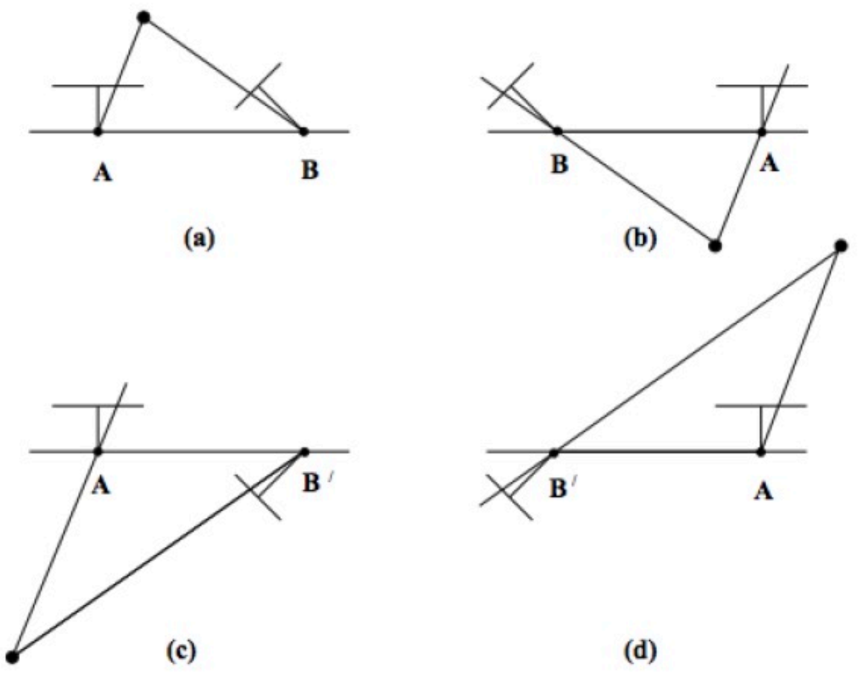
$$E = U\Sigma V^T$$

$$R = (\det UWV^T)UWV^T \text{ or } (\det UW^TV^T)UW^TV^T$$

$$t = \pm U \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \pm u_3$$

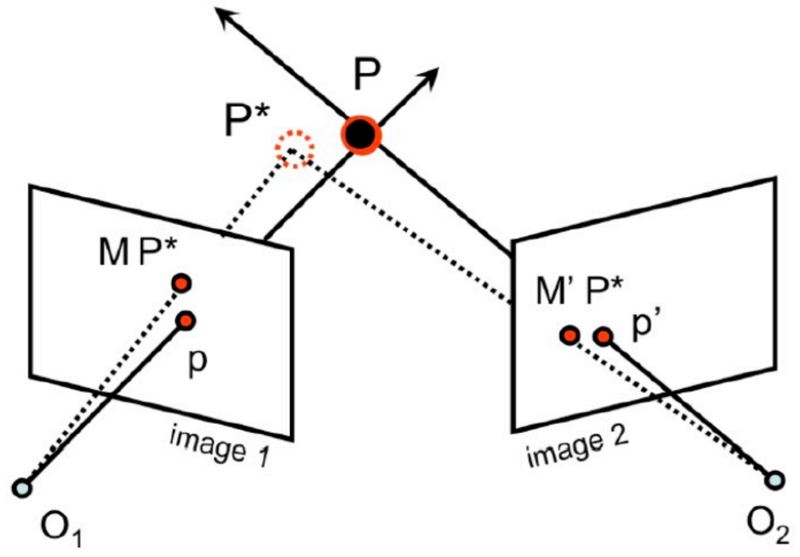
Relative Pose from Fundamental Matrix

- Essential matrix from fundamental matrix
- Relative pose from essential matrix
  - R: two potential values
  - T: two potential values
  - 3D points must be in front of both cameras
    - Reconstruct 3D points
      - using all potential pairs of R and t
    - Count the number of points in front of cameras
    - The pair giving max front points is correct



Triangulation

- Find coordinates of 3D point from its projection into two views
  - Known camera intrinsic parameters (K)
  - Known relative orientation (R) and offset (T)
  - In theory, P is ∩ of the two lines of sight
    - Straightforward and mathematically sound
    - Do not work well
      - Noisy in observation
      - K, R, T are not precise
  - A linear method for triangulation
  - A non- linear method for triangulation



A Linear Method for Triangulation

Two image points

$$p = MP = (x, y, 1)$$

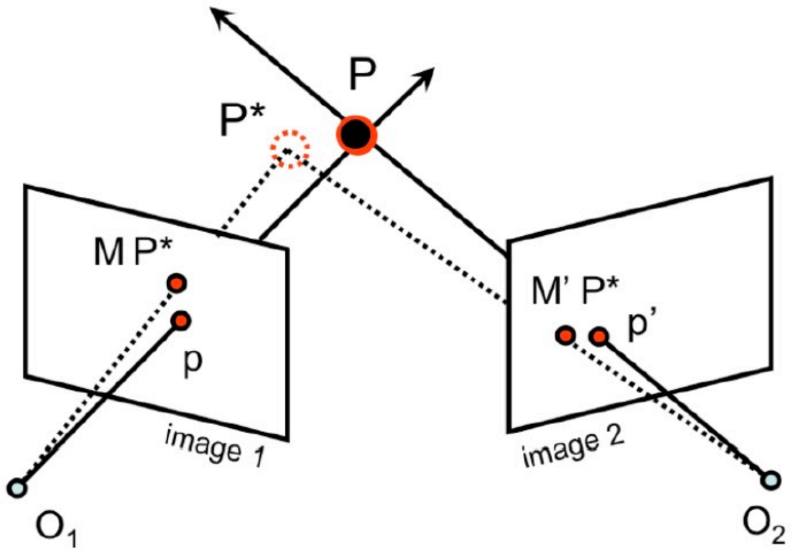
$$p' = M'P = (x', y', 1)$$

By the definition of the cross product

$$p \times (MP) = 0$$

Similar constraints can also be formulated for p' and M'.

$$x(M_3P) - (M_1P) = 0$$
$$y(M_3P) - (M_2P) = 0$$
$$x(M_2P) - y(M_1P) = 0$$



A Linear Method for Triangulation

Two image points

$$p = MP = (x, y, 1)$$

$$p' = M'P = (x', y', 1)$$

By the definition of the cross product

$$p \times (MP) = 0$$

Similar constraints can also be formulated for p' and M'.

$$x(M_3P) - (M_1P) = 0$$
$$y(M_3P) - (M_2P) = 0$$
$$x(M_2P) - y(M_1P) = 0$$

$$A = \begin{bmatrix} xM_3 - M_1 \\ yM_3 - M_2 \\ x'M_3' - M_1' \\ y'M_3' - M_2' \end{bmatrix}$$

$$AP = 0$$

A Linear Method for Triangulation

- Advantages
  - Easy to solve and very efficient
  - Any number of corresponding image points
  - Can handle multiple views
  - Used as initialization to advanced methods

$$AP = 0$$

$$A = \begin{bmatrix} xM_3 - M_1 \\ yM_3 - M_2 \\ x'M_3' - M_1' \\ y'M_3' - M_2' \end{bmatrix}$$

Structure from Motion

- Structure
  - 3D geometry of the scene/object
- Motion
  - Camera locations and orientations
- Structure from Motion
  - Simultaneously recovering structure and motion

Bundle Adjustment

- Minimize sum of squared re-projection errors:

$$g(\mathbf{X}, \mathbf{R}, \mathbf{T}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \underbrace{w_{ij}}_{\text{indicator variable}} \cdot \left\| \underbrace{\mathbf{P}(\mathbf{x}_i, \mathbf{R}_j, \mathbf{t}_j)}_{\text{predicted image points}} - \underbrace{\begin{bmatrix} u_{i,j} \\ v_{i,j} \end{bmatrix}}_{\text{observed image points}} \right\|^2$$

predicted image points

observed image points

indicator variable: is point *i* visible in image *j* ?

Advanced topics

- Lecture 7
  - Multi-view stereo
  - Deep learning for MVS
- Lecture 8
  - Machine learning basics and semantic segmentation

- **Review the lectures**
- **Assessment** 

Assignments vs. final exam

- Assignments – 40%
- Final exam – 60%
  - 09:00 – 11:30, April 12 (Monday)
  - Open book
  - Three types of questions
    - Multiple choice questions with a single correct answer
    - Multiple choice questions with >=2 correct answers
    - Open questions

Example questions

- Multiple choice questions with a single correct answer

Which of the following statement regarding camera calibration is correct?

(a) Camera calibration requires at least 6 3D-2D point correspondences.

(b) Camera calibration requires at least 8 3D-2D point correspondences.

(c) Camera calibration requires at least 8 pairs of corresponding image points.

(d) Camera calibration can be solved using SVD decomposition of the projection matrix.

Example questions

- Multiple choice questions with a single correct answer
- Multiple choice questions with >=2 correct answers

Which of the following operations will change the camera intrinsic parameters?

(a) When zooming in.

(b) When rotating the camera around its local origin.

(c) When changing the resolution of the image.

(d) When the camera is moved.

Example questions

- Multiple choice questions with a single correct answer
- Multiple choice questions with >=2 correct answers
- Open questions

For camera calibration, how do you determine the sign of the intermediate parameter ρ. How does an incorrect sign of ρ affect the subsequent calculation of the intrinsic and extrinsic camera parameters?

- Q&A session
  - 13:45 – 15:30, April 6 (Tuesday)