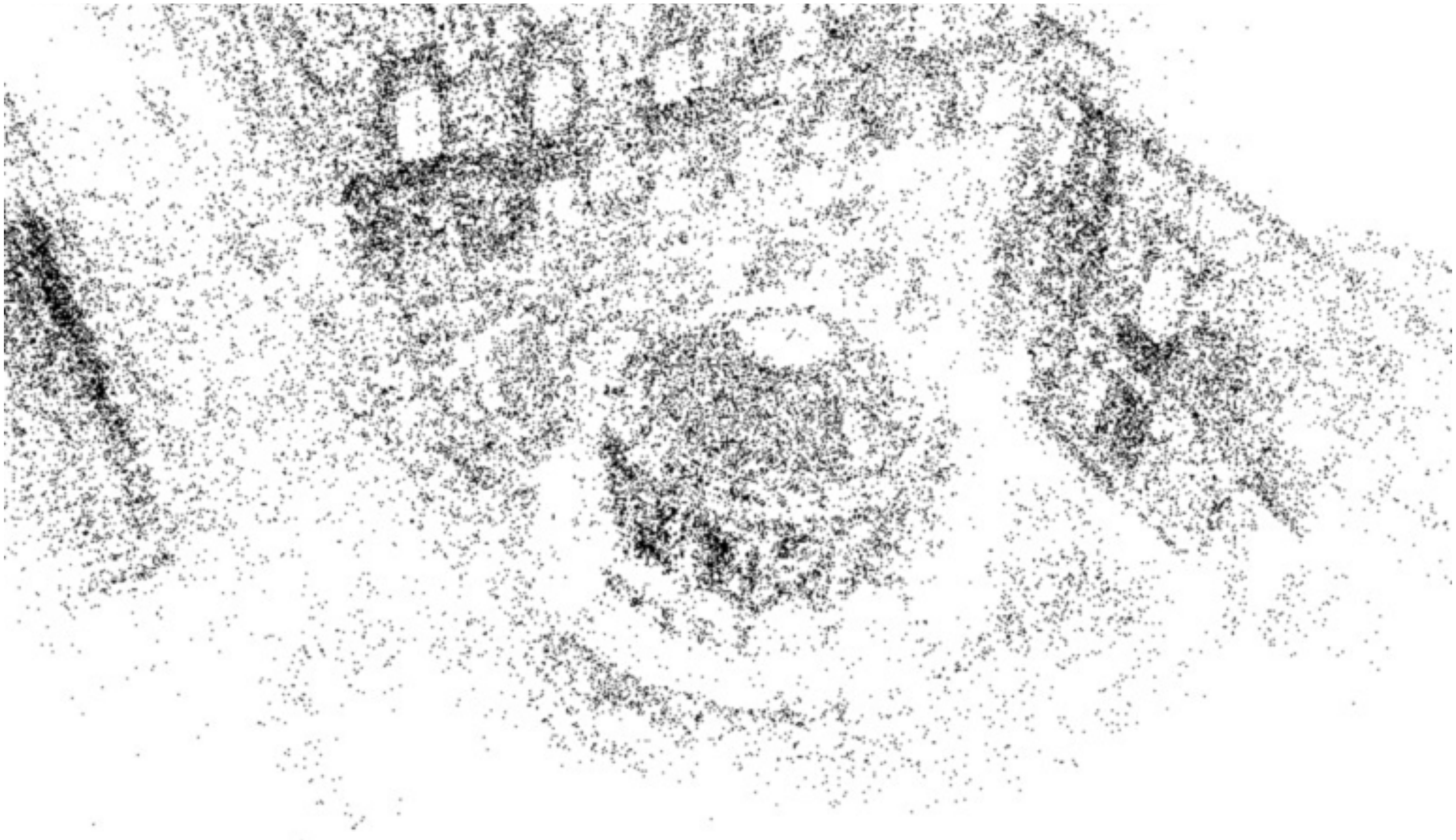


Reconstruction

16-385 Computer Vision (Kris Kitani)

Carnegie Mellon University

	Structure (scene geometry)	Motion (camera geometry)	Measurements
Pose Estimation	known	estimate	3D to 2D correspondences
Triangulation	estimate	known	2D to 2D coorespondences
Reconstruction	estimate	estimate	2D to 2D coorespondences

Reconstruction

(2 view structure from motion)

Given a set of matched points

$$\{\boldsymbol{x}_i, \boldsymbol{x}'_i\}$$

Estimate the camera matrices

$$\mathbf{P}, \mathbf{P}'$$

Estimate the 3D point

$$\mathbf{X}$$

Reconstruction

(2 view structure from motion)

Given a set of matched points

$$\{\boldsymbol{x}_i, \boldsymbol{x}'_i\}$$

Estimate the camera matrices

$$\mathbf{P}, \mathbf{P}'$$

'motion'

(of the cameras)

Estimate the 3D point

$$\mathbf{X}$$

'structure'

Procedure for Reconstruction

1. Compute the Fundamental Matrix **F** from points correspondences
   8-point algorithm

$$\boldsymbol{x}'^{\top}_m \mathbf{F} \boldsymbol{x}_m = 0$$

Procedure for Reconstruction

1. Compute the Fundamental Matrix **F** from points correspondences

   8-point algorithm

2. Compute the camera matrices **P** from the Fundamental matrix

   $\mathbf{P} = [\ \mathbf{I} \mid \mathbf{0}\ ]$ and $\mathbf{P}' = [\ [\mathbf{e}'_{x}]\mathbf{F} \mid \mathbf{e}'\ ]$

Camera matrices corresponding to the fundamental matrix **F** may be chosen as

$$\mathbf{P} = [\mathbf{I}|\mathbf{0}] \quad \mathbf{P}' = [[\boldsymbol{e}_{\times}]\mathbf{F}|\boldsymbol{e}']$$

(See Hartley and Zisserman C.9 for proof)

Decomposing **F** into **R** and **T**

If we have calibrated cameras we have $\mathbf{K}$ and $\mathbf{K}'$

Essential matrix: $\mathbf{E} = \mathbf{K}'^{\top}\mathbf{F}\mathbf{K}$

Decomposing **F** into **R** and **T**

If we have calibrated cameras we have $\mathbf{K}$ and $\mathbf{K}'$

Essential matrix: $\mathbf{E} = \mathbf{K}'^{\top}\mathbf{F}\mathbf{K}$

SVD: $\mathbf{E} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top}$ Let $\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Decomposing **F** into **R** and **T**

If we have calibrated cameras we have $\mathbf{K}$ and $\mathbf{K}'$

Essential matrix: $\mathbf{E} = \mathbf{K}'^{\top}\mathbf{F}\mathbf{K}$

SVD: $\mathbf{E} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top}$ Let $\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

We get FOUR solutions:

$$\mathbf{E} = [\mathbf{R}|\mathbf{T}]$$

$\mathbf{R}_1 = \mathbf{U}\mathbf{W}\mathbf{V}^{\top}$ $\mathbf{R}_2 = \mathbf{U}\mathbf{W}^{\top}\mathbf{V}^{\top}$ $\mathbf{T}_1 = U_3$ $\mathbf{T}_2 = -U_3$

two possible rotations

two possible translations

We get FOUR solutions:

$$\mathbf{R}_1 = \mathbf{U}\mathbf{W}\mathbf{V}^\top$$

$$\mathbf{T}_1 = U_3$$

$$\mathbf{R}_1 = \mathbf{U}\mathbf{W}\mathbf{V}^\top$$

$$\mathbf{T}_2 = -U_3$$

$$\mathbf{R}_2 = \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top$$

$$\mathbf{T}_2 = -U_3$$

$$\mathbf{R}_2 = \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top$$

$$\mathbf{T}_1 = U_3$$

Which one do we choose?

Compute determinant of R, valid solution must be equal to 1

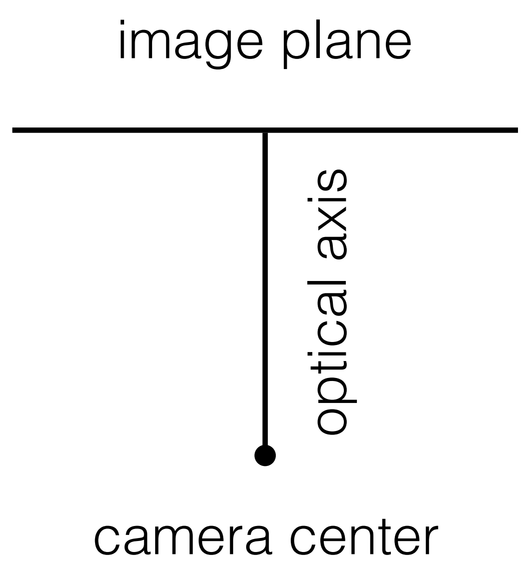
(note: det(R) = -1 means rotation and reflection)

Compute 3D point using triangulation, valid solution has positive Z value

(Note: negative Z means point is behind the camera)

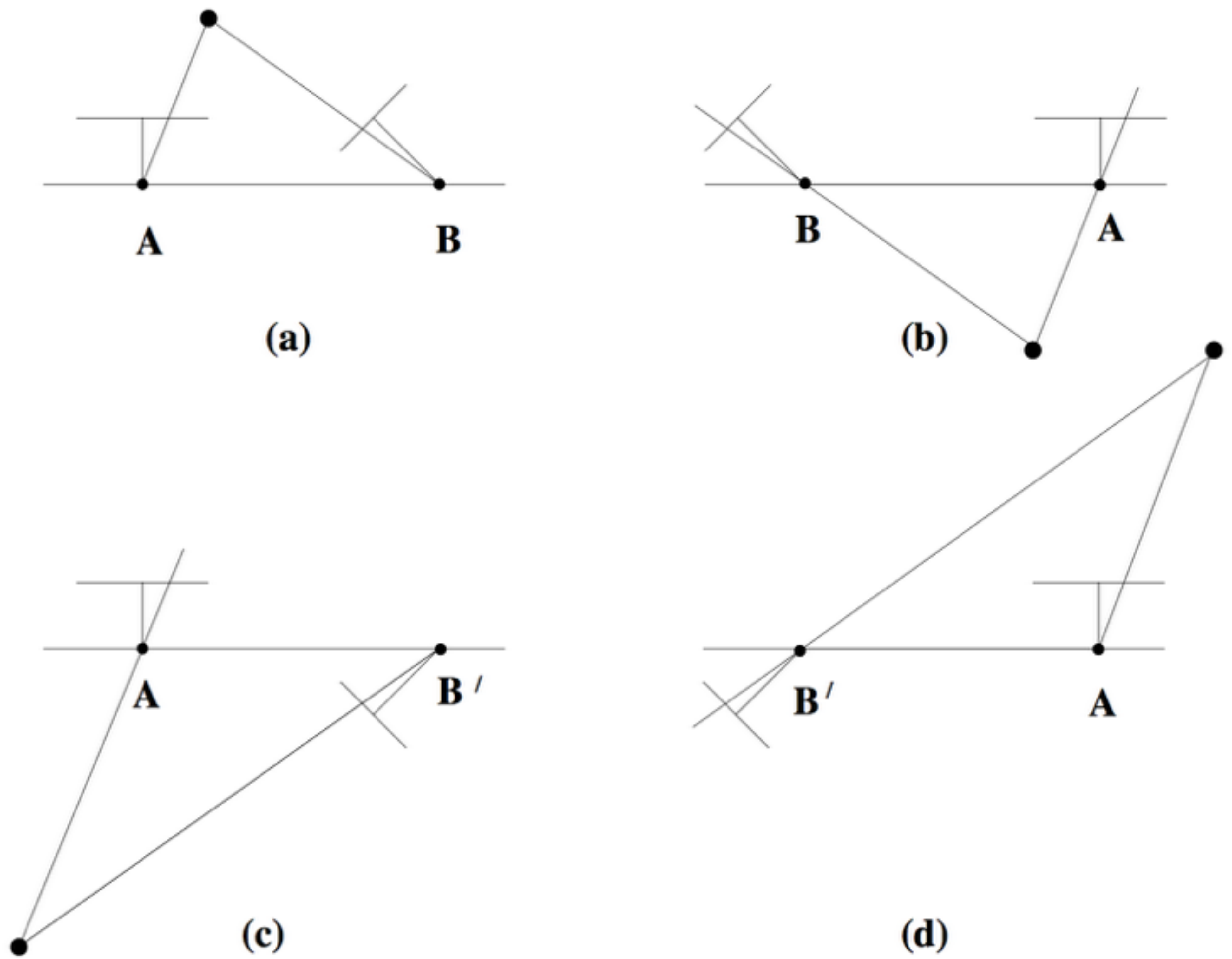
Let’s visualize the four configurations…

Camera Icon

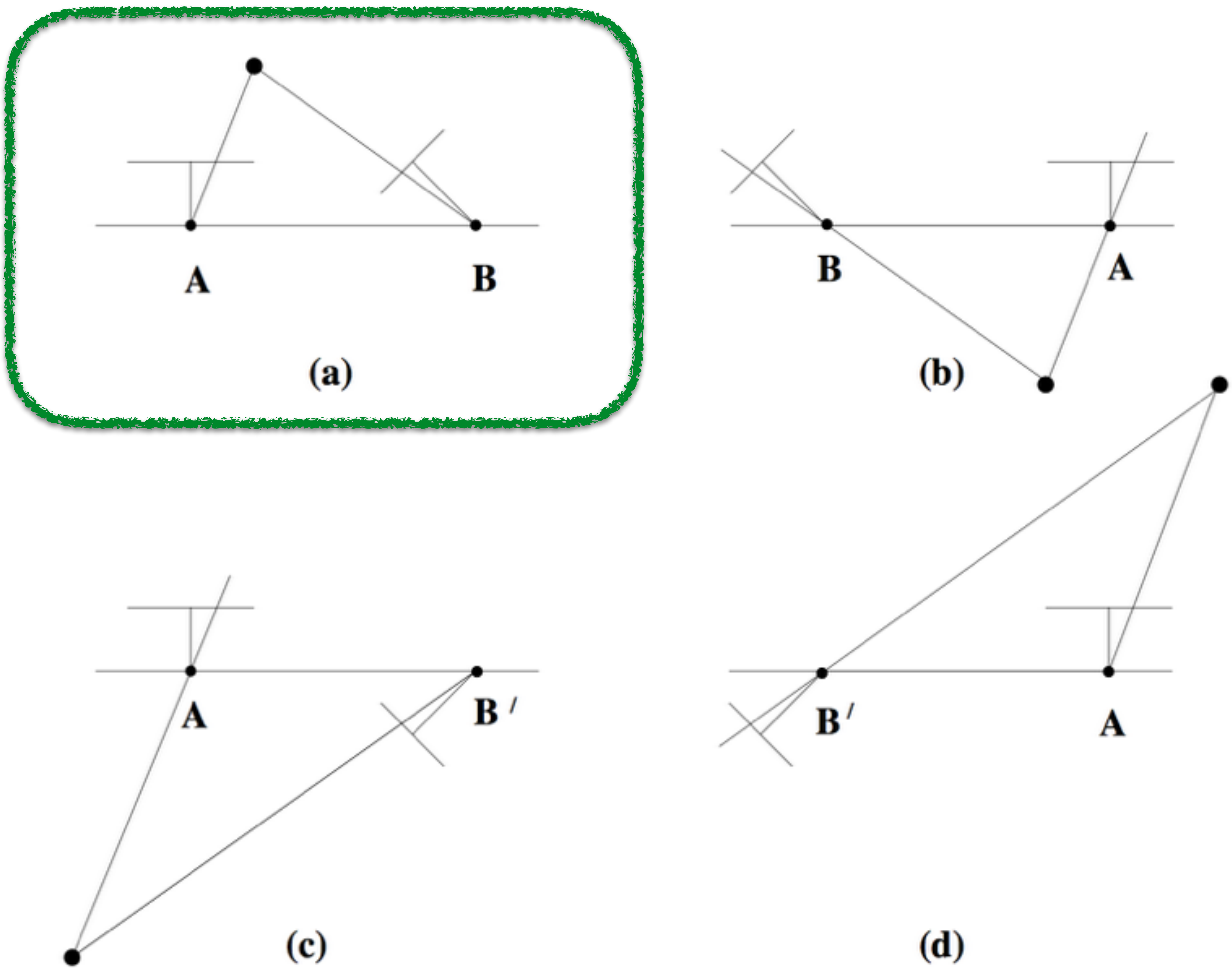


Find the configuration where the points is in front of both cameras

Find the configuration where the points is in front of both cameras



Find the configuration where the points is in front of both cameras



From points correspondences to camera displacement

1. Normalize the image points **x**,**x′** using **K**,**K′**
2. Use the 8-point algorithm to find an approximation of **E** (SVD!)
3. Project **E** to essential space (SVD!!)
   (set smallest SV to zero)
4. Recover possible solutions for **R** and **T** (SVD!!!)
5. Use point correspondence to find the correct **R**,**T** pair (don't use SVD…)

Procedure for Reconstruction

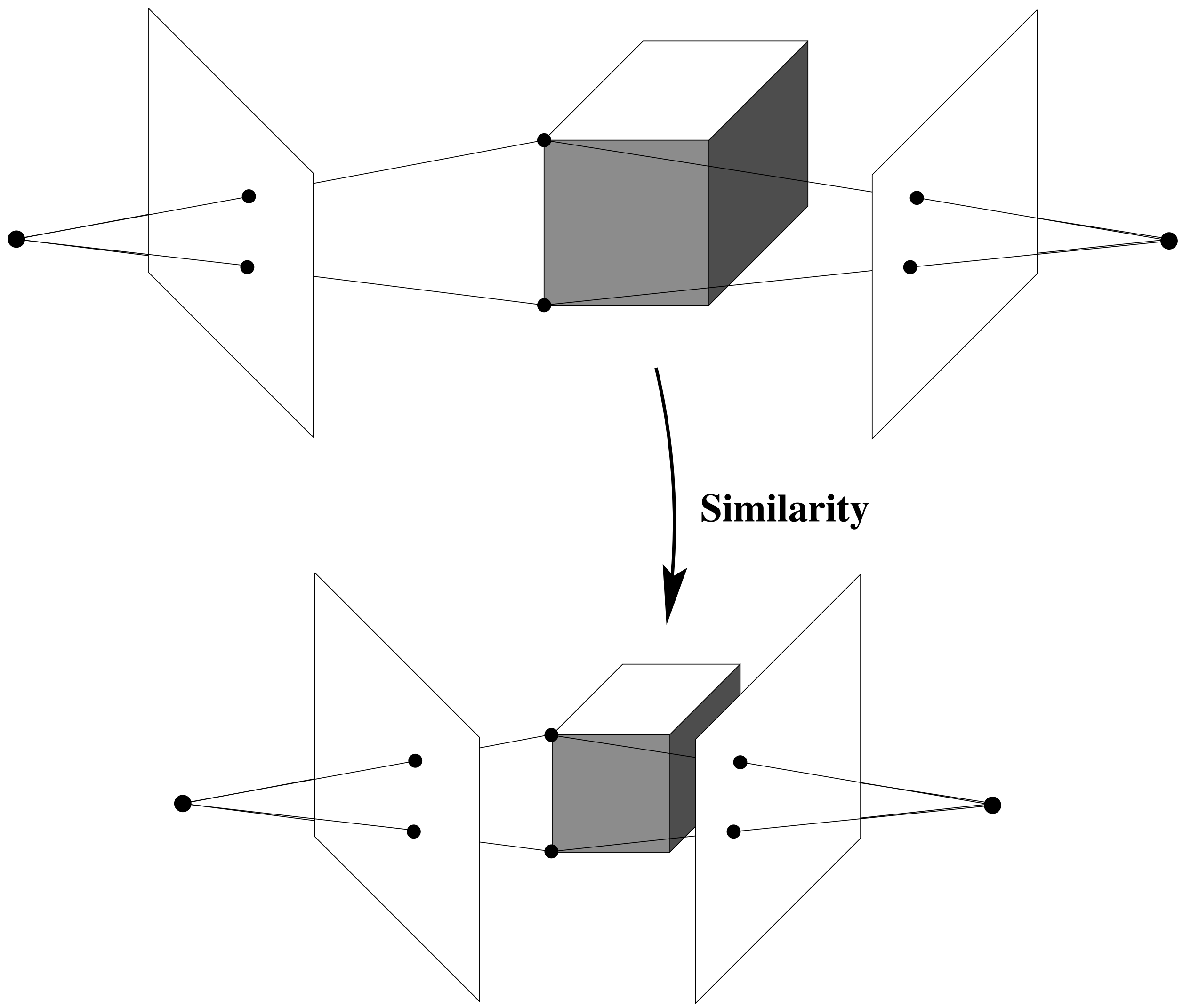
1. Compute the Fundamental Matrix **F** from points correspondences
   8-point algorithm

2. Compute the camera matrices **P** from the Fundamental matrix
   $\mathbf{P} = [\ \mathbf{I} \mid \mathbf{0}\ ]$ and $\mathbf{P}' = [\ [\mathbf{e}'_{\times}]\mathbf{F} \mid \mathbf{e}'\ ]$

3. For each point correspondence, compute the point **X** in 3D space (triangularization)
   DLT with $x = \mathbf{P}\,X$ and $x' = \mathbf{P}'\,X$

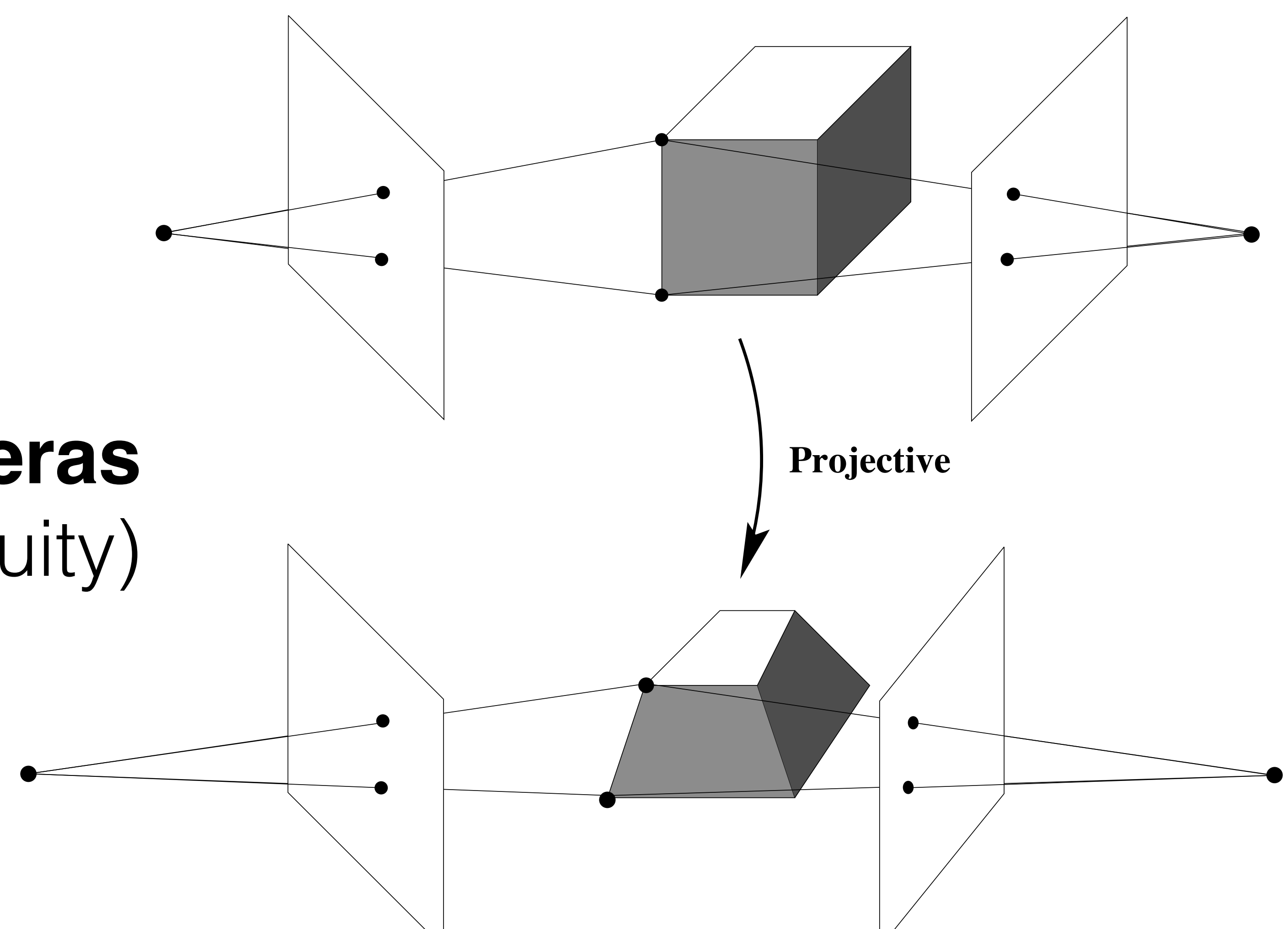
Projective Ambiguity

- Reconstruction is ambiguous by an arbitrary 3D projective transformation without prior knowledge of camera parameters



Calibrated cameras
(similarity projection ambiguity)

Uncalibrated cameras
(projective projection ambiguity)



	Structure (scene geometry)	Motion (camera geometry)	Measurements
Pose Estimation	known	estimate	3D to 2D correspondences
Triangulation	estimate	known	2D to 2D coorespondences
Reconstruction	estimate	estimate	2D to 2D coorespondences