Computer Vision

3D-Reconstruction from Image Pairs

Technische Hochschule Rosenheim
Winter 2024/25
Prof. Dr. Jochen Schmidt

Acknowledgements

Many of the slides presented here are based on the Computer Vision Courses of
Svetlana Lazebnik, University of Illinois at Urbana-Champaign https://slazebni.cs.illinois.edu/fall22
David Fouhey, University of Michigan https://web.eecs.umich.edu/~fouhey/teaching/EECS442_W23

Application

Building a 3D Reconstruction Out Of Images

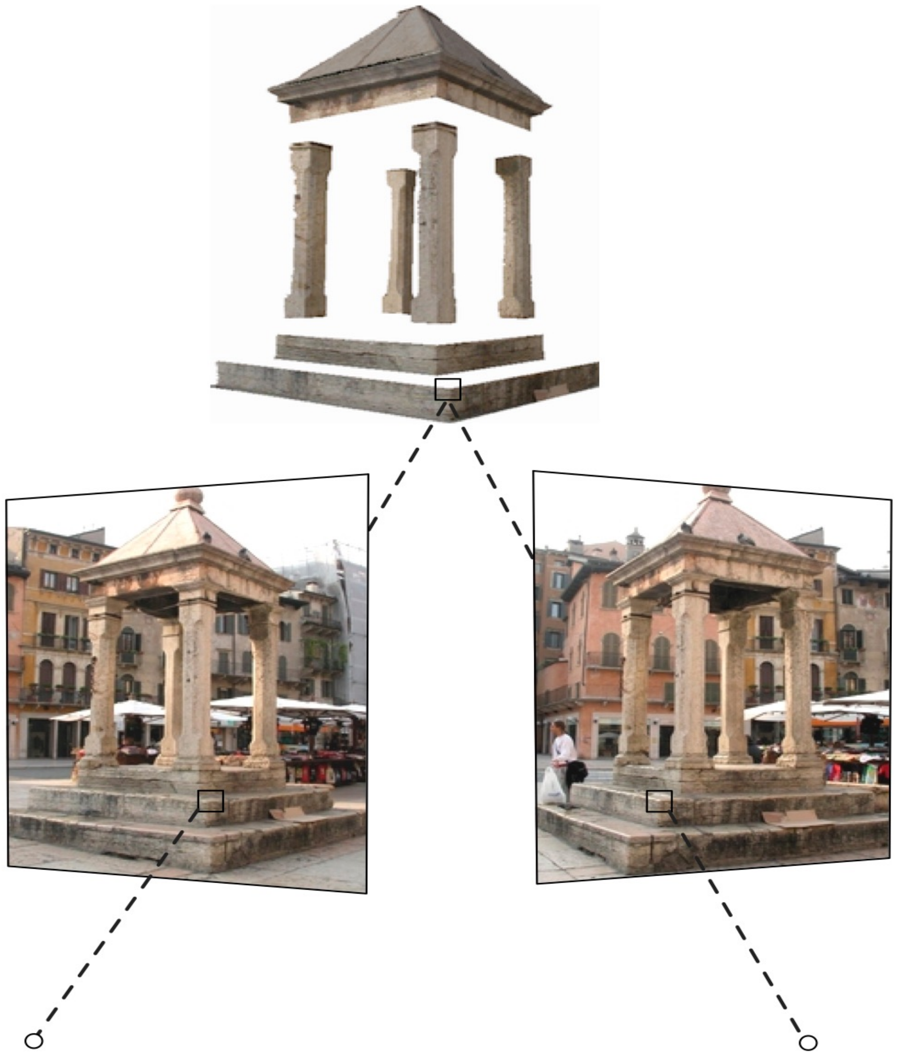
N. Snavely, S. Seitz, and R. Szeliski, [Photo tourism: Exploring photo collections in 3D](http://phototour.cs.washington.edu/), SIGGRAPH 2006.
http://phototour.cs.washington.edu/

Outline

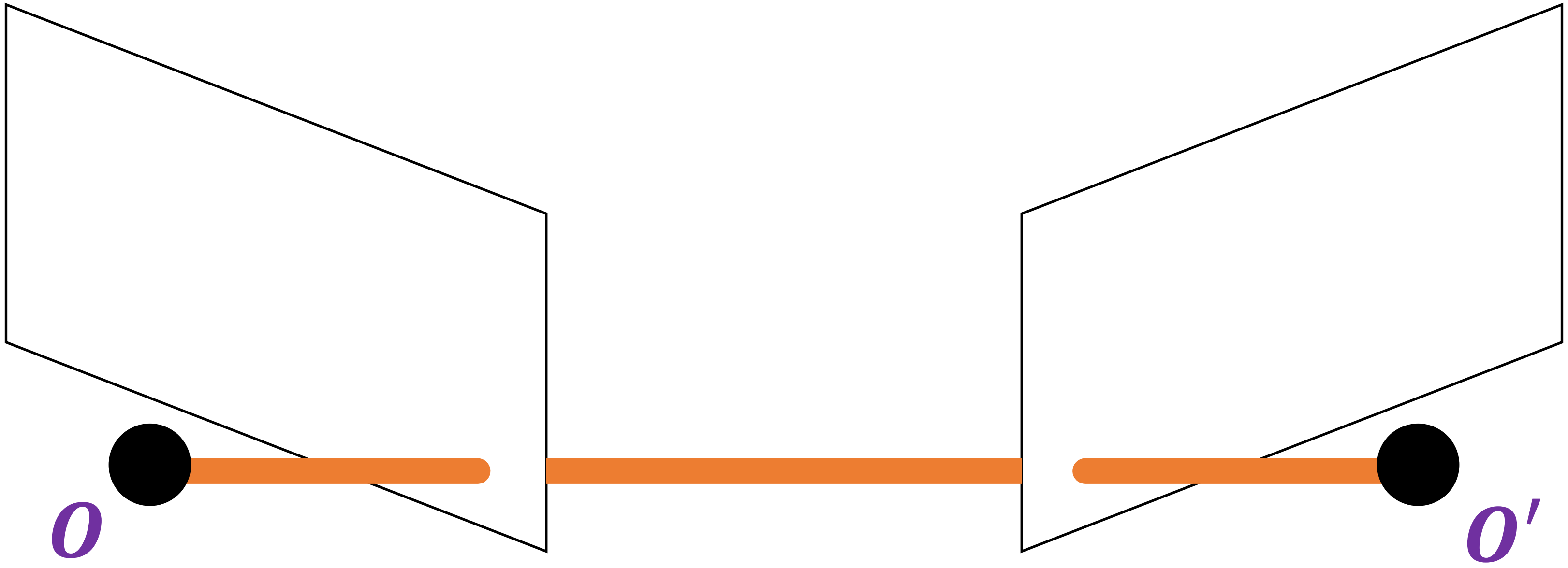
- Epipolar geometry setup
- Epipolar constraint
- Essential matrix
- Fundamental matrix
- Estimating the fundamental matrix

Consider two views of the same 3D scene

- What constraints must hold between two projections of the same 3D point?
- Given a 2D point in one view, where can we find the corresponding point in the other view?
- Given only 2D correspondences, how can we calibrate the two cameras, i.e., estimate their relative position and orientation and the intrinsic parameters?

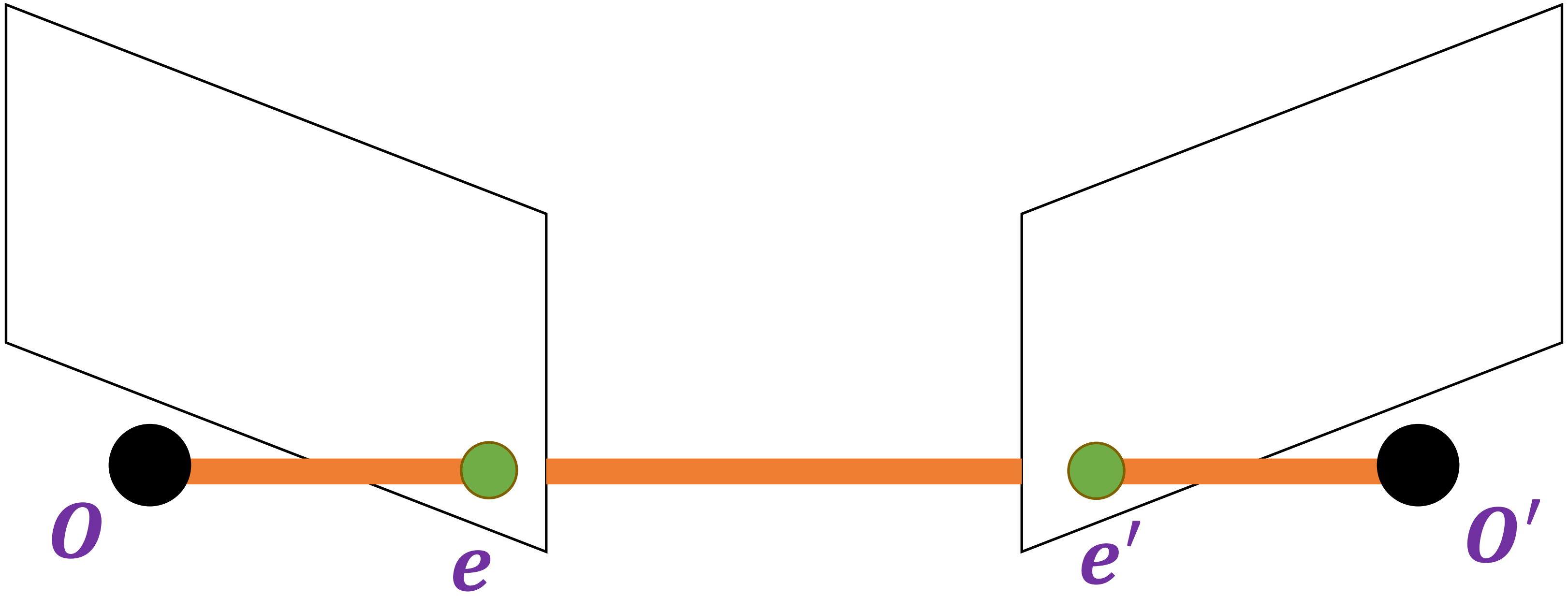


Epipolar geometry setup



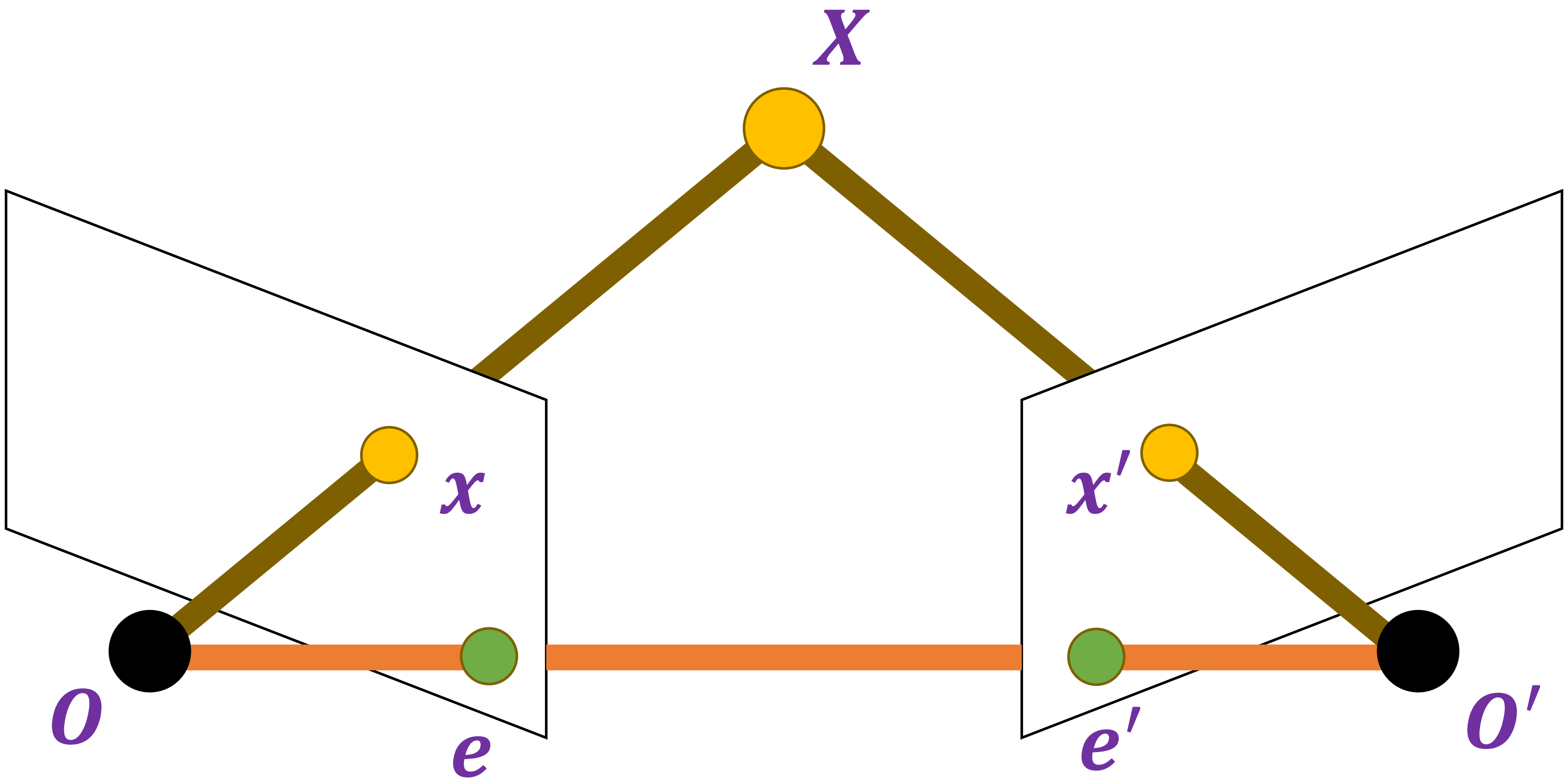
- Suppose we have two cameras with centers $\boldsymbol{O}$, $\boldsymbol{O'}$
- The **baseline** is the line connecting the origins

Epipolar geometry setup



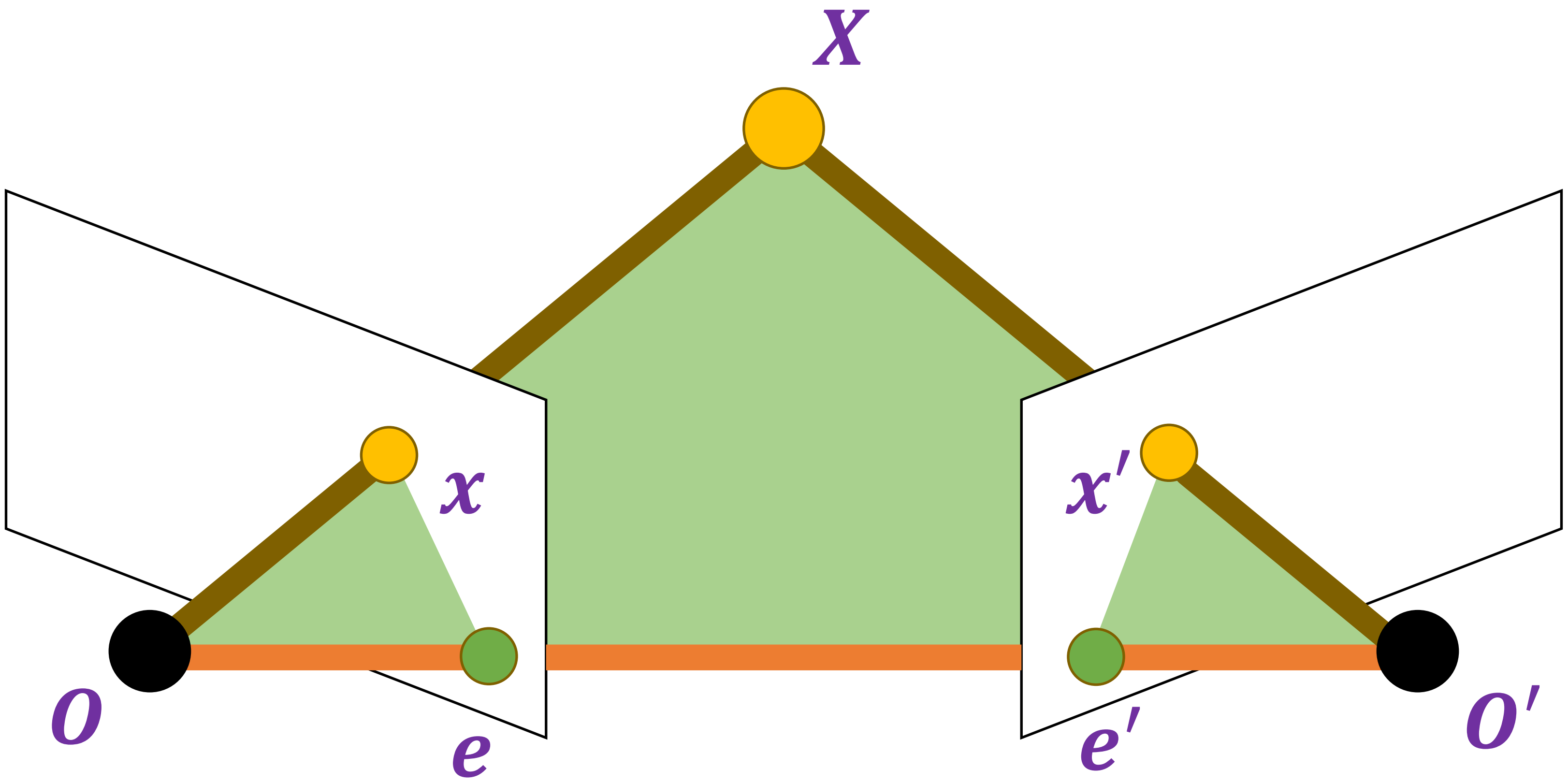
- **Epipoles** e, e' are where the baseline intersects the image planes, or projections of the other camera in each view

Epipolar geometry setup



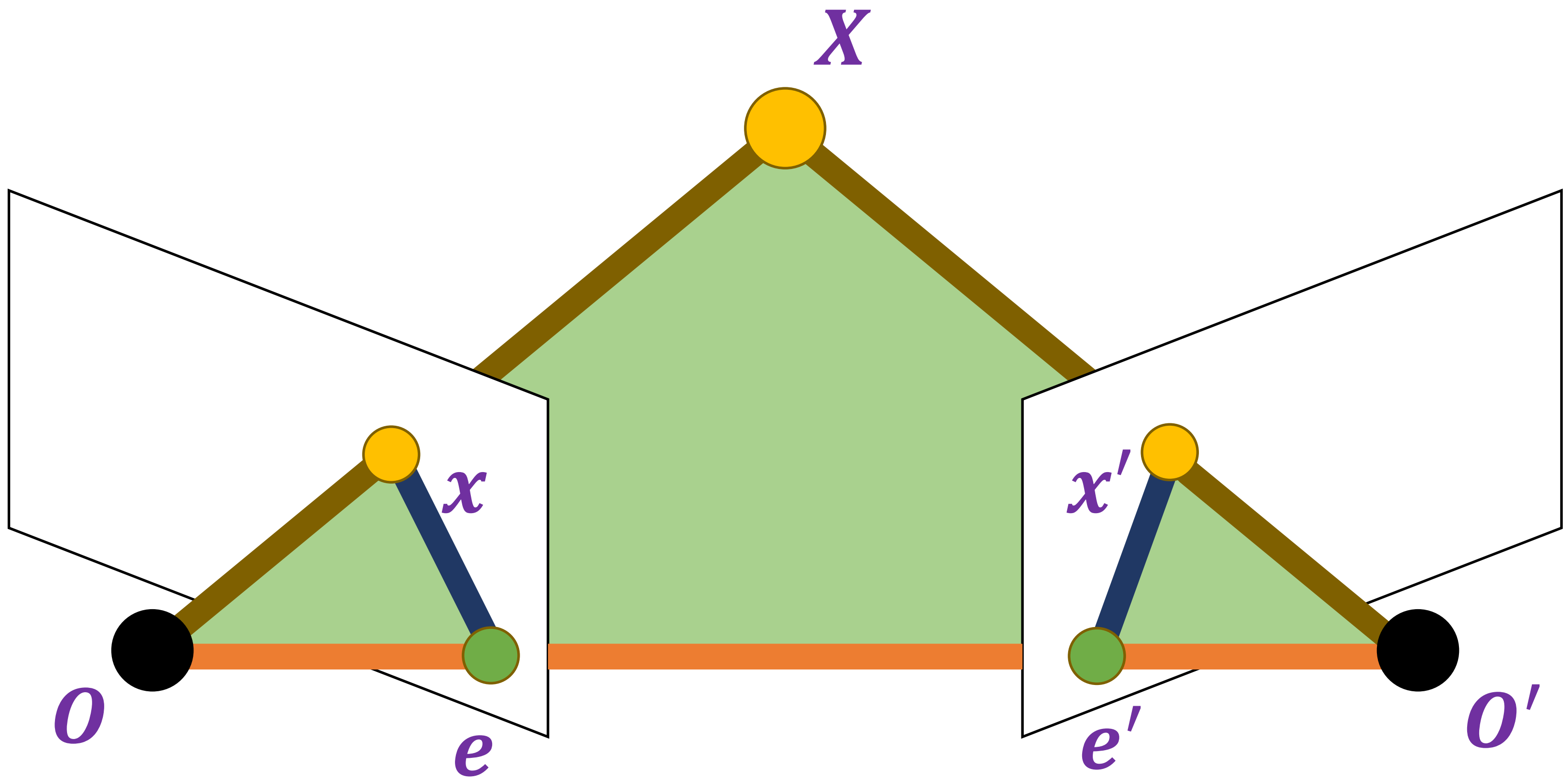
- Consider a **point** X, which projects to x and x'

Epipolar geometry setup



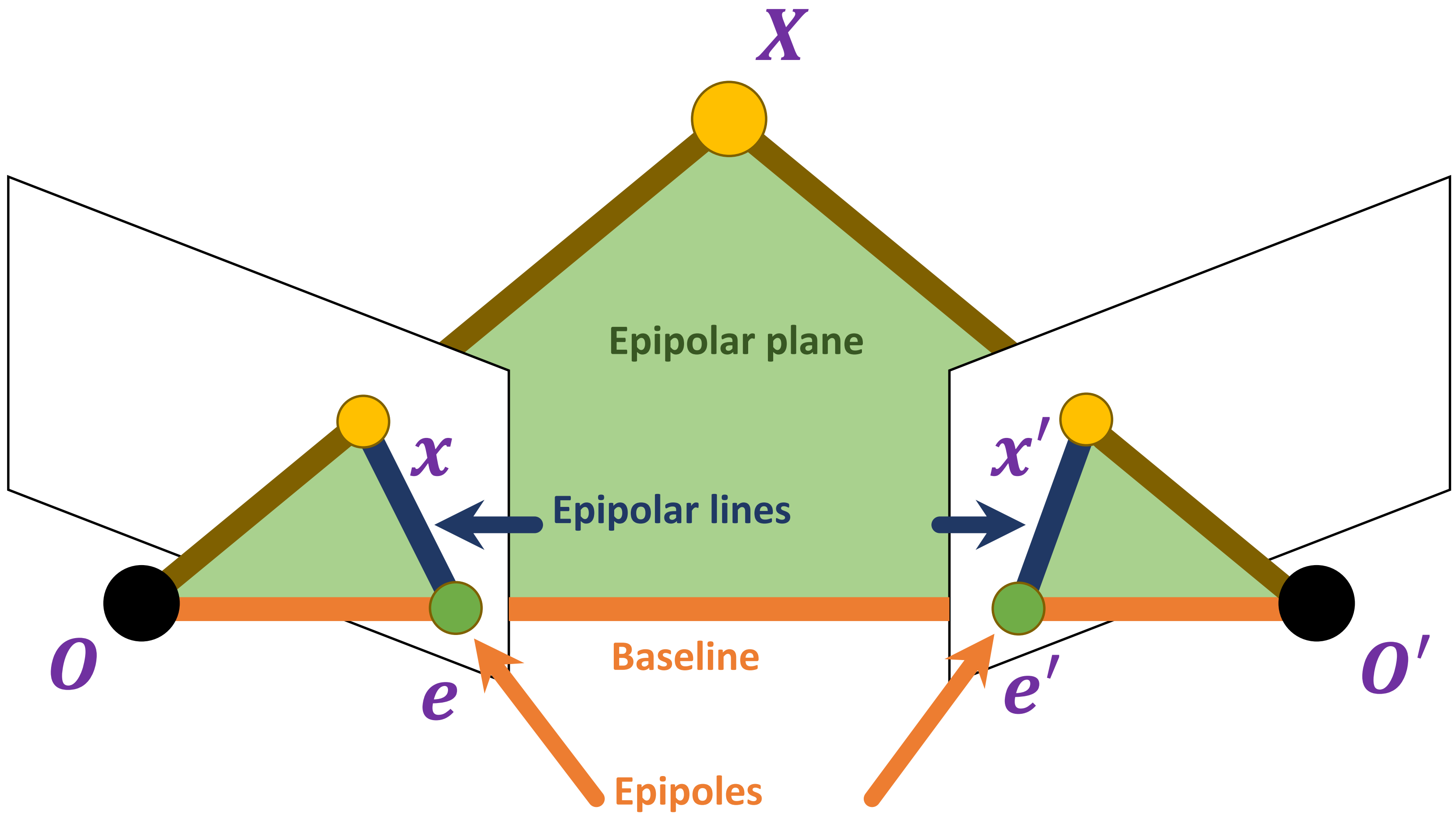
- The plane formed by X, O, and O' is called an **epipolar plane**
- There is a family of planes passing through O and O'

Epipolar geometry setup

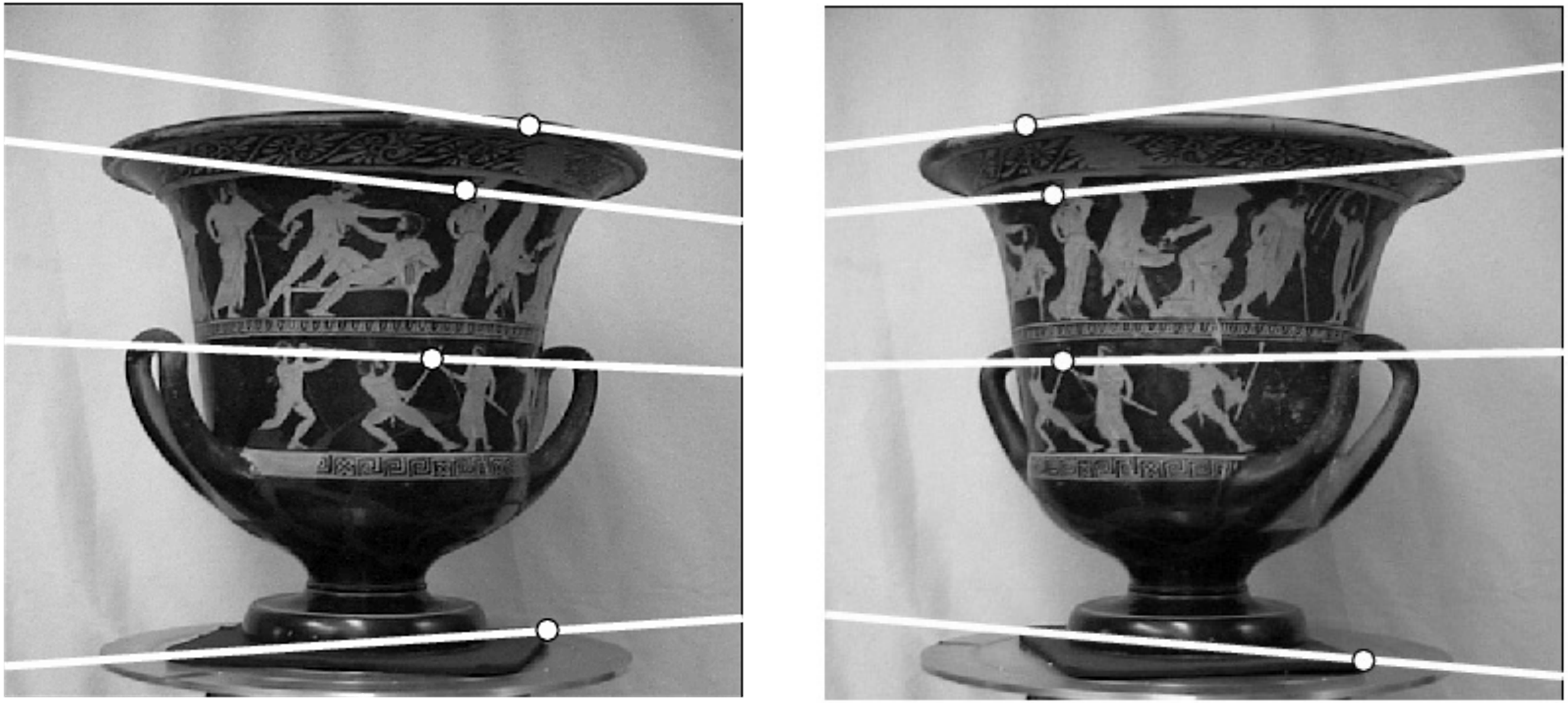


- **Epipolar lines** connect the epipoles to the projections of X
- Equivalently, they are intersections of the epipolar plane with the image planes – thus, they come in matching pairs

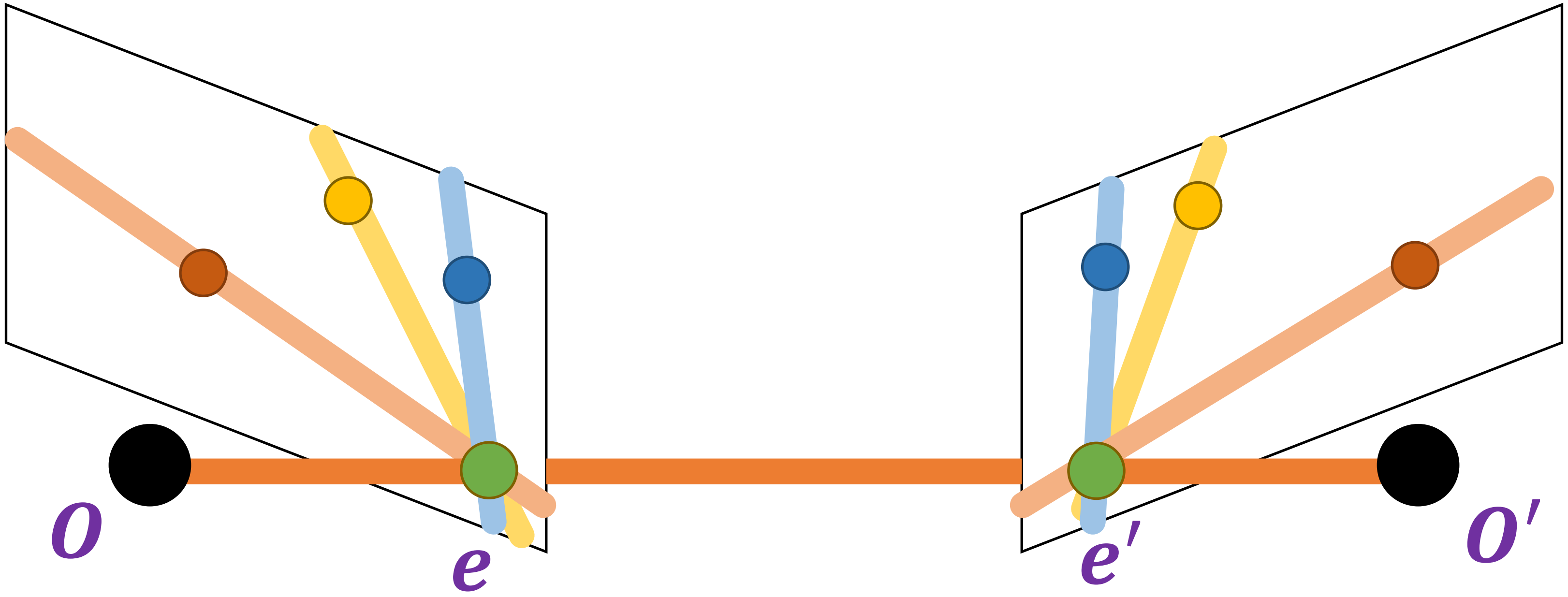
Epipolar geometry setup: Summary



Example configuration: Converging cameras

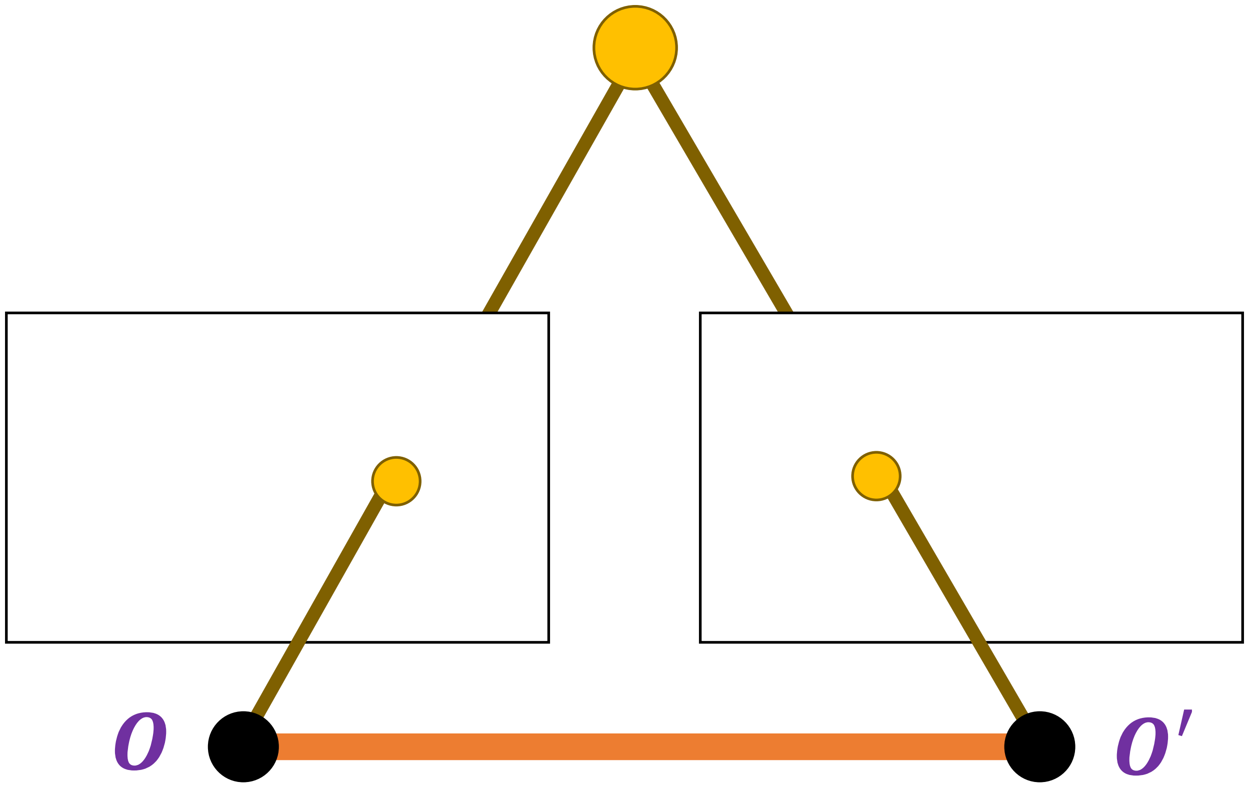


Example configuration: Converging cameras



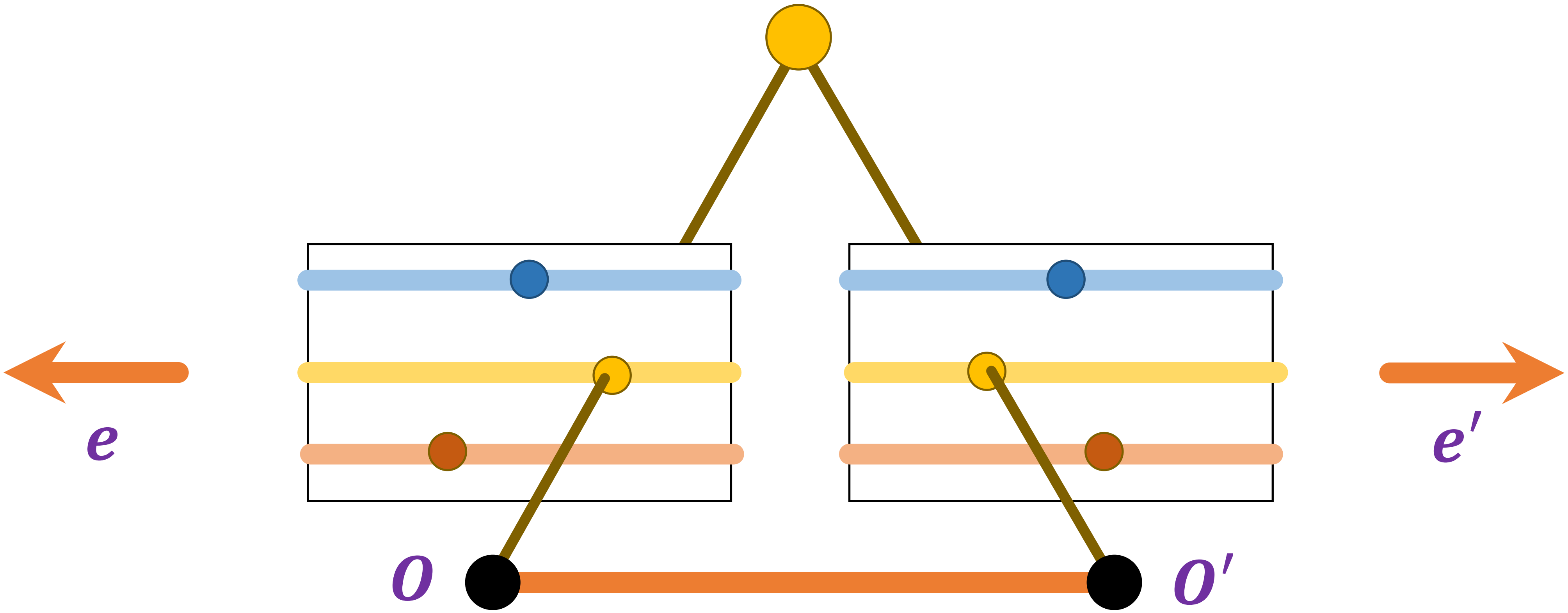
- Epipoles are finite, may be visible in the image

Example configuration: Motion parallel to image plane



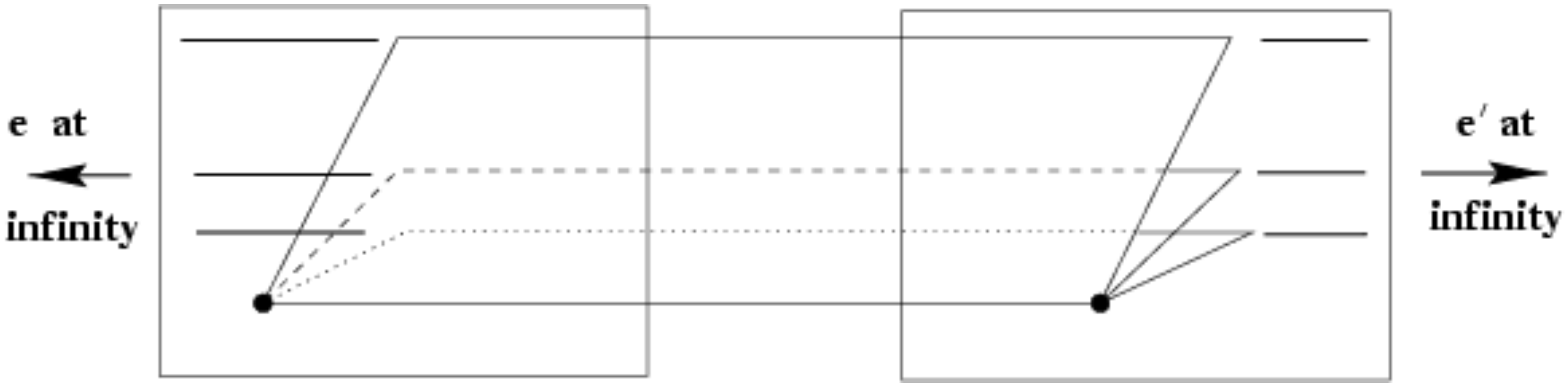
- Where are the epipoles and what do the epipolar lines look like?

Example configuration: Motion parallel to image plane



- Epipoles *infinitely* far away, epipolar lines parallel

Example configuration: Motion parallel to image plane

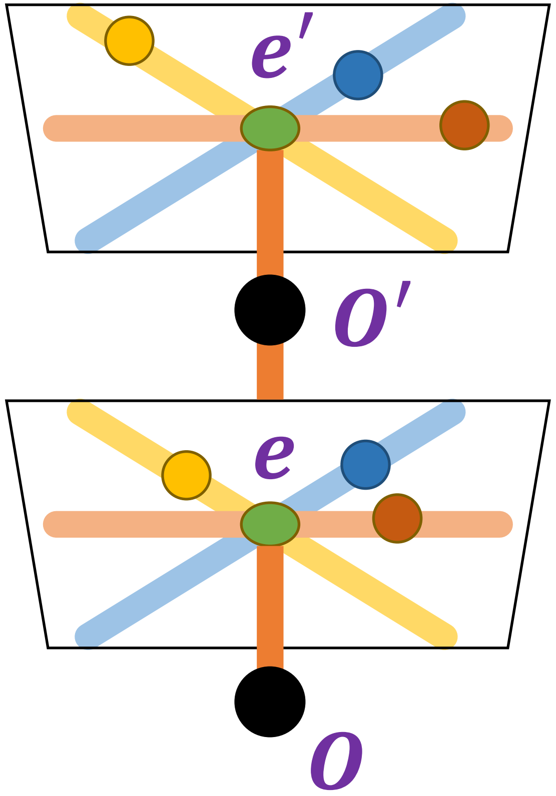


Example configuration:
Motion perpendicular to image plane

Example configuration: Motion perpendicular to image plane

Example configuration: Motion perpendicular to image plane

Technische Hochschule **Rosenheim**

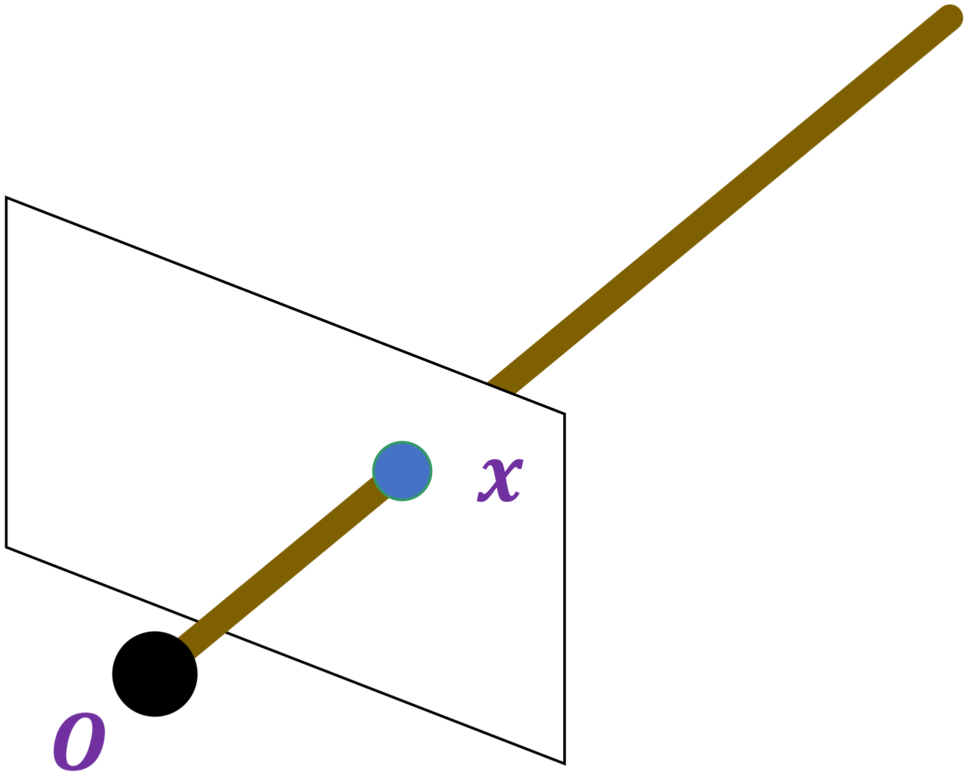


- Epipole is “focus of expansion” and coincides with the principal point of the camera
- Epipolar lines go out from principal point

Outline

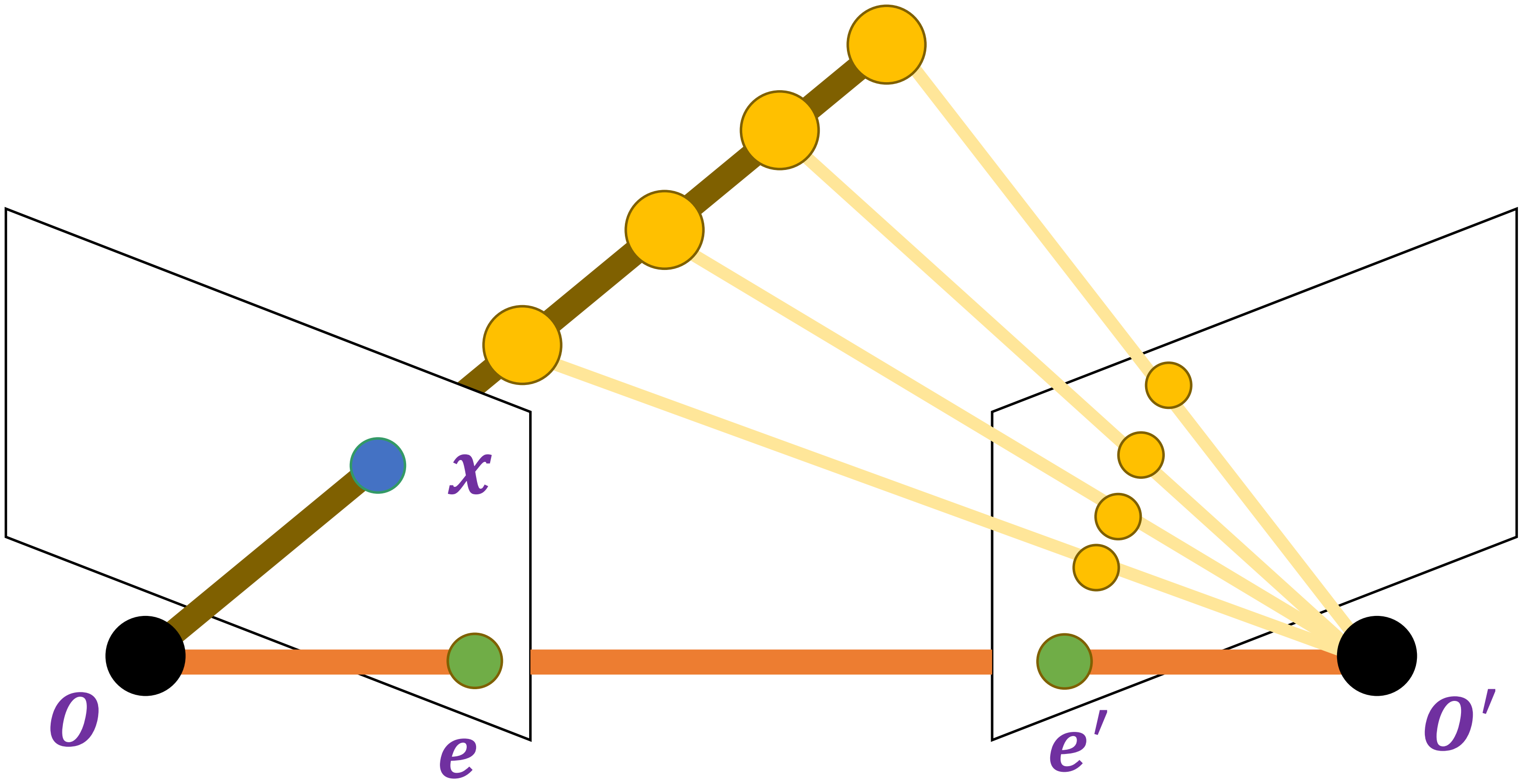
- Motivation
- Epipolar geometry setup
- Epipolar constraint

Epipolar constraint



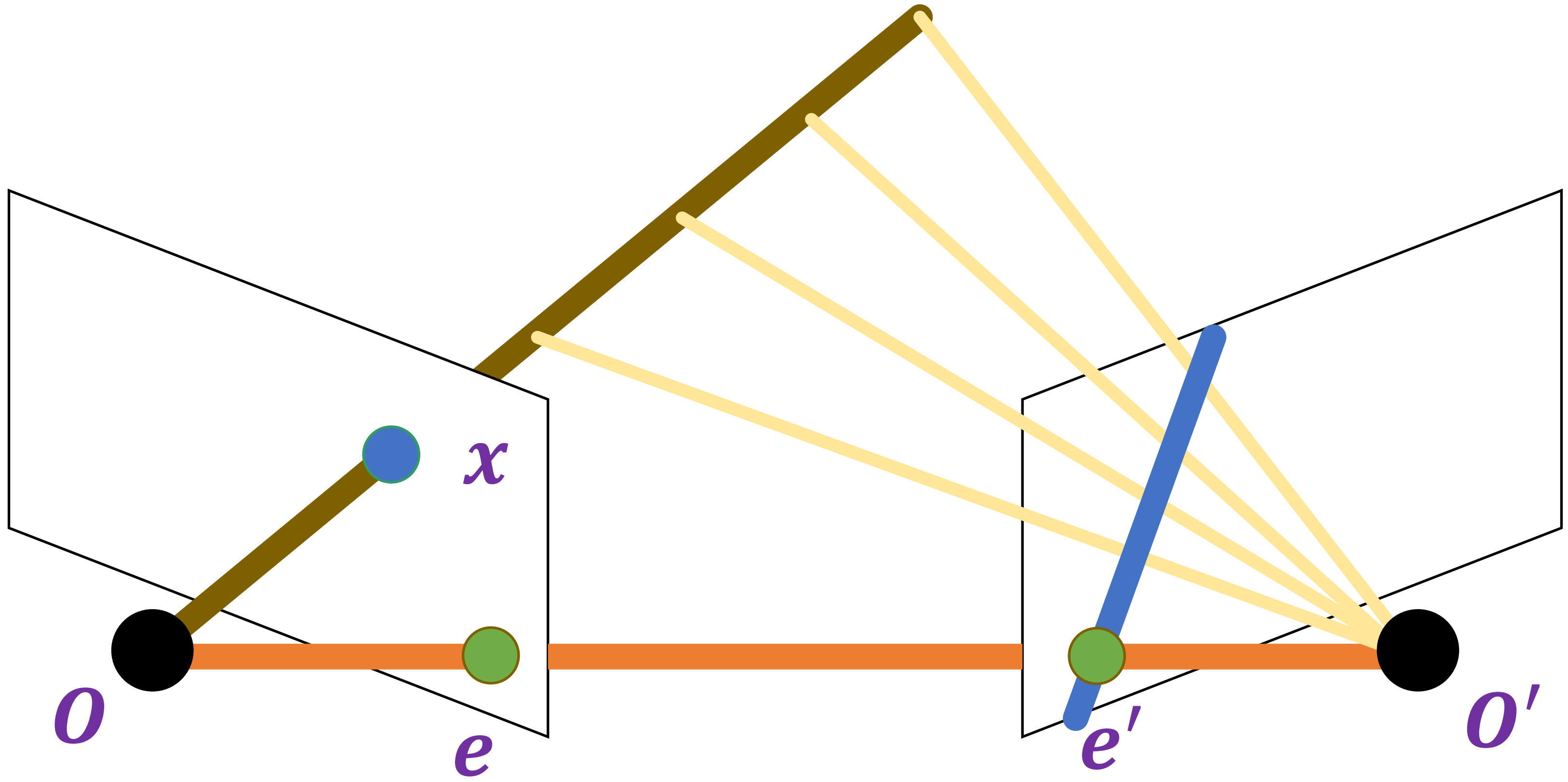
- Suppose we observe a single point $\boldsymbol{x}$ in one image

Epipolar constraint



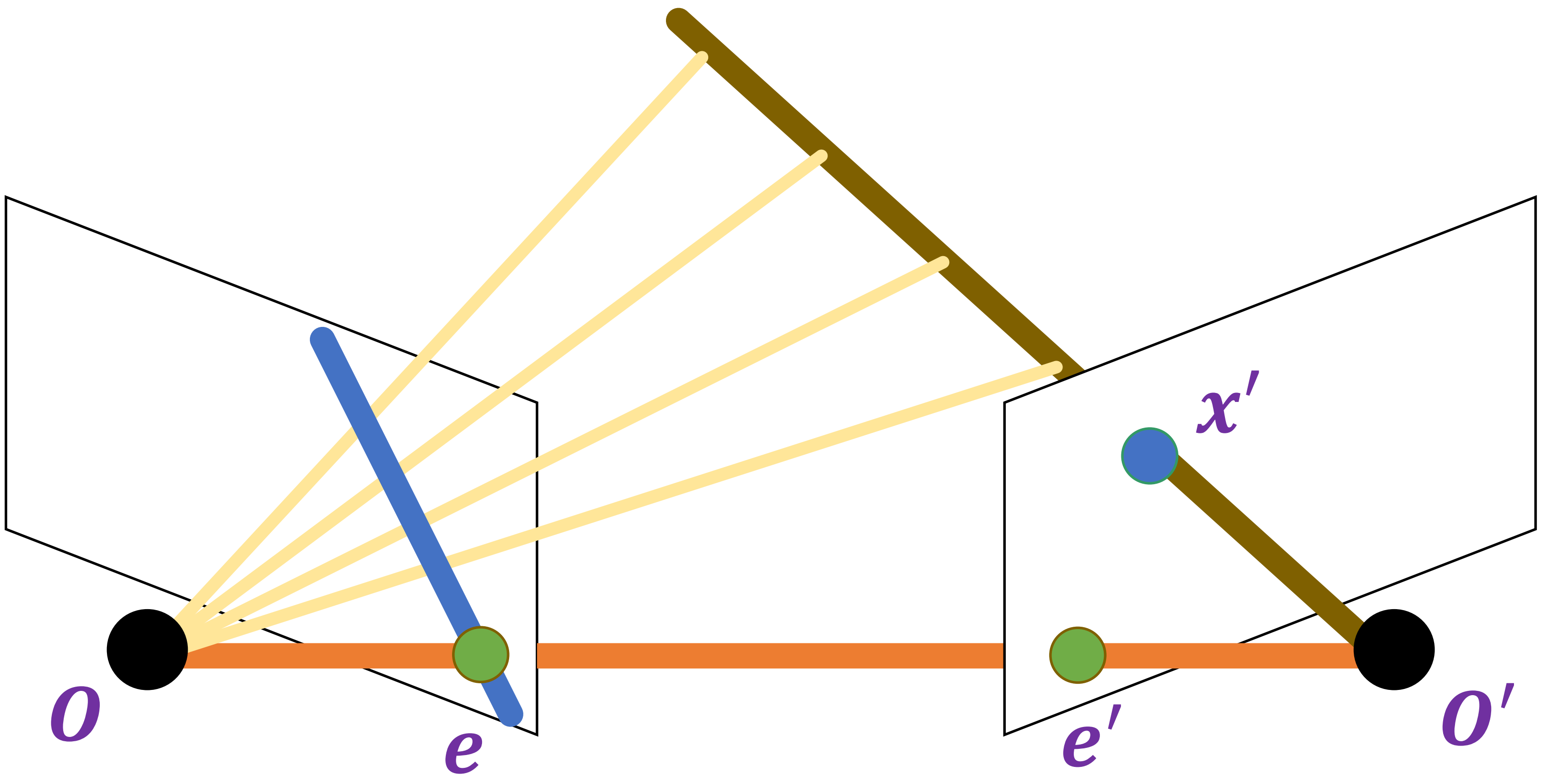
- Where can we find the x' corresponding to x in the other image?

Epipolar constraint



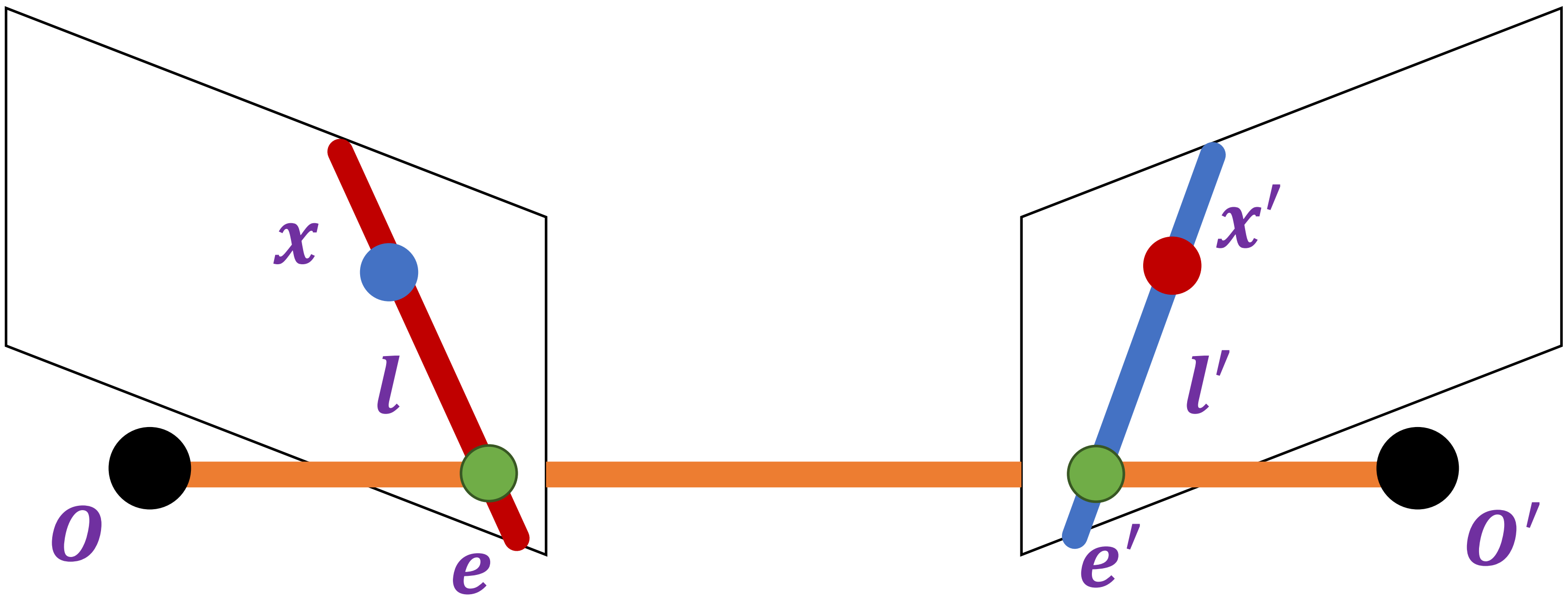
- Where can we find the $\boldsymbol{x}'$ corresponding to $\boldsymbol{x}$ in the other image?
- Along the **epipolar line** corresponding to $\boldsymbol{x}$ (projection of visual ray connecting $\boldsymbol{O}$ with $\boldsymbol{x}$ into the second image plane)

Epipolar constraint



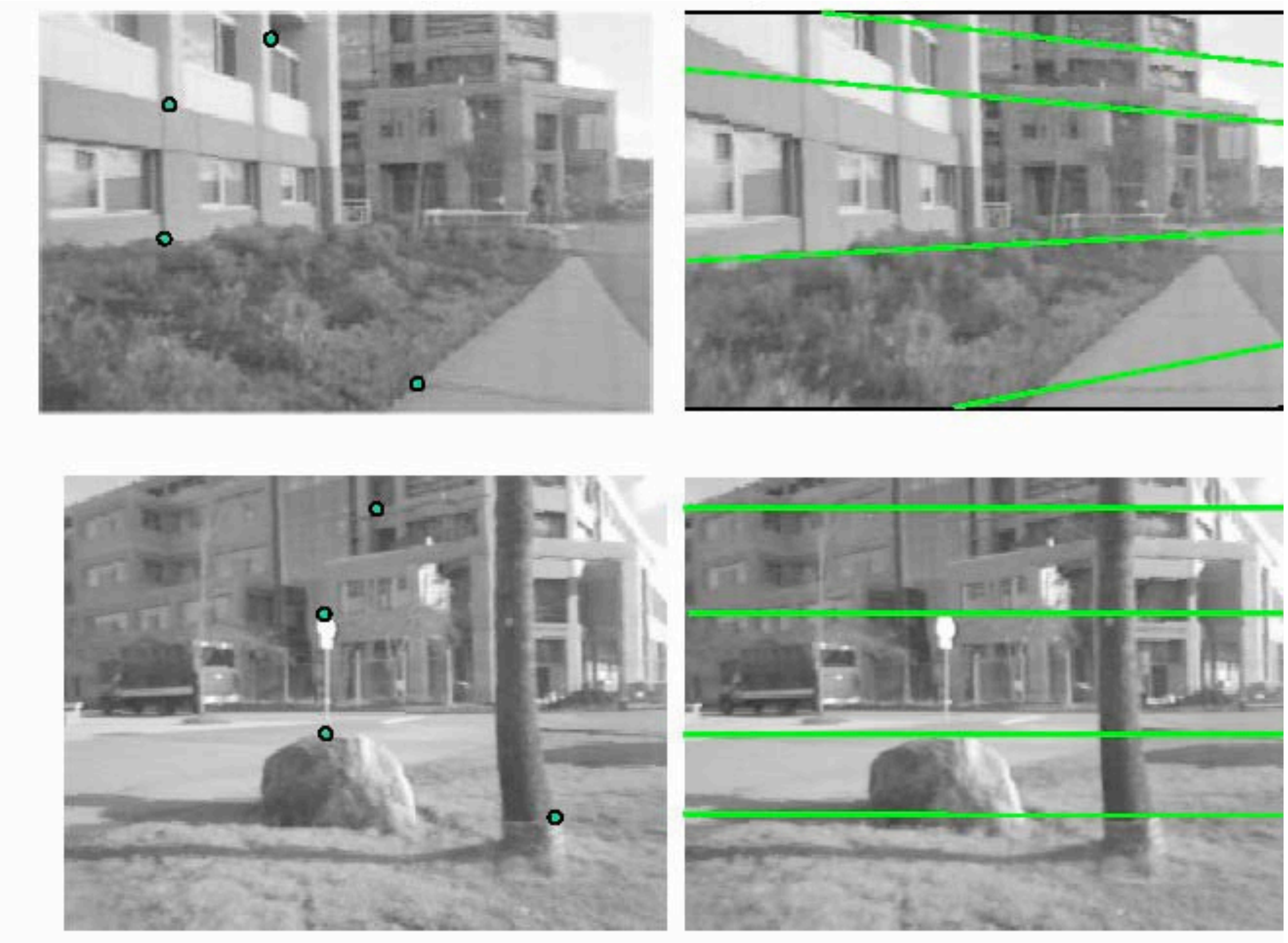
- Similarly, all points in the left image corresponding to x' have to lie along the epipolar line corresponding to x'

Epipolar constraint

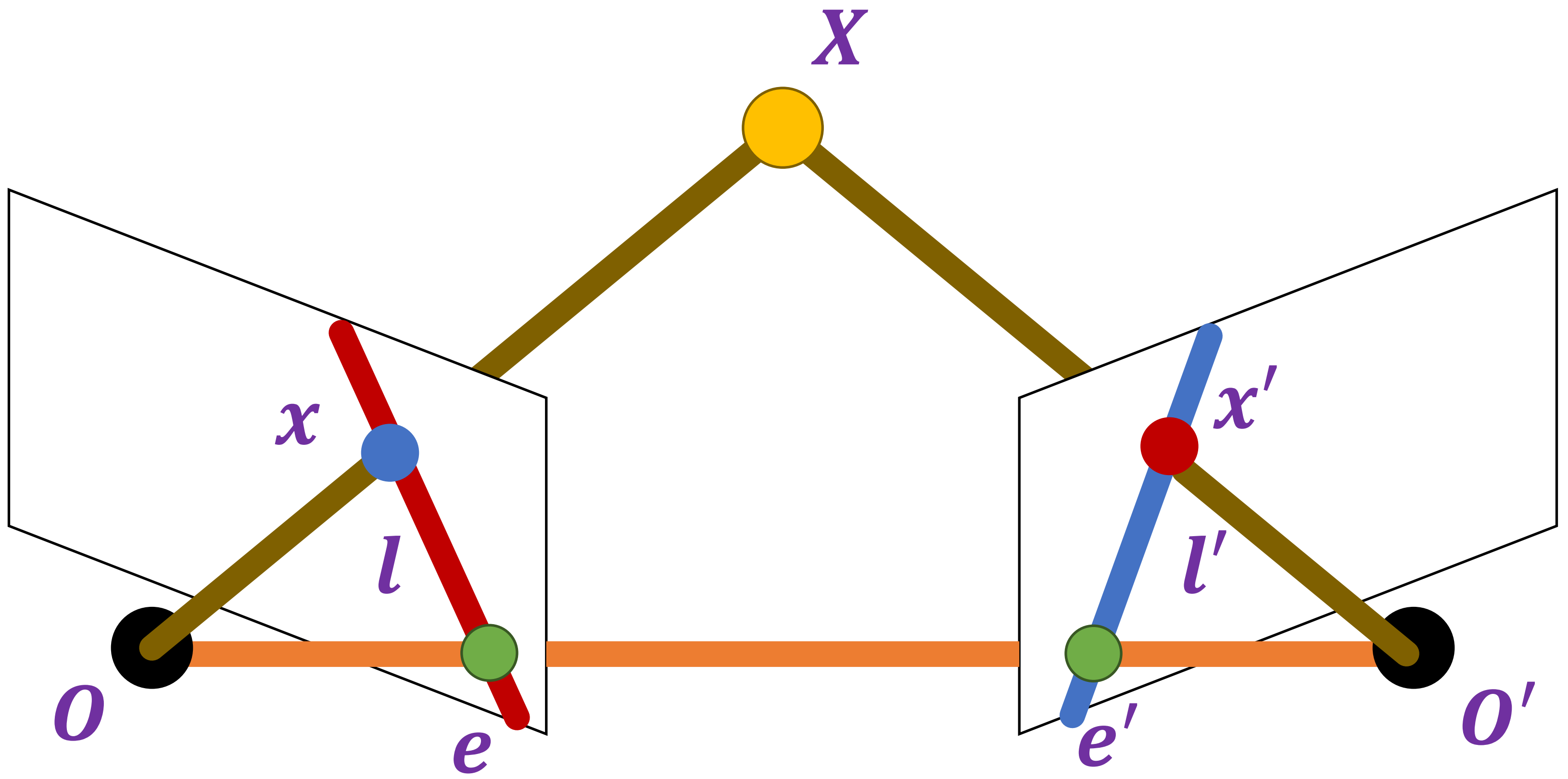


- Potential matches for x have to lie on the matching epipolar line l'
- Potential matches for x' have to lie on the matching epipolar line l

Epipolar constraint: Example

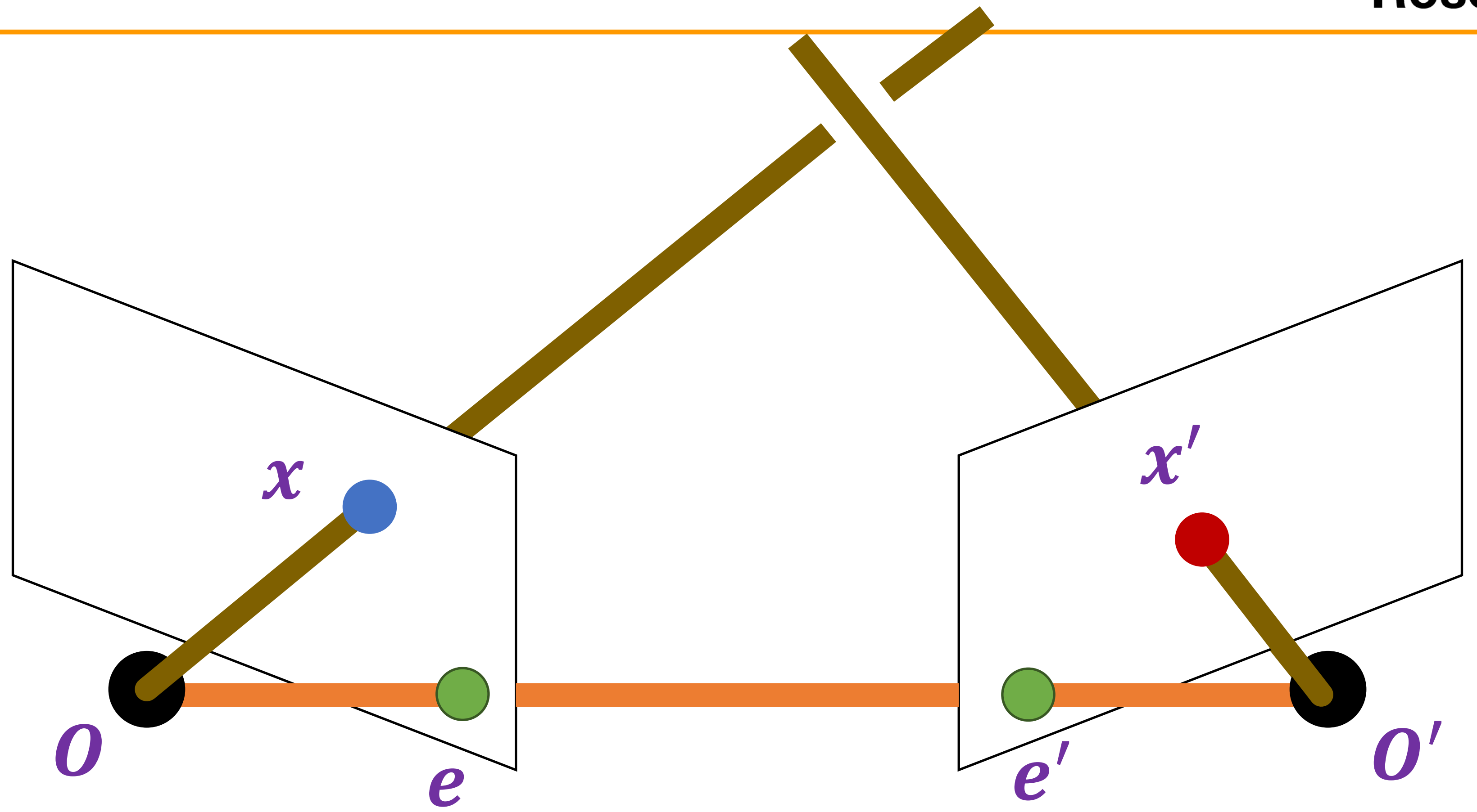


Epipolar constraint



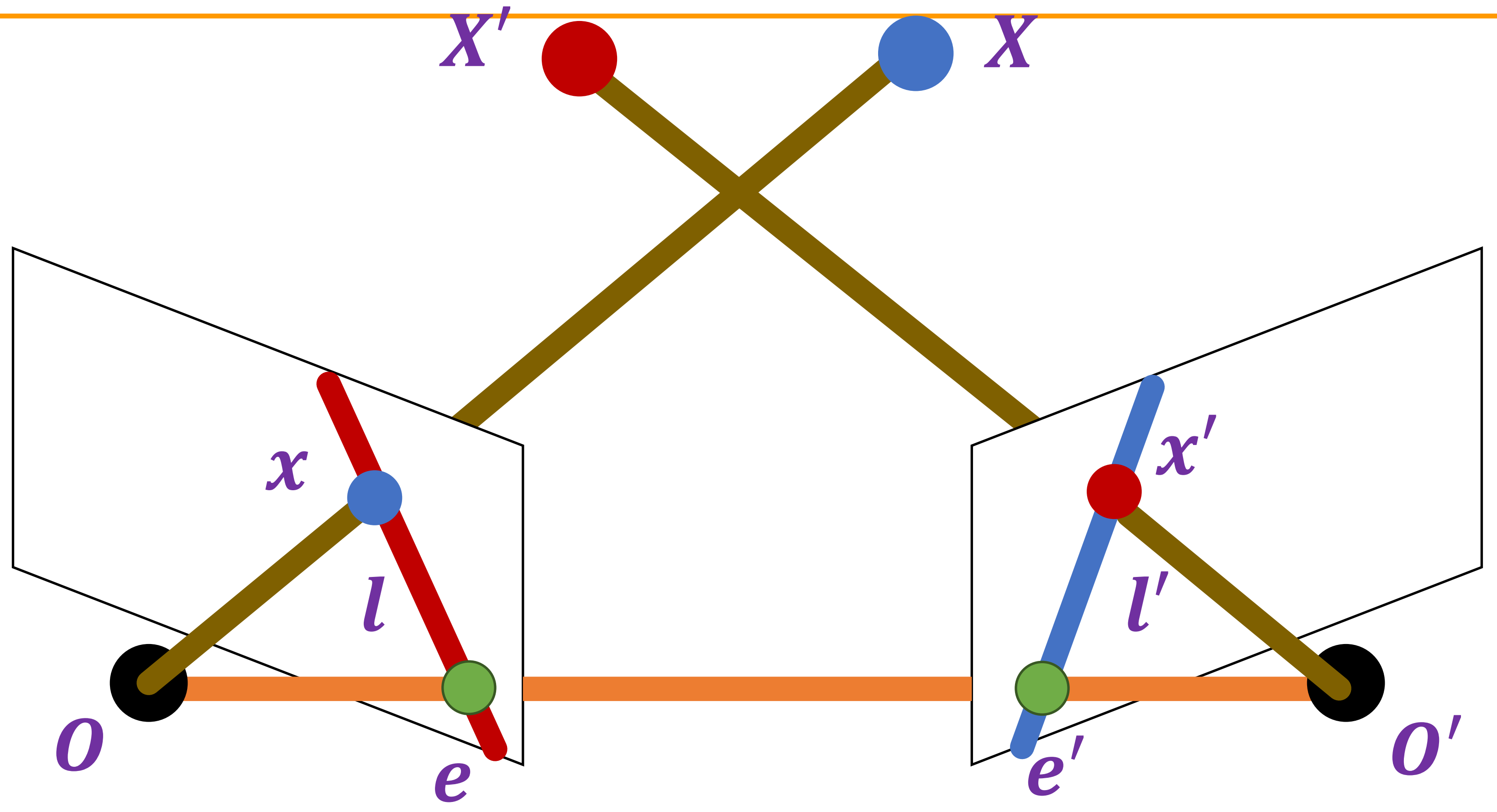
- Whenever two points x and x' lie on matching epipolar lines l and l', the visual rays corresponding to them meet in space, i.e., x and x' could be projections of the same 3D point X

Epipolar constraint



- Remember: in general, two rays *do not* meet in space!

Epipolar constraint



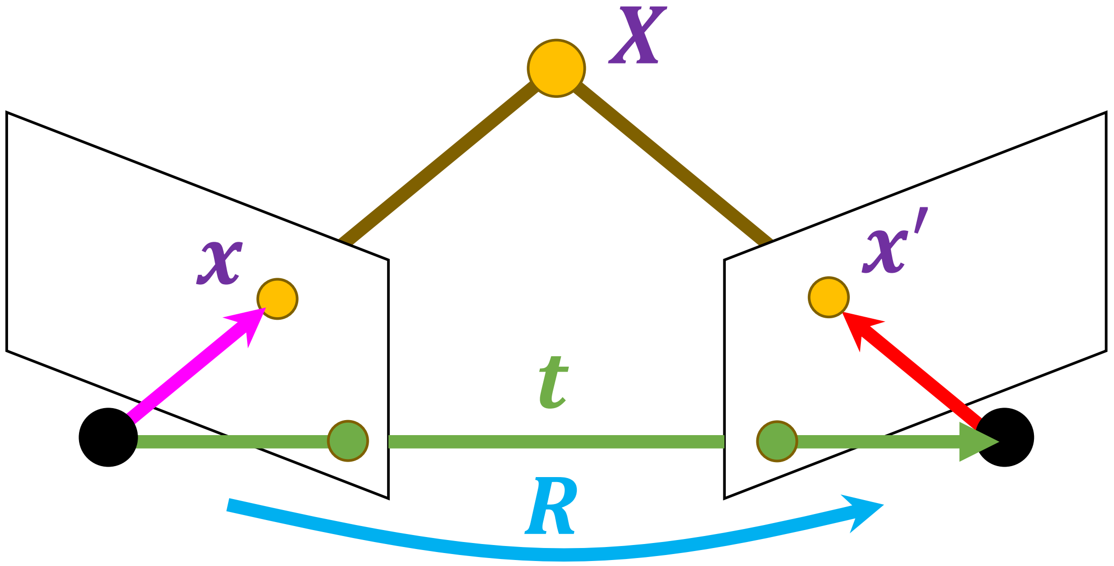
- Caveat: if x and x' satisfy the epipolar constraint, this doesn't mean they *have to be* projections of the same 3D point

Outline

- Motivation
- Epipolar geometry setup
- Epipolar constraint
- Essential matrix

Math of the epipolar constraint: Calibrated case

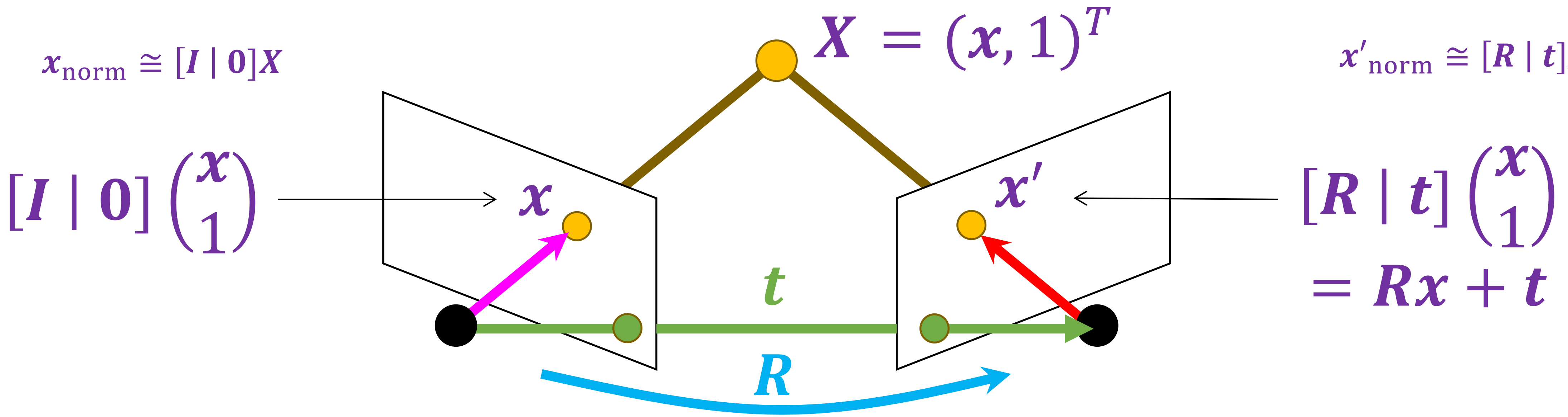
Technische Hochschule Rosenheim



- Assume the intrinsic and extrinsic parameters of the cameras are known, world coordinate system is set to that of the first camera
- Then the projection matrices are given by $K[I \mid 0]$ and $K'[R \mid t]$
- We can pre-multiply the projection matrices (and the image points) by the inverse calibration matrices to get *normalized* image coordinates

$$x_{\text{norm}} = K^{-1} x_{\text{pixel}} \cong [I \mid 0] X, \qquad x'_{\text{norm}} = K'^{-1} x'_{\text{pixel}} \cong [R \mid t] X$$

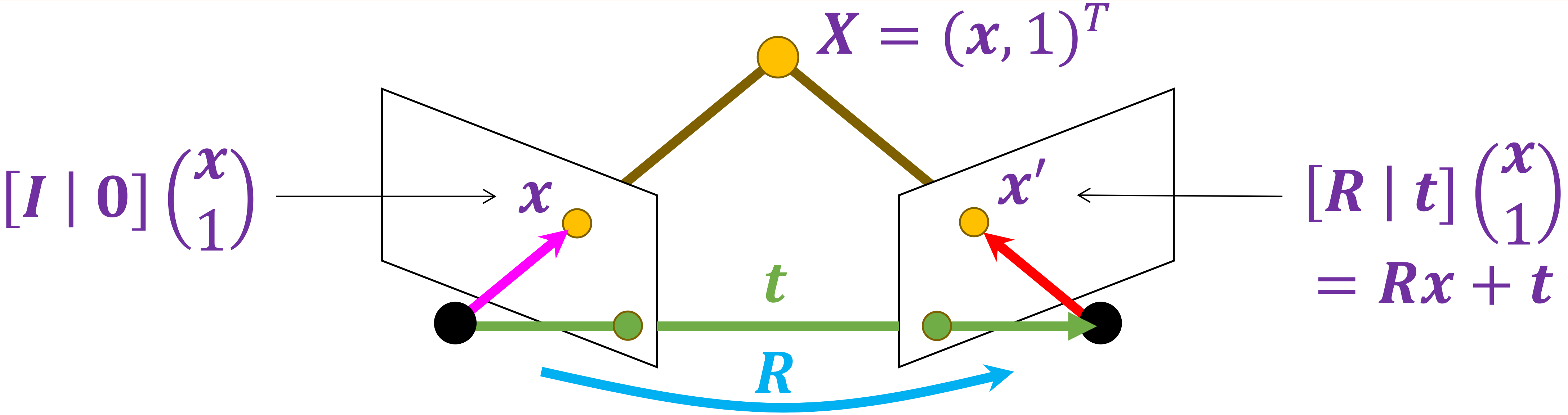
Math of the epipolar constraint: Calibrated case



- We have $\boldsymbol{x}' \cong \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$
- This means the three vectors $\boldsymbol{x}'$, $\boldsymbol{R}\boldsymbol{x}$, and $\boldsymbol{t}$ are linearly dependent
- This constraint can be written using the *triple product*

$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0$$

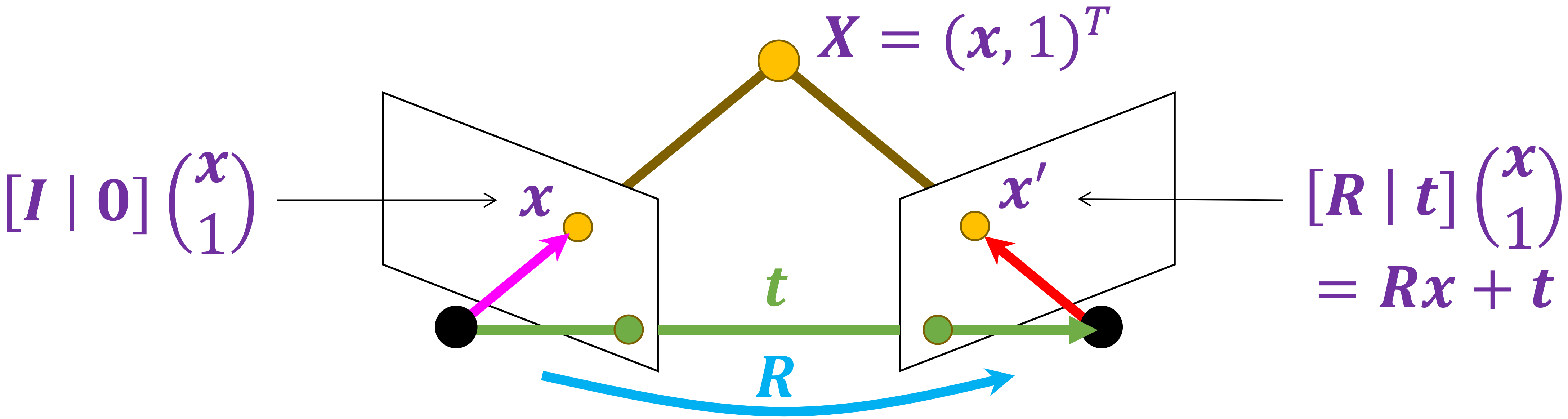
Math of the epipolar constraint: Calibrated case



$$\boldsymbol{x}' \cdot [\boldsymbol{t}\times(\boldsymbol{R}\boldsymbol{x})] = 0 \quad \Rightarrow \quad \boldsymbol{x}'^T[\boldsymbol{t}_\times]\boldsymbol{R}\boldsymbol{x} = 0$$

Recall: $\boldsymbol{a}\times\boldsymbol{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = [\boldsymbol{a}_\times]\boldsymbol{b}$

Math of the epipolar constraint: Calibrated case



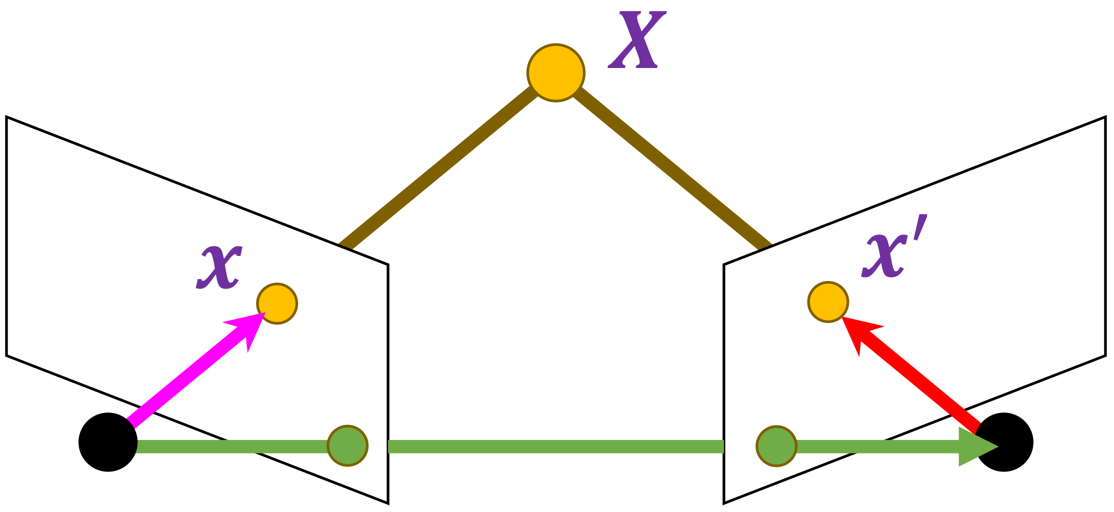
$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0 \quad \Rightarrow \quad \boldsymbol{x}'^T [\boldsymbol{t}_\times] \boldsymbol{R}\boldsymbol{x} = 0 \quad \Rightarrow \quad \boldsymbol{x}'^T \boldsymbol{E}\boldsymbol{x} = 0$$

$\Downarrow$

Essential Matrix

H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature 293 (5828): 133–135, September 1981

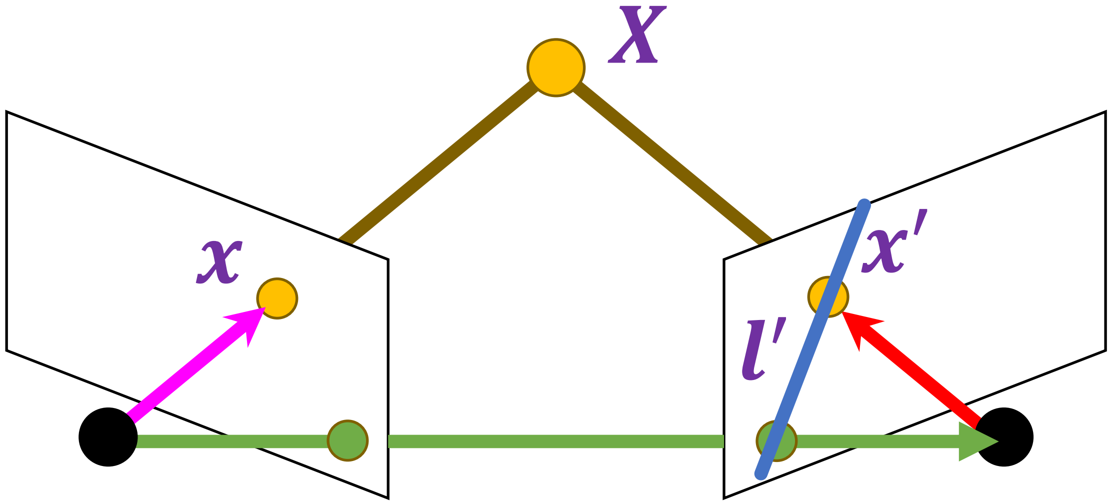
The essential matrix



$$\boxed{\boldsymbol{x}'^T \boldsymbol{E} \boldsymbol{x} = 0}$$

$$(x', y', 1) \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

The essential matrix: Properties

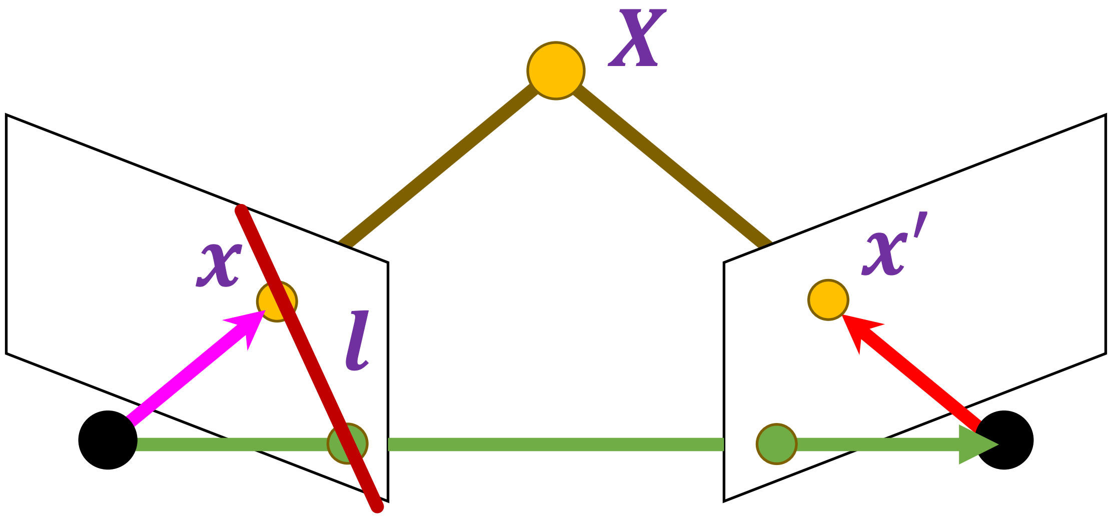


$$\boxed{\boldsymbol{x}'^T \boldsymbol{E}\boldsymbol{x} = 0}$$

- $\boldsymbol{E}\boldsymbol{x}$ is the epipolar line associated with $\boldsymbol{x}$ ($\boldsymbol{l}' = \boldsymbol{E}\boldsymbol{x}$)

Recall: a line is given by $ax + by + c = 0$ or $\boldsymbol{l}^T\boldsymbol{x} = 0$
where $\boldsymbol{l} = (a, b, c)^T$ and $\boldsymbol{x} = (x, y, 1)^T$

The essential matrix: Properties



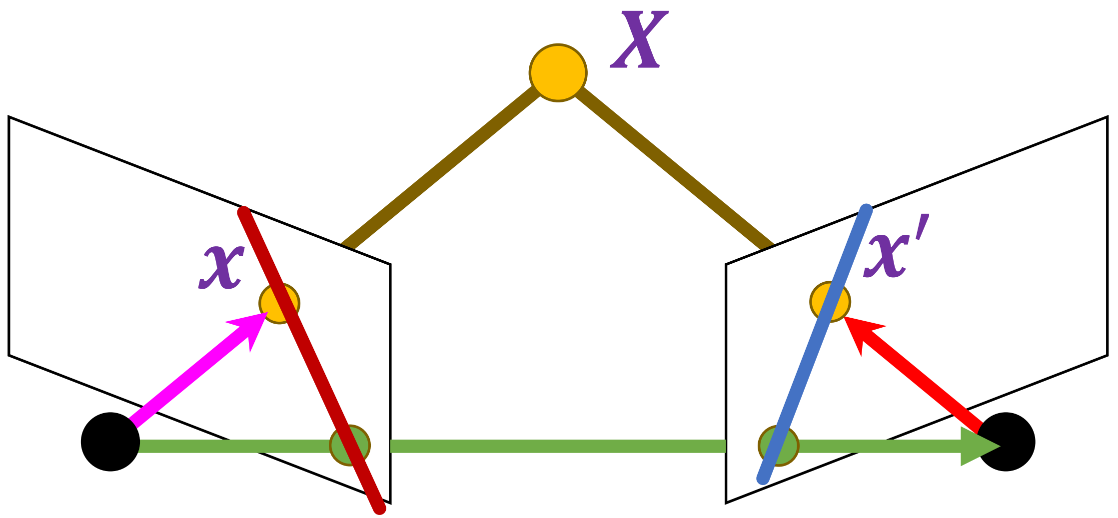
$$\boxed{x'^T E x = 0}$$

- $\boldsymbol{Ex}$ is the epipolar line associated with $\boldsymbol{x}$ ($\boldsymbol{l}' = \boldsymbol{Ex}$)
- $\boldsymbol{E}^T\boldsymbol{x}'$ is the epipolar line associated with $\boldsymbol{x}'$ ($\boldsymbol{l} = \boldsymbol{E}^T\boldsymbol{x}'$)
- $\boldsymbol{Ee} = \boldsymbol{0}$ and $\boldsymbol{E}^T\boldsymbol{e}' = \boldsymbol{0}$
- $\boldsymbol{E}$ is singular (rank two) and has five degrees of freedom

Outline

- Motivation
- Epipolar geometry setup
- Epipolar constraint
- Essential matrix
- Fundamental matrix

Epipolar constraint: Uncalibrated case

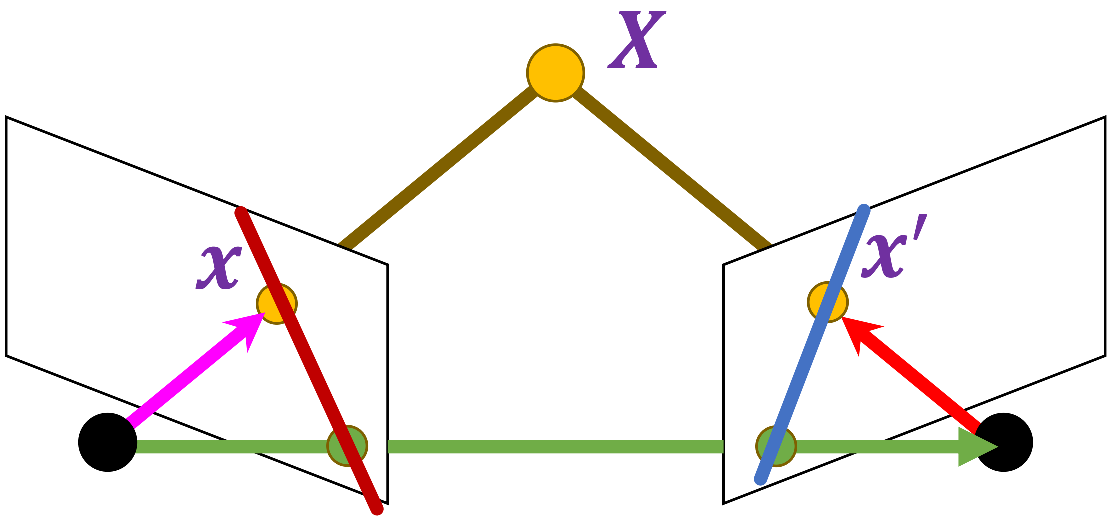


- The calibration matrices $\boldsymbol{K}$ and $\boldsymbol{K}'$ of the two cameras are unknown
- We can write the epipolar constraint in terms of *unknown* normalized coordinates:

$$\boldsymbol{x}'^{T}_{\mathrm{norm}}\boldsymbol{E}\boldsymbol{x}_{\mathrm{norm}} = 0,$$

where $\boldsymbol{x}_{\mathrm{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}, \boldsymbol{x}'_{\mathrm{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'$

Epipolar constraint: Uncalibrated case



$$\boldsymbol{x}'^{T}_{\text{norm}}\boldsymbol{E}\boldsymbol{x}_{\text{norm}} = 0$$ ⇒ $\boldsymbol{x}'^{T}\boldsymbol{F}\boldsymbol{x} = 0$, where $\boldsymbol{F} = \boldsymbol{K}'^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}$$

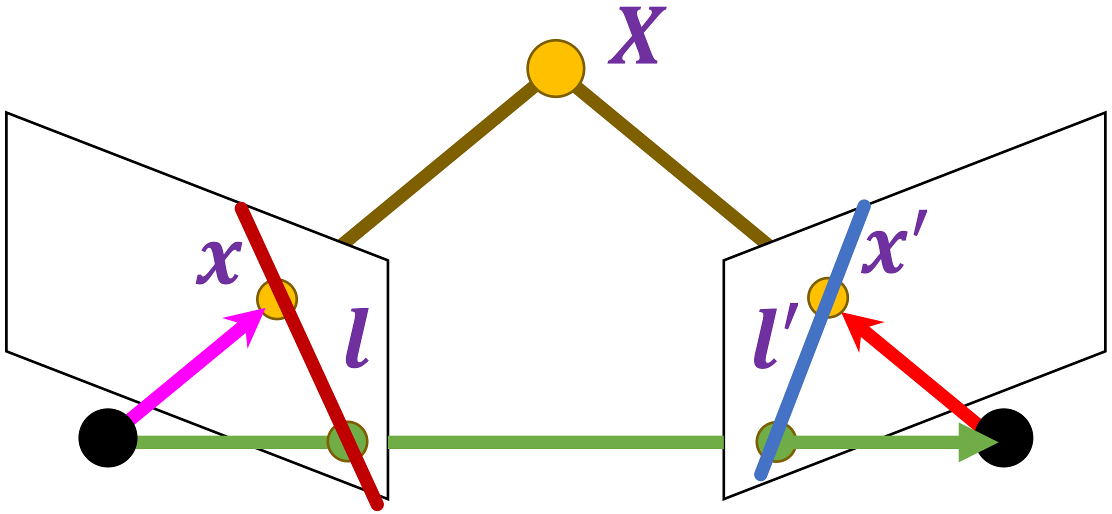
$$\boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'$$

⇓

Fundamental Matrix

Faugeras et al., (1992), Hartley (1992)

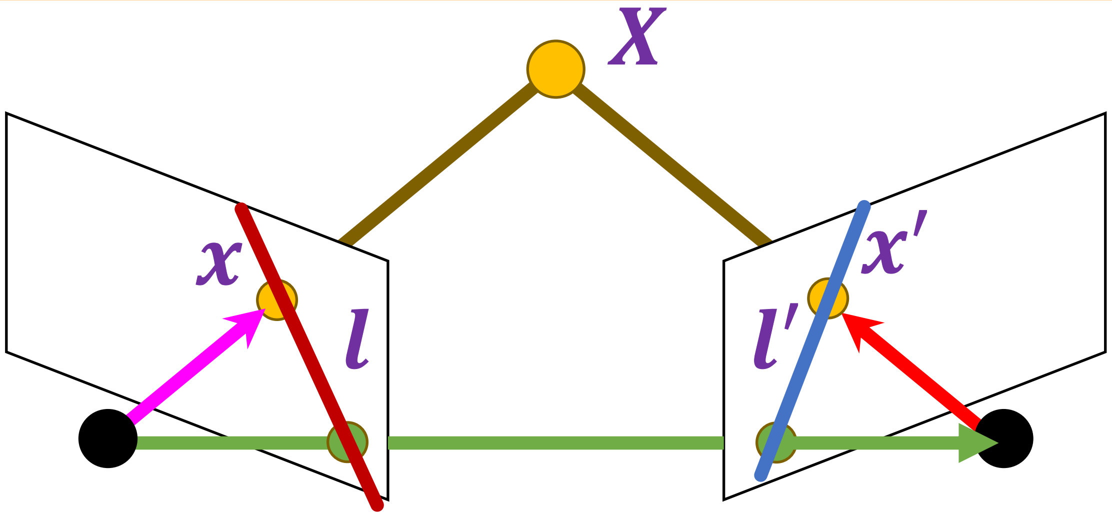
The fundamental matrix



$$\boxed{\boldsymbol{x'}^T \boldsymbol{F} \boldsymbol{x} = 0}$$

$$(x', y', 1) \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

The fundamental matrix: Properties



$$\boxed{\boldsymbol{x}'^T \boldsymbol{F} \boldsymbol{x} = 0}$$

- $\boldsymbol{F}\boldsymbol{x}$ is the epipolar line associated with $\boldsymbol{x}$ ($\boldsymbol{l}' = \boldsymbol{F}\boldsymbol{x}$)
- $\boldsymbol{F}^T\boldsymbol{x}'$ is the epipolar line associated with $\boldsymbol{x}'$ ($\boldsymbol{l} = \boldsymbol{F}^T\boldsymbol{x}'$)
- $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{0}$ and $\boldsymbol{F}^T\boldsymbol{e}' = \boldsymbol{0}$
- $\boldsymbol{F}$ is singular (rank two) and has seven degrees of freedom

Outline

- Motivation
- Epipolar geometry setup
- Epipolar constraint
- Essential matrix
- Fundamental matrix
- Estimating the fundamental matrix

Estimating the fundamental matrix

Given: correspondences $\boldsymbol{x} = (x, y, 1)^T$ and $\boldsymbol{x}' = (x', y', 1)^T$

Estimating the fundamental matrix

- Given: correspondences $\boldsymbol{x} = (x, y, 1)^T$ and $\boldsymbol{x}' = (x', y', 1)^T$
- Constraint: $\boldsymbol{x}'^T \boldsymbol{F} \boldsymbol{x} = 0$

$$(x', y', 1) \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad (x'x, x'y, x', y'x, y'y, y', x, y, 1) \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0$$

The eight point algorithm

$$\underbrace{\begin{bmatrix} & & & \vdots & & & & & \\ x'x & x'y & x' & y'x & y'y & y' & x & y & 1 \\ & & & \vdots & & & & & \end{bmatrix}}_{\boldsymbol{U}} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \mathbf{0}$$

Homogeneous least squares to find $\boldsymbol{f}$:

$$\arg\min_{\|\boldsymbol{f}\|=1} \|\boldsymbol{U}\boldsymbol{f}\|_2^2 \rightarrow$$ Eigenvector of $\boldsymbol{U}^T\boldsymbol{U}$ with smallest eigenvalue

Enforcing rank-2 constraint

- We know $\boldsymbol{F}$ needs to be singular/rank 2. How do we force it to be singular?
- Solution: take SVD of the initial estimate and throw out the smallest singular value

$$\boldsymbol{F}_{\text{init}} = \boldsymbol{U\Sigma V}^T$$

$$\downarrow$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \cancel{\sigma_3} \end{bmatrix} \longrightarrow \boldsymbol{\Sigma}' = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\downarrow$$

$$\boldsymbol{F} = \boldsymbol{U\Sigma}'\boldsymbol{V}^T$$

Enforcing rank-2 constraint

Initial $\boldsymbol{F}$ estimate

Rank-2 estimate

Normalized eight point algorithm

$$\underbrace{\begin{array}{ccccccccc} 10^6 & 10^6 & 10^3 & 10^6 & 10^6 & 10^3 & 10^3 & 10^3 & 1 \\ \end{array}}_{}$$

$$\underbrace{\begin{bmatrix} & & & \vdots & & & & & \\ x'x & x'y & x' & y'x & y'y & y' & x & y & 1 \\ & & & \vdots & & & & & \end{bmatrix}}_{\boldsymbol{U}} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \mathbf{0}$$

- Recall that x, y, x', y' are pixel coordinates. What might be the order of magnitude of each column of $\boldsymbol{U}$? Suppose the image is 1000x1000.
- This causes numerical instability!

Eight Point Algorithm – Numerical Instability

- Numbers of varying magnitude → instability

- Remember: a floating point number (float/double) isn't a "real" number: for sign, mantissa, exponent integers

$$(-1)^{\text{sign}} * \text{coefficient} * 2^{\text{exponent}}$$

- Exercise to see how this screws up: add up Gaussian noise (mean=100, std=10), divide by number you added up

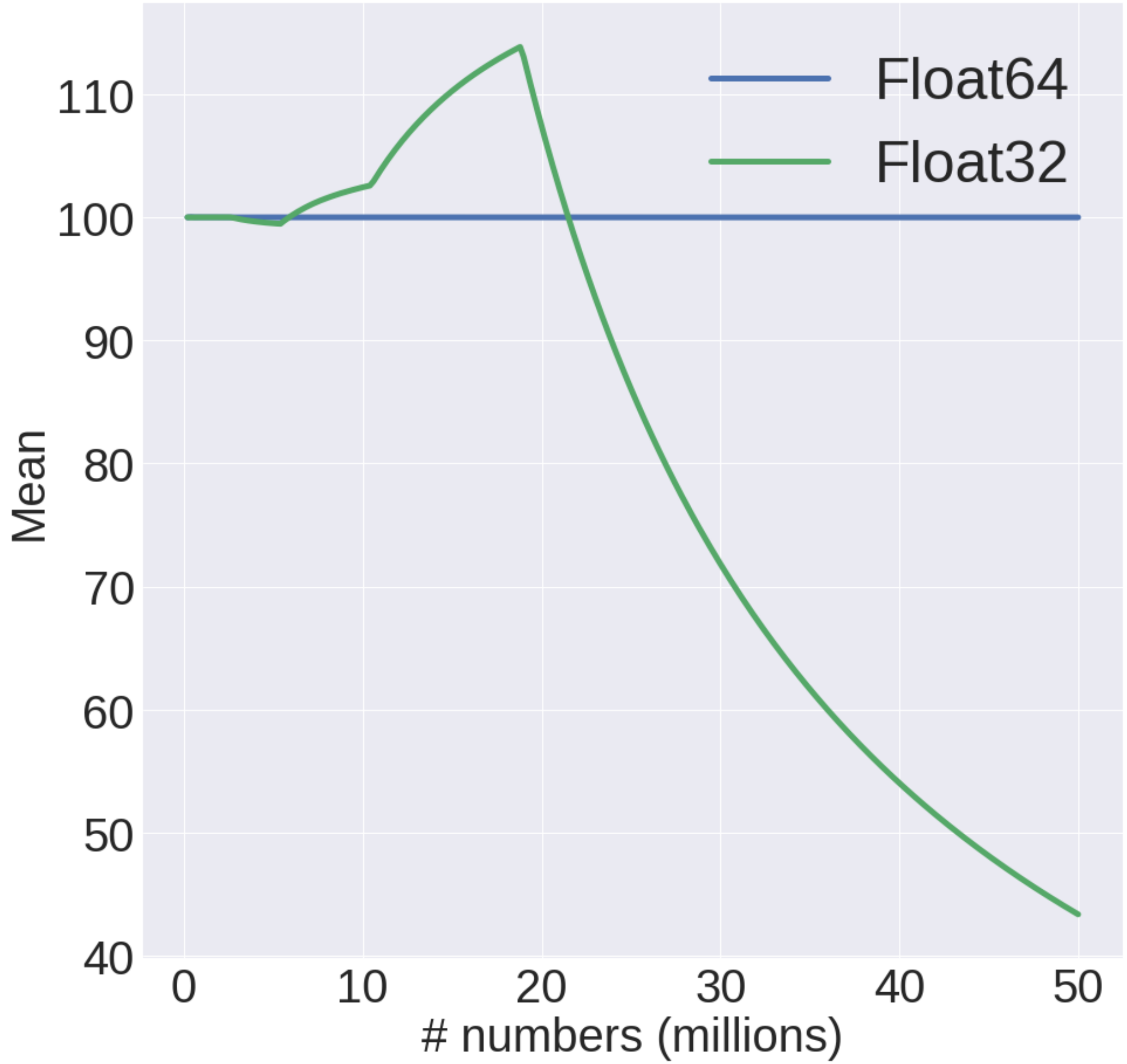
Remember Numerical Instability?

Code:

```
x += N(100,10)
i += 1
mean = x/I
```

Only change is the # of bits in accumulator x

Note: 50M is 50 1Kx1K images



The normalized eight-point algorithm

- In each image
  - center the set of points at the origin
  - scale the shifted points so the mean squared distance between the origin and the points is 2 pixels
- Use the eight-point algorithm to compute $\boldsymbol{F}$ from the normalized points
- Enforce the rank-2 constraint
- Transform fundamental matrix back to original units: if $\boldsymbol{T}$ and $\boldsymbol{T}'$ are the normalizing transformations in the two images,
  then the fundamental matrix in original coordinates is $\boldsymbol{T}'^{T}\boldsymbol{F}\boldsymbol{T}$

R. Hartley. In defense of the eight-point algorithm. TPAMI 1997

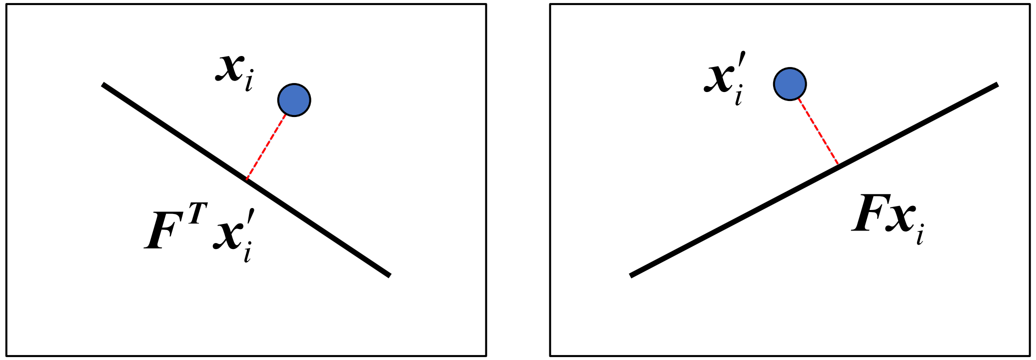
Nonlinear estimation

- Linear estimation minimizes the sum of squared *algebraic* distances between points $\boldsymbol{x}'_i$ and epipolar lines $\boldsymbol{F}\boldsymbol{x}_i$ (or points $\boldsymbol{x}_i$ and epipolar lines $\boldsymbol{F}^T\boldsymbol{x}'_i$):

$$\sum_i \left(\boldsymbol{x}_i'^T \boldsymbol{F} \boldsymbol{x}_i\right)^2$$

- Nonlinear approach: minimize sum of squared *geometric* distances

$$\sum_i \left[\operatorname{dist}(\boldsymbol{x}'_i, \boldsymbol{F}\boldsymbol{x}_i)^2 + \operatorname{dist}(\boldsymbol{x}_i, \boldsymbol{F}^T\boldsymbol{x}'_i)^2\right]$$



Comparison of estimation algorithms

	8-point	Normalized 8-point	Nonlinear least squares
Av. Dist. 1	2.33 pixels	0.92 pixel	0.86 pixel
Av. Dist. 2	2.18 pixels	0.85 pixel	0.80 pixel

Seven-point algorithm

- Set up least squares system with seven pairs of matches and solve for null space (two vectors) using SVD
- Solve for polynomial equation to get coefficients of linear combination of null space vectors that satisfies $\det(\boldsymbol{F}) = 0$

Epipolar geometry & camera calibration

- Estimating the fundamental matrix is known as “weak calibration”
- If we know the calibration matrices of the two cameras, we can estimate the essential matrix: $\boldsymbol{E} = \boldsymbol{K}'^T \boldsymbol{F} \boldsymbol{K}$
- The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters
- Alternatively, if the calibration matrices are known (or in practice, if good initial guesses of the intrinsics are available), the five-point algorithm can be used to estimate relative camera pose

- Note: These methods work well for (relatively) large baselines.
  They do not work well for very small baselines,
  e.g., when using two consecutive images of a video sequence

D. Nister. An efficient solution to the five-point relative pose problem. IEEE Trans. PAMI, 2004