Multiview Geometry

Dr. Elli Angelopoulou

Pattern Recognition Lab (Computer Science 5)

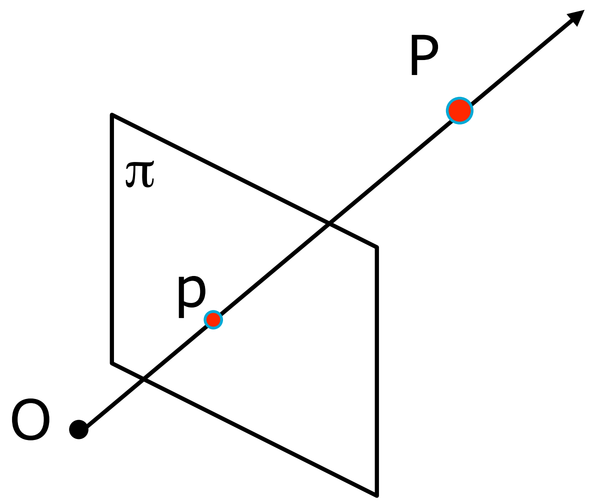
University of Erlangen-Nuremberg

Multiview Analysis

- Observing the same scene point from multiple distinct viewpoints allows the recovery of 3D structure.
- A key component of multiview analysis is finding corresponding scene regions in the different image planes – *the correspondence problem*.
- The relative shift between corresponding projections, *the disparity*, provides 3D structure information.
- Recovery of exact 3D data requires further knowledge about the camera setup.

First Camera

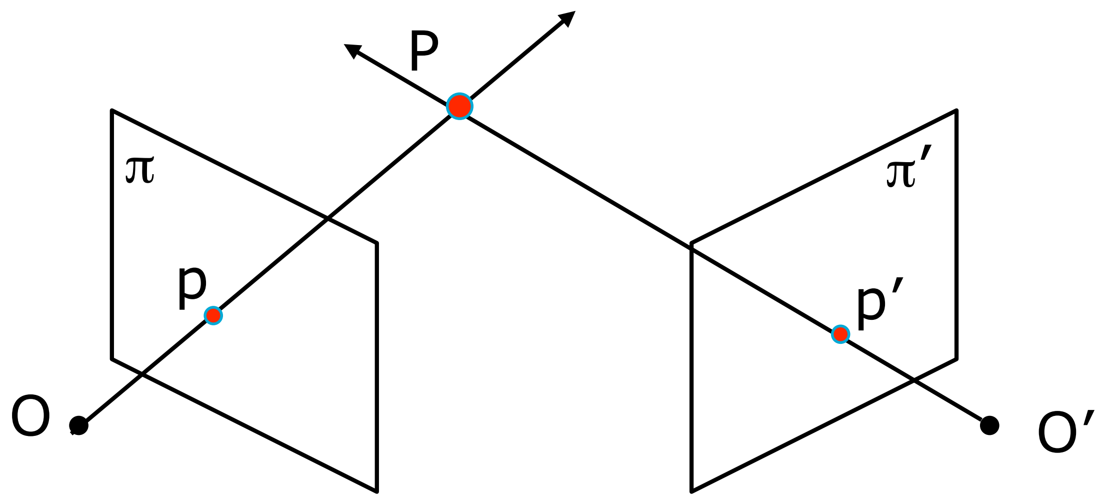
- Camera 1



- Camera 1:
  - Center of Projection *O*
  - Image plane π
  - Scene point *P* projects on point *p* on π.

Second Camera

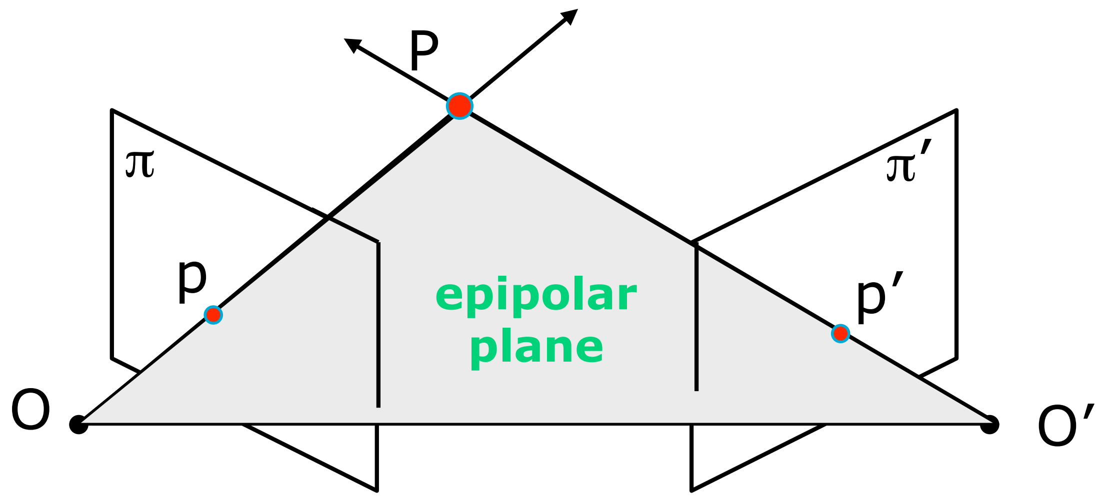
- Camera 2



- Camera 2:
  - Center of Projection O'
  - Image plane π'
  - Scene point P projects on point p' on π'.

Epipolar Plane

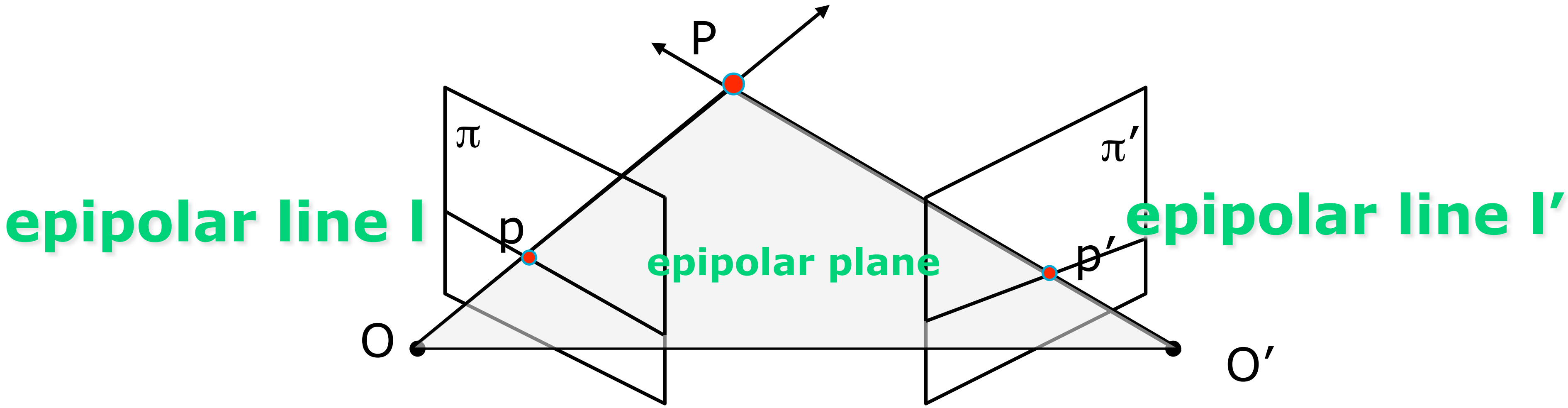
- The epipolar plane is defined by the 2 COPs *O* and *O′* and a point in the scene *P*.



  - The lines *OP* and *O′P* lie on the epipolar plane Γ.
  - Point *p* lies on the *OP* line and on the image plane π. It is the intersection of *OP* and π.
  - Point *p′* lies on the *O′P* line and on the image plane π′. It is the intersection of *O′P* and π′.

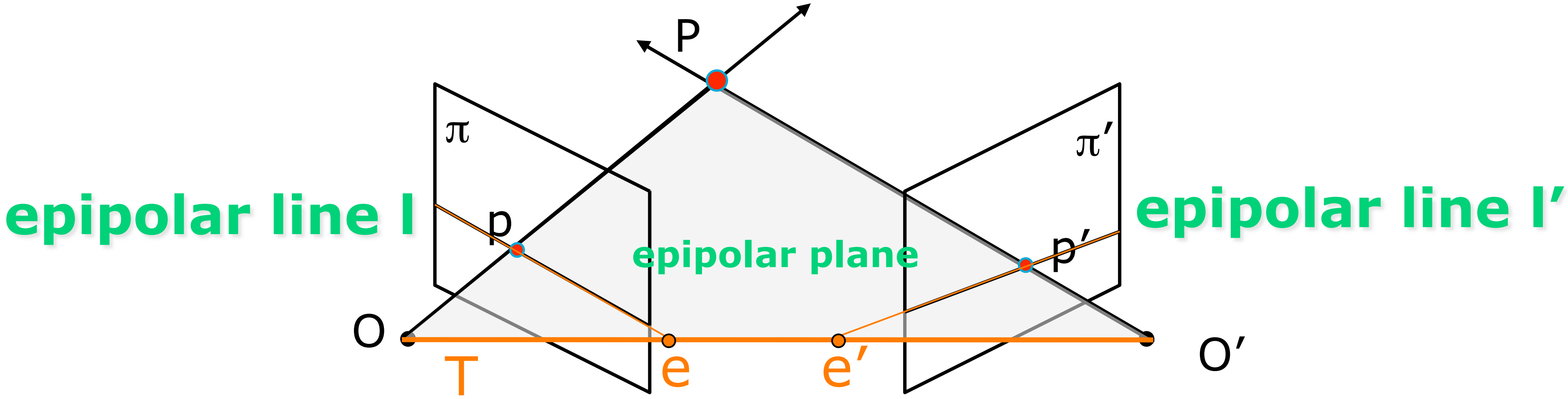
Epipolar Line

- The epipolar line is the intersection of the epipolar plane with the image plane.



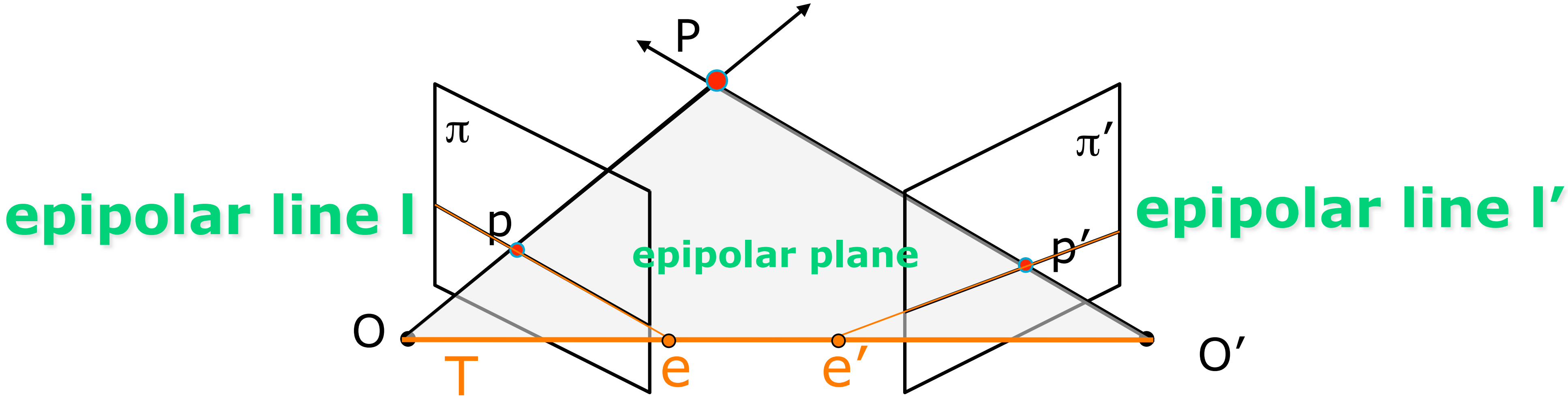
- Since point *p′* lies on the *O′P* line and on the image plane π′, it also lies on the intersection of the epipolar plane with the image plane π′, i.e. on the epipolar line l′
- Since point *p* lies on the *OP* line and on the image plane π, it also lies on the intersection of the epipolar plane with the image plane π, i.e. on the epipolar line *l*.

Epipoles



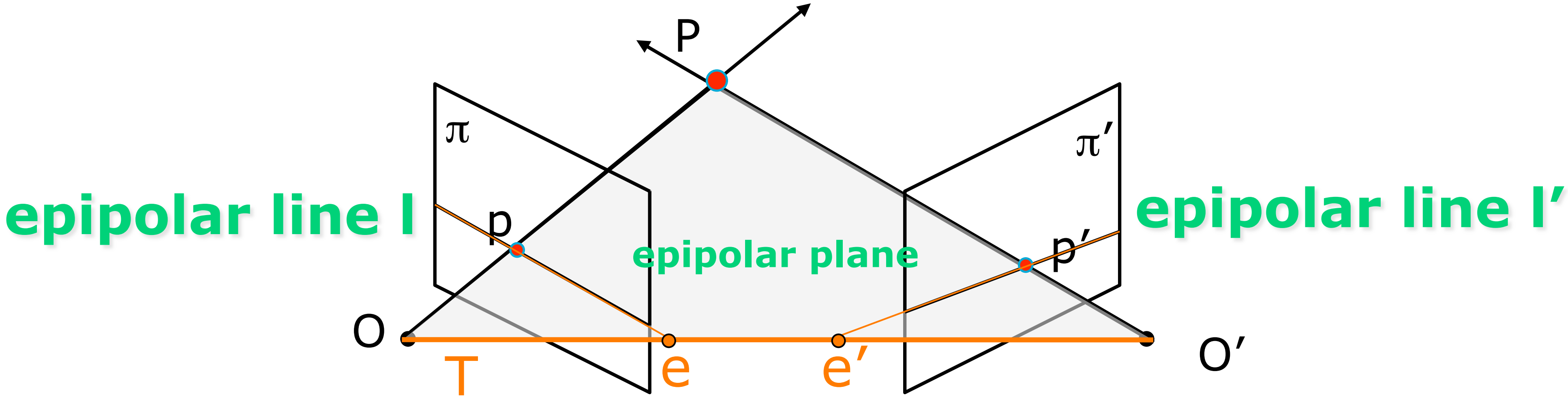
- The baseline T is the line between the 2 COPs O and O'. In verged cameras, this line intersects both plane π and π'.
- The epipole is the intersection of the baseline with the respective image plane.

Epipolar Constraint



- The epipolar line *l* passes through the epipole *e*.
- The epipolar line *l′* passes through the epipole *e′*.
- If both *p* and *p*' are projections of the same point *P*, then *p* and *p′* must lie on the same epipolar plane. They must lie on epipolar lines *l* and *l′* respectively. This is called the **epipolar constraint**.

Impact of the Epipolar Constraint



- The epipolar constraint has a fundamental role in stereo and motion analysis.
- It reduces the correspondence problem to a 1D search along *conjugate epipolar lines.*
- Given an image point *p*, one needs to only search in the epipolar line *l′* for the corresponding point *p′*.

Required Knowledge

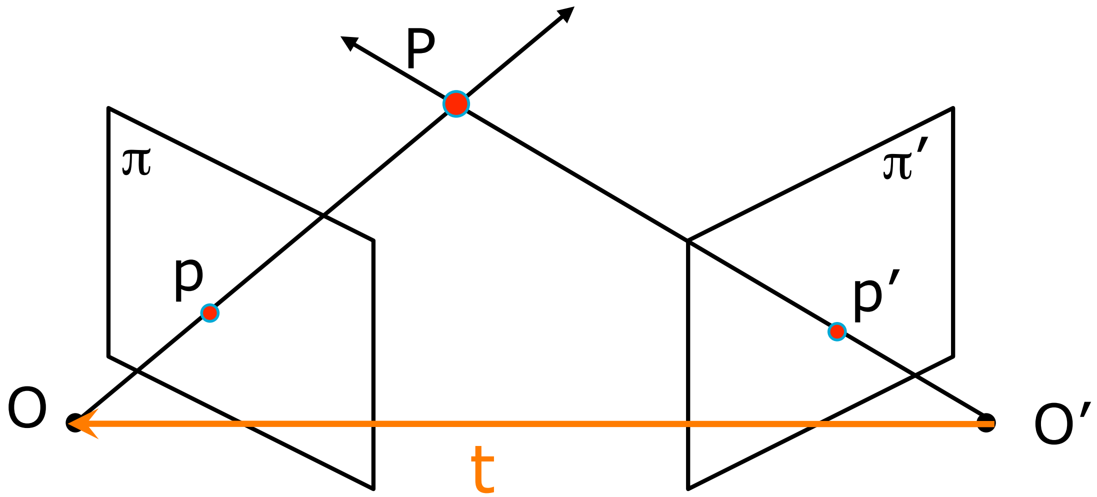
- In order to know the epipolar geometry, we need:
  - The location of the two COPs
  - The location of the two image planes
  - The orientation of the image planes
- We need to know the intrinsic and extrinsic camera characteristics.
- Intrinsic camera characteristics
  - Pixel size
  - Focal length
  - Principal point
- Extrinsic camera characteristics
  - The relative position of the 2 optical centers
  - The relative orientation of the two image planes

Epipolar Constraint – Calibrated Case

- Assume that the intrinsic parameters of each of the cameras are known, i.e. the mapping from the image coordinate system to a metric camera coordinate system.

- Goal: Express algebraically the epipolar constraint, so that it can be incorporated in our correspondence, stereo and motion algorithms.

Epipolar Plane Constraint



- The vectors *Op*, *O′p′* and *O′O* are all co-planar, i.e. they must satisfy the following equation:

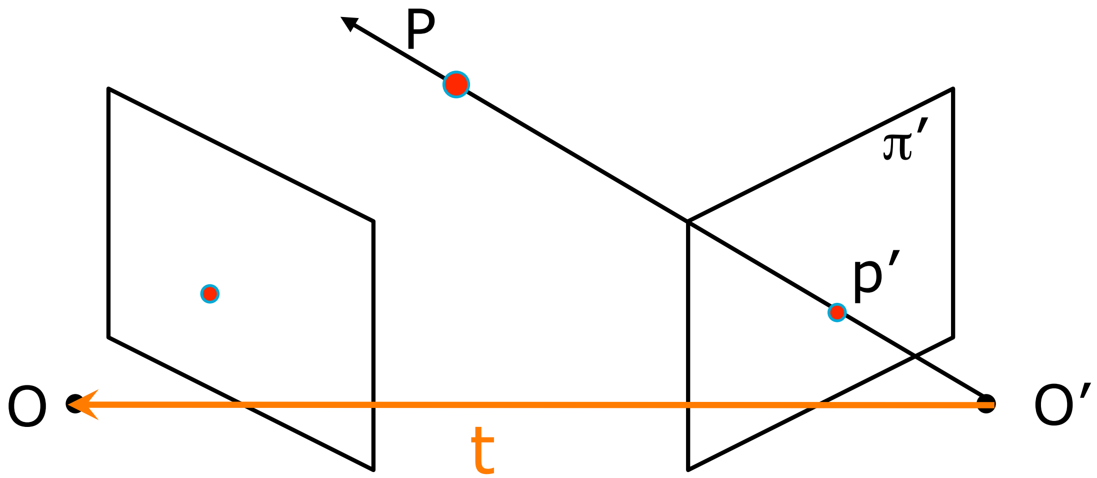
$$\overrightarrow{Op} \cdot (\overrightarrow{O'O} \times \overrightarrow{O'p'}) = 0$$

- The vector *Op* is perpendicular to the vector resulting from the cross-product of *O′O* and *O′p′*.

Relating the 2 Camera Coord. Systems

- Each image is unaware of the other camera.
- Point *p* is specified in the local coordinate system of the camera with COP *O*.
- Similarly point *p′* is specified in the local coordinate system of the camera with COP *O′*.
- We need to express everything in terms of a single coordinate system.
- Without loss of generality we choose as the reference coordinate system the one of the camera with COP *O*.

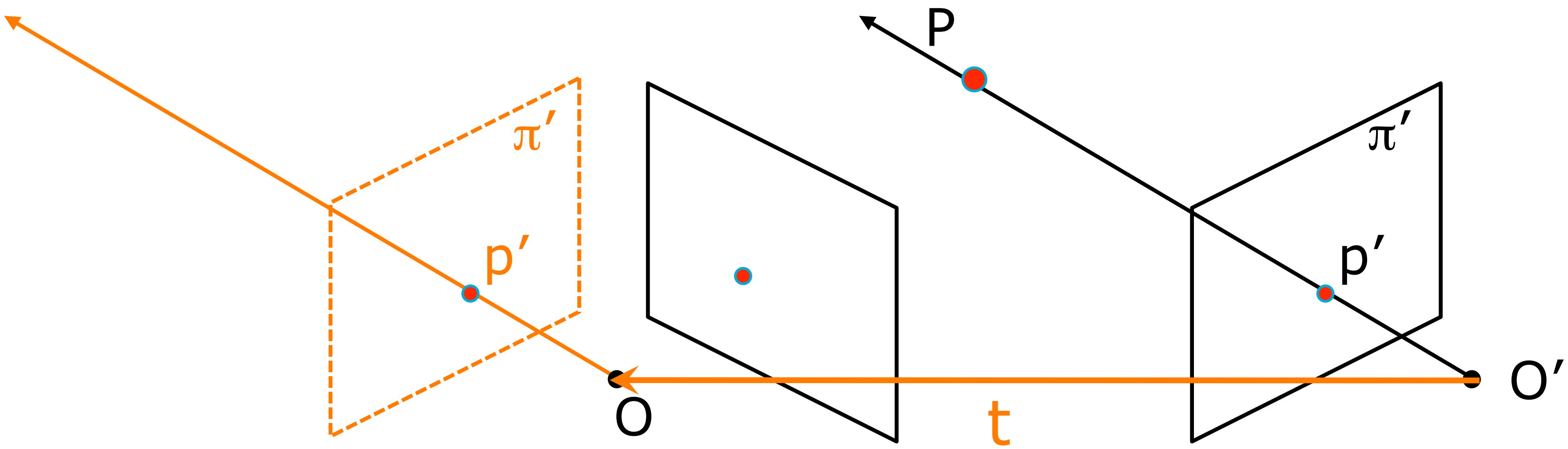
Translation



- There is a translation vector t, (the baseline T to be precise) that shows you how one can move COP O′ to COP *O*.

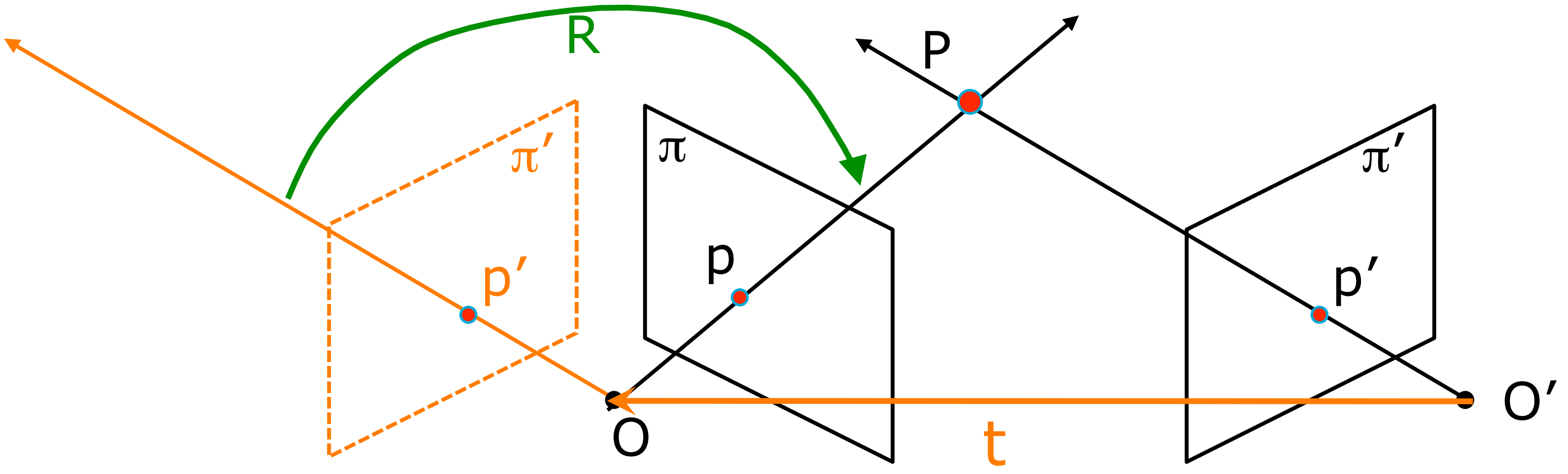
$$\vec{t} = \overrightarrow{O'O}$$

Need for Rotation



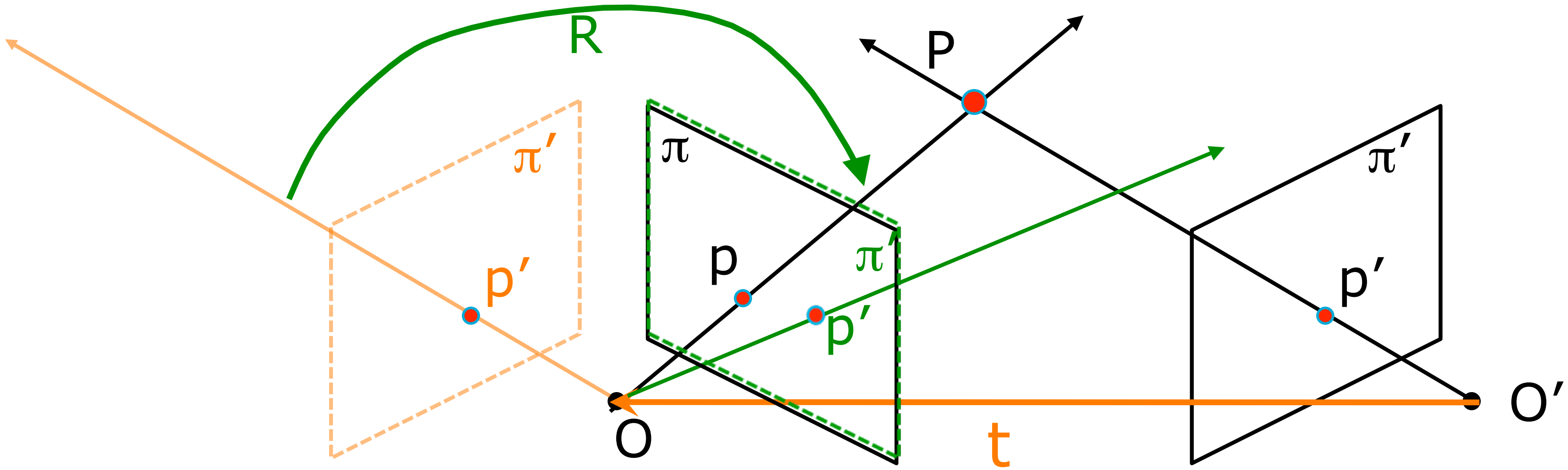
- If we apply this translation *t* to every point *p′* of the camera with COP *O′* then we will move the coordinate system with COP *O*' so that both camera coordinates are pinned to the same origin *O*.

Rotation



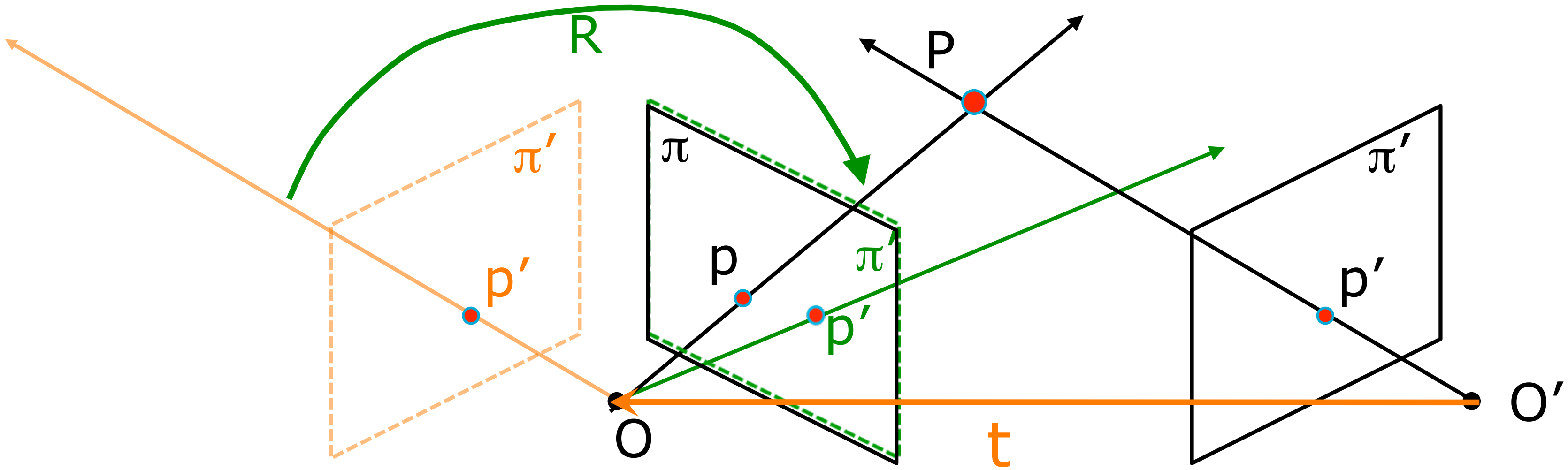
- Still the two coordinate systems can differ by a rotation. Let *R* be the rotation matrix that aligns the corresponding axes of the two camera coordinates.

Translation and Rotation



- Each point *p*' after the translation from camera *O*′ to camera *O*, is rotated by R.
- The two camera coordinate systems are now aligned.
- Everything can be expressed in terms of the coordinate system of camera *O*.

Epipolar Constraint Revisited



- Recall that vectors *Op*, *O'p'* and *O′O* are co-planar:

$$\overrightarrow{Op} \cdot (\overrightarrow{O'O} \times \overrightarrow{O'p'}) = 0$$

- Rewritten in the coordinate frame of camera *O*:

$$\vec{p} \cdot (\vec{t} \times \overrightarrow{(Rp')}) = 0$$

Epipolar Constraint – Matrix Form

- The epipolar equation can be rewritten as a series of matrix multiplications:

$$\mathbf{p}^T(\mathbf{t} \times \mathbf{R})\mathbf{p}' = 0$$

- This is often represented more compactly as:

$$\mathbf{p}^T \mathbf{E} \mathbf{p}' = 0$$

where $\mathbf{E}$ is a 3x3 matrix of the form: $\mathbf{E} = [\mathbf{t}_\times]\mathbf{R}$

and it is known as the *essential matrix*.

$[\mathbf{t}_\times]$ is a skew-symmetric matrix such that $[\mathbf{t}_\times]\mathbf{b} = \mathbf{t} \times \mathbf{b}$

$[\mathbf{t}_\times]$ is the matrix representation of the cross product with $\mathbf{t}$.

if $\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$ then $[\mathbf{t}_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$

Epipolar Constraint Equations

- The equation $\mathbf{p}^T\mathbf{E}\mathbf{p}' = 0$ is the algebraic representation of epipolar constraint.
- The vector that corresponds to the epipolar line *l* that is associated with point *p′* is $\mathbf{l} = \mathbf{E}\mathbf{p}'$.
- Similarly, the vector that corresponds to the epipolar line *l′* that is associated with point *p* is $\mathbf{l}' = \mathbf{E}^T\mathbf{p}$.
- Thus, once the essential matrix **E** is recovered, one can reduce the search space for finding the corresponding points to a 1D space.

Epipolar Constraint –Uncalibrated case

- For uncalibrated cases, the matrices (rotation **R** and translation **t**) that express point *p*' in terms of the coordinate system of camera *O* must also incorporate the intrinsic camera parameters.
- Instead of $\mathbf{p}^T\mathbf{E}\mathbf{p}' = 0$ we have:

$$\mathbf{p}^T\mathbf{K}^{-T}\mathbf{E}\mathbf{K}'^{-1}\mathbf{p}' = 0$$

$$\mathbf{p}^T\mathbf{F}\mathbf{p}' = 0$$

where $\mathbf{F} = \mathbf{K}^{-T}\mathbf{E}\mathbf{K}'^{-1}$ and $\mathbf{K}$ and $\mathbf{K}'$ are the intrinsic parameter matrices of cameras *O* and *O*' accordingly

- **F** is called the *fundamental matrix*.

Multiple Views

- For binocular setups the epipolar constraint can be represented in a 3x3 matrix form, called the *fundamental matrix*.
- When we have 3 images the epipolar constraint is represented by a 3x3x3 structure, called the *trifocal tensor*.
- When we have 4 images the epipolar constraint is represented by a 3x3x3x3 structure, called the *quadrifocal tensor*.

Key Points of Epipolar Geometry

- For each pair of corresponding points *p* and *p*' in camera coordinates (Cartesian metric coordinate. system), the following relationship holds:

$$\mathbf{p}^T \mathbf{E} \mathbf{p}' = 0$$

E is the essential matrix

- For each pair of corresponding points *q* and *q*' in pixel (image) coordinates the following relationship holds:

$$\mathbf{q}^T \mathbf{F} \mathbf{q}' = 0$$

F is the fundamental matrix

Key Points of Epipolar Geometry 2

- The epipolar line *l*' that corresponds to the point *q* has the form $l'_1 x + l'_2 y + l'_3 z = 0$, where $\mathbf{l}' = (l'_1, l'_2, l'_3)$ and is given by:

  $$\mathbf{l}' = \mathbf{F}^T \mathbf{q}$$

  where x,y,z are in the local coordinate system of camera *O'*.

- The epipolar line *l* that corresponds to the point *q'* has the form $l_1 x + l_2 y + l_3 z = 0$, where $\mathbf{l} = (l_1, l_2, l_3)$ and is given by:

  $$\mathbf{l} = \mathbf{F}\mathbf{q}'$$

  where x,y,z are in the local coordinate system of camera *O*.

The Essential Matrix in Practice

- What does the epipolar plane depend on? A point P in the scene and the camera COPs O and O'. It varies from point to point.
- What does the matrix **E** (similarly **F**) depend on? The rotation **R** and the translation **t** between the two camera coordinate systems. No dependence on the scene.
- So… recover **E** (or **F**) once, keep the camera setup stable and then reuse it for every scene point.
- How do we recover **E** (or **F**)?

Estimation of the Fundamental Matrix.

- Assume known correspondences of *n* points between the two images.
- You have *n* equations of the form:

$$\mathbf{p}_i^T \mathbf{F} \mathbf{p}_i' = 0 \ , \quad i = 1 \ldots n$$

- **F** is a 3x3 matrix => 9 unknowns.
- If you have 8 well spread correspondences, you can determine **F**.
- Why 8? The *n* equations are homogeneous linear equations, i.e. all equations have a zero as a constant in the right hand side. So the solution is unique up to a scaling factor.

Over-determined System

- If n>8, then we have an over-determined system. Use SVD (Singular Value Decomposition).
- How? Build a nx9 matrix **A** which contains the coefficients of the n equations: $\mathbf{p}_i^T \mathbf{F} \mathbf{p}_i' = 0 \ , \quad i = 1 \ldots n$
- Run SVD on **A**. It decomposes **A** to: $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$
  - D diagonal matrix; its elements are called singular values.
  - U is an n x n orthogonal matrix
  - D is an n x 9 diagonal matrix
  - V is a 9 x 9 orthogonal matrix
- In theory, the solution to **F** (the value of its 9 unknowns) is the column of **V** that corresponds to the only *null* singular value of **A**, i.e. the only zero value on the diagonal.

Estimating **F** in Practice

- In reality, due to noise, quantization, numerical errors, inaccuracies in the *n* correspondences, there is usually no null singular value.
- Thus, in practice we use the *minimum* singular value and its corresponding column in **V**.

$$\mathbf{F} = \mathbf{V}(\mathrm{Col}_m)$$

where s_m was the minimum diagonal value in **D** and was located in column *m* in D.

Estimating **F** in Practice - continued

- However, this whole process had inaccuracies. The resulting **F** may not be singular. So, run SVD again, this time on **F**.

$$\mathbf{F} = \mathbf{U}_F \mathbf{D}_F \mathbf{V}_F^{\ T}$$

- Then build the matrix $\mathbf{D}'$ from $\mathbf{D}_F$ where with the minimum singular value s_m of $\mathbf{D}_F$ is replaced by 0.
- Compute a new fundamental matrix which is singular:

$$\mathbf{F}' = \mathbf{U}_F \mathbf{D}' \mathbf{V}_F^{\ T}$$

- **F**' is a good estimate of the fundamental matrix.

Longuet-Higgins Eight-Point Algorithm

1. Let **A** be an *n*x9 matrix of the coefficients of the *n* eqs.:

$$\mathbf{p}_i^T \mathbf{F} \mathbf{p}_i' = 0 \ , \quad i = 1 \ldots n$$

2. Apply SVD on **A** and find matrices **U**, **D**, **V** such that

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$$

3. The entries of **F** are the components of the column of **V** corresponding to the least singular value of A.
4. Enforce the singularity constraint by applying SVD on **F**

$$\mathbf{F} = \mathrm{U}_F \mathbf{D}_F \mathbf{V}_F^{\ T}$$

5. and creating **D'** = $\mathbf{D}_F$ *with* the smallest singular value of $\mathbf{D}_F$ replaced by 0.
6. Get new estimate of **F**, call it **F'**, such that

$$\mathbf{F}' = \mathrm{U}_F \mathbf{D}' \mathbf{V}_F^{\ T}$$

Fundamental Matrix Video

The video is courtesy of Daniel Wedge. You can view it at the following web-site: http://danielwedge.com/fmatrix/