MAI4CAREU
Master programmes in Artificial Intelligence 4 Careers in Europe

Visual Computing Group

CYENS CENTRE OF EXCELLENCE

University of Cyprus – MSc Artificial Intelligence

MAI644 – COMPUTER VISION

Lecture 16: Stereo Vision

Melinos Averkiou

CYENS Centre of Excellence

University of Cyprus - Department of Computer Science

m.averkiou@cyens.org.cy

CYENS CENTRE OF EXCELLENCE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Last time

- Full camera model in matrix form
- Camera calibration
- Calibration – Projective camera model
- Calibration – Affine camera model

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Recovery of world position
- Triangulation
- Epipolar Geometry

[material based on Paris Kaimakis]

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today’s Agenda

- Recovery of world position
- Triangulation
- Epipolar Geometry

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Recovery of world position

- Previously we saw that the imaging process can be described as a transformation in homogeneous coordinates.

- If we can invert this transformation, the world coordinates of each pixel in the image can be computed.

- Recovering world coordinates of objects based on the projection on an image is known as **shape recovery** or **depth recovery**.

Recovery of world position

- Is this possible using a single camera ?
- The camera needs to be calibrated, i.e. we know all its parameters
- Remember the projective camera model:

$$\tilde{\boldsymbol{w}} = \boldsymbol{P}\tilde{\boldsymbol{X}}$$

$$\Leftrightarrow \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Unfortunately the transformation described in $\boldsymbol{P}$ is not invertible and the world point $\boldsymbol{X}$ cannot be uniquely determined

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

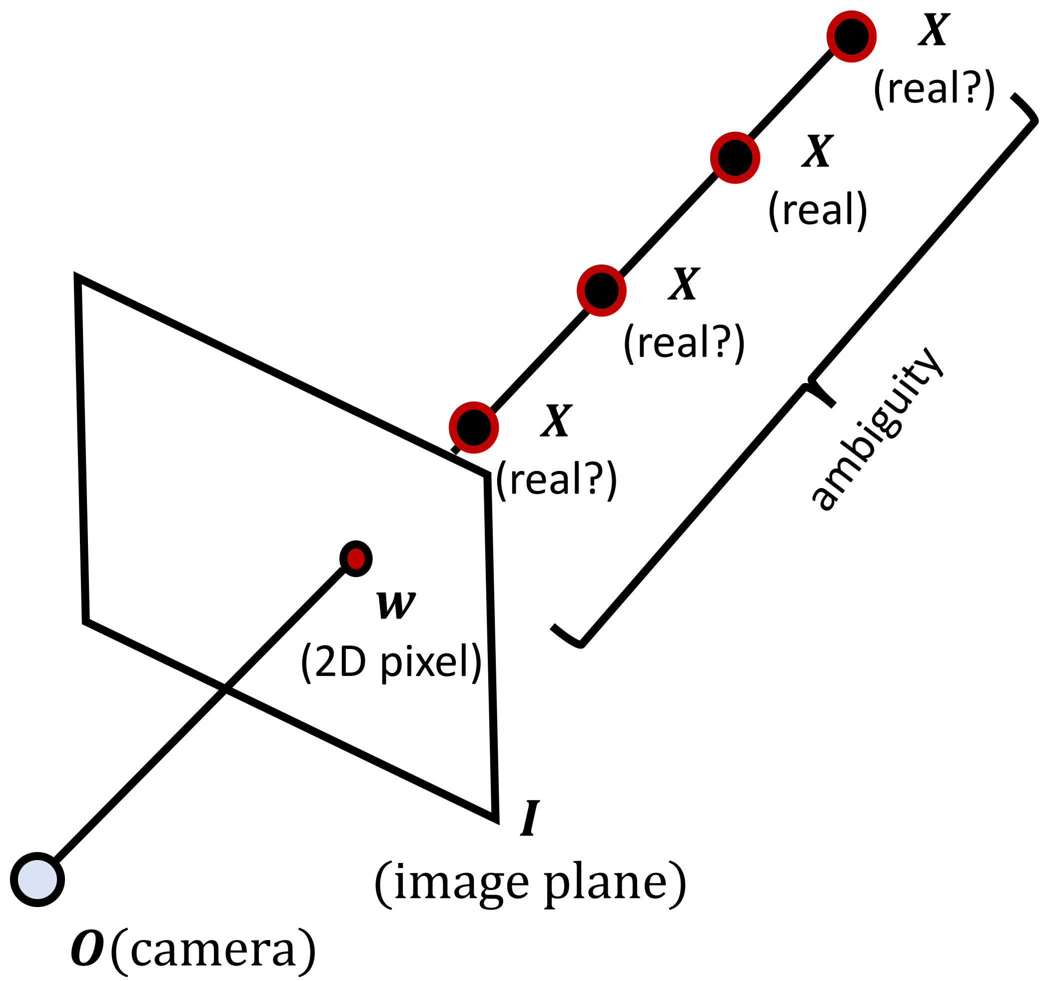
Recovery of world position

- Depth ambiguity

Courtesy slide S. Lazebnik

Recovery of world position

- Each observed feature on the image gives 2 equations with 3 unknowns and therefore defines a line (a ray) of solutions for $\boldsymbol{X}$



Recovery of world position

- Each observed feature on the image gives 2 equations with 3 unknowns and therefore defines a line (a ray) of solutions for $\boldsymbol{X}$
- This system of equations is under-constrained.
- This can be seen by the size of $\boldsymbol{P}$. There are more columns than rows.

$$\tilde{\boldsymbol{w}} = \boldsymbol{P}\tilde{\boldsymbol{X}}$$

$$\Leftrightarrow \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Recovery of world position

- Under-constrained problems never have a unique solution.

- To uniquely recover $\boldsymbol{X}$, additional views must be used, so that the transformation between $\boldsymbol{w}$ (pixel coordinates) and $\boldsymbol{X}$ (world coordinates) is *forced* to become invertible.

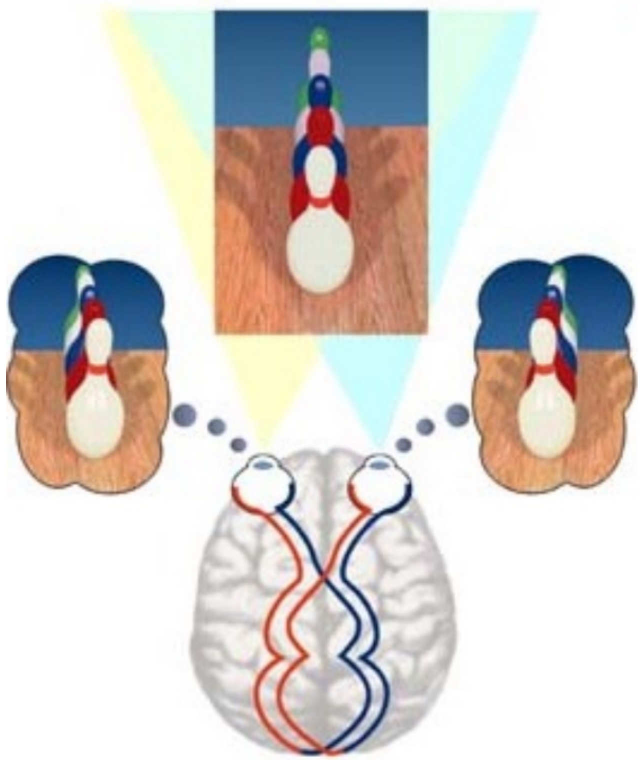
- This is the subject of **stereo vision**.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Recovery of world position

- **Two** eyes/**cameras** help.

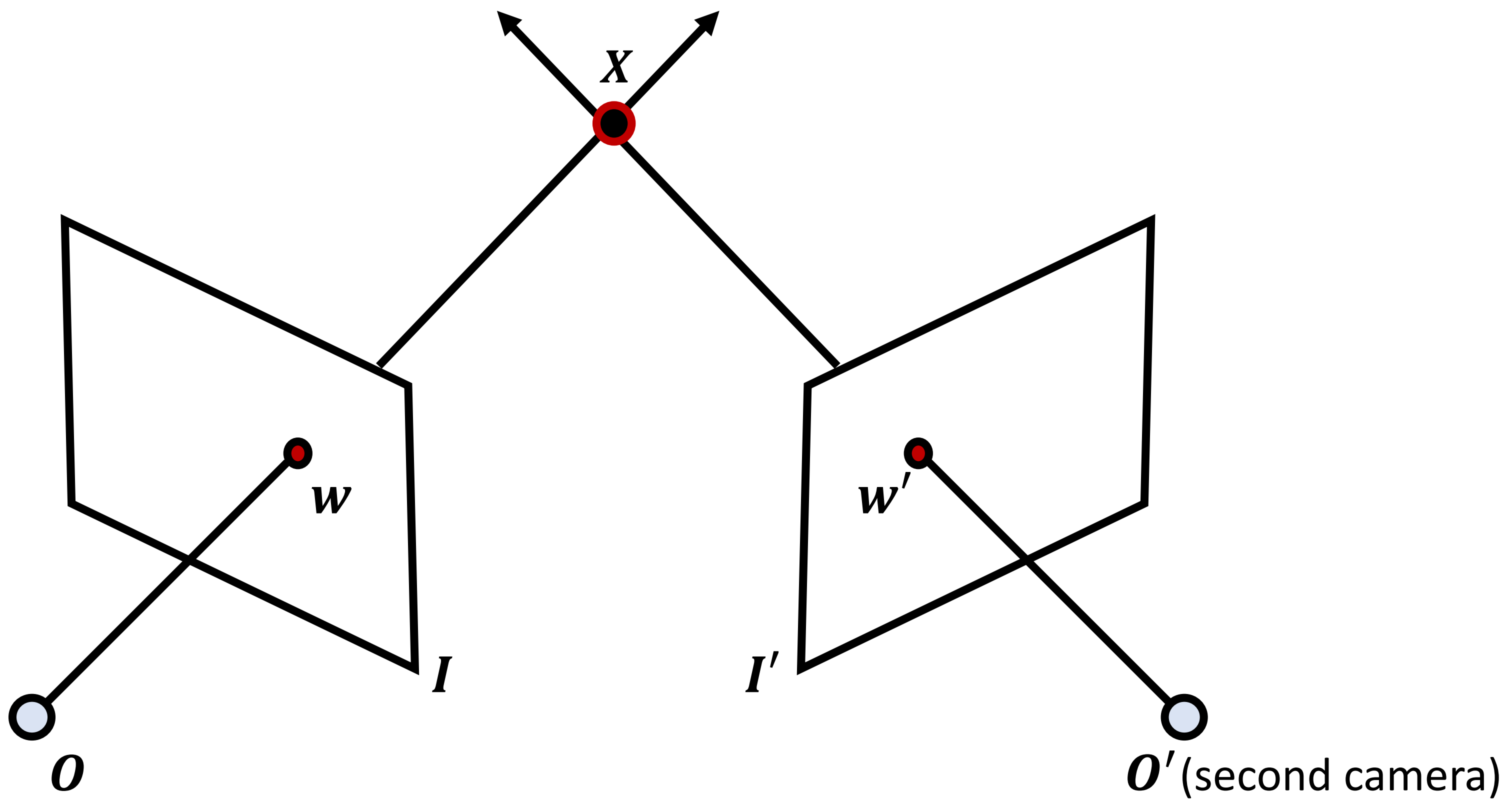


Today’s Agenda

- Recovery of world position
- Triangulation
- Epipolar Geometry

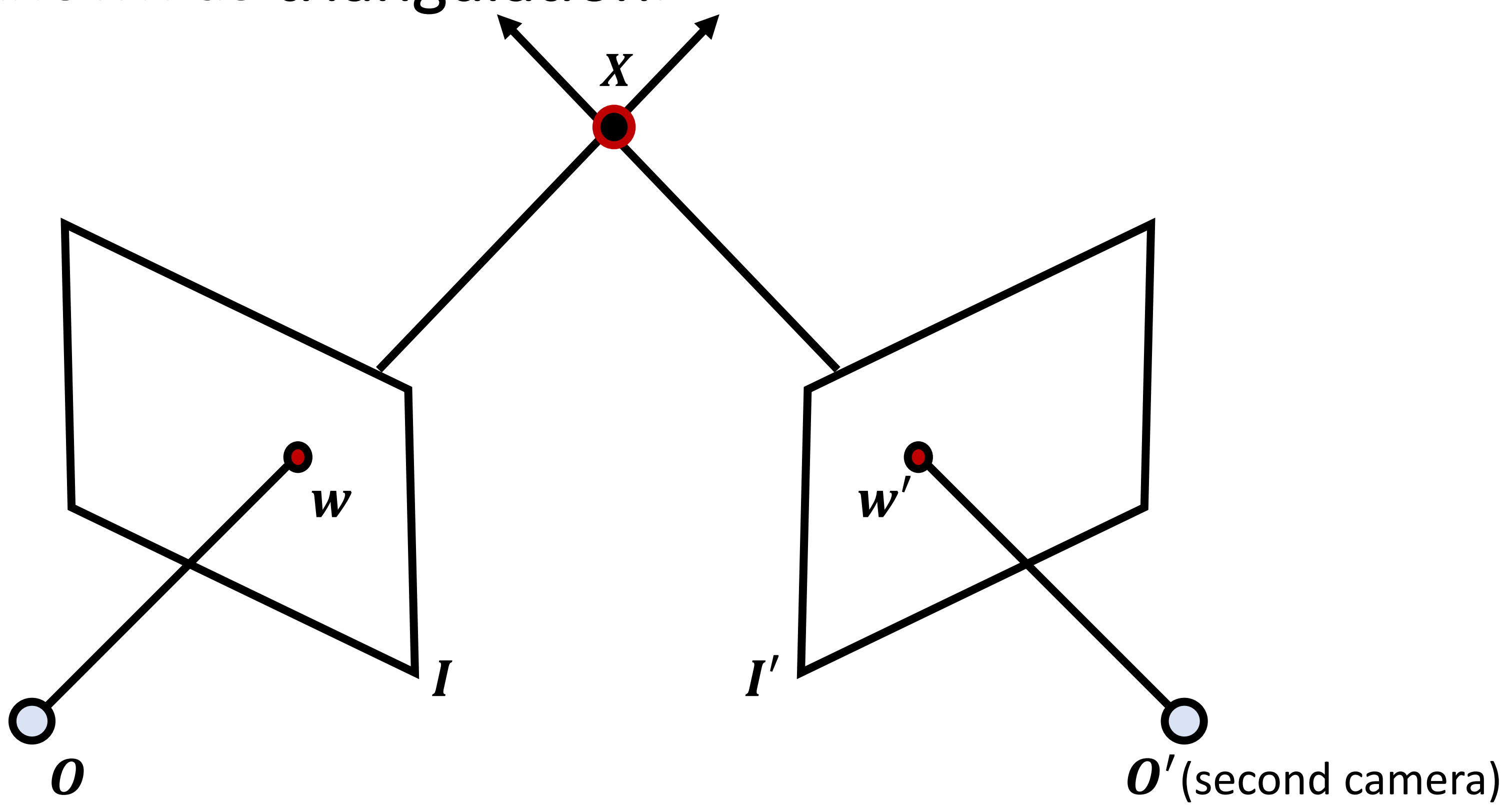
Triangulation

- In stereo vision at least two cameras are set up to view the $3D$ scene.
- Each $3D$ world location $\boldsymbol{X}$ projects to pixel $\boldsymbol{w}$ on camera 1 ($\boldsymbol{O}$) and to pixel $\boldsymbol{w}$' on camera 2 ($\boldsymbol{O}'$).



Triangulation

- If both cameras are calibrated, the $3D$ world location $\boldsymbol{X}$ projected on the pair of corresponding pixel locations $\boldsymbol{w}$ and $\boldsymbol{w}$' can be estimated via a process known as **triangulation**.



Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Consider the projection of $\boldsymbol{X}$ onto $\boldsymbol{w}$:

$$\tilde{\boldsymbol{w}} = \boldsymbol{P}\tilde{\boldsymbol{X}}$$

$$\Leftrightarrow \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Compute the equation for u and rearrange so our unknowns X, Y, Z are on the left:

$$u = \frac{su}{s} = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

$$\Rightarrow p_{11}X + p_{12}Y + p_{13}Z + p_{14} = p_{31}uX + p_{32}uY + p_{33}uZ + p_{34}u$$

$$\Rightarrow (p_{11} - p_{31}u)X + (p_{12} - p_{32}u)Y + (p_{13} - p_{33}u)Z = p_{34}u - p_{14}$$

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Compute the equation for u and rearrange so our unknowns X, Y, Z are on the left:

$$u = \frac{su}{s} = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

$$\Rightarrow p_{11}X + p_{12}Y + p_{13}Z + p_{14} = p_{31}uX + p_{32}uY + p_{33}uZ + p_{34}u$$

$$\Rightarrow (p_{11} - p_{31}u)X + (p_{12} - p_{32}u)Y + (p_{13} - p_{33}u)Z = p_{34}u - p_{14}$$

- Put this in matrix form:

$$\begin{bmatrix} p_{11} - p_{31}u & p_{12} - p_{32}u & p_{13} - p_{33}u \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{34}u - p_{14} \end{bmatrix}$$

$$\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{b}$$

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Put this in matrix form:

$$[p_{11} - p_{31}u \quad p_{12} - p_{32}u \quad p_{13} - p_{33}u] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [p_{34}u - p_{14}]$$
$$\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{b}$$

- $\boldsymbol{A}$ is a $1x3$ matrix

- The resulting system is under-constrained

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Put this in matrix form:

$$[p_{11} - p_{31}u \quad p_{12} - p_{32}u \quad p_{13} - p_{33}u] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [p_{34}u - p_{14}]$$

$$[p_{21} - p_{31}v \quad p_{22} - p_{32}v \quad p_{23} - p_{33}v] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [p_{34}v - p_{24}]$$

$$\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{b}$$

- By computing $\boldsymbol{v}$ in the same way as $\boldsymbol{u}$ and rearranging we can add a new row in $\boldsymbol{A}$ and in $\boldsymbol{b}$, making it $2x3$

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Put this in matrix form:

$$[p'_{11} - p'_{31}u' \quad p'_{12} - p'_{32}u' \quad p'_{13} - p'_{33}u'] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [p'_{34}u' - p'_{14}]$$

$$[p'_{21} - p'_{31}v' \quad p'_{22} - p'_{32}v' \quad p'_{23} - p'_{33}v'] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [p'_{34}v' - p'_{24}]$$

$$\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{b}$$

- By also considering the projection of $\boldsymbol{X}$ onto $\boldsymbol{w}'$, a new pair of rows will be added to $\boldsymbol{A}$, thus forcing it to be $4x3$, i.e. over-constrained

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Here is the resulting system

$$\begin{bmatrix} p_{11} - p_{31}u & p_{12} - p_{32}u & p_{13} - p_{33}u \\ p_{21} - p_{31}v & p_{22} - p_{32}v & p_{23} - p_{33}v \\ p'_{11} - p'_{31}u' & p'_{12} - p'_{32}u' & p'_{13} - p'_{33}u' \\ p'_{21} - p'_{31}v' & p'_{22} - p'_{32}v' & p'_{23} - p'_{33}v' \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{34}u - p_{14} \\ p_{34}v - p_{24} \\ p'_{34}u' - p'_{14} \\ p'_{34}v' - p'_{24} \end{bmatrix}$$

- Where p'_{11} etc. are the parameters inside the camera projection matrix for camera 2 ($\boldsymbol{O}'$), and $\boldsymbol{w}' = (u', v')$ are the pixel coordinates of $\boldsymbol{X}$ projected on the image plane $\boldsymbol{I}'$ of the second camera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Triangulation

- Using an additional camera has forced $\boldsymbol{A}$ to become over-constrained.

$$\begin{bmatrix} p_{11} - p_{31}u & p_{12} - p_{32}u & p_{13} - p_{33}u \\ p_{21} - p_{31}v & p_{22} - p_{32}v & p_{23} - p_{33}v \\ p'_{11} - p'_{31}u' & p'_{12} - p'_{32}u' & p'_{13} - p'_{33}u' \\ p'_{21} - p'_{31}v' & p'_{22} - p'_{32}v' & p'_{23} - p'_{33}v' \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{34}u - p_{14} \\ p_{34}v - p_{24} \\ p'_{34}u' - p'_{14} \\ p'_{34}v' - p'_{24} \end{bmatrix}$$

- Therefore we can now find a least-squares solution:

$$\boldsymbol{Ax} = \boldsymbol{b}$$
$$\Leftrightarrow \boldsymbol{x} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

Co-financed by the European Union
Connecting Europe Facility

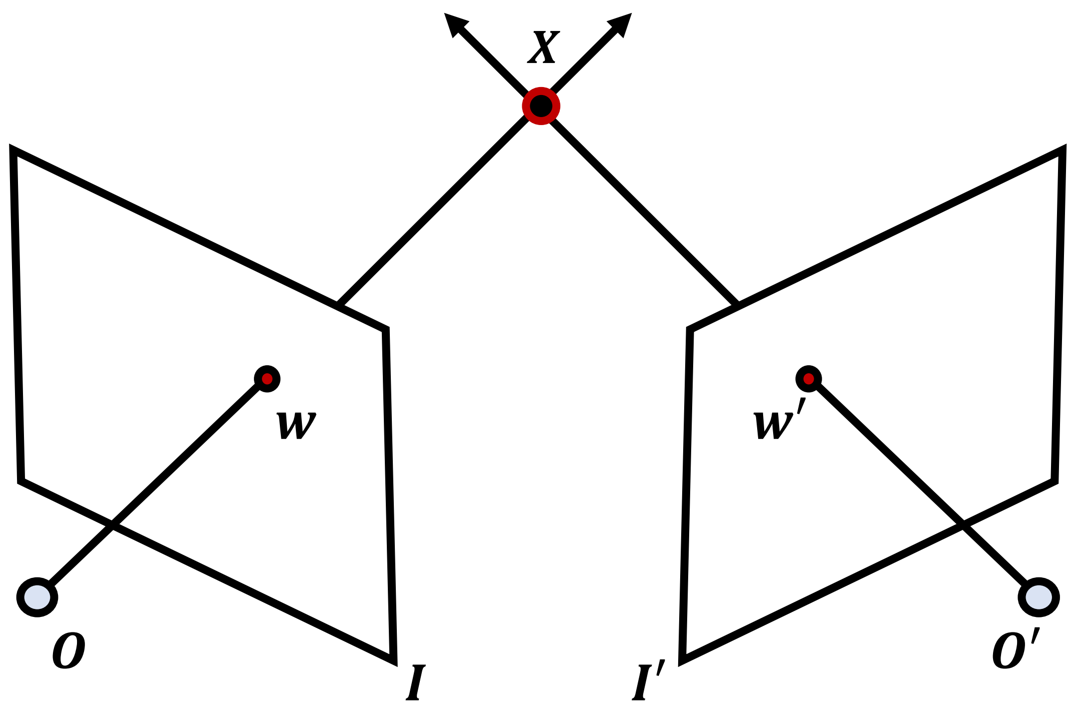
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Recovery of world position
- Triangulation
- Epipolar Geometry

Beyond triangulation

We have seen the simplest form of stereo vision: Given a pair of *calibrated* cameras observing a single feature at *corresponding* pixel locations $\boldsymbol{w} \leftrightarrow \boldsymbol{w}'$, the $3D$ position of the corresponding world location $\boldsymbol{X}$ can be estimated via triangulation



Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Beyond triangulation

How is the correspondence problem solved if there are several points $\{\boldsymbol{w_i}\}_{i=1}^{N_1}$ in image 1 and several points $\{\boldsymbol{w'}_{\boldsymbol{j}}\}_{j=1}^{N_2}$ in image 2 ?

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Beyond triangulation

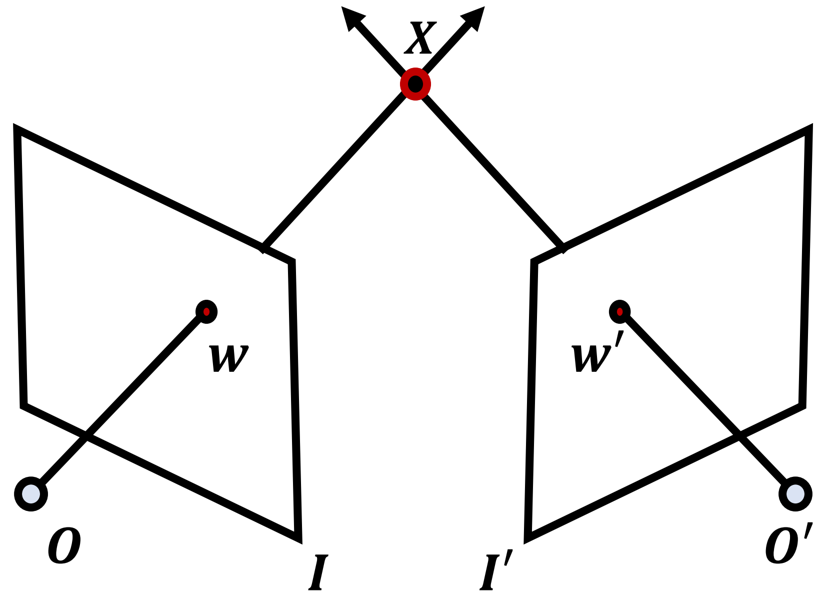
SIFT will give us a set of proposed correspondences $\{\boldsymbol{w}_i \leftrightarrow \boldsymbol{w'}_j\}$ but there will be **many** outliers in these proposals

The question is then how can we remove outliers ?

Using the **epipolar constraint**

Epipolar Geometry

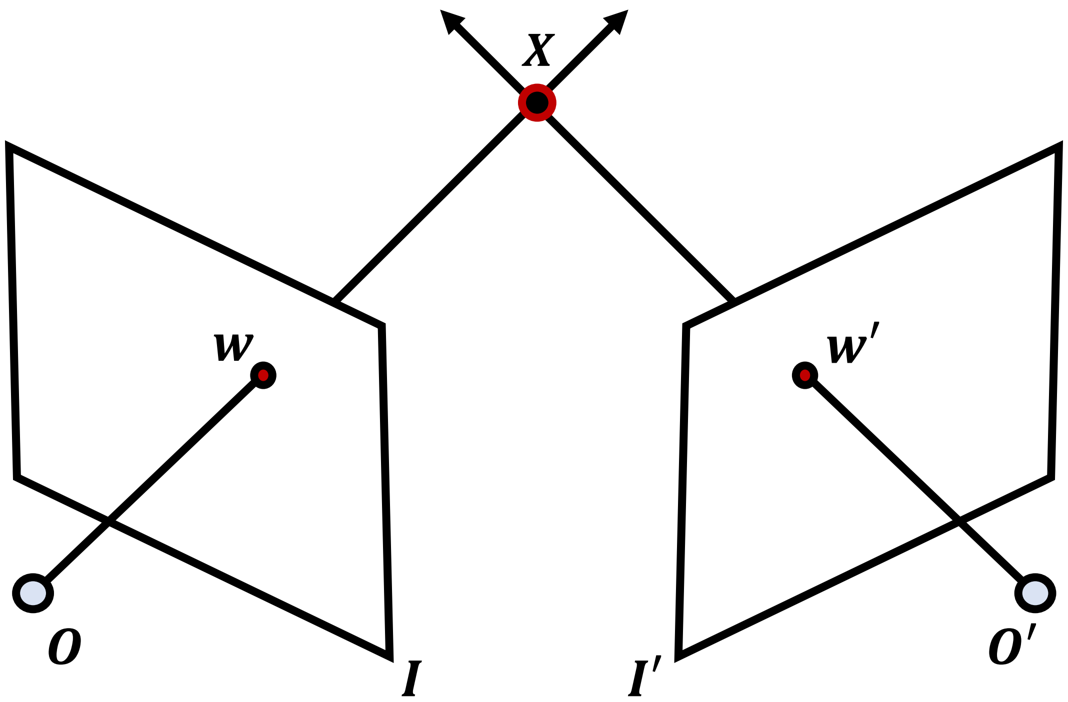
- To understand the **epipolar constraint**, we first need to understand the **geometry** that relates the
  - **cameras**
  - **points in $\mathbf{3D}$** space
  - and their corresponding observations $\{\mathbf{w}_i \leftrightarrow \mathbf{w}'_j\}$



- This type of geometry is referred to as the **epipolar geometry**

Epipolar Geometry - fundamentals

Lets revisit our **stereo pair**

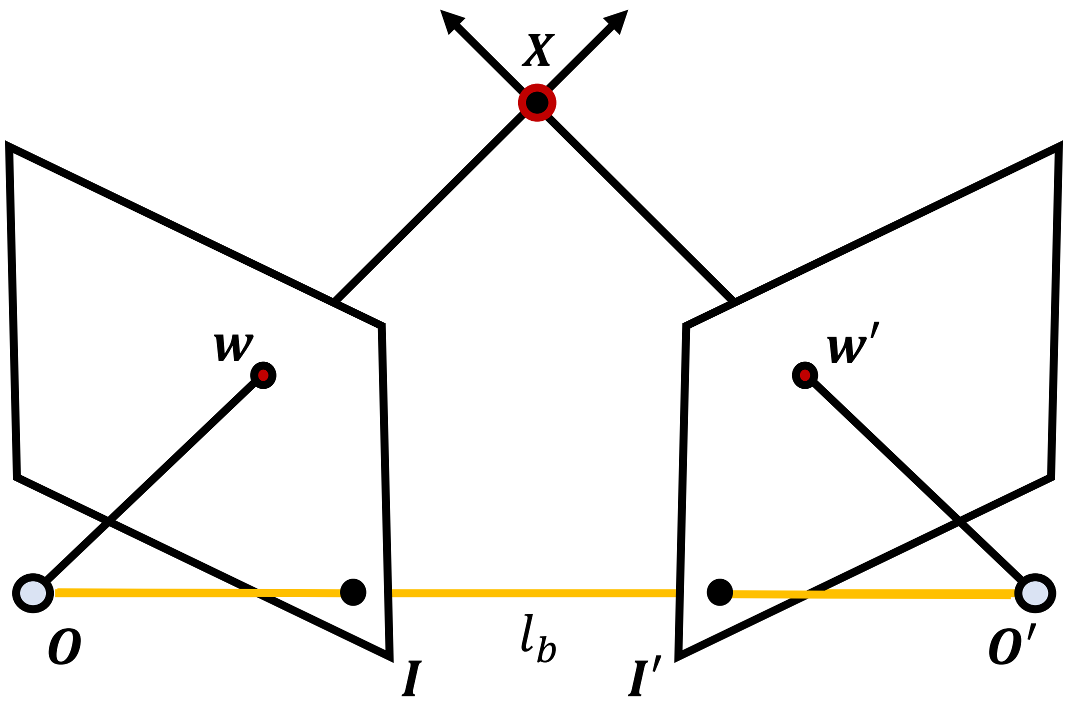


Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

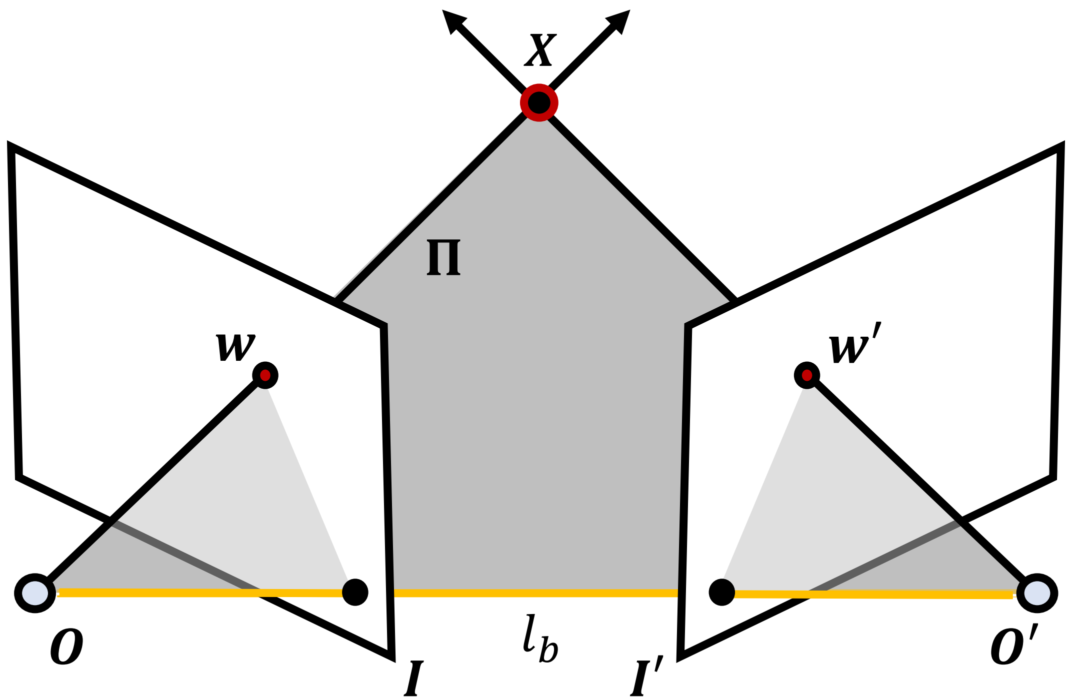
Epipolar Geometry - fundamentals

The baseline l_b is the line joining the two optical centers.



Epipolar Geometry - fundamentals

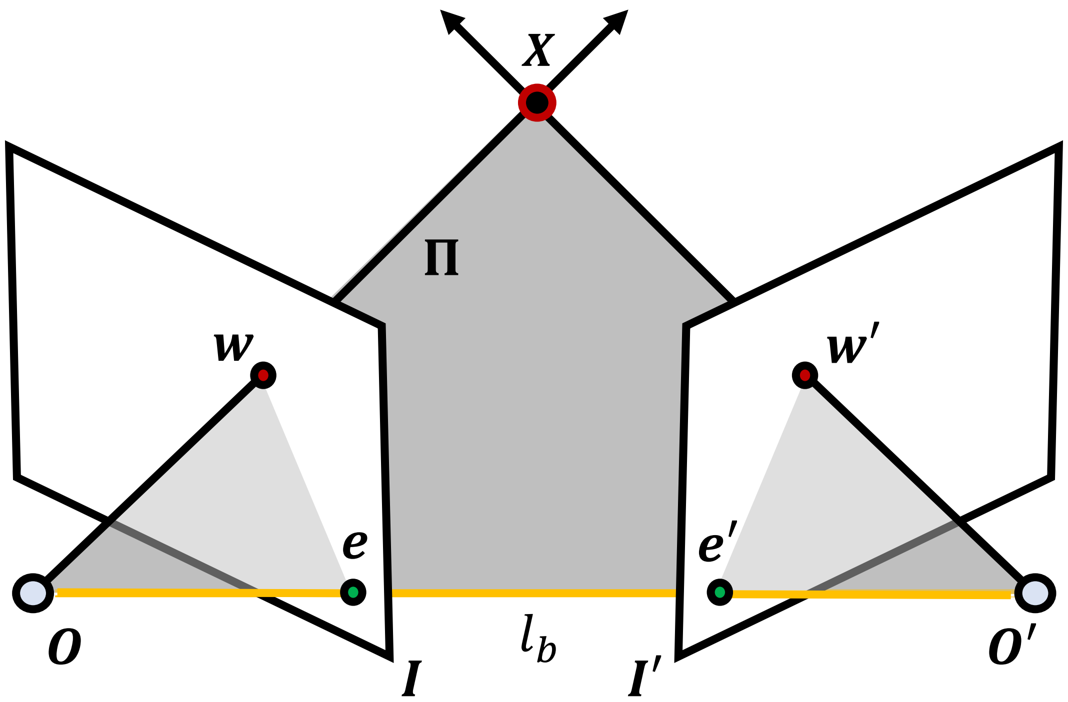
The epipolar plane $\boldsymbol{\Pi}$ is the plane defined by the $3D$ point $\boldsymbol{X}$ and the optical centers of the cameras.



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

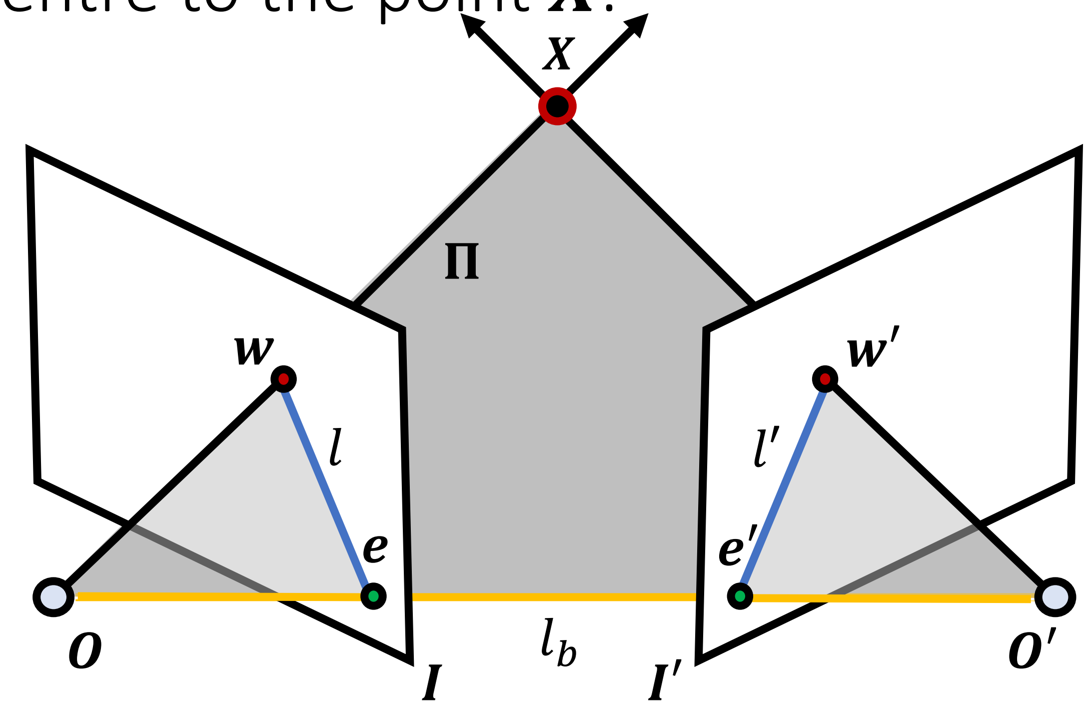
Epipolar Geometry - fundamentals

An epipole is the point of intersection of the baseline with the image plane. There are two epipoles $\boldsymbol{e}$ and $\boldsymbol{e}'$, one for each image.



Epipolar Geometry - fundamentals

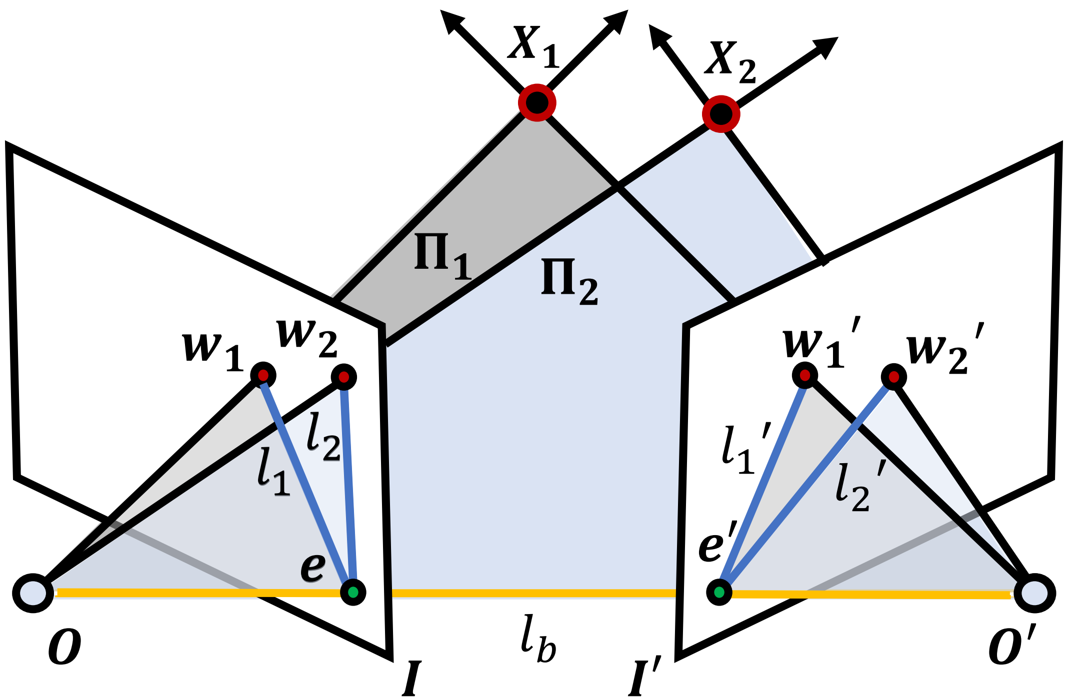
An epipolar line is a line of intersection of the epipolar plane with an image plane. It is the image, in one camera, of the ray from the other camera's optical centre to the point $\boldsymbol{X}$.



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Epipolar Geometry - fundamentals

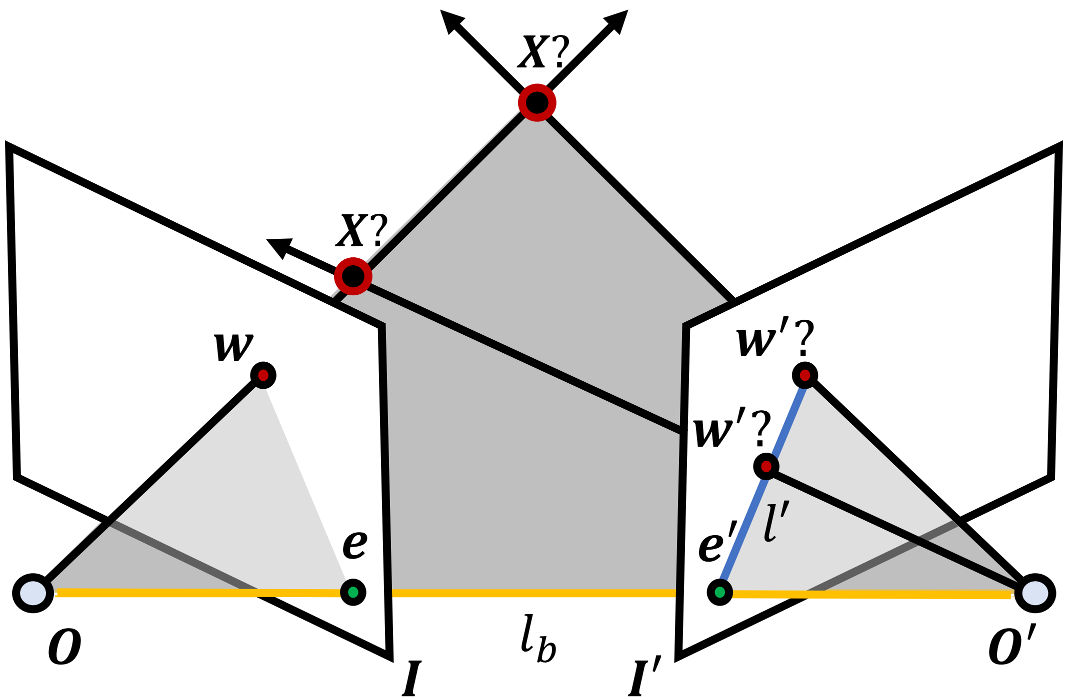
For different world points $\boldsymbol{X}$, the epipolar plane rotates about the baseline. All epipolar lines intersect at their corresponding epipole.



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

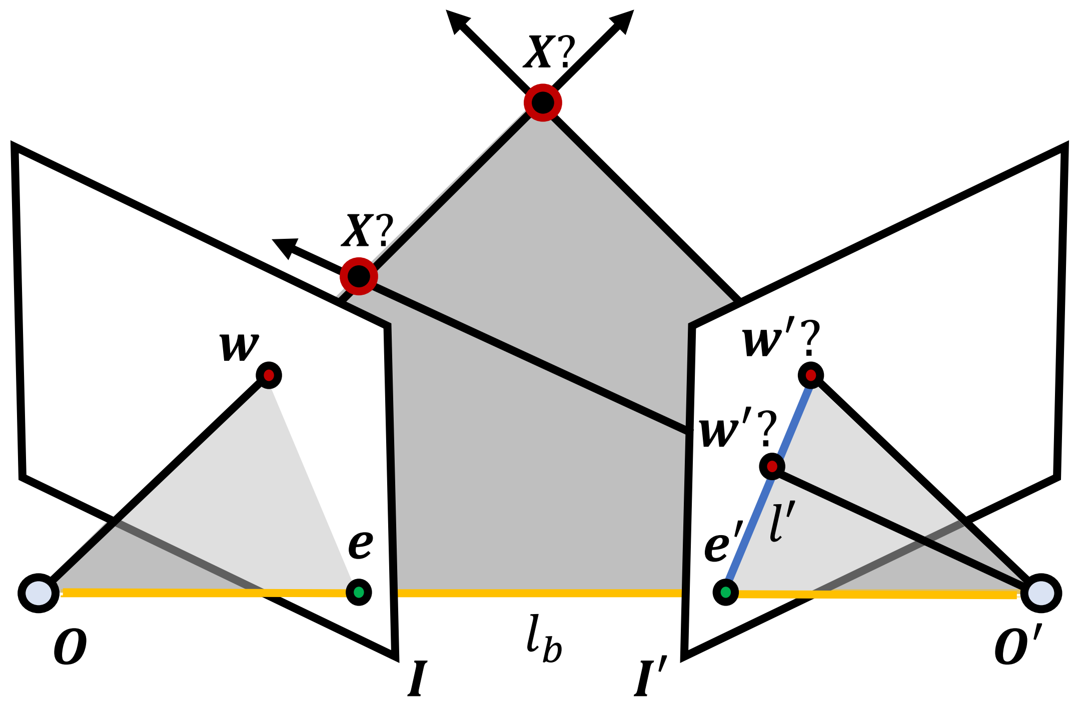
Epipolar Geometry - fundamentals

The **epipolar constraint** limits the search for correspondences, from the region of the whole image, to only the pixels spanned by the epipolar line.



Epipolar Geometry - fundamentals

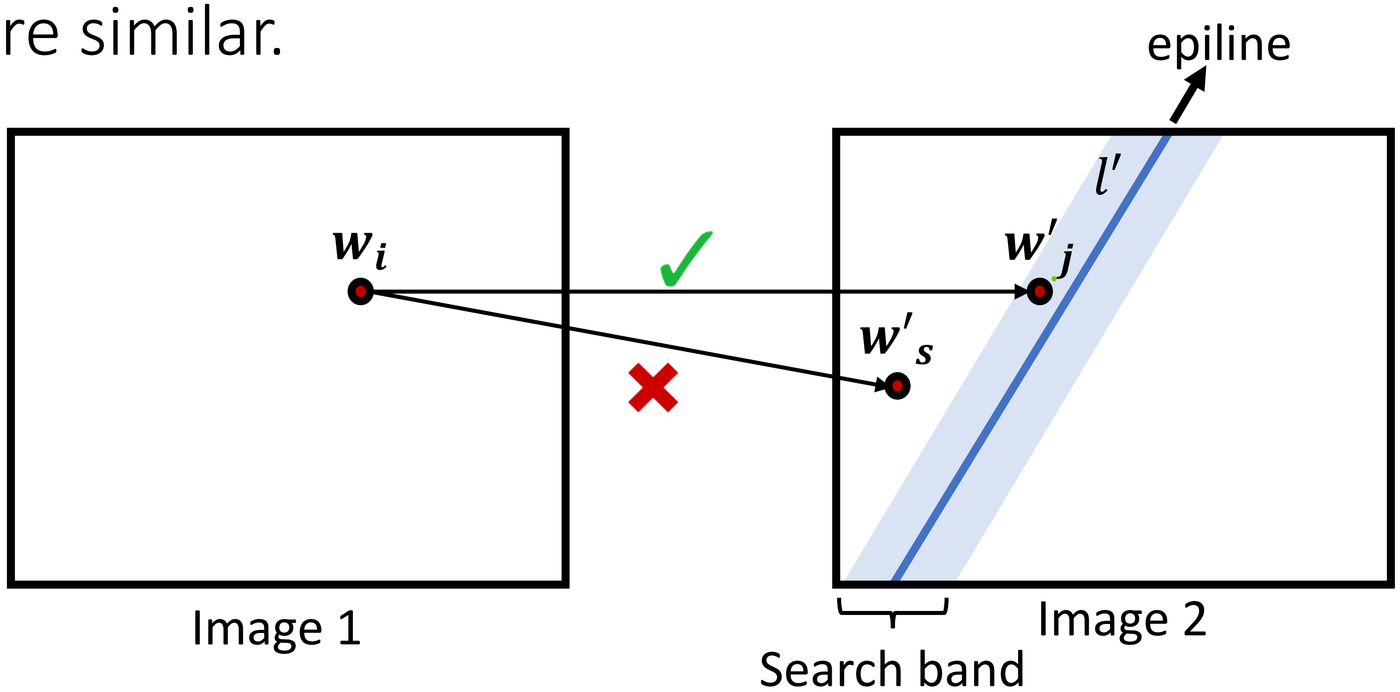
If a point feature $\boldsymbol{w}$ is observed in one image, then its location $\boldsymbol{w}'$ in the other image must lie on its corresponding epipolar line $\boldsymbol{l}'$



Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

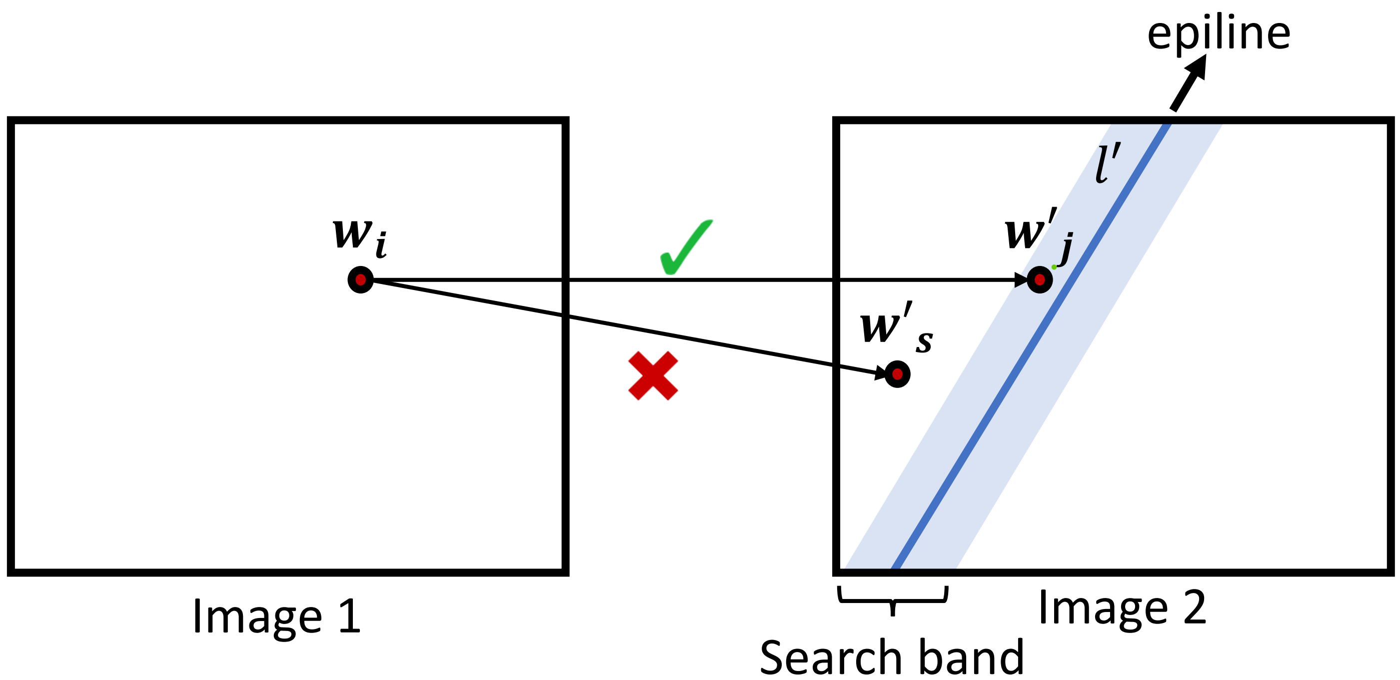
Epipolar Geometry - fundamentals

So a simple algorithm for determining correspondences is to match each feature $\boldsymbol{w}_i$ from camera 1 to a feature $\boldsymbol{w'}_j$ from camera 2 which is close to the epipolar line of camera 2, provided that the SIFT descriptors for $\boldsymbol{w}_i$ and $\boldsymbol{w'}_j$ are similar.



Epipolar Geometry - fundamentals

Therefore, we must calculate the equation of the epipolar lines.



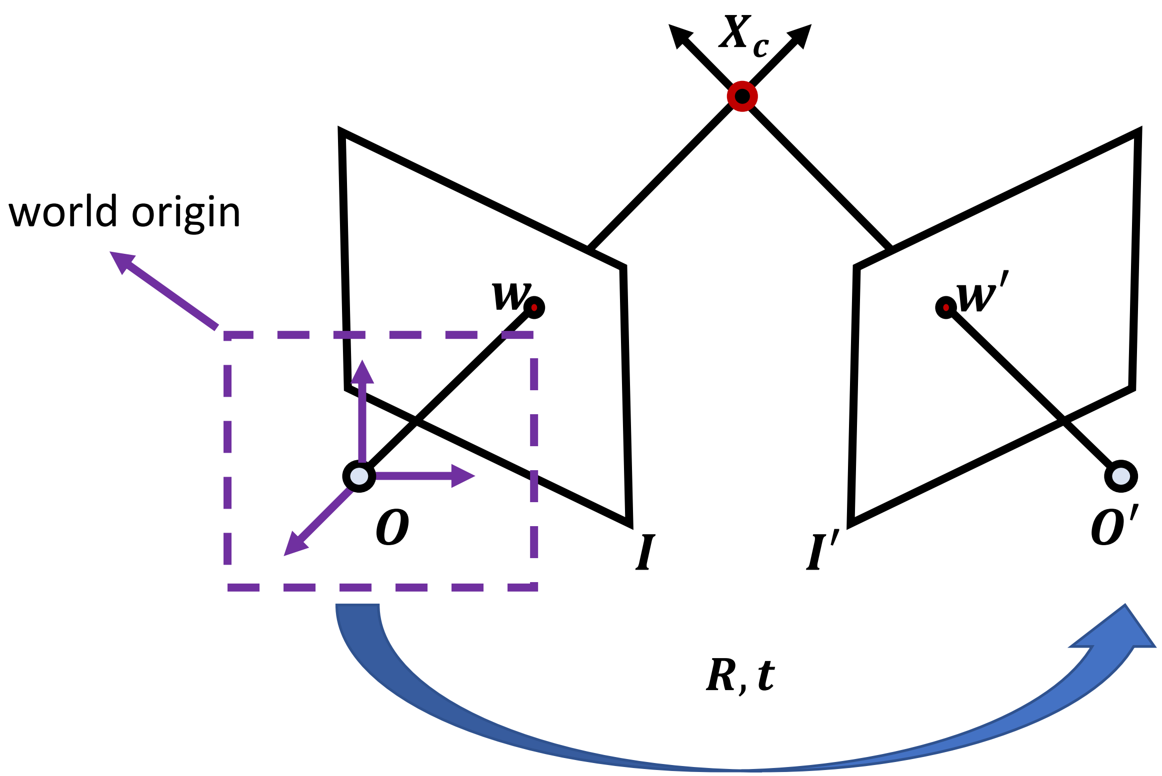
Essential Matrix

- Firstly, lets assume that the 1st camera is located at the **world origin** $\boldsymbol{O}$
- This means that every point $\boldsymbol{X}$ is expressed in the camera coordinates of the 1st camera, since its optical center coincides with the world origin $\Rightarrow \boldsymbol{X} \rightarrow \boldsymbol{X_c}$

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Essential Matrix

- If this is the case, the 2nd camera is at a location $\boldsymbol{O'}$ in world coordinates, which can be expressed by a **translation** $\boldsymbol{t}$ and **rotation** $\boldsymbol{R}$, w.r.t. 1st camera, i.e., the world origin



Essential Matrix

- Every $3D$ point $\boldsymbol{X}$ can be expressed as $\boldsymbol{X_c}$, which is the camera-centered coordinate system of the 1st camera, and as $\boldsymbol{X'_c}$, which is the camera-centered coordinate system of the 2nd camera.
- Since $\boldsymbol{X} = \boldsymbol{X_c}$, we can relate $\boldsymbol{X} \leftrightarrow \boldsymbol{X'_c}$ or $\boldsymbol{X_c} \leftrightarrow \boldsymbol{X'_c}$ using a **Euclidean transformation** composed by the **translation** $\boldsymbol{t}$ and **rotation** $\boldsymbol{R}$ of the 2nd camera, w.r.t. the 1st camera

Essential Matrix

Here is how to find an expression for the epipolar line:

$$\widetilde{X'_c} = P_e \widetilde{X_c}$$ ← This is in homogeneous coordinates

$$\Leftrightarrow X'_c = RX_c + t$$ ← Express it in cartesian coordinates

$$\Leftrightarrow t \times X'_c = t \times RX_c + \cancelto{0}{t \times t}$$ ← Apply cross product with t to both sides

$$\Leftrightarrow X'_c \cdot (t \times X'_c) = X'_c \cdot (t \times RX_c)$$ ← Apply dot product with X'_c to both sides

$$\Leftrightarrow 0 = X'_c \cdot (t \times RX_c)$$

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Essential Matrix

This can be rewritten in matrix form:

$$\boldsymbol{X}_{\boldsymbol{c}}' \cdot (\boldsymbol{t} \times \boldsymbol{R}\boldsymbol{X}_{\boldsymbol{c}}) = \boldsymbol{0}$$
$$\Leftrightarrow \boldsymbol{X}_{\boldsymbol{c}}'^{T} \boldsymbol{E} \boldsymbol{X}_{\boldsymbol{c}} = \boldsymbol{0}$$

Where $\boldsymbol{E} = \boldsymbol{T}_{\times}\boldsymbol{R}$ is the **essential matrix,** and

$$\boldsymbol{T}_{\times} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

is a matrix representing the cross product with $\boldsymbol{t}$ such that $\mathbf{t} \times \boldsymbol{v} = \boldsymbol{T}_{\times}\mathbf{v}$

Fundamental Matrix

Recall that for $\boldsymbol{X_c} \leftrightarrow \boldsymbol{w}$: $\quad \widetilde{\boldsymbol{w}} = \boldsymbol{K}\boldsymbol{X_c} \Leftrightarrow \boldsymbol{X_c} = \boldsymbol{K}^{-1}\widetilde{\boldsymbol{w}}$

and similarly, for $\boldsymbol{X'_c} \leftrightarrow \boldsymbol{w'}$: $\quad \widetilde{\boldsymbol{w}}' = \boldsymbol{K}'\boldsymbol{X}'_{\boldsymbol{c}} \Leftrightarrow \boldsymbol{X}'_{\boldsymbol{c}} = \boldsymbol{K}'^{-1}\widetilde{\boldsymbol{w}}'$

Combining the two equations yields the equation of **the two epipolar lines in pixel coordinates**:

$$\boldsymbol{X_c'^T}\,\mathbf{E}\,\mathbf{X_c} = 0$$
$$\Rightarrow (\boldsymbol{K}'^{-1}\widetilde{\boldsymbol{w}}')^T\,\boldsymbol{E}\,(\boldsymbol{K}^{-1}\widetilde{\boldsymbol{w}}) = 0$$
$$\Rightarrow \widetilde{\boldsymbol{w}}'^T(\mathbf{K}'^{-T}\,\mathbf{E}\,\mathbf{K}^{-1})\,\widetilde{\boldsymbol{w}} = 0$$
$$\Rightarrow \widetilde{\boldsymbol{w}}'^T\boldsymbol{F}\,\widetilde{\boldsymbol{w}} = \mathbf{0}$$

where $\boldsymbol{F} = \boldsymbol{K}'^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$ is the **fundamental matrix**.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Fundamental Matrix

This is the equation of the epipolar line in either camera:

$$\tilde{\boldsymbol{w}}'^{T} \boldsymbol{F}\, \tilde{\boldsymbol{w}} = \boldsymbol{0}$$

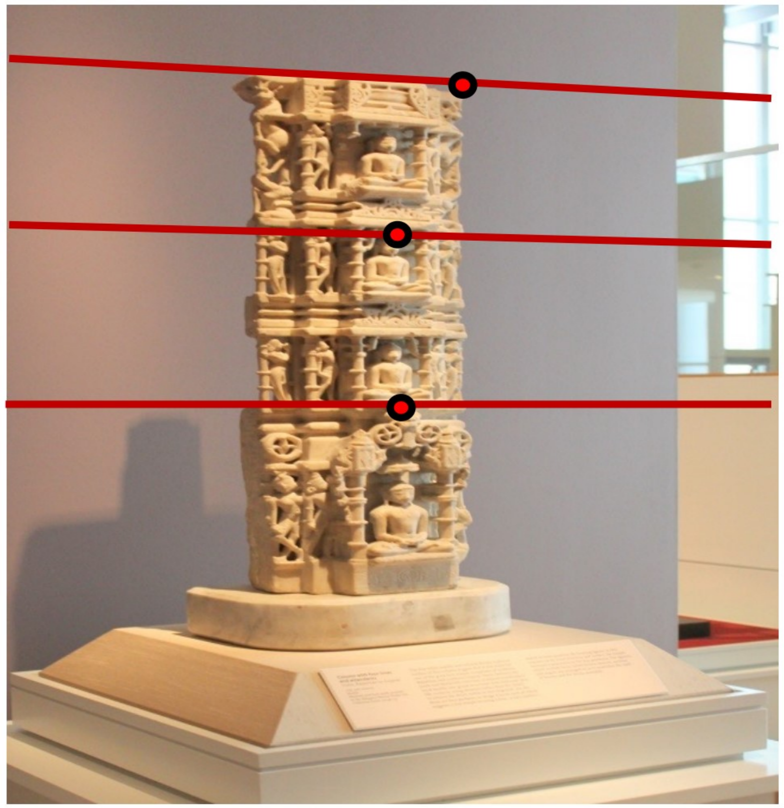
Assuming we know the fundamental matrix, for every point $\boldsymbol{w}$ in image 1, this expression gives us the line in image 2 on which the corresponding $\boldsymbol{w}'$ must lie, and *vice versa.*

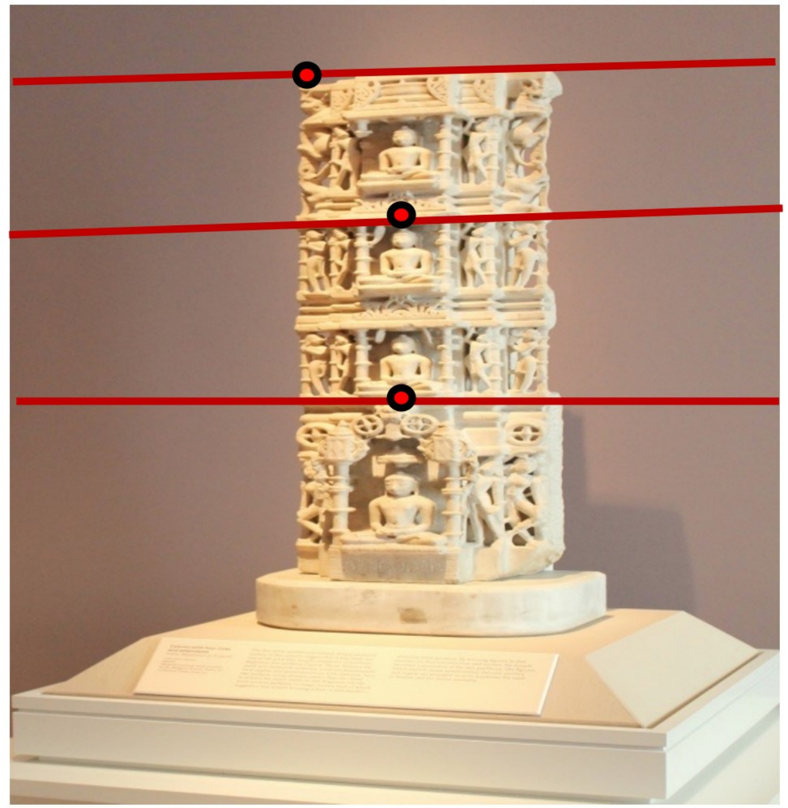
- $l' = \boldsymbol{F}\tilde{\boldsymbol{w}}$ is the epipolar line in the 2nd image, associated with $\boldsymbol{w}$
- $l = \boldsymbol{F}^{T}\tilde{\boldsymbol{w}}'$ is the epipolar line in the 1st image, associated with $\boldsymbol{w}'$
- $l_i: ax + by + c = 0$

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Examples

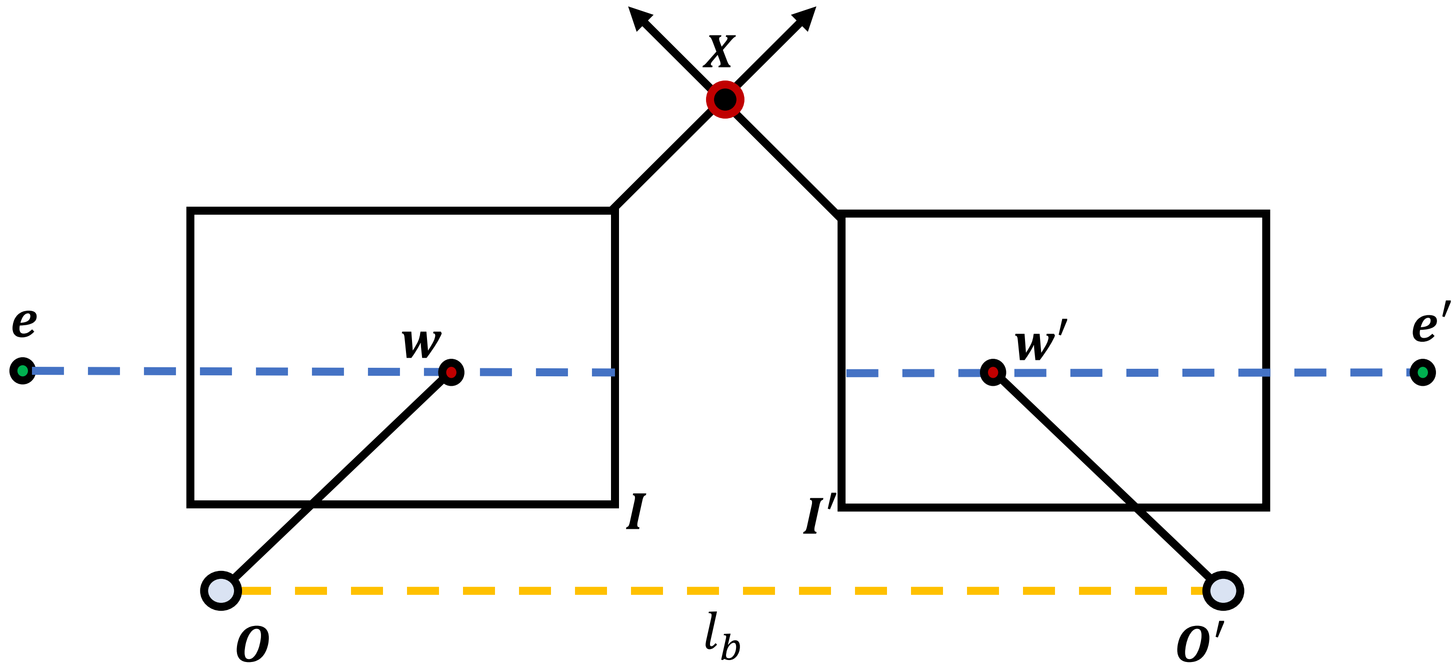
Here are a few examples of the epipolar constraint





Examples

Epipolar constraint examples: **Parallel image planes**



- Baseline intersects the image planes at infinity
- Epipoles are at infinity
- Epipolar lines are parallel to the $\boldsymbol{u}$-axis of each image plane

Examples

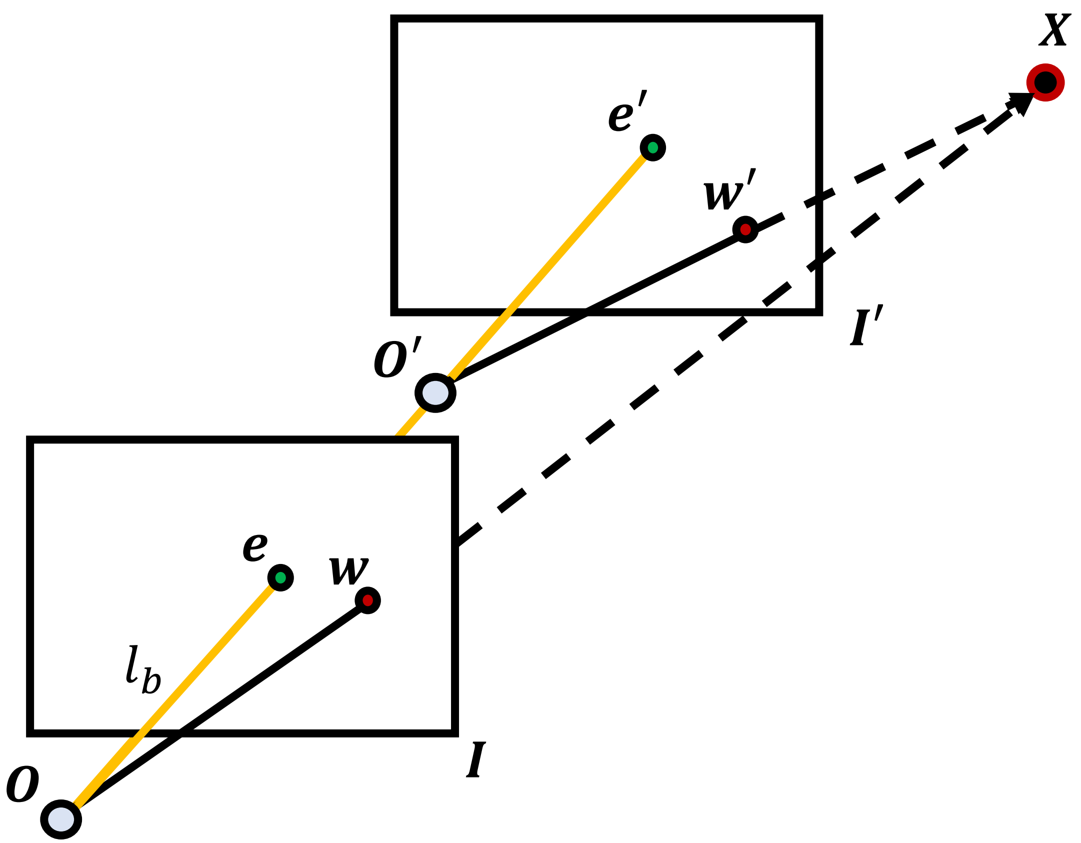
Epipolar constraint examples: **Parallel image planes**

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Examples

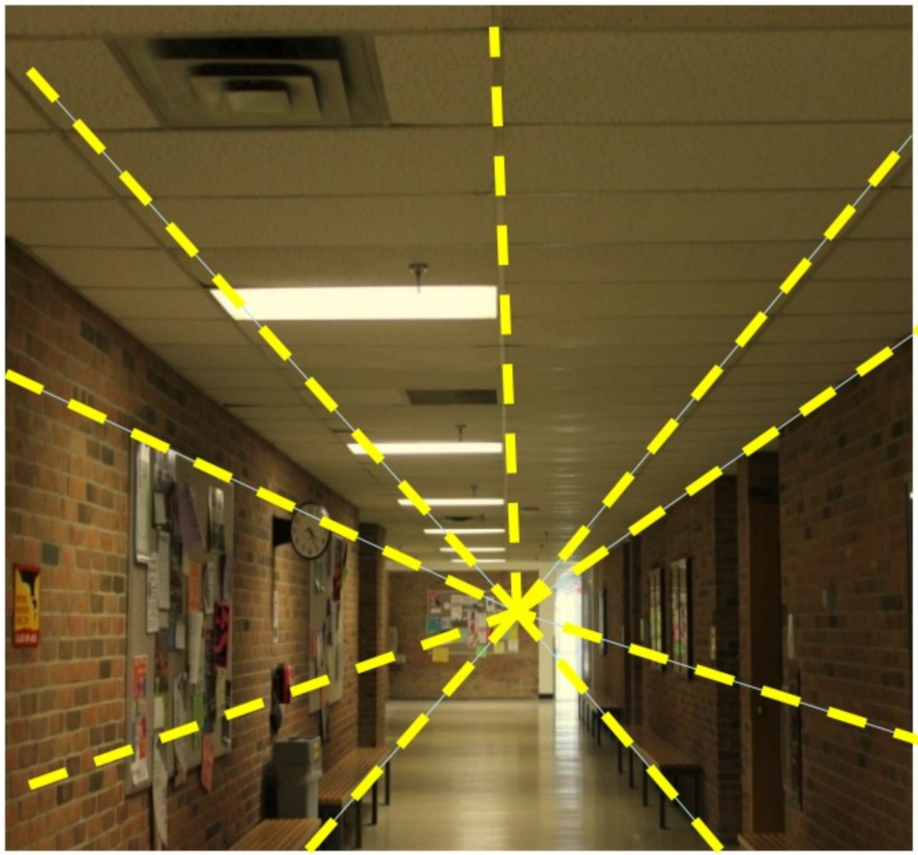
Epipolar constraint examples: **Forward translation**

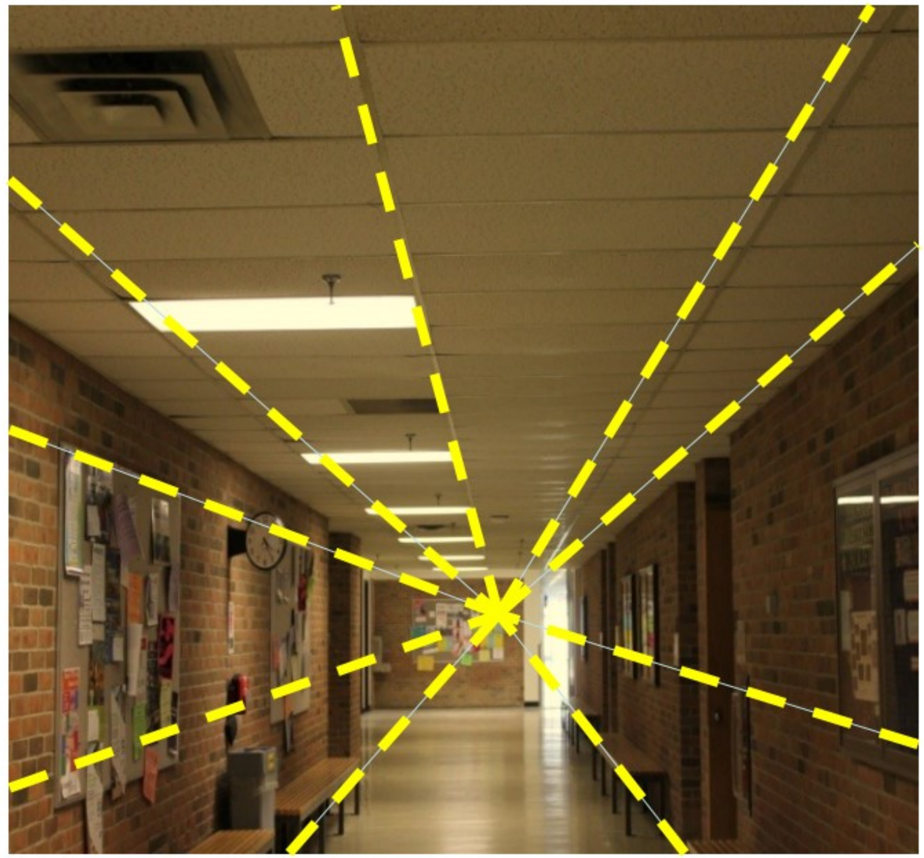


- The epipoles have the same position in both images

Examples

Epipolar constraint examples: **Forward translation**





Fundamental Matrix

In order to apply the epipolar constraint we need to know the fundamental matrix:

$$\boldsymbol{F} = \boldsymbol{K}'^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$$
$$= \boldsymbol{K}'^{-T}\boldsymbol{T}_{\times}\boldsymbol{R}\boldsymbol{K}^{-1}$$

All the parameters of $\boldsymbol{F}$ come from the calibration of the two cameras.

Remember that the perspective camera model $\boldsymbol{P}_{\boldsymbol{ps}} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{T}]$ contains this information.

If, however, these are not available (e.g., because we used a projective camera model to calibrate our cameras), $\boldsymbol{F}$ must be estimated using known image correspondences.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Estimating the fundamental matrix

To estimate the fundamental matrix, we follow a similar approach as in camera calibration

$$\tilde{\boldsymbol{w}}'^{T} \boldsymbol{F}\, \tilde{\boldsymbol{w}} = 0 \Rightarrow [u' \ \mathrm{v}' \ 1] \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0$$

Each point correspondence $\boldsymbol{w} \leftrightarrow \boldsymbol{w}'$ generates a single equation for estimating the parameters inside $\boldsymbol{F}$.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Estimating the fundamental matrix

Here are N such correspondences:

$$\begin{bmatrix} u_1u'_1 & v_1u'_1 & u'_1 & u_1v'_1 & v_1v'_1 & v'_1 & u_1 & v_1 \\ & & & \vdots & & & & \\ u_Nu'_N & v_Nu'_N & u'_N & u_Nv'_N & v_Nv'_N & v'_N & u_N & v_N \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix}$$

The minimal number for N is 8, because $\boldsymbol{F}$ has 8 degrees of freedom (we can set $f_{33} = 1$)

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Estimating the fundamental matrix

- In practise, 8 clearly indicated correspondences are never available

- What is available is a set of correspondences proposed by SIFT, for a large set of features, in which there are **always** errors

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Estimating the fundamental matrix

- We can use RANSAC to solve this problem
  1. Obtain 8 random SIFT correspondences. Use them to estimate $\boldsymbol{F}$.
  2. For every feature $\boldsymbol{w_i}$ in image 1 calculate the epipolar line in image 2. Check if the corresponding feature $\boldsymbol{w'_j}$ in image 2 proposed by SIFT falls **"close"** to the epipolar line. Count the number S of such "inliers".
  3. If $S \geq T$ where T is a threshold, then there is consensus with the random sample taken in the first step. Calculate $\boldsymbol{F}$ for all inliers and terminate here.
  4. If $S < T$ then no consensus is reached. Repeat from step 1.
  5. If after N iterations no consensus is reached, select the model that gave the highest S, calculate $\boldsymbol{F}$ using all inliers in S and terminate

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU | Master programmes in Artificial Intelligence 4 Careers in Europe

Visual Computing Group

CYENS CENTRE OF EXCELLENCE

Thank you.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423