Multi-View Geometry

Giacomo Boracchi

CVPR USI, March 13 2020

Book: HZ, chapters 9, 11, 12

Intrinsic Ambiguity

Intrinsic ambiguity of the 3D to the 2D mapping: some information is simply lost.

After projection, different depths cannot be distinguished in the image plane.

Courtesy slide S. Lazebnik

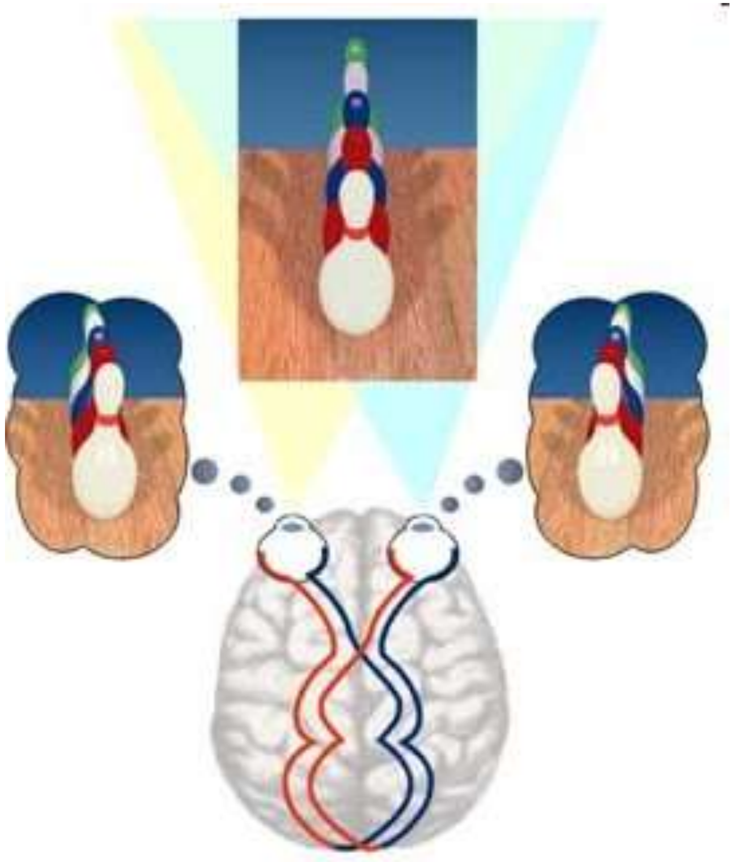
Intrinsic Ambiguity

Intrinsic ambiguity of the 3D to the 2D mapping: some information is simply lost.

After projection, different depths cannot be distinguished in the image plane

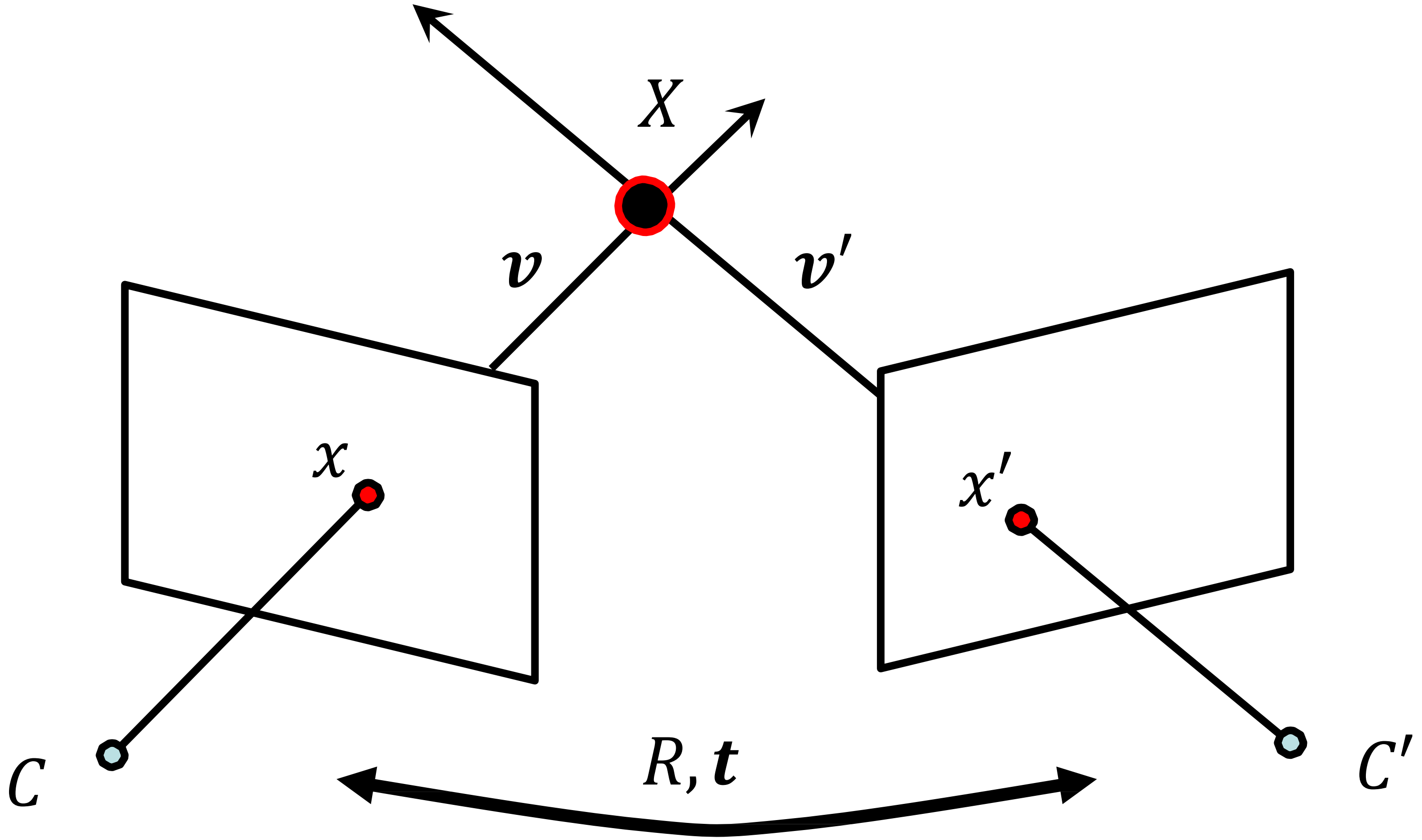
Multiple views of the same scene help us resolve these potential ambiguities

- In fact, the same scene from a different viewpoint wouldn't give rise to the same ambiguity



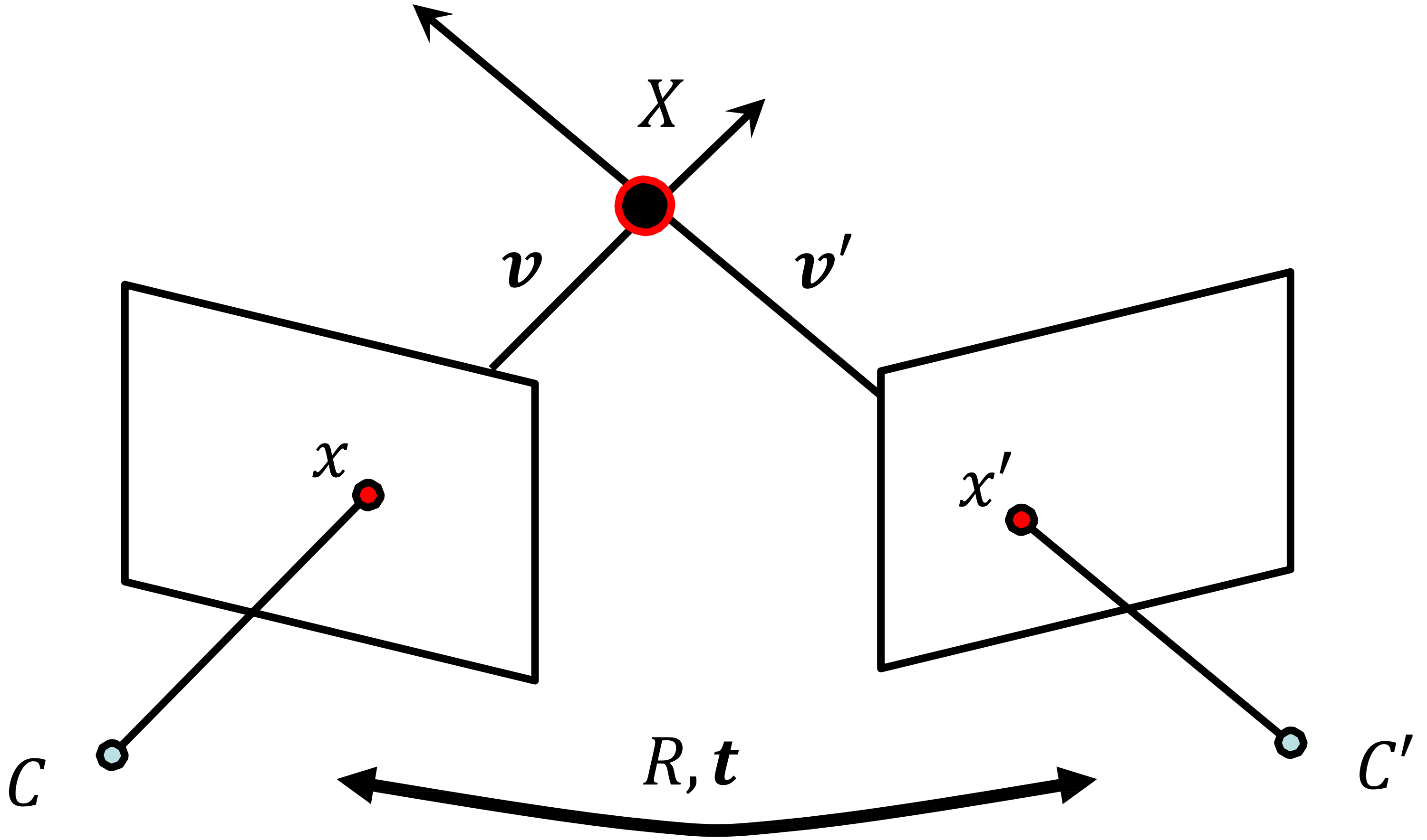
Standard Setup

Assume there are two cameras with known calibration matrices K and K', and moreover we set up a common 3D reference system



3D Reconstruction by Triangulation

Then we can compute and intersect the viewing rays through x, x' to obtain X (intersection is not the cross product since $\boldsymbol{v}$ and $\boldsymbol{v}'$ are 4-D)



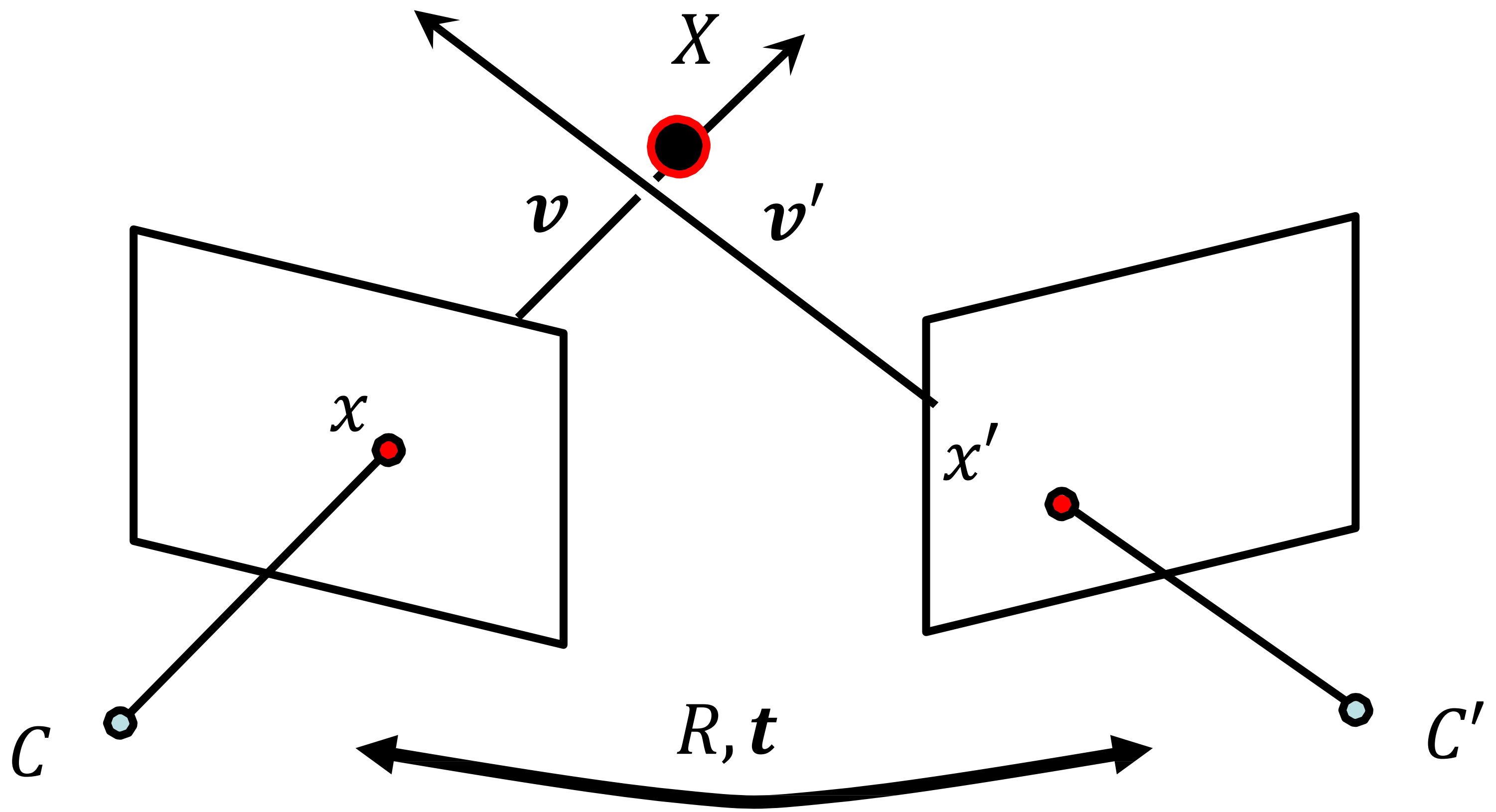
We can assume that the camera calibration matrices M and M' have been accurately estimated.

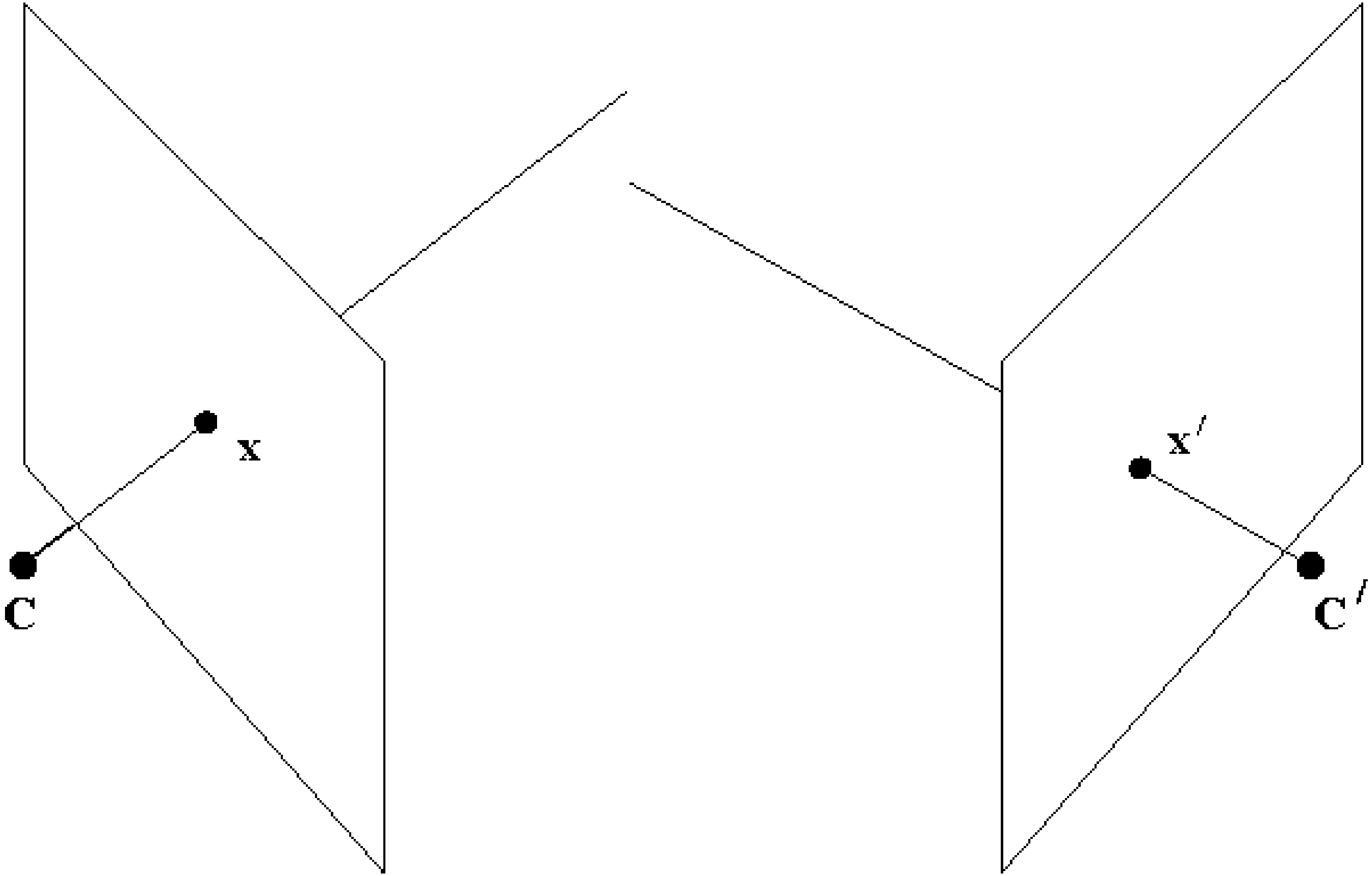
However, point correspondences might not be accurate

e.g. $\boldsymbol{x}' \neq M'X$ even though $\boldsymbol{x} = MX$

Then, in this case, the two viewing rays $\boldsymbol{v}$ and $\boldsymbol{v}'$ might be skewed ,

Thus $\boldsymbol{v}$ and $\boldsymbol{v}'$ might not intersect




x
C
x'
C'

3D Reconstruction by Triangulation

A more principled way to triangulation is solve the following system for X

$$\begin{cases} \boldsymbol{x} = MX \\ \boldsymbol{x}' = M'X \end{cases}$$

Following the same derivation for camera calibration estimation for $\boldsymbol{x} = MX$ (we are always starting from the same equation!)

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = MX = \begin{bmatrix} \boldsymbol{m_1}\, X \\ \boldsymbol{m_2}\, X \\ \boldsymbol{m_3}\, X \end{bmatrix} = \begin{bmatrix} \boldsymbol{m_1}\, X /_{\boldsymbol{m_3}\, X} \\ \boldsymbol{m_2}\, X /_{\boldsymbol{m_3}\, X} \\ 1 \end{bmatrix}$$

3D Reconstruction by Triangulation

Which leads to the following linear system when considering $\boldsymbol{x}' = M'X$

$$\begin{cases} x\,\boldsymbol{m_3}\,X - \boldsymbol{m_1}\,X = 0 \\ y\,\boldsymbol{m_3}\,X - \boldsymbol{m_2}\,X = 0 \\ x'\,\boldsymbol{m_3}'\,X - \boldsymbol{m_1}'\,X = 0 \\ y'\,\boldsymbol{m_3}'\,X - \boldsymbol{m_2}'\,X = 0 \end{cases}$$

Where $\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{m_i}, \boldsymbol{m_i}'$ are given and X is the only unknown

We can write this in a matrix

$$A = \begin{bmatrix} x\,\boldsymbol{m_3} - \boldsymbol{m_1} \\ y\,\boldsymbol{m_3} - \boldsymbol{m_2} \\ x'\,\boldsymbol{m_3}' - \boldsymbol{m_1}' \\ y'\,\boldsymbol{m_3}' - \boldsymbol{m_2}' \end{bmatrix} \in \mathbb{R}^{4\times 3}$$

3D Reconstruction by Triangulation

The triangulation corresponds to solving a over-determined system

$$AX = 0$$

It might not exist an exact and nontrivial solution, so we look for

$$\min_X ||AX||_2 \text{ s.t. } ||X||_2 = 1$$

… which algorithm does solve these problems? DLT!

Rmk This algorithm does provide a solution even in the most likely case that the two viewing rays are skewed (where simply intersecting them would not solve the problem)

Nonlinear Methods for Triangulation

However, a different approach is preferred, such that triangulation is seen as a form of estimation for 3D point location.

The approach: provided $\boldsymbol{x}, \boldsymbol{x}'$, estimate $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{x}}'$ such that

$$\widehat{\boldsymbol{x}} = M\hat{X} \text{ and } \widehat{\boldsymbol{x}}' = M'\hat{X}$$

By minimizing $\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\|_2 + \|\widehat{\boldsymbol{x}}' - \boldsymbol{x}'\|_2$

Rmk this gives pairs $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{x}}'$ corresponding to the same 3D point

… However we would like to have a better way to write down the constraint $\widehat{\boldsymbol{x}} = M\hat{X}$ and $\widehat{\boldsymbol{x}}' = M'\hat{X}$

Here is Epipolar Geometry!

Epipolar Geometry

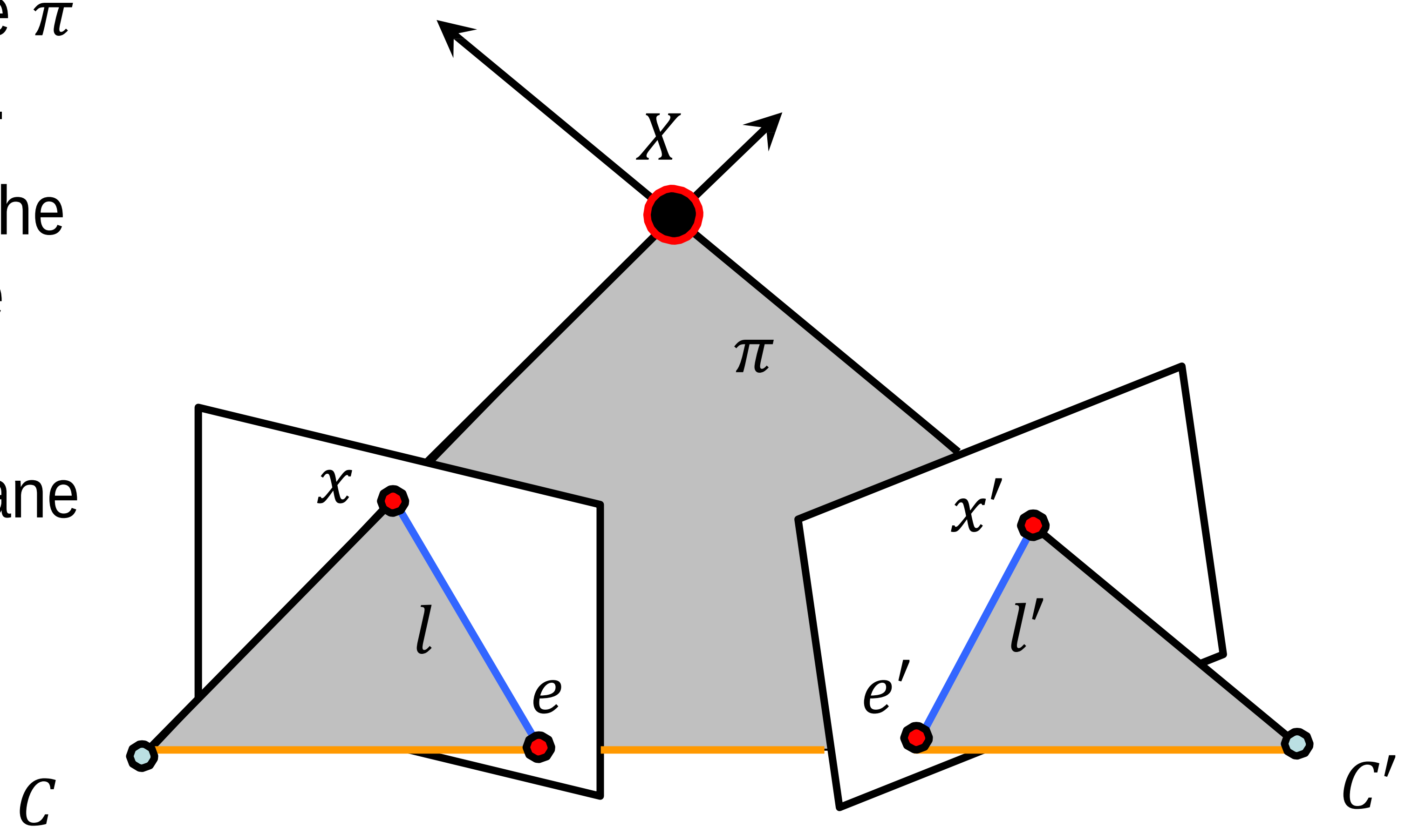
Epipolar Geometry: Setting up the Stage

Baseline: The line joining the two camera centers

Epipolar plane at X : the plane π containing X and the baseline.

Epipoles: intersection e, e' of the image planes and the baseline

Epipolar lines for X: lines l, l' intersection of the epipolar plane and the image planes.

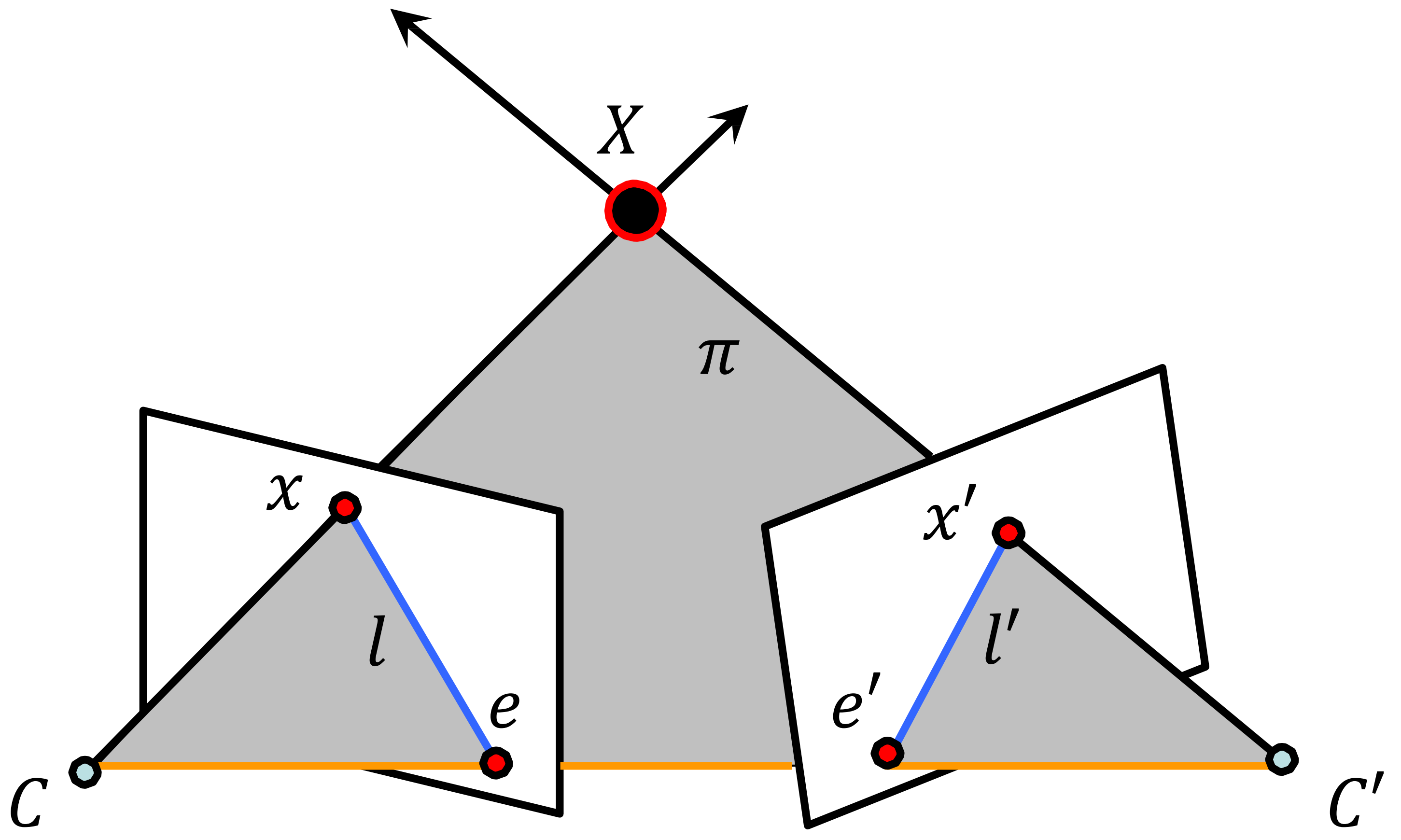


Epipolar geometry

Rmk: each epipolar line intersect the baseline at the epipole.

Rmk: therefore, the epipoles lie in the intersection of all the epipolar lines in each image

Rmk: we assume calibrated cameras



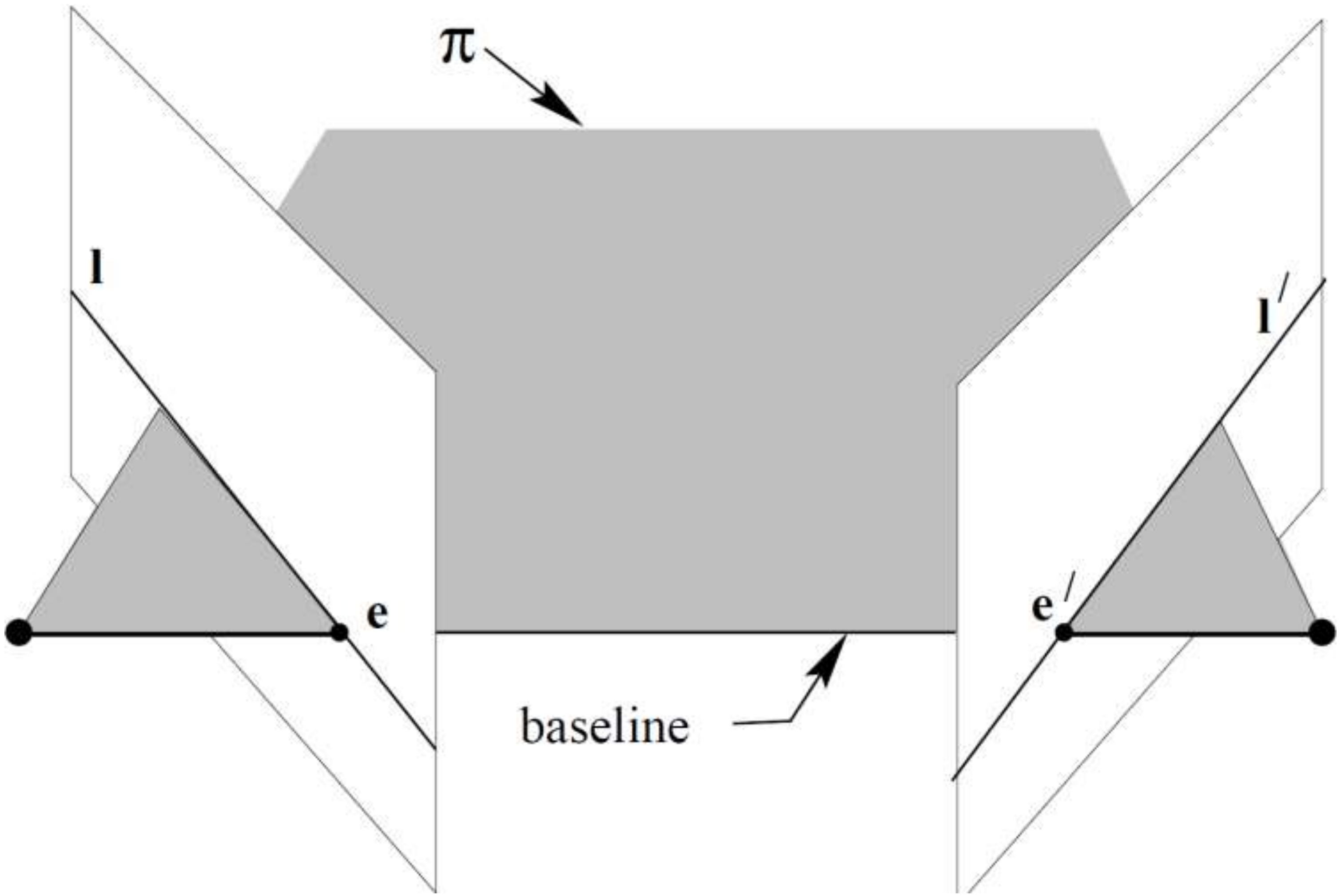
Epipoles

The epipoles can be seen as:

- the *point* of intersection of the line joining the camera centers (the baseline) with the image planes.
- the image in one view of the camera center yielding the other view.
- It is also the vanishing point of the baseline (translation) direction in each image.

Epipolar lines

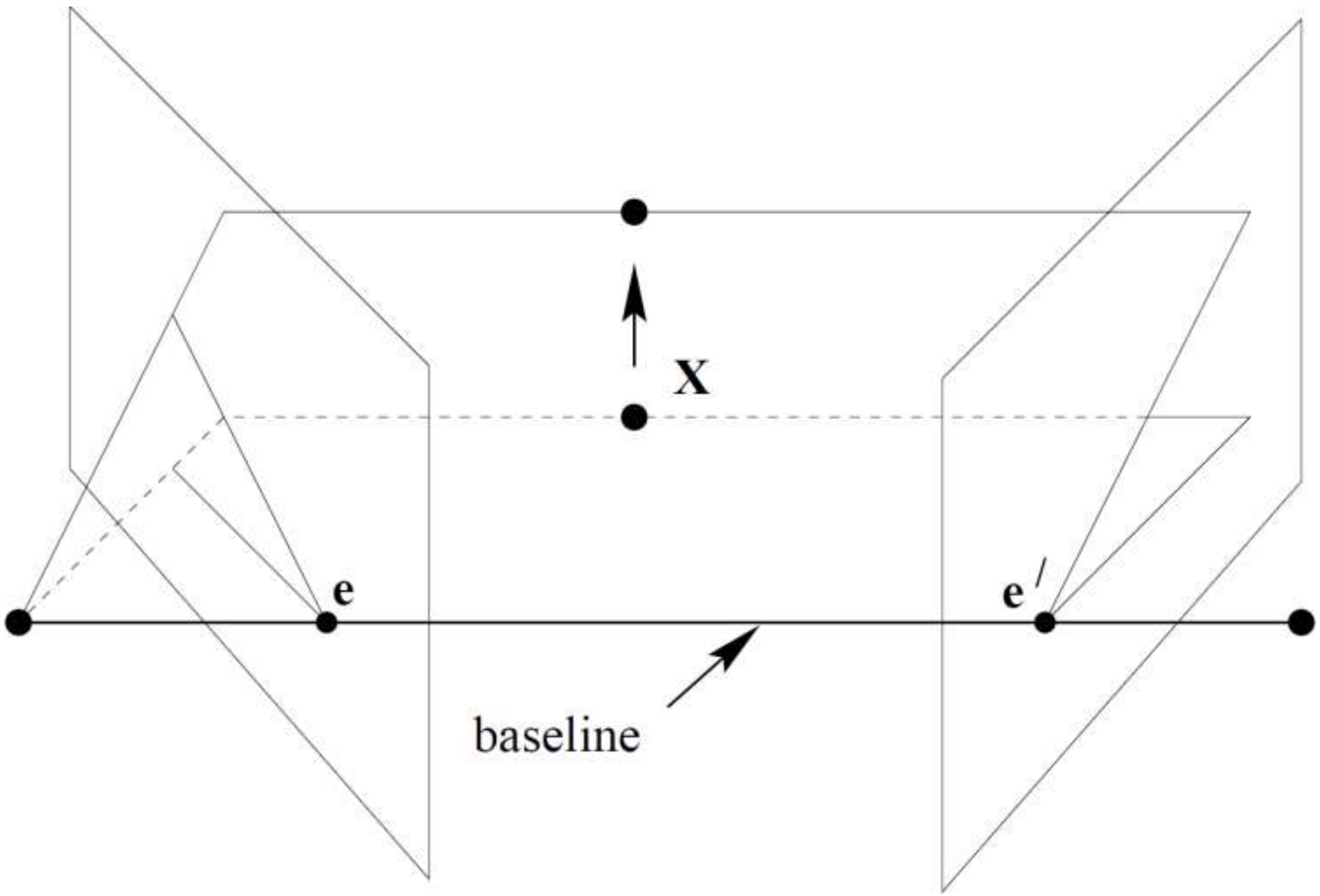
Epipolar lines can be seen as the intersection with the image plane of the pencil of planes (epipolar planes) having the baseline as axis



Hartley Zisserman Fig 9.2

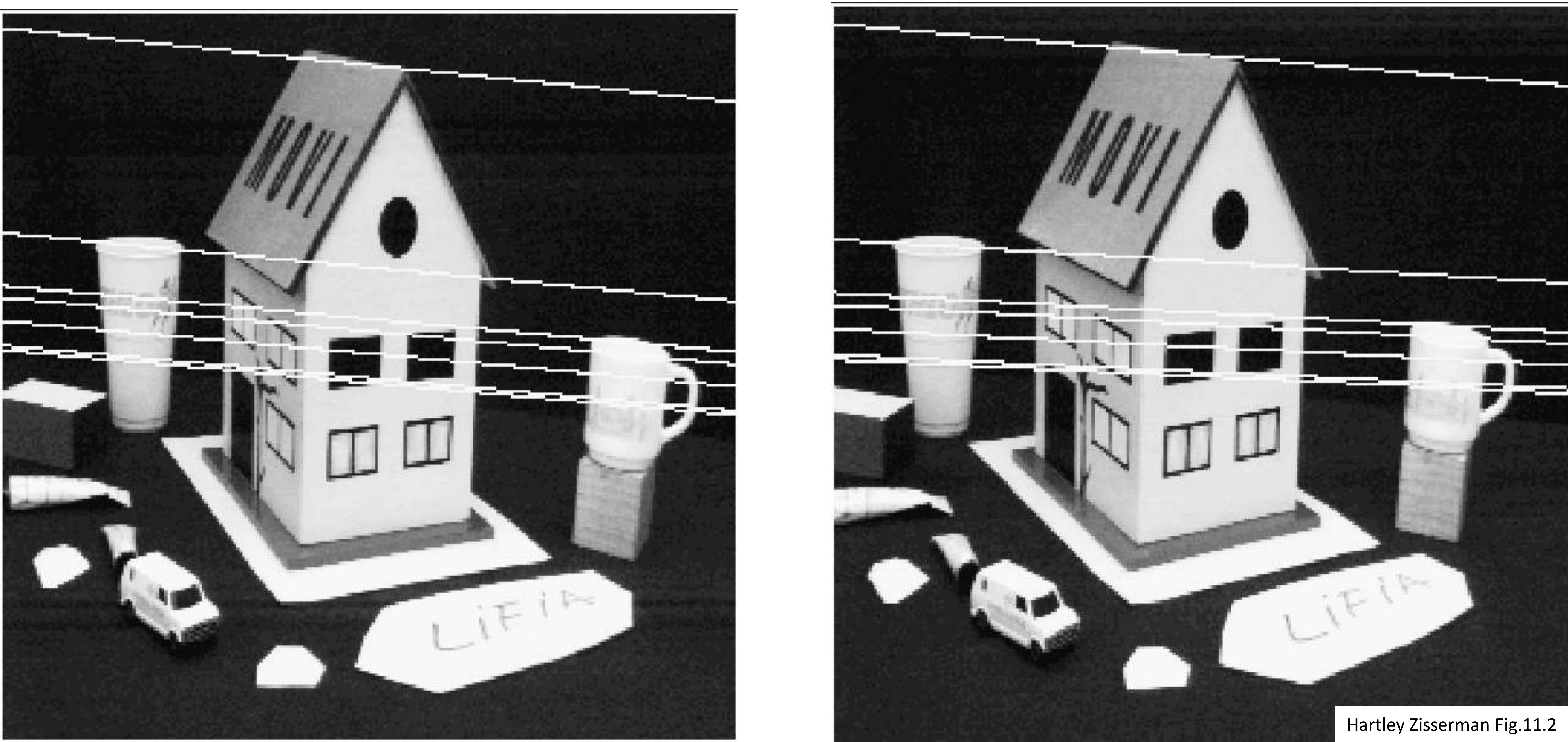
Epipolar lines

Epipolar lines can be seen as the intersection with the image plane of the pencil of planes (epipolar planes) having the baseline as axis



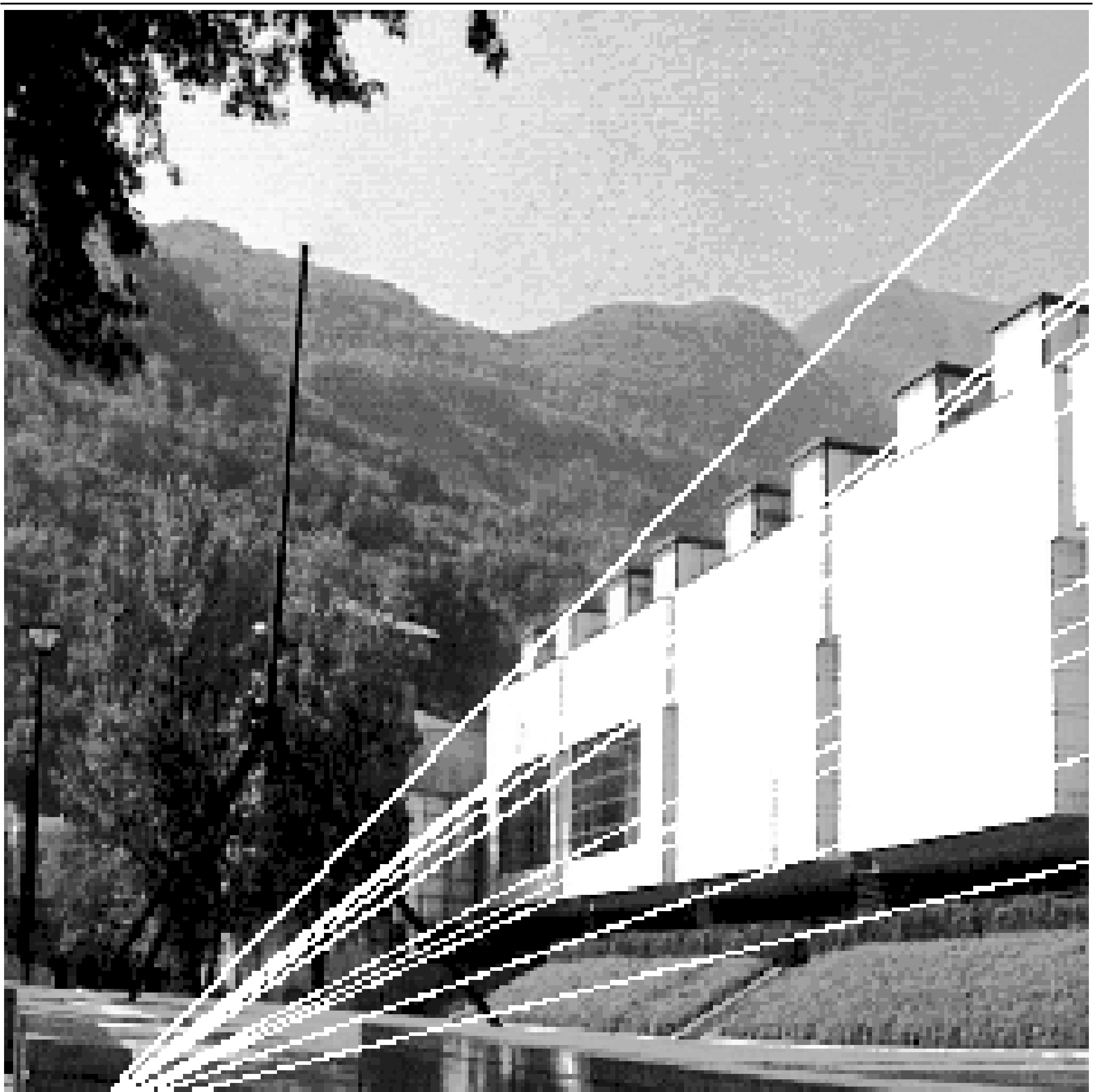
Hartley Zisserman Fig 9.2

Epipolar lines



Hartley Zisserman Fig.11.2

Epipolar lines



Hartley Zisserman Fig.11.2

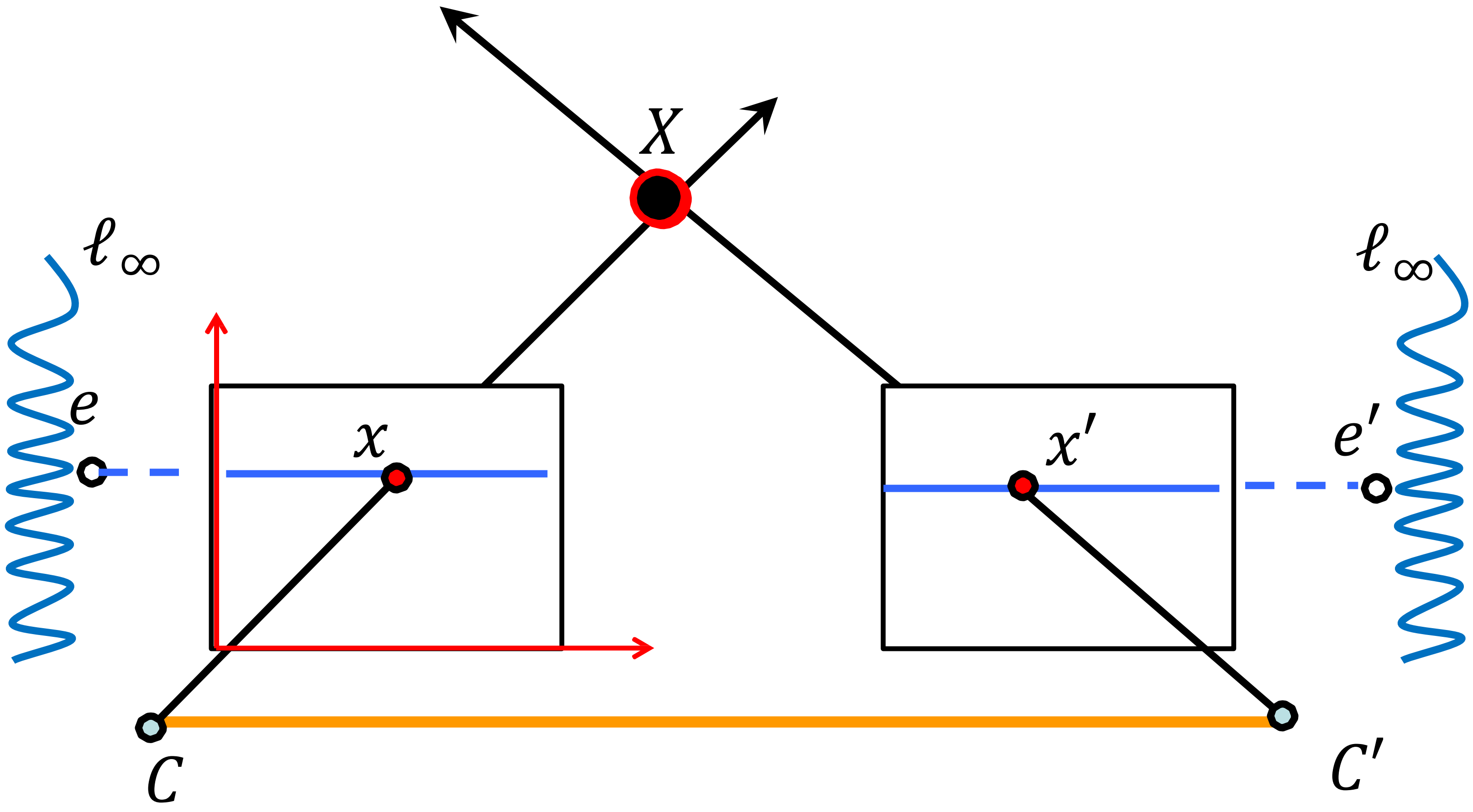
Two Interesting Cases

Parallel image planes

The baseline is parallel to both image planes

The epipoles $\boldsymbol{e}, \boldsymbol{e}'$ are located at the infinity (ideal points of the baseline direction)

All the epipolar lines are parallel to the horizontal axis of the camera



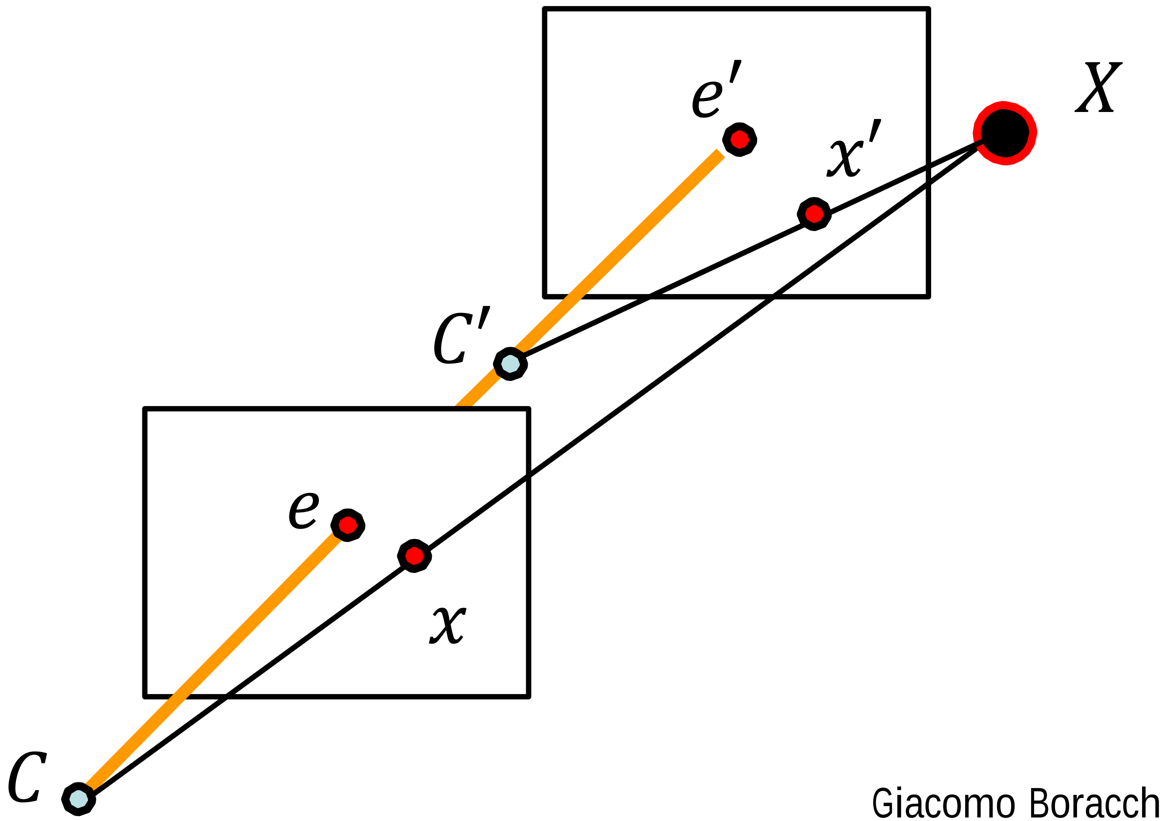
Parallel image planes

Forward Translation

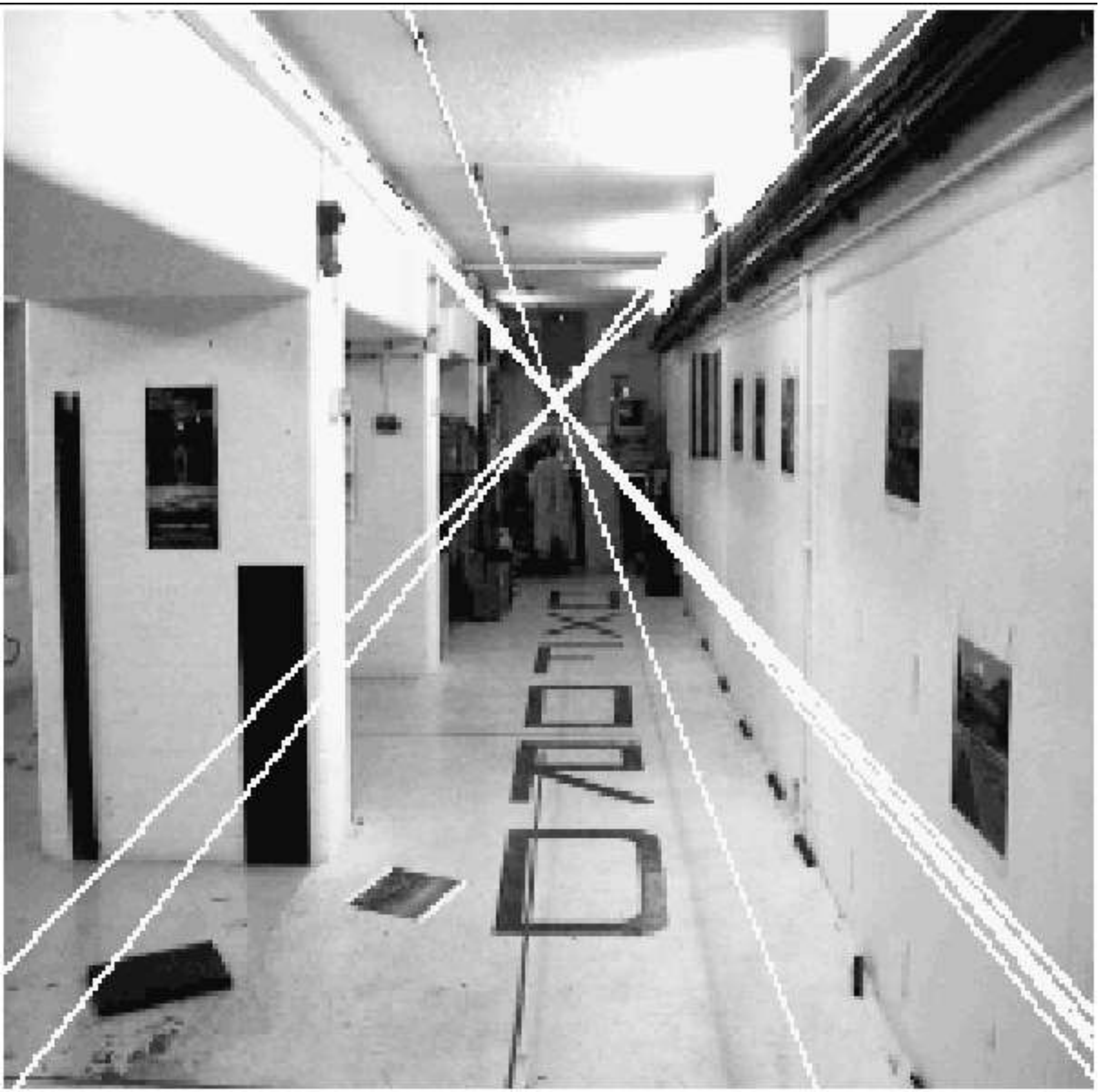
When the camera moves forward/backward, without changing orientation, the epipoles e and e' have the same coordinates in the image plane

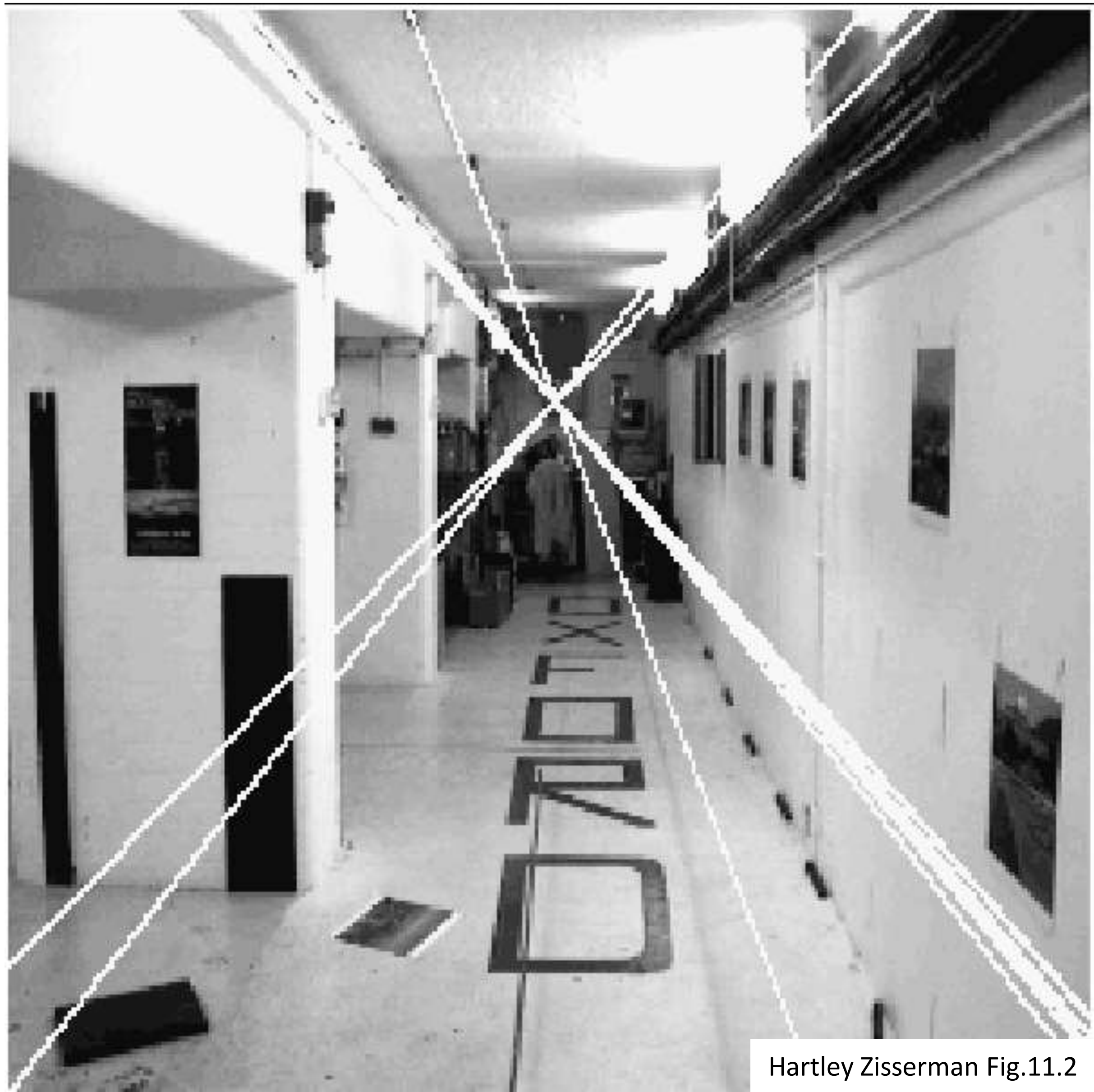
Epipolar lines are radial

Epipole is also called focus of expansion



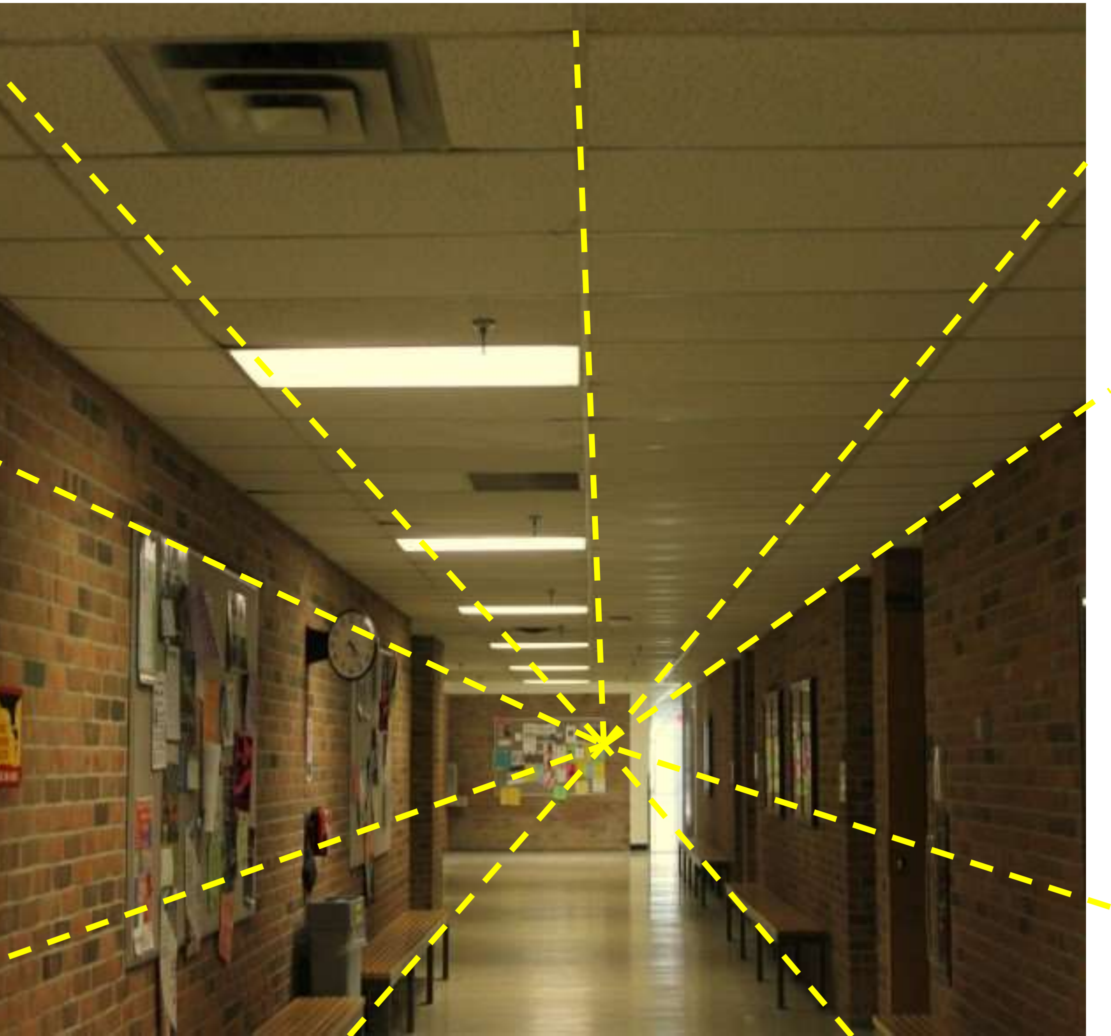
Forward Motion

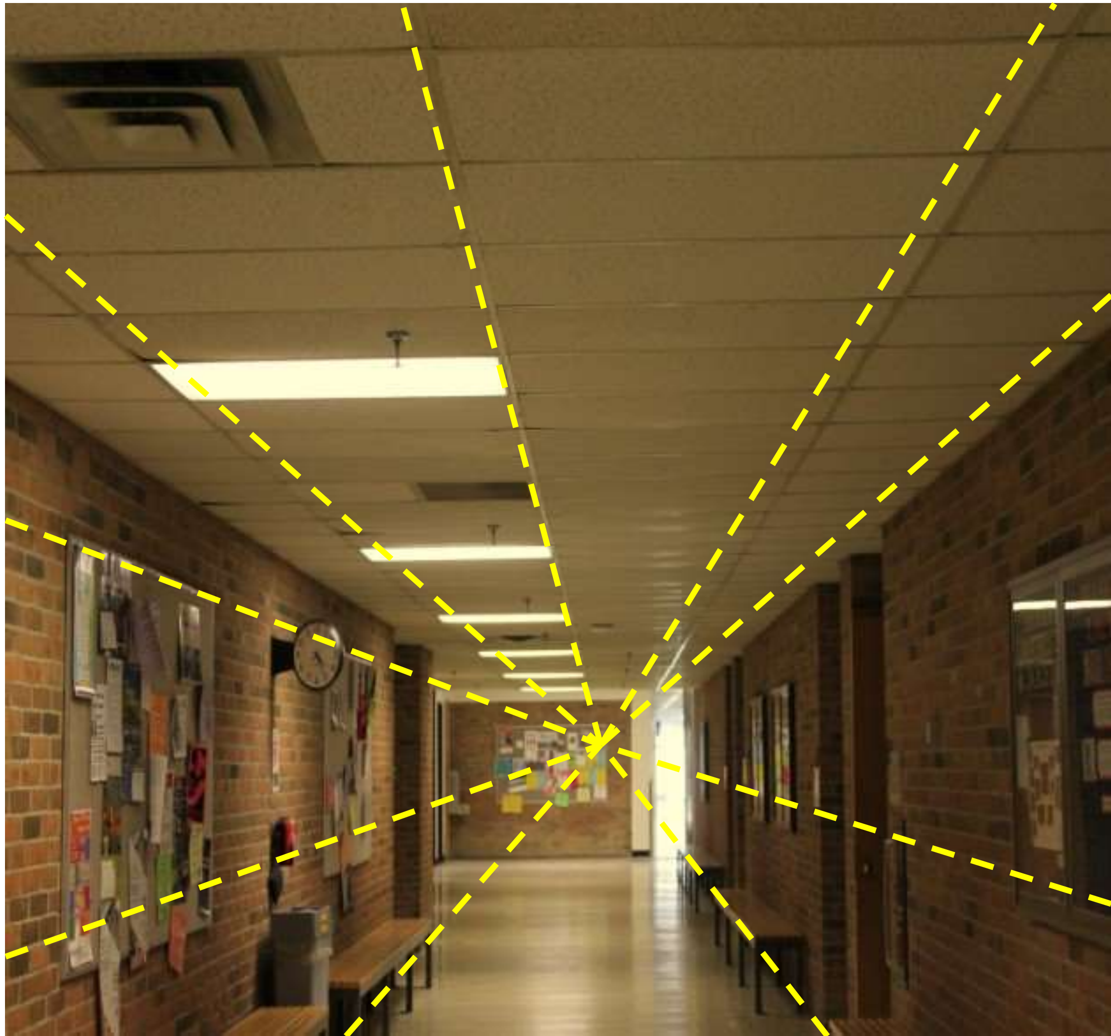




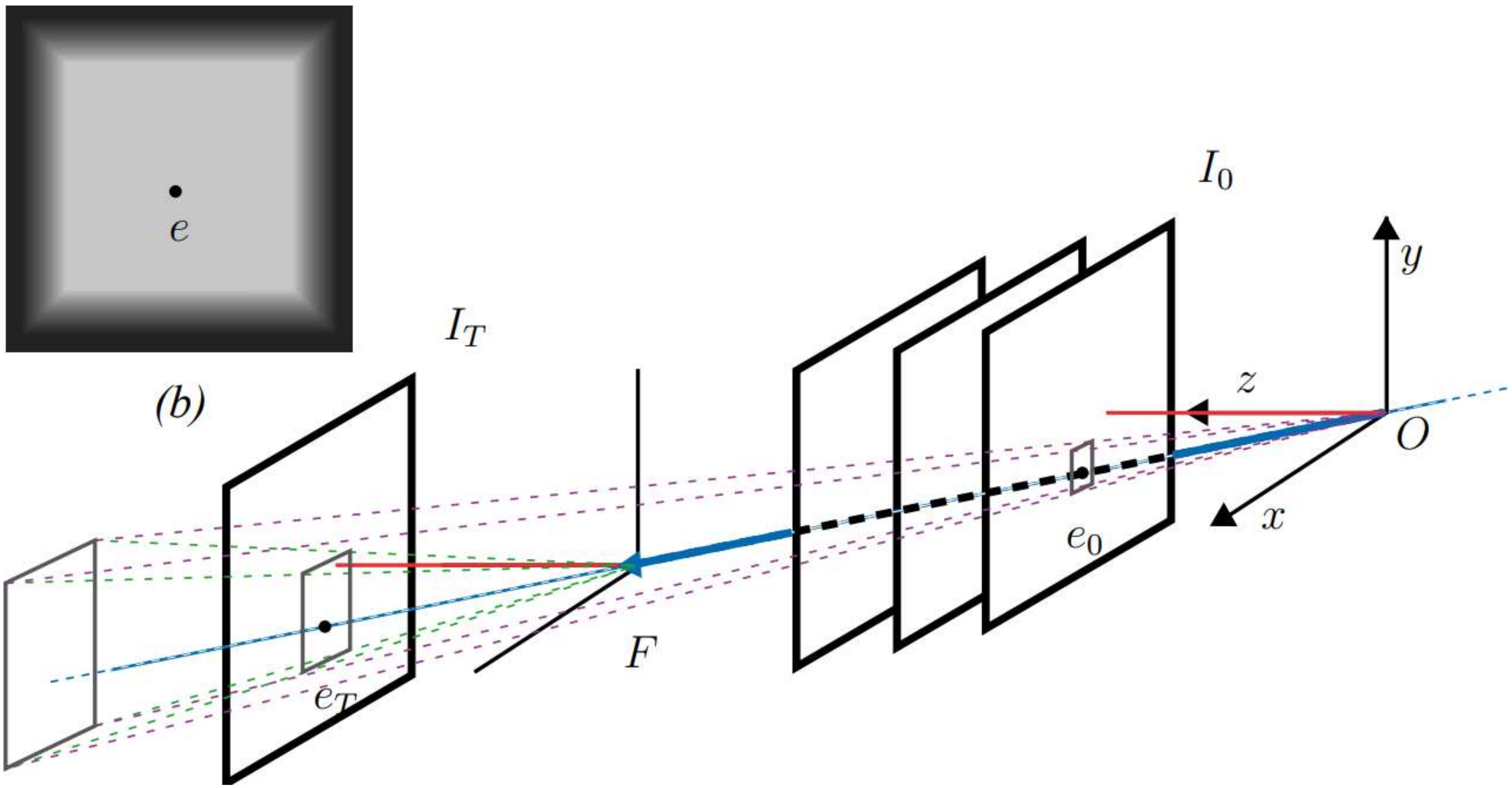
Hartley Zisserman Fig.11.2

Forward Motion





Radial Blur



Giacomo Boracchi Estimating the 3D Direction of a Translating Camera From a Single Motion-Blurred Image,, *Pattern Recognition* 2009

Giacomo Boracchi

Epipolar Constraints

Epipolar Geometry imposes geometric constraints between the two images, without any knowledge of the 3D scene

Epipolar Geometry

The intrinsic projective geometry between two views

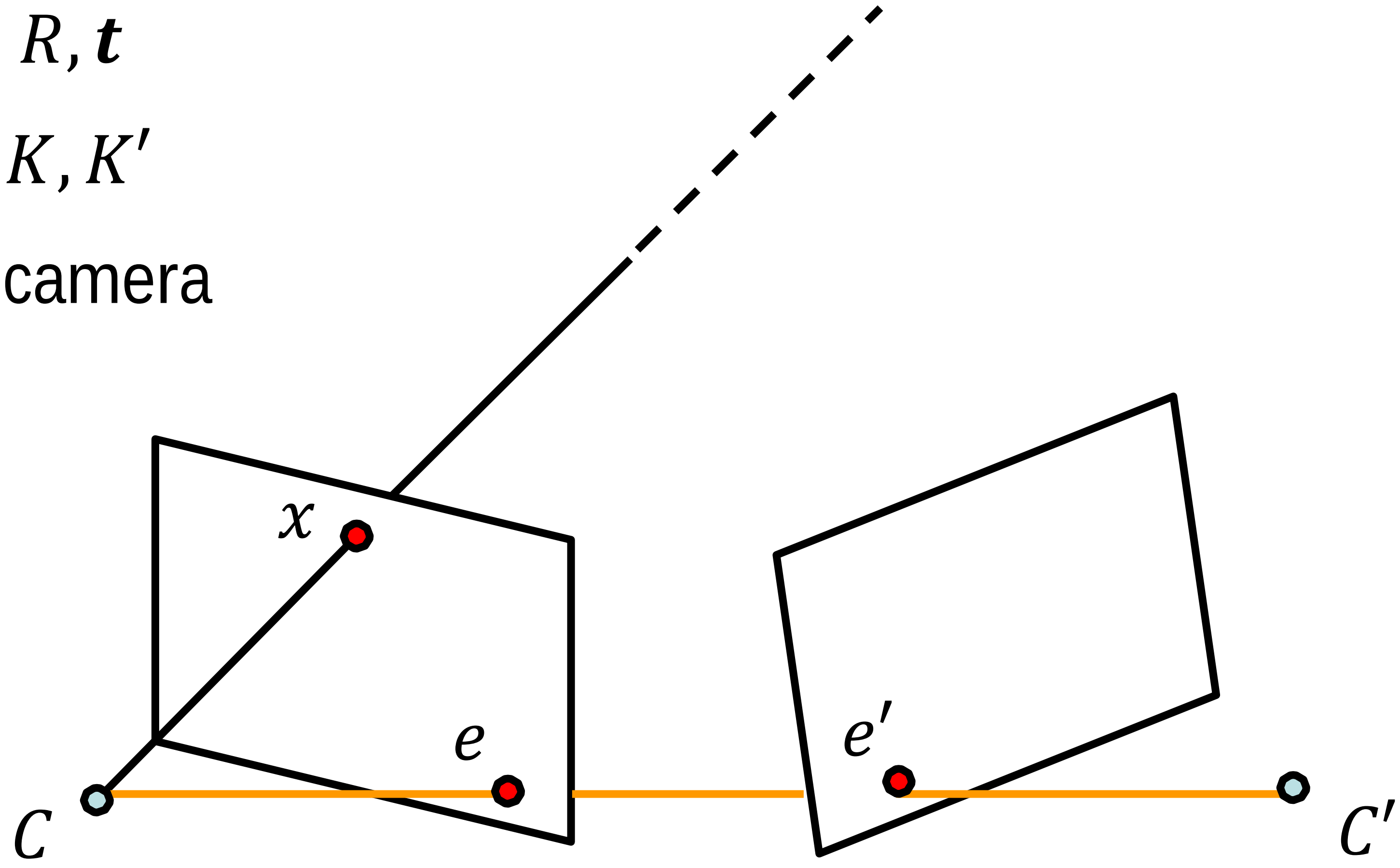
It is independent of scene structure, and only depends on the cameras' internal parameters and relative pose.

It applies to the case of two cameras observing the same scene as well as a moving camera

Epipolar Geometry Set-Up

Given two cameras observing the same scene for which we know:

- The camera centers C, C'
- The camera orientations, thus $R, \boldsymbol{t}$
- The intrinsic camera matrices K, K'
- Or alternatively, we know the camera projection matrices M and M'

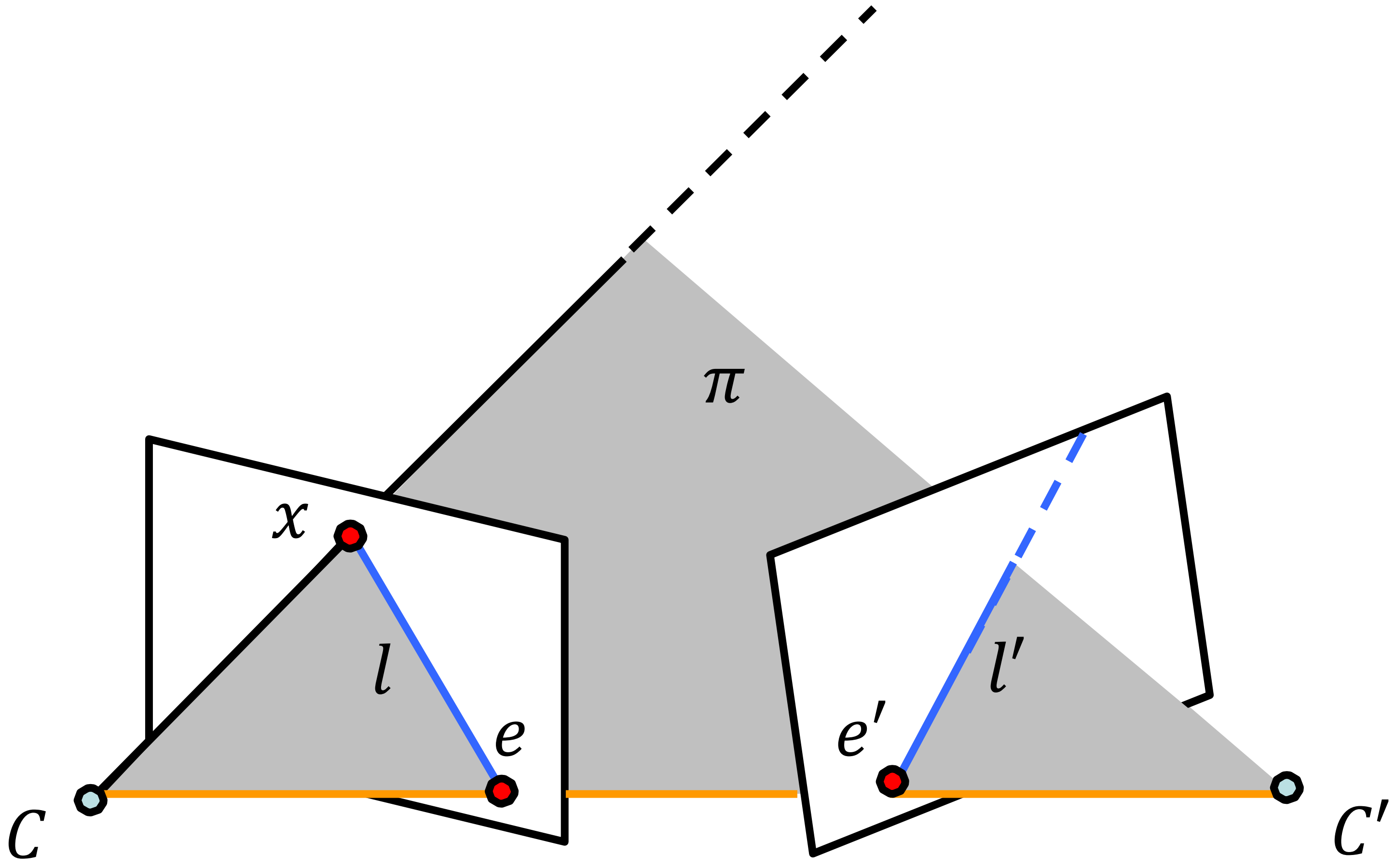


Epipolar Constraints

Then, given an image point $\boldsymbol{x}$ in the first image

1. we can compute the epipolar plane π, which does contain X and the unknown $\boldsymbol{x}'$
2. we can compute the epipolar lines $\boldsymbol{l}, \boldsymbol{l}'$ in both images (no need to know $\boldsymbol{x}'$)
3. The projection $\boldsymbol{x}'$ must be located in the epipolar line $\boldsymbol{l}'$

$$\boldsymbol{x}' \in \boldsymbol{l}'$$

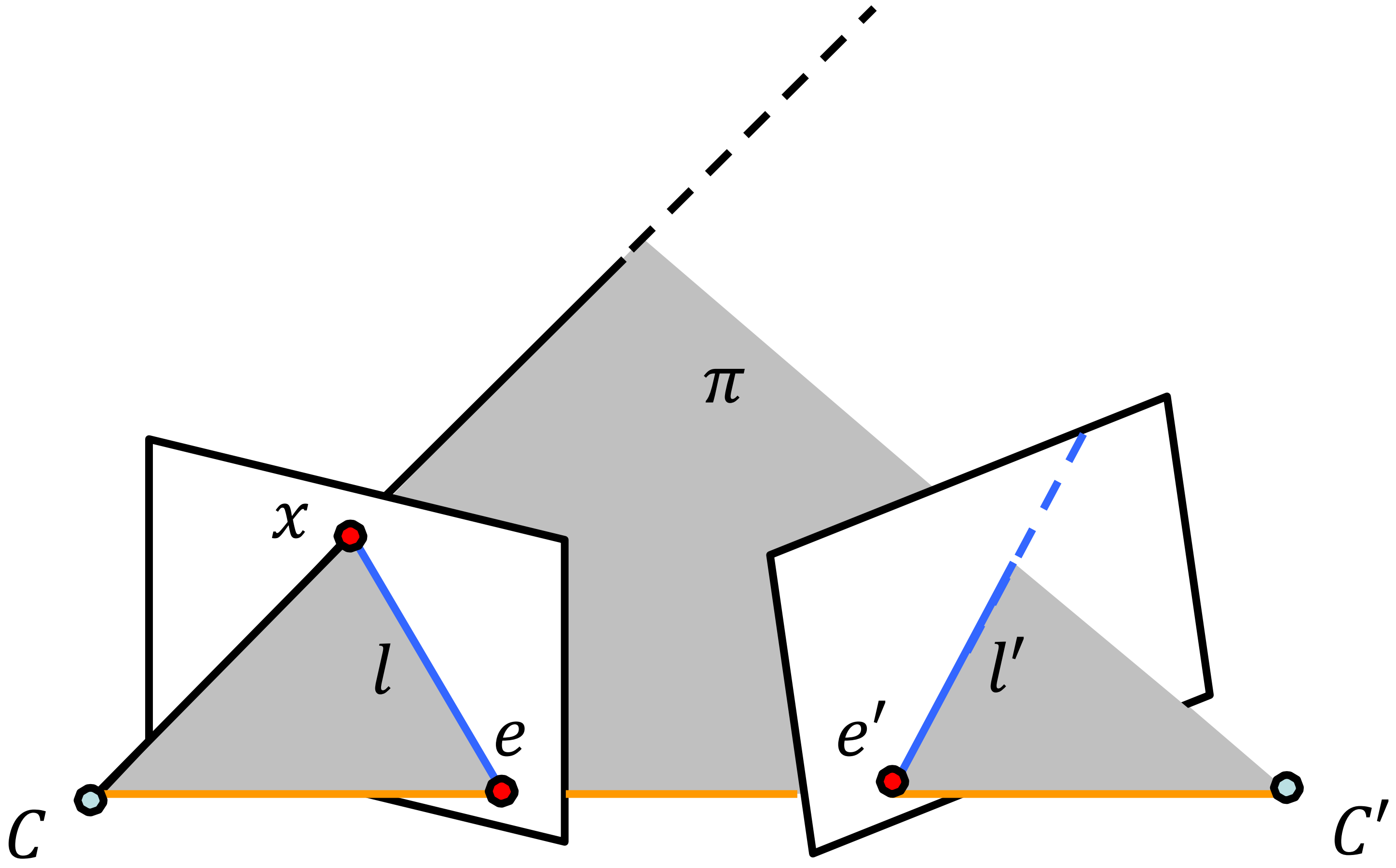


Giacomo Boracchi

Epipolar Constraints

Algorithms searching for matches between two views do not have to scan each pixel of the second image to identify the correspondence of $\boldsymbol{x}$

They just have to check the corresponding epipolar lines $\boldsymbol{l}'$



Giacomo Boracchi

The Fundamental Matrix

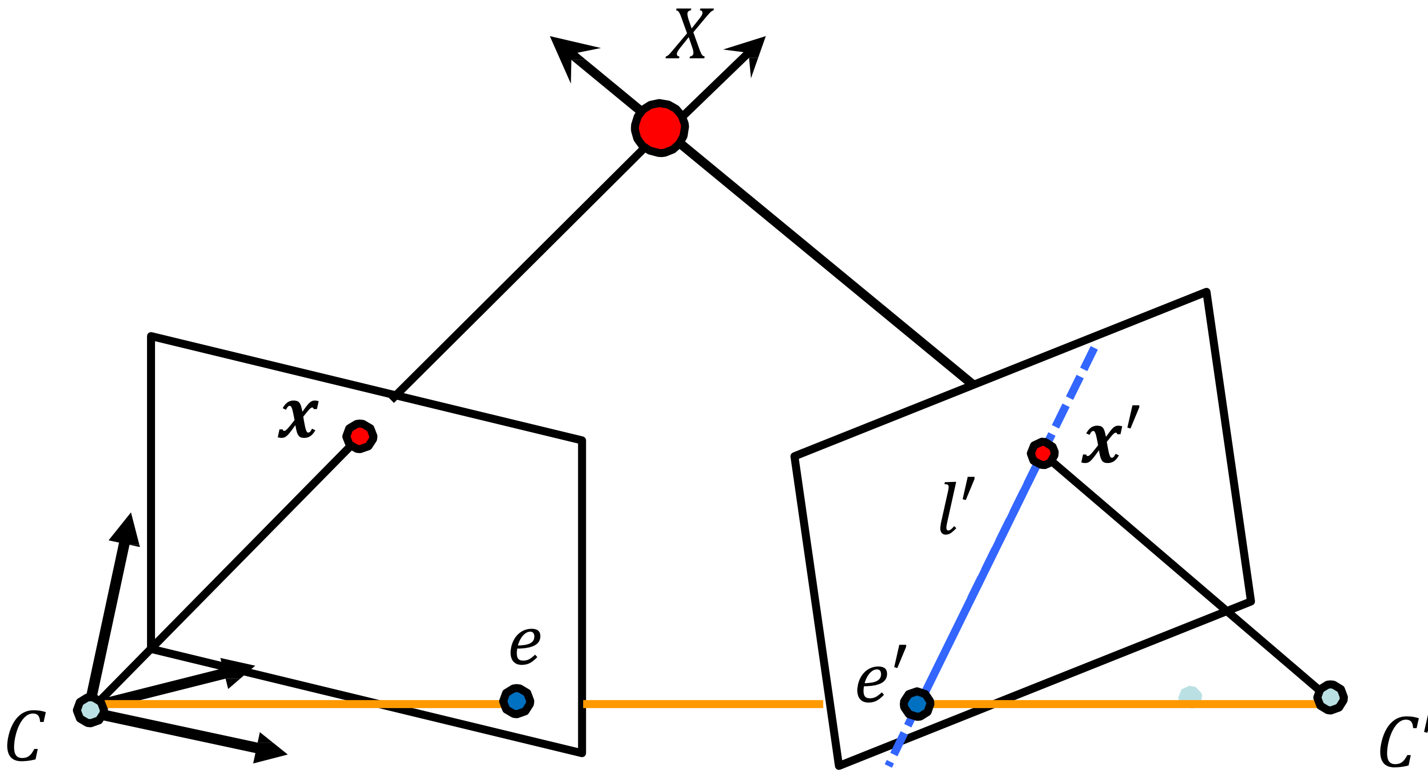
Enforcing Epipolar Constraints

Goal: Estimate an algebraic relation between $\boldsymbol{x}$ and $\boldsymbol{l}'$

$$\boldsymbol{x} \rightarrow \boldsymbol{l}'$$

Let the projection matrices M, M'

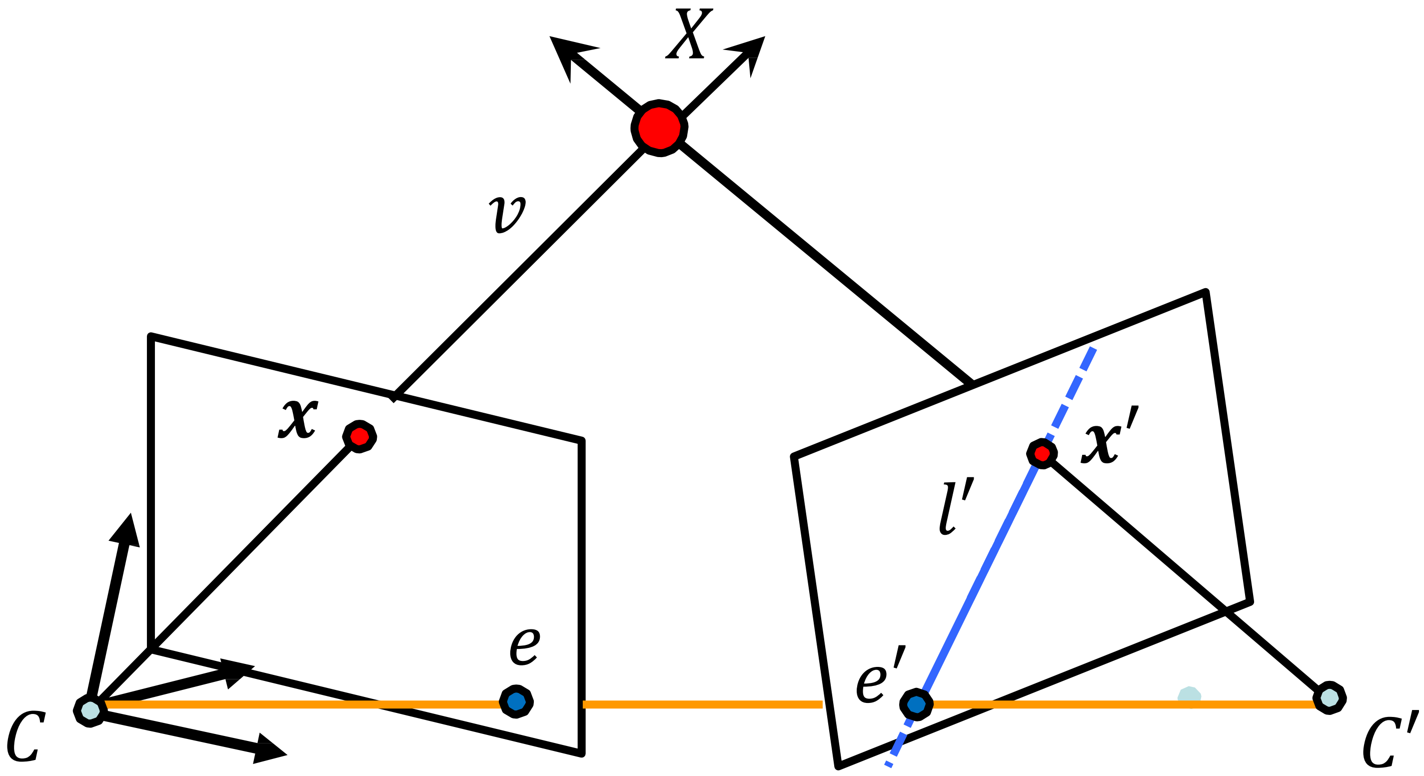
- $M = K[R, \boldsymbol{t}]$
- $M' = K'[R', \boldsymbol{t}']$



Enforcing Epipolar Constraints

Let $\boldsymbol{v}$ denote the viewing ray through $\boldsymbol{x}$.

$$\boldsymbol{v}\colon\ C + \lambda \begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}, \qquad \lambda \in \mathbb{R}$$



Enforcing Epipolar Constraints

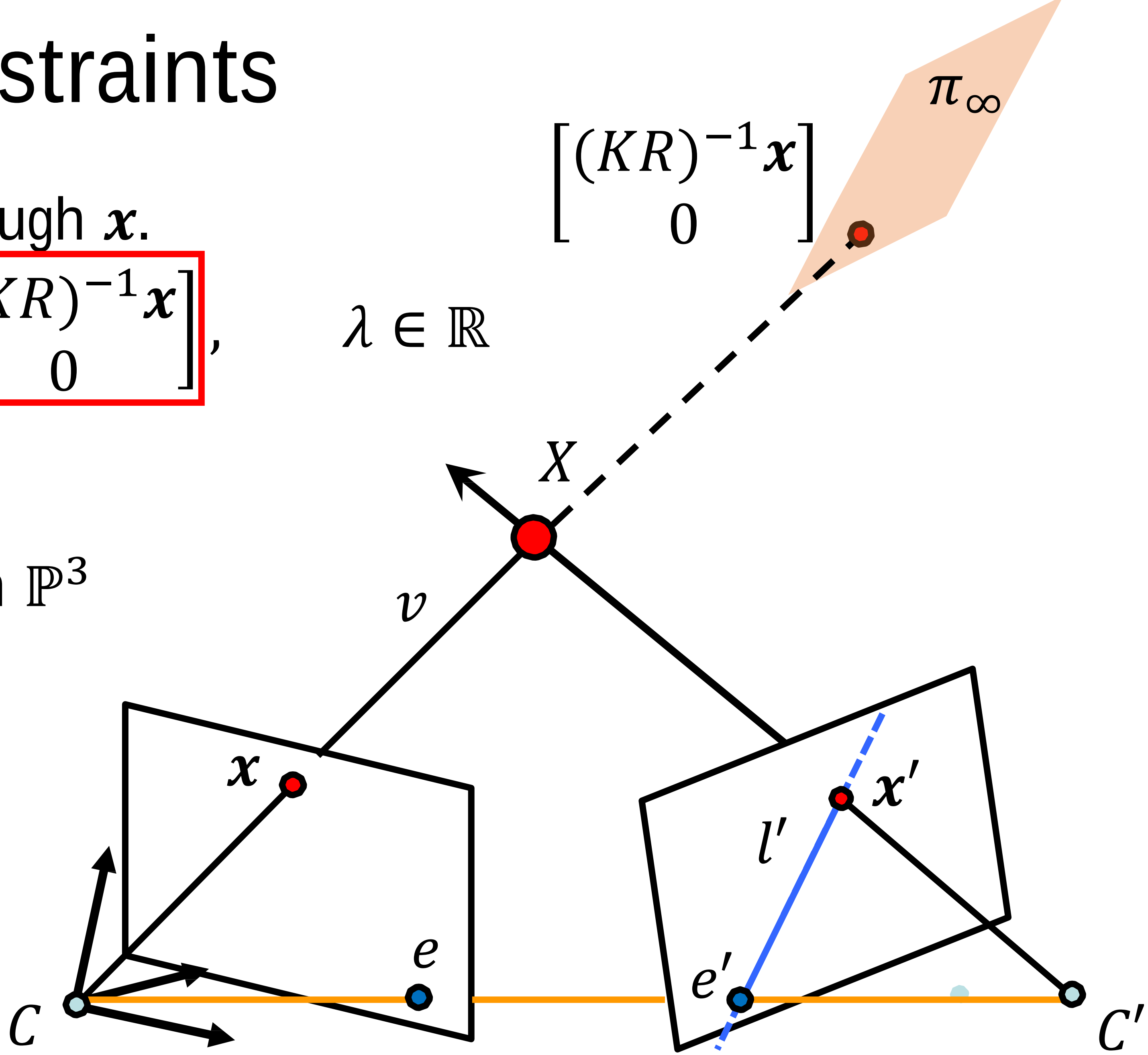
Let $\boldsymbol{v}$ denote the viewing ray through $\boldsymbol{x}$.

$$\boldsymbol{v}: \boxed{C} + \lambda \boxed{\begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}}, \qquad \lambda \in \mathbb{R}$$

there are two points involved:

- $\begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}$ is an ideal point in $\mathbb{P}^3$
- C is the camera center

any point in $\boldsymbol{v}$ can be written as linear combinaiton of these two points, since we are in $\mathbb{P}^3$

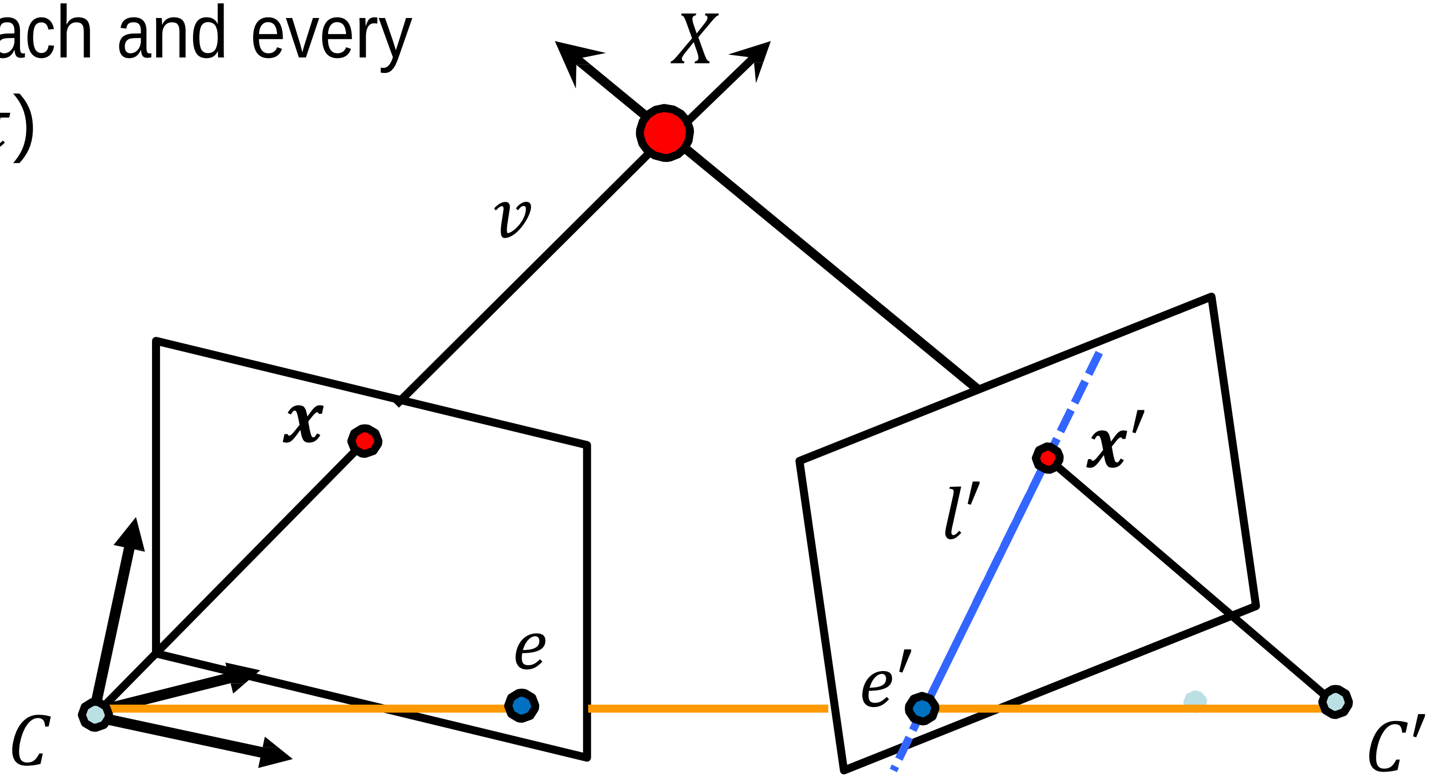


Enforcing Epipolar Constraints

Let $\boldsymbol{v}$ denote the viewing ray through $\boldsymbol{x}$

$$\boldsymbol{v}\colon\ C + \lambda \begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}, \qquad \lambda \in \mathbb{R}$$

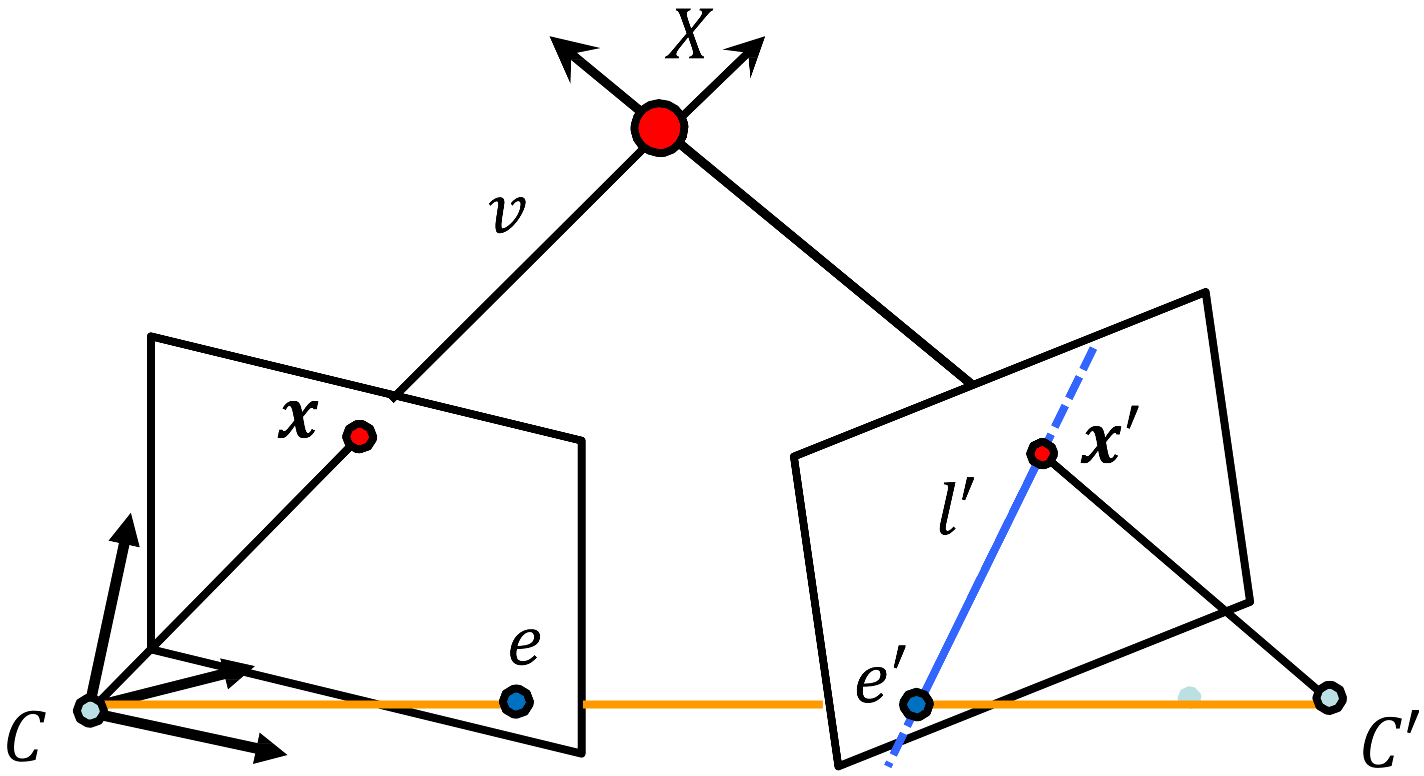
Rmk: l' is the image of $\boldsymbol{v}$ (as of each and every line living on the epipolar plane π)



Enforcing Epipolar Constraints

Now, project these two points through M'

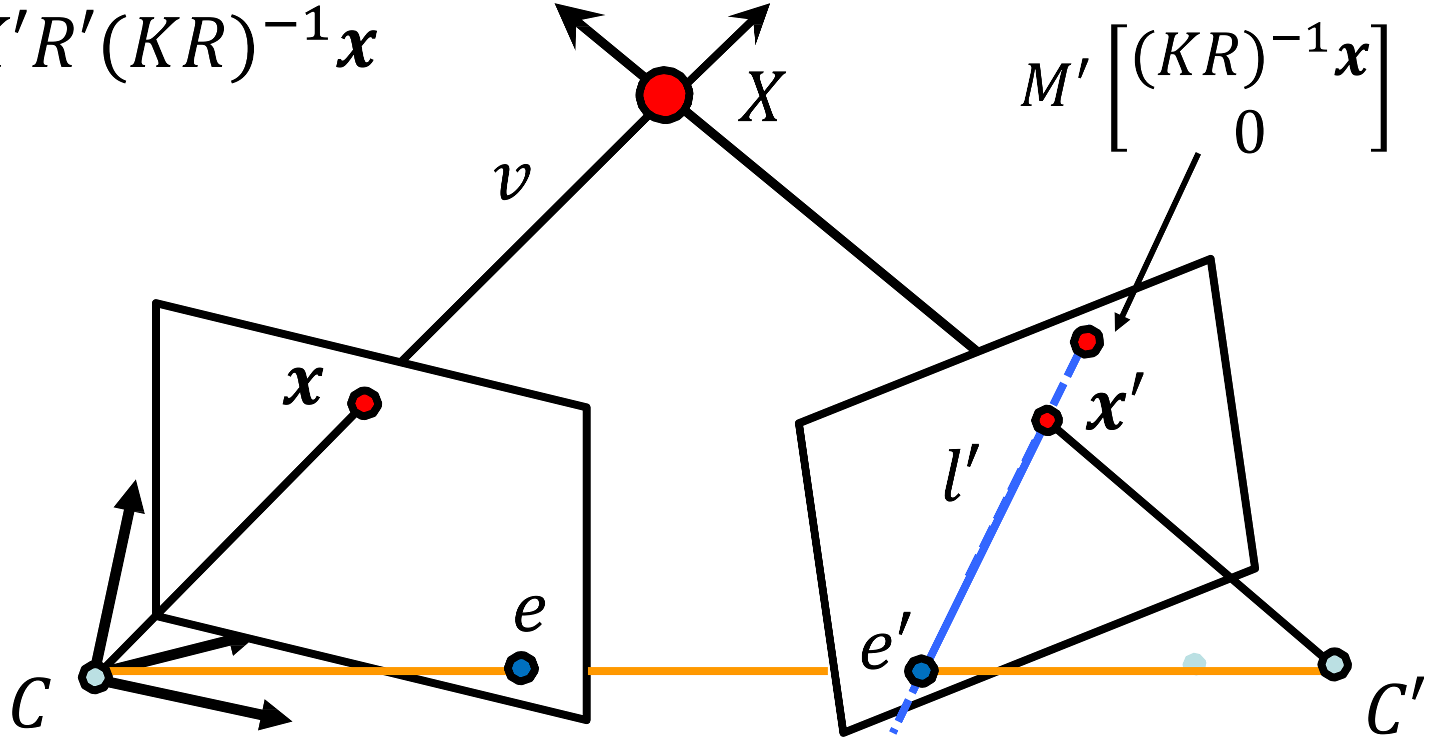
- $C \rightarrow M'C = e'$ (by definition)
- $\begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix} \rightarrow M' \begin{bmatrix} (KR)^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}$



Enforcing Epipolar Constraints

$M'\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\0\end{bmatrix}$ is the vanishing point of $\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\0\end{bmatrix}$ (it is an ideal point)

$$M'\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\0\end{bmatrix} = \lambda K'[R',\boldsymbol{t}']\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\0\end{bmatrix}$$
$$= \lambda K'R'(KR)^{-1}\boldsymbol{x}$$



Projection through M' implies

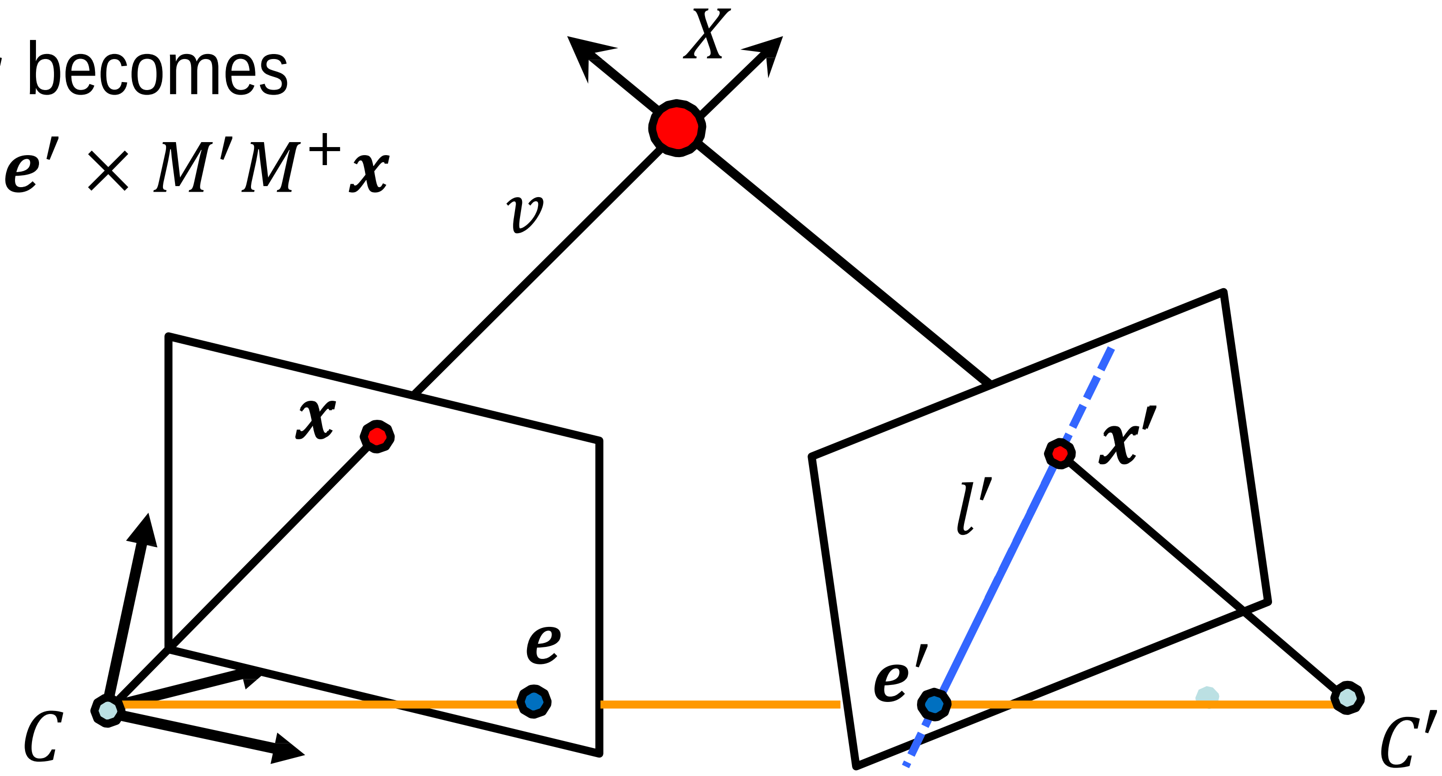
- $C \rightarrow M'C = e'$
- $\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\0\end{bmatrix} \rightarrow K'R'(KR)^{-1}\boldsymbol{x}$

Point to Line Mapping in Epipolar Geometry

Now, $K'R'(KR)^{-1}$ can be also written as $M'M^{+}$, being M^{+} the pseudo inverse of M, namely $MM^{+} = I_3$

The line $\boldsymbol{l}'$ joining $\boldsymbol{e}'$ and $\boldsymbol{M}'\boldsymbol{M}^{+}\boldsymbol{x}$ becomes

$$l' = \boldsymbol{e}' \times M'M^{+}\boldsymbol{x}$$



Cross Product as Matrix Multiplication

The cross product against can be written in as product against a special antisymmetric matrix

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}$$

Thus the matrix associated to the cross product against e' is

$$[\boldsymbol{e}']_\times = \begin{bmatrix} 0 & -e'_z & e'_y \\ e'_z & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{bmatrix}$$

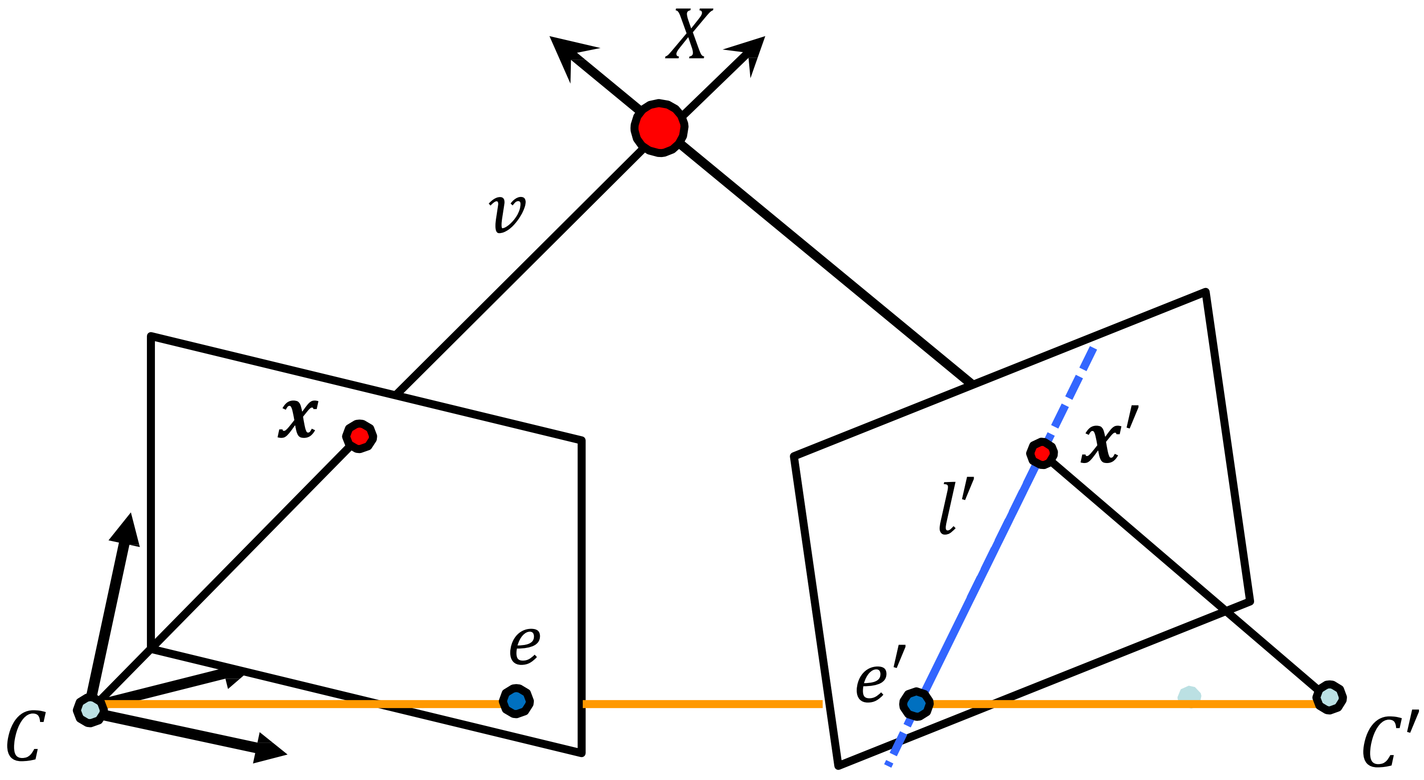
The Fundamental Matrix

So, l' is the line joining e' and $M'M^{+}\boldsymbol{x}$

$$l' = \boldsymbol{e}' \times M'M^{+}\boldsymbol{x}$$
$$l' = [\boldsymbol{e}']_{\times} M'M^{+}\boldsymbol{x}$$

The matrix $\boldsymbol{F} = [\boldsymbol{e}']_{\times}\boldsymbol{M}'\boldsymbol{M}^{+}$

is called the fundamental matrix



Epipolar Constraints Through F

The epipolar line for a point $\boldsymbol{x}$ is

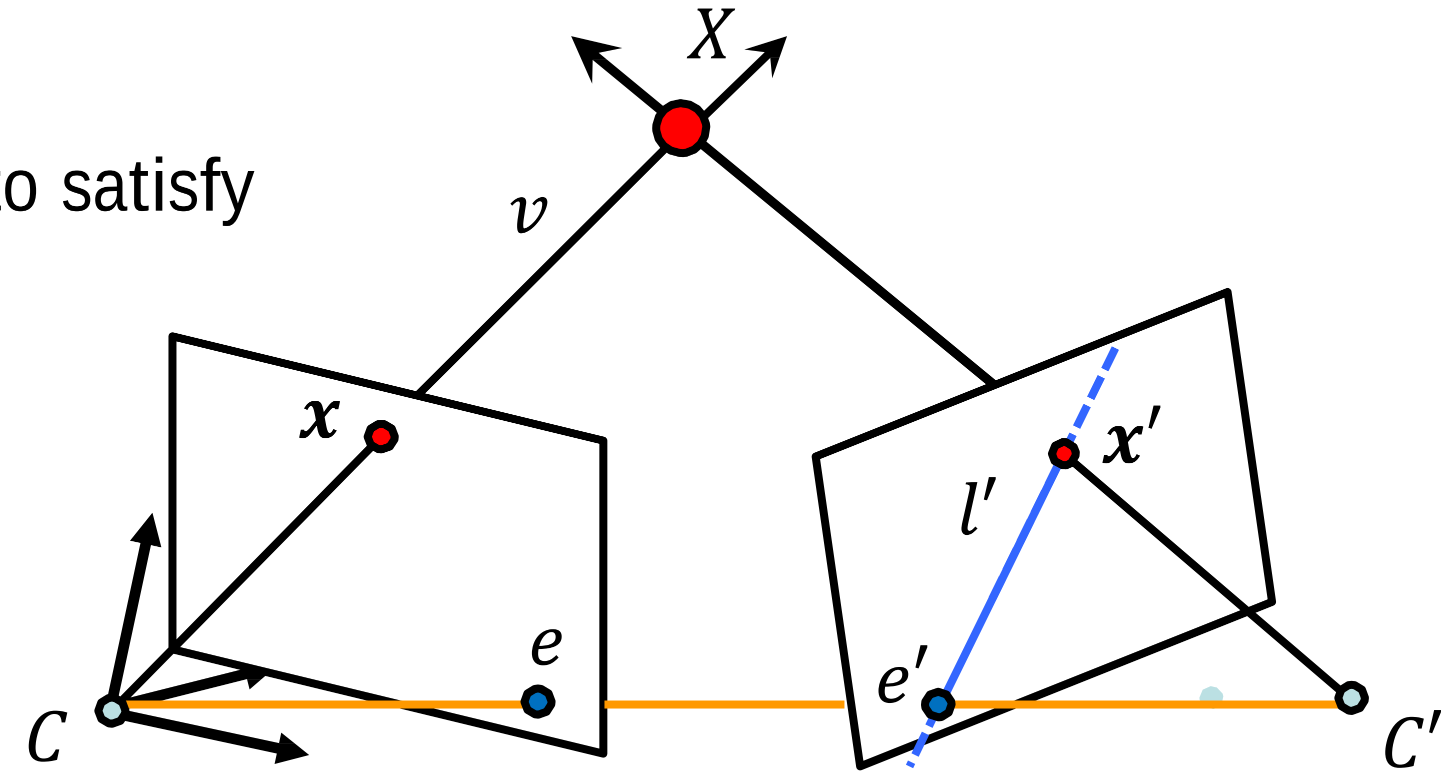
$$\boldsymbol{l}' = F\boldsymbol{x}$$

The incidence relation $\boldsymbol{x}' \in \boldsymbol{l}'$ implies

$$\boldsymbol{x}'^{\top} \boldsymbol{l}' = 0$$

Thus corresponding points have to satisfy

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



The Fundamental Matrix

Definition: the fundamental matrix F is the unique 3×3 rank 2 homogeneous matrix which satisfies

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

for all corresponding points $\boldsymbol{x}'^{\top} \leftrightarrow \boldsymbol{x}$ in the two images

Rmk: the fact that rank is not full follows from

$$F = [\boldsymbol{e}']_{\times} M' M^{+}$$

And $\det(F) = \det([\boldsymbol{e}']_{\times} M' M^{+}) = \det([\boldsymbol{e}']_{\times}) * \det(M' M^{+}) = 0$ since $\det([\boldsymbol{e}']_{\times}) = 0$

Multi-View Geometry

Giacomo Boracchi

CVPR USI, March 17 2020

Book: HZ, chapters 9, 11, 12

The Fundamental Matrix: Recap

The epipolar line for a point $\boldsymbol{x}$ is

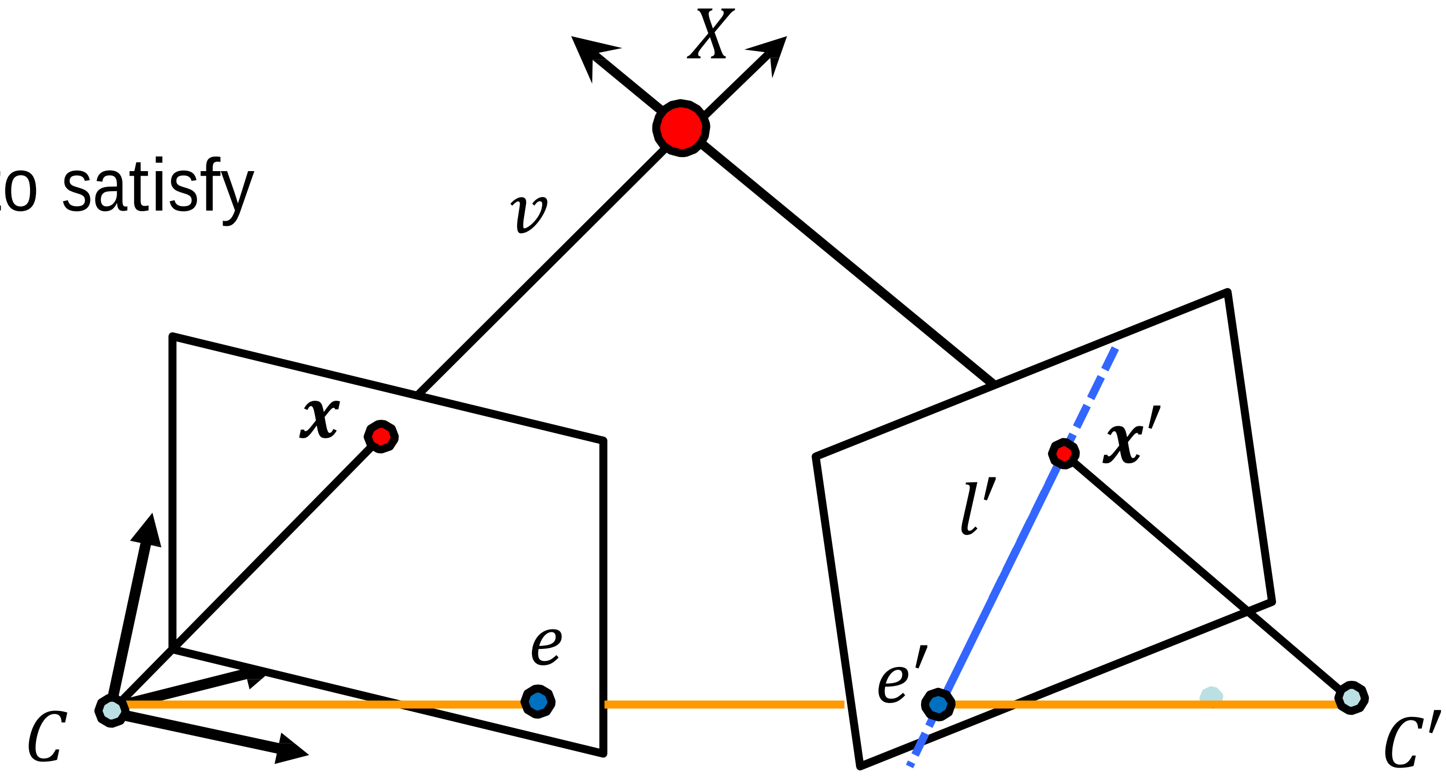
$$\boldsymbol{l}' = F\boldsymbol{x}$$

The incidence relation $\boldsymbol{x}' \in \boldsymbol{l}'$ implies

$$\boldsymbol{x}'^{\top}\, \boldsymbol{l}' = 0$$

Thus corresponding points have to satisfy

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



The Fundamental Matrix

The fundamental matrix $F = [\boldsymbol{e}']_\times M' M^+$ follows from a point-line relation between the two views based only on their camera matrices

The fundamental matrix represents the condition that corresponding points $\boldsymbol{x}$ and $\boldsymbol{x}'$ have to satisfy in the camera system

- This property enables computing $\boldsymbol{F}$ from pairs of corresponding points, without having to known $\boldsymbol{M}$ and $\boldsymbol{M}'$. So, even when cameras have not been calibrated!

The Fundamental Matrix

The fundamental matrix $F = [\boldsymbol{e}']_\times M' M^+$ follows from a point-line relation between the two views based only on their camera matrices

The fundamental matrix represents the condition that corresponding points $\boldsymbol{x}$ and $\boldsymbol{x}'$ have to satisfy in the camera system

- This property enables computing $\boldsymbol{F}$ from pairs of corresponding points, without having to known $\boldsymbol{M}$ and $\boldsymbol{M}'$. So, even when cameras have not been calibrated!

When the two centres C and C' coincides, i.e., $C = C'$, then

$$M'C = MC' = \mathbf{0}$$

and the derivation breaks down.

In this case there is an homography relating the two images!

Properties of the Fundamental Matrix

Transpose: If F is the fundamental matrix of the pair of cameras M, M', then $F^\top$ is the fundamental matrix of the pair in the opposite order: M', M

Epipolar lines: For any point $\boldsymbol{x}$ in the first image, the corresponding epipolar line is $l' = F\boldsymbol{x}$. Similarly, $l = F^\top \boldsymbol{x}'$ represents the epipolar line corresponding to $\boldsymbol{x}'$ in the second image.

Fundamental Matrix → Epipoles

The epipole: for any point $\boldsymbol{x} \neq \boldsymbol{e}$ the epipolar line $\boldsymbol{l}' = F\boldsymbol{x}$ contains the epipole $\boldsymbol{e}'$. Thus, $\boldsymbol{e}'$ satisfies the incidence equation $\forall \boldsymbol{x}$

$$\boldsymbol{e}'^{\top}\boldsymbol{l}' = 0 \quad \forall \boldsymbol{x}$$

$$\boldsymbol{e}'^{\top}(F\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x}$$

$$(\boldsymbol{e}'^{\top}F)\,\boldsymbol{x} = 0, \qquad \forall \boldsymbol{x}$$

it follows that $(\boldsymbol{e}'^{\top}F) = \boldsymbol{0}$, thus that $\boldsymbol{e}'$ is the left null-vector of $\boldsymbol{F}$.
Similarly $F\boldsymbol{e} = \boldsymbol{0}$, i.e. $\boldsymbol{e}$ is the right null-vector of F.

$$\boldsymbol{e}' = LNS(F)$$

$$\boldsymbol{e} = RNS(F)$$

Fundamental Matrix $\rightarrow$ Camera Matrix

Rmk Relations established by fundamental matrix are projectively invariant. Namely, the fundamental matrix is unchanged by projective transformations of $\mathbb{P}^3$ (the projective space)

The fundamental matrix determines the pair of camera matrices M, M' up to right-multiplication by a 3D projective transformation $H \in R^{4\times 4}$.

There are infinite many decompositions of a given F into a pair of camera matrices M, M'

Properties of Fundamental Matrix

It is common to define a specific *canonical form* for the pair of camera matrices M, M' corresponding to a given fundamental matrix F:

- The first matrix is of the simple form

$$M = [I \mid \mathbf{0}]$$

- The second matrix can be any

$$M' = \big[[\boldsymbol{e}']_\times F + \boldsymbol{e}'\boldsymbol{v}^\intercal, \lambda\boldsymbol{e}'\big]$$

where $\boldsymbol{v}$ is any 3-vector, and $\lambda \neq 0$

Among all the possible versions of M', one typically chooses

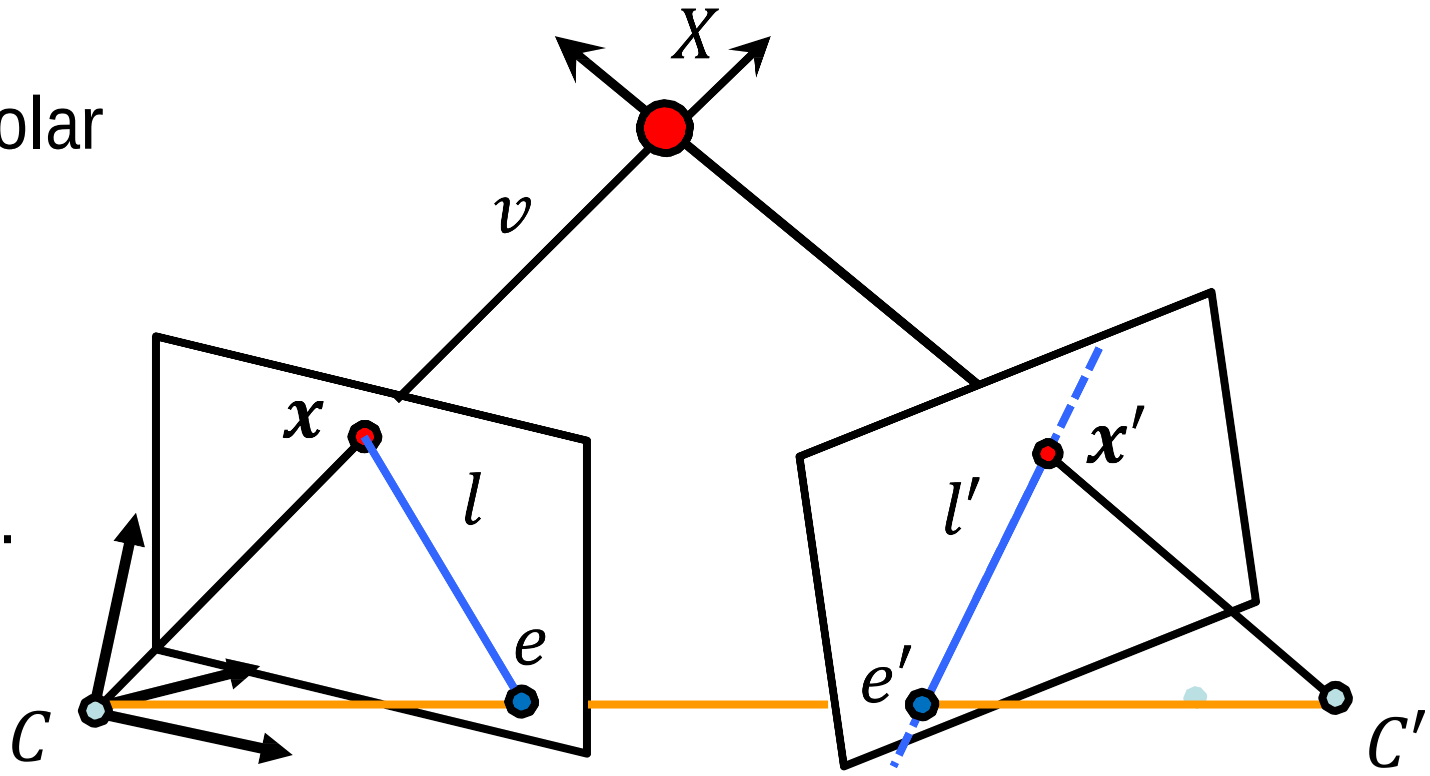
$$M' = [[\boldsymbol{e}']_\times F \mid \boldsymbol{e}']$$

The Fundamental Matrix as a correlation

F is a correlation, a projective map taking a point to a line. A point in the first image $\boldsymbol{x}$ defines the epipolar line in the second $\boldsymbol{l}' = F\boldsymbol{x}$.

If $\boldsymbol{l}$ and $\boldsymbol{l}'$ are corresponding epipolar lines then any point $\boldsymbol{x}$ on $\boldsymbol{l}$ is mapped to the same line $\boldsymbol{l}'$.

This means there is no inverse mapping, and F is not of full rank. For this reason, F is not a proper correlation (which would be invertible).



Matlab Example

Estimating The Fundamental Matrix

Estimating the Fundamental Matrix

The fundamental matrix F represents the condition for pair of corresponding points $(\boldsymbol{x_i}, \boldsymbol{x_i'})$ in the camera system

$$\boldsymbol{x_i'}^\top F \boldsymbol{x_i} = 0 \quad \forall i$$

Practical and efficient algorithms for computing $\boldsymbol{F}$ can be derived from a set of point correspondences in the two images

- Derive equations from point correspondences in a linear system where the unknown are the 9 entries of the matrix F
- No need to estimate $\boldsymbol{M}$ and $\boldsymbol{M}'$ first.

So, F can be computed and enforced even when cameras have not been calibrated!

How many degrees of freedom?

The fundamental matrix F has $3 \times 3 = 9$ parameters, plus

- It is a homogeneous matrix, so it is invariant for scaling (we can fix a term)
- F is not full rank and $\text{rank}(F) = 2$, so $\det(F) = 0$.

So, there are only 7 degrees of freedom.

How to Estimate F?

For any pair of matching points $(\boldsymbol{x_i}, \boldsymbol{x_i'})$ satisfies

$$\boldsymbol{x_i'}^\top F \boldsymbol{x_i} = 0$$

This is a single scalar equation

If we write $\boldsymbol{x_i} = [x_i; y_i; 1]$ and $\boldsymbol{x_i'} = [x_i'; y_i'; 1]$, then

$$[x_i', y_i', 1] \begin{bmatrix} f_{1,1} & f_{1,2} & f_{1,3} \\ f_{2,1} & f_{2,2} & f_{2,3} \\ f_{3,1} & f_{3,2} & f_{3,3} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0$$

Which yields a linear equation in the unknowns $\{f_{i,j}\},\ i, j = 1, \ldots 3$

We need $n > 7$ of these correspondences to compute F

Estimating F from point correspondences

Let $\boldsymbol{f} = F^{\top}(:)$ (unrolled in row-major order), the equation becomes

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1)\,\boldsymbol{f} \;=\; 0$$

From n correspondences $(\boldsymbol{x_i}, \boldsymbol{x_i'})$ we obtain an homogeneous linear system in $\boldsymbol{f}$

$$A_n\boldsymbol{f} = \begin{bmatrix} x_1'x_1, x_1'y_1, x_1', y_1'x_1, y_1'y_1, y_1', x_1, y_1, 1 \\ \dots \\ x_n'x_n, x_n'y_n, x_n', y_n'x_n, y_n'y_n, y_n', x_n, y_n, 1 \end{bmatrix}\boldsymbol{f} = \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix}$$

where $A_n \in \mathbb{R}^{n\times 9}$

The 8-Point Algorithm

Estimating $\boldsymbol{F}$ from point correspondences

For a solution to exist, matrix A_n must have rank at most 8,

When the rank is exactly 8, the solution is unique (up to scale) and can be found as the generator of the right null-space of A_n

$$\boldsymbol{f} \in RNS\,(A_8)$$

When the $\text{rank}(A_n) = 9$ we look for the nontrivial solution, by minimizing $\left||A_n \boldsymbol{f}|\right|_{\boldsymbol{2}}$ subject to the constraint $||\boldsymbol{f}|| = 1$ (its' an homogeneous matrix).

$$\boldsymbol{f} \in RNS\,(A_n) \text{ and } ||\boldsymbol{f}|| = 1$$

The DLT algorithm can solve this problem!

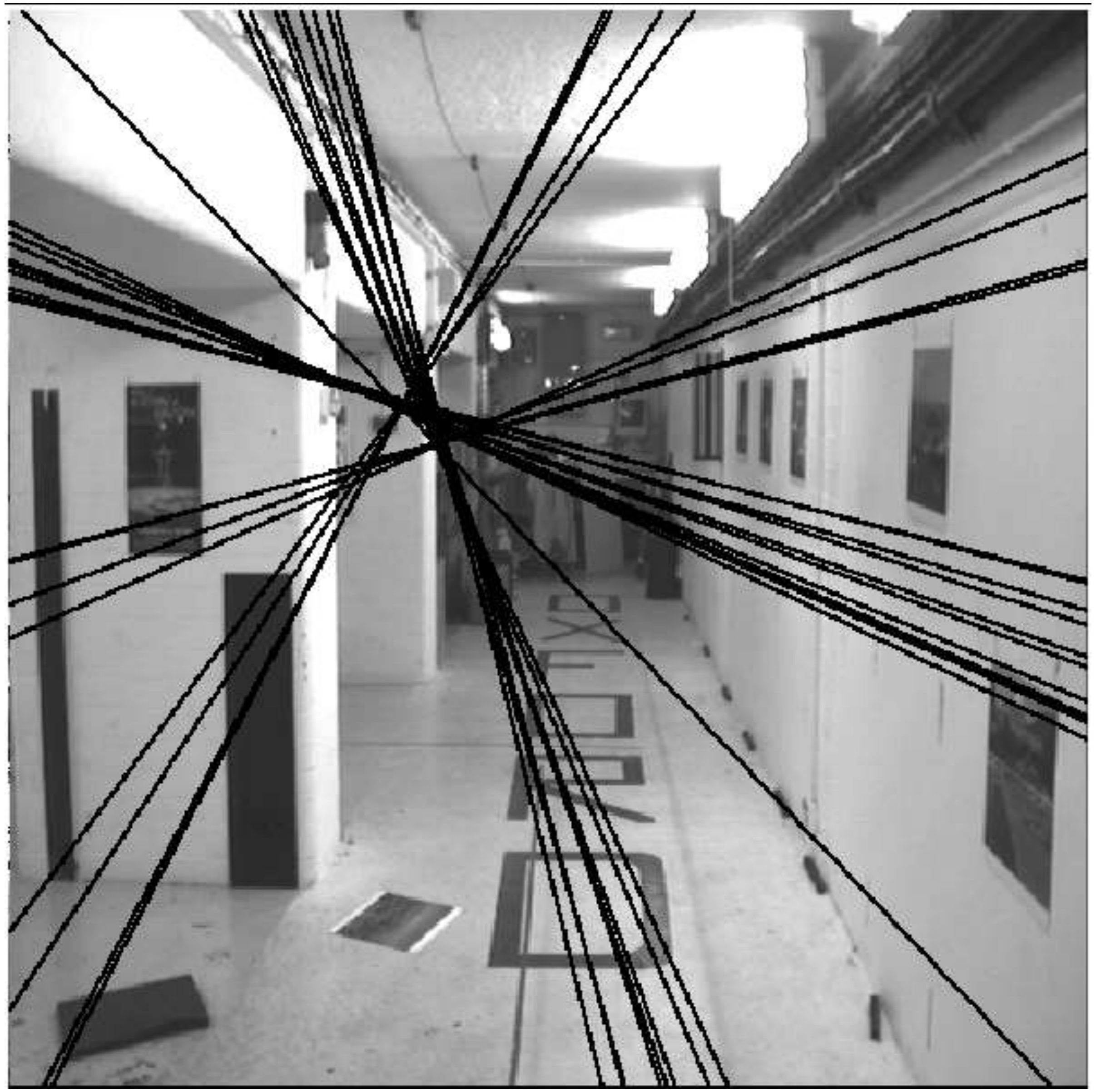
This holds also for the overdetermined case when $n > 9$

Enforcing $\mathrm{rank}(F) = 2$

The matrix F estimated by the DLT algorithm will in general have rank 3.
However applications involving F typically require $\mathrm{rank}(F) = 2$

- The $\mathrm{rank}(F) = 2$ constraint needs to be enforced

Epipolar lines without enforcing $\mathrm{rank}(F) = 2$



When $\mathrm{rank}(F) = 3$
The kernel is trivial, so it is not possible that

$$\boldsymbol{x}'^{\top} F \boldsymbol{e} = 0 \ \forall \boldsymbol{x}'$$

Thus, it is not possible for all the epipolar line to pass through $\boldsymbol{e}$

Hartley Zisserman Fig.11.2

Enforcing $\text{rank}(F) = 2$

The matrix F provided by DLT is replaced by the matrix $\tilde{F}$ such that

$$\text{argmin} \left|\left|F - \tilde{F}\right|\right|_F$$

subject to the condition $\text{rank}(\tilde{F}) = 2$

Where $||\cdot||_F$ is the Frobenius norm (the ℓ^2 norm of the matrix entries)

The best way to enforce this constraint is through the $\boldsymbol{SVD(F)}$

Enforcing Rank Constraint via SVD

Eckart-Young Theorem

Let $A \in \mathbb{R}^{m,n}$ be a matrix such that $\mathrm{rank}(A) = r$ and let $A = U_r D_r V_r^\top$. Then the closest approximation in terms of Frobenius norm to A having rank $k < r$ is

$$A_k = U_r D_k V_r^\top$$

where $D_k = \mathrm{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0)$

Moreover $\left|\left|A - A_k\right|\right|_F = \sigma_{k+1}$

Epipolar lines enforcing $\text{rank}(F) = 2$

Hartley Zisserman Fig.11.2

The Eight-point algorithm

Therefore, we can compute F as mentioned above, and then set to zero the last eigenvalue

$x_i'^T F x_i = 0$ (x_i, x_i') $i = 1 - 8$ or more

$\boxed{A_n} f = 0$

$\bar{F}$ $\det(F) = 0$

Two steps:

- Solving the linear homogeneous system via the DLT $F \in \mathbb{R}^{3 \times 3}$

$$F = \underline{\mathrm{DLT}(\boxed{A_n})}$$

- Enforcing $\underline{\mathrm{rank}(\boxed{F}) = 2}$ constraint

$$\tilde{F} = UD_2V', \text{where } F = UDV'$$

The Seven-Point Algorithm

When $\text{rank}(A_n) \geq 8$ the system is solved through the DLT by leveraging the homogeneity constraint

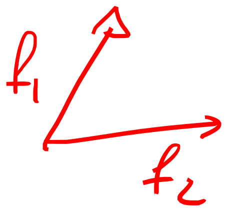
When $\text{rank}(A_n) = 7$ and in particular when $\text{rank}(A_7) = 7$ (i.e. only $n = 7$ point correspondences are given), it is still possible to solve the sytem by taking into account the singularity constraint of F

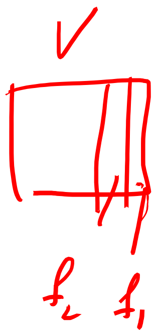
This gives rise to the seven point algorithm

The Seven-Point Algorithm

When $\mathrm{rank}(A_7) = 7$ the solutions of the system

$$\underset{7\times 9}{A_7}\,\underset{9\times 1}{\boldsymbol{f}} = \mathbf{0}$$





Form a 2-dimensional space of the type

$$F_\alpha = \boxed{\alpha F_1 + (1-\alpha)F_2} \quad \forall \alpha \in \mathbb{R}$$

where F_1 and F_2 matrixes correspond to the last and second last column of V in the SVD.

The right value of α can be found by enforcing the singularity constraint

$$\det(F_\alpha) = 0 \;\Rightarrow\; \det(\alpha F_1 + (1-\alpha)F_2) = 0$$

Which yields a third order equation in $\boxed{\alpha,}$ giving one or three real solutions (ignore the complex conjugate)

Normalization or Preconditioning

DLT and the reference system

Are the outcome of DLT independent of the reference system being used to express $\boldsymbol{x}'$ and $\boldsymbol{x}$?

Unfortunately DLT is not invariant to similarity transformations.

Therefore, it is necessary to apply a normalizing transformation to the data before applying the DLT algorithm.

Normalizing the data makes the DLT invariant to the reference system, as it is always being estimated in a canonical reference

Normalization is also called Pre-conditioning

$(x_i, x_i') \mapsto (\tilde{x}_i, \tilde{x}_i')$

DLT

Preconditioning

Needed because in homogeneous coordinate systems, components typically have very different ranges

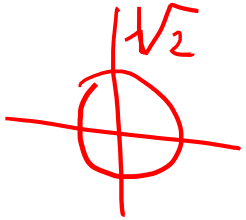
- Row and Column indexes in $\boldsymbol{x}$ ranges in $[0 - 4K]$
- Third component is 1

0 - 4K

$$\begin{bmatrix} r \\ c \\ 1 \end{bmatrix}$$

Preconditioning corresponds to defining a mapping

$$\boldsymbol{x} \rightarrow T\boldsymbol{x}$$



That brings all the points «around the origin» and rescale each component to the same range (say at an average distance $\sqrt{2}$ from the origin)

Preconditioning

I $\quad$ I'

$X = \{x \ldots\}$ $\quad$ $X' = \{x' \ldots\}$

T $\quad$ T'

$$\boldsymbol{x} \mapsto T\boldsymbol{x}$$

It's a similarity scaling in $\mathbb{P}^2$ (with $\theta = 0$ since we do not need rotations)

$$T = \begin{bmatrix} 1/s & 0 & -t_x/s \\ 0 & 1/s & -t_y/s \\ 0 & 0 & 1 \end{bmatrix}$$

The preconditioning for a set of points X of $\mathbb{P}^2$ is defined as

$$t_x = \text{mean}(X(1,:))$$
$$t_y = \text{mean}(X(2,:))$$

Which brings the barycentre of X to the origin,

The scaling coefficient becomes

$$s = \frac{\text{mean}\big(\text{std}(X,2)\big)}{\sqrt{2}}$$

Homography estimation with preconditioning

Estimate the fundamental matrix F (or homography H) between two sets of points X, X'

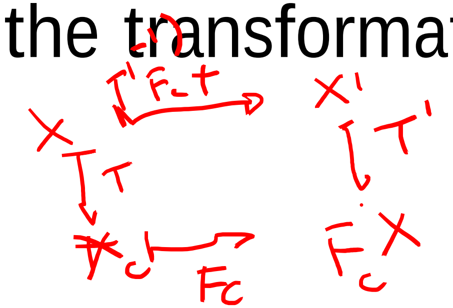
1. Compute T, T' preconditioning transformation of X, X'
2. Apply transformation

$$X_c = TX, \qquad X'_c = T'X'$$

3. Estimate the Fundamental matrix (or Homography) from X_c and X'_c,

$$\boxed{F_c} = DLT(X_c, X'_c)$$

4. Define the transformation between original points



$$F = T'^T F_c T$$

The Essential Matrix

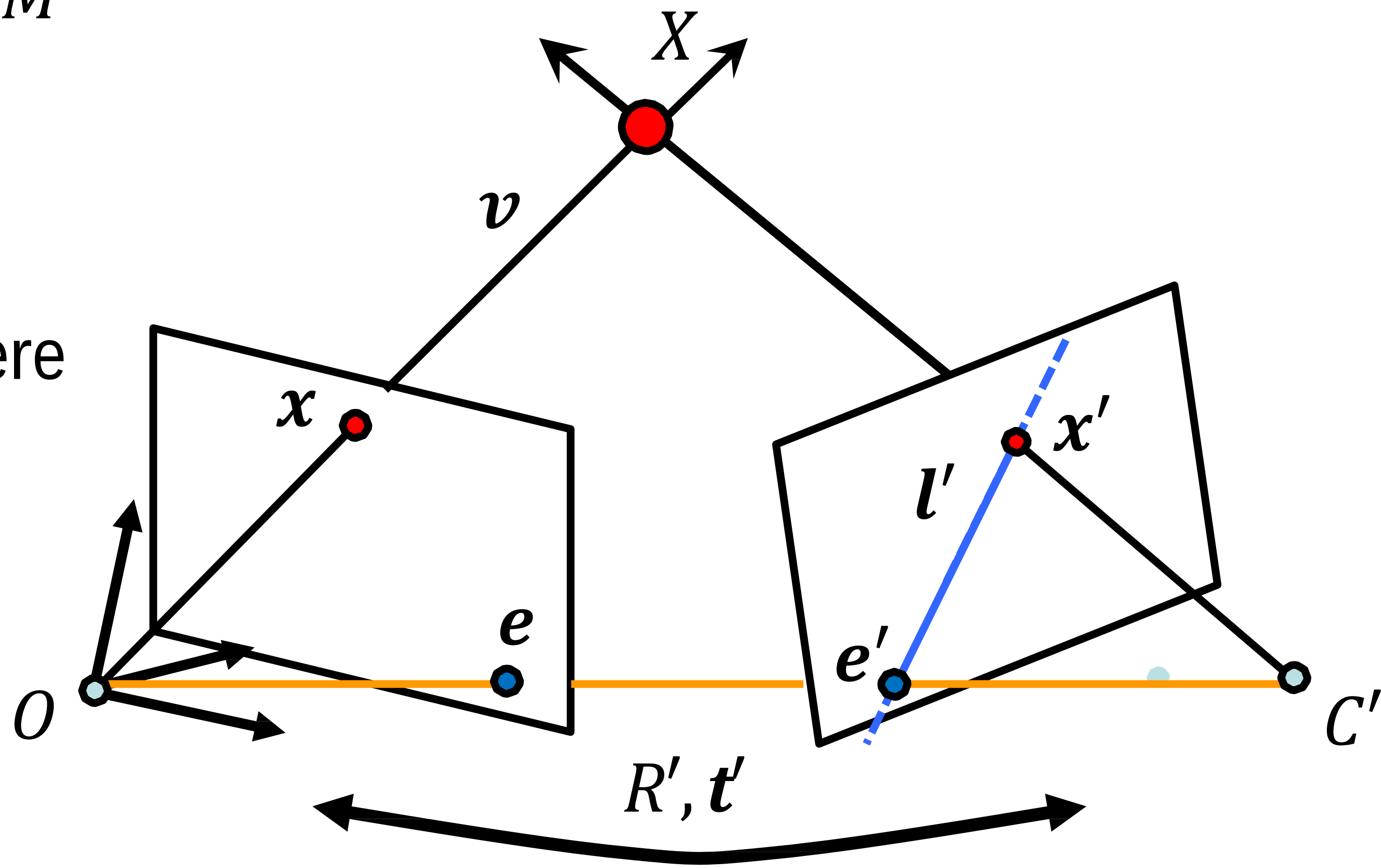
Normalized Epipolar Geometry Setup

Place the origin of world 3D reference system is in the first camera center ($O = C$)

Define the projection matrices M, M'

- $M = K[I_3, 0]$
- $M' = K'[R', \boldsymbol{t}']$

And consider the simple case where $K = K' = I_3$



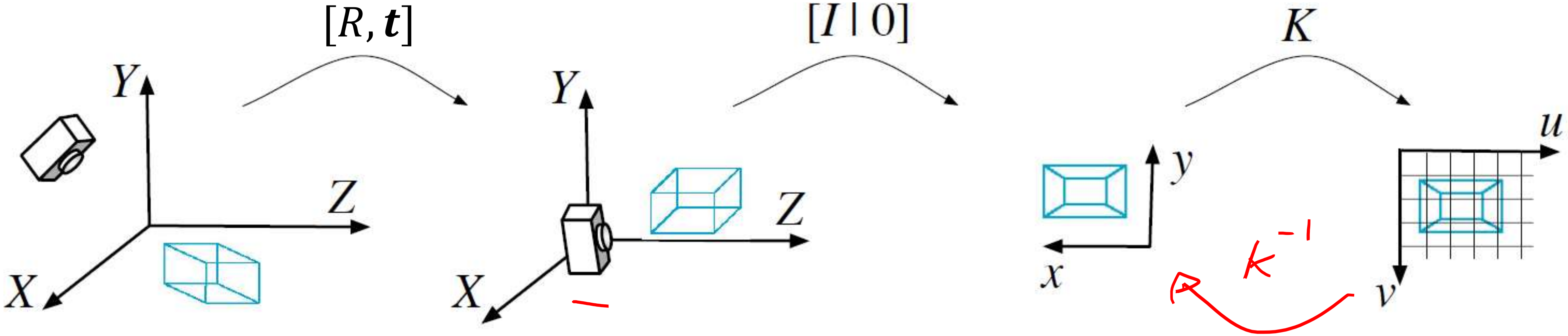
The Essential Matrix: Normalized Cameras

We can get to these settings by normalizing each camera through:

$$\hat{x} = K^{-1}x$$

Ad expressing everything w.r.t $\hat{x}$.

The Coordinate Systems Involved in Camera Projection



World 3D reference system *[meters]*

Camera 3D reference system *[meters]*

2D Normalized Coordinates *[meters]*

Image Coordinates *[pixels]*

The Essential Matrix: Normalized Cameras

We can get to these settings by normalizing each camera through:

$$\hat{x} = K^{-1}x$$

Ad expressing everything w.r.t $\hat{x}$.

The essential matrix E follows from the epipolar constraints being enforced in a pair of normalized cameras $\hat{\boldsymbol{x}} = K^{-1}\boldsymbol{x}$ and $\hat{\boldsymbol{x}}' = K'^{-1}\boldsymbol{x}'$

The epipolar setup with normalized cameras is

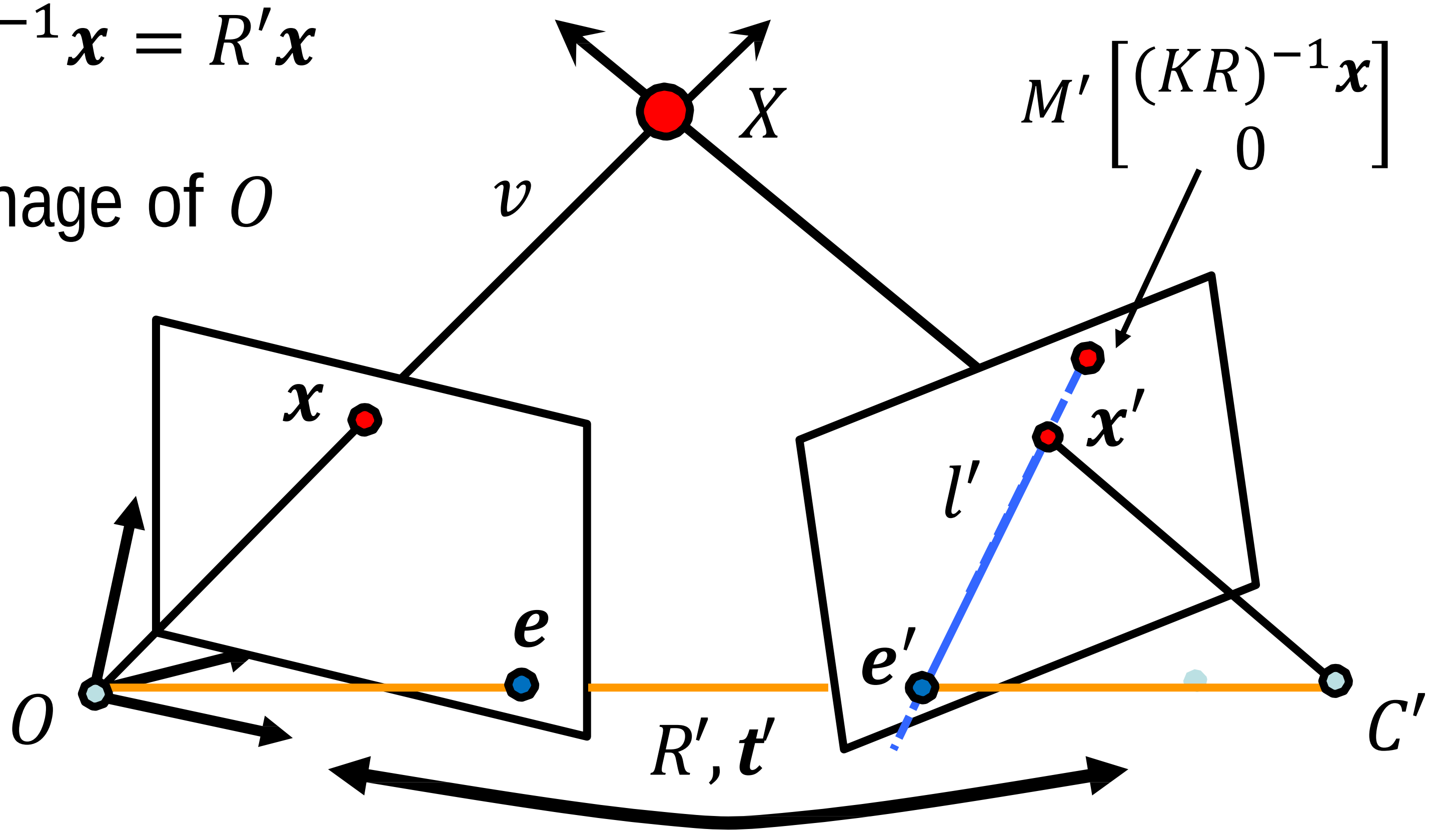
- $M = [I_3, 0]$
- $M' = [R', \boldsymbol{t}']$

The Essential Matrix: Derivation

Following up the previous calculus for the fundamental matrix, we draw the line joining $M'\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\ 0\end{bmatrix}$ and $\boldsymbol{e}'$

Being $M'\begin{bmatrix}(KR)^{-1}\boldsymbol{x}\\ 0\end{bmatrix} = K'R'(KR)^{-1}\boldsymbol{x} = R'\boldsymbol{x}$

And $\boldsymbol{e}' = \boldsymbol{t}'$ because this is the image of O through $M' = [R', \boldsymbol{t}']$



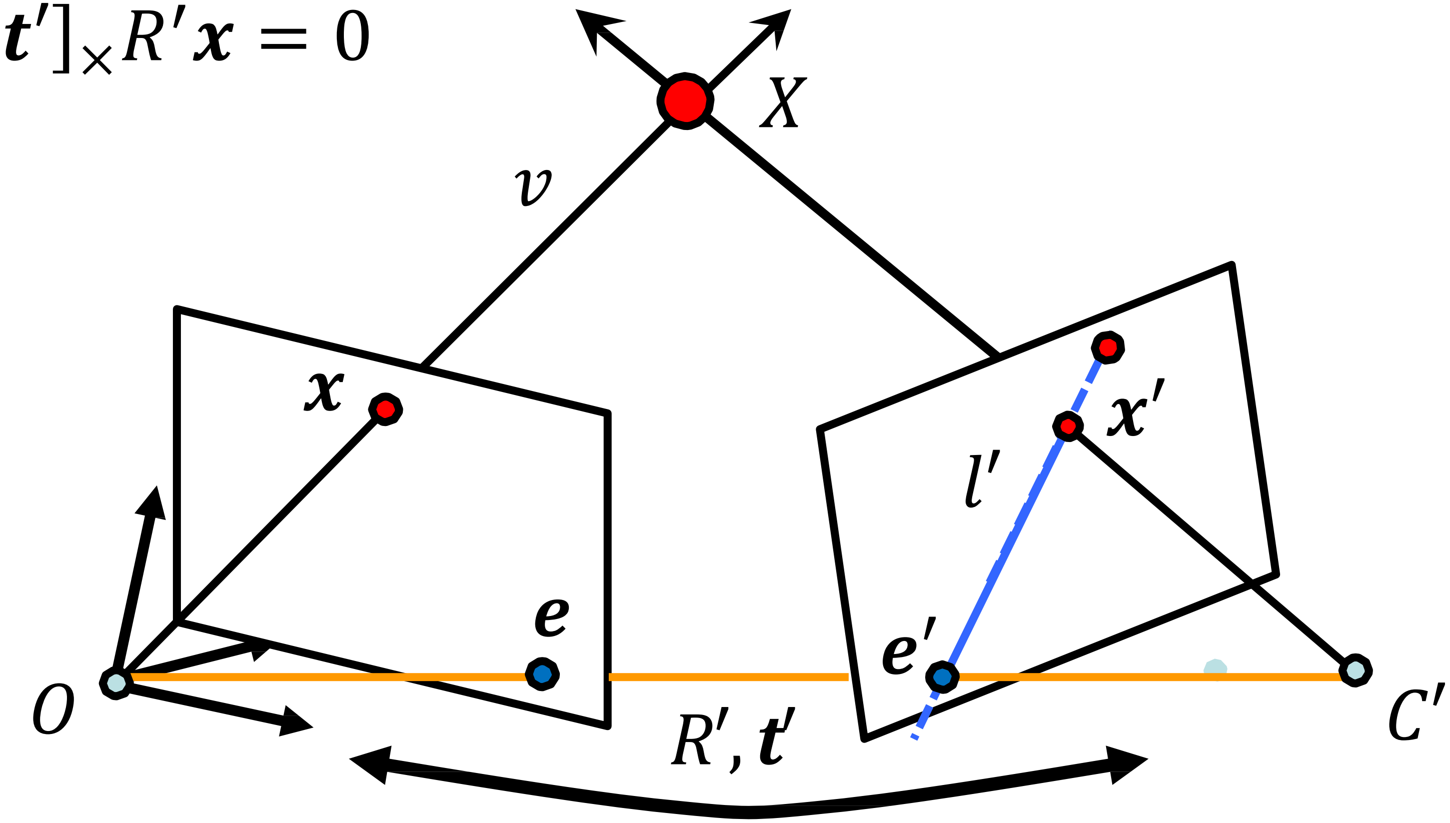
The Essential Matrix: Derivation

The essential matrix becomes

$$E = [\boldsymbol{t}']_{\times} R'$$

which installs the relation

$$\boldsymbol{x}'^{\top} [\boldsymbol{t}']_{\times} R' \boldsymbol{x} = 0$$



Essential Matrix And Fundamental Matrix

Given a pair of general cameras, we have two relations to hold

$$\boldsymbol{x}^\top F \boldsymbol{x} = 0, \text{and } \widehat{\boldsymbol{x}}'^\top E \widehat{\boldsymbol{x}} = 0$$

which by inverting normalization transformation means

$$\boldsymbol{x}^\top K'^{-\top} E K^{-1} \boldsymbol{x} = 0$$

that installs the following relation

$$\underline{E} = \underline{K'^\top F K}$$

Essential Matrix

Historically, the essential matrix was introduced before the fundamental matrix, and the fundamental matrix may be thought of as the generalization of the essential matrix in which the (inessential) assumption of calibrated cameras is removed. [HZ 9.6]

Rmk the matrix E has 5 degrees of freedom: 3 rotation angles, 3 translation components, but correspondence up to a scalar factor.

Result [9.17 HZ] A 3×3 matrix is an essential matrix if and only if two of its singular values are equal, and the third is zero

Computation of Essential Matrix

The essential matrix may be computed directly from

- matches in normalized image pairs
- From the fundamental matrix ($E = K'^{\top} F K$)

Once E is known, camera matrices may be retrieved and there are four possible solutions, except for overall scale (which cannot be determined).

Why the Fundamental Matrix?

Applications of the Fundamental Matrix

F captures information about the epipolar geometry of 2 views + camera parameters

- F gives constraints on how the scene changes under view point transformation, without reconstructing the scene!

Powerful tool in:

- 3D reconstruction (triangulation)
- Multi-view object/scene matching
- Disparity maps
- …

Nonlinear Methods for Triangulation

However, a different approach is preferred, such that triangulation is seen as a form of estimation for 3D point location.

The approach: provided $\boldsymbol{x}, \boldsymbol{x}'$, estimate $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{x}}'$ such that

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

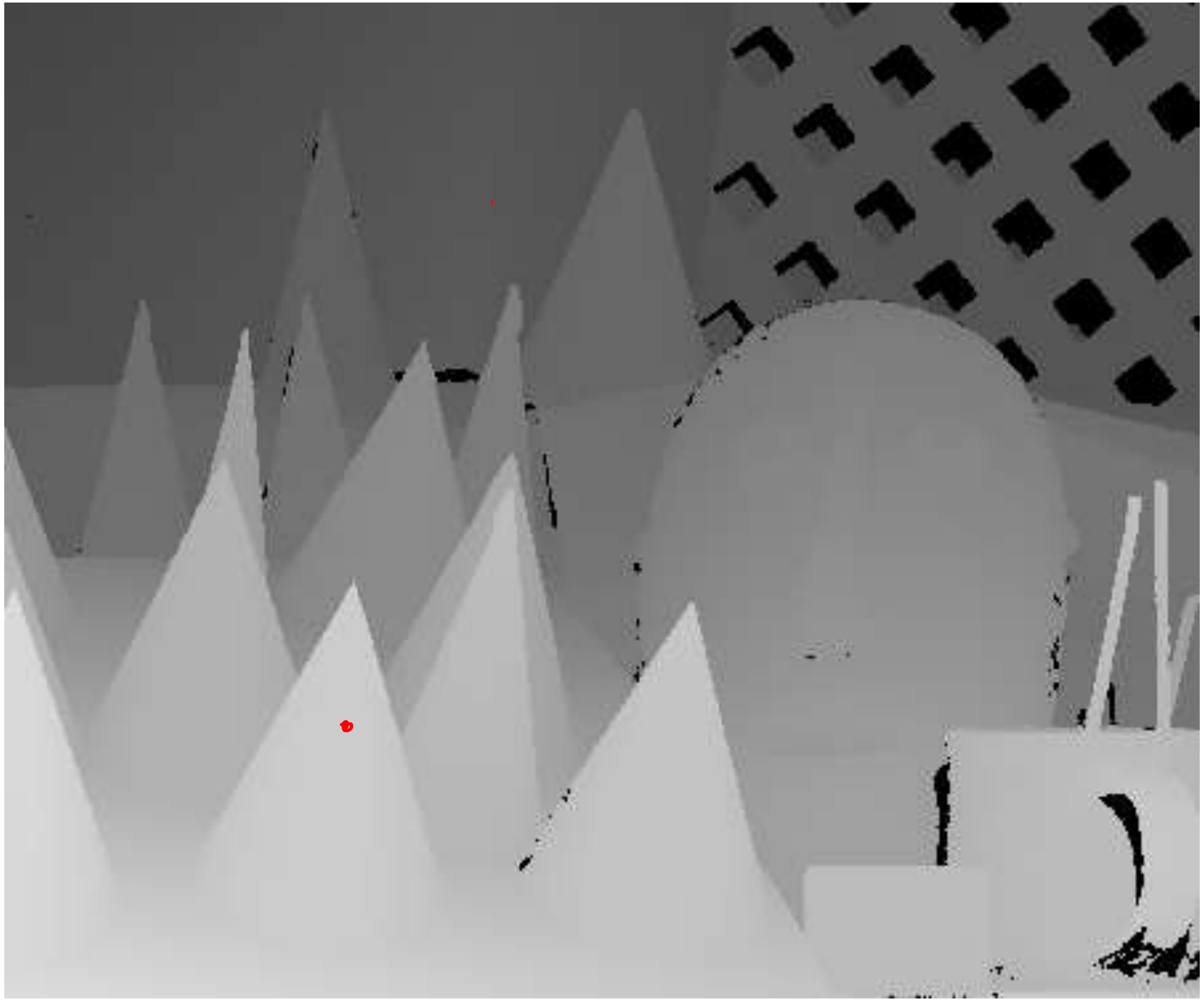
By minimizing $\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\|_2 + \|\widehat{\boldsymbol{x}}' - \boldsymbol{x}'\|_2$

Rmk this gives pairs $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{x}}'$ corresponding to the same 3D point and requires the fundamental matrix F, which can be also computed by point correspondences

Depth Estimation and Image Rectification

Stereo Pairs http://vision.middlebury.edu/stereo/data/

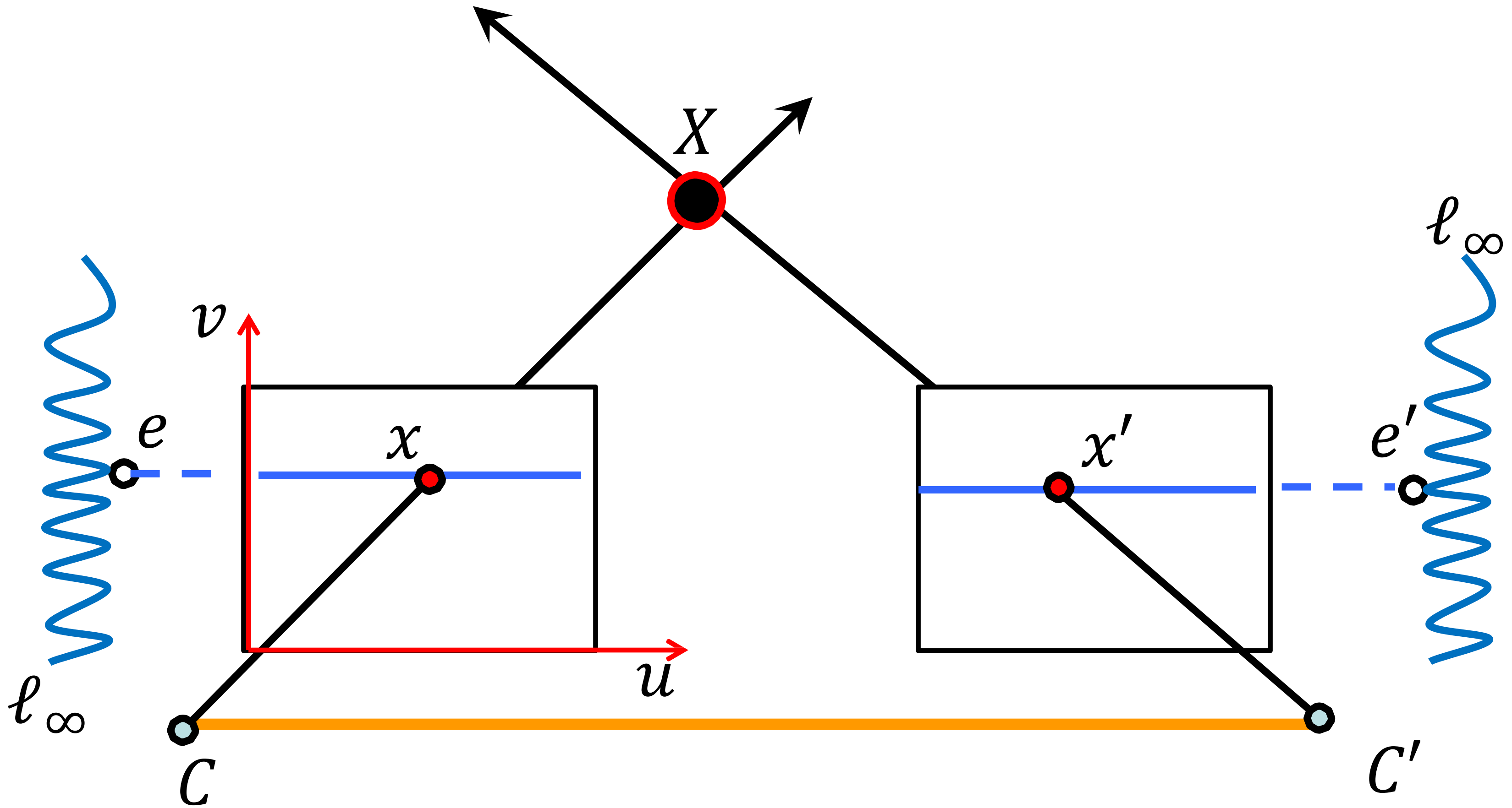
Stereo Pairs http://vision.middlebury.edu/stereo/data/



Parallel Images in a Stereo System

Epipolar lines are horizontal

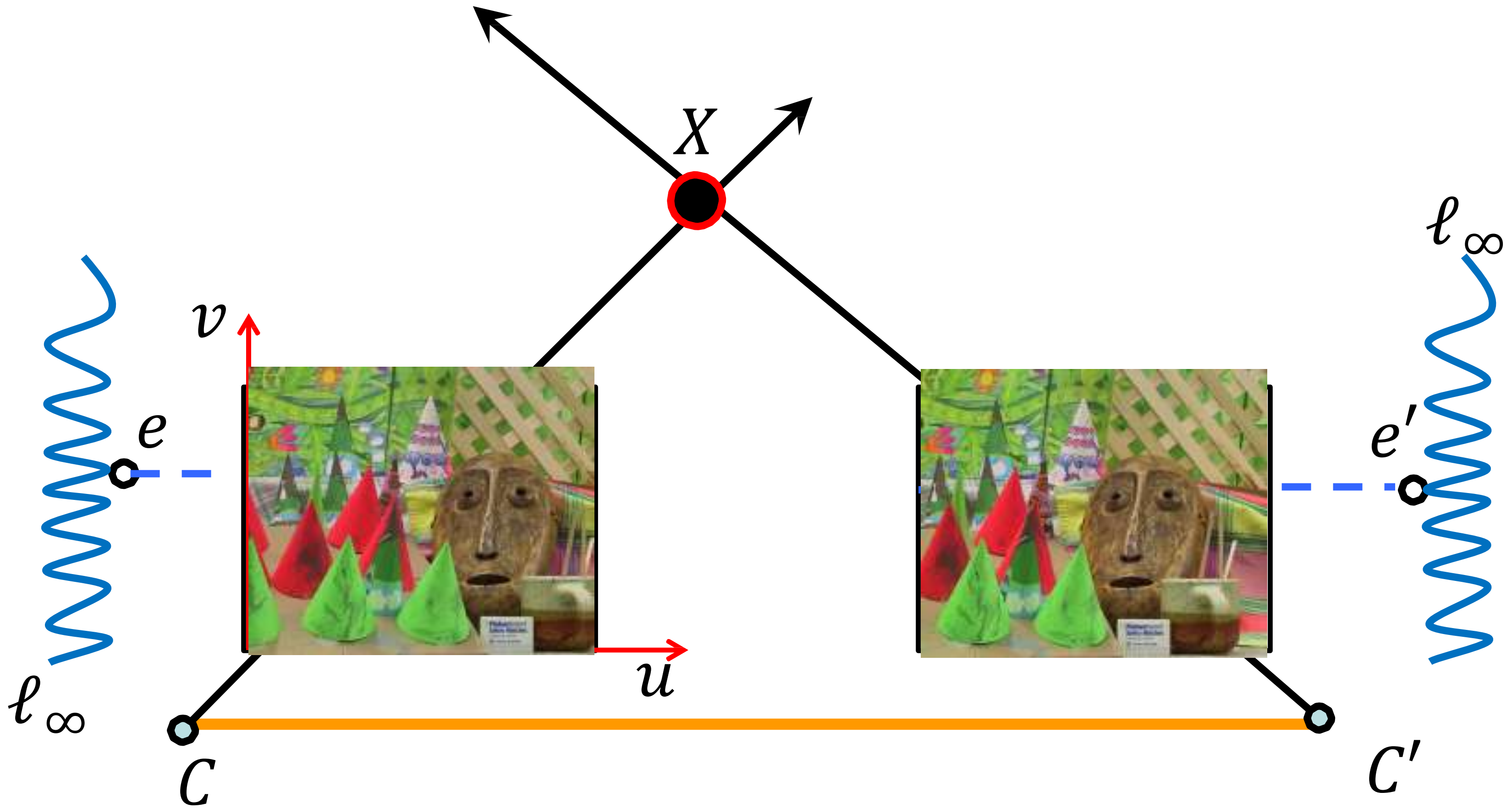
Epipoles go to infinity



Parallel Images in a Stereo System

Epipolar lines are horizontal

Epipoles go to infinity



Parallel Images in a Stereo System

The vertical coordinates $\boldsymbol{v}$ for corresponding points $\boldsymbol{x}$ and $\boldsymbol{x}'$ are the same

In fact, consider two identical and normalized cameras $K = K' = I_3$, no rotation and just an horizontal translation $\boldsymbol{t}$. Then

$$E = [\boldsymbol{t}]_\times R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{bmatrix}$$

Then

$$l = E^\top \boldsymbol{x}' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ t_x \\ -t_x v \end{bmatrix}$$

Which is an horizontal line -> corresponding points in these stereo systems live on the same row in different images

Computing Image Disparity

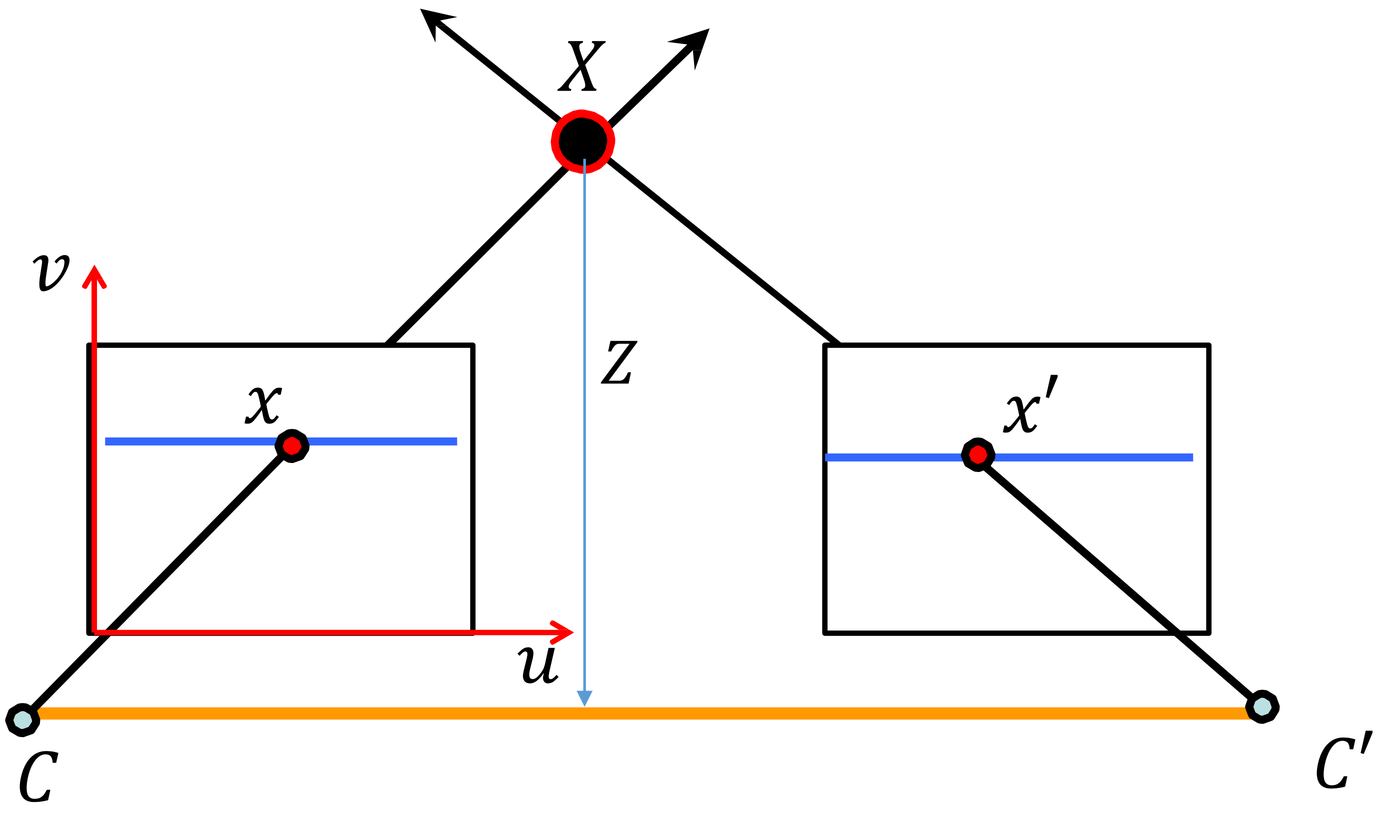
Given two corresponding points $\boldsymbol{x}$, and $\boldsymbol{x}'$, it holds $y = y'$

$$\boldsymbol{x} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \boldsymbol{x}' = \begin{bmatrix} u' \\ v \\ 1 \end{bmatrix}$$

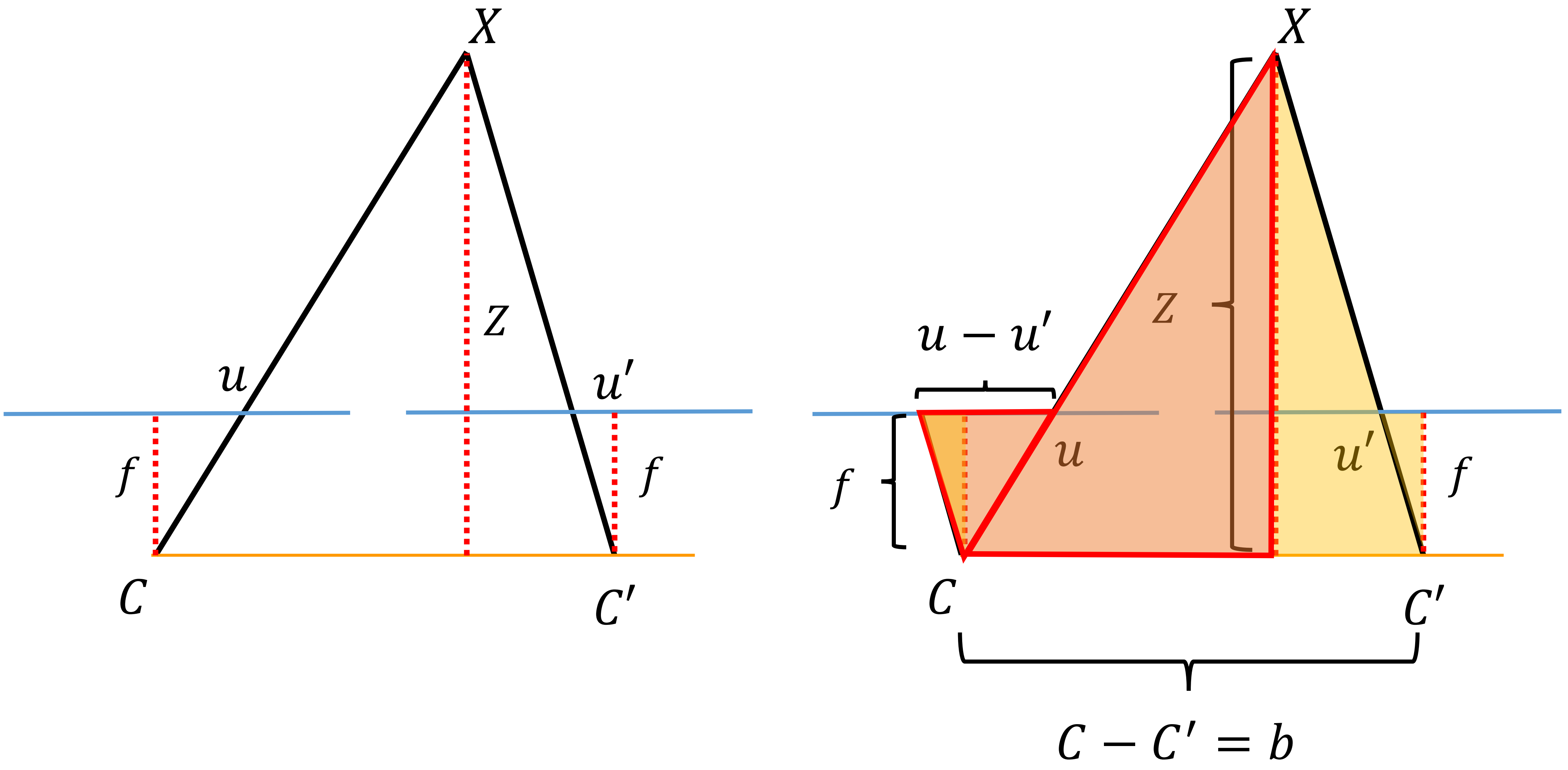
We can compute the disparity as

$$disp(\boldsymbol{x}, \boldsymbol{x}') = u - u'$$

That is inversely proportional to the depth of the point X



Disparity is Inversely Proportional to the depth



Giacomo Boracchi

Disparity is Inversely Proportional to the depth

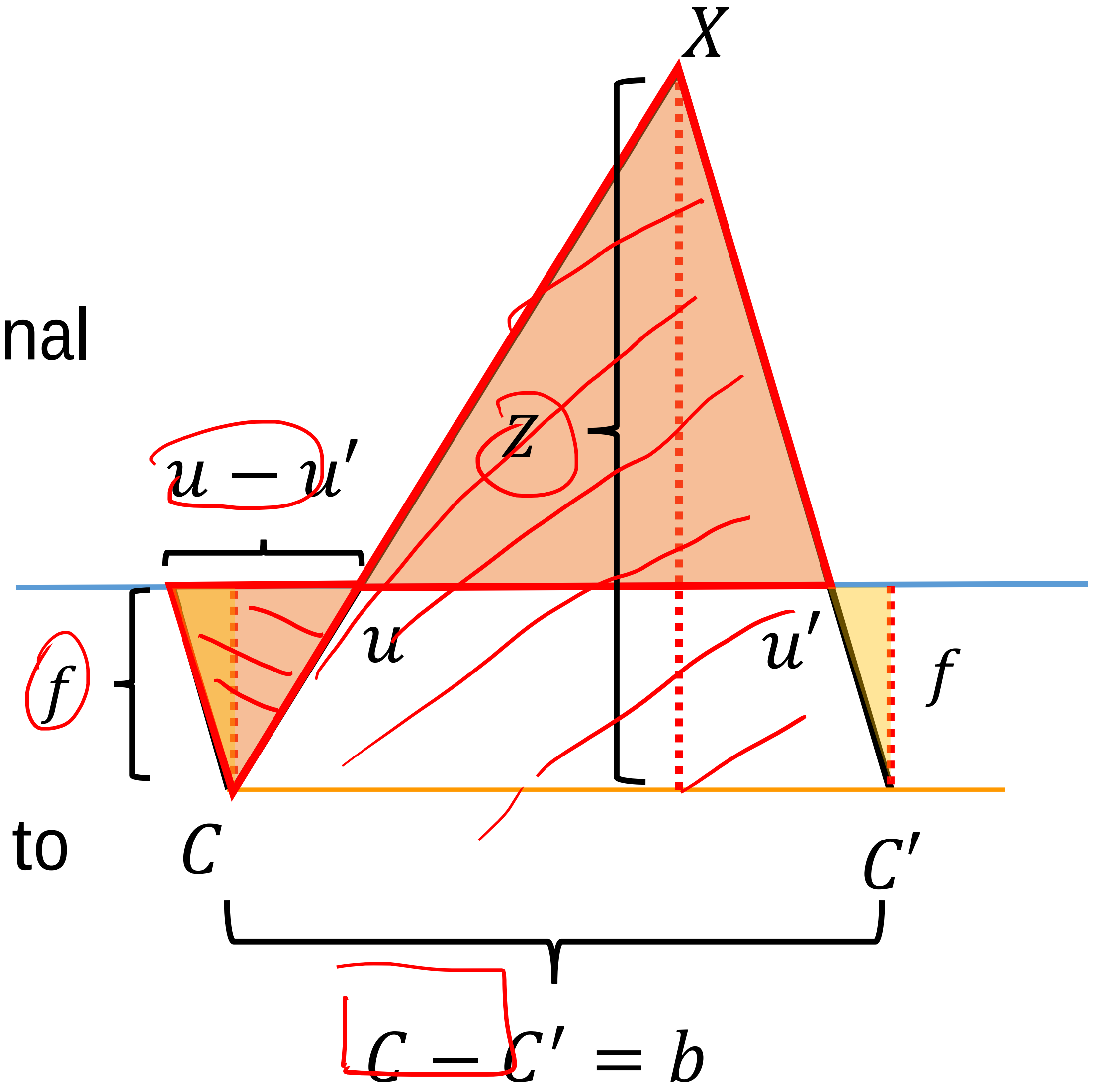
From triangle similarity it holds

$$u - u' : f = b : Z$$

Then, disparity is inversely proportional to the depth

$$u - u' \propto \frac{1}{Z}$$

Rmk to compute the depth we need to know f, thus camera has to be calibrated



Disparity Maps

Estimate at each image point $\boldsymbol{x}$, the depth of the scene point X as inversely proportional to the displacement between $\boldsymbol{x}$ and $\boldsymbol{x}'$

In general:

- Given $\boldsymbol{x}$, the corresponding point in the second image has to be searched only along the epipolar line $\boldsymbol{l}' = F\boldsymbol{x}$. This might be unpractical in many circumstances
- When the cameras are parallel, then the search is much more convenient, as it has to be performed row-wise only

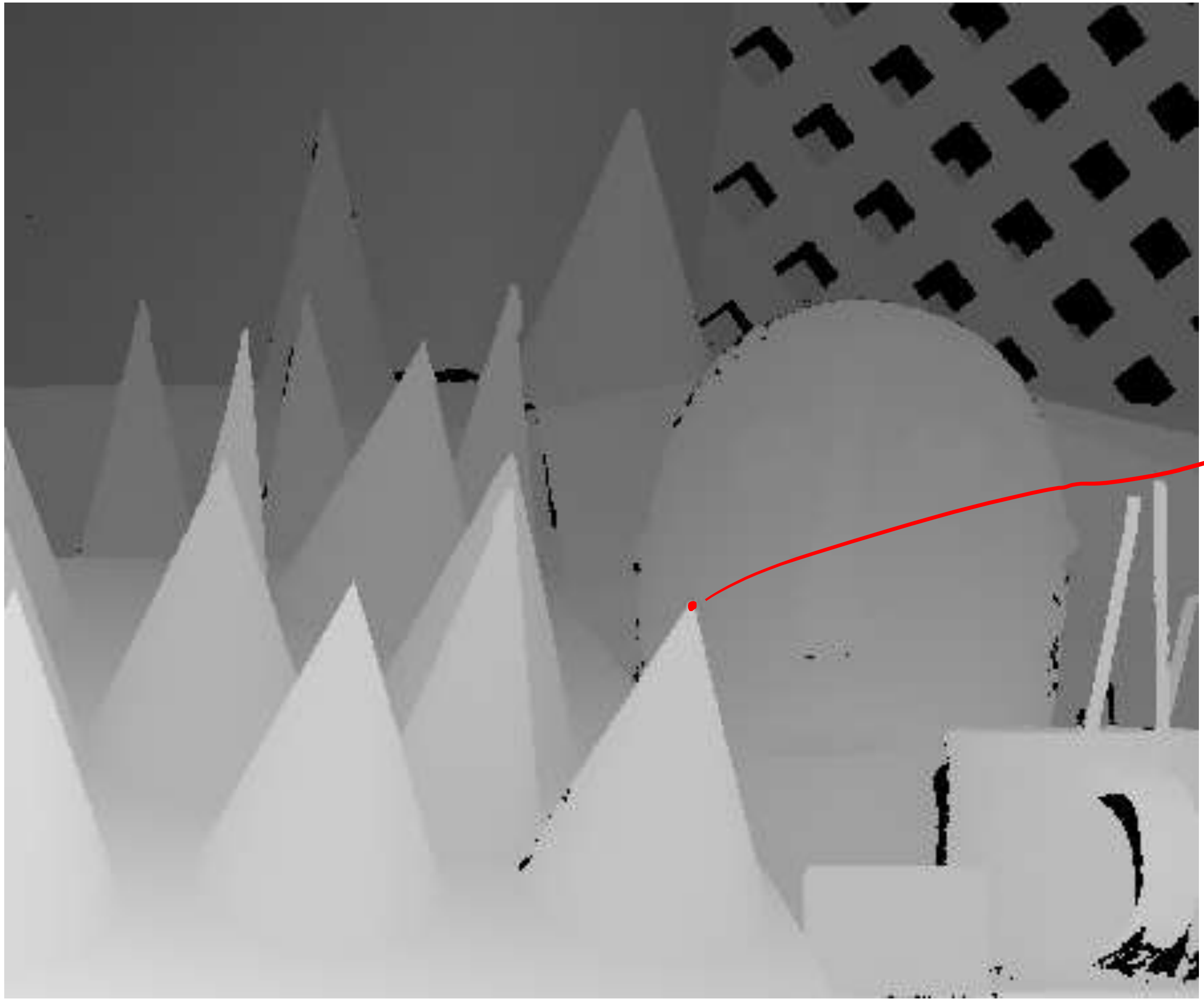
Stereo Pairs http://vision.middlebury.edu/stereo/data/

3×3

4

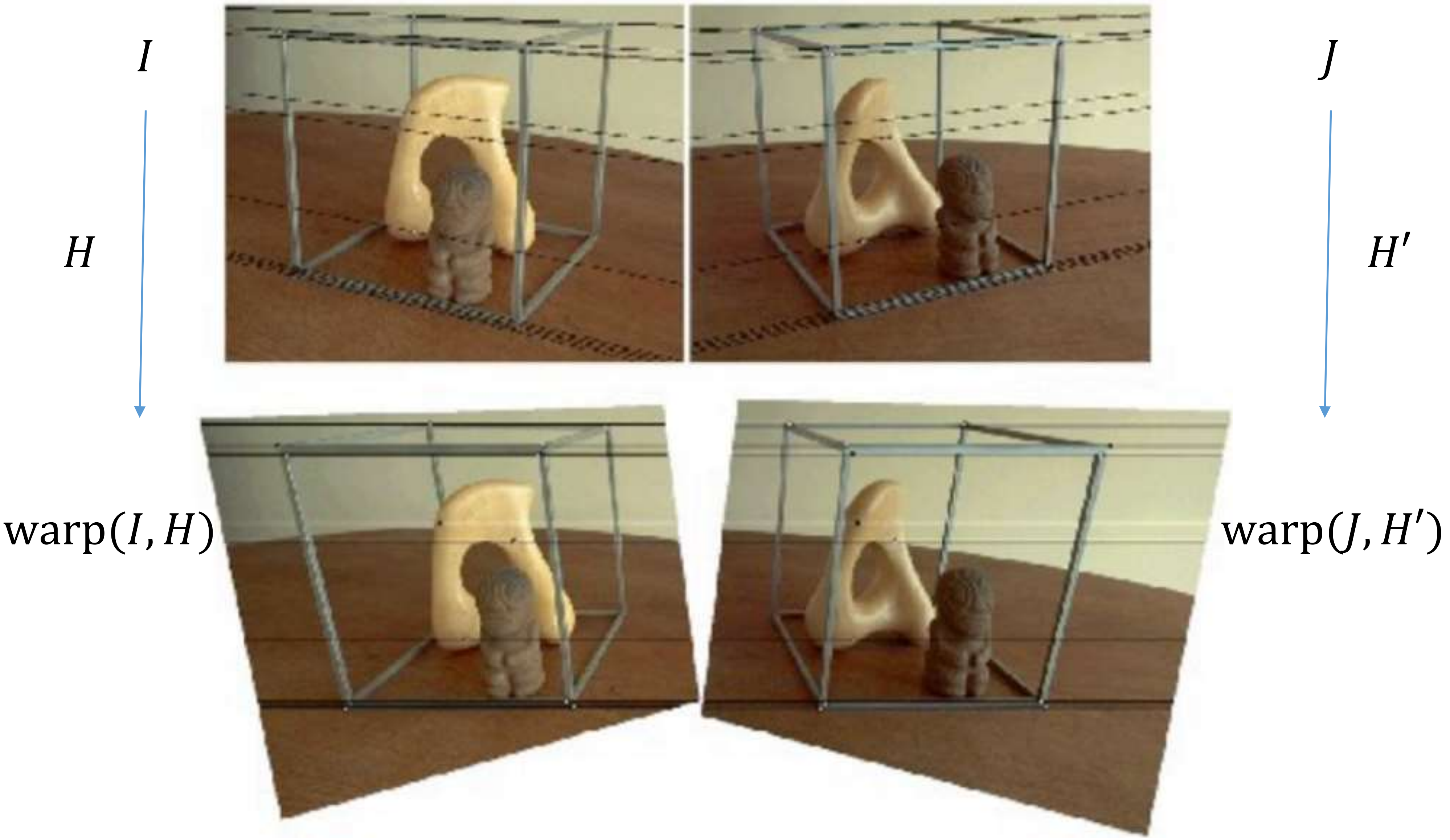
Stereo Pairs http://vision.middlebury.edu/stereo/data/

Stereo Pairs http://vision.middlebury.edu/stereo/data/



$u - u'$

Image Rectification to Ease Dispartity Maps Estimation



Basic Idea of Image Rectification

Apply two projective transformations to:

1. Make their epipolar lines parallel
2. Make matching points to have approximately the same coulimn coordinate (this latter is to avoid very large distortions)

Making epipolar lines parallel (and horizontal) means moving

$$e \rightarrow [1, 0, 0]'$$

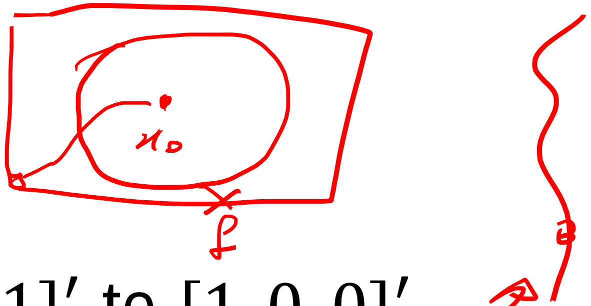
the ideal point of horizontal lines

Rmk there are infinite homographies that can do this. Most, however would introduce too large distortions

Moving Epipole to Infinity

Idea: find an homography that should be very close to a rotation near the image center. Consider the following

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/f & 0 & 1 \end{bmatrix}$$



This satisfies the condition and maps the epipole $\boldsymbol{e} = [f, 0, 1]'$ to $[1, 0, 0]'$

Thus the transformation would be: $\boxed{H = GRT,}$ where

- T: translates the image center to the image origin (where the homography should
- R: rotates the image such that the epipole goes to $\boldsymbol{e} = [f, 0, 1]'$

This solves the problem for I, $H = GRT$

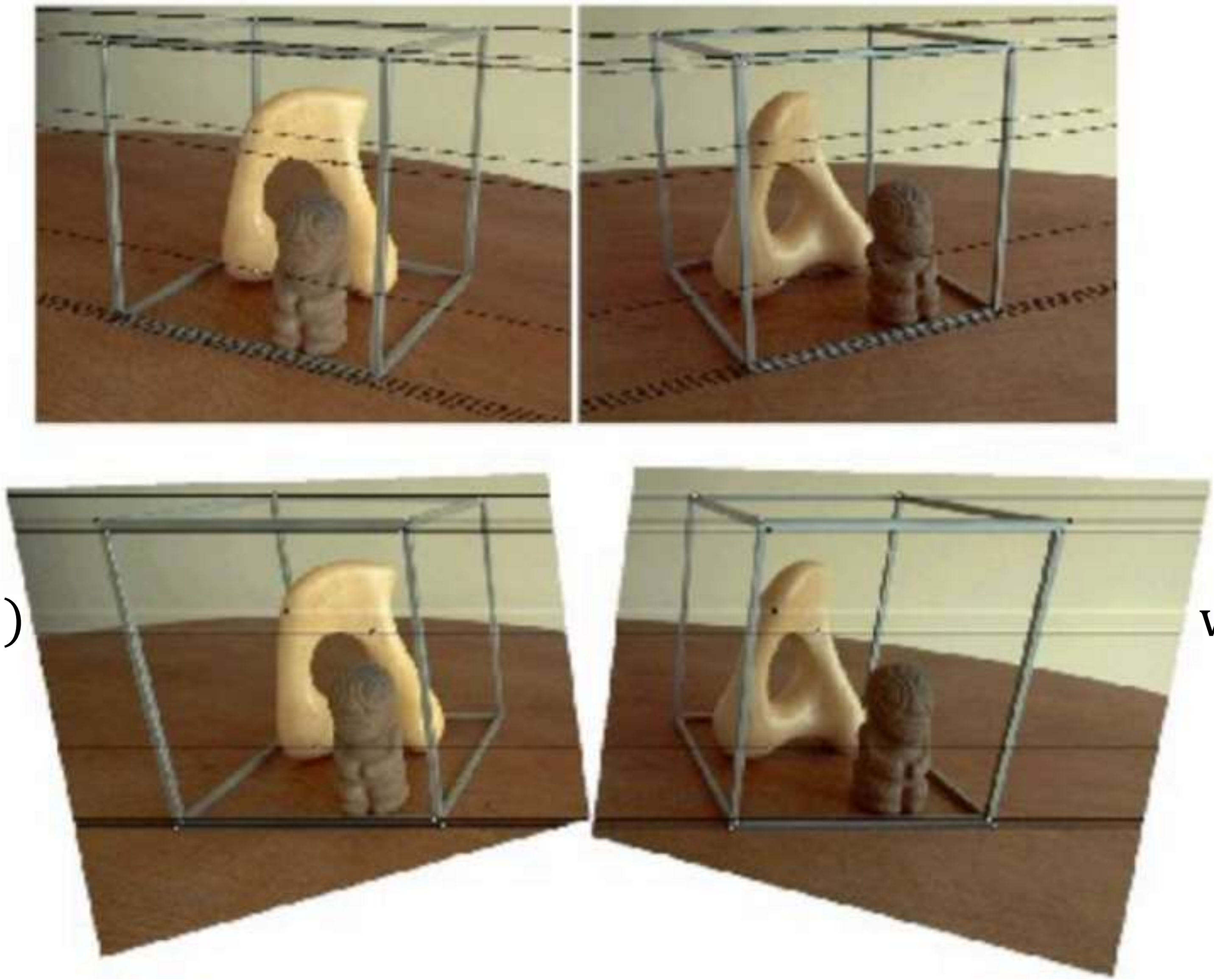
Image Rectification to Ease Dispartity Maps Estimation

I

J

$H = GRT$

H'



$\text{warp}(I, H)$

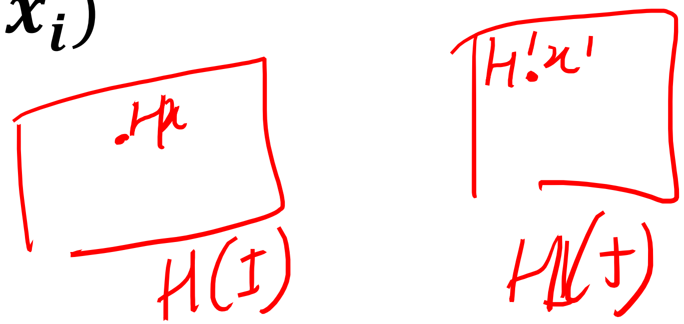
$\text{warp}(J, H')$

H': Matching the Transformation H

The first image I can be transformed by H

1. We want H' to send the epipole to infinity (as done before)
2. This leaves plenty of degrees of freedom, we choose the one minimizing the distance of corresponding points $(\boldsymbol{x}_i, \boldsymbol{x}_i')$

$$\sum_i d(H\boldsymbol{x}_i, H'\boldsymbol{x}_i')^2$$

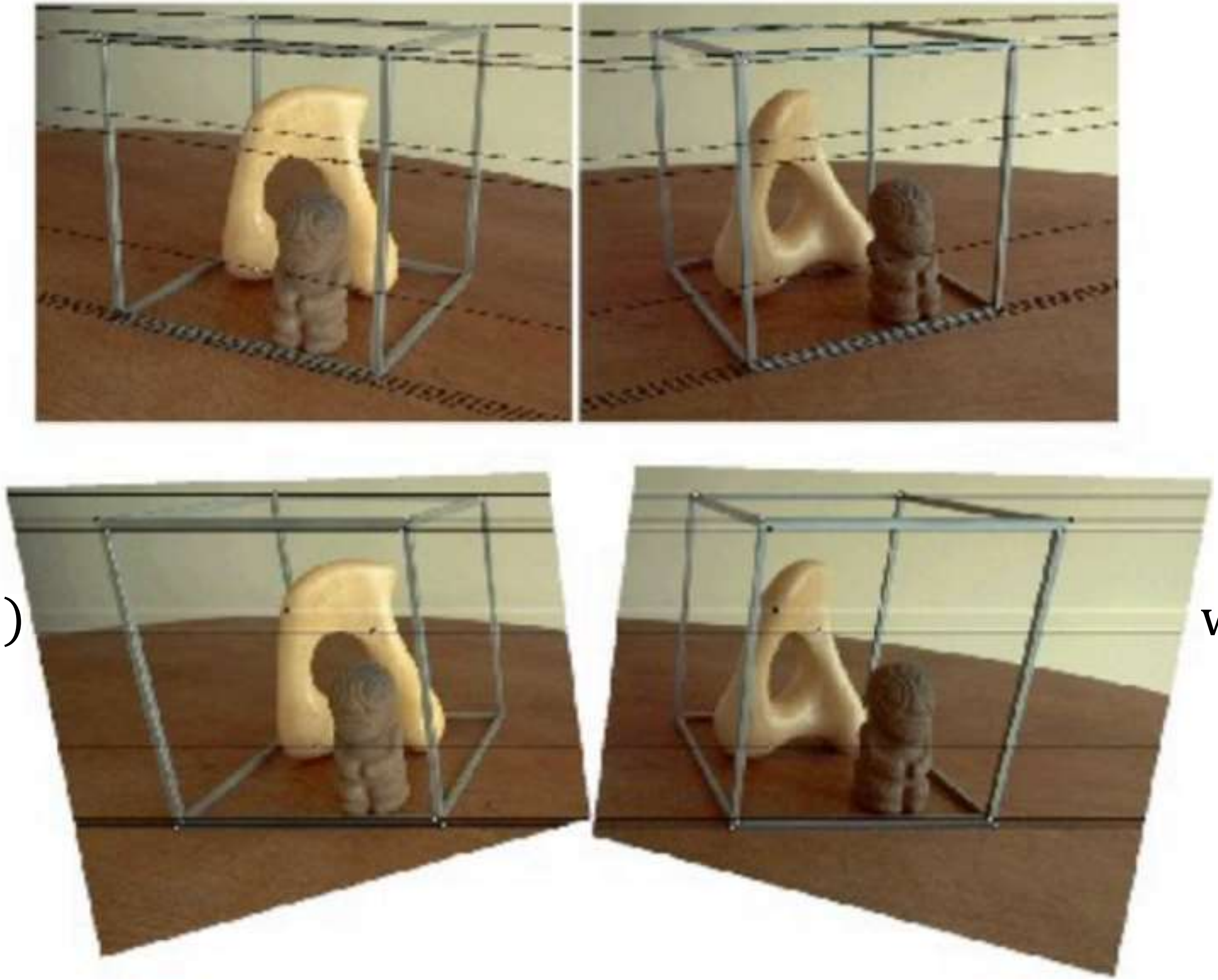


Rmk $H\boldsymbol{x}_i$ and $H'\boldsymbol{x}_i'$ do not belong to the same image!! This is a trick to force the outputs of the two transformations to be nearby..

Image Rectification to Ease Disparity Maps Estimation

I J

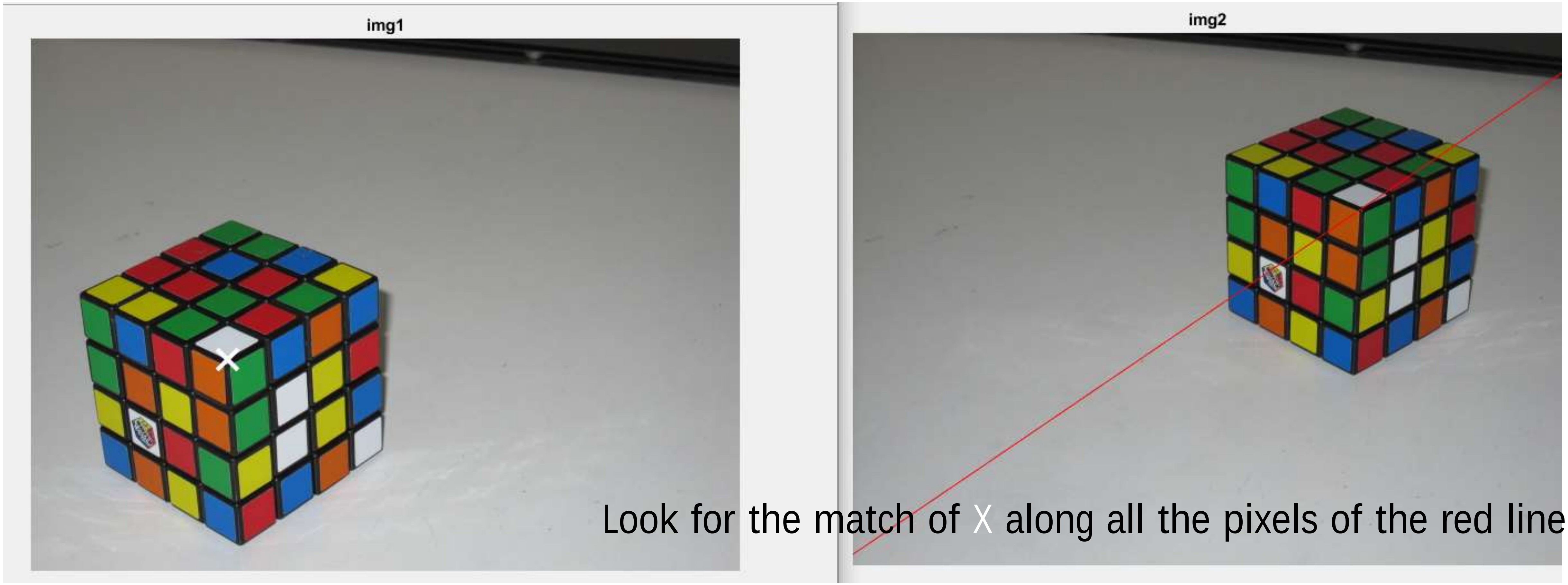
$H = GRT$ H'



$\mathrm{warp}(I, H)$ $\mathrm{warp}(J, H')$

Otherwise…

When images are not rectified, it is necessary to comute the displacement $u - u'$ for each point in the first image along its epipolar line



Look for the match of X along all the pixels of the red line

Epipolar Transfer to View Synthesis