CS 565 Computer Vision

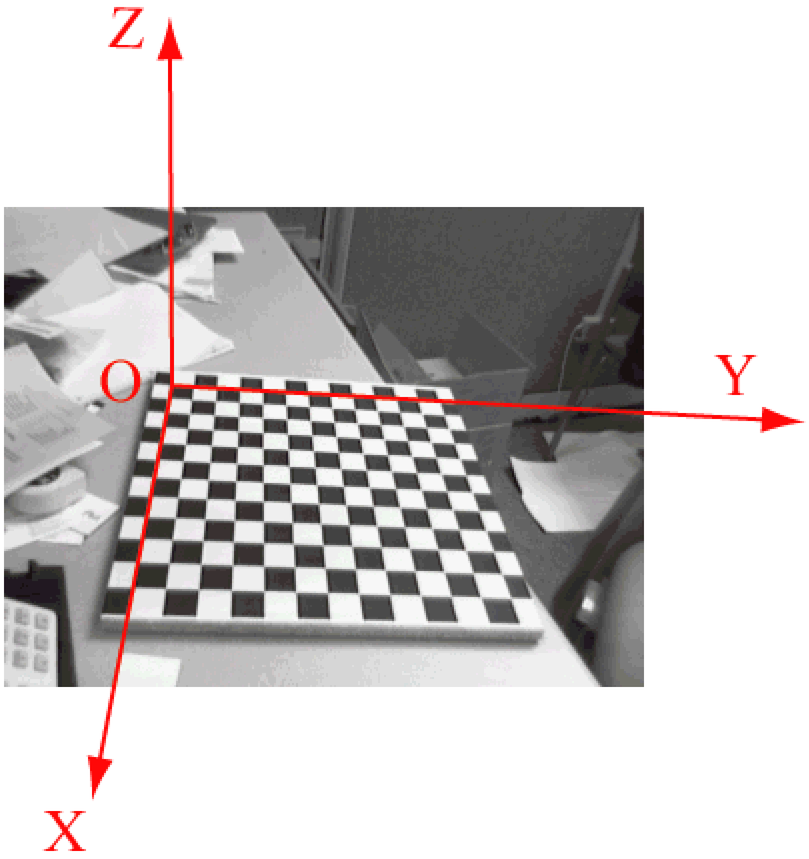
Nazar Khan
PUCIT

Lectures 19 and 20: Stereo Reconstruction

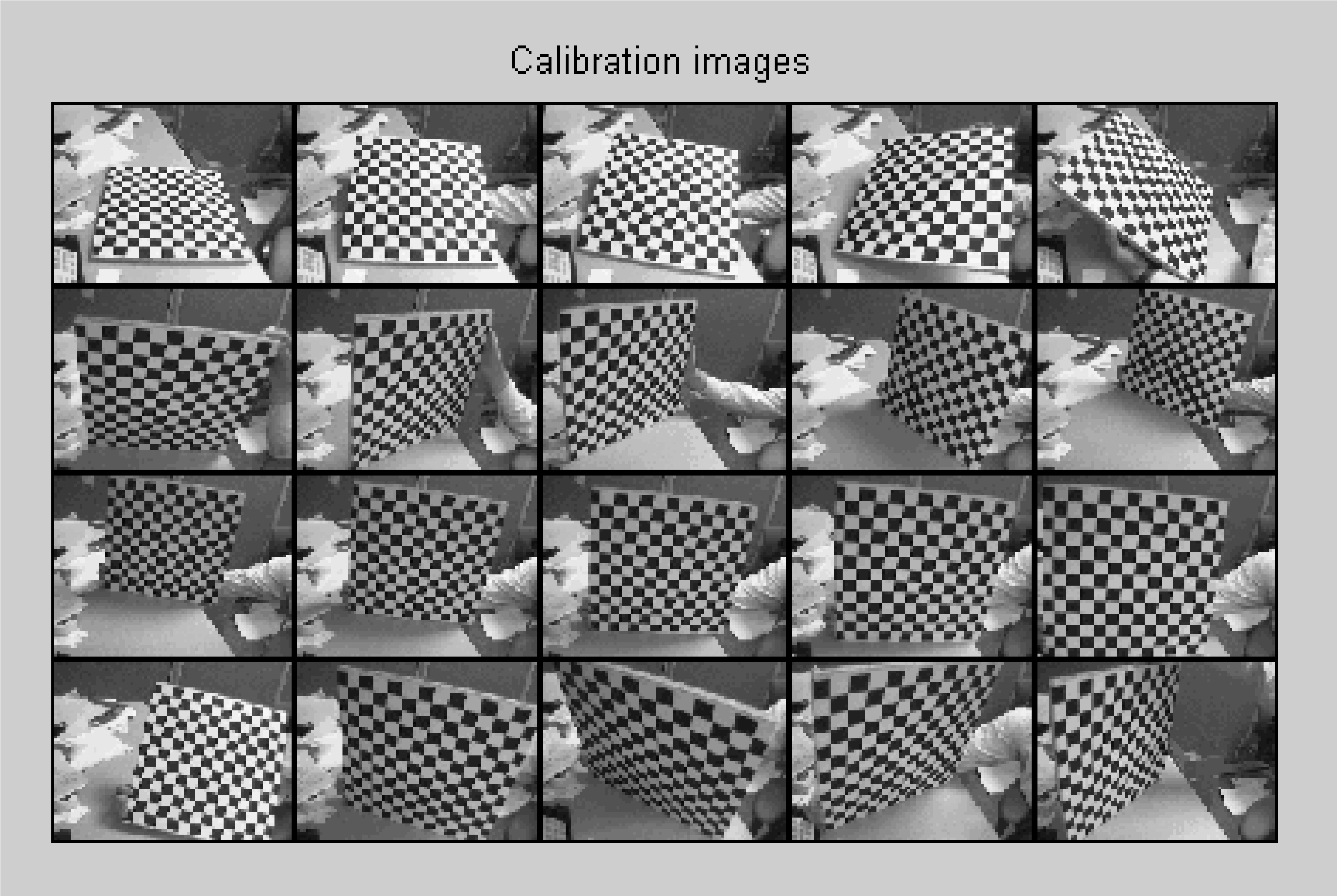
Camera Calibration

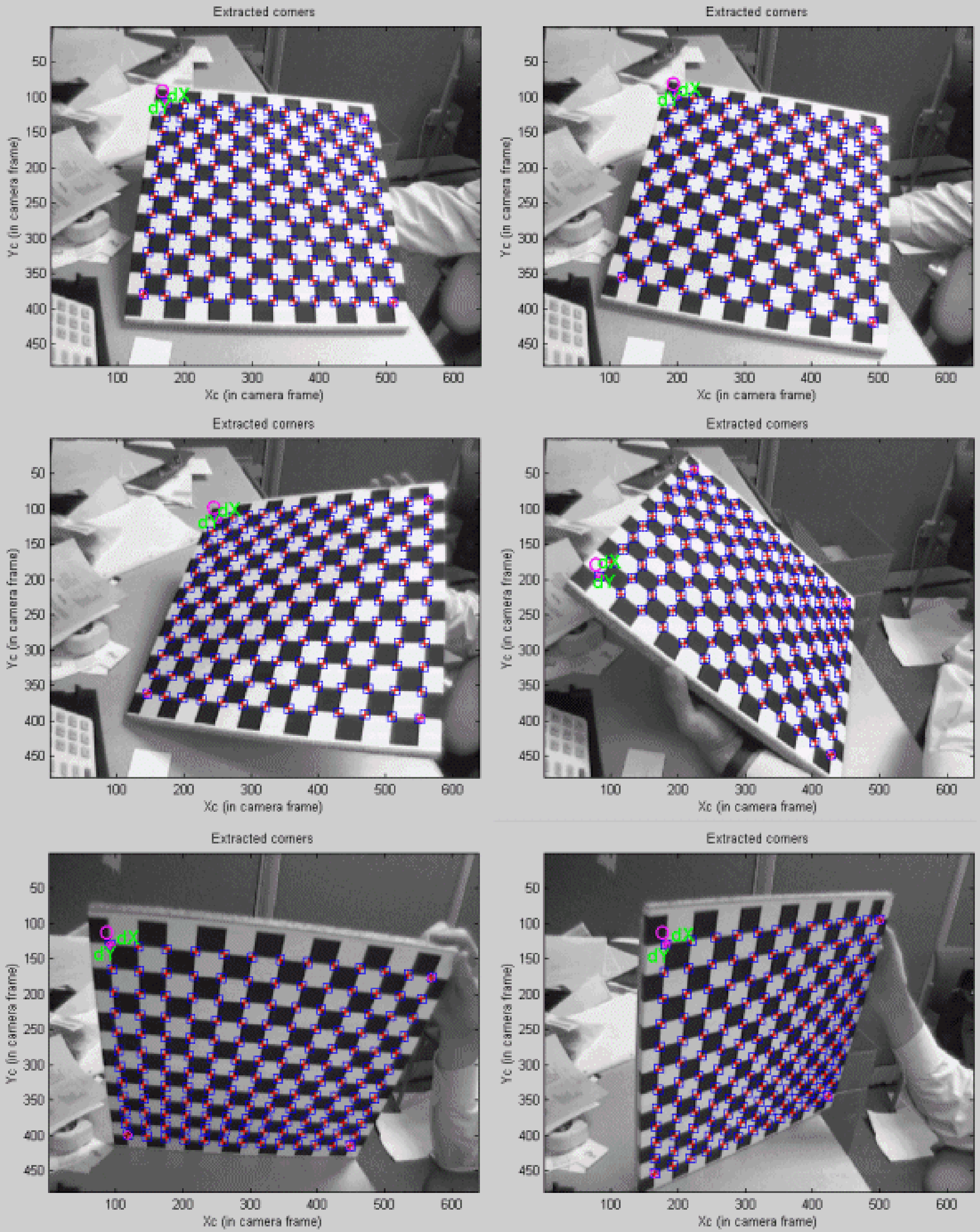
- Denotes the estimation of the 5 intrinsic and 6 extrinsic camera parameters.
- Many algorithms have been proposed in the literature.
- Basic idea: Investigate image of an object of known size and shape.

Camera Calibration



Camera Calibration




Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY
Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY
Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY
Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY
Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY
Extracted corners
Yc (in camera frame)
Xc (in camera frame)
dX
dY

Camera Calibration

- Each identified point correspondence $x_i \leftrightarrow X_i$ gives 2 constraints.
- Thus, for estimating 11 parameters, one has to find 6 corresponding points.
- Taking into account more point correspondences (e.g. in a least squares sense) makes the estimation less sensitive w.r.t. errors.

Cross Product

- A cross product of two 3-vectors $a=(a_1,a_2,a_3)^T$ and $b=(b_1,b_2,b_3)^T$ can be written as

  $a \times b = [a_2b_3-a_3b_2 \; ; \; a_3b_1-a_1b_3 \; ; \; a_1b_2-a_2b_1]$

- This can also be written in matrix form

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

- $a \times a = 0$ for all vectors a.
- **H.W: Compute $a \times a$.**

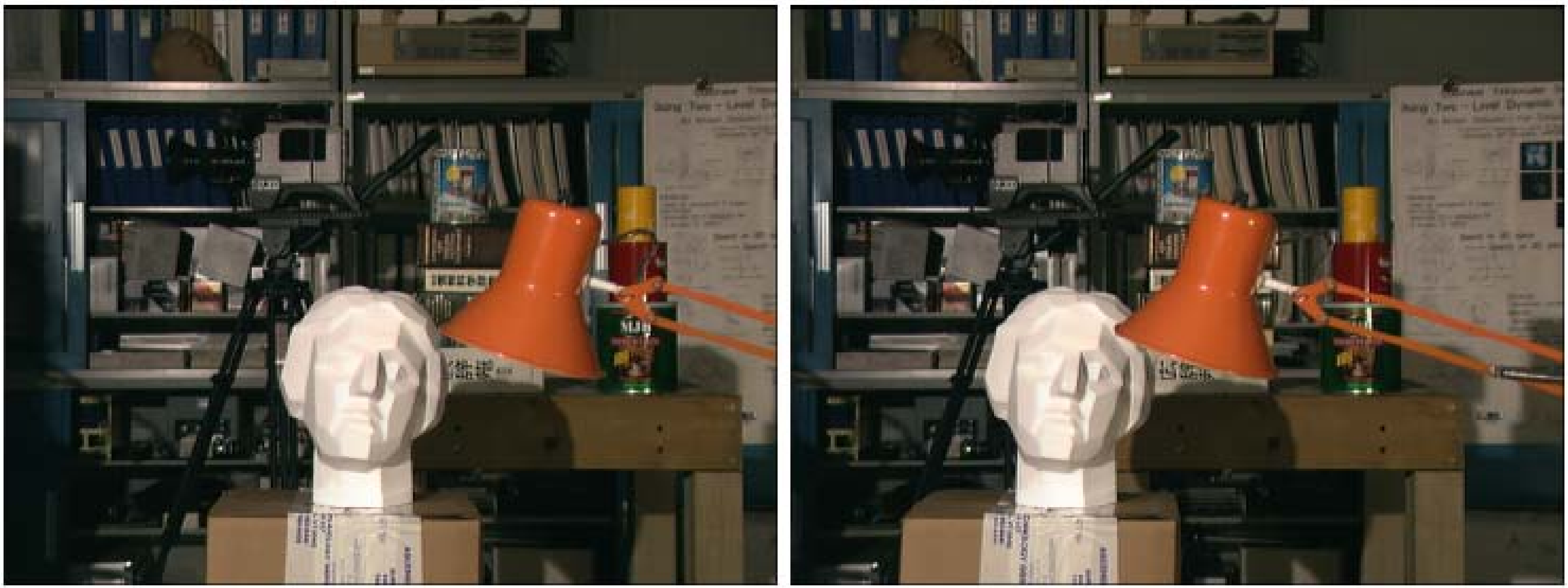
Camera Calibration

- $x_i = PX_i \Rightarrow x_i \times PX_i = 0 \Rightarrow$ $\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i^\top & y_i\mathbf{X}_i^\top \\ w_i\mathbf{X}_i^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0}$
- A set of such image-world point correspondences leads to a linear system Av=0.
- Solve for the 12x1 vector v and rearrange to form the camera matrix P.
- We have already looked at solving systems of type Av=0 when we studied homography estimation.
  - DLT (see end of lectures 16-18)

Stereo Reconstruction

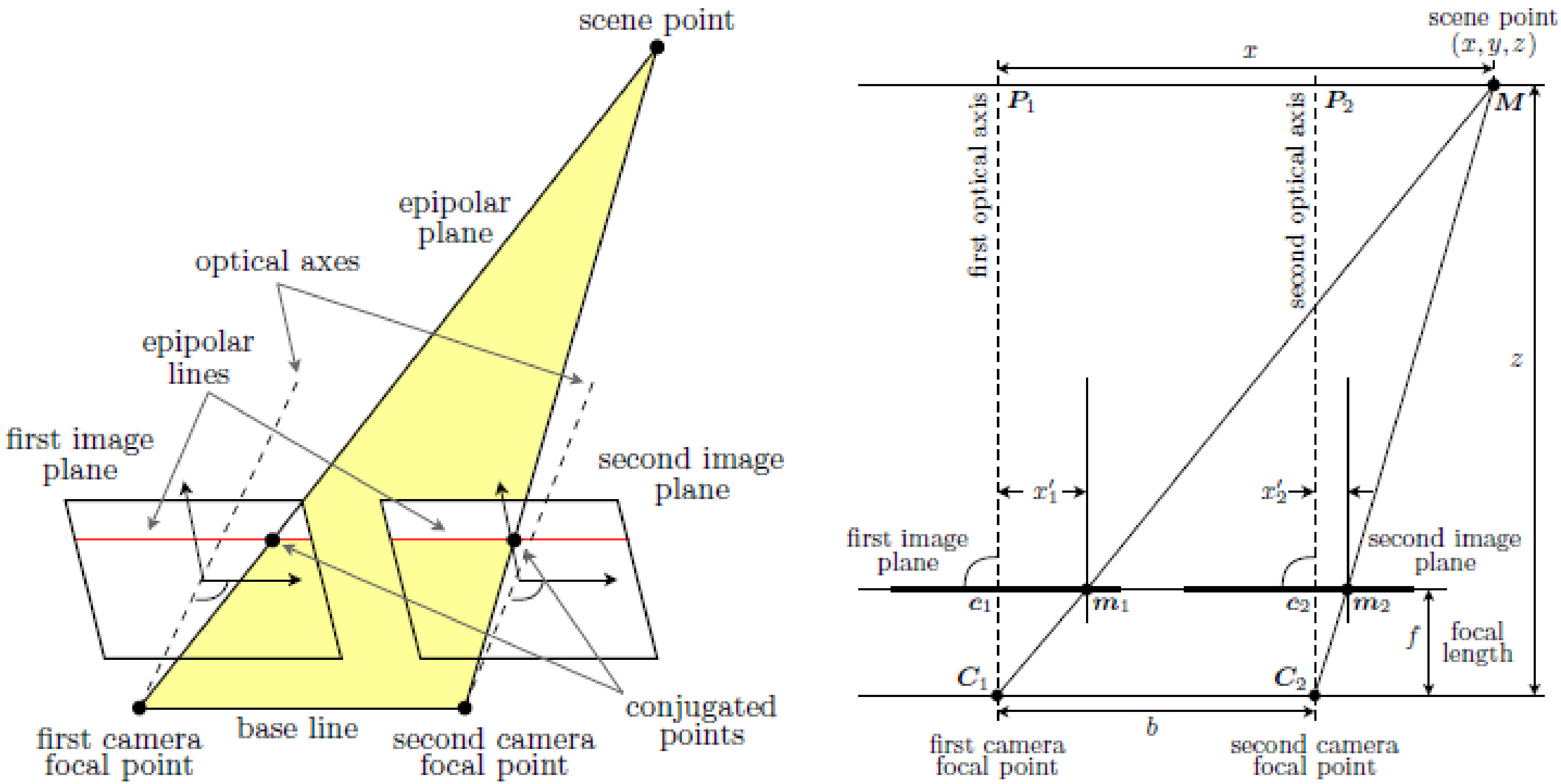
- So far, we have only investigated the projective geometry in the monocular case with a single pinhole camera.
- Considering two cameras allows us to reconstruct the depth of a scene from the displacements between the two stereo images.
- To do this, we will study **stereo geometry** (also called **epipolar geometry**).

Stereo Reconstruction



A stereo image pair from the Middlebury web page. The goal is to reconstruct the 3-D scene. Source: `http://cat.middlebury.edu/stereo/data.html`.

Simplified Model: Stereo Geometry for Orthoparallel Cameras



Stereo geometry for two identical pinhole cameras with parallel optical axes (orthoparallel cameras). Author: M. Mainberger (2010).

Simplified Model: Stereo Geometry for Orthoparallel Cameras

Terminology

- **orthoparallel cameras**: two identical cameras with parallel optical axes.
- **base line**: connecting line between both optical centres (focal points)
- **base line distance b**: distance between both optical centres
- **conjugated points**: two points in different images that result from the same 3-D scene point
- **epipolar plane**: plane through the scene point and both optical centres
- **epipolar lines**: intersecting lines of the epipolar plane with both image planes; contain conjugated points
- **epipole**: image of camera centre in the other image plane
- **disparity**: distance between two conjugated points, if both images are superimposed

Depth Computation

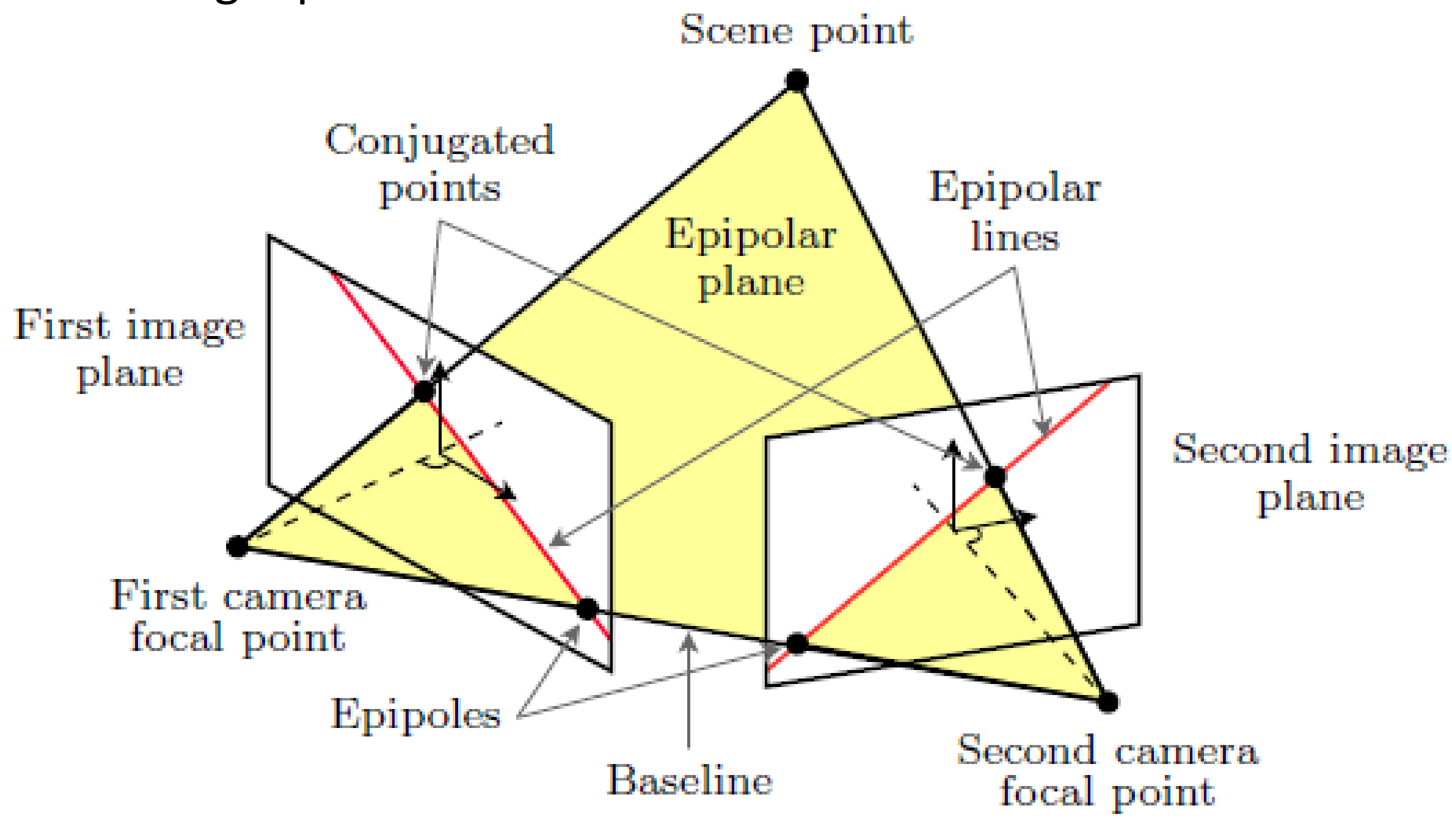
- Place the origin of the coordinate system in the left camera lens centre C_1.
- From the similarity of the triangles P_1MC_1 and $c_1m_1C_1$ it follows that $x/z=x_1'/f$
- From the similarity of the triangles P_2MC_2 and $c_2m_2C_2$ one obtains $(x - b)/z=x_2'/f$
- Eliminating *x* in both equations and using the fact that $x_1' > x_2'$ gives $z = bf/(x_1' - x_2')$. **(H.W: Show that this formula is correct.)**

Simplified Model: Stereo Geometry for Orthoparallel Cameras

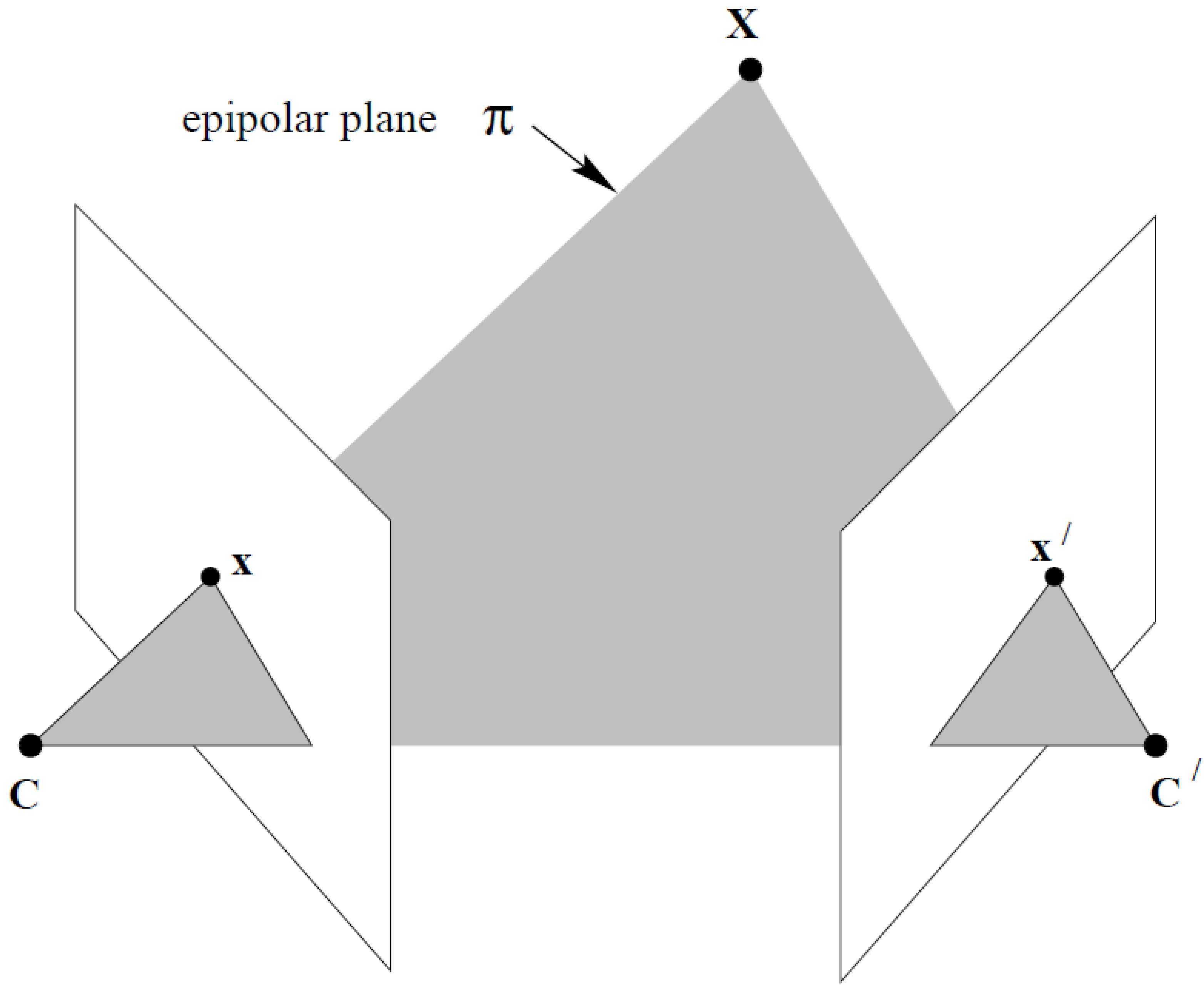
- If the baseline distance *b* and the focal length *f* are known in the orthoparallel case, the disparity $|x_1' - x_2'|$ allows to compute the depth z.
- The main problem is the reliable estimation of the disparity:
  - Often disparities can only be measured with pixel precision. This suggests to choose a large baseline distance.
  - On the other hand, this may lead to more occlusions and makes it more difficult to find correspondences between both images.

Stereo Geometry for Converging Cameras

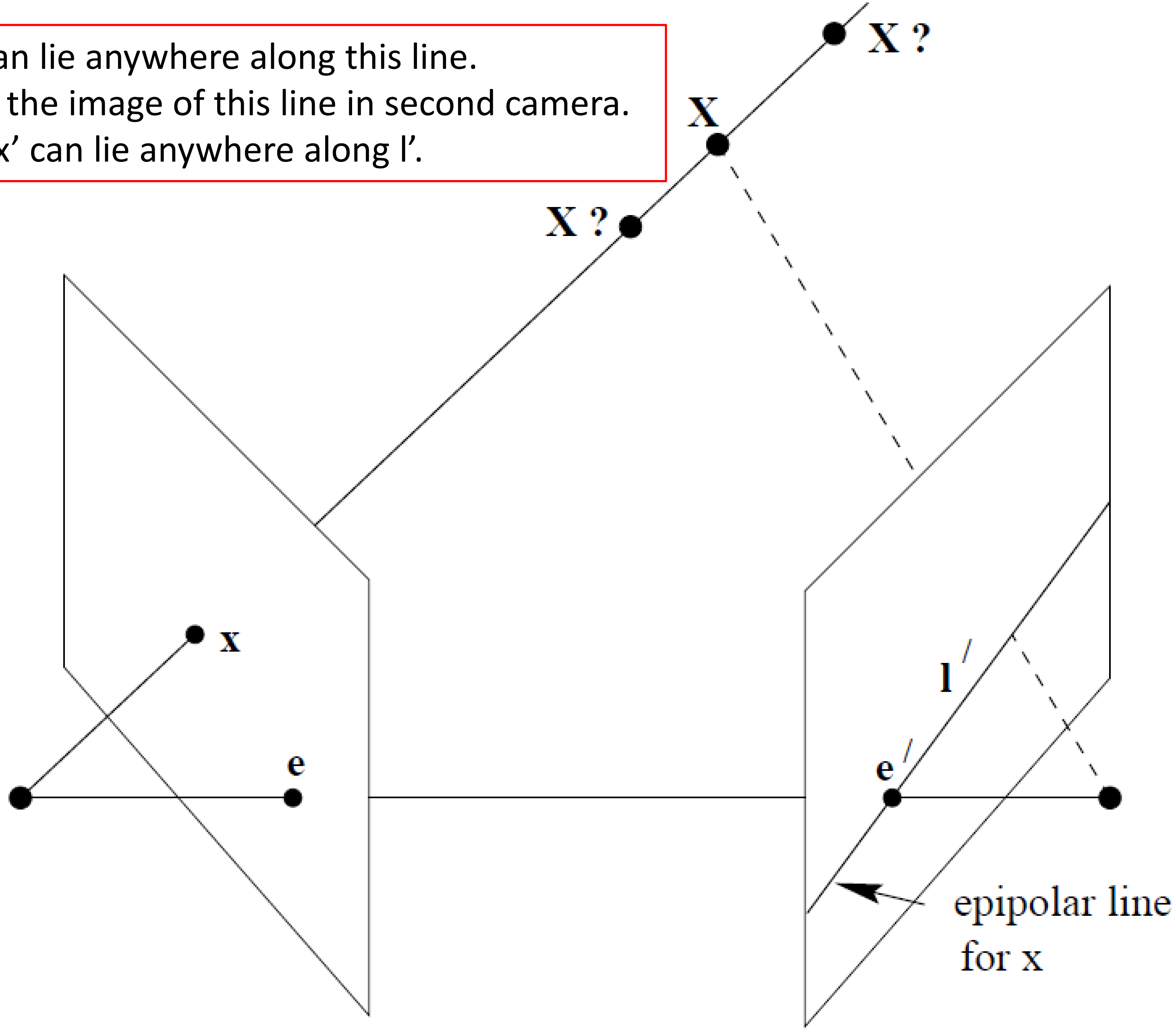
Conjugated points still lie along the epipolar lines. However, the two epipolar lines are no longer parallel.

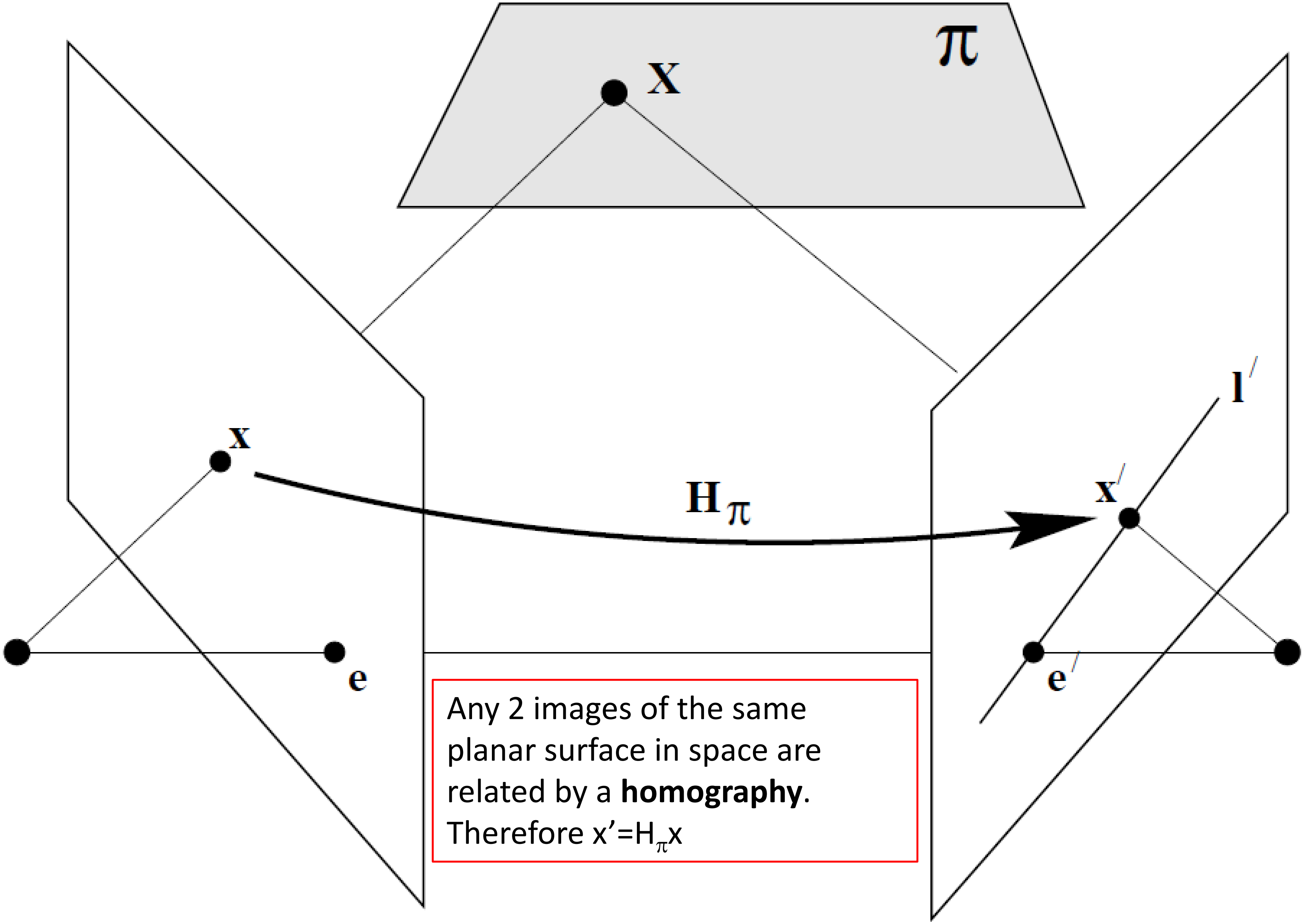


Two converging cameras in arbitrary position and orientation. Author: M. Mainberger (2010).


X
epipolar plane
π
x
x /
C
C /

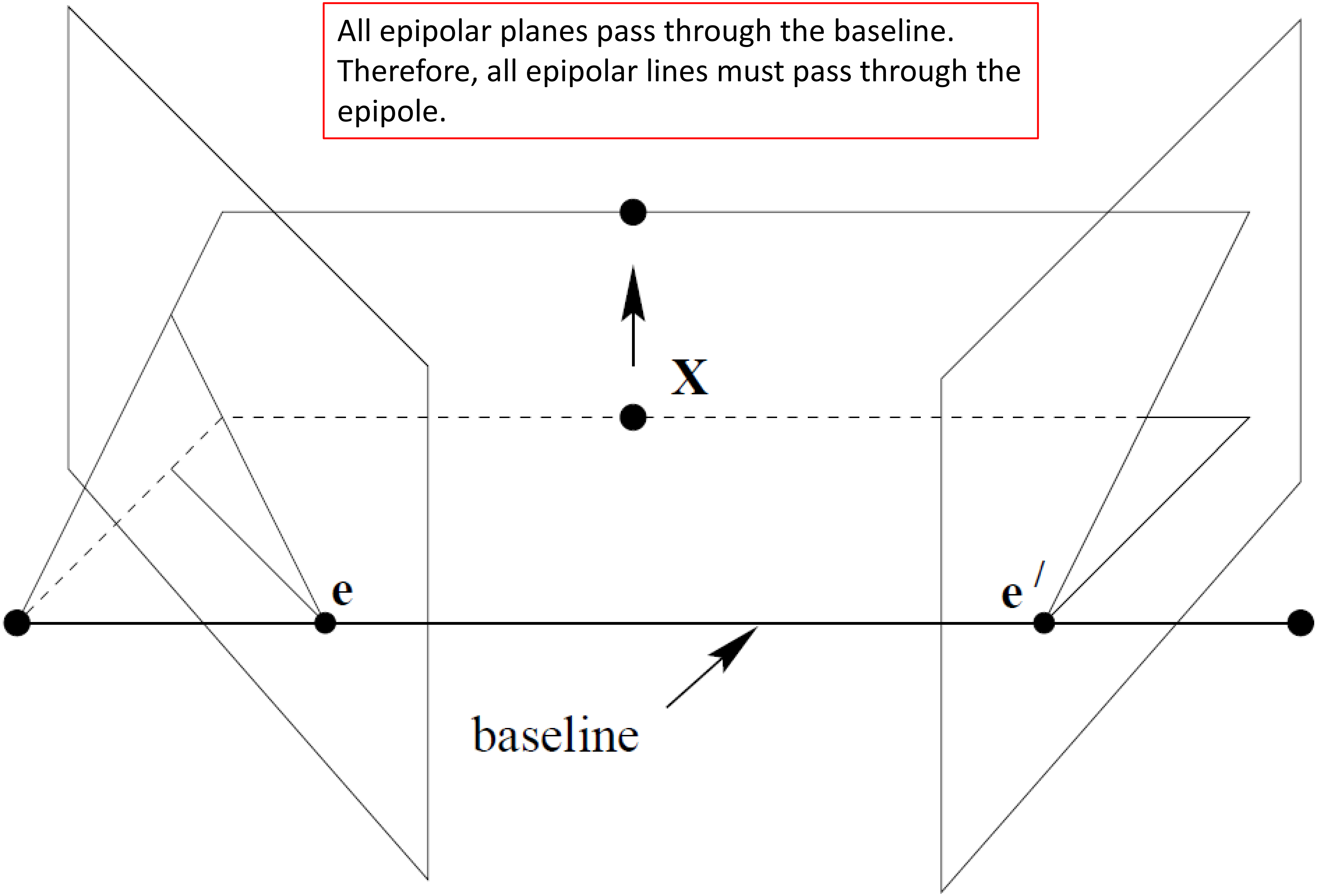
X can lie anywhere along this line.
l’ is the image of this line in second camera.
So x’ can lie anywhere along l’.





Any 2 images of the same planar surface in space are related by a **homography**. Therefore $x' = H_\pi x$

All epipolar planes pass through the baseline. Therefore, all epipolar lines must pass through the epipole.



Epipolar Constraint and Fundamental Matrix

- Epipolar line l' passes through x' and epipole e'
  - $l' = e' \times x' = [e']_x \, x'$
- Since $x'=H_\pi x$, we can write
  - $l' = [e']_x \, H_\pi x = Fx$ where $F = [e']_x \, H_\pi$ is the so-called **fundamental matrix**.
  - Rank(F)=2 (because $[e']_x$ is rank 2 and H_π is rank 3).
- **Fundamental matrix F maps points in camera 1 to corresponding epipolar lines in camera 2.**
  - **$l'=Fx$**

Epipolar Constraint and Fundamental Matrix

- Fundamental matrix F maps points in camera 1 to corresponding epipolar lines in camera 2.
  - $l' = Fx$
- Since x' lies on the epipolar line l', we must have $x'^T l' = 0$.
- This gives us the **epipolar constraint**
  - $x'^T F x = 0$

Epipolar Constraint and Fundamental Matrix

- F has rank 2. Thus, it is not invertible.
- F offers 7 degrees of freedom: 9 minus 2 for
  - rank,
  - and scale (if F satisfies epipolar constraint, then αF also satisfies it).
- A system where only the fundamental matrix is known is called **weakly calibrated**.

Epipolar Constraint and Fundamental Matrix

- For a weakly calibrated system, one can compute for each pixel m_1 in the first frame the corresponding epipolar line l_2 in the second frame and vice versa:

$$\ell_2 = \boldsymbol{F}\,\tilde{m}_1 \quad \Longrightarrow \quad \tilde{m}_2^\top \ell_2 = 0\,,$$
$$\ell_1 = \boldsymbol{F}^\top \tilde{m}_2 \quad \Longrightarrow \quad \tilde{m}_1^\top \ell_1 = 0\,.$$

- In this notation, a vector $l_i = (a, b, c)^T$ describes the epipolar line ax+by +c = 0.
- This creates a reduced search space (1-D) for a stereo matching algorithm (search along epipolar lines).
- Our earlier example of Orthoparallel cameras yielded horizontal epipolar lines (search in x-direction).

Estimation of the Fundamental Matrix

- If the fundamental matrix is not known (uncalibrated system), one can estimate it from point correspondences. Let us study now how this can be done.

Estimation of the Fundamental Matrix

Let us consider the epipolar constraint given by the equation

$$
\begin{aligned}
0 \;=\; \tilde{m}_2^\top \boldsymbol{F}\, \tilde{m}_1 \;=\; \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}^\top \begin{pmatrix} f_{1,1} & f_{1,2} & f_{1,3} \\ f_{2,1} & f_{2,2} & f_{2,3} \\ f_{3,1} & f_{3,2} & f_{3,3} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} \\
= \quad x_1 x_2\, f_{1,1} + y_1 x_2\, f_{1,2} + x_2\, f_{1,3} \\
+ x_1 y_2\, f_{2,1} + y_1 y_2\, f_{2,2} + y_2\, f_{2,3} \\
+ x_1\, f_{3,1} + y_1\, f_{3,2} + f_{3,3}.
\end{aligned}
$$

Defining the following two vectors for correspondences and matrix entries,

$$
\begin{aligned}
s &:= (x_1 x_2,\, y_1 x_2,\, x_2,\, x_1 y_2,\, y_1 y_2,\, y_2,\, x_1,\, y_1,\, 1)^\top, \\
\boldsymbol{f} &:= (f_{1,1},\, f_{1,2},\, f_{1,3},\, f_{2,1},\, f_{2,2},\, f_{2,3},\, f_{3,1},\, f_{3,2},\, f_{3,3})^\top,
\end{aligned}
$$

we can write the epipolar constraint as an inner product:

$$
0 = \tilde{m}_2^\top \boldsymbol{F}\, \tilde{m}_1 = s^\top \boldsymbol{f}.
$$

Estimation of the Fundamental Matrix

- Take N>=8 conjugated point pairs.
- Sum up the squared deviations from the N constraints and minimise the resulting quadratic form

$$E(f) = \sum_{i=1}^{N} (s_i^\top f)^2 = f^\top \left(\sum_{i=1}^{N} s_i s_i^\top \right) f$$

  with explicit constraint $||f|| = 1$ to avoid the trivial solution $f = 0$.
- We want to minimise $f^T A f$ with the constraint that vector f has unit norm.
  – This must be familiar to you now!

Estimation of the Fundamental Matrix

- The solution to this problem is given by the normalised eigenvector to the smallest eigenvalue of the symmetric 9 × 9 matrix

$$\sum_{i=1}^{N} s_i s_i^{\top}$$

Finding Conjugated Points

- Correlation-based Methods:
  - Move along epipolar line and find the point where **correlation coefficient is maximised**.
- Variational Methods:
  - A family of much more elegant methods with many other applications.

Correlation Coefficient

- $$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$
- Measures similarity of patches X and Y.
- Just a normalised inner product with values in the the range [-1,1].
- Higher value implies that X and Y are similar.

Basic Stereo Algorithm Outline

- Step 1: Find some correspondences (conjugated points) and estimate fundamental matrix F.
- Step 2: For every point x in image 1, compute corresponding epipolar line l' in image 2 using l'=Fx.
- Compute correlation coefficients between patch P around x and patches along l'.
  - x' is the location with max correlation coefficient .

Stereo Reconstruction

Stereo Reconstruction