Reconstruction/Triangulation
Old book Ch11.1 F&P
New book Ch7.2 F&P

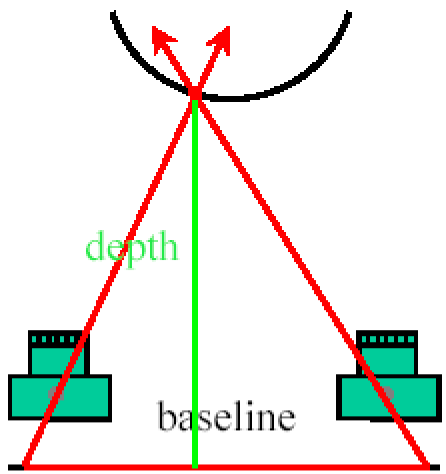
Guido Gerig

CS 6320, S 2015

(modified from original slides by J. Ponce and by Marc Pollefeys)

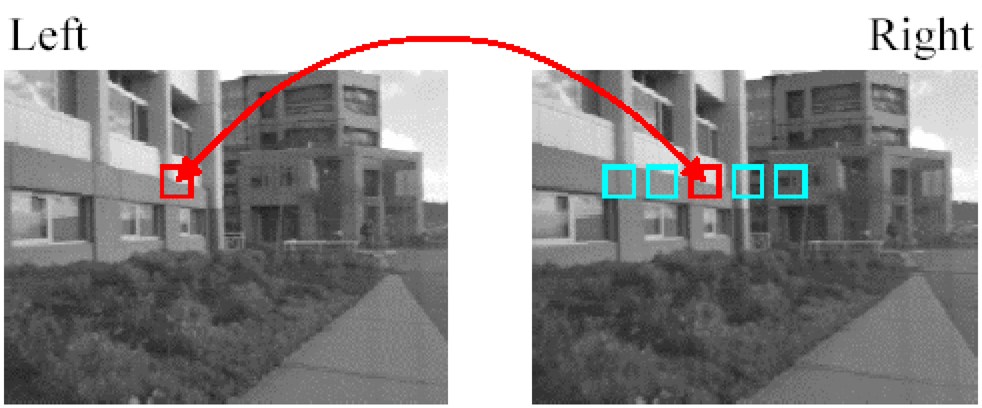
Credits: J. Ponce, M. Pollefeys, A. Zisserman & S. Lazebnik

Reconstruction



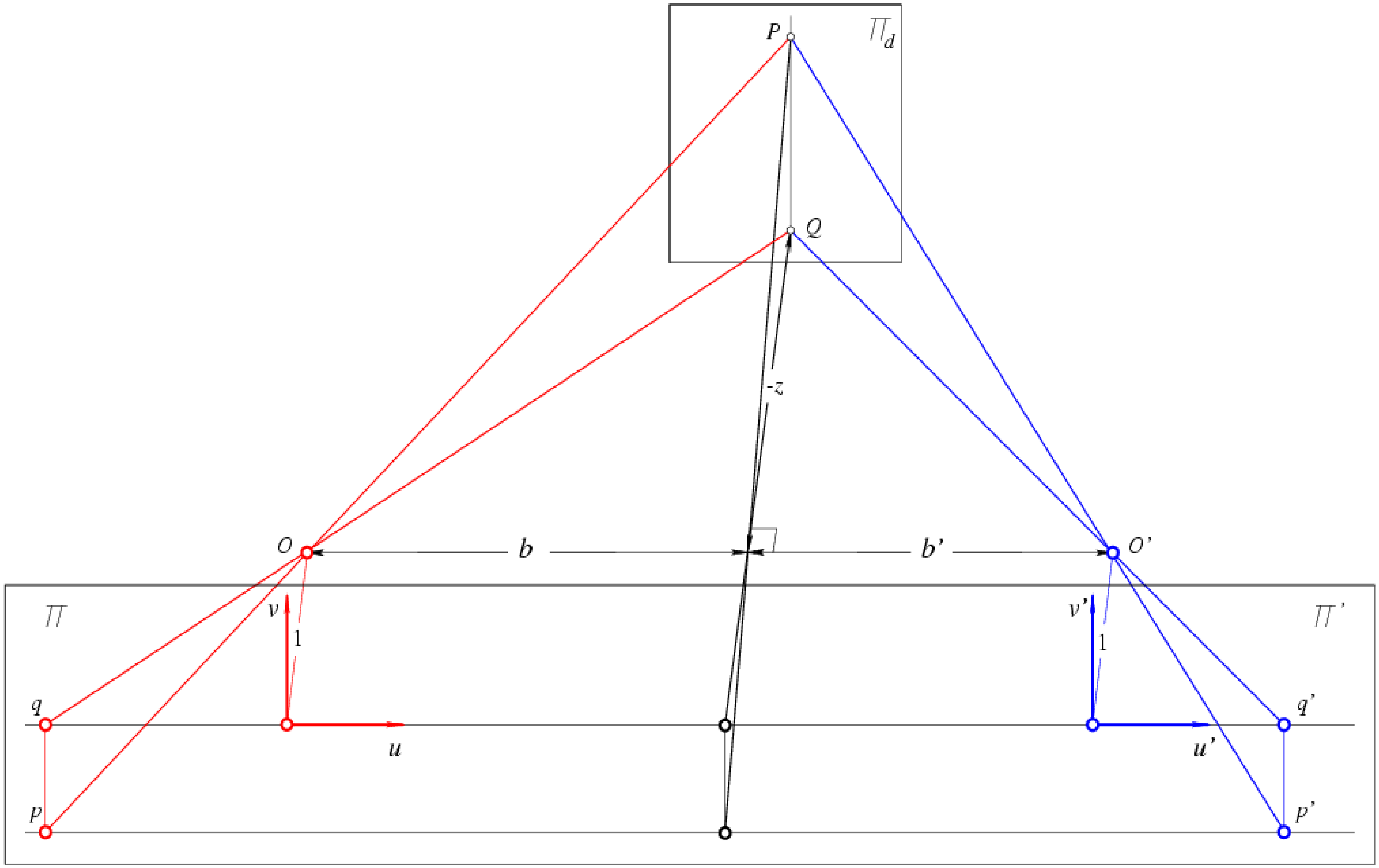
Triangulate on two images of the same point to recover depth.

- Feature matching across views
- Calibrated cameras



Only need to match features across epipolar lines

Reconstruction from Rectified Images

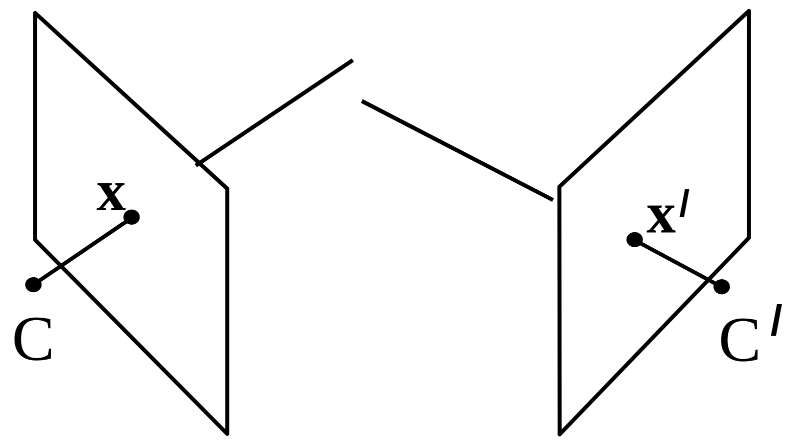


Disparity: $d=u'-u$. Depth: $z = -B/d$.

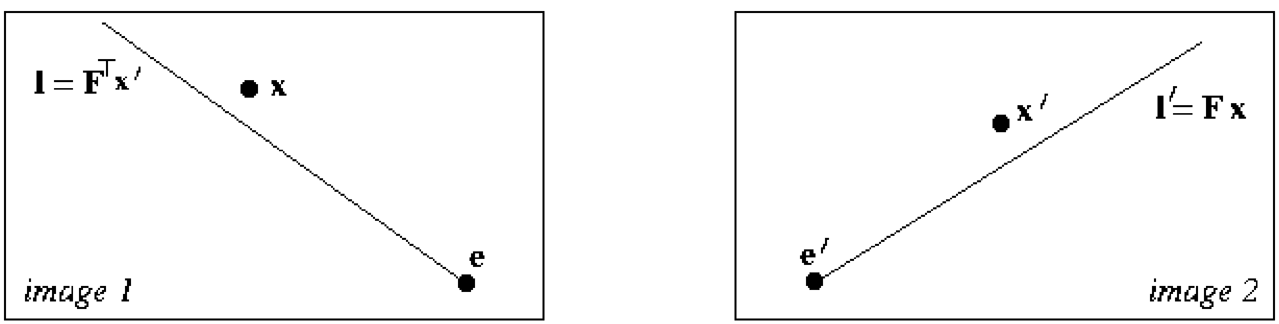
Problem statement

<u>Given:</u> corresponding measured (i.e. noisy) points $\mathbf{x}$ and $\mathbf{x}^{\prime}$, and cameras (exact) P and P$^{\prime}$, compute the 3D point $\mathbf{X}$

Problem: in the presence of noise, back projected rays do not intersect



rays are skew in space



Measured points do not lie on corresponding epipolar lines

Problem statement

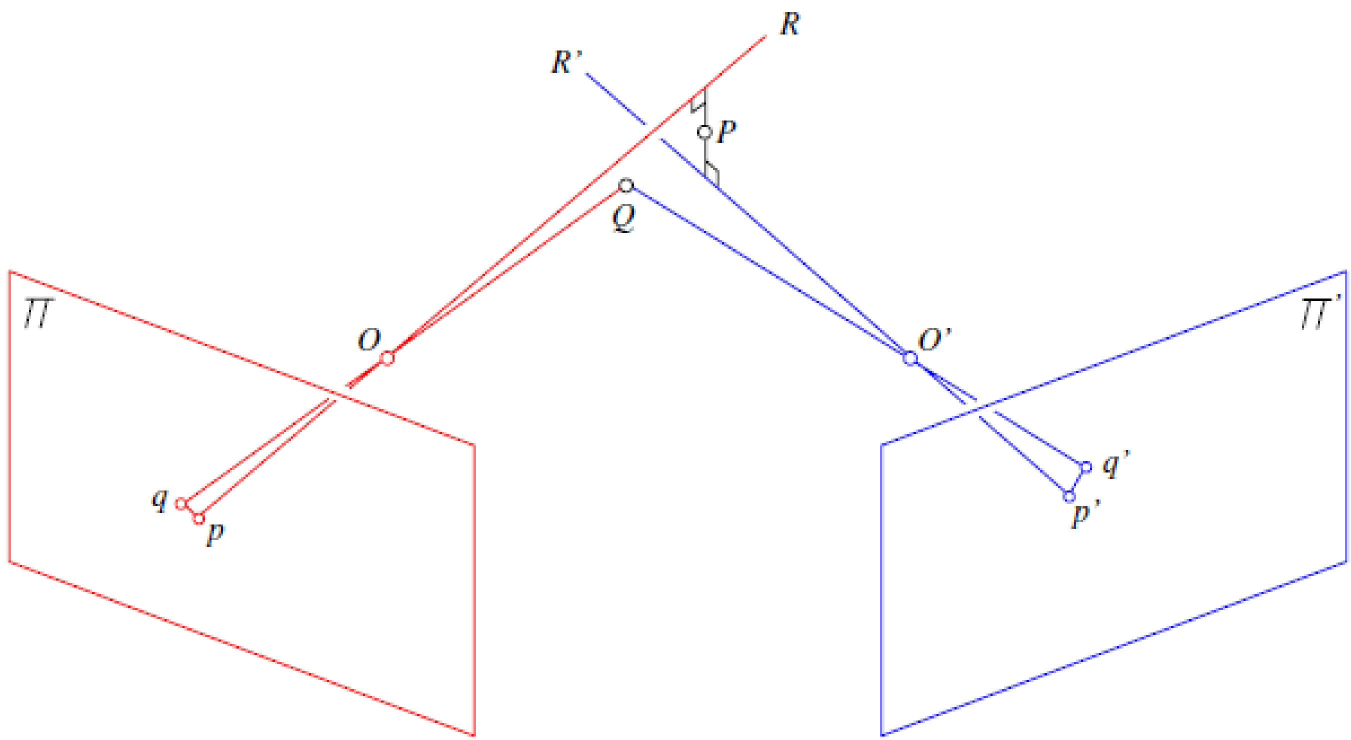
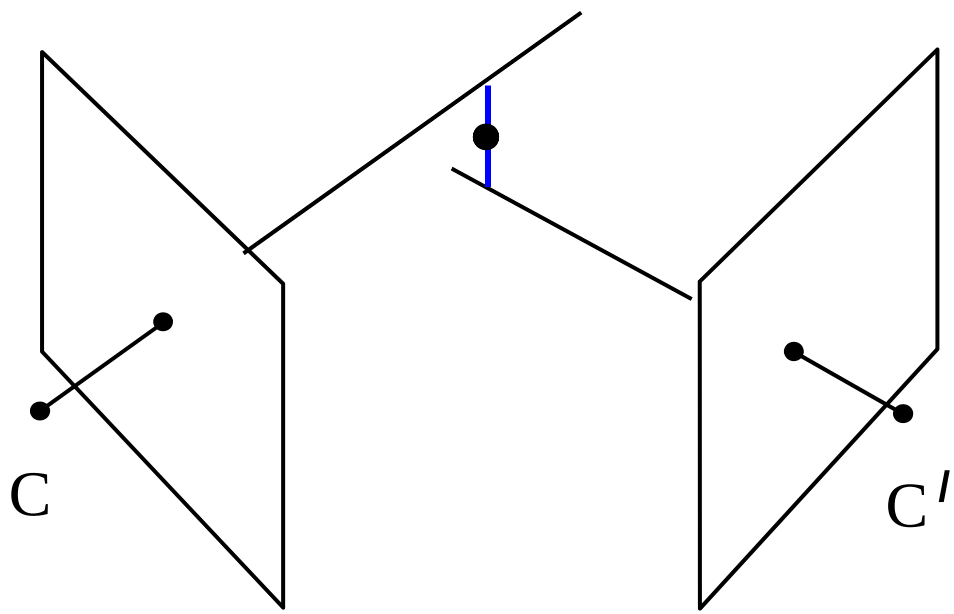


Figure 13.4. Triangulation in the presence of measurement errors. See text for details.

1. Vector solution



Compute the mid-point of the shortest line between the two rays

2. Linear triangulation (algebraic solution)

Use the equations $\mathbf{x} = \mathtt{P}\mathbf{X}$ and $\mathbf{x}' = \mathtt{P}'\mathbf{X}$ to solve for $\mathbf{X}$

For the first camera:

$$\mathtt{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^{1\top} \\ \mathbf{p}^{2\top} \\ \mathbf{p}^{3\top} \end{bmatrix}$$

where $\mathbf{p}^{i\top}$ are the rows of $\mathtt{P}$

- eliminate unknown scale in $\lambda\mathbf{x} = \mathtt{P}\mathbf{X}$ by forming a cross product $\mathbf{x} \times (\mathtt{P}\mathbf{X}) = \mathbf{0}$

$$\begin{aligned} x(\mathbf{p}^{3\top}\mathbf{X}) - (\mathbf{p}^{1\top}\mathbf{X}) &= 0 \\ y(\mathbf{p}^{3\top}\mathbf{X}) - (\mathbf{p}^{2\top}\mathbf{X}) &= 0 \\ x(\mathbf{p}^{2\top}\mathbf{X}) - y(\mathbf{p}^{1\top}\mathbf{X}) &= 0 \end{aligned}$$

- rearrange as (first two equations only)

$$\begin{bmatrix} x\mathbf{p}^{3\top} - \mathbf{p}^{1\top} \\ y\mathbf{p}^{3\top} - \mathbf{p}^{2\top} \end{bmatrix} \mathbf{X} = \mathbf{0}$$

Similarly for the second camera:

$$\begin{bmatrix} x'\mathbf{p}'^{3\top} - \mathbf{p}'^{1\top} \\ y'\mathbf{p}'^{3\top} - \mathbf{p}'^{2\top} \end{bmatrix} \mathbf{X} = \mathbf{0}$$

Collecting together gives

$$\mathtt{A}\mathbf{X} = \mathbf{0}$$

where $\mathtt{A}$ is the 4×4 matrix

$$\mathtt{A} = \begin{bmatrix} x\mathbf{p}^{3\top} - \mathbf{p}^{1\top} \\ y\mathbf{p}^{3\top} - \mathbf{p}^{2\top} \\ x'\mathbf{p}'^{3\top} - \mathbf{p}'^{1\top} \\ y'\mathbf{p}'^{3\top} - \mathbf{p}'^{2\top} \end{bmatrix}$$

from which $\mathbf{X}$ can be solved up to scale.

Problem: does not minimize anything meaningful

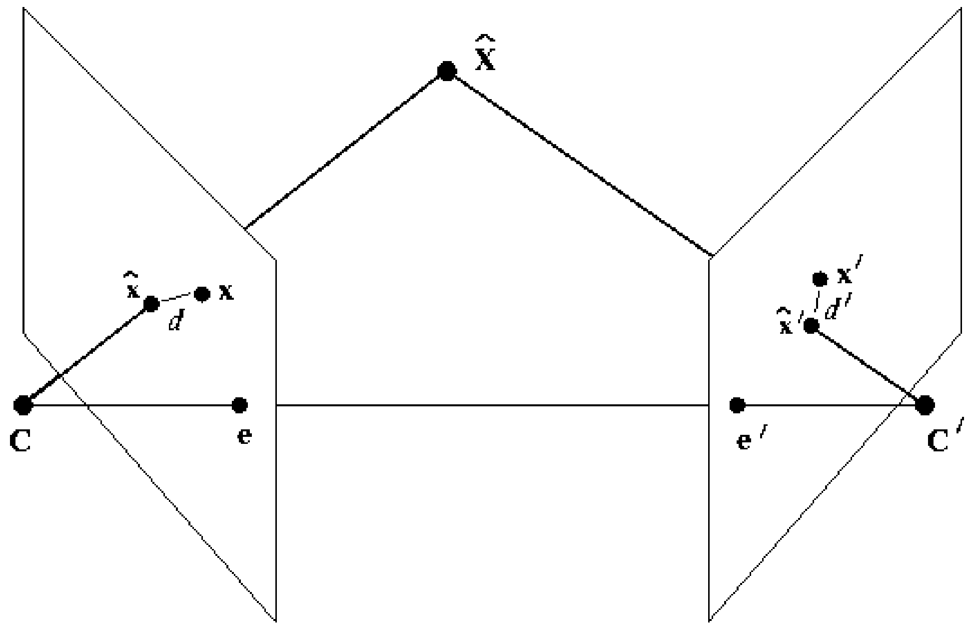
Advantage: extends to more than two views

3. Minimizing a geometric/statistical error

The idea is to estimate a 3D point $\widehat{\mathbf{X}}$ which exactly satisfies the supplied camera geometry, so it projects as

$$\hat{\mathbf{x}} = \mathtt{P}\widehat{\mathbf{X}} \qquad \hat{\mathbf{x}}' = \mathtt{P}'\widehat{\mathbf{X}}$$

and the aim is to estimate $\widehat{\mathbf{X}}$ from the image measurements $\mathbf{x}$ and $\mathbf{x}'$.



$$\min_{\widehat{\mathbf{X}}} \quad \mathcal{C}(\mathbf{x}, \mathbf{x}') = d(\mathbf{x}, \hat{\mathbf{x}})^2 + d(\mathbf{x}', \hat{\mathbf{x}}')^2$$

where $d(*, *)$ is the Euclidean distance between the points.

- It can be shown that if the measurement noise is Gaussian mean zero, $\sim N(0, \sigma^2)$, then minimizing geometric error is the Maximum Likelihood Estimate of X

- The minimization appears to be over three parameters (the position X), but the problem can be reduced to a minimization over one parameter

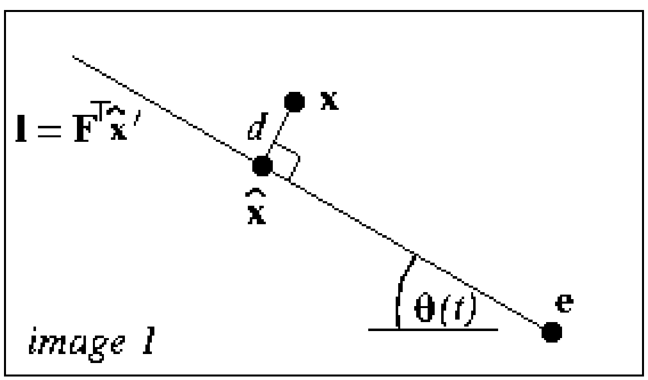
Different formulation of the problem

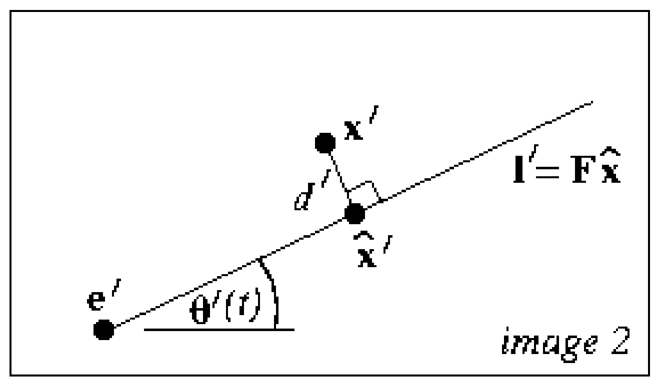
The minimization problem may be formulated differently:

- Minimize

$$d(\mathbf{x}, \mathbf{l})^2 + d(\mathbf{x}', \mathbf{l}')^2$$

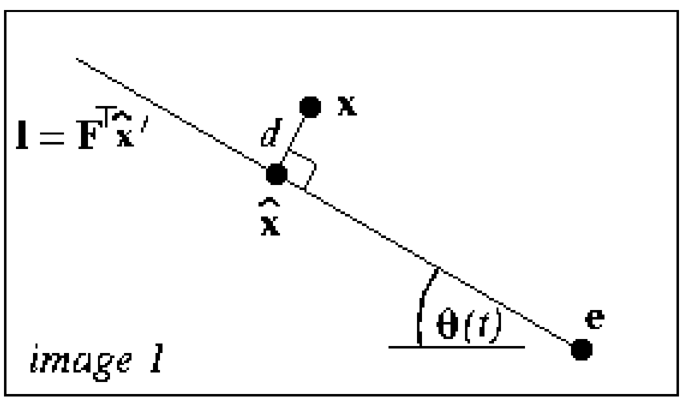
- $\mathbf{l}$ and $\mathbf{l}'$ range over all choices of corresponding epipolar lines.
- $\hat{\mathbf{x}}$ is the closest point on the line $\mathbf{l}$ to $\mathbf{x}$.
- Same for $\hat{\mathbf{x}}'$.

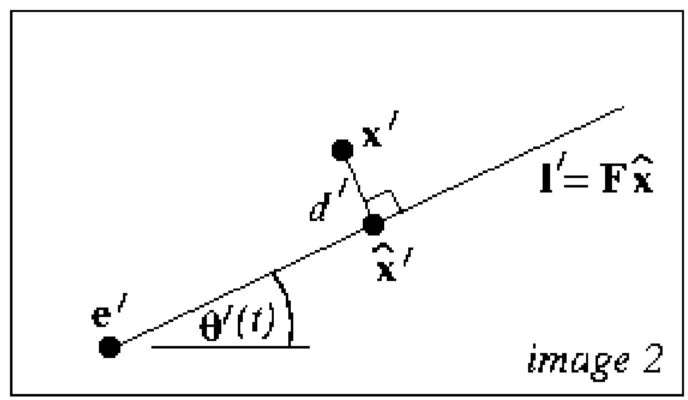




Minimization method

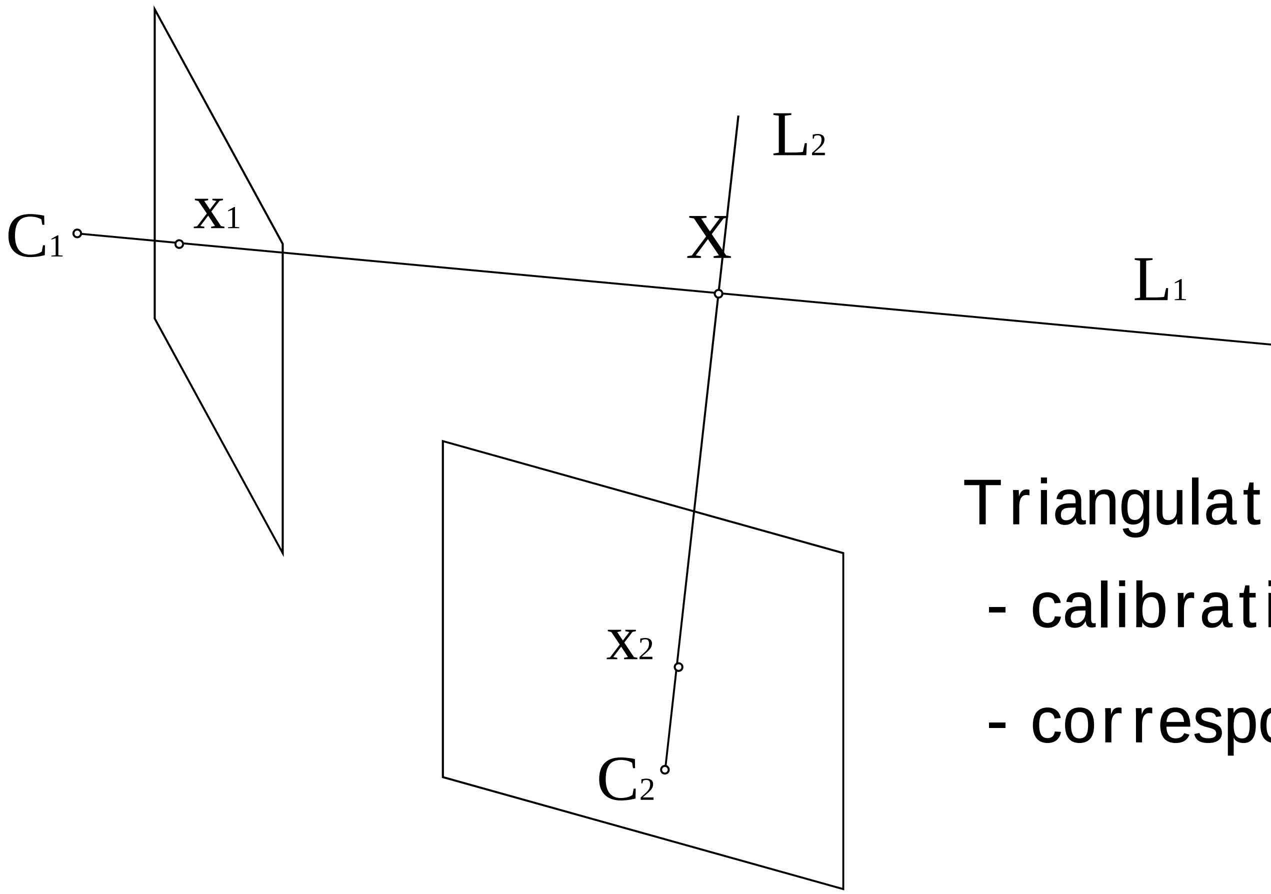
- Parametrize the pencil of epipolar lines in the first image by t, such that the epipolar line is $\mathbf{l}(t)$
- Using F compute the corresponding epipolar line in the second image $\mathbf{l}'(t)$
- Express the distance function $d(\mathbf{x}, \mathbf{l})^2 + d(\mathbf{x}', \mathbf{l}')^2$ explicitly as a function of t
- Find the value of t that minimizes the distance function
- Solution is a 6th degree polynomial in t





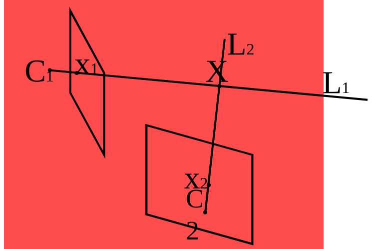
More slides for self-study.

Triangulation (finally!)



Triangulation

- calibration
- correspondences



Triangulation

- Backprojection

$$\lambda \mathrm{x} = \mathbf{P}\mathrm{X} \qquad \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathrm{P}_1 \\ \mathrm{P}_2 \\ \mathrm{P}_3 \end{bmatrix} \mathrm{X}$$

$$\begin{aligned} \mathrm{P}_3\mathrm{X}x &= \mathrm{P}_1\mathrm{X} \\ \mathrm{P}_3\mathrm{X}y &= \mathrm{P}_2\mathrm{X} \end{aligned} \qquad \begin{bmatrix} \mathrm{P}_3 x - \mathrm{P}_1 \\ \mathrm{P}_3 y - \mathrm{P}_2 \end{bmatrix} \mathrm{X} = 0$$

- Triangulation

$$\begin{bmatrix} \mathrm{P}_3 x - \mathrm{P}_1 \\ \mathrm{P}_3 y - \mathrm{P}_2 \\ \mathrm{P}'_3 x' - \mathrm{P}'_1 \\ \mathrm{P}'_3 y' - \mathrm{P}'_2 \end{bmatrix} \mathrm{X} = 0 \qquad \begin{bmatrix} \frac{1}{\mathrm{P}_3\tilde{\mathrm{X}}} \begin{pmatrix} \mathrm{P}_3 x - \mathrm{P}_1 \\ \mathrm{P}_3 y - \mathrm{P}_2 \end{pmatrix} \\ \frac{1}{\mathrm{P}'_3\tilde{\mathrm{X}}} \begin{pmatrix} \mathrm{P}'_3 x - \mathrm{P}'_1 \\ \mathrm{P}'_3 y - \mathrm{P}'_2 \end{pmatrix} \end{bmatrix} \mathrm{X} = 0$$

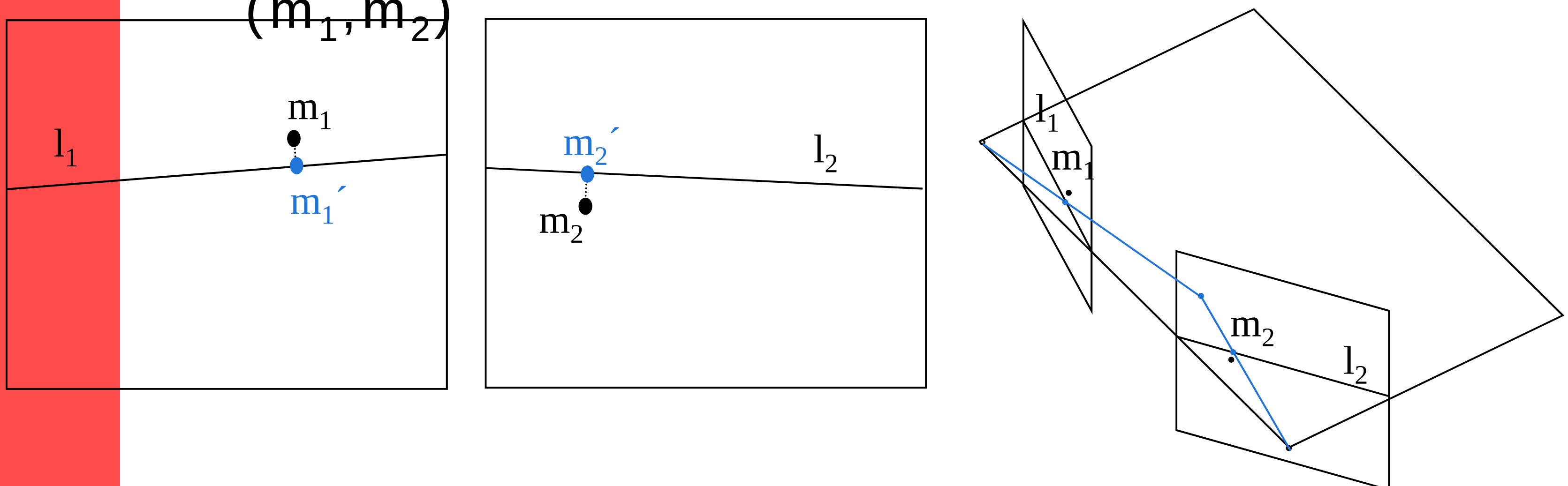
Iterative least-squares

- Maximum Likelihood Triangulation

$$\arg\min_{\mathrm{X}} \sum_i \left(\mathrm{x}_i - \lambda^{-1} \mathbf{P}_i \mathrm{X} \right)^2$$

Optimal 3D point in epipolar plane

- Given an epipolar plane, find best 3D point for (m_1, m_2)



Select closest points (m_1', m_2') on epipolar lines

Obtain 3D point through exact triangulation

Guarantees minimal reprojection error (given this epipolar plane)

Non-iterative optimal solution

- Reconstruct matches in projective frame by minimizing the reprojection error

$$D(\mathbf{m}_1, \mathbf{P}_1\mathbf{M})^2 + D(\mathbf{m}_2, \mathbf{P}_2\mathbf{M})^2 \quad \textbf{3DOF}$$

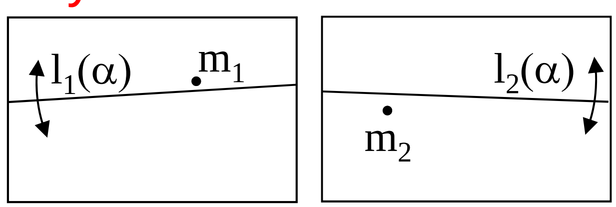
- Non-iterative method

Determine the epipolar plane for reconstruction (Hartley and Sturm, CVIU´97)

$$D(\mathbf{m}_1, \mathbf{l}_1(\alpha))^2 + D(\mathbf{m}_2, \mathbf{l}_2(\alpha))^2 \text{ (polynomial of degree 6)}$$

Reconstruct optimal point from selected epipolar plane

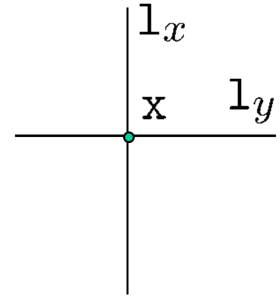
Note: only works for two views



1DOF

Backprojection

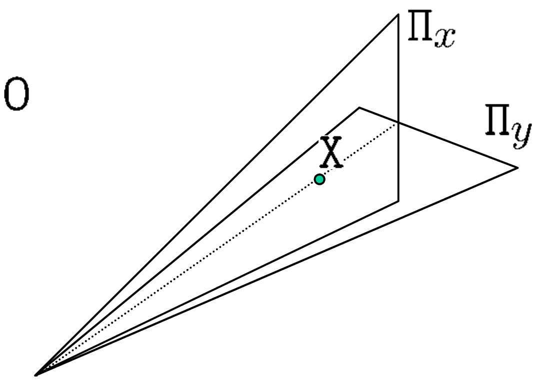
- Represent point as intersection of row and column



$$\mathtt{x} = \mathtt{l}_x \times \mathtt{l}_y \text{ with } \mathtt{l}_x = \begin{bmatrix} -1 \\ 0 \\ x \end{bmatrix}, \mathtt{l}_y = \begin{bmatrix} 0 \\ -1 \\ y \end{bmatrix}$$

$$\Pi = \mathtt{P}^\top \mathtt{l}$$

$$\begin{bmatrix} \Pi_x^\top \\ \Pi_y^\top \end{bmatrix} \mathtt{X} = 0 \qquad \begin{bmatrix} \mathtt{l}_x^\top \mathtt{P} \\ \mathtt{l}_y^\top \mathtt{P} \end{bmatrix} \mathtt{X} = 0$$



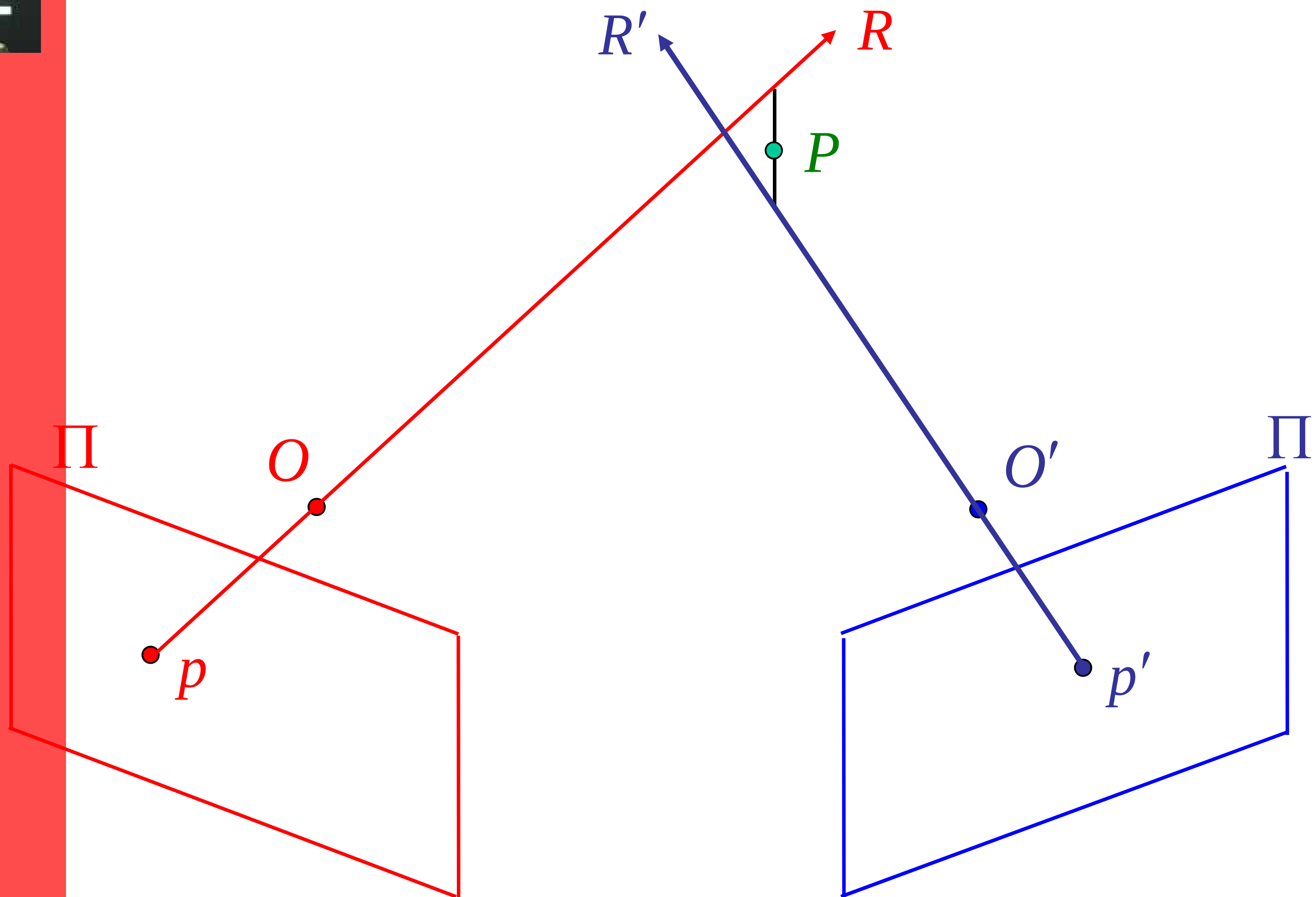
- Condition for solution?

$$\det \begin{bmatrix} \mathtt{l}_x^\top \mathtt{P} \\ \mathtt{l}_y^\top \mathtt{P} \\ \mathtt{l}_{x'}^\top \mathtt{P}' \\ \mathtt{l}_{y'}^\top \mathtt{P}' \end{bmatrix} = 0$$

Useful presentation for deriving and understanding multiple view geometry (notice 3D planes are linear in 2D point coordinates)

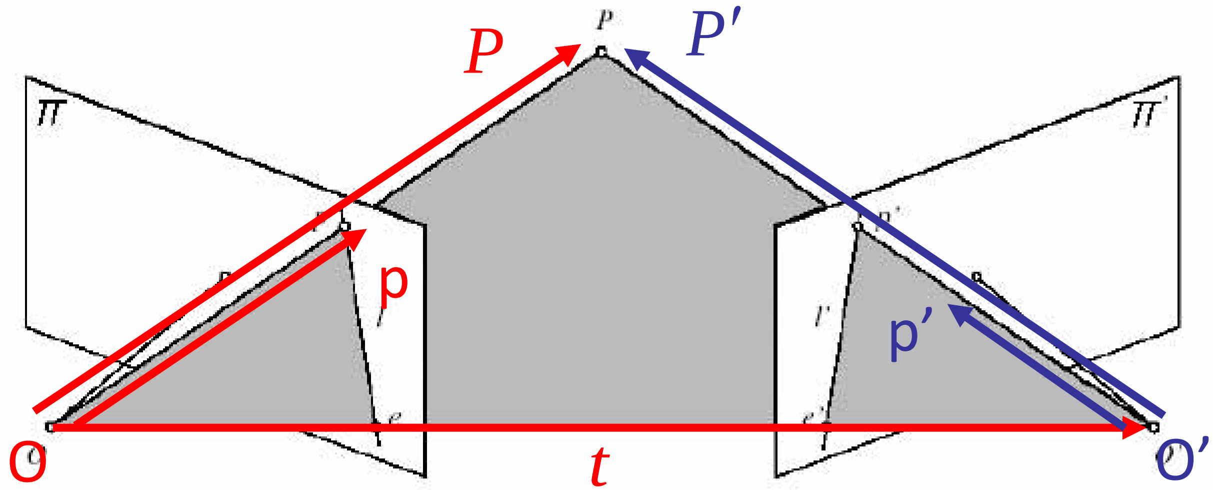
Reconstruction

Geometric Reconstruction



Reconstruction

$$p = f\frac{P}{Z}$$



$$p' = f'\frac{P'}{Z'}$$

FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

$$P = RP' + t$$

$$P' = R^{-1}(P - t) = R^{T}(P - t)$$

Reconstruction

$$p' = f' \frac{P'}{Z'}$$

$$P' = R^T(P-t) = R'(P-t)$$

$$R' = \begin{bmatrix} R_1'^T \\ R_2'^T \\ R_3'^T \end{bmatrix}$$

$$p' = f' \frac{R'(P-t)}{R_3'^T(P-t)}$$

$$x' = f' \frac{R_1'^T(P-t)}{R_3'^T(P-t)}$$ Equation 1

$$p = f \frac{P}{Z} \Rightarrow P = \frac{pZ}{f}$$ Equation 2

$$\boxed{Z = f \frac{(x'R_3' - f'R_1')^T t}{(x'R_3' - f'R_1')^T p}}$$

(From equations 1 and 2)

Reconstruction up to a Scale Factor

- Assume that intrinsic parameters of both cameras are known
- Essential Matrix is known up to a scale factor (for example,
estimated from the 8 point algorithm).

Reconstruction up to a Scale Factor

$$\mathcal{E} = k[t_\times]R$$

$$\mathcal{E}\mathcal{E}^T = k^2[t_\times]RR^T[t_\times]^T = k^2[t_\times][t_\times]^T = \begin{bmatrix} k^2\left(T_Y^2+T_Z^2\right) & -k^2T_XT_Y & -k^2T_XT_Z \\ -k^2T_XT_Y & k^2\left(T_X^2+T_Z^2\right) & -k^2T_YT_Z \\ -k^2T_XT_Z & -k^2T_YT_Z & k^2\left(T_X^2+T_Y^2\right) \end{bmatrix}$$

$$Trace\left[\mathcal{E}\mathcal{E}^T\right] = 2k^2\left(T_X^2+T_Y^2+T_Z^2\right) = 2k^2\|t\|^2$$

$$\frac{\mathcal{E}}{|k|\|t\|} = \operatorname{sgn}(k)\frac{[t_\times]}{\|t\|}R = \operatorname{sgn}(k)\left[\left(\frac{t}{\|t\|}\right)_\times\right]R = \operatorname{sgn}(k)\left[\hat{t}_\times\right]R = \hat{E}$$

$$\hat{E}\hat{E}^T = \left[\hat{t}_\times\right]\left[\hat{t}_\times\right]^T = \begin{bmatrix} 1-\hat{T}_X^2 & -\hat{T}_X\hat{T}_Y & -\hat{T}_X\hat{T}_Z \\ -\hat{T}_X\hat{T}_Y & 1-\hat{T}_Y^2 & -\hat{T}_Y\hat{T}_Z \\ -\hat{T}_X\hat{T}_Z & -\hat{T}_Y\hat{T}_Z & 1-\hat{T}_Z^2 \end{bmatrix}$$

Reconstruction up to a Scale Factor

$$\hat{E} = \begin{bmatrix} \hat{E}_1^T \\ \hat{E}_2^T \\ \hat{E}_3^T \end{bmatrix} \qquad R = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix}$$

Let $w_i = \hat{E}_i \times \hat{t},\ i \in \{1,2,3\}$

It can be proved that

$$R_1 = w_1 + w_2 \times w_3$$
$$R_2 = w_2 + w_3 \times w_1$$
$$R_3 = w_3 + w_1 \times w_2$$

Reconstruction up to a Scale Factor

We have two choices of **t**, ($\mathbf{t}^+$ and $\mathbf{t}^-$) because of sign ambiguity
and two choices of **E,** (E^+ and E^-).

This gives us four pairs of translation vectors and rotation matrices.

Reconstruction up to a Scale Factor

Given $\hat{E}$ and $\hat{t}$

1. Construct the vectors **w**, and compute R
2. Reconstruct the Z and Z' for each point
3. If the signs of Z and Z' of the reconstructed points are
   a) both negative for some point, change the sign of $\hat{t}$ and go to step 2.
   b) different for some point, change the sign of each entry of $\hat{E}$ and go to step 1.
   c) both positive for all points, exit.

$$Z = f\frac{\left(x'R_3' - f'R_1'\right)^T t}{\left(x'R_3' - f'R_1'\right)^T p}$$

$$Z' = -f'\frac{\left(xR_3 - fR_1\right)^T (t)}{\left(xR_3 - fR_1\right)^T p'}$$

3D Reconstruction

[Trucco pp. 161]

- Three cases:
  - a) intrinsic and extrinsic parameters known: Solve reconstruction by triangulation: ray intersection
  - b) only intrinsic parameters known: estimate essential matrix E up to scaling
  - c) intrinsic and extrinsic parameters not known: estimate fundamental matrix F, reconstruction up to global, projective transformation

Run Example

Demo for stereo reconstruction:

http://research.microsoft.com/en-us/um/people/zhang/inria/calibenv/calibenv.html