Introduction to computer vision X

Instructors: Jean Ponce and Matthew Trager
jean.ponce@inria.fr, matthew.trager@cims.nyu.edu

TAs: Jiachen (Jason) Zhu and Sahar Siddiqui
jiachen.zhu@nyu.edu, ss12414@nyu.edu

Slides will be available after class at:
https://mtrager.github.io/introCV-fall2019/

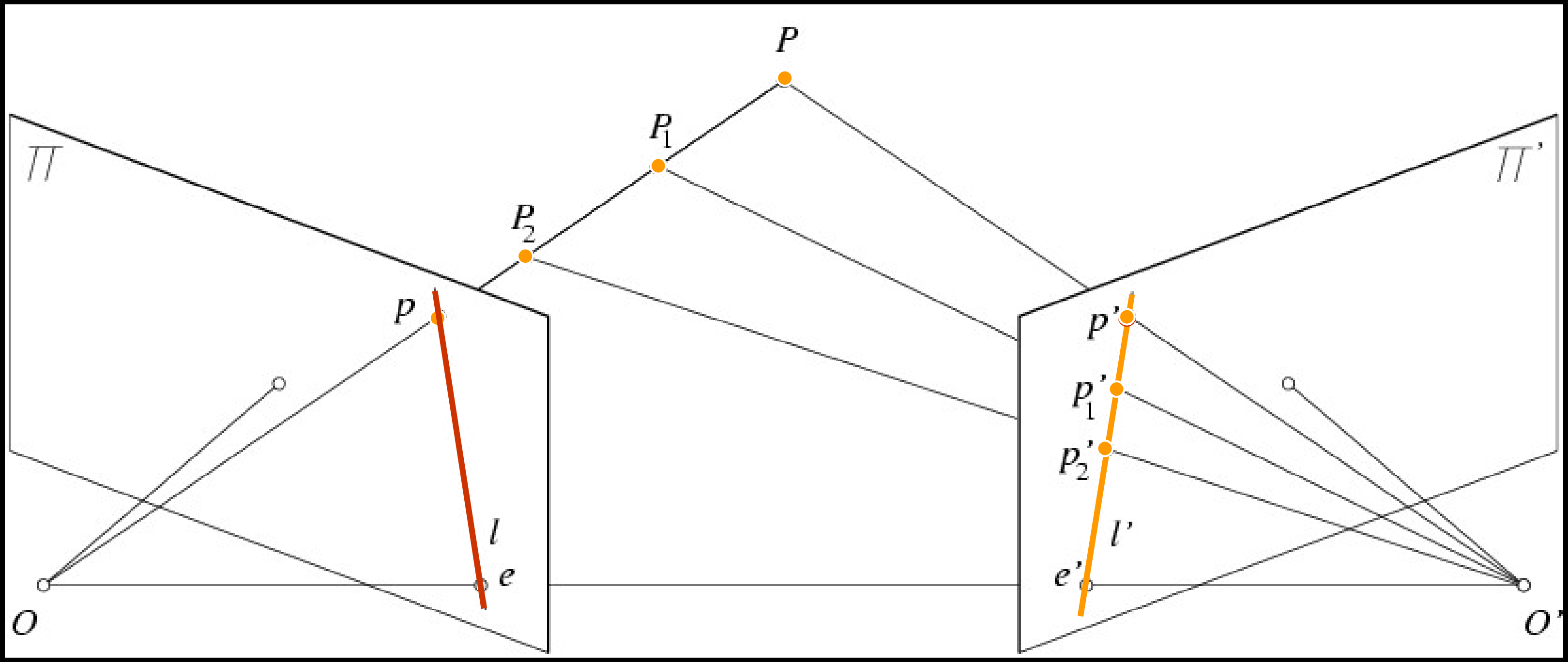
Stereo

- Essential and fundamental matrices
- 8-point agorithm
- Rectification
- Triangulation
- Fusion algorithms

Structure from motion

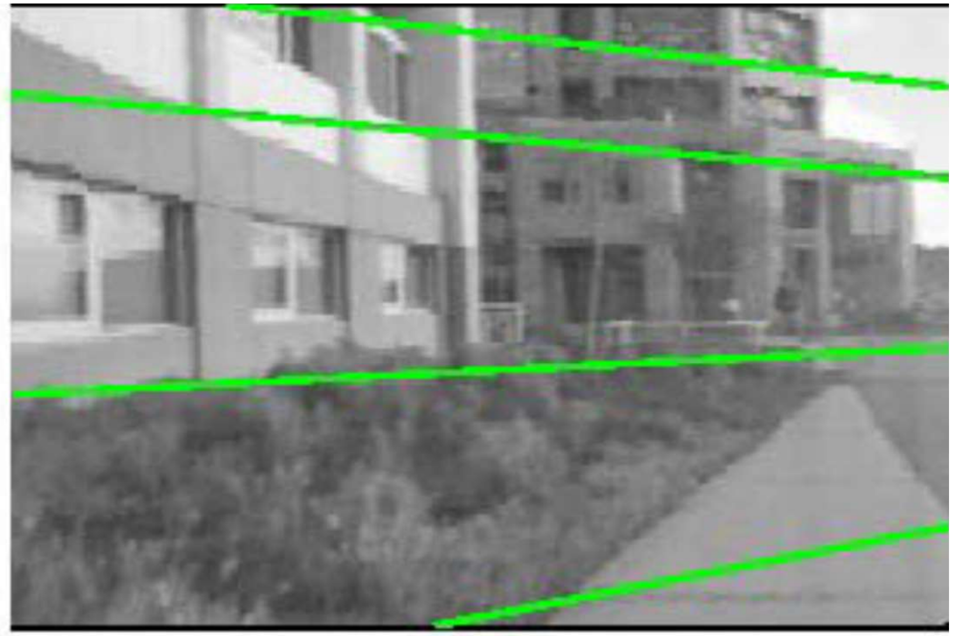
- Problem definition
- Ambiguities
- Euclidean SFM from the essential matrix
- Affine SFM from two views
- Affine SFM from multiple views
- Projective SFM

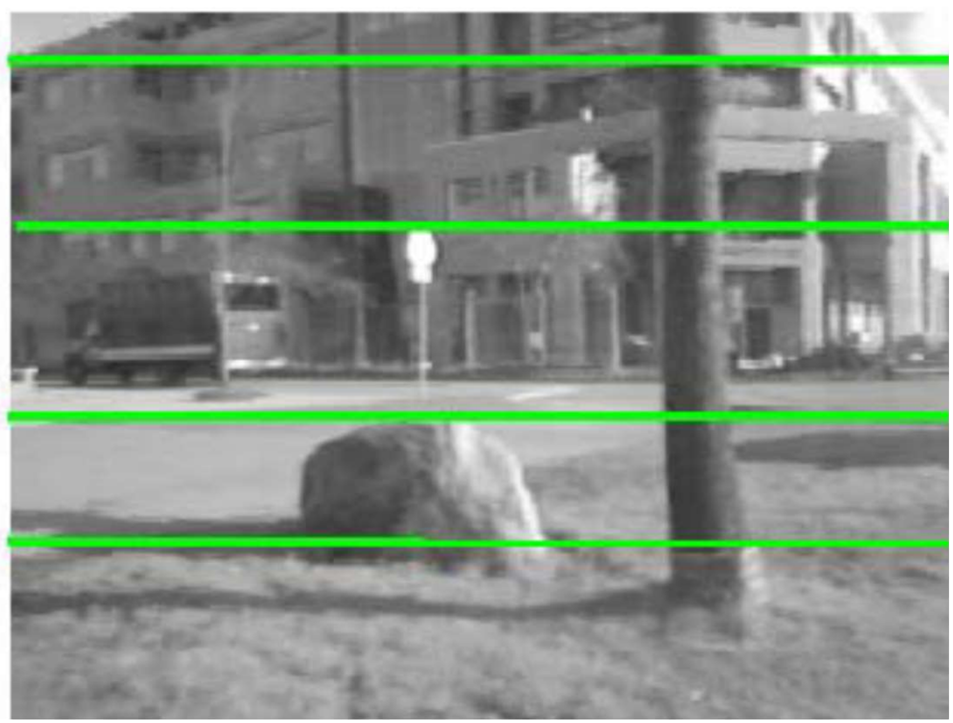
Epipolar Constraint

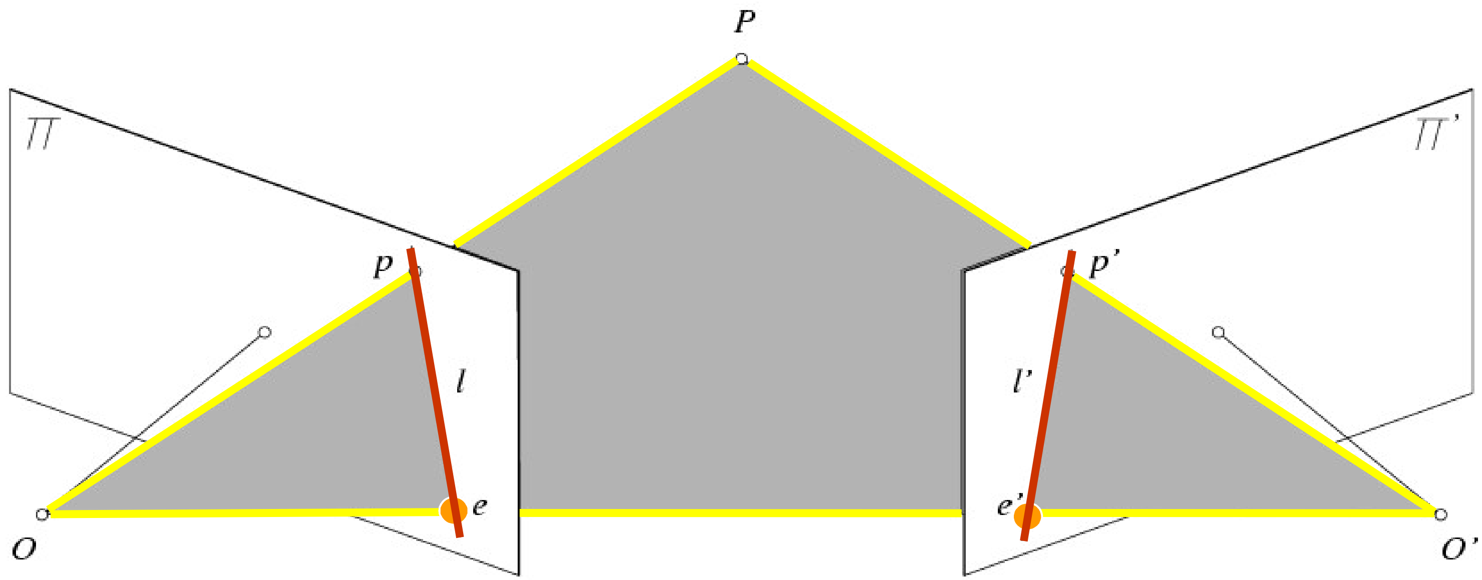


- Potential matches for *p* have to lie on the corresponding epipolar line *l'*.

- Potential matches for *p'* have to lie on the corresponding epipolar line *l*.

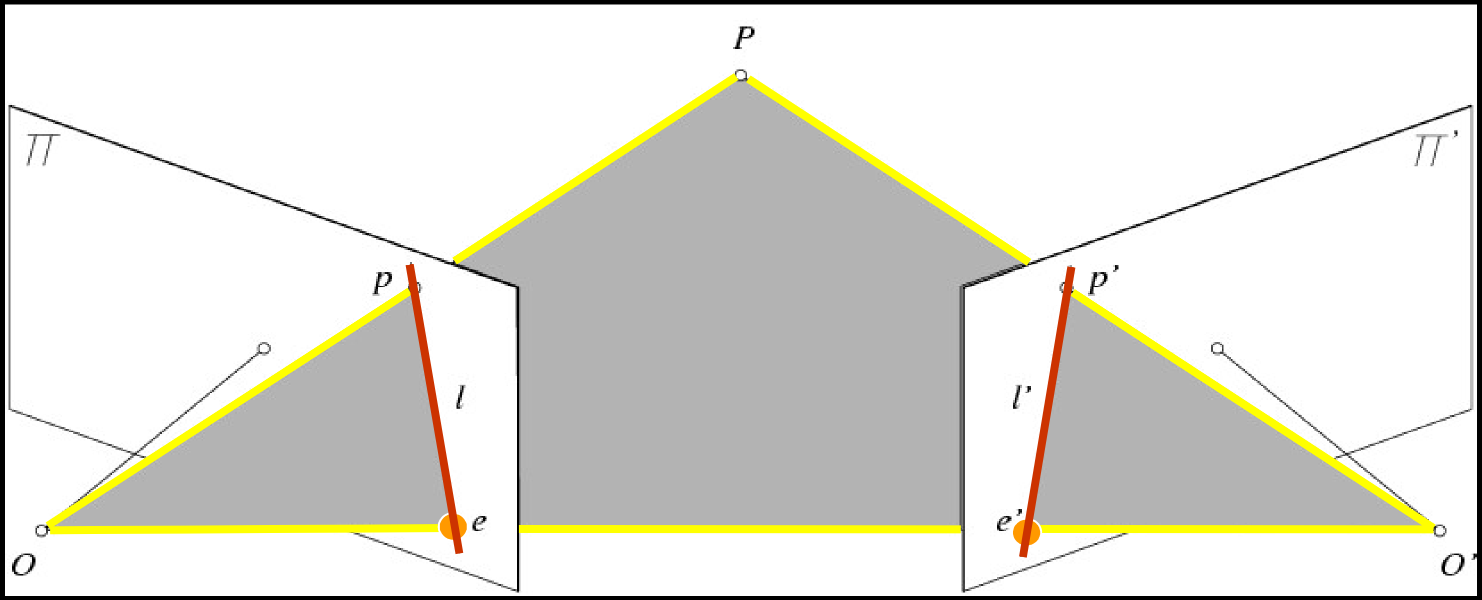







$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$

Epipolar Constraint: Calibrated Case



$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0 \quad \text{with} \begin{cases} \boldsymbol{p} = (u, v, 1)^T \\ \boldsymbol{p}' = (u', v', 1)^T \\ \mathcal{M} = (\mathrm{Id} \quad \boldsymbol{0}) \\ \mathcal{M}' = (\mathcal{R}^T, -\mathcal{R}^T\boldsymbol{t}) \end{cases}$$

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0 \quad \text{with} \quad \mathcal{E} = [\boldsymbol{t}_\times]\mathcal{R}$$

Essential Matrix
(Longuet-Higgins, 1981)

Properties of the Essential Matrix

- $E\,p'$ is the epipolar line associated with p'.
- $E^T p$ is the epipolar line associated with p.
- $E\,e'=0$ and $E^T e=0$.
- E is singular.
- E has two equal non-zero singular values (Huang and Faugeras, 1989).

Epipolar Constraint: Small Motions

$$\mathcal{R}(\boldsymbol{a},\theta) = e^{\theta[\boldsymbol{a}_\times]} \stackrel{\text{def}}{=} \sum_{i=0}^{+\infty} \frac{1}{i!}(\theta[\boldsymbol{a}_\times])^i$$

To First-Order:

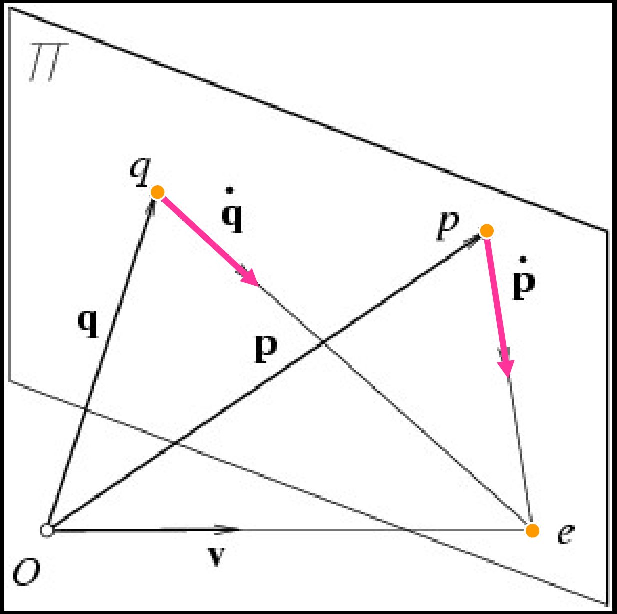
$$\begin{cases} \boldsymbol{t} = \delta t\, \boldsymbol{v} \\ \mathcal{R} = \mathrm{Id} + \delta t\, [\boldsymbol{\omega}_\times] \\ \boldsymbol{p}' = \boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}} \end{cases}$$

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0 \quad \text{with} \quad \mathcal{E} = [\boldsymbol{t}_\times]\mathcal{R}$$

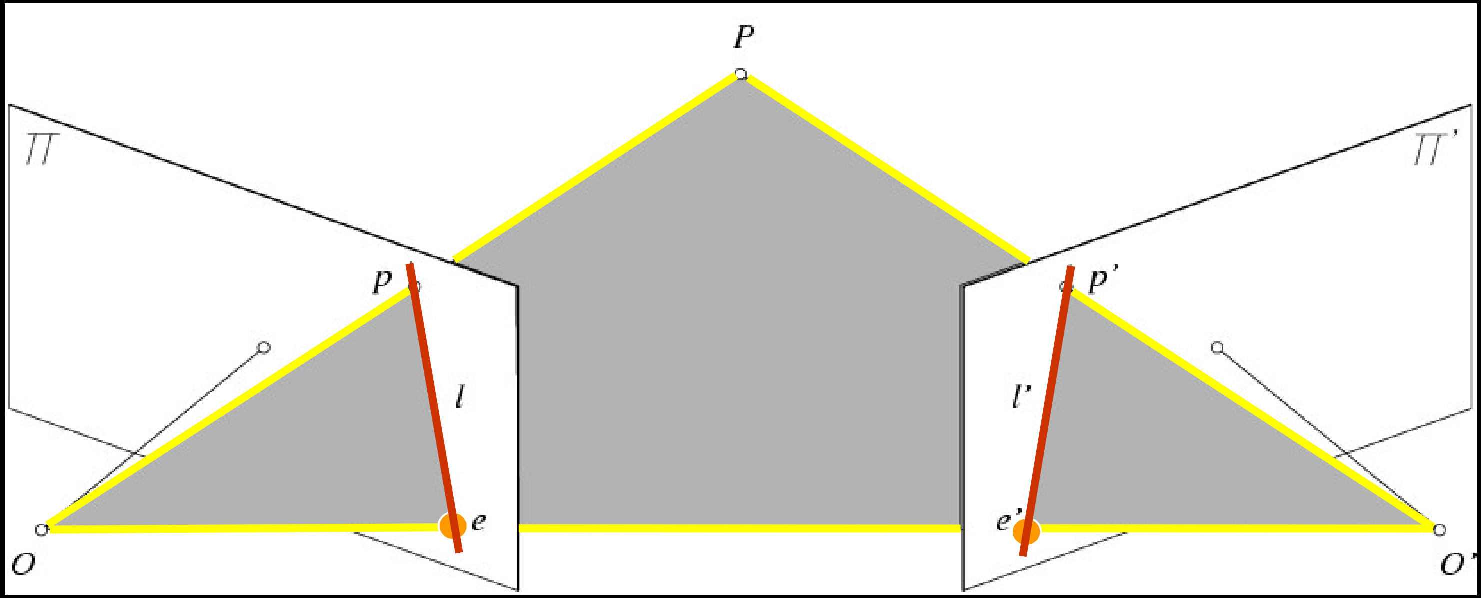
$$\boldsymbol{p}^T [\boldsymbol{v}_\times](\mathrm{Id} + \delta t\, [\boldsymbol{\omega}_\times])(\boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}}) = 0$$

$$\boldsymbol{p}^T([\boldsymbol{v}_\times][\boldsymbol{\omega}_\times])\boldsymbol{p} - (\boldsymbol{p} \times \dot{\boldsymbol{p}}) \cdot \boldsymbol{v} = 0$$

Case of pure translation:
Focus of expansion



Epipolar Constraint: Uncalibrated Case



$$\hat{\boldsymbol{p}}^T \mathcal{E} \hat{\boldsymbol{p}}' = 0$$

$$\boldsymbol{p} = \mathcal{K} \hat{\boldsymbol{p}}$$

$$\boldsymbol{p}' = \mathcal{K}' \hat{\boldsymbol{p}}'$$

$$\boldsymbol{p}^T \mathcal{F} \boldsymbol{p}' = 0 \quad \text{with} \quad \mathcal{F} = \mathcal{K}^{-T} \mathcal{E} \mathcal{K}'^{-1}$$

Fundamental Matrix
(Faugeras and Luong, 1992)

Properties of the Fundamental Matrix

- $F p'$ is the epipolar line associated with p'.
- $F^T p$ is the epipolar line associated with p.
- $F e' = 0$ and $F^T e = 0$.
- F is singular.

The Eight-Point Algorithm (Longuet-Higgins, 1981)

$$(u, v, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0$$

$$(uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

$$\begin{pmatrix} u_1u'_1 & u_1v'_1 & u_1 & v_1u'_1 & v_1v'_1 & v_1 & u'_1 & v'_1 \\ u_2u'_2 & u_2v'_2 & u_2 & v_2u'_2 & v_2v'_2 & v_2 & u'_2 & v'_2 \\ u_3u'_3 & u_3v'_3 & u_3 & v_3u'_3 & v_3v'_3 & v_3 & u'_3 & v'_3 \\ u_4u'_4 & u_4v'_4 & u_4 & v_4u'_4 & v_4v'_4 & v_4 & u'_4 & v'_4 \\ u_5u'_5 & u_5v'_5 & u_5 & v_5u'_5 & v_5v'_5 & v_5 & u'_5 & v'_5 \\ u_6u'_6 & u_6v'_6 & u_6 & v_6u'_6 & v_6v'_6 & v_6 & u'_6 & v'_6 \\ u_7u'_7 & u_7v'_7 & u_7 & v_7u'_7 & v_7v'_7 & v_7 & u'_7 & v'_7 \\ u_8u'_8 & u_8v'_8 & u_8 & v_8u'_8 & v_8v'_8 & v_8 & u'_8 & v'_8 \end{pmatrix}\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Minimize:

$$\sum_{i=1}^{8}(\boldsymbol{p}_i^T \mathcal{F} \boldsymbol{p}'_i)^2$$

under the constraint

$$|\mathcal{F}|^2 = 1.$$

Problem with eight-point algorithm

250906.36	183269.57	921.81	200931.10	146766.13	738.21	272.19	198.81
2692.28	131633.03	176.27	6196.73	302975.59	405.71	15.27	746.79
416374.23	871684.30	935.47	408110.89	854384.92	916.90	445.10	931.81
191183.60	171759.40	410.27	416435.62	374125.90	893.65	465.99	418.65
48988.86	30401.76	57.89	298604.57	185309.58	352.87	846.22	525.15
164786.04	546559.67	813.17	1998.37	6628.15	9.86	202.65	672.14
116407.01	2727.75	138.89	169941.27	3982.21	202.77	838.12	19.64
135384.58	75411.13	198.72	411350.03	229127.78	603.79	681.28	379.48

$$\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Poor numerical conditioning

Can be fixed by rescaling the data

The Normalized Eight-Point Algorithm (Hartley, 1995)

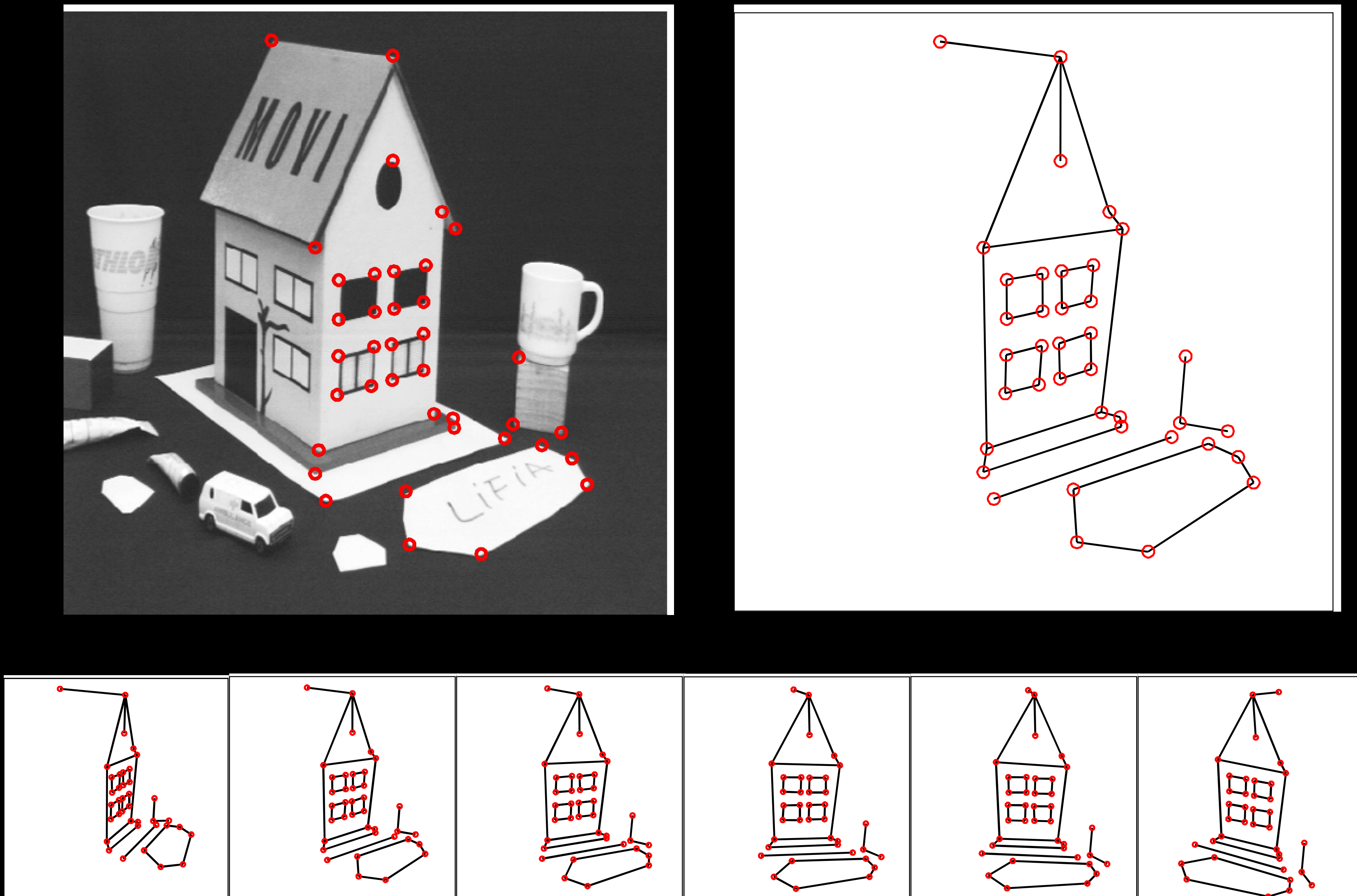
- Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels: $q_i = T p_i$, $q'_i = T' p'_i$
- Use the eight-point algorithm to compute F from the points q_i and q'_i .
- Enforce the rank-2 constraint.
- Output $T^T F T'$.

Non-Linear Least-Squares Approach (Luong et al., 1993)

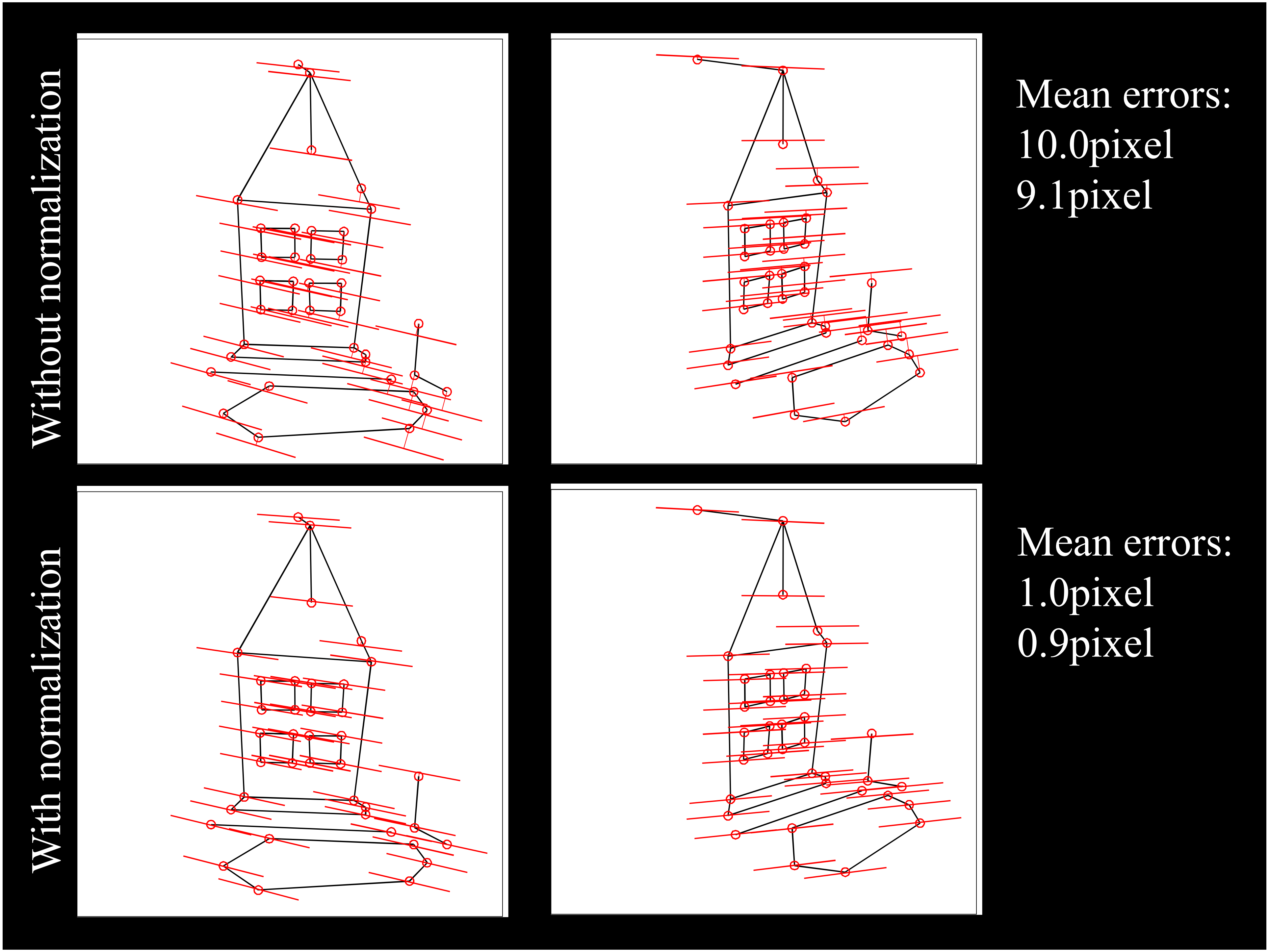
Minimize

$$\sum_{i=1}^{n} [\mathrm{d}^2(\boldsymbol{p}_i, \mathcal{F}\boldsymbol{p}'_i) + \mathrm{d}^2(\boldsymbol{p}'_i, \mathcal{F}^T\boldsymbol{p}_i)]$$

with respect to the coefficients of F , using an appropriate rank-2 parameterization.



Data courtesy of R. Mohr and B. Boufama.


Without normalization
Mean errors:
10.0pixel
9.1pixel
With normalization
Mean errors:
1.0pixel
0.9pixel