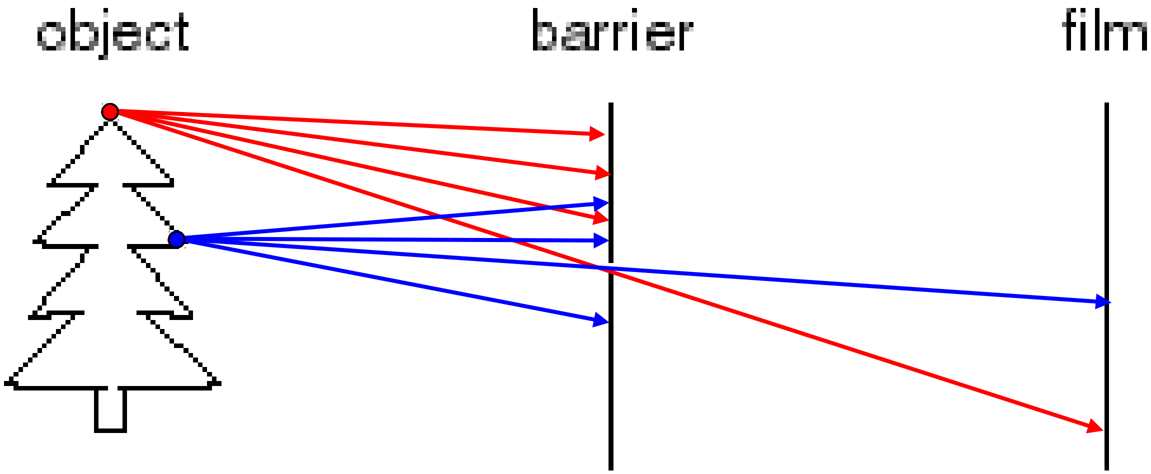
מבוא לראייה ממוחשבת – 22928
2016א

מנחה: אמיר אגוזי
egozi5@gmail.com

מפגש מס' 6 (ואחרון...)

Camera model

Pinhole camera

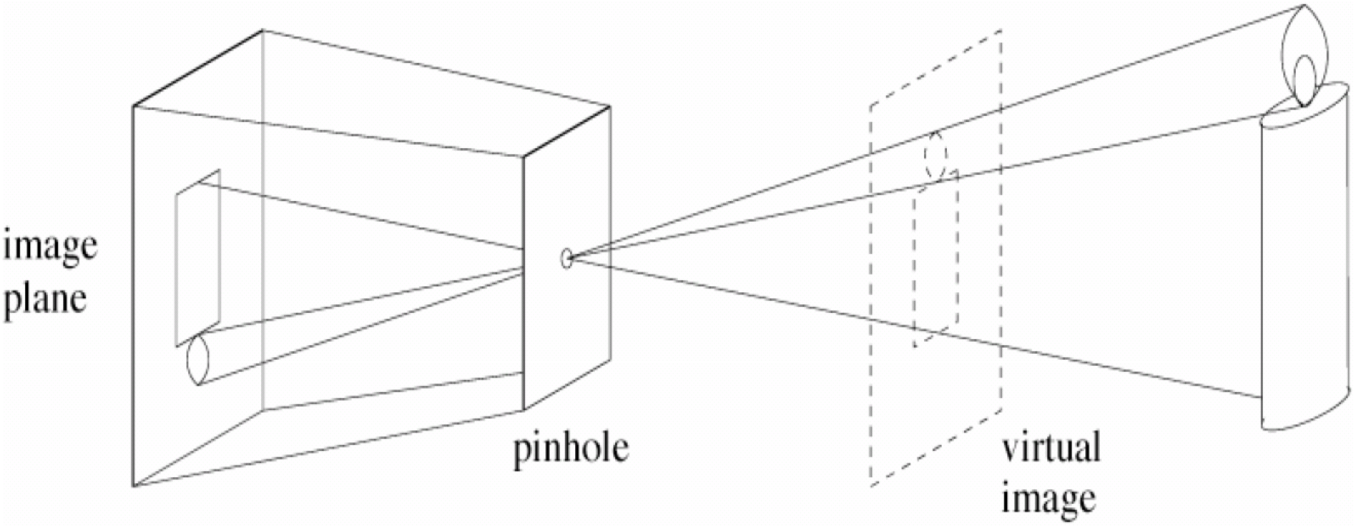


- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**
  - How does this transform the image?

Slide by Steve Seitz

Pinhole camera

Pinhole camera is a simple model to approximate imaging process, perspective **projection**.



If we treat pinhole as a point, only one ray from any given point can enter the camera.

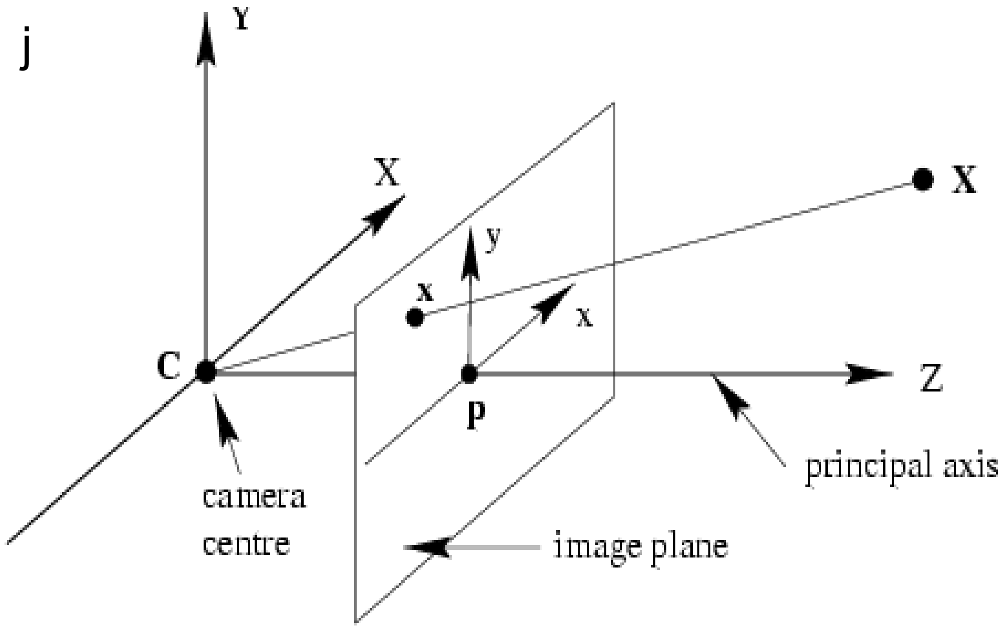
Fig from Forsyth and Ponce

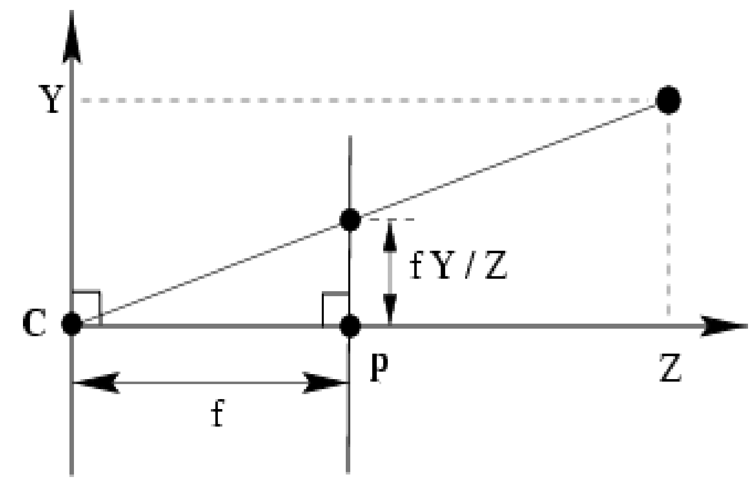
Perspective effects

Perspective effects

Pinhole camera model

j



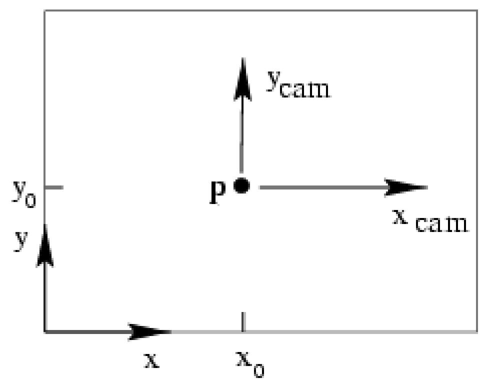


$$\begin{bmatrix} X & Y & Z \end{bmatrix}^T \mapsto \begin{bmatrix} fX/Z & fY/Z \end{bmatrix}^T$$

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Note – all coordinates are represented in camera reference frame

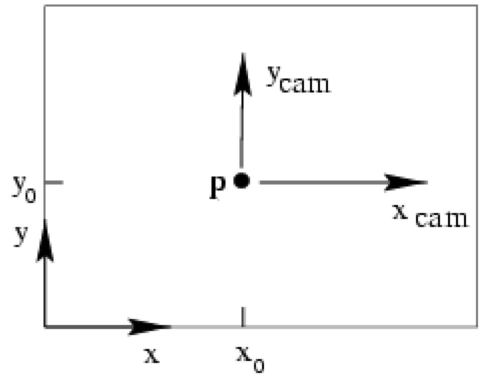
Principle point offset



$$\begin{bmatrix} X & Y & Z \end{bmatrix}^T \mapsto \begin{bmatrix} fX/Z + p_x & fY/Z + p_y \end{bmatrix}^T$$

$$\begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Principle point offset



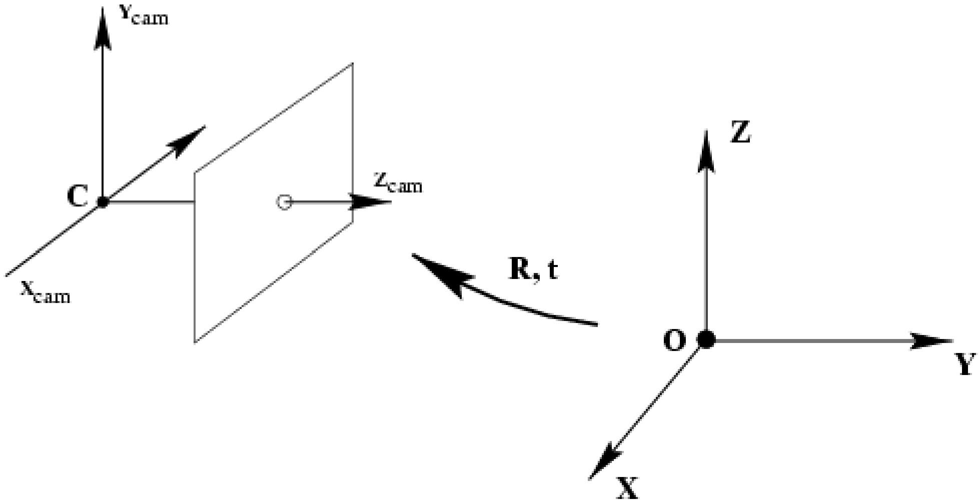
$$\mathbf{x} = K[\,I \mid \mathbf{0}]\mathbf{X}_{cam}$$

$$\begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}$$

Calibration matrix

Camera rotation and translation



- R - is a 3x3 rotation matrix
- C - is the camera center in the world coordinate system.

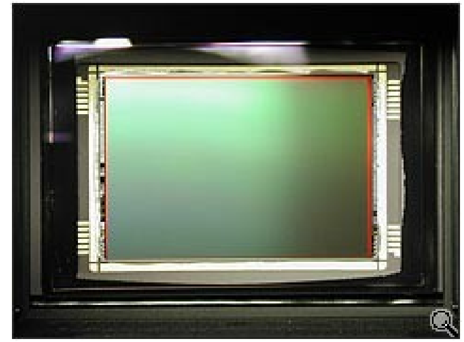
$$\mathbf{X}_{cam} = R(\mathbf{X} - C)$$

$$\mathbf{X}_{cam} = \begin{bmatrix} R & -RC \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & -RC \\ 0 & 1 \end{bmatrix} \mathbf{X}$$

$$\mathbf{x} = K[\,I \mid \mathbf{0}]\mathbf{X}_{cam} = K[\,R \mid -RC]\mathbf{X} = P\mathbf{X}$$

$$P = K[\,R \mid \mathbf{t}], \quad \mathbf{t} = -RC$$

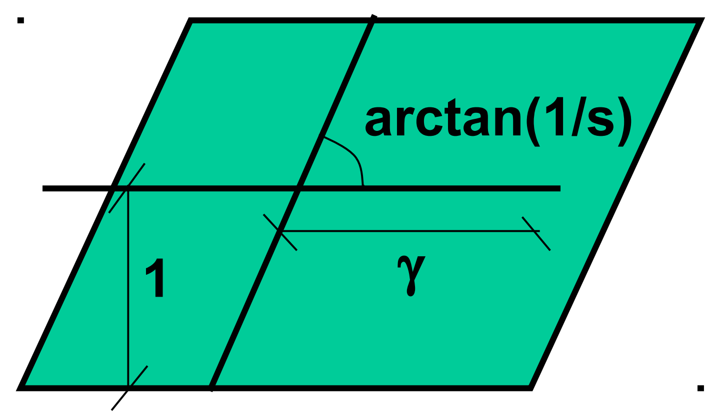
CCD camera



$$K = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

When skew angle is not zero:

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$



Note - After this transformation all K's entries are in pixel units.

Finite projective camera

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

$$P = K[\,R \mid \mathbf{t}], \quad \mathbf{t} = -RC$$

11 DOF (5 + 3 + 3)

תרגיל

- מצלמת חריר אידיאלית (ideal pinhole camera) בעלת מרחק מוקד של 7mm.
- גודל כל פיקסל הוא 0.03mm X 0.02mm.
- ומרכז הצילום נמצא ב- 550×650 כאשר הקורדינטות מתחילות מפינה שמאלית עליונה ב- (0,0).
- מהי מטריצת הקליברציה של המצלמה?

תרגיל - תשובה

- מצלמת חריר אידיאלית (ideal pinhole camera) בעלת מרחק מוקד של 7mm.
- גודל כל פיקסל הוא 0.02mm X 0.03mm.
- ומרכז הצילום נמצא ב-550×650 כאשר הקורדינטות מתחילות מפינה שמאלית עליונה ב- (0,0).
- מהי מטריצת הקליברציה של המצלמה?

<u>תשובה</u>:

$$K = \begin{bmatrix} f \cdot k_u & 0 & x_0 \\ 0 & f \cdot k_v & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 \cdot \frac{1}{0.02} & 0 & 550 \\ 0 & 7 \cdot \frac{1}{0.03} & 650 \\ 0 & 0 & 1 \end{bmatrix}$$

סטראו

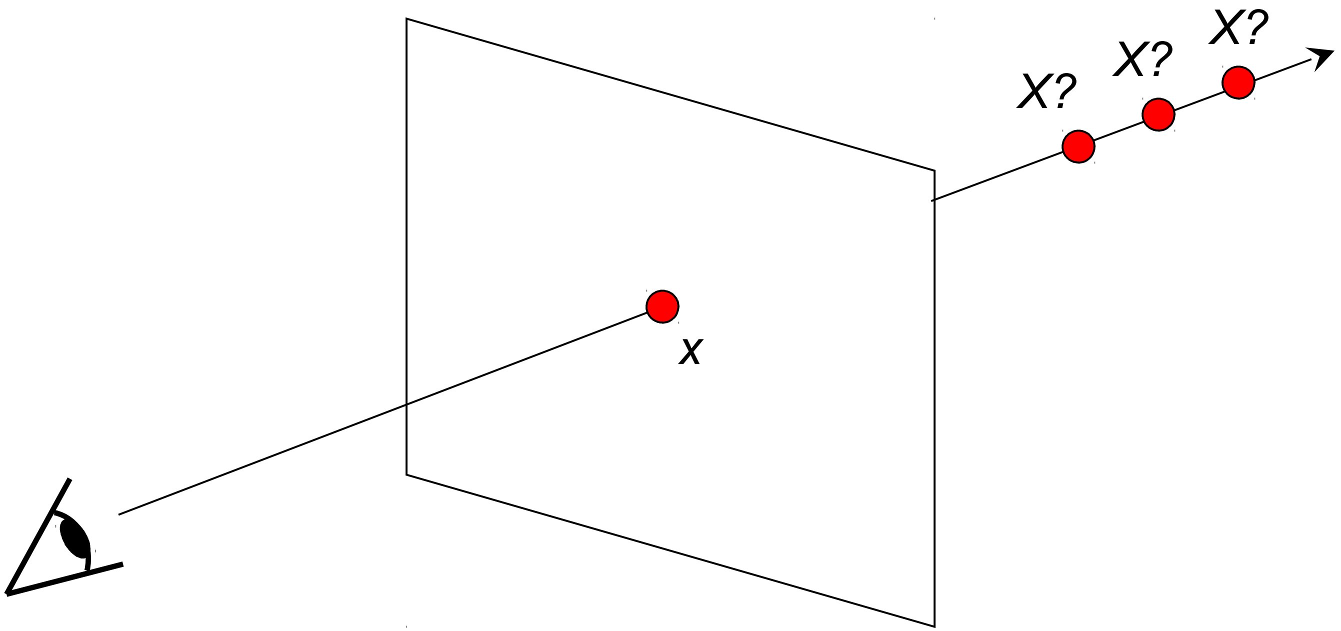
Geometric vision

- Goal: Recovery of 3D structure
  - What cues in the image allow us to do this?

Slide credit: Svetlana Lazebnik

Our Goal: Recovery of 3D Structure

- **We will focus on perspective and motion**
- **We need *multi-view geometry* because recovery of structure from one image is inherently ambiguous**



Slide credit: Svetlana Lazebnik

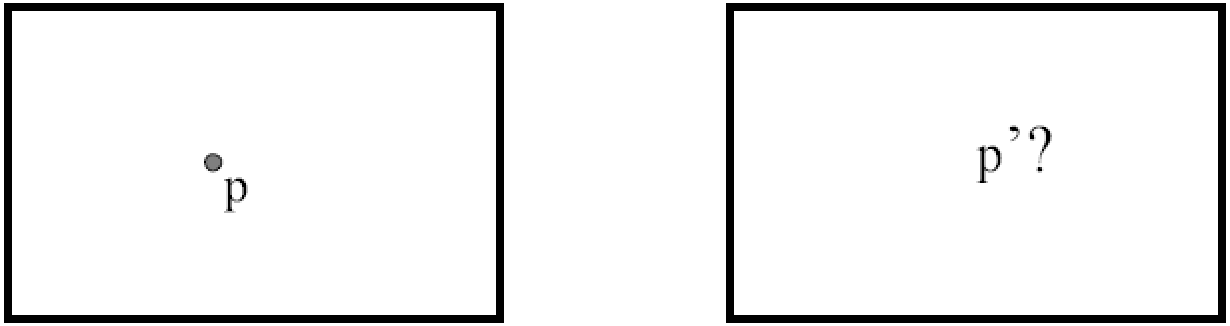
Three questions:

(i) **Correspondence geometry:** Given an image point x in the first view, how does this constrain the position of the corresponding point x' in the second image?

(ii) **Camera geometry (motion):** Given a set of corresponding image points $\{x_i \leftrightarrow x'_i\}$, i=1,…,n, what are the cameras P and P' for the two views?

(iii) **Scene geometry (structure):** Given corresponding image points $x_i \leftrightarrow x'_i$ and cameras P, P', what is the position of (their pre-image) X in space?

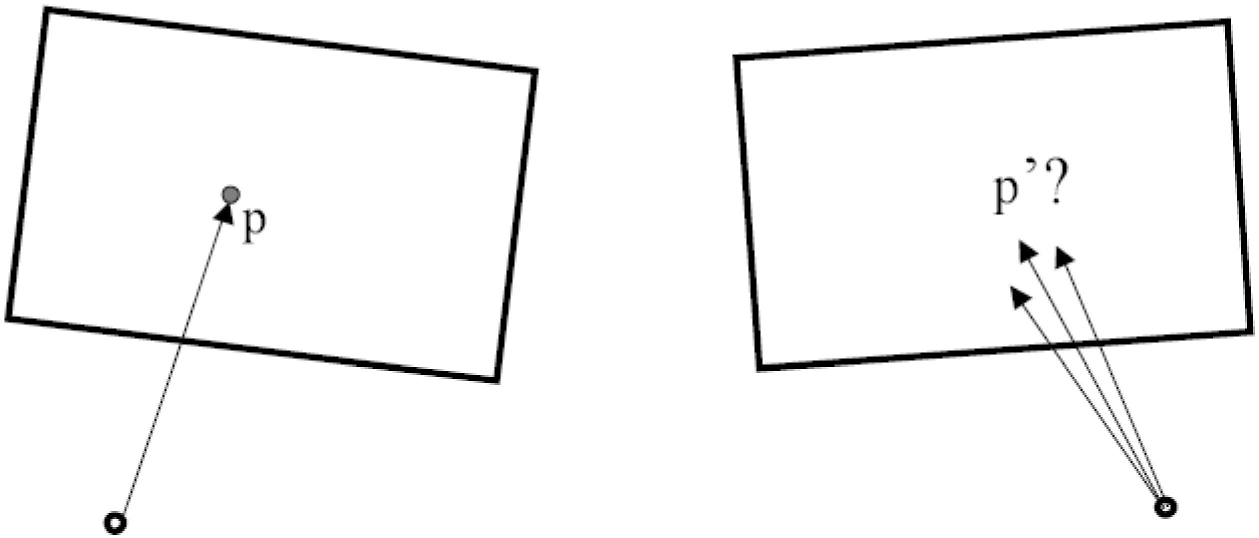
Stereo Correspondence Constraints



- **Given p in the left image, where can the corresponding point p' in the right image be?**

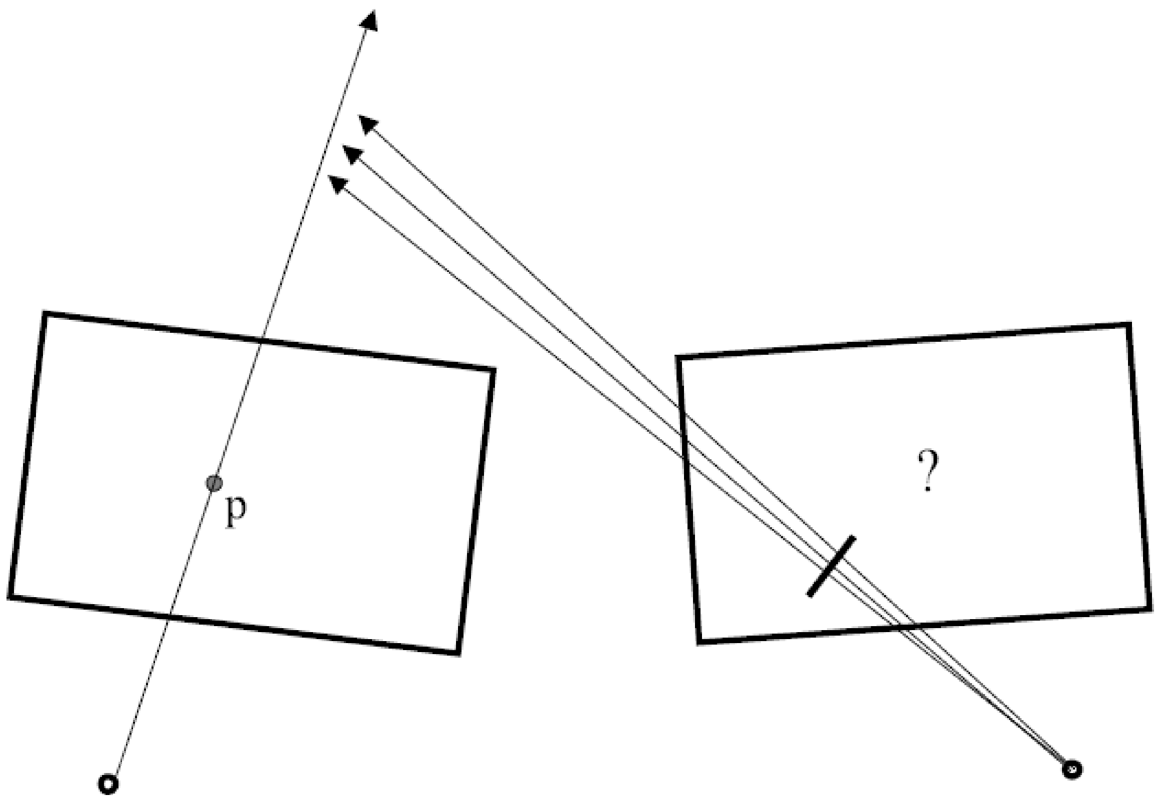
Slide credit: Kristen Grauman

Stereo Correspondence Constraints

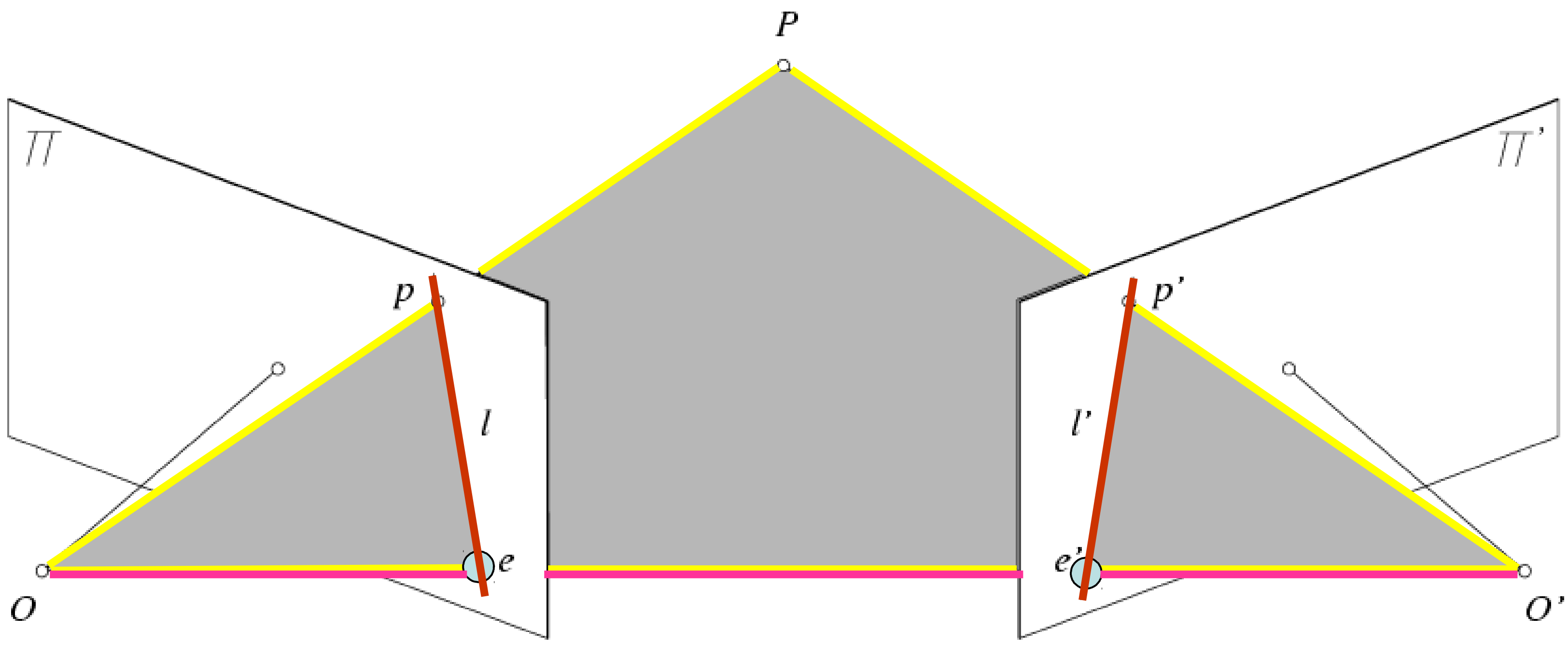


- **Given p in the left image, where can the corresponding point p' in the right image be?**

Stereo Correspondence Constraints



Epipolar Geometry



- Epipolar Plane
- Baseline
- Epipoles
- Epipolar Lines

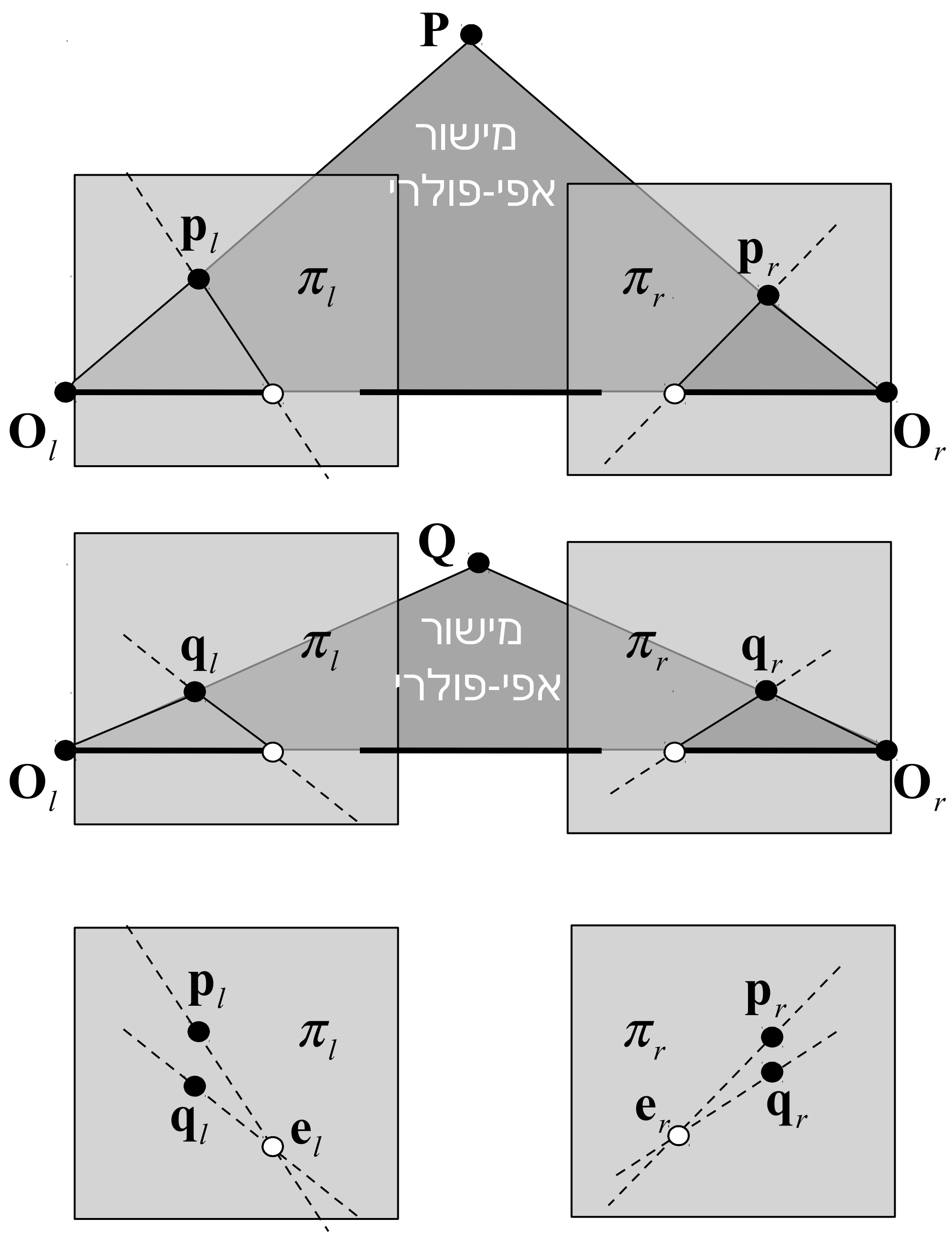
Slide adapted from Marc Pollefeys

Epipolar Geometry: Terms

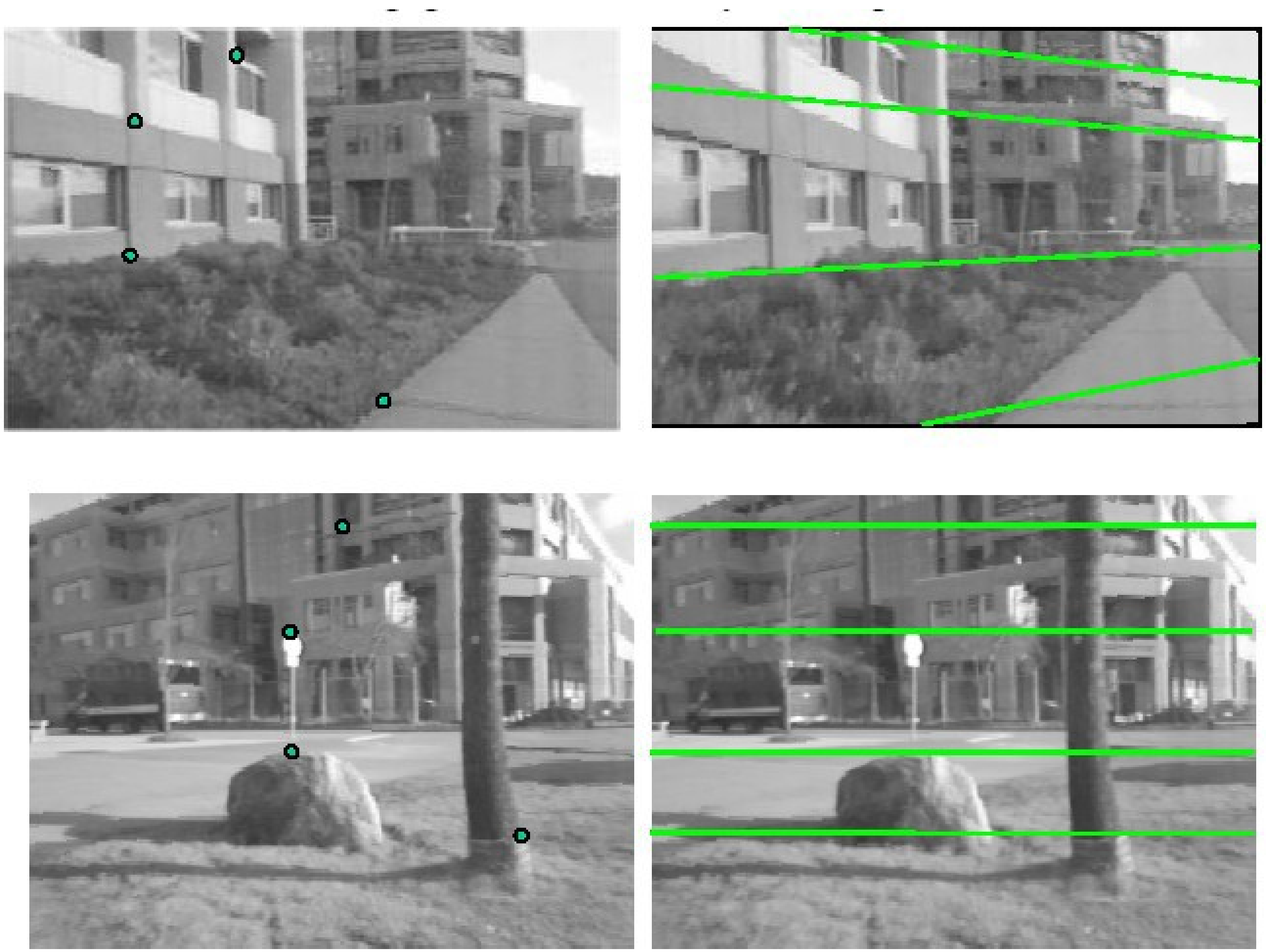
- ***Baseline*: line joining the camera centers**
- ***Epipole*: point of intersection of baseline with the image plane**
- ***Epipolar plane*: plane containing baseline and world point**
- ***Epipolar line*: intersection of epipolar plane with the image plane**

- All epipolar lines intersect at the epipole.
- An epipolar plane intersects the left and right image planes in epipolar lines.

Slide credit: Marc Pollefeys


P
מישור
אפי-פולרי
$\mathbf{p}_l$
π_l
π_r
$\mathbf{p}_r$
$\mathbf{O}_l$
$\mathbf{O}_r$
Q
מישור
אפי-פולרי
$\mathbf{q}_l$
π_l
π_r
$\mathbf{q}_r$
$\mathbf{O}_l$
$\mathbf{O}_r$
$\mathbf{p}_l$
π_l
$\mathbf{q}_l$
$\mathbf{e}_l$
π_r
$\mathbf{p}_r$
$\mathbf{q}_r$
$\mathbf{e}_r$

Example



Slide credit: Kristen Grauman

The fundamental matrix F

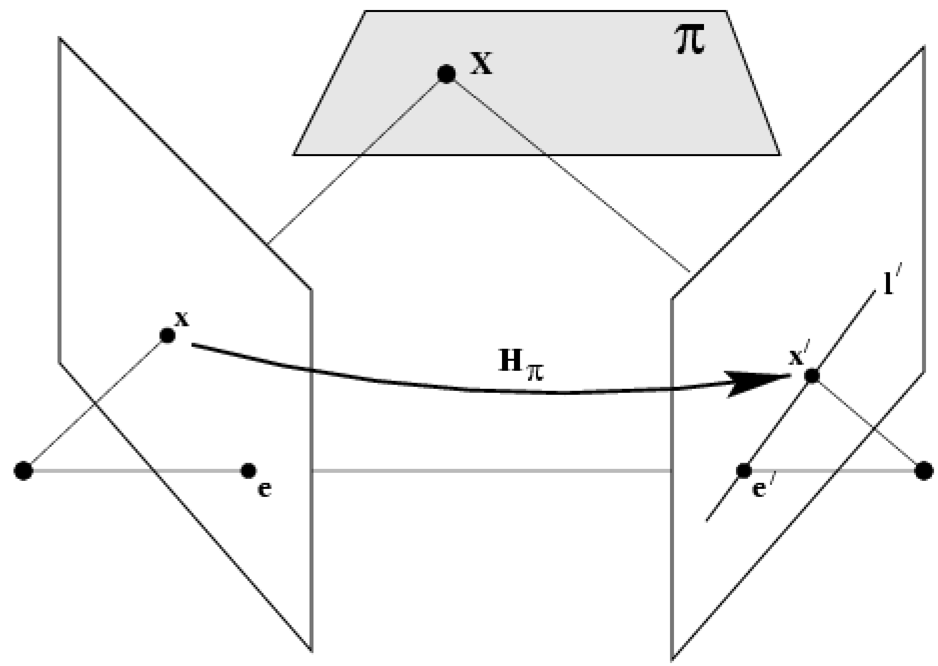
algebraic representation of epipolar geometry

$$\mathbf{x}^T \cdot \mathbf{l} = 0$$

we will see that mapping is (singular) correlation
(i.e. projective mapping from points to lines)
represented by the fundamental matrix ***F***

The fundamental matrix F

geometric derivation



$$x' = H_\pi x$$

$$l' = e' \times x' = [e']_\times H_\pi x = Fx$$

mapping from 2-D to 1-D family (rank 2)

The fundamental matrix F

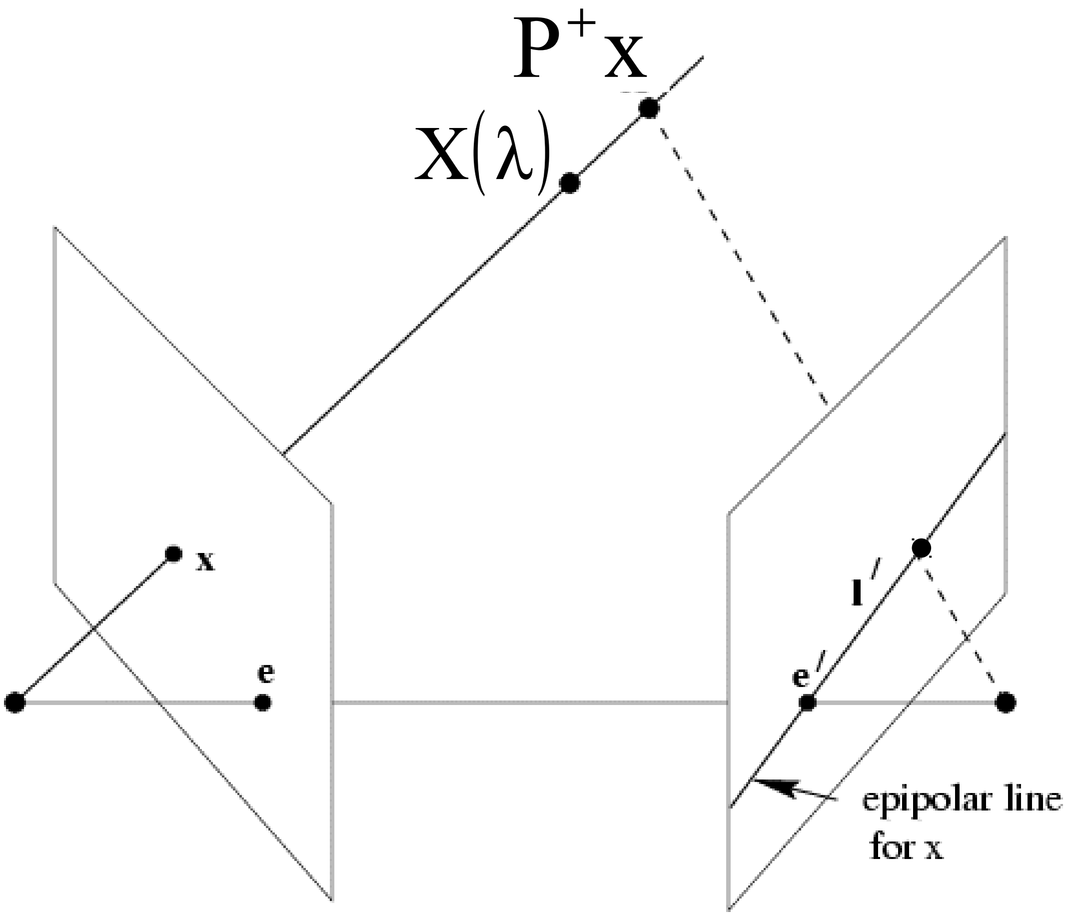
algebraic derivation

$$X(\lambda) = P^{+}x + \lambda C$$

$$l = P'C \times P'P^{+}x$$

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

$$\left(P^{+}P = I\right)$$



(note: doesn’t work for C=C’ $\Rightarrow$ F=0)

The fundamental matrix F

correspondence condition

The fundamental matrix satisfies the condition that for any pair of corresponding points x↔x' in the two images $$\mathbf{x}'^T F \mathbf{x} = 0 \qquad (x'^T \mathbf{l}' = 0)$$

The fundamental matrix F

F is the unique 3×3 rank 2 matrix that satisfies $\mathbf{x}'^T F \mathbf{x} = 0$ for all $\mathbf{x}' \leftrightarrow \mathbf{x}$

(i) **Transpose:** if F is fundamental matrix for (P,P'), then F^T is fundamental matrix for (P',P)
(ii) **Epipolar lines:** $l' = Fx$ & $l = F^T x'$
(iii) **Epipoles:** on all epipolar lines, thus $e'^T Fx = 0, \forall x$ $\Rightarrow e'^T F = 0$, similarly $Fe = 0$
(iv) **F** has 7 d.o.f. , i.e. 3x3-1(homogeneous)-1(rank2)
(v) **F** is a correlation, projective mapping from a point x to a line l'=Fx (not a proper correlation, i.e. not invertible)

The essential matrix

~fundamental matrix for calibrated cameras (remove K)

$$E = [t]_\times R = R[R^T t]_\times$$

$$\hat{x}'^T E \hat{x} = 0 \qquad \left(\hat{x} = K^{-1}x;\ \hat{x}' = K^{-1}x'\right)$$

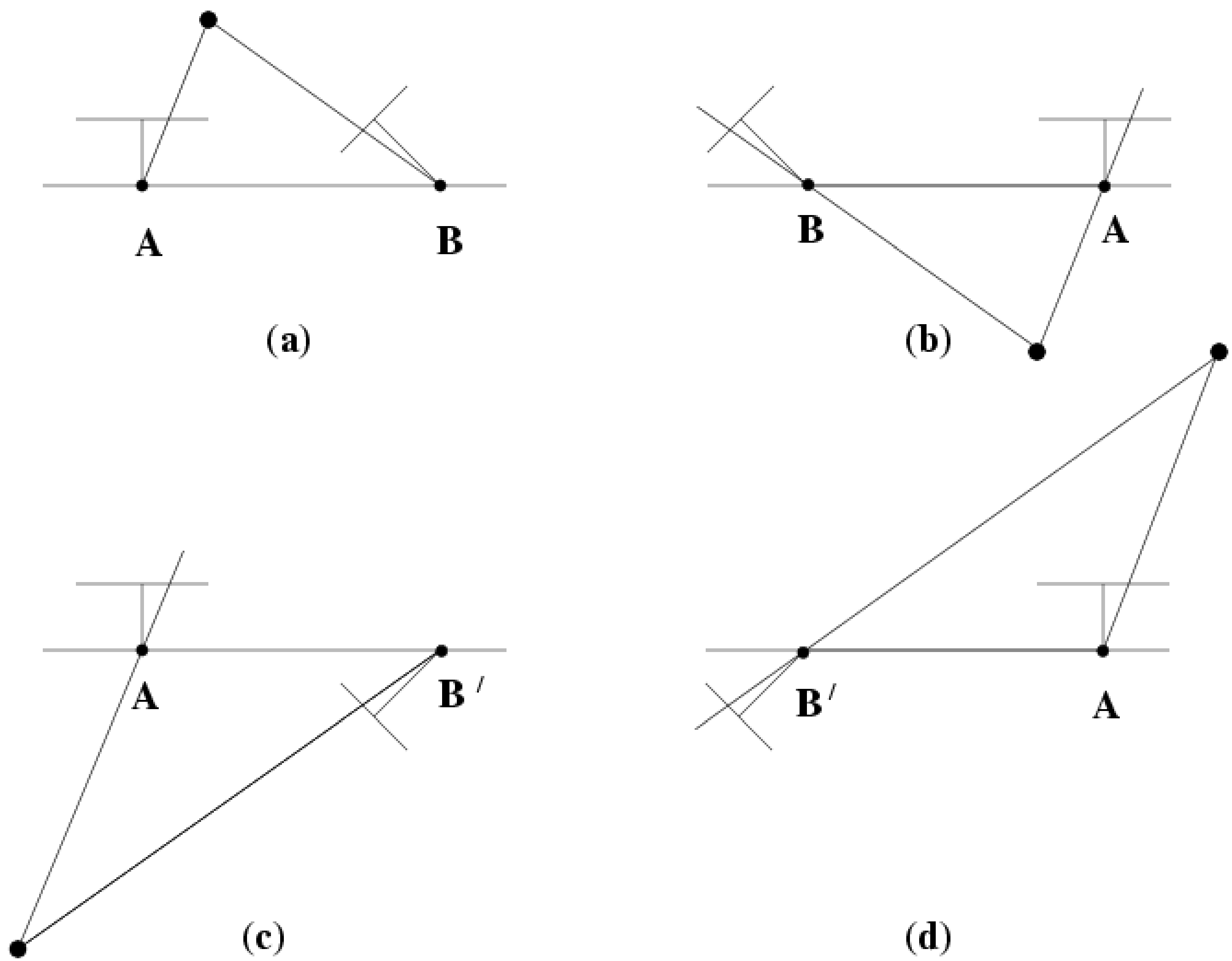
$$E = K'^T F K$$

5 d.o.f. (3 for R; 2 for t up to scale)

E is essential matrix if and only if
two singularvalues are equal (and third=0)

$$E = U\mathrm{diag}(1,1,0)V^T$$

Four possible reconstructions from E



(only one solution where points is in front of both cameras)

Fundamental matrix

- Relates pixel coordinates in the two views
- More general form than essential matrix: we remove need to know intrinsic parameters
- If we estimate fundamental matrix from correspondences in pixel coordinates, can reconstruct epipolar geometry without intrinsic or extrinsic parameters

Grauman

Epipolar geometry: basic equation

$$\mathrm{x'}^{\mathrm{T}}\mathrm{Fx} = 0$$

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

separate known from unknown

$$[x'x, x'y, x', y'x, y'y, y', x, y, 1][f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}]^{\mathrm{T}} = 0$$

(data) (unknowns)

(linear)

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathrm{f} = 0$$

$$\mathrm{Af} = 0$$

8-point algorithm

1. Solve a system of homogeneous linear equations
   a. Write down the system of equations
   b. Solve $\boldsymbol{f}$ from $A\boldsymbol{f}=0$ using SVD

Matlab:

```
[U, S, V] = svd(A);
f = V(:, end);
F = reshape(f, [3 3])';
```

the singularity constraint

$$e'^{T} F = 0 \qquad Fe = 0 \qquad \det F = 0 \qquad \operatorname{rank} F = 2$$

SVD from linearly computed F matrix (rank 3)

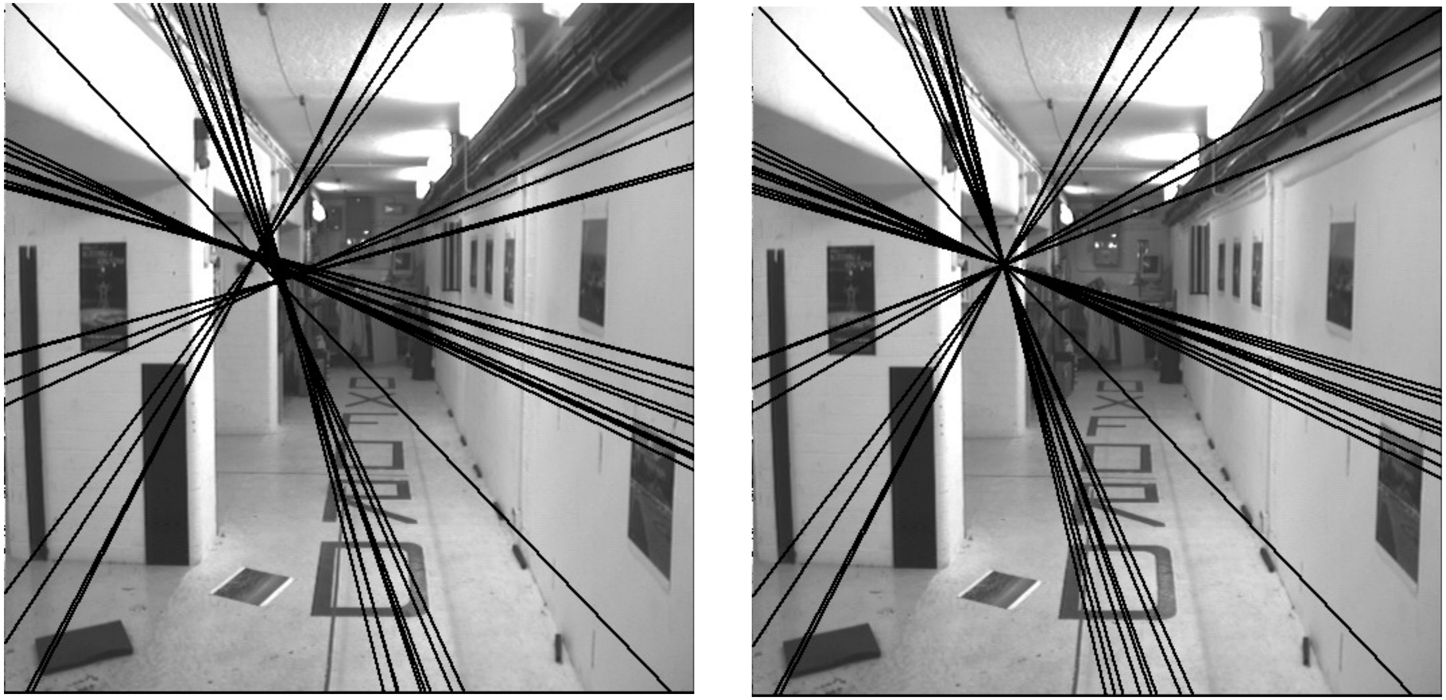
$$F = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T + U_3 \sigma_3 V_3^T$$

Compute closest rank-2 approximation $\min \| F - F' \|_F$

$$F' = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T$$

Need to enforce singularity constraint

Fundamental matrix has rank 2: $\det(F) = 0$



Left – uncorrected F – epipolar lines are not coincident

Right – epipolar lines from corrected F

8-point algorithm

1. Solve a system of homogeneous linear equations
   a. Write down the system of equations
   b. Solve $\boldsymbol{f}$ from $A\boldsymbol{f}=0$ using SVD

      Matlab:
      ```
      [U, S, V] = svd(A);
      f = V(:, end);
      F = reshape(f, [3 3])';
      ```

2. Resolve $\det(F) = 0$ constraint using SVD

      Matlab:
      ```
      [U, S, V] = svd(F);
      S(3,3) = 0;
      F = U*S*V';
      ```

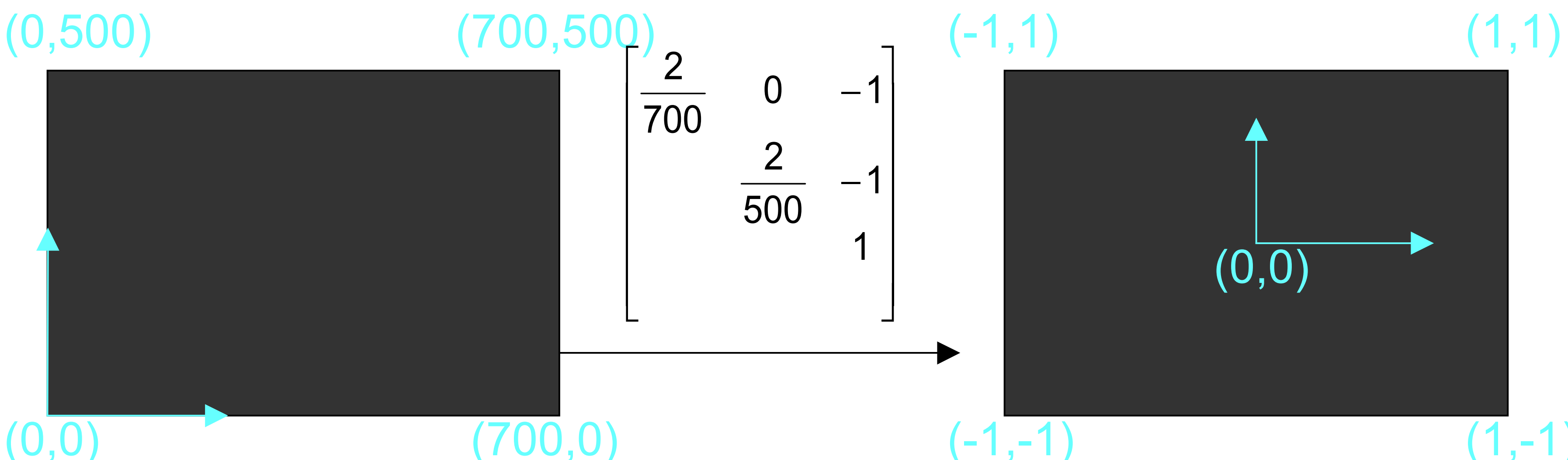
the NOT normalized 8-point algorithm

$$
\begin{array}{c}
\begin{bmatrix}
x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \\
x_2x_2' & y_2x_2' & x_2' & x_2y_2' & y_2y_2' & y_2' & x_2 & y_2 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_nx_n' & y_nx_n' & x_n' & x_ny_n' & y_ny_n' & y_n' & x_n & y_n & 1
\end{bmatrix} \\
\begin{matrix} \sim 10000 & \sim 10000 & \sim 100 & \sim 10000 & \sim 10000 & \sim 100 & \sim 100 & \sim 100 & 1 \end{matrix}
\end{array}
\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0
$$

Orders of magnitude difference
Between column of data matrix
$\rightarrow$ least-squares yields poor results

the normalized 8-point algorithm

Transform image to ~[-1,1]x[-1,1]



$$\begin{bmatrix} \frac{2}{700} & 0 & -1 \\ & \frac{2}{500} & -1 \\ & & 1 \end{bmatrix}$$

Least squares yields good results (Hartley, PAMI´97)

Recommendations

- Do not use unnormalized algorithms
- Quick and easy to implement: 8-point normalized
- Better: enforce rank-2 constraint during minimization
- Best: Maximum Likelihood Estimation (minimal parameterization, sparse implementation)

$$\sum_i d\left(\mathrm{x}_i, \hat{\mathrm{x}}_i\right)^2 + d\left(\mathrm{x}'_i, \hat{\mathrm{x}}'_i\right)^2 \quad \text{subject to } \hat{\mathrm{x}}'^{\mathrm{T}} \mathrm{F} \hat{\mathrm{x}} = 0$$

Initialize: normalized 8-point, (P,P‘) from F, reconstruct X_i

Stereo pipeline with weak calibration

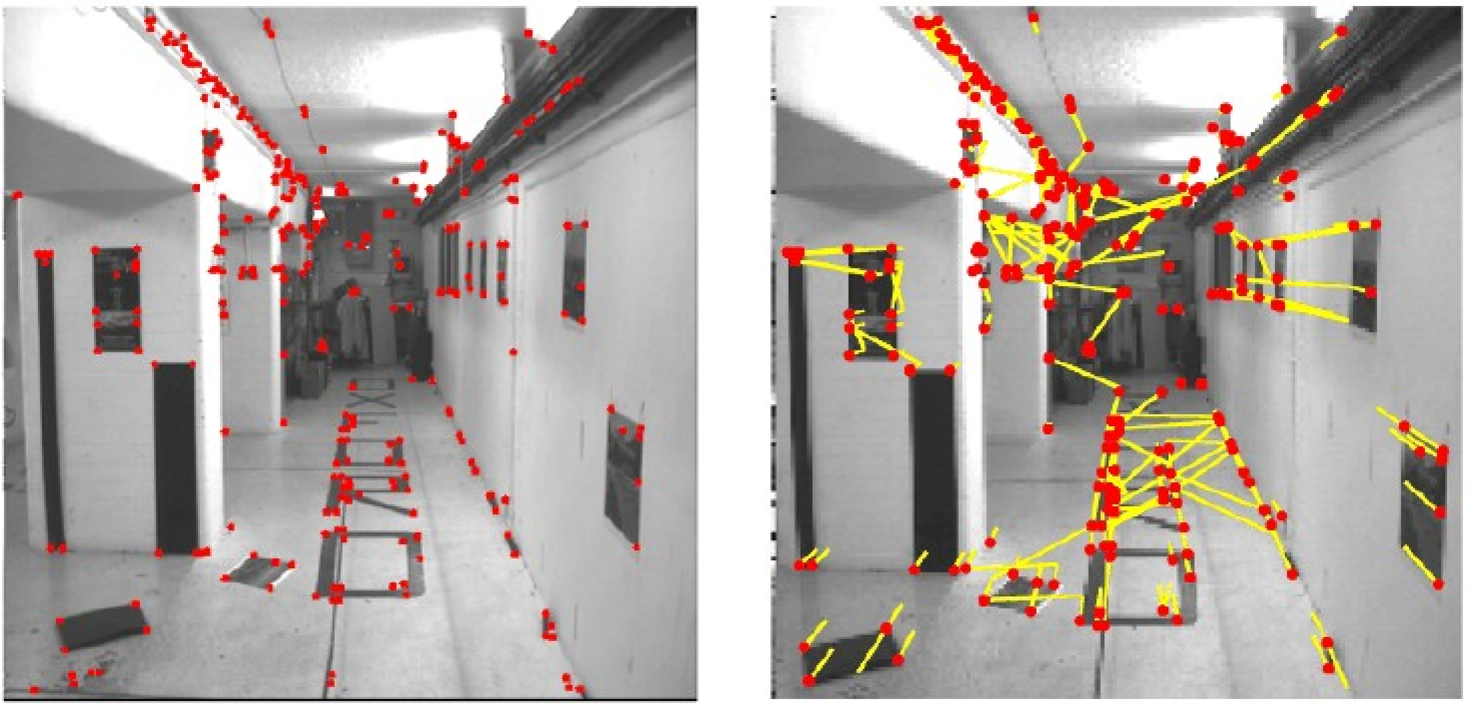
So, where to start with uncalibrated cameras?

Need to find fundamental matrix F **and** the correspondences (pairs of points (u',v') $\leftrightarrow$ (u,v)).

1) Find interest points in image
2) Compute correspondences
3) Compute epipolar geometry
4) Refine

Example from Andrew Zisserman

Putative matches based on correlation search



• Many wrong matches (10-50%), but enough to compute F

Grauman

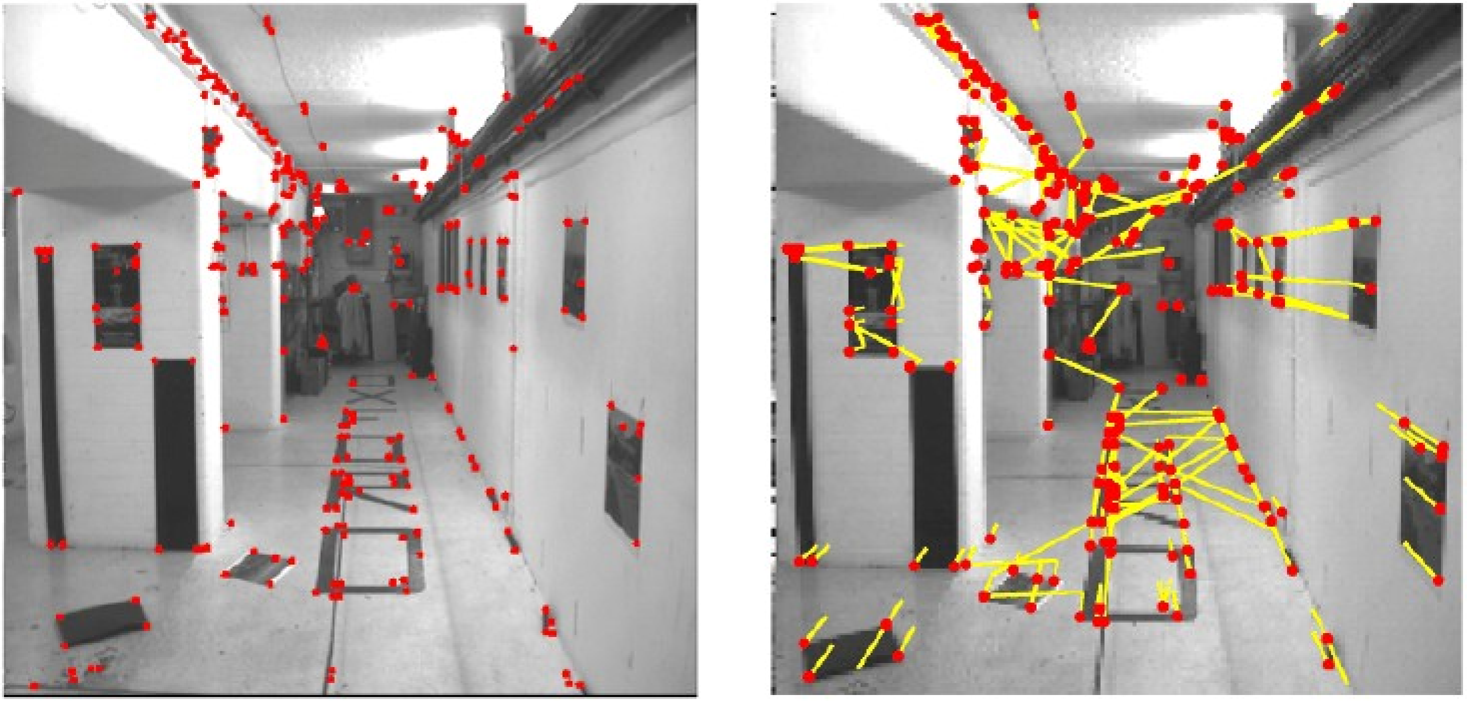
RANSAC for robust estimation of the fundamental matrix

- Select random sample of correspondences
- Compute F using them
  - This determines epipolar constraint
- Evaluate amount of support – inliers within threshold distance of epipolar line

- Choose F with most support (inliers)

Grauman

Putative matches based on correlation search

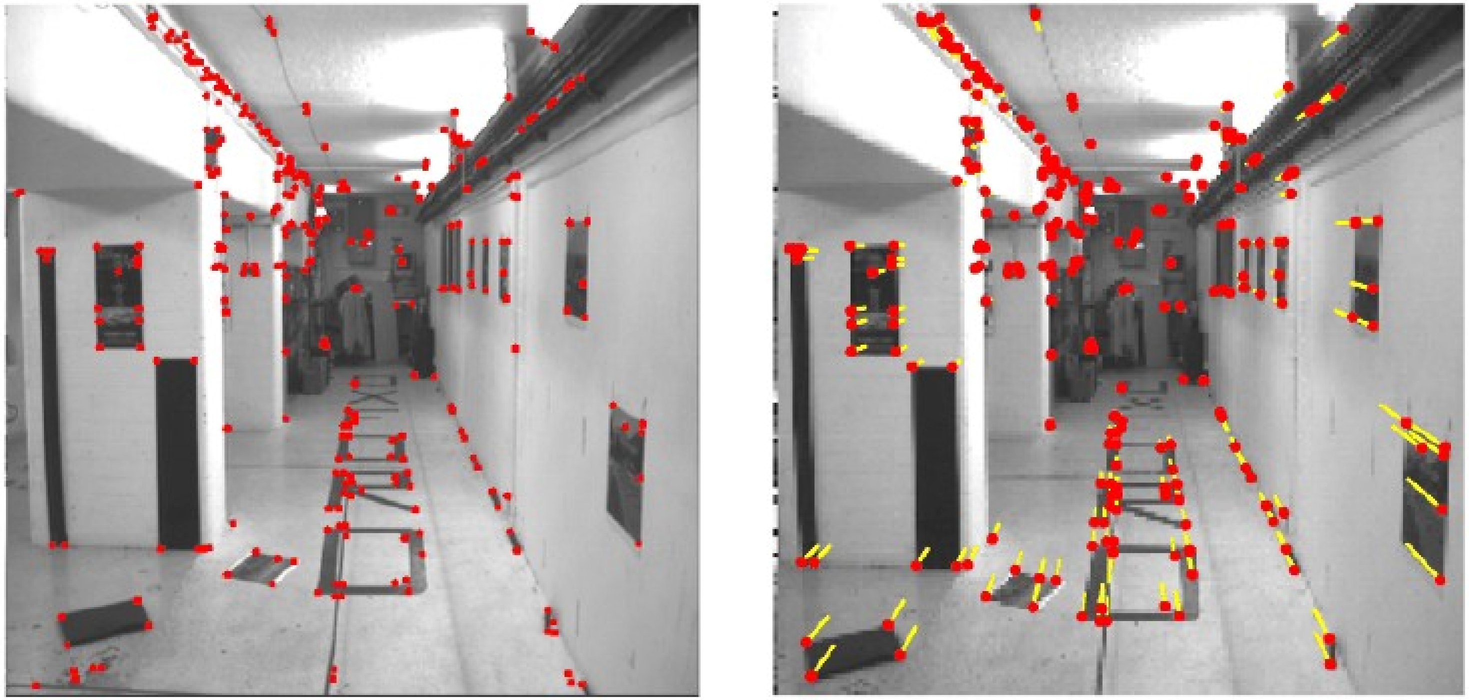


- Many wrong matches (10-50%), but enough to compute F

Grauman

Pruned matches

- Correspondences consistent with epipolar geometry



Grauman

- Resulting epipolar geometry

Grauman

Conclusion

- Camera model
- Stereo-vision
  - Epipolar geometry
  - Fundamental matrix
  - Essential matrix
  - (normalized) 8-pt algorithm.

This is it…

- Next semester – vision seminar
- Thesis??

בהצלחה