Epipolar geometry
Class 5

3D photography course schedule
(tentative)

	Lecture	Exercise
Sept 26	Introduction	-
Oct. 3	Geometry & Camera model	Camera calibration
Oct. 10	Single View Metrology	Measuring in images
Oct. 17	Feature Tracking/matching (Friedrich Fraundorfer)	Correspondence computation
Oct. 24	Epipolar Geometry	F-matrix computation
Oct. 31	Shape-from-Silhouettes (Li Guan)	Visual-hull computation
Nov. 7	Stereo matching	Project proposals
Nov. 14	Structured light and active range sensing	Papers
Nov. 21	Structure from motion	Papers
Nov. 28	Multi-view geometry and self-calibration	Papers
Dec. 5	Shape-from-X	Papers
Dec. 12	3D modeling and registration	Papers
Dec. 19	Appearance modeling and image-based rendering	Final project presentations

Optical flow

- Brightness constancy assumption

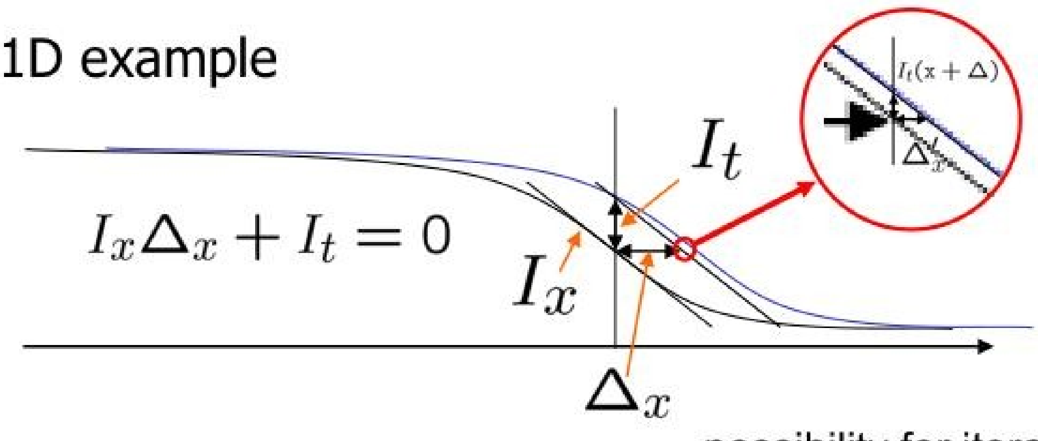
$$I(x + \Delta_x, y + \Delta_y, t + 1) = I(x, y, t)$$

$$I(x+u, y+v, t+1) = I(x, y, t) + \mathtt{I}_x \Delta_x + \mathtt{I}_y \Delta_y + \mathtt{I}_t \quad \text{(small motion)}$$

$$I_x \Delta_x + I_y \Delta_y + I_t = 0$$

- 1D example

$$I_x \Delta_x + I_t = 0$$



possibility for iterative refinement

Optical flow

- Brightness constancy assumption

$$I(x+\Delta_x, y+\Delta_y, t+1) = I(x,y,t)$$

$$I(x+u, y+v, t+1) = I(x,y,t) + \mathrm{I}_x\Delta_x + \mathrm{I}_y\Delta_y + \mathrm{I}_t \quad \text{(small motion)}$$

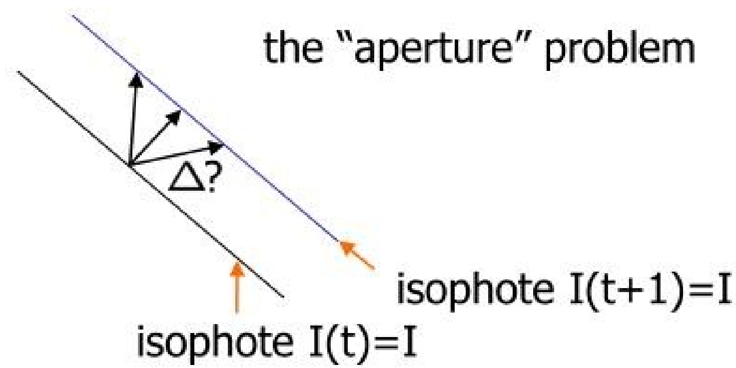
$$I_x\Delta_x + I_y\Delta_y + I_t = 0$$

- 2D example

$$I_x\Delta_x + I_y\Delta_y + I_t = 0$$

(1 constraint)

Δ_x, Δ_y (2 unknowns)



Optical flow

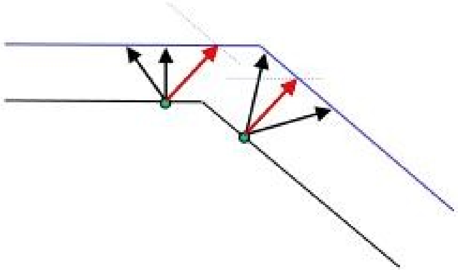
- How to deal with aperture problem?

$$R_x\Delta_x + R_y\Delta_y + R_t = 0 \quad G_x\Delta_x + G_y\Delta_y + G_t = 0 \quad B_x\Delta_x + B_y\Delta_y + B_t = 0$$

(3 constraints if color gradients are different)

Assume neighbors have same displacement

$$I_x(\mathrm{x})\Delta_x + I_y(\mathrm{x})\Delta_y + I_t(\mathrm{x}) = 0 \quad I_x(\mathrm{x}')\Delta_x + I_y(\mathrm{x}')\Delta_y + I_t(\mathrm{x}') = 0 \quad \cdots$$



Lucas-Kanade

Assume neighbors have same displacement

least-squares:

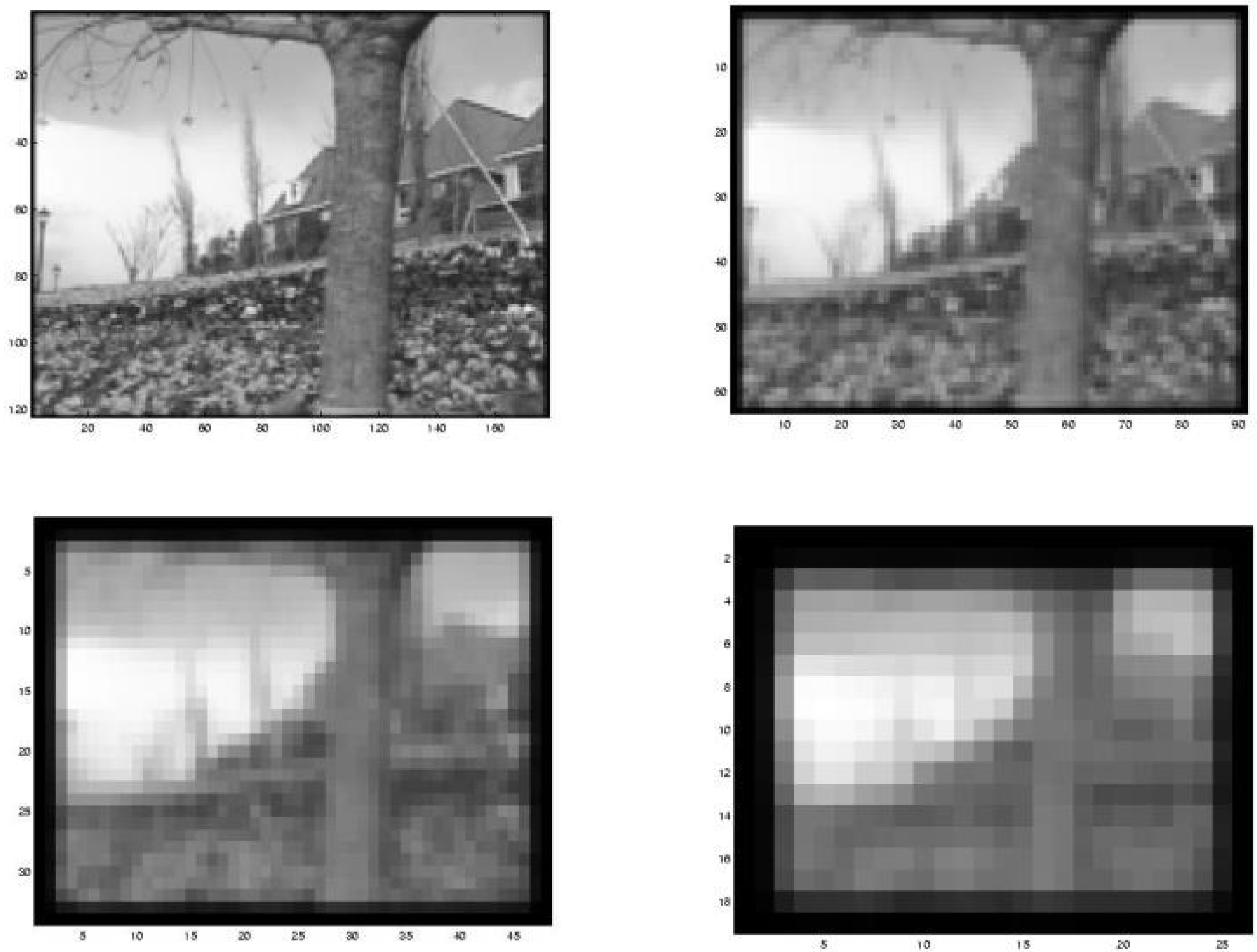
$$\begin{bmatrix} I_x(\mathrm{x}) & I_y(\mathrm{x}) \\ I_x(\mathrm{x}) & I_y(\mathrm{x}) \\ I_x(\mathrm{x}) & I_y(\mathrm{x}) \end{bmatrix} \Delta = \begin{bmatrix} -I_t(\mathrm{x}) \\ -I_t(\mathrm{x}') \\ -I_t(\mathrm{x}'') \end{bmatrix} \qquad \mathbf{A}\Delta = \mathbf{b}$$

Revisiting the small motion assumption

- Is this motion small enough?
  - Probably not—it's much larger than one pixel (2^{nd} order terms dominate)
  - How might we solve this problem?

* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

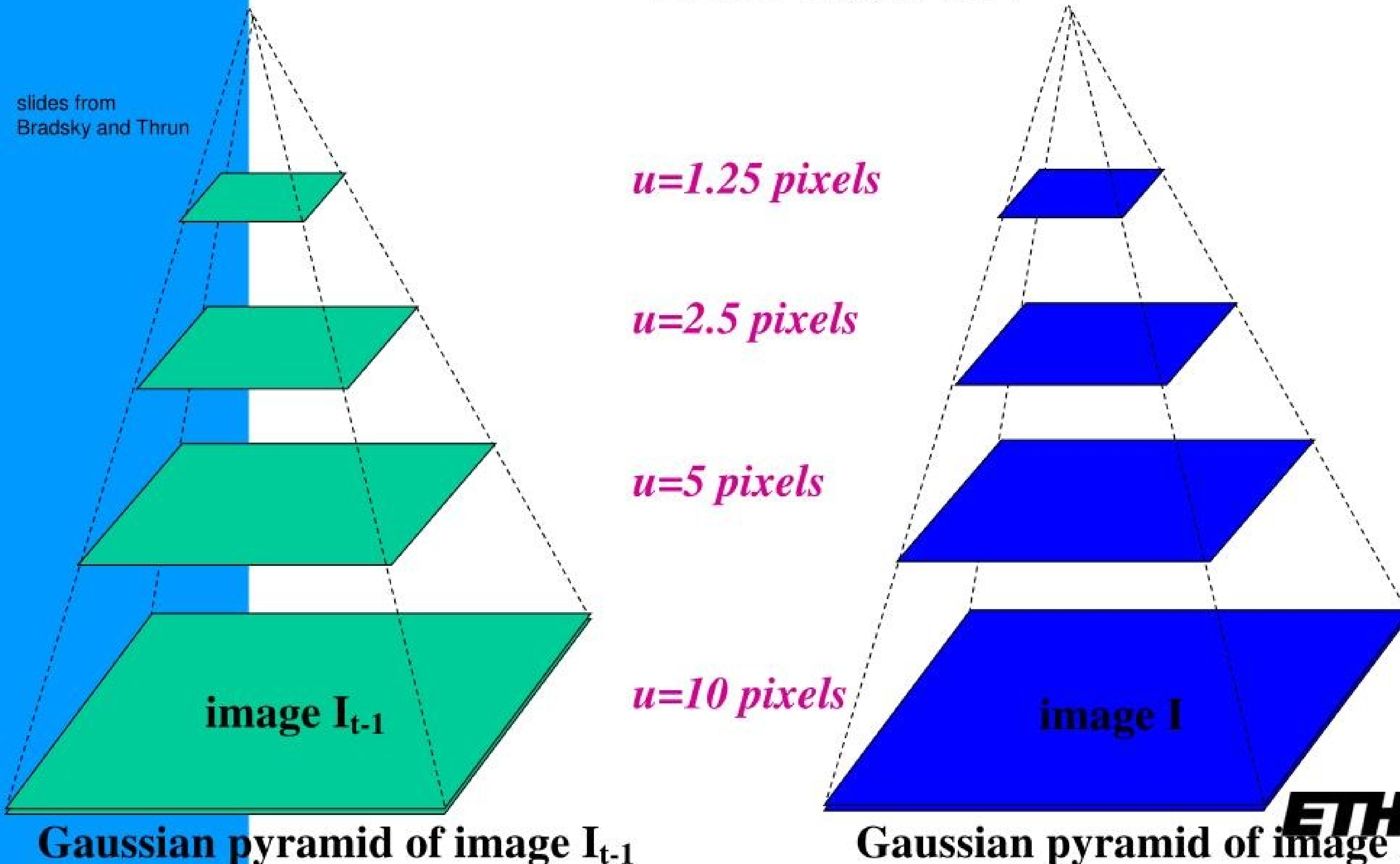
Reduce the resolution!



* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

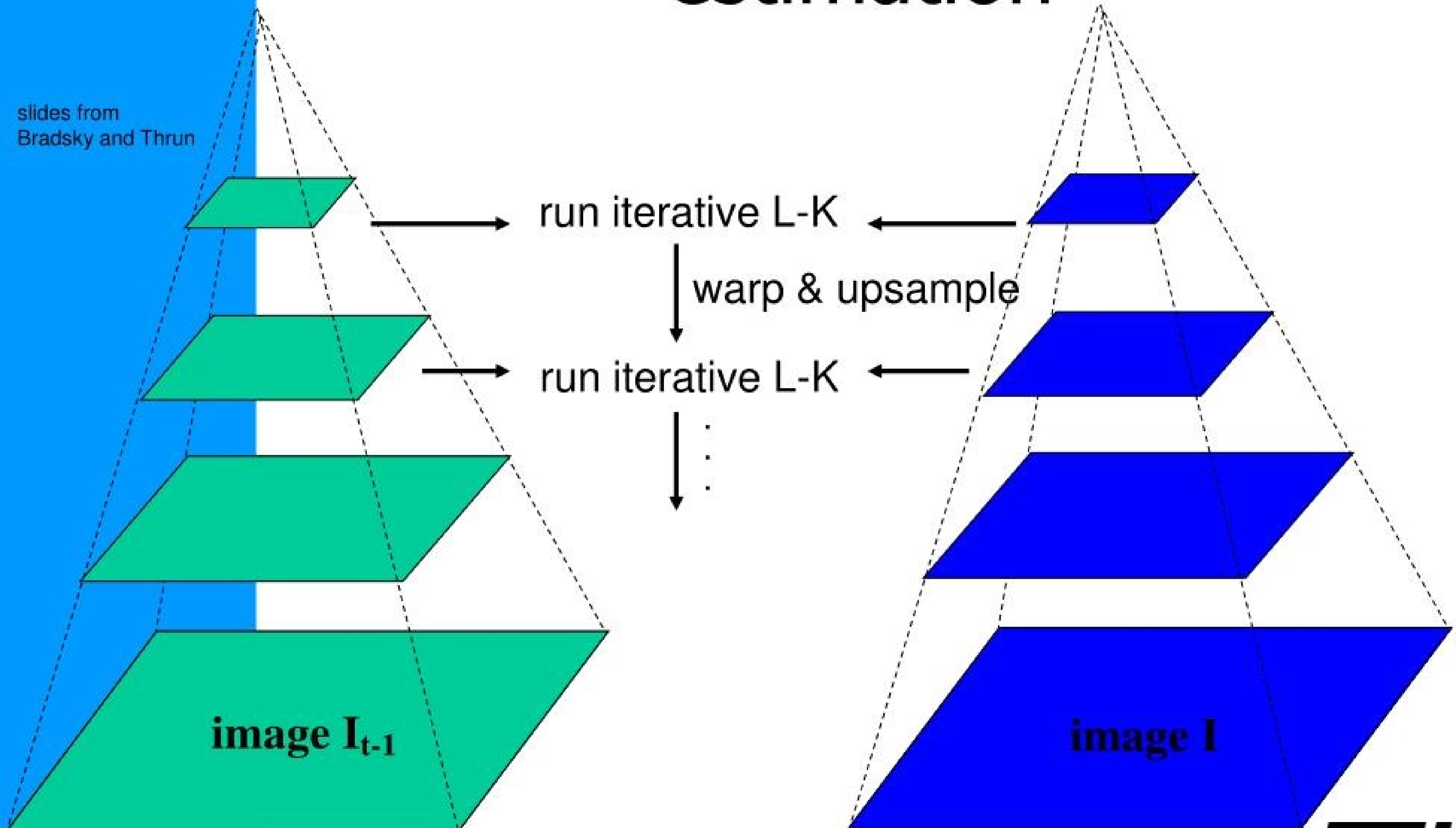
Coarse-to-fine optical flow estimation

slides from
Bradsky and Thrun



Coarse-to-fine optical flow estimation

slides from
Bradsky and Thrun



Gaussian pyramid of image I_{t-1}

Gaussian pyramid of image I

Feature tracking

- Identify features and track them over video
  - Small difference between frames
  - potential large difference overall
- Standard approach:
  KLT (Kanade-Lukas-Tomasi)

ETH

Good features to track

- Use same window in feature selection as for tracking itself

$$\text{with } \mathbf{M} = \iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} w(x,y) \mathrm{d}x\mathrm{d}y$$

- Compute motion assuming it is small

$$min \iint_W \left(I + \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \Delta - J\right)^2 w(x,y) \mathrm{d}x\mathrm{d}y$$

$$\text{differentiate: } \iint_W 2 \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \left(I + \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \Delta - J\right) w(x,y) \mathrm{d}x\mathrm{d}y$$

$$\iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} w(x,y) \mathrm{d}x\mathrm{d}y \Delta = \iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} (J - I) w(x,y) \mathrm{d}x\mathrm{d}y$$

Affine is also possible, but a bit harder (6x6 in stead of 2x2)

Example

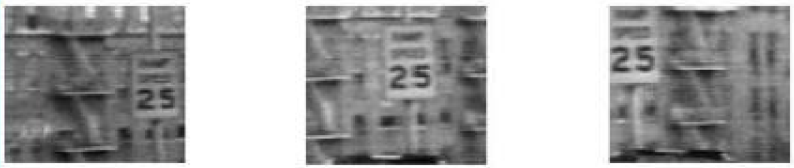
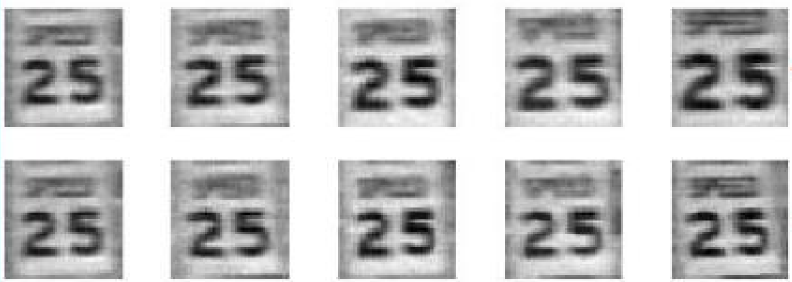
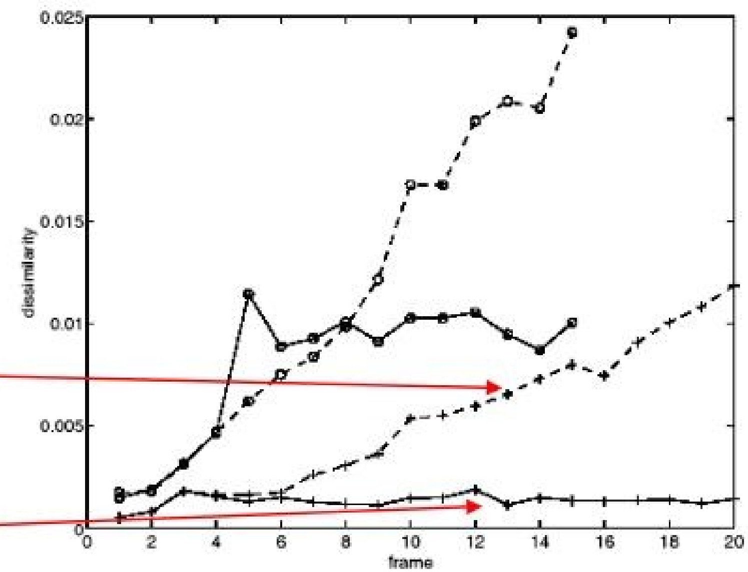


Figure 1: Three frame details from Woody Allen's *Manhattan*. The details are from the 1st, 11th, and 21st frames of a subsequence from the movie.





Simple displacement is sufficient between consecutive frames, but not to compare to reference template

Example

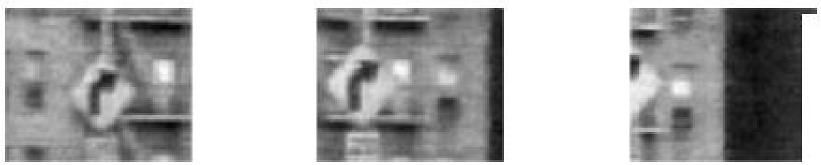
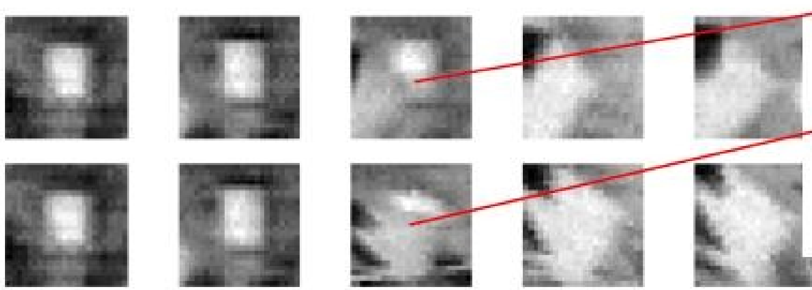
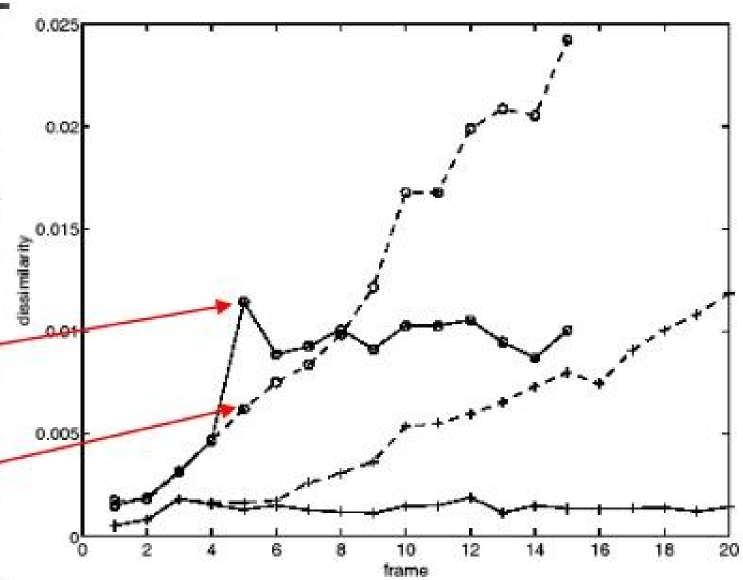
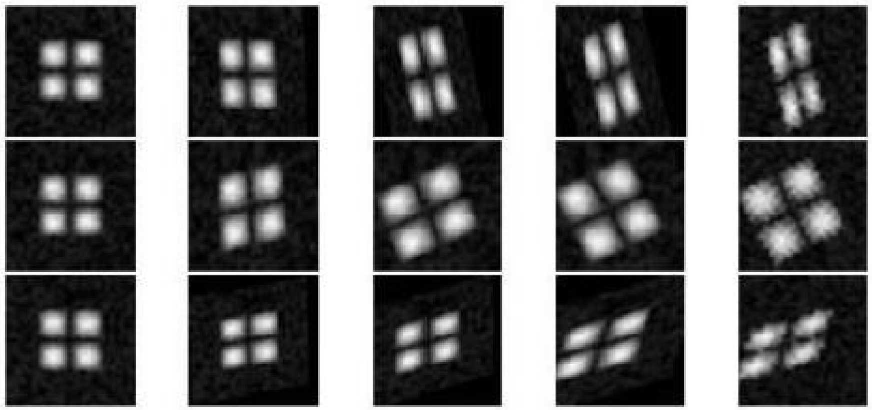


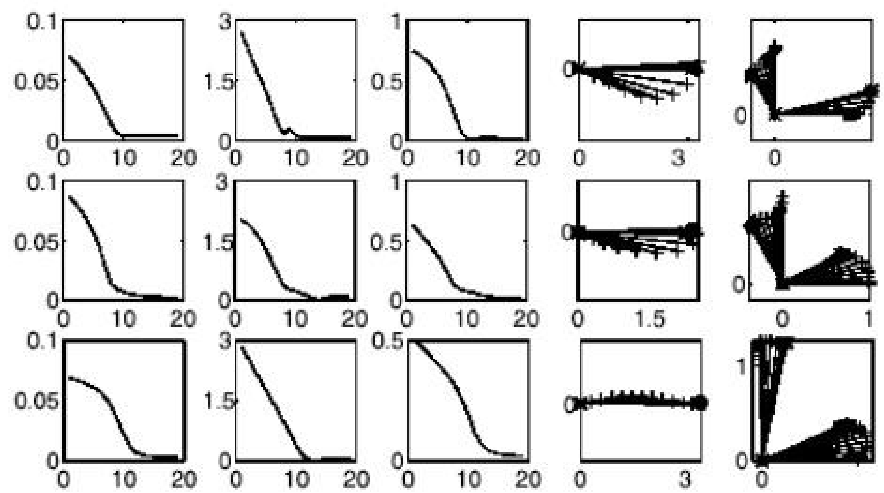
Figure 4: Three more frame details from *Manhattan*
The feature tracked is the bright window on the back
ground, on the right of the traffic sign.



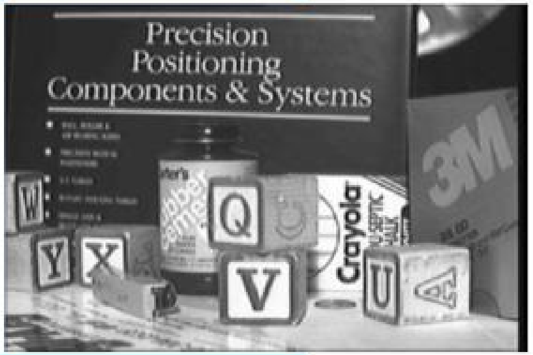


Synthetic example

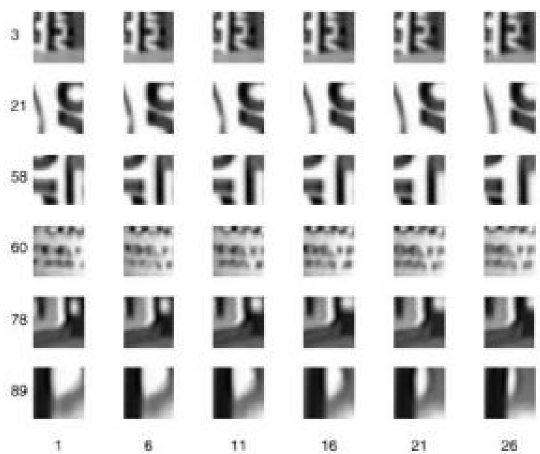


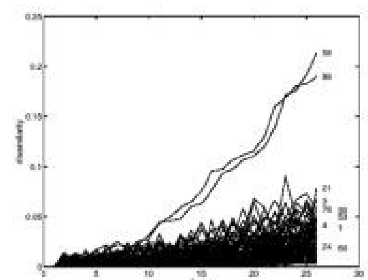


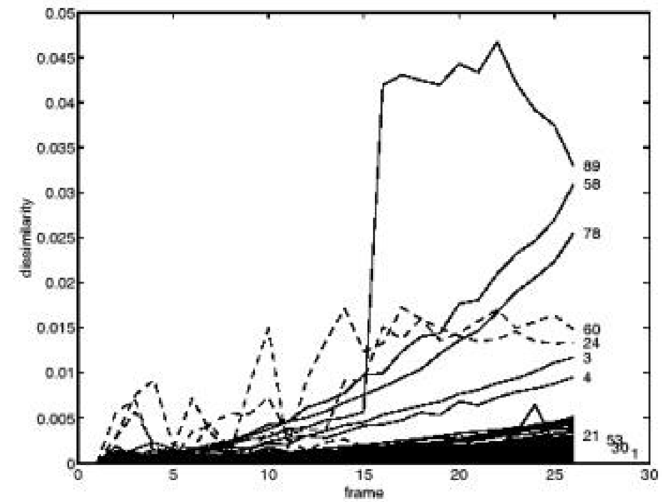
Good features to keep tracking



Perform affine alignment between first and last frame
Stop tracking features with too large errors








gpu_vis
1024 x 768 video, Time: 28.243 msec, Features: [MAX 1000] (Tracked 343 of 344) (Added 0)
STOP


gpu_vis
640 x 480 video, Time: 20.532 msec, Features: [MAX 1000] (Tracked 308 of 327) (Added 0)

Two-view geometry

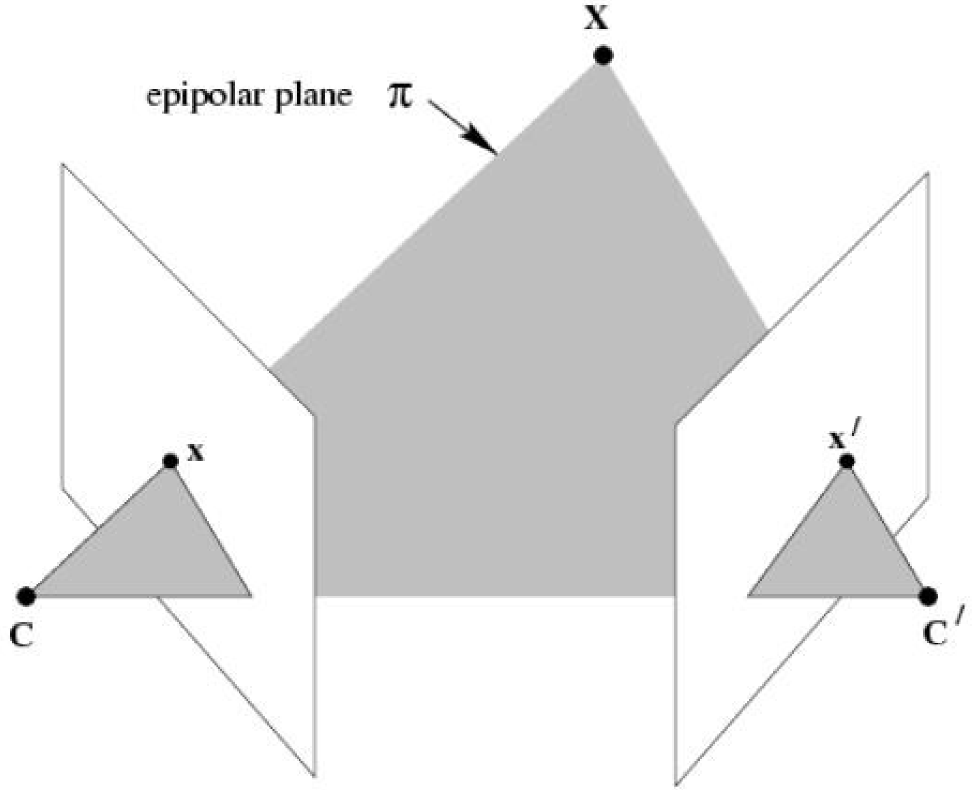
Three questions:

(i) **Correspondence geometry:** Given an image point x in the first image, 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 $\{x_i \leftrightarrow x'_i\}$, $i=1,\ldots,n$, what are the cameras P and P' for the two views?

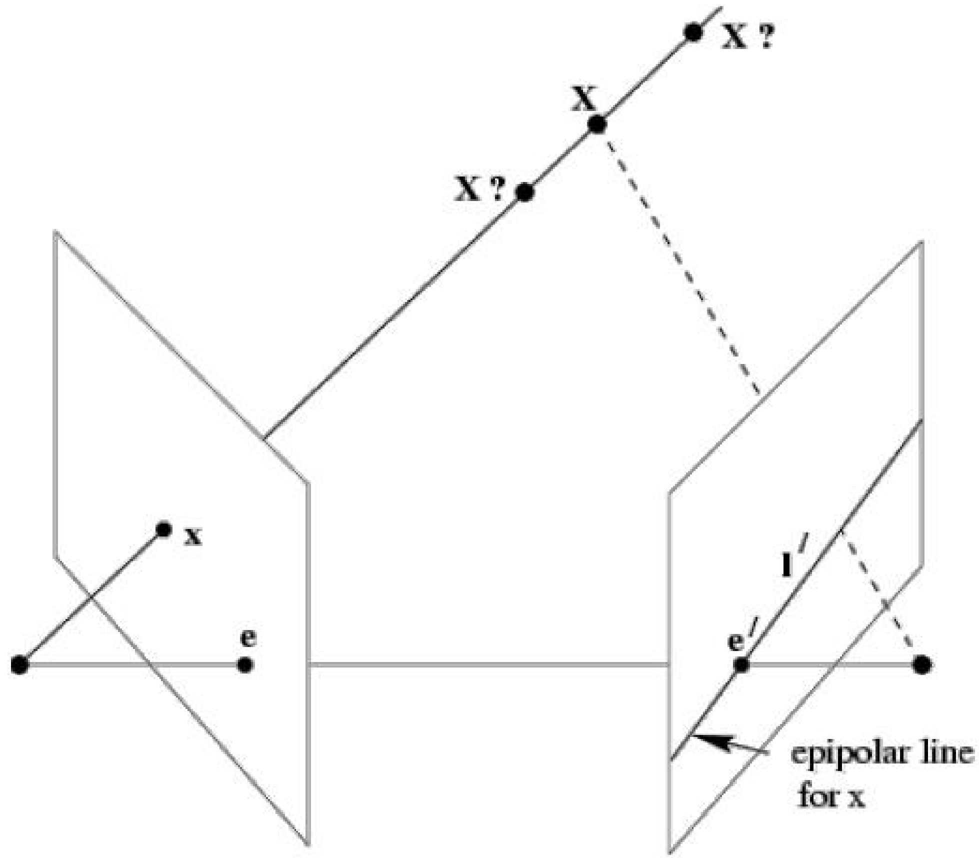
(iii) **Scene geometry (structure):** Given corresponding image points $x_i \leftrightarrow x'_i$ and cameras P, P', what is the position of (their pre-image) X in space?

The epipolar geometry



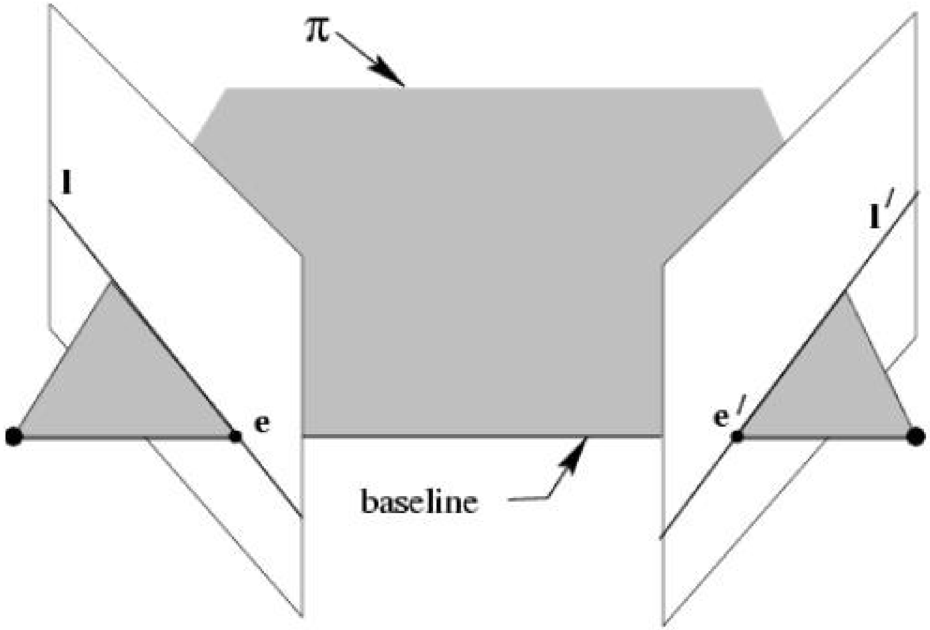
C,C',x,x' and X are coplanar

The epipolar geometry



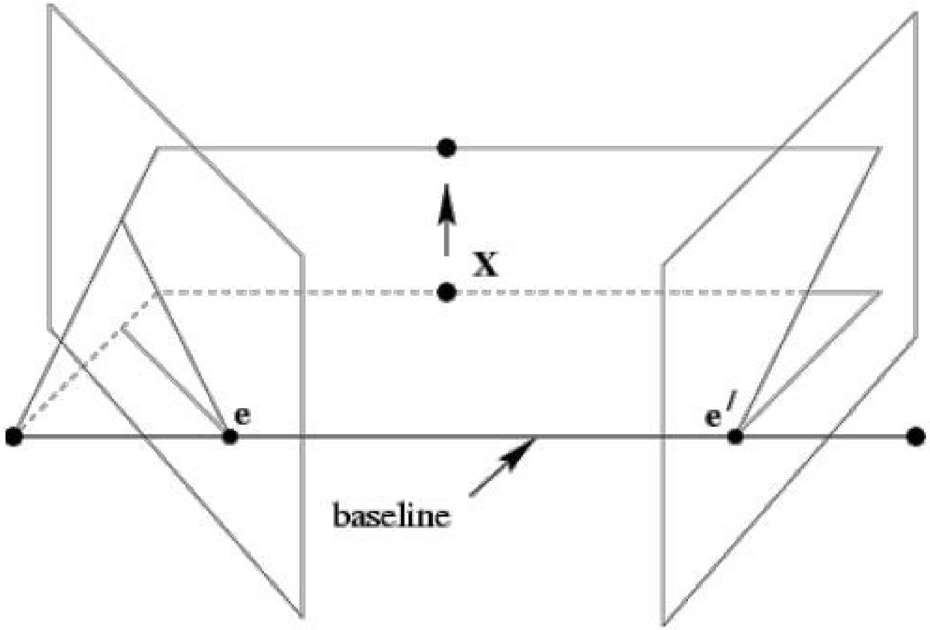
What if only C,C',x are known?

The epipolar geometry



All points on π project on l and l'

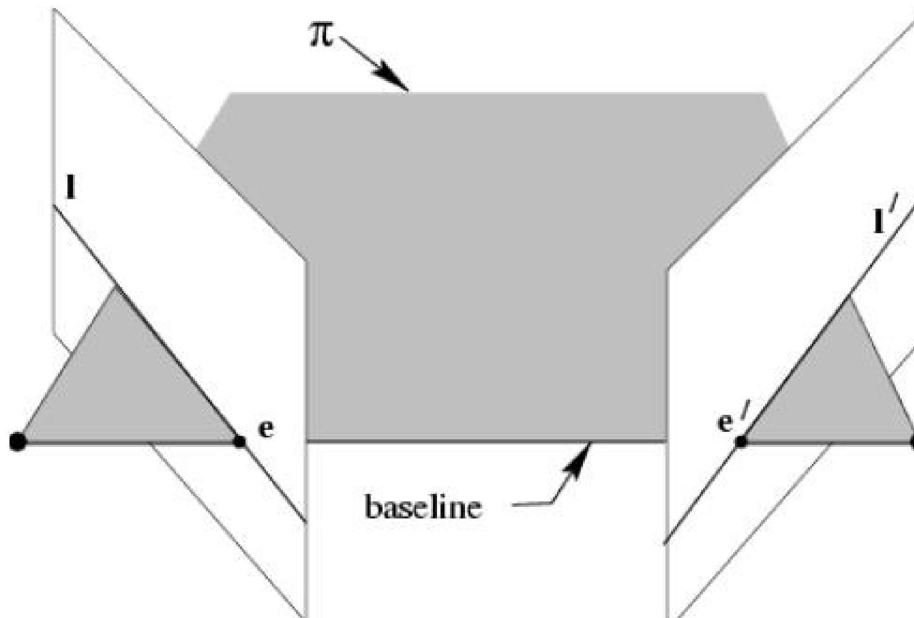
The epipolar geometry



Family of planes π and lines l and l'
Intersection in e and e'

The epipolar geometry

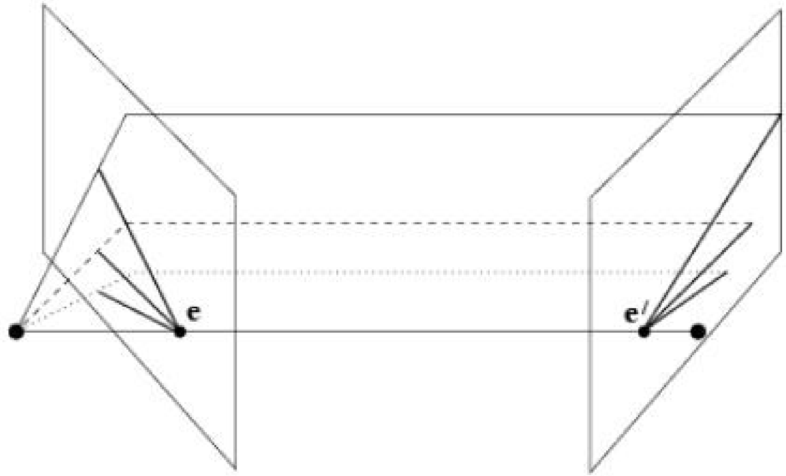
epipoles e,e'
= intersection of baseline with image plane
= projection of projection center in other image
= vanishing point of camera motion direction



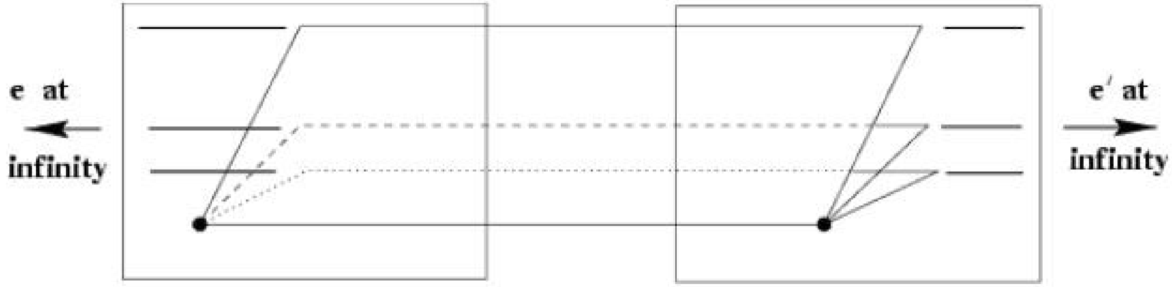
an epipolar plane = plane containing baseline (1-D family)

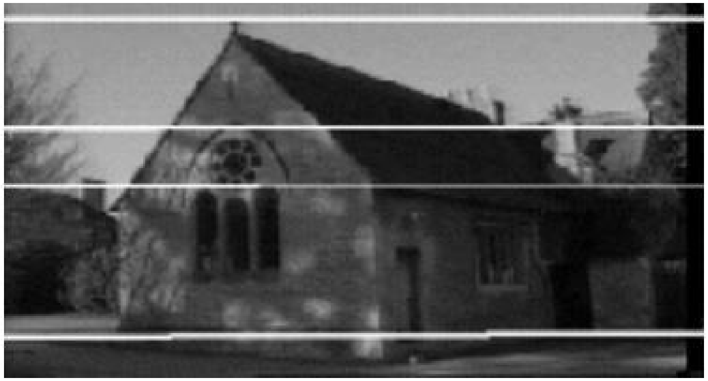
an epipolar line = intersection of epipolar plane with image
(always come in corresponding pairs)

Example: converging cameras



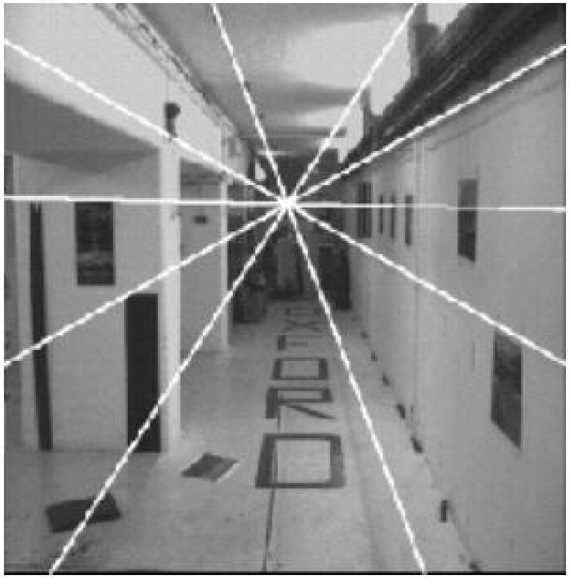
Example: motion parallel with image plane

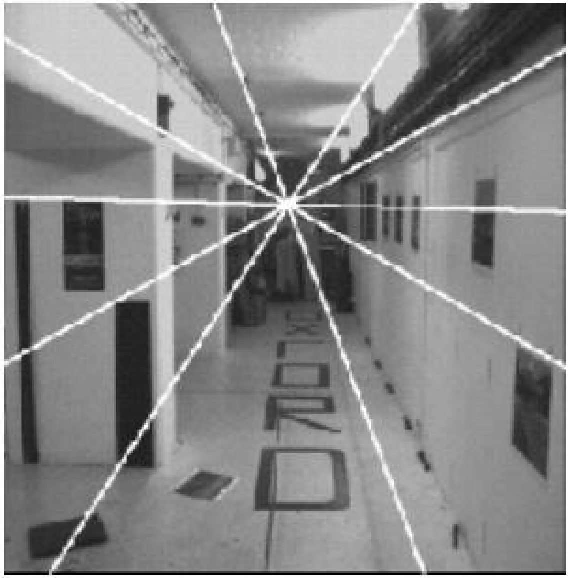


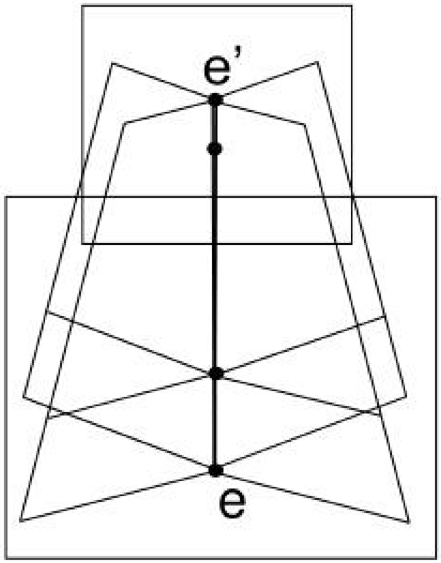


(simple for stereo → rectification)

Example: forward motion







The fundamental matrix F

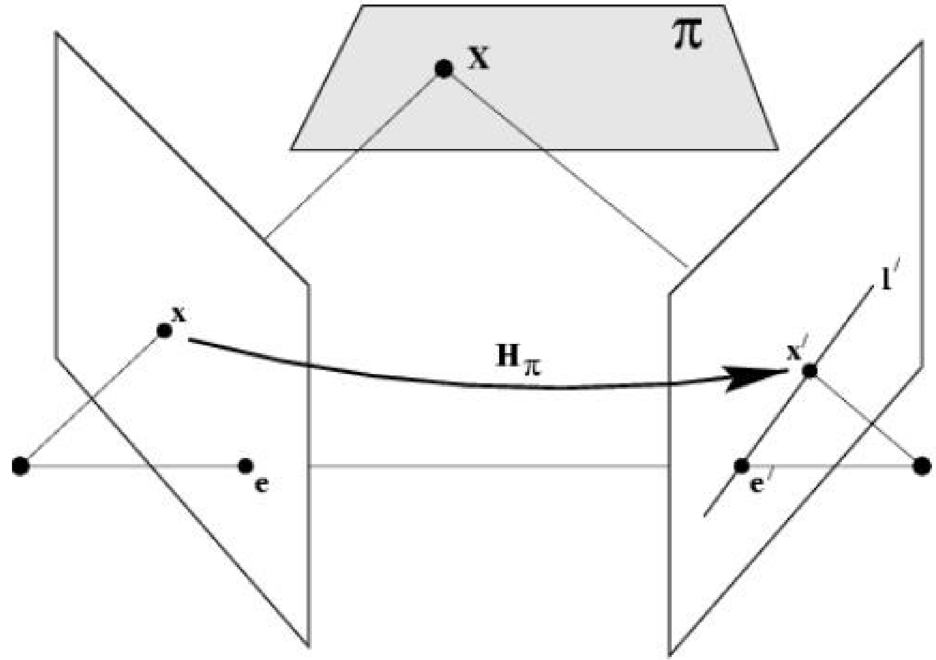
algebraic representation of epipolar geometry

$$x \mapsto l'$$

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)

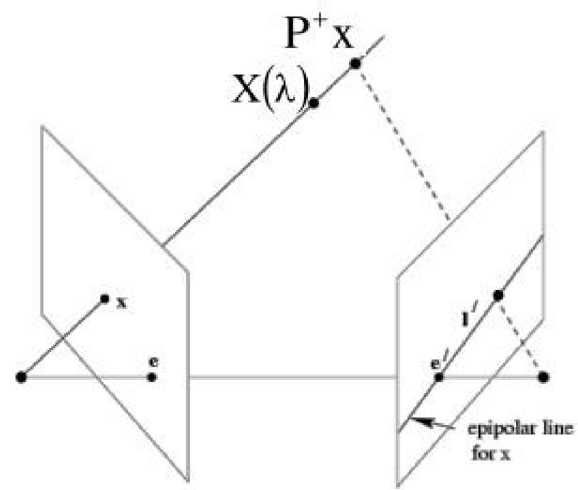
The fundamental matrix F

algebraic derivation

$$X(\lambda) = P^{+}x + \lambda C \qquad \left(PP^{+} = I\right)$$

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

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



(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 $x \leftrightarrow x'$ in the two images

$$x'^{T} F x = 0 \qquad \left(x'^{T} l' = 0 \right)$$

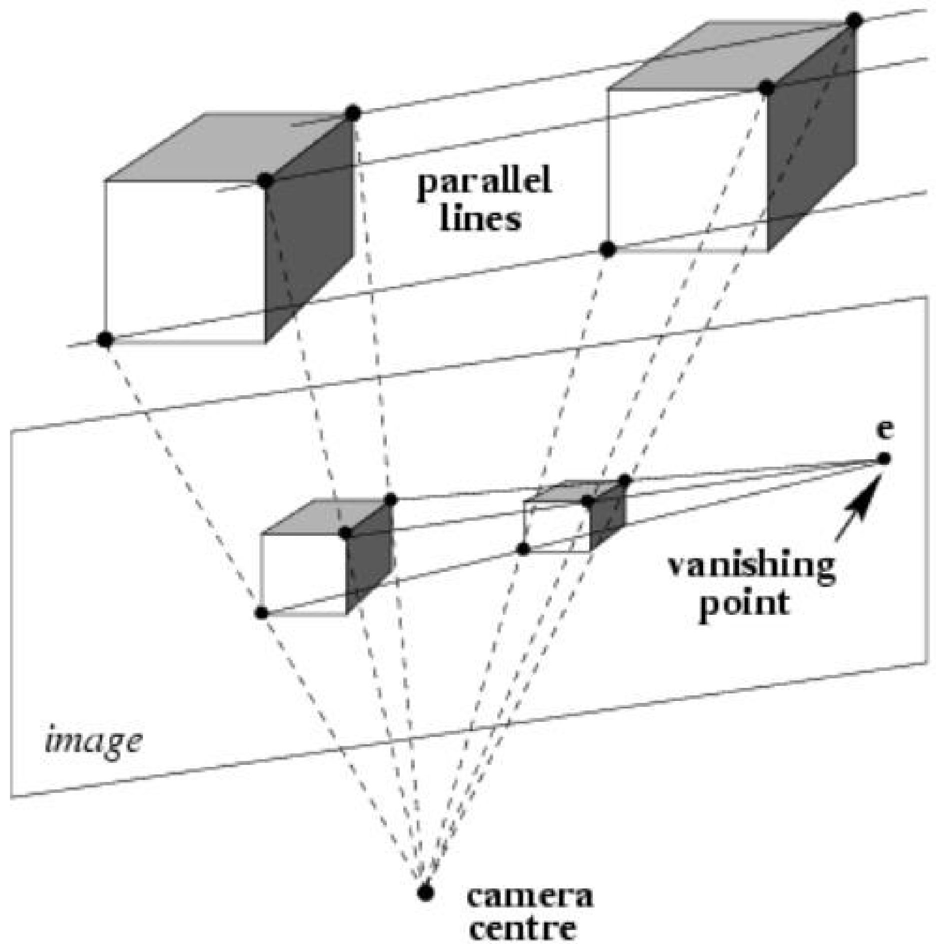
The fundamental matrix F

F is the unique 3x3 rank 2 matrix that satisfies $x'^T F x = 0$ for all $x \leftrightarrow x'$

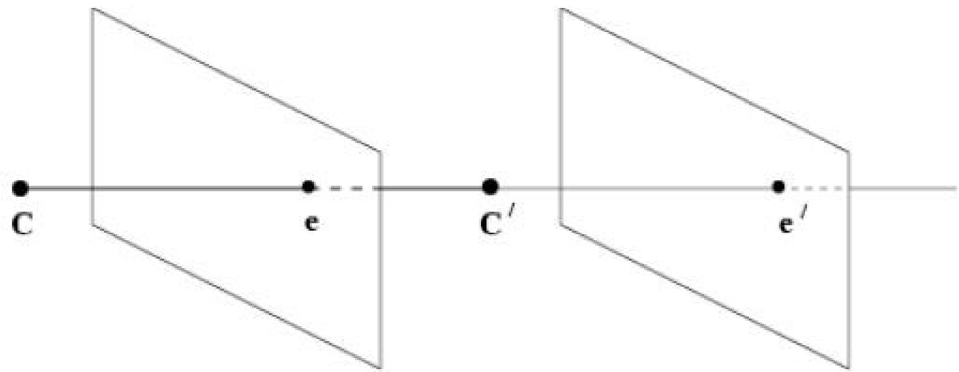
(i) **Transpose:** if F is fundamental matrix for (P,P'), then F^T is fundamental matrix for (P',P)
(ii) **Epipolar lines:** $l' = Fx$ & $l = F^T x'$
(iii) **Epipoles:** on all epipolar lines, thus $e'^T F x = 0, \forall x$ $\Rightarrow e'^T F = 0$, similarly $Fe = 0$
(iv) **F** has 7 d.o.f. , i.e. 3x3-1(homogeneous)-1(rank2)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

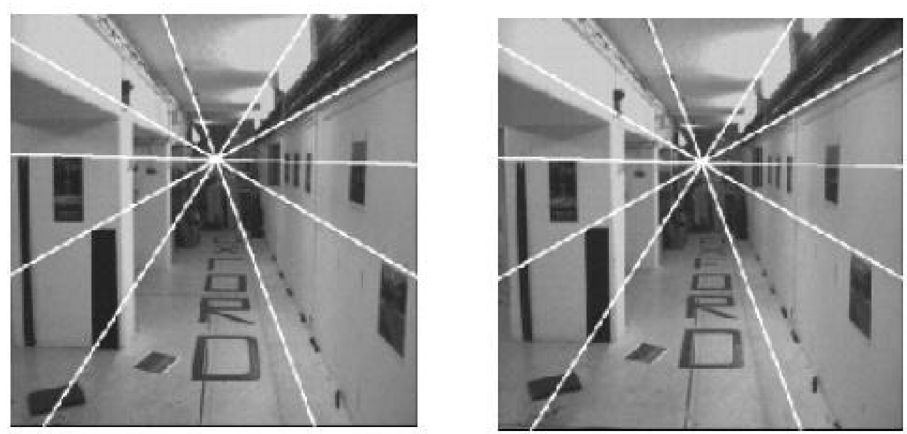
ETH

Fundamental matrix for pure translation



Fundamental matrix for pure translation





Fundamental matrix for pure translation

General motion

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

Pure translation

$$P = K[I \mid 0] \qquad P^+ = \begin{bmatrix} K^{-1} \\ 0 \end{bmatrix}$$

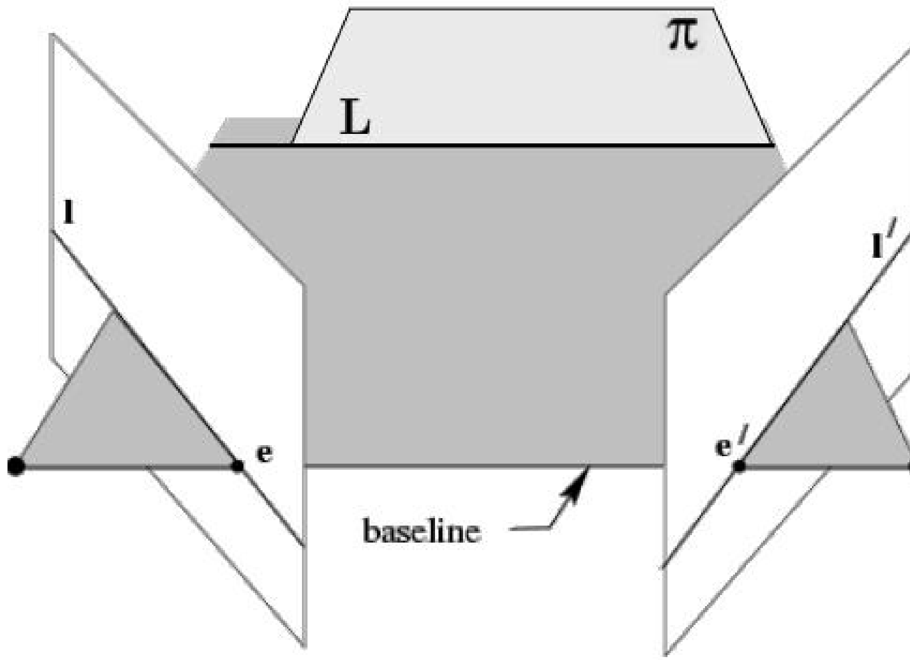
$$P' = K[I \mid t]$$

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

for pure translation F only has 2 degrees of freedom

The fundamental matrix F

relation to homographies

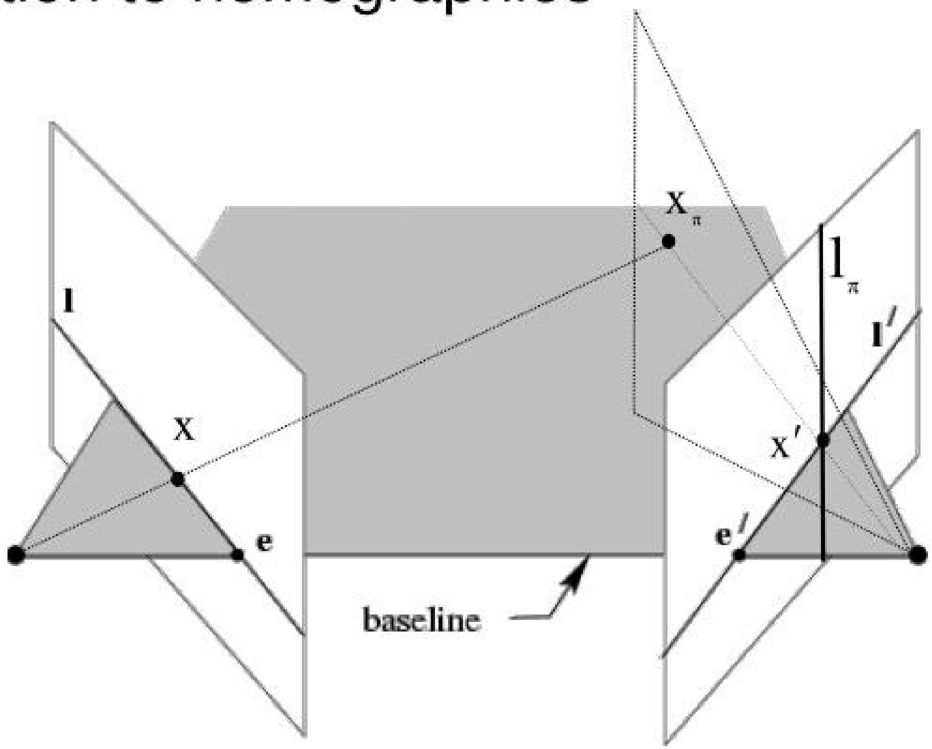


$$[e']_\times H_\pi = F \qquad l' = H_\pi^{-T} l \qquad e' = H_\pi e$$

valid for all plane homographies

The fundamental matrix F

relation to homographies



$$x' = H_\pi x = [l_\pi]_\times F x$$

requires $l_\pi^T e' \neq 0$

$$\text{e.g. } H = [e']_\times F$$

$$(e'^T e' \neq 0)$$

Projective transformation and invariance

Derivation based purely on projective concepts

$$\hat{x} = Hx,\ \hat{x}' = H'x' \Rightarrow \hat{F} = H'^{-T} F H^{-1}$$

F invariant to transformations of projective 3-space

$$x = PX = (PH)(H^{-1}X) = \hat{P}\hat{X}$$
$$x' = P'X = (P'H)(H^{-1}X) = \hat{P}'\hat{X}$$

$(P, P') \mapsto F$ unique

$F \mapsto (P, P')$ not unique

canonical form

$$P = [I \mid 0] \qquad F = [m]_\times M \qquad \left(F = [e']_\times P'P^+\right)$$
$$P' = [M \mid m]$$

ETH

Projective ambiguity of cameras given F

previous slide: at least projective ambiguity
this slide: not more!

Show that if F is same for (P,P') and ($\tilde{P}$,$\tilde{P}$'),
there exists a projective transformation H so that
$\tilde{P}$=HP and $\tilde{P}$'=HP'

$$P=[I\,|\,0] \quad P'=[A\,|\,a]$$
$$\tilde{P}=[I\,|\,0] \quad \tilde{P}'=[\tilde{A}\,|\,\tilde{a}]$$
$$F=[a]_\times A=[\tilde{a}]_\times \tilde{A}$$

<u>lemma:</u> $\tilde{a}=ka$ and $\tilde{A}=k^{-1}\left(A+av^T\right)$

$$aF=a[a]_\times A=0=\tilde{a}F \overset{\text{rank 2}}{\Longrightarrow} \tilde{a}=ka$$
$$[a]_\times A=[\tilde{a}]_\times \tilde{A} \Rightarrow [a]_\times\left(k\tilde{A}-A\right)=0 \Rightarrow \left(k\tilde{A}-A\right)=av^T$$

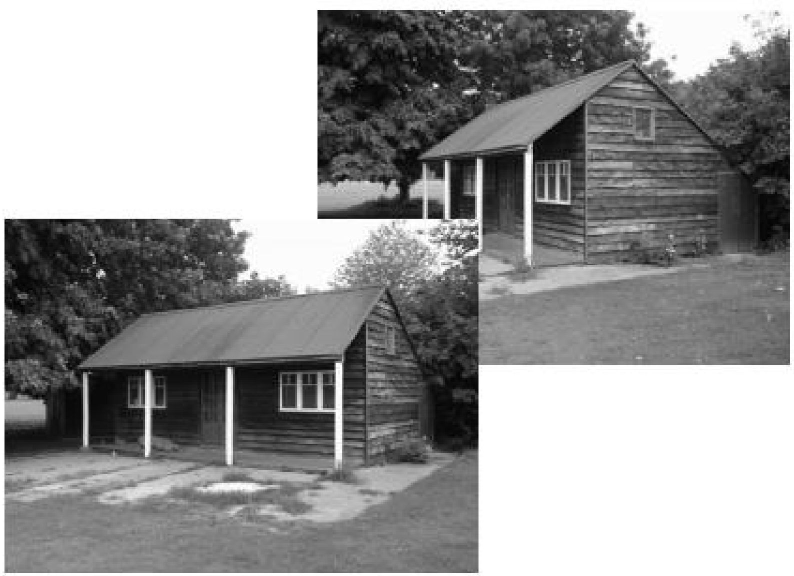
$$H=\begin{bmatrix} k^{-1}I & 0 \\ k^{-1}v^T & k \end{bmatrix} \quad P'H=[A\,|\,a]\begin{bmatrix} k^{-1}I & 0 \\ k^{-1}v^T & k \end{bmatrix}$$
$$=[k^{-1}\left(A-av^T\right)|\,ka]=\tilde{P}'$$

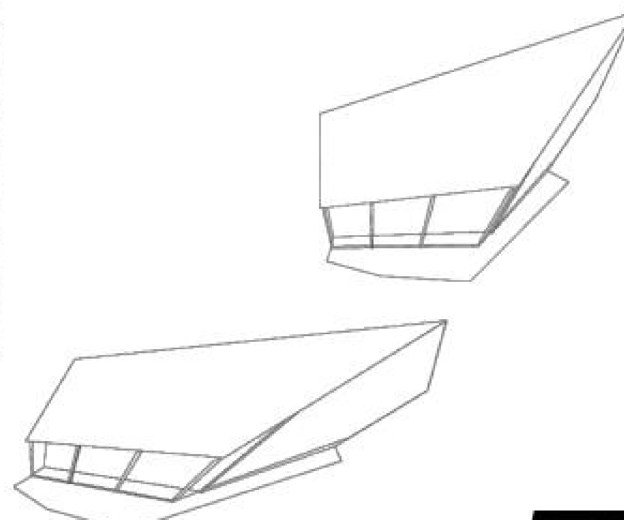
(22-15=7, ok)

The projective reconstruction theorem

If a set of point correspondences in two views determine the fundamental matrix uniquely, then the scene and cameras may be reconstructed from these correspondences alone, and any two such reconstructions from these correspondences are projectively equivalent

allows reconstruction from pair of uncalibrated images!

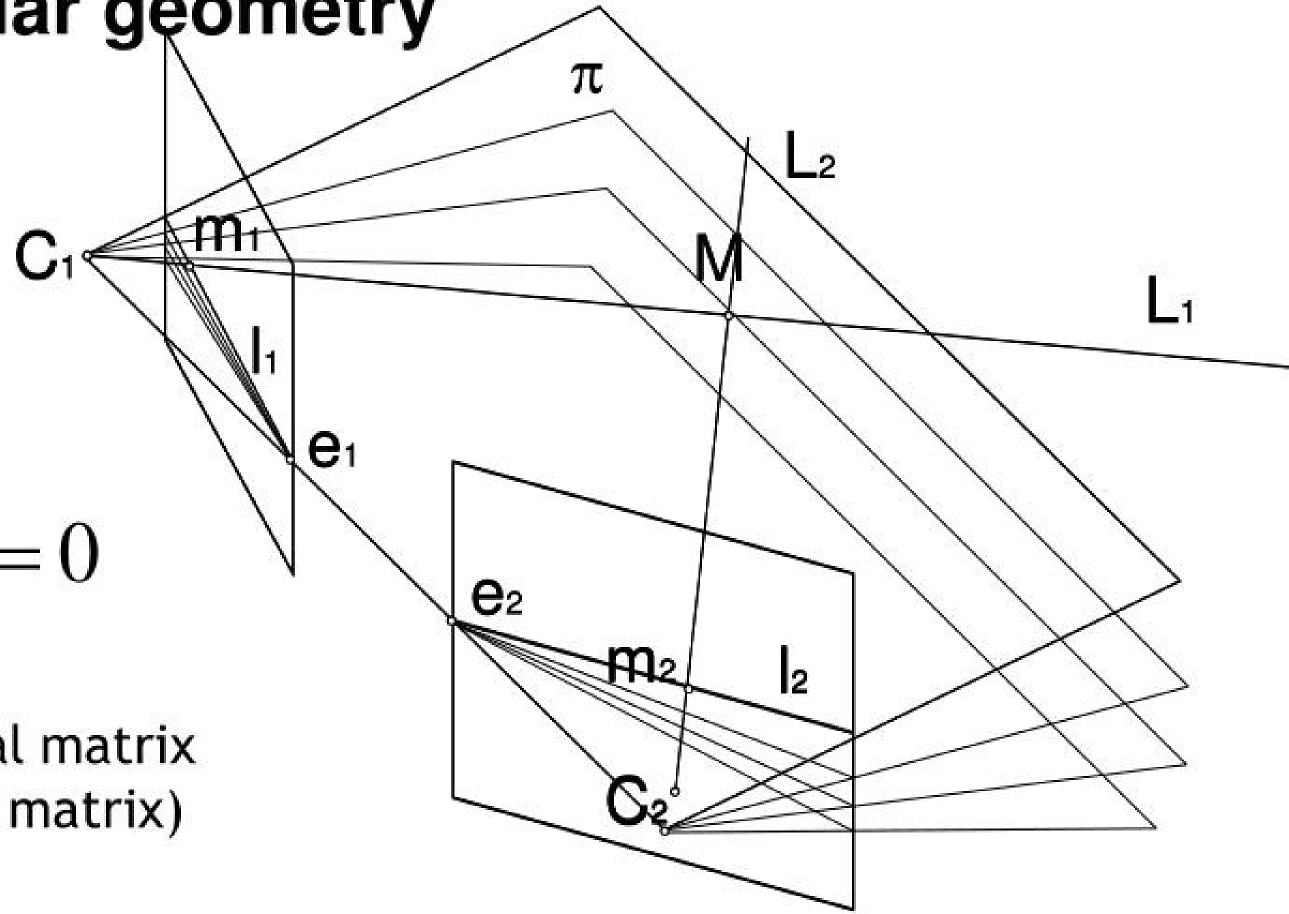




ETH

Epipolar geometry

Underlying structure in set of matches for rigid scenes



$l^T_1 \quad l_2$

$$m_2^T \mathbf{F}\, m_1 = 0$$

Fundamental matrix (3x3 rank 2 matrix)

1. Computable from corresponding points
2. Simplifies matching
3. Allows to detect wrong matches
4. Related to calibration

Canonical representation:

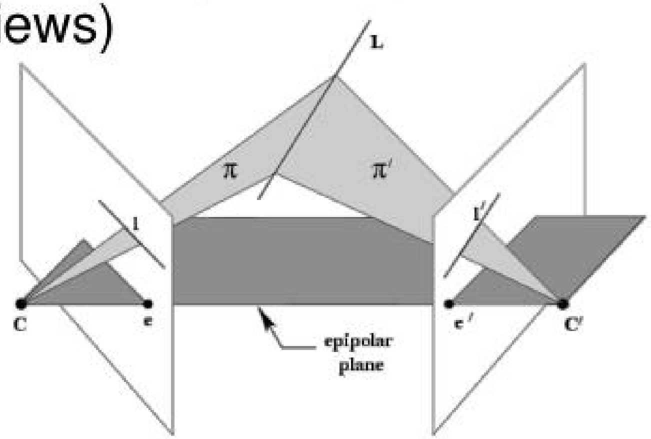
$$P = [I \mid 0] \quad P' = [[e']_\times F + e' v^T \mid \lambda e']$$

Epipolar geometry?

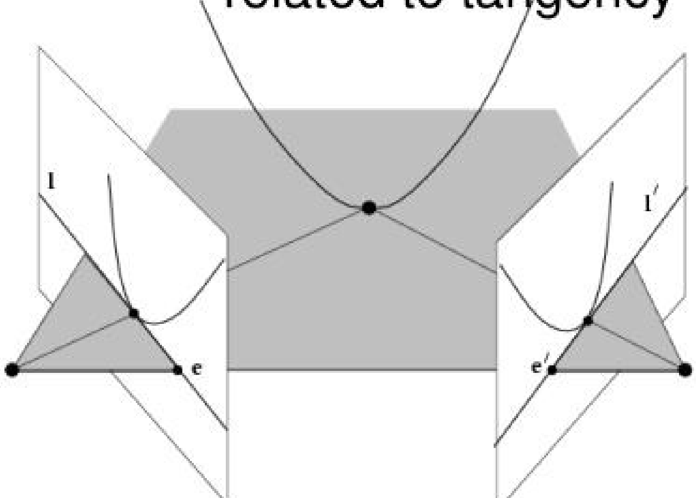
courtesy Frank Dellaert

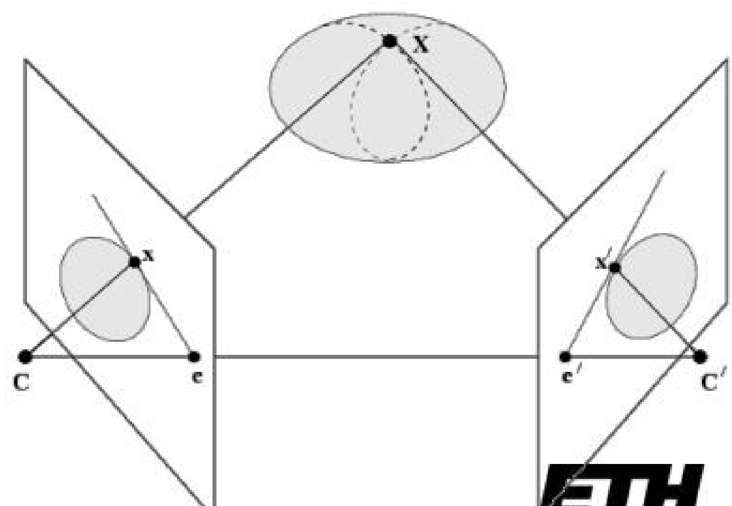
Other entities besides points?

Lines give no constraint for two view geometry
(but will for three and more views)



Curves and surfaces yield some constraints
related to tangency





(e.g. Sinha et al. CVPR'04)

Computation of F

- Linear (8-point)
- Minimal (7-point)
- Robust (RANSAC)
- Non-linear refinement (MLE, …)

- Practical approach

Epipolar geometry: basic equation

$$\mathrm{x'}^{\mathrm{T}}\,\mathrm{F}\mathrm{x} = 0$$

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

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)

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

$$\mathrm{Af} = 0$$

the NOT normalized 8-point algorithm

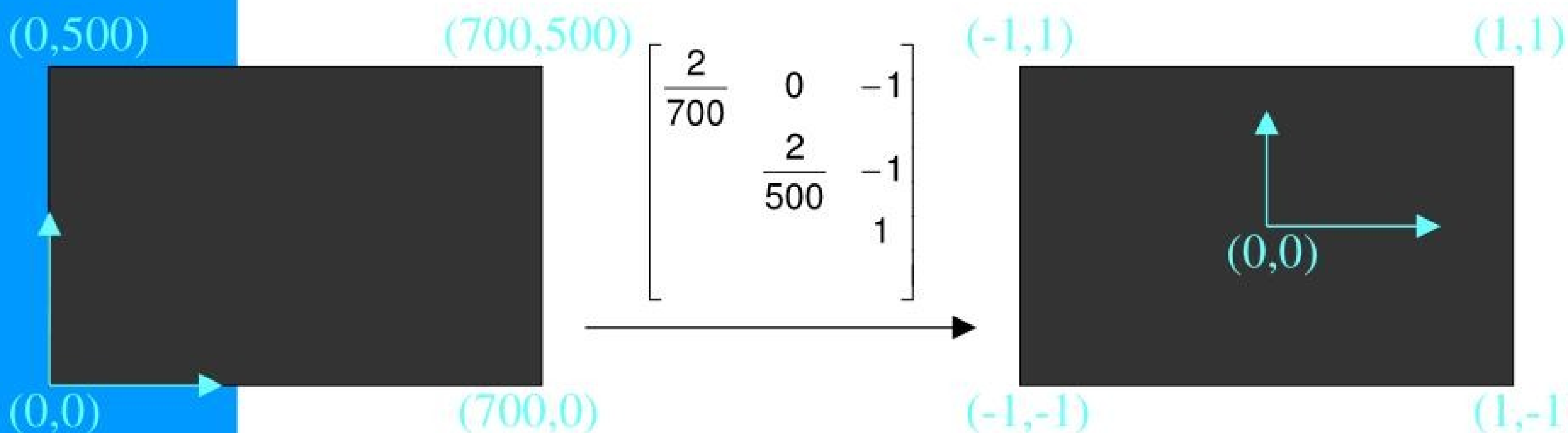
$$\begin{bmatrix} x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \\ x_2x_2' & y_2x_2' & x_2' & x_2y_2' & y_2y_2' & y_2' & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx_n' & y_nx_n' & x_n' & x_ny_n' & y_ny_n' & y_n' & x_n &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

$$\begin{array}{ccccccccc} \sim 10000 & \sim 10000 & \sim 100 & \sim 10000 & \sim 10000 & \sim 100 & \sim 100 & \sim 100 & 1 \end{array}$$

! 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]



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

normalized least squares yields good results

(Hartley, PAMI´97)

the singularity constraint

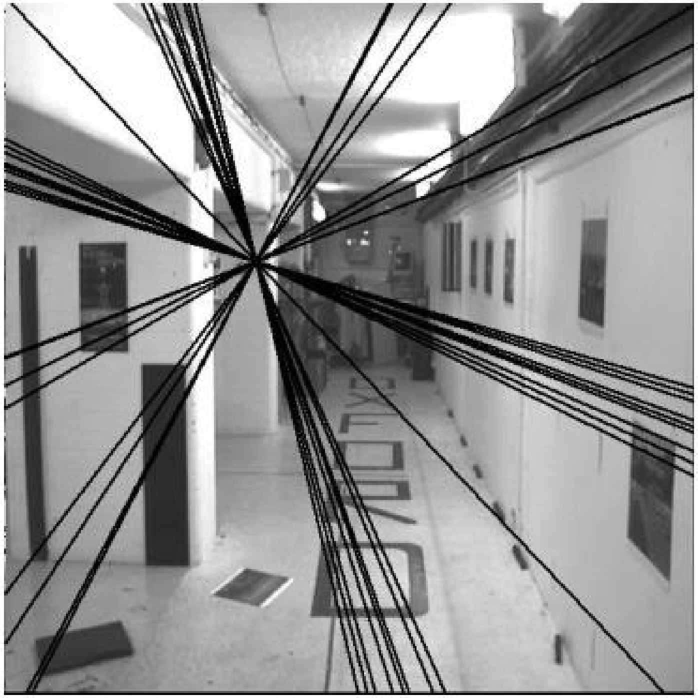
$$e'^{T} F = 0 \qquad Fe = 0 \qquad \det F = 0 \qquad \mathrm{rank}\ F = 2$$

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 $\min \|F - F'\|_F$

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



the minimum case – 7 point correspondences

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_7 x_7 & x'_7 y_7 & x'_7 & y'_7 x_7 & y'_7 y_7 & y'_7 & x_7 & y_7 & 1 \end{bmatrix} \mathrm{f} = 0$$

$$\mathrm{A} = \mathrm{U}_{7\mathrm{x}7} \mathrm{diag}(\sigma_1, \ldots, \sigma_7, 0, 0) \mathrm{V}_{9\mathrm{x}9}{}^{\mathrm{T}}$$

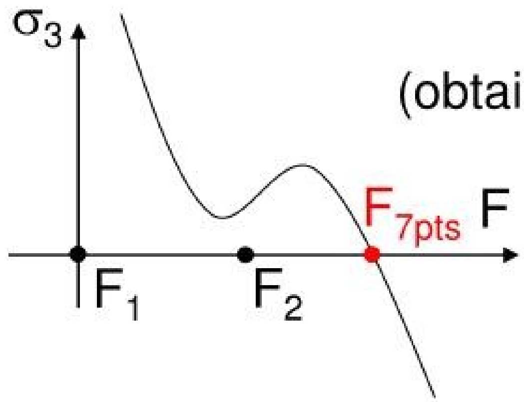
$$\Rightarrow \mathrm{A}[\mathrm{V}_8 \mathrm{V}_9] = 0_{9\mathrm{x}2} \qquad \left(\mathrm{e.g.} \mathrm{V}^{\mathrm{T}} \mathrm{V}_8 = [000000010]^{\mathrm{T}}\right)$$

$$\mathrm{x}_i{}^{\mathrm{T}} (\mathrm{F}_1 + \lambda \mathrm{F}_2) \mathrm{x}_i = 0, \forall i = 1 \ldots 7$$

one parameter family of solutions

but $F_1+\lambda F_2$ not automatically rank 2

the minimum case – impose rank 2



(obtain 1 or 3 solutions)

$$\det(F_1+\lambda F_2)=a_3\lambda^3+a_2\lambda^2+a_1\lambda+a_0=0$$ (cubic equation)

$$\det(F_1+\lambda F_2)=\det F_2 \det(F_2^{-1}F_1+\lambda I)=0$$ $(\det(AB)=\det(A).\det(B))$

Compute possible λ as eigenvalues of $F_2^{-1}F_1$
(only real solutions are potential solutions)

Automatic computation of F

Step 1. Extract features

Step 2. Compute a set of potential matches

RANSAC

Step 3. do

- Step 3.1 select minimal sample (i.e. 7 matches)
- Step 3.2 compute solution(s) for F

(generate hypothesis)

- Step 3.3 determine inliers (verify hypothesis)

until $\Gamma(\#inliers, \#samples) < 95\%$

Step 4. Compute F based on all inliers

Step 5. Look for additional matches

Step 6. Refine F based on all correct matches

$$\Gamma = 1 - \left(1 - \left(\frac{\#inliers}{\#matches}\right)^7\right)^{\#samples}$$

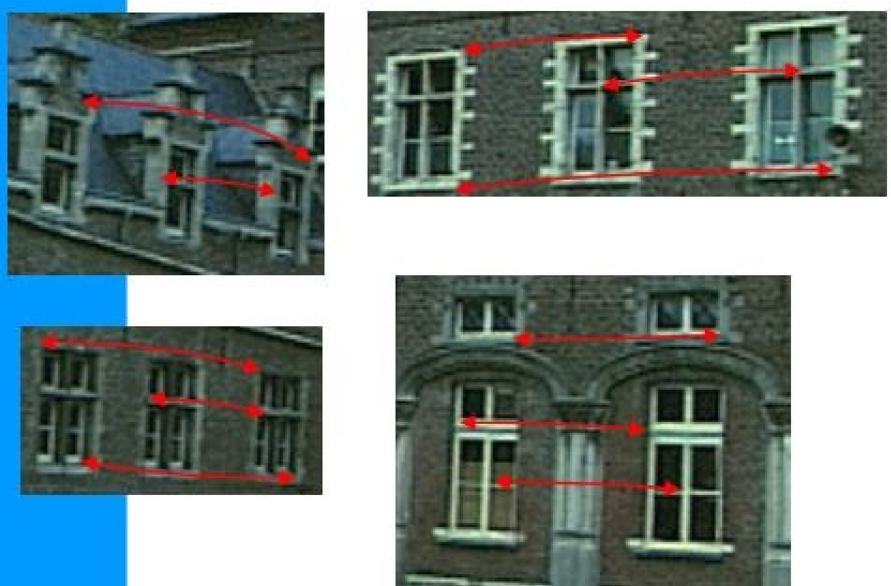
#inliers	90%	80%	70%	60%	50%
#samples	5	13	35	106	382

Finding more matches

restrict search range to neighborhood of epipolar line
(e.g. $\pm$1.5 pixels)
relax disparity restriction (along epipolar line)

Issues:

- (Mostly) planar scene (see next slide)
- Absence of sufficient features (no texture)
- Repeated structure ambiguity



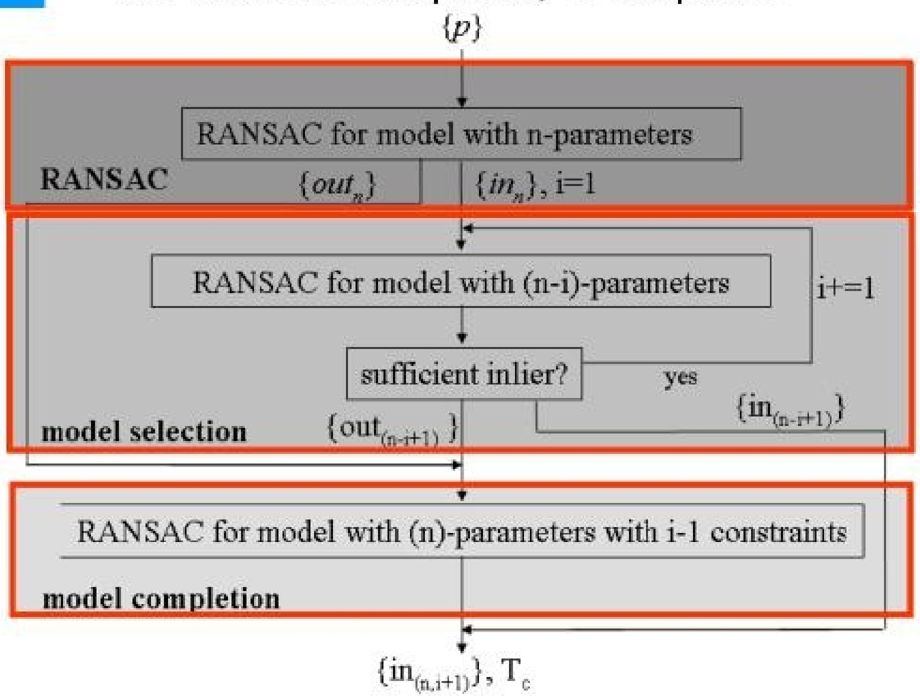
- Robust matcher also finds support for wrong hypothesis
- solution: detect repetition

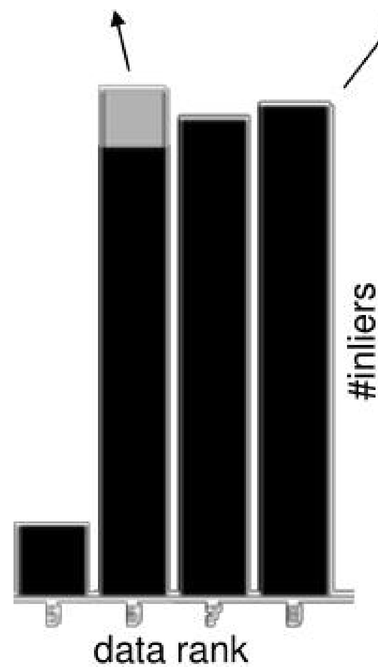
(Schaffalitzky and Zisserman, BMVC'98)

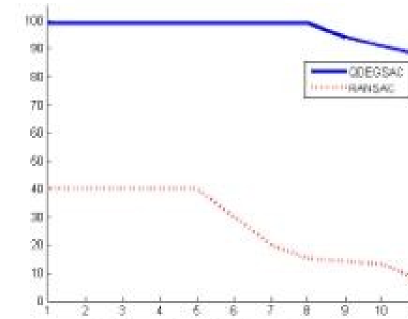
ETH

Computing F for quasi-planar scenes QDEGSAC

337 matches on plane, 11 off plane



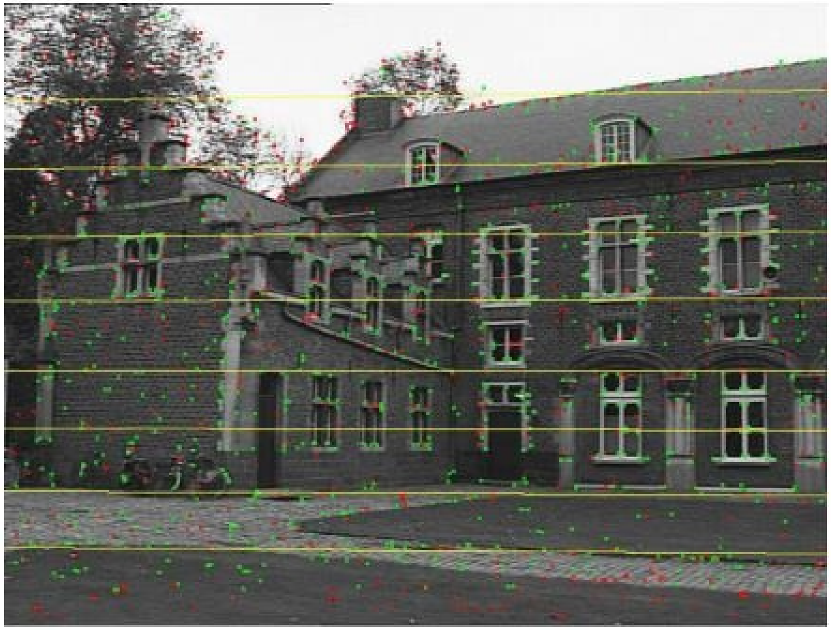




%inclusion of out-of-plane inliers

17% success for RANSAC
100% for QDEGSAC

two-view geometry



geometric relations between two views is fully described by recovered 3x3 matrix F