55:148 Digital Image Processing

Chapter 11

3D Vision, Geometry

Topics:

- Basics of projective geometry
  - Points and hyperplanes in projective space
  - Homography
  - Estimating homography from point correspondence
- The single perspective camera
  - An overview of single camera calibration
  - Calibration of one camera from the known scene
- Scene reconstruction from multiple views
  - Triangulation
  - Projective reconstruction
  - Matching constraints
  - Bundle adjustment
- Two cameras, stereopsis
  - The geometry of two cameras. The fundamental matrix
  - Relative motion of the camera; the essential matrix
  - Estimation of a fundamental matrix from image point correspondences
  - Camera Image rectification
  - Applications of the epipolar geometry in vision
- Three and more cameras
  - Stereo correspondence algorithms

Epipolar geometry and Fundamental matrix

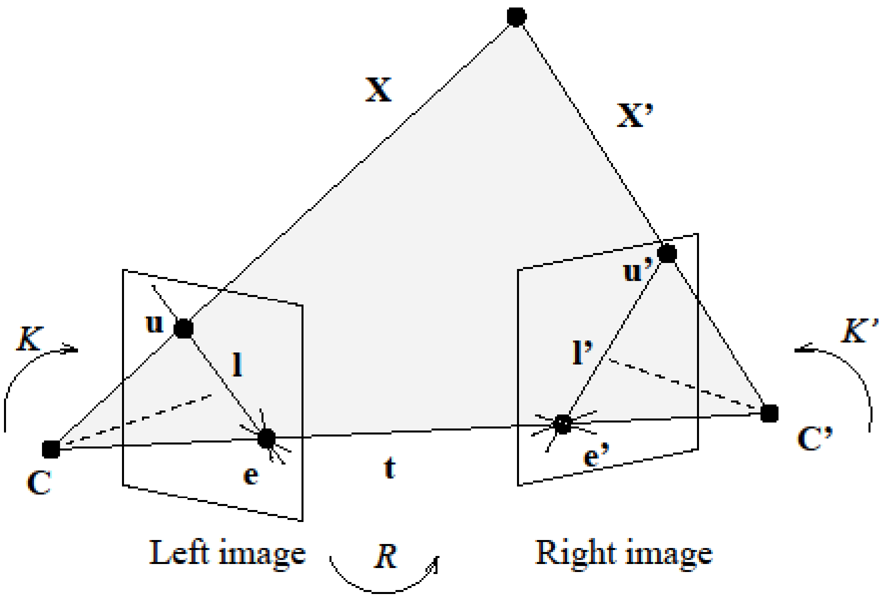
Fundamental matrix relates corresponding points in two stereo images

$$\mathbf{u}'^{\mathrm{T}} F \mathbf{u} = 0$$

What does it mean?

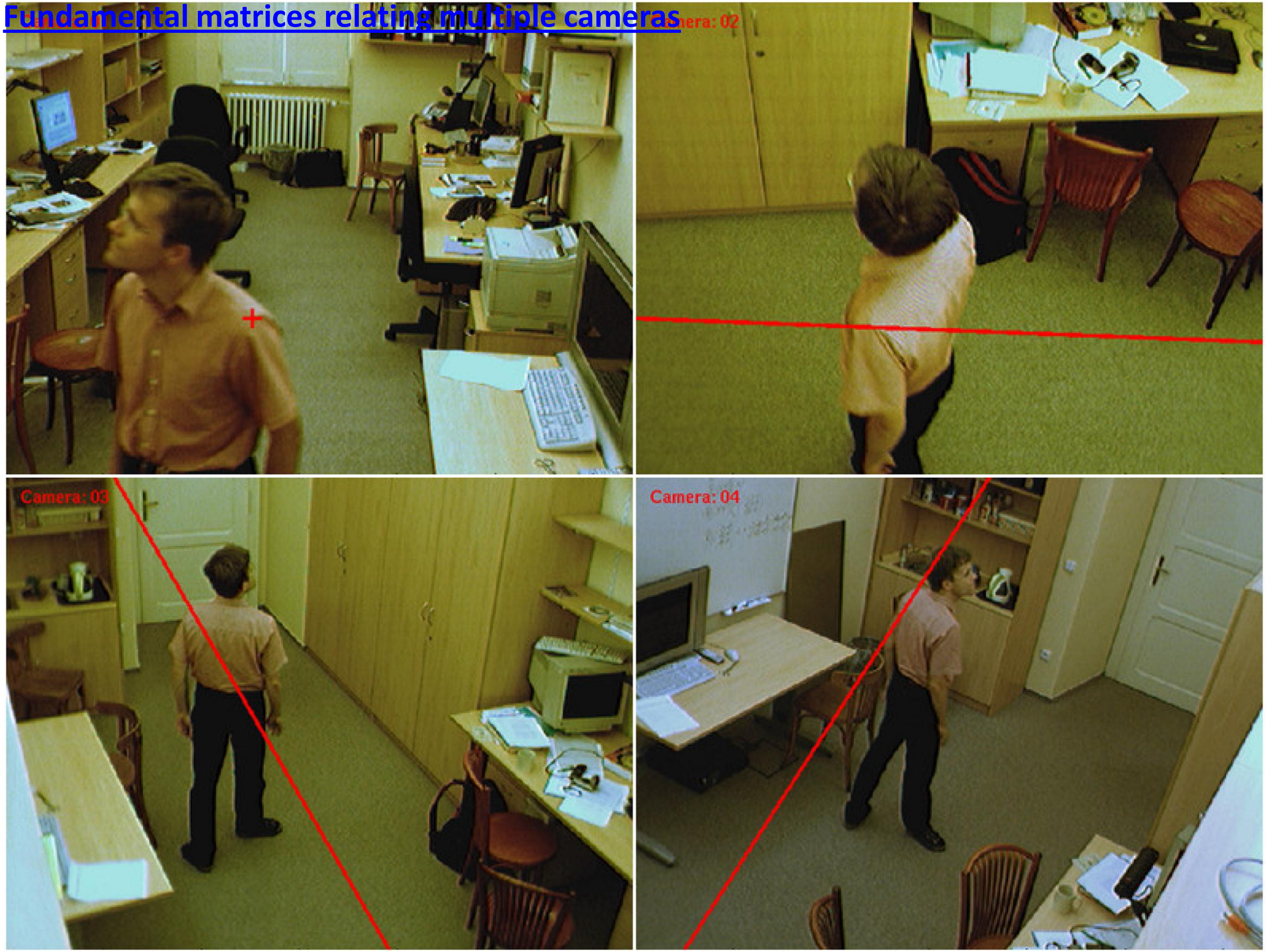
A point on the left image ≈ a line on the right image

What is this line called



Fundamental matrices relating multiple cameras

Fundamental matrices relating multiple cameras



Fundamental matrices relating multiple cameras

Fundamental matrices relating multiple cameras

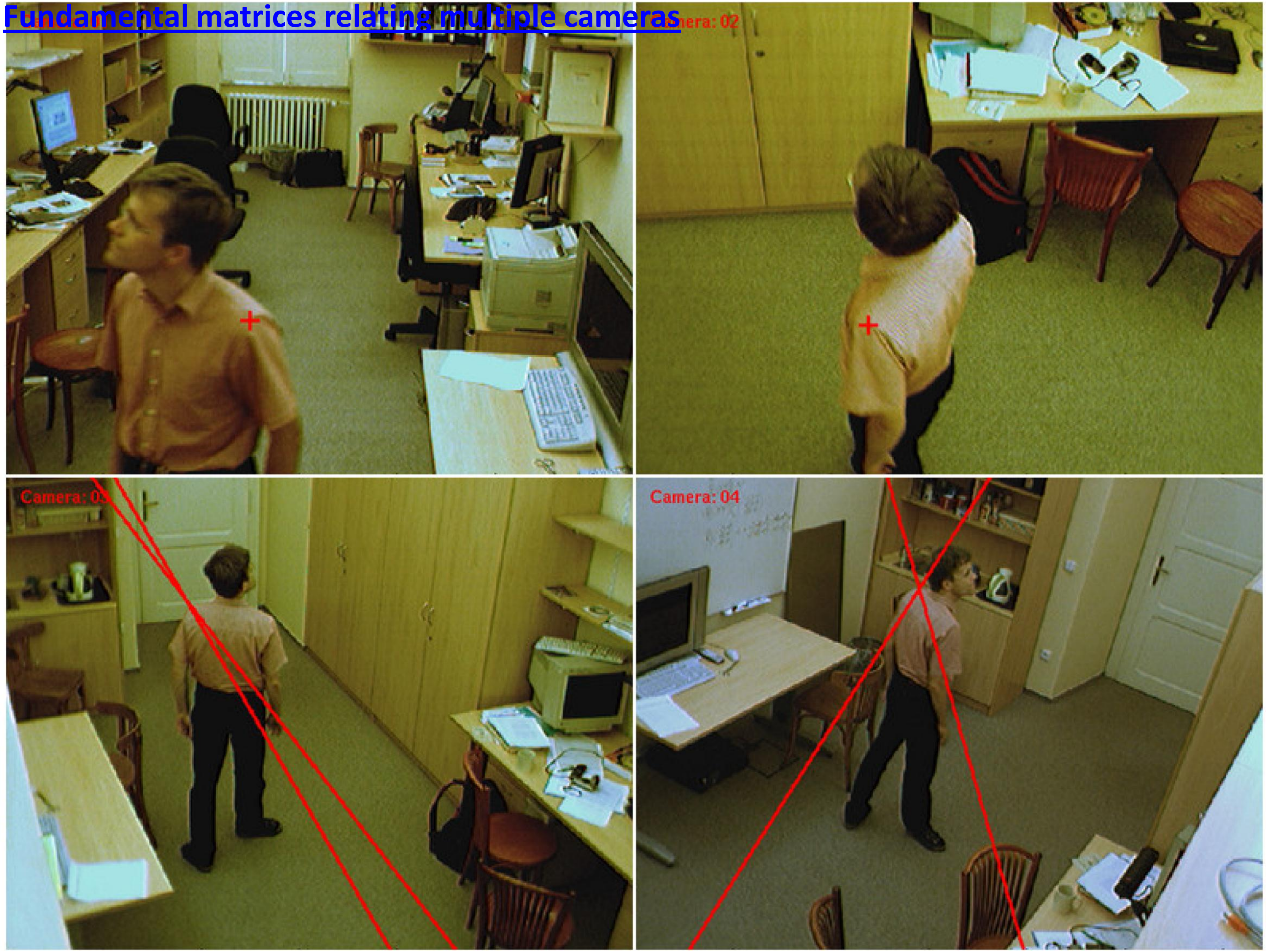


Image rectification (before)

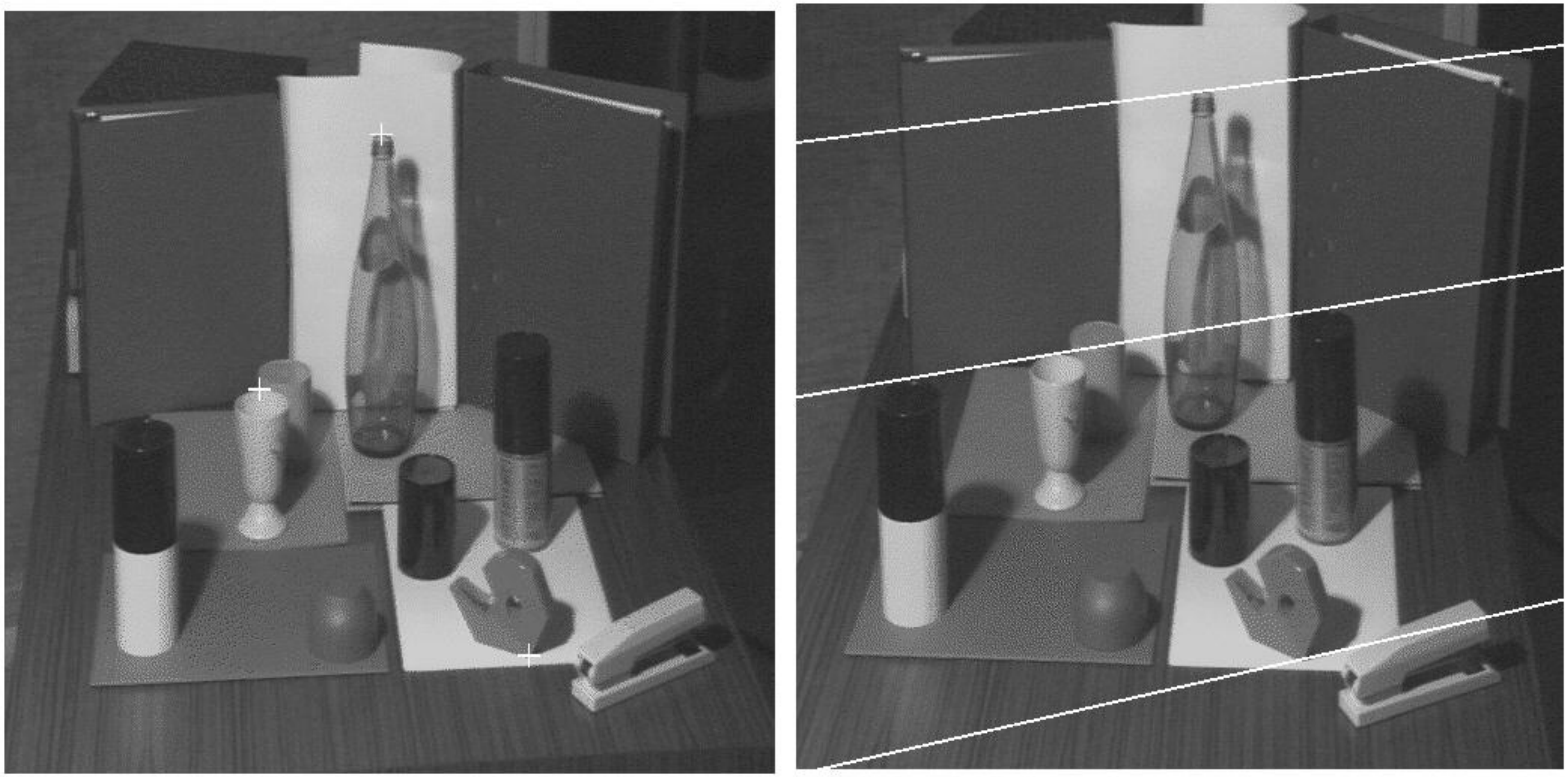


Image rectification (after)

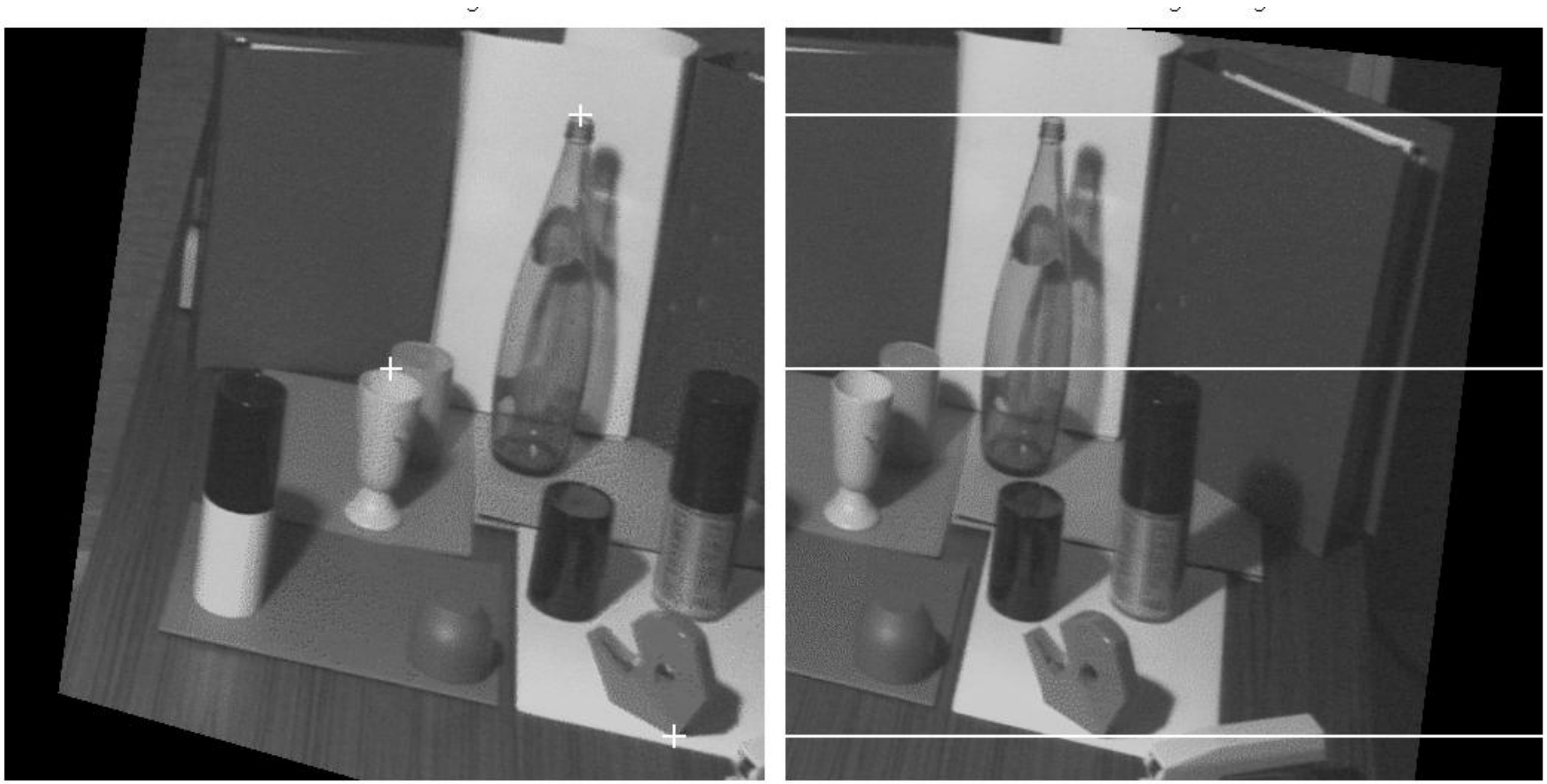


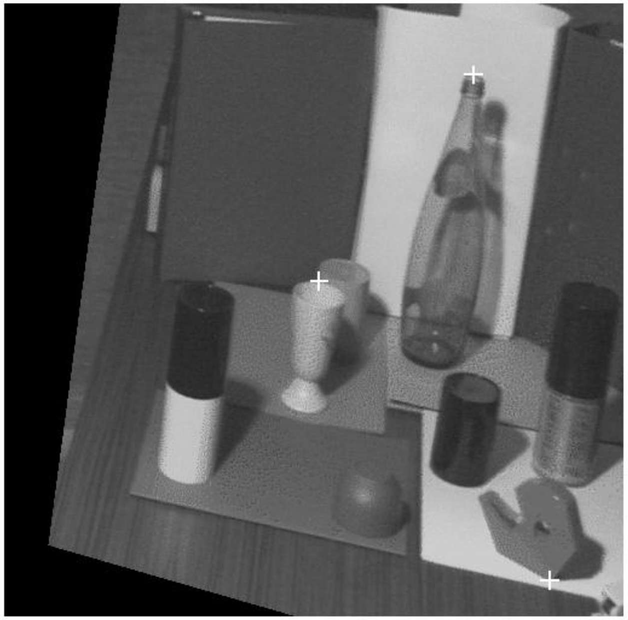
Image rectification

What happens in terms of epipolar geometry?

Where are the two epipoles?

What is the relation between the baseline and the camera matrix?

Can we solve it using a homographic transformation on each camera image?



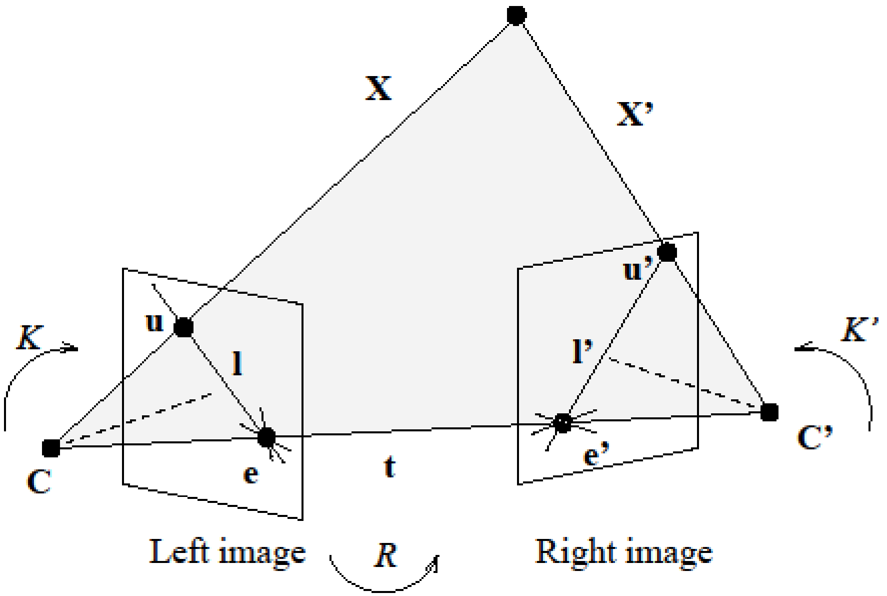


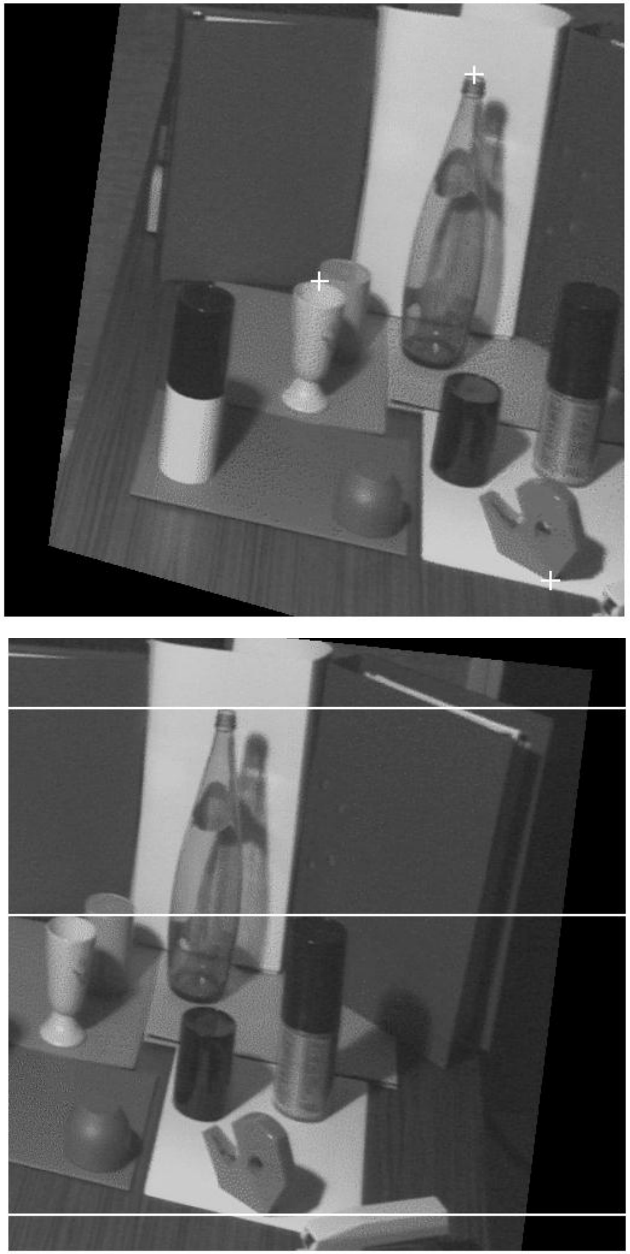
Image rectification

What happens in terms of epipolar geometry?

Where are the two epipoles?

What is the relation between the baseline and the camera matrix?

Can we solve it using a homographic transformation on each camera image?



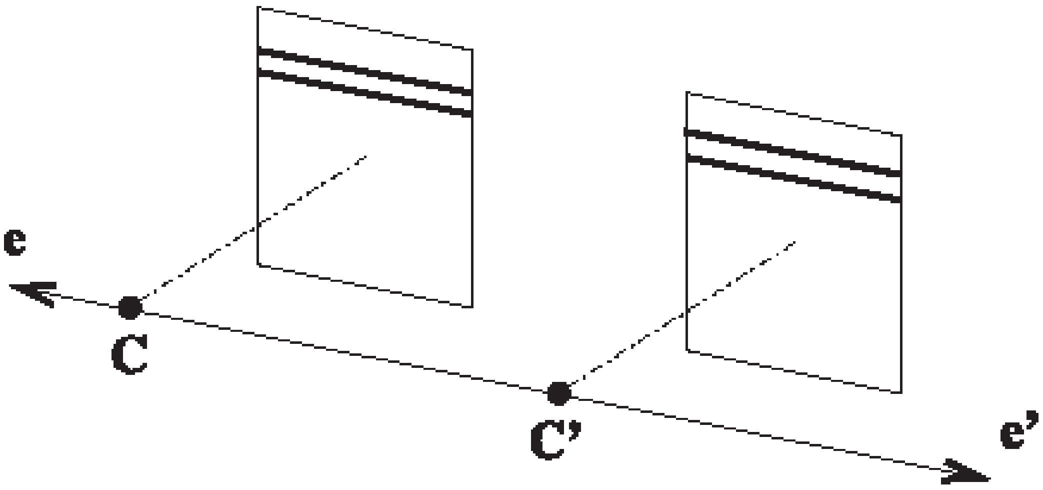


Image rectification: advantages

3D reconstruction becomes easier

Image stitching to generate a panoramic view

Panoramic view

So, how to accomplish image rectification?

- Learn how to determine the fundamental matrix
- Relative camera motion and essential matrix
- Relation between fundamental matrix and camera matrix
- Compute image rectification

Relative camera motion and essential matrix

In the previous class, we have seen: $F = K'^{-\mathrm{T}} R S(\mathbf{t}) K^{-1}$

K' and K are intrinsic camera parameters that maps Euclidean image plane to image pixels; primarily plays a role to correct the shear distortion between the x- and y-axes.

It's very difficult to determine K' and K without use of a known 3D scene and just by using the correspondence between two acquired images

Thus, if we ignore this shear component, the epipolar constraint in the image Euclidean plane translates to

$$\mathbf{u}_i'^{\mathrm{T}} R S(\mathbf{t}) \mathbf{u}_i = 0 \Rightarrow \mathbf{u}_i'^{\mathrm{T}} E \mathbf{u}_i = 0, \qquad \text{where } E = RS(\mathbf{t})$$

E is called the **essential matrix** that defines the relative motion between two camera position

Application: Determine camera movements from a video image (http://www.2d3.com/)

Relation between fundamental matrix and essential matrix (when we know K' and K)

$$E = K'^{\mathrm{T}} F K$$

Decomposition of essential matrix

Note that the vector $\mathbf{t}$ in the essential matrix $E = RS(\mathbf{t})$ tells us about the relative location of the two optical centers. i.e., the baseline.

Also, assuming that the camera matrix $M = [I \mid \mathbf{0}]$ for the first camera, R and $\mathbf{t}$ together determine M' -- the camera matrix of the second camera

Now, assume that, somehow, we have computed the essential matrix E
But, it does not immediately give us the translation vector $\mathbf{t}$ or the rotation matrix R

So, we need to decompose E

Singular value decomposition of E gives $E = UDV^{\mathrm{T}}$, U and V are rotation matrices.

Following that the rows of $S(\mathbf{t})$ are coplanar (why), it has a rank of two and the two singular values are equal (follows from the formulation of $S(\mathbf{t})$); so

$$D = diag[\sigma, \sigma, 0]$$

We will later see that scale factor in the actual computation of E is arbitrarily set

Decomposition of the essential matrix *continued …*

Denote

$$\bar{\mathbf{t}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \bar{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then the translation vector is given by

$$S(\mathbf{t}) = VS(\bar{\mathbf{t}})V^{\mathrm{T}}$$

The rotation matrix is not given uniquely, we have

$$R = U\bar{R}V^{\mathrm{T}} \qquad \text{or} \qquad R = U\bar{R}^{\mathrm{T}}V^{\mathrm{T}}$$

Before getting into image rectification, we need to learn

- Relation between the fundamental matrix and the camera matrix
- How to compute the fundamental matrix

Camera matrices:

$$M = [I \mid \mathbf{0}]$$

$$M' = [S(\mathbf{e}')F \mid \mathbf{e}']$$

Computation of the fundamental matrix using point correspondence

Number of unknowns:
9 parameters in F minus one for scale standardization minus one for rank of F is two

$$9-1-1=7$$

So, we can solve F with $m \geq 8$ corresponding point pairs in two images.

We have to solve the following linear system:

$$\mathbf{u}_i'^{\mathrm{T}} F \mathbf{u}_i = 0, \qquad i = 1,2,\dots,m$$

Use Kronecker product identity: $AB\mathbf{c} = \left(\mathbf{c}^{\mathrm{T}} \otimes A\right)\mathbf{b}$

$$\mathbf{u}_i'^{\mathrm{T}} F \mathbf{u}_i = \left[\mathbf{u}_i^{\mathrm{T}} \otimes \mathbf{u}_i'^{\mathrm{T}}\right]\mathbf{f} = 0$$

Put together all point correspondences

$$\begin{bmatrix} \mathbf{u}_{i,1}^{\mathrm{T}} \otimes \mathbf{u}_{i,1}'^{\mathbf{T}} \\ \vdots \\ \mathbf{u}_{i,m}^{\mathrm{T}} \otimes \mathbf{u}_{i,m}'^{\mathbf{T}} \end{bmatrix} \mathbf{f} = W\mathbf{f} = 0$$

Compute $W^T W$ and apply singular value decomposition; choose $\mathbf{f}$ along the eigenvector corresponding to the smallest eigenvalue

Computation of the fundamental matrix using maximum likelihood estimation

$$\min_{F,u_i,v_i,u_i',v_i'} [(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2 + (u_i' - \hat{u}_i')^2 + (v_i' - \hat{v}_i')^2]$$

Given $[u_i', v_i', 1] F [u_i, v_i, 1]^{\mathrm{T}} = 0$ and $\det F = 0$

Use **Lagrange multiplier**

$$\text{maximize } f(x,y), \text{ given } g(x,y) = c$$

is equivalent to optimizing the Lagrange function

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda \cdot (g(x,y) - c)$$

where λ is the new variable called Lagrange multiplier