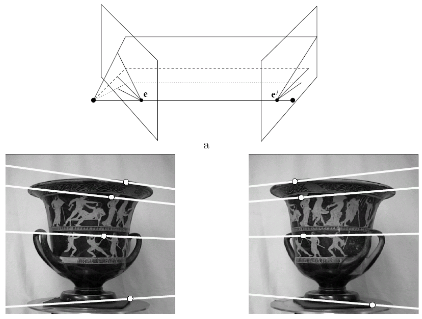
Multiple views



Hartley and Zisserman

stereo vision
structure from motion
optical flow

1

Why multiple views?

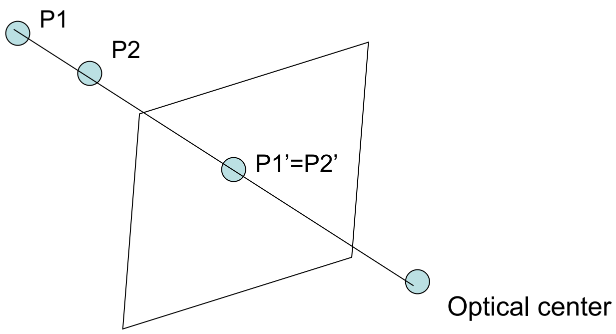
Images from Lana Lazebnik

2

3

Why multiple views?

- Structure and depth are inherently ambiguous from single views.

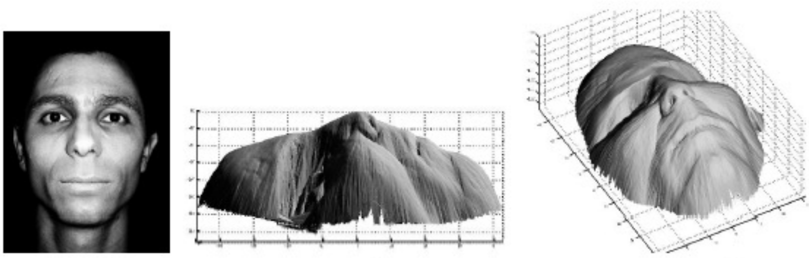


4

- What cues help us to perceive 3d shape and depth?

5

Shading

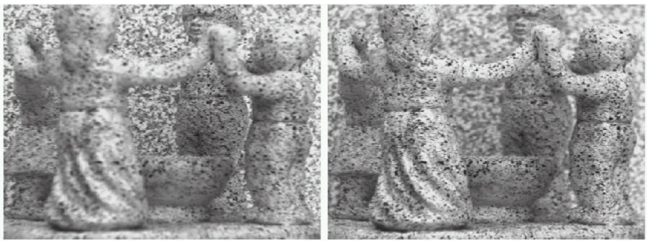


a) b) c)

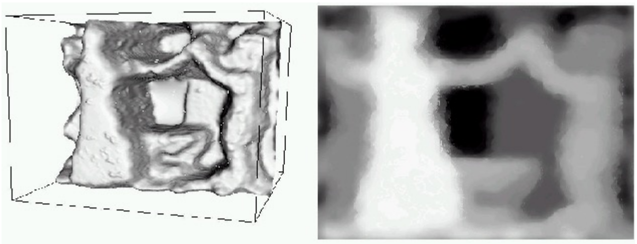
[Figure from Prados & Faugeras 2006]

6

Focus/defocus



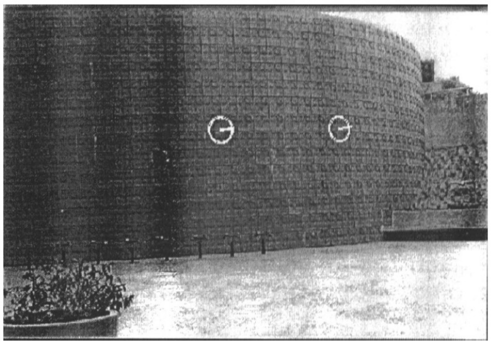
Images from same point of view, different camera parameters

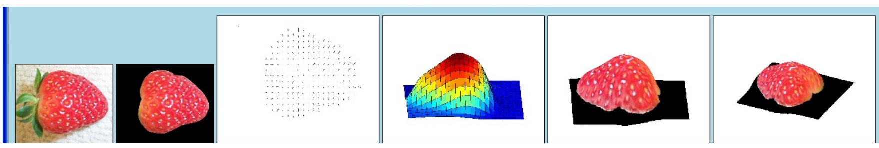


3d shape / depth estimates

[figs from H. Jin and P. Favaro, 2002]

Texture





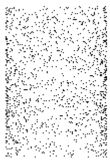
[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]

Perspective effects

Image credit: S. Seitz

9

Motion



Figures from L. Zhang

http://www.brainconnection.com/teasers/?main=illusion/motion-shape

10

Occlusion

Rene Magritt'e famous painting *Le Blanc-Seing* (literal translation: "The Blank Signature") roughly translates as "free hand". 1965

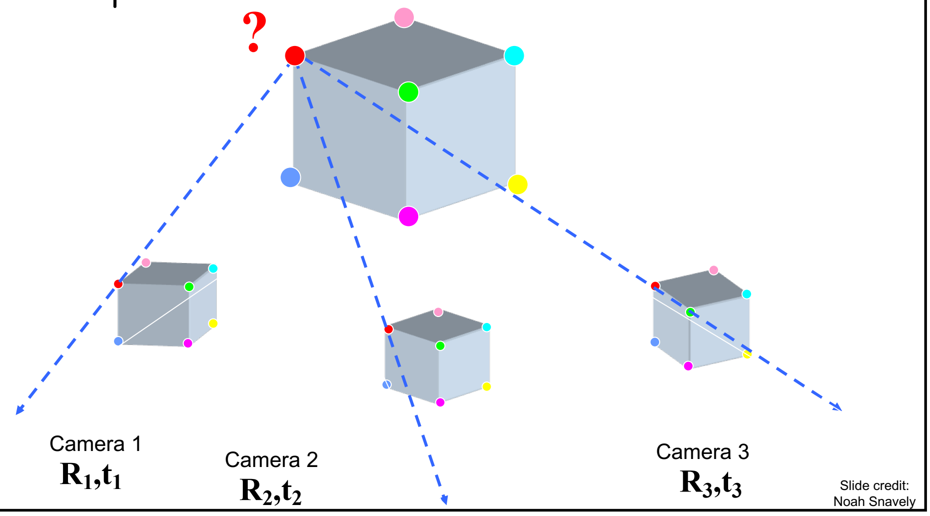
11

If stereo were critical for depth perception, navigation, recognition, etc., then this would be a problem

12

Multi-view geometry problems

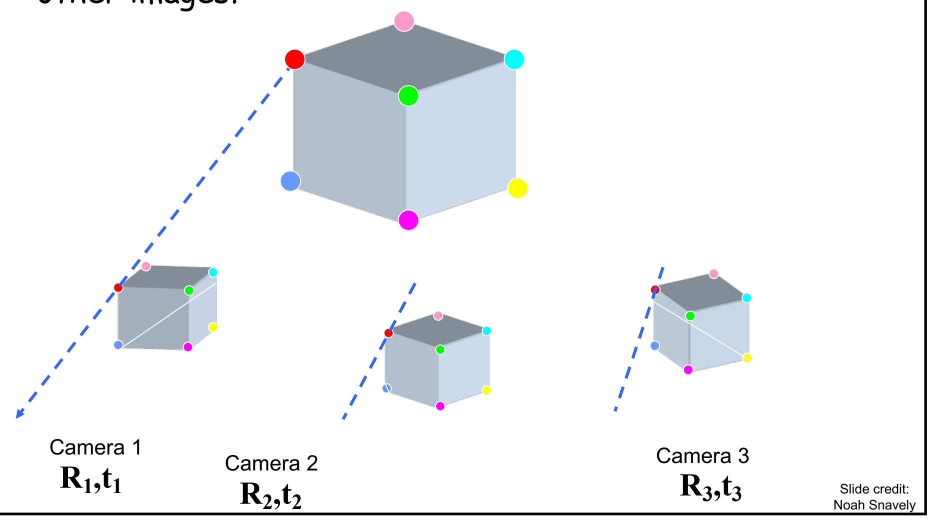
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



13

Multi-view geometry problems

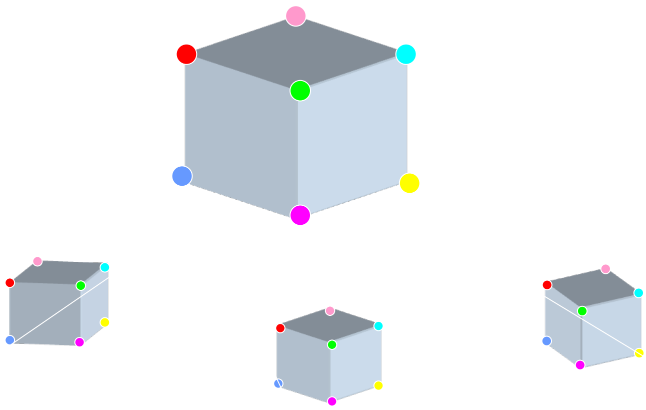
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



14

Multi-view geometry problems

- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters



Camera 1 $\mathbf{R}_1, \mathbf{t}_1$? Camera 2 $\mathbf{R}_2, \mathbf{t}_2$? ? Camera 3 $\mathbf{R}_3, \mathbf{t}_3$

Slide credit: Noah Snavely

15

Stereo photography and stereo viewers

Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.

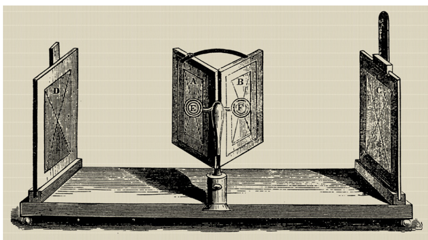
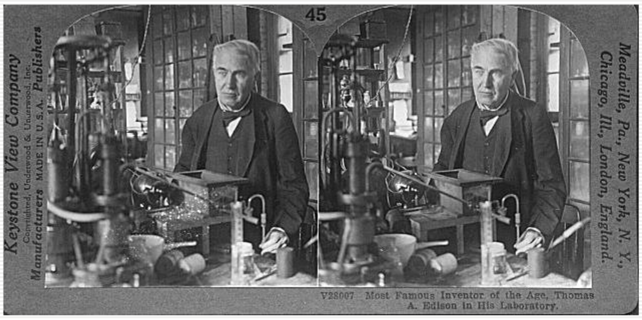


Image from fisher-price.com

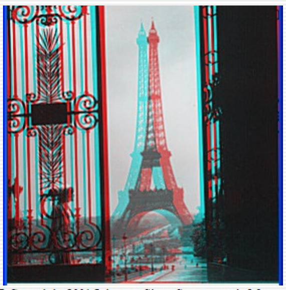
25

© Copyright 2001 Johnson-Shaw Stereoscopic Museum



http://www.johnsonshawmuseum.org

26

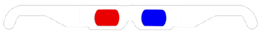


© Copyright 2001 Johnson-Shaw Stereoscopic Museum

http://www.johnsonshawmuseum.org

27

Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

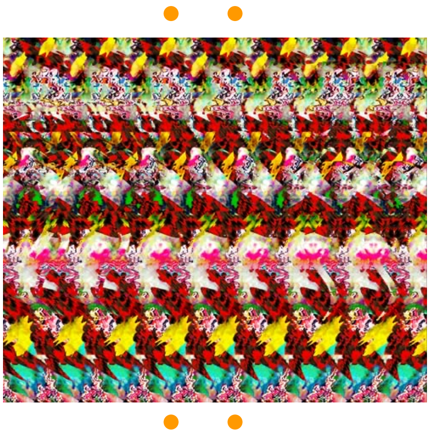


28

http://www.well.com/~jimg/stereo/stereo_list.html

29

Autostereograms



Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

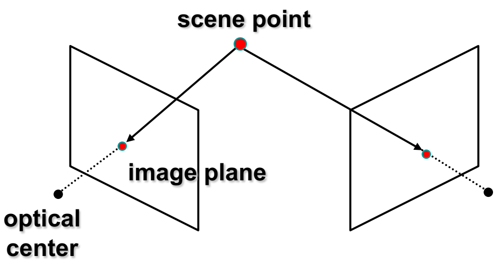
Images from magiceye.com

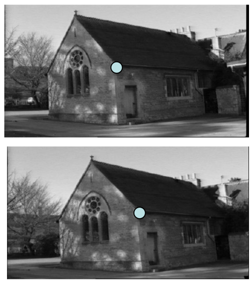
Autostereograms

Images from magiceye.com

Estimating depth with stereo

- **Stereo**: shape from "motion" between two views
- We'll need to consider:
  - Info on camera pose ("calibration")
  - Image point correspondences



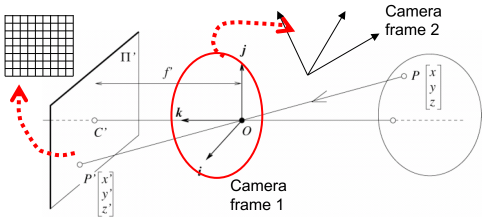


Stereo vision

Two cameras, simultaneous views

Single moving camera and static scene

Camera parameters



Extrinsic parameters:
Camera frame 1 ←→ Camera frame 2

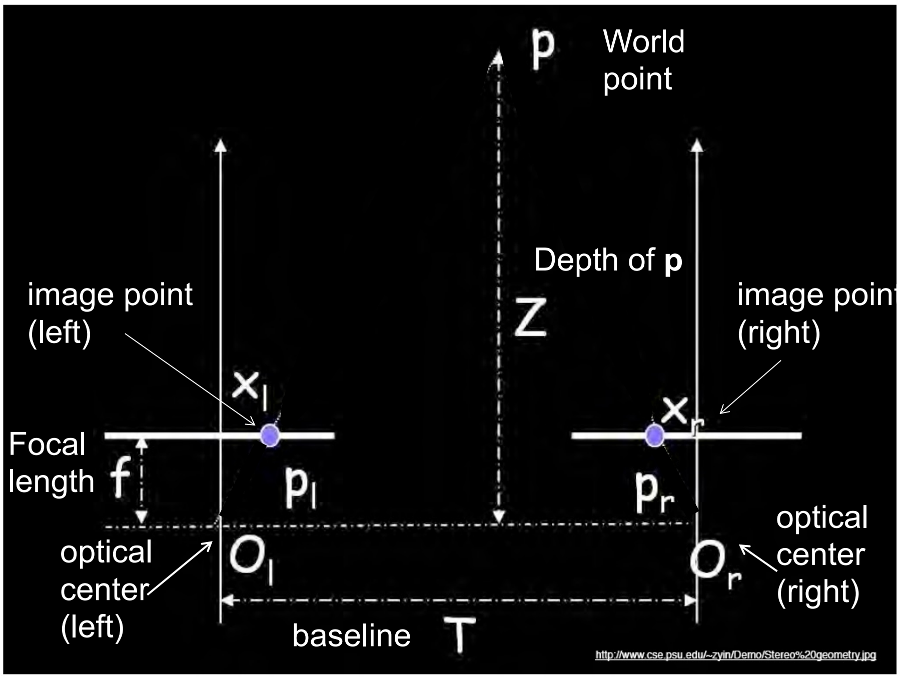
Intrinsic parameters:
Image coordinates relative to camera ←→ Pixel coordinates

- *Extrinsic* params: rotation matrix and translation vector
- *Intrinsic* params: focal length, pixel sizes (mm), image center point, radial distortion parameters

We'll assume for now that these parameters are given and fixed.

Geometry for a simple stereo system

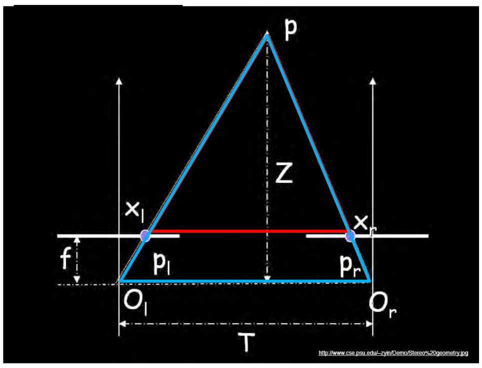
- First, assuming parallel optical axes, known camera parameters (i.e., calibrated cameras):



36

Geometry for a simple stereo system

- *Assume parallel optical axes, known camera parameters (i.e., calibrated cameras).* ***What is the expression for Z?***



Similar triangles (p_l, P, p_r) and (O_l, P, O_r):

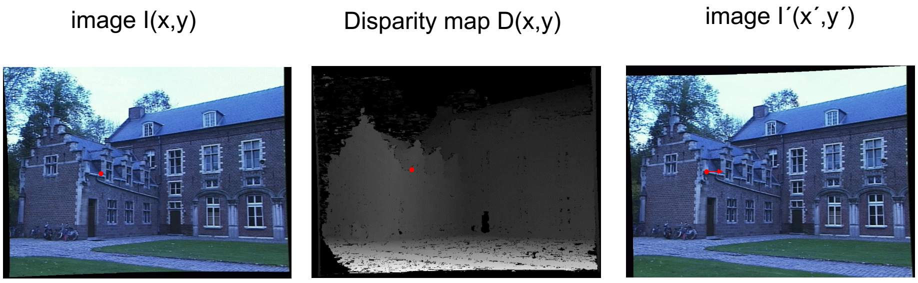
$$\frac{T - x_l + x_r}{Z - f} = \frac{T}{Z}$$

$$Z = f \frac{T}{x_r - x_l}$$

disparity

37

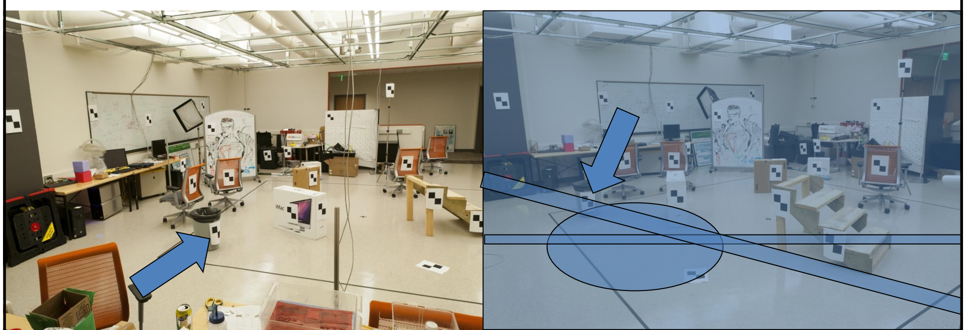
Depth from disparity



(x´,y´)=(x+D(x,y), y)

So if we could find the **corresponding points** in two images, we could **estimate relative depth**…

Where do we need to search?



To be continued…

Degrees of freedom

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

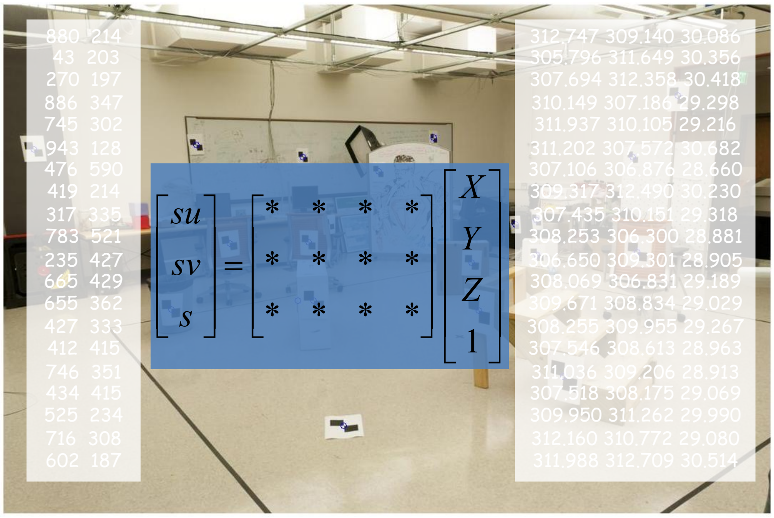
$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \overset{5}{\begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{6}{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

How to calibrate the camera?

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

How do we calibrate a camera?



Method 1 – homogeneous linear system

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Solve for m's entries using linear least squares

Ax=0 form

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 & -u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 & -v_1 \\ & & & & & & \vdots & & & & & \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_nX_n & -u_nY_n & -u_nZ_n & -u_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_nX_n & -v_nY_n & -v_nZ_n & -v_n \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

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

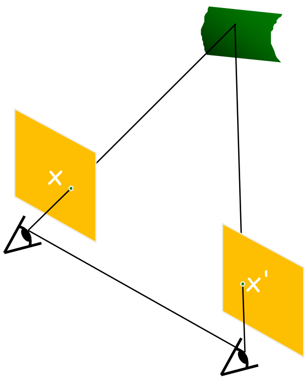
Epipolar Geometry and Stereo Vision

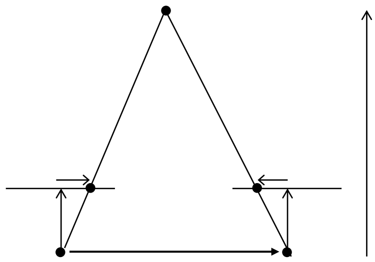
Chapter 12 in Szeliski

- Epipolar geometry
  - Relates cameras from two positions

Depth from Stereo

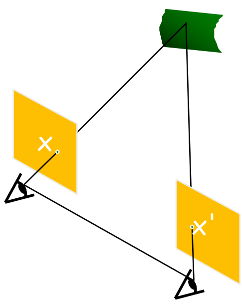
- Goal: recover depth by finding image coordinate x' that corresponds to x



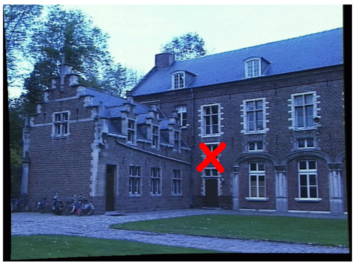


Depth from Stereo

- Goal: recover depth by finding image coordinate x' that corresponds to x
- Sub-Problems
  1. Calibration: How do we recover the relation of the cameras (if not already known)?
  2. Correspondence: How do we search for the matching point x'?

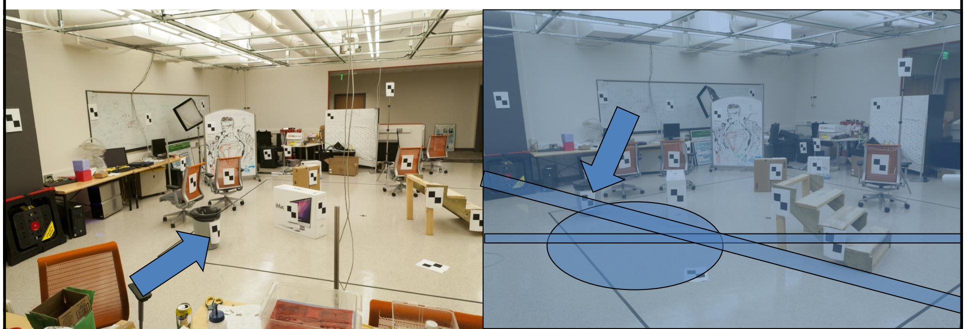


Correspondence Problem



- We have two images taken from cameras with different intrinsic and extrinsic parameters
- How do we match a point in the first image to a point in the second? How can we constrain our search?

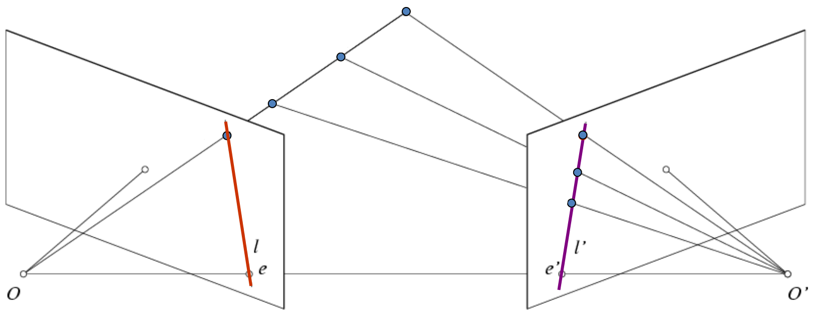
Where do we need to search?



Key idea: Epipolar constraint

50

Key idea: Epipolar constraint

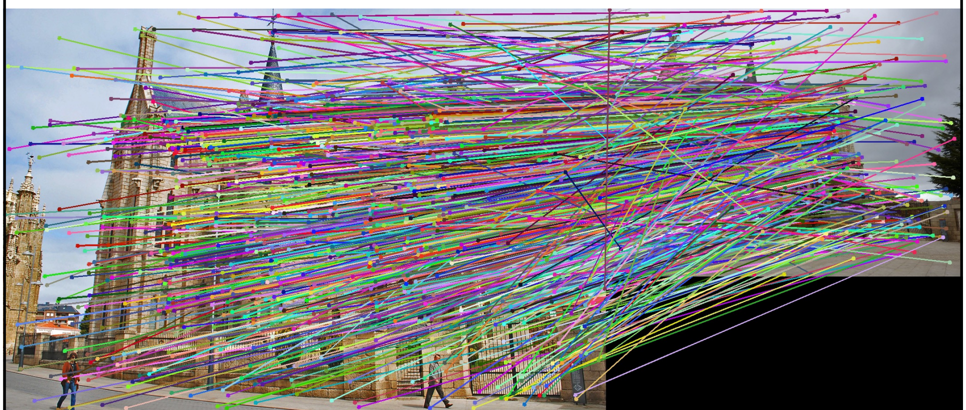


51

Wouldn't it be nice to know where matches can live? To constrain our 2d search to 1d.

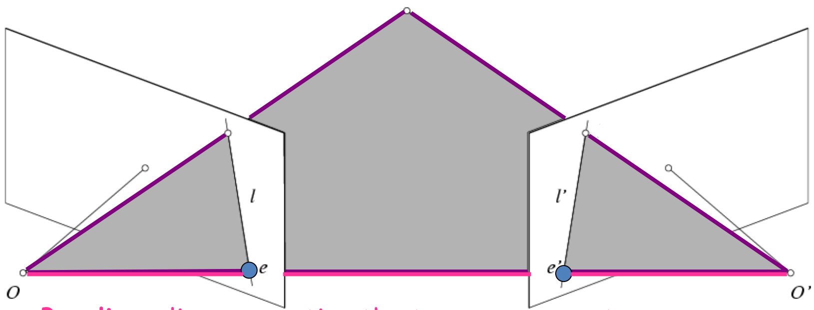
52

VLFeat's 800 most confident matches among 10,000+ local features.



53

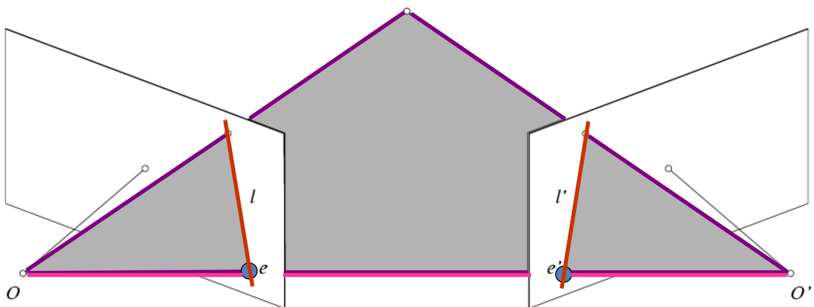
Epipolar geometry: notation



- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)

54

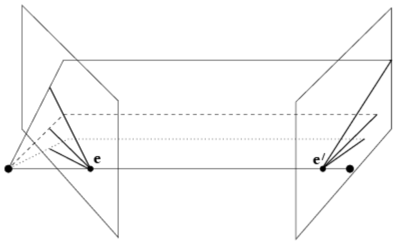
Epipolar geometry: notation

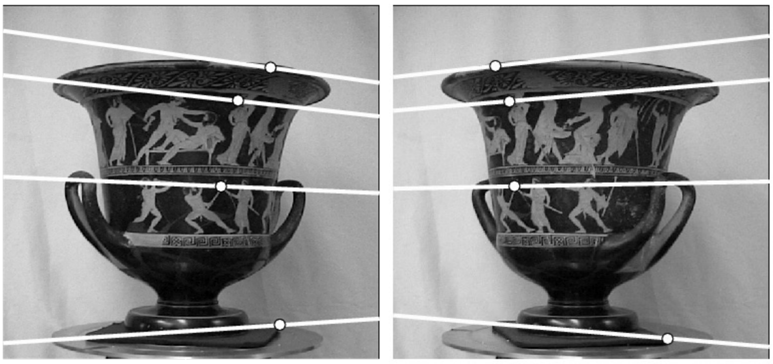


- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

55

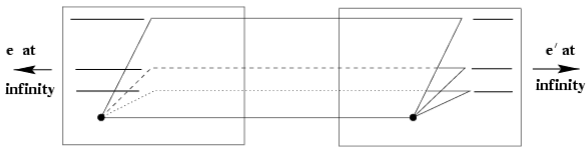
Example: Converging cameras

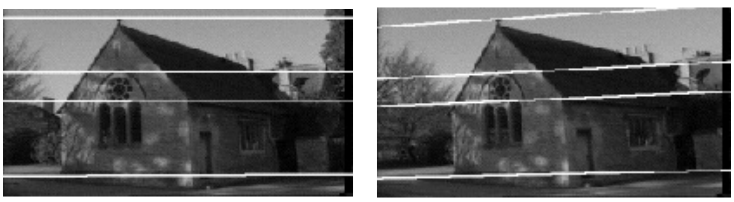




56

Example: Motion parallel to image plane



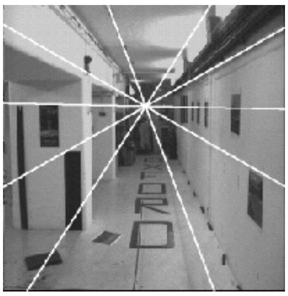


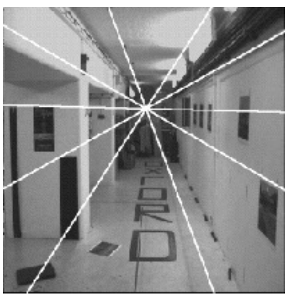
57

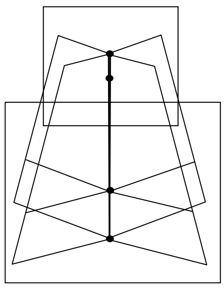
Example: Forward motion

What would the epipolar lines look like if the camera moves directly forward?

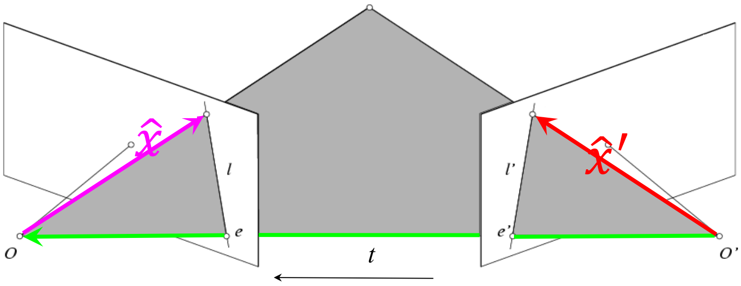
Example: Forward motion







Epipolar constraint: Calibrated case

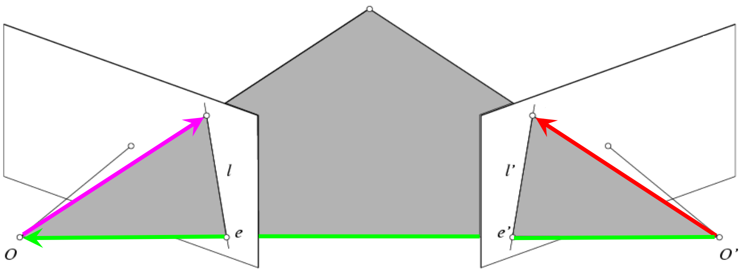


$$\hat{x} = K^{-1}x = X \qquad \hat{x}' = K'^{-1}x' = X'$$

$$\hat{x} \cdot [t \times (R\hat{x}')] = 0$$

60

Essential matrix



$$\hat{x} \cdot [t \times (R\hat{x}')] = 0 \quad \Rightarrow \quad \hat{x}^T E \hat{x}' = 0 \quad \text{with} \quad E = [t]_\times R$$

Essential Matrix
(Longuet-Higgins, 1981)

61

The Fundamental Matrix

Without knowing K and K', we can define a similar relation using *unknown* normalized coordinates

$$\hat{x}^T E \hat{x}' = 0$$

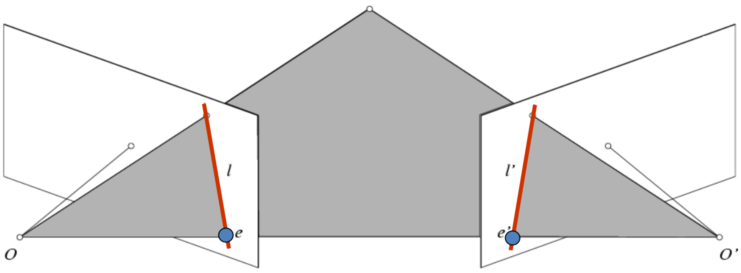
$$\hat{x} = K^{-1} x$$

$$\hat{x}' = K'^{-1} x'$$

$$\Rightarrow \quad x^T F x' = 0 \quad \text{with} \quad F = K^{-T} E K'^{-1}$$

Fundamental Matrix
(Faugeras and Luong, 1992)

Properties of the Fundamental matrix



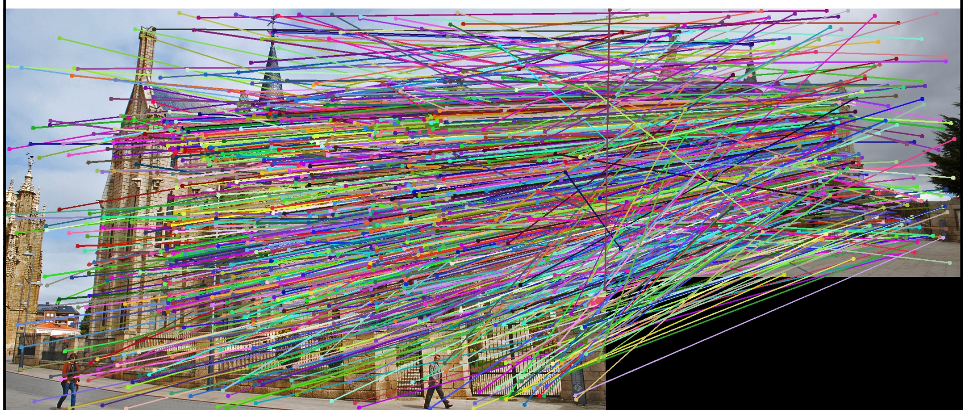
$$x^T F x' = 0 \quad \text{with} \quad F = K^{-T} E K'^{-1}$$

- $F x' = 0$ is the epipolar line associated with x'
- $F^T x = 0$ is the epipolar line associated with x
- $F e' = 0$ and $F^T e = 0$
- F is singular (rank two): det(F)=0
- F has seven degrees of freedom: 9 entries but defined up to scale, det(F)=0

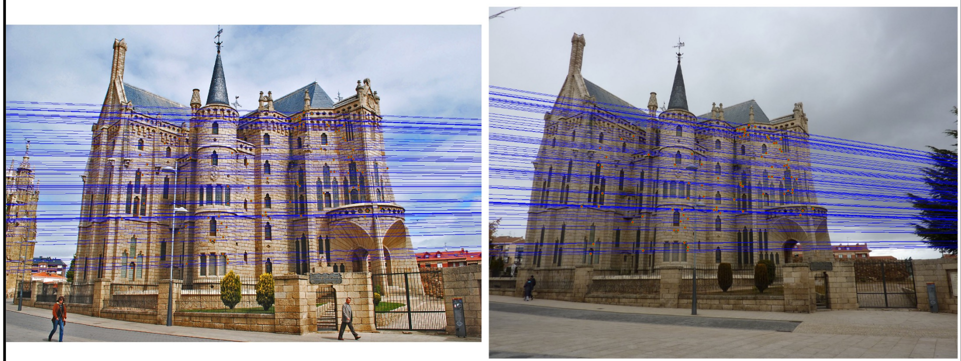
Estimating the Fundamental Matrix

- 8-point algorithm
  - Least squares solution using SVD on equations from 8 pairs of correspondences
  - Enforce det(F)=0 constraint using SVD on F
- 7-point algorithm
  - Use least squares to solve for null space (two vectors) using SVD and 7 pairs of correspondences
  - Solve for linear combination of null space vectors that satisfies det(F)=0
- Minimize reprojection error
  - Non-linear least squares

VLFeat's 800 most confident matches among 10,000+ local features.

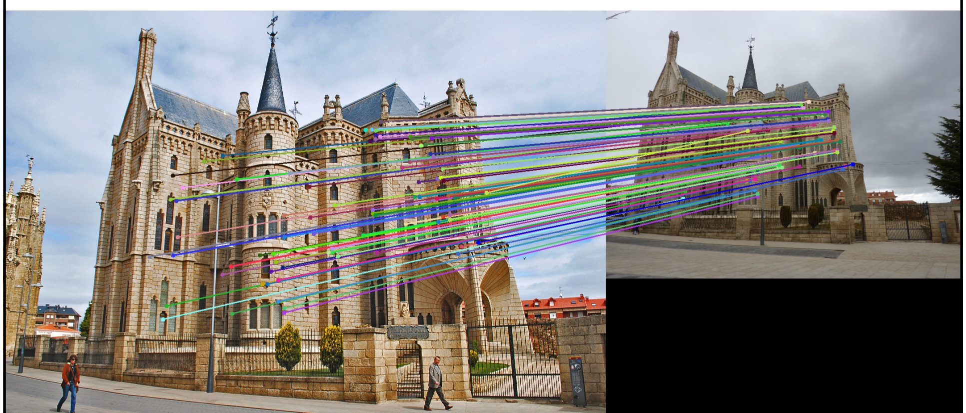


Epipolar lines



66

Keep only the matches at are "inliers" with respect to the "best" fundamental matrix



67