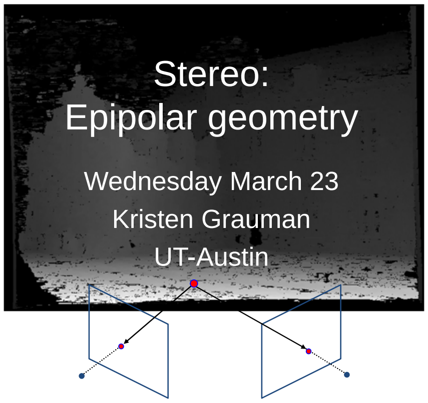


Announcements

- Reminder: Pset 3 due next Wed, March 30

Last time

- Image formation affected by geometry, photometry, and optics.
- Projection equations express how world points mapped to 2d image.
- Parameters (focal length, aperture, lens diameter,…) affect image obtained.

Review

- How do the perspective projection equations explain this effect?

http://www.mzephotos.com/gallery/mammals/rabbit-nose.html

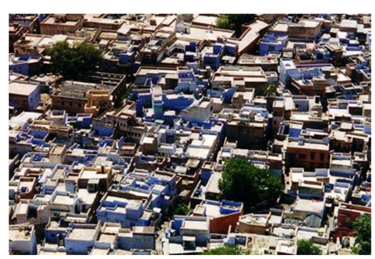
flickr.com/photos/lungstruck/434631076/

Miniature faking

In close-up photo, the depth of field is limited.

http://en.wikipedia.org/wiki/File:Jodhpur_tilt_shift.jpg

Miniature faking

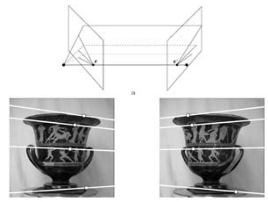


Miniature faking

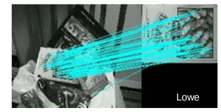
http://en.wikipedia.org/wiki/File:Oregon_State_Beavers_Tilt-Shift_Miniature_Greg_Keene.jpg

Multiple views

Multi-view geometry, matching, invariant features, stereo vision



Hartley and Zisserman



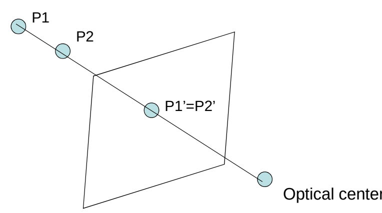
Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Images from Lana Lazebnik

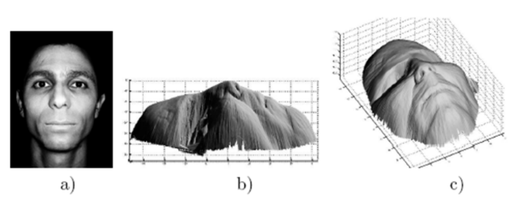
Why multiple views?

- Structure and depth are inherently ambiguous from single views.



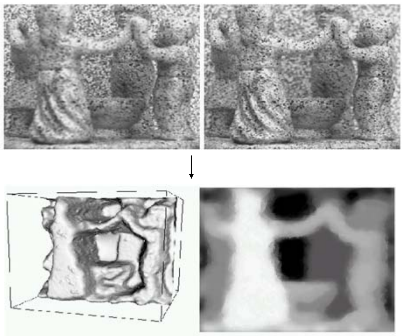
- What cues help us to perceive 3d shape and depth?

Shading



[Figure from Prados & Faugeras 2006]

Focus/defocus

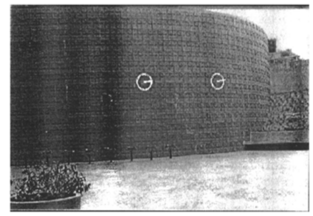


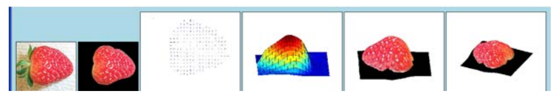
Images from same point of view, different camera parameters

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

Motion

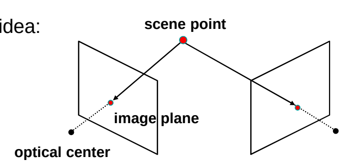
Figures from L. Zhang

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

Estimating scene shape

- “Shape from X”: Shading, Texture, Focus, Motion…
- **Stereo**:
  - shape from “motion” between two views
  - infer 3d shape of scene from two (multiple) images from different viewpoints

Main idea:

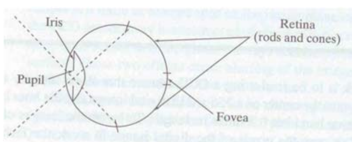


Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras

Human eye

Rough analogy with human visual system:



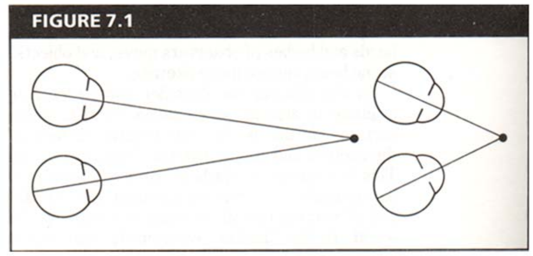
Pupil/Iris – control amount of light passing through lens

Retina - contains sensor cells, where image is formed

Fovea – highest concentration of cones

Fig from Shapiro and Stockman

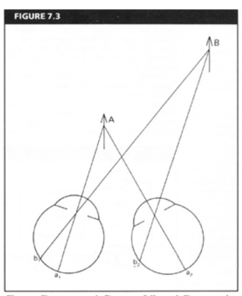
Human stereopsis: disparity



From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

Human eyes **fixate** on point in space – rotate so that corresponding images form in centers of fovea.

Human stereopsis: disparity

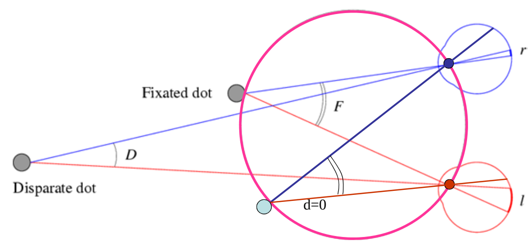


From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

Disparity occurs when eyes fixate on one object; others appear at different visual angles

Adapted from David Forsyth, UC Berkeley

Human stereopsis: disparity



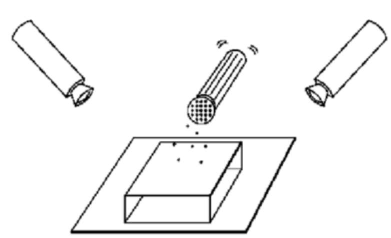
Disparity: $d = r\text{-}l = D\text{-}F$.

Forsyth & Ponce

Random dot stereograms

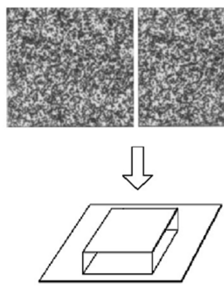
- Julesz 1960: Do we identify local brightness patterns before fusion (monocular process) or after (binocular)?
- To test: pair of synthetic images obtained by randomly spraying black dots on white objects

Random dot stereograms



Forsyth & Ponce

Random dot stereograms



Random dot stereograms

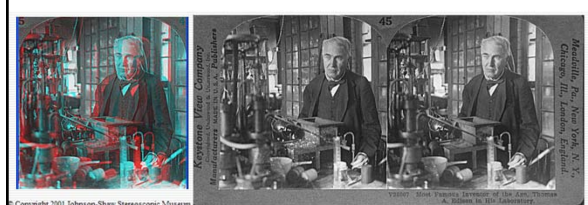
- When viewed monocularly, they appear random; when viewed stereoscopically, see 3d structure.
- Conclusion: human binocular fusion not directly associated with the physical retinas; must involve the central nervous system
- Imaginary "*cyclopean retina*" that combines the left and right image stimuli as a single unit

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.

Invented by Sir Charles Wheatstone, 1838

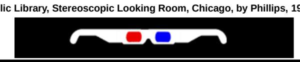
Image from fisher-price.com



http://www.johnsonshawmuseum.org

http://www.johnsonshawmuseum.org

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



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

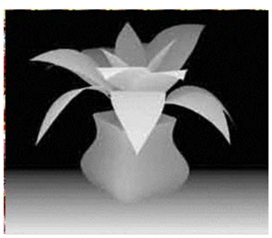
Autostereograms

Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com

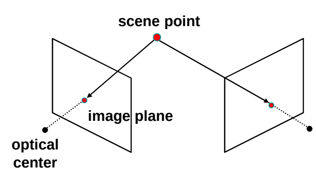
Autostereograms



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

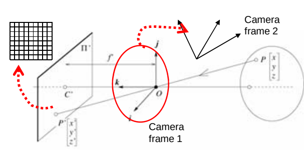


Stereo vision

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

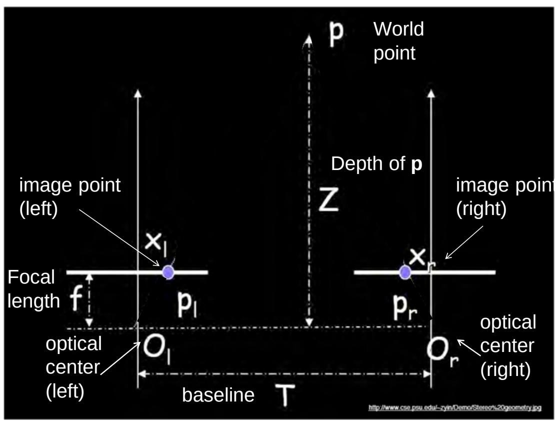
We'll assume for now that these parameters are given and fixed.

Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras

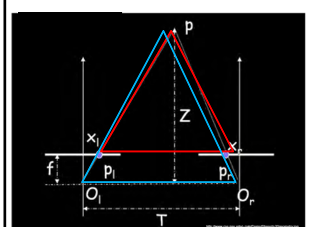
Geometry for a simple stereo system

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



Geometry for a simple stereo system

- Assume parallel optical axes, known camera parameters (i.e., calibrated cameras). **What is 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

Depth from disparity

image I(x,y)

Disparity map D(x,y)

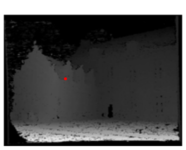


image I´(x´,y´)

(x´,y´)=(x+D(x,y), y)

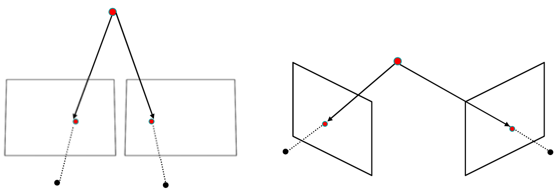
So if we could find the **corresponding points** in two images, we could **estimate relative depth**…

Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras

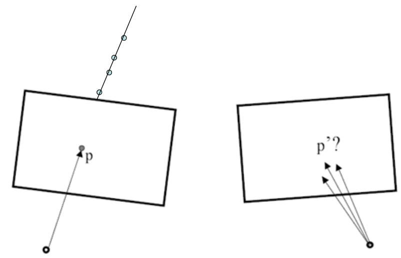
General case, with calibrated cameras

- The two cameras need not have parallel optical axes.



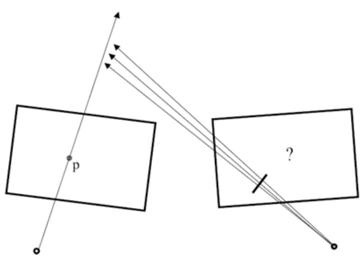
Vs.

Stereo correspondence constraints

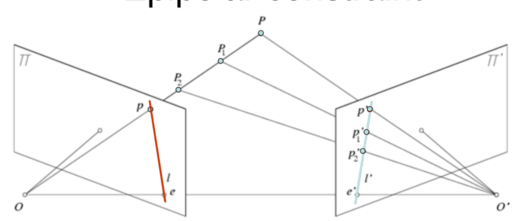


- Given p in left image, where can corresponding point p' be?

Stereo correspondence constraints



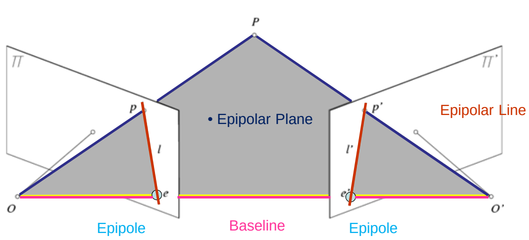
Epipolar constraint



Geometry of two views constrains where the corresponding pixel for some image point in the first view must occur in the second view.

- It must be on the line carved out by a plane connecting the world point and optical centers.

Epipolar geometry



http://www.ai.sri.com/~luong/research/Meta3DViewer/EpipolarGeo.html

Epipolar geometry: terms

- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with 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

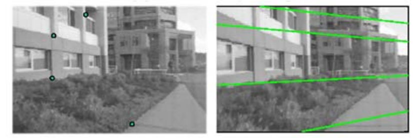
Why is the epipolar constraint useful?

Epipolar constraint

This is useful because it reduces the correspondence problem to a 1D search along an epipolar line.

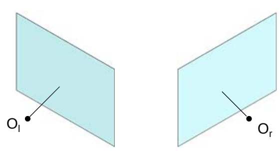
Image from Andrew Zisserman

Example

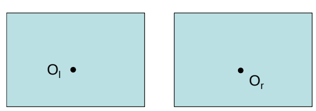


What do the epipolar lines look like?

1.



2.



Example: converging cameras

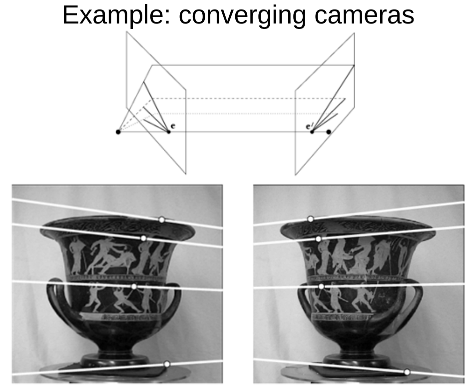
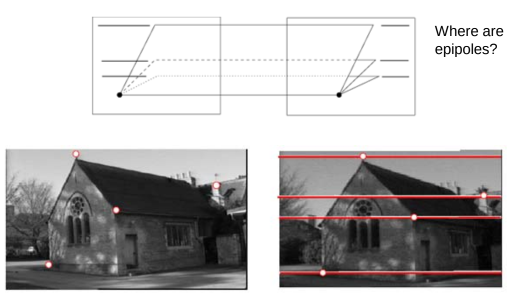


Figure from Hartley & Zisserman

Example: parallel cameras

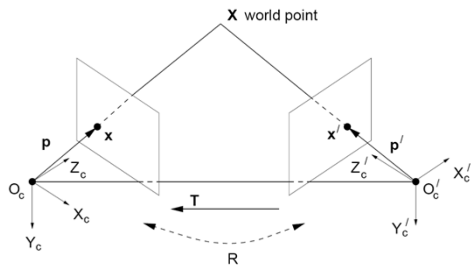


Where are the epipoles?

Figure from Hartley & Zisserman

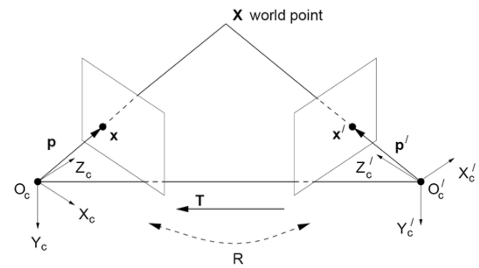
- So far, we have the explanation in terms of geometry.
- Now, how to express the epipolar constraints algebraically?

Stereo geometry, with calibrated cameras



Main idea

Stereo geometry, with calibrated cameras

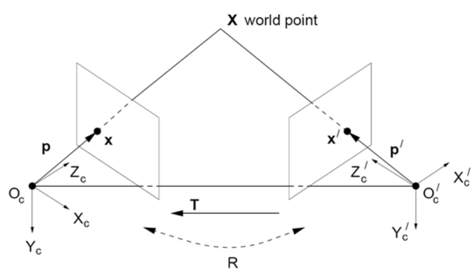


If the stereo rig is calibrated, we know :

how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2.

Rotation: 3 x 3 matrix **R**; translation: 3 vector **T**.

Stereo geometry, with calibrated cameras



If the stereo rig is calibrated, we know :

how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2. $\mathbf{X}'_c = \mathbf{R}\mathbf{X}_c + \mathbf{T}$

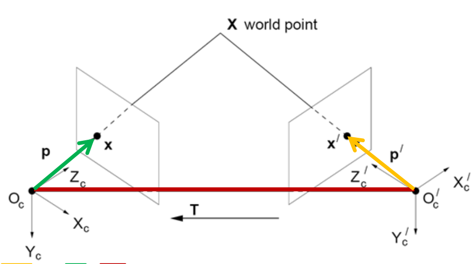
An aside: cross product

$$\vec{a} \times \vec{b} = \vec{c} \qquad \begin{aligned} \vec{a} \cdot \vec{c} &= 0 \\ \vec{b} \cdot \vec{c} &= 0 \end{aligned}$$

Vector cross product takes two vectors and returns a third vector that's perpendicular to both inputs.

So here, c is perpendicular to both a and b, which means the dot product = 0.

From geometry to algebra



$\mathbf{X}' = \mathbf{R}\mathbf{X} + \mathbf{T}$

$\mathbf{T} \times \mathbf{X}' =$ (Normal to the plane)

$= \mathbf{T} \times \mathbf{R}\mathbf{X}$

$$\mathbf{X}' \cdot (\mathbf{T} \times \mathbf{X}') = \mathbf{X}' \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X}) = 0$$

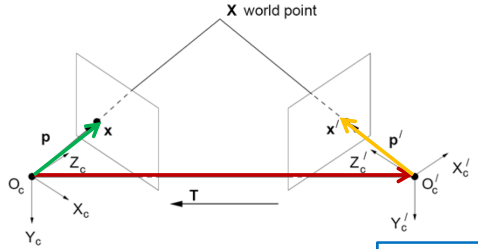
Another aside: Matrix form of cross product

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{c} \qquad \begin{aligned} \vec{a} \cdot \vec{c} &= 0 \\ \vec{b} \cdot \vec{c} &= 0 \end{aligned}$$

Can be expressed as a matrix multiplication.

$$[a_x] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \qquad \boxed{\vec{a} \times \vec{b} = [a_x]\vec{b}}$$

From geometry to algebra



$$\mathbf{X}' = \mathbf{R}\mathbf{X} + \mathbf{T}$$

$$\mathbf{T} \times \mathbf{X}' = \mathbf{T} \times \mathbf{R}\mathbf{X} + \mathbf{T} \times \mathbf{T}$$

Normal to the plane

$$= \mathbf{T} \times \mathbf{R}\mathbf{X}$$

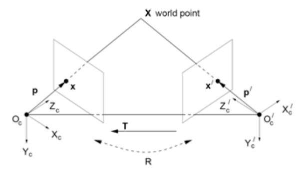
$$\mathbf{X}' \cdot (\mathbf{T} \times \mathbf{X}') = \mathbf{X}' \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X})$$

$$= 0$$

Essential matrix

$$\mathbf{X}' \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X}) = 0$$

$$\mathbf{X}' \cdot ([\mathrm{T}_x]\mathbf{R}\mathbf{X}) = 0$$



Let $\mathbf{E} = [\mathrm{T}_x]\mathbf{R}$

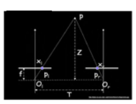
$$\mathbf{X}'^T \mathbf{E}\mathbf{X} = 0$$

E is called the **essential matrix**, and it relates corresponding image points between both cameras, given the rotation and translation.

If we observe a point in one image, its position in other image is constrained to lie on line defined by above.

Note: these points are in **camera coordinate systems**.

Essential matrix example: parallel cameras



$\mathbf{R} =$

$\mathbf{T} =$

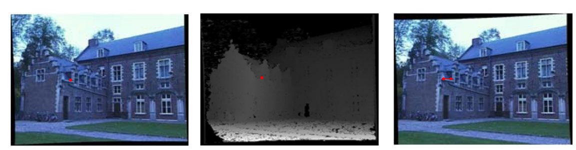
$\mathbf{E} = [\mathrm{T}_x]\mathbf{R} =$

$\mathbf{p} = [x, y, f]$

$\mathbf{p}' = [x', y', f]$

$$\mathbf{p}'^{T}\mathbf{E}\mathbf{p} = 0$$

For the parallel cameras, image of any point must lie on same horizontal line in each image plane.

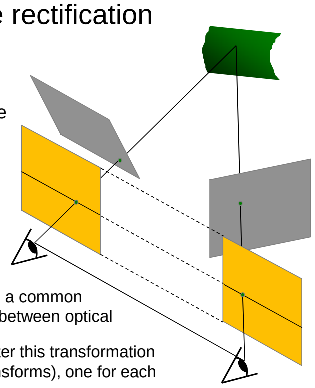


(x´,y´)=(x+D(x,y),y)

What about when cameras' optical axes are not parallel?

Stereo image rectification

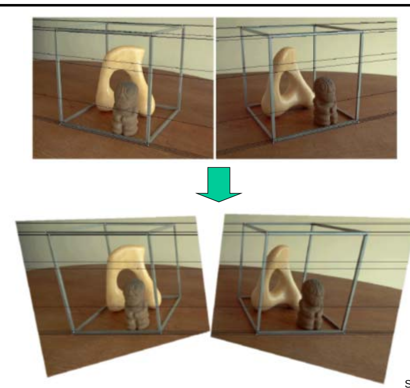
In practice, it is convenient if image scanlines (rows) are the epipolar lines.



- reproject image planes onto a common plane parallel to the line between optical centers
- pixel motion is horizontal after this transformation
- two homographies (3x3 transforms), one for each input image reprojection

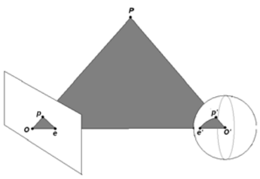
Slide credit: Li Zhang

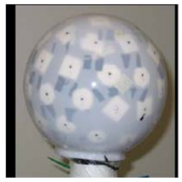
Stereo image rectification: example



Source: Alyosha Efros

An audio camera & epipolar geometry





Spherical microphone array

Adam O' Donovan, Ramani Duraiswami and Jan Neumann
Microphone Arrays as Generalized Cameras for Integrated Audio Visual Processing, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, 2007

An audio camera & epipolar geometry

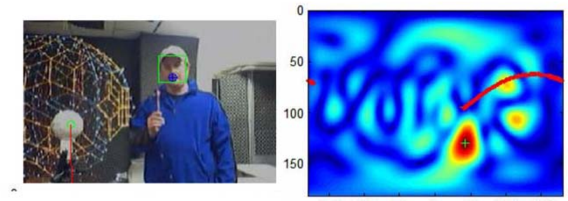


Figure 4. An example of the use of the system in speaker tracking with noise suppression. The bright red spot on the sound image (marked with a +) corresponds to the dominant source. The less dominant source however lies on the epipolar line in the sound image induced by the location of the mouth in the camera image, and this source is beamformed.

Summary

- Depth from stereo: main idea is to triangulate from corresponding image points.
- Epipolar geometry defined by two cameras
  - We've assumed known extrinsic parameters relating their poses
- Epipolar constraint limits where points from one view will be imaged in the other
  - Makes search for correspondences quicker

- **Terms**: epipole, epipolar plane / lines, disparity, rectification, intrinsic/extrinsic parameters, essential matrix, baseline

Coming up

– Computing correspondences
– Non-geometric stereo constraints
– Weak calibration