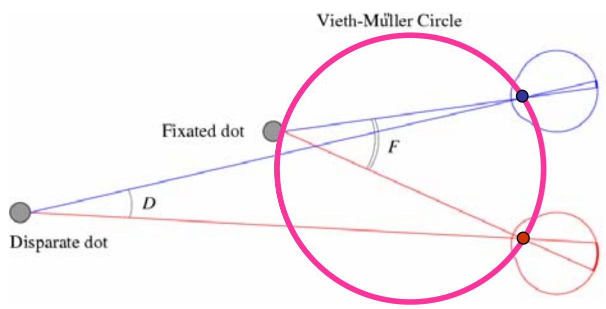
Lecture 11: Stereo II

Thursday, Oct 4

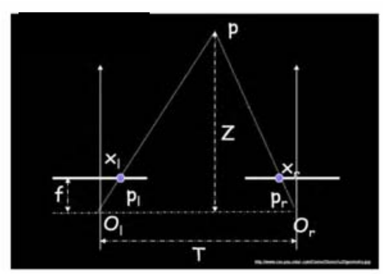
CS 378/395T
Prof. Kristen Grauman

Last time: Disparity

- Disparity: difference in retinal position of same item



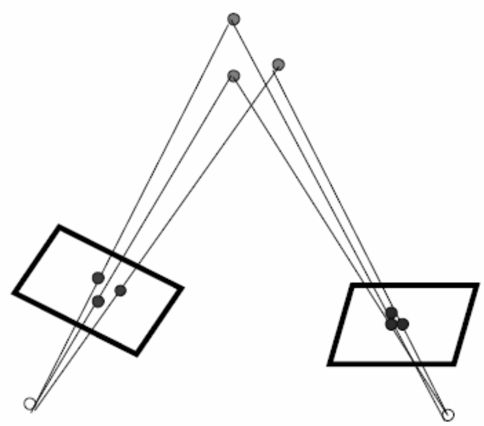
- Case of stereo rig for parallel image planes and calibrated cameras: depth (Z) is inversely related to disparity (xr-xl).



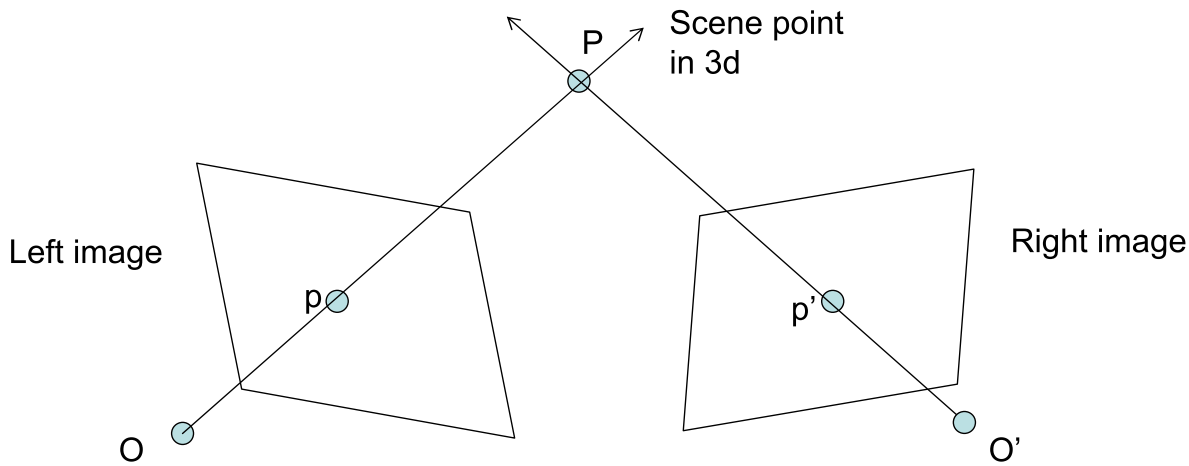
Last time: Multi-view geometry

Relate

- 3-D points
- Camera centers
- Camera orientation
- Camera intrinsics



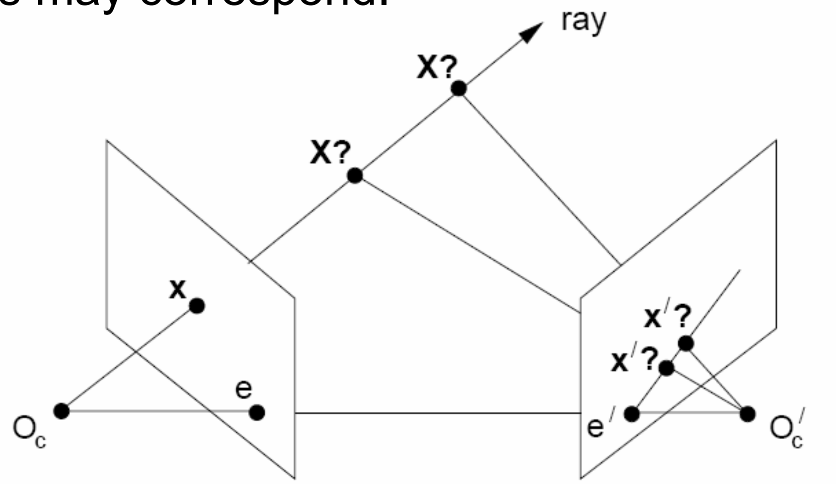
Last time: Triangulation



Estimate scene point based on camera relationships and *correspondence*.

Last time: Epipolar geometry

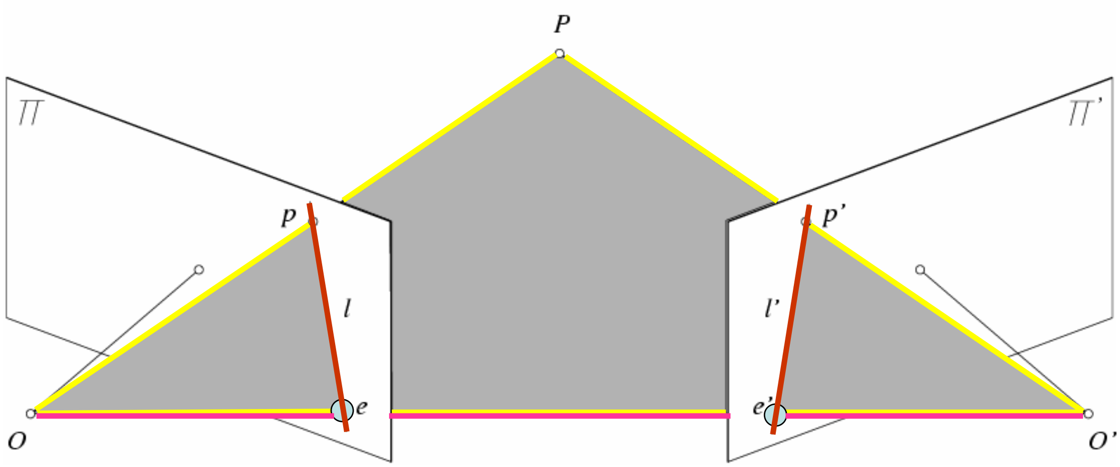
Key idea: geometry imposes constraints on which points may correspond.



If a point feature **x** is observed in one image, its location **x'** in the other image must lie on the epipolar line.

Figure from Gee & Cipolla 1999

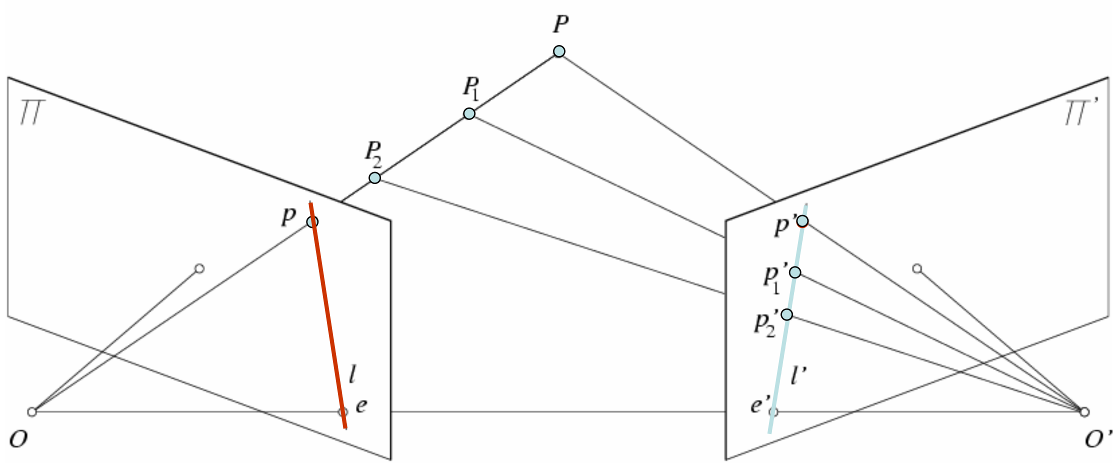
Last time: Epipolar geometry



- Epipolar Plane
- Baseline
- Epipoles
- Epipolar Lines

Adapted from M. Pollefeys, UNC

Last time: Epipolar constraint



- Potential matches for *p* have to lie on the corresponding epipolar line *l'*.
- Potential matches for *p'* have to lie on the corresponding epipolar line *l*.

Slide credit: M. Pollefeys, UNC

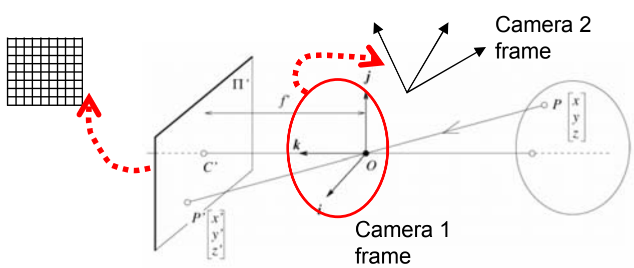
Today

- How do we compute those epipolar lines?
- How do we relate corresponding points algebraically?
  - Essential matrix
- What other constraints can we use besides geometry?

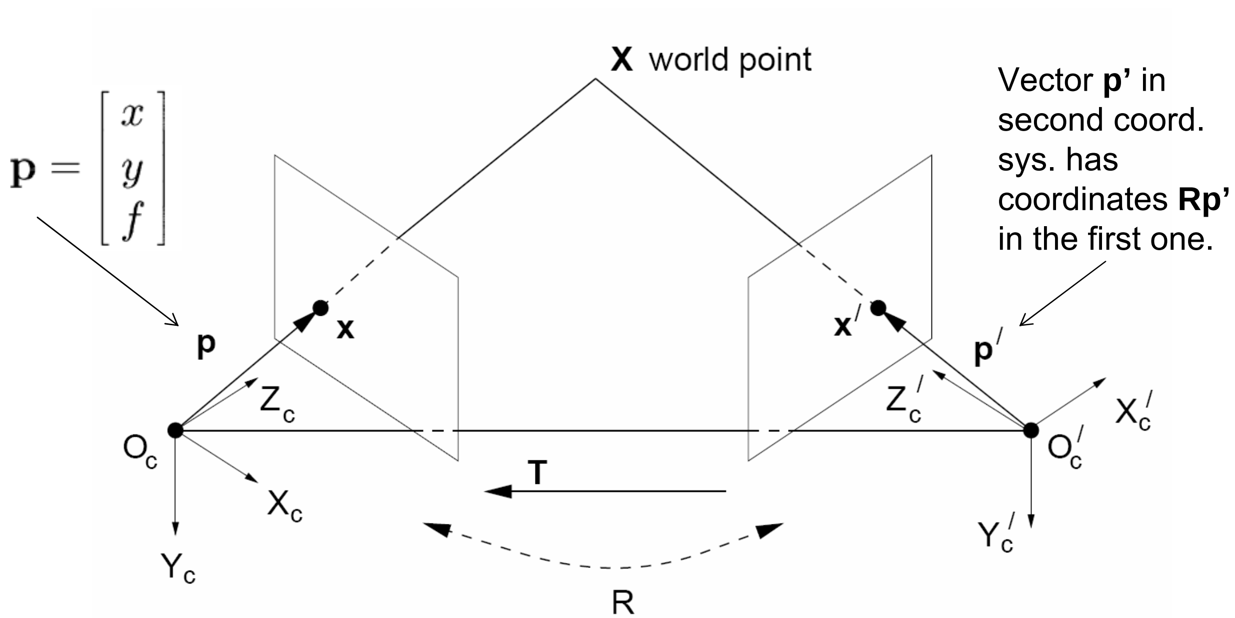
- Still assuming *calibrated cameras* for now.

Calibrated cameras

- If fully calibrated, we know
  - how to rotate and translate camera reference frame 1 to get to camera reference frame 2.
  - how to map pixel coordinates to image plane coordinates



Stereo geometry, with calibrated cameras



Camera-centered coordinate systems are related by known rotation **R** and translation **T**.

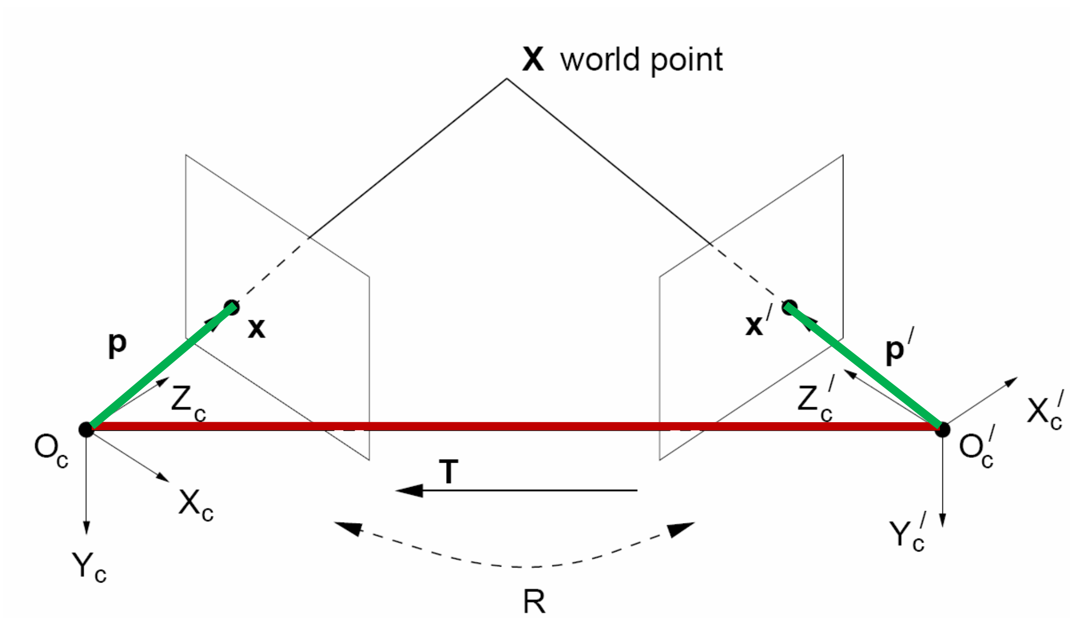
Recall: Cross product

$$\vec{a} \times \vec{b} = \vec{c}$$

$$\vec{a} \cdot \vec{c} = 0$$
$$\vec{b} \cdot \vec{c} = 0$$

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

From geometry to algebra



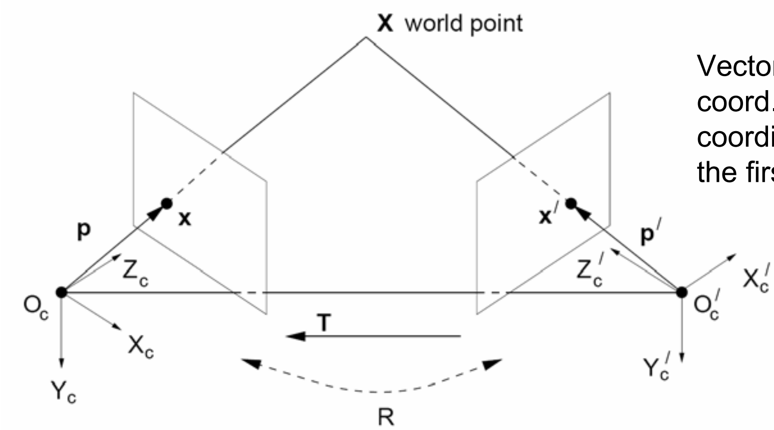
$$\overrightarrow{Op} \cdot \overrightarrow{OO'} \quad \overrightarrow{O'p'}$$

Coplanar vectors

Also coplanar, so dot product with normal is 0

Normal to this plane

From geometry to algebra



Vector **p'** in second coord. sys. has coordinates **Rp'** in the first one.

$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$

$$p \cdot \left[T \times \left(Rp'\right)\right] = 0$$

Matrix form of cross product

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c} \qquad \begin{matrix} \vec{a} \cdot \vec{c} = 0 \\ \vec{b} \cdot \vec{c} = 0 \end{matrix}$$

Can be expressed as a matrix multiplication.

$$[a_x] = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

$$\boxed{\vec{a} \times \vec{b} = [a_x]\vec{b}}$$

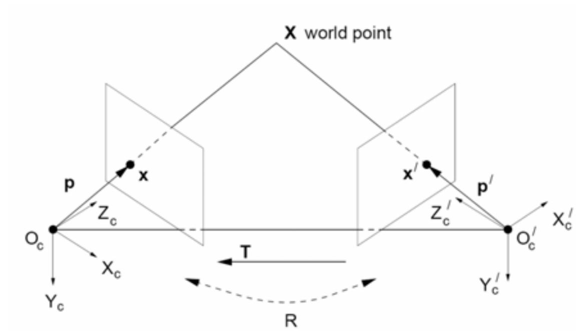
From geometry to algebra

$$\mathbf{p} \cdot [\mathbf{T} \times (\mathbf{R}\mathbf{p}')] = 0$$

$$\mathbf{p} \cdot [\mathbf{T}_{\mathbf{x}}]\mathbf{R}\mathbf{p}' = 0$$

Let $$\mathbf{E} = [\mathbf{T}_{\mathbf{x}}]\mathbf{R}$$

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



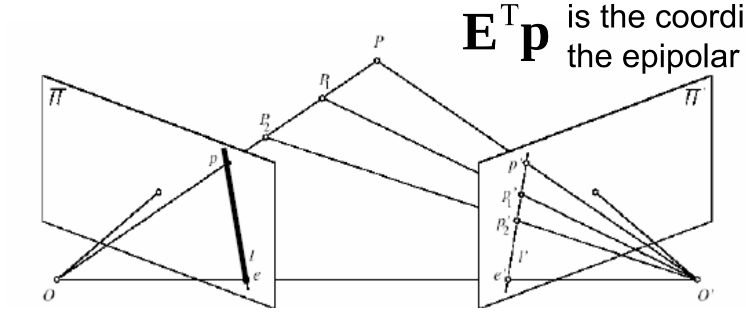
E is the **essential matrix**, which relates corresponding image points [Longuet-Higgins 1981]

Essential matrix and epipolar lines

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

Epipolar constraint: if we observe point **p** in one image, then its position **p'** in second image must satisfy this equation.

$\mathbf{E}^{\mathrm{T}}\mathbf{p}$ is the coordinate vector representing the epipolar line for point p

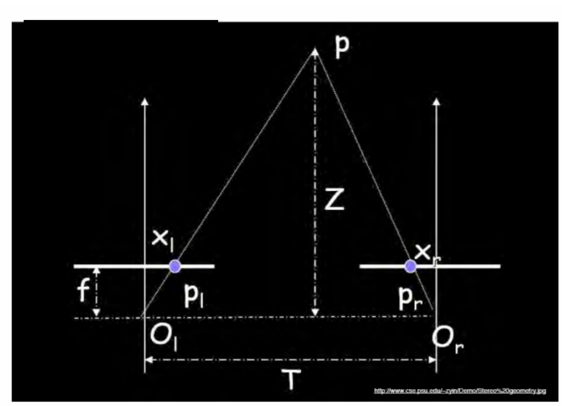


$\mathbf{E}\mathbf{p}'$ is the coordinate vector representing the epipolar line for point p'

Essential matrix properties

- Relates image of corresponding points in both cameras, given rotation and translation
- Assuming intrinsic parameters are known

Essential matrix example: parallel cameras

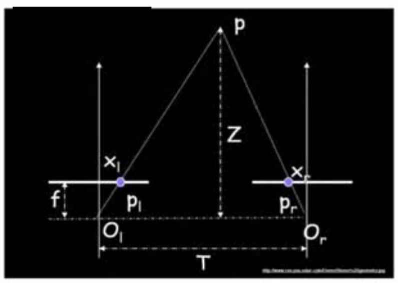


$$\mathbf{R} = \mathbf{I}$$

$$\mathbf{T} = [-T, 0, 0]^{\mathrm{T}}$$

$$\mathbf{E} = [\mathbf{T}_{\mathbf{x}}]\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & T \\ 0 & -T & 0 \end{pmatrix}$$

Essential matrix example: parallel cameras



$$\mathbf{R} = \mathbf{I}$$

$$\mathbf{T} = [-T, 0, 0]^{\mathrm{T}}$$

$$\mathbf{E} = [\mathbf{T_x}]\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & T \\ 0 & -T & 0 \end{pmatrix}$$

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

$$\begin{bmatrix} x & y & f \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & T \\ 0 & -T & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ f \end{bmatrix} = 0$$

$$\Leftrightarrow \begin{bmatrix} x & y & f \end{bmatrix} \begin{bmatrix} 0 \\ Tf \\ -Ty' \end{bmatrix} = 0$$

Image of any point must lie on same horizontal line in each image plane!

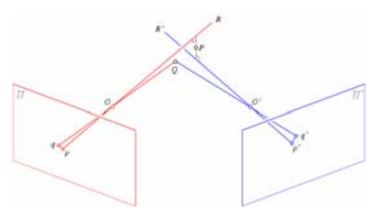
$$\Leftrightarrow y = y'$$

Stereo reconstruction for fully calibrated cameras

- Image pair
- Detect some features
- Compute E from R and T
- Match features using the epipolar and other constraints (coming up)
- Triangulate for 3d structure

Left Right

Left Right



Disparity, depth maps

image I(x,y)

Disparity map D(x,y)

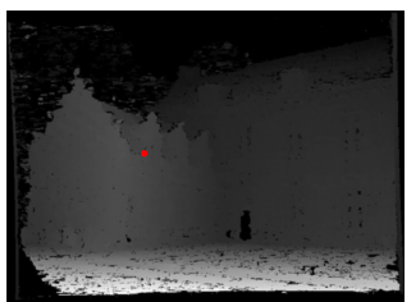
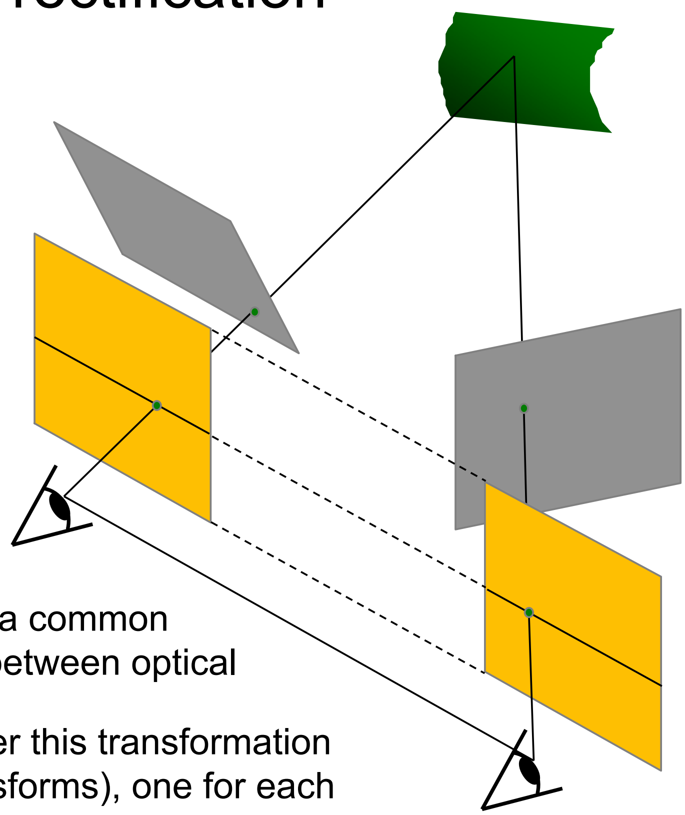


image I´(x´,y´)

$$(x´,y´)=(x+D(x,y),y)$$

Stereo image rectification

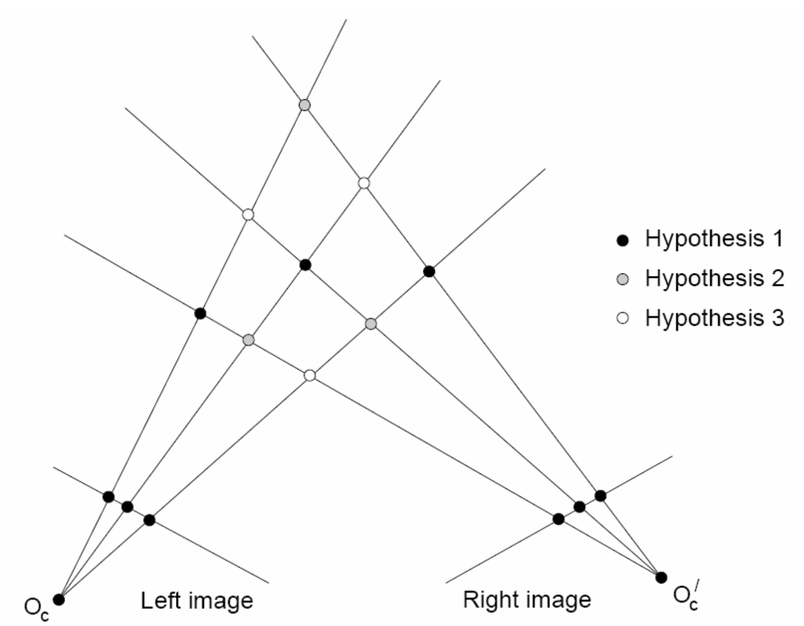
Motivation: make the lines to be searched correspond to scanlines in images



- 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

Adapted from Li Zhang

Correspondence problem



Multiple match hypotheses satisfy epipolar constraint, but which is correct?

Figure from Gee & Cipolla 1999

Correspondence problem

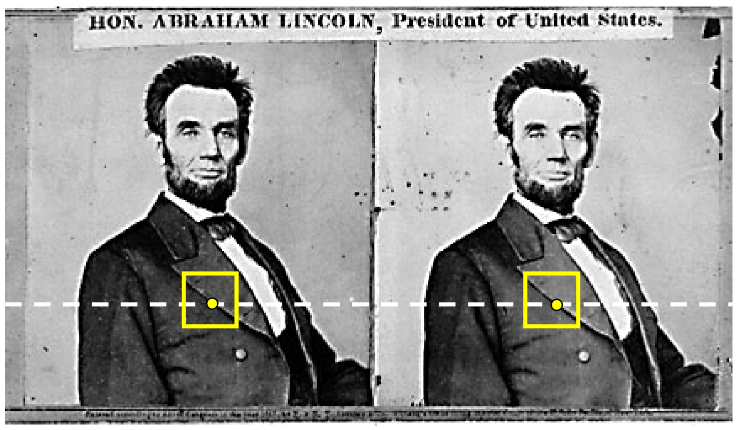
- To find matches in the image pair, we will assume
  - Most scene points visible from both views
  - *Image regions for the matches are similar in appearance*

- Ok when distance of fixation point >> baseline
- (But, we can’t guarantee)

Additional correspondence constraints

- Similarity
- Uniqueness
- Ordering
- Figural continuity
- Disparity gradient

Dense correspondence search



For each epipolar line

For each pixel / window in the left image

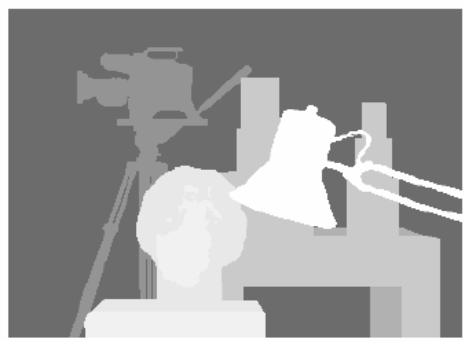
- compare with every pixel / window on same epipolar line in right image
- pick position with minimum match cost (e.g., SSD, correlation)

Adapted from Li Zhang

Example: window search

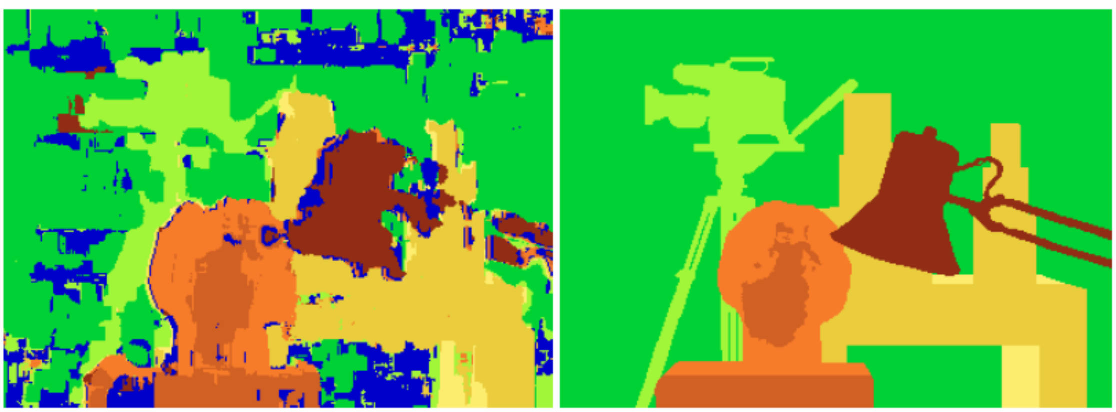
Data from University of Tsukuba

Scene



Ground truth

Example: window search

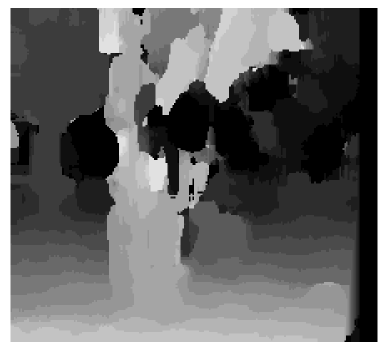


Window-based matching
(best window size)

Ground truth

Effect of window size

W = 3

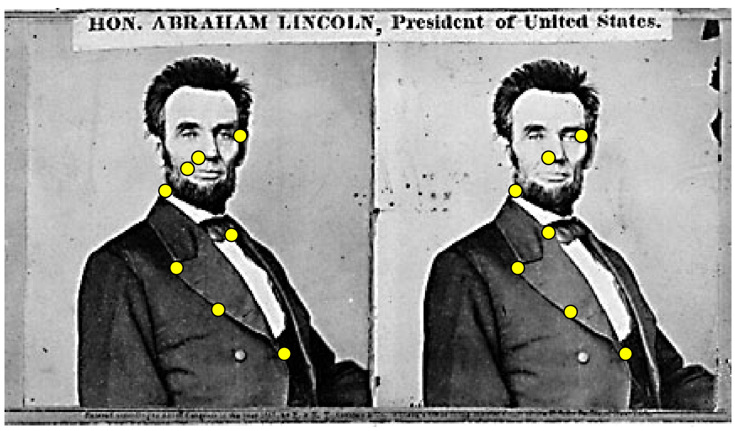


W = 20

Want window large enough to have sufficient intensity variation, yet small enough to contain only pixels with about the same disparity.

Figures from Li Zhang

Sparse correspondence search



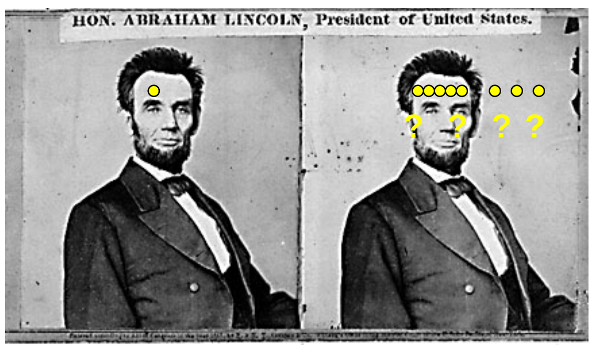
- Restrict search to sparse set of detected features
- Rather than pixel values (or lists of pixel values) use *feature descriptor* and an associated *feature distance*
- Still narrow search further by epipolar geometry

What would make good features?

Dense vs. sparse

- Sparse
  - Efficiency
  - Can have more reliable feature matches, less sensitive to illumination than raw pixels
  - …But, have to know enough to pick good features; sparse info
- Dense
  - Simple process
  - More depth estimates, can be useful for surface reconstruction
  - …But, breaks down in textureless regions anyway, raw pixel distances can be brittle, not good with very different viewpoints

Difficulties in similarity constraint



Untextured surfaces

Occlusions

Uniqueness

- For opaque objects, up to one match in right image for every point in left image

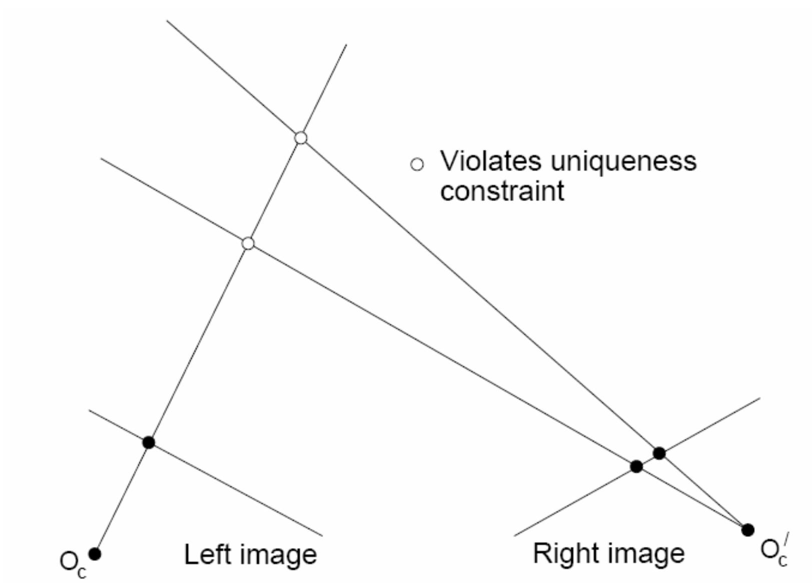


Figure from Gee & Cipolla 1999

Ordering

- Points on **same surface** (opaque object) will be in same order in both views

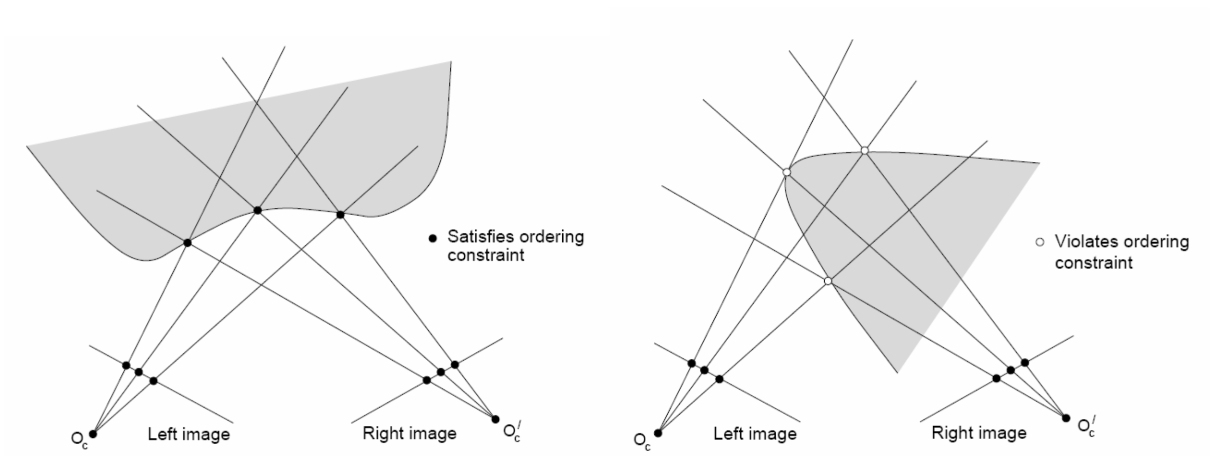


Figure from Gee & Cipolla 1999

Figural continuity

- When interest points lie on image contours

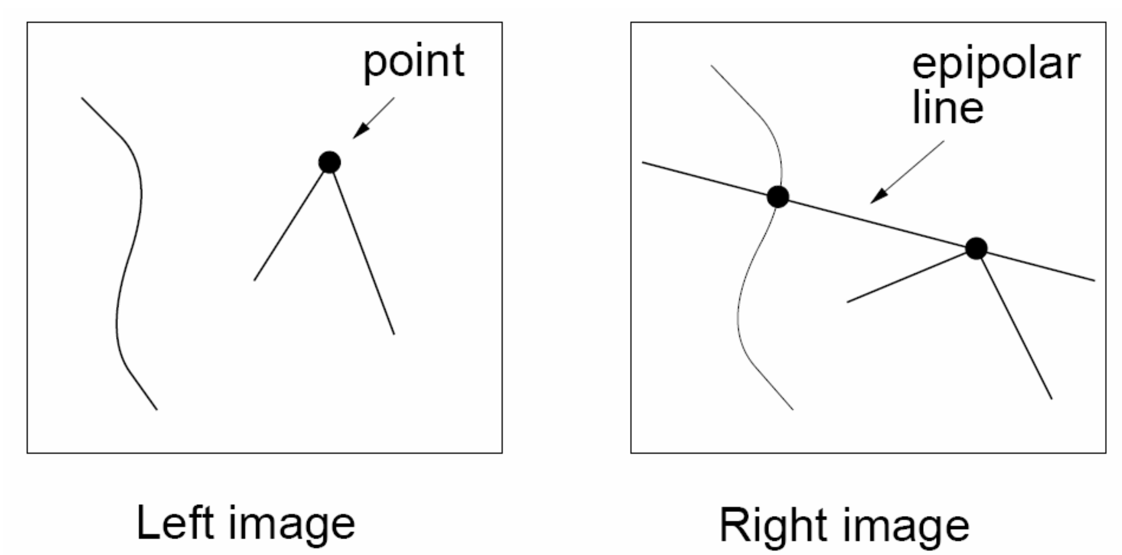
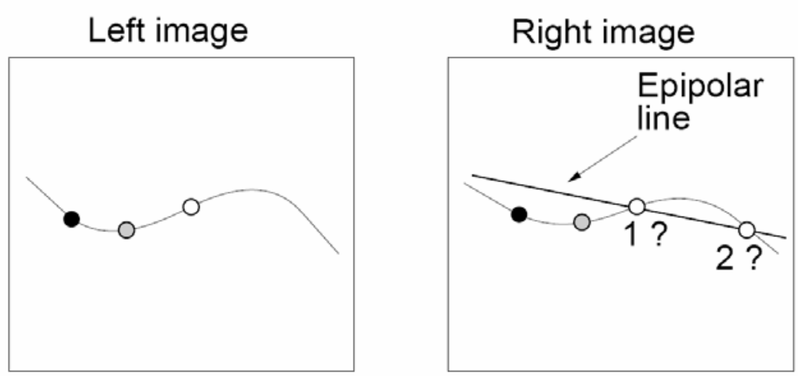


Figure from Gee & Cipolla 1999

Disparity gradient

- Assume piecewise continuous surface, so want disparity estimates to be locally smooth



Given matches ● and ◎, point ○ in the left image must match point 1 in the right image. Point 2 would exceed the disparity gradient limit.

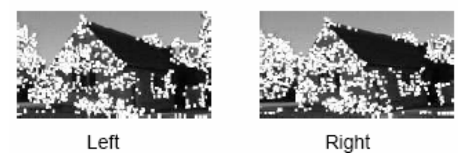
Figure from Gee & Cipolla 1999

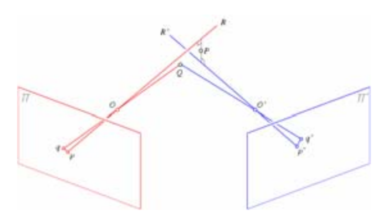
Additional correspondence constraints

- Similarity
- Uniqueness
- Ordering
- Figural continuity
- Disparity gradient

Stereo reconstruction for fully calibrated cameras

- Image pair
- Detect some features
- Compute E from R and T
- Match features using the epipolar and other constraints
- Triangulate for 3d structure





Sources of error in correspondences

- Low-contrast / textureless image regions
- Occlusions
- Camera calibration errors
- Poor image resolution
- Violations of brightness constancy (specular reflections)
- Large motions

Model-based body tracking, stereo input

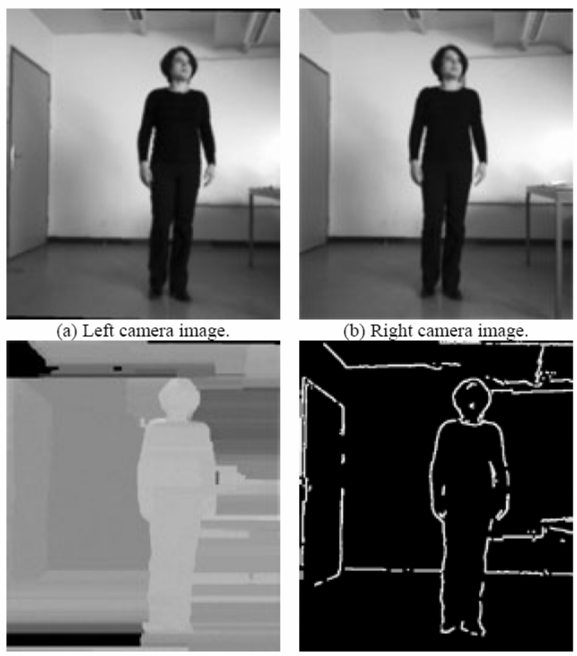
Fitting!

David Demirdjian, MIT Vision Interface Group

Model-based body tracking, stereo input

David Demirdjian, MIT Vision Interface Group

Depth for segmentation



(a) Left camera image. (b) Right camera image.

(c) Depth image. (d) Edge combination image.

Figure 3 Stereo video frames with computed depth map and edge combination result.

Edges in disparity in conjunction with image edges enhances contours found

Danijela Markovic and Margrit Gelautz, Interactive Media Systems Group, Vienna University of Technology

Depth for segmentation

(a) Original image with snake initialization.

(b) Final snake on original image.

(c) Final snake on depth image.

(d) Original image with snake from (c) overlaid.

(e) Final snake on edge combination image.

(f) Original image with snake from (e) overlaid.

Danijela Markovic and Margrit Gelautz, Interactive Media Systems Group, Vienna University of Technology

Uncalibrated case

- What if we don’t know the extrinsic camera parameters?
- What if we don’t even know the intrinsic parameters?
- We can still reconstruct 3d structure, up to certain ambiguities, if we can find correspondences between points…

Coming up

- Exam Tuesday Oct 9 (next class)

- Thursday (Oct 11):
  - Finish up multi-view geometry and stereo

- Following week (Oct 16 and 18):
  - Guest lectures
    - Dana Ballard
    - Michael Ryoo & Shalini Gupta