Single-view Metrology and Camera Calibration

Computer Vision

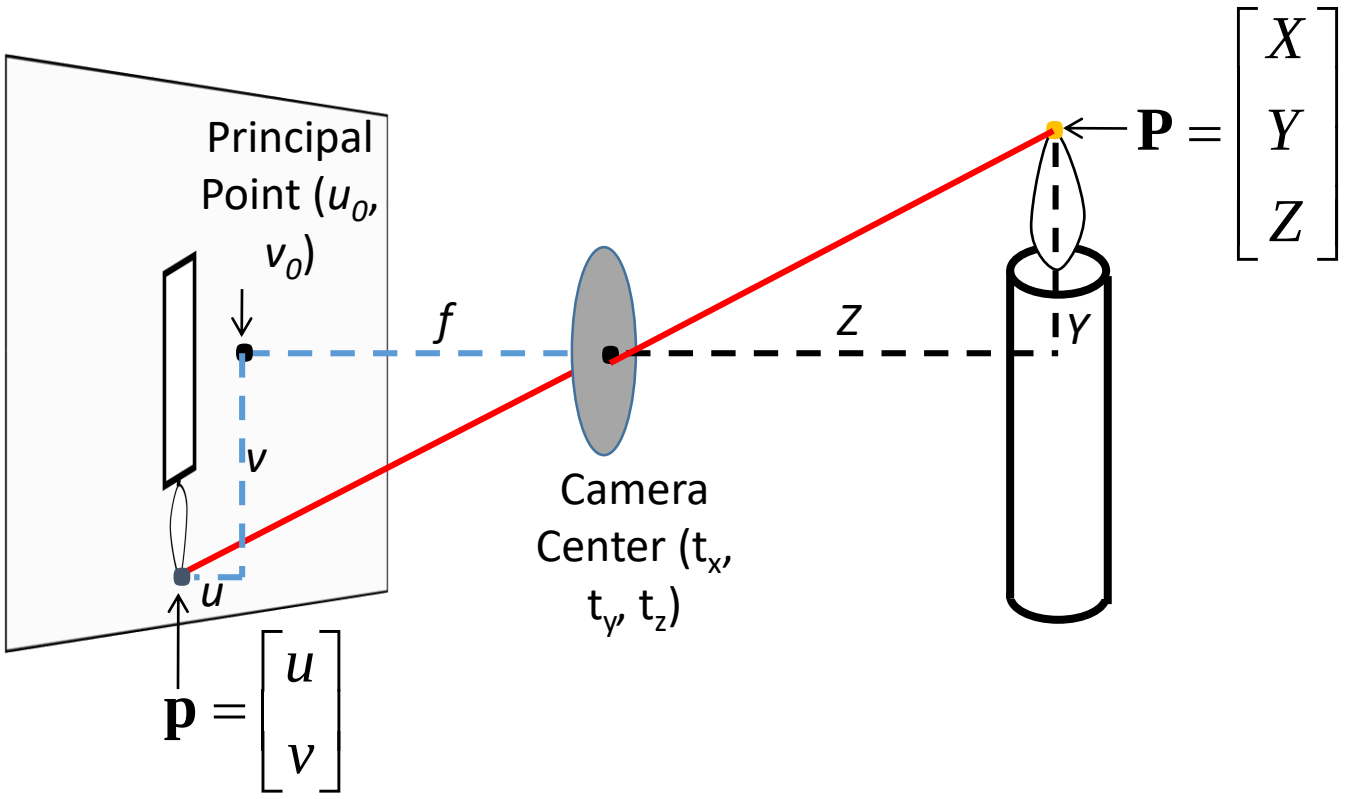
Jia-Bin Huang, Virginia Tech

Many slides from S. Seitz and D. Hoiem

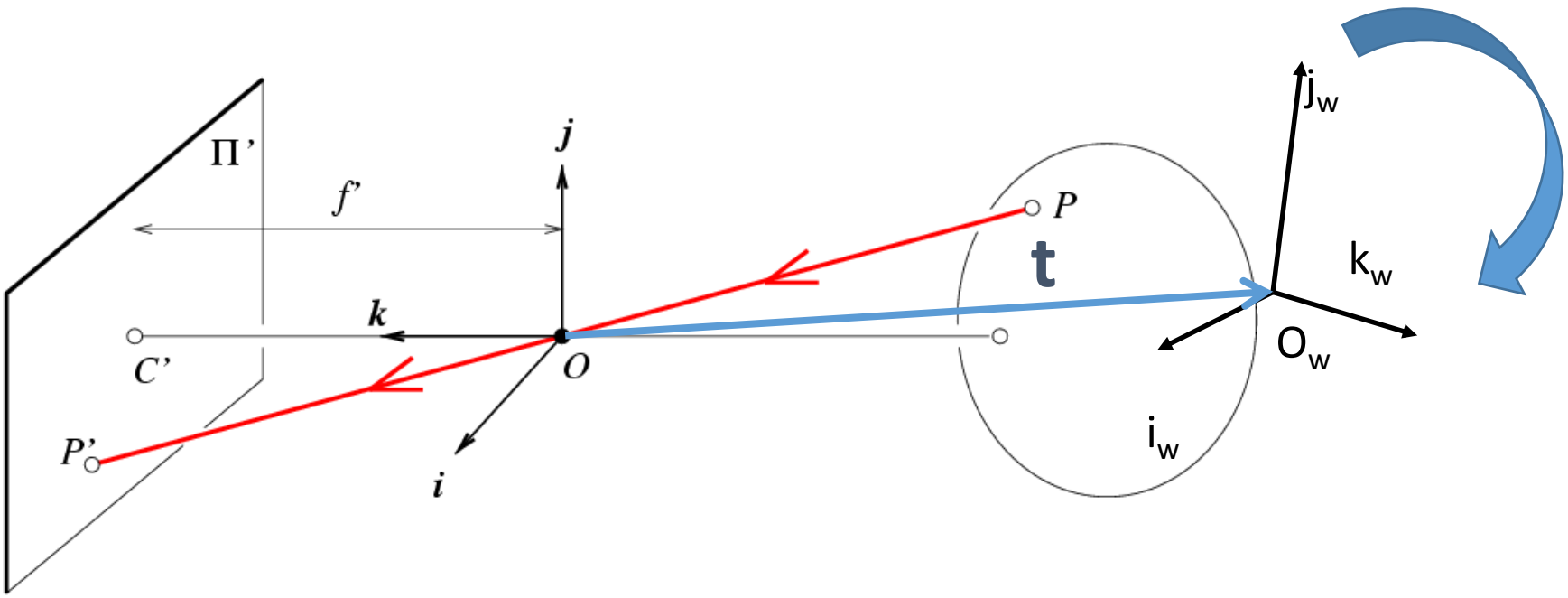
Administrative stuffs

- HW 2 due 11:59 PM on Oct 3rd
- HW 2 competition on shape alignment
  - Submit your results here
- Ask/discuss questions on Piazza

Last Class: Pinhole Camera

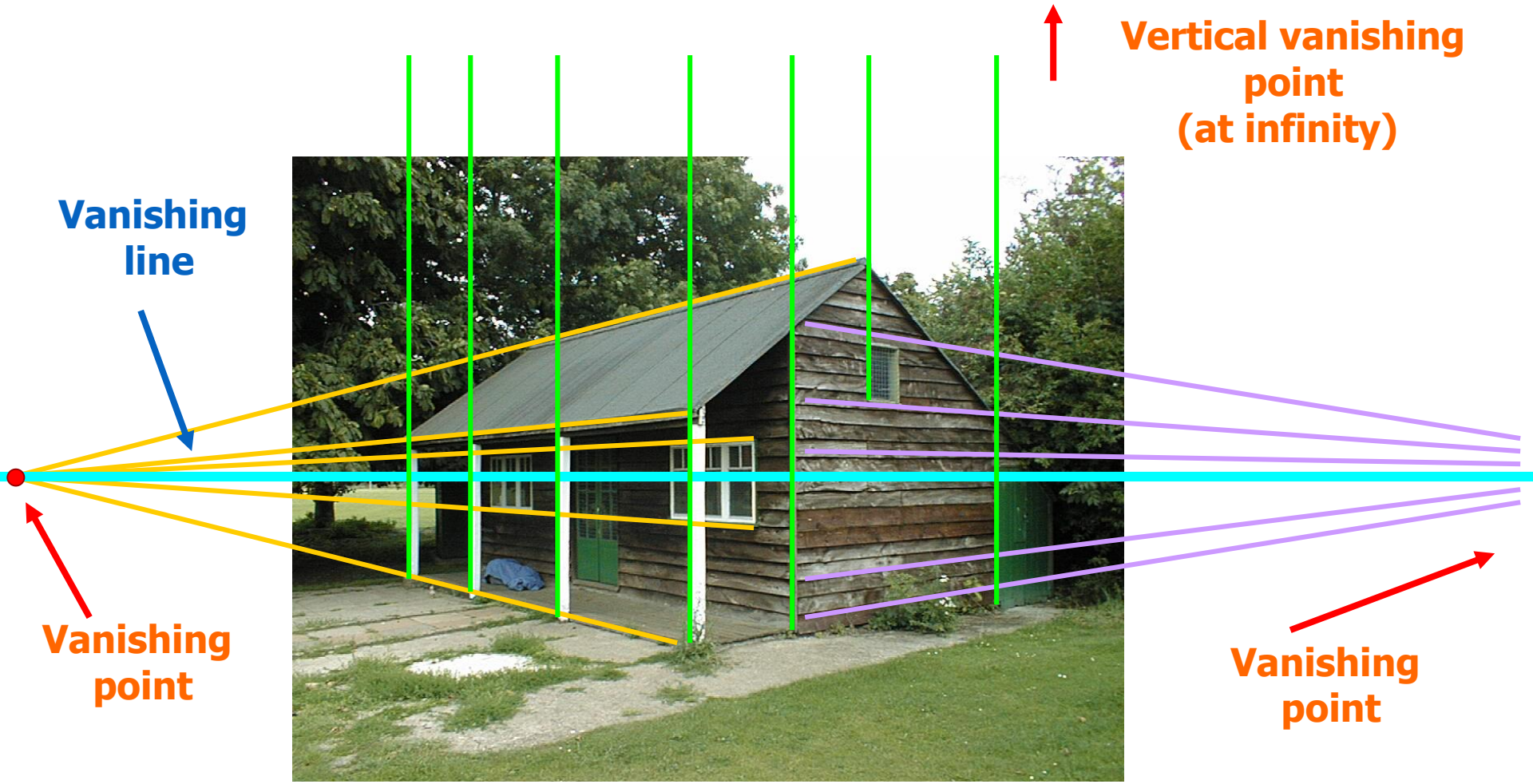


Last Class: Projection Matrix



$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\ v\\ 1\end{bmatrix} = \begin{bmatrix}f & s & u_0\\ 0 & \alpha f & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13} & t_x\\ r_{21} & r_{22} & r_{23} & t_y\\ r_{31} & r_{32} & r_{33} & t_z\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

Last class: Vanishing Points



Slide from Efros, Photo from Criminisi

This class

- How can we calibrate the camera?
- How can we measure the size of objects in the world from an image?
- What about other camera properties: focal length, field of view, depth of field, aperture, f-number?

How to calibrate the camera?

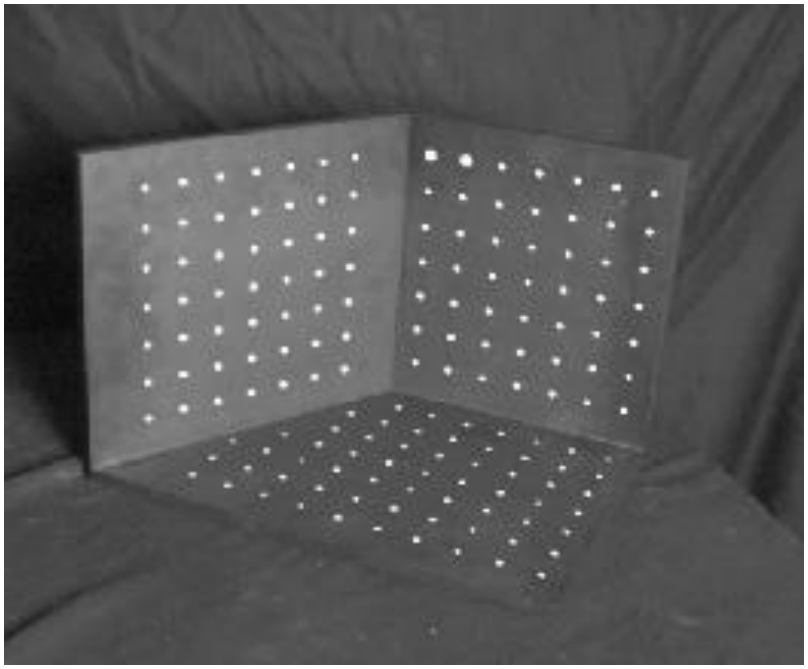
$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Calibrating the Camera

Method 1: Use an object (calibration grid) with known geometry

- Correspond image points to 3d points
- Get least squares solution (or non-linear solution)



Known 2d image coordinates

Known 3d locations

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

Unknown Camera Parameters

Unknown Camera Parameters

Known 2d image coords

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

Known 3d locations

$$m_{31}uX + m_{32}uY + m_{33}uZ + m_{34}u = m_{11}X + m_{12}Y + m_{13}Z + m_{14}$$
$$m_{31}vX + m_{32}vY + m_{33}vZ + m_{34}v = m_{21}X + m_{22}Y + m_{23}Z + m_{24}$$

$$0 = m_{11}X + m_{12}Y + m_{13}Z + m_{14} - m_{31}uX - m_{32}uY - m_{33}uZ - m_{34}u$$
$$0 = m_{21}X + m_{22}Y + m_{23}Z + m_{24} - m_{31}vX - m_{32}vY - m_{33}vZ - m_{34}v$$

Unknown Camera Parameters

Known 2d image coords

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

Known 3d locations

$$0 = m_{11}X + m_{12}Y + m_{13}Z + m_{14} - m_{31}uX - m_{32}uY - m_{33}uZ - m_{34}u$$

$$0 = m_{21}X + m_{22}Y + m_{23}Z + m_{24} - m_{31}vX - m_{32}vY - m_{33}vZ - m_{34}v$$

- Method 1 – homogeneous linear system. Solve for m's entries using linear least squares

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

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

Unknown Camera Parameters

Known 2d image coords

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

Known 3d locations

- Method 2 – nonhomogeneous linear system. Solve for m's entries using linear least squares

Ax=b form

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

```
M = A\Y;
M = [M;1];
M = reshape(M,[],3)';
```

Calibration with linear method

- Advantages
  - Easy to formulate and solve
  - Provides initialization for non-linear methods

- Disadvantages
  - Doesn’t directly give you camera parameters
  - Doesn’t model radial distortion
  - Can’t impose constraints, such as known focal length
  - Doesn’t minimize projection error

- Non-linear methods are preferred
  - Define error as difference between projected points and measured points
  - Minimize error using Newton’s method or other non-linear optimization

Can we factorize M back to K [R | T]?

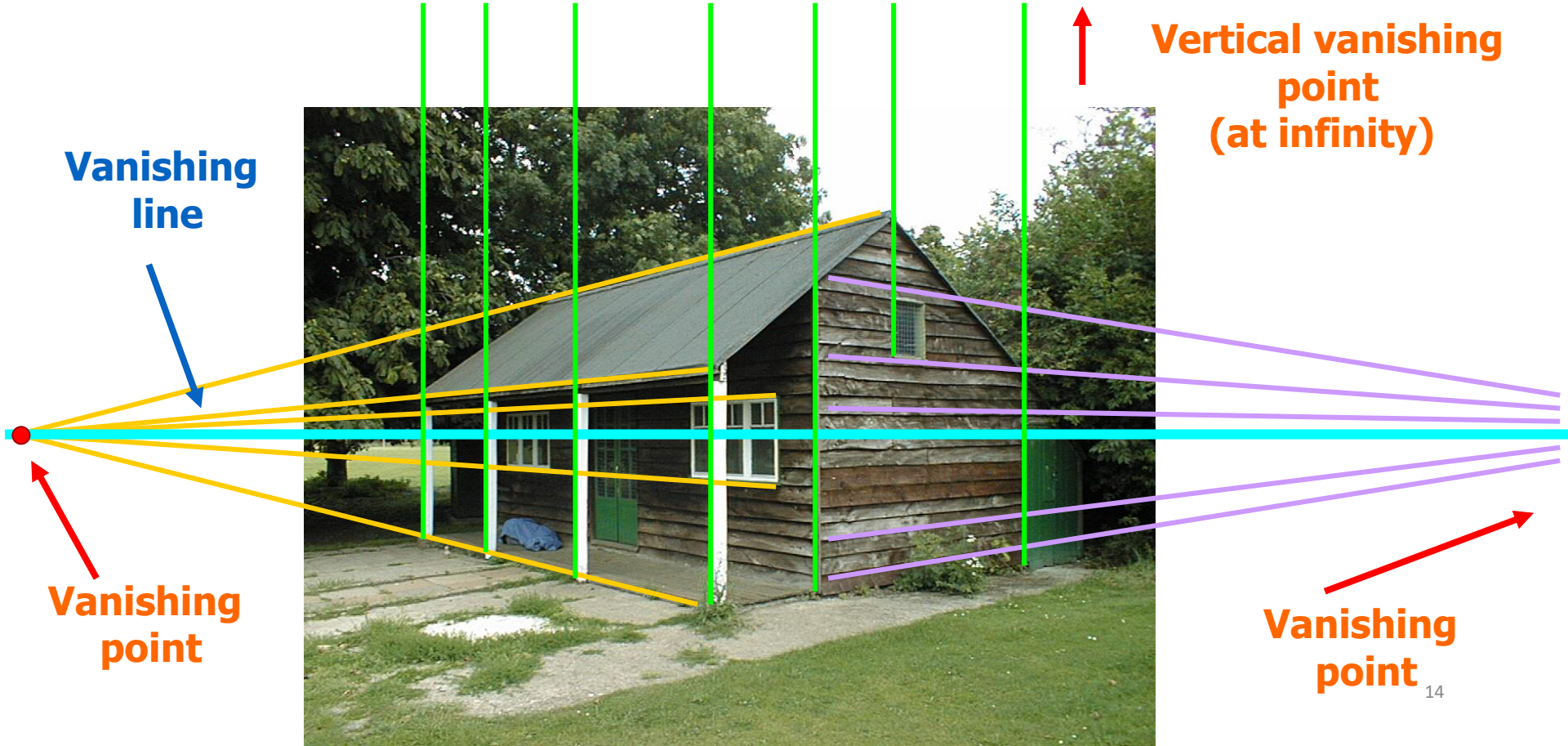
- Yes!
- You can use *RQ* factorization (note – not the more familiar *QR* factorization). *R* (right diagonal) is K, and *Q* (orthogonal basis) is R. T, the last column of [R | T], is inv(K) * last column of M.
  - But you need to do a bit of post-processing to make sure that the matrices are valid. See http://ksimek.github.io/2012/08/14/decompose/

Slide credit: J. Hays

Calibrating the Camera

Method 2: Use vanishing points

- Find vanishing points corresponding to orthogonal directions



Calibration by orthogonal vanishing points

- Intrinsic camera matrix
  - Use orthogonality as a constraint
  - Model K with only f, u_0, v_0

For vanishing points

$$\mathbf{p}_i = \mathbf{K}\mathbf{R}\mathbf{X}_i \qquad \mathbf{X}_i^T \mathbf{X}_j = 0$$

$$\mathbf{X_i} = \mathbf{R^{-1}K^{-1}p_i}$$

$$\mathbf{p_i^\top}\left(\mathbf{K^{-1}}\right)^\top(\mathbf{R})(\mathbf{R^{-1}})(\mathbf{K^{-1}})\mathbf{p_i} = \mathbf{0}$$

- What if you don't have three finite vanishing points?
  - Two finite VP: solve f, get valid u_0, v_0 closest to image center
  - One finite VP: u_0, v_0 is at vanishing point; can't solve for f

Calibration by vanishing points

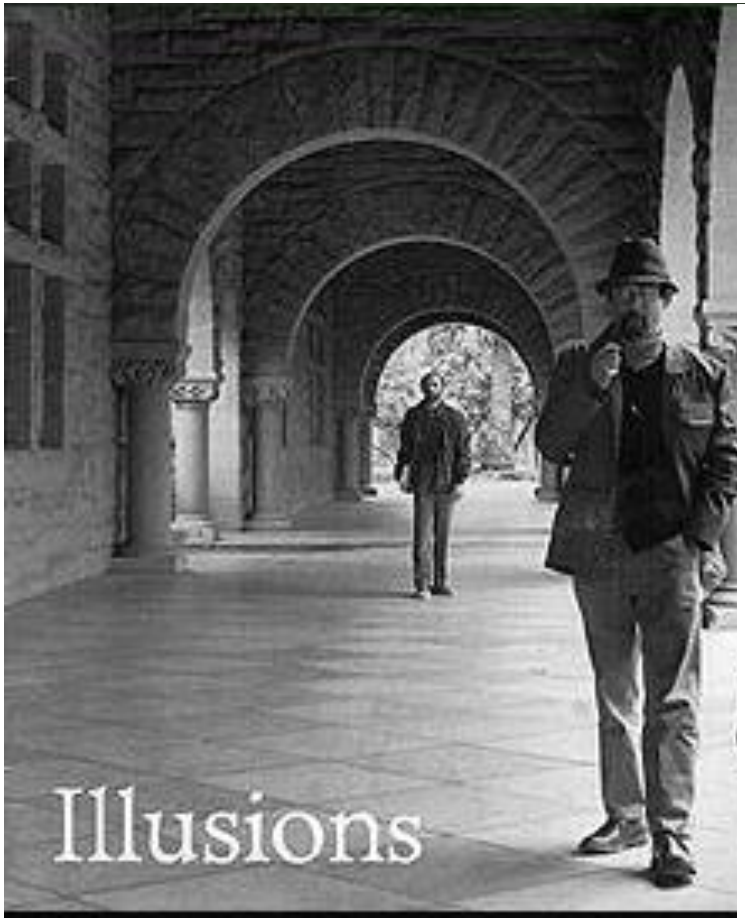
- Intrinsic camera matrix

$$\mathbf{p}_i = \mathbf{K}\mathbf{R}\mathbf{X}_i$$

- Rotation matrix
  - Set directions of vanishing points
    - e.g., $\mathbf{X}_1 = [1, 0, 0]$
  - Each VP provides one column of **R**
  - Special properties of **R**
    - inv(**R**)=$\mathbf{R}^T$
    - Each row and column of **R** has unit length

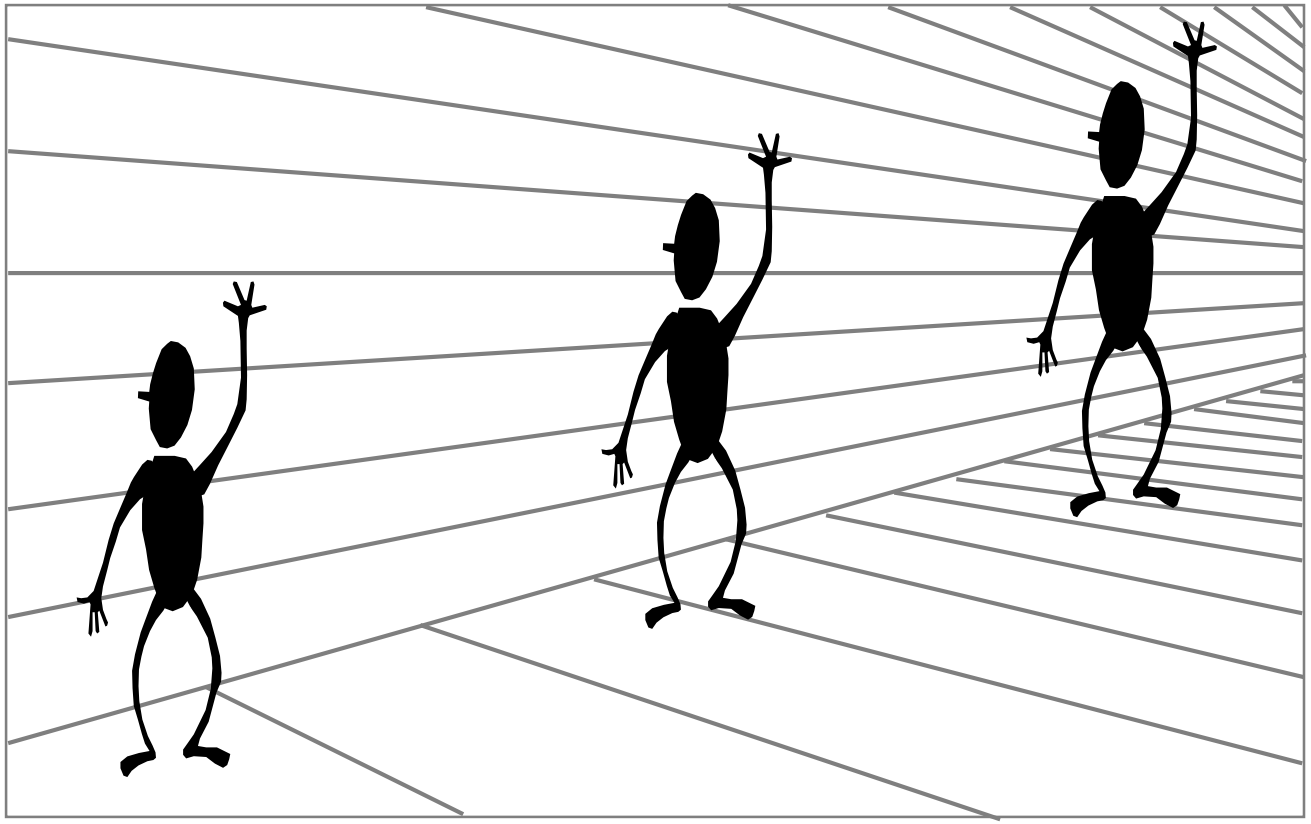
$$\mathbf{p_i} = \mathbf{K}\mathbf{r_i}$$

How can we measure the size of 3D objects from an image?

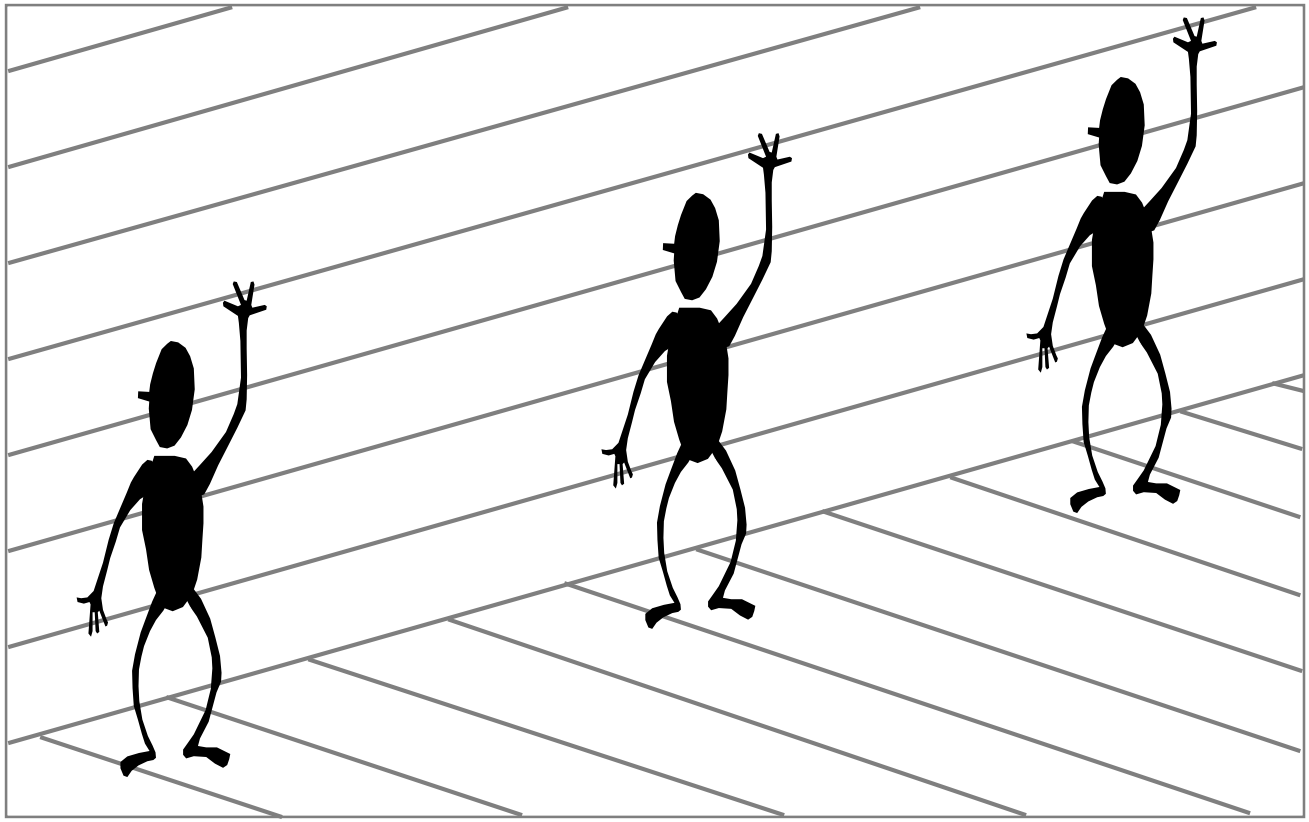


Slide by Steve Seitz

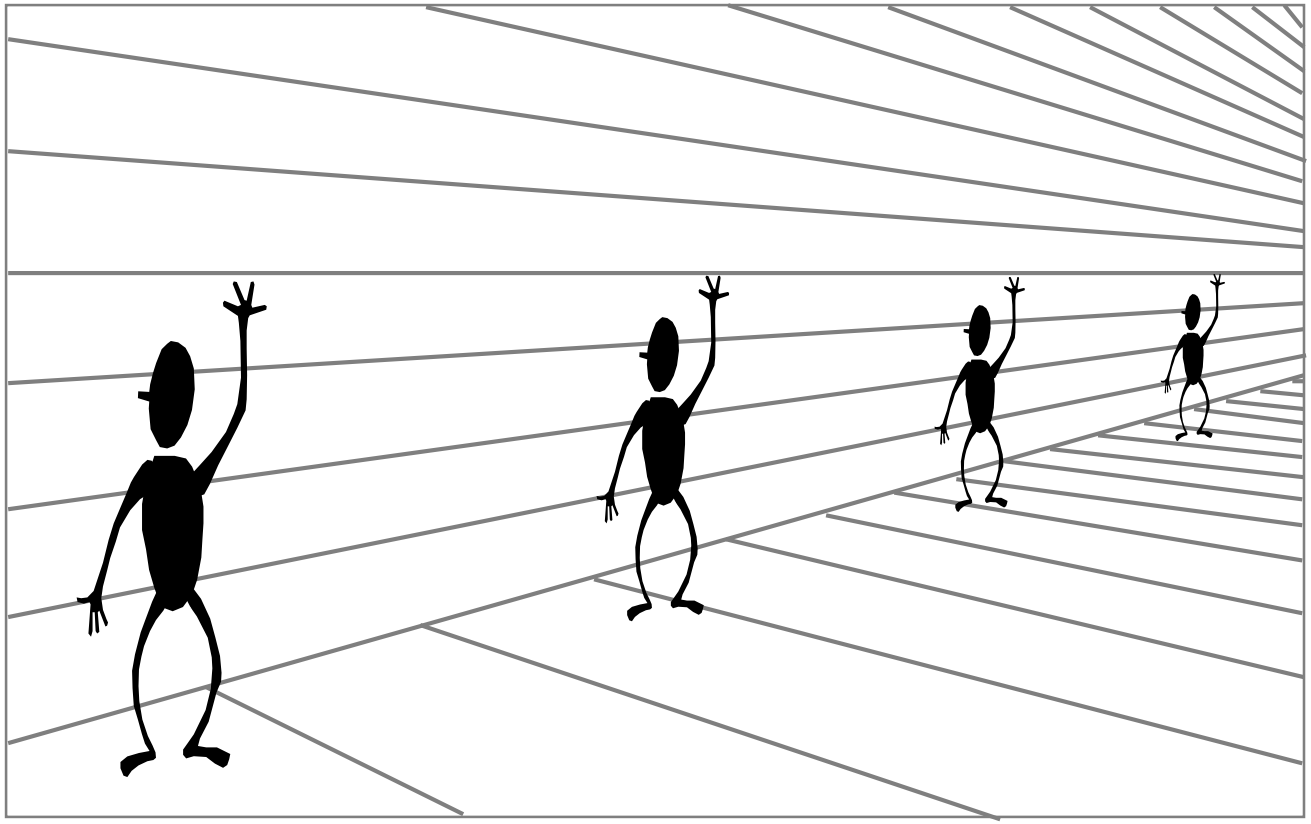
Perspective cues



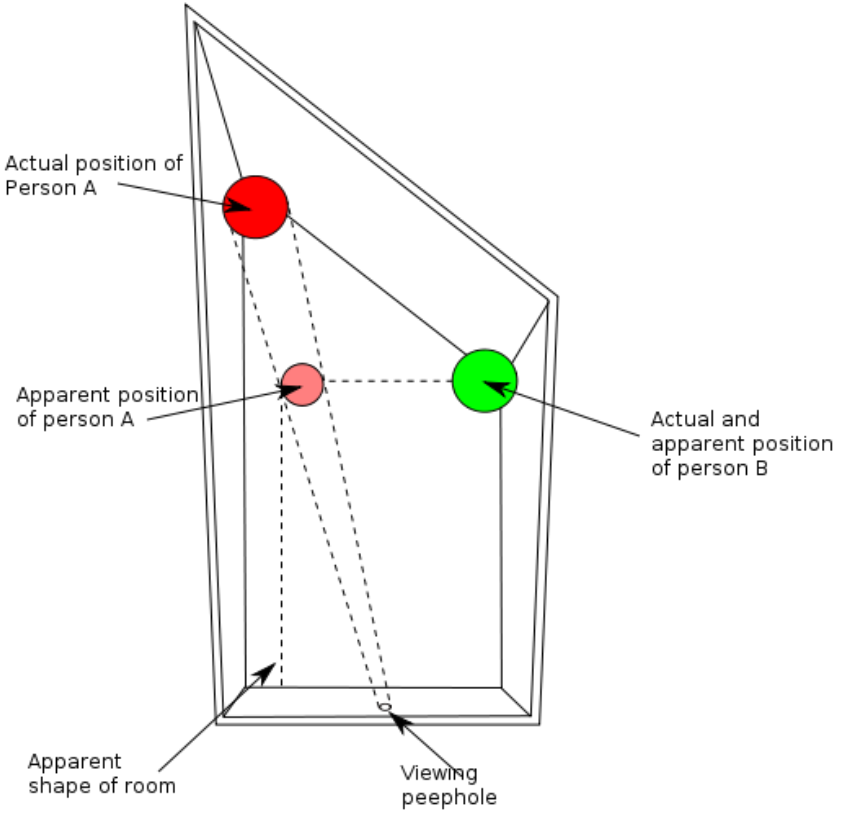
Perspective cues



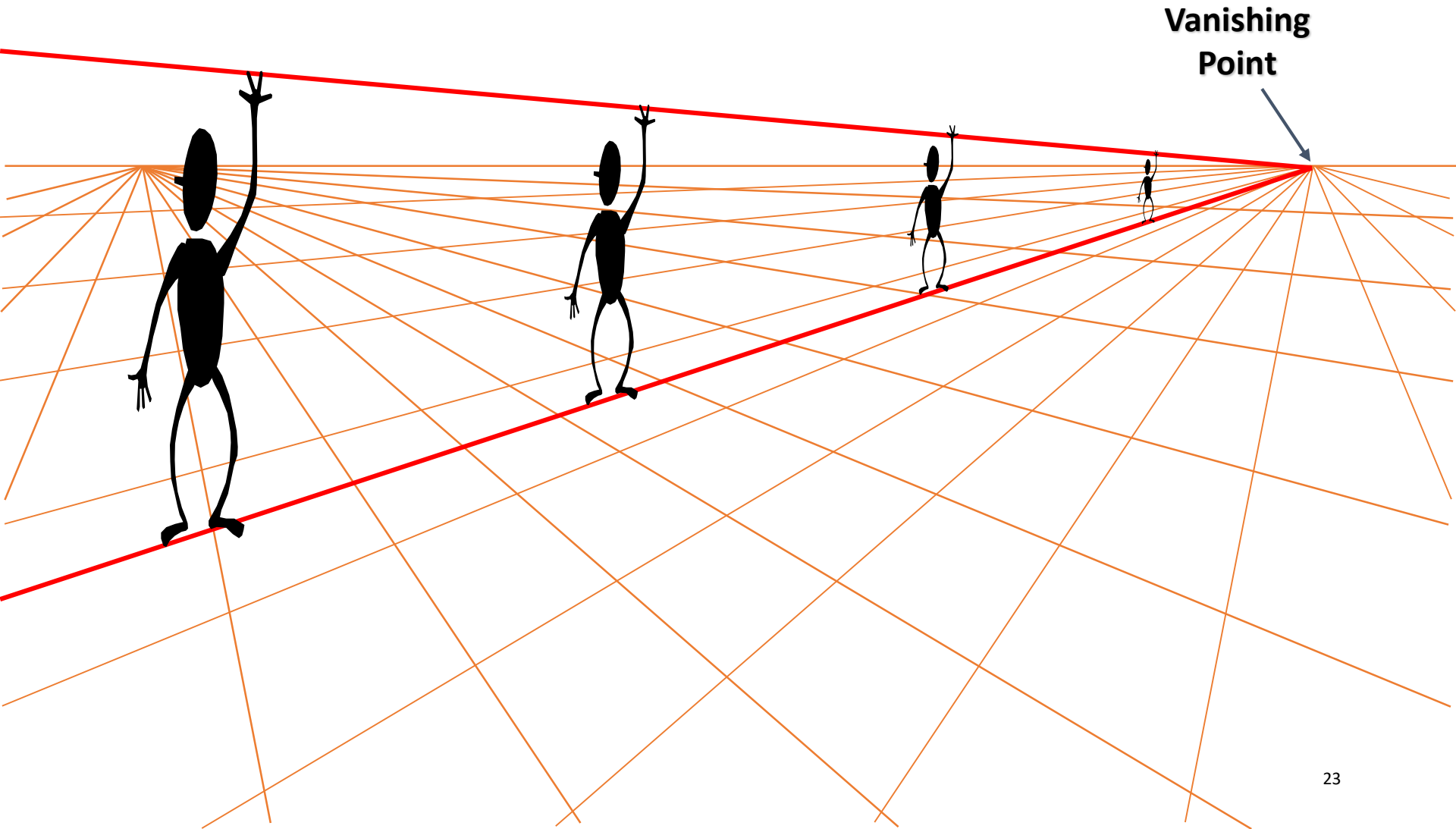
Perspective cues



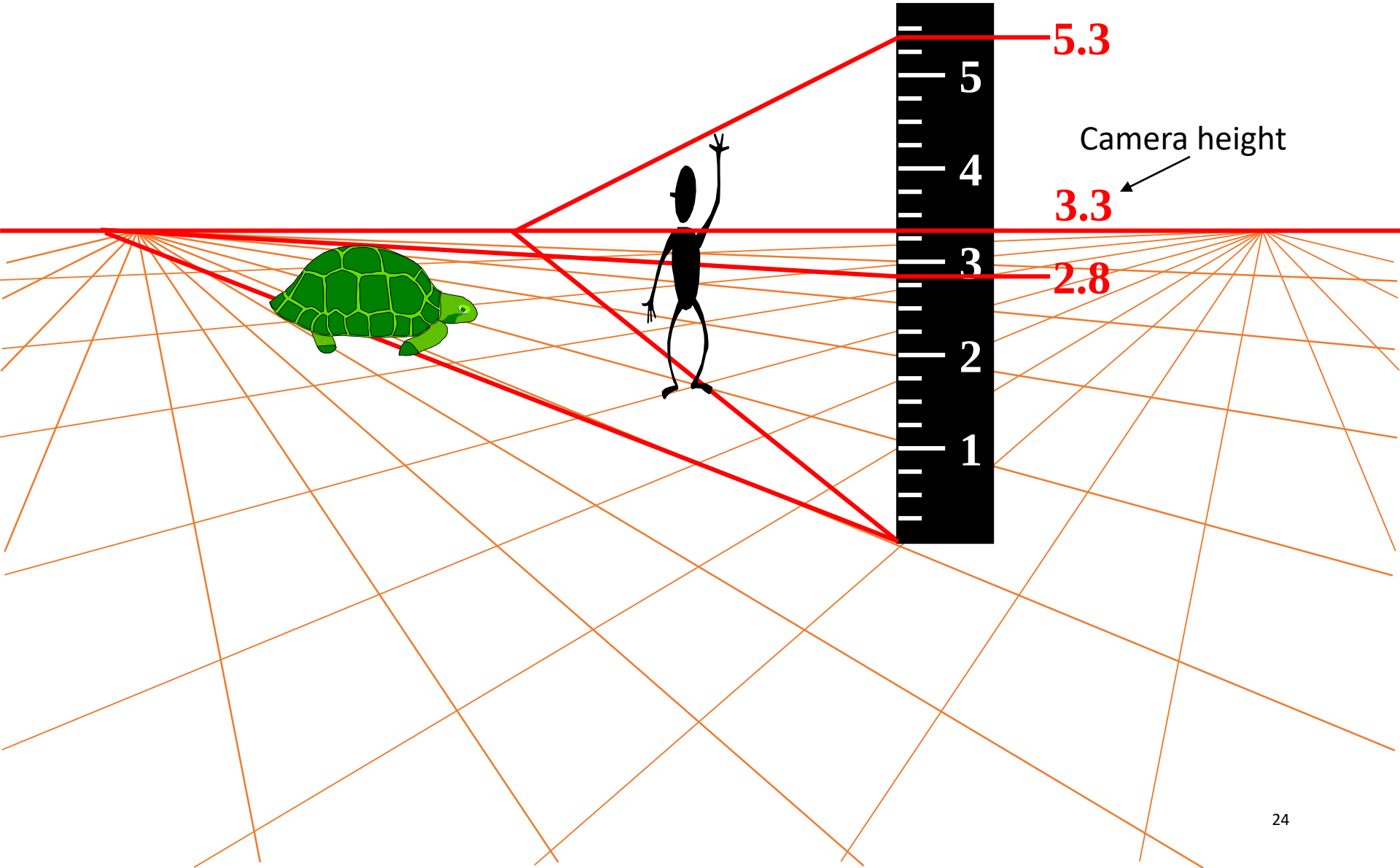
Ames Room


Actual position of Person A
Apparent position of person A
Actual and apparent position of person B
Apparent shape of room
Viewing peephole

Comparing heights



Measuring height

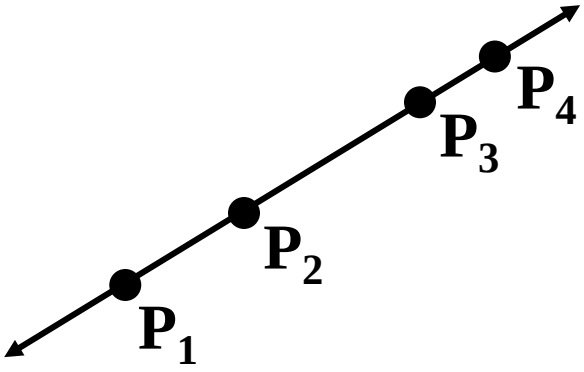


Which is higher – the camera or the man in the parachute?

The cross ratio

- A Projective Invariant
  - Something that does not change under projective transformations (including perspective projection)

The cross-ratio of 4 collinear points



$$\frac{\|\mathbf{P}_3 - \mathbf{P}_1\| \, \|\mathbf{P}_4 - \mathbf{P}_2\|}{\|\mathbf{P}_3 - \mathbf{P}_2\| \, \|\mathbf{P}_4 - \mathbf{P}_1\|}$$

$$\mathbf{P}_i = \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

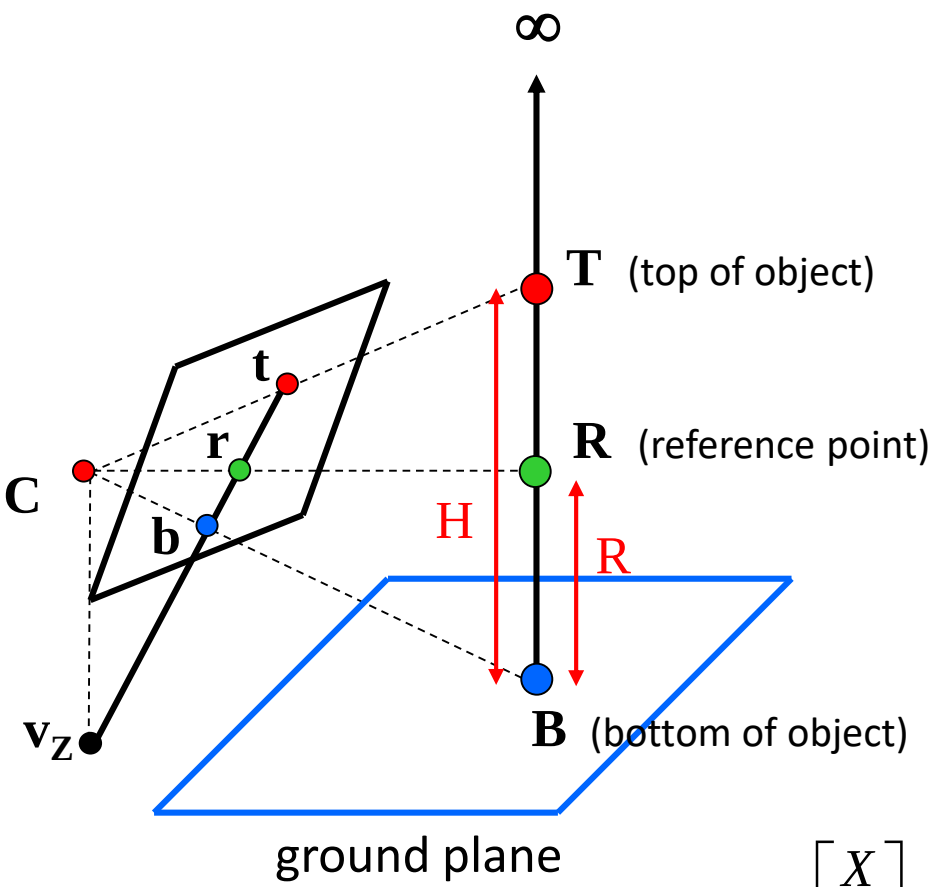
Can permute the point ordering

$$\frac{\|\mathbf{P}_1 - \mathbf{P}_3\| \, \|\mathbf{P}_4 - \mathbf{P}_2\|}{\|\mathbf{P}_1 - \mathbf{P}_2\| \, \|\mathbf{P}_4 - \mathbf{P}_3\|}$$

- 4! = 24 different orders (but only 6 distinct values)

This is the fundamental invariant of projective geometry

Measuring height



$$\frac{\|\mathbf{B}-\mathbf{T}\|\,\|\infty-\mathbf{R}\|}{\|\mathbf{B}-\mathbf{R}\|\,\|\infty-\mathbf{T}\|} = \frac{H}{R}$$

scene cross ratio

$$\frac{\|\mathbf{b}-\mathbf{t}\|\,\|\mathbf{v}_Z-\mathbf{r}\|}{\|\mathbf{b}-\mathbf{r}\|\,\|\mathbf{v}_Z-\mathbf{t}\|} = \frac{H}{R}$$

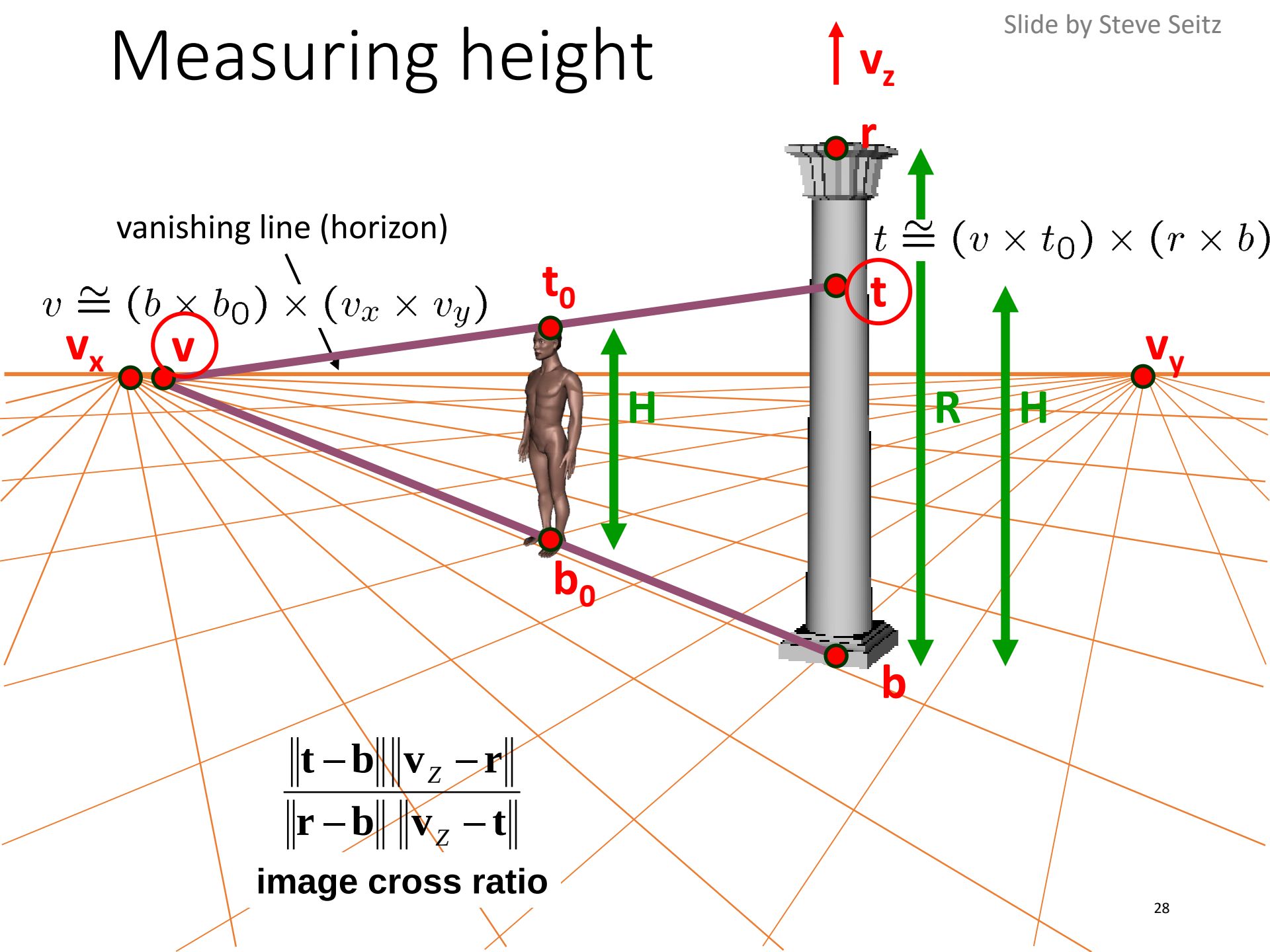
image cross ratio

scene points represented as $\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$

image points as $\mathbf{p} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

Measuring height

Slide by Steve Seitz

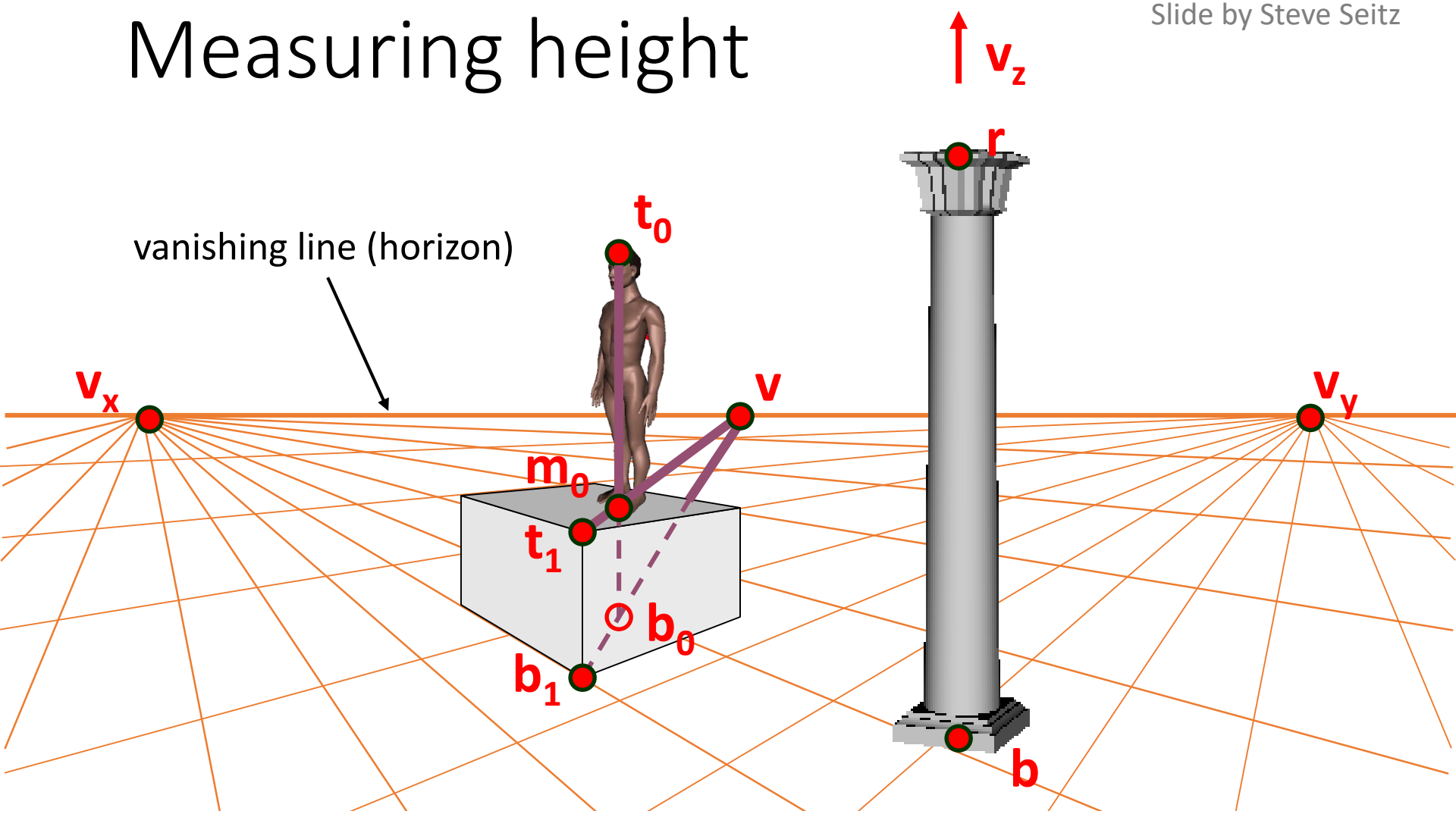


$$\frac{\|\mathbf{t}-\mathbf{b}\|\|\mathbf{v}_z-\mathbf{r}\|}{\|\mathbf{r}-\mathbf{b}\|\|\mathbf{v}_z-\mathbf{t}\|}$$

image cross ratio

Measuring height

Slide by Steve Seitz

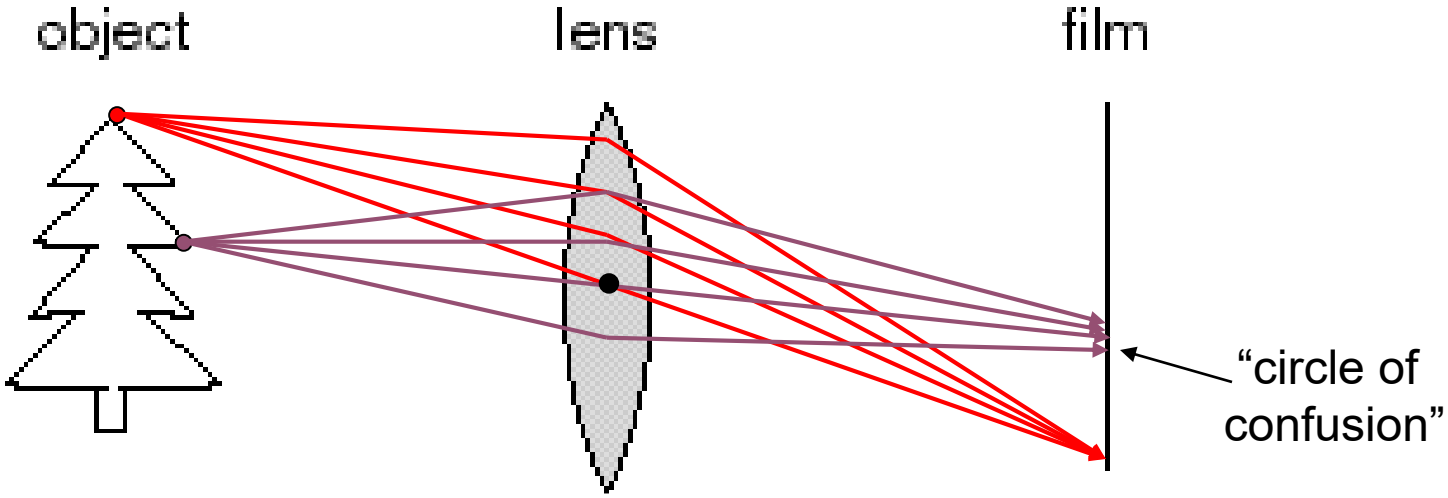


What if the point on the ground plane $\mathbf{b}_0$ is not known?

- Here the guy is standing on the box, height of box is known
- Use one side of the box to help find $\mathbf{b}_0$ as shown above

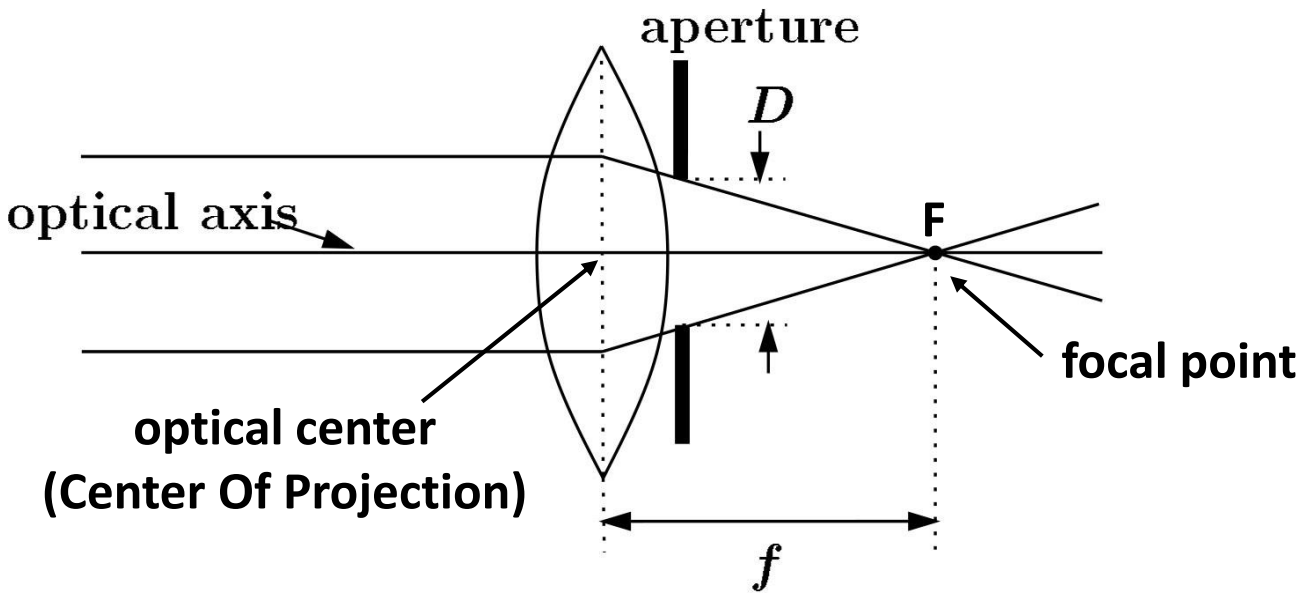
What about focus, aperture, DOF, FOV, etc?

Adding a lens



- A lens focuses light onto the film
  - There is a specific distance at which objects are "in focus"
    - other points project to a "circle of confusion" in the image
  - Changing the shape of the lens changes this distance

Focal length, aperture, depth of field

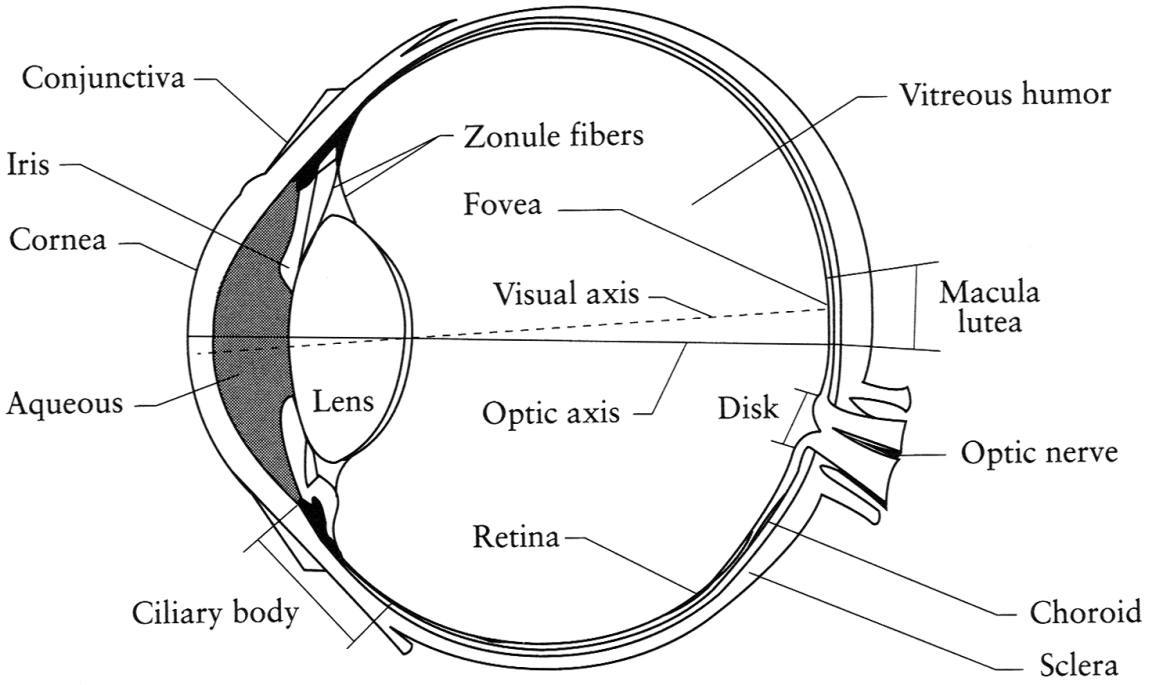


A lens focuses parallel rays onto a single focal point

- focal point at a distance *f* beyond the plane of the lens
- Aperture of diameter D restricts the range of rays

Slide source: Seitz

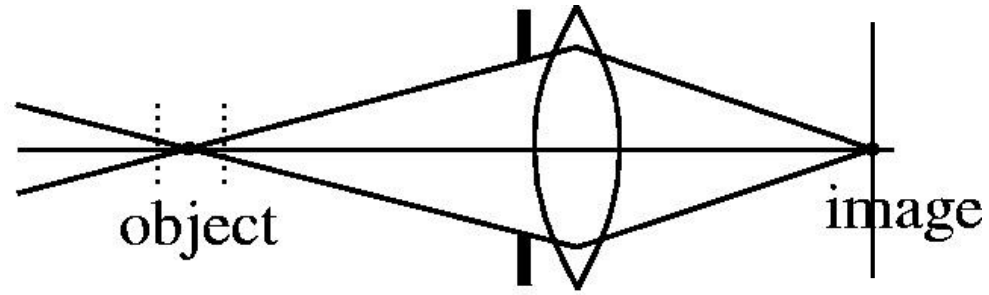
The eye



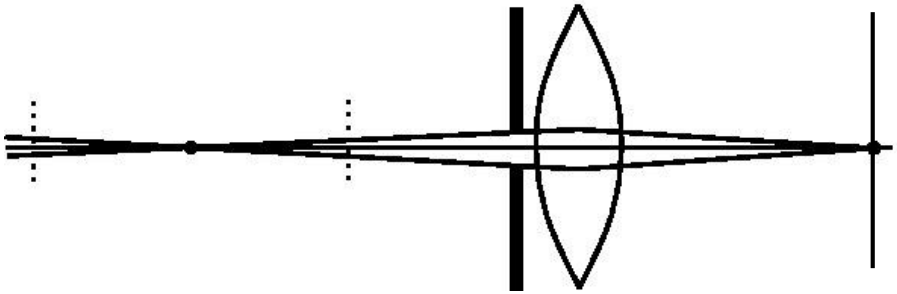
- The human eye is a camera
  - **Iris** - colored annulus with radial muscles
  - **Pupil** (aperture) - the hole whose size is controlled by the iris
  - **Retina** (film): photoreceptor cells (rods and cones)

Slide source: Seitz

Depth of field



f / 5.6



f / 32

Changing the aperture size or focal length affects depth of field

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

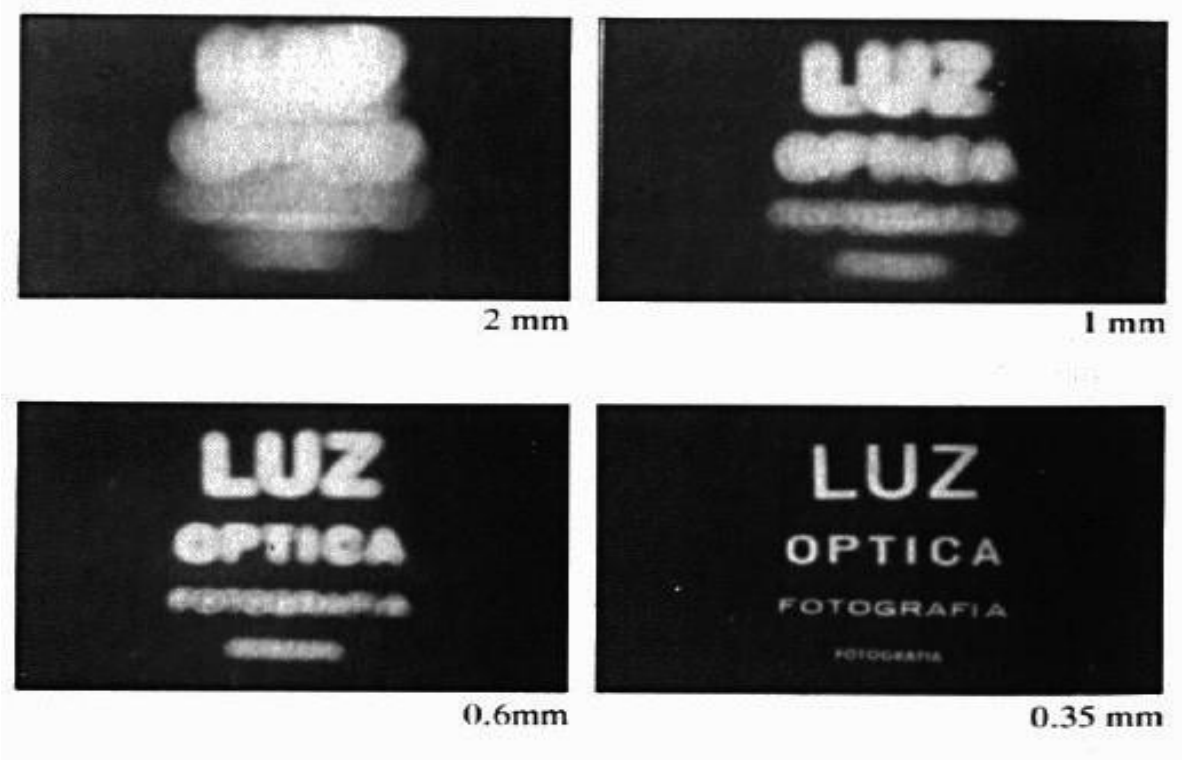
Slide from Efros

Varying the aperture

Large aperture = small DOF

Small aperture = large DOF

Shrinking the aperture



- Why not make the aperture as small as possible?
  - Less light gets through
  - Diffraction effects

Slide by Steve Seitz

Shrinking the aperture

Slide by Steve Seitz

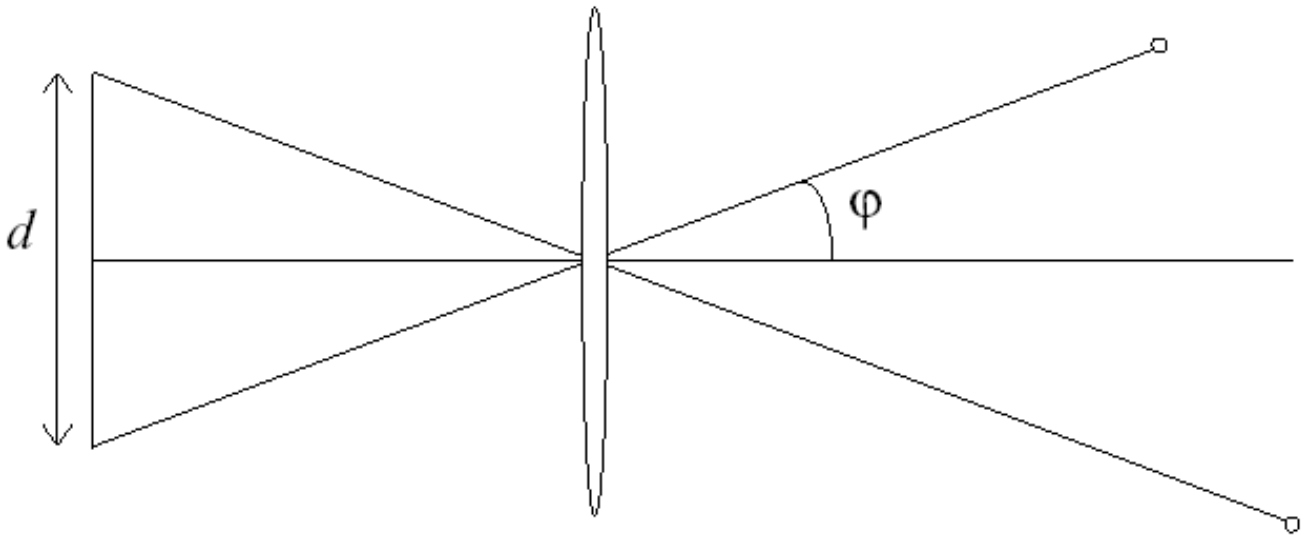
Relation between field of view and focal length

Field of view (angle width)

Film/Sensor Width

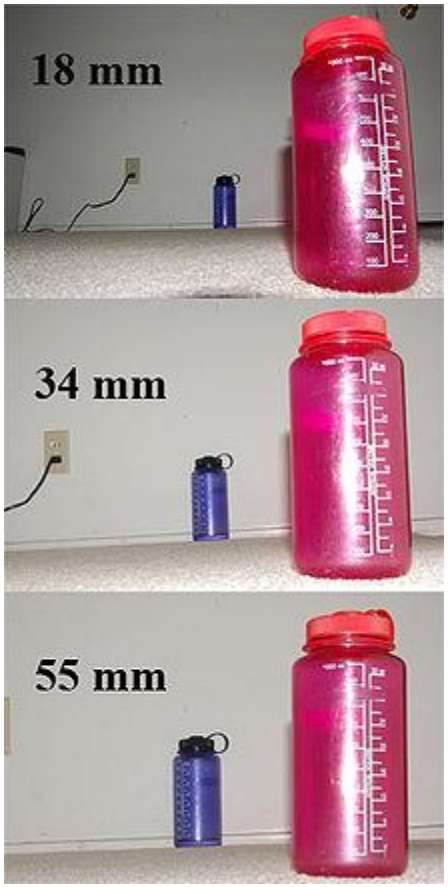
$$fov = \tan^{-1} \frac{d}{2f}$$

Focal length



Dolly Zoom or “Vertigo Effect”

http://www.youtube.com/watch?v=NB4bikrNzMk



How is this done?

Zoom in while moving away

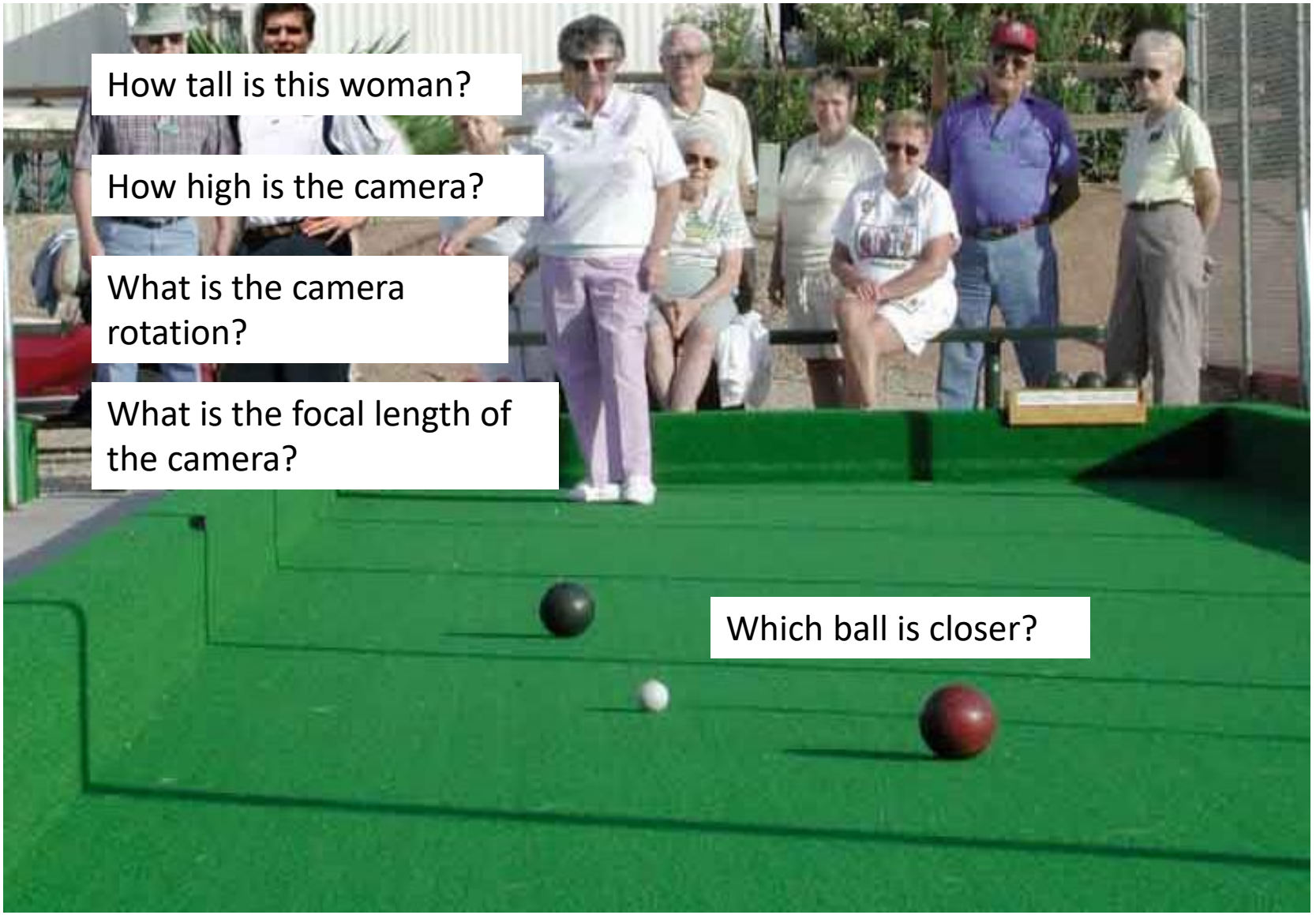
http://en.wikipedia.org/wiki/Focal_length

Poltergeist (1982)

Variables that affect exposure

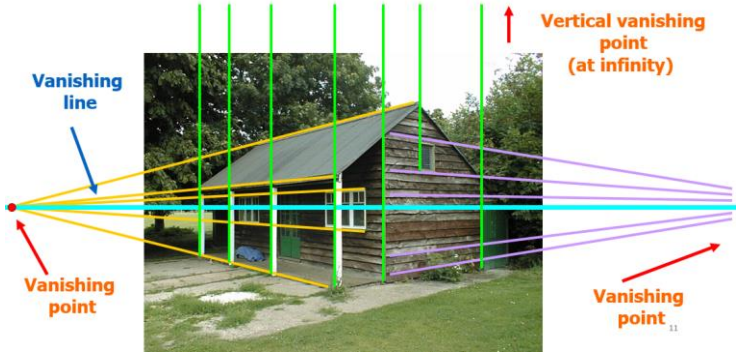
- http://graphics.stanford.edu/courses/cs178-10/applets/exposure.html

Review

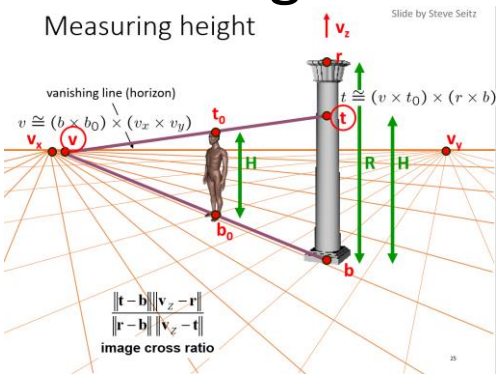

How tall is this woman?
How high is the camera?
What is the camera rotation?
What is the focal length of the camera?
Which ball is closer?

Things to remember

- Calibrate the camera?
  - Use an object with known geometry
  - Use vanishing points



- Measure the size of objects in the world from an image?
  - Use perspective cues



- Camera properties
  - focal length,
  - field of view,
  - depth of field,
  - aperture,
  - f-number?

Next class

- Image stitching