# Lecture 16: Image formation

# **Announcements: coronavirus edition**

- Coming to class is optional now!
- - May stream in future
- Project proposal due
- PS7 out tonight
- PS4 grades out

# • Will post lecture recordings after class



- Camera models
- Projection equations

# Today

### This section of the course: physical models

### The structure of ambient light











#### Ρ

X, Y, Z) Eye position





Angle

P(θ, φ, X, Y, Z)





P (θ, φ, λ, t, X, Y, Z) Wavelength, time





"The complete set of all convergence points constitutes the permanent possibilities of vision." Gibson 9

P (θ,  $\phi$ ,  $\lambda$ , t, X, Y, Z)





# Measuring the Plenoptic function



10 Source: Freeman, Torralba, Isola



### Image formation



#### Let's design a camera - Idea 1: put a piece of film in front of an object - Do we get a reasonable image? – No. This is a bad camera.





Add a barrier to block off most of the rays This reduces blurring - The opening known as the aperture – How does this transform the image?





scene to strike each point of the paper.

Source: Freeman, Torralba, Isola























http://www.foundphotography.com/PhotoThoughts/archives/2005/04/pinhole\_camera\_2.html











# Shrinking the aperture



2 mm



0.6mm

- - Less light gets through
  - *Diffraction* effects...

1 mm

0.35 mm

# • Why not make the aperture as small as possible?



## Shrinking the aperture





0.6mm



0.15 mm

1 mm

0.35 mm



# 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



# The eye



#### The human eye is a camera

- Iris colored annulus with radial muscles
- Pupil the hole (aperture) whose size is controlled by the iris
  What's the "film"?
  - photoreceptor cells (rods and cones) in the retina



#### **Eyes in nature:** eyespots to pinhole camera





http://upload.wikimedia.org/wikipedia/commons/6/6d/Mantis\_shrimp.jpg





Fifth row: 1 Great horned owl. 2 Mountain lion. 3 Boa constrictor. 4 Pufferfish. 5 African crested crane.



http://www.telegraph.co.uk/news/earth/earthpicturegalleries/7598120/Animal-eyes-quiz-Can-you-work-out-which-creatures-these-are-from-their-eyes.html?image=25 <u>ge=25</u> 25 Source: N. Snavely



- Top row: 1 Bengal tiger. 2 Asian elephant. 3 Zebra. 4 Chimpanzee. 5 Flamingo.
- Second row: 1 Domestic cat. 2 Hairless sphynx cat. 3 Grey wolf. 4 Booted eagle. 5 Iguana.
- Fourth row: 1 Lioness. 2 Bearded dragon (a type of lizard). 3 Leaf-tailed gecko. 4 Macaroni penguin. 5

















# Accidental pinhole camera



Source: Freeman, Torralba, Isola







#### Window turned into a pinhole



#### View outside







#### Window open



#### Window turned into a pinhole

Source: Freeman, Torralba, Isola








See Zomet, A.; Nayar, S.K. CVPR 2006 for a detailed analysis.

Source: Freeman, Torralba, Isola





Source: Freeman, Torralba, Isola



## Mixed accidental pinhole and anti-pinhole cameras



## Mixed accidental pinhole and anti-pinhole cameras



Source: Freeman, Torralba, Isola



### Mixed accidental pinhole and anti-pinhole cameras

Room with a window







41

Source: Freeman, Torralba, Isola



# Mixed accidental pinhole and anti-pinhole cameras

Body as the occluder



View outside the window









Source: Freeman, Torralba, Isola

### Looking for a small accidental occluder

### Body as the occluder



### Hand as the occluder



### View outside the window





### **Dimensionality Reduction Machine (3D to 2D)**

### 3D world



Point of observation

### What have we lost?

- Angles
- Distances (lengths)  $\bullet$



Slide by A. Efros Figures © Stephen E. Palmer, 2002



## Projection





## Projection







### Müller-Lyer Illusion



http://www.michaelbach.de/ot/sze\_muelue/index.html



### Geometric Model: A Pinhole Camera

**Principal point** 





Source: N. Snavely Figure credit: Peter Hedman



## Modeling projection

### • The coordinate system

- We use the pinhole model as an approximation
- Put the optical center (aka Center of Projection, or COP) at the origin
- Put the Image Plane (aka Projection Plane) in front of the COP
- The camera looks down the *positive* z-axis, and the y-axis points down
  - we like this if we want right-handed-coordinates
  - other versions are possible (e.g., OpenGL)





## Modeling projection

### Projection equations

- Compute intersection with PP of ray from (x,y,z) to COP
- Derived using similar triangles \_\_\_\_

$$(x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z}, f)$$

 We get the projection by throwing out the last coordinate:

$$(x, y, z) \to (f\frac{x}{z}, f\frac{y}{z})$$



51 Source: N. Snavely



### Perspective projection



### 

Similar triangles: y / f = Y / Z

y = f Y/Z

How can we represent this more compactly?

Source: Freeman, Torralba, Isola



### Homogeneous coordinates for affine transformations

## Trick: add one more coordinate: $(x, y) \Rightarrow \begin{vmatrix} x \\ y \\ 1 \end{vmatrix}$

homogeneous image coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$







### Translation

Solution: homogeneous coordinates to the rescue





### **Affine transformations**

# $\mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \longleftrightarrow \text{ any transformation represented by a 3x3}$



matrix with last row [001 ] we call an *affine* 

Source: N. Shavely



### **Basic affine transformations**

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$
Translate

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

2D *in-plane* rotation

$$\begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ 1 \end{bmatrix}$$
  
Scale

$$\begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & sh_x & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ 1 \end{bmatrix}$$
Shear

Source: N. Shavely



## **Modeling projection**

- Is this a linear transformation?
  - no—division by z is nonlinear

Homogeneous coordinates to the rescue!

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

### homogeneous image coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$
  
homogeneous scene

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$



### **Perspective Projection**

## Projection is a matrix multiply using homogeneous coordinates: $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} =$

### This is known as **perspective projection**

- The matrix is the projection matrix
- this)

$$= \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

divide by third coordinate

(Can also represent as a 4x4 matrix – OpenGL does something like



## **Perspective Projection**

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$
Scale by f: 
$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Scaling a projection matrix produces an equivalent projection matrix!

How does scaling the projection matrix change the transformation?

$$= \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

$$= \begin{bmatrix} fx\\ fy\\ z \end{bmatrix} \Rightarrow (f\frac{x}{z}, f\frac{y}{z})$$



## Orthographic projection

- Special case of perspective projection
  - Distance from the COP to the PP is infinite



- Good approximation for telephoto optics
- Also called "parallel projection":  $(x, y, z) \rightarrow (x, y)$
- What's the projection matrix?



ction nfinite

tics ⁄, z) → (x,



## Orthographic projection



![](_page_60_Picture_2.jpeg)

![](_page_60_Picture_3.jpeg)

61 Source: N. Snavely

![](_page_60_Picture_5.jpeg)

### Perspective projection

![](_page_61_Picture_1.jpeg)

![](_page_61_Picture_2.jpeg)

![](_page_61_Picture_3.jpeg)

## **Projection properties**

- Many-to-one: any points along same ray map to same point in image
- Points  $\rightarrow$  points
- Lines  $\rightarrow$  lines (collinearity is preserved)
  - But line through focal point projects to a point
- Planes  $\rightarrow$  planes (or half-planes) But plane through focal point projects to line

![](_page_62_Picture_6.jpeg)

## **Projection properties**

- Parallel lines converge at a vanishing point
  - Each direction in space has its own vanishing point
  - But lines parallel to the image plane remain parallel

![](_page_63_Figure_4.jpeg)

### a vanishing point ts own vanishing point e plane remain parallel

![](_page_63_Picture_6.jpeg)

![](_page_63_Picture_7.jpeg)

### **Camera parameters**

How can we model the geometry of a camera?

![](_page_64_Picture_2.jpeg)

Three important coordinate systems:

- World coordinates
- *Camera* coordinates 2.
- Image coordinates З.

How do we project a given world point (x, y, z) to an image point?

![](_page_64_Picture_9.jpeg)

### "The World"

![](_page_64_Picture_11.jpeg)

![](_page_65_Figure_0.jpeg)

### World coordinates

### **Camera coordinates**

### Image coordinates

Source: N. Snavely Figure credit: Peter Hedman

![](_page_65_Picture_5.jpeg)

### **Camera parameters**

- First transform (x, y, z) into camera coordinates
- Need to know
  - Camera position (in world coordinates)
  - Camera orientation (in world coordinates)
- Then project into the image plane to get *image (pixel)* coordinates
  - Need to know camera *intrinsics*

## To project a point (x, y, z) in world coordinates into a camera

![](_page_66_Picture_10.jpeg)

### **Camera parameters**

A camera is described by several parameters

- Rotation R of the image plane  $\bullet$
- focal length f, principal point (c<sub>x</sub>, c<sub>y</sub>), pixel aspect size α
- $\bullet$

**Projection equation** 

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

- The projection matrix models the cumulative effect of all ulletparameters
- Useful to decompose into a series of operations •  $\begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ intrinsics \end{bmatrix} prc$$

- $\bullet$ 
  - especially intrinsics—varies from one book to another

### Translation T of the optical center from the origin of world coords blue parameters are called "extrinsics," red are "intrinsics" x' $\neg [X]$

$$\begin{array}{c|c} * & Y \\ * & Z \\ * & 1 \end{array} = \mathbf{\Pi} \mathbf{X}$$

![](_page_67_Figure_17.jpeg)

![](_page_67_Picture_18.jpeg)

### **Projection matrix**

![](_page_68_Figure_1.jpeg)

![](_page_68_Figure_2.jpeg)

![](_page_68_Picture_3.jpeg)

### Extrinsics

 How do we get the camera to "canonical form"? points up, z-axis points backwards)

![](_page_69_Figure_2.jpeg)

# - (Center of projection at the origin, x-axis points right, y-axis

Step 1: Translate by -c

![](_page_69_Figure_5.jpeg)

![](_page_69_Picture_6.jpeg)

### Extrinsics

 How do we get the camera to "canonical form"? points up, z-axis points backwards)

![](_page_70_Figure_2.jpeg)

# - (Center of projection at the origin, x-axis points right, y-axis

### Step 1: Translate by -c

How do we represent translation as a matrix multiplication?

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{3\times3} & -\mathbf{C} \\ 0 & 0 & 0 \end{bmatrix}$$

![](_page_70_Picture_7.jpeg)

### Extrinsics

y-axis points up, z-axis points backwards)

![](_page_71_Figure_2.jpeg)

### How do we get the camera to "canonical form"? - (Center of projection at the origin, x-axis points right,

![](_page_71_Figure_4.jpeg)

![](_page_71_Picture_5.jpeg)
### Extrinsics

y-axis points up, z-axis points backwards)



# How do we get the camera to "canonical form"? - (Center of projection at the origin, x-axis points right,

Step 1: Translate by -c Step 2: Rotate by **R** 



(with extra row/column of [0 0 0 1])





- $\mathcal{U}$ : **aspect ratio** (1 unless pixels are not square)
- S : **skew** (0 unless pixels are shaped like rhombi/parallelograms)

Source: N. Snavely  $(c_x, c_y)$ : principal point ((w/2,h/2) unless optical axis doesn't intersect projection plane at image ceriter)

# $\begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

(converts from 3D rays in camera (intrinsics) coordinate system to pixel coordinates)

in general,  $\mathbf{K} = \begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix}$  (upper triangular matrix)



# **Typical intrinsics matrix** $\mathbf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}$

- length) and a translation by  $(c_x, c_y)$  (principal point)
- Maps 3D rays to 2D pixels

# 2D affine transform corresponding to a scale by f (focal



### Focal length

#### Can think of as "zoom"



24mm



#### 200mm Also related to field of view





50mm







#### Wide angle

#### Standard

#### Telephoto





http://petapixel.com/2013/01/11/how-focal-length-affects-your-subjects-apparent-weight-as-seen-with-a-cat/



# **Projection matrix**





### **Projection matrix**







### **Projection matrix**







### **Perspective distortion**

 Problem for architectural photography: converging verticals





82

# Perspective distortion

 Problem for architectural photography: converging vertic

> Tilting the camera upwards results in converging verticals

#### Solution: view camera (lens shifted w.r.t. film)

http://en.wikipedia.org/wiki/Perspective\_correction\_lens



Keeping the camera level, with an ordinary lens, captures only the bottom portion of the building



Shifting the lens upwards results in a picture of the entire subject



# **Perspective distortion**

What does a sphere project to?



Image source: F. Durand



### Distortion



No distortion

- Radial distortion of the image
  - Caused by imperfect lenses
  - edge of the lens

Pin cushion

Barrel

- Deviations are most noticeable for rays that pass through the









# **Modeling distortion**

 $(\hat{x}, \hat{y}, \hat{z})$ Project to "normalized" image coordinates

Apply radial distortion

Apply focal length translate image center

To model lens distortion

$$\begin{array}{rcl} x'_n &=& \widehat{x}/\widehat{z} \\ y'_n &=& \widehat{y}/\widehat{z} \end{array}$$

$$r^{2} = x'_{n}^{2} + y'_{n}^{2}$$
  

$$x'_{d} = x'_{n}(1 + \kappa_{1}r^{2} + \kappa_{2}r^{4})$$
  

$$y'_{d} = y'_{n}(1 + \kappa_{1}r^{2} + \kappa_{2}r^{4})$$

$$\begin{aligned} x' &= fx'_d + x_c \\ y' &= fy'_d + y_c \end{aligned}$$

- Use above projection operation instead of standard projection matrix multiplication



# **Correcting radial distortion**





#### from <u>Helmut Dersch</u>



# Next lecture: More geometry!

