## Intro \& <br> Cameras

EECS 442 - David Fouhey
Winter 2023, University of Michigan

https://web.eecs.umich.edu/~fouhey/teaching/EECS442_W23/

## Welcome!

## Goals of Computer Vision

## Get a computer to understand



## Goal: Naming



## Goal: Naming



The picture shows a building with many windows and grass in front of it. There is a person walking on the righto.

## Goal: 3D



## Goal: Actions



## Seems Obvious, Right?

- Key concept to keep in mind throughout the course: you see with both your eyes and your brain.

Why is it Hard?


Why is it Hard?


## Why is it Hard?

| 097 | 097 | 097 | 097 | 097 | 097 | 097 | 097 | 096 | 097 | 097 | 096 | 096 | 096 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 100 | 100 | 100 | 100 | 100 | 100 | 101 | 101 | 102 | 101 | 100 | 100 | 100 | 099 |
| 105 | 105 | 105 | 105 | 105 | 105 | 105 | 103 | 102 | 102 | 101 | 103 | 104 | 105 |
| 109 | 109 | 109 | 109 | 109 | 110 | 107 | 118 | 145 | 132 | 120 | 112 | 106 | 103 |
| 113 | 113 | 113 | 112 | 112 | 113 | 110 | 129 | 160 | 160 | 164 | 162 | 157 | 151 |
| 118 | 117 | 118 | 123 | 119 | 118 | 112 | 125 | 142 | 134 | 135 | 139 | 139 | 175 |
| 123 | 121 | 125 | 162 | 166 | 157 | 149 | 153 | 160 | 151 | 150 | 146 | 137 | 168 |
| 127 | 127 | 125 | 168 | 147 | 117 | 139 | 135 | 126 | 147 | 147 | 149 | 156 | 160 |
| 133 | 130 | 150 | 179 | 145 | 132 | 160 | 134 | 150 | 150 | 111 | 145 | 126 | 121 |
| 138 | 134 | 179 | 185 | 141 | 090 | 166 | 117 | 120 | 153 | 111 | 153 | 114 | 126 |
| 144 | 151 | 188 | 178 | 159 | 154 | 172 | 147 | 159 | 170 | 147 | 185 | 105 | 122 |
| 152 | 157 | 184 | 183 | 142 | 127 | 141 | 133 | 137 | 141 | 131 | 147 | 144 | 147 |
| 130 | 147 | 185 | 180 | 139 | 131 | 154 | 121 | 140 | 147 | 107 | 147 | 120 | 128 |
| 035 | 102 | 194 | 175 | 149 | 140 | 179 | 128 | 146 | 168 | 096 | 163 | 101 | 125 |

## Despite This, We’ve Made Progress

- Few of these problems are solved (and there are lots of dangers to pretending things are solved when they aren't)
- But we do have systems with performance ranging from non-embarrassing to super-human (with the right caveats)


## Look at Your Phone



## Graphics

## https://affinelayer.com/pixsrv/

INPUT


## Stable Diffusion



## Recognition



He et al. Mask RCNN. ICCV 2017.
Video Credit: Karol Majek (https://www.youtube.com/watch?v=OOT3UIXZztE)

## 3D

## Administrivia

## Meeting and Communication

- Class: Mon/Wed 10:30AM-Noon, STAMPS and Zoom + Recordings (but please give it a minute!)
- Discussions (6), attend any, starting Monday January 9. One will be recorded.
- Office Hours (a bunch, mainly in-person): Poll out this afternoon - please fill in
- Piazza: Sign up (link on canvas). We monitor but don't guarantee instant responses.
- Direct Email: Avoid. Why?


## Class Size, Waitlist, Modalities

- Thing to keep in mind: there are 280 students, 1 professor, 6 GSIs and 1 IA
- Waitlist: I'm limited by staff. I don't reorder the waitlist - each person has a good reason and it's a zero-sum game. Contact the advising office



## Office Hours + Piazza

- 140 hours of work available for the course across me, GSIs, IAs. Needs to cover:
- Lectures + Discussion: ~10 hours of standing-there time
- Writing, testing, debugging, and setting up grading for HW; preparing for lectures + discussions; conceptual Qs; course administration
- Helping students learn the material + do HW
- Lining up for a personal zoom room with a GSI for debugging / debugging via piazza is inefficient and makes everyone want to tear their hair out. Why?


## Let's Work Together

- Please watch the lectures + discussions and then do the HWs; please don't do the reverse.
- We'll try to teach you how to debug better. Debugging well is a hard skill, but it will be important for your career.
- Work together from the beginning.
- I really want to do and encourage modalities for OH that are more efficient


## Zoom, Recordings, etc.

- I'll attempt to monitor zoom and in-person at the same time. It's trickier than it looks!
- Recordings may take a little bit of time to get up.
- Anticipate at least one tech screw-up during the semester


## Websites

- Canvas: Links to other stuff and turning in big files
- Course website: slides, assignments, syllabus https://web.eecs.umich.edu/~fouhey/teaching/EEC S442 W23/.
- Piazza: ask questions, answer questions https://piazza.com/umich/winter2023/eecs442
- Gradescope: details later with release of HW1


## Evaluation

- Homework ( $5 \times 12 \%$ ) - five mini-project homeworks with a writeup and code
- Midterm (20\%) - One in-class midterm
- Project (2\% [proposal] +18\% [report+presentation]) - a semester-long project done in a team


## Homework Philosophy

- Tons of fantastic resources for many things so teaching you to setup stable diffusion is a waste
- Few components remain the same in the long run. I want to teach these.
- Some work may be frustrating or hard to wrap your head around; please expect some of that
- This may be your first ULCS. Be patient and build in some transition time
- Homework =/= spec. We'll often ask you to interpolate details, figure things out, etc.


## Evaluation: Homework Late Policy

- Penalty: 1\% per hour, rounded to nearest
- Example:
- Due: 5pm Tuesday.
- Submitted 5:15pm Tue: No penalty!
- Submitted 12:15am Wed: 7\% penalty (i.e., 90\% -> 83\%)
- Everyone gets 120 free late hours, applied automatically and optimally. These waive late penalties. Assume you will need these
- Exceptions for exceptional situations. Contact us.


## Copying

- Read the syllabus - don't look at peoples' code, no pair programming
- We will run MOSS
- Submit it late (that's why we have late days), halfworking (that's why we have partial credit), or take the zero on the homework - I guarantee you won't care about one bad homework in a year
- If you're overwhelmed, talk to us


## Evaluation: Term Project

- Work in a team of 3-4 to do something cool
- There will be a piazza thread for pairing up
- Could be:
- Applying vision to a problem you care about
- Independent re-implementation of a paper
- Trying to build and extend an approach
- An idea that we give you
- Should be 1 homeworks worth of work per person


## Doing Well in 442 - Work Together

Enjoy each topic like a story. Discuss with friends about homework. Check piazza for similar questions, ask for help or get inspiring ideas on piazza as well. Go to discussions - Siyi Chen
Exchanging advice and discussing concepts with your peers is a valuable learning opportunity and something that we want you to do (as long as it's within the limits of the Honor Code) - Ahmed Khan
+1 on finding people to work with. It made the class much more enjoyable. - Jacob Skwirsk

Don't have friends in the class?
We'll introduce you to people (stay tuned)

## Doing Well in 442 - Start Early



Start early. - Nikhil Devraj

Start early, especially for any projects using colab and gpus. - Rahul Gupta

We're trying to figure out how to optimally handle the GPU situation.

## How To Have a Bad Time in 442

- Bad time: start homeworks with three days to go; have conceptual issues and debugging fun
- Bad time: ignore lectures; just-in-time lecturewatching by scrubbing through lectures while doing homeworks
- Bad time: debug by piazza +OH queue


## Your Todos

- Office hours poll if you have preferences
- Study buddy matching poll if you'd like us to help you match
- Background poll if you'd like to help me tailor the course (and future offerings)
- Get to know you poll if you'd like
- First homework out next week. No need to do anything. Just relax


## Cameras

## Let's Take a Picture!



## Idea 1: Just use film

Result: Junk

## Let's Take a Picture!



## Idea 2: add a barrier

## Let's Take a Picture!



## Idea 2: add a barrier

## Let's Take a Picture!



Film captures all the rays going through a point (a pencil of rays). Result: good in theory!

## Camera Obscura



- Basic principle known to Mozi (470-390 BCE), Aristotle (384-322 BCE)
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)


## Camera Obscura



After scouting rooms and reserving one for at least a day, Morell masks the windows except for the aperture. He controls three elements: the size of the hole, with a smaller one yielding a sharper but dimmer image; the length of the exposure, usually eight hours; and the distance from the hole to the surface on which the outside image falls and which he will photograph. He used $4 \times 5$ and $8 \times 10$ view cameras and lenses ranging from 75 to 150 mm .

After he's done inside, it gets harder. "I leave the room and I am constantly checking the weather, I'm hoping the maid reads my note not to come in, l'm worrying that the sun will hit the plastic masking and it will fall down, or that I didn't trigger the lens."

From Grand Images Through a Tiny Opening, Photo District News, February 2005


## Camera Obscura

## Useful for viewing solar eclipses!



Photo Credit: Justin Johnson

## Camera Obscura

## Useful for viewing solar eclipses!



Put your eye here

Pinhole: aluminum foil with a tiny hole


Justin on 8/21/2017

Photo Credit: Justin Johnson

## Camera Obscura

## Useful for viewing solar eclipses!



Photo of the sun

Photo Credit: Justin Johnson


View in the box


Justin on 8/21/2017

## Projection



## Projection



How do we find the projection $P$ of a point $X$ ?
Form visual ray from $X$ to camera center and intersect it with camera plane

## Projection



Both $X$ and $X^{\prime}$ project to $P$. Which appears in the image?

Are there points for which projection is undefined?

## Quick Aside: Remember This?



$$
\frac{a}{b}=\frac{d}{c} \longrightarrow a=\frac{b d}{c}
$$

## Projection Equations



Coordinate system: $\mathbf{O}$ is origin, XY in image, Z sticks out. $X Y$ is image plane, $Z$ is optical axis.
$(x, y, z)$ projects to ( $f x / z, f y / z$ ) via similar triangles

## Some Facts About Projection



3D lines project to 2D lines
The projection of any 3D parallel lines converge at a vanishing point

Distant objects are smaller


## Some Facts About Projection

Let's try some fake images

## Some Facts About Projection



## Some Facts About Projection



## Some Facts About Projection



## What's Lost?



## Is she shorter or further away?

Are the orange lines we see parallel / perpendicular / neither to the red line?

## What's Lost?

# Is she shorter or further away? 



## Are the orange lines we see parallel / perpendicular / neither to the red line?

## What's Lost?

Be careful of drawing conclusions:

- Projection of 3D line is 2D line; NOT 2D line is 3D line.
- Can you think of a counter-example (a 2D line that is not a 3D line)?
- Projections of parallel 3D lines converge at VP; NOT any pair of lines that converge are parallel in 3D.
- Can you think of a counter-example?


## Do You Always Get Perspective?



## Do You Always Get Perspective?



Y location of blue and red dots in image:

$$
\frac{f y}{z_{2}} \quad \frac{f y}{z_{1}} \quad \frac{f y}{z} \quad \frac{f y}{z}
$$

## Do You Always Get Perspective?



When plane is fronto-parallel (parallel to camera plane), everything is:

- scaled by $\mathrm{f} / \mathrm{z}$
- otherwise is preserved.


## What's This Useful For?



Things looking different when viewed from different angles seems like a nuisance. It's also a cue. Why?

## Projection Equation



I promised you linear algebra: is this linear? Nope: division by $z$ is non-linear

## Homogeneous Coordinates (2D)

## Trick: add a dimension!

This also clears up lots of nasty special cases

Physical
Point

$\left[\begin{array}{l}u \\ v \\ w\end{array}\right] \quad \underset{\text { Divide }}{ }$

Homogeneous
Point

by w

Physical Point
$\left[\begin{array}{l}u / w \\ v / w\end{array}\right]$

$$
w=1
$$

Concat

## Homogeneous Coordinates

$$
\lambda[x, y, w]
$$

Triple /
Equivalent
$\left[\begin{array}{l}u \\ v \\ w\end{array}\right] \equiv\left[\begin{array}{c}u^{\prime} \\ v^{\prime} \\ w^{\prime}\end{array}\right]$

Double /
Equals
$[x, y, w]$
$\leftrightarrow\left[\begin{array}{c}u \\ v \\ w\end{array}\right]=\lambda\left[\begin{array}{c}u^{\prime} \\ v^{\prime} \\ w^{\prime}\end{array}\right]$

$$
\lambda \neq 0
$$

Two homogeneous coordinates are equivalent if they are proportional to each other. Not = !

## Benefits of Homogeneous Coords

General equation of 2D line:

$$
a x+b y+c=0
$$

Homogeneous Coordinates

$$
\boldsymbol{l}^{T} \boldsymbol{p}=0, \quad \boldsymbol{l}=\left[\begin{array}{l}
a \\
b \\
c
\end{array}\right], \boldsymbol{p}=\left[\begin{array}{l}
x \\
y \\
1
\end{array}\right]
$$

## Benefits of Homogeneous Coords

- Lines (3D) and points (2D $\rightarrow 3 \mathrm{D}$ ) are now the same dimension.
- Use the cross ( x ) and dot product for:
- Intersection of lines I and $\mathbf{m}$ : I x m
- Line through two points $\mathbf{p}$ and $\mathbf{q}: \mathbf{p} \times \mathbf{q}$
- Point pon line I: $\mathbf{I}^{\top} p$
- Parallel lines, vertical lines become easy (compared to $y=m x+b$ )


## Benefits of Homogeneous Coords

What's the intersection?

[0,1,-2] x [1,0,-1] = [-1,-2,-1]
Converting back (divide by -1 )
$(1,2)$

## Benefits of Homogeneous Coords



Intersection of $\mathrm{y}=2, \mathrm{y}=1$

$$
[0,1,-2] \times[0,1,-1]=[1,0,0]
$$

Does it lie on $\mathbf{y}=3$ ? Intuitively?

$$
[0,1,-3]^{\top}[1,0,0]=0
$$

## Benefits of Homogeneous Coords

Translation is now linear / matrix-multiply

$$
\left.\begin{array}{rl}
\text { If } w=1 & {\left[\begin{array}{c}
u^{\prime} \\
v^{\prime} \\
w^{\prime}
\end{array}\right]}
\end{array}\right]\left[\begin{array}{lll}
1 & 0 & t_{x} \\
0 & 1 & t_{y} \\
0 & 0 & 1
\end{array}\right]\left[\begin{array}{c}
u \\
v \\
1
\end{array}\right]=\left[\begin{array}{c}
u+t_{x} \\
v+t_{y} \\
1
\end{array}\right] .\left[\begin{array}{c}
1 \\
\text { Generically }
\end{array}\left[\begin{array}{c}
u^{\prime} \\
v^{\prime} \\
w^{\prime}
\end{array}\right]=\left[\begin{array}{ccc}
1 & 0 & t_{x} \\
0 & 1 & t_{y} \\
0 & 0 & 1
\end{array}\right]\left[\begin{array}{c}
u \\
v \\
w+w t_{x} \\
v+w t_{y} \\
w
\end{array}\right] .\right.
$$

Rigid body transforms (rot + trans) now linear

$$
\left[\begin{array}{c}
u^{\prime} \\
v^{\prime} \\
w^{\prime}
\end{array}\right]=\left[\begin{array}{ccc}
r_{11} & r_{12} & t_{x} \\
r_{21} & r_{22} & t_{y} \\
0 & 0 & 1
\end{array}\right]\left[\begin{array}{c}
u \\
v \\
w
\end{array}\right]
$$

## 3D Homogeneous Coordinates

Same story: add a coordinate, things are equivalent if they're proportional

$$
\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right] \longrightarrow\left[\begin{array}{l}
u \\
v \\
w \\
t
\end{array}\right] \longrightarrow\left[\begin{array}{l}
u / t \\
v / t \\
w / t
\end{array}\right]
$$

## Projection Matrix

Projection ( $\mathrm{fx} / \mathrm{z}, \mathrm{fy} / \mathrm{z}$ ) is matrix multiplication


## Projection Matrix

Projection ( $f x / z, f y / z$ ) is matrix multiplication


$$
\text { Why } \equiv \neq=
$$



## Project $X$ and $X$ ' to the image and compare them

YES $\left[\begin{array}{c}f x \\ f y \\ z\end{array}\right] \equiv\left[\begin{array}{c}f x^{\prime} \\ f y^{\prime} \\ z^{\prime}\end{array}\right] \quad$ NO $\left[\begin{array}{c}f x \\ f y \\ z\end{array}\right]=\left[\begin{array}{c}f x^{\prime} \\ f y^{\prime} \\ z^{\prime}\end{array}\right]$

## Typical Perspective Model

P: 2D homogeneous point (3D)

X: 3d homogeneous point (4D)


## Typical Perspective Model

$\mathbf{R}$ : rotation between world system and camera
$P \equiv$
t: translation
between world system and camera

$\left[\begin{array}{lll}\boldsymbol{R}_{3 x 3} & \boldsymbol{t}_{3 x 1}\end{array}\right] \quad \boldsymbol{X}_{4 x 1}$

## Typical Perspective Model

f focal length

u0,v0: principal point (image coords of camera origin on retina)
$\left[\begin{array}{lll}\boldsymbol{R}_{3 x 3} & \boldsymbol{t}_{3 x 1}\end{array}\right] \quad \boldsymbol{X}_{4 x 1}$

## Typical Perspective Model

$$
\left.\begin{array}{c}
\begin{array}{c}
\text { Intrinsic } \\
\text { Matrix K }
\end{array} \\
\boldsymbol{P} \equiv \begin{array}{c}
\text { Extrinsic } \\
\text { Matrix }[\mathbf{R}, \mathbf{t}]
\end{array} \\
\left.\begin{array}{|ccc}
f & 0 & u_{0} \\
0 & f & v_{0} \\
0 & 0 & 1
\end{array}\right]
\end{array} \begin{array}{ll}
\boldsymbol{R}_{3 \times 3} & \boldsymbol{t}_{3 \times 1}
\end{array}\right] \boldsymbol{X}_{4 \times 1}
$$

## Other Cameras - Orthographic

Orthographic Camera (z infinite)

$$
\boldsymbol{P}=\left[\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{array}\right] \quad \boldsymbol{X}_{3 \times 1}
$$



## Other Cameras - Orthographic

Why does this make things easy and why is this popular in old games?

$$
\boldsymbol{P}=\left[\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{array}\right]\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right]
$$

## The Big Issue



Film captures all the rays going through a point (a pencil of rays).

## How big is a point?

## Math vs. Reality

- Math: Any point projects to one point
- Reality:
- Don't image points behind the camera / objects
- Don't have an infinite amount of sensor material
- Other issues
- Light is limited
- Spooky stuff happens with infinitely small holes


## Limitations of Pinhole Model

## Ideal Pinhole

-1 point generates 1 image
-Low-light levels


Finite Pinhole
-1 point generates region
-Blurry.
Why is it blurry?


## Limitations of Pinhole Model



## Adding a Lens



- A lens focuses light onto the film
- Thin lens model: rays passing through the center are not deviated (pinhole projection model still holds)


## Adding a Lens



- All rays parallel to the optical axis pass through the focal point


## What's The Catch?



- There's a distance where objects are "in focus"
- Other points project to a "circle of confusion"


## Thin Lens Formula

We care about images that are in focus. When is this true?
When two paths from a point hit the same image location.


## Thin Lens Formula

Let's derive the relationship between object distance $D$, image plane distance D', and focal length f.


## Thin Lens Formula



## Thin Lens Formula



## Thin Lens Formula

Set them equal:

$$
\frac{D^{\prime}}{D}=\frac{D-f}{f} \rightarrow \frac{1}{D}+\frac{1}{D^{\prime}}=\frac{1}{f}
$$



## Thin Lens Formula

Suppose I want to take a picture of a lion with D big? Which of $D, D^{\prime}, f$ are fixed?

How do we take pictures of things at different distances?


## Depth of Field


http://www.cambridgeincolour.com/tutorials/depth-of-field.htm

## Controlling Depth of Field



Changing the aperture size affects depth of field A smaller aperture increases the range in which the object is approximately in focus

## Controlling Depth of Field



If a smaller aperture makes everything focused, why don't we just always use it?

## Varying the Aperture



Small aperture = large DOF


Large aperture = small DOF

## Varying the Aperture



## Field of View (FOV)


$\tan ^{-1}$ is monotonic increasing. How can I get the FOV bigger?

## Field of View



Slide Credit: A. Efros

## Field of View



Slide Credit: A. Efros

## Field of View and Focal Length




Large FOV, small $f$
Camera close to car


Small FOV, large $f$
Camera far from the car

## Field of View and Focal Length


wide-angle

standard

telephoto

## Dolly Zoom

## Change $f$ and distance at the same time



## More Bad News!

- First a pinhole...
- Then a thin lens model....



## Lens Flaws: Radial Distortion

Lens imperfections cause distortions as a function of distance from optical axis


Less common these days in consumer devices

## Vignetting



## What happens to the light between the black and red lines?

## Vignetting



## Lens Flaws: Chromatic Abberation

Lens refraction index is a function of the wavelength. Colors "fringe" or bleed


## Lens Flaws: Chromatic Abberation

## Researchers tried teaching a network about objects by forcing it to assemble jigsaws.



Initial layout, with sampled patches in red
 is discarded


We can recover image layout automatically

## From Photon to Photo



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

- Each cell in a sensor array is a light-sensitive diode that converts photons to electrons
- Dominant in the past: Charge Coupled Device (CCD)
- Dominant now: Complementary Metal Oxide Semiconductor (CMOS)


## From Photon to Photo

## Rolling Shutter: pixels read in sequence Can get global reading, but \$\$\$



## Preview of What's Next

## Bayer grid



Demosaicing:

Estimation of missing components from neighboring values



Human Luminance Sensitivity Function

## For the Curious

- Cut in the interest of time


## Radial Distortion Correction



Ideal

$$
y^{\prime}=f \frac{y}{z} \quad y^{\prime}=\left(1+k_{1} r^{2}+\cdots\right) \frac{y}{z}
$$

## Lens Flaws: Spherical Abberation

Lenses don't focus light perfectly! Rays farther from the optical axis focus closer


## Historic milestones

- Pinhole model: Mozi (470-390 BCE), Aristotle (384-322 BCE)
- Principles of optics (including lenses):

Alhacen (965-1039 CE)

- Camera obscura: Leonardo da Vinci (1452-1519), Johann Zahn (1631-1707)
- First photo: Joseph Nicephore Niepce (1822)
- Daguerréotypes (1839)
- Photographic film (Eastman, 1889)
- Cinema (Lumière Brothers, 1895)
- Color Photography (Lumière Brothers, 1908)
- Television (Baird, Farnsworth, Zworykin, 1920s)
- First consumer camera with CCD

Sony Mavica (1981)

- First fully digital camera: Kodak DCS100 (1990)


Alhacen's notes


Niepce, "La Table Servie," 1822


Old television camera

## First digitally scanned photograph

- 1957, 176x176 pixels



## Historic Milestone

Sergey Prokudin-Gorskii (1863-1944) Photographs of the Russian empire (1909-1916)


## Historic Milestone



## Future Milestone

## Your job in homework 1: Make the left look like the right.



