CS4501: Introduction to Computer Vision

Cameras and Image Formation

Various slides from previous courses by:

D.A. Forsyth (Berkeley / UIUC), I. Kokkinos (Ecole Centrale / UCL). S. Lazebnik (UNC / UIUC), S. Seitz (MSR / Facebook), J. Hays (Brown / Georgia Tech), A. Berg (Stony Brook / UNC), D. Samaras (Stony Brook) . J. M. Frahm (UNC), V. Ordonez (UVA).

Make computers understand images and video

What kind of scene?

Where are the cars?

How far is the building?

…

Slide by James Hays.

Today’s Class

What is a camera?

Who invented cameras?

Image Formation

Brief Introduction to Projective Geometry (Computer Graphics)

About the Course

CS4501-001: Introduction to Computer Vision

- Instructor: Vicente Ordóñez
- Email: vicente@virginia.edu
- Website: http://vicenteordonez.com/vision/
- Class Location: **Olsson Hall 120** (Capacity 148)
- Class Times: **Monday-Wednesday 5pm - 6:15pm**
- Piazza: https://piazza.com/virginia/fall2019/cs4501001
- UVA Collab (for submitting assignments / quizzes / etc)

Pre-requisites

- Python programming skills
- Calculus / Linear Algebra / Probability

Office Hours

Vicente Ordonez
(vicente at virginia.edu)

Office Hours:
Tuesdays from
4pm to 6pm
(Rice Hall 310).

Ziyan Yang
(zy3cx at virginia.edu)

Office Hours:
Thursdays from
12:30pm to 2:30pm
(Rice Hall 442).

Paola Cascante-Bonilla
(pc9za at virginia.edu)

Office Hours:
Fridays from
2pm to 4pm
(Rice Hall 442)

Grading

- Assignments: 50% (5 assignments)
  (10% + 10% + 10% + 10% + 10%)
- Quiz: 20%   (2 quizzes)
  (10% + 10%)
- Final project: 30% (group project – 3 people max)

Late Submission Policy

- You shouldn’t submit late assignments
    - That being said there is a 3-day grace period, if you submit an assignment late (by one day), you will not be penalized. Any fraction of a day counts as a full day. You can have up to 3 days under the 3-day grace period. This doesn’t apply to the final project. It is 3-days for the entire semester not for individual assignments. These exceptions will only be accounted for at the end of the semester. If this policy is confusing please just don’t submit late assignments. Also, please don’t send emails about submitting late assignments.

Textbook

- *Computer Vision: Algorithms and Applications*
  by Richard Szeliski. http://szeliski.org/Book/

Computer Vision: A Modern Approach, Forsyth and Ponce, Prentice Hall 2011.

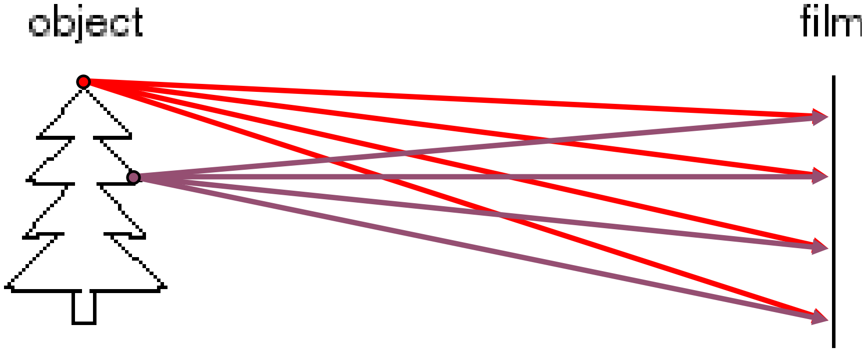
Introductory Techniques for 3D Computer Vision, Trucco and Verri, Prentice Hall 1998.

Computer Vision: Models, Learning, and Inference by S. Prince, Cambridge University Press, 2012: draft at http://www.computervisionmodels.com/

Cameras

What do you need to make a camera from scratch?

Image formation

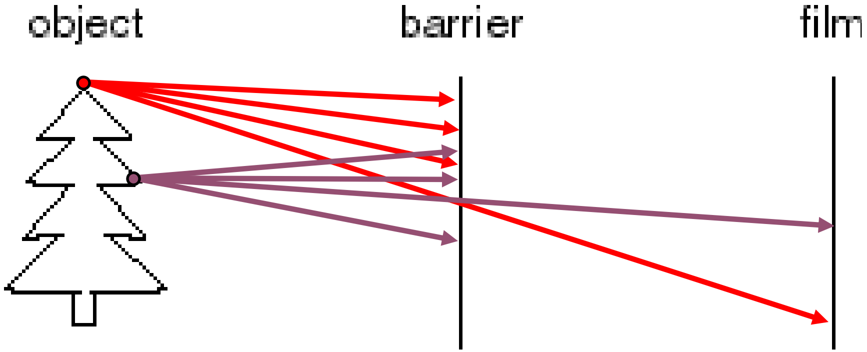


Let’s design a camera

  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Slide source: Seitz

Pinhole camera



Idea 2: add a barrier to block off most of the rays

- This reduces blurring
- The opening known as the **aperture**

Slide source: Seitz

Accidental Cameras

Accidental Pinhole and Pinspeck Cameras
Revealing the scene outside the picture.
Antonio Torralba, William T. Freeman

Accidental Cameras

a) Input (occluder present)

b) Reference (occluder absent)

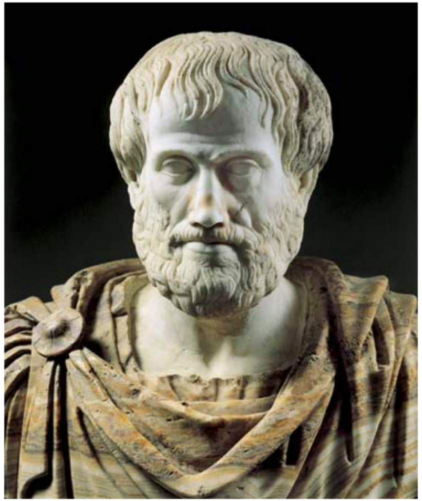
c) Difference image (b-a)

d) Crop upside down

e) True view

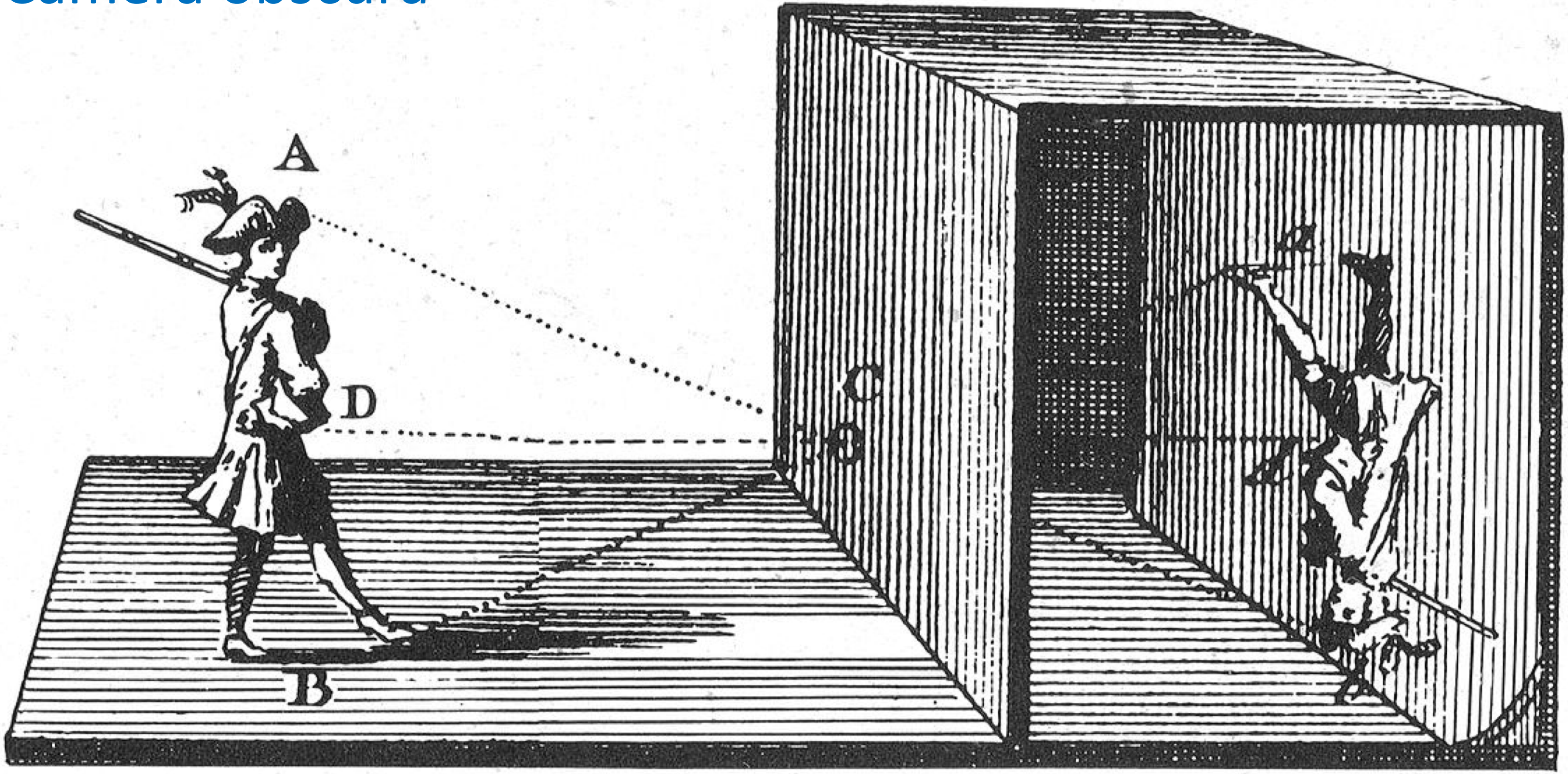
Known by the Greeks and the Chinese 470BC-322BC

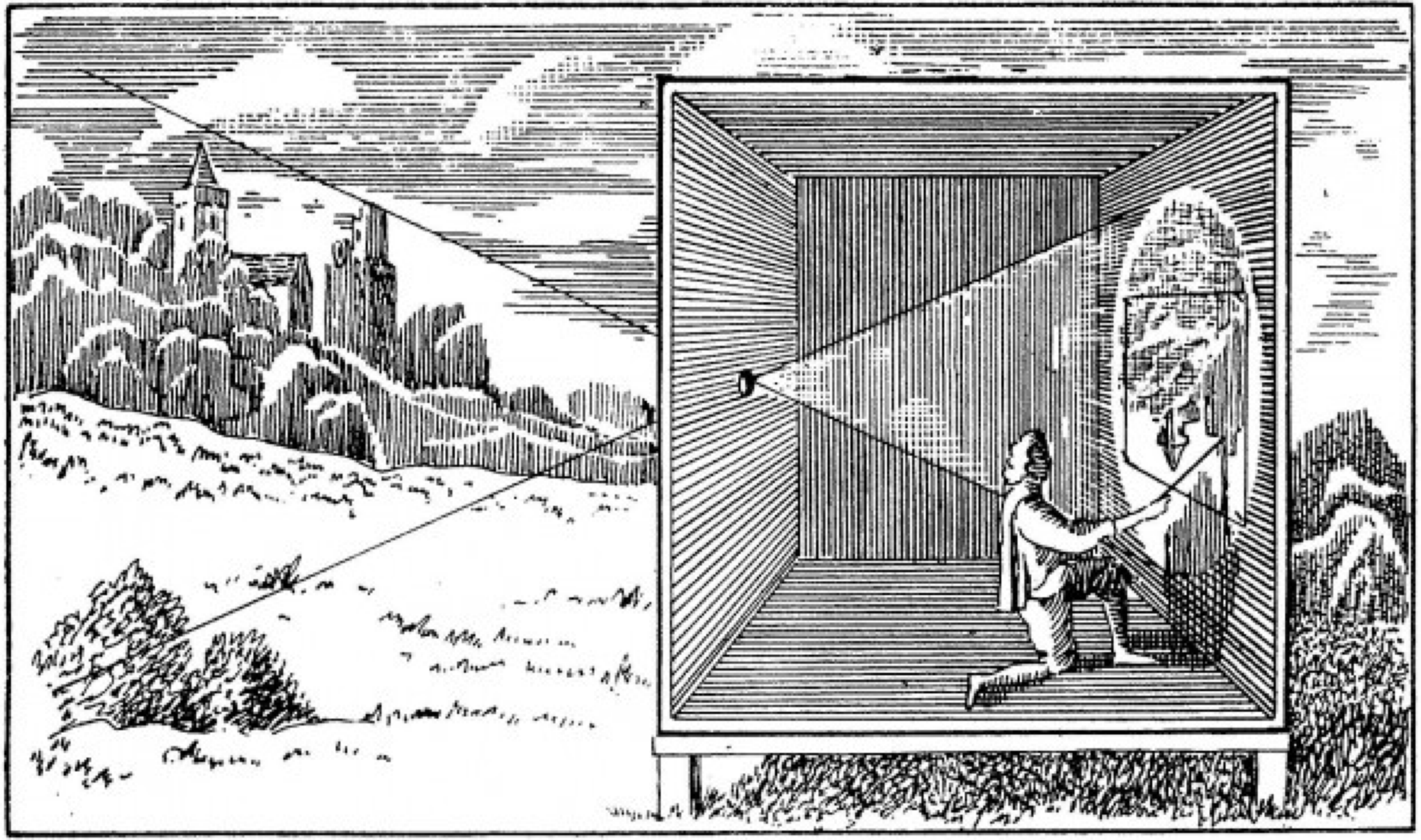
Recorded in writings by Chinese Philosopher Mozi (墨子)
470-390BCE



Recorded in writings by Aristotle or one of his disciples
384-322BCE

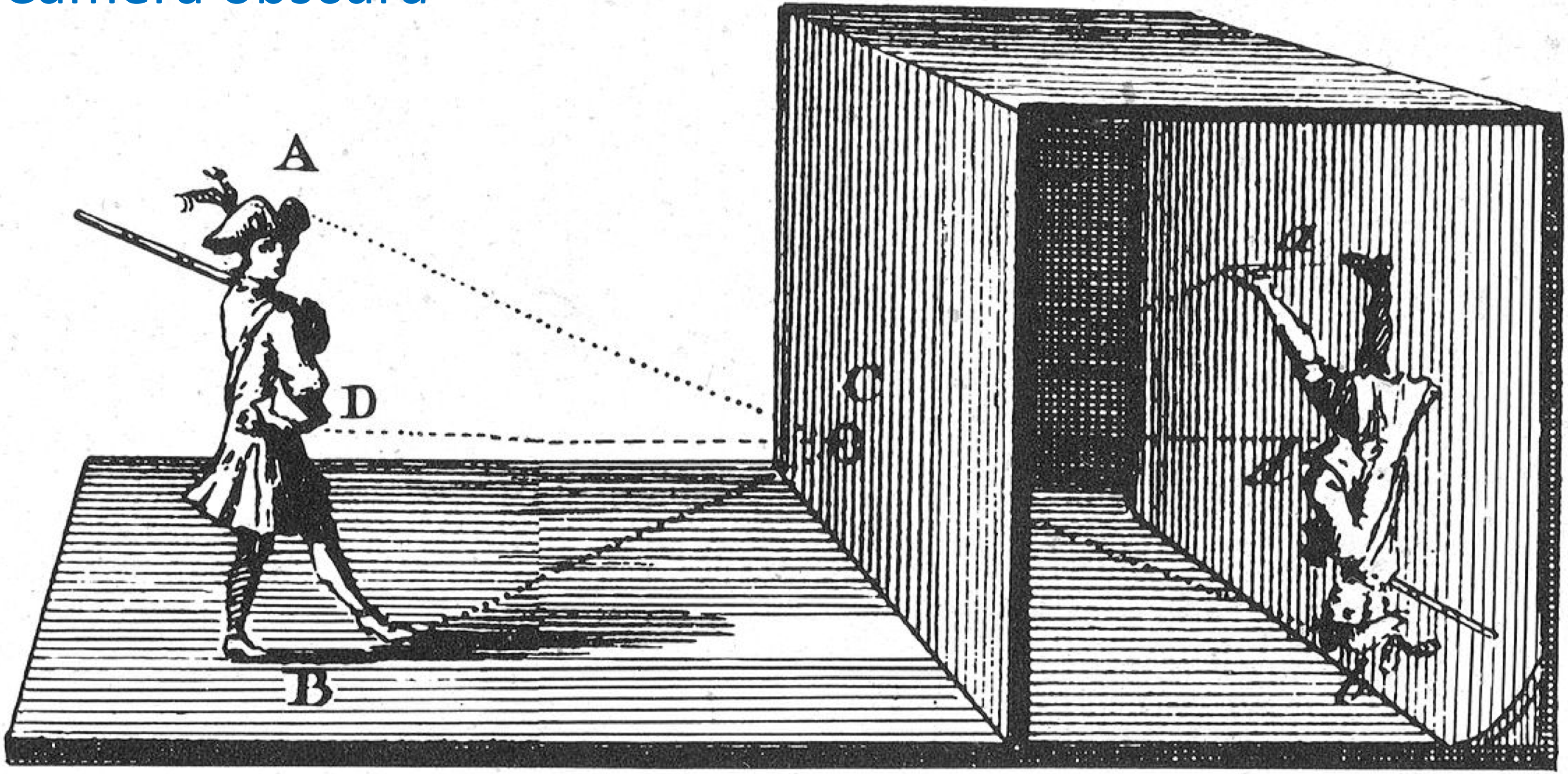
Camera obscura





https://pixsylated.com/blog/stepping-inside-havana-camera-obscura/

Camera obscura



Camera obscura: the pre-camera

Freestanding camera obscura at UNC Chapel Hill

Photo by Seth Ilys

Camera Obscura used for Tracing

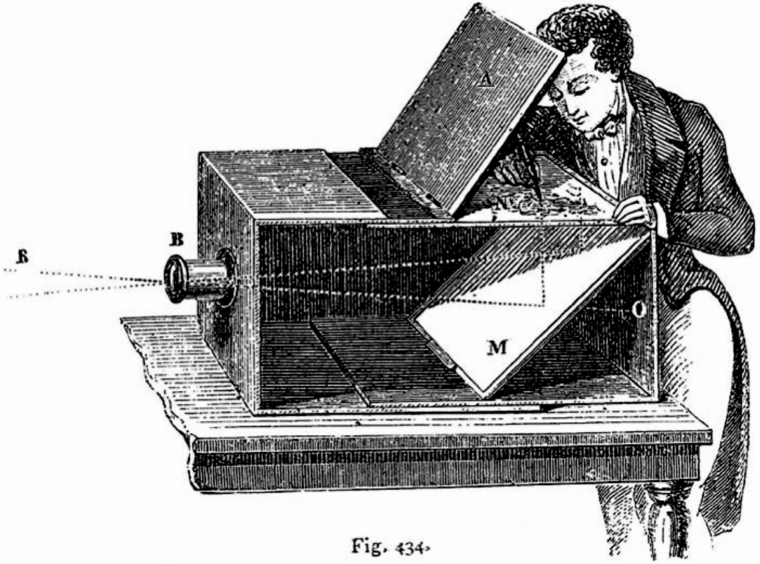
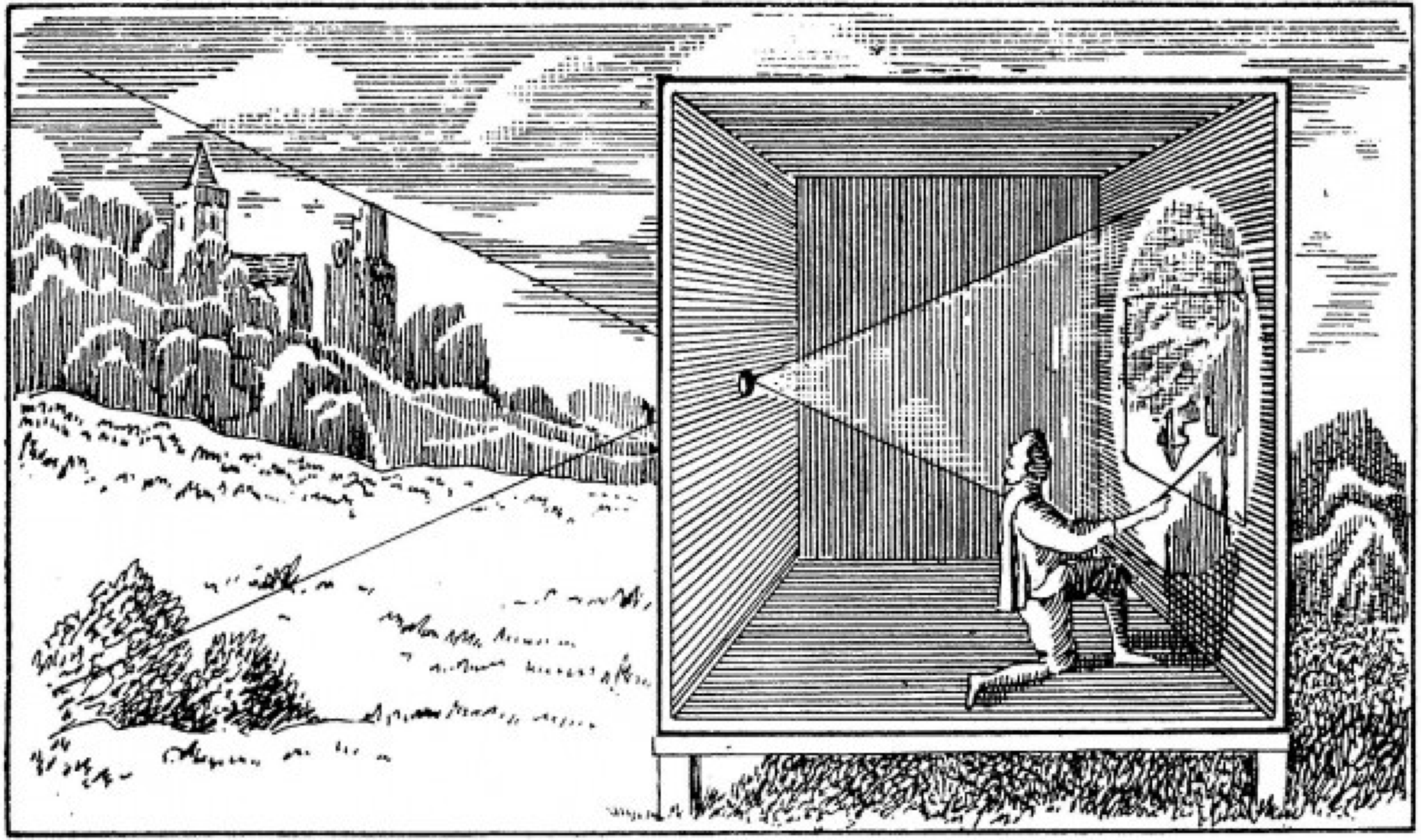


Fig. 434.

Lens Based Camera Obscura, 1568



https://pixsylated.com/blog/stepping-inside-havana-camera-obscura/

First Photograph

Oldest surviving photograph

- Took 8 hours on pewter plate

Photograph of the first photograph

Joseph Niepce, 1826

Stored at UT Austin

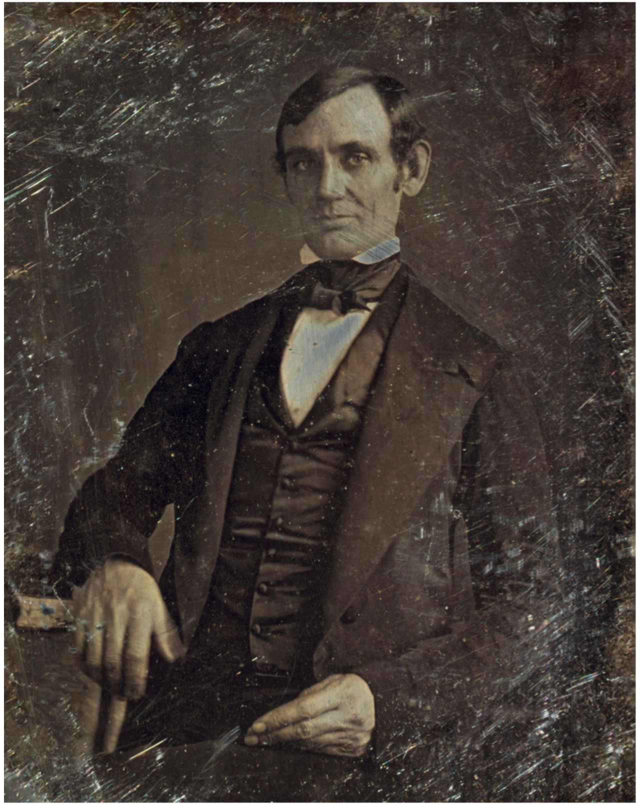
Bitumen of Judea: Naturally occurring asphalt that is photo-sensitive
https://en.wikipedia.org/wiki/Bitumen_of_Judea

From Joseph Niepce to Louis Daguerre

- Act 1: Joseph Niepce tells his nice idea to Louis Daguerre
- Act 2: Good friend Louis Daguerre improves idea and names it Daguerrotypes
- Act 3: Louis Daguerre makes history (and money).

1846 Daguerrotype of a young Abraham Lincoln

Daguerreotype camera built by La Maison Susse Frères in 1839, with a lens by Charles Chevalier

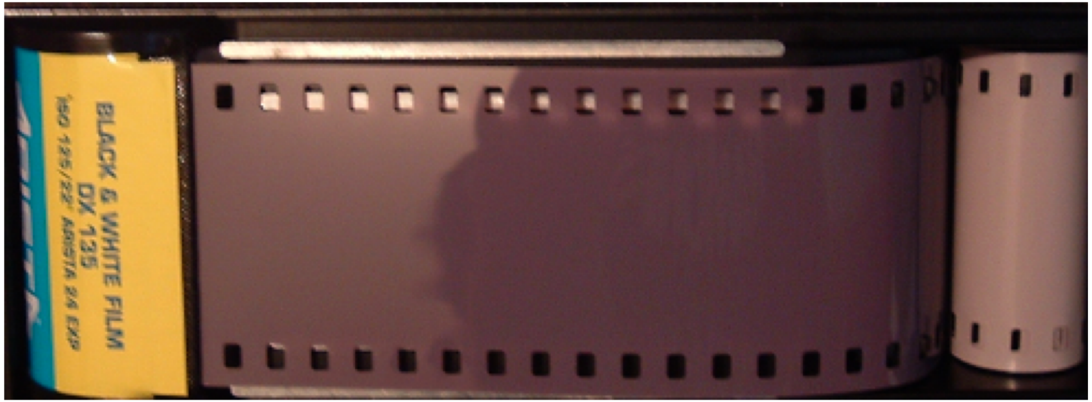


Hercules Florence’s *Photographie*

- Brazilian painter and inventor
- Before Daguerre but after Niepce
- Included the idea of negatives

George Eastman 1885 (Rochester, NY)

- Founder of pioneering Eastman Kodak Company (Kodak)
- Popularization of film photography (nitrate film)



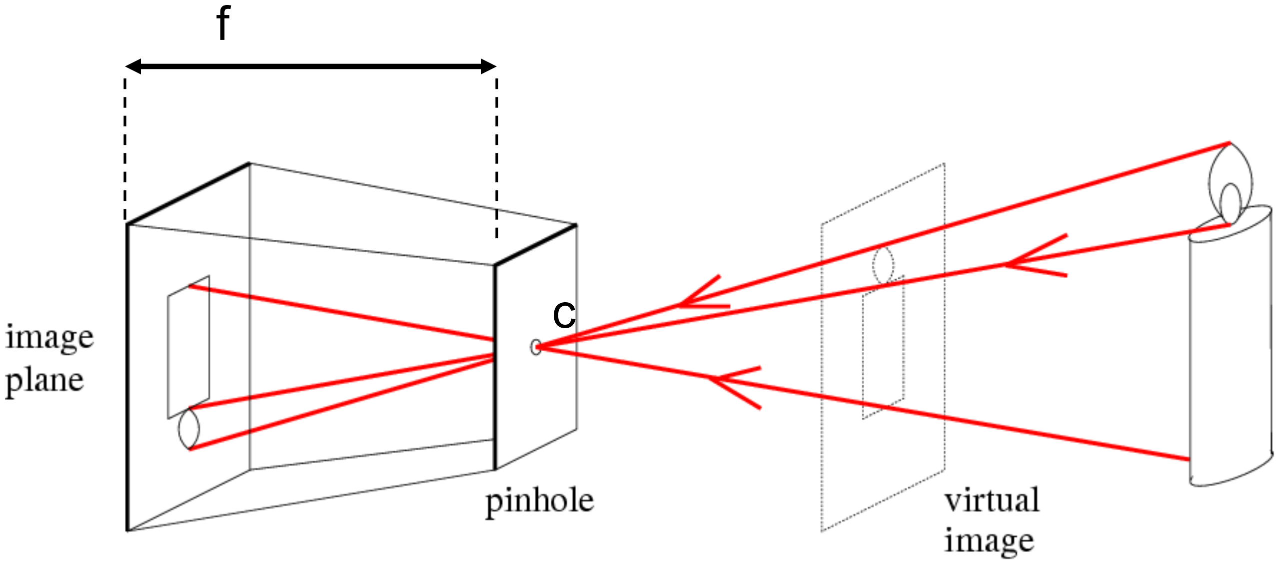
Should look familiar if you were born in the 80’s or earlier or if you are a true modern-day hipster!

- Photo negatives

So, who invented cameras?

Maybe the wrong question to ask?

Back to the Pinhole camera



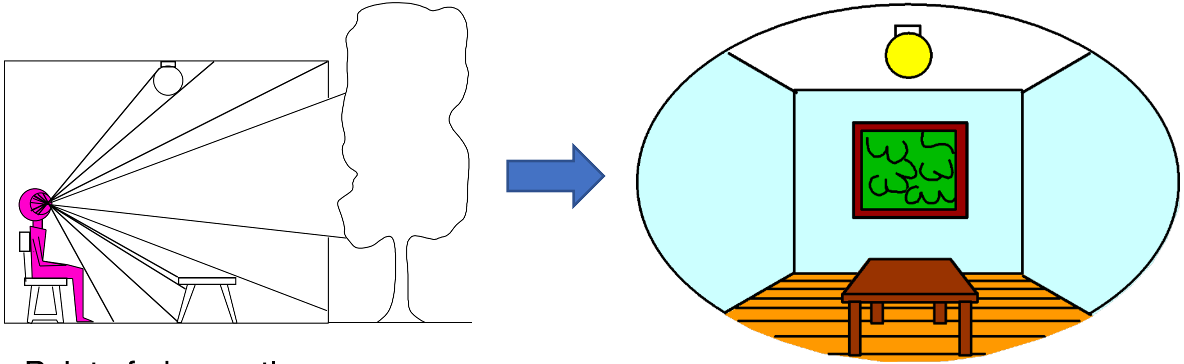
f = focal length
c = center of the camera

Figure from Forsyth

Dimensionality Reduction Machine (3D to 2D)

3D world

2D image

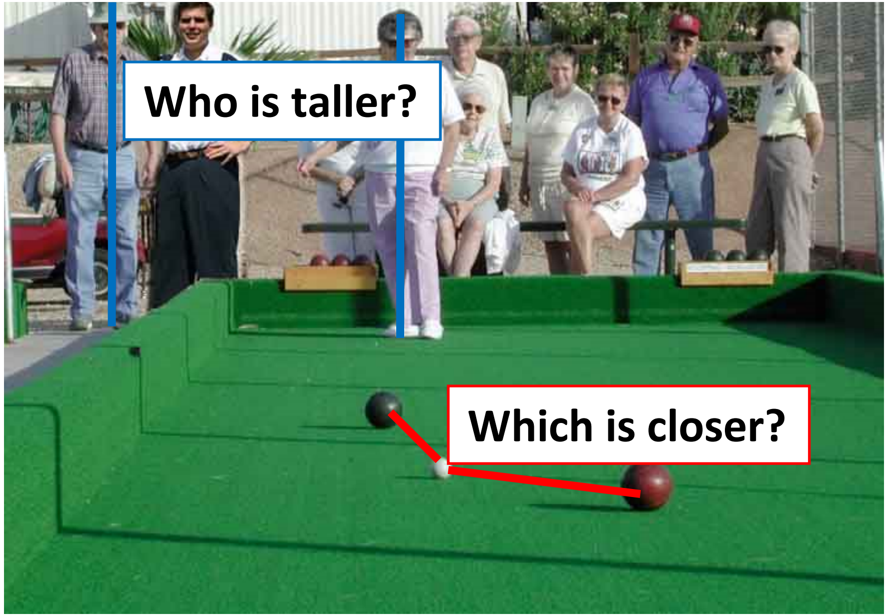


Point of observation

Projective Geometry

What is lost?

- Length



Length and area are not preserved

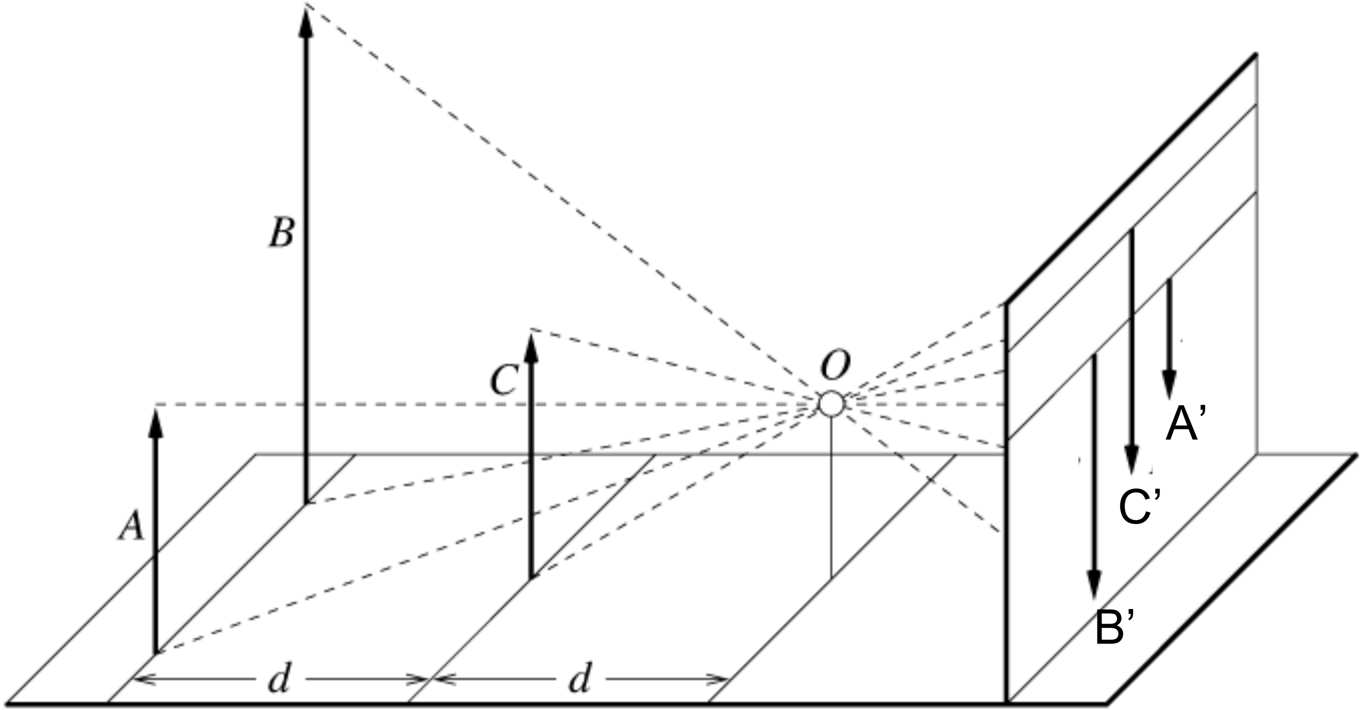
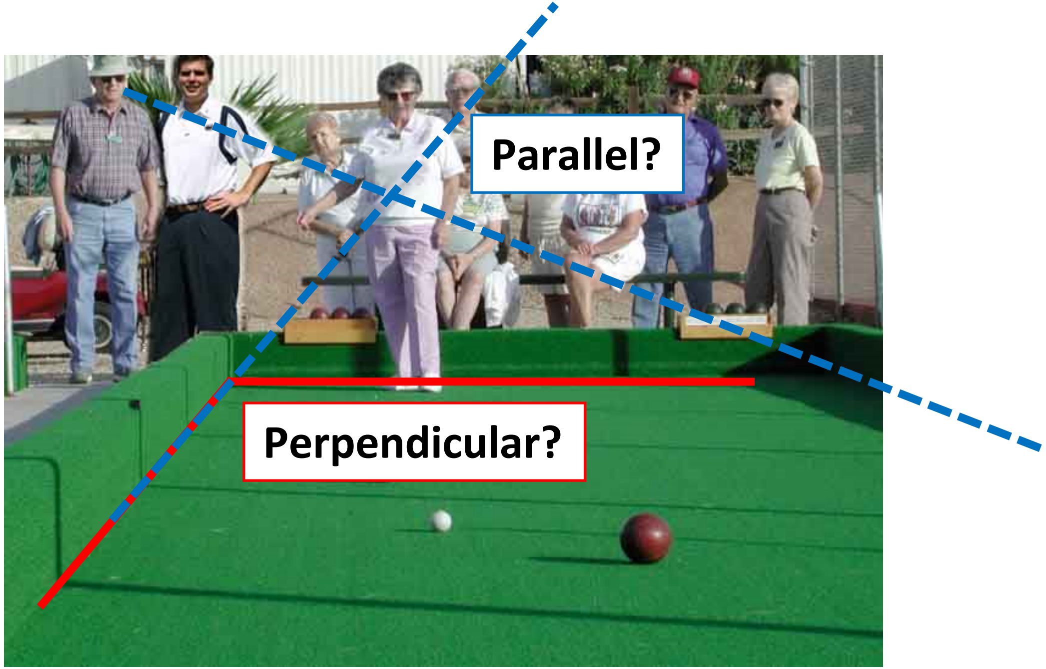


Figure by David Forsyth

Projective Geometry

What is lost?

- Length
- Angles



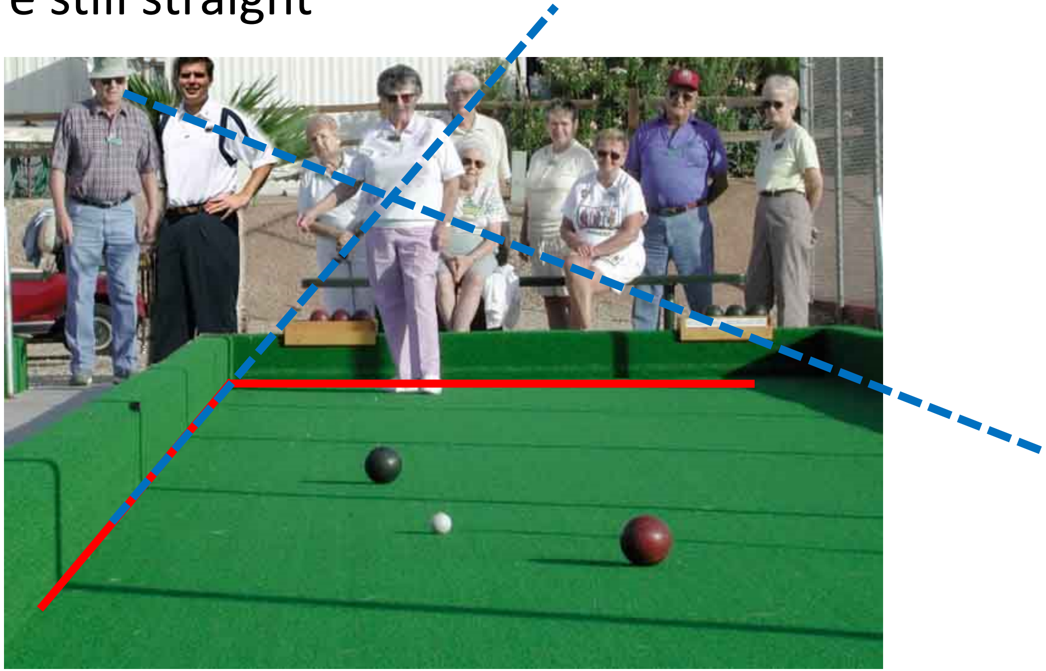
Parallel?

Perpendicular?

Projective Geometry

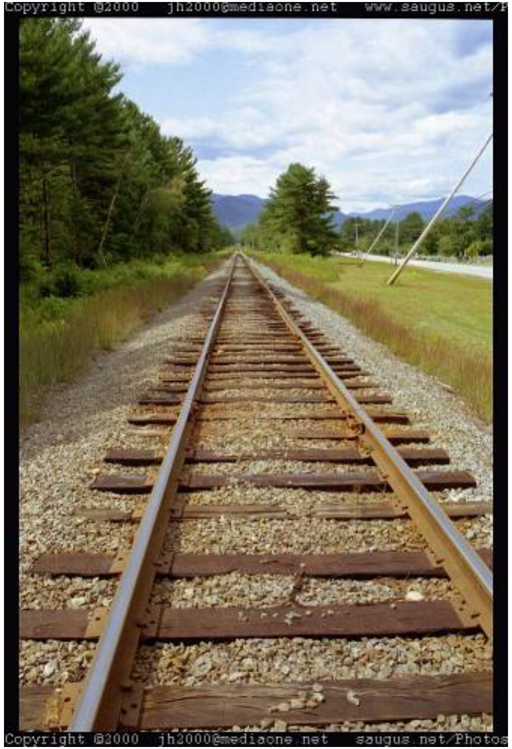
What is preserved?

- Straight lines are still straight

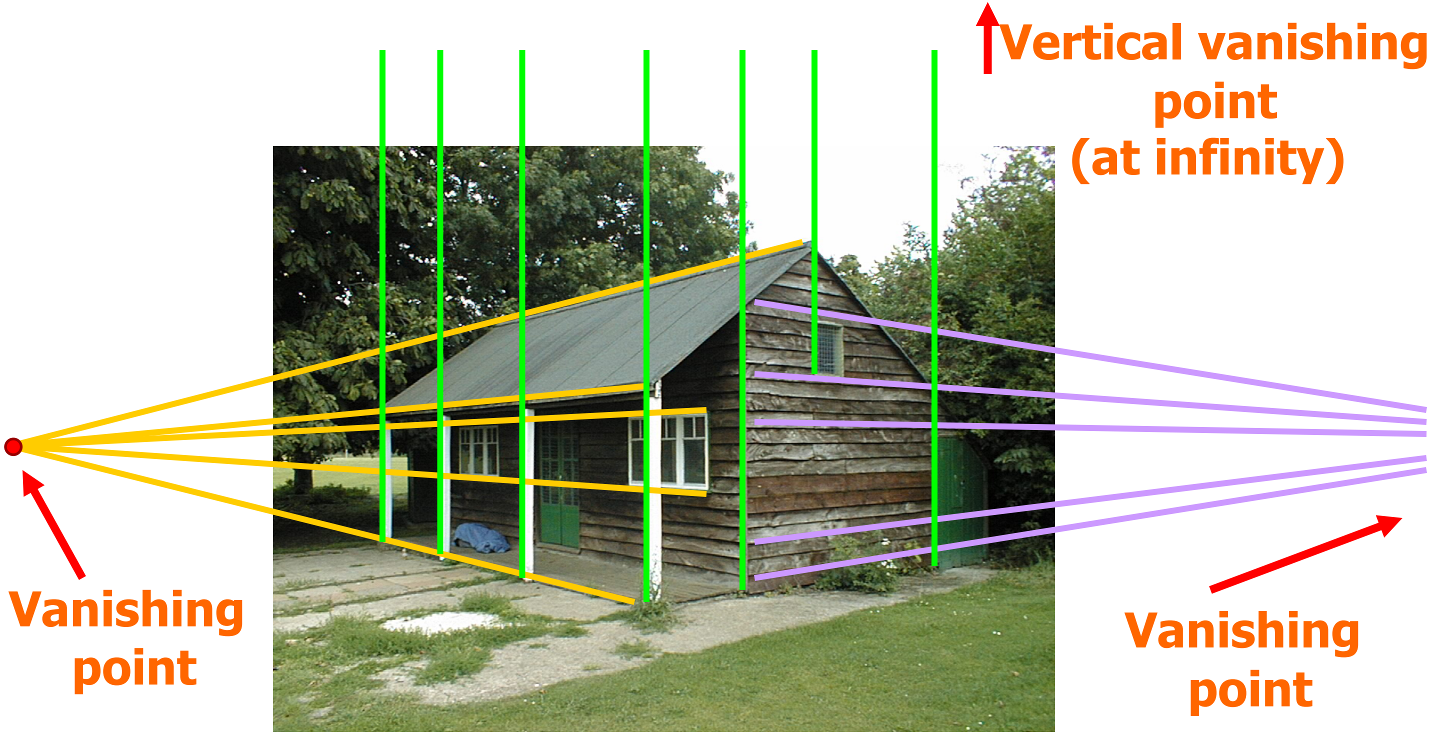


Vanishing points and lines

Parallel lines in the world intersect in the image at a “vanishing point”



Vanishing points and lines



Slide from Efros, Photo from Criminisi

Limitations of the Pinhole camera

- Not easy to produce a perfect pinhole in practice

- To photograph some objects you would really need a room sized camera (”camera”means literally room in latin)
- Depending on the type of photo-sensitive material, you would need to keep the light passing through the pinhole a long time (hours? days?).

Solution? Lenses!

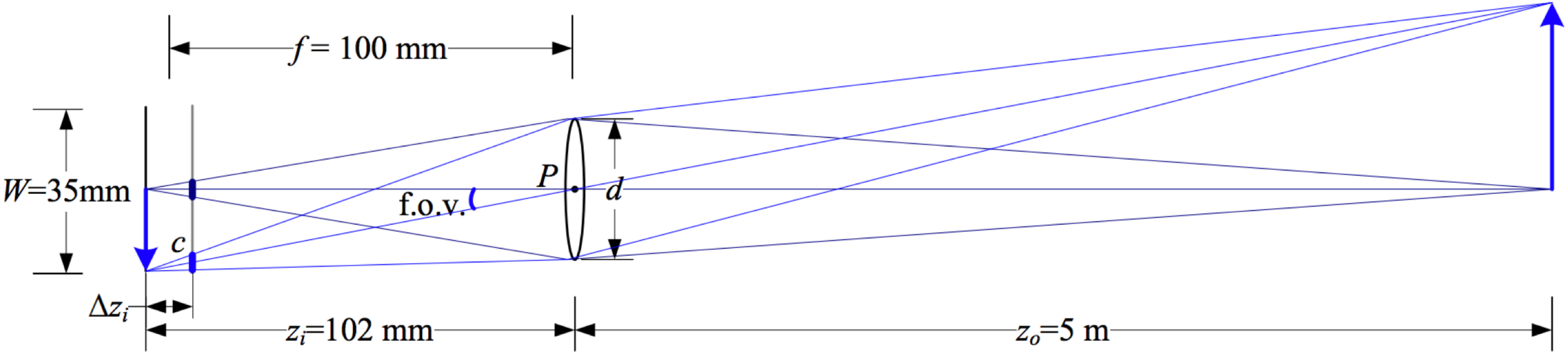


Figure from Szeliski for thin lens

We will mostly use the pinhole camera model however as an approximation of actual cameras.

Digital camera

- A digital camera replaces film with a sensor array
  - Each cell in the array is light-sensitive diode that converts photons to electrons
  - Two common types
    - Charge Coupled Device (CCD)
    - CMOS
  - http://electronics.howstuffworks.com/digital-camera.htm

Slide by Steve Seitz

Sensor Array

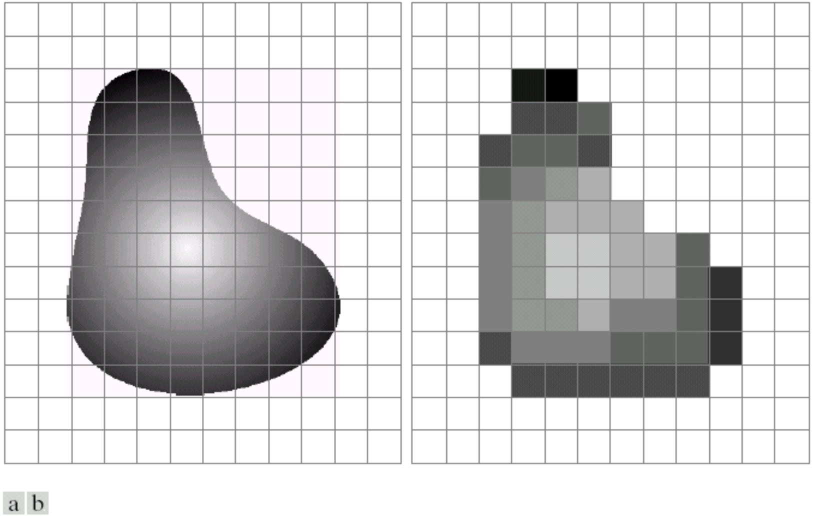
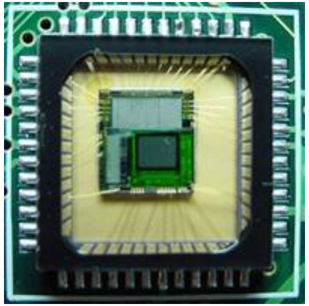
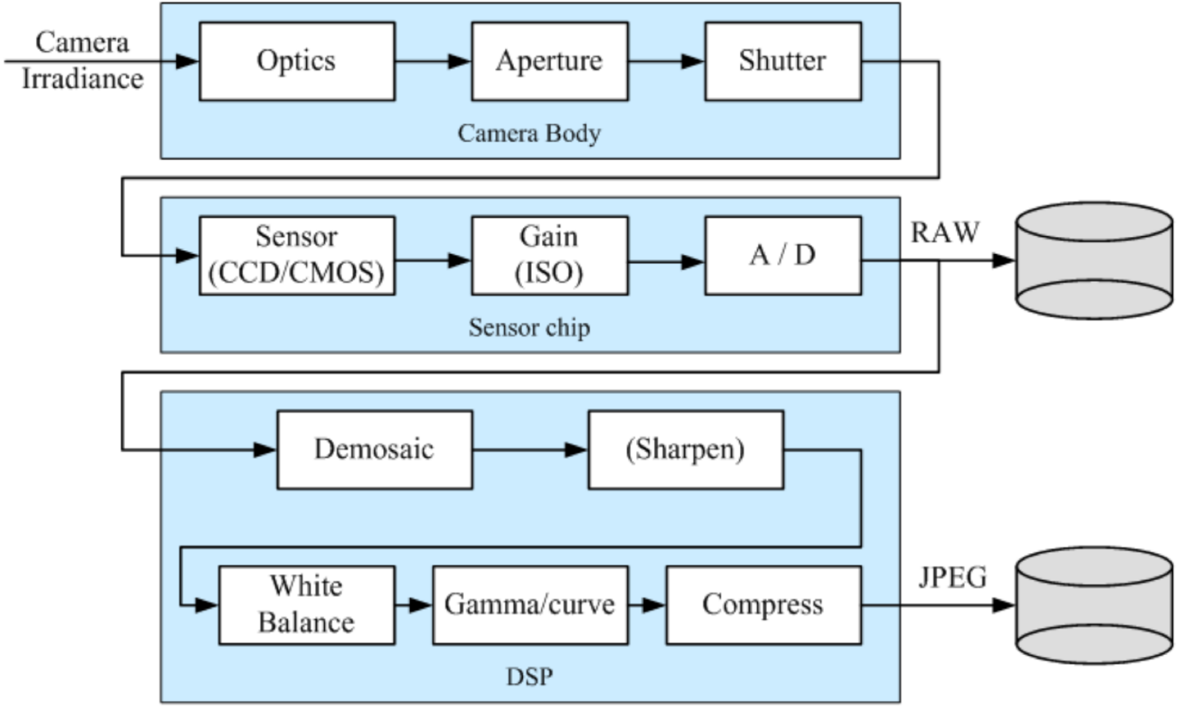


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



CMOS sensor

Digital Camera Pipeline



Slide by Steve Seitz

Cameras

Vivitar 20.1MP
$30

Polaroid
$100

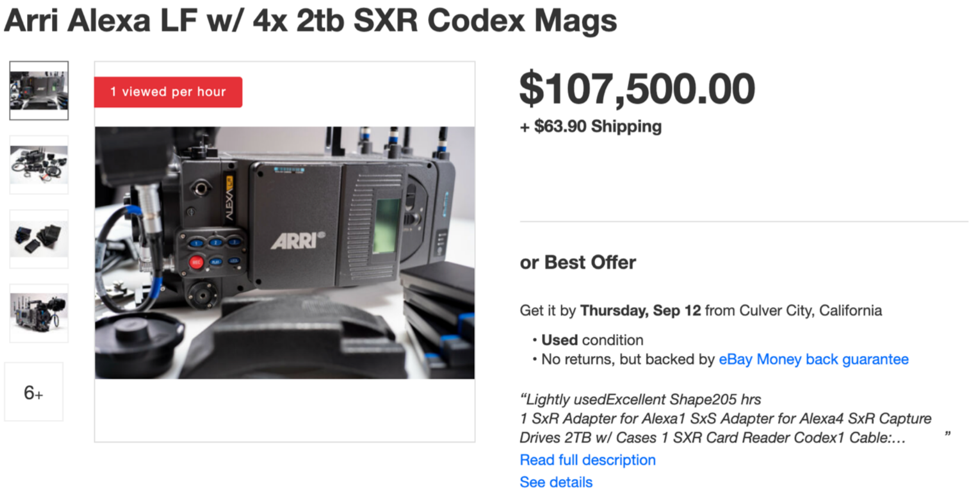
EOS Rebel T6i
$900

Canon EOS C300
$40,000

Movie Quality

Netflix Approved Cameras for Content Producers

https://backlothelp.netflix.com/hc/en-us/articles/217237077-Production-and-Post-Production-Requirements-v2-1



Arri Alexa LF w/ 4x 2tb SXR Codex Mags

$107,500.00

+ $63.90 Shipping

or Best Offer

Get it by **Thursday, Sep 12** from Culver City, California

- **Used** condition
- No returns, but backed by eBay Money back guarantee

"Lightly usedExcellent Shape205 hrs 1 SxR Adapter for Alexa1 SxS Adapter for Alexa4 SxR Capture Drives 2TB w/ Cases 1 SXR Card Reader Codex1 Cable:... "

Read full description

See details

Cameras

Apple’s Iphone X
$1000

Google’s Pixel 2 XL
$800

Cameras

Nikon D90

$1200

Nikon D3300

$700

How to Shoot Photos in Manual?

- Shutter time
- Aperture
- ISO
- Focus / Auto-focus (Yes, you can shoot in manual and also probably should focus in manual)

Small Shutter Time / Speed

http://www.photographymad.com/pages/view/shutter-speed-a-beginners-guide

Long Shutter Time

Long Shutter Time

Very Large Shutter Time – 25 seconds

https://www.davemorrowphotography.com/shutter-speed-chart

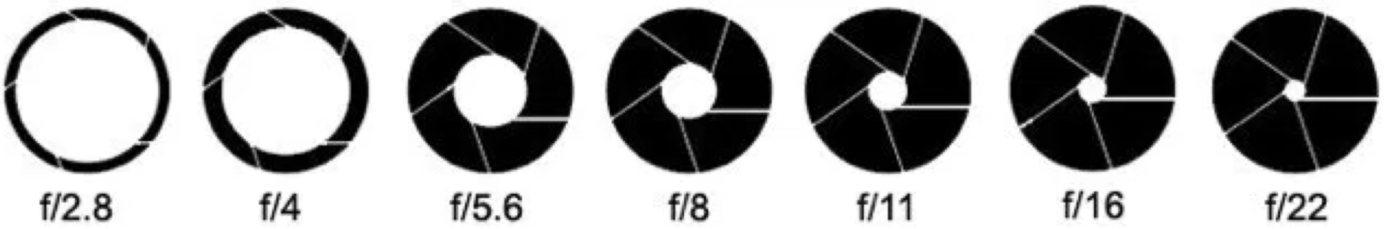
Long Shutter Time? Think of Buying a Tripod

Aluminum Tripod $140

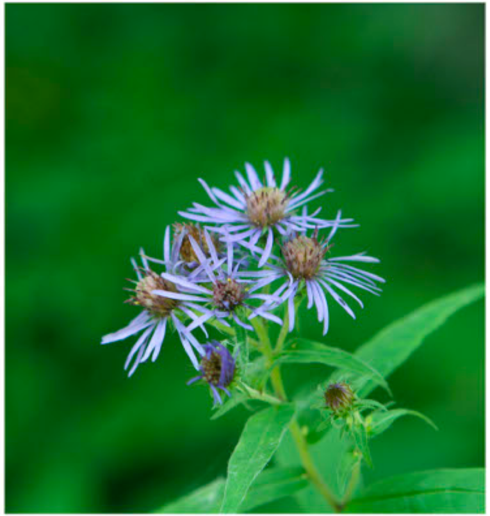
Carbon Fiber Tripod $200

Manfrotto Mountaineer
Carbon Fiber Tripod
$1300

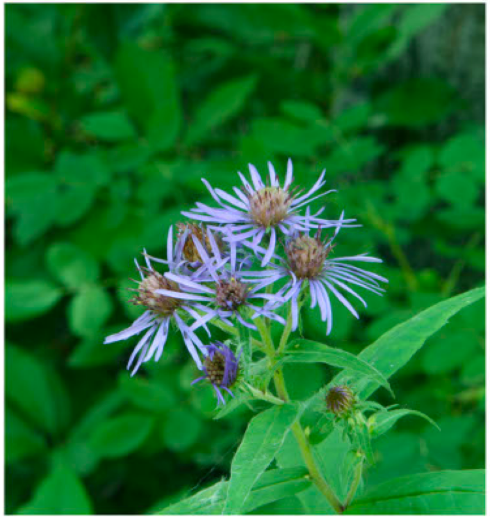
Aperture



Large vs Small Aperture + Focus Control



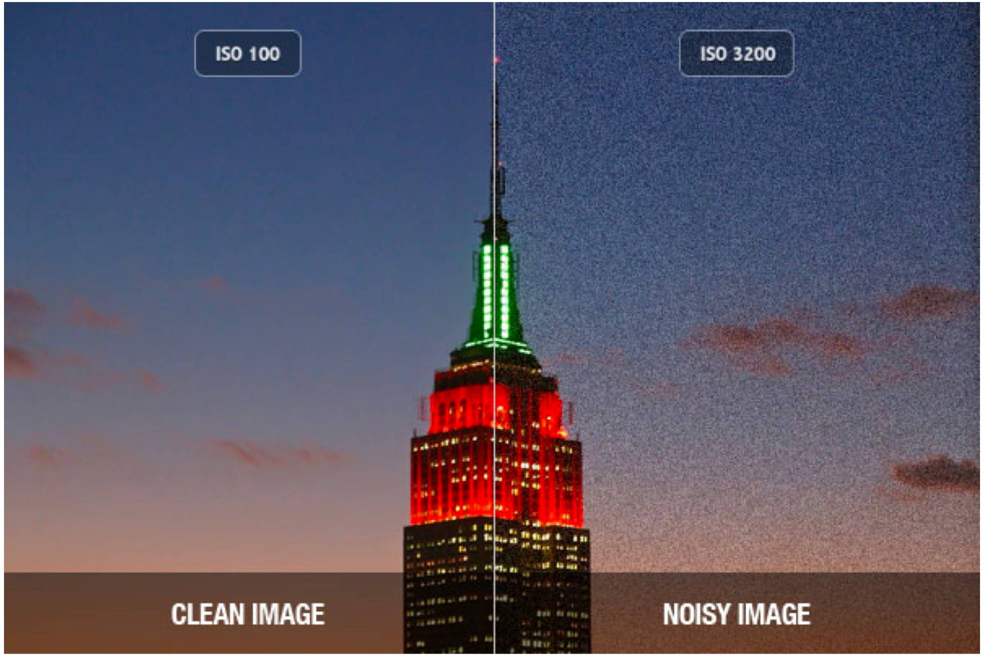
Large Aperture (F4.0)
Background nicely blurred



Small Aperture (F22)
Background is distracting

http://www.pgphotoclub.com/articles/aperture.html

ISO – Should be small ideally



https://www.exposureguide.com/iso-sensitivity/

Trade-offs (We need light to capture a photo)

- Small Aperture leads to less light
  (but allows more focus on objects)

- Small Shutter speed leads to less light
  (but allows capturing fast moving objects)

- Small ISO leads to less light
  (but produces less noisy “grainy” output)

Most difficult picture to take?

- One where we want focus on a small object in the scene
  (this requires small aperture)

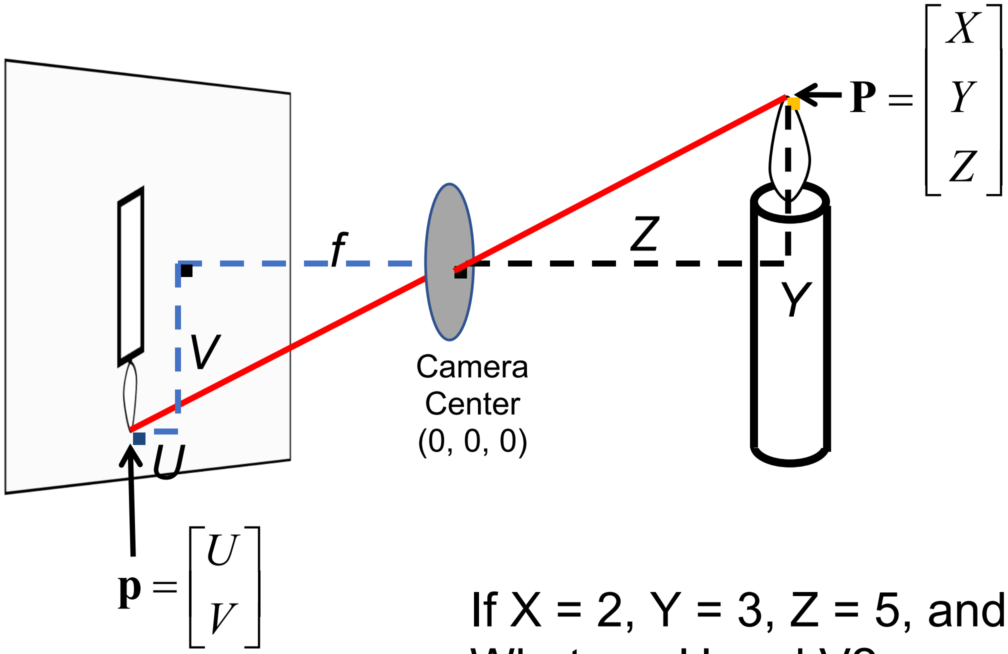
- AND the object is moving fast
  (this requires fast shutter speed)

- AND the object is in a rather poorly illuminated room
  (but if we use a high ISO the picture will be noisy)

Final Thoughts - Take with grain of salt

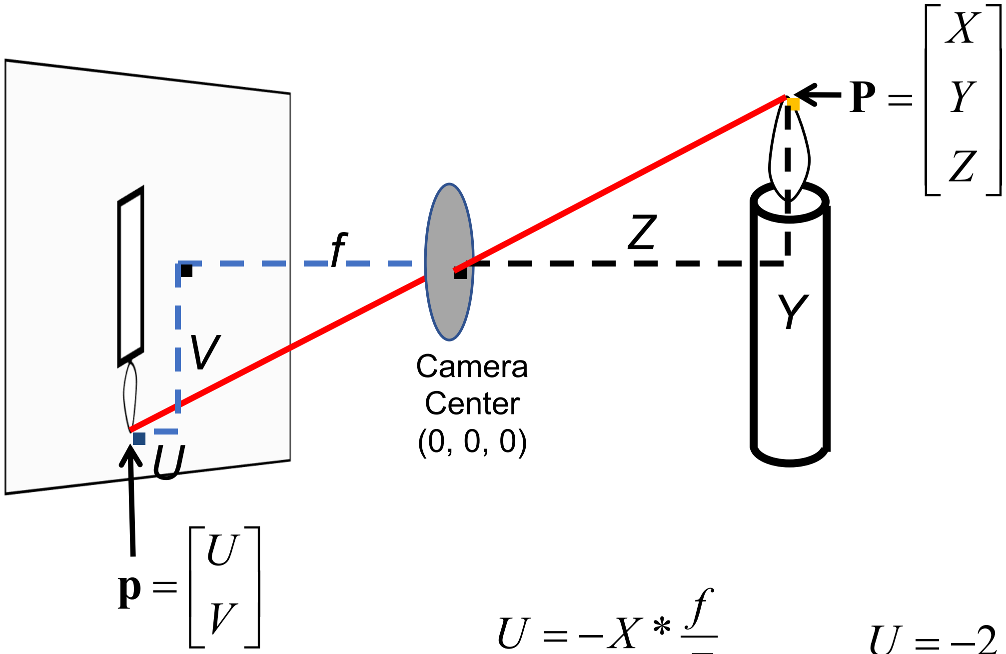
- Shooting in Automatic, especially in low light conditions will often go the easy route of just increasing the ISO all the way up
- Sometimes in low light conditions instead you want to increase the shutter time to compensate the low light, or increase the aperture. (or use Flash)
- No shame in using Automatic in a clear day, unless trying to achieve some effect.

Projection: world coordinates→image coordinates



If X = 2, Y = 3, Z = 5, and f = 2
What are U and V?

Projection: world coordinates→image coordinates

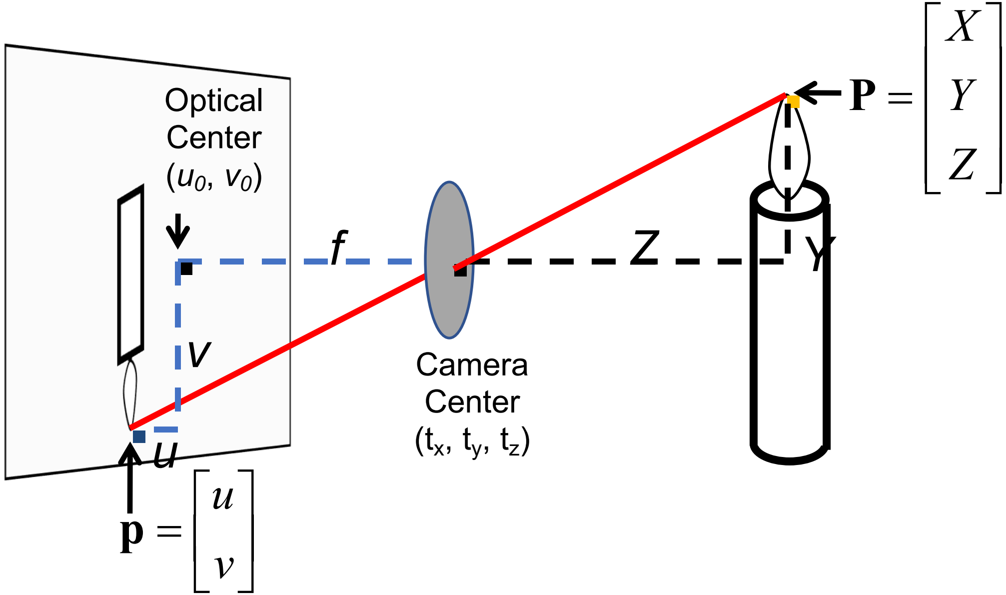


$$U = -X * \frac{f}{Z} \qquad U = -2 * \frac{2}{5}$$

$$V = -Y * \frac{f}{Z} \qquad V = -3 * \frac{2}{5}$$

Sanity check, what if f and Z are equal?

Projection: world coordinates→image coordinates



Homogeneous coordinates

Conversion

Converting to *homogeneous* coordinates

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

homogeneous image coordinates

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

homogeneous scene coordinates

Converting *from* homogeneous coordinates

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

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

Homogeneous coordinates

Invariant to scaling

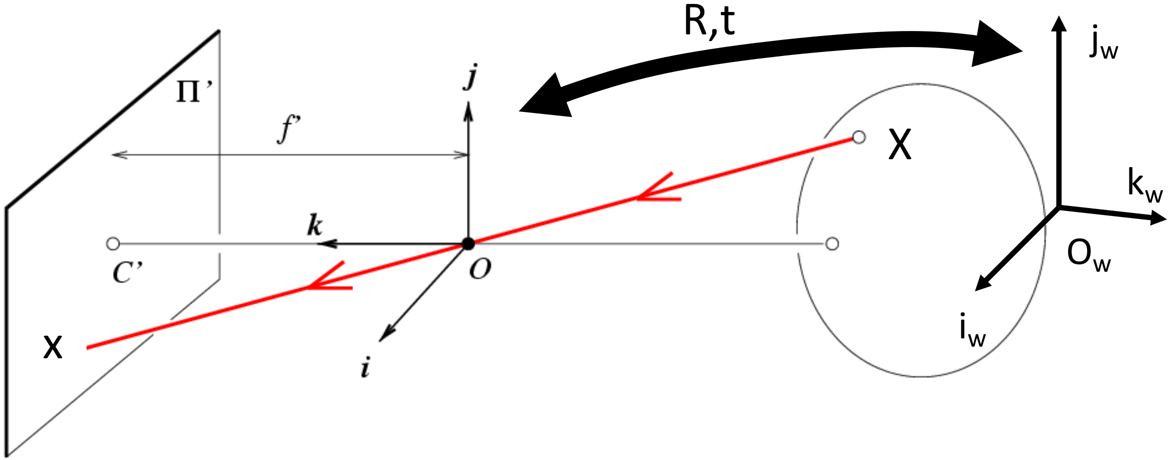
$$k\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} kx \\ ky \\ kw \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{kx}{kw} \\ \frac{ky}{kw} \end{bmatrix} = \begin{bmatrix} \frac{x}{w} \\ \frac{y}{w} \end{bmatrix}$$

Homogeneous Coordinates

Cartesian Coordinates

Point in Cartesian is ray in Homogeneous

Projection matrix (World Coordinates to Image Coordinates)



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

x: Image Coordinates: (u,v,1)

K: Intrinsic Matrix (3x3)

R: Rotation (3x3)

t: Translation (3x1)

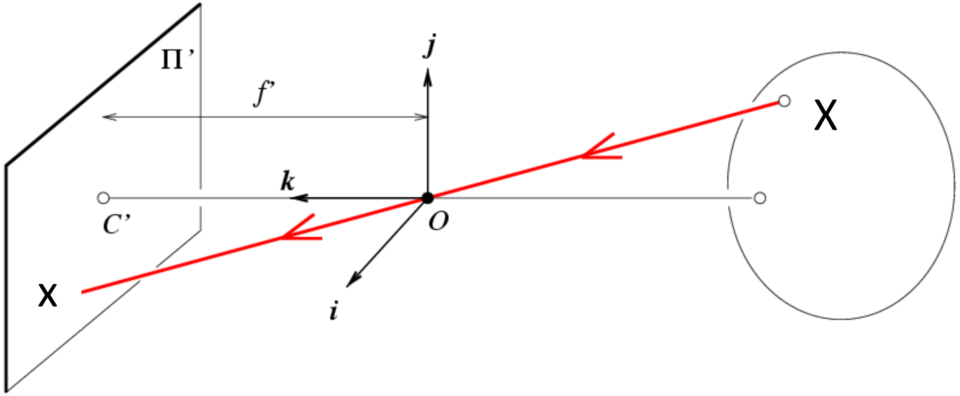
X: World Coordinates: (X,Y,Z,1)

Intrinsic Camera Properties: **K**

Extrinsic Camera Properties: **[R t]**

Slide Credit: Savarese

Projection matrix



Intrinsic Assumptions

- Unit aspect ratio
- Optical center at (0,0)
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

$$\mathbf{x}=\mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\ v\\ 1\end{bmatrix}=\left[\begin{matrix}\overset{\mathbf{K}}{\boxed{\begin{matrix}f & 0 & 0\\ 0 & f & 0\\ 0 & 0 & 1\end{matrix}}} & \begin{matrix}0\\ 0\\ 0\end{matrix}\end{matrix}\right]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

Slide Credit: Savarese

Remove assumption: known optical center

Intrinsic Assumptions

- Unit aspect ratio
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\v\\1\end{bmatrix} = \left[\boxed{\begin{matrix}f & 0 & u_0\\0 & f & v_0\\0 & 0 & 1\end{matrix}}\;\begin{matrix}0\\0\\0\end{matrix}\right]\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

Remove assumption: square pixels

Intrinsic Assumptions

- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \left[\boxed{\begin{matrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{matrix}}\;\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}\right]\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Remove assumption: non-skewed pixels

Intrinsic Assumptions

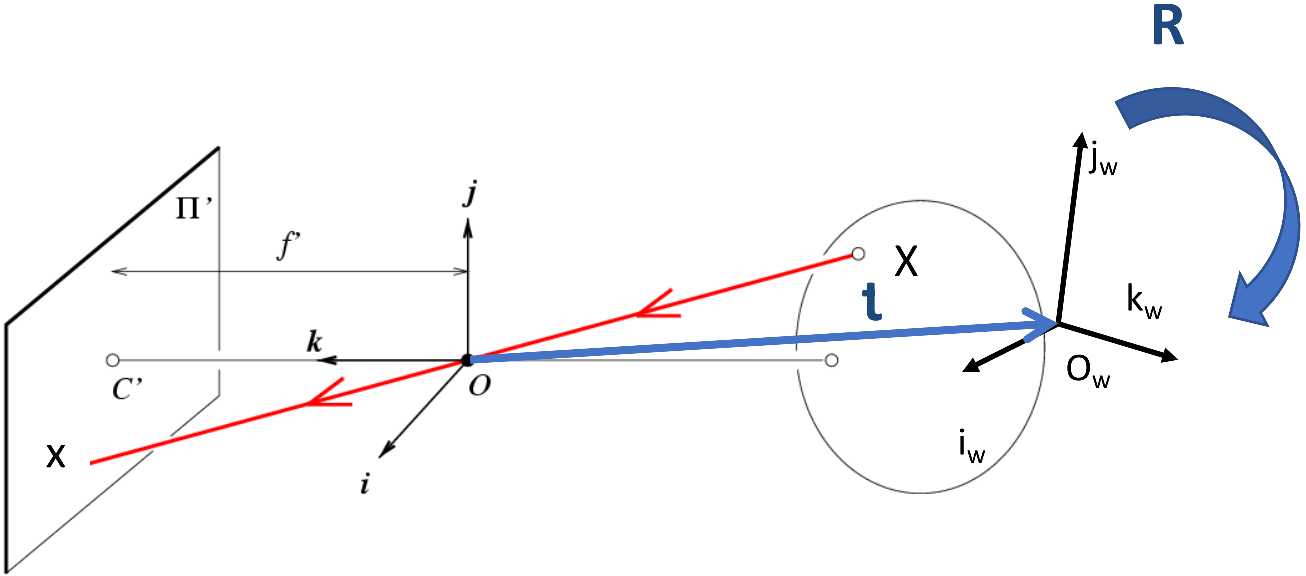
Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\ v\\ 1\end{bmatrix} = \left[\begin{array}{cc}\boxed{\begin{matrix}\alpha & s & u_0\\ 0 & \beta & v_0\\ 0 & 0 & 1\end{matrix}} & \begin{matrix}0\\ 0\\ 0\end{matrix}\end{array}\right]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

Note: different books use different notation for parameters

Oriented and Translated Camera



Allow camera translation

Intrinsic Assumptions

Extrinsic Assumptions

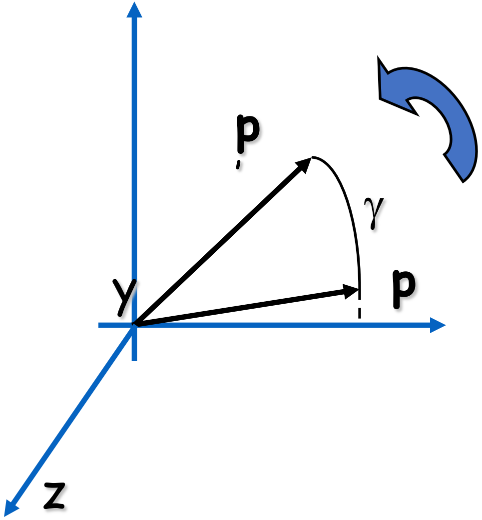
- No rotation

$$\mathbf{x} = \mathbf{K}\left[\mathbf{I} \quad \mathbf{t}\right]\mathbf{X} \Rightarrow w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Slide Credit: Saverese

3D Rotation of Points

Rotation around the coordinate axes, counter-clockwise:



$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Allow camera rotation

$$\mathbf{x} = \mathbf{K}\left[\mathbf{R} \quad \mathbf{t}\right]\mathbf{X}$$

$$\Downarrow$$

$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Degrees of freedom

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

⬇

$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \overset{5}{\begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{6}{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Things to Remember for Quiz

- Pinhole camera model
- Focal length in the pinhole camera model
- Shutter Time / Aperture / ISO
- Homogeneous Coordinates
- Extrinsic Camera Properties and Intrinsic Camera Properties
- Describe mathematically (and intuitively) the conversion process from World Coordinates to Image Coordinates

Next Class: Light

- What determines the color of a pixel?

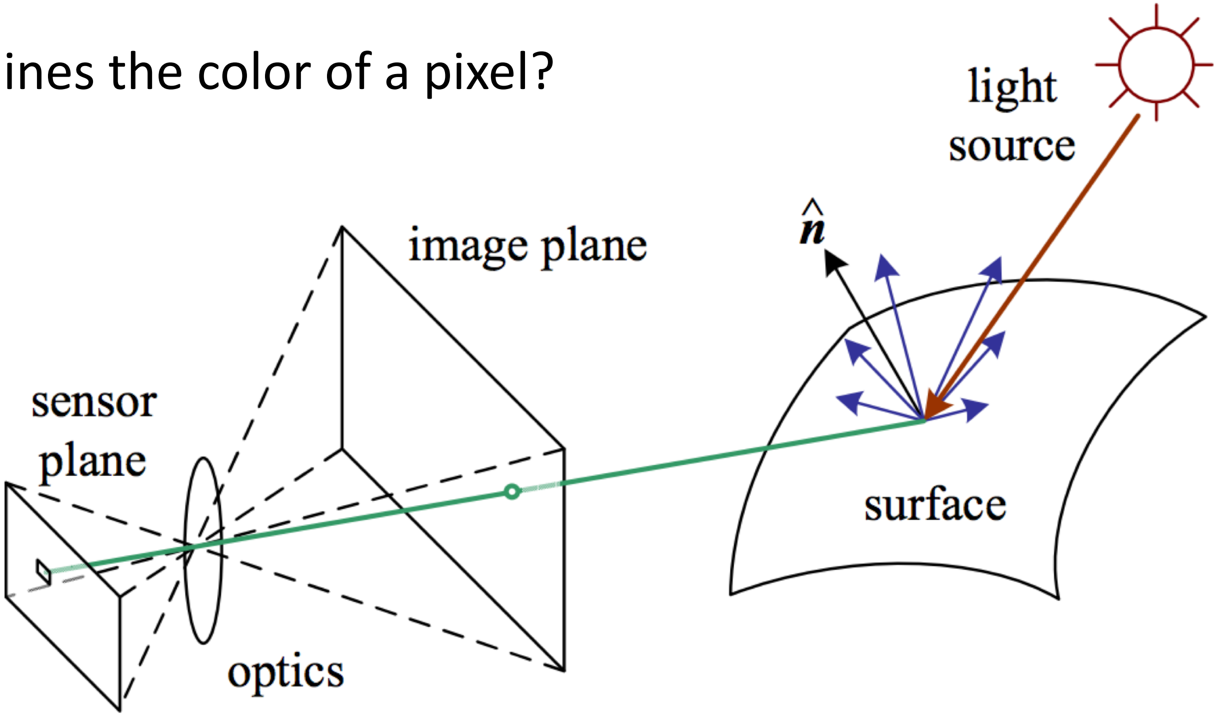
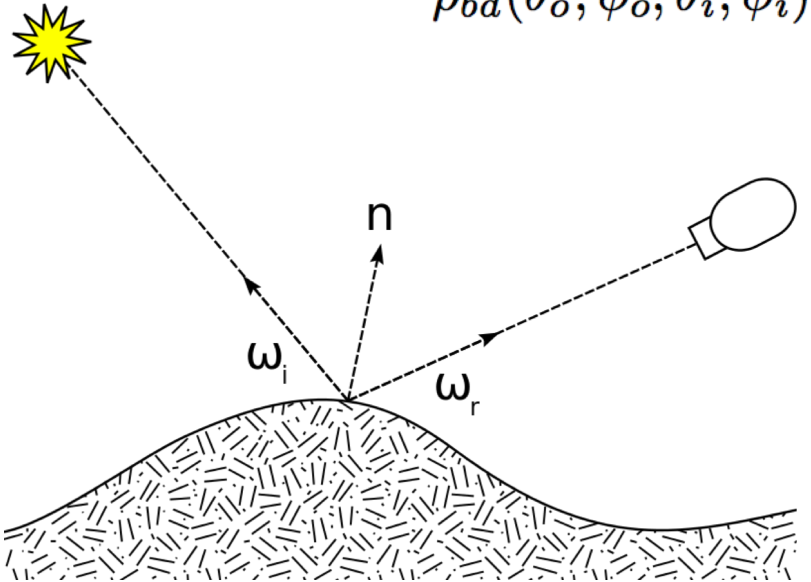


Figure from Szeliski

BRDF (Bidirectional reflectance distribution function)

$$\rho_{bd}(\theta_o, \phi_o, \theta_i, \phi_i) = \frac{L_o(\boldsymbol{x}, \theta_o, \phi_o)}{L_i(\boldsymbol{x}, \theta_i, \phi_i) \cos \theta_i d\omega}$$



Questions?