CIS 580

# Machine Perception

Or *Geometric Computer Vision*

Instructor: Lingjie Liu

Lec 1: Jan 15, 2025

Based on slides by Dinesh Jayaraman and Kostas Daniilidis

# Learning Outcomes

- Firm knowledge of fundamentals of geometric computer vision from single and multiple cameras, some image processing, and (if time permits) some deep learning for geometry
- Understanding of challenges: why algorithms work or do not work
- Ability to perform as a vision engineer in vision and robotics companies

**Prerequisites:** mainly undergraduate linear algebra (e.g. vector & matrix products, inverses, determinants, solving systems of linear equations, eigenvectors), and some high school Euclidean geometry

# CIS 580 Machine Perception Spring 2025

- Resources:
  - Canvas
    - Slides & readings
    - (-> Recordings, Ed, Gradescope)

# Instructor Introduction

Lingjie Liu

Assistant Professor, CIS

Penn CG lab and GRASP lab

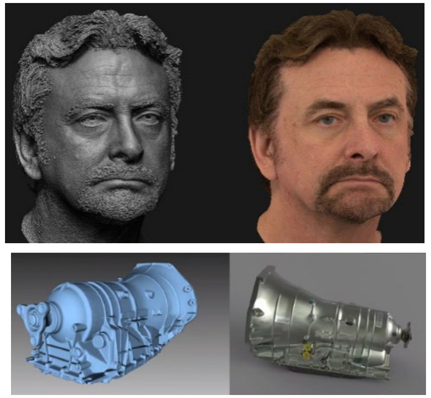
https://lingjie0206.github.io/

# My Research Interests

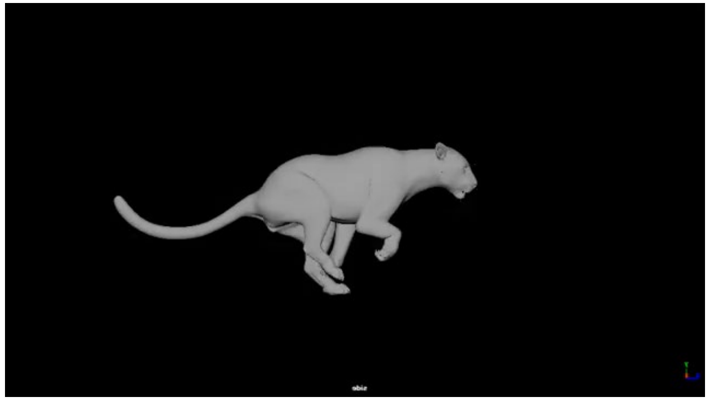
- **Reconstruction** of Real-world Dynamic Scenes.

# My Research Interests

- **Reconstruction** of Real-world Dynamic Scenes.



Geometry
+ Appearance



Motion
+ Deformation

# My Research Interests

- **Reconstruction** of Real-world Dynamic Scenes.

# My Research Interests

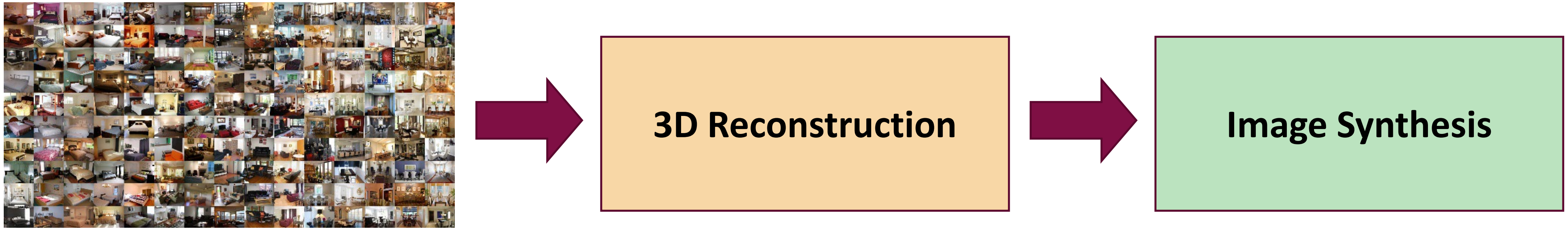
- **Image Synthesis** of Real-world Scenes with 3D Control.

# My Research Interests

- **Image Synthesis** of Real-world Scenes with 3D Control.

# My Research Interests

- **Image Synthesis** of Real-world Scenes with 3D Control.

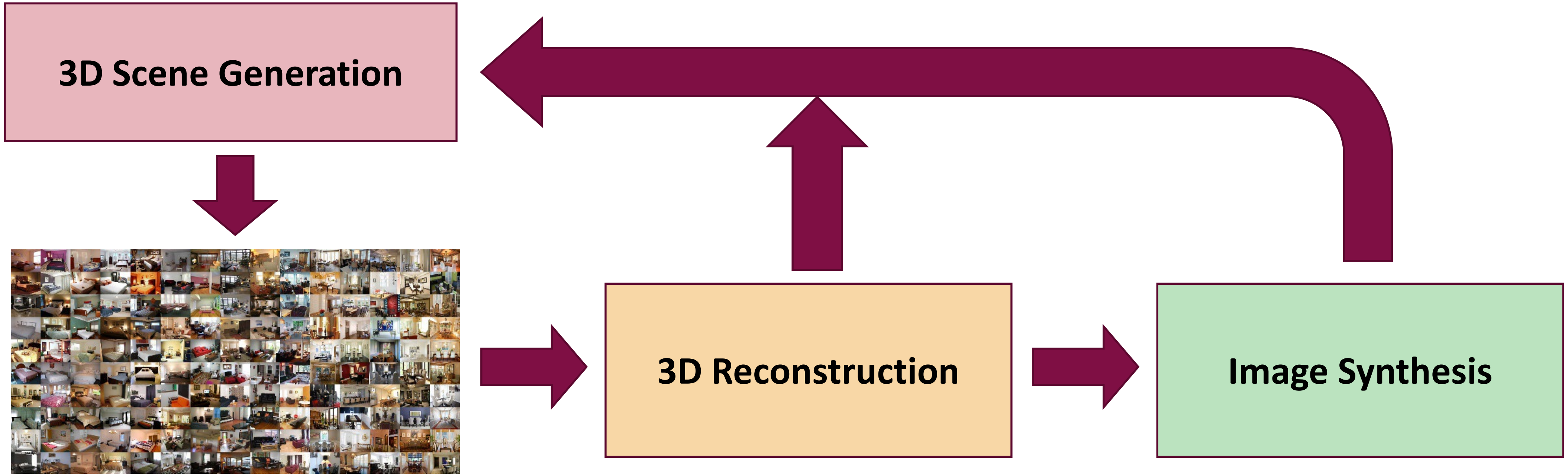


# My Research Interests

- Large-scale 3D **Scene Generation**

# My Research Interests

- Large-scale 3D Scene **Generation**



# Grading Scheme

- 5x Homeworks: 60%
- Midterm: 20%
- Final Exam: 20%
- Possible bonus for extensive in-class participation (up to 5%)
  - I need your interaction and questions. Do not let the communication become one-way! Interrupt at any time.

- Lenient letter grades. The only way to fail this class is to violate the honor code (cheat, plagiarize)
  - Zero Tolerance. Think before you act: One homework or one question in a midterm is unlikely to affect your final letter grade.

# Waitlisty

- A delay of 2-3 days for the system to reflect the approval on your end after I process your request. Please be patient.

# Course Team Spring' 25

Qiao Feng
fengqiao@seas.upenn.edu

Aishwarya Balaji
abalaji3@seas.upenn.edu

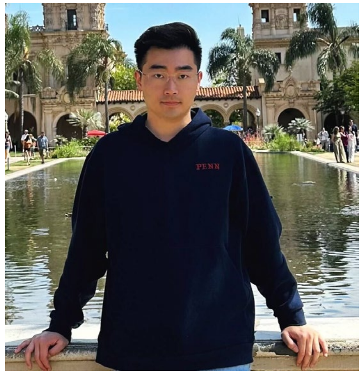
Bryan Alfaro
balfaro@seas.upenn.edu

Pengyu Chen
chpengyu@seas.upenn.edu

Yiming Huang
ymhuang9@seas.upenn.edu

Prakriti Prasad
prasadpr@seas.upenn.edu

Zi-Yan
lzi@seas.upenn.edu



Xiangyu Han
hanxy@seas.upenn.edu

Xuyi Meng
mengxuyi@seas.upenn.edu

Yicong Wang
yicongw@seas.upenn.edu

Quan A. Pham
quanpham@seas.upenn.edu

Chuhao Chen
chuhaoc@seas.upenn.edu

Paisley Hou
jingxhou@seas.upenn.edu

# How to Speak With The Course Team

Go to office hours!

- Each of the TAs will have 1 hour of OH each week.

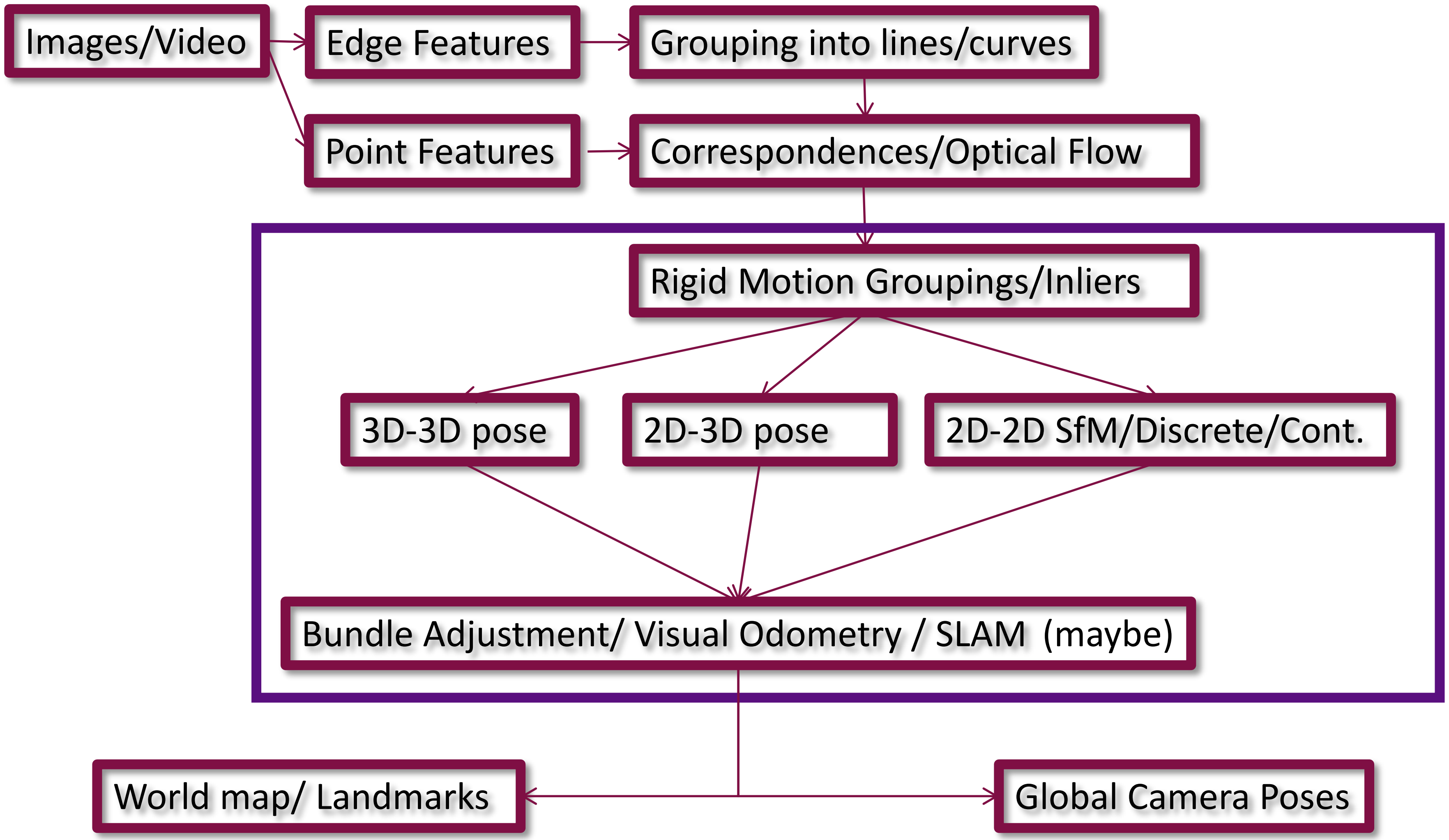
- I will also have 1 hour of OH each week.

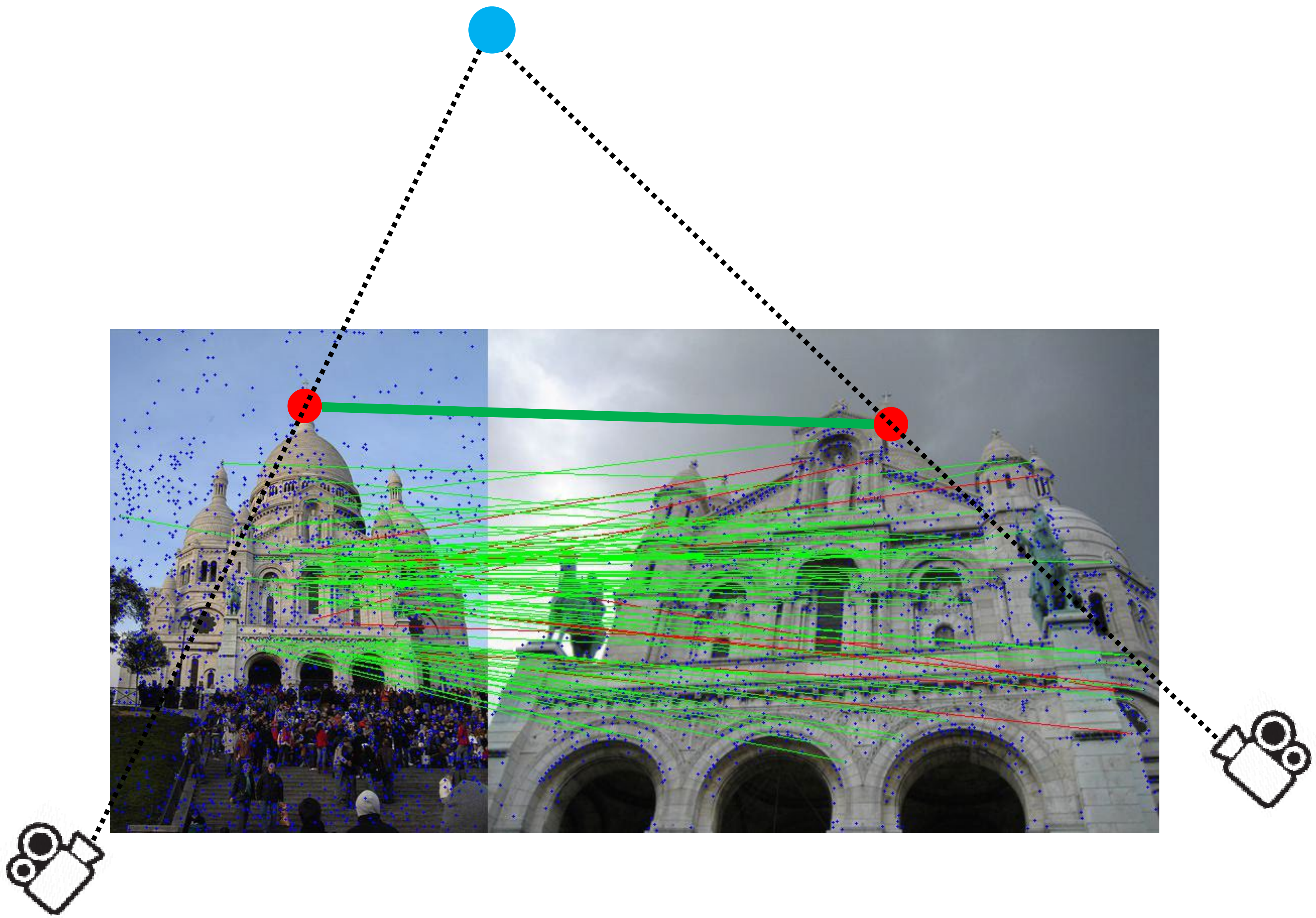
- Starting from the week of Jan 27

(PS: No class next Monday)

# Lec 1: Perspective Projection

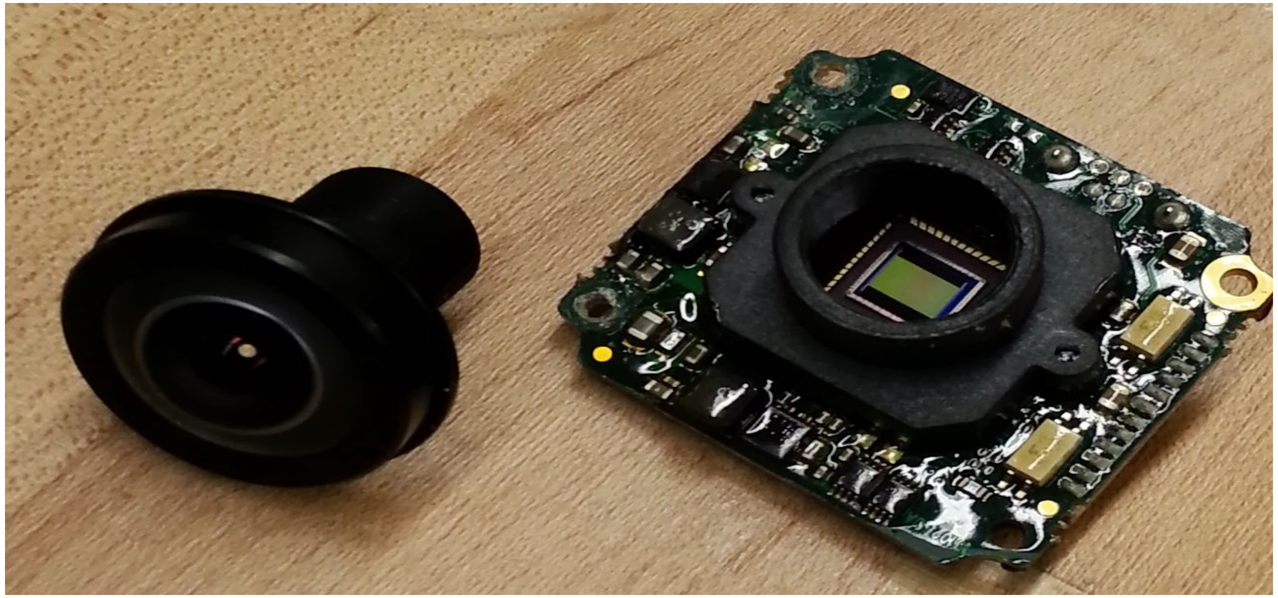
# Geometric Perception Pipeline





# What is a camera?


=

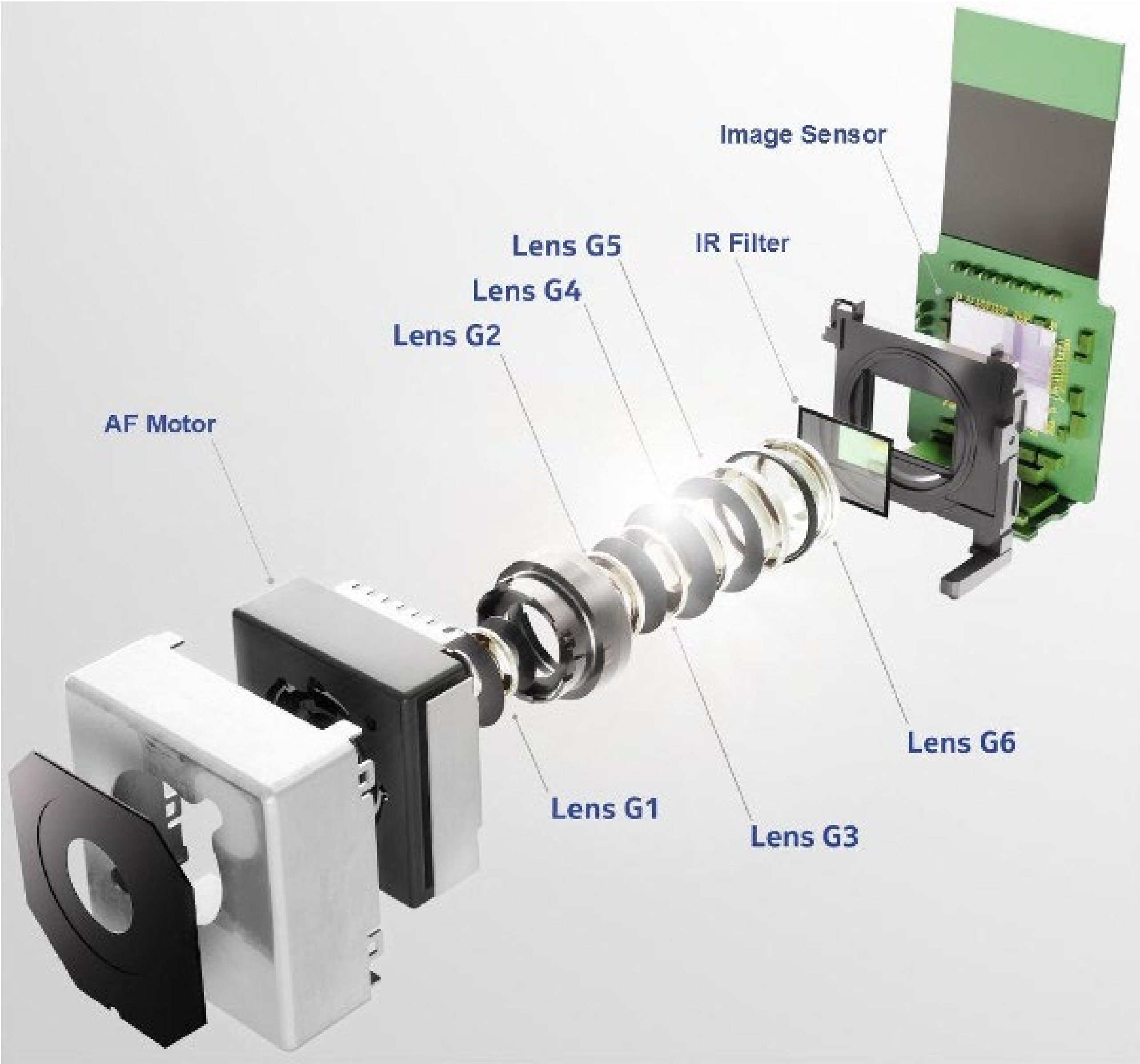


Most basically, an imaging sensor, and a lens

# But real cameras hide many complex details!

- Compound lenses containing many plastic / glass components
- Moving lenses for autofocus
- Lens distortions etc.

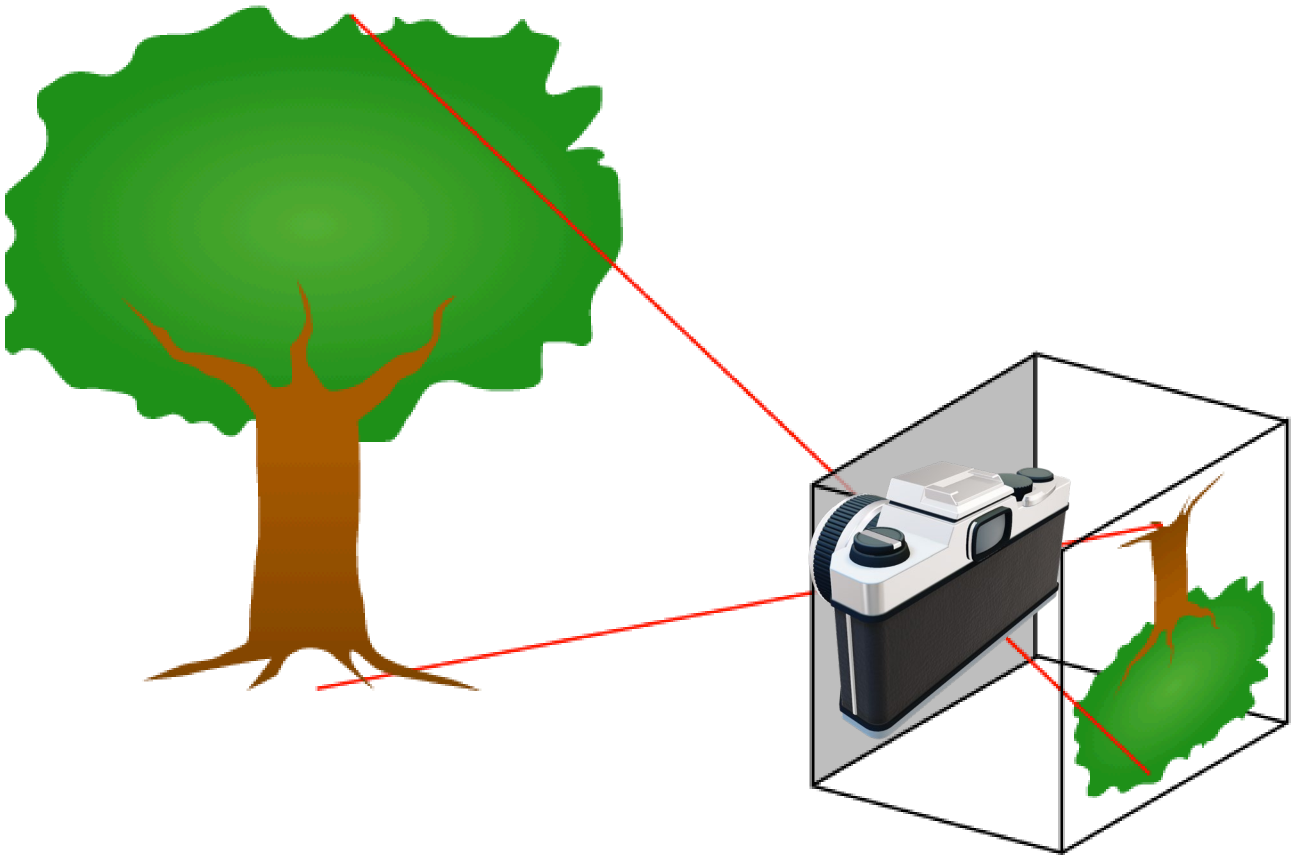
For the most part, we can ignore these complex optics when studying geometric vision, and then deal with them through minor corrections afterwards.




https://insightsolutionsglobal.com/camera-module-definition-and-types/

# Introduction to Perspective Projection

# Perspective Projection => Pinhole Camera Model



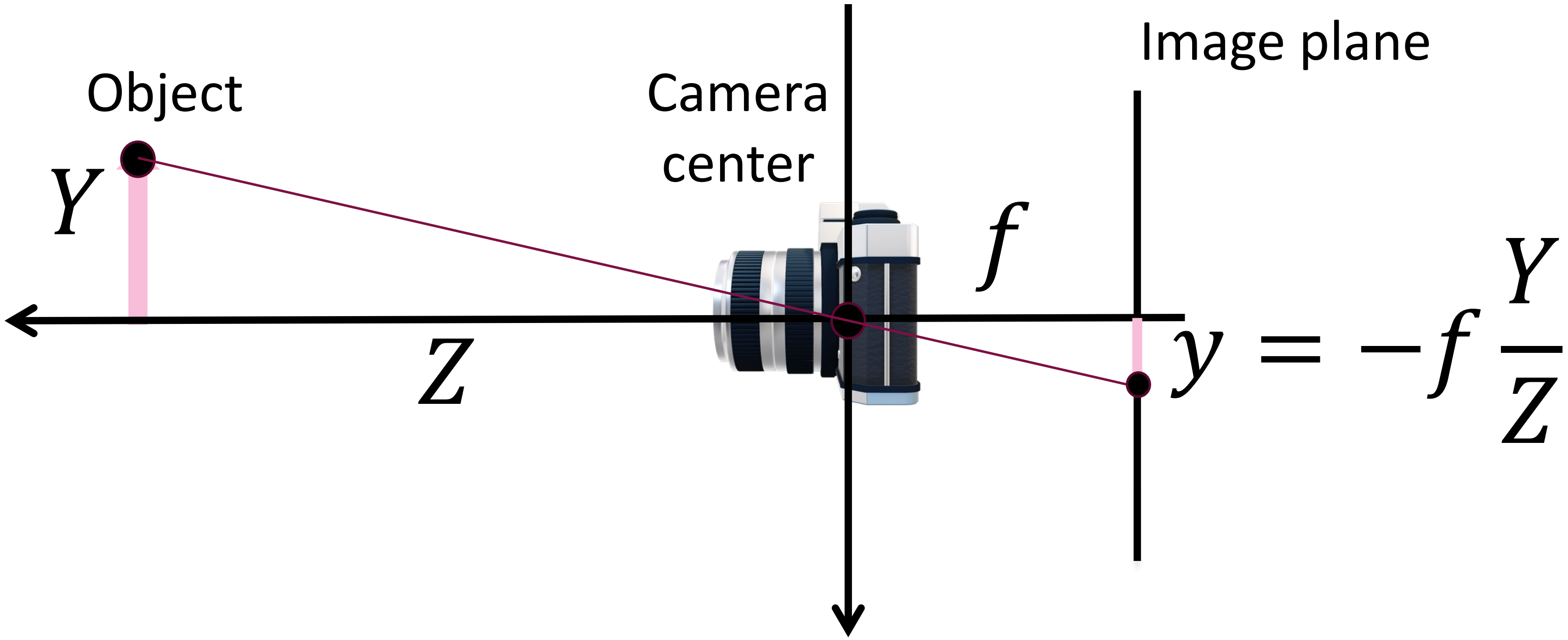
This is a good model for a camera with a lens,
as long as the "aperture" is small.

Wikipedia

"Alhazen" Ibn al-Haytham, 12th century Iraq

# Pinhole camera model (visualized in 2D)

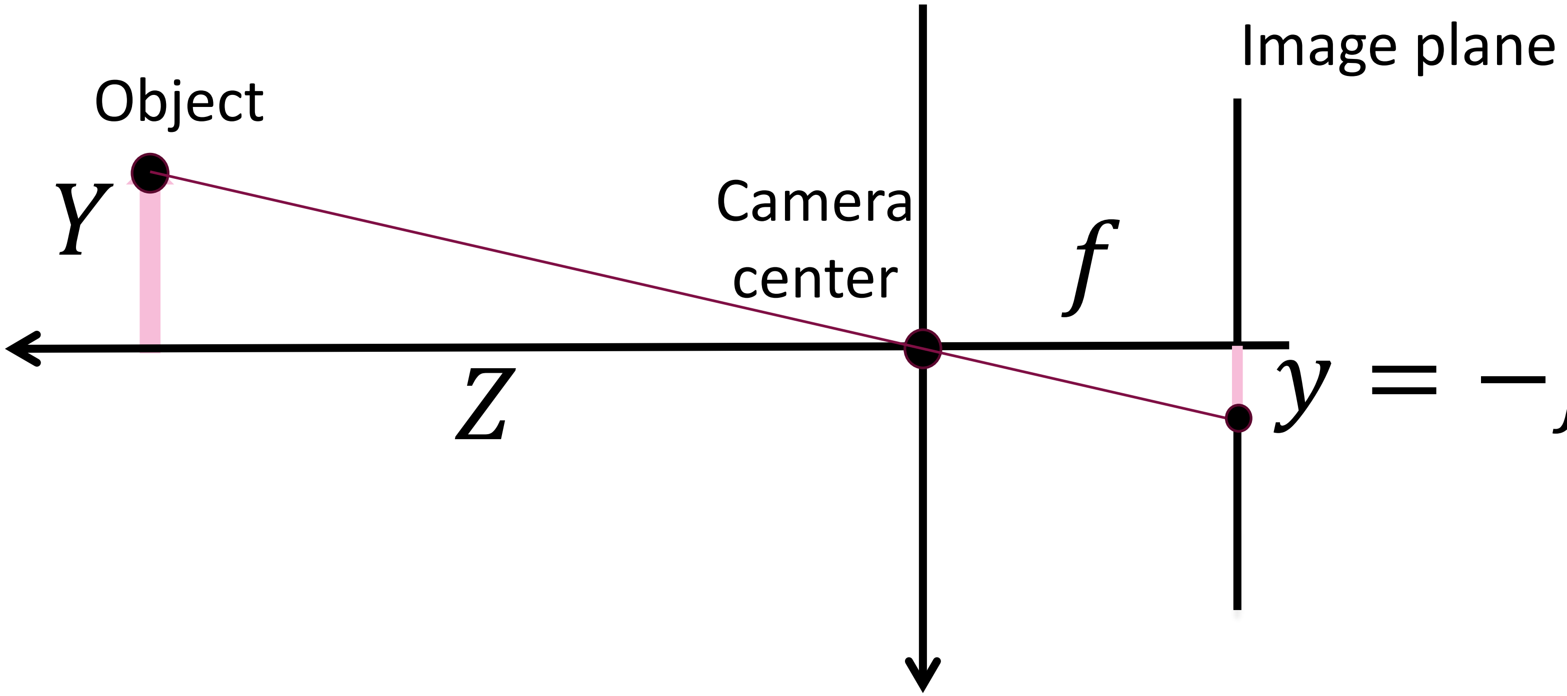
f = distance of camera center from image plane




Note how we are expressing all positions like $Y, Z, y$ in coordinates tied to the camera. We will revisit this choice soon.

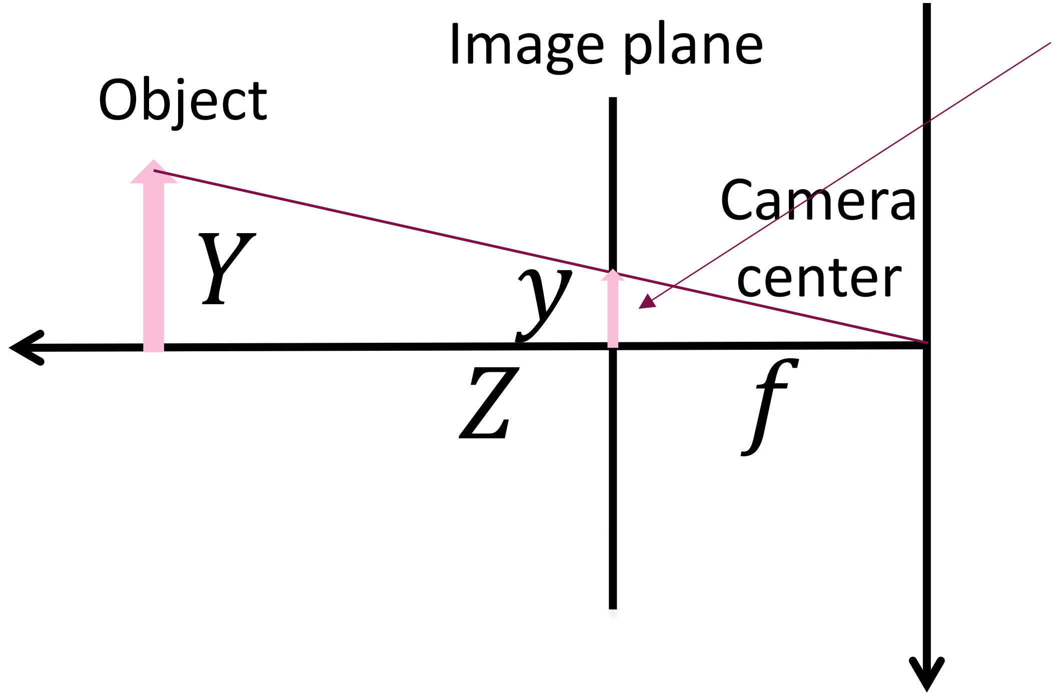
# Pinhole camera model (visualized in 2D)

f = distance of camera center from image plane




$$y = -f\frac{Y}{Z}$$

# Pinhole camera model with the image plane in front

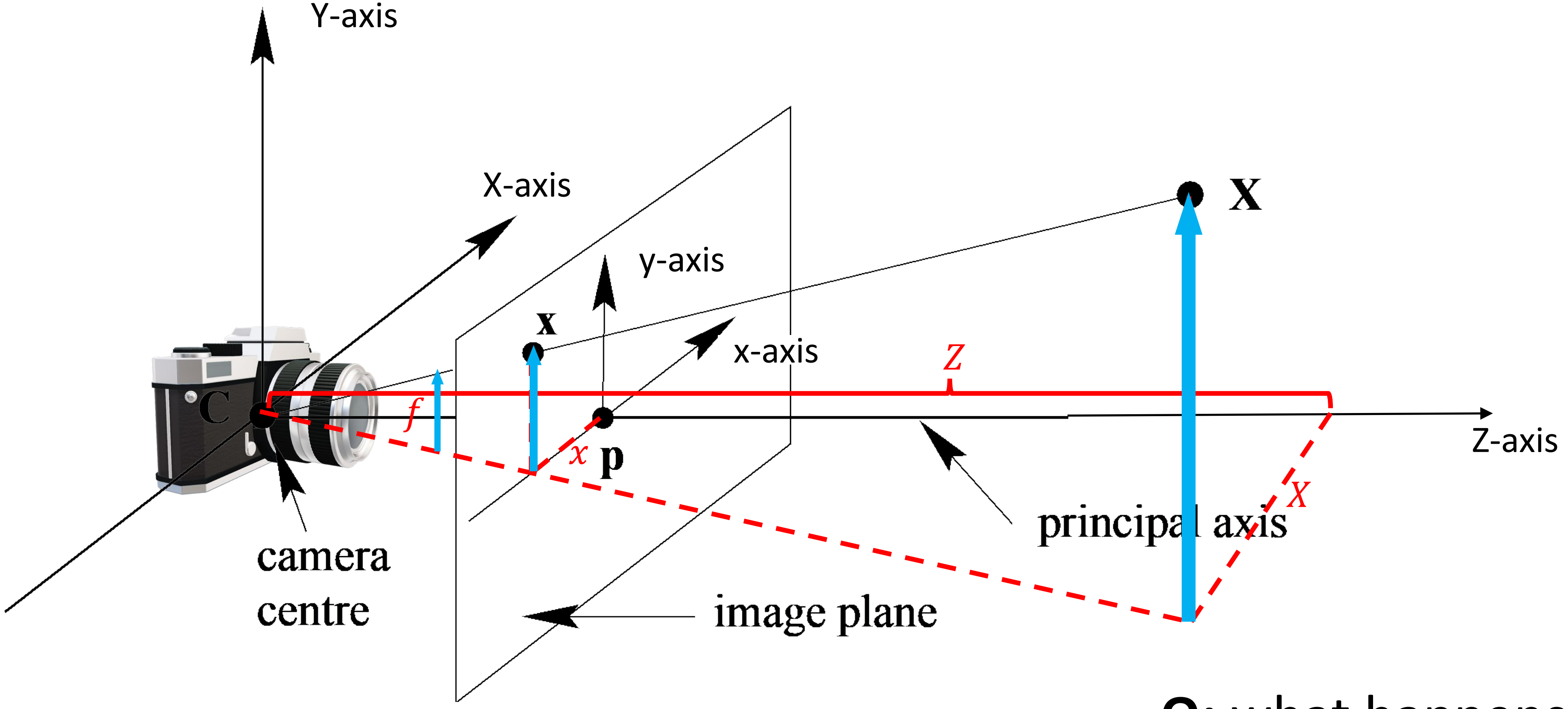



$$y = f\frac{Y}{Z}$$

For convenience move the image plane in front of camera so that objects appear upright.

Can get rid of the pesky negative sign.

# Basic Perspective Projection Equations, now in 3D

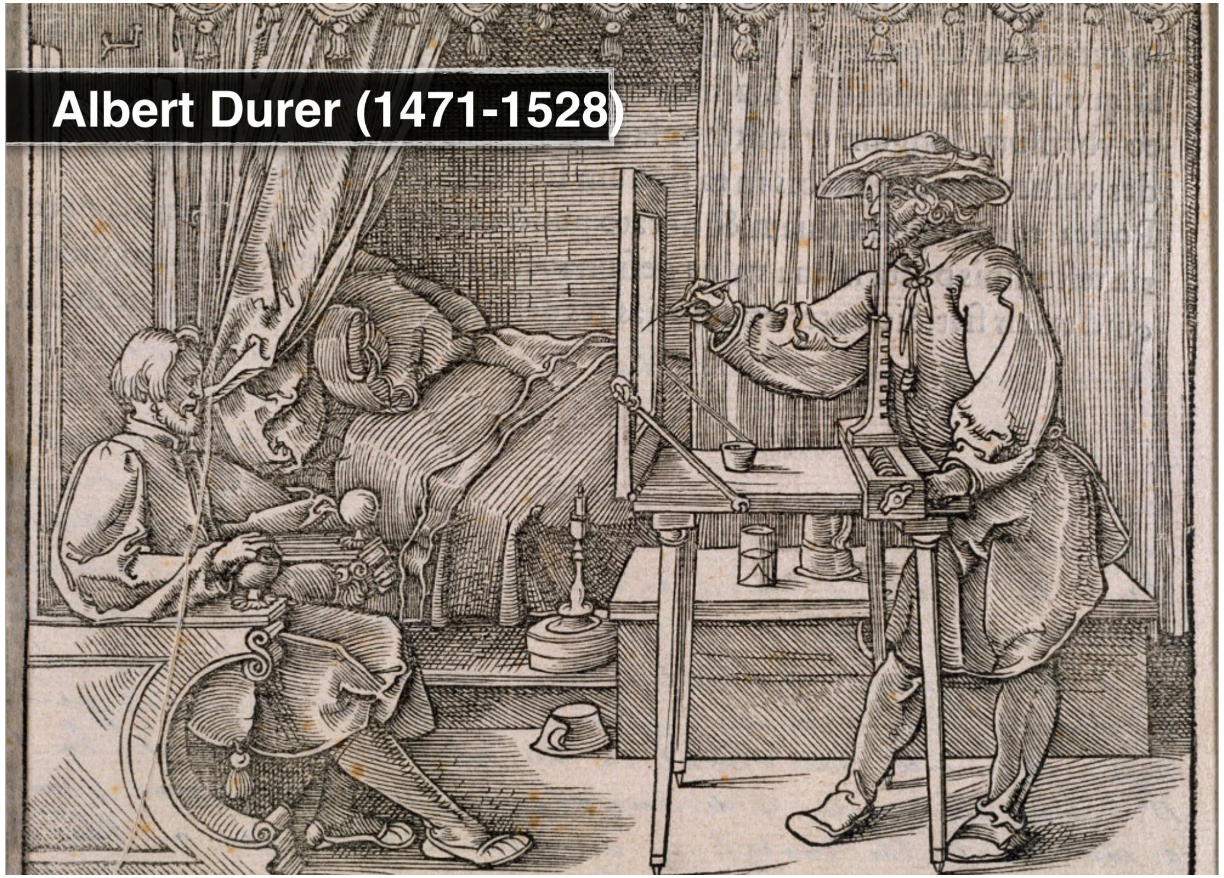



$$x = f\frac{X}{Z}, y = f\frac{Y}{Z}$$

**Q:** what happens to sizes of objects as you move the image plane close or far?

Z&H Ch6

# Perspective Projection in Renaissance Art



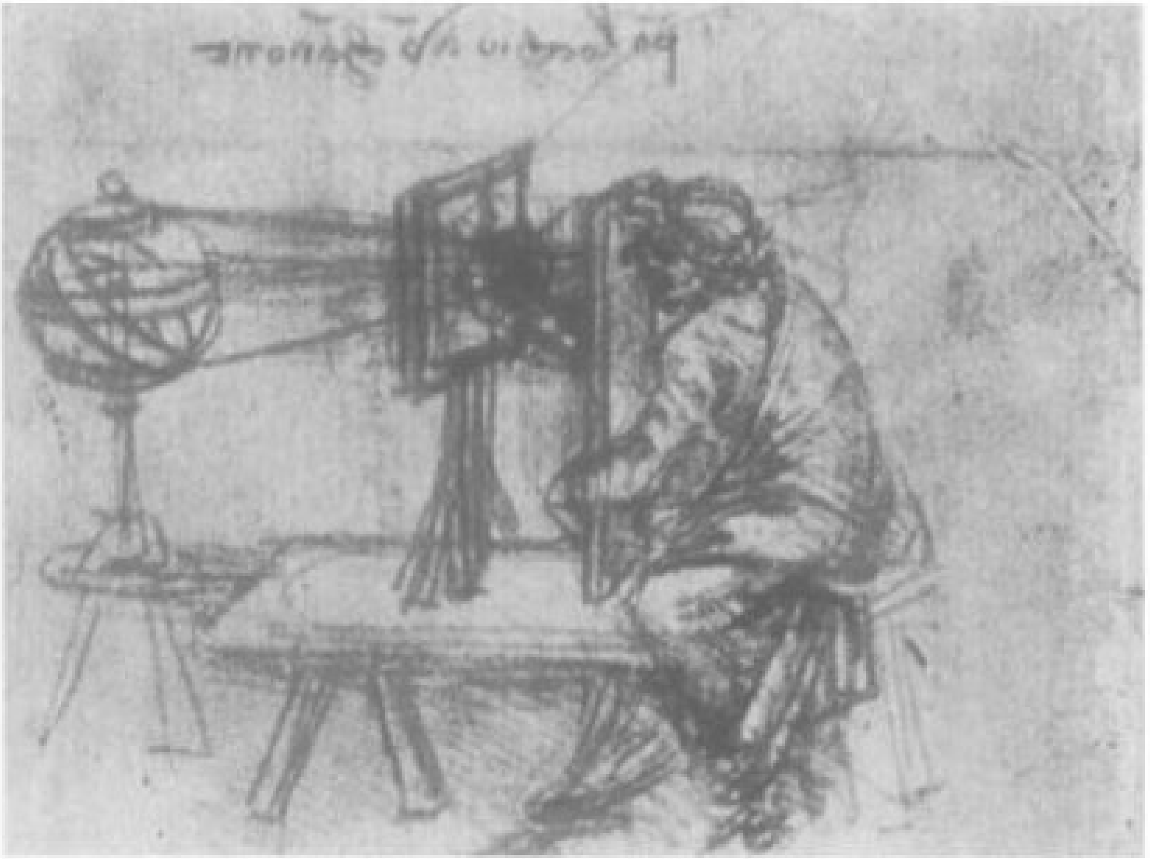



*Figure 2.1.* Leonardo's technique for making a perspectival drawing of the sphere of the macrocosm (CA 1 ra bis).


LMSC

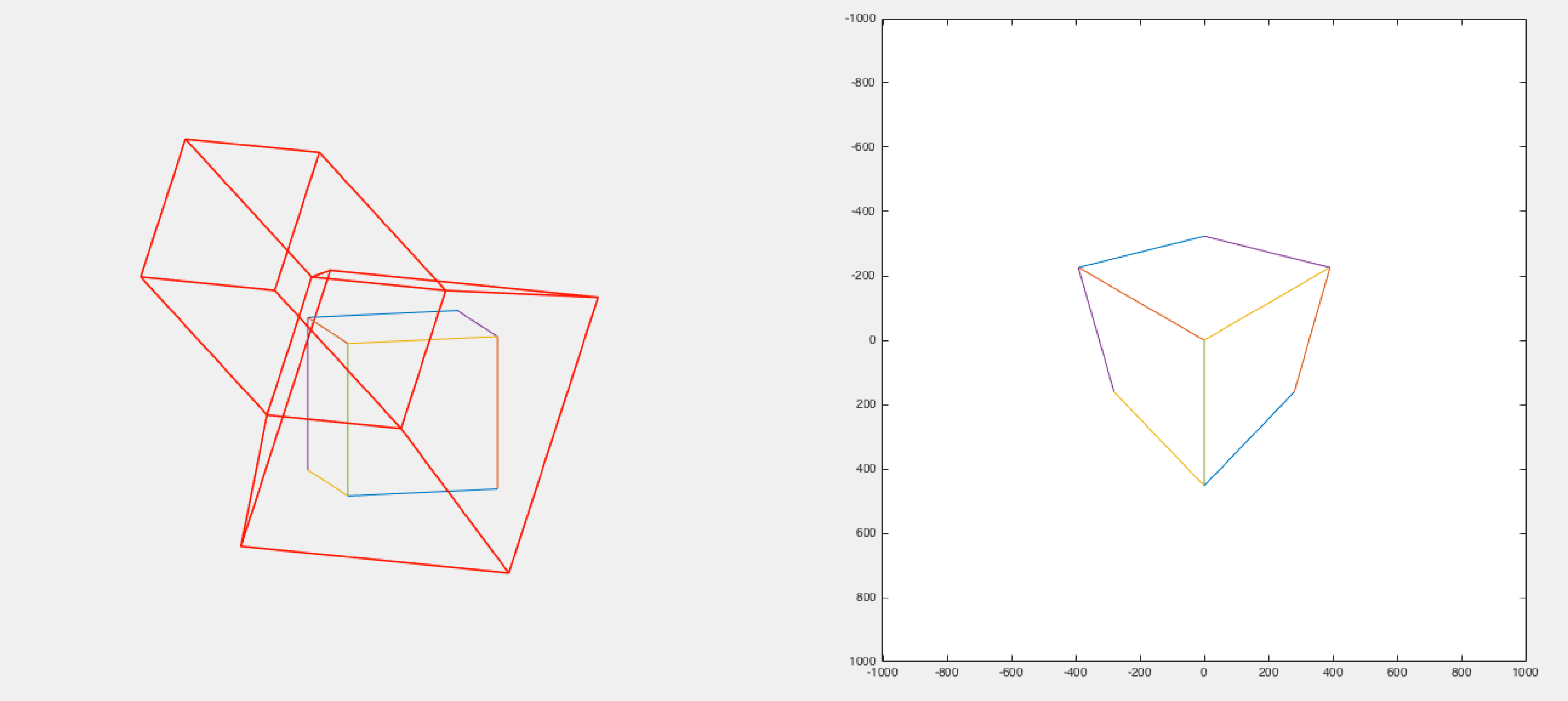
# What properties are preserved by perspective projections?

- Shape?
  - e.g. square, circle, etc.?
- Parallelness / Angles?
- Lengths?
- Ratio of lengths?

**None of the above!**

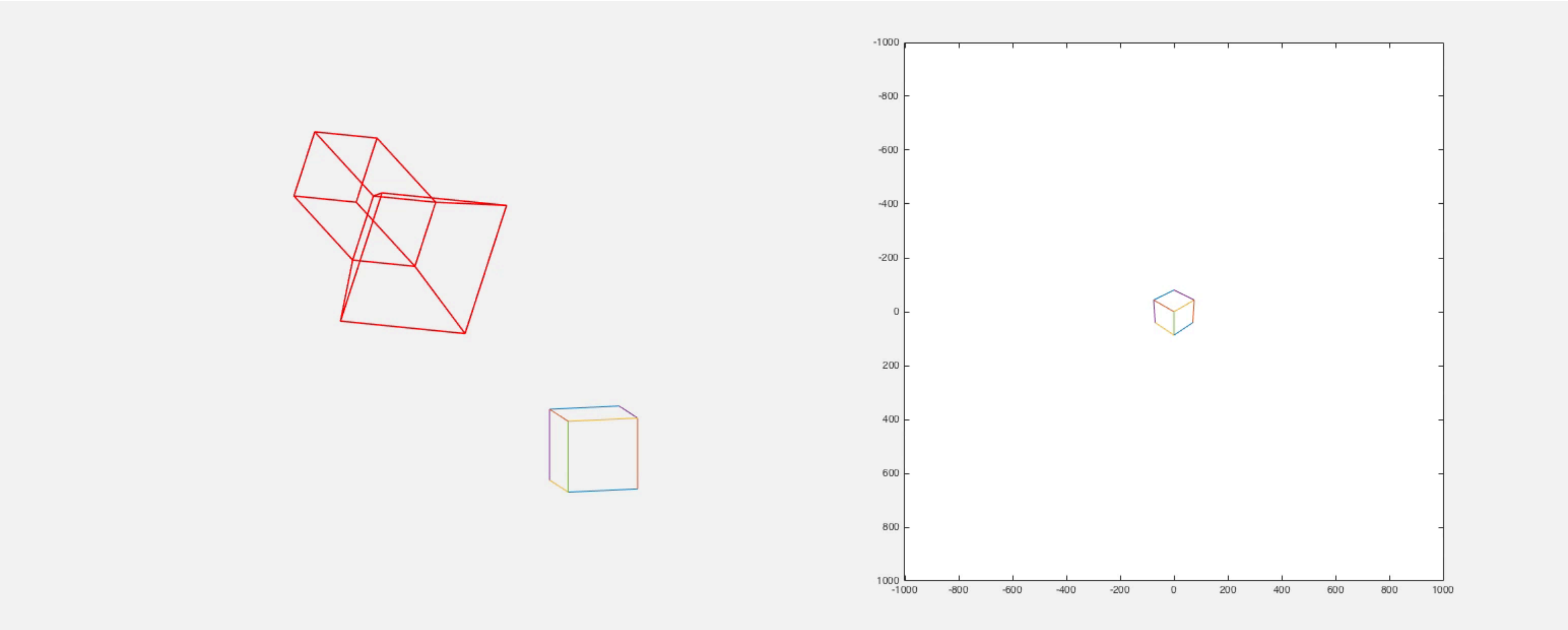
A projection is only required to preserve “straightness”/”collinearity”.

# Perspective effects

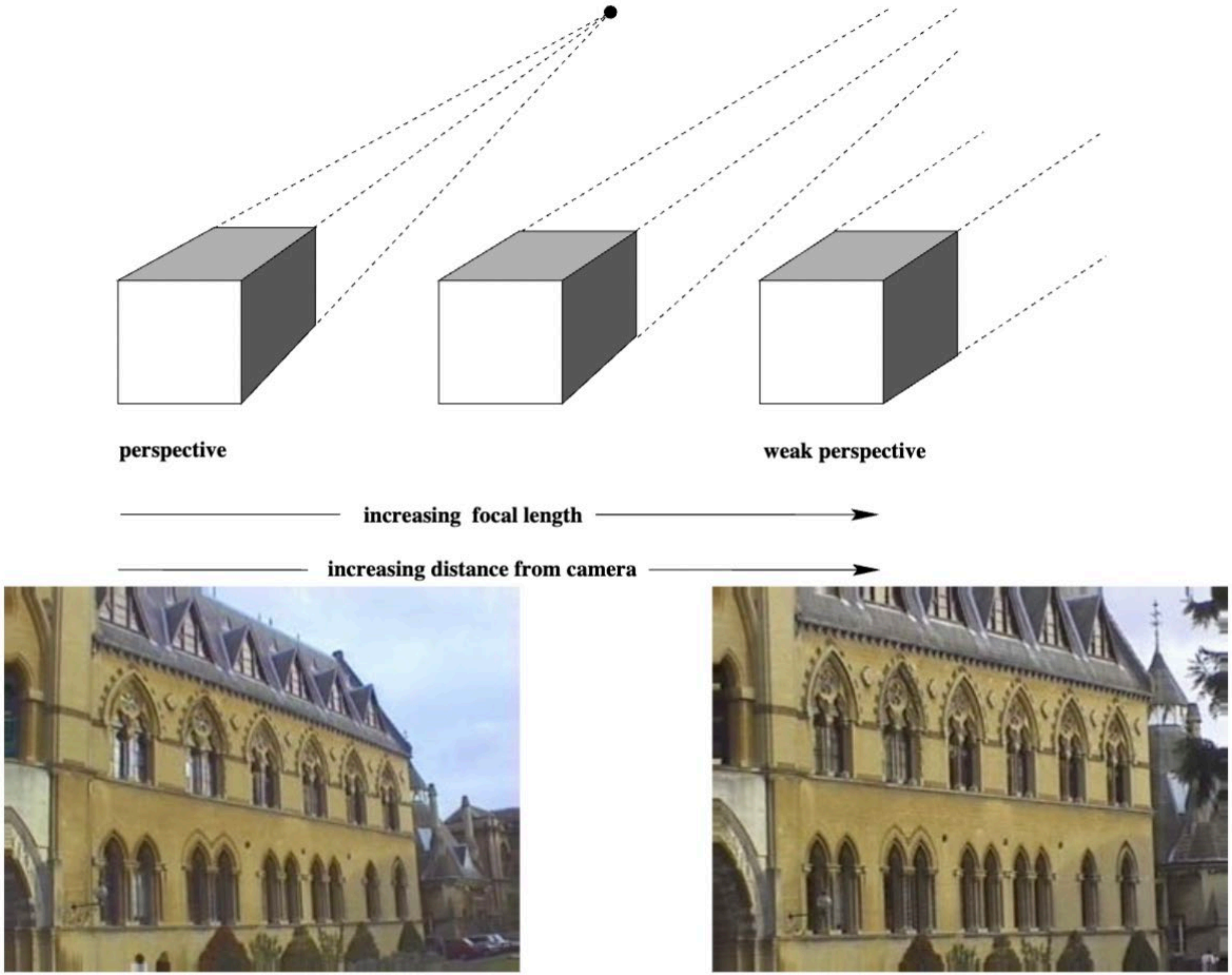



Parallel lines do not remain parallel !

# Perspective effects




Objects decrease in size with distance

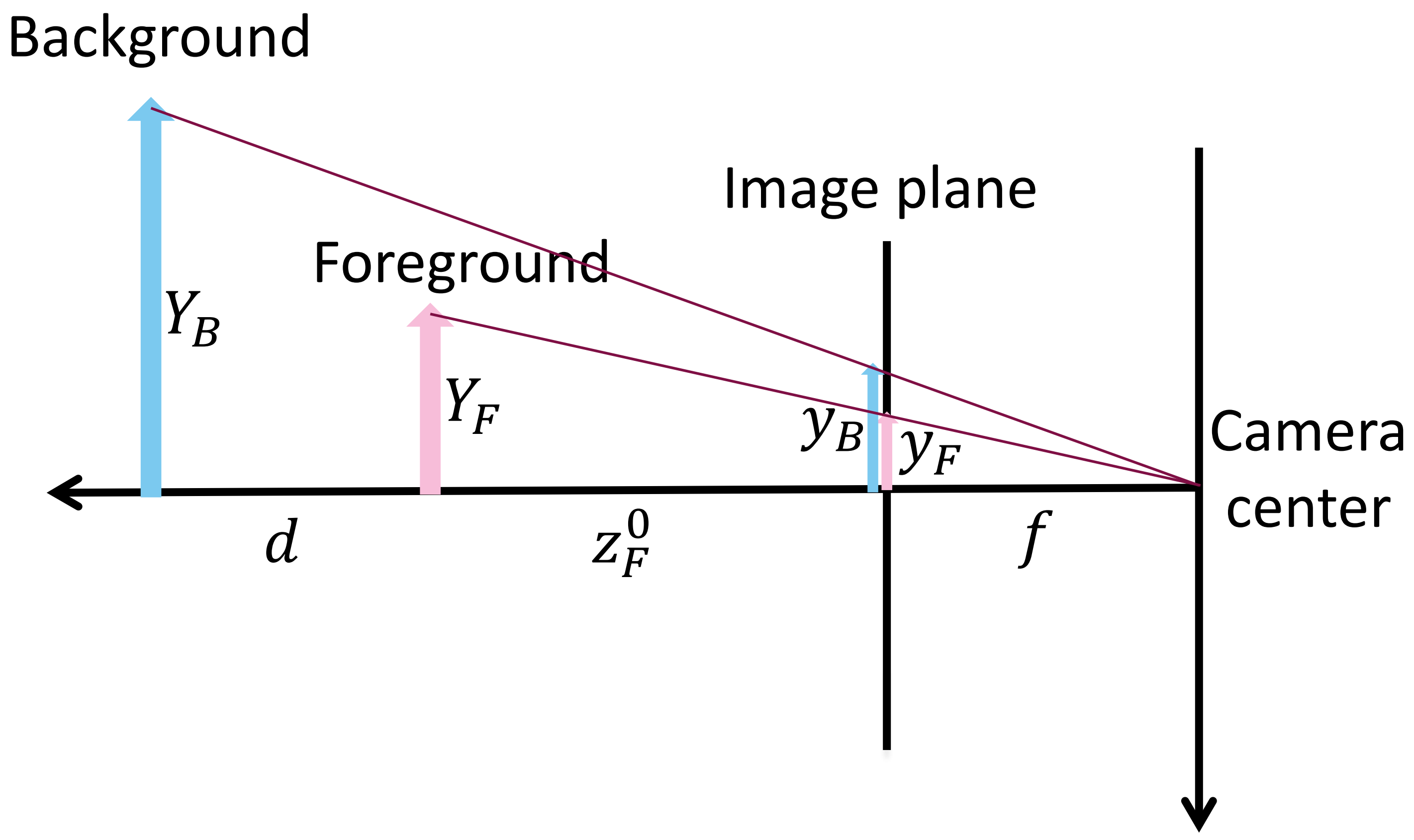

perspective
weak perspective
increasing focal length
increasing distance from camera

ZH Ch6

# Dolly Zoom or “Vertigo” Effect: z-motion and Zoom


Goodfellas, Martin Scorsese, 1990

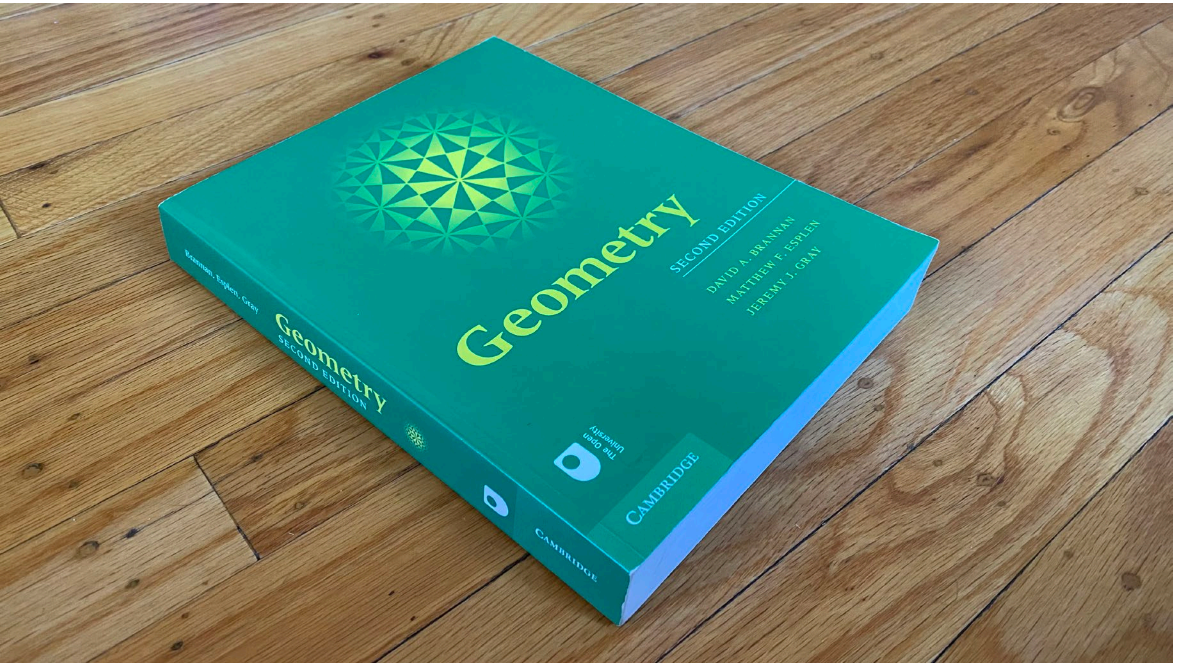
# Dolly Zoom or "Vertigo" Effect: z-motion and Zoom

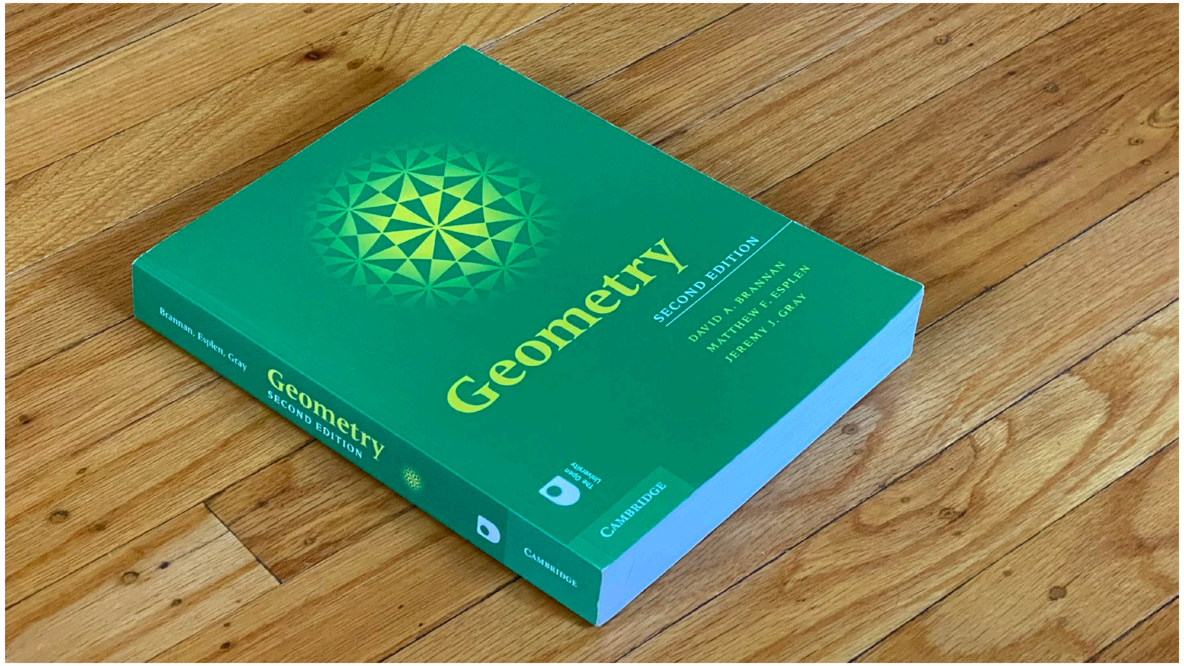


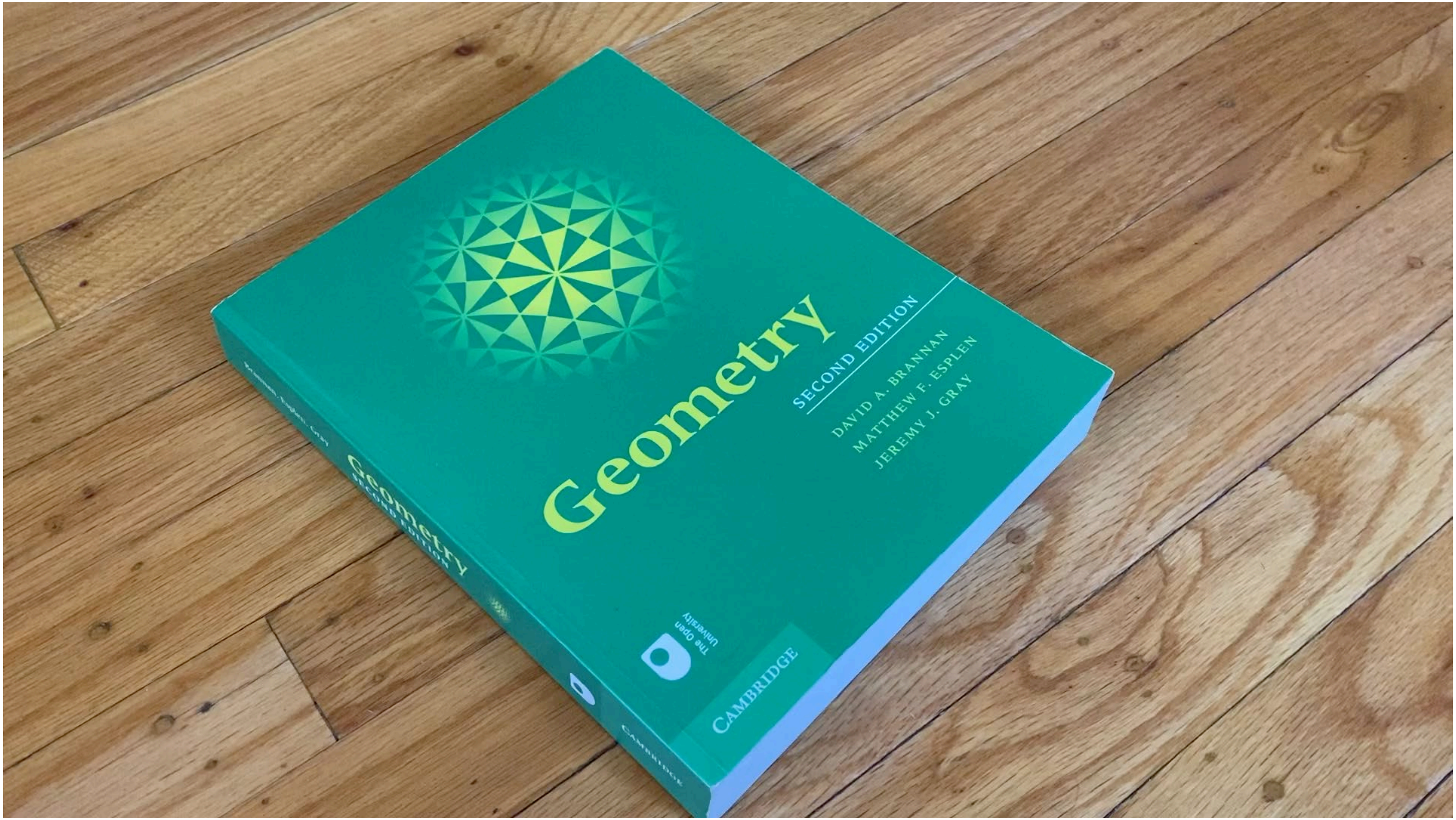

$$y_F = f \frac{Y_F}{Z_F}$$

$$y_B = f \frac{Y_B}{Z_B}$$

# Zoom vs distance…

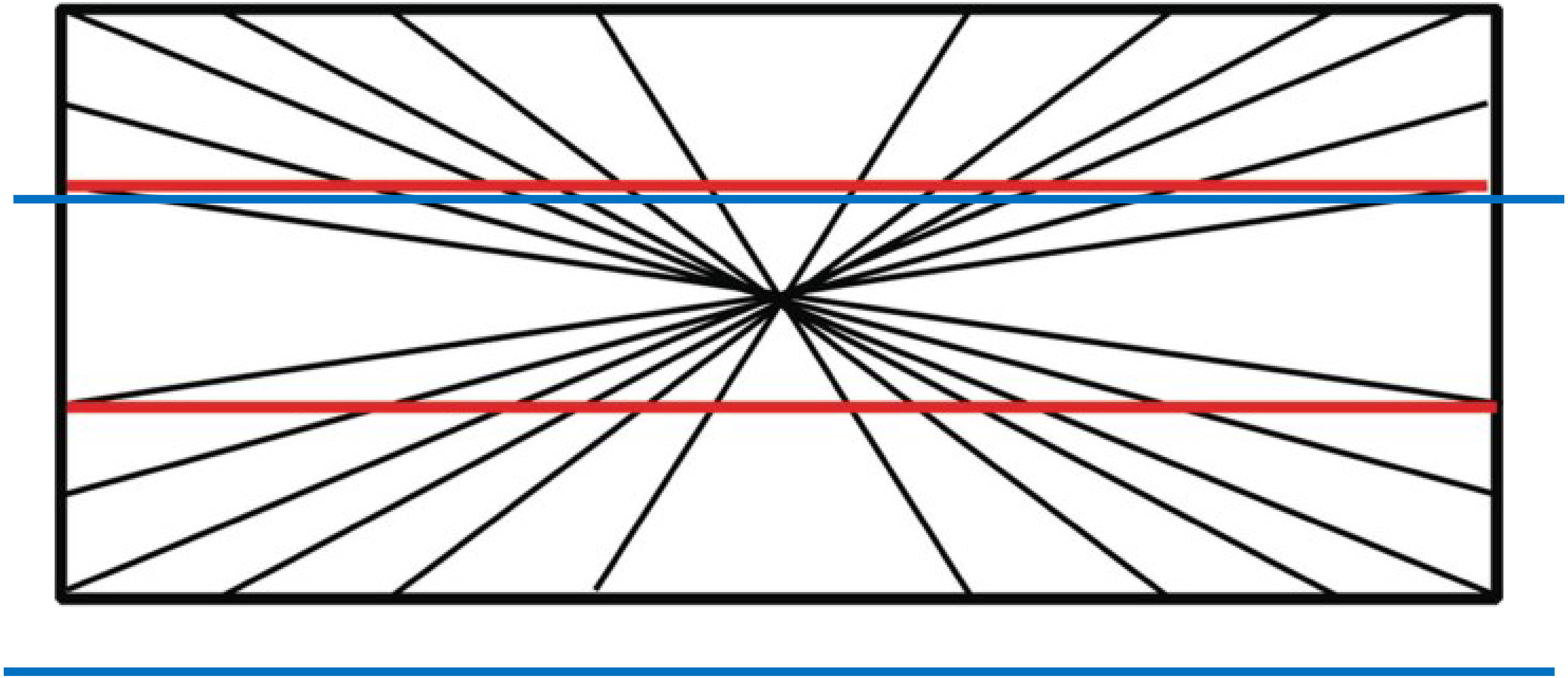


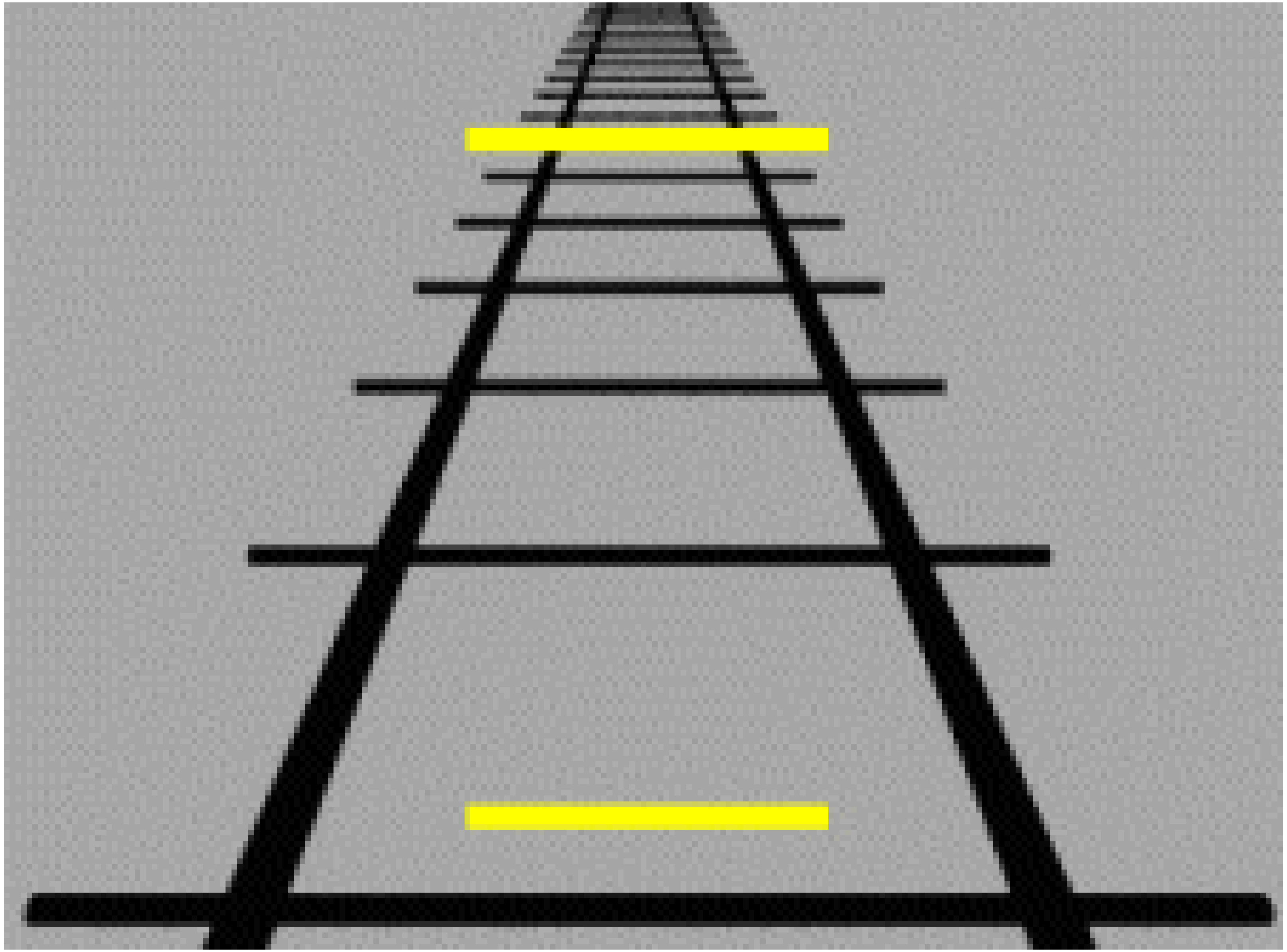



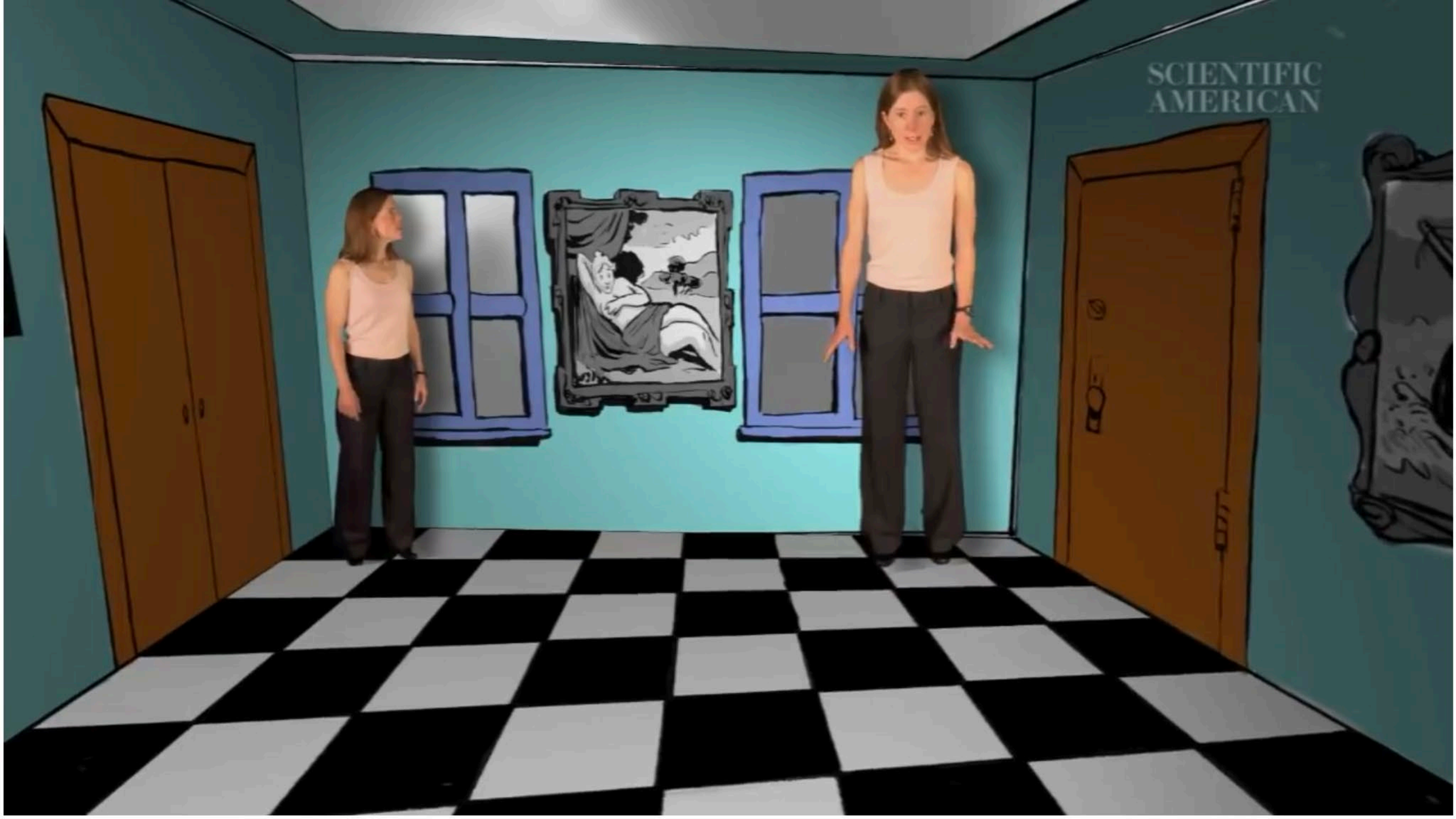

Geometry
SECOND EDITION
DAVID A. BRANNAN
MATTHEW F. ESPLEN
JEREMY J. GRAY
The Open University
CAMBRIDGE
Geometry
SECOND EDITION
CAMBRIDGE

# Human visual system is aware of perspective effects!

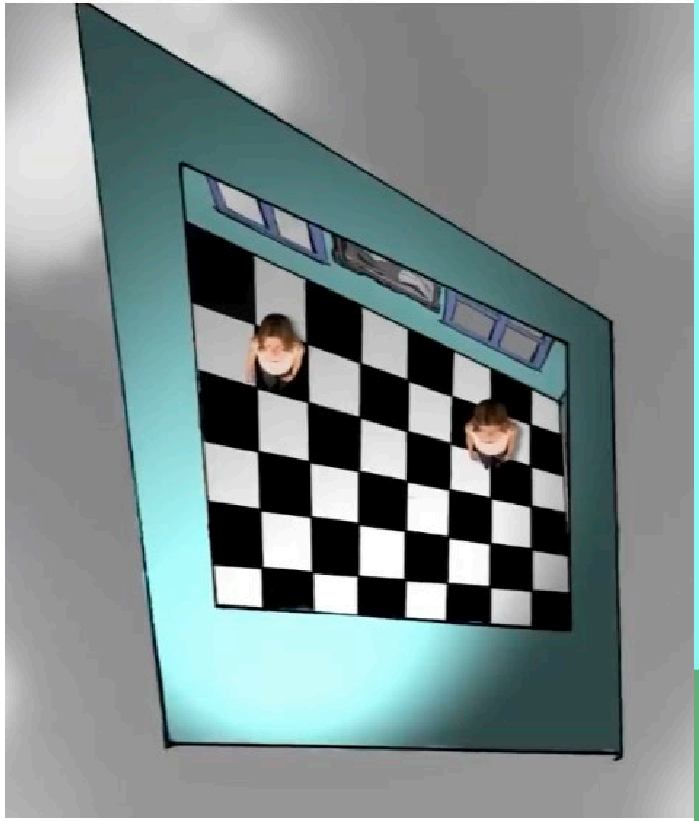
# Visual Illusion

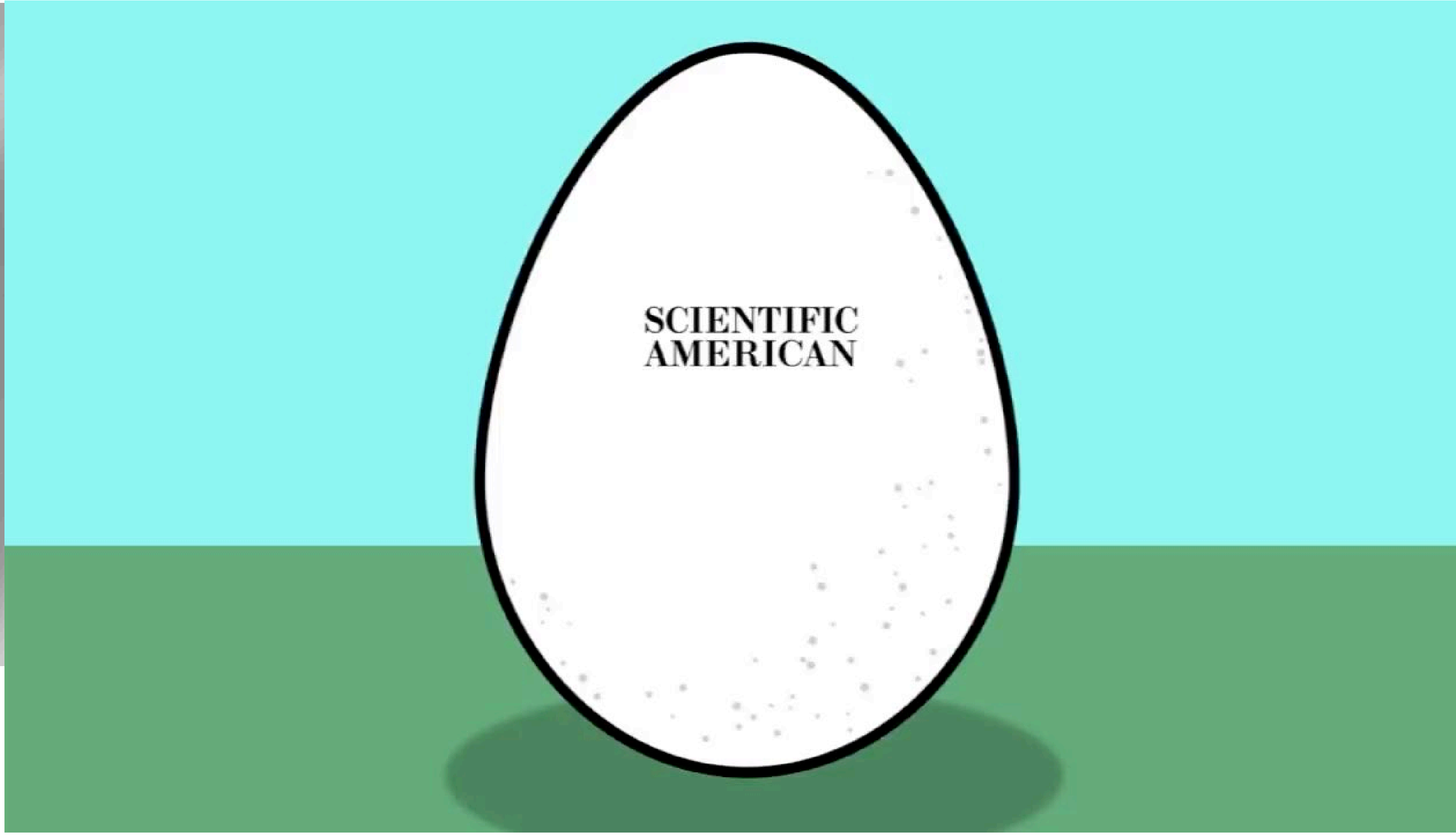





SCIENTIFIC
AMERICAN

# Ames Room Illusion






https://www.scientificamerican.com/video/instant-egghead-what-is-the-ames2012-10-09/

# Ames Room Illusion

https://www.scientificamerican.com/video/instant-egghead-what-is-the-ames2012-10-09/

# Vector Algebra Recap: Dot Products

- $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^T \boldsymbol{y} =< \boldsymbol{x}, \boldsymbol{y} > = \sum_i \boldsymbol{x}_i \boldsymbol{y}_i = ||\boldsymbol{x}||||\boldsymbol{y}|| \cos \boldsymbol{\theta}_{xy}$
  - Output is a scalar
  - Measures angles if applied to unit vectors.
  - Measures projection of one vector onto another: $||\boldsymbol{y}|| \cos \boldsymbol{\theta}_{xy}$ is projection of $\boldsymbol{y}$ onto the direction of $\boldsymbol{x}$

- Orthogonality: Two vectors are orthogonal iff $\boldsymbol{x}^T \boldsymbol{y} = \boldsymbol{0}$
  - The equation of a plane in 3D is $\boldsymbol{\pi}^T \boldsymbol{x} = \boldsymbol{0}$
    - This means the vector $\boldsymbol{\pi}$ is normal to the plane
    - (same reasoning also also applies for a 2D line $\boldsymbol{l}$)
- $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$

# Vector Algebra Recap: Matrix Products

- $M\boldsymbol{x} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}_{2\times 2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{2\times 1} = \begin{bmatrix} m_{11}x_1 + m_{12}x_2 \\ m_{21}x_1 + m_{22}x_2 \end{bmatrix}_{2\times 1}$

- Can be reframed in terms of dot products.
  - $\boldsymbol{m}_1 = [m_{11}, m_{12}]$
  - $\boldsymbol{m}_2 = [m_{21}, m_{22}]$

  - Then $M\boldsymbol{x} = \begin{bmatrix} \boldsymbol{m}_1 \cdot \boldsymbol{x} \\ \boldsymbol{m}_2 \cdot \boldsymbol{x} \end{bmatrix}_{2\times 1}$