Image Formation: Pinhole Model, Perspective Projection, and Binocular Stereo

Lecture by Margrit Betke, CS 585, February 6, 2024

BOSTON UNIVERSITY

Pinhole camera

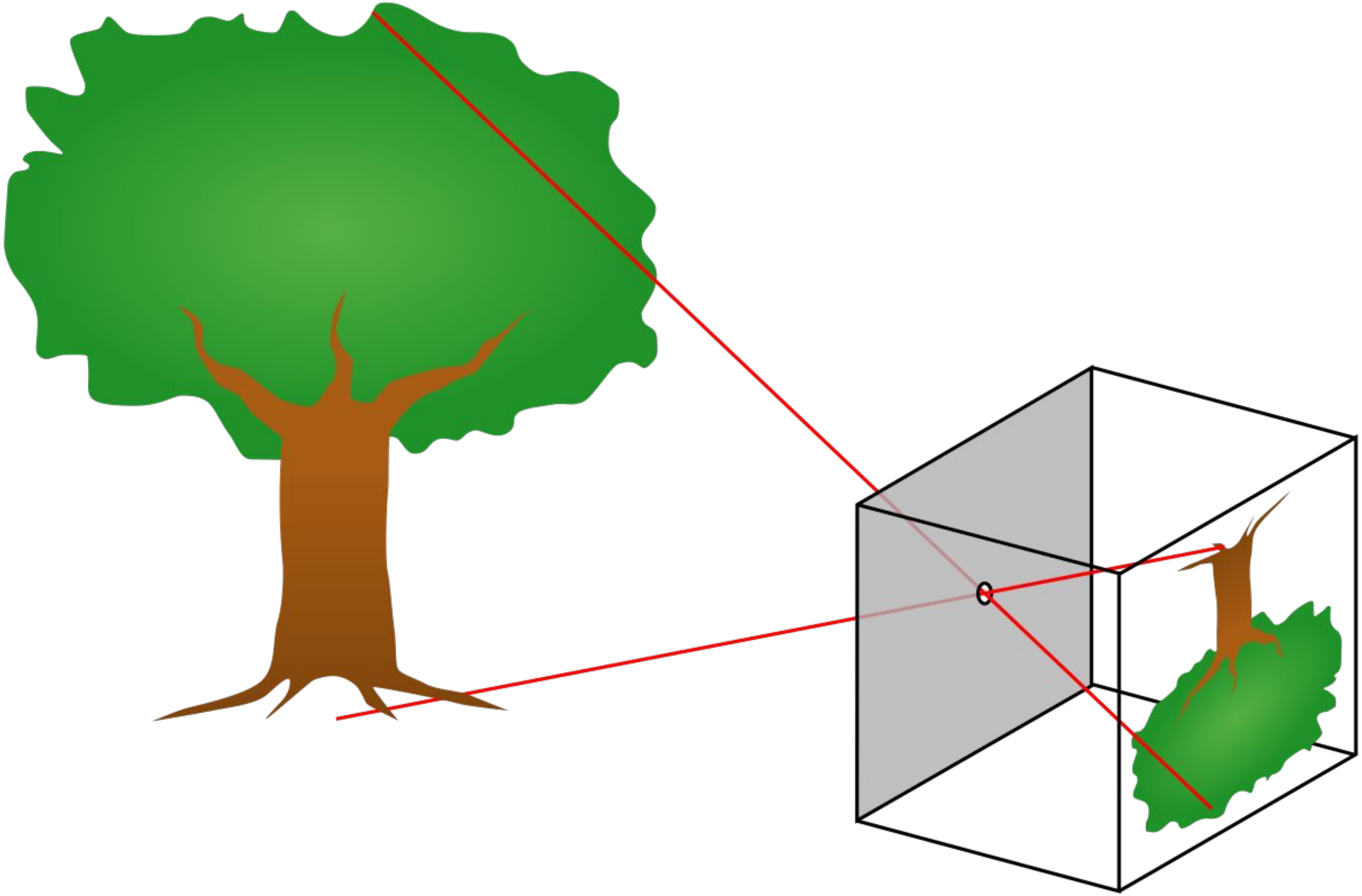


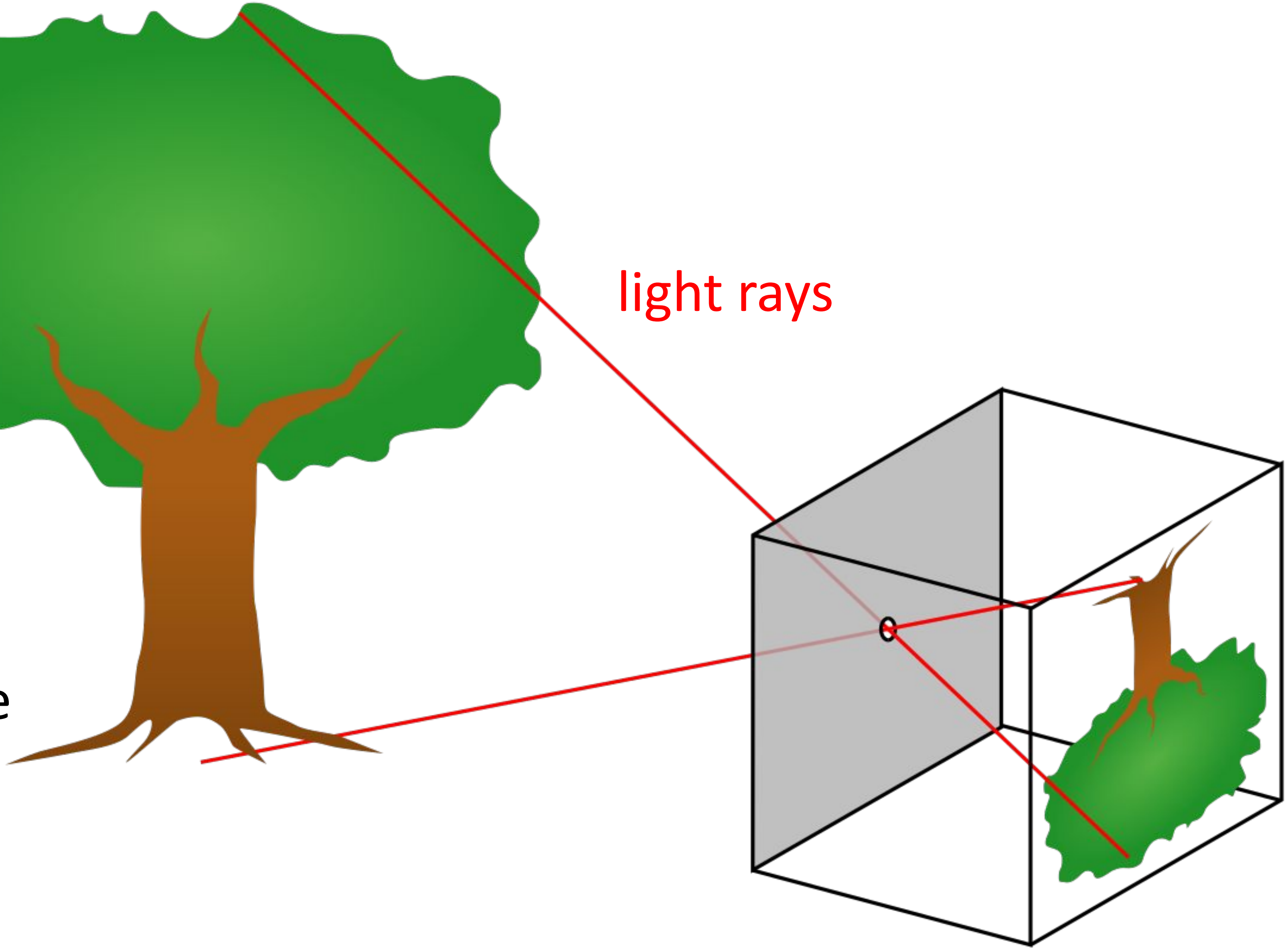
Image credit: public domain, Wikipedia

Pinhole camera

=

Box with a hole

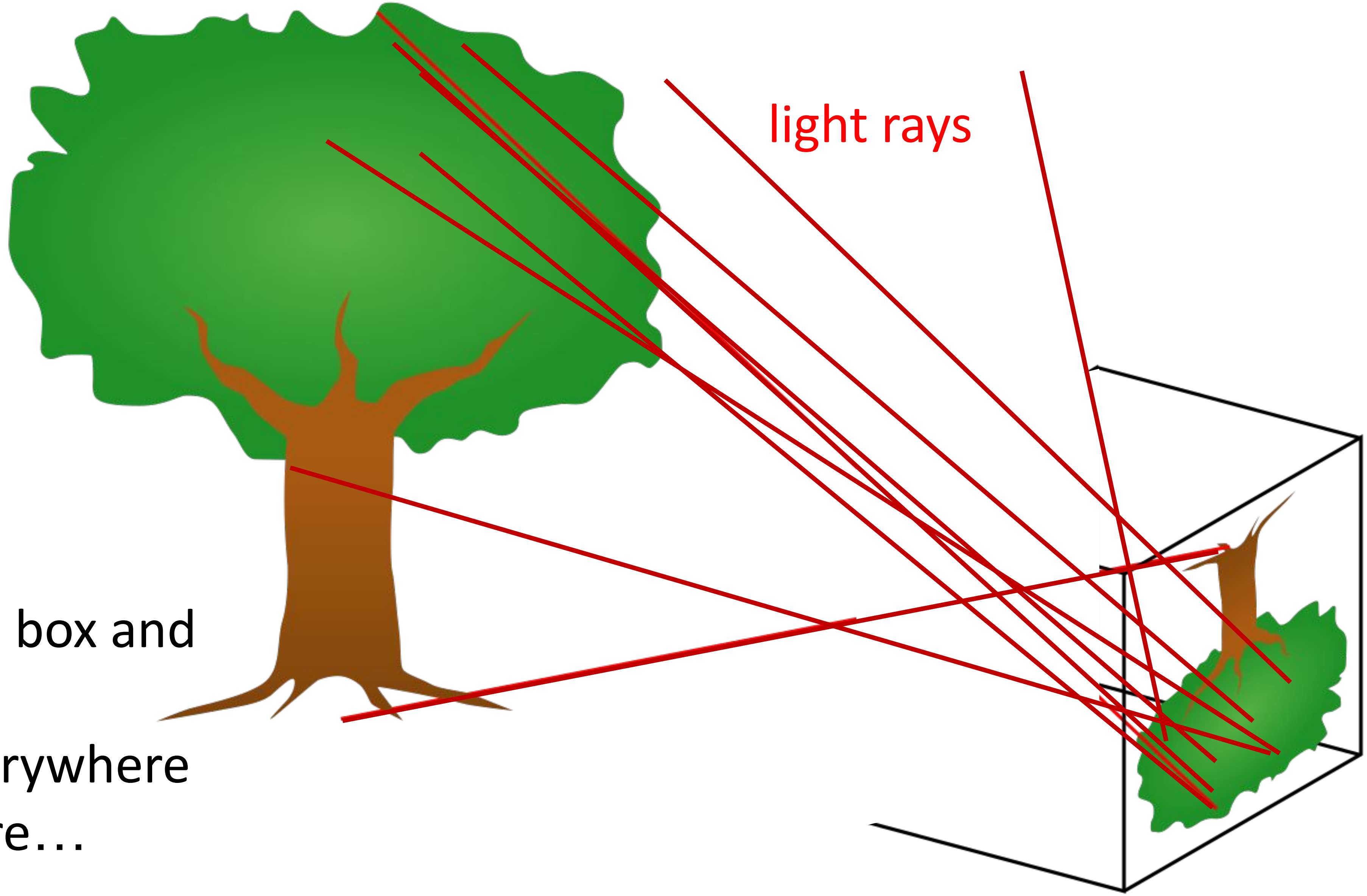
Image is projected upside down on side opposite to hole.



Pinhole camera

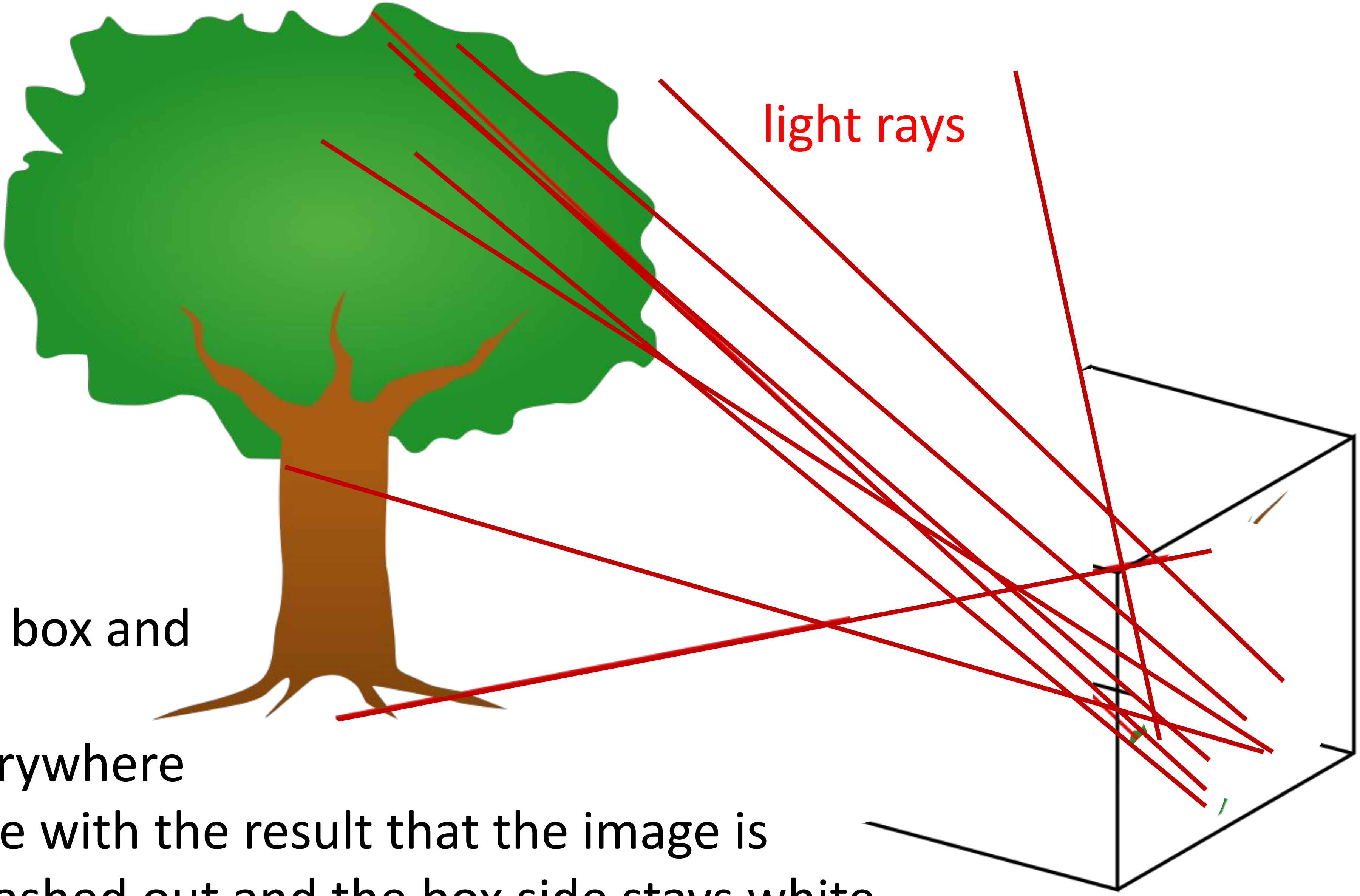
Why is a hole needed?

With an open box and no hole, light rays from everywhere go everywhere…



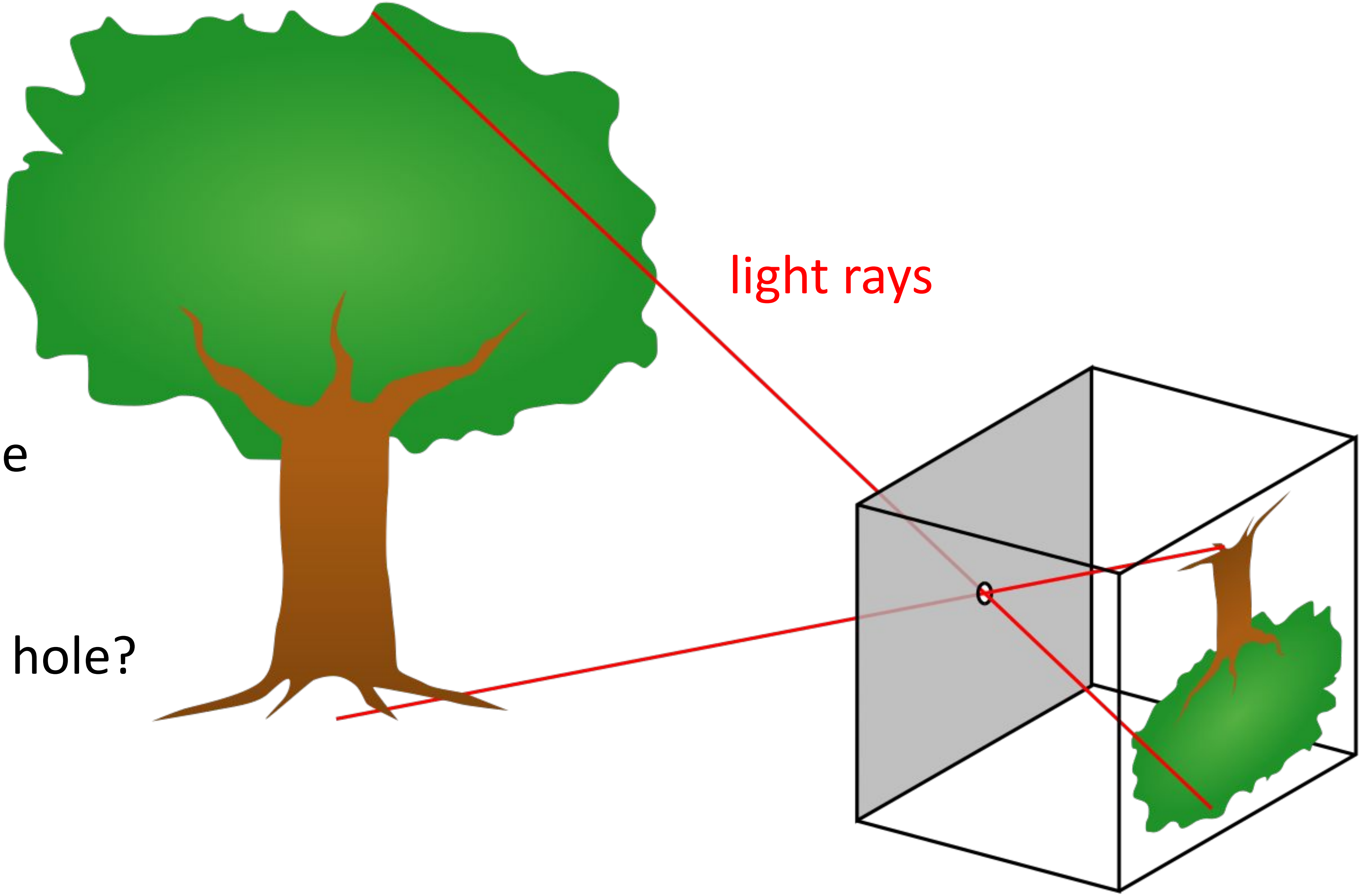
Pinhole camera

Why is a hole needed?



With an open box and no hole, light rays from everywhere go everywhere with the result that the image is completely washed out and the box side stays white

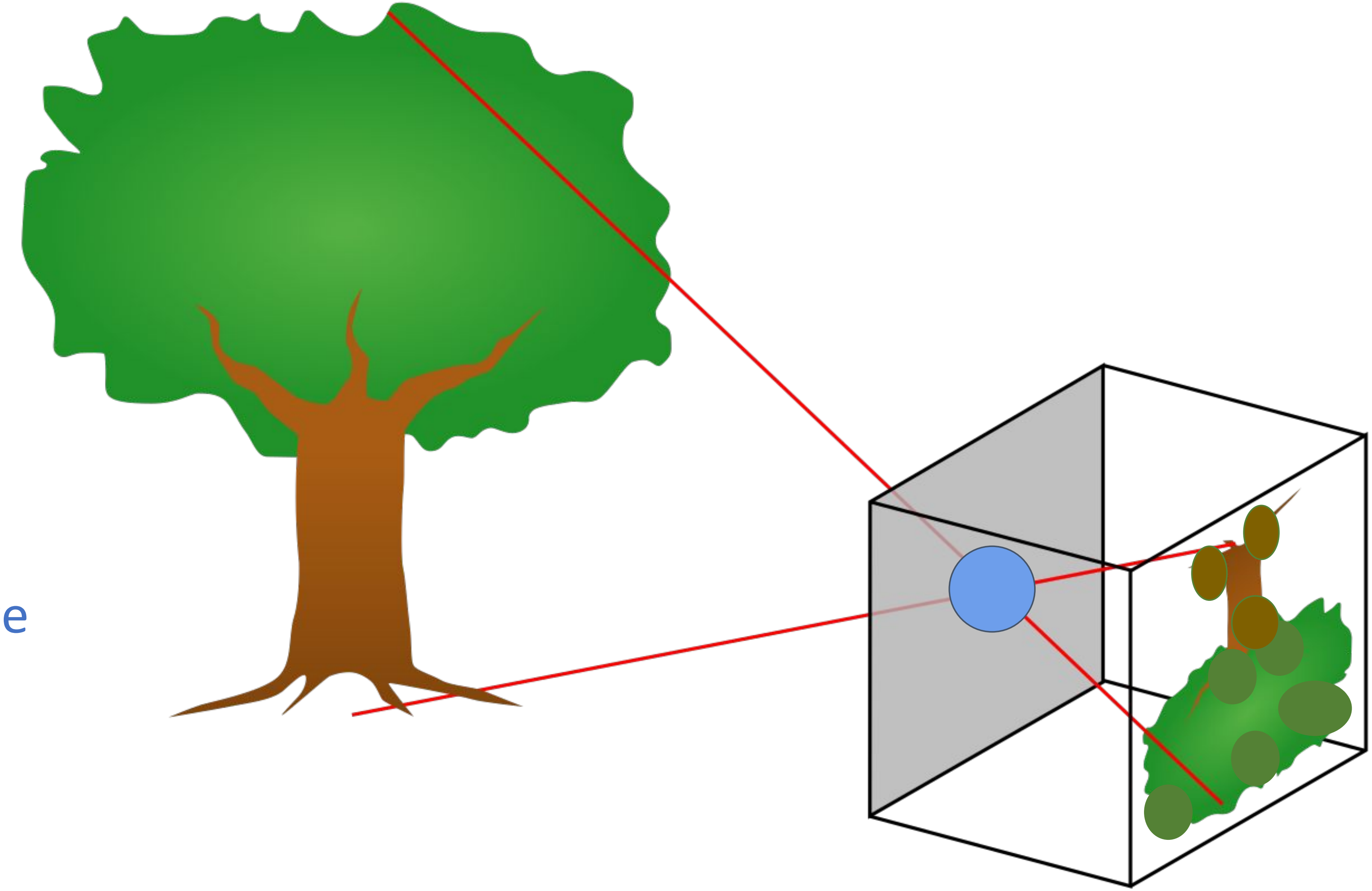
Pinhole camera



Ok, so the hole is crucial.

But how big a hole?

Pinhole camera



Larger hole:

Blurrier image

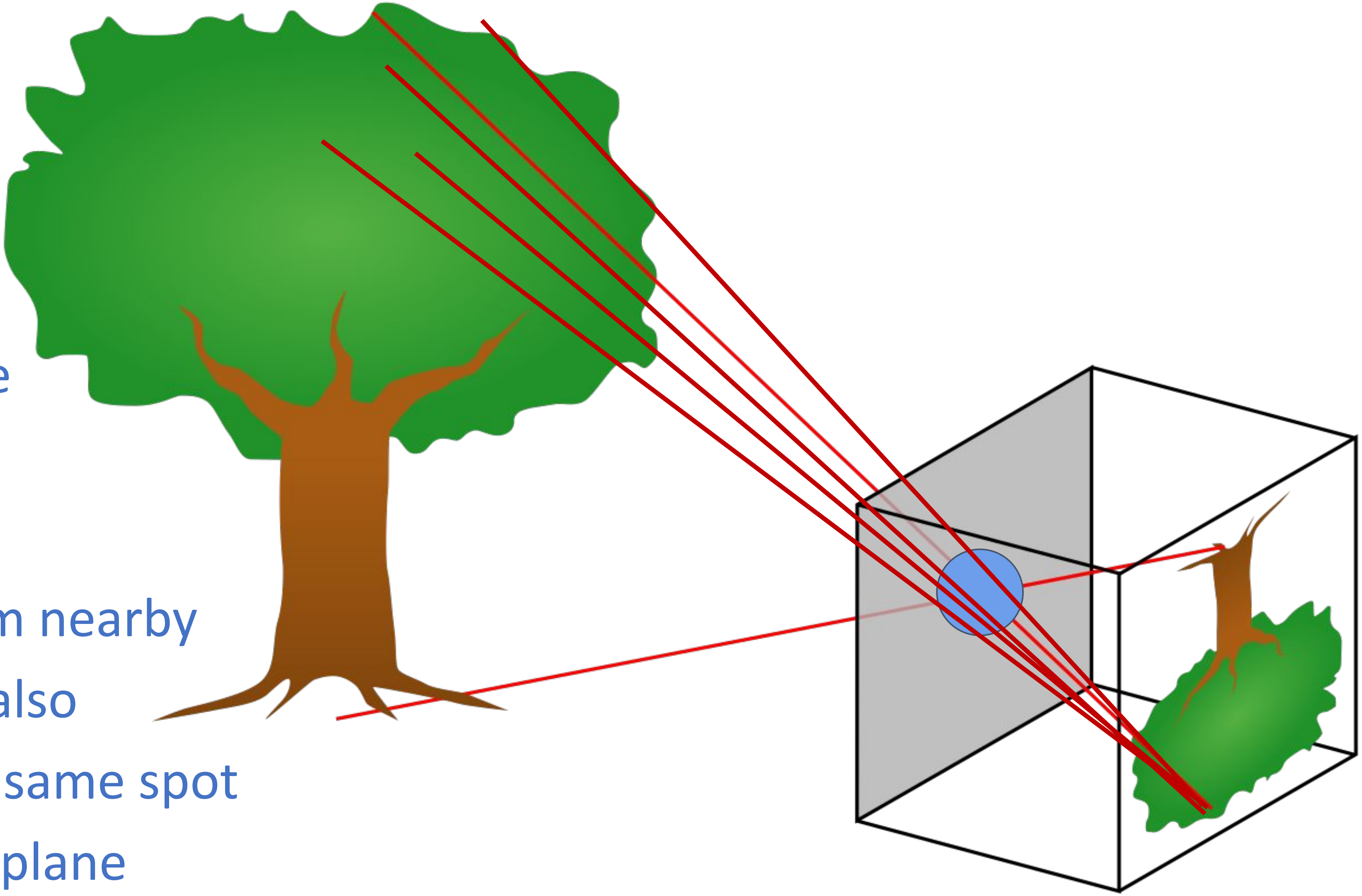
Pinhole camera

Larger hole:

Blurrier image

Why?

Light rays from nearby scene points also end up in the same spot on the image plane



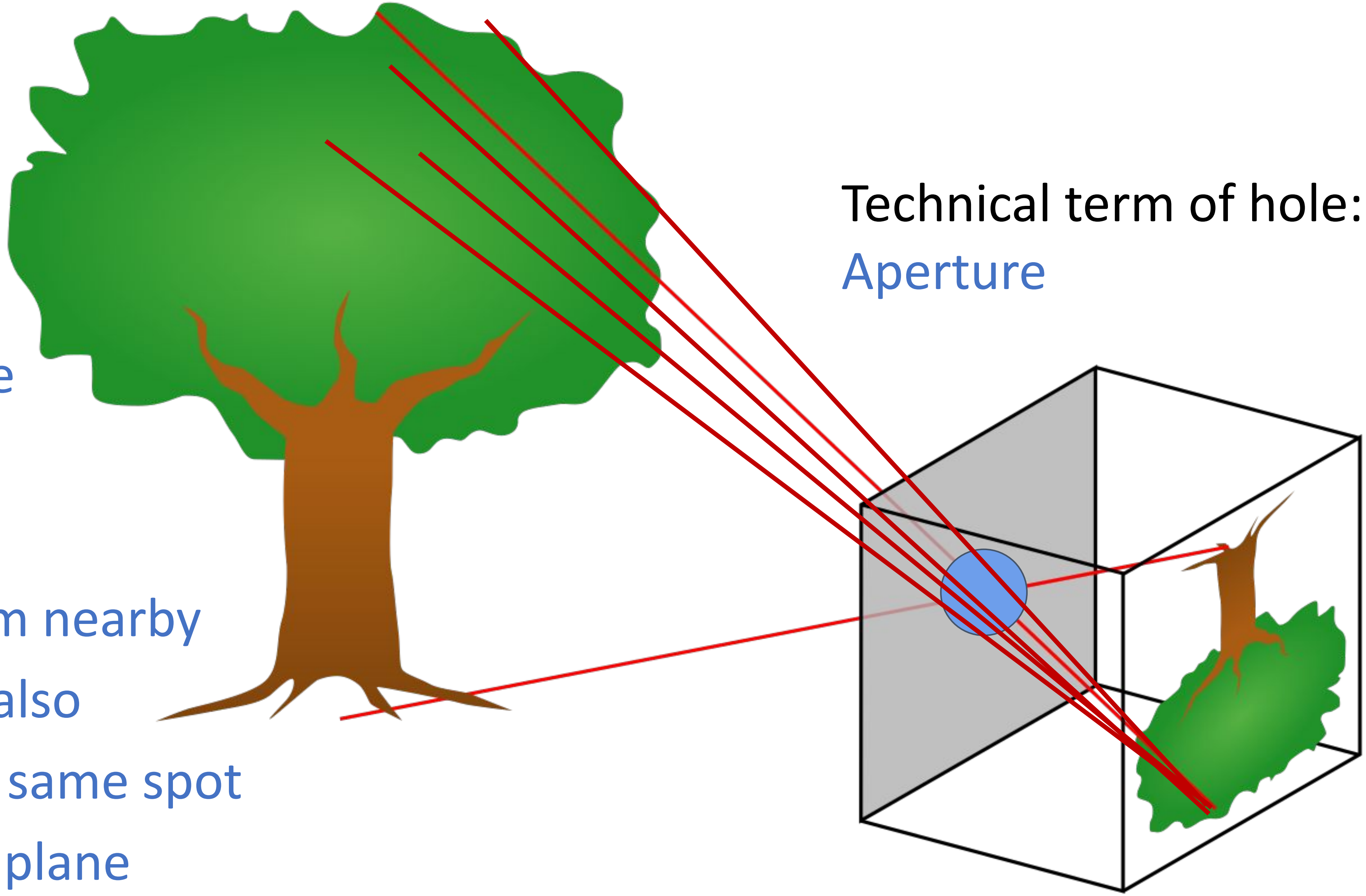
Pinhole camera

Larger ~~hole~~:

Blurrier image

Why?

Light rays from nearby scene points also end up in the same spot on the image plane

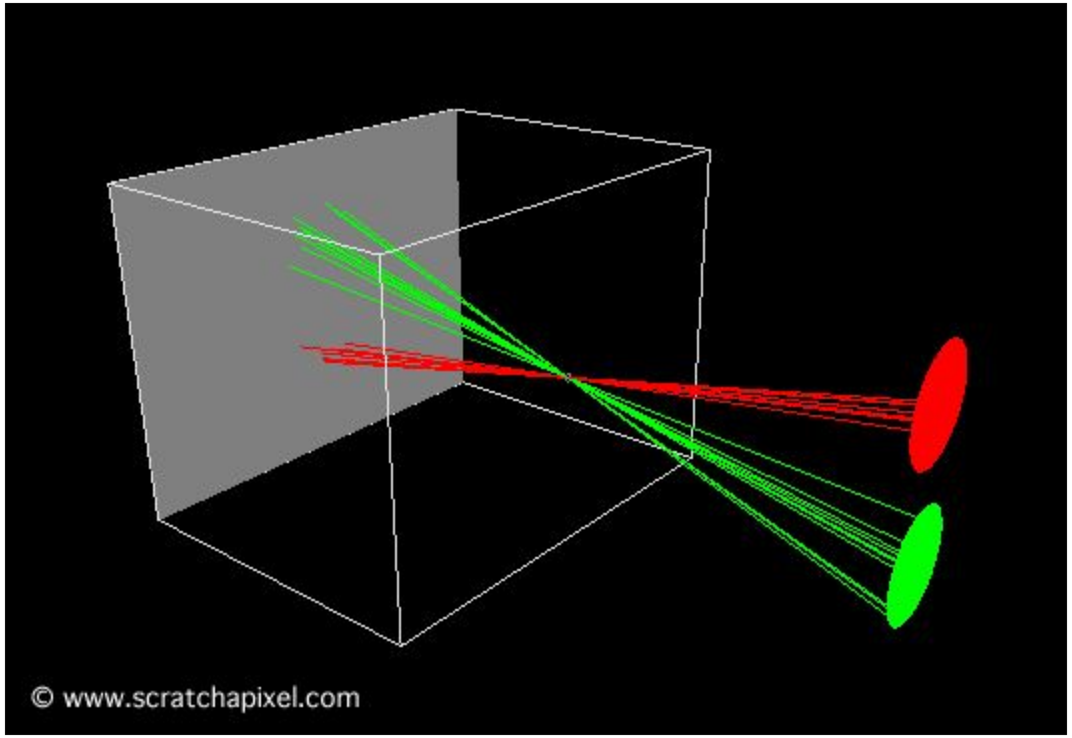


Technical term of hole:

Aperture

BOSTON UNIVERSITY

Small Aperture



Large Aperture

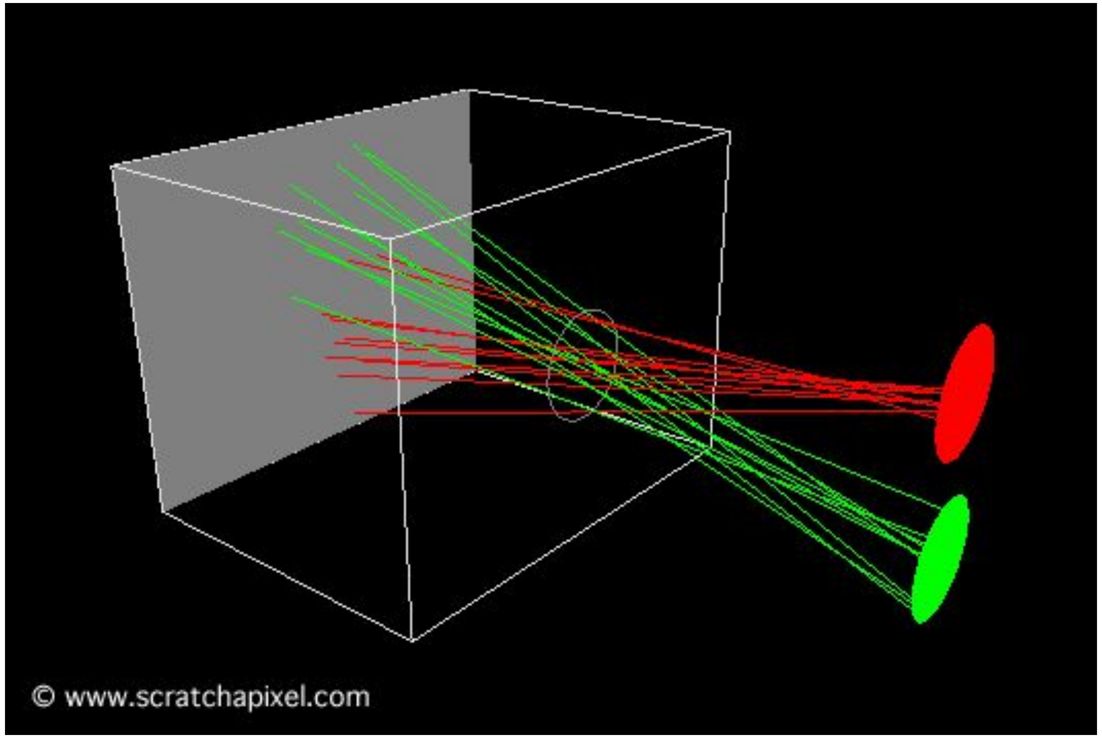


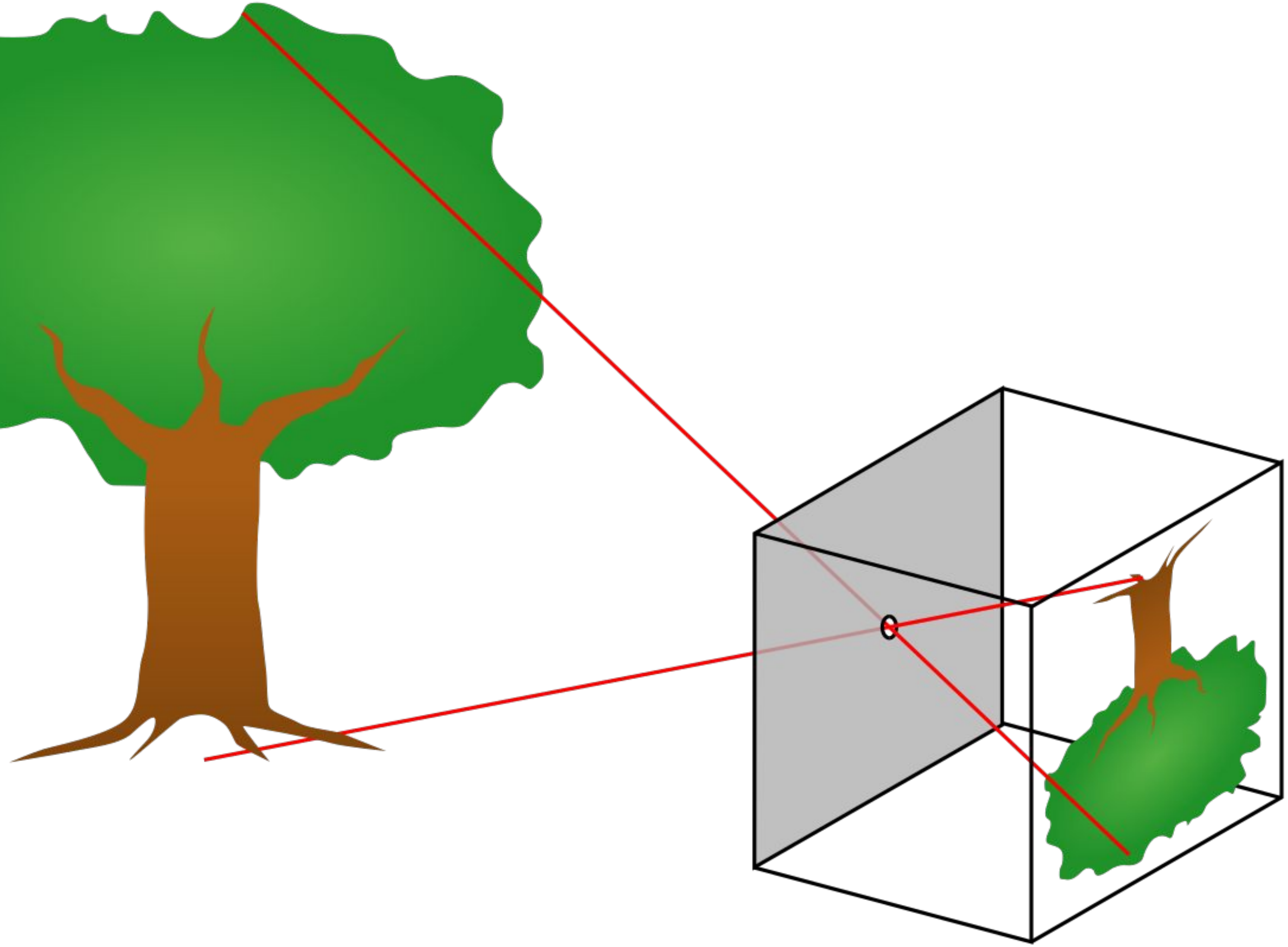
Image credit: www.scratchpixel.com

Pinhole camera

Smaller aperture

Fainter image

because less light travels through pinhole

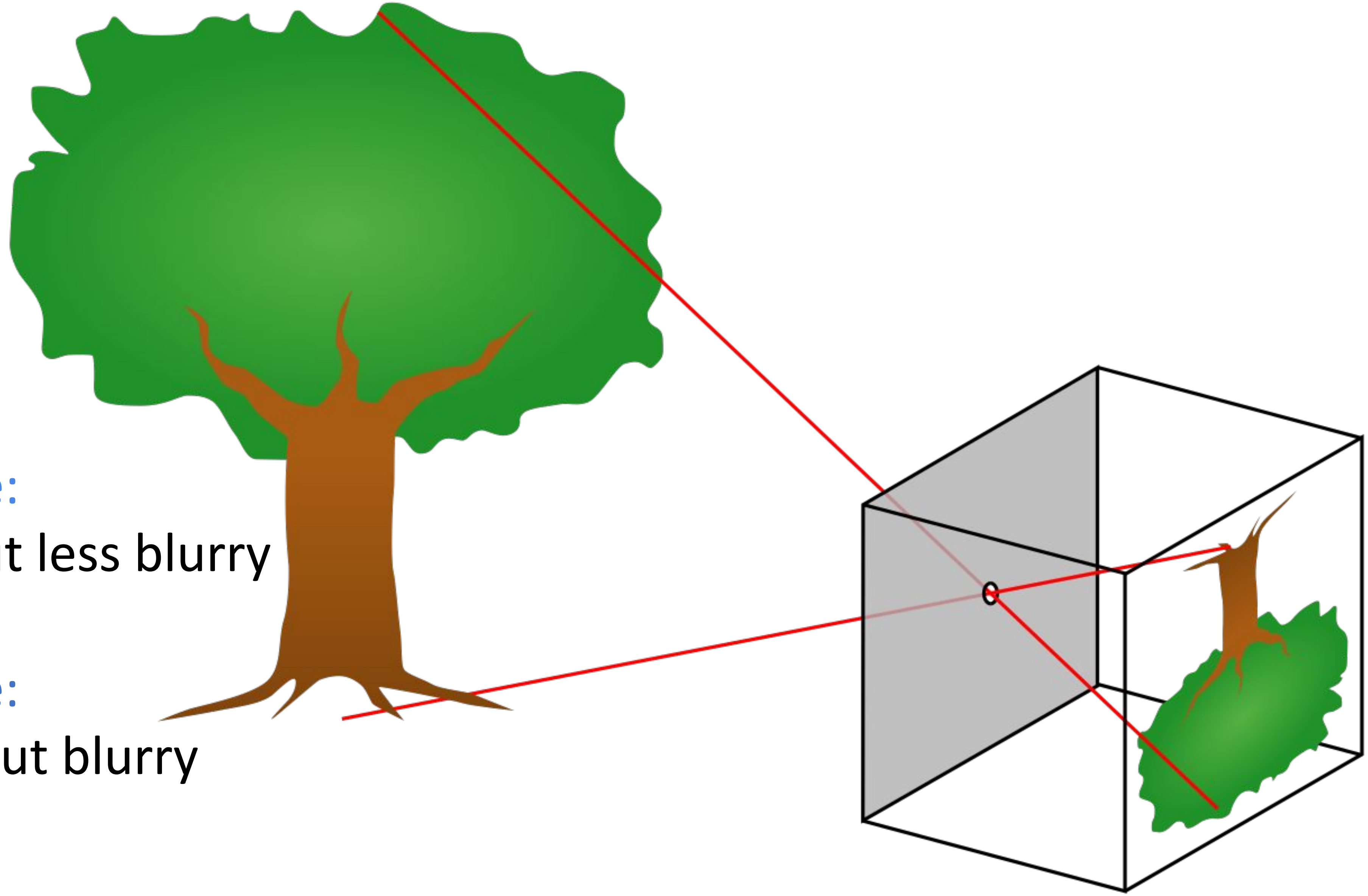


Pinhole camera

Trade off:

Small aperture:
Faint image but less blurry

Large aperture:
Bright image but blurry



Watch Steve Seitz' Video:

https://www.youtube.com/watch?v=F5WA26W4JaM&list=PLWfDJ5nla8UpwShx-lzLJqcp575fKpsSO&index=11

until 2:58

Pinhole camera = Camera obscura

Camera means ‘room’ and obscura means ‘dark’ in Latin.

Historic descriptions in

- Chinese Mozi writings (~ 500 BCE),
- Aristotelian Problems (~ 300 BCE),
- Arab writings (~1000 CE).

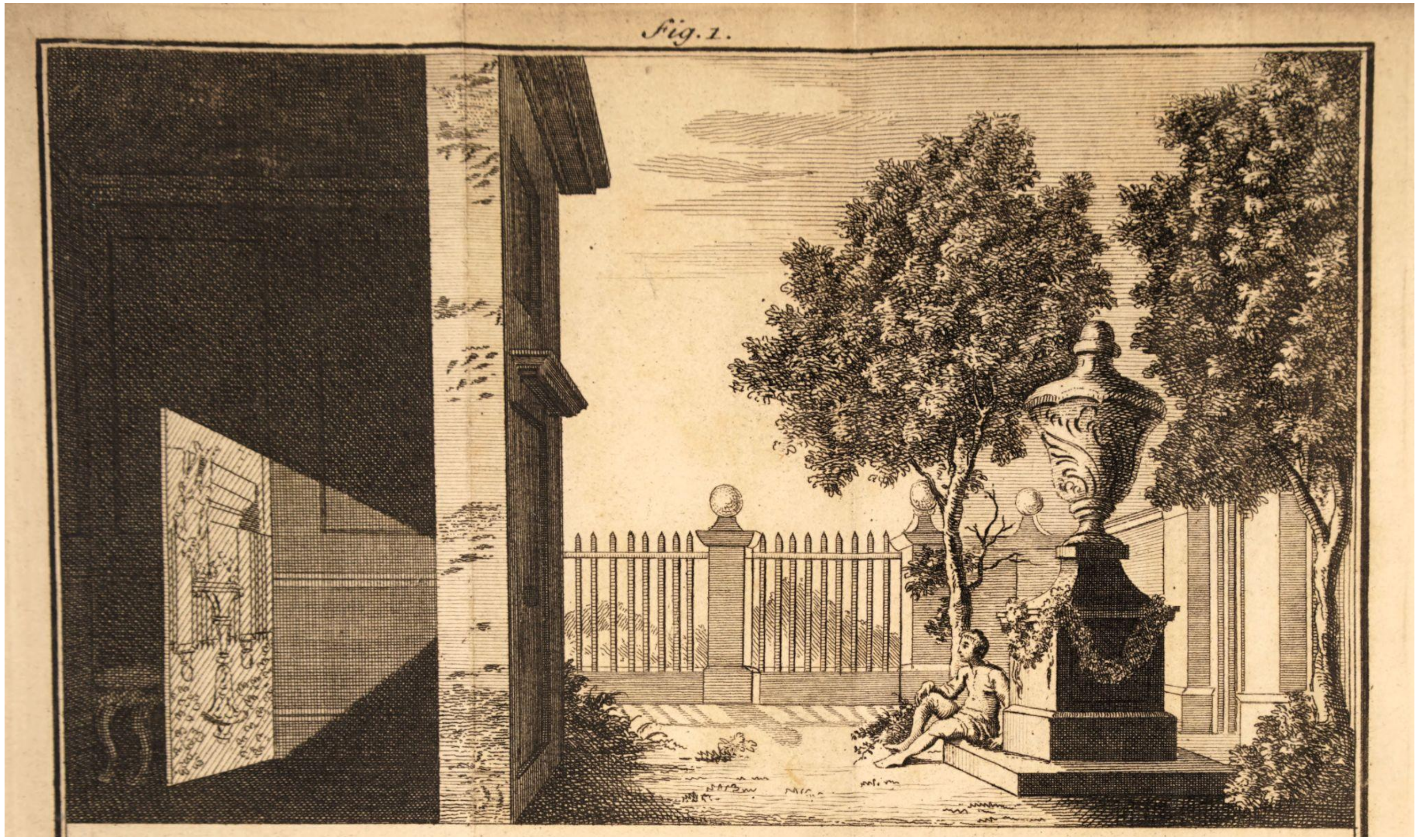


Image credit: James Ayscough, 1755

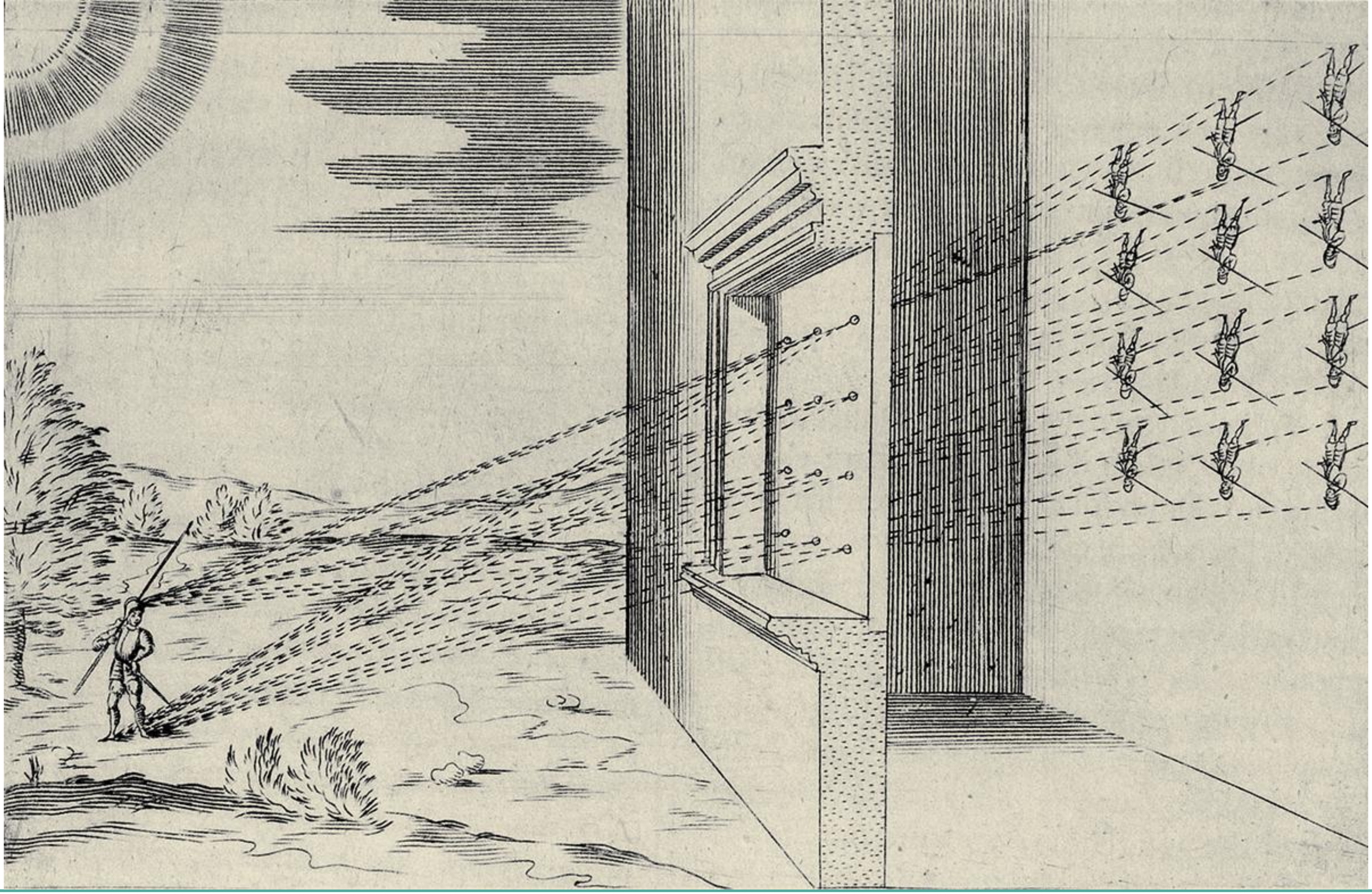
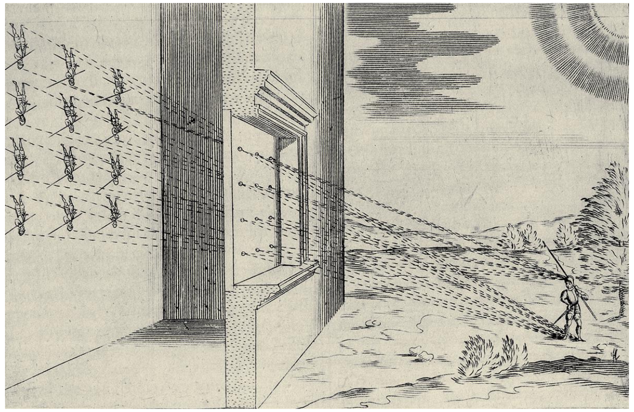
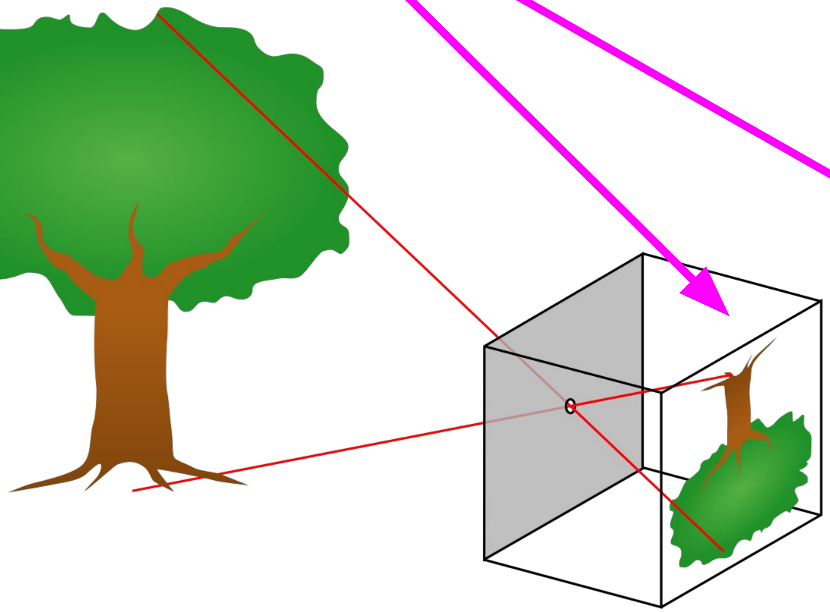


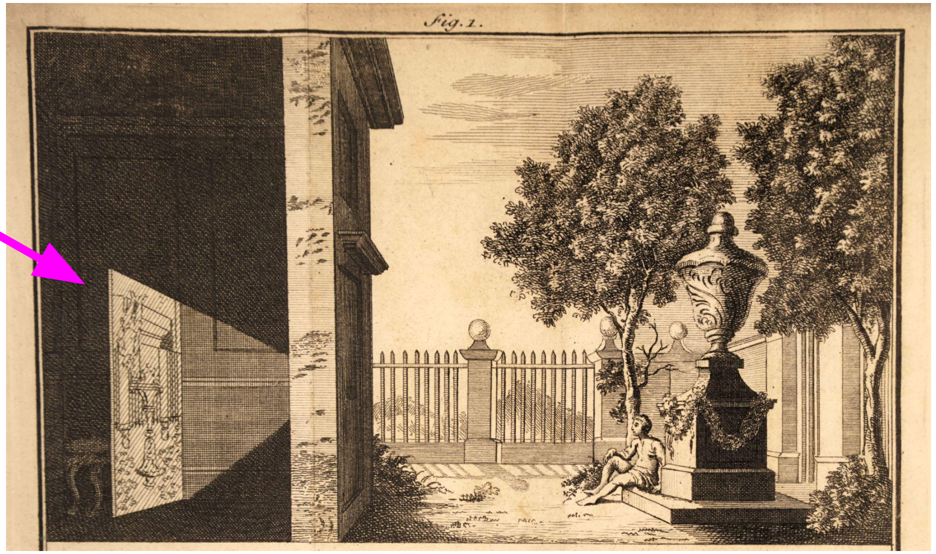
Image credit: Mario Bettini, 1642

Pinhole camera = Camera obscura



Problem: Real-time images cannot be stored!





Development of Camera Obscura to Modern Camera

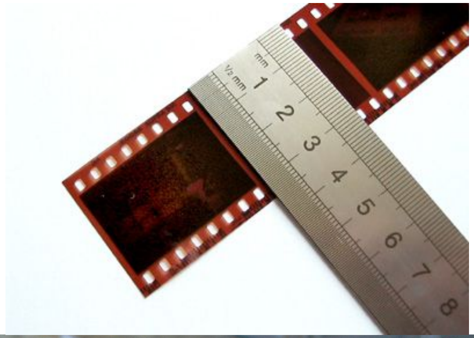
The first permanent photoetching was an image produced in 1822 by the French inventor Nicéphore Niépce.

Image credit: Jonnychiwa, Wikipedia

Development of Camera Obscura to Modern Camera

Film as a storage medium:

The first flexible photographic roll film was marketed by George Eastman, founder of Kodak in 1885.

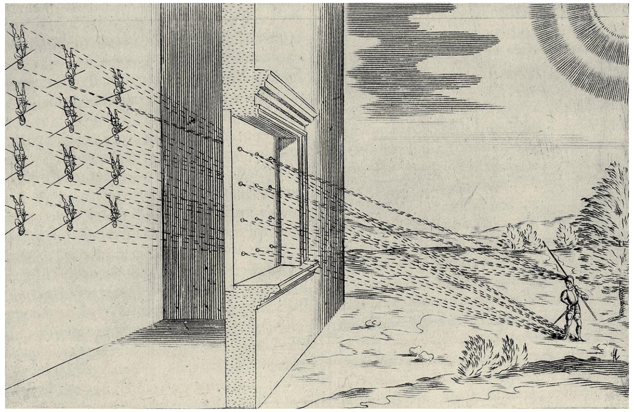


Array of linked capacitors as storage medium:

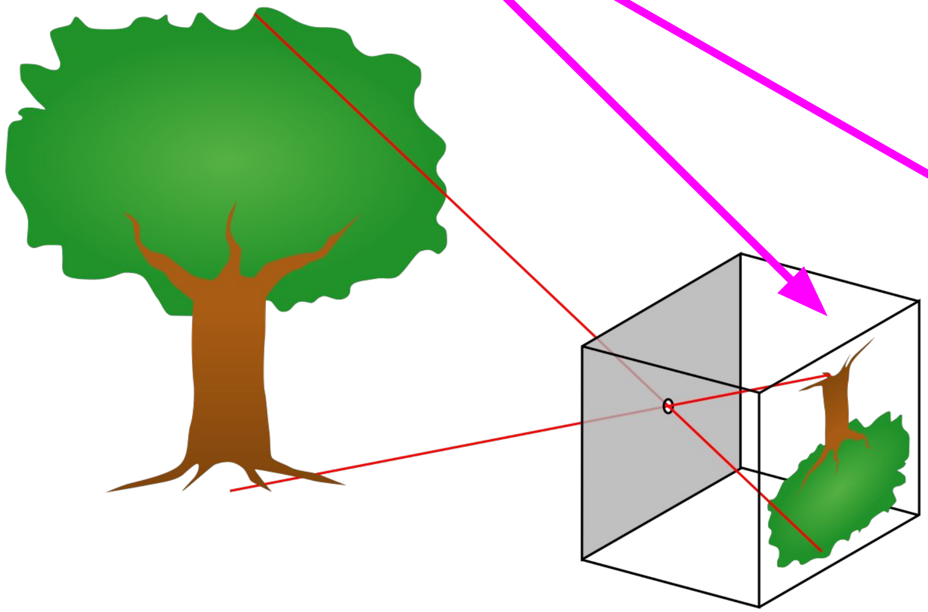
Sony unveiled the first consumer camera (Mavica) to use a charge-coupled device (CCD) for imaging, eliminating the need for film, in 1981.

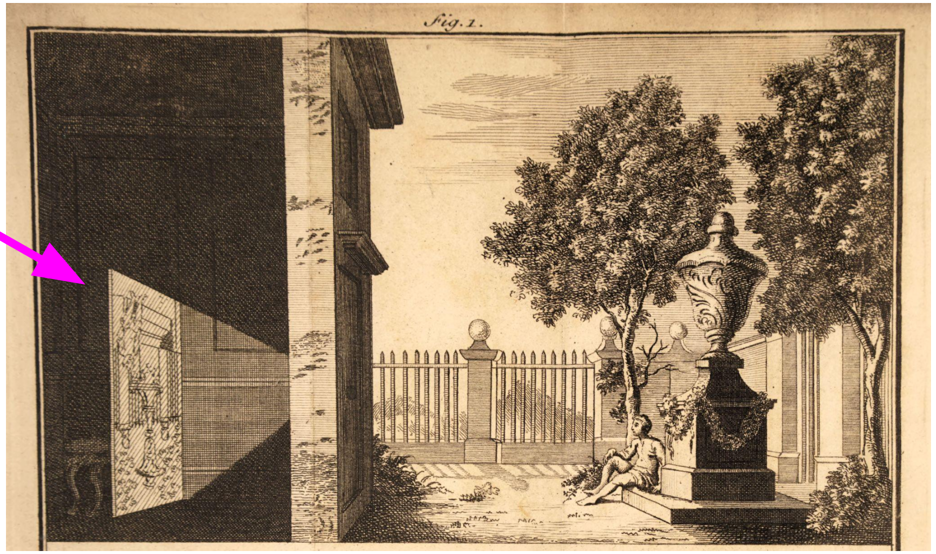
Image credits: Morio, Airport-1975, Wikipedia

Modern Film or Digital Cameras



Problem solved: Real-time images can be stored!





So then, why do we care about pinhole cameras in CS 585?

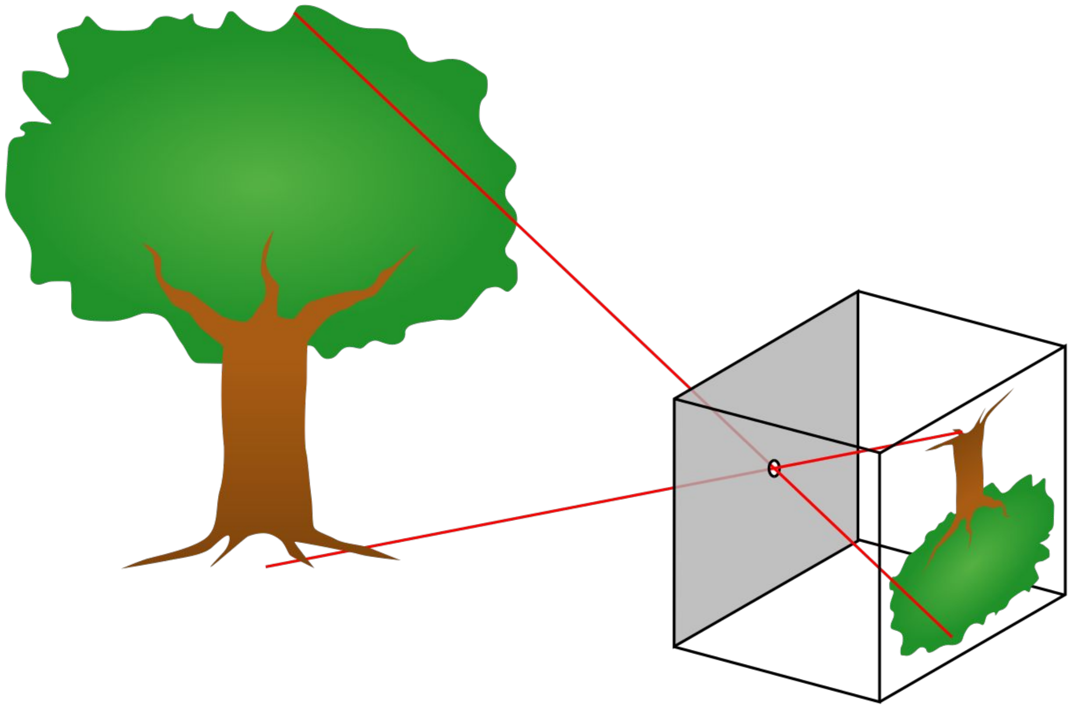
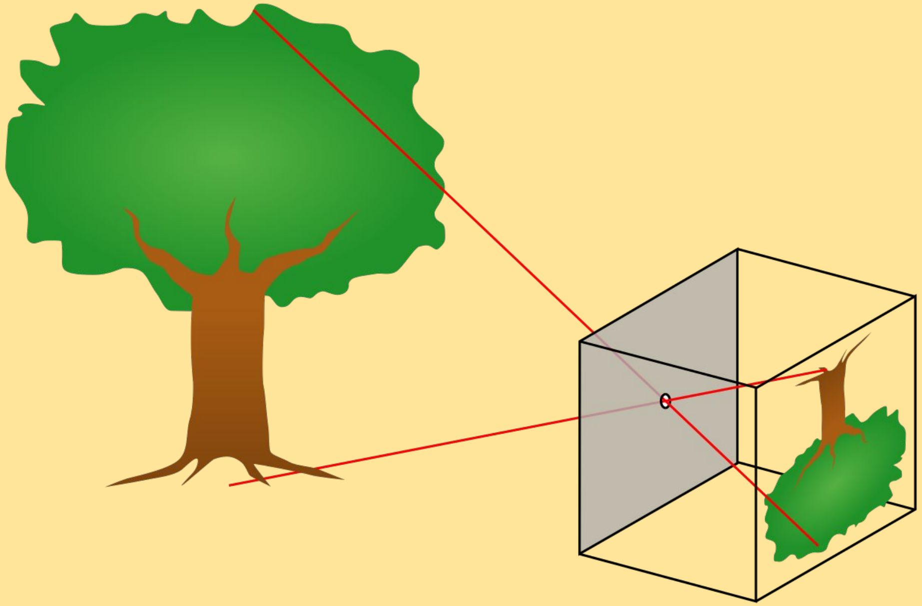


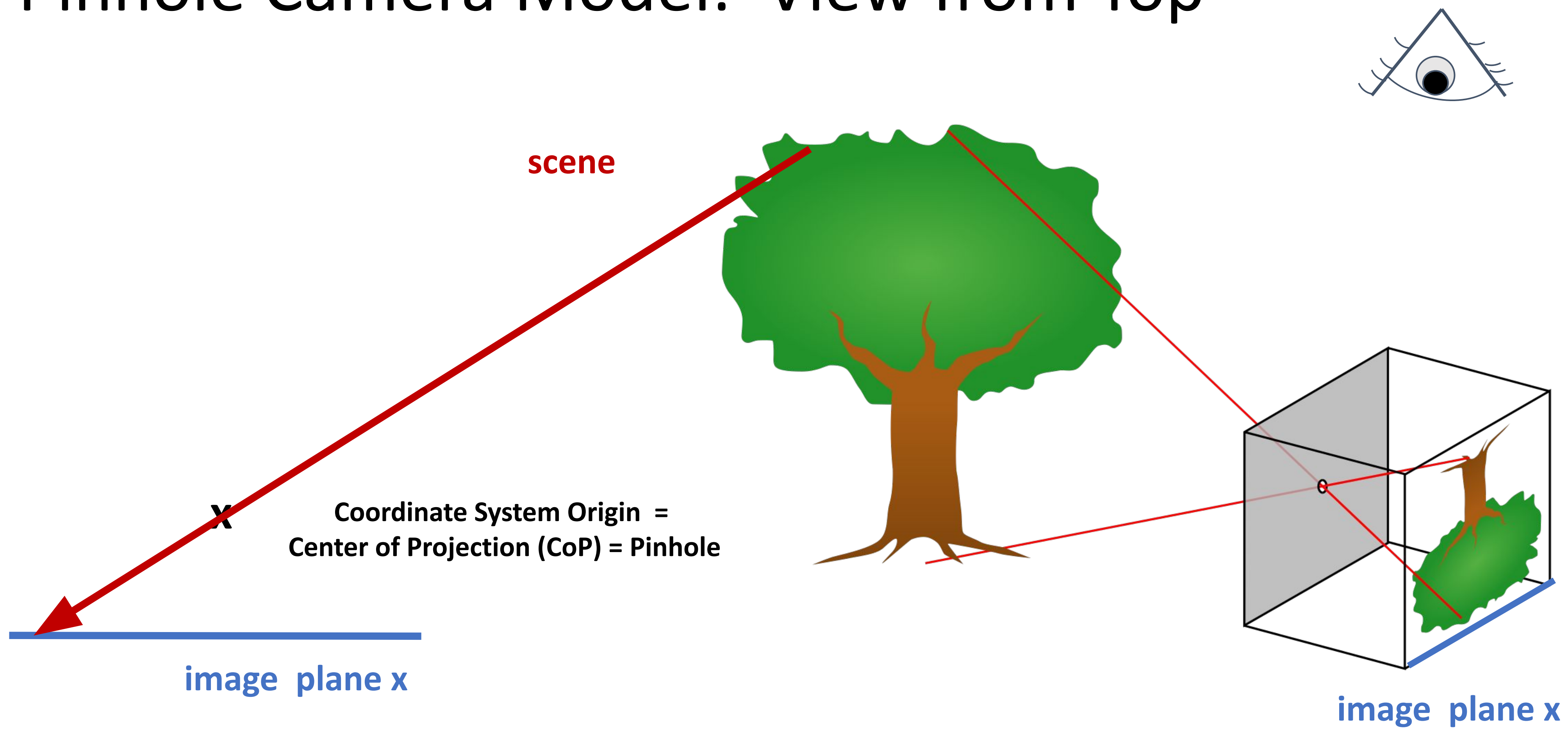
Image credit: public domain, Wikipedia

Poll: So then, why do we care about pinhole cameras in CS 585?

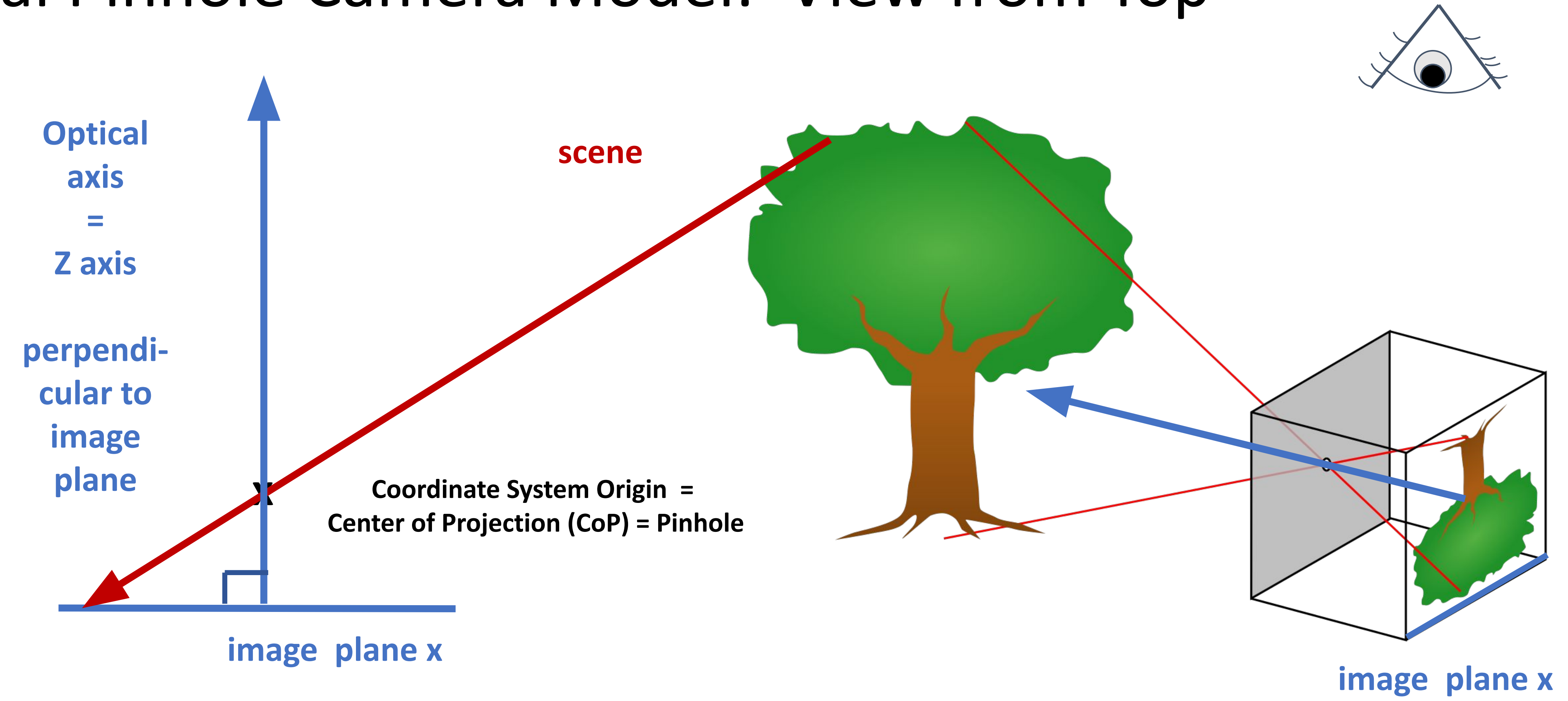
1. Historic reason: We need to learn how computer vision started as a research field.
2. Mathematical reason: Real cameras have complicated lens systems, not pin holes. We can simplify the geometry of image formation mathematically by ignoring the lenses.
3. Computational reason: Modern cameras post-process the projected images as if they were collected by a pinhole camera.



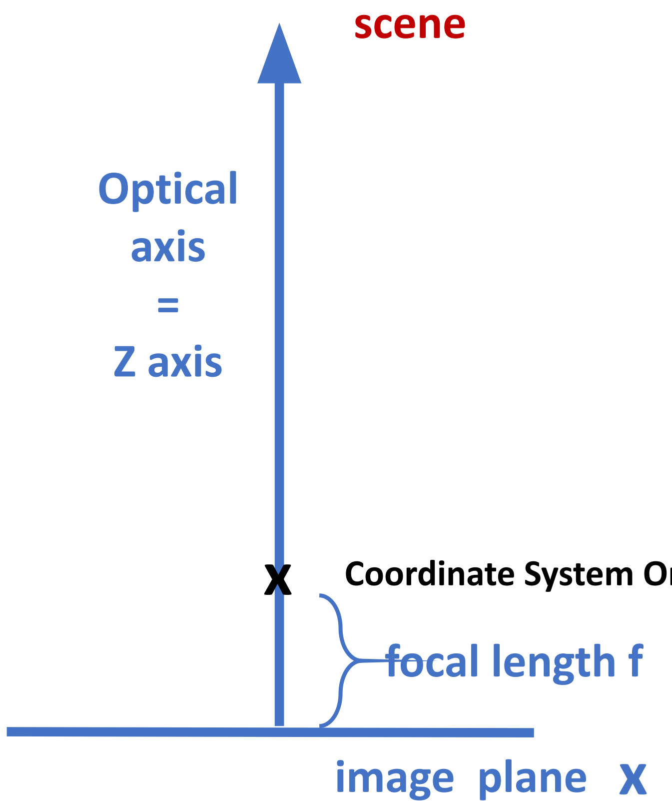
Ideal Pinhole Camera Model: View from Top



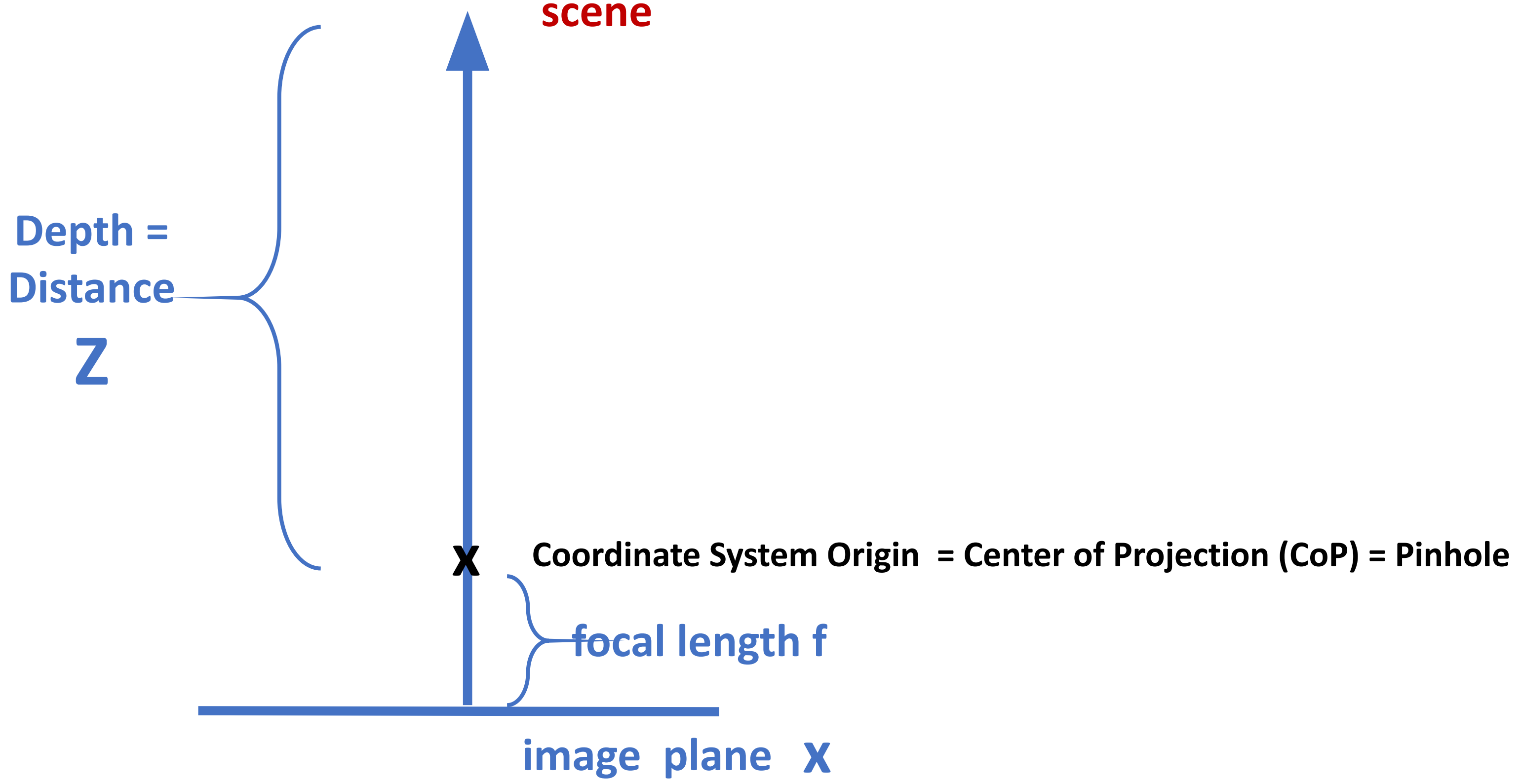
Ideal Pinhole Camera Model: View from Top



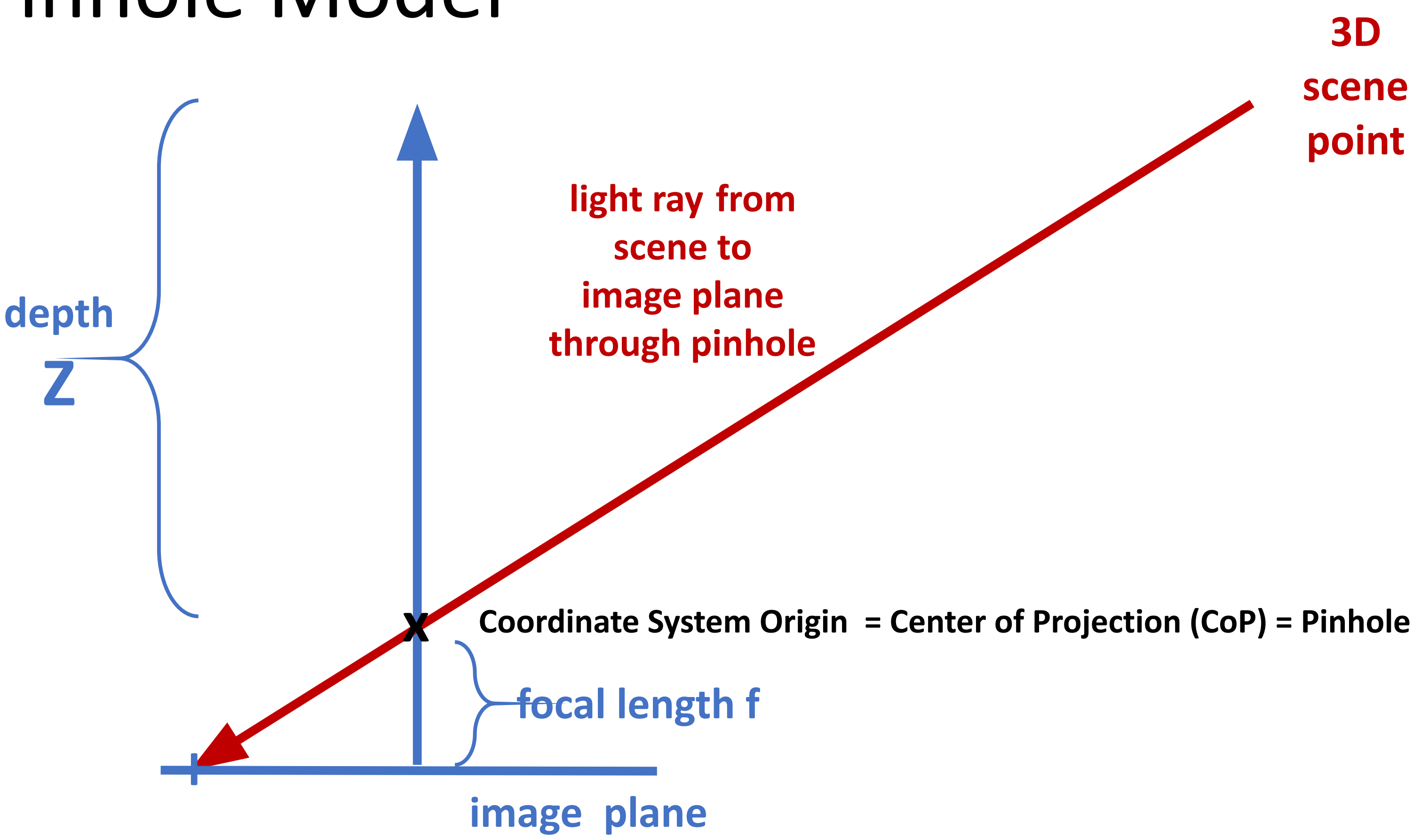
Ideal Pinhole Camera Model: View from Top



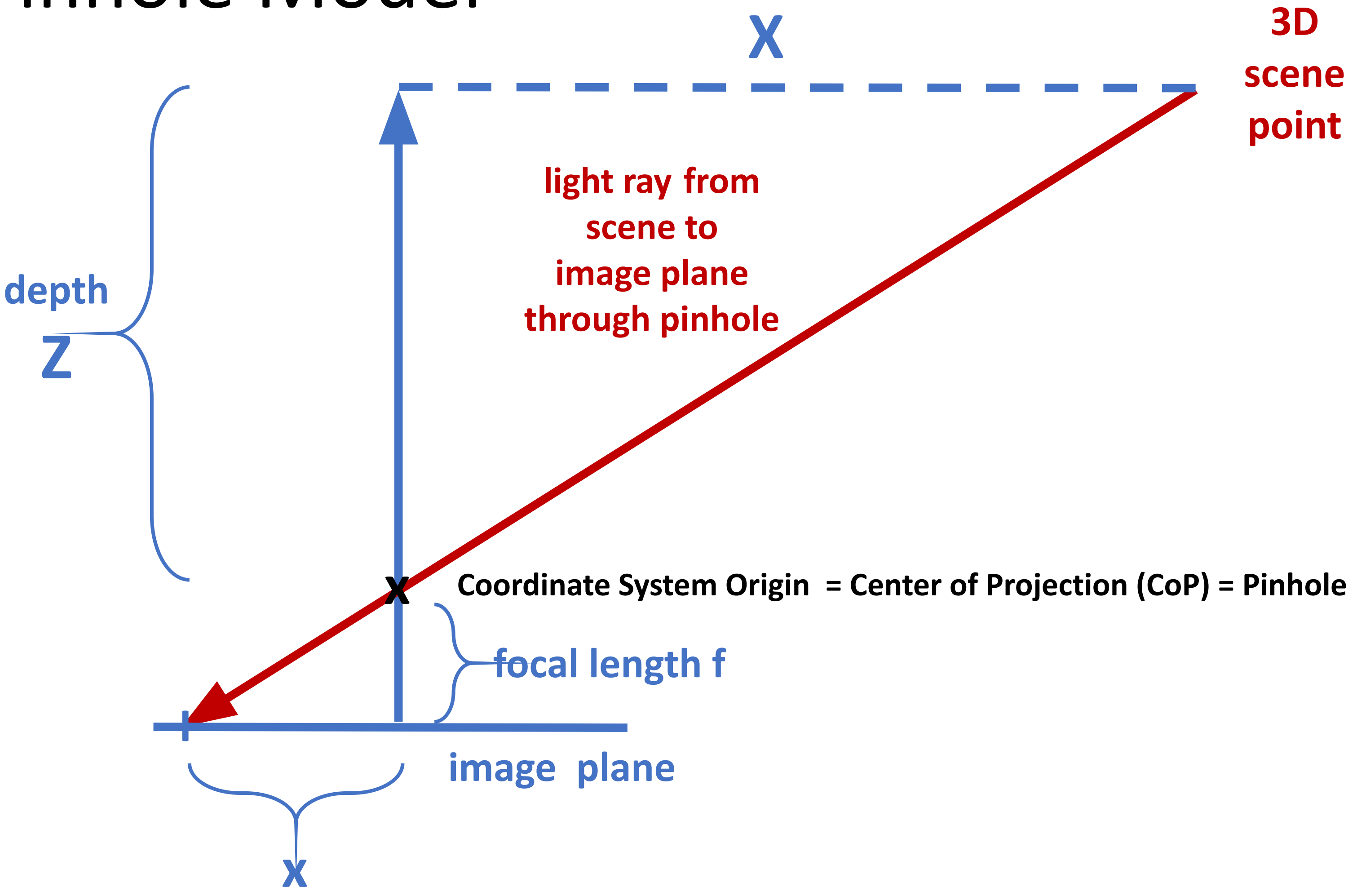
Ideal Pinhole Camera Model: View from Top



Pinhole Model

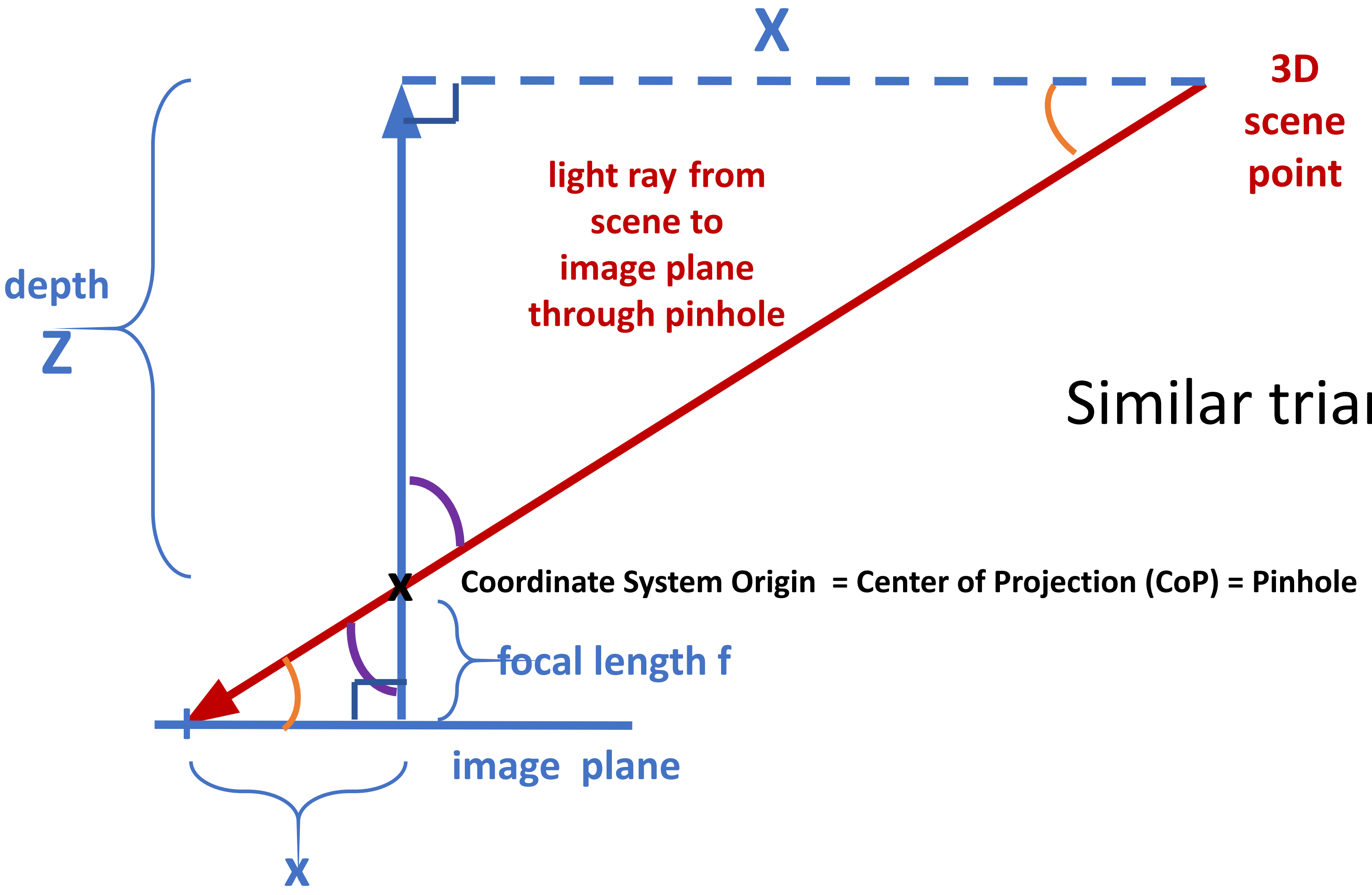


Pinhole Model



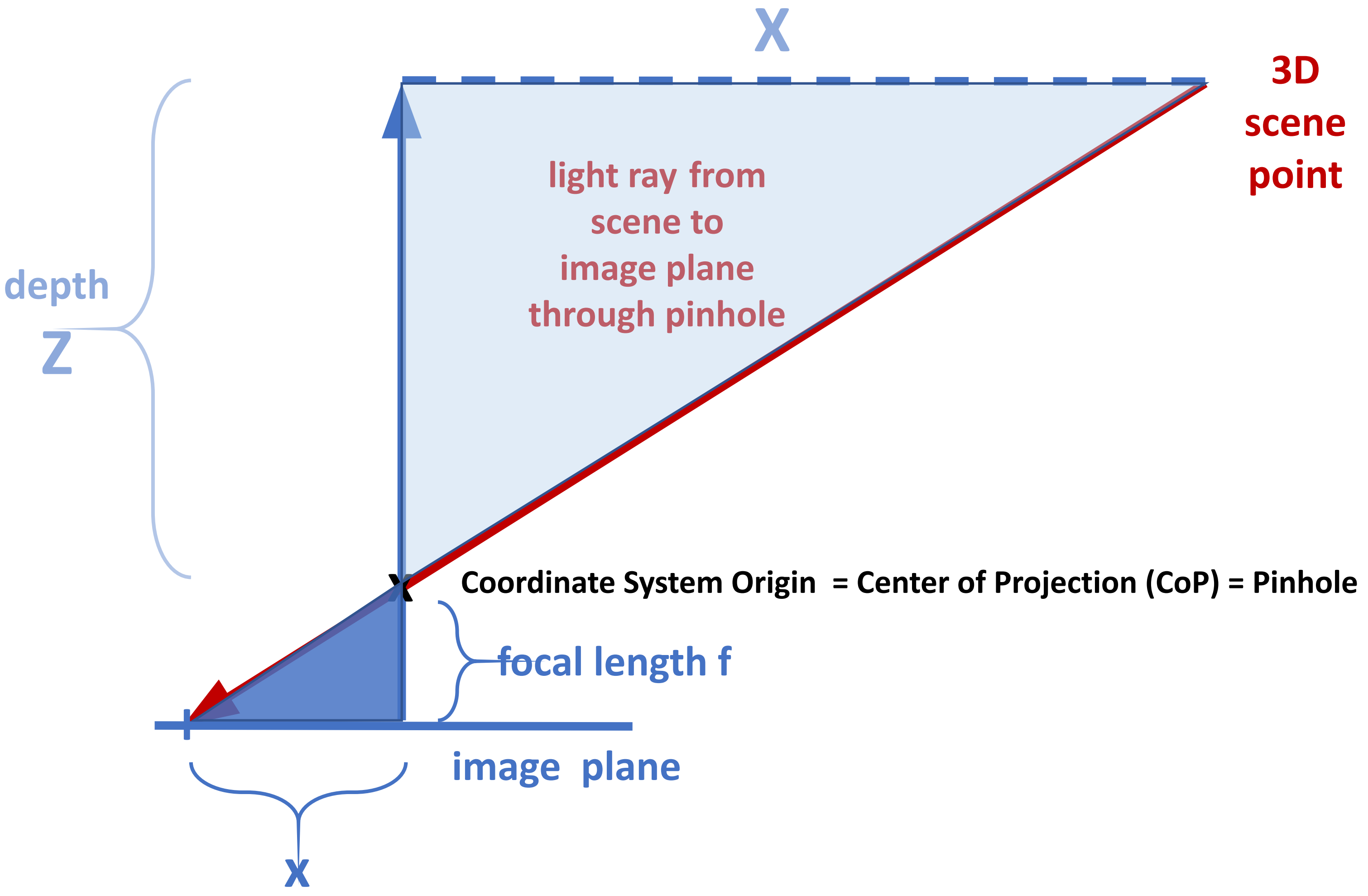
We can relate the scene and image plane coordinates X and x using a perspective projection equation.

Derivation of the Perspective Projection Equation

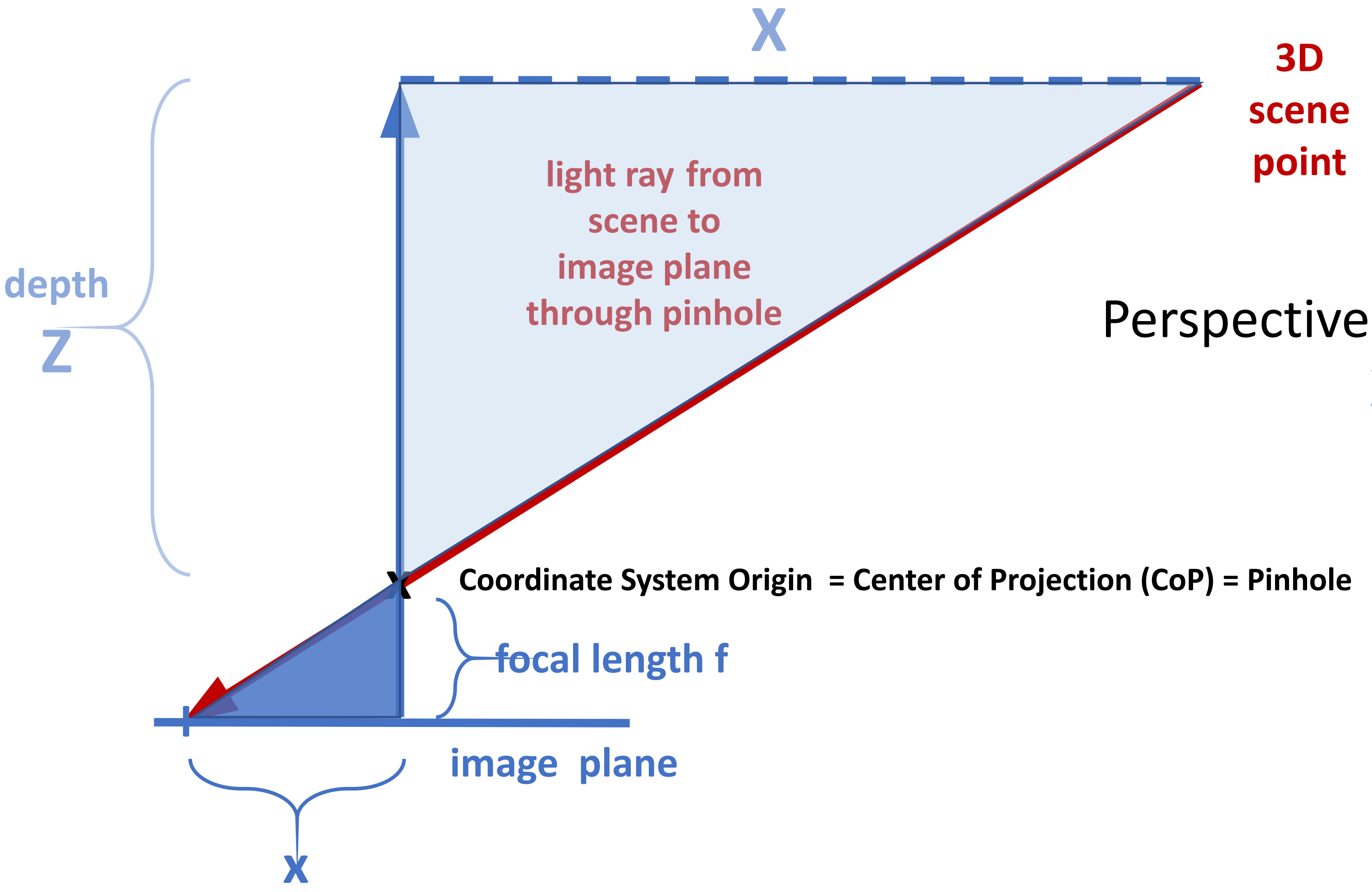


Similar triangles have same angles

Derivation of the Perspective Projection Equation

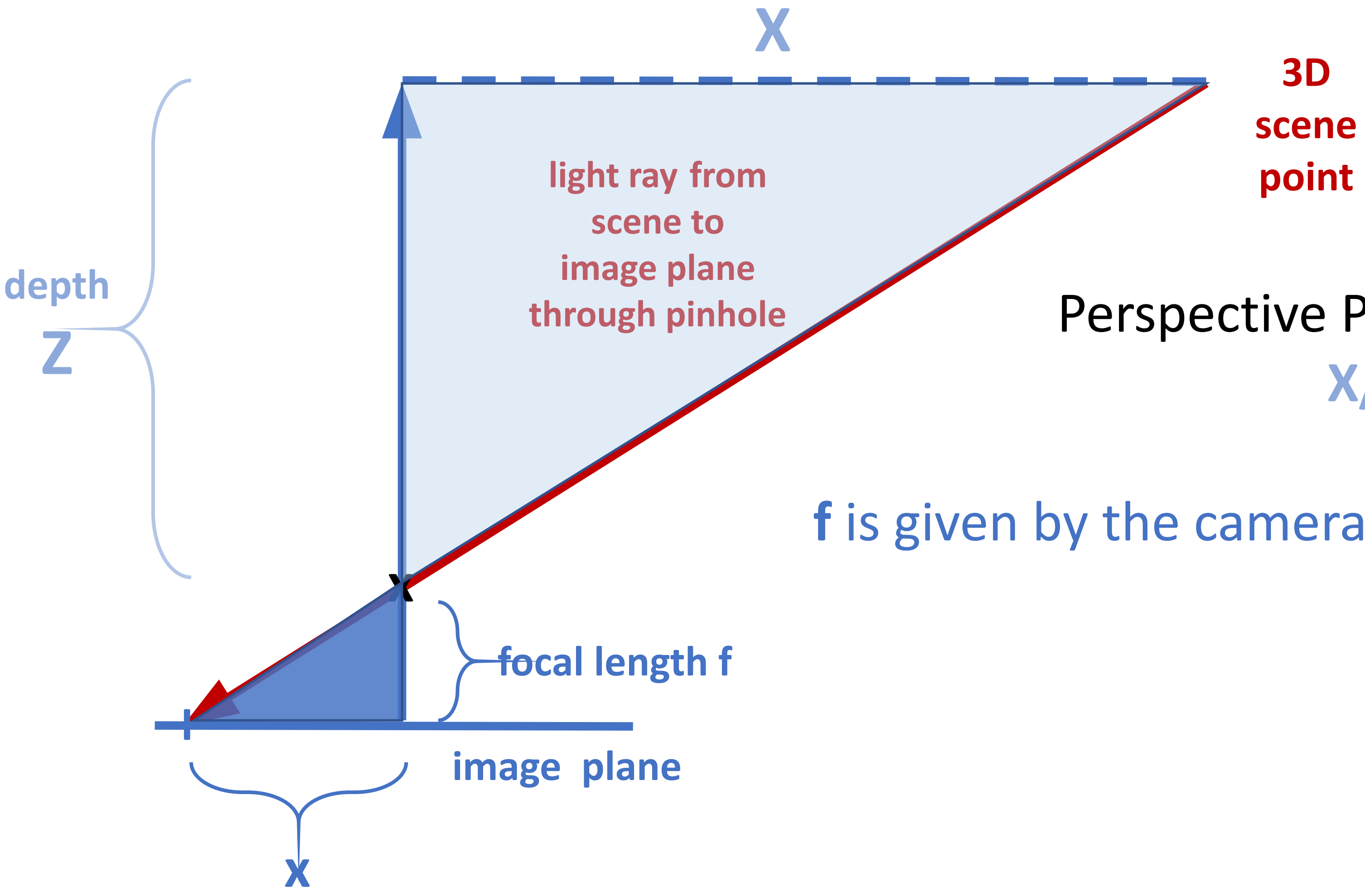


Derivation of the Perspective Projection Equation



Perspective Projection Equation:

$X/Z = x/f$

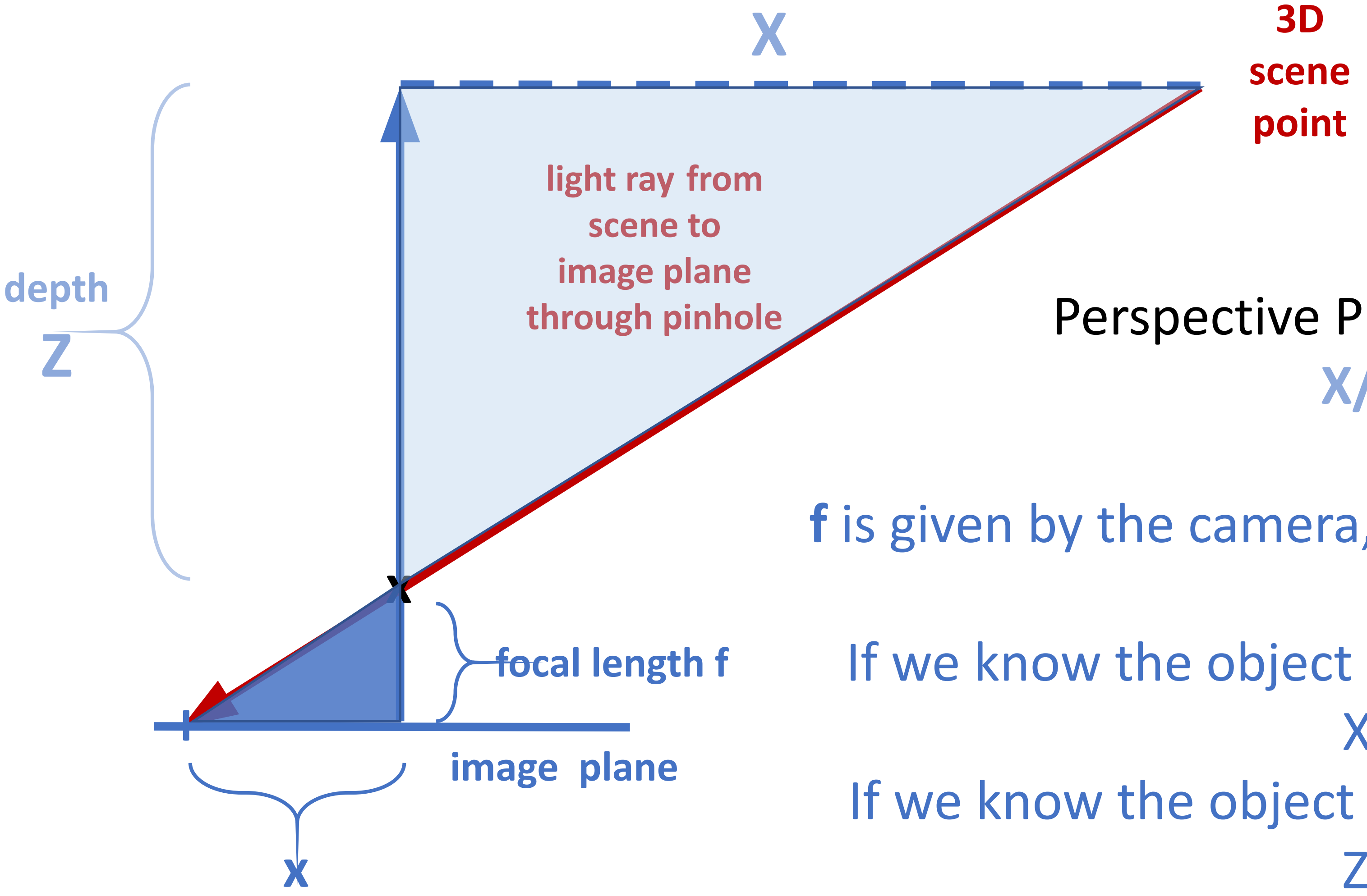


Perspective Projection Equation:

$X/Z = x/f$

f is given by the camera, **x** is our image measurement

Usefulness of the Perspective Projection Equation



Perspective Projection Equation:

$X/Z = x/f$

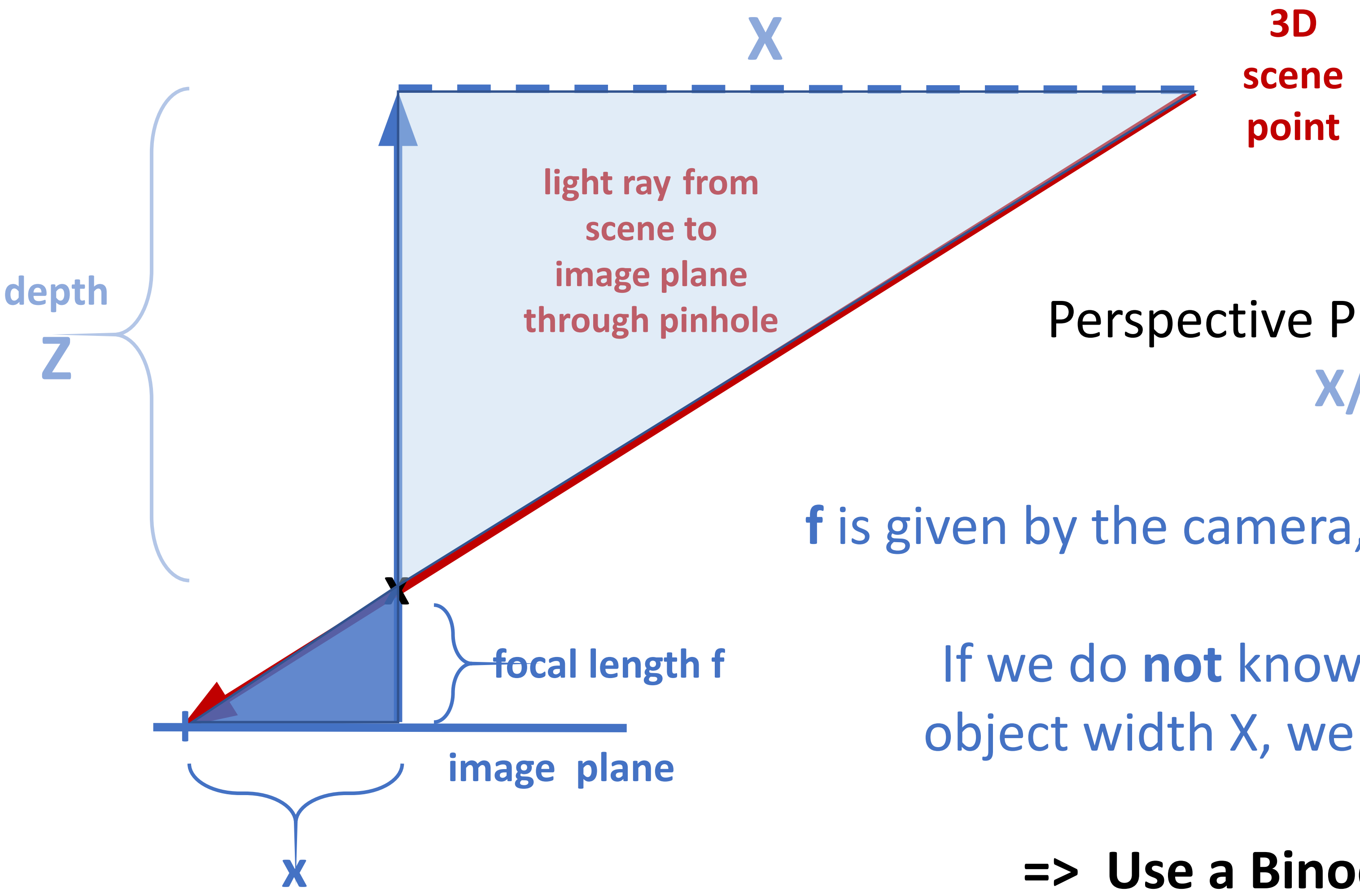
f is given by the camera, **x** is our image measurement

If we know the object depth Z, we can compute X:

$X = xZ/f$

If we know the object width X, we can compute Z:

$Z = Xf/x$



Perspective Projection Equation:

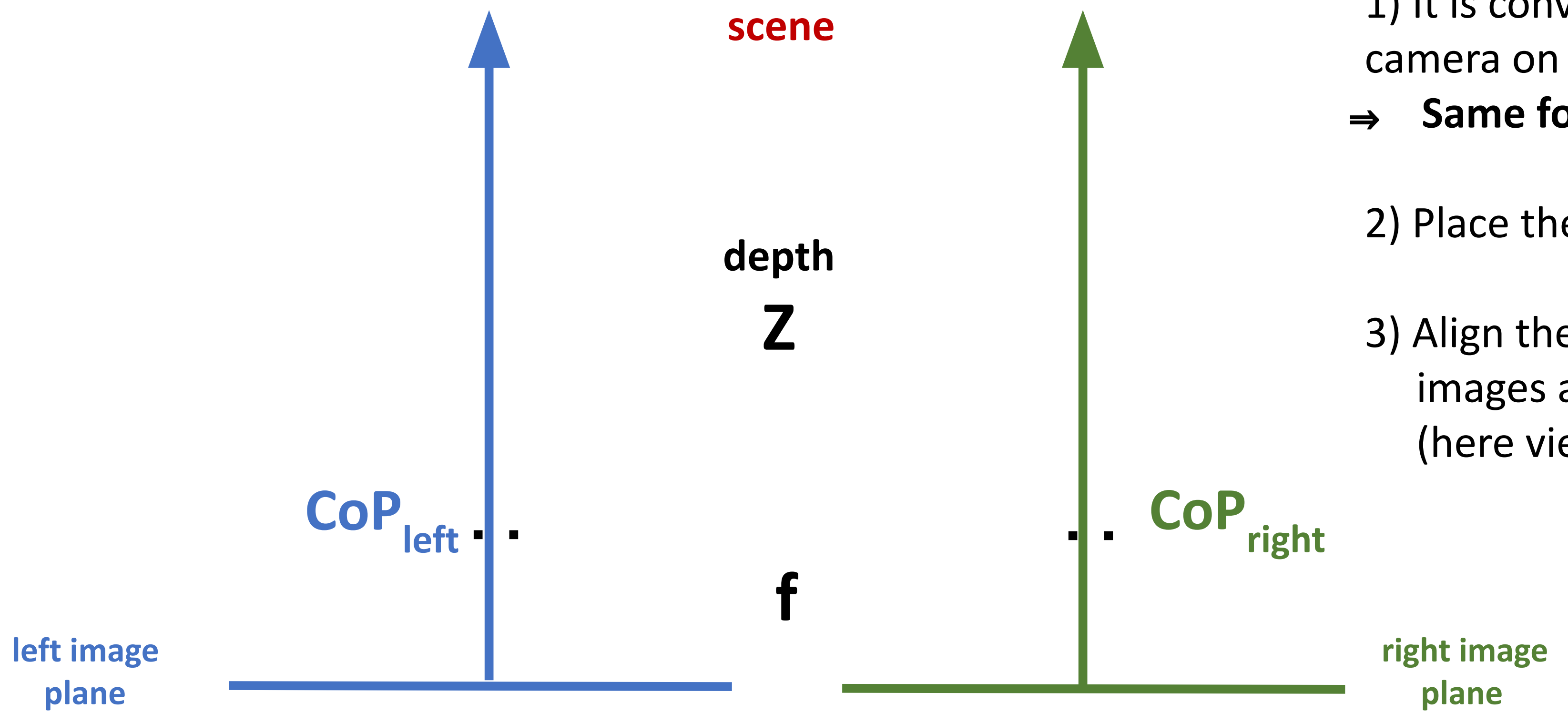
$X/Z = x/f$

f is given by the camera, **x** is our image measurement

If we do **not** know the object depth Z **and** object width X, we need another equation!

=> Use a Binocular Stereo System

Binocular Stereo



Special Considerations:

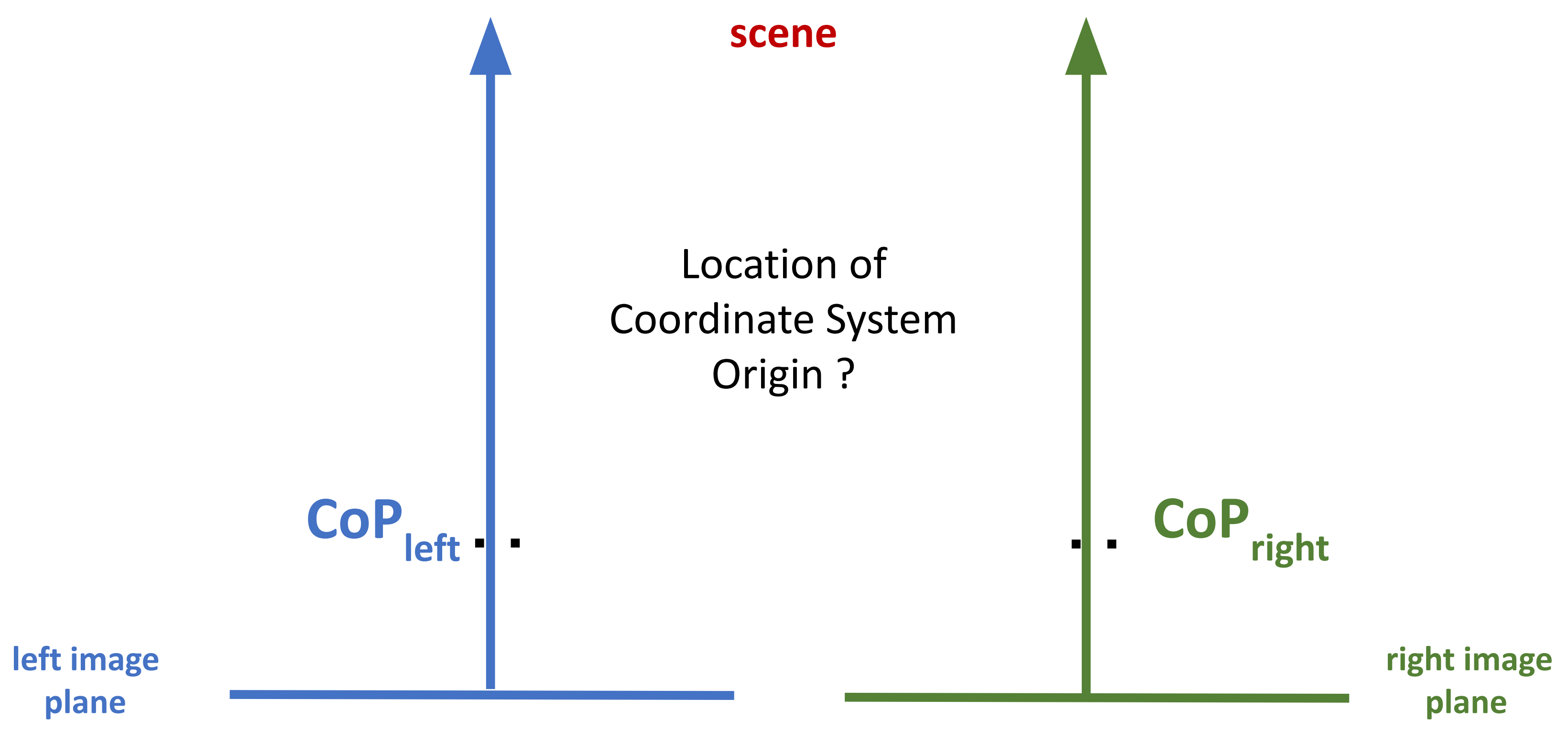
1) It is convenient to use the same type of camera on the left and the right

⇒ **Same focal length f**

2) Place the optical axes **parallel** to each other

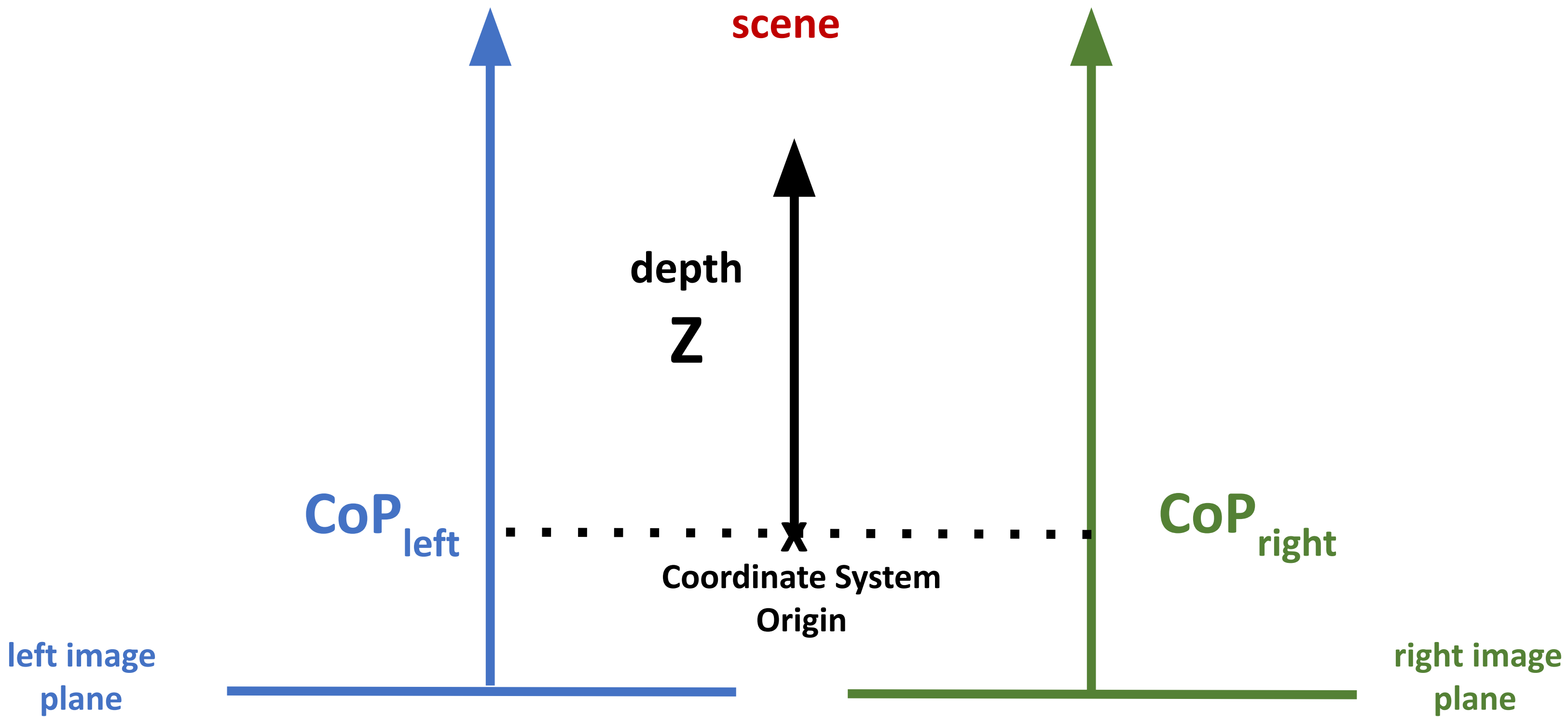
3) Align the cameras so that the left and right images are in the **same plane** (here viewed in cross section)

Binocular Stereo



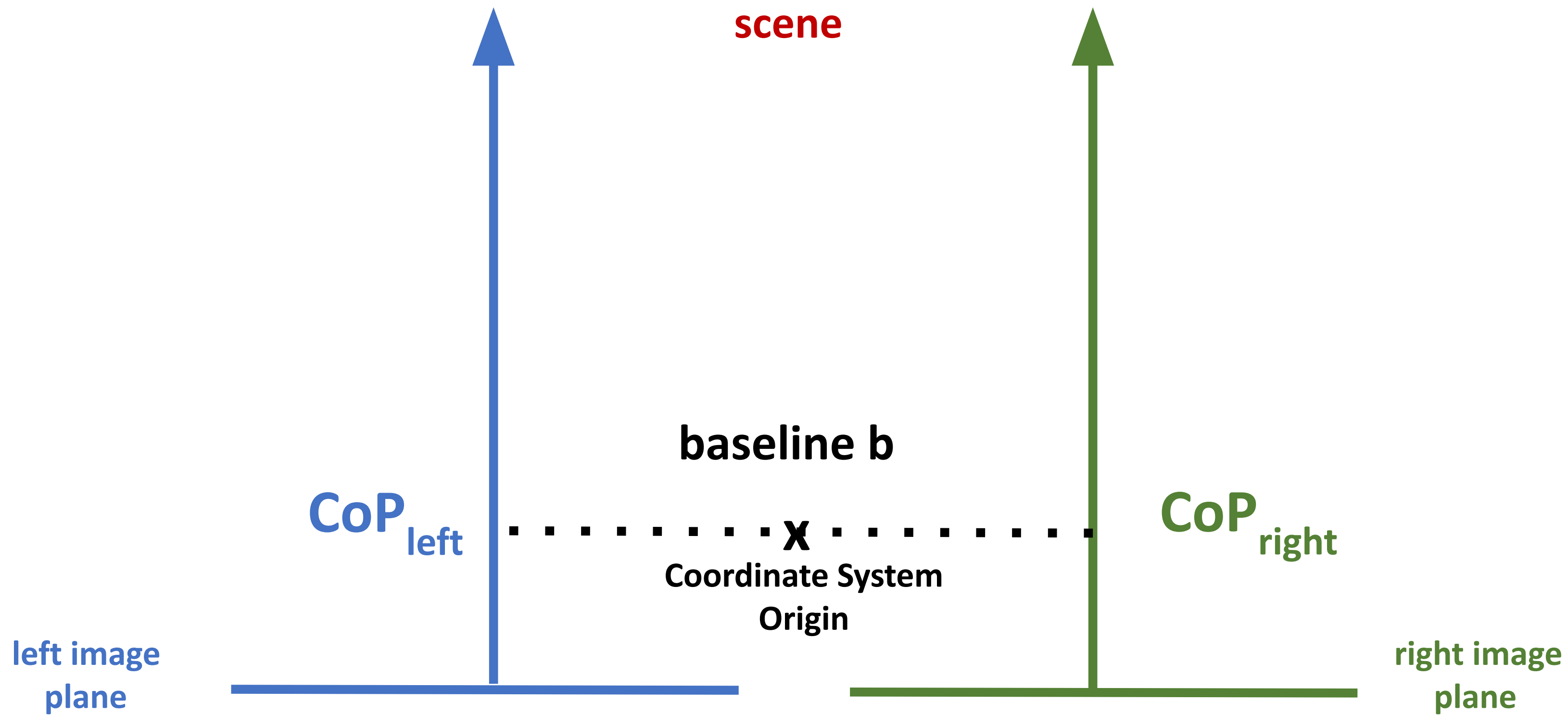
In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole

Binocular Stereo



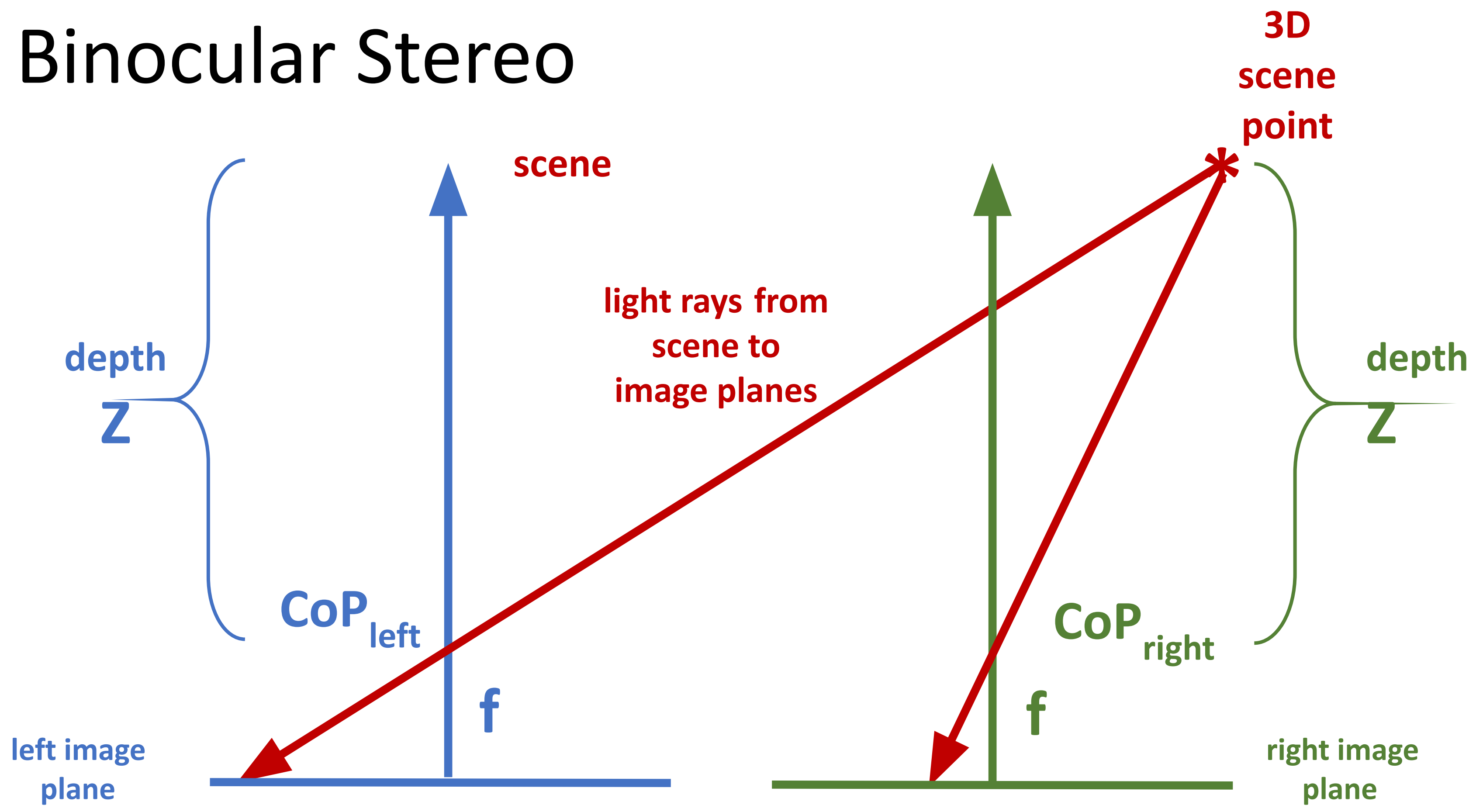
In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole
In a Binocular System: Coordinate System Origin in the middle between CoPs

Binocular Stereo

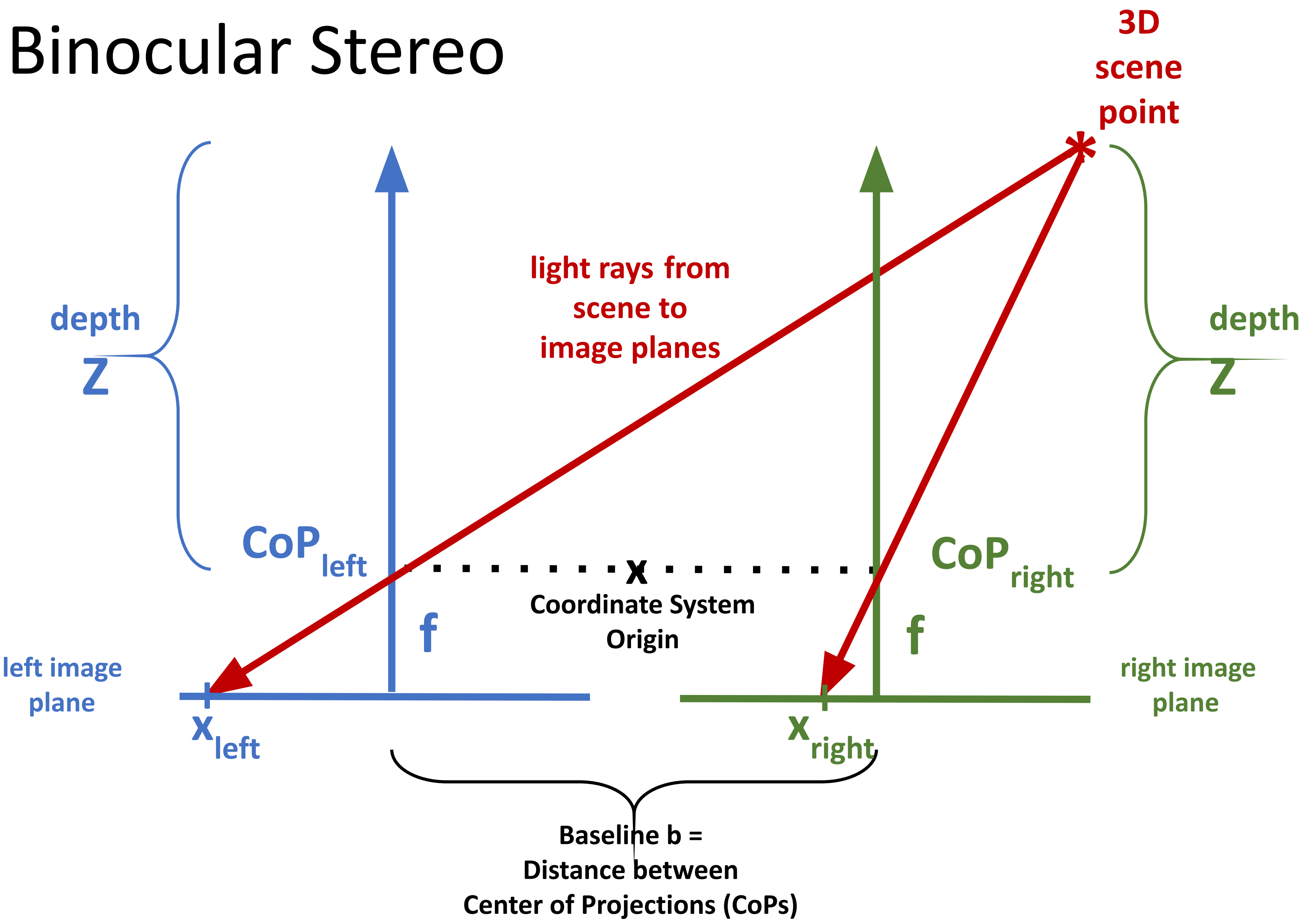


In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole
In a Binocular System: Coordinate System Origin in the middle between CoPs

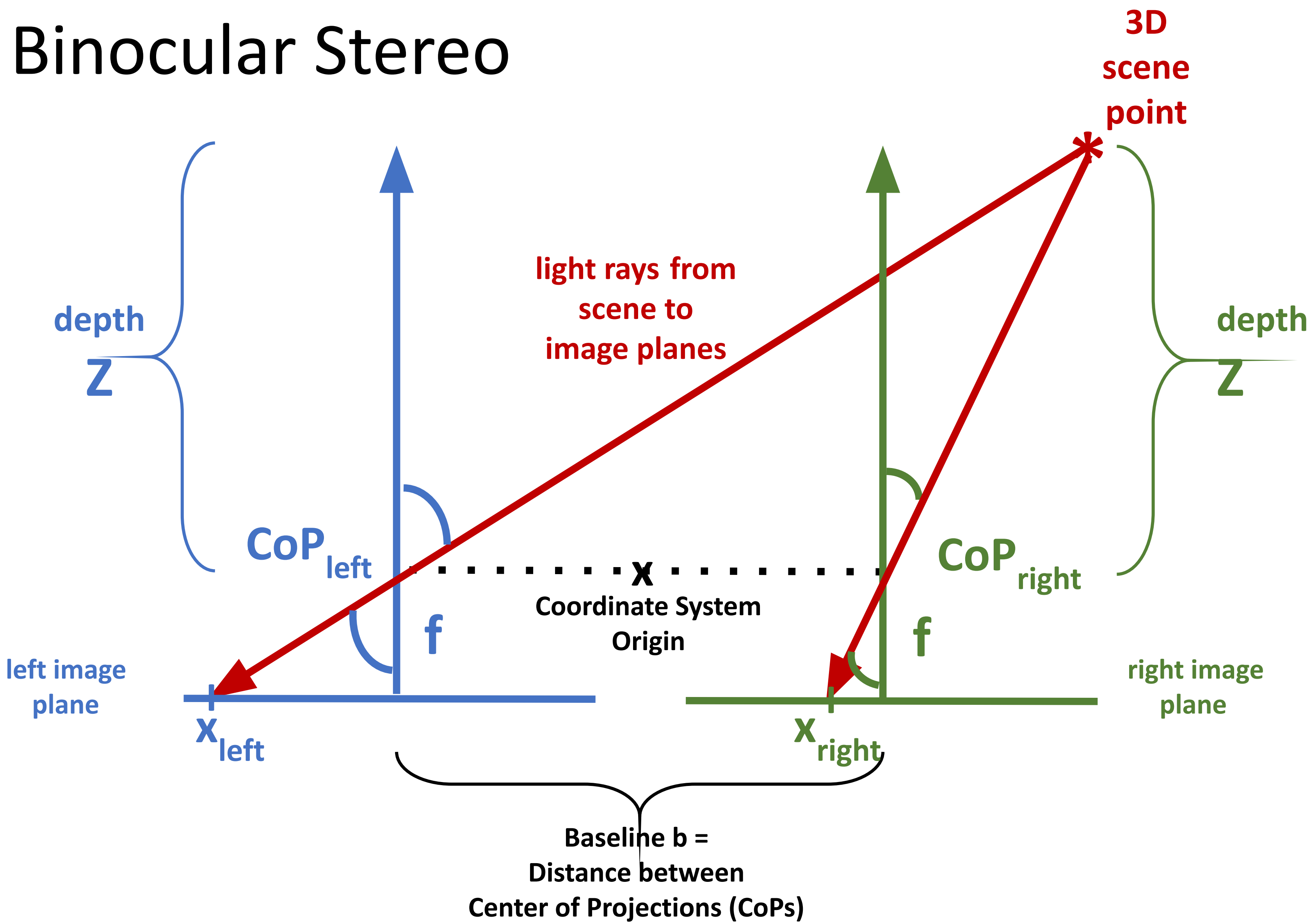
Binocular Stereo



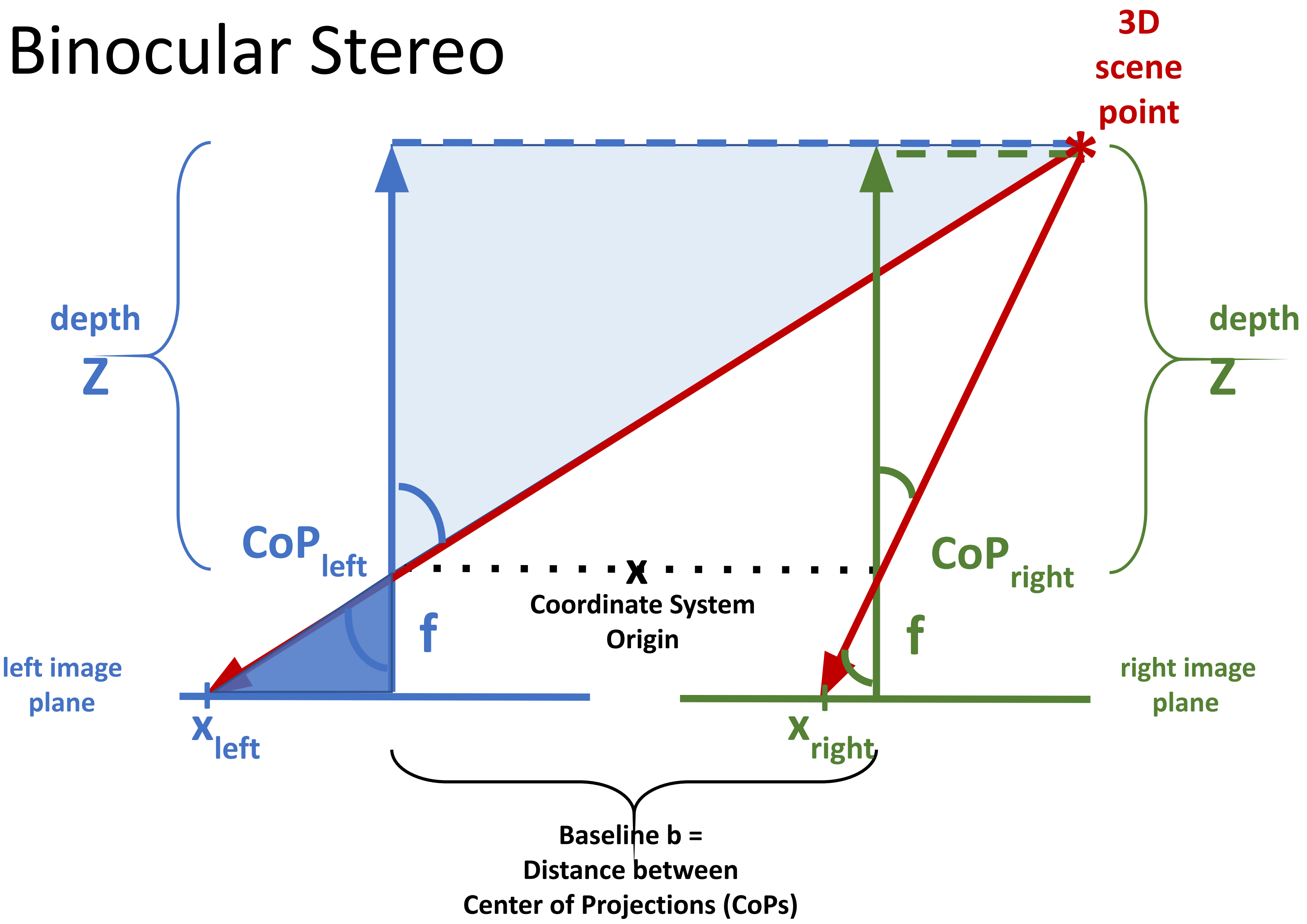
Binocular Stereo



Binocular Stereo



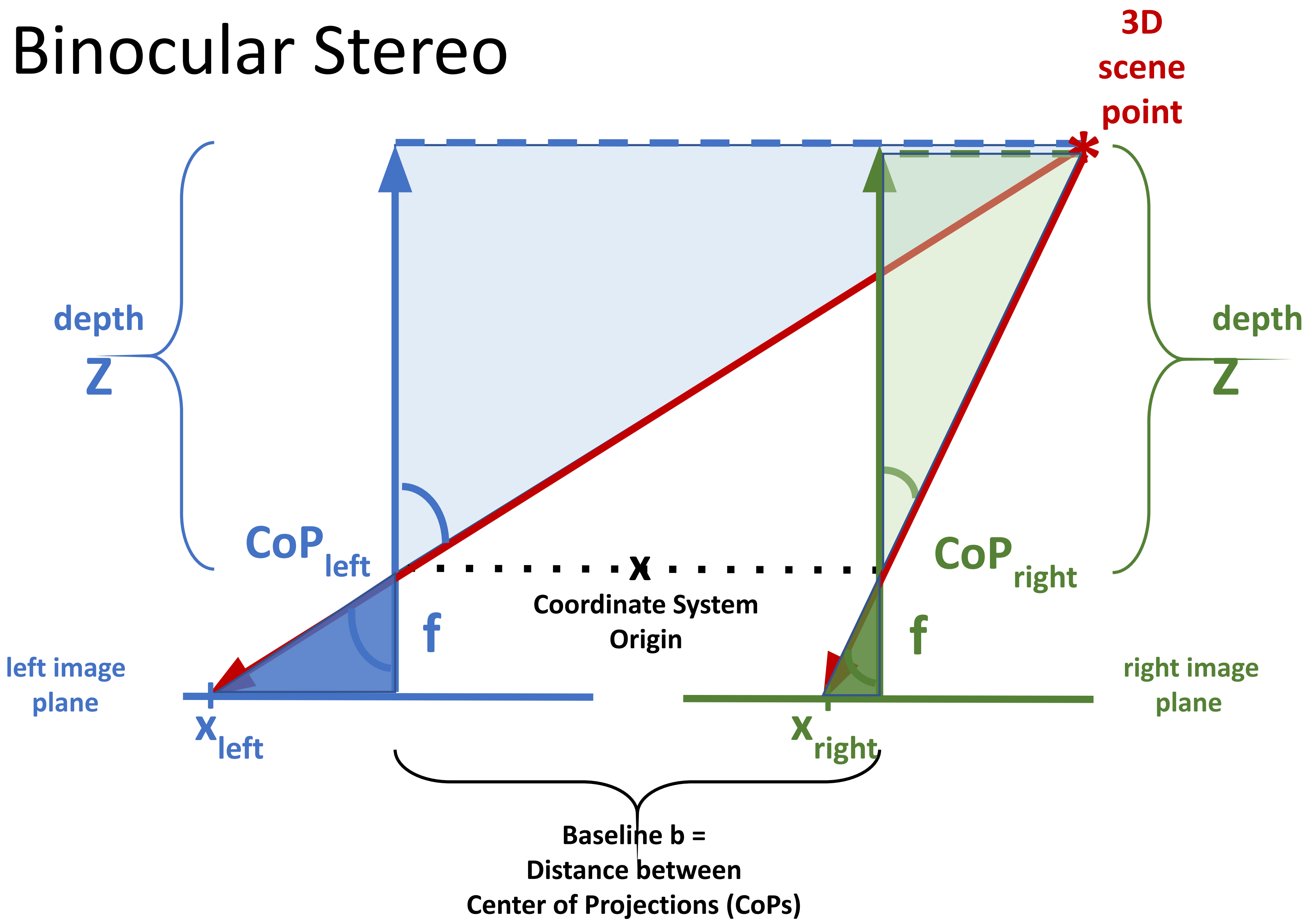
Binocular Stereo



Projection Equation:

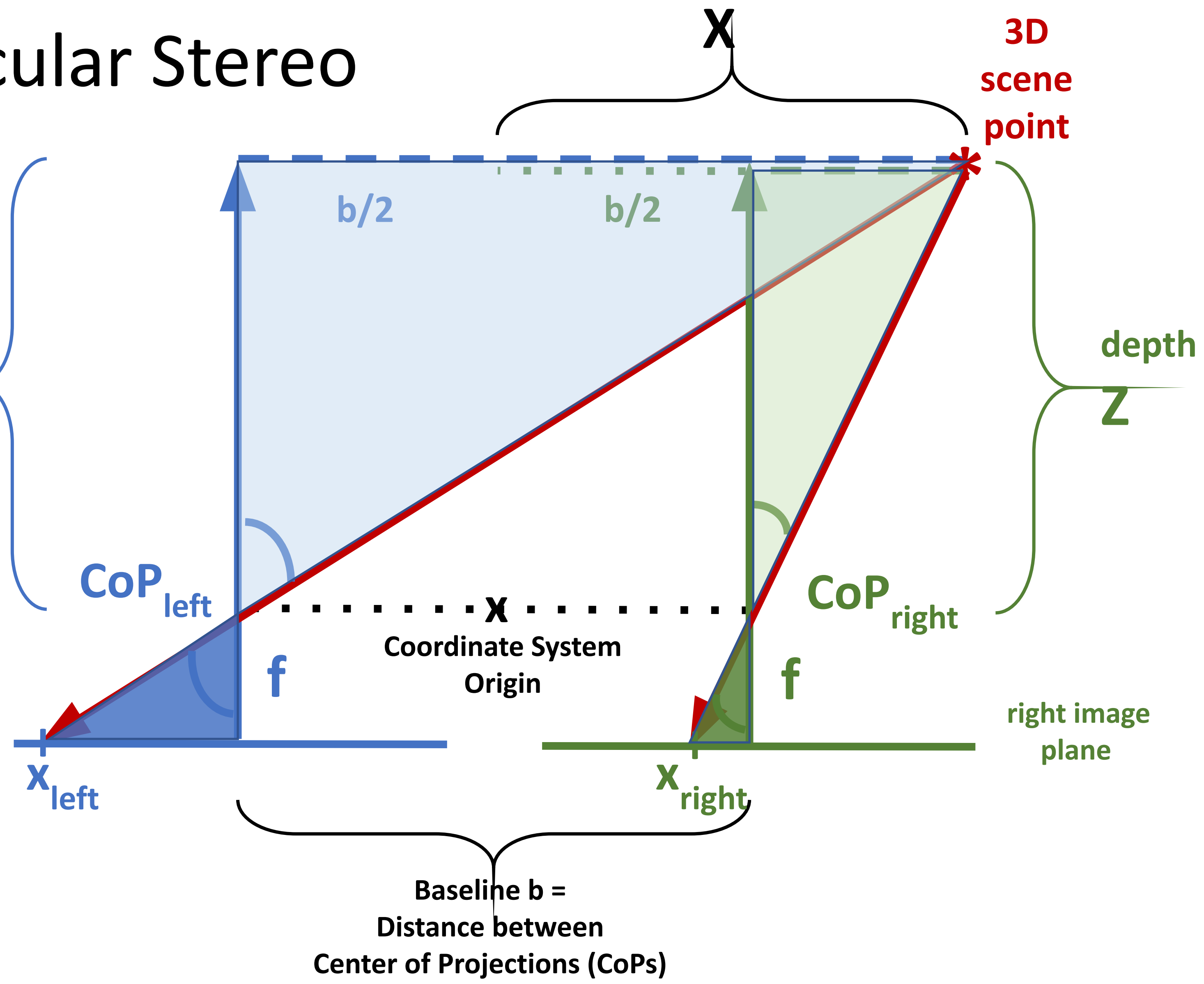
x_{left}/f = (something)/Z

Binocular Stereo

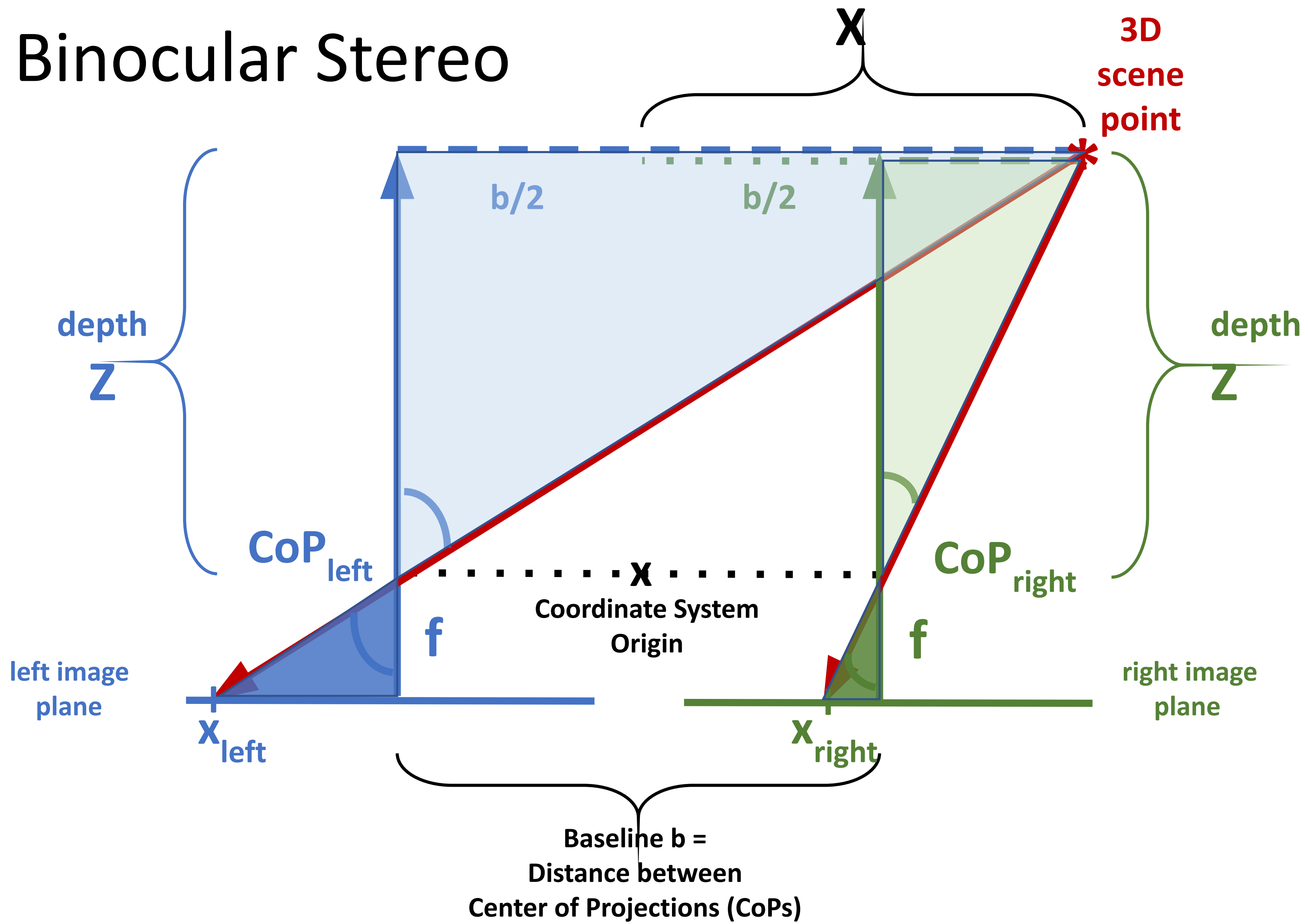


x_{right}/f = (something else) $/Z$

Binocular Stereo



Binocular Stereo

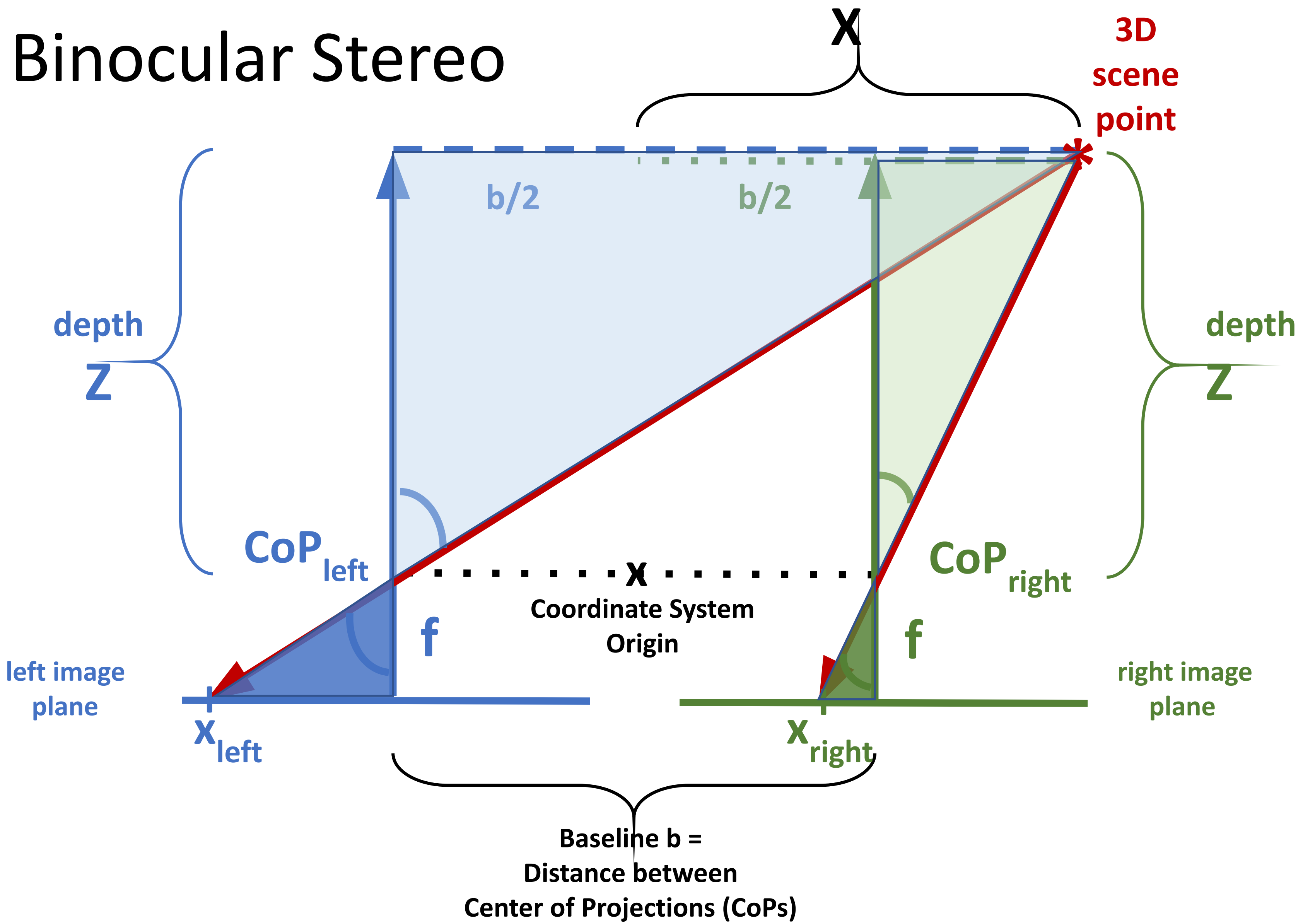


Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

Binocular Stereo

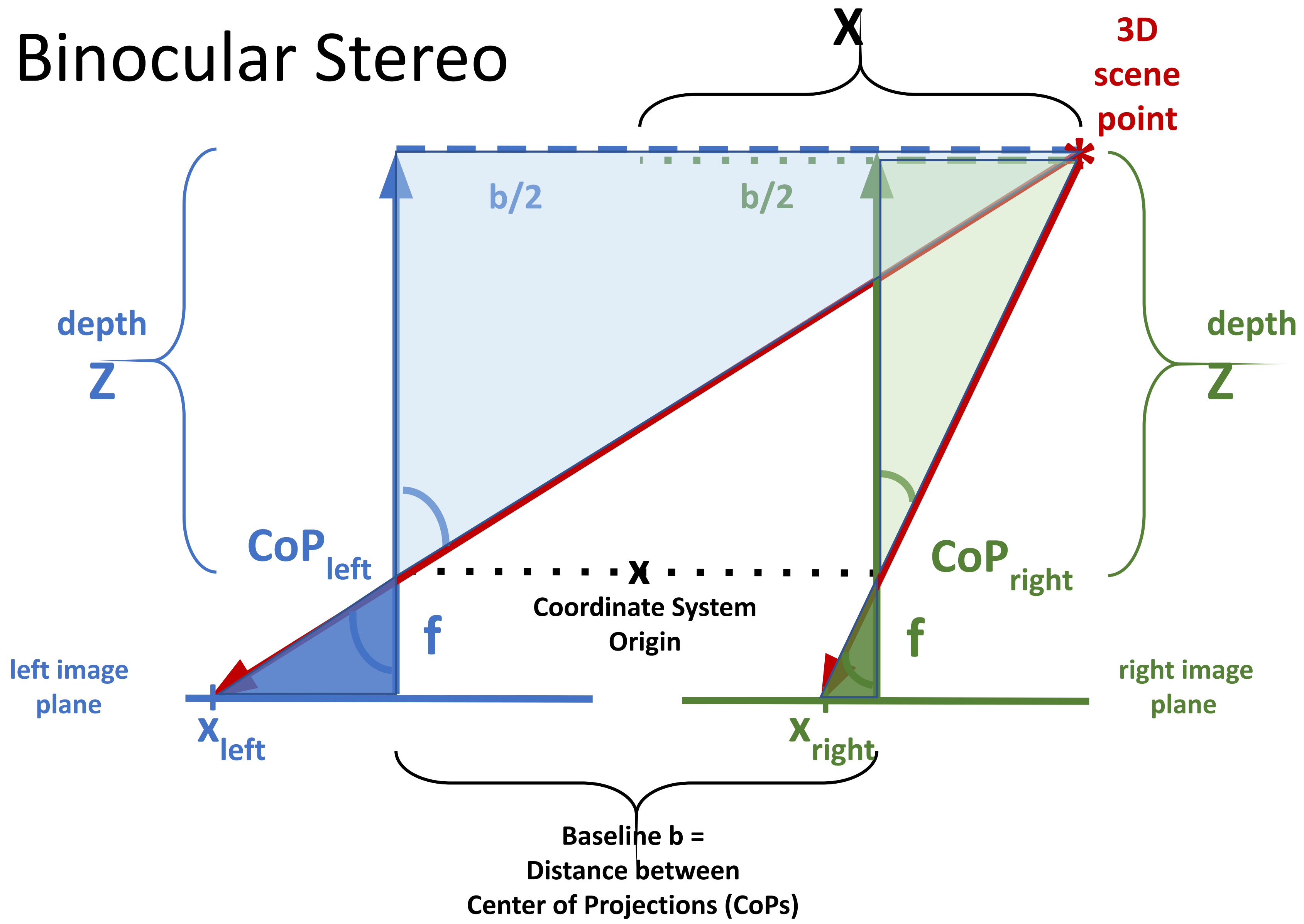


Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

Binocular Stereo



Projection Equations:

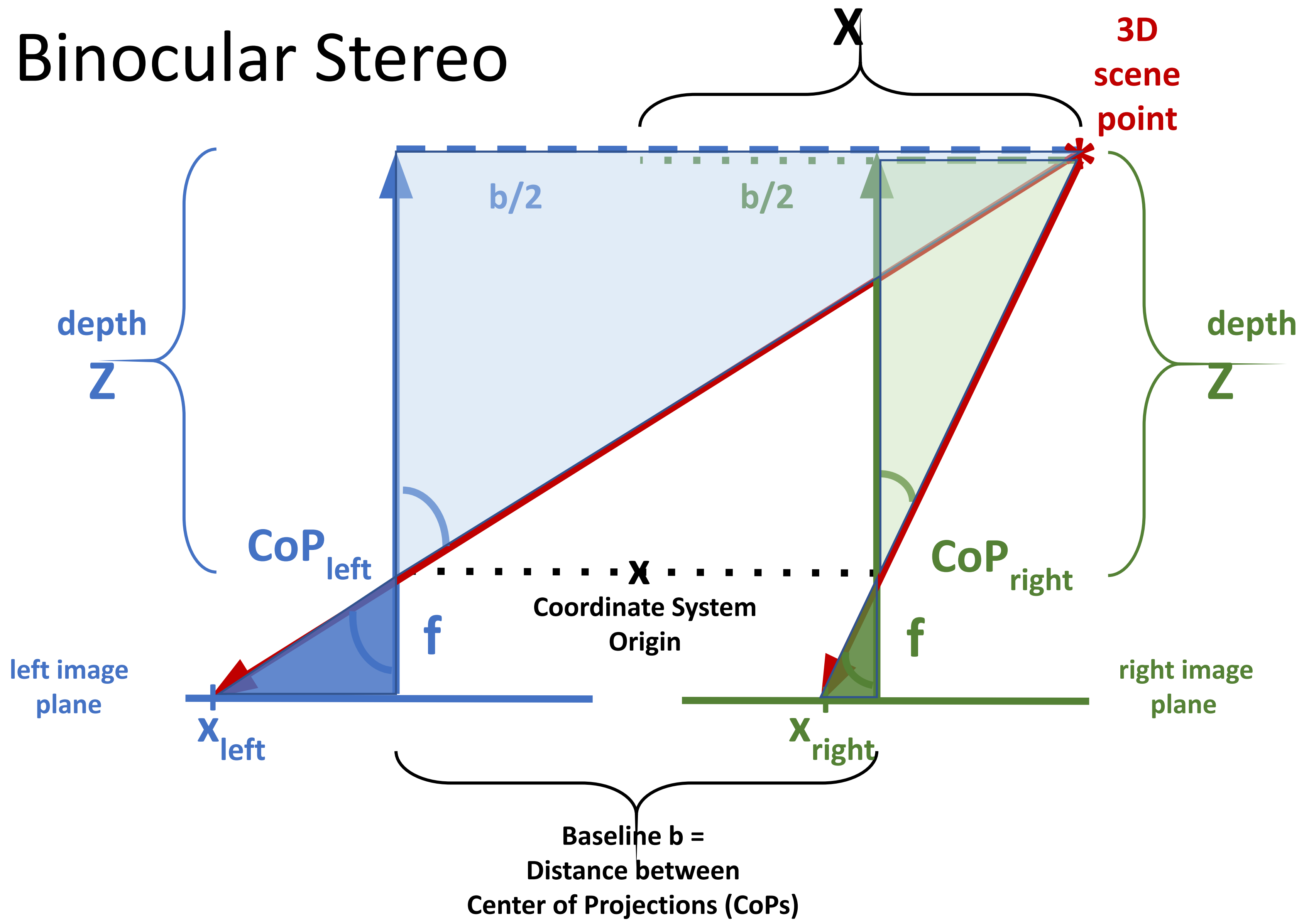
$$x_{left}/f = (X+b/2)/Z$$

$$x_{right}/f = (X-b/2)/Z$$

Subtract the 2nd equation from the 1st equation:

$$(x_{left} - x_{right})/f = (X+b/2-X+b/2))/Z$$

Binocular Stereo



Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

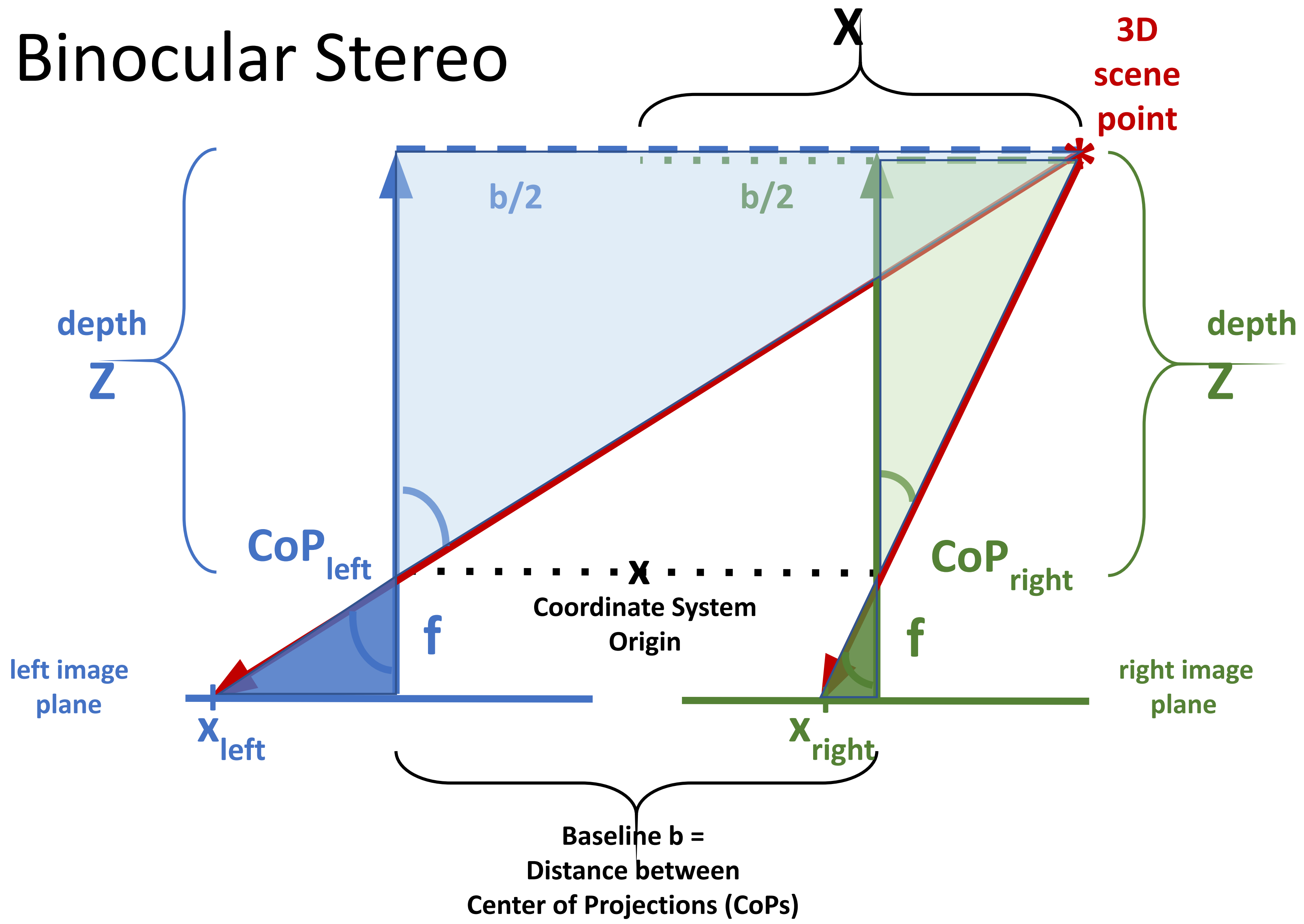
Subtract the 2nd equation from the 1st equation:

$(x_{left} - x_{right})/f = (X+b/2-X+b/2))/Z$

which results in:

$(x_{left} - x_{right})/f = b/Z$

Binocular Stereo



Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

Subtract the 2[nd] equation from the 1[st] equation:

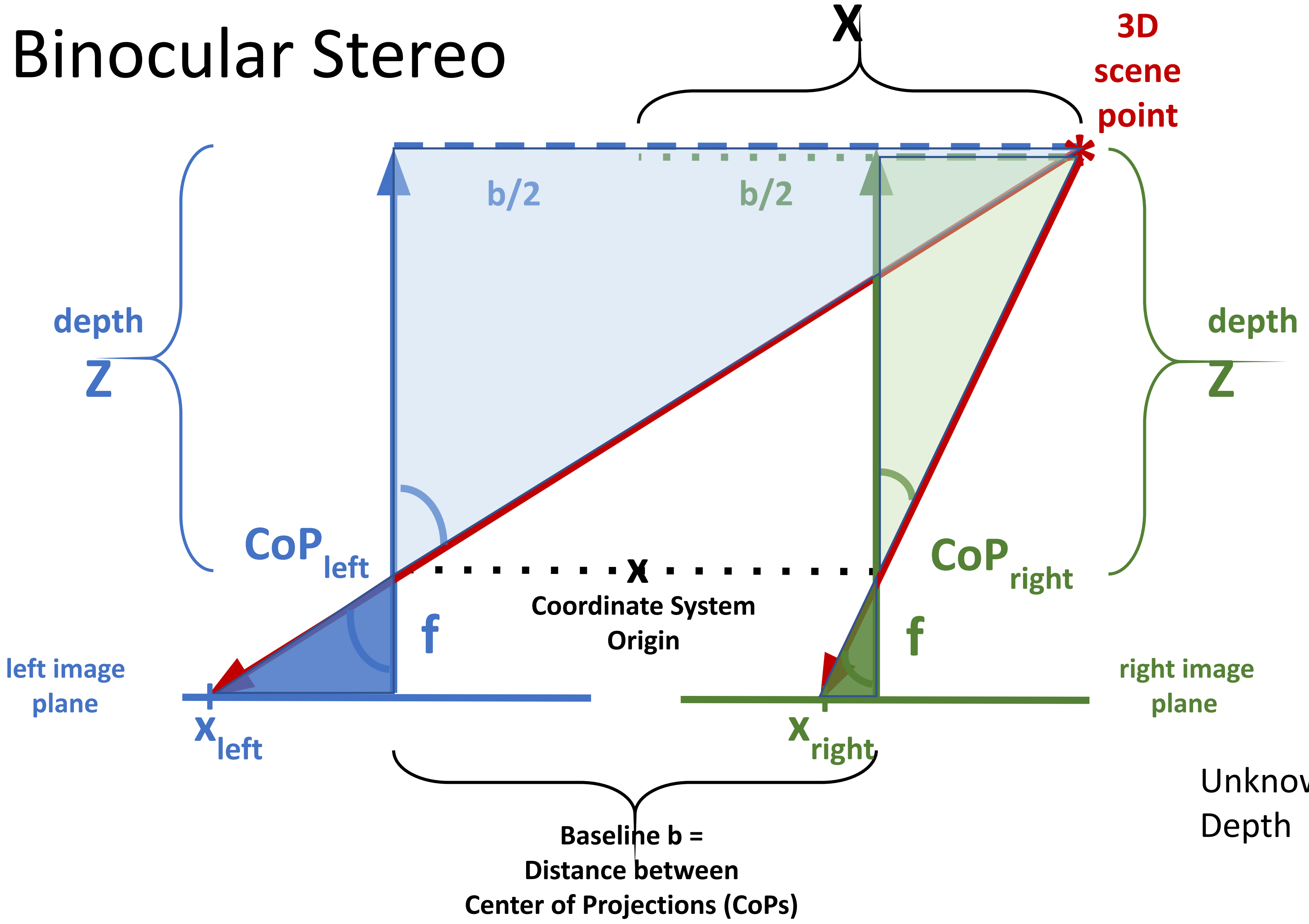
$(x_{left} - x_{right})/f = (X+b/2-X+b/2))/Z$

which results in:

$(x_{left} - x_{right})/f = b/Z$ or:

$Z = bf/(x_{left} - x_{right})$

Binocular Stereo



Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

Subtract the 2nd equation from the 1st equation:

$(x_{left} - x_{right})/f = (X+b/2-X+b/2))/Z$

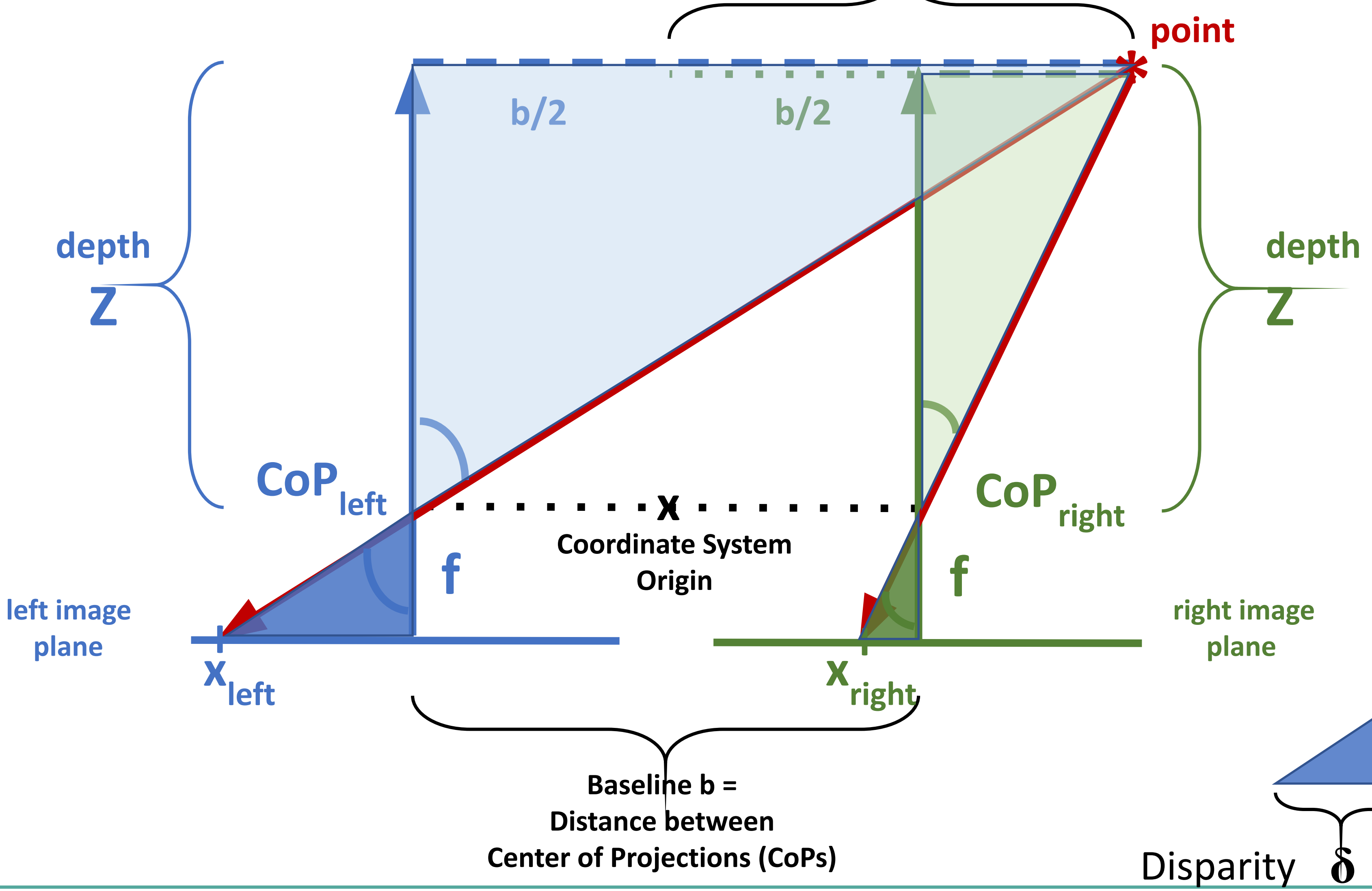
which results in:

$(x_{left} - x_{right})/f = b/Z$ or:

$Z = bf/(x_{left} - x_{right})$

Unknown Depth → Z; b, f: known; $x_{left} - x_{right}$: measured

Binocular Stereo



Projection Equations:

$x_{left}/f = (X+b/2)/Z$

$x_{right}/f = (X-b/2)/Z$

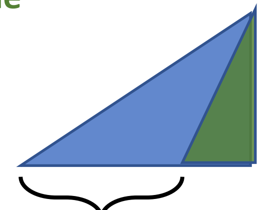
$(x_{left} - x_{right})/f = (X+b/2-X+b/2))/Z$

which results in:

$(x_{left} - x_{right})/f = b/Z$ or:

$Z = bf/\,(x_{left} - x_{right})$ or:

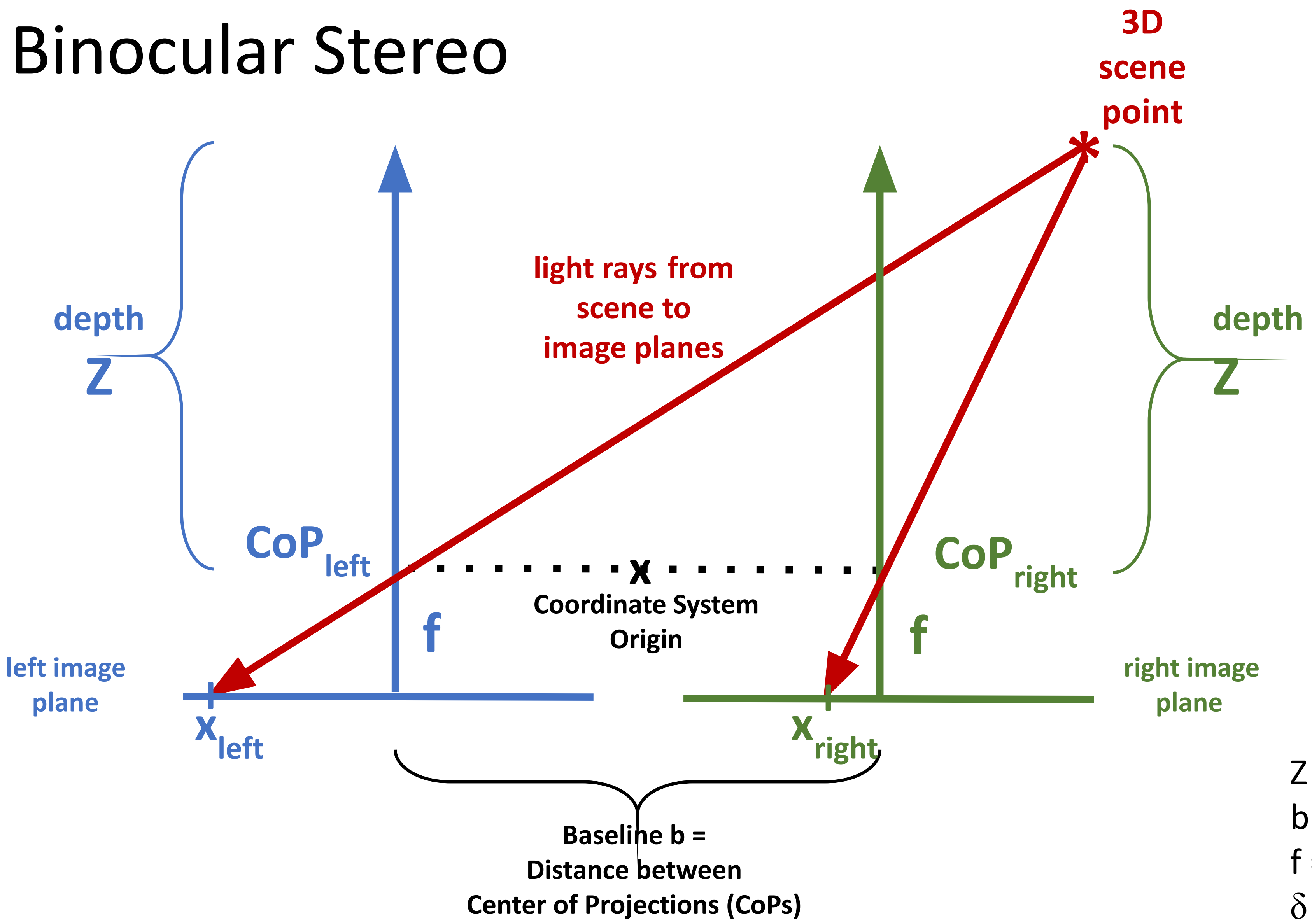
$\underbrace{x_{left} - x_{right}}_{\delta}$



Disparity δ

$Z = bf/\,\delta$

Binocular Stereo



Key Equation:

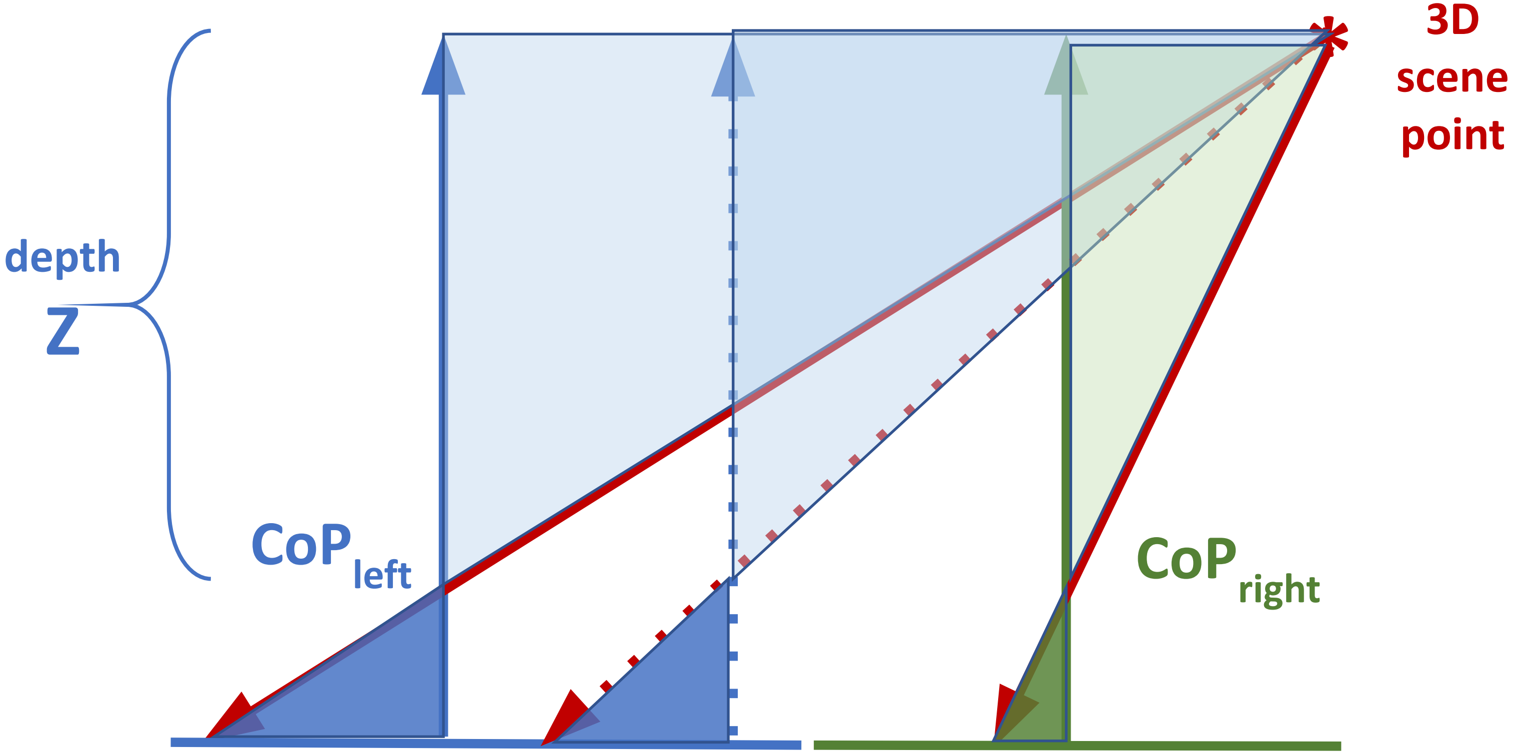
$$Z = \frac{b\,f}{x_{left} - x_{right}}$$

or

$$Z = \frac{b\,f}{\delta}$$

Z = Unknown depth
b = selected baseline
f = focal length of camera
δ = disparity of imaged scene point

What happens with the images when we make the baseline smaller?



original baseline:

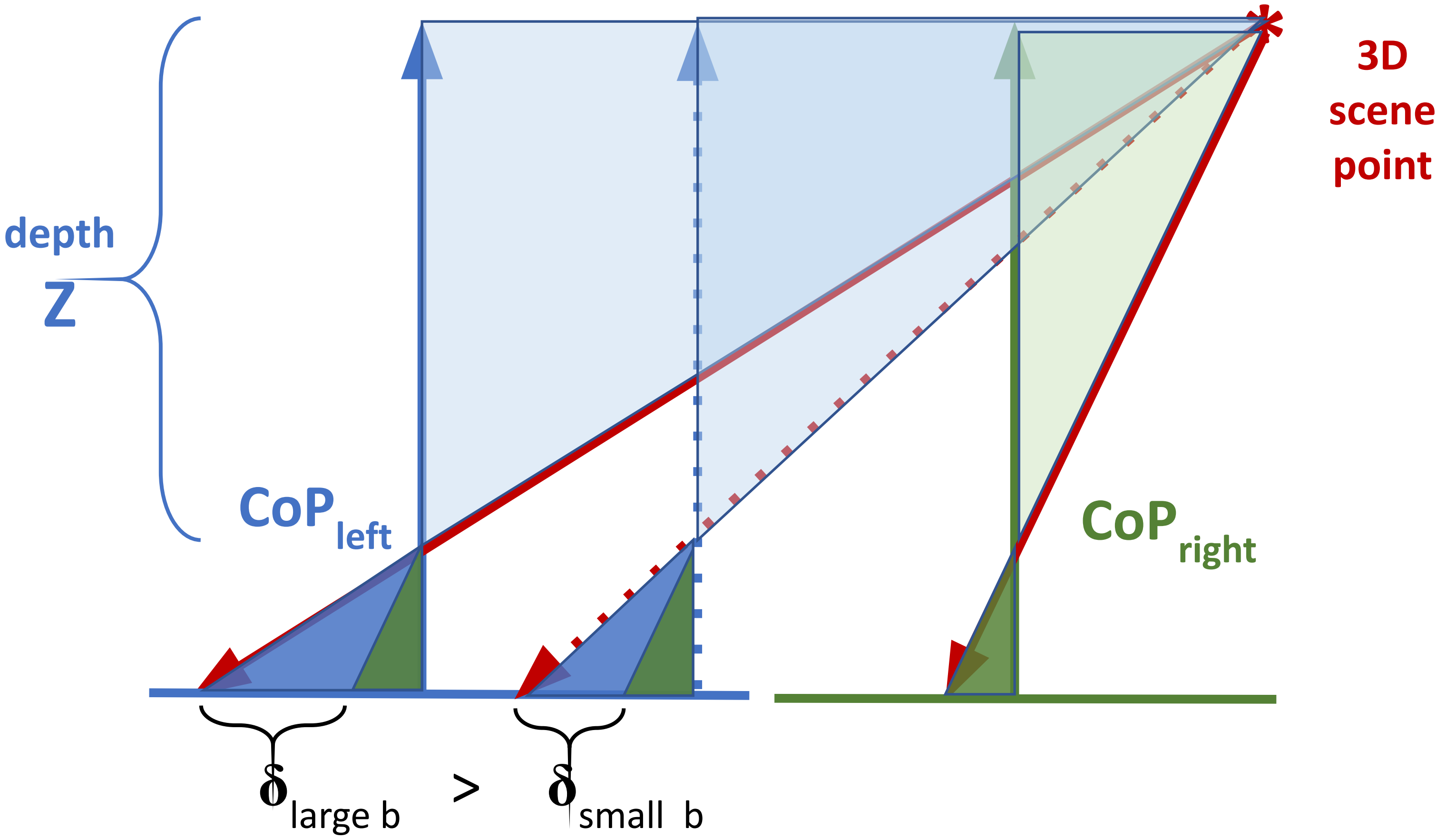
smaller baseline:

Poll: What happens with the images when we make the baseline **smaller**?

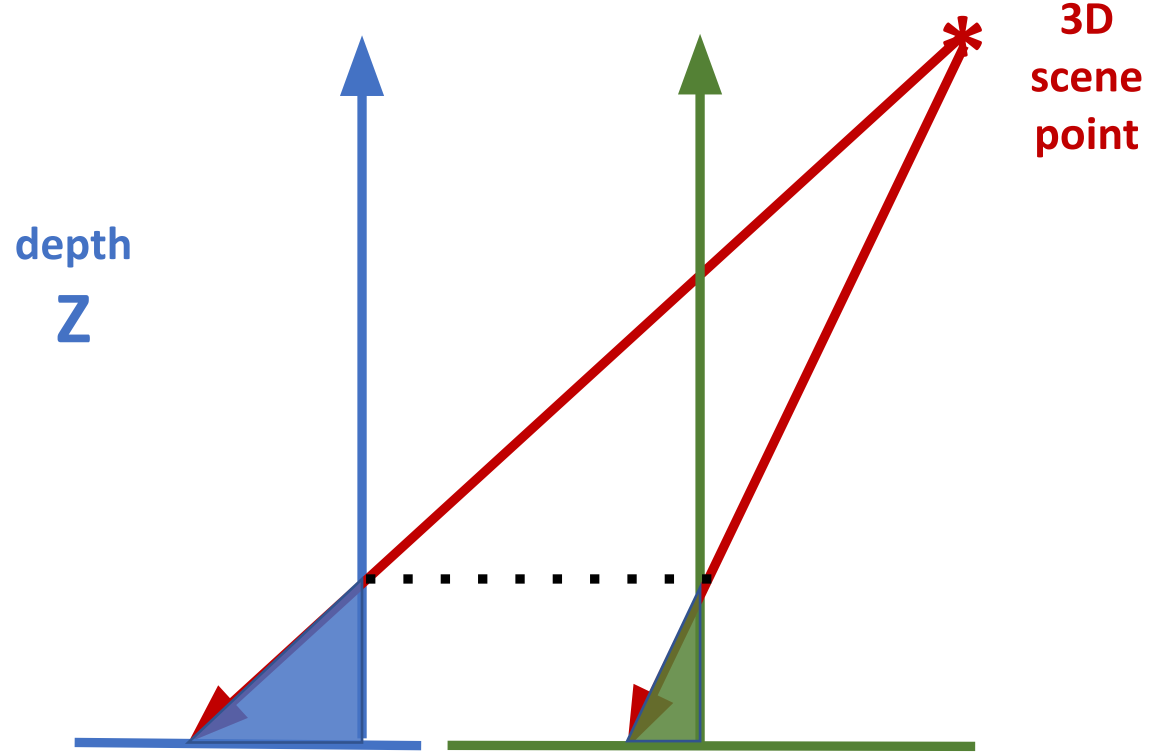
The disparity

1. increases
2. stays the same
3. decreases

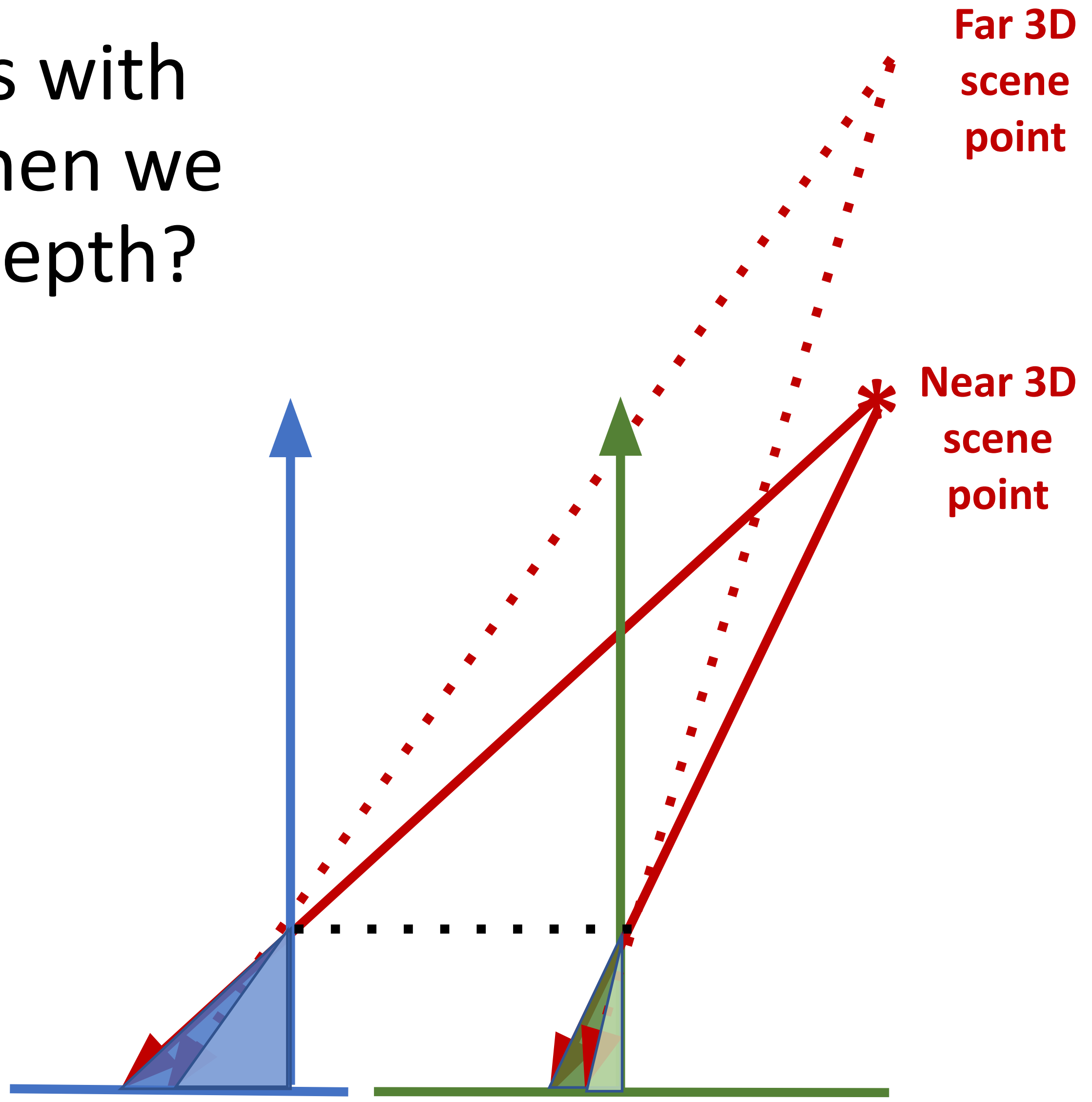
Making the Baseline Smaller Reduces the Disparity



What happens when we increase the depth = distance between scene and image plane?



What happens with the images when we increase the depth?

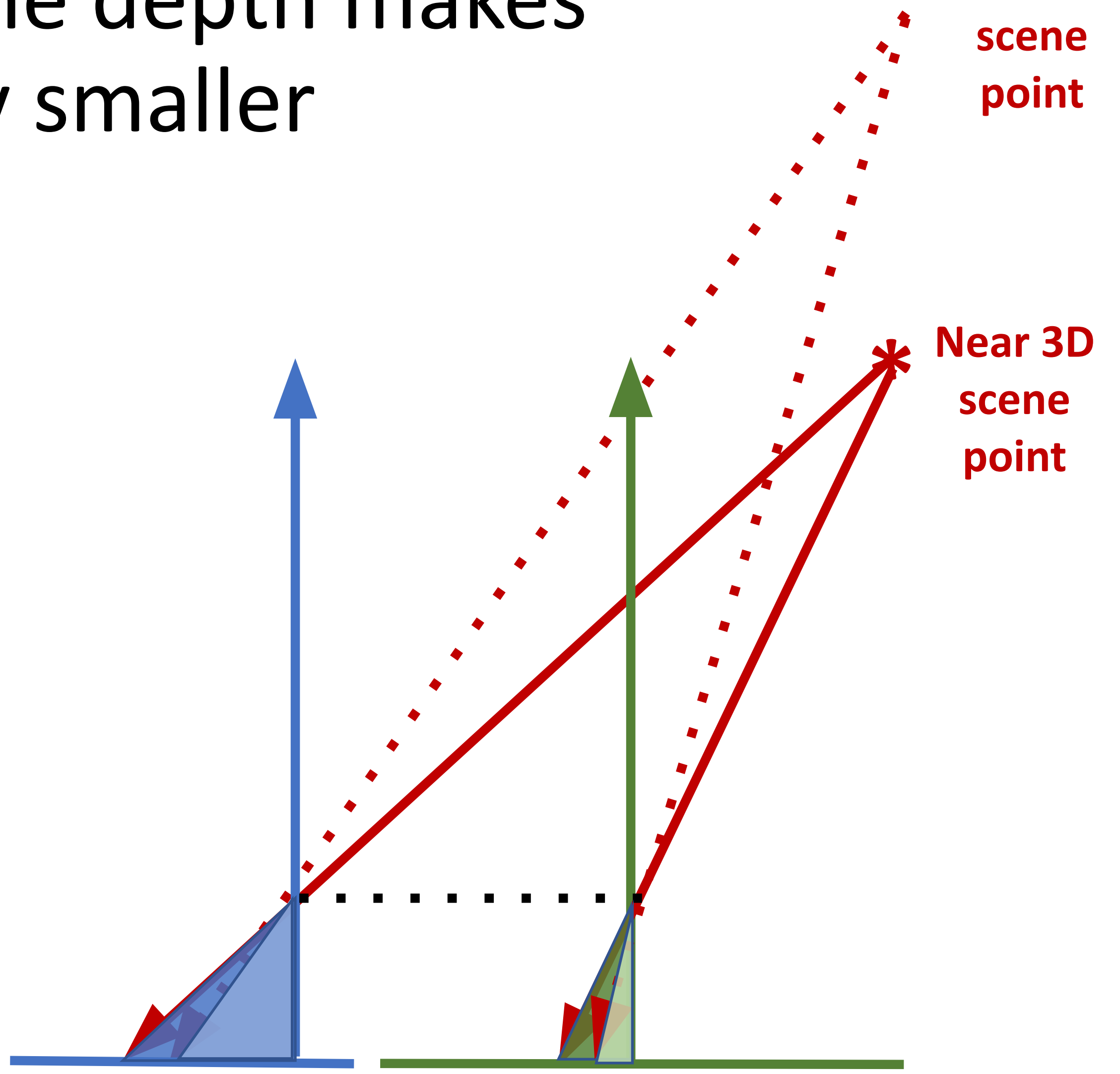


What happens with the images when we **increase** the depth?

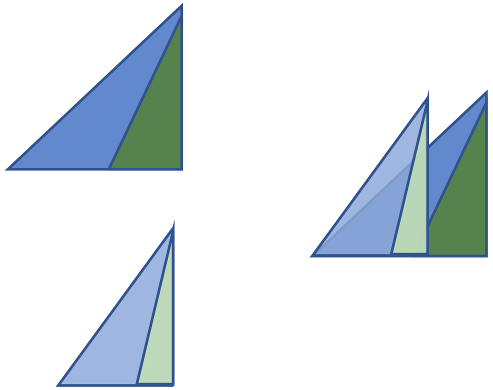
The disparity

1. increases
2. stays the same
3. decreases

Increasing the depth makes the disparity smaller



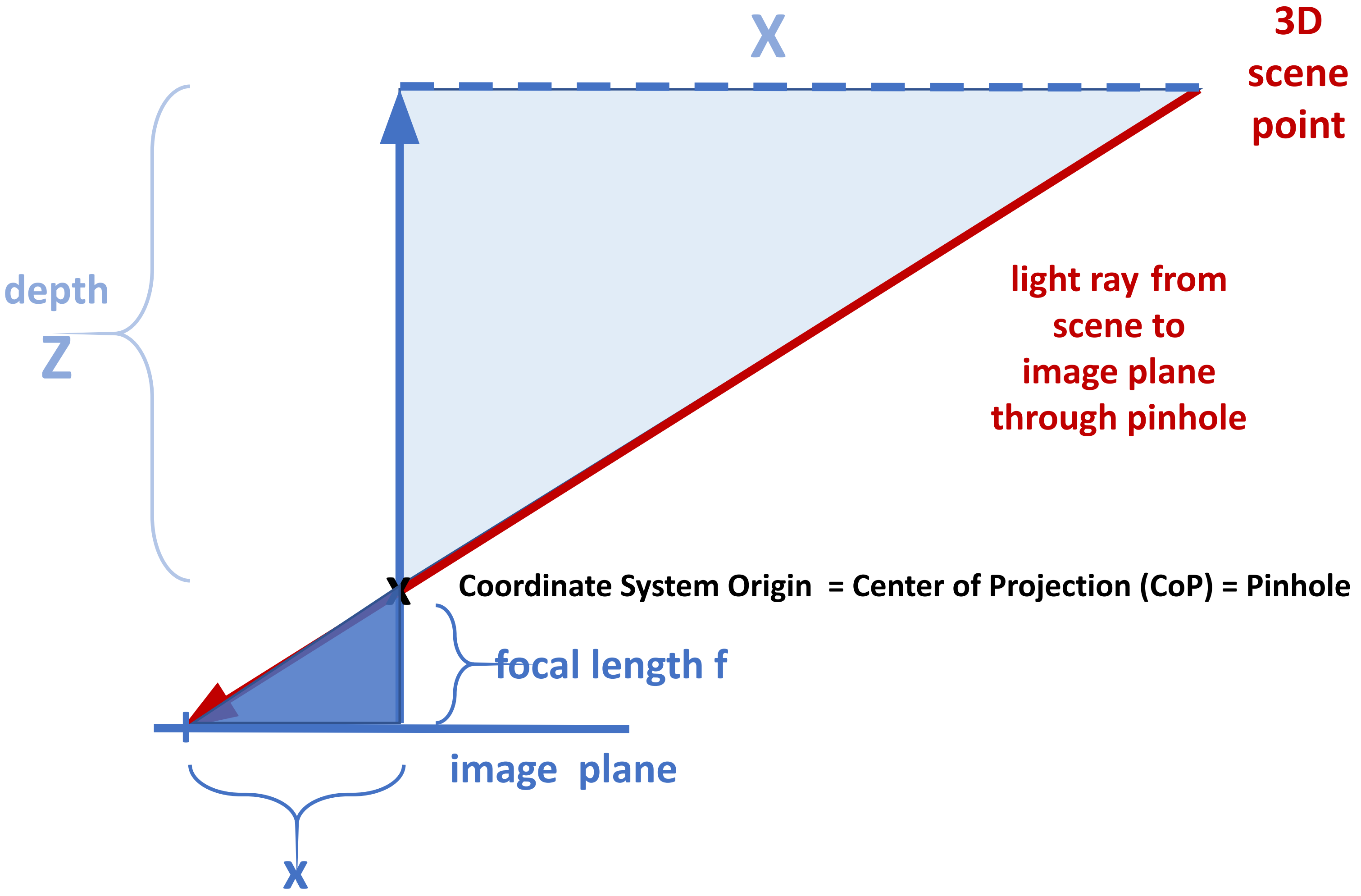
$\delta_{\text{near Z}} > \delta_{\text{far Z}}$



Summary of Concepts: Binocular Stereo

- Today we considered a special case:
  - parallel optical axes
  - image planes aligned
  - same focal length
- Combining perspective projection equations for both cameras yields formula $Z = bf/\delta$
- We discussed how disparity δ changes with changes in b or Z

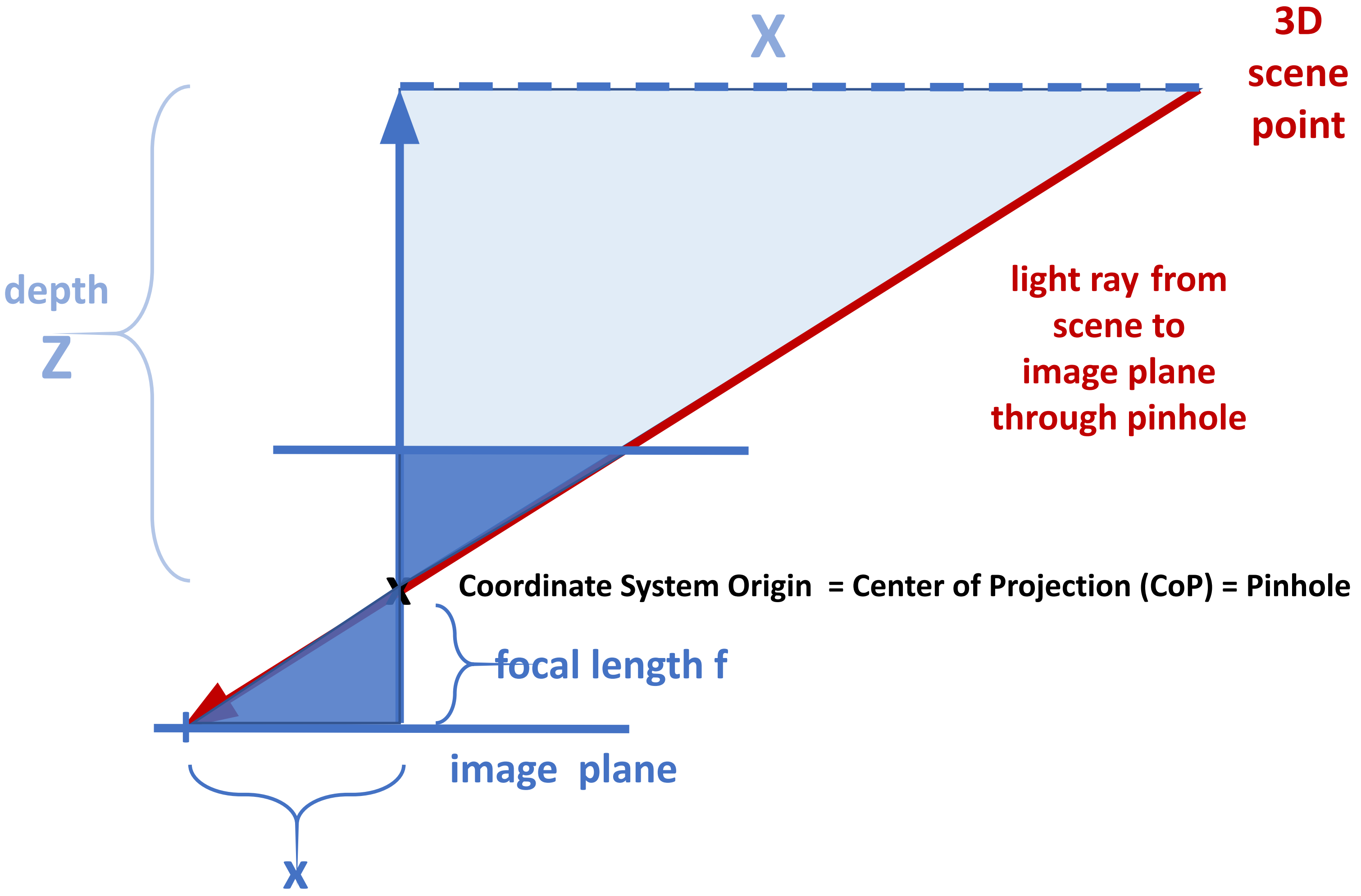
Back to the Single Camera Pinhole Model



Projection Equation:

$X/Z = x/f$

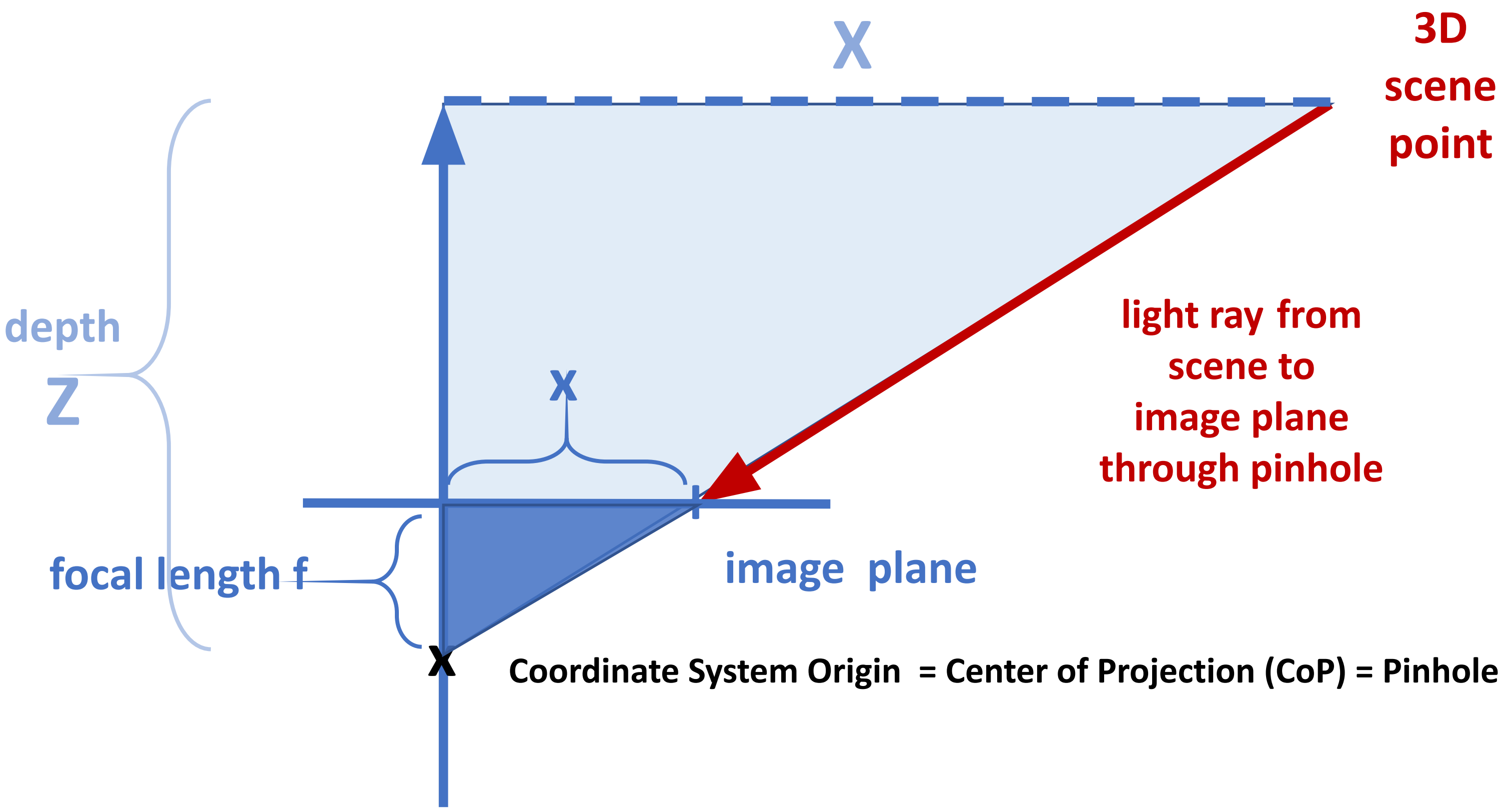
Placing Image Plane in Front of Pinhole



Projection Equation:

$X/Z = x/f$

Placing Image Plane in Front of Pinhole

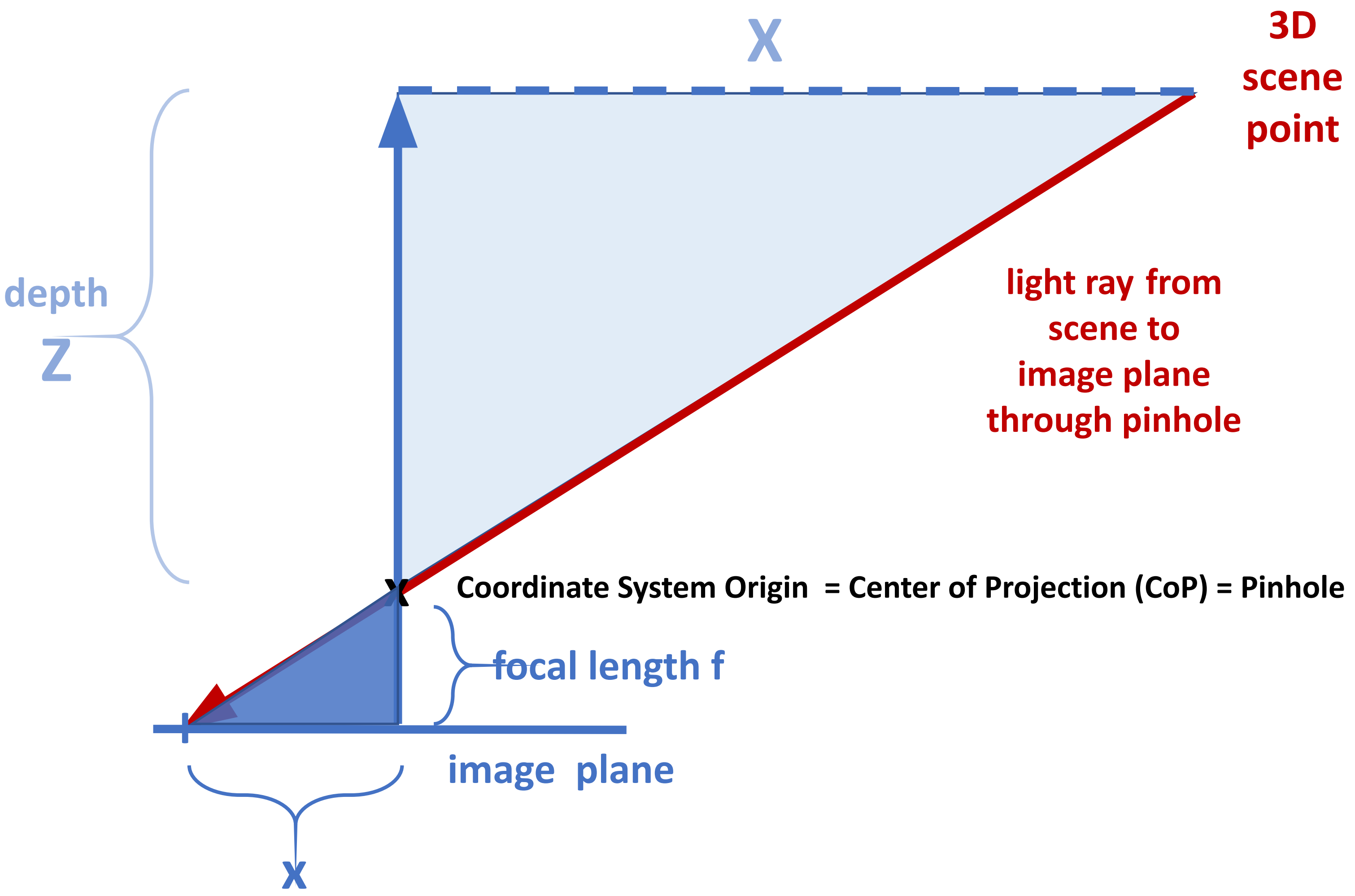


Projection Equation:

$X/Z = x/f$

This is done for **mathematical convenience**: Image x and scene X are measured in the same direction on the x-axis (a positive x means a positive X). The focal length is often set to f=1.

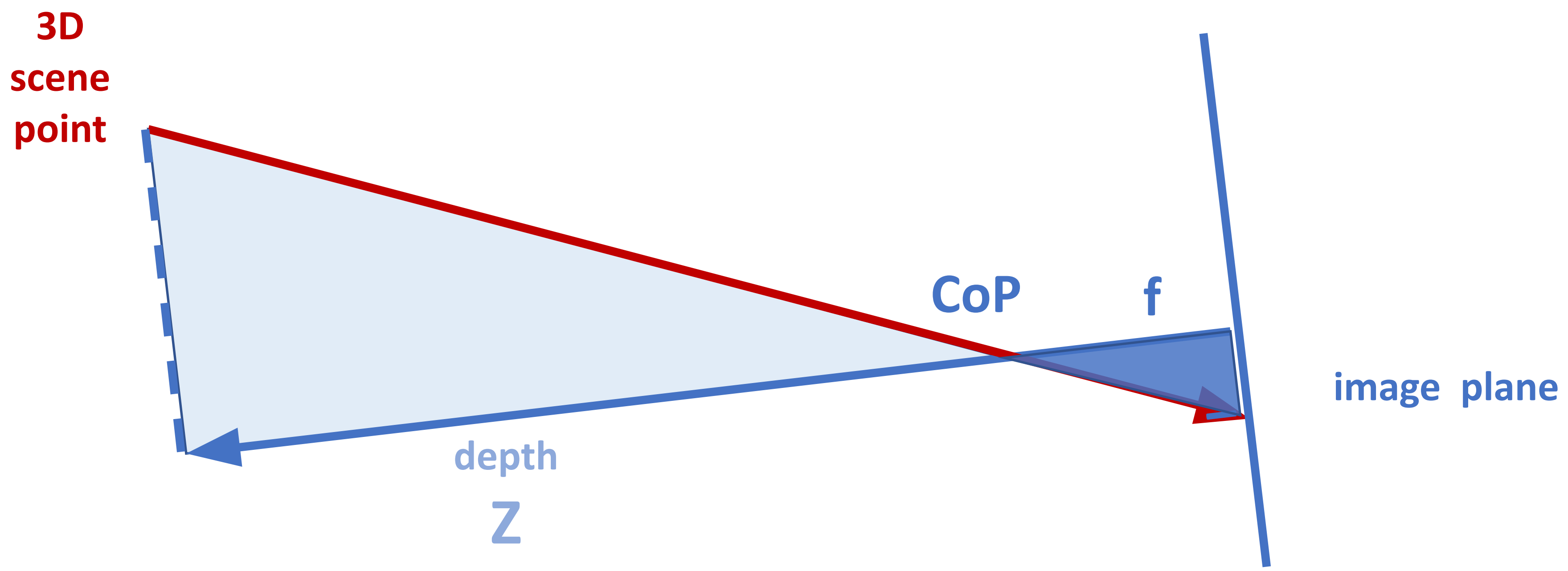
Back to the Original Single Camera Pinhole Model



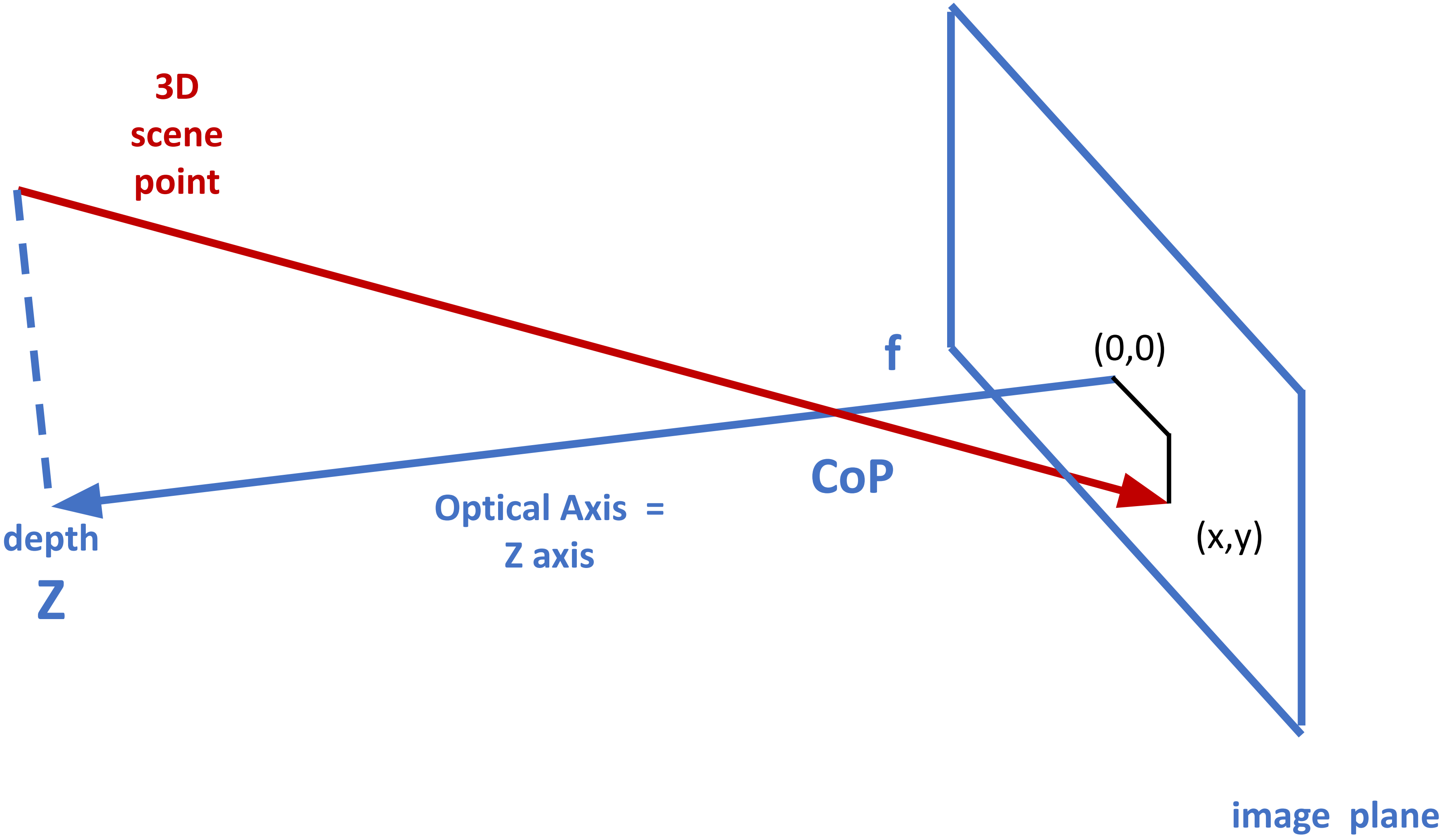
Projection Equation:

$X/Z = x/f$

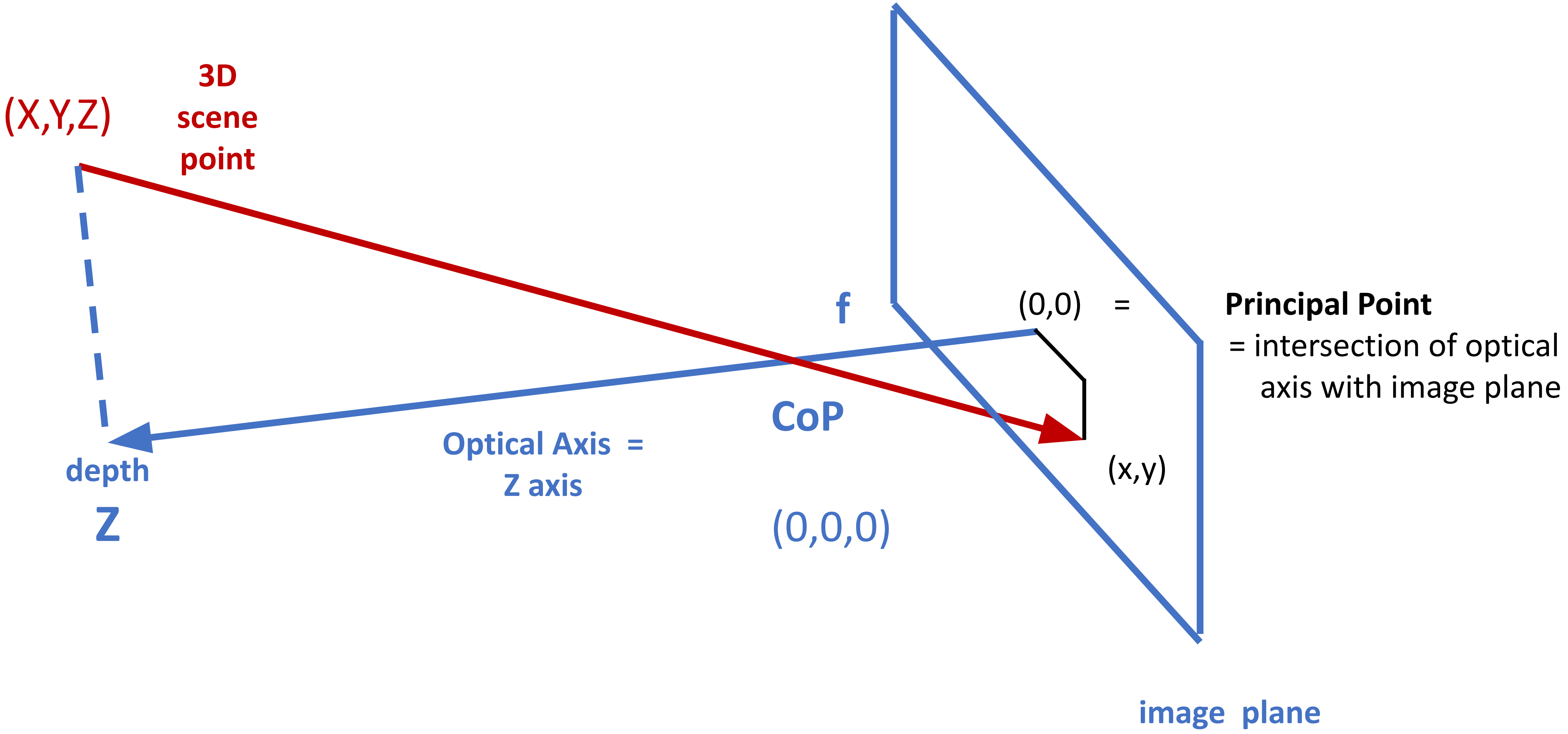
Let's rotate the camera counterclockwise:



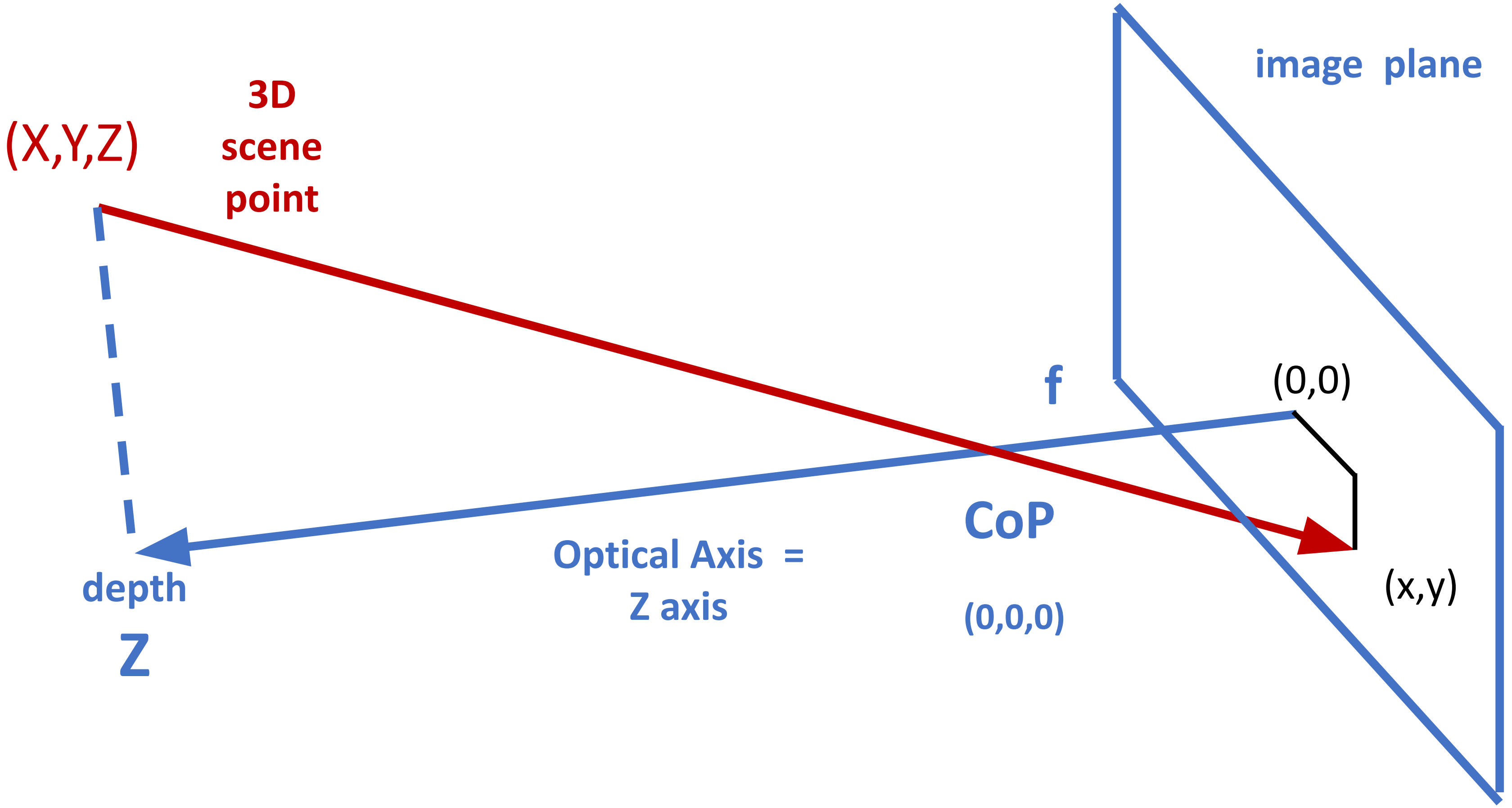
Vertical Dimension Included:



Vertical Dimension Included:



Vertical Dimension Included:



Perspective Projection Equations:

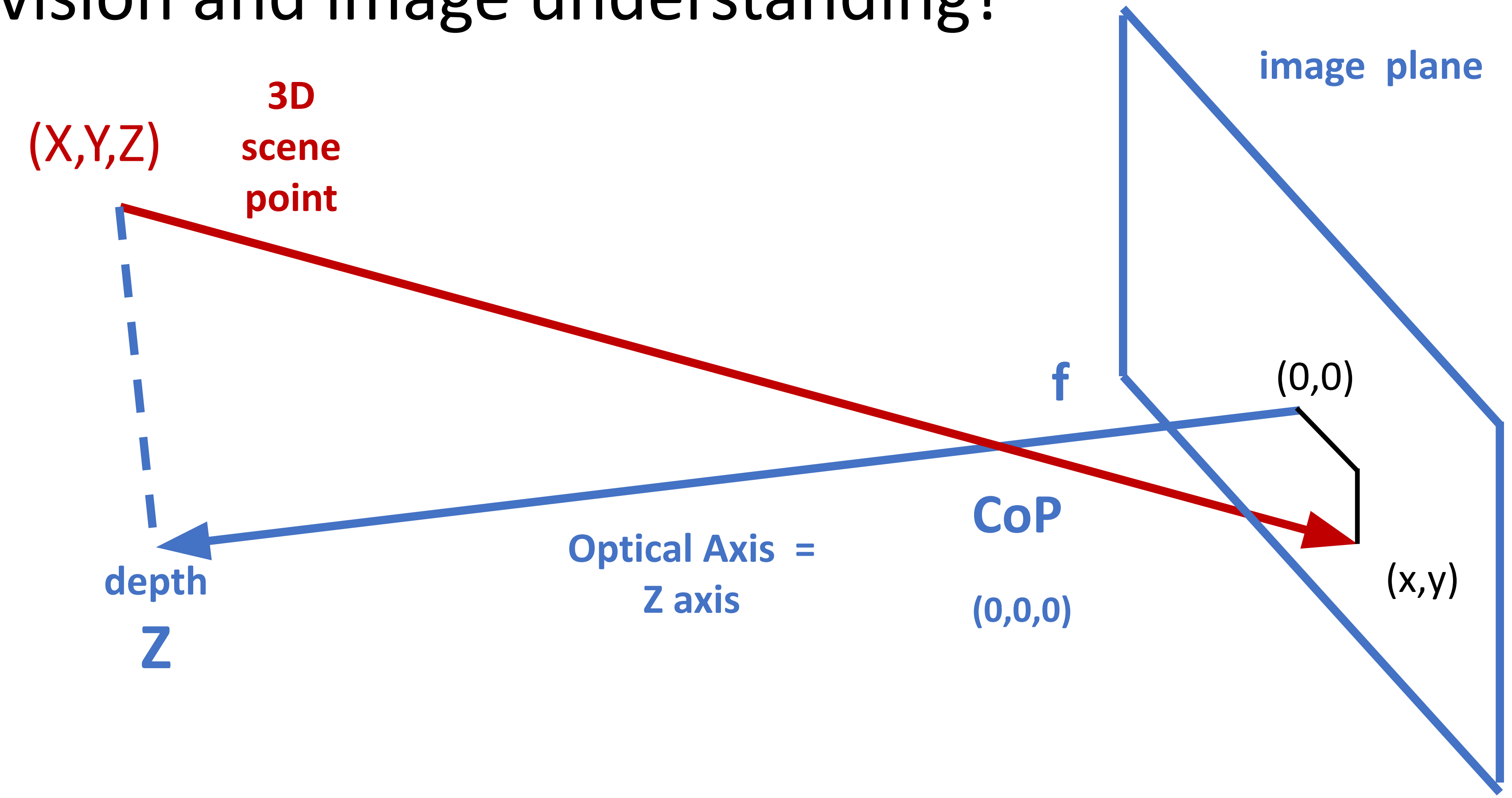
$X/Z = x/f$

and

$Y/Z = y/f$

relate scene (X,Y,Z) and image (x,y) coordinates

Why are perspective projection equations important for computer vision and image understanding?



Perspective Projection Equations:

$X/Z = x/f$

and

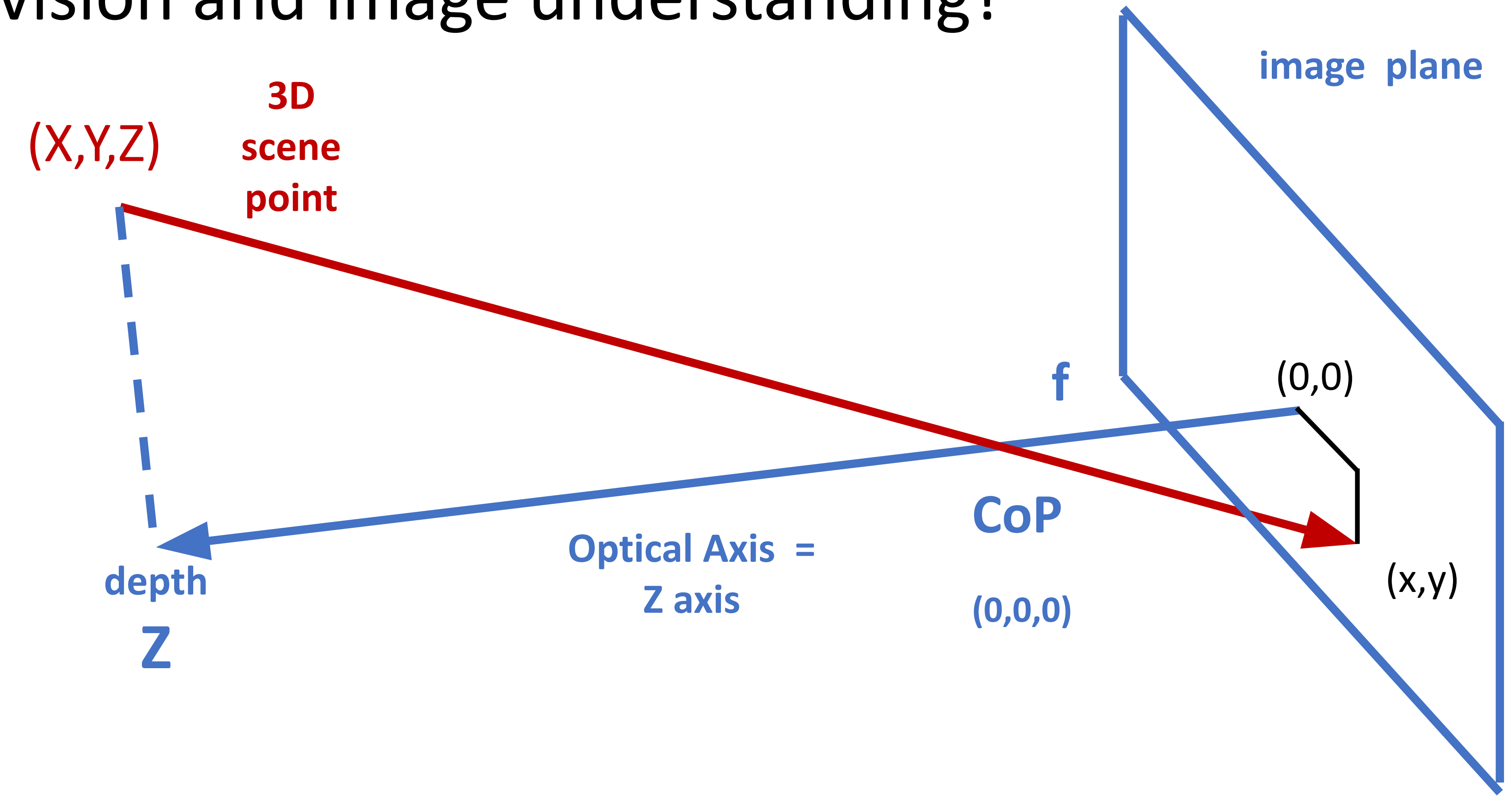
$Y/Z = y/f$

relate scene (X,Y,Z) and image coordinates (x,y)

Poll: Why are perspective projection equations important for computer vision and image understanding?

1. Mathematical reason: We need to learn about the mathematical underpinnings of computer vision.
2. Historical reason: All computer vision students need to learn about these fundamental equations.
3. Computational reason: We can process measurements in the image to infer information about the scene.

Why are perspective projection equations important for computer vision and image understanding?



Perspective Projection Equations:

$X/Z = x/f$

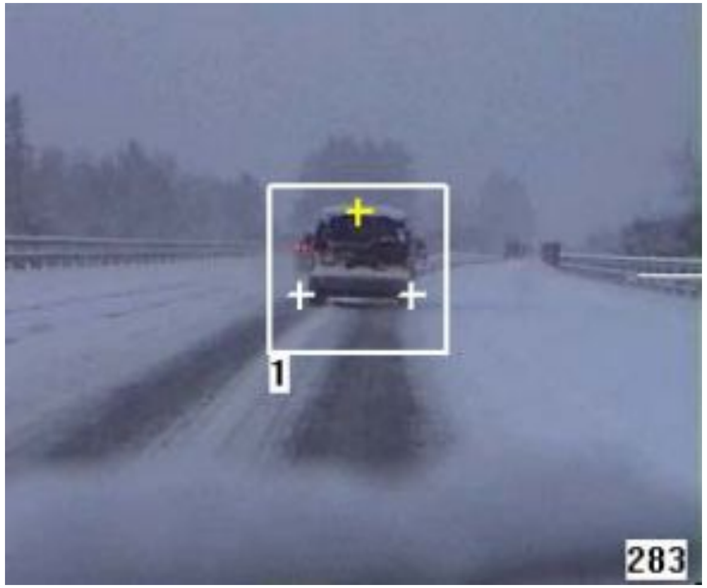
and

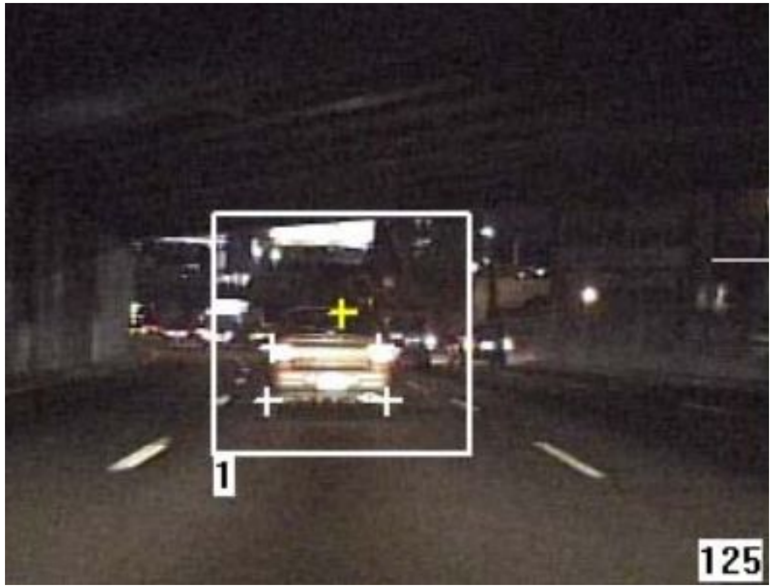
$Y/Z = y/f$

Use measured image coordinates (x,y) to interpret the scene (X,Y,Z)

Example:
Self-driving Cars:

Estimate distance to car in front



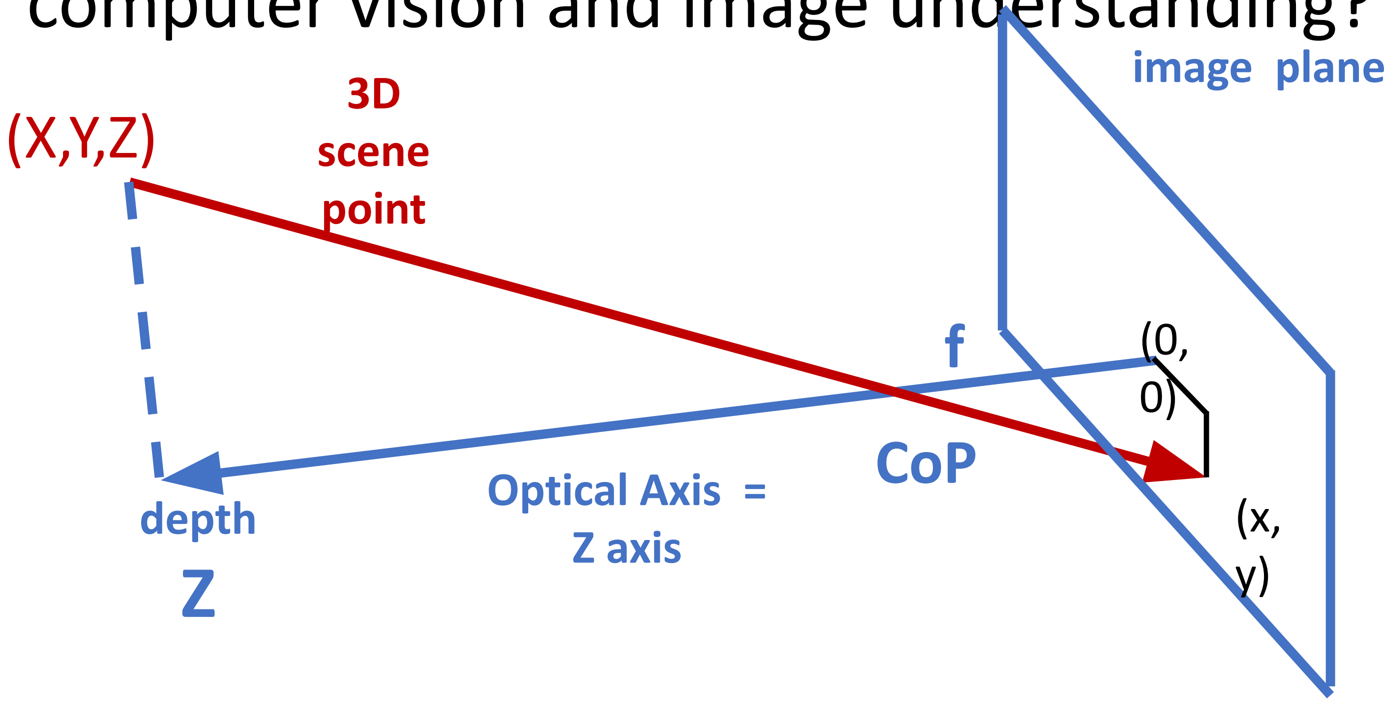


Distance in meters: $Z = c_{horiz} f\, W / w = (22 \text{ pix/mm})(50 \text{ mm})(1.77 \text{ m})/(100 \text{ pix})$
$= 19.47 \text{ m}$

- Typical width of a car: $W = 1.77$ m
- Car width measured in image: $w = 100$ pixel
- focal length $f = 50$ mm
- 35-mm camera: pixel-to-mm conversion $c_{horiz} = 22$ pixel/mm

Images from Betke & Nguyen, IEEE International Conference on Intelligent Vehicles, **1998**

Why are perspective projection equations important for computer vision and image understanding?



Perspective Projection Equations:

$X/Z = x/f$

and

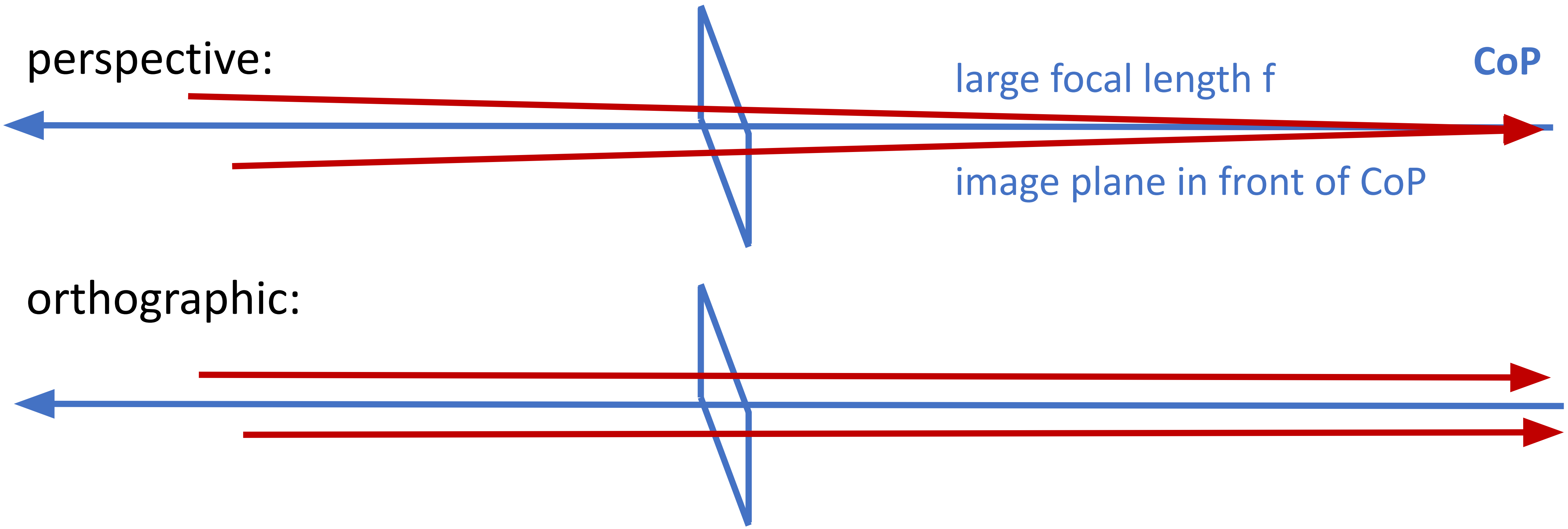
$Y/Z = y/f$

In the self-driving car example:

Use measured image coordinates (w,y) of car in front to interpret its distance: (W,Y, Z)

Orthographic Projection

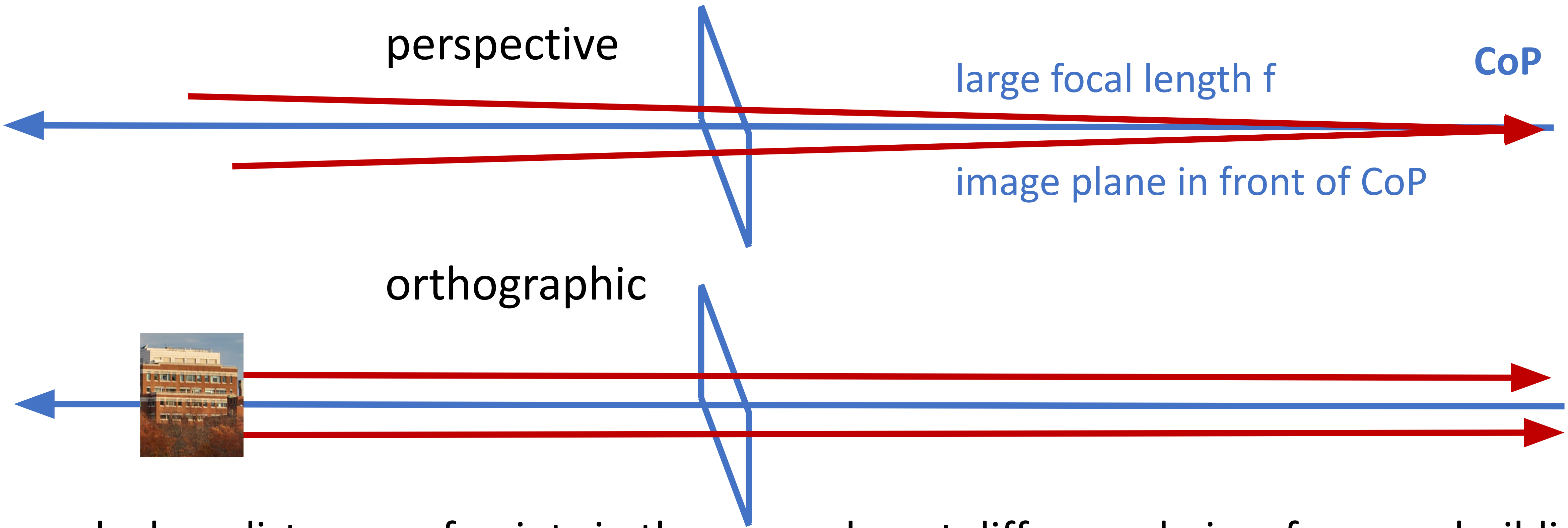
Alternative to perspective projection when imaged objects are far away



light rays assumed parallel to optical axis

Orthographic Projection

Alternative to perspective projection when imaged objects are far away



used when distances of points in the scene do not differ much, i.e., far away building

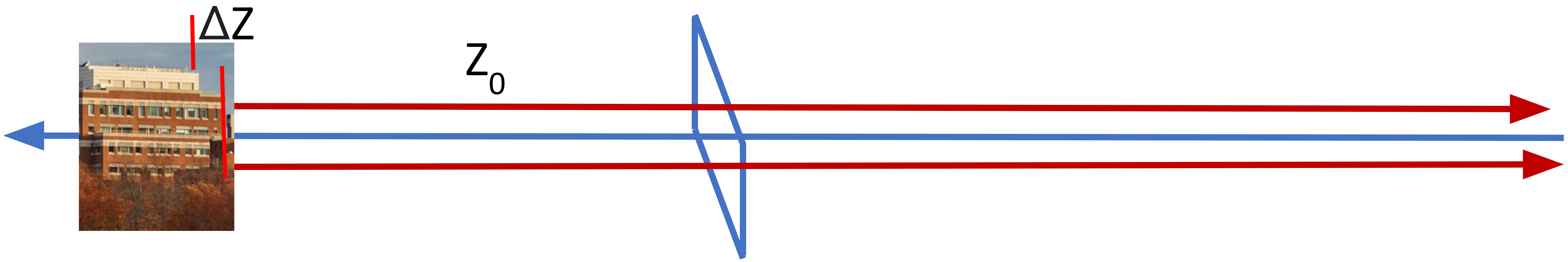
Orthographic Projection

Assume all Z's are approximately at fixed distance Z_0 and $\frac{|\Delta Z|}{Z_0} << 1$

i.e., distances of points in the scene do not differ much

Then $x = (f/Z_0)X$ and $y = (f/Z_0)Y$.

Or assume $x = X$ and $y = Y$. => Simplifies image analysis to 2D problem.



Vanishing Point

Definition:

Point at which receding parallel lines viewed in perspective appear to converge

Photo by Wikipedia user: MikKBDFJKGeMalak

Vanishing Point

Definition:

Point at which receding parallel lines viewed in perspective appear to converge

train tracks

Vanishing Point

Definition:

Point at which receding parallel lines viewed in perspective appear to converge

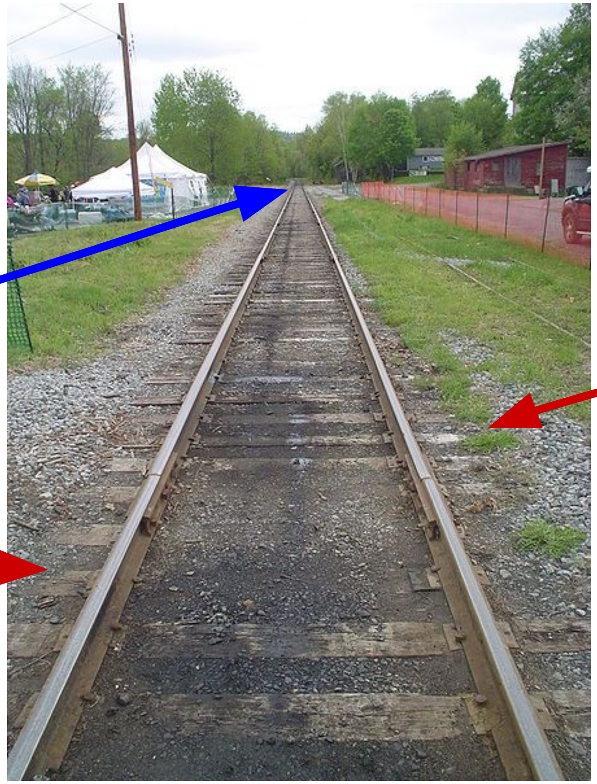


Photo by Wikipedia user: [MikKBDFJKGeMalak]

Why do parallel lines intersect when projected?

Why do parallel lines intersect when projected?

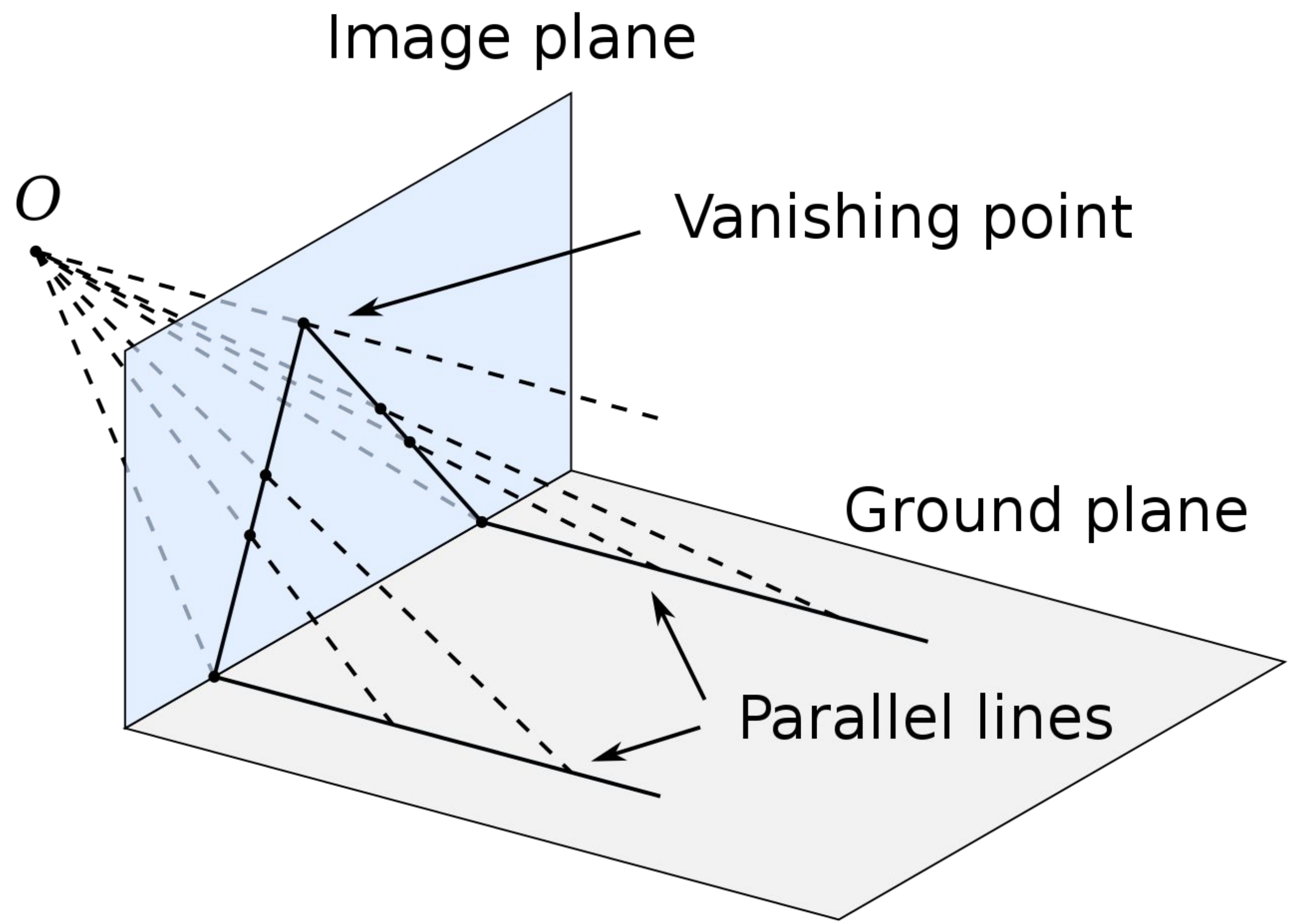


Illustration by Wikipedia user: Mgunyho

One Vanishing Point

Image Source: Betke

Learning Objectives

Be able to explain:

- Pinhole model, camera obscura, center of projection, aperture, principal point, optical axis, focal length, depth, perspective projection, disparity
- What happens to an image if the aperture is increased or decreased?
- What is the impact of changing the baseline or the distance to the scene on the images in a binocular camera system?
- What is the difference between the perspective and orthographic projection models and when should you use them?
- How could you use a perspective projection equation to estimate the distance of a car in front?
- What is a vanishing point? Why could it be useful for highway scene analysis?