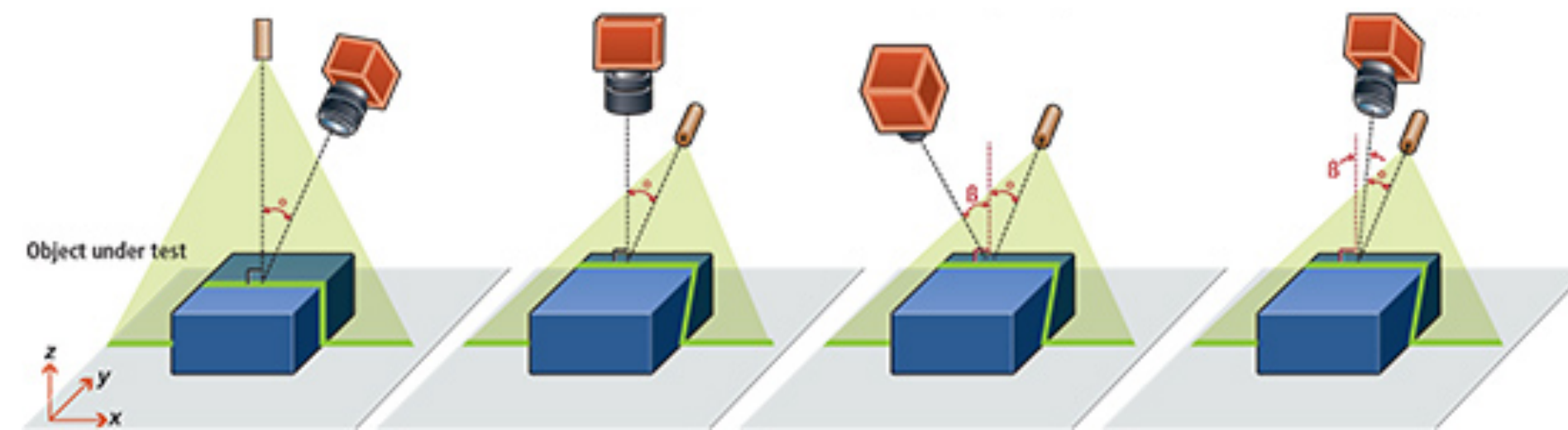


CPSC 425: Computer Vision



Lecture 2: Image Formation

(unless otherwise stated slides are taken or adopted from **Bob Woodham, Jim Little** and **Fred Tung**)

Menu for Today (January 9, 2020)

Updates:

- **Assignment 0** is updated (optional)
- **All assignments** release and due dates are posted
- **Midterm** is scheduled for (right after the break)

Reminders:

- **WWW**: assignments, lecture notes, readings
- **Piazza**: discussion (lecture notes and assignment will also be posted here)
- **Canvas**: assignment hand in and grading

Menu for Today (January 9, 2020)

Topics:

- Image Formation
- Cameras and Lenses
- Projection
- Human eye (as camera)

Readings:

- **Today's** Lecture: Forsyth & Ponce (2nd ed.) 1.1.1 — 1.1.3
- **Next** Lecture: Forsyth & Ponce (2nd ed.) 4.1, 4.5

Reminders:

- Complete **Assignment 0** (ungraded) by Tuesday, **January 14**
- **WWW**: <http://www.cs.ubc.ca/~lsgal/teaching.html>
- **Piazza**: piazza.com/ubc.ca/winterterm22020/cpsc425201/home

Today's "fun" Example



Photo credit: reddit user [Liammm](#)

Today's "fun" Example: **Eye Sink Illusion**

Pereidolia



Photo credit: reddit user [Liammm](#)

Today's "fun" Example: **Eye Sink Illusion**



“Tried taking a picture of a sink draining, wound up with a picture of an eye instead”

Photo credit: reddit user [Liammm](#)

Lecture 1: Re-cap

Types of computer vision **problems**:

- Computing properties of the 3D world from visual data (***measurement***)
- Recognition of objects and scenes (***perception and interpretation***)
- Search and interact with visual data (***search and organization***)
- Manipulation or creation of image or video content (***visual imagination***)

Computer vision **challenges**:

- Fundamentally **ill-posed**
- Enormous **computation** and **scale**
- Lack of fundamental understanding of how **human perception** works

Lecture 1: Re-cap

Computer vision technologies have moved **from research labs into commercial products and services**. Examples cited include:

- broadcast television sports
- electronic games (Microsoft Kinect)
- biometrics
- image search
- visual special effects
- medical imaging
- robotics

... many others

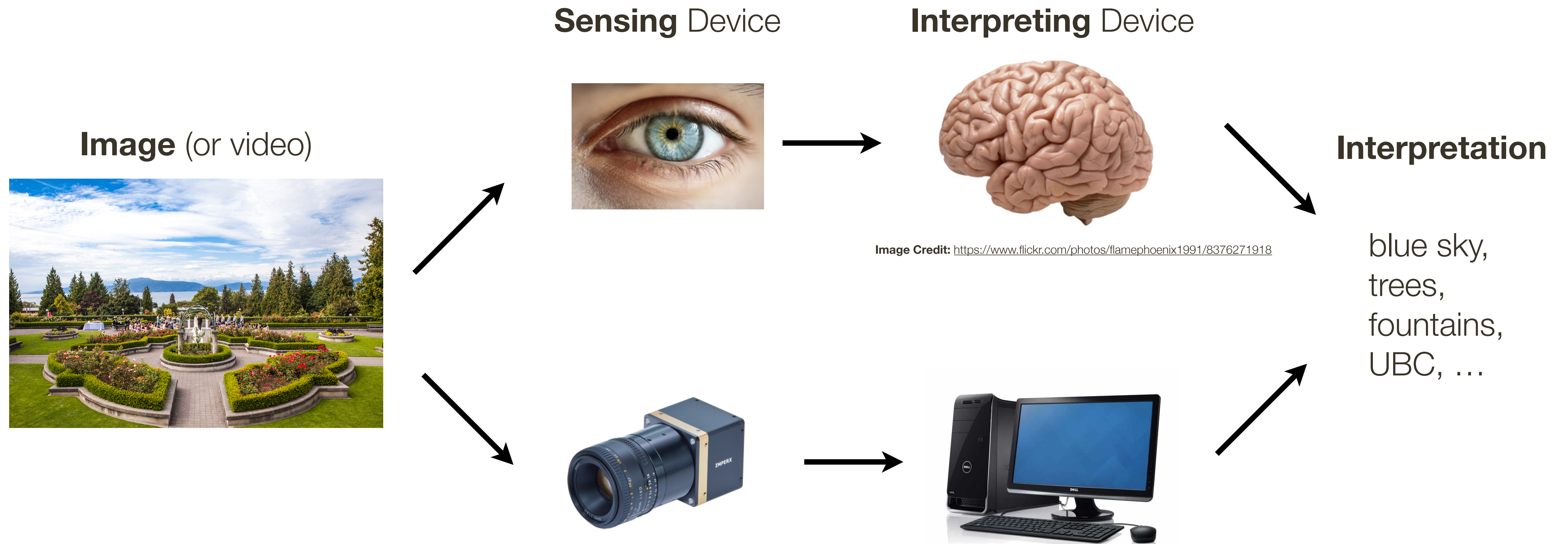
Lecture 2: Goal

To understand how images are formed

(and develop relevant mathematical
concepts and abstractions)

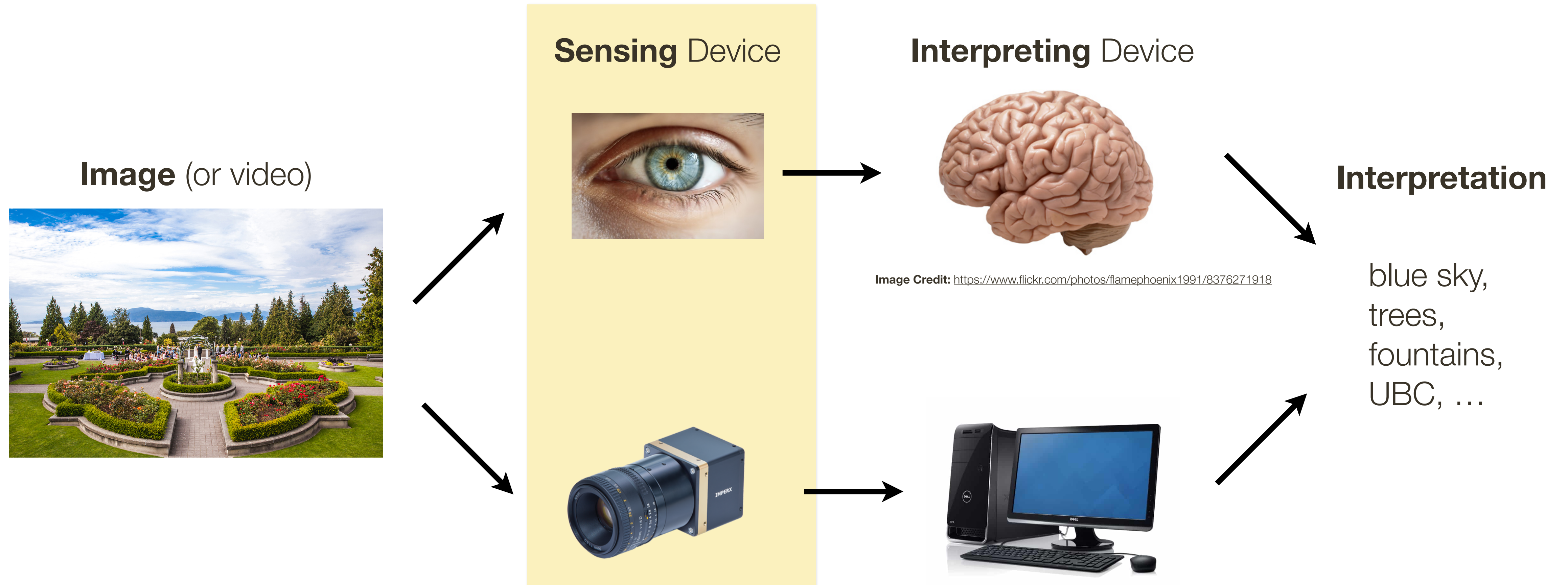
What is **Computer Vision**?

Computer vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.



What is **Computer Vision**?

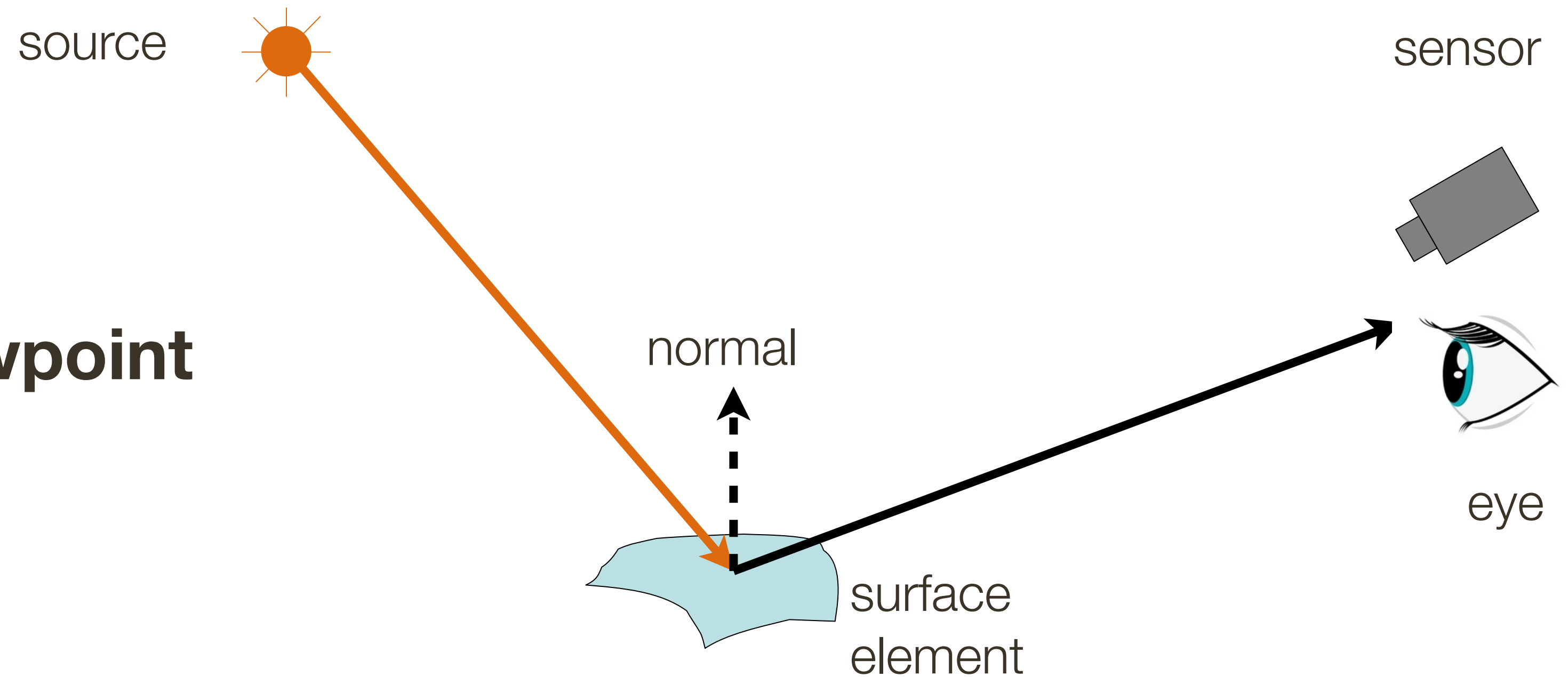
Computer vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.



Overview: Image Formation, Cameras and Lenses

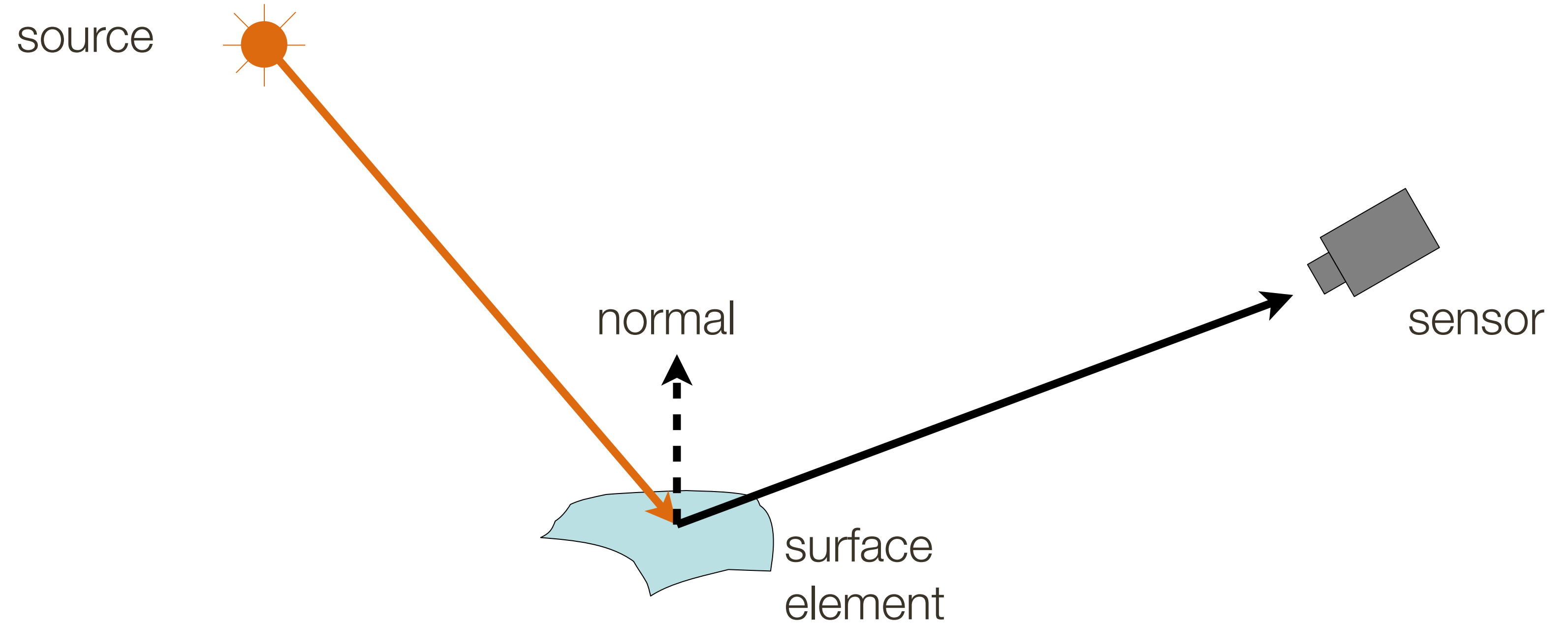
The **image formation process** that produces a particular image depends on

- **Lighting** condition
- Scene **geometry**
- **Surface** properties
- Camera **optics** and **viewpoint**

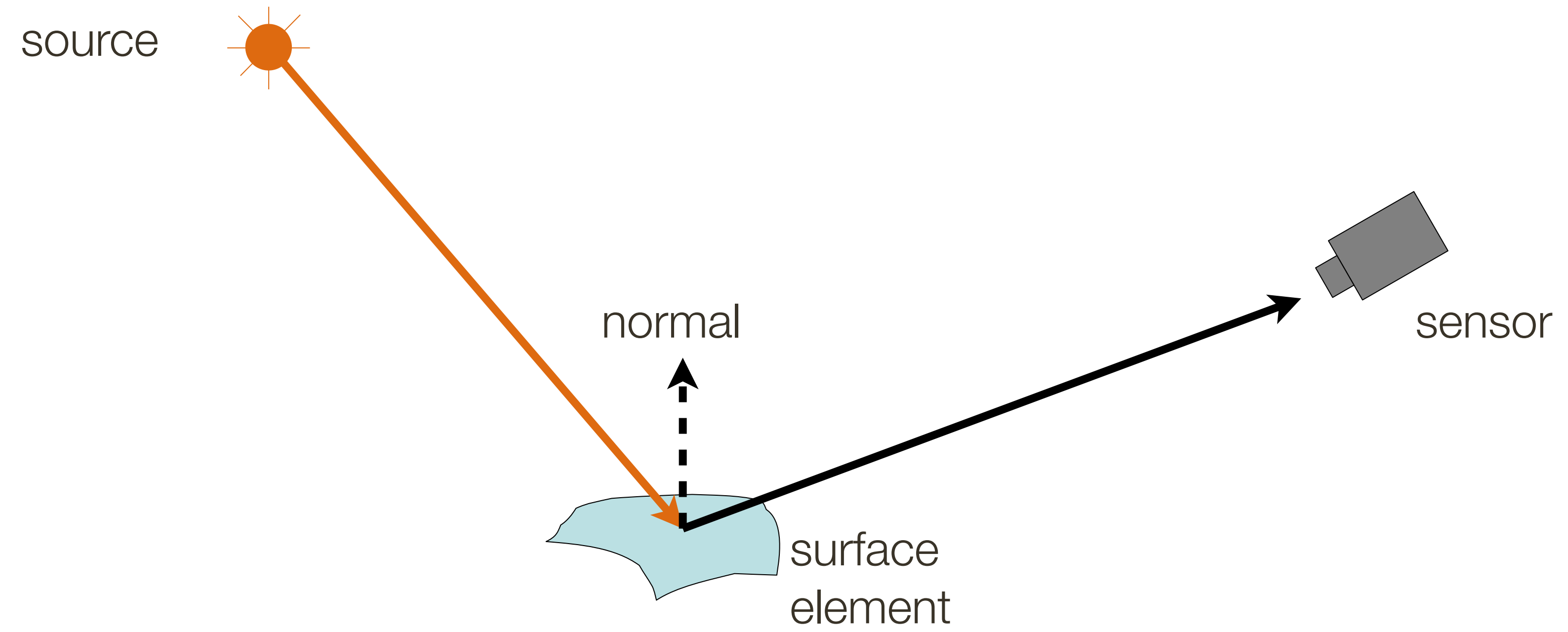


Sensor (or eye) **captures amount of light** reflected from the object

(small) Graphics Review

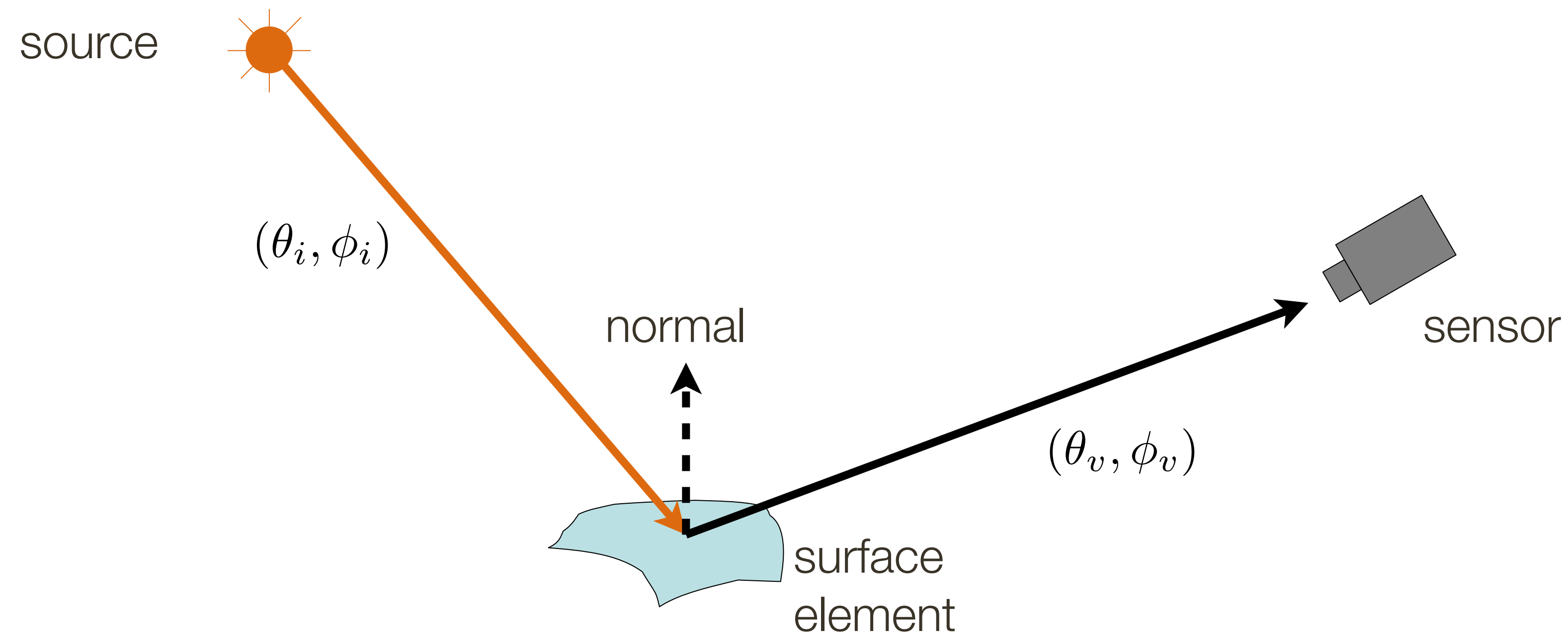


(small) Graphics Review



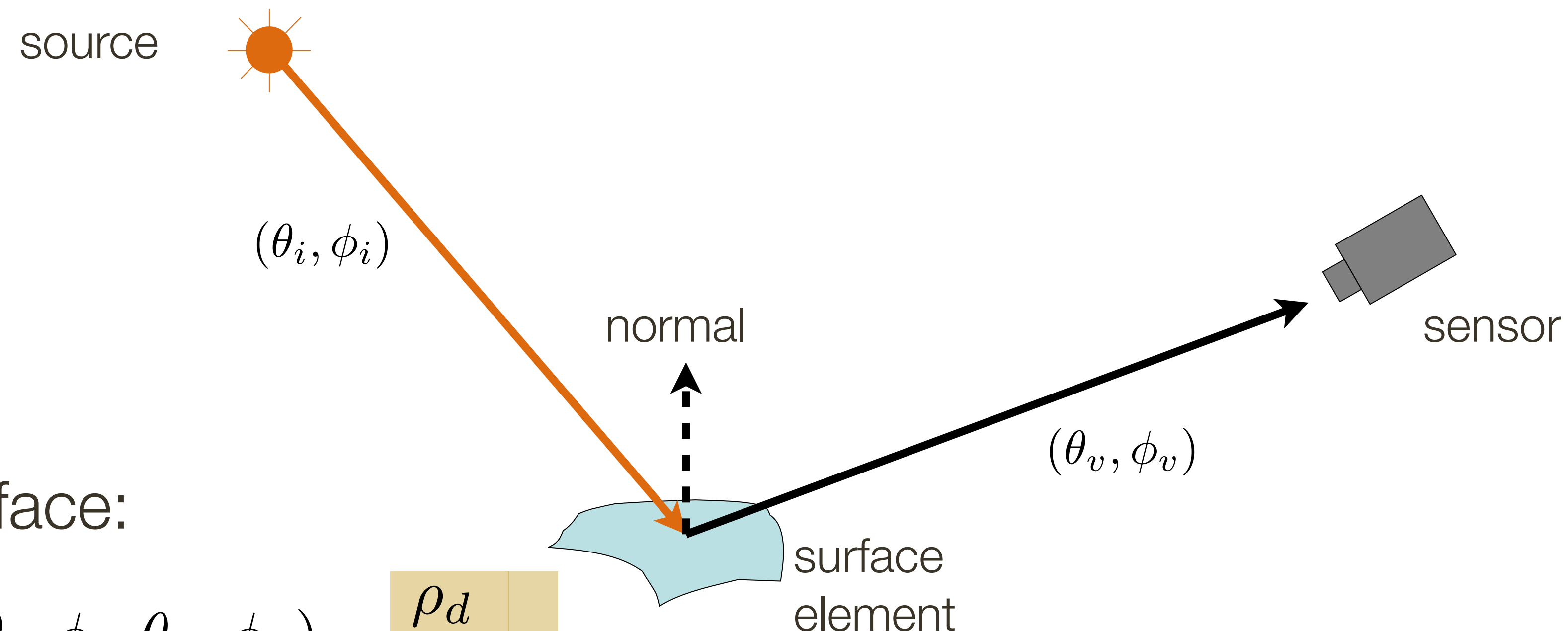
(small) **Graphics** Review

Surface reflection depends on both the **viewing** (θ_v, ϕ_v) and **illumination** (θ_i, ϕ_i) direction, with Bidirectional Reflection Distribution Function: **BRDF** $(\theta_i, \phi_i, \theta_v, \phi_v)$



(small) Graphics Review

Surface reflection depends on both the **viewing** (θ_v, ϕ_v) and **illumination** (θ_i, ϕ_i) direction, with Bidirectional Reflection Distribution Function: **BRDF** $(\theta_i, \phi_i, \theta_v, \phi_v)$



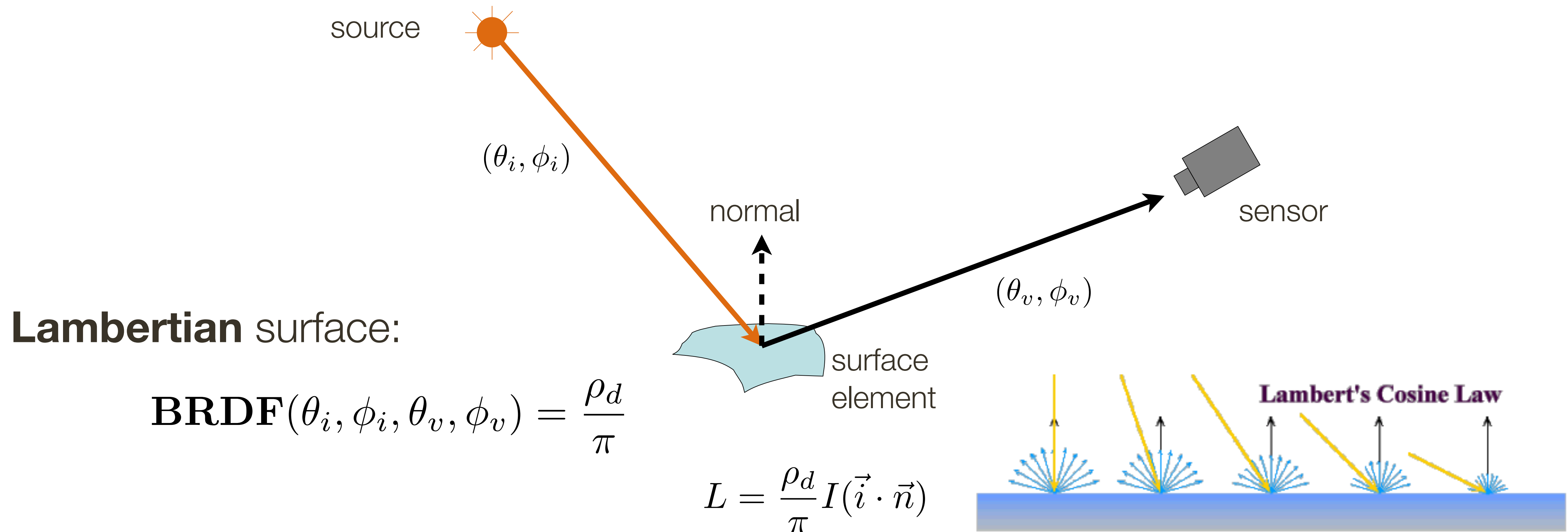
Lambertian surface:

$$\text{BRDF}(\theta_i, \phi_i, \theta_v, \phi_v) = \frac{\rho_d}{\pi}$$

constant, called **albedo**

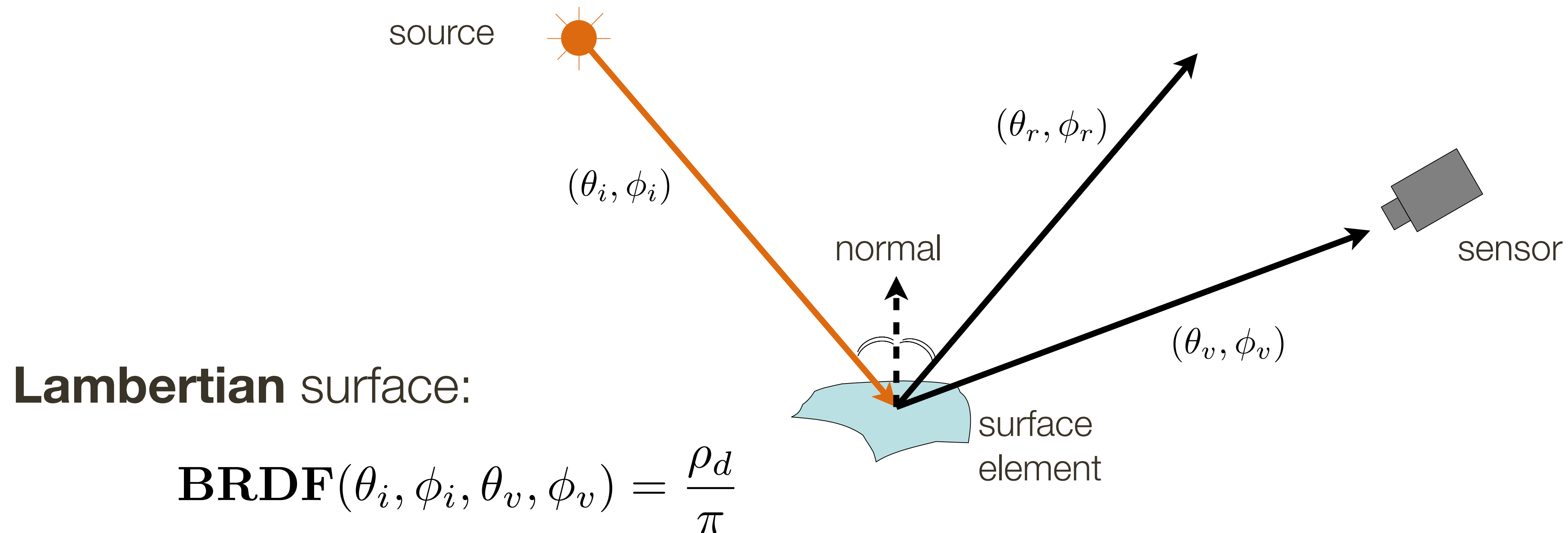
(small) Graphics Review

Surface reflection depends on both the **viewing** (θ_v, ϕ_v) and **illumination** (θ_i, ϕ_i) direction, with Bidirectional Reflection Distribution Function: **BRDF** $(\theta_i, \phi_i, \theta_v, \phi_v)$



(small) Graphics Review

Surface reflection depends on both the **viewing** (θ_v, ϕ_v) and **illumination** (θ_i, ϕ_i) direction, with Bidirectional Reflection Distribution Function: **BRDF** $(\theta_i, \phi_i, \theta_v, \phi_v)$



Mirror surface: all incident light reflected in one directions $(\theta_v, \phi_v) = (\theta_r, \phi_r)$

Cameras

Old school **film** camera



Digital CCD/CMOS camera



Let's say we have a **sensor** ...

Digital CCD/CMOS camera



Let's say we have a **sensor** ...

Digital CCD/CMOS camera



Let's say we have a **sensor** ...

Digital CCD/CMOS camera



digital sensor
(CCD or
CMOS)

... and the **object** we would like to photograph

What would an image taken like this look like?

real-world
object



digital sensor
(CCD or
CMOS)



Bare-sensor imaging

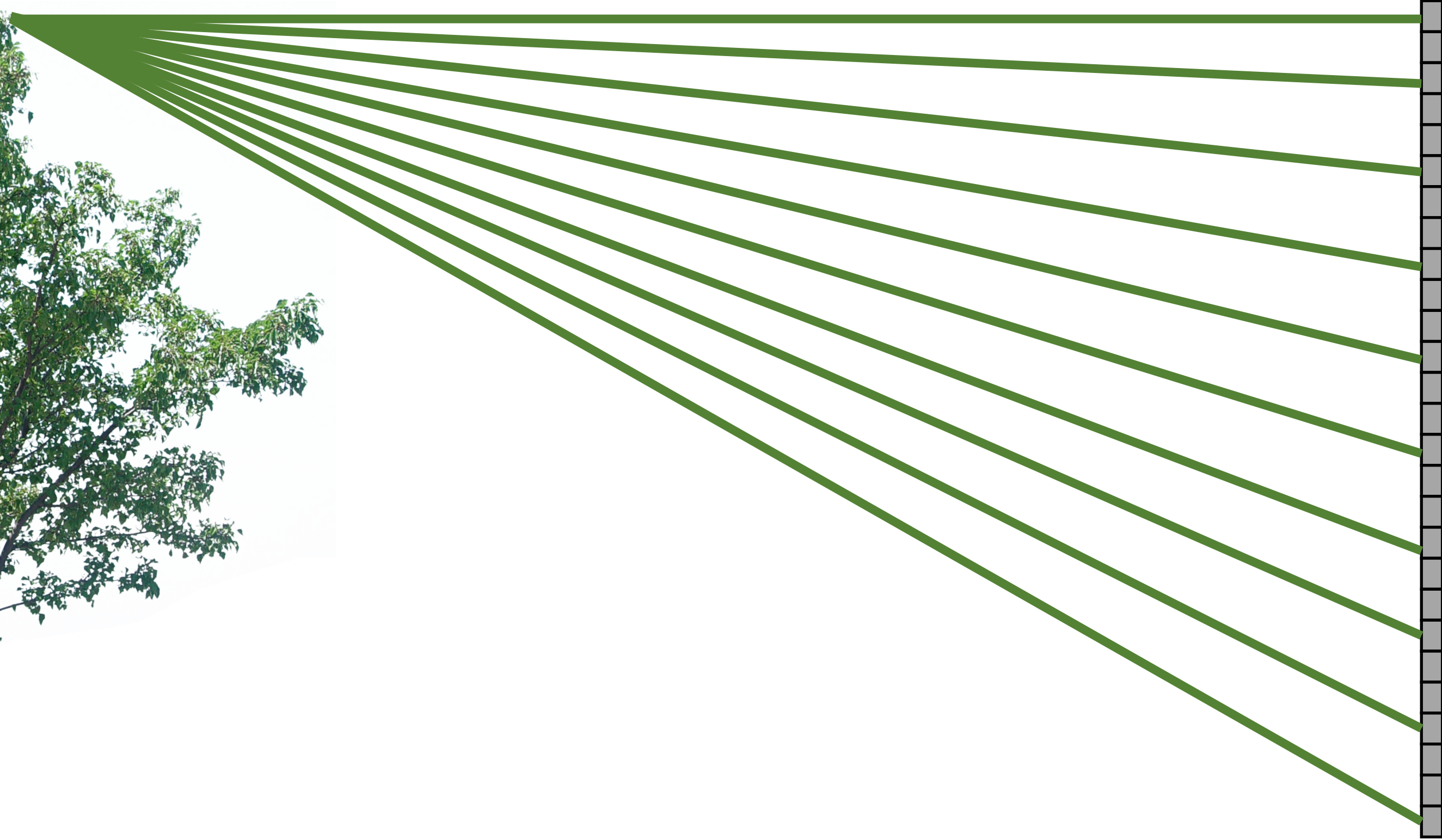
real-world
object



digital sensor
(CCD or
CMOS)

Bare-sensor imaging

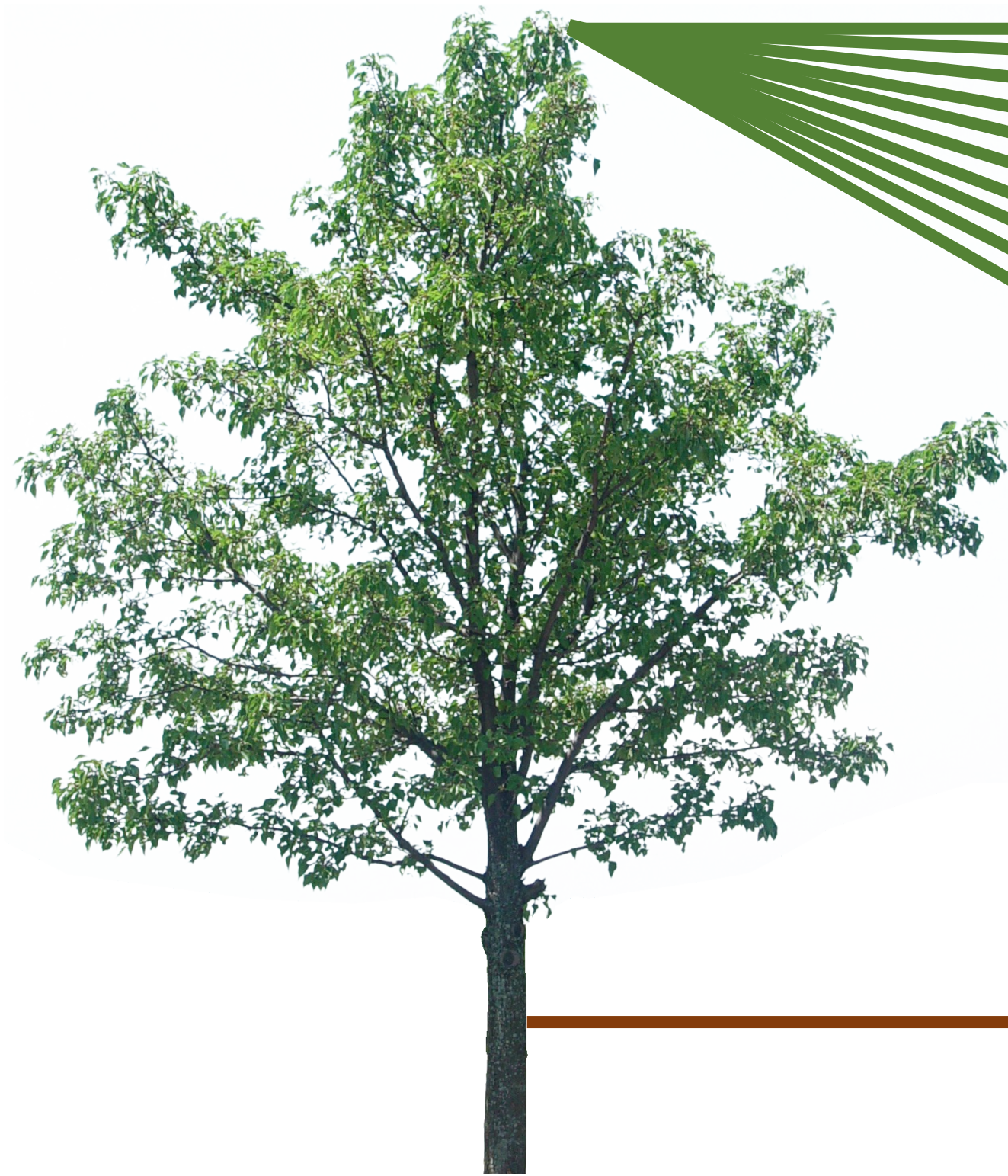
real-world
object



digital sensor
(CCD or
CMOS)

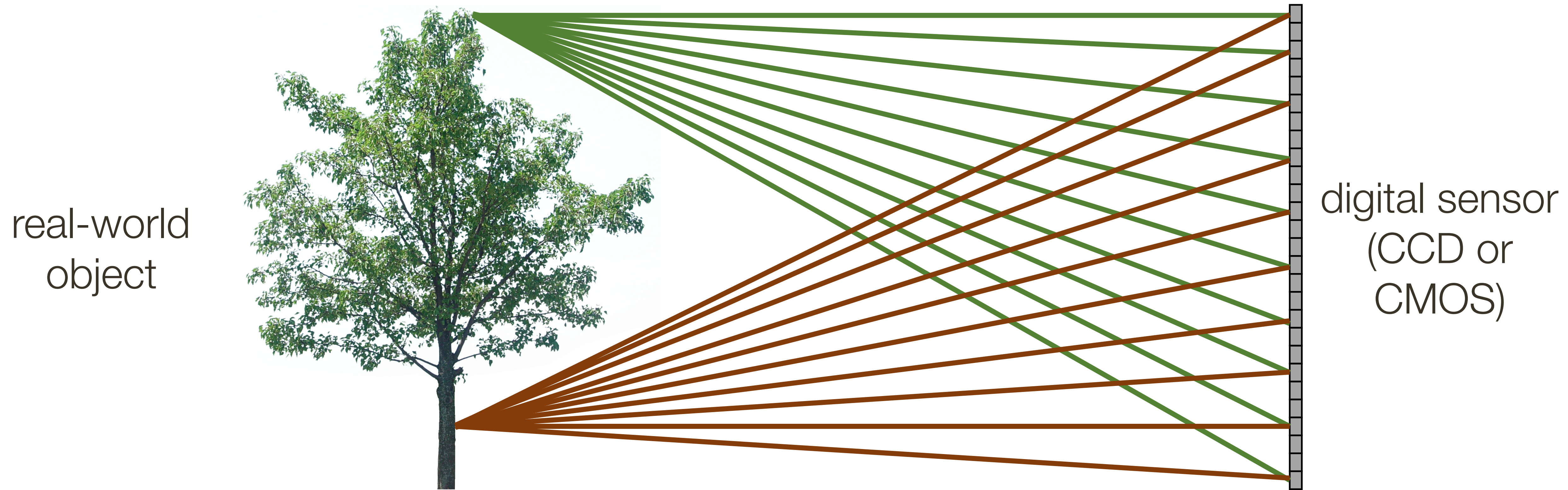
Bare-sensor imaging

real-world
object



digital sensor
(CCD or
CMOS)

Bare-sensor imaging



All scene points contribute to all sensor pixels

Bare-sensor imaging



All scene points contribute to all sensor pixels

Pinhole Camera

real-world
object



barrier (diaphragm)



pinhole
(aperture)



digital sensor
(CCD or
CMOS)

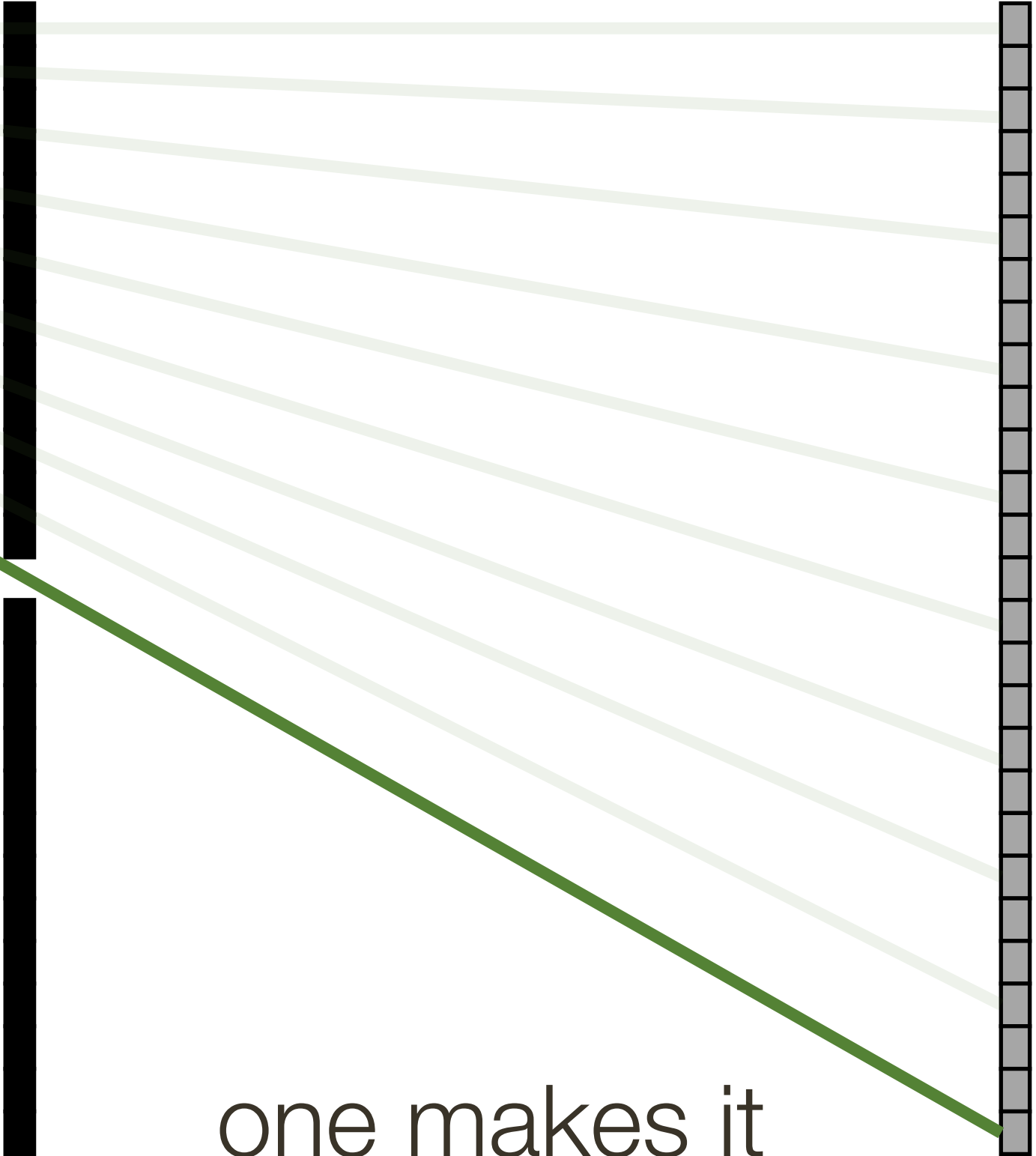
What would an image taken like this look like?

Pinhole Camera

real-world
object



most rays are
blocked

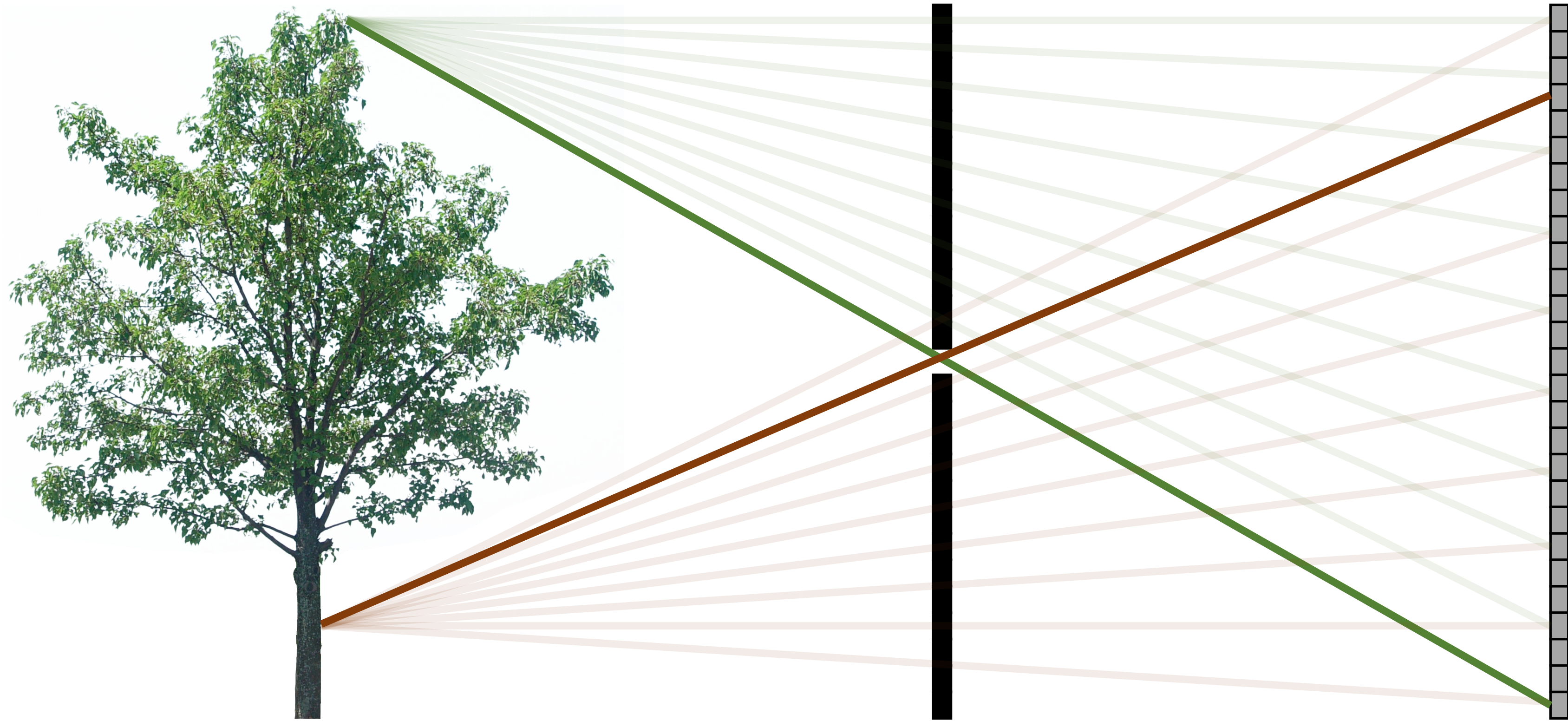


digital sensor
(CCD or
CMOS)

one makes it
through

Pinhole Camera

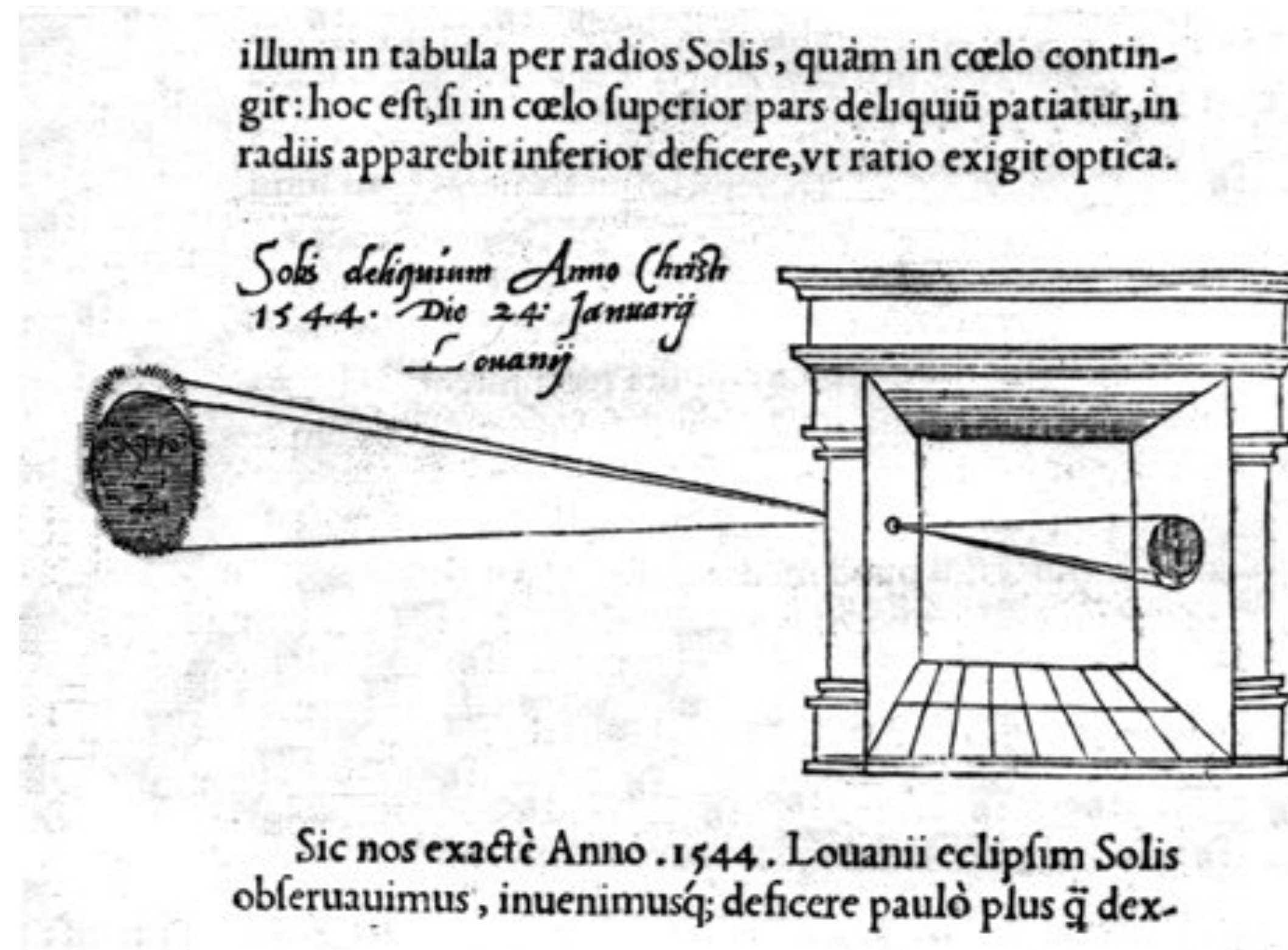
real-world
object



digital sensor
(CCD or
CMOS)

Each scene point contributes to only one sensor pixel

Camera Obscura (latin for “dark chamber”)



Reinerus Gemma-Frisius observed an eclipse of the sun at Louvain on January 24, 1544. He used this illustration in his book, “De Radio Astronomica et Geometrica,” 1545. It is thought to be the first published illustration of a camera obscura.

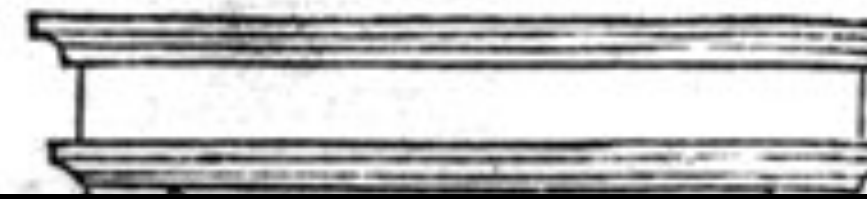
Credit: John H., Hammond, “Th Camera Obscure, A Chronicle”

Camera Obscura (latin for “dark chamber”)



illum in tabula per radios Solis, quam in cælo contin-
git: hoc est, si in cælo superior pars deliquiū patiatur, in
radiis apparebit inferior deficere, vt ratio exigat optica.

*Solis deliquium Anno Christi
1544. Die 24. Januarij
Louanij*



principles behind the pinhole camera or camera obscura were first mentioned by Chinese philosopher Mozi (Mo-Ti) (470 to 390 BCE)



Sic nos exactè Anno .1544. Louanii eclipsim Solis
obseruauimus, inuenimusq; deficere paulò plus q̄ dex-

Reinerus Gemma-Frisius observed an eclipse of the sun at Louvain on January 24, 1544. He used this illustration in his book, “De Radio Astronomica et Geometrica,” 1545. It is thought to be the first published illustration of a camera obscura.

Credit: John H., Hammond, “Th Camera Obscure, A Chronicle”

First **Photograph** on Record

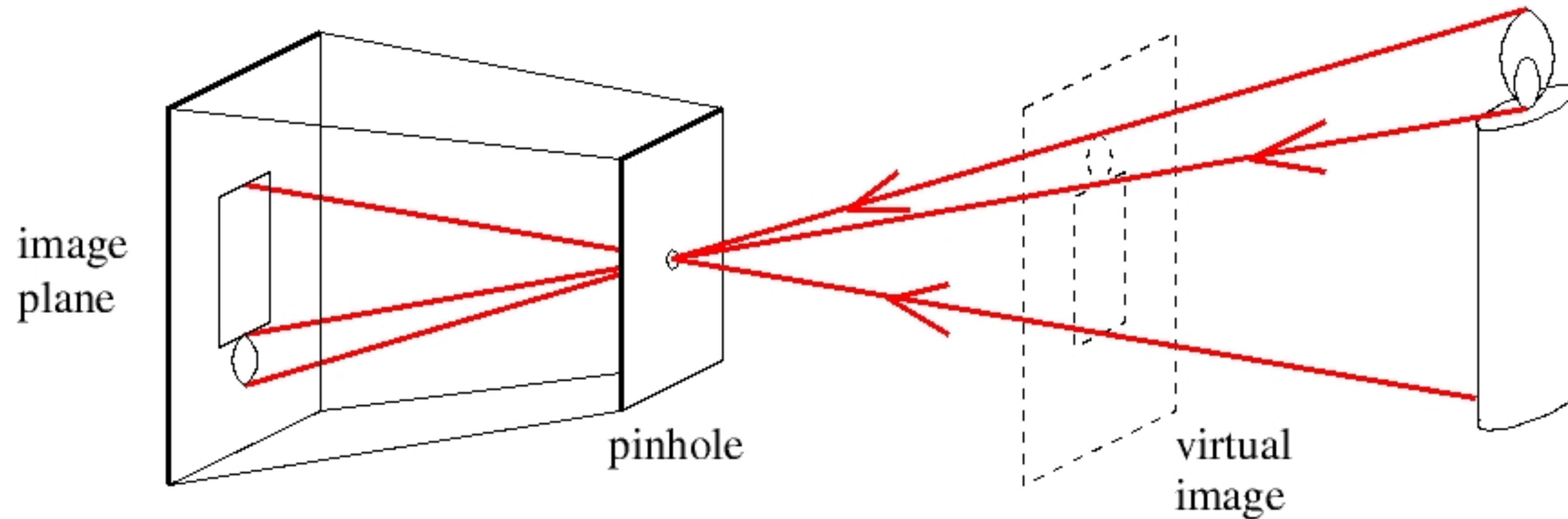
La table servie



Credit: Nicéphore Niepce, 1822

Pinhole Camera

A pinhole camera is a box with a small hole (**aperture**) in it



Forsyth & Ponce (2nd ed.) Figure 1.2

Image Formation



Forsyth & Ponce (2nd ed.) Figure 1.1

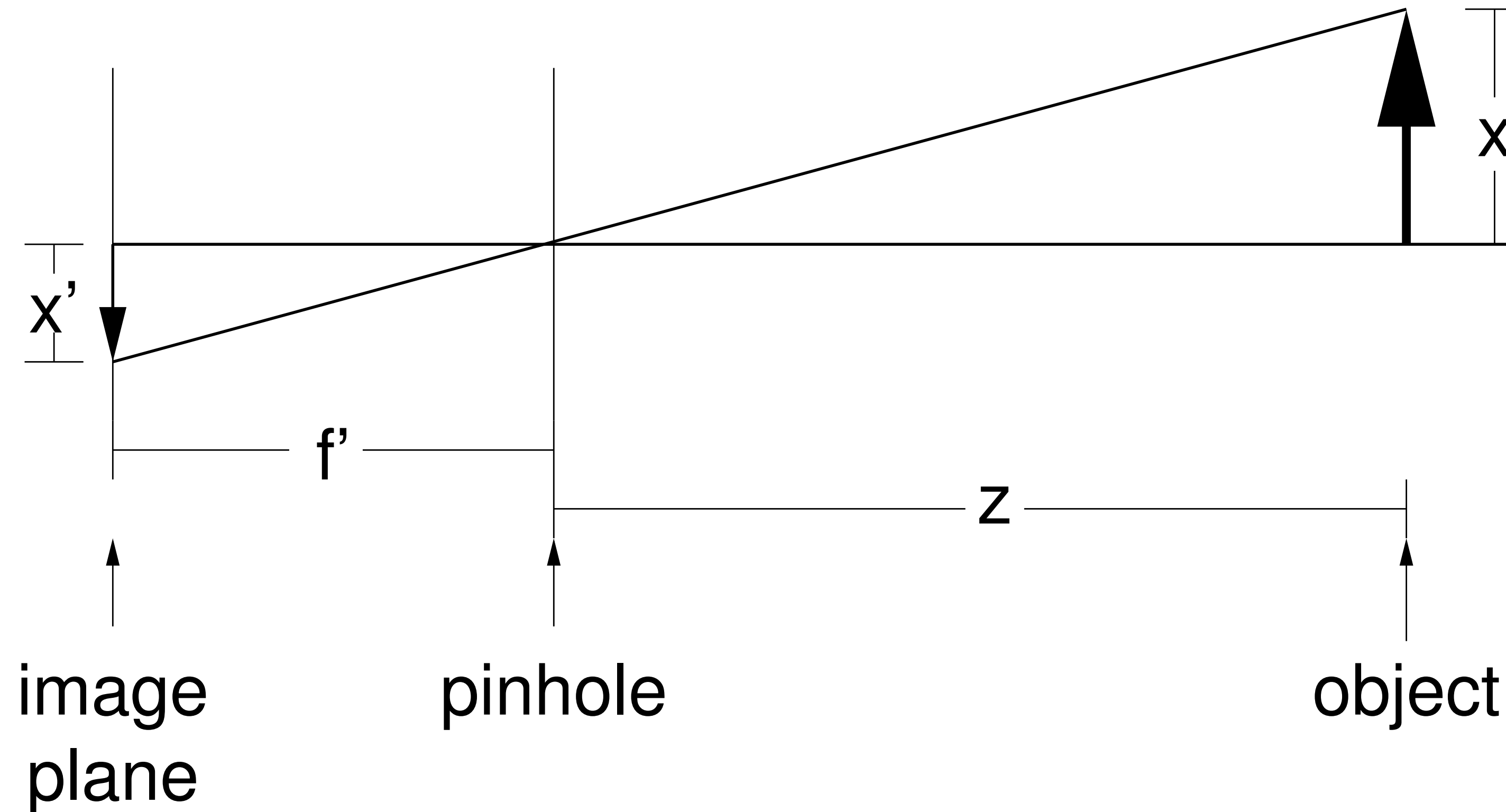
Accidental Pinhole Camera



Image Credit: Ioannis (Yannis) Gkioulekas (CMU)

Pinhole Camera (Simplified)

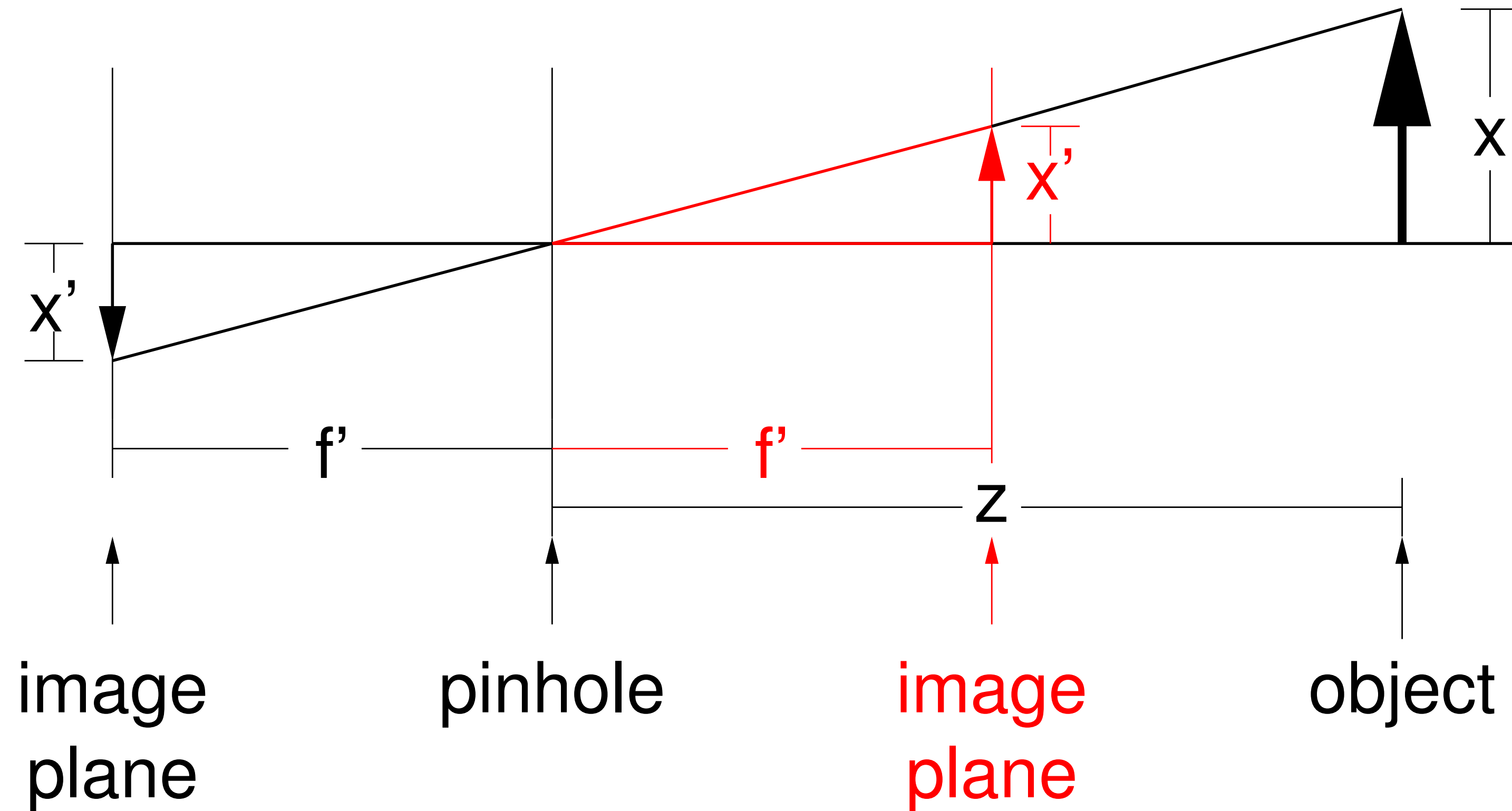
f' is the **focal length** of the camera



Note: In a pinhole camera we can adjust the focal length, all this will do is change the **size** of the resulting image

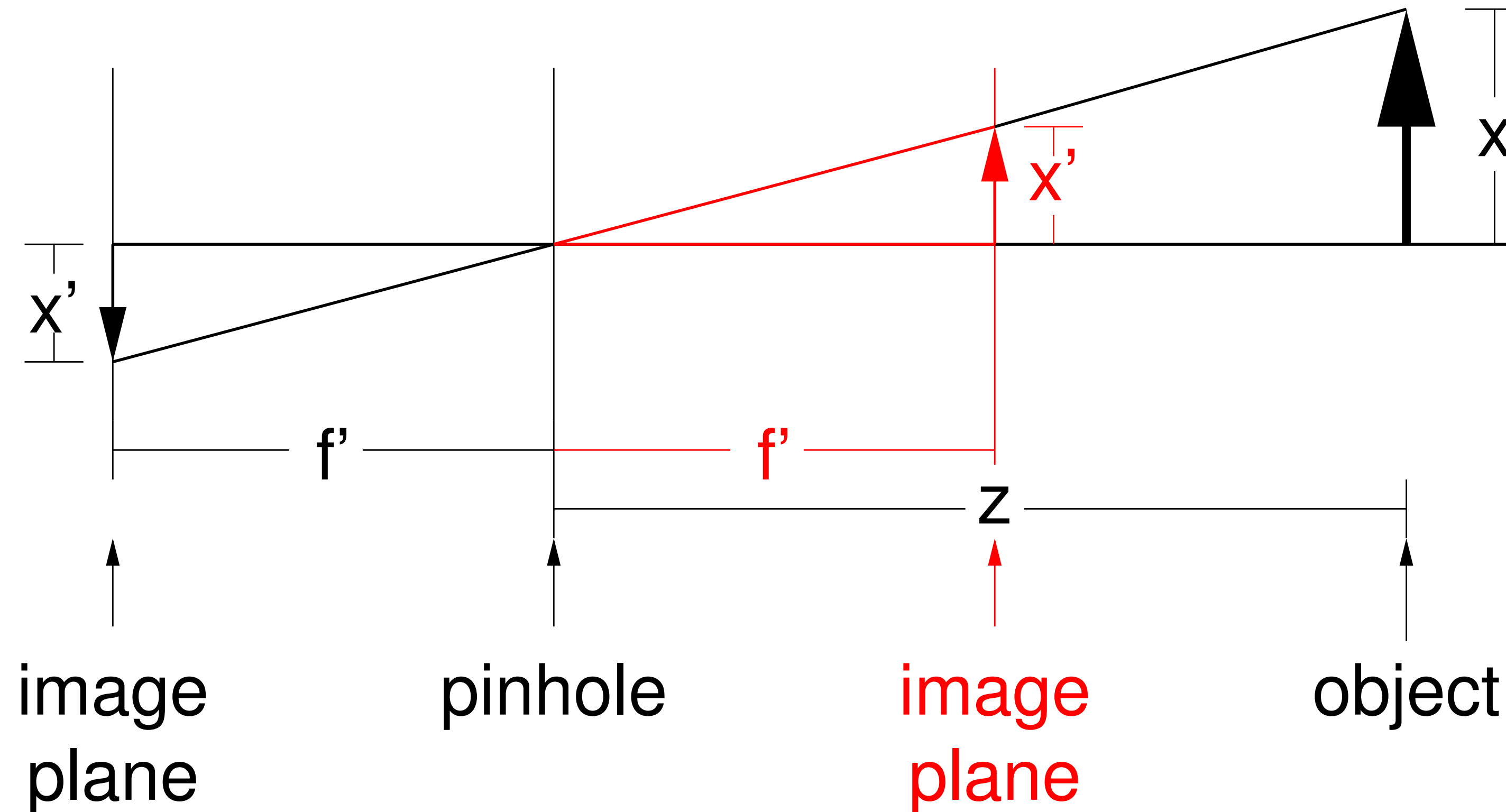
Pinhole Camera (Simplified)

It is convenient to think of the **image plane** which is in front of the pinhole



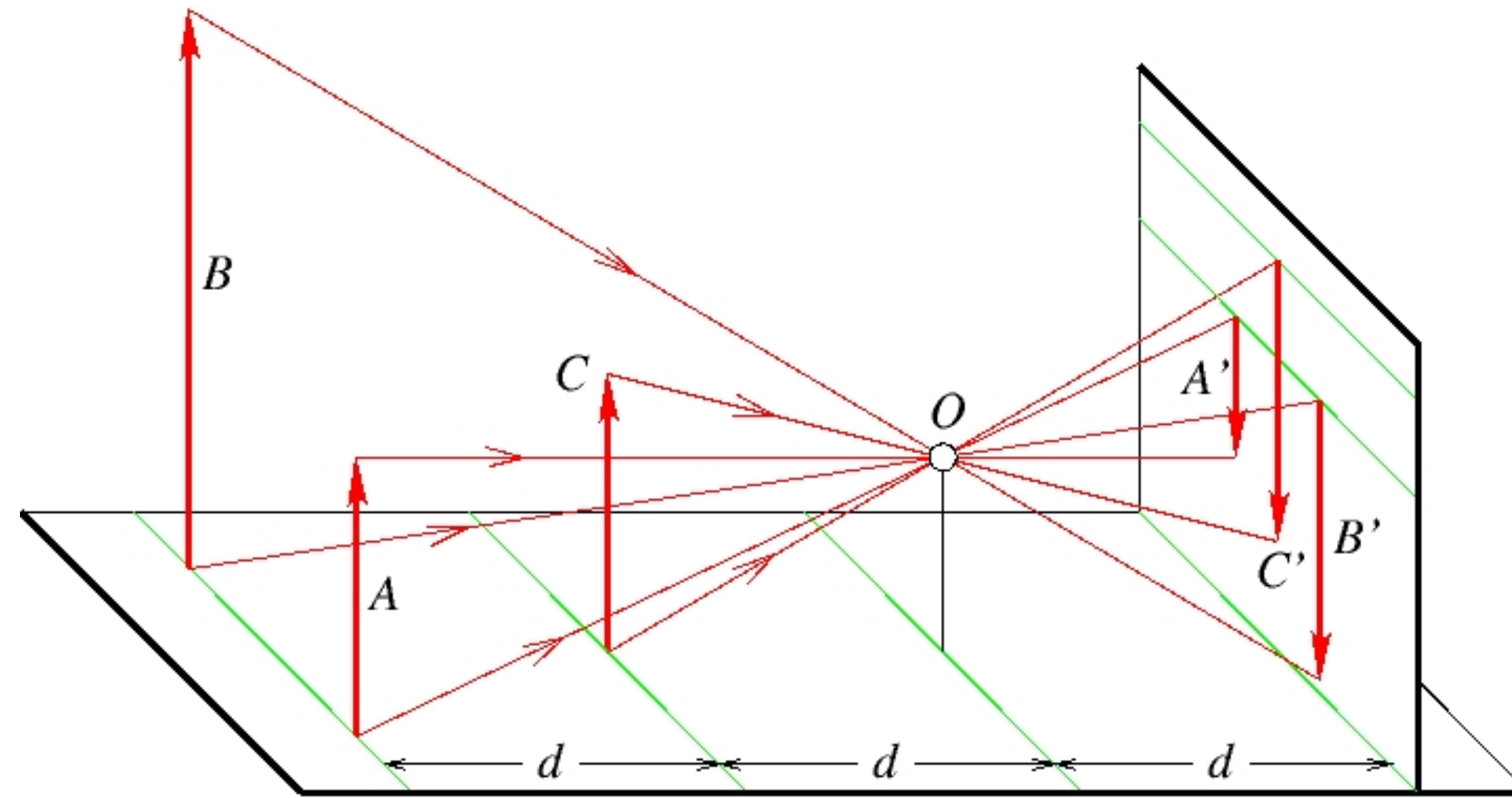
Pinhole Camera (Simplified)

It is convenient to think of the **image plane** which is in front of the pinhole



What happens if object moves towards the camera? Away from the camera?

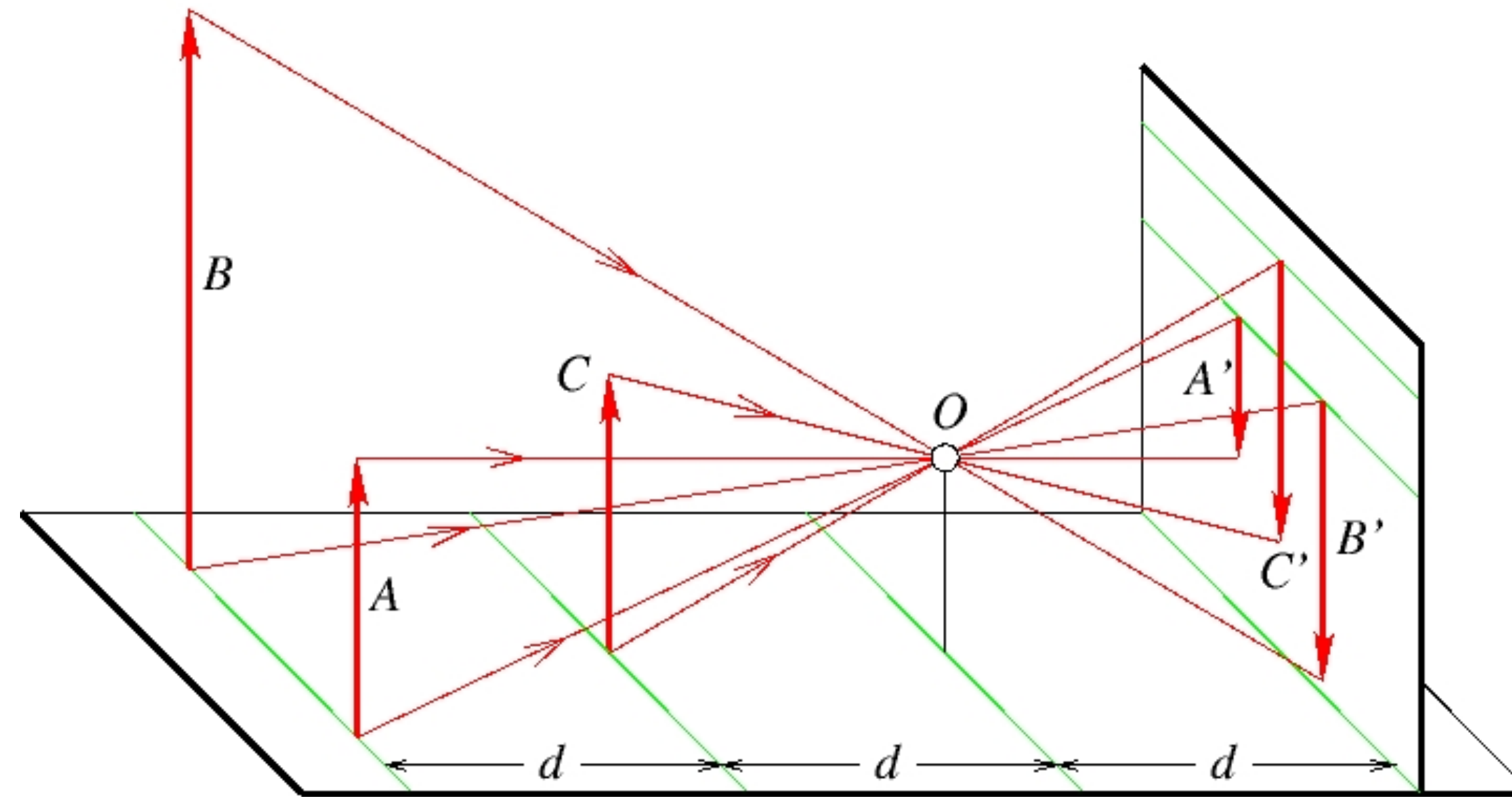
Perspective Effects



Forsyth & Ponce (1st ed.) Figure 1.3a

Perspective Effects

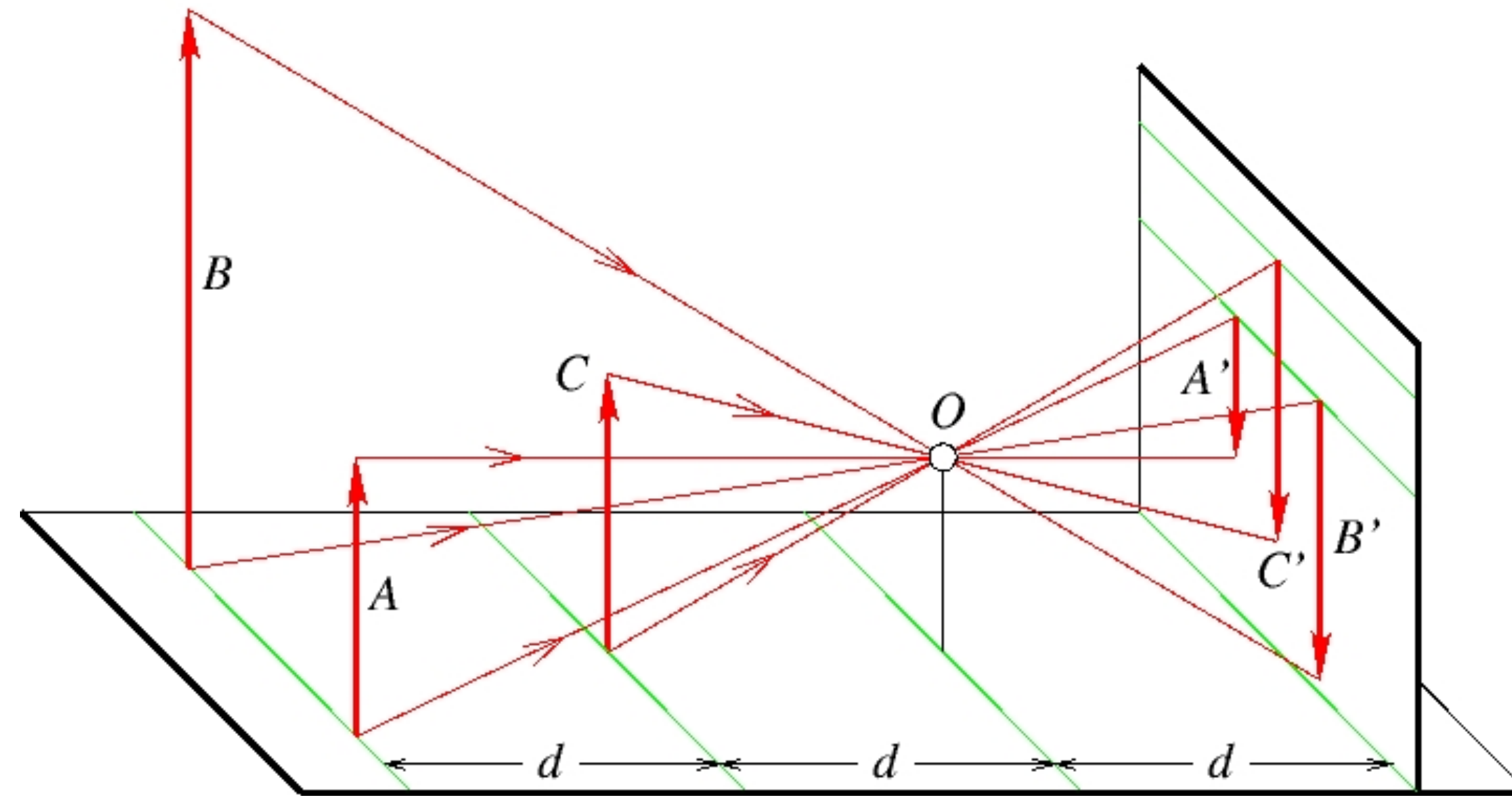
Far objects appear **smaller** than close ones



Forsyth & Ponce (1st ed.) Figure 1.3a

Perspective Effects

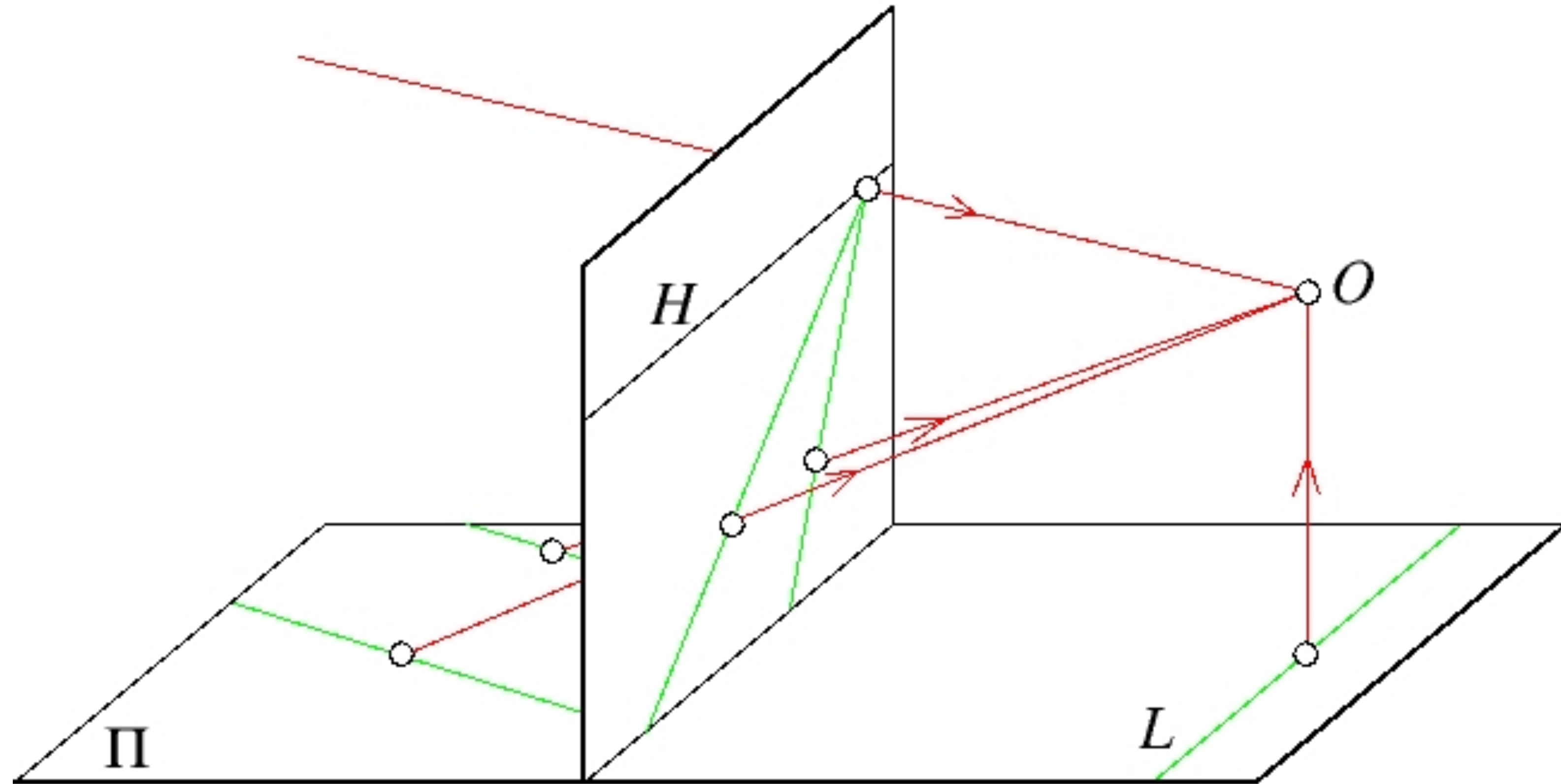
Far objects appear **smaller** than close ones



Forsyth & Ponce (1st ed.) Figure 1.3a

Size is **inversely** proportions to distance

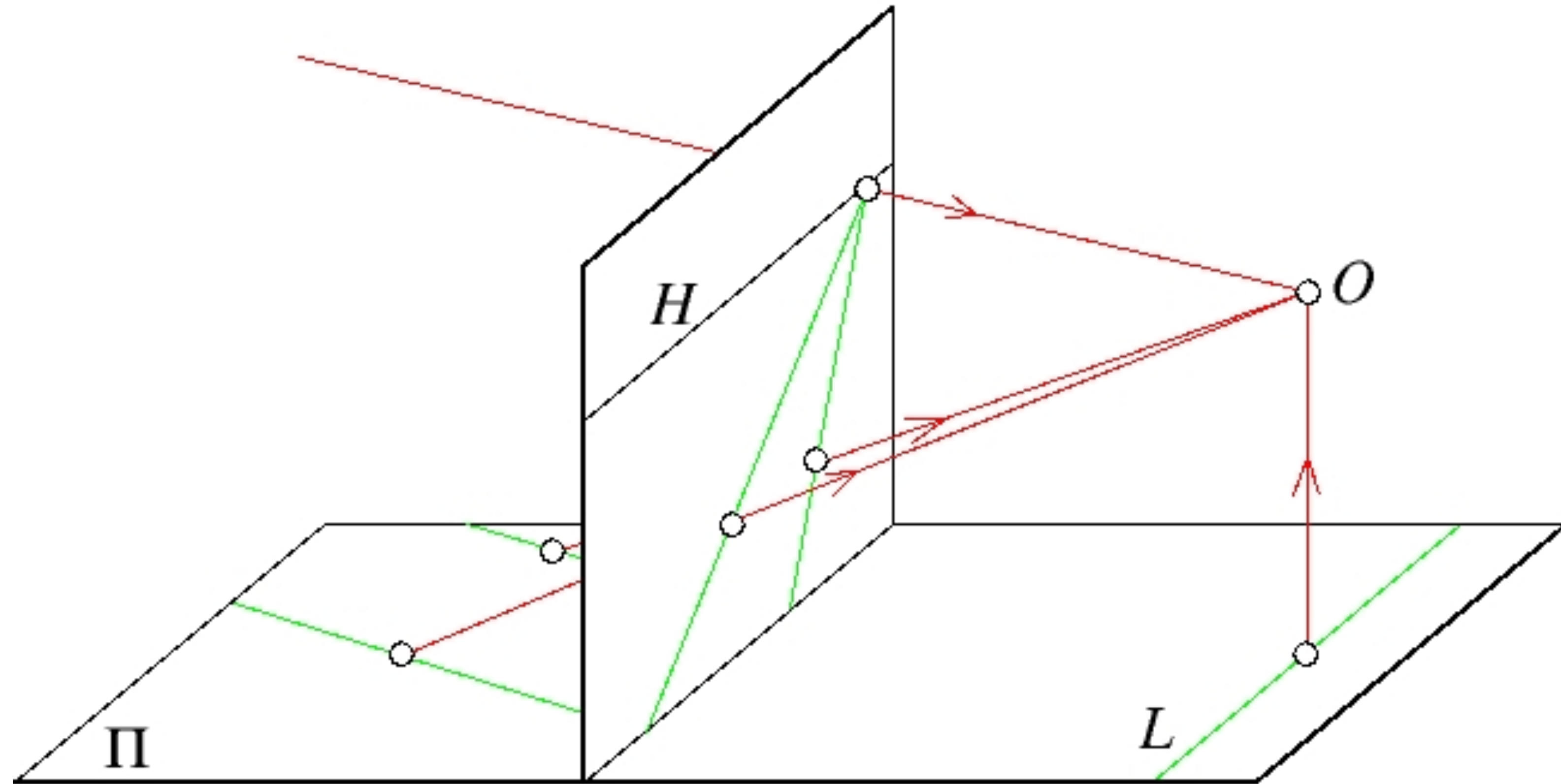
Perspective Effects



Forsyth & Ponce (1st ed.) Figure 1.3b

Perspective Effects

Parallel lines meet at a point (**vanishing point**)



Forsyth & Ponce (1st ed.) Figure 1.3b

Vanishing Points

Each set of parallel lines meet at a different point

— the point is called **vanishing point**

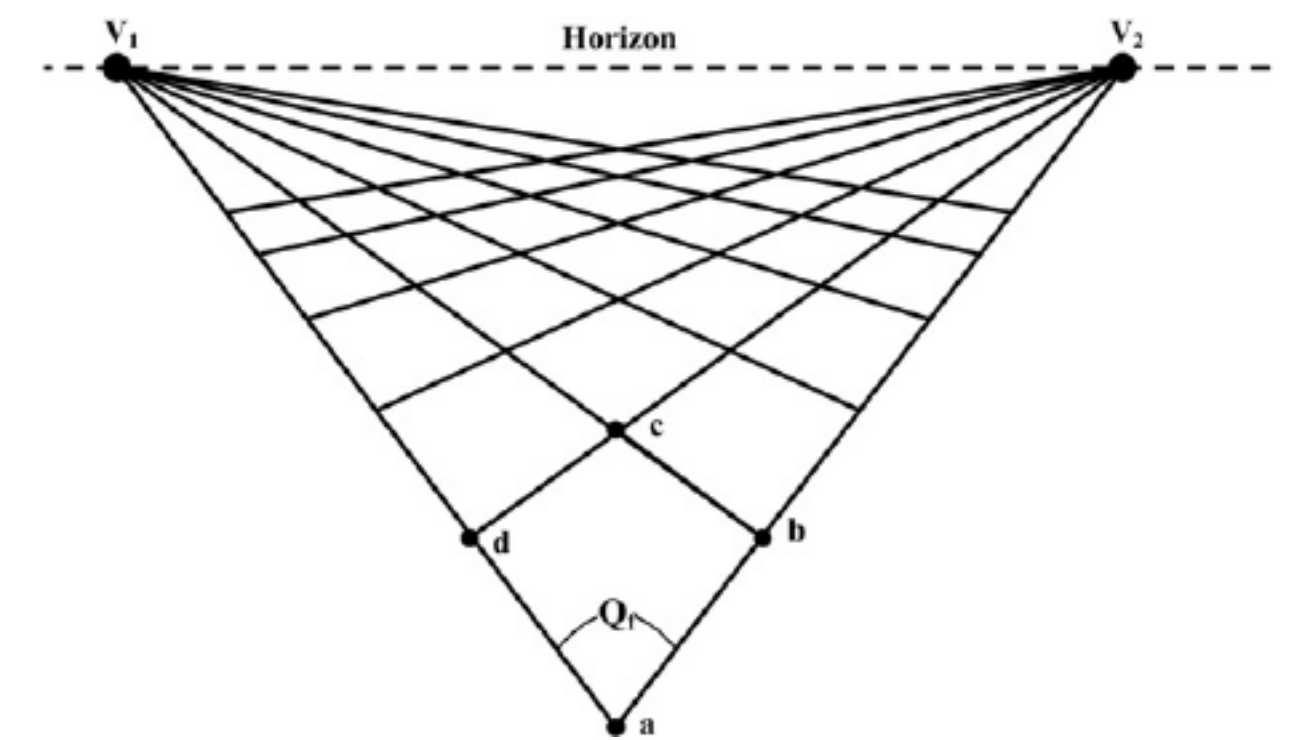
Vanishing Points

Each set of parallel lines meet at a different point

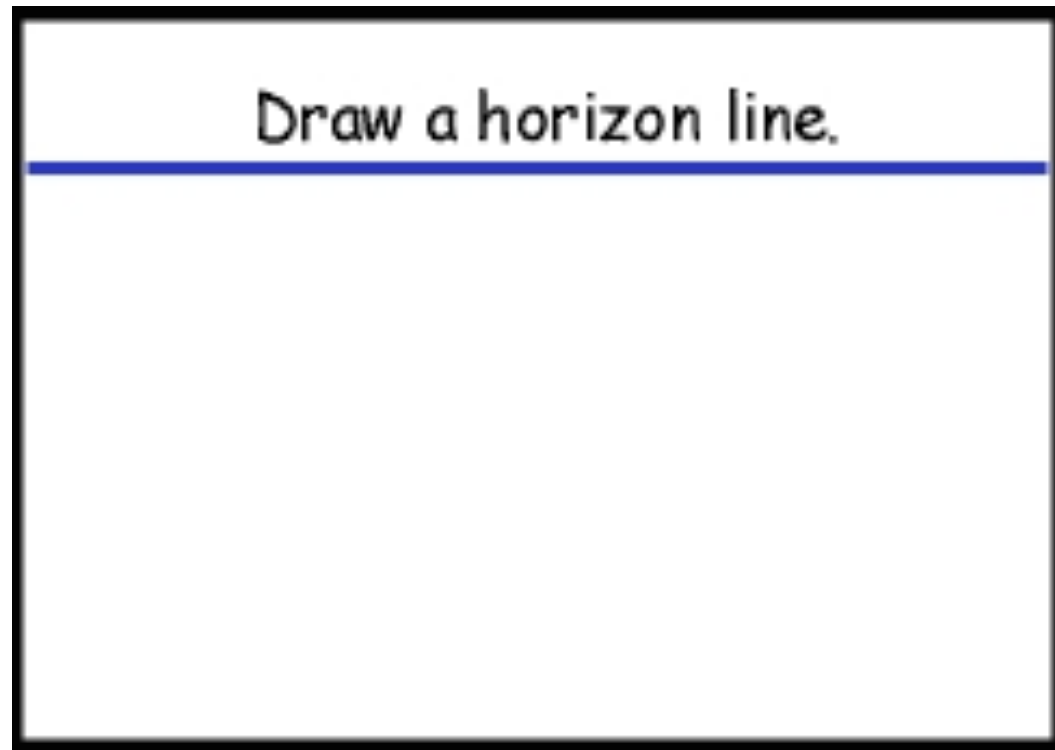
— the point is called **vanishing point**

Sets of parallel lines on the same plane lead to **collinear** vanishing points

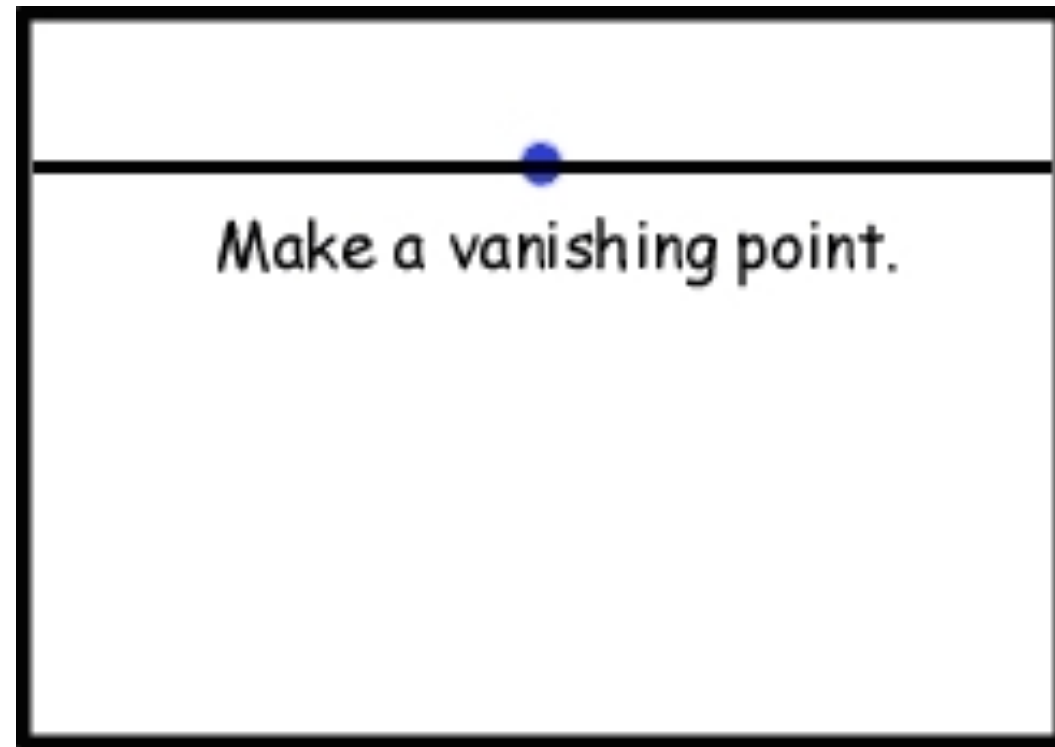
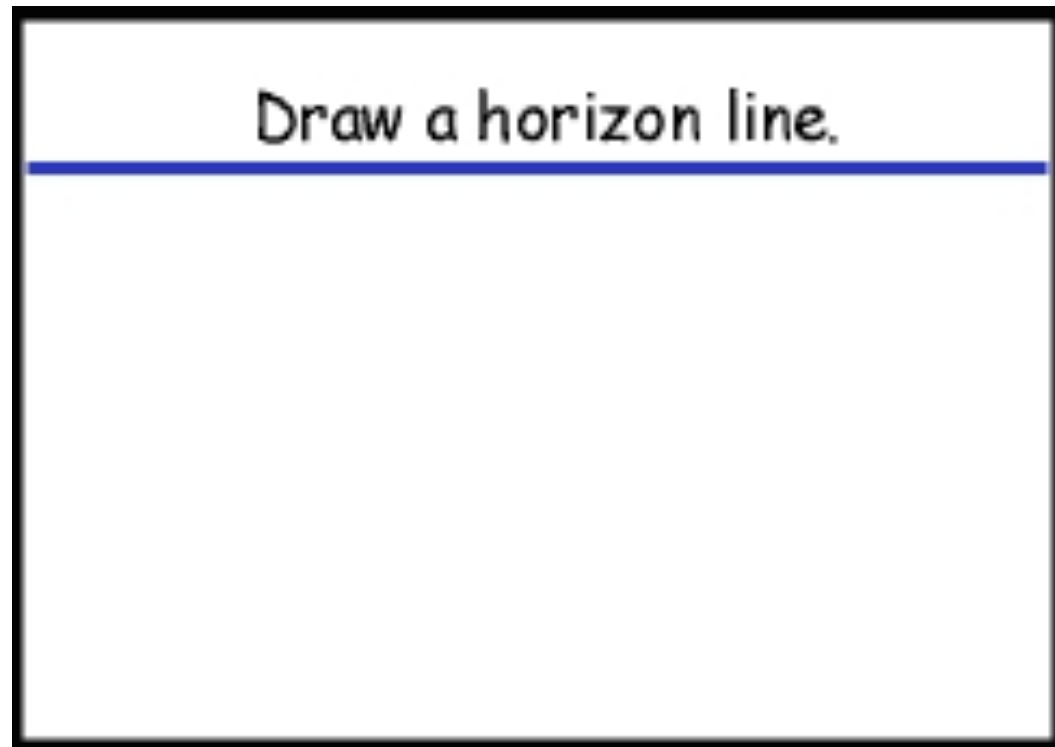
— the line is called a **horizon** for that plane



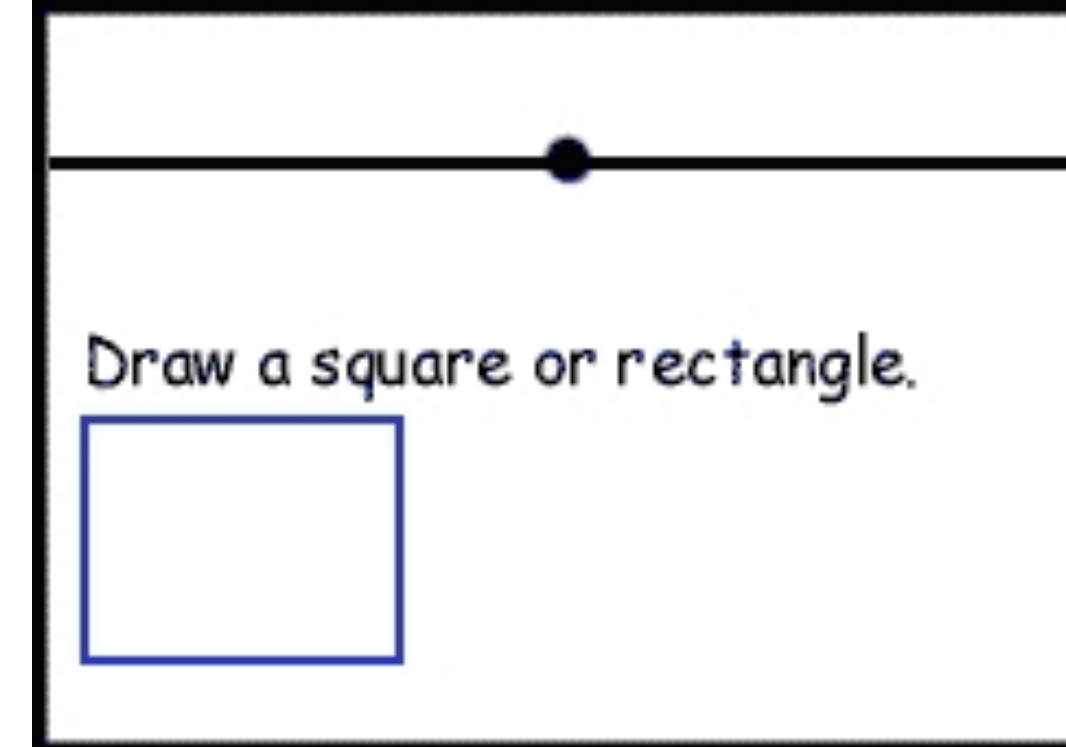
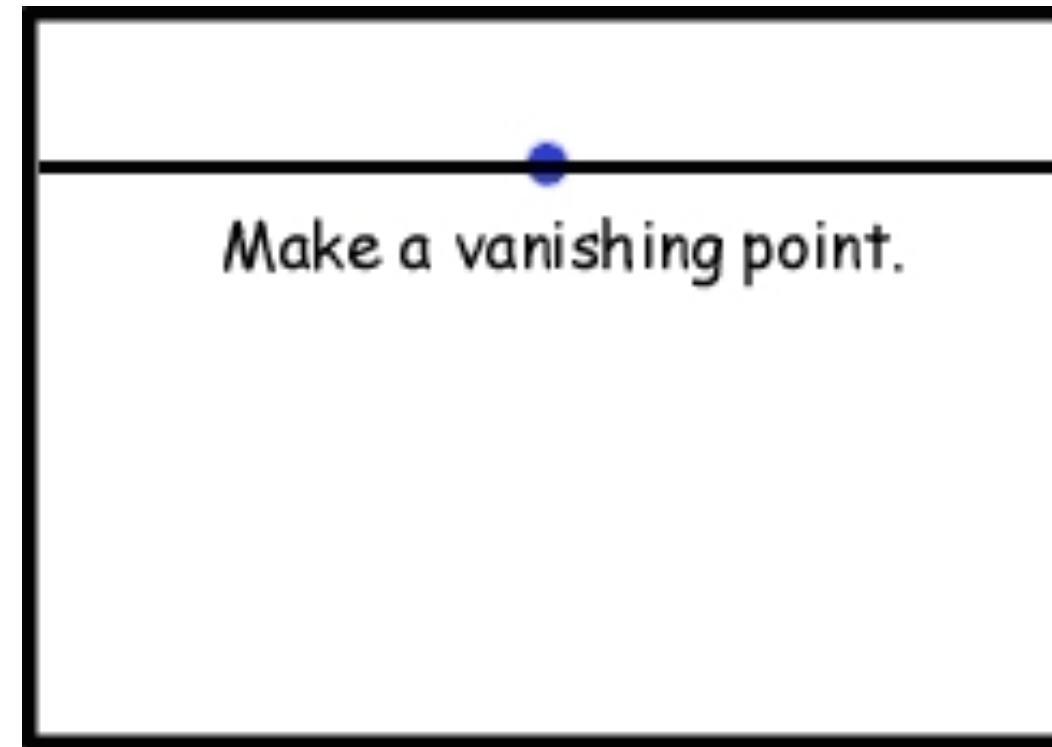
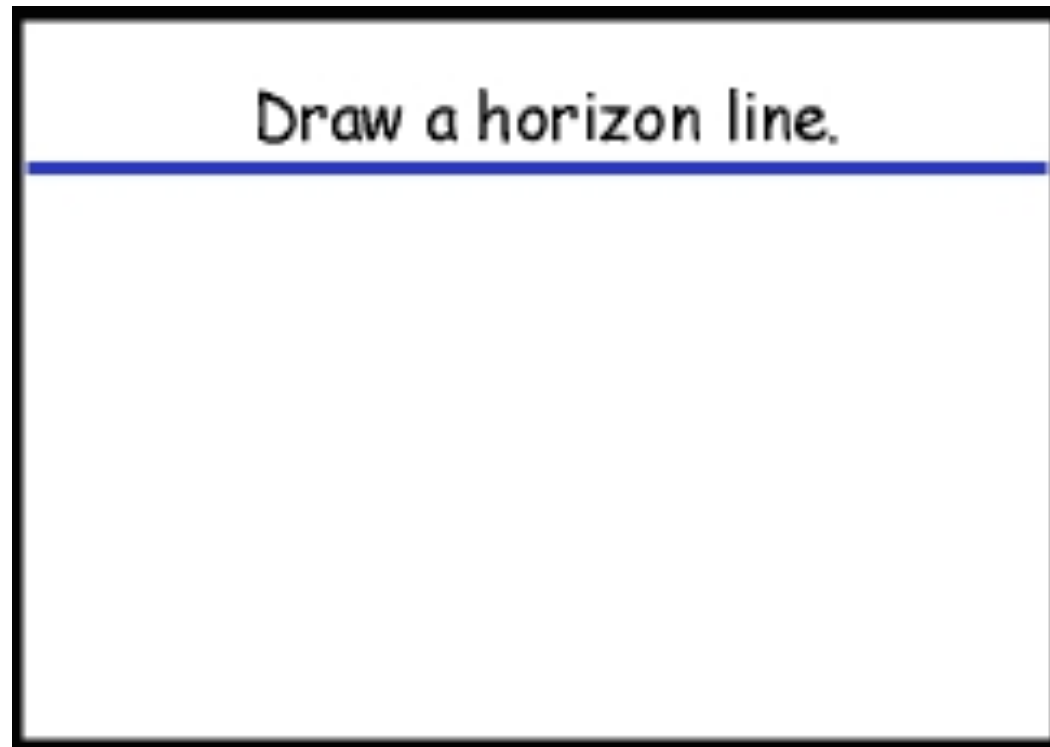
Vanishing Points



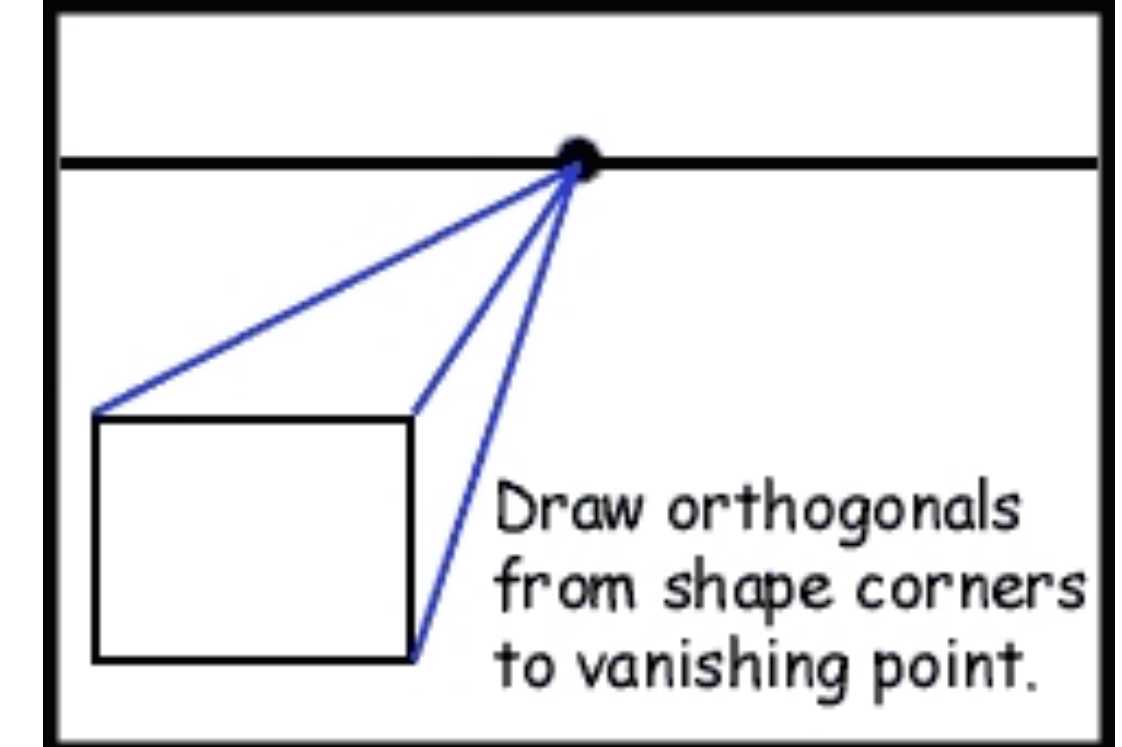
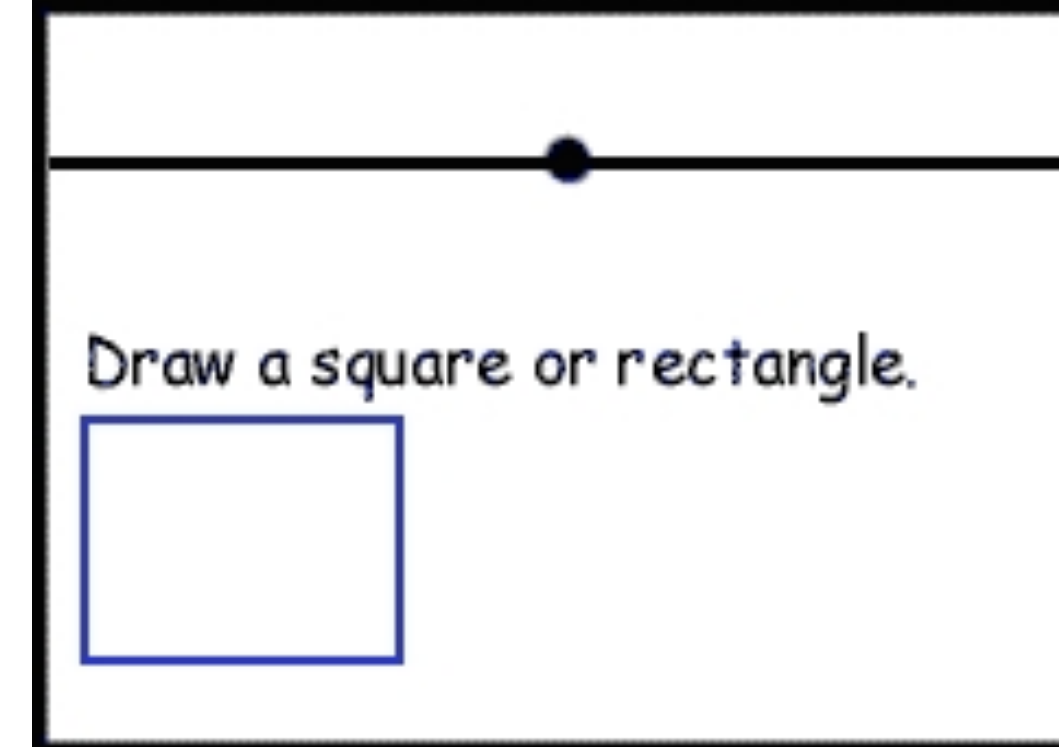
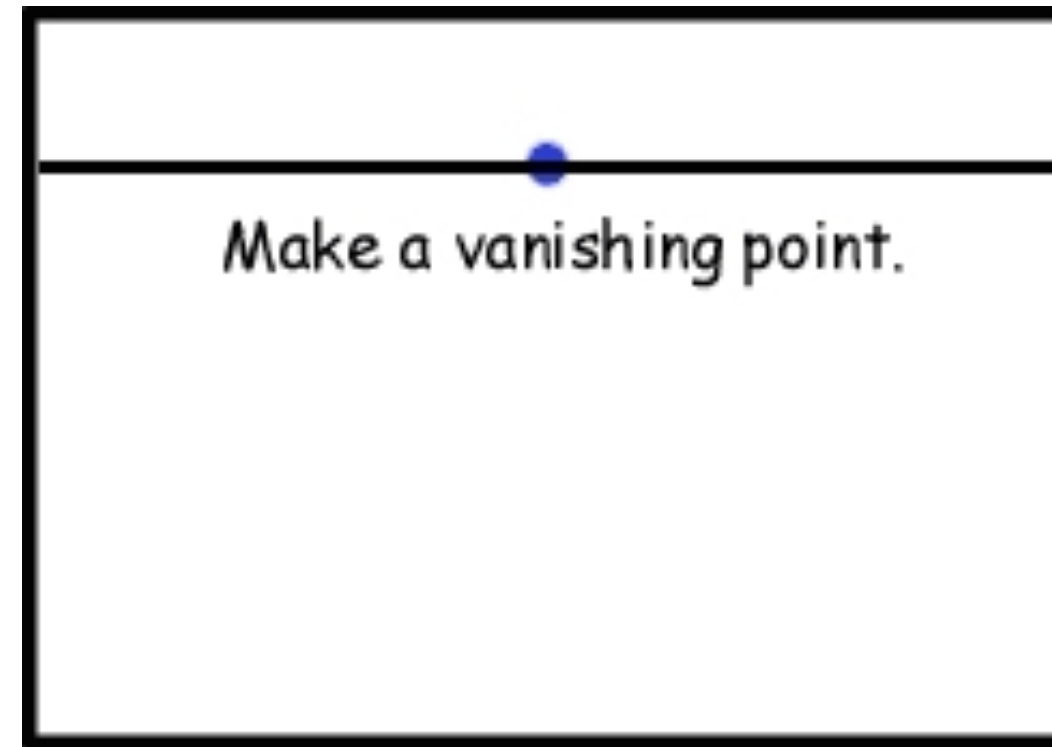
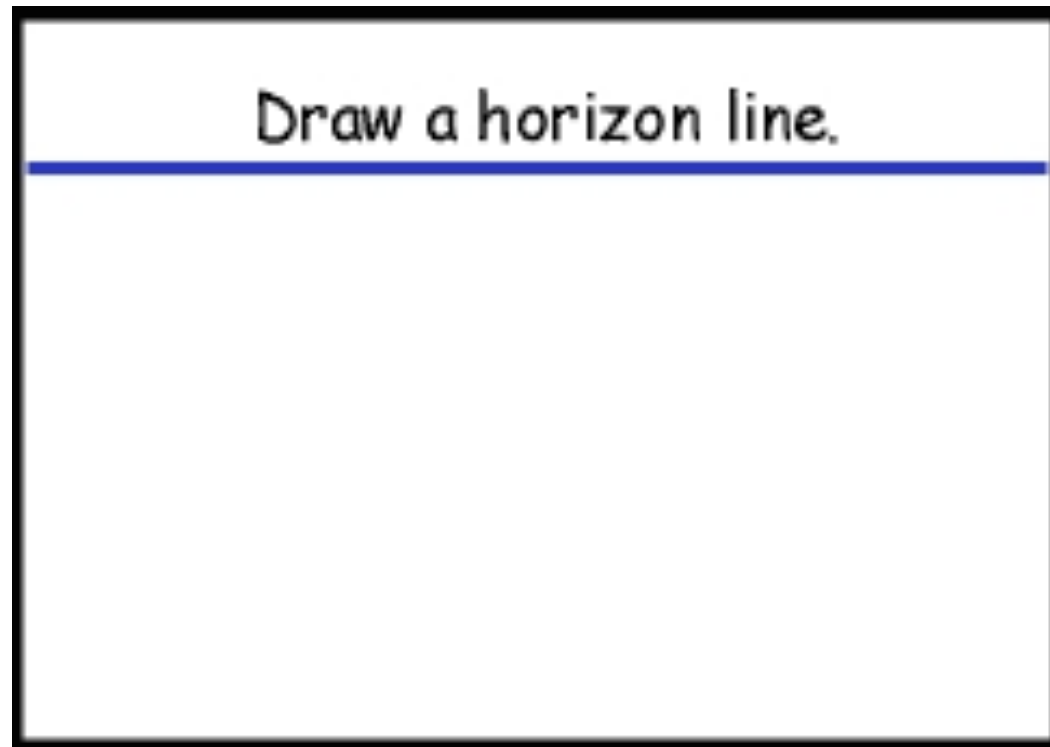
Vanishing Points



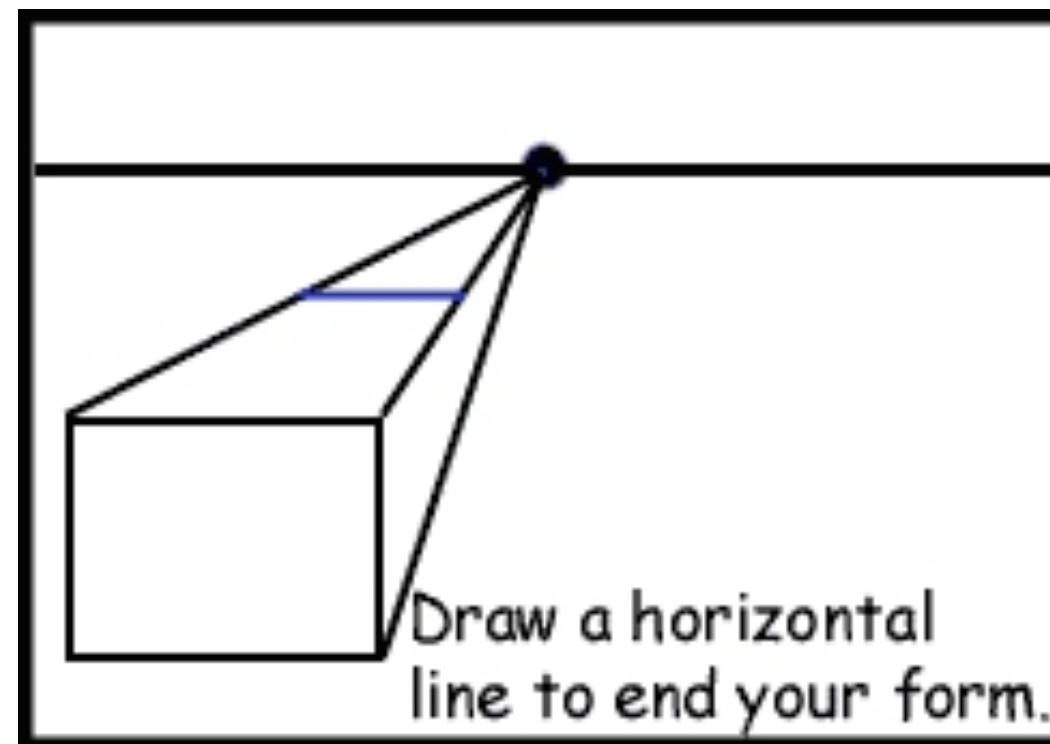
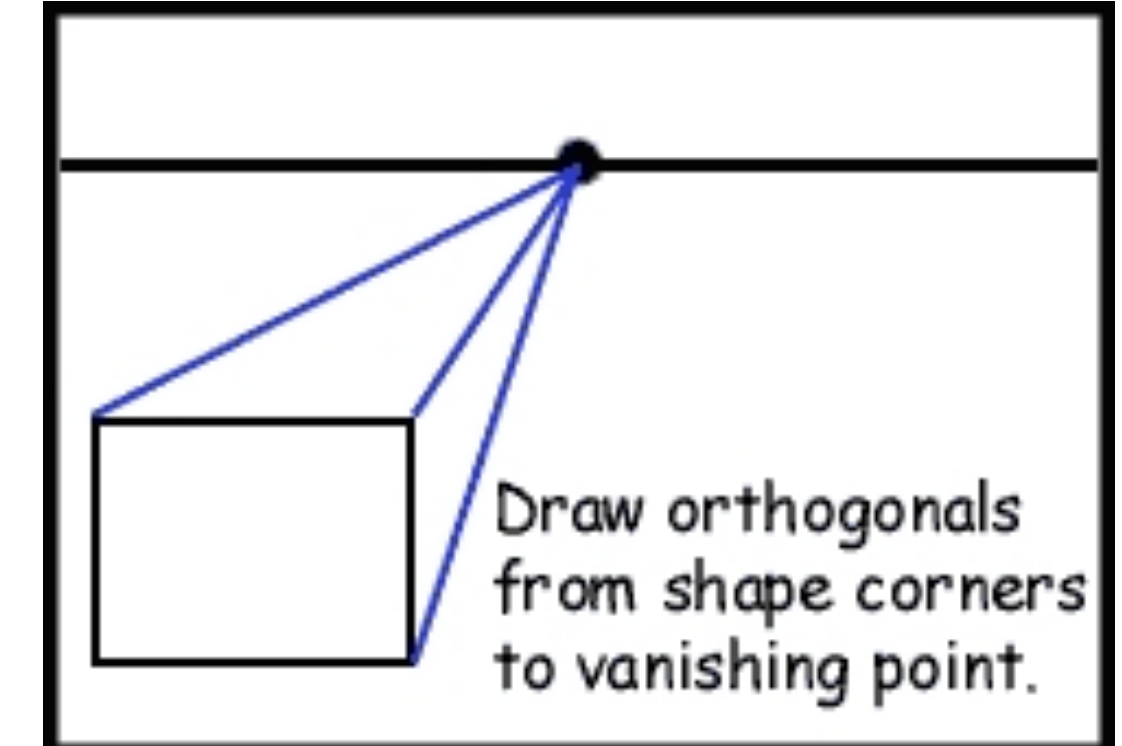
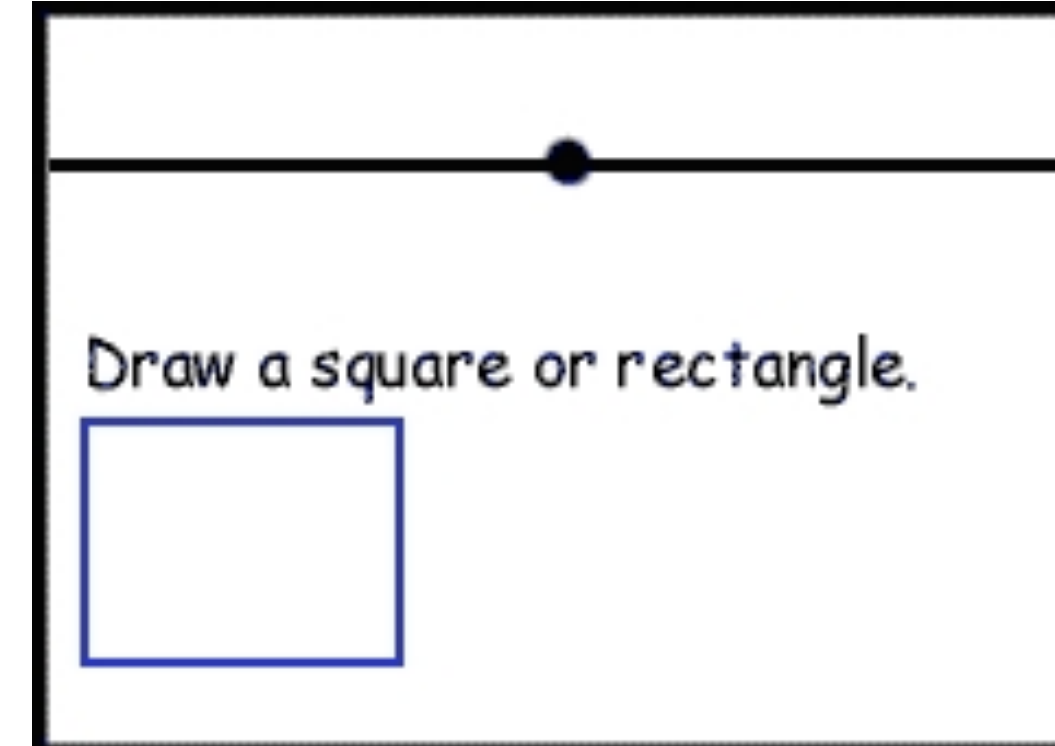
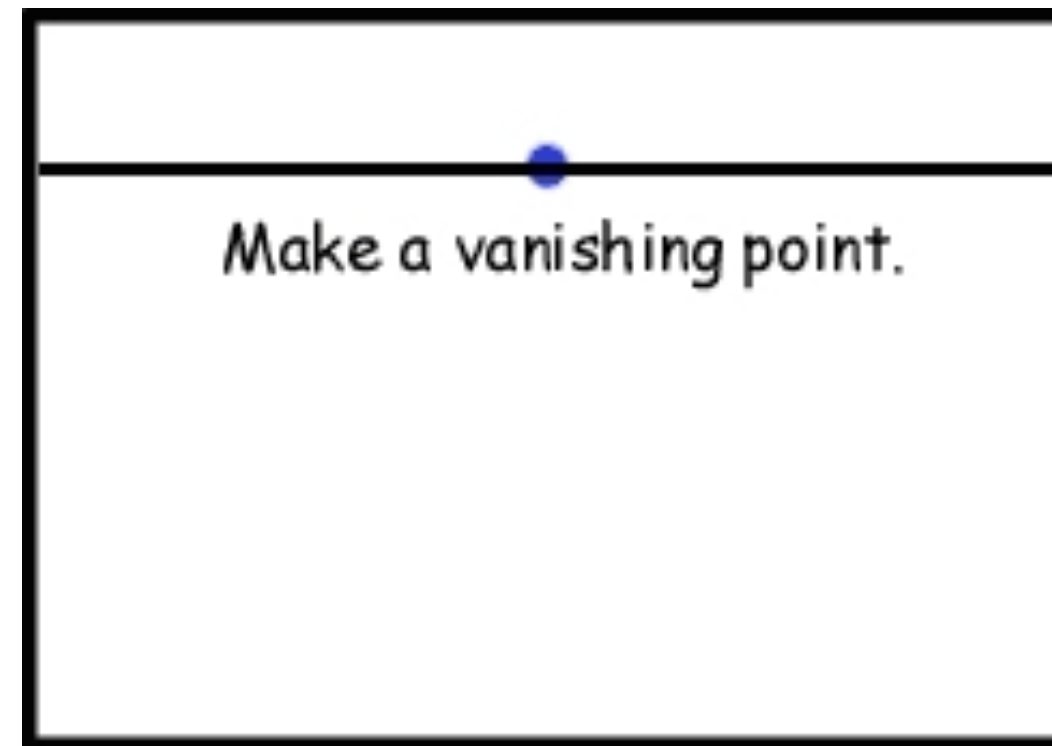
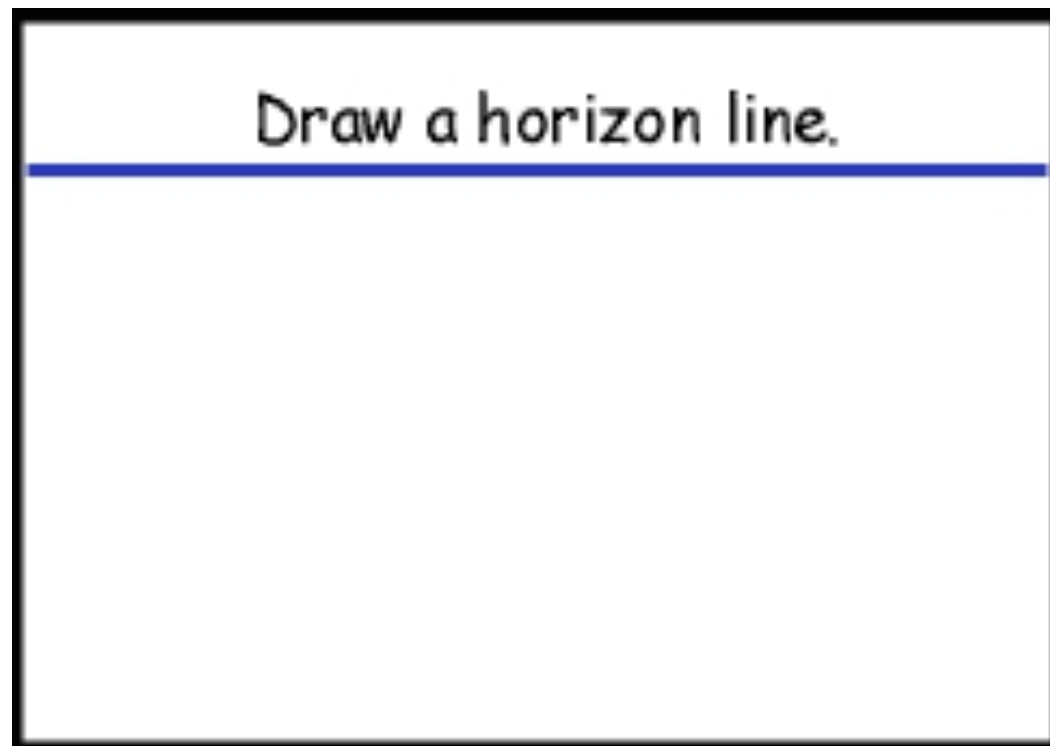
Vanishing Points



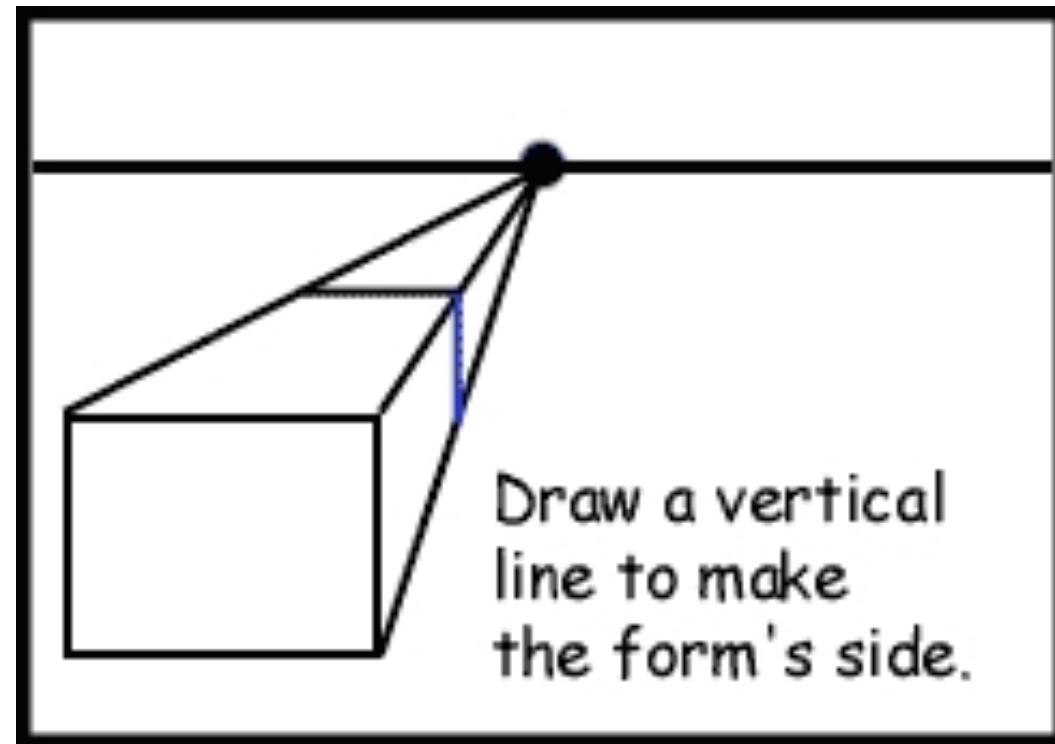
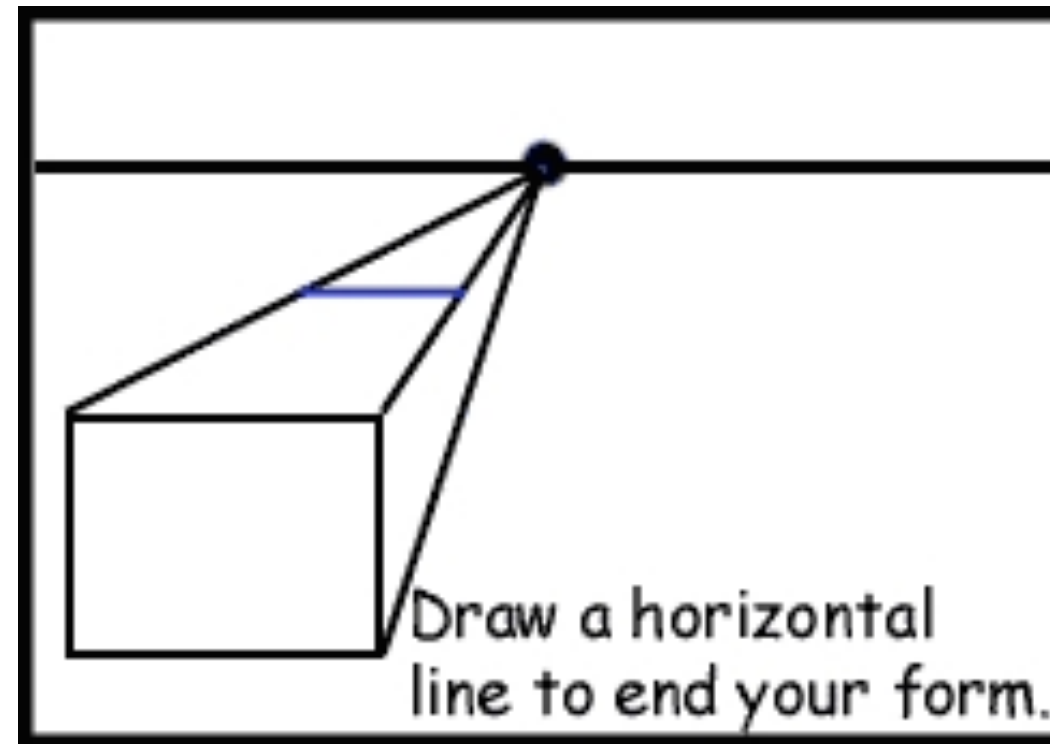
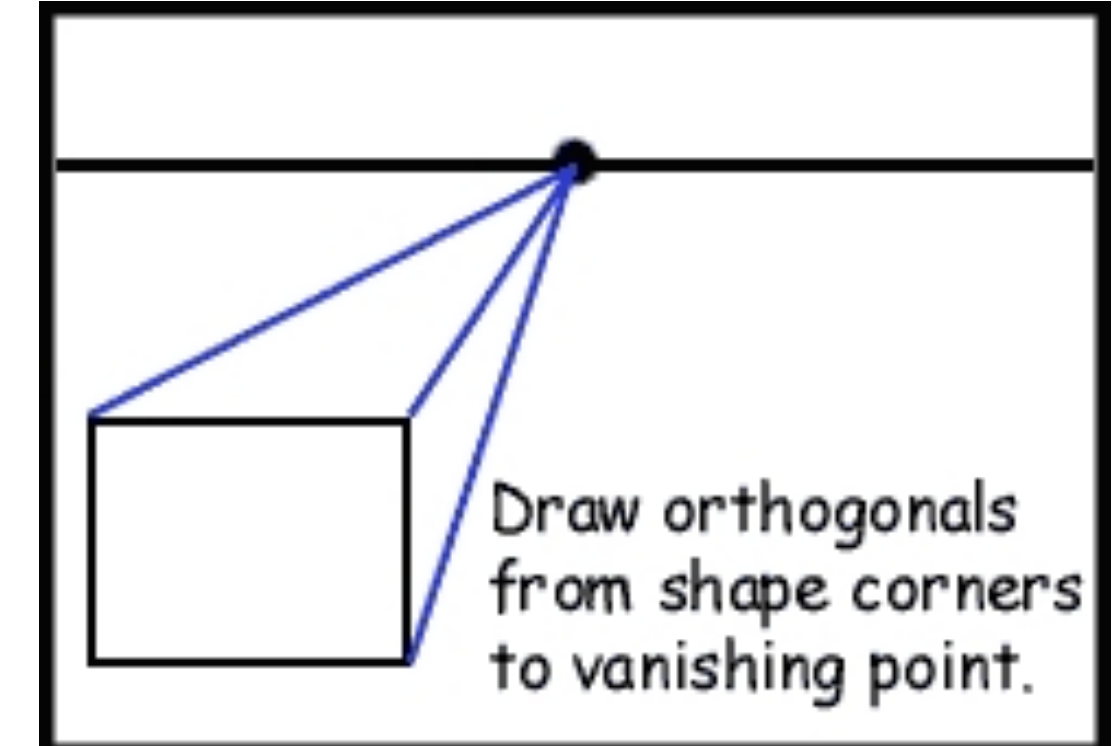
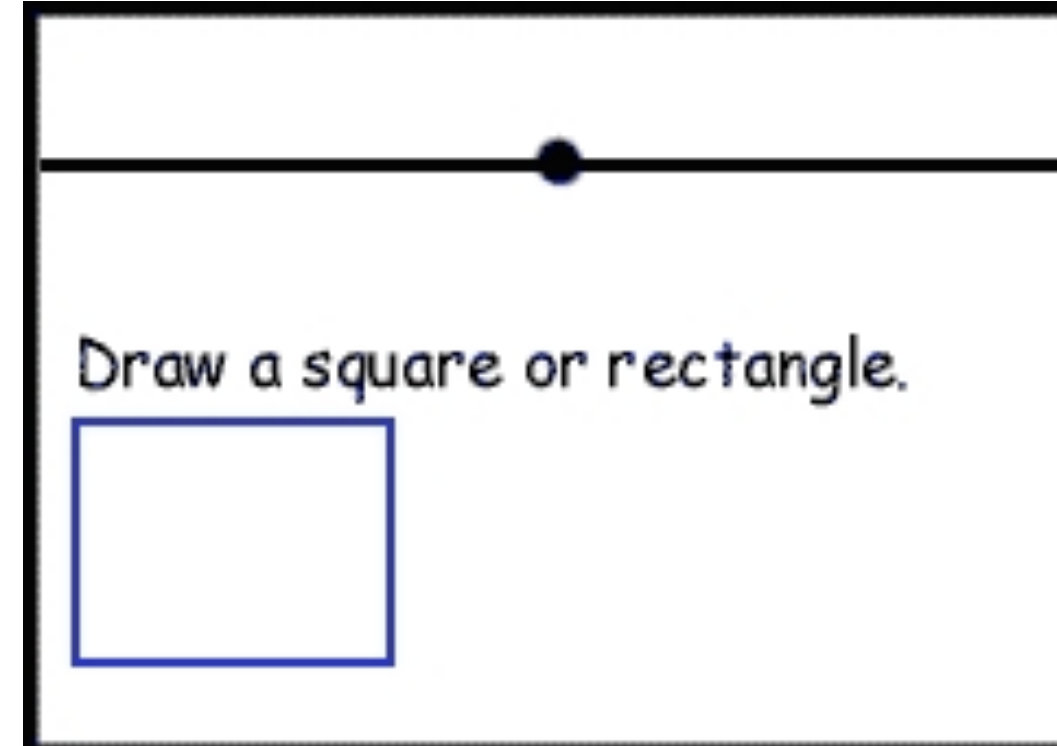
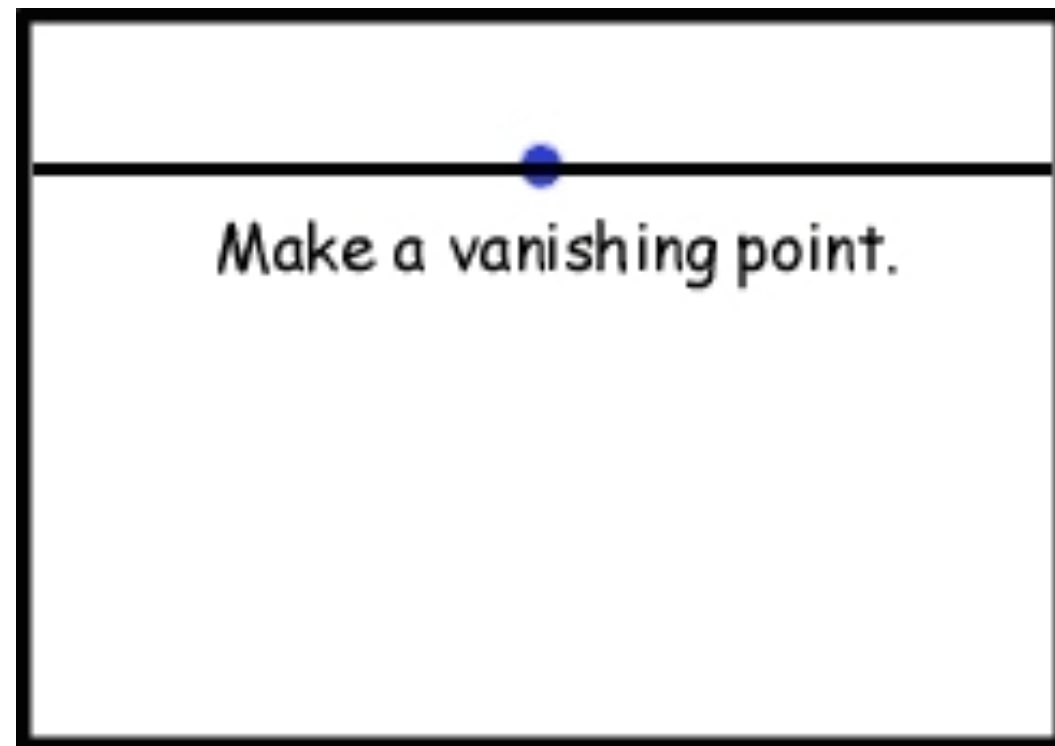
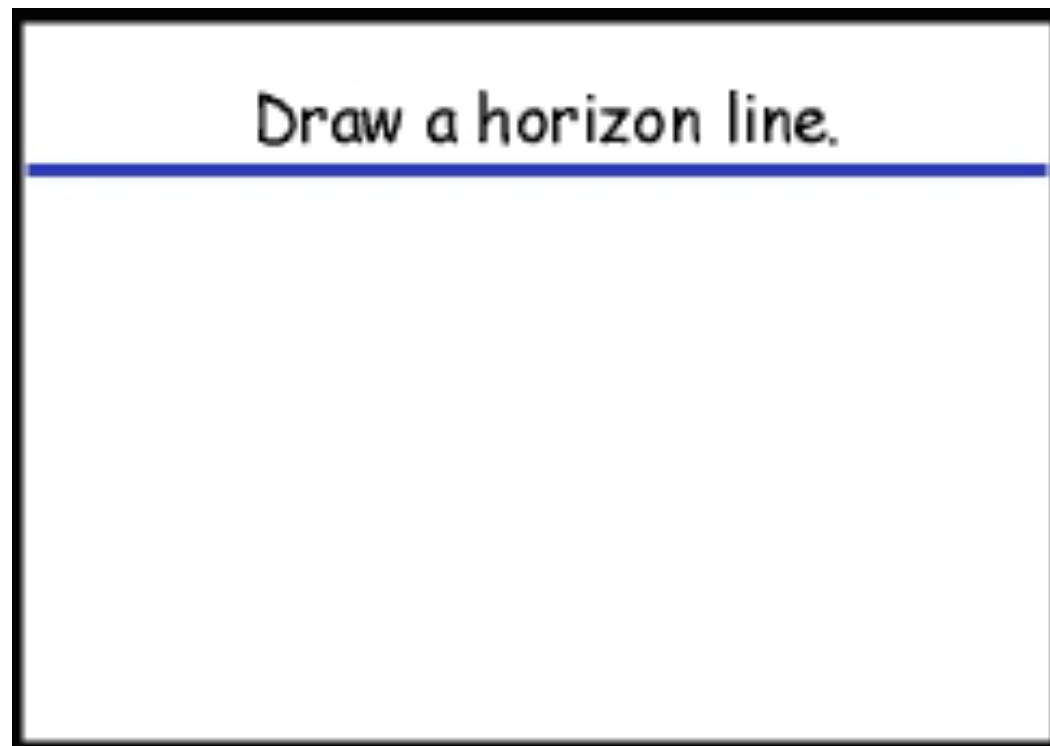
Vanishing Points



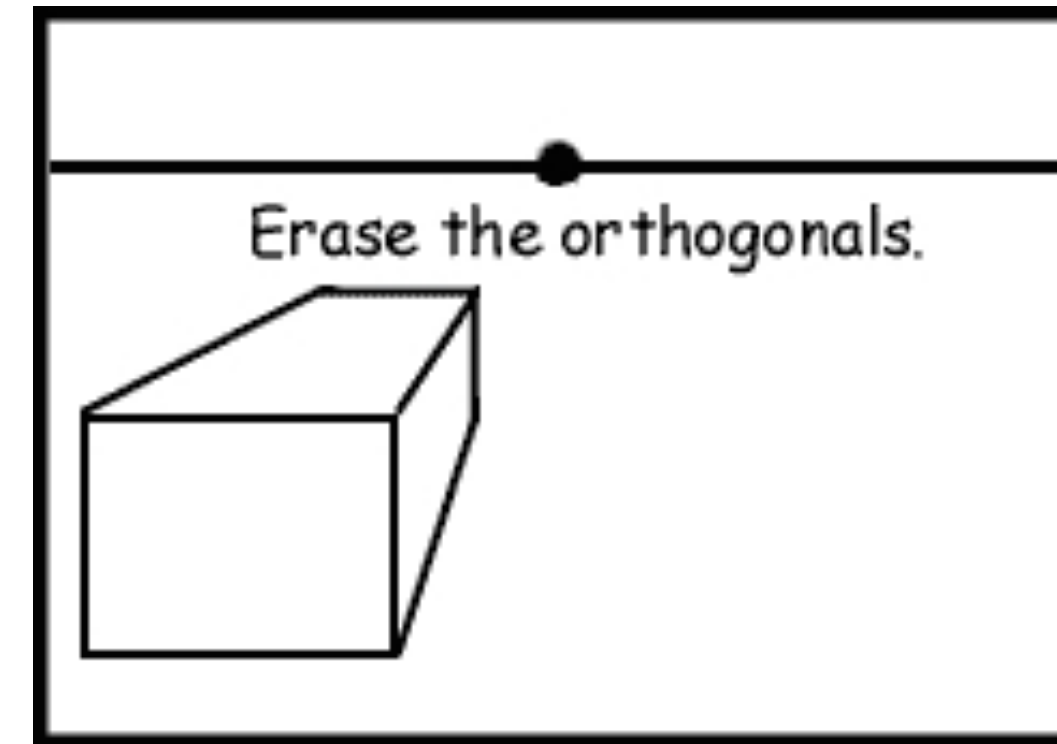
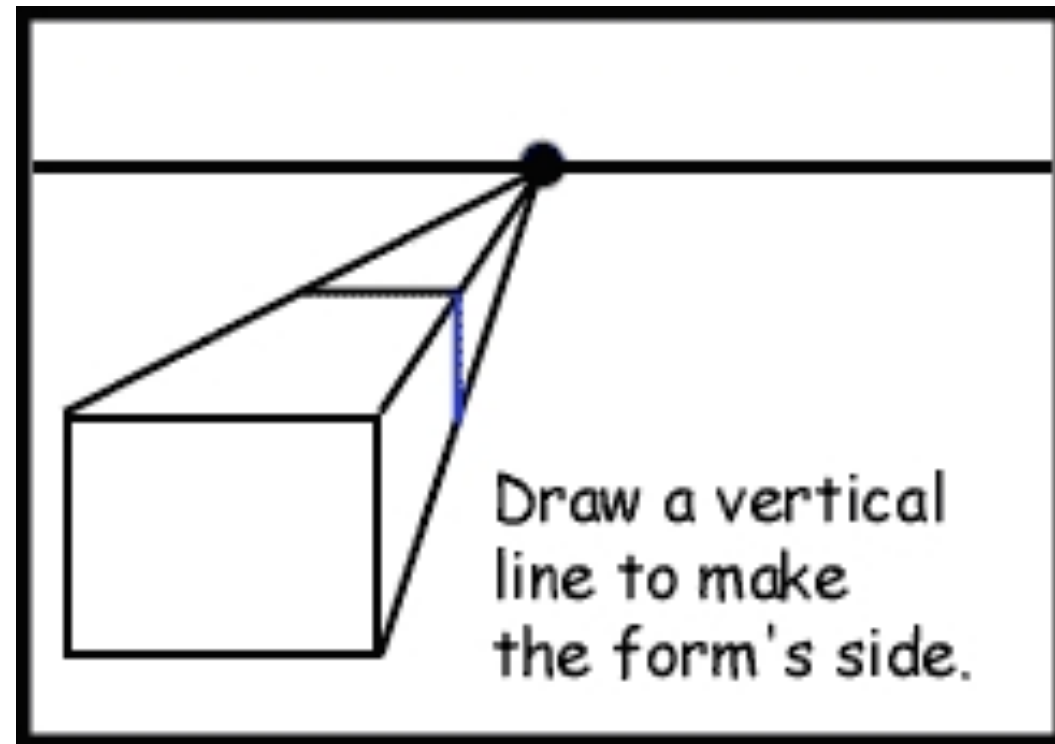
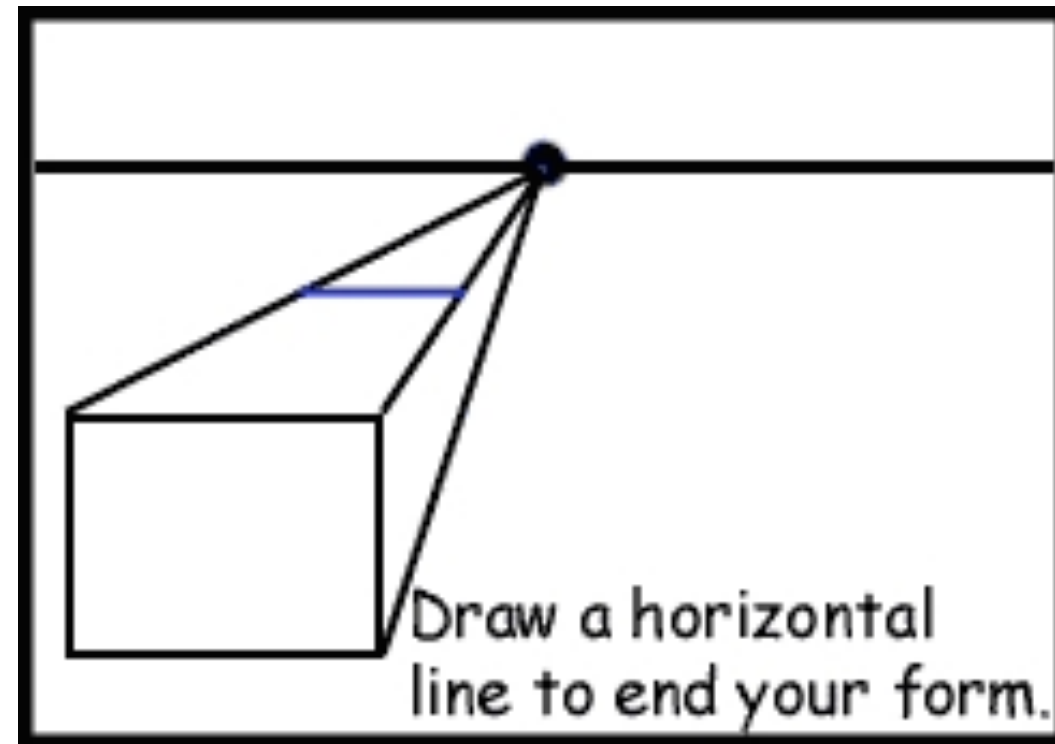
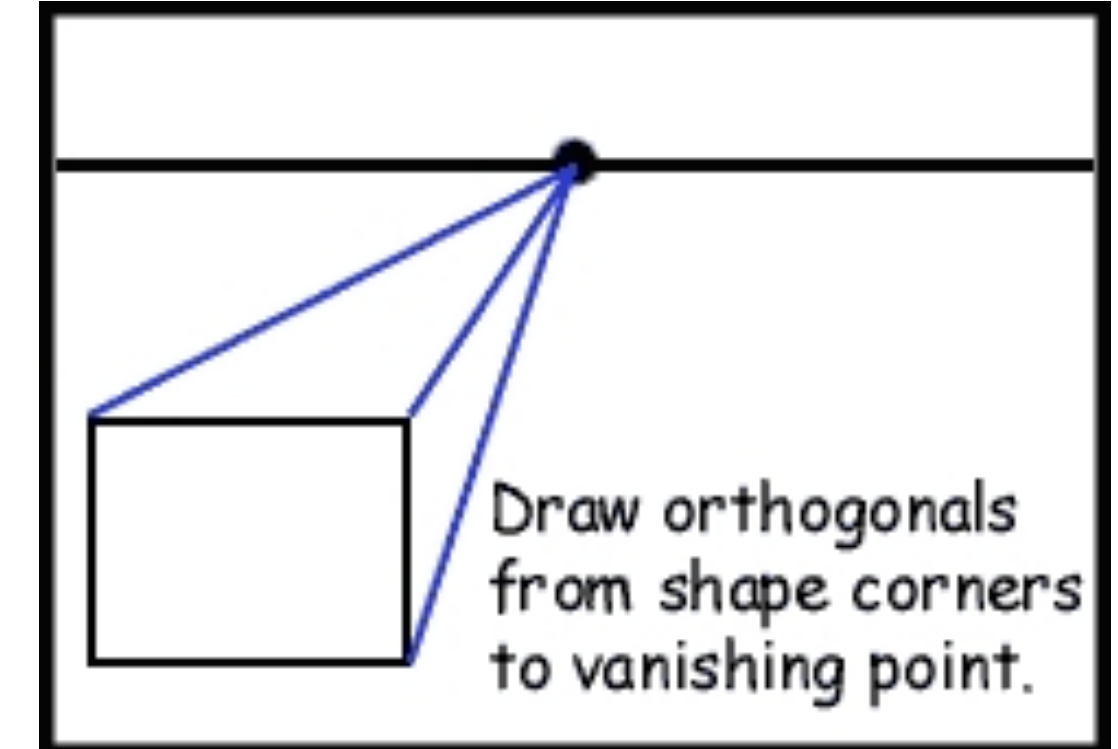
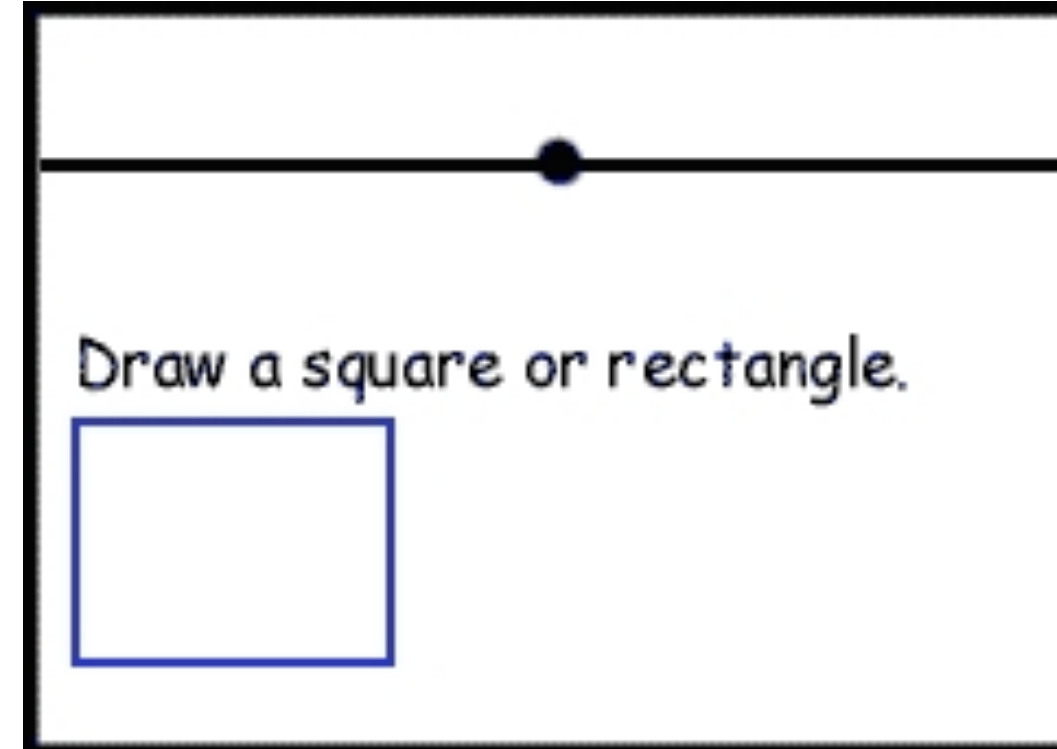
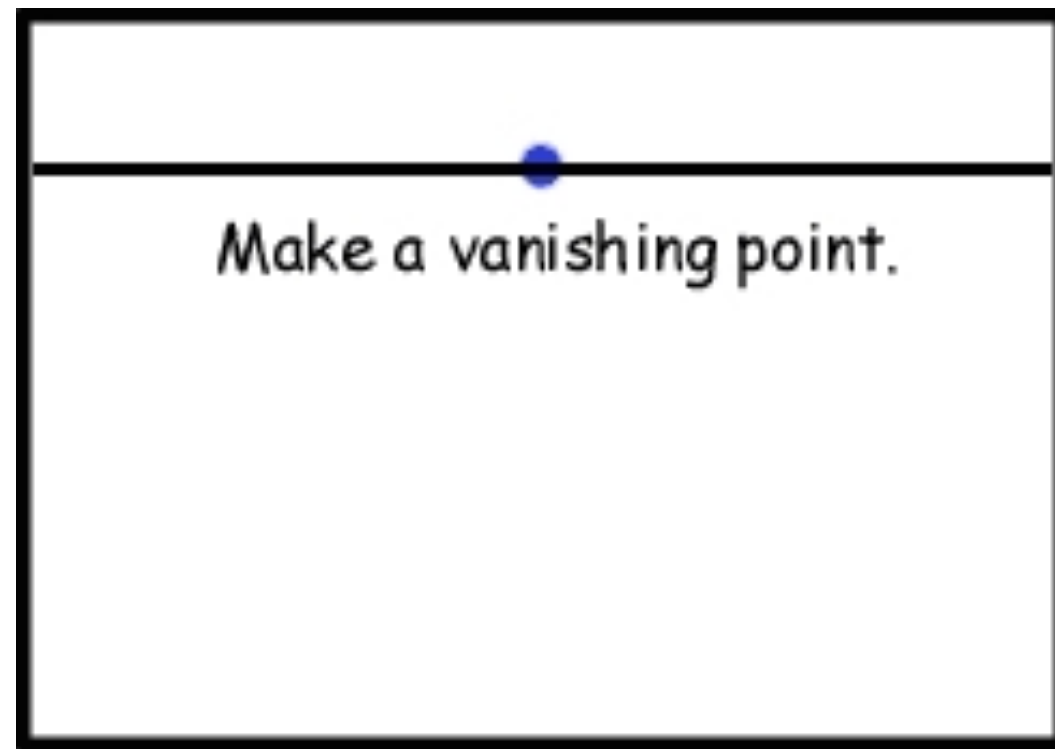
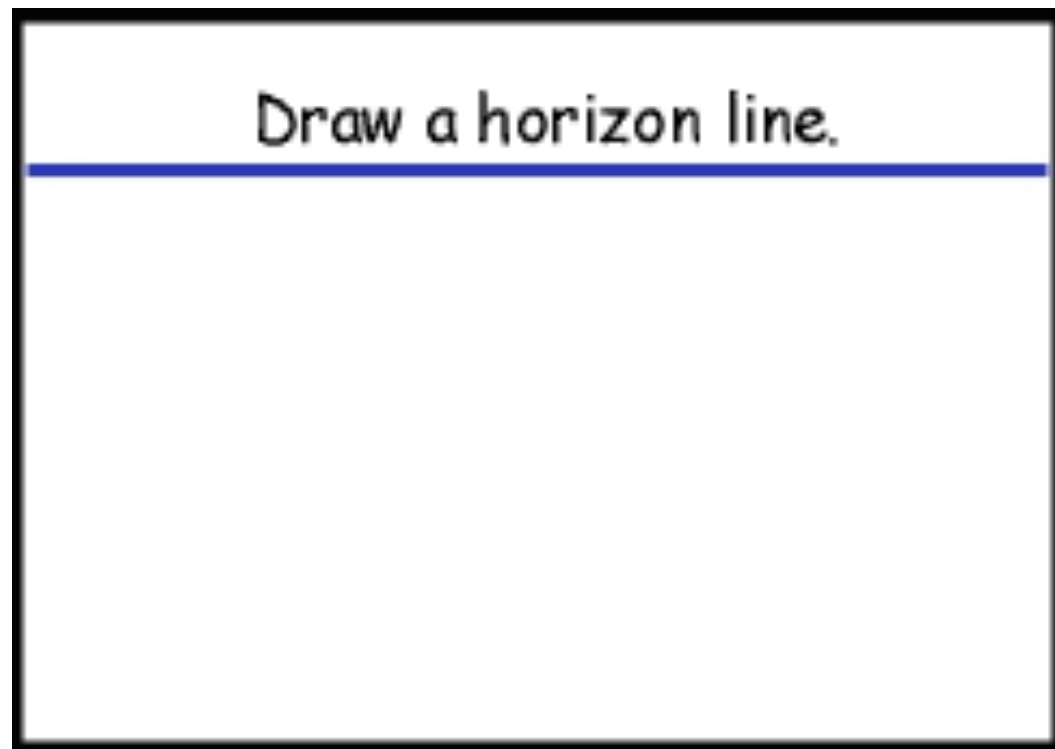
Vanishing Points



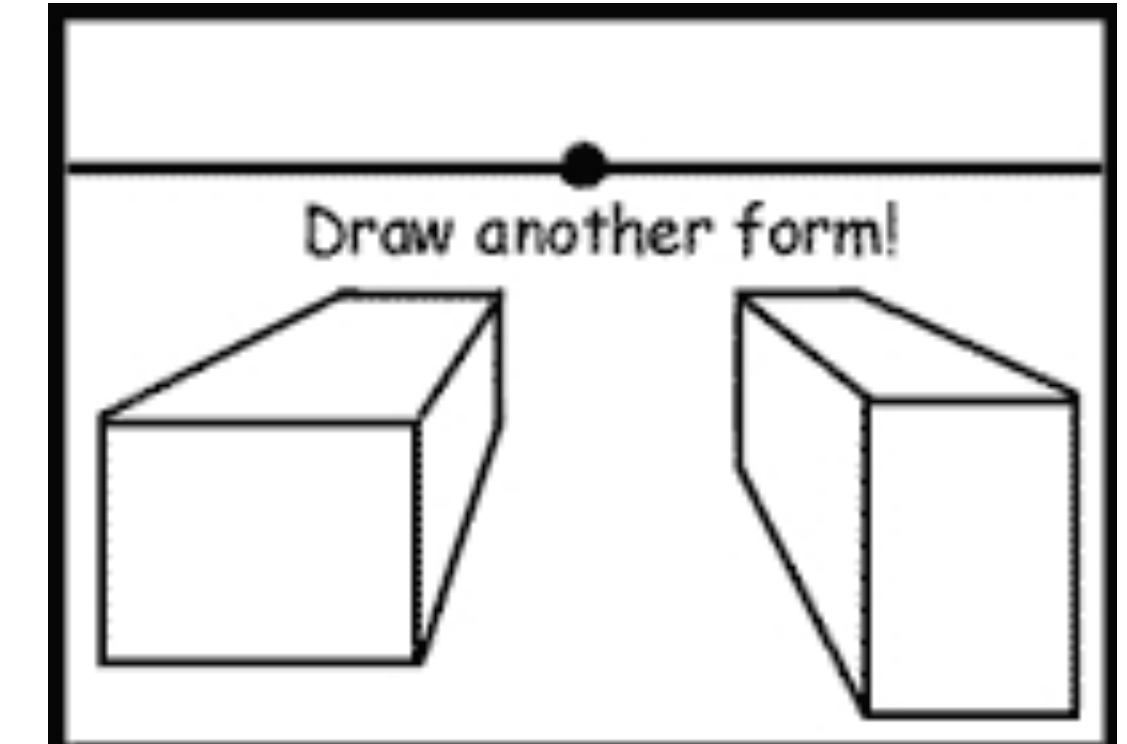
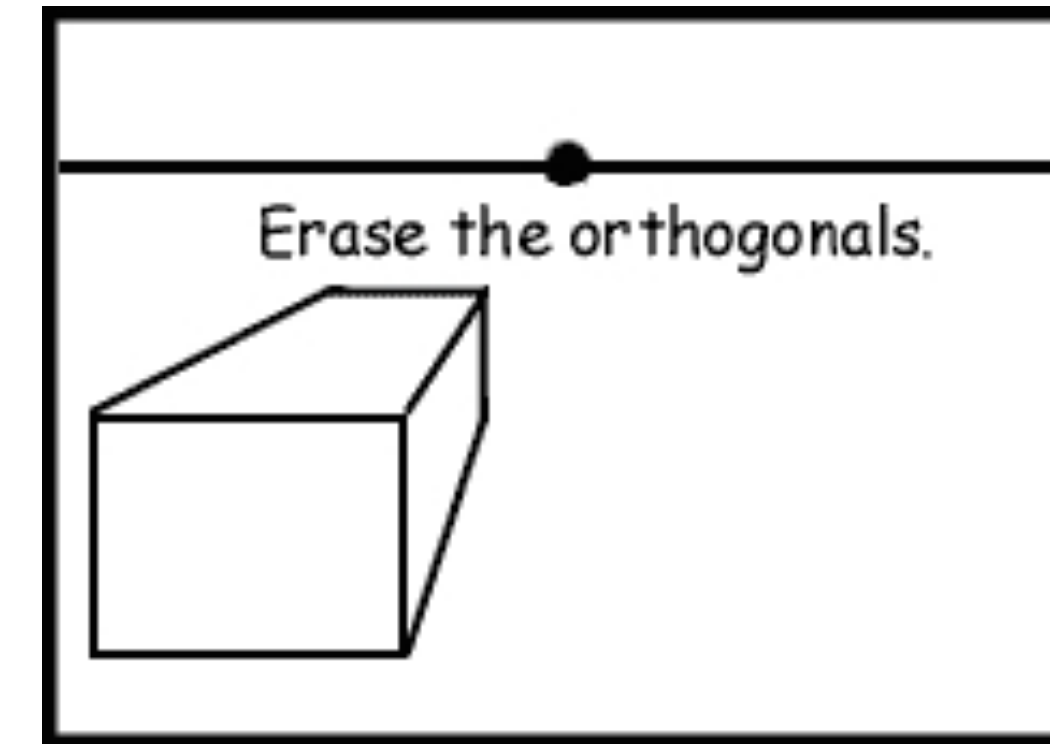
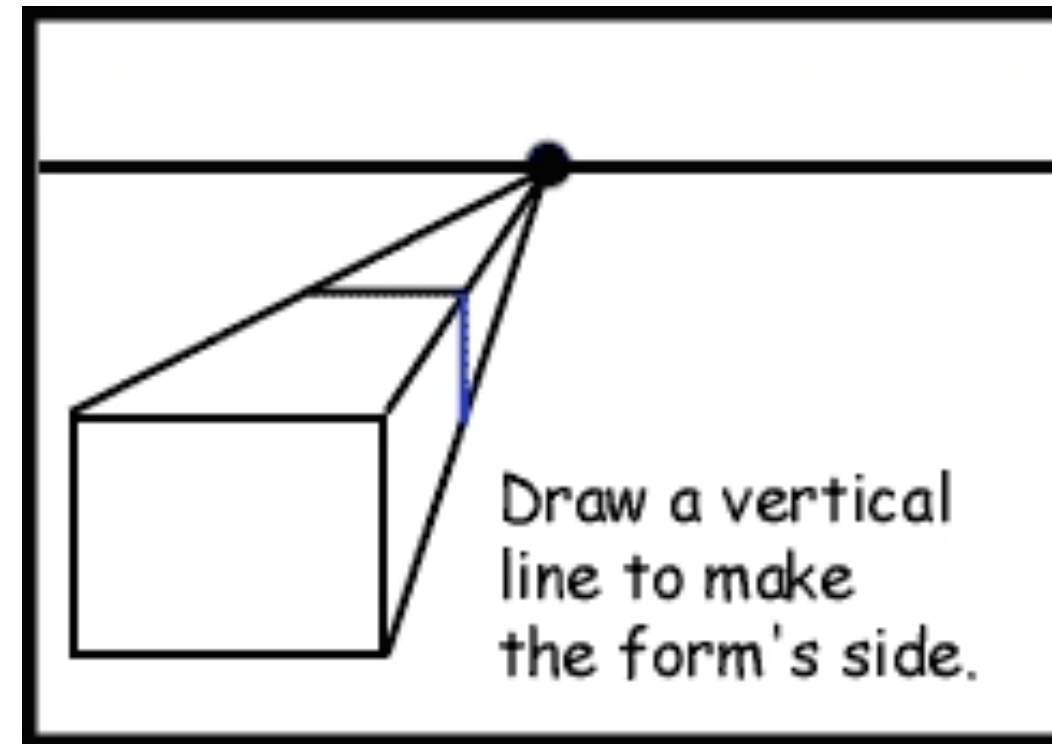
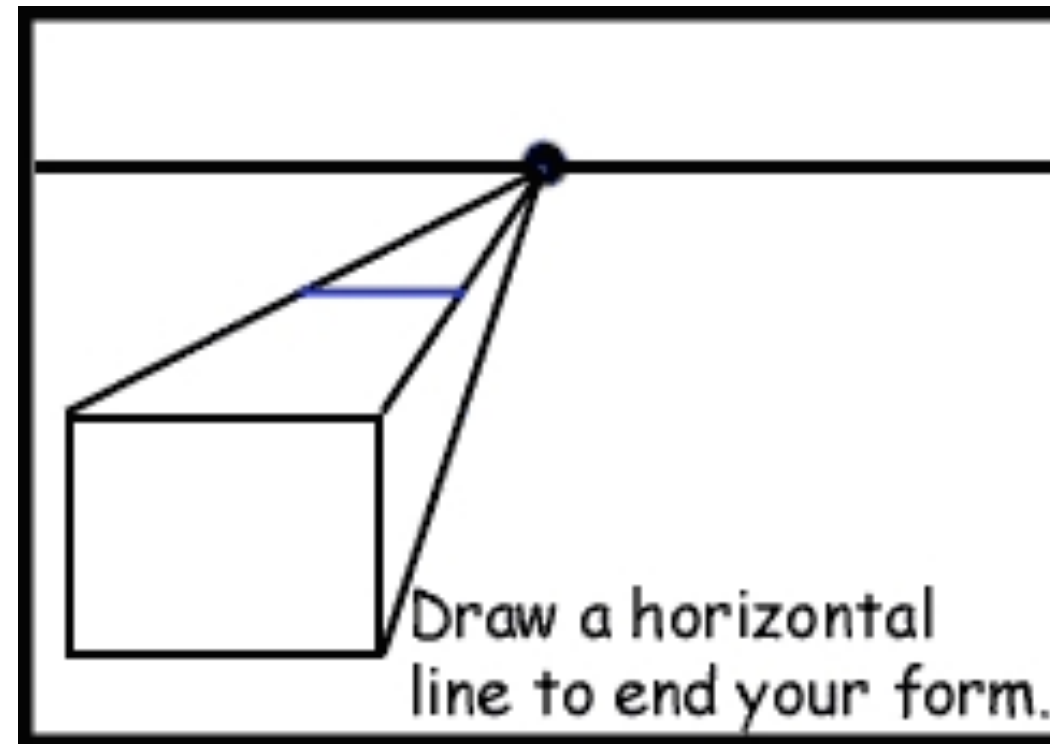
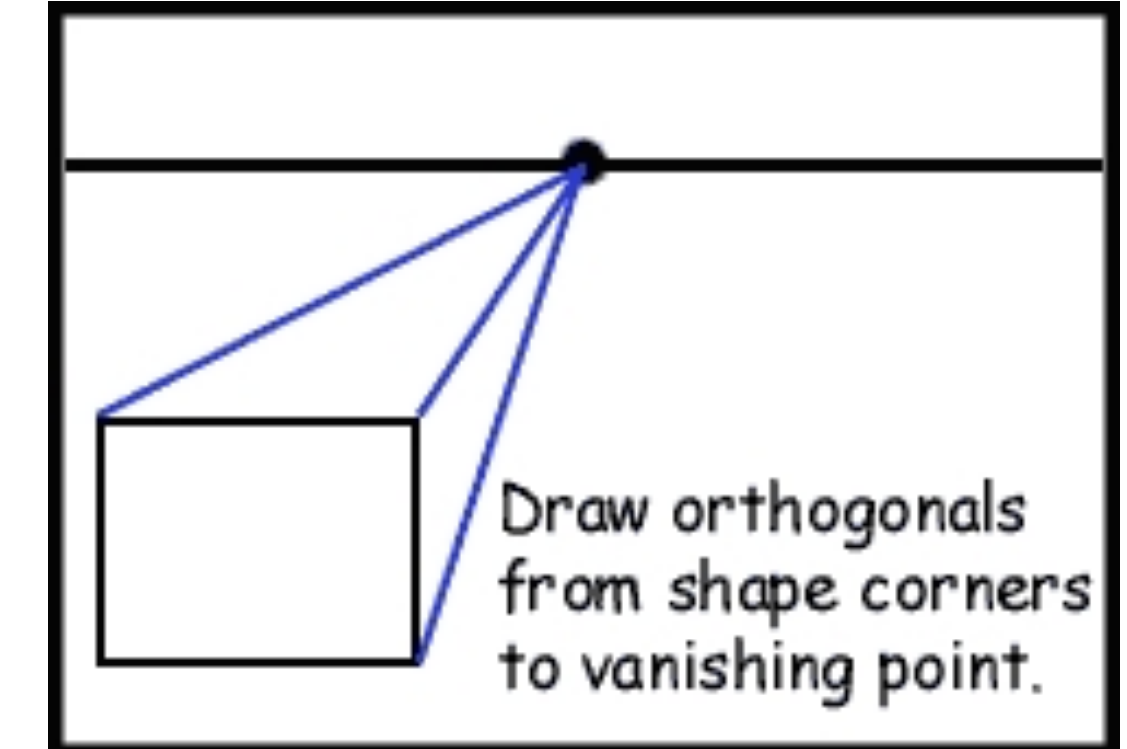
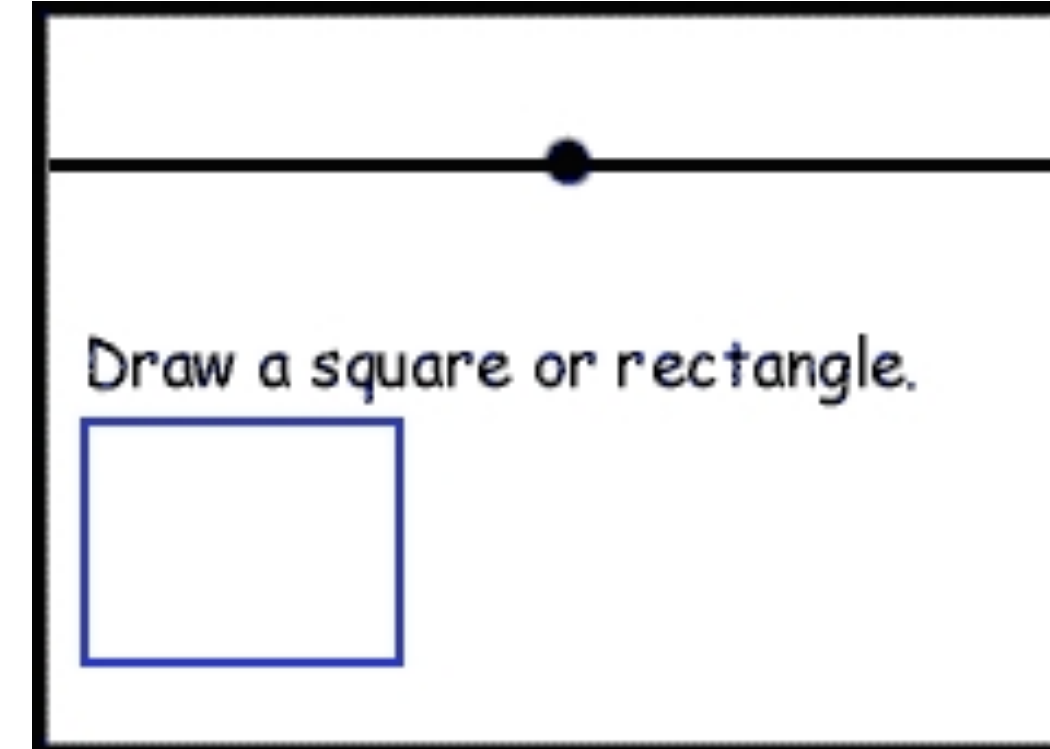
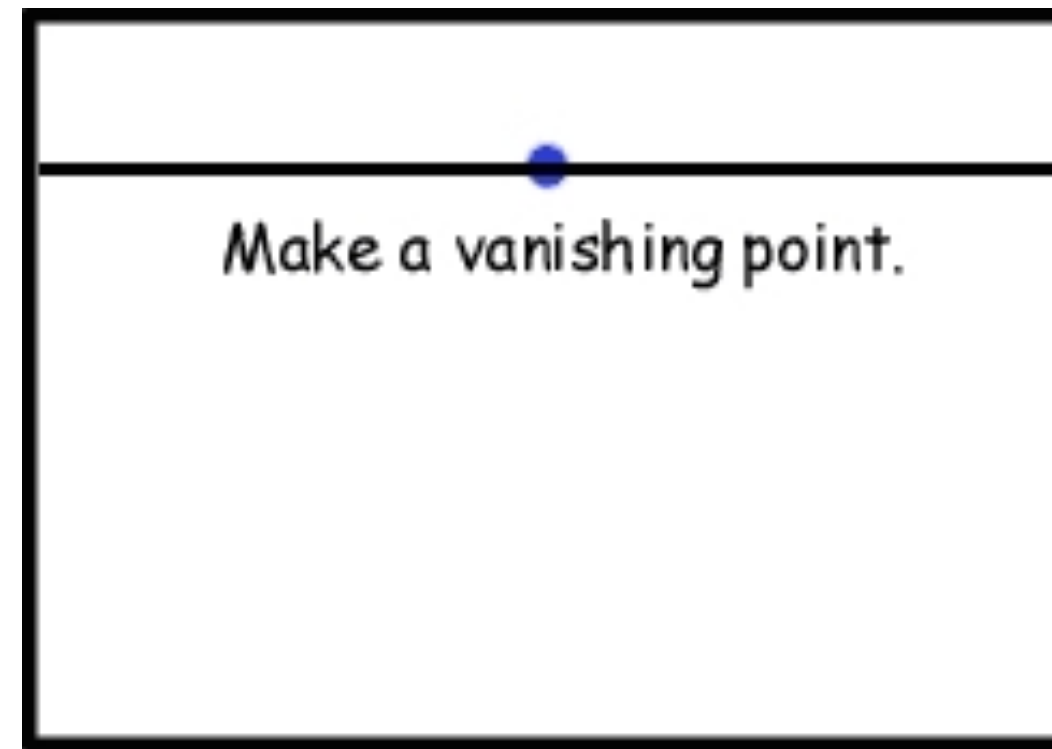
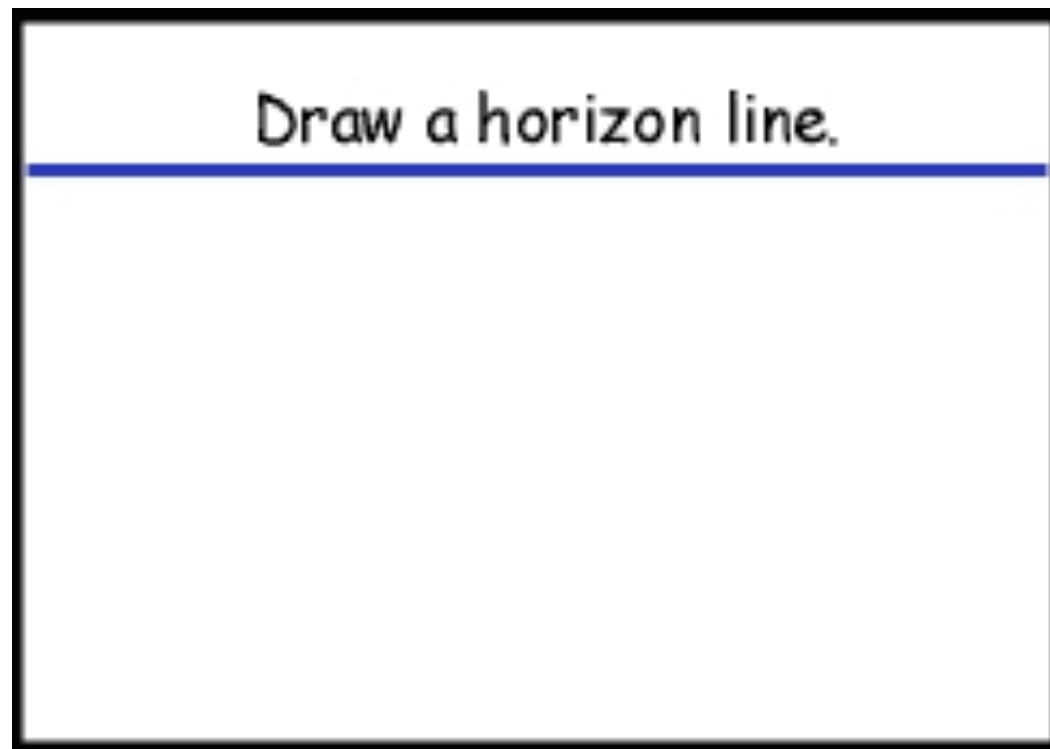
Vanishing Points



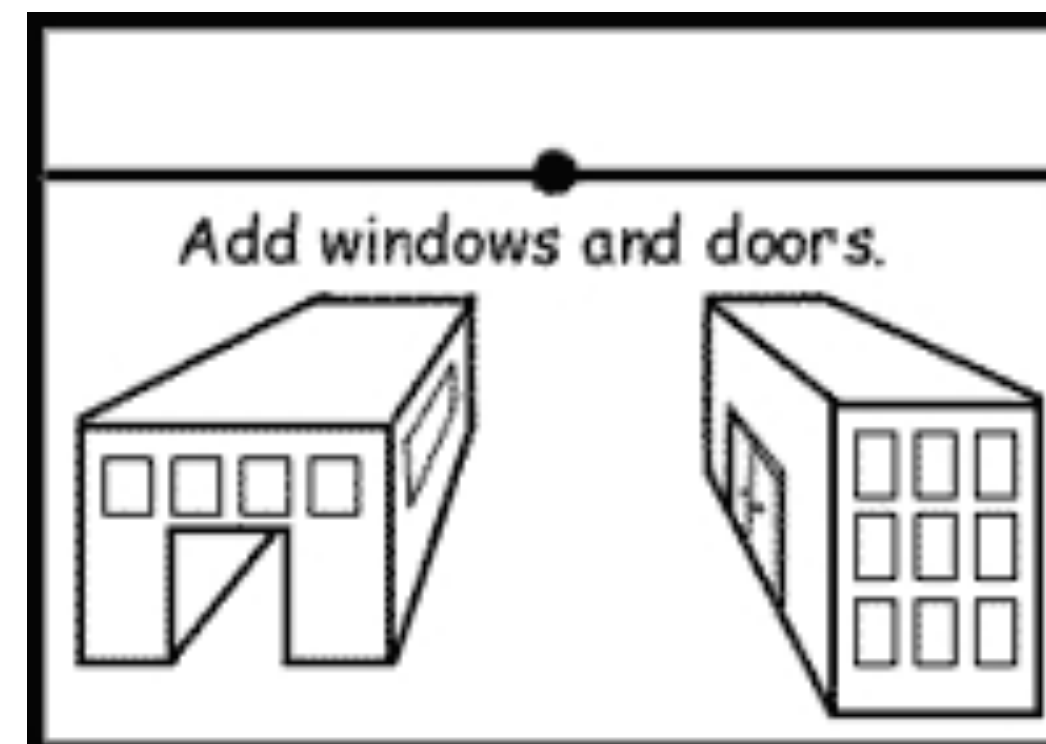
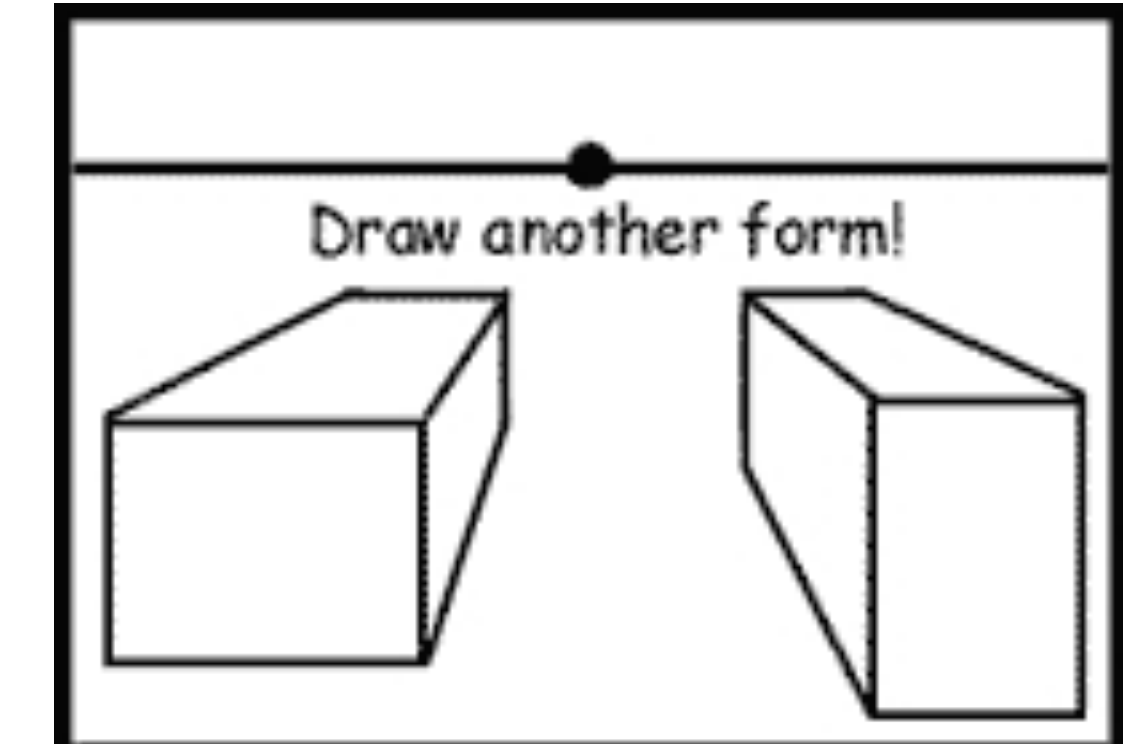
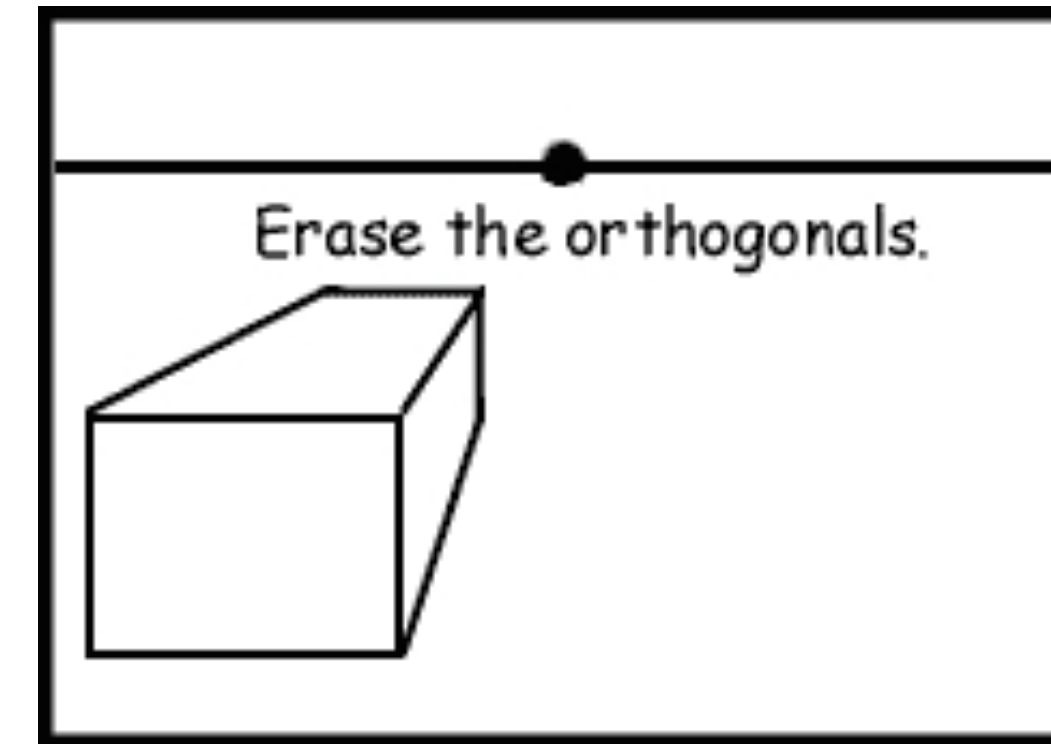
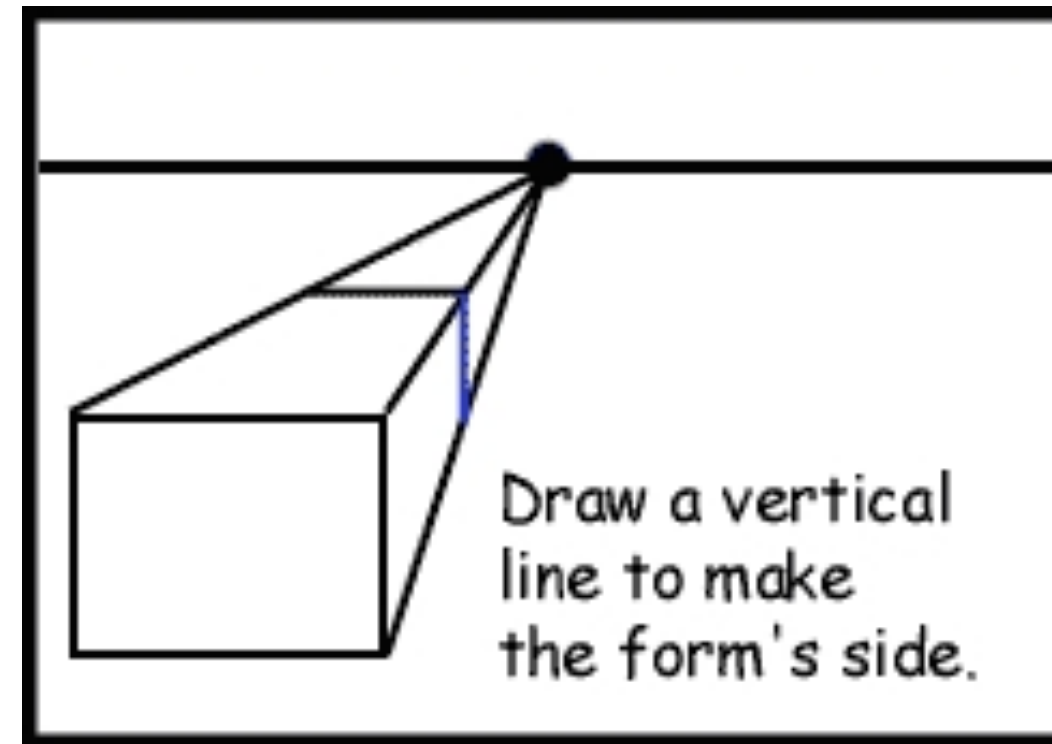
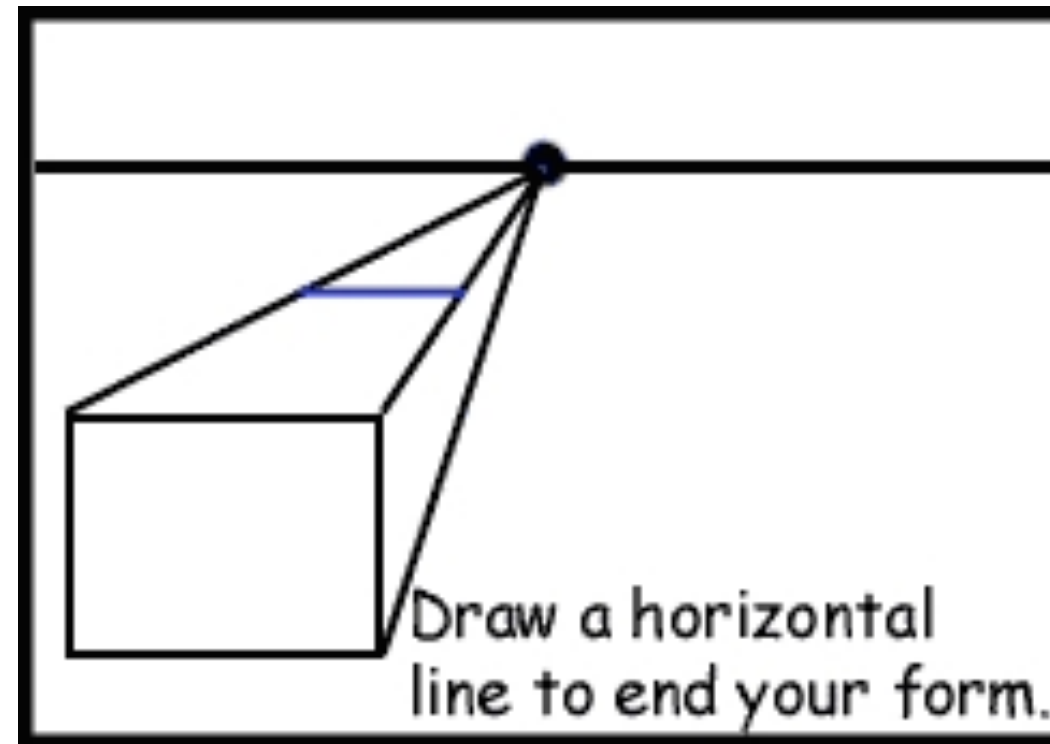
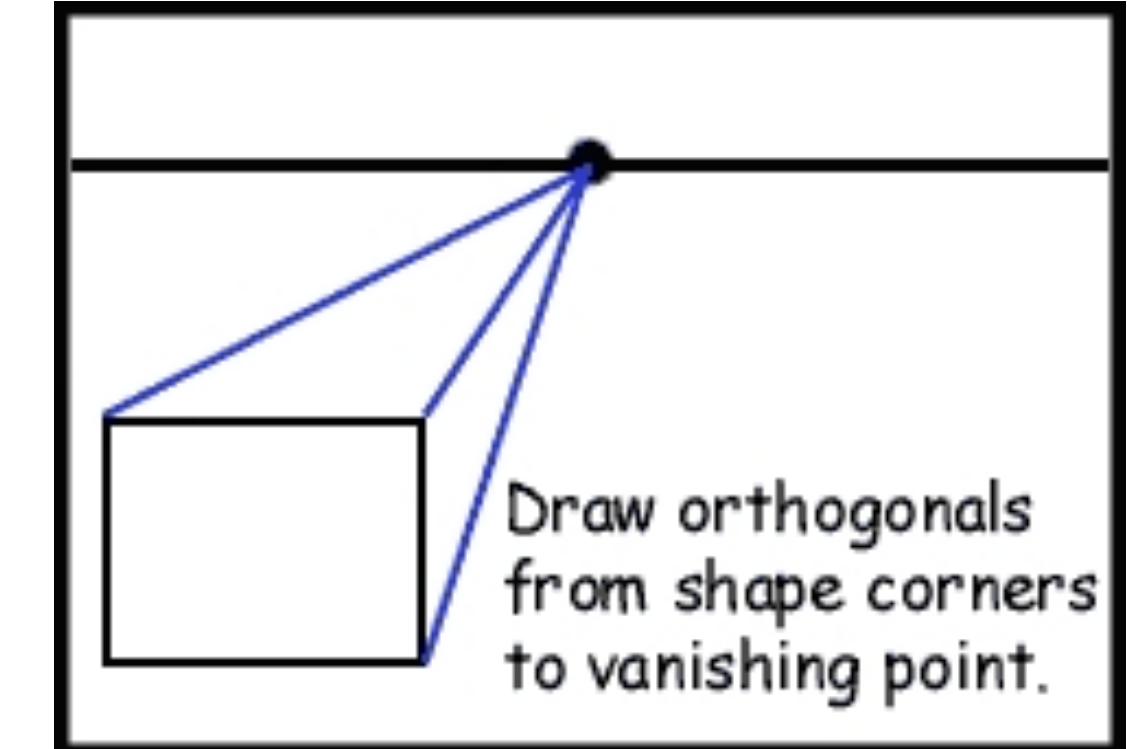
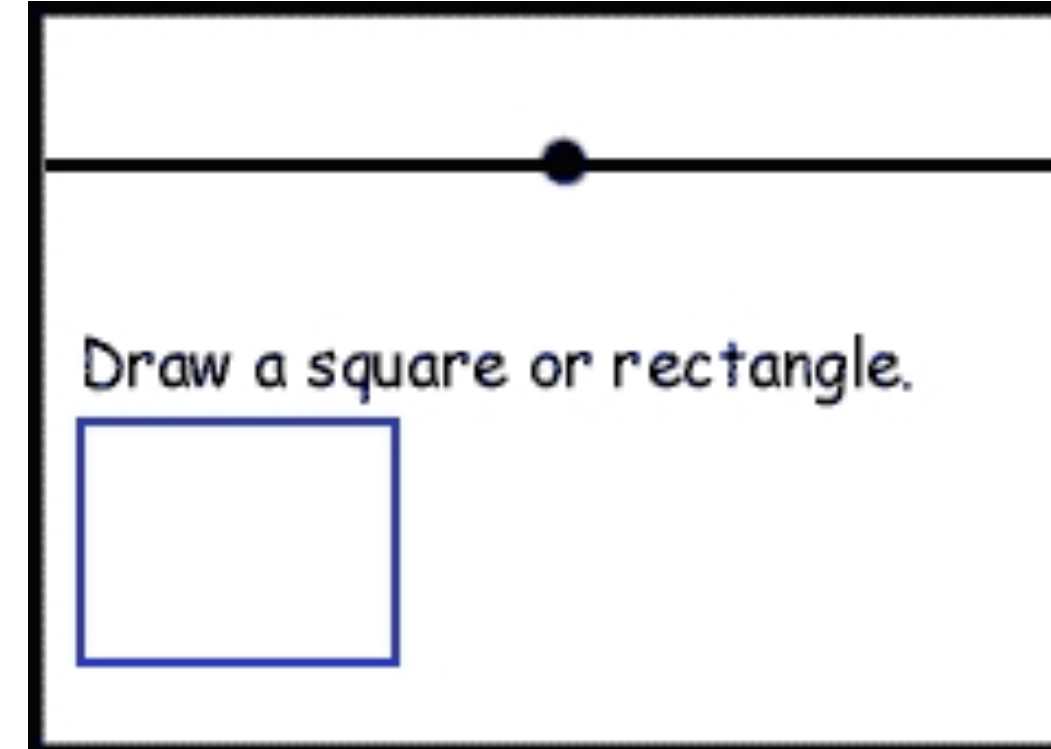
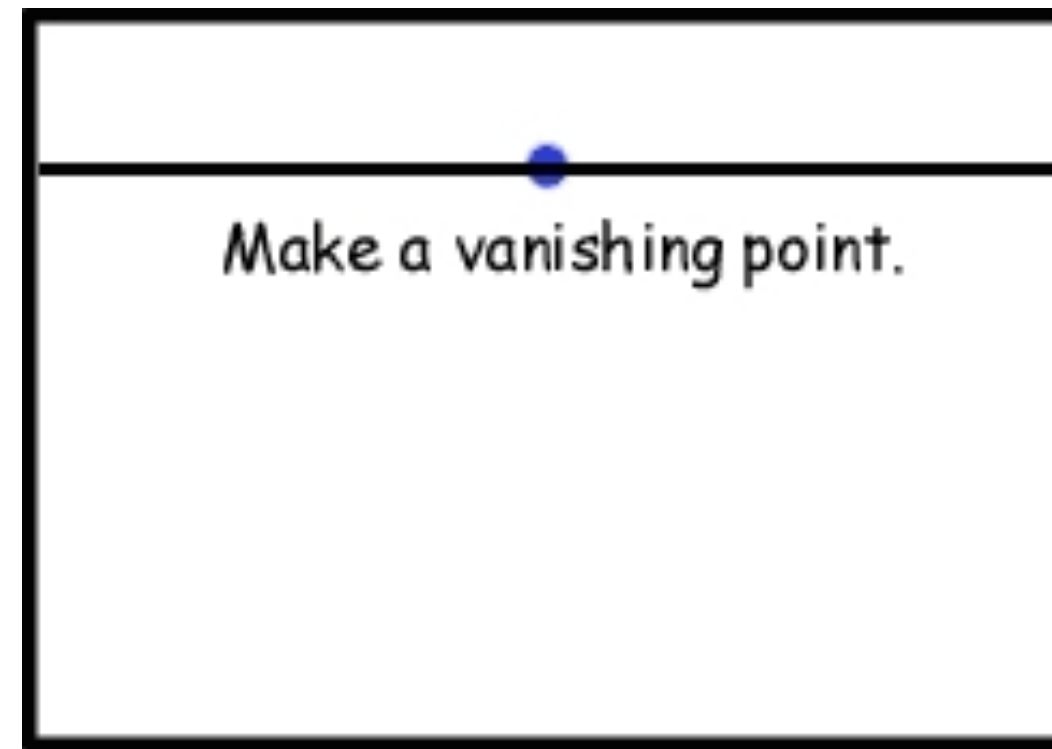
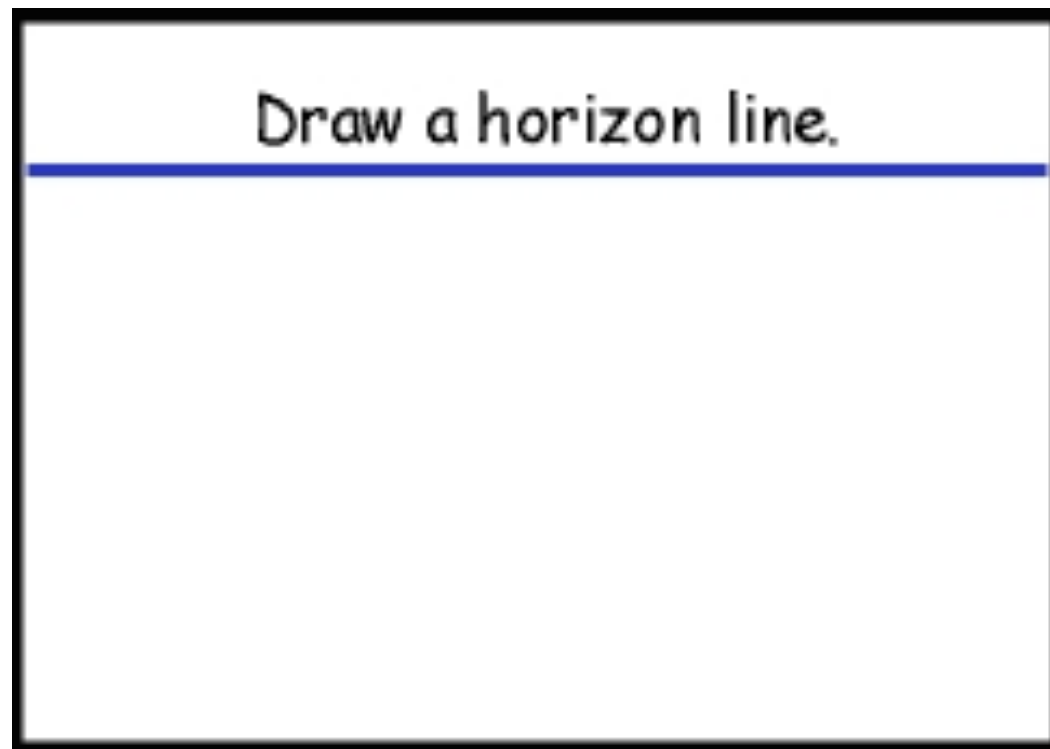
Vanishing Points



Vanishing Points



Vanishing Points



Vanishing Points

Each set of parallel lines meet at a different point

— the point is called **vanishing point**

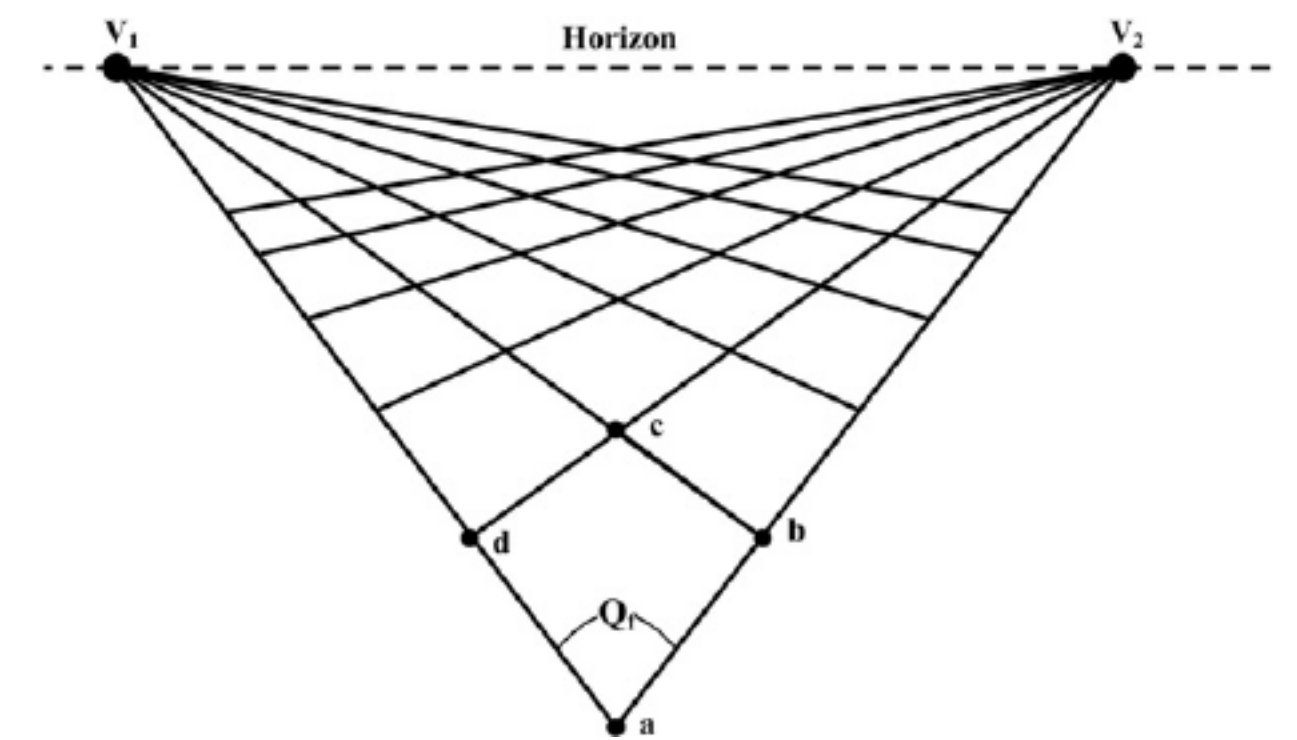
Sets of parallel lines on the same plane lead to **collinear** vanishing points

— the line is called a **horizon** for that plane

Good way to **spot fake images**

— scale and perspective do not work

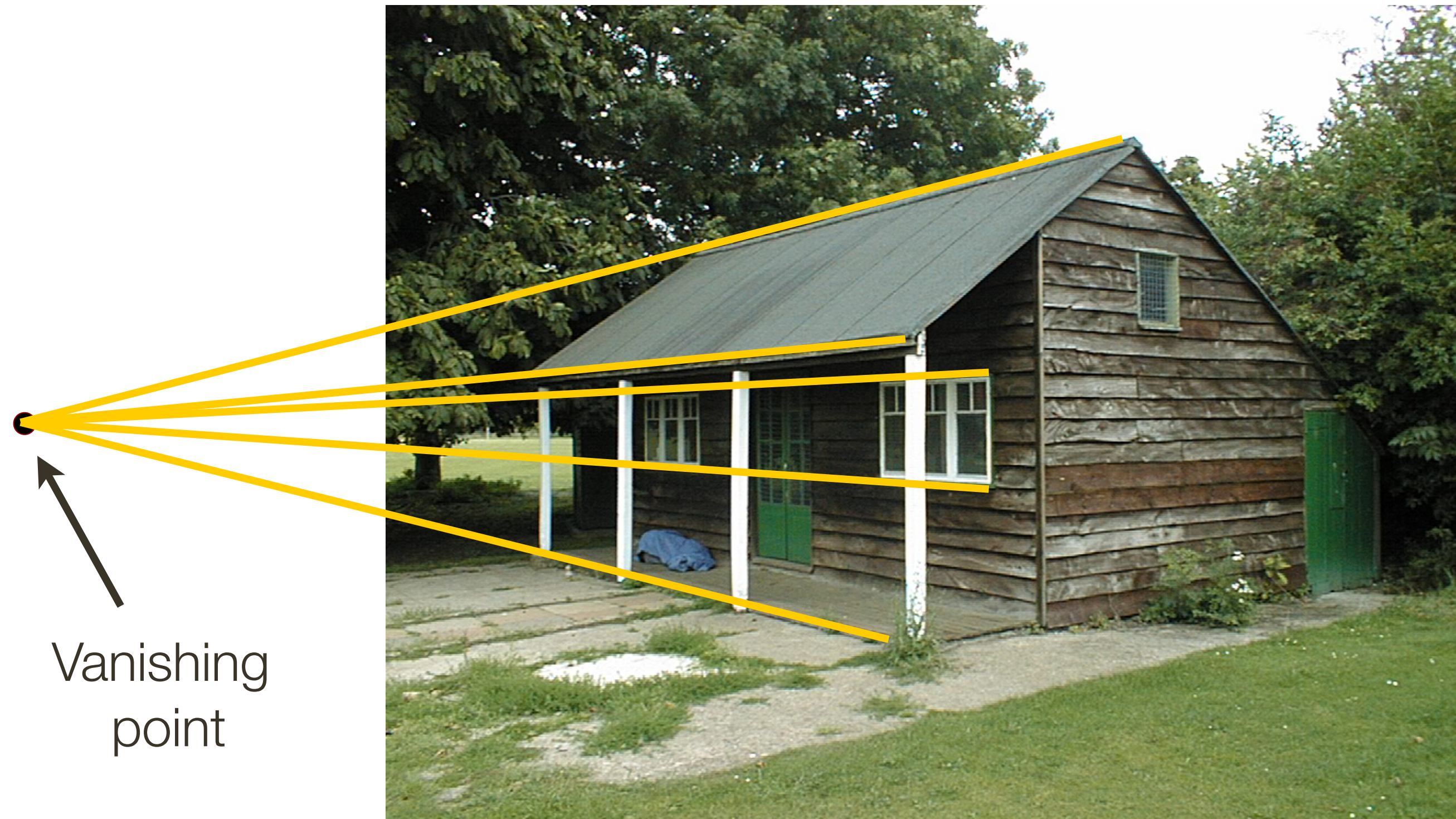
— vanishing points behave badly



Vanishing Points



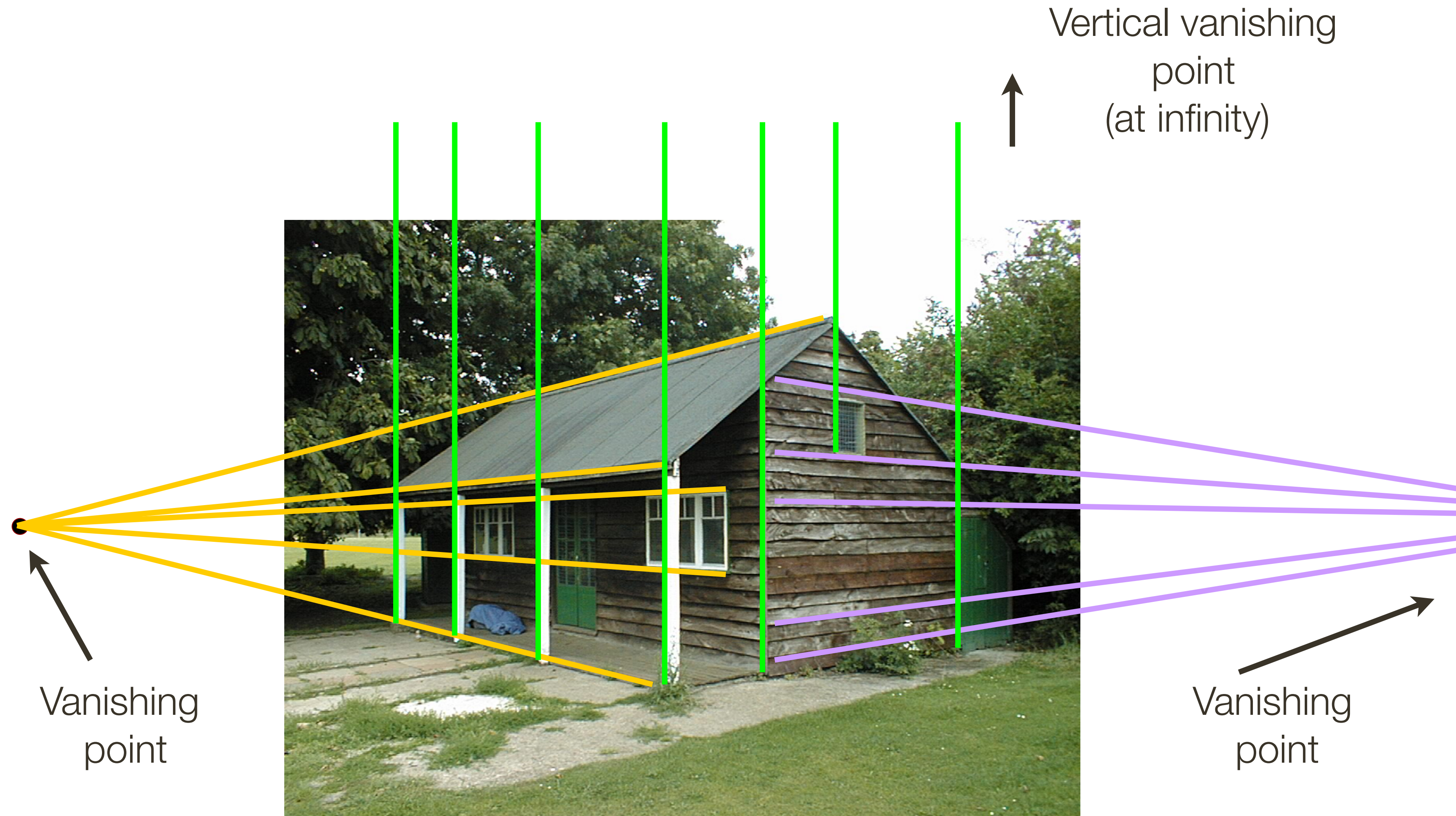
Vanishing Points



Vanishing Points

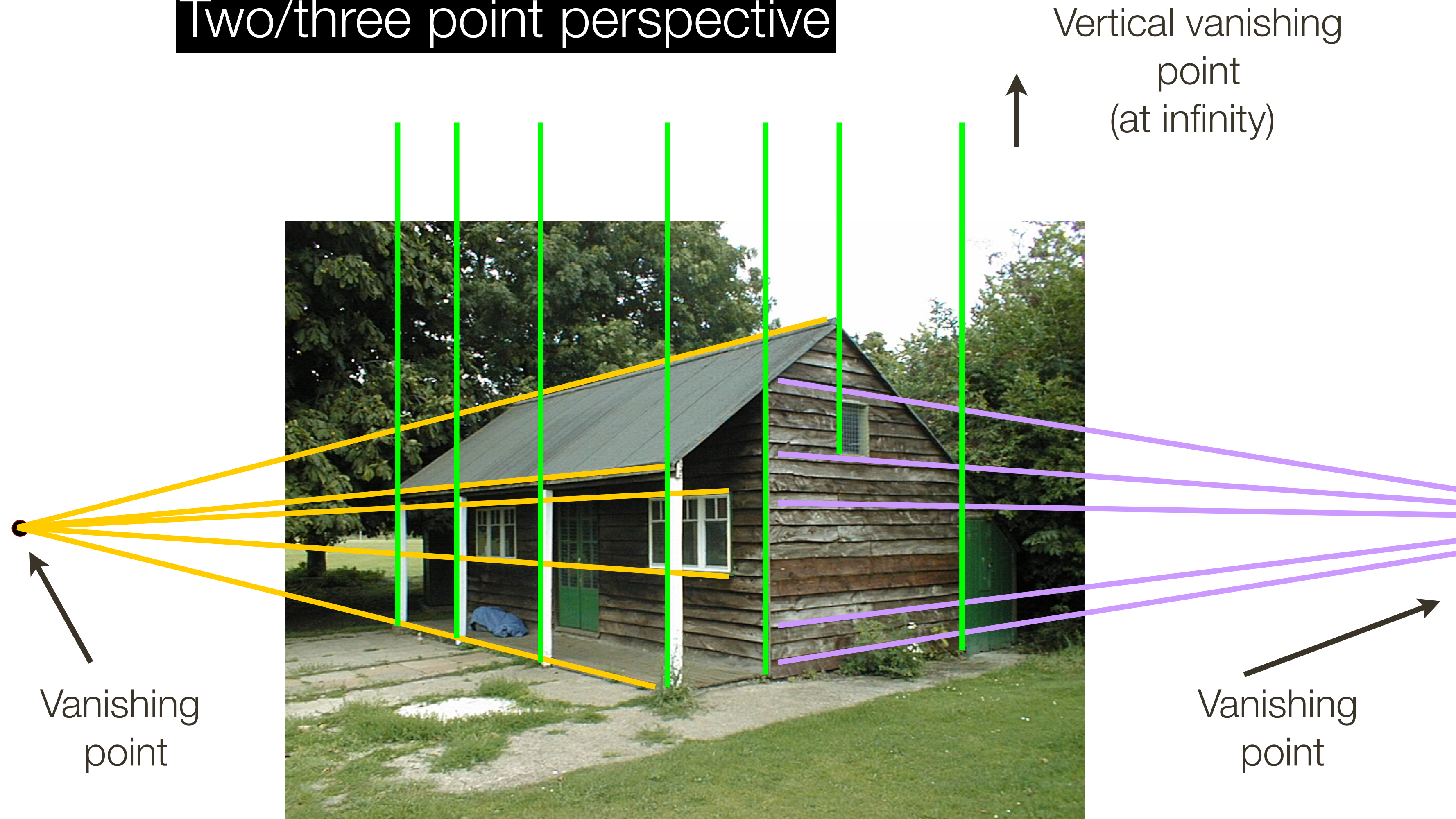


Vanishing Points



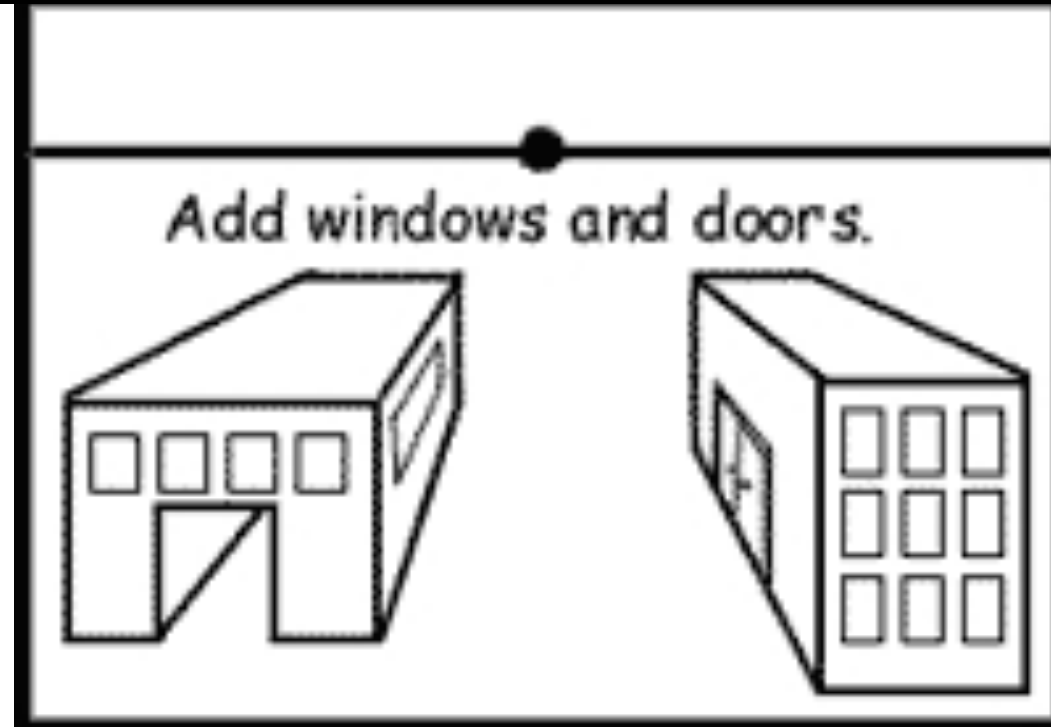
Vanishing Points

Two/three point perspective

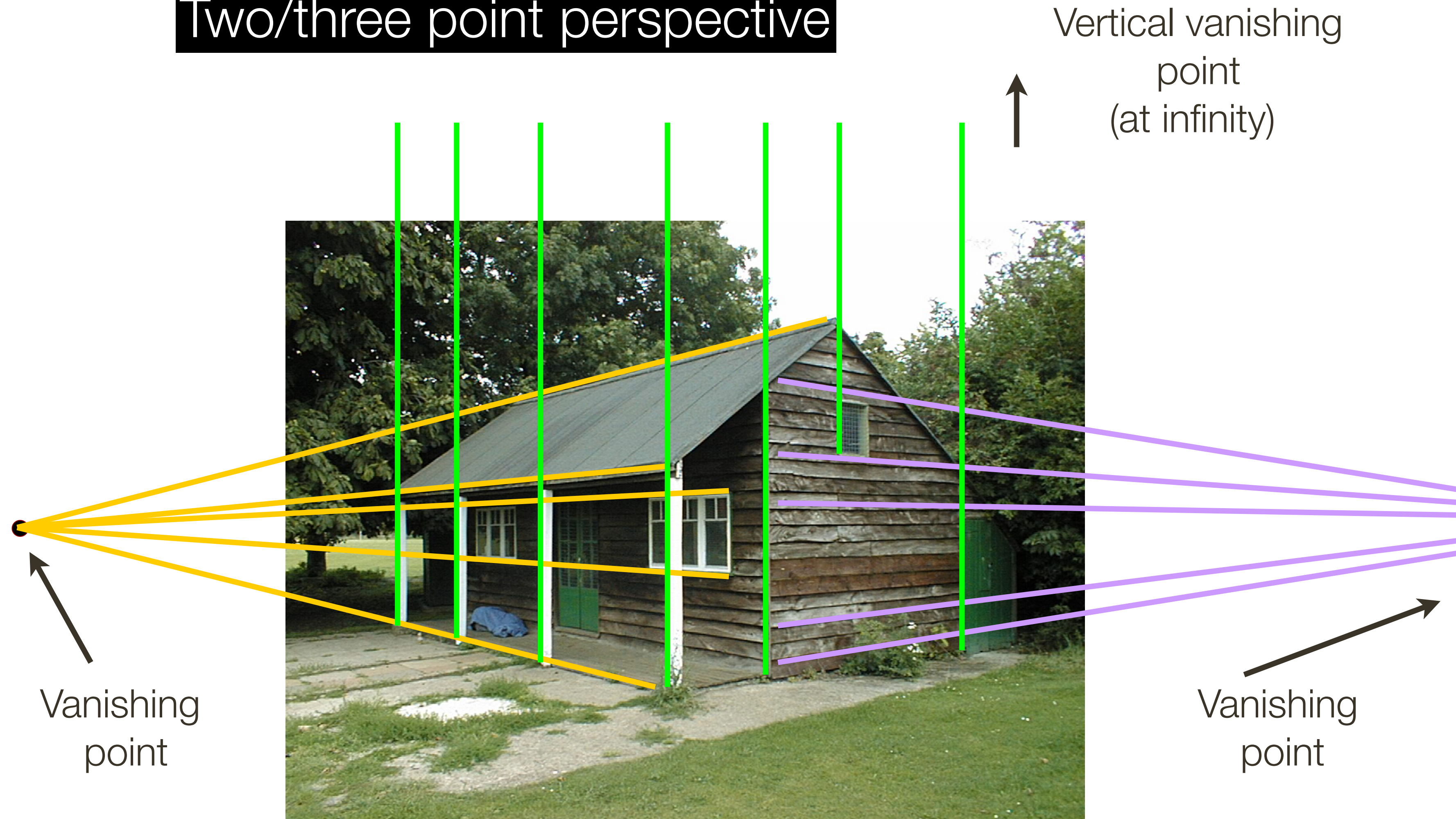


Vanishing Points

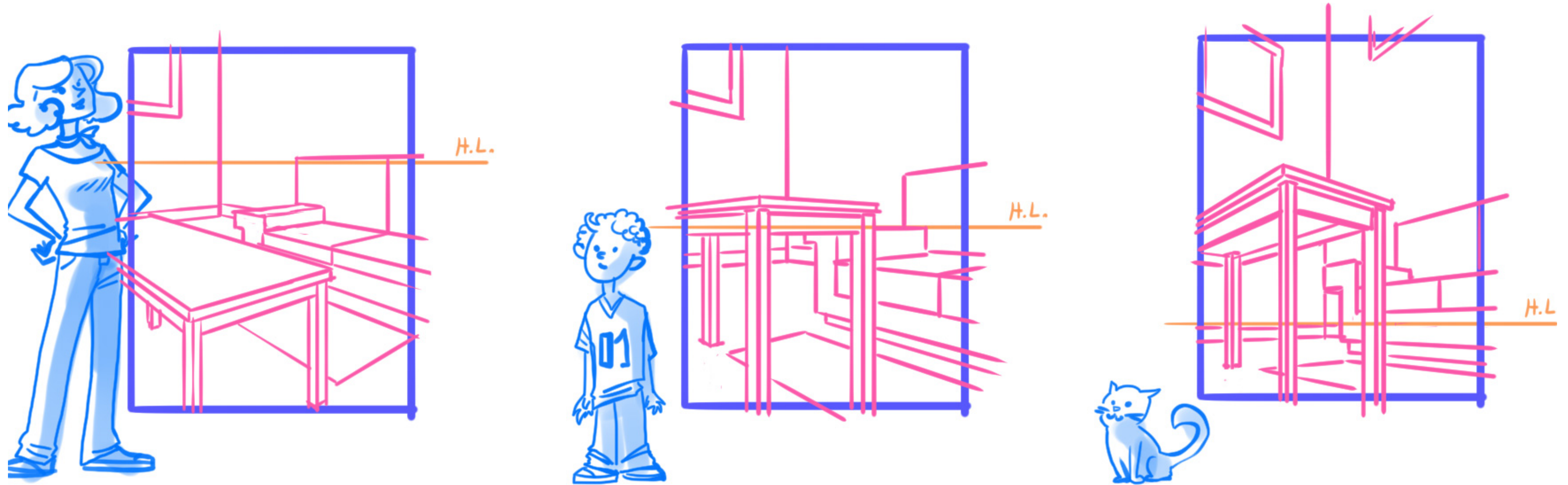
One point perspective



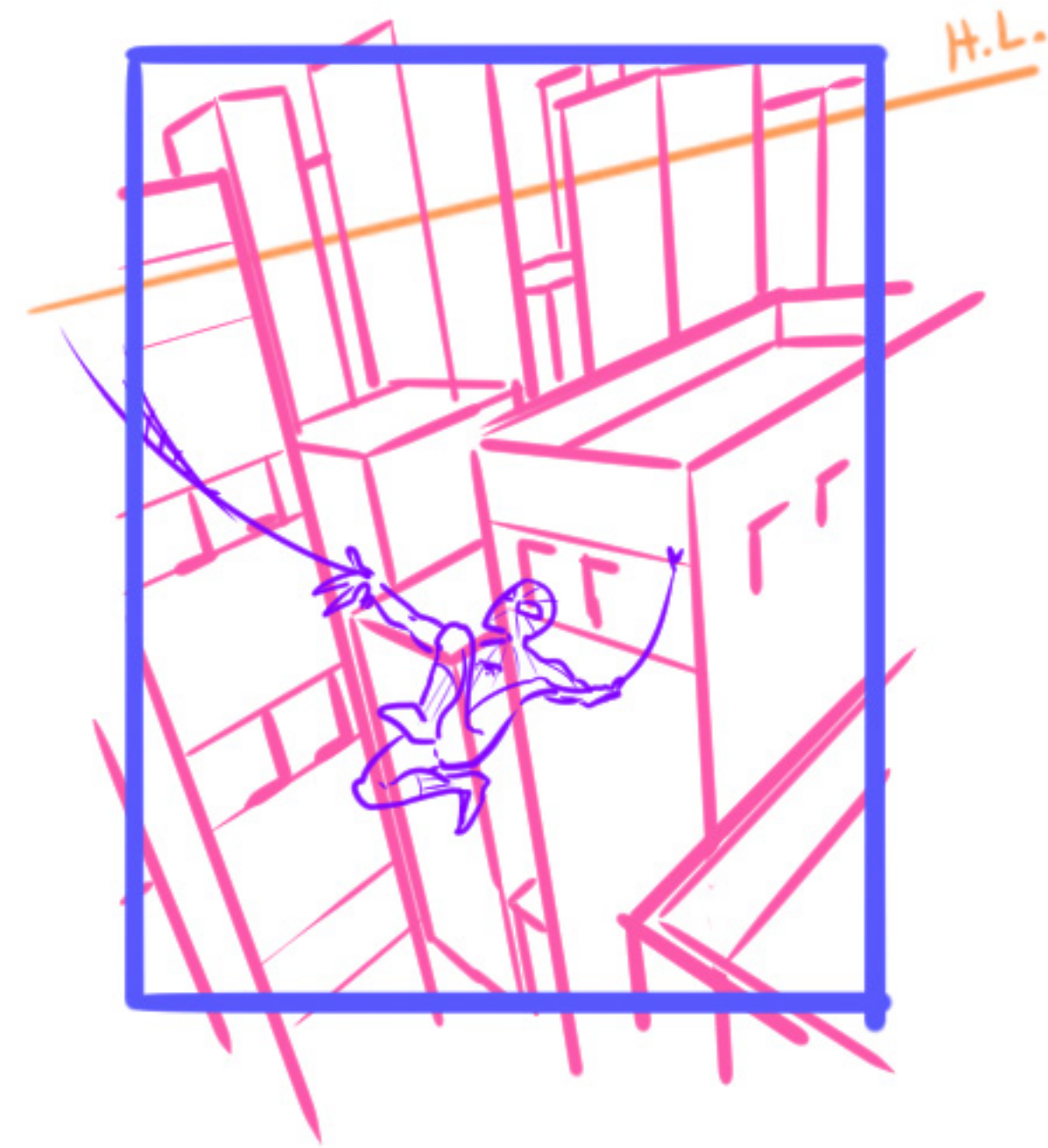
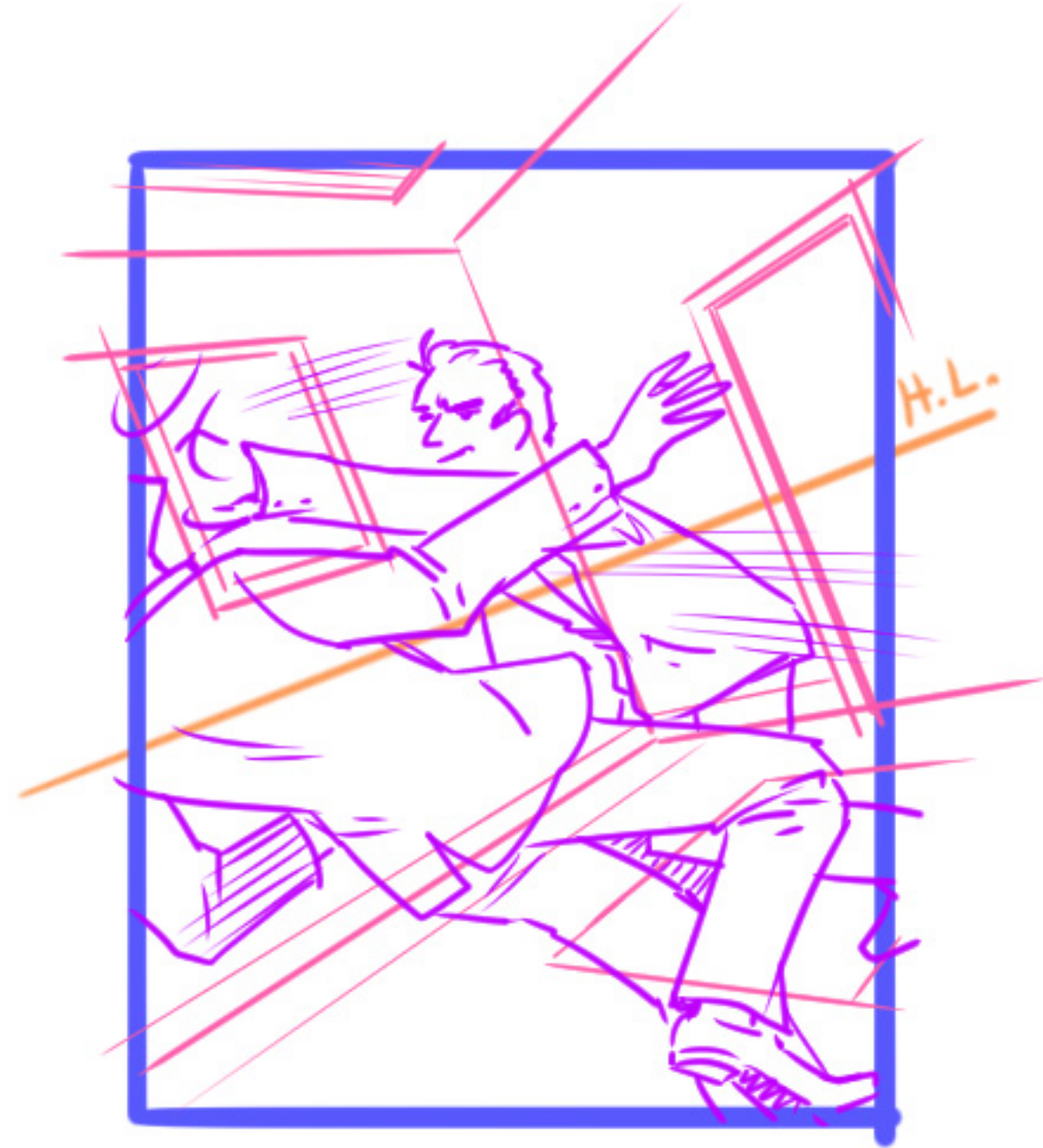
Two/three point perspective



Perspective Aside



Perspective Aside



Properties of Projection

- **Points** project to **points**
- **Lines** project to **lines**
- **Planes** project to the **whole** or **half** image
- Angles are **not** preserved

Properties of Projection

- **Points** project to **points**
- **Lines** project to **lines**
- **Planes** project to the **whole** or **half** image
- Angles are **not** preserved

Degenerate cases

- Line through focal point projects to a point
- Plane through focal point projects to a line

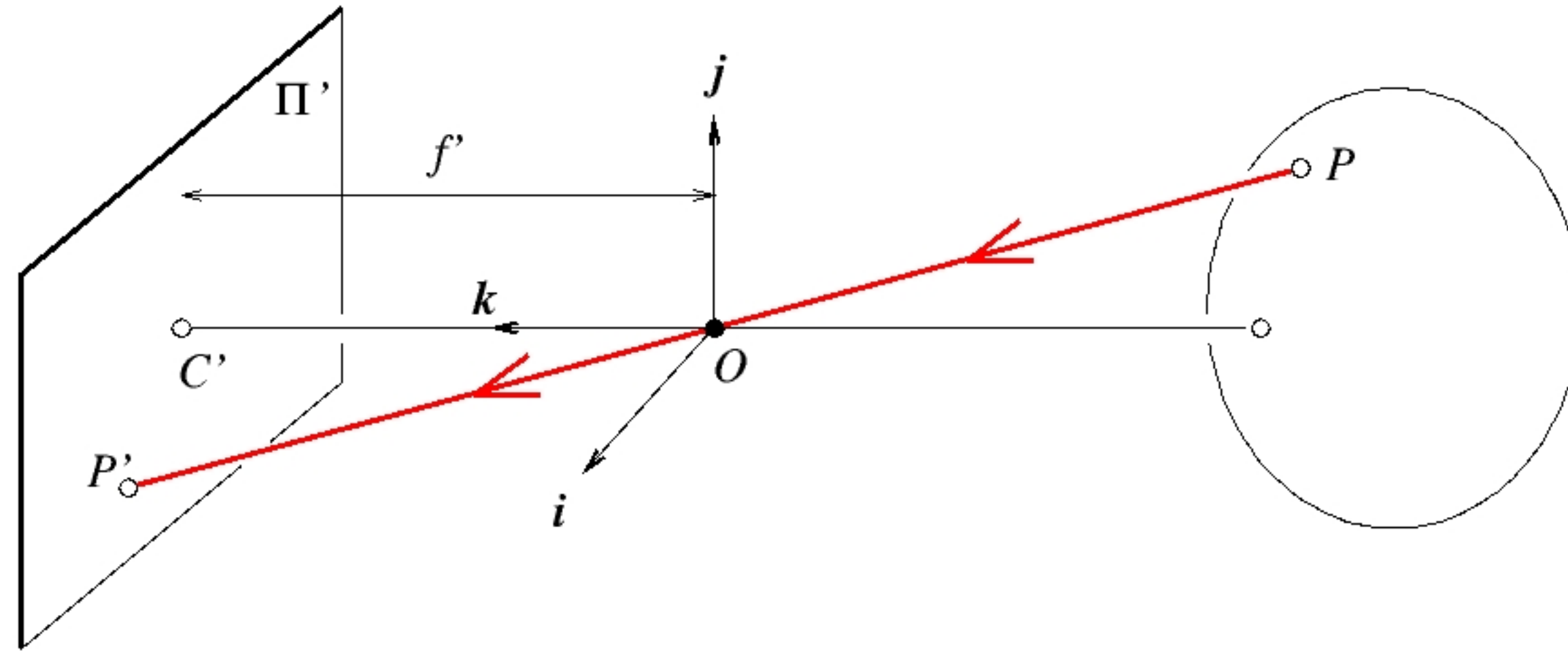
Projection Illusion



Projection Illusion



Perspective Projection



3D object point

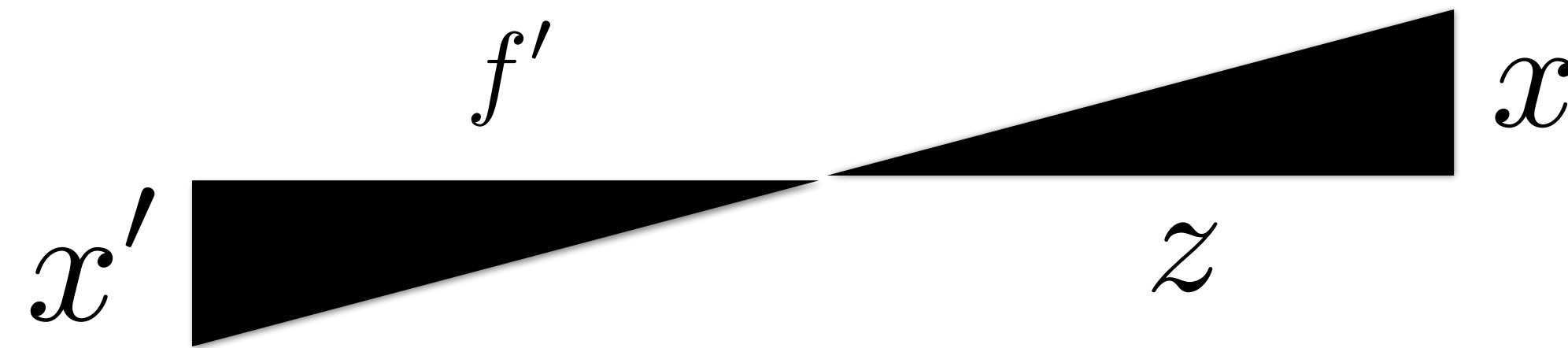
Forsyth & Ponce (1st ed.) Figure 1.4

$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ where

$$\begin{aligned} x' &= f' \frac{x}{z} \\ y' &= f' \frac{y}{z} \end{aligned}$$

Note: this assumes world coordinate frame at the optical center (pinhole) and aligned with the image plane, image coordinate frame aligned with the camera coordinate frame

Perspective Projection: Proof



3D object point

Forsyth & Ponce (1st ed.) Figure 1.4

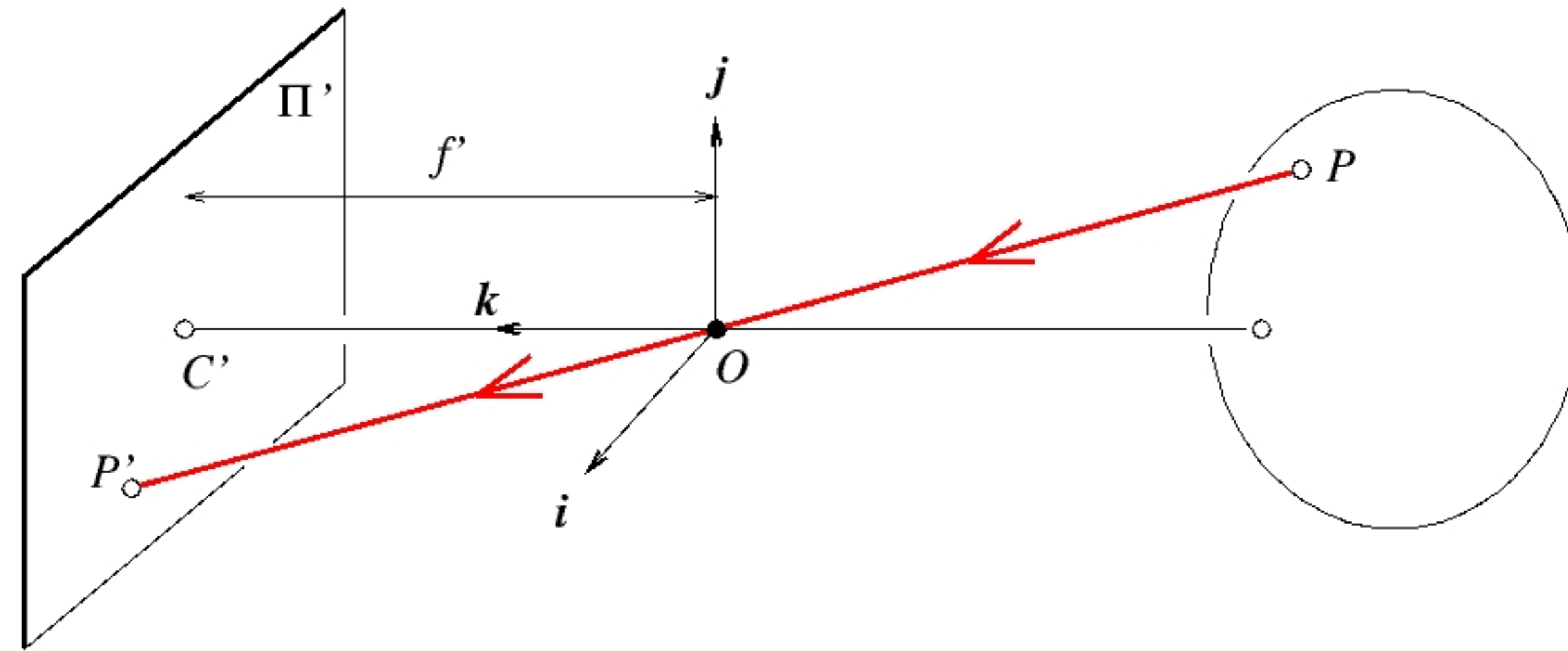
$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ where

$$\begin{aligned} x' &= f' \frac{x}{z} \\ y' &= f' \frac{y}{z} \end{aligned}$$

Note: this assumes world coordinate frame at the optical center (pinhole) and aligned with the image plane, image coordinate frame aligned with the camera coordinate frame

Perspective Projection: Proof

Camera Matrix



$$\mathbf{C} = \begin{bmatrix} f' & 0 & 0 & 0 \\ 0 & f' & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

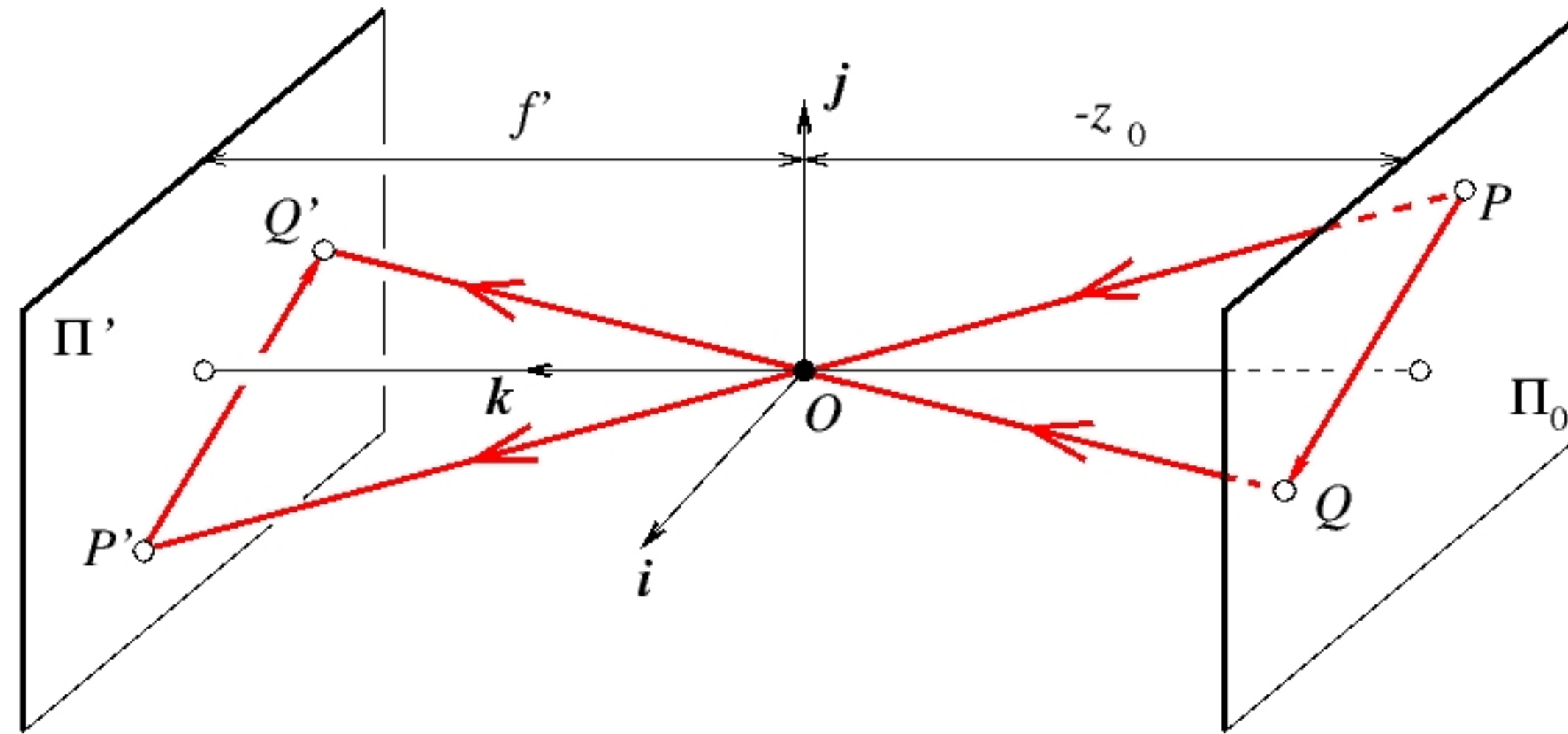
3D object point

Forsyth & Ponce (1st ed.) Figure 1.4

$P = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$ projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$ where $P' = \mathbf{C}P$

Note: this assumes world coordinate frame at the optical center (pinhole) and aligned with the image plane, image coordinate frame aligned with the camera coordinate frame

Weak Perspective

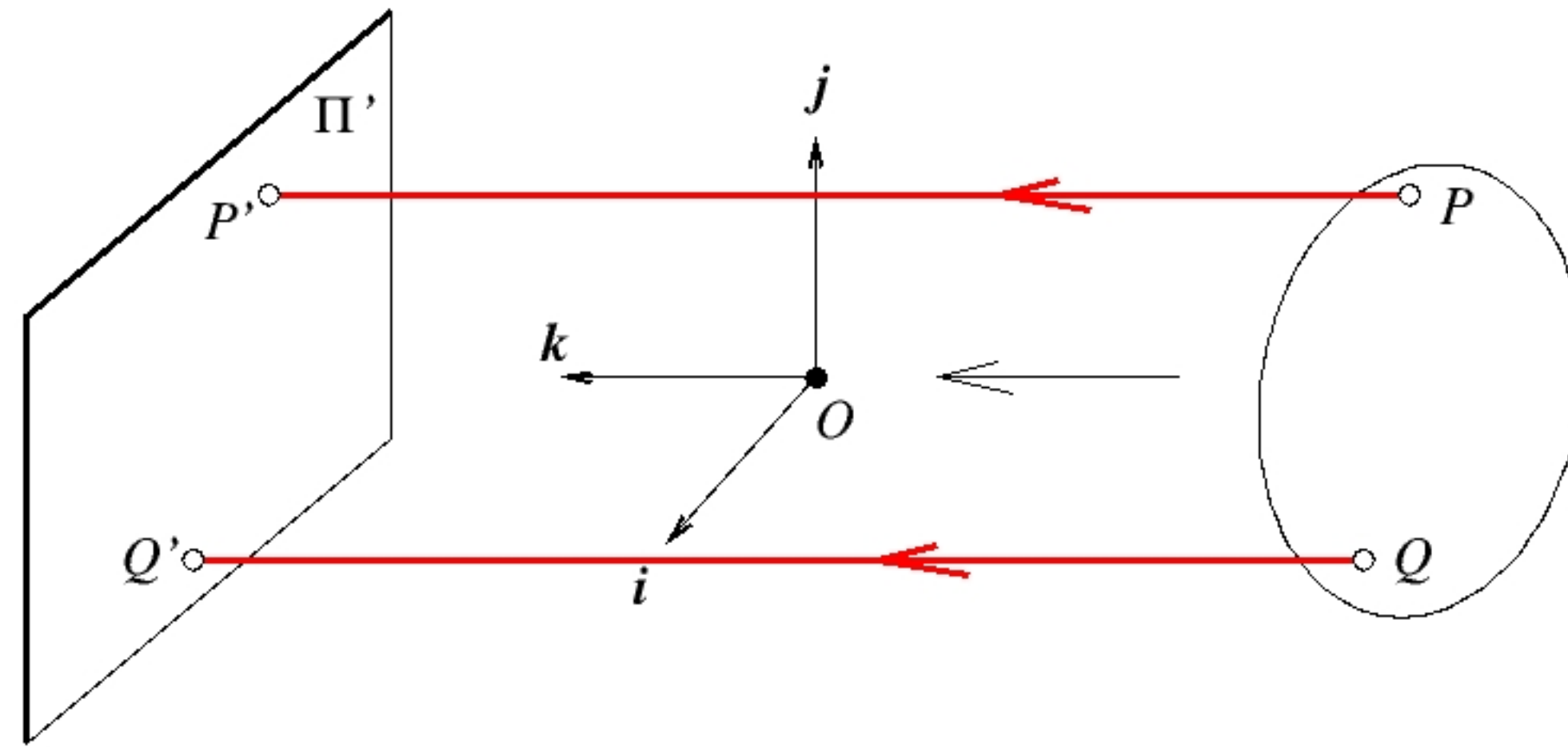


Forsyth & Ponce (1st ed.) Figure 1.5

3D object point $P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in Π_0 projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$

where $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} mx \\ my \end{bmatrix}$ and $m = \frac{f'}{z_0}$

Orthographic Projection



Forsyth & Ponce (1st ed.) Figure 1.6

3D object point $P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$

where

$$\begin{array}{l} x' = x \\ y' = y \end{array}$$

Summary of **Projection Equations**

3D object point $P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ projects to 2D image point $P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ where

Perspective

$$\begin{aligned} x' &= f' \frac{x}{z} \\ y' &= f' \frac{y}{z} \end{aligned}$$

Weak Perspective

$$\begin{aligned} x' &= m x \\ y' &= m y \end{aligned} \quad m = \frac{f'}{z_0}$$

Orthographic

$$\begin{aligned} x' &= x \\ y' &= y \end{aligned}$$

Projection Models: Pros and Cons

Weak perspective (including orthographic) has simpler mathematics

- accurate when object is small and/or distant
- useful for recognition

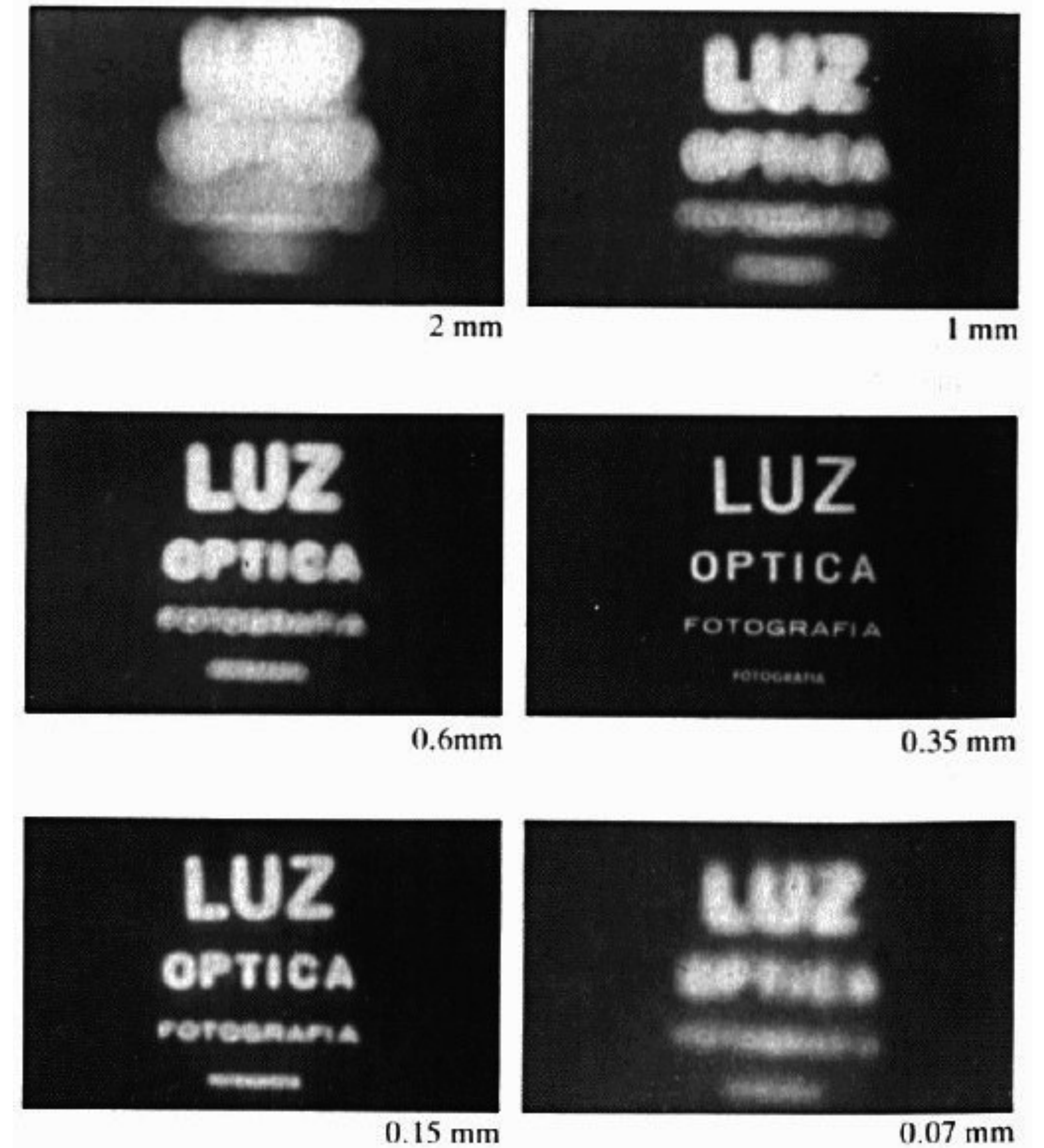
Perspective is more accurate for real scenes

When **maximum accuracy** is required, it is necessary to model additional details of a particular camera

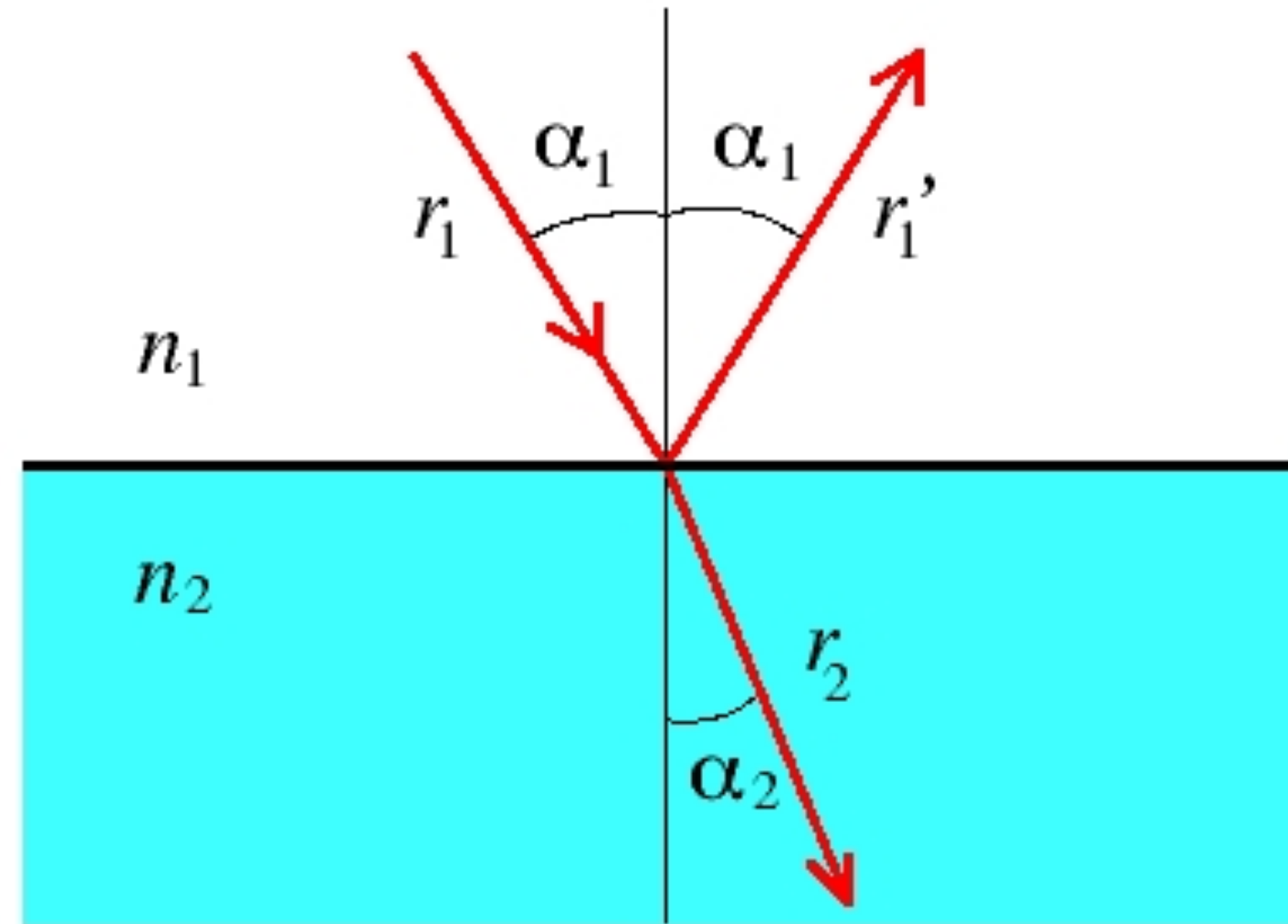
- use perspective projection with additional parameters (e.g., lens distortion)

Why **Not** a Pinhole Camera?

- If pinhole is **too big** then many directions are averaged, blurring the image
- If pinhole is **too small** then diffraction becomes a factor, also blurring the image
- Generally, pinhole cameras are **dark**, because only a very small set of rays from a particular scene point hits the image plane
- Pinhole cameras are **slow**, because only a very small amount of light from a particular scene point hits the image plane per unit time

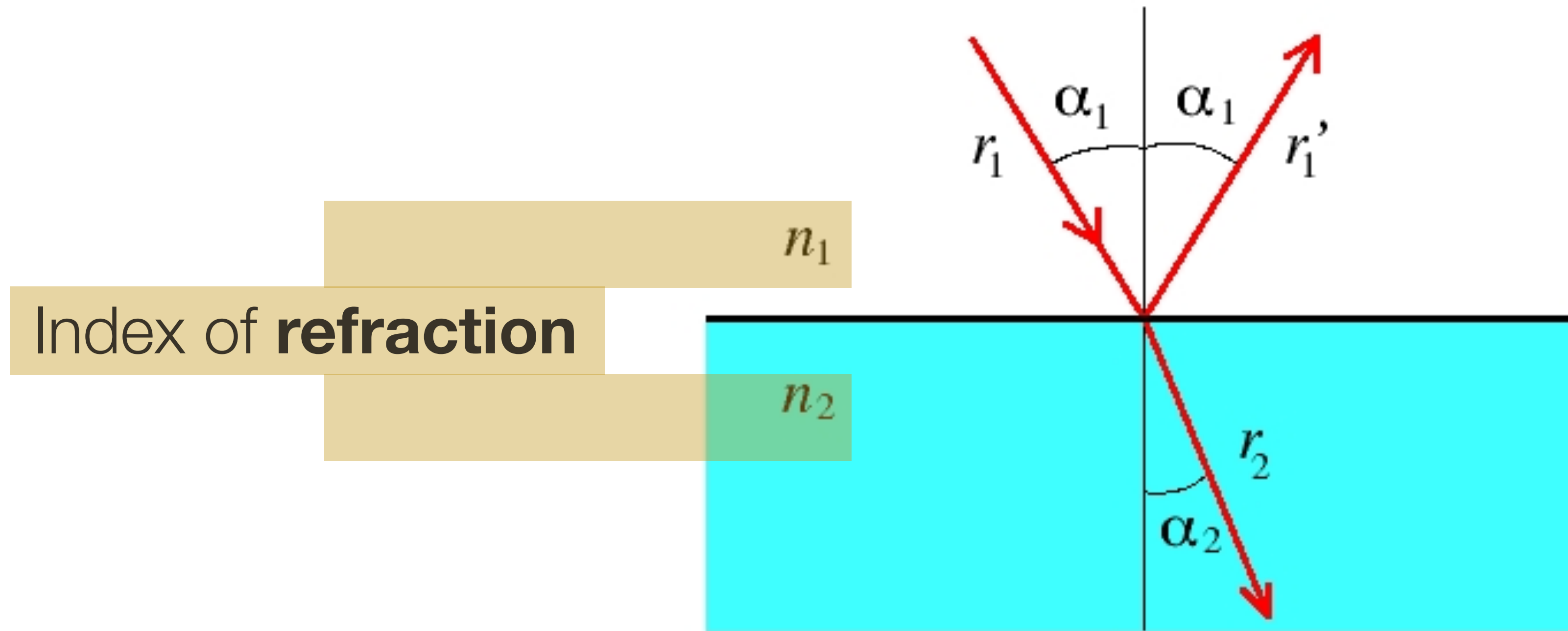


Snell's Law



$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$$

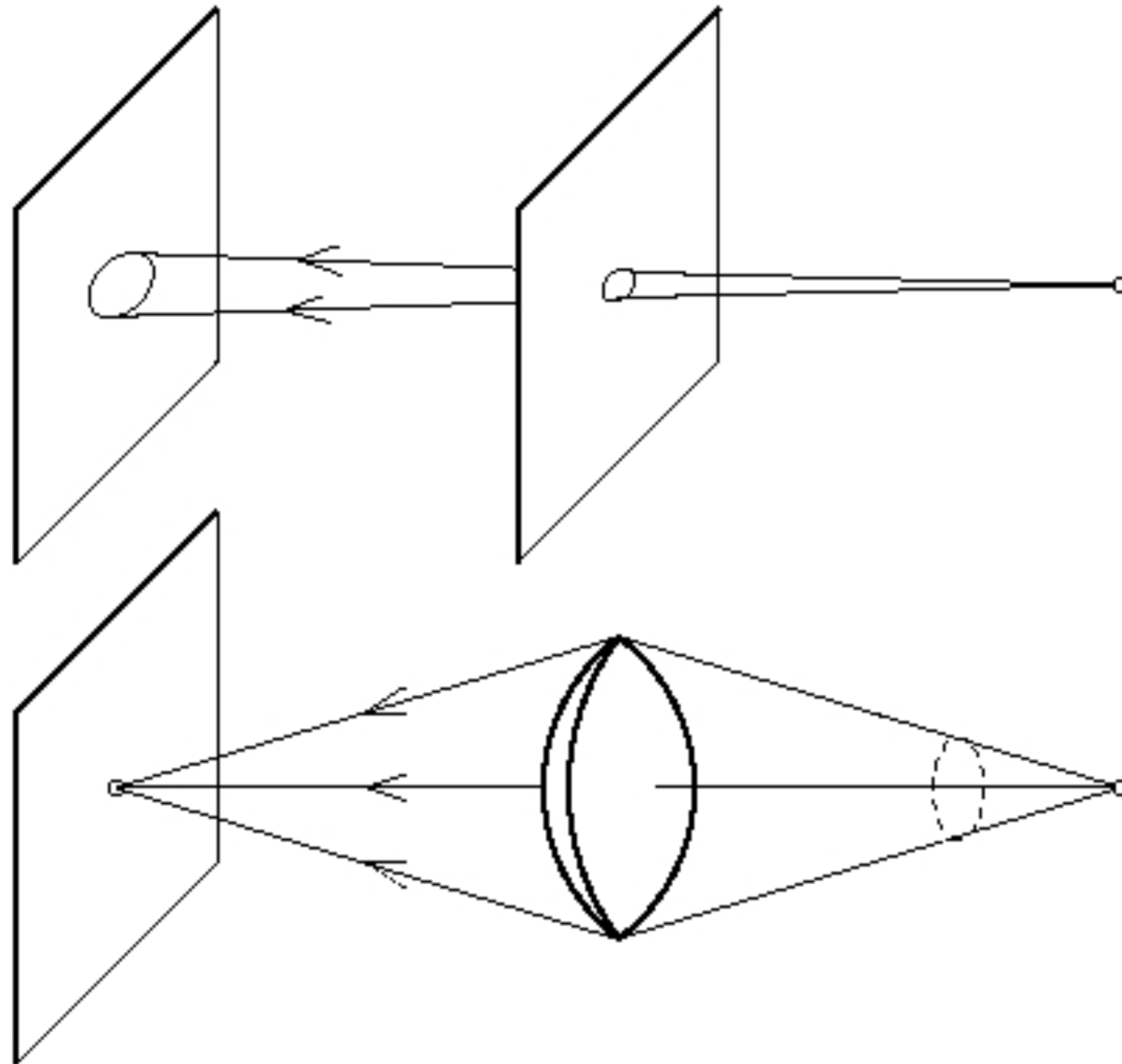
Snell's Law



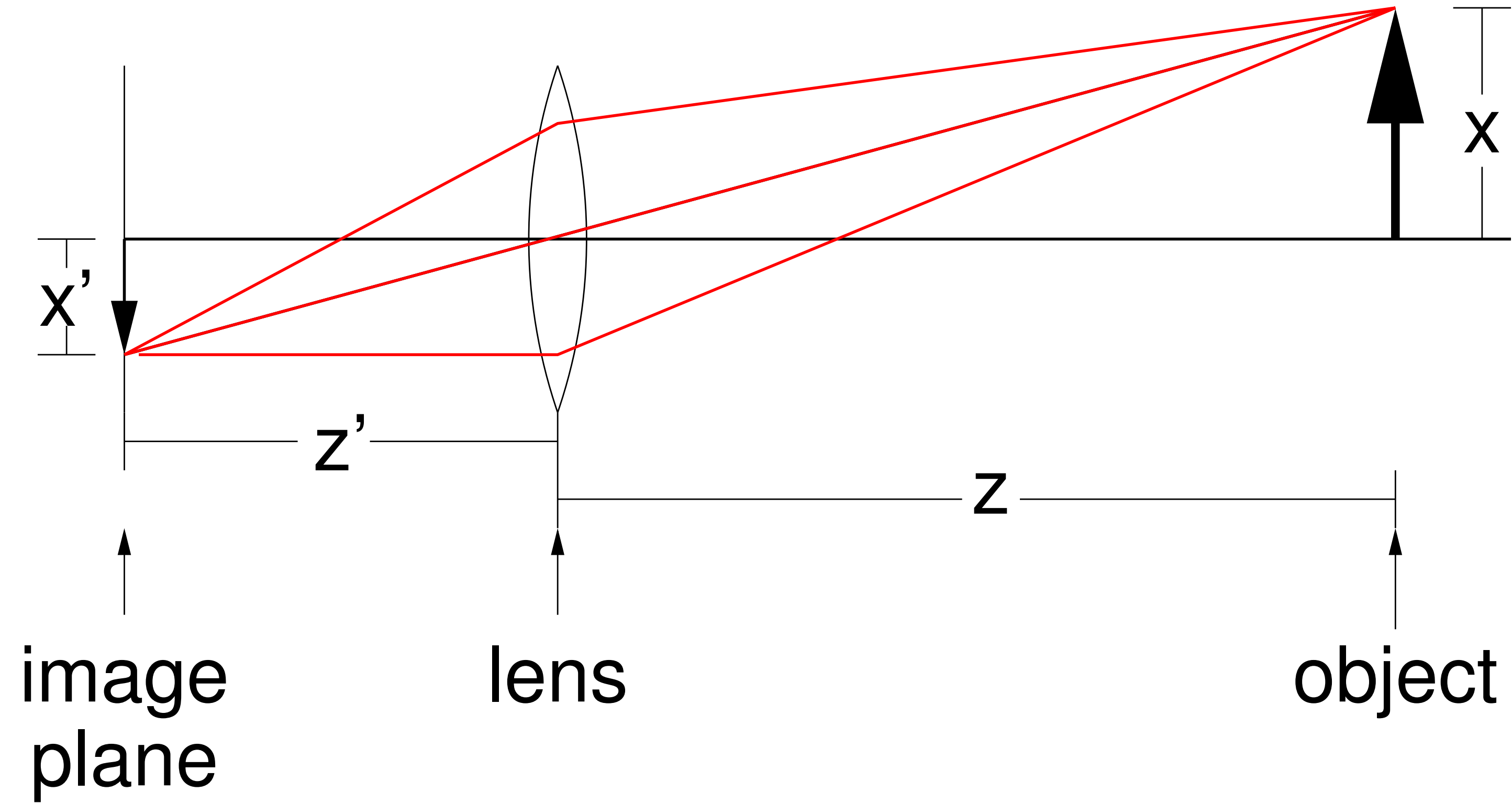
$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$$

Reason for **Lenses**

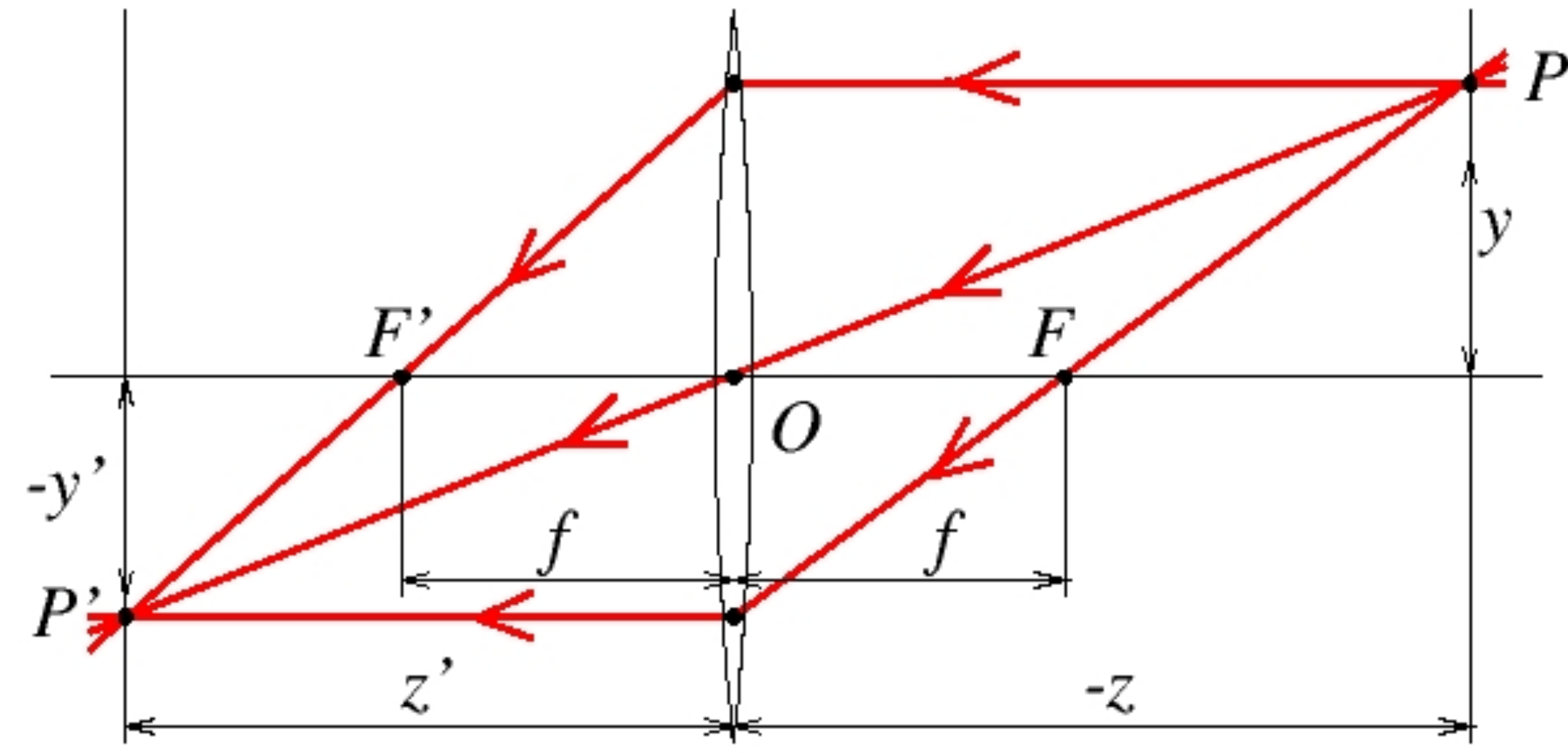
The role of a lens is to **capture more light** while preserving, as much as possible, the abstraction of an ideal pinhole camera.



Pinhole Model (Simplified) **with Lens**



Thin Lens Equation

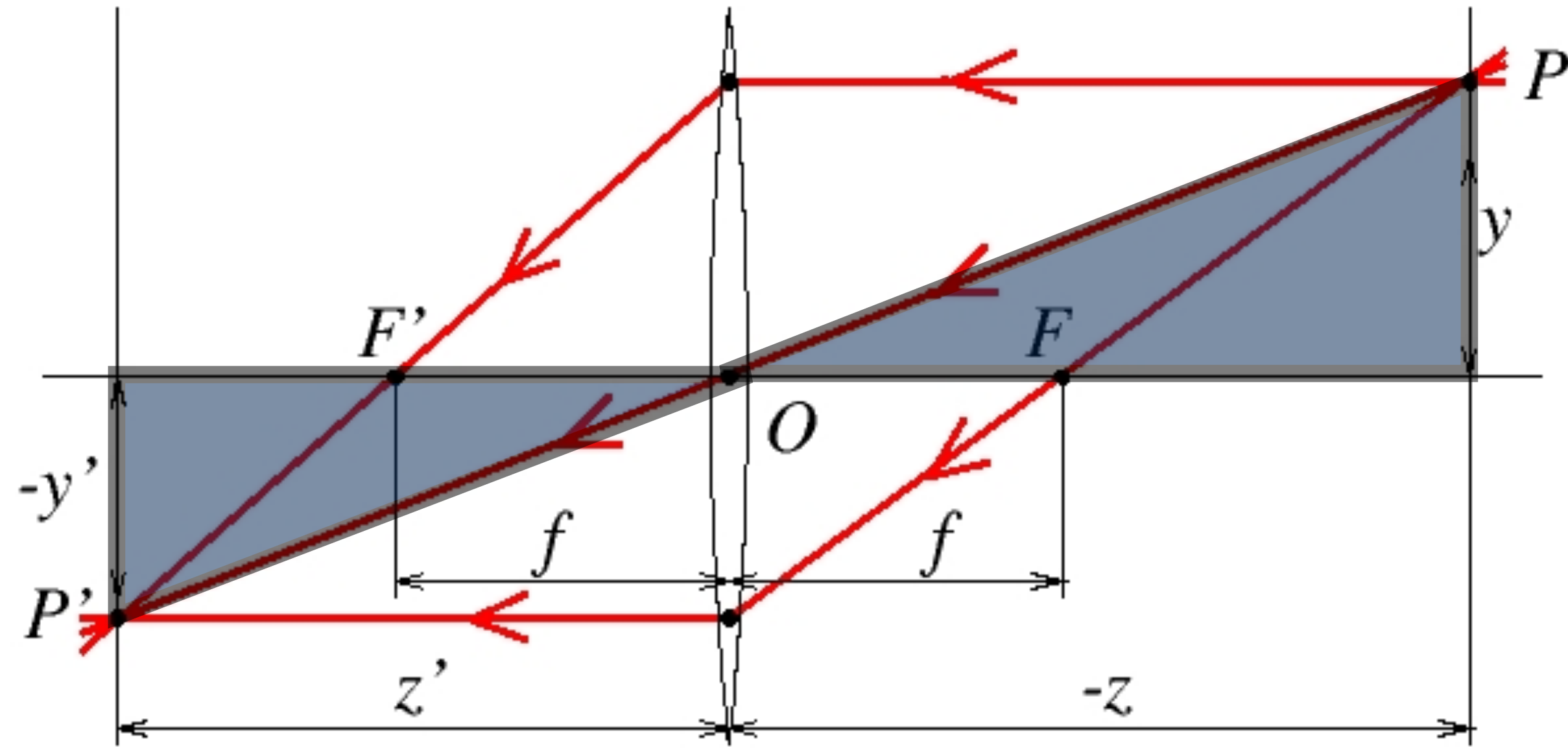


Forsyth & Ponce (1st ed.) Figure 1.9

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Thin Lens Equation: Derivation

$$\frac{y}{-z} = \frac{-y'}{z'}$$
$$\frac{y}{y'} = \frac{z}{z'}$$



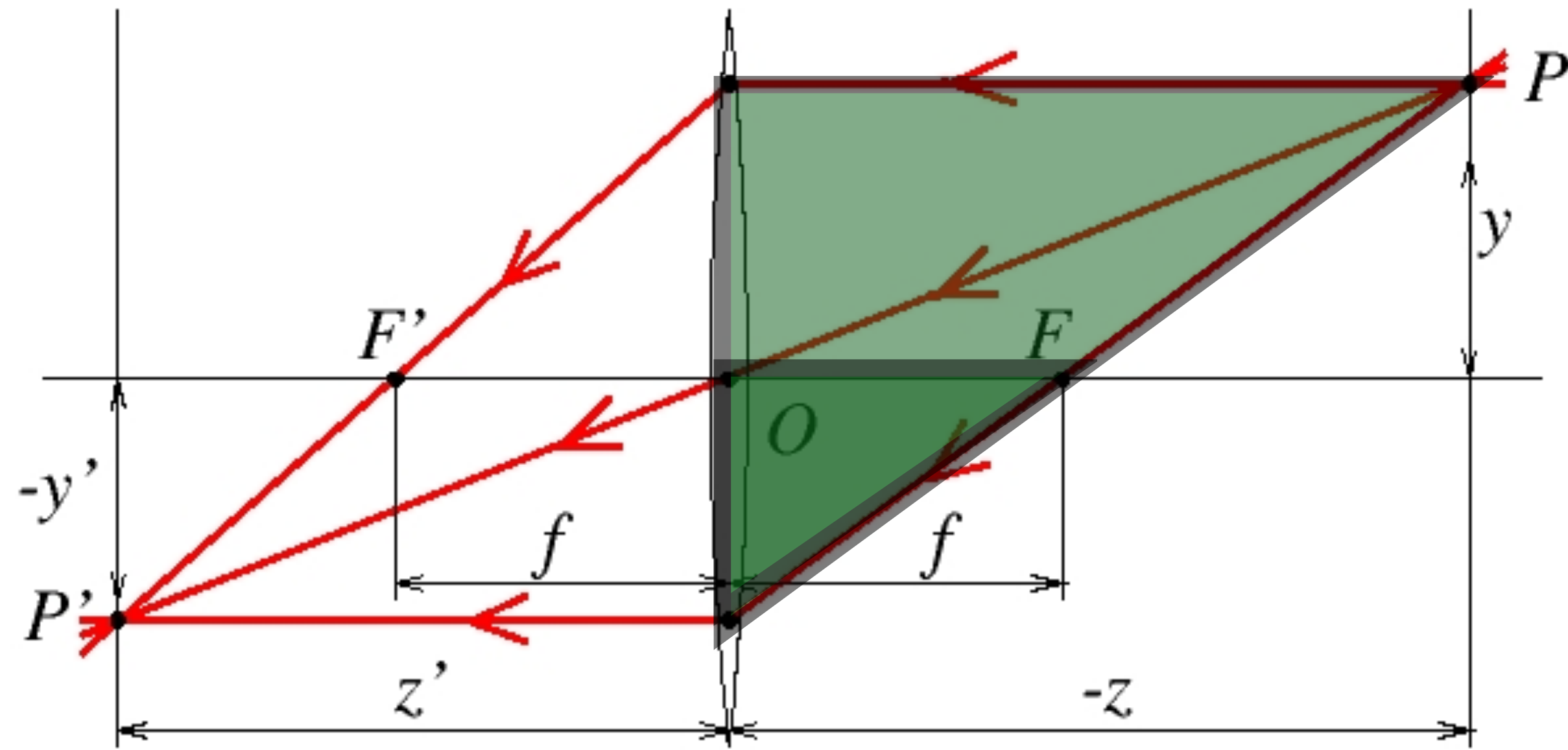
Forsyth & Ponce (1st ed.) Figure 1.9

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Thin Lens Equation: Derivation

$$\frac{y}{-z} = \frac{-y'}{z'}$$

$$\frac{y}{y'} = \frac{z}{z'}$$



Forsyth & Ponce (1st ed.) Figure 1.9

$$\frac{-y'}{f} = \frac{y - y'}{-z}$$

$$\frac{1}{f} = \frac{y - y'}{zy'}$$

$$= \frac{y}{zy'} - \frac{y'}{zy'}$$

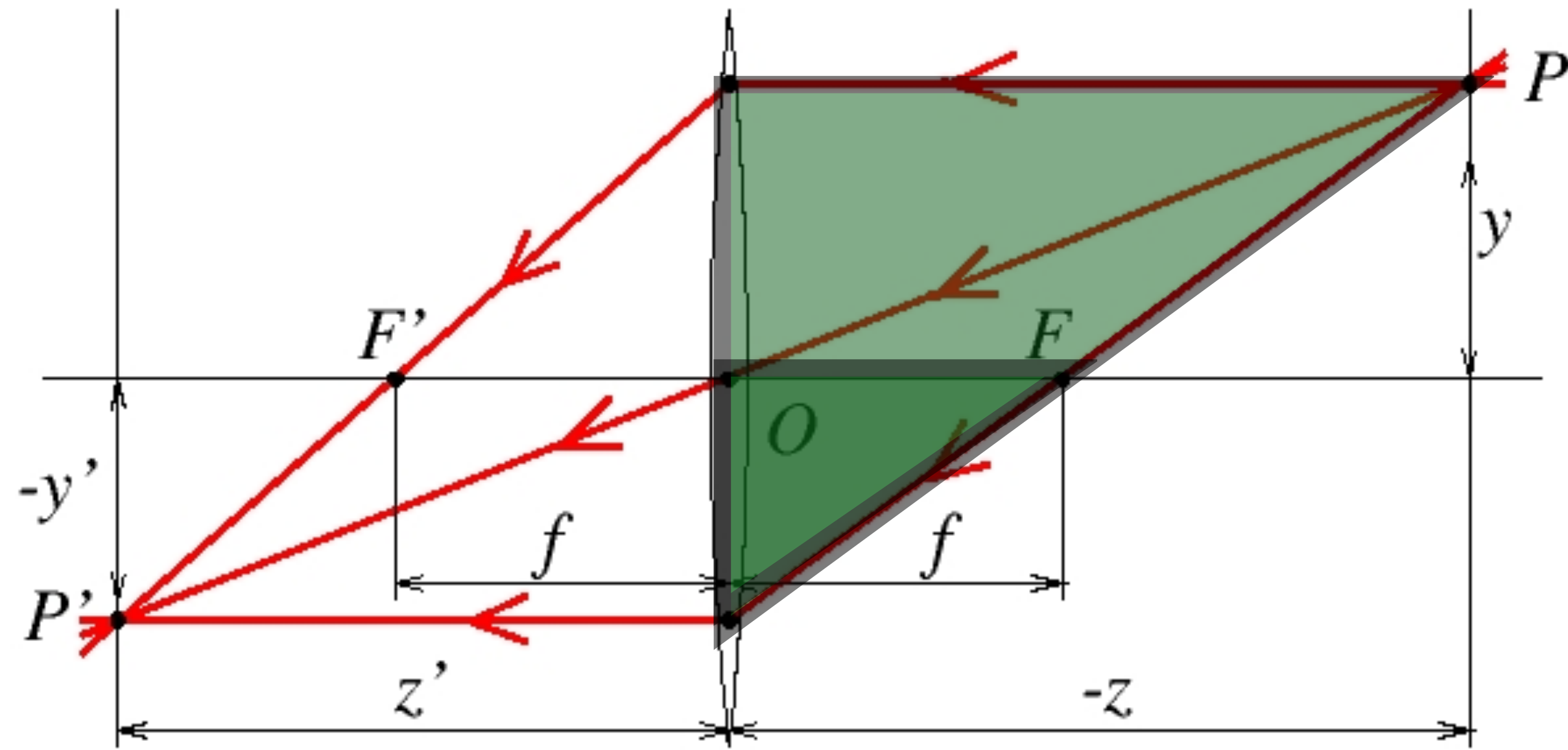
$$= \frac{y}{zy'} - \frac{1}{z}$$

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Thin Lens Equation: Derivation

$$\frac{y}{-z} = \frac{-y'}{z'}$$

$$\frac{y}{y'} = \frac{z}{z'}$$



Forsyth & Ponce (1st ed.) Figure 1.9

$$\frac{-y'}{f} = \frac{y - y'}{-z}$$

$$\frac{1}{f} = \frac{y - y'}{zy'}$$

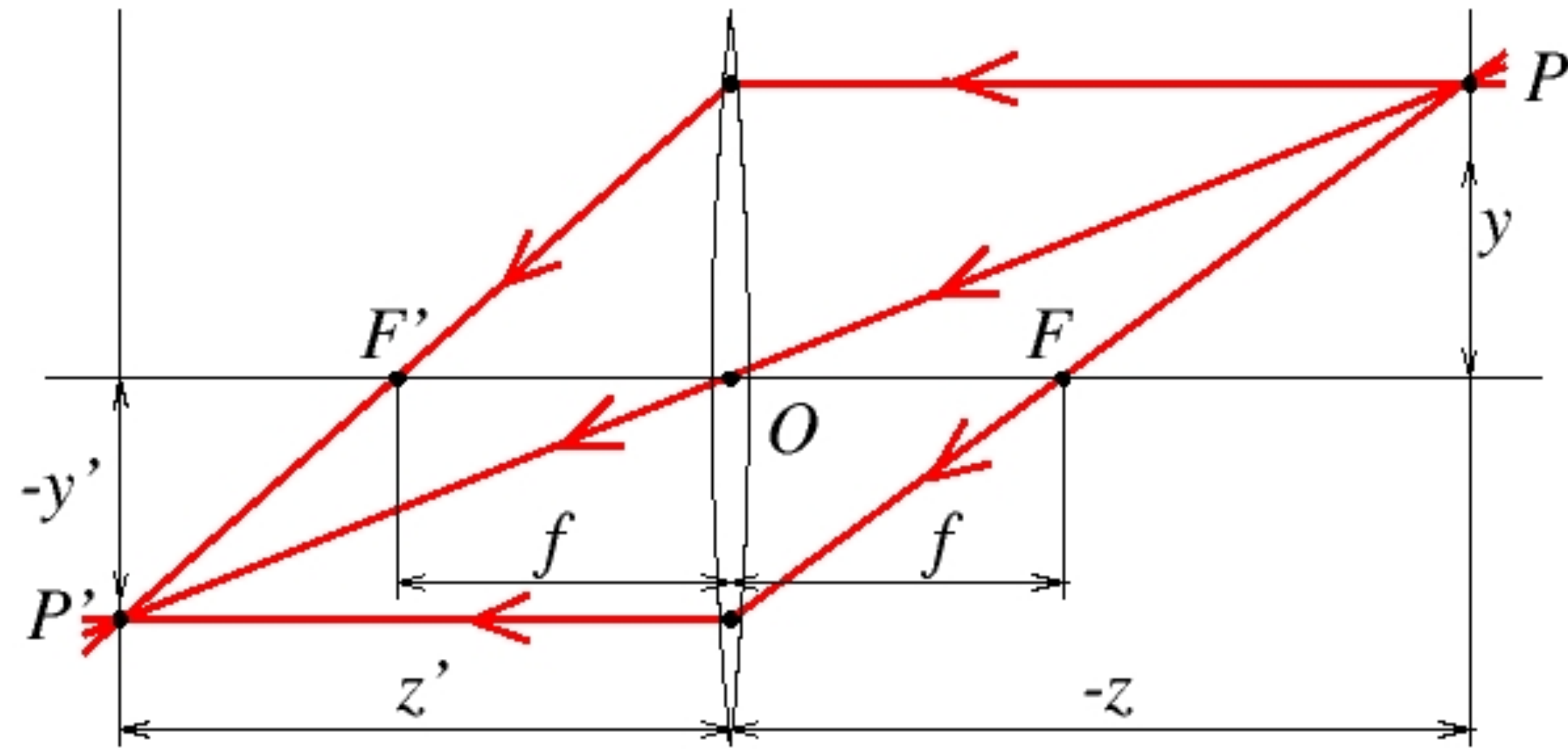
$$= \frac{y}{zy'} - \frac{y'}{zy'}$$

$$= \frac{y}{zy'} - \frac{1}{z}$$

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Substitute: $\frac{1}{f} = \frac{1}{z} - \frac{1}{z'}$

Possible Uses of Thin Lens Abstraction

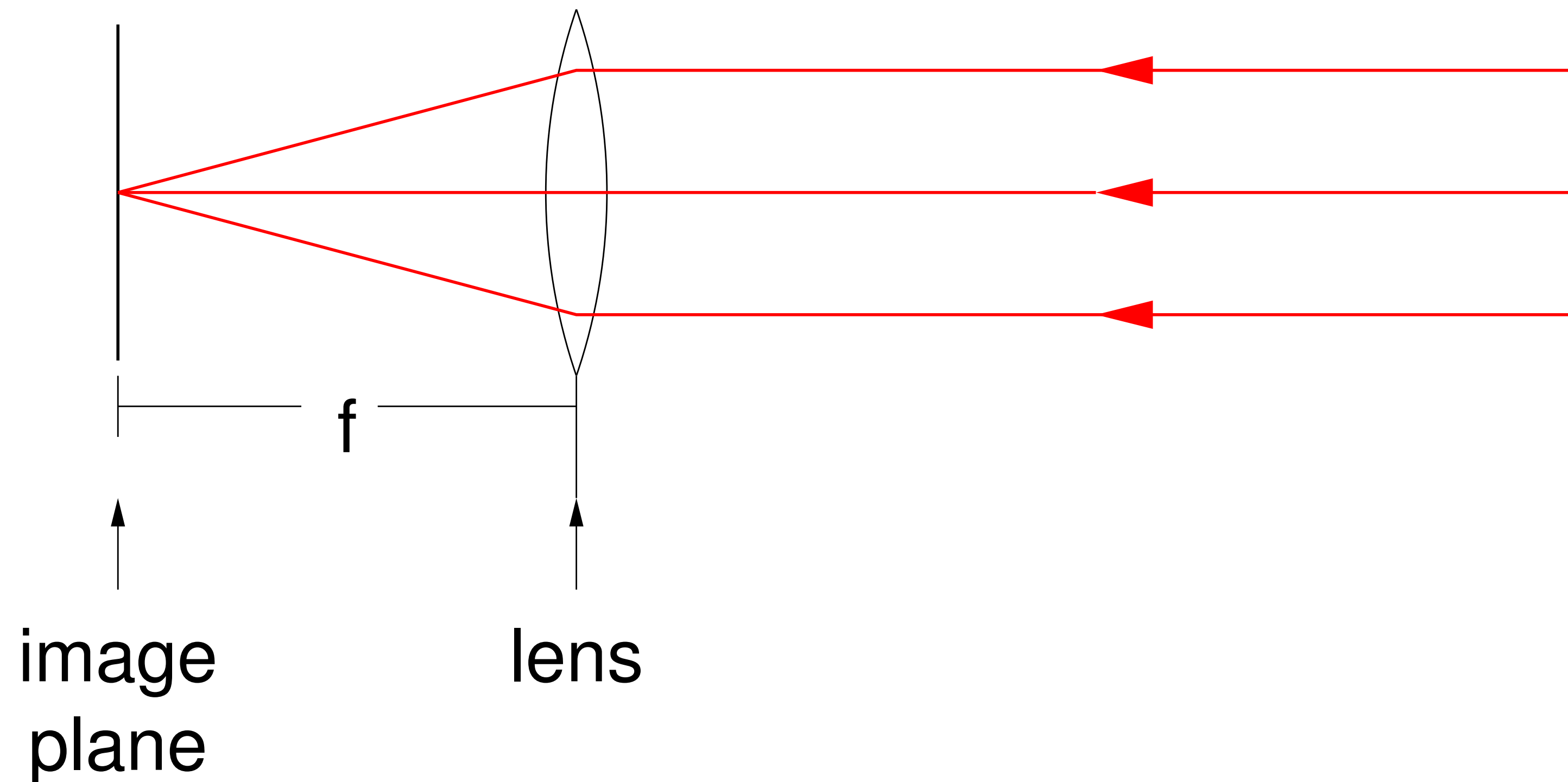


Forsyth & Ponce (1st ed.) Figure 1.9

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

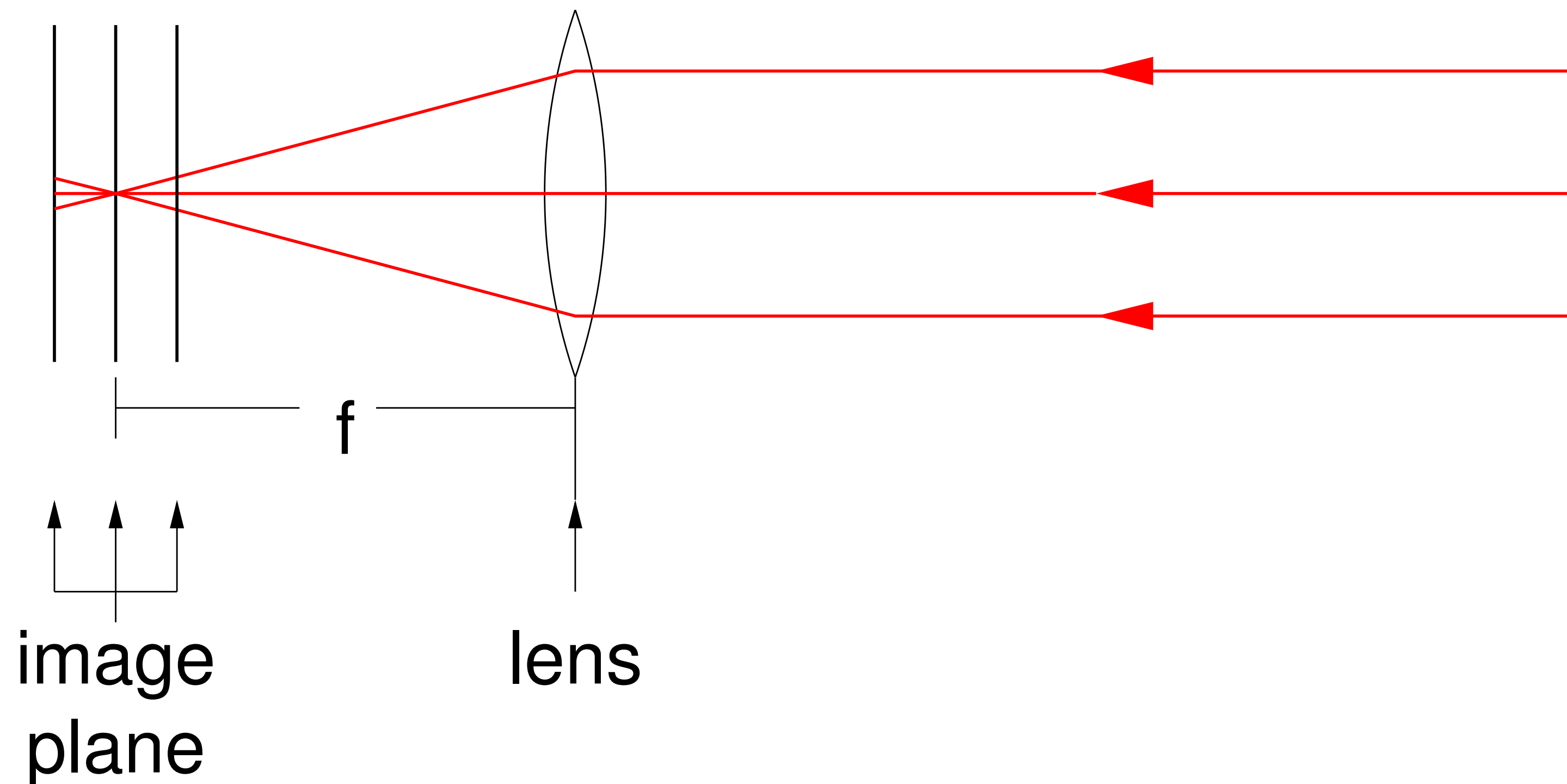
Focal Length

Another way of looking at the **focal length** of a lens. The incoming rays, parallel to the optical axis, **converge to a single point a distance f behind the lens**. This is where we want to place the image plane.

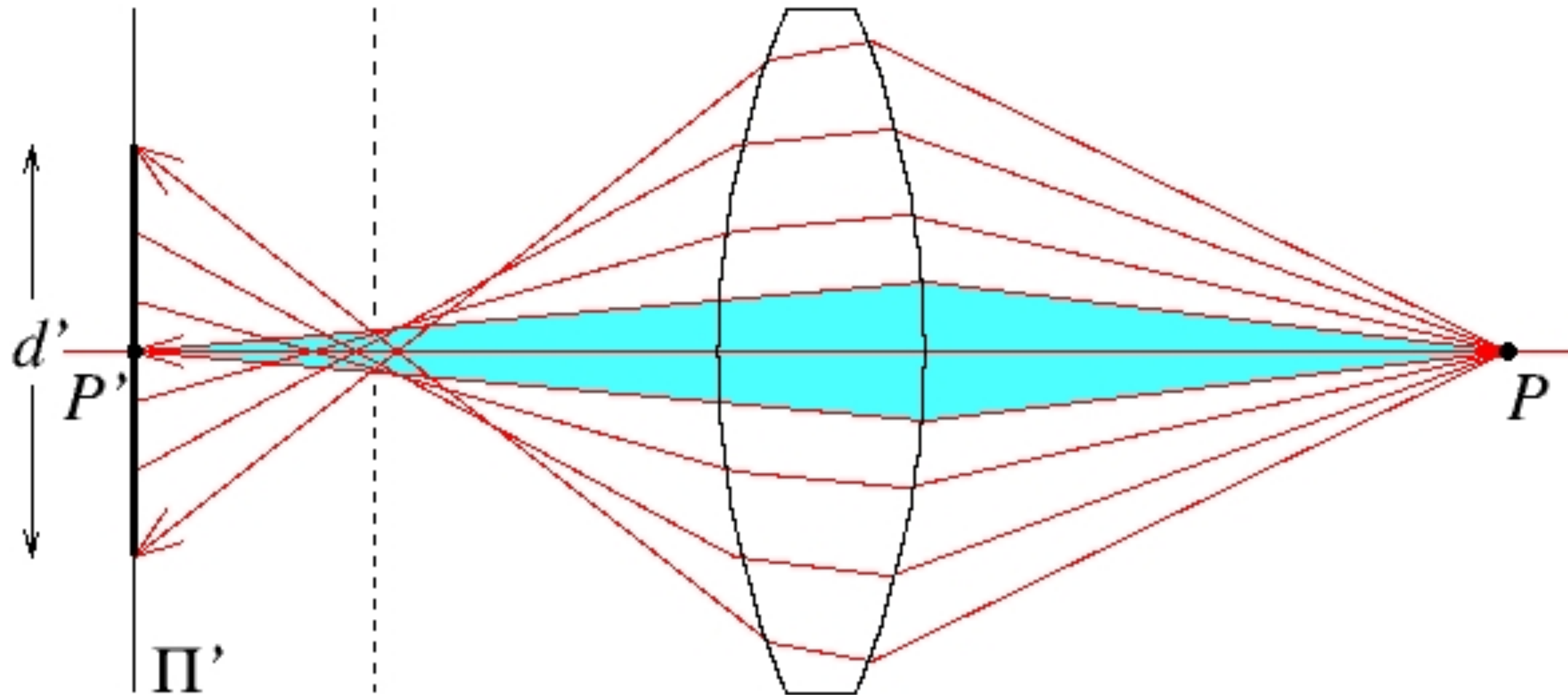


Out-of-Focus

The image plane is in the wrong place, either slightly closer than the required focal length, f , or slightly further than the required focal length, f .



Spherical Aberration



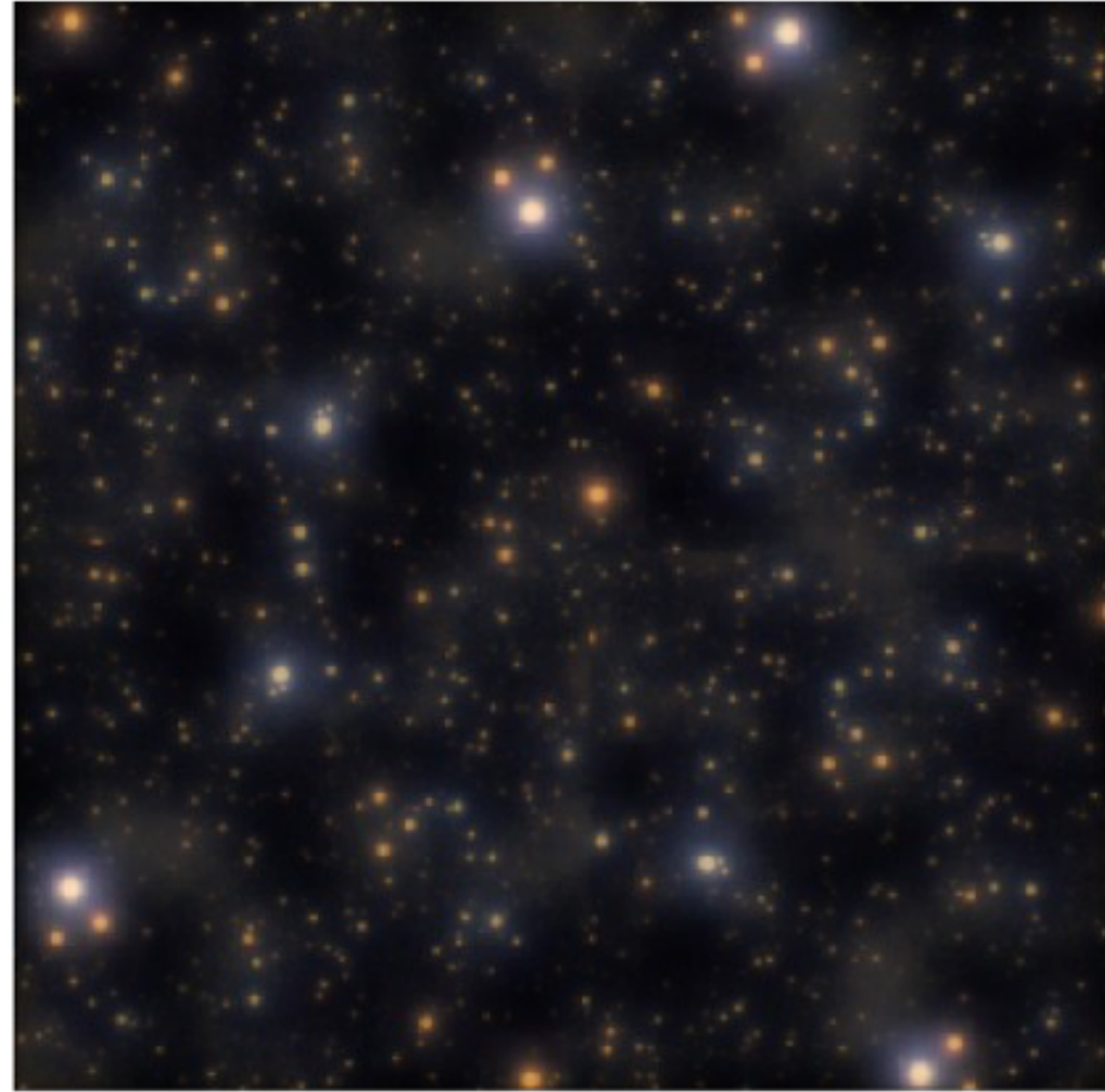
Forsyth & Ponce (1st ed.) Figure 1.12a

Spherical **Aberration**

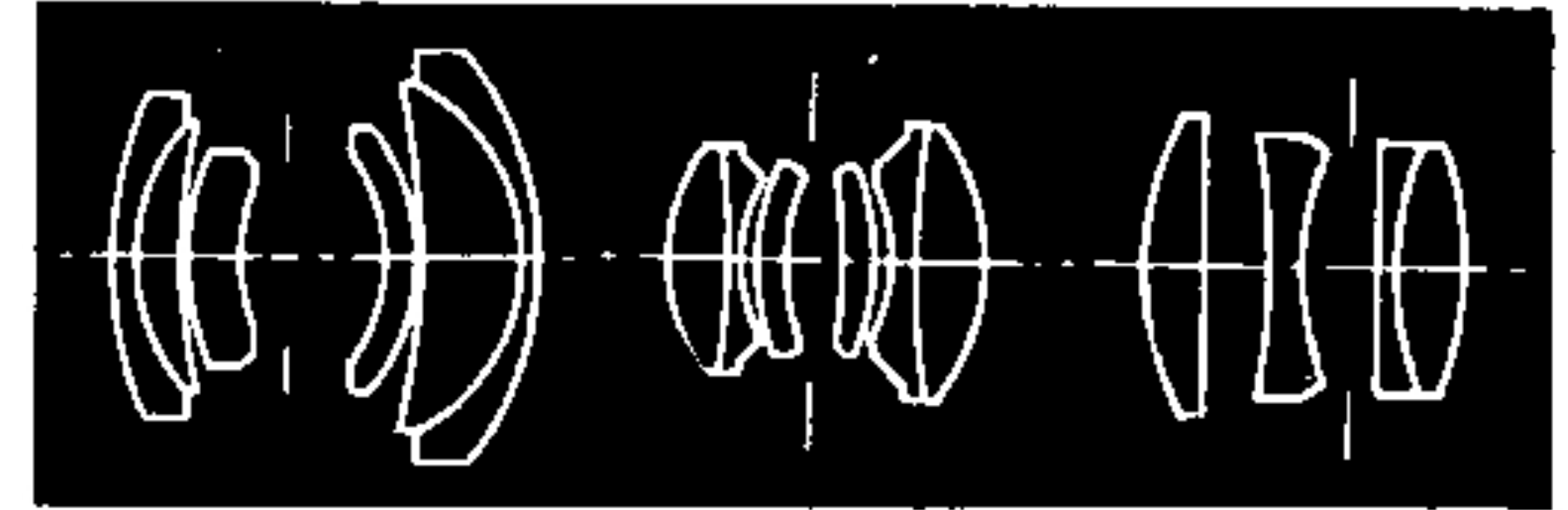
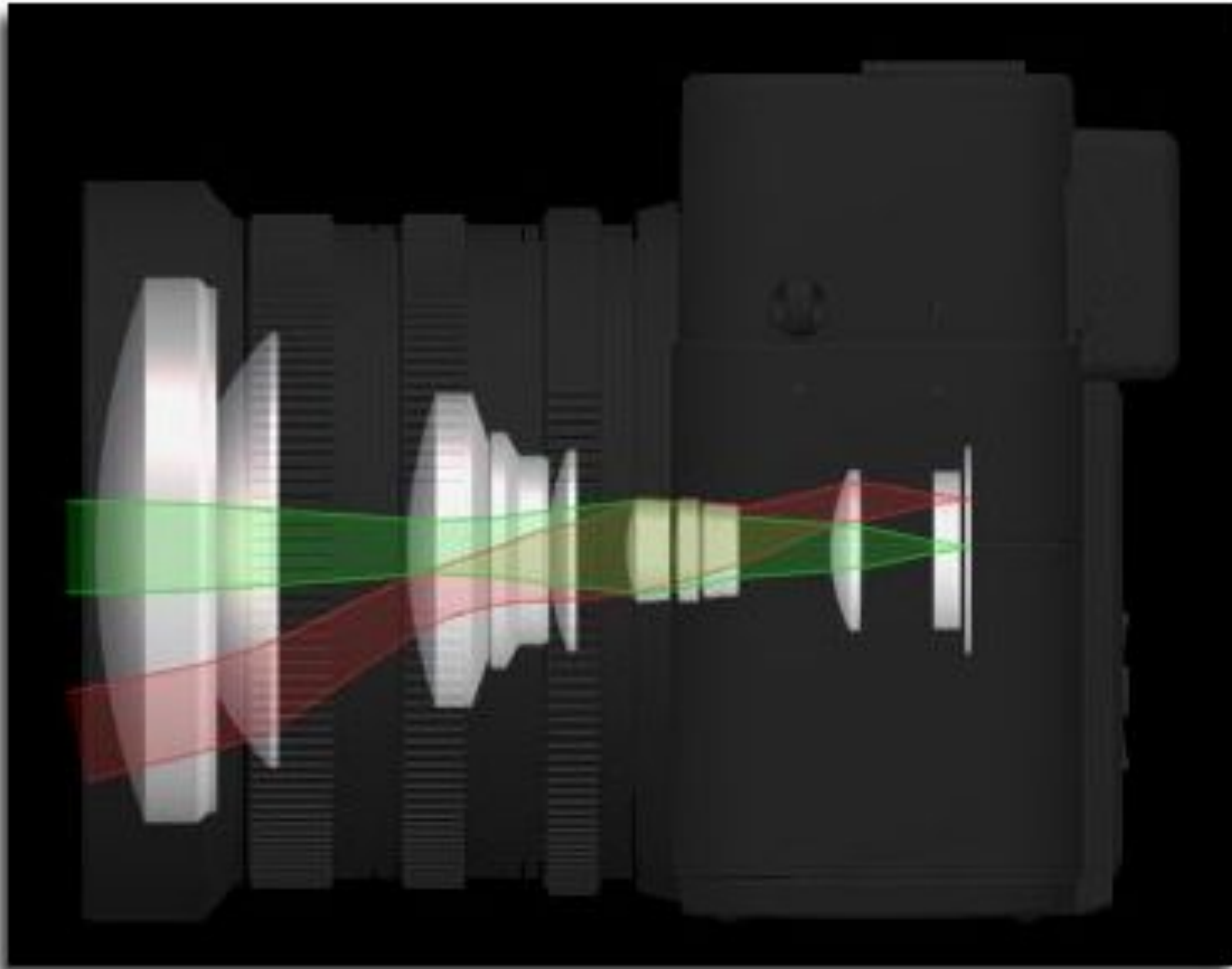
Un-aberrated image



Image from lens with Spherical Aberration



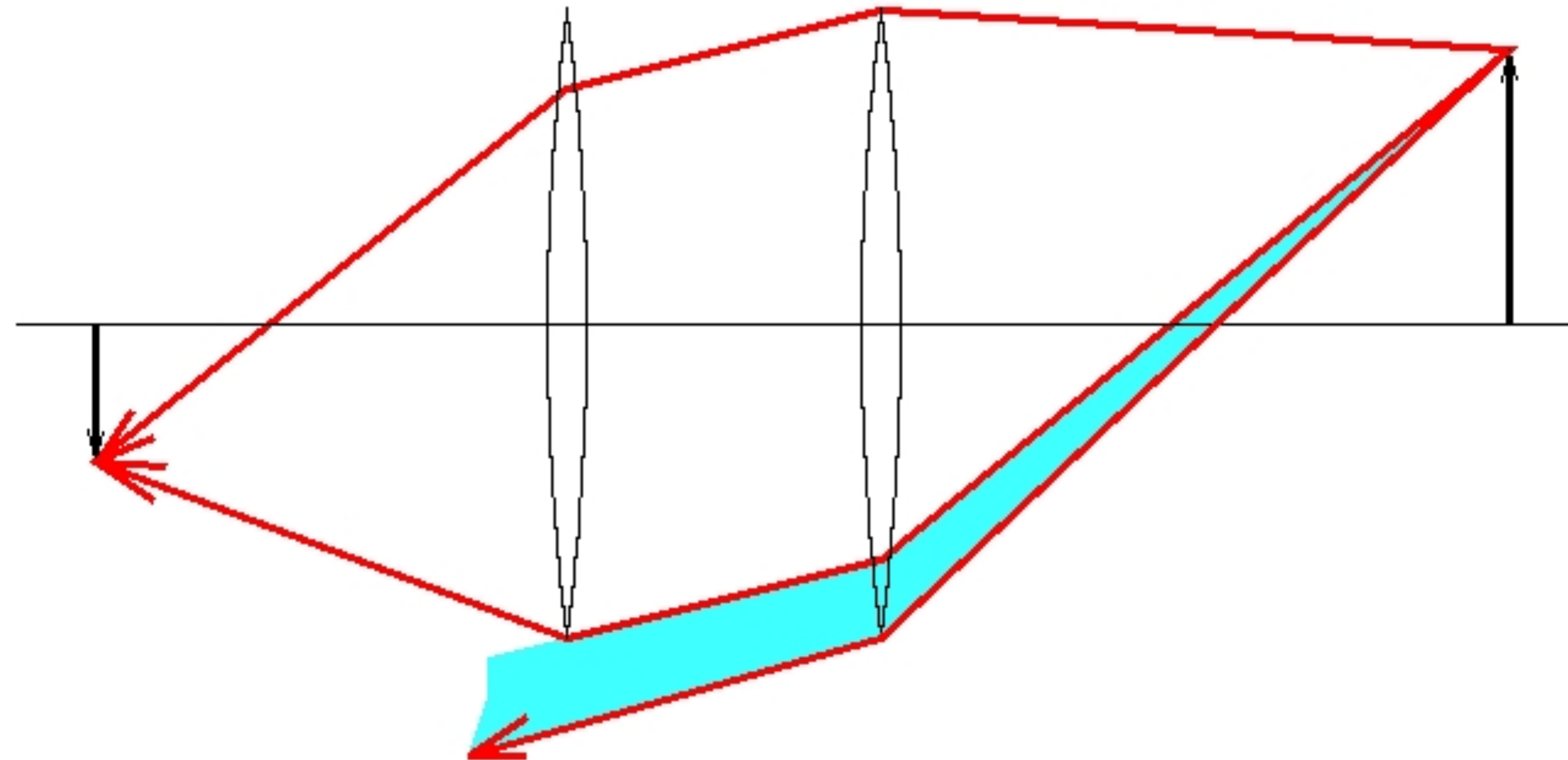
Compound **Lens Systems**



A modern camera lens may contain multiple components, including aspherical elements

Vignetting

Vignetting in a two-lens system



Forsyth & Ponce (2nd ed.) Figure 1.12

The shaded part of the beam **never reaches** the second lens

Vignetting



Chromatic Aberration

- Index of **refraction depends on wavelength**, λ , of light
- Light of different colours follows different paths
- Therefore, not all colours can be in equal focus

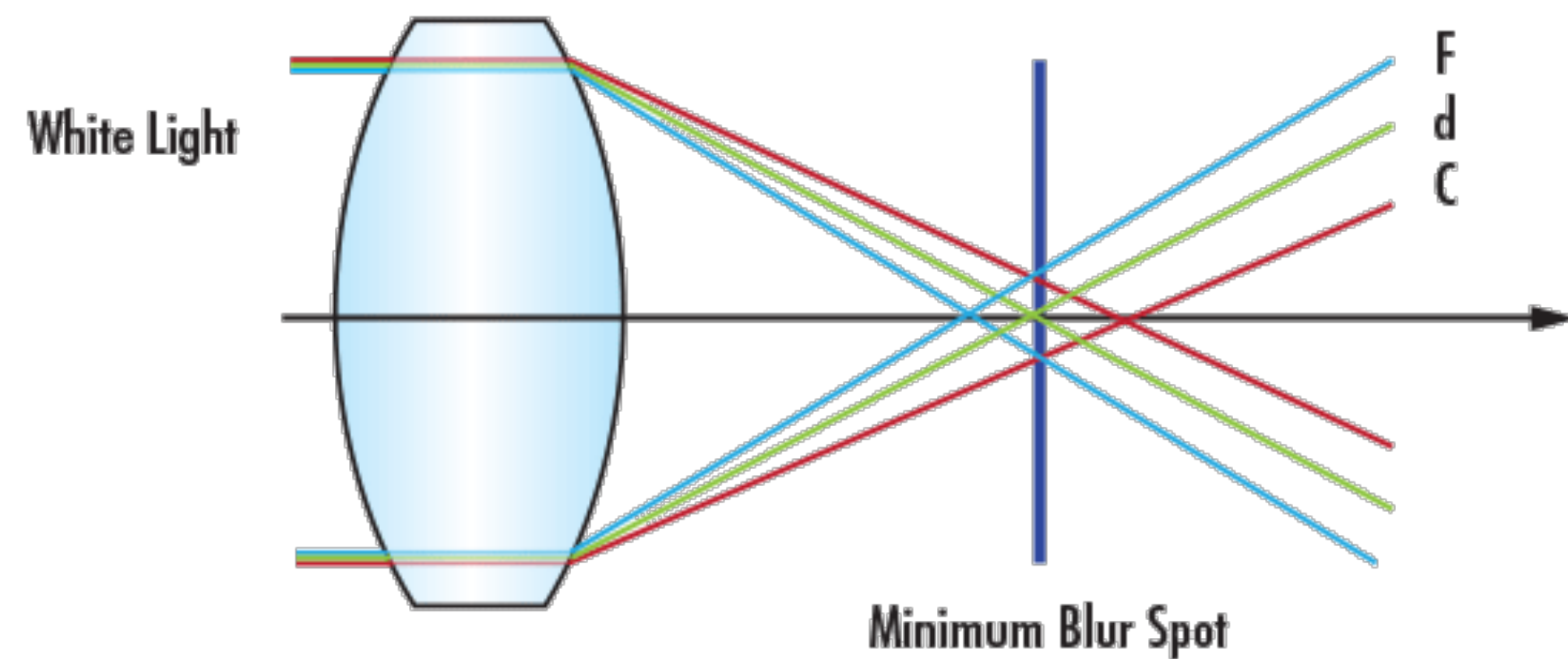


Image Credit: Trevor Darrell

Other (Possibly Significant) **Lens Effects**

Chromatic **aberration**

- Index of refraction depends on wavelength, λ , of light
- Light of different colours follows different paths
- Therefore, not all colours can be in equal focus

Scattering at the lens surface

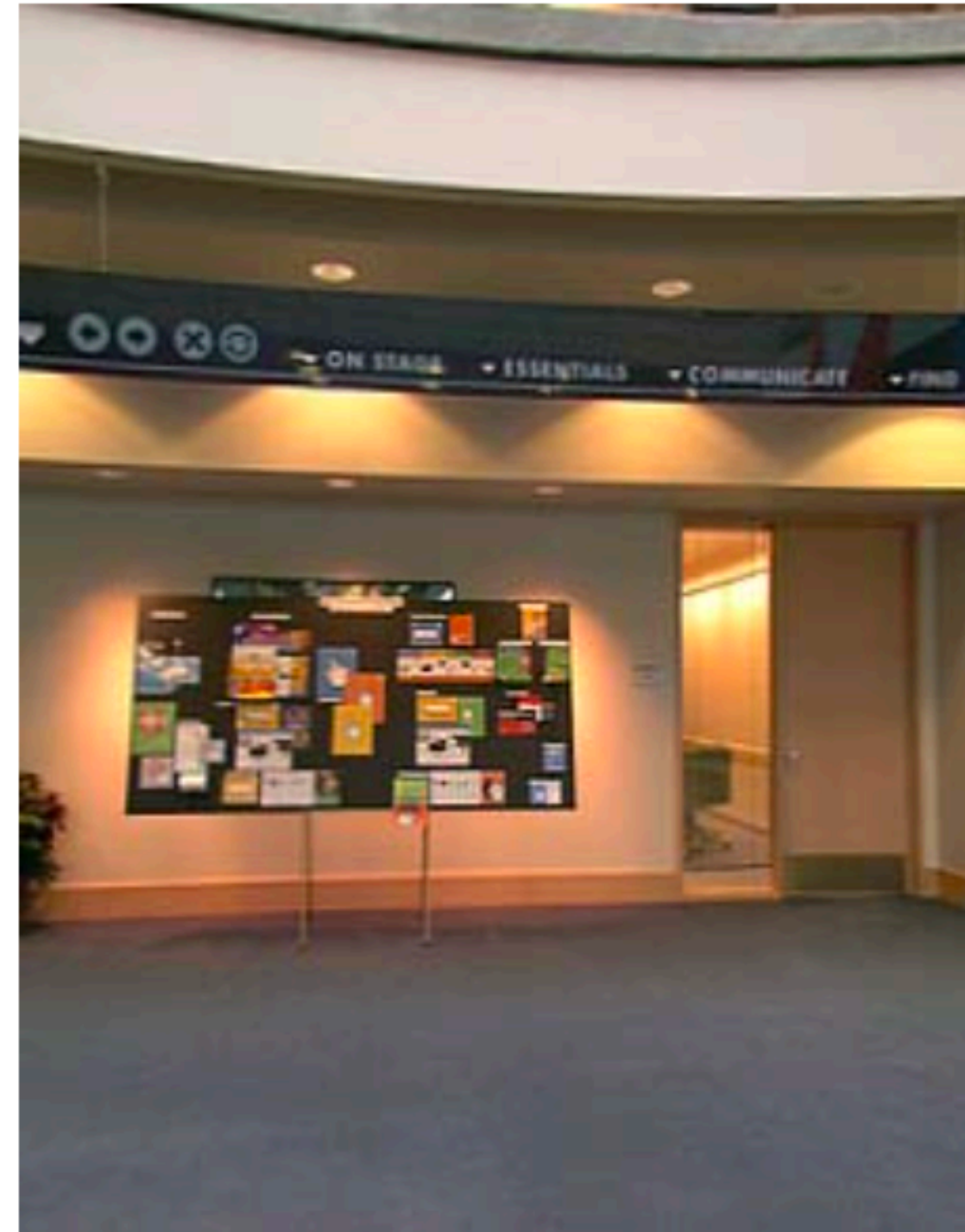
- Some light is reflected at each lens surface

There are other **geometric phenomena/distortions**

- pincushion distortion
- barrel distortion
- etc

Lens Distortion

Fish-eye Lens

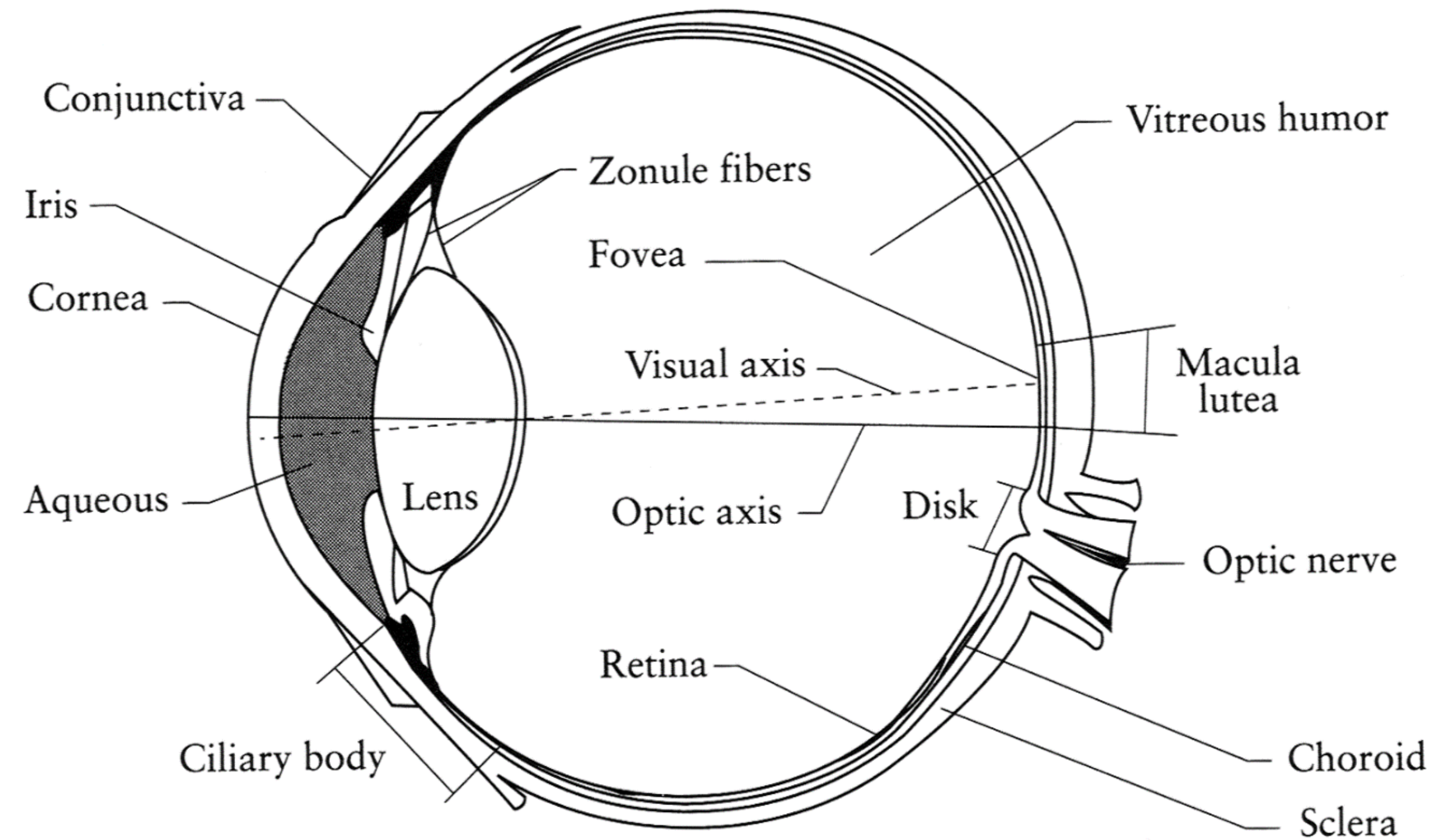


Szeliski (1st ed.) Figure 2.13

Lines in the world are no longer lines on the image, they are curves!

Human Eye

- The eye has an iris (like a camera)
- Focusing is done by changing shape of lens
- When the eye is properly focused, light from an object outside the eye is imaged on the retina
- The retina contains light receptors called rods and cones

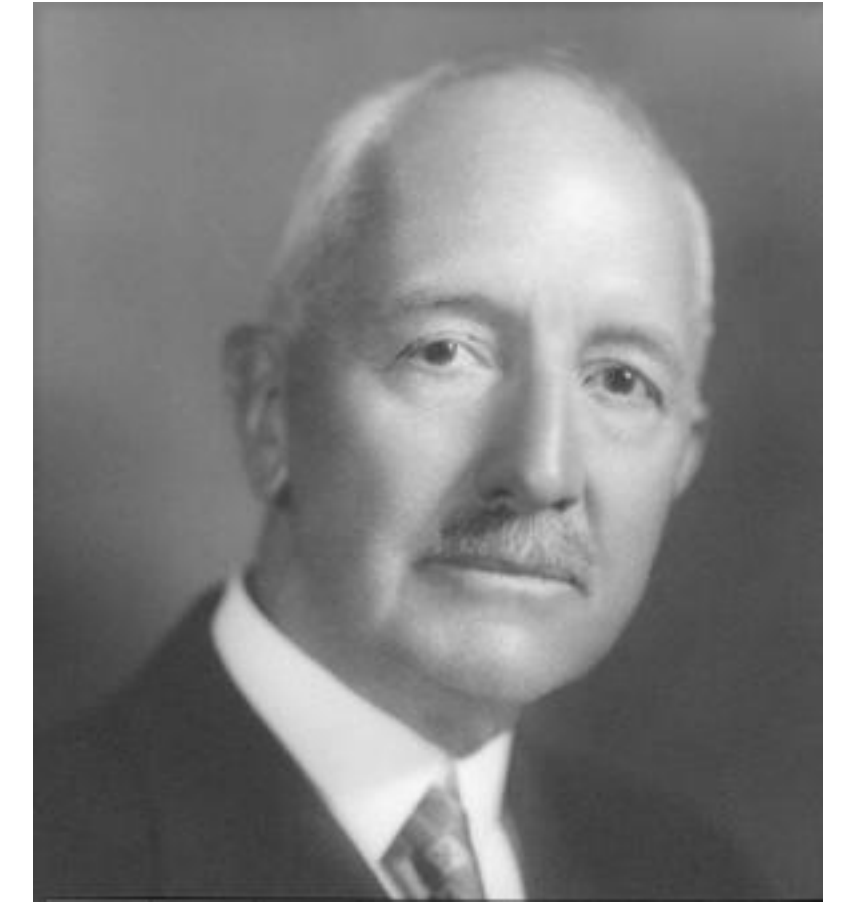


pupil = pinhole / aperture

retina = film / digital sensor

Slide adopted from: Steve Seitz

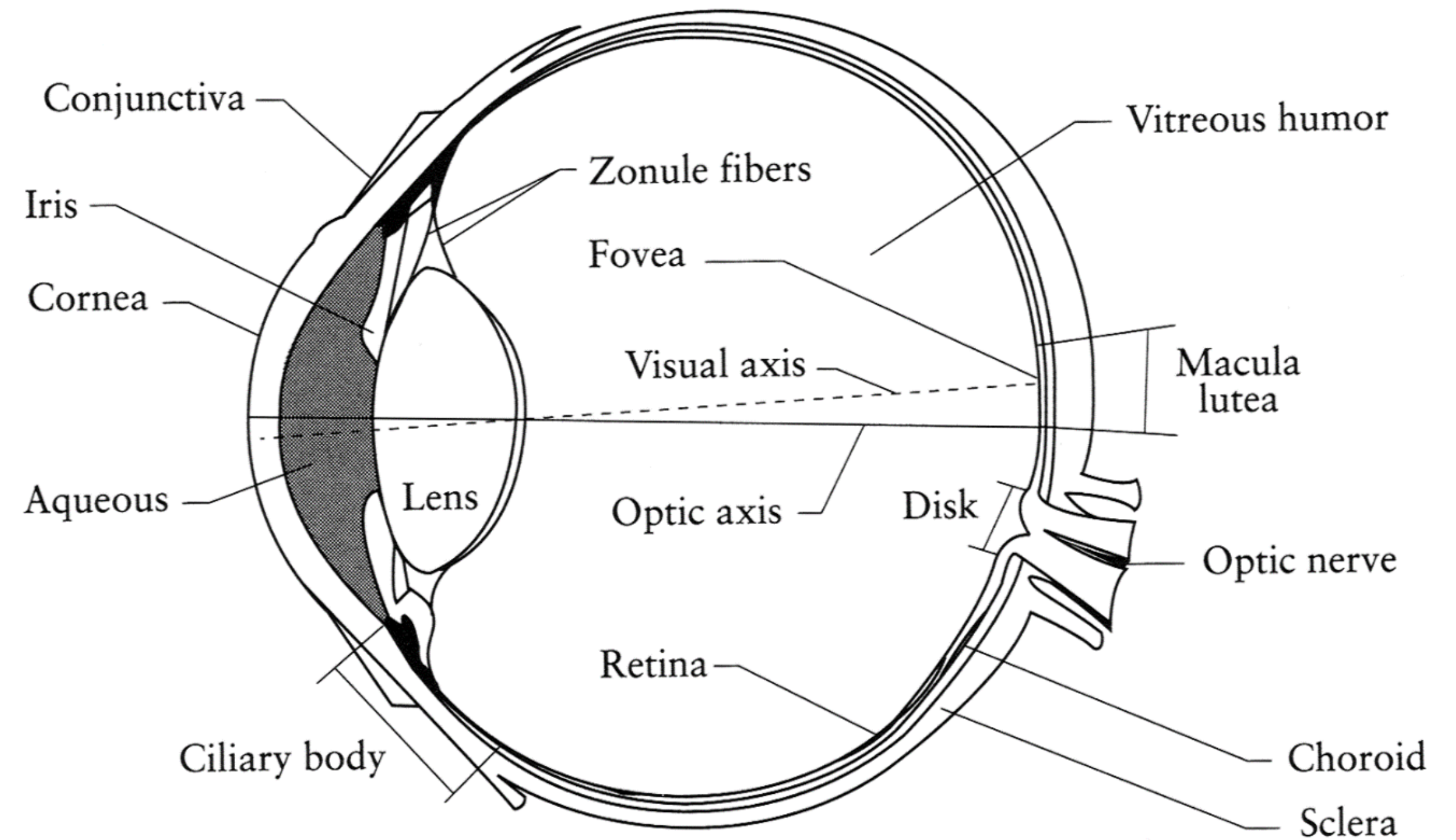
Fun **Aside**



George M. Stratton

Human Eye

- The eye has an iris (like a camera)
- Focusing is done by changing shape of lens
- When the eye is properly focused, light from an object outside the eye is imaged on the retina
- The retina contains light receptors called rods and cones



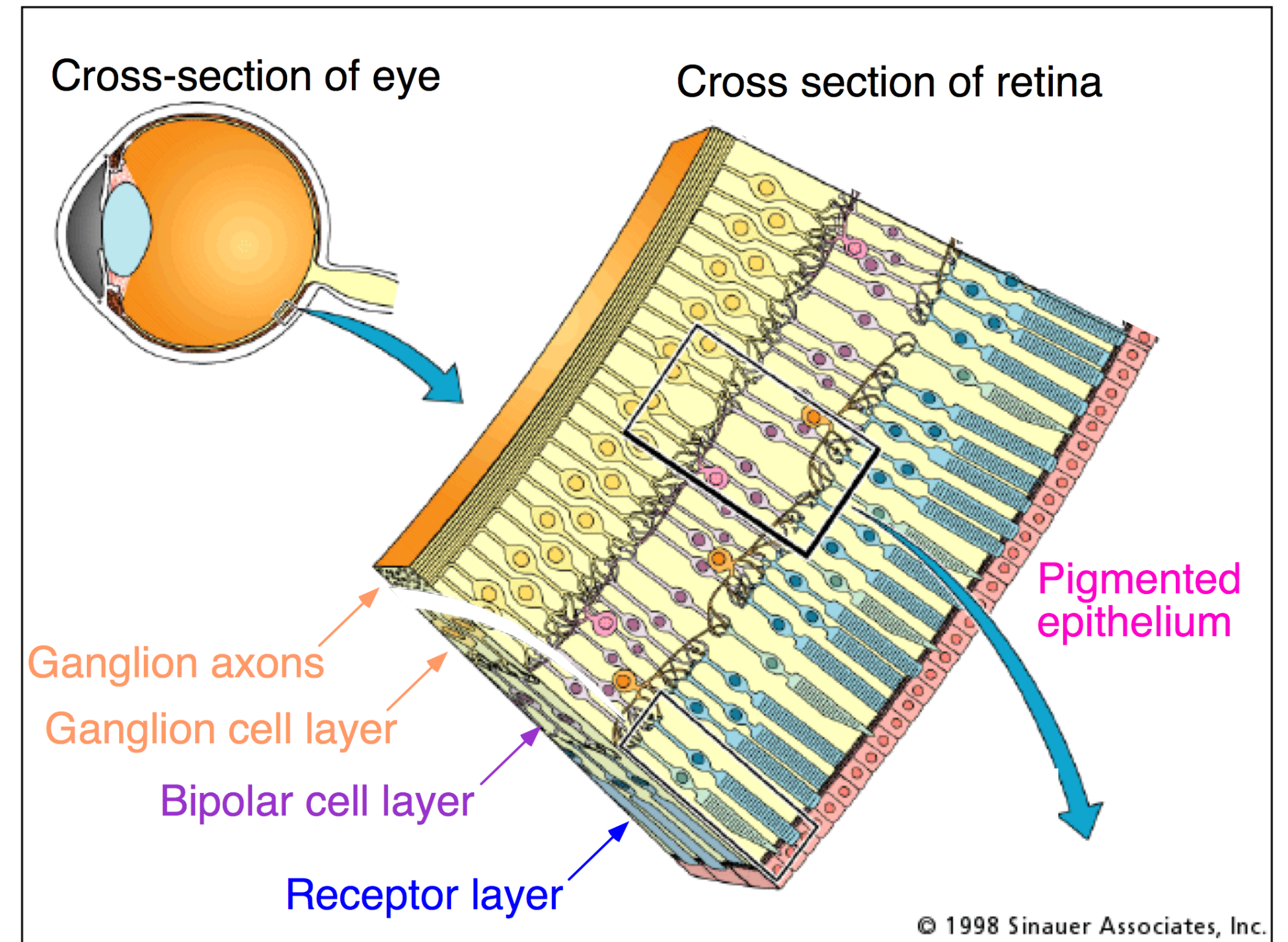
pupil = pinhole / aperture

retina = film / digital sensor

Slide adopted from: Steve Seitz

Human Eye

- The eye has an iris (like a camera)
- Focusing is done by changing shape of lens
- When the eye is properly focused, light from an object outside the eye is imaged on the retina.
- The retina contains light receptors called rods and cones



pupil = pinhole / aperture

retina = film / digital sensor

Slide adopted from: Steve Seitz

Two-types of **Light Sensitive Receptors**

Rods

75-150 million rod-shaped receptors

not involved in color vision, gray-scale vision only

operate at night

highly sensitive, can responding to a single photon

yield relatively poor spatial detail

Cones

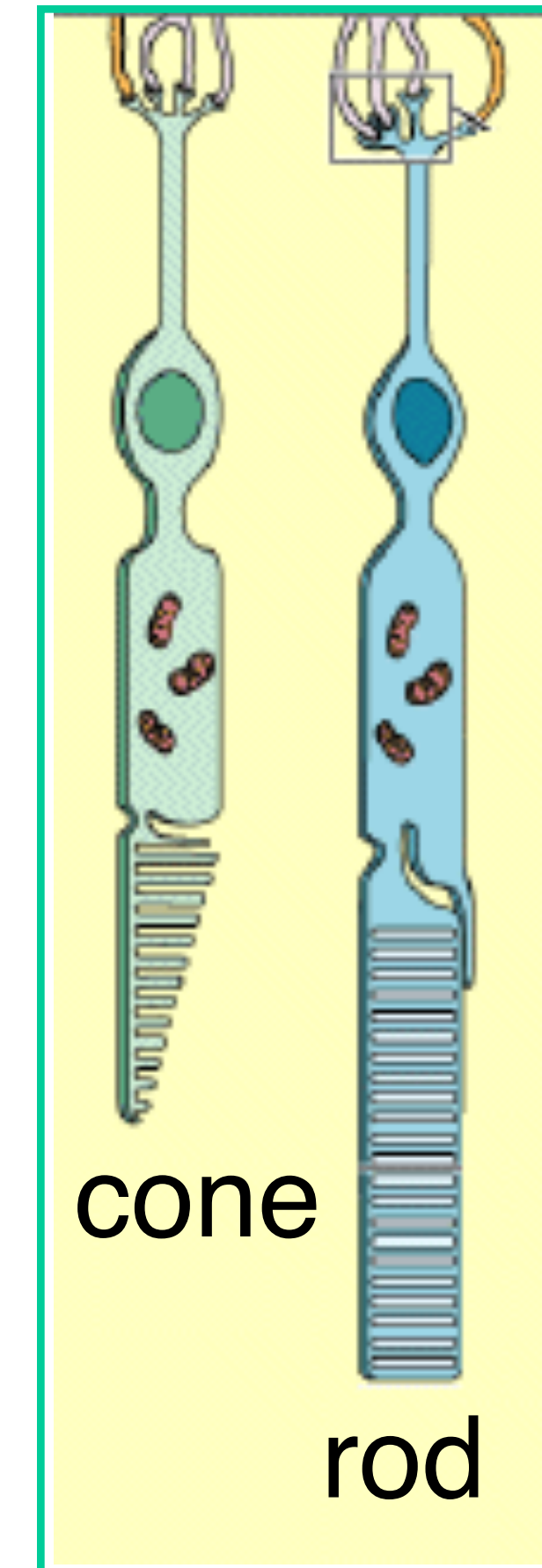
6-7 million cone-shaped receptors

color vision

operate in high light

less sensitive

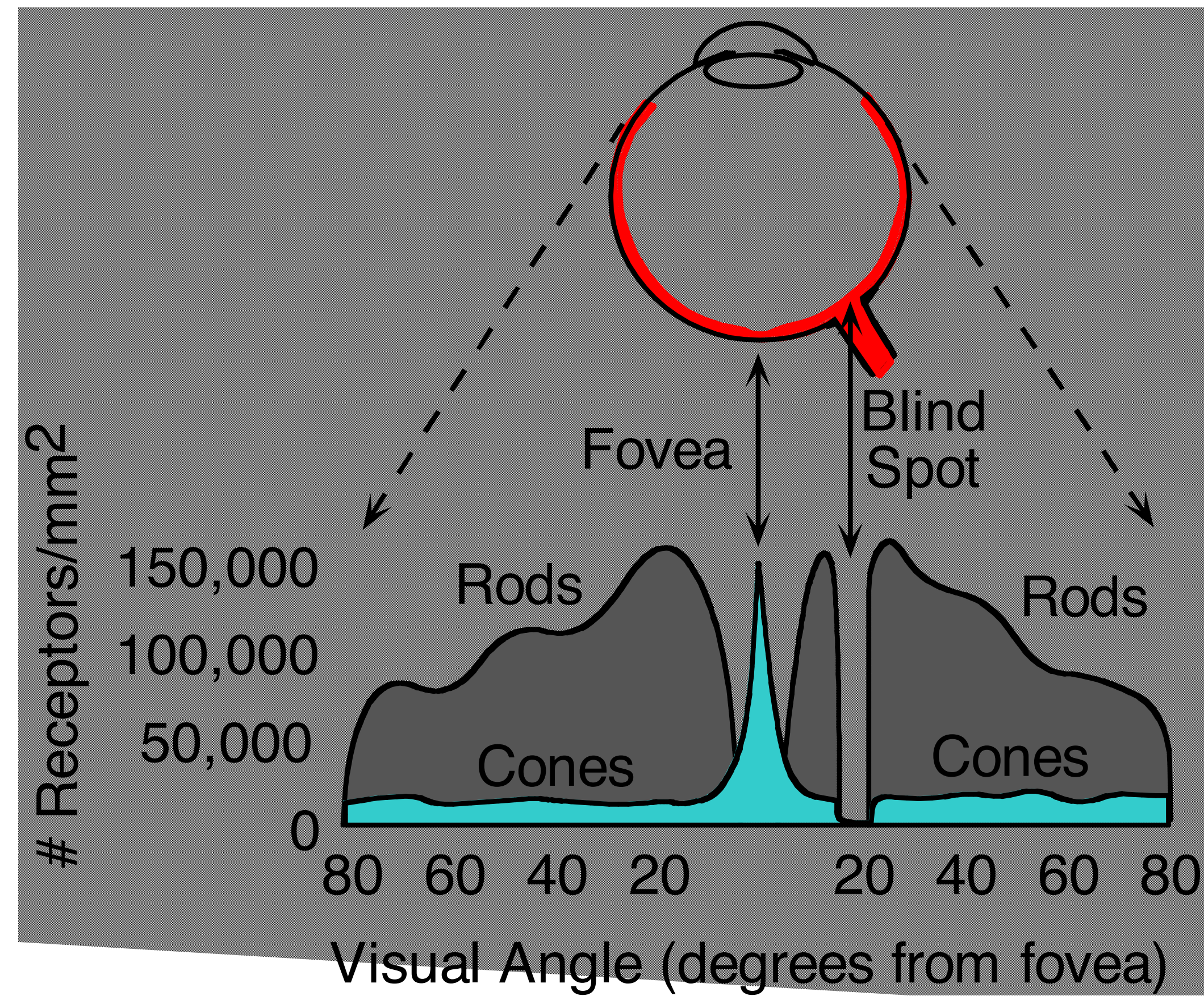
yield higher resolution



Slide adopted from: James Hays

Human Eye

Density of rods and cones



Lecture **Summary**

- We discussed a “physics-based” approach to image formation. Basic abstraction is the **pinhole camera**.
- **Lenses overcome limitations** of the pinhole model while trying to preserve it as a useful abstraction
- Projection equations: **perspective**, weak perspective, orthographic
- Thin lens equation
- Some “aberrations and **distortions**” persist (e.g. spherical aberration, vignetting)
- The **human eye** functions much like a camera

Reminders

Readings:

- **Today's** Lecture: Forsyth & Ponce (2nd ed.) 1.1.1 — 1.1.3
- **Next** Lecture: Forsyth & Ponce (2nd ed.) 4.1, 4.5

Reminders:

- Complete **Assignment 0** (ungraded) by Tuesday, **January 14**
- **WWW:** <http://www.cs.ubc.ca/~lsigal/teaching.html>
- **Piazza:** piazza.com/ubc.ca/winterterm22020/cpsc425201/home