EECS 4422/5323 Computer Vision

Topic 1: Image Formation

Calden Wloka

September 9, 2019

Finding the Website

- For anyone who hasn't found the course website yet, it is: `https://wiki.eecs.yorku.ca/course_archive/2019-20/F/4422/`
- As a general point of interest, course websites through EECS are usually easily found through the course archive: `https://www.eecs.yorku.ca/course_archive/`
  - Navigate to the current session (academic year), semester (*e.g.* F or W), and then find the course
  - Cross-listed courses (like this one) are usually listed by undergraduate course code

Labs Start Today!

Your labs start today (or Wednesday, depending on which section you are in)!

- Labs will usually be run by the course TA, Abdullah
- This week's lab will review useful tools and resources for this course
  - OpenCV
  - LaTeX

Abdullah the course TA

Lecture Outline

- Light Interacting with Surfaces
- Image Capture
- Optics
- Sensors
- Eyes

Light from Surfaces

Excluding light sources, image capture relies on surfaces reflecting light onto an image sensor.

The manner in which light is reflected (blue arrows) depends on a number of properties, including:

- The angle and intensity of the incoming light (red vector)

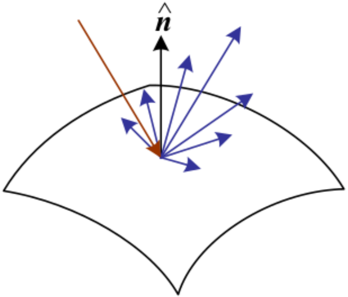


Image source: Szeliski, 2011

Light from Surfaces

Excluding light sources, image capture relies on surfaces reflecting light onto an image sensor.

The manner in which light is reflected (blue arrows) depends on a number of properties, including:

- The angle and intensity of the incoming light (red vector)
- The properties of the material being struck by light

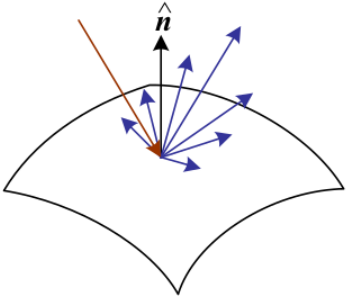


Image source: Szeliski, 2011

Light from Surfaces

Excluding light sources, image capture relies on surfaces reflecting light onto an image sensor.

The manner in which light is reflected (blue arrows) depends on a number of properties, including:

- The angle and intensity of the incoming light (red vector)
- The properties of the material being struck by light
- The orientation of the surface (represented mathematically as the surface normal)

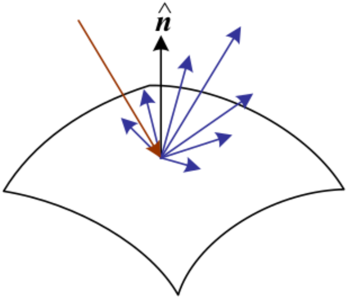


Image source: Szeliski, 2011

Modeling Light Scattering

Surface interactions can be very complicated!

For the purposes of this course, we will cover only a basic model with two primary forms of scattering:

- Diffuse
- Specular

If this topic interests you, computer graphics is a great way to delve deeper.

Image source: Original source unknown.

BRDFs

Bidirectional Reflectance Distribution Functions (BRDFs) are a class of functions which are used to characterize the interaction of light with a surface.

The general form of a BRDF is:

$$f_r(\theta_i, \phi_i, \theta_r, \phi_r)$$

where

- θ_i is the zenith angle and ϕ_i is the azimuth angle of incident light
- θ_r is the zenith angle and ϕ_r is the azimuth angle of outgoing light

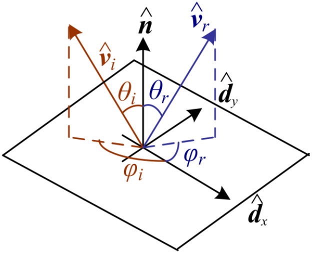


Image source: Szeliski, 2011

General BRDF Properties

There are three main properties of a realistic BRDF:

- Positivity: $f_r(\theta_i, \phi_i, \theta_r, \phi_r) \geq 0$ You can't have negative light.

General BRDF Properties

There are three main properties of a realistic BRDF:

- Positivity: $f_r(\theta_i, \phi_i, \theta_r, \phi_r) \geq 0$ You can't have negative light.
- Helmholtz reciprocity: $f_r(\theta_i, \phi_i, \theta_r, \phi_r) = f_r(\theta_r, \phi_r, \theta_i, \phi_i)$ "a ray of light and its reverse ray encounter matched optical adventures" [Wikipedia]

General BRDF Properties

There are three main properties of a realistic BRDF:

- Positivity: $f_r(\theta_i, \phi_i, \theta_r, \phi_r) \geq 0$ You can't have negative light.
- Helmholtz reciprocity: $f_r(\theta_i, \phi_i, \theta_r, \phi_r) = f_r(\theta_r, \phi_r, \theta_i, \phi_i)$ "a ray of light and its reverse ray encounter matched optical adventures" [Wikipedia]
- Energy conservation: if you integrate over all outgoing rays, you can't end up with more light than you started with.

Diffuse Reflectance

Diffuse reflection (also known as *matte*) scatters light uniformly in all directions. A perfectly diffuse surface follows the *Lambertian* model of reflectance, which collapses the BRDF to a constant.

$$f_d(\theta_i, \phi_i, \theta_r, \phi_r) = \frac{\rho_d}{\pi}$$

where ρ_d is the surface albedo.

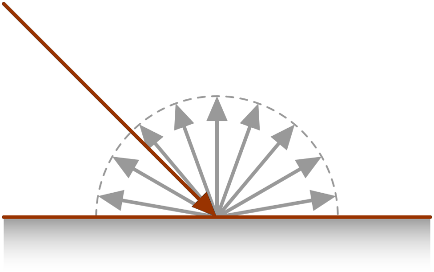


Image source: Modified from Wikipedia

Diffuse Reflectance as a Function of Incident Angle

The intensity of reflected light is still a function of the incident angle of the source.

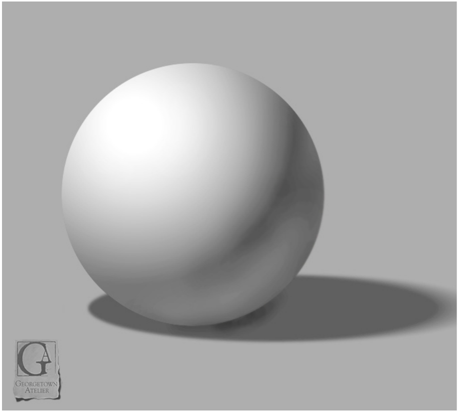


Image source: Georgetown Atelier

The surface radiance, L, is given by:

$L = \frac{\rho_d}{\pi} I \cos^+ \theta_i$

where I is the source intensity.

Mirror Reflection

A completely specular reflection occurs when all incident light for a given incoming source is reflected in a single outgoing direction rotated 180° about the azimuth.

The BRDF collapses to a delta function:

$$f_s(\theta_i, \phi_i, \theta_r, \phi_r) = \rho_s \delta(\theta_i - \theta_r)\delta(\phi_i + \pi - \phi_r)$$

where ρ_s is the specular albedo.

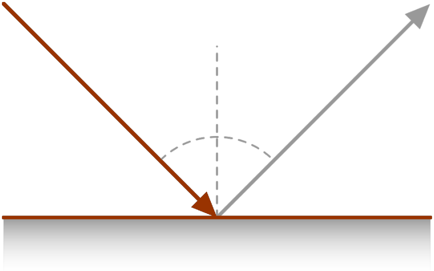


Image source: Modified from Wikipedia

This model is only truly valid for highly polished surfaces, such as mirrors.

Glossy Reflection

More common specular reflections are better represented as *glossy* reflections in which there is some scattering of the light, but it is still highly directional.

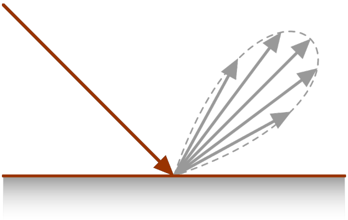


Image source: Modified from Wikipedia

There are a number of different behaviours and corresponding models which fall under glossy reflection. Commonly used models include the *Phong* model, *Blinn-Phong* model, *Torrance-Sparrow* model, among others.

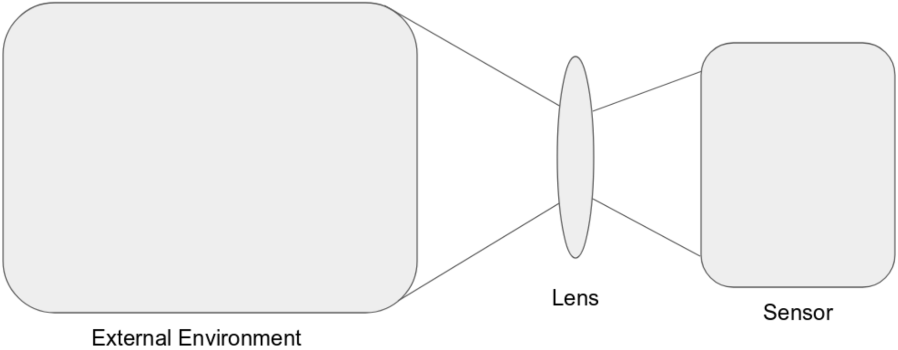
Many Objects Exhibit a Mixture of Behaviour

Image source: Rakuten Global Market

Image source: CARRcrete

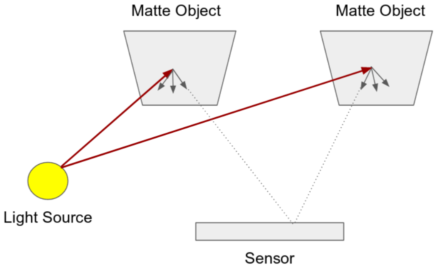
Basic Image Capture Components

At its most basic, capturing an image from the light rays available in a given environment requires two basic components: a lens and a sensor.



Controlling Access to the Sensor

An image sensor records the quantity of light incident on its surface. It does not know the source or direction of the light ray, so if more than one object happens to reflect light to the same location on the sensor, then the image becomes confused.



A schematic overview of an exposed sensor combining information from two objects.

Even with a properly functioning camera, this can still happen if something in the environment interferes with the normal camera function (like a window), creating a trick image.

A real example of light capture gone wrong.

Pinhole Camera

The simplest way to ensure light rays don't get mixed up on the sensor is to block most of the light and provide only a single point of access.

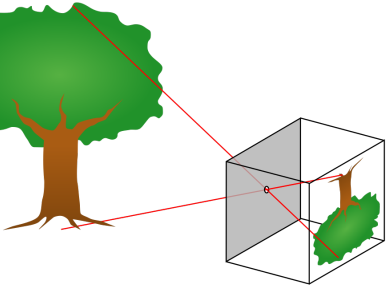


Image source: Wikimedia Commons

Pinhole Camera

The simplest way to ensure light rays don't get mixed up on the sensor is to block most of the light and provide only a single point of access.

- Any point on the sensor has only a single possible path for light to travel from the field of view, preventing interference

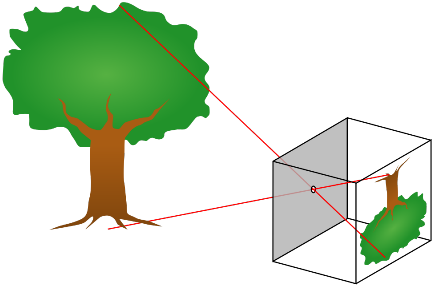


Image source: Wikimedia Commons

Pinhole Camera

The simplest way to ensure light rays don't get mixed up on the sensor is to block most of the light and provide only a single point of access.

- Any point on the sensor has only a single possible path for light to travel from the field of view, preventing interference
- Most of the light gets blocked; we need either very long exposure times or a very high ambient brightness to collect enough light on our sensor

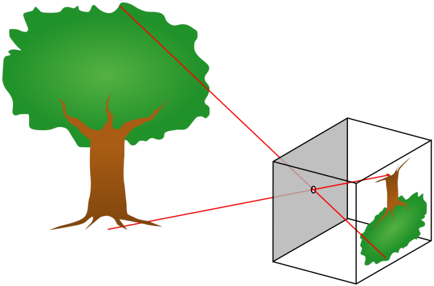
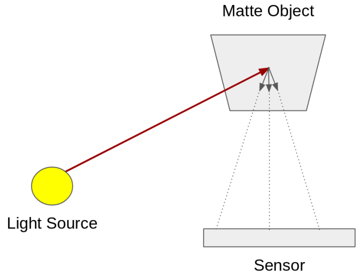


Image source: Wikimedia Commons

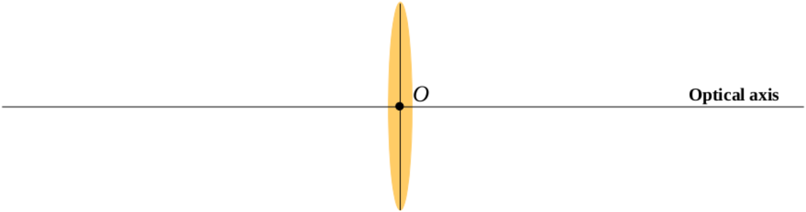
Getting More Light

So far we talked about interference between multiple objects, but we skipped one of the other major problems: creating a spatial correspondence between the real world and image coordinates.



To extract a clear image, we want all light rays entering our sensor from real world point p_e to land on some corresponding sensor point p_s.

Basic optics: Thin lens - a simple idealization

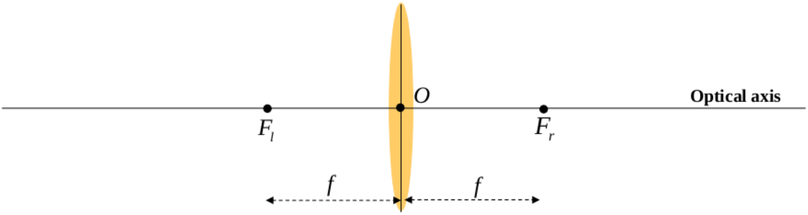


Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to its plane.

Note: This slide and the next several have been adapted from Richard Wildes' slides for this course.

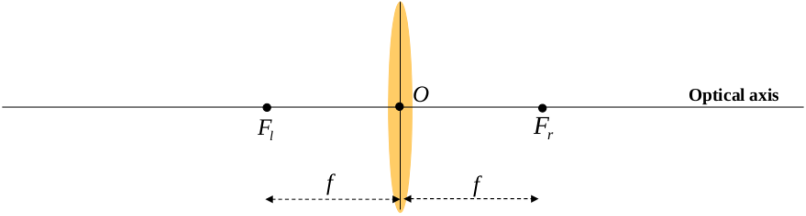
Basic optics: Thin lens - a simple idealization



Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to its plane.
2. Two points, F_l and F_r, called left and right focus along the optical axis at a distance, f, the focal length, from the lens center.

Basic optics: Thin lens - a simple idealization

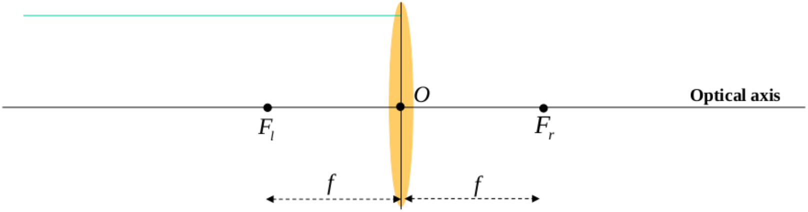


Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to its plane.
2. Two points, F_l and F_r, called left and right focus along the optical axis at a distance, f, the focal length from the lens center.

Two basic properties

Basic optics: Thin lens - a simple idealization



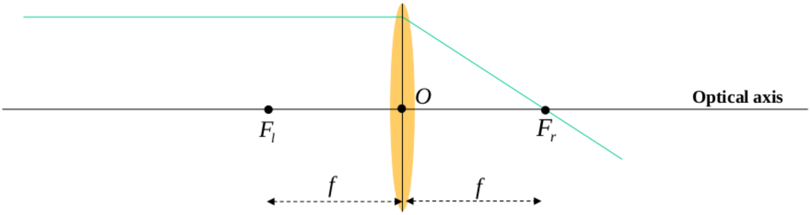
Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to the image plane.
2. Two points, F_l and F_r, called left and right focus along the optical axis at a distance, f, the focal length from the lens center.

Two basic properties

1. Any ray entering the lens parallel to the axis on one side goes through the focus on the other side.

Basic optics: Thin lens - a simple idealization



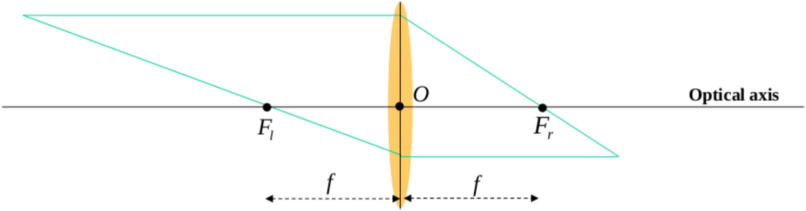
Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to the image plane.
2. Two points, F_l and F_r, called left and right focus along the optical axis at a distance, f, the focal length from the lens center.

Two basic properties

1. Any ray entering the lens parallel to the axis on one side goes through the focus on the other side.

Basic optics: Thin lens - a simple idealization

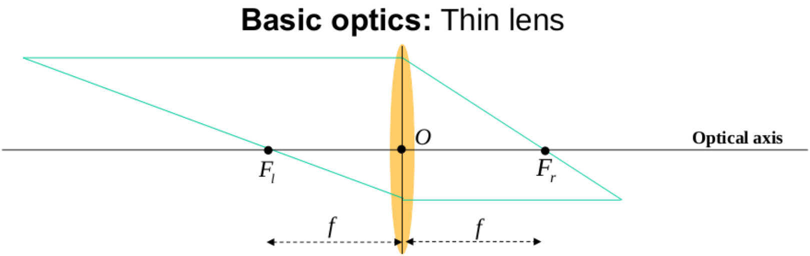


Two characterizing elements

1. Optical axis going through lens center, O, and perpendicular to the image plane.
2. Two points, F_l and F_r, called left and right focus along the optical axis at a distance, f, the focal length from the lens center.

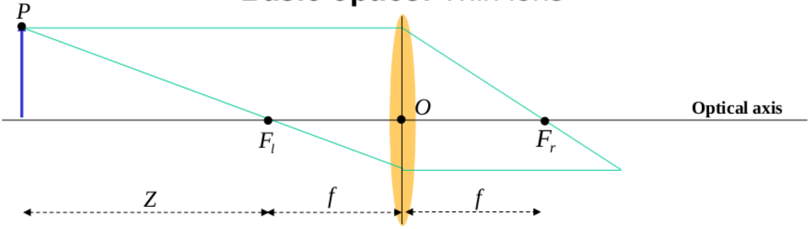
Two basic properties

1. Any ray entering the lens parallel to the axis on one side goes through the focus on the other side.
2. Any ray entering the lens from the focus on one side emerges parallel to the axis on the other side.



Fundamental equation of thin lens: Derived from the two basic properties.

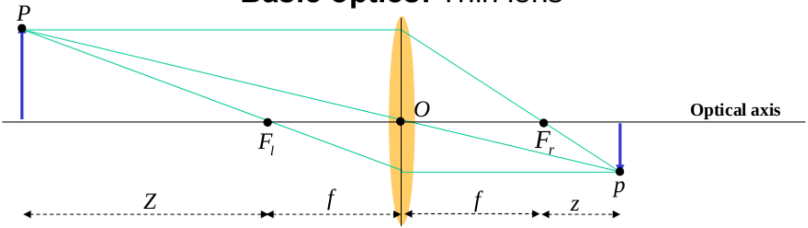
Basic optics: Thin lens



Fundamental equation of thin lens: Derived from the two basic properties.

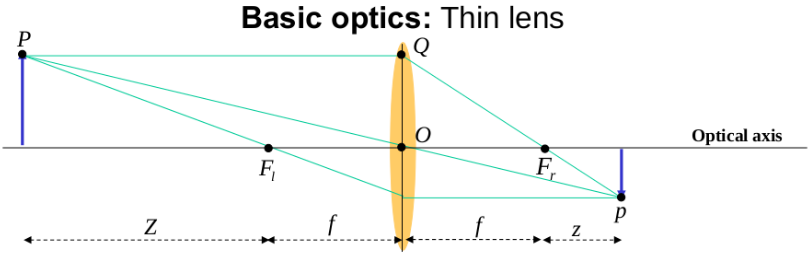
- Consider a point P at a distance $Z + f$ from the lens along the optical axis.

Basic optics: Thin lens



Fundamental equation of thin lens: Derived from the two basic properties.

- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

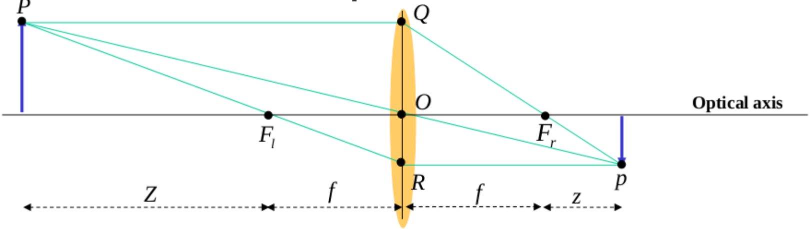


Fundamental equation of thin lens: Derived from the two basic properties.

- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

1. PQ goes through F_r

Basic optics: Thin lens



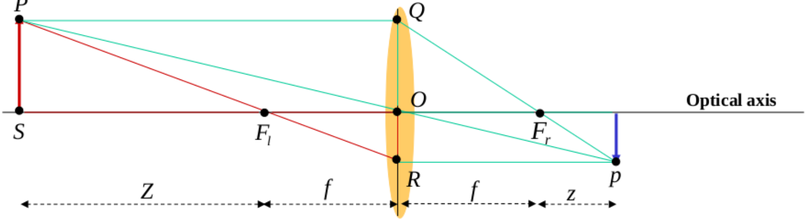
Fundamental equation of thin lens: Derived from the two basic properties.

- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

1. PQ goes through F_r
2. PR emerges parallel to the optical axis

that intersect at p, the image of P.

Basic optics: Thin lens



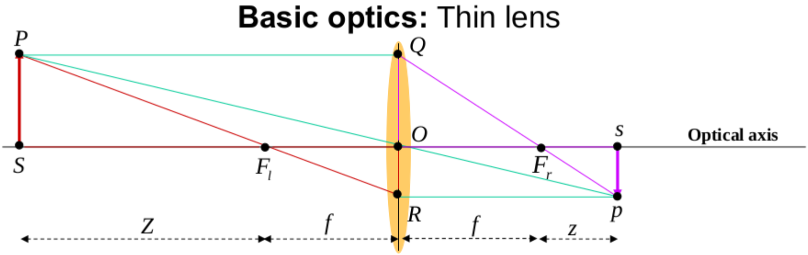
Fundamental equation of thin lens: Derived from the two basic properties.

- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

  1. PQ goes through F_r
  2. PR emerges parallel to the optical axis

  that intersect at p, the image of P.
- From similar triangles PF_lS & RF_lO

$$\frac{Z}{f} = \frac{PS}{OR}$$



Fundamental equation of thin lens: Derived from the two basic properties.

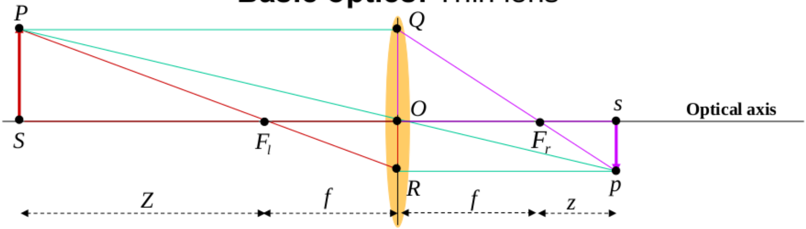
- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

  1. PQ goes through F_r
  2. PR emerges parallel to the optical axis

  that intersect at p, the image of P.
- From similar triangles PF_lS & RF_lO and psF_r & QOF_r we have

$$\frac{Z}{f} = \frac{PS}{OR} \quad \& \quad \frac{QO}{sp} = \frac{f}{z}$$

Basic optics: Thin lens



Fundamental equation of thin lens: Derived from the two basic properties.

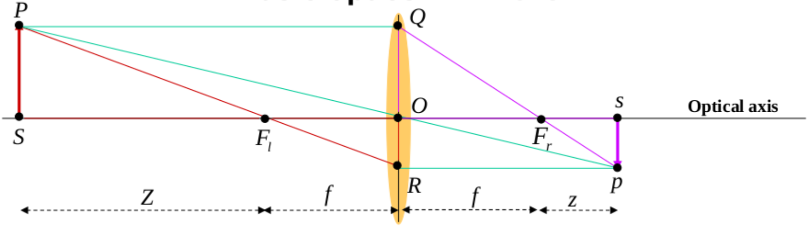
- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

  1. PQ goes through F_r
  2. PR emerges parallel to the optical axis

  that intersect at p, the image of P.
- From similar triangles PF_lS & RF_lO and psF_r & QOF_r we have

$$\frac{Z}{f} = \frac{PS}{OR} = \frac{QO}{sp} = \frac{f}{z}$$

Basic optics: Thin lens



Fundamental equation of thin lens: Derived from the two basic properties.

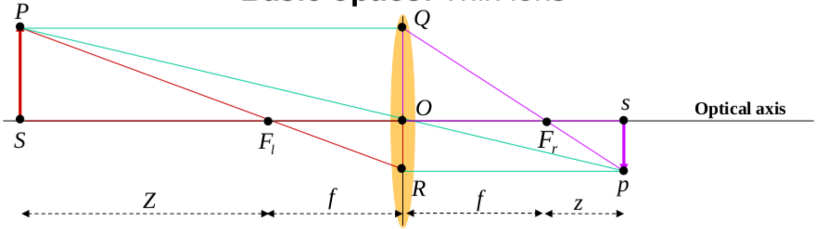
- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

  1. PQ goes through F_r
  2. PR emerges parallel to the optical axis

  that intersect at p, the image of P.
- From similar triangles PF_lS & RF_lO and psF_r & QOF_r we have

$$\frac{Z}{f} = \frac{PS}{OR} = \frac{QO}{sp} = \frac{f}{z}$$

Basic optics: Thin lens



Fundamental equation of thin lens: Derived from the two basic properties.

- Consider a point P at a distance $Z + f$ from the lens along the optical axis.
- All rays from P are focussed to the same point p; consider 2:

  1. PQ goes through F_r
  2. PR emerges parallel to the optical axis

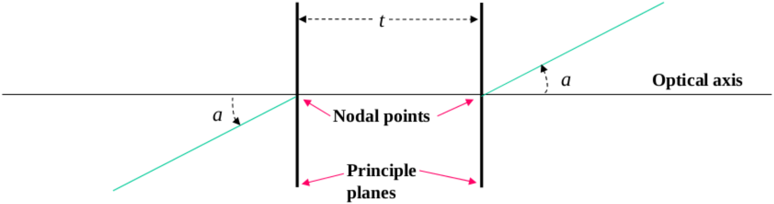
  that intersect at p, the image of P.
- From similar triangles PF_lS & RF_lO and psF_r & QOF_r we have

$$\frac{Z}{f} = \frac{PS}{OR} = \frac{QO}{sp} = \frac{f}{z}$$

- Letting $\hat{Z} = Z + f$ and $\hat{z} = z + f$ yields

$$\frac{1}{\hat{Z}} + \frac{1}{\hat{z}} = \frac{1}{f}$$

Basic optics: Thick lens - a more realistic model



Motivation

- Any simple lens will have number of optical defects.
- For better imaging it is customary to combine several simple lenses by aligning their optical axes to yield a compound lens.
- The thick lens provides a reasonable model of such systems.

Two basic characterizing elements

1. A pair of principle planes parallel to the common optical axis.
2. A pair of nodal points, separated by a distance t – the thickness, where the planes intersect the optical axis.

Fundamental properties

- A ray entering at one nodal point exits at the other without changing direction.
- Produces the same projection as an ideal thin lens, but with an additional offset, t.
- A thin lens is a thick lens where the two nodal points coincide.

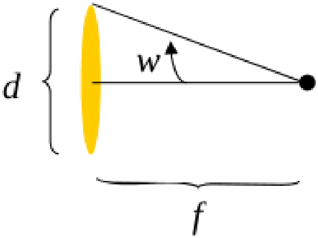
Field of View

Field of View (FOV) is an angular measure of the portion of space captured by a camera.

Let d be the diameter of the of the lens reachable by light rays.

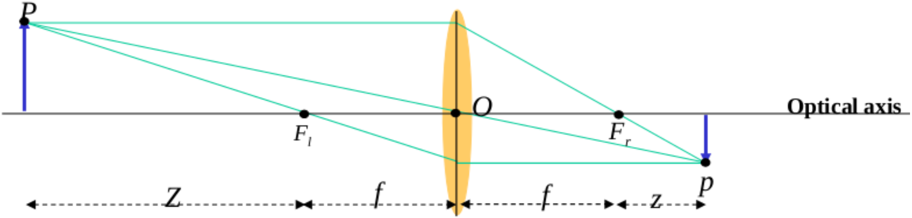
The field of view, w, is equal to half the angle subtended by the lens diameter as seen from the focus.

$\tan w = \frac{d}{2f}$



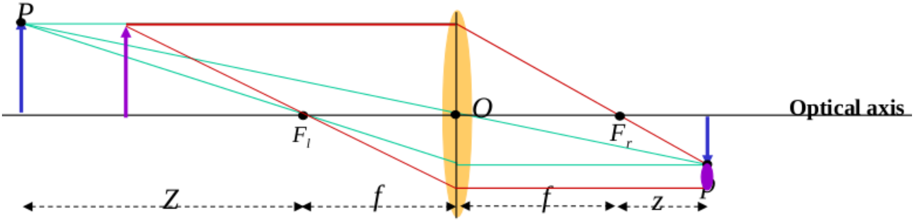
Depth of Field

Returning to our thin lens, we know that a point at distance Z will be focused at z.



Depth of Field

Points at other distances will instead be imaged as (small) circles.



Aperture and Depth of Field

As we saw with the pinhole camera, restricting the available paths of incoming light can mitigate loss of focus. Even in a non-pin hole setting, the size of the opening through which light can reach the sensor will affect how rapidly points off the focal plane lose focus.

Aperture and Depth of Field

As we saw with the pinhole camera, restricting the available paths of incoming light can mitigate loss of focus. Even in a non-pin hole setting, the size of the opening through which light can reach the sensor will affect how rapidly points off the focal plane lose focus.

- Smaller apertures mean a greater range of depths will be (at least practically) in focus.

Aperture and Depth of Field

As we saw with the pinhole camera, restricting the available paths of incoming light can mitigate loss of focus. Even in a non-pin hole setting, the size of the opening through which light can reach the sensor will affect how rapidly points off the focal plane lose focus.

- Smaller apertures mean a greater range of depths will be (at least practically) in focus.
- Larger apertures mean more light can reach the sensor in a given length of time; this is important for low light settings or high speed photography.

Aberrations

Imperfections or limitations in materials and form can lead to performance issues.

Aberrations

Imperfections or limitations in materials and form can lead to performance issues.

- Spherical aberrations: defocusing of rays further from the optical center

Aberrations

Imperfections or limitations in materials and form can lead to performance issues.

- Spherical aberrations: defocusing of rays further from the optical center
- Chromatic aberrations: differential defocusing as a function of light wavelength

Aberrations

Imperfections or limitations in materials and form can lead to performance issues.

- Spherical aberrations: defocusing of rays further from the optical center
- Chromatic aberrations: differential defocusing as a function of light wavelength
- Vignetting: loss of image intensity near the periphery (*e.g.* due to a complex aperture system causing some occlusion)

Lens Distortions

In addition to aberrations, lenses can introduce *radial distortions* in an image. This manifests as curvature of what should be straight lines, and tends to grow in prominence as field of view increases.

Image Source: Szeliski, 2011

Some example distortions are shown above: barrel (left), pincushion (middle), and fisheye (right). Note that the fisheye image approaches a 180° FOV.

The Sensor

The sensor can be formed from a variety of photosensitive materials. In traditional photography, image capture is an analog process with light impacting a photosensitive film and inducing chemical changes.

Modern photography (and computer vision) is dominated by digital image capture. In digital image capture a photoelectric sensor converts incoming light signals to an electrical charge which is quantized into a digital signal.

Common Digital Sensors

The two common classes of digital sensor:

Charge-Coupled Device (CCD):

Complementary Metal-Oxide-Semiconductor (CMOS):

Common Digital Sensors

The two common classes of digital sensor:

Charge-Coupled Device (CCD):

1. Generally higher quality, more precise imaging

Complementary Metal-Oxide-Semiconductor (CMOS):

1. Generally cheaper to manufacture

Common Digital Sensors

The two common classes of digital sensor:

Charge-Coupled Device (CCD):

1. Generally higher quality, more precise imaging
2. Higher power consumption

Complementary Metal-Oxide-Semiconductor (CMOS):

1. Generally cheaper to manufacture
2. Lower power consumption

Common Digital Sensors

The two common classes of digital sensor:

Charge-Coupled Device (CCD):

1. Generally higher quality, more precise imaging
2. Higher power consumption
3. Captures images using capacitive bins
   - Better packing of light sensitive elements
   - Global shutter

Complementary Metal-Oxide-Semiconductor (CMOS):

1. Generally cheaper to manufacture
2. Lower power consumption
3. Each pixel has an individual photodetector and amplifier
   - Better control of "blooming" effects
   - Typically captures pixels one row at a time

Digital Cameras Allow On-board Pre-processing

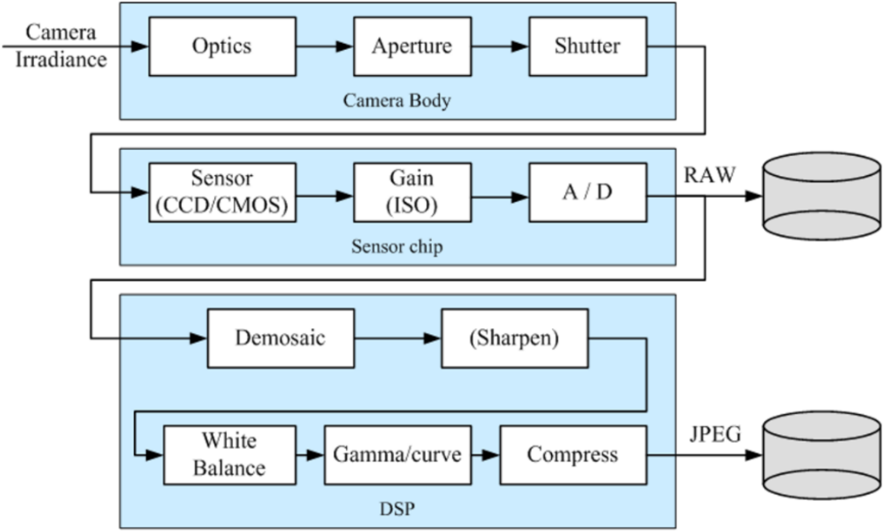
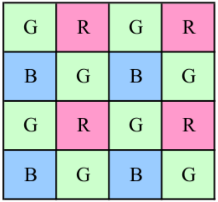
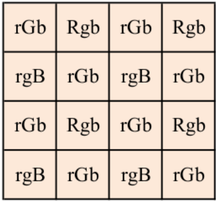


Image Source: Szeliski, 2011

Colour Sensing

To capture colour we need to differentiate sensor response to different wavelengths of light.





A common colour sensitive pattern is the Bayer pattern (left). This, however, means that many pixel colour values are interpolated (lower case letters in the array on the right).

Image Source: Szeliski, 2011

Colour Sensing

Many aspects of the camera pipeline are proprietary and not publicly available. It is sometimes possible to reverse engineer or experimentally determine these properties.

Professor Michael S. Brown here at York does a lot of research on digital image formation and capture, and several of his students work on specific aspects of camera pipeline modeling.

Image Source: York University web pages

Cameras are not Eyes!

- At a certain level of abstraction, cameras are very similar to eyes: both capture light from the environment to provide visual sensory input

Cameras are not Eyes!

- At a certain level of abstraction, cameras are very similar to eyes: both capture light from the environment to provide visual sensory input
- Due to the nature of light capture, they also share some common features (*e.g.* a lens)

Cameras are not Eyes!

- At a certain level of abstraction, cameras are very similar to eyes: both capture light from the environment to provide visual sensory input
- Due to the nature of light capture, they also share some common features (*e.g.* a lens)
- However, there are a number of important differences in function which are useful to keep in mind

Basic Eye Physiology - The Optics

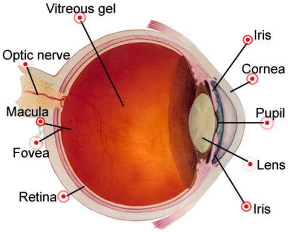


Image source: Original source unknown

- The cornea is an outer protective layer, but also interacts with light to provide focus

Basic Eye Physiology - The Optics

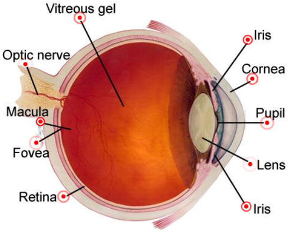


Image source: Original source unknown

- The cornea is an outer protective layer, but also interacts with light to provide focus
- The lens serves as a final step in focal adjustment

Basic Eye Physiology - The Optics

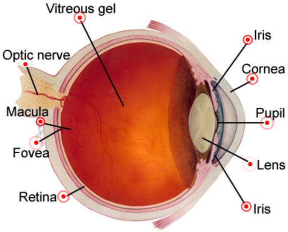


Image source: Original source unknown

- The cornea is an outer protective layer, but also interacts with light to provide focus
- The lens serves as a final step in focal adjustment
- The pupil provides the path for light, similar to an aperture

Basic Eye Physiology - The Optics

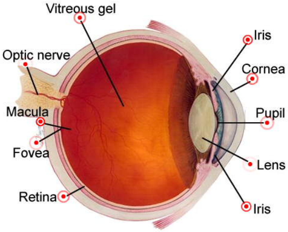


Image source: Original source unknown

- The cornea is an outer protective layer, but also interacts with light to provide focus
- The lens serves as a final step in focal adjustment
- The pupil provides the path for light, similar to an aperture
- The iris controls the size of the pupil

Basic Eye Physiology - The Retina

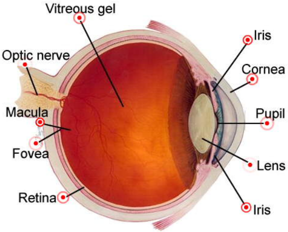


Image source: Original source unknown

The sensor of the eye is the retina.

- Photoreceptors convert light to electrochemical signals

Basic Eye Physiology - The Retina

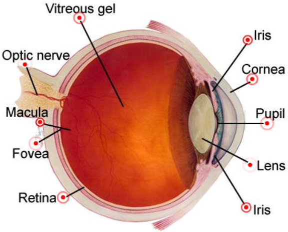


Image source: Original source unknown

The sensor of the eye is the retina.

- Photoreceptors convert light to electrochemical signals
- Retinal ganglion cells transmit information from the retina to the brain

Basic Eye Physiology - The Retina

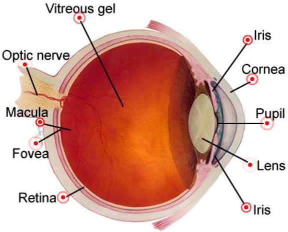


Image source: Original source unknown

The sensor of the eye is the retina.

- Photoreceptors convert light to electrochemical signals
- Retinal ganglion cells transmit information from the retina to the brain
- Highly anisotropic - photoreceptor density and ratio of ganglion cell to photoreceptor inputs varies

Basic Eye Physiology - The Retina

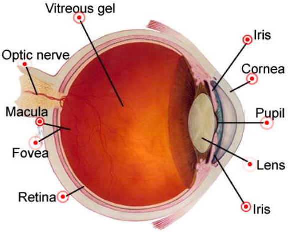


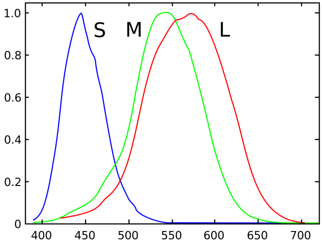
Image source: Original source unknown

The sensor of the eye is the retina.

- Photoreceptors convert light to electrochemical signals
- Retinal ganglion cells transmit information from the retina to the brain
- Highly anisotropic - photoreceptor density and ratio of ganglion cell to photoreceptor inputs varies
- Highest visual acuity is in the fovea at the center of the retina

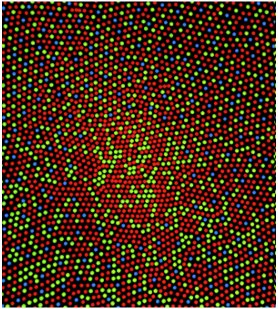
Human Colour Sensing

Humans sense colour with cone cells. There are three types: Long (L), Medium (M), and Short (S).



Normalized response curves to light wavelength (nm)

Image source: Wikipedia



Retinal mosaic showing a normal distribution of foveal cone cells

Image source: Wikipedia

Some Major Differences

Cameras:

Eyes:

Some Major Differences

Cameras:

1. Mostly isotropic field of view

Eyes:

1. Highly anisotropic field of view

Some Major Differences

Cameras:

1. Mostly isotropic field of view
2. Synchronous shutter, frame-based acquisition

Eyes:

1. Highly anisotropic field of view
2. Asynchronous, continuous input

Some Major Differences

Cameras:

1. Mostly isotropic field of view
2. Synchronous shutter, frame-based acquisition
3. Typically operated in a stabilized manner

Eyes:

1. Highly anisotropic field of view
2. Asynchronous, continuous input
3. Complicated interplay between stabilization and eye movements