Computer Vision

Lecture Set 02
Digital Snapshots

Huei-Yung Lin

lin@ntut.edu.tw

Robot Vision Lab
Department of Computer Science and Information Engineering
National Taipei University of Technology
Taipei 106, Taiwan

September 26, 2022

Digital Images

- Digital image
  - 2-D array (matrix) of numbers
  - Numbers can be light intensity, distance (range), etc.
  - Intensity image, range image
- Intensity image
  - Photograph-like images encoding light intensities, acquired by cameras
  - Measure the amount of light impinging on a photosensitive device
- Range image
  - Encoding shape and distance acquired by special sensors like sonar or laser scanners (time-of-flight)
  - Estimate directly the 3-D structure of the viewed scene through variety of (sensing or vision-based) techniques

Digital Images

- Digital image
  - 2-D array (matrix) of numbers
  - Numbers can be light intensity, distance (range), etc.
  - Intensity image, range image
- Intensity image
  - Photograph-like images encoding light intensities, acquired by cameras
  - Measure the amount of light impinging on a photosensitive device
- Range image
  - Encoding shape and distance acquired by special sensors like sonar or laser scanners (time-of-flight)
  - Estimate directly the 3-D structure of the viewed scene through variety of (sensing or vision-based) techniques

Digital Images

- Digital image
  - 2-D array (matrix) of numbers
  - Numbers can be light intensity, distance (range), etc.
  - Intensity image, range image
- Intensity image
  - Photograph-like images encoding light intensities, acquired by cameras
  - Measure the amount of light impinging on a photosensitive device
- Range image
  - Encoding shape and distance acquired by special sensors like sonar or laser scanners (time-of-flight)
  - Estimate directly the 3-D structure of the viewed scene through variety of (sensing or vision-based) techniques

Intensity Image Formation

- Optical parameters – Characterize the sensor's optics
  - Lens type, focus length, field of view, angular aperture
- Photometric parameters – Appear in models of the light energy reaching the sensor after being reflected from the objects in the scene
  - Type, intensity, direction of illumination
  - Reflectance properties of the viewed surface
  - Effects of the sensor's structure on the amount of light reaching the photoreceptors
- Geometric parameters – Determine the image position on which a 3-D point is projected
  - Type of projections (perspective, affine, orthographic, etc.)
  - Position and orientation of camera in space
  - Perspective distortion introduced by the imaging process

Intensity Image Formation

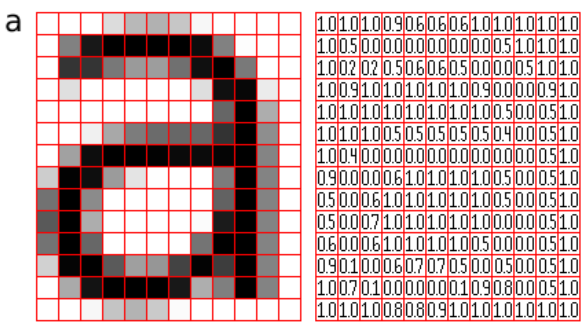
- Optical parameters – Characterize the sensor's optics
  - Lens type, focus length, field of view, angular aperture
- Photometric parameters – Appear in models of the light energy reaching the sensor after being reflected from the objects in the scene
  - Type, intensity, direction of illumination
  - Reflectance properties of the viewed surface
  - Effects of the sensor's structure on the amount of light reaching the photoreceptors
- Geometric parameters – Determine the image position on which a 3-D point is projected
  - Type of projections (perspective, affine, orthographic, etc.)
  - Position and orientation of camera in space
  - Perspective distortion introduced by the imaging process

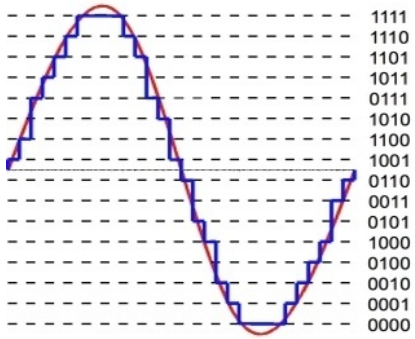
Intensity Image Formation

- Optical parameters – Characterize the sensor's optics
  - Lens type, focus length, field of view, angular aperture
- Photometric parameters – Appear in models of the light energy reaching the sensor after being reflected from the objects in the scene
  - Type, intensity, direction of illumination
  - Reflectance properties of the viewed surface
  - Effects of the sensor's structure on the amount of light reaching the photoreceptors
- Geometric parameters – Determine the image position on which a 3-D point is projected
  - Type of projections (perspective, affine, orthographic, etc.)
  - Position and orientation of camera in space
  - Perspective distortion introduced by the imaging process

Digital Image Representations

- A digital image is an array of numbers indicating the image irradiance at various points on the image plane
- Image intensities are spatially sampled
- Intensity values are quantized (8-bits, 10-bit, 12-bits)

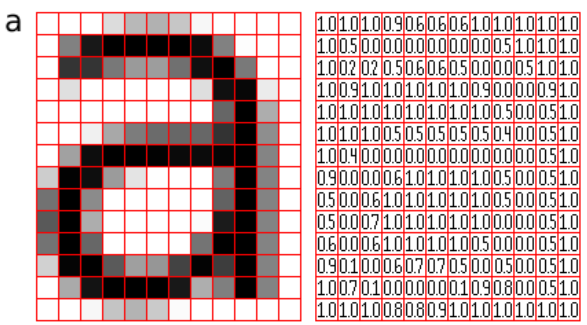


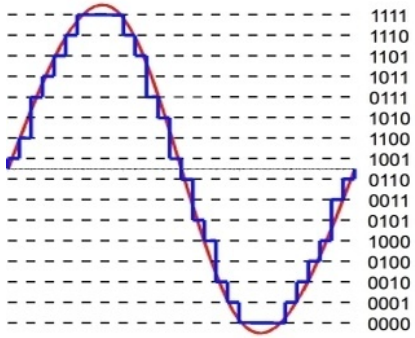


- Acquisition: CCD or CMOS arrays
- Storage: Usually in computer memory
- Display: Computer hardware (video boards) and monitors

Digital Image Representations

- A digital image is an array of numbers indicating the image irradiance at various points on the image plane
- Image intensities are spatially sampled
- Intensity values are quantized (8-bits, 10-bit, 12-bits)





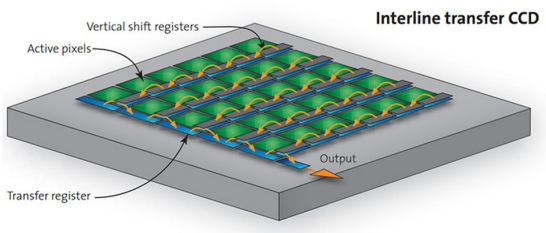
- Acquisition: CCD or CMOS arrays
- Storage: Usually in computer memory
- Display: Computer hardware (video boards) and monitors

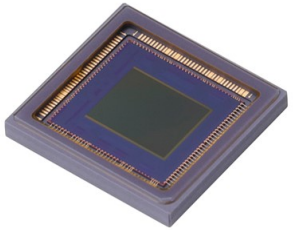
Image Sensors

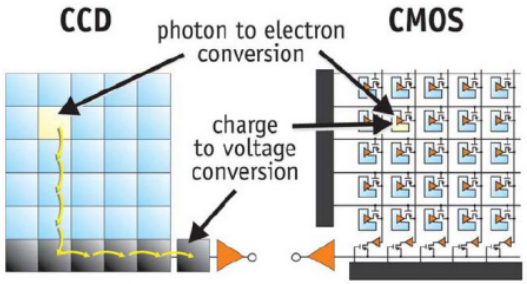
- Images are formed by the interaction of the incident image irradiance with light sensitive elements on the image plane
- Light sensitive elements
  - Film
  - CCD (Charge Coupled Device)
  - CMOS Imaging element
  - ...

Image Sensors

- Images are formed by the interaction of the incident image irradiance with light sensitive elements on the image plane
- Light sensitive elements
  - Film
  - CCD (Charge Coupled Device)
  - CMOS Imaging element
  - ...

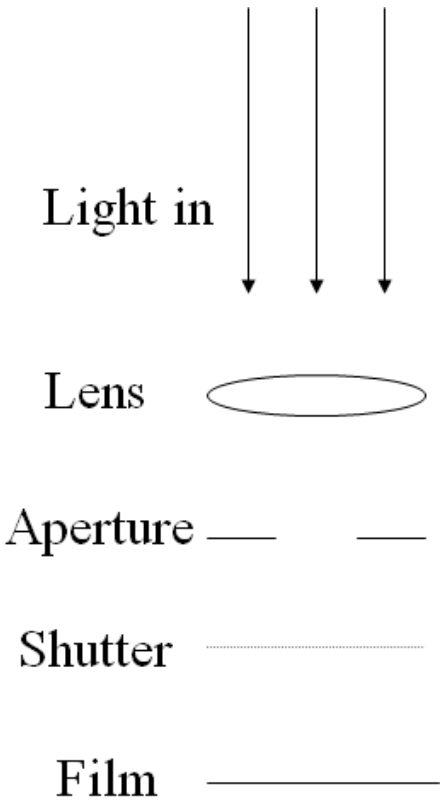


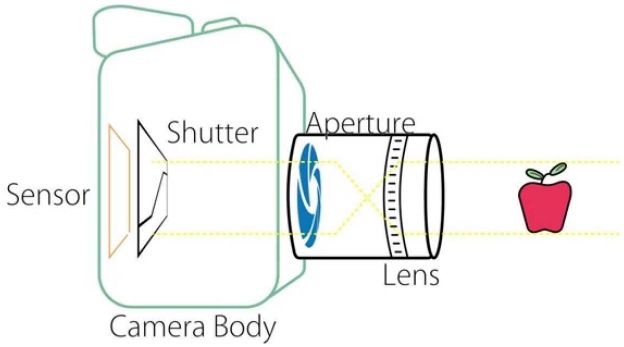




Cameras

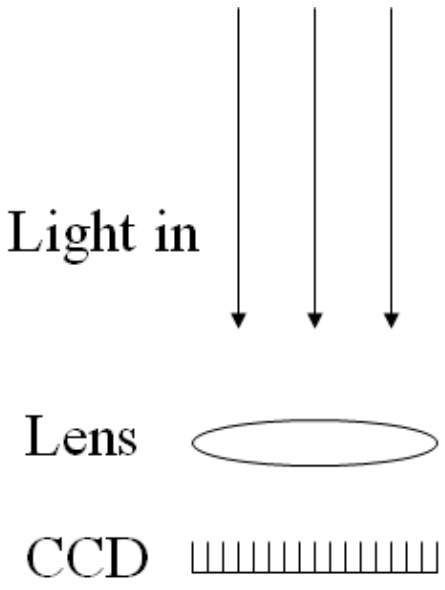
- A camera is a device for mediating the way light strikes film
- Lens lets light in while maintaining focus
- Aperture controls proportion of the light that gets to the film
- Shutter controls how long light is allowed to get to the film





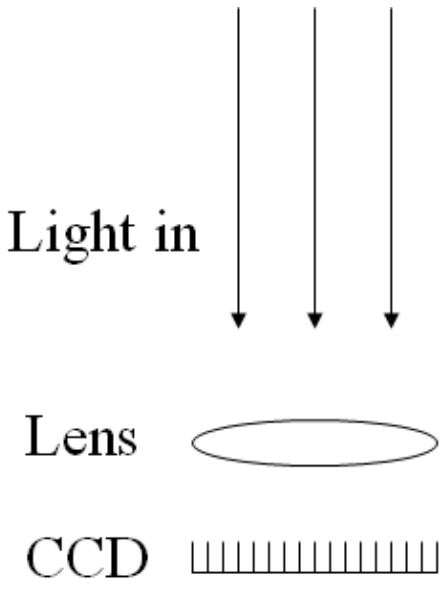
Digital Cameras

- CCD stores a charge each time a photon hits it
  - "Bins" have discrete area, one per pixel
  - Spatially discrete
- Camera "reads" the charges out of the bins at some frequency
- Convert charges to discrete value
- Store values in memory – the image
- Still have issues of motion blur, depth of field, dynamic range, etc.



Digital Cameras

- CCD stores a charge each time a photon hits it
  - "Bins" have discrete area, one per pixel
  - Spatially discrete
- Camera "reads" the charges out of the bins at some frequency
- Convert charges to discrete value
- Store values in memory – the image
- Still have issues of motion blur, depth of field, dynamic range, etc.



Digital Images

- Computers work with discrete pieces of information
- How do we digitize a continuous image?
  - Break the continuous space into small areas, pixels
  - Use a single value for each pixel – the pixel value
  - No longer continuous in space or intensity

Digital Images

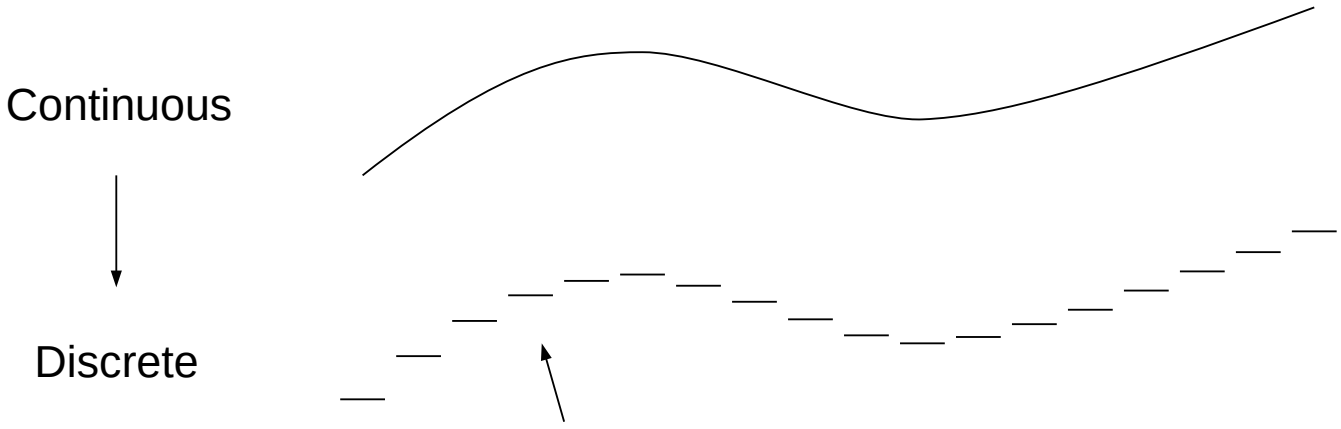
- Computers work with discrete pieces of information
- How do we digitize a continuous image?
  - Break the continuous space into small areas, pixels
  - Use a single value for each pixel – the pixel value
  - No longer continuous in space or intensity

Digital Images

- Computers work with discrete pieces of information
- How do we digitize a continuous image?
  - Break the continuous space into small areas, pixels
  - Use a single value for each pixel – the pixel value
  - No longer continuous in space or intensity

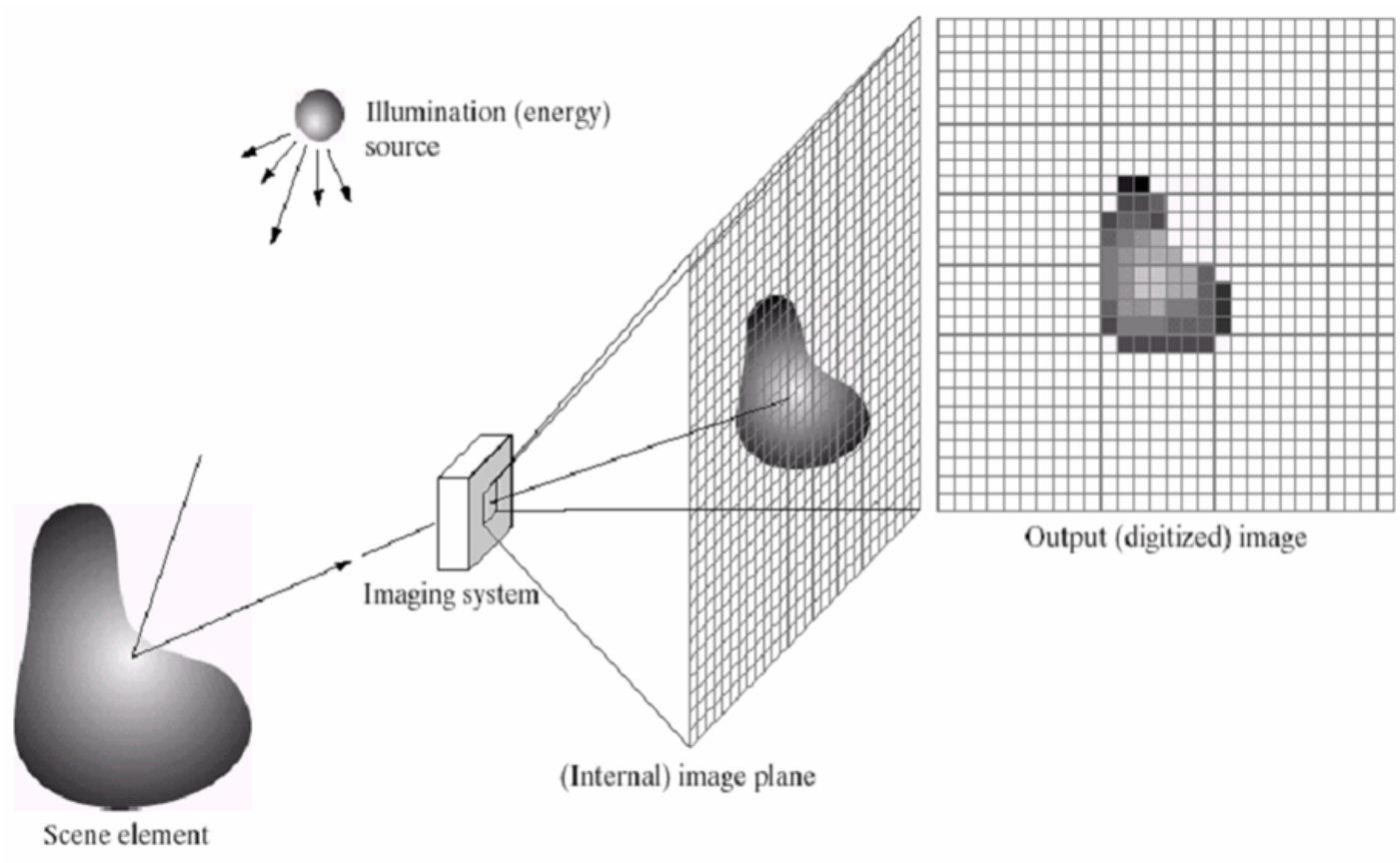
Digital Images

- Computers work with discrete pieces of information
- How do we digitize a continuous image?
  - Break the continuous space into small areas, pixels
  - Use a single value for each pixel – the pixel value
  - No longer continuous in space or intensity

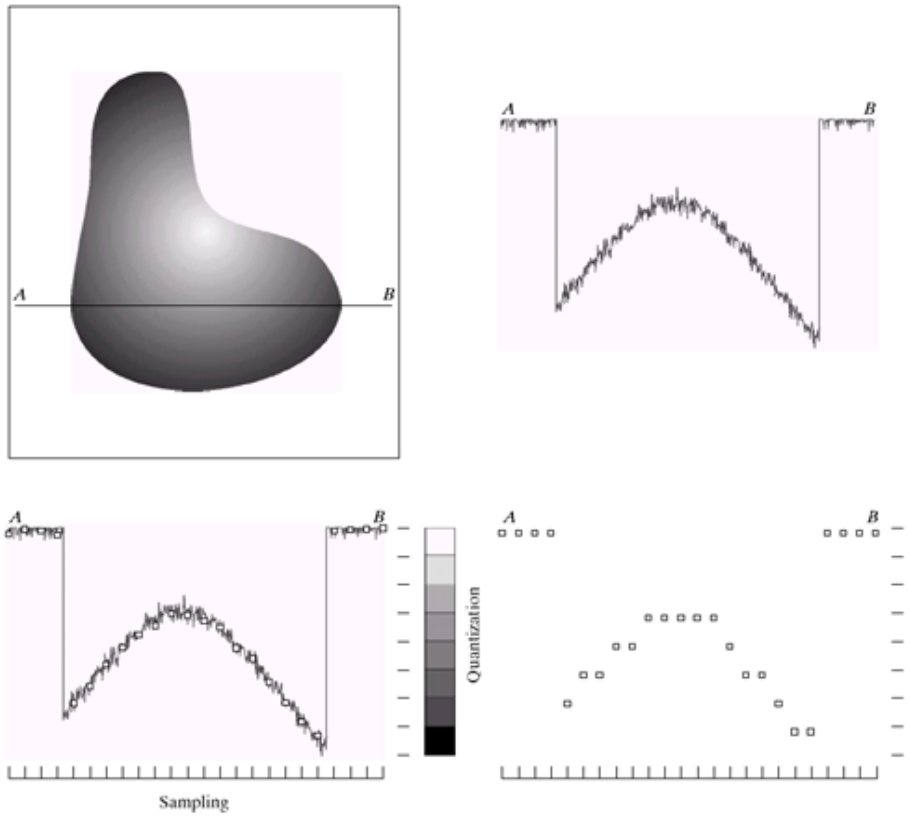


Pixel: Picture Element

Digital Image Acquisition Process



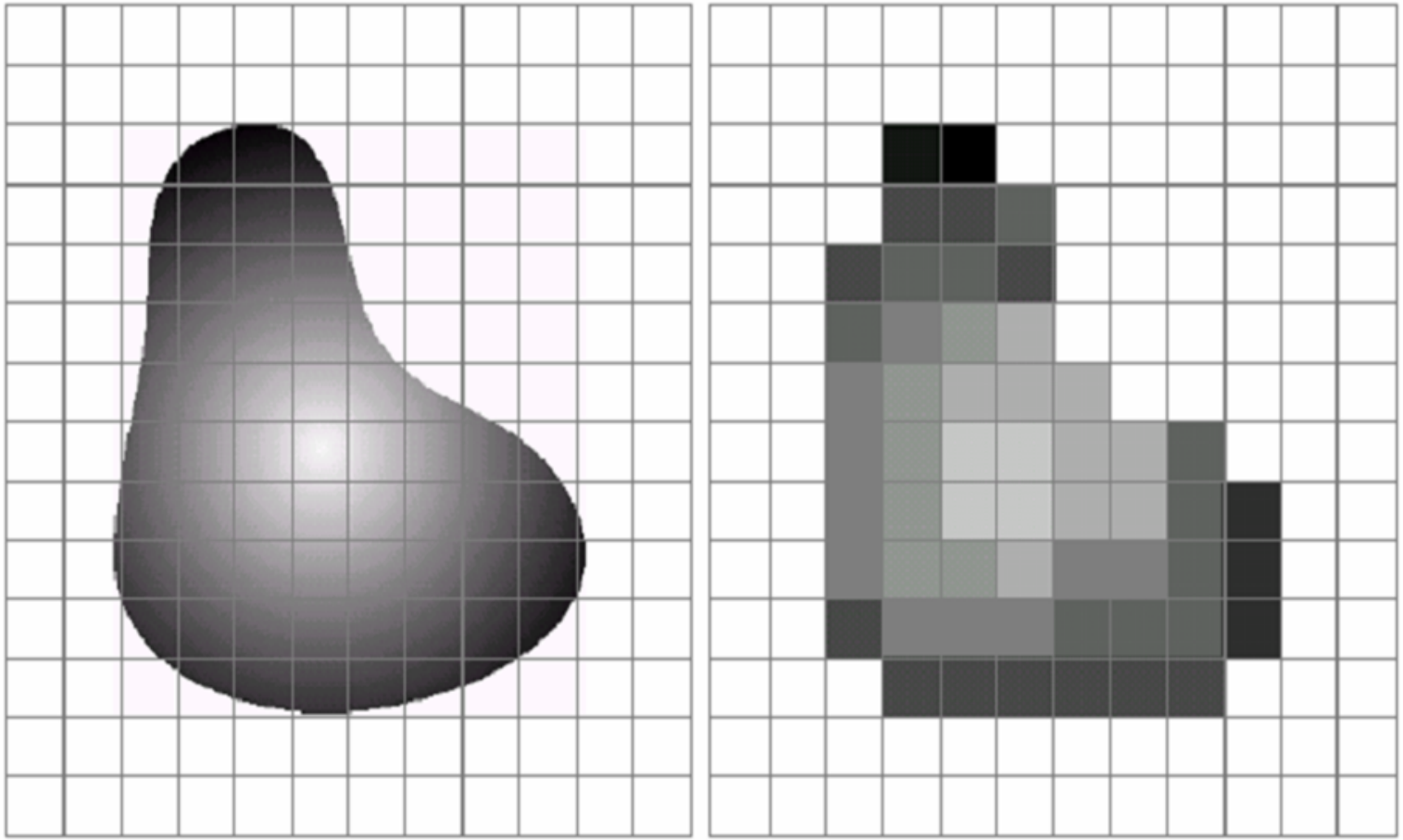
Generating a Digital Image



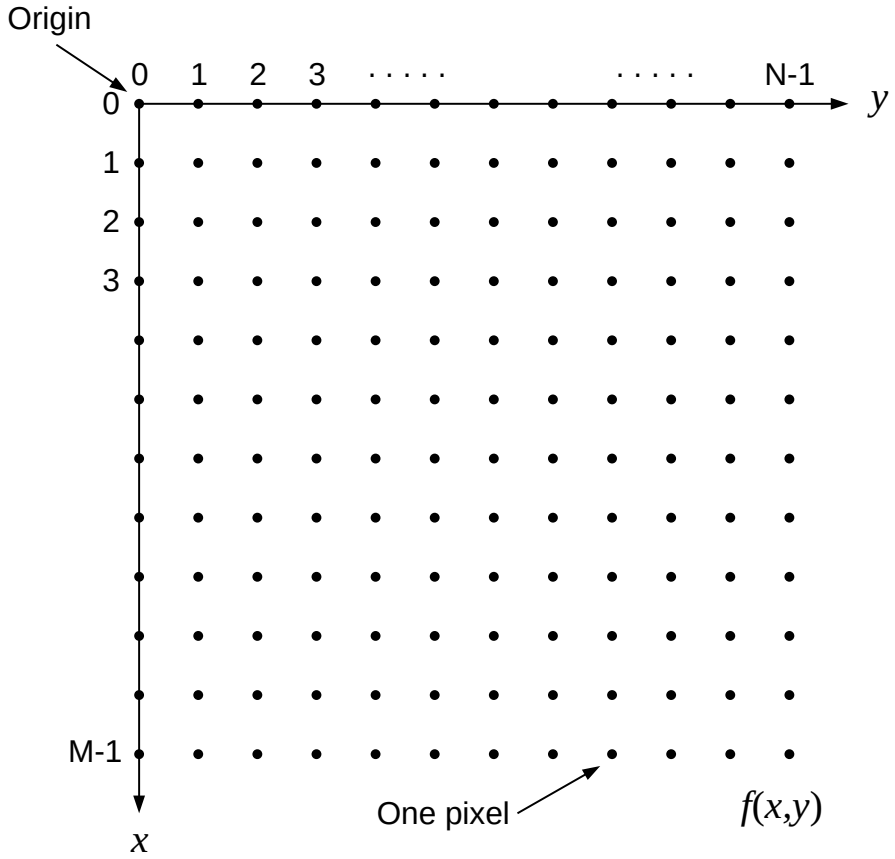
a b
c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Image Sampling and Quantization



Coordinate Conversion



Video Images

- Images are taken sequentially by opening and closing the shutter 30 times per second (might be more or less)
  - Temporal aliasing
  - Motion blur
- Sensing Color
  - In a 3 CCD video camera the light path is split into 3 components which are passed through colored light filters and then imaged
  - As a result – a color image contains three channels of information: red, green and blue image intensities

Video Images

- Images are taken sequentially by opening and closing the shutter 30 times per second (might be more or less)
  - Temporal aliasing
  - Motion blur
- Sensing Color
  - In a 3 CCD video camera the light path is split into 3 components which are passed through colored light filters and then imaged
  - As a result – a color image contains three channels of information: red, green and blue image intensities

Video Images

- Images are taken sequentially by opening and closing the shutter 30 times per second (might be more or less)
  - Temporal aliasing
  - Motion blur
- Sensing Color
  - In a 3 CCD video camera the light path is split into 3 components which are passed through colored light filters and then imaged
  - As a result – a color image contains three channels of information: red, green and blue image intensities

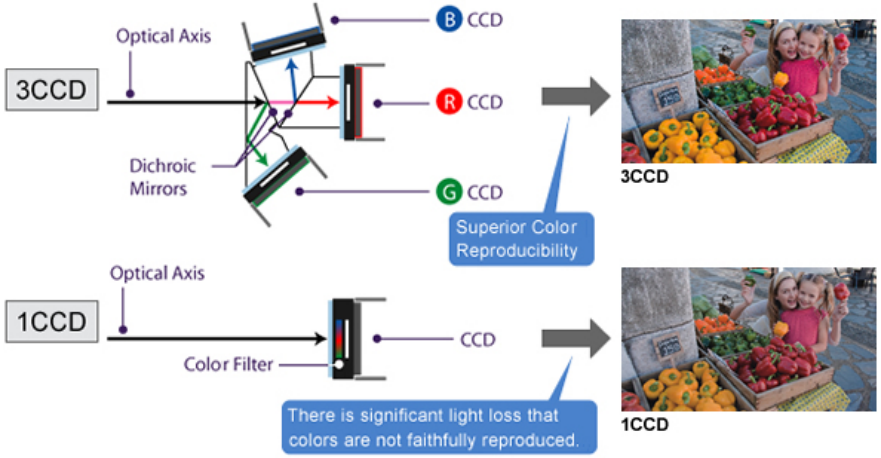


Image Storage

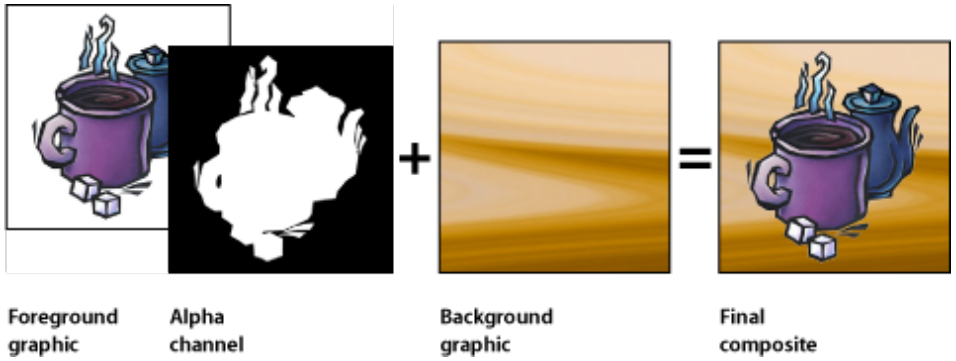
- Images are stored as continuous string of bytes
- As a two-dimensional array of pixels
- Color images can be stored using either 24 or 8 bits
- The way of storage can affect performance of algorithms
- 4-bytes per pixel is more efficient than 3, because it aligns better with integer boundaries in memory
- 4th pixel is often the alpha component between 0 and 1
- Final color is [aR aG aB]

Image Storage

- Images are stored as continuous string of bytes
- As a two-dimensional array of pixels
- Color images can be stored using either 24 or 8 bits
- The way of storage can affect performance of algorithms
- 4-bytes per pixel is more efficient than 3, because it aligns better with integer boundaries in memory
- 4th pixel is often the alpha component between 0 and 1
- Final color is [aR aG aB]

Image Storage

- Images are stored as continuous string of bytes
- As a two-dimensional array of pixels
- Color images can be stored using either 24 or 8 bits
- The way of storage can affect performance of algorithms
- 4-bytes per pixel is more efficient than 3, because it aligns better with integer boundaries in memory
- 4th pixel is often the alpha component between 0 and 1
- Final color is [aR aG aB]



Announcements – 9/26/22

- Homework I description will be given next Monday (10/3), in class, due 10/17.
- The TA will explain the homework details online using Google Meet.
- No late homework.

Basic Optics

- Light rays which enter the camera through an angular aperture (pupil) and hit the image plane
- Light rays are the result of reflections of the rays emitted by the light sources and hitting object surfaces
- Lenses are used to increase the light gathering power of the instrument (otherwise, it is called a pinhole camera)

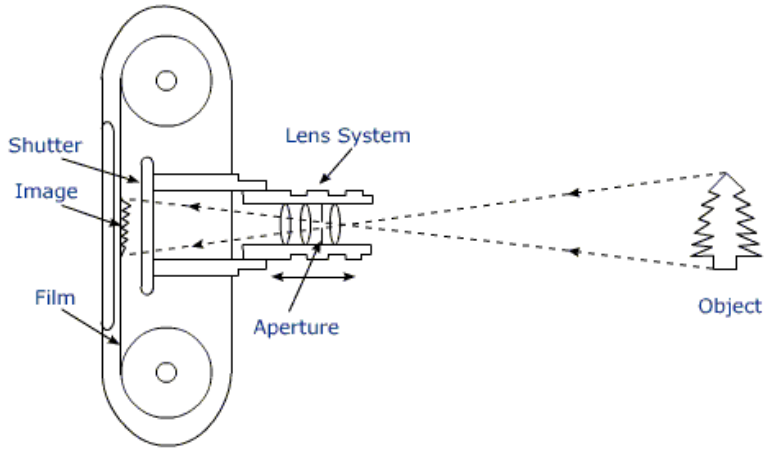
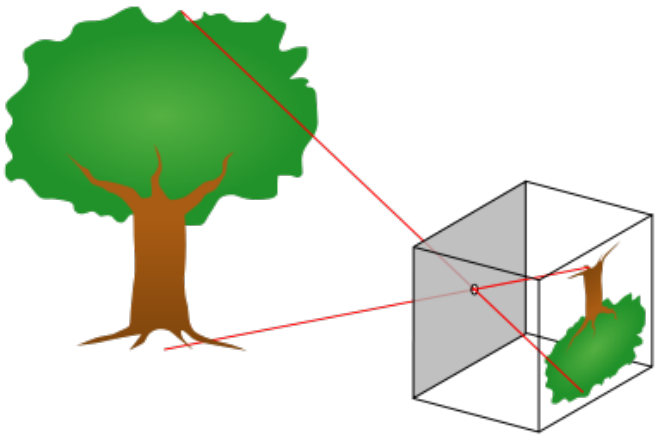


Image Focusing

- A scene point is in focus if all rays coming from the point converges to a single point on the image plane
  - Pinhole – exposure time
  - Optical system – lenses, shutter, etc.



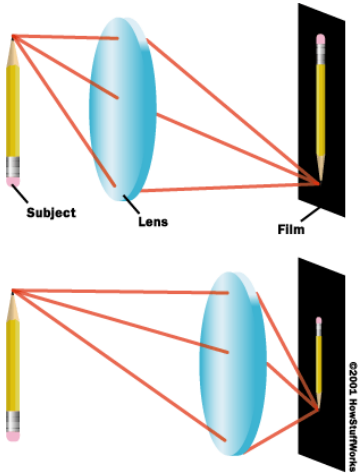
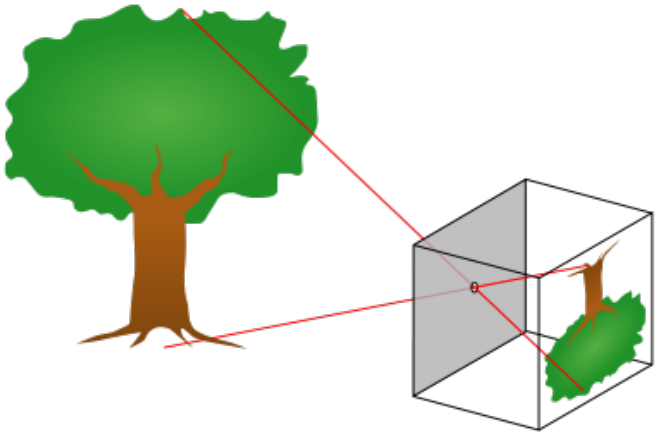
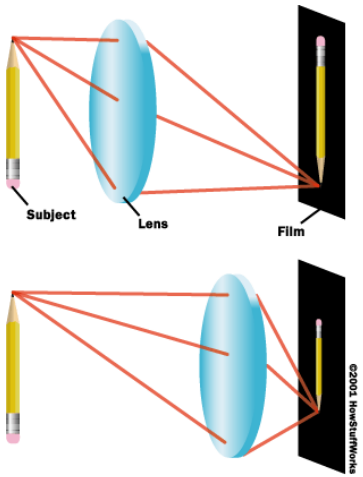


Image Focusing

- Thin lens (model)
  - Any ray entering the lens parallel to the axis on one side goes through the focus on the other side
  - Any ray entering the lens from the focus on one side merges parallel to the axis on the other side





Thin Lens Equation

- The fundamental equation of thin lenses:

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

- When $z \to \infty, z' \to f$
  - Focal length is the distance of image plane (IP) from the optical origin when parallel rays are focused to a single point in the IP
  - z' is called the camera constant (distance between lens and IP)
- When the lens is focused on a point *not* at infinity
  - $z' < f$
  - Using f to approximate z' overestimates the camera constant
  - It needs to change the image plane position (or lens position)

Thin Lens Equation

- The fundamental equation of thin lenses:

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

- When $z \to \infty, z' \to f$
  - Focal length is the distance of image plane (IP) from the optical origin when parallel rays are focused to a single point in the IP
  - z' is called the camera constant (distance between lens and IP)
- When the lens is focused on a point *not* at infinity
  - $z' < f$
  - Using f to approximate z' overestimates the camera constant
  - It needs to change the image plane position (or lens position)

Thin Lens Equation

- The fundamental equation of thin lenses:

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

- When $z \to \infty, z' \to f$
  - Focal length is the distance of image plane (IP) from the optical origin when parallel rays are focused to a single point in the IP
  - z' is called the camera constant (distance between lens and IP)
- When the lens is focused on a point *not* at infinity
  - $z' < f$
  - Using f to approximate z' overestimates the camera constant
  - It needs to change the image plane position (or lens position)

Depth of Field

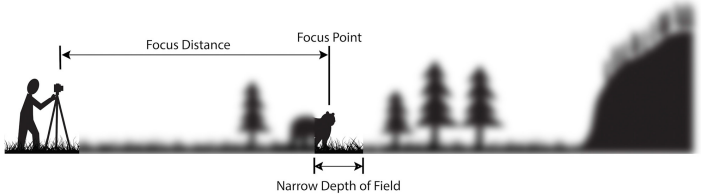
- Aperture size
  - Large – more light, faster shutter speed, small depth of field
  - Small – less light, large depth of field
- According to the lens equation, for a particular image plane at z', only points at distance z are in focus
- However, some amount of defocusing below the resolution of the imaging device can be tolerated

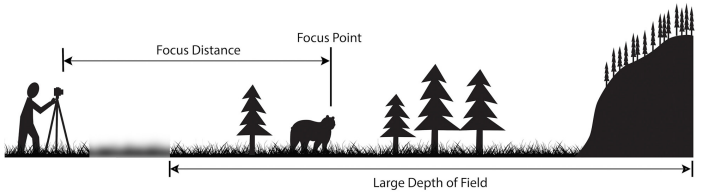
Depth of Field

- Aperture size
  - Large – more light, faster shutter speed, small depth of field
  - Small – less light, large depth of field
- According to the lens equation, for a particular image plane at z', only points at distance z are in focus
- However, some amount of defocusing below the resolution of the imaging device can be tolerated

Depth of Field

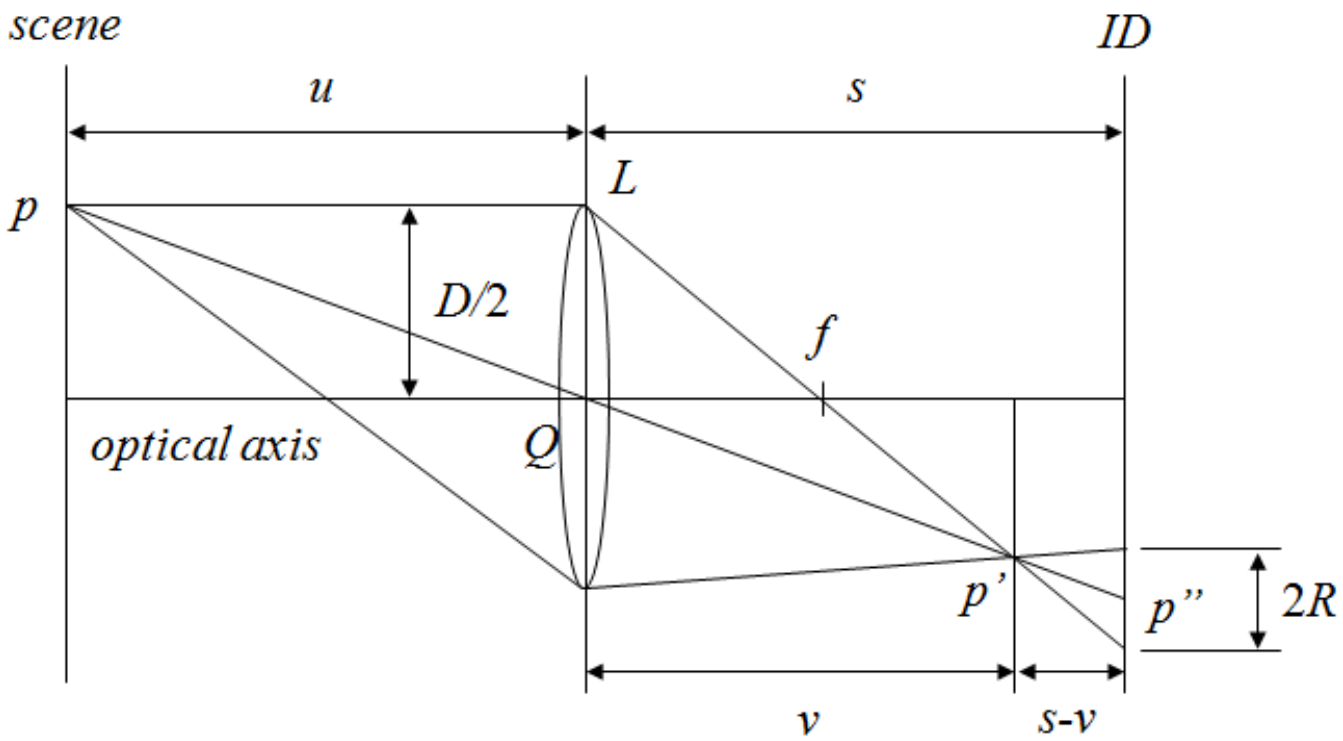
- Depth of field: A range of image plane distance z' with an acceptable level of defocusing and a corresponding range of scene distance z





Defocus

- When a scene point is out of focus, it creates a blur circle instead of a single point on the image plane



Exposure

- Exposure
  - The amount of light collected by the camera

$$\varepsilon = Et$$

  - E: image irradiance (the intensity of light falling on the image plane)
- F-number (F-stop)
  - The ratio of the focal and aperture diameter
  - F-number = f/d
  - Increases with multiple of $\sqrt{2}$ (light increases with multiple of 2)
  - 2.8, 4, 5.6, 8, 11, etc.

Exposure

- Exposure
  - The amount of light collected by the camera

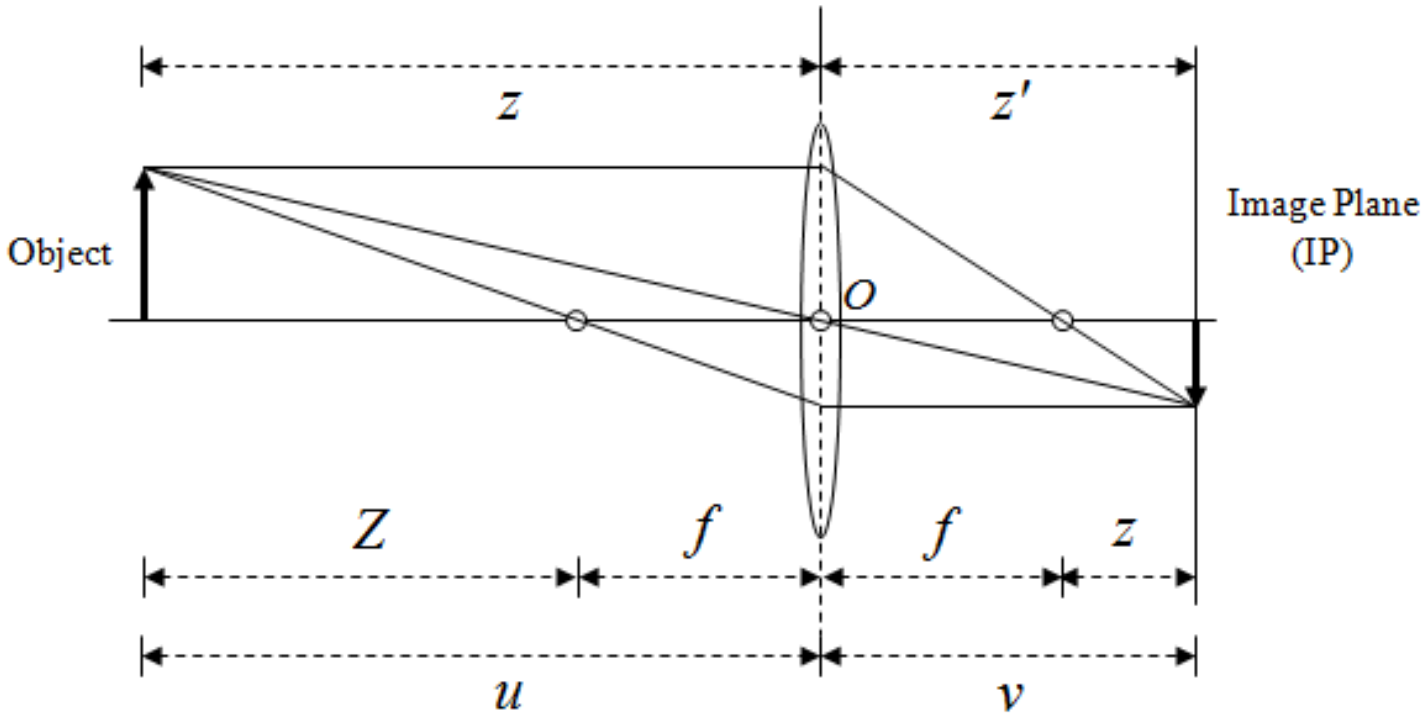
$$\varepsilon = Et$$

  - E: image irradiance (the intensity of light falling on the image plane)
- F-number (F-stop)
  - The ratio of the focal and aperture diameter
  - F-number = f/d
  - Increases with multiple of $\sqrt{2}$ (light increases with multiple of 2)
  - 2.8, 4, 5.6, 8, 11, etc.

Commonly Used Lens Model

- The fundamental equation of thin lenses:

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}, \qquad u = Z + f, \quad v = z + f$$



Field of View

- Field of view:
  - An angular measure of the portion of 3-D space actually seen by the camera
  - Determined by the effective lens diameter and the focal length

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

- Some other problems for optics (thin lenses)
  - Spherical aberration
    - Defocusing of non-paraxial rays
  - Chromatic aberration
    - Different defocusing of rays of different colors

Field of View

- Field of view:
  - An angular measure of the portion of 3-D space actually seen by the camera
  - Determined by the effective lens diameter and the focal length

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

- Some other problems for optics (thin lenses)
  - Spherical aberration
    - ★ Defocusing of non-paraxial rays
  - Chromatic aberration
    - ★ Different defocusing of rays of different colors

Light (and Matter)

- A form of electromagnetic radiation
  - Speed = wavelength $\times$ frequency
- Speed of light in a vacuum = 3×10^8 m/sec
- Visible spectrum
  - 400 nm – 700 nm
  - Longer wavelengths correspond to redder colors and shorter wavelengths to bluer color
- EM radiation comes in discrete packets called photons
- The interaction between light and matter can take many forms
  - Reflection
  - Refraction
  - Diffraction
  - Absorption
  - Scattering

Light (and Matter)

- A form of electromagnetic radiation
  - Speed = wavelength $\times$ frequency
- Speed of light in a vacuum = 3×10^8 m/sec
- Visible spectrum
  - 400 nm – 700 nm
  - Longer wavelengths correspond to redder colors and shorter wavelengths to bluer color
- EM radiation comes in discrete packets called photons
- The interaction between light and matter can take many forms
  - Reflection
  - Refraction
  - Diffraction
  - Absorption
  - Scattering

Basic Radiometry

- Radiometry
  - The essential part of image formation concerned with the relation among the amounts of light energy emitted from light sources, reflected from surfaces, and registered by sensors
- What for?
  - Modeling how much the illuminating light is reflected by object surfaces
  - Modeling how much of the reflected light actually reaches the image plane of the camera

Radiometry

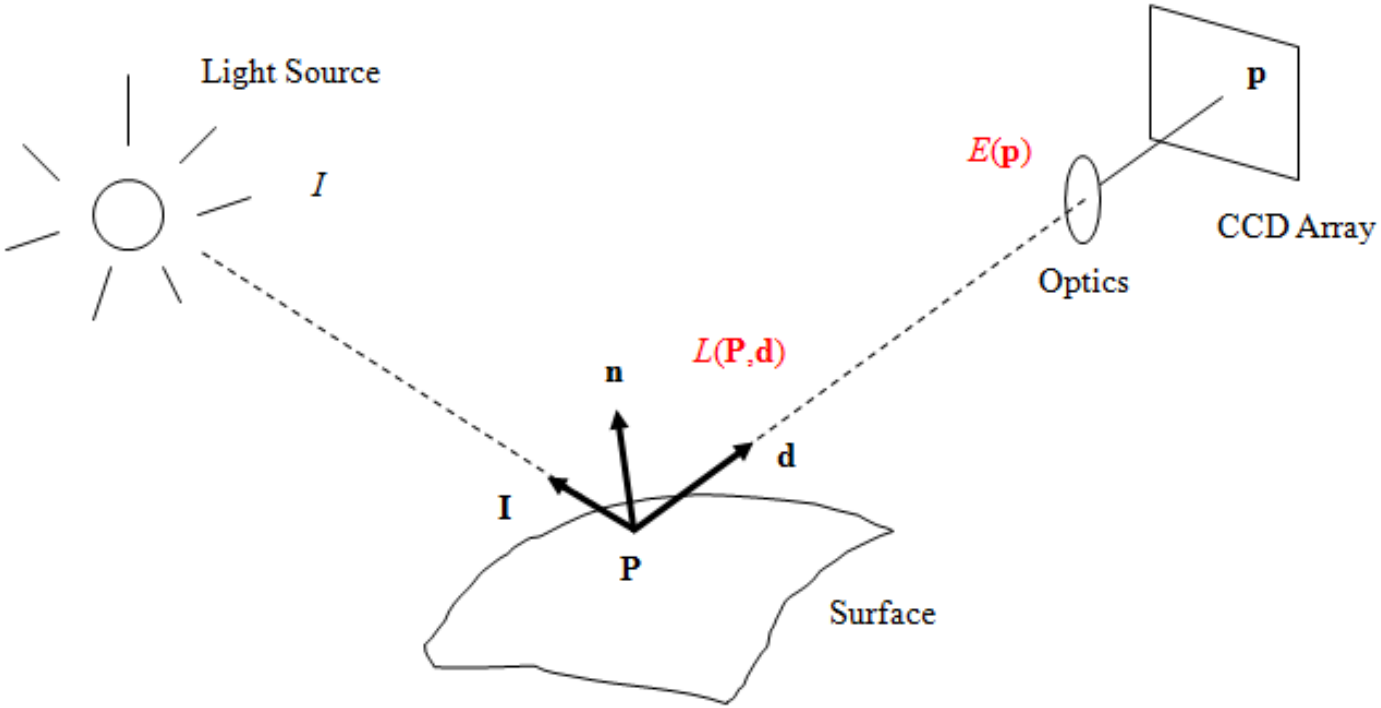


Image Irradiance & Scene Radiance

- Image irradiance
  - The amount of energy falling per unit time, per unit from a specific direction at a point $\mathbf{p}$ of the image plane
  - Measured using an area perpendicular to that direction
  - The energy incident on a surface
- Scene radiance
  - The amount of energy emitted per unit time per unit solid angle per unit foreshortened area by a point $\mathbf{P}$ of a surface in 3-D space in a specific direction $\mathbf{d}$
  - The total radiance emitted by a point
  - The energy radiated from a surface (emitted or reflected)

Image Irradiance & Scene Radiance

- Image irradiance
  - The amount of energy falling per unit time, per unit from a specific direction at a point $\mathbf{p}$ of the image plane
  - Measured using an area perpendicular to that direction
  - The energy incident on a surface
- Scene radiance
  - The amount of energy emitted per unit time per unit solid angle per unit foreshortened area by a point $\mathbf{P}$ of a surface in 3-D space in a specific direction $\mathbf{d}$
  - The total radiance emitted by a point
  - The energy radiated from a surface (emitted or reflected)

BRDF

- Bidirectional Reflectance Distribution Function (BRDF) relates surface properties and light source properties
- Expresses the ratio between the incident light and the emitted light from the surface

$$f(\lambda, \theta_i, \phi_i, \theta_e, \phi_e) = \frac{L(\lambda, \theta_e, \phi_e)}{I(\lambda, \theta_i, \phi_i)}$$

  - f: Bidirectional Reflectance Distribution Function (BRDF)
  - L: outgoing scene radiance in a given direction
  - I: incoming irradiance in a particular direction
  - The wavelength parameter λ is there because surface may reflect differently under different wavelengths of light

BRDF

- Bidirectional Reflectance Distribution Function (BRDF) relates surface properties and light source properties
- Expresses the ratio between the incident light and the emitted light from the surface

$$f(\lambda, \theta_i, \phi_i, \theta_e, \phi_e) = \frac{L(\lambda, \theta_e, \phi_e)}{I(\lambda, \theta_i, \phi_i)}$$

  - f: Bidirectional Reflectance Distribution Function (BRDF)
  - L: outgoing scene radiance in a given direction
  - I: incoming irradiance in a particular direction
  - The wavelength parameter λ is there because surface may reflect differently under different wavelengths of light

Lambertian Surfaces

- A model of the way a surface reflects incident light is called a surface reflectance model
- Surfaces which obey Lambert's Law radiate light equally in all directions regardless of viewing direction
- For such surfaces the BRDF is a constant
- It is called Lambertian model or Lambertian surface

Lambertian Surfaces

- A model of the way a surface reflects incident light is called a surface reflectance model
- Surfaces which obey Lambert's Law radiate light equally in all directions regardless of viewing direction
- For such surfaces the BRDF is a constant
- It is called Lambertian model or Lambertian surface

Lambertian Surfaces

- A model of the way a surface reflects incident light is called a surface reflectance model
- Surfaces which obey Lambert's Law radiate light equally in all directions regardless of viewing direction
- For such surfaces the BRDF is a constant
- It is called Lambertian model or Lambertian surface

Albedo

- The albedo (whiteness) ρ of a surface refers to the fraction of incident light that the surface reflects
- It depends on the surface's material
- For Lamertian surfaces the scene radiance of a point illuminated by a single distant light source can be computed from

$$L = \rho \mathbf{I} \cdot \mathbf{n}$$

  where ρ is a constant and $\mathbf{I} \cdot \mathbf{n}$ is positive (surface faces the light source, check the previous figure)
- Independent of $\mathbf{d}$ (the viewing direction)

Albedo

- The albedo (whiteness) ρ of a surface refers to the fraction of incident light that the surface reflects
- It depends on the surface's material
- For Lamertian surfaces the scene radiance of a point illuminated by a single distant light source can be computed from

$$L = \rho \mathbf{I} \cdot \mathbf{n}$$

  where ρ is a constant and $\mathbf{I} \cdot \mathbf{n}$ is positive (surface faces the light source, check the previous figure)
- Independent of $\mathbf{d}$ (the viewing direction)

Observation on Lambertian Model

- Lambert's Law basically holds for matte surfaces like paper, chalk, plastics
- The amount of light reaching any surface is always proportional to the cosine of the angle between the illuminant and surface normal $\mathbf{n}$
- According to the model, a Lambertian surface reflects light in a given direction $\mathbf{d}$ proportionally to the cosine of θ, the angle between $\mathbf{d}$ and $\mathbf{n}$
- Since the surface's area seen from the direction θ is inversely proportional to $\cos\theta$, the two $\cos\theta$ factors are cancelled out[1]

[1] You should check this and write down the equations on your own.

Observation on Lambertian Model

- Lambert's Law basically holds for matte surfaces like paper, chalk, plastics
- The amount of light reaching any surface is always proportional to the cosine of the angle between the illuminant and surface normal $\mathbf{n}$
- According to the model, a Lambertian surface reflects light in a given direction $\mathbf{d}$ proportionally to the cosine of θ, the angle between $\mathbf{d}$ and $\mathbf{n}$
- Since the surface's area seen from the direction θ is inversely proportional to $\cos\theta$, the two $\cos\theta$ factors are cancelled out[1]

[1] You should check this and write down the equations on your own.

Observation on Lambertian Model

- Lambert's Law basically holds for matte surfaces like paper, chalk, plastics
- The amount of light reaching any surface is always proportional to the cosine of the angle between the illuminant and surface normal $\mathbf{n}$
- According to the model, a Lambertian surface reflects light in a given direction $\mathbf{d}$ proportionally to the cosine of θ, the angle between $\mathbf{d}$ and $\mathbf{n}$
- Since the <u>surface's area</u> seen from the direction θ is inversely proportional to $\cos\theta$, the two $\cos\theta$ factors are cancelled out[1]

[1] You should check this and write down the equations on your own.

Specular Surfaces

- Specular surfaces act like mirrors
- The BRDF for such a surface looks like an impulse function, or an impulse function and a lobe
- Comprehensive models mostly used in graphics might contain
  - An ambient component
  - A diffuse component
  - A specular component

Specular Surfaces

- Specular surfaces act like mirrors
- The BRDF for such a surface looks like an impulse function, or an impulse function and a lobe
- Comprehensive models mostly used in graphics might contain
  - An ambient component
  - A diffuse component
  - A specular component

Problems with BRDF

- It does not account for systems that absorb energy in one band of wavelengths and emit it in another (fluorescence)
- Best viewed as an approximation to the aggregate behavior of the material, can change with scale
- A truly correct model would have to capture the interaction between photons and atom, both at the boundary and in the objects interior

Reality Is Not So Simple

- In real uncontrolled scenes the incident light on various surfaces may come from many different directions
- In addition, surfaces may reflect light onto each other (inter-reflections)
- As a result, it is very difficult to compute what the scene radiance would be for an actual environment
- Non-point Light Sources
  - For distributed light sources – such as the sky – it is important to take into consideration how much light arrives on a given surface in order to properly compute the scene radiance
  - Example: The floor of a well will appear darker than flat surfaces on the ground plane

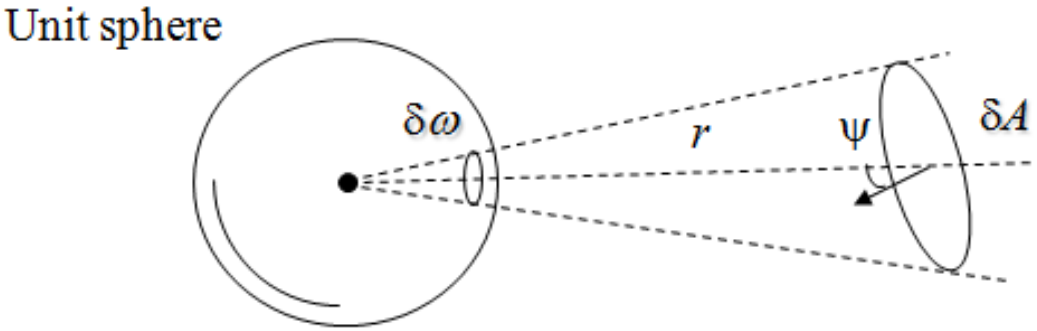
Reality Is Not So Simple

- In real uncontrolled scenes the incident light on various surfaces may come from many different directions
- In addition, surfaces may reflect light onto each other (inter-reflections)
- As a result, it is very difficult to compute what the scene radiance would be for an actual environment
- Non-point Light Sources
  - For distributed light sources – such as the sky – it is important to take into consideration how much light arrives on a given surface in order to properly compute the scene radiance
  - Example: The floor of a well will appear darker than flat surfaces on the ground plane

Solid Angle

- The solid angle of a cone of directions is the area cut out by the cone on the unit sphere centered in the cone's vertex
- The solid angle $\delta\omega$ subtended by a small, planar patch of area δA at distance r from the origin is

$$\delta\omega = \frac{\delta A \cos \Psi}{r^2}$$

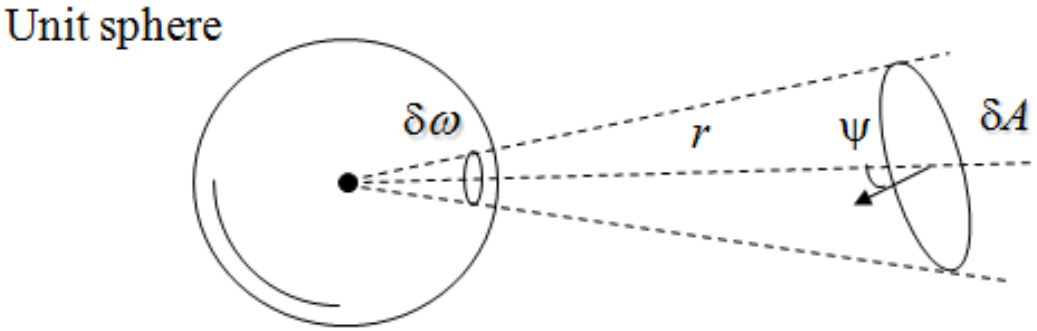


ψ : the angle between the normal to δA and the ray that points from the origin to δA

Solid Angle

- The solid angle of a cone of directions is the area cut out by the cone on the unit sphere centered in the cone's vertex
- The solid angle $\delta\omega$ subtended by a small, planar patch of area δA at distance r from the origin is

$$\delta\omega = \frac{\delta A \cos \Psi}{r^2}$$



ψ : the angle between the normal to δA and the ray that points from the origin to δA

Announcements – 10/3/22

- Description for Homework I will be given by TA later.

Image Irradiance

- Image irradiance E is defined as $E = \delta P / \delta I$
  - δP: the power of light over a small image patch
  - δI: the area of the small image patch
- Consider a scene point $\mathbf{P}$ and its image point $\mathbf{p}$
  - δO: the area of a small surface patch around $\mathbf{P}$
  - L: the scene radiance at $\mathbf{P}$ in the direction toward the lens
  - $\Delta\Omega$: the solid angle subtended by the lens
  - θ: angle between the normal to the viewed surface at $\mathbf{P}$ and the principal ray

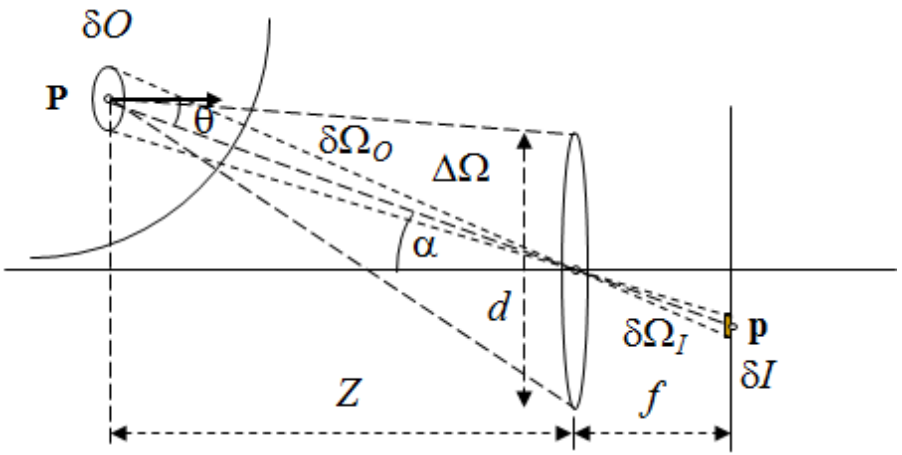
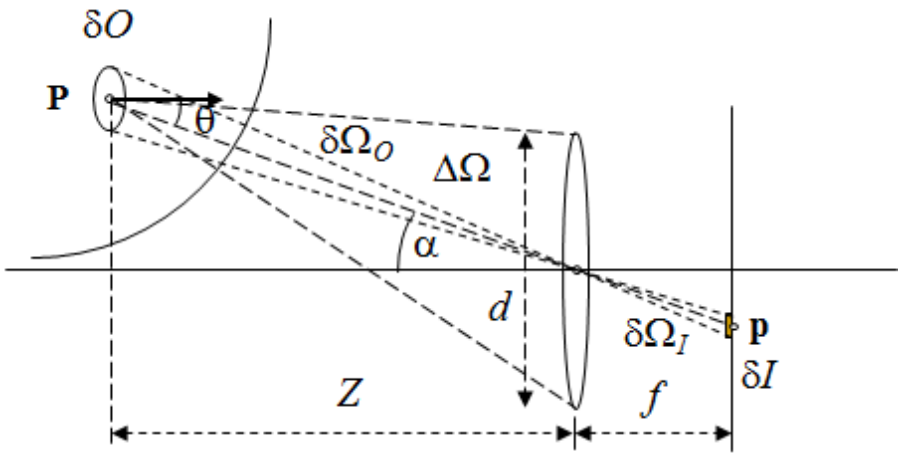


Image Irradiance

- Image irradiance E is defined as $E = \delta P / \delta I$
  - δP: the power of light over a small image patch
  - δI: the area of the small image patch
- Consider a scene point $\mathbf{P}$ and its image point $\mathbf{p}$
  - δO: the area of a small surface patch around $\mathbf{P}$
  - L: the scene radiance at $\mathbf{P}$ in the direction toward the lens
  - $\Delta\Omega$: the solid angle subtended by the lens
  - θ: angle between the normal to the viewed surface at $\mathbf{P}$ and the principal ray



Radiometry of Image Formation

- The power δP is given by $\delta P = L\delta O \cos\theta \Delta\Omega$
- From the solid angle $\Delta\Omega$, $\delta A = \pi d^2/4$, $\Psi = \alpha$, $r = Z/\cos\alpha$, we have

$$\Delta\Omega = \frac{\pi}{4} d^2 \frac{\cos^3\alpha}{Z^2}$$

- For the solid angle $\delta\Omega_I$, $\delta A = \delta I$, $\Psi = \alpha$, $r = f/\cos\alpha$, we have

$$\delta\Omega_I = \frac{\delta I \cos\alpha}{(f/\cos\alpha)^2}$$

- Similarly,

$$\delta\Omega_o = \frac{\delta O \cos\theta}{(Z/\cos\alpha)^2}$$

Fundamental Equation

- Since $\delta\Omega_I = \delta\Omega_O$,

$$1 = \frac{\delta\Omega_O}{\delta\Omega_I} = \frac{\delta O \cos\theta}{(Z/\cos\alpha)^2} \frac{(f/\cos\alpha)^2}{\delta I \cos\alpha} = \frac{\delta O}{\delta I} \frac{f^2}{Z^2} \frac{\cos\theta}{\cos\alpha} \Rightarrow \frac{\delta O}{\delta I} = \frac{\cos\alpha}{\cos\theta} \left(\frac{Z}{f}\right)^2$$

- Thus,

$$E = \frac{\delta P}{\delta I} = \frac{L\delta O \cos\theta \Delta\Omega}{\delta I} = \frac{\cos\alpha}{\cos\theta} \left(\frac{Z}{f}\right)^2 L\frac{\pi}{4} d^2 \frac{\cos^3\alpha}{Z^2} \cos\theta = L\frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\alpha$$

- Fundamental equation of radiometric image formation:

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\alpha$$

- Points to note: $E(\mathbf{p})$ linearly related to $L(\mathbf{P})$, F-number f/d, $\cos^4\alpha$ terms

Fundamental Equation

- Since $\delta\Omega_I = \delta\Omega_O$,

$$1 = \frac{\delta\Omega_O}{\delta\Omega_I} = \frac{\delta O \cos\theta}{(Z/\cos\alpha)^2} \frac{(f/\cos\alpha)^2}{\delta I \cos\alpha} = \frac{\delta O}{\delta I} \frac{f^2}{Z^2} \frac{\cos\theta}{\cos\alpha} \Rightarrow \frac{\delta O}{\delta I} = \frac{\cos\alpha}{\cos\theta} \left(\frac{Z}{f}\right)^2$$

- Thus,

$$E = \frac{\delta P}{\delta I} = \frac{L\delta O \cos\theta \Delta\Omega}{\delta I} = \frac{\cos\alpha}{\cos\theta} \left(\frac{Z}{f}\right)^2 L\frac{\pi}{4} d^2 \frac{\cos^3\alpha}{Z^2} \cos\theta = L\frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\alpha$$

- Fundamental equation of radiometric image formation:

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\alpha$$

- Points to note: $E(\mathbf{p})$ linearly related to $L(\mathbf{P})$, F-number f/d, $\cos^4\alpha$ terms

Fundamental Equation

- From the fundamental equation of radiometric image formation

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- The illumination of the image at $\mathbf{p}$ decreases as the fourth power of the cosine of the angle formed by the principal ray through $\mathbf{p}$ with the optical axis
  - This can be neglected in the case of small angular aperture
  - Therefore, the image irradiance can be regarded as uniformly proportional to the scene radiance over the whole image plane
- The non-uniform illumination predicted by the fundamental equation is hard to notice in ordinary images
- This is because the major component of brightness changes is usually due to the spatial gradient of the image irradiance

Fundamental Equation

- From the fundamental equation of radiometric image formation

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- The illumination of the image at $\mathbf{p}$ decreases as the fourth power of the cosine of the angle formed by the principal ray through $\mathbf{p}$ with the optical axis
  - This can be neglected in the case of small angular aperture
  - Therefore, the image irradiance can be regarded as uniformly proportional to the scene radiance over the whole image plane
- The non-uniform illumination predicted by the fundamental equation is hard to notice in ordinary images
- This is because the major component of brightness changes is usually due to the spatial gradient of the image irradiance

Fundamental Equation

- From the fundamental equation of radiometric image formation

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- The illumination of the image at $\mathbf{p}$ decreases as the fourth power of the cosine of the angle formed by the principal ray through $\mathbf{p}$ with the optical axis
  - This can be neglected in the case of small angular aperture
  - Therefore, the image irradiance can be regarded as uniformly proportional to the scene radiance over the whole image plane
- The non-uniform illumination predicted by the fundamental equation is hard to notice in ordinary images
- This is because the major component of brightness changes is usually due to the spatial gradient of the image irradiance

Fundamental Equation

- From the fundamental equation of radiometric image formation

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- The illumination of the image at $\mathbf{p}$ decreases as the fourth power of the cosine of the angle formed by the principal ray through $\mathbf{p}$ with the optical axis
  - This can be neglected in the case of small angular aperture
  - Therefore, the image irradiance can be regarded as uniformly proportional to the scene radiance over the whole image plane
- The non-uniform illumination predicted by the fundamental equation is hard to notice in ordinary images
- This is because the major component of brightness changes is usually due to the spatial gradient of the image irradiance

Some Comments

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- Intensity is the grey level recorded by an image and is linked to irradiance by a monotonic relation depending on the sensor
- Brightness often indicates the subjective human perception of the image intensity
- F-number influences how much light is collected by the camera
- The small the F-number, the larger the fraction of L which reaches the image plane

Some Comments

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- Intensity is the grey level recorded by an image and is linked to irradiance by a monotonic relation depending on the sensor
- Brightness often indicates the subjective human perception of the image intensity
- F-number influences how much light is collected by the camera
- The small the F-number, the larger the fraction of L which reaches the image plane

Some Comments

$$E(\mathbf{p}) = L(\mathbf{P})\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

- Intensity is the grey level recorded by an image and is linked to irradiance by a monotonic relation depending on the sensor
- Brightness often indicates the subjective human perception of the image intensity
- F-number influences how much light is collected by the camera
- The small the F-number, the larger the fraction of L which reaches the image plane

Geometric Image Formation

- The aim of geometric image formation is to link the scene points (or object points) with the corresponding image points

- Several different camera models are used to achieve this
  - Perspective projection model (pinhole camera model)
  - Weak-perspective projection model
  - Orthographic projection model

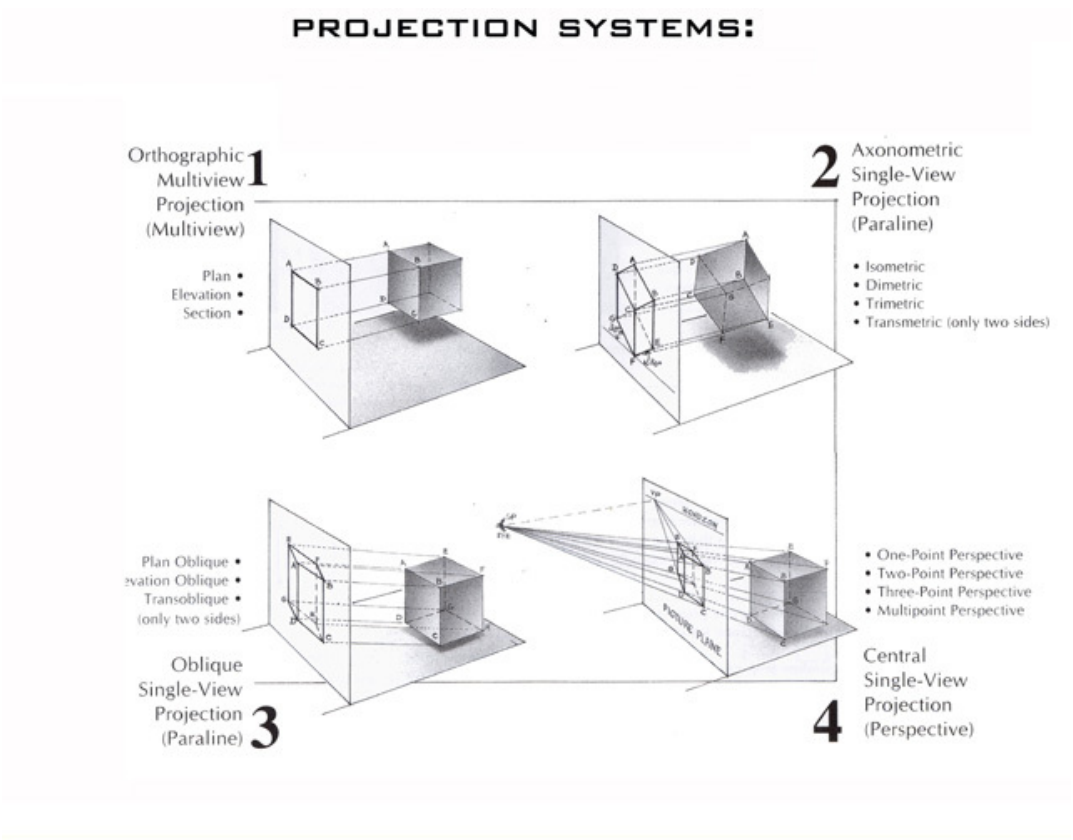
Geometric Image Formation

- The aim of geometric image formation is to link the scene points (or object points) with the corresponding image points

- Several different camera models are used to achieve this
  - Perspective projection model (pinhole camera model)
  - Weak-perspective projection model
  - Orthographic projection model

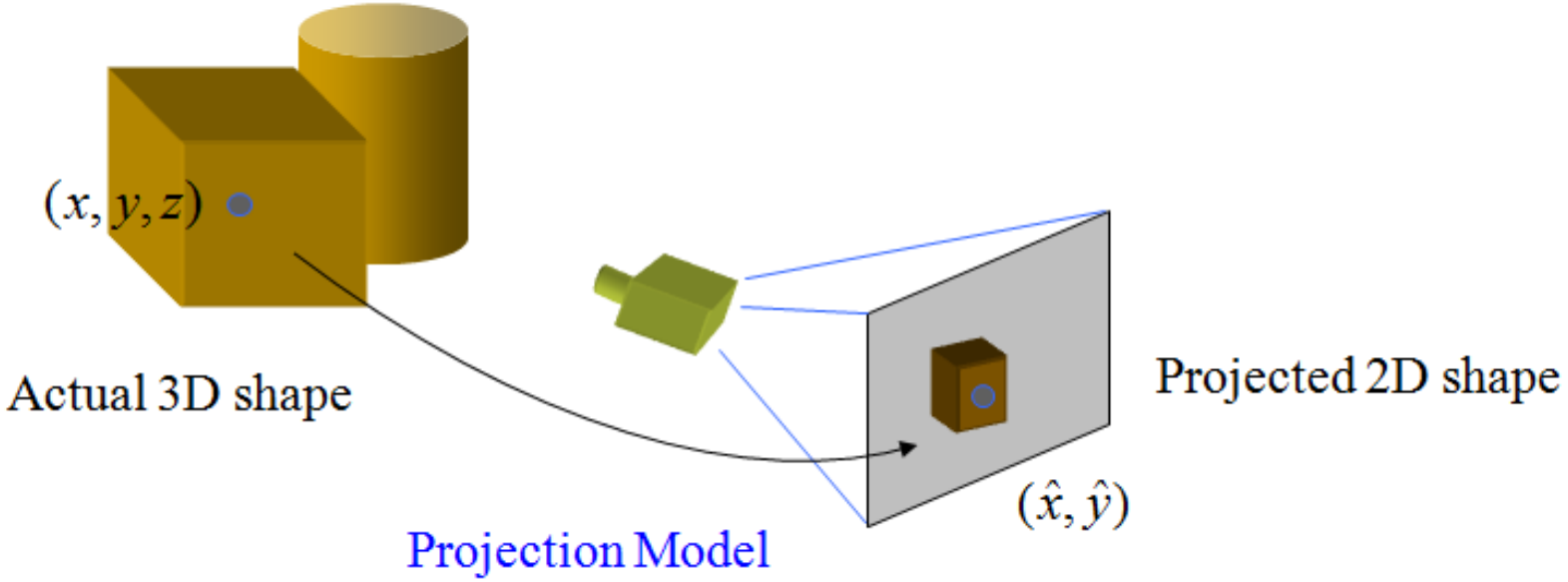
Projection Models

- Projection models are geometrical models that allow us to determine the projection of a 3D point onto a 2D photograph



Perspective Projection

- Points are projected onto the image plane along rays that pass through a pinhole camera



Emulating Perspective Projection

Woodcut from Durer's "A Treatise on Measurement" (late 15th century)

Image Geometry

- Model for the projection of scene point to image point
  - Center of projection (COP)
  - Focal length
  - Line of sight

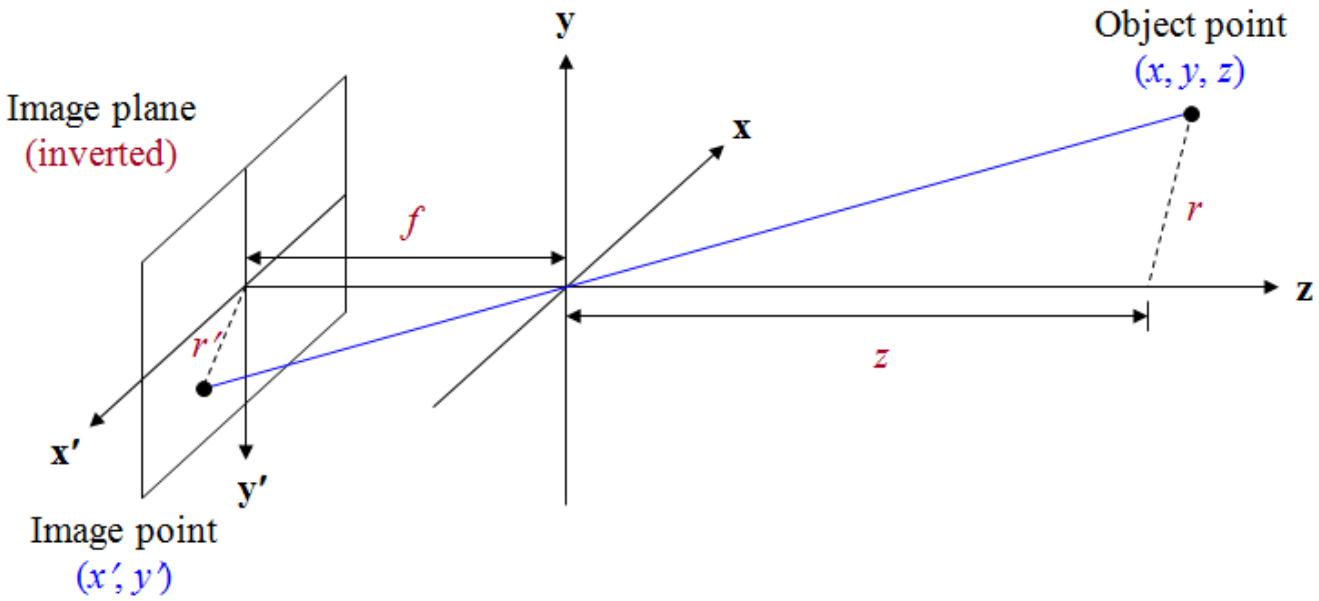
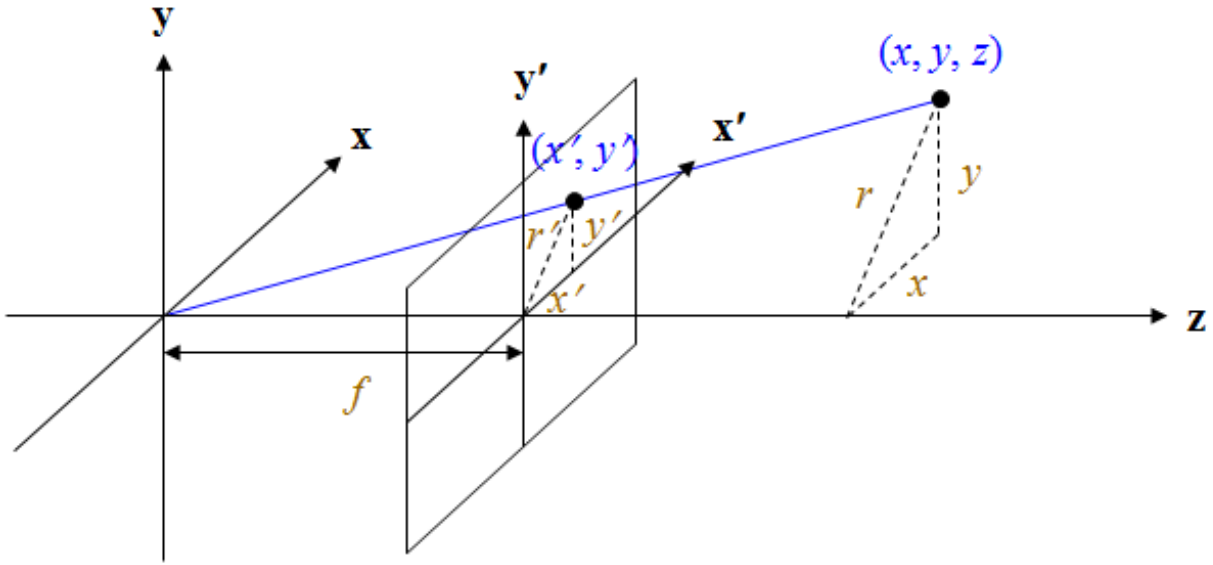


Image Geometry

- Avoid inversion by placing image plane in front of center of projection
- Invert $\mathbf{x}'$ and $\mathbf{y}'$ also



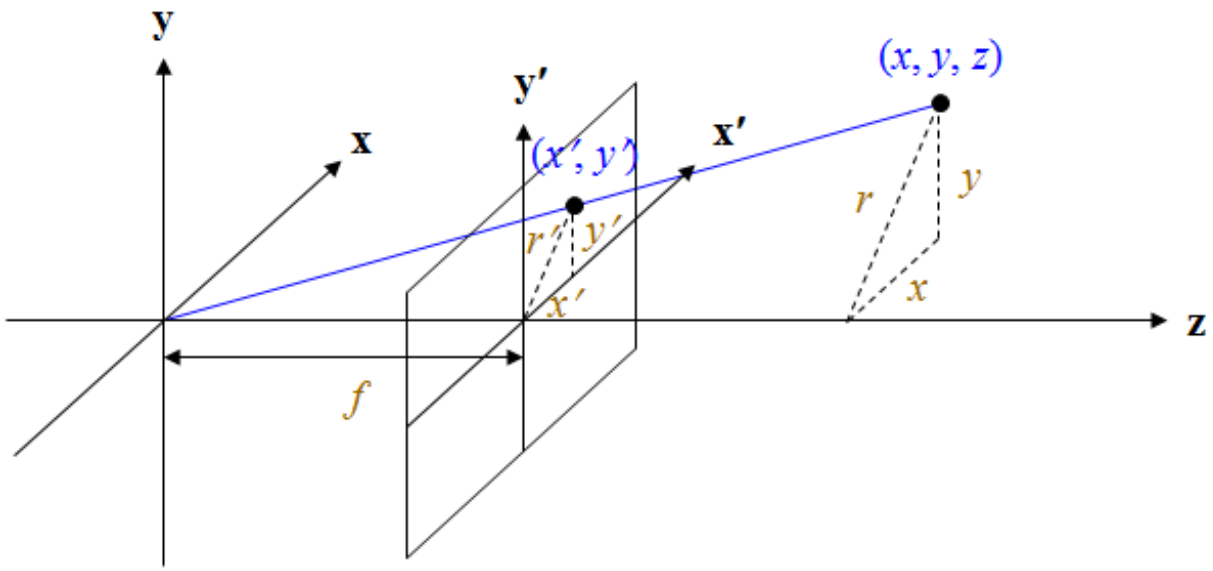
Perspective Projection

- From similar triangles:

$$\frac{f}{z} = \frac{r'}{r} \quad \text{and} \quad \frac{x'}{x} = \frac{y'}{y} = \frac{r'}{r}$$

- We have

$$x' = \frac{f}{z}x \quad \text{and} \quad y' = \frac{f}{z}y$$



Coordinate Systems

- There are four coordinate systems
  - World coordinate system
    - Also called absolute or object coordinates
    - (x_w, y_w, z_w) or (x_a, y_a, z_a)
  - Object coordinate system
    - Also called model coordinates
  - Camera coordinate system
    - (x_c, y_c, z_c)
  - Image coordinate system
    - (x', y')

- To use the image geometry, coordinates have to be transformed first!

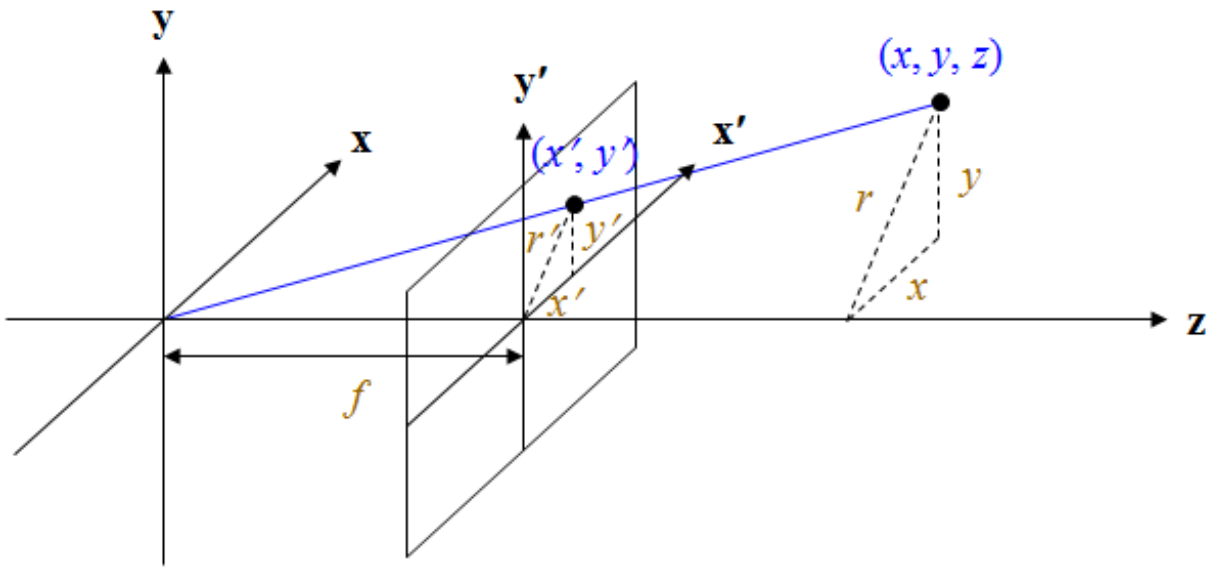
Fundamental Equations

- From similar triangles:

$$\frac{f}{Z} = \frac{r}{R} \quad \text{and} \quad \frac{x}{X} = \frac{y}{Y} = \frac{r}{R}$$

We have

$$x = f\frac{X}{Z} \quad \text{and} \quad y = f\frac{Y}{Z}$$



Weak Perspective Model

- Appropriate when the depth of field of the relevant objects is small compared to the distance of those objects from the center of projection

- Suppose the average distance is $\bar{Z}$, then

$$x = f\frac{X}{Z} \approx f\frac{X}{\bar{Z}} \tag{1}$$

$$y = f\frac{Y}{Z} \approx f\frac{Y}{\bar{Z}} \tag{2}$$

Weak Perspective Model

- Appropriate when the depth of field of the relevant objects is small compared to the distance of those objects from the center of projection

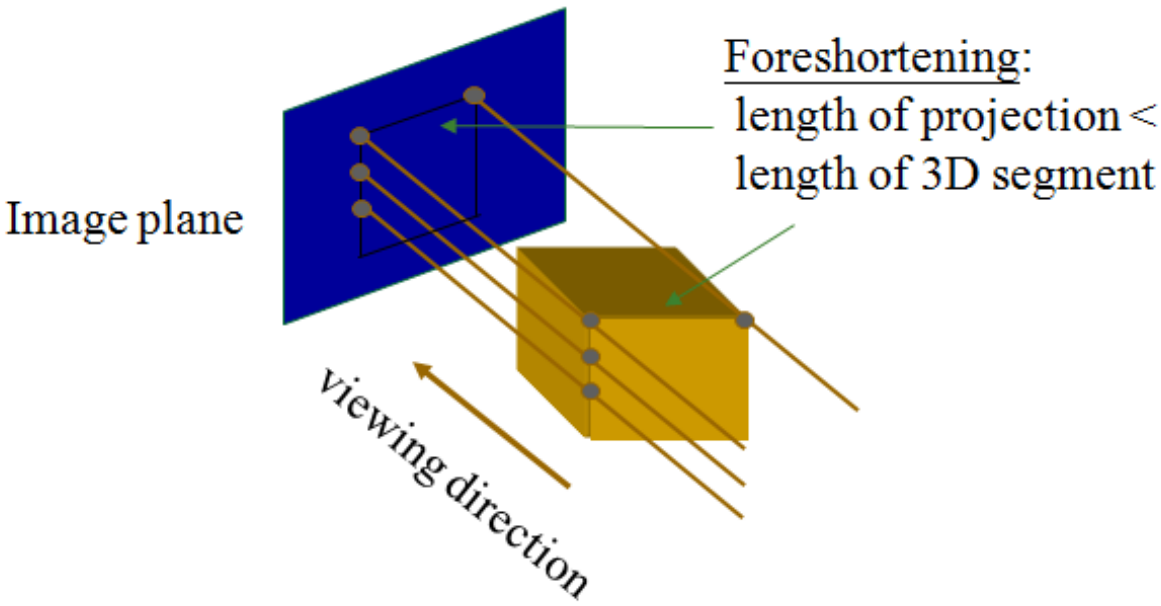
- Suppose the average distance is $\bar{Z}$, then

$$x = f\frac{X}{Z} \approx f\frac{X}{\bar{Z}} \tag{1}$$

$$y = f\frac{Y}{Z} \approx f\frac{Y}{\bar{Z}} \tag{2}$$

Orthographic Projection

- 3D points are projected along parallel rays onto the image plane



Summary

Perspective camera model

In the perspective camera model (nonlinear), the coordinates (x, y) of a point $\mathbf{p}$, image of the 3-D point $\mathbf{P} = [X, Y, Z]^T$, are given by

$$x = f\frac{X}{Z} \qquad \text{and} \qquad y = f\frac{Y}{Z} \tag{3}$$

Weak-perspective camera model

If the average depth of the scene, $\bar{Z}$, is much larger than the relative distance between any two scene points along the optical axis, the weak-perspective camera model (linear) holds:

$$x = f\frac{X}{\bar{Z}} \qquad \text{and} \qquad y = f\frac{Y}{\bar{Z}} \tag{4}$$

Summary

Perspective camera model

In the perspective camera model (nonlinear), the coordinates (x, y) of a point $\mathbf{p}$, image of the 3-D point $\mathbf{P} = [X, Y, Z]^T$, are given by

$$x = f\frac{X}{Z} \qquad \text{and} \qquad y = f\frac{Y}{Z} \tag{3}$$

Weak-perspective camera model

If the average depth of the scene, $\bar{Z}$, is much larger than the relative distance between any two scene points along the optical axis, the weak-perspective camera model (linear) holds:

$$x = f\frac{X}{\bar{Z}} \qquad \text{and} \qquad y = f\frac{Y}{\bar{Z}} \tag{4}$$

Camera Parameters

- The fundamental equations links the 3-D object point with the corresponding 2-D image point

$$x = f\frac{X}{Z} \quad \text{and} \quad y = f\frac{Y}{Z}$$

- The equations are written in the camera reference frame and assume
  - The camera reference frame can be located w.r.t. some other known reference frame (e.g. the world reference frame)
  - The coordinates of the image points in the camera reference frame can be obtained from pixel coordinates (in pixels)
- The above assumptions are equivalent to knowing the camera's characteristics, i.e., camera's parameters (intrinsic and extrinsic parameters)

Camera Parameters

- The fundamental equations links the 3-D object point with the corresponding 2-D image point

$$x = f\frac{X}{Z} \quad \text{and} \quad y = f\frac{Y}{Z}$$

- The equations are written in the camera reference frame and assume
  - The camera reference frame can be located w.r.t. some other known reference frame (e.g. the world reference frame)
  - The coordinates of the image points in the camera reference frame can be obtained from pixel coordinates (in pixels)
- The above assumptions are equivalent to knowing the camera's characteristics, i.e., camera's parameters (intrinsic and extrinsic parameters)

Camera Parameters

- The fundamental equations links the 3-D object point with the corresponding 2-D image point

$$x = f\frac{X}{Z} \quad \text{and} \quad y = f\frac{Y}{Z}$$

- The equations are written in the camera reference frame and assume
  - The camera reference frame can be located w.r.t. some other known reference frame (e.g. the world reference frame)
  - The coordinates of the image points in the camera reference frame can be obtained from pixel coordinates (in pixels)
- The above assumptions are equivalent to knowing the camera's characteristics, i.e., camera's parameters (intrinsic and extrinsic parameters)

Camera Parameters

- Extrinsic parameters:
  - Define the location and orientation of the camera reference frame w.r.t. a known world reference frame
- Intrinsic parameters:
  - Link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame

- The problem of estimating the value of these parameters is called camera calibration

Camera Parameters

- Extrinsic parameters:
  - Define the location and orientation of the camera reference frame w.r.t. a known world reference frame
- Intrinsic parameters:
  - Link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame

- The problem of estimating the value of these parameters is called camera calibration

Camera Calibration

- Why camera calibration?
  - A direct measurement of the parameters (intrinsic and extrinsic) of a given and positioned camera in 3D scene space is usually impossible or not feasible

- The parameters are calculated indirectly in practice using a suitable calibration technique
- To achieve the best possible correspondence between the used camera model and the realized image acquisition with the given camera

Camera Calibration

- Why camera calibration?
  - A direct measurement of the parameters (intrinsic and extrinsic) of a given and positioned camera in 3D scene space is usually impossible or not feasible

- The parameters are calculated indirectly in practice using a suitable calibration technique
- To achieve the best possible correspondence between the used camera model and the realized image acquisition with the given camera

Extrinsic Parameters

- The world reference frame can be assigned by ourselves
- To use the fundamental equation of perspective projection, we need to know the camera reference frame (which is unknown!)

- To find the camera reference frame, we need to determine the location and orientation of this camera reference frame w.r.t. the world reference frame using only image information
- That is, to determine the camera's extrinsic parameters:
  - Any set of geometric parameters that identify uniquely the transformation between the unknown camera reference frame and a known world reference frame

Extrinsic Parameters

- The world reference frame can be assigned by ourselves
- To use the fundamental equation of perspective projection, we need to know the camera reference frame (which is unknown!)

- To find the camera reference frame, we need to determine the location and orientation of this camera reference frame w.r.t. the world reference frame using only image information
- That is, to determine the camera's extrinsic parameters:
  - Any set of geometric parameters that identify uniquely the transformation between the unknown camera reference frame and a known world reference frame

Extrinsic Parameters

- The transformation between camera and world reference frames can be specified by
  - A 3-D translation vector $\mathbf{T}$:
    - Describes the relative positions of the origins of the camera and world reference frames
  - A 3×3 rotation matrix $\mathtt{R}$:
    - Brings the corresponding axes of the camera and world reference frames onto each other

- The rotation matrix $\mathtt{R}$ is orthogonal ($\mathtt{R}^T\mathtt{R} = \mathtt{R}\mathtt{R}^T = I$), thus the degrees of freedom is only 3 (not 9)

Extrinsic Parameters

- The transformation between camera and world reference frames can be specified by
  - A 3-D translation vector $\mathbf{T}$:
    - Describes the relative positions of the origins of the camera and world reference frames
  - A 3×3 rotation matrix $\mathtt{R}$:
    - Brings the corresponding axes of the camera and world reference frames onto each other

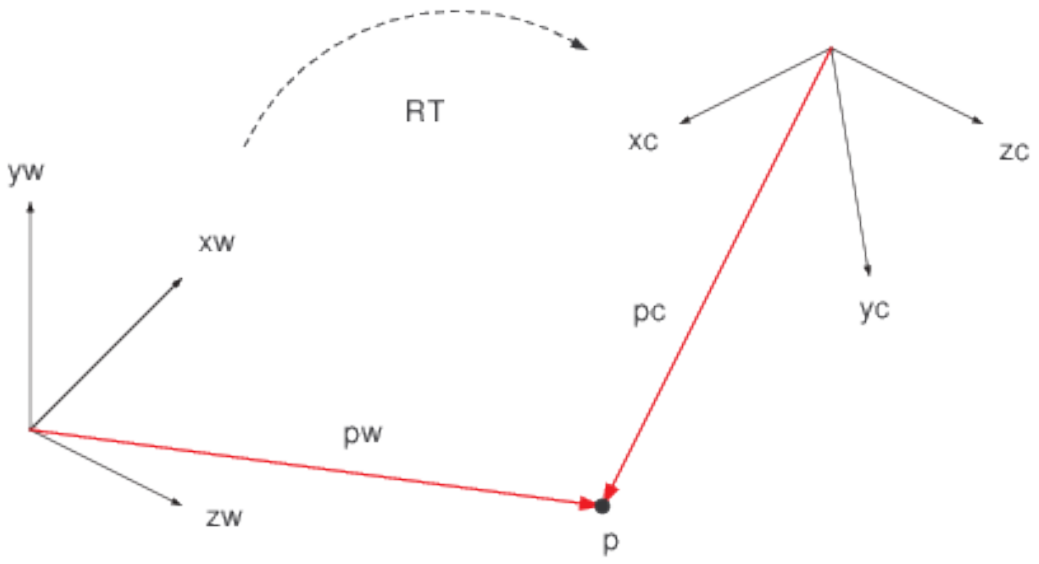
- The rotation matrix $\mathtt{R}$ is orthogonal ($\mathtt{R}^T\mathtt{R} = \mathtt{R}\mathtt{R}^T = I$), thus the degrees of freedom is only 3 (not 9)

World and Camera Frames

- For the same point $\mathbf{P}$, the relation between world and camera reference frames, $\mathbf{P}_w$ and $\mathbf{P}_c$ respectively is given by

$$\mathbf{P}_c = \mathtt{R}(\mathbf{P}_w - \mathbf{T})$$

- Camera's extrinsic parameters:
  - Translation vector $\mathbf{T}$ and rotation matrix R, which specify the transformation between camera and world reference frames



Intrinsic Parameters

- The intrinsic parameters are used to characterize the optical, geometric, and digital characteristics of the viewing camera
- For a pinhole camera, there are three sets of intrinsic parameters
  - The perspective projection (focal length)
  - The transformation between camera frame coordinates and pixel coordinates
  - The geometric distortion introduced by the optics

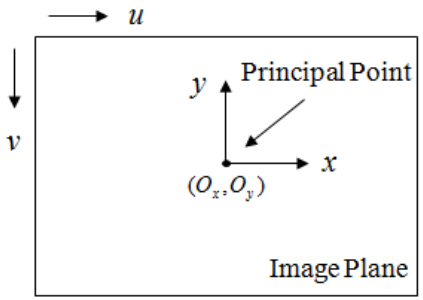
Camera and Image Frames

- Let (x_{im}, y_{im}) be the image (or pixel) coordinate (in pixel) and (x, y) be the same point in camera reference frame
- Let (o_x, o_y) be the image center coordinate in pixel and (s_x, s_y) be the effective pixel size in mm in horizontal and vertical directions
- The transformation between camera and image frame coordinates is given by

$$x = -s_x(x_{im} - o_x) \tag{5}$$

$$y = -s_y(y_{im} - o_y) \tag{6}$$

- Observe the sign change!



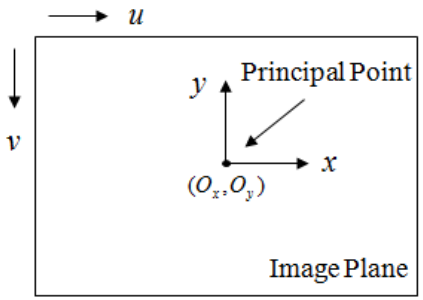
Camera and Image Frames

- Let (x_{im}, y_{im}) be the image (or pixel) coordinate (in pixel) and (x, y) be the same point in camera reference frame
- Let (o_x, o_y) be the image center coordinate in pixel and (s_x, s_y) be the effective pixel size in mm in horizontal and vertical directions
- The transformation between camera and image frame coordinates is given by

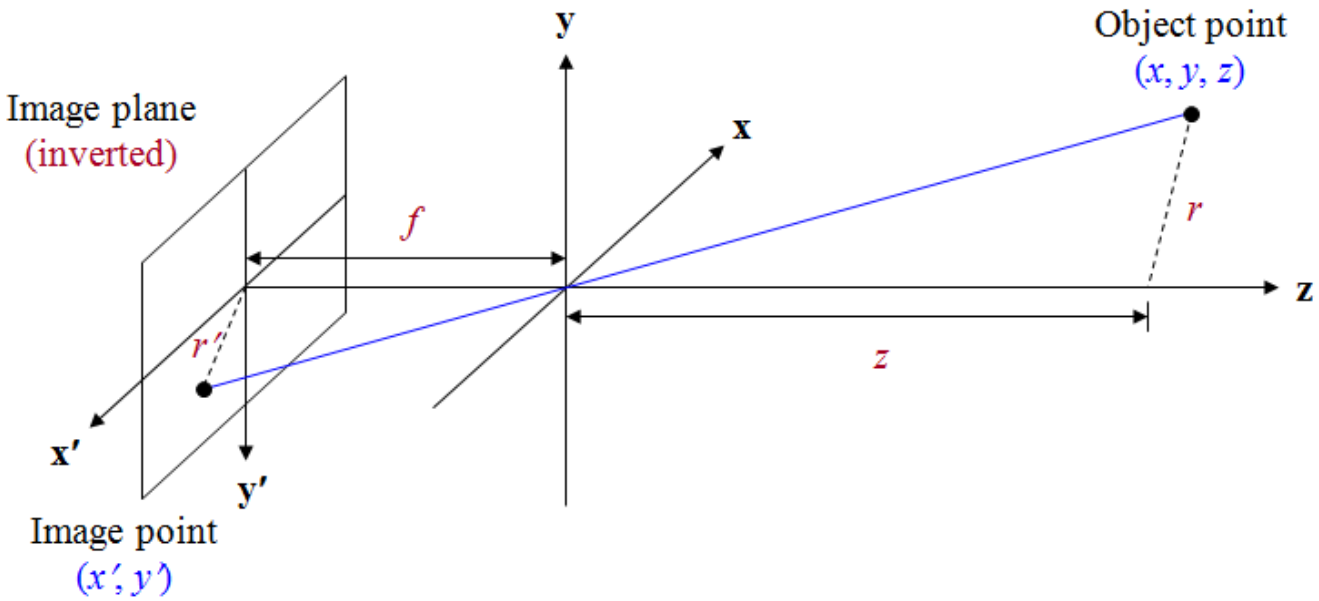
$$x = -s_x(x_{im} - o_x) \tag{5}$$

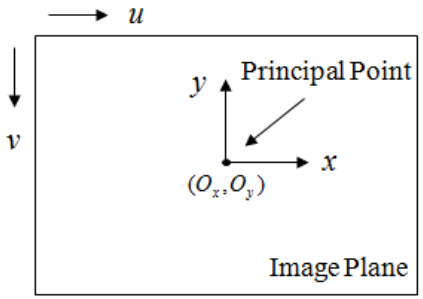
$$y = -s_y(y_{im} - o_y) \tag{6}$$

- Observe the sign change!



Camera and Image Frames





Radial Distortion

- The optics introduces image distortions at the periphery of the image

Radial Distortion

- The distortions are modeled by radial distortions:

$$x = x_d(1 + k_1 r^2 + k_2 r^4) \tag{7}$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4) \tag{8}$$

  where (x_d, y_d) is the distorted point and $r^2 = x_d^2 + y_d^2$
- The distortion is a radial displacement of image point
  - Null at the image center
  - Increases with the distance of the point from the image center
- Thus, k_1 and k_2 are two more intrinsic parameters
  - k_1 and k_2 are usually small and can be ignored unless high accuracy is required
  - Even if radial distortion has to be considered, k_2 is usually set to 0 since $k_2 << k_1$

Radial Distortion

- The distortions are modeled by radial distortions:

$$x = x_d(1 + k_1 r^2 + k_2 r^4) \tag{7}$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4) \tag{8}$$

  where (x_d, y_d) is the distorted point and $r^2 = x_d^2 + y_d^2$
- The distortion is a radial displacement of image point
  - Null at the image center
  - Increases with the distance of the point from the image center
- Thus, k_1 and k_2 are two more intrinsic parameters
  - k_1 and k_2 are usually small and can be ignored unless high accuracy is required
  - Even if radial distortion has to be considered, k_2 is usually set to 0 since $k_2 << k_1$

Radial Distortion

- The distortions are modeled by radial distortions:

$$x = x_d(1 + k_1 r^2 + k_2 r^4) \tag{7}$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4) \tag{8}$$

  where (x_d, y_d) is the distorted point and $r^2 = x_d^2 + y_d^2$
- The distortion is a radial displacement of image point
  - Null at the image center
  - Increases with the distance of the point from the image center
- Thus, k_1 and k_2 are two more intrinsic parameters
  - k_1 and k_2 are usually small and can be ignored unless high accuracy is required
  - Even if radial distortion has to be considered, k_2 is usually set to 0 since $k_2 << k_1$

Summary – Intrinsic Parameters

- For a pinhole camera, there are three sets of intrinsic parameters:
  - The focal length f
  - The location of the image center in pixel coordinates, (o_x, o_y)
  - The effective pixel size in the horizontal and vertical direction (s_x, s_y)
  - The radial distortion coefficients, k_1, k_2

Camera Parameter Matrices

- If we don't consider the radial distortion, the equations

$$-s_x(x_{im} - o_x) = f\frac{\mathbf{R}_1^T(\mathbf{P}_w - \mathbf{T})}{\mathbf{R}_3^T(\mathbf{P}_w - \mathbf{T})} \tag{9}$$

$$-s_y(y_{im} - o_y) = f\frac{\mathbf{R}_2^T(\mathbf{P}_w - \mathbf{T})}{\mathbf{R}_3^T(\mathbf{P}_w - \mathbf{T})} \tag{10}$$

can be written as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int}M_{ext}\begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

Camera Parameter Matrices

- where

$$M_{int} = \begin{bmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

depends only on the intrinsic parameters

$$M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -\mathbf{R}_1^T\mathbf{T} \\ r_{21} & r_{22} & r_{23} & -\mathbf{R}_2^T\mathbf{T} \\ r_{31} & r_{32} & r_{33} & -\mathbf{R}_3^T\mathbf{T} \end{bmatrix}$$

depends only on the extrinsic parameters

Camera Parameter Matrices

- From the equation below, the image coordinates are given by the ratios x_1/x_3 and x_2/x_3

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

- That is, $(x_{im}, y_{im}) = (\frac{x_1}{x_3}, \frac{x_2}{x_3})$. Or we can write the equation as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

  where λ is a scale factor

- Thus, the perspective projection is unique up to an arbitrary scale!

Camera Parameter Matrices

- From the equation below, the image coordinates are given by the ratios x_1/x_3 and x_2/x_3

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

- That is, $(x_{im}, y_{im}) = (\frac{x_1}{x_3}, \frac{x_2}{x_3})$. Or we can write the equation as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

  where λ is a scale factor
- Thus, the perspective projection is unique up to an arbitrary scale!

Perspective Projection Matrix

- From the perspective projection matrix equation

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

  we have separated the perspective projection from world coordinates to image coordinates to two steps:
  - M_{ext} performs the transformation between the world and the camera reference frame
  - M_{int} performs the transformation between the camera and the image reference frame
- This is something related to projective line, projective plane, and projective geometry (Appendix)

Perspective Projection Matrix

- From the perspective projection matrix equation

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

  we have separated the perspective projection from world coordinates to image coordinates to two steps:
  - M_{ext} performs the transformation between the world and the camera reference frame
  - M_{int} performs the transformation between the camera and the image reference frame
- This is something related to projective line, projective plane, and projective geometry (Appendix)

Camera Model via Projection Matrix

- From the equation

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} = M \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

  the matrix $M = M_{ext} M_{int}$ is called projection matrix
- Various camera models can be derived by setting appropriate constraints on the projection matrix M

Camera Model via Projection Matrix

- The perspective camera model (with $o_x = o_y = 0$, $s_x = s_y = 1$)

$$M = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & f\mathbf{R}_1^T\mathbf{T} \\ -fr_{21} & -fr_{22} & -fr_{23} & -f\mathbf{R}_2^T\mathbf{T} \\ r_{31} & r_{32} & r_{33} & -\mathbf{R}_3^T\mathbf{T} \end{bmatrix}$$

- The weak-perspective camera model

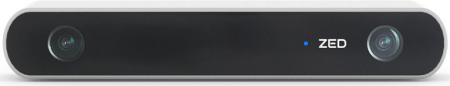
$$M = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & f\mathbf{R}_1^T\mathbf{T} \\ -fr_{21} & -fr_{22} & -fr_{23} & -f\mathbf{R}_2^T\mathbf{T} \\ 0 & 0 & 0 & -\mathbf{R}_3^T(\bar{\mathbf{P}} - \mathbf{T}) \end{bmatrix}$$

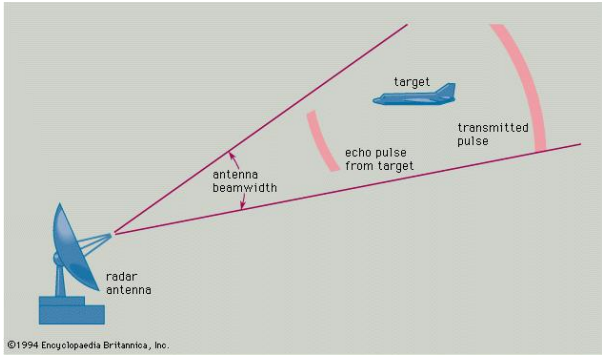
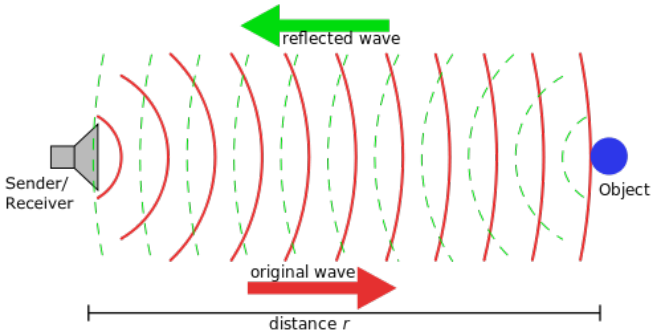
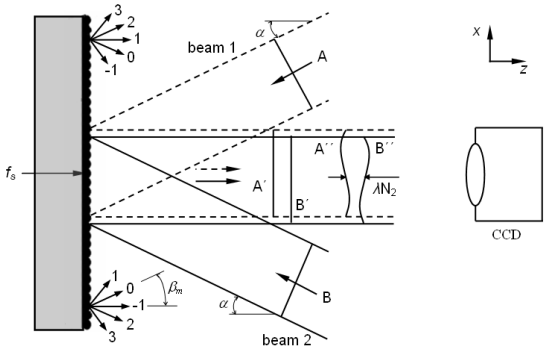
Range Image Representation

- Range image can be represented in two basic forms
  - A list of 3-D coordinates in a given reference frame, called xyz form or point cloud, for which no specific order is required
  - A matrix of depth values of points along the direction of the x, y image axes, called the r_{ij} form, which makes spatial information explicit
- Range images are also called
  - Depth images, depth maps, xyz maps, surface profiles, 2.5-D images
- r_{ij} data can always be visualized as a normal intensity image

Range Sensors

- An optical range sensor is a device using optical phenomena to acquire range images
  - Active range sensors – project energy (e.g., a pattern of light, sonar pulses) on the scene and detect its position to perform the measure; or exploit the effect of controlled changes of some sensor parameters (e.g., focus)
  - Passive range sensors – rely only on intensity images to reconstruct depth (e.g., stereopsis)



Active Range Sensors

- Radar and Sonars
- Moire Interferometry
- Active Focusing/Defocusing
  - These methods infer range from two or more images of the same scene, acquired under varying focus settings