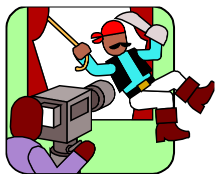
Image Formation

CSc I6716
Spring 2011



Topic 1 of Part I
Image Formation

Zhigang Zhu, City College of New York zhu@cs.ccny.cuny.edu

Acknowledgements

The slides in this lecture were kindly provided by

Professor Allen Hanson
University of Massachusetts at Amherst

3D Computer Vision
and Video Computing

Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

3D Computer Vision
and Video Computing

Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

3D Computer Vision
and Video Computing

Abstract Image

- An image can be represented by an image function whose general form is **f(x,y)**.
- **f(x,y)** is a vector-valued function whose arguments represent a pixel location.
- The value of **f(x,y)** can have different interpretations in different kinds of images.

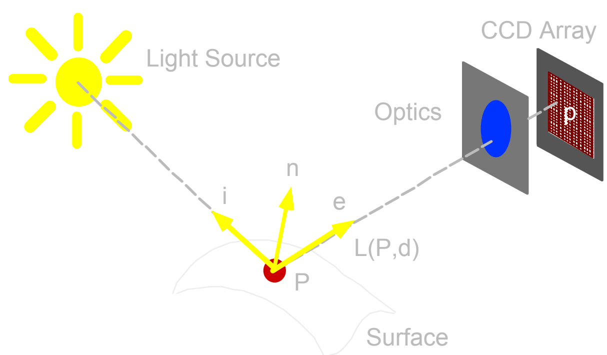
Examples

Intensity Image	- f(x,y) = intensity of the scene
Range Image	- f(x,y) = depth of the scene from imaging system
Color Image	- **f**(x,y) = $\{f_r(x,y), f_g(x,y), f_b(x,y)\}$
Video	- f(x,y,t) = temporal image sequence

3D Computer Vision
and Video Computing

Basic Radiometry

- Radiometry is the 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.



Light and Matter

- The interaction between light and matter can take many forms:
  - Reflection
  - Refraction
  - Diffraction
  - Absorption
  - Scattering

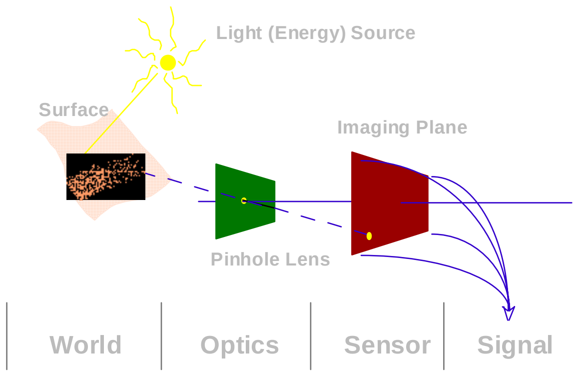
Lecture Assumptions

- Typical imaging scenario:
  - visible light
  - ideal lenses
  - standard sensor (e.g. TV camera)
  - opaque objects
- Goal

> To create 'digital' images which can be processed to recover some of the characteristics of the 3D world which was imaged.

3D Computer Vision
and Video Computing

Image Formation



Sensor	Signal
B&W Film	Silver Density
Color Film	Silver density in three color layers
TV Camera	Electrical

3D Computer Vision
and Video Computing

Steps

World → Optics → Sensor →
Signal → Digitizer →
Digital Representation

World	reality
Optics	focus {light} from world on sensor
Sensor	converts {light} to {electrical energy}
Signal	representation of incident light as continuous electrical energy
Digitizer	converts continuous signal to discrete signal
Digital Rep.	final representation of reality in computer memory

Factors in Image Formation

- Geometry
  - concerned with the relationship between points in the three-dimensional world and their images
- Radiometry
  - concerned with the relationship between the amount of light radiating from a surface and the amount incident at its image
- Photometry
  - concerned with ways of measuring the intensity of light
- Digitization
  - concerned with ways of converting continuous signals (in both space and time) to digital approximations

Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

Geometry

- Geometry describes the projection of:

three-dimensional (3D) world → two-dimensional (2D) image plane.

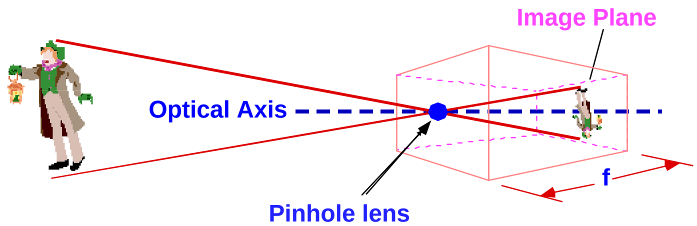
- Typical Assumptions
  - Light travels in a straight line
- **Optical Axis**: the axis perpendicular to the image plane and passing through the pinhole (also called the central projection ray)
- Each point in the image corresponds to a particular direction defined by a **ray** from that point through the pinhole.
- Various kinds of projections:
  - - perspective - oblique
  - - orthographic - isometric
  - - spherical

Basic Optics

- Two models are commonly used:
  - Pin-hole camera
  - Optical system composed of lenses
- **Pin-hole** is the basis for most graphics and vision
  - Derived from physical construction of early cameras
  - Mathematics is very straightforward
- **Thin lens** model is first of the lens models
  - Mathematical model for a physical lens
  - Lens gathers light over area and focuses on image plane.

3D Computer Vision
and Video Computing

Pinhole Camera Model



- World projected to 2D Image
  - Image inverted
  - Size reduced
  - Image is dim
  - No direct depth information
- f called the focal length of the lens
- Known as perspective projection

3D Computer Vision
and Video Computing

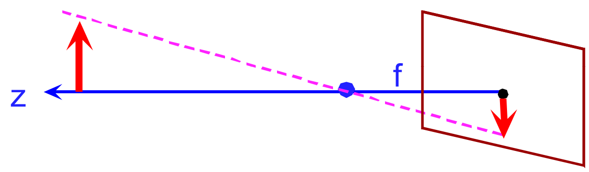
Pinhole camera image

Amsterdam

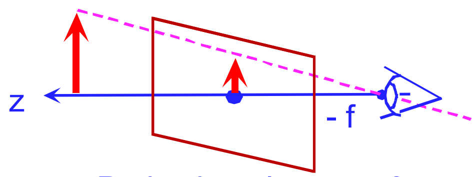
Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

3D Computer Vision and Video Computing — Equivalent Geometry

- Consider case with object on the optical axis:



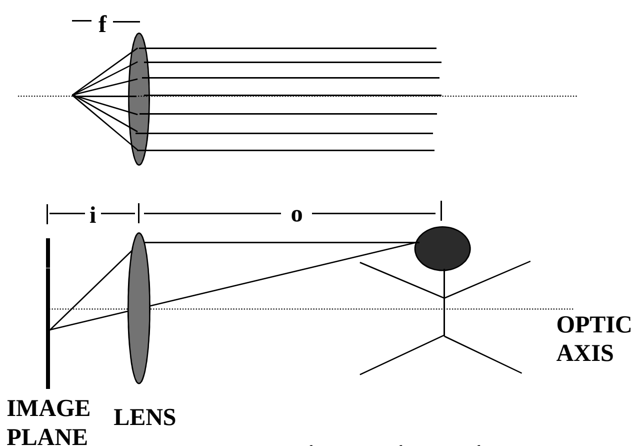
- More convenient with upright image:



- Equivalent mathematically

3D Computer Vision and Video Computing — Thin Lens Model

- Rays entering parallel on one side converge at focal point.
- Rays diverging from the focal point become parallel.

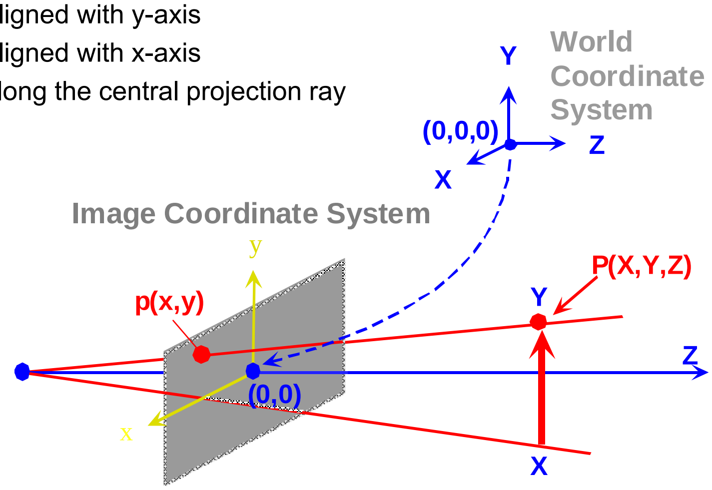


$$\frac{1}{f} = \frac{1}{i} + \frac{1}{o}$$ 'THIN LENS LAW'

3D Computer Vision
and Video Computing

Coordinate System

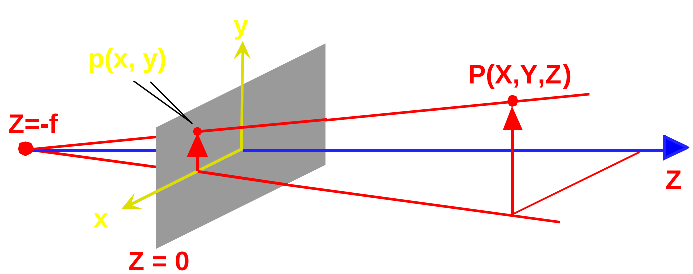
- Simplified Case:
  - Origin of world and image coordinate systems coincide
  - Y-axis aligned with y-axis
  - X-axis aligned with x-axis
  - Z-axis along the central projection ray



3D Computer Vision
and Video Computing

Perspective Projection

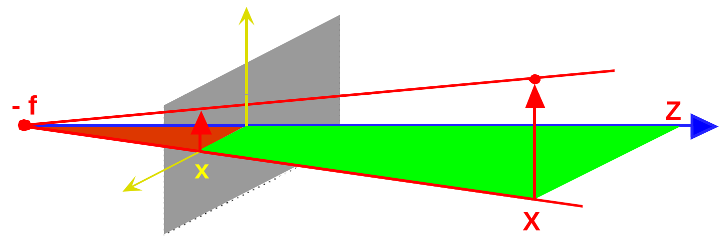
- Compute the image coordinates of p in terms of the world coordinates of P.



- Look at projections in x-z and y-z planes

3D Computer Vision
and Video Computing

X-Z Projection

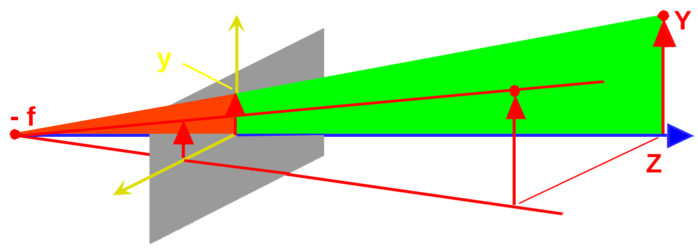


- By similar triangles: $\frac{x}{f} = \frac{X}{Z+f}$

$$x = \frac{fX}{Z+f}$$

3D Computer Vision
and Video Computing

Y-Z Projection



- By similar triangles: $\frac{y}{f} = \frac{Y}{Z+f}$

$$y = \frac{fY}{Z+f}$$

Perspective Equations

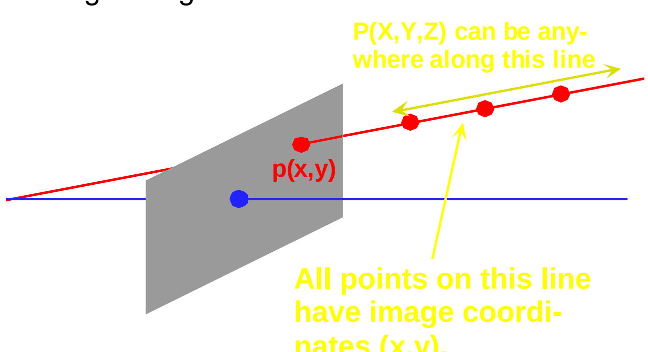
- Given point P(X,Y,Z) in the 3D world
- The two equations:

$$x = \frac{fX}{Z+f} \qquad y = \frac{fY}{Z+f}$$

- transform world coordinates (X,Y,Z)
  into image coordinates (x,y)
- Question:
  - What is the equation if we select the origin of both coordinate systems at the nodal point?

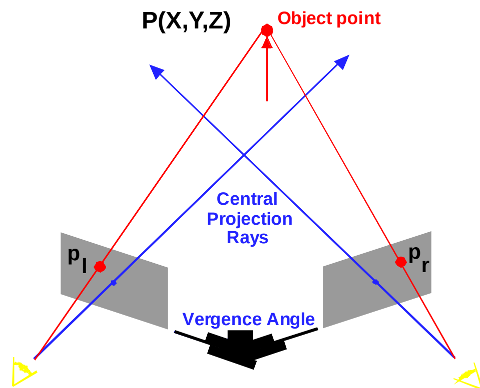
Reverse Projection

- Given a center of projection and image coordinates of a point, it is not possible to recover the 3D depth of the point from a single image.



In general, at least two images of the same point taken from two different locations are required to recover depth.

Stereo Geometry



- Depth obtained by triangulation
- Correspondence problem: p_l and p_r must correspond to the left and right projections of P, respectively.

Lecture Outline

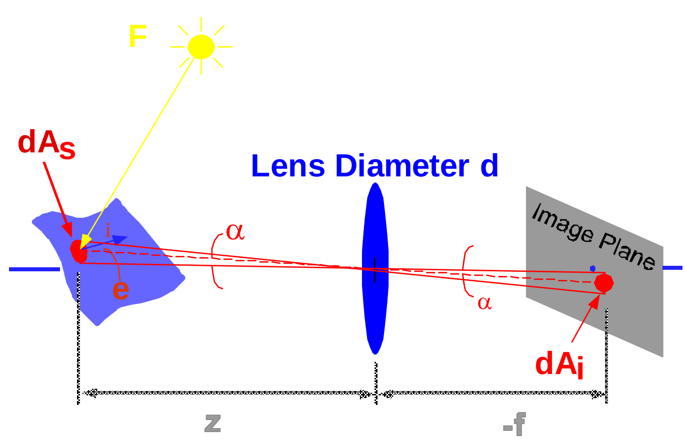
- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

Radiometry

- Image: two-dimensional array of 'brightness' values.
- Geometry: where in an image a point will project.
- Radiometry: what the brightness of the point will be.
  - Brightness: informal notion used to describe both scene and image brightness.
  - Image brightness: related to energy flux incident on the image plane: => IRRADIANCE
  - Scene brightness: brightness related to energy flux emitted (radiated) from a surface: => RADIANCE

Radiometry & Geometry

- Goal: Relate the radiance of a surface to the irradiance in the image plane of a simple optical system.



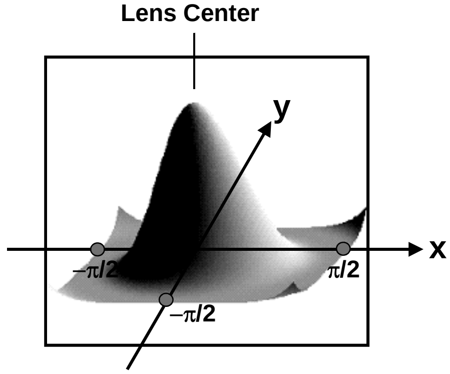
Radiometry Final Result

$$E_i = L_s \frac{\pi}{4} \left(\frac{d}{-f} \right)^2 \cos^4 \alpha$$

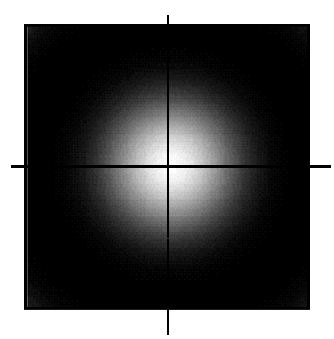
- Image irradiance is proportional to:
  - Scene radiance Ls
  - Focal length of lens f
  - Diameter of lens d
    - f/d is often called the **f-number** of the lens
  - Off-axis angle α

$Cos^4 \alpha$ Light Falloff

Lens Center



Top view shaded by height



Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

Photometry

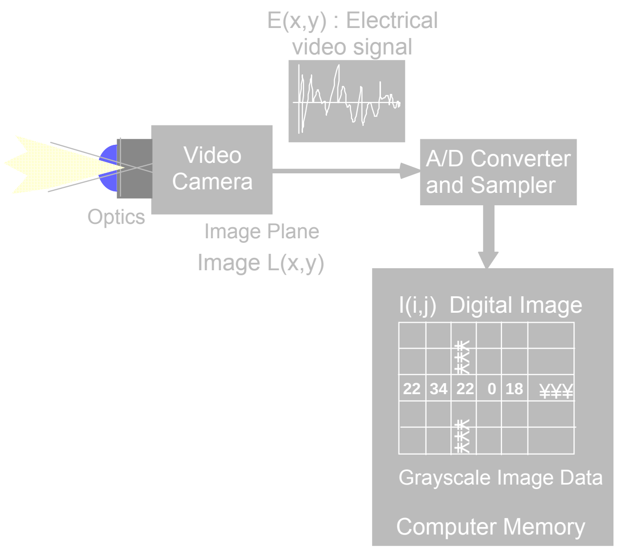
- Photometry:

  Concerned with mechanisms for converting light energy into electrical energy.

World → Optics → Sensor → Signal → Digitizer → Digital Representation

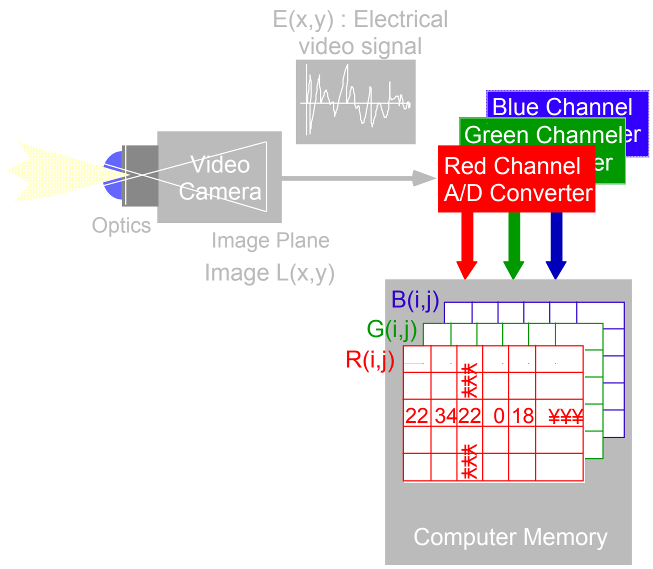
3D Computer Vision
and Video Computing

B&W Video System



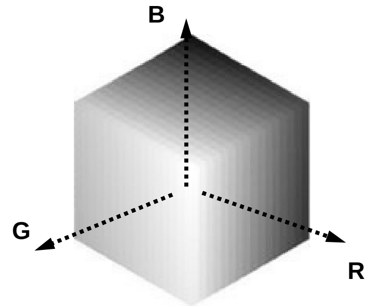
3D Computer Vision
and Video Computing

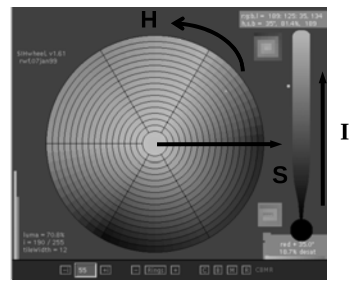
Color Video System



Color Representation

- Color Cube and Color Wheel





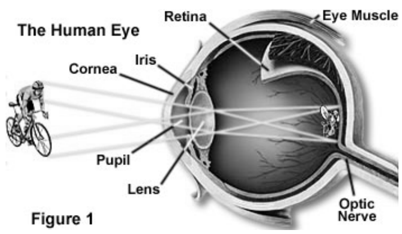
- For color spaces, please read
  - Color Cube http://www.morecrayons.com/palettes/webSmart/
  - Color Wheel http://r0k.us/graphics/SIHwheel.html
  - http://www-viz.tamu.edu/faculty/parke/ends489f00/notes/sec1_4.html

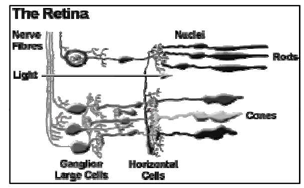
Digital Color Cameras

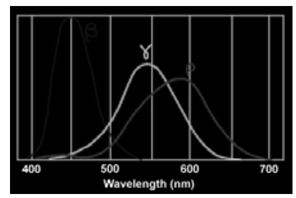
- Three CCD-chips cameras
  - R, G, B separately, AND digital signals instead analog video
- One CCD Cameras
  - Bayer color filter array
  - http://www.siliconimaging.com/RGB%20Bayer.htm

3D Computer Vision
and Video Computing

Human Eyes & Color Perception







- Visit a cool site with Interactive Java tutorial:
  - **Human Vision and Color Perception**
- Another site about human color perception:
  - **Color Vision**

3D Computer Vision
and Video Computing

Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

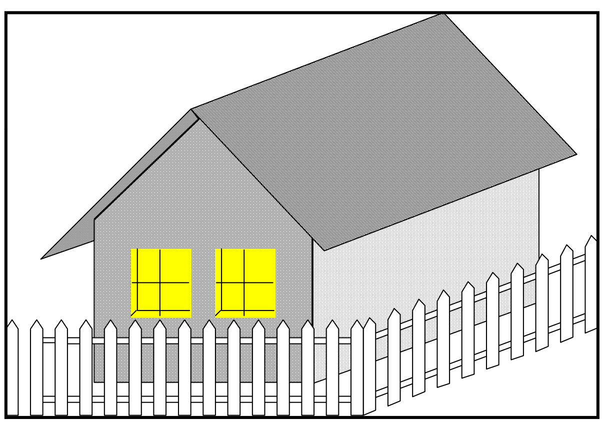
Digitization

World → Optics → Sensor →
Signal → Digitizer →
Digital Representation

- Digitization: conversion of the continuous (in space and value) electrical signal into a digital signal (digital image)
- Three decisions must be made:
  - Spatial resolution (how many samples to take)
  - Signal resolution (dynamic range of values- quantization)
  - Tessellation pattern (how to 'cover' the image with sample points)

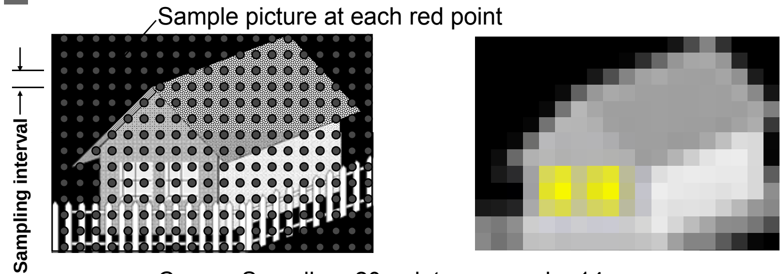
3D Computer Vision
and Video Computing

Digitization: Spatial Resolution

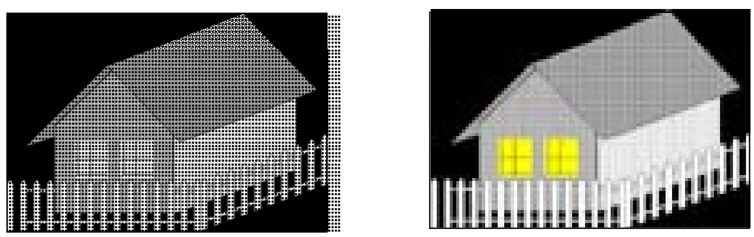


- Let's digitize this image
  - Assume a square sampling pattern
  - Vary density of sampling grid

Spatial Resolution



Coarse Sampling: 20 points per row by 14 rows

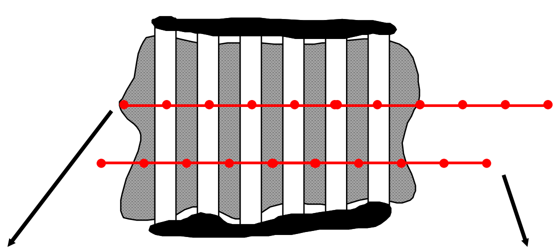


Finer Sampling: 100 points per row by 68 rows

Effect of Sampling Interval - 1

- Look in vicinity of the picket fence:

Sampling Interval:



100	100	100	100	100	100
100	100	100	100	100	100
100	100	100	100	100	100
100	100	100	100	100	100

NO EVIDENCE OF THE FENCE!

40	40	40	40	40	40
40	40	40	40	40	40
40	40	40	40	40	40
40	40	40	40	40	40

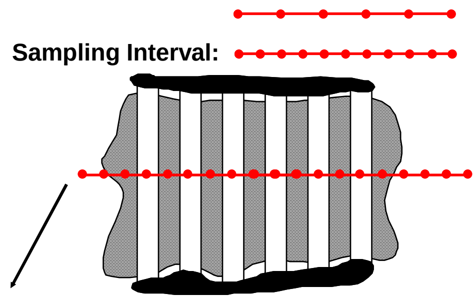
White Image!

Dark Gray Image!

Effect of Sampling Interval - 2

- Look in vicinity of picket fence:

Sampling Interval:



40	100	40	100	40
40	100	40	100	40
40	100	40	100	40
40	100	40	100	40

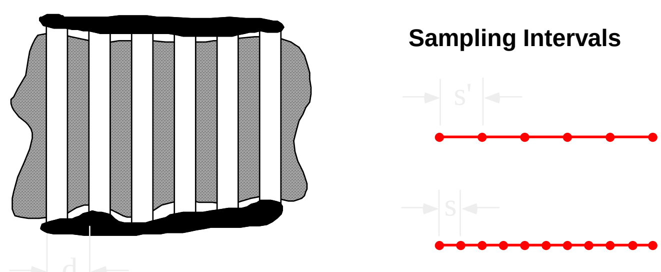
[redacted]

Now we've got a fence!

The Missing Fence Found

- Consider the repetitive structure of the fence:

Sampling Intervals



Case 1: s' = d	The sampling interval is equal to the size of the repetitive structure	**NO FENCE**
Case 2: s = d/2	The sampling interval is one-half the size of the repetitive structure	**FENCE**

The Sampling Theorem

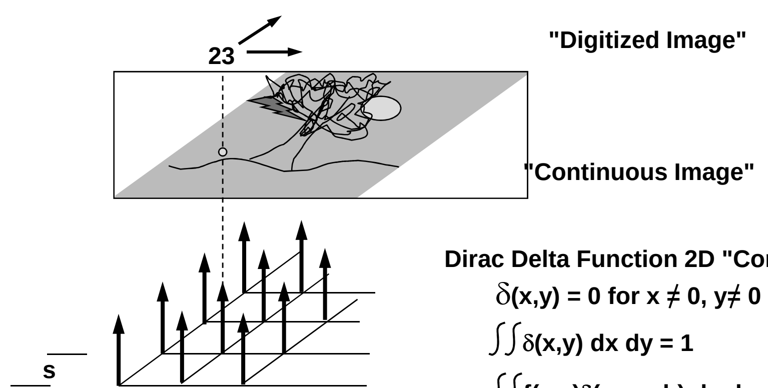
- IF: the size of the smallest structure to be preserved is d
- THEN: the sampling interval must be smaller than d/2

- Can be shown to be true mathematically
- Repetitive structure has a certain frequency
  - To preserve structure must sample at twice the frequency
  - Holds for images, audio CDs, digital television….
- Leads naturally to Fourier Analysis (optional)

3D Computer Vision
and Video Computing

Sampling

- **Rough Idea: Ideal Case**



Dirac Delta Function 2D "Comb"

$\delta(x,y) = 0$ for $x \neq 0, y \neq 0$

$\int\int \delta(x,y)\, dx\, dy = 1$

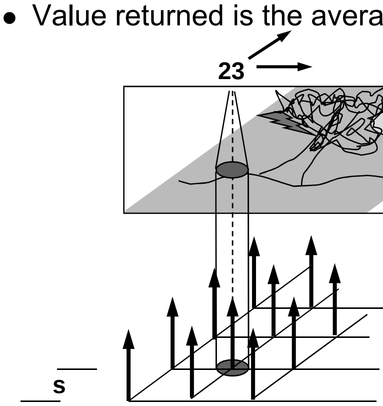
$\int\int f(x,y)\delta(x-a,y-b)\, dx\, dy = f(a,b)$

$\delta(x-ns,y-ns)$ for n = 1….32 (e.g.)

3D Computer Vision
and Video Computing

Sampling

- Rough Idea: Actual Case
  - Can't realize an ideal point function in real equipment
  - "Delta function" equivalent has an area
  - Value returned is the average over this area

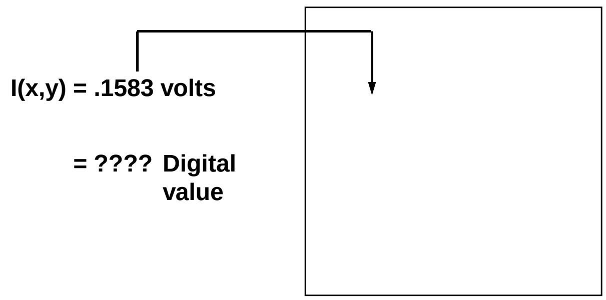


3D Computer Vision
and Video Computing

Mixed Pixel Problem

Signal Quantization

- Goal: determine a mapping from a continuous signal (e.g. analog video signal) to one of K discrete (digital) levels.



Quantization

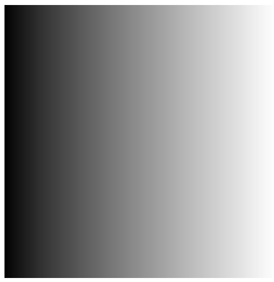
- I(x,y) = continuous signal: $0 \leq I \leq M$
- Want to quantize to K values 0,1,....K-1
- K usually chosen to be a power of 2:

K	#Levels	#Bits
2	2	1
4	4	2
8	8	3
16	16	4
32	32	5
64	64	6
128	128	7
256	256	8

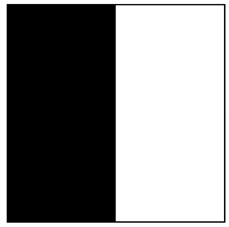
- Mapping from input signal to output signal is to be determined.
- Several types of mappings: uniform, logarithmic, etc.

Choice of K

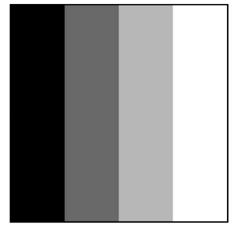
Original



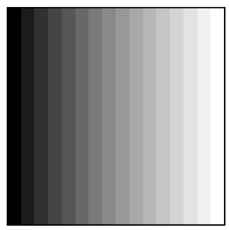
Linear Ramp



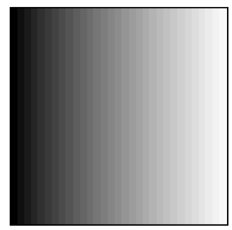
K=2



K=4

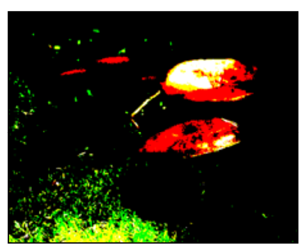


K=16



K=32

Choice of K



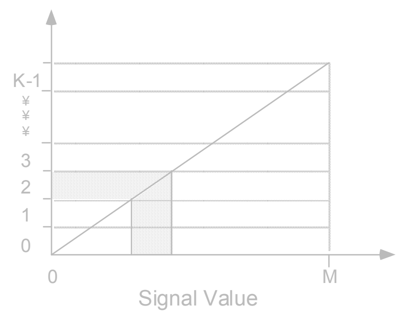
K=2 (each color)

K=4 (each color)

3D Computer Vision
and Video Computing

Choice of Function: Uniform

- Uniform quantization divides the signal range [0-M] into K equal-sized intervals.
- The integers 0,...K-1 are assigned to these intervals.
- All signal values within an interval are represented by the associated integer value.
- Defines a mapping:



3D Computer Vision
and Video Computing

Logarithmic Quantization

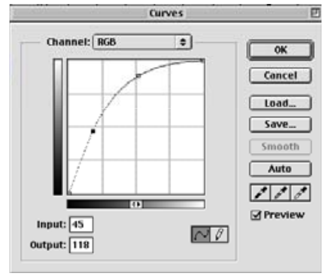
- Signal is log I(x,y).
- Effect is:

- Detail enhanced in the low signal values at expense of detail in high signal values.

Logarithmic Quantization

Original

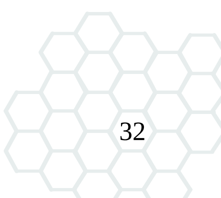
Quantization Curve



Logarithmic Quantization

3D Computer Vision
and Video Computing

Tesselation Patterns



Hexagonal

Triangular

Rectangular

Typical

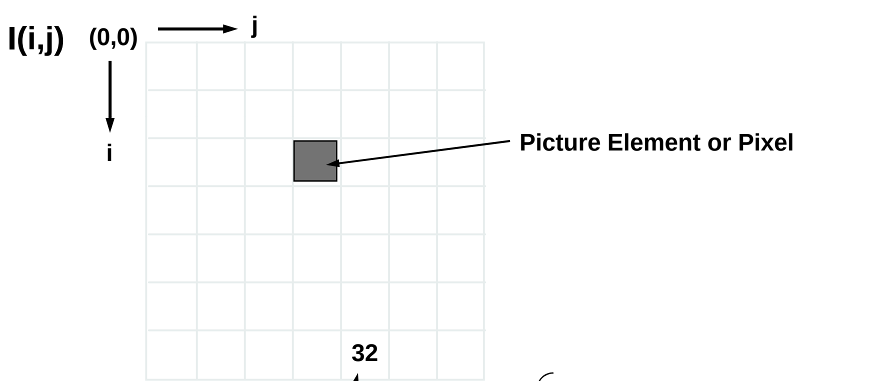
3D Computer Vision
and Video Computing

Lecture Outline

- Image Formation Basic Steps
- Geometry
  - Pinhole camera model & Thin lens model
  - Perspective projection & Fundamental equation
- Radiometry
- Photometry
  - Color, human vision, & digital imaging
- Digitalization
  - Sampling, quantization & tessellations
- More on Digital Images
  - Neighbors, connectedness & distances

3D Computer Vision
and Video Computing

Digital Geometry

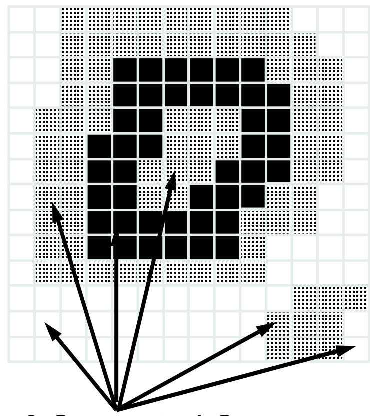


Pixel value I(I,j) =
- 0,1 Binary Image
- 0 - K-1 Gray Scale Image
- Vector: Multispectral Image

- Neighborhood
- Connectedness
- Distance Metrics

Connected Components

- Binary image with multiple 'objects'
- Separate 'objects' must be labeled individually

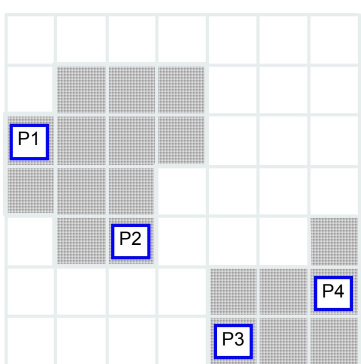


6 Connected Components

3D Computer Vision
and Video Computing

Finding Connected Components

- Two points in an image are 'connected' if a path can be found for which the value of the image function is the same all along the path.



P_1 connected to P_2

P_3 connected to P_4

P_1 not connected to P_3 or P_4

P_2 not connected to P_3 or P_4

P_3 not connected to P_1 or P_2

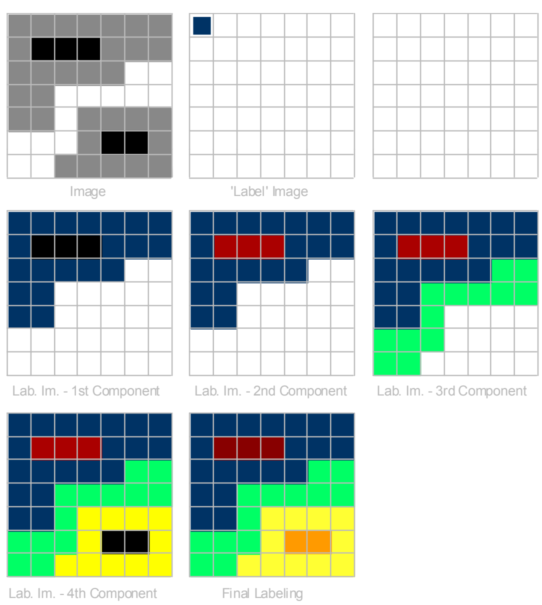
P_4 not connected to P_1 or P_2

Algorithm

- Pick any pixel in the image and assign it a label
- Assign same label to any neighbor pixel with the same value of the image function
- Continue labeling neighbors until no neighbors can be assigned this label
- Choose another label and another pixel not already labeled and continue
- If no more unlabeled image points, stop.

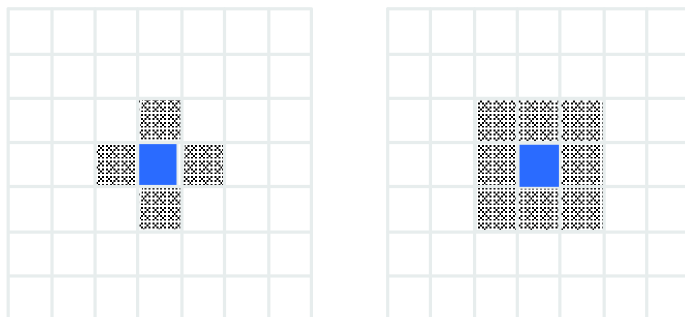
Who's my neighbor?

Example



Neighbor

- Consider the definition of the term 'neighbor'
- Two common definitions:



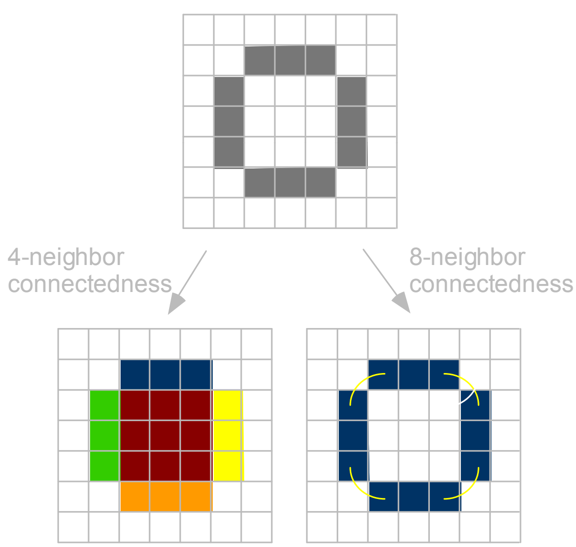
Four Neighbor

Eight Neighbor

- Consider what happens with a closed curve.
- One would expect a closed curve to partition the plane into two connected regions.

3D Computer Vision
and Video Computing

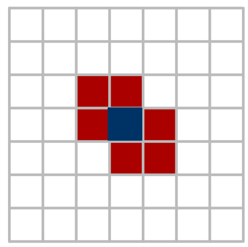
Alternate Neighborhood Definitions



Neither neighborhood definition satisfactory!

Possible Solutions

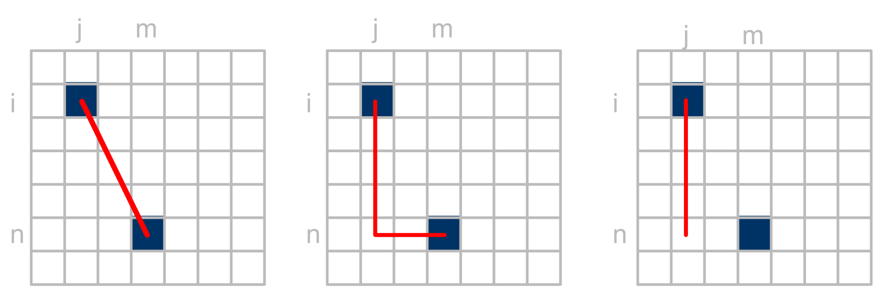
- Use 4-neighborhood for object and 8-neighborhood for background
  - requires a-priori knowledge about which pixels are object and which are background
- Use a six-connected neighborhood:



3D Computer Vision
and Video Computing

Digital Distances

- Alternate distance metrics for digital images



Euclidean Distance

$= \sqrt{(i-n)^2 + (j-m)^2}$

City Block Distance

$= |i-n| + |j-m|$

Chessboard Distance

$= \max[\ |i-n|, |j-m|\ ]$

Next:
Feature Extraction

- Homework #1 online, Due Feb 22 before class