Geospatial Vision and Visualization

Xin Chen, Ph.D.

Lecture 2

Today's Topics

- Geo-referenced street level imagery
  - Street views
  - Image formation
  - Basic Image Processing

Geo-referenced Street Level Imagery

Why street level imagery?

- To make maps
  - What are maps: locations and attributes
  - Conventional data sources
    - Ground survey and positioning
    - Remote sensing
    - Census and sampling
  - Street level imagery
    - Navigation attributes: signs, road attributes, POI, etc.

Why street level imagery? (cont.)

- **Street level photorealistic visualization**
  - NAVTEQ Street View (USA) http://www.here.net
  - Yahoo Street View (USA) http://maps.yahoo.com
  - Microsoft Bing Map Street Side (USA) http://www.bing.com/maps
  - Google Street View (USA) http://maps.google.com/help/maps/streetview
  - Mapquest (USA) http://www.mapquest.com
  - Mapjack (China) http://www.mapjack.com
  - Everyscape (USA) http://www.everyscape.com
  - Daum's Road View (South Korea) http://local.daum.net/map
  - Wuhan City Map (China) http://map.wuhan.net.cn/wh/index/index.aspx
  - City8 (China) http://www.city8.com
  - Seety (UK) http://www.seety.co.uk/
  - NORC (Romania) http://www.norc.ro/
  - Driveme.in (India) http://www.driveme.in/
  - Cyclomedia (Dutch) http://www.cyclomedia.com
  - Location View (Japan) https://www.locaview.com
  - Earthmine (USA) http://www.earthmine.com/index
  - Soso (China) http://map.soso.com

Demos

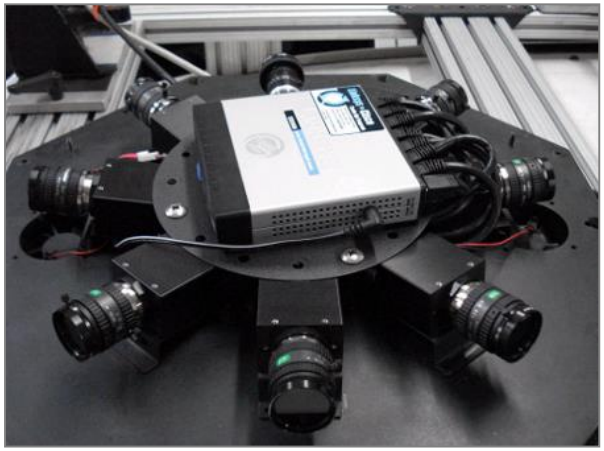
- Maps.google.com
- Bing.com/maps/
- Map.qq.com
- Hongkong.edushi.com

Why street level imagery? (cont.)

- For fun – community based photos
  - Flickr http://www.flickr.com/map/
  - Photosynth (part of Bing map)
    - Snavely, et al. SIGGRAPH 2006. 3D Photo tourism
    - Agarwal, et al. ICCV 2009. Building Rome in a day
    - Furukawa, et al. Reconstructing building interiors from images
  - PhotoSphere (part of Google Maps)

Imagery Data Collection

- **Cameras**
  - **High resolution video (> 1 MP per image)**
  - **Panoramic**



Imagery Data Collection (cont.)

- **Positioning – INS (Inertial Navigation System)**
  - **GPS**
  - **IMU (Inertial Measurement Unit)**
  - **DMI (Distance Measurement Indicator)**
  - **Software**

How does GPS work?
Beyond GPS

Imagery Data Collection Platforms

Privacy Concerns and Solutions

- GDPR (General Data Protection Regulation) – May 25, 2018
- Faces, license plates, houses, etc.
- Automatic blur
- User report
- Google
  - Large-scale Privacy Protection in Google Street View. In ICCV 2008
  - US: no privacy safeguard in place; EU: some countries have such legislations
  - Blur in all countries

Novel Street View

- MIT Infrared street view
- Microsoft Street Slide
- Route panorama
- Non-vehicle-accessible places (pedestrian areas, plants, etc.) - Behind the scenes with Trekker

HW1 – Automatic Lens Smear Detection

- **Input: a sequence of street view images**
- **Output: a mask of the smear on the lens**
- **Submit by email:**
  - **1. PPT slides of your approach and both intermediate and final results**
  - **2. Source code and executable/scripts with appropriate documentation**
- **Resources**
  - **"Removing Image Artifacts Due to Dirty Camera Lenses and Thin Occluders,"**

    J. Gu, R. Ramamoorthi, P.N. Belhumeur and S.K. Nayar,

    ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)

Image and Camera

Image Formation

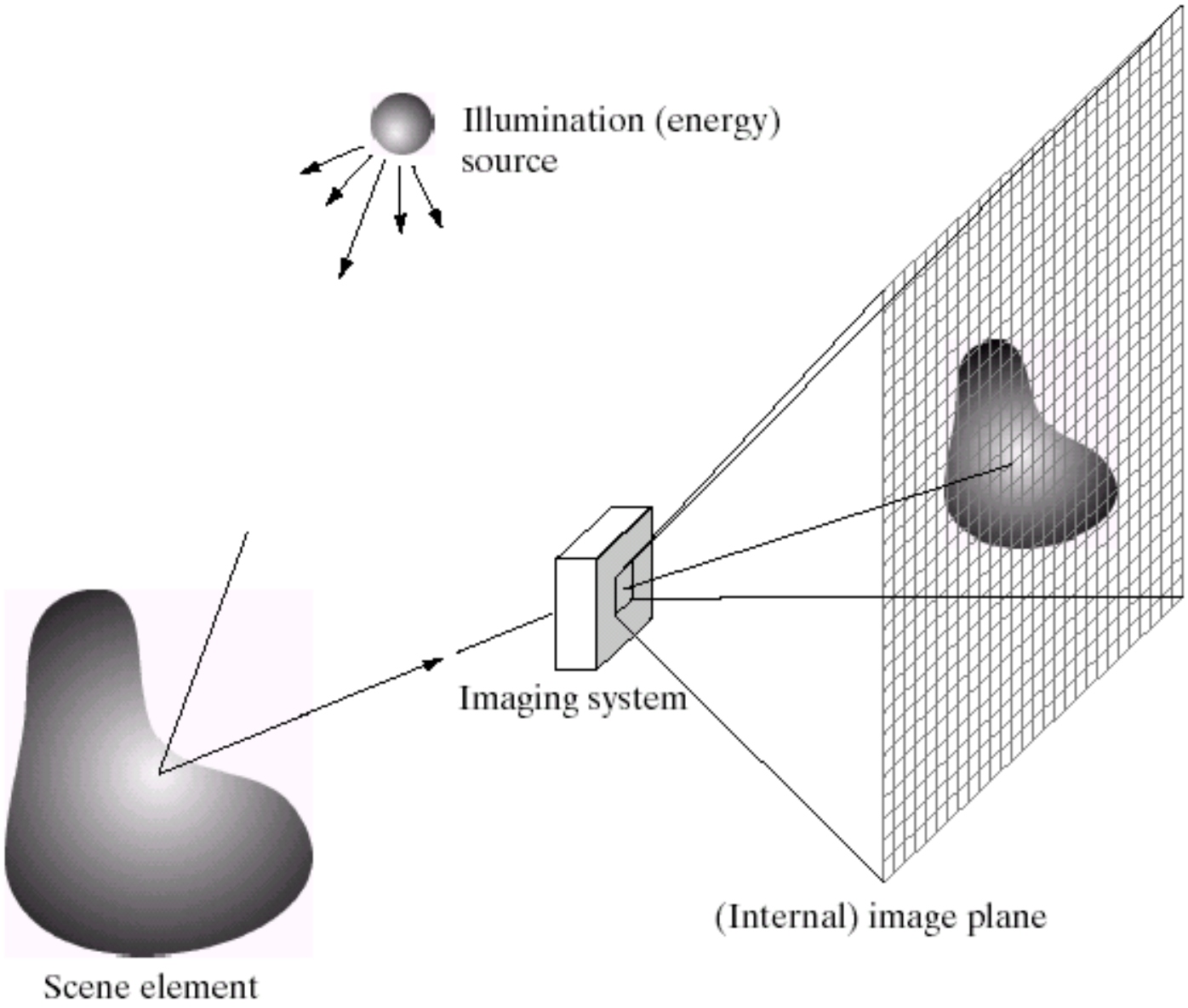
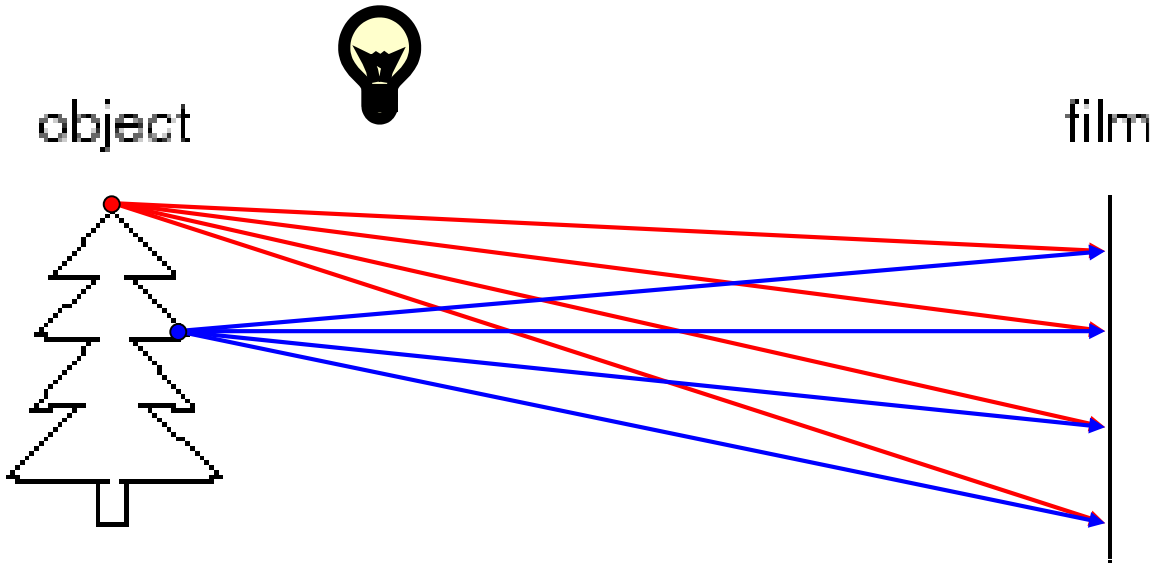
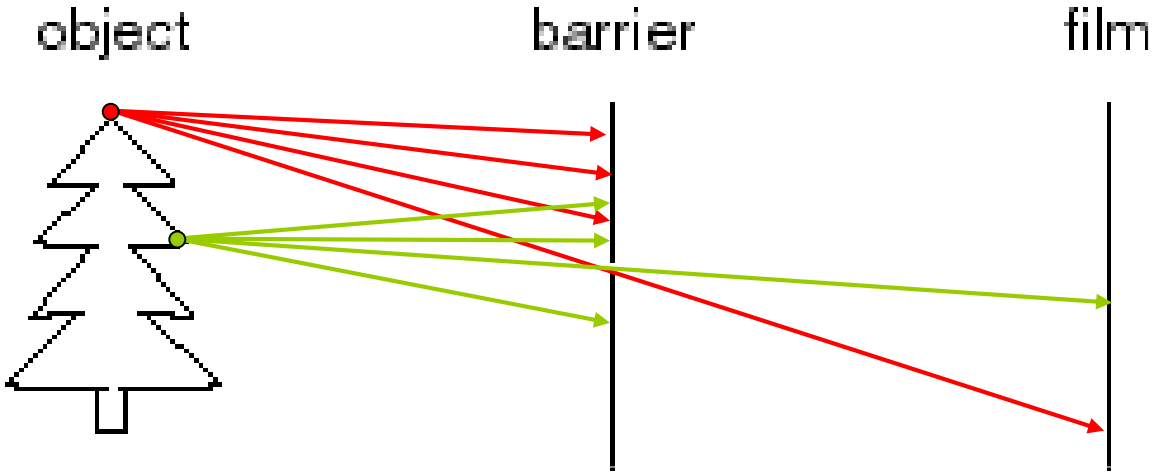


Image Formation (cont.)

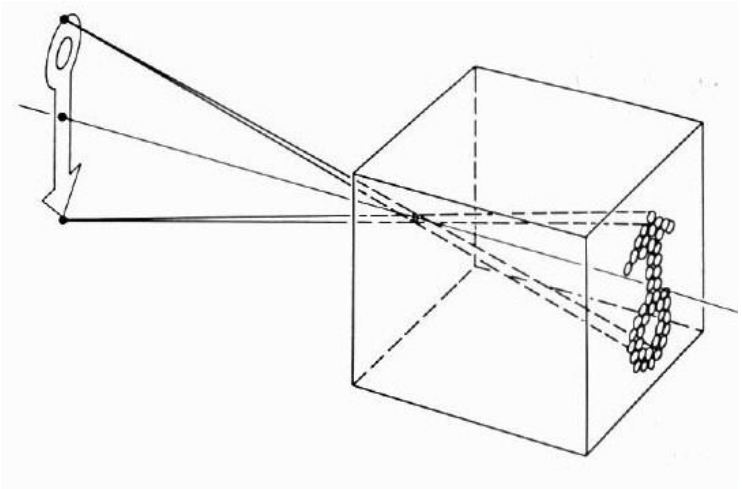
- **Does this form an image? Why?**



Pinhole Camera

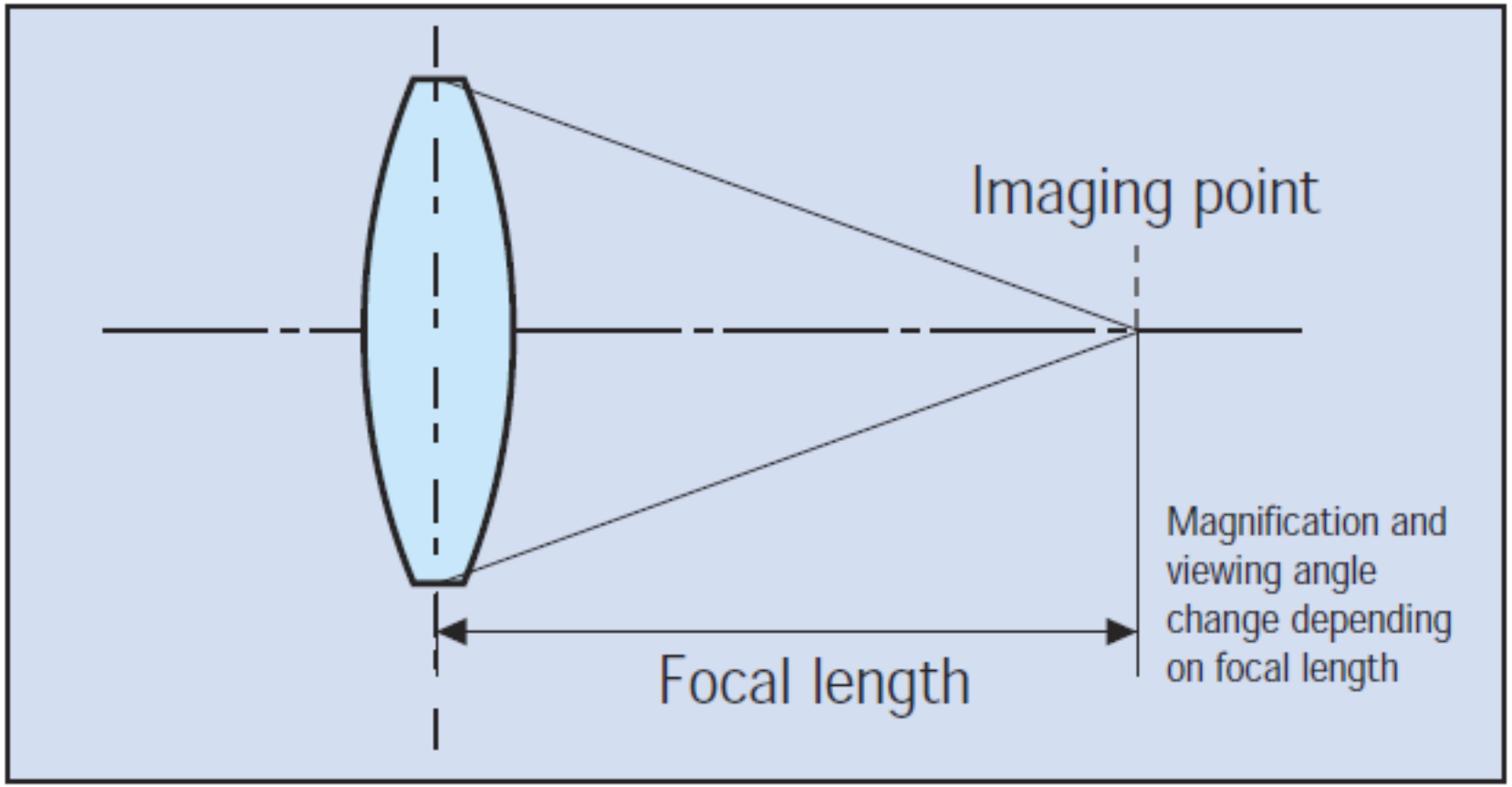


- Add a barrier to block off most of the rays
- Aperture
- Camera obscura



Pinhole Camera Images

What is focal length?



Aperture Size

Sharpest image is obtained when:

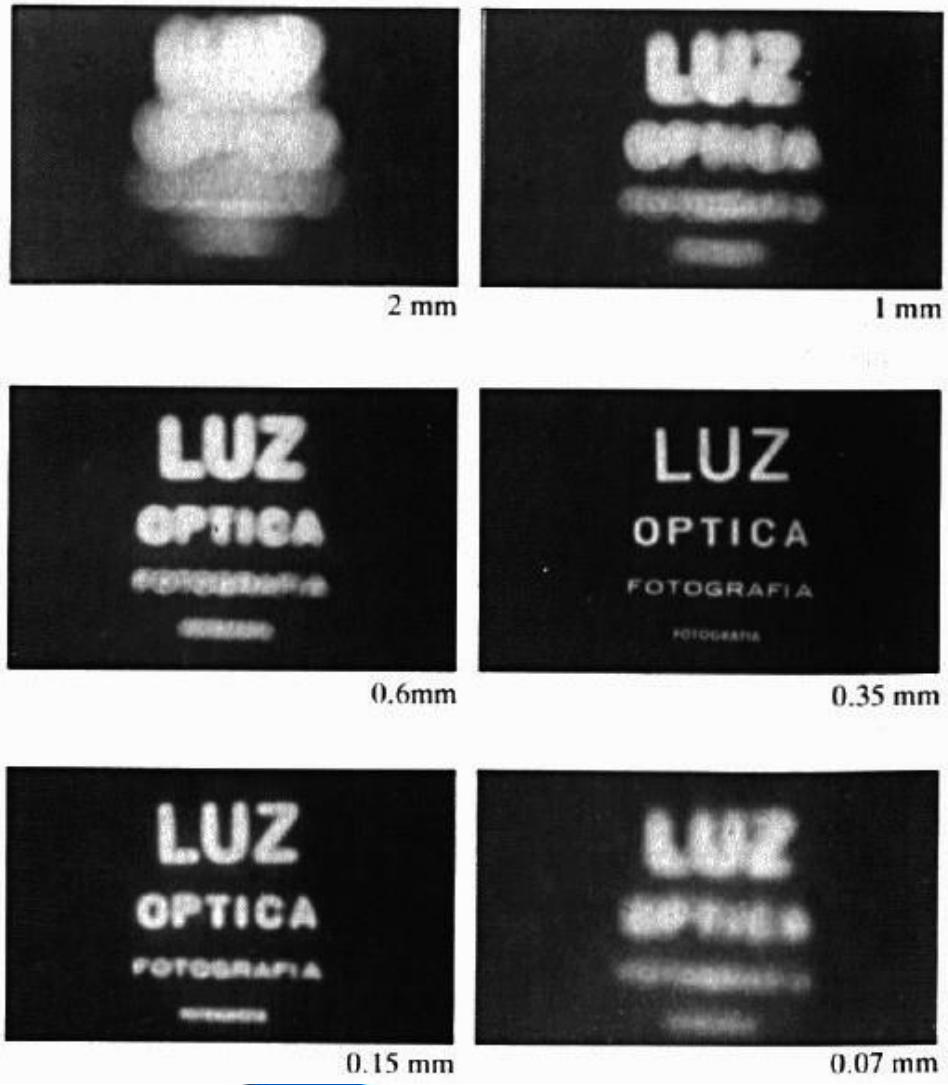
$$d = 2\sqrt{f\lambda}$$

d is diameter,
f is distance from hole to film
λ is the wavelength of light,
all given in meters.

Example: If *f = 50mm,*

λ = 600nm (red),

d = 0.36mm



Drawback of Pinhole Cameras

- **Low incoming light -> Long exposure -> Tripod**

KODAK Film or Paper	Bright Sun	Cloudy Bright
TRI-X Pan	1 or 2 seconds	4 to 8 seconds
T-MAX 100 Film	2 to 4 seconds	8 to 16 seconds
KODABROMIDE Paper, F2	2 minutes	8 minutes

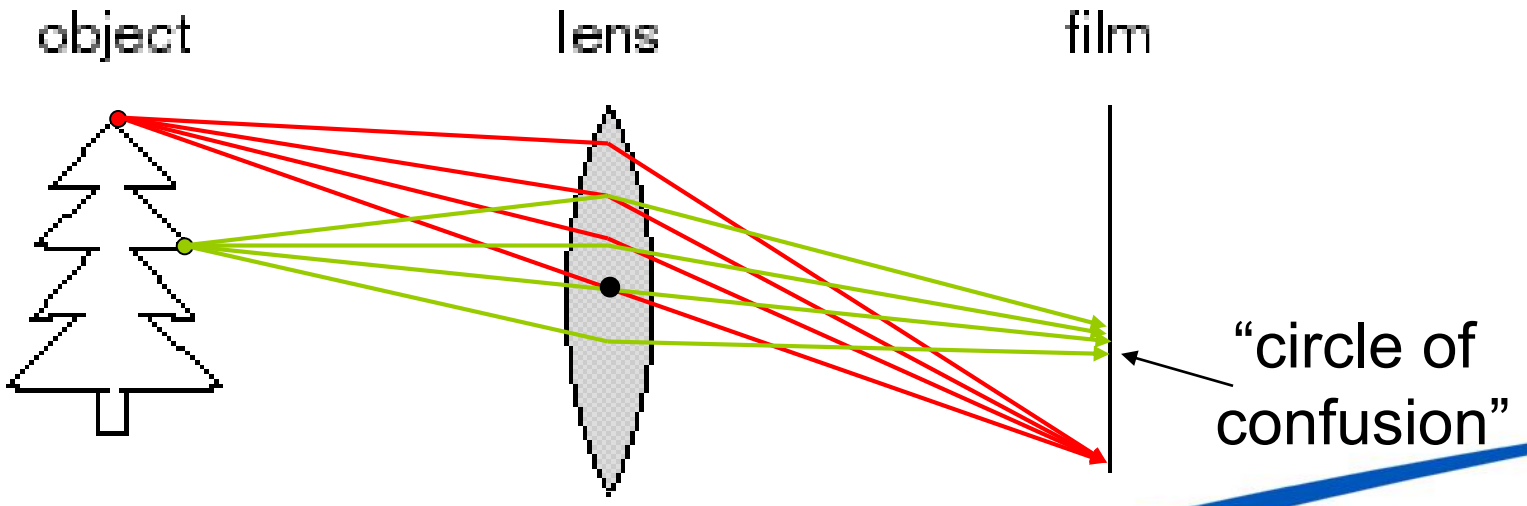
Drawback of Pinhole Cameras (cont.)

Pinhole Camera Recap

- **Pinhole size (aperture) must be "very small" to obtain a clear image.**
- **However, as pinhole size is decreased, less light is received by image plane.**
- **If pinhole size is comparable to wavelength of incoming light, *DIFFRACTION* effects blur the image!**

Optical Lenses

- *Intuitively, an optical system can be regarded as a device that aims at producing the same image obtained by a pinhole aperture, but by means of* ***a much larger aperture*** *and* ***a shorter exposure time****. Moreover, an optical system enhances the light gathering power.*

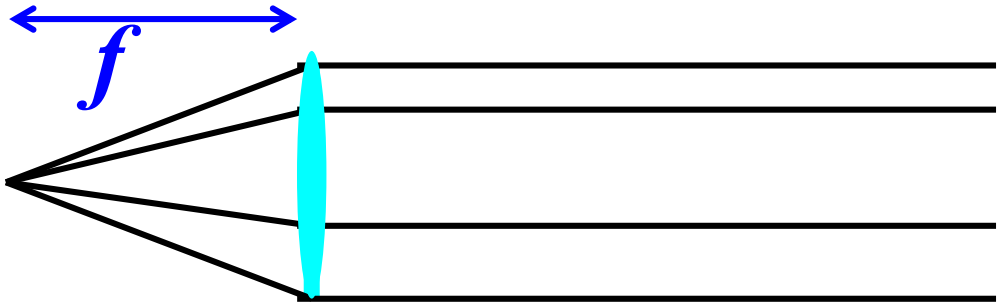


Optical Lenses (cont.)

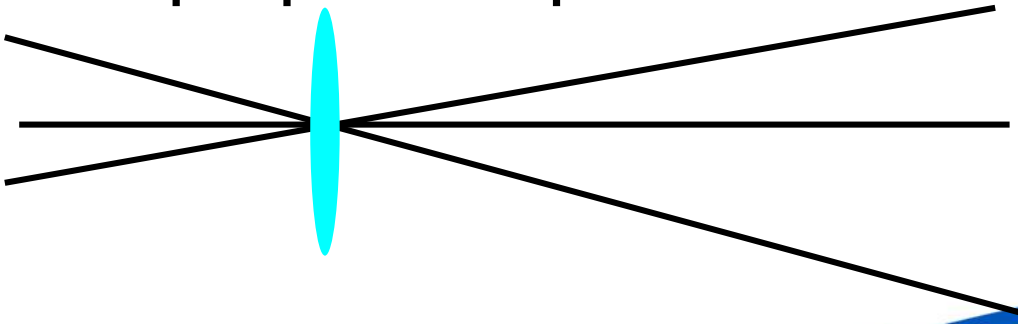
- A lens focuses light onto the film
  - There is a specific distance at which objects are “in focus”
  - Other points project to a “circle of confusion” in the image
  - Changing the shape of the lens changes this distance

Thin lenses

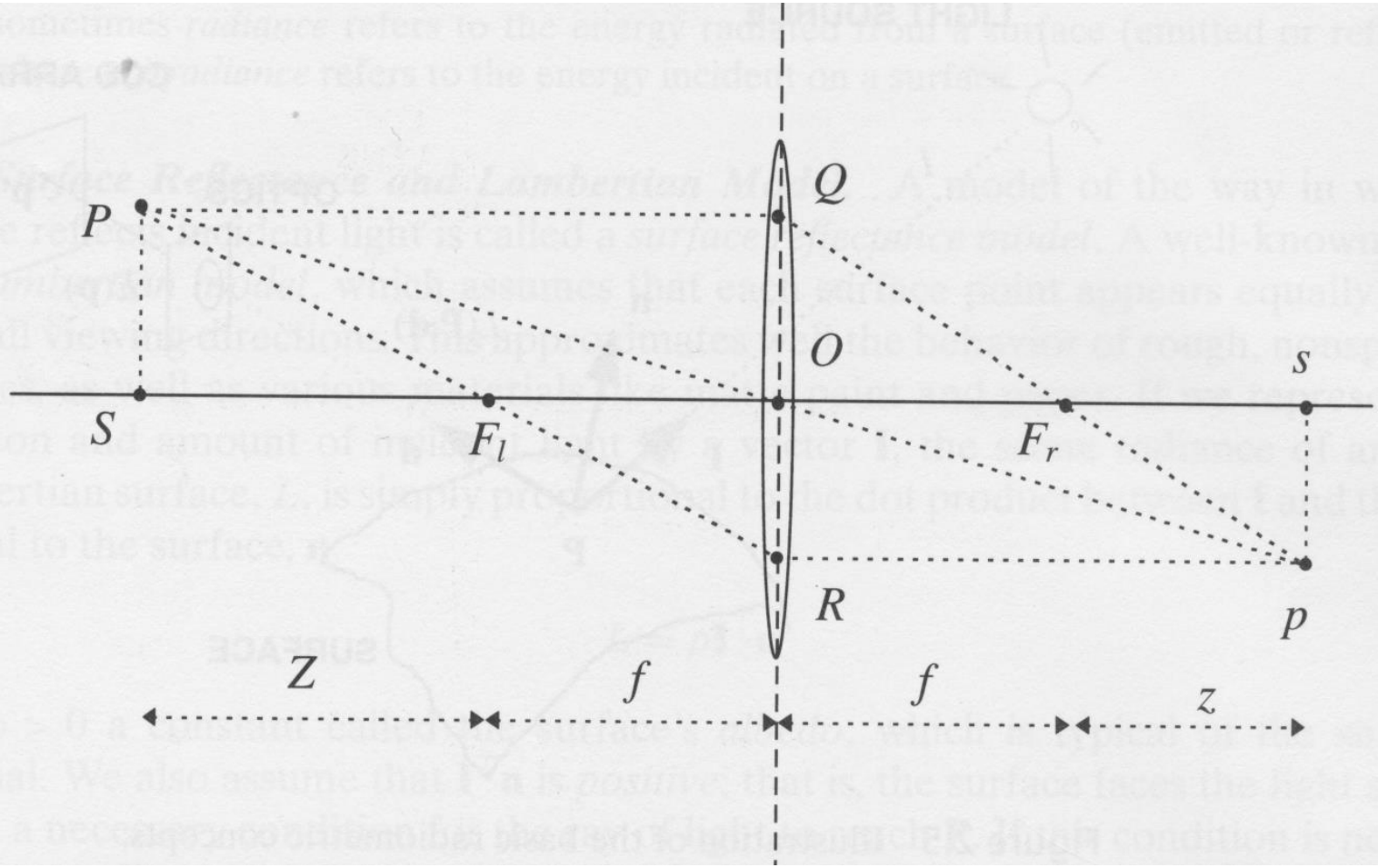
- Simplification of geometrical optics for well-behaved lenses
- All parallel rays converge to one point on a plane located at the focal length *f*



- All rays going through the center are not deviated
  - Hence same perspective as pinhole



Fundamental Equation of Thin Lenses

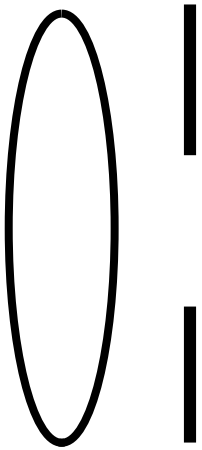


$$1/Ź + 1/ź = 1/f$$

Where $Ź = Z + f$, $ź = z + f$

Film Camera

scene

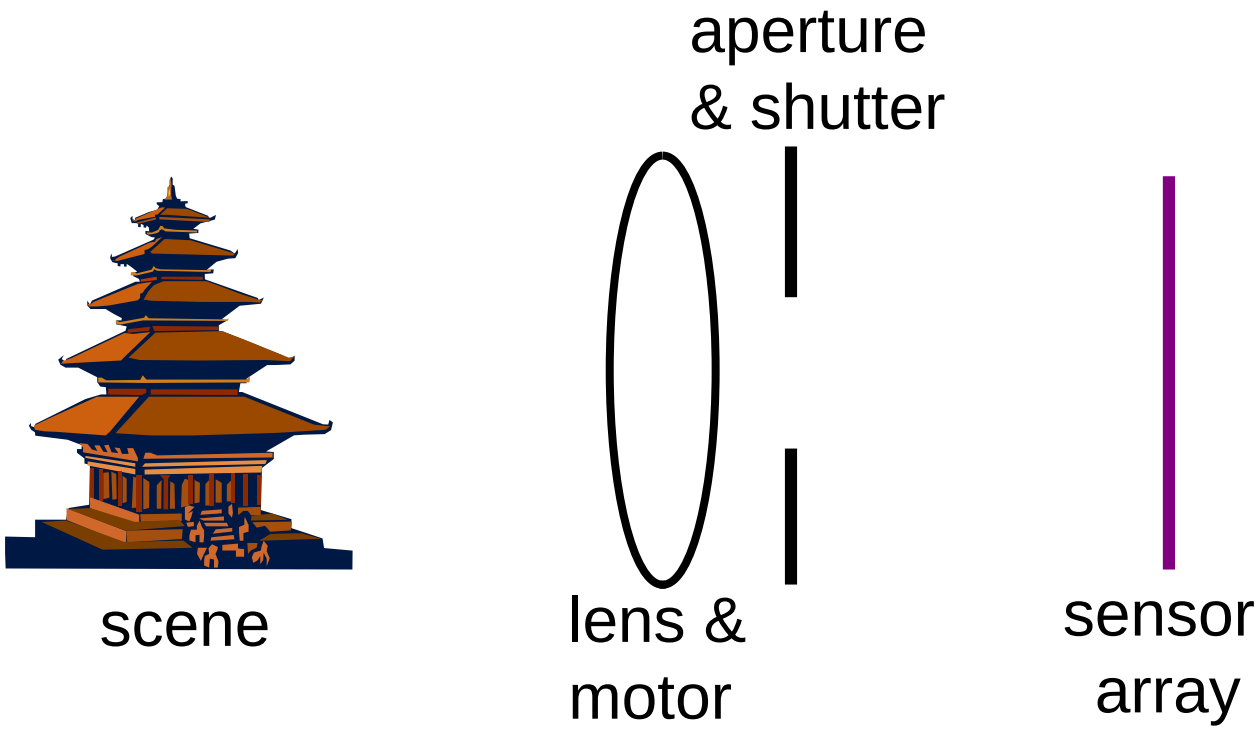


aperture
& shutter

lens &
motor

film

Digital Camera



- **A digital camera replaces film with a sensor array**
- **Each cell in the array is a light-sensitive diode that converts photons to electrons**

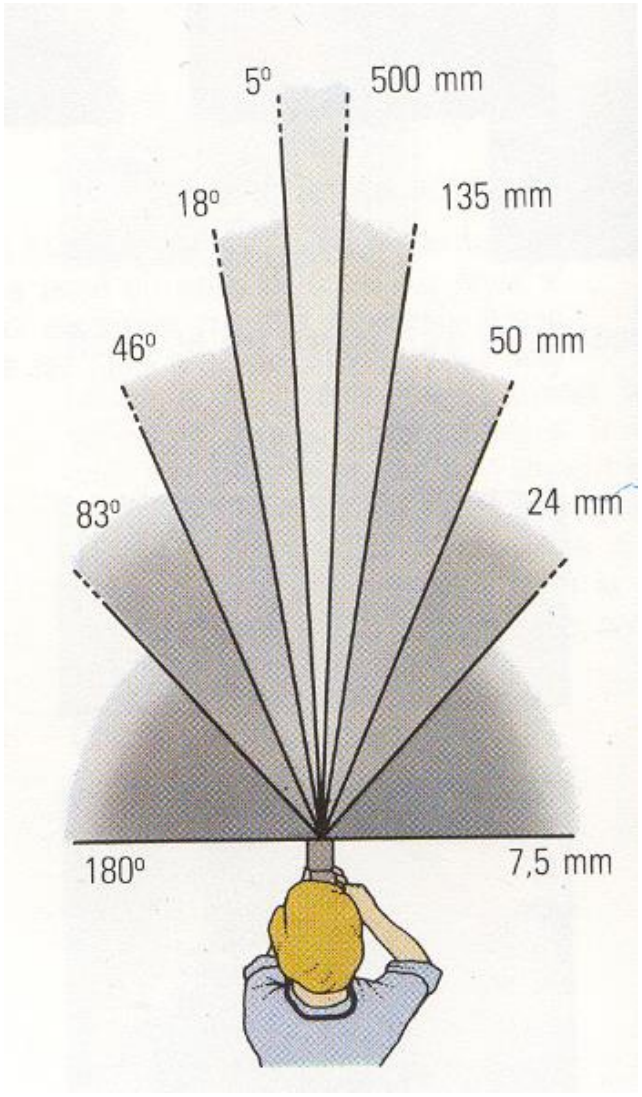
Field of View (FoV)

- **An angular measure of the portion of 3D space actually seen by the camera**

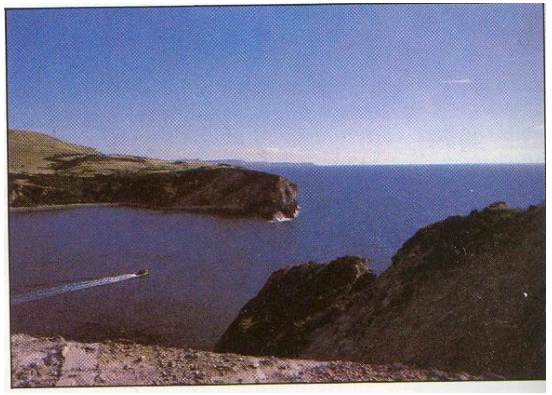
$$\tan \omega = d / 2f$$

where d is the effective diameter of the lenses

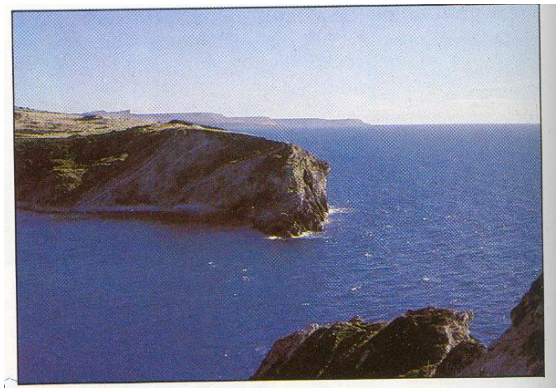
Field of View (FoV) vs Focal Length



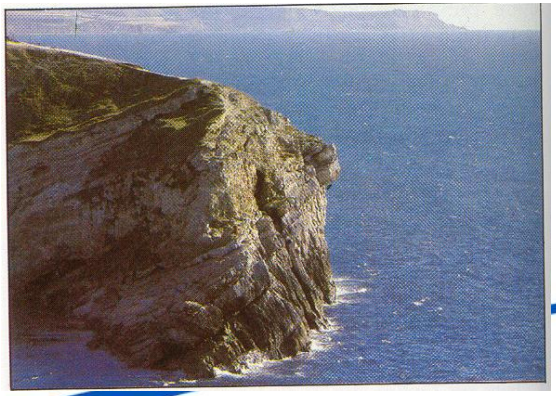
24mm



50mm



135mm



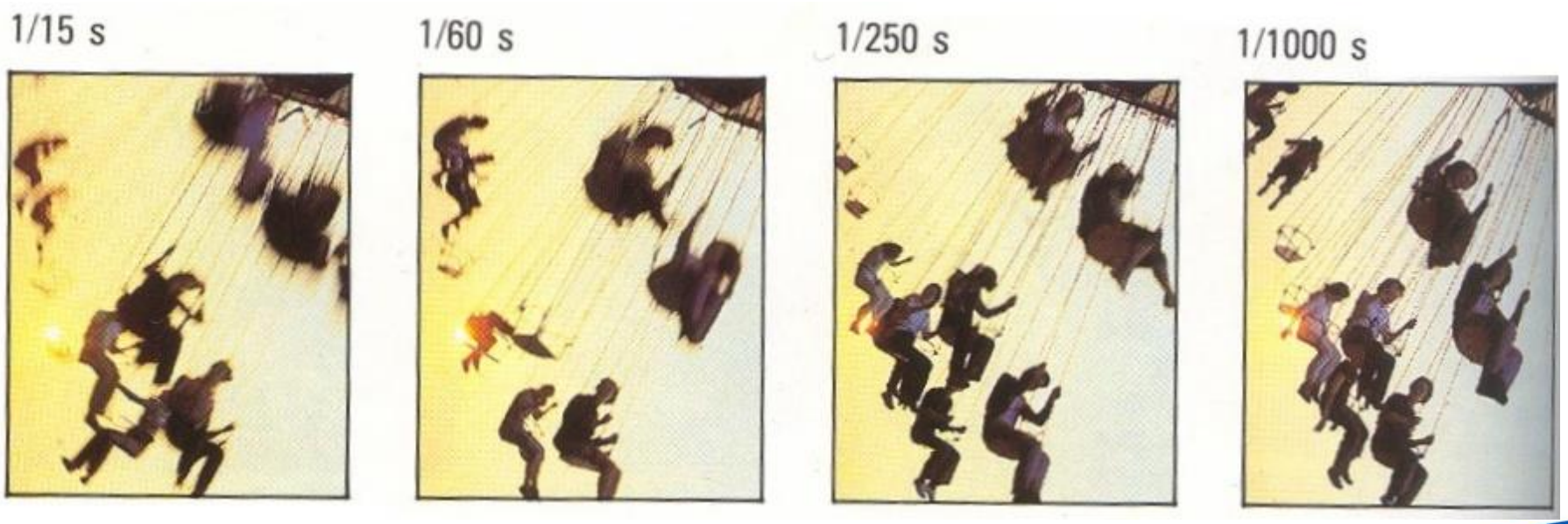
Exposure (Filling a Bucket with Water)

- **Shutter speed (the length of time you pour water)**
- **F-stops (the size of the stream of water)**
- **Film speed (the size of the bucket) - fixed**

http://www.wonkeedonkeetools.co.uk/water-pressure-gauges/how-do-you-use-a-water-pressure-gauge/

Shutter Speed

- Doubling and halving the amount of light reaching the film in the following sequence:
  - 8 seconds 4 seconds 2 seconds 1 second 1/2 second 1/4 1/8 1/15 1/30 1/60 1/125 1/250 1/500 1/1000



F-stop

- Doubling and halving the amount of light reaching the film in the following sequence:
  - 1.4 2.0 2.8 4 5.6 8 11 16 22

F-stop (cont.)

- **Quantitative measure of lens speed**
  - **Exposure setting's aperture value**
  - **Overall brightness of the lens**

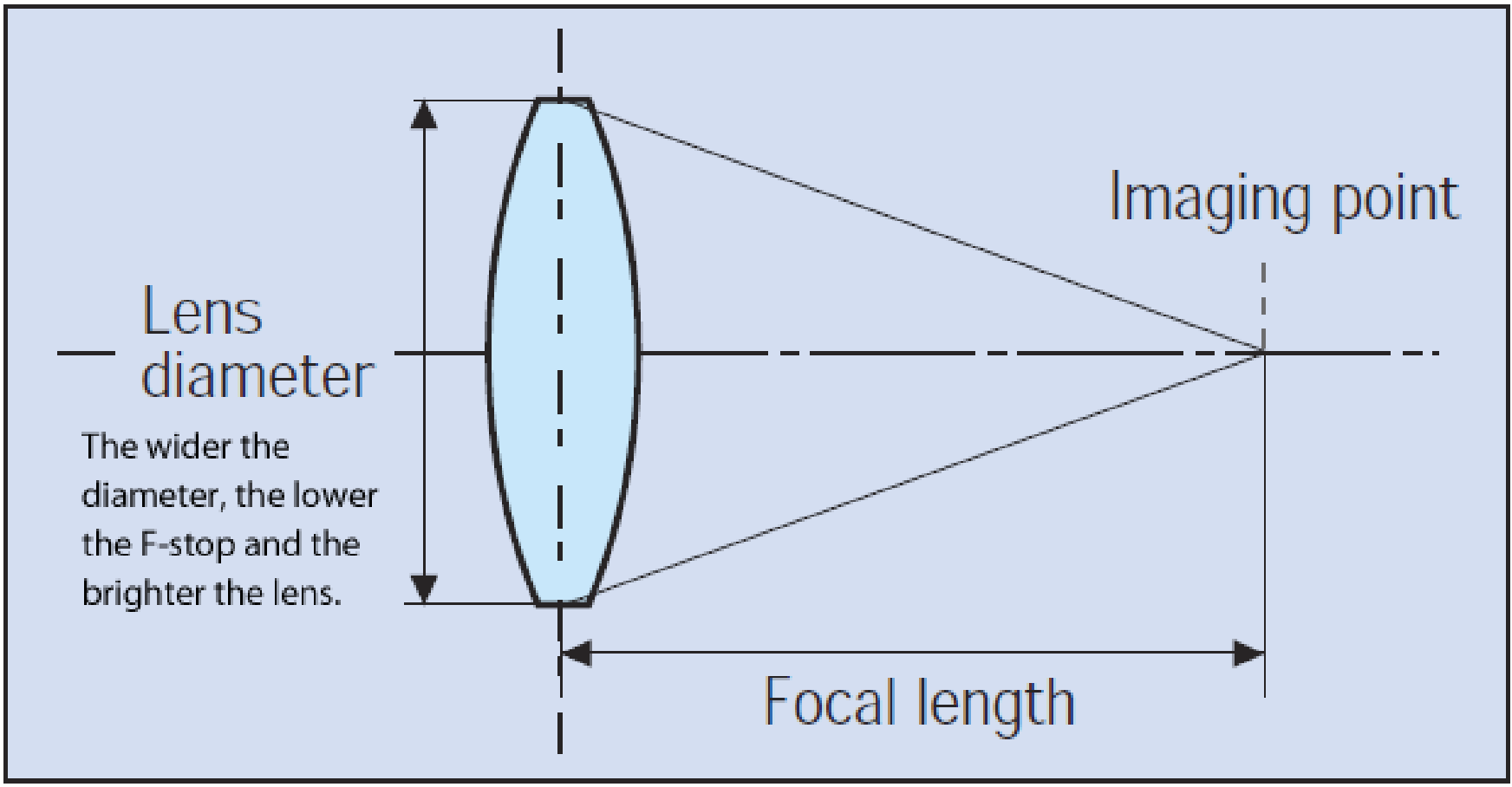
Canon EF-S 60mm f/2.8

Canon EF 100mm f/2.8

Canon EF 180mm f/3.5

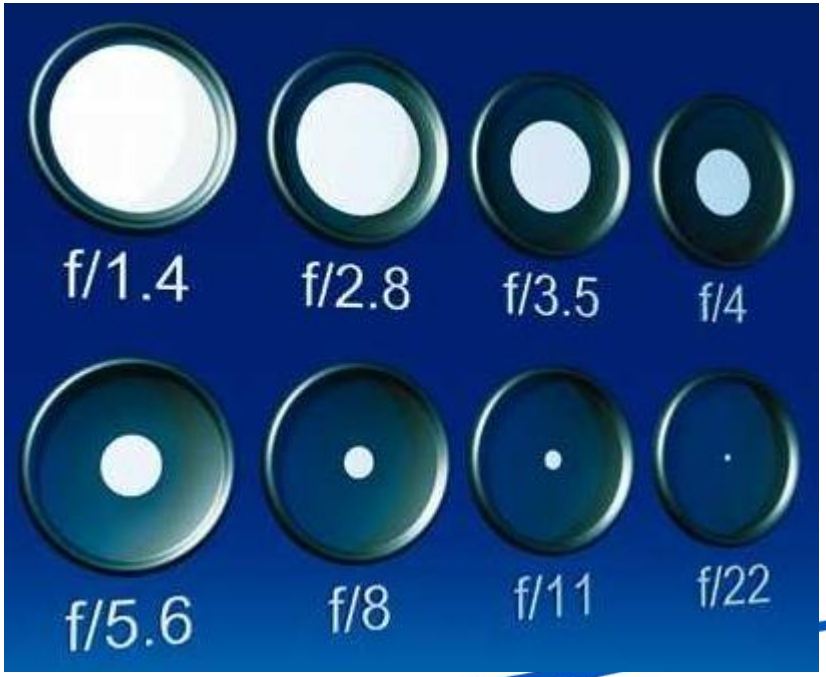
F-stop and Focal Length

$$f\text{-stop } \# = \frac{f}{d}$$



Aperture

- Aperture is the diameter of the lens opening, usually specified by f-stop, f/D, a fraction of the focal length.
  - f/2.0 on a 50mm means that the aperture is 25mm
  - f/2.0 on a 100mm means that the aperture is 50mm
- Lower f-stop, more light (larger lens opening)
- Higher f-stop, less light (smaller lens opening)



For a 50mm lens:

f/stop	Diameter of aperture (mm)	Area of Aperture (sq. mm)
f/1.0	50.0	1,963
f/1.4	35.7	1,002
f/2.0	25.0	491
f/2.8	17.9	250
f/4	12.5	123
f/5.6	8.9	63
f/8	6.3	31
f/11	4.5	16
f/16	3.1	8
f/22	2.3	4
(As shown on lens)	(50mm divided by f/stop)	(pi X the radius squared)

Why not just use area of aperture?

- *"I shot this with my 50mm at 1/125th of a second and an aperture area of 63 square millimeters“ vs. “I shot this at 1/125th at f/5.6"*
- *Focal length + Area of Aperture vs. f-stop*

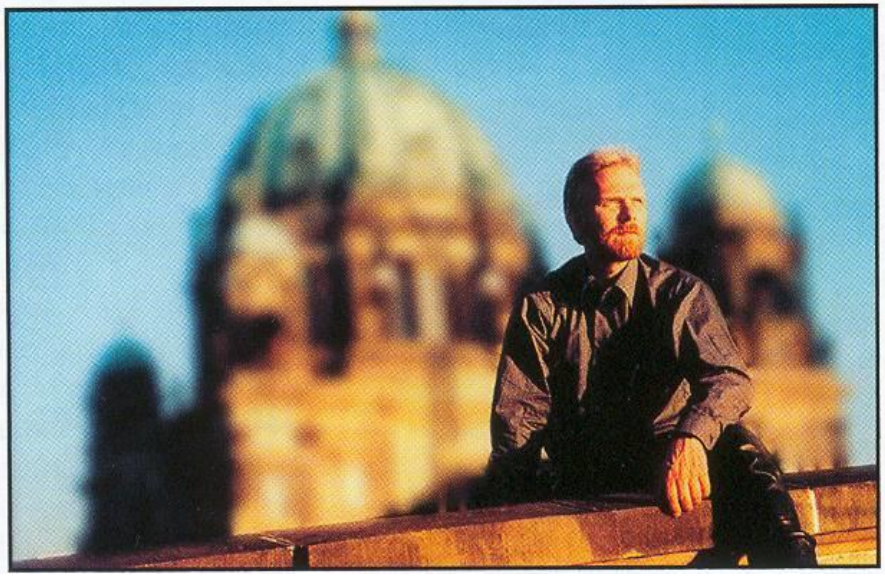
Is f/8 the same brightness of illumination on the film/sensor regardless of focal length of the lens?

- **YES**

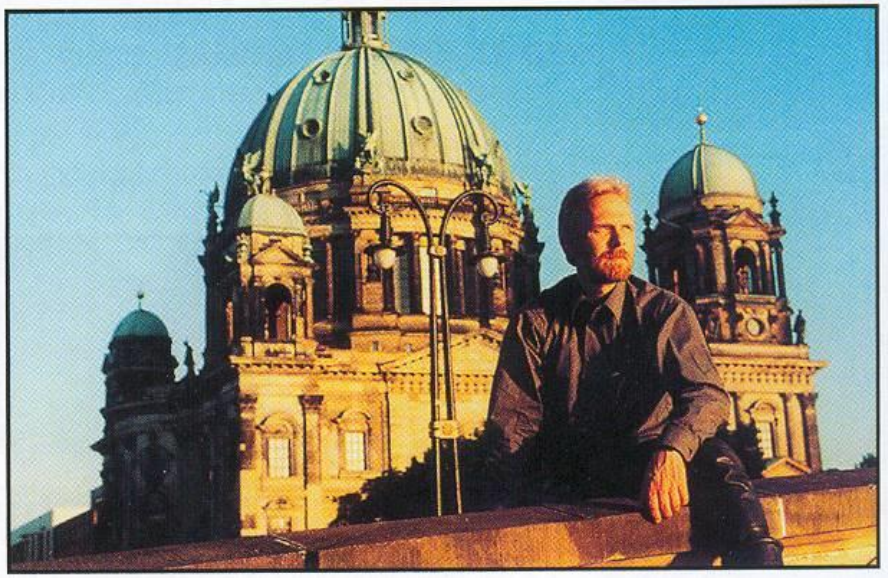
Depth of Field

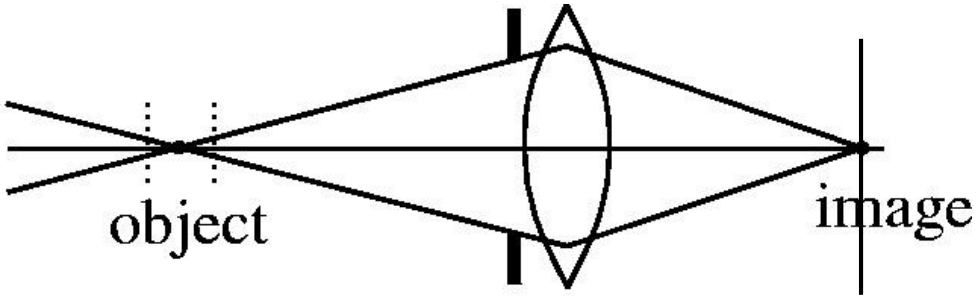
- The range of distances within which all points are imaged on or close to the image plane and therefore acceptably in focus.

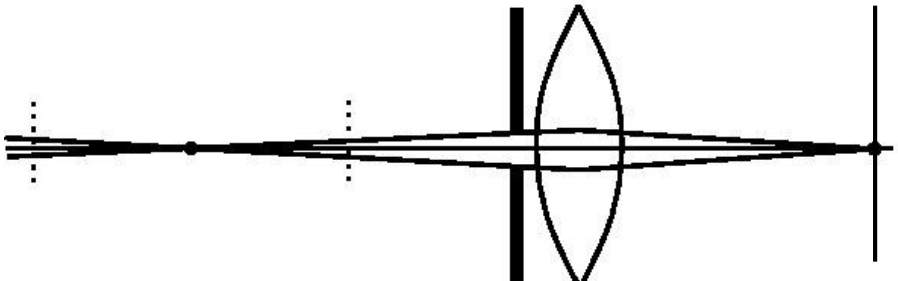
Large aperture opening



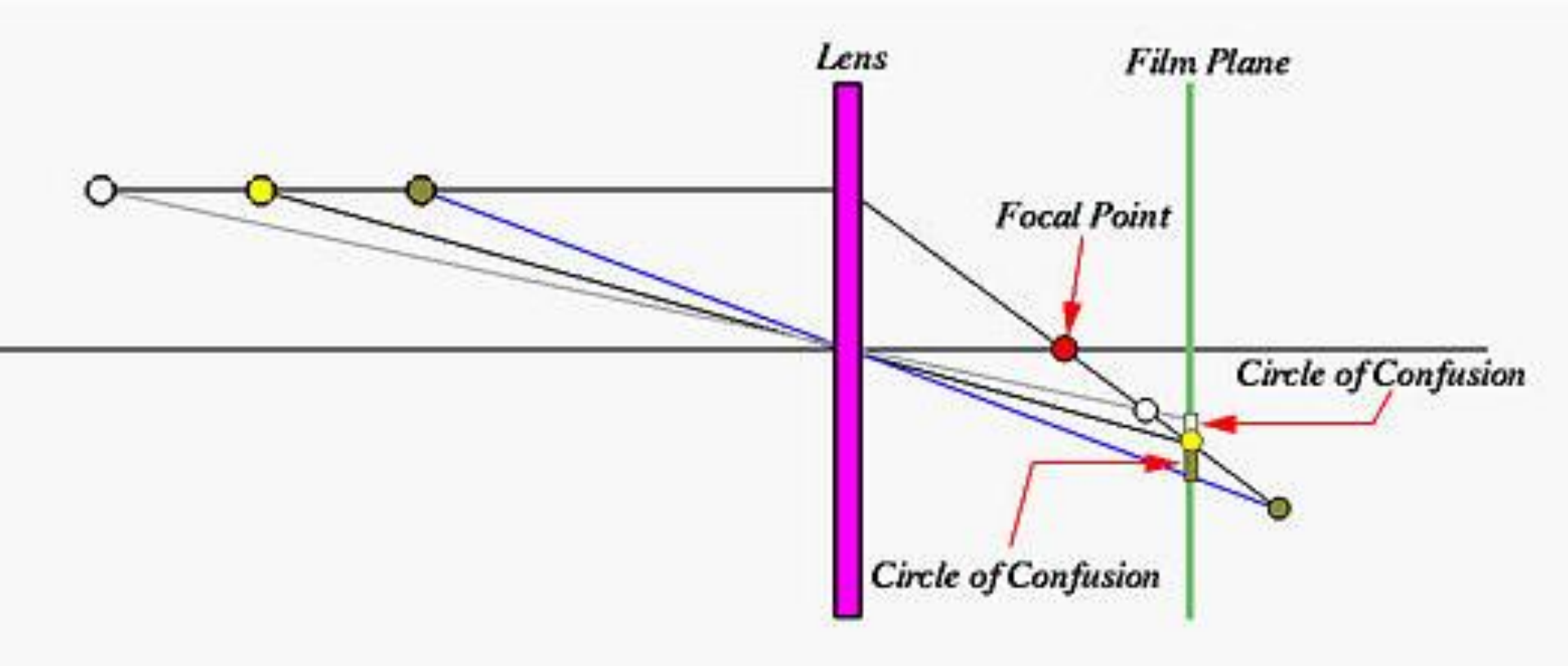
Small aperture opening







Circle of Confusion

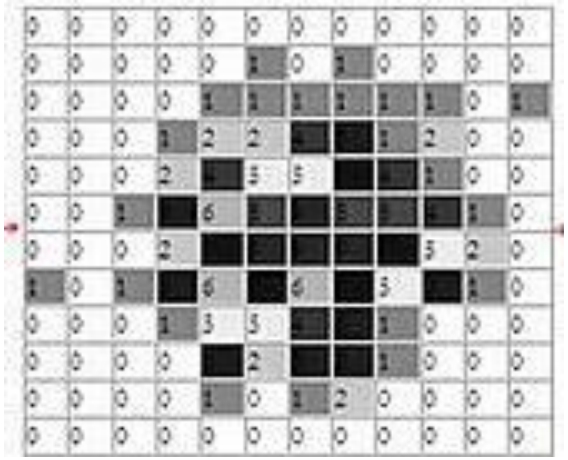


Exposure recap

- Two main parameters:
  - Aperture/focal length (in f stop)
  - Shutter speed (in fraction of a second)

Optical Image

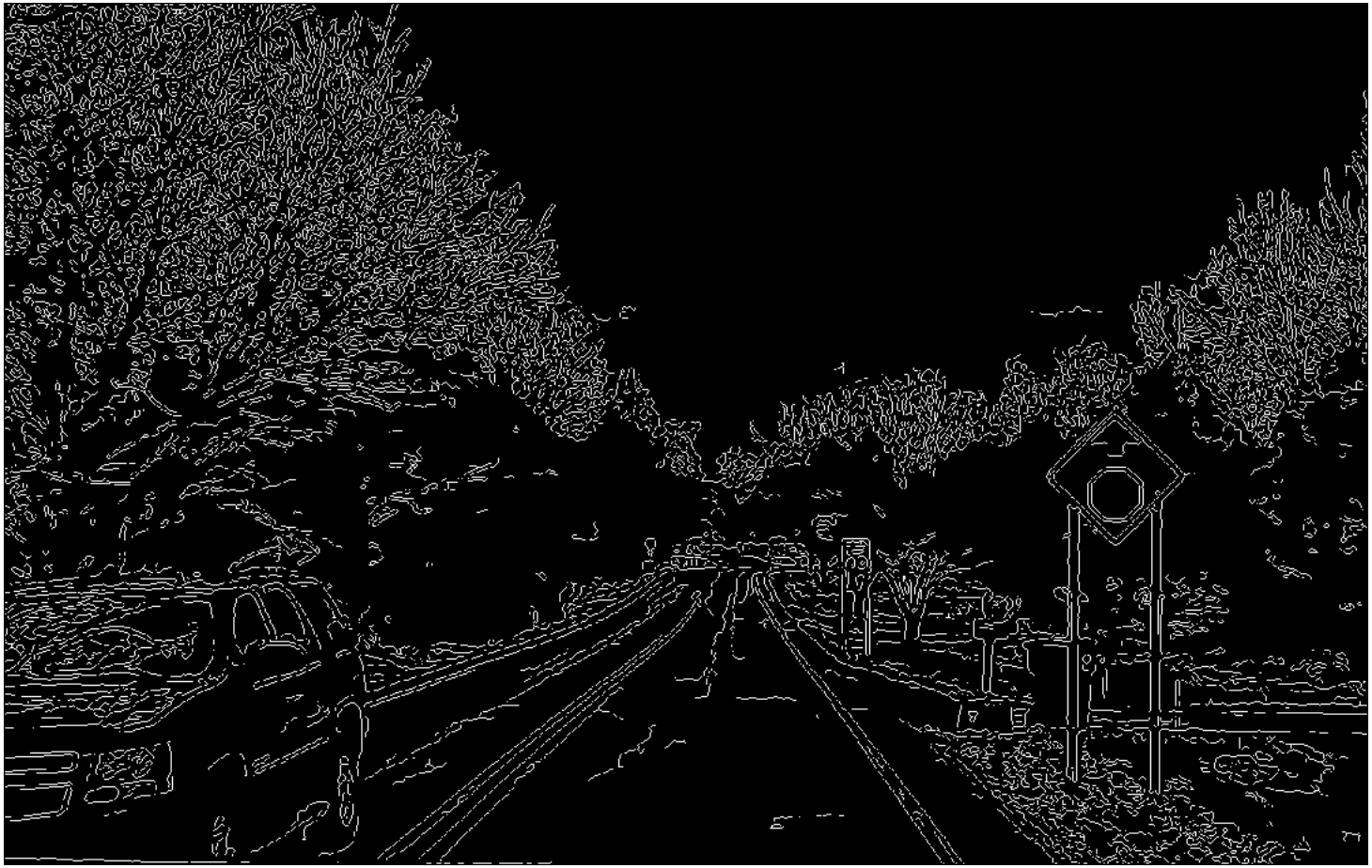
- A collection of spatially distributed light energy
- A 2D array (matrix) of numbers
- Pixel – picture element
- Types
  - Binary
  - Grayscale
  - Color
  - Other spectrum



https://apps.carleton.edu/campus/observatory/observing/knowledge/topics/ipr

Binary Image

- 0s and 1s,1 bit/pixel



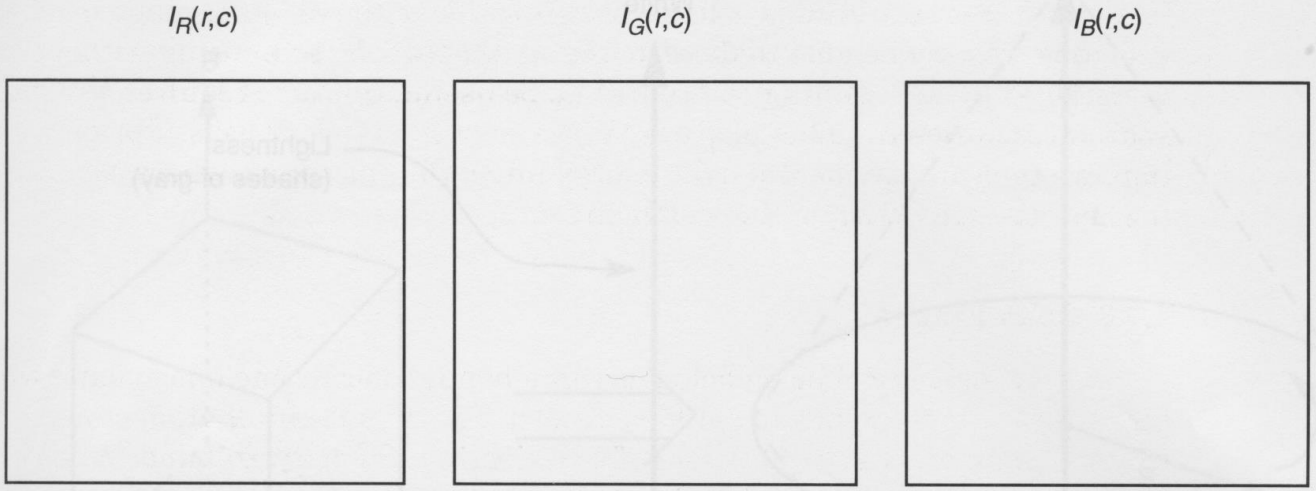
Grayscale Image

- Brightness, no color
- Typically 8 bits/pixel -> 256 different brightness (gray) levels

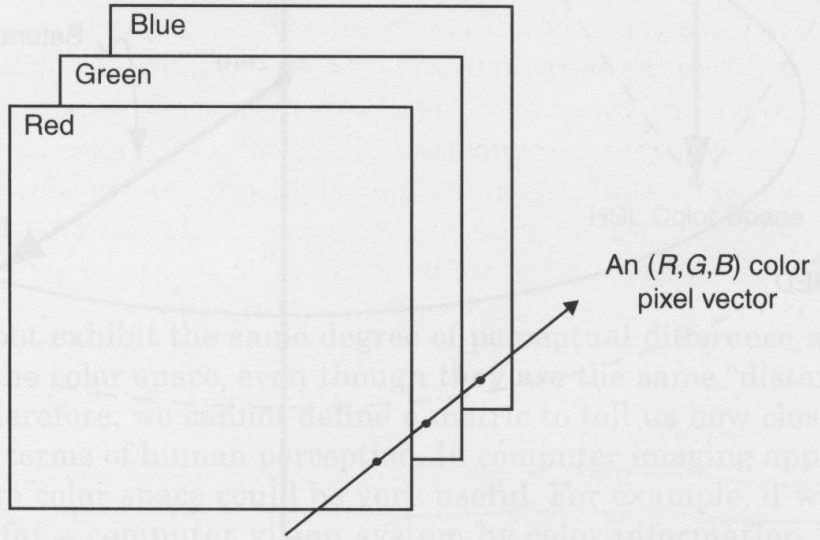
Color Image

- Three-band monochrome image
- Typically 24 bits/pixel (8 bits per band), RGB

RGB



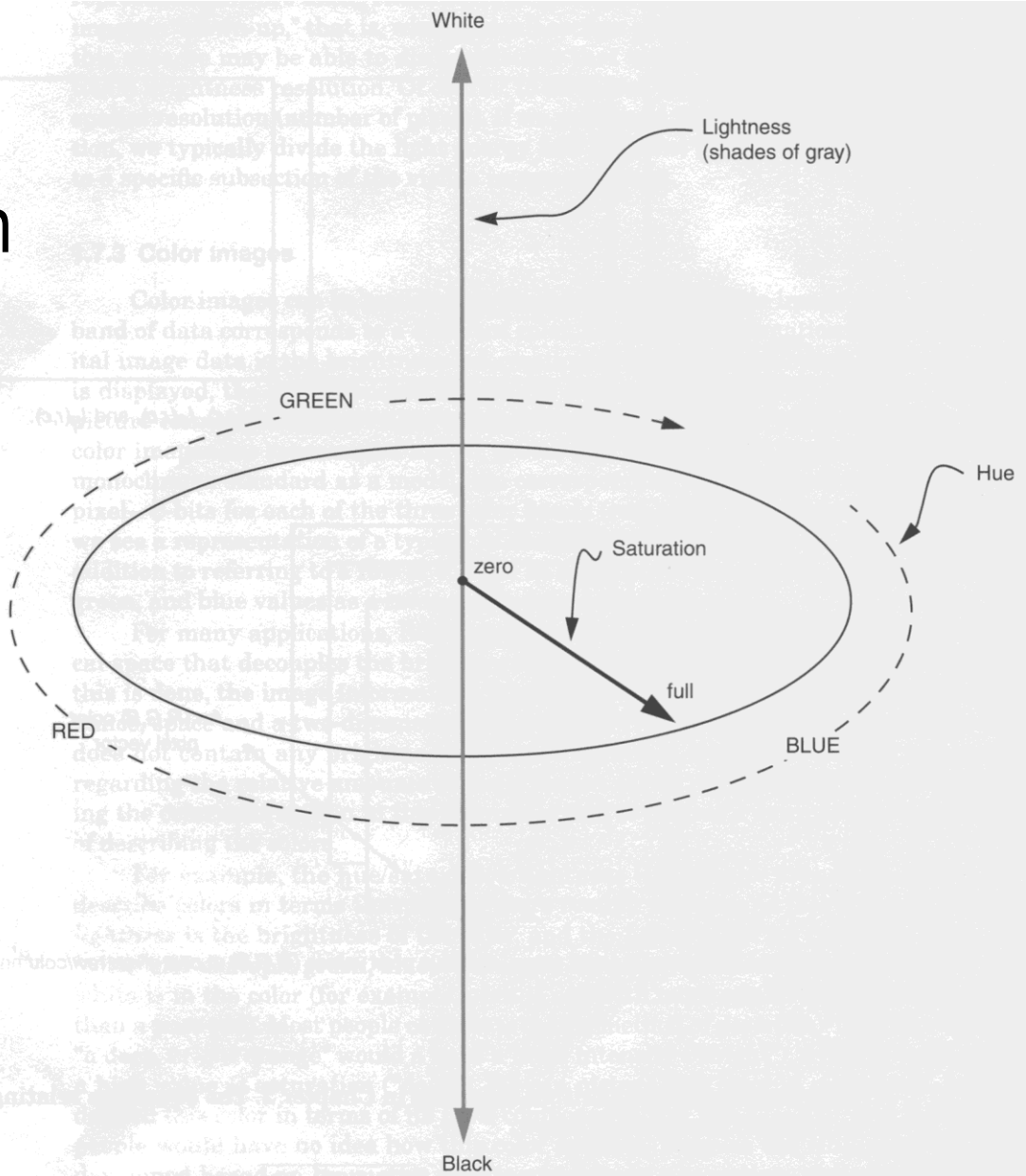
a. A typical RGB color image can be thought of as three separate images: $I_R(r,c)$, $I_G(r,c)$, and $I_B(r,c)$.



b. A color pixel vector consists of the red, green, and blue pixel values (*R,G,B*) at one given row/column pixel coordinate (*r,c*).

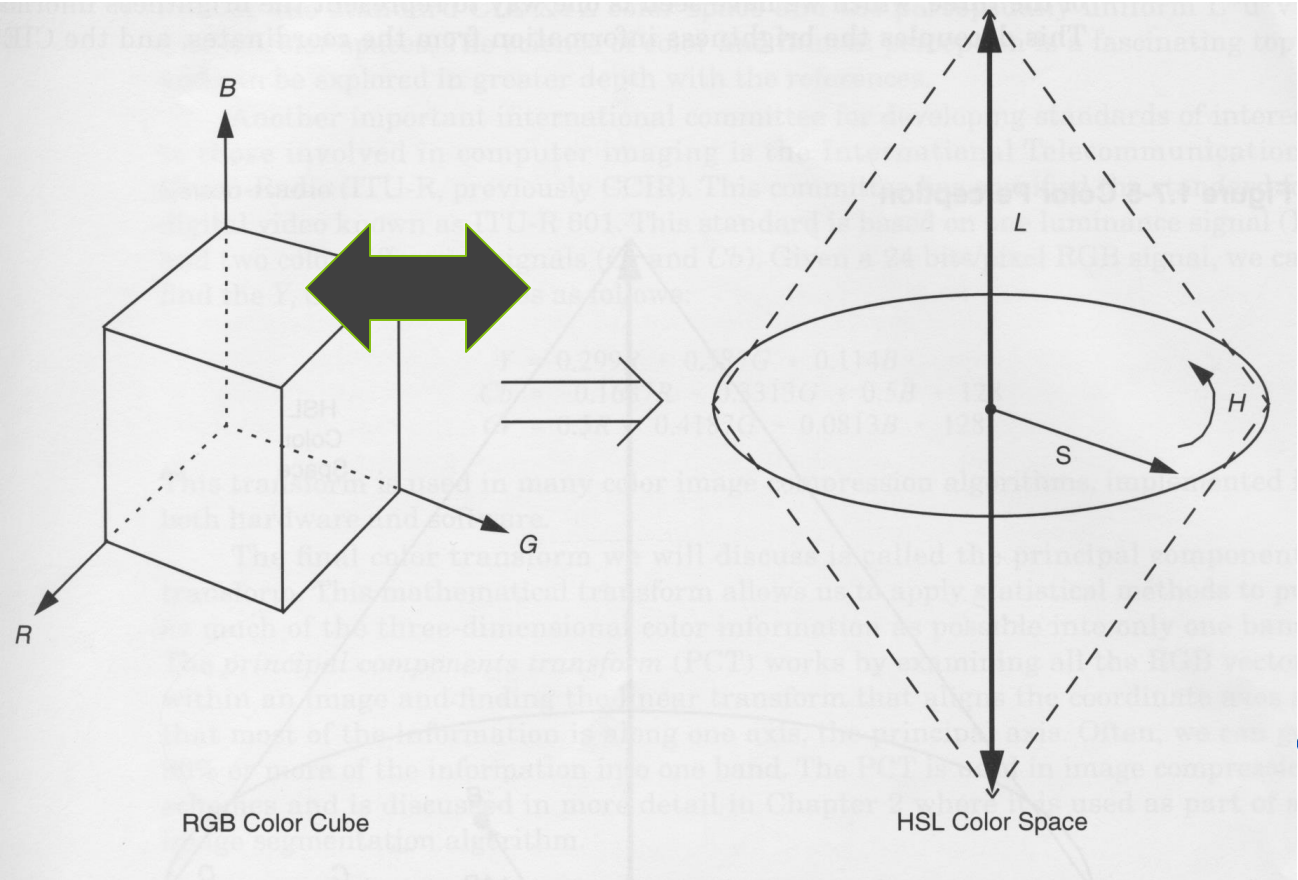
HSL

- **Hue** – color
- **Saturation** – how much white is in the color
- **Lightness** – brightness



Why HSL?

- "A deep, bright orange" vs. "R=245, G=110, B=20"
  - HSL was developed based on human perception
- RGB <-> HSL



Other Spectral Image

- Information outside visible domain: sonar, radar, infrared
- Remote (satellite) sensing

Pixel Depth

- How many bits/pixel are required to create 100 gray levels?

Pixel Depth (cont.)

- 3 bits/pixel – false contours

Pixel Depth (cont.)

- How many bit(s)/pixel are required for a black and white image?

Module Review Questions

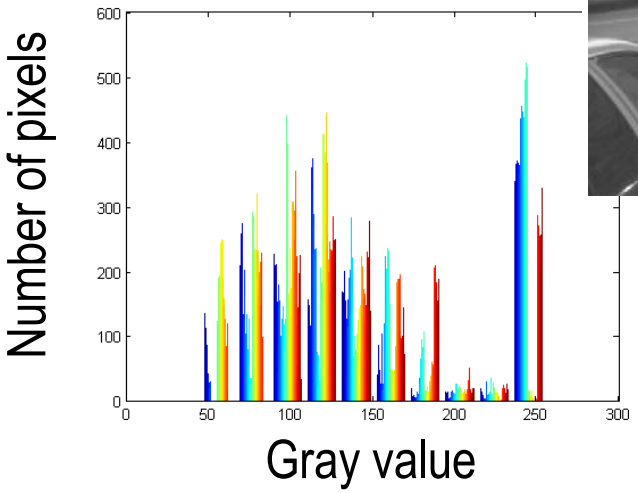
- What are the following terms? How would you set them up in street view collection?
  - Field of View (FoV)
  - Exposure: shutter speed, f-stop
  - Depth of Field
- What is pixel depth?

Image Enhancement

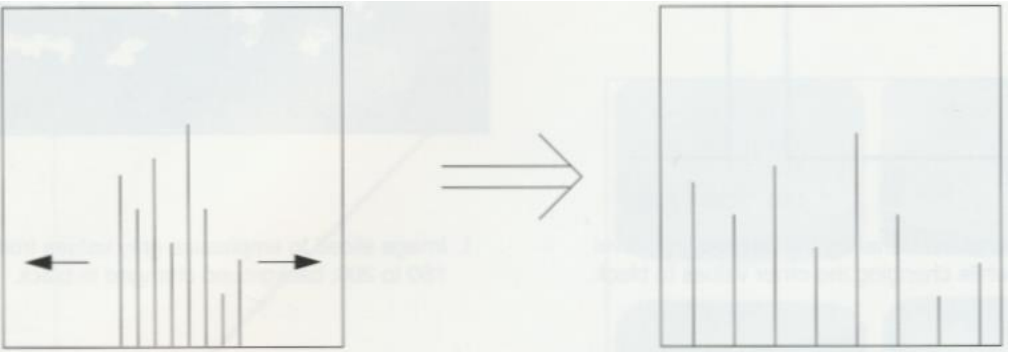
Motivation

- Visually desirable image
- A preprocessing step in computer vision applications

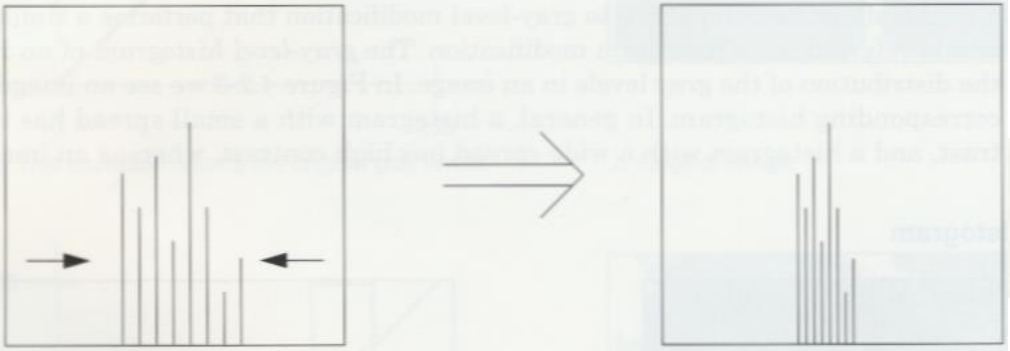
Image Histogram



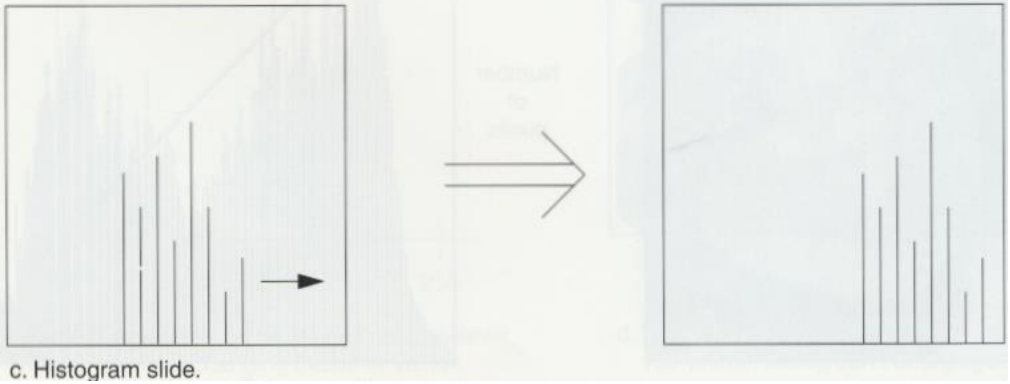
Histogram Modification



a. Histogram stretch.



b. Histogram shrink.



c. Histogram slide.

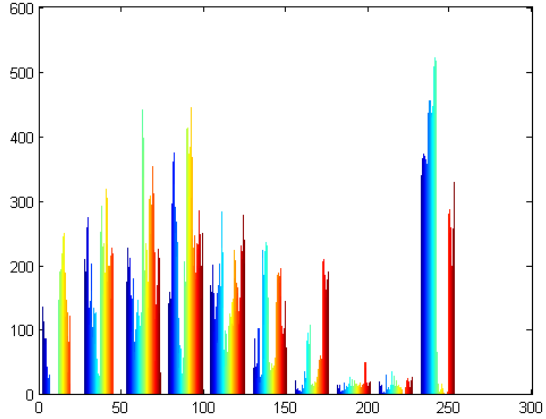
$$\text{Stretch}(I(r,c)) = \left[\frac{I(r,c) - I(r,c)_{\text{MIN}}}{I(r,c)_{\text{MAX}} - I(r,c)_{\text{MIN}}}\right][\text{MAX} - \text{MIN}] + \text{MIN}$$

Histogram Equalization

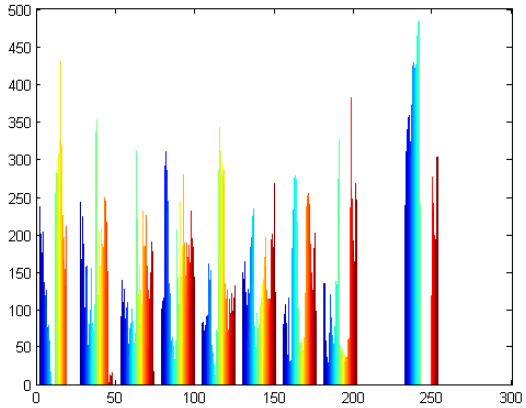
- Histogram of the resultant image is as flat as possible (vs. histogram stretching where the overall shape remains the same)
- Provides more visually pleasing results across a wider range of images

Histogram Equalization (cont.)

Histogram stretch



Histogram equalization



Contrast Limited Adaptive Histogram Equalization (CLAHE)

- Operates on small regions in the image, called *tiles*
- After performing the equalization, CLAHE combines neighboring tiles using bilinear interpolation to eliminate artificially induced boundaries.

CLAHE

Origin

Histogram Equalization

CLAHE (cont.)

- More examples

White Balance

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255/R'_w & 0 & 0 \\ 0 & 255/G'_w & 0 \\ 0 & 0 & 255/B'_w \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

Auto White Balance

The auto white balance was unable to find a white reference, resulting in dull and artificial colors.

The auto white balance got it right this time in a very similar scene because it could use the clouds as its white reference.

Street Level Imagery Quality Issues

- **Wash-out**
  - **Overexposure due to the aperture/shutter speed adjustment slower than lighting condition change**
- **Low light (may be corrected with histogram modification)**
- **Blur**
  - **Motion blur**
  - **Not much out-of-focus**
- **Poor White Balance**
- **Dirt on lens**

Image Processing

Motivation for Geospatial Industry - Automation

Motivation for Geospatial Industry - Automation (cont.)

Image Addition

- Application1: image morphing

- Application2: modeling additive noise for image restoration

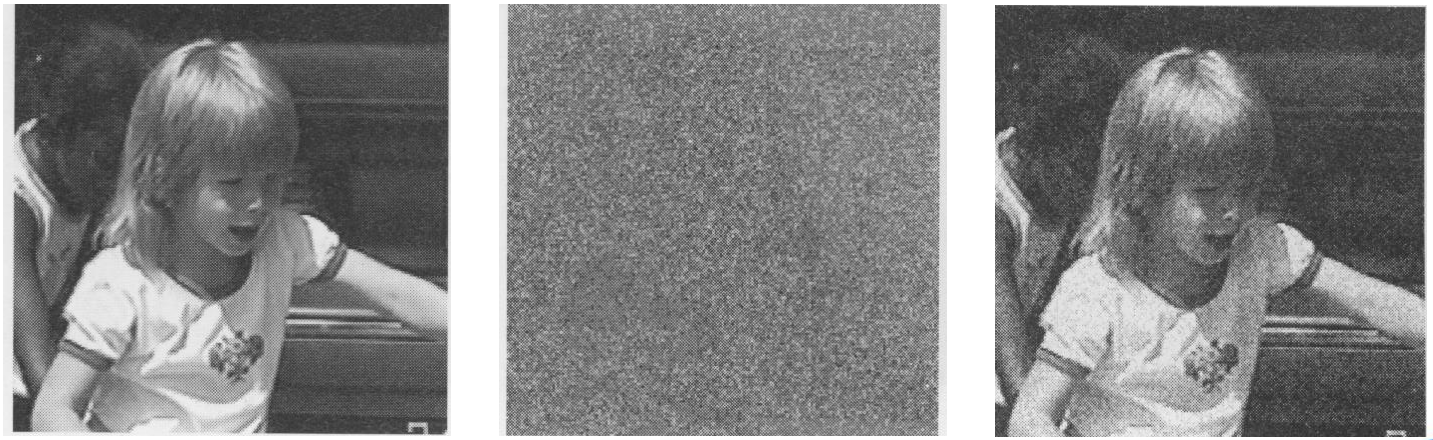
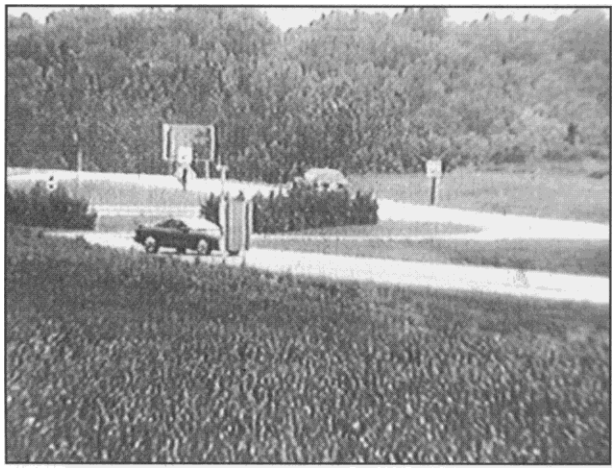


Image Subtraction

- Motion/Change Detection
- Foreground/background detection



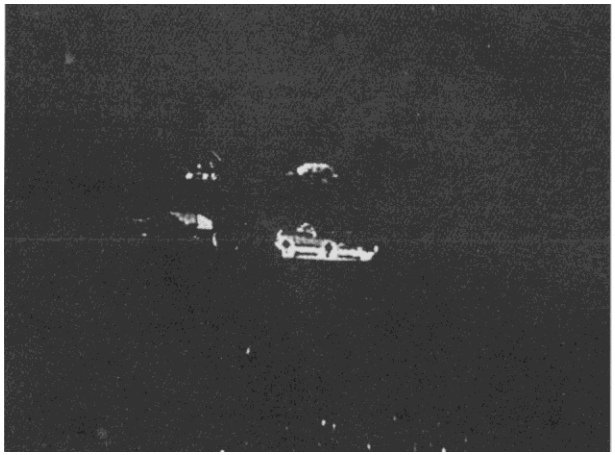


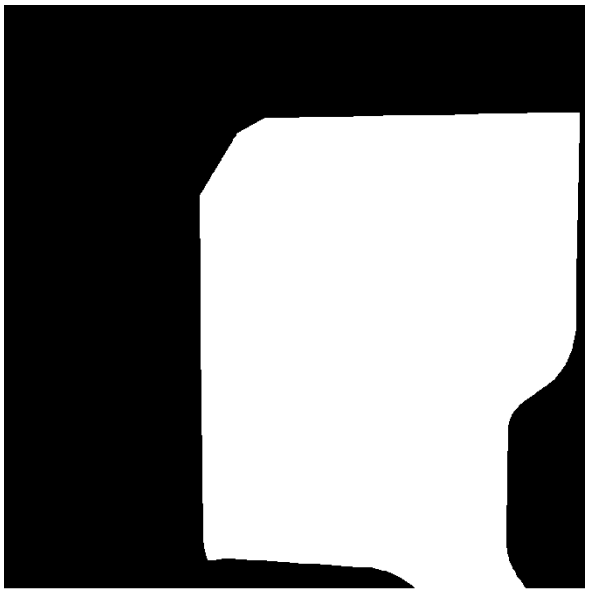
Image Multiplication/Division

- Adjust image brightness

Image AND/OR Operation

- Masking

&

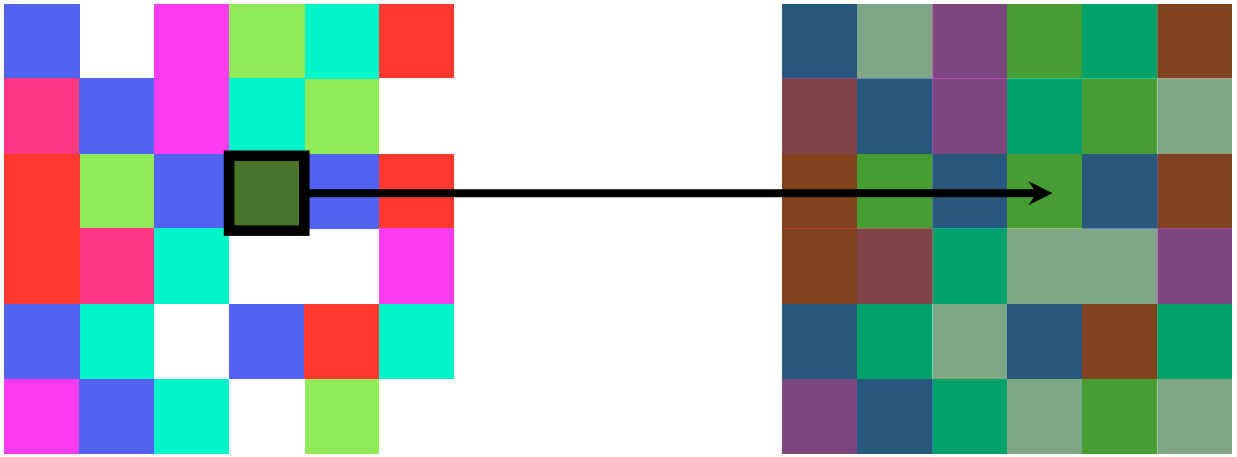


=

Image NOT Operation

- Negative of the image

Pixel Processing

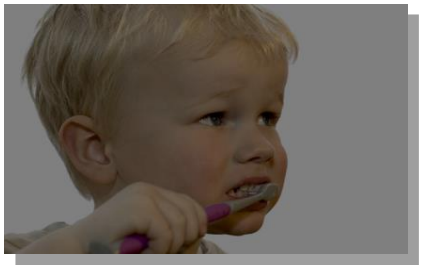


Pixel Processing Examples

Original

Darken

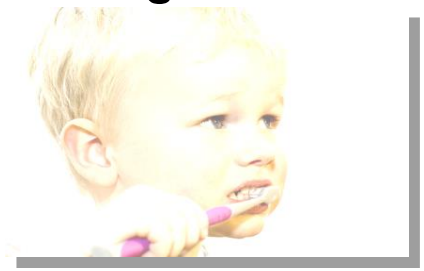
Lower Contrast



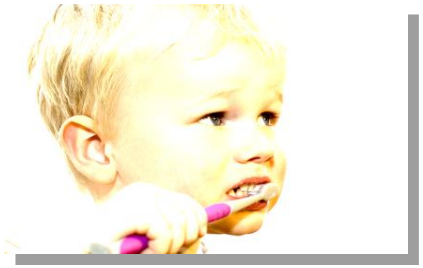
Nonlinear Lower Contrast

Invert

Lighten



Raise Contrast



Nonlinear Raise Contrast

Pixel Processing Examples (cont.)

Original

`x`

Darken

`x - 128`

Lower Contrast

`x / 2`

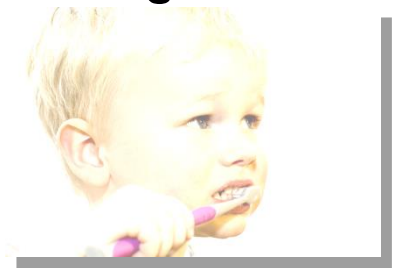
Nonlinear Lower Contrast

`((x / 255.0) ^ 0.33) * 255.0`

Invert

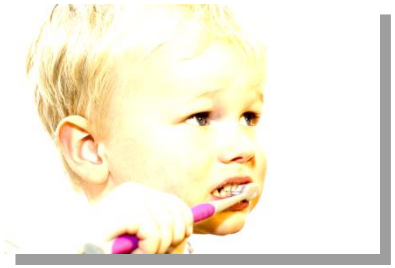
`255 - x`

Lighten



`x + 128`

Raise Contrast

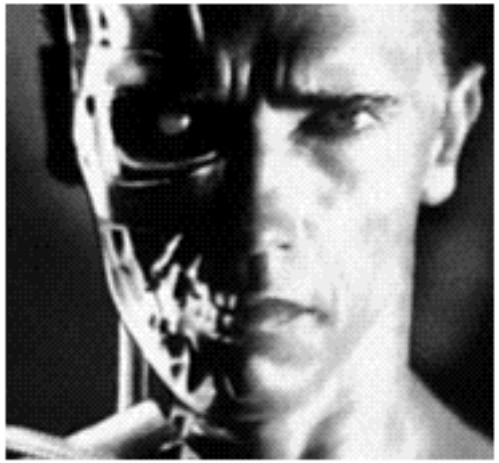


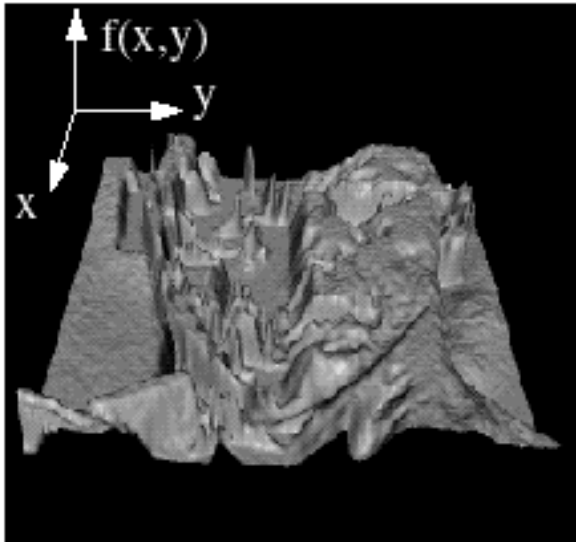
`x * 2`

Nonlinear Raise Contrast

`((x / 255.0) ^2) * 255.0`

Image as a Discrete Function





Represented by a matrix

i ↓ , j →

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

Image as a Discrete Function (cont.)

- We can think of an image as a function, ***f***,

$$f: R^2 \rightarrow R$$

  - $f(x, y)$ gives the intensity at position (x, y)
  - Realistically, we expect the image only to be defined over a rectangle, with a finite range:

$$f: [a,b]\text{x}[c,d] \rightarrow [0,1]$$

- A color image is just three functions pasted together.

Neighborhood Processing

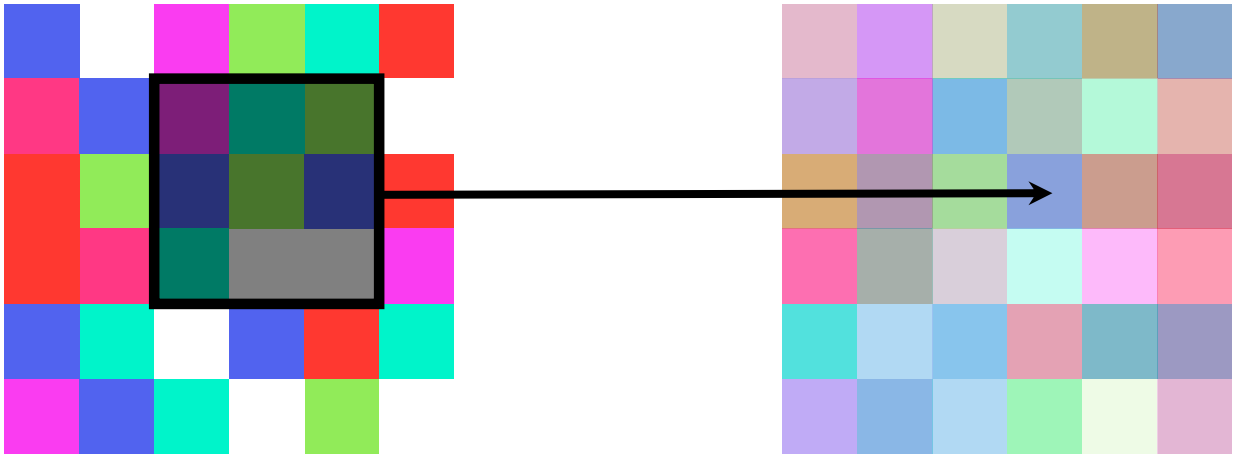
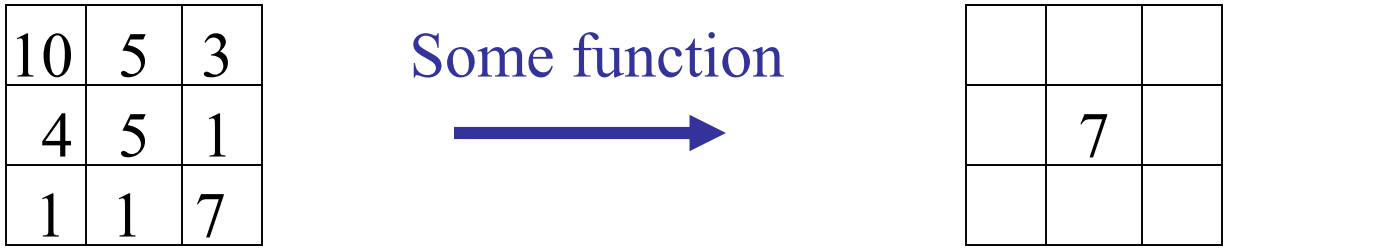


Image Filtering

- **Modify the pixels in an image based on some function of a local neighborhood of the pixels.**



Local image data

Modified image data

Linear Functions

- **Simplest: linear filtering**
  - **Replace each pixel by a linear combination of its neighbors.**
- **The prescription for the linear combination is called the "convolution kernel."**

10	5	3
4	5	1
1	1	7

Local image data

0	0	0
0	0.5	0
0	1	0.5

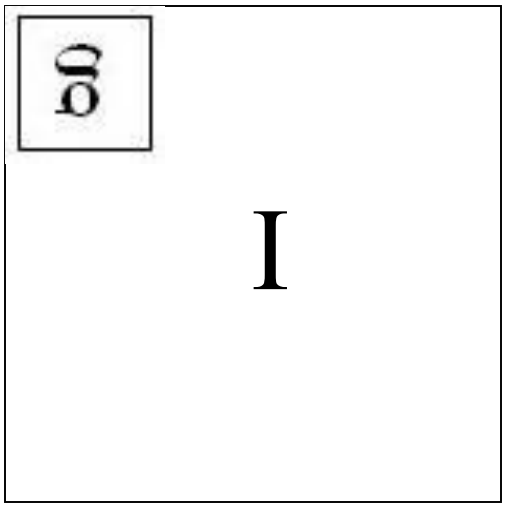
kernel

	7	

Modified image data

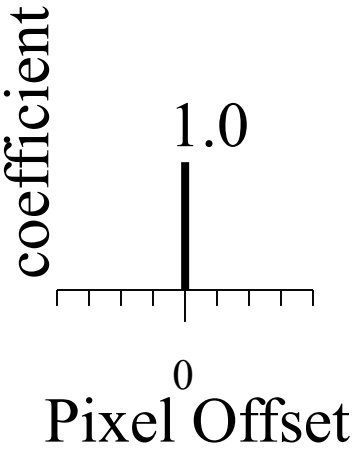
Convolution

$$f[m,n] = I \otimes g = \sum_{k,l} I[m-k, n-l] g[k,l]$$



Linear Filtering

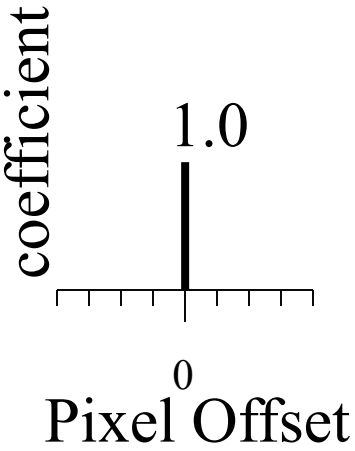
Original



?

Linear Filtering (cont.)

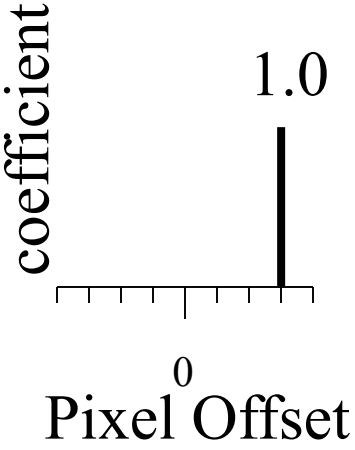
Original



Filtered
(no change)

Linear Filtering (cont.)

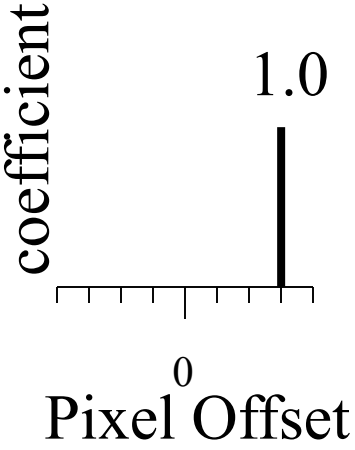
Original

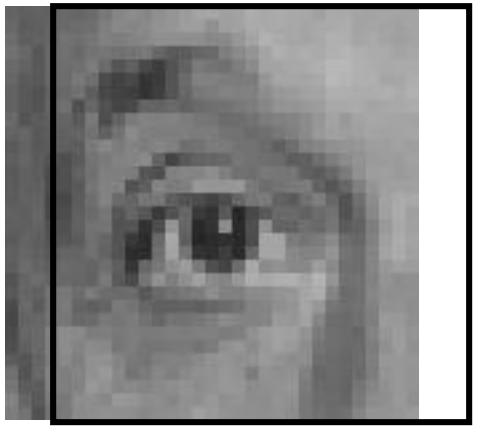


?

Shift

Original

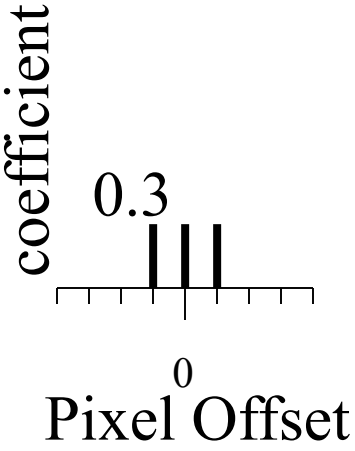




Shifted

Linear Filtering (cont.)

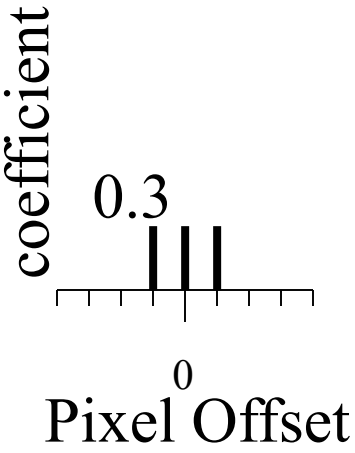
Original



?

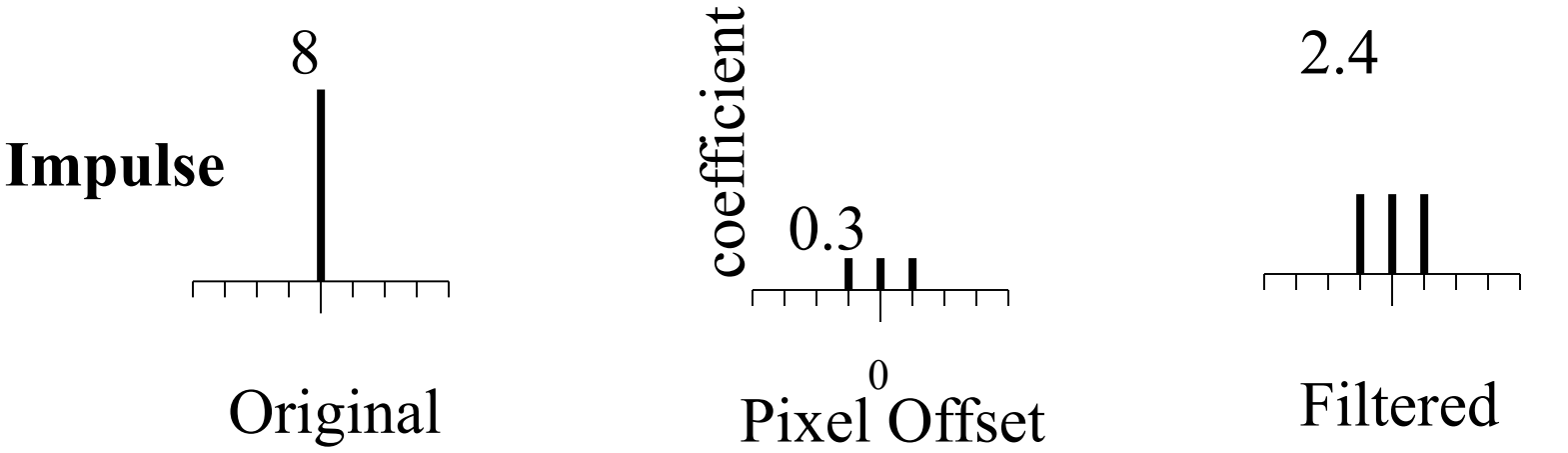
Blurring

Original



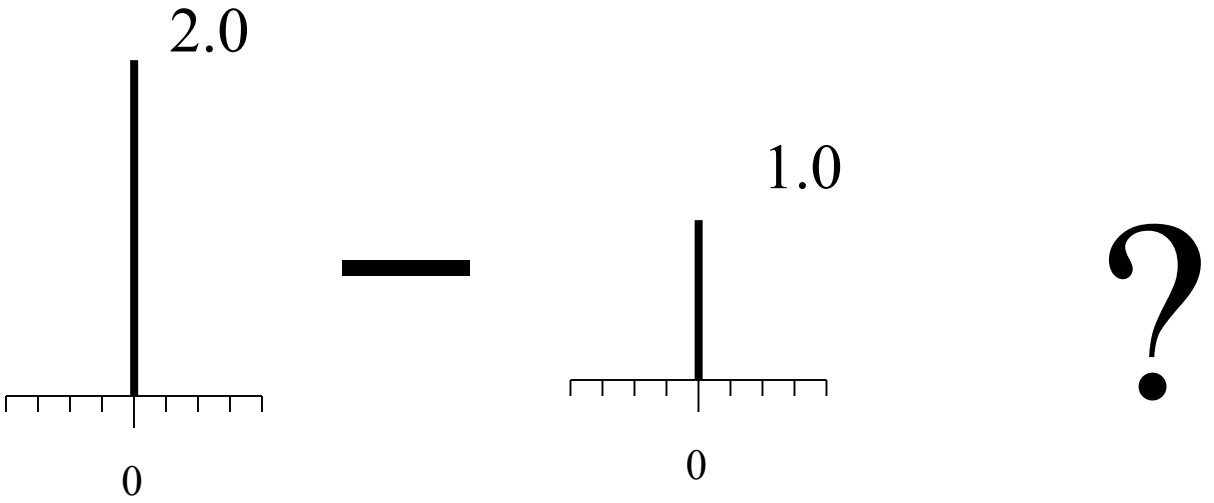
Blurred (filter applied in both dimensions).

Blur Examples

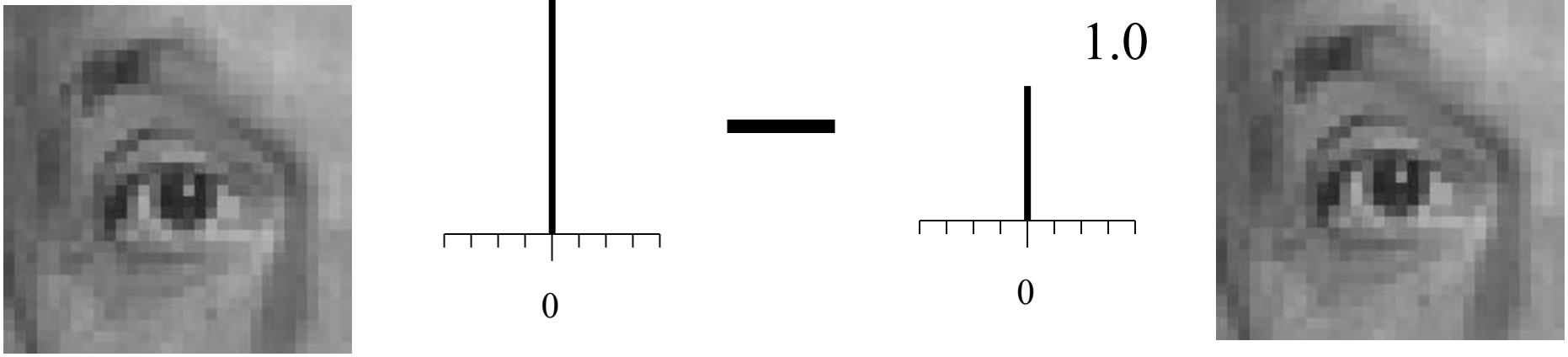


Linear Filtering (cont.)

Original



Linear Filtering (cont.)

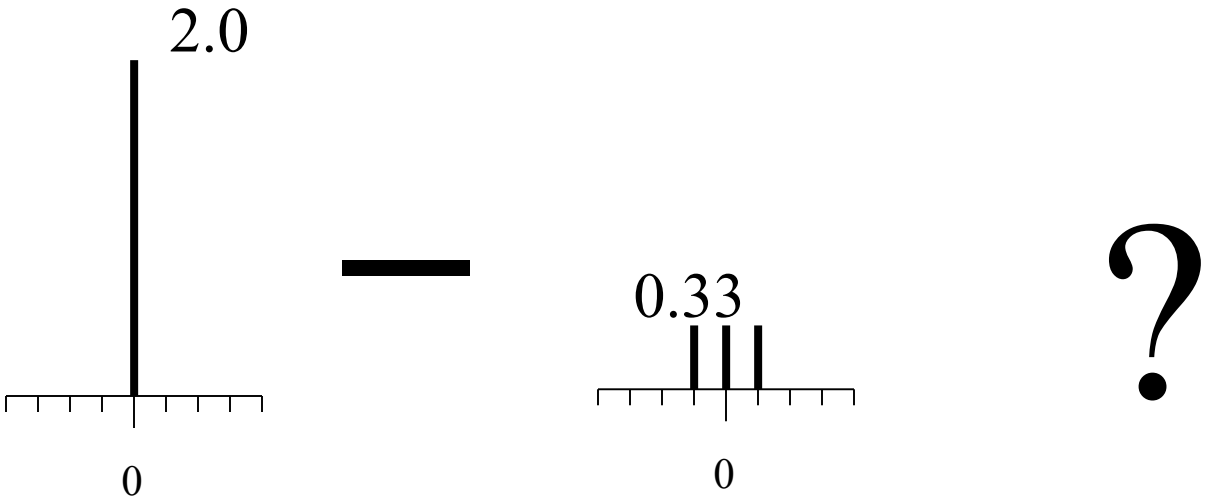


Original

Filtered
(no change)

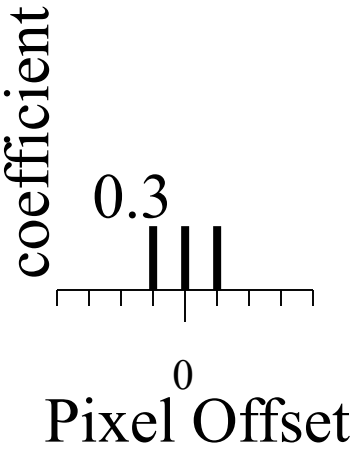
Linear Filtering (cont.)

Original



Recall Blurring

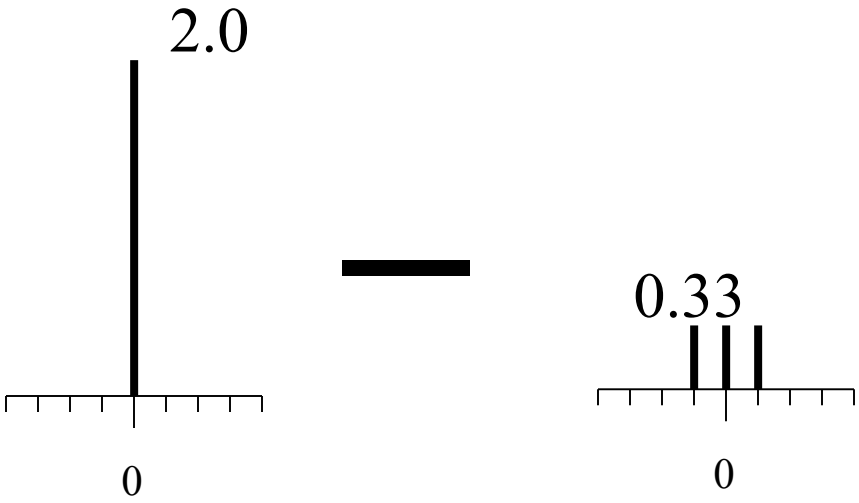
Original

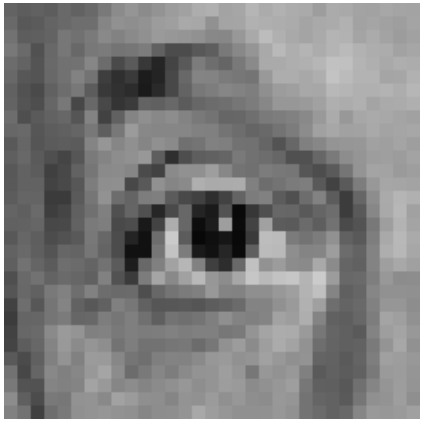


Blurred (filter applied in both dimensions).

Sharpening

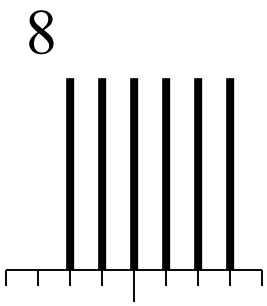
original



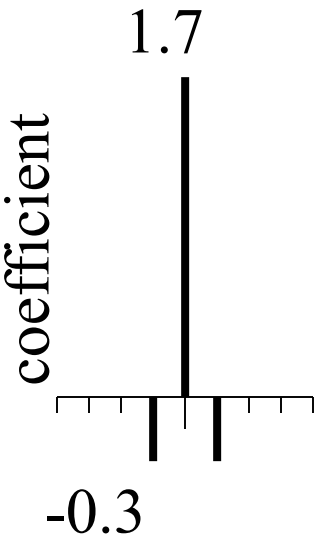


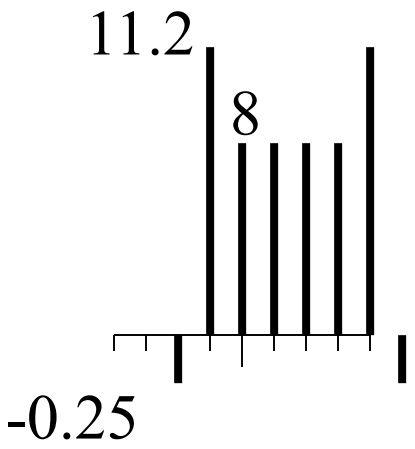
Sharpened
original

Sharpening Example



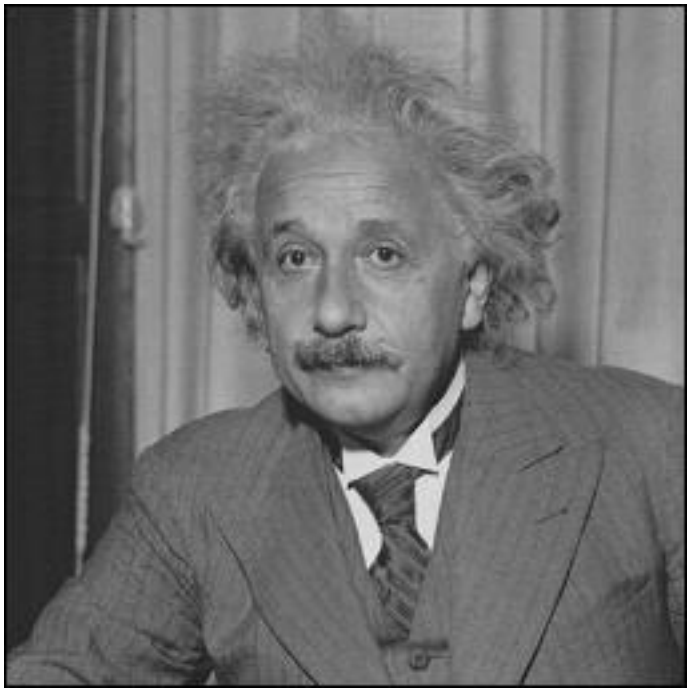
original



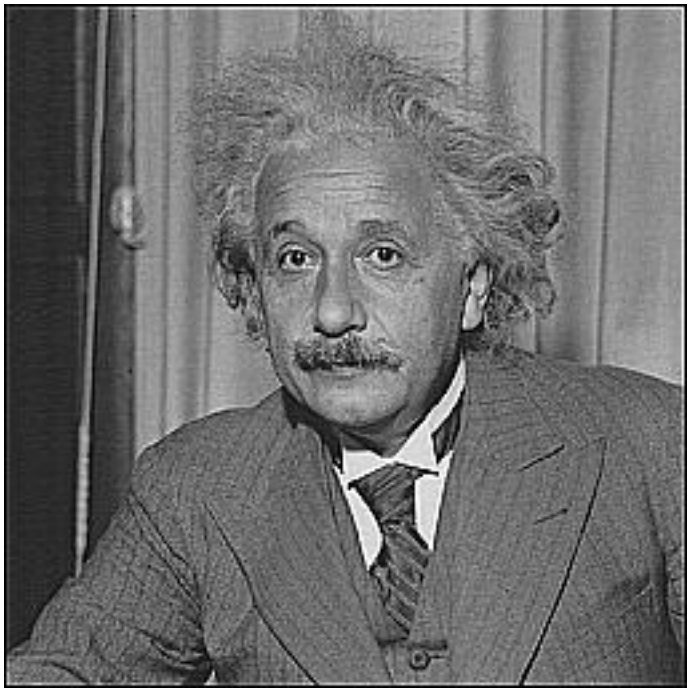


Sharpened (differences are accentuated; constant areas are left untouched).

Image Sharpening



before



after

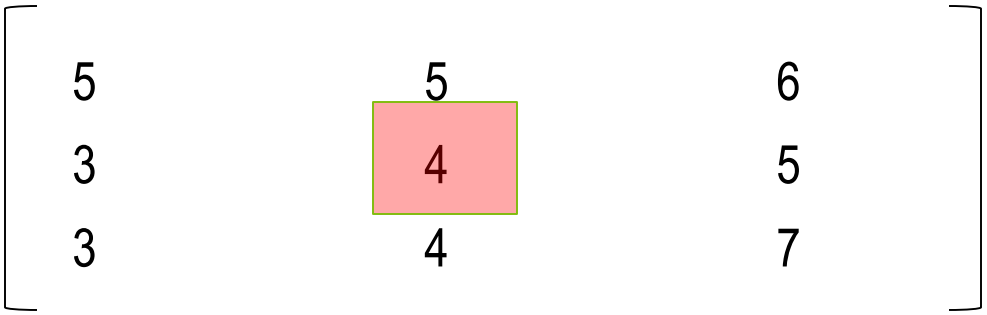
Image Filters

- **Mean filter**

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

Convolution kernel

- **Median filter (non-linear)**

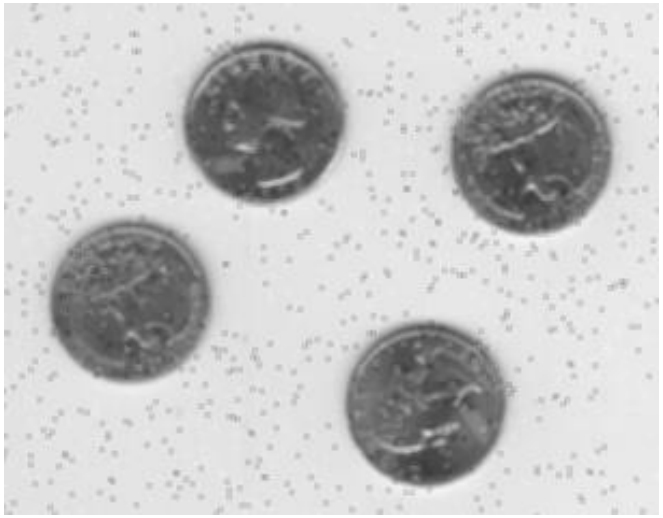


Neighborhood of the original image

3, 3, 4, 4, 5, 5, 5, 6, 7

Image Filter Examples

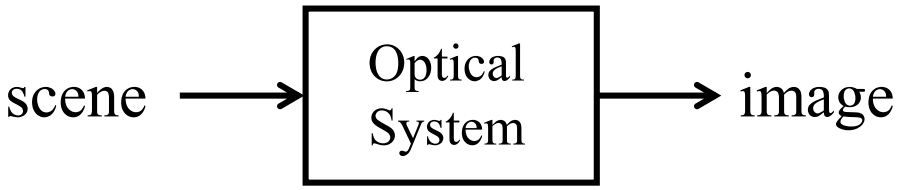
Origin image



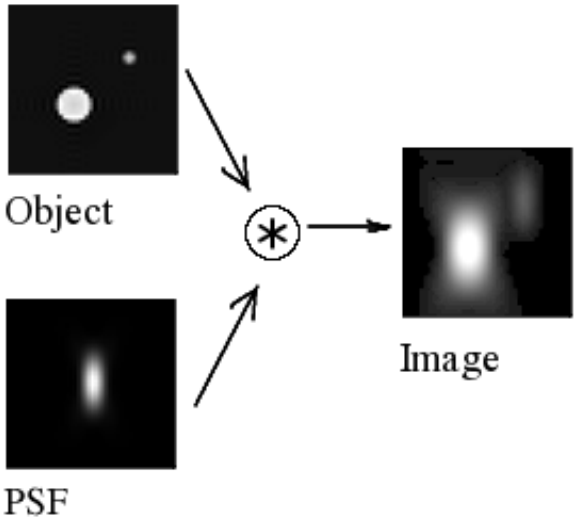
Mean filtered image

Median filtered image

Point Spread Function



- **Optical system is never perfect**
  - **The image can be seen as a convolution of the true object and the PSF**



Point Spread Function (cont.)

normal vision

myopia

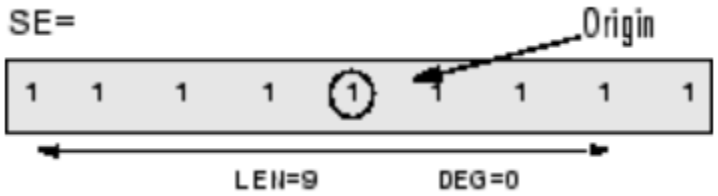
hyperopia

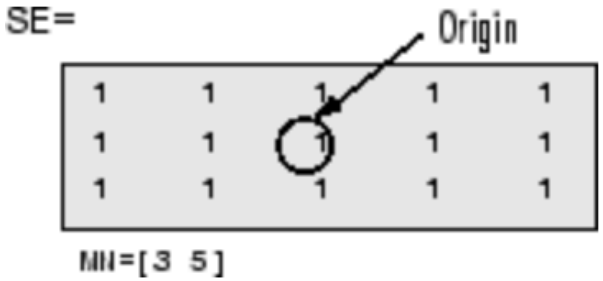
astigmatism

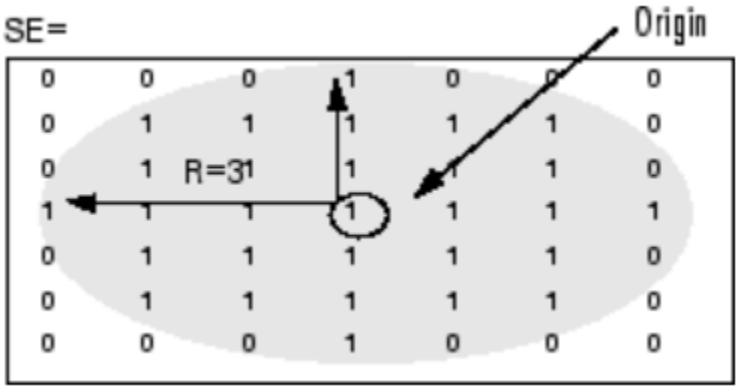
Morphological Image Processing

Image Dilation

- Two inputs
  - A binary image
  - A structuring element (similar to convolution kernel)







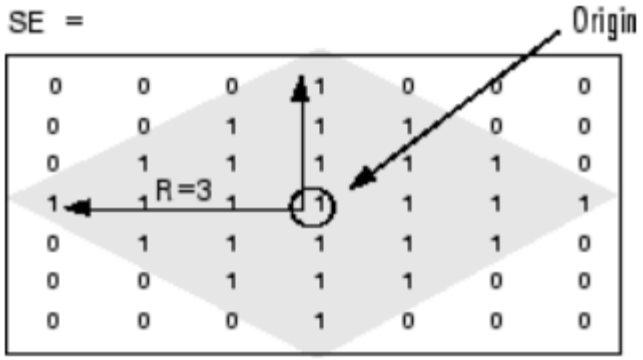
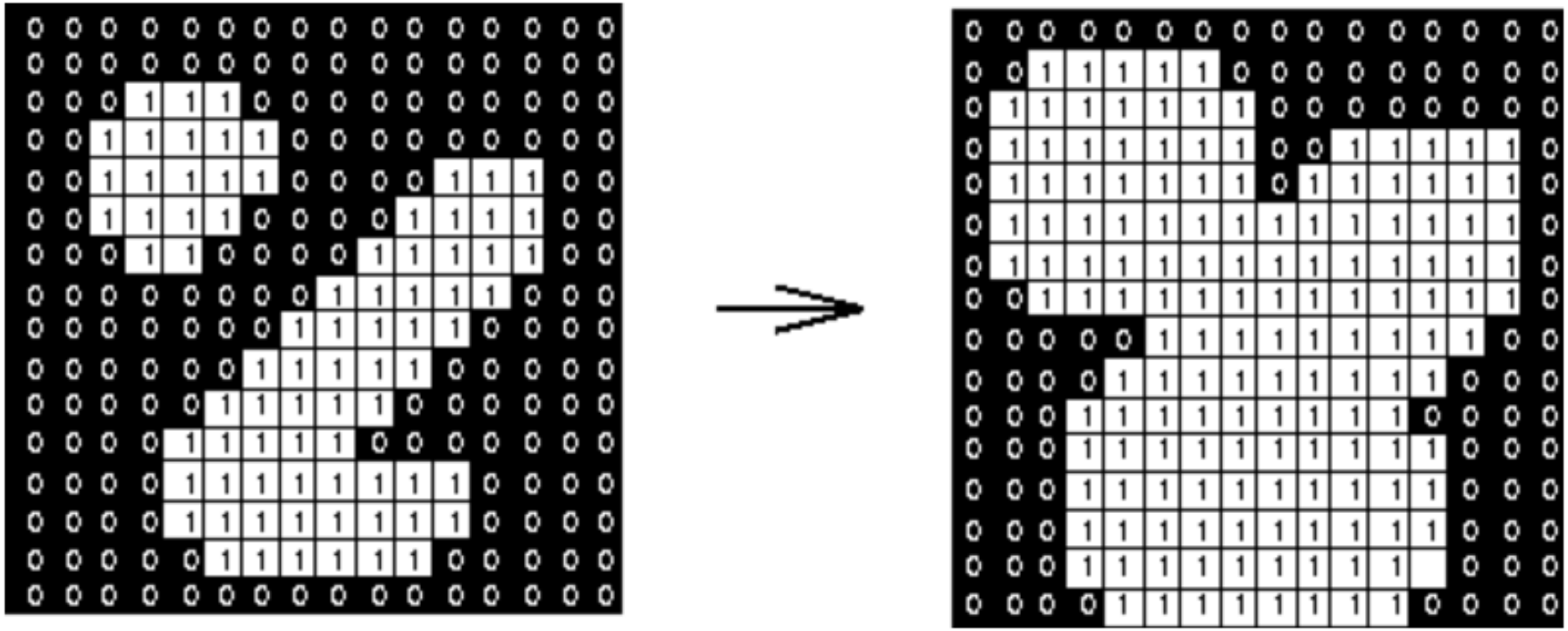


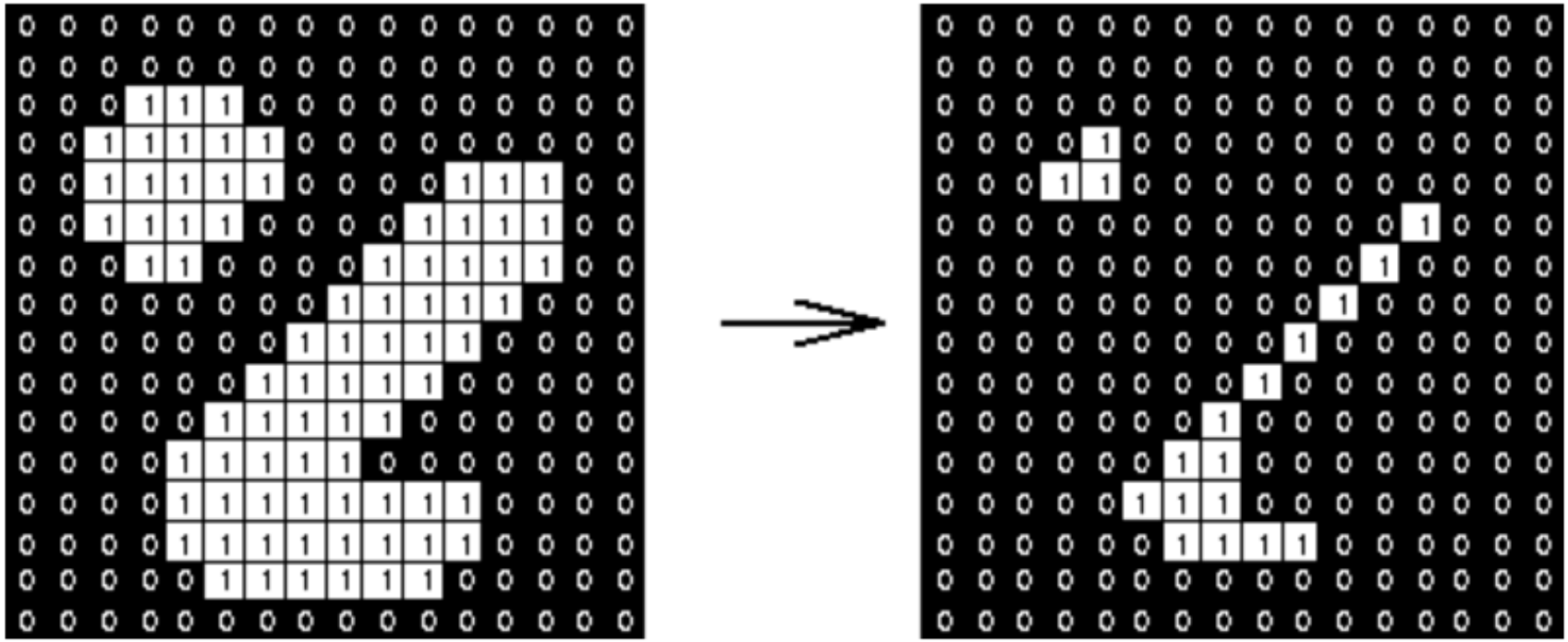
Image Dilation (cont.)



3x3 square structuring element

Image Erosion

- The opposite of dilation



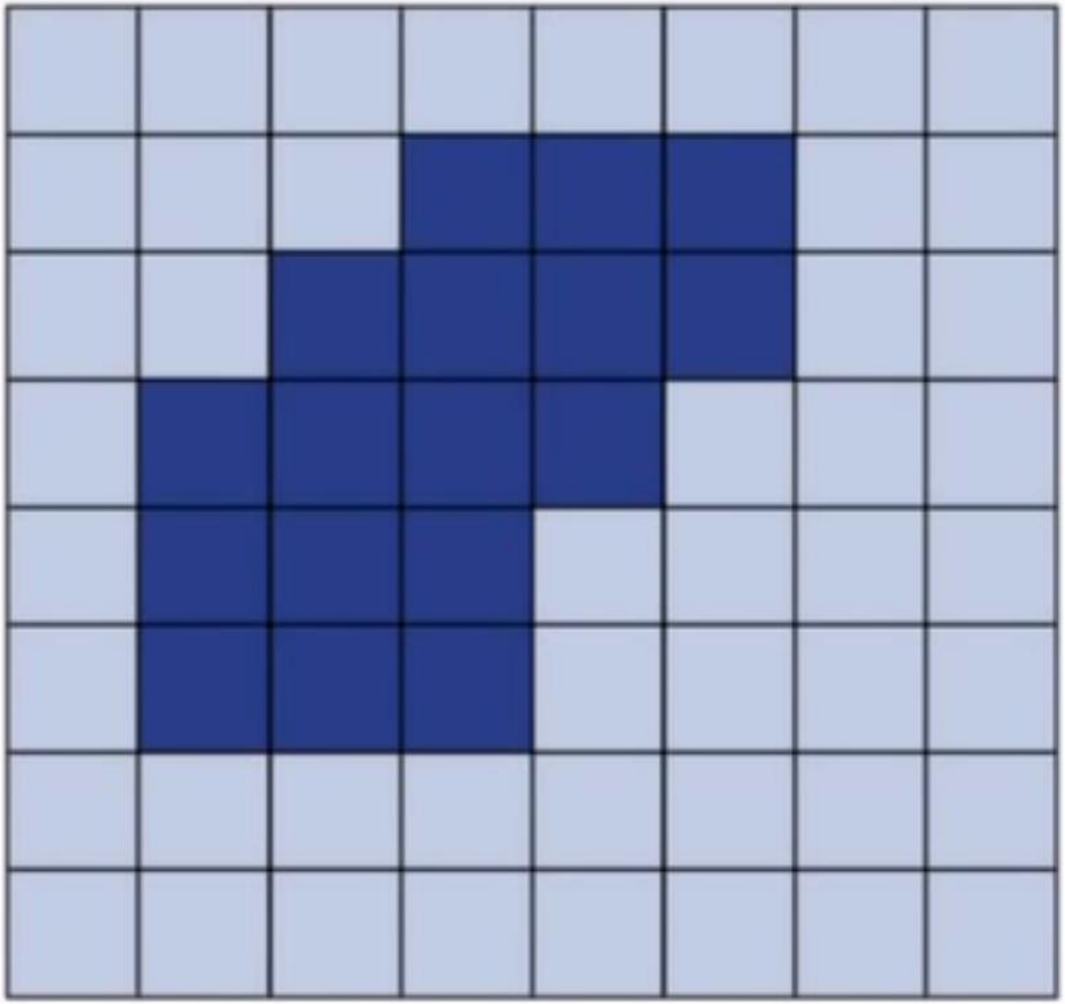
3x3 square structuring element

An Erosion Example

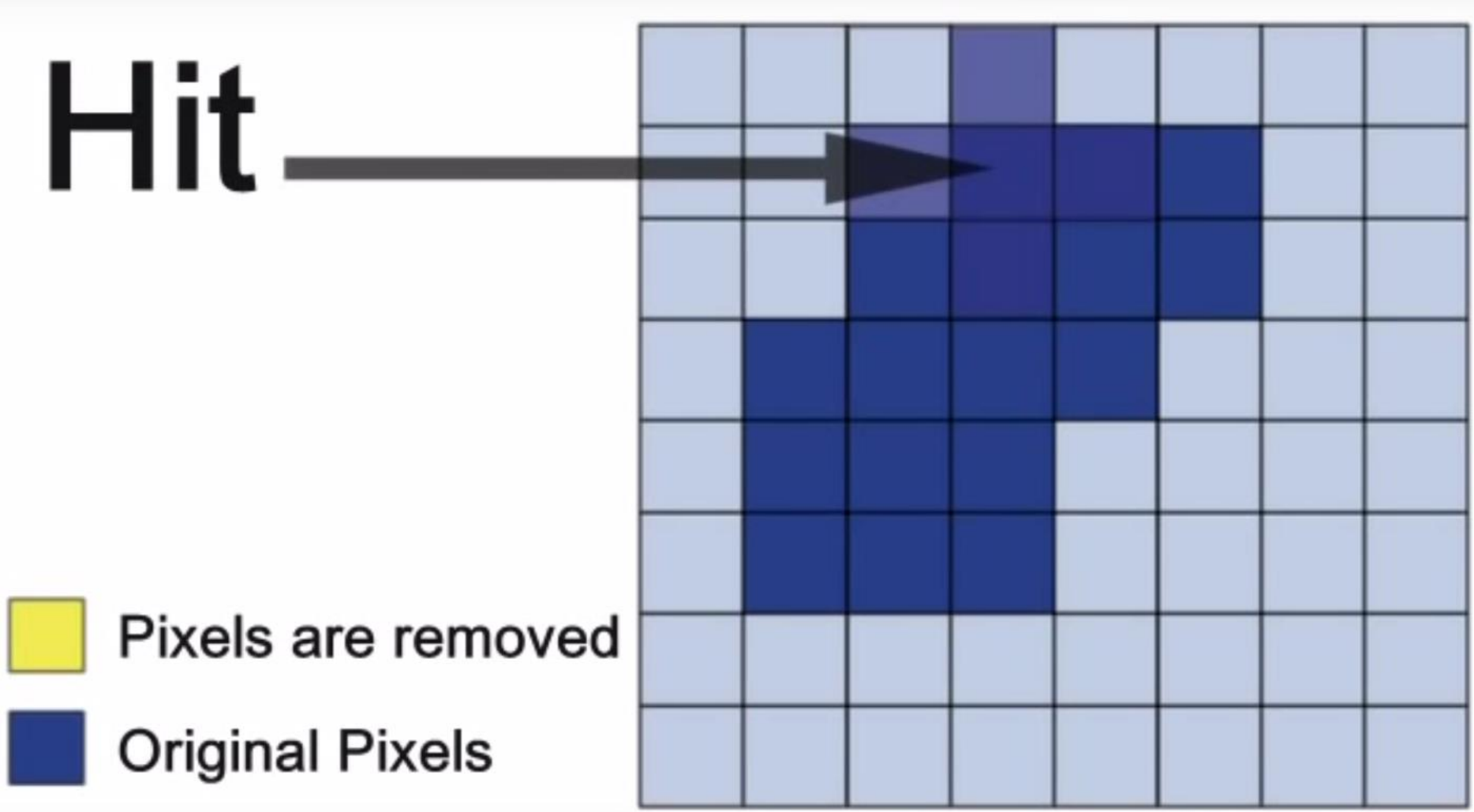
Structuring Element

Pixels are removed

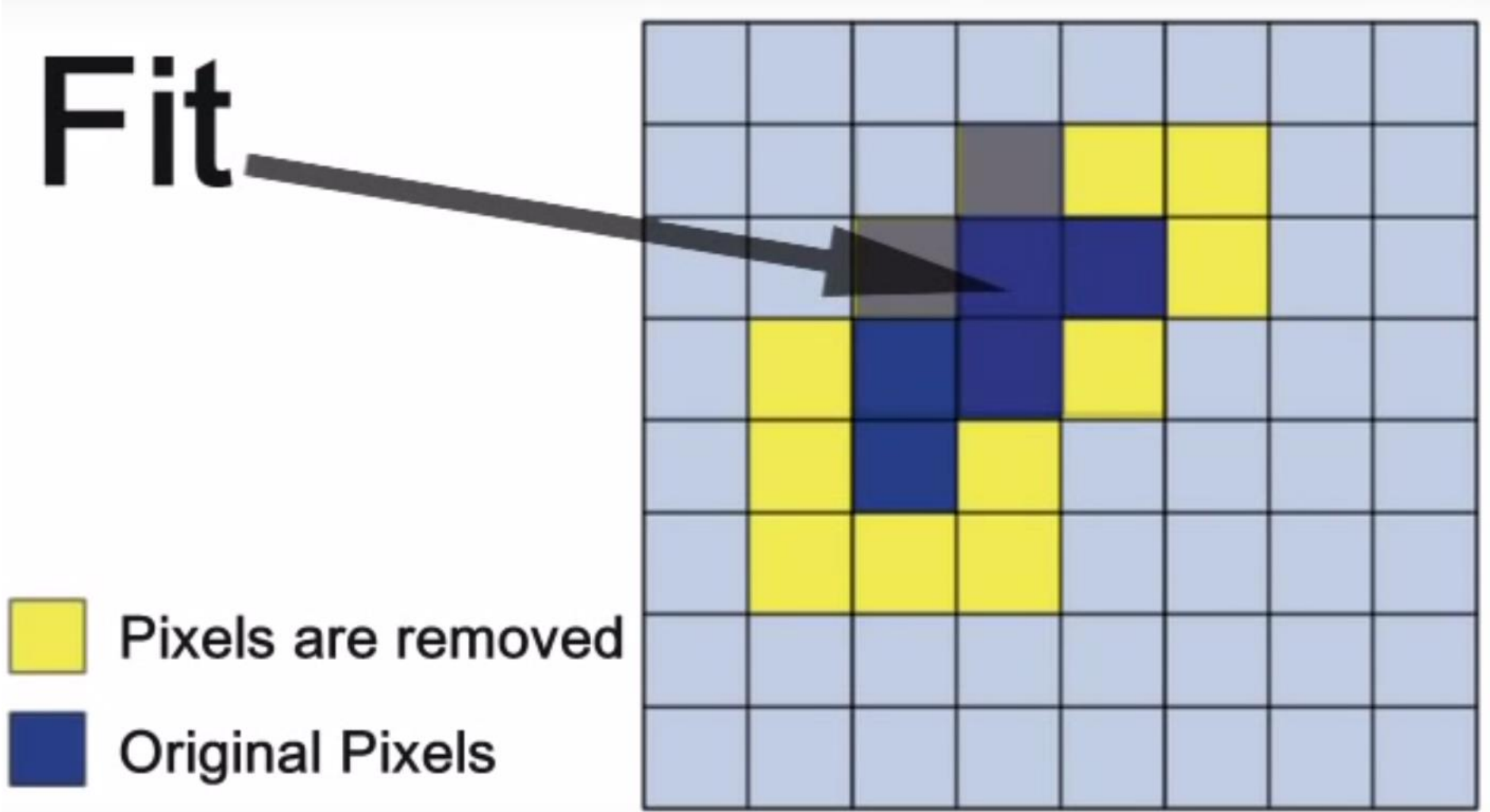
Original Pixels



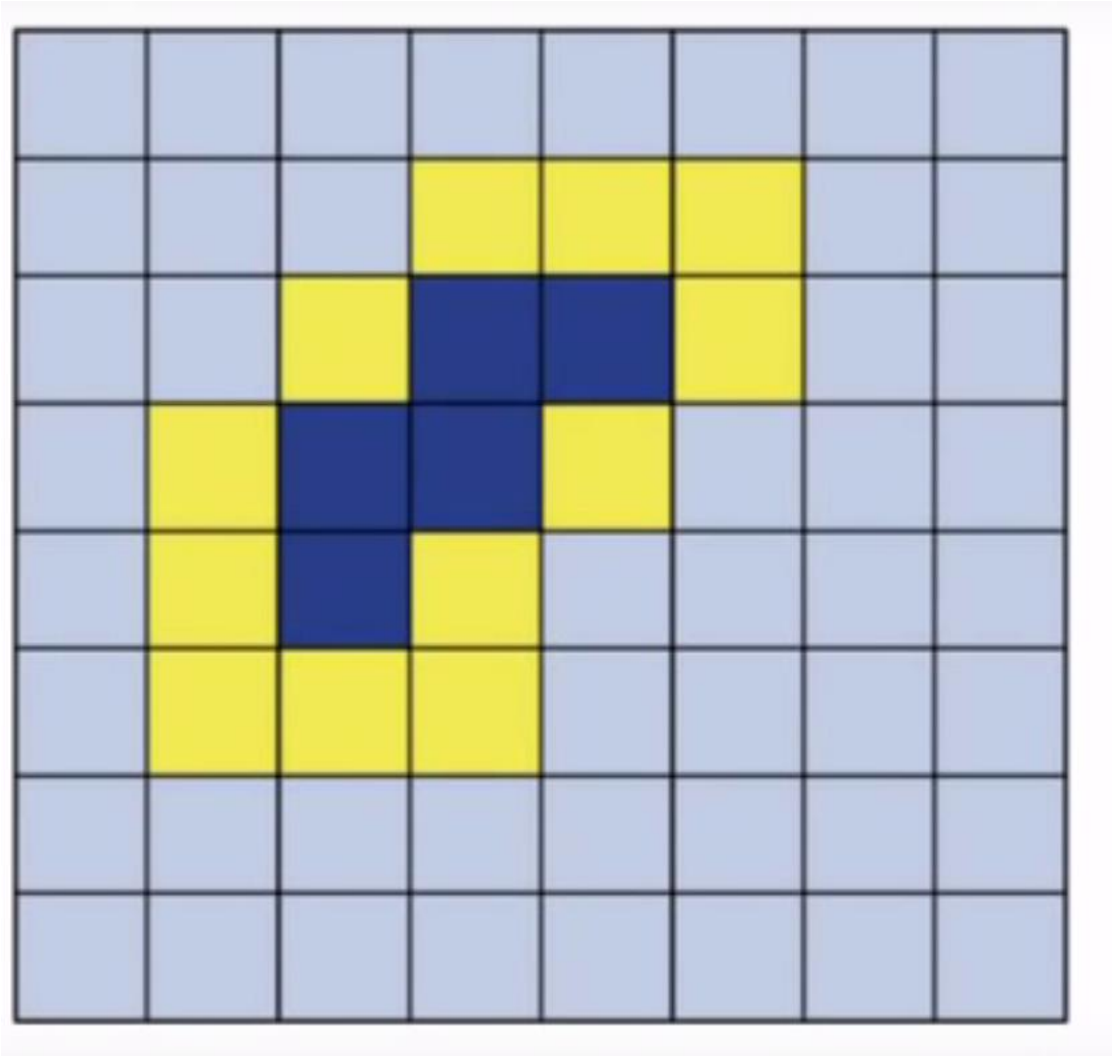
An Erosion Example



An Erosion Example



An Erosion Example

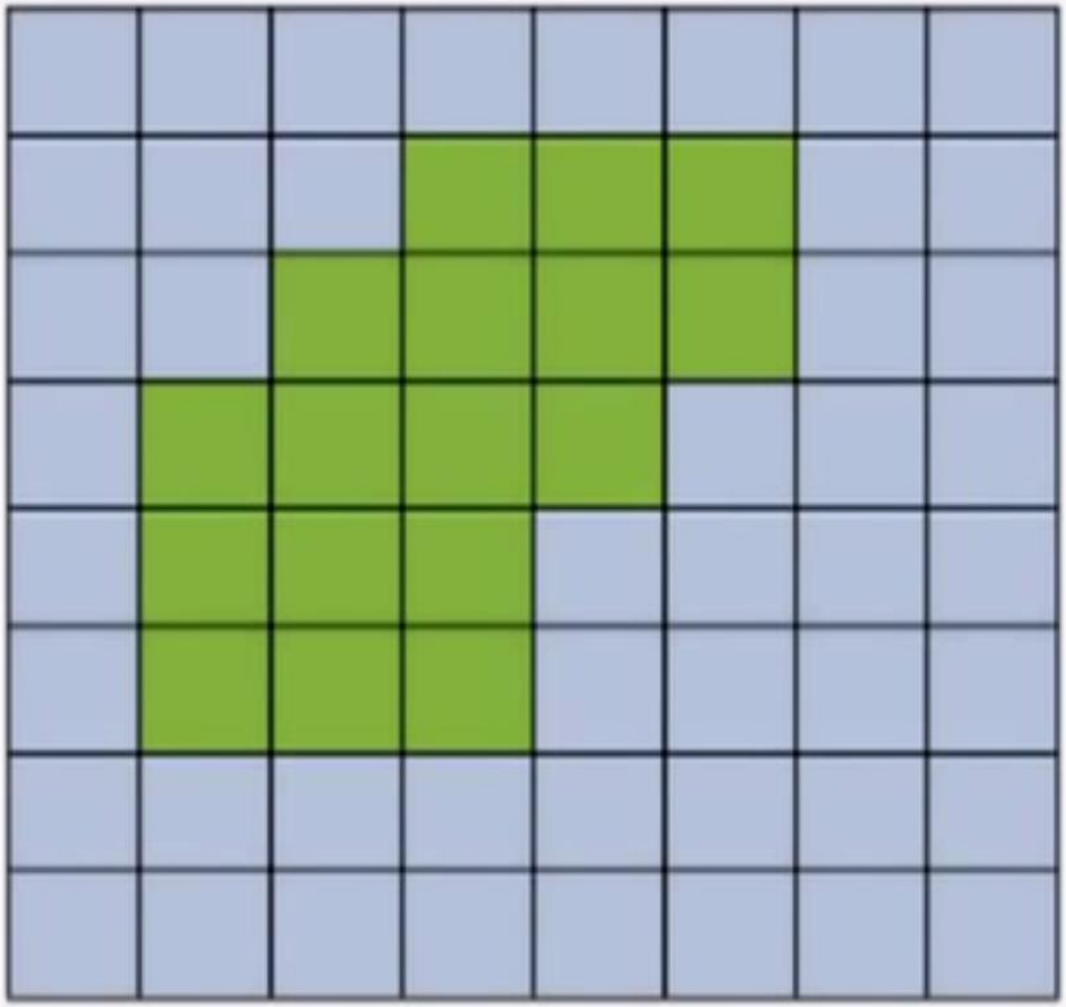


A Dilation Example

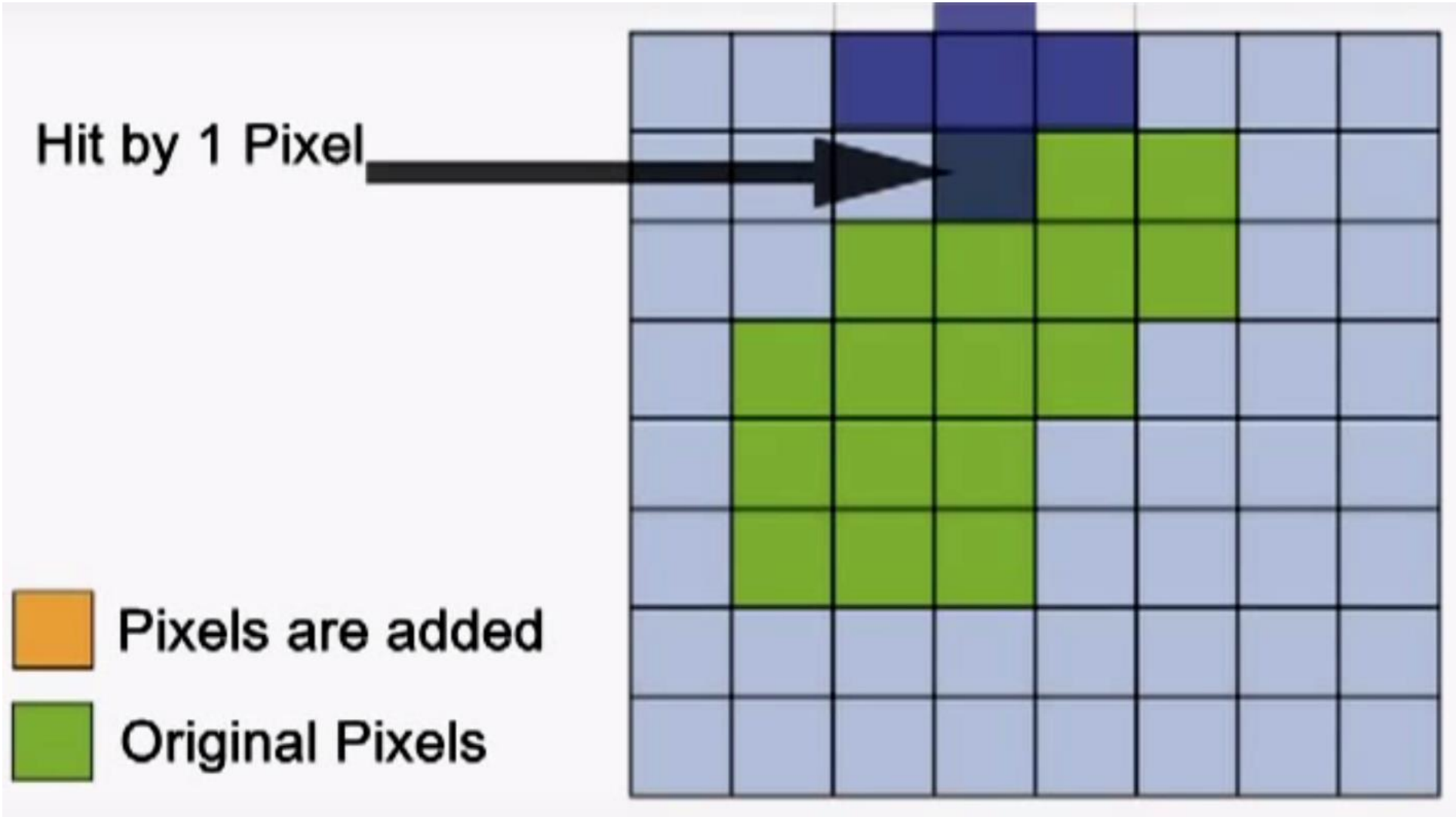
Structuring Element

Pixels are added

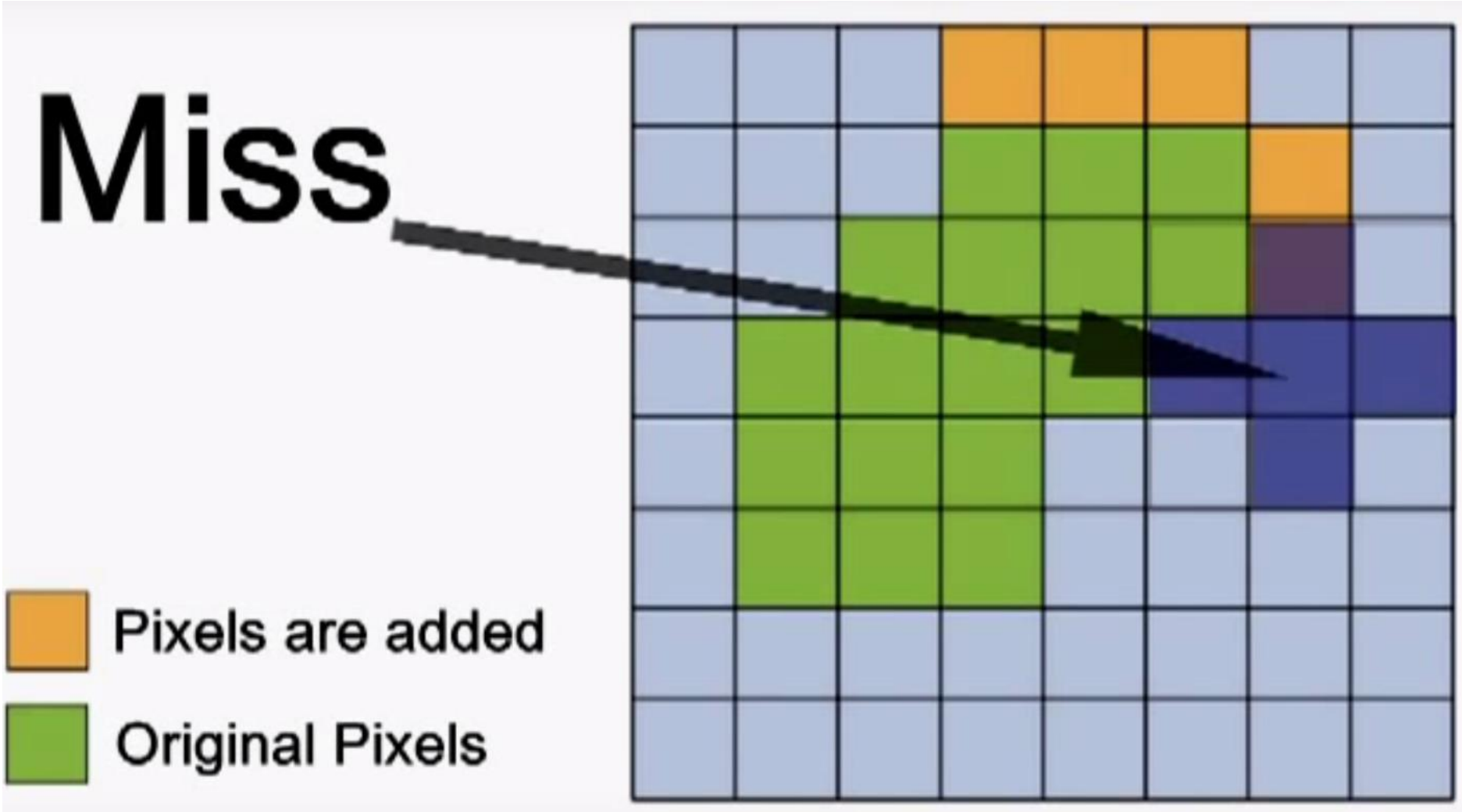
Original Pixels



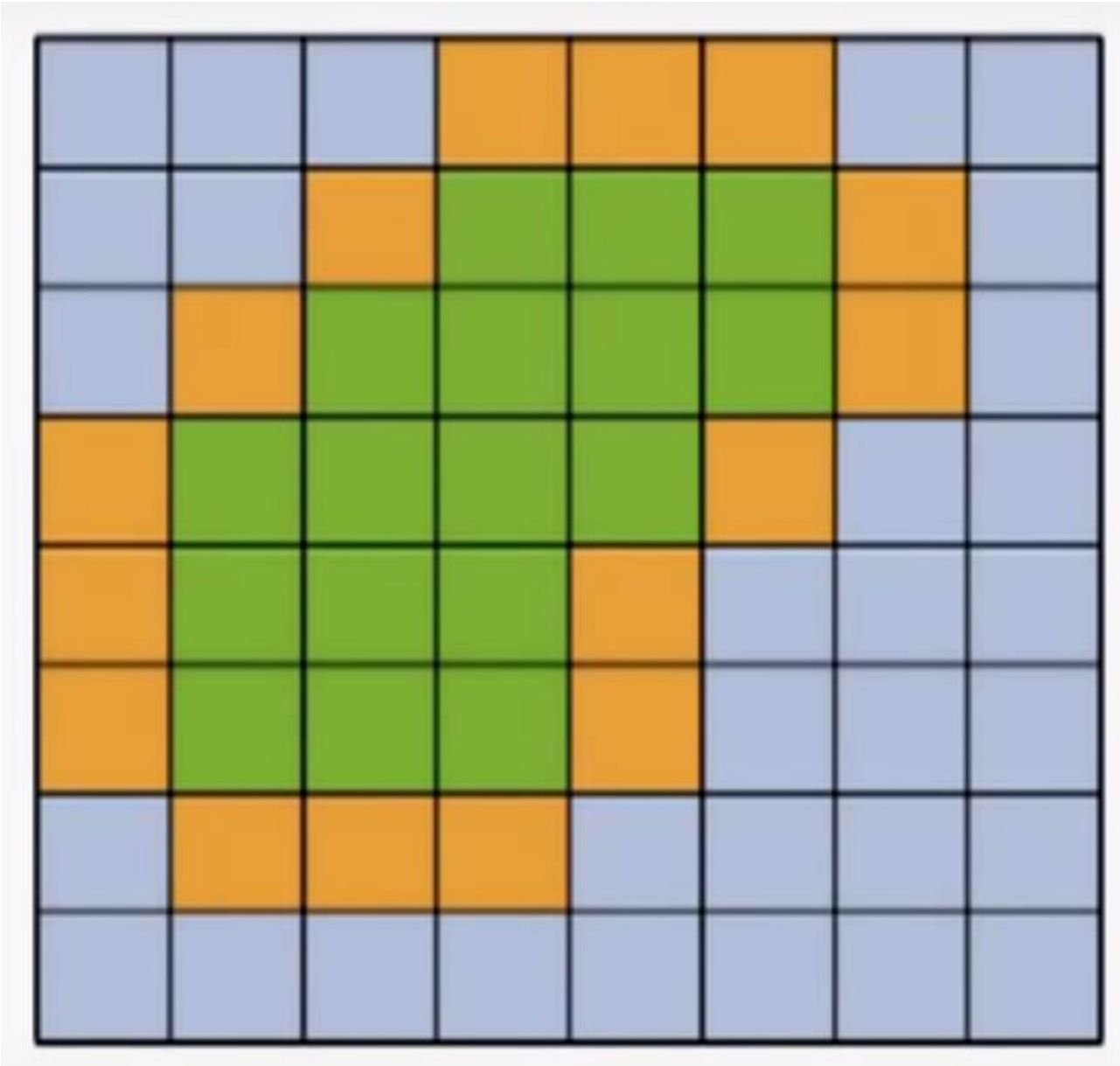
A Dilation Example



A Dilation Example

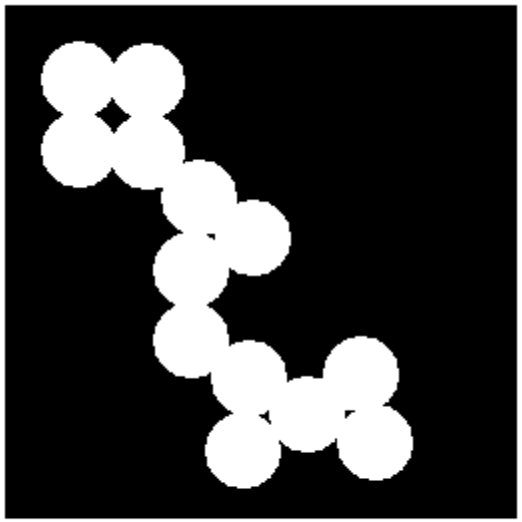


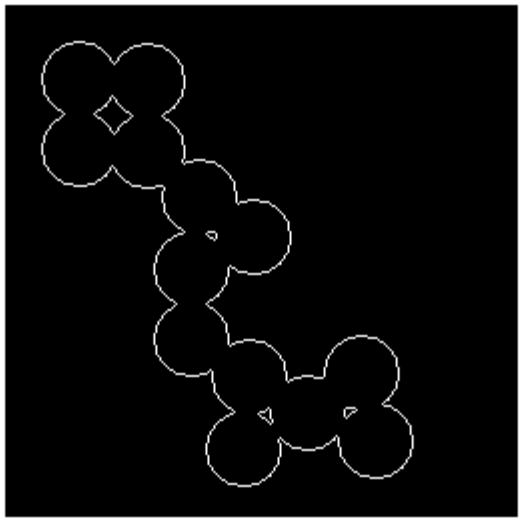
A Dilation Example

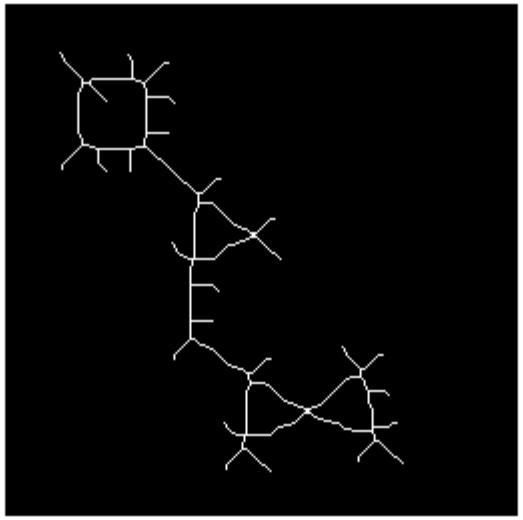


Demos: Other Mathematical Morphologies

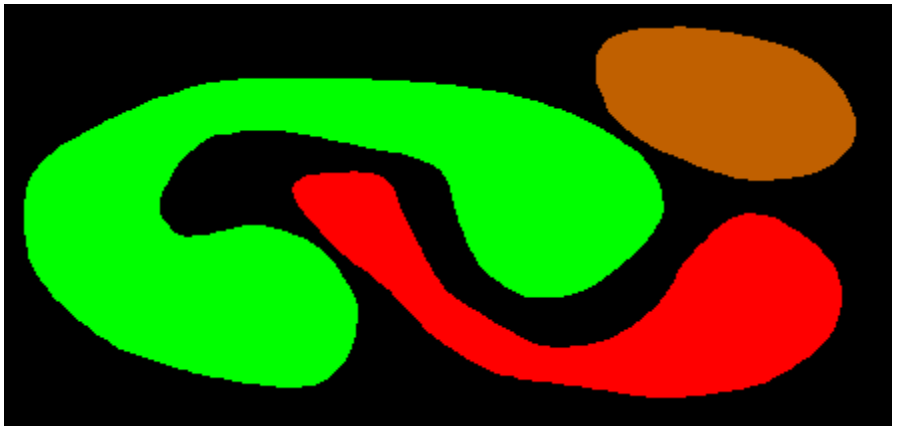
- **Matlab: bwmorph operator**



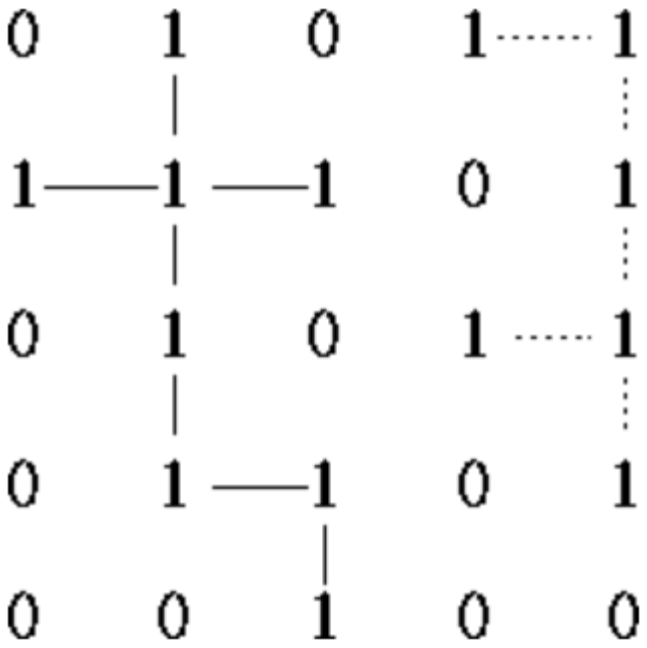




Connected Component Analysis



- Pixel connectivity
  - 4-neighbor
  - 8-neighbor



Connected Component Analysis (cont.)

- Remove the largest letter from the image on the left