Probabilities, Greyscales, and Histograms: Chapter 3a G&W

Ross Whitaker

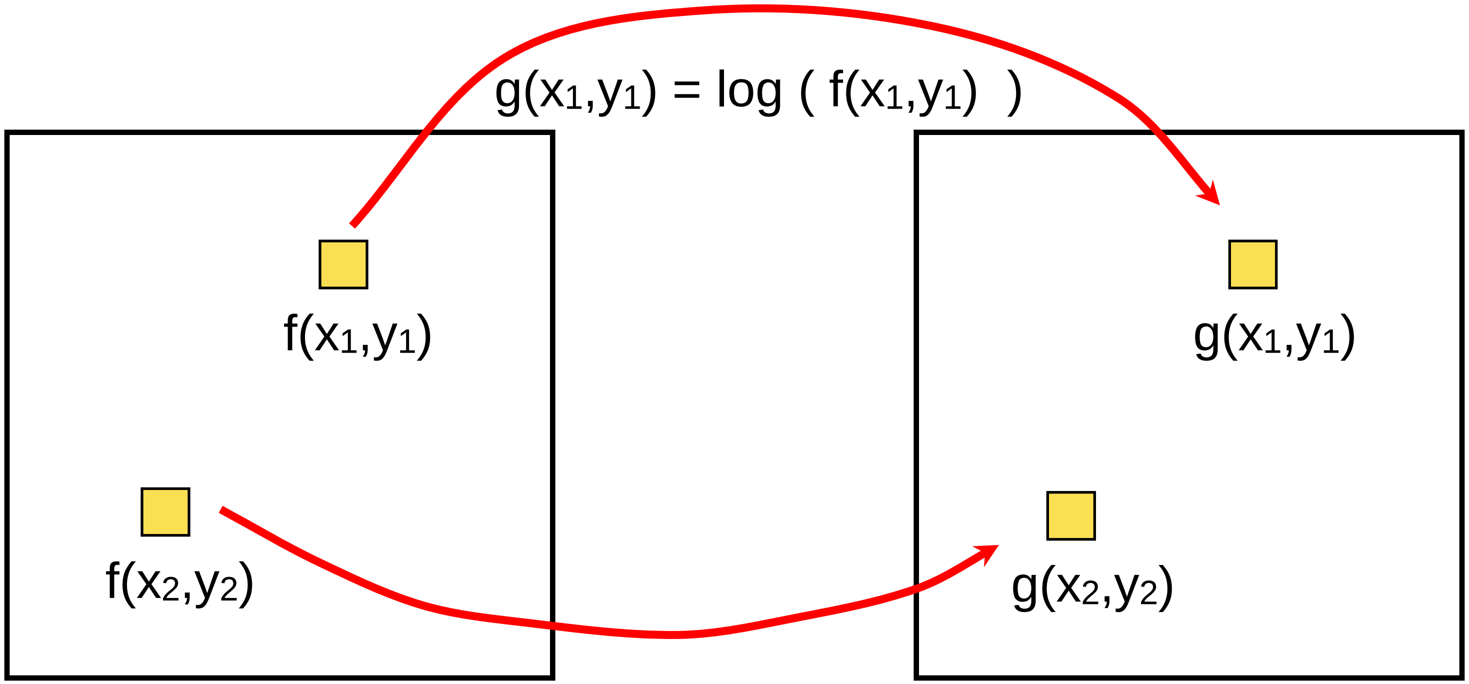
(slightly modified by Guido Gerig)

SCI Institute, School of Computing

University of Utah

Intensity transformation example

$g(x,y) = \log(f(x,y))$



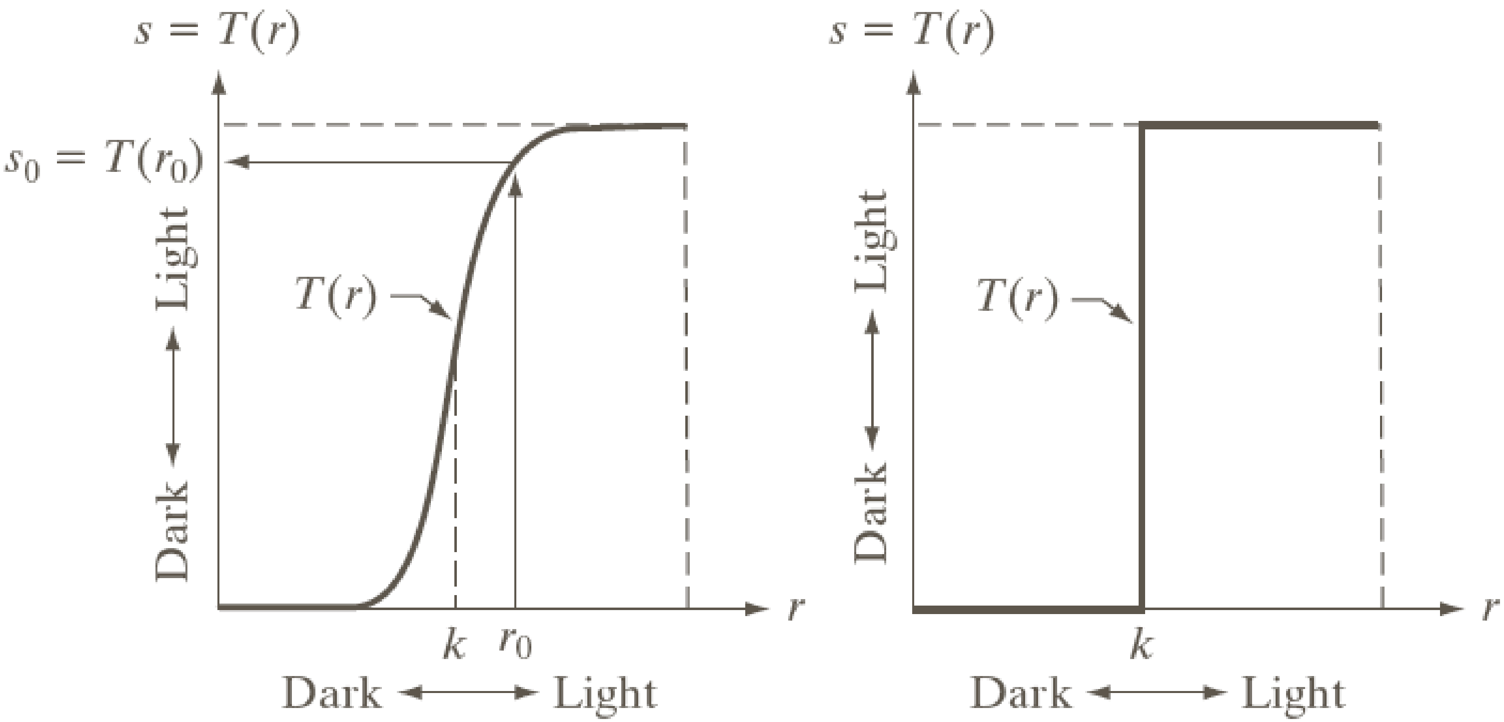
$g(x_2,y_2) = \log(\, f(x_2,y_2) \,)$

- We can **drop the (x,y)** and represent this kind of filter as an intensity transformation s=T(r). In this case s=log(r)

-s: output intensity

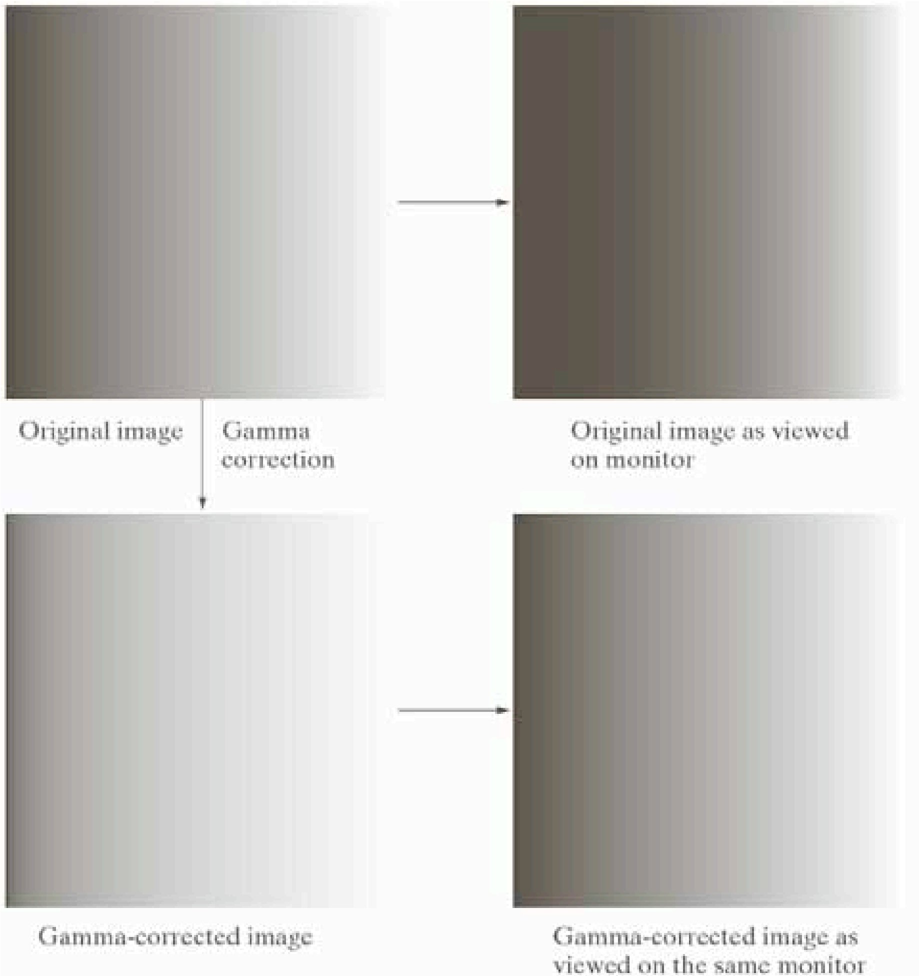
-r: input intensity

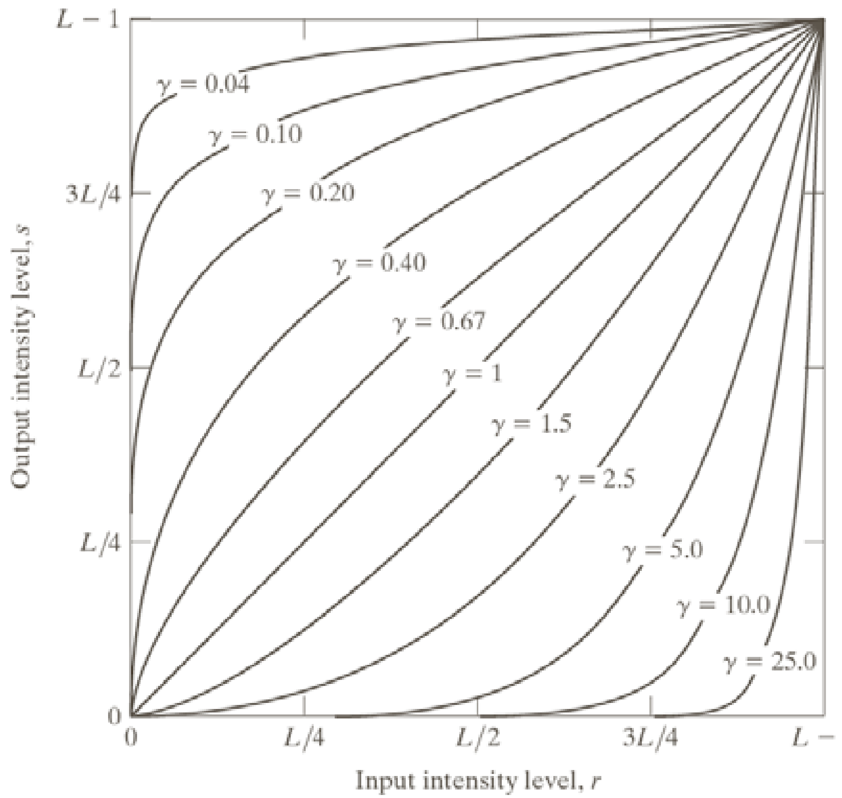
Intensity transformation



$$s = T(r)$$

Gamma correction





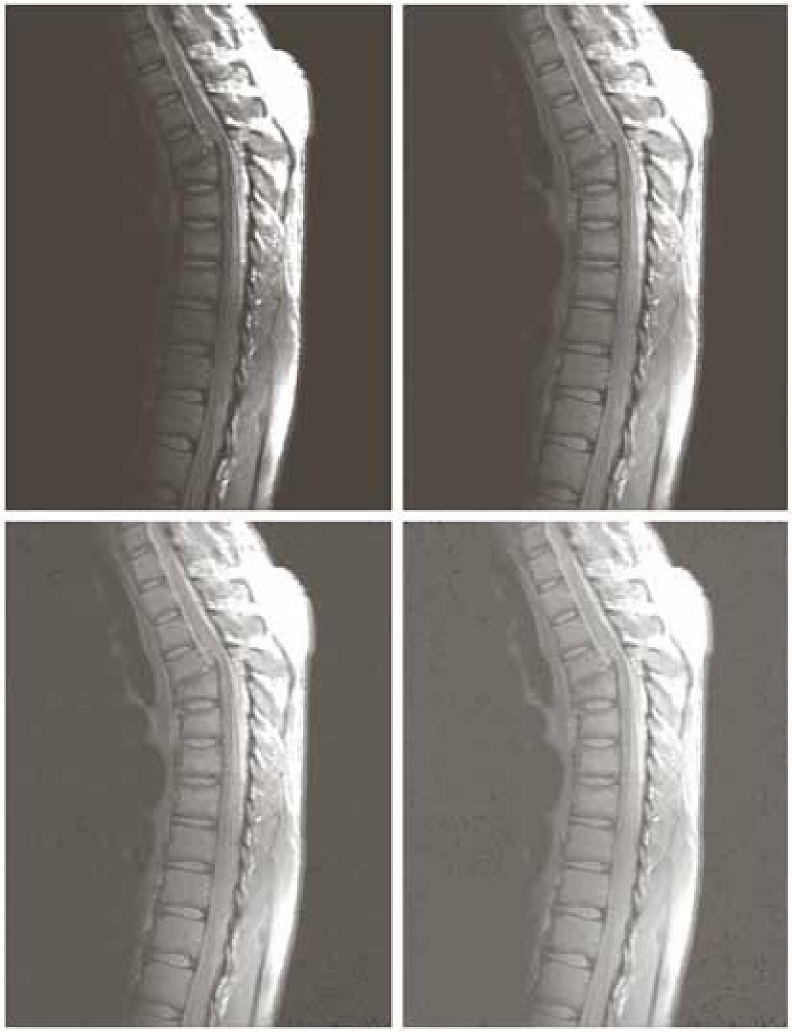
$$s = cr^{g}$$

Gamma transformations

a	b
c	d

FIGURE 3.9 (a) Aerial image. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 3.0, 4.0,$ and 5.0, respectively. (Original image for this example courtesy of NASA.)

Gamma transformations

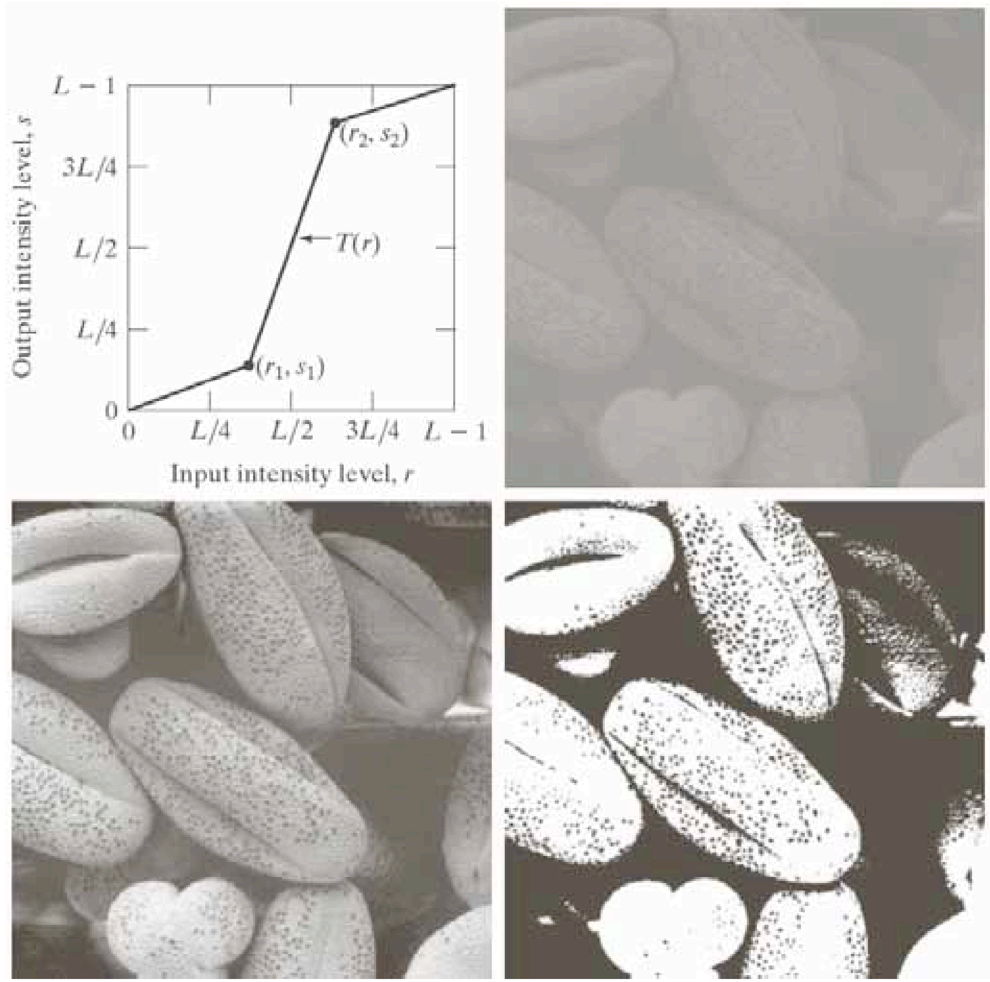


a b
c d

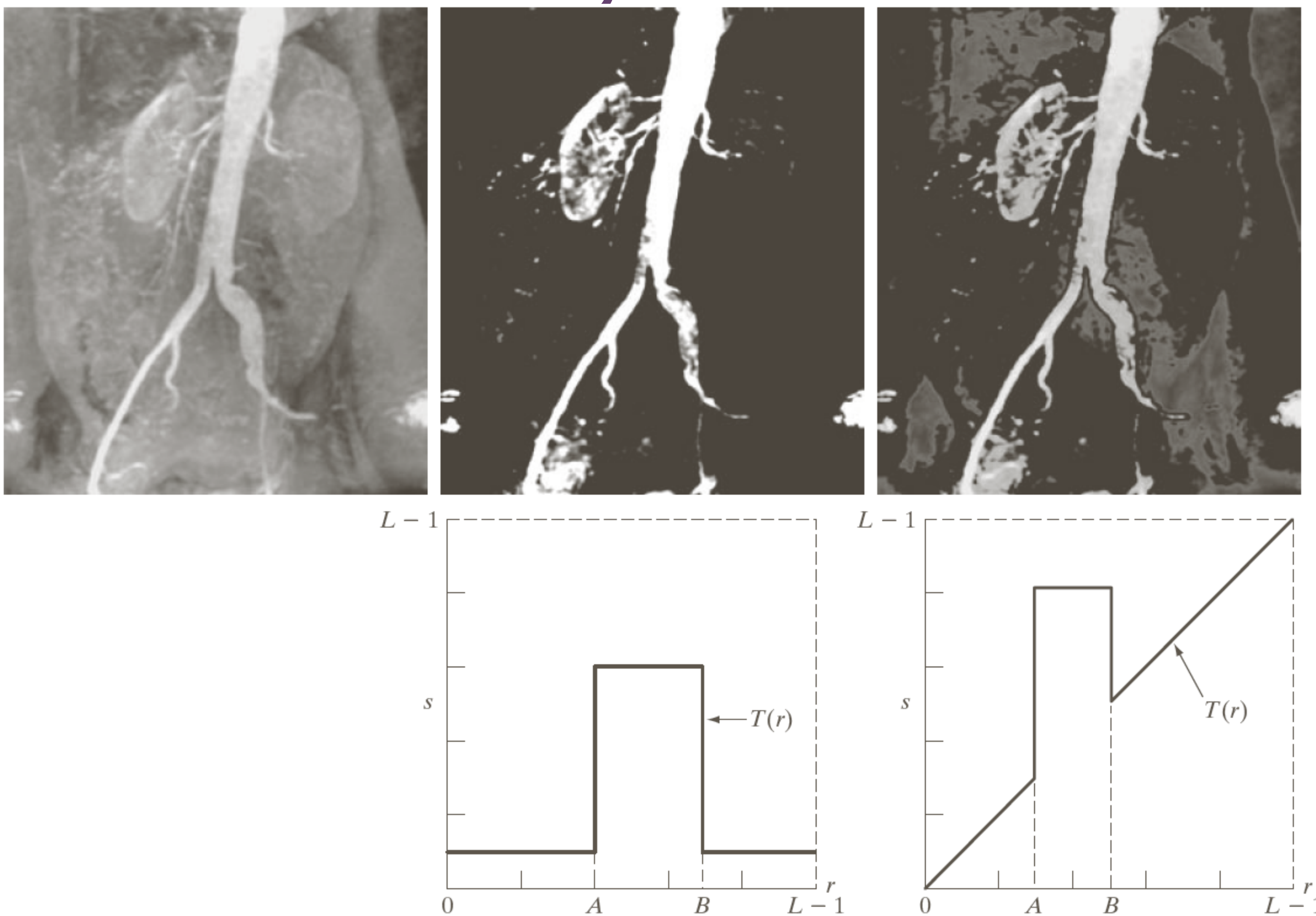
FIGURE 3.8 (a) Magnetic resonance image (MRI) of a fractured human spine. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4,$ and 0.3, respectively. (Original image courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

Piecewise linear intensity transformation

- More control
- But also more parameters for user to specify
  - Graphical user interface can be useful

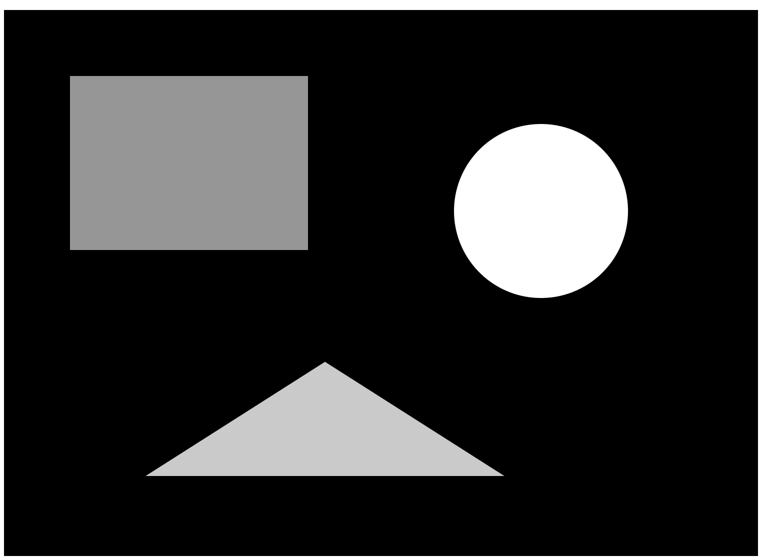


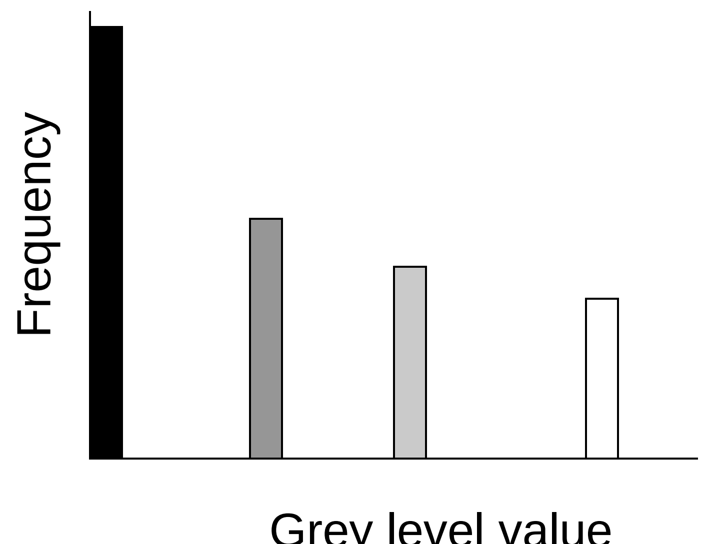
More intensity transformations



Histogram of Image Intensities

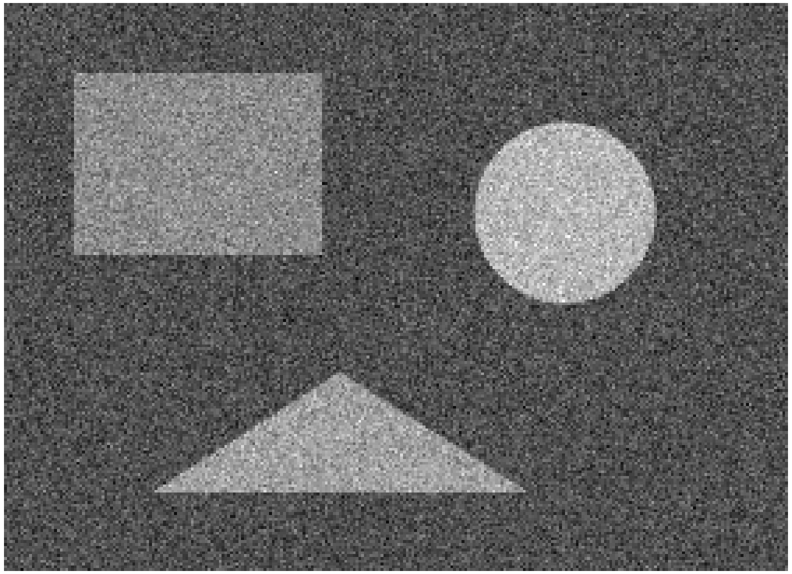
- Create bins of intensities and count number of pixels at each level
  - Normalize or not (absolution vs % frequency)

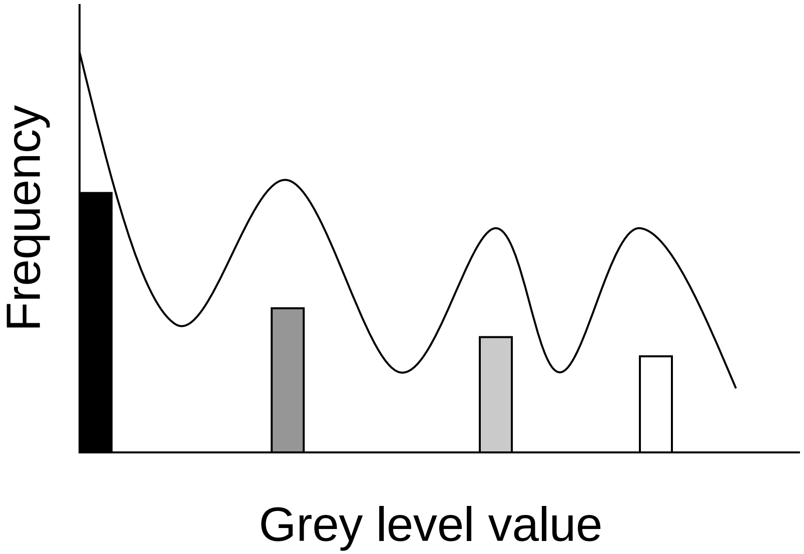




Histograms and Noise

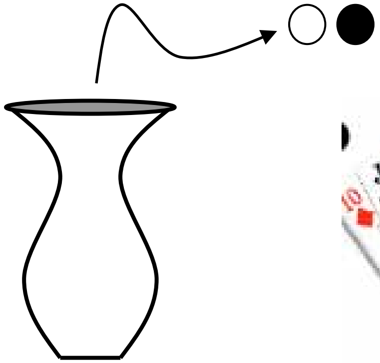
- What happens to the histogram if we add noise?
  - $g(x, y) = f(x, y) + n(x, y)$





Sample Spaces

- S = <u>Set</u> of possible outcomes of a random event
- Toy examples
  - Dice
  - Urn
  - Cards



- Probabilities

$$P(S) = 1 \qquad A \in S \Rightarrow P(A) \geq 0$$

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) \;\text{ where }\; \mathrm{A_i} \cap \mathrm{A_j} = \emptyset$$

$$\cup_{i=1}^{n} A_i = S \Rightarrow \sum_{i=1}^{n} P(A_i) = 1$$

Conditional Probabilities

- Multiple events
  - S2 = SxS Cartesian produce - sets
  - Dice - (2, 4)
  - Urn - (black, black)
- P(A|B) - probability of A in second experiment knowledge of outcome of first experiment
  - This quantifies the effect of the first experiment on the second
- P(A,B) - probability of A in second experiment and B in first experiment
- $P(A,B) = P(A|B)P(B)$

Independence

- P(A|B) = P(A)
  - The outcome of one experiment does not affect the other
- Independence -> P(A,B) = P(A)P(B)
- Dice
  - Each roll is unaffected by the previous (or history)
- Urn
  - Independence -> put the stone back after each experiment
- Cards
  - Put each card back after it is picked

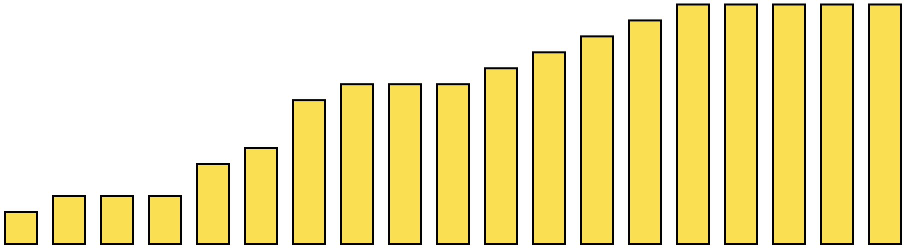
Random Variable (RV)

- Variable (number) associated with the outcome of an random experiment
- Dice
  - E.g. Assign 1-6 to the faces of dice
- Urn
  - Assign 0 to black and 1 to white (or vise versa)
- Cards
  - Lots of different schemes - depends on application
- A function of a random variable is also a random variable

Cumulative Distribution Function (cdf)

- F(x), where x is a RV
- F(-infty) = 0, F(infty) = 1
- F(x) non decreasing

$$F(x) = \sum_{i=-\infty}^{x} P(i)$$

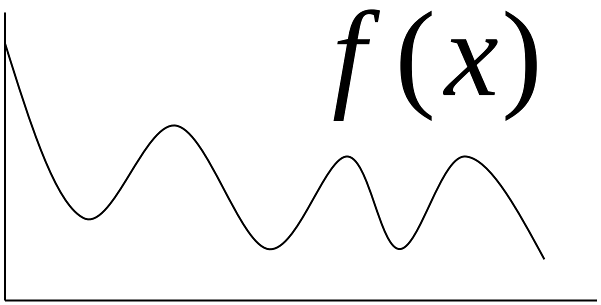


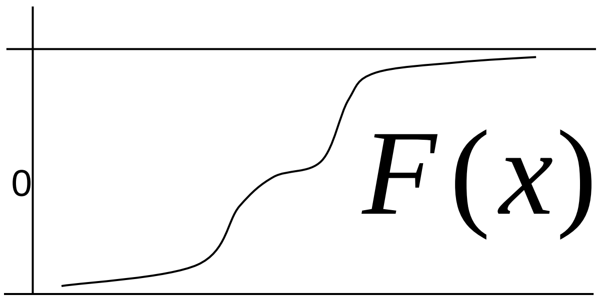
Continuous Random Variables

- f(x) is pdf (normalized to 1)
- F(x) – cdf continuous
  - – –> x is a continuous RV

$$F(x) = \int_{-\infty}^{x} f(q)dq$$

$$f(x) = \left.\frac{dF(q)}{dq}\right|_{x} = F'(x)$$





Probability Density Functions

- f(x) is called a probability density function (pdf)

$$\int_{-\infty}^{\infty} f(x) = 1 \qquad f(x) \geq 0 \;\; \forall \;\; x$$

- A probability density is not the same as a probability
- The probability of a specific value as an outcome of continuous experiment is (generally) zero
  - To get meaningful numbers you must specify a range

$$P(a \leq x \leq b) = \int_{a}^{b} f(q)dq = F(b) - F(a)$$

Expected Value of a RV

$$E[x] = \sum_{i=-\infty}^{\infty} i\ p(i)$$

$$E[x] = \int_{-\infty}^{\infty} q\ f(q)\ dq$$

- Expectation is linear
  - E[ax] = aE[x] for a scalar (not random)
  - E[x + y] = E[x] + E[y]
- Other properties
  - E[z] = z ——— if z is not random

Mean of a PDF

- Mean: E[x] = m
  - also called "m"
  - The mean is <u>not a random variable</u>–it is a fixed value for any PDF
- Variance: $E[(x - m)^2] = E[x^2] - 2E[mx] + E[m^2] = E[x^2] - m^2 = E[x^2] - E[x]^2$
  - also called "s^2"
  - Standard deviation is s
  - If a distribution has zero mean then: $E[x^2] = s^2$

Sample Mean

- Run an experiments
  - Take N samples from a pdf (RV)
  - Sum them up and divide by N
- Let M be the result of that experiment
  - M is a random variable

$$M = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$E[M] = E[\frac{1}{N} \sum_{i=1}^{N} x_i] = \frac{1}{N} \sum_{i=1}^{N} E[x_i] = m$$

Sample Mean

- How close can we expect to be with a sample mean to the true mean?
- Define a new random variable: D = (M - m)$^{2.}$
  - Assume independence of sampling process

$$D = \frac{1}{N^2}\sum_i x_i \sum_j x_j - \frac{1}{N}2m\sum_i x_i + m^2$$

Independence –> E[xy] = E[x]E[y]

$$\begin{aligned} e[D] &= \tfrac{1}{N^2}E[\textstyle\sum_i x_i \sum_j x_j] - \tfrac{1}{N}2mE[\sum_i x_i] + m^2 \\ &= \tfrac{1}{N^2}E[\textstyle\sum_i x_i \sum_j x_j] - m^2 \end{aligned}$$

Number of terms off diagonal

$$\frac{1}{N^2}E[\sum_i x_i \sum_j x_j] = \frac{1}{N^2}\sum_i E[x_i^2] + \frac{1}{N^2}\sum_i\sum_j E[x_i x_j] = \frac{1}{N}\cancel{\sum_i} E[x^2] + \frac{N(N-1)}{N^2}m^2$$

$$E[D] = \frac{1}{N}E[x^2] + \frac{N(N-1)}{N^2}m^2 - \frac{N^2}{N^2}m^2 = \frac{1}{N}\left(E[x^2] - m^2\right) = \frac{1}{N}\sigma^2$$

Root mean squared difference between true mean and sample mean is stdev/sqrt(N).
As number of samples –> infty, sample mean –> true mean.

Application: Noisy Images

- Imagine N images of the same scene with random, independent, zero-mean noise added to each one
  - Nuclear medicine–radioactive events are random
  - Noise in sensors/electronics
- Pixel value is s+n

True pixel value

Random noise:

•Independent from one image to the next

•Variance = S

Application: Noisy Images

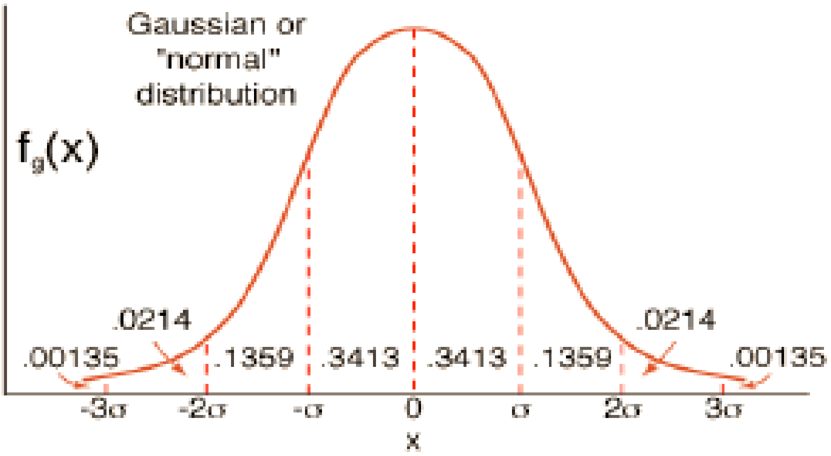
- If you take multiple images of the same scene you have
  - $s_i = s + n_i$
  - $S = (1/N)\ Ss_i = s + (1/N)\ Sn_i$
  - $E[(S - s)^2] = (1/N)\ E[n_i^{\ 2}] = (1/N)\ E[n_i^{\ 2}] - (1/N)\ E[n_i]^2 = (1/N)s^{\ 2}$
  - Expected **root mean squared error** is s/sqrt(N)

Zero mean

- Application:
  - Digital cameras with large gain (high ISO, light sensitivity)
  - Not necessarily random from one image to next
    - Sensors CCD irregularity
  - How would this principle apply

Gaussian Distribution

- “Normal” or “bell curve”
- Two parameters: m- mean, s – standard deviation



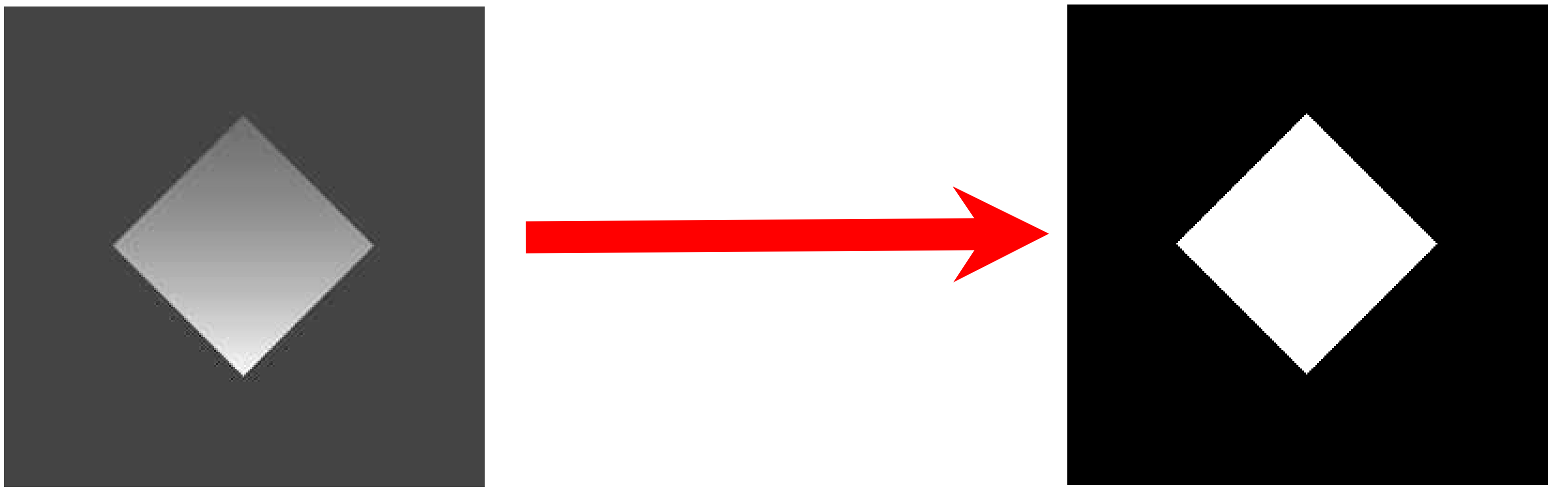
$$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Gaussian Properties

- Best fitting Guassian to some data is gotten by mean and standard deviation of the samples
- Occurrence
  - Central limit theorem: result from lots of random variables
  - Nature (approximate)
    - Measurement error, physical characteristic, physical phenomenon
    - Diffusion of heat or chemicals

What is image segmentation?

- Image segmentation is the process of subdividing an image into its constituent regions or objects.
- Example segmentation with two regions:

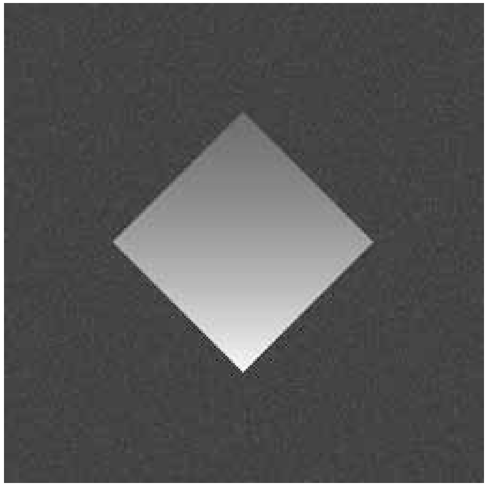


Input image
intensities 0-255

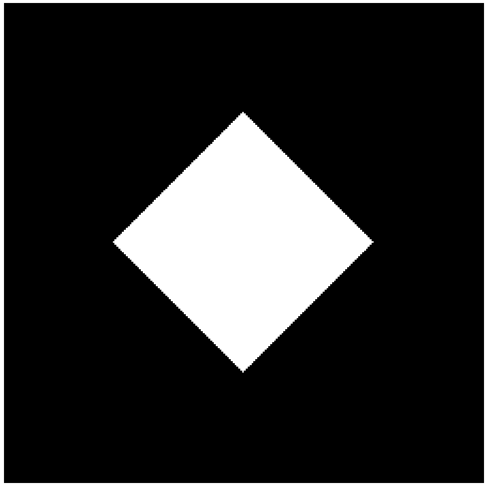
Segmentation output
0 (background)
1 (foreground)

Thresholding

$$g(x,y) = \begin{cases} 1 & if \quad f(x,y) > T \\ 0 & if \quad f(x,y) \leq T \end{cases}$$



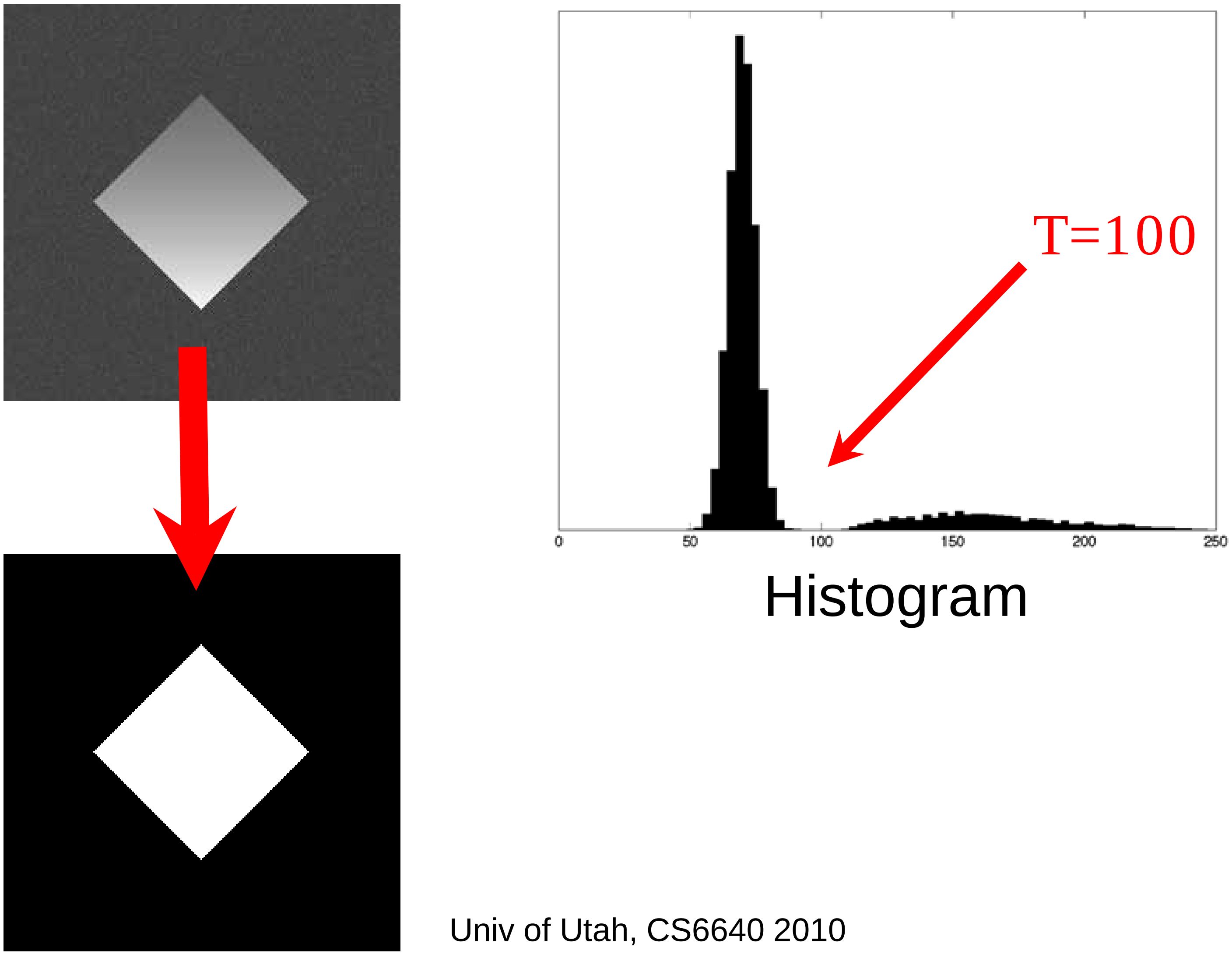
Input image f(x,y)
intensities 0-255



Segmentation output g(x,y)
0 (background)
1 (foreground)

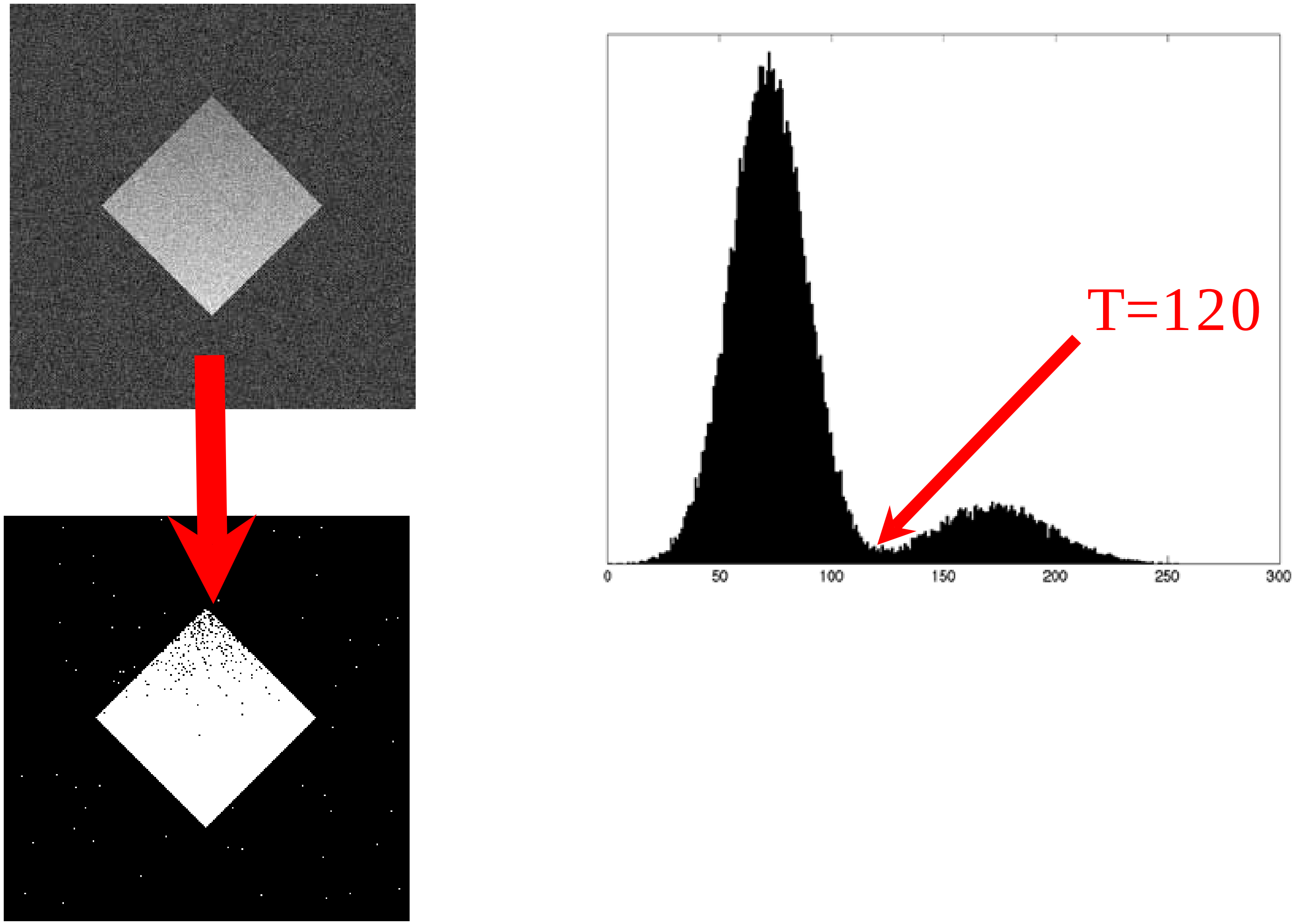
- How can we choose T?
  - Trial and error
  - Use the histogram of f(x,y)

Choosing a threshold

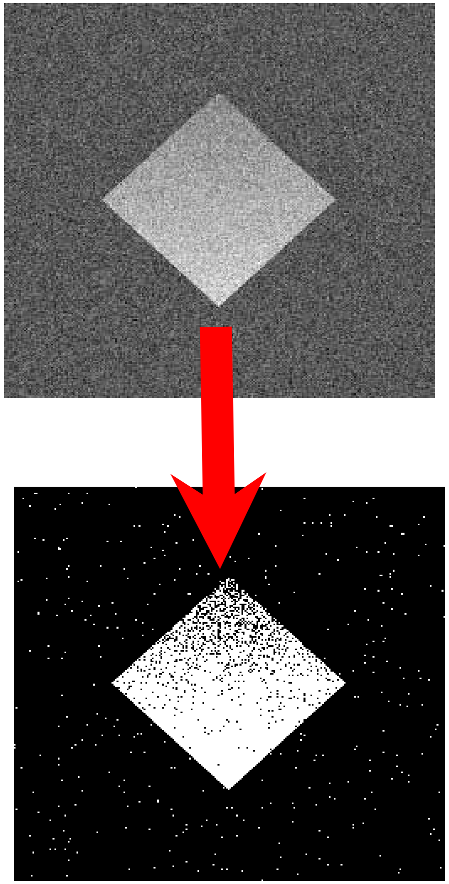


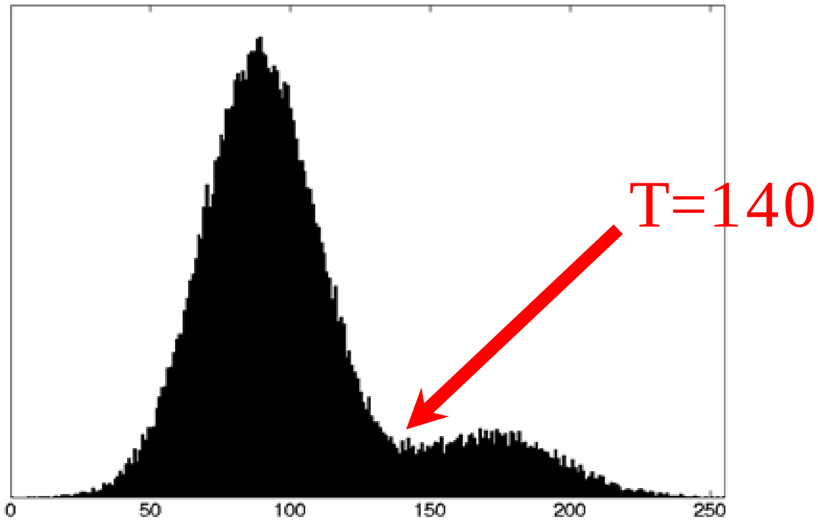
Histogram

Role of noise



Low signal-to-noise ratio

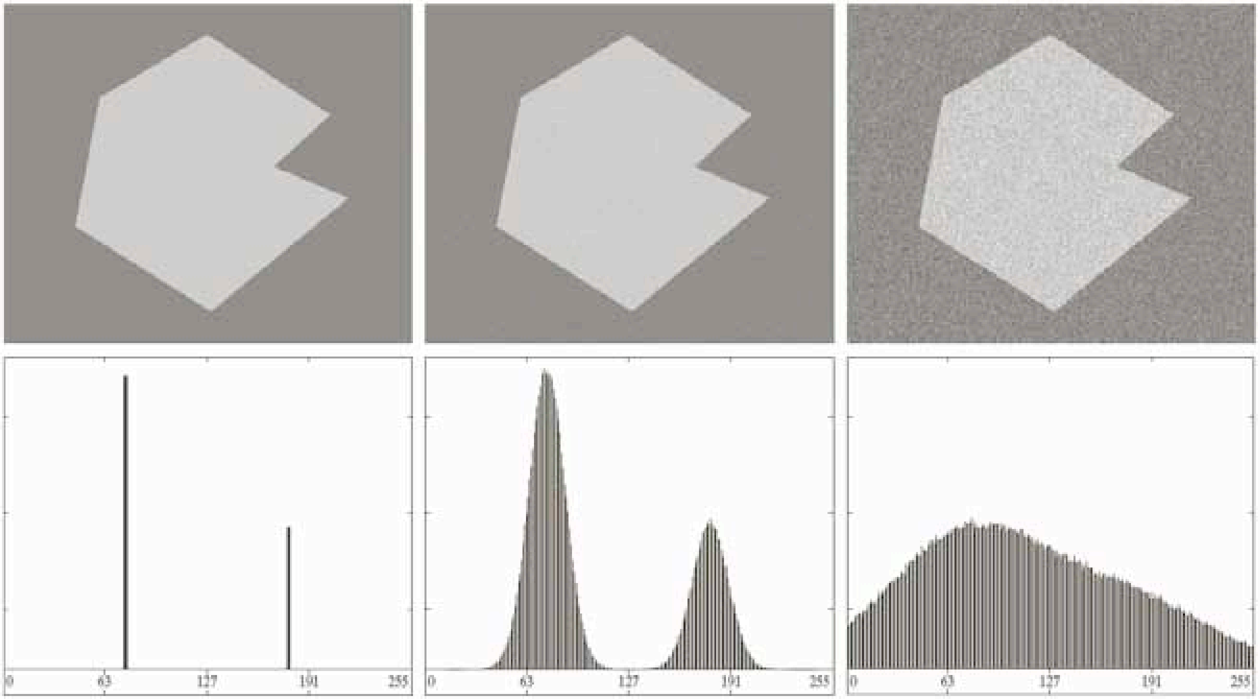




Effect of noise on image histogram

Images

Histograms

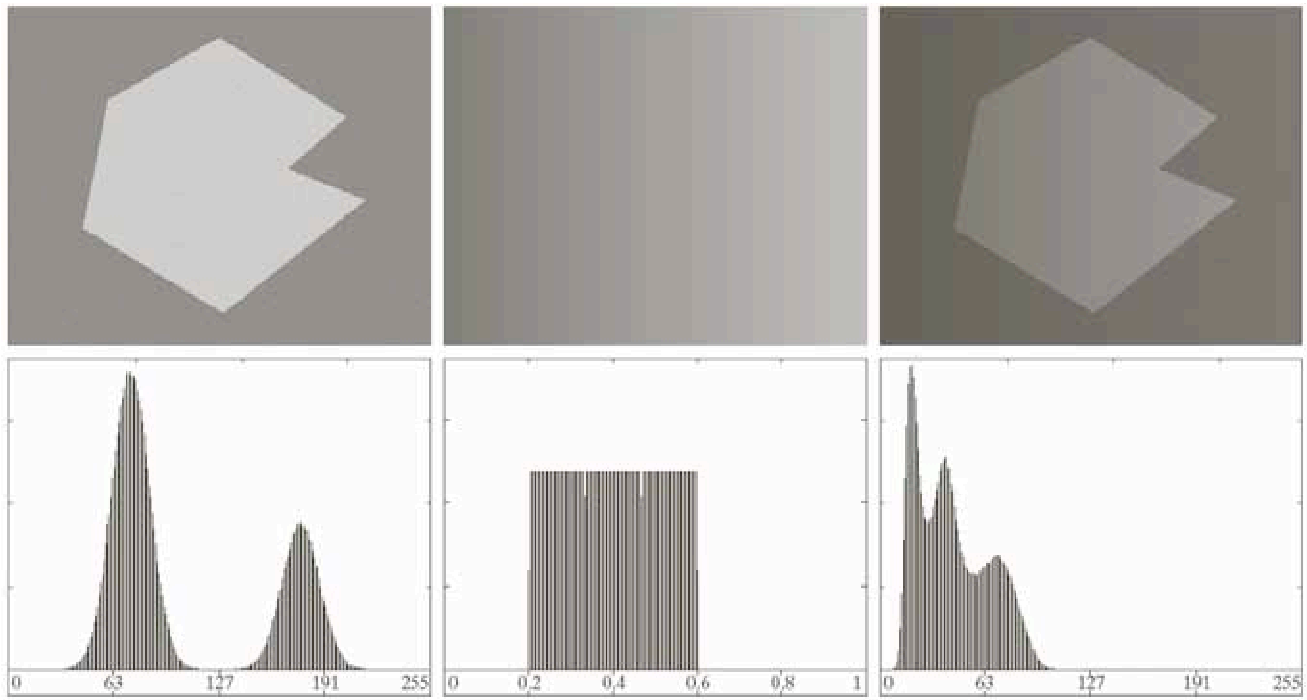


No noise With noise More noise

Effect of illumination on histogram

Images

Histograms



f x g = h

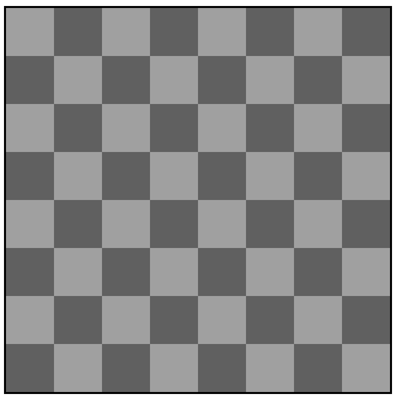
Original image

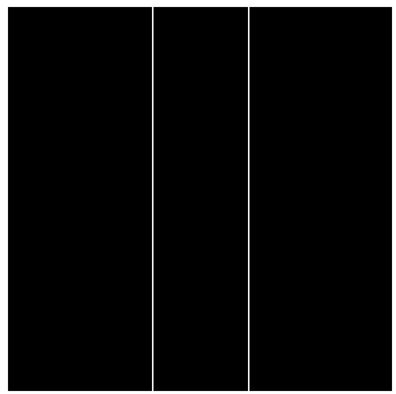
Illumination image

Final image

Histogram of Pixel Intensity Distribution

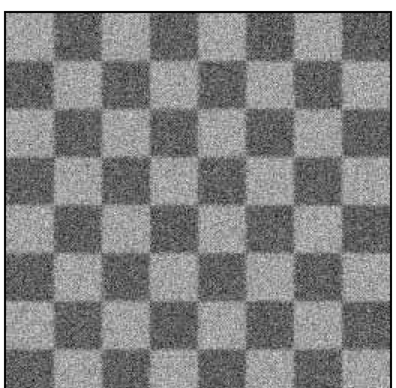
Histogram: Distribution of intensity values $p(v)$
(count #pixels for each intensity level)

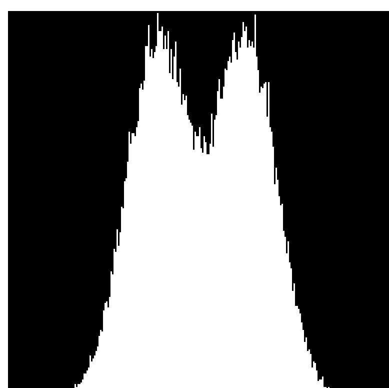




Checkerboard with values 96 and 160.

Histogram:
- horizontal: intensity
- vertical: # pixels





Checkerboard with additive Gaussian noise (sigma 20).

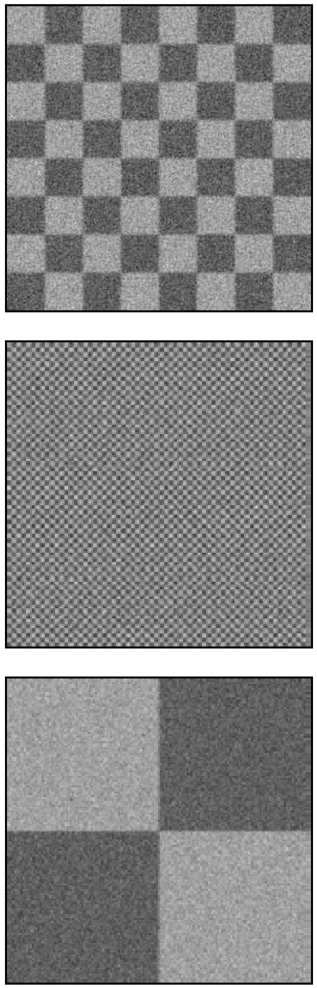
Regions: 50%b,50%w

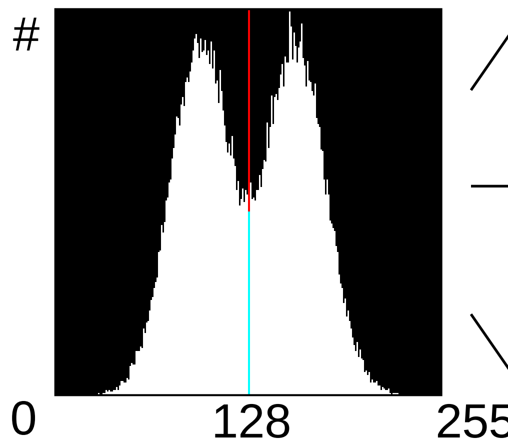
Classification by Thresholding

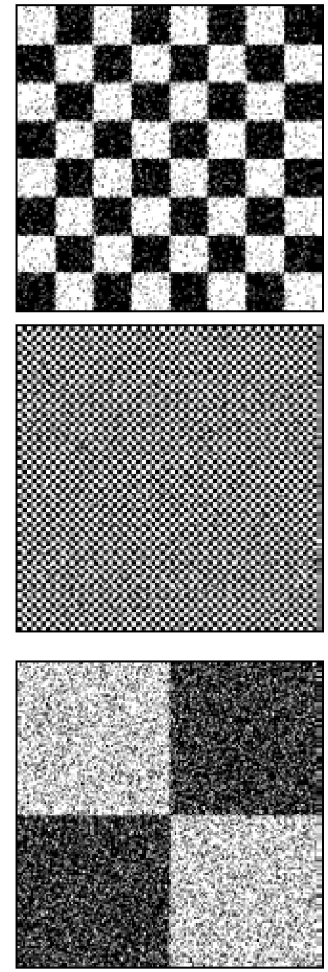
50%b
50%w

50%b
50%w

50%b
50%w







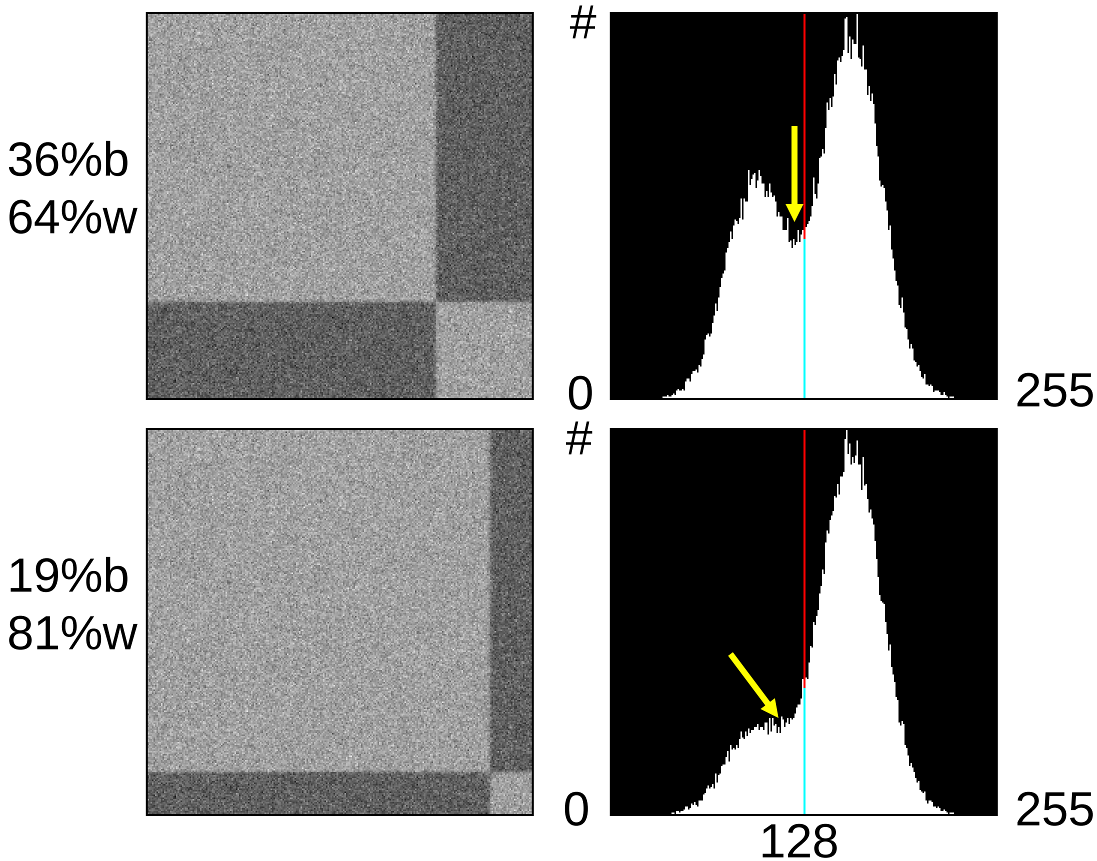
Optimal Threshold:

$$\frac{m_1 + m_2}{2} = \frac{96+160}{2} = 128$$

Important!

- Histogram does not represent image structure such as regions and shapes, but only distribution of intensity values
- Many images share the same histogram

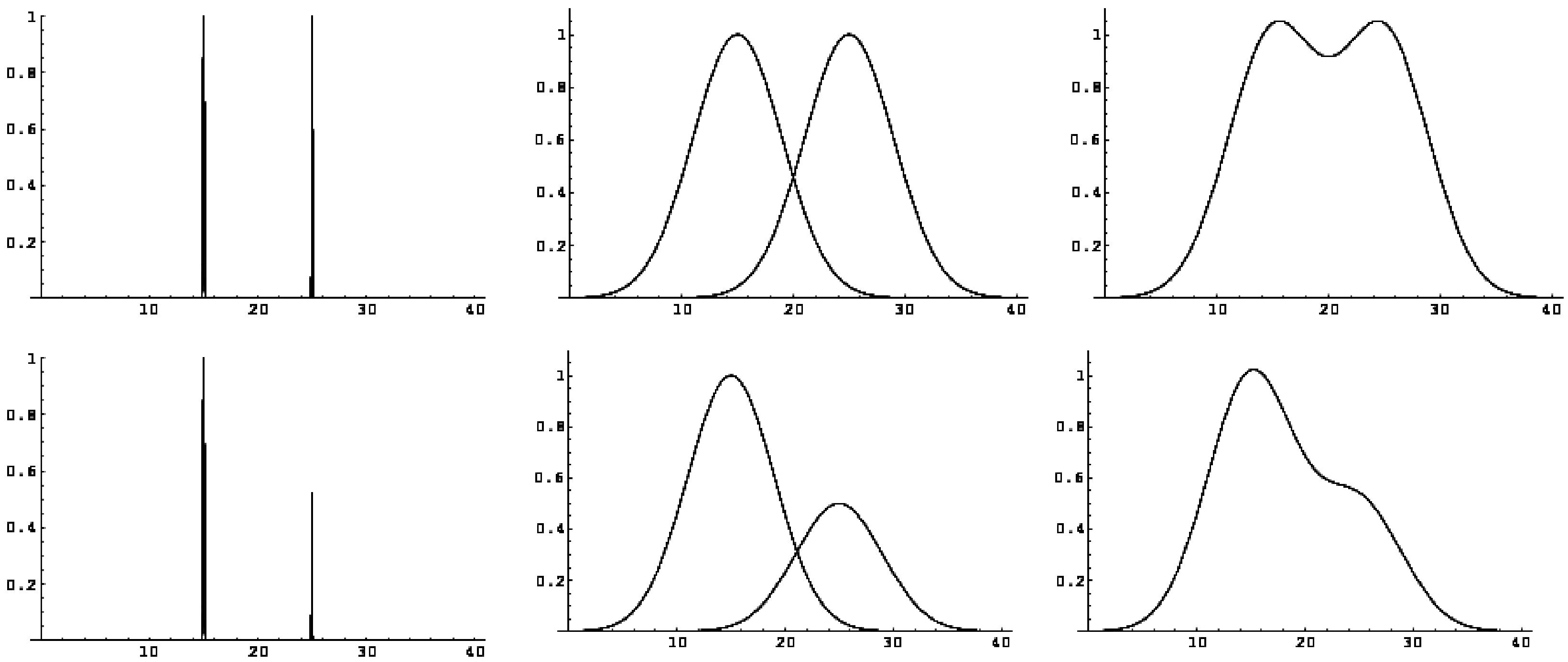
Is the histogram suggesting the right threshold?



Proportions of bright and dark regions are different Þ Peak presenting bright regions becomes dominant.

Threshold value 128 does not match with valley in distribution.

Histogram as Superposition of PDF's
(probability density functions)



Regions with
2 brightness levels,
different proportions

Corruption with
Gaussian noise,
individual distributions

Histogram:
Superposition of
distributions

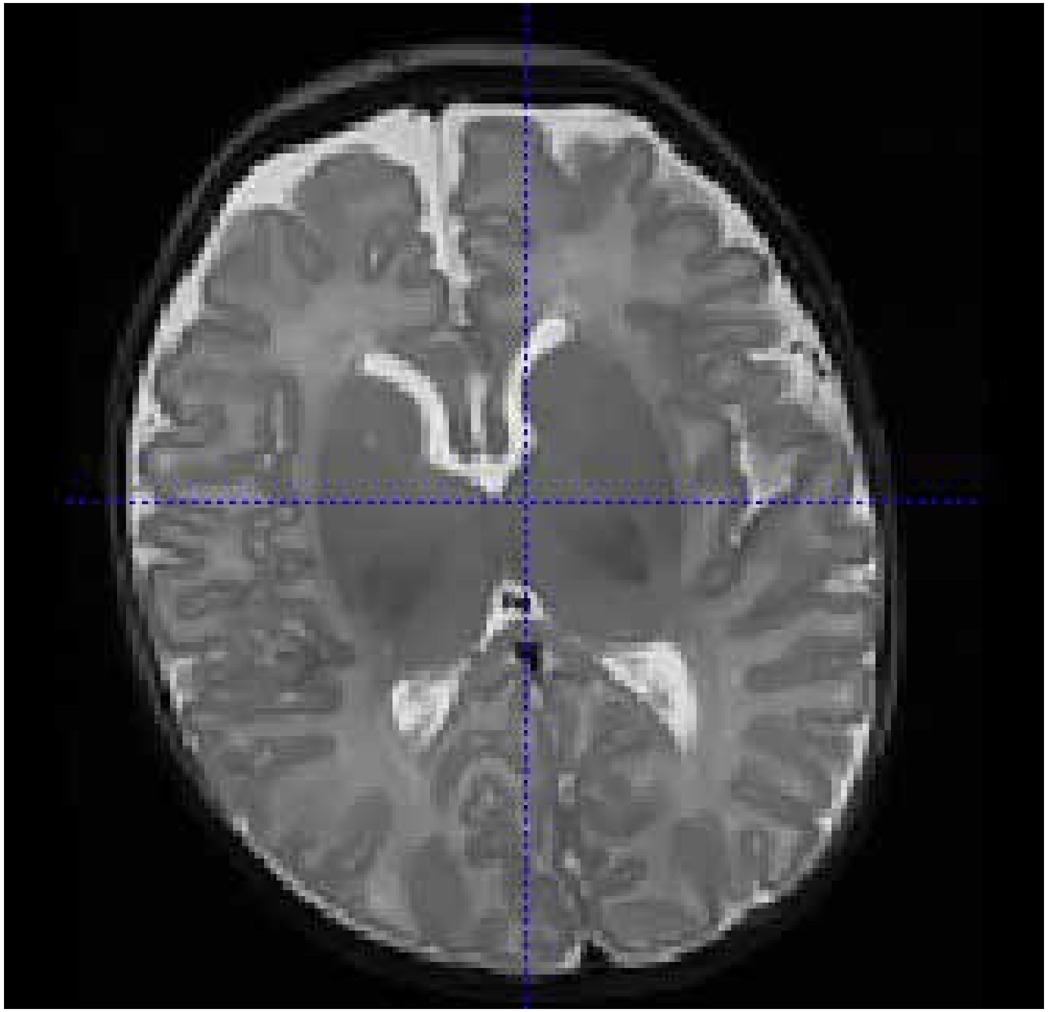
Gaussian Mixture Model

$$hist = a_1 G(\mu_1, \sigma_1) + a_1 G(\mu_1, \sigma_2)$$

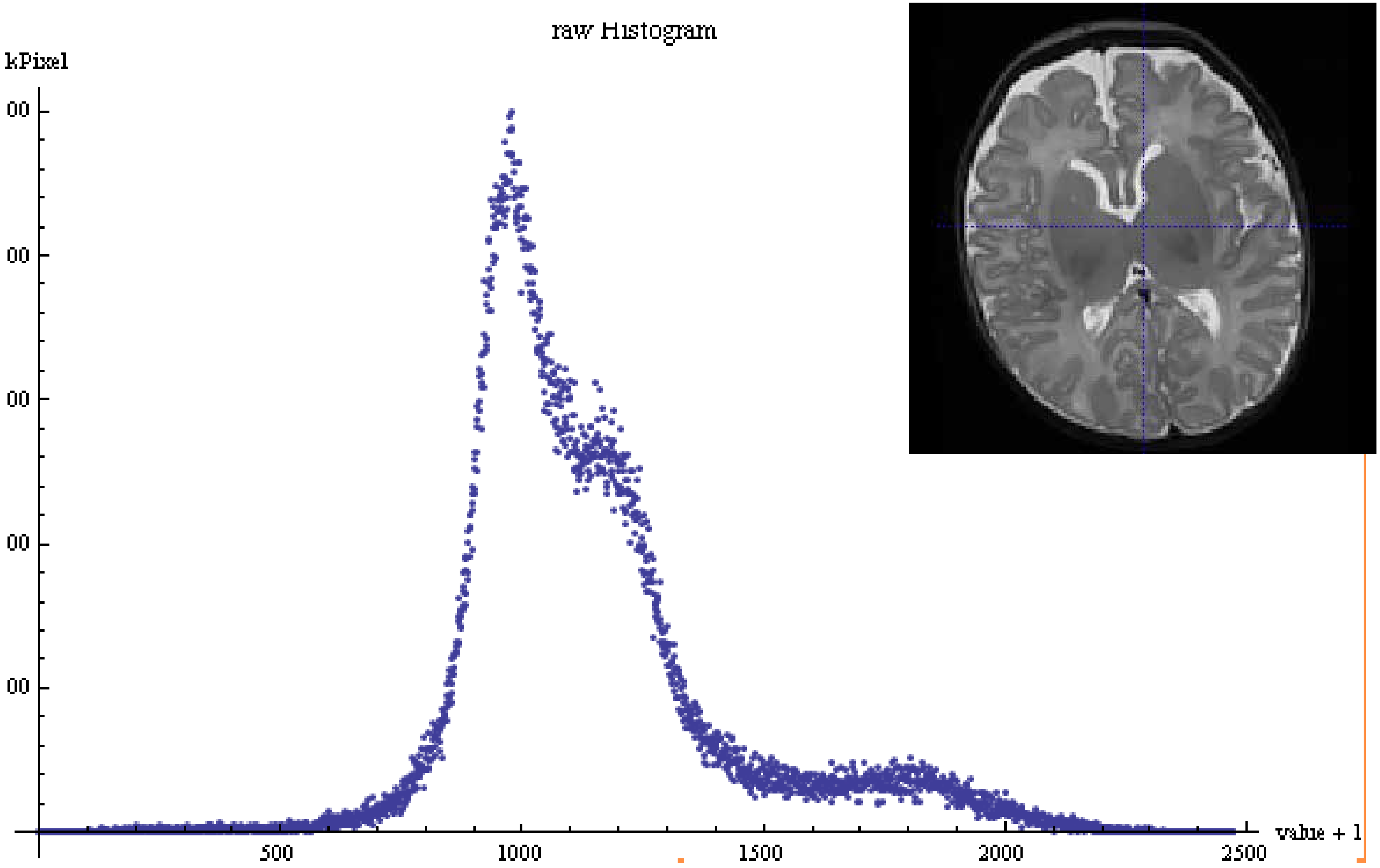
$$more\ general\ with\ k\ classes:$$

$$hist = \sum_k a_k G(\mu_k, \sigma_k)$$

Example: MRI

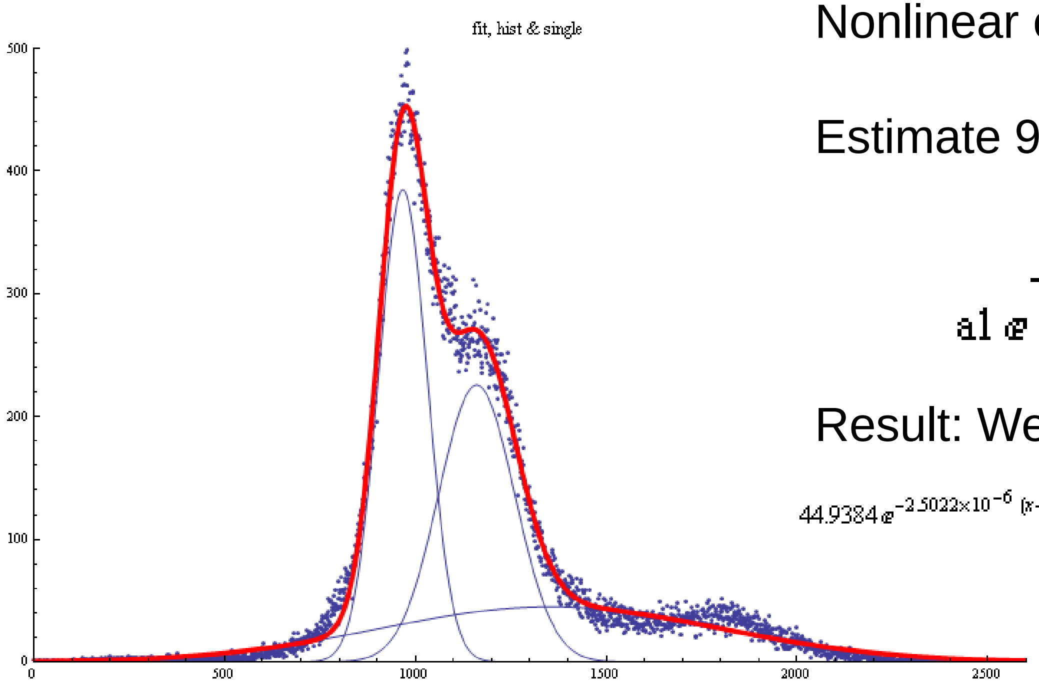


Example: MRI



raw Histogram
kPixel
00
00
00
00
00
500
1000
1500
2000
2500
value + 1

Fit with 3 weighted Gaussians



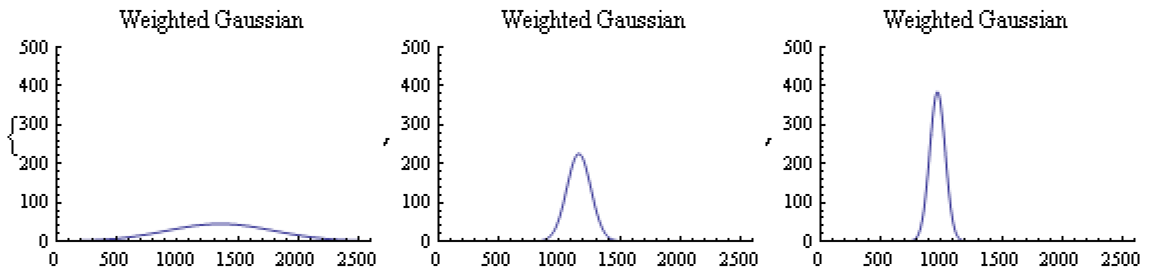
Nonlinear optimization

Estimate 9 parameters for best fit:

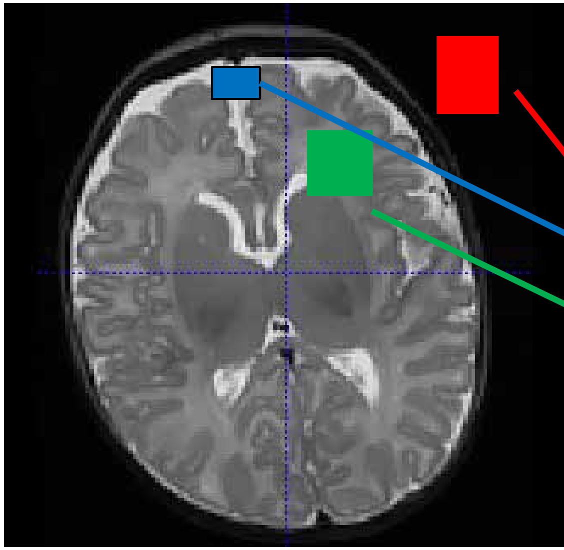
$$a1\, e^{-\frac{(x-my1)^2}{2\, sig1^2}} + a2\, e^{-\frac{(x-my2)^2}{2\, sig2^2}} + a3\, e^{-\frac{(x-my3)^2}{2\, sig3^2}}$$

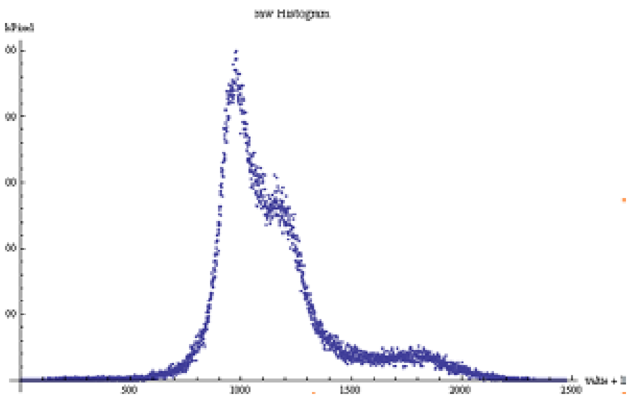
Result: Weighted sum of Gaussians (pdf’s):

$$44.9384\, e^{-2.5022\times 10^{-6}\,(x-1353.63)^2} + 225.575\, e^{-0.0000503733\,(x-1160.5)^2} + 384.58\, e^{-0.000122748\,(x-967.112)^2}$$

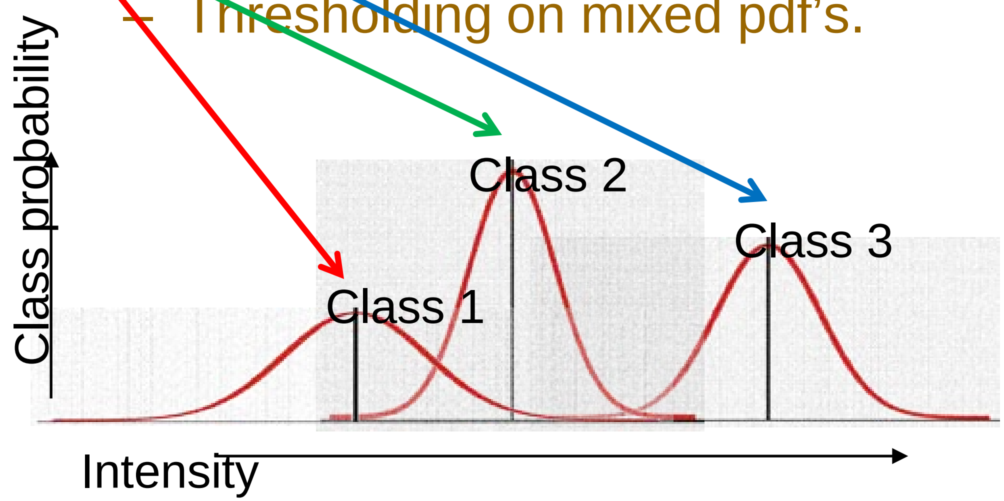


Segmentation: Learning pdf's





- We learned: histogram can be misleading due to different size of regions.
- **Solution**:
  - Estimate class-specific pdf's via training (or nonlinear optimization)
  - Thresholding on mixed pdf's.



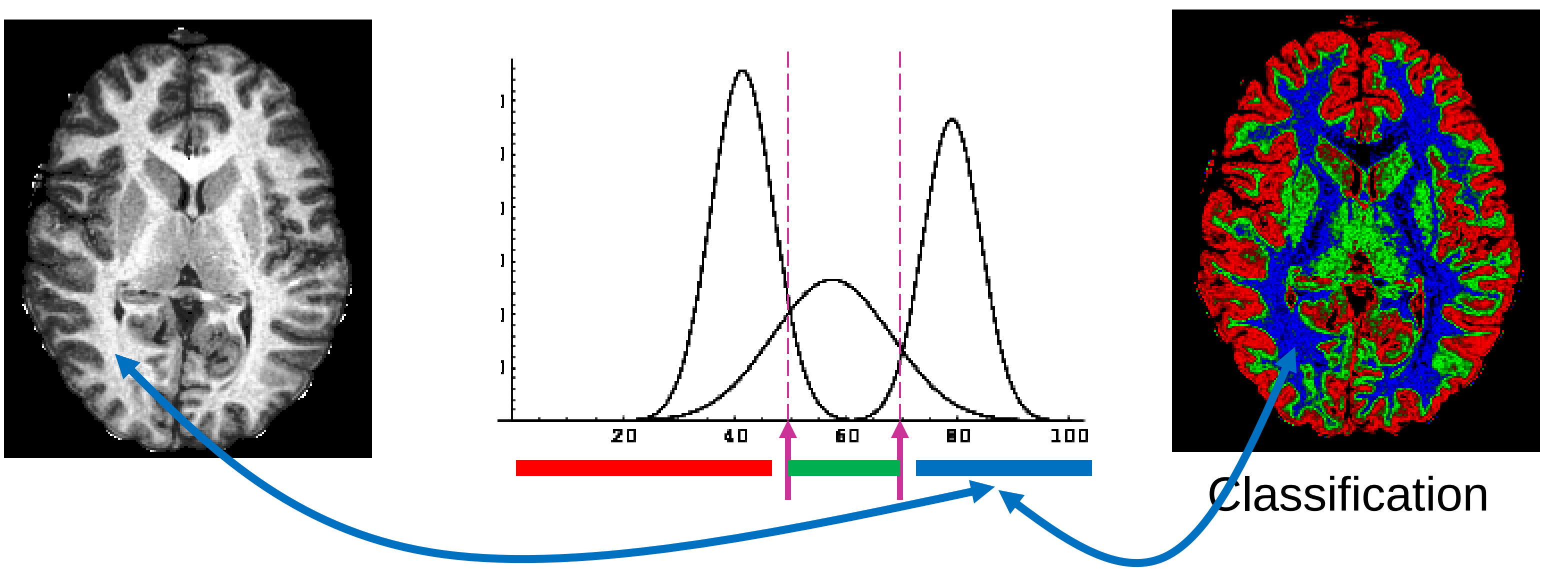
Segmentation: Learning pdf's

$$set\ of\ \ pdf's:$$

$$G_k(m_k, s_k \mid k), \quad (k = 1, \dots, n)$$

$$calculate \quad thresholds$$

$$assign\ \ pixels\ to\ categories$$

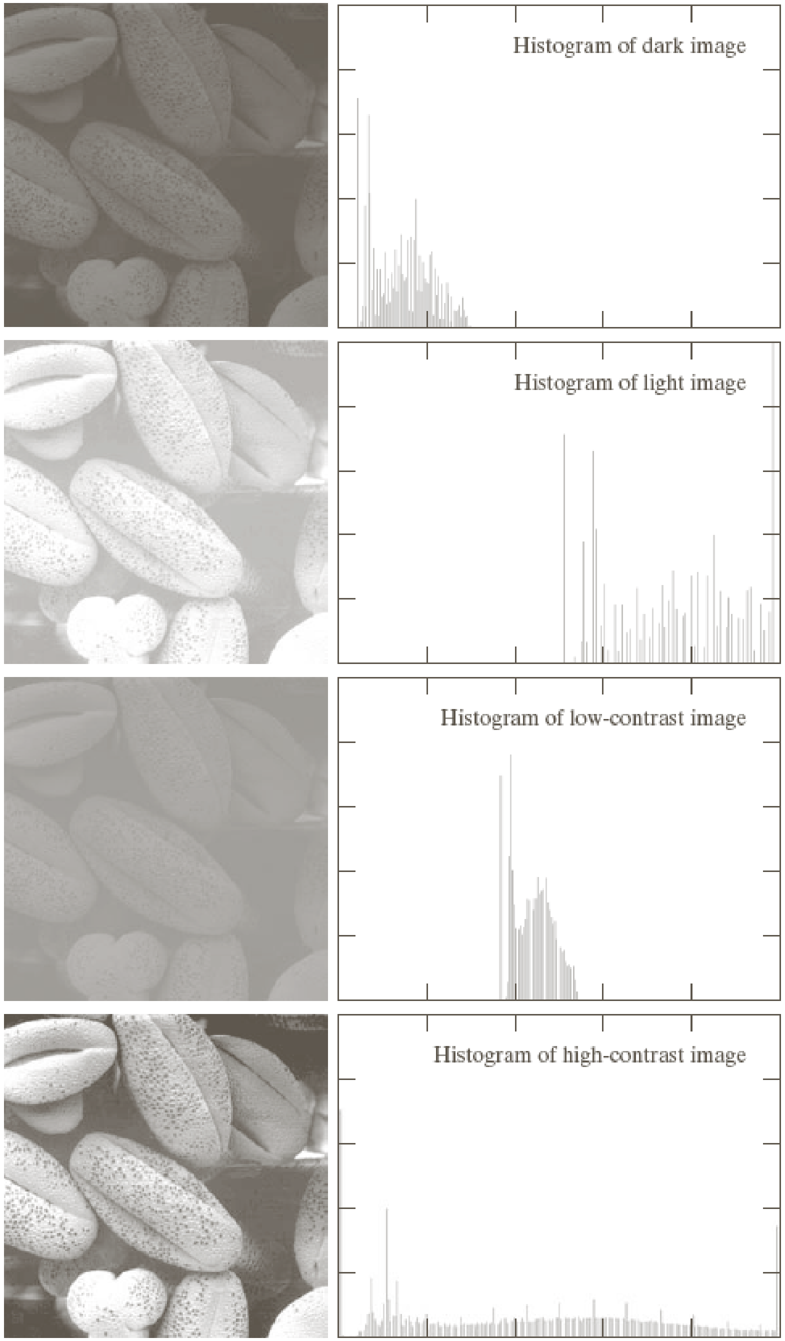


Histogram Processing and Equalization

- Notes

Histograms

- $h(r_k) = n_k$
  - Histogram: number of times intensity level r_k appears in the image
- $p(r_k)= n_k/NM$
  - normalized histogram
  - also a probability of occurence



Histogram equalization

- Automatic process of enhancing the contrast of any given image

Histogram Equalization


Next Class

- Chapter 3 G&W second part on “Spatial Filtering”
- Also read chapter 2, section 2.6.5. as introduction