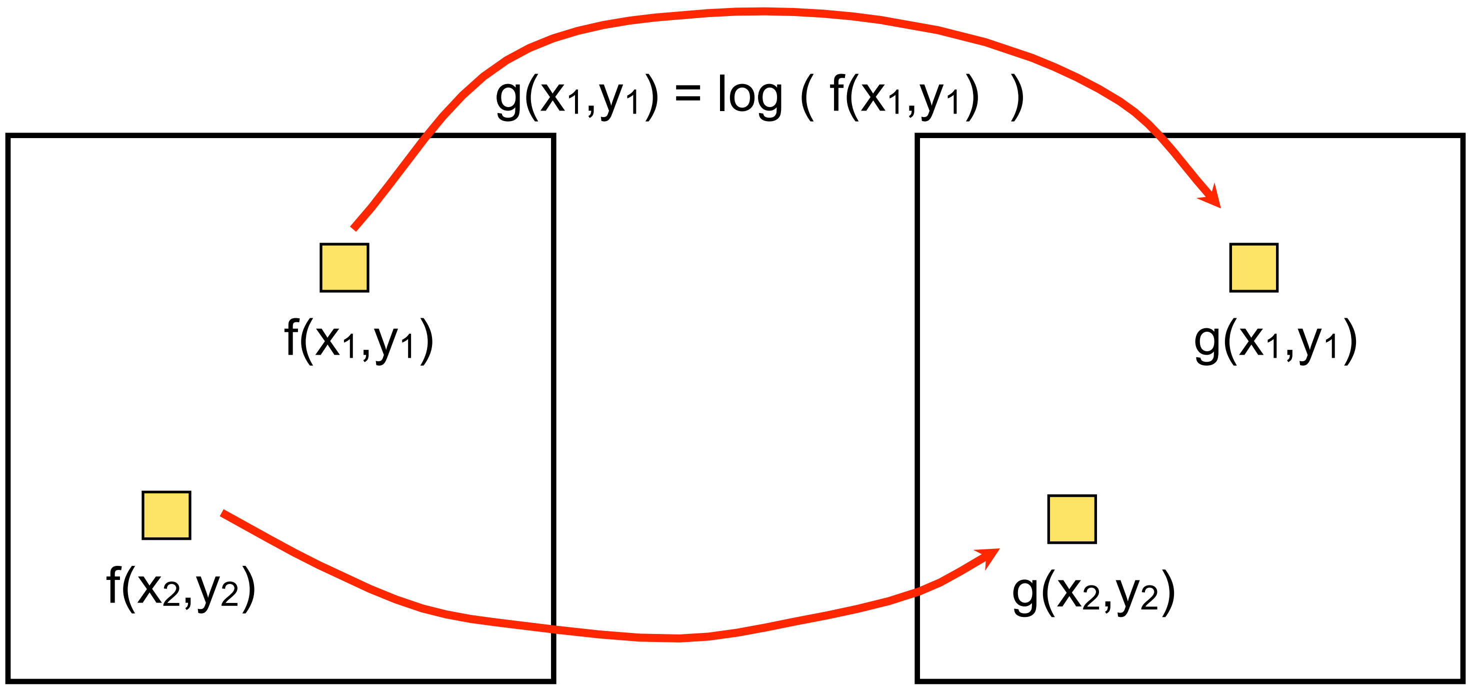
Greyscales, Histograms, and Probabilities

Ross Whitaker
SCI Institute, School of Computing
University of Utah

Intensity Transformation Example (log)

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



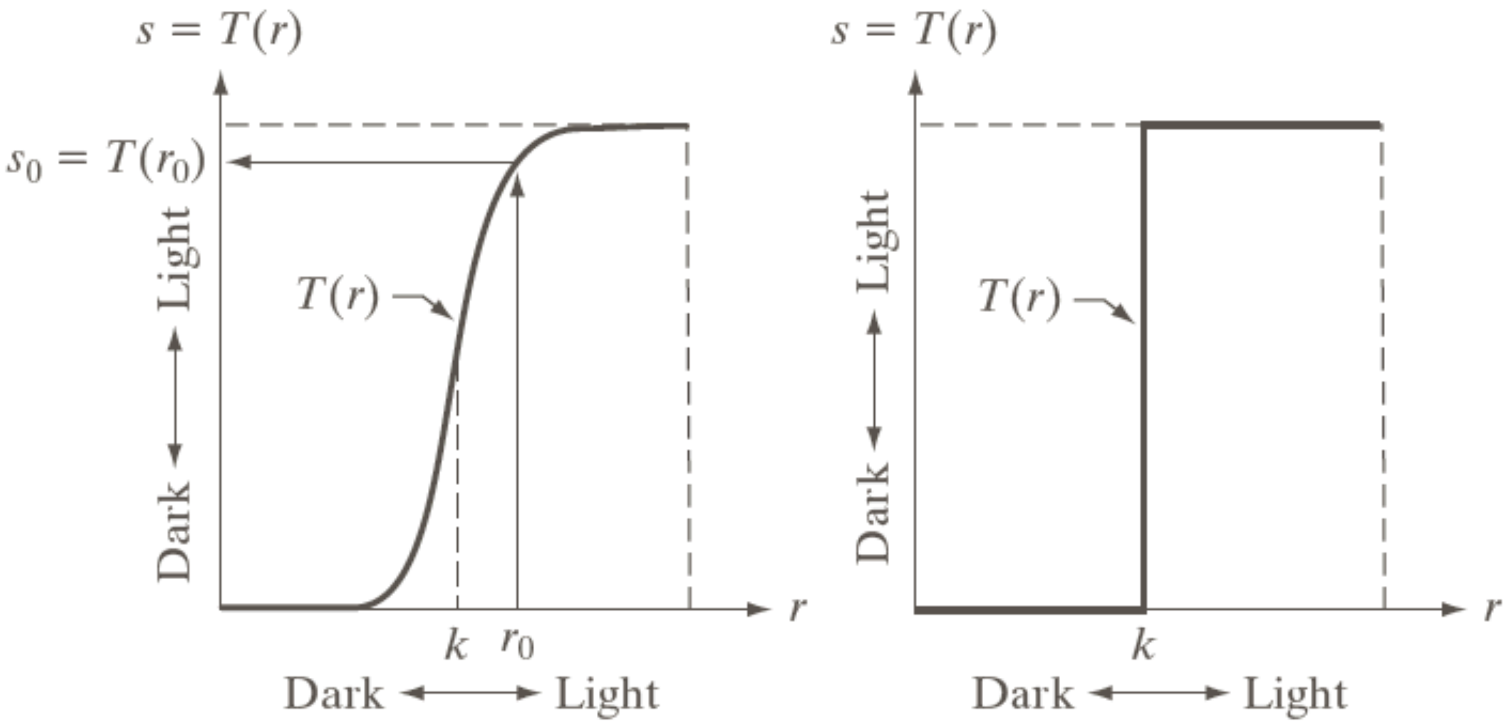
$g(x_2,y_2) = \log(f(x_2,y_2))$

- 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

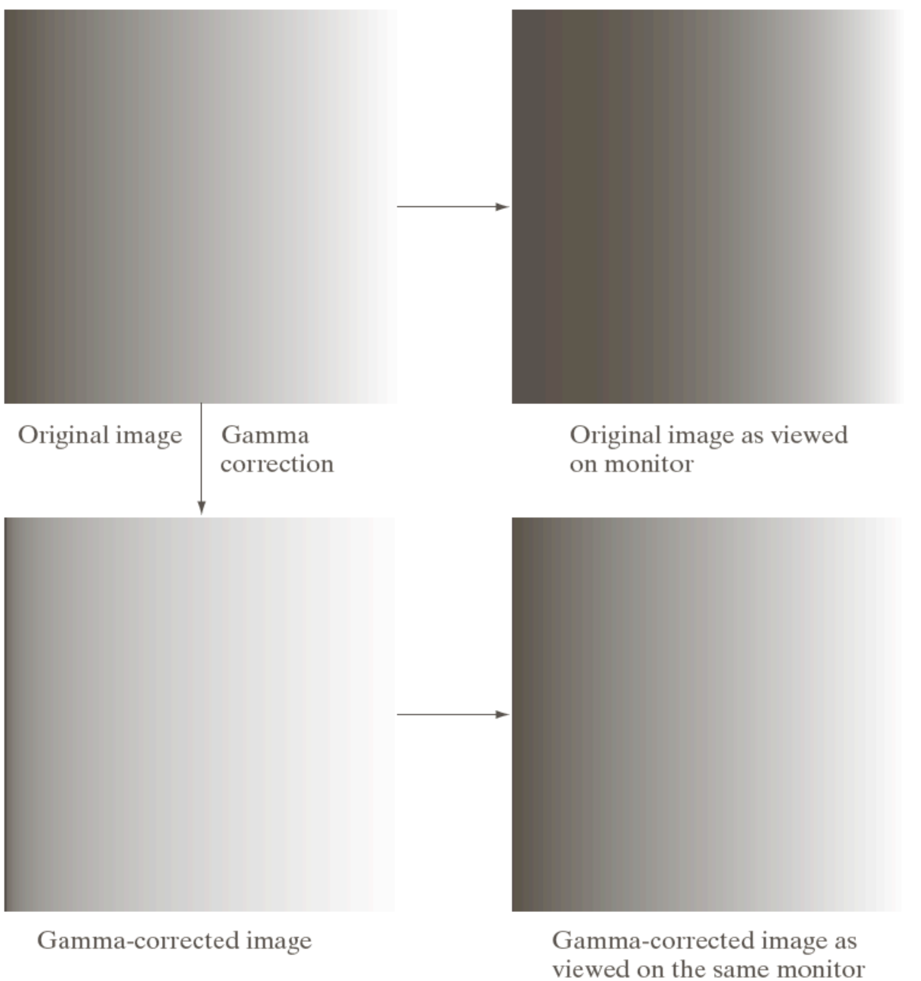
-r: input intensity

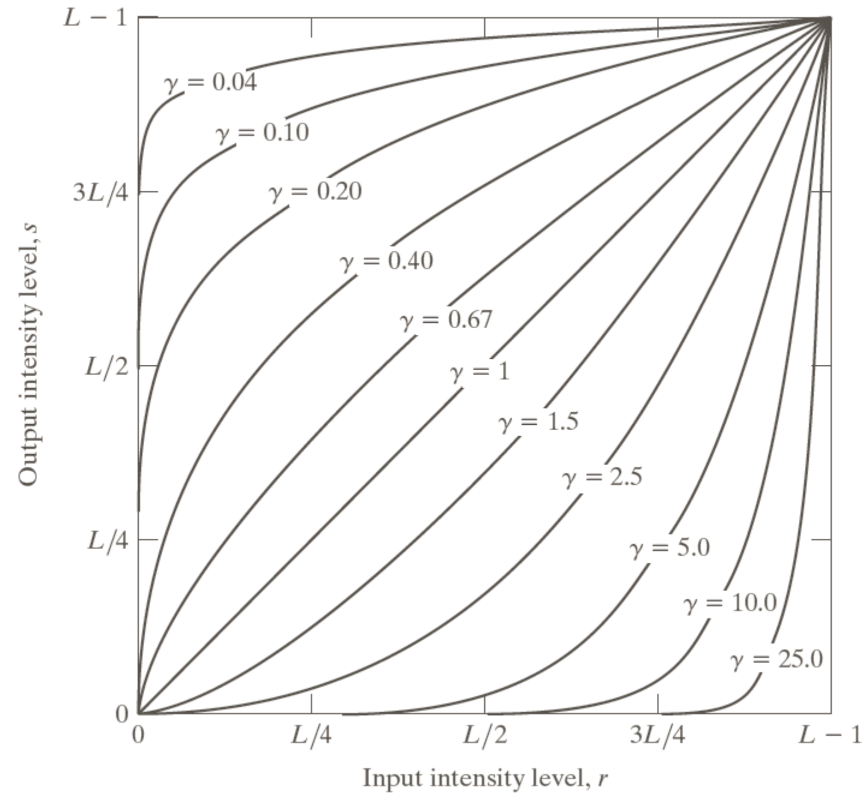
Intensity transformation



$$s = T(r)$$

Gamma correction





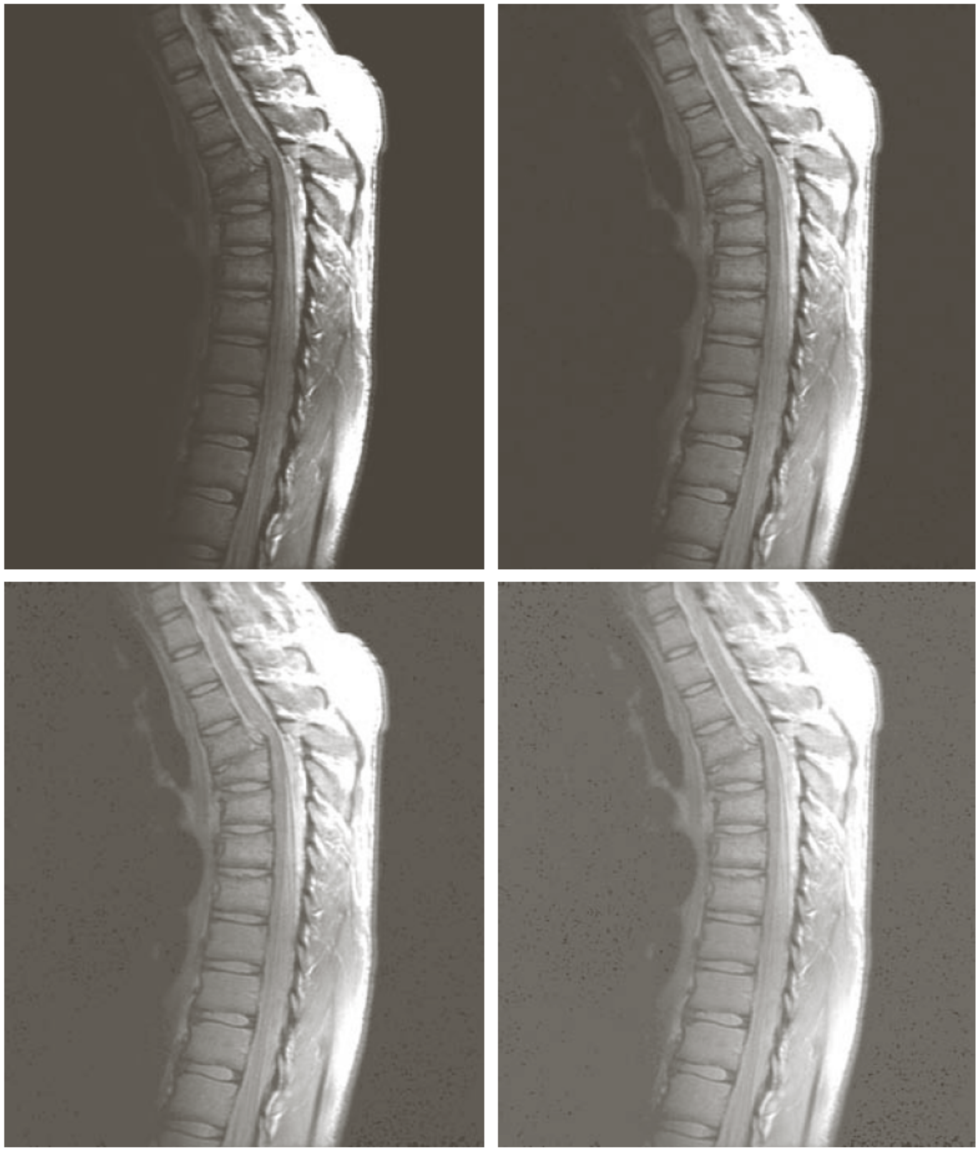
$$s = cr^{\gamma}$$

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.)

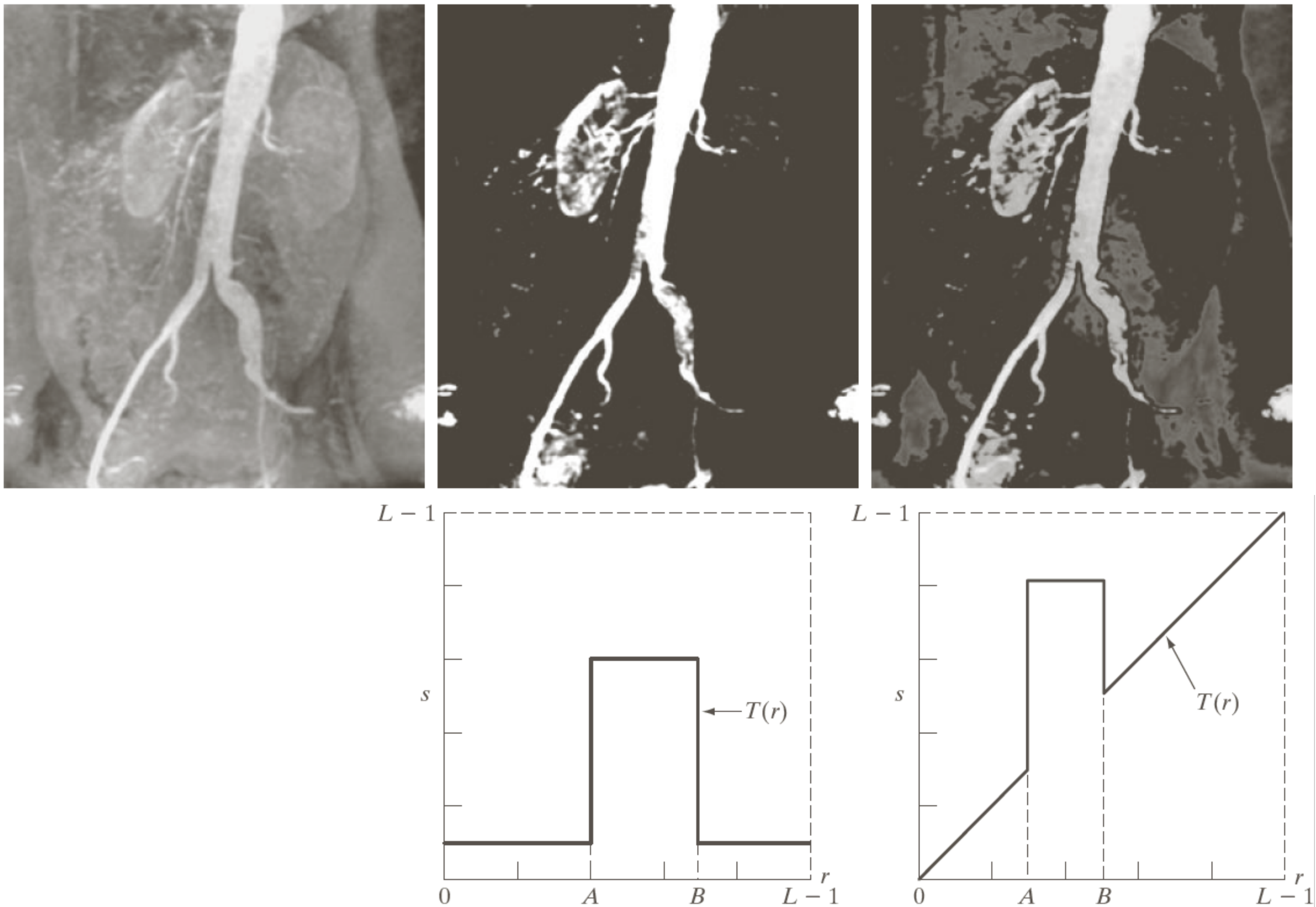
Gamma transformations



a b
c d

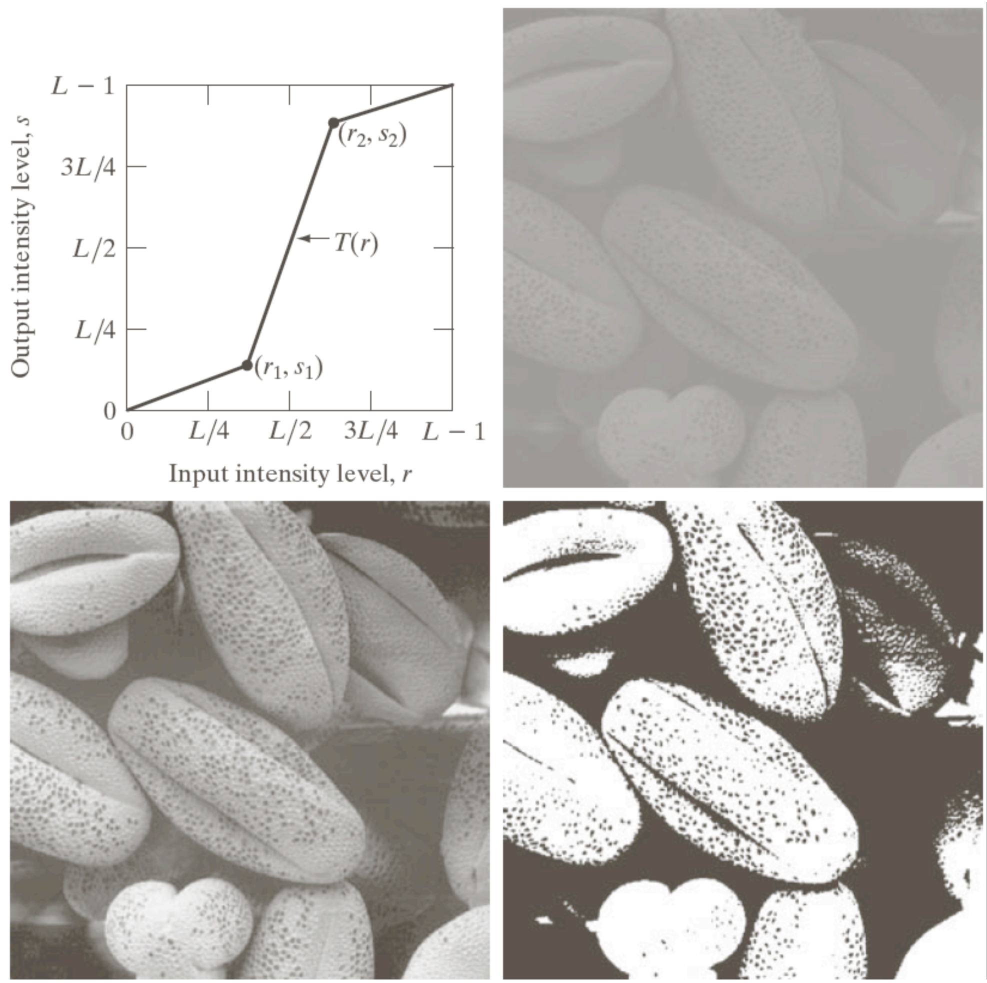
FIGURE 3.8 (a) Magnetic resonance image (MRI) of a fractured human spine. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4,$ and 0.3, respectively. (Original image courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt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

More Intensity Transformations



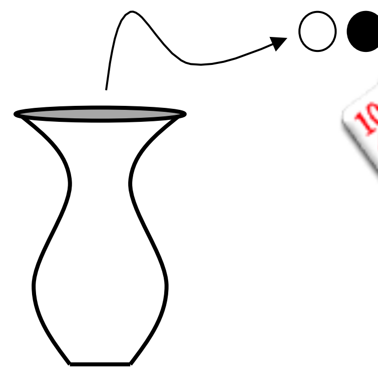
Piecewise linear intensity transformation

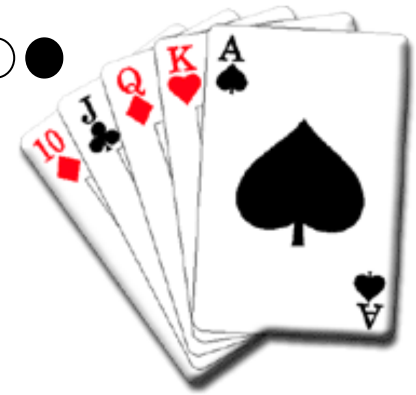
- More control
- But also more parameters for user to specify
  - Graphical user interface can be useful



Sample Spaces

- S = Set 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
  - SxS Cartesian product - sets
  - 2 throws of Dice - (2, 4)
  - 2 picks from an urn - (black, black)
- P(B|A) - probability of B in second experiment given outcome (A) of first experiment
  - This quantifies the effect of the first experiment on the second
- P(A,B) - probability of A in first experiment and B in second experiment
- P(A,B) = P(A) P(B|A)

Independence

- $P(B|A) = P(B)$
  - 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: replace stone after each experiment
- Cards
  - Replace card after it is picked

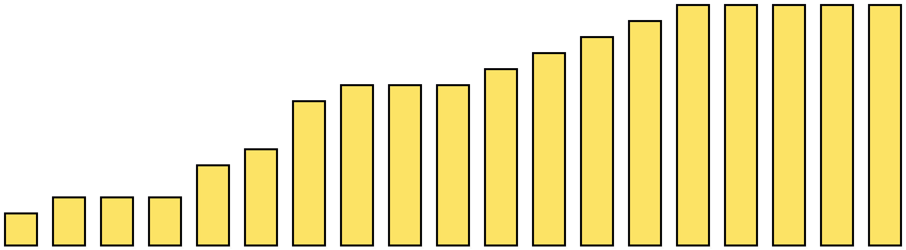
Random Variable (RV)

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

Cumulative Distribution Function (cdf)

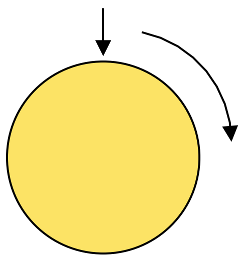
- F(x), where x is a RV
- F(-infty) = 0, F(infty) = 1
- F(x) non decreasing

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



Continuous Random Variables

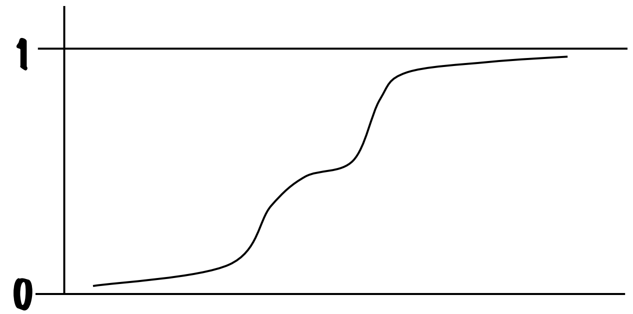
- Example: spin a wheel and associate value with angle



- 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)

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

- A probability density is not the same as a probability

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

  - To get meaningful numbers you must specify a range

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 a constant

Mean of a PDF

- Mean = E[x]
  - also called "μ"
- Variance = $E[(x - \mu)2]$
- = $E[x2] - E[2\mu x] + E[\mu 2]$
- = $E[x2] - \mu 2$
  - also called "$\sigma 2$"
  - Standard deviation is σ
  - For a distribution having zero mean: $E[x2] = \sigma 2$

Sample Mean

- Run N experiments (independent)
  - Draw N sample points from a single pdf
  - Sum them up and divide by N
- Resulting M is called the sample mean
  - 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?
- Consider variance of sample mean (M)
- Define a new random variable: D = (M - m)2

Independence -> E[xy] = E[x]E[y]

Number of terms off diagonal

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

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

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

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

As number of samples -> infty, sample mean -> true mean

Application: Denoising 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
- At pixel (x,y): g(x,y) = s(x,y) + n(x,y)

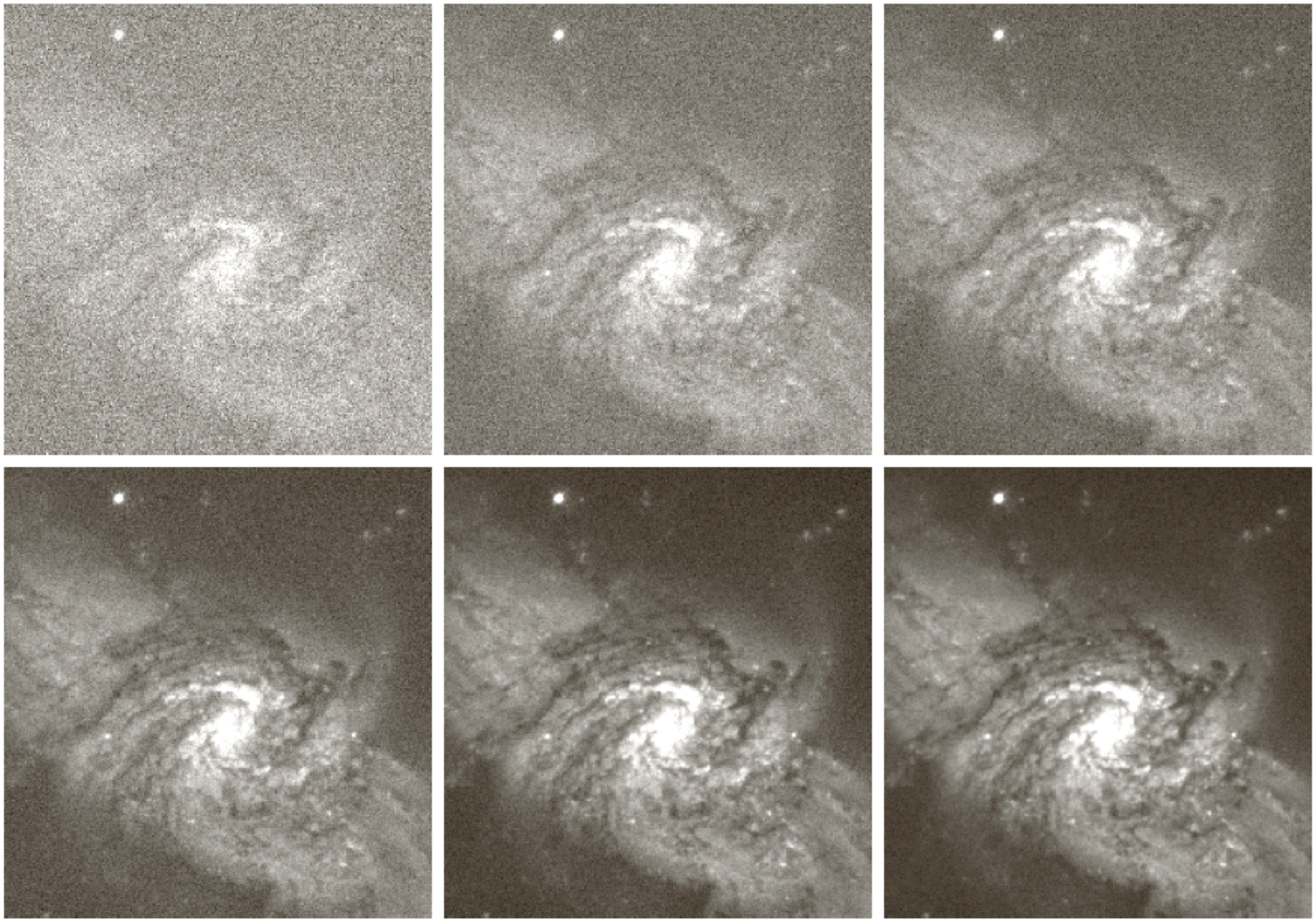
True pixel value

Random zero-mean noise:
•Independent from one image to the next
•Variance = σ

Application: Denoising Images

- Take multiple images of the same scene
  - $g_i = s + n_i$
  - Mean $[n_i] = 0$; Variance $[n_i] = \sigma^2$
  - Mean $[g_i] = s$; Variance $[g_i] = \sigma^2$
  - Sample mean $= M = (1/N) \Sigma g_i = s + (1/N) \Sigma n_i$
  - Mean $[M] = s$; Variance $[M] = (1/N) \sigma^2$
- Application:
  - Digital cameras with large gain (high ISO, light sensitivity)
  - Astronomy imagery

Averaging Noisy Images Can Improve Quality

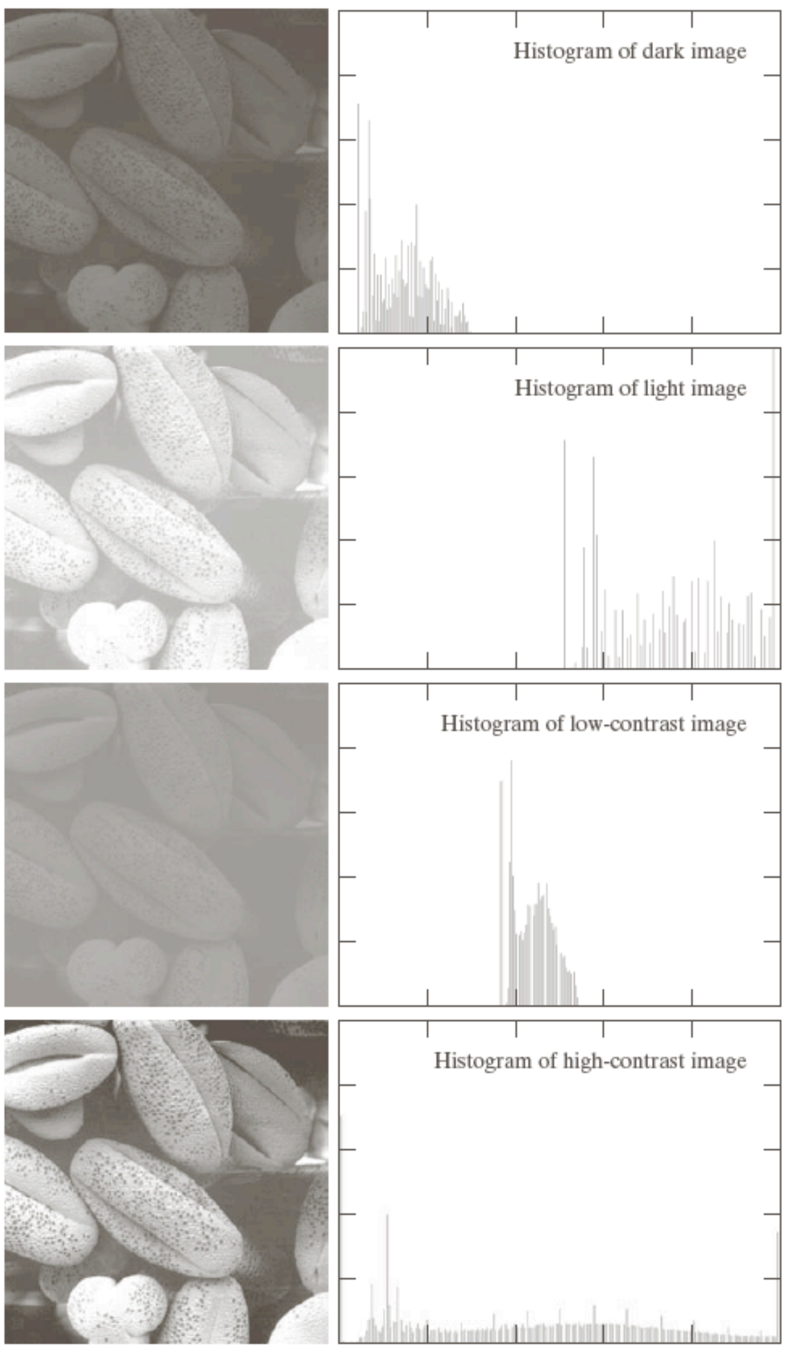


a b c
d e f

FIGURE 2.26 (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)–(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)

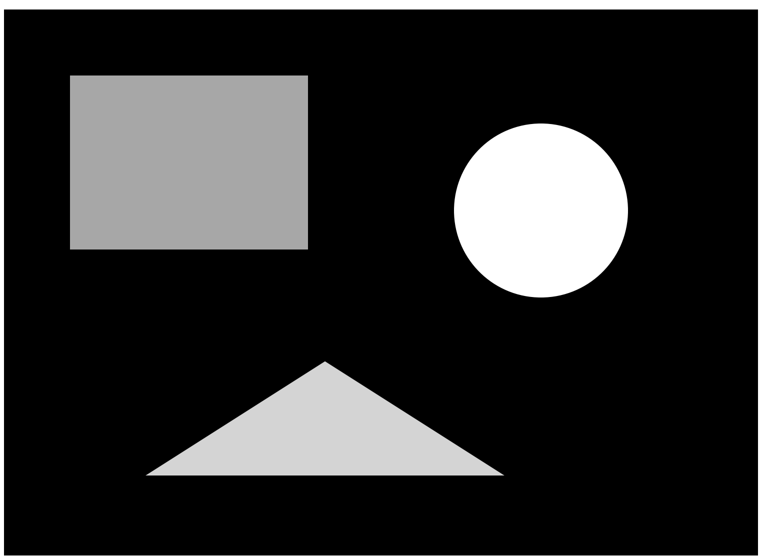
Histograms

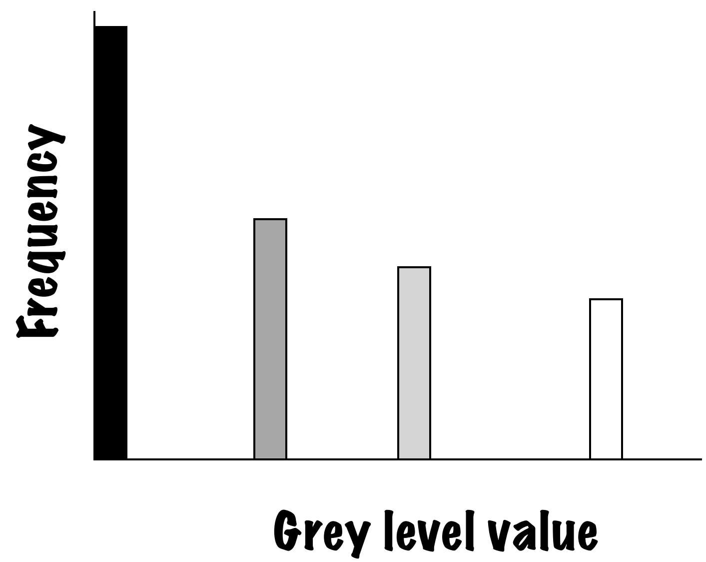
- $h(r_k) = n_k$
  - Histogram: number of times intensity level rk appears in the image
- $p(r_k)= n_k/NM$
  - normalized histogram
  - also a probability of occurence



Histogram of Image Intensities

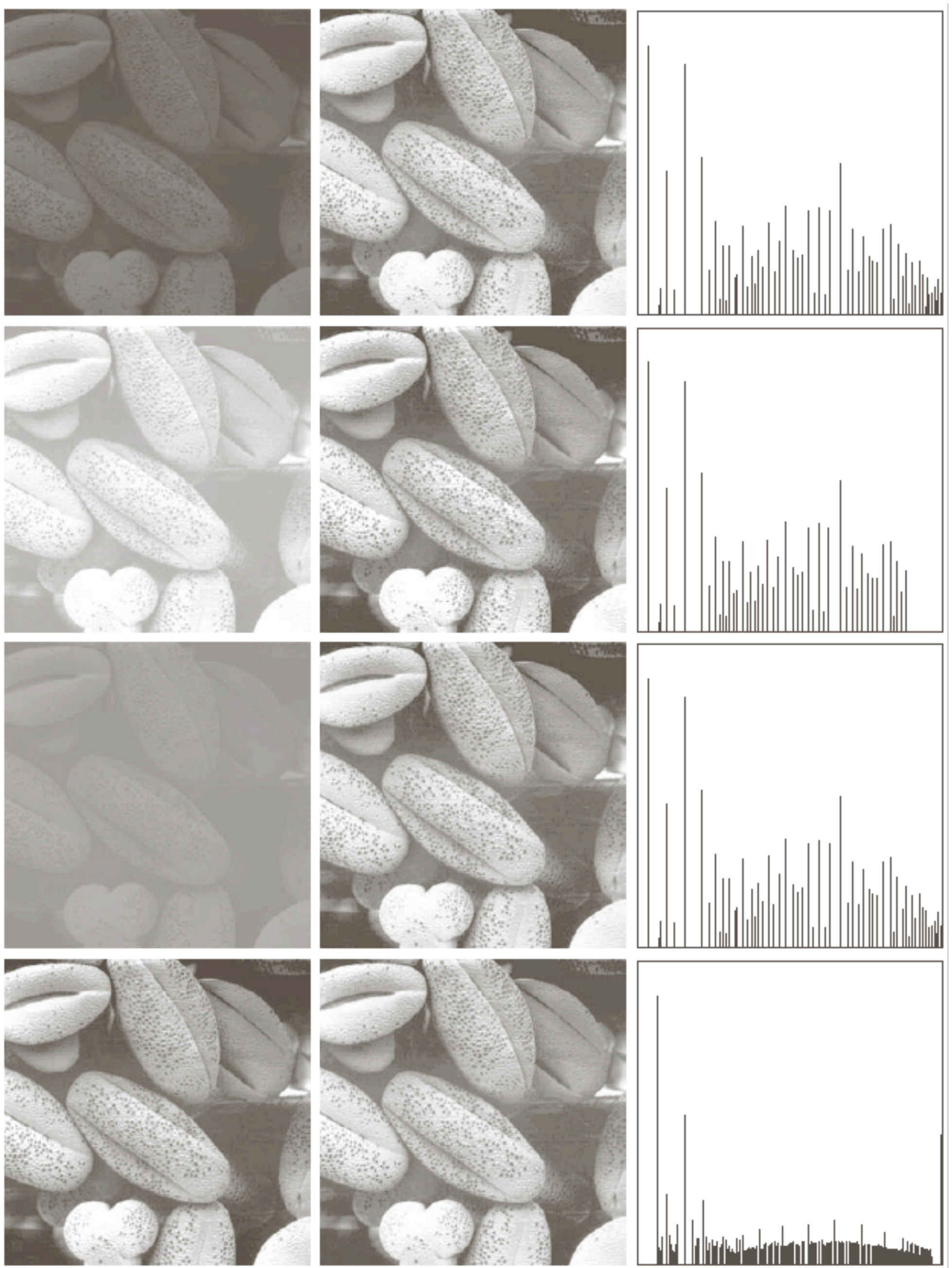
- Create bins of intensities and count number of pixels at each level
  - Normalized (divide by total # pixels)





Histogram Equalization

- Automatic process of enhancing the contrast of any given image



Histogram Equalization

Tuning Down Hist. Eq.

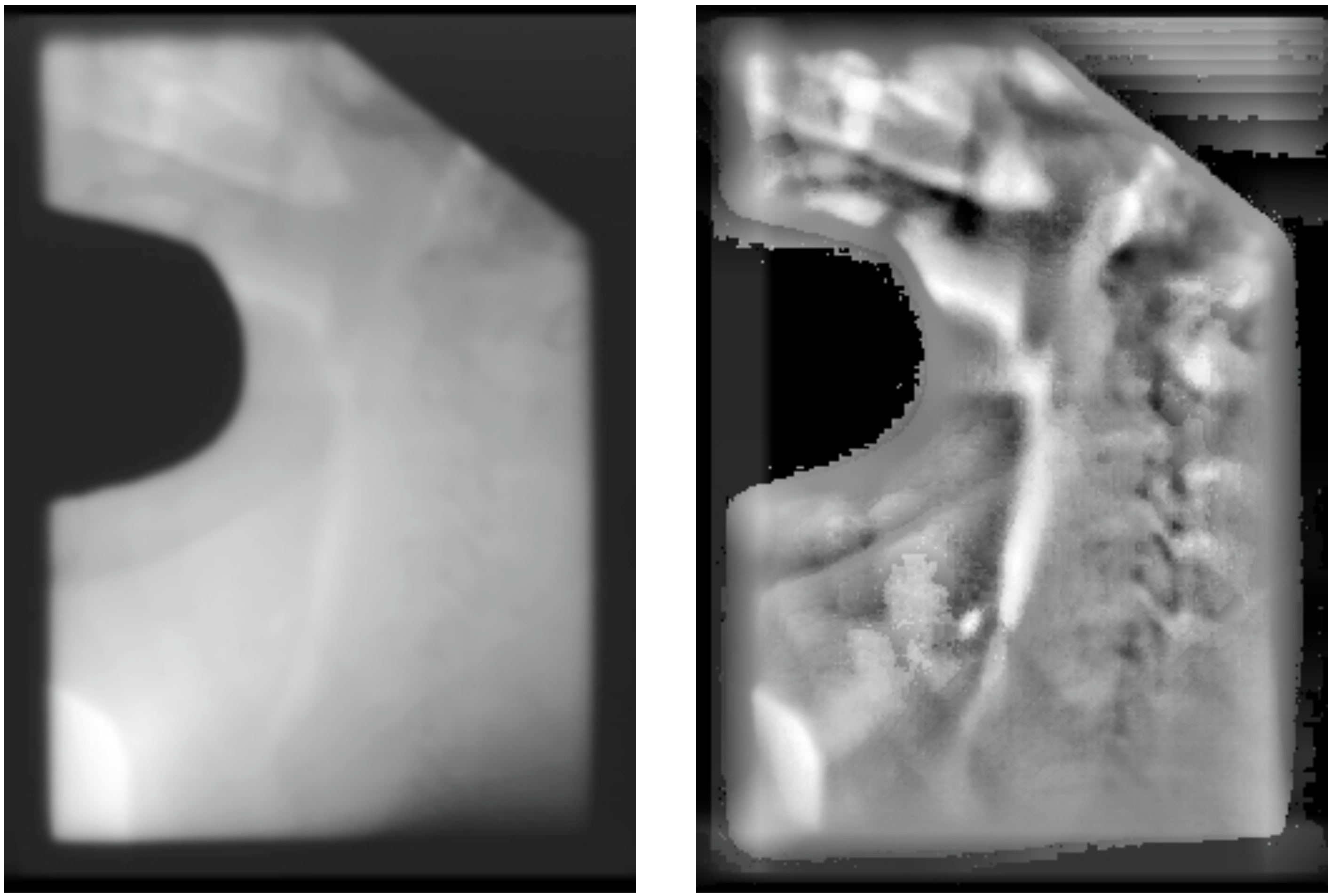
- Transformation is weighted combination of CDF and identity with parameter alpha

$$t(s) = (1 - \alpha)s + \alpha A(s)$$

α = 0.0 α = 0.2 α = 0.4

α = 0.6 α = 0.8 α = 1.0

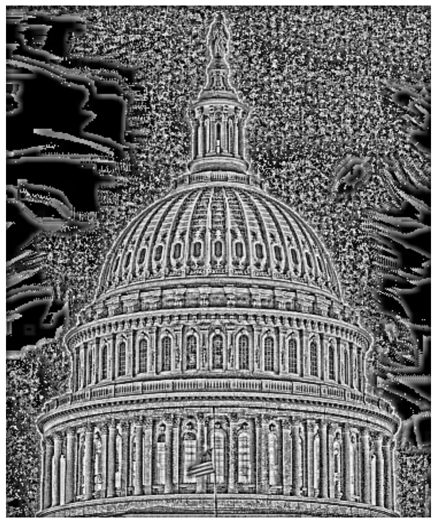
Adaptive Histogram Equalization (AHE)



AHE Gone Bad…

Effect of Window Size

Orig



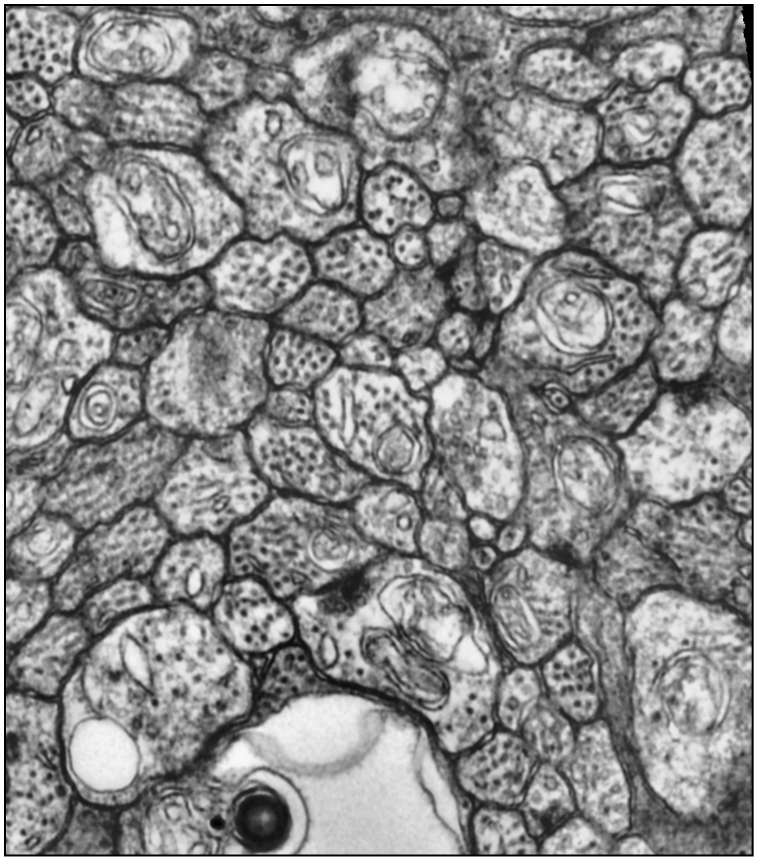
10x10

25x25

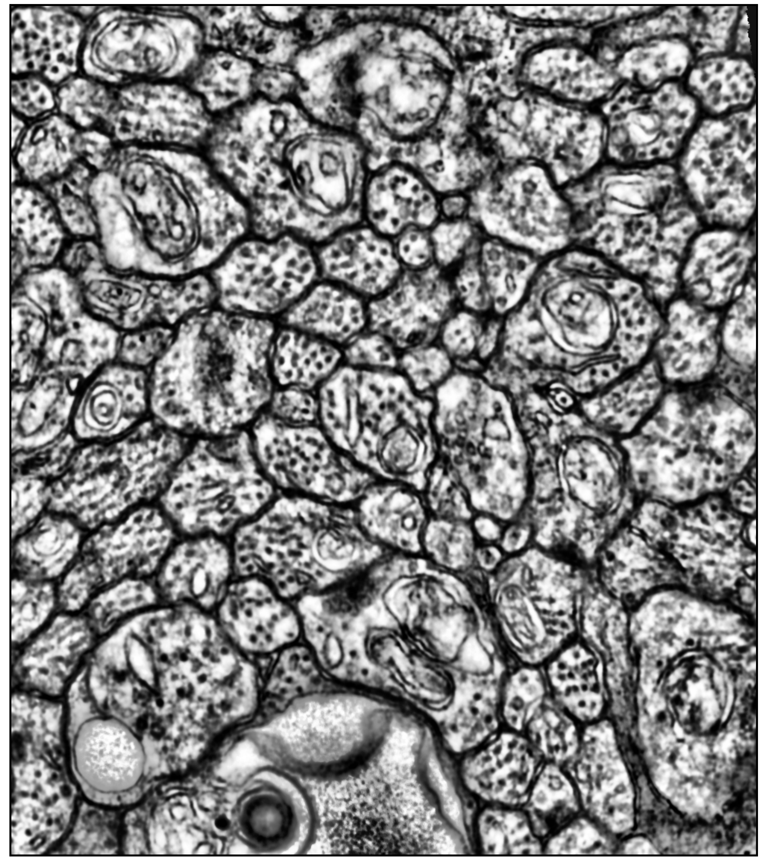
50x50

AHE Application: Microscopy Imaging

Original

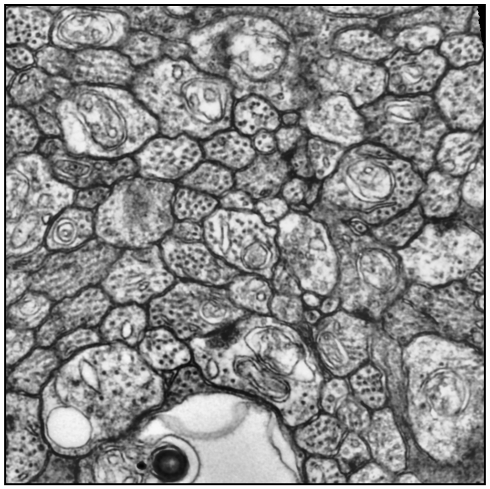


AHE

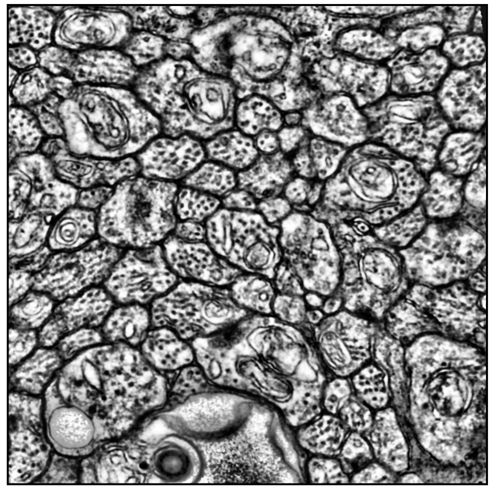


AHE Application: Microscopy Imaging

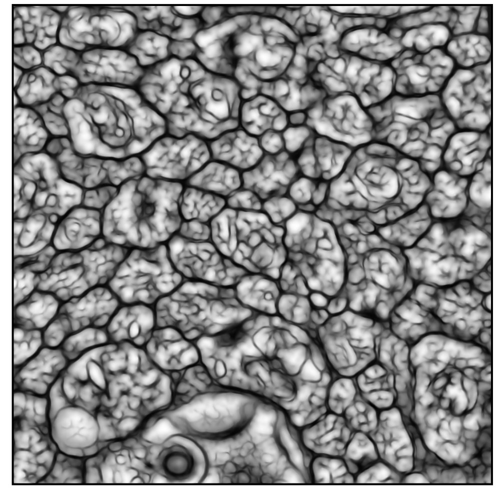
Original

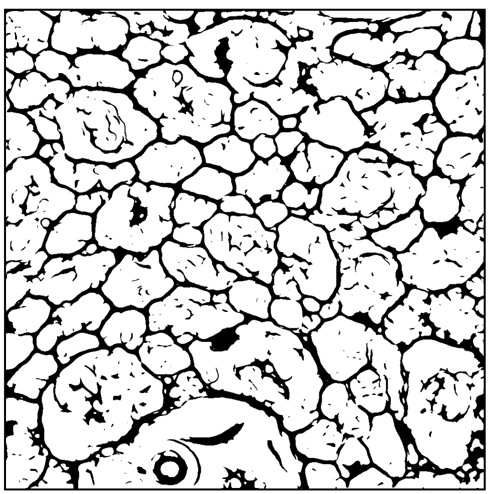


AHE

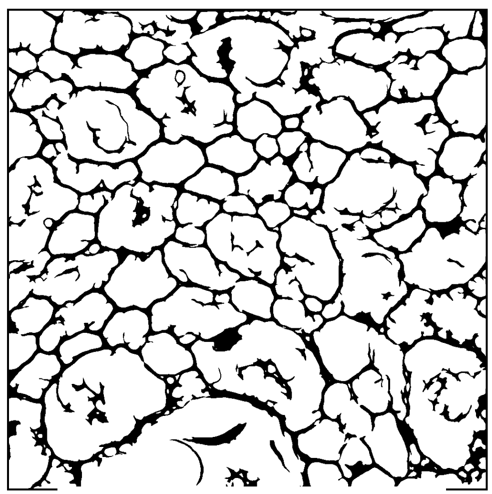


Adaptive Filtering

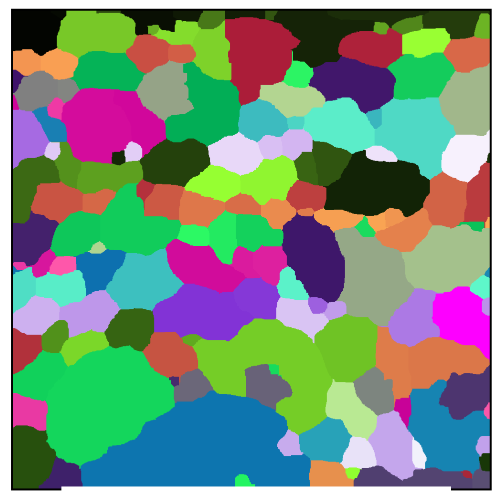




Threshold



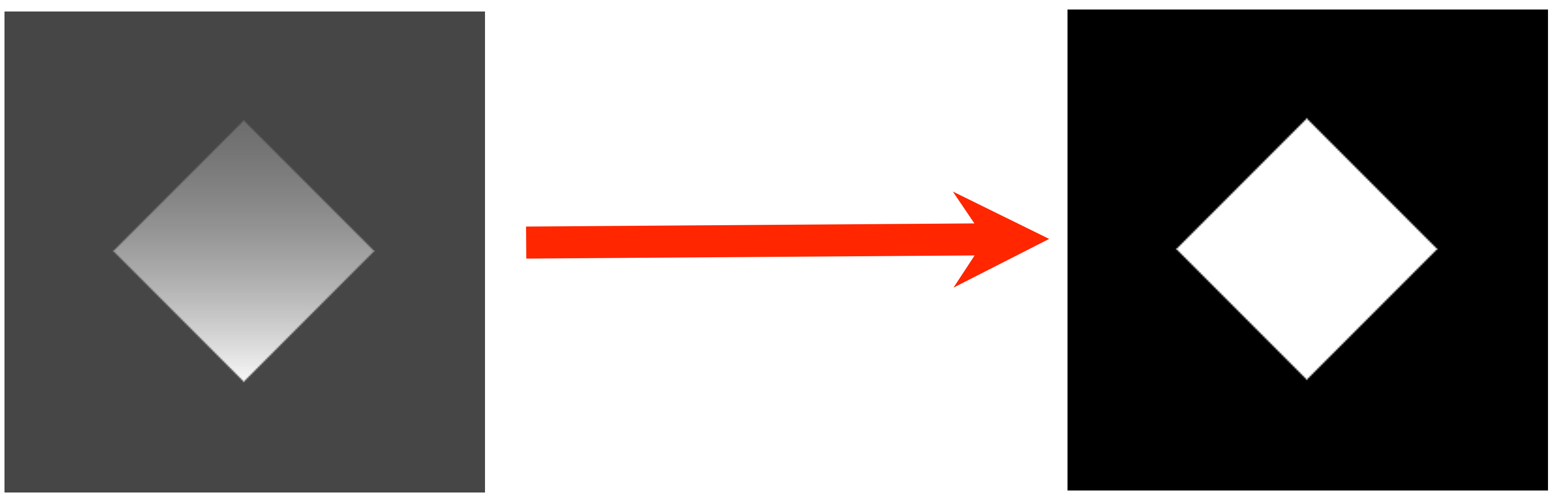
CC Analysis/Morphology



CC Analysis/Watersheds

What is image segmentation?

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



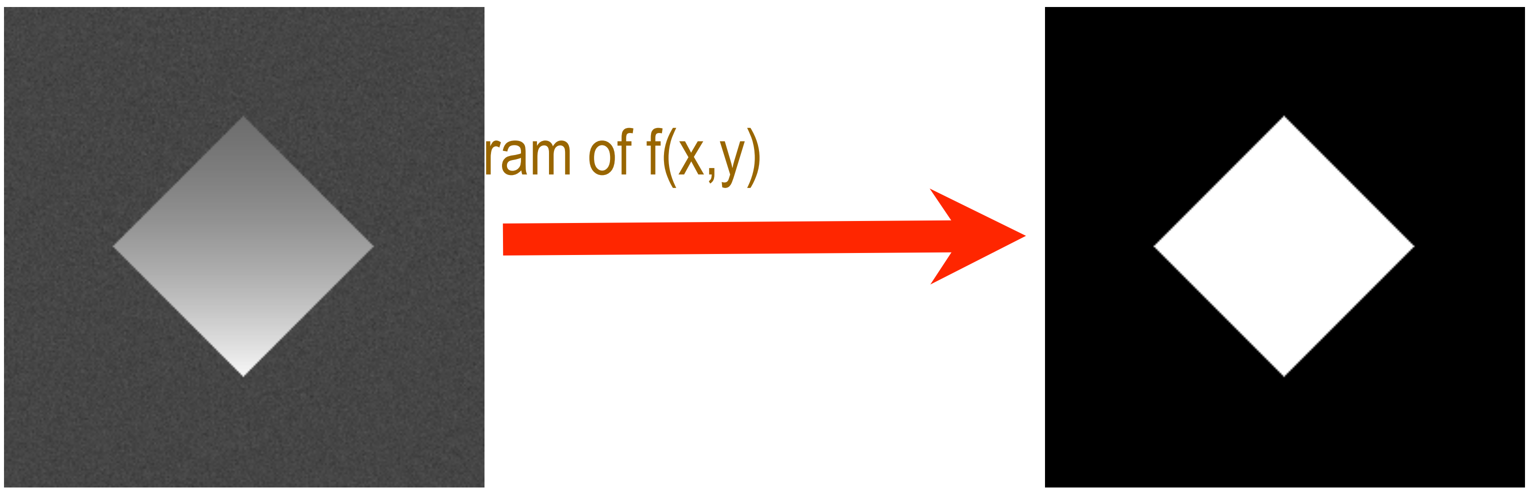
Input image
intensities 0-255

Segmentation output
0 (background)
1 (foreground)

Thresholding

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

- How can we choose T?

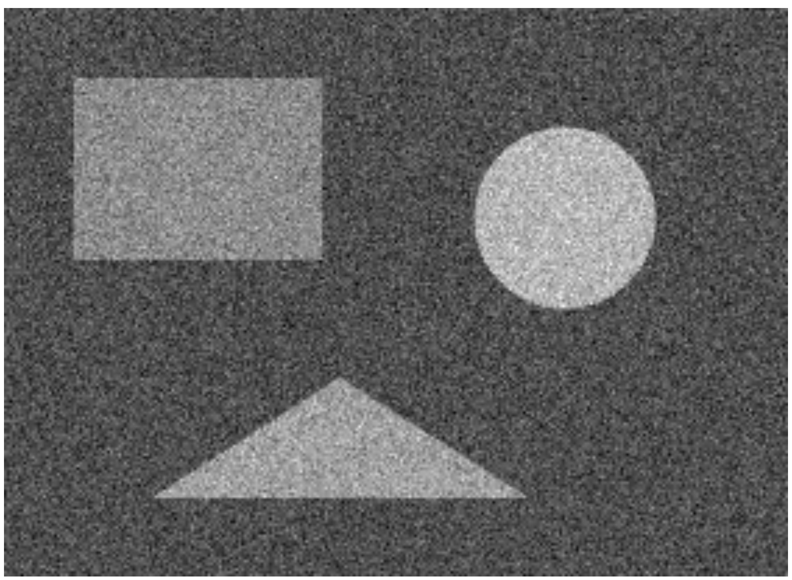


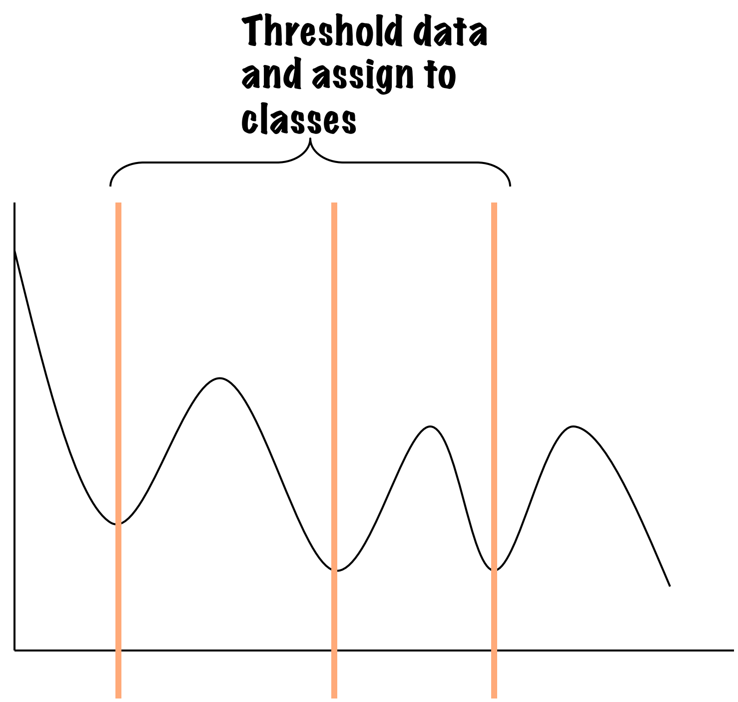
Input image f(x,y)
intensities 0-255

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

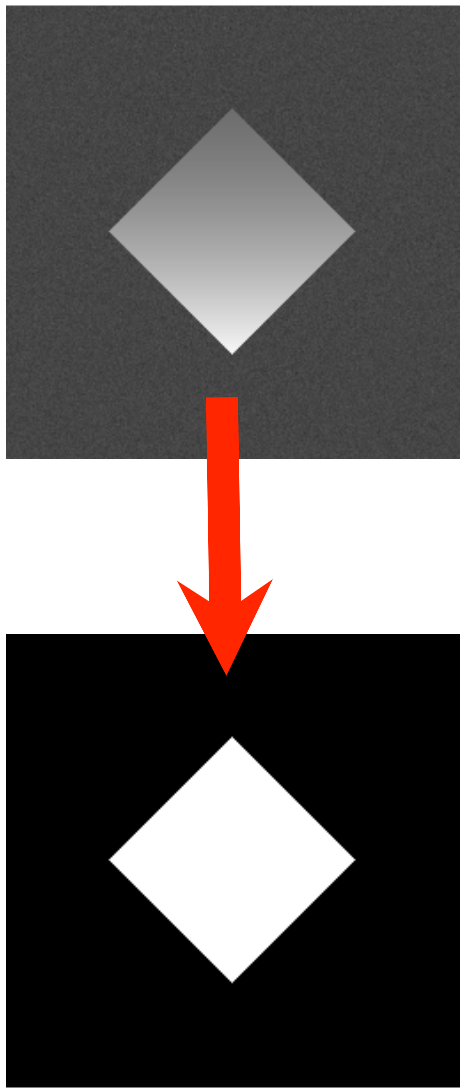
Histograms and Noise

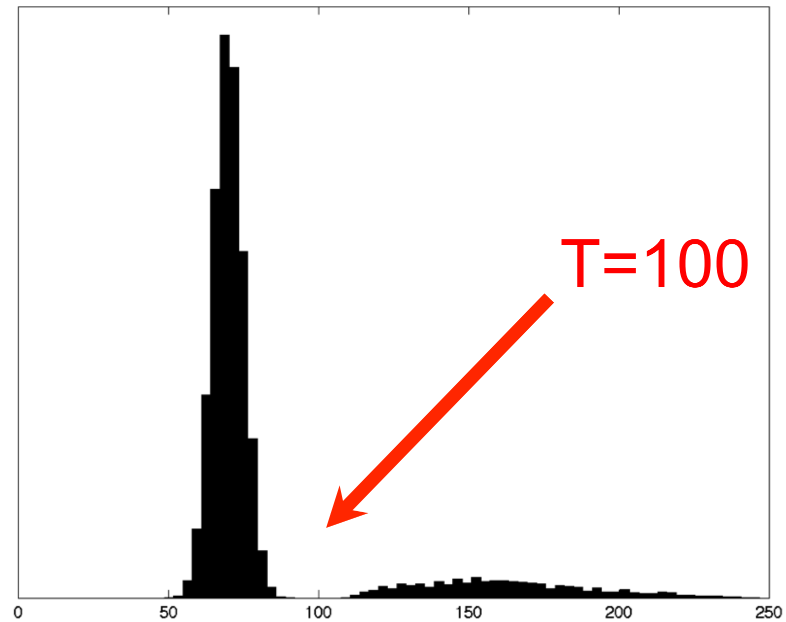
- What happens to the histogram if we add noise?
  - $g(x, y) = f(x, y) + n(x, y)$





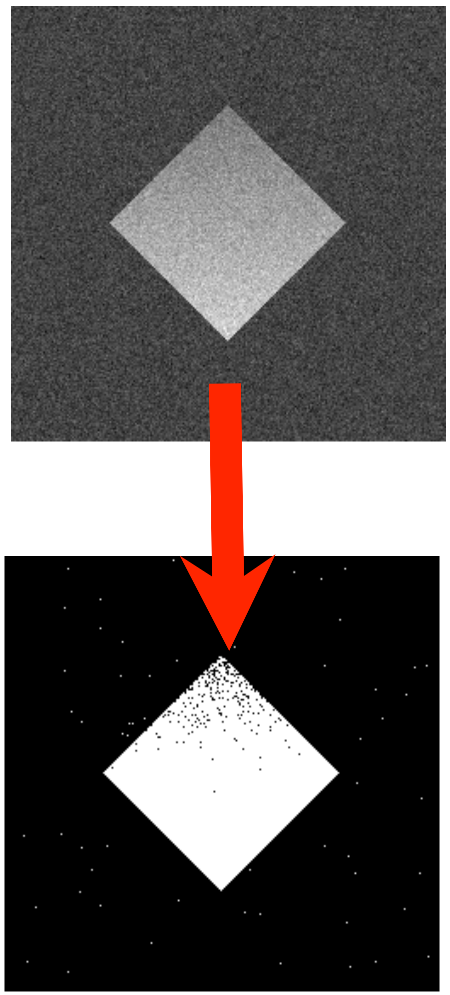
Choosing a threshold

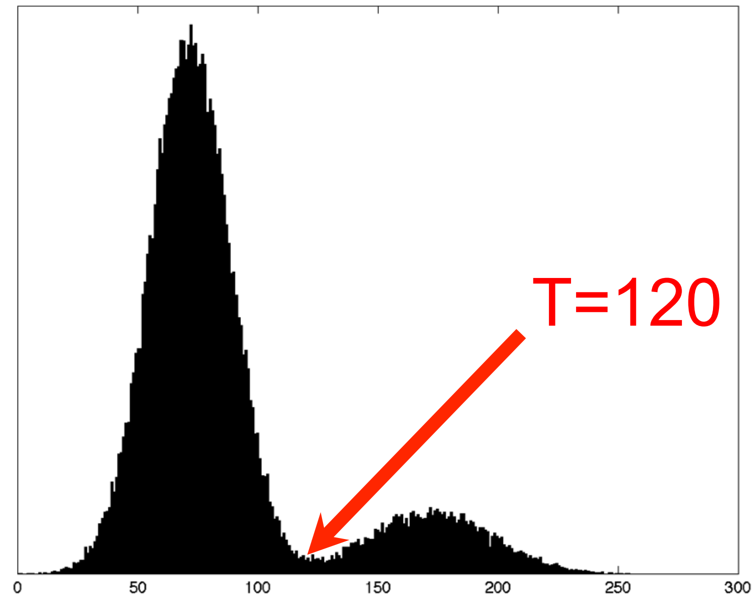




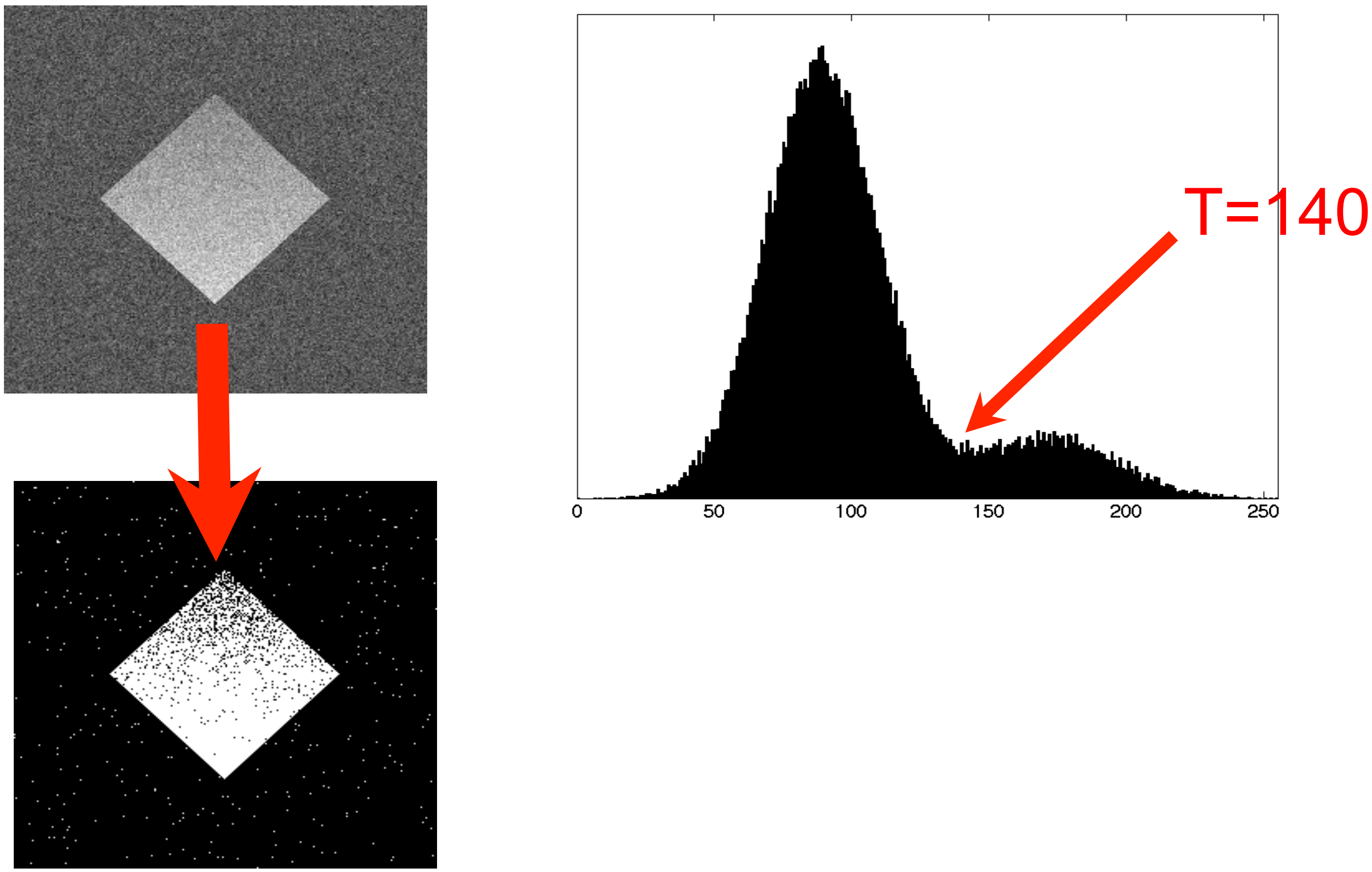
Histogram

Role of noise





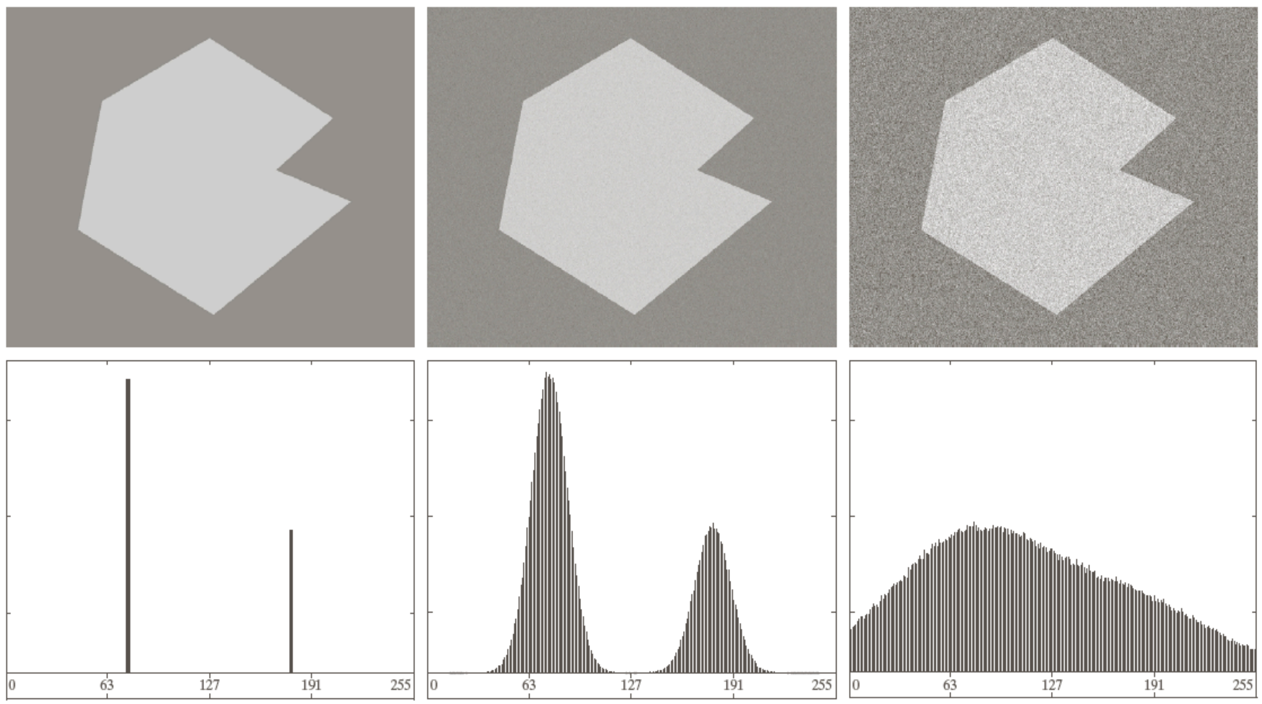
Low signal-to-noise ratio



Effect of noise on image histogram

Images

Histograms

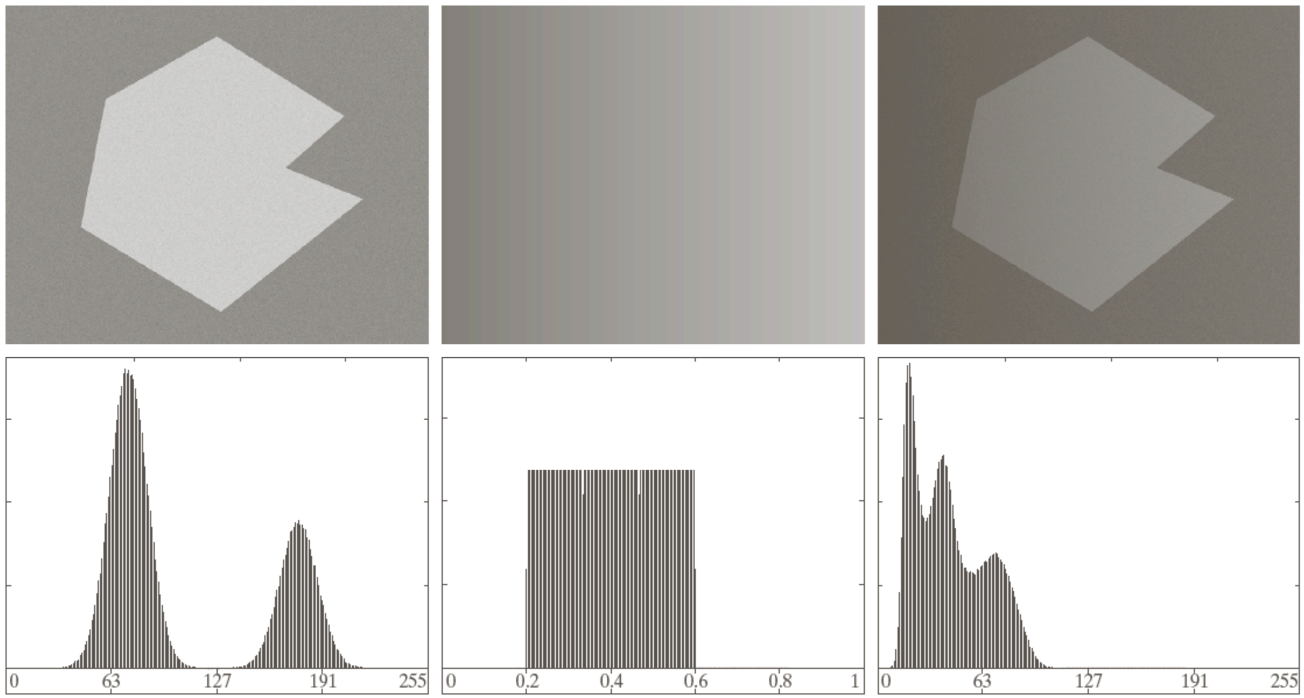


No noise With noise More noise

Effect of illumination on image histogram

Images

Histograms



f x g = h

Original image

Illumination image

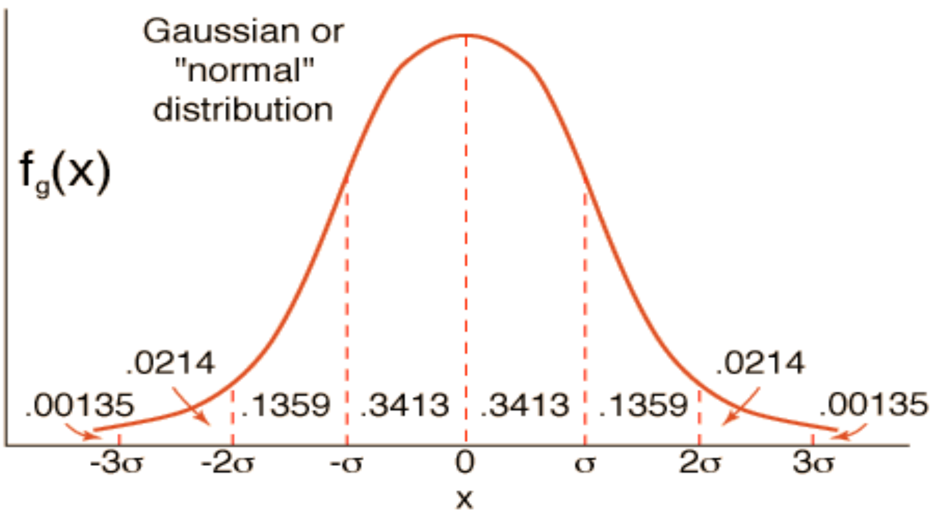
Final image

Some Extra Things

- Gaussian/normal distribution
- Weighted means

Gaussian Distribution

- "Normal" or "bell curve"
- Two parameters
  - μ = mean, σ = standard deviation



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

Gaussian Properties

- Best fitting Gaussian to some data is gotten by mean and standard deviation of the samples
- Occurrence
  - Central limit theorem: mean of lots of independent & identically-distributed RVs
  - Nature (approximate)
    - Measurement error, physical characteristic, physical phenomenon
    - Diffusion of heat or chemicals

Weighted Mean from Samples

- Suppose
  - We want to compute the sample mean of a “class” of things (or we want to reduce it’s influence)
  - We are not sure if the ith item belongs to this class or not - “partially belongs”
    - probability w_i, random variable r_i

Sample mean (no weights)

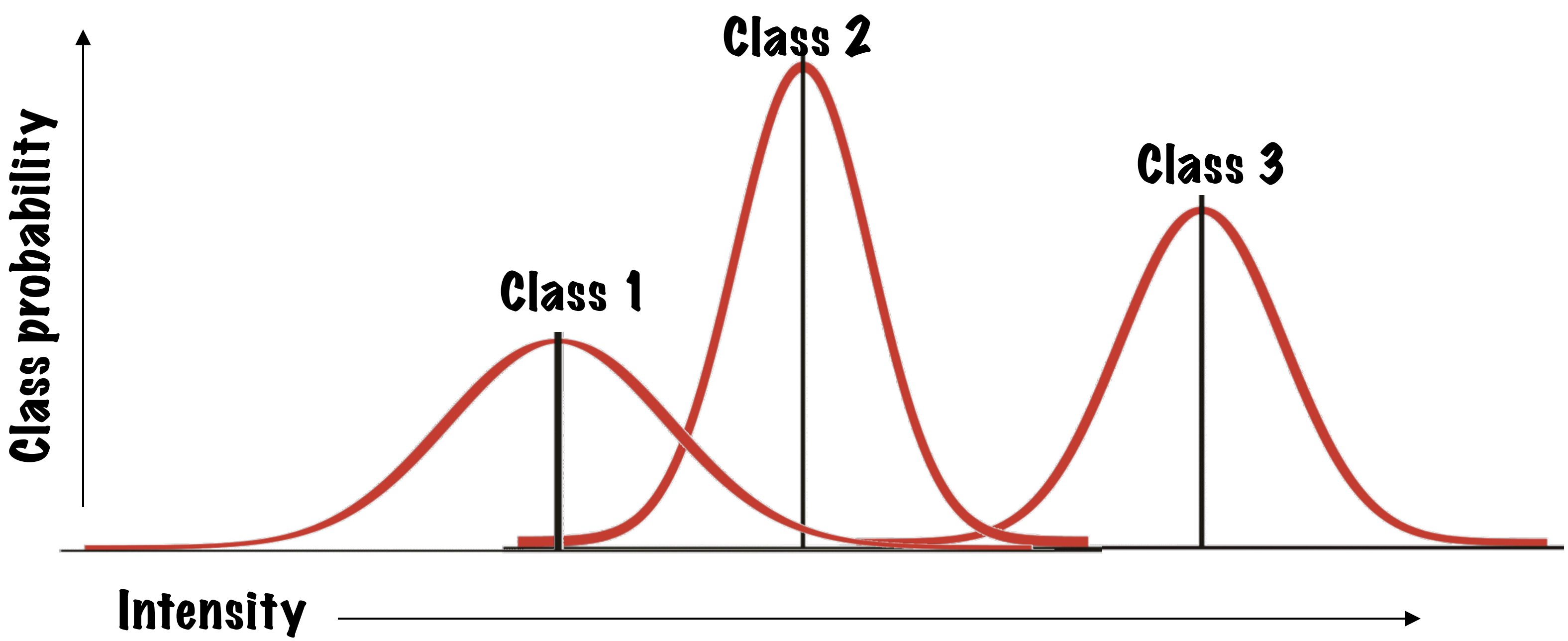
$$E[r] = \frac{1}{N} \sum_{i=1}^{N} r_i$$

Weighted sample mean

$$E[r] = \frac{1}{\sum_{i=1}^{N} w_i} \sum_{i=1}^{N} w_i r_i$$

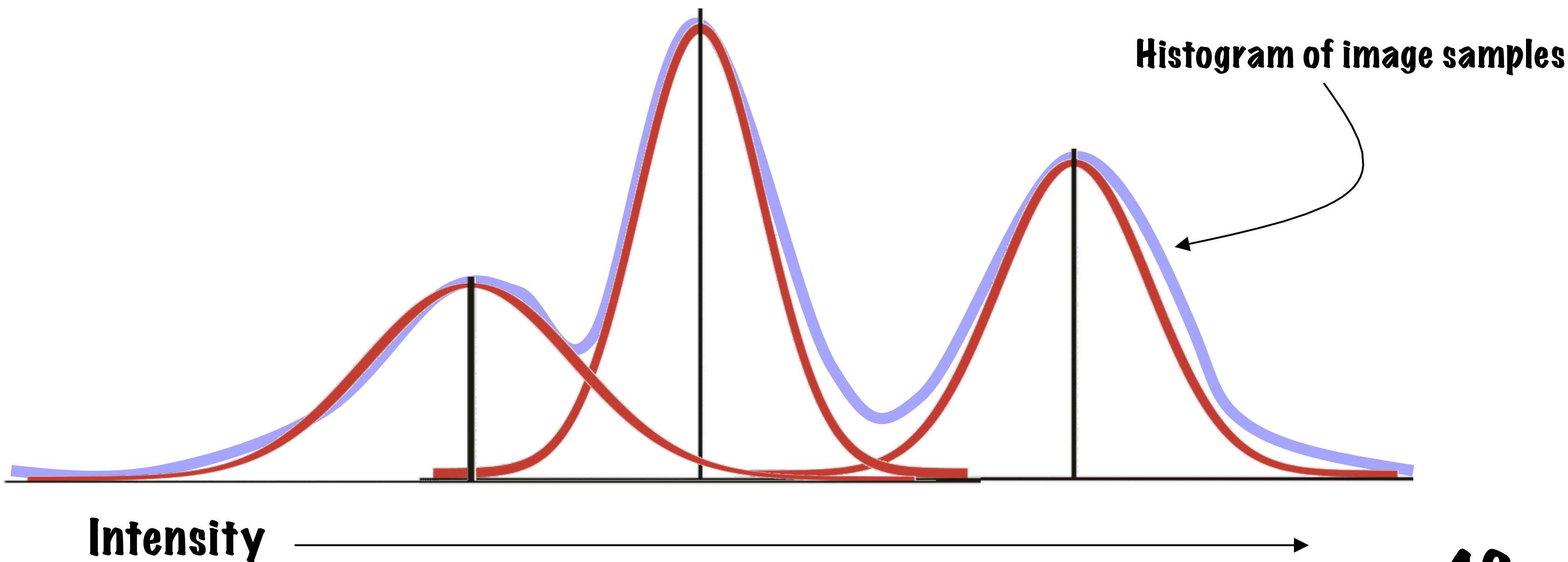
Gaussian Mixture Modeling of Image Histograms

- K classes, N samples



Problem Statement

- Goal: assign pixels to classes based on intensities (output = label image)
- Problem: can we simultaneously learn the class structure and assign the class labels?



Crisp vs. Soft Class Assignment

- If we knew the pdfs (Gaussians) of the classes, we could assign class labels to each data point/ pixel
  - Assume equal overall probabilities of classes

Crisp Assign

$$C_i = \text{argmax}_j P_j(r_i)$$

Find class that has max probability for given intensity r at pixel I. Assign that class label to that pixel

Soft Assign

$$w_i^j = P(C_i = j | r_i) = \frac{1}{\sum_{l=1}^{K} P_l(r_i)} P_j(r_i)$$

For each pixel and each class, assign a (conditional) probability that that pixel belongs to that class

Simultaneously Estimate Class PDFs and Pixel Labels – Iterative Algorithm

- Start with initial estimate of class models

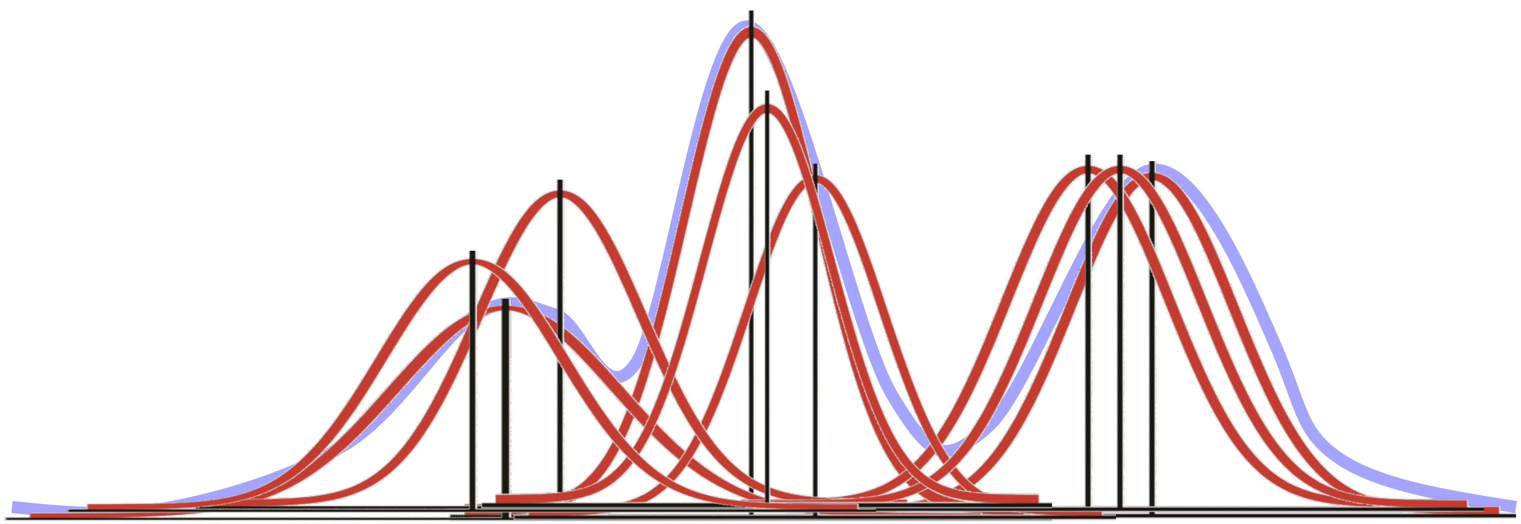
$$\mu_j^0, \sigma_j^0 \text{ for } j = 1 \ldots K$$

- Compute matrix of soft assignments

$$w_i^j = \frac{1}{\sum_{l=1}^{K} P_l(r_i)} P_j(r_i)$$

- Use soft assignments to compute new weighted mean and standard deviation for each class μ_j^1, σ_j^1
- Use new mean and standard deviation to compute new soft assignments and repeat (until change in parameters is very small)

EM Algorithm – Example



MRI Brain Example