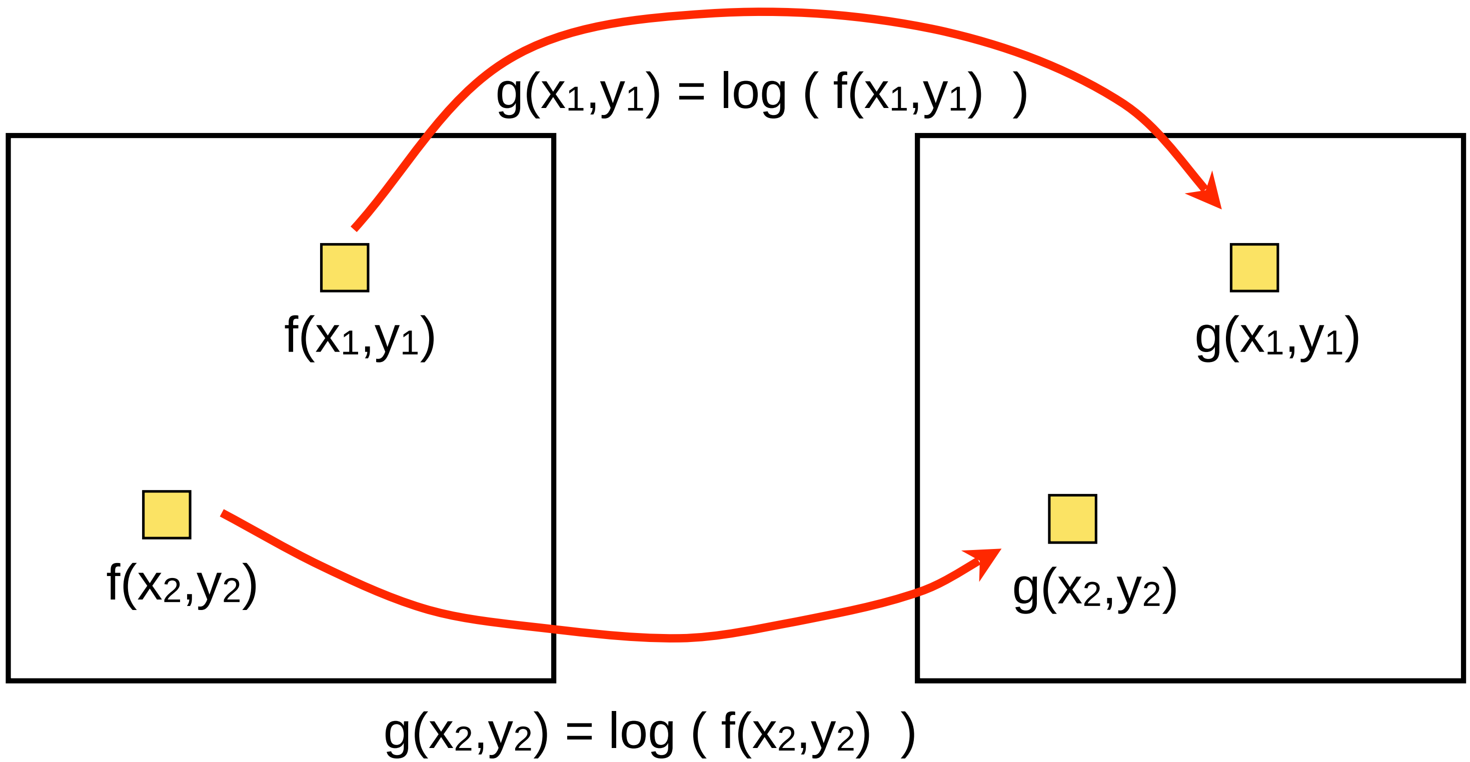
Probabilities, Greyscales, and Histograms

Ross Whitaker
SCI Institute, School of Computing
University of Utah

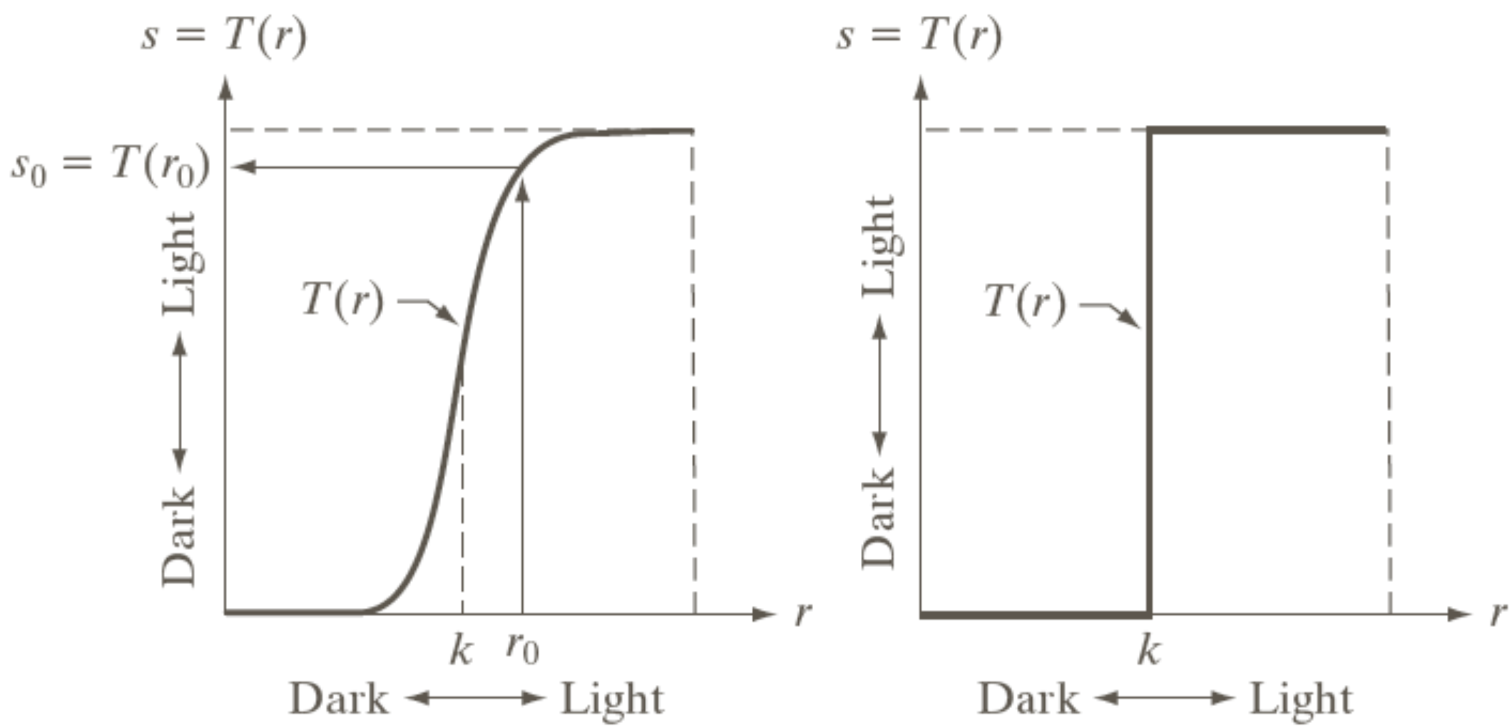
Intensity transformation example

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



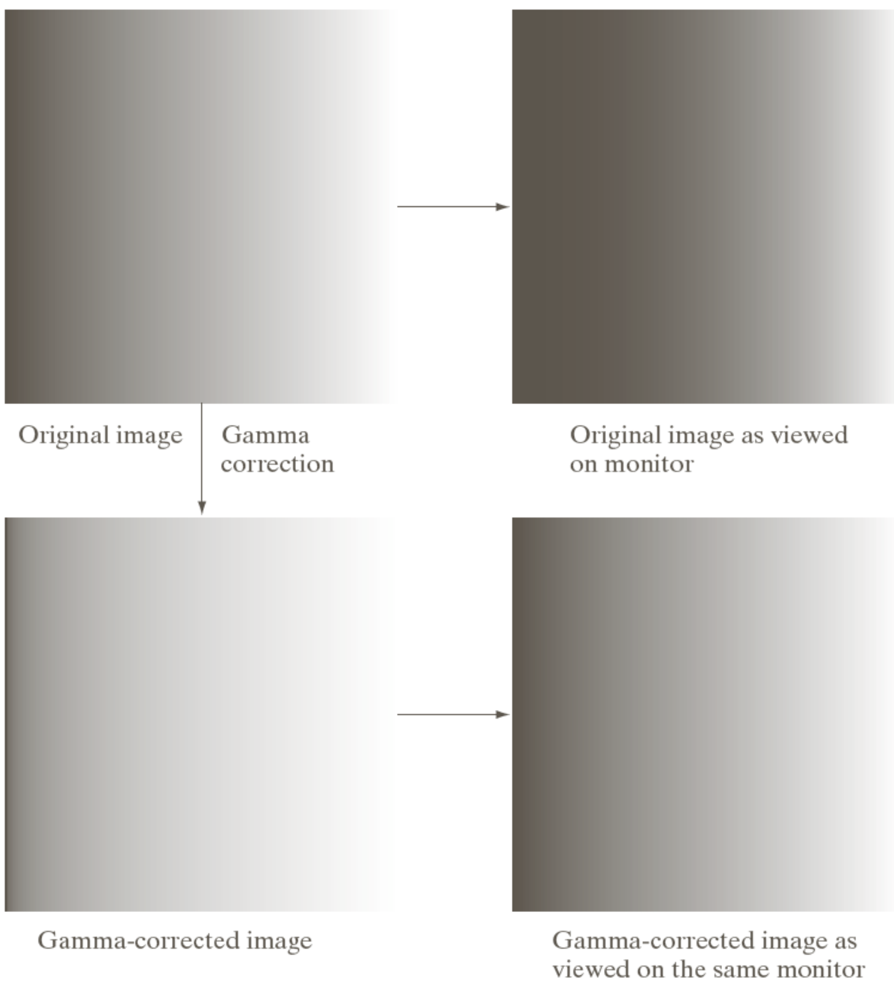
- We can drop the (x,y) and represent this kind of filter as an intensity transformation $s=T(r)$. In this case $s=\log(r)$
  - s: output intensity
  - r: input intensity

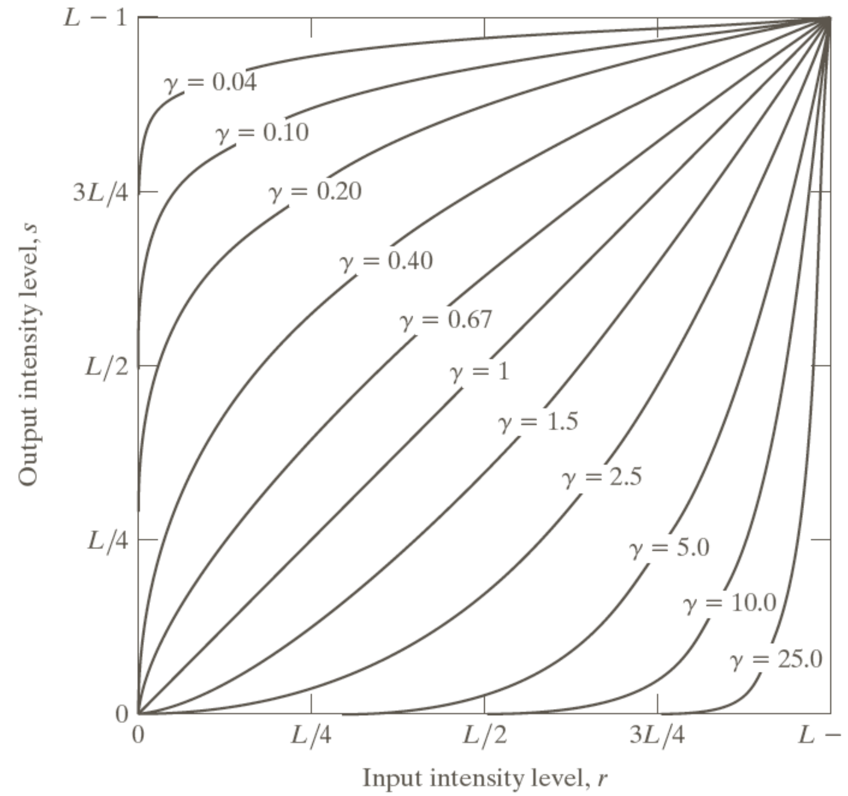
Intensity transformation



$$s = T(r)$$

Gamma correction





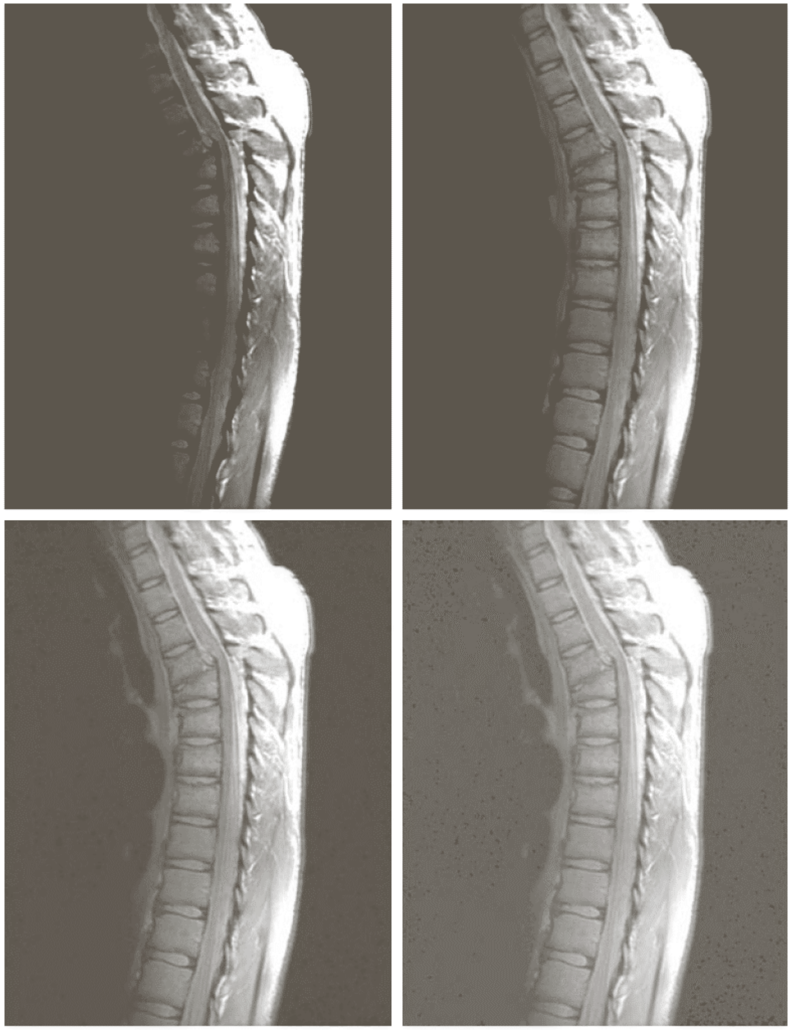
$$s = cr^{\gamma}$$

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

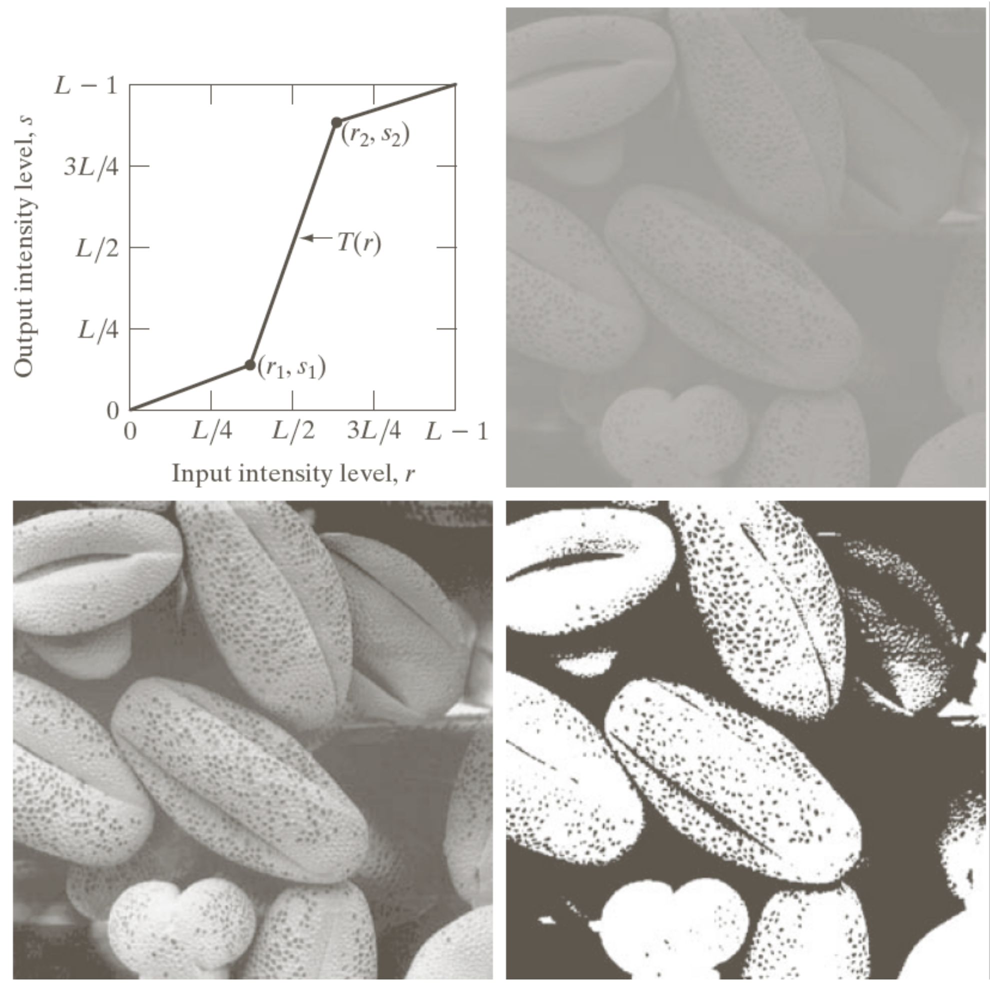


a b
c d

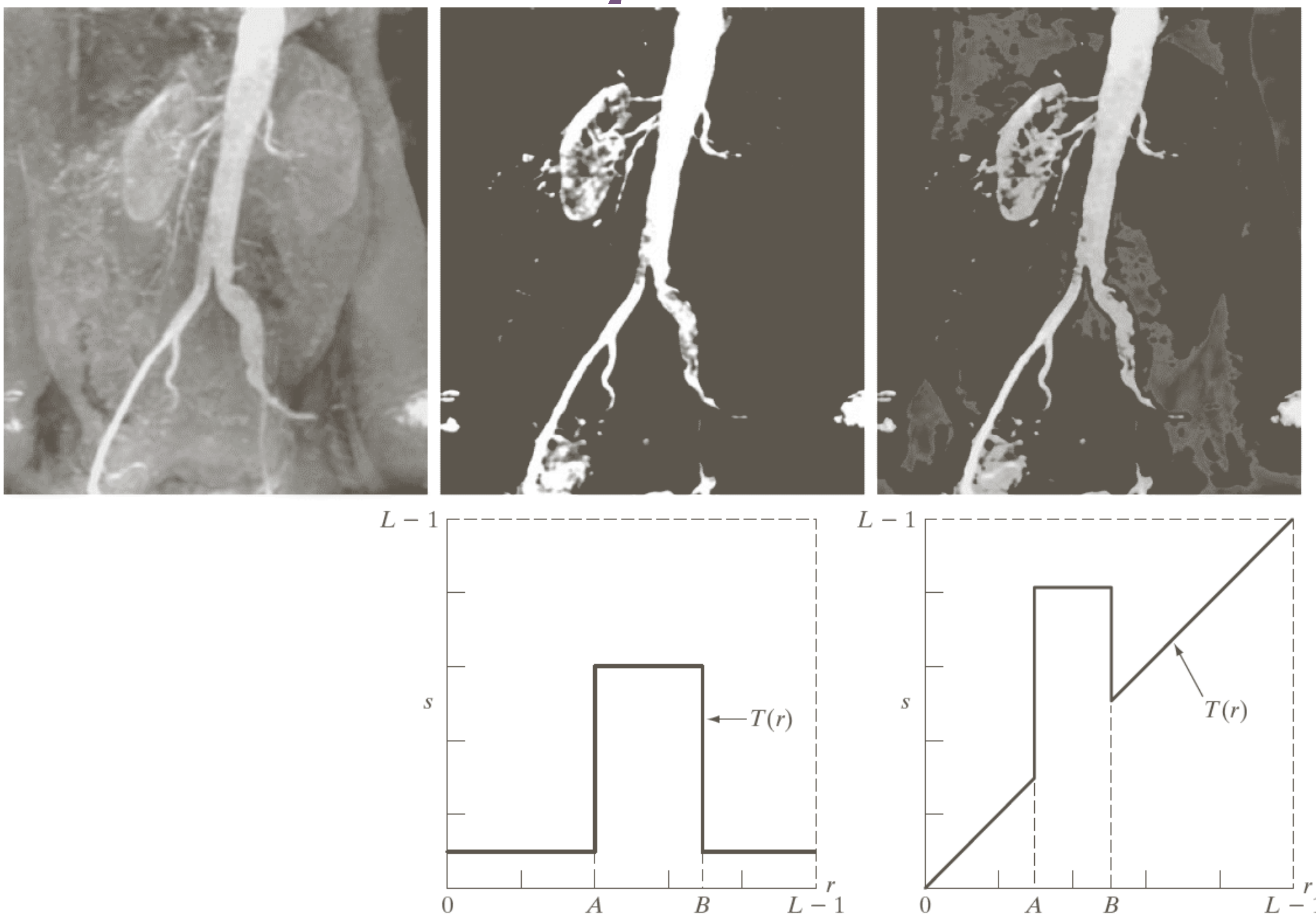
FIGURE 3.8 (a) Magnetic resonance image (MRI) of a fractured human spine. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4$, and 0.3, respectively. (Original image courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt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

Piecewise linear intensity transformation

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

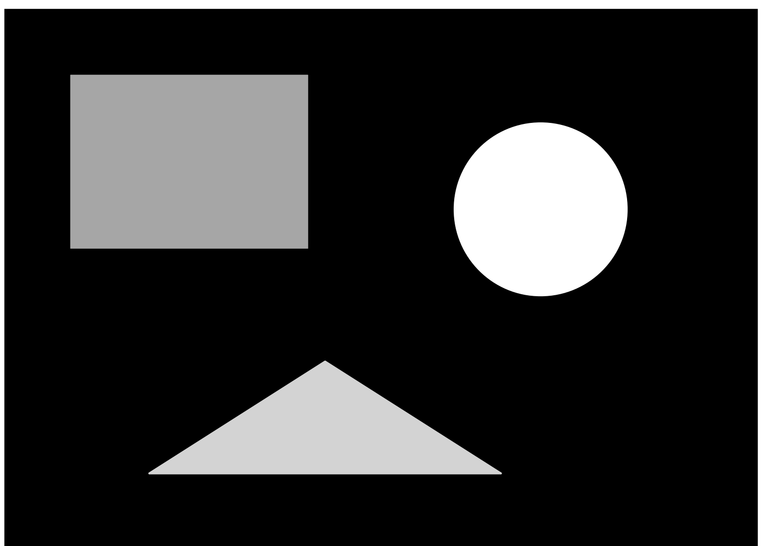


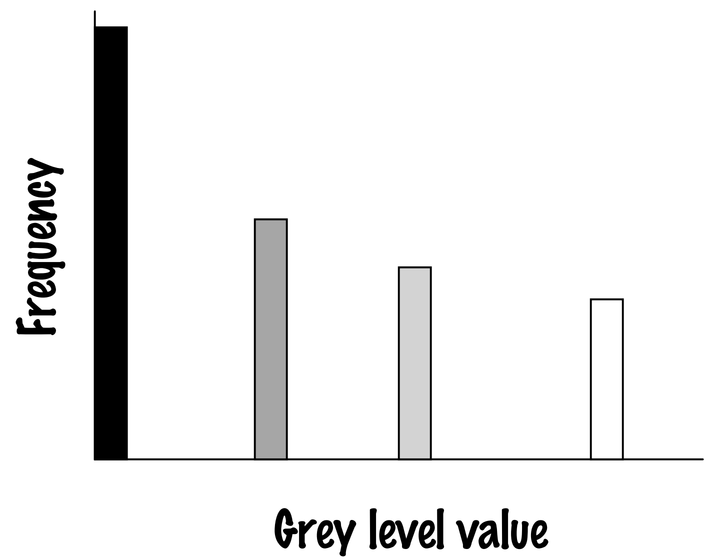
More intensity transformations



Histogram of Image Intensities

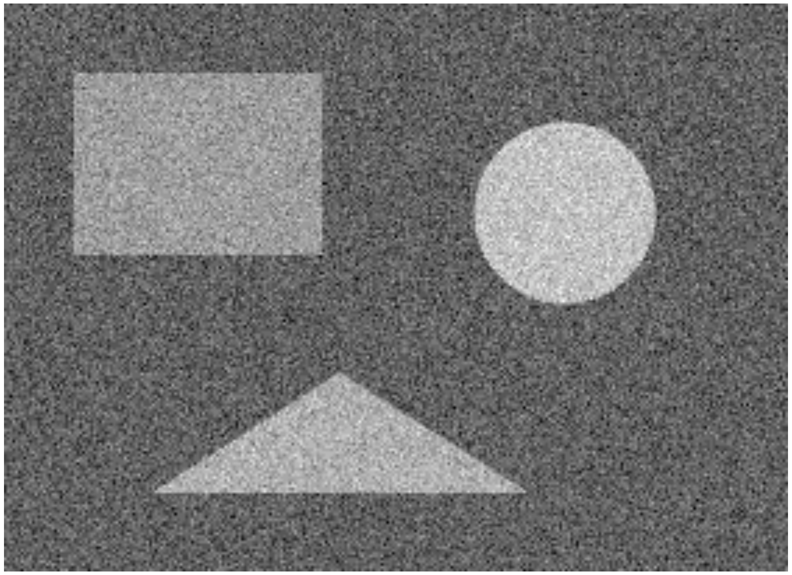
- Create bins of intensities and count number of pixels at each level
  - Normalize or not (absolution vs % frequency)

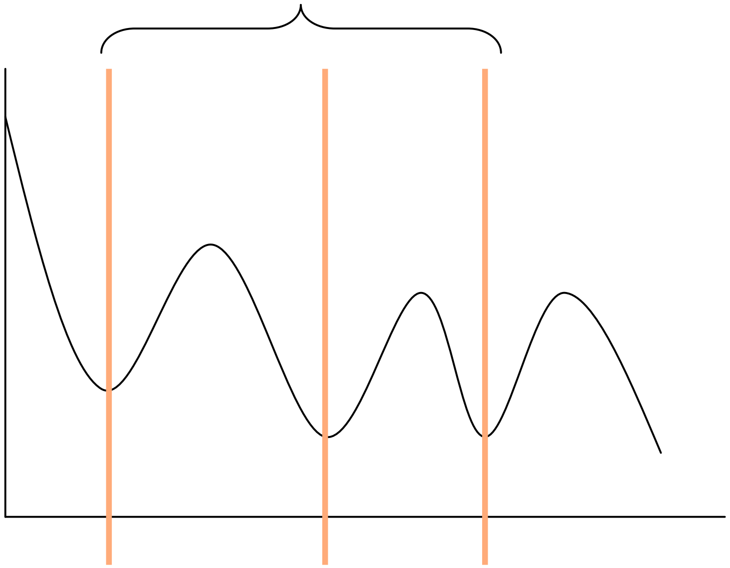




Histograms and Noise

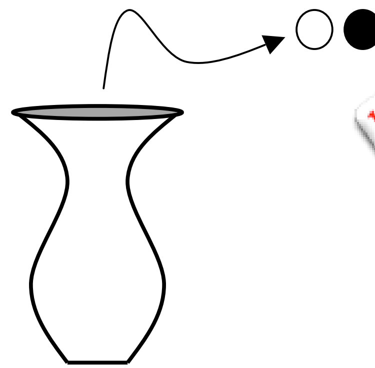
- What happens to the histogram if we add noise?
  - $g(x, y) = f(x, y) + n(x, y)$

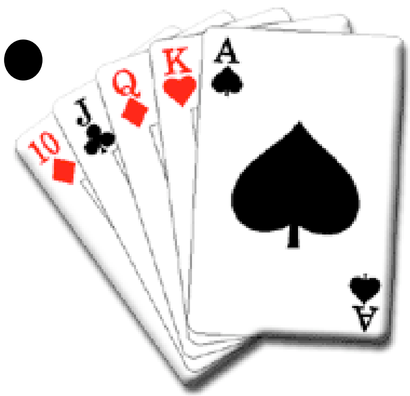




Sample Spaces

- 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

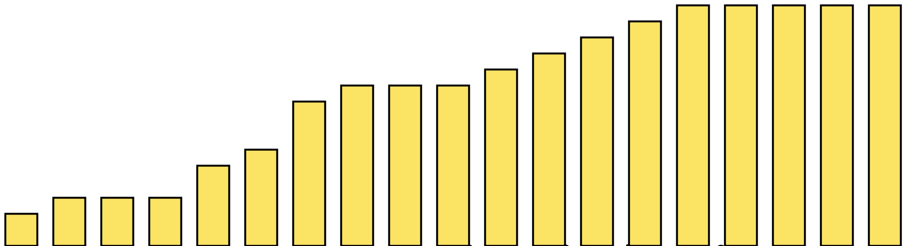
Random Variable (RV)

- Variable (number) associated with the outcome of an random experiment
- Dice
  - E.g. Assign 1-6 to the faces of die
- 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)

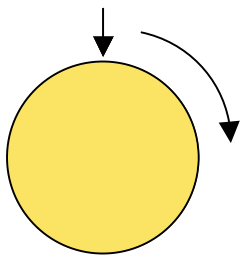
- F(x), where x is a RV
- F(-infty) = 0, F(infty) = 1
- F(x) non decreasing

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

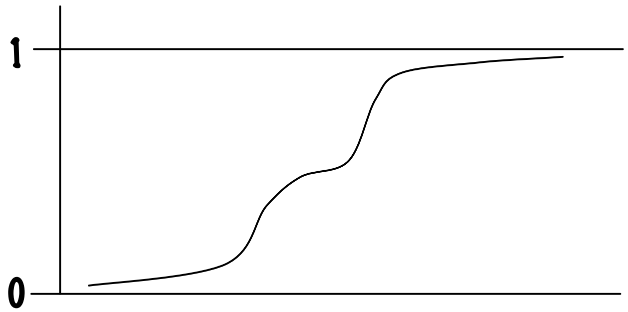


Continuous Random Variables

- 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 <u>not</u> 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 "μ"
  - The mean is not a random variable–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 "σ^2"
  - Standard deviation is σ
  - If a distribution has zero mean then: $E[x^2] = \sigma^2$

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

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

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

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 = σ

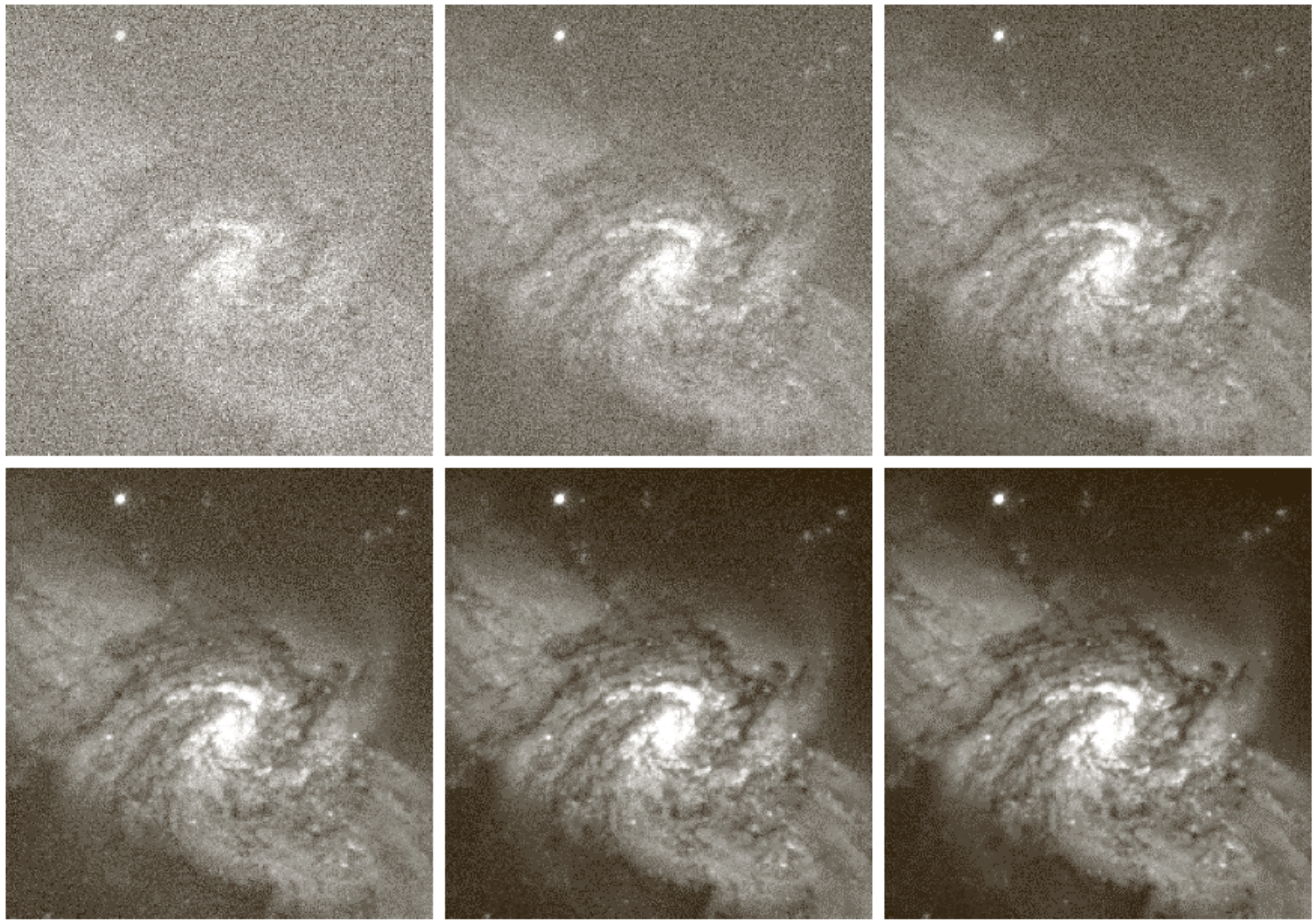
Application: Noisy Images

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

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

Averaging Noisy Images Can Improve Quality

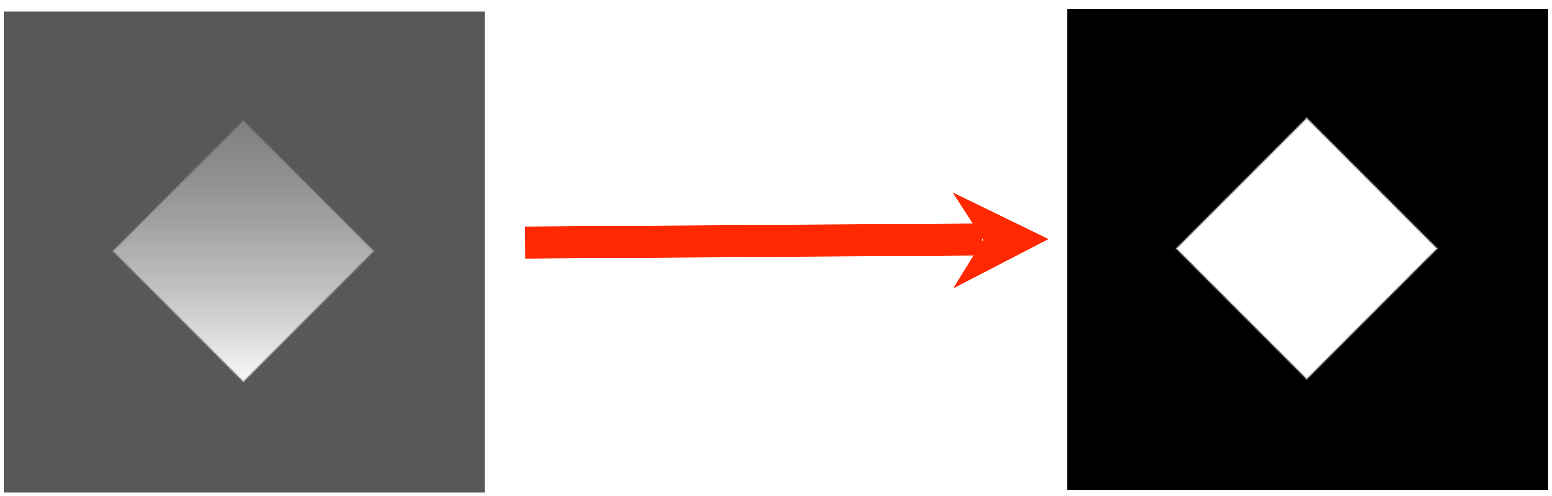


a b c
d e f

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.)

What is image segmentation?

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

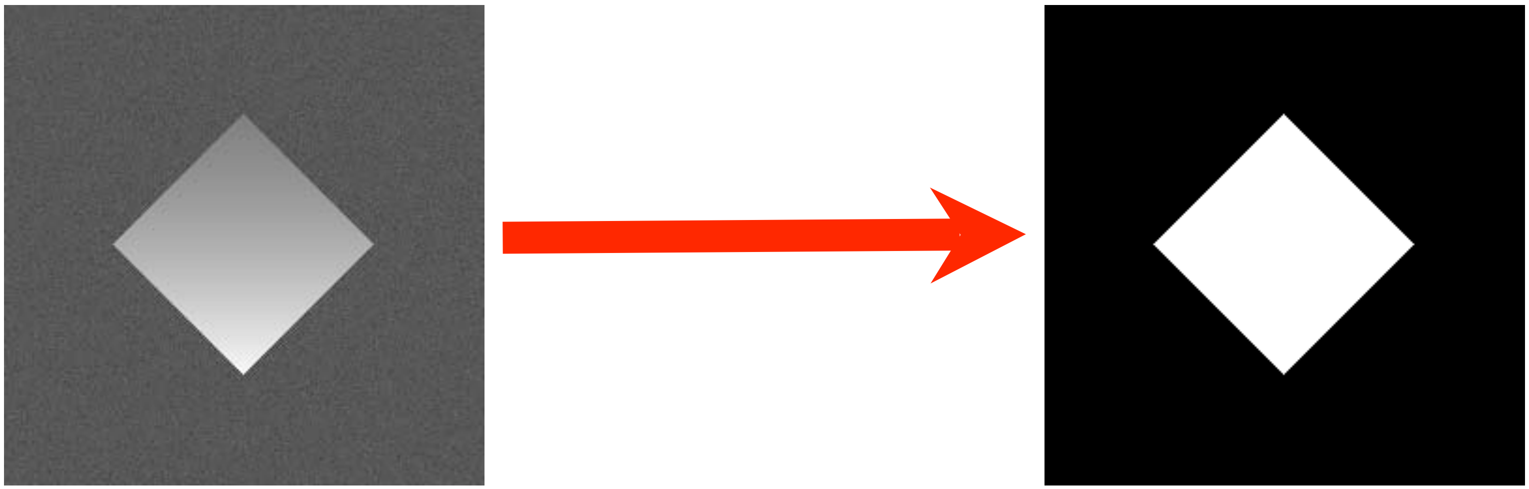


Input image
intensities 0-255

Segmentation output
0 (background)
1 (foreground)

Thresholding

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

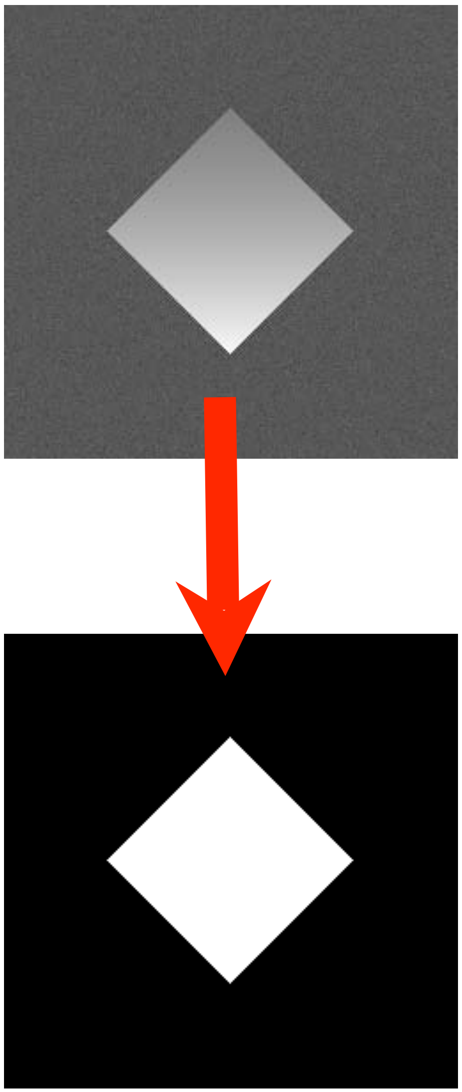


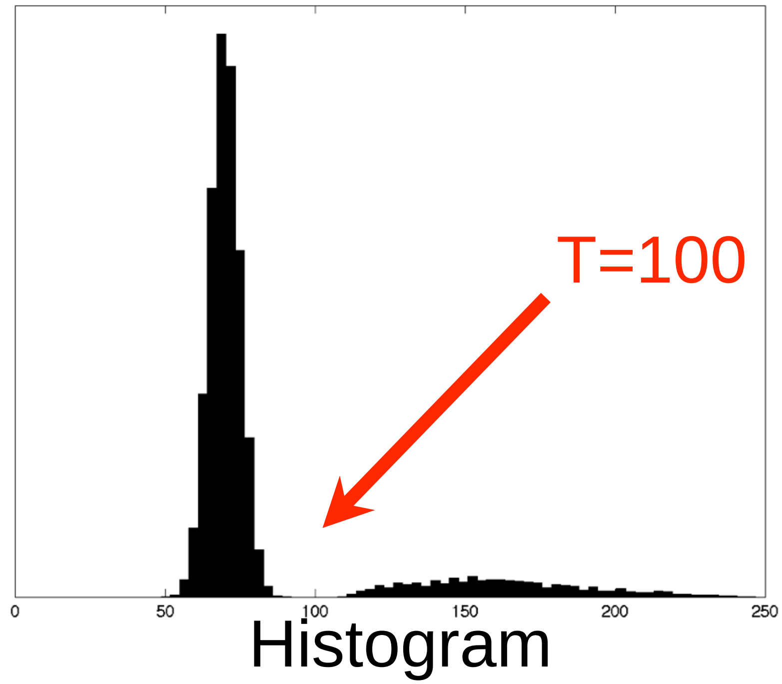
Input image f(x,y)
intensities 0-255

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

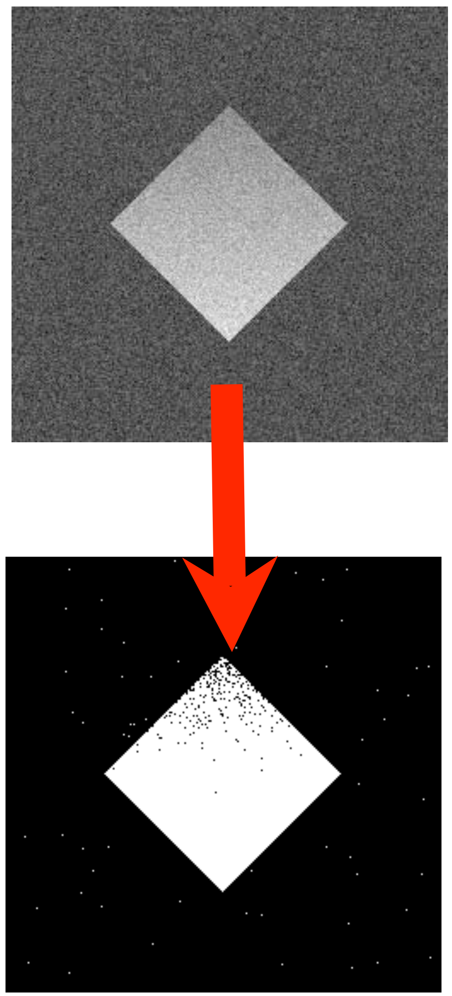
- How can we choose T?
  - Trial and error
  - Use the histogram of f(x,y)

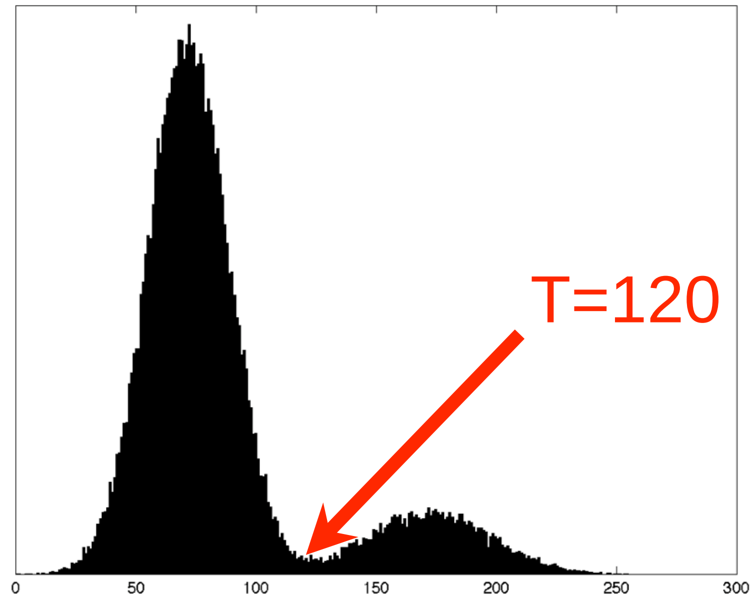
Choosing a threshold



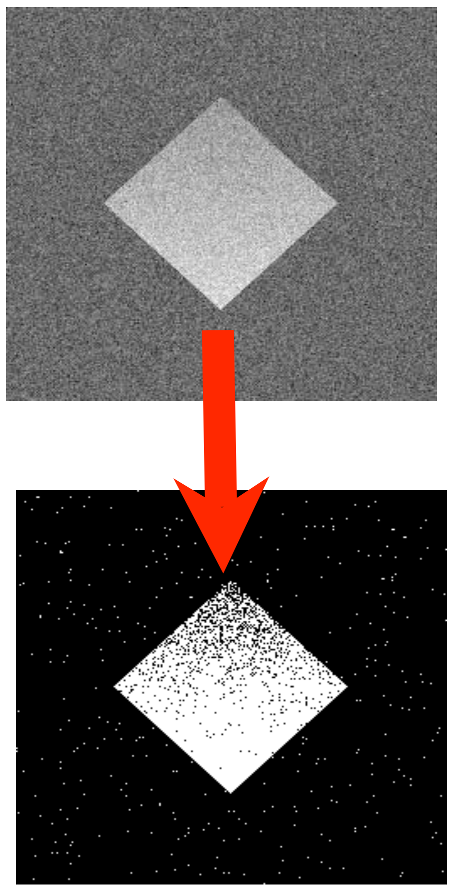


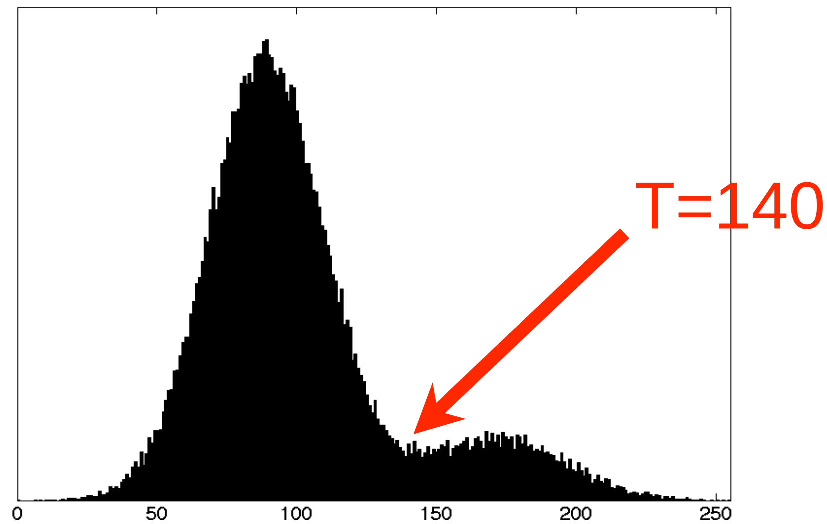
Role of noise





Low signal-to-noise ratio

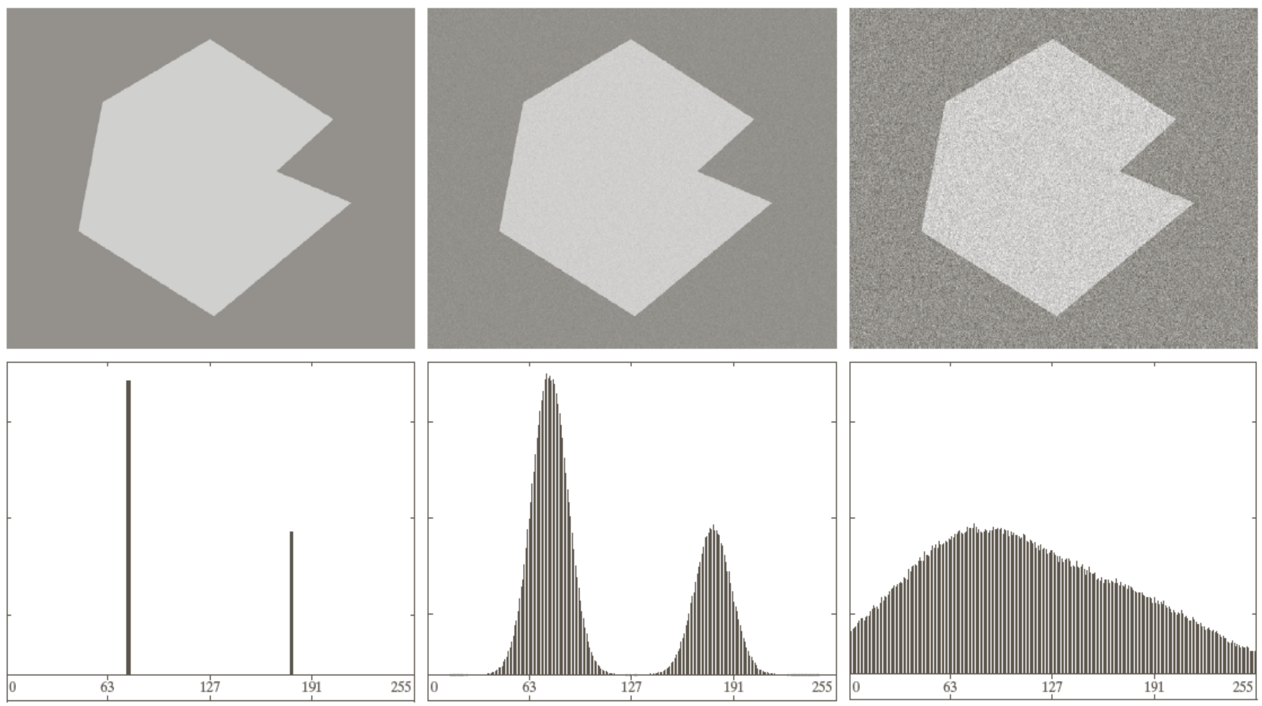




Effect of noise on image histogram

Images

Histograms

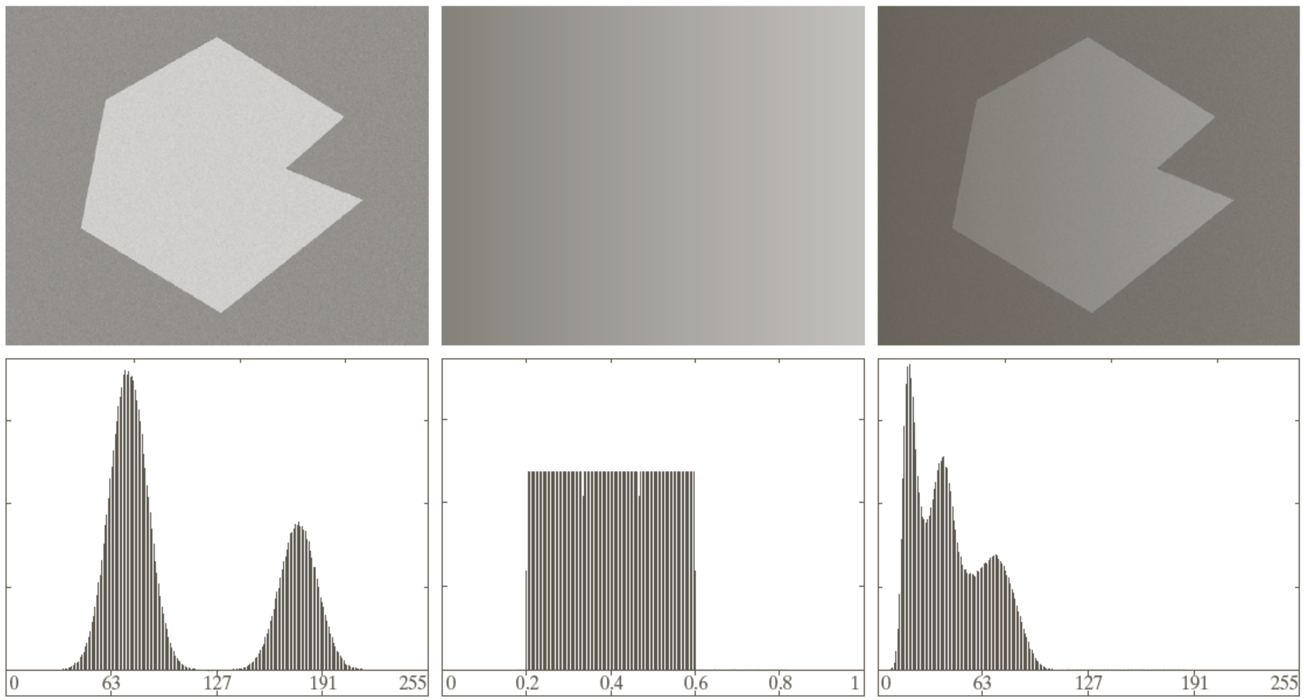


No noise With noise More noise

Effect of illumination on histogram

Images

Histograms



f x g = h

Original image

Illumination image

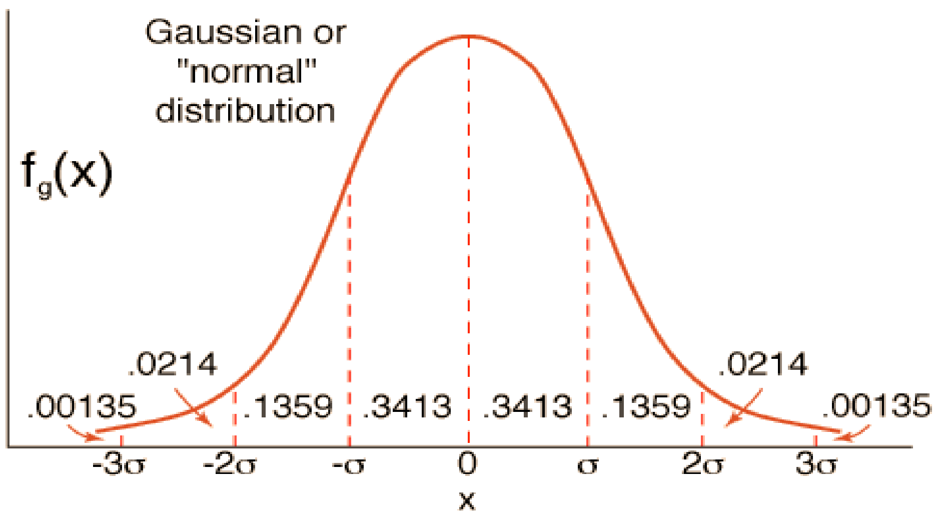
Final image

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 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

Weighted Expectation 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 *i*th item belongs to this class or not - "partially belongs"
    - probability w_i, random variable r_i

Sample mean (no weights)

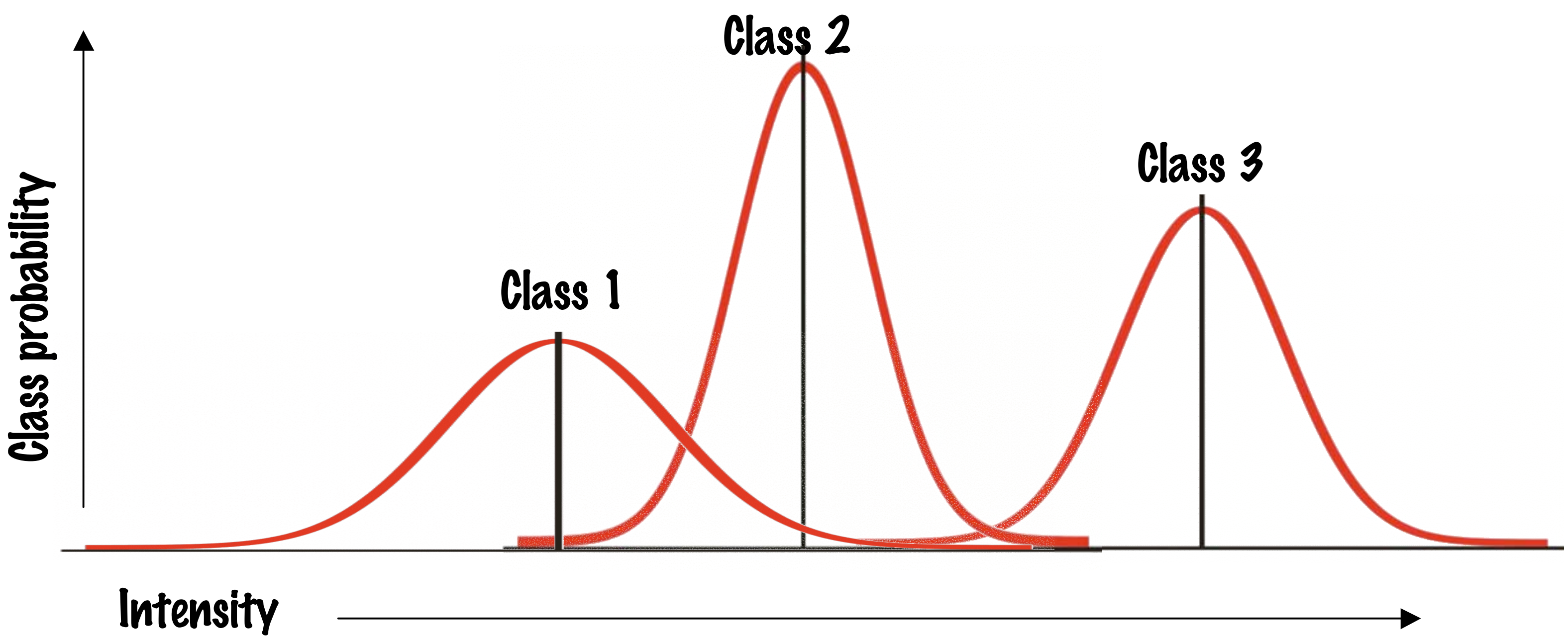
$$E[r] = \frac{1}{N}\sum_{i=1}^{N} r_i$$

Weighted sample mean

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

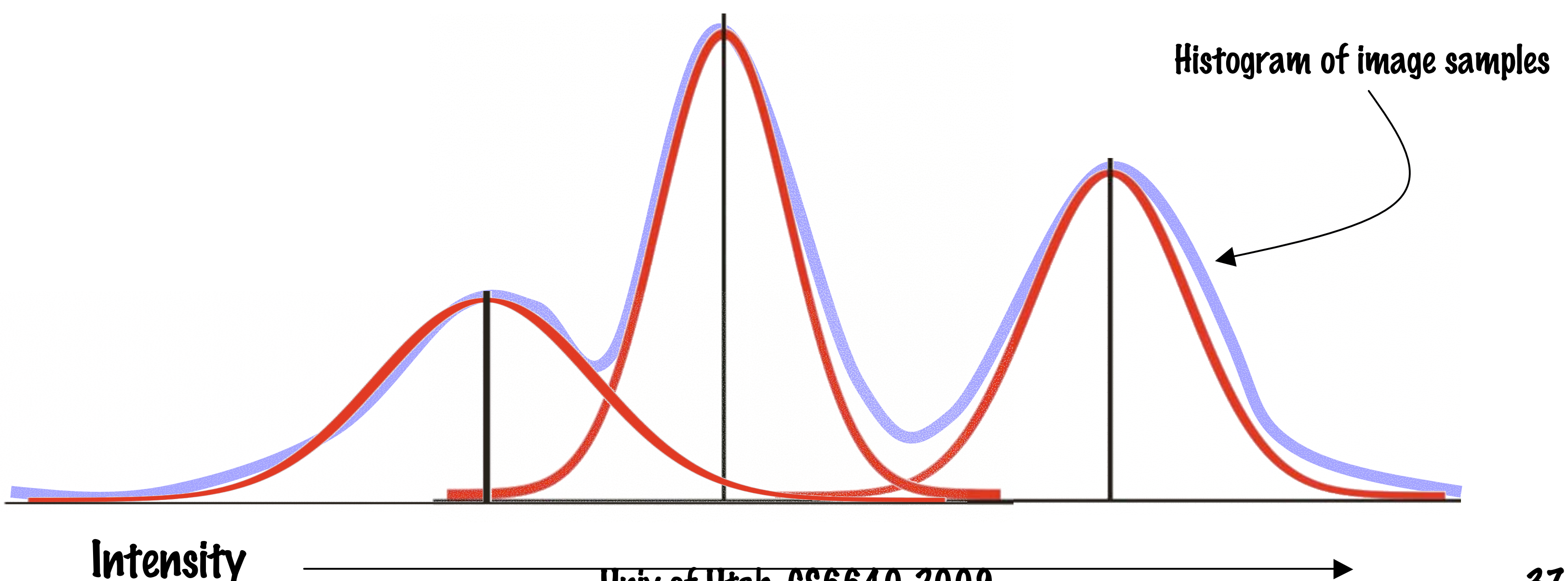
Gaussian Mixture Modeling of Image Histograms

- K classes, N samples



Problem Statement

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



Hard vs Soft Assign

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

Hard 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

Simultaneous Estimate of Class Probabilities and Pixel Labels – Iterative Algorithm

- Start with initial estimate of class models

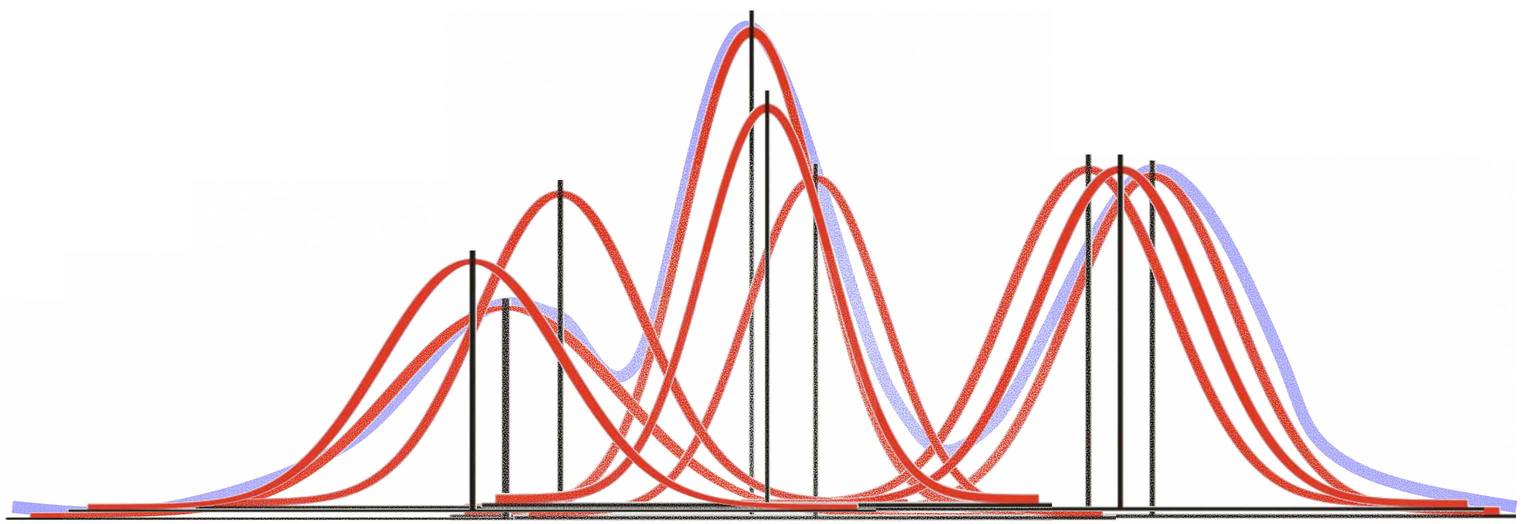
$$\mu_j^0, \sigma_j^0 \text{ for } j = 1 \ldots K$$

- Compute matrix of soft assignments

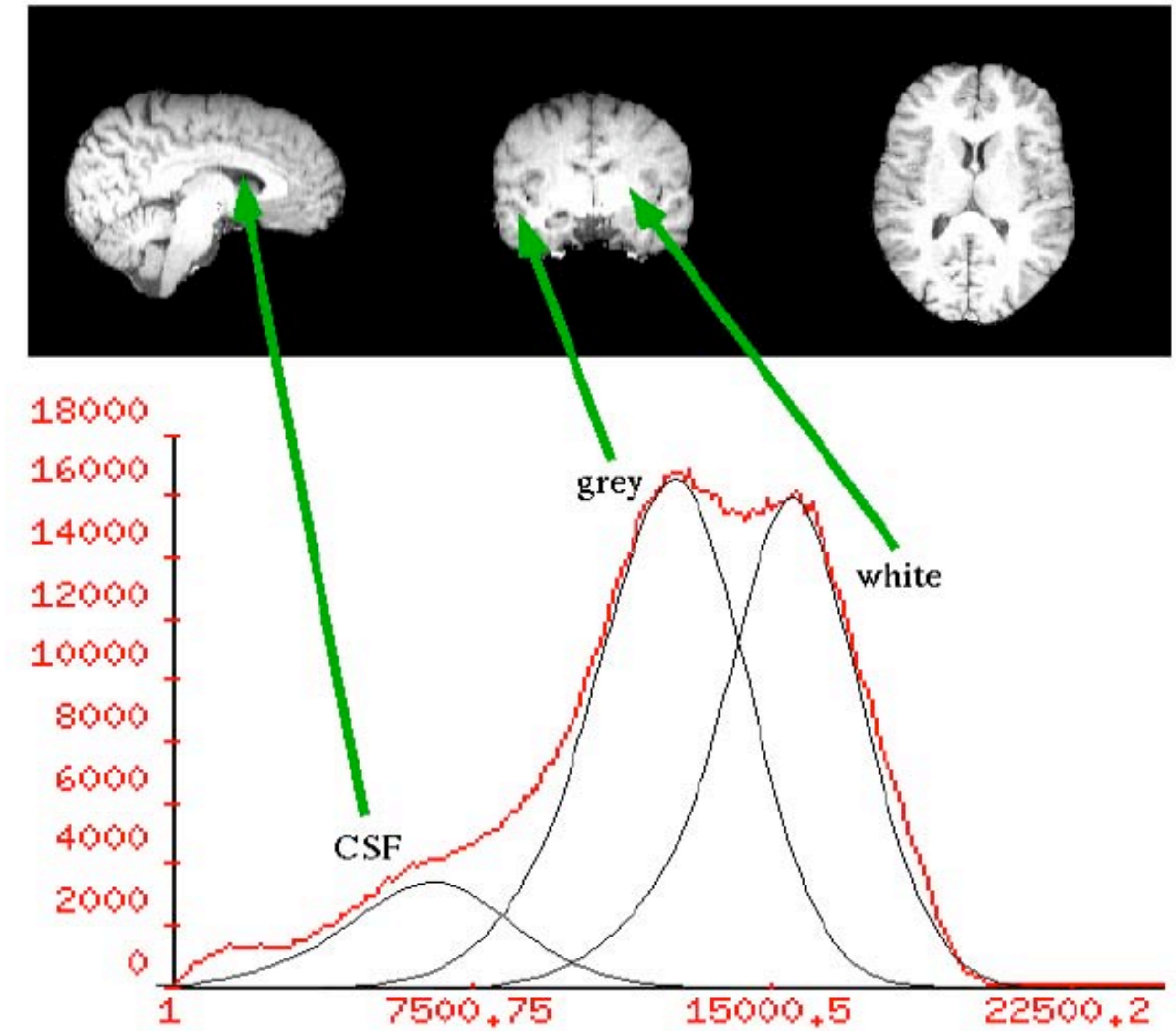
$$w_i^j = \frac{1}{\sum_{l=1}^{K} P_l(r_i)} P_j(r_i)$$

- Use soft assignments to compute new <u>weighted</u> 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

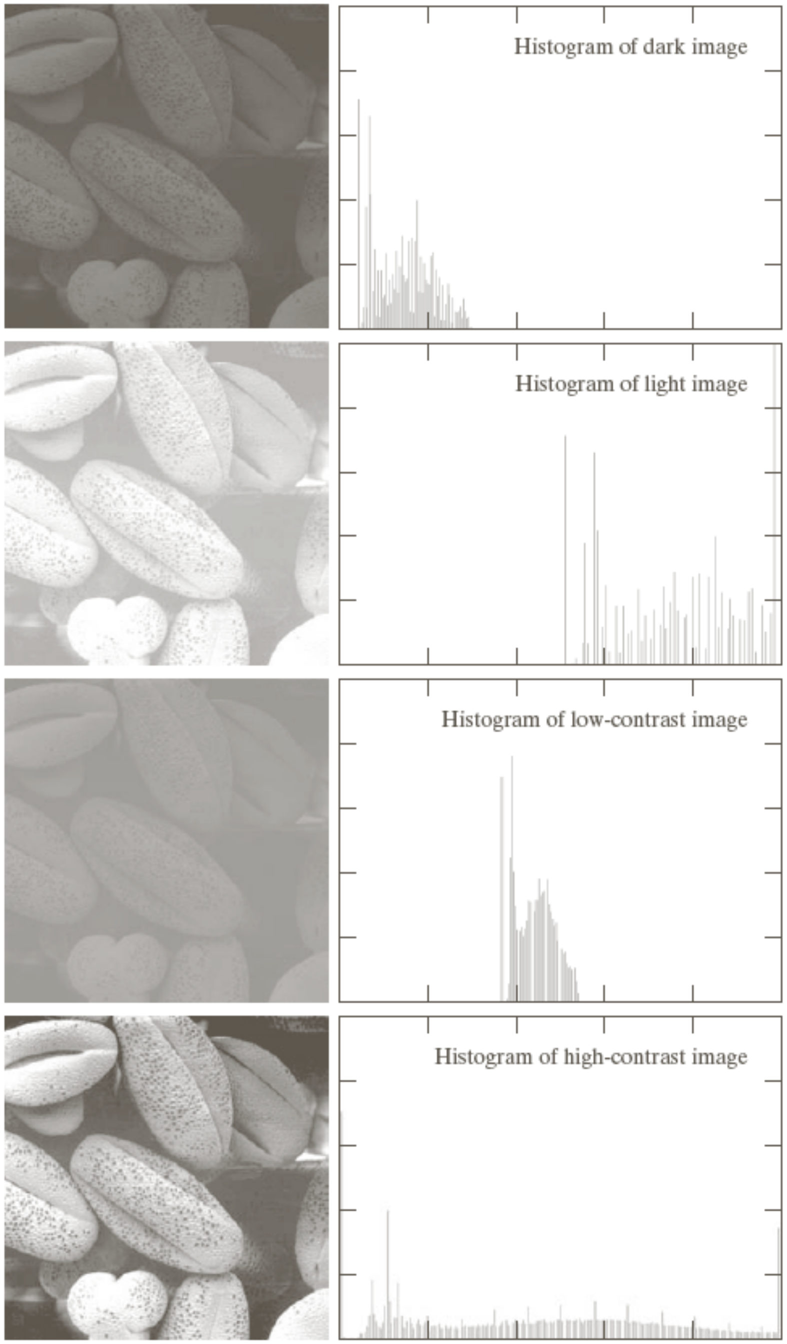


Histogram Processing and Equalization

- 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

Tuning It Down

- 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

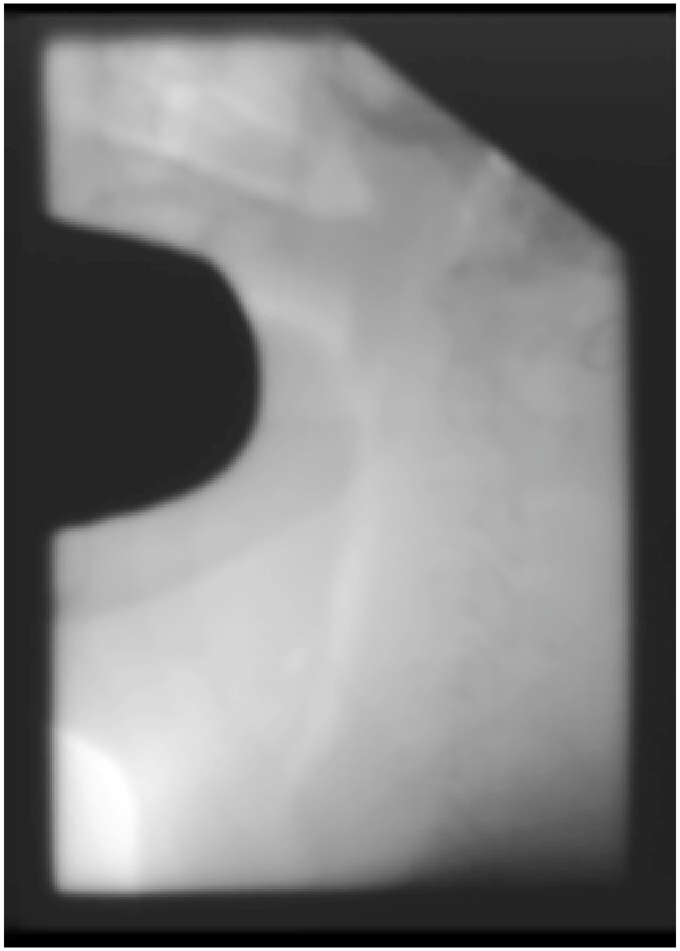
α = 0.4

α = 0.6

α = 0.8

α = 1.0

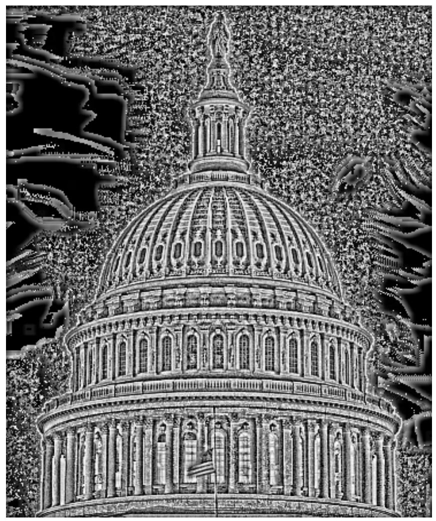
Adaptive Histogram Equalization



AHE Gone Bad...

Effect of Window Size

Orig



10x10

25x25

50x50

AHE Application: Cell Segmentation