Introduction to Digital Image Processing

Class overview and procedures

General info in Syllabus, Handout #1 and on web
→ WEB SITE: http://www.stanford.edu/class/ee168

Class materials
*Text: Umbaugh, Computer Vision and Image Processing
Handouts: Usually every lecture, will supplement material in text. Also contain more specialized topics (2nd half of course)
Computer resources:
- can use any machine you like for homework
-
- Leland accounts augmented by 2000 Mb when registered
- Lab exercises on class web page
- homework handed in digitally when applicable
- should know programming basics (CS 106 or equiv.)

Grading
Homework/Lab exercises: 30-35%
Midterm: On or about Second week in Feb., 25%
Final project and report: Class presentations: Week 10 / Wrtten report: Monday, Finals week } 40%

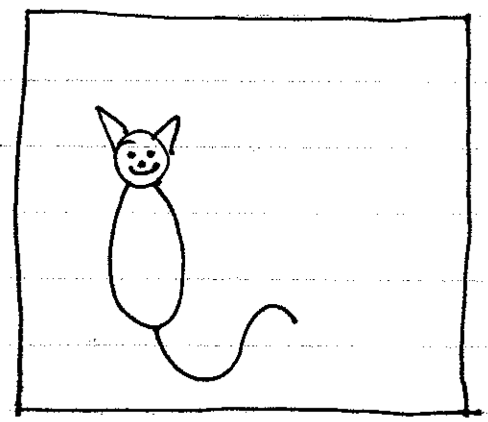
Other/extra credit: 0-5%

Student backgrounds
Courses, majors
Computer familiarity

*Text recommended only

What are images?

What do we mean by "images" and "imaging"? Consider an "image" of a cat:



What do each of the following have in common? How are they different?

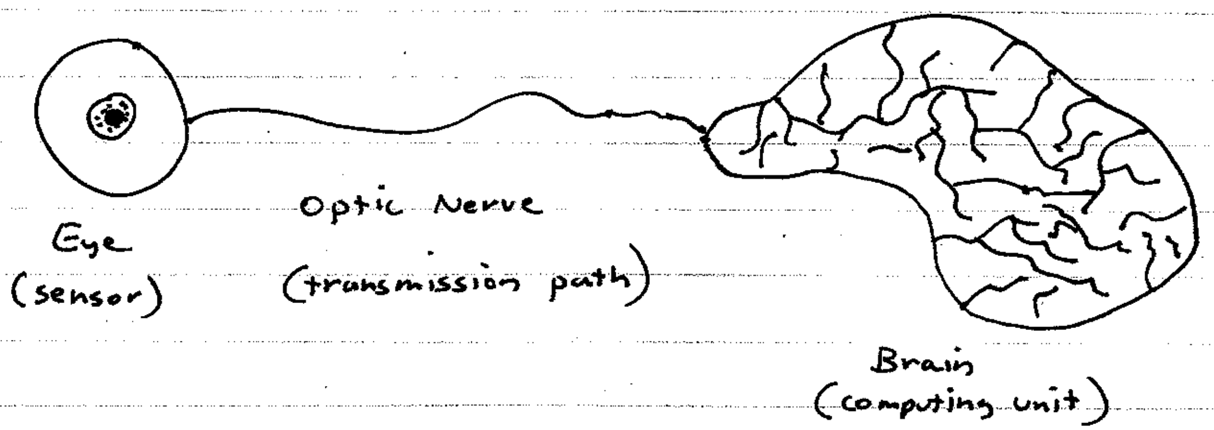
A photograph of a cat
A tv image of a cat
A picture on a friend's web page of a cat?
An MRI scan of a cat

We will define here an image as a two (or more) dimensional distribution of some quantity that can be, and usually is, displayed to be sensed by the human eye. When such a function is represented by numbers and stored/manipulated/displayed on a computer, we can refer to it as a digital image. In this course we will investigate the creation and manipulation of digital images.

Organic image processing – The human visual system

Since many images are intended to be displayed for our eyes to view, we need to understand some properties of the human visual system. Our eyes have certain physiological properties that drive the optical processing system in the brain, where images are eventually interpreted.

System diagram:

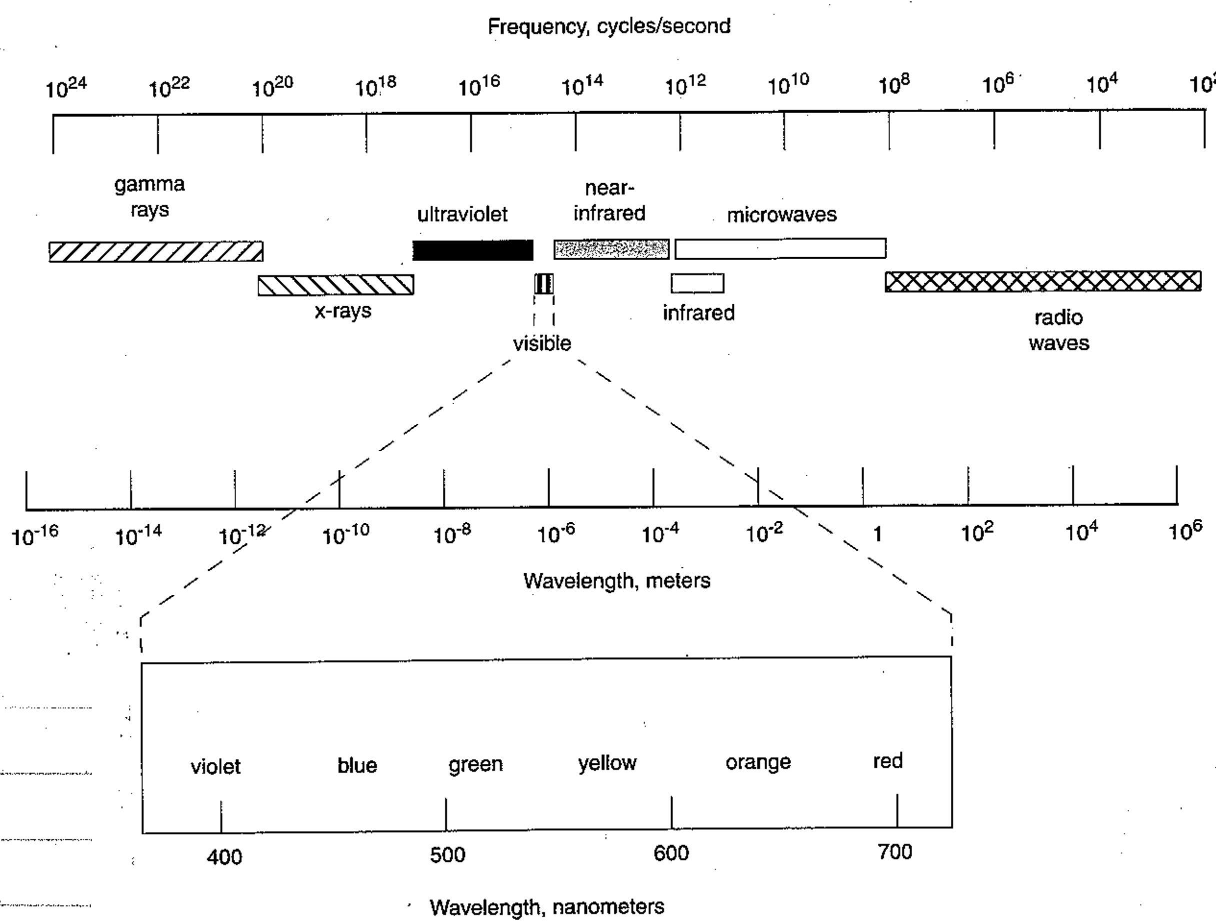


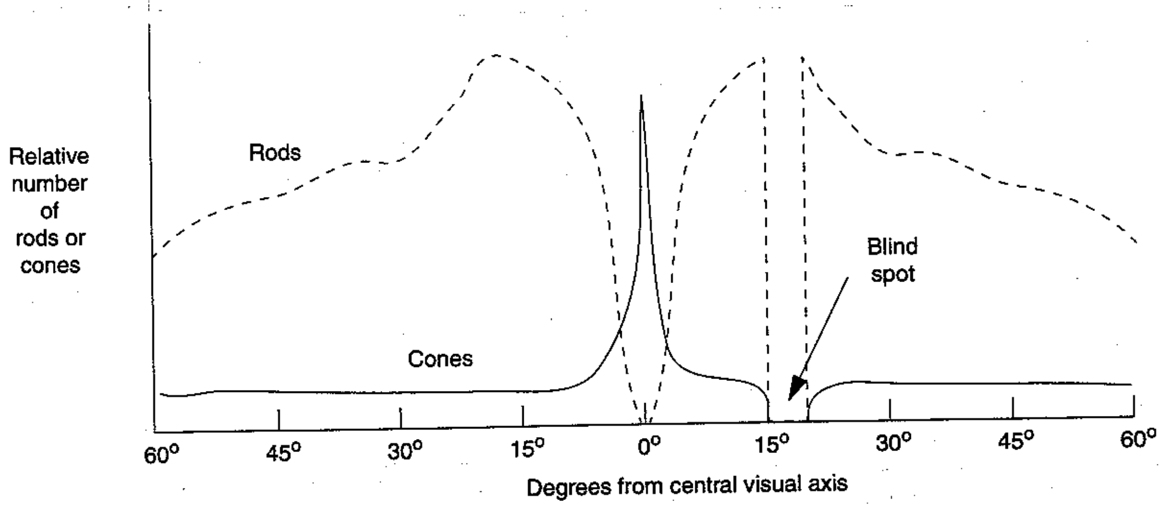
The eye is the sensor that converts an angular distribution of incident light to signals that can be transmitted via the optic nerve to the brain. The eye is sensitive to a limited portion of the electromagnetic spectrum called the visible spectrum, and can detect light with a wavelength ranging from about 0.38 μm – 0.70 μm. We sense the shortest wavelengths as blue in color, and the longest as red.

The actual photoreceptors in the eye are two specialized nerve cells called rods and cones. The cones, used mostly for daylight vision, are concentrated in the back of the eye, and come in three types, sensitive to red, green, and blue wavelengths, respectively.

The cones are also very densely packed, so they can distinguish very fine details. The rods, on the other hand, are less densely packed, insensitive to color, but can detect very low light levels. They are used mainly at night. This is why we can't easily distinguish colors at low light levels.

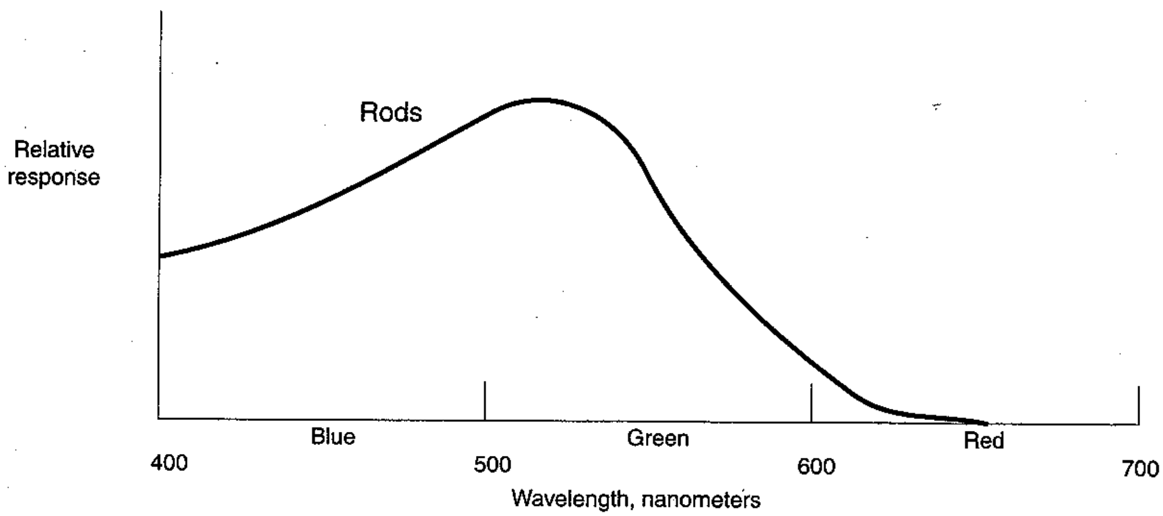
The Electromagnetic Spectrum



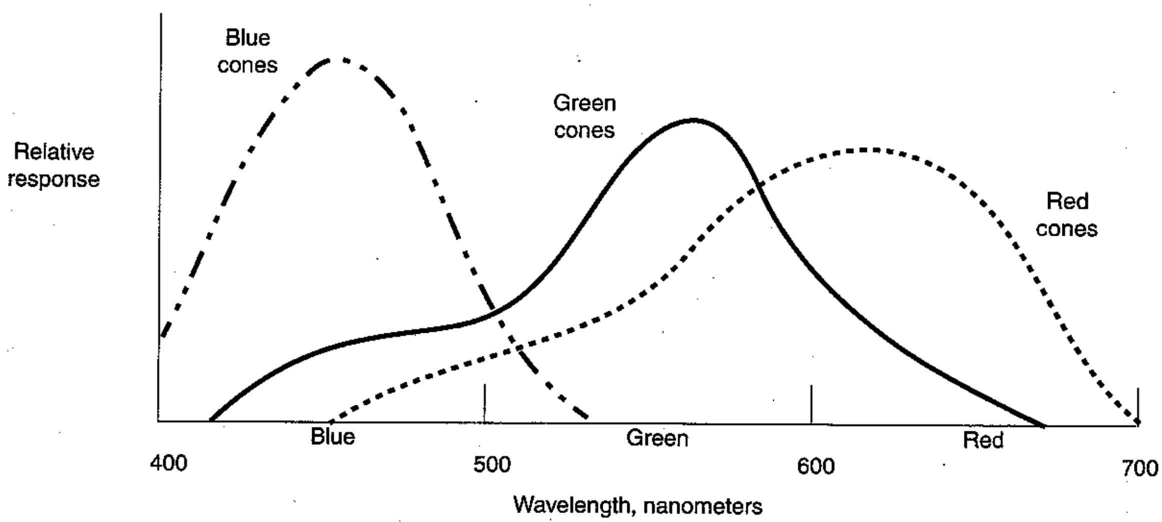


Concentration of rods and cones across retina.

Relative Responses of Rods and Cones



a. Rods react even in low light levels but see only a single spectral band; they cannot distinguish colors.



b. Cones react only to high light intensities; the three types enable us to see colors.

Once the data are sensed by the eye, they are sent to the brain and interpreted as useful information. While the brain can simply report how intensity or color changes as a function of angle, it often adds in ancillary information to aid in understanding of images. Sometimes the brain can be fooled into mis-interpreting information. We usually call this effect an <u>optical illusion</u>.

Methods of representation

We have said that an image is a two-d distribution of a quantity, such as brightness. Represent the image as a function of two variables $f(x,y)$ (we'll use Cartesian coordinates for now).

For a photograph, $f(x,y)$ varies from complete reflectance ($f=1$) to completely black ($f=0$).

How white is white?

paper ~0.8

snow ~0.9 (very white snow)

Black paper: ~0.1 - 0.2 (sometimes less)

The moon has an average ~~att~~ albedo (reflectance) of just under 10%, hence it would appear black if not surrounded by very dark ($f \approx 0$) space!

We call the range of reflectances the contrast ratio or dynamic range of an image. Thus for photographs a range of $^{0.8}/_{0.1}$ or 8:1 is attainable. As EE's we express ratios as dB, 8:1 gives a 9 dB dynamic range.

In principle, the 8:1 or 9 dB of dynamic range can be divided up any way we please, so that a photo may be used to convey an unlimited amount of information. But we know that all real signals contain "noise", and that will limit the number of independent levels we can express in an image.

For example, suppose that the noise or uncertainty in a brightness measurement from the above example is about 10% in reflectance. Then the total range of 10% to 80% results in 7 "grey levels" achievable in the image:

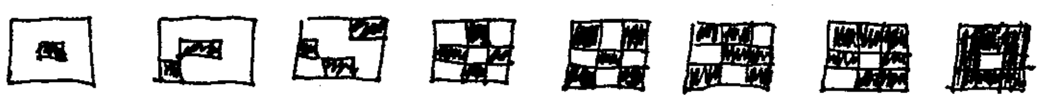
$$\frac{80\% - 10\%}{10\%} \quad \text{or} \quad \frac{\text{Max reflectance} - \text{min reflectance}}{\text{uncertainty in reflectance}}$$

→ How many gray levels in principle are representable using an 8-bit computer display?

Half-toning

A laser printer can only write white or black dots on a piece of paper. How then can we form images using them? The dot density and placement can be used to generate in-between shades of grey.

Example: 8 grey levels in a 3x3 box of dots

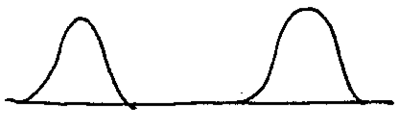


A random statistical method works also, with dot on probabilities set by reflectance values.

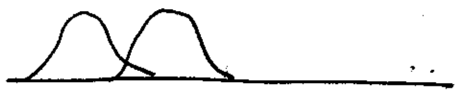
Magazines and newspapers vary the dot sizes in addition to the densities.

Resolution

Another very important image property is resolution, or how fine an object may be distinguished in the image. We can think of resolution as how close two objects may be and yet be interpreted as separate entities. A 1-D example:



Two easily separable objects



Somewhat separable objects

Not easily separable

Often we use the width of an object at its half-power point (the 3-dB width since $-3\,dB \approx \frac{1}{2}$) as a measure of resolution:

← point response, impulse response

Some resolutions of display systems:

Photographic emulsion:	1 μm grain size (binary)
Laser printer:	600-1200 dots/inch (40 μm)
Computer monitor:	0.25 mm dots size (8 bit brightness)

For optical systems we often use angular resolutions, where angular resolution $\Delta\theta$ is related to linear resolution Δl

by $\Delta\theta = \frac{\Delta l}{\rho}$, ρ is the imaging distance.

The human eye can distinguish $\Delta\theta \approx \frac{1}{3000}$ radian $\approx \frac{1}{50}$ degree

Atmosphere-limited telescope $\Delta\theta \approx 10^{-6}$ radian $\approx \frac{1}{20000}$ degree

Theoretical Palomar telescope $\Delta\theta \approx 10^{-7} m \approx \frac{1}{200000}$ degree

For an atmosphere-limited telescope, what can we see on the moon?

$$10^{-6} \text{ radian} \times 400000 \text{ km} = 400 \text{ m resolution}$$

How many "pixels" form the moon at this rate?

Area of moon: $\pi \cdot a^2 = \pi \cdot (1738 \text{ km})^2 = 9.5 \times 10^{12} \text{ m}^2$

Area of pixel: $\pi \cdot 200^2 = 1.26 \times 10^5 \text{ m}^2$

$$\# \text{ pixels} = \frac{9.5 \times 10^{12}}{1.26 \times 10^5} \approx 75{,}000{,}000 \text{ pixels}$$

A typical digital image is 1000×1000 or 10^6 pixels, so a full moon image would be quite large.

<u>Resels vs. pixels</u>

You are probably aware of the term pixel (picture element) to describe the point spacing in a digital image. How does this relate to the resolution element we have been discussing?

The resel size is the width of the impulse response, in units appropriate to the image.

The pixel size is how often the intensity (or other) distribution is sampled.

In well-designed systems, resel ≈ pixel so the storage is used efficiently without undersampling the distribution. More on this when we discuss aliasing later on.

<u>Raster displays, matrix displays</u>

How can we store and display 2-D intensity distributions? First, assume we have quantized our function $f(x,y)$ into $f(x_i, y_i)$. By knowing the line ~~l~~ length of an image we can go from an ordered list in a computer to a raster display.

Consider this sequence:

0 1 1 1 1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 1 1 1 0 0 0 0 0 1 0
0 1 0 0 1 0 0 1 1 1 1 0

This "meaningless" string becomes an instantly recognizable picture when we know the line length and display it properly. If instead of ones and zeros, we use shades of grey we can create very complex and informative pictures.

Sometimes we (or others) include "header" information with each displayed line. This may, for example, include a <u>line count</u> so that lost data may be detected, or to ensure that we can "line up" missing data.