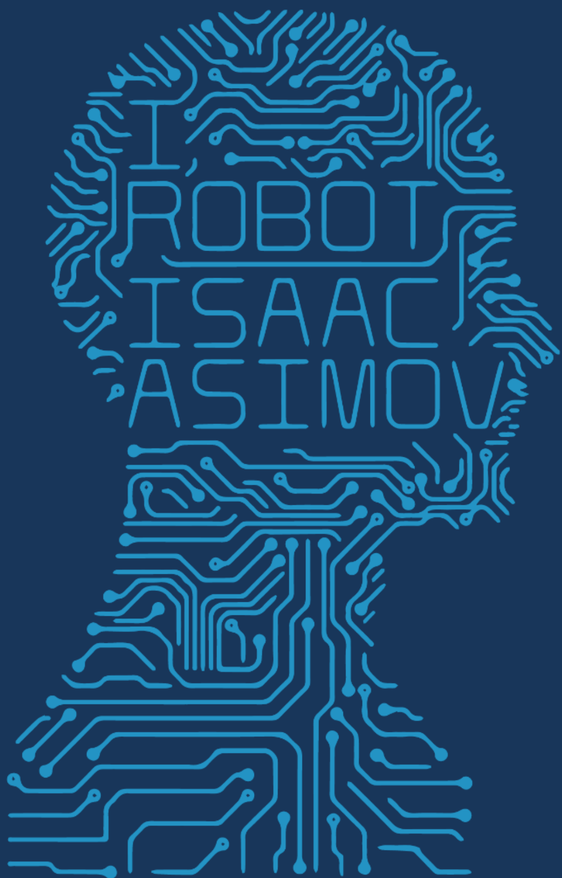
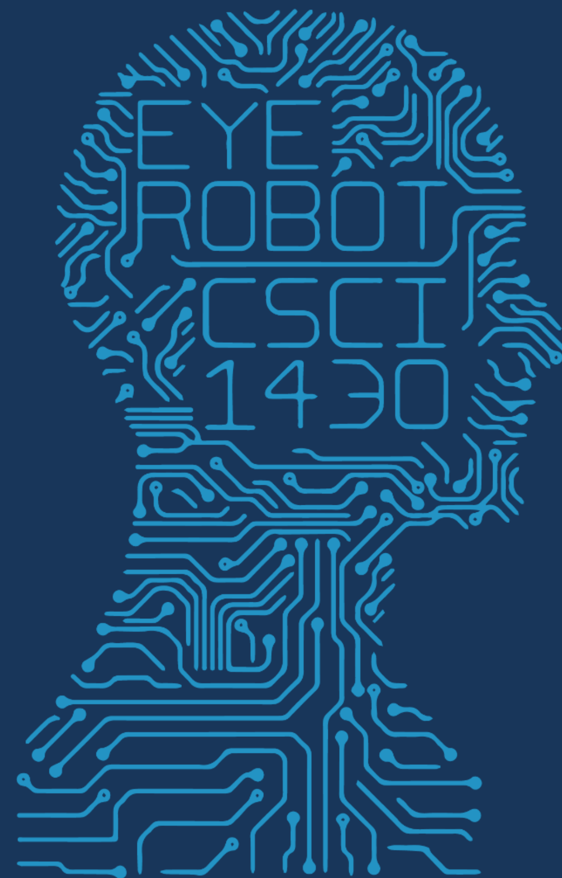


1950

Future Vision



2017 MWF 1pm 368

Computer Vision

Bela Borsodi

Bela Borsodi

Oversubscription

- Sorry, not fixed yet.
- We’ll let you know as soon as we can.

CS 143 – James Hays

- Continuing his course – many materials, courseworks, based from him + previous staff – serious thanks!

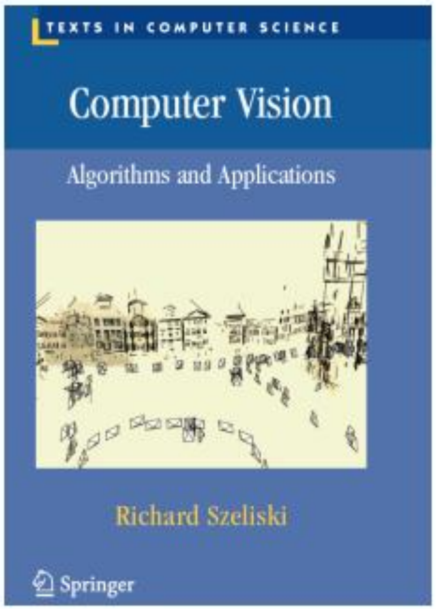
Reminder: the books

- Lectures have associated readings
- Szeliski 2.2 and 2.3 for today

Textbook

Computer Vision: Algorithms and Applications

© 2010 [Richard Szeliski](), Microsoft Research



http://szeliski.org/Book/

James Hayes

Textbook

Deep Learning

An MIT Press book

Ian Goodfellow and Yoshua Bengio and Aaron Courville

- Can I get a PDF of this book?

  No, our contract with MIT Press forbids distribution of too easily copied electronic formats of the book.

- Why are you using HTML format for the web version of the book?

  This format is a sort of weak DRM required by our contract with MIT Press. It's intended to discourage unauthorized copying/editing of the book.

- What is the best way to print the HTML format?

  Printing seems to work best printing directly from the browser, using Chrome. Other browsers do not work as well.

Class visual computing experience

- Linear algebra
- Probability

- Graphics course?
- Vision/image processing course before?
- Machine learning?

WHAT IS AN IMAGE?

First MATLAB

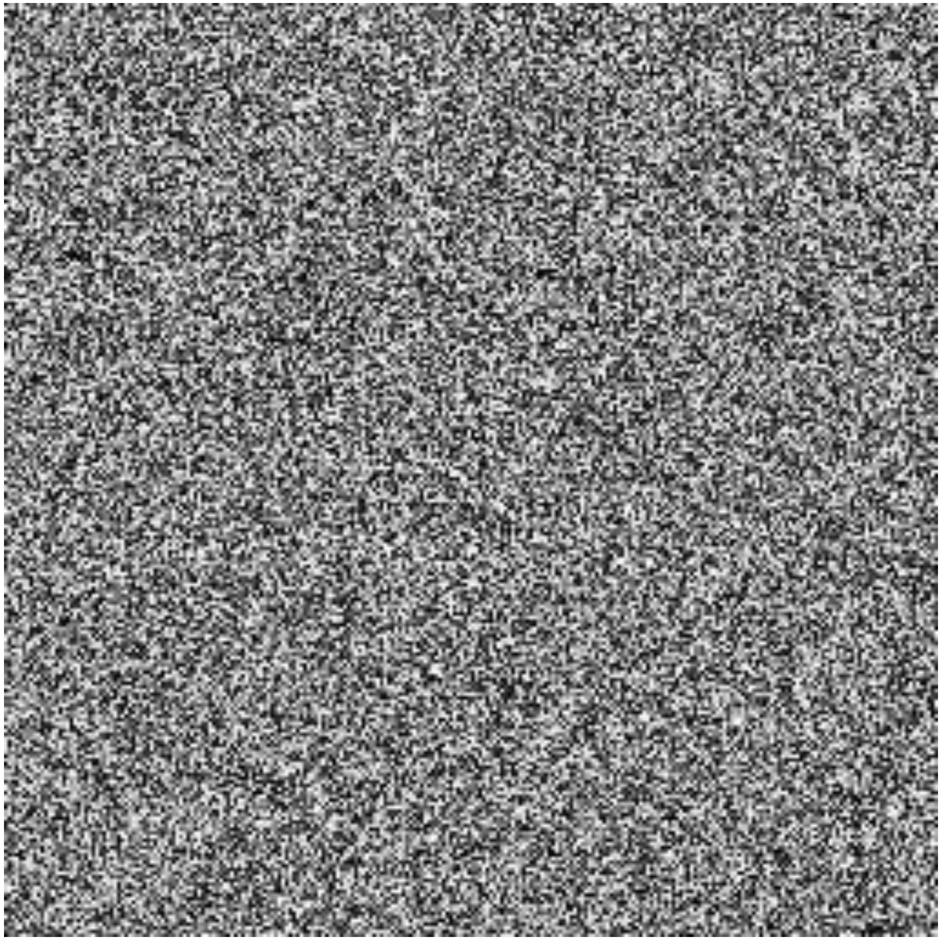
```
>> I = rand(256,256);
```

Think-Pair-Share:

- What is this?
- How many values can it take?
- Is it an image?

First MATLAB: What is this?

```
>> I = rand(256,256);
>> imshow(I);
```



Danny Alexander

Dimensionality of an Image

- @ 8bit = 256 values ^ 65,536
  - Computer says ‘Inf’ combinations.

- Some depiction of all possible scenes would fit into this memory.

Dimensionality of an Image

- @ 8bit = 256 values ^ 65,536
  - Computer says ‘Inf’ combinations.

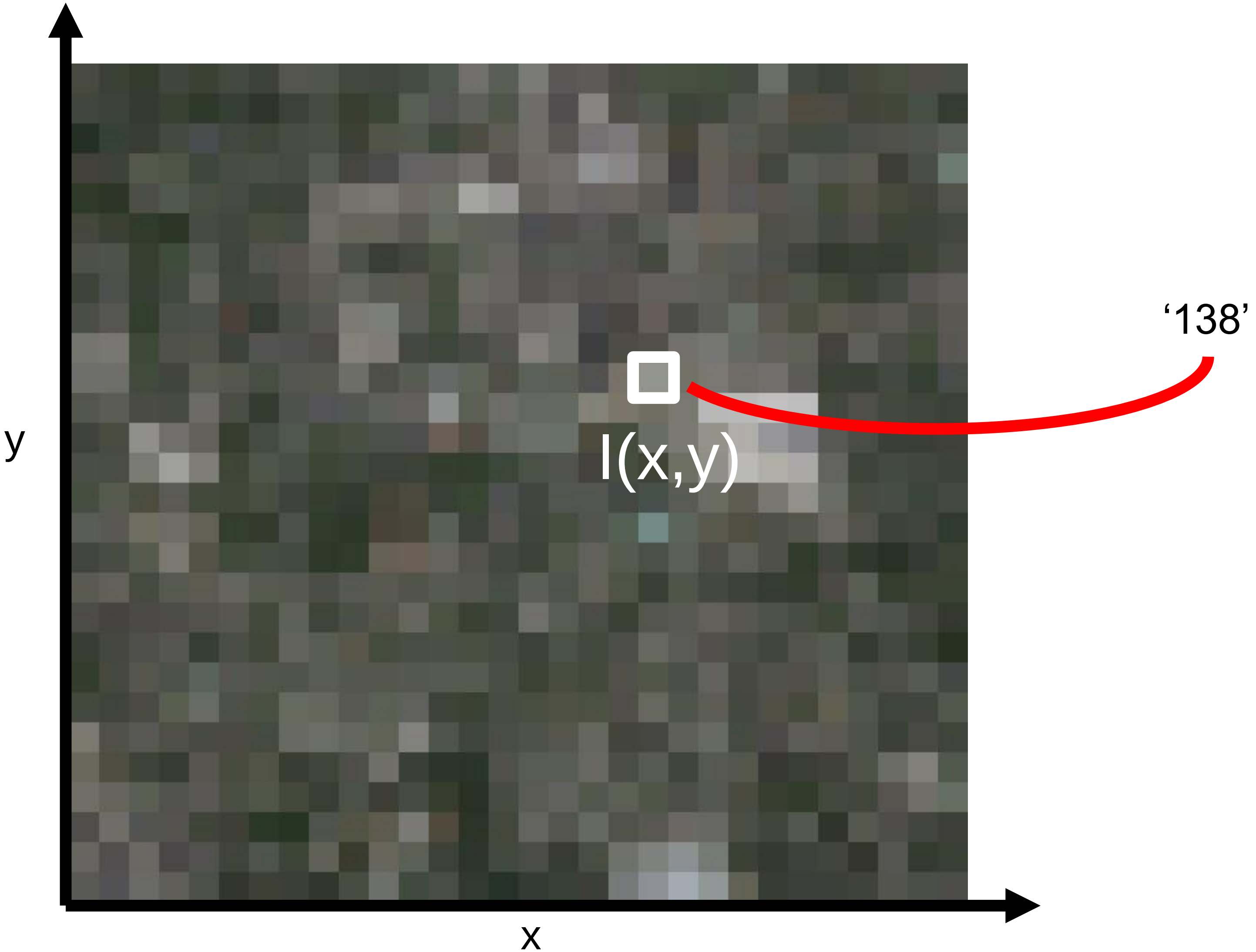
- Some depiction of all possible scenes would fit into this memory.
- Computer vision as making sense of an extremely high-dimensional space.
  - Subspace of ‘natural’ images.
  - Deriving low-dimensional, explainable models.

What is each part of an image?

- What does it represent in terms of cameras?

What is each part of an image?

- Pixel -> picture element



Perhaps a pixel is not a little square?

"A Pixel Is **Not** ***A Little Square,***
A Pixel Is **Not** ***A Little Square,***
A Pixel Is **Not** ***A Little Square!***
(And a Voxel is **Not** ***a Little Cube)"***

- **Alvy Ray Smith,**
- **MS Tech Memo 6, 1995.**

A Pixel Is *Not* A Little Square,
A Pixel Is *Not* A Little Square,
A Pixel Is *Not* A Little Square!
(And a Voxel is *Not* a Little Cube)[1]

Technical Memo 6

Alvy Ray Smith
July 17, 1995

Abstract

My purpose here is to, once and for all, rid the world of the misconception that a pixel is a little geometric square. This is not a religious issue. This is an issue that strikes right at the root of correct image (sprite) computing and the ability to correctly integrate (converge) the discrete and the continuous. The little square model is simply incorrect. It harms. It gets in the way. If you find yourself thinking that a pixel is a little square, please read this paper. I will have succeeded if you at least understand that you are using the model and why it is permissible in your case to do so (is it?).

Everything I say about little squares and pixels in the 2D case applies equally well to little cubes and voxels in 3D. The generalization is straightforward, so I won't mention it from hereon[1].

I discuss why the *little square model* continues to dominate our collective minds. I show why it is wrong in general. I show when it is appropriate to use a little square in the context of a pixel. I propose a discrete to continuous mapping—because this is where the problem arises—that always works and does not assume too much.

I presented some of this argument in Tech Memo 5 ([Smith95]) but have encountered a serious enough misuse of the little square model since I wrote that paper to make me believe a full frontal attack is necessary.

The Little Square Model

The little square model pretends to represents a pixel (picture element) as a geometric square[2]. Thus pixel (i, j) is assumed to correspond to the area of the plane bounded by the square $\{(x, y) \mid i-.5 \le x \le i+.5, j-.5 \le y \le j+.5\}$.

[1] Added November 11, 1996, after attending the Visible Human Project Conference 96 in Bethesda, MD.

[2] In general, a little rectangle, but I will normalize to the little square here. The little rectangle model is the same mistake.

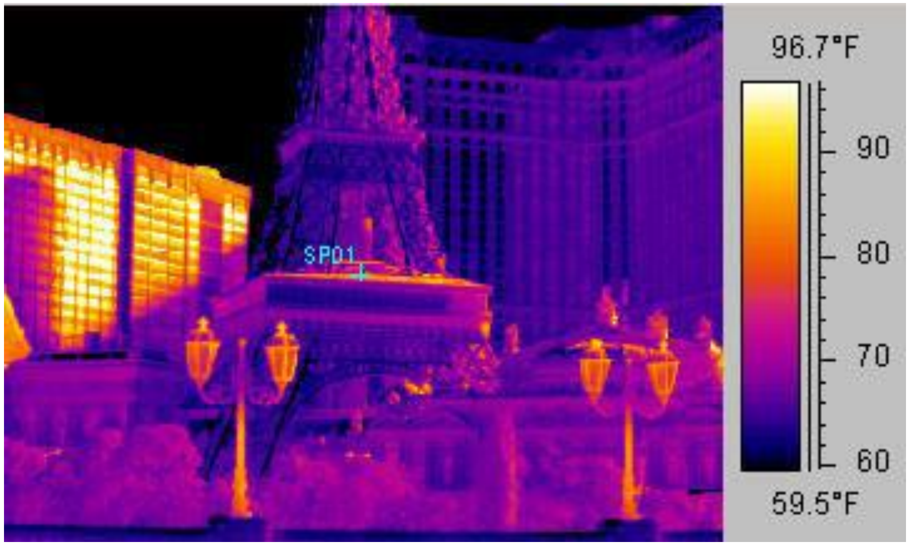
Image as a 2D sampling of signal

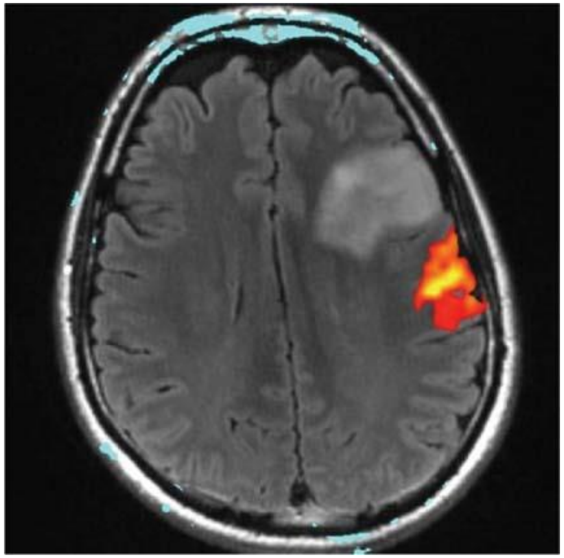
- Signal: function depending on some variable with physical meaning

- Image: sampling of that function
  - 2 variables: xy coordinates
  - 3 variables: xy + time (video)
  - 'Brightness' is the value of the function for visible light

- Can be other physical values too: temperature, pressure, depth …

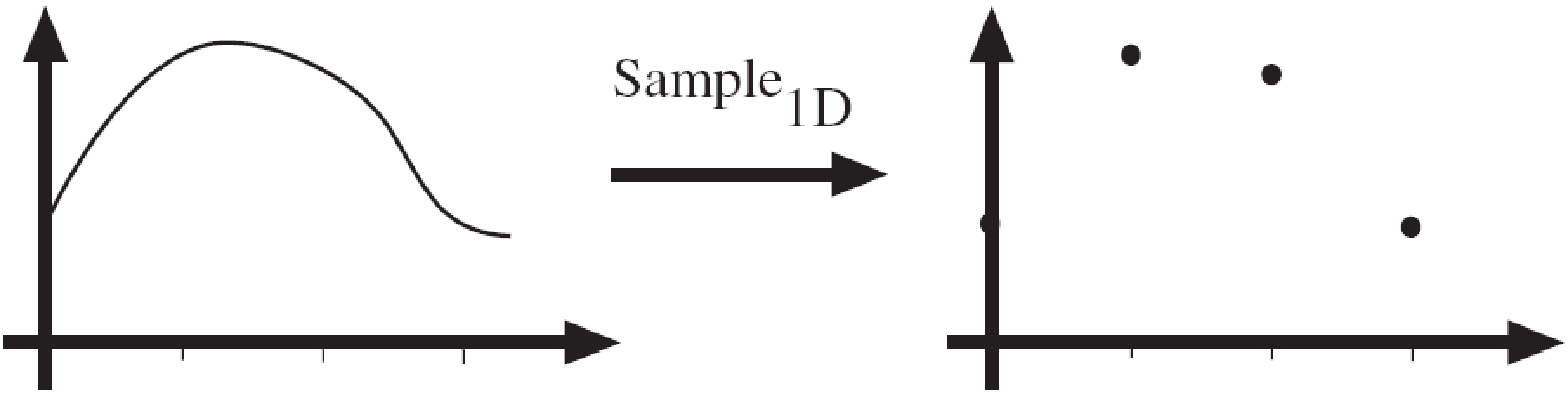
Example 2D Images





Danny Alexander

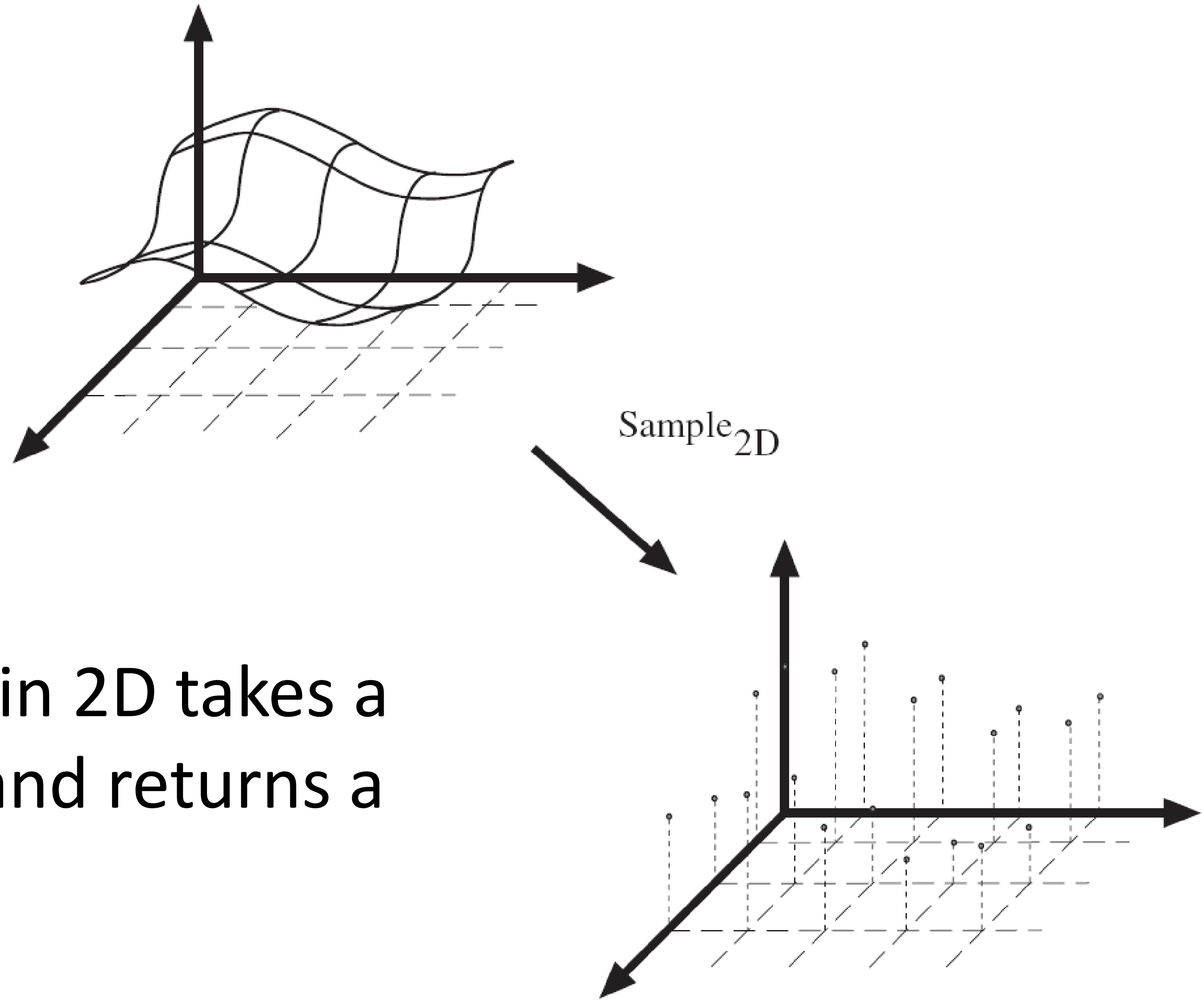
Sampling in 1D



- Sampling in 1D takes a function, and returns a vector whose elements are values of that function at the sample points.

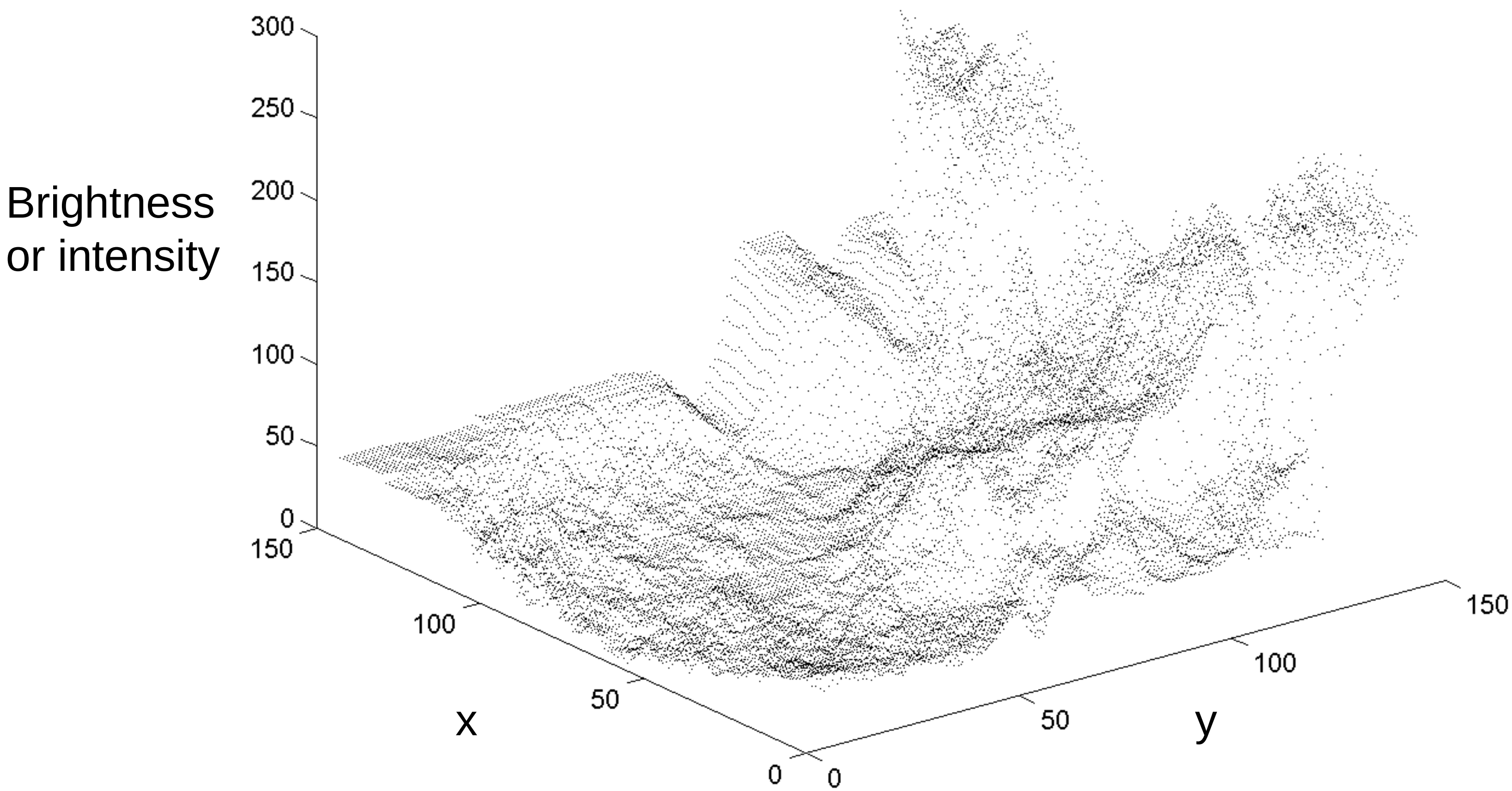
Danny Alexander

Sampling in 2D



- Sampling in 2D takes a function and returns a matrix.

Grayscale Digital Image



What is each part of an image?

- Pixel -> picture element

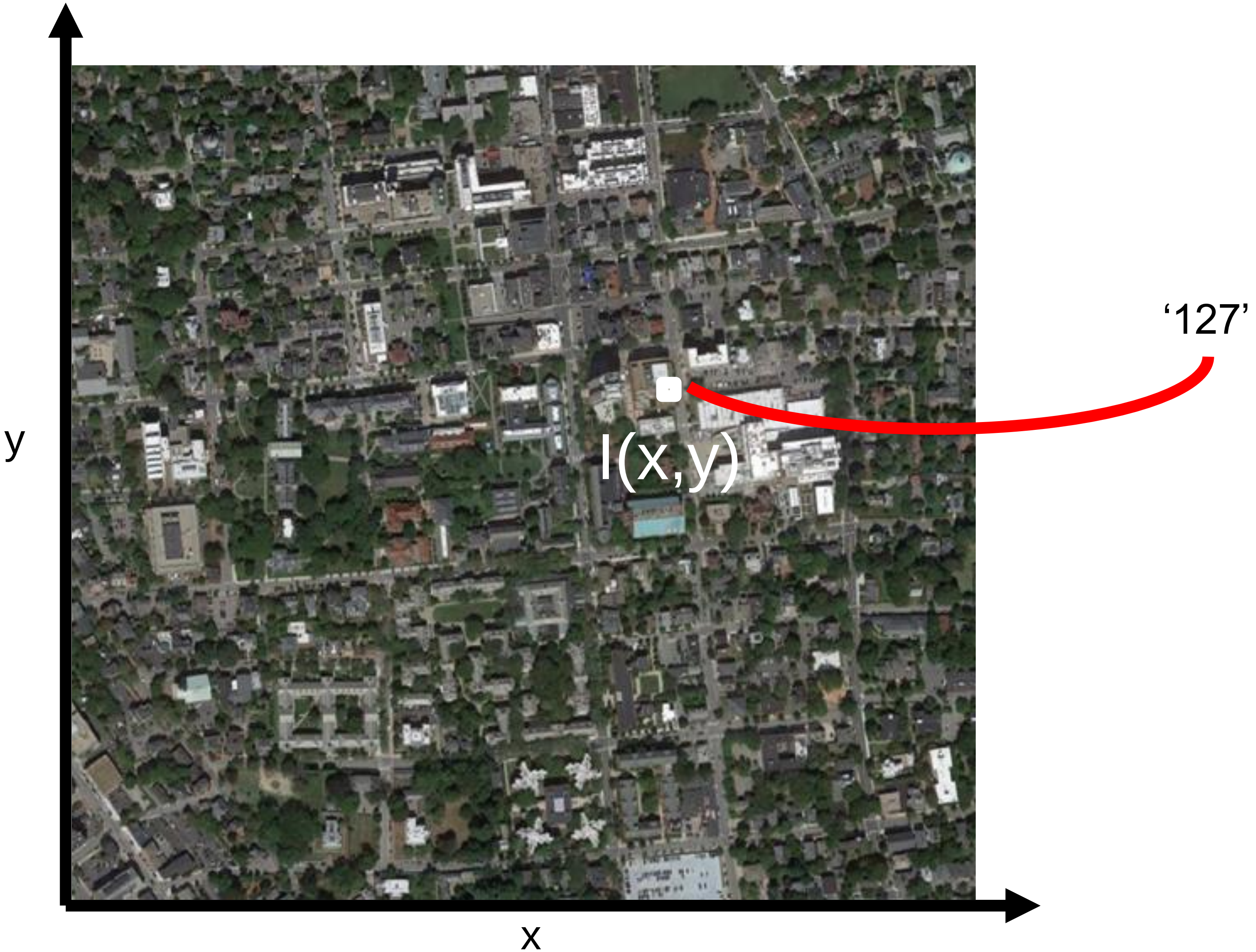
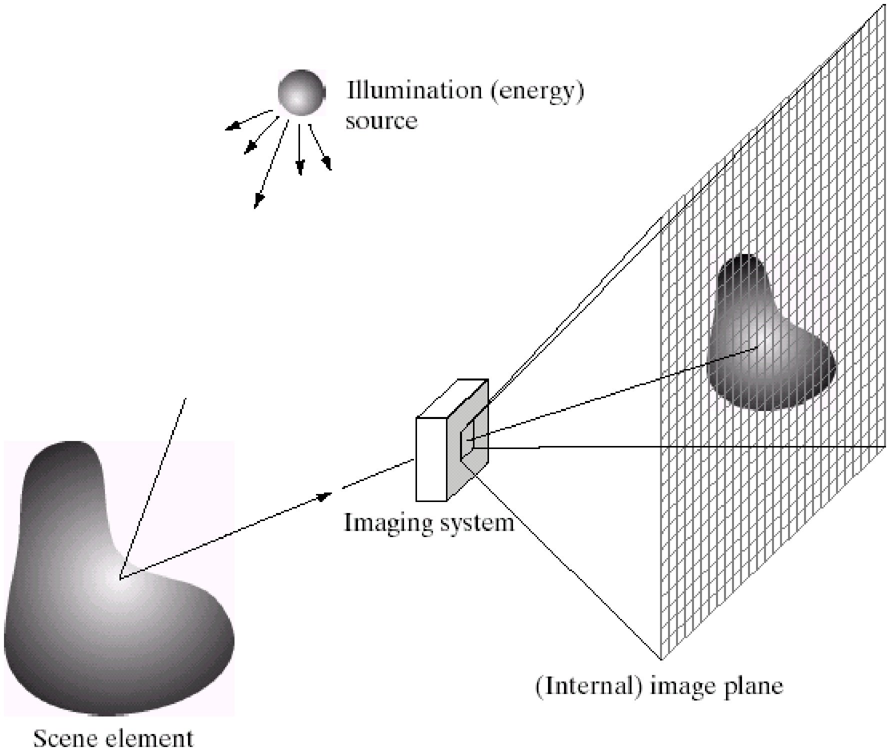
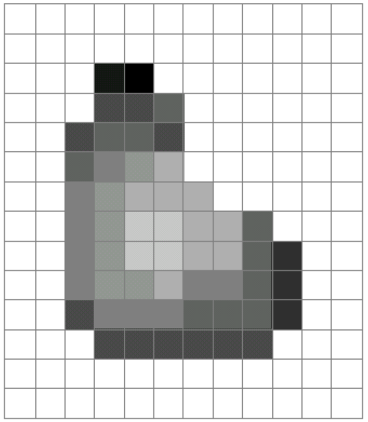


Image Formation



Camera Sensor

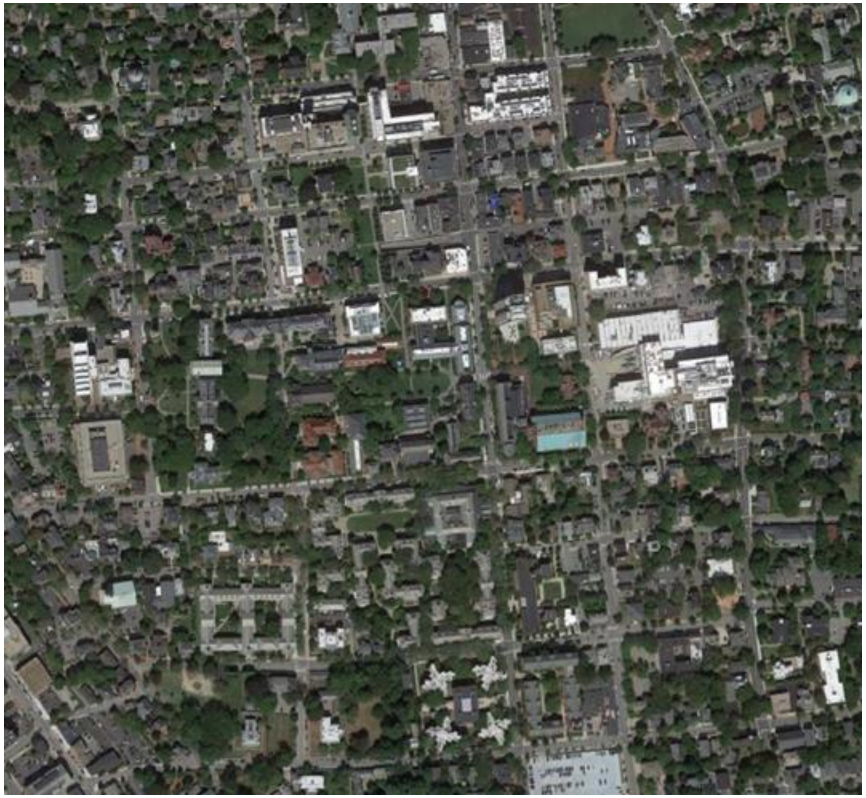


Output Image

James Hays

Resolution – geometric vs. spatial resolution

Both images are ~500x500 pixels



Quantization

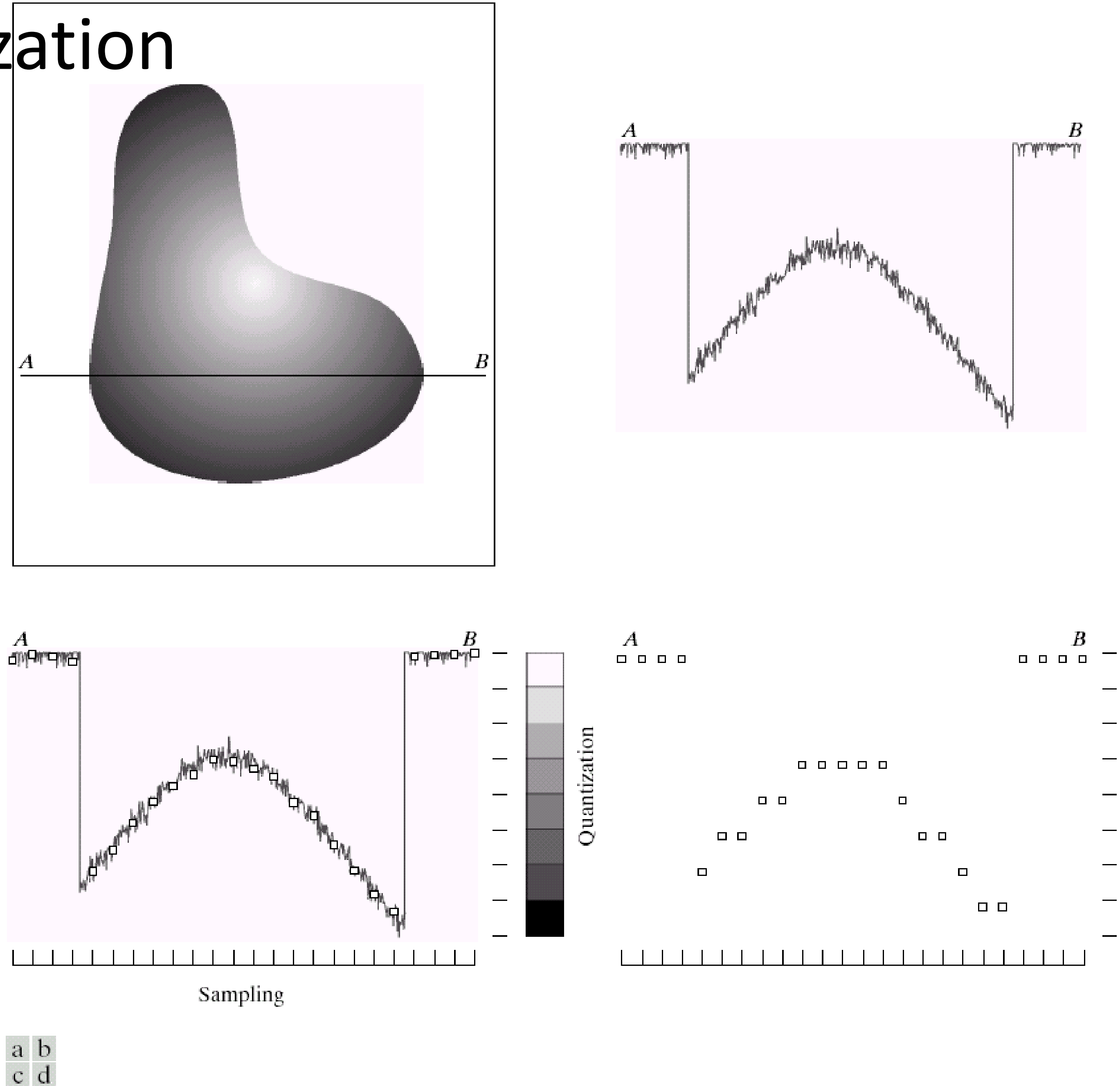
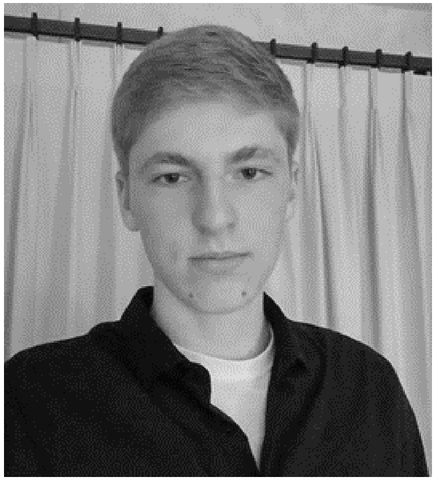


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

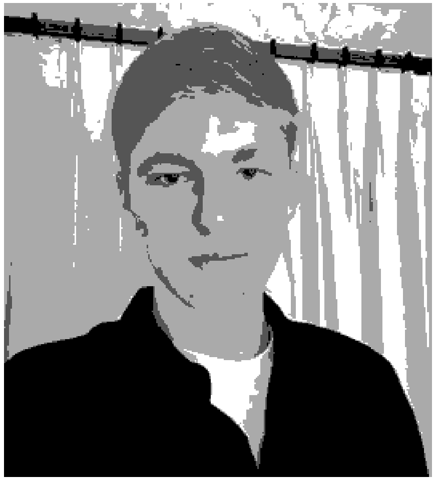
James Hays

Quantization Effects – Radiometric Resolution



8 bit – 256 levels

4 bit – 16 levels

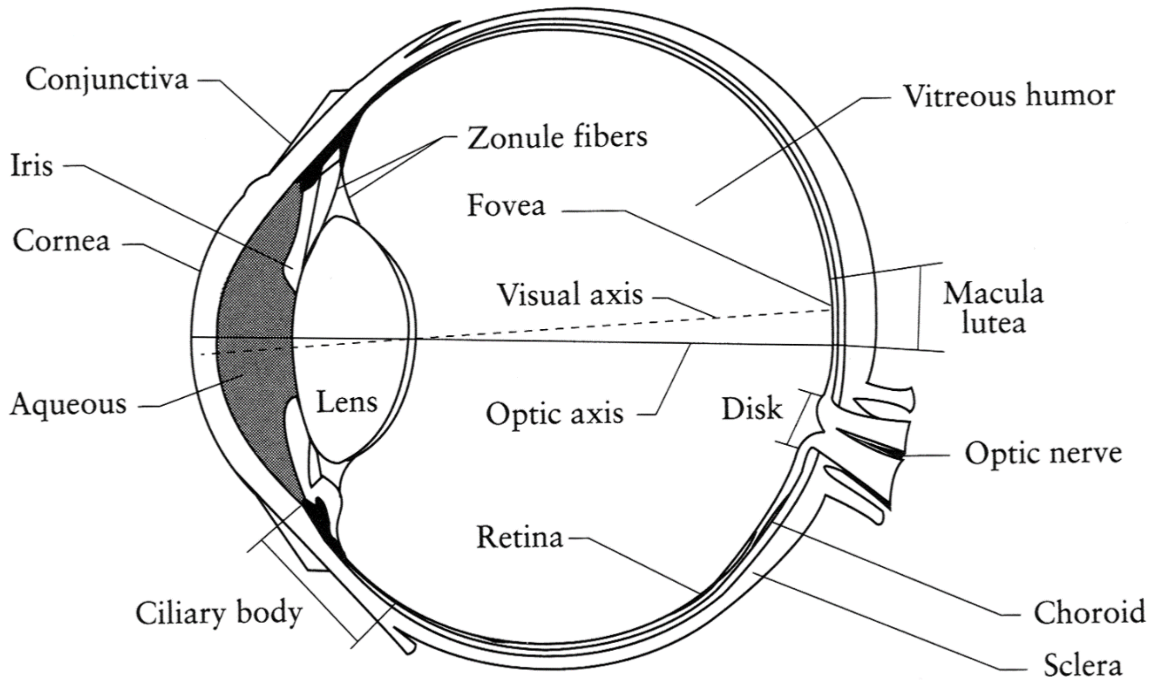


2 bit – 4 levels

1 bit – 2 levels

ANATOMY

The Eye

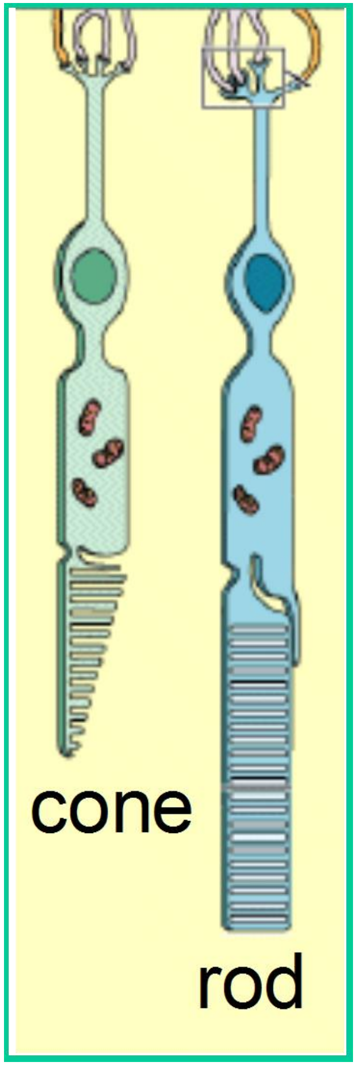


- The human eye is a camera
  - **Iris** - colored annulus with radial muscles
  - **Pupil** - the hole (aperture) whose size is controlled by the iris
  - What's the sensor?
    - photoreceptor cells (rods and cones) in the **retina**

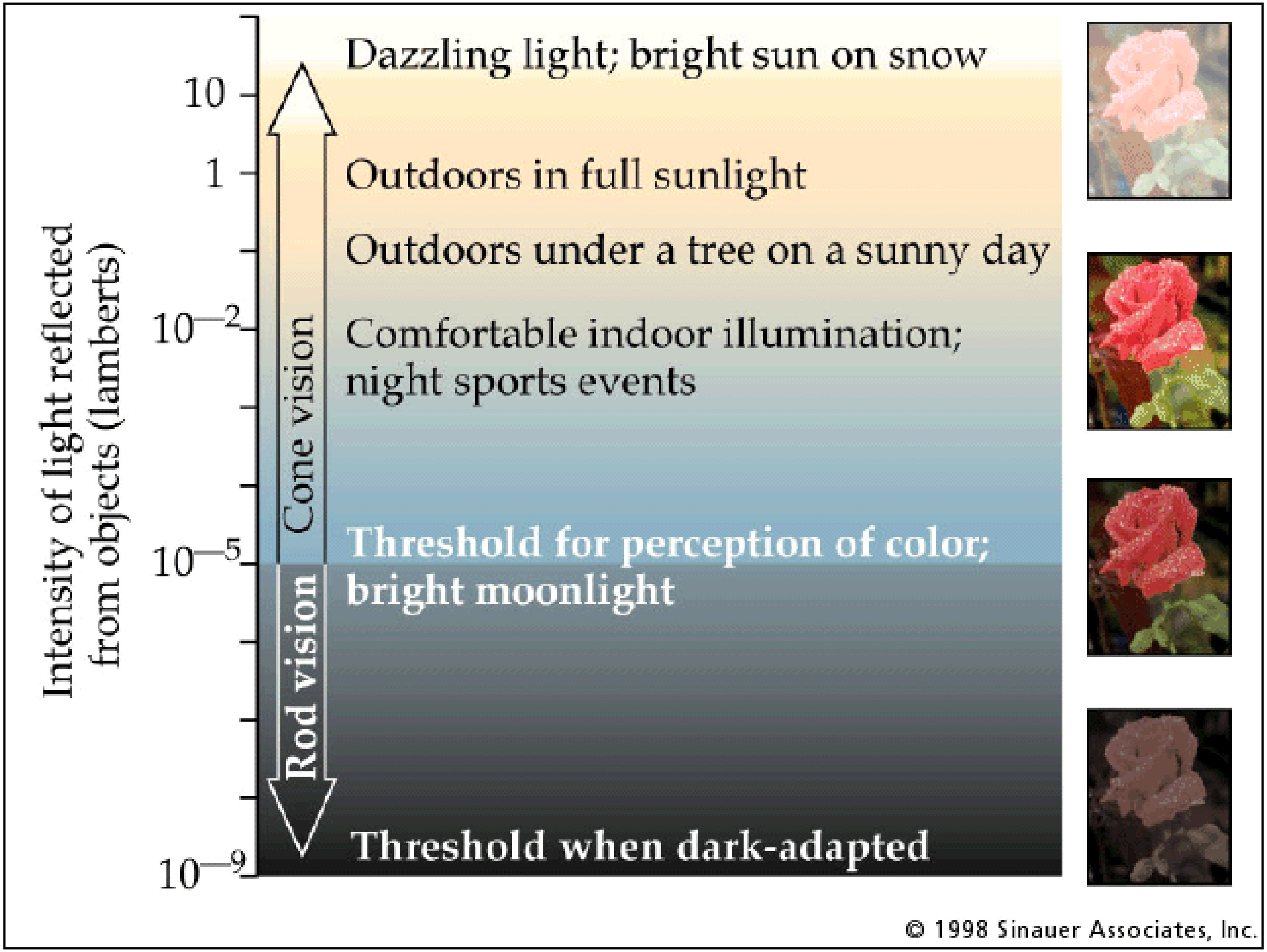
Two types of light-sensitive receptors

Cones
- cone-shaped
- less sensitive
- operate in high light
- color vision

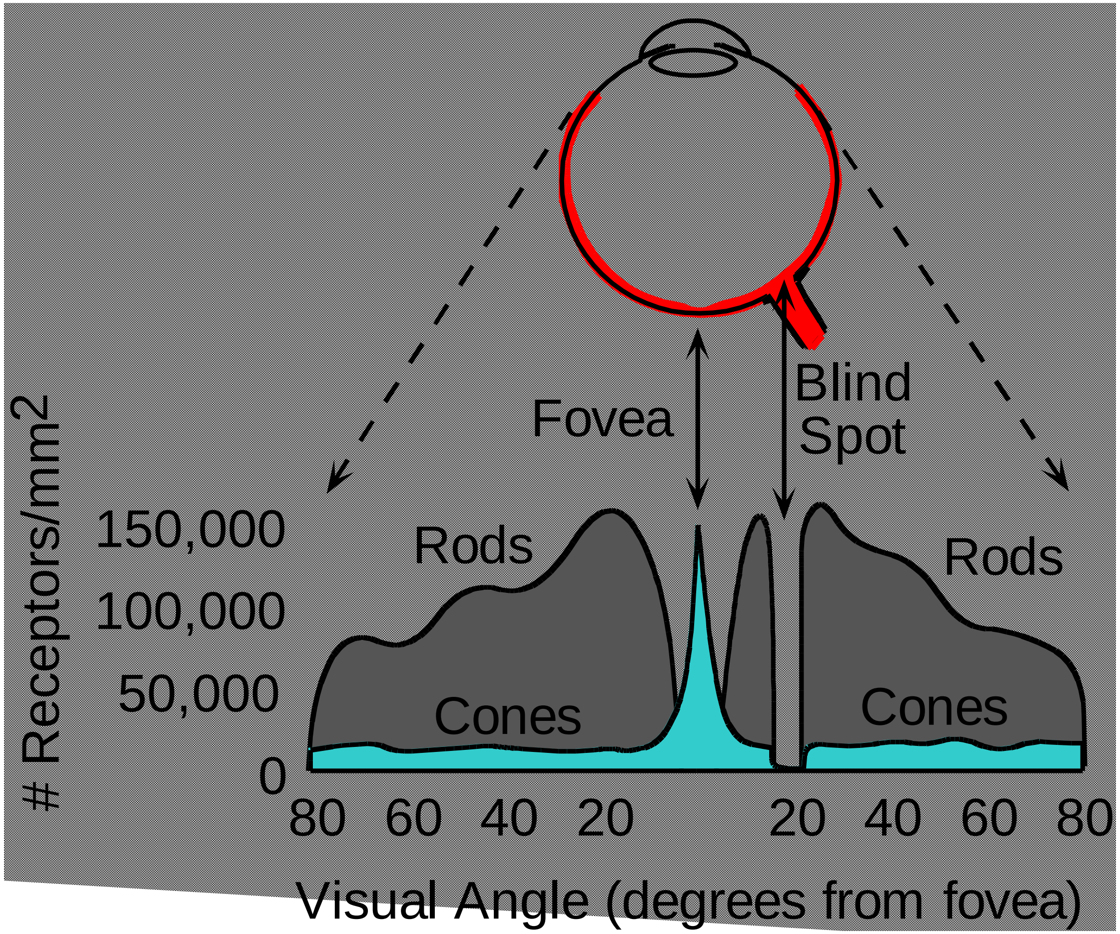
Rods
- rod-shaped
- highly sensitive
- operate at night
- gray-scale vision



Rod / Cone sensitivity



Distribution of Rods and Cones

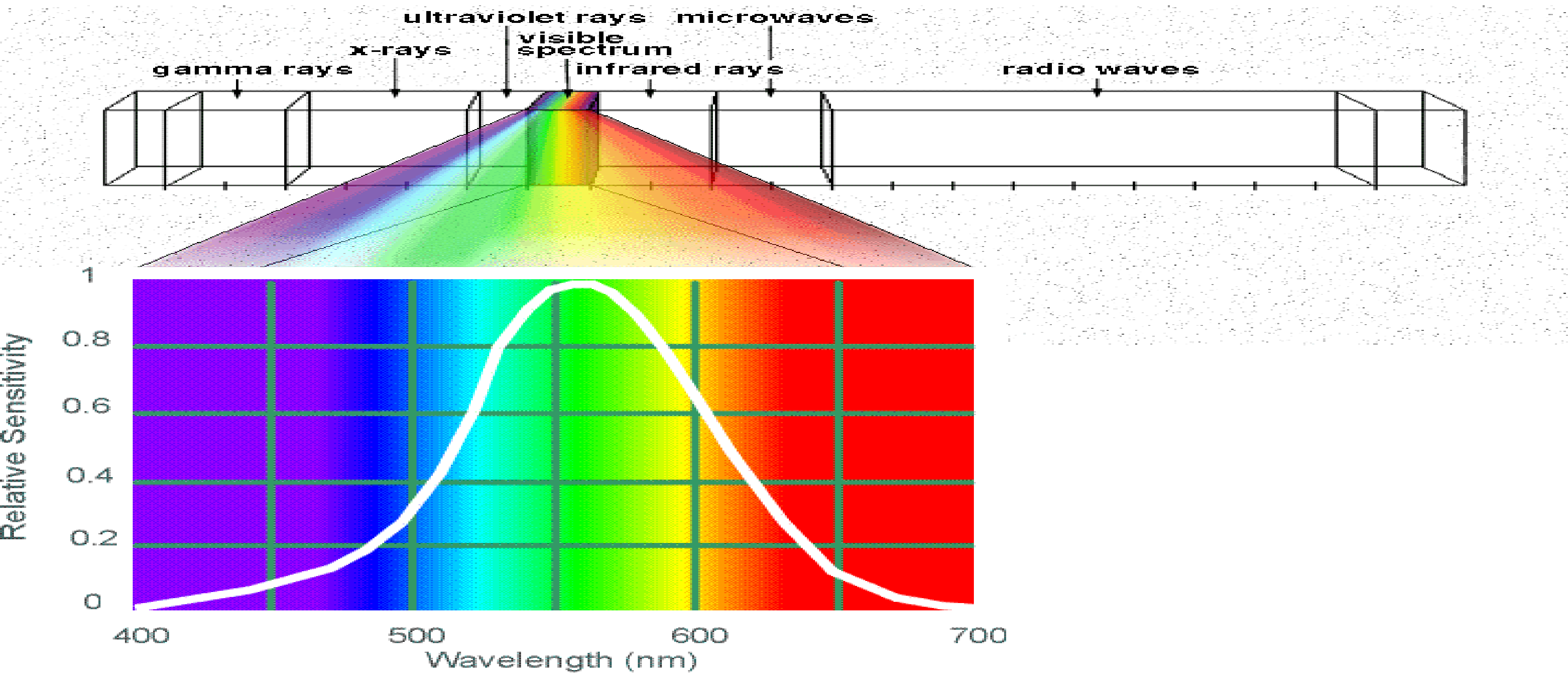


Night Sky: why are there more stars off-center?

Averted vision: http://en.wikipedia.org/wiki/Averted_vision

James Hays

Electromagnetic Spectrum

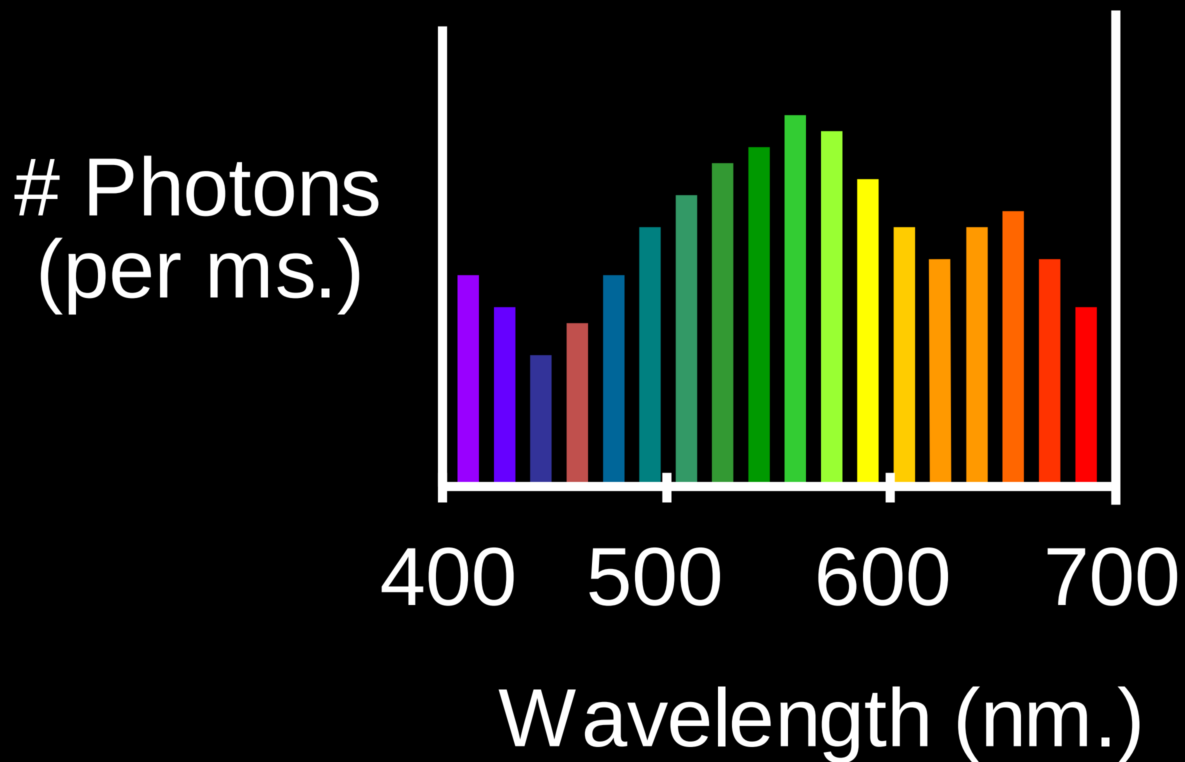


Human Luminance Sensitivity Function

http://www.yorku.ca/eye/photopik.htm

The Physics of Light

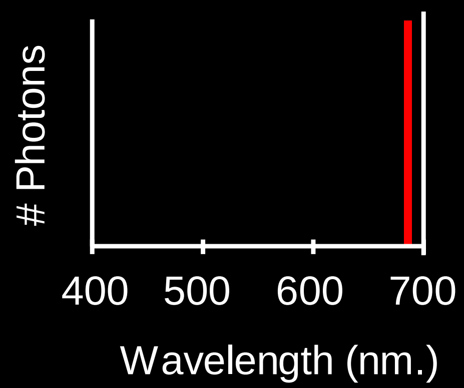
Any patch of light can be completely described physically by its spectrum: the number of photons (per time unit) at each wavelength 400 - 700 nm.



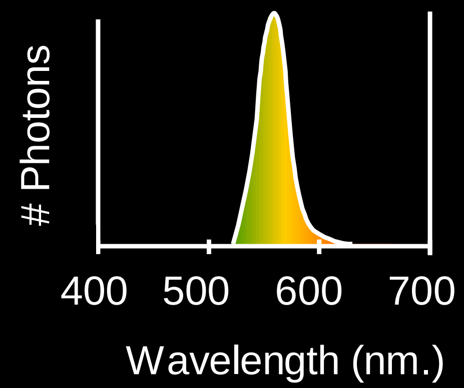
The Physics of Light

Some examples of the spectra of light sources

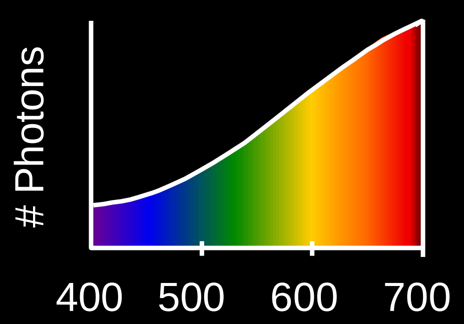
A. Ruby Laser



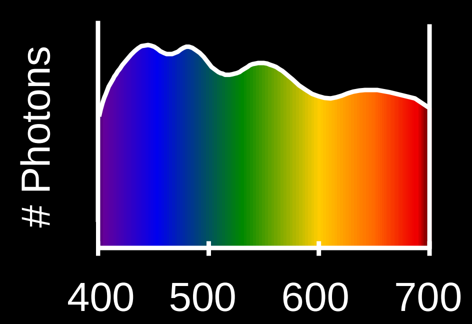
B. Gallium Phosphide Crystal



C. Tungsten Lightbulb

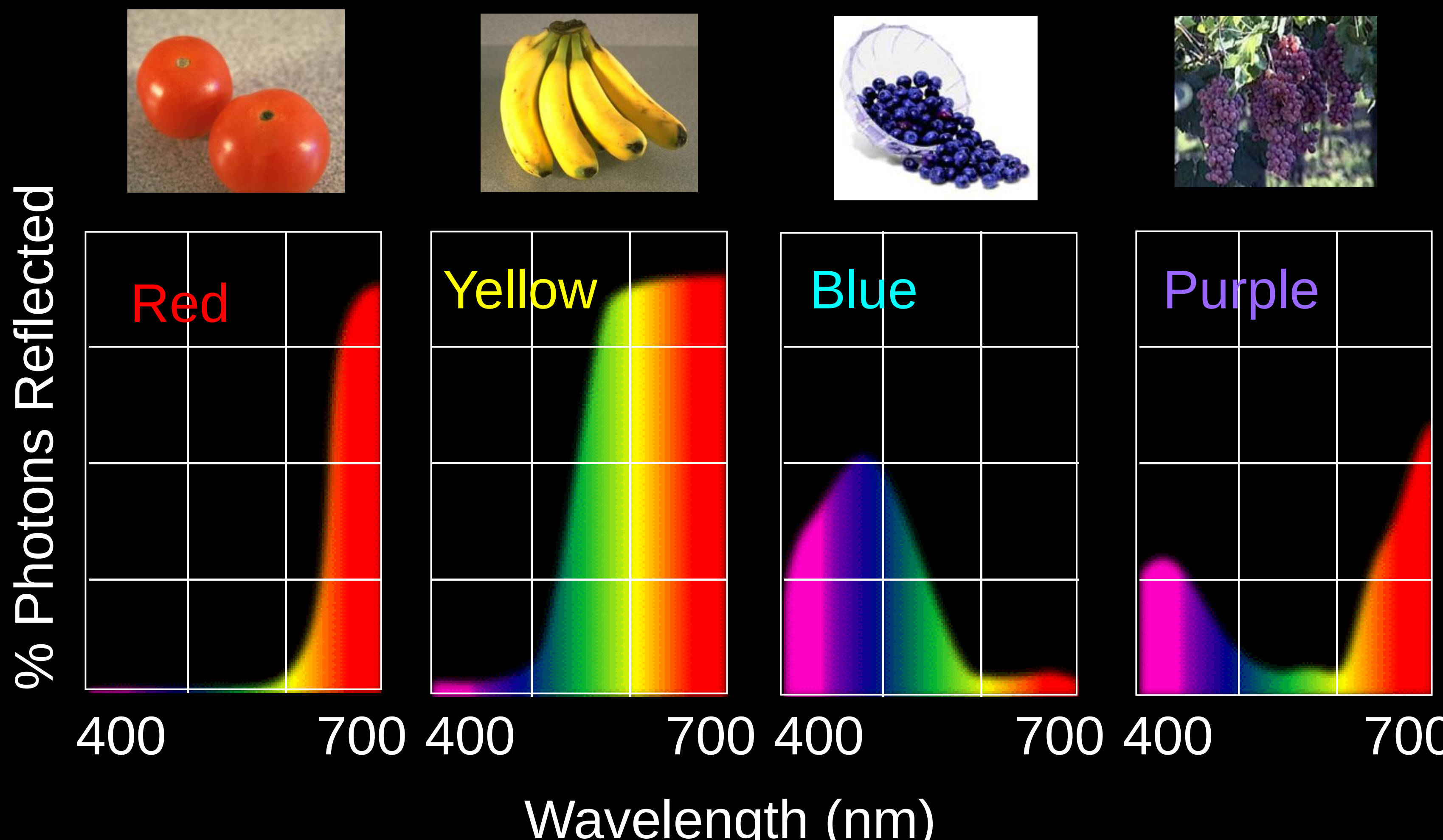


D. Normal Daylight



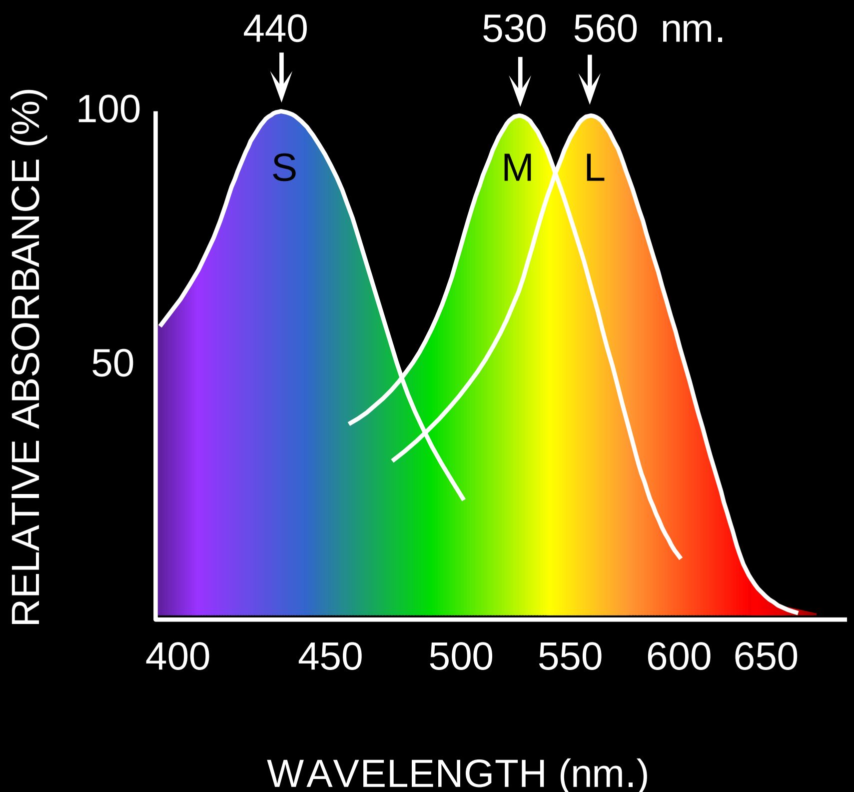
The Physics of Light

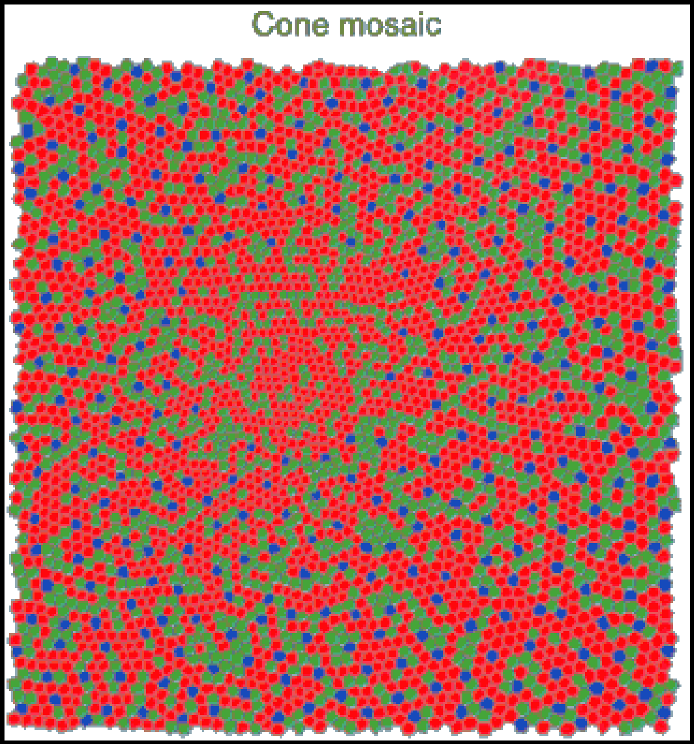
Some examples of the reflectance spectra of surfaces



Physiology of Color Vision

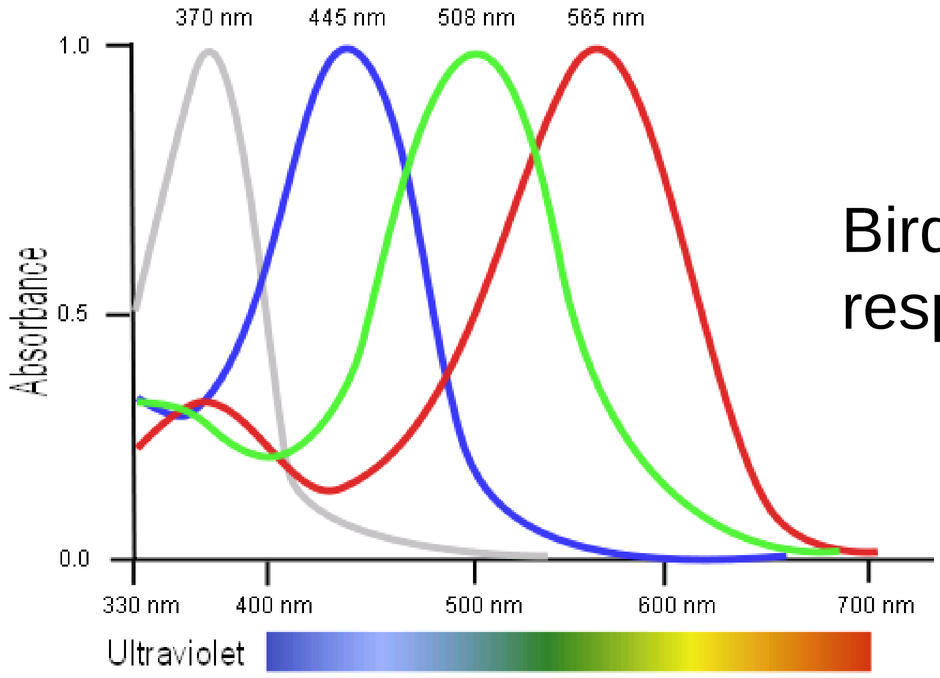
Three kinds of cones:





- Why are M and L cones so close?
- Why are there 3?

Tetrachromatism



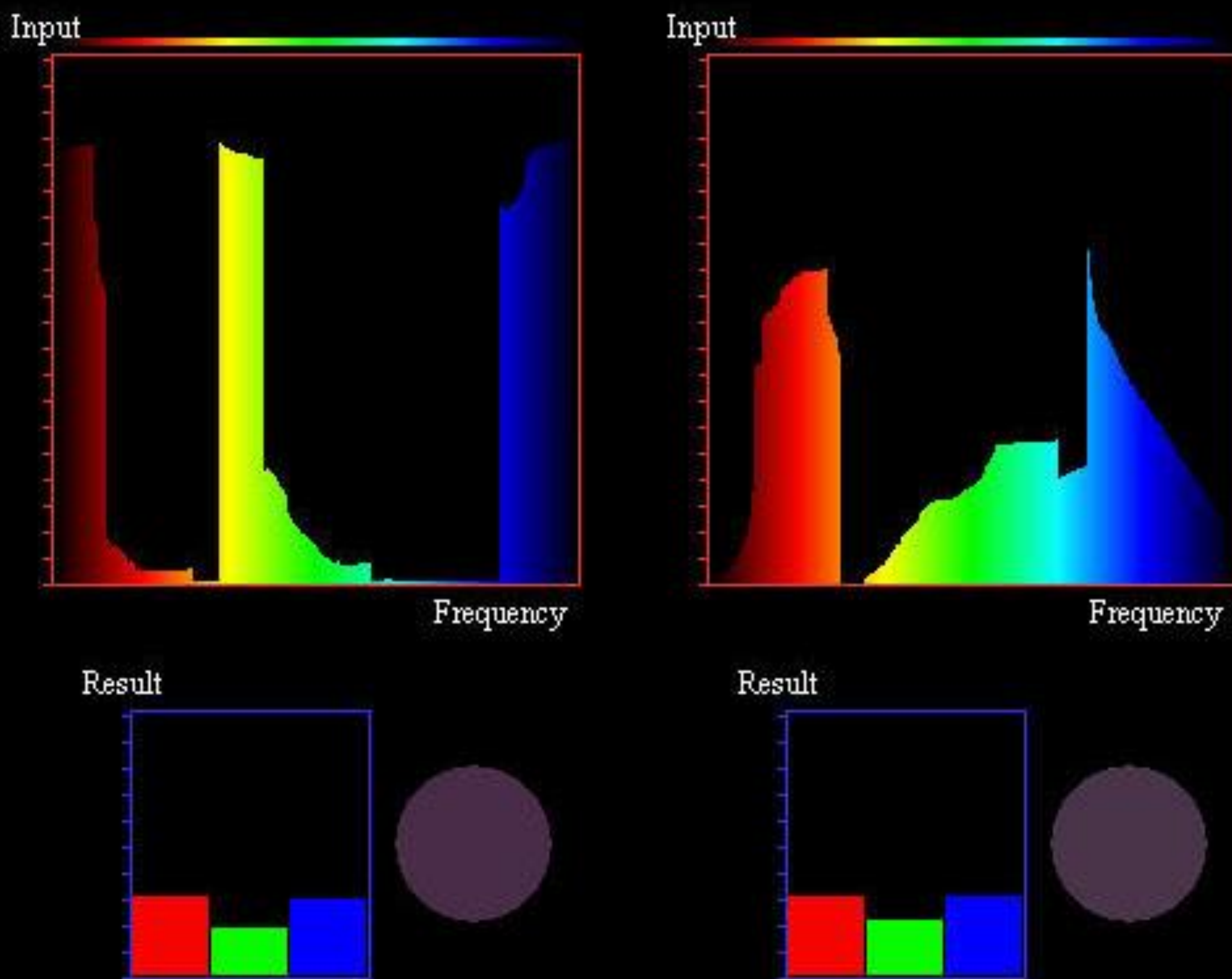
Bird cone responses

- Most birds, and many other animals, have cones for ultraviolet light.
- Some humans seem to have four cones (12% of females).
- True tetrachromatism is _rare_; requires learning.

James Hays

Bee vision

Metamers

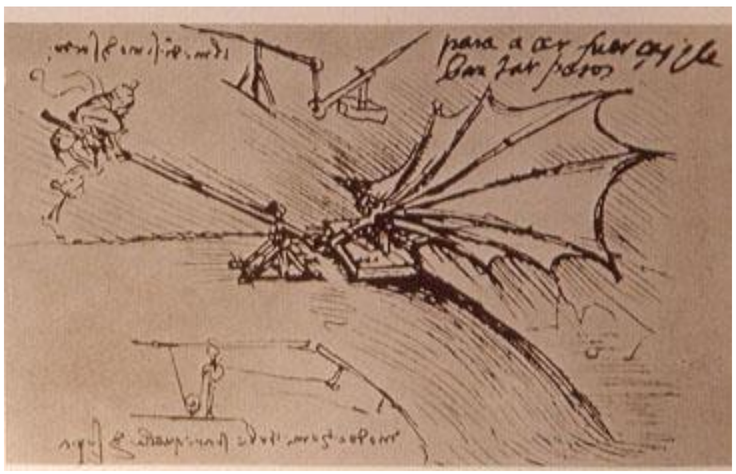


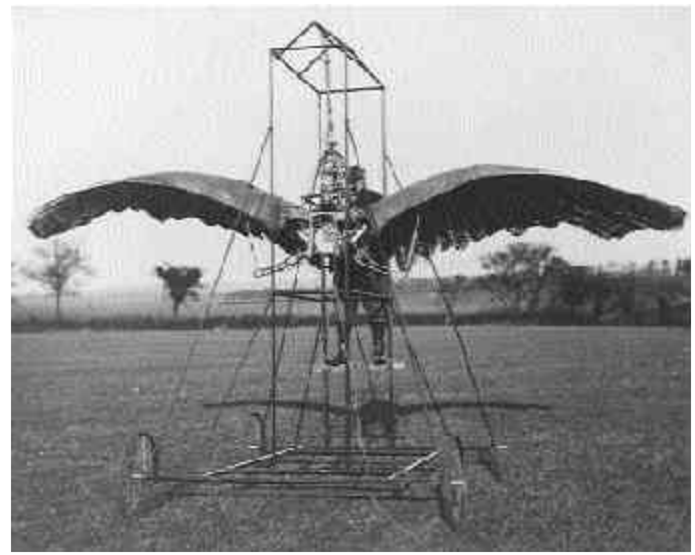
by Jeff Beall, Adam Doppelt and John F. Hughes

Does color exist?

Do we care about human vision
in this class?

Ornithopters





James Hays

Why do we care about human vision?

- We don’t, necessarily.

- But cameras imitate the frequency response of the human eye, so we should know that much.
- Computer vision wouldn’t get as much scrutiny if biological vision (especially human vision) hadn’t proved that it was possible to make important judgements from images.

Does computer vision “understand” images?

"Can machines fly?" The answer is yes, because airplanes fly.

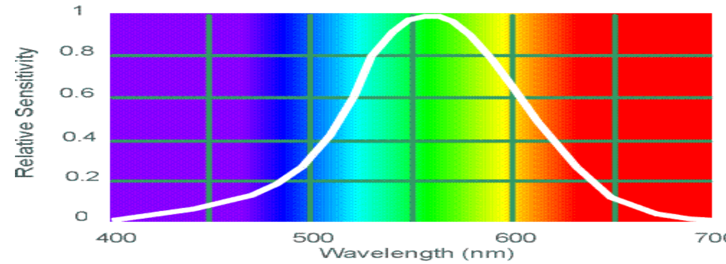
"Can machines swim?" The answer is no, because submarines don't swim.

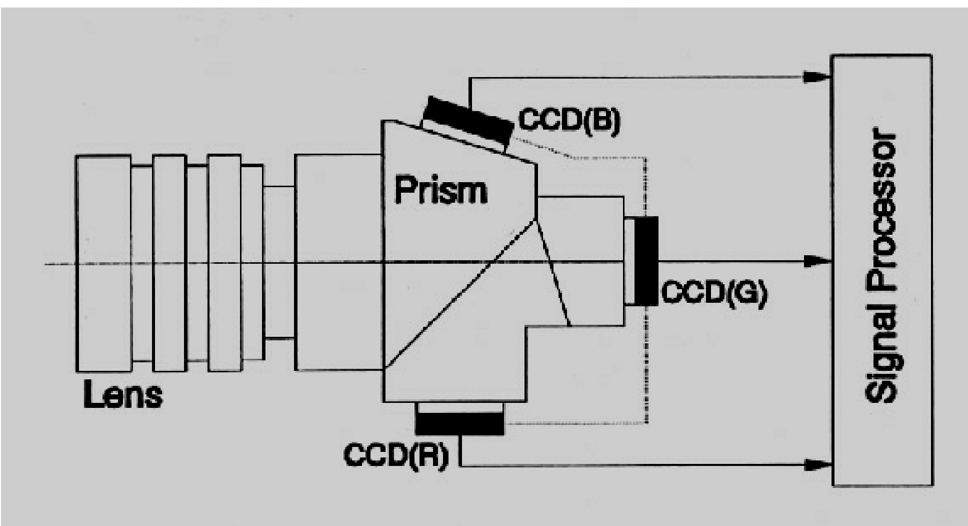
"Can machines think?" Is this question like the first, or like the second?

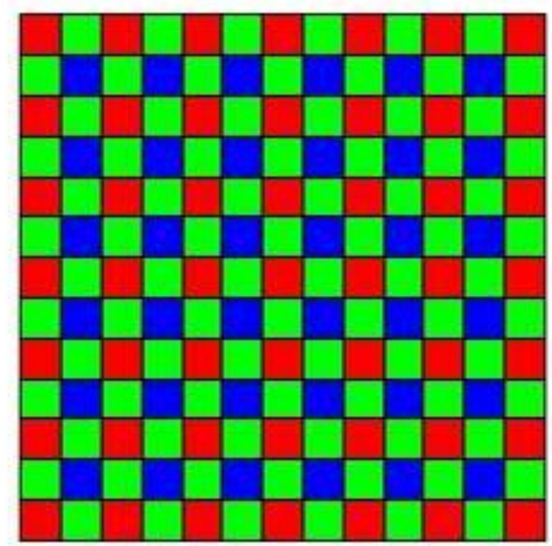
Source: Norvig

Color Sensing in Camera (RGB)

- 3-chip vs. 1-chip: quality vs. cost
- Why more green?







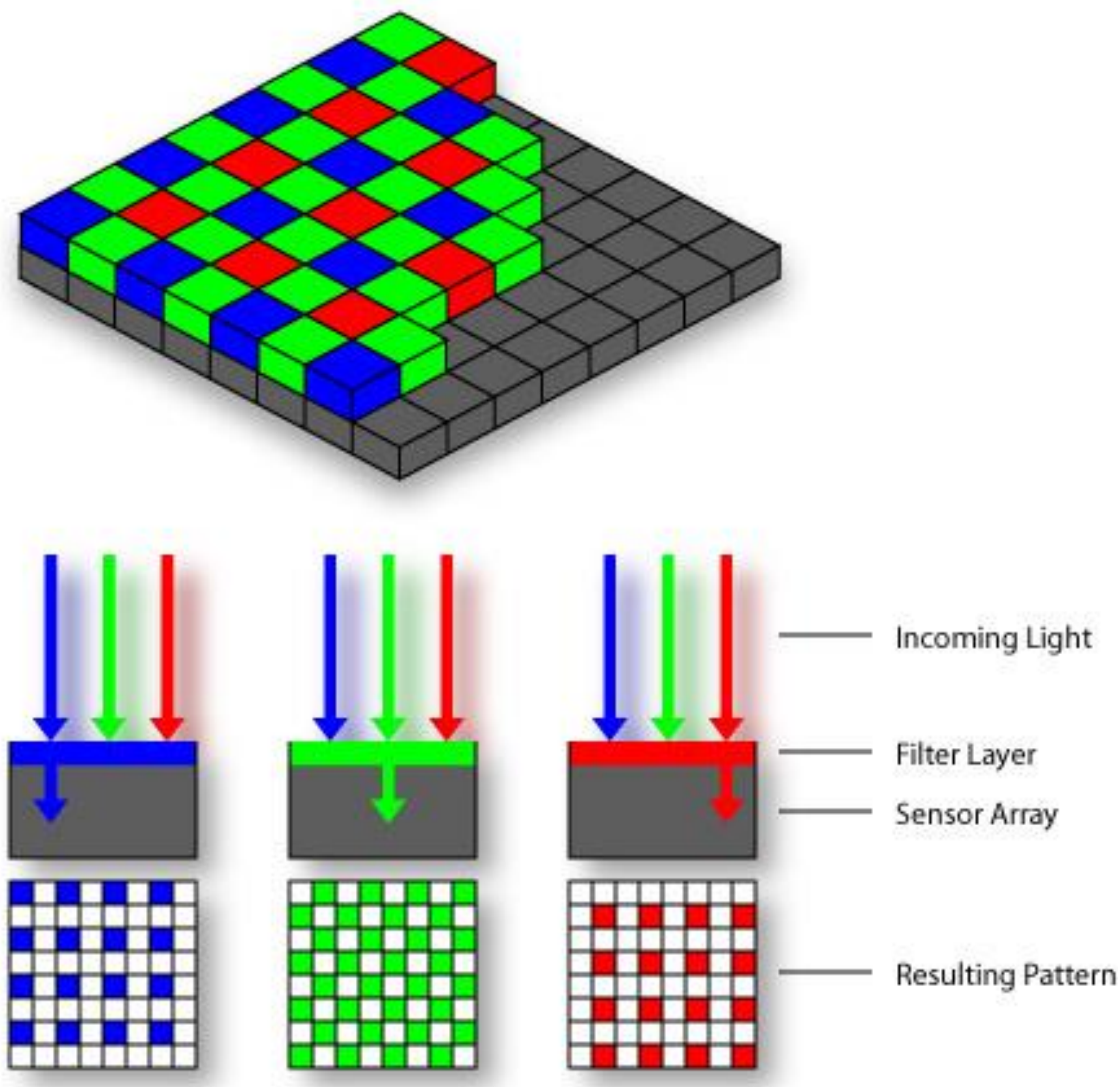
Bayer filter

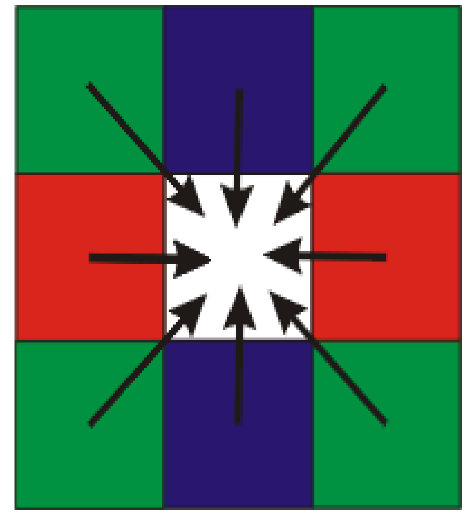
Why 3 colors?

http://www.cooldictionary.com/words/Bayer-filter.wikipedia

Slide by Steve Seitz

Practical Color Sensing: Bayer Grid

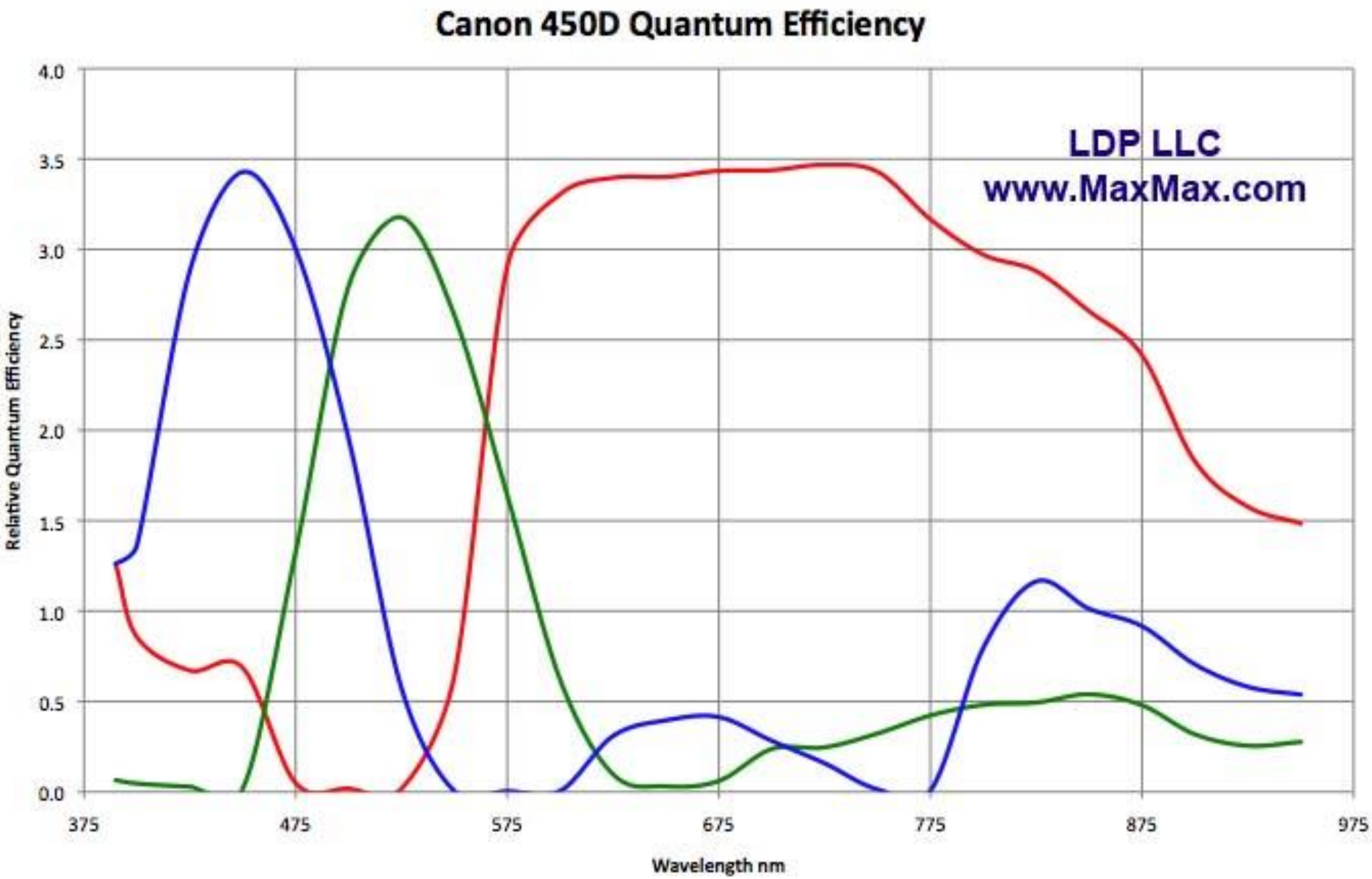




- Estimate RGB at 'G' cells from neighboring values

Slide by Steve Seitz

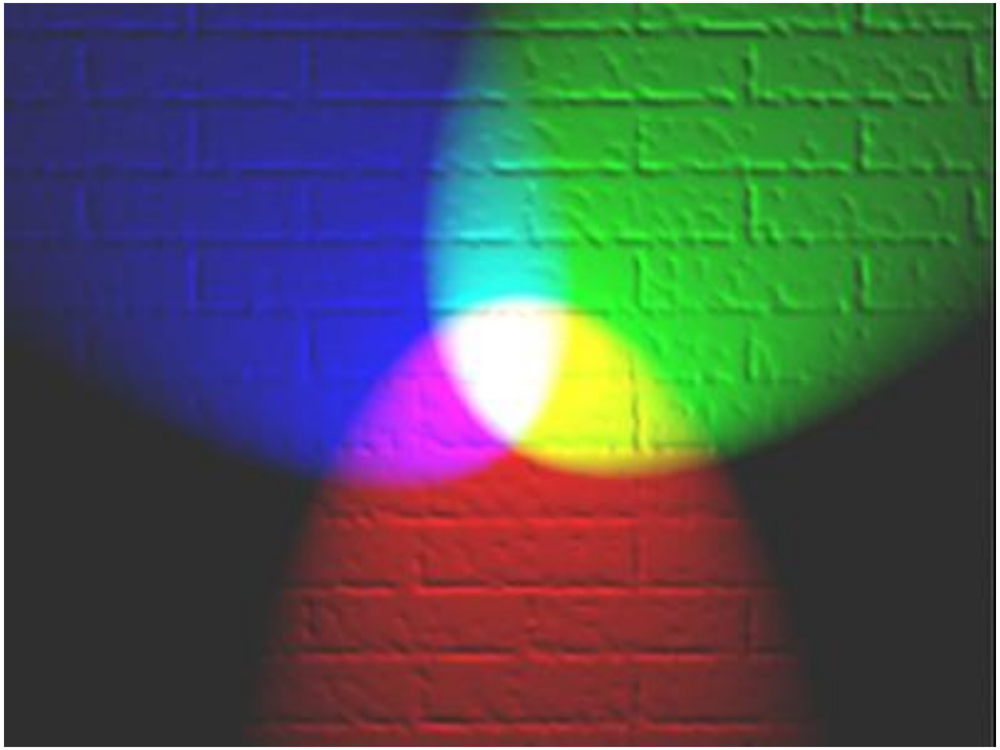
Camera Color Response



MaxMax.com

Color spaces

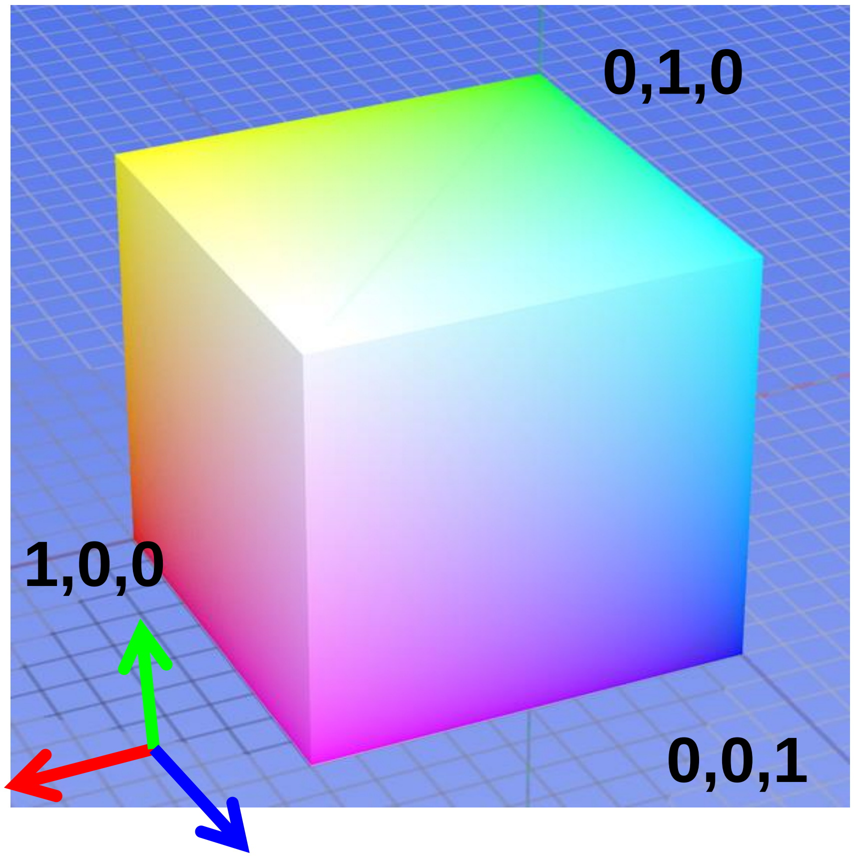
- How can we represent color?



http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Color spaces: RGB

Default color space



R = 1
(G=0,B=0)

G = 1
(R=0,B=0)

B = 1
(R=0,G=0)

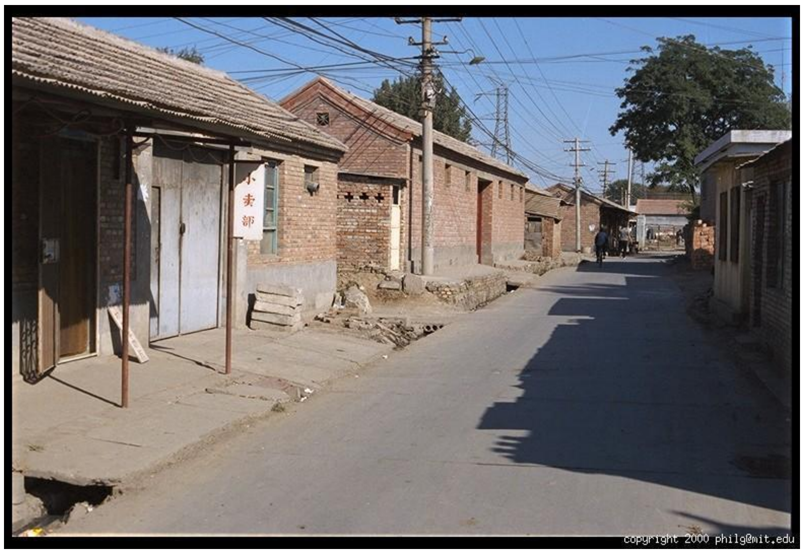
Any color = r*R + g*G + b*B
- Strongly correlated channels
- Non-perceptual

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Got it. $C = r*R + g*G + b*B$

IS COLOR A VECTOR SPACE?

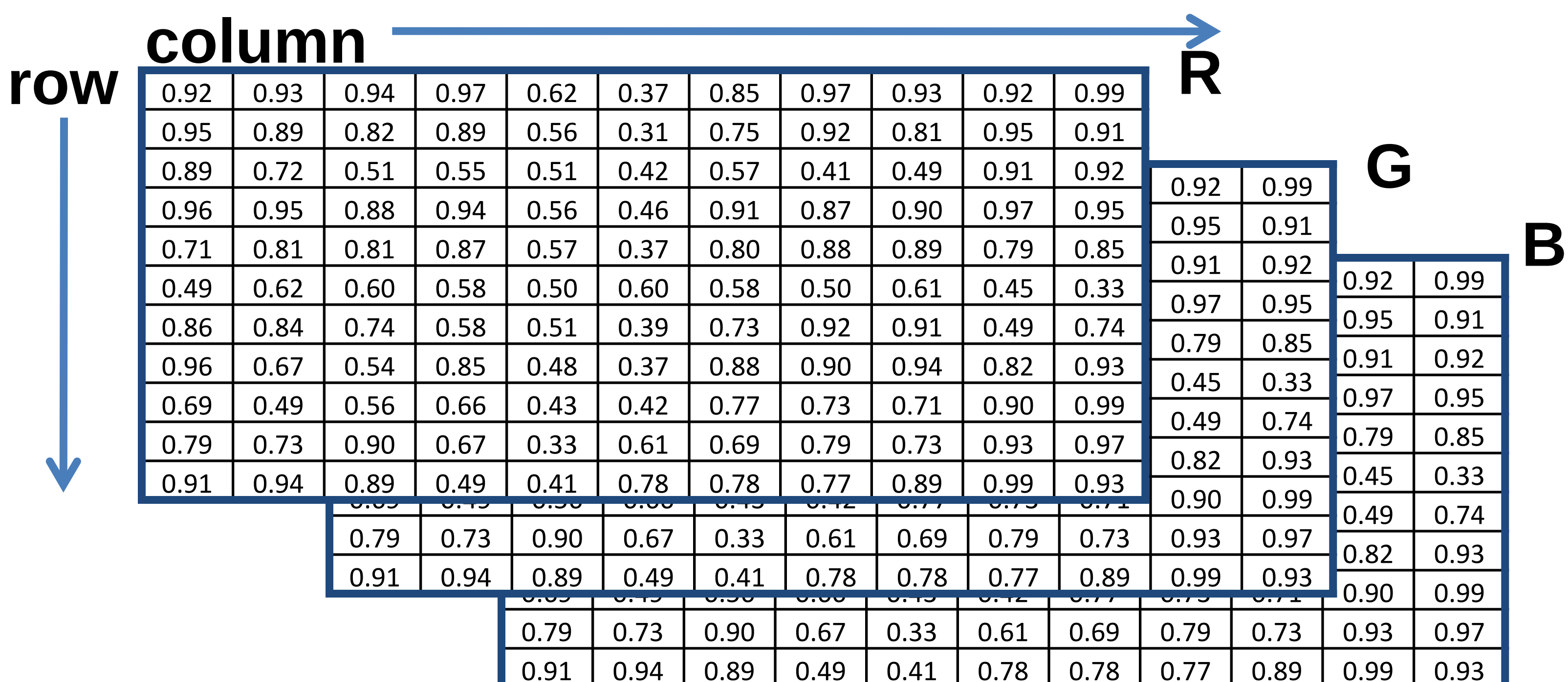
Color Image



James Hays

Images in Matlab

- Images represented as a matrix
- Suppose we have a NxM RGB image called “im”
  - im(1,1,1) = top-left pixel value in R-channel
  - im(y, x, b) = y pixels down, x pixels to right in the b^th channel
  - im(N, M, 3) = bottom-right pixel in B-channel
- imread(filename) returns a uint8 image (values 0 to 255)
  - Convert to double format (values 0 to 1) with im2double

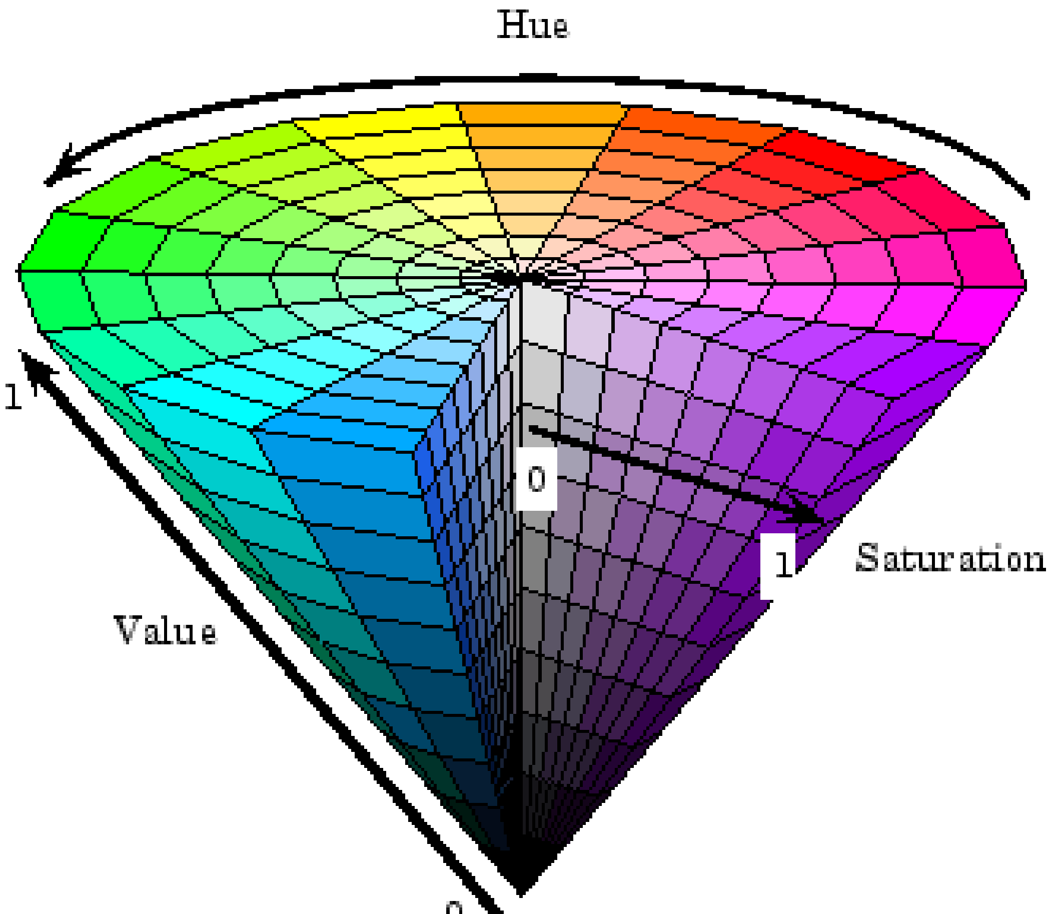


0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

James Hays

Color spaces: HSV

Intuitive color space



If you had to choose, would you rather go without:

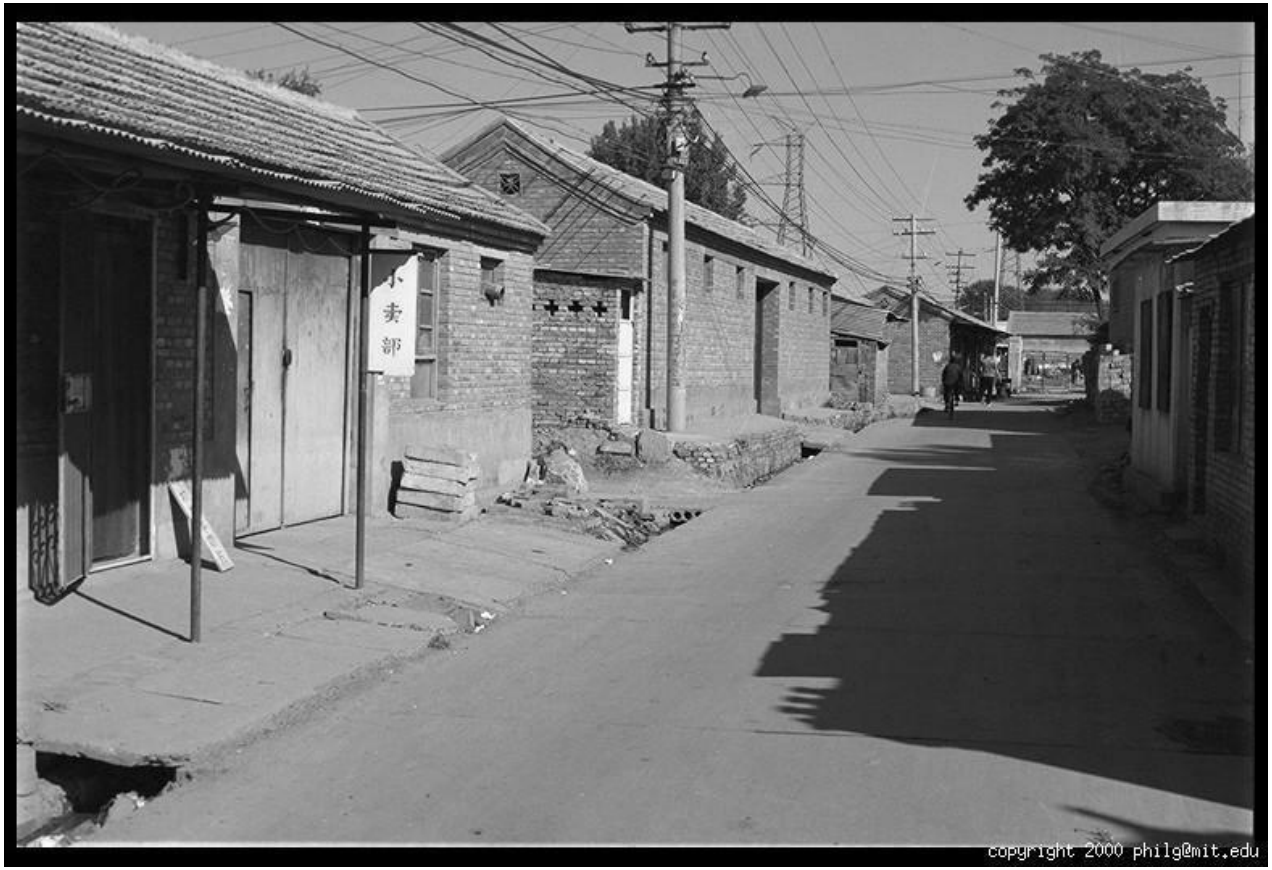
- intensity (‘value’), or
- hue + saturation (‘chroma’)?

Think-Pair-Share

Most information in intensity

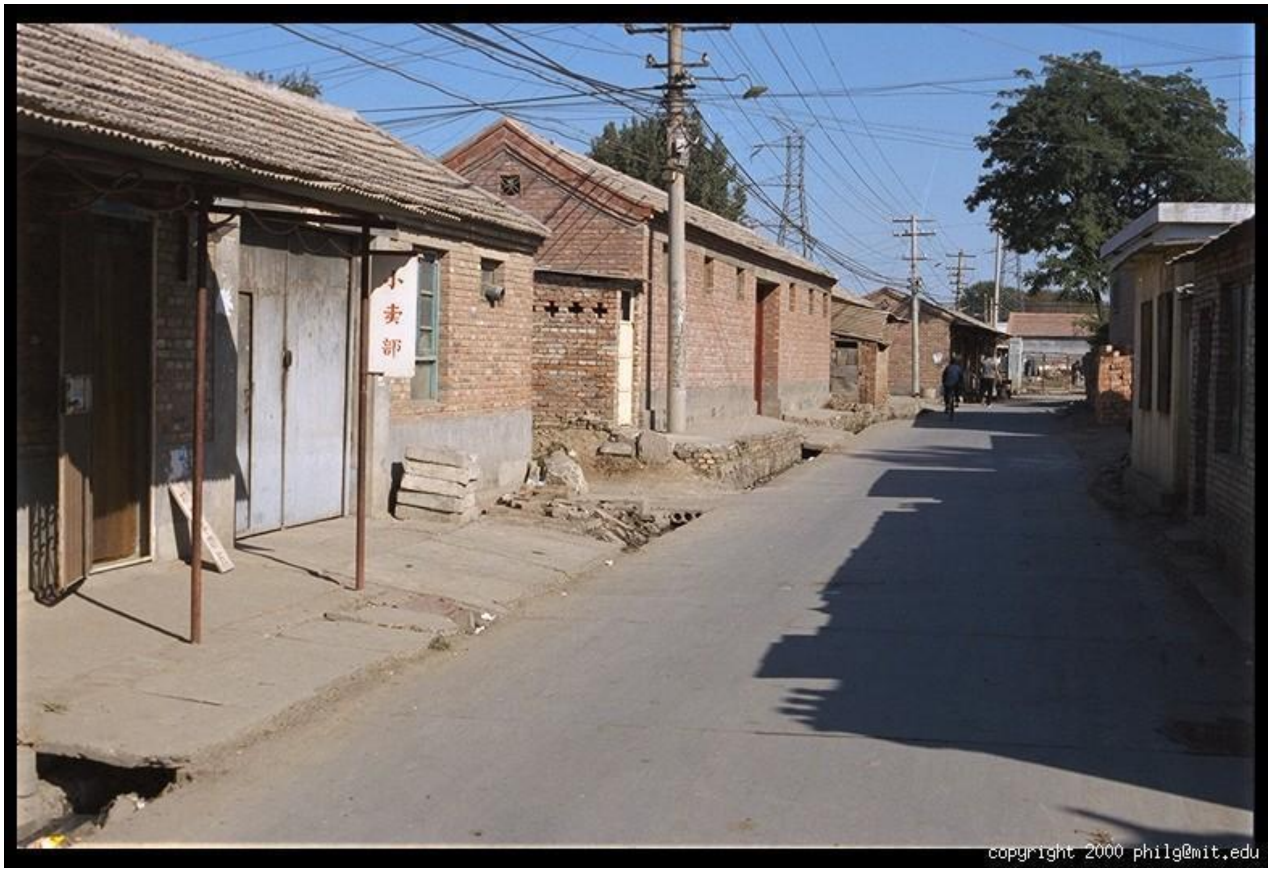
Only color shown – constant intensity

Most information in intensity



Only intensity shown – constant color

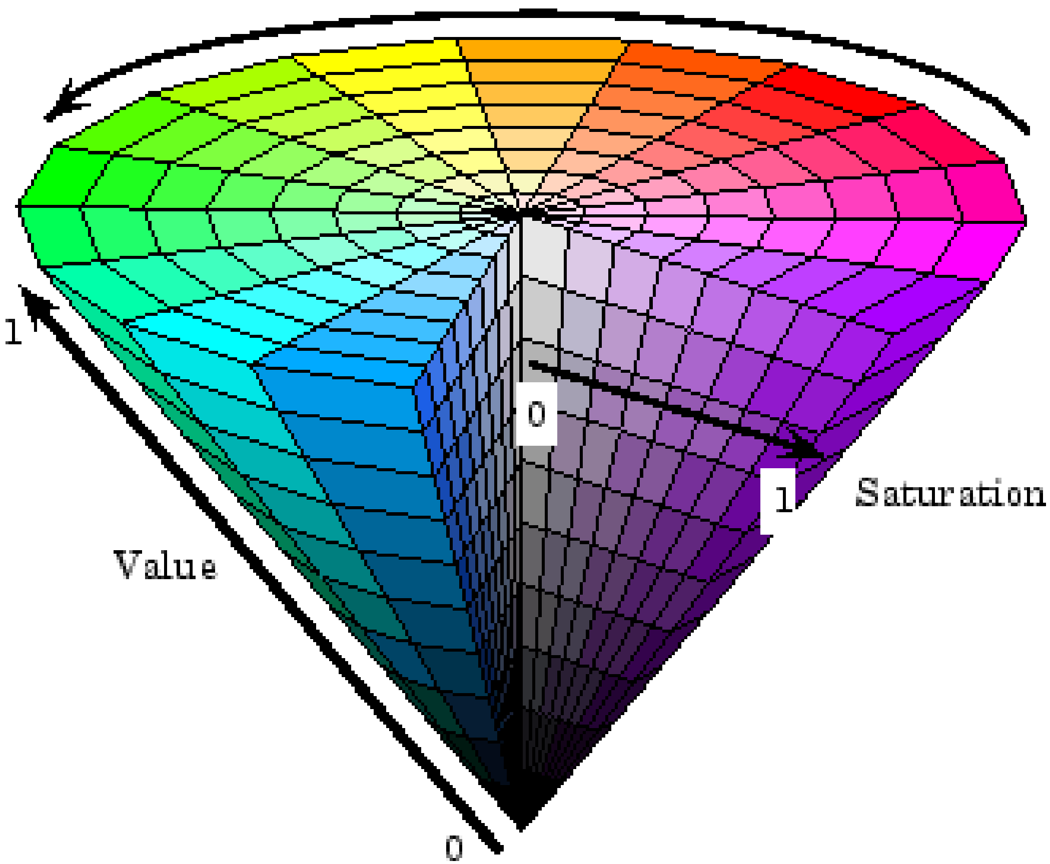
Most information in intensity

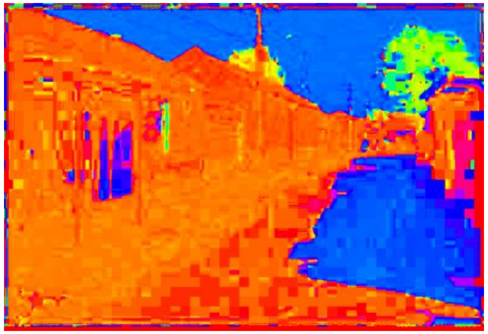


Original image

Color spaces: HSV

Intuitive color space





H
(S=1,V=1)

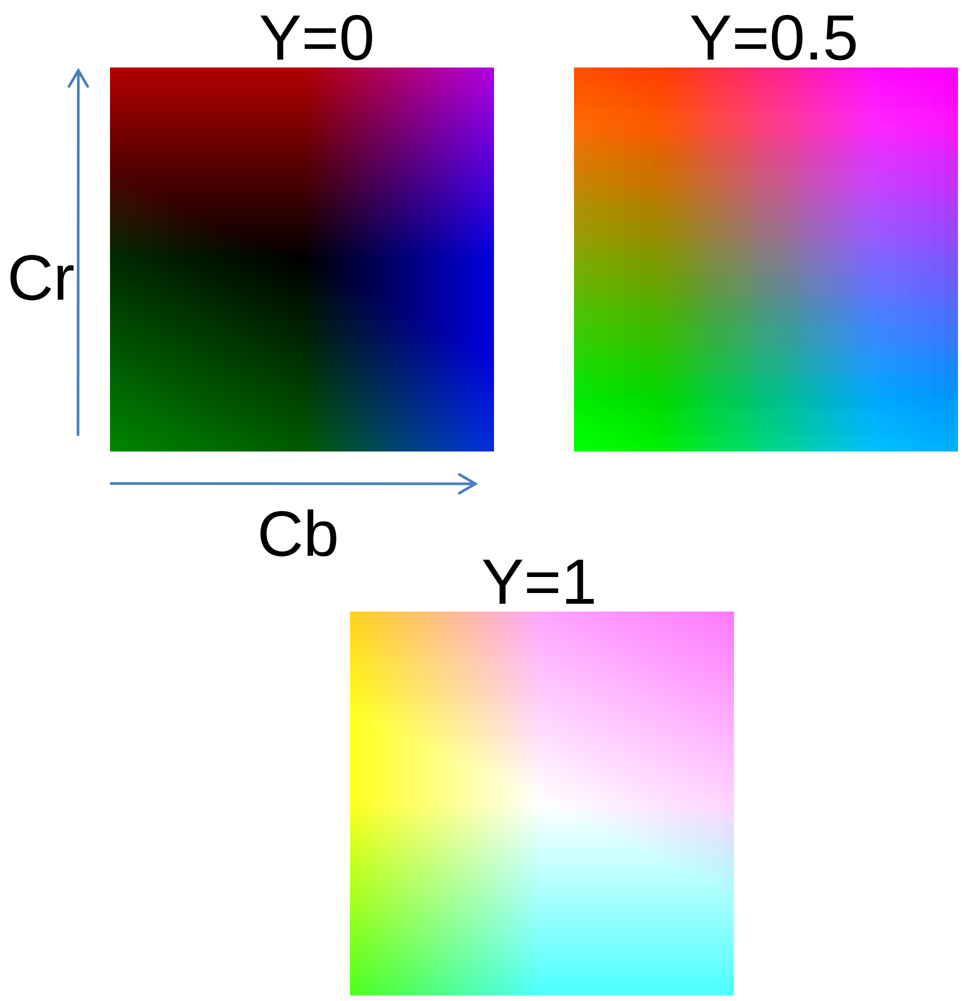
S
(H=1,V=1)

V
(H=1,S=0)

James Hays

Color spaces: YCbCr

Fast to compute, good for compression, used by TV



Y
(Cb=0.5,Cr=0.5)

Cb
(Y=0.5,Cr=0.5)

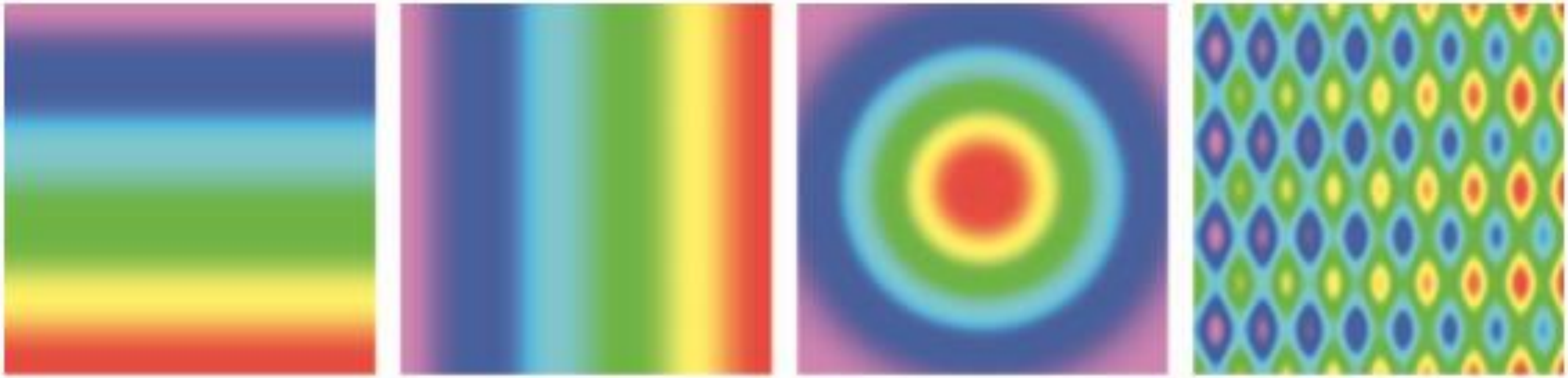
Cr
(Y=0.5,Cb=05)

James Hays

Most JPEG images & videos subsample chroma

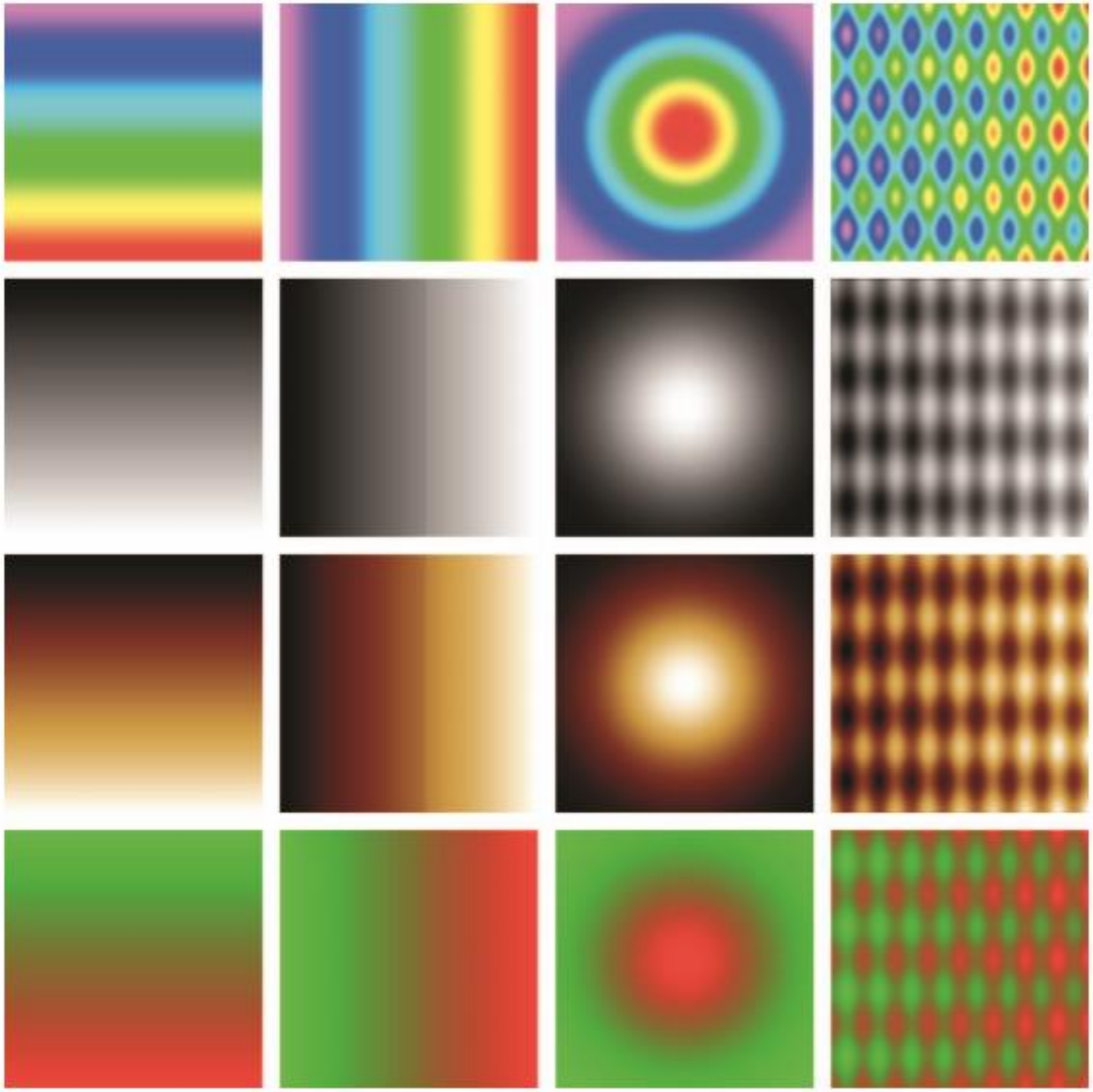
PSP Comp 3
2x2 Chroma subsampling
285K

Original
1,261K lossless
968K PNG



Rainbow color map considered harmful

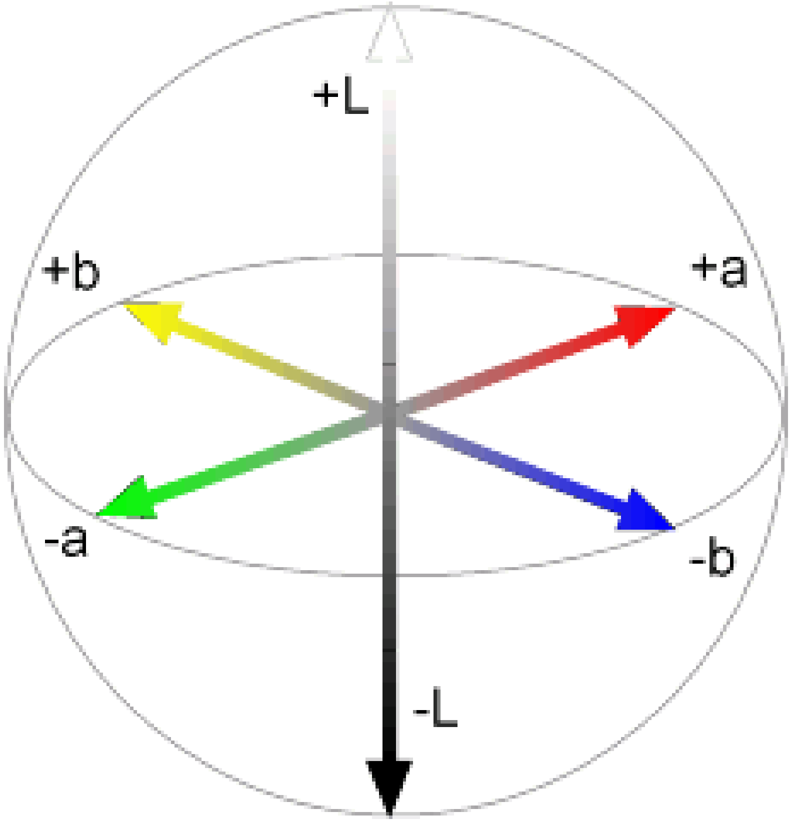
Borland and Taylor



IS COLOR PERCEPTION A VECTOR SPACE?

Color spaces: L*a*b*

"Perceptually uniform"* color space



L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)

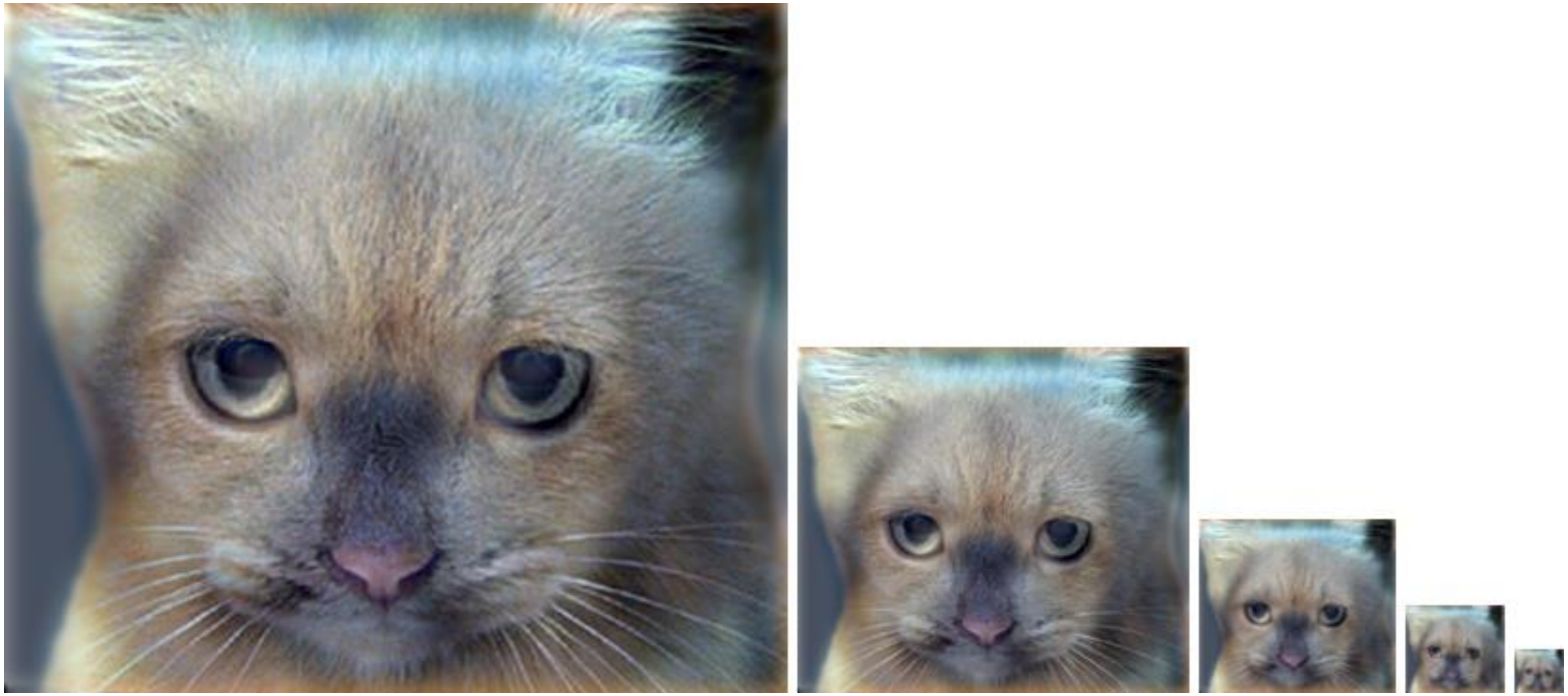
James Hays

Next week

- Convolution
- Filtering
- Image Pyramids
- Frequencies

Proj 1: Image Filtering and Hybrid Images

- Implement image filtering to separate high and low frequencies.
- Combine high frequencies and low frequencies from different images to create a scale-dependent image.



James Hays