Images

CS 4620 Lecture 38

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

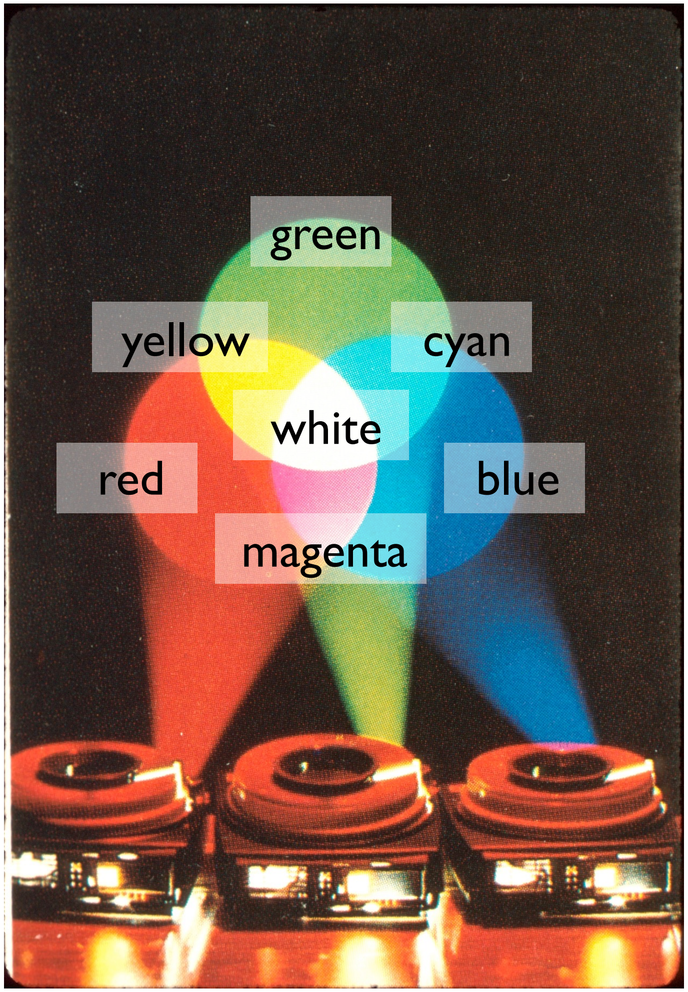
Announcements

- A7 extended by 24 hours

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

Color displays

- Operating principle: humans are trichromatic
  - match any color with blend of three
  - therefore, problem reduces to producing 3 images and blending
- Additive color
  - blend images by sum
  - e.g. overlapping projection
  - e.g. unresolved dots
  - R, G, B make good primaries

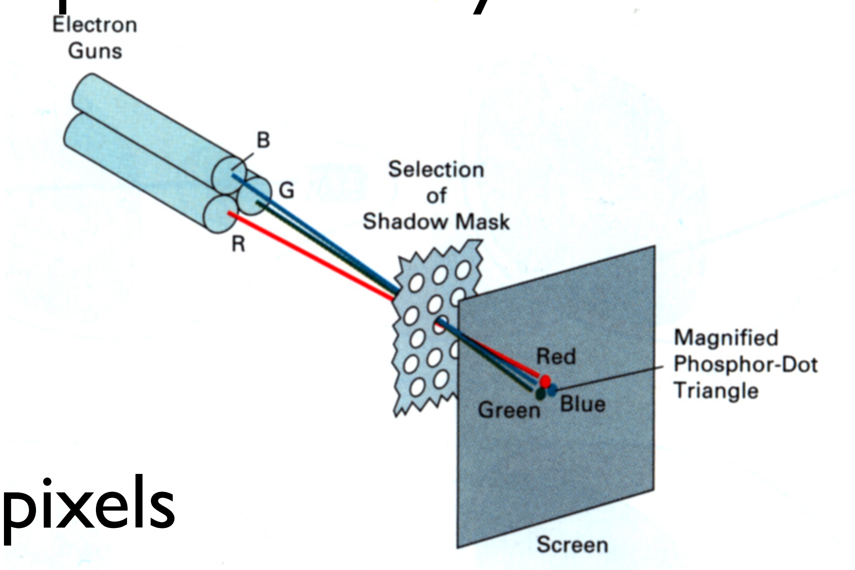


[cs417 S02 slides]

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

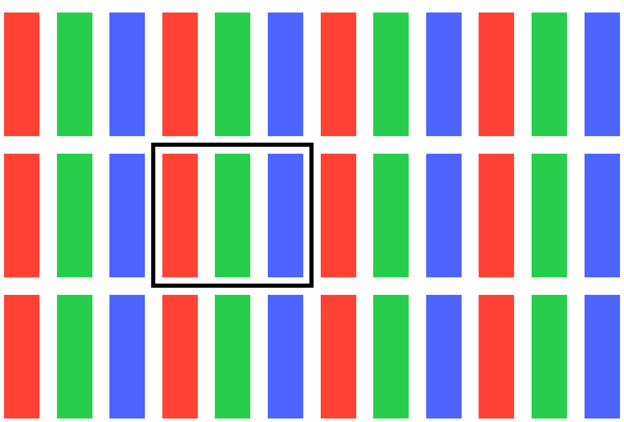
Color displays

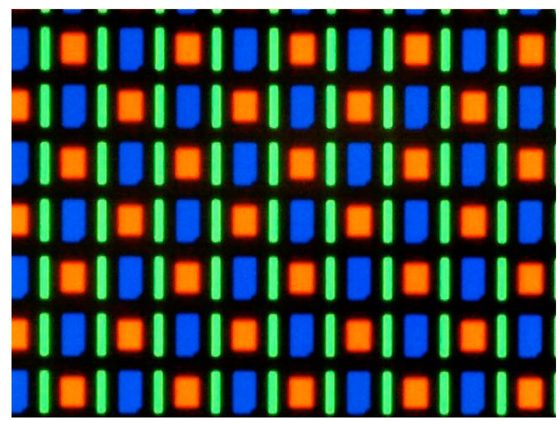
- CRT: phosphor dot pattern to produce finely interleaved color images



[H&B fig. 2-10]

- LCD, LED: interleaved R,G,B pixels



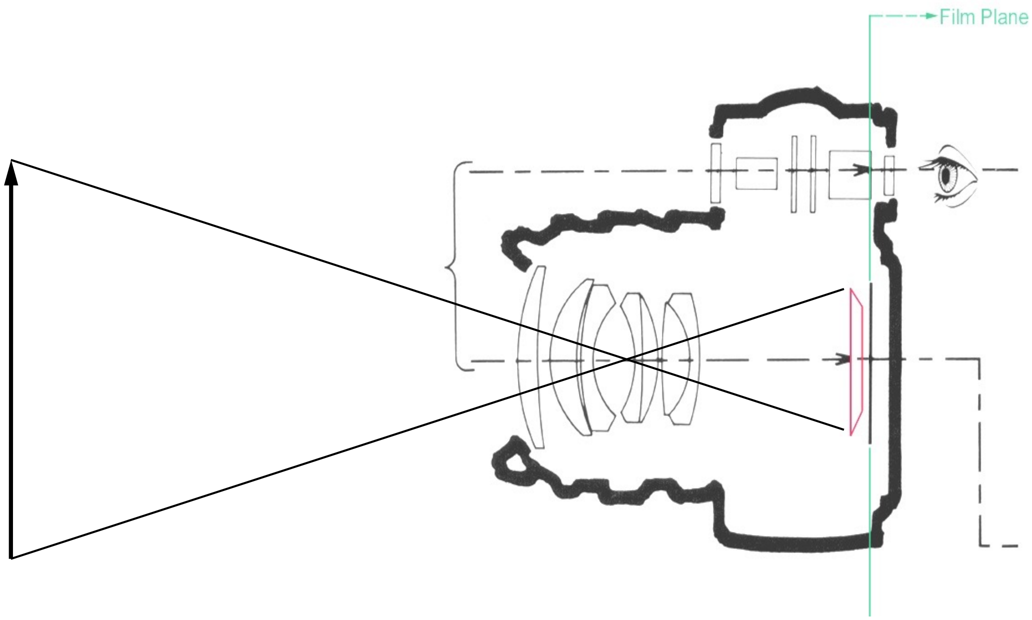


[Wikimedia Commons]

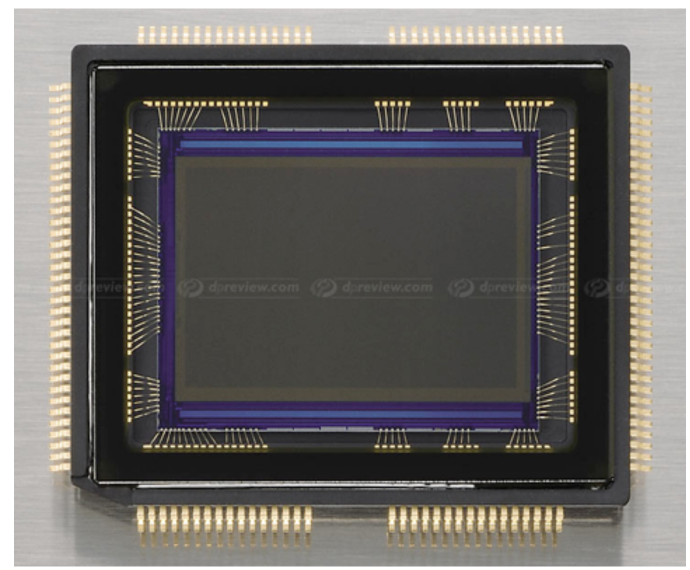
© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

Digital camera

- A raster input device
- Image sensor contains 2D array of photosensors



[CS 417 Spring 2002]

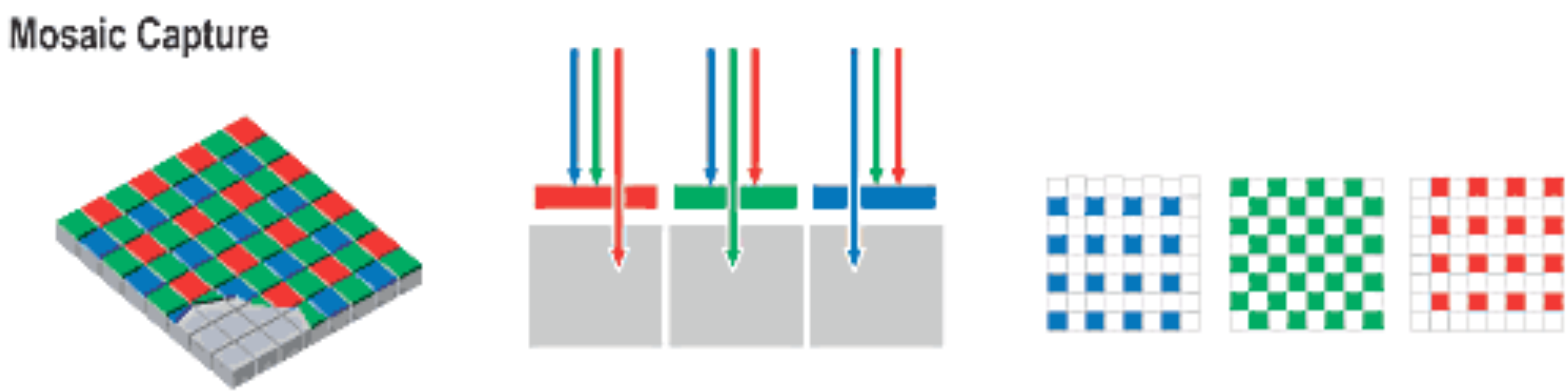


[dpreview.com]

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

Digital camera

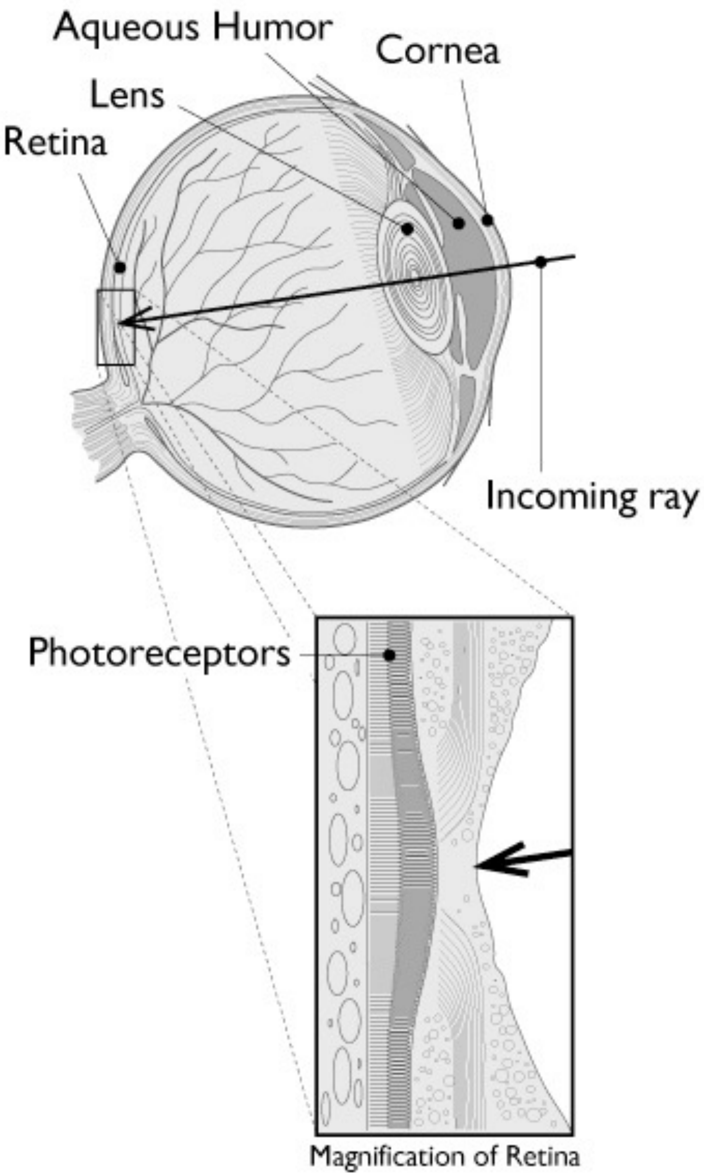
- Color typically captured using color mosaic



[Foveon]

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

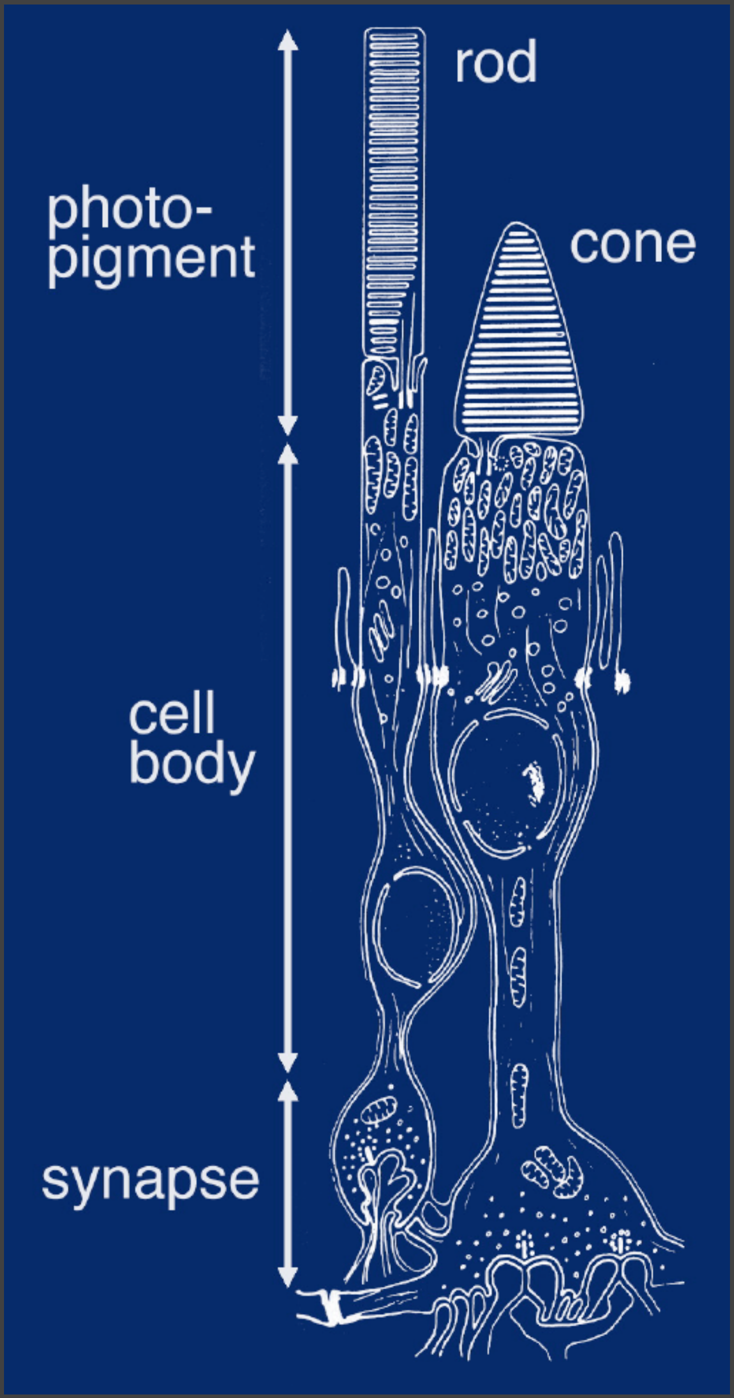
The eye as a measurement device



[Greger et al. 1995]

- We can model the low-level behavior of the eye by thinking of it as a light-measuring machine
  - –its optics are much like a camera
  - –its detection mechanism is also much like a camera
- Light is measured by the *photoreceptors* in the retina
  - –they respond to visible light
  - –different types respond to different wavelengths

© 2012 Kavita Bala •
(with previous instructors James/Marschner)



Photoreceptors

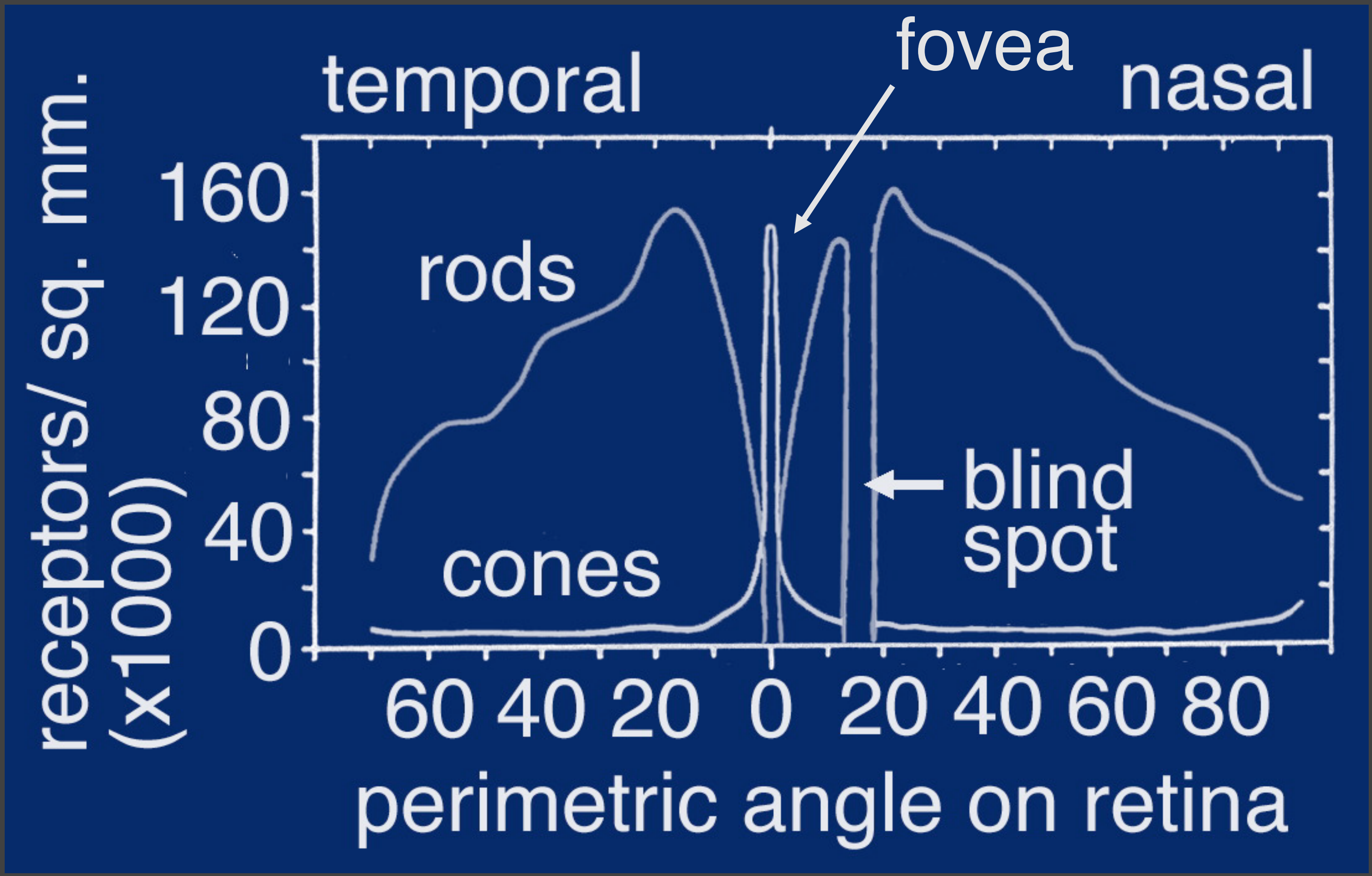
- 120 million rods
- 7-8 million cones in each eye

- rods: scotopic
- cones: photopic

© Kavita Bala, Computer Science, Cornell University

Adapted from Levine, Vision in Man and Machine © McGraw-Hill, 1985.

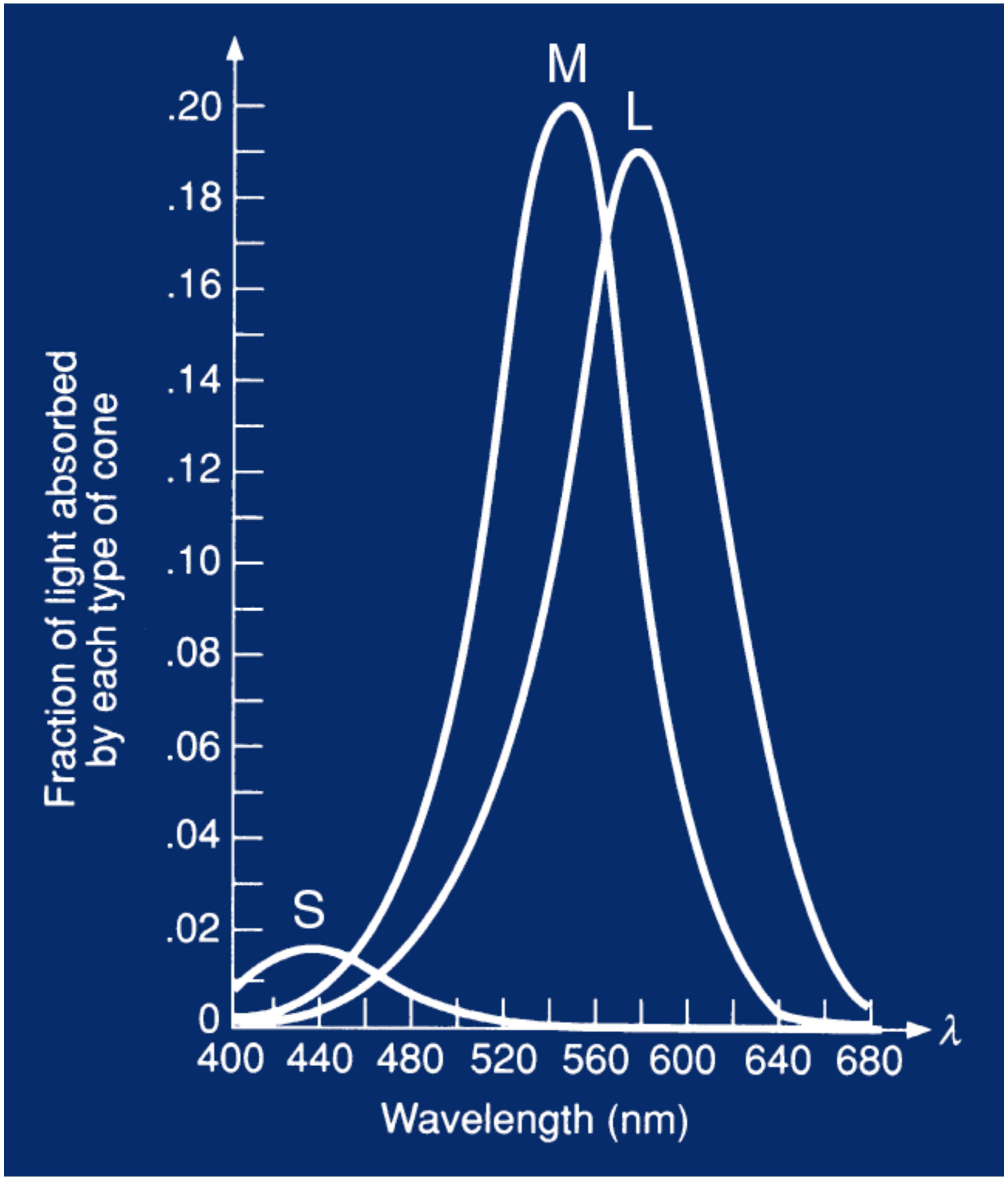
Receptor distribution



© Kavita Bala, Computer Science, Cornell University

Adapted from Levine, Vision in Man and Machine © McGraw-Hill, 1985.

Cone Responses



- S,M,L cones have broadband spectral sensitivity
- Results in a trichromatic visual system
- S, M, and L are *tristimulus values*

[source unknown]

© 2012 Kavita Bala

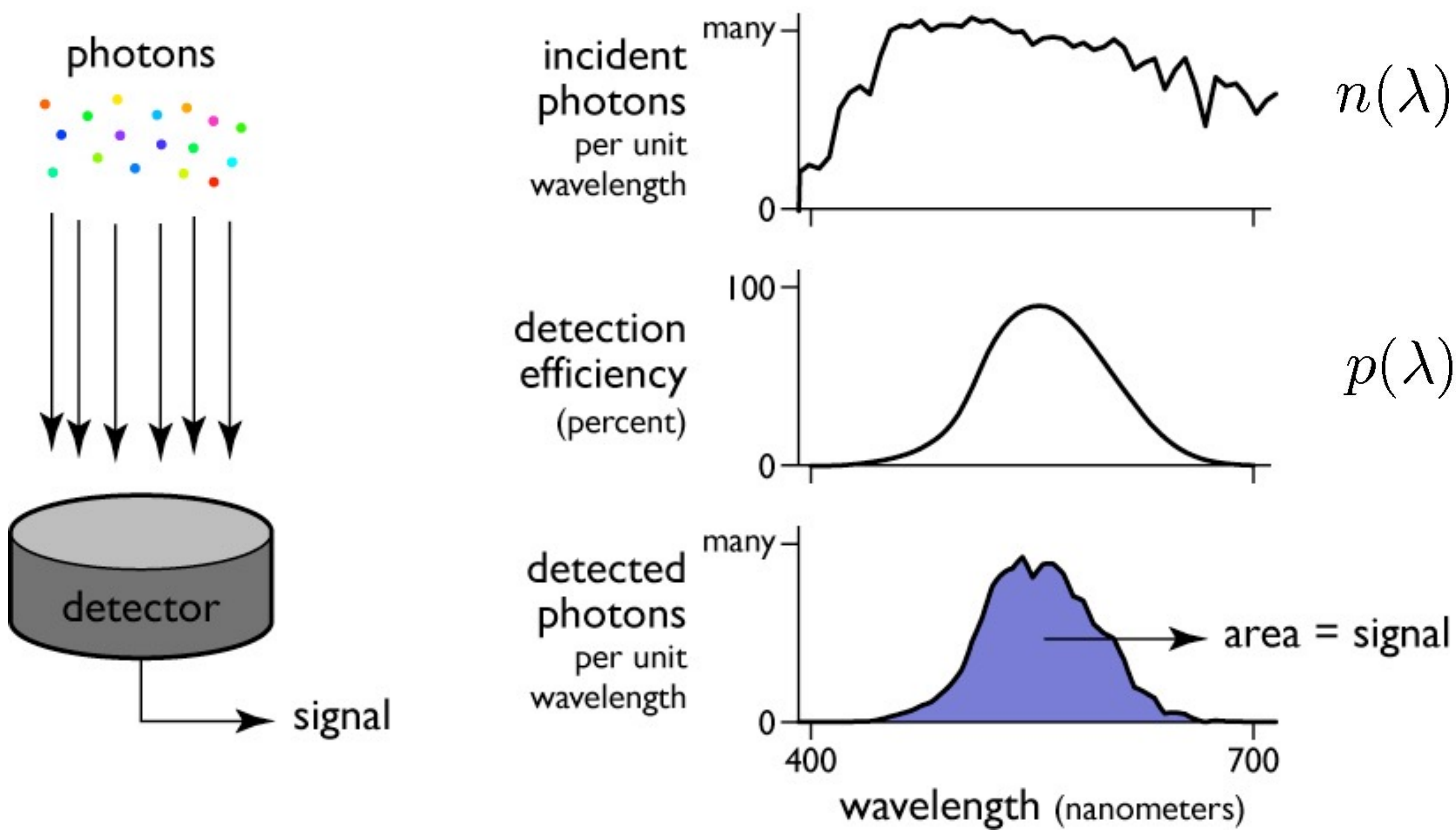
(with previous instructors James/Marschner)

A simple light detector

- Produces a scalar value (a number) when photons land on it
  - –this value depends strictly on the number of photons detected
  - –each photon has a probability of being detected that depends on the wavelength
  - –there is no way to tell the difference between signals caused by light of different wavelengths: there is just a number
- This model works for many detectors:
  - –based on semiconductors (such as in a digital camera)
  - –based on visual photopigments (such as in human eyes)

© 2012 Kavita Bala • 11
(with previous instructors James/Marschner)

A simple light detector



$$X = \int n(\lambda)p(\lambda)\,d\lambda$$

© 2012 Kavita Bala • 12
(with previous instructors James/Marschner)

Light detection math

- Same math carries over to power distributions
  - –spectum entering the detector has its spectral power distribution (SPD), $s(\lambda)$
  - –detector has its *spectral sensitivity* or *spectral response*, $r(\lambda)$

$$X = \int s(\lambda) r(\lambda) \, d\lambda$$

measured signal

detector's sensitivity

input spectrum

© 2012 Kavita Bala

(with previous instructors James/Marschner)

Light detection math

$$X = \int s(\lambda) r(\lambda)\, d\lambda \qquad \text{or} \qquad X = s \cdot r$$

- If we think of *s* and *r* as vectors, this operation is a dot product (aka inner product)
  - in fact, the computation is done exactly this way, using sampled representations of the spectra.
    - let λ_i be regularly spaced sample points $\Delta\lambda$ apart; then:

$$\tilde{s}[i] = s(\lambda_i); \tilde{r}[i] = r(\lambda_i)$$

$$\int s(\lambda) r(\lambda)\, d\lambda \approx \sum_i \tilde{s}[i] \tilde{r}[i]\, \Delta\lambda$$

- this sum is very clearly a dot product

© 2012 Kavita Bala •

(with previous instructors James/Marschner)

Cone responses to a spectrum *s*

$$S = \int r_S(\lambda)s(\lambda)\,d\lambda = r_S \cdot s$$

$$M = \int r_M(\lambda)s(\lambda)\,d\lambda = r_M \cdot s$$

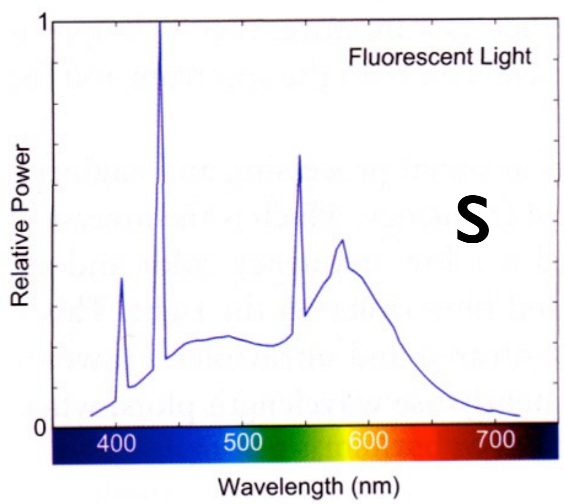
$$L = \int r_L(\lambda)s(\lambda)\,d\lambda = r_L \cdot s$$

© 2012 Kavita Bala •

(with previous instructors James/Marschner)

Colorimetry: mapping light to signals

- Want to map a *Physical light description* to a *Perceptual color sensation*
- Basic solution was known and standardized by 1930



$$S = r_S \cdot s$$
$$M = r_M \cdot s$$
$$L = r_L \cdot s$$

Physical

Perceptual

[Stone 2003]

© 2012 Kavita Bala

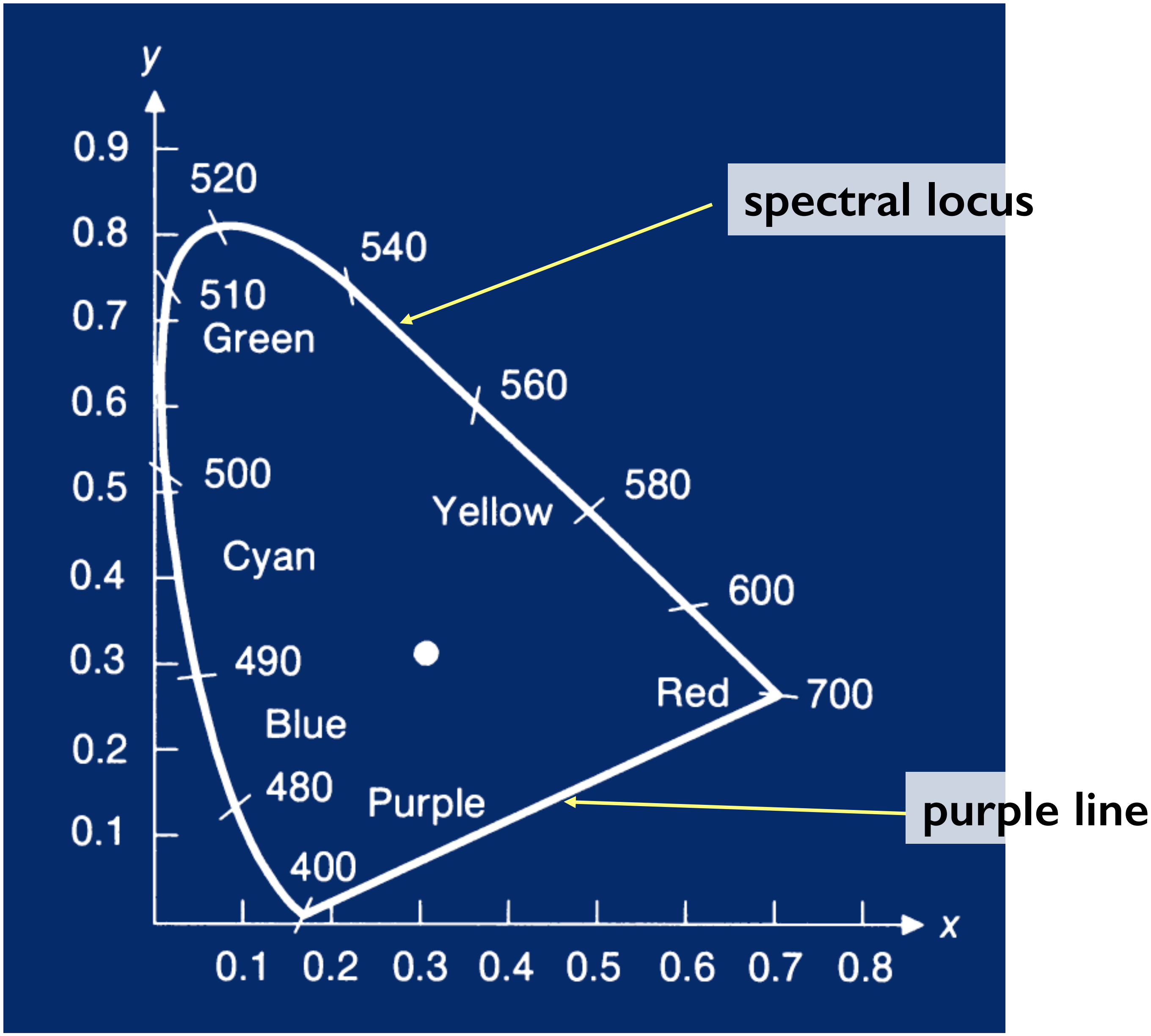
(with previous instructors James/Marschner)

Basic fact of colorimetry

- Take a spectrum (which is a function)
- Eye produces three numbers
- This throws away a lot of information!
  - Quite possible to have two different spectra that have the same S, M, L tristimulus values
  - Two such spectra are *metamers*

© 2012 Kavita Bala • 17
(with previous instructors James/Marschner)

Chromaticity Diagram

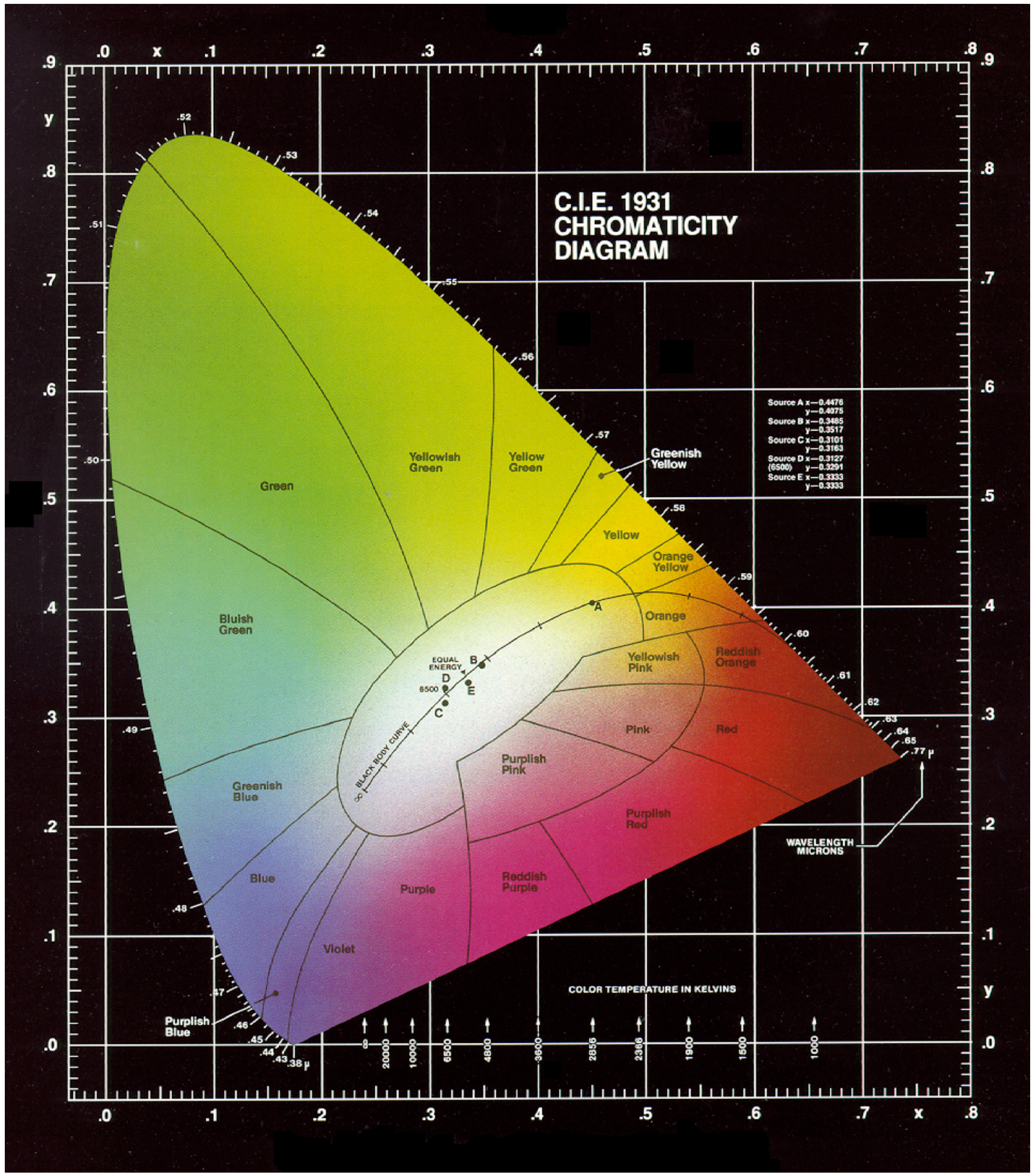


[source unknown]

© 2012 Kavita Bala

(with previous instructors James/Marschner)

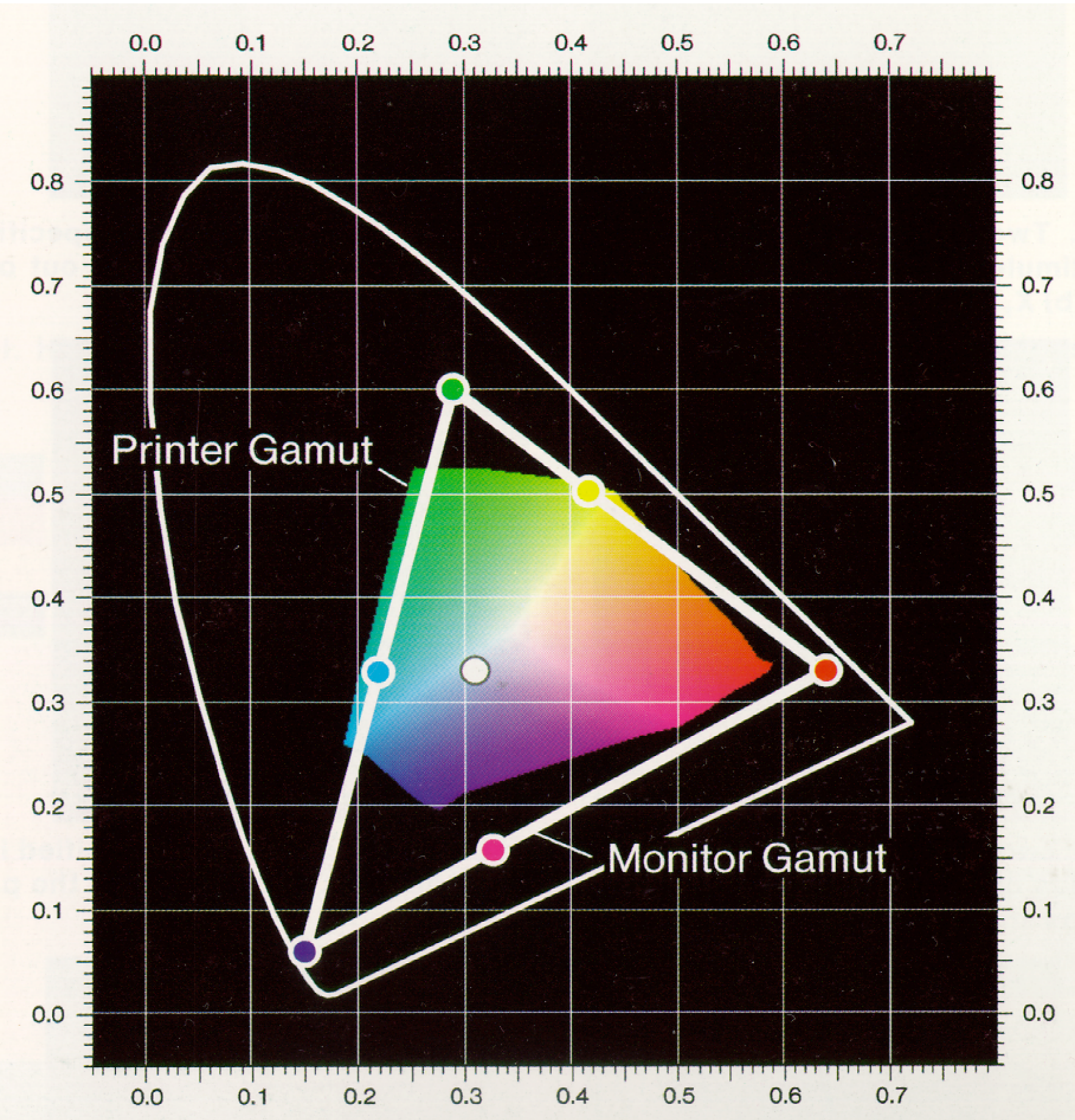
Chromaticity Diagram



[source unknown]

© 2012 Kavita Bala • 19
(with previous instructors James/Marschner)

Color Gamuts



[source unknown]

Monitors/printers can’t produce all visible colors

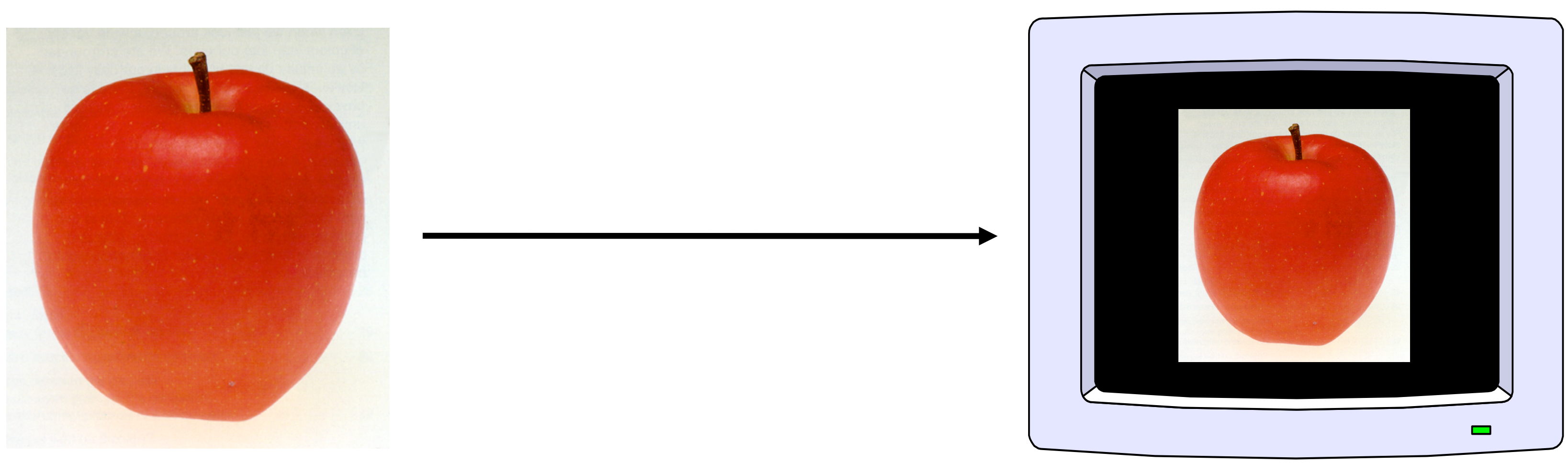
Reproduction is limited to a particular domain

For additive color (e.g. monitor) gamut is the triangle defined by the chromaticities of the three primaries.

© 2012 Kavita Bala • 20
(with previous instructors James/Marschner)

Color reproduction

- Have a spectrum *s*; want to match on RGB monitor
  - "match" means it looks the same
  - any spectrum that projects to the same point in the visual color space is a good reproduction
- Must find a spectrum that the monitor *can* produce that is a metamer of *s*



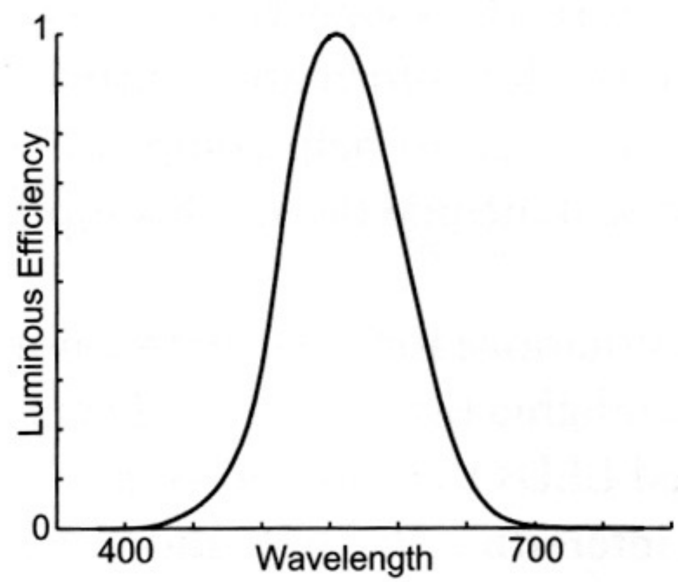
R, G, B?

[cs417—Greenberg]

© 2012 Kavita Bala • 21
(with previous instructors James/Marschner)

Basic colorimetric concepts

- Luminance
  - the overall magnitude of the the visual response to a spectrum (independent of its color)
    - corresponds to the everyday concept "brightness"
  - determined by product of SPD with the *luminous efficiency function* V_λ that describes the eye's overall ability to detect light at each wavelength
  - e.g. lamps are optimized to improve their luminous efficiency (tungsten vs. fluorescent vs. sodium vapor)



[Stone 2003]

© 2012 Kavita Bala

(with previous instructors James/Marschner)

Luminance, mathematically

- Y just has another response curve (like *S*, *M*, and *L*)

$$Y = r_Y \cdot s$$

  – r_Y is really called "V_λ"

- V_λ is a linear combination of *S*, *M*, and *L*
  – Has to be, since it's derived from cone outputs

© 2012 Kavita Bala

(with previous instructors James/Marschner)

More basic colorimetric concepts

- Chromaticity
  - –what's left after luminance is factored out (the color without regard for overall brightness)
  - –scaling a spectrum up or down leaves chromaticity alone
- Dominant wavelength
  - –many colors can be matched by white plus a spectral color
  - –correlates to everyday concept "hue"
- Purity
  - –ratio of pure color to white in matching mixture
  - –correlates to everyday concept "colorfulness" or "saturation"

© 2012 Kavita Bala

(with previous instructors James/Marschner)

Datatypes for raster images

- Bitmaps: `boolean` per pixel (1 bpp):
  - interp. = black and white; e.g. fax
- Grayscale: integer per pixel:
  - interp. = shades of gray; e.g. black-and-white print
  - precision: usually `byte` (8 bpp); sometimes 10, 12, or 16 bpp
- Color: 3 integers per pixel:
  - interp. = full range of displayable color; e.g. color print
  - precision: usually `byte[3]` (24 bpp)
  - sometimes 16 (5+6+5) or 30 or 36 or 48 bpp
- Floating point: more abstract, because no output device has infinite range
  - provides *high dynamic range* (HDR)
  - represent real scenes independent of display
  - becoming the standard intermediate format in graphics processor

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

Intensity encoding in images

- What do the numbers in images (pixel values) mean?
  - they determine how bright that pixel is
  - for floating point pixels, they directly give the intensity (in some units) — they are *linearly* related to the intensity
  - for pixels encoded in integers, this mapping is **not direct**
- *Transfer function*: function that maps input pixel value to luminance of displayed image

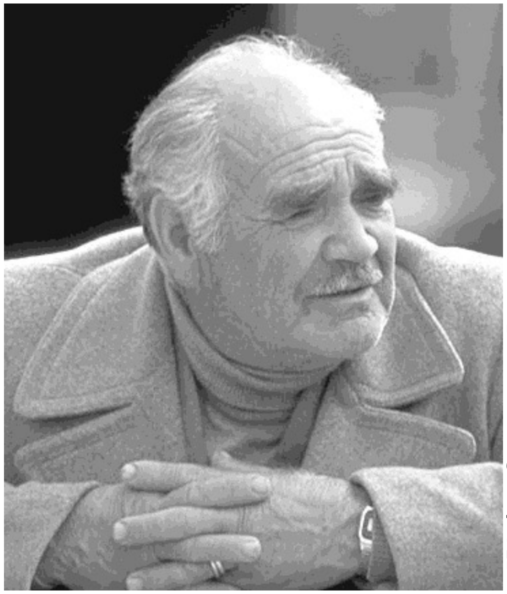
$$I = f(n) \quad f : [0, N] \to [I_{\min}, I_{\max}]$$

- What determines this function?
  - physical constraints of device or medium
  - desired visual characteristics

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Transfer function shape

- Desirable property: the change from one pixel value to the next highest pixel value should not produce a visible contrast
  - otherwise smooth areas of images will show visible bands



[Philip Greenspun]

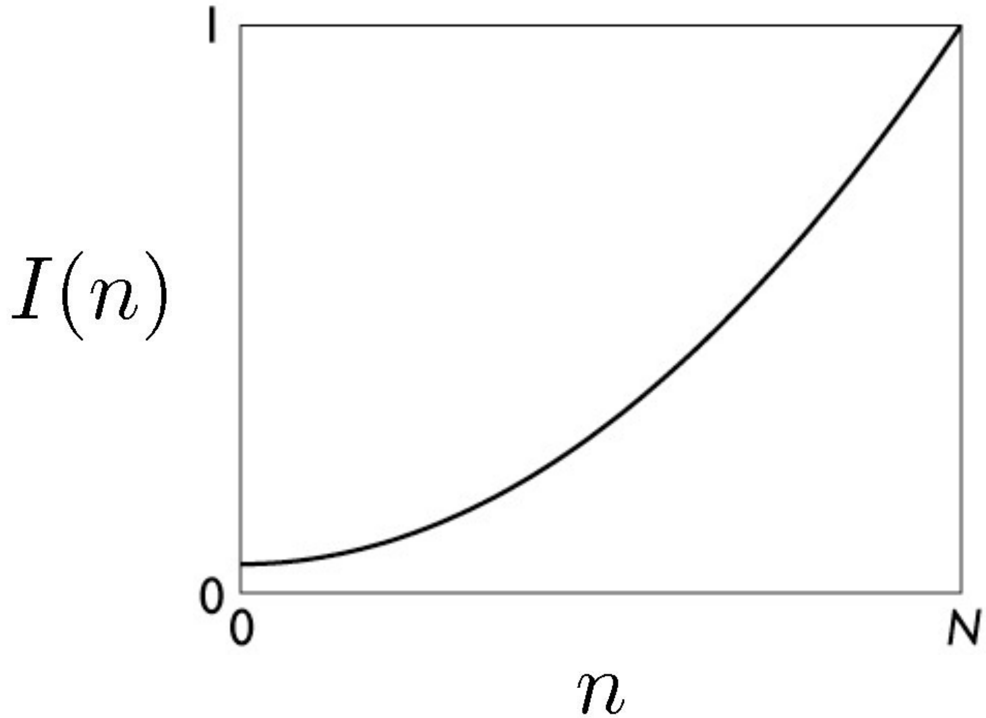
an image with severe *bandin*

- What contrasts are visible?
  - rule of thumb: under good conditions we can notice a 2% change in intensity
  - therefore we generally need smaller quantization steps in the darker tones than in the lighter tones
  - most efficient quantization is logarithmic

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Transfer function

- Something like this:



© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Constraints on transfer function

- Maximum displayable intensity, I_{max}
  - how much power can be channeled into a pixel?
    - LCD: backlight intensity, transmission efficiency (<10%)
    - projector: lamp power, efficiency of imager and optics
- Minimum displayable intensity, I_{min}
  - light emitted by the display in its "off" state
    - e.g. stray electron flux in CRT, polarizer quality in LCD
- Viewing flare, *k*: light reflected by the display
  - very important factor determining image contrast in practice
    - 5% of I_{max} is typical in a normal office environment [sRGB spec]
    - much effort to make very black CRT and LCD screens
    - all-black decor in movie theaters

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Dynamic range

- Dynamic range $R_d = I_{max} / I_{min}$, or $(I_{max} + k) / (I_{min} + k)$
  - determines the degree of image contrast that can be achieved
  - a major factor in image quality
- Ballpark values
  - Desktop display in typical conditions: 20:1
  - Photographic print: 30:1
  - Desktop display in good conditions: 100:1
  - High-end display under ideal conditions: 1000:1
  - Digital cinema projection: 1000:1
  - Photographic transparency (directly viewed): 1000:1
  - High dynamic range display: 10,000:1

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

How many levels are needed?

- Depends on dynamic range
  - 2% steps are most efficient:

$$0 \mapsto I_{\min}; 1 \mapsto 1.02 I_{\min}; 2 \mapsto (1.02)^2 I_{\min}; \ldots$$

  - log 1.02 is about 1/120, so 120 steps per decade of dynamic range
    - 240 for desktop display
    - 480 to drive HDR display
- If we want to use linear quantization (equal steps)
  - one step must be < 2% (1/50) of $I_{\min}$
  - need to get from ~0 to $I_{\min} \cdot R_d$, so need about 50 R_d levels
    - 1500 for a print; 5000 for desktop display; 500,000 for HDR display
- Moral: 8 bits is just barely enough for low-end applications
  - but only if we are careful about quantization

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Intensity quantization in practice

- Option 1: linear quantization $I(n) = (n/N)\, I_{\max}$
  - pro: simple, convenient, amenable to arithmetic
  - con: requires more steps (wastes memory)
  - need 12 bits for any useful purpose; more than 16 for HDR
- Option 2: power-law quantization $I(n) = (n/N)^{\gamma}\, I_{\max}$
  - pro: fairly simple, approximates ideal exponential quantization
  - con: need to linearize before doing pixel arithmetic
  - con: need to agree on exponent
  - 8 bits are OK for many applications; 12 for more critical ones

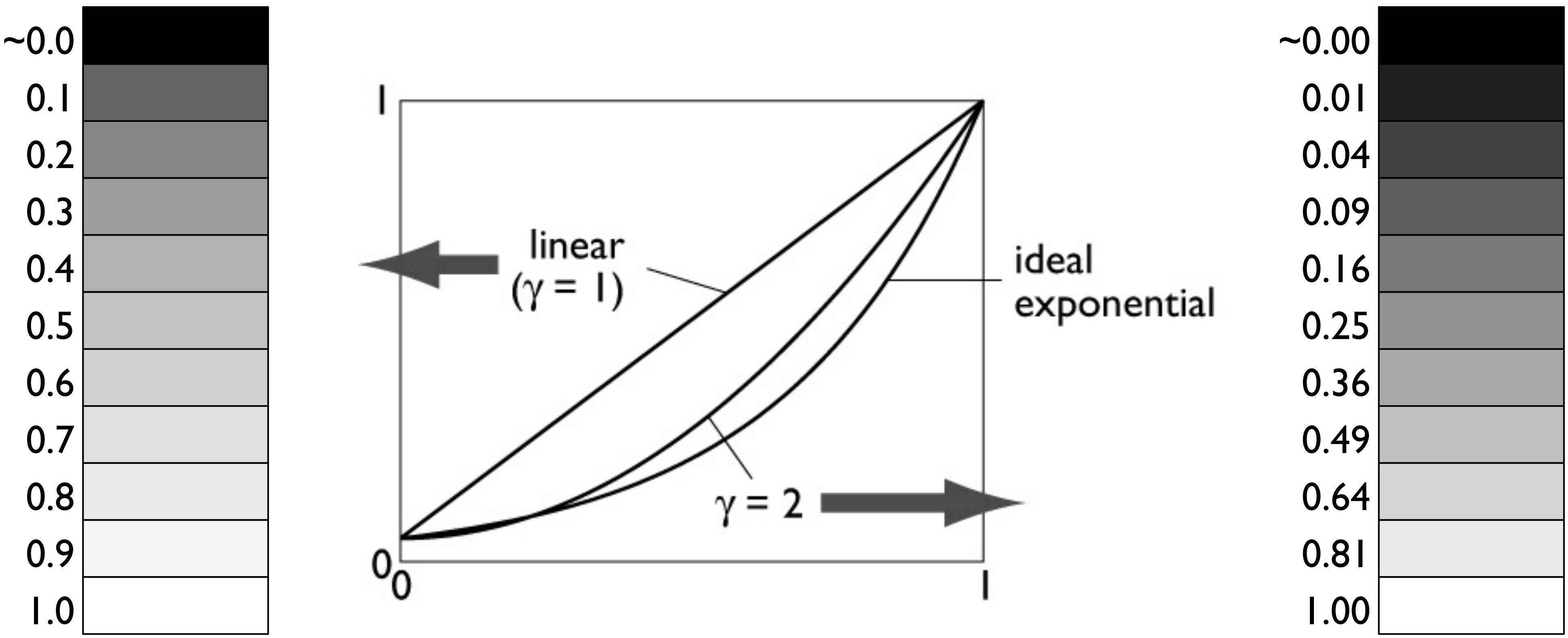
© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Why gamma?

- Power-law quantization, or *gamma correction* is most popular
- Original reason: CRTs are like that
  - intensity on screen is proportional to (roughly) voltage^2
- Continuing reason: inertia + memory savings
  - inertia: gamma correction is close enough to logarithmic that there's no sense in changing
  - memory: gamma correction makes 8 bits per pixel an acceptable option

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Gamma quantization



- Close enough to ideal perceptually uniform exponential

© 2015 Kavita Bala
w/ prior instructor Steve Marschner •

Gamma correction

- Sometimes (often, in graphics) we have computed intensities *a* that we want to display linearly
- In the case of an ideal monitor with zero black level,

$$I(n) = (n/N)^{\gamma}$$

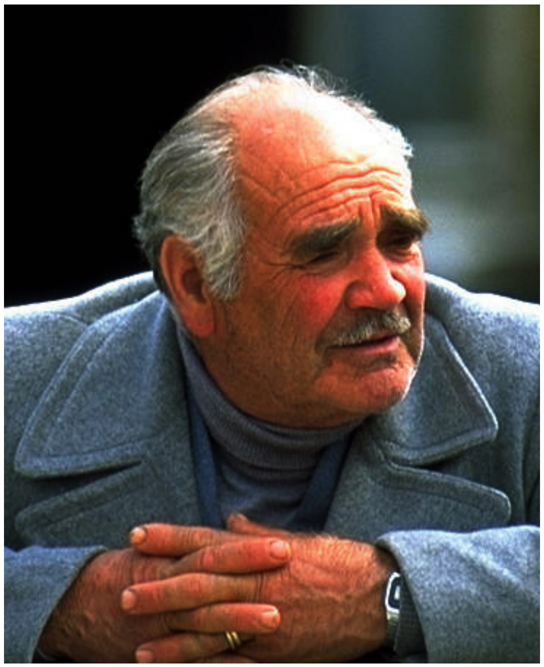
(where N = $2^n - 1$ in *n* bits). Solving for *n*:

$$n(I) = NI^{\frac{1}{\gamma}}$$

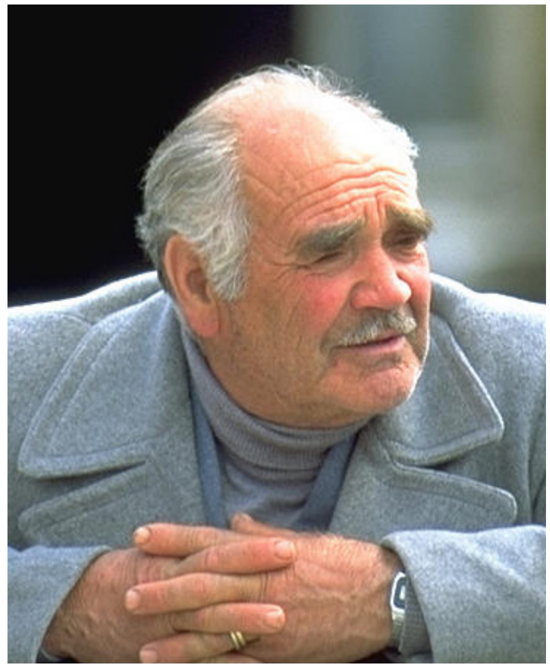
- This is the "gamma correction" recipe that has to be applied when computed values are converted to 8 bits for output
  - failing to do this (implicitly assuming gamma = 1) results in dark, oversaturated images

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

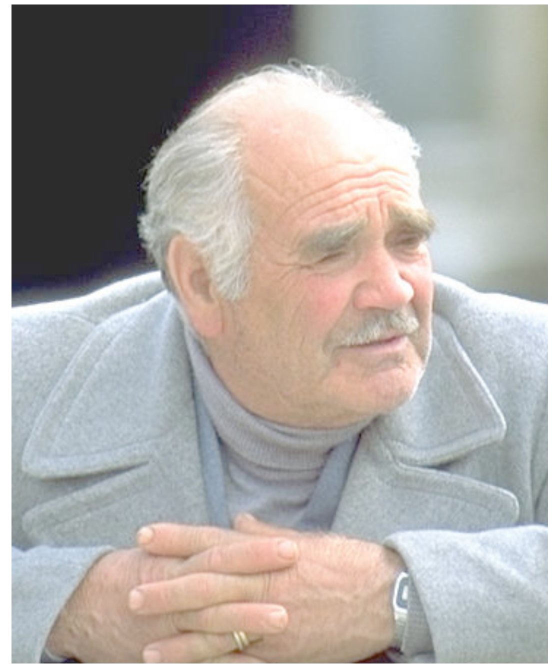
Gamma correction



corrected for γ lower than display



OK



corrected for γ higher than display

[Philip Greenspun]

© 2015 Kavita Bala

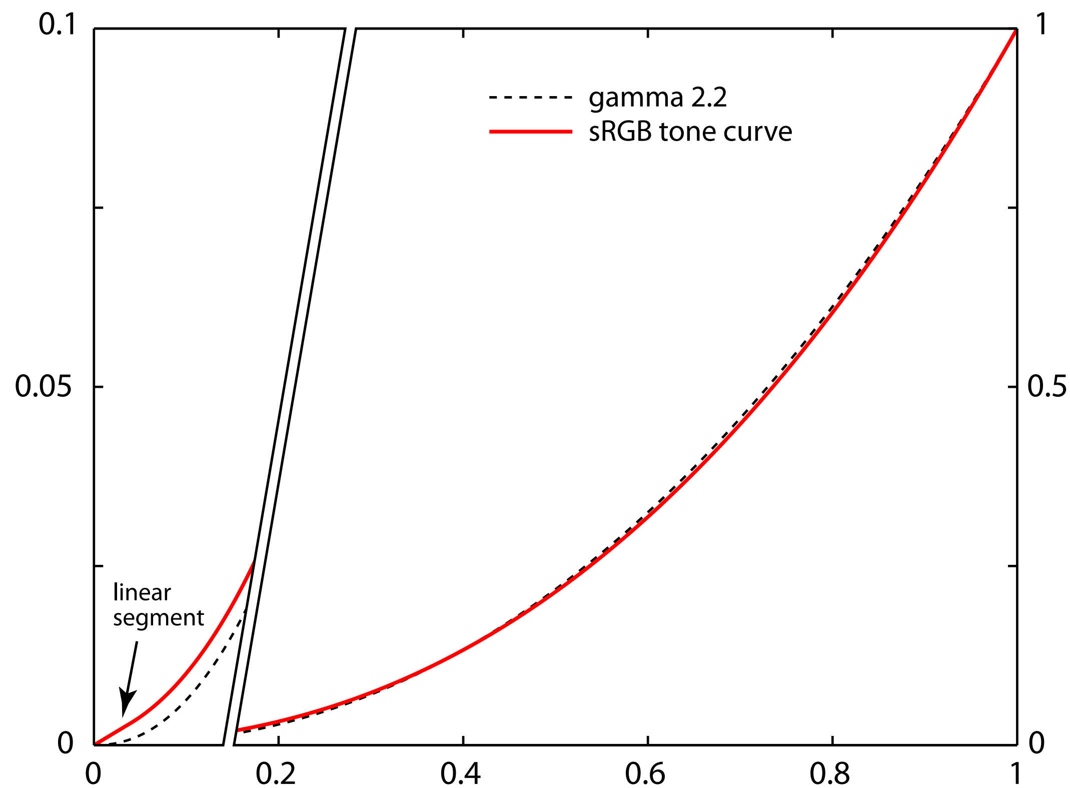
sRGB quantization curve

- The predominant standard for "casual color" in computer displays
  - consistent with older typical practice
  - designed to work well under imperfect conditions
  - these days all monitors are calibrated to sRGB by default
  - in practice, usually defines what your pixel values mean

$$I(C) = \begin{cases} \frac{C}{12.92}, & C \leq 0.04045 \\ \left(\frac{C+a}{1+a}\right)^{2.4}, & C > 0.04045 \end{cases}$$

$$C = n/N$$

$$a = 0.055$$



[derived from a figure by Dick Lyon]

© 2015 Kavita Bala
w/ prior instructor Steve Marschner

Converting from HDR to LDR

- "High dynamic range" — pixels can be arbitrarily bright or dark
- "Low dynamic range" — there are limits on the min and max

- Simplest solution: just scale and clamp
- More flexible: introduce a contrast control

- Scale factor *a* is "exposure"
  - often quoted on a power-of-2 scale

© 2015 Kavita Bala
w/ prior instructor Steve Marschner