Capturing Light… in man and machine

CS194: Image Manipulation, Comp. Vision, and Comp. Photo
Alexei Efros, UC Berkeley, Spring 2020
Computer Vision
Algorithms and Applications

Richard Szeliski

Springer

http://szeliski.org/Book/
General Comments

Prerequisites

• Linear algebra!!! (EE16A, Math 54, or Math 110)
• Good programming skills (at least CS61B)
• Machine Learning experience helpful

Emphasis on programming projects!

• Building something from scratch

Graduate Version:

• Final project required (not pre-canned), including conference-style report paper

“No Screens” Policy:

• No laptops, no cell phones, no smartphones, etc.
Getting help outside of class

Course Web Page
  * http://inst.eecs.berkeley.edu/~cs194-26/

Discussion board:
  * piazza.com

Office hours
  * Me: after class plus TBA
  * see webpage and piazza

Homework Parties
  * Weekly, evenings in Woz (TBA)

Python Image Tutorial
  * Next Tuesday here at 8pm (hopefully)
Administrative Stuff

Grading

- Programming Project (60%)
- Exam + popup quizzes (20%) \(\Rightarrow\) Have pen and paper
- Final Project (20%)
- Class Participation: priceless

Late Policy

- Five (5) emergency late days for semester, to be spent wisely
- Max 10% of full credit afterwards

Extra Points

- Most projects will have optional “bells & whistles”
- These extra points could be used to pad scores on other projects (but not exams!)
Academic Integrity

• Can discuss projects, but don’t share code

• Don’t look up code or copy from a friend

• If you’re not sure if it’s allowed, ask

• Acknowledge any inspirations

• If you get stuck, come talk to us
Why you should NOT take this class

• Project-based class
  • No canned problem sets
  • Not theory-heavy (but will read a few research papers)
  • No clean rubrics
  • Open-ended by design

• Need time to think, not just hack
  • Creativity is a class requirement

• Lots of work…There are easier classes if
  • you just need some units
  • you care more about the grade than about learning stuff

• Not worth it if you don’t enjoy it
Now… reasons TO take this class

• It’s your reward after 3 grueling years 😊
• You get work with pictures, unleash your creative potential
• Interested in grad school? 😊
Other Questions?
Etymology

PHOTOGRAPHY

light
drawing / writing
Image Formation

Digital Camera

The Eye
Sensor Array

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.
Sampling and 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.
Interlace vs. progressive scan


Slide by Steve Seitz
Progressive scan

Rolling Shutter

http://en.wikipedia.org/wiki/Rolling_shutter
Saccadic eye movement
Saccadic eye movement
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 “film”?
  - photoreceptor cells (rods and cones) in the retina
The Retina

Cross-section of eye

Cross section of retina

Ganglion axons
Ganglion cell layer
Bipolar cell layer
Receptor layer

Pigmented epithelium

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
The famous sock-matching problem...
Distribution of Rods and Cones

# Night Sky: why are there more stars off-center?
3.4 THE SPATIAL MOSAIC OF THE HUMAN CONES. Cross sections of the human retina at the level of the inner segments showing (A) cones in the fovea, and (B) cones in the periphery. Note the size difference (scale bar = 10 μm), and that, as the separation between cones grows, the rod receptors fill in the spaces. (C) Cone density plotted as a function of distance from the center of the fovea for seven human retinas; cone density decreases with distance from the fovea. Source: Curcio et al., 1990.
Electromagnetic Spectrum

Human Luminance Sensitivity Function

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

Why do we see light of these wavelengths?

...because that's where the Sun radiates EM energy
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

© Stephen E. Palmer, 2002
The Physics of Light

Some examples of the reflectance spectra of surfaces

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>% Photons Reflected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>400 - 700</td>
</tr>
<tr>
<td>Yellow</td>
<td>400 - 700</td>
</tr>
<tr>
<td>Blue</td>
<td>400 - 700</td>
</tr>
<tr>
<td>Purple</td>
<td>400 - 700</td>
</tr>
</tbody>
</table>

© Stephen E. Palmer, 2002
The Psychophysical Correspondence

There is no simple functional description for the perceived color of all lights under all viewing conditions, but ……

A helpful constraint:
Consider only physical spectra with normal distributions

<table>
<thead>
<tr>
<th>Wavelength (nm.)</th>
<th># Photons</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td></td>
</tr>
</tbody>
</table>

© Stephen E. Palmer, 2002
The Psychophysical Correspondence

Mean \leftrightarrow \text{Hue}

# Photons

Wavelength

blue, green, yellow
The Psychophysical Correspondence

Variance ↔ Saturation

# Photons

Wavelength

© Stephen E. Palmer, 2002
The Psychophysical Correspondence

Area \leftrightarrow \text{Brightness}

# Photons

Wavelength

© Stephen E. Palmer, 2002
Physiology of Color Vision

Three kinds of cones:

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

Rods and cones act as filters on the spectrum

- To get the output of a filter, multiply its response curve by the spectrum, integrate over all wavelengths
  - Each cone yields one number

- How can we represent an entire spectrum with 3 numbers?
- We can’t! Most of the information is lost
  - As a result, two different spectra may appear indistinguishable
    » such spectra are known as metamers
More Spectra

The graph shows the reflectance spectra of various objects, including
- yellow flower
- orange flower
- white flower
- orange berry
- violet flower
- blue flower

The graph is labeled with wavelengths in nanometers (nm) on the x-axis and reflectance on the y-axis.
Color Constancy

The “photometer metaphor” of color perception:
Color perception is determined by the spectrum of light on each retinal receptor (as measured by a photometer).
Color Constancy

The “photometer metaphor” of color perception: Color perception is determined by the spectrum of light on each retinal receptor (as measured by a photometer).
The “photometer metaphor” of color perception:
Color perception is determined by the spectrum of light on each retinal receptor (as measured by a photometer).
Color Constancy

Do we have constancy over all global color transformations?

60% blue filter

Complete inversion
**Color Constancy**: the ability to perceive the invariant color of a surface despite ecological Variations in the conditions of observation.

Another of these hard *inverse problems*: Physics of light emission and surface reflection underdetermine perception of surface color.
Camera White Balancing

- Manual
  - Choose color-neutral object in the photos and normalize

- Automatic (AWB)
  - Grey World: force average color of scene to grey
  - White World: force brightest object to white
Color Sensing in Camera (RGB)

3-chip vs. 1-chip: quality vs. cost

Why more green?

Why 3 colors?

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

Slide by Steve Seitz
Green is in!
Practical Color Sensing: Bayer Grid

Estimate RGB at ‘G’ cells from neighboring values

http://www.cooldictionary.com/words/Bayer-filter.wiki
Color Image
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
Color spaces

How can we represent color?

Color spaces: RGB

Default color space

RGB cube

- Easy for devices
- But not perceptual
- Where do the grays live?
- Where is hue and saturation?

Hue, Saturation, Value (Intensity)

- RGB cube on its vertex

Decouples the three components (a bit)

Use rgb2hsv() and hsv2rgb() in Matlab
Color spaces: HSV

Intuitive color space

H (S=1, V=1)
S (H=1, V=1)
V (H=1, S=0)
Color spaces: \(L^*a^*b^*\)

“Perceptually uniform”* color space

\[L (a=0, b=0)\]

\[a (L=65, b=0)\]

\[b (L=65, a=0)\]
Programming Project #1

Prokudin-Gorskii’s Color Photography (1907)
Programming Project #1
Programming Project #1

• How to compare R,G,B channels?
• No right answer
  • Sum of Squared Differences (SSD):
    \[ sssd(u, v) = \sum_{(x,y) \in N} [I(u + x, v + y) - P(x, y)]^2 \]
  • Normalized Correlation (NCC):
    \[ ncc(u, v) = \frac{\sum_{(x,y) \in N} [I(u + x, v + y) - \bar{I}] [P(x, y) - \bar{P}]}{\sqrt{\sum_{(x,y) \in N} [I(u + x, v + y) - \bar{I}]^2} \sqrt{\sum_{(x,y) \in N} [P(x, y) - \bar{P}]^2}} \]