Color

- Used heavily in human vision
- Color is a pixel property, making some recognition problems easy
- Visible spectrum for humans is 400nm (blue) to 700 nm (red)
- Machines can “see” much more; ex. X-rays, infrared, radio waves
Imaging Process (review)
Factors that Affect Perception

- **Light:** the spectrum of energy that illuminates the object surface
- **Reflectance:** ratio of reflected light to incoming light
- **Specularity:** highly specular (shiny) vs. matte surface
- **Distance:** distance to the light source
- **Angle:** angle between surface normal and light source
- **Sensitivity:** how sensitive is the sensor
Some physics of color

- White light is composed of all visible frequencies (400-700)
- Ultraviolet and X-rays are of much smaller wavelength
- Infrared and radio waves are of much longer wavelength
Coding methods for humans

- RGB is an additive system (add colors to black) used for displays
- CMY[K] is a subtractive system for printing
- HSV is a good perceptual space for art, psychology, and recognition
- YIQ used for TV is good for compression
Comparing color codes

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>CMY</th>
<th>HSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>(255, 0, 0)</td>
<td>(0, 255, 255)</td>
<td>(0.0, 1.0, 255)</td>
</tr>
<tr>
<td>YELLOW</td>
<td>(255, 255, 0)</td>
<td>(0, 0, 255)</td>
<td>(1.05, 1.0, 255)</td>
</tr>
<tr>
<td></td>
<td>(100, 100, 50)</td>
<td>(155, 155, 205)</td>
<td>(1.05, 0.5, 100)</td>
</tr>
<tr>
<td>GREEN</td>
<td>(0, 255, 0)</td>
<td>(255, 0, 255)</td>
<td>(2.09, 1.0, 255)</td>
</tr>
<tr>
<td>BLUE</td>
<td>(0, 0, 255)</td>
<td>(255, 255, 0)</td>
<td>(4.19, 1.0, 255)</td>
</tr>
<tr>
<td>WHITE</td>
<td>(255, 255, 255)</td>
<td>(0, 0, 0)</td>
<td>(-1.0, 0.0, 255)</td>
</tr>
<tr>
<td>GREY</td>
<td>(192, 192, 192)</td>
<td>(63, 63, 63)</td>
<td>(-1.0, 0.0, 192)</td>
</tr>
<tr>
<td></td>
<td>(127, 127, 127)</td>
<td>(128, 128, 128)</td>
<td>(-1.0, 0.0, 127)</td>
</tr>
<tr>
<td></td>
<td>(63, 63, 63)</td>
<td>(192, 192, 192)</td>
<td>(-1.0, 0.0, 63)</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>BLACK</td>
<td>(0, 0, 0)</td>
<td>(255, 255, 255)</td>
<td>(-1.0, 0.0, 0)</td>
</tr>
</tbody>
</table>

- Convenient to scale values in the range 0 to 1 in algorithms
- HSI values are computed from RGB values using Alg.
- $H \in [0.0, 2\pi)$, $S \in [0.0, 1.0]$ and $I \in [0.0, 255]$.
- Equal proportions of RGB yield grey.
- Equal proportions of R and G yield yellow.
RGB color cube

- R, G, B values normalized to (0, 1) interval
- Human perceives gray for triples on the diagonal
- “Pure colors” on corners
Color palette and normalized RGB

\[
\begin{align*}
\text{intensity } I &= (R + G + B)/3 \\
\text{normalized red } r &= R/(R + G + B) \\
\text{normalized green } g &= G/(R + G + B) \\
\text{normalized blue } b &= B/(R + G + B)
\end{align*}
\]

- Color triangle for normalized RGB coordinates.
- Blue (’b’) axis is out of page perpendicular to ’r’ and ’g’ axes.
- Triangle is a slice through the points \([1,0,0],[0,1,0],[0,0,1]\).
Color hexagon for HSI (HSV)

Color is coded relative to the diagonal of the color cube. Hue is encoded as an angle, saturation is the relative distance from the diagonal, and intensity is height.
(Left) Image of food originating from a digital camera; 
(center) saturation value of each pixel decreased 20%; 
(right) saturation value of each pixel increased 40%.
Properties of HSI (HSV)

- Separates out intensity $I$ from the coding
- Two values ($H \& S$) encode chromaticity
- Convenient for designing colors

- Hue $H$ is defined by an angle

- Saturation $S$ models the *purity* of the color
  - $S=1$ for a completely pure or saturated color
  - $S=0$ for a shade of “gray”
YIQ and YUV for TV signals

- Have better compression properties
- Luminance Y encoded using more bits than chrominance values I and Q; humans more sensitive to Y than I,Q
- NTSC TV uses luminance Y; chrominance values I and Q
- Luminance used by black/white TVs
- All 3 values used by color TVs
- YUV encoding used in some digital video and JPEG and MPEG compression
Conversion from RGB to YIQ

An approximate linear transformation from RGB to YIQ is

\[
\text{luminance } Y = 0.30R + 0.59G + 0.11B \\
R - \text{cyan } I = 0.60R - 0.28G - 0.32B \\
magenta - \text{green } Q = 0.21R - 0.52G + 0.31B
\]

We often use this for color to gray-tone conversion.
Colors can be used for image segmentation into regions

- Can cluster on color values and pixel locations

- Can use connected components and an approximate color criteria to find regions

- Can train an algorithm to look for certain colored regions – for example, skin color
Color Clustering by K-means Algorithm

Form K-means clusters from a set of n-dimensional vectors

1. Set ic (iteration count) to 1

2. Choose randomly a set of K means m1(1), ..., mK(1).

3. For each vector xi, compute D(xi,mk(ic)), k=1,...,K and assign xi to the cluster Cj with nearest mean.

4. Increment ic by 1, update the means to get m1(ic),...,mK(ic).

5. Repeat steps 3 and 4 until Ck(ic) = Ck(ic+1) for all k.
K-means Clustering Example

Original RGB Image

Color Clusters by K-Means
Extracting “white regions”

- Program learns white from training set of sample pixels.
- Aggregate similar neighbors to form regions.
- Components might be classified as characters.
- (Work contributed by David Moore.)

(Left) input RGB image

(Right) output is a labeled image.

CSE 803 Fall 2015
Skin color in RGB space

Purple region shows skin color samples from several people. Blue and yellow regions show skin in shadow or behind a beard.
Finding a face in video frame

- (left) input video frame
- (center) pixels classified according to RGB space
- (right) largest connected component with aspect similar to a face (all work contributed by Vera Bakic)
Color histograms can represent an image

- Histogram is fast and easy to compute.

- Size can easily be normalized so that different image histograms can be compared.

- Can match color histograms for database query or classification.
Histograms of two color images
Retrieval from image database

Top left image is query image. The others are retrieved by having similar color histogram (See Ch 8).
How to make a color histogram

- Make 3 histograms and concatenate them

- Create a single pseudo color between 0 and 255 by using 3 bits of R, 3 bits of G and 2 bits of B (which bits?)

- Can normalize histogram to hold frequencies so that bins total 1.0
Apples versus oranges

Separate HSI histograms for apples (left) and oranges (right) used by IBM’s VeggieVision for recognizing produce at the grocery store checkout station (see Ch 16).
Swain and Ballard’s Histogram Matching for Color Object Recognition

Opponent Encoding:

- \( wb = R + G + B \)
- \( rg = R - G \)
- \( by = 2B - R - G \)

Histograms: \( 8 \times 16 \times 16 = 2048 \) bins

Intersection of image histogram and model histogram:

\[
\text{intersection}(h(I), h(M)) = \sum_{j=1}^{\text{numbins}} \min\{h(I)[j], h(M)[j]\}
\]

Match score is the normalized intersection:

\[
\text{match}(h(I), h(M)) = \frac{\text{intersection}(h(I), h(M))}{\sum_{j=1}^{\text{numbins}} h(M)[j]}
\]
Histogram of Oriented Gradient

Figure 2. (a) The parametrization of a cell; (b) The gradient map; (c) The HOG of a block; (d) The HOG features of positive and negative samples.


https://www.youtube.com/watch?v=QaeJQTVnB-Y
Models of Reflectance

We need to look at models for the physics of illumination and reflection that will
1. help computer vision algorithms extract information about the 3D world,
and
2. help computer graphics algorithms render realistic images of model scenes.

Physics-based vision is the subarea of computer vision that uses physical models to understand image formation in order to better analyze real-world images.
The Lambertian Model: Diffuse Surface Reflection

A diffuse reflecting surface reflects light uniformly in all directions.

Uniform brightness for all viewpoints of a planar surface.
Real matte objects

Figure 6.16  Diffuse reflection from Lambertian objects—a vase and an egg—and a plot of intensities across the highlighted row. The intensities are closely related to the object shape. (Image courtesy of Deborah Trytten.)
Specular reflection is highly directional and mirrorlike.

R is the ray of reflection
V is direction from the surface toward the viewpoint
$\alpha$ is the shininess parameter

$$R = 2N(N \circ (-S)) \oplus S$$

specular reflected $i \sim (R \circ V)^\alpha$
Real specular objects

- Chrome car parts are very shiny/mirror like
- So are glass or ceramic objects
- And waxy plant leaves
Phong reflection model

- Reasonable realism, reasonable computing
- Uses the following components
  - (a) ambient light
  - (b) diffuse reflection component
  - (c) specular reflection component
  - (d) darkening with distance

Components (b), (c), (d) are summed over all light sources.

- Modern computer games use more complicated models.
Phong shading model uses

1. the reflective properties of surface elt imaging at \( I[x, y] \)
   
   - \( K_{d\lambda} \) is diffuse reflectivity
   - \( K_{s\lambda} \) for specular reflectivity
   - \( K_{q\lambda} \) is a vector of coefficients of reflection for different wavelengths \( \lambda \)

2. the position and characteristics of all \( M \) light sources
Phong model for intensity at wavelength lambda at pixel \([x, y] \]

\[
I_{\lambda}[x, y] = I_{a\lambda}K_{d\lambda} + \sum_{m=1}^{M} \left( \frac{1}{cd_m^2} I_{m\lambda}[K_{d\lambda}(n \circ s) + K_{s\lambda}(R_m \circ V)^\alpha] \right)
\]

\(I_{m\lambda}\) is the intensity of the light source \(m\) for wavelength \(\lambda\). (RGB)

The \(m - th\) light source is a distance \(d_m\) from the surface element and makes reflection ray \(R_m\) off the surface element.