Computer Vision

James L. Crowley

M2R MoSIG option GVR | Fall Semester
8 October 2009

Lesson 3

Color Spaces and Contrast Description

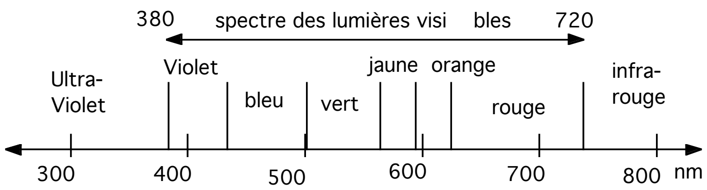
Lesson Outline:

1 Color Spaces and Color Models

1.1 Color Perception

The human retina is a tissue composed of rods, cones and bi-polar cells.
Cones are responsible for daytime vision.
Bi-polar cells perform initial image processing in the retina.

Rods provide night vision. Night vision is achromatique. It does not provide color perception. Night vision is low acuity - Rods are dispersed over the entire retina.

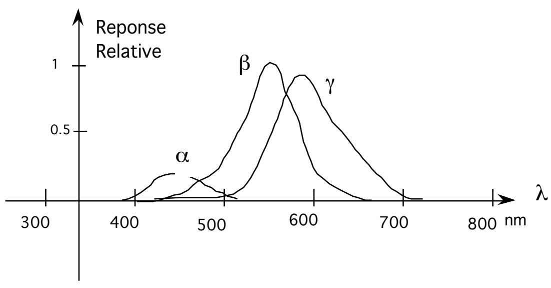


Rods are responsible for perception of very low light levels and provide night vision. Rods employ a very sensitive pigment named "rhodopsin".

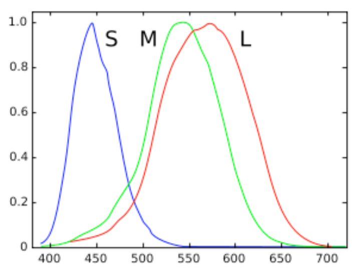
Rodopsin is sensitive to a large part of the visible spectrum of with a maximum sensitivity around 510 nano-meters.

Rhodopsin sensitive to light between 0.1 and 2 lumens, (typical moonlight) but is destroyed by more intense lights.

Rhodopsin can take from 10 to 20 minutes to regenerate.



Relative Sensitivities

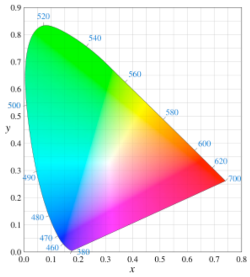


Normalised Sensitivities

Cones provide our chromatique "day vision". Human Cones employ 3 pigments :

cyanolabe α 400–500 nm peak at 420–440 nm
chlorolabe β 450–630 nm peak at 534–545 nm
erythrolabe γ 500–700 nm peak at 564–580 nm

Perception of cyanolabe is low probability, hence poor sensitivity to blue.
Perception of Chlorolabe and erythrolabe are more sensitive.



The three pigments give rise to a color space shown here (CIE model).

Note, these three pigments do NOT map directly to color perception.
Color perception is MUCH more complex, and includes a difficult to model phenomena known as "color constancy".

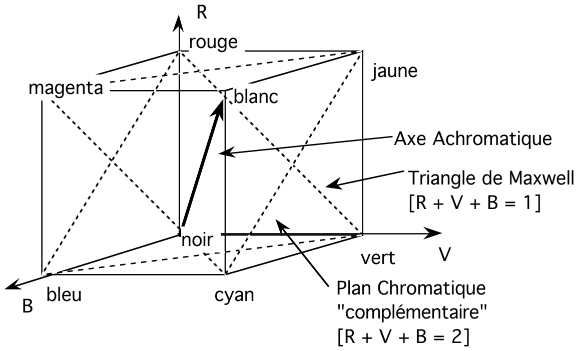
For example, yellow is always yellow, despite changes to the spectrum of an ambiant source

Many color models have been proposed but each has its strengths and weaknesses.

1.2 The RGB Color Model

One of the oldest color models, originally proposed by Isaac Newton. This is the model used by most color cameras.

The RGB model "pretends" that Red, Green and Blue are orthogonal (independent) axes of a Cartesian space.



The achromatic axis is R=G=B.
Maxwell's triangle is the surface defined when R+G+B = 1.
A complementary triangle exists when R+G+B = 2.

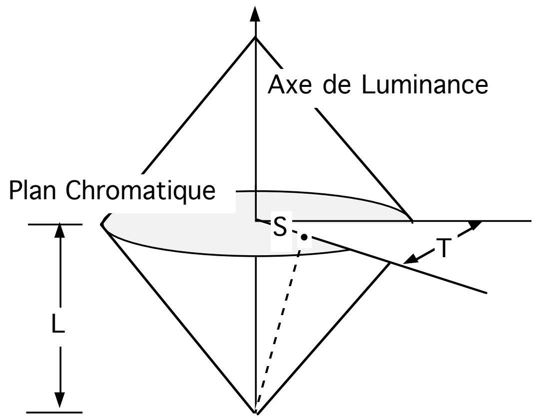
For printers (subtractive color) this is converted to CMY (Cyan, Magenta, Yellow).

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} R_{\max} \\ G_{\max} \\ B_{\max} \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

1.3 The HLS color model

HLS: Hue Luminance Saturation - called TLS in French.

Often used by artists.
HLS is a polar coordinate model for and hue (perceived color) and saturation.
The polar space is placed on a third axis. The size of the disc corresponds to the range of saturation values available.



One (of many possible) mappings from RGB:

Luminance : L = (R + B + B)

Saturation : 1 - 3*min(R, G, B)/L

Hue : $x = \cos^{-1}\left(\dfrac{\frac{1}{2}(R-G)+(R-B)}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right)$

if B>G then H = x else H = 2π–x.

1.4 Color Opponent Model

Color Constancy: The subjective perception of color is independent of the spectrum of the ambient illumination.

Subjective color perception is provide by "Relative" color and not "absolute" measurements.

This is commonly modeled using a Color Opponent space.

The opponent color theory suggests that there are three opponent channels: red versus green, blue versus yellow, and black versus white (the latter type is achromatic and detects light-dark variation, or luminance).

This can be computed from RGB by the following transformation:

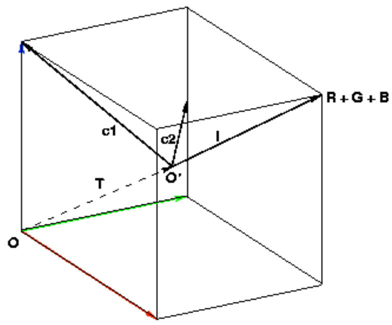
Luminance : L = R+G+B

Chrominance: C1 = (R-G)/2

C2 = B – (R+G)/2

as a matrix :

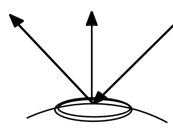
$$\begin{pmatrix} L \\ C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -0.5 & -0.5 & 1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$



Such a vector can be "steered" to accommodate changes in ambient illumination.

1.5 Separating Specular and Lambertian Reflection.

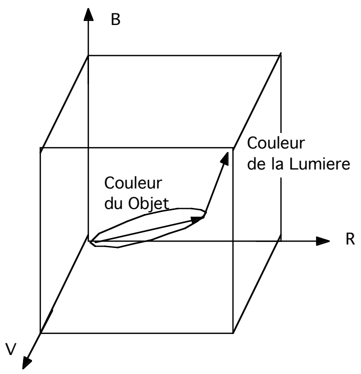
Consider what happens at a specular reflection.



The specularity has the same spectrum as the illumination.
The rest of the object has a spectrum that is the product of illumination and pigments.

This scan be seen in a histogram of color:

$$\forall \vec{C}(i,j) \quad : \quad H(\vec{C}(i,j)) = H(\vec{C}(i,j)) + 1$$

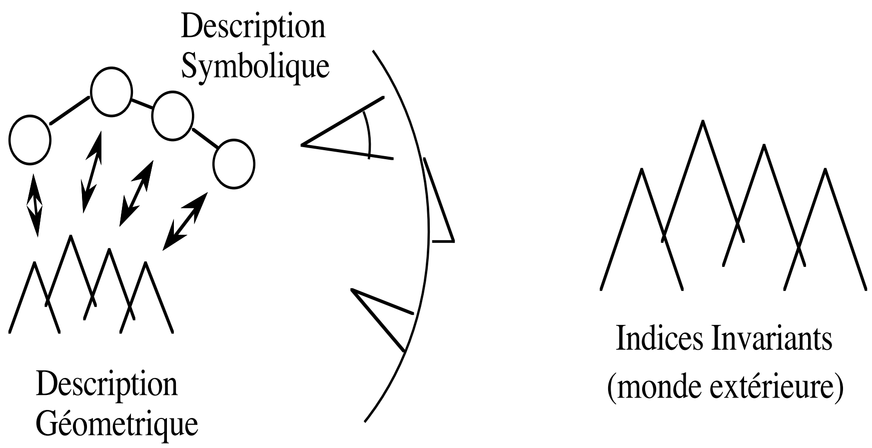


Two clear axes emerge:
One axis from the origin to the RGB of the product of the illumination and the source.
The other axis towards the RGB representing the illumination.

2 Describing Contrast

An image is simply a large table of numerical values (pixels).
The "information" in the image may be found in the colors of regions of pixels, and the variations in intensity of pixels (contrast).

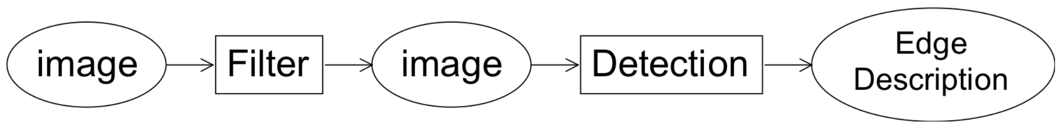
Extracting information from an image requires organizing these values into patterns that are "invariant" to changes in illumination and viewing direction.



Color provides information about regions of constant pigment.
Contrast provides information about 3D shape, as well as surface markings.

Contours of high contrast are referred to as "edges".

Edge detection is typically organized in two steps
1) contrast filtering
2) edge point detection and linking.



Two classic contrast detection operators are:

1) Roberts Cross Operator, and
2) The Sobel edge detector.

2.1 **Roberts Cross Edge Detector**

One of the earliest methods fore detecting image contrast (edges) was proposed by Larry Roberts in his 1962 Stanford Thesis.

Note, in this same thesis, Roberts introduced the use of homogeneous coordinates for camera models, as well as wire frame scene models. Roberts subsequently went to work for DARPA where he managed the program that created the Arpanet (now known as the internet).

Roberts Cross operator employs two simple image filters:

$$m_1(i,j) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad m_2(i,j) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

These two operators are used as filters. They are convolved with the image.

Convolution (or filtering) : for n = 1, 2

$$E_n(i, j) = m_n * p(i,j) = \sum_{k=0}^{1}\sum_{l=0}^{1} m_n(k,l)p(i-k, j-l)$$

The contrast is the module of each pixel :

$$E(i,j) = \left\|\vec{E}(i,j)\right\| = \sqrt{E_1(i,j)^2 + E_2(i,j)^2}$$

The direction of maximum contrast is the phase

$$\varphi(i,j) = Tan^{-1}\left(\frac{E_2(i,j)}{E_1(i,j)}\right) + \frac{\pi}{4}$$

Because of its small size and simplicity, the Roberts detector is VERY sensitive to high spatial-frequency noise. This is exactly the noise that is most present in images.

To reduce such noise, it is necessary to "smooth" the image with a low pass filter.
We can better understand the Roberts operators by looking at their Fourier Transform.

$$M_n(u,v) = \sum_{k=0}^{1}\sum_{l=0}^{1} m_n(k,l)e^{-j(ku+lv)}$$

$$M_1(u,v) = (+1)\cdot e^{-j(-(0)u+-(0)v)} + (-1)\cdot e^{-j(u+v)} = 2j\ Sin(0.5u+0.5v)$$

$$M_2(u,v) = (+1)\cdot e^{-j(-(0)u+-(1)v)} + (-1)\cdot e^{-j(-(1)u+-(0)v)} = 2j\ Sin(0.5u-0.5v)$$

2.2 The Sobel Detector

Invented by Irwin Sobel in his 1964 Doctoral thesis, this edge detector was made famous by the the text book of R. Duda adn P. Hart published in 1972.

It is perhaps the most famous and widely used edge detector:

$$m_1(i,j) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \qquad m_2(i,j) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

Convolution (or filtering) : for n = 1, 2

$$E_n(i, j) = m_n * p(i,j) = \sum_{k=-1}^{1}\sum_{l=-1}^{1} m_n(k,l)\,p(i-k, j-l)$$

The contrast is the module of each pixel :

$$E(i,j) = \left\| \vec{E}(i,j) \right\| = \sqrt{E_1(i,j)^2 + E_2(i,j)^2}$$

The direction of maximum contrast is the phase

$$\varphi(i,j) = Tan^{-1}\left(\frac{E_2(i,j)}{E_1(i,j)} \right)$$

Sobel's Edge Filters can be seen as a composition of a image derivative and a smoothing filter.

$$m_1(i,j) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$m_2(i,j) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

The filter $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$ is a form of image derivative.

The filter $\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$ is a binomial smoothing filter.

2.3 Difference Operators: Derivatives for Sampled Signals

For the function, s(x) the derivative can be defined as :

$$\frac{\partial s(x)}{\partial x} = \lim_{\Delta x \to 0} \left\{ \frac{s(x+\Delta x) - s(x)}{\Delta x} \right\}$$

For a sampled signal, s(n), an the equivalent is $\frac{\Delta s(n)}{\Delta n}$

the limit does not exist, however we can observe

$$\Delta n = 1: \quad \frac{\Delta s(n)}{\Delta n} = \frac{s(n+1) - s(n)}{1} = s(n) * \begin{bmatrix} -1 & 1 \end{bmatrix}$$

$$\Delta n = 0: \quad \frac{\Delta s(n)}{\Delta n} = \frac{0}{0}$$

This is the operator used by Roberts.

If we use a Symmetric definition for the derivative:

$$\frac{\partial s(x)}{\partial x} = \lim_{\Delta x \to 0} \left\{ \frac{s(x+\Delta x) - s(x-\Delta x)}{\Delta x} \right\}$$

then

$$\Delta n = 1: \quad \frac{\Delta s(n)}{\Delta n} = \frac{s(n+1) - s(n-1)}{1} = s(n) * \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

This is the operator used by Sobel.

Note that a derivative is equivalent to convolution!
We can define derivation in the fourier domain as follows:

$$F\left\{ \frac{\partial s(x)}{\partial x} \right\} = -j\omega \cdot F\{s(x)\}$$

and thus

$$\frac{\partial s(x)}{\partial x} = F^{-1}\{-j\omega\} * s(x)$$

If we can determine d(x) = $F^{-1}\{-j\omega\}$ then we have our derivative operator. If we "sample" d(x) to produce d(n) we have a sampled derivative operator.

Unfortunately, $F^{-1}\{-j\omega\}$ has an infinite duration in x, and thus d(n) is an infinite series. However, the first term of d(n) is [-1 0 1].

Thus we can define the first "difference" operator as a first order approximation for the derivative of a discrete signal.

$$\Delta_i p(i,j) = \Delta p(i,j)/\Delta i = p(i,j) * [-1, 0, 1]$$

$$\Delta_i p(i,j) = \frac{\Delta p(i,j)}{\Delta i} = p(i,j) * \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

$$\Delta_j p(i,j) = \frac{\Delta p(i,j)}{\Delta j} = p(i,j) * \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

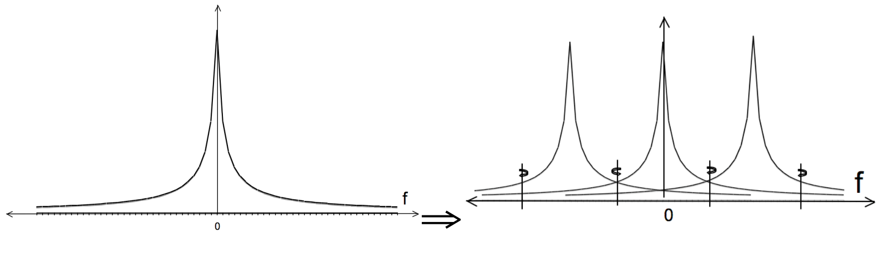
$$\nabla P(i,j) = \begin{pmatrix} \Delta_i p(i,j) \\ \Delta_j p(i,j) \end{pmatrix}$$

$$E(i,j) = \left\| \nabla P(i,j) \right\|$$

$$\vartheta(i,j) = Tan^{-1}\left(\frac{\Delta_j p(i,j)}{\Delta_i p(i,j)} \right)$$

This works fine, except that such a derivative operator amplifies sampling noise.

When a signal s(x) is sampled to create s(n), sampling noise is introduced



Sampling adds repeated copies of the spectrum at periods of two times the nyquist frequency $2F_n = 2/T$. The result amplifies high frequency noise.

The first difference filter $d_1(n)$ = [1, 0, -1] has a Fourier transform:

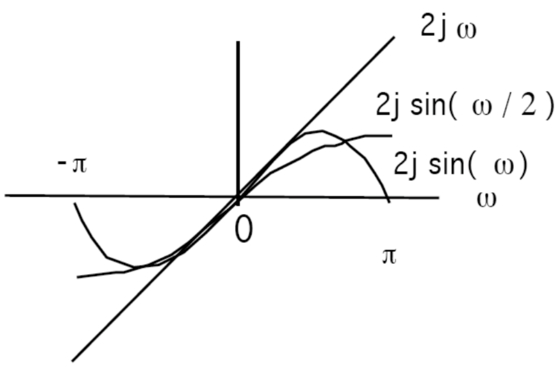
$$D(\omega) = \sum_{n=-1}^{1} d(n) e^{-j\omega n}$$
$$D(\omega) = 1e^{-j\omega(-1)} + 0e^{-j\omega 0} + (-1)e^{-j\omega(1)}$$
$$D(\omega) = e^{j\omega} - e^{-j\omega}$$
$$D(\omega) = -2j\sin(\omega)$$

Calculation of a derivative is the same as convolution with the filter [1, 0, -1], which is the same as multiplication of the spectrums.

$$d(n) * s(n) \quad \Leftrightarrow D(\omega) \cdot S(\omega)$$

The filter d(n) = [1, -1] is even worse. Its Fourier transform is

$$D(\omega) = -2j\sin(\omega/2)$$



Sobel uses the optimal local derivative filter.