Computer vision

23-09-07

Physics of color

The white light is composed of almost equal energy in all wavelengths of the visible spectrum. Actually, *the white is a combination of red, blue and green*. When the light passes through the triangular prism, you see a rainbow. This is because the prism separates the colors by a different angle (by varying refractive index).

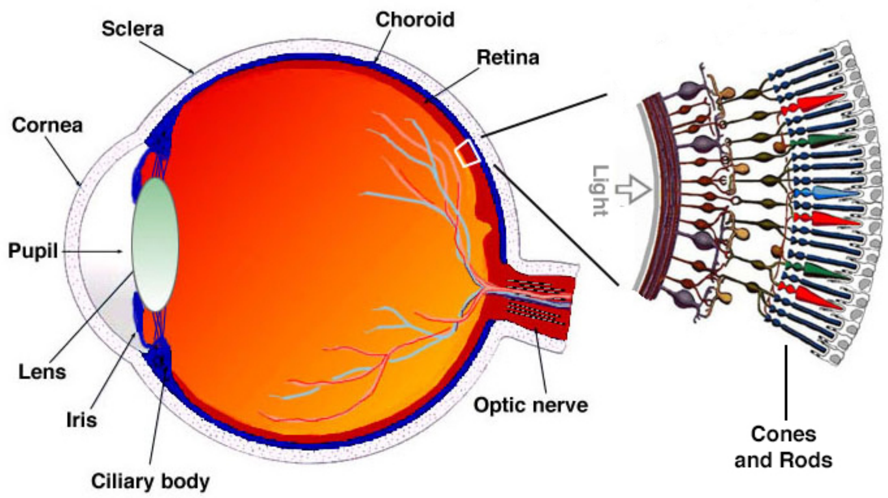
Prism experiments

In simple words, **color** is the result of the interaction between physical light and our visual system. Sientifically (and broader), reflected color is the result of interaction of light source spectrum with surface reflectance. Spectral radiometry is a field of science that deals with the measurement and analysis of light or electromagnetic radiation.

Amaizing math behind colors

Human eye

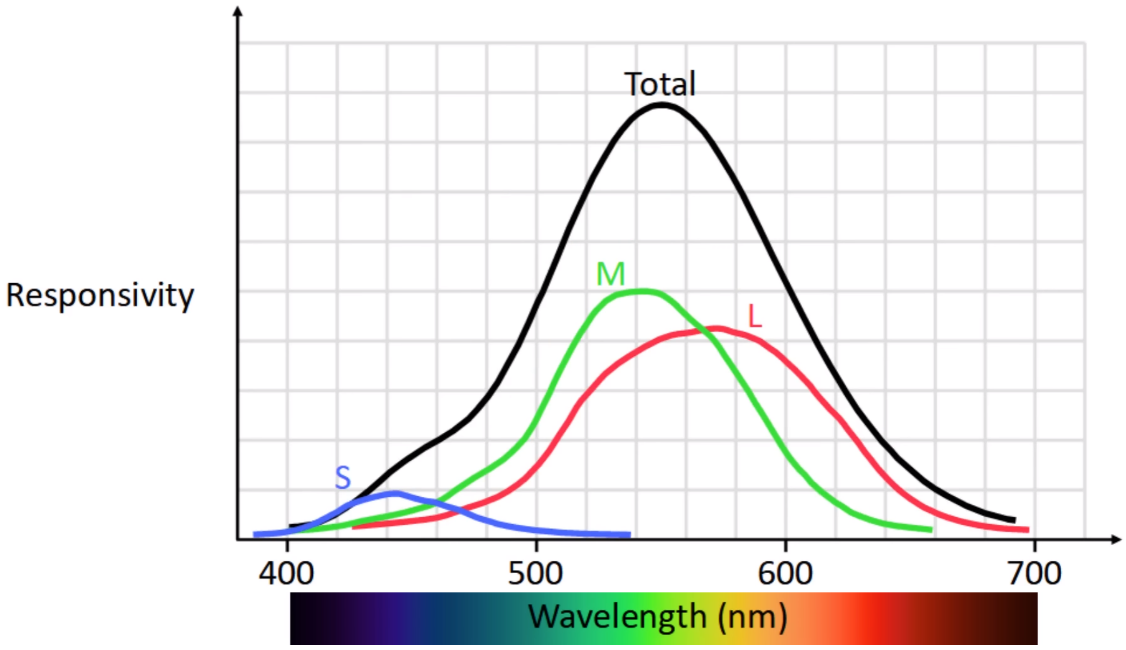
Cones and rods are the two types of light-sensitive receptors (dwell in the retina).



Rods are rod-shaped and highly sensitive to light. They are responsible for the gray-scale (black & white) vision. Rods are bad at detecting colors. They are more suited for detecting motion and shapes under low-light conditions.

Cones are cone-shaped, responsible for color vision and provide high-resolution, detailed images. Cones are less sensitive to light, meaning they require more light to become activated.

There are three types of cones: short-wavelength cones (blue), medium-wavelength cones (green), and long-wavelength cones (red). Each type detects the light of definite wavelength (the range can be found in the pic below). Depending on the wavelength, different amounts of cone types are activated. The brain combines the signals from the cones to create our perception of colors.



You can also notice that cones of different types have different responsivity. If we add three functions together, it gives us a function of a total responsivity. Its maximum is in the range of green and yellow wavelengths. It explains why we see colors around yellow as a lot brighter than others.

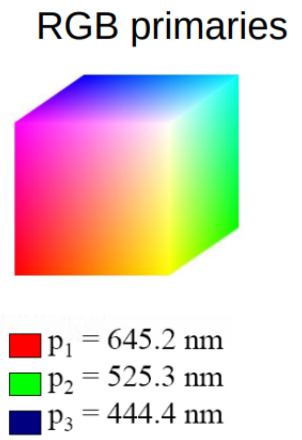
Color spaces

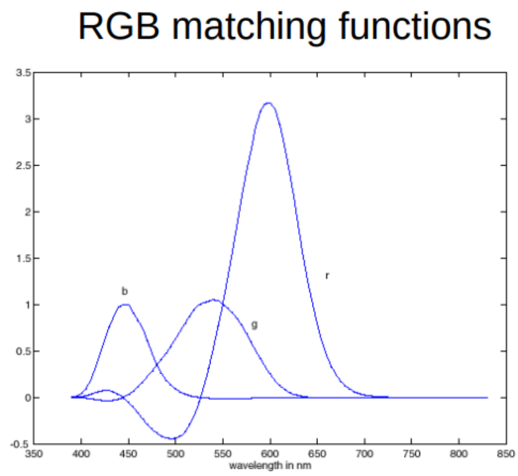
Linear color space

Linear color space is defined by three primaries. The coordinates of a color are given by the weights of the primaries used to match it.

One of the most well-known examples of a linear color space is CIELAB.

RGB space





RGB space describes how to express every color as a combination of red, green and blue. Primaries used in RGB space are monochromatic lights (consist of a single wavelength).

HSV space \todo

HSV color space is a nonlinear collor space. HSV stands for:

- Hue (0 to 360) - chromaticity angle
- Saturation (0 to 100) - chromaticity distance from the center of the [chromaticity diagram]
- Value (0 to 100) - controls luminance. Basically, it is the scale of the RGB cube

HSV space can be converted into RGB space:

- (R, G, B) is the color in the RGB space
- $r = \frac{R}{255},\ g = \frac{G}{255},\ b = \frac{B}{255}$
- $m_1 = \min(r, g, b),\ m_2 = \max(r, g, b)$
- $V = m_2 \cdot 100$
- $S = \frac{m_2 - m_1}{m_2} \cdot 100$

- \todo

White balancing

White balance is the process of removing unnatural color casts.

The von Kries adaptation hypothesis suggests that the human eye adapts to changes in illumination by adjusting its sensitivity to different wavelengths of light. Cones adapt to changes by multiplying their responses to incoming light by a scaling factor.

So, the idea of von Kries adaptation is to multiply each channel by a gain factor. One way to find the factor is to use the "gray card" method:

- take a picture of a neutral object (white or gray)
- deduce the weight of each channel. If the neutral object is recorded as r_w, g_w, b_w, then weights for custom white balance settings will be $\frac{1}{r_w}, \frac{1}{g_w}, \frac{1}{b_w}$.

23-09-14

Vector math

A **vector norm** is any function that satisfies 4 properties:

1. $\forall x \in R^n \ : \ f(x) \geqslant 0$
2. $f(x) = 0 \Rightarrow x = 0$
3. $\forall x \in R^n \ : \ f(tx) = |t| f(x)$
4. $\forall x, y \in R^n \ : \ f(x+y) \leqslant f(x) + f(y)$

$$||x|| = \sqrt{\sum_1^n x_i^2}$$

$$||x||_p = (\sum_1^n x_i^p)^{\frac{1}{p}}$$

$$||x||_\infty = \max_i |x_i|$$

Dot product (inner product)

$x^T y = \sum_1^n x_i y_i$

$x^T y = |x| \cdot |y| \cdot \cos\phi$ where ϕ is the angle between x and y

The result of $x \cdot y$ gives the length of the component of x which is parallel to y:

- positive $\Rightarrow \phi < 90^\circ$ and the component lies in the same direction as y.
- negative $\Rightarrow \phi > 90^\circ$ and the component lies in the opposite direction.
- zero $\Rightarrow x$ and y are perpendicular

Cross product

The result is a vector perpendicular to the plane created by x and y.

$x \times y = |x| \cdot |y| \cdot \sin\phi$

Matrix math

Multiplication

- associative: $(AB)C = A(BC)$
- distributive: $A(B + C) = AB + AC$
- non-commutative: $AB \neq BA$

Transpose

- $(ABC)^T = C^T B^T A^T$

Determinant

- $\det AB = \det BA$
- $\det A^{-1} = \frac{1}{\det A}$
- $\det A^T = \det A$
- $detA = 0 \Rightarrow A$ is singular (a square matrix that is not invertible)

Besically, a determinant of a matrix $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ is a square of a parallelogram formed by vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$.

Trace

$tr(A) =$ sum of elements on the diagonal.

- $\mathrm{tr}AB = \mathrm{tr}BA$

- $\mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$

Matrix norm

Frobenius norm (or Euclidean norm): $||A||_F = \sqrt{\sum_j \sum_i a_{ij}^2} = \sqrt{Tr(A^T A)}$

Rotation

R is a rotation matrix if it satisfies properties:

- $R \cdot R^T = E$
- $\det R = 1$

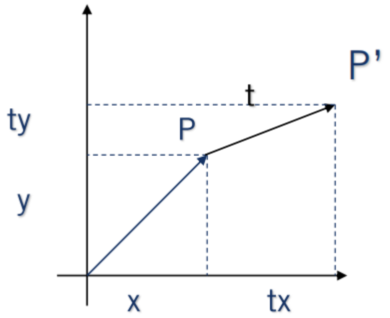
Rotating a system by ϕ is equivalent to rotating a point within that system by the angle $-\phi$. When you rotate a system by ϕ, you change the directions of the axes. This operation affects all points that are *within that system*. On the other hand, when you rotate a point within the same coordinate system by the angle $-\phi$, you are changing the position of that point while keeping the system's orientation unchanged.

Homogeneous coordinates

Translation

The matrix have the form of $T = \begin{pmatrix} 1 & 0 & \dots & t_1 \\ 0 & 1 & \dots & t_2 \\ 0 & 0 & \dots & t_3 \\ & \dots & & \\ 0 & 0 & \dots & 1 \end{pmatrix}$.

$\begin{pmatrix} t_1 \\ \dots \\ t_n \end{pmatrix}$ is an addition of a constants to the initial coordinates.



$$\mathbf{P} = (x, y) \rightarrow (x, y, 1)$$
$$\mathbf{t} = (t_x, t_y) \rightarrow (t_x, t_y, 1)$$

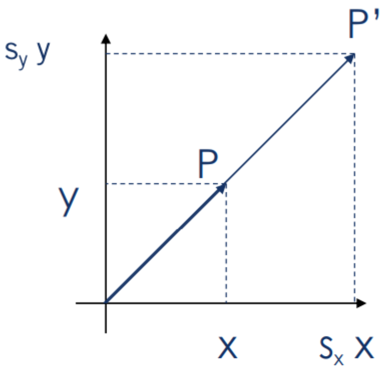
$$\mathbf{P}' \rightarrow \begin{bmatrix} x+t_x \\ y+t_y \\ 1 \end{bmatrix} = \begin{array}{c} \begin{array}{ccc} & & \mathrm{t} \end{array} \\ \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \end{array} \cdot \begin{array}{c} \mathrm{P} \\ \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \cdot \mathbf{P} = \mathbf{T} \cdot \mathbf{P}$$

Scaling

The matrix has a form of $\begin{pmatrix} s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & s_n & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$.

$s_1, \dots, s_n$ are scalars for the initial coordinates.



$$\mathbf{P} = (\mathrm{x}, \mathrm{y}) \rightarrow \mathbf{P}' = (\mathrm{s_x x}, \mathrm{s_y y})$$

$$\mathbf{P} = (x, y) \rightarrow (x, y, 1)$$
$$\mathbf{P}' = (s_x x, s_y y) \rightarrow (s_x x, s_y y, 1)$$

$$\mathbf{P}' \rightarrow \begin{bmatrix} s_x x \\ s_y y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{S}} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{S}' & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \cdot \mathbf{P} = \mathbf{S} \cdot \mathbf{P}$$

The order of scaling and translation is important. `Scalling * Translating` ends up in a bigger image, then `Translating * Scalling`.

Eigenvectors and eigenvalues

Given a matrix $\mathcal{A}$ of size $n \times n$.

$v \in V/\{0\}$ is an *eigenvector* if $\exists \lambda \ : \ \mathcal{A}(v) = \lambda v$.

λ is an *eigenvalue*.

$p_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda E)$ is a *characteristic polynomial* of matrix $\mathcal{A}$.

λ is an eigenvalue $\Leftrightarrow$ λ is the root of the characteristic polynomial, which means $\det(\mathcal{A} - \lambda E) = p_{\mathcal{A}}(\lambda) = 0$.

$tr(A) = \sum\limits_i \lambda_i$ (λ_i are eigenvalues)

Example

$\mathcal{A} = \begin{pmatrix} 3 & -1 \\ -3 & 5 \end{pmatrix}$.

$p_{\mathcal{A}}(\lambda) = \det \begin{pmatrix} 3-\lambda & -1 \\ -3 & 5-\lambda \end{pmatrix} = \lambda^2 - 8\lambda + 12 = 0.$

$\lambda_1 = 6, \ \lambda_2 = 2.$

$\mathcal{A}v_1 = \lambda_1 v_1 \ \Rightarrow \ v_1 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$. Normalized $\frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ -3 \end{pmatrix}$.

$\mathcal{A}v_2 = \lambda_2 v_2 \ \Rightarrow \ v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Normalized $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Singular Value Deomposition (SVD)

SVD decomposes a matrix into three other matrices, allowing to represent the original matrix in a more interpretable and compact form.

$$A = U \Sigma V^T$$

where:

- A is the original matrix of size $m \times n$
- U is an orthogonal matrix (meaning $U^T U = E$). Its size is $m \times m$.

Columns of U are called the left singular vectors of A.

- Σ is a diagonal matrix with non-negative real numbers, called the singular values of A. They are typically arranged in descending order. $rank(\Sigma) = rank(A)$.
- V is an orthogonal matrix of size $n \times n$.

  Rows of V^T are called the right singular vectors of A.

U and V are always rotation matrices. They represent rotations and reflections in the original space that transform the columns of A into the coordinate axes. These transformations preserve the length of vectors.

Σ works like a scalling matrix. The values on the diagonal are the singular values $\sigma_1, \ldots, \sigma_m$ of the matrix A (singular value is a root of eigenvalue for $A^T A$). The values are typically ordered in descending order, meaning the largest singular value is at the top-left corner. Larger singular valueiis correspond to more significant contributions, meaning the karger coefficient of the corresponding component.

Example

Given the $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, find its SVD.

- Find eigenvalues and eigenvectors of $A^T A$:

  $$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  $$\lambda_1 = 2,\ \lambda_2 = 1,\ \lambda_3 = 0$$

  $$w_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\ w_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\ w_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

- Values of Σ are roots of eigenvalues of $A^T A$:

  $$\sigma_1 = \sqrt{2},\ \sigma_2 = 1;\ \ \Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- Columns of V are normalized eigenvectors of $A^T A$:

  $$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix},\ v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\ v_3 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

  $$V = [v_1 \quad v_2 \quad v_3]$$

- $u_1 = \dfrac{1}{\sigma_1} A v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \ u_2 = \dfrac{1}{\sigma_2} A v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

  $$U = [u_1 \quad u_2]$$

Gradient \todo