Self-Driving Car Research Studio

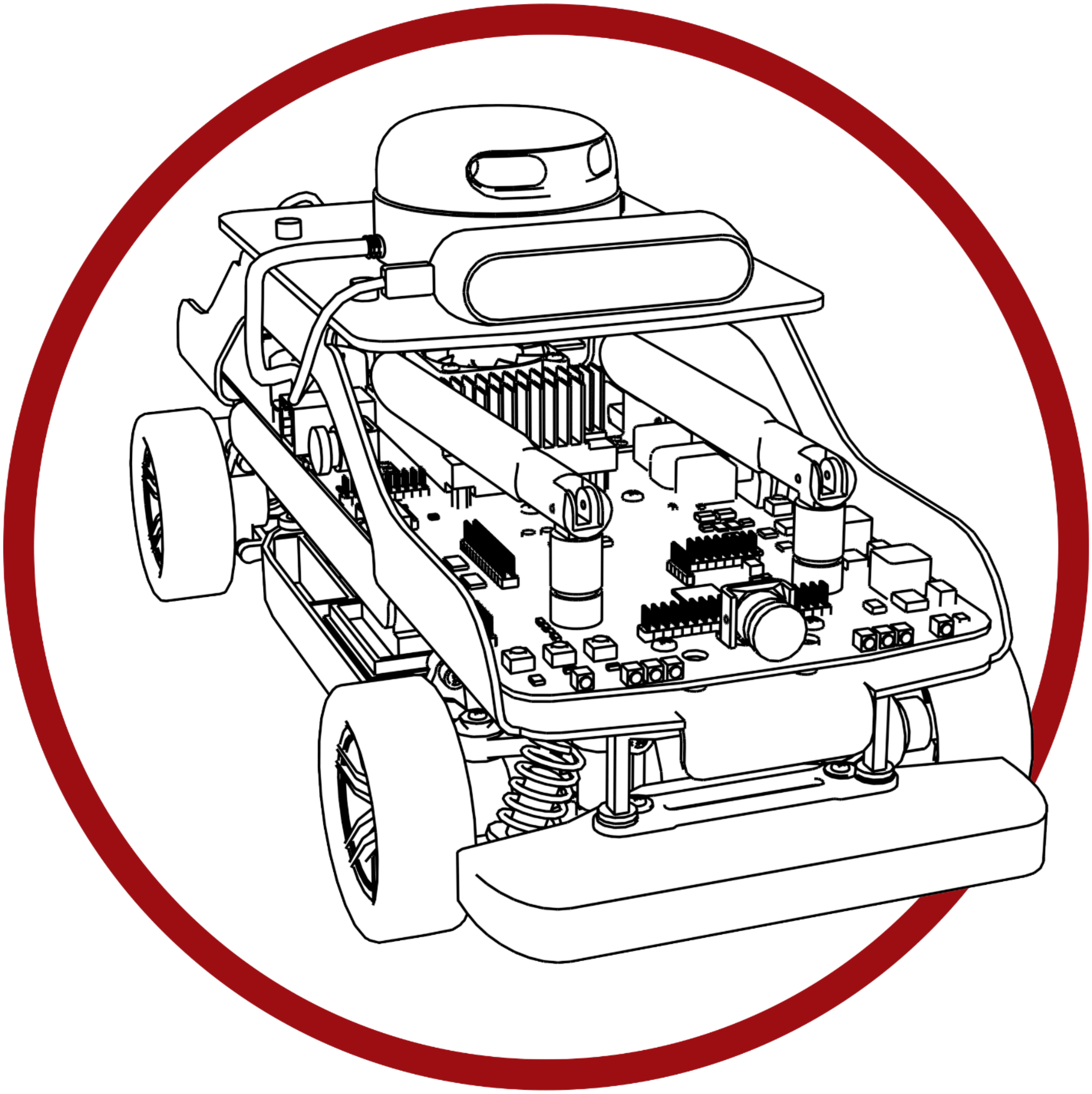

Image Color Spaces

V 1.1 (November 2020)

Table of Contents

Images as Matrices

Images captured by cameras can be represented as mathematical constructs in the form of matrices. The spatial spread of pixels in the image plane is captured by 2 dimensions representing the **rows** and **columns** as shown in Figure 1.

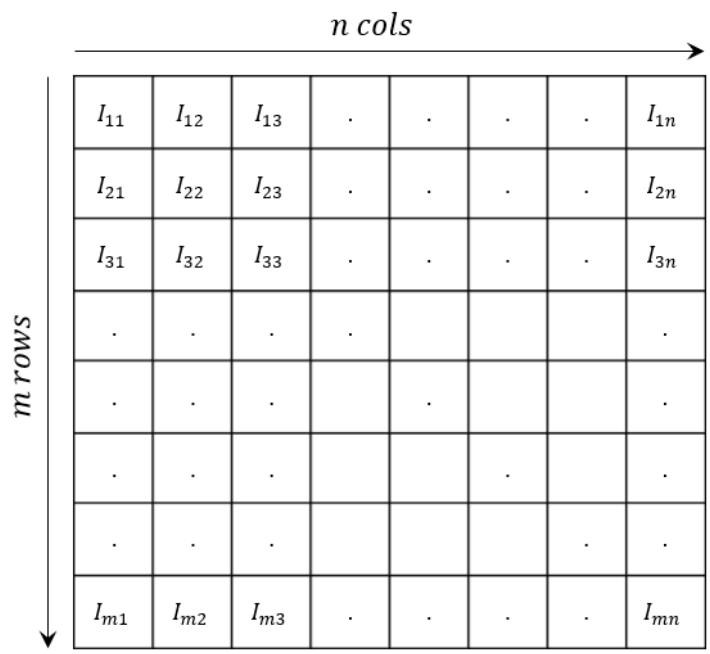

Figure 1. Spread of pixels on the image plane

Information such as color, intensity etc. of light at a location in the matrix specified by the (row, col) index tuple, is specified by using third dimensions, or **channels**. For example, consider a grayscale image with a resolution of m x n. Any given pixel would have an integer intensity $I_{i,j}$ where $i \in 1, m$ and $j \in 1, n$, where the intensity ranges from 0 (specifying black) up to 255 (specifying white) - an 8-bit unsigned integer value range. Values within the range specify varying intensities of gray. This would correspond to a single channel, or overall, an **m rows x n cols x 1 channel** image.

Representing colors requires more channels, as described in sections below.

Color Spaces

RGB

A color space based on the additive RGB color model uses the red, green and blue color primaries to describe the spectrum of colors in an image. In this case, we continue to use the first 2 dimensions to represent the spatial spread of pixels, but use a third dimension with 3 channels to allow the specification of the red, green and blue intensities. This would correspond to a 3 channel, or over all, an **m rows x n cols x 3 channels** image, as shown in Figure 2.

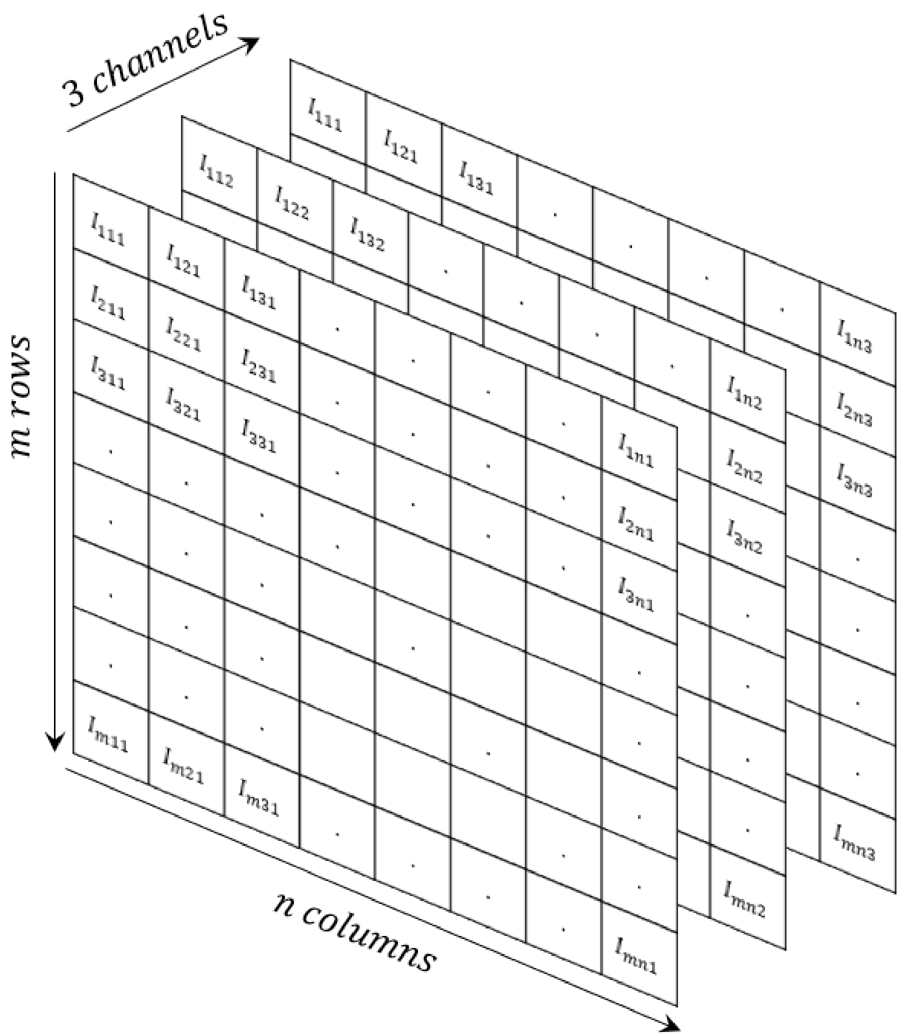

Figure 2. RGB images as 3D matrices

The red, green and blue intensities also vary between 0 and 255 each, with values within that range specify varying intensities of that particular primary color. Together, they create various colors as shown in Figure 3.

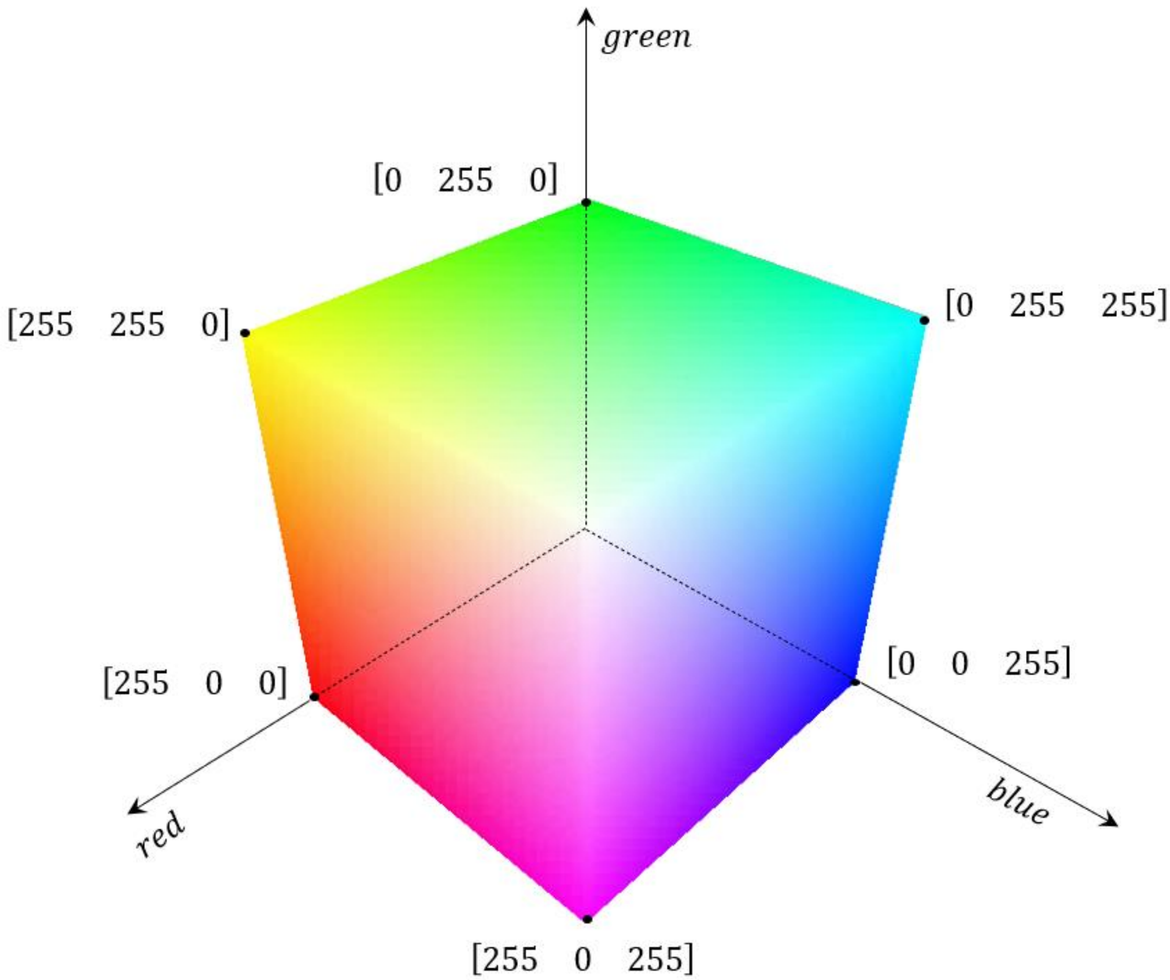

Figure 3. RGB primary colors

A bright pure red would correspond to an intensity of 255 in the red channel/plane, and 0 in the green and blue channels, and vice versa for green and blue. Mixing intensities in different channels creates additive colors. For example, white corresponds to a 255 in all three channels and black all 0s. Red and green intensities at 255 and blue at 0 would correspond to a yellow.

The RGB color space, although intuitive, does not describe observed colors under different viewing conditions well. It combines luma (or light intensity) with the chroma (or color intensity). A red stop sign in daylight may correspond to a RGB intensity of [255, 0, 0] but at night, may correspond to [50, 0, 0], even though the color is still red.

HSV

The HSV (hue, saturation and value) space distinguishes the luma (light intensity information) from the chroma (color information). The coordinate transformation between the RGB and HSV color spaces is portrayed in Figure 4.

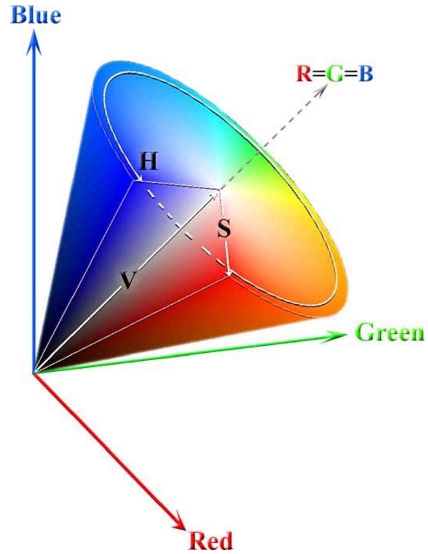

Figure 4. RGB to HSV coordinate transformation

(Pekel, J.-F & Ceccato, Pietro & Vancutsem, Christelle & Cressman, Keith & Vanbogaert, Eric & Defourny, Pierre. (2011). Development and Application of Multi-Temporal Colorimetric Transformation to Monitor Vegetation in the Desert Locust Habitat. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of. 4. 318 - 326. 10.1109/JSTARS.2010.2052591.)

Taking a cross-section of the color cone shown in Figure 4 portrays the Hue representation shown in Figure 5 as a color wheel, which ranges from red at 0˚, to green at 120˚, blue at 240˚ and back to red at 360˚. The **hue** parameter of an object remains stationary regardless of the light intensities (represented using **saturation** and **value** parameters).

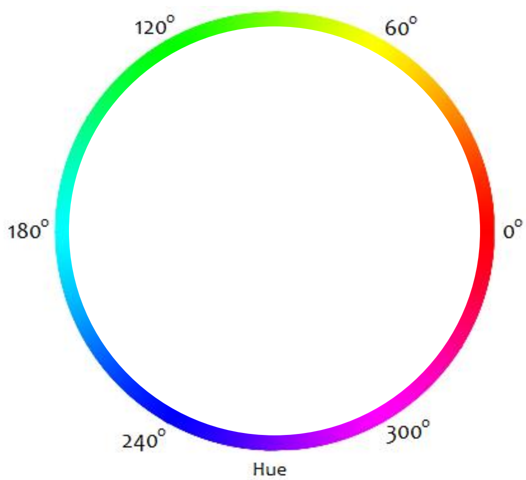

Figure 5. Hue color wheel

The saturation represents the amount of gray in a pixel. For example, consider an RGB color [100, 100, 100]. This represents a medium to dark gray, with equal red, green and blue intensities. This would correspond to the lowest saturation of 0. As the red, green and blue intensities vary, the saturation becomes larger, up to 255 (or 100% in some cases).

Lastly, the **value** represents the overall lightness of the pixel. Thus, a stop sign in daylight would correspond to [255, 0, 0] in the RGB space or [0, 255, 255] in the HSV space. This refers to a hue of red at 0˚, maximum saturation (unique color) and value (light intensity). The same stop sign at night could refer to [10, 0, 0] in the RGB space or [0, 255, 10] in the HSV space. This refers to the same hue of red at 0˚, maximum saturation (unique color) and low value (light intensity).

Note: RGB intensities all fall in the range [0, 255], or [0, 1] in some representations. However, HSV intensities fall in the range [0, 360] for Hue, and [0, 255] for saturation/value. HSV intensities in the range [0, 1] for all three are also used.

RGB values can be converted to HSV values using the **rgb2hsv()** function in MATLAB and the **cvtColor()** method in **opencv** for Python.

Grayscale

Various formulas can be used to convert RGB images to a grayscale format. A weighted sum is a common approach,

$$I_{gray} = \frac{I_r + I_g + I_b}{3}$$

Not using equal weights, but rather those in the equation below represent the intensity perception of typical humans to primary colors, as we have most sensitivity to green and least to blue.

$$I_{gray} = 0.2126 I_r + 0.7152 I_g + 0.0722 I_b$$

For grayscale from HSV, use the **value** plane itself.

Bitonal

Bitonal or binary images have binary intensities being high (1 or 255) and low (0). These are the result of typical thresholding algorithms or are used as masks to highlight different regions of images.