

Introduction to Computer Vision

Alvin Lin

August 2018 - December 2018

Image Processing

Pixels are derived from the phrase “picture element” and are indexed as (x,y) or (col,row). Color images are 3D tensors and videos are n-dimensional. Arithmetic operations on image pixels light addition and subtraction translate to generic brightness adjustments and image blending, while multiplying and dividing corresponds to contrast adjustment though it is not very common in image processing.

Image Transforms

- Logarithmic transforms increase the dynamic range of the dark region but decreases the range of the bright regions.
- Exponential transforms are the inverse of the log transform.
- Power-law (gamma) transforms can either enhance contrast in high value regions at the expense of low value ones when $\gamma > 1$ or do the reverse when $\gamma < 1$.

Histograms

If we count all pixels with similar intensity values, we get the histogram image. This gives us a good visual representation of the image distribution.

Adaptive Thresholding

Adaptive thresholding was designed to overcome the limitations of global thresholding by using a different threshold value at each pixel. The threshold is determined by values in the neighborhood.

Color

Color is a psychological property of a visual experience when we look at objects and lights. It is the result of an interaction between the physical light in an environment with our physical system. Color is the property possessed by an object to produce different sensations on the eye as a result of the way the object reflects and emits light.

The Human Eye

The human eye is a camera with many parts:

- Lens: changes shape using ciliary muscles to focus on objects at different distances.
- Pupil: the hole (aperture) in your eye whose size is controlled by the iris.
- Iris: the colored annulus with radial muscles
- Retina: photoreceptor cells (cones and rods) in the back of your eye

The cones in your retina are responsible for color vision and are less sensitive. They operate in high light conditions, unlike the rods, which operate in low light conditions and are highly sensitive in order to provide gray-scale vision. Rods and cones are non-uniformly distributed on the retina.

Computing Color Matches

Computing color matches for any given color are useful because we want the colors in the world, on a monitor, and in print format to all look the same. We want to match the skin color of a person in a photograph printed on an ink jet printer to their true skin color.

Since it is hard to reproduce color exactly, it is important to know whether a color difference would be noticeable to a human viewer.

Linear Color Spaces

LAB color space expresses color as three numerical values:

- L stands for lightness
- A represents green-red color components
- B represents blue-yellow color components

Nonlinear Color Spaces

HSV has perceptually meaningful dimensions:

- H stands for hue
- S stands for saturation
- V stands for value (intensity)

Color Based Image Retrieval

Given a collection of images, we can extract and store a color histogram per image. We can then compare these to a histogram for a new query image and compute the intersection to determine the most similar image. Color can also be used for simpler and faster model tracking.

You can find all my notes at <http://omgimanerd.tech/notes>. If you have any questions, comments, or concerns, please contact me at alvin@omgimanerd.tech