Introduction to Computer Vision

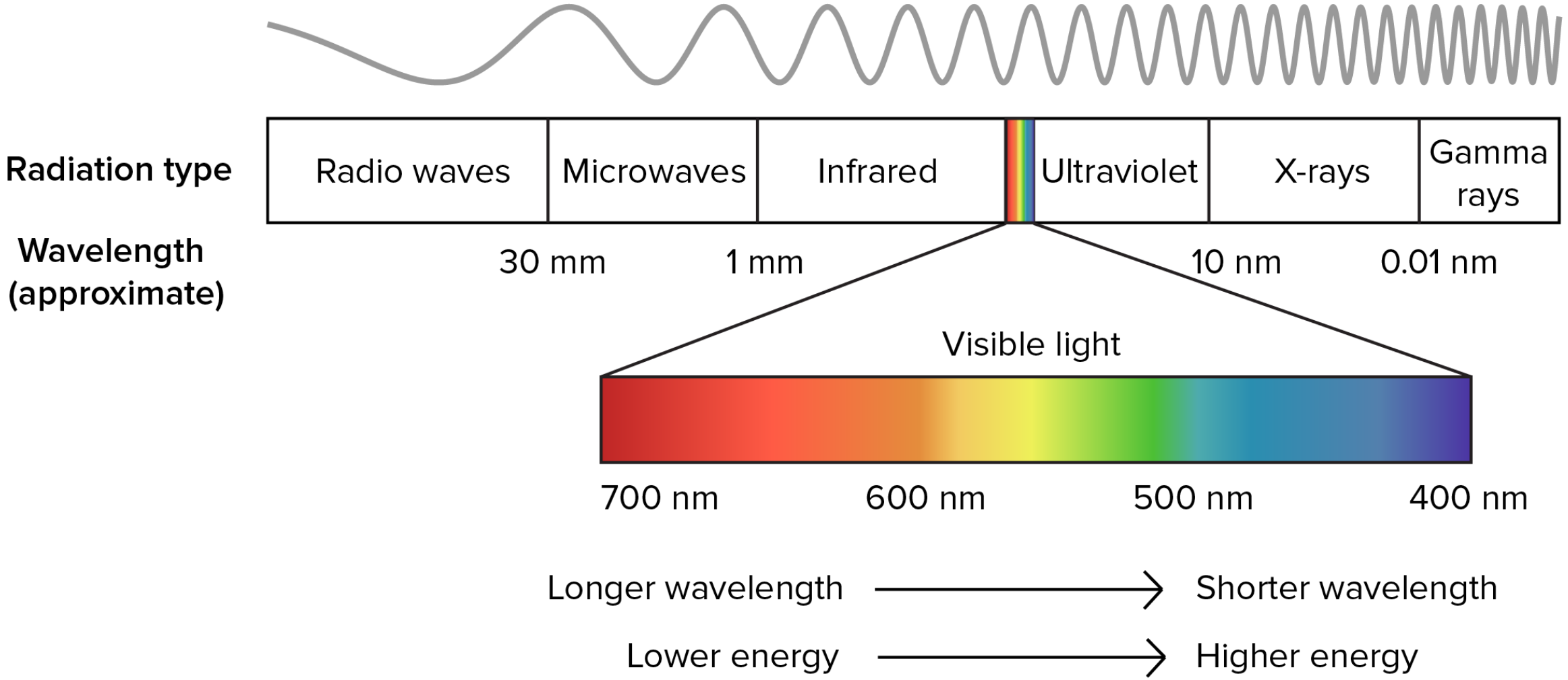
Zamaliev Eduard

eduard.zamaliev@intel.com

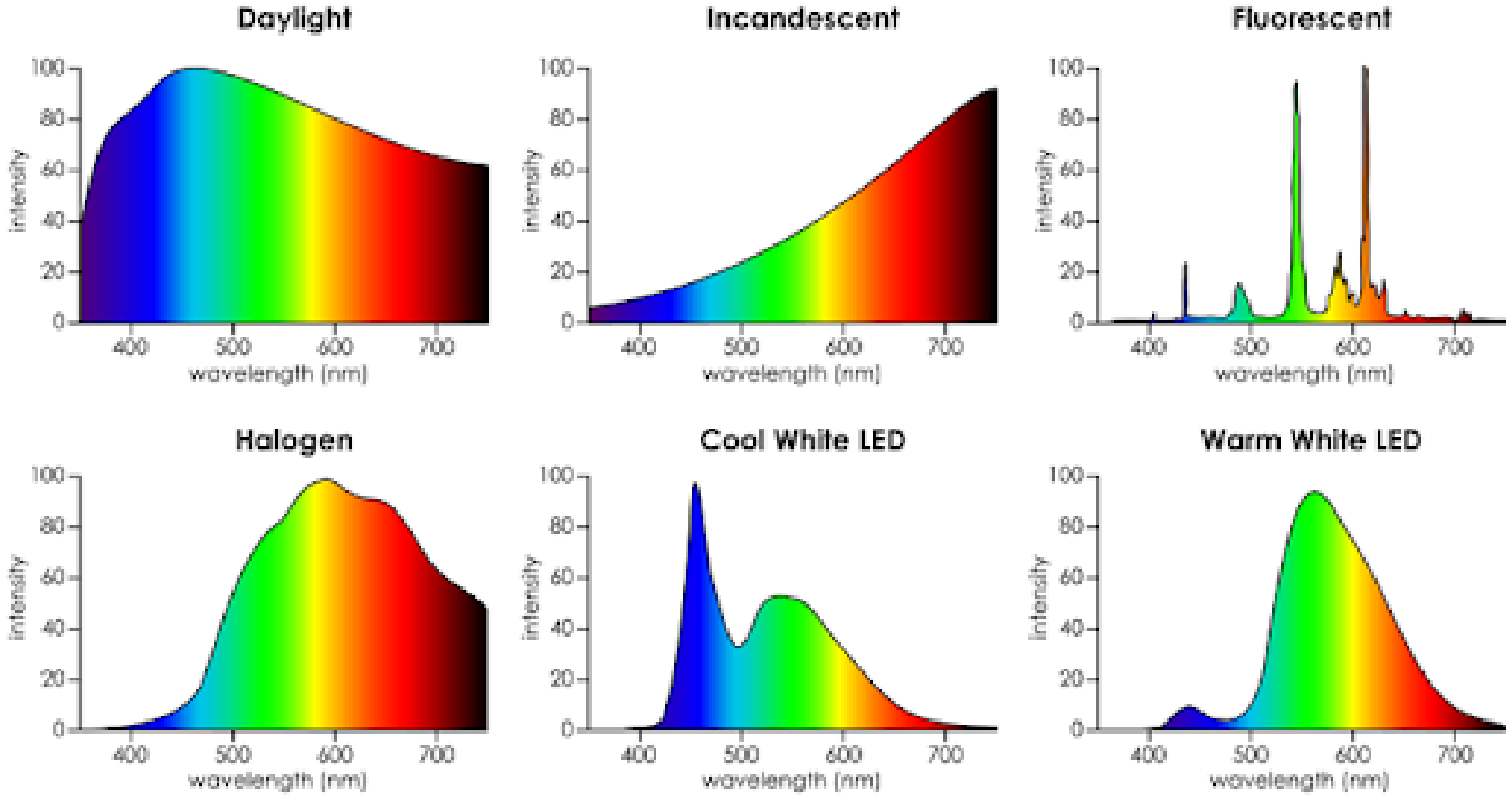
Plan

- Introduction to natural and digital vision
- Color spaces
- Image manipulations
  - Low level
  - Mid level
  - High level

Introduction: Visible light

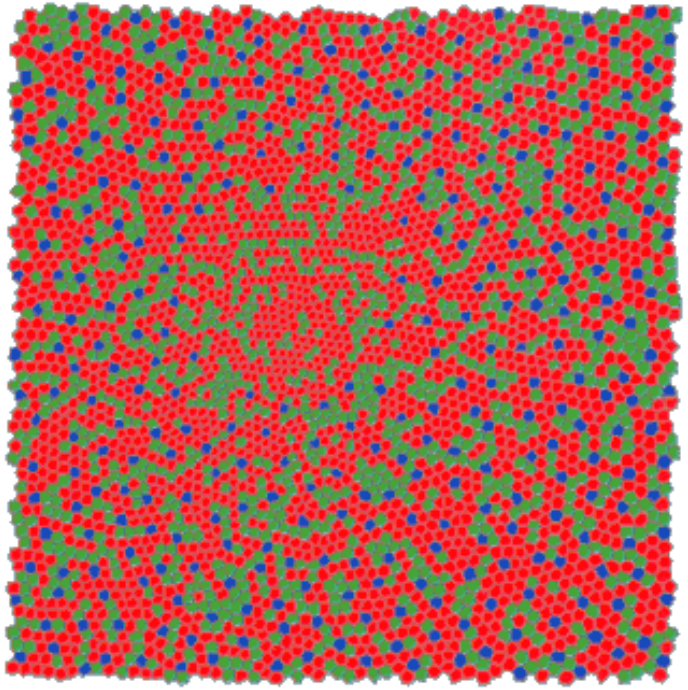


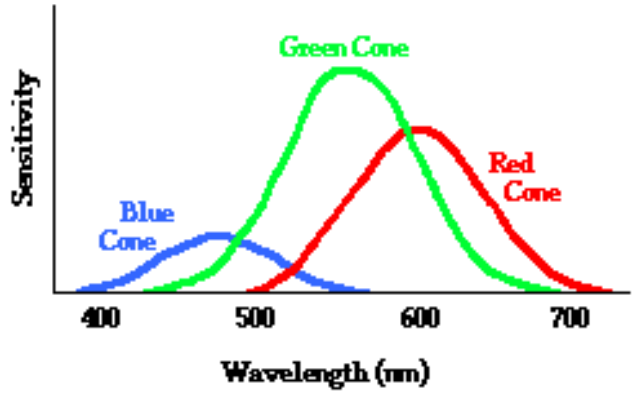
Introduction: Visible light



Introduction: Human eye

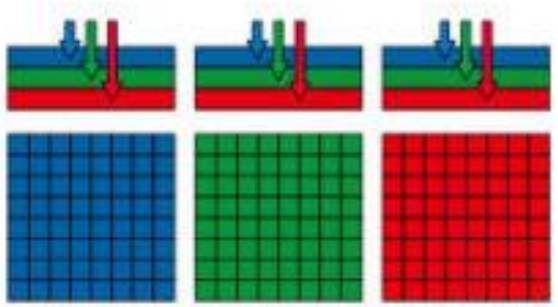
- Cone and rod cells are a base photoreceptor
- Receptor has responsiveness curve
- Rod cell peak ~490nm
- 3 types of cone cells:
  - Short: peak ~440 nm
  - Medium: peak ~ 540 nm
  - Long: peak ~570 nm



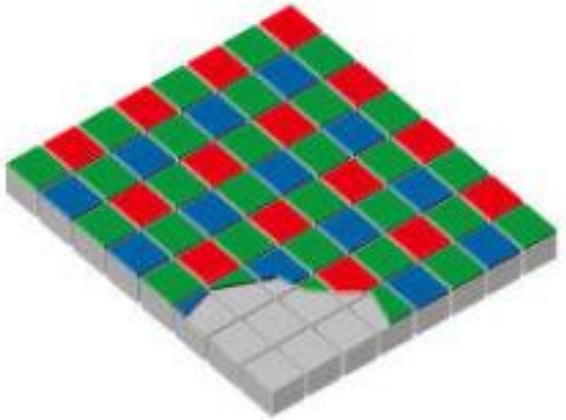


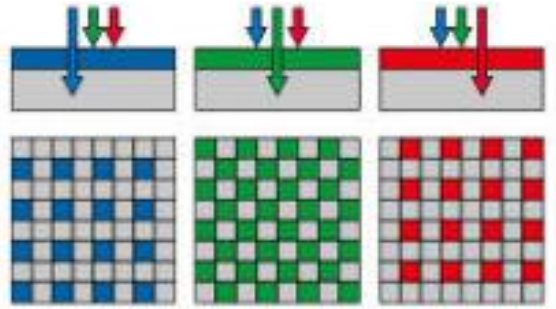
Introduction: Digital vision

Foveon X3
direct image sensor



Color filter array sensor
(Bayer filter sensor)





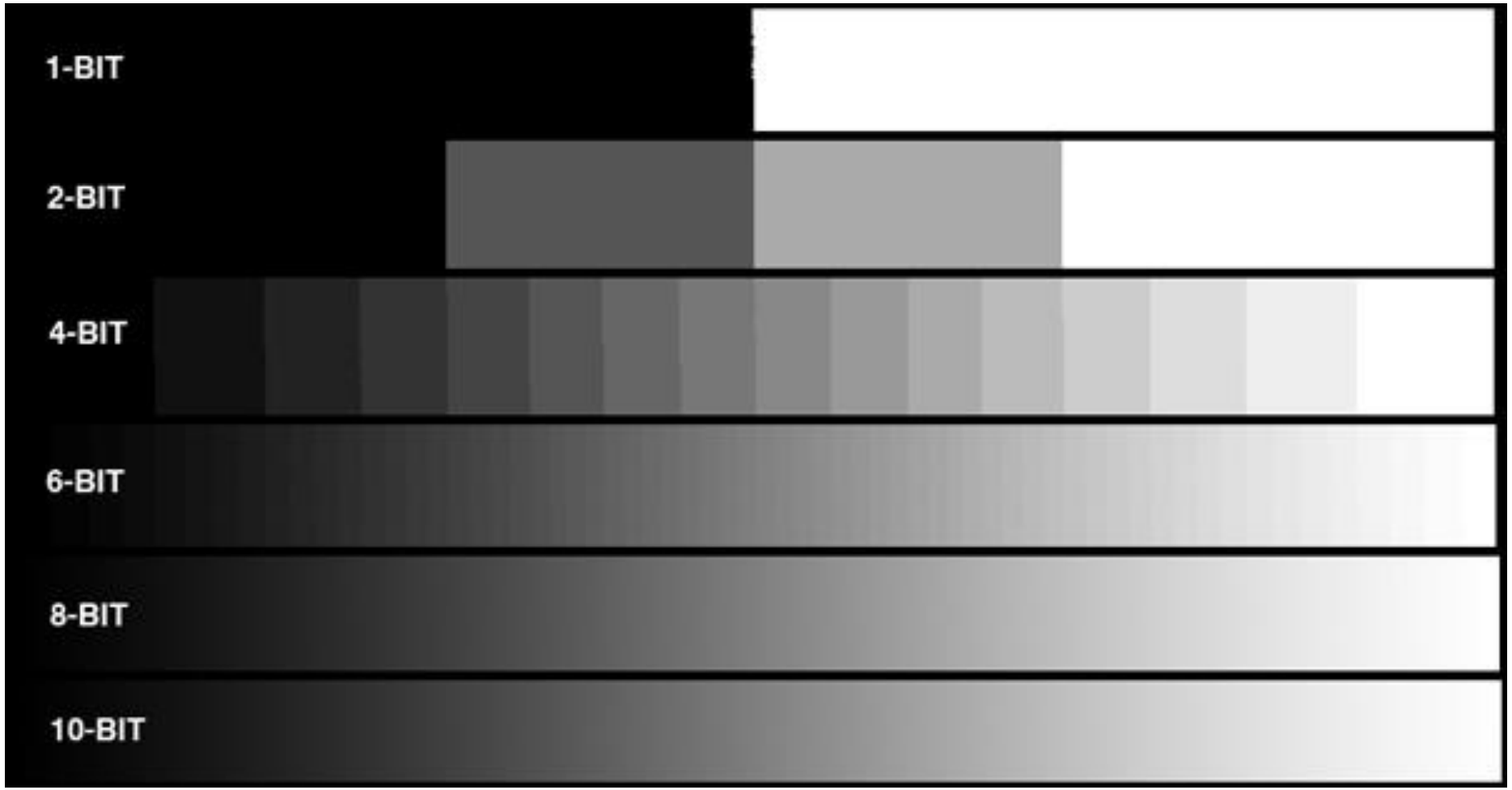
Introduction: CV tasks

- Receive image
- Process image
- Visualize image

Color space

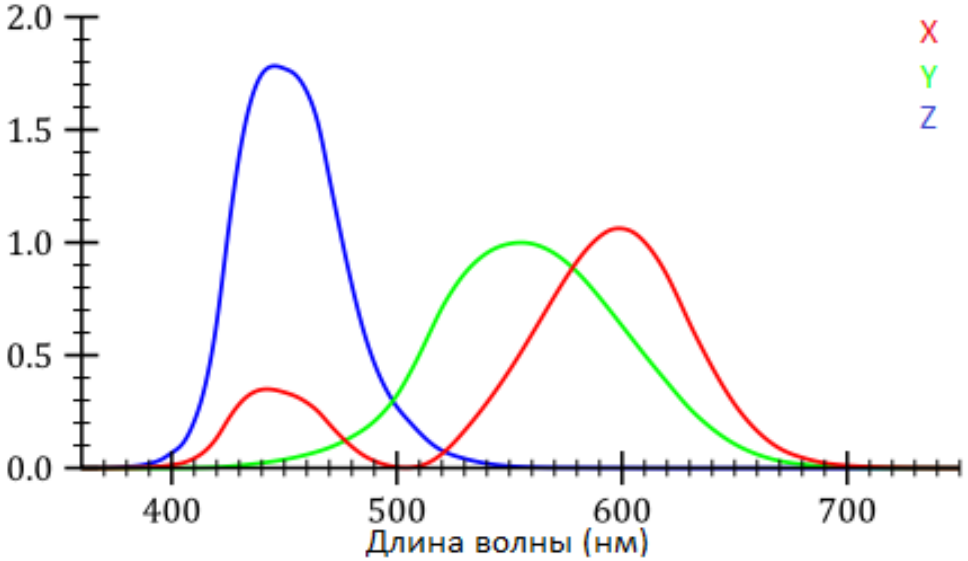
- Grayscale
- RGB, RGBA, BGR
- LUV, LAB, HSV
- YUV, YUV 420, YUV422, YUYV
- CMY, CMYK
- etc.

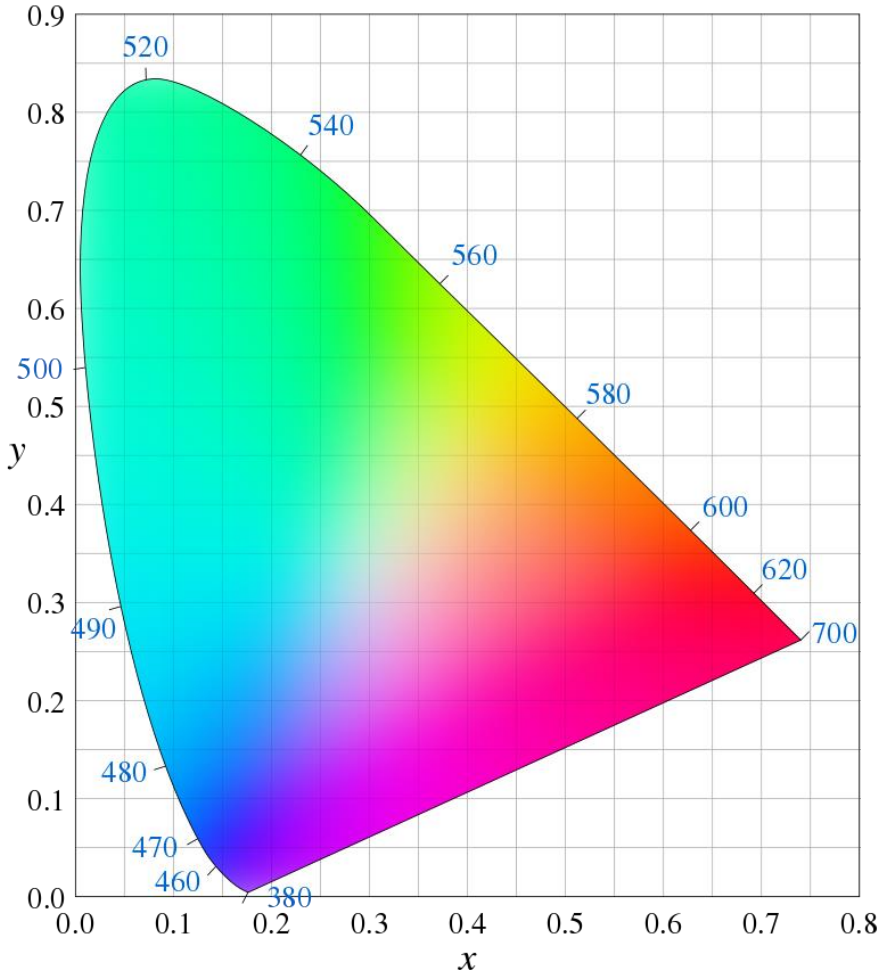
Color space: Grayscale



Color space: XYZ

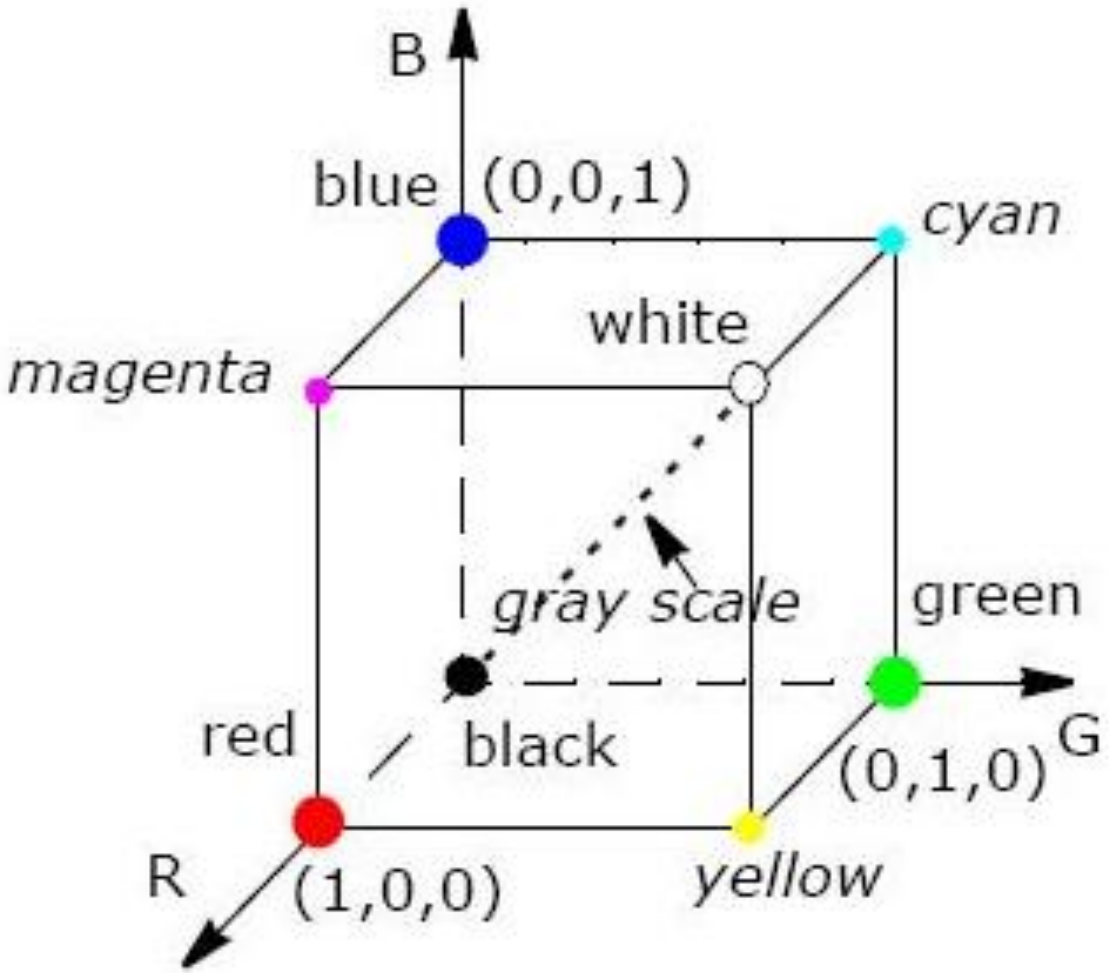
- Nonlinear color model
- Based on cone cells responsiveness curve





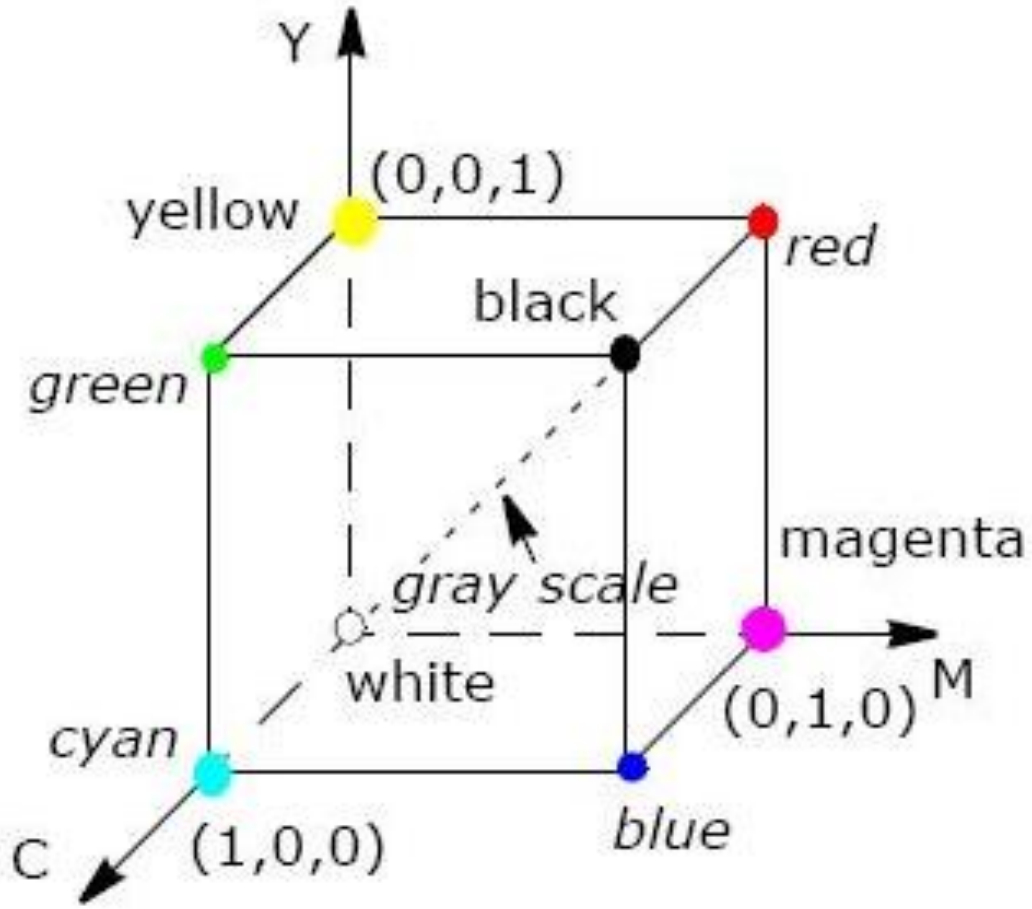
Color space: RGB

- Additive color model
- 3 channels: red, green and blue



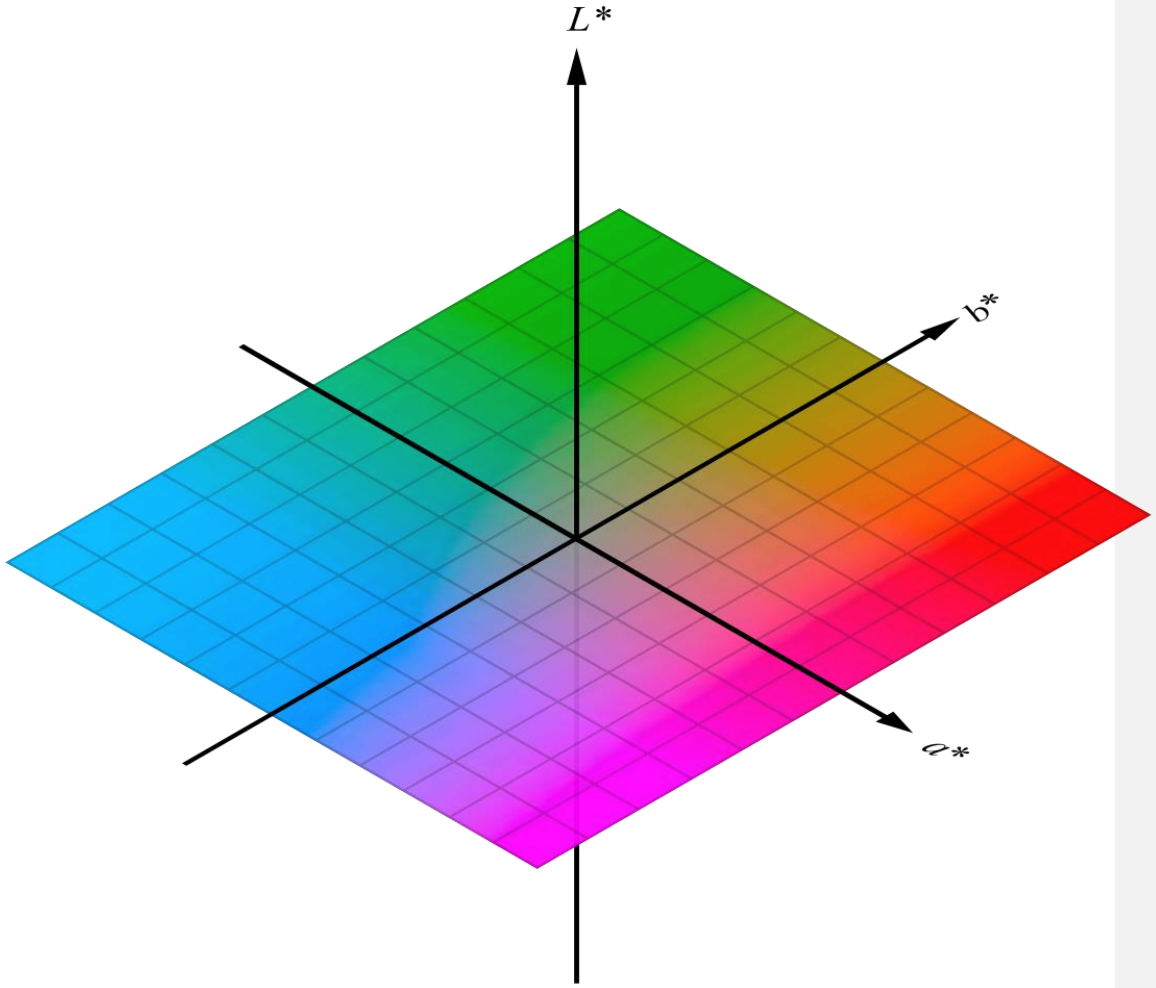
Color space: CMY and CMYK

- Subtractive color model
- Channels: cyan, magenta, yellow and black (for CMYK)
- Black stands for key color
- Typographical color space



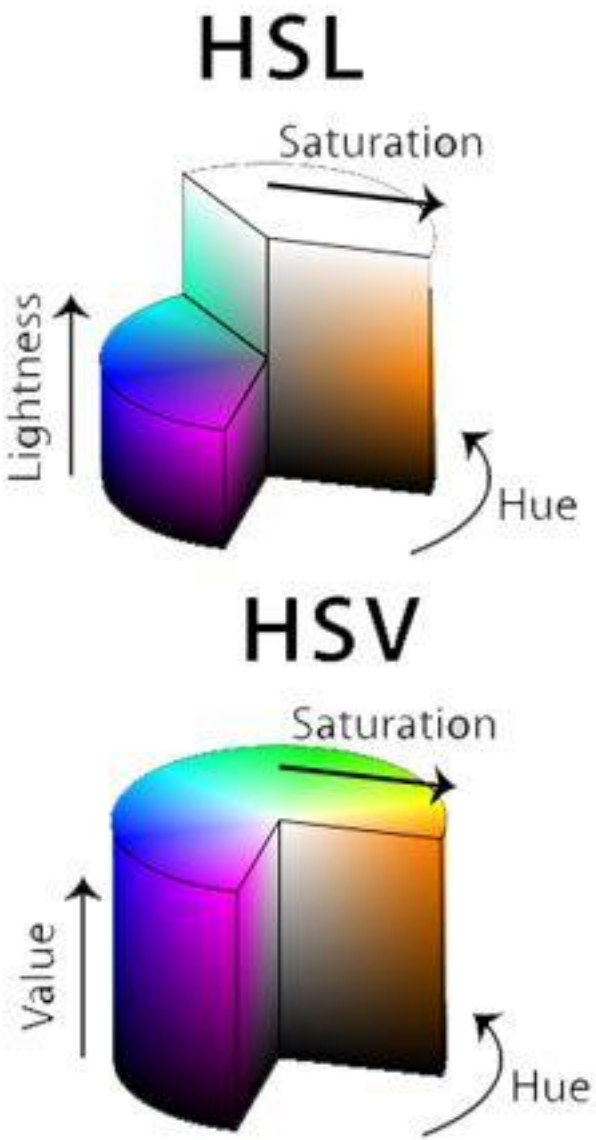
Color space: LAB

- L – lightness (from black to white)
- A – from green(-) to red(+)
- B – from blue(-) to yellow(+)



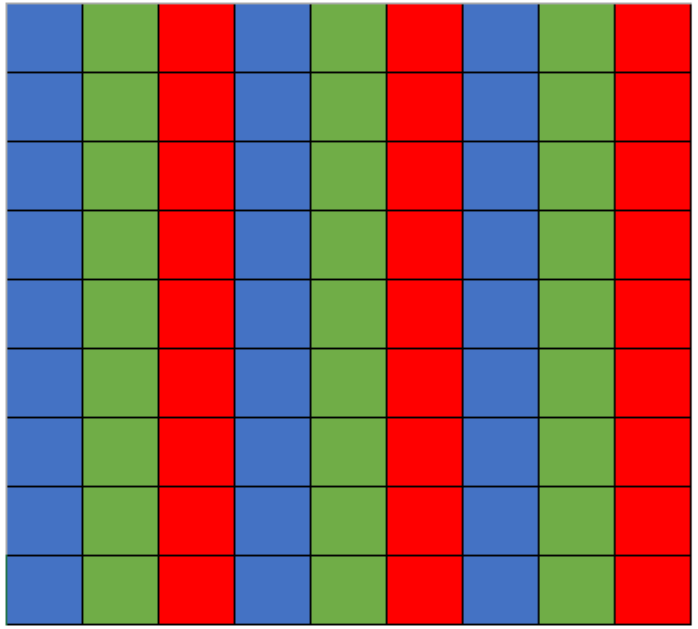
Color space: HSL and HSV

- Cylindrical-coordinate color model
- Channels: hue, saturation and lightness (HSL) or value (HSV)
- Hue changes from 0° to 360°
- Red – 0°, yellow – 60°, green – 120°, cyan – 180°, blue – 220°, magenta – 300°



Color space: Channel order

Packed storage



Planar storage

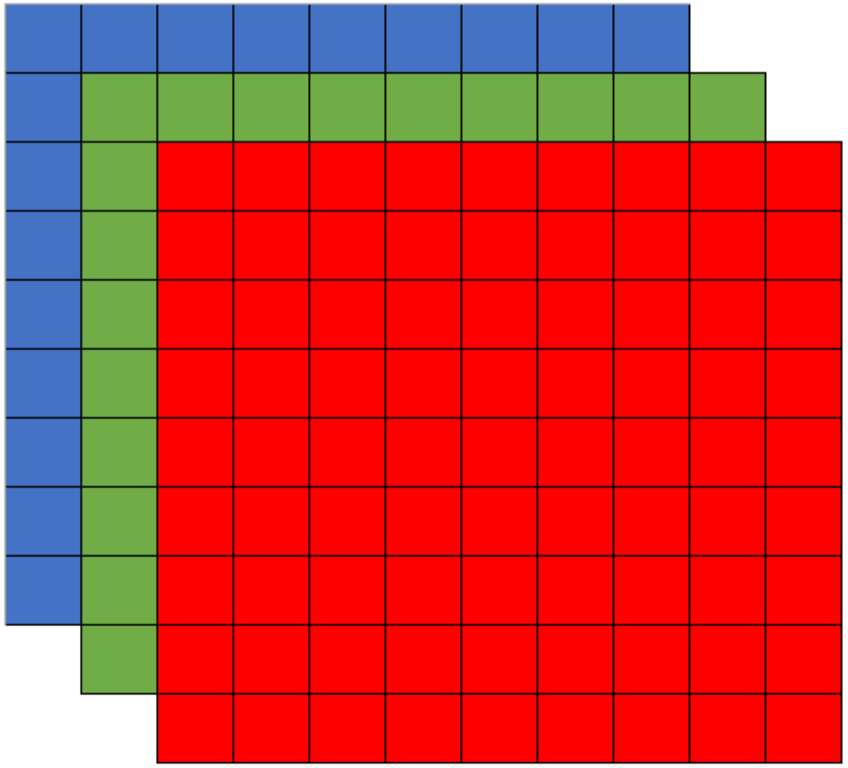


Image processing

- Low-level
  - Operations over pixels
- Mid-level
  - Operations over image or several images
- High-level
  - Operations over image content

Low level: Resize

Low level: Resize

No interpolation

Bilinear interpolation

Bicubic interpolation

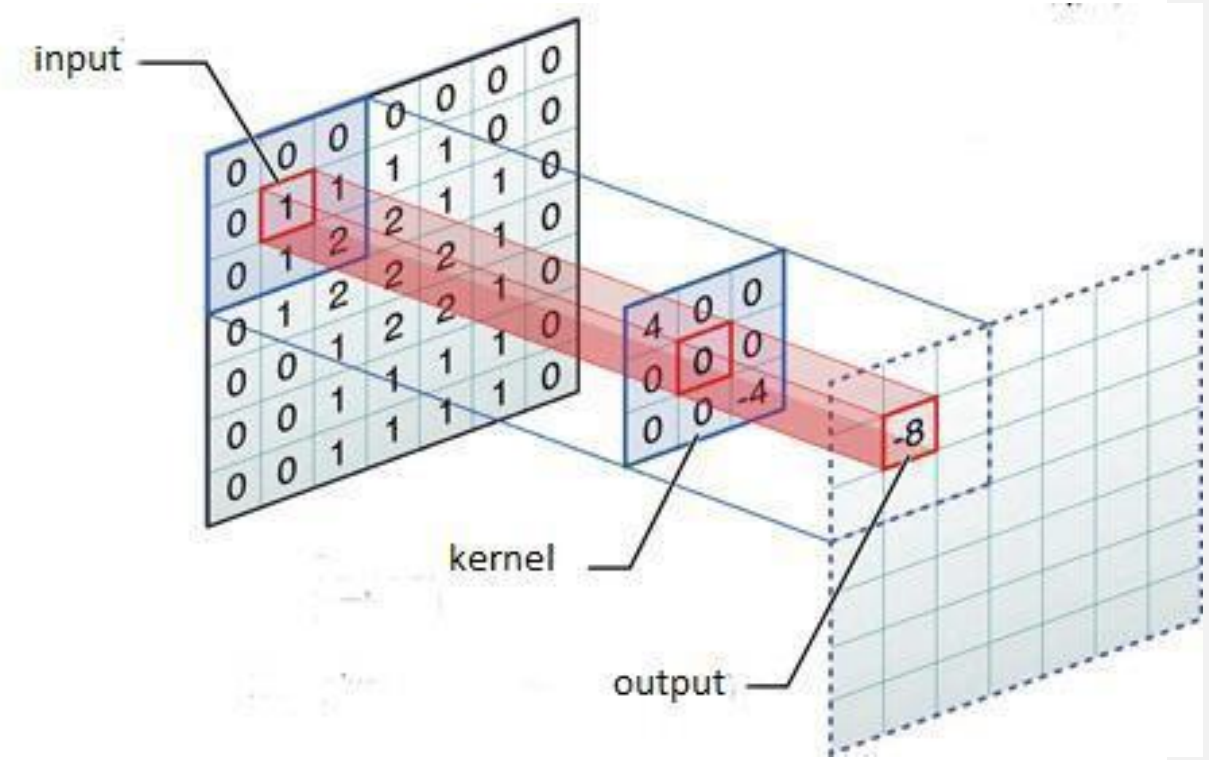
Low level: Convolution

$$(f \overset{-\infty}{\underset{+\infty}{\otimes}} g)(t)$$

$$= \int f(\tau)g(t-\tau)d\tau$$

$$(f \overset{-\infty}{\underset{+\infty}{\otimes}} g)[n]$$

$$= \sum_{m=-\infty} f[m]g[n-m]$$



Low level: Convolution – blur

$$\frac{1}{9}\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}$$

$$\frac{1}{25}\begin{pmatrix}1 & 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1 & 1\end{pmatrix}$$

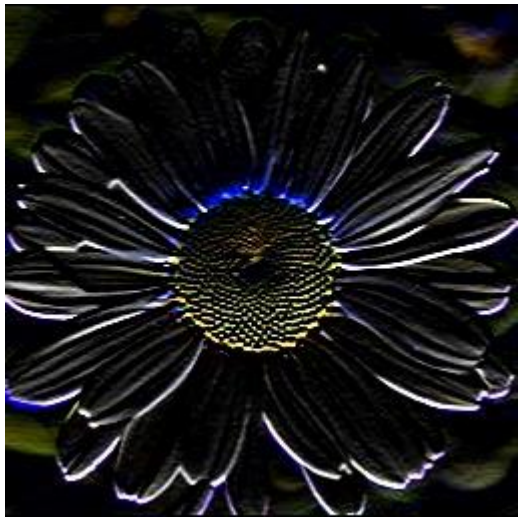
Low level: Convolution – contrast

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Low level: Convolution – edges

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$



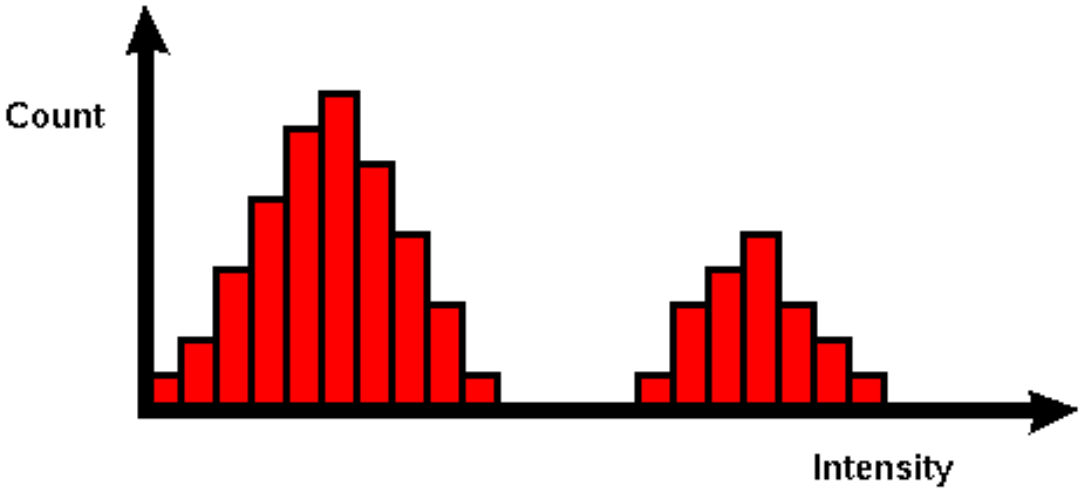
$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

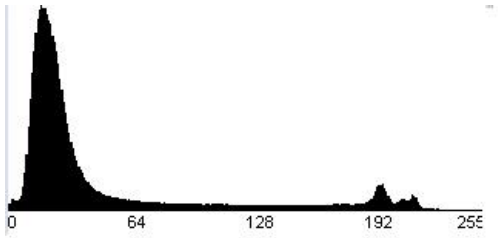
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

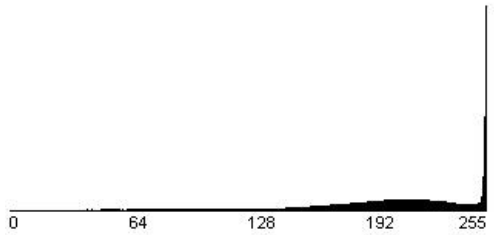
Low level: Color segmentation

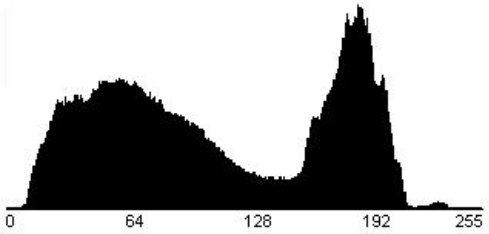
Mid level: Histogram

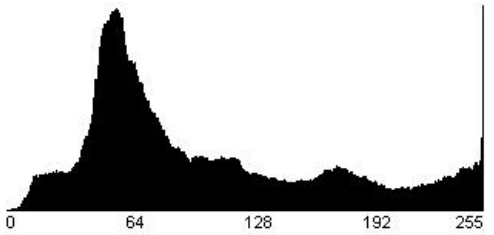
$$Hist[i] = \sum_{x,y} \begin{cases} 1, & Pixel(x,y) = i \\ 0, & Pixel(x,y) \neq i \end{cases}$$





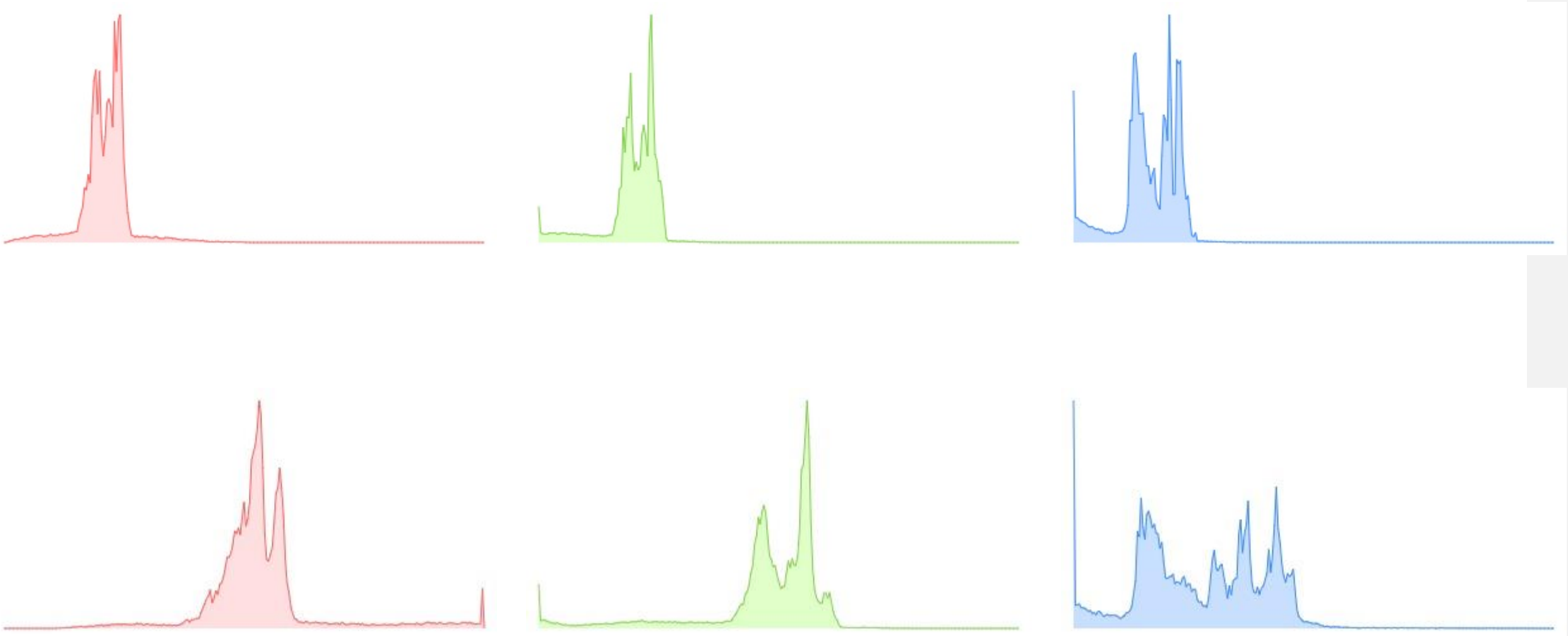






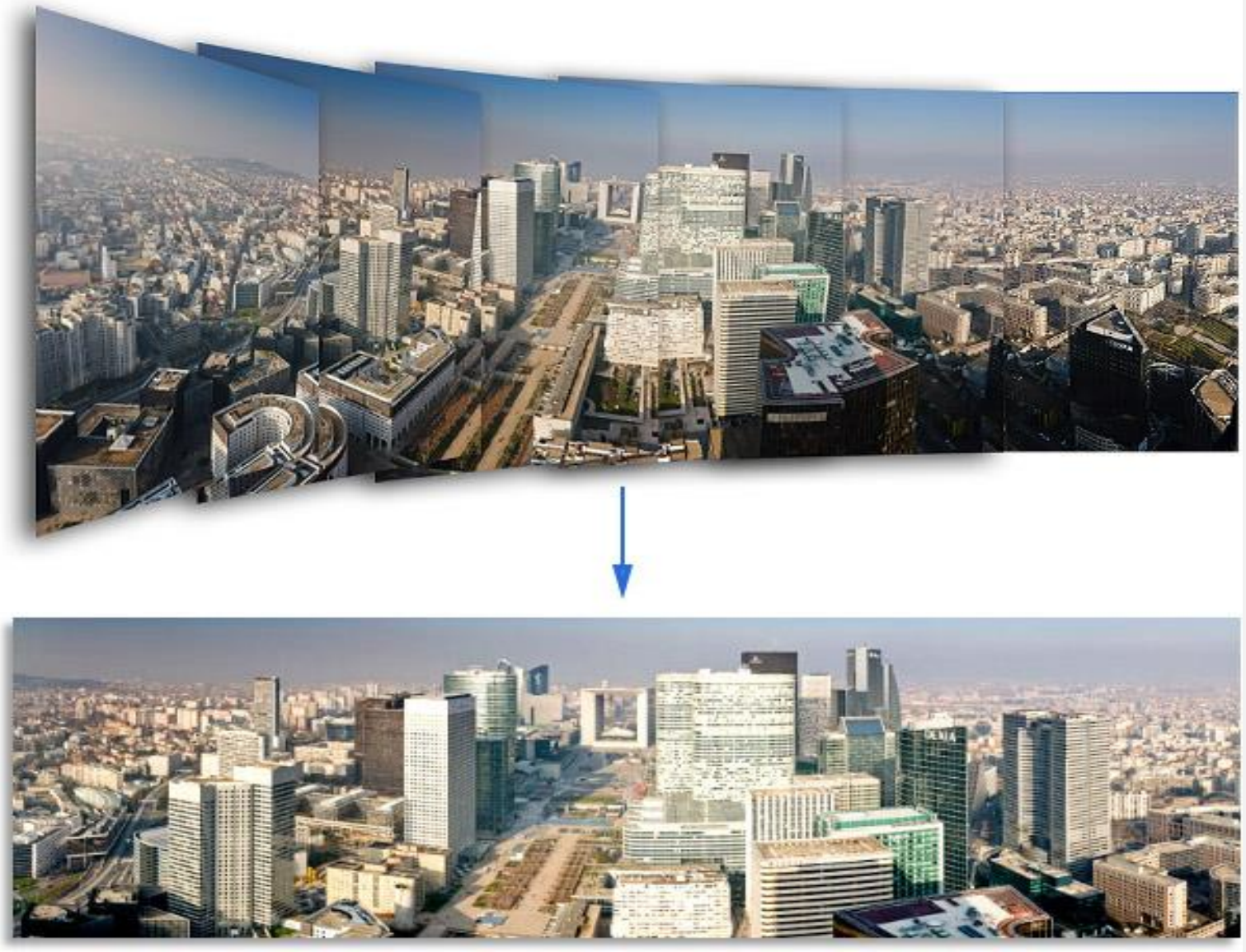
Mid level: Histogram equalization

Mid level: Histogram equalization



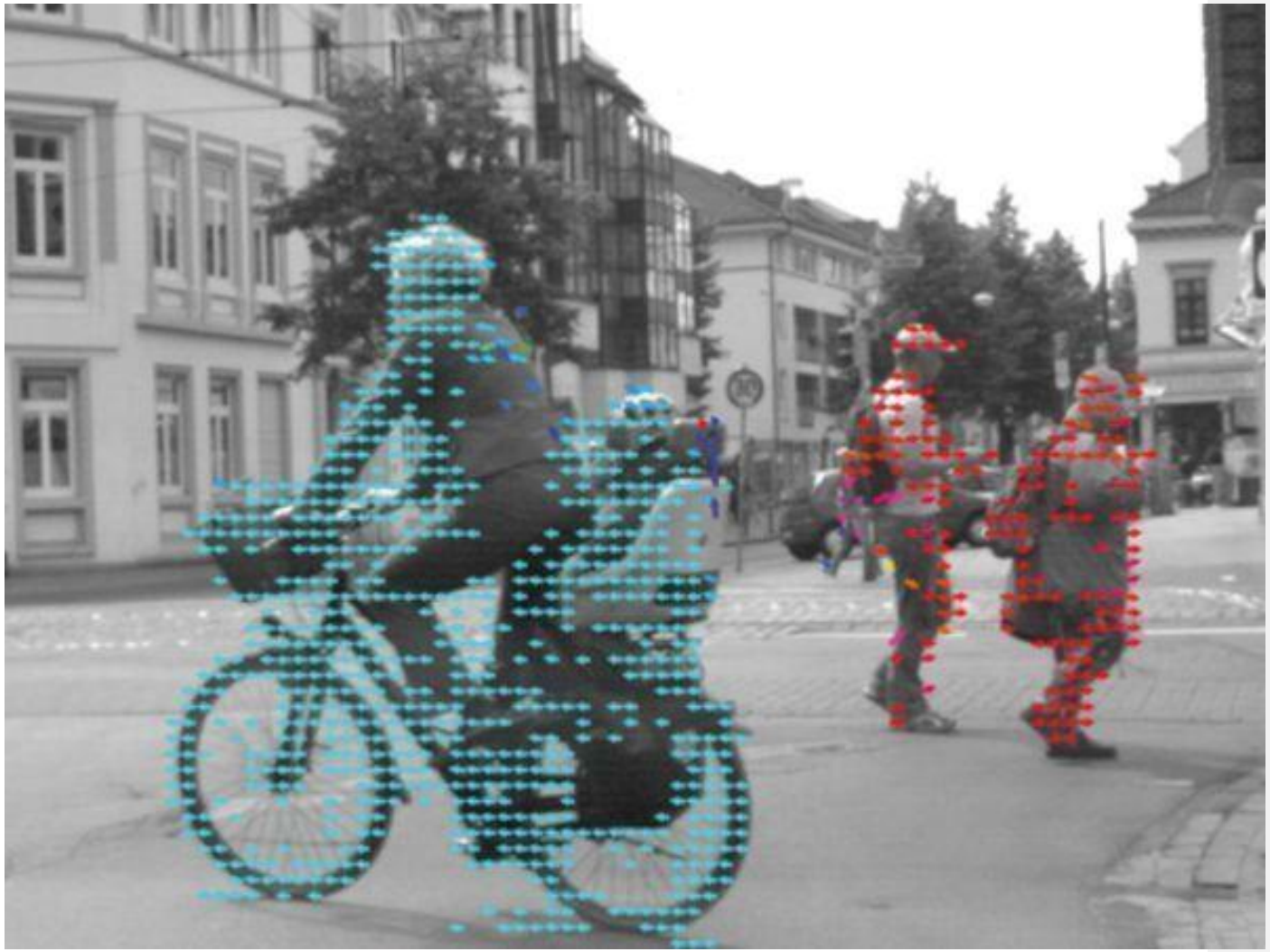
Mid level: Panorama

- Several images
- Stitching by keypoints



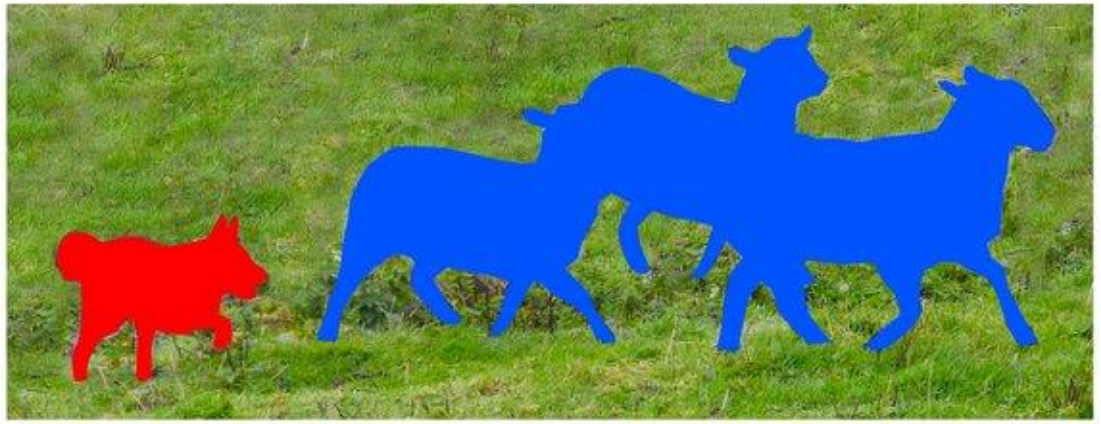
Mid level: Optical flow

- Temporal image sequence
- Tracking pixels changes

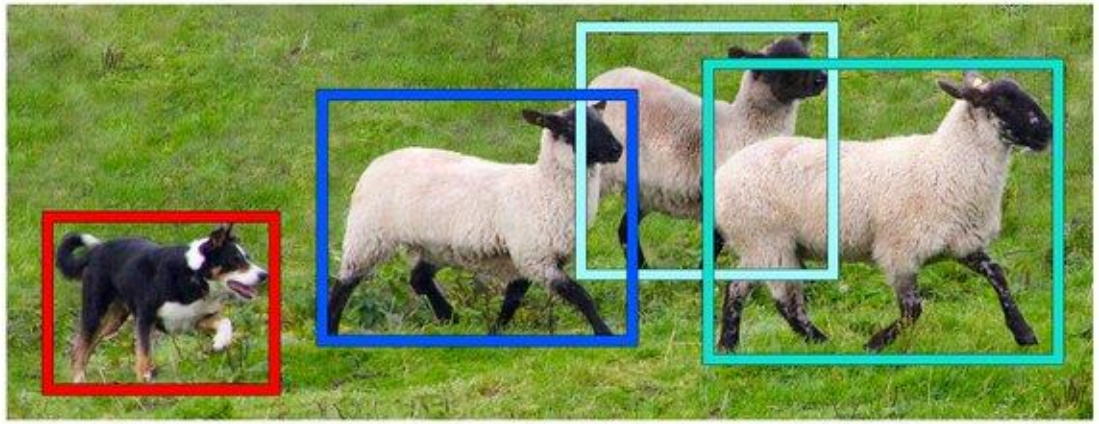


High level

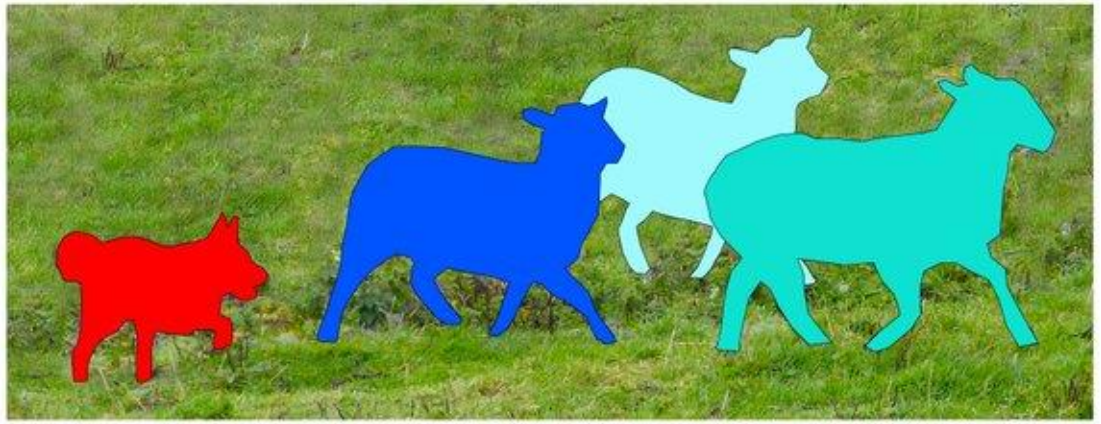
Image Recognition



Semantic Segmentation



Object Detection



Instance Segmentation

High level

- Classification
- Detection
- Segmentation
- Super-resolution
- Image generation
- etc.


intel