Image Processing

01

Image Processing

Dr. Márton Szemenyei
Associate Professor
2024

Lecturer

Dr. Márton Szemenyei

Associate Professor

szemenyei@iit.bme.hu

IB411

(463) 4025

Subject information

Subject focus: Teams & Moodle

- Slides
- Notes
- Questions

Data sheet

https://portal.vik.bme.hu/kepzes/targyak/VIIIAD01

Suggested literature

- Online published notes (pdf)

Requirements

Midterms

27th November (Wednesday) 18-20, Q1

Essay questions

30% (18/60p) needed

Retake: 11th December (Wednesday) 10-12, ?

Practicals

12 Practicals, in IB410

You need to show up at least 8 times (NO REDO OPPORTUNITIES!)

Online Test at the end of each practical, we take the best 8

30% (18/60p) needed, can retake 3 on the 14th week (but only if you were present)

Final grade

Midterms+Practicals, 40% (46/120p) needed

Grade goes up every 18 points

Scheduling plan

Lectures 1-8

- Traditional computer vision

Lectures 9-13

- Intelligent vision systems

Midterms, 27th November, 18:00

Retake, 11th December, 10:00

Scheduling plan

Week	Wednesday	Room	Thursday	Room
1	Introduction	IB410	Python basics	IB410
2	Image Correction	IB410	Color tracking	IB410
3	Frequency Domain	IB410	OCR	IB410
4	Feature Detection I	IB410	OCR in Frequency	IB410
5	Feature Detection II	IB410	Stereo Vision	IB410
6	Segmentation	IB410	Document orientation	IB410
7	Binary Images	IB410	Disparity, distance	IB410
8	National Holiday		Motion detection	IB410
9	Neural Network Basics	IB410	Object separation	IB410
10	Convolutional Networks	IB410	Motion tracking	IB410
11	Deep Learning in Practice	IB410	Neural Network basics	IB410
12	Detection, Segmentation	IB410	Holiday	
13	Unsupervised Learning	IB410	Classification	IB410
14	Detection	IB410	Practical Test Retakes	IB410

Course Goals

How do we call the function? → "**Google**"

Connections

Why

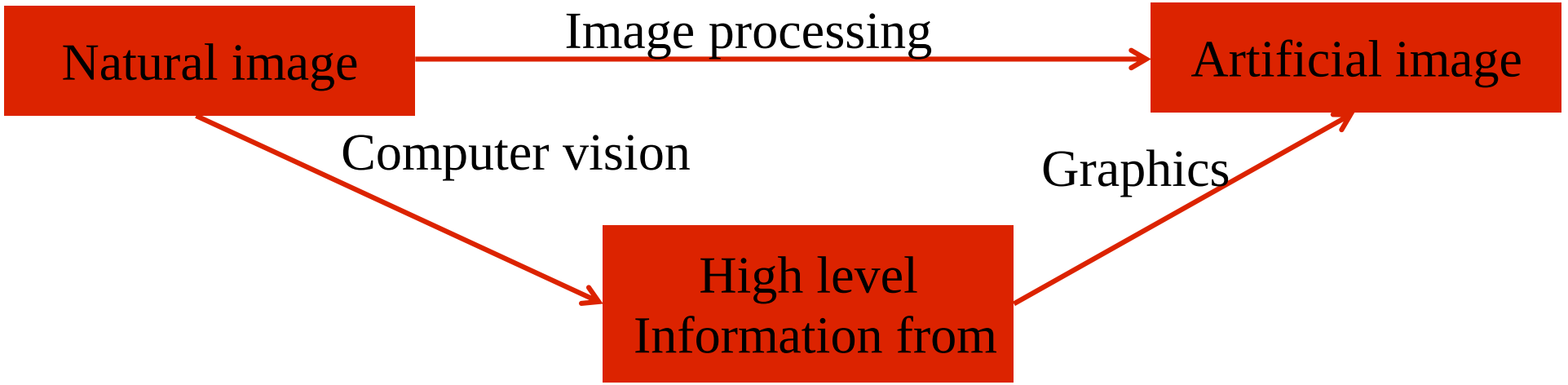
How

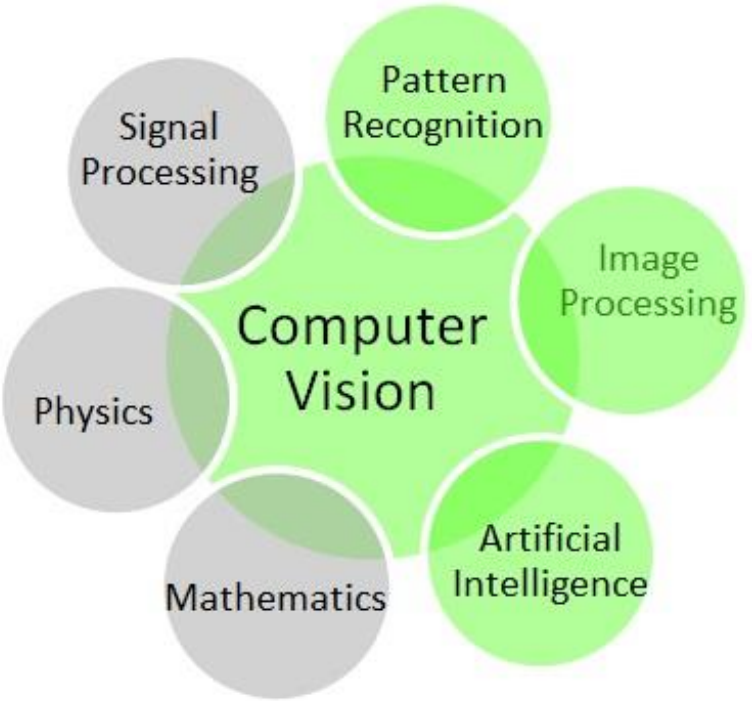
What

Which

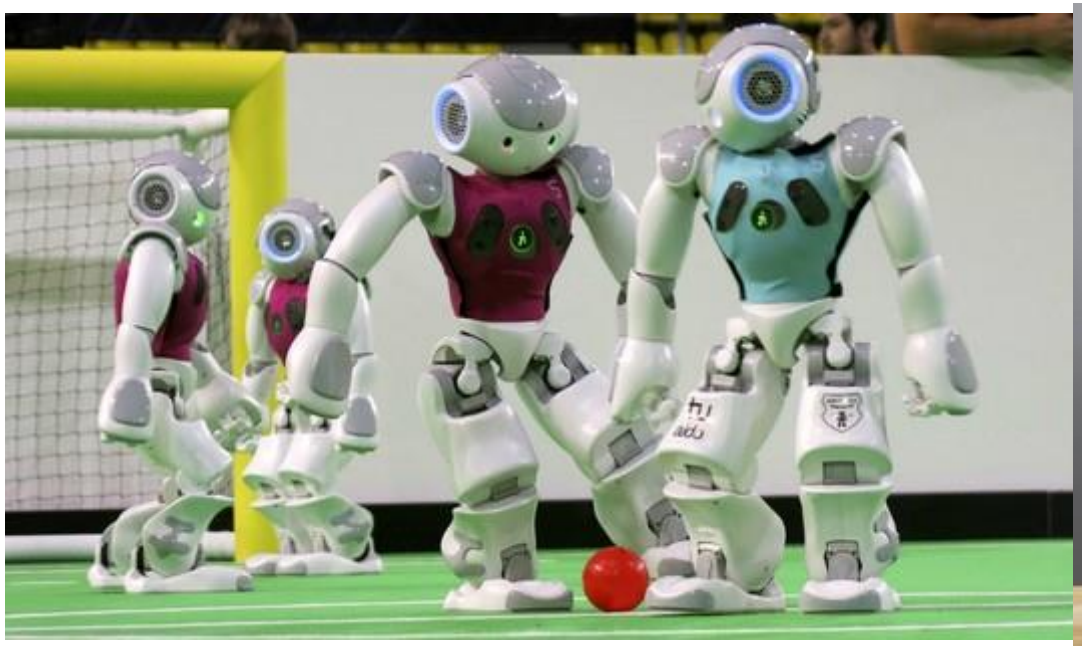
Goal: Vision

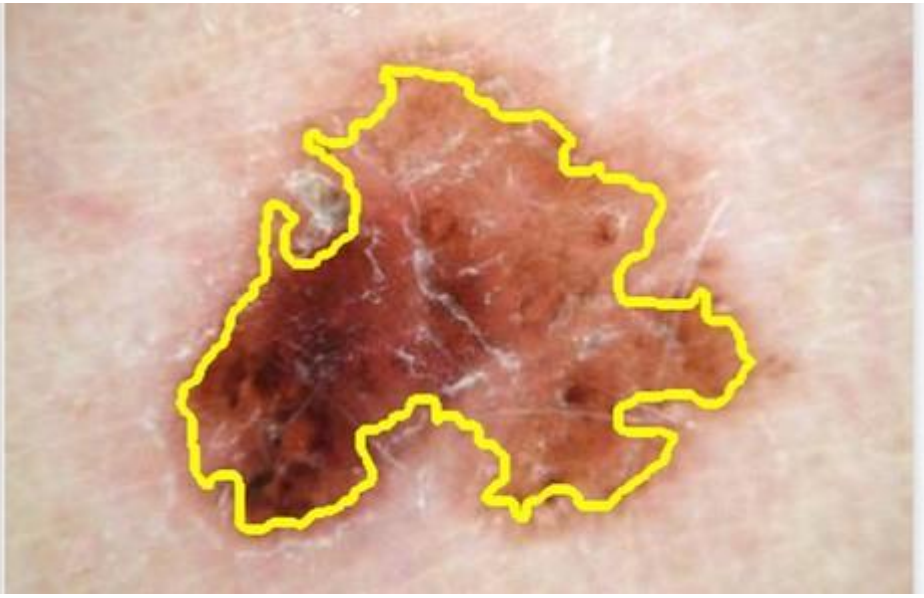
Computer vision





Computer vision

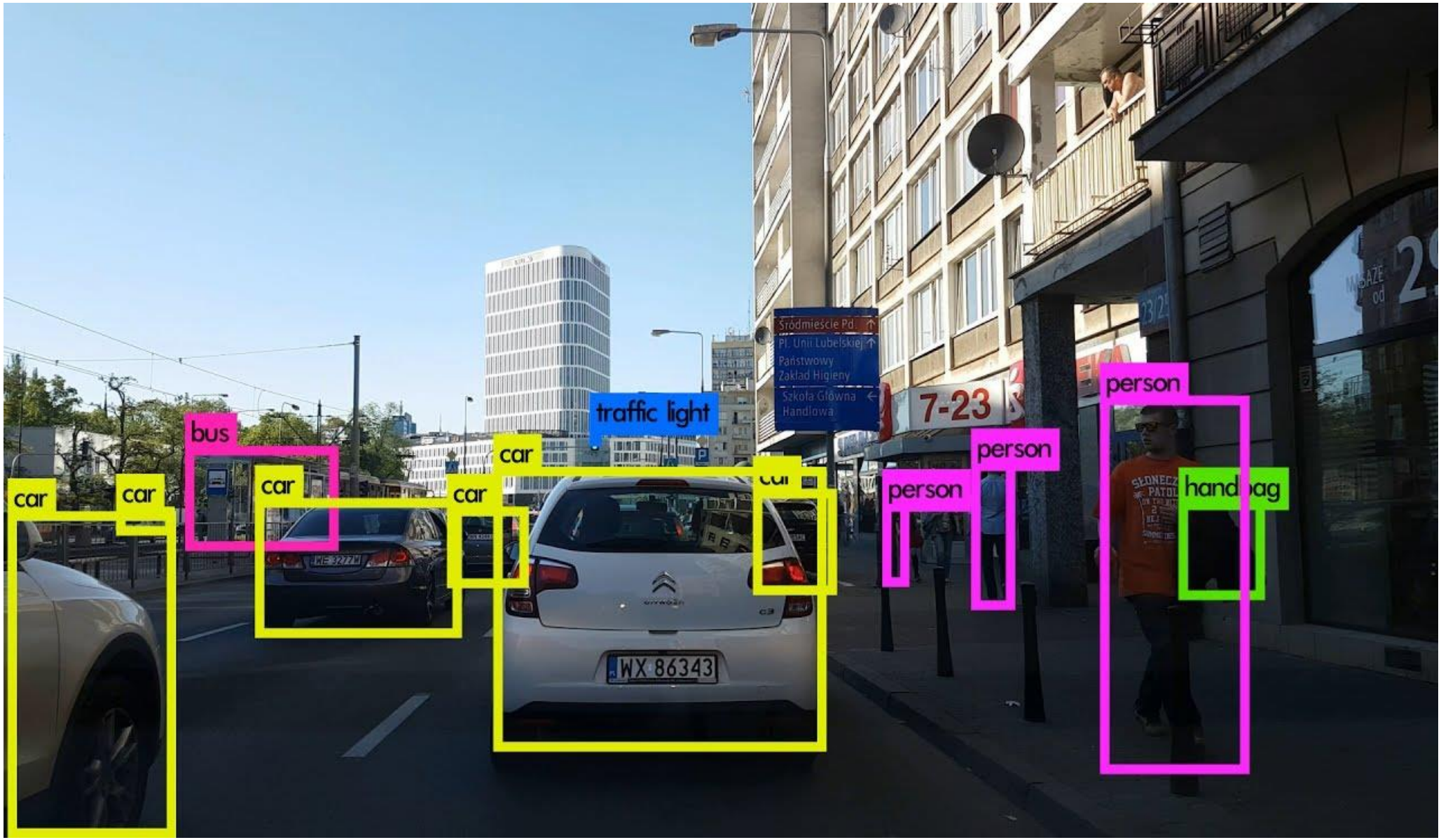




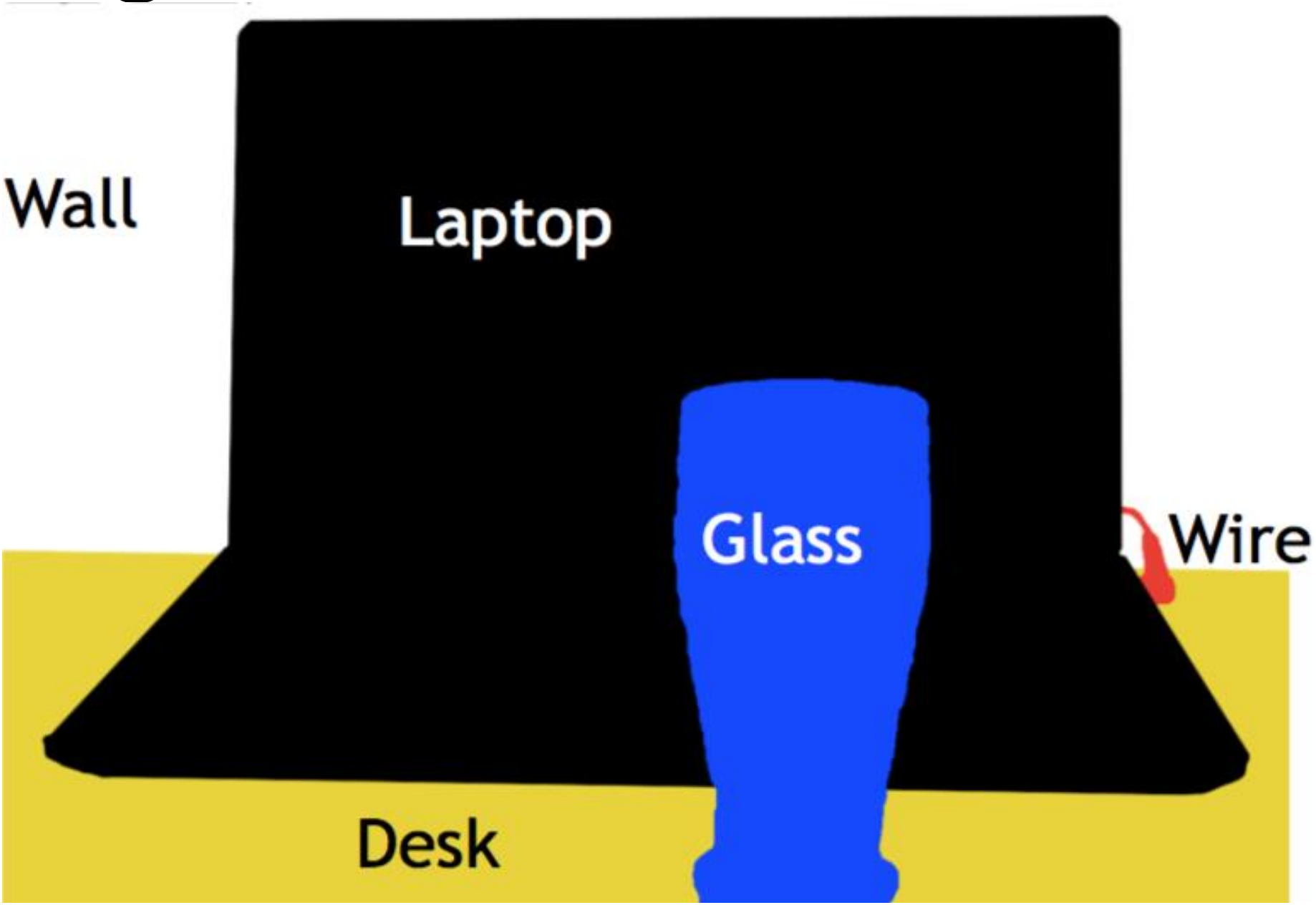
Classification

→ Cat

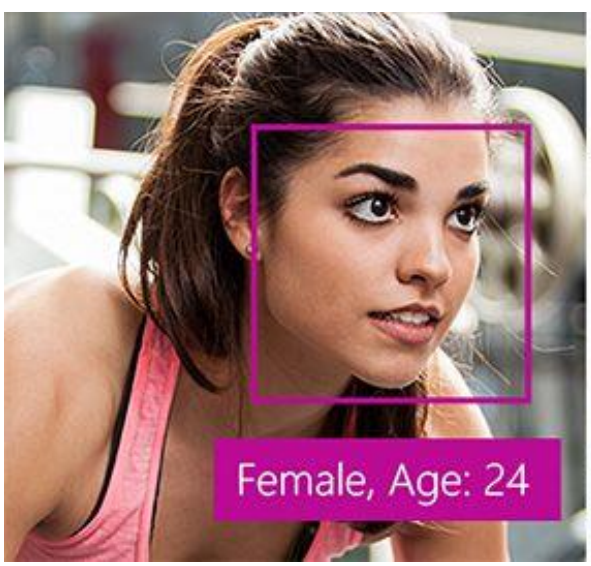
Detection

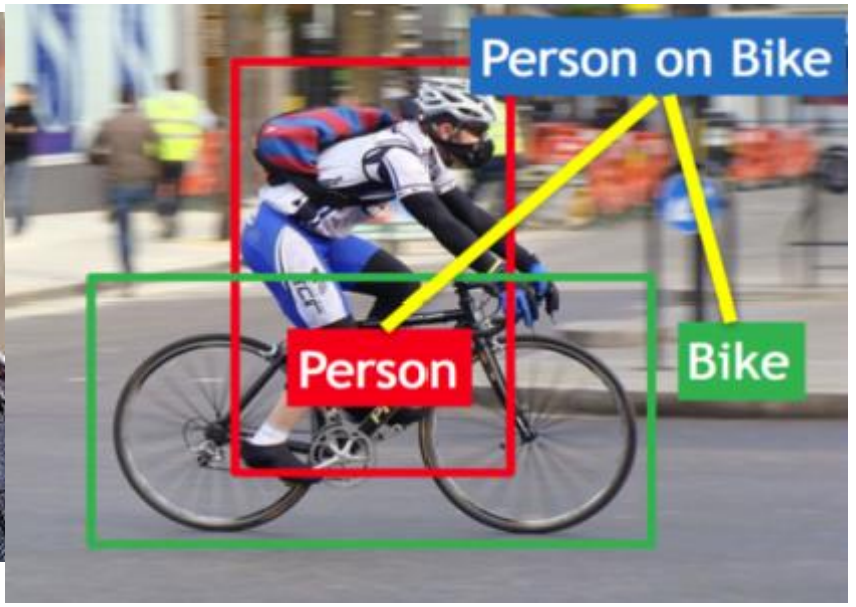

car
car
bus
car
car
traffic light
car
person
person
person
handbag
WX 86343
7-23

Segmentation



Computer Vision





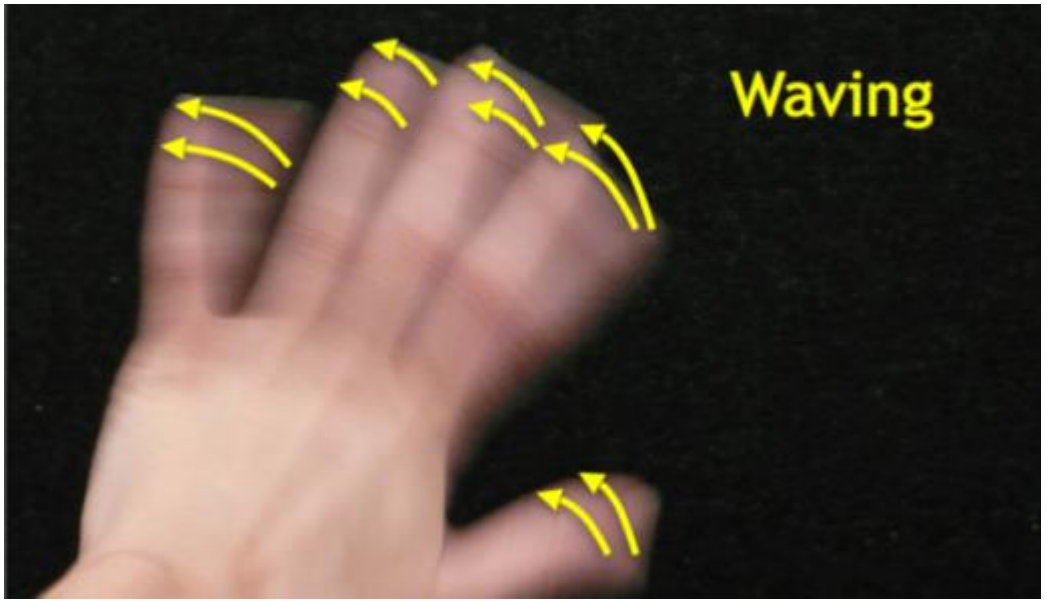
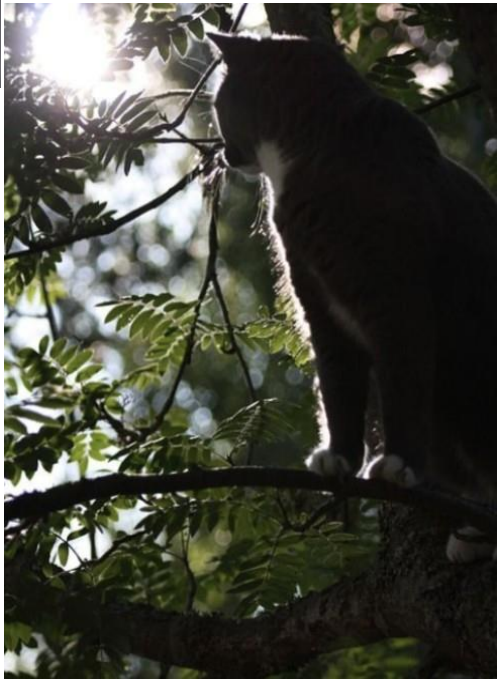


Image Understanding

Problem: Semantic Gap

```
[[105 112 108 111 104  99 106  99  96 103 112 119 104  97  93  87]
 [ 91  98 102 106 104  79  98 103  99 105 123 136 110 105  94  85]
 [ 76  85  90 105 128 105  87  96  95  99 115 112 106 103  99  85]
 [ 99  81  81  93 120 131 127 100  95  98 102  99  96  93 101  94]
 [106  91  61  64  69  91  88  85 101 107 109  98  75  84  96  95]
 [114 108  85  55  55  69  64  54  64  87 112 129  98  74  84  91]
 [133 137 147 103  65  81  80  65  52  54  74  84 102  93  85  82]
 [128 137 144 140 109  95  86  70  62  65  63  63  60  73  86 101]
 [125 133 148 137 119 121 117  94  65  79  80  65  54  64  72  98]
 [127 125 131 147 133 127 126 131 111  96  89  75  61  64  72  84]
 [115 114 109 123 150 148 131 118 113 109 100  92  74  65  72  78]
 [ 89  93  90  97 108 147 131 118 113 114 113 109 106  95  77  80]
 [ 63  77  86  81  77  79 102 123 117 115 117 125 125 130 115  87]
 [ 62  65  82  89  78  71  80 101 124 126 119 101 107 114 131 119]
 [ 63  65  75  88  89  71  62  81 120 138 135 105  81  98 110 118]
 [ 87  65  71  87 106  95  69  45  76 130 126 107  92  94 105 112]
 [118  97  82  86 117 123 116  66  41  51  95  93  89  95 102 107]
 [164 146 112  80  82 120 124 104  76  48  45  66  88 101 102 109]
 [157 170 157 120  93  86 114 132 112  97  69  55  70  82  99  94]
 [130 128 134 161 139 100 109 118 121 134 114  87  65  53  69  86]
 [128 112  96 117 150 144 120 115 104 107 102  93  87  81  72  79]
 [123 107  96  86  83 112 153 149 122 109 104  75  80 107 112  99]
 [122 121 102  80  82  86  94 117 145 148 153 102  58  78  92 107]
 [122 164 148 103  71  56  78  83  93 103 119 139 102  61  69  84]]
```

Difficulty: lighting

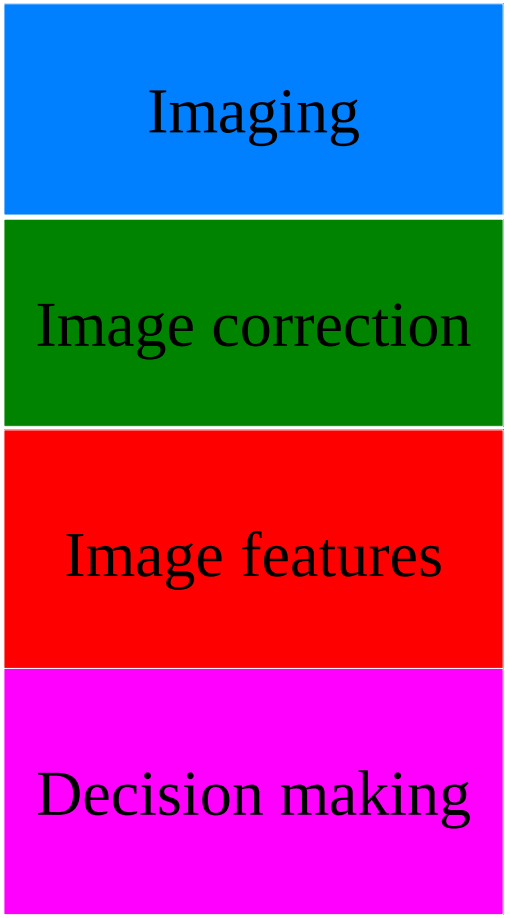


Difficulty: deformation

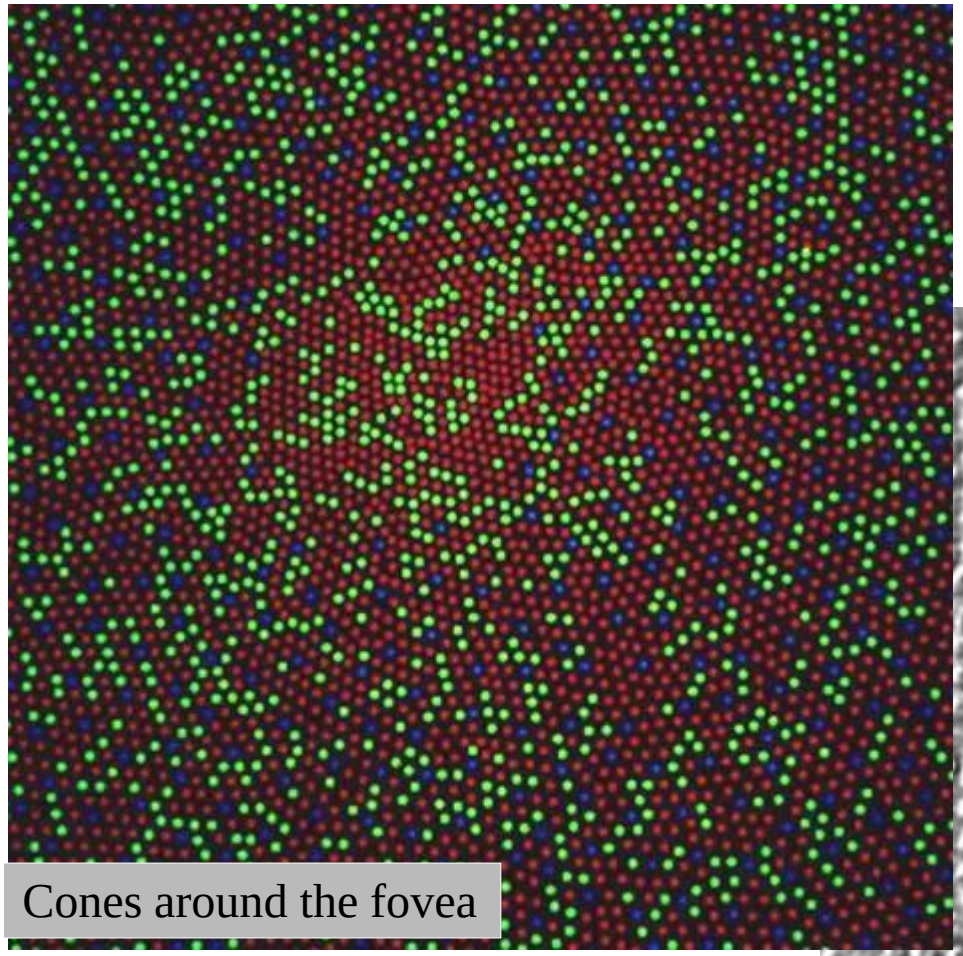
Difficulty: occlusion

Difficulty: intra-class variation

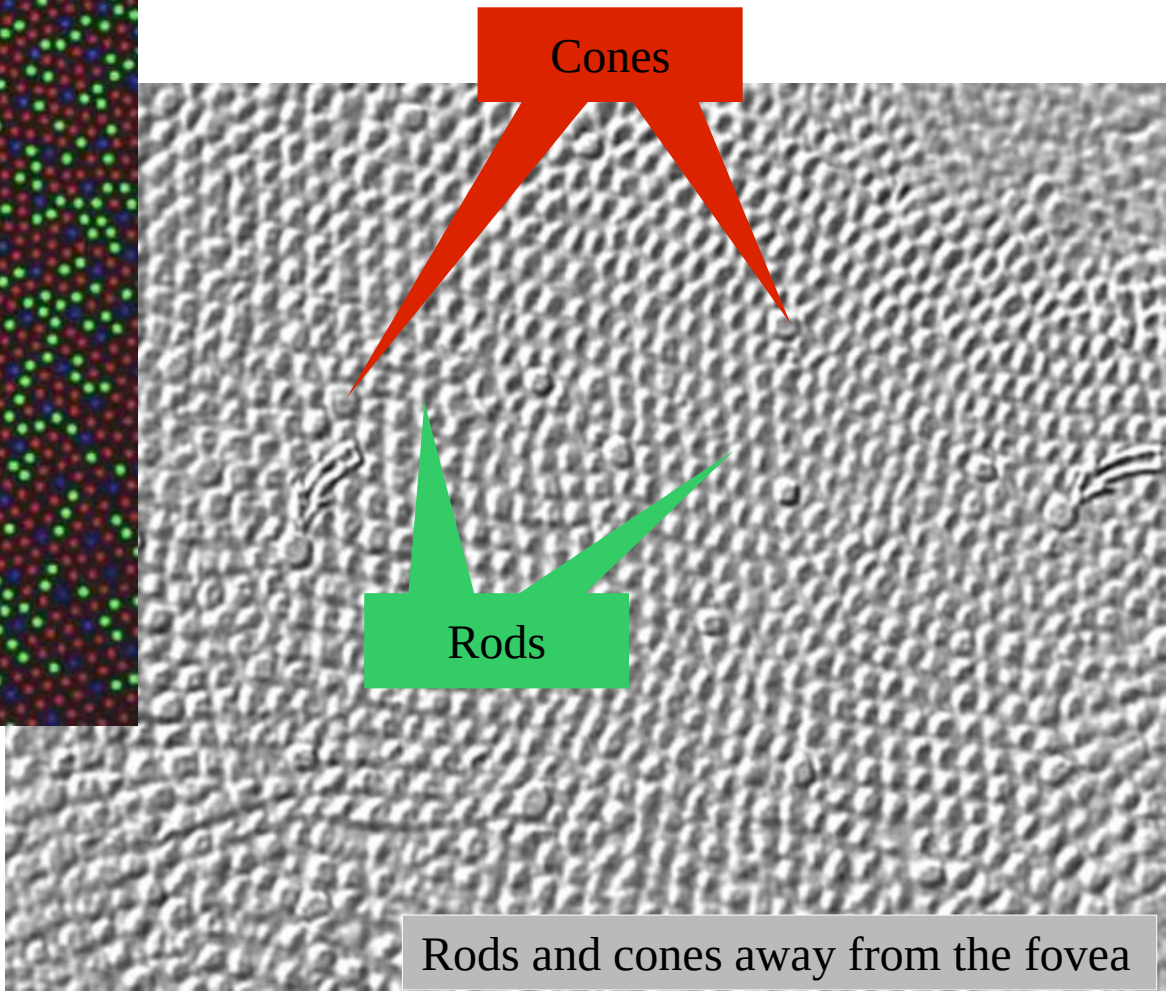
Traditional vision



Human Vision



Cones around the fovea



Rods and cones away from the fovea

Human colour vision

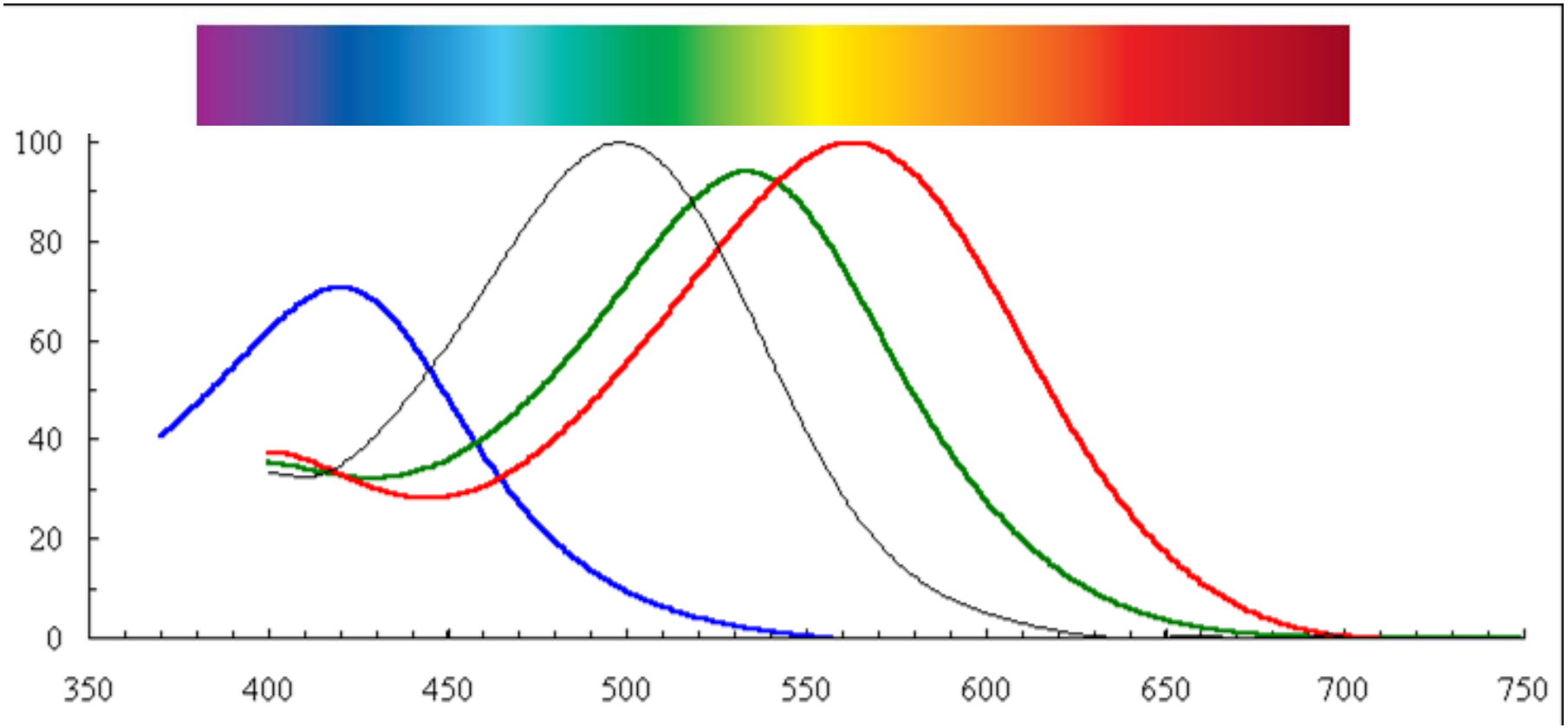


Photo diode

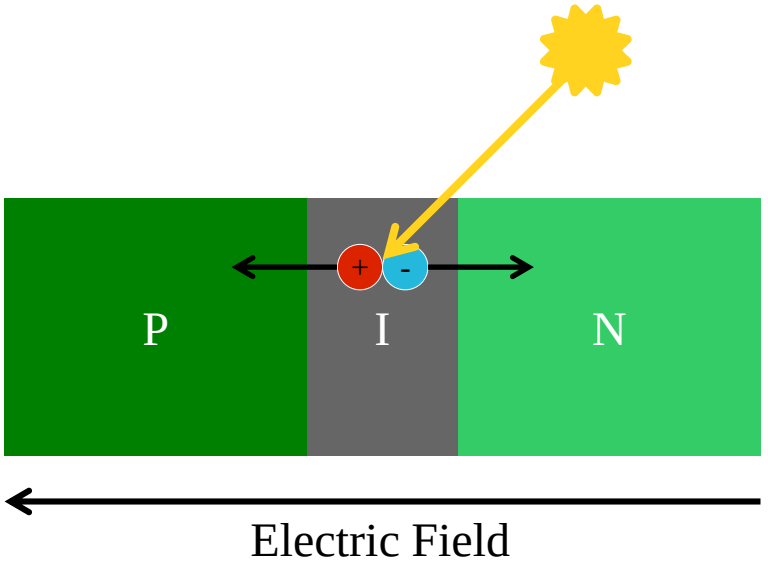
All diodes are sensors

Photovoltaic mode

Photoelectric mode

- Faster reaction
- More noise

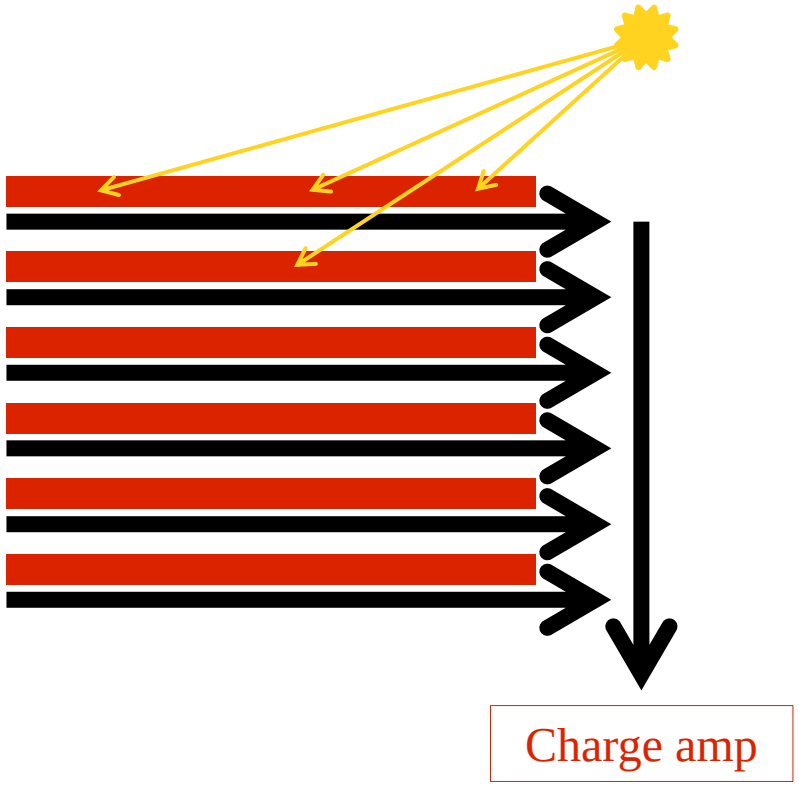
Phototransistor

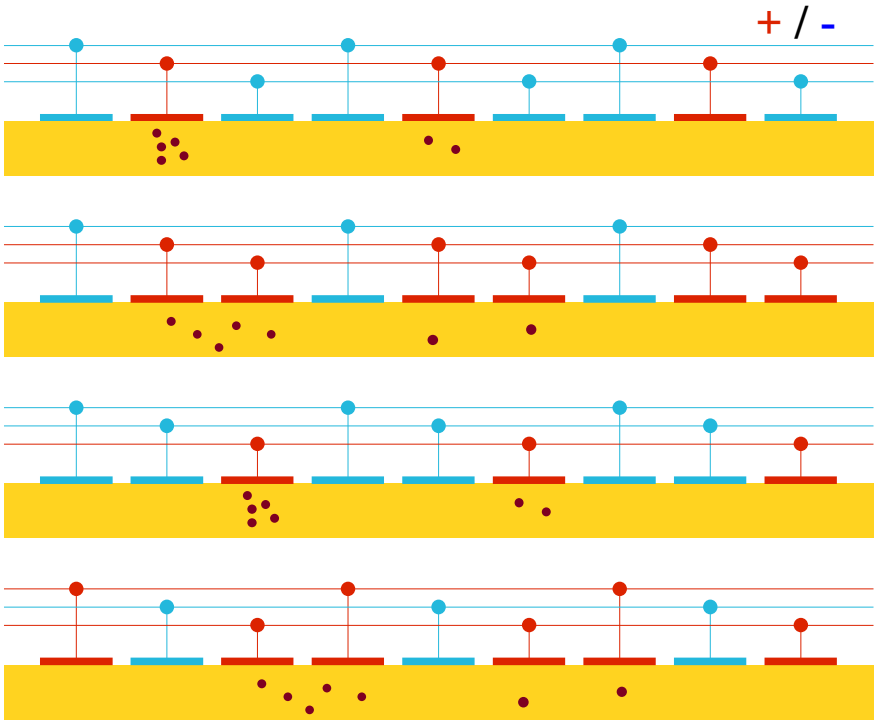


CCD

Charge-Coupled Device (CCD)

Analog shift registers





CCD types

Full frame CCD

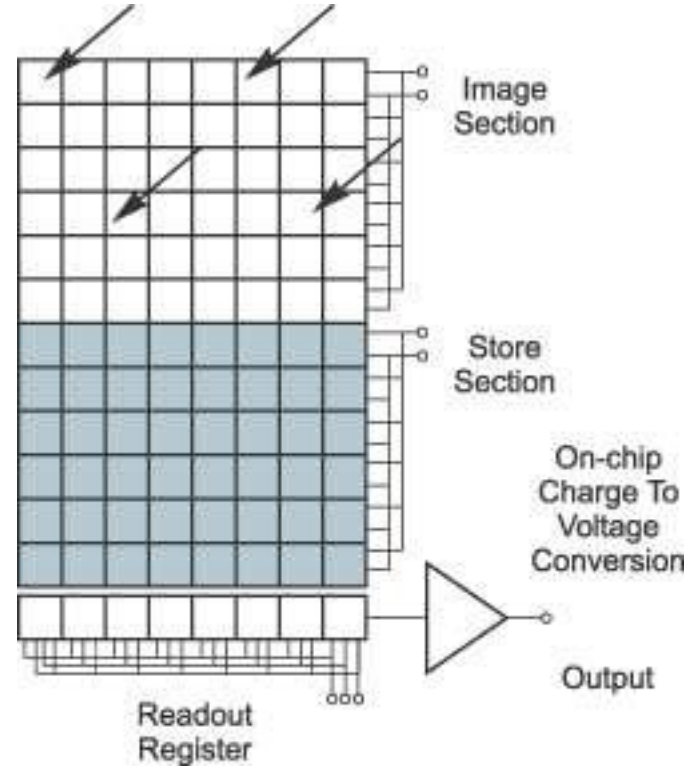
No per-pixel storage
External shutter required

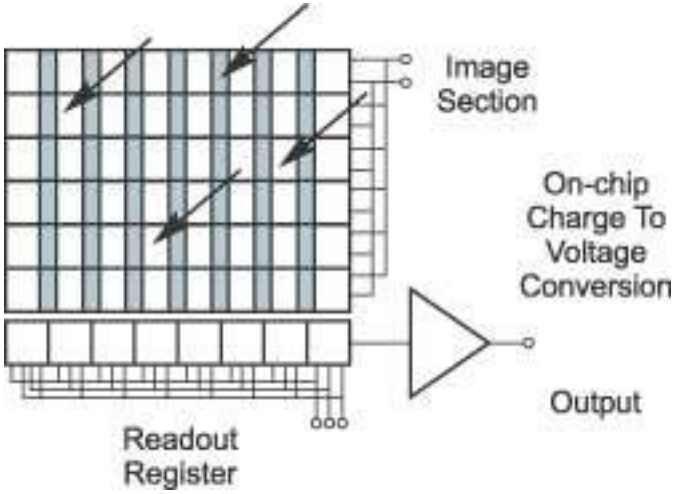
Frame-transfer CCD

CCD behind the photo sensors

Interline CCD

CCD between photoactive columns





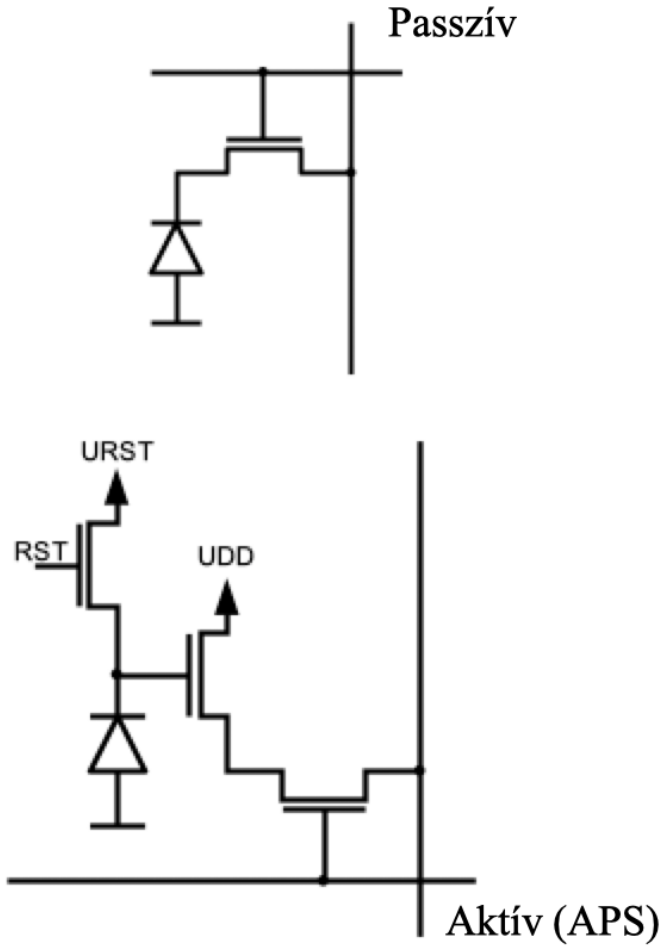
CMOS

Advantages and disadvantages

- Less consumption
- Smaller delay
- Cheaper design
- Higher speed
- More noise
- Less sensitivity

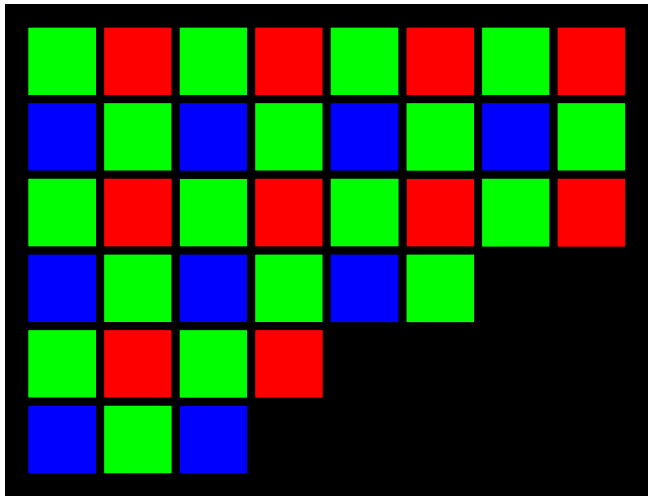
Pixel-level pre-processing

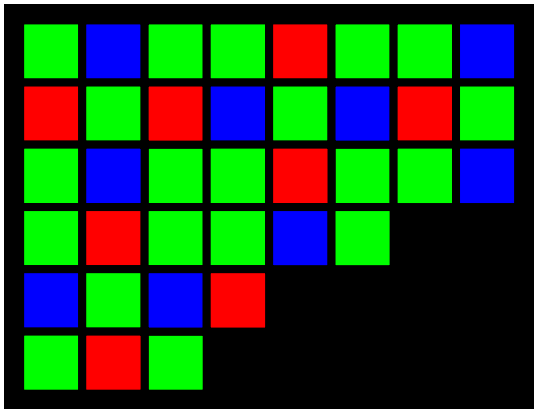
- CNN, AER...



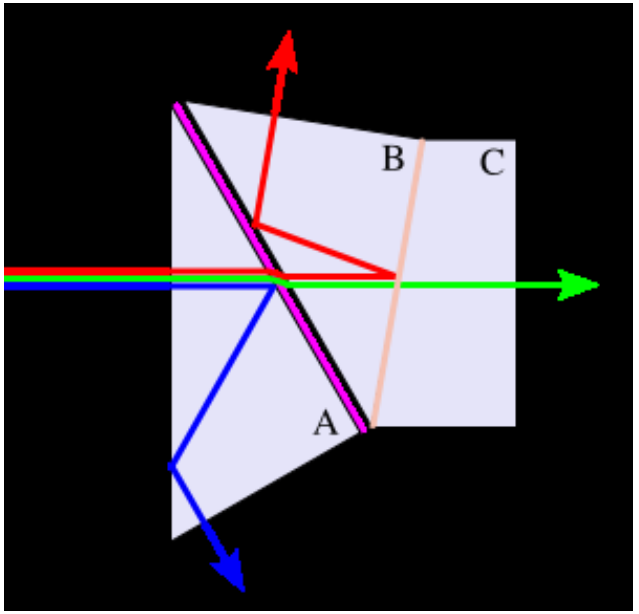
Colour cameras

Bayer





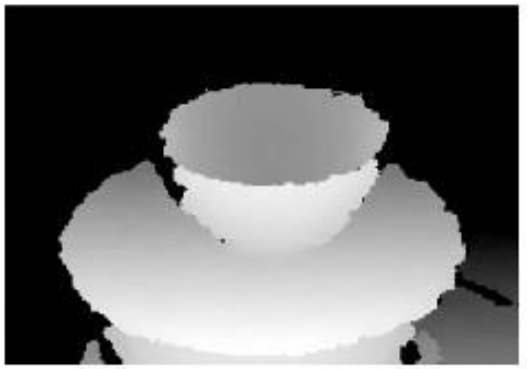
3CCD



Line CCD

Camera types

- Sensitivity range
  - Visible range
  - Infrared
- Spatiality
  - Traditional (2D)
  - Line camera (1D)
  - Depth sensors (3D)
    - Stereo camera
    - Structured light
    - LIDAR



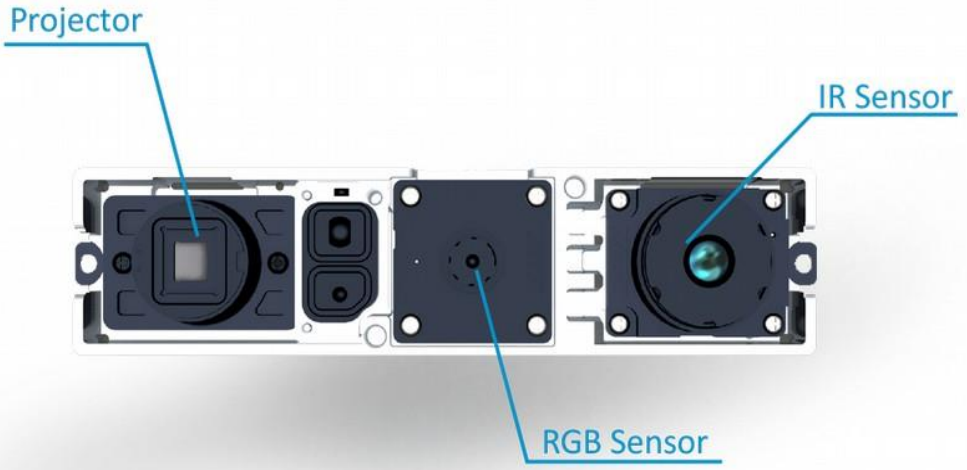
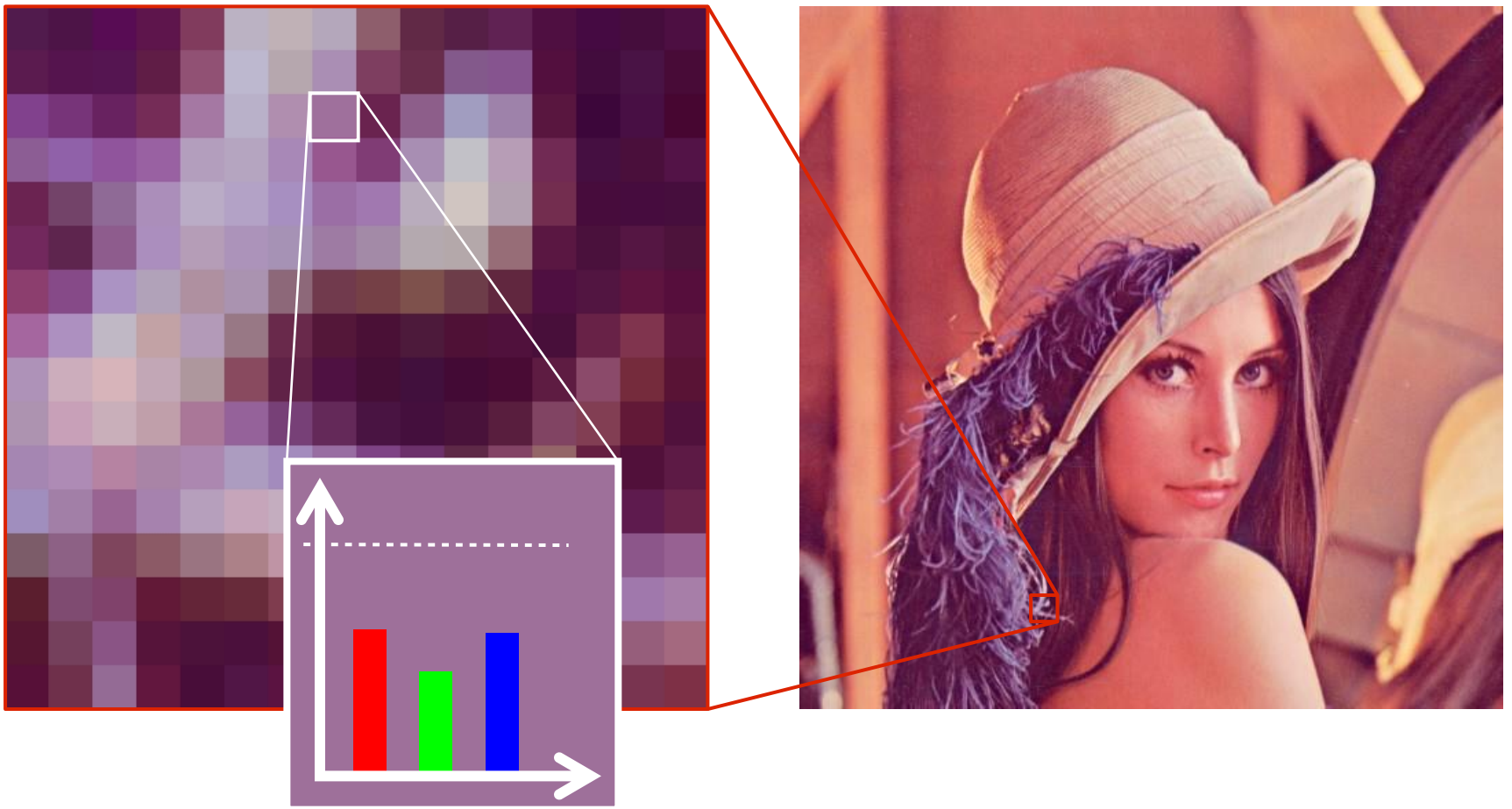


Image basics

Typically: 1D/2D/3D array (pixel/voxel array)



Resolution

Bit depth

1Bit image, Dithering

Digital noise

8x averaging

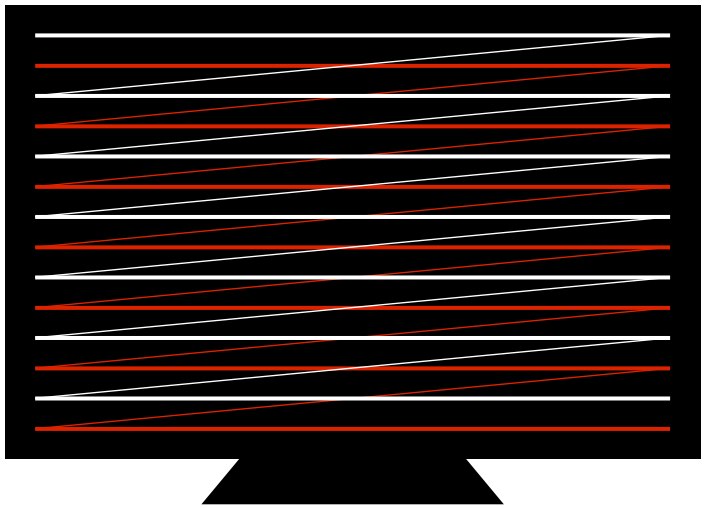
Video stream

Typical: A series of pictures (BUT...)

Interlace

Goal: Double the temporal resolution

TV standards!



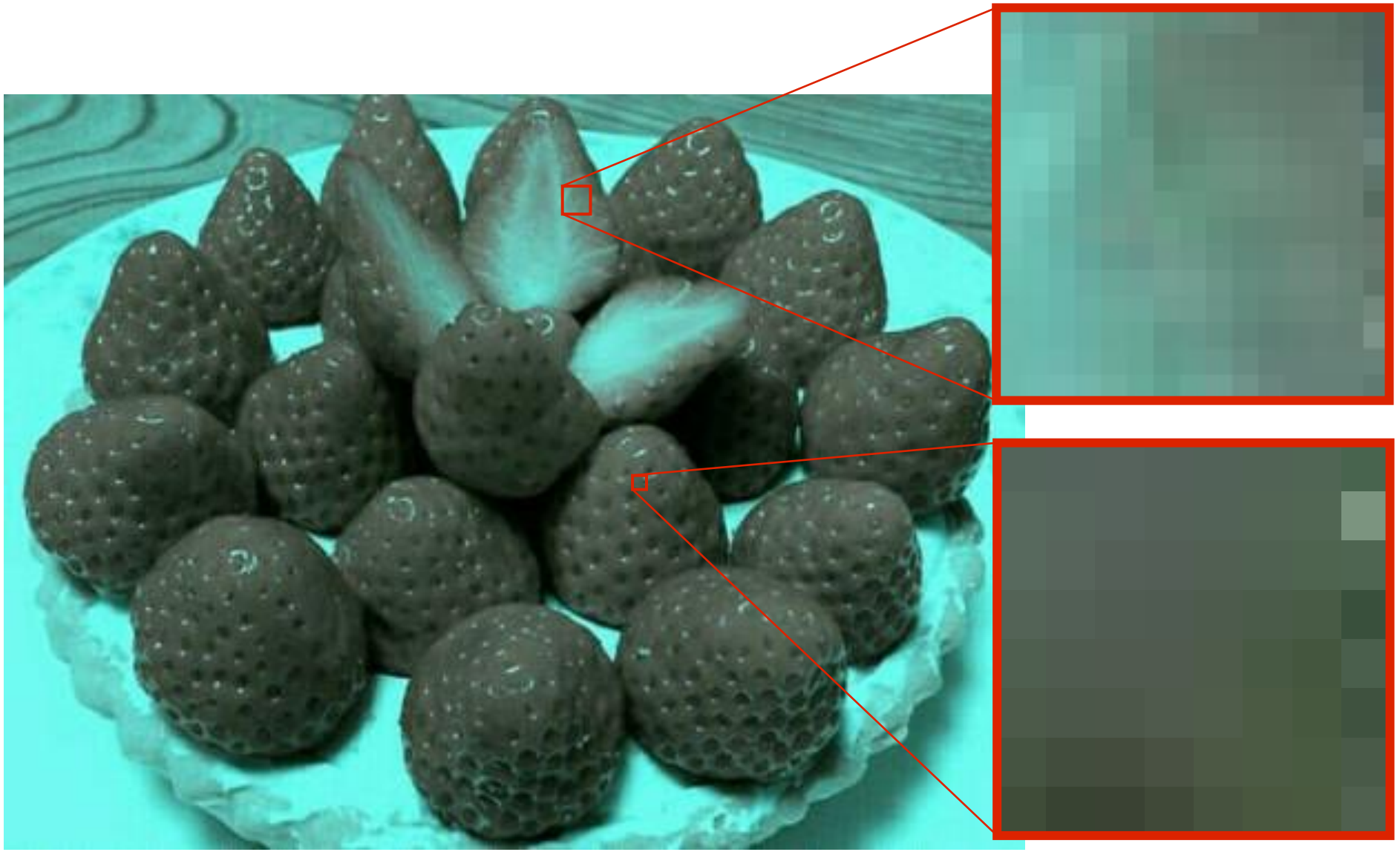
1st frame

2nd frame

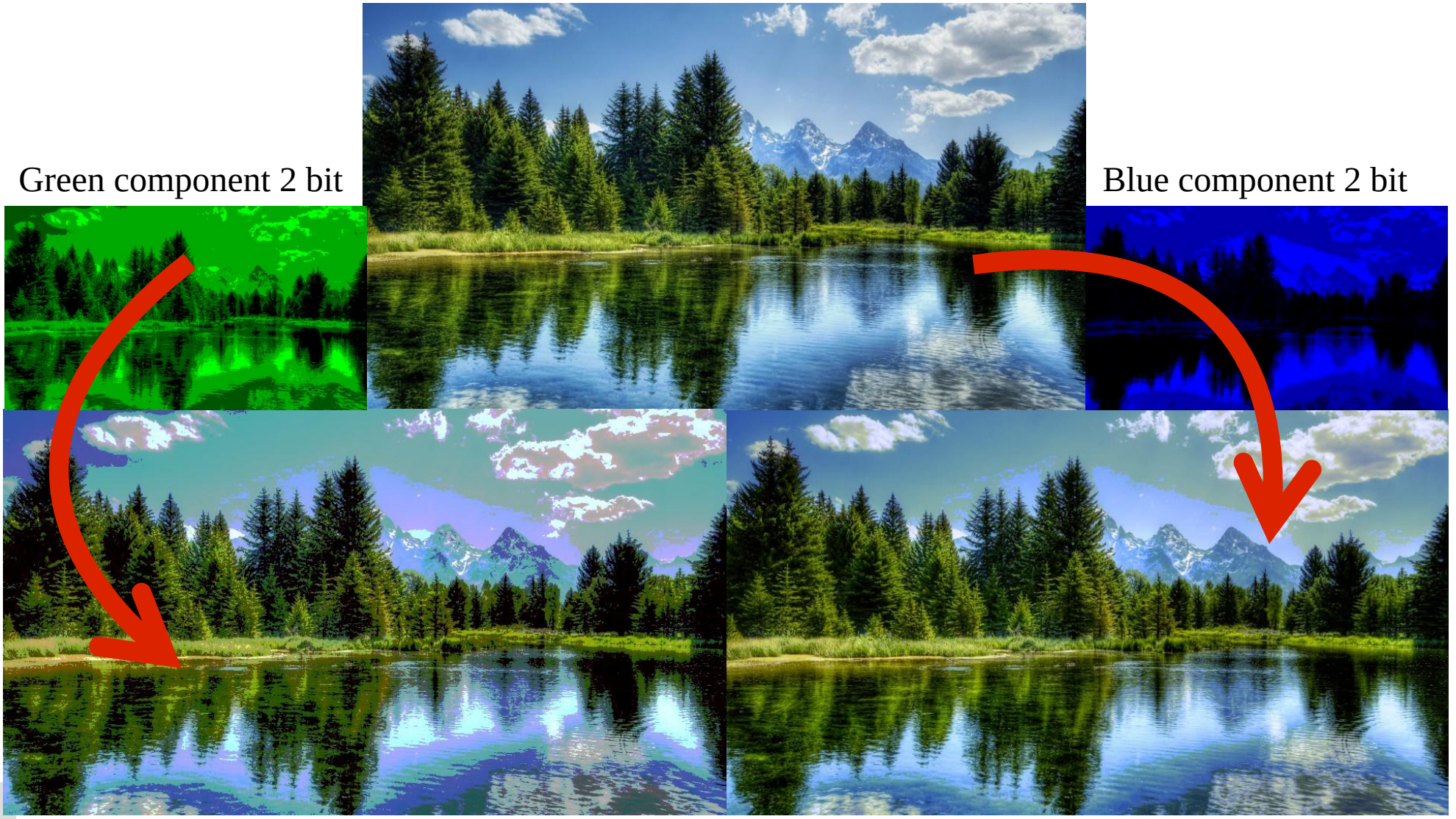
Motion

Focus

Perceived colour



Role of colour components



Storage

Image (3648 x 2736 - 10MPx, colour): 30MB

Video (HD1080, 30fps, 1 hour): 672GB (1.5Gbit/sec)

CT (1000^3 ,12 bit): 1GB

Compression: **human vision!**

HDD: 100-300 MB/sec

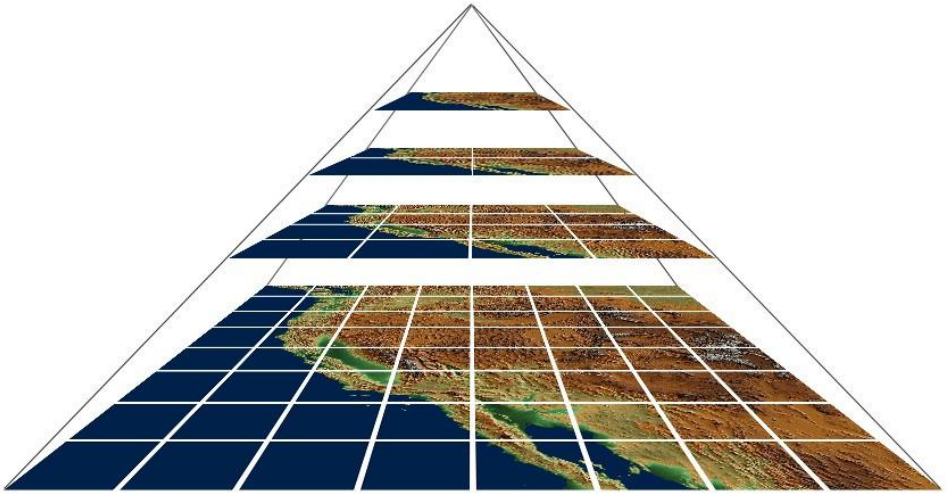
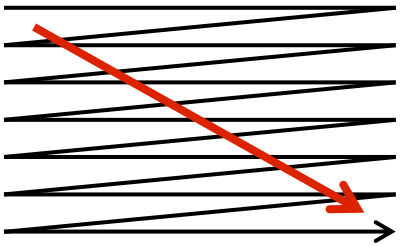
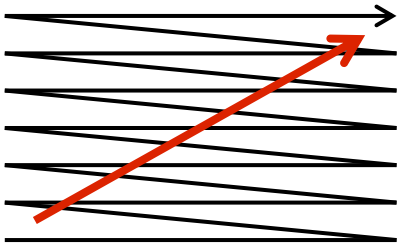
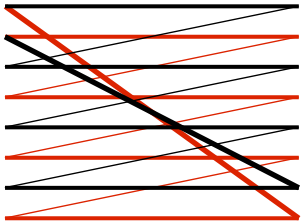


Image content







[I_0 I_1 I_2 I_3] - block: 1B

[R_0 G_0 B_0 R_1 G_1 B_1 R_2 G_2 B_2 ...] - block: 3B

[R_0 G_0 B_0 - R_1 G_1 B_1 - R_2 G_2 B_2 - ...] - block: 4B

[R_0 G_0 B_0 A_0 R_1 G_1 B_1 A_1 R_2 G_2 B_2 A_2 ...] - block: 4B

[R_0 G_0 B_0 D_0 R_1 G_1 B_1 D_1 R_2 G_2 B_2 D_2 ...] - block: 4B

[B_0 G_0 R_0 B_1 G_1 R_1 B_2 G_2 R_2 ...] - block: 3B **OpenCV!**

[{ R_{5bit} G_{6bit} B_{5bit} }$_0$, { R_{5bit} G_{6bit} B_{5bit} }$_1$ ...] - block: 2B

Lookup table

RGB

(R, G, B)

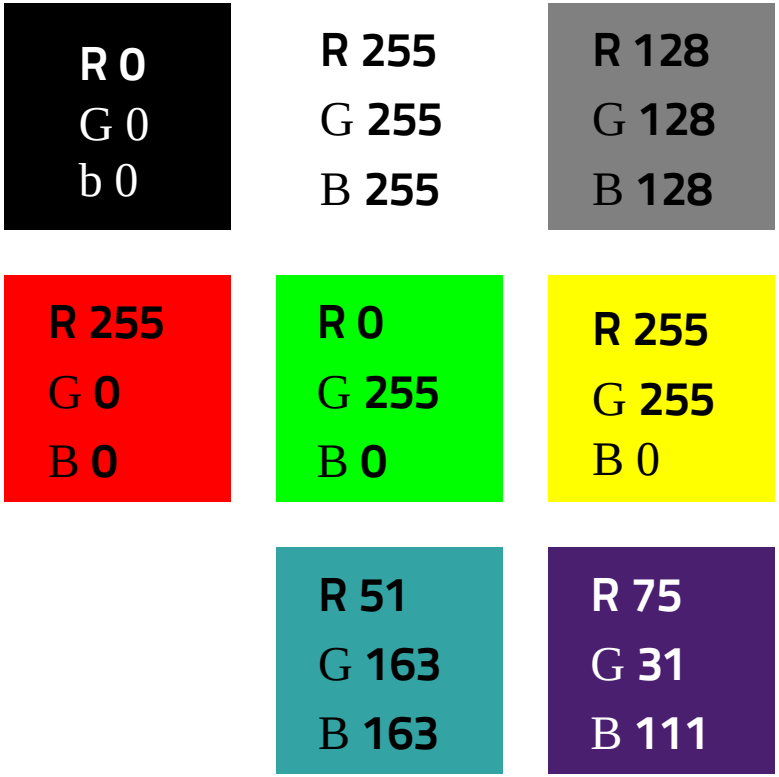
8 bits 0-255

0.0-1.0

0%-100%

#rrggbb

Hexadecimal value



Colorspace conversion

Grey shading

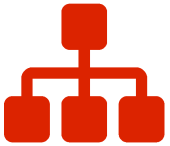
$$I = \mathrm{G_c} \cdot [R \quad G \quad B]$$

Multi-mode images

$$\begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \mathrm{C} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathrm{C^{-1}} \cdot \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix}$$

$$\begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ 1 \end{bmatrix} = \mathrm{C_{4x4}} \cdot \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} \quad \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} = \mathrm{C_{4x4}^{-1}} \cdot \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ 1 \end{bmatrix}$$

Greyscale



Luma (Y) **[0.299 0.587 0.114]** **[0.30 0.59 0.11]**

BT709 **[0.2125 0.7154 0.0721]**

RMI **[0.5 0.419 0.081]**

Intensity **[1/3 1/3 1/3]**

Channel **R/G/B channel**

Value $V = \max(R, G, B)$

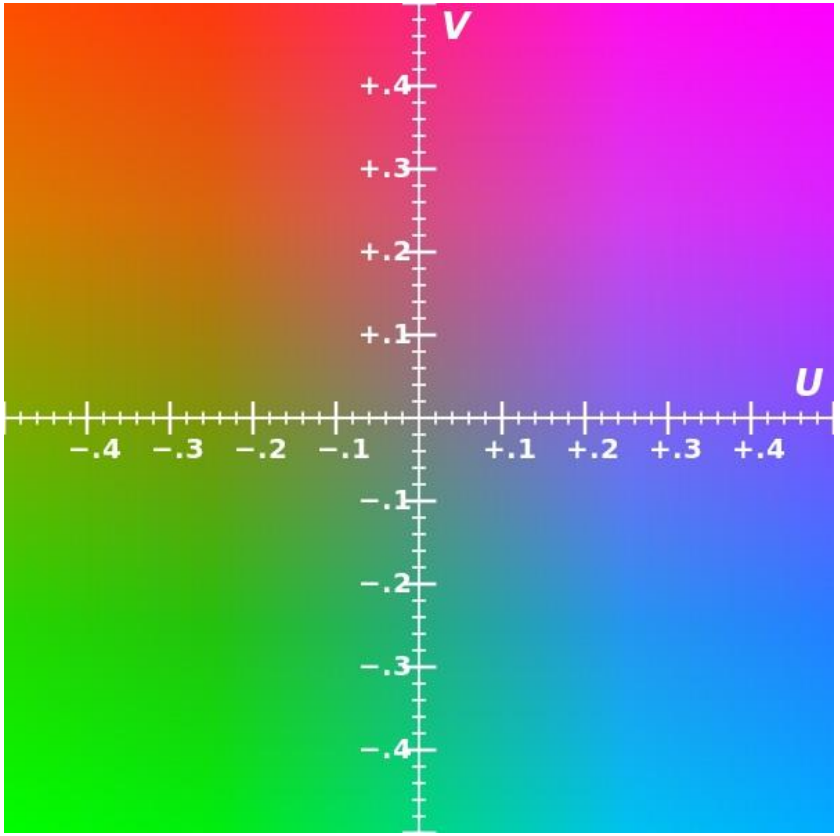
Luminance $L = \frac{\max(R, G, B) + \min(R, G, B)}{2}$

Camera colorspaces

$$C_{YUV} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix}$$

$$C_{YIQ} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix}$$

$$R = Y + 1{,}140 \cdot V$$
$$G = Y - 0{,}394 \cdot U - 0{,}581 \cdot V$$
$$B = Y + 2{,}032 \cdot U$$

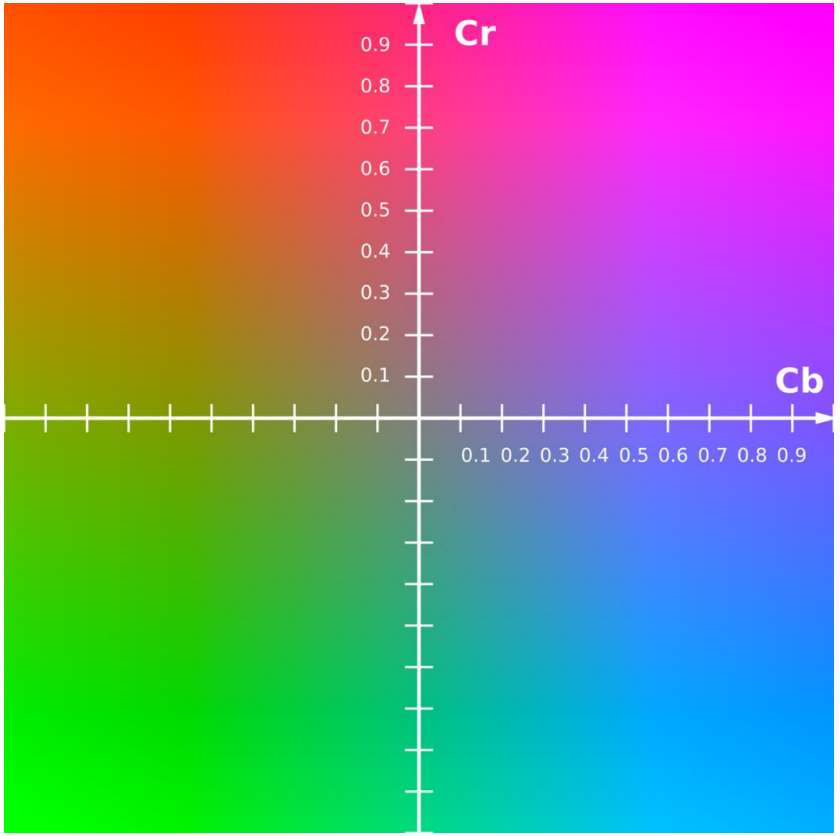


Compression (JPEG)

$$\mathbf{C_{YCbCr}} = \begin{bmatrix} 0.299 & 0.587 & 0.114 & 0 \\ -0.168736 & -0.331264 & 0.5 & 128 \\ 0.5 & -0.418688 & -0.081312 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Y,Cb,Cr: 8 bit positive)

$$R = Y + 1{,}402 \cdot (C_R - 128)$$
$$G = Y - 0{,}34414 \cdot (C_B - 128) - 0{,}71414 \cdot (C_R - 128)$$
$$B = Y + 1.772 \cdot (C_B - 128)$$



HSV

$$V = \max(R, G, B)$$

$$\text{if } V \stackrel{?}{=} 0 \text{ then } H = S = 0; \text{ exit}$$

$$m = \min(R, G, B)$$

$$S = \frac{V - m}{V}$$

$$\text{if } S \stackrel{?}{=} 0 \text{ then } H = 0; \text{ exit}$$

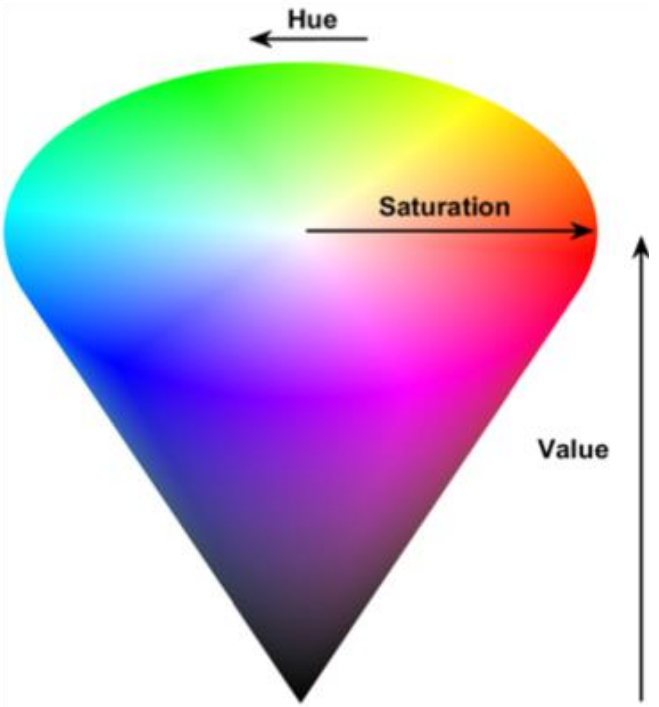
$$c_r = \frac{V - R}{V - m}; c_g = \frac{V - G}{V - m}; c_b = \frac{V - B}{V - m}$$

$$\text{if } R \stackrel{?}{=} V \text{ then } H = 60 \cdot (c_b - c_g) \bmod 360; \text{ exit}$$

$$\text{if } G \stackrel{?}{=} V \text{ then } H = 120 + 60 \cdot (c_r - c_b); \text{ exit}$$

$$\text{if } B \stackrel{?}{=} V \text{ then } H = 240 + 60 \cdot (c_g - c_r); \text{ exit}$$

H 0° S 0.000 V 0.000	H 0° S 0.000 V **1**.000	H 0° S 0.000 V **0**.500
H 0° S **1**.000 V **1**.000	H 120° S **1**.000 V **1**.000	H 60° S **1**.000 V **1**.000



HSL

$$mx = \max(R, G, B); mn = \min(R, G, B)$$

$$L = \frac{mx + mn}{2}$$

$$\text{if } mx \stackrel{?}{=} mn \text{ then } H = S = 0; \text{ exit}$$

$$S = \frac{mx - mn}{1 - |2L - 1|}$$

$$c_r = \frac{G - B}{mx - mn}; c_g = \frac{B - R}{mx - mn}; c_b = \frac{R - G}{mx - mn}$$

$$\text{if } R \stackrel{?}{=} mx \text{ then } H = 60 \cdot c_r \bmod 360; \text{ exit}$$

$$\text{if } G \stackrel{?}{=} mx \text{ then } H = 120 + 60 \cdot c_g; \text{ exit}$$

$$\text{if } B \stackrel{?}{=} mx \text{ then } H = 240 + 60 \cdot c_b; \text{ exit}$$

H 0°
S 0.000
L 0.000

H 0°
S 0.000
L **1**.000

H 0°
S 0.000
L **0**.500

H 0°
S **1**.000
L **0**.500

H 120°
S **1**.000
L **0**.500

H 60°
S **1**.000
L **0**.500

H 180°
S **0**.523
L **0**.420

H 273°
S 0.563
L 0.278

HSI

$$m = \min(R, G, B)$$

$$I = \frac{R + G + B}{3}$$

$$\text{if } I \stackrel{?}{=} 0 \text{ then } H = S = 0; \text{ exit}$$

$$S = 1 - m \cdot \frac{3}{R + G + B}$$

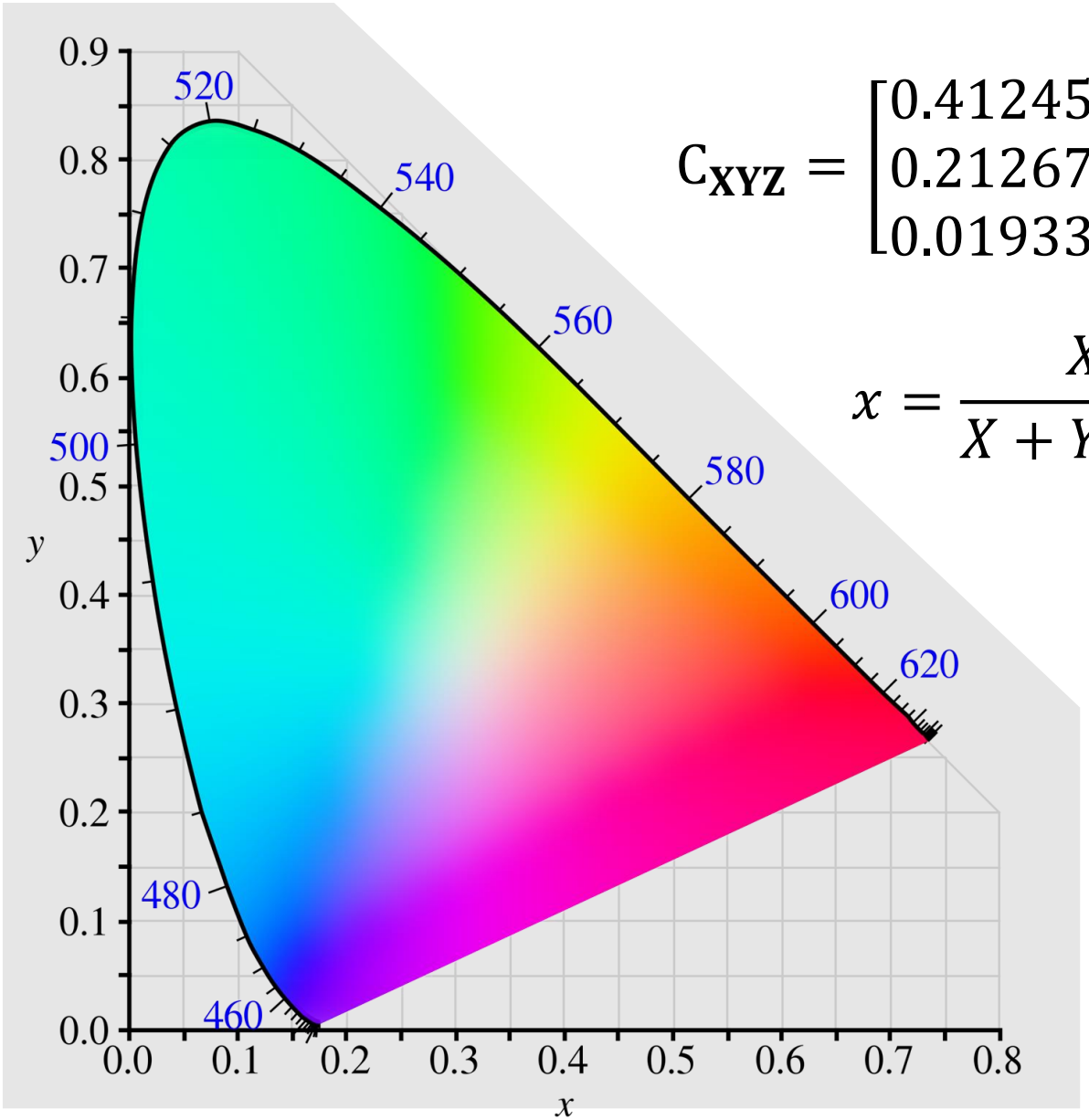
$$\text{if } S \stackrel{?}{=} 0 \text{ then } H = 0; \text{ exit}$$

$$H = \begin{cases} \cos^{-1} z & \text{ha } G \geq R \\ 2\pi - \cos^{-1} z & \text{ha } G < R \end{cases}$$

$$z = \frac{2B - G - R}{2 \cdot \sqrt{R^2 + G^2 + B^2 - RG - RB - GB}}$$

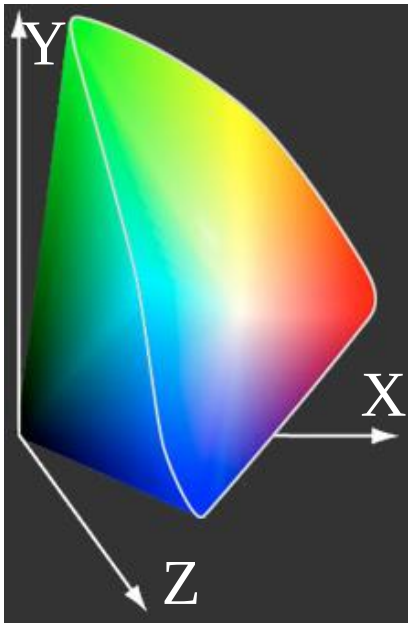
H 0° S 0.000 I 0.000	H 0° S 0.000 I **1**.000	H 0° S 0.000 I **0**.500
H 0° S **1**.000 I **0**.333	H 120° S **1**.000 I **0**.333	H 60° S **1**.000 I **0**.667
	H 180° S **0**.594 I **0**.494	H 273° S **0**.571 I **0**.282

CIE XYZ

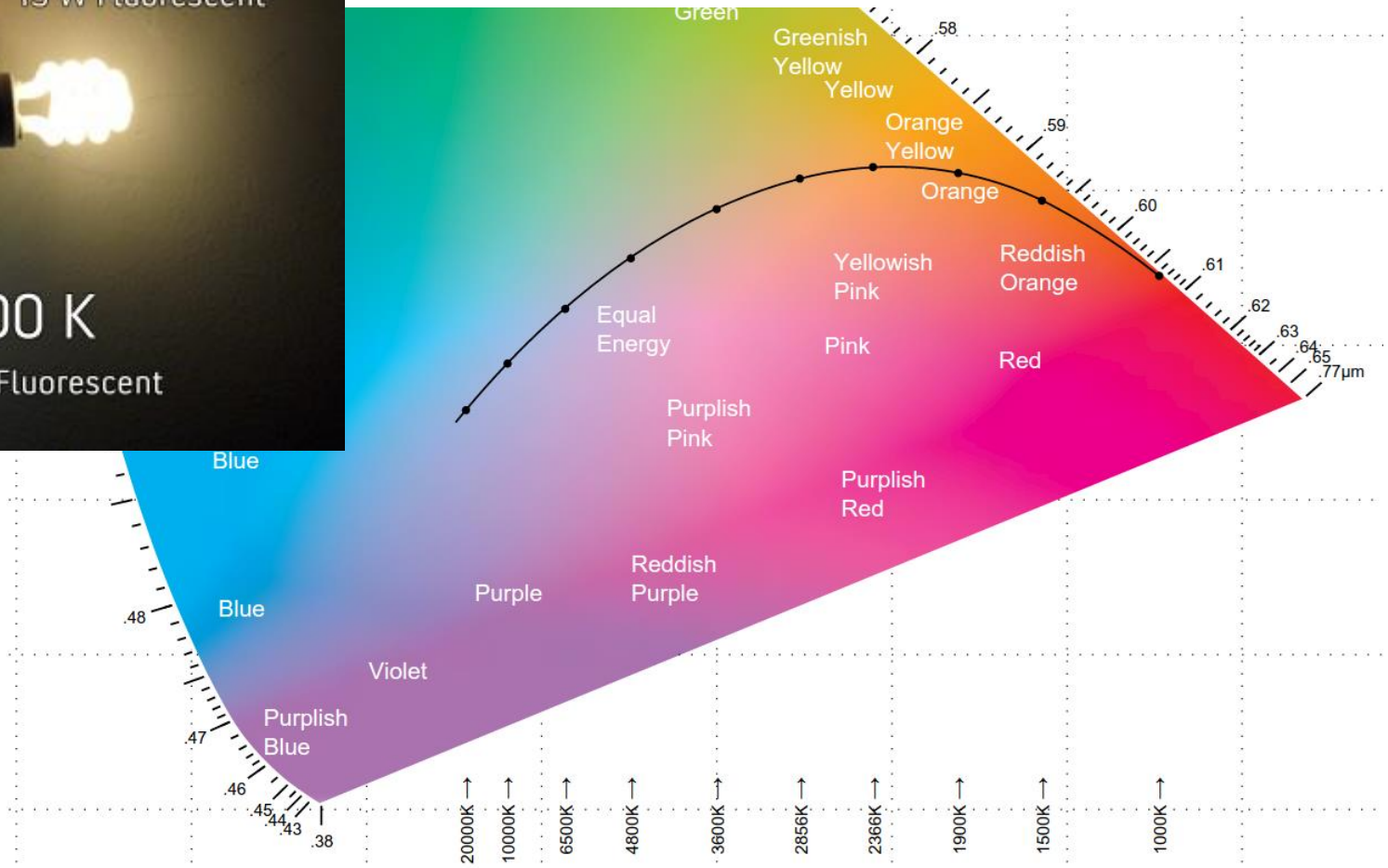


$$C_{XYZ} = \begin{bmatrix} 0.412453 & 0.35758 & 0.180423 \\ 0.212671 & 0.71516 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix}$$

$$x = \frac{X}{X+Y+Z};\ y = \frac{Y}{X+Y+Z}$$



Colour temperature



Use of the space