

# Számítógépes

# Látás

# 01

**Számítógépes látórendszerek**

**Dr. Szemenyei Márton**

**Adjunktus**

**2022**



# Előadó

---

Dr. Szemenyei Márton

Adjunktus

**[szemenyei@iit.bme.hu](mailto:szemenyei@iit.bme.hu)**

IB411

(463) 4025



# Tantárgyinfók

---

## Tárgy központi oldala: Teams & Moodle

Diák

Jegyzet

Kérdések

Adatlap

<https://portal.vik.bme.hu/kepzes/targyak/VIIIMA07>

Javasolt irodalom

Online közzétett jegyzet (html és pdf)



# Számonkérések

---

ZH

Április 11. (hétfő) 18-20, jelenléti (feltehetően)

Igaz-hamis, sorbarendezés, párosítás, rajz, felsorolás

40% (32/80p) kell az aláíráshoz

Pót: Április 27. (szerda) 18-20, jelenléti (feltehetően)

**Pótpót csak annak van, aki legalább egyszer eljött ZH-ra!!!**

Vizsga

80p + 20p a ZH-ról

40% (40/100p) kell (15 pontonként 1 jegy)



# Házi feladat

## Önvezető autós feladatok

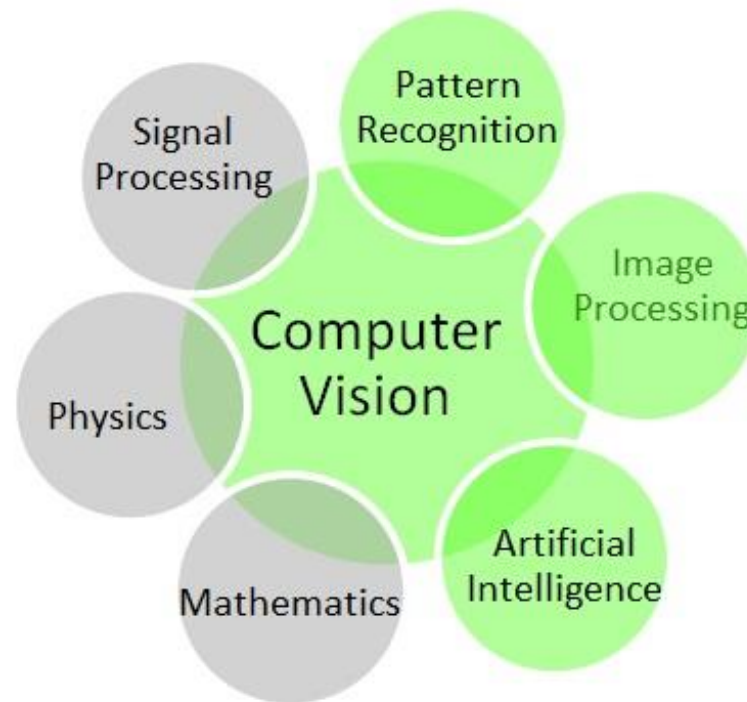
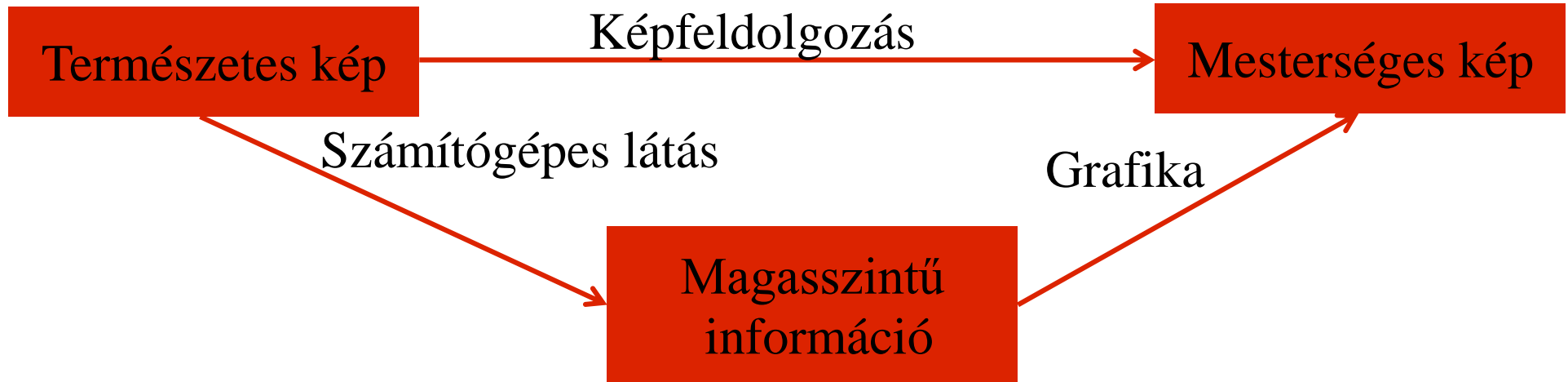
Szegmentálás, Objektumdetektálás, 3D követés

### Fontos paraméterek:

- Opcionális
- 4 fős csapatok
- Megajánlott jegy: Kiváltja a ZH-t és a Vizsgát
- Bemutató pótlási héten
- Jelentkezni 2. hét vasárnapig lehet (teams/email, csapattagok neve, Neptun kód)
- Házira való jelentkezés automatikusan pótpótra jogosít



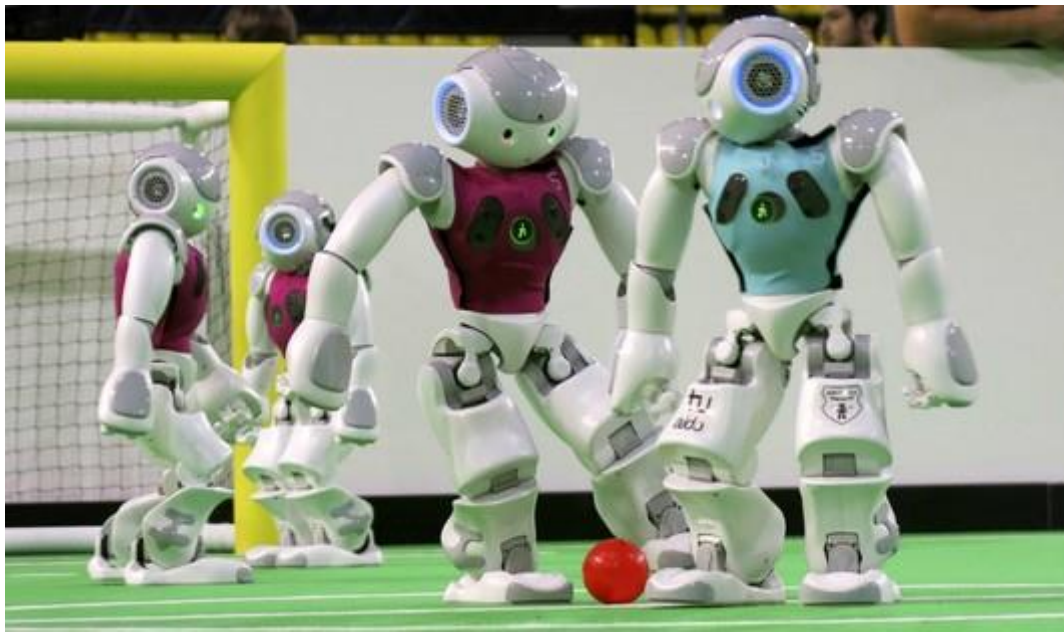
# Számítógépes látás



# Számítógépes látás



# Számítógépes látás





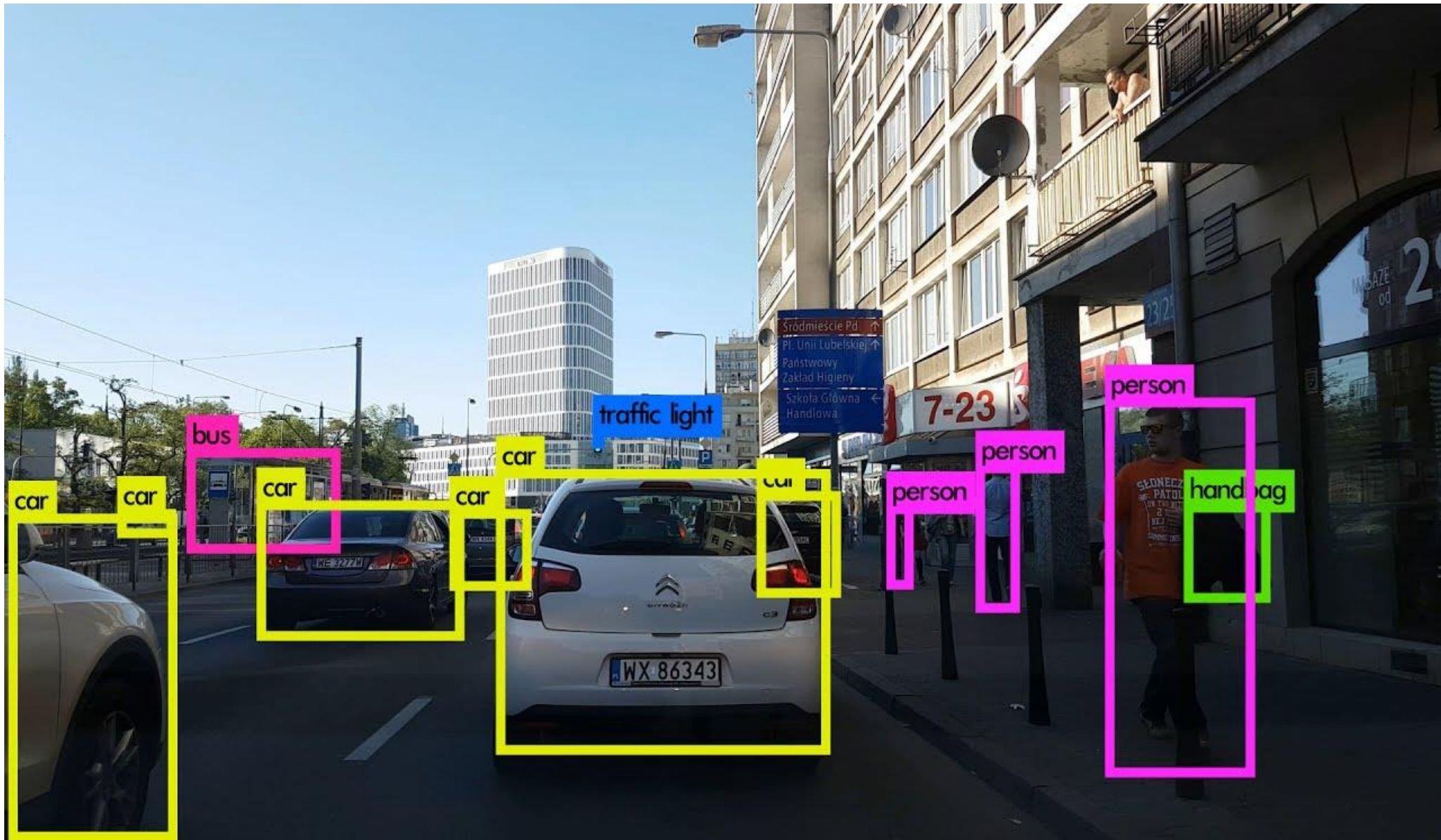
# Osztályozás

---

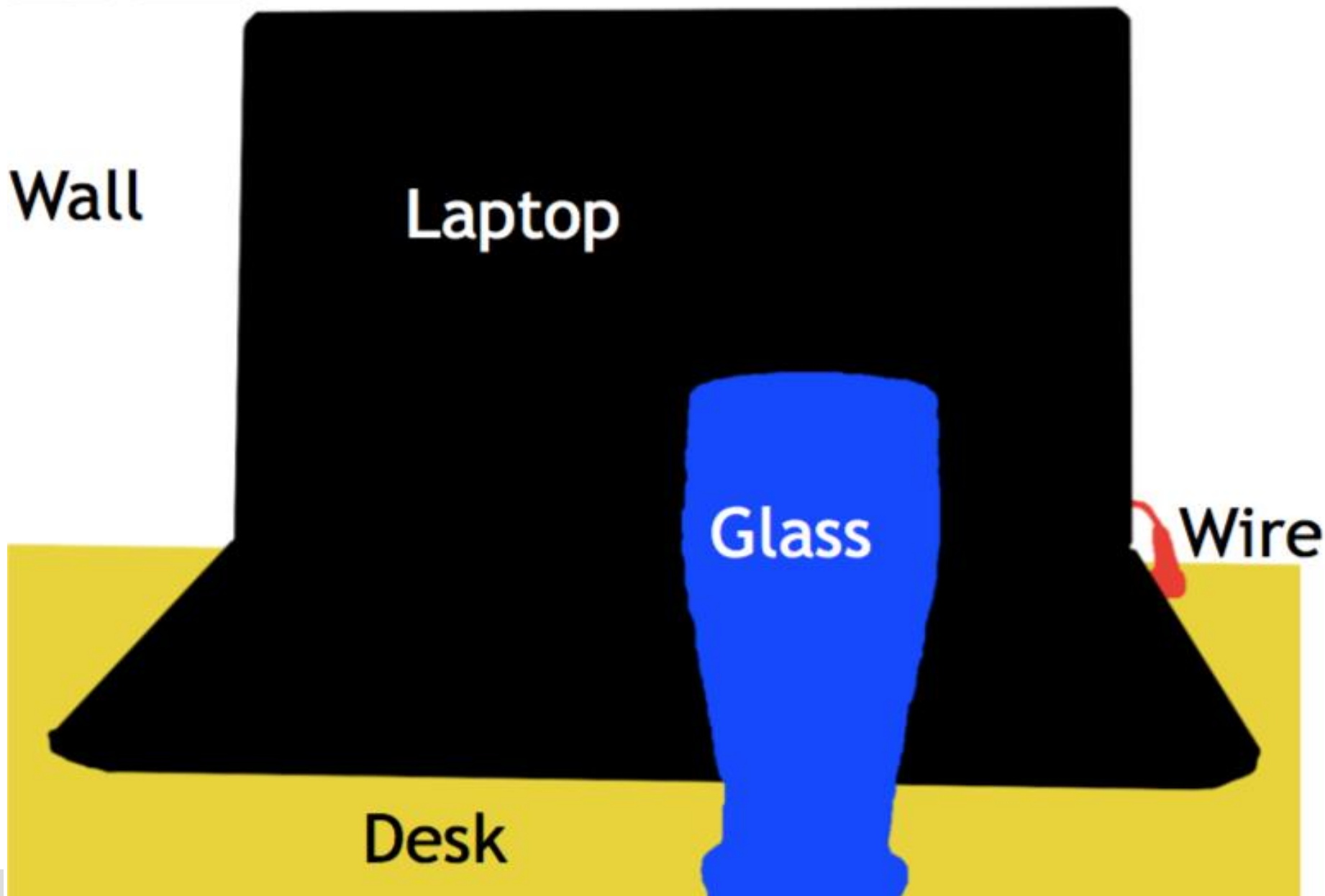


→ Macska

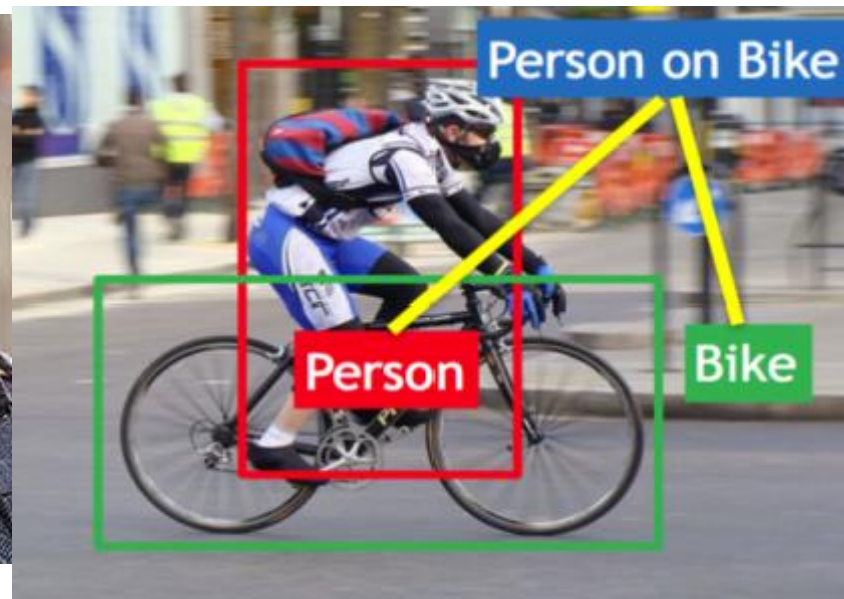
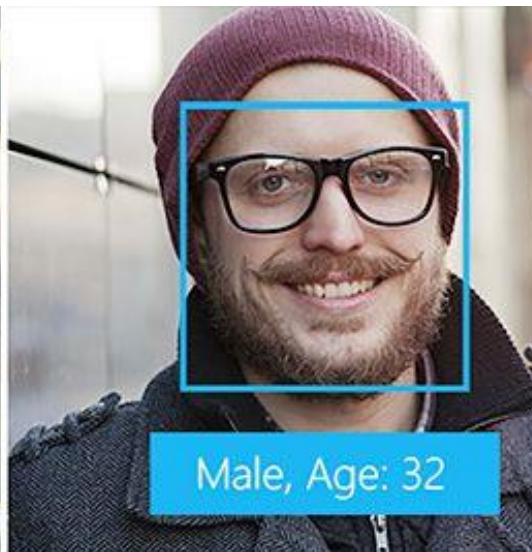
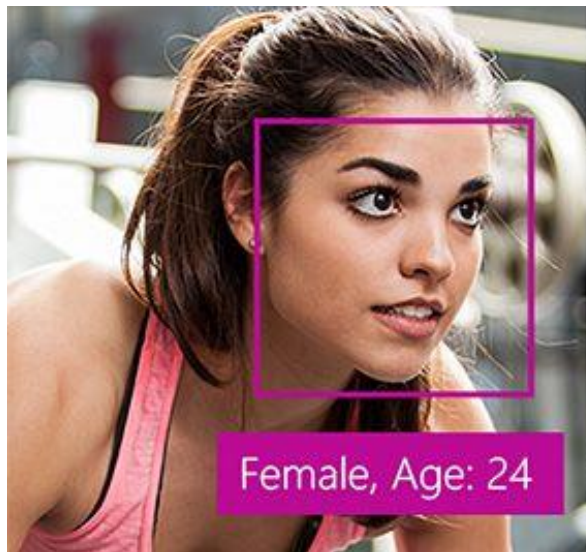
# Detekció



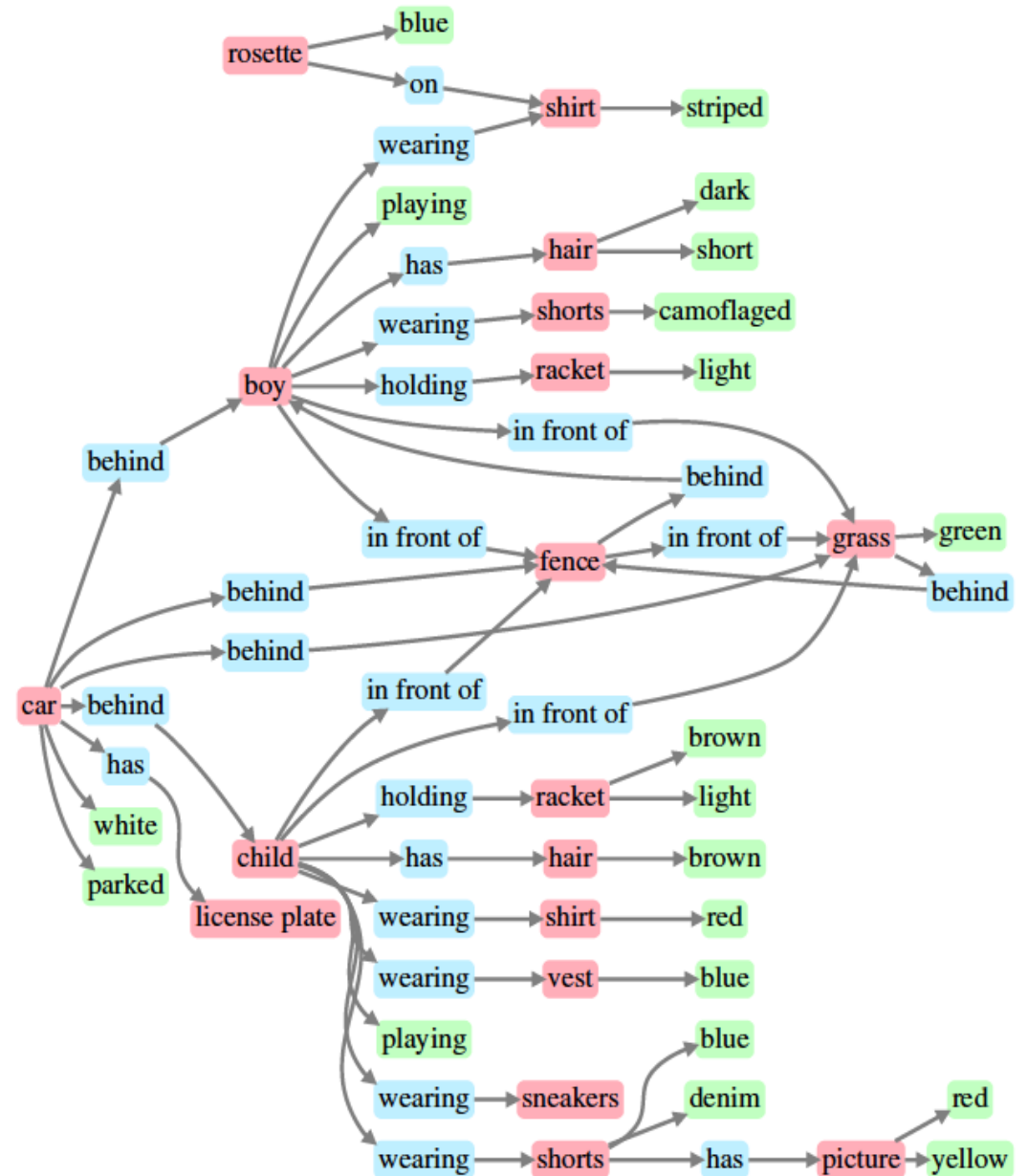
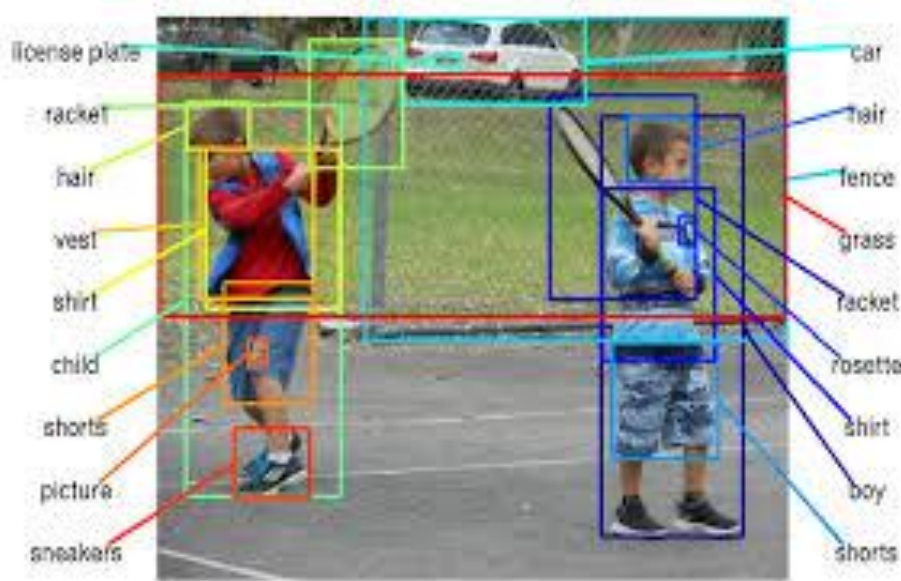
# Szegmentáció



# Számítógépes Látás



# Értelmezés



# Értelmezés

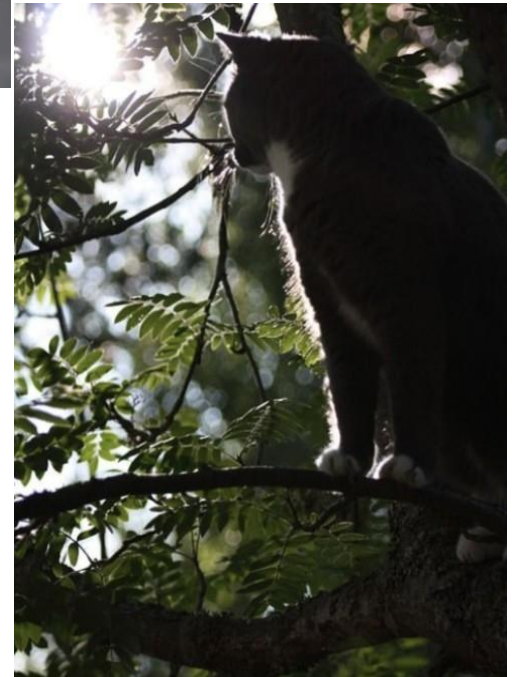


# Probléma: Szemantikus Gát



```
[[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]]
```

# Nehézség: Megvilágítás





# Nehézség: Deformáció



# Nehézség: Takarás



# Nehézség: Osztályon belüli variáció





# Tárgy célja

---

Hogy hívjuk meg a függvényt? → „Google”

**Összefüggések**

**Miérték**

**Hogyanok**

**Mitek**

**Melyiketek**



**Cél: Látás**



# Ütemezés terve

---

1-7. előadás + 1 gyakorlat

Hagyományos számítógépes látás

8-11. előadás + 1 gyakorlat

Intelligens látórendszerek

**Zárthelyi dolgozat, április 11.**

12-14. előadás + 1 gyakorlat

3D látás

15-16. előadás + 1 gyakorlat

Valós idejű látás eszközei

# Ütemezés terve



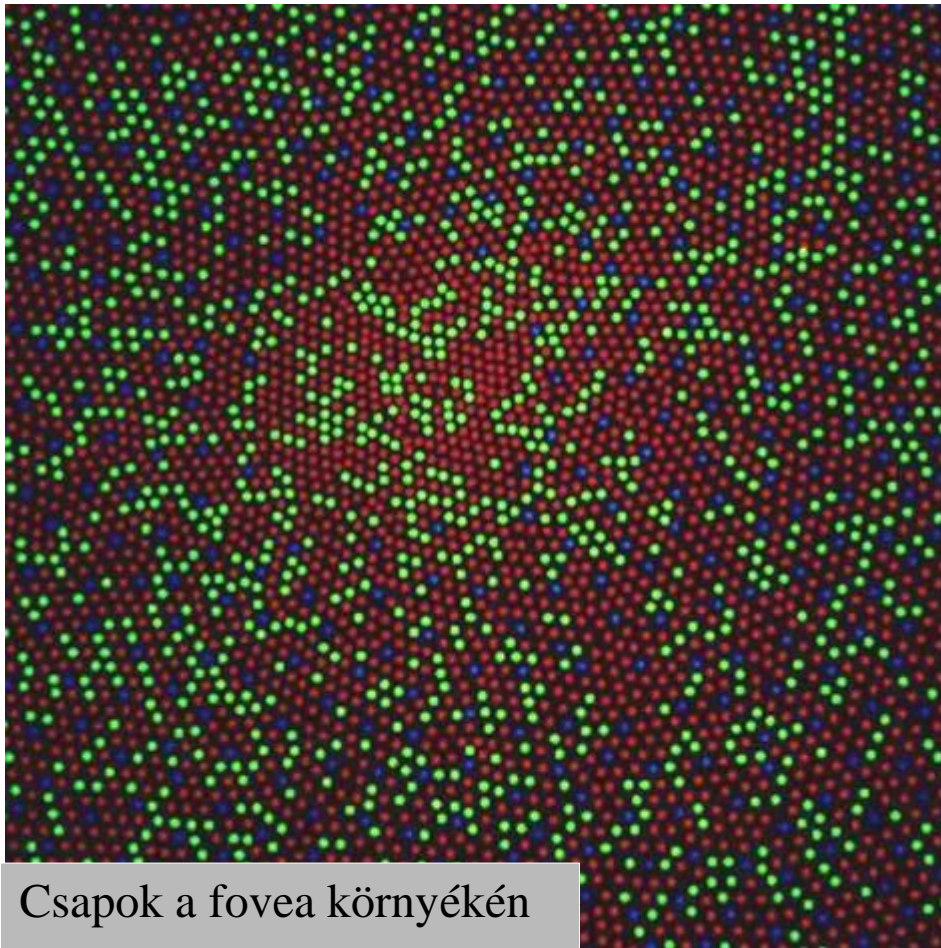
Hét	Kedd	Csütörtök
1	Bevezetés	
2	Képjavítás	Frekvenciatartomány
3	Képjellemzők	
4	Detektálás, Követés	Szegmentálás
5	Bináris Képek	
6	Neurális Hálózatok	Hagyományos Látás
7	Konvolúciós Neurális Hálók	
8	Deep Learning a Gyakorlatban	Magas szintű Látás
9	Deep Learning	
10	Tavaszi szünet	Kameramodell és Kalibráció
11	3D Rekonstrukció	
12	3D Feldolgozás	3D Látás
13	Hardverek	
14	TPU, FPGA	CUDA

# Hagyományos látás

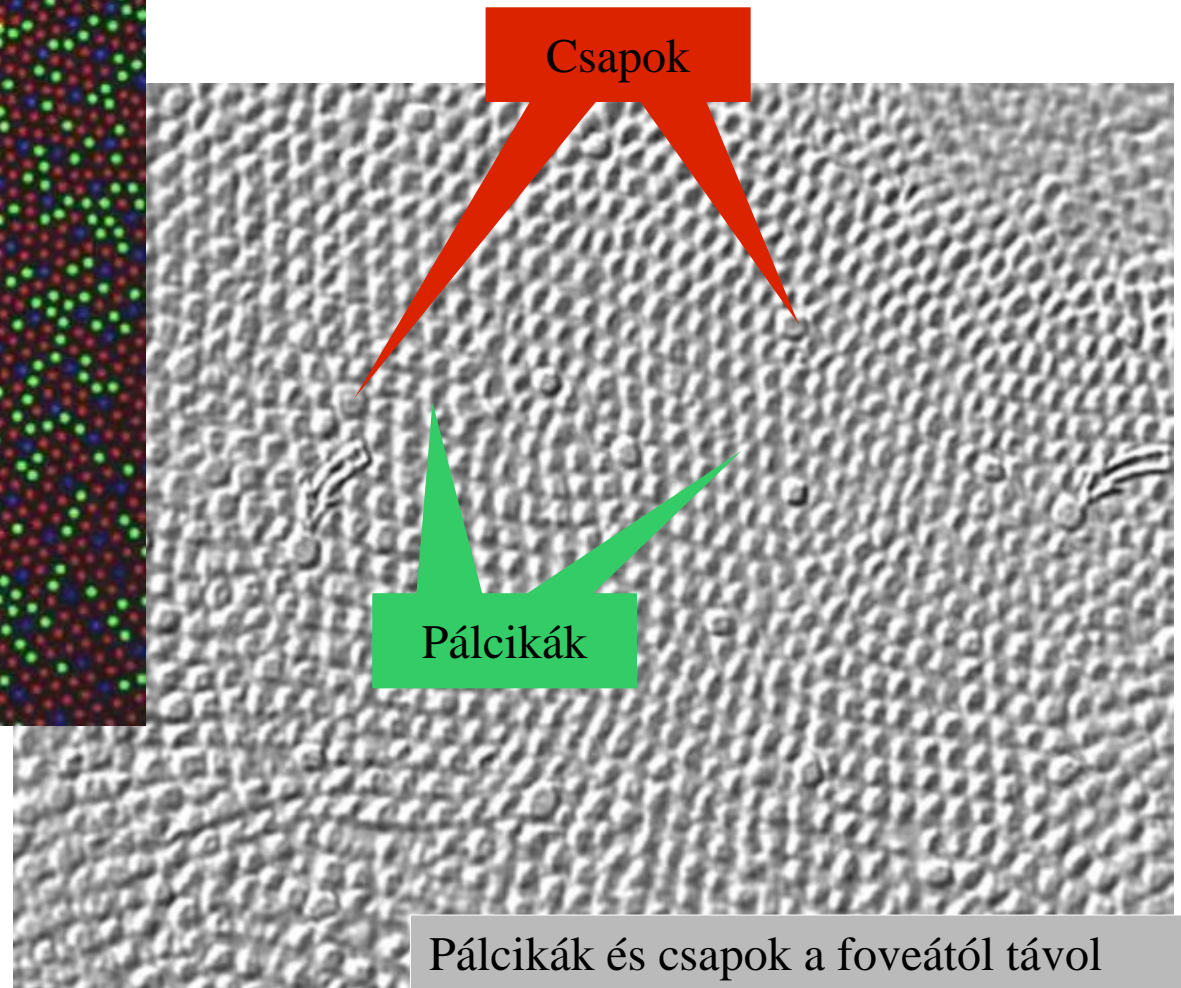
---



# Kép, ahogy az ember látja



Csapok a fovea környékén



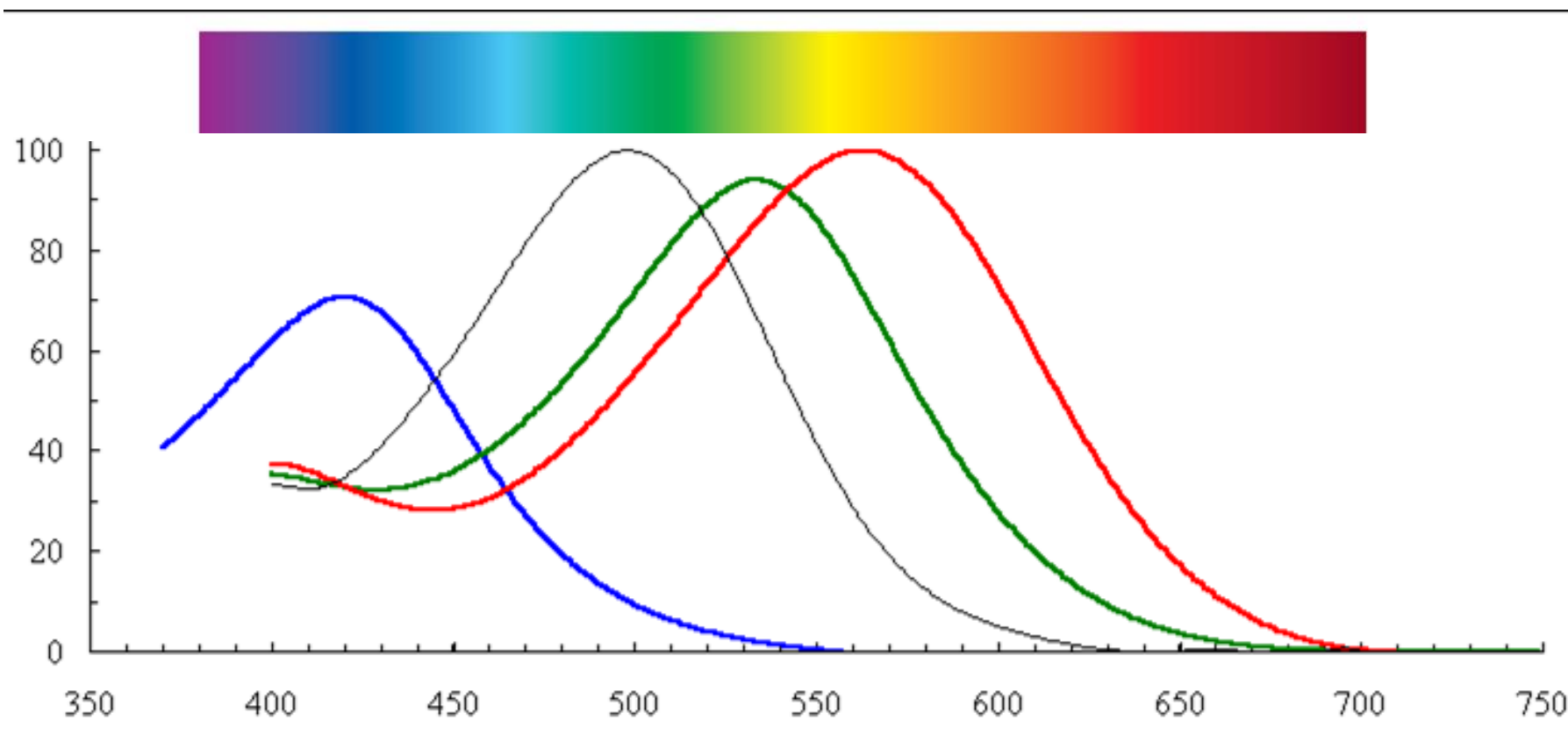
Csapok

Pálcikák

Pálcikák és csapok a foveától távol



# Emberi színlátás





# Fotodióda

**Minden dióda érzékelő**

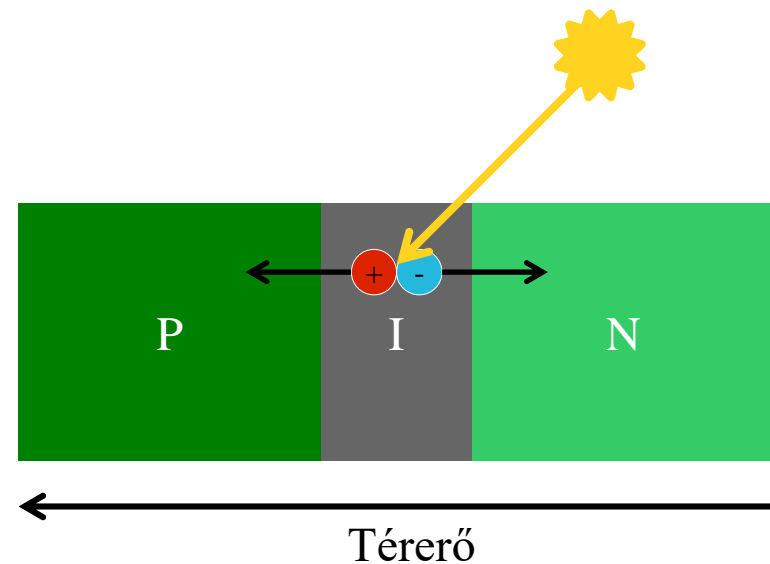
Fotovoltaikus mód

Fotovezető mód

Gyorsabb reakció

Nagyobb zaj

Fototranzisztor

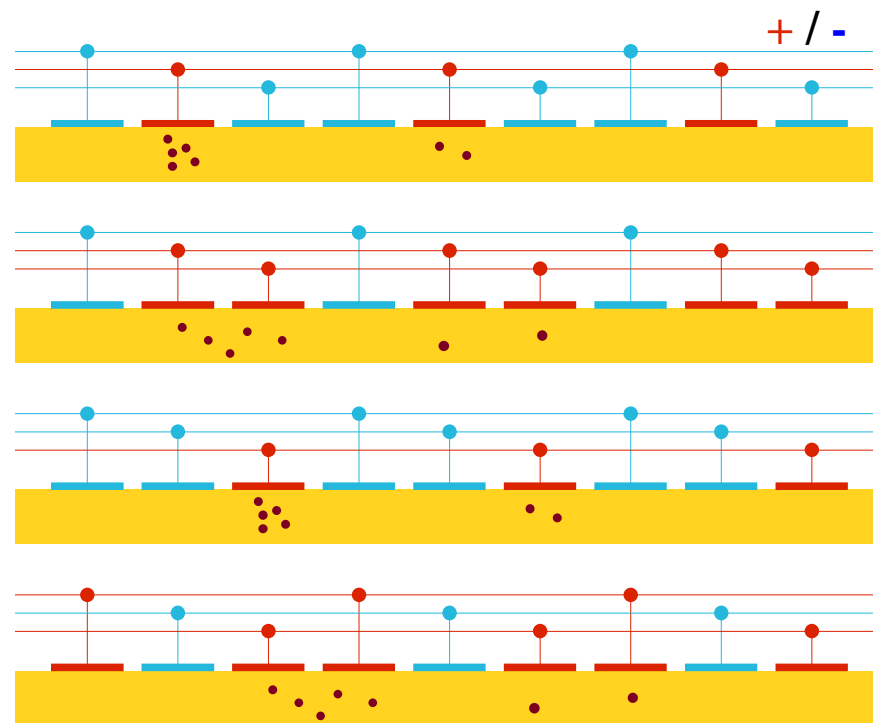
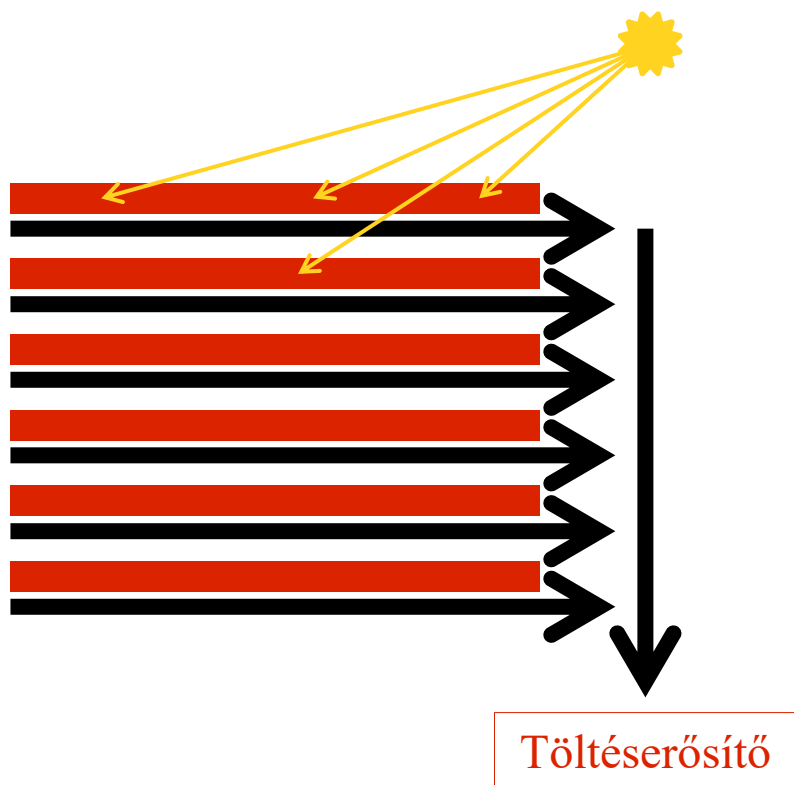


# CCD



## Töltéscsatolt eszközök (Charge-Coupled Device)

### Analóg léptető regiszterek



# CCD típusok

---



## Full frame CCD

Nincs pixelenkénti tárolás  
Külső zár szükséges

## Frame-transfer CCD

CCD a fotoérzékelők mögött

## Interline CCD

CCD a fotoaktív oszlopok között (rosszabb felbontás)

# CMOS



## Előnyök és hátrányok

Kisebb fogyasztás

Kisebb késleltetés

Olcsóbb kivitel

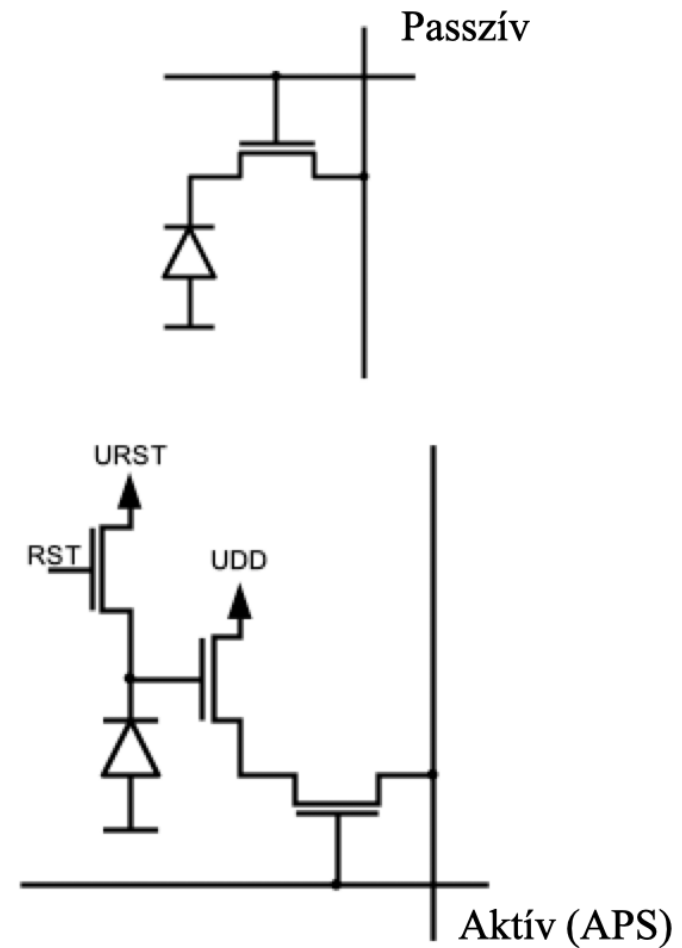
Nagyobb sebesség

Nagyobb zaj

Kisebb érzékenység

Pixelszintű előfeldolgozás

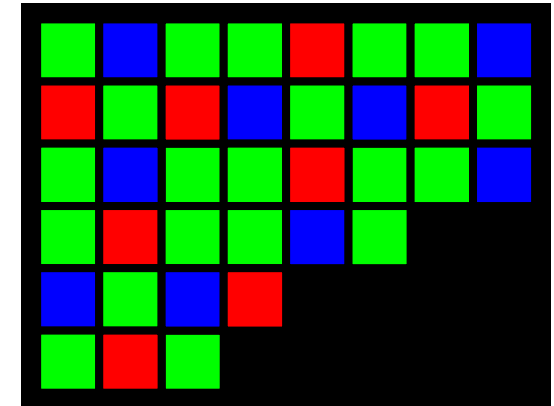
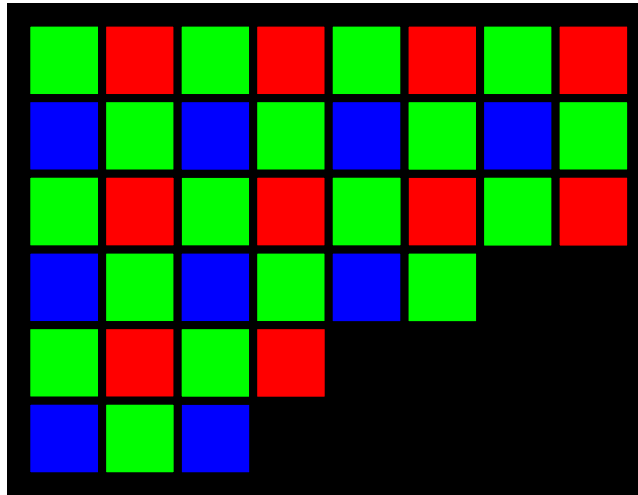
CNN, AER...



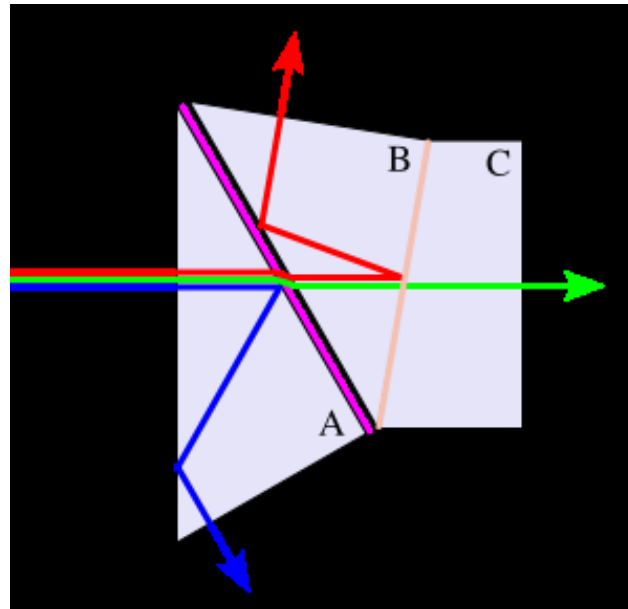
# Színes kamerák



Bayer



3CCD



Vonal CCD





# Kameratípusok

Érzékenységi tartomány

Látható tartomány

Infra

Térbeliség

Hagyományos (2D)

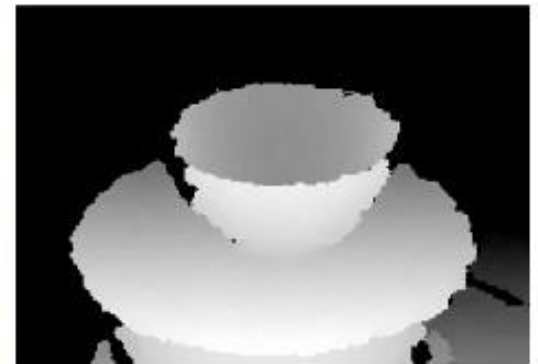
Vonalkamera (1D)

Mélység szenzorok (3D)

Sztereó kamera

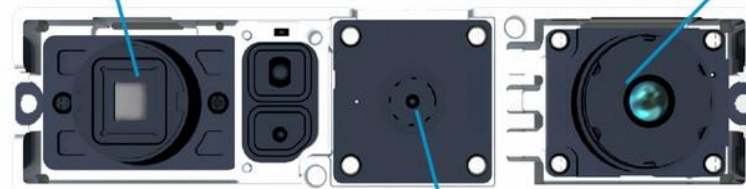
Strukturált fény

LIDAR



Projector

IR Sensor

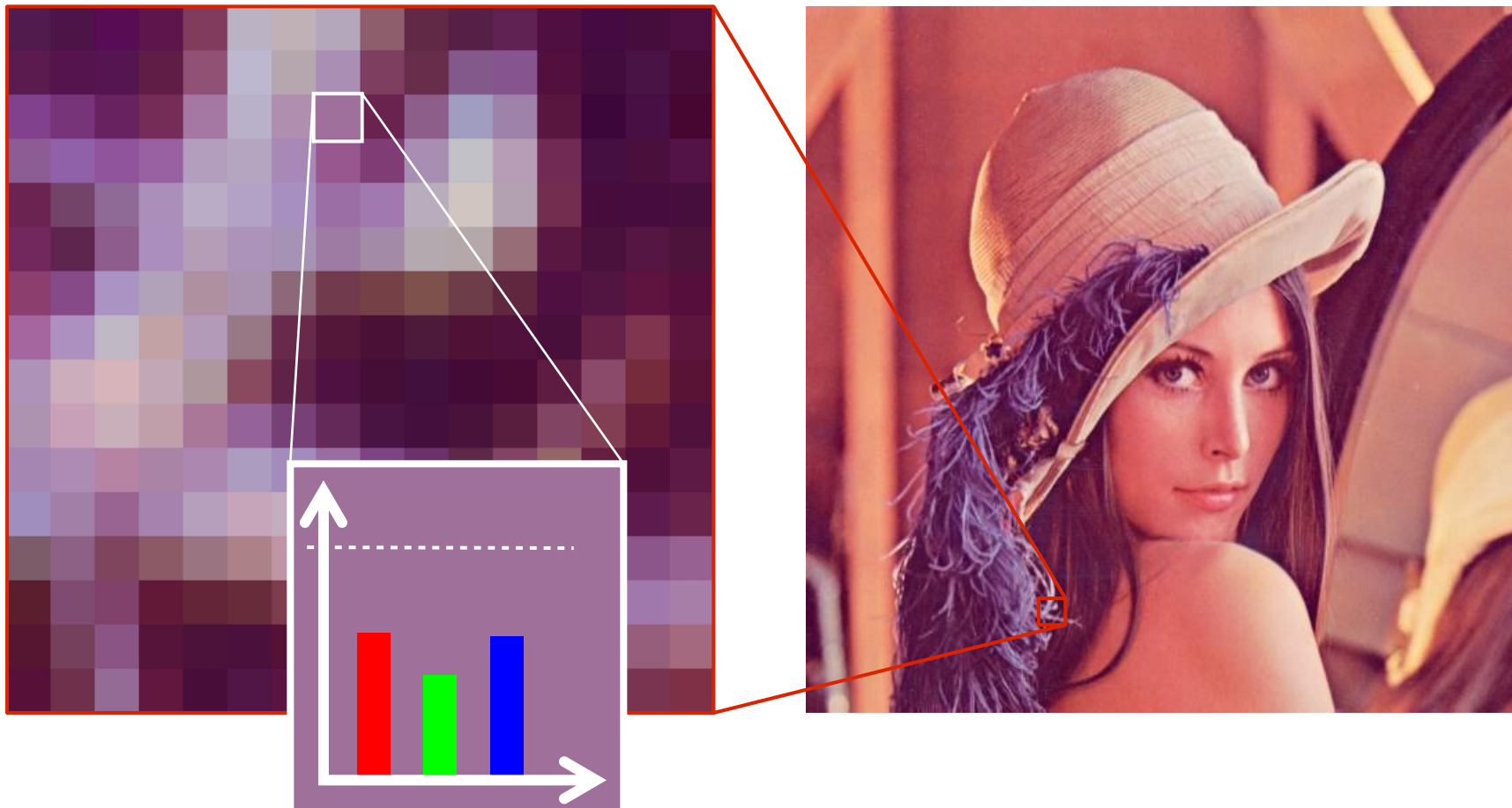


RGB Sensor

# Kép

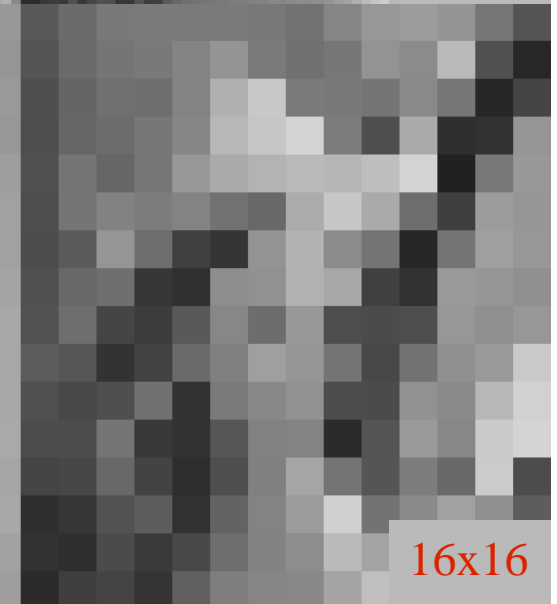
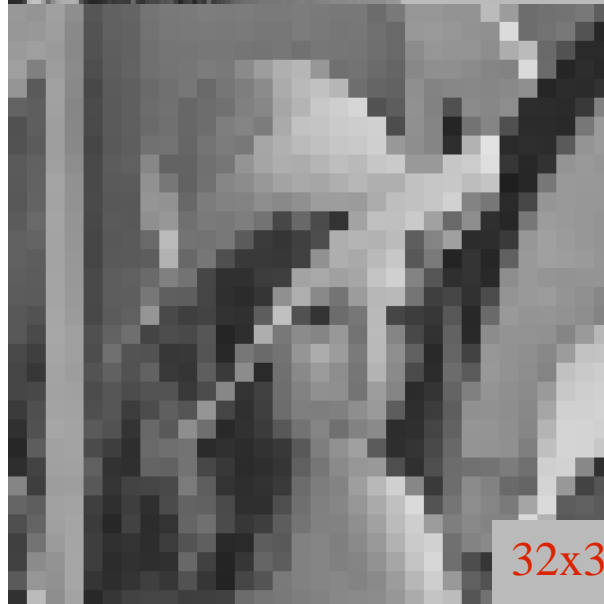


Tipikusan: 1D/2D/3D tömb (pixel-/voxeltömb)





# Felbontás



# Bitmélység



# 1Bites kép, Dithering



# Digitális zaj



8x-os átlagolás

# Videofolyam



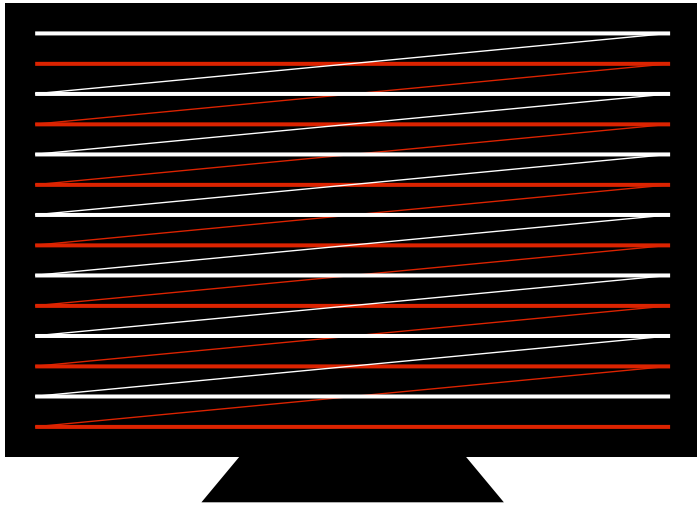
Tipikusan: Képek sorozata (DE...)



# Interlace



Cél: Időbeli felbontás duplázása  
TV szabványok!



1. keret

2. keret

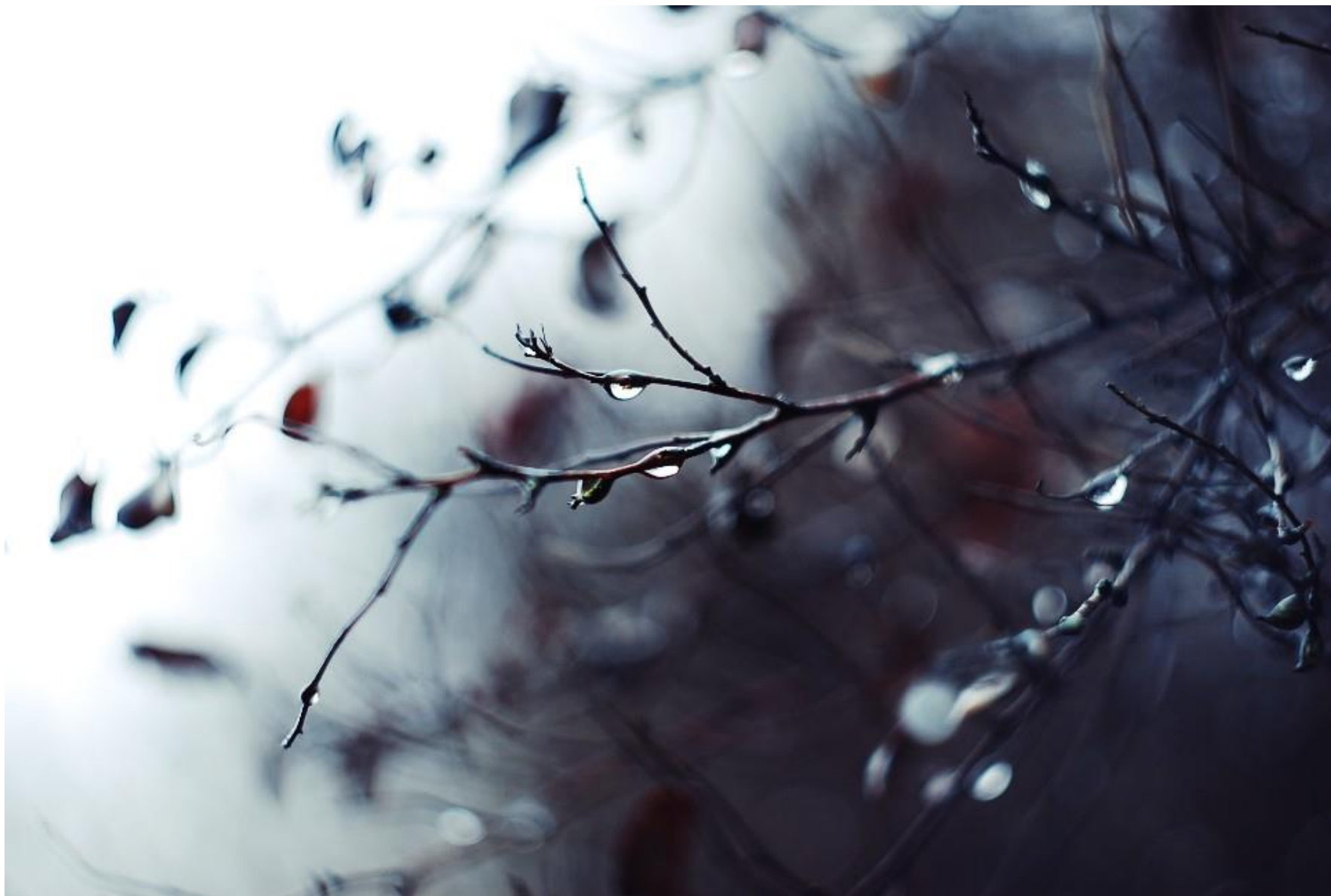


# Mozgás



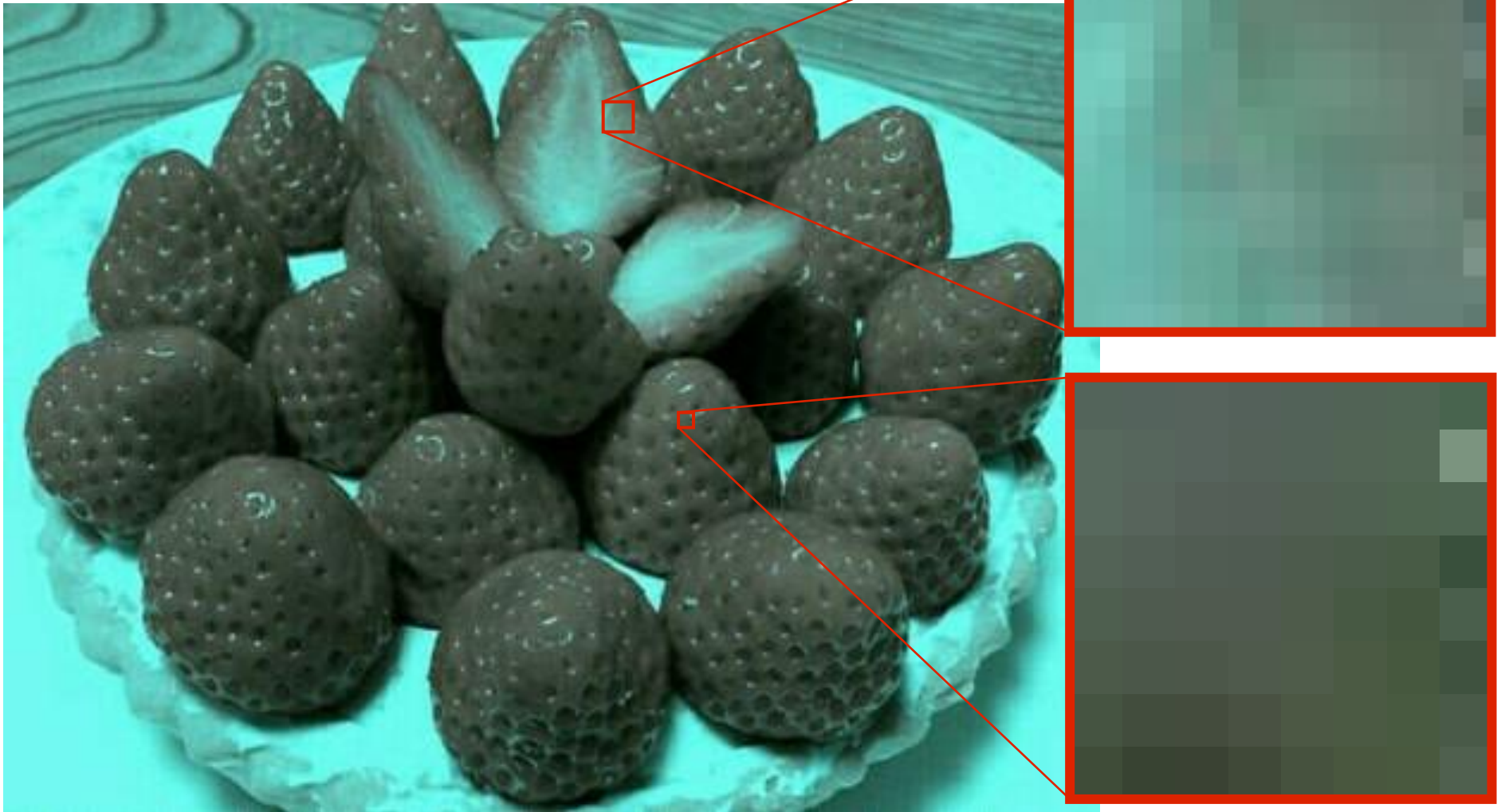
# Fókusz

---





# Érzékelt szín



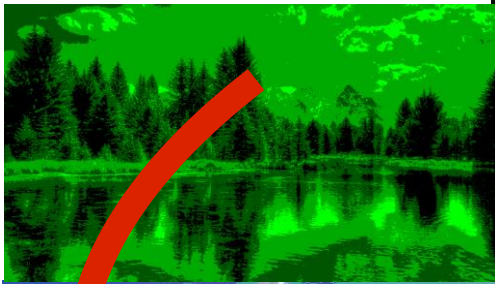
# Színkomponensek szerepe



# Színterminológusok szerepe



Zöld komponens 2 bit



Kék komponens 2 bit





# Tárolás

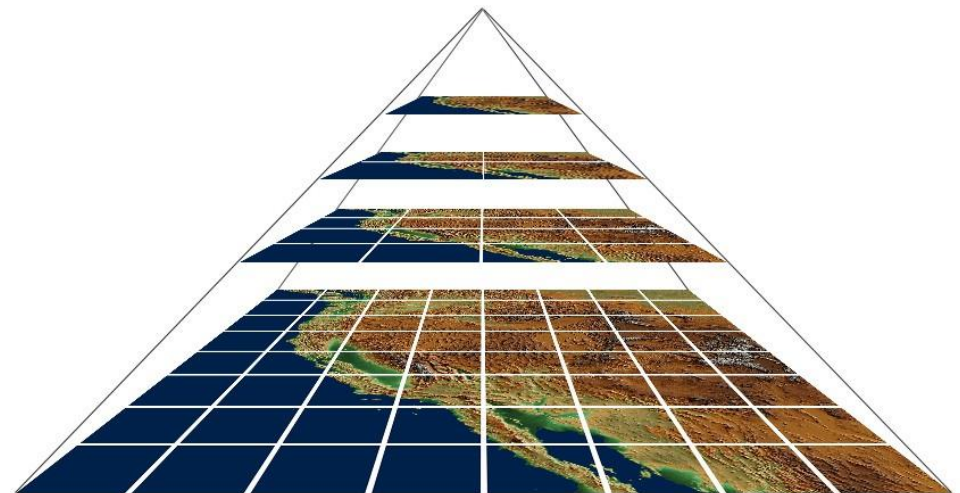
Kép (3648 x 2736 – 10MPx, színes): 30MB

Video (HD1080, 30fps, 1 óra): 672GB (1,5Gbit/sec)

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

Tömörítés: **emberi látás!**

HDD: 100-300 MB/sec



# Képtartalom



[  $I_0 I_1 I_2 I_3$  ] – block: 1B

[  $R_0 G_0 B_0 R_1 G_1 B_1 R_2 G_2 B_2 \dots$  ] – block: 3B

[  $R_0 G_0 B_0 - R_1 G_1 B_1 - R_2 G_2 B_2 - \dots$  ] – block: 4B

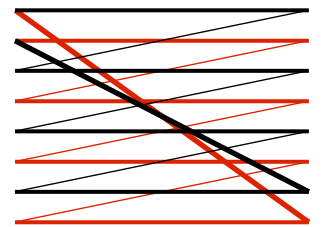
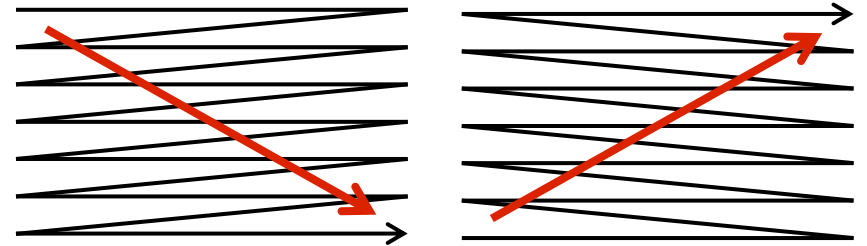
[  $R_0 G_0 B_0 A_0 R_1 G_1 B_1 A_1 R_2 G_2 B_2 A_2 \dots$  ] – block: 4B

[  $R_0 G_0 B_0 D_0 R_1 G_1 B_1 D_1 R_2 G_2 B_2 D_2 \dots$  ] – block: 4B

[  $B_0 G_0 R_0 B_1 G_1 R_1 B_2 G_2 R_2 \dots$  ] – block: 3B **OpenCV!**

[ {  $R_{5bit} G_{6bit} B_{5bit}$  }<sub>0</sub>, {  $R_{5bit} G_{6bit} B_{5bit}$  }<sub>1</sub>... ] – block: 2B

Kereső tábla (lookup table)





# Formátumok

---

Tömörítetlen:

BMP, PNM

Veszteség nélkül:

GIF, JPEG-LS/2000, PNG, TIFF, *H.264 lossless...*

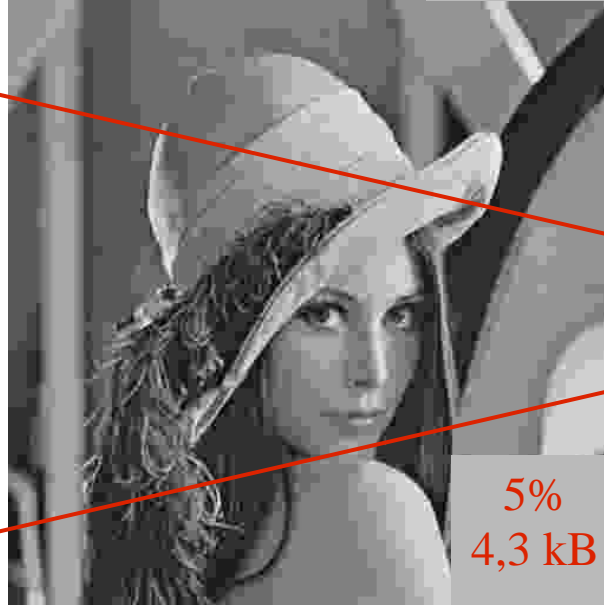
Veszteséges:

JPEG, PNG, *H.264...*

Konténer formátumok:

(PNG), TIFF, AVI, MOV, MP4, OGG, WMV...

# JPEG tömörítés



# RGB



(R, G, B)

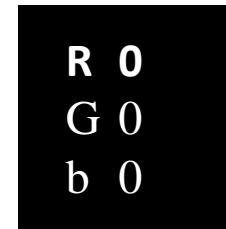
8 biten 0-255

0.0-1.0

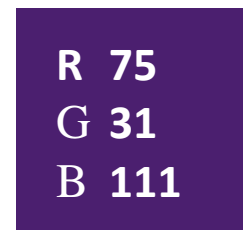
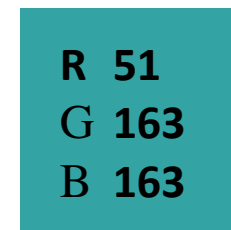
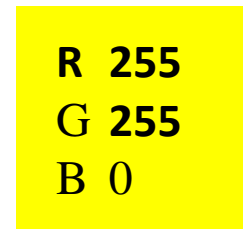
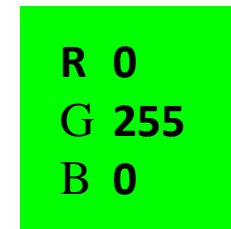
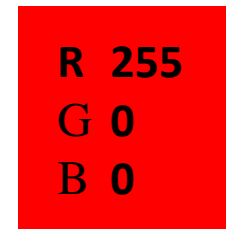
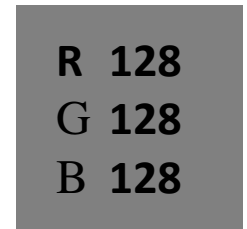
0%-100%

#rrggbb

Hexadecimális érték

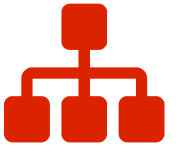


R 255  
G 255  
B 255





# Színtérkonverzió



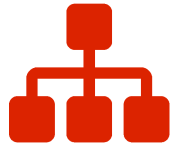
Szürkeárnyalatosítás

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

Többszínű képek

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

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



# Szürkeárnyalatosítás

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** ]

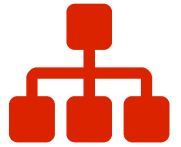
Intenzitás [ **1/3 1/3 1/3** ]

Csatorna **R/G/B** csatorna

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

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

# Kamera színterek



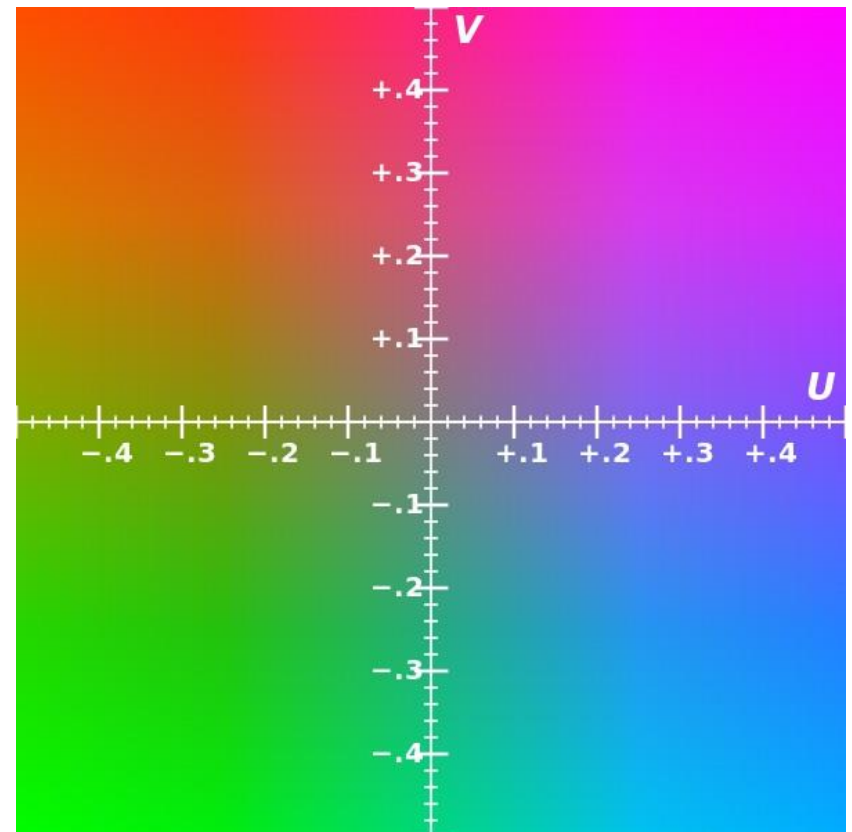
$$C_{YUV} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix}$$

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

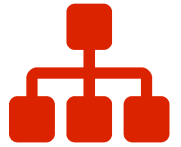
$$R = Y + 1,140 \cdot V$$

$$G = Y - 0,394 \cdot U - 0,581 \cdot V$$

$$B = Y + 2,032 \cdot U$$



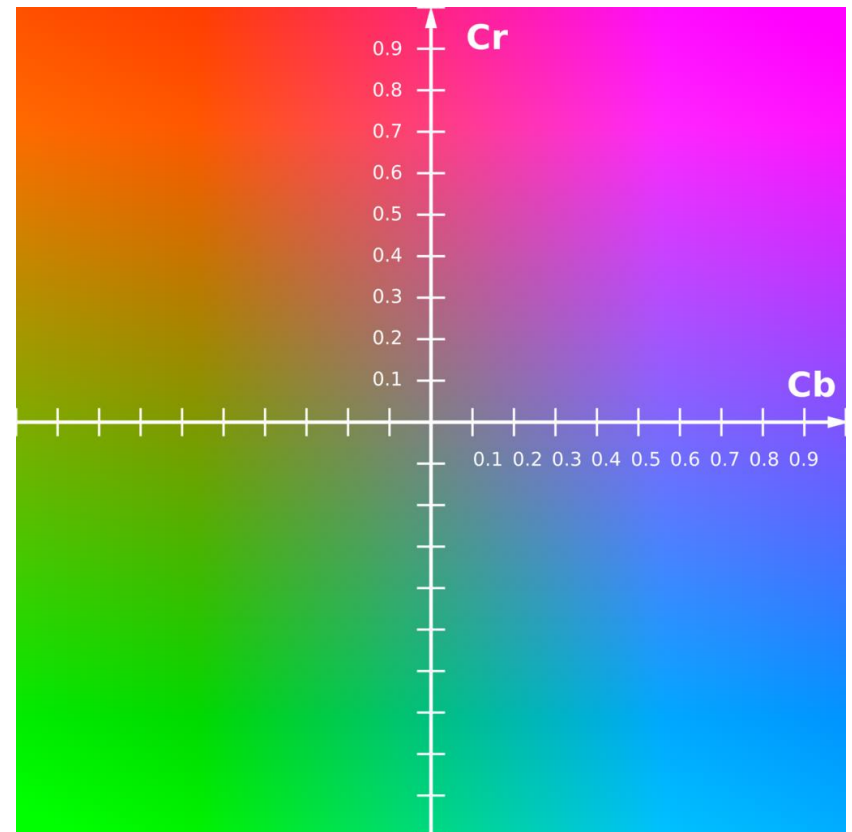
# Tömörítés (JPEG)



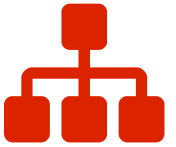
$$C_{YCbCr} = \begin{bmatrix} 0.299 & 0.587 & 0.114 & 0 \\ -0.168736 & -0.331264 & 0.5 & 128 \\ 0.5 & -0.418688 & -0.081312 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

*(Y,Cb,Cr: 8 bit pozitív)*

$$\begin{aligned} R &= Y + 1,402 \cdot (C_R - 128) \\ G &= Y - 0,34414 \cdot (C_B - 128) \\ &\quad - 0,71414 \cdot (C_R - 128) \\ B &= Y + 1,772 \cdot (C_B - 128) \end{aligned}$$



# HSV



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

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

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

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

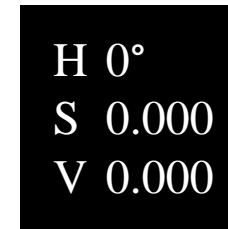
if  $S \stackrel{?}{=} 0$  then  $H = 0$ ; exit

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

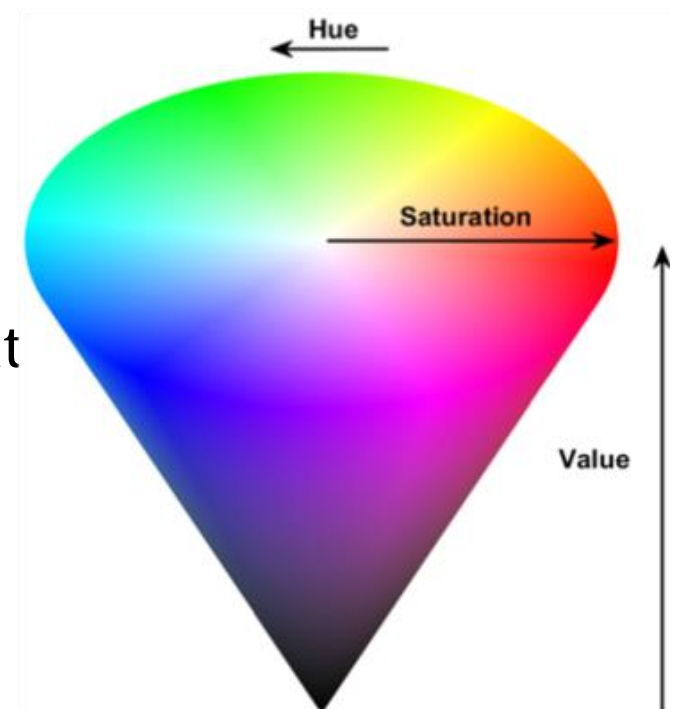
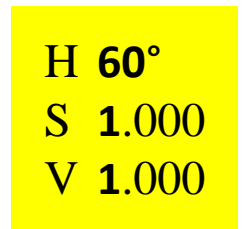
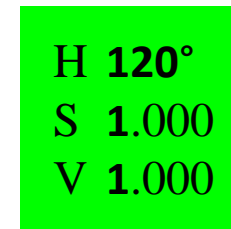
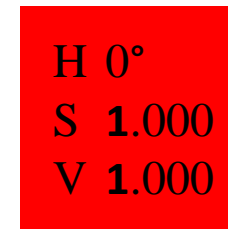
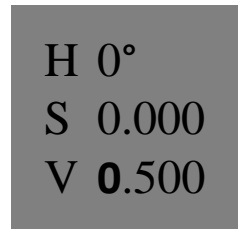
if  $R \stackrel{?}{=} V$  then  $H = 60 \cdot (c_b - c_g) \bmod 360$ ; exit

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

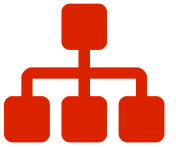
if  $B \stackrel{?}{=} V$  then  $H = 240 + 60 \cdot (c_g - c_r)$ ; exit



H 0°  
S 0.000  
V 1.000



# HSL



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

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

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

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

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

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

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

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

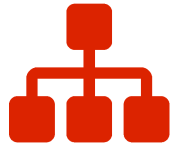
H 120°  
S 1.000  
L 0.500

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



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

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

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

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

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

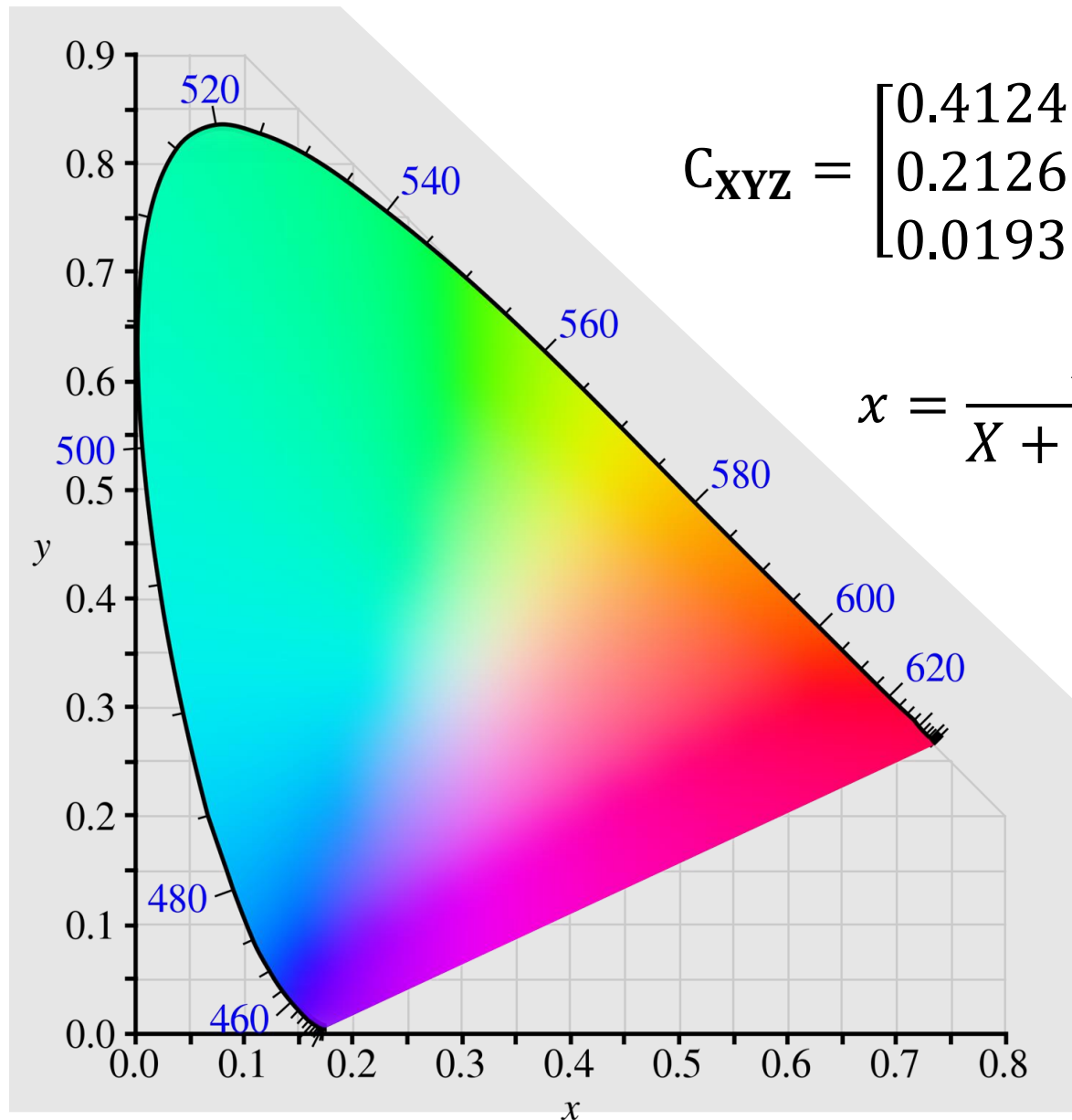
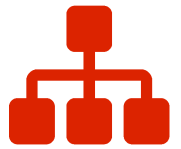
H 120°  
S 1.000  
I 0.333

H 60°  
S 1.000  
I 0.667

H 180°  
S 0.594  
I 0.494

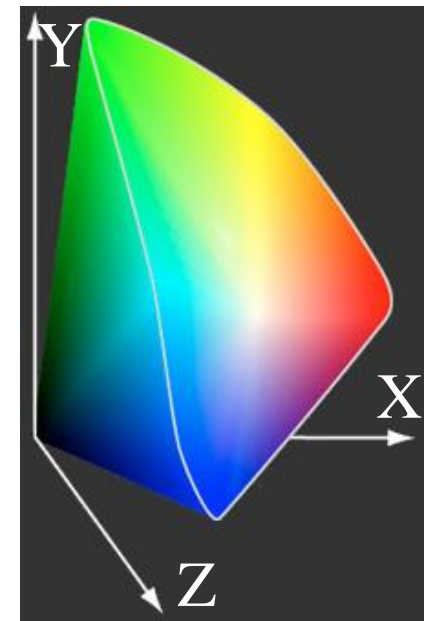
H 273°  
S 0.571  
I 0.282

# CIE XYZ



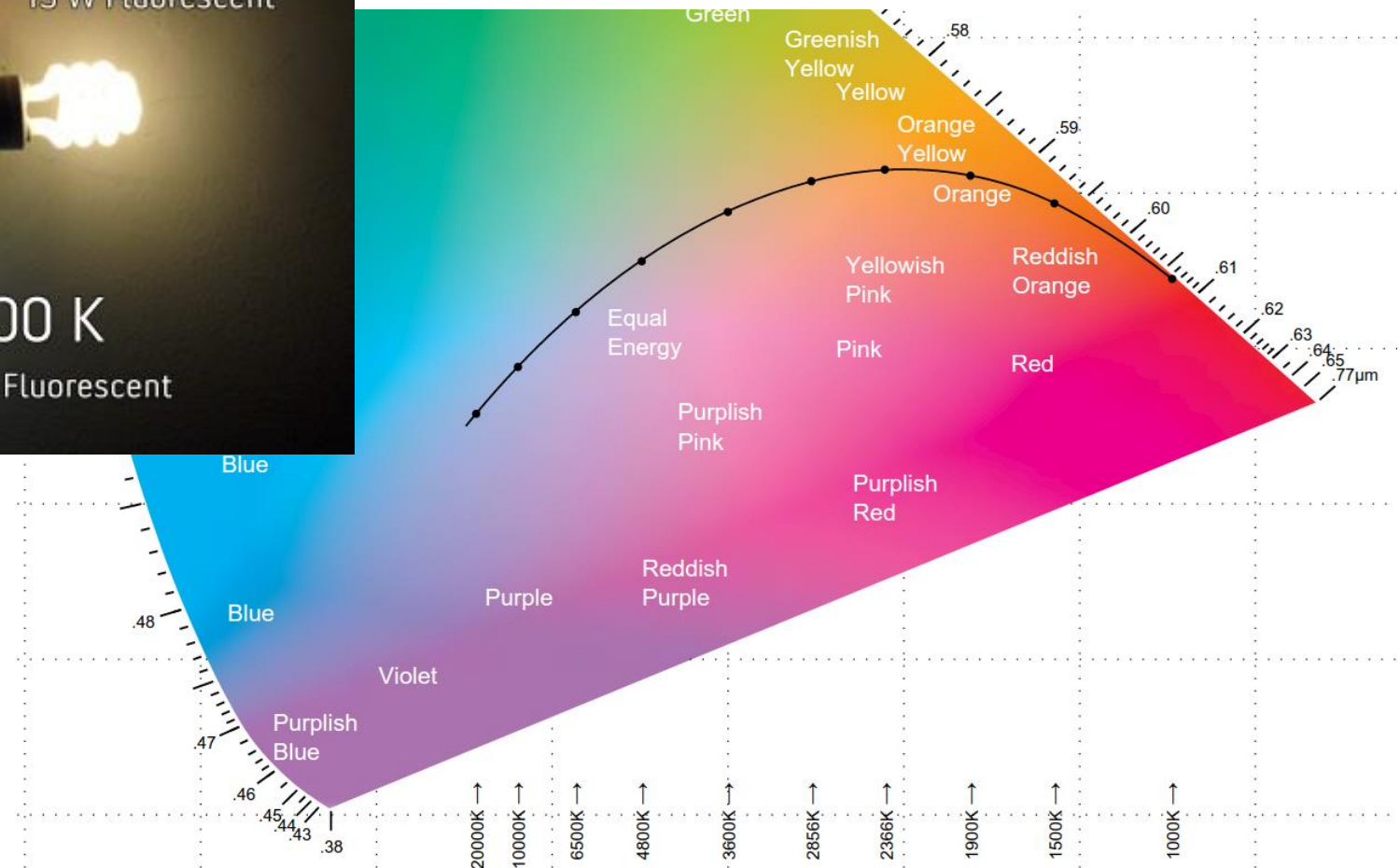
$$C_{XYZ} = \begin{bmatrix} 0.412453 & 0.35758 & 0.180423 \\ 0.212671 & 0.71516 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix}$$

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





# Színhőmérséklet



# Szín tér kihasználtsága

