

Visión por Computadora I

Ing. Andrés F. Brumovsky
(abrumov@fi.uba.ar)

Laboratorio de Sistemas Embebidos -FIUBA

PROGRAMA SUGERIDO

- **Clase 1:** Introducción a imágenes y OpenCV
- **Clase 2:** Op. de píxel, histogramas, binarización, White patch, coord. cromáticas
- **Clase 3:** Filtros: Lineales, separables, padding, DoG, Fourier, Bordes (Canny)
- **Clase 4:** Bordes. Harris. Transformada de Hough. Pirámides.
- **Clase 5:** Extracción de características. SIFT, SURF, ORB, FAST.
- **Clase 6:** Segmentación: K-Means, watersheed, mean-shift, texturas, graph-cut.
- **Clase 7:** Introducción a clasificación y detección: PCA, SVMs, AdaBoost.
- **Clase 8:** Seguimiento. Kalman, filtro de partículas.



RÉGIMEN DE APROBACIÓN

- Trabajos prácticos a implementarse y entregarse luego de las clases
- Examen teórico final online

Dinámica esperada para las clases:

- ↗ ▪ 90 minutos de teoría
- ↗ ▪ 10 minutos de descanso
- ↗ ▪ 80 minutos de práctica



HERRAMIENTAS PARA LA CURSADA

- Lenguaje de programación

- Python 3.8 *Python*

- Librerías de código

- Numpy 1.18

- OpenCV 3.4.2

- Matplotlib 3.1.3

- Entorno de programación

- Jupyter notebook 6.0.3

- IDE recomendado

- Anaconda Navigator 1.9.12

- <https://www.anaconda.com/distribution/>



BIBLIOGRAFÍA SUGERIDA

- La bibliografía es de referencia y no será obligatorio el uso de la misma.

- Computer Vision: Algorithms and Applications | Szeliski | Springer

- <http://szeliski.org/Book>

- Computer Vision: A Modern Approach | Forsyth, Ponce | Pearson

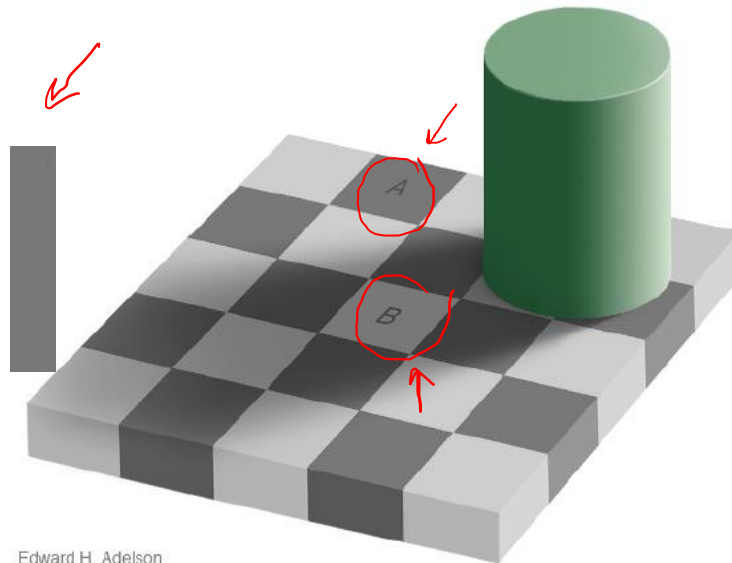
- Computer Vision | Shapiro | Pearson

- Learning OpenCV | Bradski, Kaehler | O'Reilly

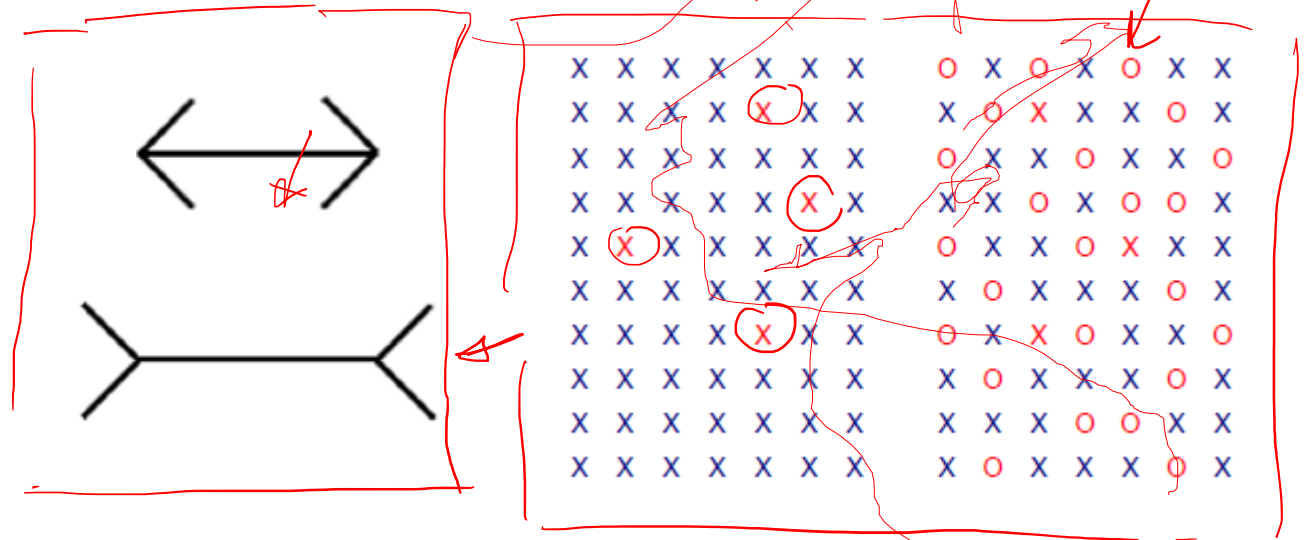


DESAFÍOS

- ¿Por qué puede la visión por computadora resultar compleja?
 - Lo humanos reconocemos nuestro entorno con relativa facilidad
 - Podemos segmentar objetos de un entorno aparentemente sin esfuerzo
 - Podemos mirar una foto y nombrar a las personas que reconocemos
 - Inclusive podemos intuir sus emociones a partir de las expresiones
 - Podemos seguir una persona que se mueve a través de un entorno complejo



Edward H. Adelson



Visión por computadora: Replicar el ojo humano + Replicar el cerebro humano



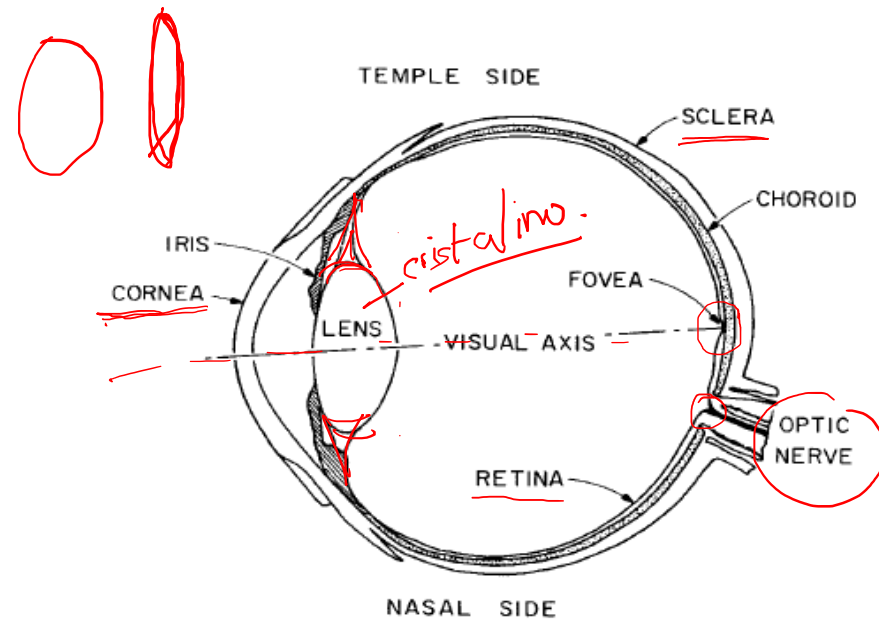
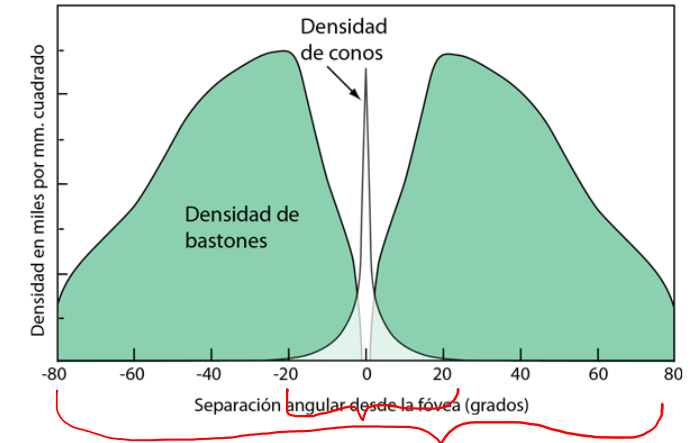
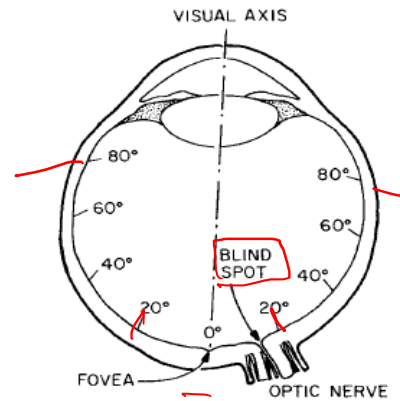


FIGURE 2.2-1. Eye cross section.



PERCEPCIÓN DE LA LUZ POR EL OJO

- **Cornea:** Superficie transparente que envuelve al ojo
- **Sclera:** Lo que resta de la cobertura exterior
- **Coroide:** Capa que contiene capilares sanguíneos, dentro de esta capa está la retina
- **Retina:** Contiene dos tipos de células receptoras de luz, conos y bastones
- **Nervio óptico:** Conjunto de nervios que provienen de la retina
- **Cristalino:** Lente que cambia de forma mediante control muscular



PERCEPCIÓN DE LA LUZ POR EL OJO

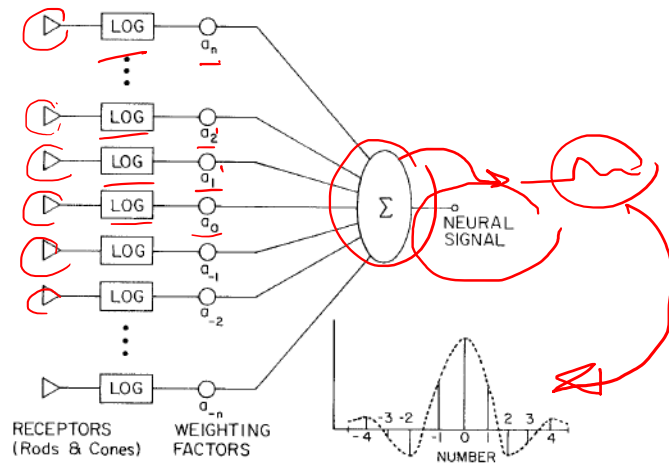
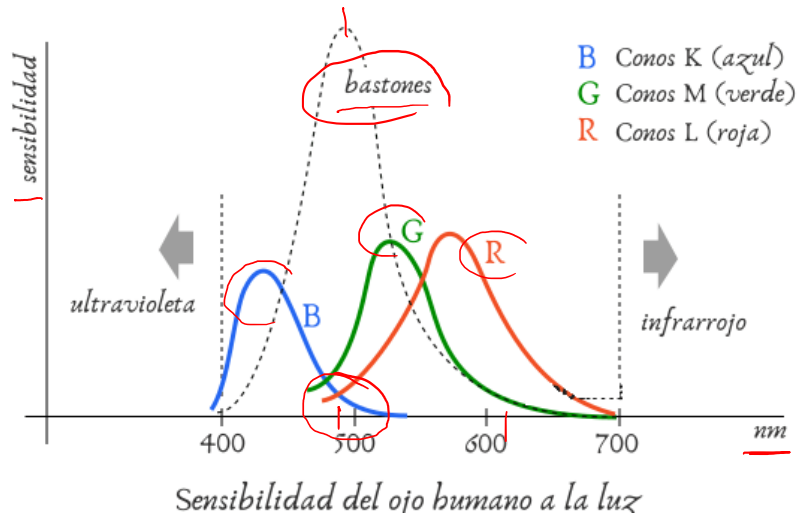
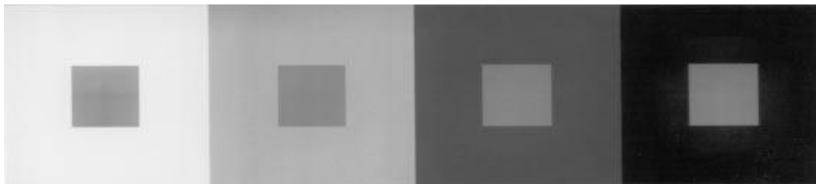


FIGURE 2.4-6. Lateral inhibition effect.



- Bastones: Intensidad (responden poco al rojo). 120 millones
- Conos: Color (concentrados en la mácula). 6~7 millones
 - Rojos – 64%
 - Verdes – 32 %
 - Azules – 2 %
- Respuesta logarítmica. Fracción de Weber (0.02)
- Daltonismo (8% hombres / 1% mujeres)



MODELOS DE COLOR

- RGB: Commission Internationale d'Eclairage (CIE) en 1930

- Rojo: 700 nm
- Verde: 546,1 nm
- Azul: 435,8 nm

- XYZ: Resuelve el color negativo

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0,17697} \begin{bmatrix} 0,49 & 0,31 & 0,20 \\ 0,17697 & 0,81240 & 0,01063 \\ 0,00 & 0,01 & 0,99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

CIE XYZ

Además, permite separar crominancia de luminancia

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

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

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$

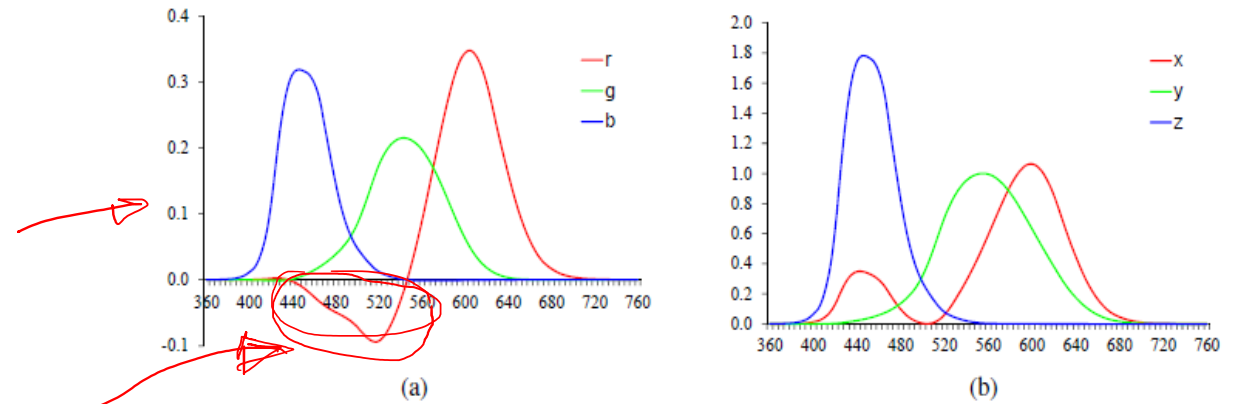
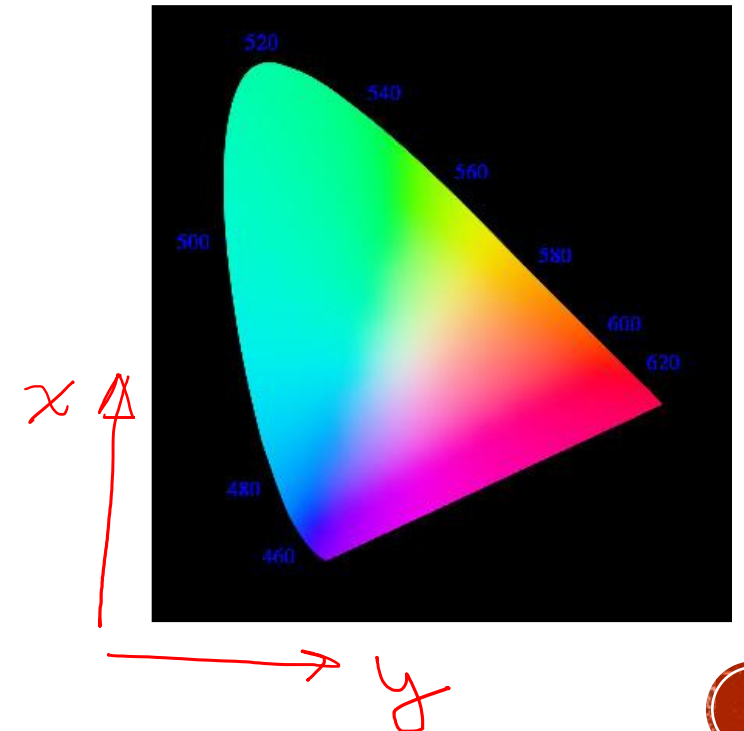
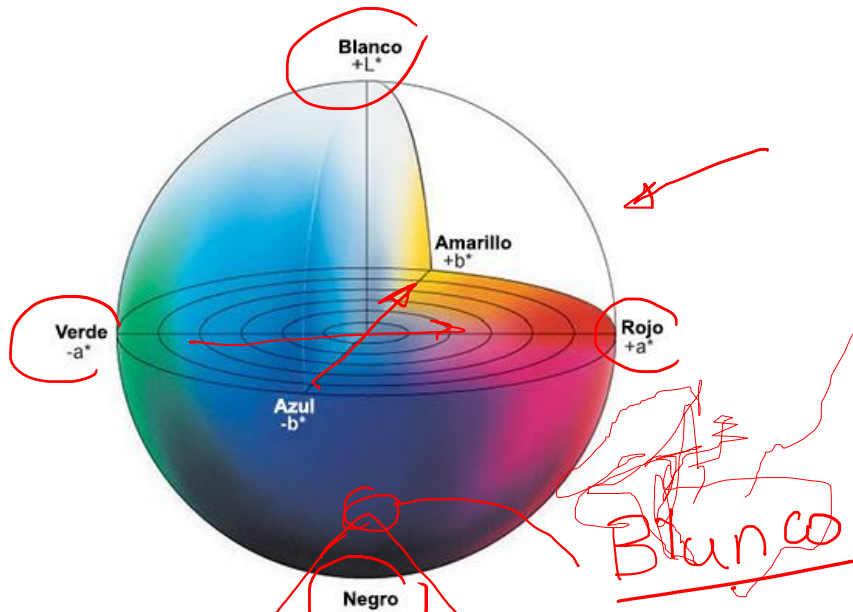


Figure 2.28 Standard CIE color matching functions: (a) $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ color spectra obtained from matching pure colors to the R=700.0nm, G=546.1nm, and B=435.8nm primaries; (b) $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ color matching functions, which are linear combinations of the $(\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda))$ spectra.



OTROS ESPACIOS DE COLOR

▪ CIELAB ($L^*a^*b^*$)



$$L^* = 116 f\left(\frac{Y}{Y_n}\right); \quad a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right]; \quad b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right]$$

$$f(t) = \begin{cases} t^{1/3} & \text{si } t > \delta^3 \\ \frac{t}{3\delta^2} + \frac{2\delta}{3} & \text{otro caso} \end{cases}$$

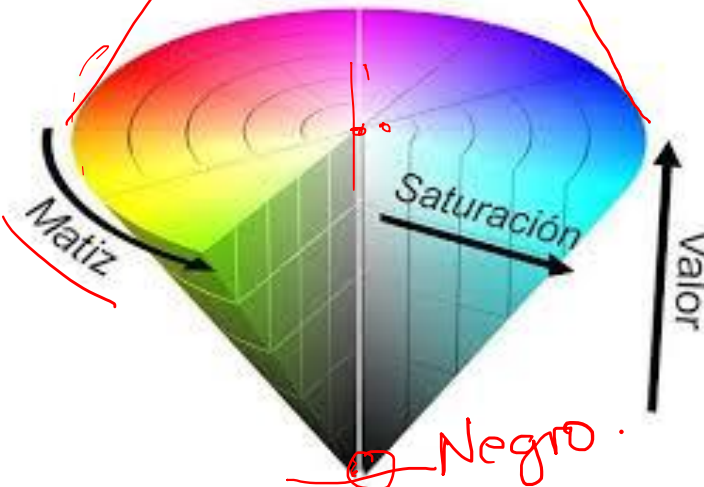
▪ HSV

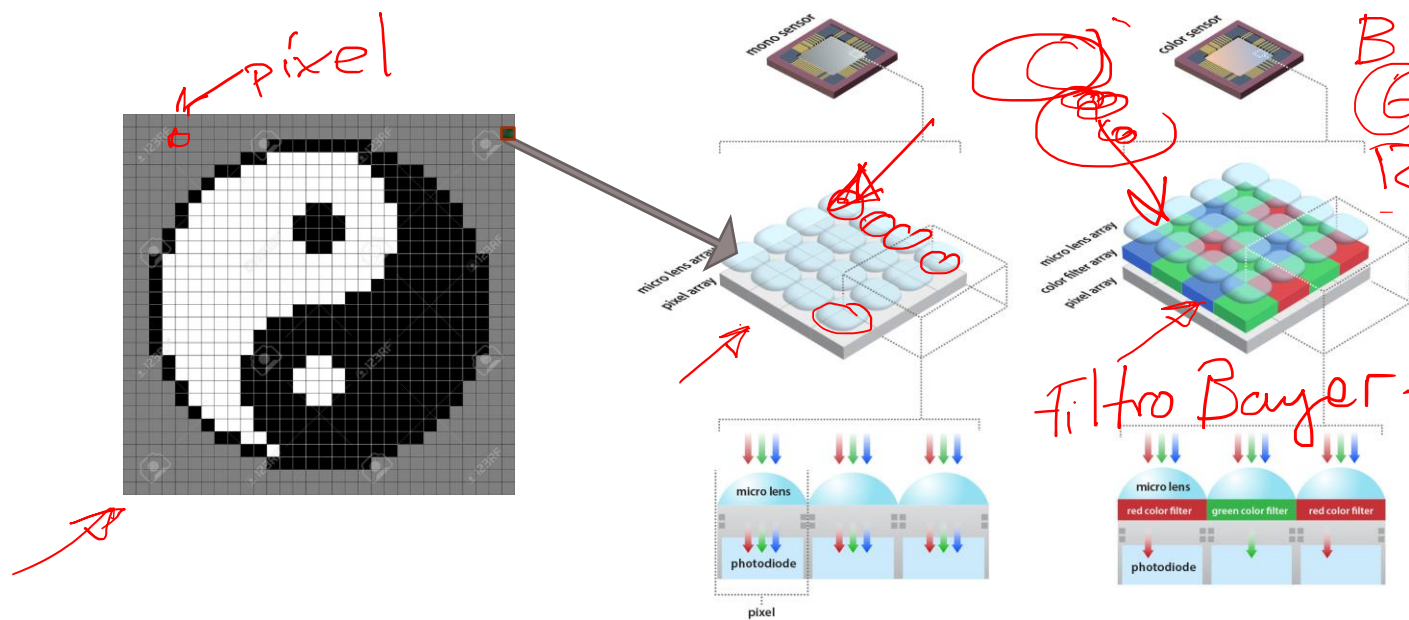
- **Hue:** Dirección alrededor de la rueda de color, en grados $\in [0, 360]$
- **Saturation:** Distancia escalada desde la diagonal $\in [0, 1]$
- **Value:** Promedio o máximo valor de color $\in [0, 1]$

Conversión RGB \rightarrow HSV

- $V = M = \max(R, G, B); \quad m = \min(R, G, B)$
- $S = (M - m)/M \quad (S = 0, \text{ si } V = 0)$
- $H = 60 \times \begin{cases} 0 & , \text{ si } (M - m) = 0 \\ 0 + (G - B)/(M - m), & \text{ si } \max = R \\ 2 + (B - R)/(M - m), & \text{ si } \max = G \\ 4 + (R - G)/(M - m), & \text{ si } \max = B \end{cases}$

$$H = H + 360, \text{ si } H < 0$$





Handwritten notes and formulas:

- $B_1 G_2 R_1 (G_1 + G_2 + G_3 + G_4) / 4$
- $(B_1 + B_2) / 2$
- $G_1 R_1 G_3$
- $R_1 G_1 B_2$

TIPOS DE SENSORES

▪ **CCD** (charge-couple device):
Willard Boyle y George E. Smith
– Laboratorios Bell – 1969

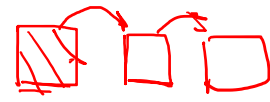
- A/D central 20 a 75MHz
- Blooming/ Smearing
- Mayor sensibilidad

▪ **CMOS** (complementary metal oxide semiconductor)

- Conversión en el fotositio
- Rolling shutter
- Microlentes

▪ **Otros parámetros a considerar**

- Eficiencia cuántica
- Capacidad de pozo
- Binning
- ROI
- LUTs



Handwritten text: 640 x 320



Nombre	Resolución	Nro. píxeles
VGA	640x480	0,3 MP (Megapíxel)
SVGA	800x600	0,48 MP (Megapíxel)
XGA	1024x768	0,78 MP (Megapíxel)
SXGA	1280x1024	1,3 MP (Megapíxel)
UXGA	1600x1200	1,9 (2) MP (Megapíxel)
SUXGA	2048x1536	3,1 MP (Megapíxel)
-	2048x2048	4,0 MP (Megapíxel)
-	2452x2054	5,0 MP (Megapíxel)
QUXGA	3200x2400	7,7 MP (Megapíxel)
HD	1280x720	0,92 MP (Megapíxel)
Full HD	1920x1080	2,1 MP (Megapíxel)