

Tema 1: Introducción al Proceso Digital de Imágenes

Introducción



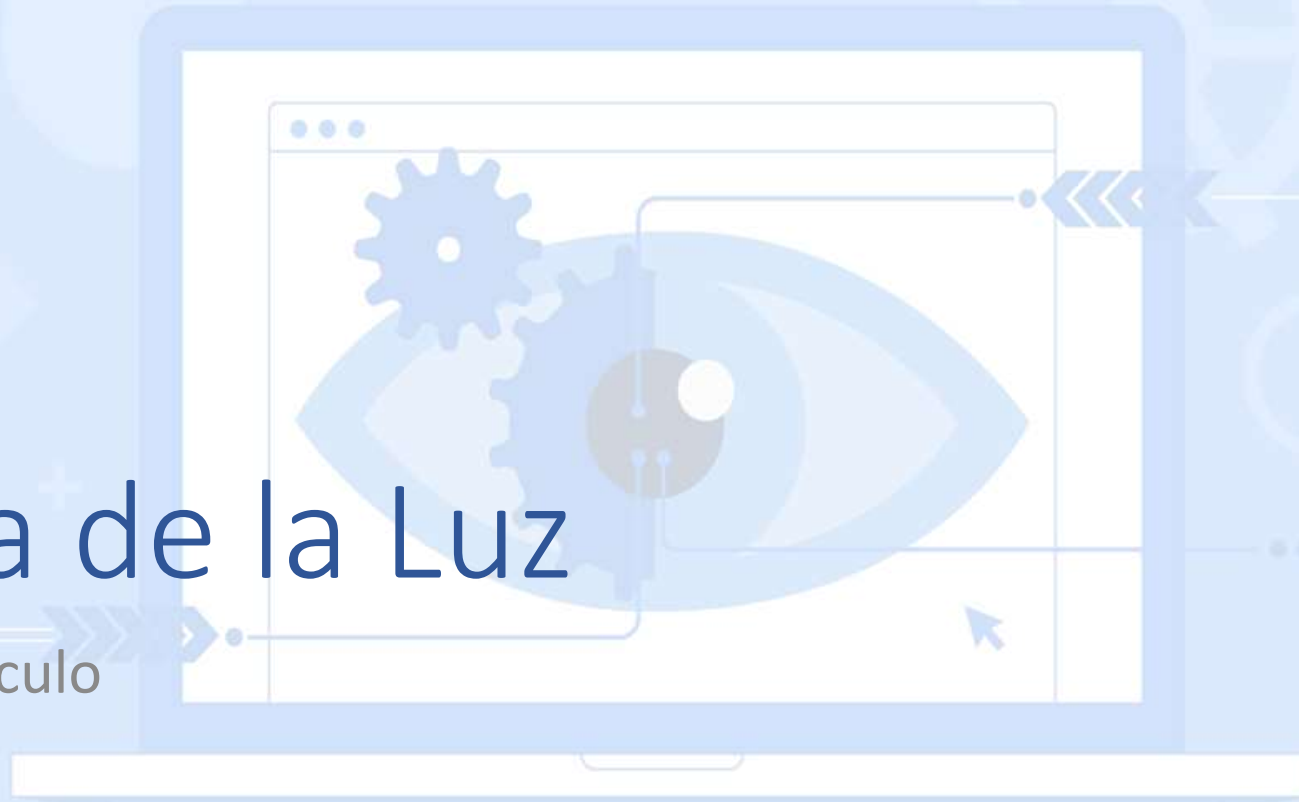
1. Naturaleza de la Luz
2. El color
3. Formación de la Imagen
4. Obtención de la Imagen
5. Fundamentos de Óptica
6. Imagen Digital
7. Tratamiento de Imagen
8. Filtros y Convoluciones



Naturaleza de la Luz

Dualidad onda-corpúsculo

Efecto fotoeléctrico



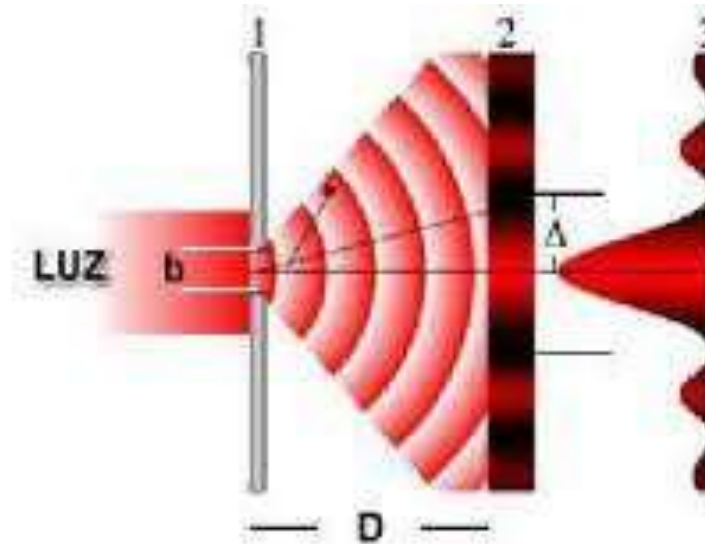
Naturaleza de la luz

Teoría **corpúscular** (Newton, 1669):
 explica fenómenos como la reflexión, la
 refracción y dispersión de la luz en un
 prisma

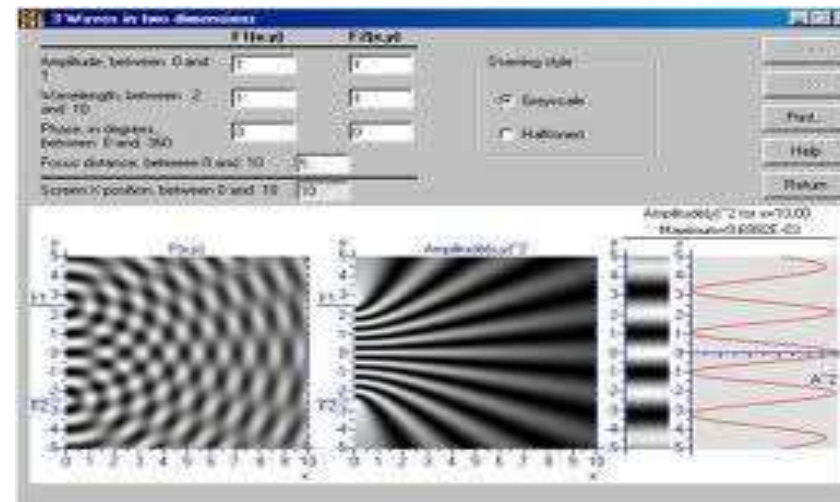
Teoría **ondulatoria** (Huygens, 1690):
 explica fenómenos de difracción e
 interferencia luminosa

Posteriores trabajos de Young (s. XIX)
 confirman la explicación anterior
 (experimento de la doble rendija)

Fresnel contribuye en 1816 a esa
 confirmación



1. Rendija 2. Luz recogida en la pantalla 3. Distribución de intensidades



- Teoría **electromagnética** (Maxwell, s. XIX): teórico del modelo ondulatorio, unificación de luz, electricidad y magnetismo.
- Cada tipo de radiación electromagnética se caracteriza por su longitud de onda λ

$$\lambda = c/\nu$$

c velocidad de la luz, ν frecuencia

- Teoría **cuántica** (Planck, 1900): la radiación electromagnética emitida se realiza en “cuantos de energía” o fotones

$$E = h\nu$$

(h cte. de Planck $6,62 \cdot 10^{-34} \text{ J}\cdot\text{s}$)

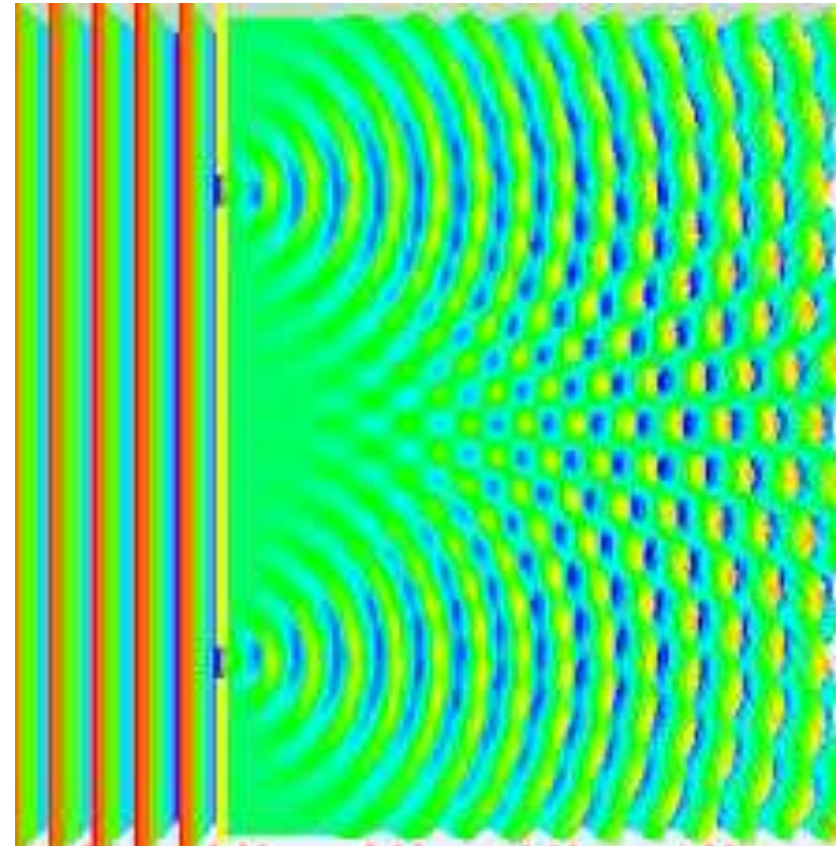
- Vuelta al modelo corpuscular (Nobel 1918)



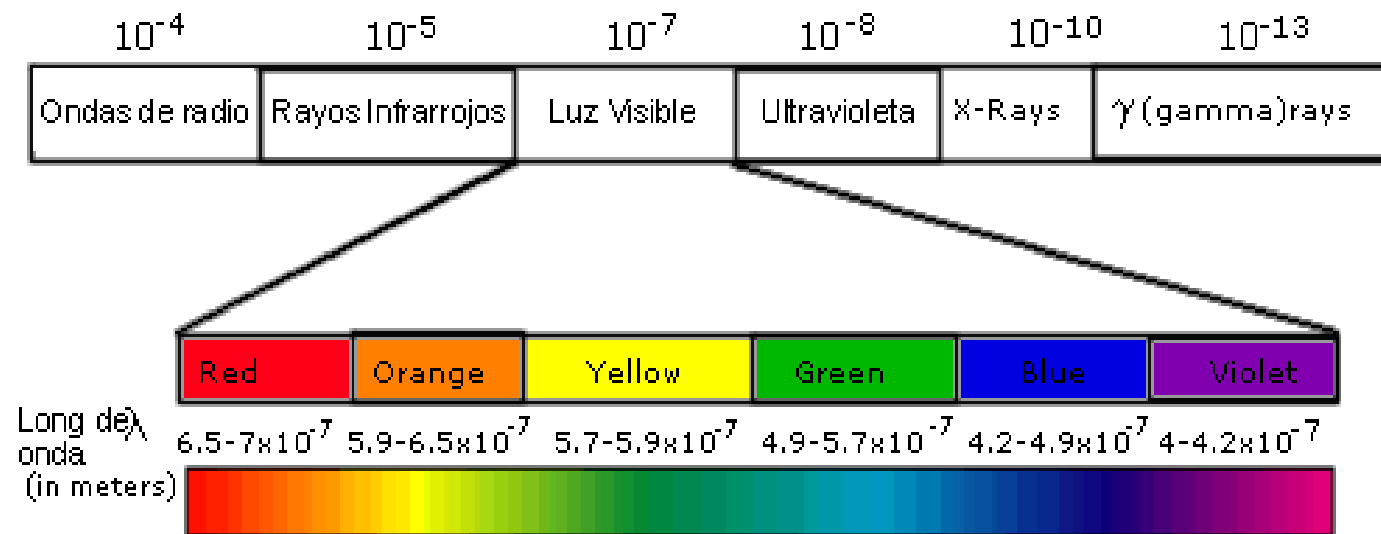
http://es.wikipedia.org/wiki/James_Clerk_Maxwell

Naturaleza de la luz

- Efecto **fotoeléctrico** (Einstein, 1905): “emisión de electrones en una superficie metálica era producida por la absorción de cuantos de luz”.
 - Se explica mediante la teoría cuántica
 - (Nobel 1921)
- Teoría **onda-corpúsculo** (de Broglie-Bohr, 1925): la luz tiene naturaleza ondulatoria y corpuscular
 - Relacionó la cantidad de movimiento con la longitud de onda
 - Naturaleza ondulatoria del electrón (hipótesis de De Broglie) Nobel 1929

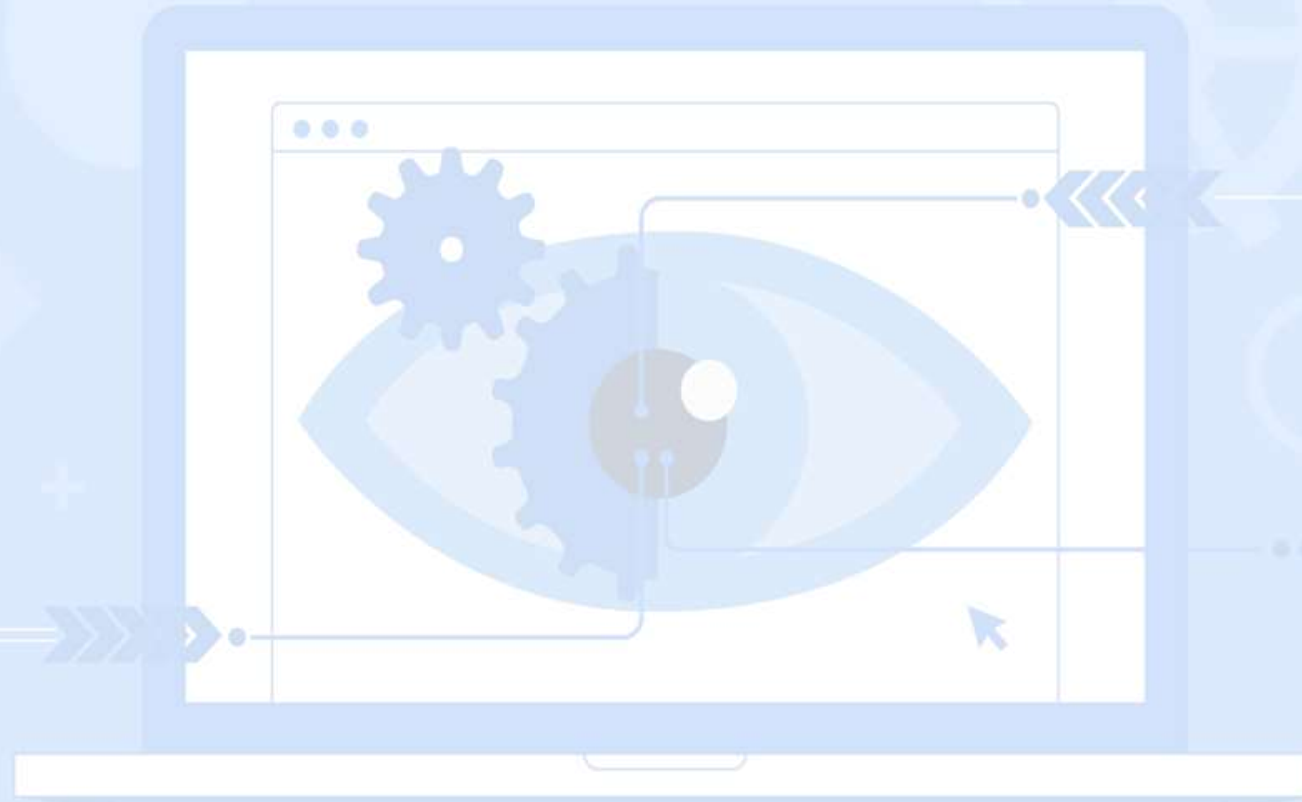


- Una parte pequeña del espectro electromagnético es la zona visible (luz visible)
- Comprende longitudes de onda aproximadamente desde los 400 hasta los 700 nm

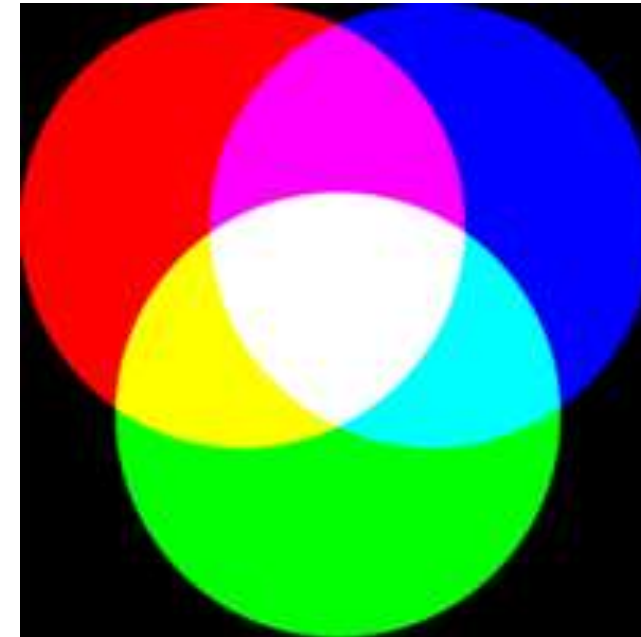


El Color

Modelos de color
 RGB, CMYK, HSV ...

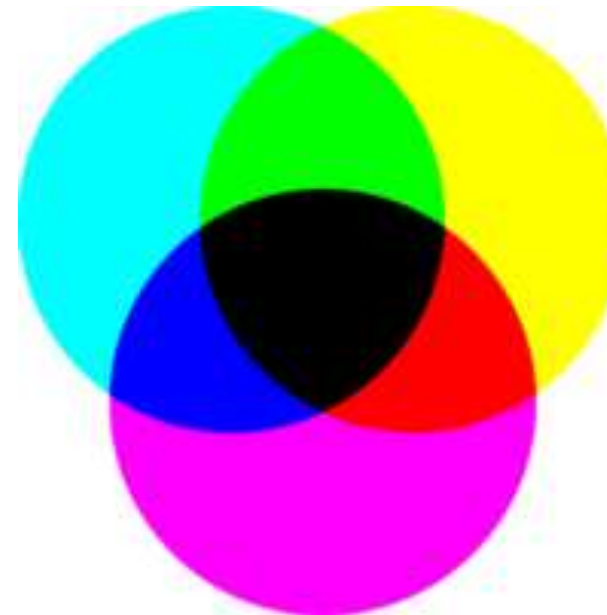


- El significado diferente del color en la luz (espectro emitido) y en los objetos (espectro reflejado) da lugar a dos modos de ver el proceso de mezcla de colores: modelo aditivo y substractivo.
- Modelo **aditivo** mezcla:
 - Corresponde a los colores luz.
 - La mezcla de dos colores se obtiene sumando los espectros asociados a ambos colores.
 - Podemos seleccionar un número reducido de colores primarios y obtener los demás con ellos.
 - Por ejemplo, si tomamos como primarios rojo, verde y azul, los secundarios son amarillo, magenta y cian.



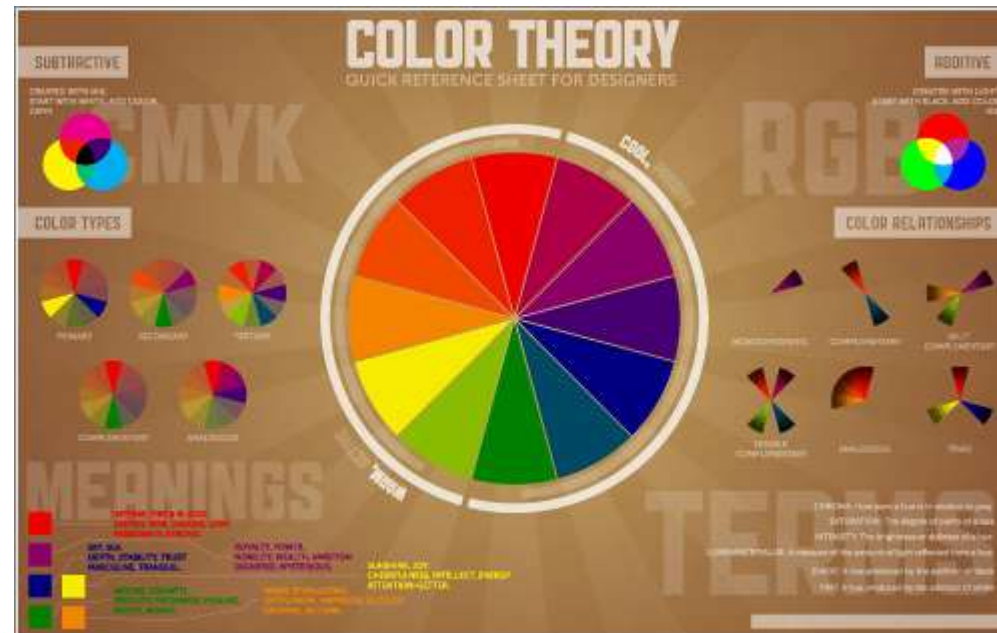
http://en.wikipedia.org/wiki/Color_space

- Modelo **subtractivo** de mezcla:
 - Corresponde a los colores pigmento.
 - La mezcla de dos colores se obtiene tomando el producto de los espectros asociados a ambos colores.
 - Si tomamos como colores primarios cian, magenta y amarillo, los colores secundarios son rojo, azul y verde.
 - Colores terciarios: los que surgen de combinar un primario con un secundario.



http://en.wikipedia.org/wiki/Color_space

- El estudio de los modelos de color es importante, porque el modelo condiciona cómo se captura, almacena, procesa, transmite y genera el color.
 - Existen muchos modelos de color. Algunos son mejores para ciertas aplicaciones. No todos son completos.
 - **RGB** es el que más se ajusta al modo de captura (filtros de color) y de generación (píxeles del monitor) en imagen digital.
 - **CMYK** se relaciona con la generación de color en impresoras.
 - **YIQ** y **YUV** separan crominancia (color) y luminancia (brillo).



El Color

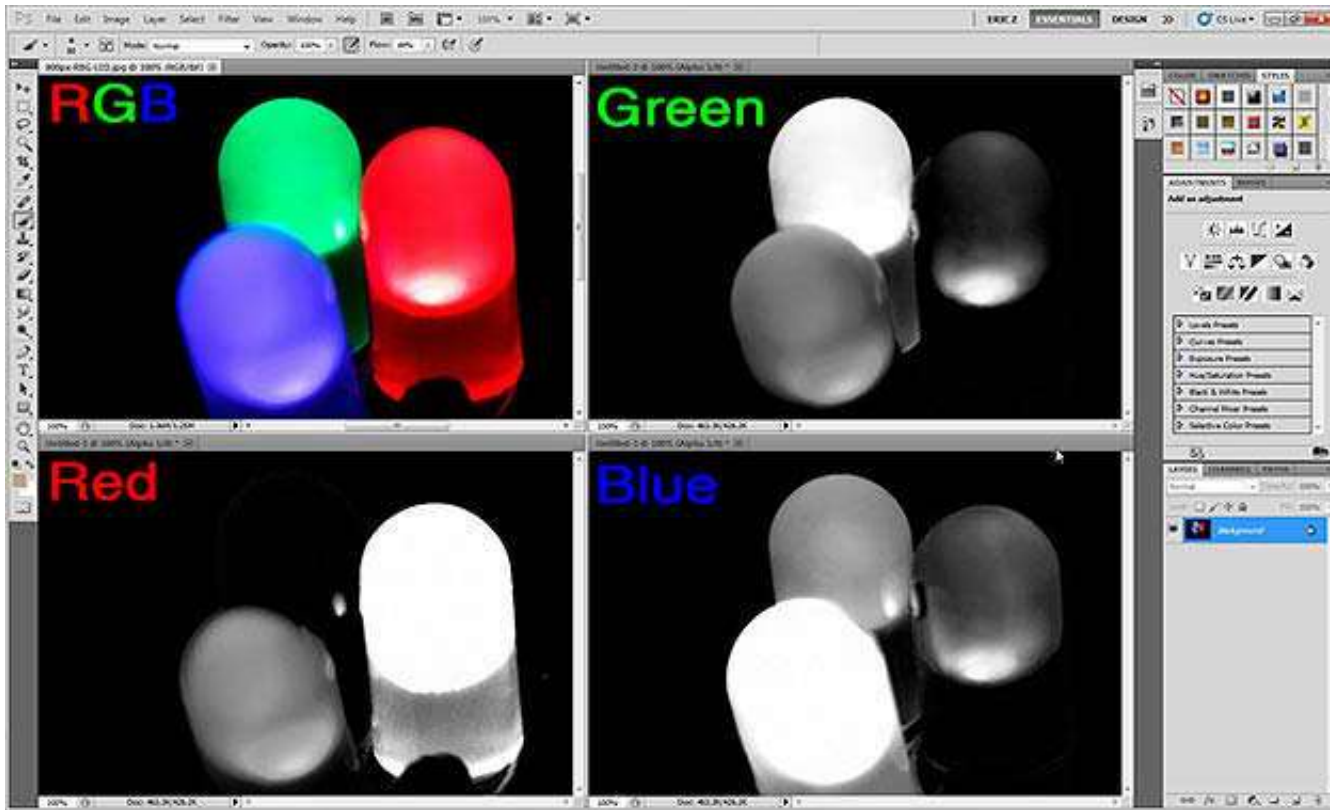
El modelo RGB:

Es el modelo más usado en aplicaciones digitales de imagen.

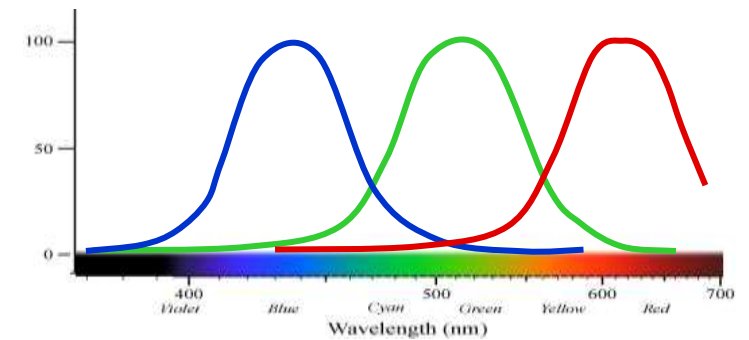
Basado en un modelo aditivo de mezcla, con 3 colores primarios: R-rojo, G-verde, B-azul.

La combinación aditiva de estos colores primarios produce todo el rango de colores representables en RGB.

(R,G,B)	Long. de onda	Frecuencia
(255,0,0)	~ 625-740 nm	~ 480-405 THz
(255,128,0)	~ 590-625 nm	~ 510-480 THz
(255,255,0)	~ 565-590 nm	~ 530-510 THz
(0,255,0)	~ 500-565 nm	~ 600-530 THz
(0,255,255)	~ 485-500 nm	~ 620-600 THz
(0,0,255)	~ 440-485 nm	~ 680-620 THz
(0,139,255)	~ 380-440 nm	~ 790-680 THz



Cada imagen de gris representa la intensidad de cada color por separado.



El Color

El modelo CMY:

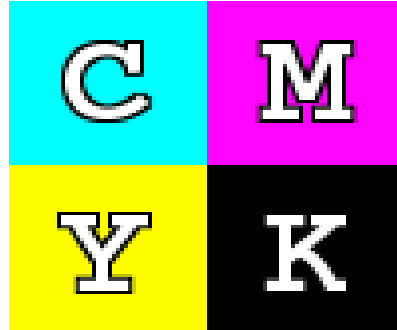
En ciertas aplicaciones, como por ejemplo impresión de imágenes, se utiliza más el modelo CMY (o CMYK).

CMY está basado en un modelo sustractivo de mezcla, con 3 colores primarios: C-cian, M-magenta, Y-amarillo.

La combinación sustractiva (tintas de color) de estos colores primarios produce todo el rango de colores representables en CMY.

En la práctica, la mezcla de C, M e Y no llega a producir negro, sino una especie de gris marengo.

El modelo **CMYK** soluciona el problema, añadiendo el negro como color primario.



<http://en.wikipedia.org/wiki/CMYK>



El Color

Los modelos HLS y HSV:

Los modelos HLS (o HSL) y HSV están pensados para ser fácilmente interpretables y legibles por un humano, usan términos más familiares cuando hablamos de color.

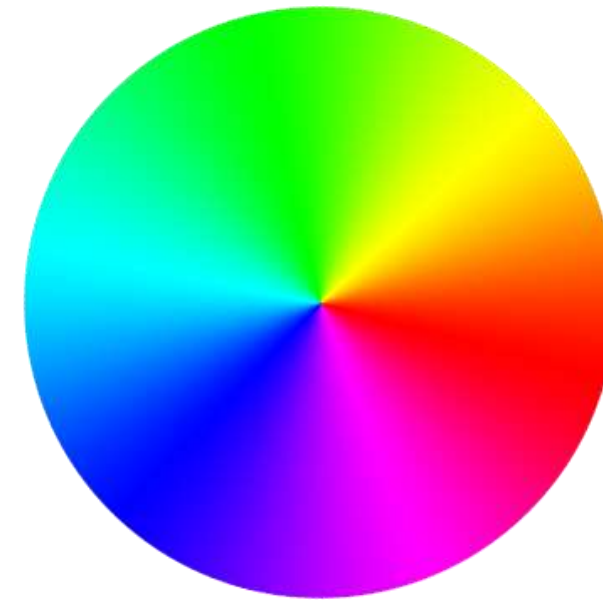
Luminosidad o intensidad de un color: cualidad de ser más claro o más oscuro.

Saturación: diferencia del color respecto a un gris con la misma intensidad. Cuanto más diferente, más saturado.

Matiz de un color: su ángulo dentro de la rueda cromática.

También, se puede definir como la frecuencia dominante del espectro.

Rueda cromática



http://en.wikipedia.org/wiki/HSV_color_space

El Color

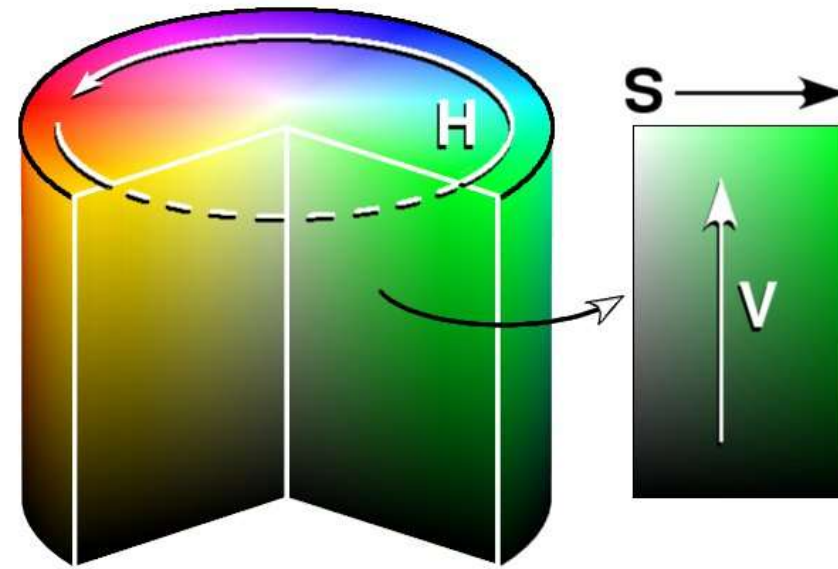
HSV consta de los componentes: **H-matiz** (hue), **S-saturación**, **V-valor de intensidad**.

HLS consta de: **H-matiz**, **L-luminosidad**, **S-saturación**.

Ambos son transformaciones no lineales del RGB.

La definición de H es igual en ambos. La diferencia se encuentra en la forma de calcular la saturación, S, y la intensidad, V o L.

El espacio HSV se suele representar como un cono o como un cilindro.



http://en.wikipedia.org/wiki/HSV_color_space



Canales S y V de HSV



Canal H

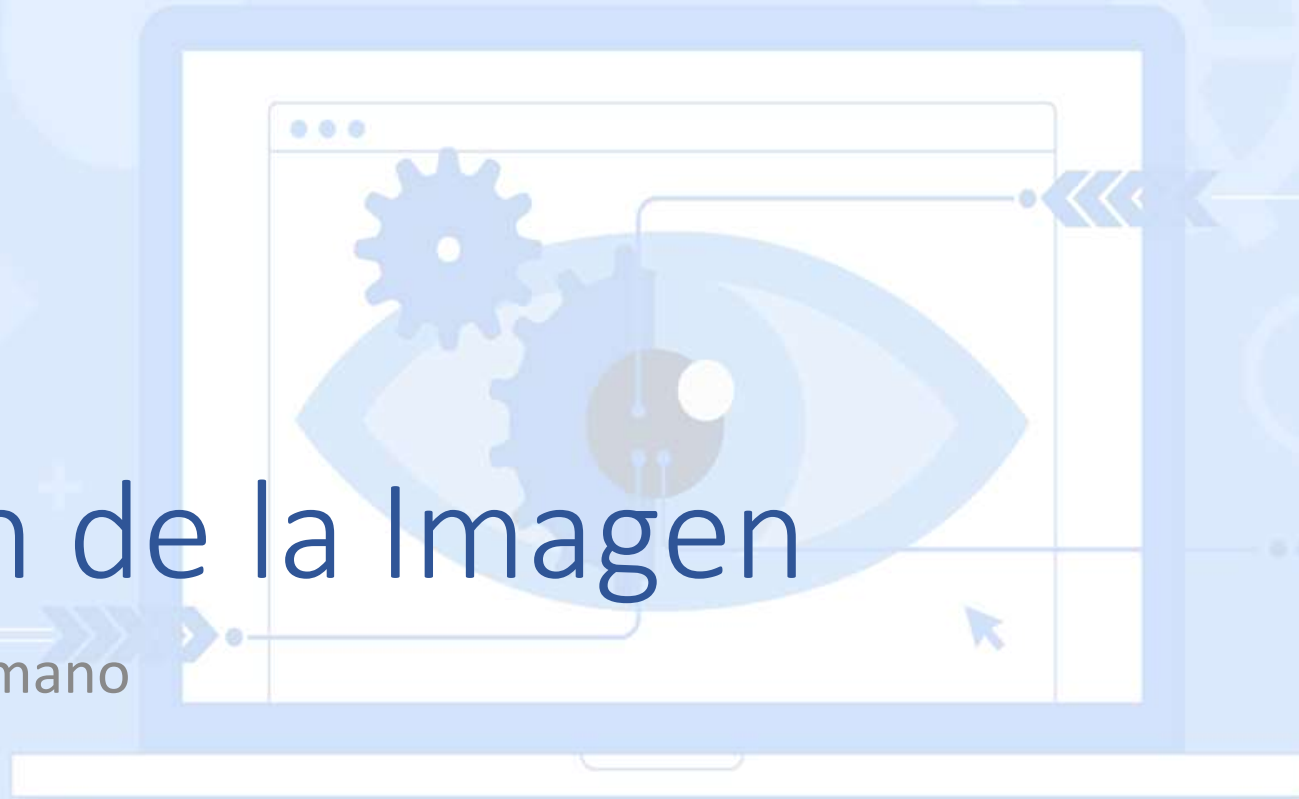


Canales de S y L de HLS

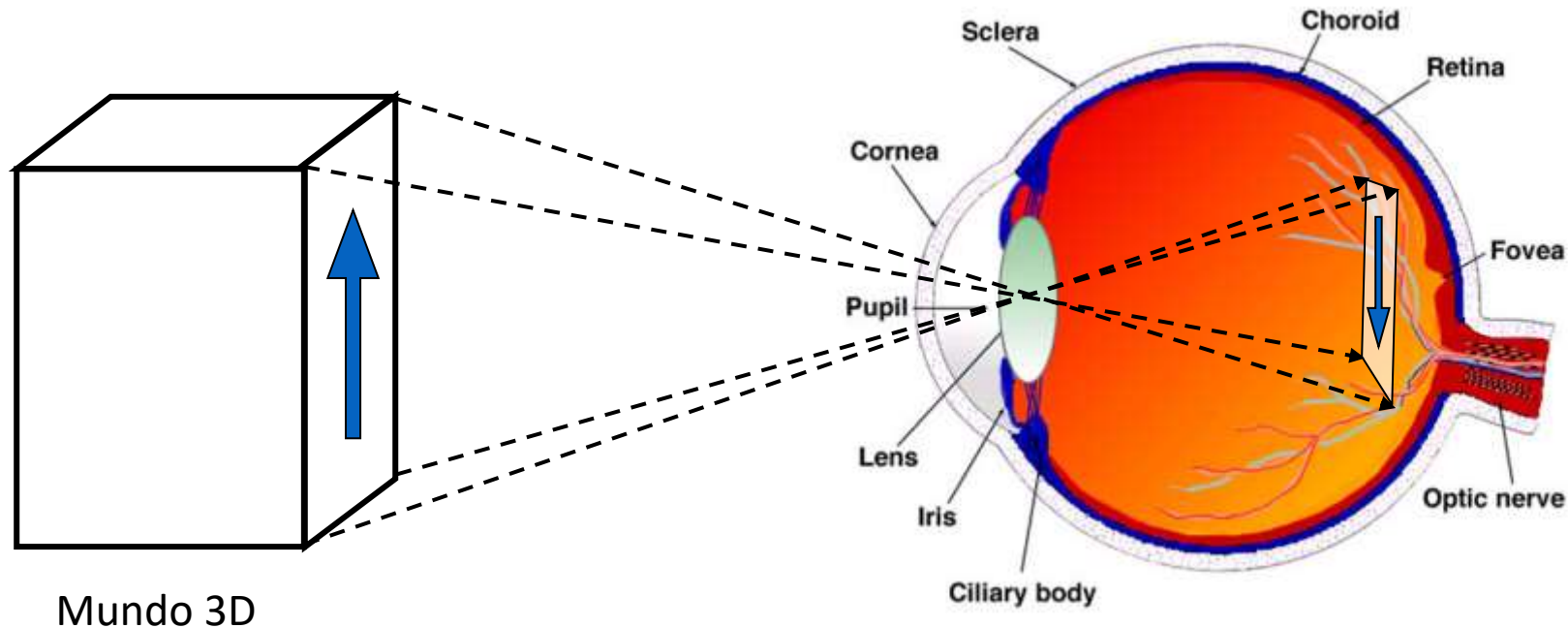
Formación de la Imagen

Percepción del ojo humano

Cámaras pinhole



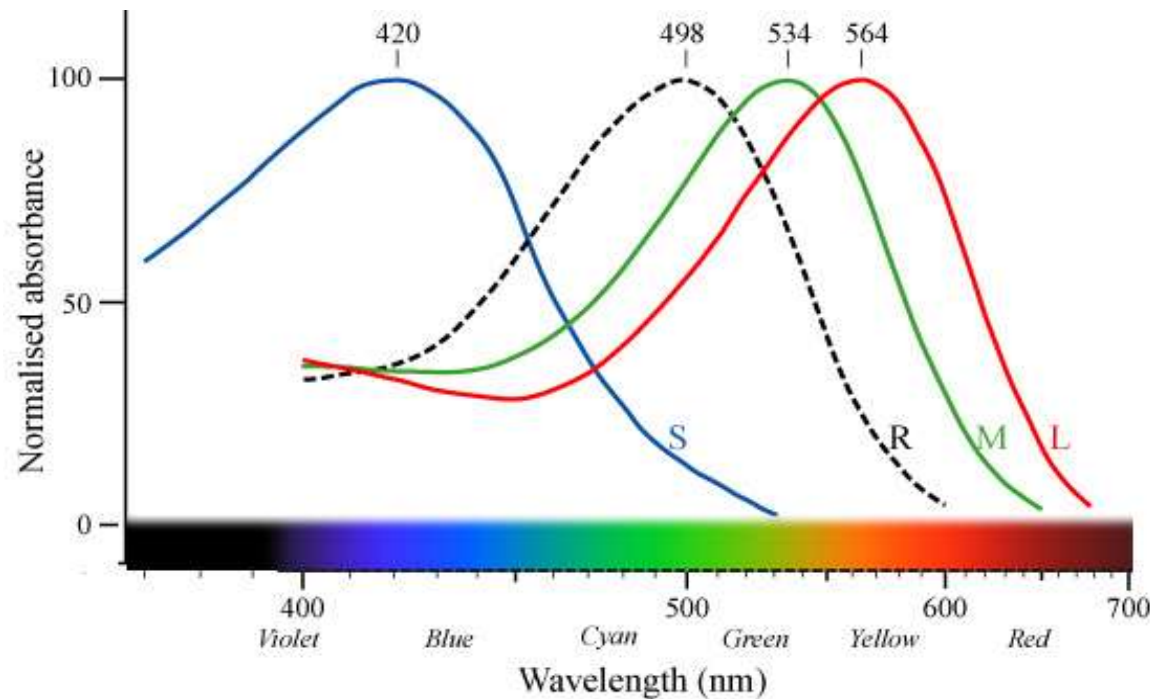
- La formación de imágenes es un proceso mediante el cual una información luminosa 3D (la escena) es proyectada en un plano 2D (la imagen).



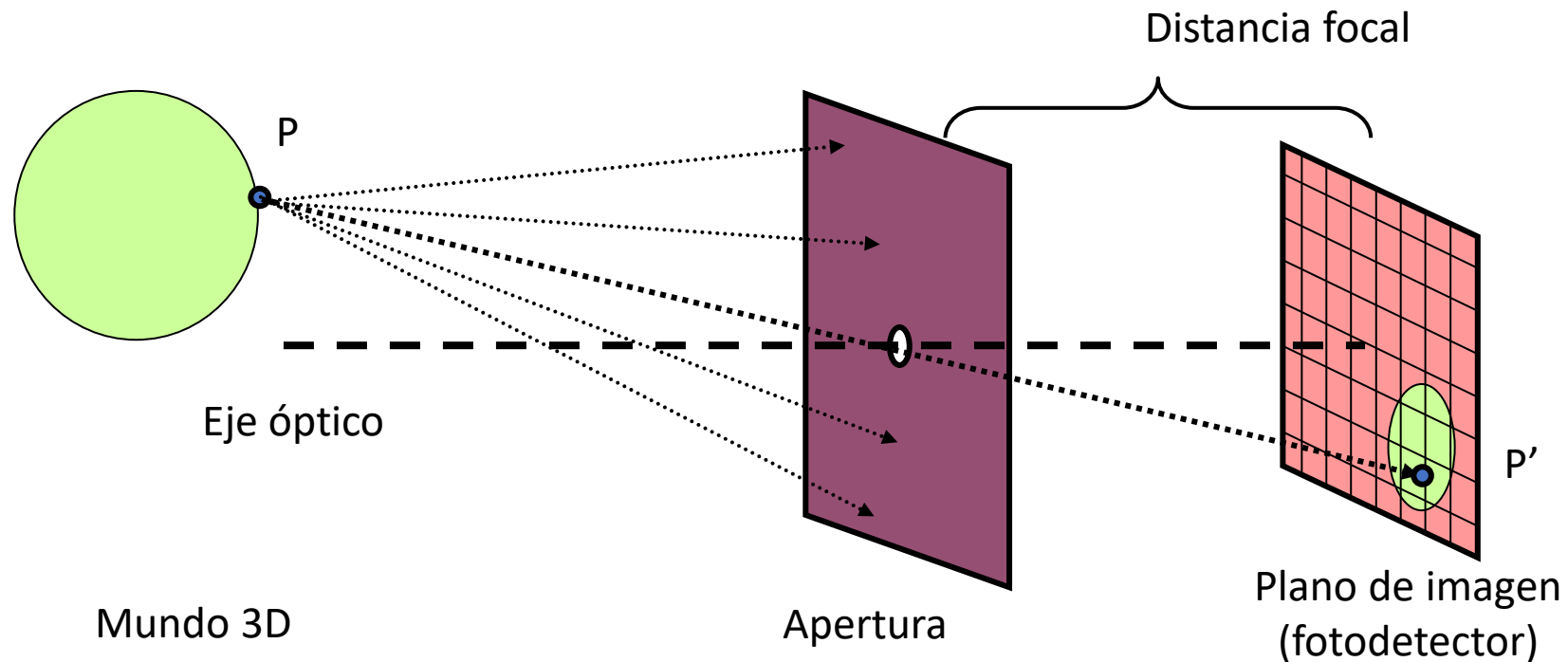
<http://webvision.med.utah.edu/anatomy.html>

- Los conos se clasifican según la longitud de onda de la radiación que captan en: bajos (S), medios (M) y altos (L).
- El grado de excitación de un receptor (cono o bastón) sería una suma, para cada frecuencia, del producto de la absorción por la intensidad de luz entrante.

Absorción espectral de los conos (S,M,L) y los bastones (R)



- Modelo ideal de cámara (pinhole).
- Una superficie mate emite luz en todas las direcciones.
- Cuando la apertura es muy pequeña, desde cualquier punto sólo pasa luz con una dirección.
- Todos los puntos están bien definidos: imagen enfocada.
- Circulo de confusión pequeño



- La imagen estará enfocada cuando el tamaño del círculo de confusión sea menor que el tamaño de cada celda del foto detector.
- Sin embargo, el modelo *pinhole* tiene muchas limitaciones y es poco usado.
- Apertura muy pequeña → Entra muy poca luz → La imagen sale muy oscura.

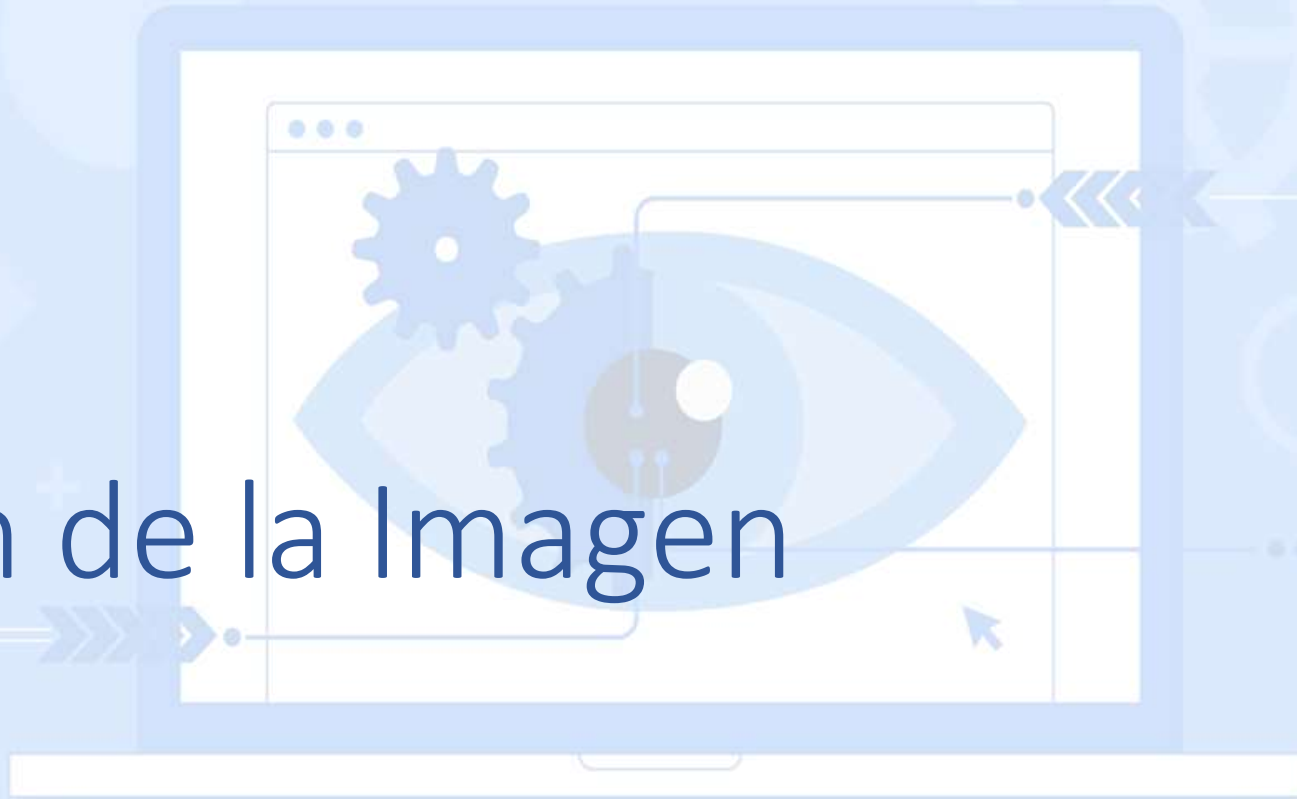


<http://www.pinholeday.org/>

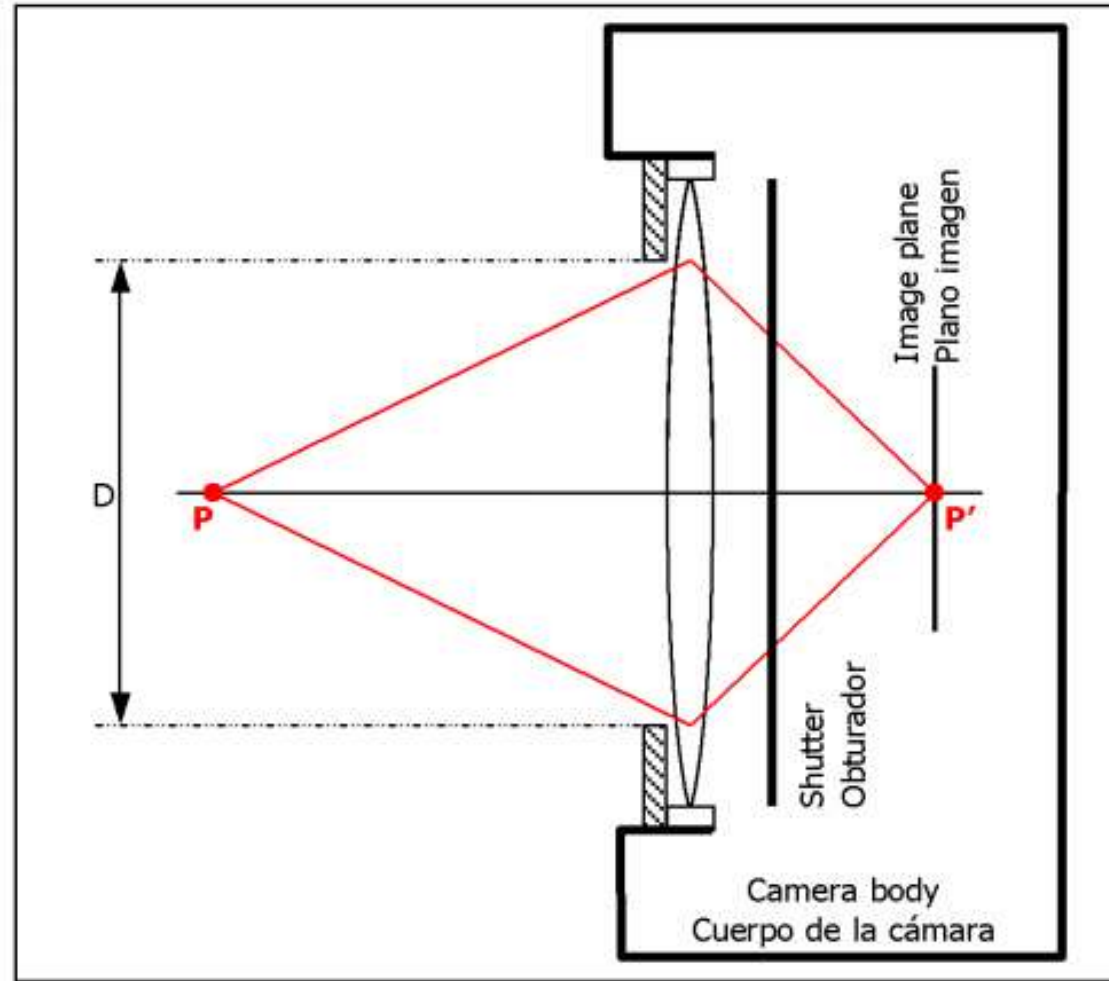
Obtención de la Imagen

Modelo de cámara

Sensores CCD, CMOS



- Cámaras
 - Diafragma
 - Objetivo
 - Obturador
 - Plano de imagen

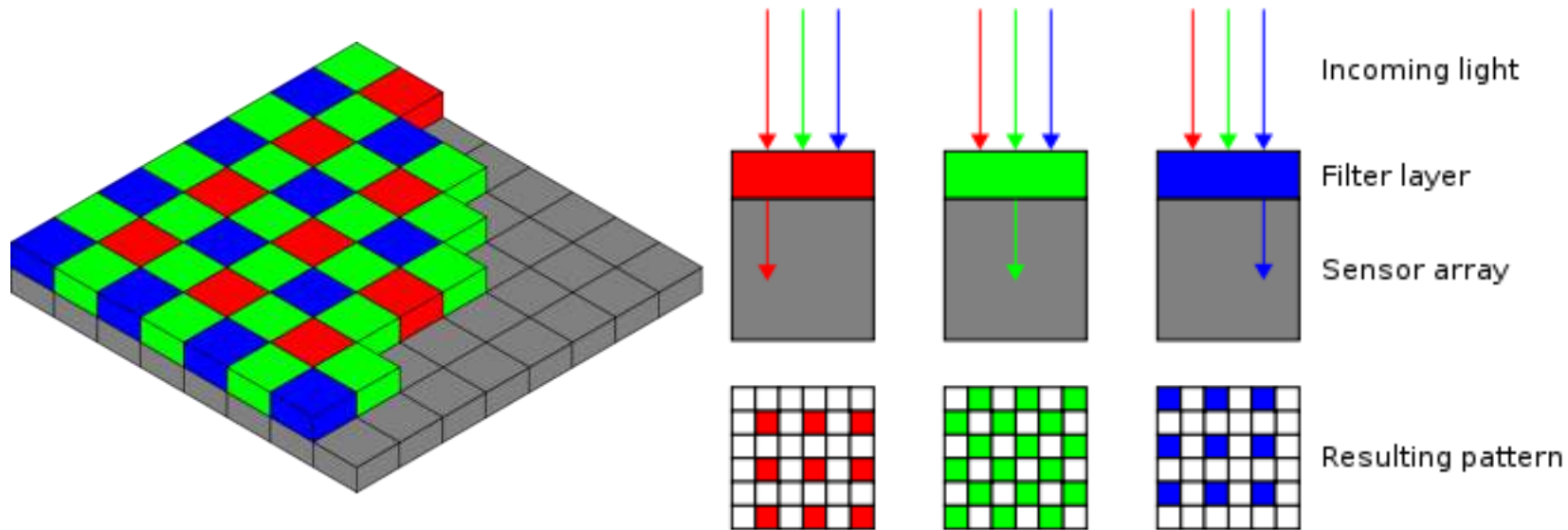


- **Diafragma** (iris). Sistema mecánico que forma un orificio circular por el que penetra la luz. Su tamaño se puede cambiar, con lo que se controla la cantidad de luz que entra en la cámara.
- **Objetivo**. Conjunto de lentes por donde pasa la luz que, procedente de un objeto, entra por el diafragma formando la imagen.
- **Obturador**. Sistema electrónico que bloquea o permite el paso de la luz. Permite controlar el intervalo de tiempo durante el cual el plano imagen recibe la luz que procede del objeto. Tiempos de exposición típicos: $1/25s$, $1/50s$, $1/120s$, $1/250s$

Europa: video PAL $1/50s$, 50 campos/s, 25 imágenes/s

EEUU: video NTSC $1/59,94s$, 29,97 imágenes/s

- **Plano imagen.** Plano donde se capta la imagen. Es donde se sitúa el sensor CCD o CMOS.
- **CCD:** matriz de sensores. Relación de aspecto 4:3
- Mayor nº sensores \Rightarrow mayor resolución.
- Cada píxel almacena como carga eléctrica la información sobre el nivel de luz que ha incidido sobre su superficie.
- Si solo hay un CCD, cada píxel tiene zonas sensibles a cada color
- Sistema **PAL** genera 3 señales : Y, U, V
- Y (luminancia), nivel de brillo, genera la imagen en sistema de blanco y negro
- coordenadas U, V, con las que se reconstruye el color



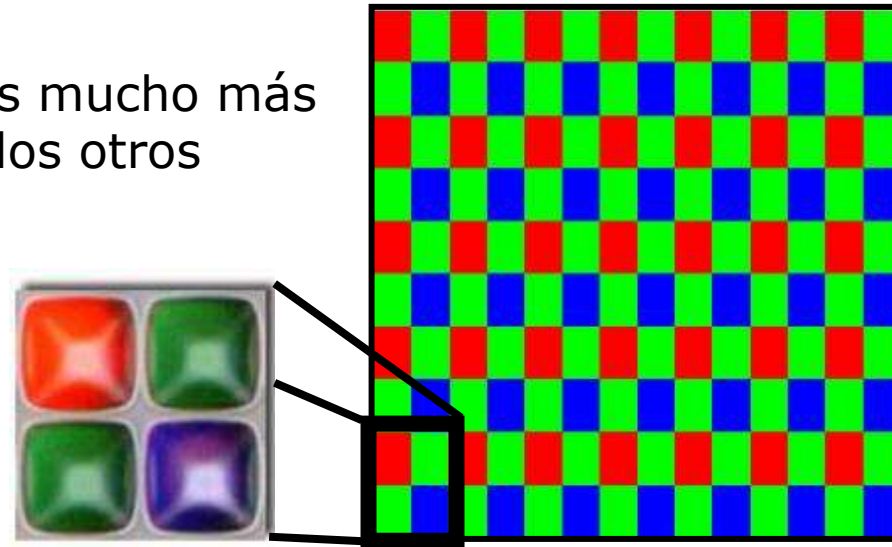
Usan un Filtro Bayer: array de filtros de color
(mosaico con los 3 colores)

Patrón de Bayer:

Distribución típica de los filtros de color en el CCD

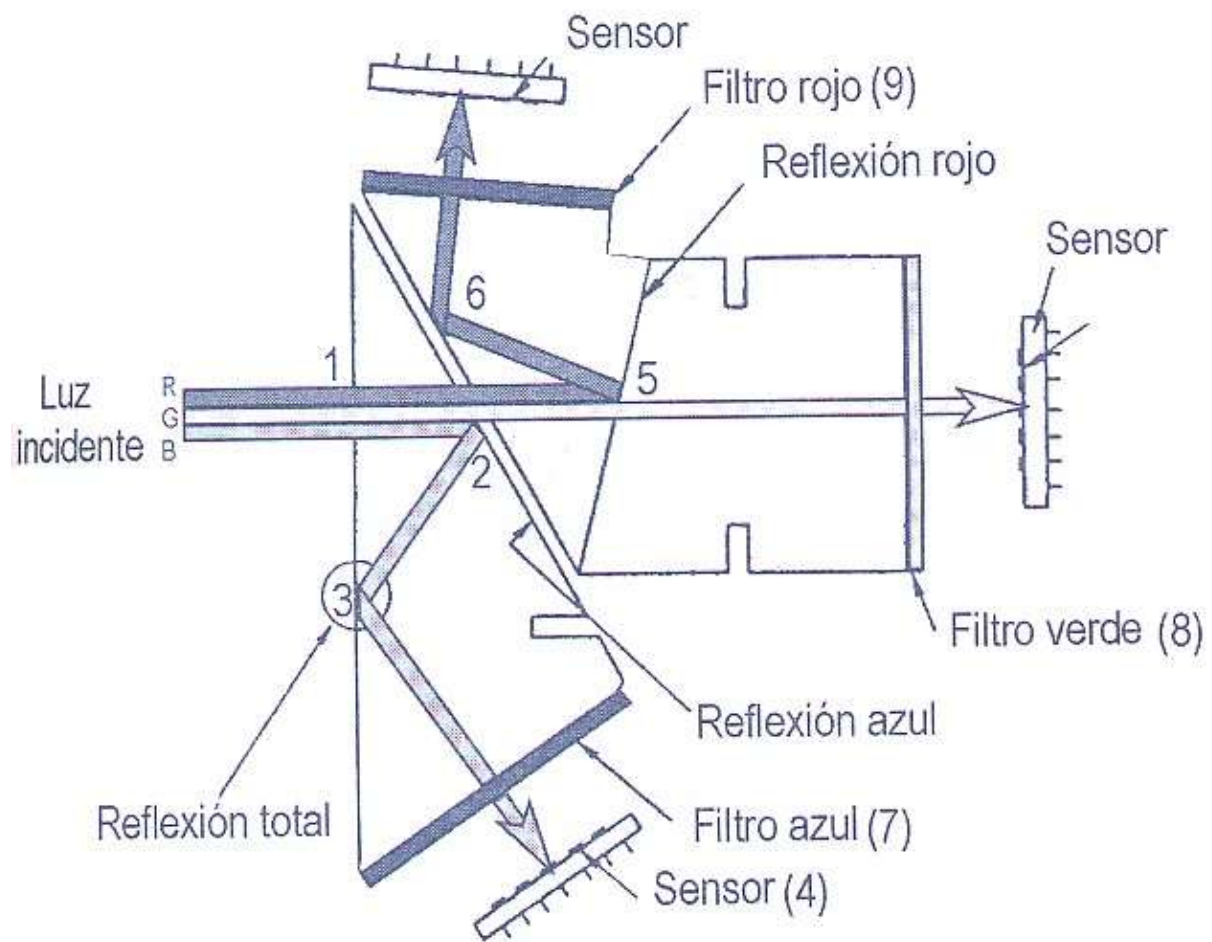
Existen el doble de detectores de verde que de rojo y de azul.

Razón: el ojo humano es mucho más sensible al verde que a los otros colores.



duncantech.com

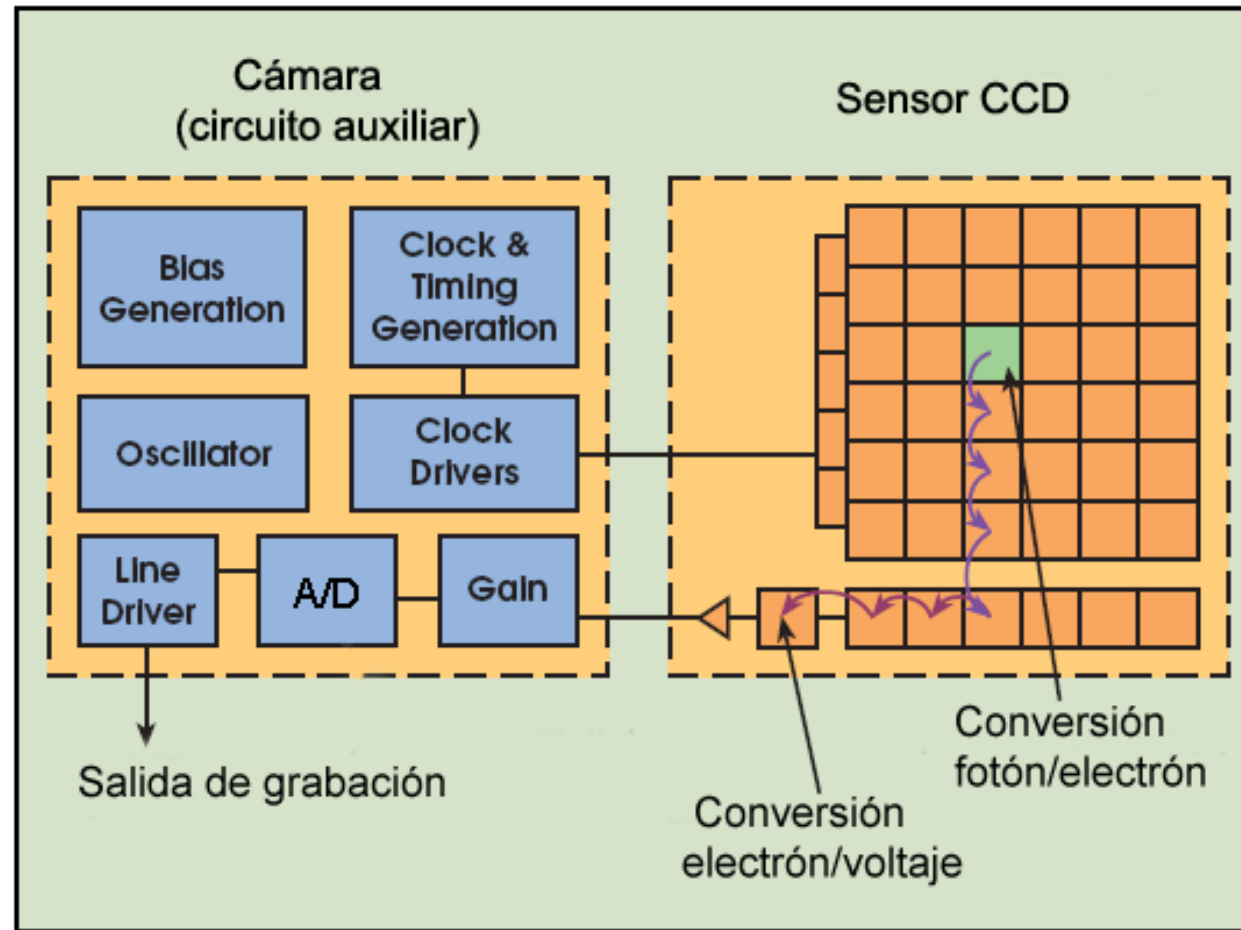
Cámaras con 3 CCD



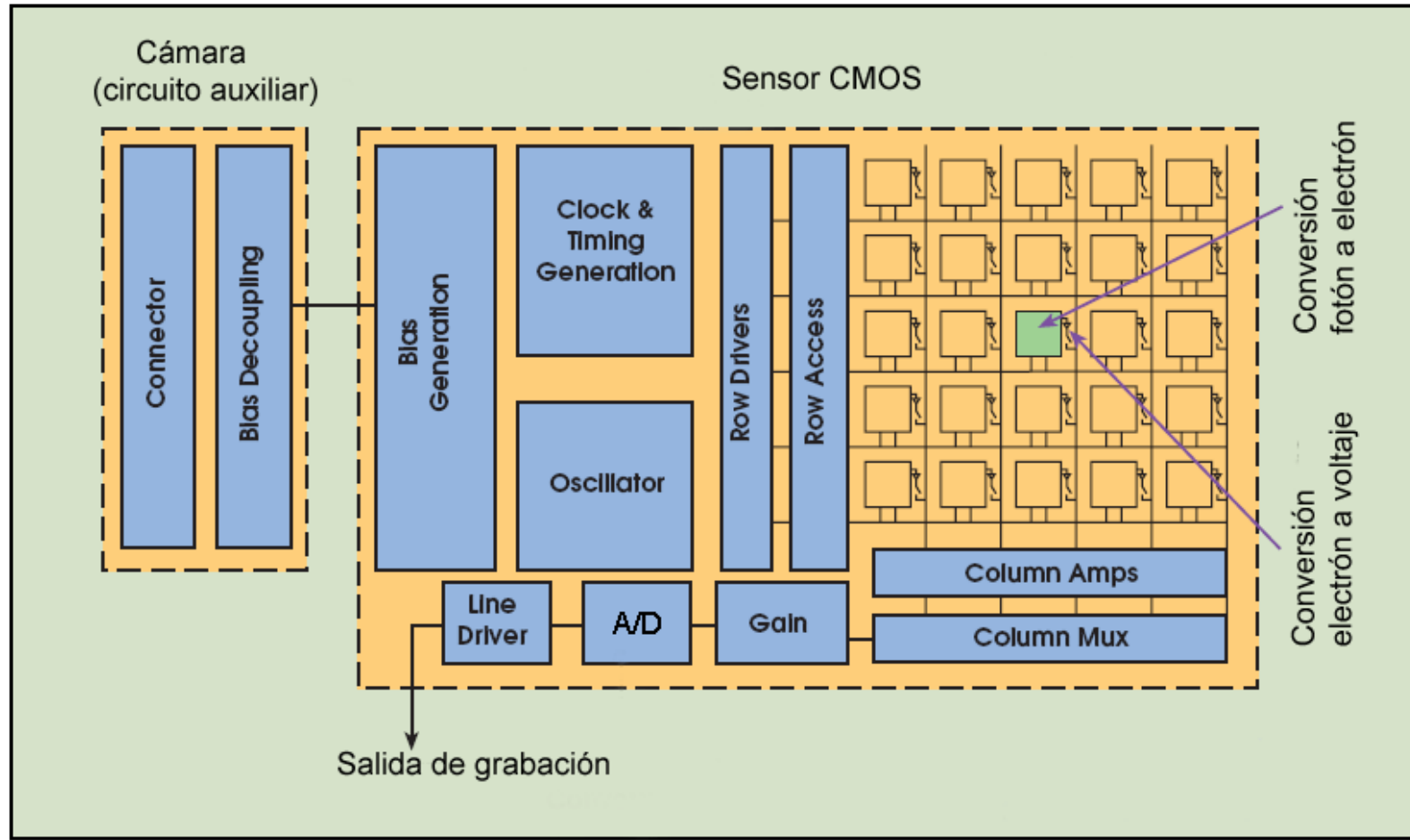
Dos tecnologías maduras diferentes:

- En una **CCD** cada píxel transfiere su carga a un nodo, se convierte en voltaje (señal analógica de salida).
- En un **CMOS** cada píxel tiene su conversor carga-voltaje. El sensor incluye toda la electrónica necesaria (amplificadores, correctores de ruido, digitalización) para que la salida sea señal digital. Esto aumenta la complejidad y reducen el área disponible para la captura de luz.

Esquema CCD



Esquema CMOS



CCD vs. CMOS

- CCD mejor calidad de la imagen => aplicaciones científicas e industriales
- CMOS más integración (sistemas de menor tamaño) y menor disipación de potencia. Más velocidad por la posibilidad de paralelizar
- Actualmente son tecnologías complementarias



CCD



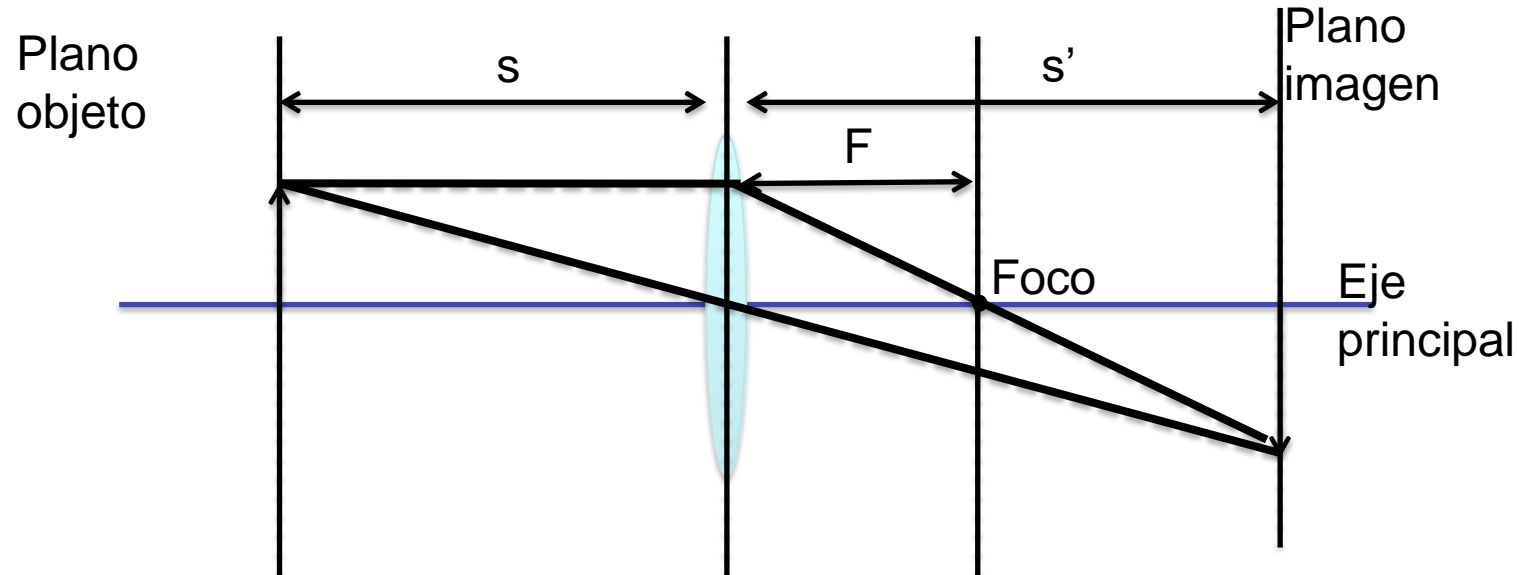
CMOS

Fundamentos de Óptica

Lente delgada

Apertura, profundidad de campo, ángulo de visión ...

Fundamentos de óptica



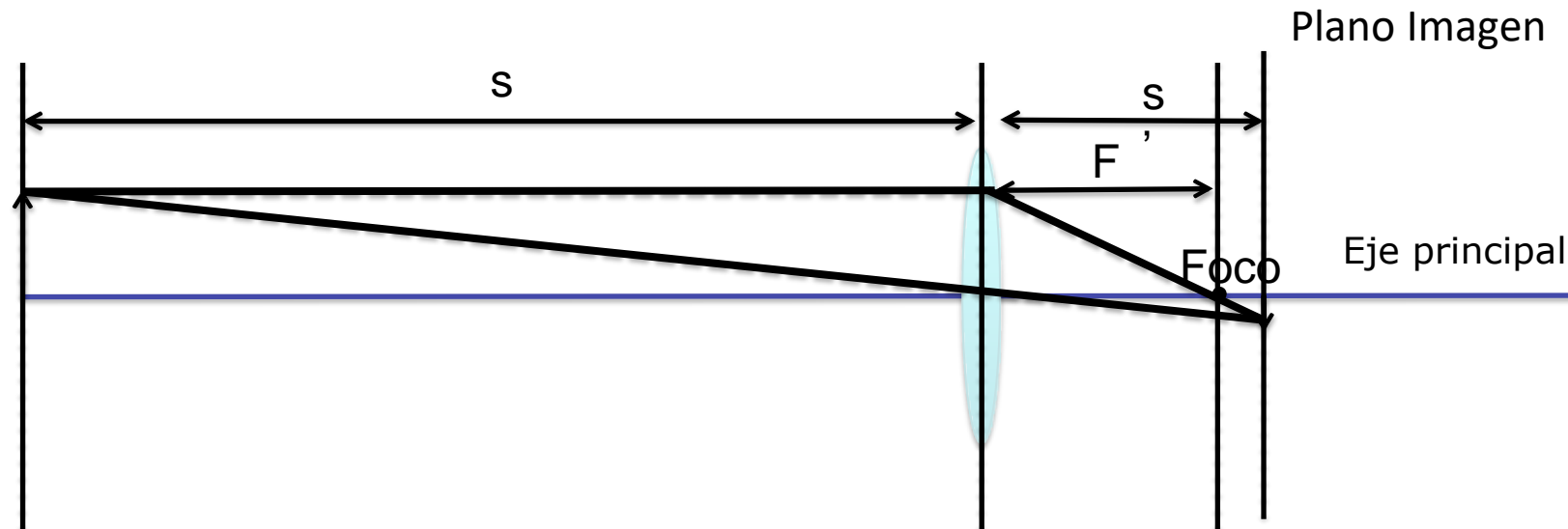
Centro óptico: punto sobre el eje principal, dentro de la lente donde los rayos no se desvían

Eje Principal: recta que une los centros de curvatura de la lente

Distancia Focal (F): distancia del foco principal al centro óptico

Foco Principal: punto al que concurren los rayos que llegan paralelos a la lente

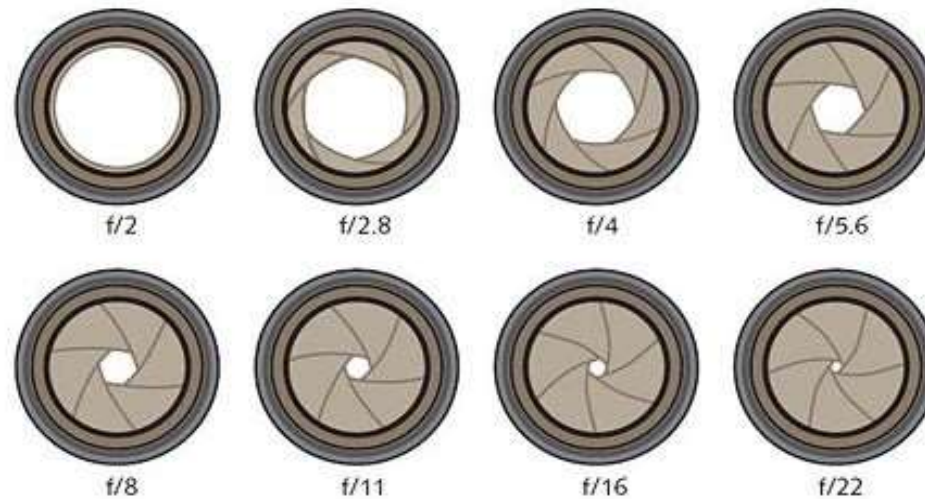
- Si s aumenta (el objeto se aleja), s' disminuye.
- Su posición mínima es la distancia focal (si $s = \infty$, $s' = F$)



- Distancia focal (F)
 - Distancia entre el centro óptico de la lente y el foco.
 - La luz de un objeto en el infinito intercepta el eje óptico en el pto. F. En cámaras CCD el sensor se situa en este punto focal de la lente.
 - Las distancias focales de las cámaras digitales son más pequeñas que las de las analógicas porque el sensor es más pequeño que la película de 35 mm.
- Apertura
 - Medida de la máxima cantidad de luz que pasa a través de la óptica. Se controla con el diafragma de diámetro variable. Se caracteriza por el número f.
 - Números f : $f/1$, $f/1.4$, $f/2$, $f/2.8$, $f/4$, $f/5.6$
 - Al pasar de un número-f al siguiente se disminuye a la mitad el área y por tanto el brillo.
 - Óptica grande \Rightarrow diámetro grande \Rightarrow apertura grande \Rightarrow menor número f \Rightarrow precio elevado

- **Apertura**

- Partiendo de un diámetro = 1, la sucesión típica es: 1, 1.4, 2, 2.8, 4, 5.6, 8, 11, 16, 22



Fundamentos de Óptica

Montura C:

Más común en cámaras CCD para aplicaciones industriales

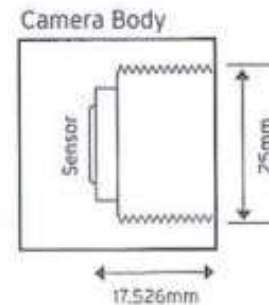
Montura CS:

Para cámaras de vigilancia.

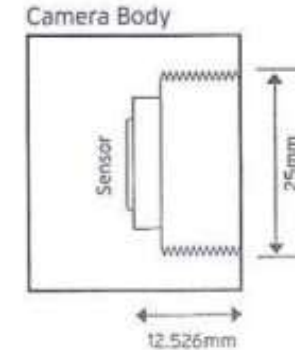
Montura F Nikon y montura 42 mm:

Para cámaras con sensores de gran formato, lineales y matriciales

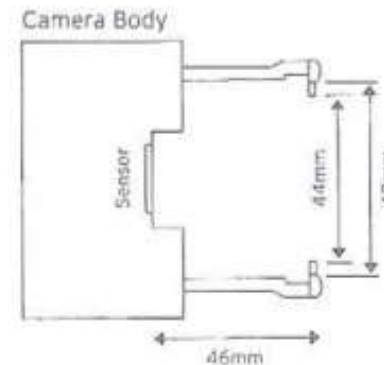
C- MOUNT



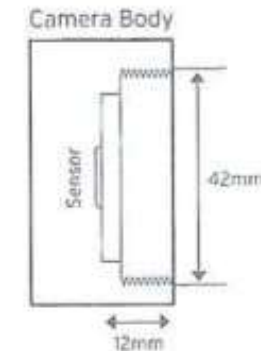
CS- MOUNT



NIKON F- MOUNT



42mm THREAD

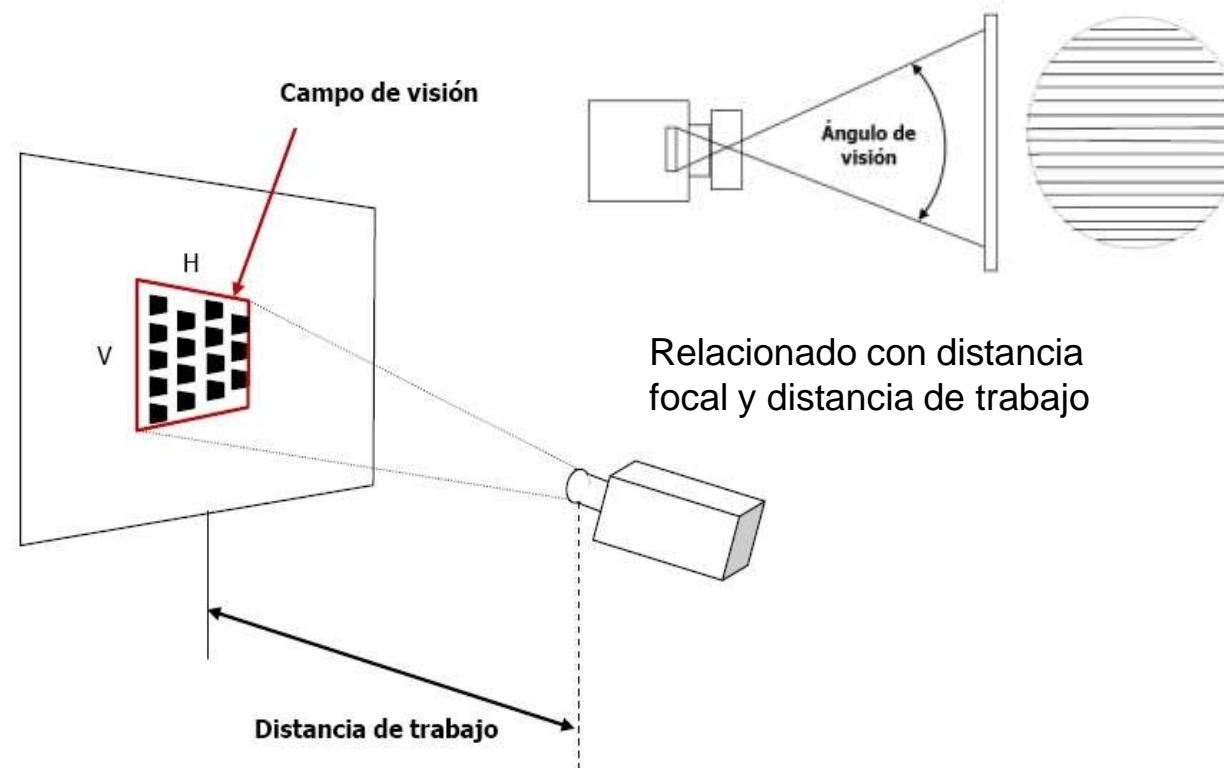


- **Profundidad de Campo**

- Mide la distancia en que los objetos aparecen enfocados.
- Depende del tamaño de la lente.
- A menor tamaño mayor es la profundidad de campo, pero más luz se necesita para ver el objeto.
- Aumenta cuando disminuye la apertura



- **Campo de Visión**, ángulo de visión y distancia de trabajo



• Resolución

- Capacidad del sistema para reproducir los detalles de un objeto. Es la porción más pequeña del objeto que se puede distinguir y procesar por el sistema.
- Resolución (lp/mm) = $500 / \text{tamaño del píxel } (\mu\text{m})$
- Por ejemplo, si el píxel tiene un tamaño de 10 μm , la resolución es de 50 pares de línea por mm (caben 100 líneas en un mm)
- Resolución de una pantalla: nº líneas horizontales y verticales.

VGA	640	x	480
SVGA	800	x	600
XGA	1024	x	768
SXGA	1280	x	1024
UXGA	1600	x	1200
Full HD	1920	x	1080
Quad HD	3840	x	2160
4K	4096	x	2160



Fundamentos de Óptica

Distorsión

- Cambio en la geometría de un objeto en el plano de la imagen
Ejemplo: un rectángulo puede aparecer con los lados hacia dentro (distorsión almohadilla) o hacia fuera (distorsión barril).
- La distorsión de una lente se expresa en %.
- Causada por la difracción de la luz en la lente (superficie curvada). Es máxima en los bordes.
- Es más importante en las ópticas gran angular (con baja distancia focal).
- Se puede minimizar utilizando un sistema de lentes múltiples

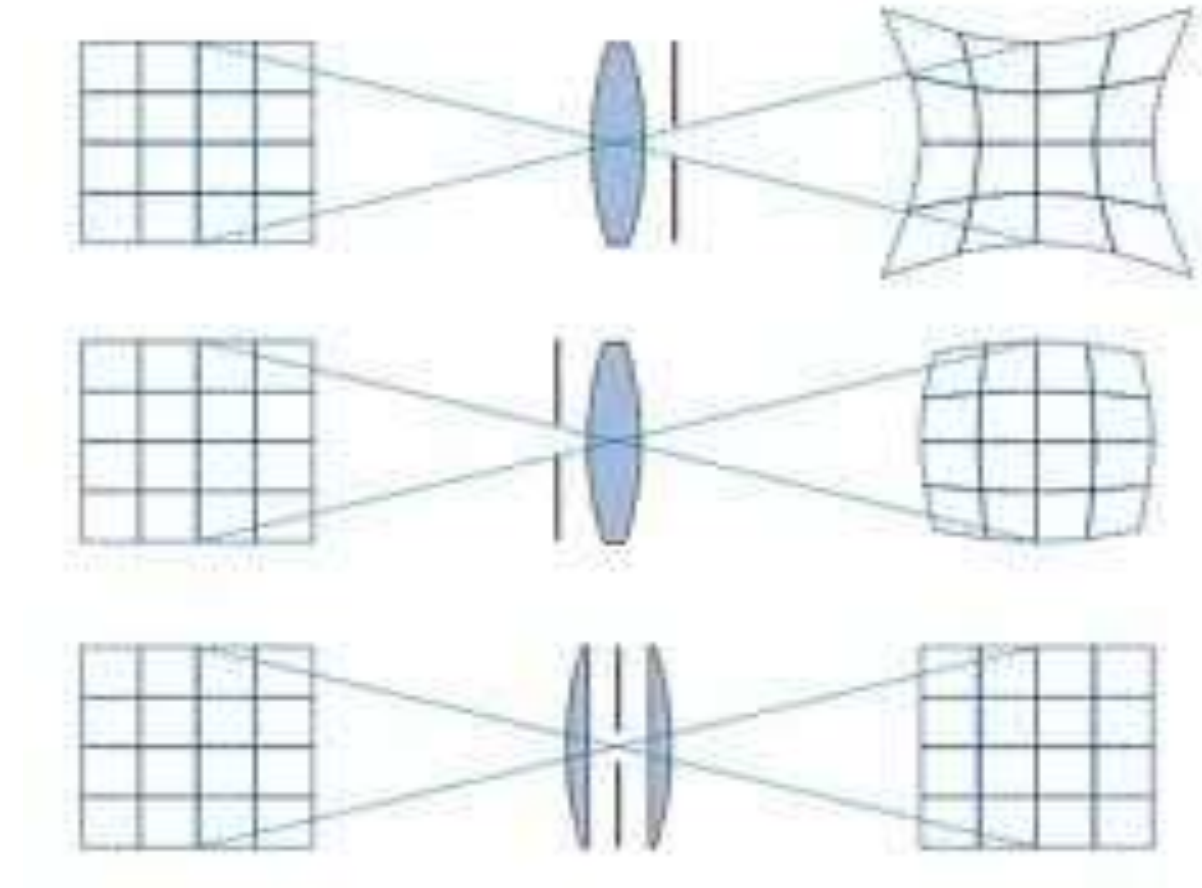
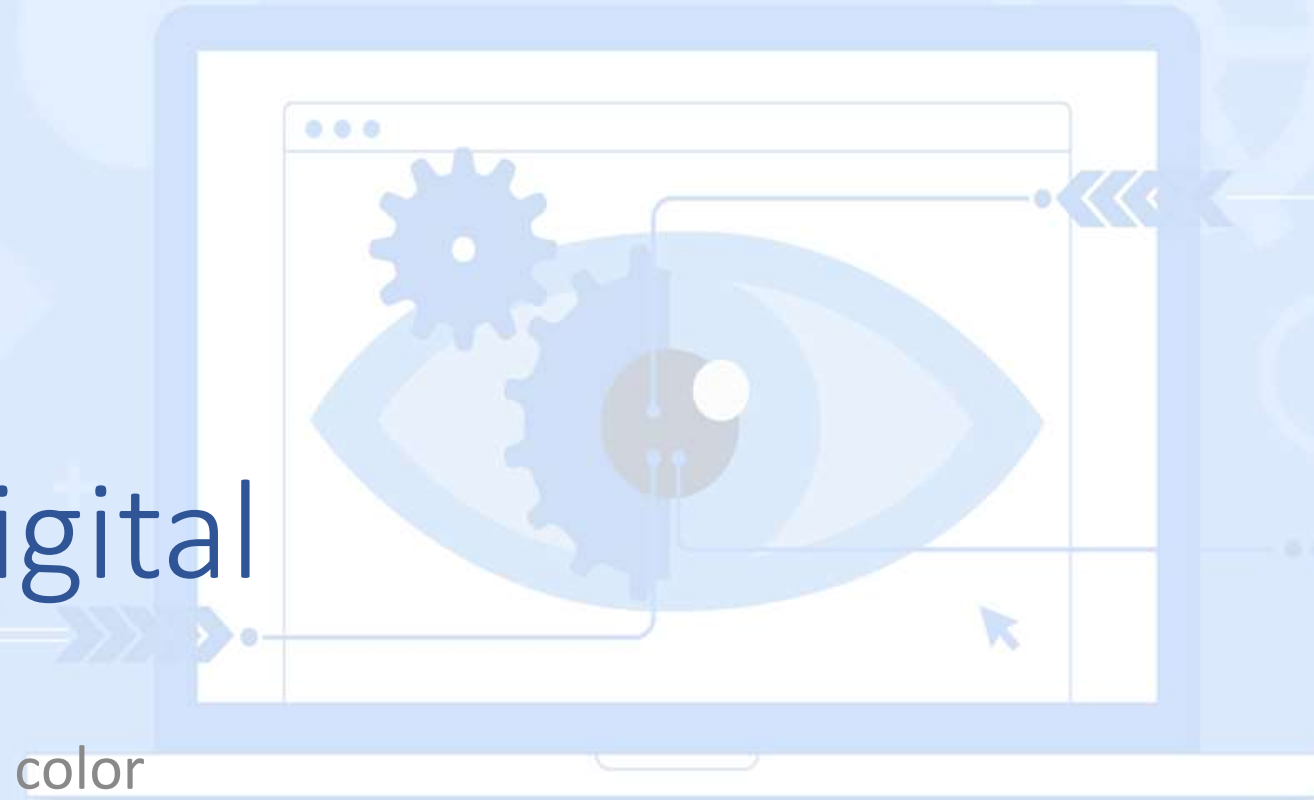


Imagen Digital

Pixel

Resolución, planos de color

Profundidad de color, ...

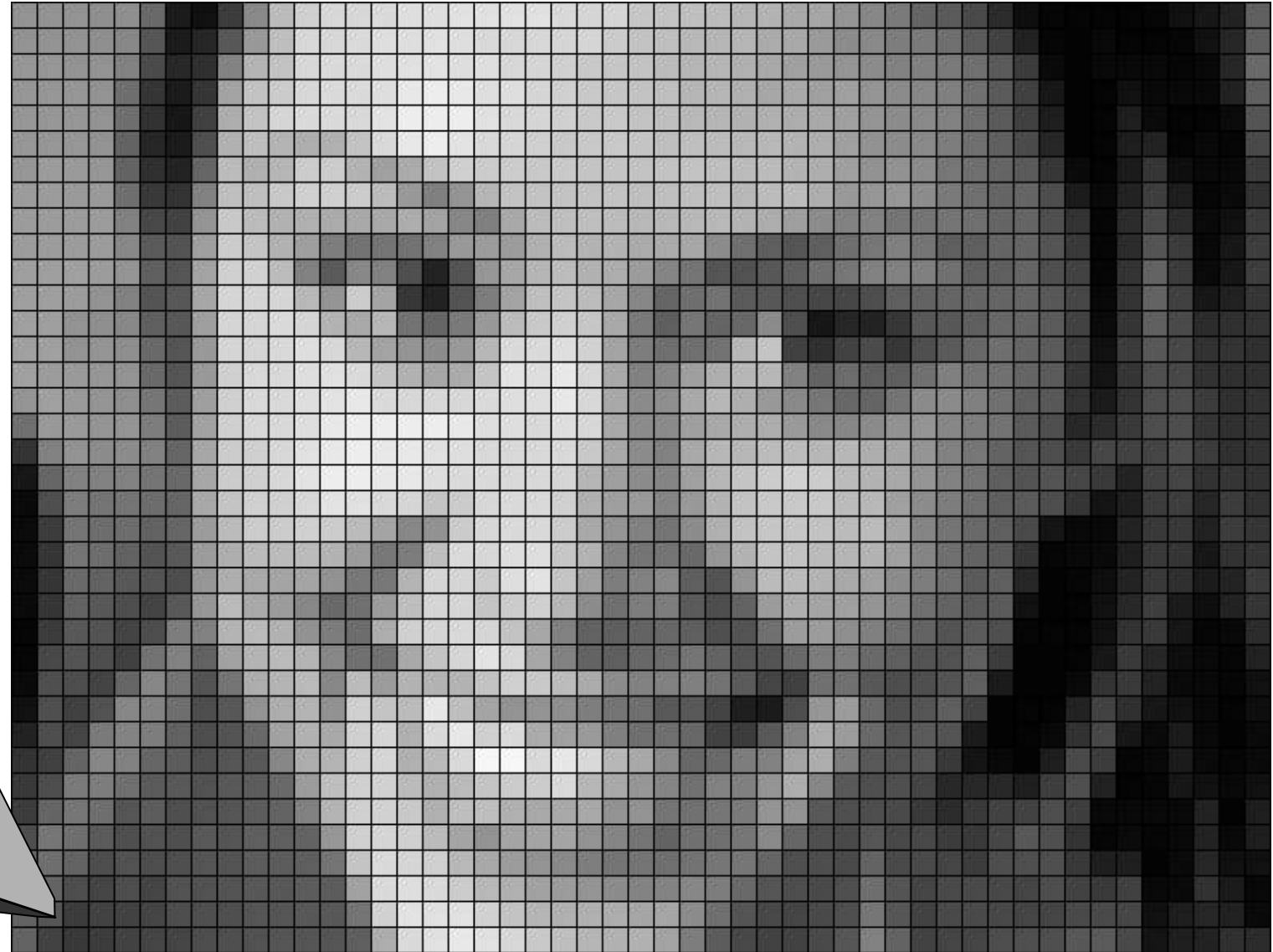
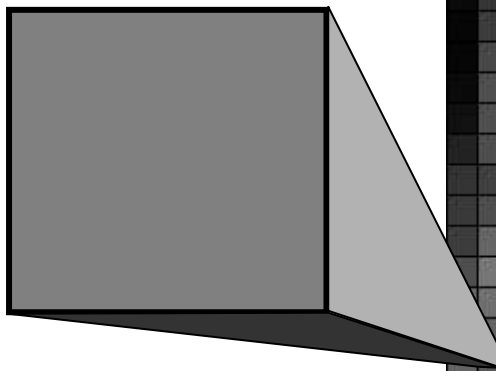


- Una imagen digital es una matriz, o array bi-dimensional, de números.
- Cada celda de la matriz es un PIXEL (Picture Element).
- Ejemplo. Imagen de 20x15.

Un PIXEL

90	67	68	75	78	98	185	180	153	139	132	106	70	80	81	69	69	67	35	34
92	87	73	78	82	132	180	152	134	120	102	106	95	75	72	63	75	42	19	29
63	102	89	76	98	163	166	164	175	159	120	103	132	96	68	42	49	46	17	22
45	83	109	80	130	158	166	174	158	134	105	71	82	121	80	51	12	50	31	17
39	69	92	115	154	122	144	173	155	105	98	86	82	106	83	76	17	29	41	19
34	80	73	132	144	110	142	181	173	122	100	88	141	142	111	87	33	18	46	36
37	93	88	136	171	164	137	171	190	149	110	137	168	161	132	96	56	23	48	49
66	117	106	147	188	202	198	187	187	159	124	151	167	158	138	105	80	55	59	54
127	136	107	144	188	197	188	184	192	172	124	151	138	108	116	114	84	46	67	54
143	134	99	143	188	172	129	127	179	167	106	118	111	54	70	95	90	46	69	52
141	137	96	146	167	123	91	90	151	156	121	93	78	82	97	91	87	45	66	39
139	137	80	131	162	145	131	129	154	161	158	149	134	122	115	99	84	35	52	30
137	133	56	104	165	167	174	181	175	169	165	162	158	142	124	103	67	19	31	23
135	132	65	86	173	186	200	198	181	171	162	153	145	135	121	104	53	14	15	33
132	132	88	50	149	182	189	191	186	178	166	157	148	131	106	78	28	10	15	44

Un píxel



- **Intensidad:** Nivel de luz que recibe cada pixel, se expresa en niveles de gris, y normalmente son $2^8 = 256$ [0, 255] niveles de gris.
- En algunas aplicaciones, sobre todo médicas se utilizan imágenes de más niveles de gris $2^{10} = 1024$ [0,1023] o incluso $2^{12} = 4096$ [0,4095]



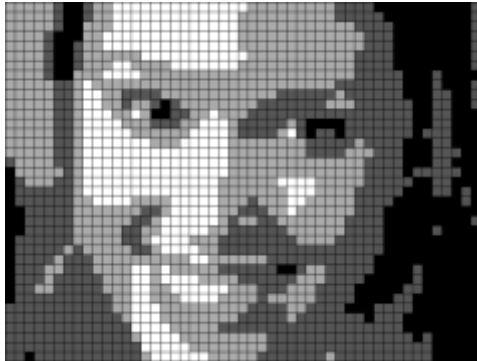
- Hi-color: método reducido para representar colores

1 píxel = 2 bytes

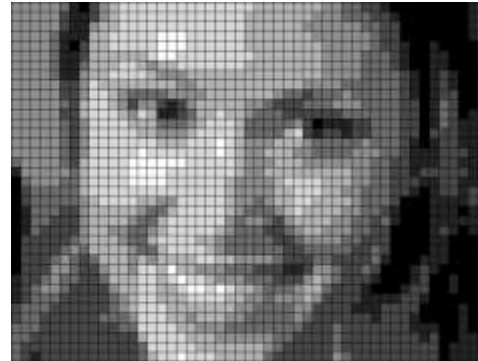
5 bits por cada color (Rojo, Verde, Azul)

- Imágenes en punto flotante: útiles en procesos intermedios

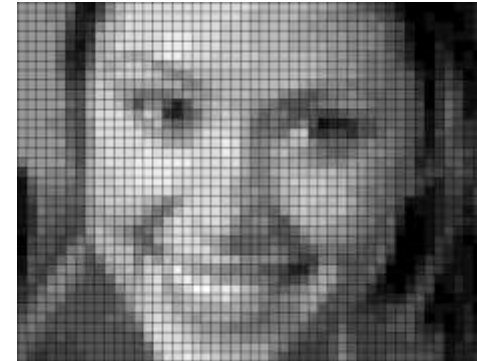
1 píxel = 1 float o un double



2 bits por píxel



3 bits por píxel



4 bits por píxel

- Nomenclatura

- Nº de columnas de la matriz:
ancho de la imagen (width).
- Nº de filas de la matriz:
alto de la imagen (height).
- Eje horizontal: eje x.
- Eje vertical: eje y.

Normalmente el tamaño
de la imagen se
expresa como:
ancho x alto



- Cuando los píxeles representan magnitudes en distintos dominios físicos, decimos que la imagen es multicanal.
- Ejemplo. Imagen en color → Imagen con 3 canales:



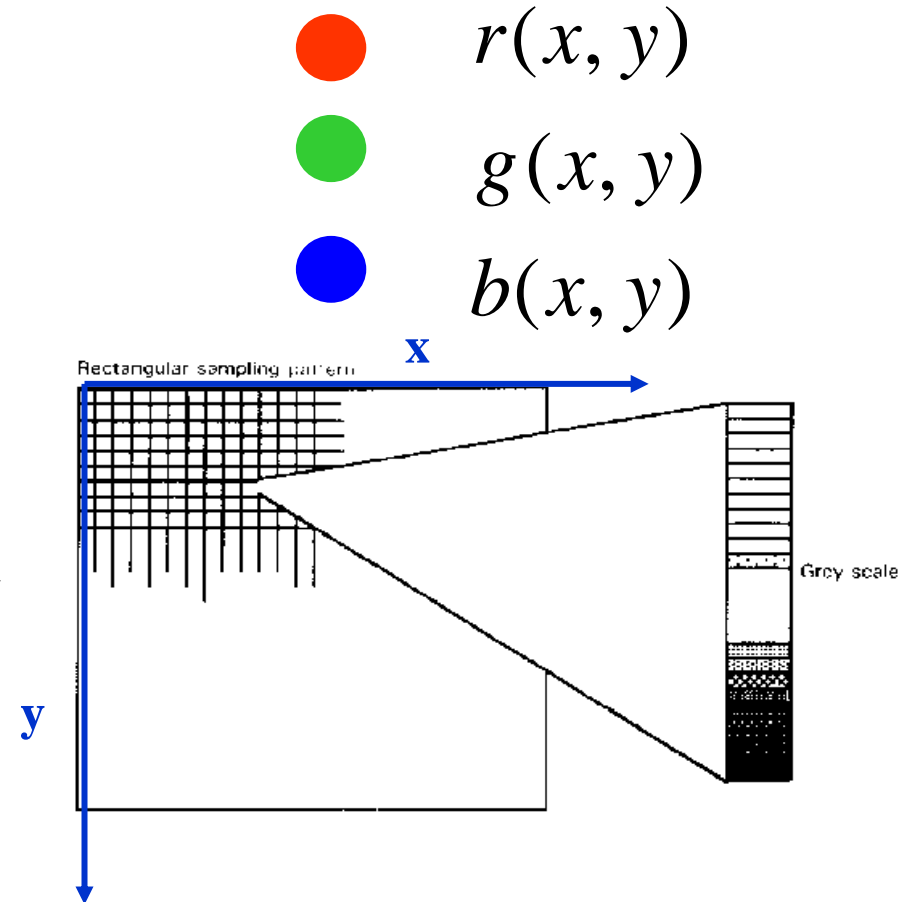
canal R (rojo)

canal G (verde)

canal B (azul).

Concepto de Imagen

$$ZxZ \xrightarrow{f(x,y)} Z$$



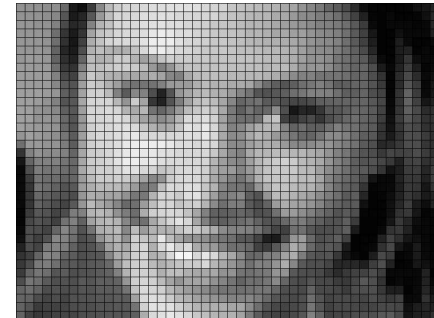
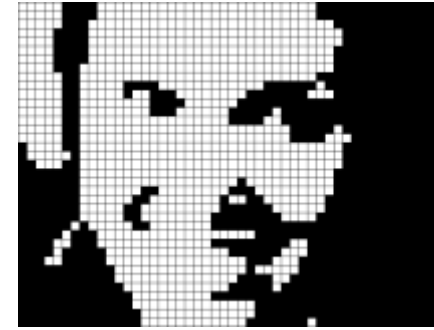
Muestreo: Digitalización de las coordenadas espaciales (x,y)



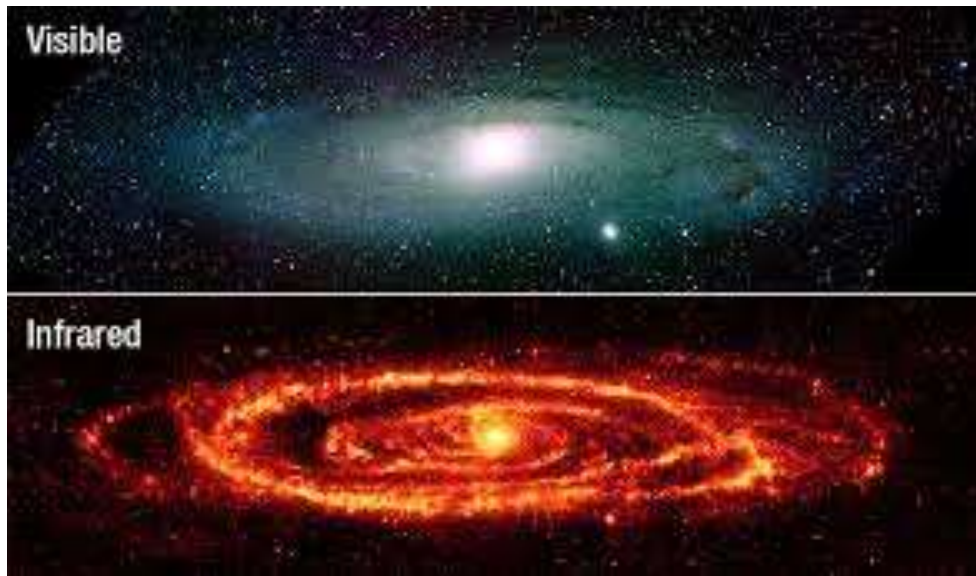
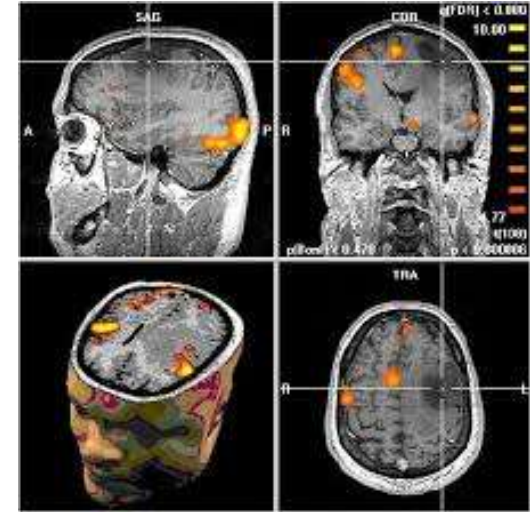
Cuantificación: Digitalización del nivel de luminosidad $f(x,y)$

¿De qué tipo de datos es cada celda de la matriz?

- Imagen **binaria**:
1 píxel = 1 bit
- 0 = negro; 1= blanco
- Imagen en **escala de grises**:
1 píxel = 1 byte
- Permite 256 niveles de gris
- 0 = negro; 255 = blanco
- Imagen en **color**:
1 píxel = 3 bytes
- Cada píxel consta de 3 valores:
(Rojo, Verde, Azul)
- Un byte por color
- 16,7 millones de colores posibles



- Pero no siempre usamos el espectro visible, en algunas ocasiones es interesante utilizar imágenes con otros “colores”



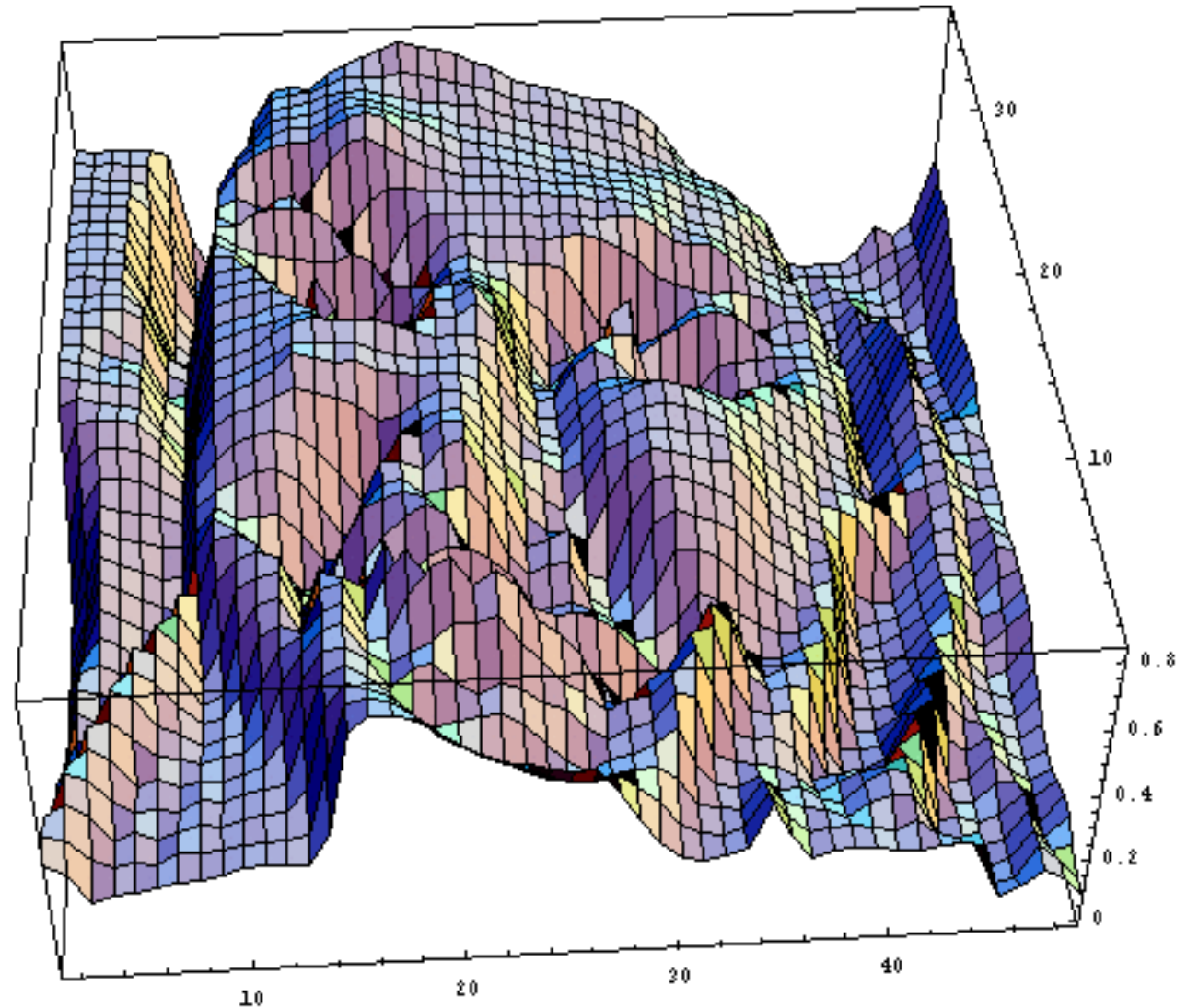
as it looks to the eye

in ultraviolet (spurious-color)

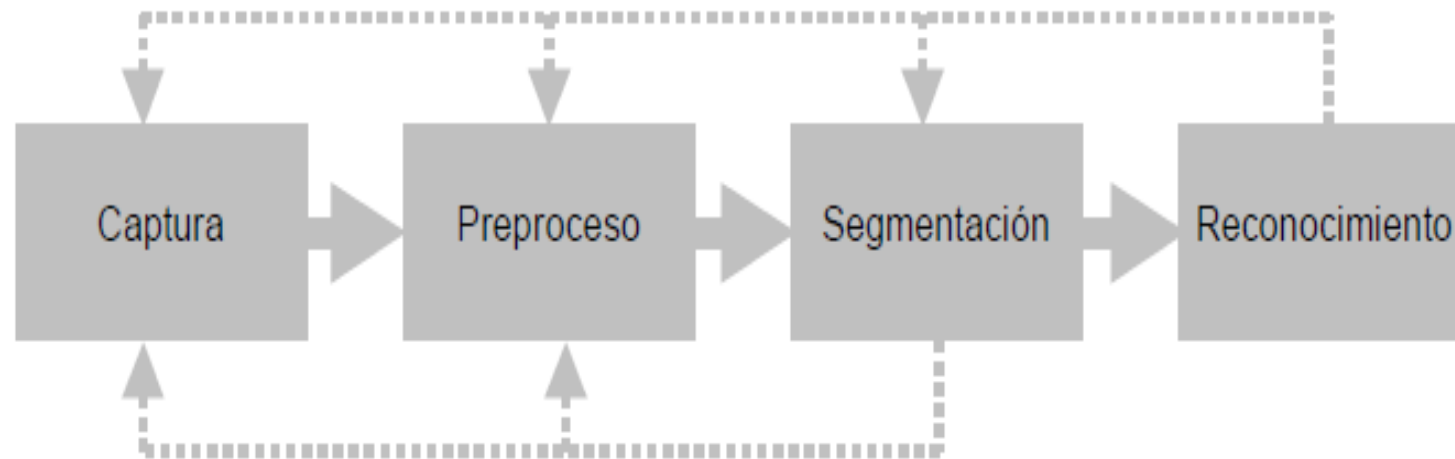
proper B&W appearance in ultraviolet

Una imagen puede ser interpretada como una superficie bidimensional.

En el plano horizontal se representa la posición (x,y) del pixel, y en el plano vertical z , se representa el nivel de gris del mismo.



Etapas de un sistema de visión artificial



Tratamiento de Imagen

Matrices, operaciones globales

Operaciones geométricas

Transformaciones lineales, umbralización

Filtros, convoluciones

- Imagen digital = **Matriz numérica**
- Tratamiento de la Imagen

90	67	68	75	78
92	87	73	78	82
63	102	89	76	98
45	83	109	80	130
39	69	92	115	154

- Por tanto podemos hacer todas las operaciones que podemos hacer con matrices, como sumar, restar, multiplicar, dividir, and, or, máximo, mínimo, derivadas, etc ...

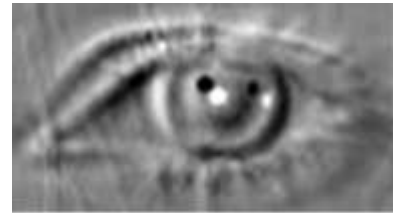
- Grupos de procesamiento



Operaciones **globales**: Cada pixel es tratado de forma independiente

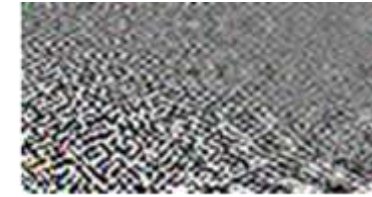


Filtros y Convoluciones: Se toman en consideración los píxeles vecinos.



Transformaciones **geométricas**:
Modifican el tamaño y forma de las imágenes.

Transformaciones **lineales**:
Operaciones más complejas como Transformada de Fourier, wavelets, etc.



- Operaciones de procesamiento global
 - **Aritméticas:** sumar, restar, multiplicar, máximo, etc.
 - Unarias: una sola imagen y un valor constante.
 - Binarias: con dos imágenes.
 - **Booleanas:** and, or, not, etc.
 - Unarias: una sola imagen y una constante.
 - Binarias: con dos imágenes.
 - **Otras transformaciones generales:**
 - Transformaciones de histograma.
 - Transformaciones de color.
 - Binarización, etc.
- Cada operación tendrá un significado, utilidad y aplicaciones específicos.

$$R(x,y) := f(A(x,y)), \forall (x,y)$$

- **Filtros y convoluciones:** el valor de un píxel depende de la vecindad local de ese píxel:

$$R(x,y) := f(A(x-k,y-k), \dots, A(x,y), \dots, A(x+k,y+k))$$

- **Transformaciones geométricas:** el valor de un píxel depende de píxeles situados en otras posiciones:

$$R(x,y) := A(f1(x,y), f2(x,y))$$

- **Transformaciones lineales:** el valor de un píxel puede depender de todos los píxeles de la imagen:

$$R(x,y) := f(A, x, y)$$

- Supongamos una imagen de entrada A y una imagen resultado R.
- Una operación global (píxel a píxel) se puede expresar como una función:

$$R(x,y) := f(A(x,y))$$

- Ejemplo. Invertir. $R(x,y) := 255 - A(x,y)$

Imagen A

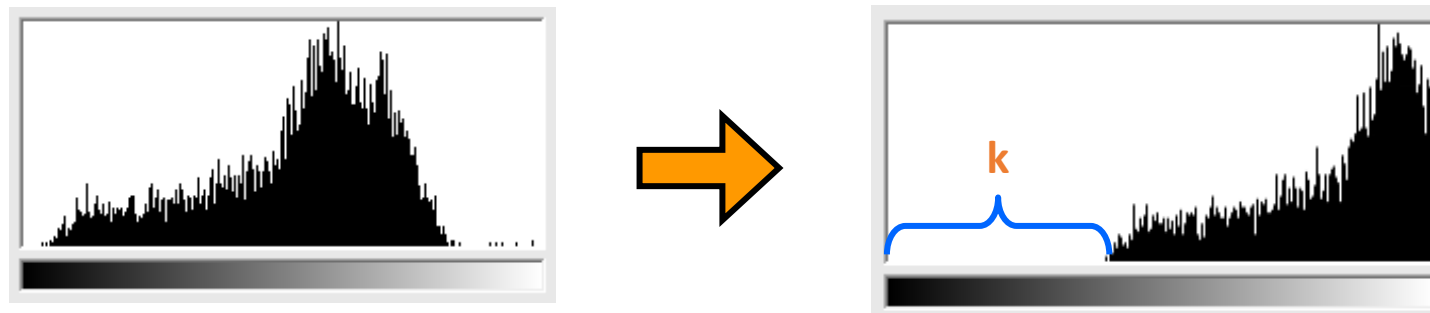


Imagen R



Operaciones unitarias con píxeles

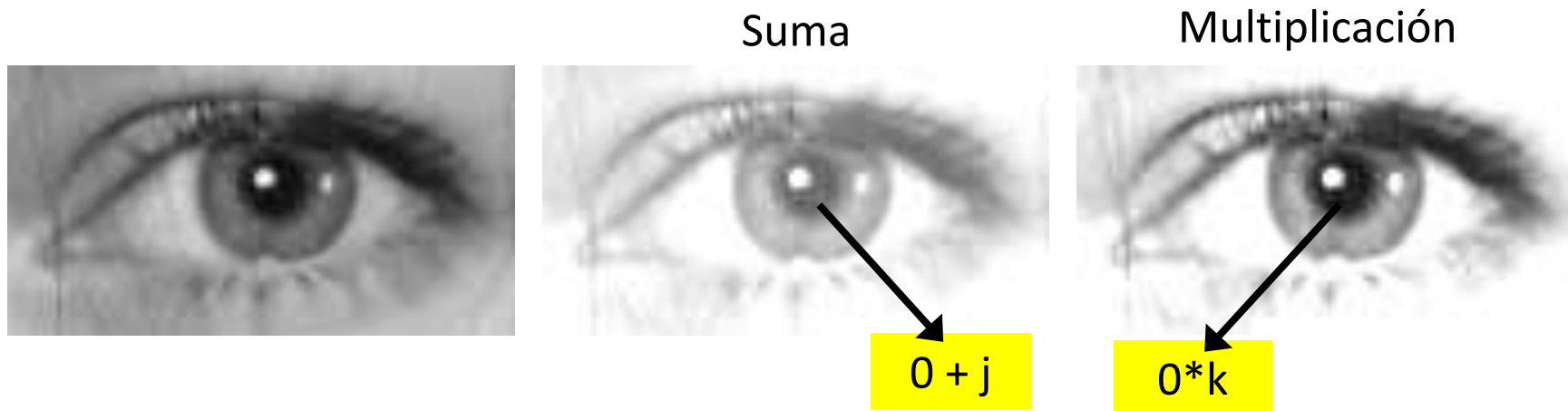
- **Sumar** una constante: $R(x,y) := A(x,y) + k$
Efecto: incrementa el brillo de la imagen
El histograma se desplaza a la derecha k píxeles
Hay que controlar el overflow $R(x,y) > 255$



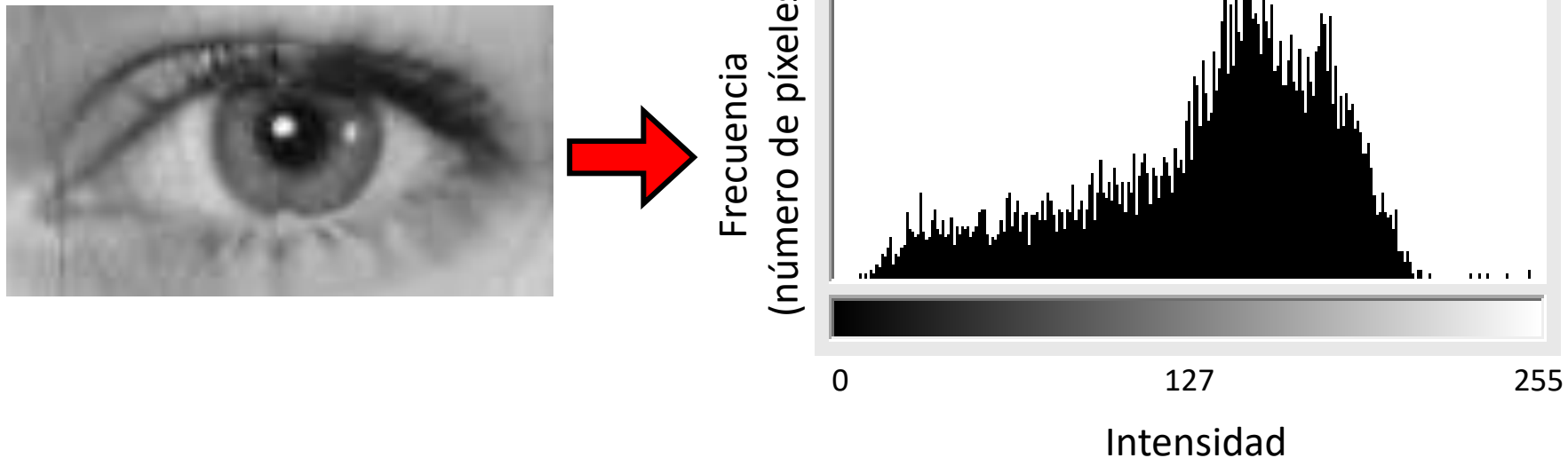
- **Multiplicar** por una constante: $R(x, y) := k \cdot A(x, y)$
Significado: aumentar la intensidad de la imagen en k .
El histograma se “estira” hacia la derecha.



- Tanto en la suma como en la multiplicación, se aumenta el nivel de gris de los píxeles, pero de forma distinta.
- En la suma, el parámetro j (entero) indica el número de niveles de gris a aumentar: de -255 a 255.
- En el producto, el parámetro k (real) indica el factor a multiplicar.
- $k=1 \rightarrow$ Ningún cambio
- $k=2 \rightarrow$ Se duplica el valor de gris. Los px >127 se saturan.
- $k=0,5 \rightarrow$ Se “encoge” a la mitad el histograma.



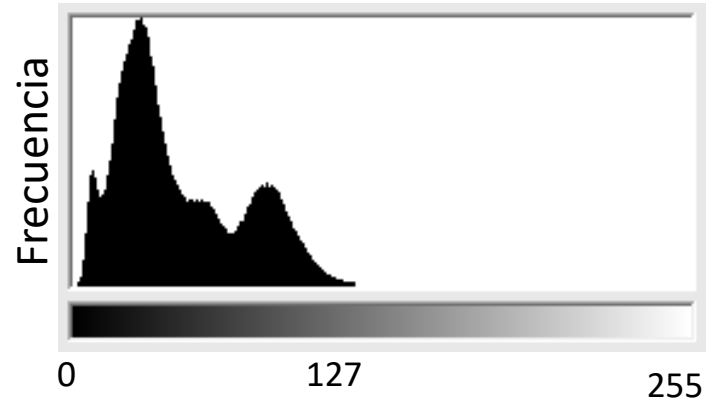
- Para comprender el significado de muchas transformaciones y saber cuál conviene aplicar se usan histogramas.
- Un **histograma** representa gráficamente una distribución de frecuencias.
- **Histograma de una imagen:** representa las frecuencias de los diferentes valores de gris en la imagen.



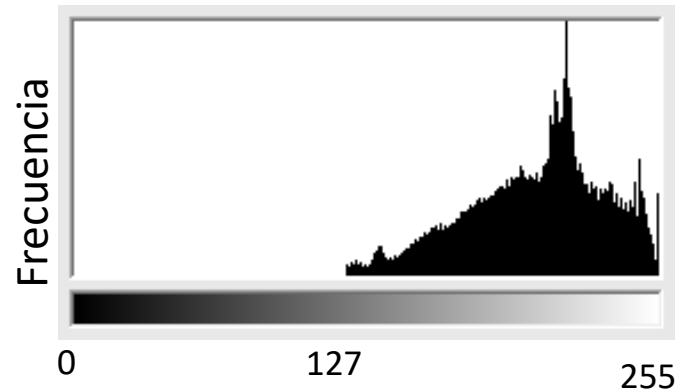
- Los histogramas son una herramienta importante en análisis de imágenes: ¿es buena la calidad de una imagen?, ¿sobra luz?, ¿falta contraste?
- Ayudan a decidir cuál es el procesamiento más adecuado para mejorar la calidad de una imagen...
- Tanto cualitativamente (qué operación aplicar),
- Como cuantitativamente (en qué cantidad).
- En principio, una buena imagen debe producir un histograma más o menos uniforme y repartido en todo el rango de valores.



Ejemplo 1. La imagen es muy oscura. Falta luz.



Ejemplo 2. La imagen es muy clara. Sobra brillo.

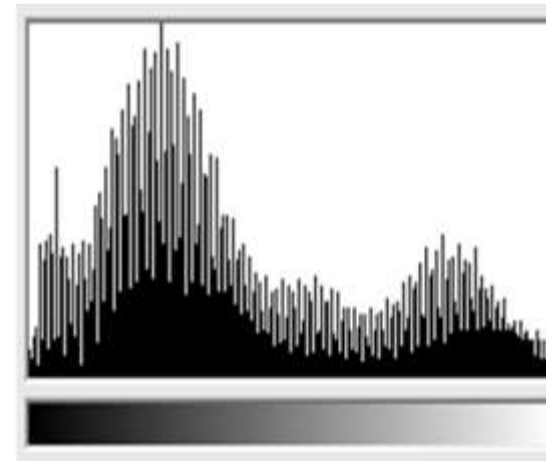


- **Histogramas** para mejorar la calidad de una imagen
- **Rango Dinámico**

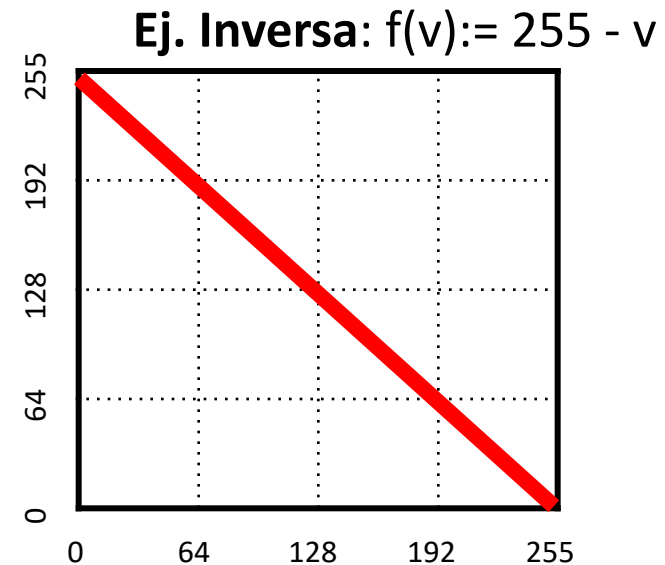
Problema. El histograma indica tonos muy oscuros.



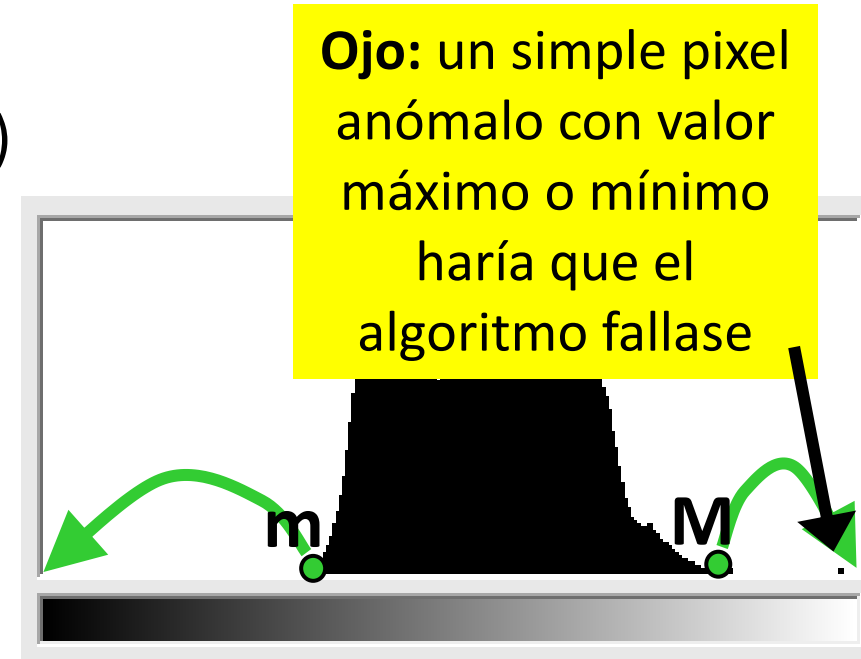
Solución.
Aplicar un operador de normalización del histograma.



- En general podemos definir una **transformación lineal** de la forma:
$$F(v) := a \cdot v + b$$
- Podemos también diseñar **transformaciones no lineales**: cuadráticas, polinomiales, exponenciales, etc ...
- Para saber cual usar, nos basamos en el histograma



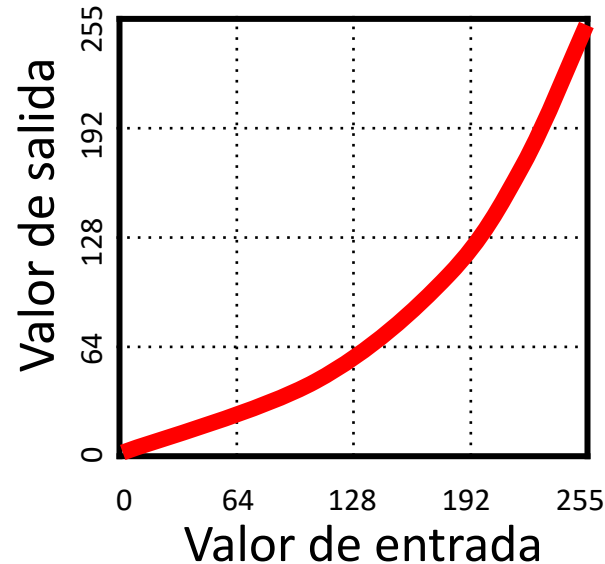
- Normalmente interesa que el histograma ocupe todo el rango de valores (aplicaciones para el ojo humano)
- **Ejemplo:** Un estiramiento de histograma
 - Obtener el mínimo m
 - Obtener el máximo M
 - $F(v) := (v-m)*255/(M-m)$



- La **transformación de histograma** puede tomar cualquier forma (no necesariamente lineal).

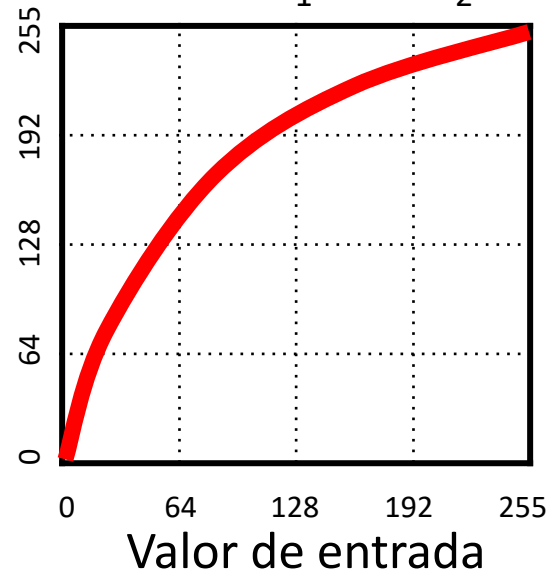
Ejemplos:

Parábola: $c_1v^2 + c_2v + c_3$



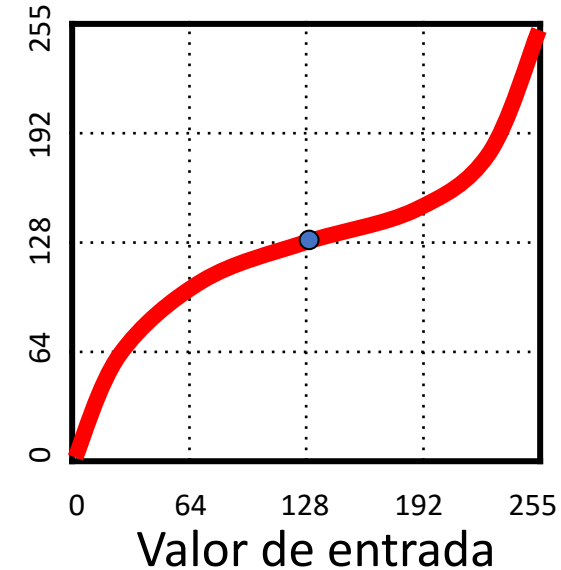
Resultado: oscurecer los medios tonos.

Raíz: $c_1v^{0.5} + c_2$



Resultado: aclarar los medios tonos.

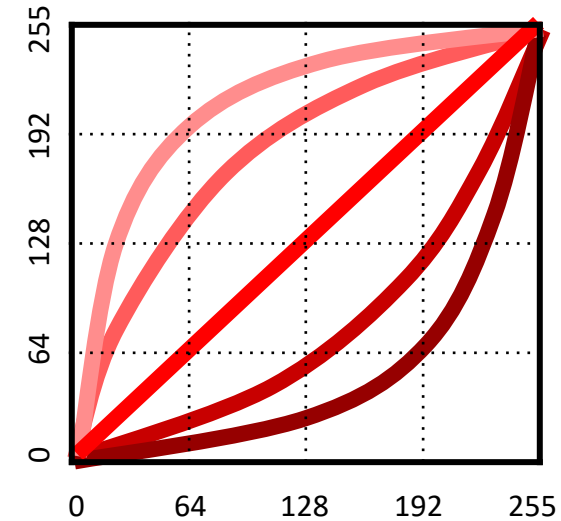
**Dos trozos de curva
(parábola y raíz)**



Resultado: aclarar tonos oscuros y oscurecer los claros.

- Elevar a 2, elevar a $\frac{1}{2}$...
- **Transformación de gama:**

$$f(v) := 255 \cdot (v/255)^{1/\text{GAMA}}$$



Gama 0,5



Gama 0,75



Gama 1



Gama 2



Gama 4

- **Ecualización:** La idea es que haya un número igual de píxeles en cada segmento de histograma.

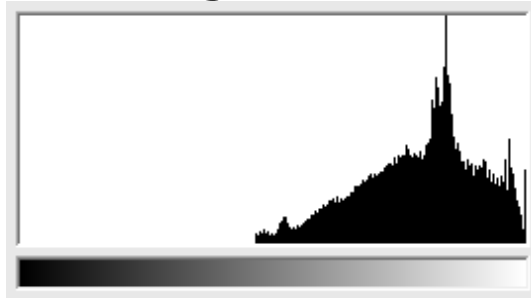
Imagen de entrada (A)



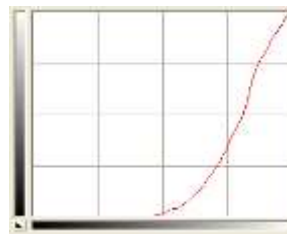
Imagen ecualizada (R)



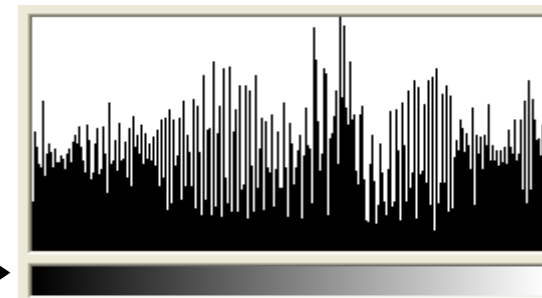
Histograma de A



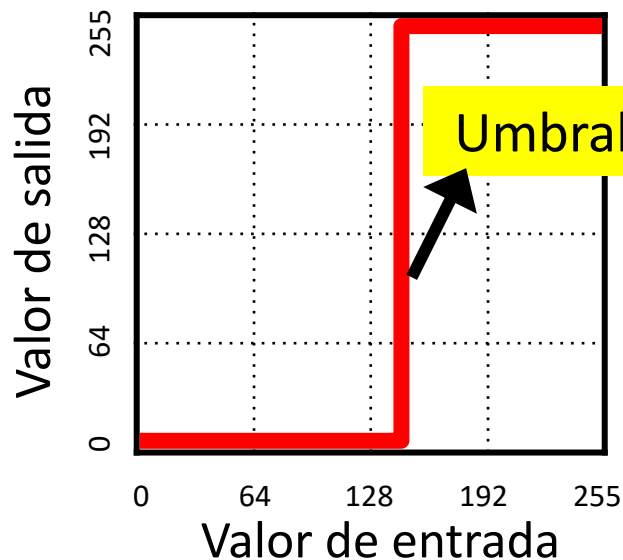
Función f



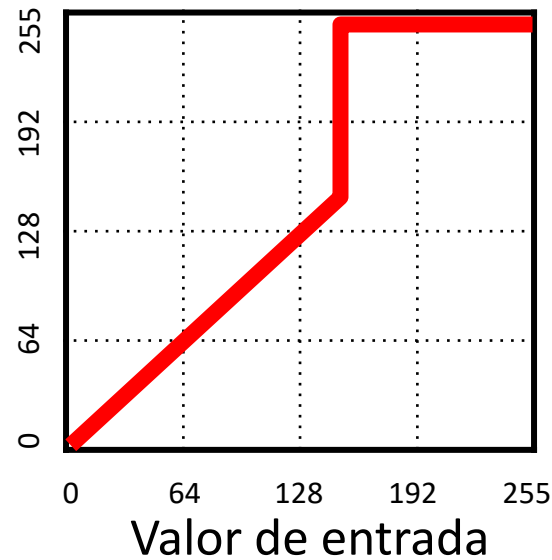
Histograma de R



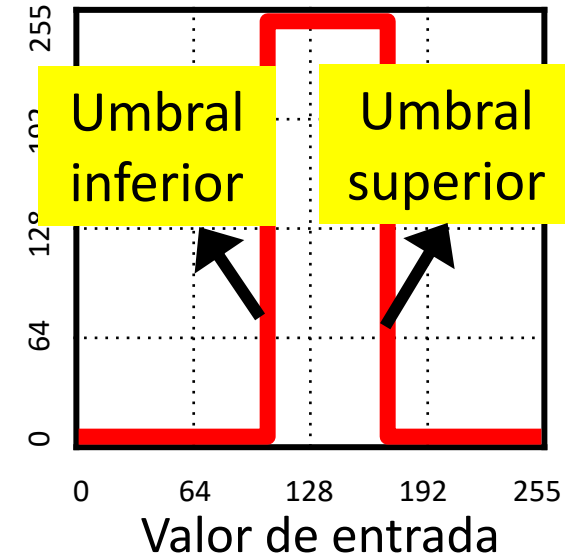
- **Umbralización:** Consiste en convertir la imagen en binaria, pero a partir de un valor umbral, o recortando cierto rango de valores.



Umbralizar la imagen
con valor cte.



Cortar un rango y
mantener el resto



Seleccionar un
rango

- Las funciones serán del estilo:
$$f(v) := \begin{array}{ll} \text{si } v > \text{umbral1} & \text{entonces } g(v) \\ \text{sino} & h(v) \end{array}$$
- Transformación de **binarización** (saturar a 0 ó 255).
$$f(v) := \begin{array}{ll} \text{si } v < \text{umbral} & \text{entonces } 0 \\ \text{sino} & 255 \end{array}$$

Ejemplo 1. La binarización se suele aplicar en OCR.

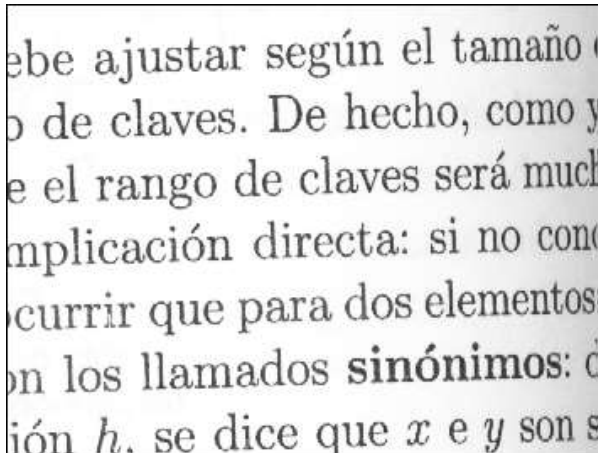
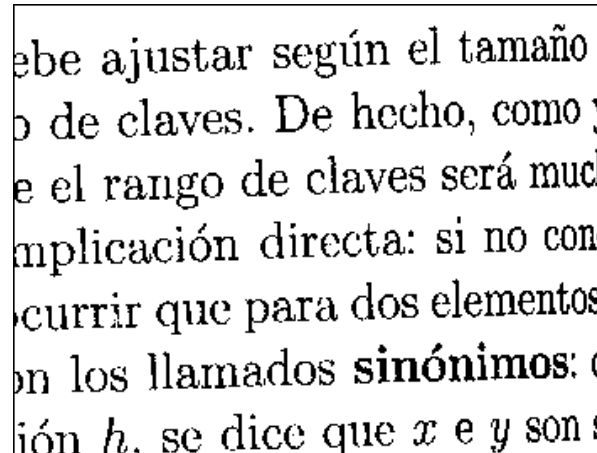
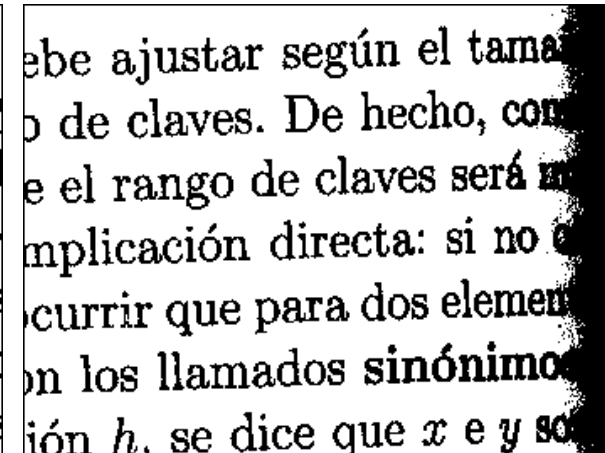


Imagen de entrada
(256 grises)



Umbral = 160



Umbral = 215

Ejemplo 2. Umbralización.



Imagen de entrada



- La separación del objeto del fondo se llama **segmentación**.
- La umbralización se puede usar para segmentar...
- ... aunque por sí sola no suele funcionar muy bien.

Repasamos:

- Una **transformación elemental** se puede ver desde distintas perspectivas:
 - Como una **función** unidimensional: $f: N \rightarrow N$
 - Como una **curva tonal**.
 - Como una **modificación del histograma**.
- La característica fundamental es que cada píxel se trata **independientemente** de los demás.
- Los histogramas son útiles para encontrar la transformación adecuada.
- En imágenes RGB, aplicamos la misma operación a los 3 canales para que se mantenga el color.

- **Combinación de imágenes:**

- utilizar dos o más imágenes de entrada para producir una imagen de salida.
- Entrada: imágenes A y B.
- Salida: imagen R.

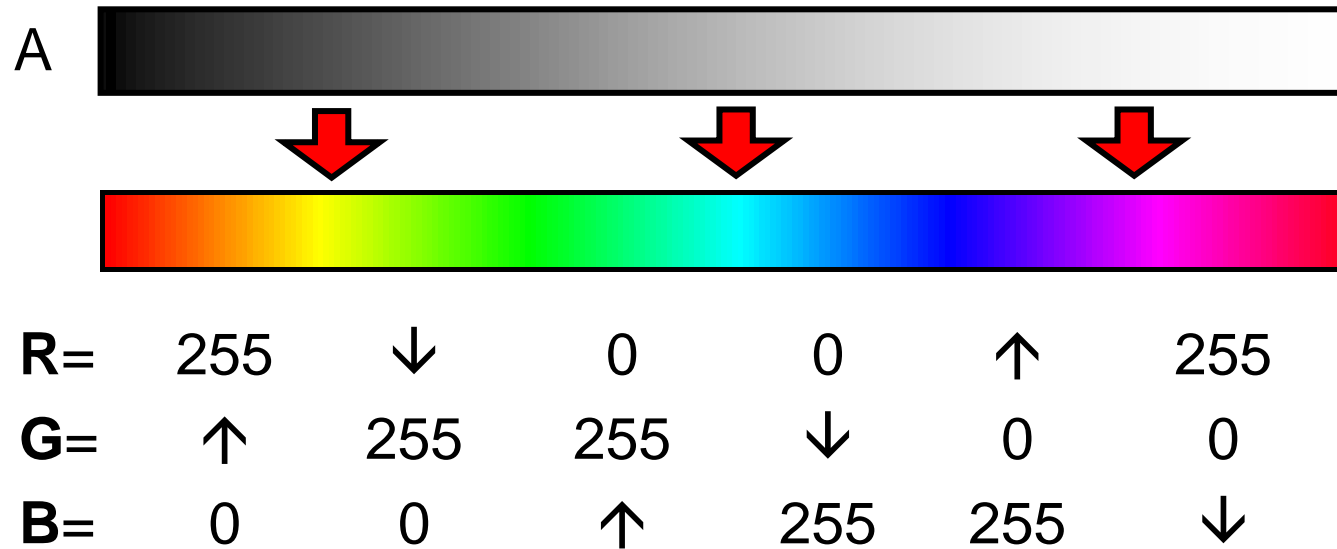
$$R(x, y) := f(A(x, y), B(x, y))$$

- Posibles operaciones de combinación:

- Booleanas: and, or, xor, not
- Aritméticas: suma, resta, producto/división, media
- Relacionales: máximo, mínimo

- En los puntos anteriores la transformación era la misma para todos los canales (R, G y B).
- Si es **distinta**, hablamos de **transformación de color**:
 - $R(x, y).R := f_1(A(x, y).R, A(x, y).G, A(x, y).B)$
 - $R(x, y).G := f_2(A(x, y).R, A(x, y).G, A(x, y).B)$
 - $R(x, y).B := f_3(A(x, y).R, A(x, y).G, A(x, y).B)$
- Posibilidades:
 - Aplicar las mismas transformaciones que antes (suma, producto, ajuste de histograma, etc.), pero con distintos parámetros para cada canal.
 - Transformaciones basadas en modelos de color. Cambiar el modelo de color (RGB, HSV, HLS, XYZ, YUV, etc.) y aplicar la función en ese modelo.

- Transformación de **falso color**
- Es una transformación de la misma familia, cuyo objetivo es hacer más visibles las **pequeñas variaciones del nivel de gris**.
- Se define una paleta de salida adecuada y una transformación de cada valor de gris en la paleta.



Ejemplo. Transformación de falso color.

Las transformaciones de este tipo son comunes en imágenes médicas y de satélite.

- En estas aplicaciones, la profundidad del canal puede ser fácilmente mayor que 1 byte. Al usar sólo 256 grises se pierde información.



Imagen de entrada

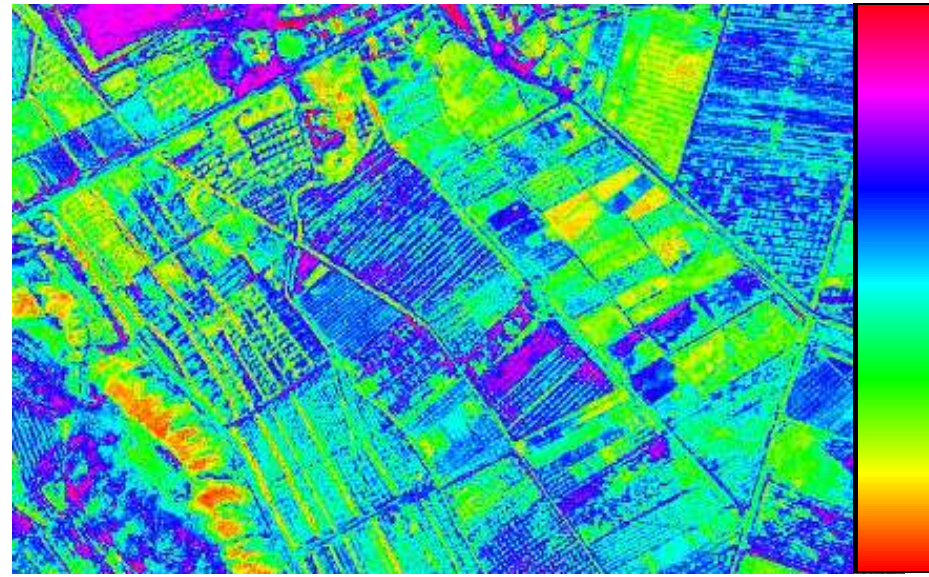


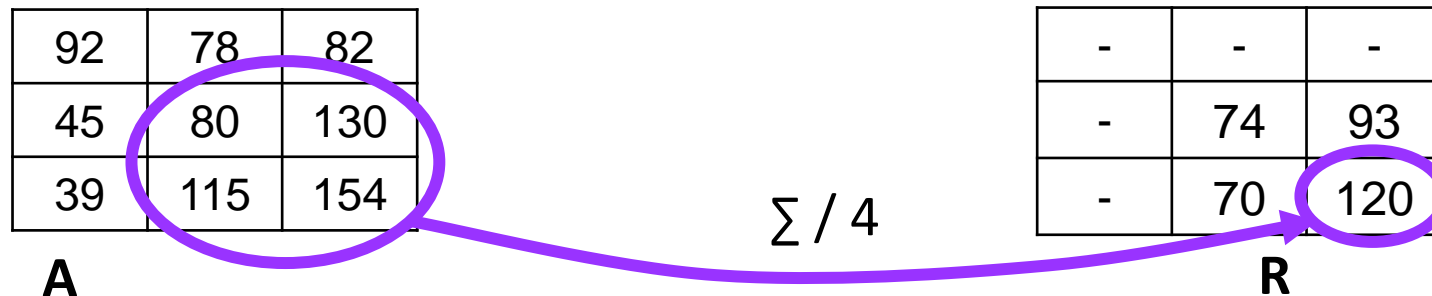
Imagen con falso color

Recordamos:


- Las transformaciones globales se pueden realizar igual en todos los canales o con valores distintos.
- En el primer caso, habrá un cambio en la intensidad. En el segundo, puede haber también un cambio de color.
- **Balance de blancos:** compensar los canales para obtener los colores más realistas posibles.
- **Procesamiento global:** el valor de un píxel de salida depende del píxel (o píxeles) correspondientes de la imagen de entrada.
- Operaciones aritméticas, lógicas, etc.
- Distintas aplicaciones: mejora del histograma, reducción de ruido, composición de imágenes, ajuste del color, etc.
- Normalmente no aparecen solas, sino combinadas con otros tipos de operaciones


- Un filtro también es llamado convolución por el tipo de operador matemático que se utiliza.
- Los filtros son operaciones locales: El valor de cada pixel depende de si mismo y también de los valores de los pixeles vecinos
- Ejemplo si hacemos una media de cuatro vecinos


$$R(x,y) := (A(x-1,y-1) + A(x,y-1) + A(x-1,y) + A(x,y)) / 4$$






- Conceptos:
 - **Convolución:** operación matemática que suma una función f consigo misma repetidas veces en todo el dominio de otra función k , utilizando en cada suma como valor de escala el valor de k en ese punto de su dominio
 - **Kernel:** Es la matriz que se usa para transformar la imagen original $I(x,y)$ en otra $J(x,y) = I(x, y)*k$

Identidad	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
-----------	---	---

Nitidez	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
---------	---	---

Desenfoque gaussiano	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	
----------------------	--	--

Los filtros se aplican mediante un *kernel*, que es la matriz por la que multiplicamos cada uno de los píxeles de la imagen original, para obtener la imagen resultado.

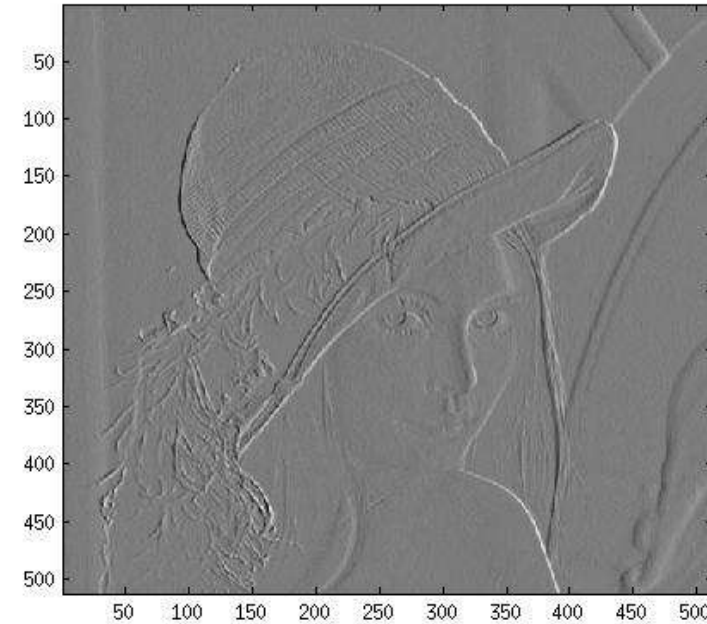
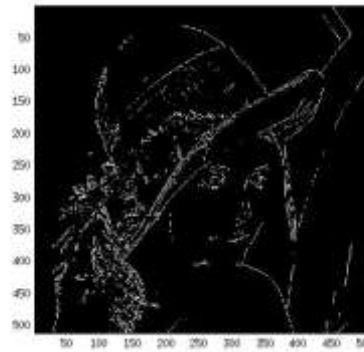
Detección de Bordes	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	

Los filtros de detección de bordes son diferentes formas calcular gradientes en la imagen.
Equivalente a derivar la imagen original

Filtro de Sobel en la dirección x



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



- Laplaciana

Representa la segunda derivada de la imagen



Original

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



Laplaciana

$$\Delta(I(x, y)) = \nabla(\nabla(I(x, y))) = \frac{\partial^2 I}{\partial x^2} \bar{u}_x + \frac{\partial^2 I}{\partial y^2} \bar{u}_y$$

- Transformada Discreta de Fourier Directa

$$\mathcal{F}(f, n) = \frac{1}{N} \sum_{k=0}^{N-1} f(k) \cdot e^{-i \frac{2\pi}{N} kn} \quad \forall n = 0, 1, \dots, N-1$$

N: número de datos

k: k-ésimo dato

$$e^{ix} = \cos x + i \sin x$$

$e^{\frac{-i2\pi}{N}}$ = W Factor de rotación

- Transformada Discreta de Fourier Inversa

$$f(x) = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{F}(f, n) \cdot e^{i \frac{2\pi}{N} kn} \quad \forall n = 0, 1, \dots, N-1$$

- Simplificando para dos dimensiones

Directa

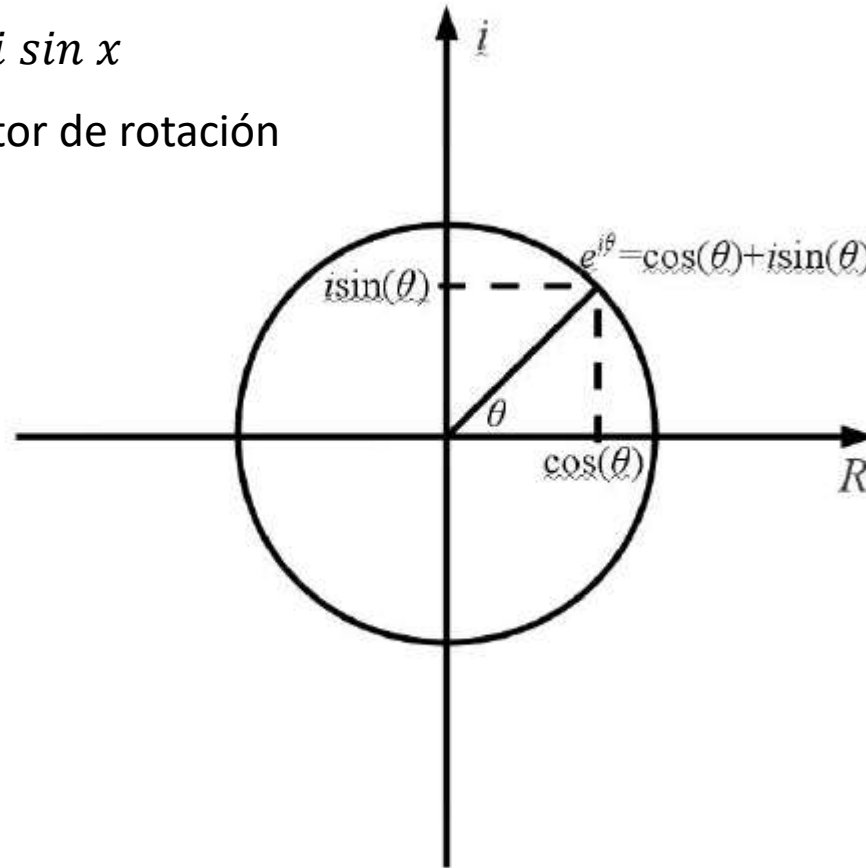
$$\mathcal{F}(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-i2\pi(\frac{xu}{M} + \frac{yv}{N})}$$

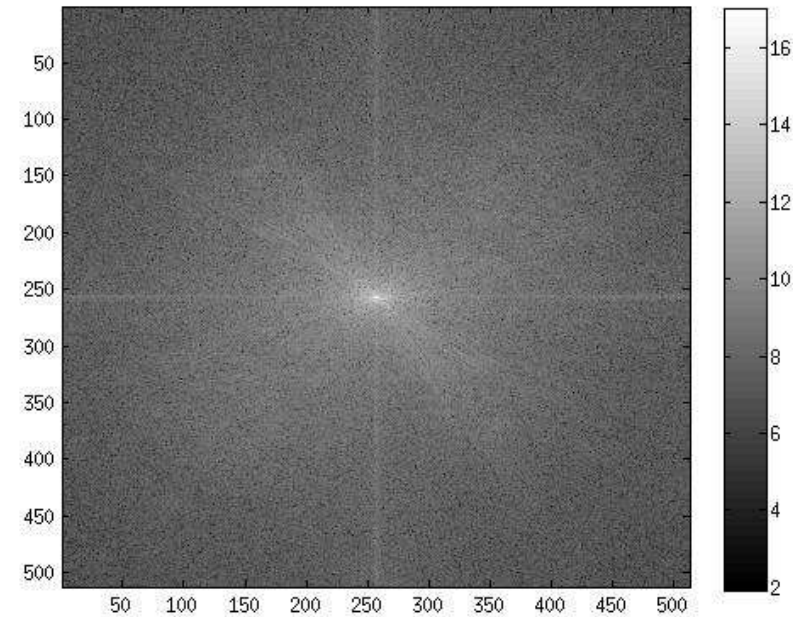
Inversa

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) e^{i2\pi(\frac{xu}{M} + \frac{yv}{N})}$$

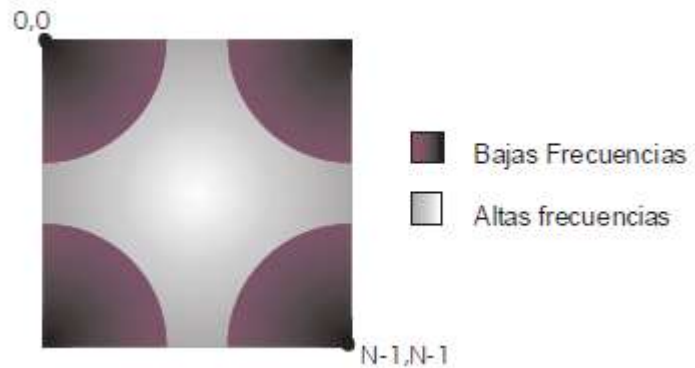
$$e^{ix} = \cos x + i \sin x$$

$$e^{\frac{-i2\pi}{N}} = W \text{ Factor de rotación}$$





Ambas imágenes contienen la misma información

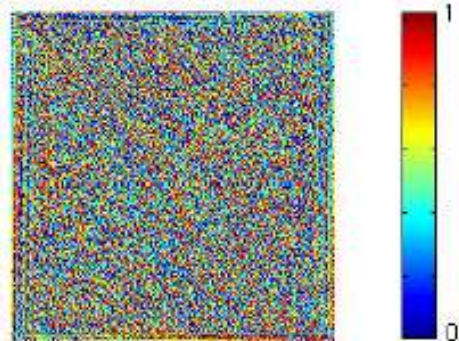


Modificación en el espacio de Fourier

- (a) Poniendo a cero el módulo correspondiente a las altas frecuencias
- (b) Manteniendo las fases
- (c) El resultado es una imagen suavizada



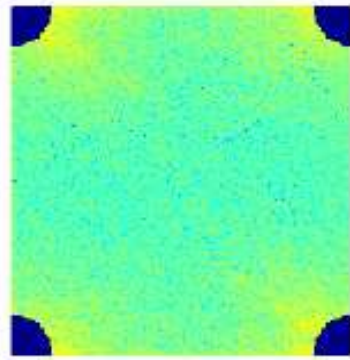
(a)



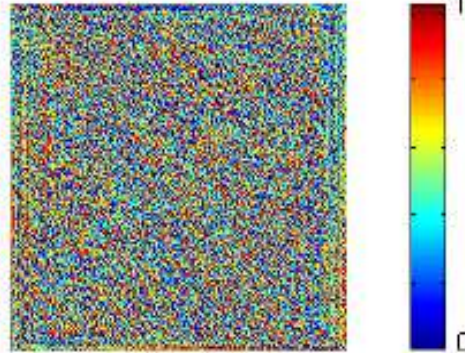
(b)



(c)



(a)



(b)



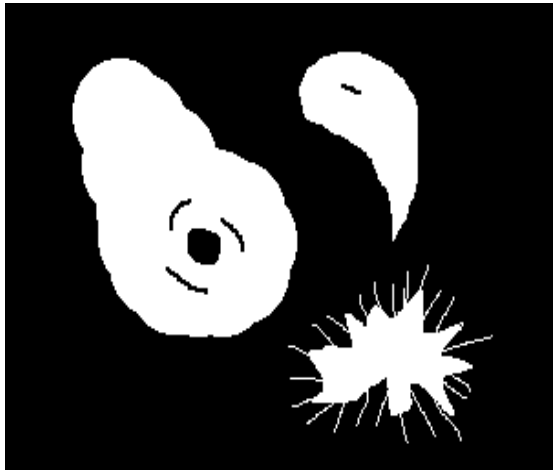
(c)

Modificación en el espacio de Fourier

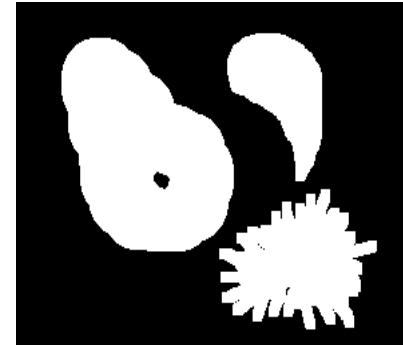
- (a) Poniendo a cero el módulo correspondiente a las bajas frecuencias
- (b) Manteniendo las fases
- (c) El resultado es una imagen de contornos y ruido de alta frecuencia

- **Operaciones Morfológicas**
 - **Dilatación:** $A \oplus B$. Esta operación produce el efecto de dilatar el aspecto del elemento A , usando para ello B
 - **Erosión:** $A \ominus B$. La erosión “adelgaza” la imagen A con el patrón B .
 - **Apertura:** $A \odot B = (A \ominus B) \oplus B$. Intuitivamente equivale a re-estructurar A con B .

Imagen de entrada



Dilatación 1



Dilatación 3

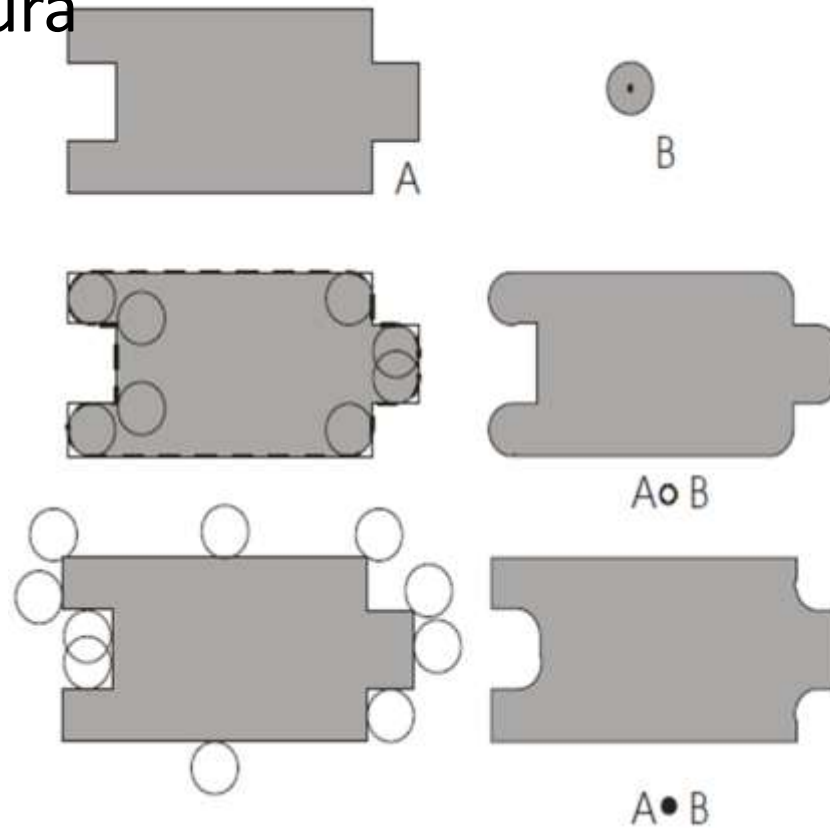


Erosión 1



Erosión 3

Apertura



FIN TEMA 1