



Instituto Tecnológico de La Paz

Maestría en Sistemas Computacionales

Introducción a la Visión por Computadora

Dr. Alejandro Israel Barranco Gutiérrez

13 de enero de 2014



Definiciones

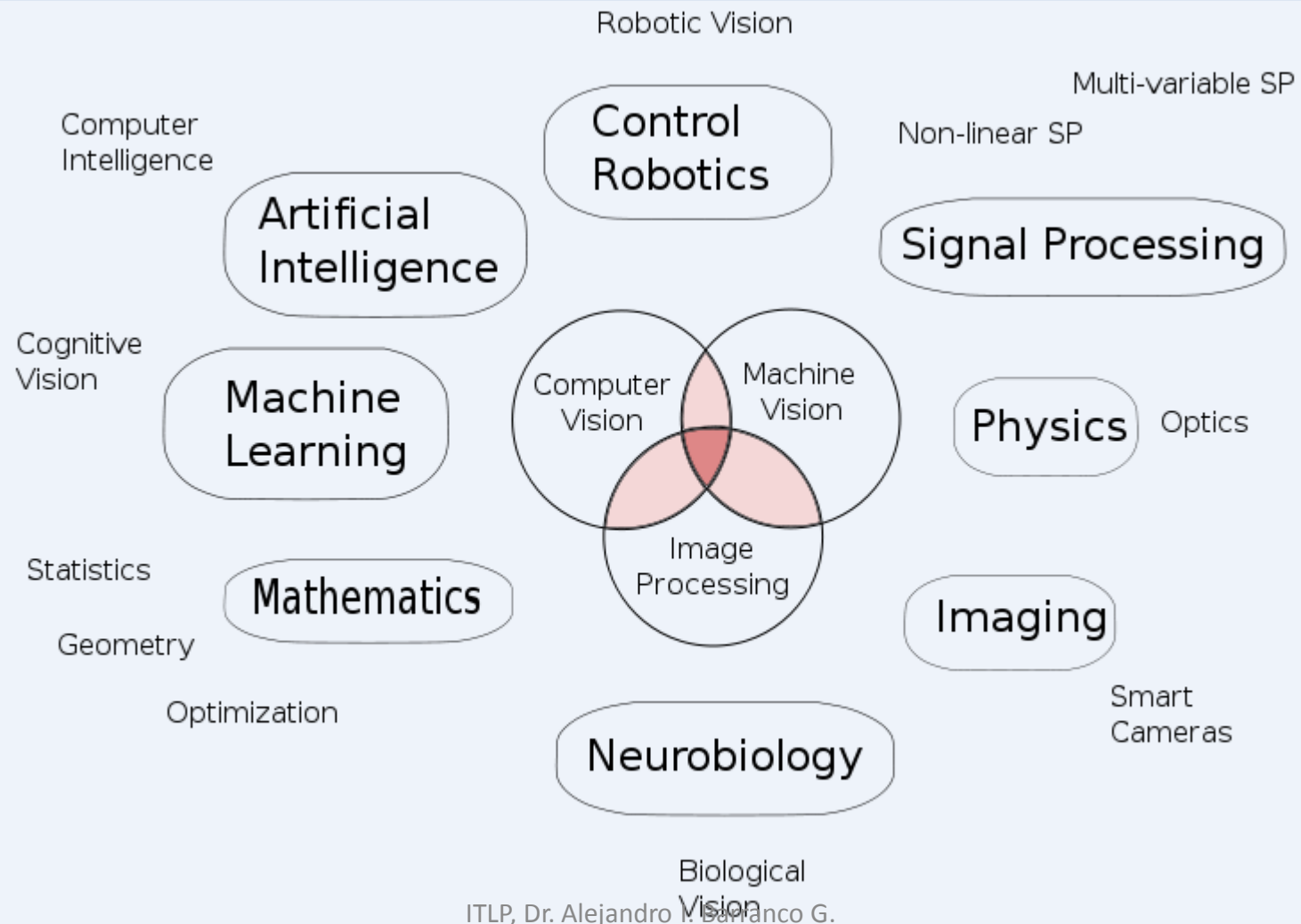


- Computer Vision = Geometry + Measurement + Interpretation
(S. Nagabhushana).

- Ciencia que estudia la interpretación de imágenes mediante computadores digitales.
(José Andrés Somolinos Sánchez).

- Consiste en la adquisición, procesamiento, clasificación y reconocimiento de imágenes digitales.
(José Jaime Esquema Elizondo).

Algunas relaciones de la visión por computadora con otras disciplinas



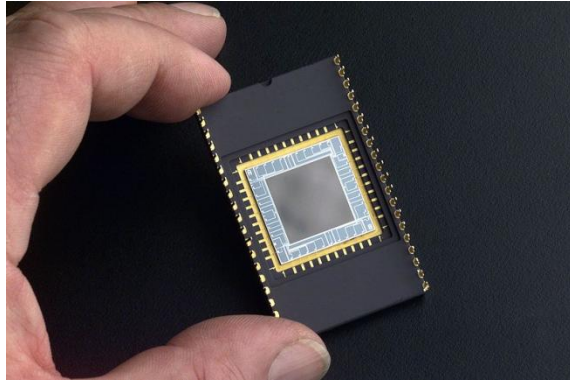
Subcampos de la visión por computadora

- Adquisición de imágenes
- Restauración de imágenes
- Fotogrametría
- Procesamiento de imágenes
- Reconocimiento de objetos



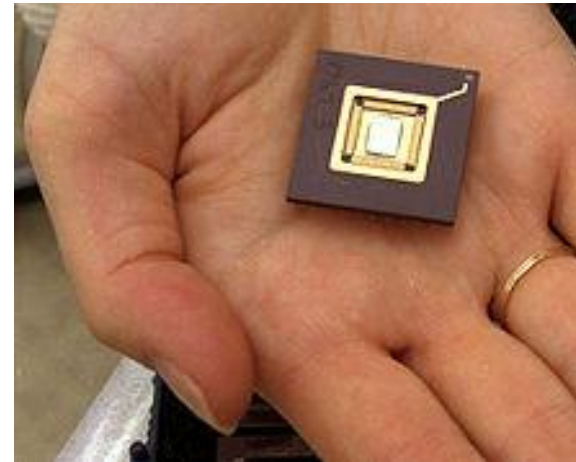
Elementos de Percepción visual (Cámaras)

CCD, APS y MEMS



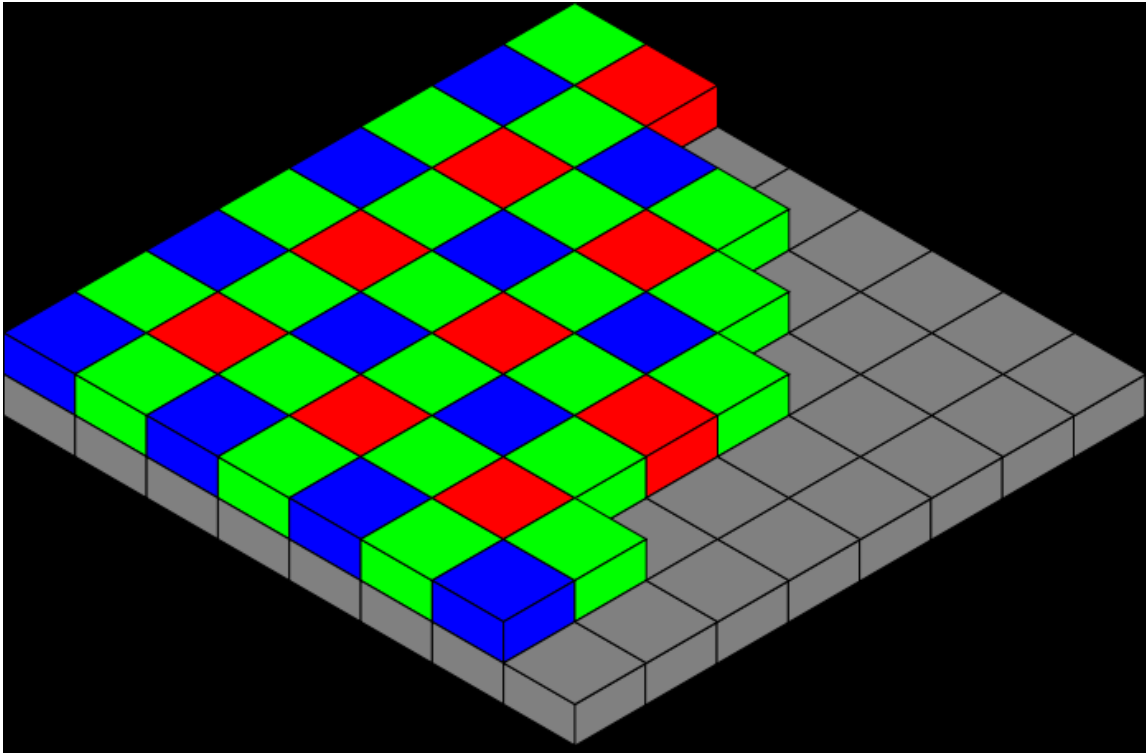
Charge-Coupled Device es un circuito integrado que contiene un número determinado de condensadores acoplados. Bajo el control de un circuito interno.

Active Pixel Sensor. Sensor que detecta la luz basado en tecnología CMOS y por ello más conocido como Sensor CMOS.



Microelectromechanical Systems. Es la tecnología electromecánica, micrométrica y sus productos.

Filtro de Bayer

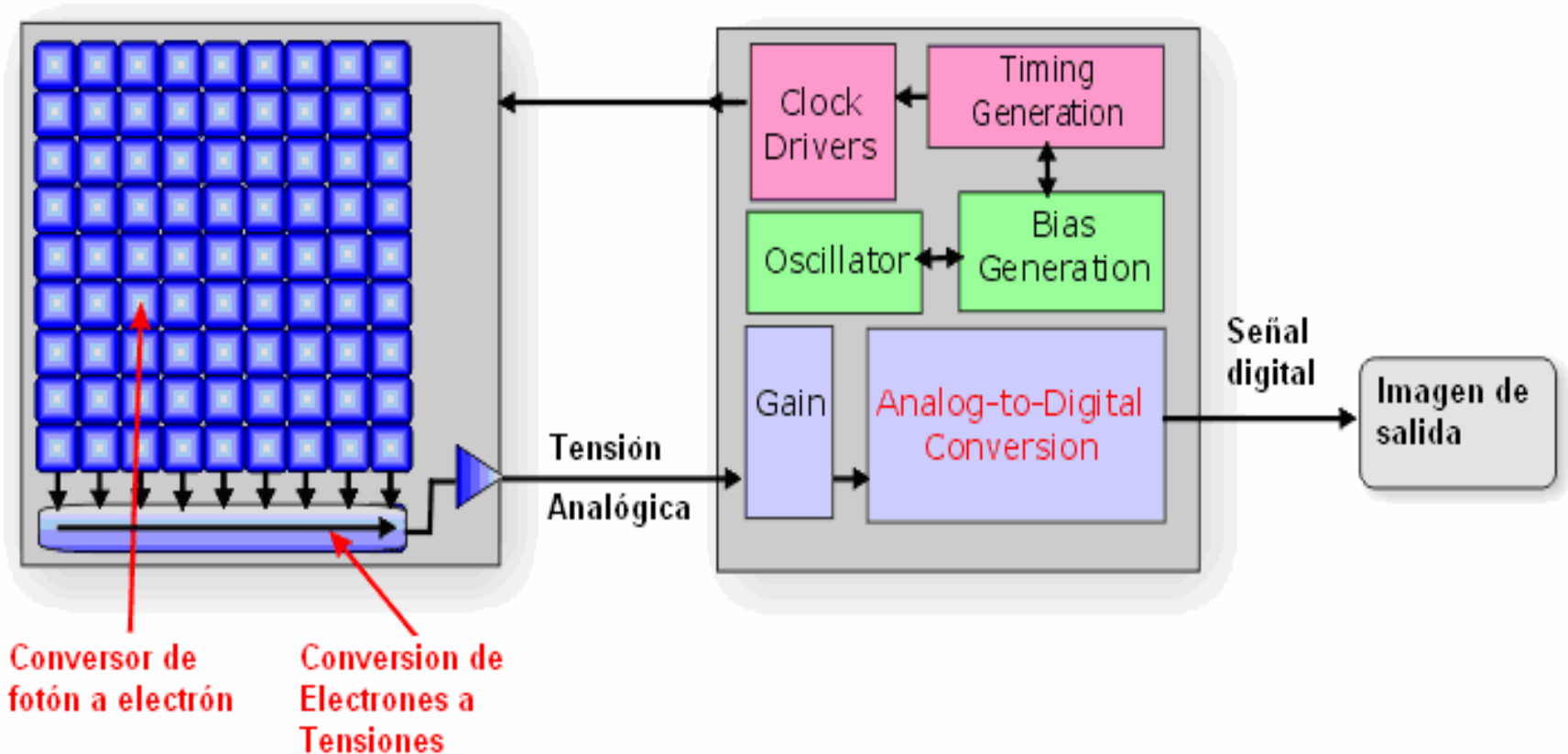


El filtro, máscara o mosaico de Bayer es un tipo de matriz de filtros, rojos verdes y azules, que se sitúa sobre un sensor digital de imagen (CCD o APS) para hacer llegar a cada fotodiodo la información de luminosidad correspondiente a una sección de los distintos colores primarios. Se llama así por su creador, Bryce Bayer, de la empresa Eastman Kodak.

Arquitectura CCD

Matriz de sensores de Luz

Circuito integrado de la cámara



Comparación APS vs CCD

Ventajas APS

- Consumo eléctrico muy inferior
- Económico (necesita pocos componentes externos)
- Lectura simultánea de mayor número de píxeles
- El conversor digital puede estar integrado en el mismo chip
- Escaso Blooming ("Smear") o inexistente
- Mayor flexibilidad en la lectura (Previsualización más rápida, vídeo,...)
- Los píxeles pueden ser expuestos y leídos simultáneamente
- Otras topologías posibles (el sensor SuperCCD de Fujifilm emplea una construcción en forma de panel (octogonal) para los píxeles)
- Distintos tipos de píxeles (según tamaño y sensibilidad) combinables
- Muy alta frecuencia de imagen en comparación a un CCD del mismo tamaño

Desventajas

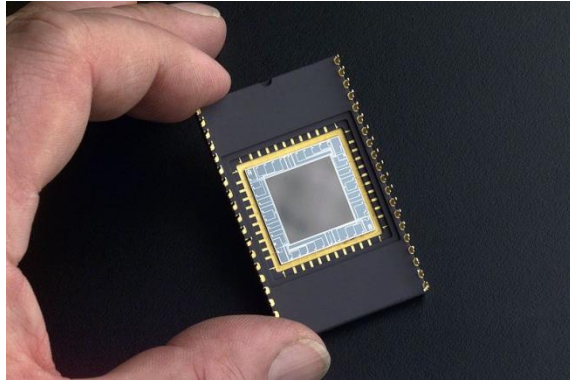
- Menor superficie receptora de la luz por píxel
- Menor uniformidad de los píxeles (mayor ruido de patrón fijo-FPN)
- Efecto "jelly" o inestabilidad en la imagen con movimientos rápidos o flashes debido (se tuerce el video) al tipo de obturación giratoria que utiliza.

Cámaras omnidireccionales



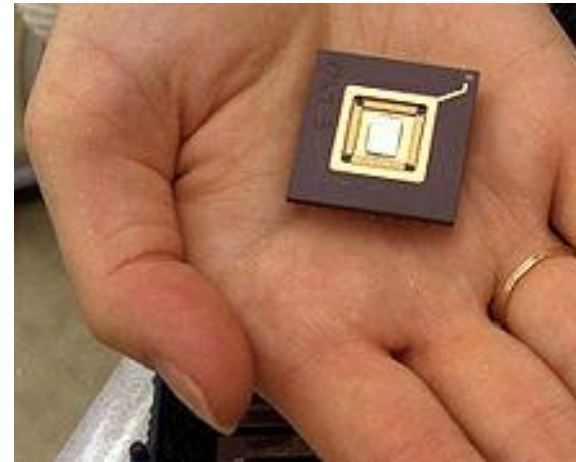
Elementos de Percepción visual (Cámaras)

CCD, APS y MEMS



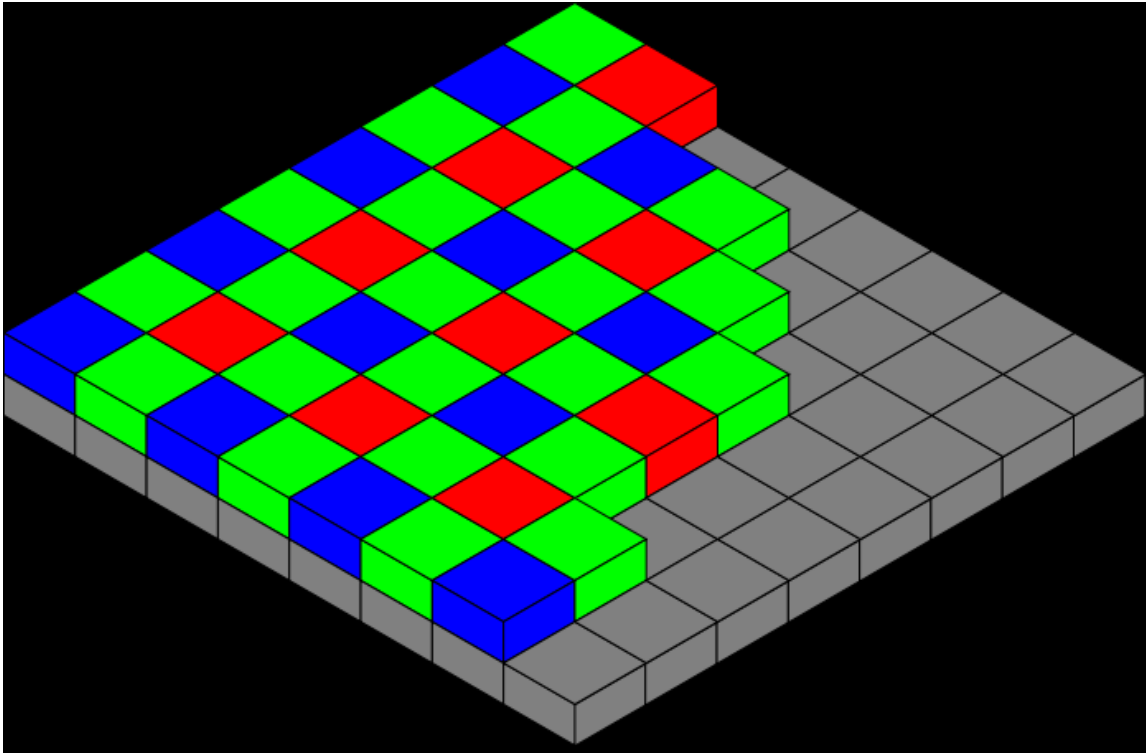
Charge-Coupled Device es un circuito integrado que contiene un número determinado de condensadores acoplados. Bajo el control de un circuito interno.

Active Pixel Sensor. Sensor que detecta la luz basado en tecnología CMOS y por ello más conocido como Sensor CMOS.



Microelectromechanical Systems. Es la tecnología electromecánica, micrométrica y sus productos.

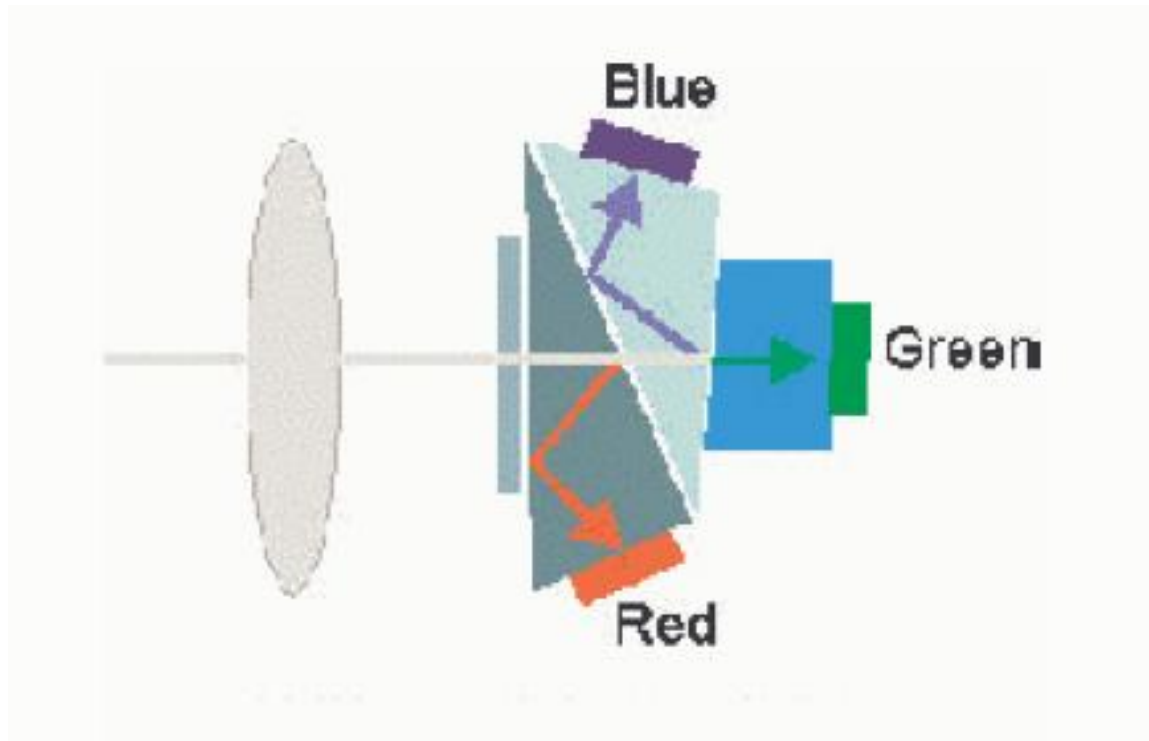
Filtro de Bayer



El filtro, máscara o mosaico de Bayer es un tipo de matriz de filtros, rojos verdes y azules, que se sitúa sobre un sensor digital de imagen (CCD o APS) para hacer llegar a cada fotodiodo la información de luminosidad correspondiente a una sección de los distintos colores primarios. Se llama así por su creador, Bryce Bayer, de la empresa Eastman Kodak.

Cámara de color con prisma

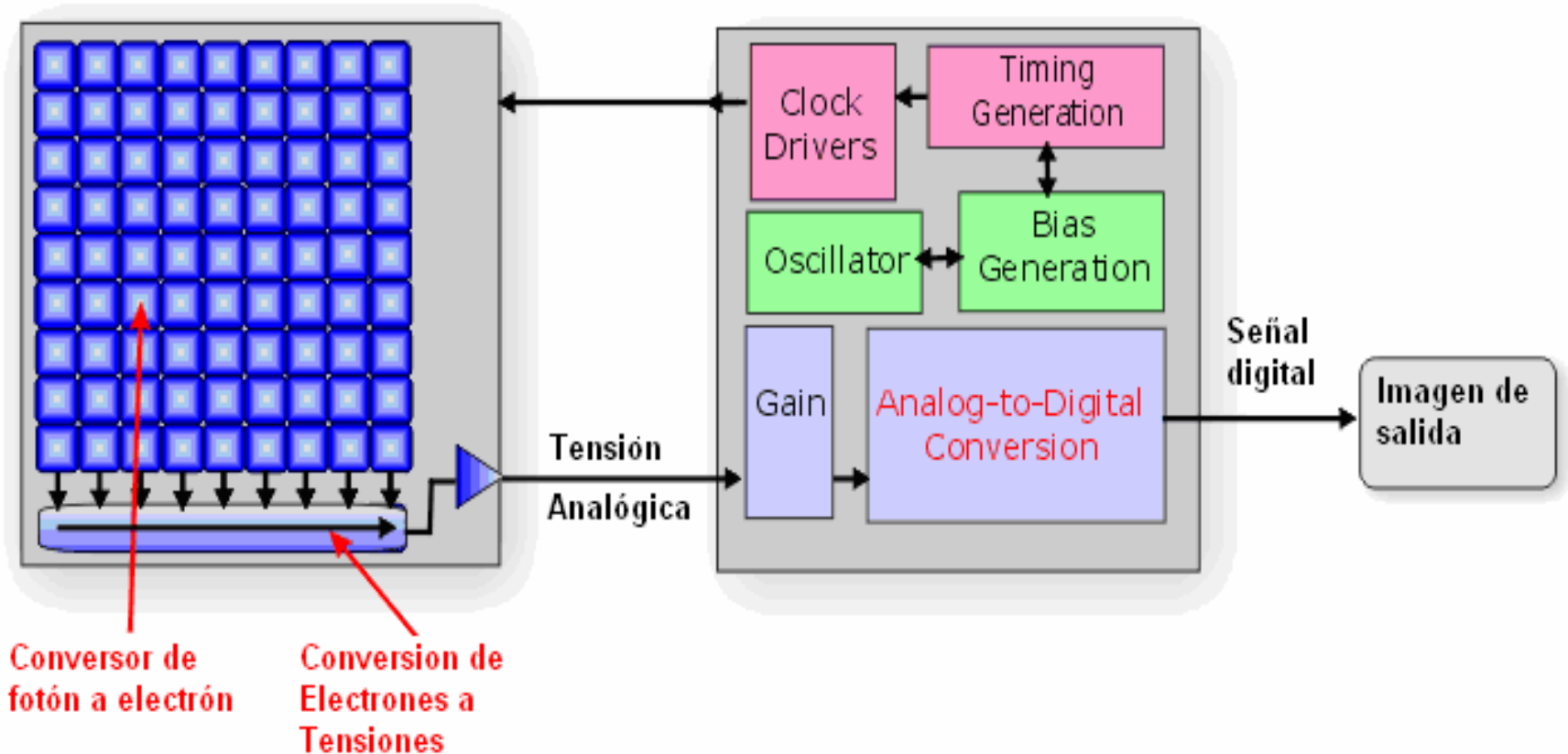
Con un prisma separa la luz en 3 rayos de colores distintos, requiere alineación precisa.



Arquitectura CCD

Matriz de sensores de Luz

Circuito integrado de la cámara



Comparación APS vs CCD

Ventajas APS

- Consumo eléctrico muy inferior al CCD
- Económico (necesita pocos componentes externos)
- Lectura simultánea de mayor número de píxeles
- El conversor digital puede estar integrado en el mismo chip
- Escaso Blooming ("Smear") o inexistente
- Mayor flexibilidad en la lectura (Previsualización más rápida, vídeo,...)
- Los píxeles pueden ser expuestos y leídos simultáneamente
- Otras topologías posibles (el sensor SuperCCD de Fujifilm emplea una construcción en forma de panel (octogonal) para los píxeles)
- Distintos tipos de píxeles (según tamaño y sensibilidad) combinables
- Muy alta frecuencia de imagen en comparación a un CCD del mismo tamaño

Desventajas

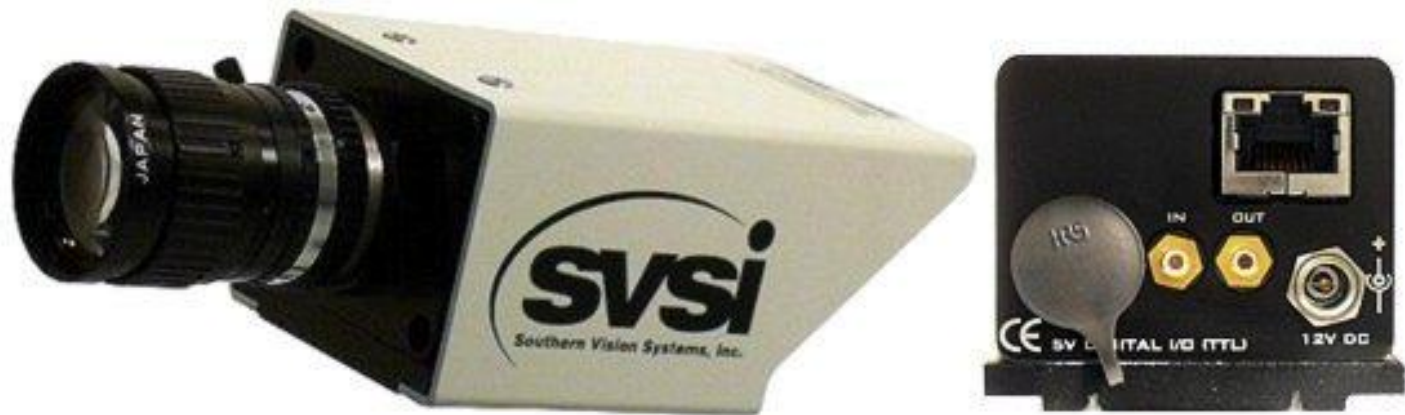
- Menor superficie receptora de la luz por píxel
- Menor uniformidad de los píxeles (mayor ruido de patrón fijo-FPN)
- Efecto "jelly" o inestabilidad en la imagen con movimientos rápidos o flashes debido (se tuerce el video) al tipo de obturación giratoria que utiliza.

Cámaras omnidireccionales

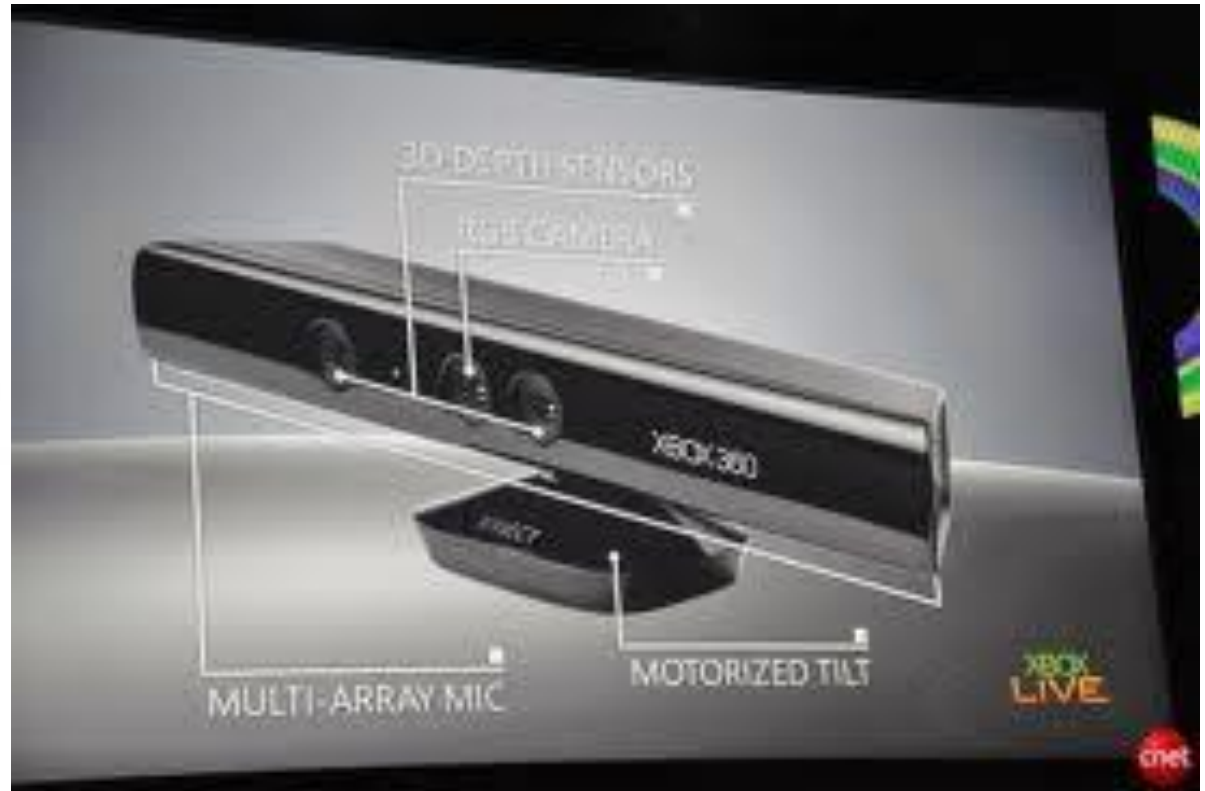


Cámaras Rápidas

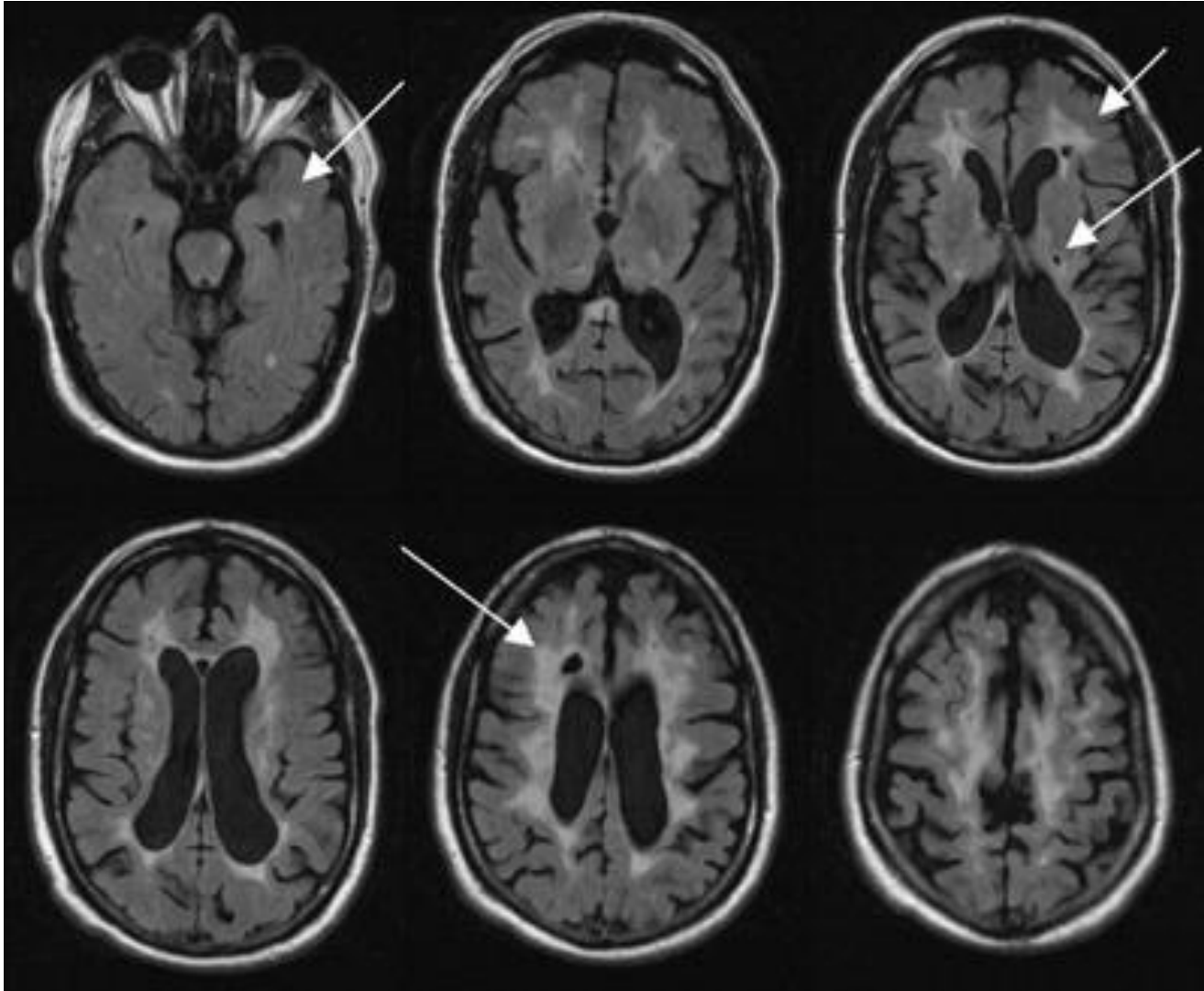
969 fps@640x64pixeles



Cámaras con lámpara



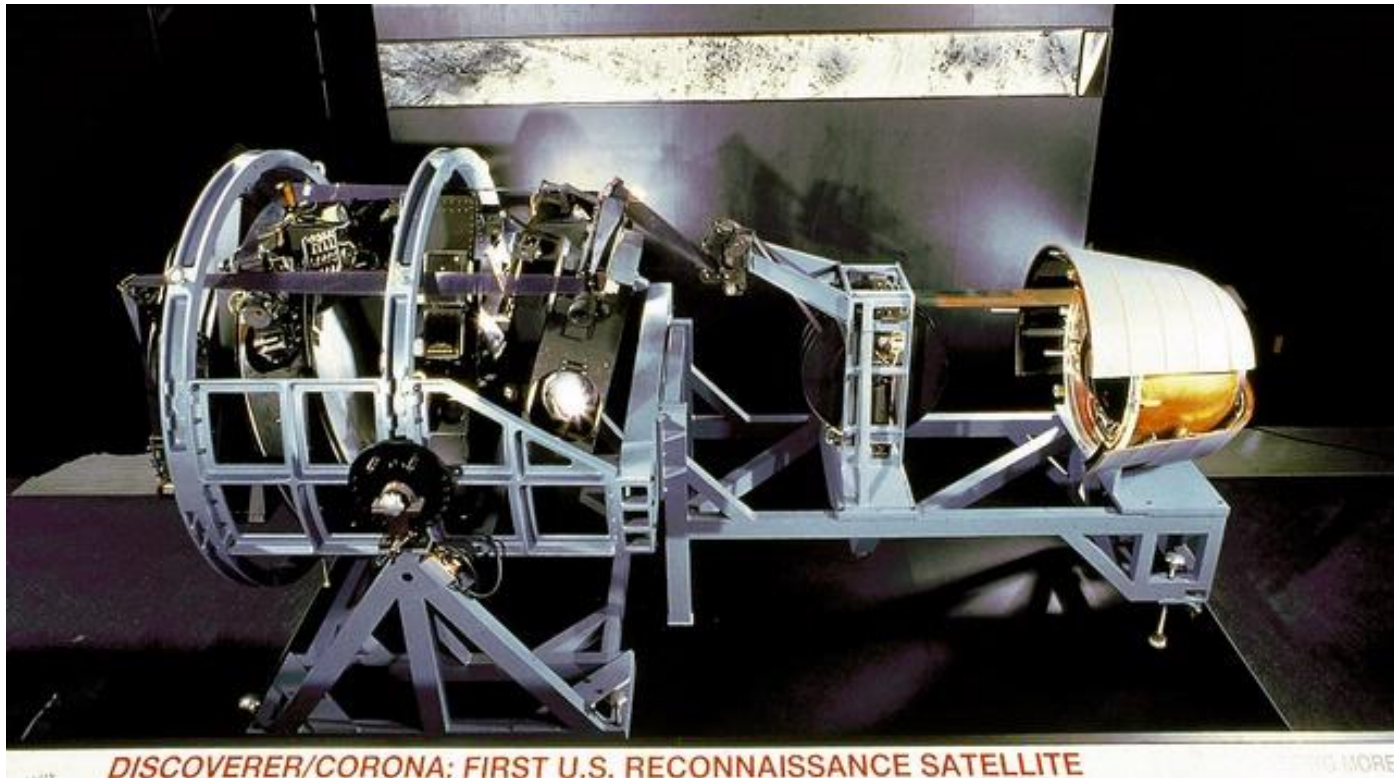
Neuroimagen



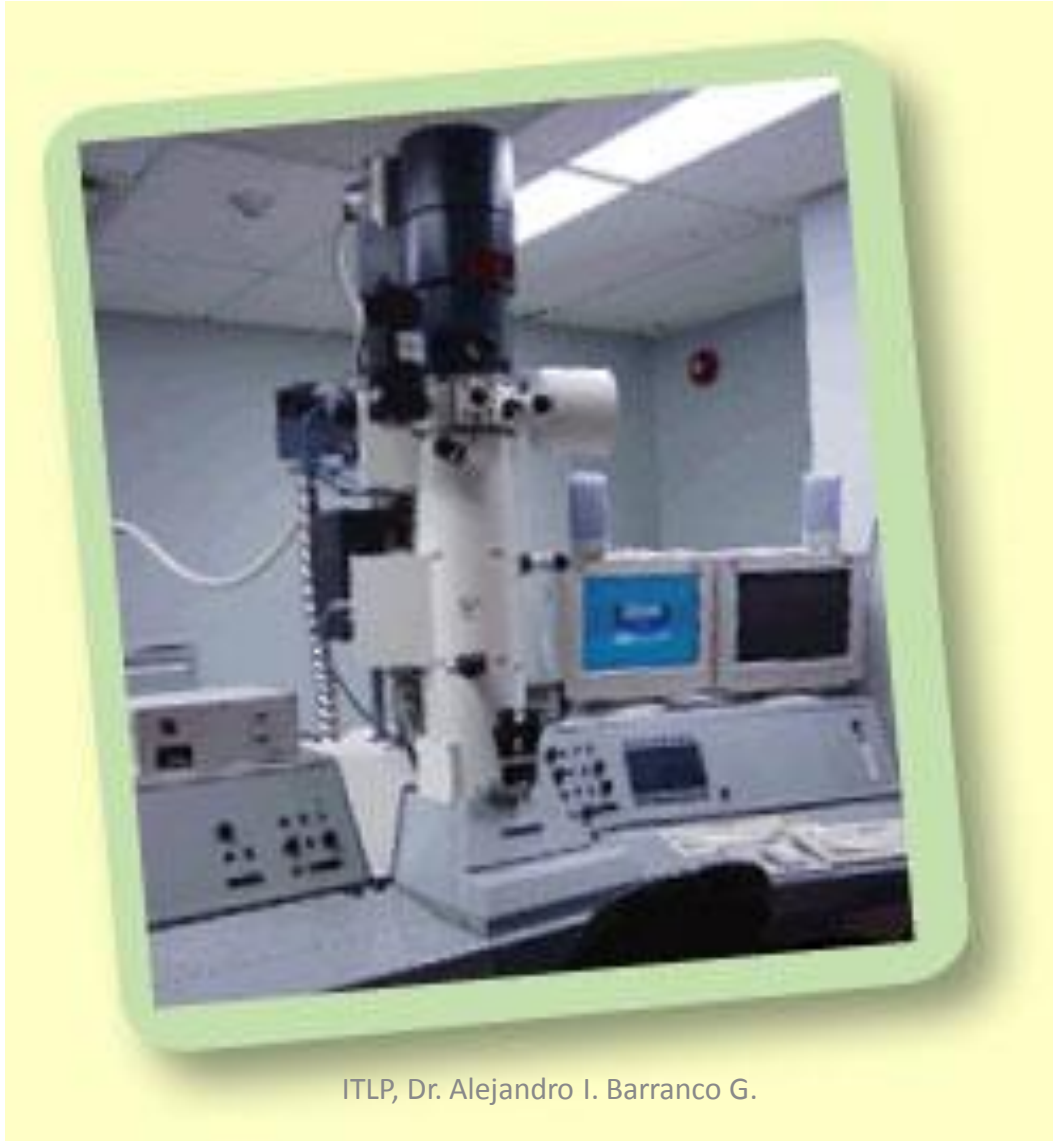
Cámaras para trabajo en el agua



Cámara Satelital



Microscopio Electrónico

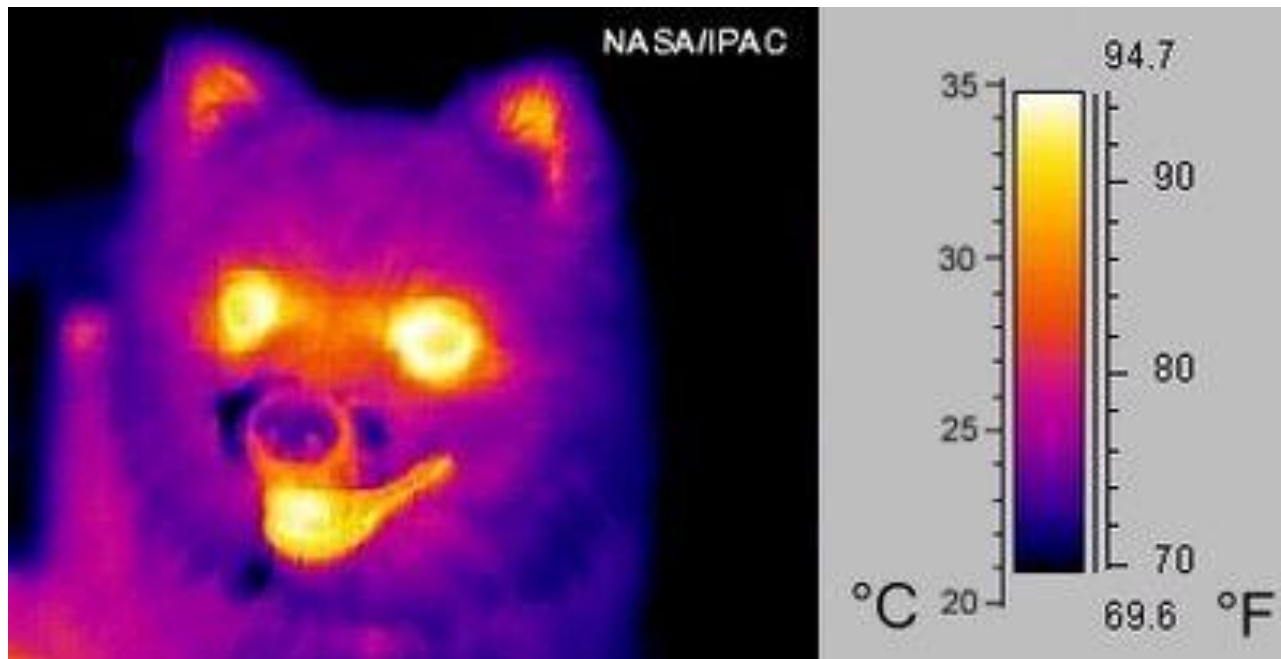


ITLP, Dr. Alejandro I. Barranco G.

Cámara Telescopio



Cámaras Térmicas



Ultrasonido



Rayos X



Representación de imágenes

Imagen representada en una matriz

Las cámaras digitales a color nos entregan 3 matrices

La combinación de las tres matrices (RGB) nos da el efecto de una matriz a color

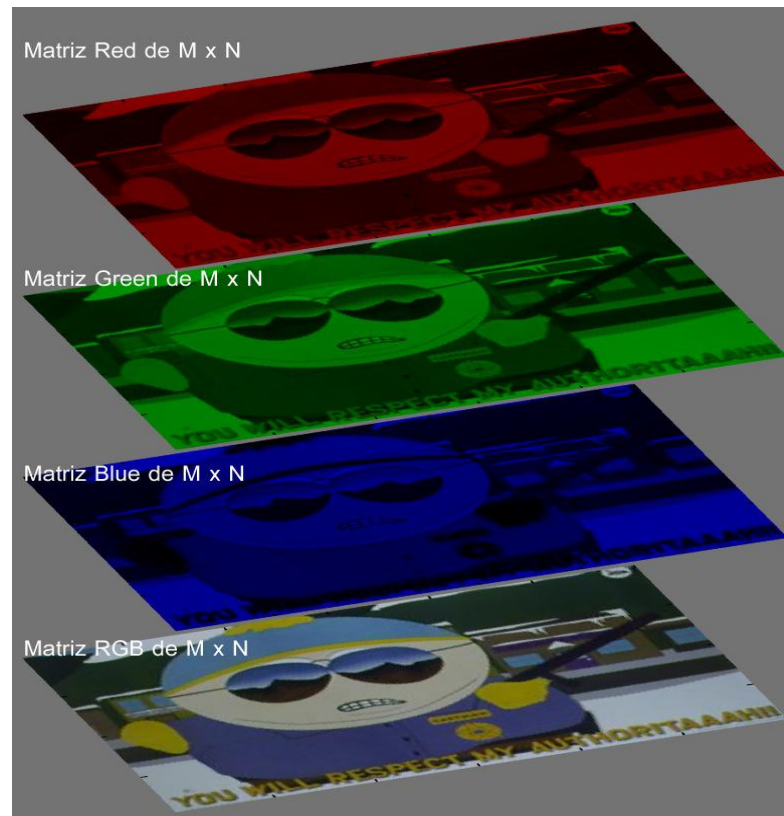


Imagen a color

Definimos a una imagen de color en RGB como:

$$I(x, y, z) \in \{0 \leq \mathbb{Z} \leq 255\}$$

$$\forall \{x|0 \leq x \leq k\}, \{y|0 \leq y \leq l\}, \{z|0 \leq z \leq 2\}$$

Donde: x, y, z son indices en las 3 dimensiones de la imagen y \mathbb{Z} es el conjunto de los números enteros.

Imagen en escala de grises

- A partir de una imagen a color RGB, podemos obtener una versión de la misma en escala de grises:

$$I_g(x, y) = \frac{(I_w(x, y, 1) + I_w(x, y, 2) + I_w(x, y, 3))}{3}$$

Manejo de imágenes en MATLAB

```
I=imread('PNPC2012.jpg')  
subplot(4,1,1)  
imshow(I)
```

```
subplot(4,1,2)  
imshow(I(:,:,1))
```

```
subplot(4,1,3)  
imshow(I(:,:,2))
```

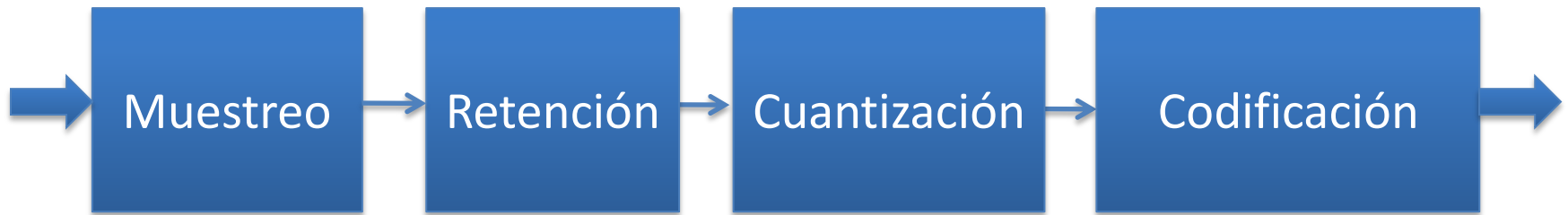
```
subplot(4,1,4)  
imshow(I(:,:,3))
```

```
G(:,:,:)=(I(:,:,1)+I(:,:,2)+I(:,:,3))/3;  
figure(2)  
imshow(G)
```

```
J = rgb2gray(I);  
figure(3)  
imshow(J)
```


Digitalización de señales

Diagrama a bloques básico de la digitalización

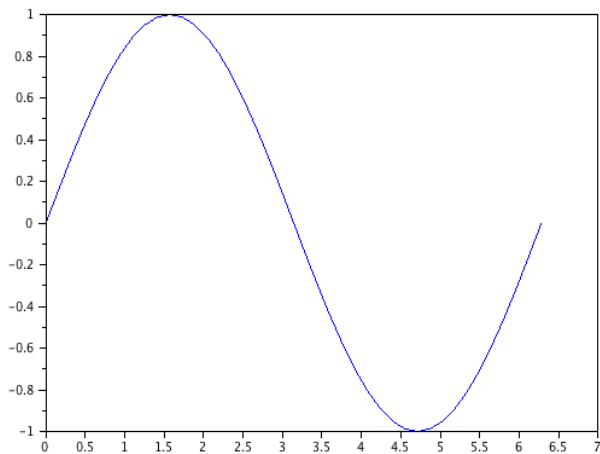


Este esquema recibe como entrada una señal analógica y entrega la versión digitalizada de esa señal de entrada.

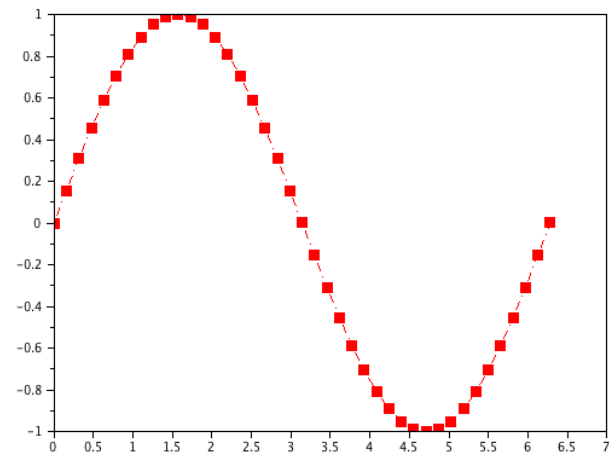
Muestreo

Muestreo: El muestreo (en inglés, *sampling*) consiste en tomar muestras periódicas de la amplitud de onda. La velocidad con que se toman esta muestra, es decir, el número de muestras por segundo, es lo que se conoce como frecuencia de muestreo.

Variable analógica



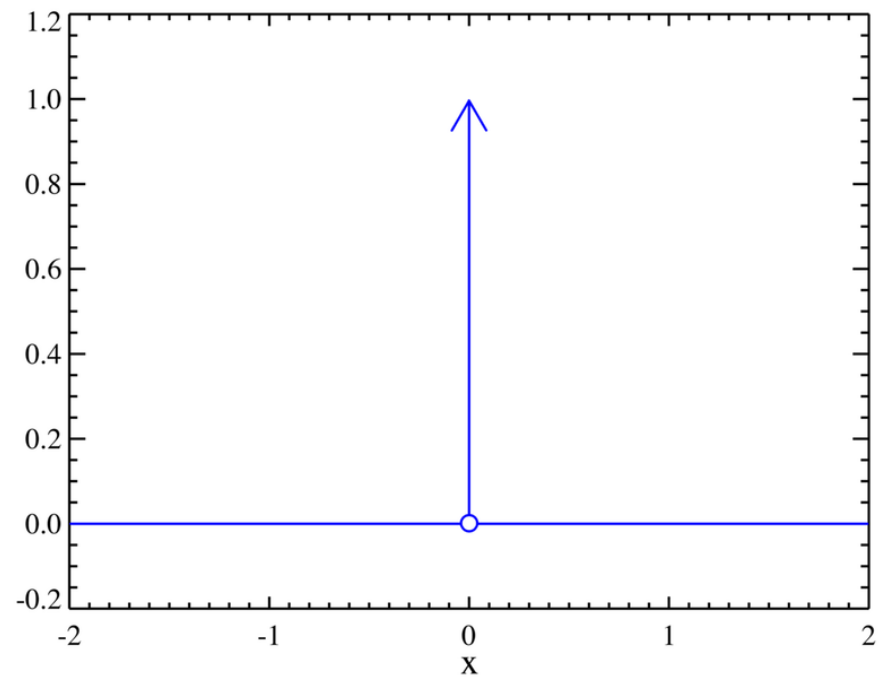
Variable discreta



$$y(k) = \sum_{n=1}^N \delta(t - nT) y(nT)$$

Función impulso

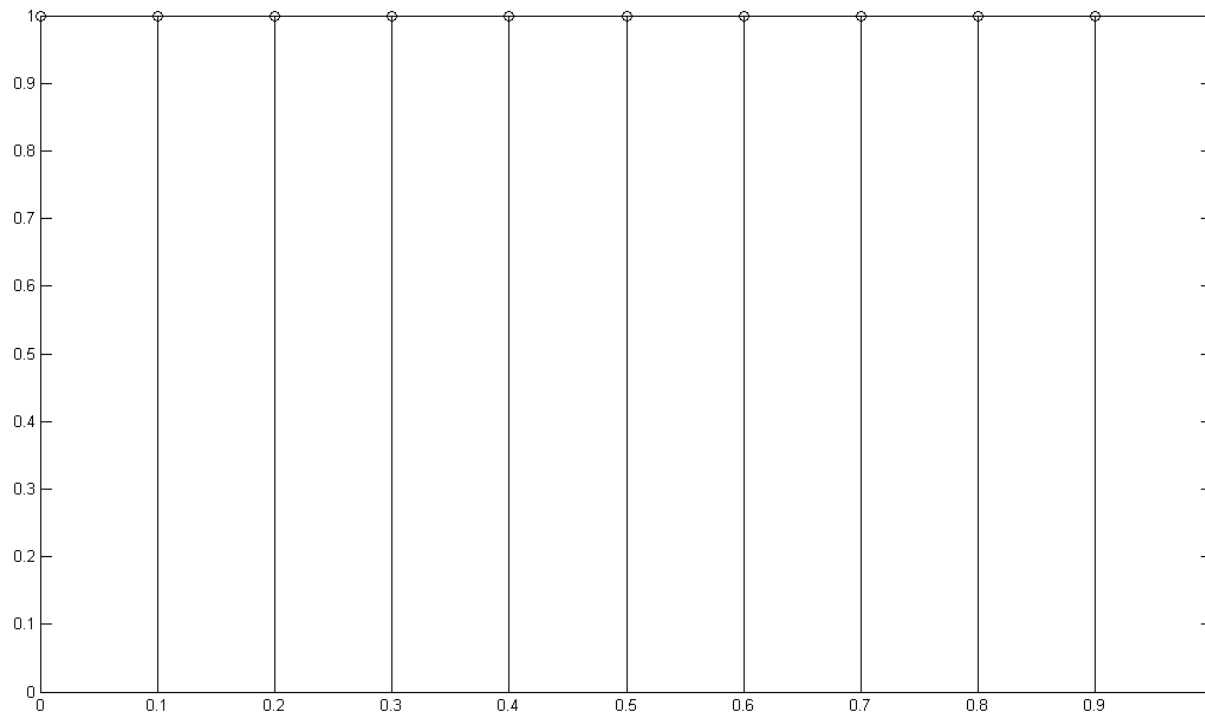
$$\delta(t - a) = \begin{cases} 1 & \text{para } t = a \\ 0 & \text{para } t \neq a \end{cases}$$



Muestreo

Se considera un tiempo de muestreo constante por ejemplo de T_s segundos entonces la frecuencia de muestreo es:

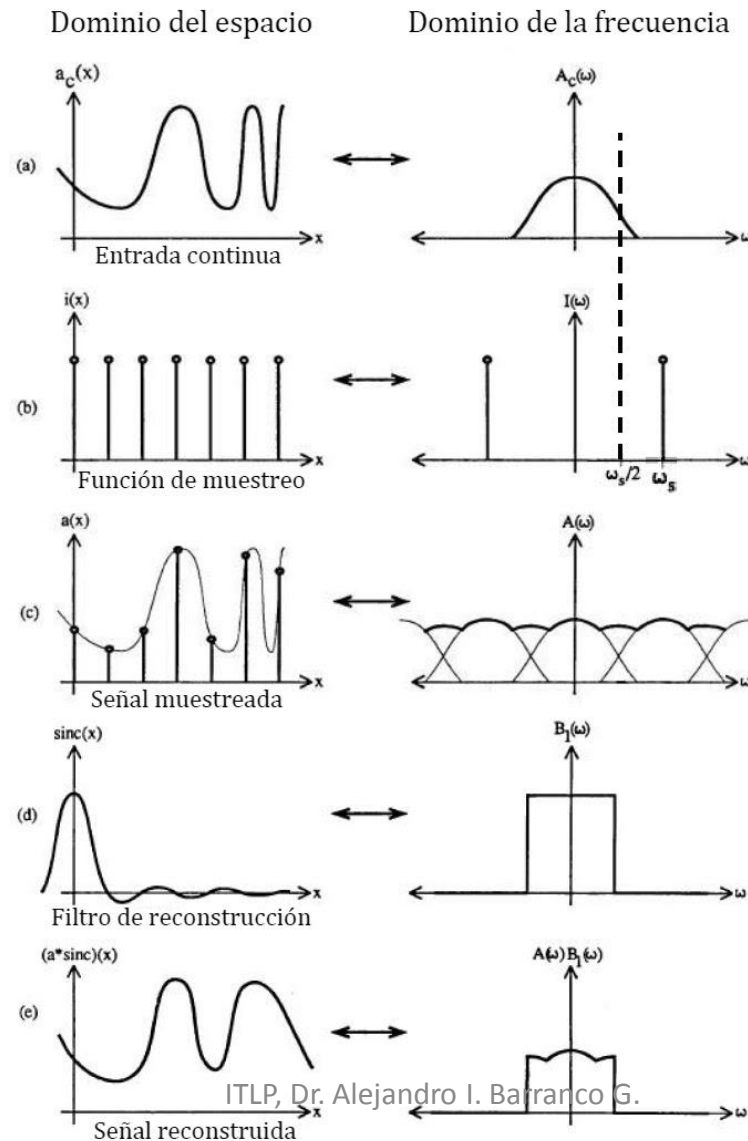
$$f_s = \frac{1}{T_s} \text{ (Muestras por segundo)}$$



```
Fs=10;  
t=0:1/Fs:1;  
x=ones(1,length(t));  
stem(t,x)
```

$$\sum_{n=0}^{10} \delta(t - nT_s)$$

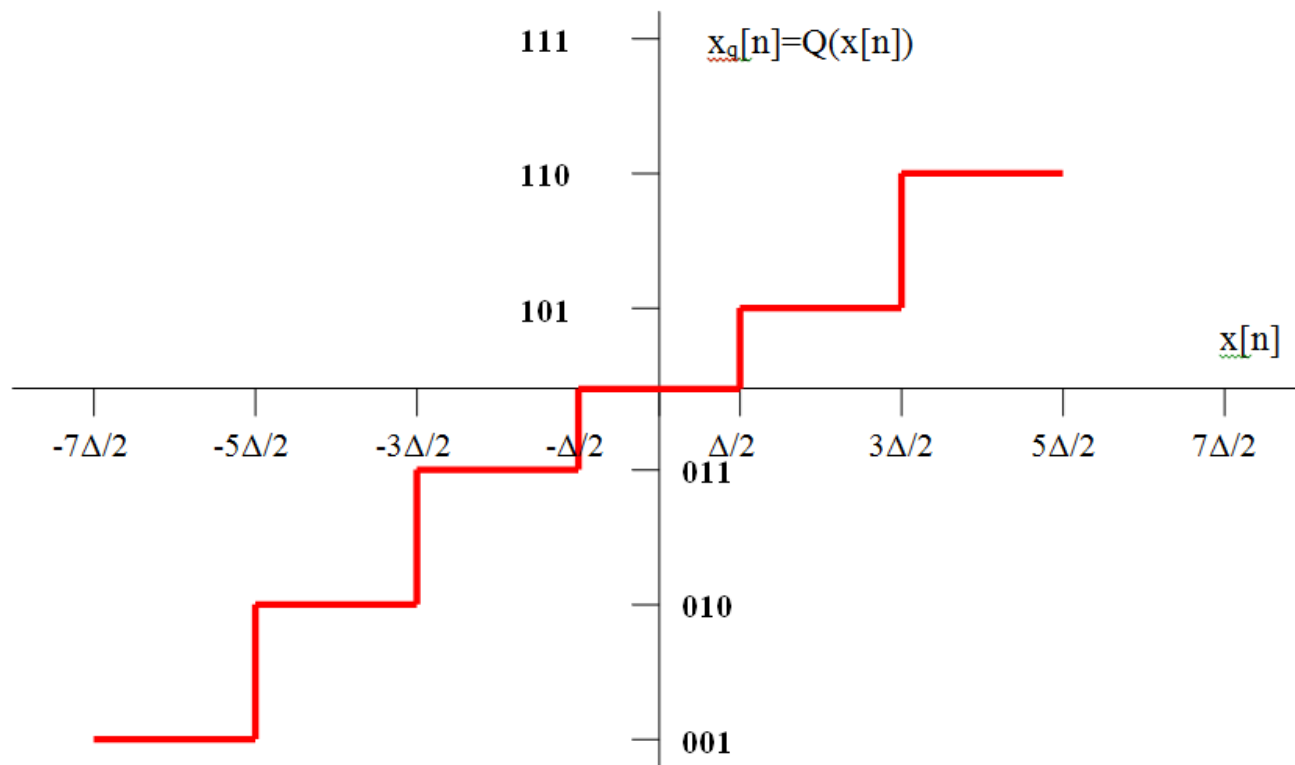
Muestreo en el dominio del tiempo y de la frecuencia



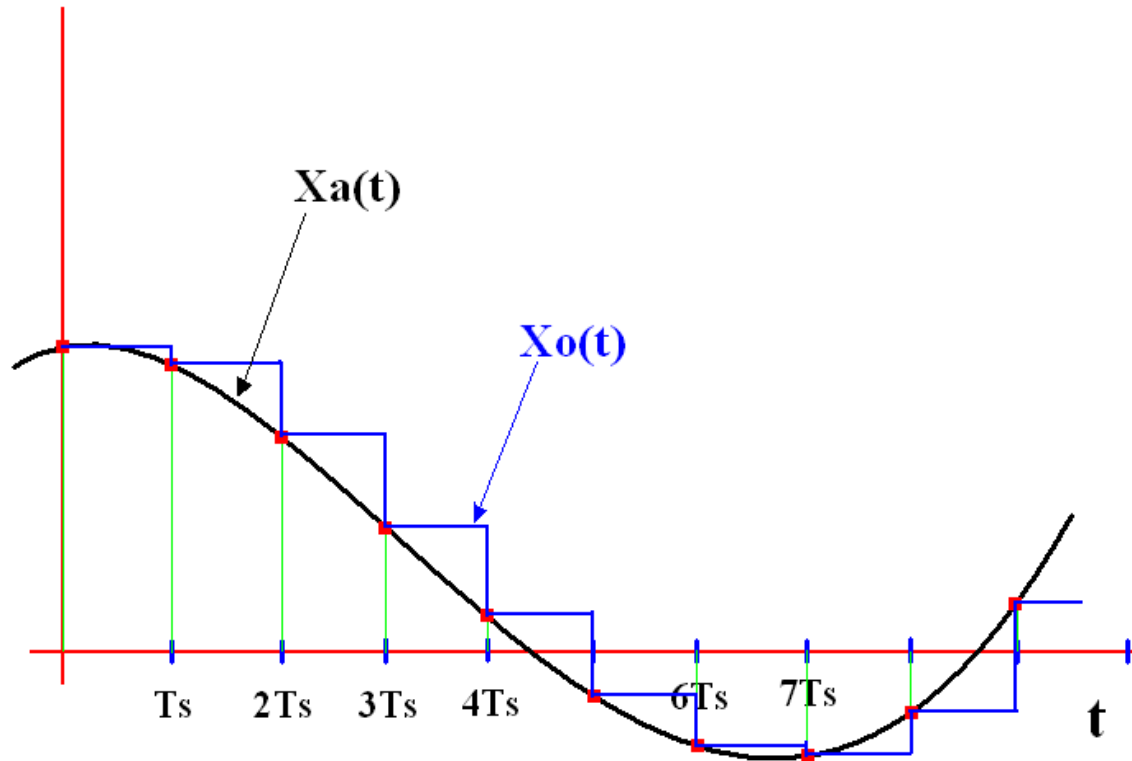
Retención

- *Retención:* (En ingles, *Hold*): Las muestras tomadas han de ser retenidas (retención) por un [circuito](#) de retención (Hold), el tiempo suficiente para permitir evaluar su nivel (cuantificación).
- z^0 , de orden cero.
- z^{-1} , de orden uno.
- z^{-2} , de orden dos.

Ejemplo de un cuantificador de 3 bits



Cuantificación (Cuantización)



Cuantificación: En el proceso de cuantificación, a un rango de valores de cada una de las muestras se le asigna un único valor de salida.

Codificación

Tensión	Código
3 volts	0111
2.5 volts	0101
2 volts	0100
1.5 volts	0011
1 volt	0010
0.5 volts	0001
0 volts	0000

Codificación: La codificación consiste en traducir los valores obtenidos durante la cuantificación a un código; por ejemplo: Binario, HDB3, AMI, entre otros.

Criterio de Nyquist

If a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $1/(2B)$ seconds apart.

Si una función $x(t)$ no contiene frecuencias mas altas que B , es completamente determinada dando sus ordenadas en series de puntos espaciados por $1/(2B)$ segundos.

En otras palabras:

Una señal $x(t)$ de ancho de banda B (en banda base), debe ser muestreada a una frecuencia mínima de $2B$ para poder ser recuperada.

$$\text{frecuencia de muestreo} = f_s = 2B$$

Recuperación de la señal

$x(t)$ se puede recuperar totalmente a partir de sus muestras mediante la siguiente función de interpolación:

$$x(k) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{f_s}\right) g\left(k - \frac{n}{f_s}\right)$$

Donde:

$$g(t) = \frac{\text{sen}(2\pi Bt)}{2\pi Bt}$$

En resumen

- Cuando digitalizamos una señal de ancho de banda B . Podrá ser reconstruida a señal analógica si es muestreada a $2B$ como mínimo.
- La alta resolución de los niveles de cuantización reduce el nivel de ruido de esta etapa, aunque aumenta la cantidad de información a almacenar.

Muestreo y reconstrucción

```
% El programa muestrea una señal senoidal adelantada 90 grados y la  
% reconstruye
```

```
frec=2; %Frecuencia de la señal a muestrear  
fs=4; %Frecuencia de muestreo  
Ts=1/(fs); %Periodo de muestreo  
t=0:Ts:1; %Tiempo discreto  
seno_s=sin(2*pi*frec*t+(pi/2));  
plot(t,seno_s)  
hold on  
stem(t,seno_s)
```

```
% Funcion Sa
```

```
figure  
ti=0:Ts/30:1  
sa=(sin(2*pi*frec*(ti))./(2*pi*frec*(ti)));  
plot(ti,sa)
```

```
MuestrasR=50 % Cantidad de muestras para reconstruir
```

```
for k=1:MuestrasR  
    for n=1:length(t)  
        if(n==1)  
            x(k)=seno_s(n)*Sa(frec,((k-1)/MuestrasR)-n*Ts);  
        else  
            x(k)=x(k)+seno_s(n)*Sa(frec,((k-1)/MuestrasR)-n*Ts);  
        end  
    end  
end  
k=0:1/MuestrasR:1-(1/MuestrasR)  
figure  
plot(k,x)
```

```
function g=Sa(B,tiempo)
```

```
if(tiempo==0)  
    g=1;  
else
```

```
g=sin(2*pi*B*tiempo)/(2*pi*B*tiempo);  
end
```

Procesamiento en el dominio espacial

Procesamiento Básico

- Identidad
- Inversa (negativo)
- Umbral
- Umbral Binario (biumbral)
- Umbral Binario Invertido (biumbral invertido)
- Umbral de escala de grises
- Adición
- Sustracción
- Operaciones lógicas
- Operaciones de vecindad
- Histograma

Identidad

El procesamiento de imágenes generalmente se expresa como una función de transformación T :

$$q(x, y) = T(p(x, y))$$

donde:

$p(x, y)$: pixel de la imagen de entrada

$q(x, y)$: pixel de la imagen de salida

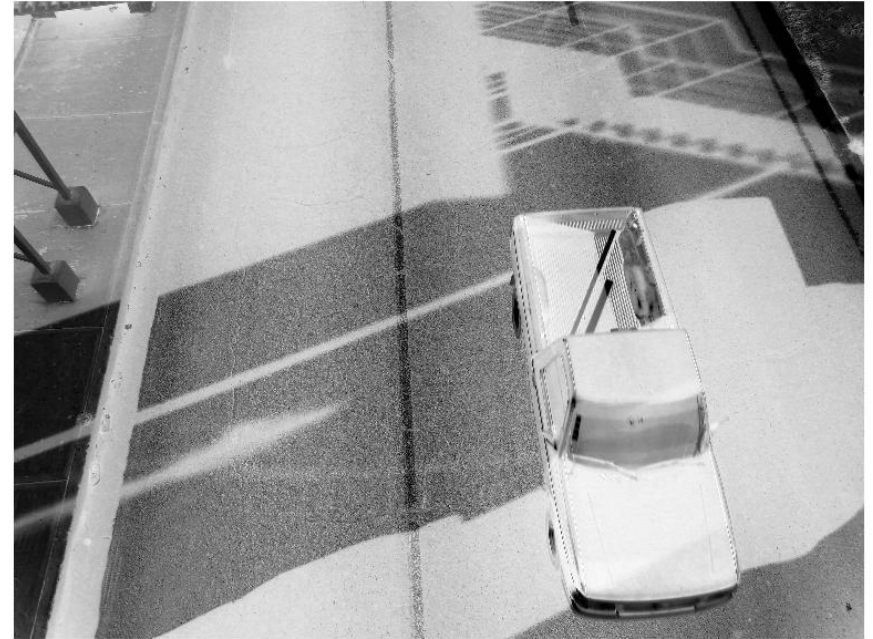
Entonces la transformación identidad crea una imagen idéntica a la imagen de entrada.

$$q = p$$

Inversa o negativo

$$q = 255 - p$$

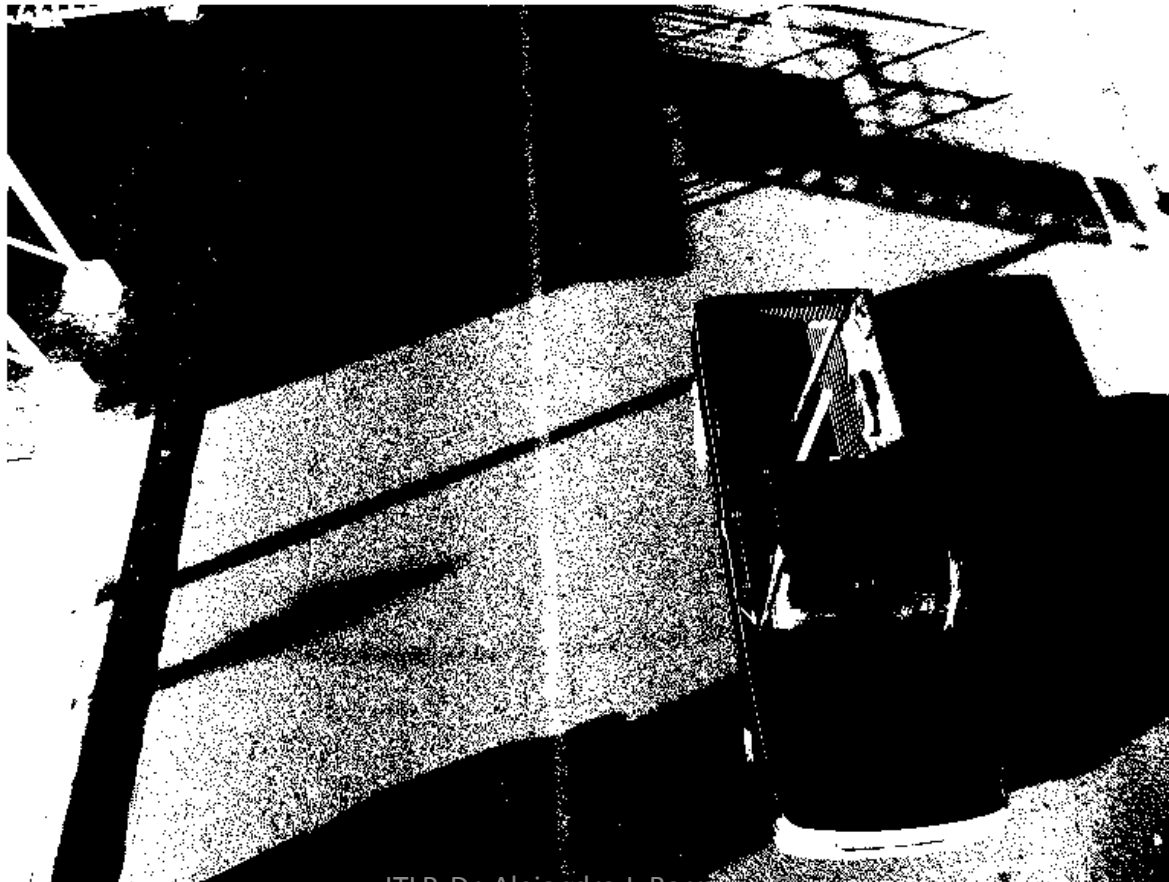
NEG=255-IR;



Umbralización(Thresholding)

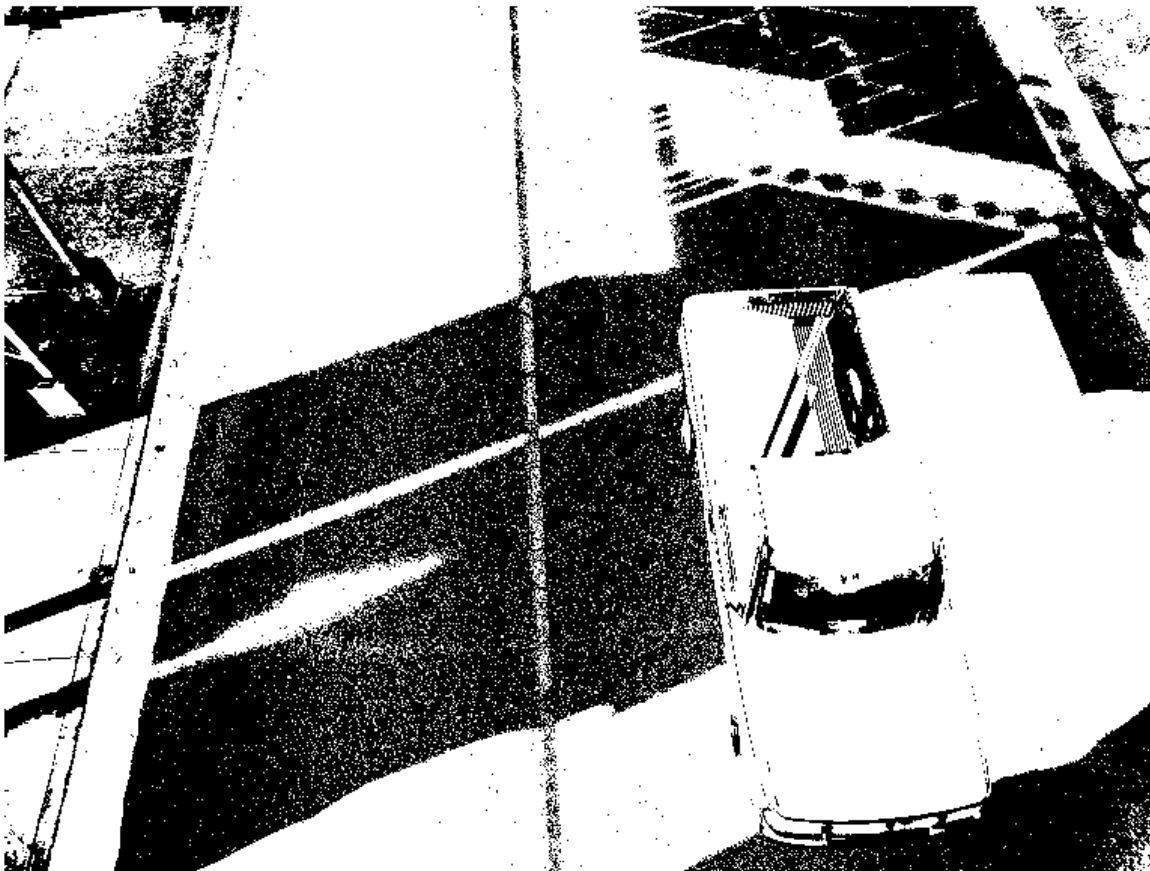
$$b(x, y) = \begin{cases} 0 & \forall f(x, y) \leq \mu_0 \\ 1 & \forall f(x, y) > \mu_0 \end{cases}$$

```
umbralada=im2bw(IR,0.5);  
imshow(umbralada);
```



Bi-umbral

$$q = \begin{cases} 255 & \forall p(x,y) \leq \mu_1 \text{ or } p(x,y) \geq \mu_2 \\ 0 & \forall \mu_1 < p(x,y) < \mu_2 \end{cases}$$

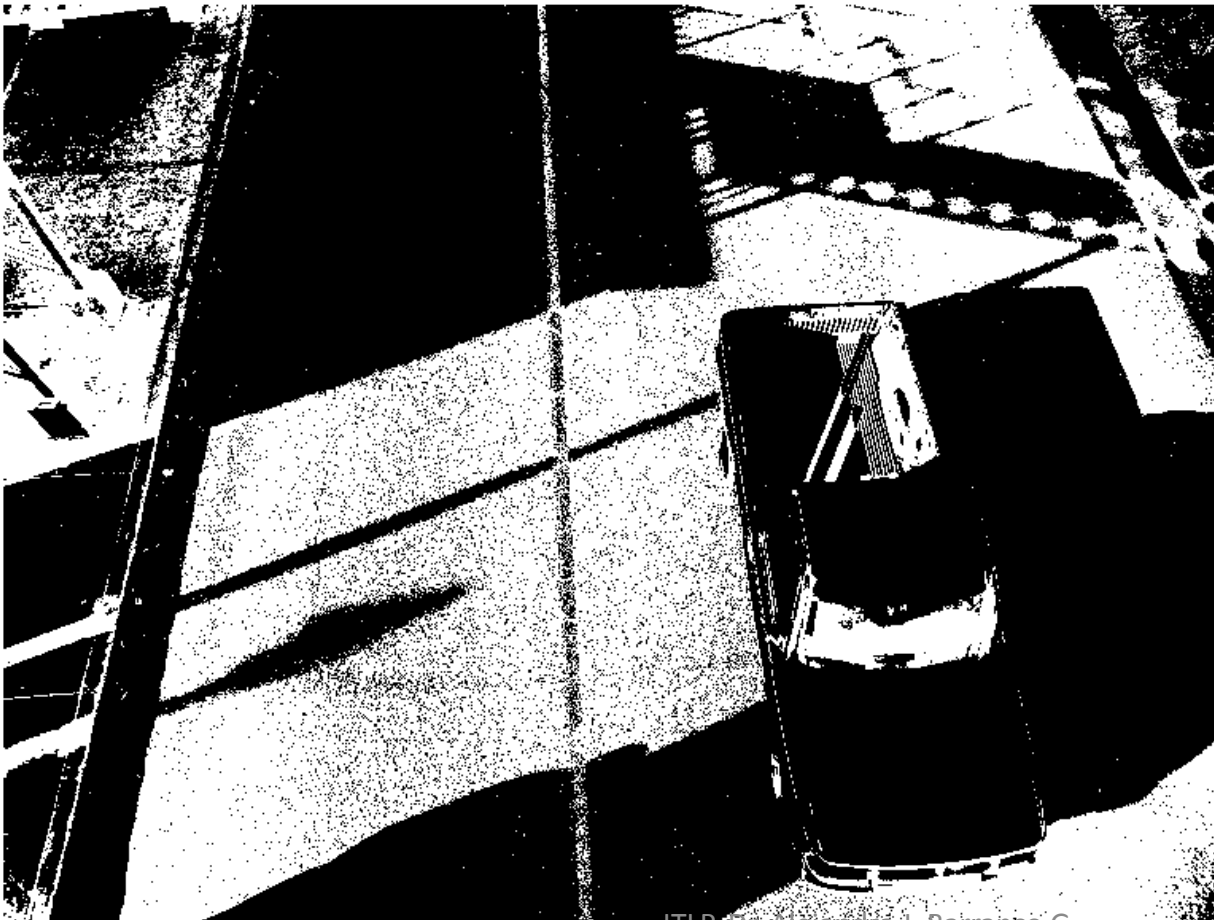


```
[M,N]=size(IR);  
biumbralada=ones(M,N);
```

```
for i=1:M  
    for j=1:N  
        if (IR(i,j)<=55 || IR(i,j)>=200)  
            biumbralada(i,j)=255;  
        else  
            biumbralada(i,j)=0;  
        end  
    end  
end
```

Umbral Binario Invertido

$$q = \begin{cases} 0 & \forall p(x, y) \leq \mu_1 \text{ or } p(x, y) \geq \mu_2 \\ 255 & \forall \mu_1 < p(x, y) < \mu_2 \end{cases}$$

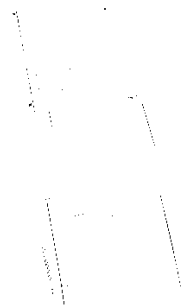


```
for i=1:M
    for j=1:N
        if (IR(i,j)<=100 || IR(i,j)>=200)
            biumbralada(i,j)=0;
        else
            biumbralada(i,j)=255;
        end
    end
end
```

Umbral de escala de grises

$$q = \begin{cases} 255 & \forall p(x, y) \geq \mu_2 \\ p(x, y) & \text{otro caso} \end{cases}$$

Umbral grises



```
for i=1:M
    for j=1:N
        if (IR(i,j)>=254)
            umbraladagrises(i,j)=255;
        else
            umbraladagrises(i,j)=IR(i,j);
        end
    end
end
```


Adición



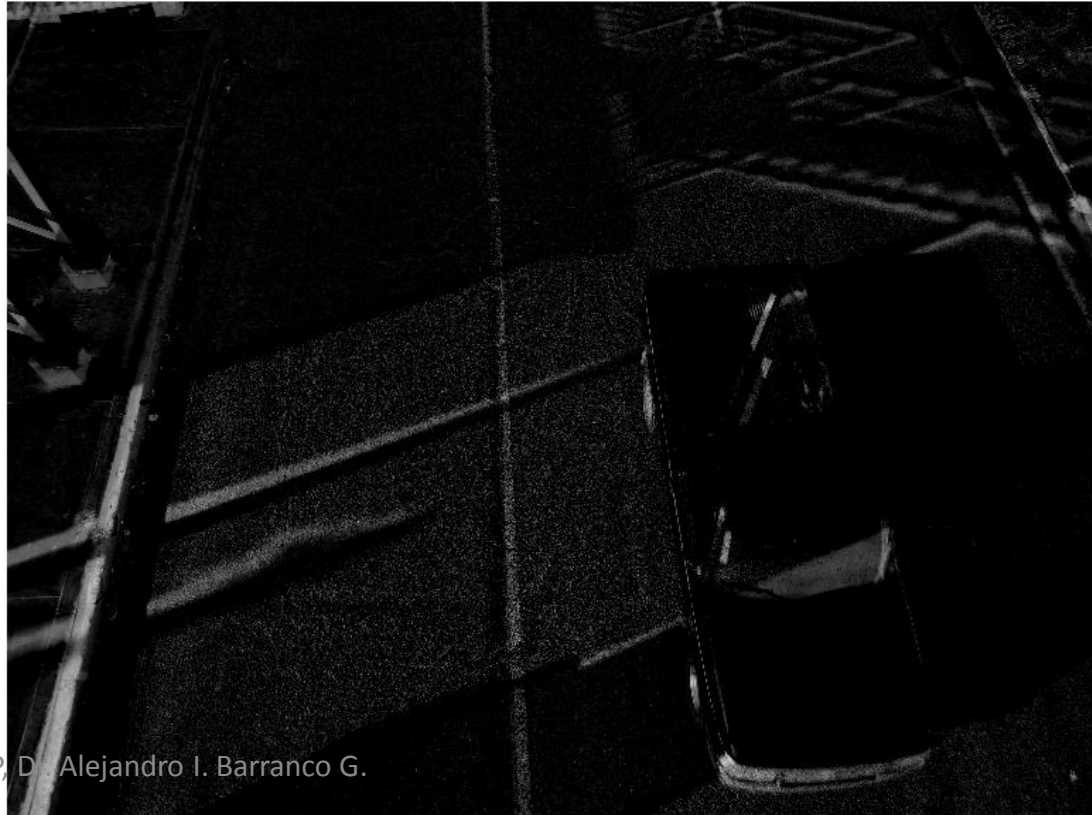
$$S = p(x, y) + q(x, y)$$

$$S = IR + IR2;$$



Sustracción

$$R(x, y) = p(x, y) - q(x, y)$$



Operaciones lógicas

Para realizar estas operaciones, es necesario contar con imágenes binarias (valores Falso y Verdadero). Las tablas de verdad de las operaciones lógicas son las siguientes:

NOT

p1	Ps
0	1
1	0

AND

p1	p2	ps
0	0	0
0	1	0
1	0	0
1	1	1

OR

p1	p2	ps
0	0	0
0	1	1
1	0	1
1	1	1

Operaciones de vecindad

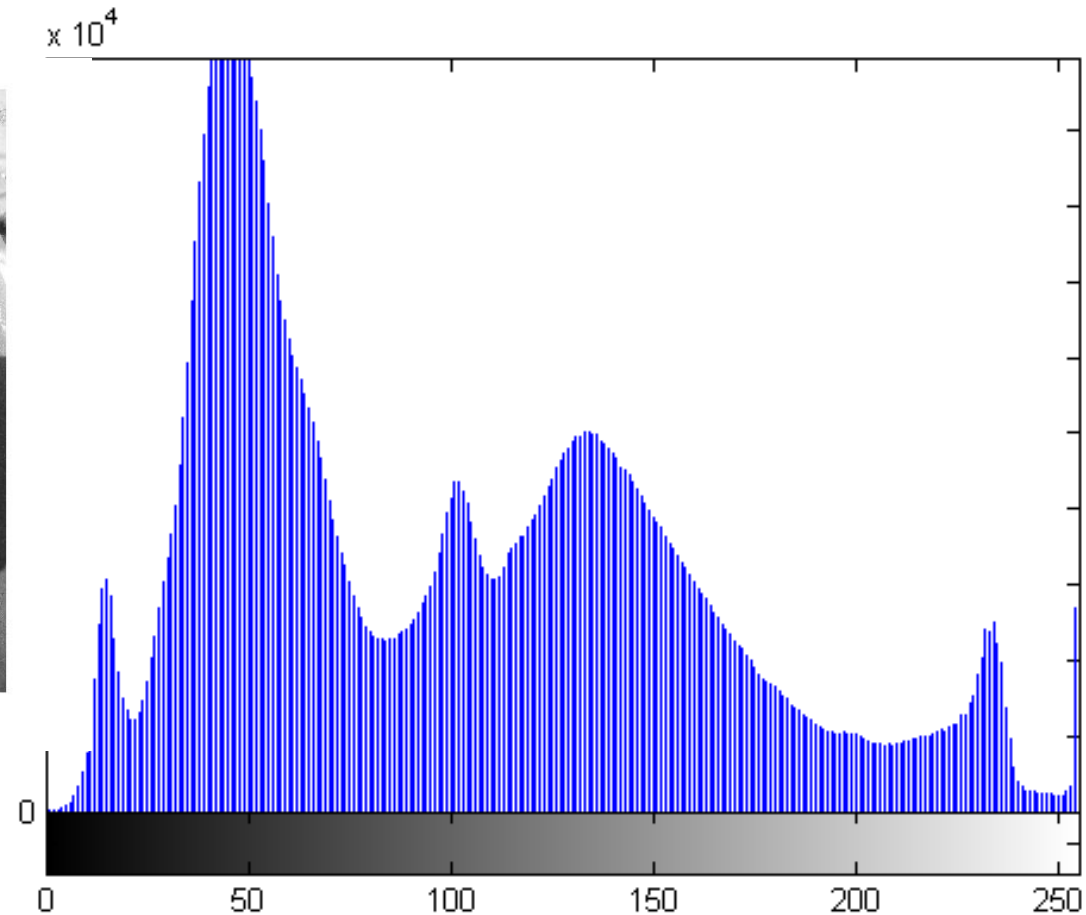
$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\begin{bmatrix} p(x-1, y-1) & p(x, y-1) & p(x+1, y-1) \\ p(x-1, y) & p(x, y) & p(x+1, y) \\ p(x-1, y+1) & p(x, y+1) & p(x+1, y+1) \end{bmatrix}$$

$$q(x, y) = 1 * p(x-1, y-1) + 2 * p(x, y-1) + 1 * p(x+1, y-1) + 0 * p(x-1, y) + 0 * p(x, y) + 0 * p(x+1, y) + (-1) * p(x-1, y+1) + (-2) * p(x, y+1) + (-1) * p(x+1, y+1)$$



Histograma



```
figure  
imhist(IR)
```

El histograma cuenta la cantidad de pixeles en cada nivel de gris de la imagen.

ITLP, Dr. Alejandro I. Barranco G.