

Tratamiento de Imágenes por computadora

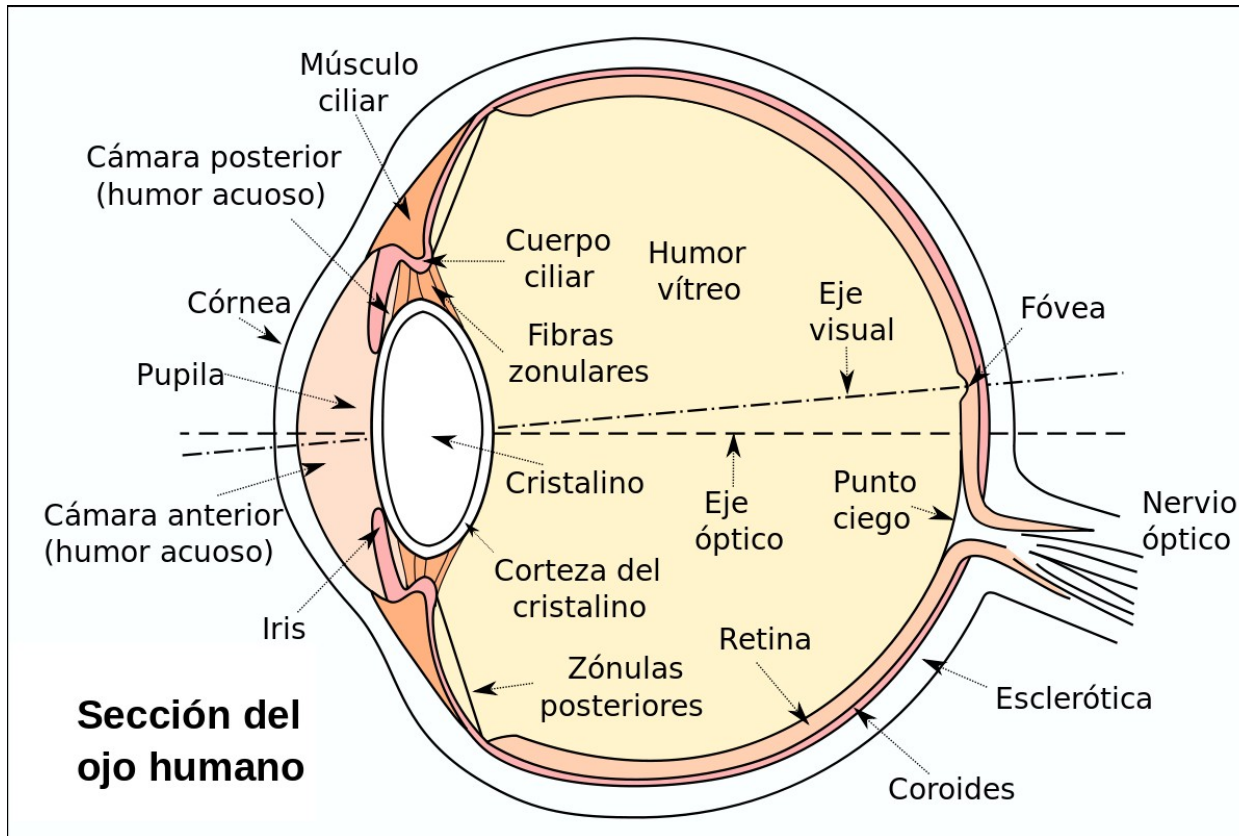
# Conceptos básicos

Facultad de Ingeniería, Universidad de la República

# Agenda

- Ojo humano
- Modelo de cámara pinhole
- Lentes
- Formación de imagen
- Digitalización

# El ojo humano

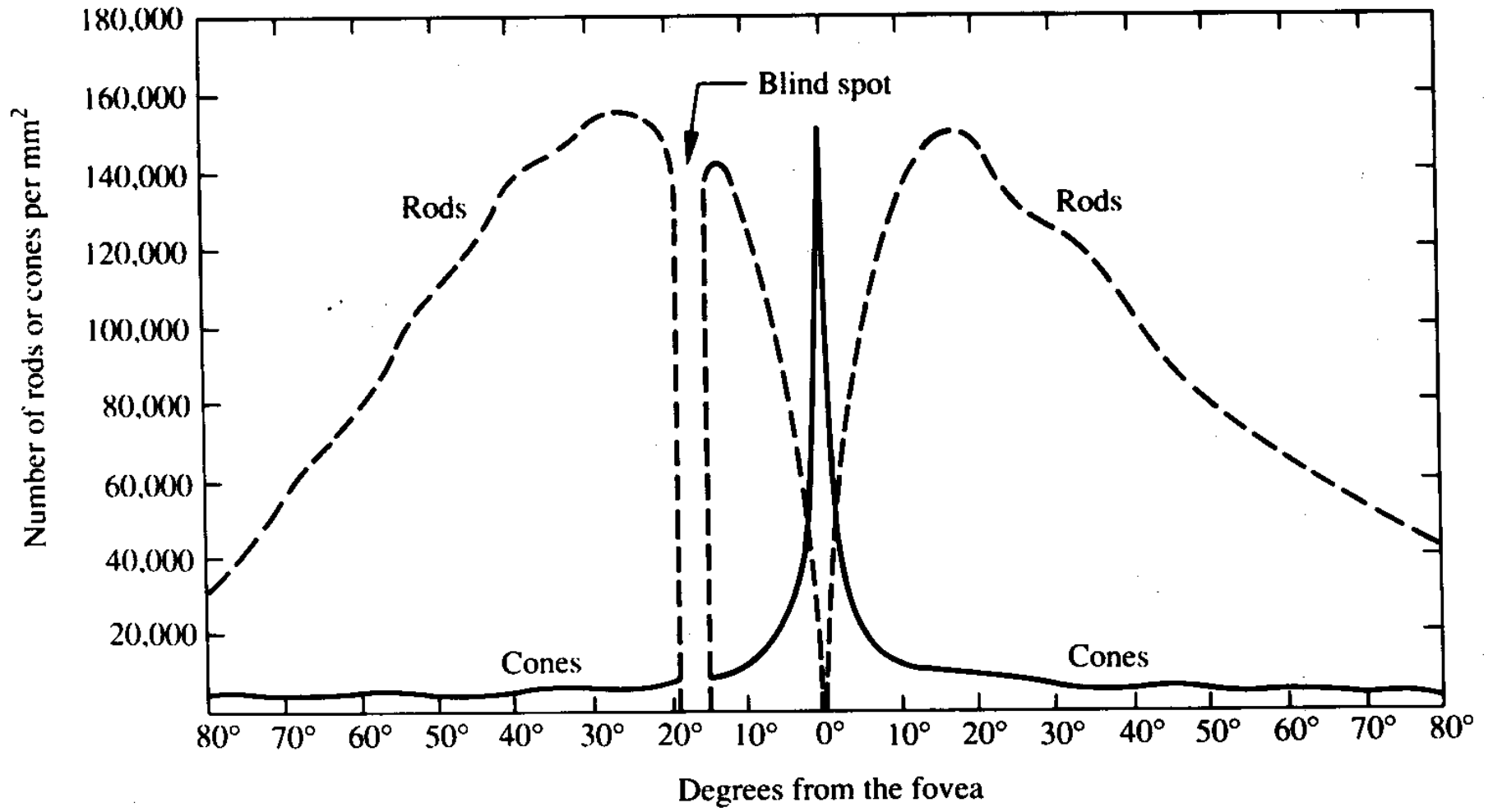


Sacado de wikipedia

- Un sistema de cámara oscura con un campo visual de  $160^\circ$  (ancho) x  $135^\circ$  (alto):
  - **Córnea**: protección externa, refracción de la luz
  - **Iris y pupila**: diafragma (1 a 8 mm), controla la cantidad de luz. Dilata si la iluminación es baja. En condiciones normales de iluminación se contrae para reducir la aberración esférica.
  - **Cristalino**: lente adaptativo y filtro
  - **Retina**: arreglo de receptores fotosensibles, donde se forma la imagen
  - Otros: humor vítreo, tejido conjuntivo (nutrientes e hidratación)

# Receptores de luz

- **Conos** (6-7 millones):
  - Son sensibles al color: 3 tipos de conos con distinta sensibilidad espectral debido a fotoreceptores específicos (fotopsinas).
  - Concentrados en la zona central: *macula lutea*, *fóvea*, zona de enfoque fino.
  - Permiten ver detalles finos girando el ojo.
- **Bastones** (75-150 millones):
  - No son sensibles al color.
  - Dan una visión general pero poco detalle.
  - Fotoreceptores extremadamente sensibles, gran rango de iluminación



# Distribución de receptores

- Distribución circular.
- **Fóvea:**
  - región central con densidad de conos de unos 150000 conos por  $\text{mm}^2$ .
  - Un CCD de  $7 \text{ mm}^2$  tiene ese número de elementos...
  - 120 conos por grado, i.e.  $\sim 2$  micras de espaciamento

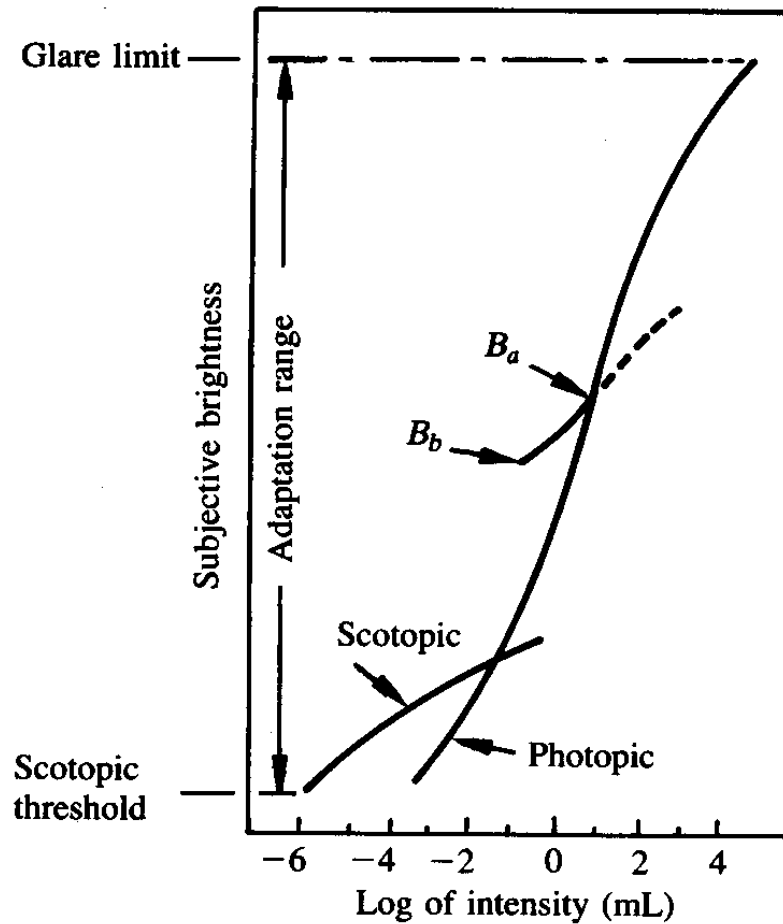
# Distribución de receptores

- Se detectan objetos con la visión periférica (sobre todo bastones)
- Se examinan objetos de interés paseando la fovea (conos) con movimientos rápidos (*sacadas oculares*)



# Adaptación a la iluminación

- Rango dinámico enorme:
  - $10^{10}$  desde el umbral escóptico al deslumbramiento (luz de una estrella a luz del mediodía)
- Se adapta a un nivel y resuelve un entorno.
- En realidad se discriminan pocos niveles de gris en un entorno local.

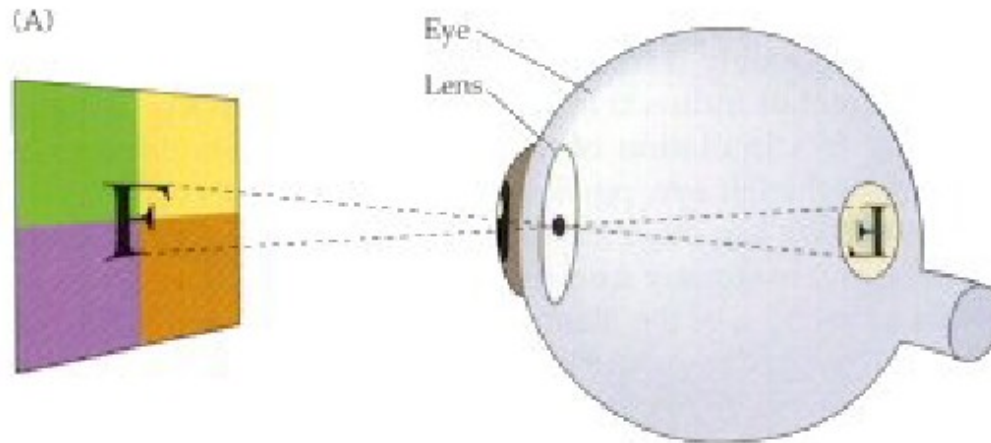


*f subjective brightness sensations showing a particular*

Sacado de "Digital Image Processing" Rafael Gonzalez & Richard Woods

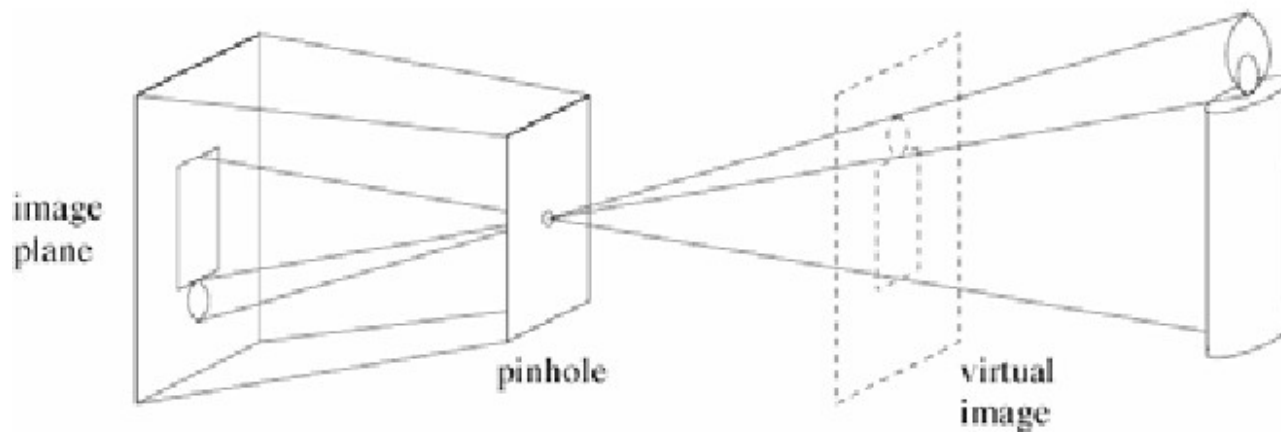
# Modelo de cámara pinhole

- El ojo se puede modelar como una cámara pinhole:



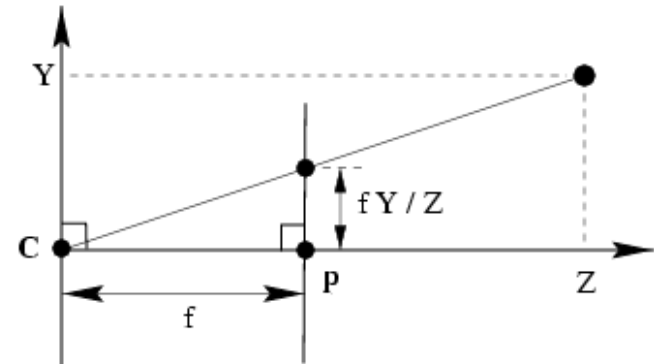
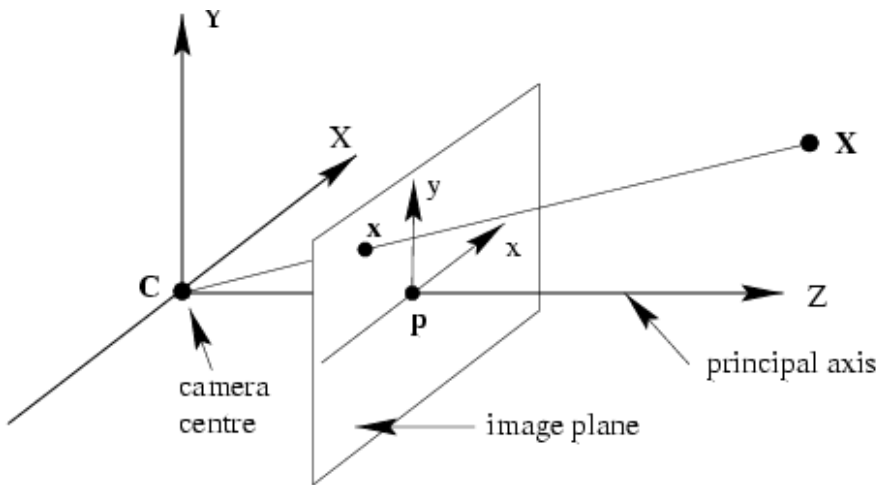
# Modelo de cámara pinhole

- Los rayos de luz pasan por un orificio
- Se forma una imagen invertida en el plano de imagen
- Imagen virtual no invertida



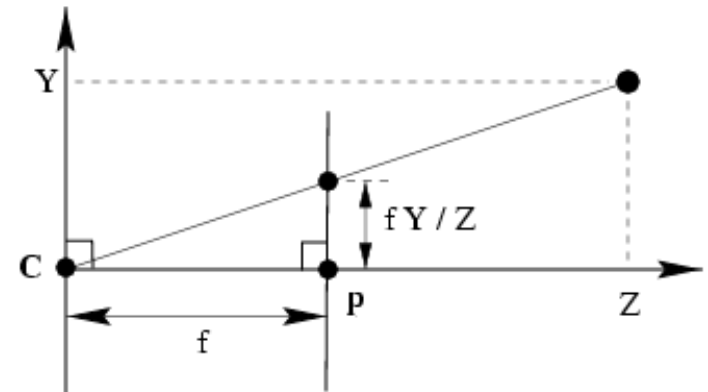
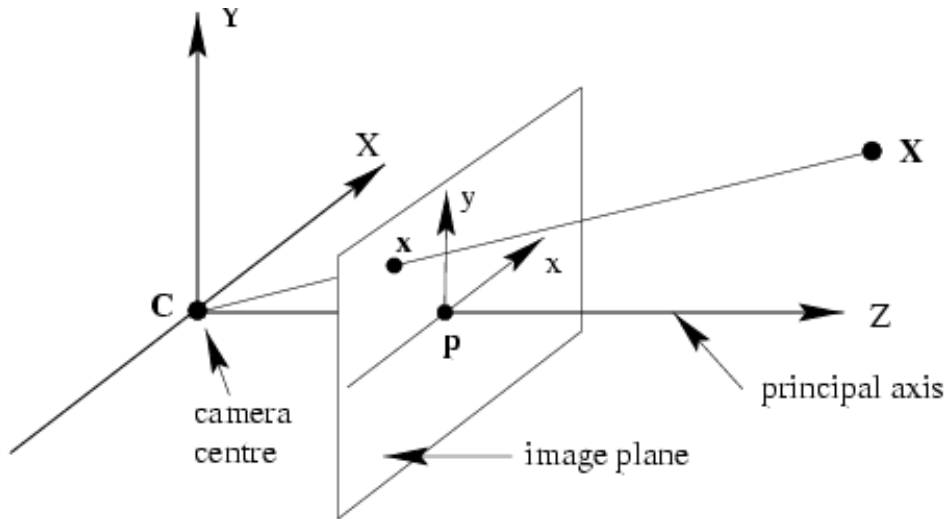
# Terminología

- Centro de la cámara
- Plano de imagen, plano principal
- Eje principal
- Distancia focal



# Proyección perspectiva

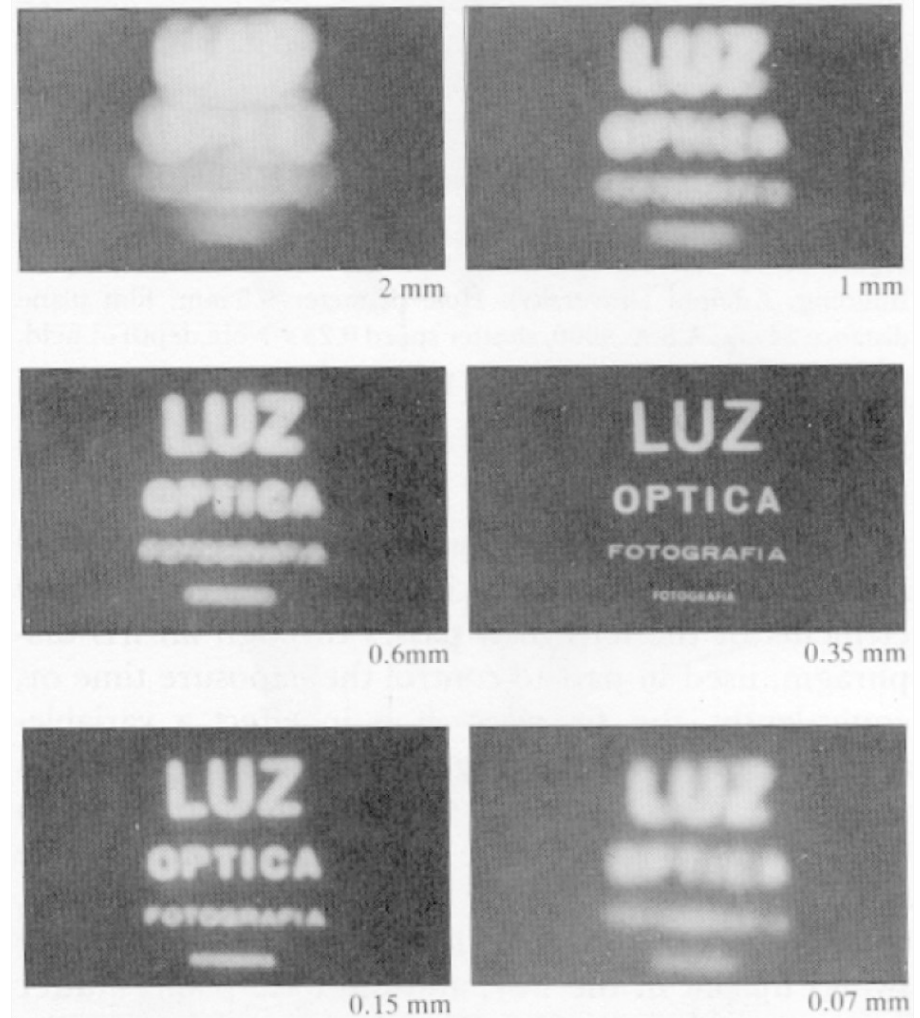
- Punto del espacio  $(X, Y, Z)$
- Punto de la imagen  $(x, y)$
- $x = f X / Z$  ,  $y = f Y / Z$



# Cámara pinhole en la práctica

- Se puede implementar
- El orificio permite pasar poca luz
- Problemas de fuera de foco (orificio muy grande), difracción (muy chico)
- Para tener una cámara utilizable hay que usar lentes. Permiten formar imagen con mayor ingreso de luz

- De izq. a der., de arriba a abajo:  
2mm, 1mm,  
0.6mm, 0.35mm,  
0.15mm, 0.07mm





# Lentes... ¿para que?

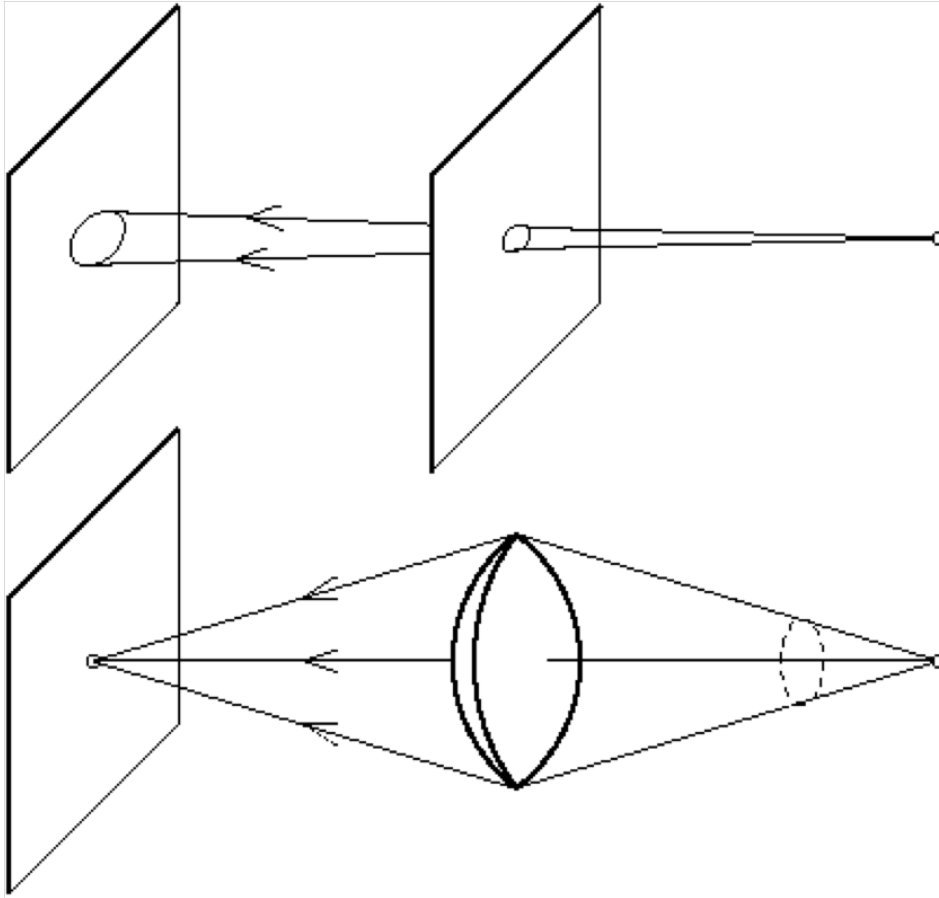


Figura cortesía de D. Forsyth

# Lentes

- Idealmente una lente fina. Forma la imagen prácticamente como el pinhole
- En la práctica las lentes introducen problemas:
  - Aberraciones geométricas
  - Distorsiones geométricas
  - Vignetting (menor pasaje de luz en los bordes del lente)
  - Aberración cromática (las diferentes longitudes de onda en la luz siguen caminos diferentes)

- Ej.: Aberración esférica

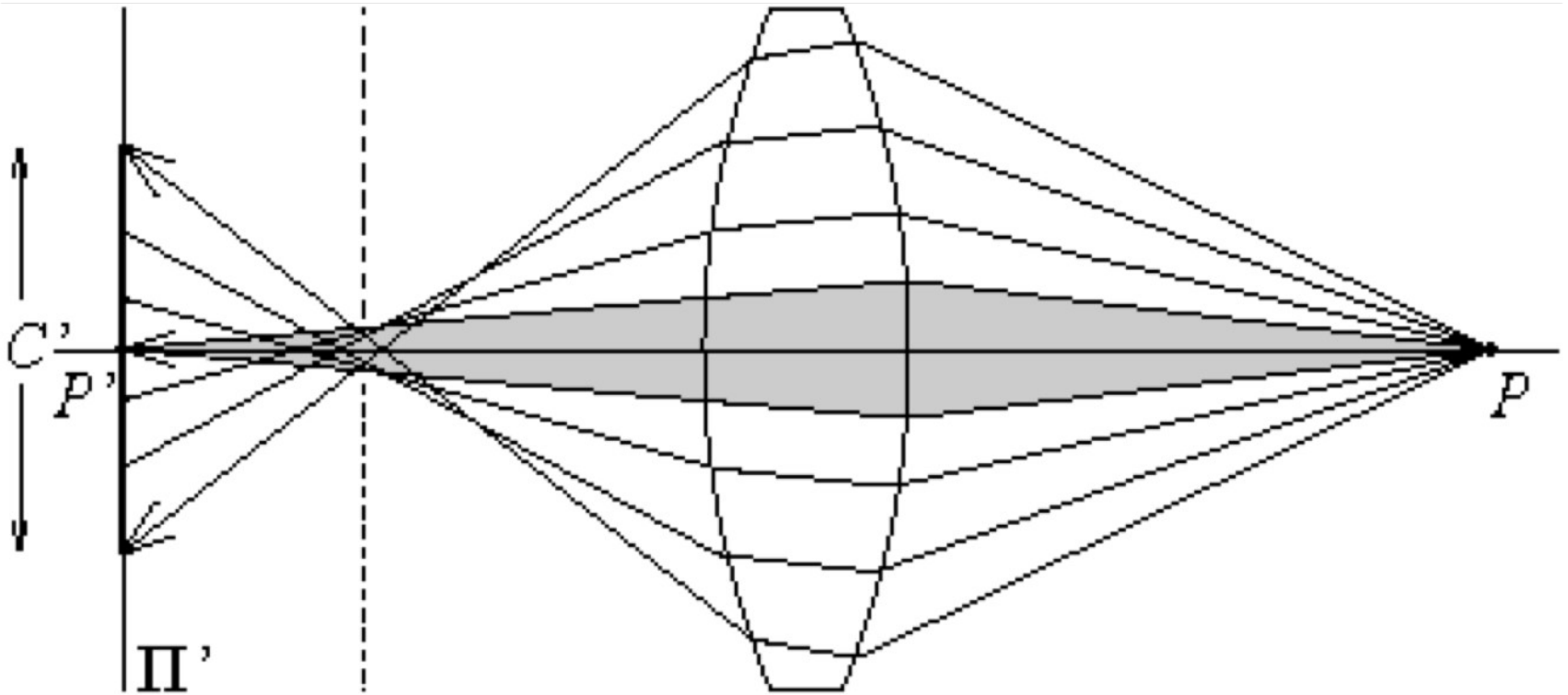


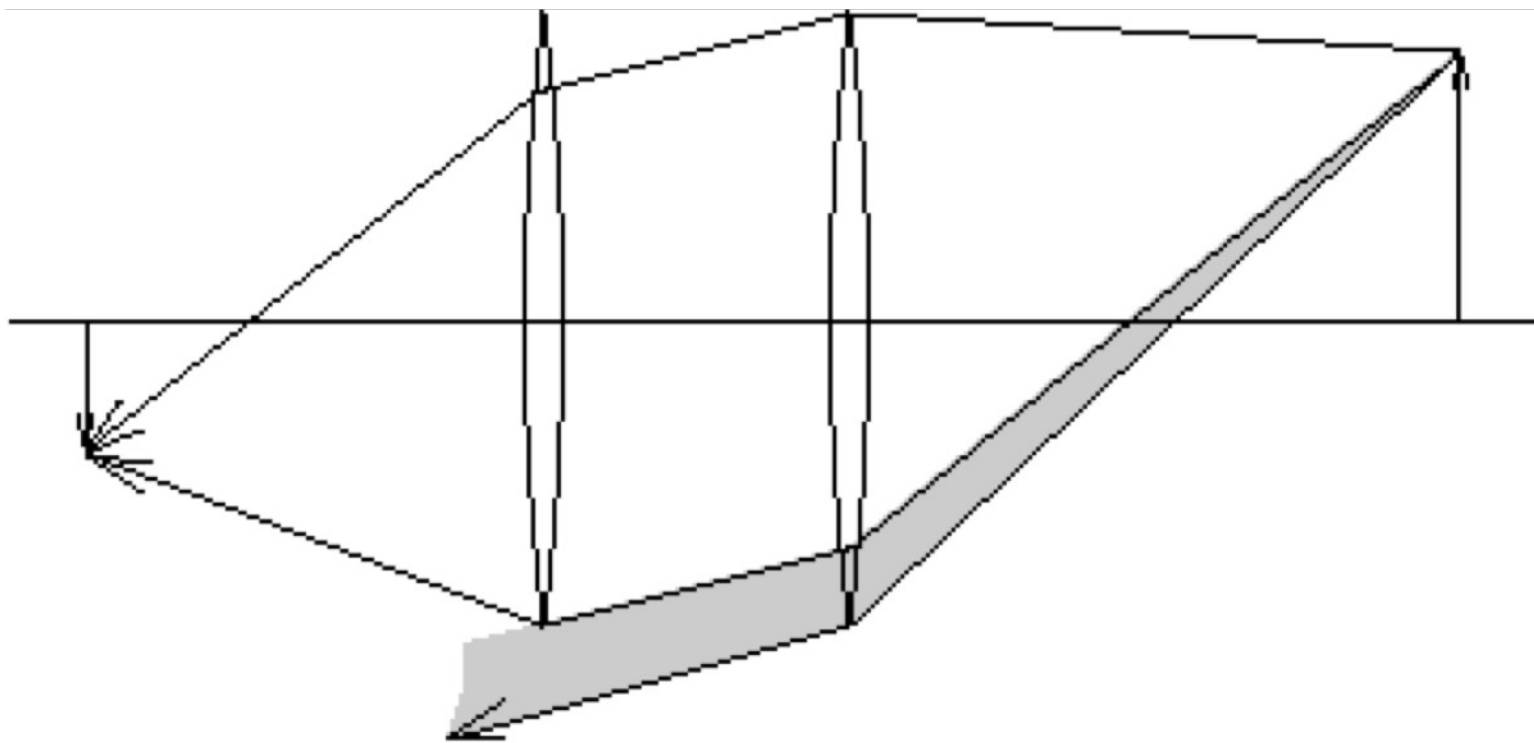
Figura cortesía de D. Forsyth

# Distorsión radial y corrección



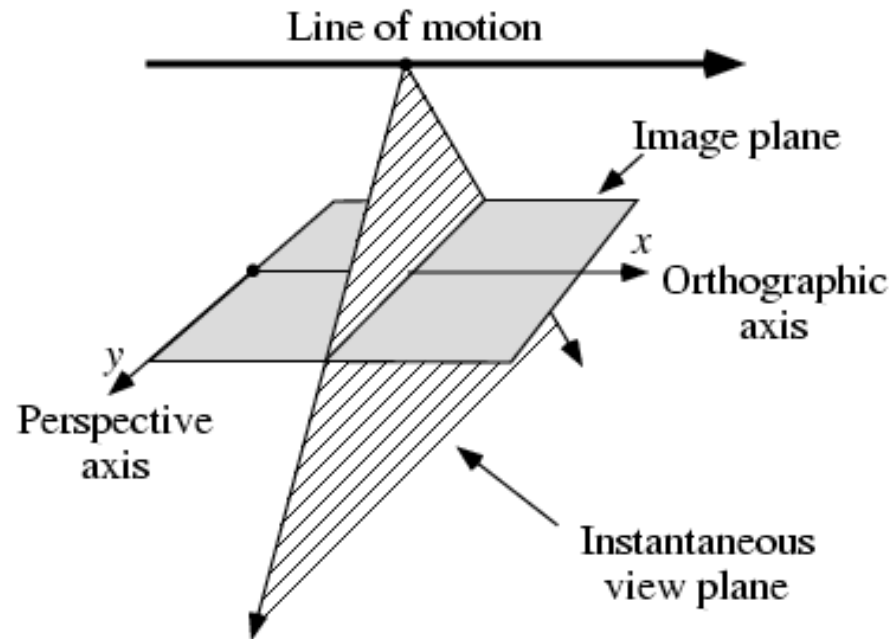
Sacado de proyecto de fin de carrera, Musé y Azar

## Vignetting (sistemas de dos o más lentes)



# Otros tipos de cámaras

- Satélites SPOT – cámara de barrido



# Modelo de formación de la imagen

- Valor de la imagen  $f$  en el espacio es la intensidad de la iluminación en ese punto.
- Es una energía: positiva y finita.

$$0 < f(x, y, \lambda) < \infty \quad 350\text{nm} < \lambda < 780 \text{ nm}$$

- Luz incidente y luz reflejada:

$$f(x, y, \lambda) = i(x, y, \lambda) r(x, y, \lambda)$$

$$0 < i(x, y, \lambda) < \infty \quad \text{y} \quad 0 < r(x, y, \lambda) < 1$$

(absorb. + transmit. + reflejada (+difundida) = 1)

# Formación de imagen

- Existe proporcionalidad entre la cantidad de luz emitida por un punto de la escena y la cantidad de luz incidente en la imagen de ese punto



# Formación de imagen

- La cantidad de luz emitida por un punto de la escena depende de:
  - Iluminación que llega a ese punto
  - Propiedades reflectivas de la superficie del objeto

# Luz incidente y reflejada

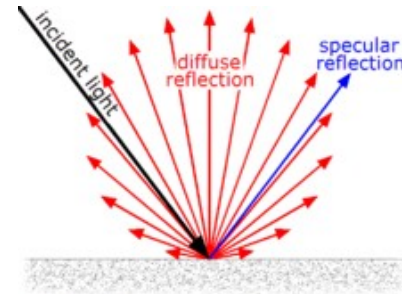
- Rangos típicos de  $i(x,y,\lambda)$ :
  - Día claro: 9000
  - Día nublado: 1000
  - Luna llena: 0.01
  - Oficina comercial iluminada: 100
- Rangos típicos de  $r(x,y,\lambda)$ :
  - terciopelo negro: 0.01
  - Acero inoxidable: 0.65
  - Muro blanco: 0.8
  - Lámina metálica plateada: 0.9
  - nieve: 0.93

# Superficies lambertianas

- Para algunas superficies no importa de donde se las ilumine reflejan igual en todas las direcciones
- Superficies perfectamente difusivas

# Superficies especulares

- Superficies tipo espejo.
  - Reflejan en la dirección especular a la de incidencia con respecto a la normal
  - En superficies reales es más bien un lóbulo en torno a la dirección especular (**modelo de phong**)



Fuente: wikipedia

# Lambertiano + especular

- En general se considera una mezcla ponderada de los dos
- Hay muchas superficies que no lo cumplen (ej CD)

# Digitalización

- Se pasa de una imagen “óptica” a una imagen “eléctrica”

# Del mundo analógico al digital...

- Se produce una discretización espacial y del valor de la función en cada punto.
- Rango dinámico: rango útil de sensibilidad del sensor, en general menor que el rango de variación del fenómeno físico.
- Resolución: detalles distinguibles una vez producida la discretización.

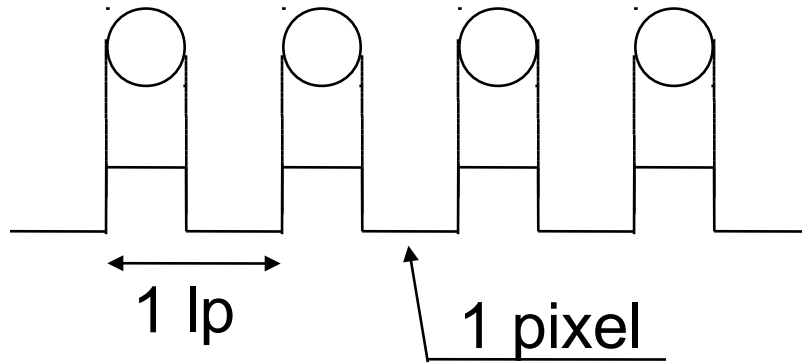
# Contenido de bits.

- Determinado por el número de elementos de imagen (pixels) y el número de pasos de discretización de cada pixel.
- Ruido (grano fotográfico, ruido electrónico, etc.).
- No es interesante un paso de discretización menor que el ruido presente...



# Resolución espacial.

- Se expresa en pares de líneas por milímetro (lp/mm).



# Qué resolución?

- Es necesario una resolución mayor que el doble del menor detalle a observar.
- En principio podemos resolver detalles que son mayores a 2 veces un par de líneas.

# Sensores

- La retina puede ser discreta: un conjunto de sensores (CCD).
- La retina puede ser continua: film fotográfico, tubo de rayos catódicos, placas de fósforo luminiscente, etc.

# sensores

- Un punto más chico que un pixel será visto?
- Su energía se difundirá en uno o varios píxeles dando una señal más débil pero eventualmente perceptible.

# Ruido

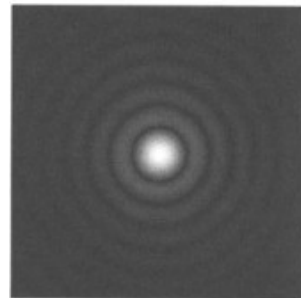
- Cuántos fotones impactan el sensor por unidad de área y de tiempo? Esto es probabilístico.
- Crece la relación entre pixel y objeto y crece la probabilidad de detección.
- Crece el tiempo de integración y crece la probabilidad de detección.

# PSF (Point Spread Function)

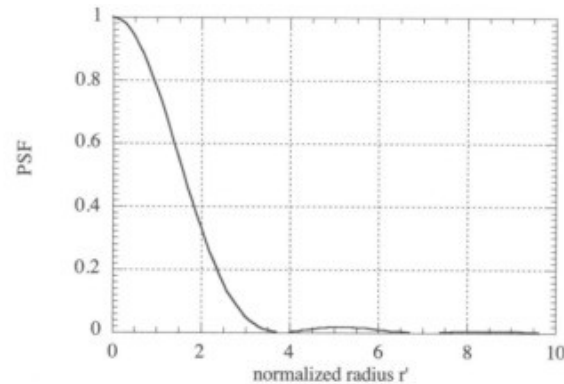
- Es la respuesta al impulso del dispositivo óptico (cámara).
- En un dispositivo óptico ideal, la PSF es el núcleo de difracción introducido por una apertura de diámetro  $D$ .
- **Airy Pattern** (para onda monocromática de longitud  $\lambda$  y distancia focal  $f$ ) :  $\rho = 1.22 \lambda f / D$
- Donde  $\rho$  es la separación de 2 objetos en la imagen para verlos sin blur.
  - Descripción de lo que produce un spot de luz solo limitado por la difracción.

# Airy Pattern

- Medida de resolución: representa el grado de detalle que puede resolver la cámara



*Airy Pattern*



# MTF (Modulation Transfer Function)

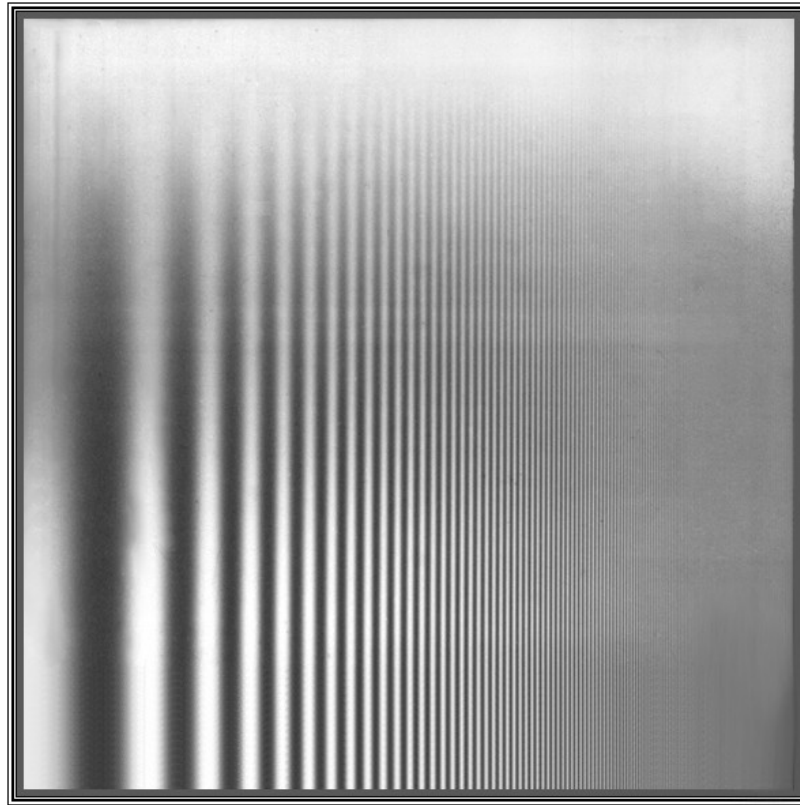
- **Módulo de la Transformada de Fourier de la PSF.** Caracteriza la respuesta del sistema de adquisición en su conjunto. Es una mejor caracterización que la resolución.
- Es la respuesta en amplitud en función de la frecuencia espacial.
- Se normaliza estimando que la respuesta de amplitud es 100% a la frecuencia 0.



# MTF

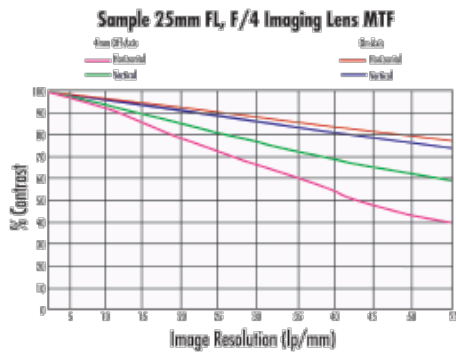
- **¿Cómo se mide?**: Se excita el sistema con un patrón sinusoidal cuya frecuencia varía y se mide su respuesta. Se puede aproximar por una onda cuadrada.
- Rayos X: patrón hecho de barras de tungsteno espaciadas de forma decreciente.
- Microscopio: patrón de barras de espaciamiento decreciente observado.

# MTF del sistema visual

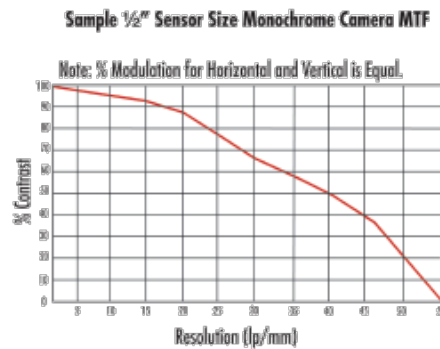


Tomado de <https://www.appliedimage.com/product/sine-lsm-1-tm-g/>

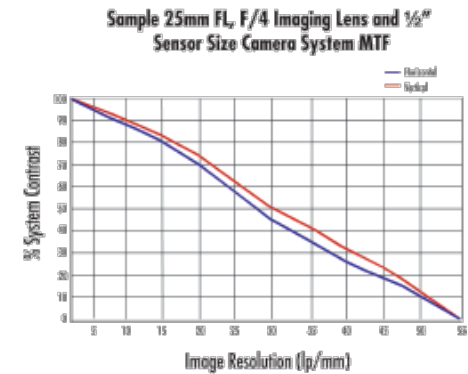
# MTF de una cámara/lente



X



||



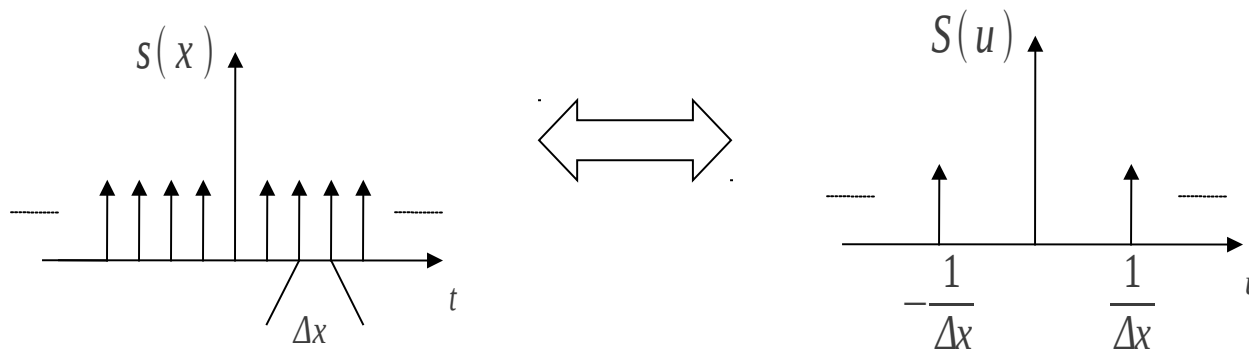
Extraído de <https://www.edmundoptics.com/knowledge-center/application-notes/optics/introduction-to-modulation-transfer-function/>

# Muestreo en 2D

- En una imagen discretizada espacialmente y en niveles de gris.
  - imagen 2D: pixel.
  - imagen 3D: voxel.
- Una imagen puede describirse de varias formas: lista de píxeles, líneas de nivel, regiones homogéneas, etc.

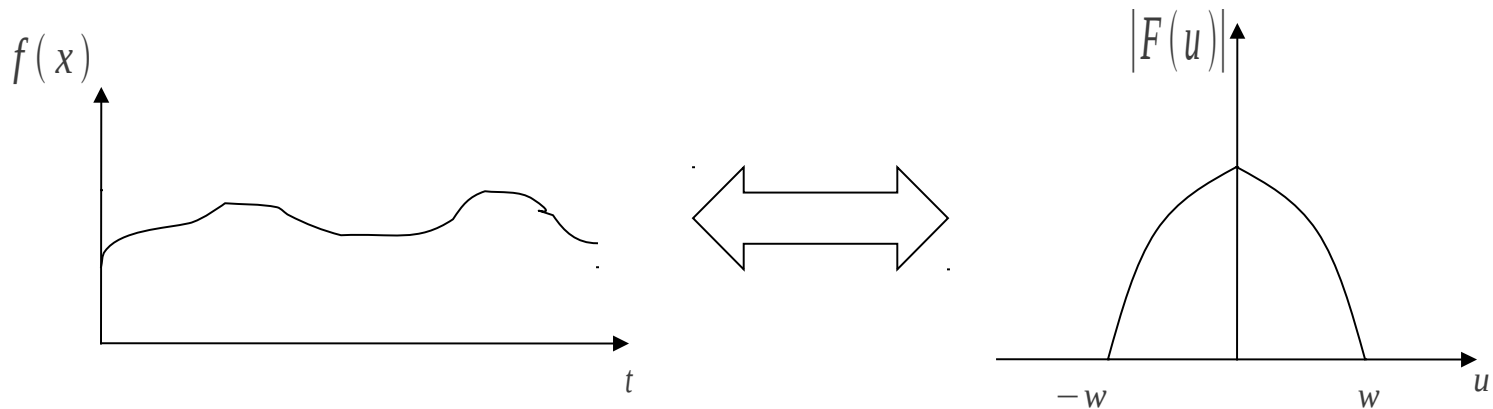
# Delta de Dirac

- $\delta(x - x_0)$  : señal cuya área es 1 en un entorno infinitesimalmente pequeño en torno a  $x_0$  y cero en el resto del espacio.
- $s(x)$  : tren de impulsos separados  $\Delta x$



# Función a banda limitada

- Señal cuya TF es nula para valores fuera de un intervalo  $[-W, W]$



# En dos dimensiones...

- Imagen muestreada. Producto de un tren de deltas de Dirac por la función. Sobrevive  $f$  en el soporte de los deltas:

$$f_s(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m\Delta x, n\Delta y) \delta(x - m\Delta x, y - n\Delta y)$$

- TF de la imagen muestreada:

$$F_s(\xi_1, \xi_2) = \xi_{xs} \xi_{ys} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F(\xi_1 - k\xi_{xs}, \xi_2 - l\xi_{ys})$$

# Imagen muestreada

- La TF de la imagen muestreada es una replicación de la TF de la imagen de entrada en una grilla definida por  $(\xi_{xs}, \xi_{ys})$
- Si el espectro de la imagen es a banda limitada podemos recuperarla interpolando a partir de las muestras.

- Condición:

$$\xi_{xs} > 2\xi_{x0} \quad \xi_{ys} > 2\xi_{y0}$$

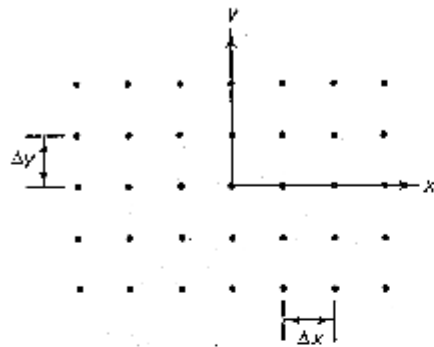
- Donde:

$$\xi_{xs} = \frac{1}{\Delta x} \quad \xi_{ys} = \frac{1}{\Delta y}$$

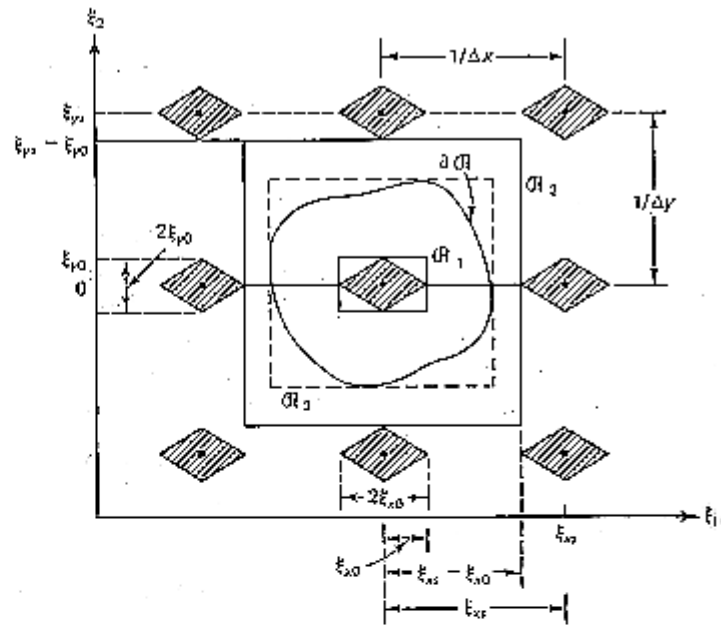


# Nyquist

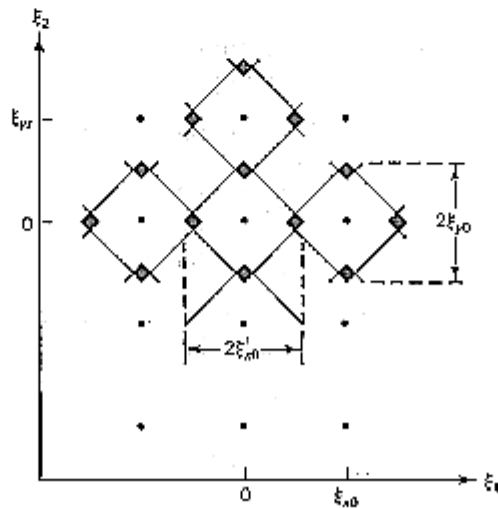
- Se puede recuperar una imagen a banda limitada si está muestreada a frecuencias superiores al doble del ancho de banda  $\xi_{x0}$   $\xi_{y0}$
- En caso contrario aparece el aliasing (superposición de frecuencias)



(a) Sampling grid.



(b) Sampled image spectrum.



(c) Aliasing and foldover frequencies (shaded areas).

# Aliasing

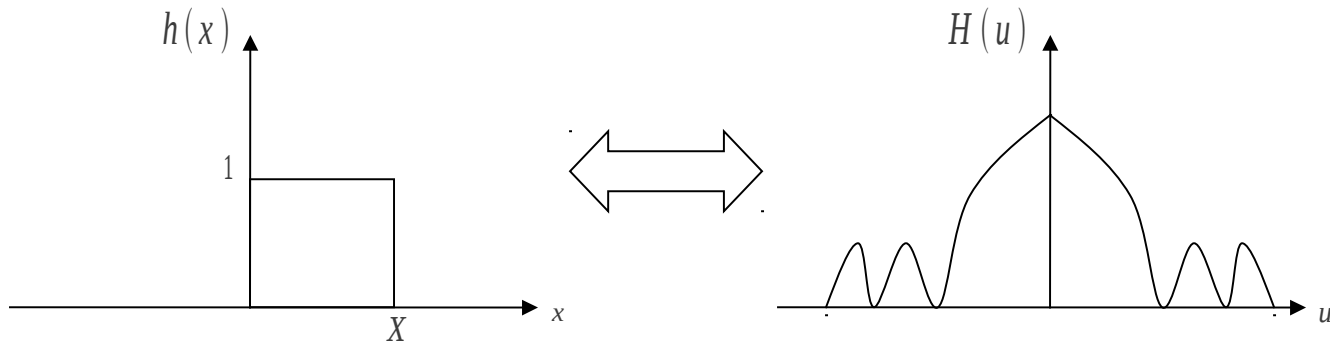
- Si se satisface el teorema de Shannon entonces es posible recuperar la señal original con un filtro pasabajo que deje pasar las frecuencias menores que  $W$ .
- Si no se satisface Shannon entonces en la señal recuperada aparecen contaminaciones por componentes de frecuencia de los periodos adyacentes. Filtro de interpolación

# Temas

- Las señales no son a banda limitada
- La frecuencia de muestreo impone límites a la resolución.
- El filtro pasa bajos de interpolación no es ideal. Hay un compromiso entre error de interpolación y resolución.

# Trabajamos con señales finitas...

- Muestreamos solo un intervalo. Esto equivale a multiplicar la función  $f(x)$  por una “ventana” rectangular:



# Ventanas

- La señal recuperada es el producto de la señal, el tren de impulsos y la ventana:

$$f_r(x) = h(x)[s(x)f(x)]$$

- La ventana participa del proceso de discretización.
- Cuanto más abrupta es la ventana, más componentes de alta frecuencia tiene, que al convolucionarse con la señal distorsiona todo el contenido espectral de la señal.
- Se usan ventanas con caída suave: hamming, etc.

# Imagen digital

- Señal analógica a la cual de le aplican 2 procesos:
  - Muestreo: Discretizar el espacio.
  - Cuantificación: Discretizar la amplitud.
- En general se utilizan potencias de 2 tanto para el número de líneas y columnas como para el número de niveles de gris.

# Muestreo y cuantificación

- Discretización espacial: muestreo ( $N \times M$ )
- Discretización de amplitud: cuantificación del nivel de gris ( $G$  niveles de gris)
- En general  $N, M$  y  $G$  son potencias de 2.
- Pixel: elemento de imagen. Es una muestra de la distribución « continua » adquirida.

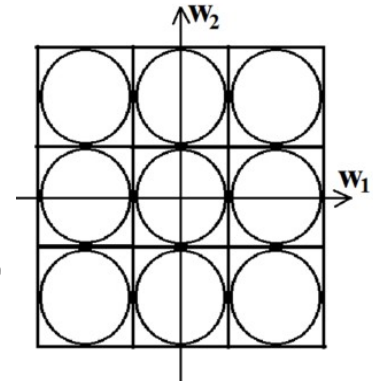


# Resolución

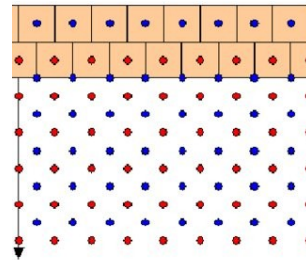
- Qué valores de  $N$ ,  $M$  y  $G$  para una buena resolución?
- Compromiso con potencia de cálculo, memoria, etc.
- Medidas subjetivas de calidad.

# Grillas de muestreo

- Rectangulares
  - Pasos diferentes, relación de aspecto



- *Quincunx*



Latry, C. & Rougé, Bernard. (2003). Super resolution: quincunx sampling and fusion processing. 315 - 317 vol.1

- Hexagonales: tema isotropía.

