



Visión Computacional

Posgrado en Ciencias e Ingeniería de la Computación

UNAM

Semestre 2019-I

Dr. Boris Escalante

Dra. Jimena Olveres

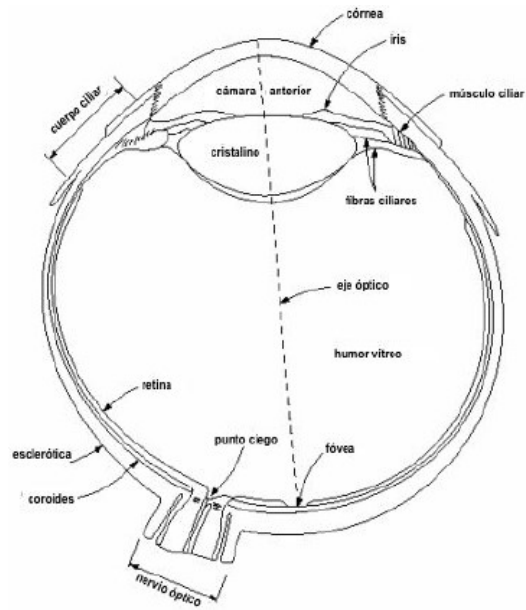
IIMAS



Sistema de vision Humana

- La luz entrante al ojo es absorbida por pigmentos fotosensibles alojados en células en la retina, llamadas bastones y conos, y son los detonantes de una cascada de eventos que crean impulsos eléctricos que viajan a través del nervio óptico al cerebro produciendo la visión.

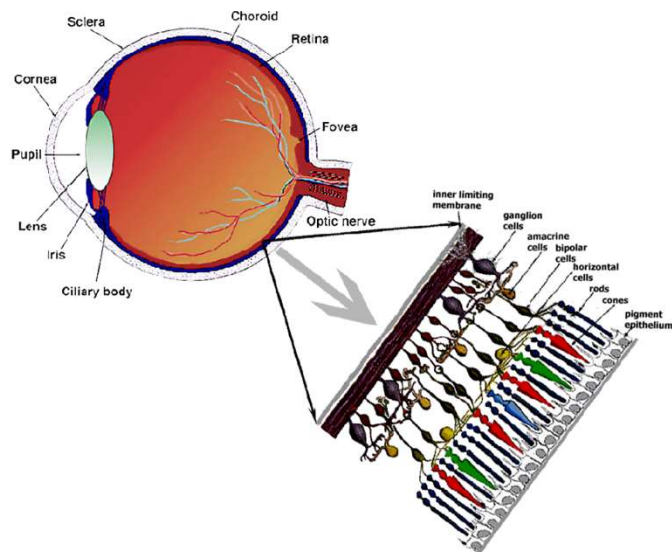
Estructura del ojo




- El ojo humano es casi una esfera, de aproximadamente 20 mm de diámetro. Tres membranas lo cubren: *Córnea, Esclerótica y Retina.*
- *Córnea:* Junto con la esclerótica forma la membrana exterior del ojo. Es un tejido transparente que cubre la superficie frontal del ojo.
- *Esclerótica:* Es una membrana opaca que cubre el resto del globo ocular.
- *Coroides:* Contiene una red de vasos sanguíneos que son la principal fuente de nutrición del ojo. La coroides está fuertemente pigmentada para reducir la cantidad de luz exterior que entra al ojo y la reflexión interna dentro de él. En su parte frontal, se divide en el cuerpo ciliar y el iris.




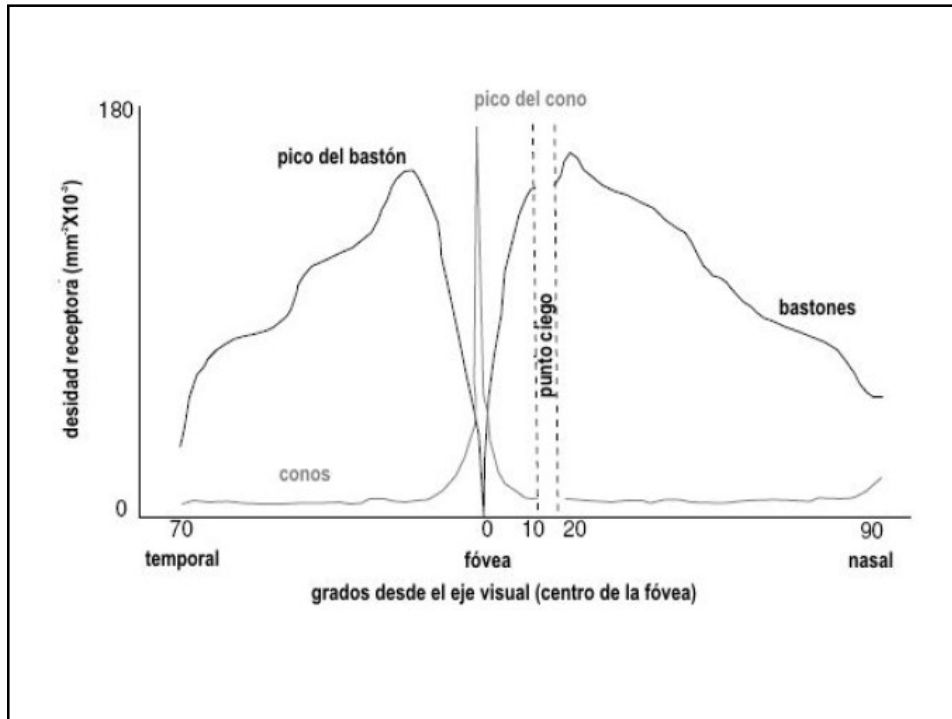
- *Cuerpo ciliar: A él están unidos los músculos ciliares, que controlan el cristalino.*
- *Diafragma Iris: Este se expande o contrae para controlar la cantidad de luz que entra en el ojo. La apertura central del iris, llamada pupila, varía su diámetro de 2 a 8 mm. El frente del iris contiene el pigmento visible del ojo, y la parte trasera contiene un pigmento negro.*
- *Cristalino: Está suspendido por fibras que lo atan al cuerpo ciliar. Tiene una pigmentación amarillosa que puede aumentar con la edad. El cristalino absorbe aproximadamente 8% del espectro visible, especialmente en longitudes de onda cortas. La luz ultravioleta e infrarroja son absorbidas apreciablemente. El cristalino es "la lente" del ojo y sirve para enfocar, ayudado por los músculos ciliares.*




A schematic section through the human eye with an enlargement of the retina; the gray arrow indicates the light direction. This picture was modified with permission of Prof. Dr. Helga Kolb [Nava, Rodrigo. Over complete Image representations for Texture Analysis. Thesis UNAM. 2013]

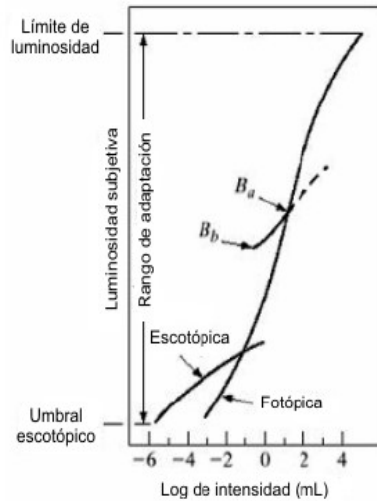
- 
- *Retina: Es la membrana interior posterior del ojo. Cuando el ojo está enfocado, la imagen de un objeto exterior es proyectada en la retina. Sobre su superficie se encuentran los receptores que permiten la percepción visual: los conos y los bastones.*
 - Los conos son entre 6 y 7 millones, se localizan principalmente en la parte central de la retina, llamada fovea, y son muy sensibles al color. Los músculos mueven al ojo hasta que la región de interés se localiza en la fovea. Los humanos somos capaces de percibir detalles muy finos gracias a que cada uno de los conos está conectado a una terminal nerviosa. La visión que utiliza los conos es llamada fotópica, o de luz brillante.

- 
- Los bastones son muchos más, entre 75 y 150 millones y están distribuidos sobre la superficie retinal. Esta área de distribución tan grande y el hecho de que varios están conectados a una misma terminal nerviosa reduce su cantidad de detalle discernible. Sirven para dar una imagen general del campo de visión. No participan en la visión del color y son sensitivos a bajos niveles de iluminación (visión escotópica, o de luz tenue).



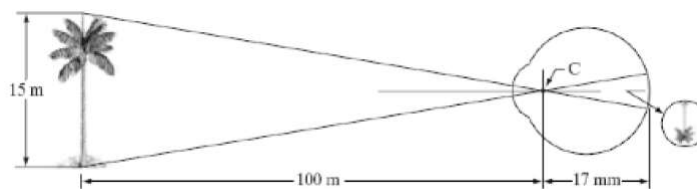
- 
- Existe una región del ojo en la que no hay receptores, ya que es la región donde surge el nervio óptico. Se denomina punto ciego. Fuera de este punto, la distribución de receptores es radial con respecto a la fóvea.
 - *La fóvea: Región circular de la retina de aproximadamente 1.5 mm de diámetro. Con ciertas libertades de interpretación, podemos pretender que la fóvea es un arreglo cuadrado de sensores de 1.5 X 1.5 mm. La densidad de conos ahí es de aproximadamente 150,000 elementos por mm^2 . Basado en esas aproximaciones, el número de conos en la región de más agudeza en el ojo es de cerca de 337,000 elementos*

Rango de sensaciones de iluminación subjetiva respecto a intensidad.



- El ojo no opera en todo este rango de manera simultánea.
 - Se logra esta gran variación con pequeños cambios de sensibilidad. Adaptación a la iluminación.
- LINEA PUNTEADA**

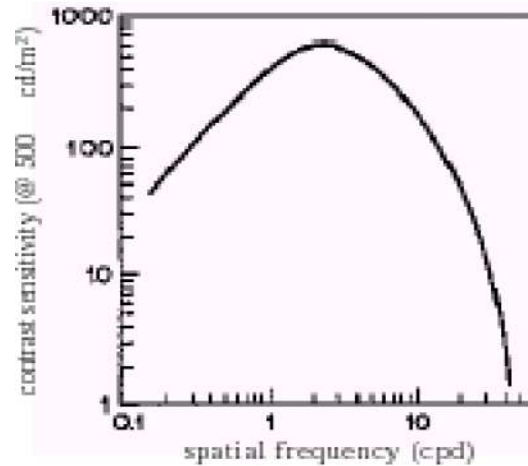
Formación de la imagen en el ojo



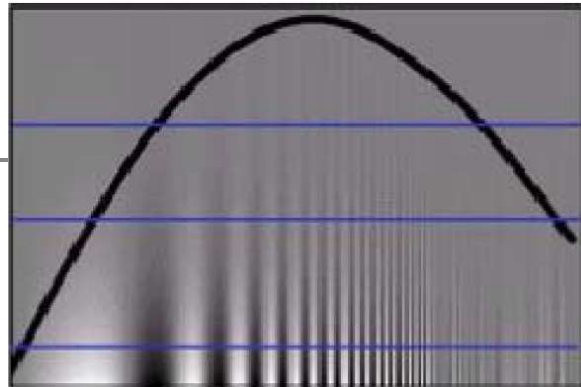
Representación gráfica de la mirada del ojo en una palmera. El punto C es el centro óptico de las lentes.

Función de sensibilidad de contraste

Se basa en la capacidad de resolver o diferenciar líneas a una determinada frecuencia espacial.



MTF

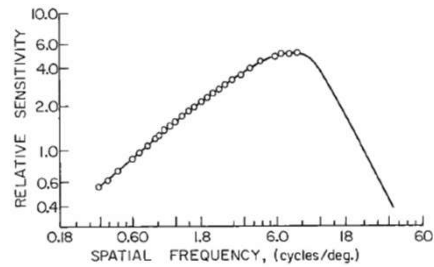


- La MTF puede interpretarse como la capacidad de un sistema óptico para percibir contraste, es decir, la capacidad de resolver o diferenciar líneas a una determinada frecuencia espacial. Para entender de manera simplificada el concepto de frecuencia espacial, hay que notar que las frecuencias espaciales bajas equivalen a repeticiones de patrones muy separados por el espacio (por ejemplo, líneas muy separadas) y que las frecuencias espaciales altas equivalen a repeticiones más compactas (por ejemplo, líneas muy con poca separación). Hay que considerar también que todo sistema de visión tiene un límite superior a partir del cual ya no puede distinguir más detalles. Este límite está directamente relacionado con la frecuencia de Nyquist.
- La MTF (Modulation Transfer Function) o función de transferencia de modulación se ha convertido en una herramienta muy utilizada para especificar el rendimiento y la resolución de toda clase de sistemas ópticos y de visión que van desde lentes, cintas magnéticas y películas hasta telescopios, la atmósfera y el ojo humano. La MTF se usada para caracterizar sistemas lineales e invariantes en el espacio, sin embargo, a pesar de que el Sistema Visual Humano (SVH) no cumple con estas dos propiedades, la MTF se ha utilizado para caracterizarlo bajo condiciones de iluminación controladas.

MTF

- La forma que tendrá la MTF estará determinada por la función de entrada, en este caso es una señal senoidal, en donde las frecuencias espaciales aumentan en una dirección de las coordenadas y el contraste aumenta en la otra dirección.

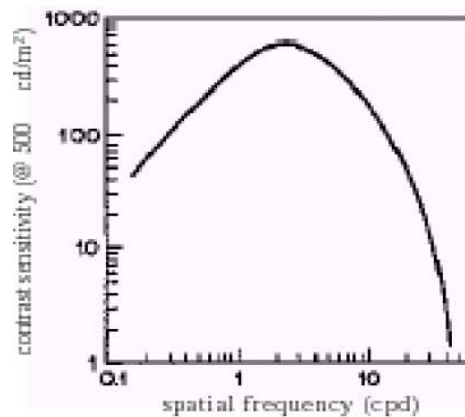
- Se ha demostrado que la respuesta no lineal del ojo humano a las variaciones de intensidad es de forma logarítmica y ocurre al inicio del sistema de procesamiento de información visual, es decir, antes de que ocurra la interacción espacial entre las señales visuales.



- **MTF con una señal senoidal con comportamiento exponencial**
- Este modelo lineal/logarítmico del sistema visual humano permite predecir de manera bastante cercana la respuesta visual en un amplio rango de intensidades.

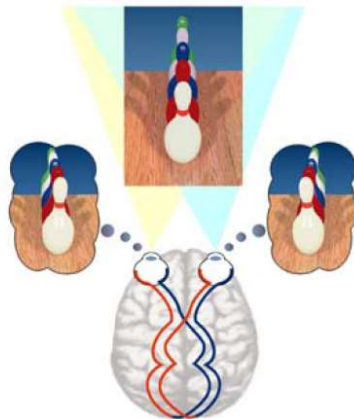
Función de sensibilidad de contraste

- En los estudios de sensibilidad del SVH a base de experimentación psicofísica, se ha encontrado que el MTF humano consiste en una función del tipo paso-banda con un pico o frecuencia espacial máxima en el rango de los 2 a 6 ciclos/grado



Visión estéreo

<http://www.vision3d.com/stereo.html>



Formación de la imagen

- Fuente de Luz
- Cámara o sensor (parámetros intrínsecos y extrínsecos)
- Escena (Reflejo y forma de la superficie)

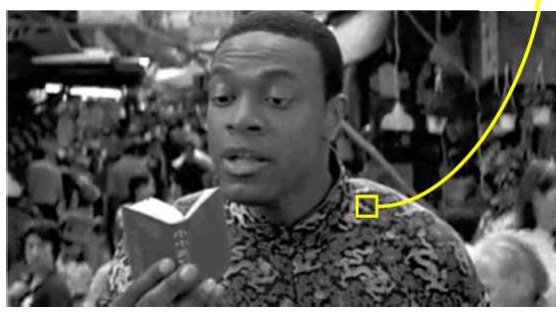
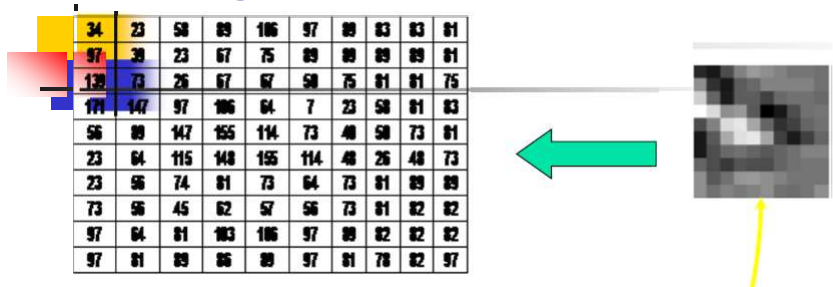
Imagen

- 2-D array of numbers (intensity values,
 - gray levels)
 - Gray levels 0 (black) to 255 (white)
 - Color image is 3 2-D arrays of numbers
 - Red
 - Green
 - Blue

Resolution (number of rows and columns)

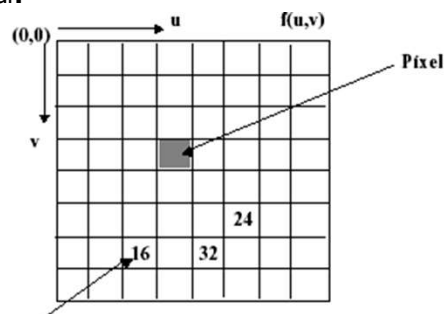
- 128X128
- 256X256
- 512X512
- 640X480

Imagen



Caracterización matemática de las Imágenes

- Una imagen puede ser definida como una función de dos dimensiones $f(x,y)$ donde x y y son las coordenadas espaciales (plano) y la amplitud de la función f en algún par de coordenadas (x,y) es llamada intensidad o nivel de gris de la imagen en ese punto. Cuando x , y y los valores de la amplitud de la función f son cantidades discretas finitas, a dicha imagen se le llama imagen digital.



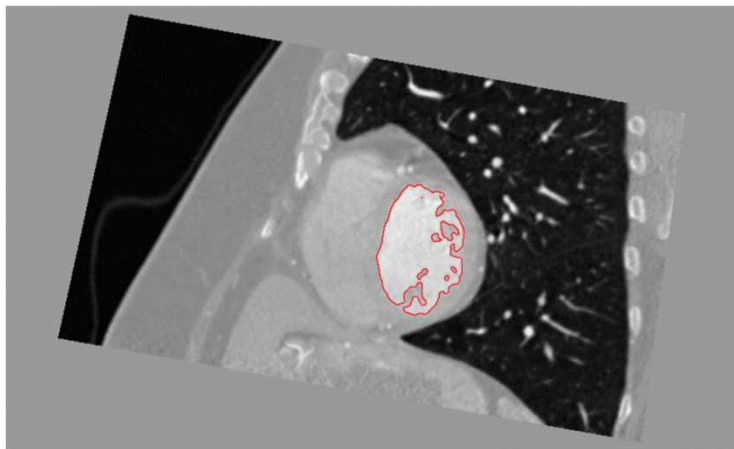
Formatos de la Imagen

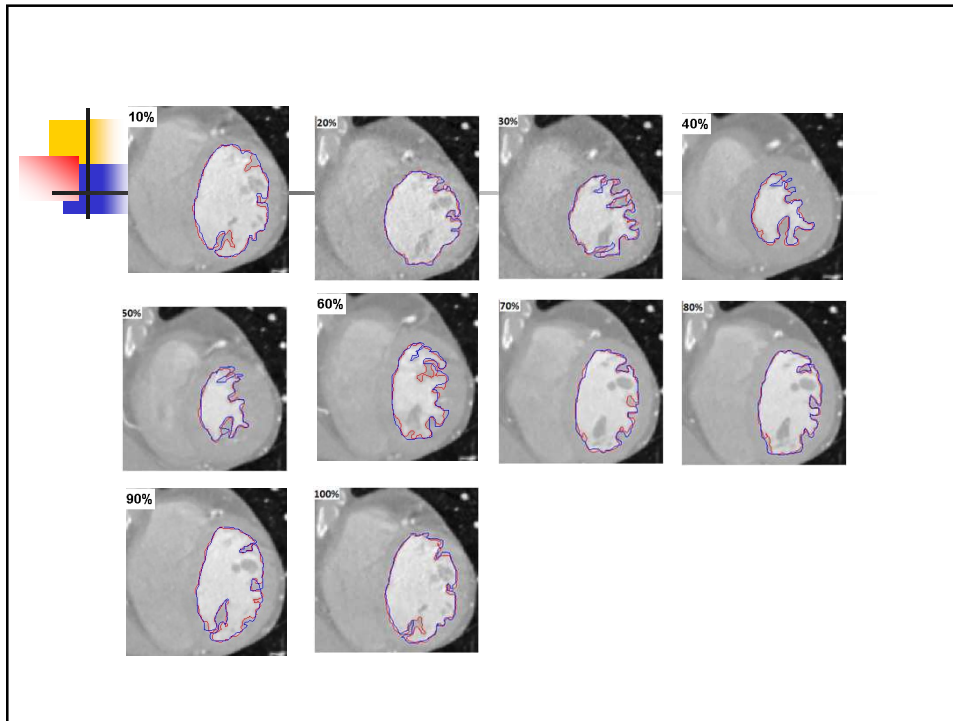
- TIFF
- PNG
- BMP
- GIF
- JPEG



Video

- Secuencia de frames
- 30 frames por segundo
- Formatos:
 - AVI
 - MPEG
 - QUICK TIME





¿Cómo representan las imágenes las computadoras?

- Una imagen puede ser representada por una función: $f(u,v)$.
- El argumento de $f(u,v)$ representa la localización de cada pixel en el plano imagen.
- El valor de $f(u,v)$ puede tener diferentes interpretaciones en diferentes tipos de imágenes. **Ejemplos:**

Imágenes de intensidad:

$f(u,v)$ = intensidad de la escena

Imágenes de distancia:

$f(u,v)$ = distancia desde la escena al sistema de captación.

Imágenes en color:

$f(u,v) = \{f_r(u,v), f_g(u,v), f_b(u,v)\}$

Video:

$f(u,v,t)$ = secuencia temporal de imágenes

