

Recuperación de Información Multimedia



Índice Rec. Inf. Mult.

Recuperación de Imágenes

Color: Modelado y Recuperación

Forma: Modelado y Recuperación

Textura: Modelado y Recuperación

Relaciones Espaciales: Modelado y Recuperación

Otras Dimensiones en la Información Multimedia



Recuperación de Imágenes



Información Visual

El avance de la tecnología multimedia en los últimos años (desde finales del siglo XX) ha favorecido el aumento exponencial de la información visual

El **valor real** de este tipo de información depende directamente de la **facilidad con la cual se pueda recuperar** aquella información que se necesite en cada momento



S.R.I.B.C.

S: Sistemas

R: de Recuperación

I: de Imagen

B: Basados

C: en Contenido



Objetivo genérico:

Diseñar sistemas de recuperación eficientes, para bases de datos de imágenes, capaces de **ofrecer información relevante** ante unos **requerimientos de información visual** definidos por el usuario

Dos generaciones de SRIBC

Primera Generación:

Basada directamente en los sistemas clásicos de recuperación de texto

Trabajo previo de anotación manual y/o semiautomática

Segunda Generación:

Basada en la extracción automática de propiedades perceptuales sobre la imagen



Dos generaciones de SRIBC

Primera Generación:

Basada directamente en los sistemas clásicos de recuperación de texto

Trabajo previo de anotación manual y/o semiautomática

Casa
Cesped
Piscina
Hamaca
...

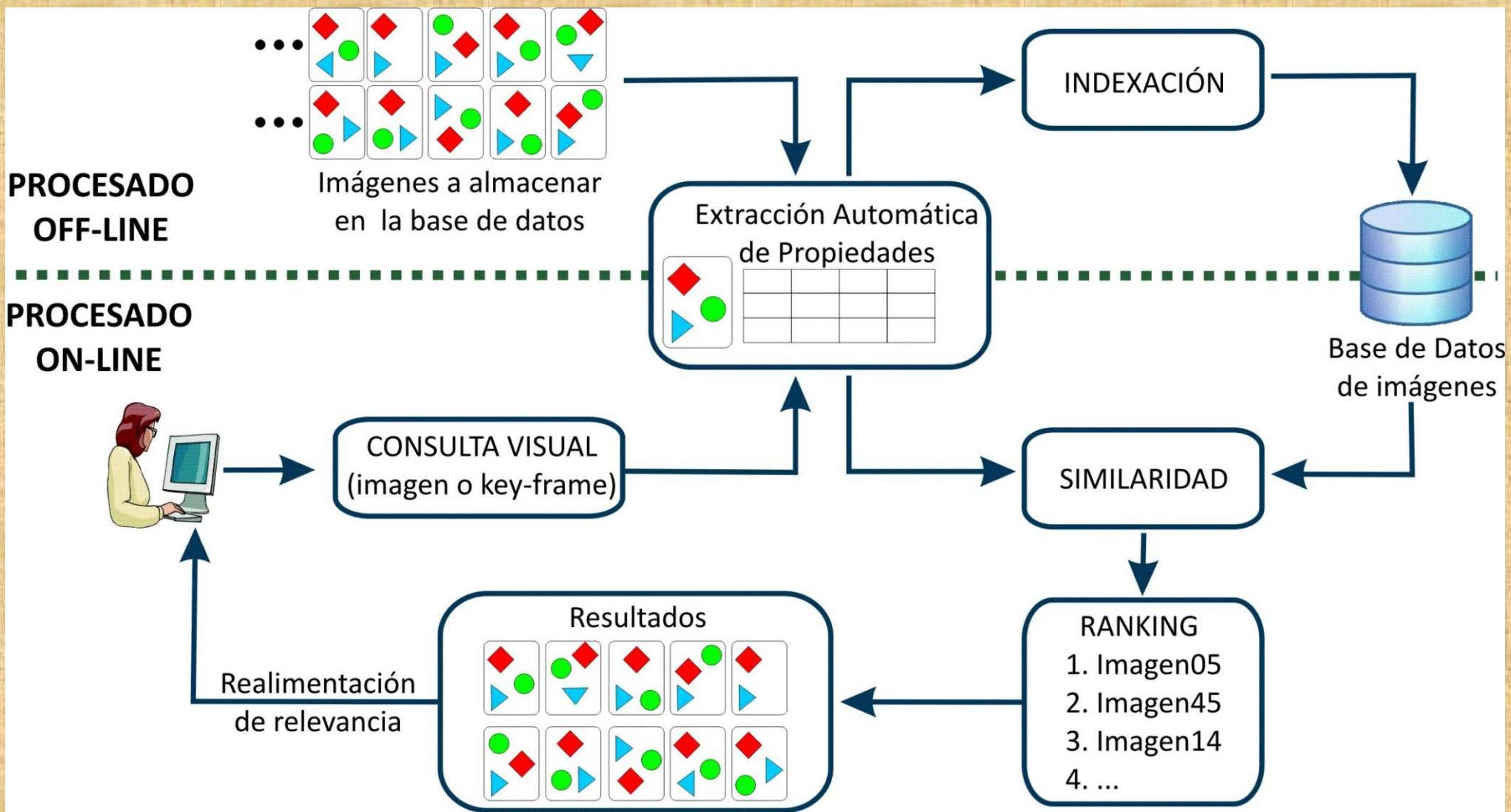


Segunda Generación:

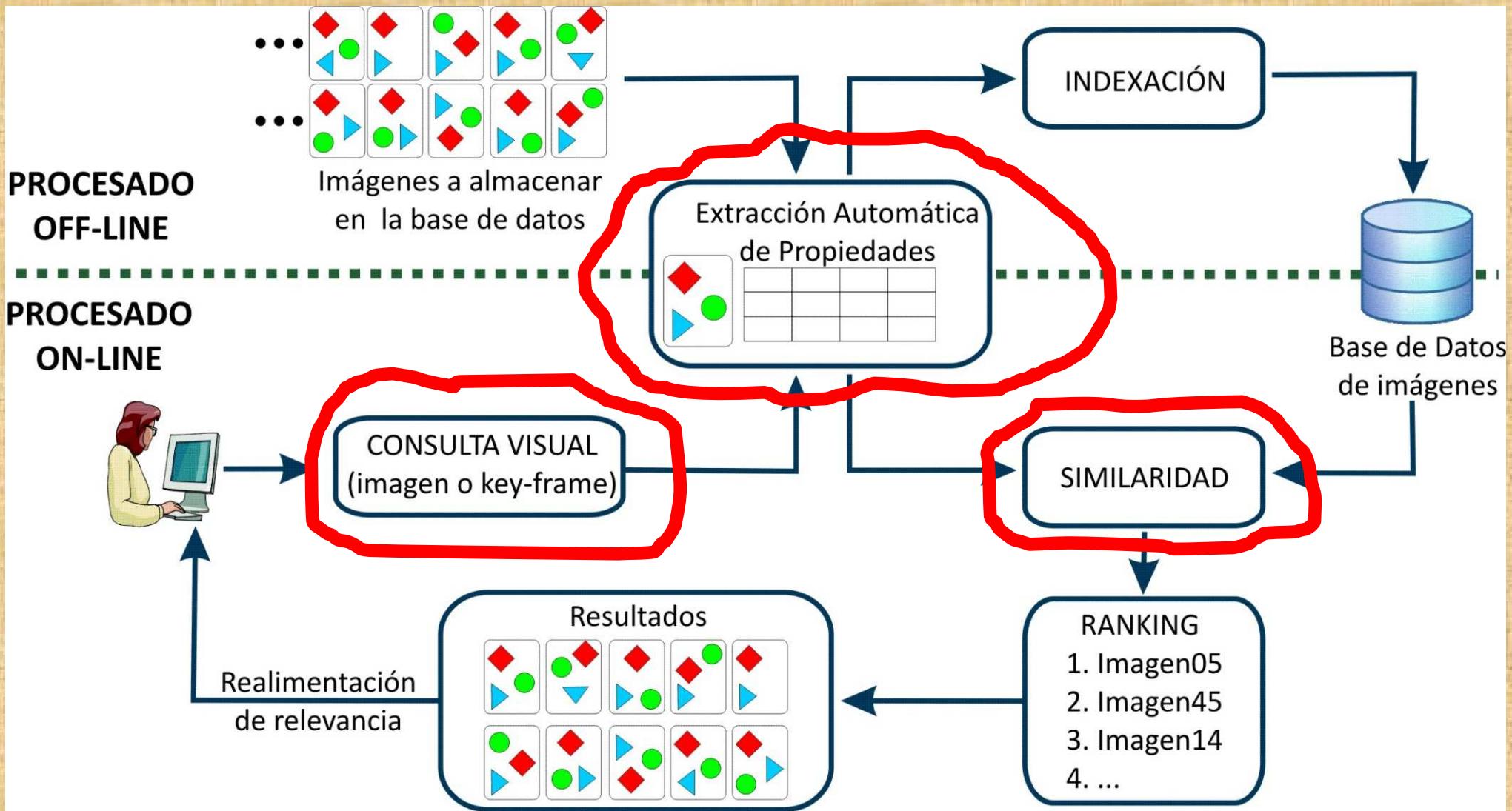
Basada en la extracción automática de propiedades perceptuales sobre la imagen

Colores
Formas
Texturas
...

Arquitectura SRIBC 2^a Generación



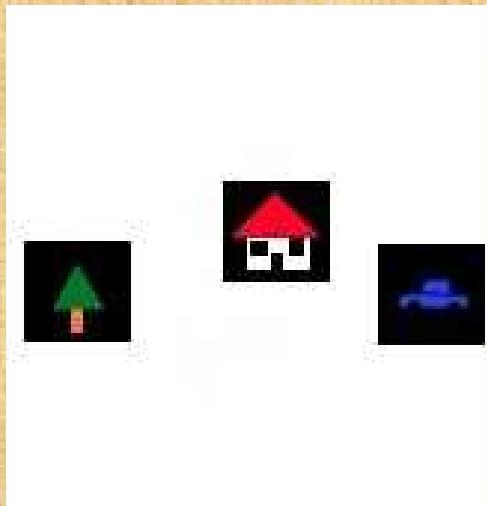
Arquitectura SRIBC 2^a Generación



Consulta Basada en Contenido

En los SRIBC de segunda generación, la **introducción de las necesidades de información** se debe poder realizar de forma perceptual, sin recurrir a descripciones mediante texto

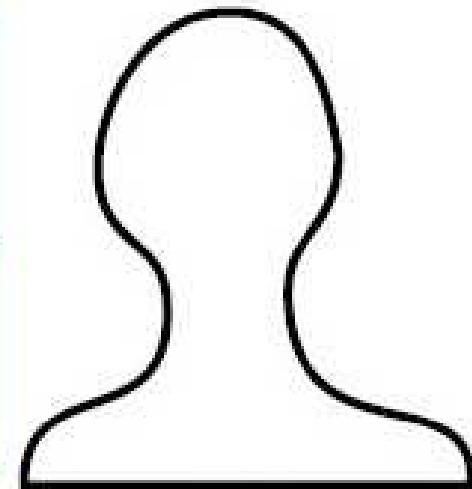
de iconos



de pintado



de boceto



de imagen ejemplo



Representación del Contenido

Cuando un observador ve una imagen, **percibe en su cerebro parámetros perceptuales psicofísicos** como pueden ser el contraste, los bordes, las formas, el color, la textura, la situación espacial, etc

La **representación** y la **comparabilidad** (similaridad) de las propiedades perceptuales de las imágenes es el problema fundamental de la recuperación de imágenes



Dominios de Aplicación

- ✓ Procesado y Análisis de Imágenes
- ✓ Reconocimiento de Patrones
- ✓ Visión por Ordenador
- ✓ Modelado de Datos Multidimensionales
- ✓ Indexación y Visualización de Datos

Periodismo



Medicina



Arte



Logotipos



Recuperación de Imágenes

El Gap Semántico

En los dominios imagen amplios (bases de datos de imágenes formadas por un número grande de elementos) la variabilidad es ilimitada e impredecible, incluso para un mismo significado semántico

El gap entre los descriptores de propiedades y la interpretación semántica es muy importante



Los Ordenadores Ven

Proporcionar a los ordenadores la capacidad de “**ver**” no es una tarea sencilla

Nosotros vivimos en un mundo **tridimensional**



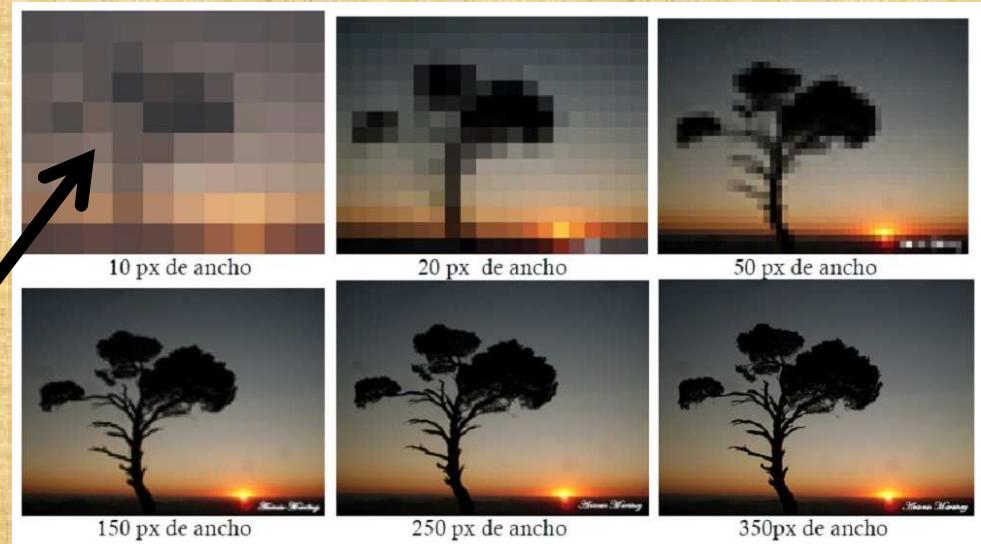
Los sensores visuales disponibles permiten obtener imágenes **bidimensionales**

La **proyección de datos** en un número menor de dimensiones produce una **enorme pérdida de información**

Imagen

Una **imagen** se puede definir como un valor particular de la **distribución de probabilidad conjunta** de un conjunto de **NxM variables aleatorias** ordenadas espacialmente sobre una rejilla de NxM posiciones

Cada variable aleatoria representa el **valor de la propiedad física que forma la imagen** en la posición espacial de dicha variable aleatoria en la rejilla de NxM posiciones

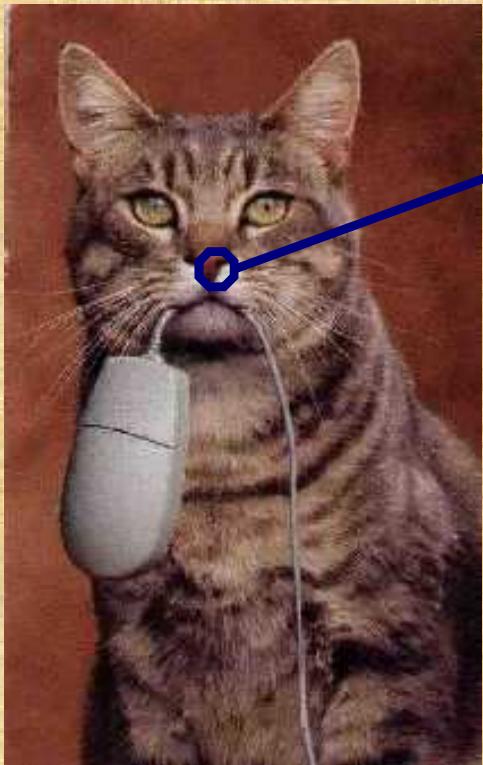


PIXEL (Picture Element)

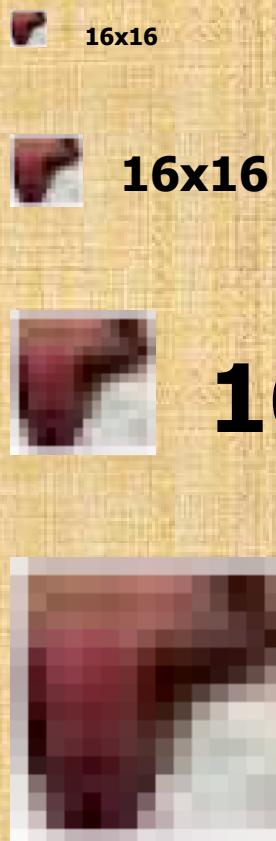


¿Cuánto “Mide” 1 Píxel?

¿Cuáles son las **dimensiones físicas** de un píxel?



211 x 337



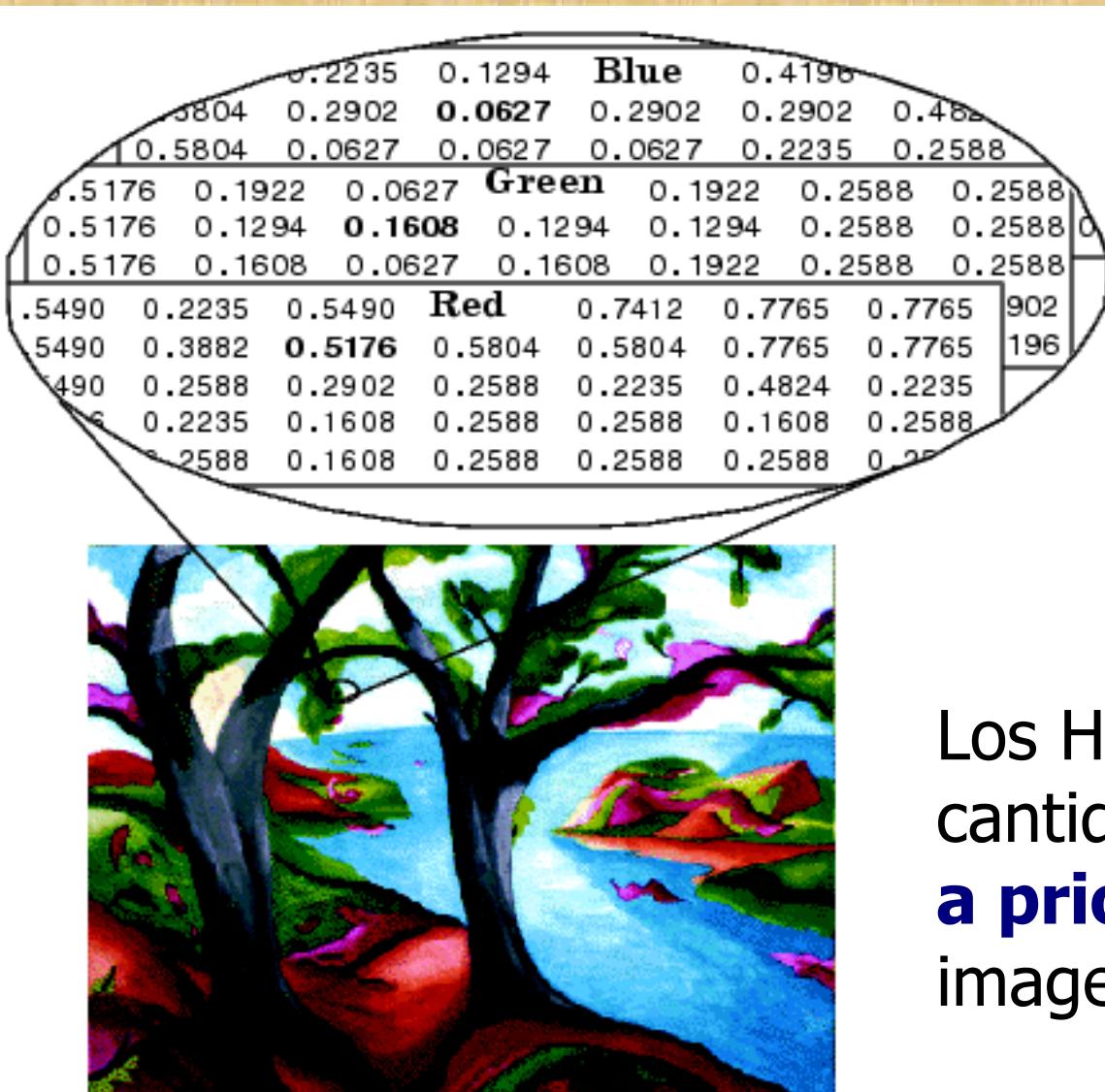
16x16



16x16



Representaciones de Datos



Los Ordenadores sólo necesitan una, o varias, **matrices de datos**

Los Humanos usan gran cantidad de **información a priori** para interpretar la imagen que ve

Modelado del Color

El Color es la propiedad visual más
destacada de una imagen



Imágenes en Color

El **color** es una **propiedad muy importante** en los procesos humanos de visión

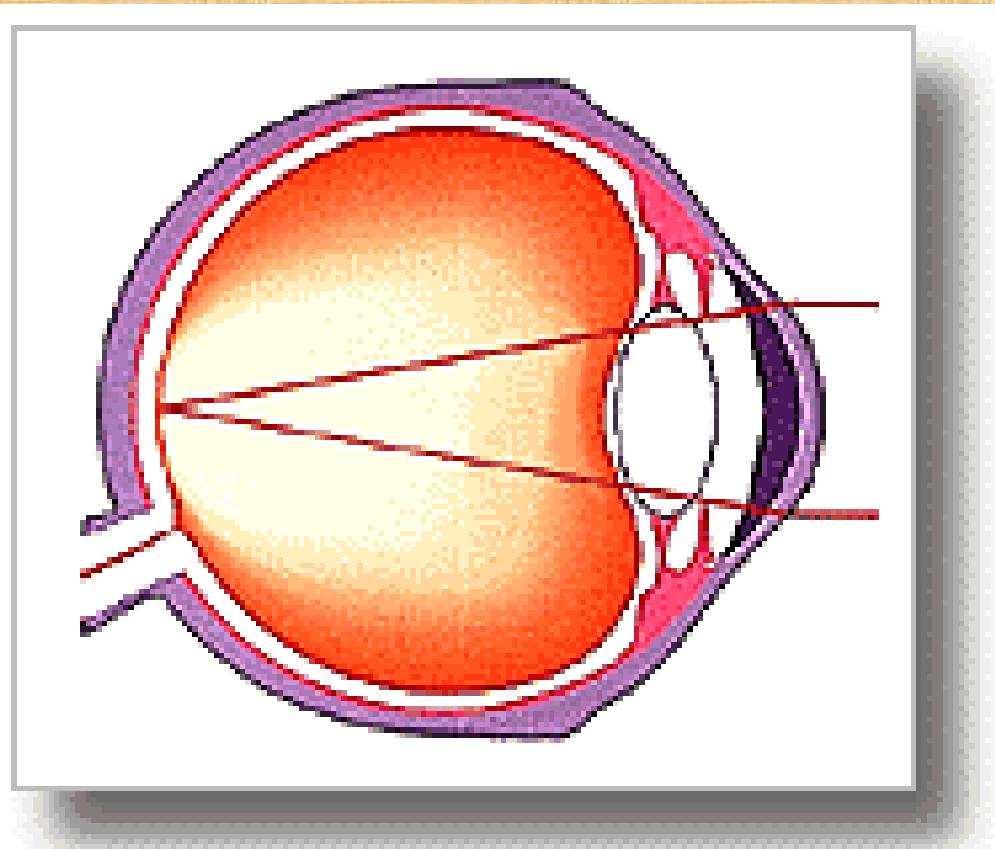
Históricamente no se le ha dedicado demasiada importancia en Procesado de Imagen debido al **alto coste**, durante muchos años, de los equipos para procesar color



Actualmente ya no es un problema

Fisiología

El color es el efecto principal que se percibe cuando una onda electromagnética incide sobre **los conos y bastones** del sistema de visión humano



Conos vs Bastones

CONOS (6-7 millones): Muy sensibles al color, sirven para apreciar detalles finos de la escena

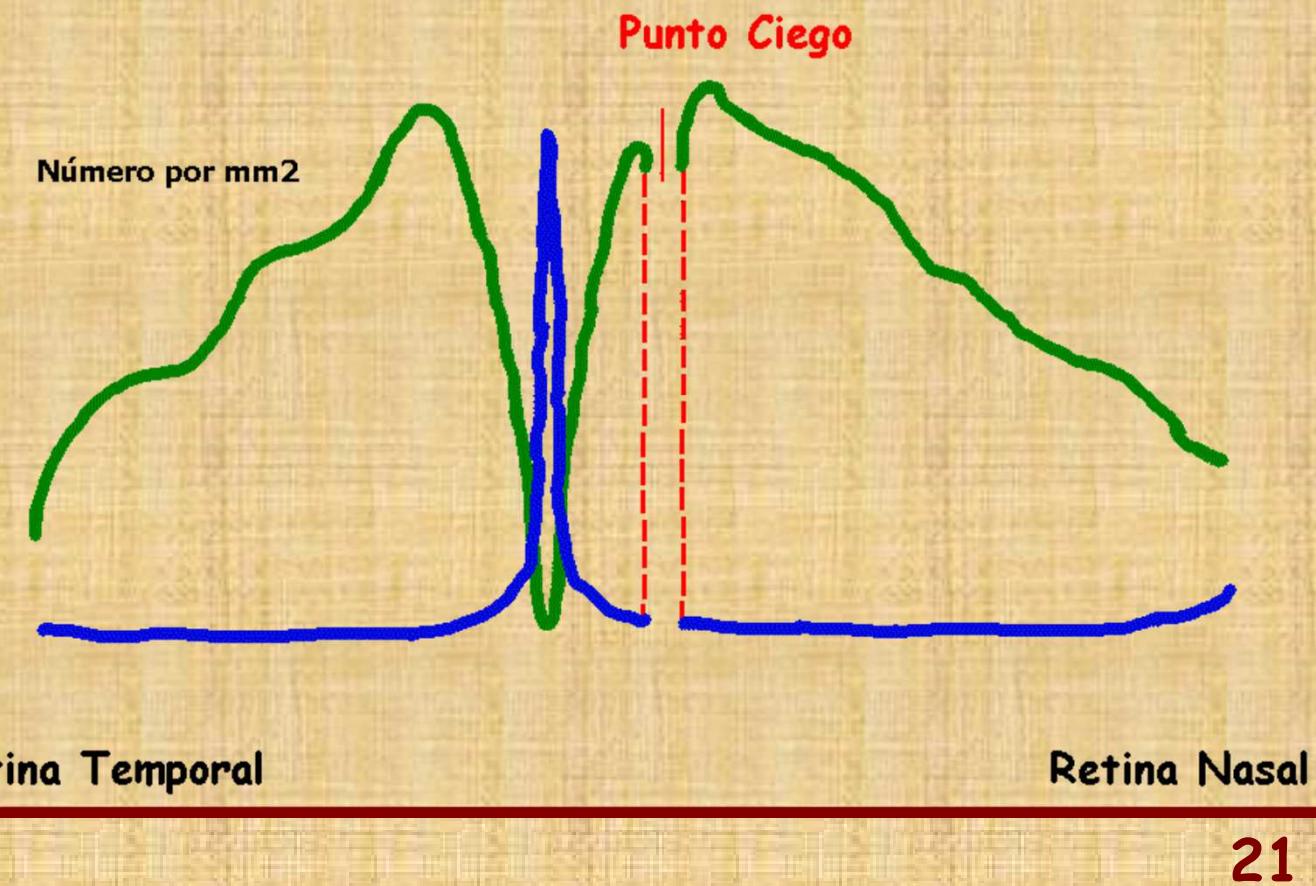
BASTONES (75-150 millones): Sensibles a niveles de iluminación bajos, sirven para una visión general de la escena

Visión Fotópica

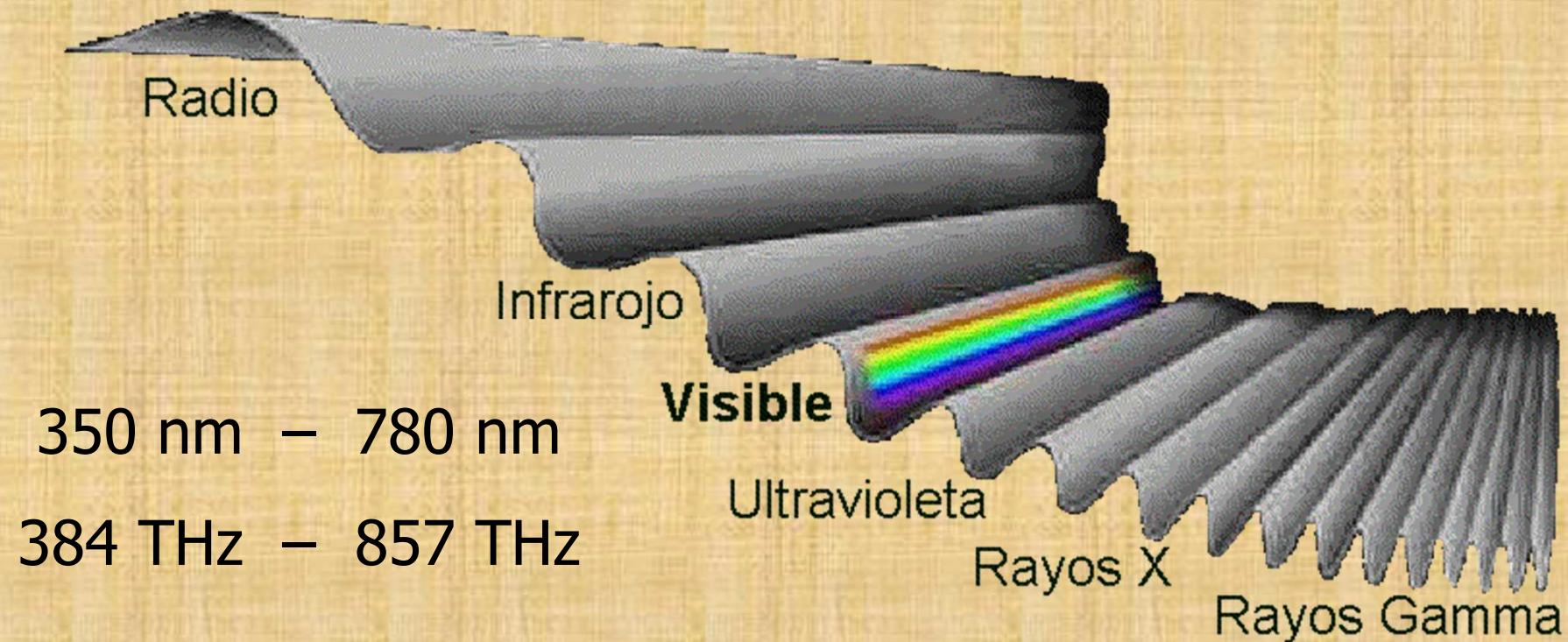
Visión de Luz Brillante

Visión Escotópica

Visión de Tenue



Rango de Visión

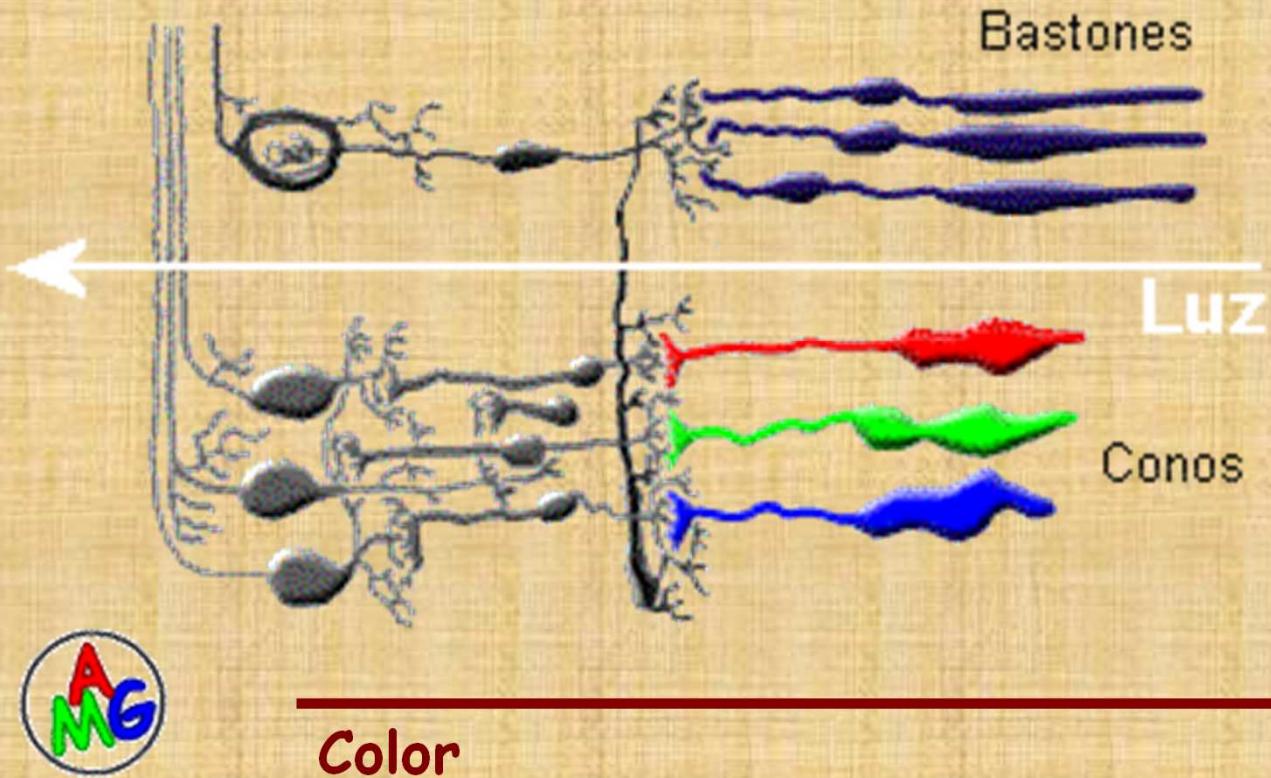


Tipos de Conos

En la retina humana, los conos son los responsables de la percepción del color

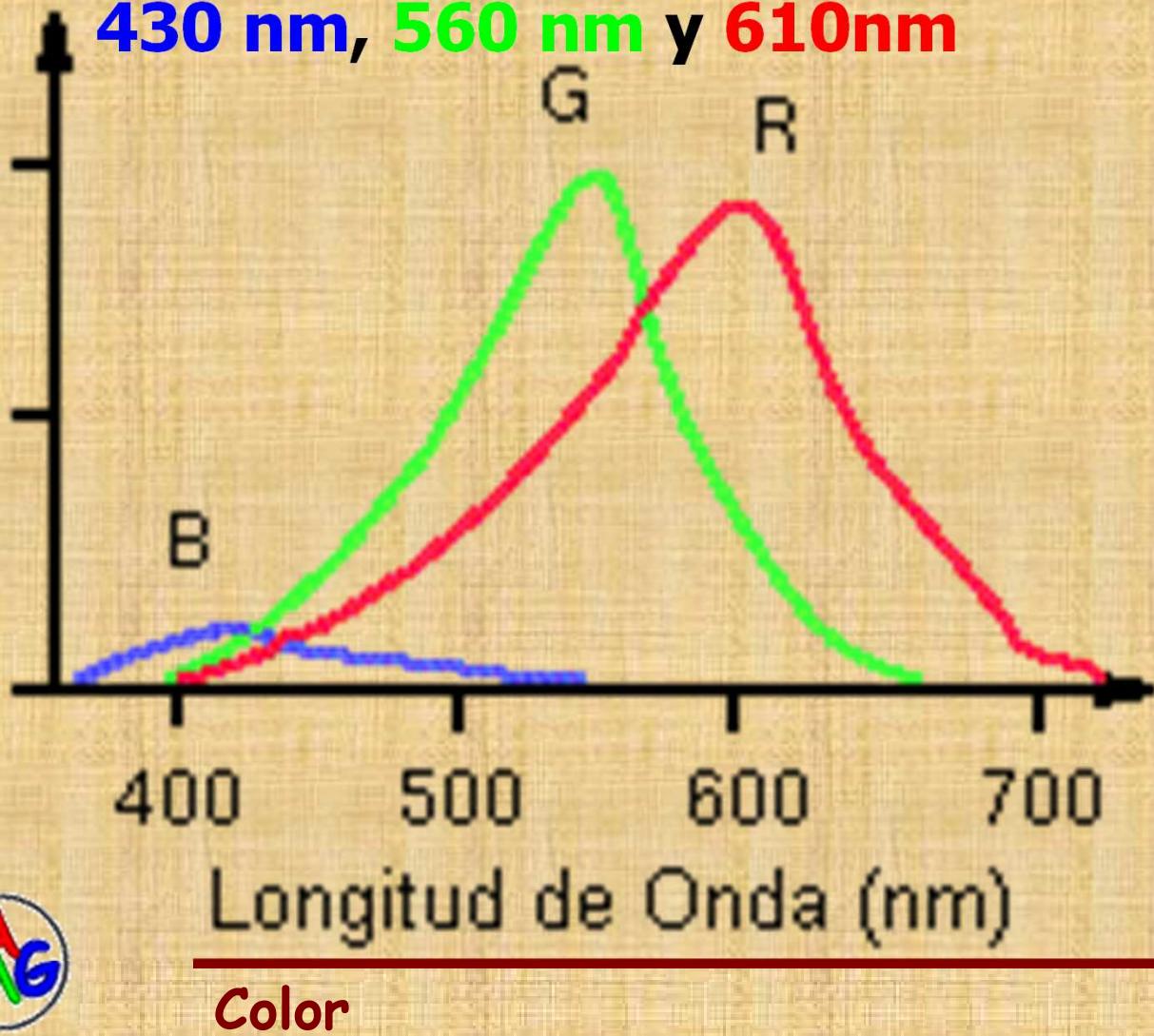
En la retina existen **tres tipos de conos** denominados

B, G y R
(ó S, M y L)



Sensibilidad de los Conos

Los **picos de sensibilidad** de los conos B, G y R se localizan para un observador humano promedio en los **430 nm, 560 nm y 610nm**



Muchas especies de animales tienen más o menos conos

Funcionamiento de los Conos

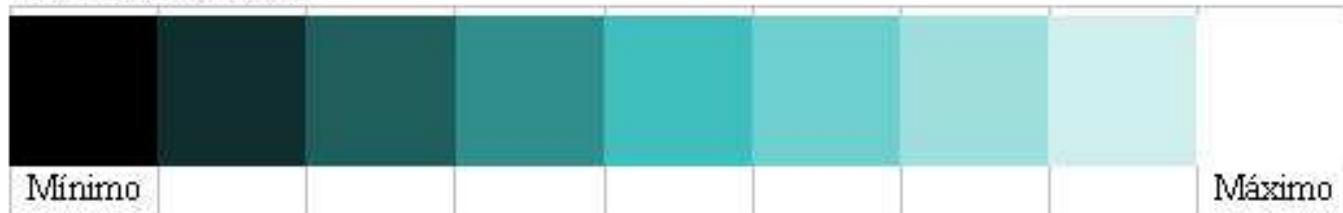
El **Modelo Ocular** de la interpretación del color en función **de la activación de los conos** postula la existencia de **8 colores elementales** en el **Sistema Cromático Humano**



Percepción Cerebral del Color

La percepción de los colores en el cerebro humano se basa en los **descriptores perceptuales de la luz**

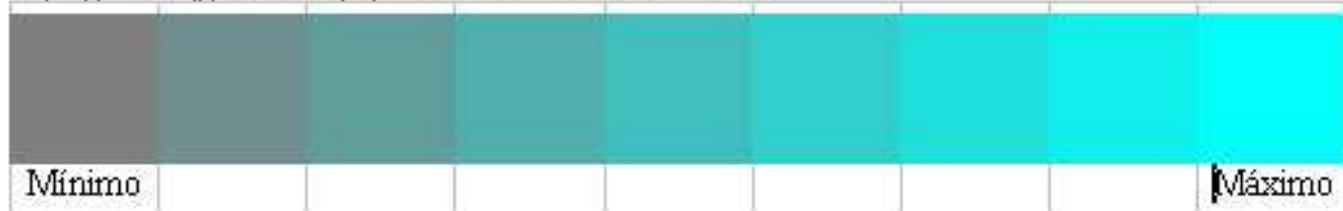
- Brillo. Dadas dos fuentes de luz con la misma distribución espectral, aquella que tenga mayor intensidad física parece perceptualmente más brillante o clara.



- Tono. El tono es el color puro que define la luz. La longitud de onda, o período espacial de la señal, no es una buena medida del color porque algunos colores no están contenidas en el arco iris que produce un prisma al descomponer luz blanca (por ejemplo, morado = rojo + azul).

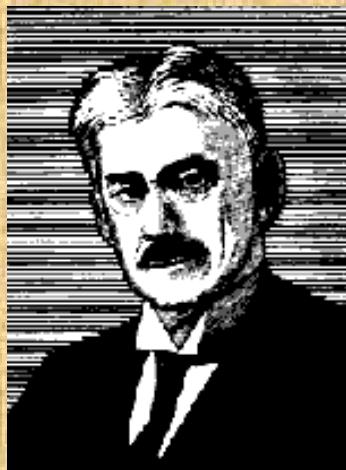


- Saturación. La saturación es la responsable de distinguir entre un tono apagado (grisáceo) y un tono claro bien definido.



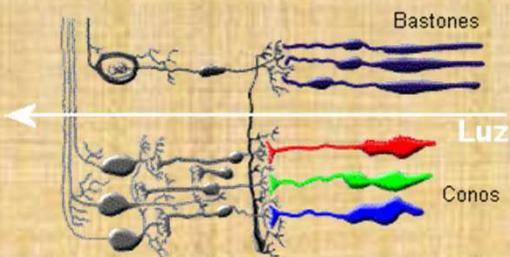
Albert Munsell (1905)

Primeros intentos de **organizar los colores**

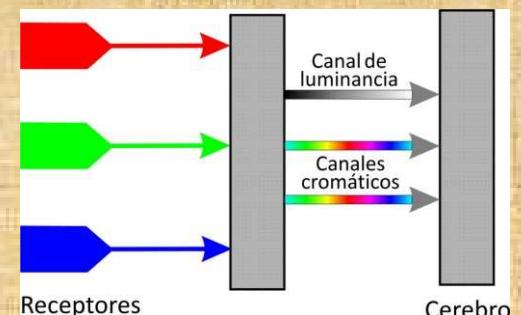


Modelos de Color

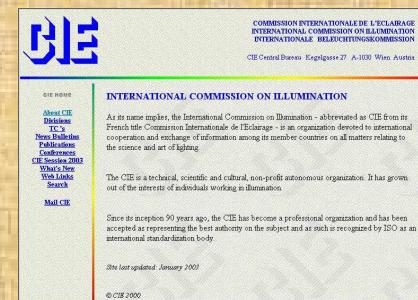
Los modelos de color se introducen para dotar al color de una **representación numérica**



Representan los estímulos de color como puntos de un **espacio tridimensional**



En 1931 la “**Commission Internationale de l’Éclairage**” definió un modelo que permite representar todos los colores mediante síntesis aditiva de tres fuentes primarias



Modelo CIE XYZ

Fuentes primarias recomendadas (Triestímulos):

$$P_x = 600.0 \text{ nm}$$

$$P_y = 546.1 \text{ nm}$$

$$P_z = 435.8 \text{ nm}$$

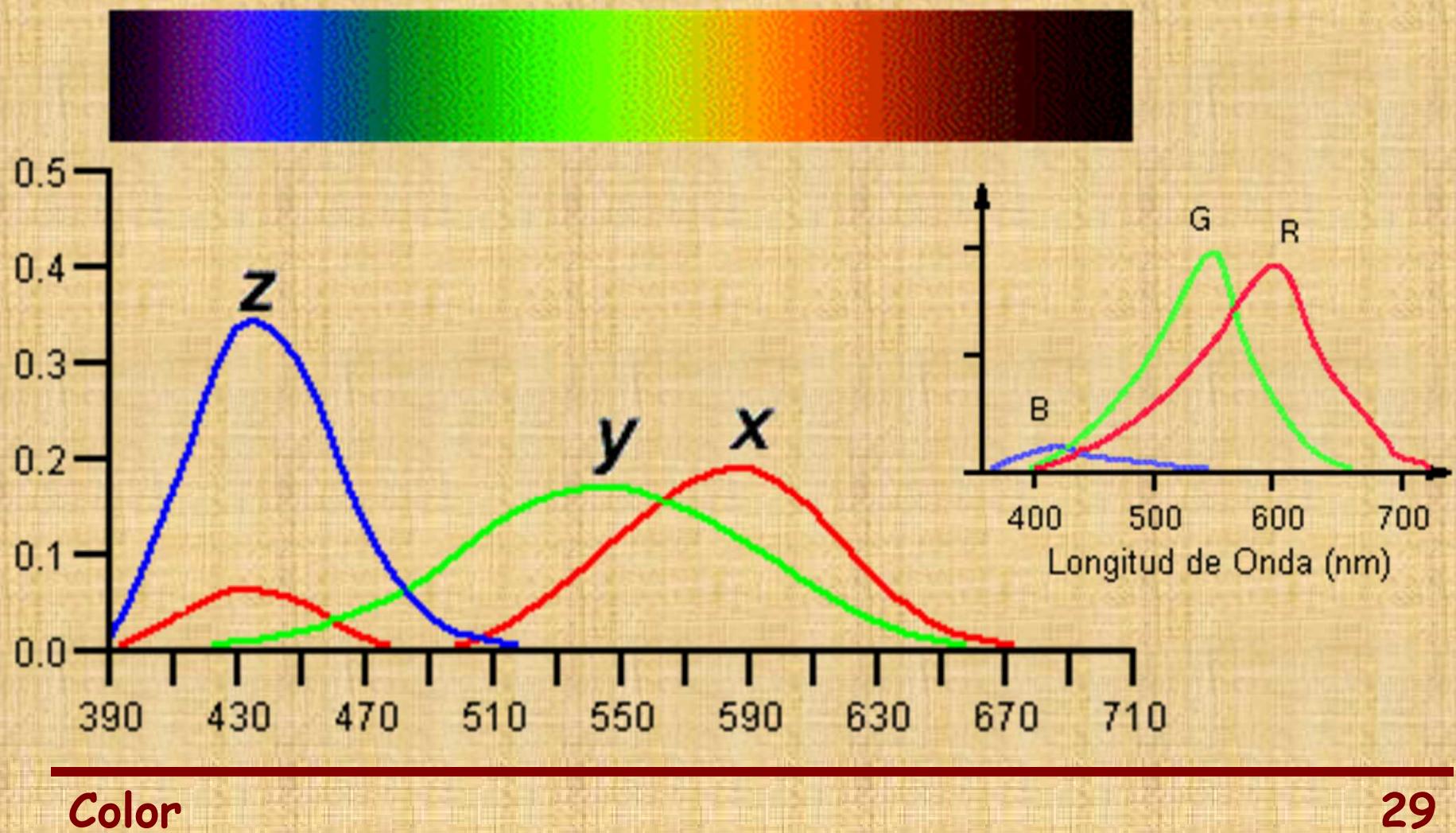
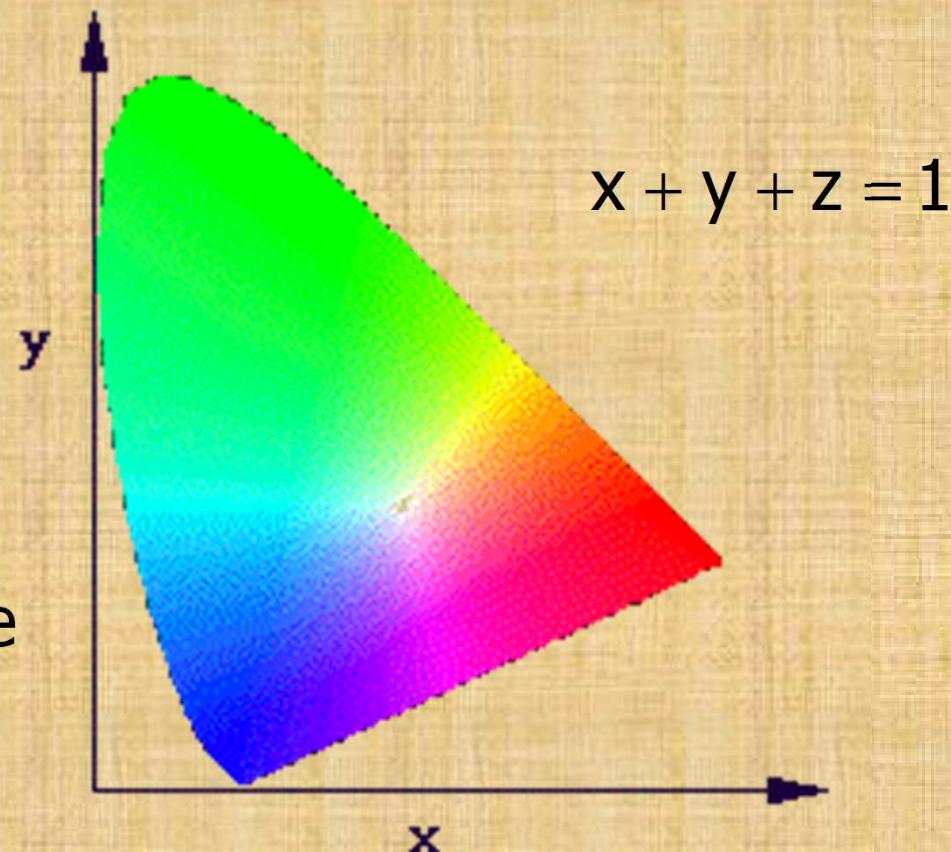


Diagrama de Cromaticidad

Si se normalizan los triestímulos

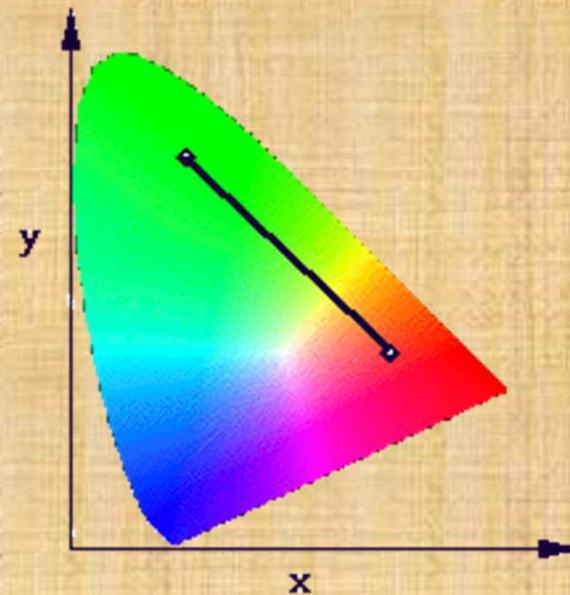
$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z}$$



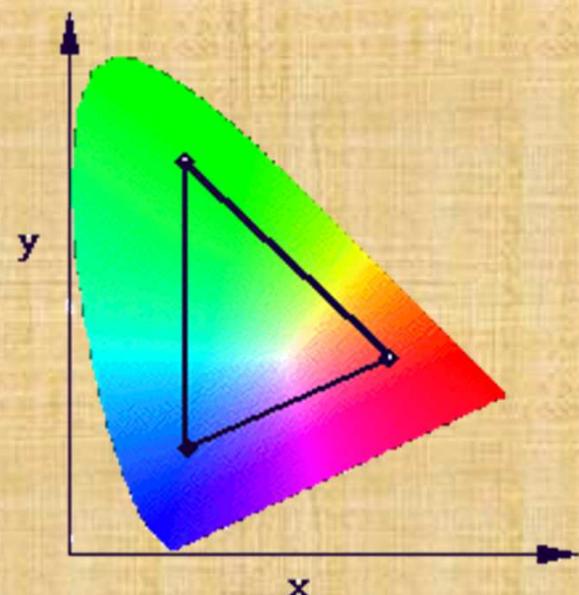
Cada color realizable se puede representar mediante un punto en el plano

Propiedades

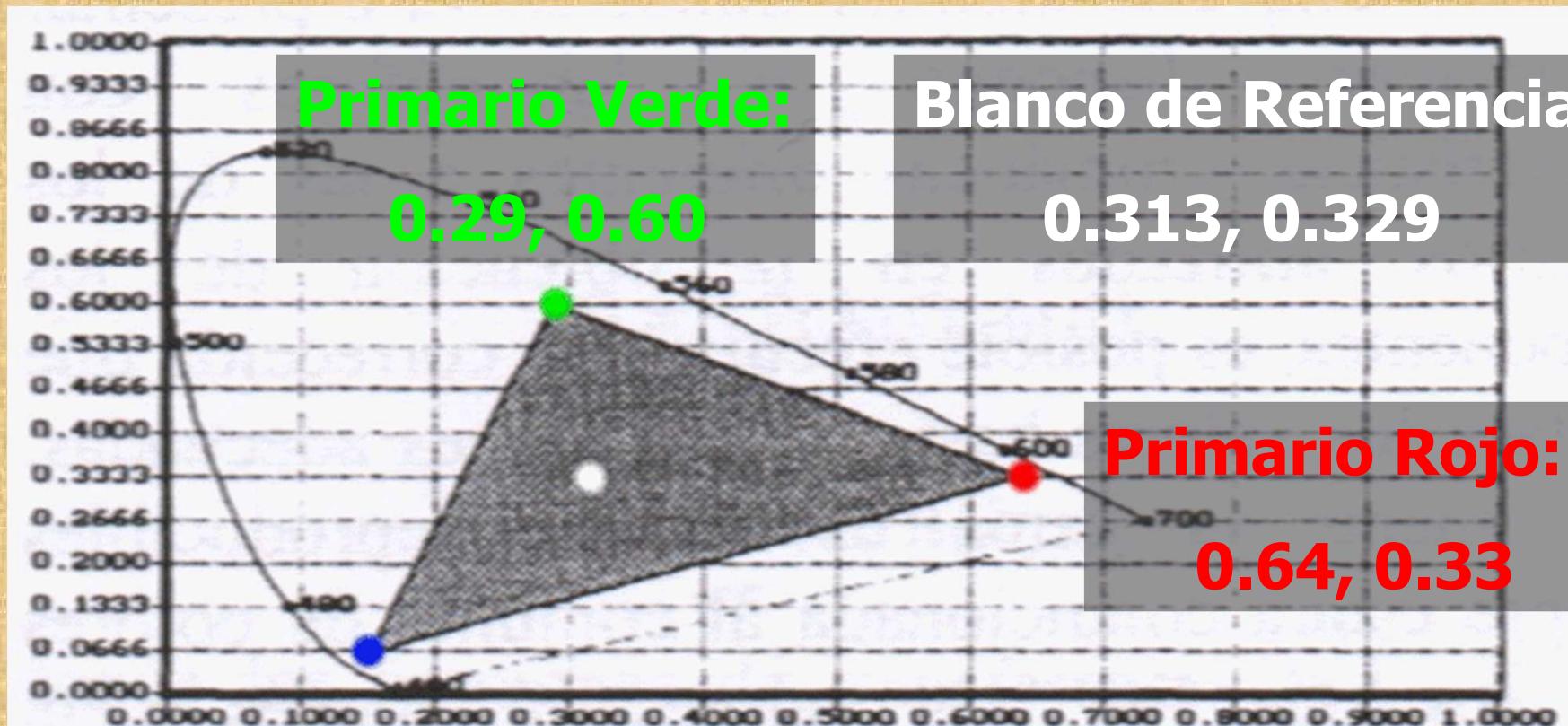
Dados **dos** colores primarios, la gama de colores que pueden generar se encuentran en el **segmento** que une esos colores primarios



Dados **tres** colores primarios, la gama de colores que pueden generar se encuentra en el **triángulo** cuyos vértices son esos colores primarios



Sistema PAL para TV



Primario Azul:

0.15, 0.06

COEFICIENTES X,Y,Z DE LOS PRIMARIOS (R**) (G**) (B**) :

$$\begin{cases} x = 0.639999 \\ y = 0.330000 \\ z = 0.029999 \end{cases}$$

$$\begin{cases} x = 0.289999 \\ y = 0.600000 \\ z = 0.110000 \end{cases}$$

$$\begin{cases} x = 0.150000 \\ y = 0.059999 \\ z = 0.790000 \end{cases}$$

COEFICIENTES X,Y,Z DEL BLANCO (D63) :

$$x = 0.3130$$

$$y = 0.3290$$

$$z = 0.3580$$

Modelos OH

Los **modelos de color orientados al hardware** están basados en la fisiología de la retina humana

Colores Primarios: Se definen como los correspondientes a la respuesta máxima de los conos del ojo a estímulos monocromáticos puros (o combinaciones lineales de estos)

Gama de colores: Se obtienen mediante la suma de los colores primarios (primarios aditivos)

Son los modelos que se emplean en los dispositivos hardware



RGB

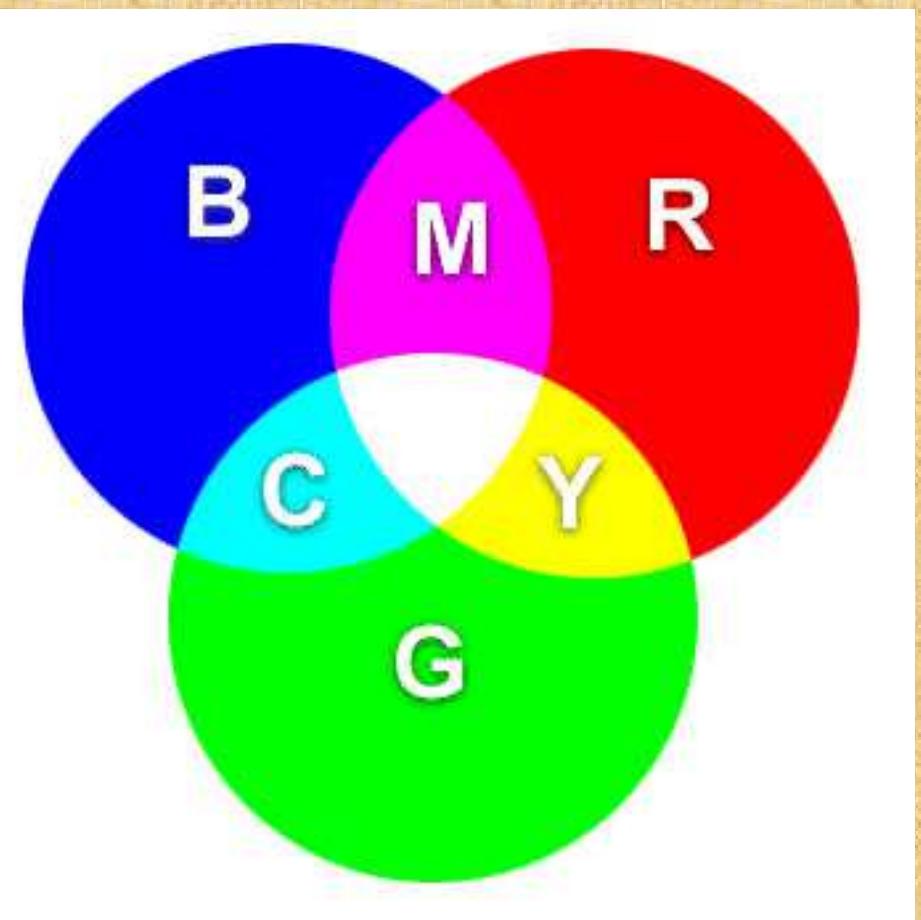
Modelo Aditivo

$$X = 0.490R + 0.310G + 0.200B$$

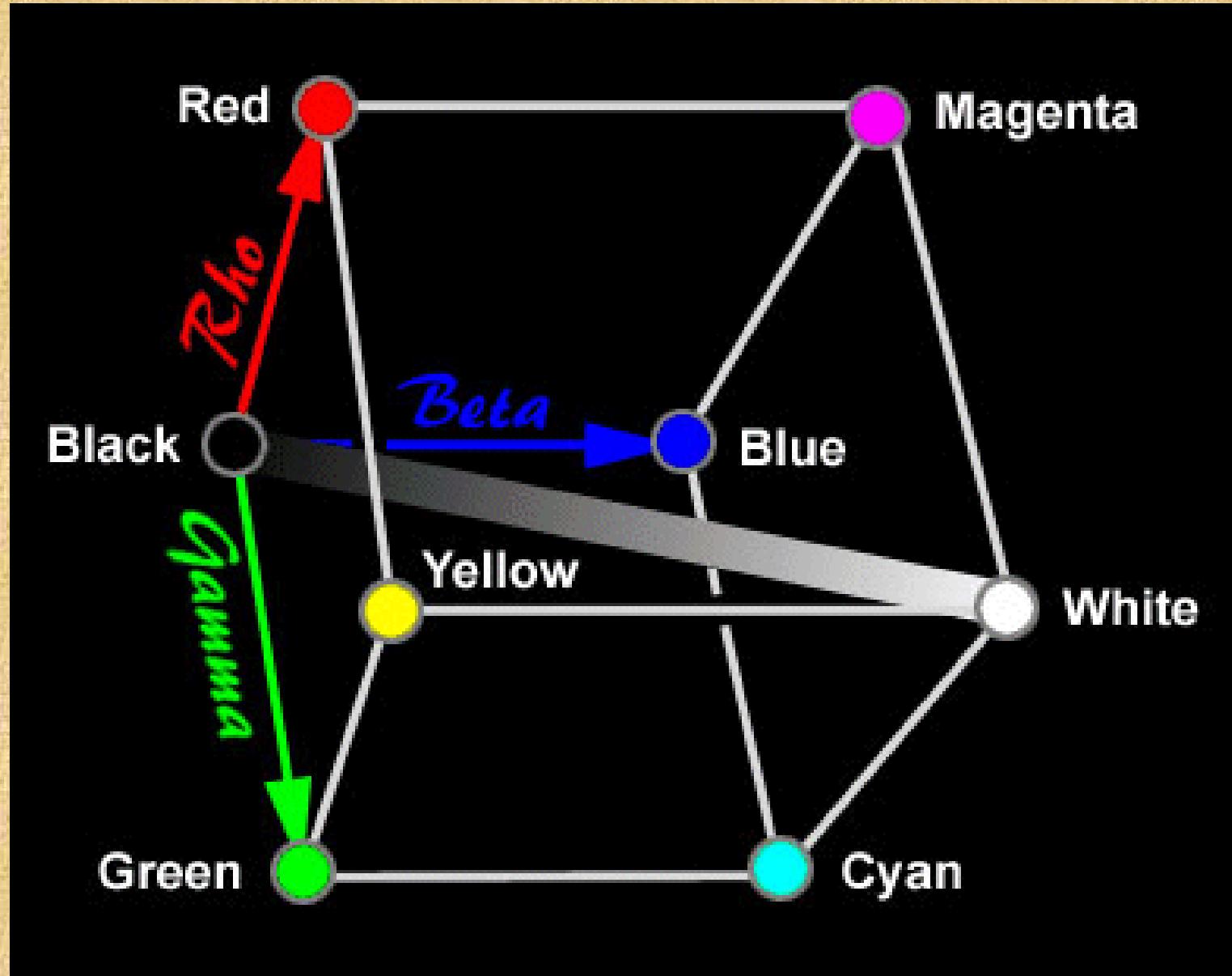
$$Y = 0.177R + 0.813G + 0.010B$$

$$Z = 0.000R + 0.010G + 0.990B$$

Rojo, Verde, Azul



RGB



Descomposición RGB



CMY

Modelo Sustractivio

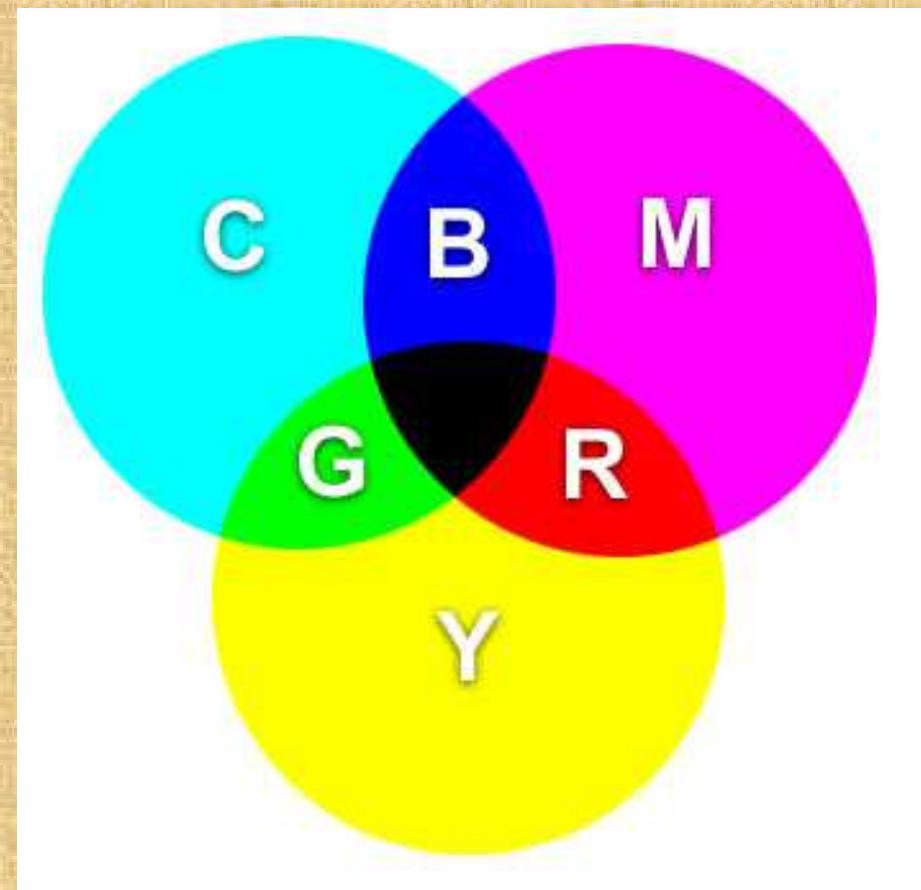
$$C = 1 - R$$

$$M = 1 - G$$

$$Y = 1 - B$$

Cian, Magenta, Amarillo

$$CMYK = CMY + \text{Negro}$$



Descomposición CMY



Estándar de
las emisiones
de Televisión

Luminancia, Fase, Cuadratura

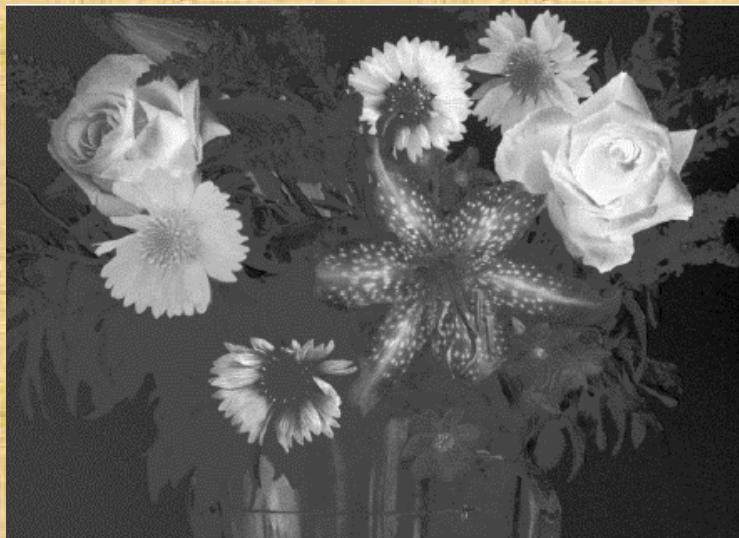
$$Y = 0.299R + 0.587G + 0.114B$$

$$I = 0.596R - 0.275G - 0.321B$$

$$Q = 0.212R - 0.523G + 0.311B$$



Descomposición YIQ



Otros Modelos OH

Se trata de combinaciones lineales del modelo RGB

$$I_1 = \frac{R + G + B}{3}$$

$$I_2 = \frac{R - B}{2}$$

$$I_3 = \frac{2G - R - B}{4}$$

$$I'_1 = \frac{R + G + B}{3}$$

$$I'_2 = R - B$$

$$I'_3 = \frac{2G - R - B}{2}$$



Modelos OU

Los **modelos de color orientados al usuario** están basados en la percepción de los colores

Modelos no uniformes: Separan la luminancia del resto de componentes; sus componentes derivan de combinaciones, lineales o no, del modelo RGB o el modelo CIE XYZ (Ejemplo: HLS, HSV, HSB)

Modelos de color uniformes: La diferencia de color percibida por un observador humano se puede aproximar por la distancia euclídea entre dos puntos en el espacio de color (Ejemplos: MTM, L*u*v*, L*a*b*, L*C*h*)



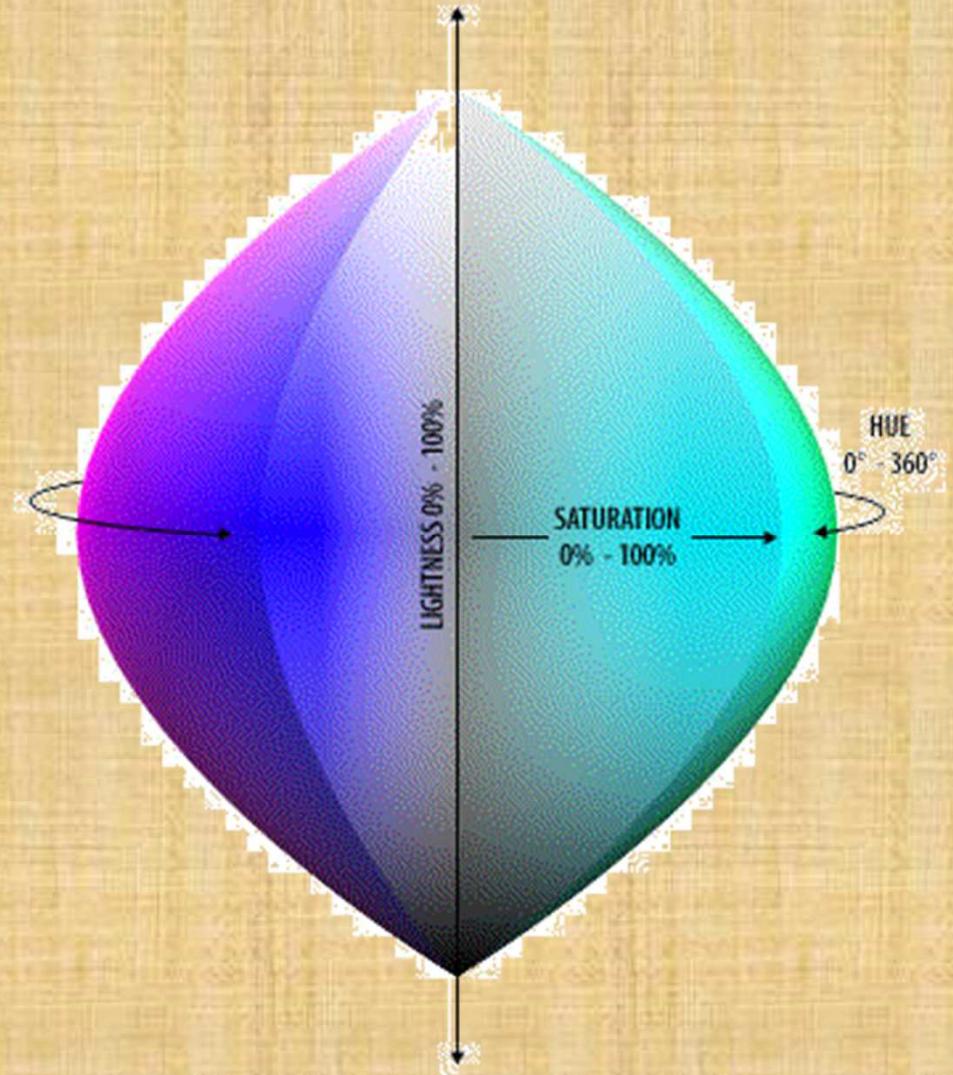
HLS

Tono, Intensidad, Saturación

$$H = \arctan \left[\frac{3^{1/3}(G - B)}{(R - G) + (R - B)} \right]$$

$$L = R + G + B$$

$$S = 1 - \left(\frac{\min(R, G, B)}{R + G + B} \right)$$



HSV

Tono, Saturación, Valor

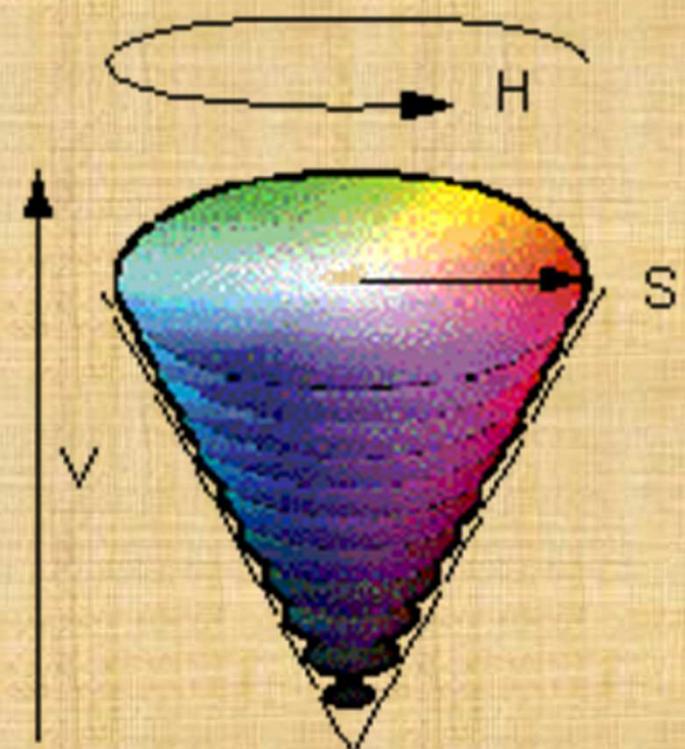
$$H = 180 \frac{0.5(R - G) + (R - B)}{\left((R - G)^2 + (R - B)(G - B)\right)^{1/2}}$$

$H = \text{no definido} \quad \text{si} \quad S = 0$

$$H = 360 - H \quad \text{si} \quad B > G$$

$$S = 1 - \frac{3\min(R, G, B)}{R + G + B}$$

$$V = \frac{R + G + B}{3}$$

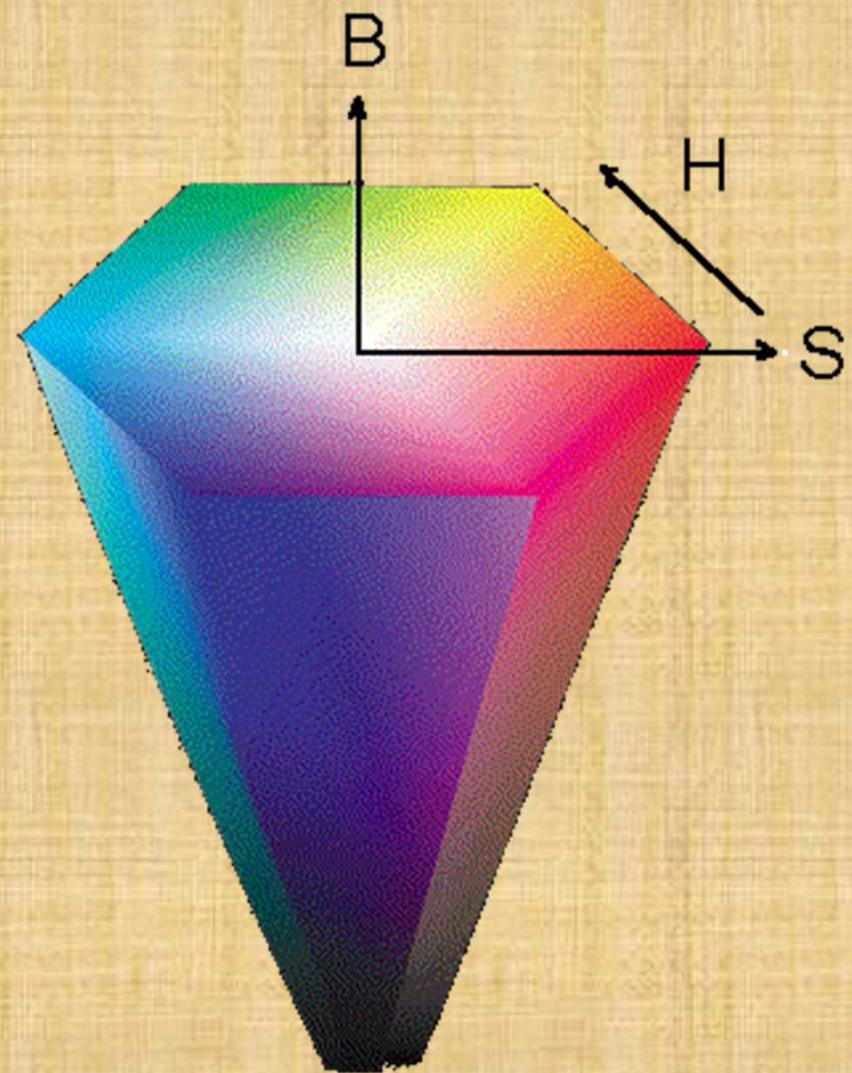


Descomposición HSV



HSB

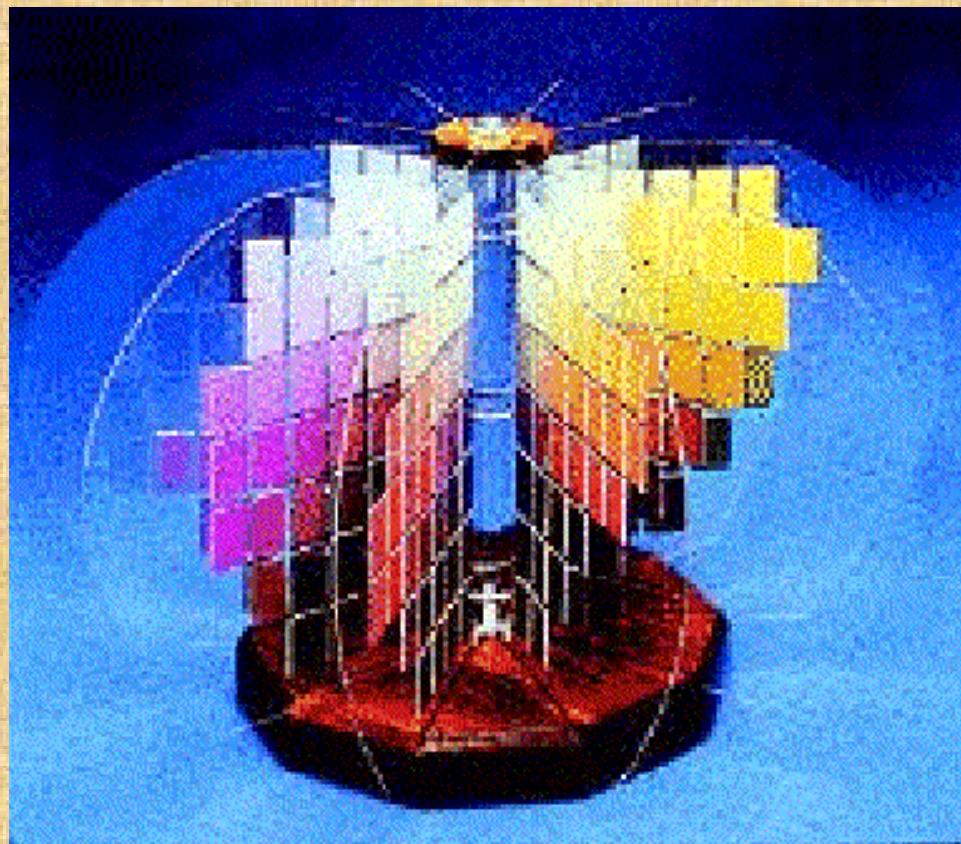
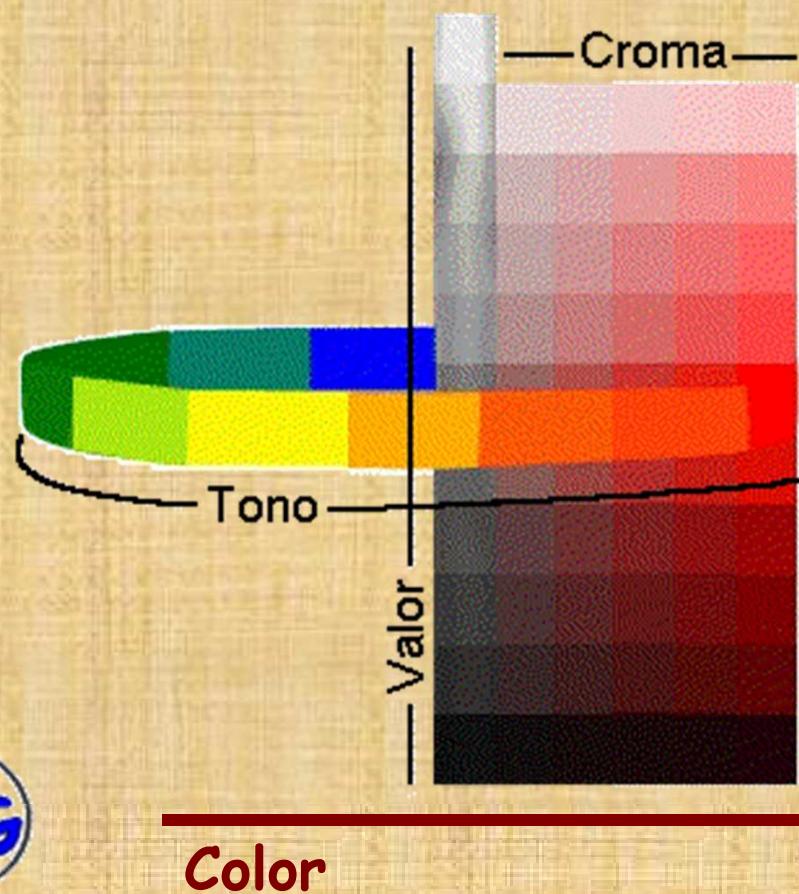
Tono, Saturación, Brillo



Modelo Munsell

Es un esquema de organización perceptual uniforme

Es uno de los primeros intentos de organizar la observación del color



550 Colores

11 Niveles de Gris

L*U*V*

Recomendado por la CIE para cuantificar diferencias de color en monitores en condiciones de **fuentes de iluminación aditivas**

$$L^* = 116 \left[\frac{Y}{Y_0} \right]^{1/3} - 16 \quad \text{para} \quad \frac{Y}{Y_0} > 0.008856$$

$$L^* = 903.3 \left[\frac{Y}{Y_0} \right] \quad \text{para} \quad \frac{Y}{Y_0} \leq 0.008856$$

$$u^* = 13L^* (u' - u'_0) \quad \text{donde} \quad u' = \frac{4X}{X + 15Y + 3Z}$$

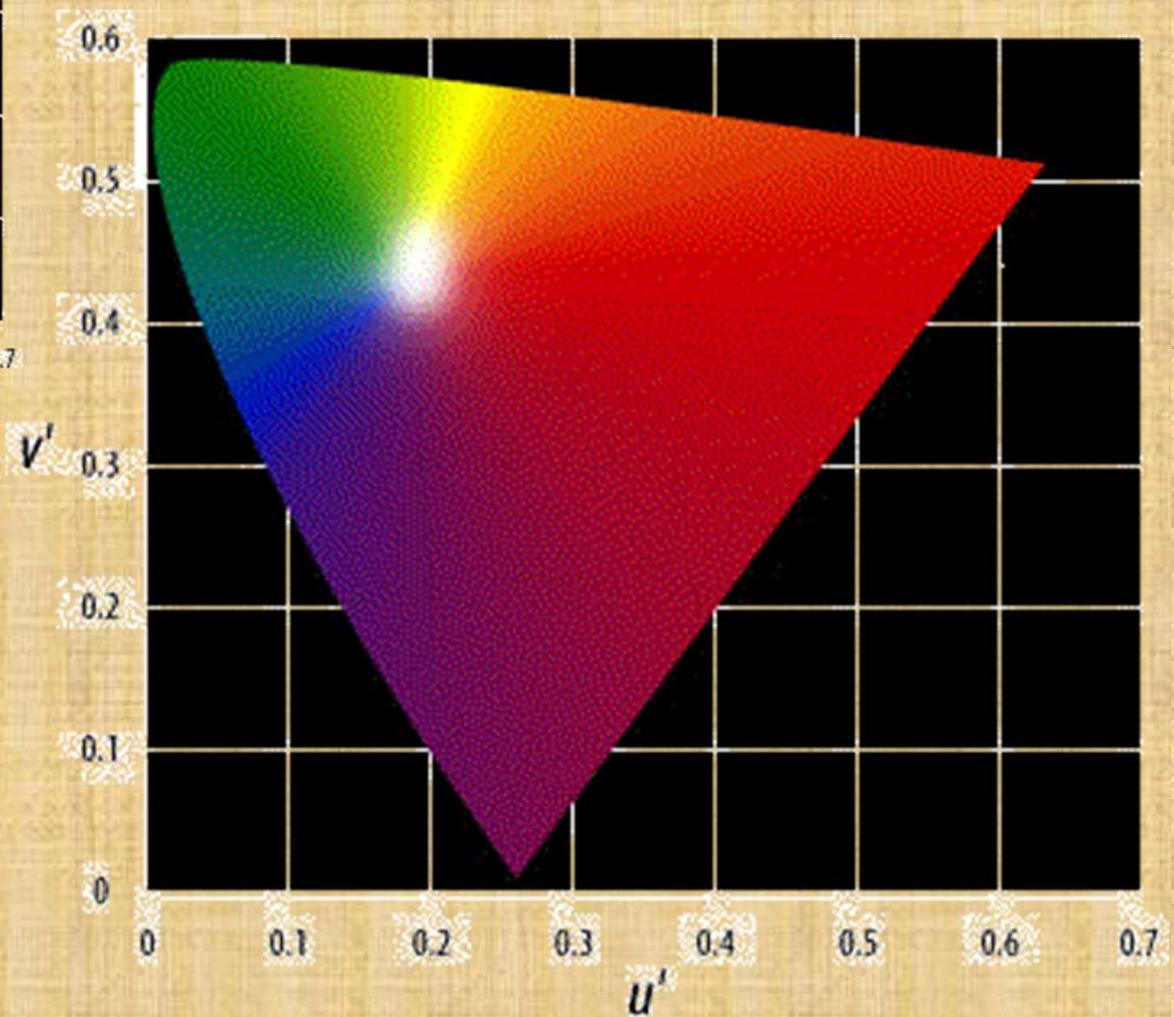
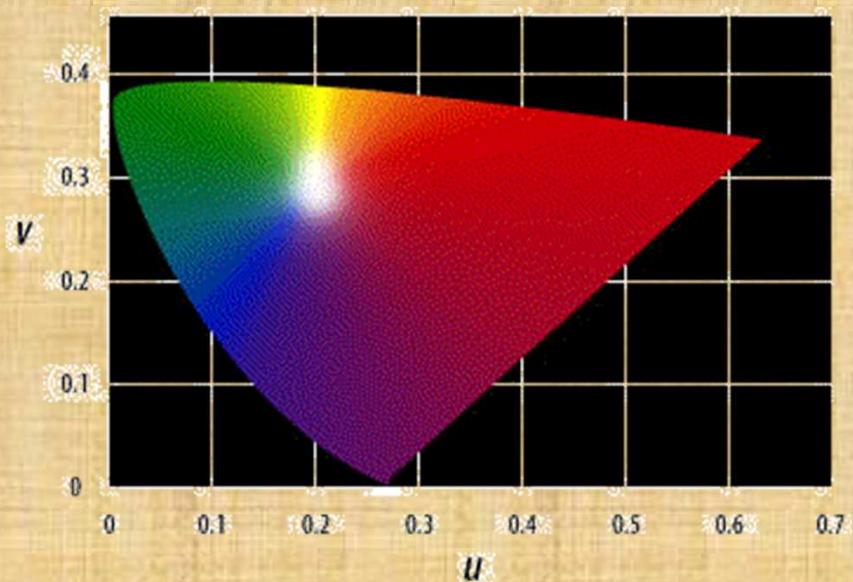
$$v^* = 13L^* (v' - v'_0) \quad \text{donde} \quad v' = \frac{9Y}{X + 15Y + 3Z}$$

Y_0, u'_0, v'_0 son Y, u', v' para la referencia del blanco

Color



$L^*U^*V^*$



$L^*a^*b^*$

Recomendado por la CIE para cuantificar diferencias de color en condiciones de **fuentes de iluminación próximas a la luz solar**

$$L^* = 116 \left[\frac{Y}{Y_0} \right]^{1/3} - 16 \quad \text{para} \quad \frac{Y}{Y_0} > 0.008856$$

$$L^* = 903.3 \left[\frac{Y}{Y_0} \right] \quad \text{para} \quad \frac{Y}{Y_0} \leq 0.008856$$

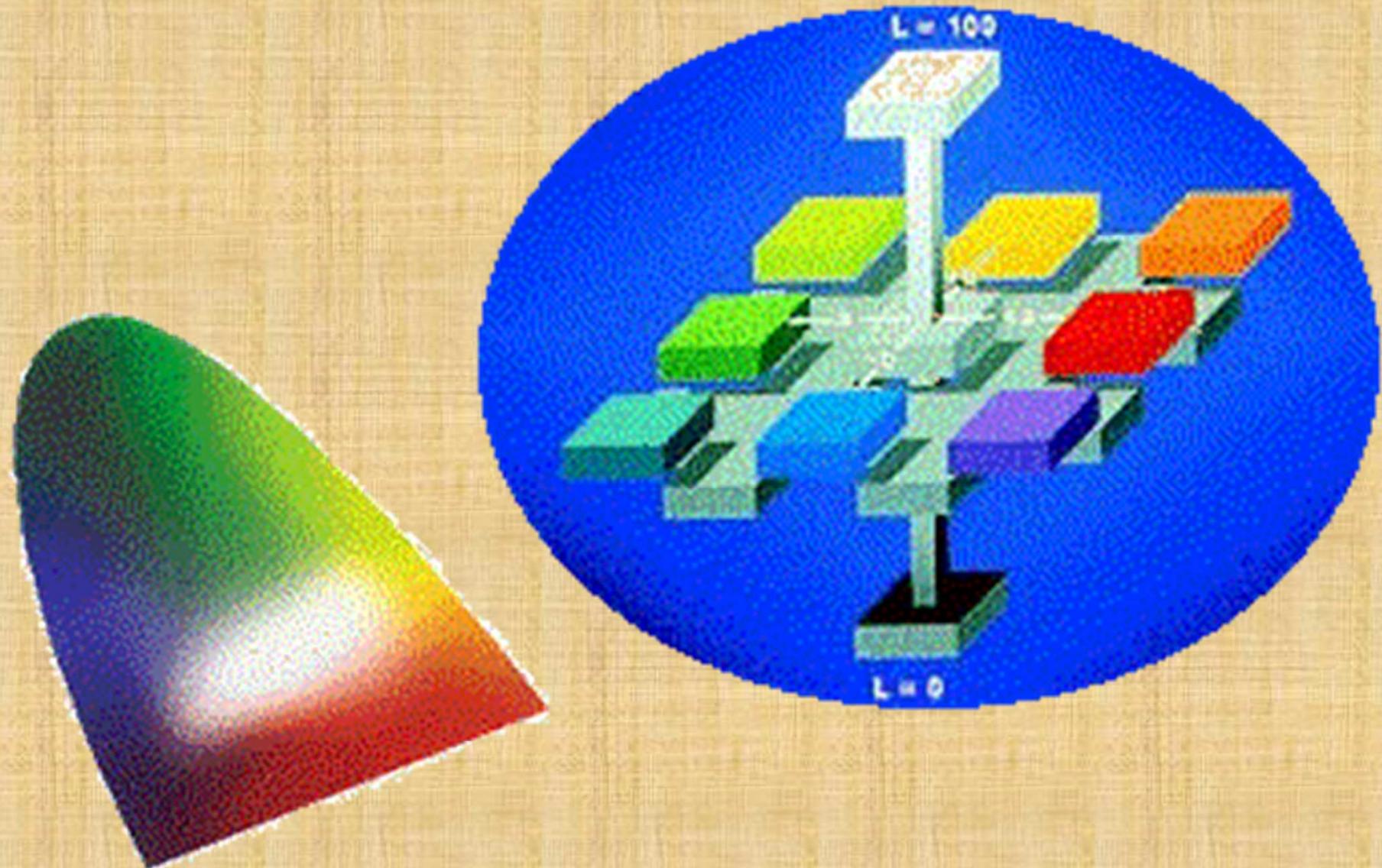
$$a^* = 500 \left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right) \right]$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right) \right]$$

$$f(t) = \begin{cases} t^{1/3} & \text{para } t > 0.008856 \\ 7.787t + \frac{16}{116} & \text{para } t \leq 0.008856 \end{cases}$$

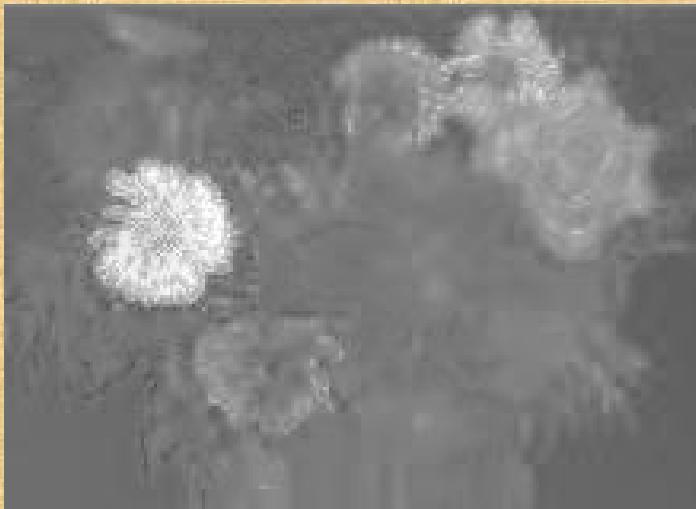
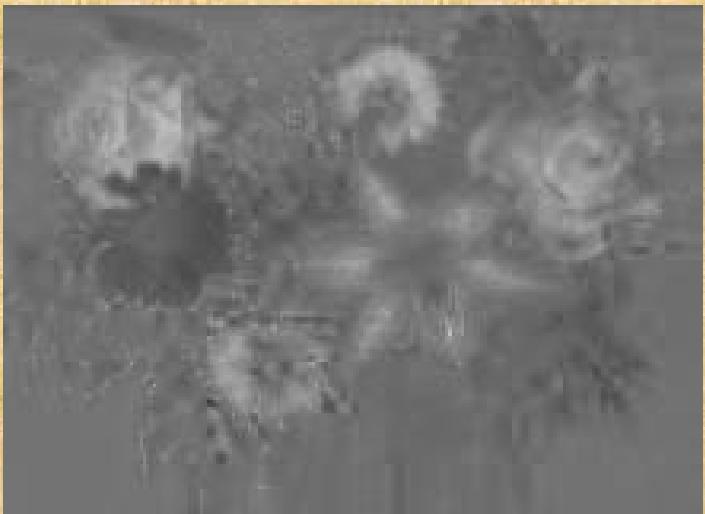
X_0, Y_0, Z_0 son X, Y, Z para la referencia del blanco

$L^*a^*b^*$



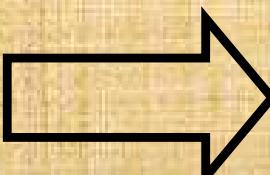
Color

Descomposición $L^*a^*b^*$



Representación del Color

Una vez elegido el modelo de color que se va a utilizar, el siguiente paso es obtener una **propiedad que describa el contenido** en color de las imágenes

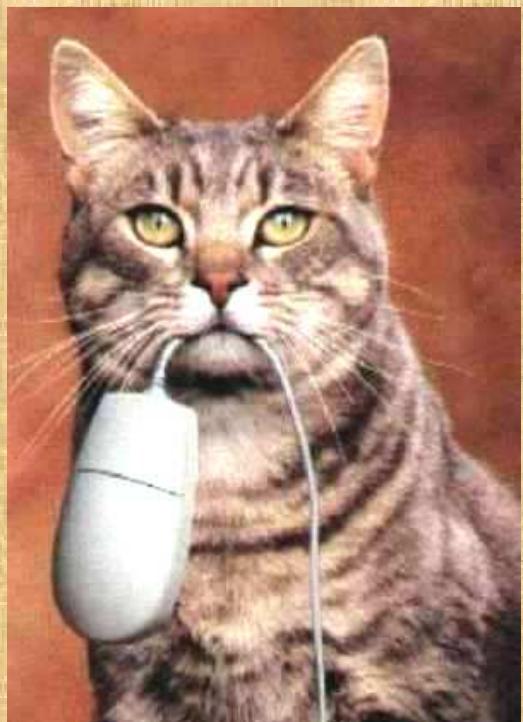
Modelo RGB con 256 niveles por banda de color  **16.777.216** colores diferentes

¿Cómo construir una representación manejable estadísticamente?



Tamaño de Imagen

Las representaciones de las imágenes necesitan, habitualmente, una **gran cantidad de datos**



123, 34, 145	129, 30, 146	111, 67, 146	...
132, 34, 153	131, 31, 154	119, 57, 144	...
127, 35, 146	120, 19, 136	120, 60, 140	...
...

300 x 419 píxeles

Color Modelo RGB

**377100
números enteros
(300 x 419 x 3)**

Cuantización

División del espacio de color en colores representativos

Debe ser suficientemente fina como para poder distinguir **suficientes** colores

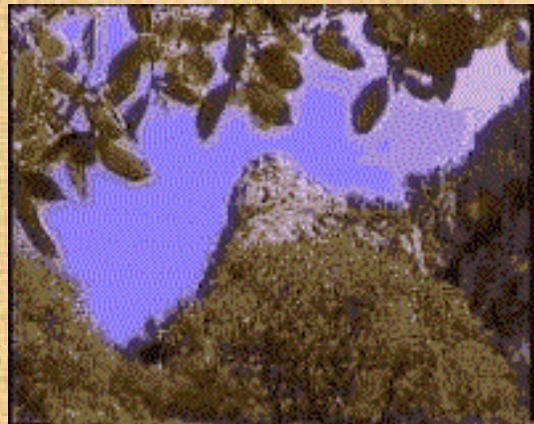
19000



256



16



Realización de la Cuantización

La elección del **número y valor** de los colores representativos afecta a la **percepción de la similaridad** en las imágenes

Espacios de color no uniformes (RGB, HSV, ...):
Cuantización **no uniforme** realizada de forma heurística

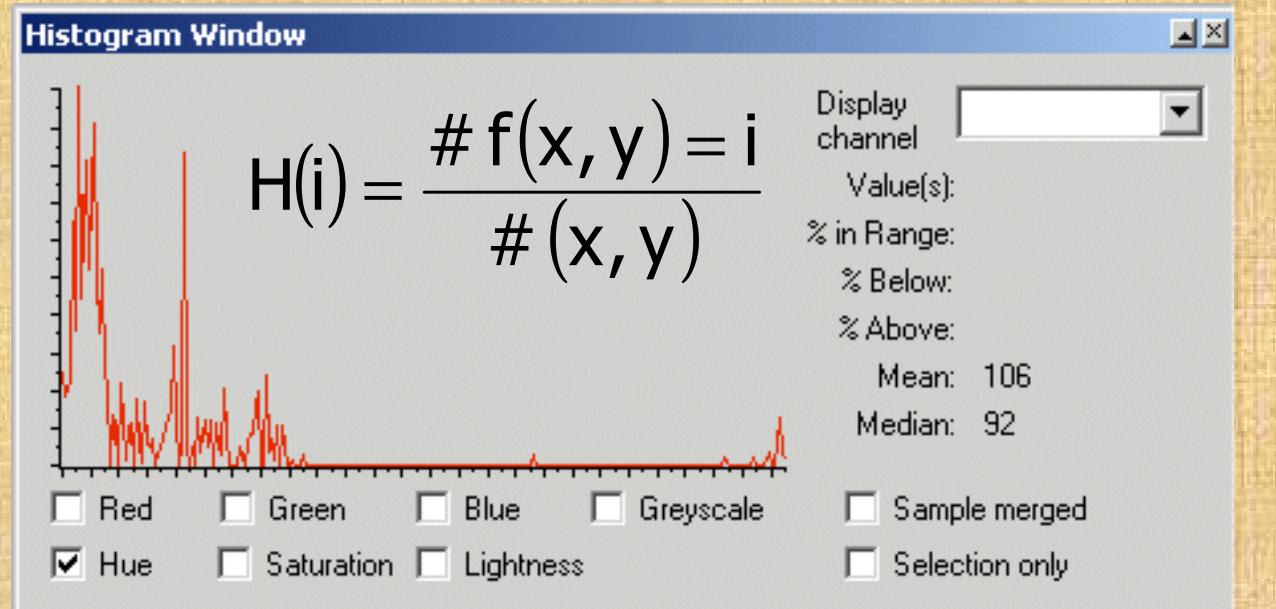
Espacios de color uniformes ($L^*u^*v^*$, $L^*a^*b^*$, ...):
Cuantización **uniforme** realizada mediante clusterización



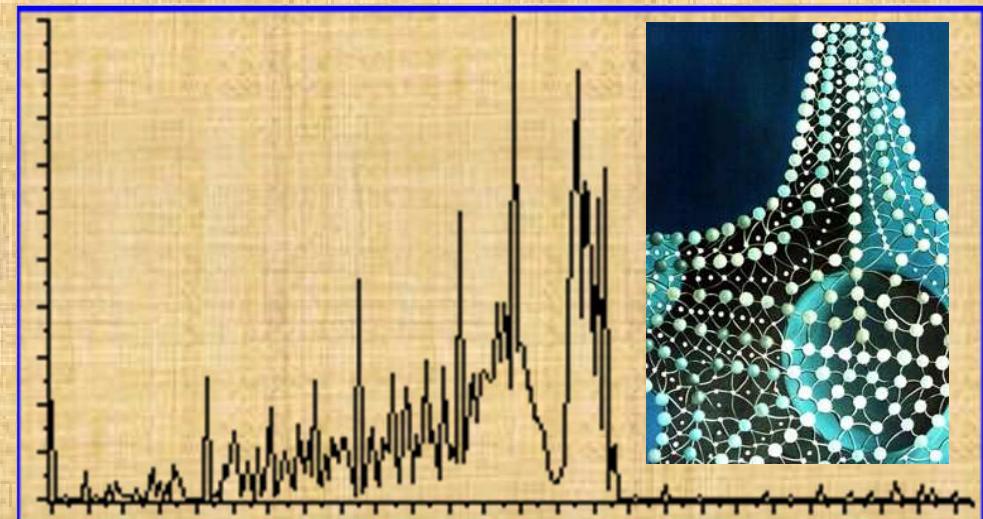
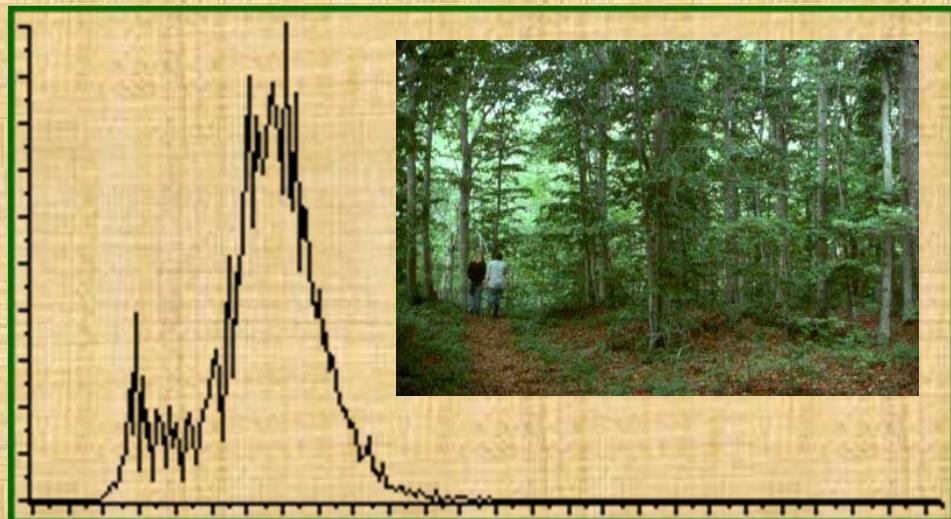
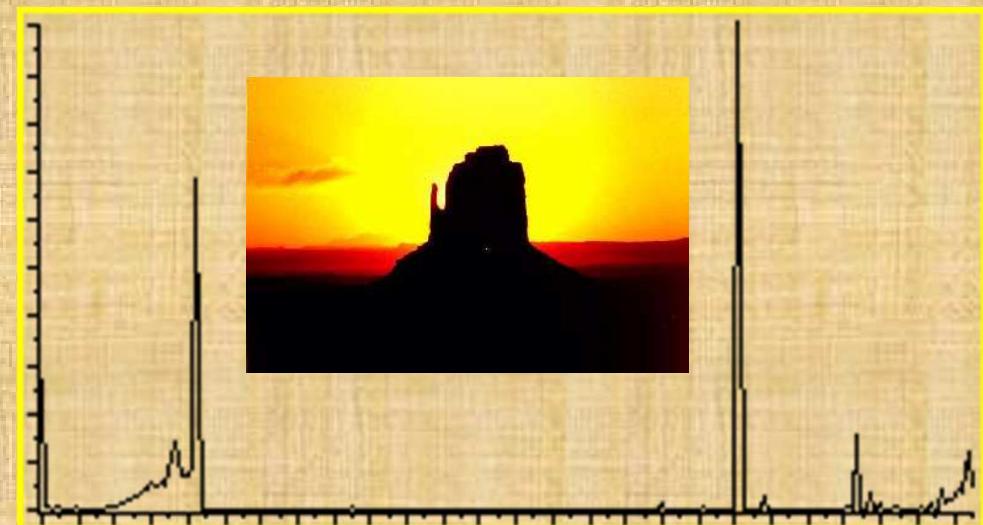
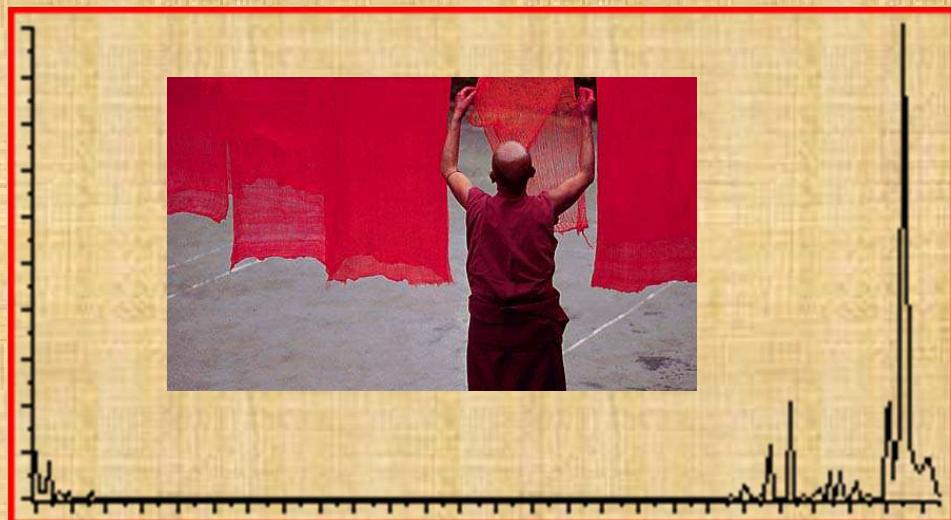
Histograma

Es la forma usual de representar los colores

Una vez elegido el número de colores n , el histograma de una imagen es un **vector H de dimensión n** , tal que **$H(i)$** representa la **fracción de píxeles de color i** presentes en la imagen



Histograma



Histograma



A



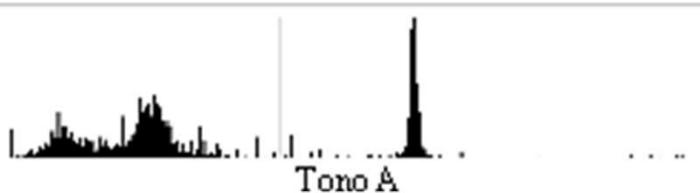
B



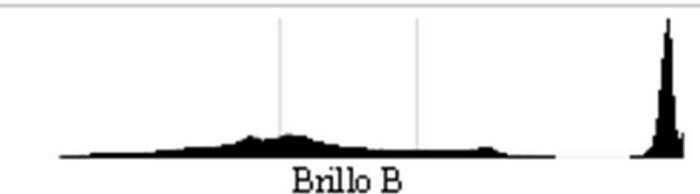
C



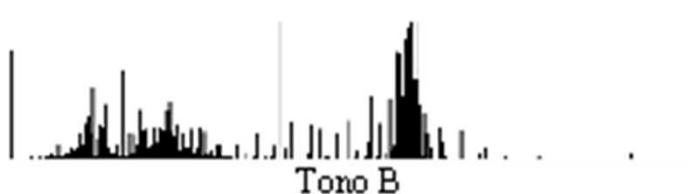
Brillo A



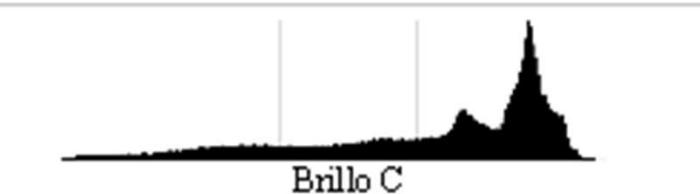
Tono A



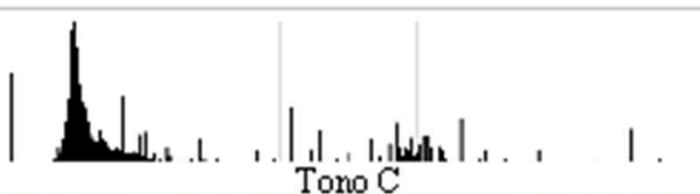
Brillo B



Tono B



Brillo C



Tono C

Ventajas / Problemas

Es la forma más habitual de **describir propiedades de bajo nivel** asociadas al color de las imágenes

Su cálculo es trivial



Insensitive a rotaciones, zoom y cambios de resolución

Sensible a cambios de iluminación

Conjuntos de Colores

Los conjuntos de colores introducen una
representación simplificada del histograma

El contenido de la imagen se representa mediante un
vector binario cuyos valores (1 ó 0) corresponden a
la **presencia significativa o no**, respectivamente,
de un color

Los colores que forman el conjunto de color se
toman de un **diccionario de nombres asociados**
a los colores básicos (asociar nombres a colores
permite crear imágenes mentales)



Colores Básicos

Su aplicación no debe estar restringida a clases de objetos muy específicas (color oliva no es un color válido)

Su significado no debe ser predecible a partir del significado de una parte del objeto (color de las hojas no es un color válido)

Su significado no puede estar incluido en otro color (color verdoso no es un color válido)

Deben destacar psicológicamente



Nombres de Colores

Problema fundamental: Buscar colores básicos que sean iguales para todas las lenguas y culturas

Berlin y Kay (basados en las cartas de color de Munsell)

Blanco Negro Gris Rojo Verde Azul
Amarillo Marrón Rosa Morado Naranja

Se usa una función Gaussiana para aproximar las funciones de pertenencia de un estímulo de color a un color básico



Color Local

La descripción de las **características cromáticas locales** de una imagen tiene como objetivo poder distinguir zonas con contenidos cromáticos diferentes en la imagen

Partición uniforme de la imagen

Partición adaptativa de la imagen (segmentación)



Partición Uniforme

La posibilidad más simple es subdividir la imagen en un **número prefijado de bloques** y describir el contenido global de cada uno de esos bloques

Si los bloques son **pequeños** las regiones con colores uniformes resultan muy subdivididas

Si los bloques son **grandes** varias regiones con colores uniformes pueden estar dentro de un bloque



Segmentación

Por segmentación de una imagen se entiende su **particionamiento en regiones disjuntas** que sean uniformes respecto a alguna característica (por ejemplo, el color)

En general es un proceso difícil



Descripción de Regiones

Una vez segmentada la imagen, cada región individual puede ser descrita mediante los siguientes **atributos**:

Baricentro, Área, Color Medio

Cada región queda descrita por un vector de dimensión 6



Recuperación mediante Color



Búsqueda de Similaridad

El problema fundamental es diseñar **métodos de búsqueda rápidos** que sean capaces de localizar objetos, **exacta o aproximadamente**, similares a una consulta dada

Proporciona una medida de la distancia entre los objetos de la base de datos

La medida de distancia debe estar relacionada con la percepción de similaridad humana



Consideraciones Geométricas

Resultado Experimental: El modelo de color de los triestímulos XYZ puede considerarse como un espacio de Riemann con métrica

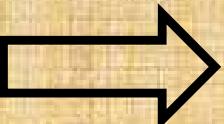
$$d^2 = \sum_{i,j=1}^3 c_{ij} dX_i dY_i$$

donde los **c_{ij} corresponden a la percepción humana de la diferencia** entre los colores X_i y $X_i dX_i$ (diferencias pequeñas)

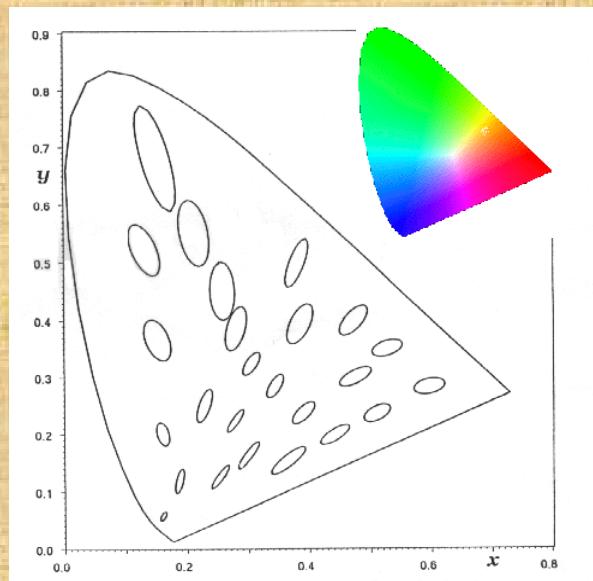


Uniformidad Perceptual

Si los **c_{ij} son constantes**
a lo largo del espacio de
color



El espacio de color se
puede considerar como
un **espacio Euclídeo**



La distancia entre dos colores será
proporcional a la distancia entre ellos

**El espacio de los triestímulos XYZ no
verifica la uniformidad perceptual**

Distancia entre Colores

Las transformaciones que definen **los modelos L*u*v*** y **L*a*b*** **aproximan un espacio de Riemann a un espacio Euclídeo** (verifican la uniformidad perceptual)

$$(\Delta S)^2 = (\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2$$
$$(\Delta S)^2 = (\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2$$

Las aproximaciones matemáticas utilizadas introducen desviaciones en algunas regiones del espacio de color



Propiedades de la Distancia

Si d es una medida de distancia y s_1 , s_2 y s_3 son tres estímulos genéricos en un espacio de color perceptual

Autosimilaridad: $d(s_1, s_1) = d(s_2, s_2) = 0$

Minimalidad: $d(s_1, s_2) \geq d(s_1, s_1) = 0$

Simetría: $d(s_1, s_2) = d(s_2, s_1)$

Desigualdad triangular: $d(s_1, s_2) + d(s_2, s_3) \geq d(s_1, s_3)$



Conjuntos de Colores

La distancia entre el conjunto de colores consulta c_q y el conjunto de colores imagen c_d viene dada por

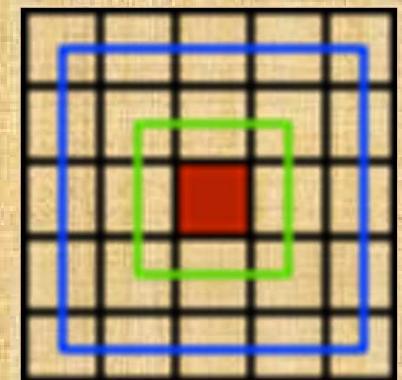
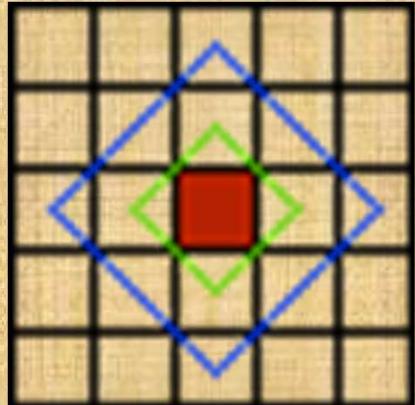
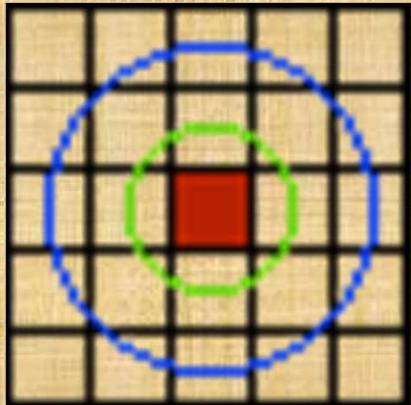
$$D = \sqrt{(c_q - c_d)^t A (c_q - c_d)}$$



Distancia de Histogramas

Dado un **histograma consulta $H(I_q)$** y un **histograma imagen $H(I_d)$** , ambos con n bins, su distancia de intersección de histograma viene dada por

$$d_h(I_q, I_d) = \left[\sum_{j=1}^n (H(I_q, j) - H(I_d, j))^p \right]^{1/p}$$



L_p Norma

p=1 distancia city-block

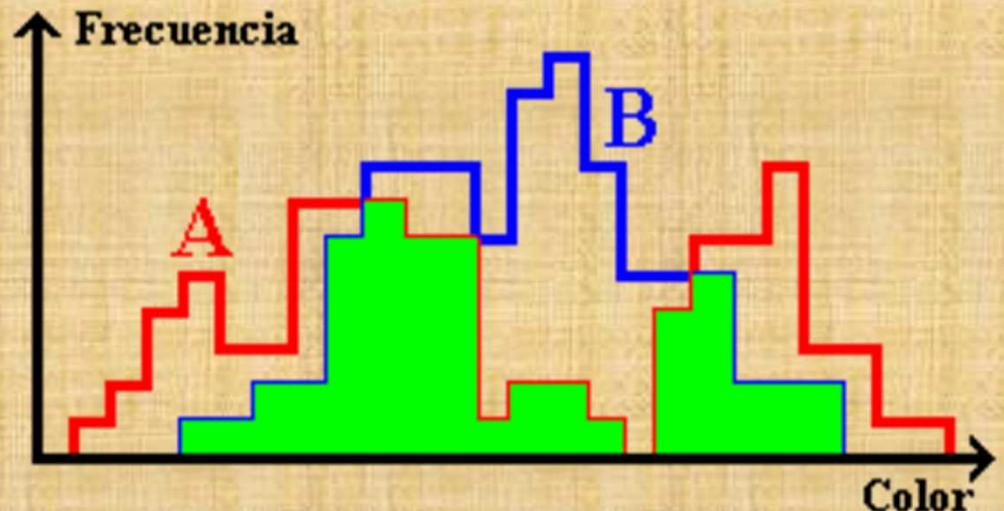
p=2 distancia Eculídea

Intersección de Histogramas

Sólo contribuyen a la distancia los **colores presentes en la imagen consulta**

$$d_h(I_q, I_d) = \frac{\sum_{j=1}^n \min(H(I_q, j), H(I_d, j))}{\sum_{j=1}^n (H(I_q, j))}$$

Dados dos histogramas **A** y **B**, la frecuencia de color mínima entre bins homólogos viene dada por la intersección **$A \cap B$**



Forma Cuadrática

Dado un **histograma consulta $H(I_q)$** y un **histograma imagen $H(I_d)$** su distancia de histograma viene dada por la **distancia de Mahalanobis**

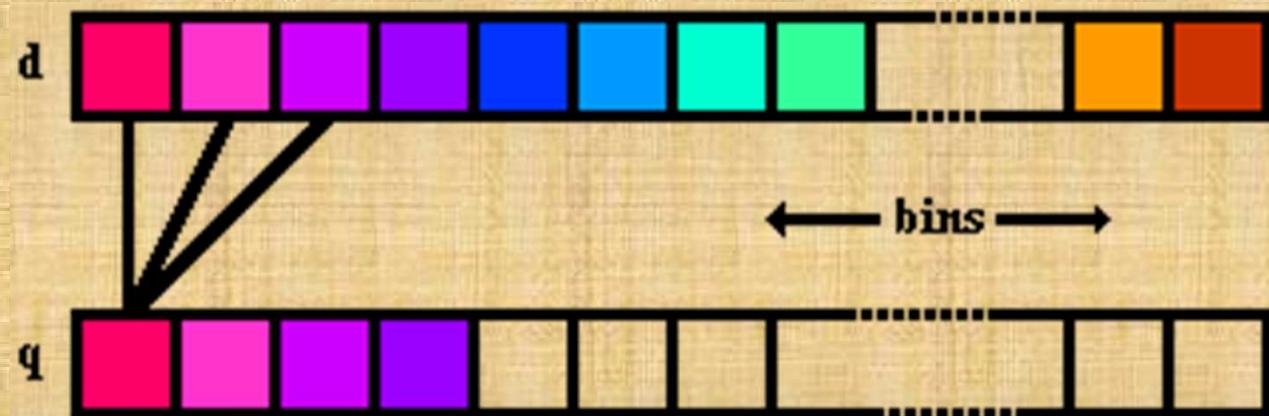
$$d_h(I_q, I_d) = \sqrt{(H(I_q, i) - H(I_d, j))^t A (H(I_q, i) - H(I_d, j))}$$

donde **$A=[a_{ij}]$ es la matriz de similaridad** que pondera la similitud entre los bins i y j



Cross-Talk

El cálculo de la distancia mediante una forma cuadrática implica “**cross-talk**” sobre las propiedades



La **distancia entre colores** no depende sólo de la **distancia entre bins homólogos** del histograma sino también de la **distancia entre colores vecinos**

Recuperación mediante Color



Resultados de Recuperación



Problemas

Dimensionalidad: El número de bins en el histograma puede ser grande (64, 128, 256, ...) y esto **reduce** drásticamente **la eficiencia** del método de acceso

Complejidad: Dada la naturaleza cuadrática de la distancia (hay experimentos que demuestran que es la medida que mejor aproxima la percepción humana de la similaridad de colores) la **complejidad computacional crece** rápidamente



Modelado de la Forma

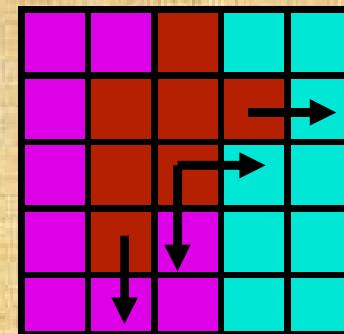


Definiciones

Un **eje** es un **cambio significativo en una propiedad** de la imagen (usualmente, la intensidad o el color)

Un **contorno** es un **grupo de ejes unidos** espacialmente bajo alguna condición

Un **objeto** se define a través de un **contorno cerrado** que lo envuelve



Definiciones

Los **descriptores de forma** son medidas que **cuantifican** los objetos

Forma global: sólo necesitan la imagen de ejes

Forma local: asumen conocido el contorno del objeto a describir



Hipótesis muy restrictiva ya que disponer de curvas cerradas para cada objeto de la imagen implica que se puedan separar automáticamente todos los objetos presentes en la imagen del fondo

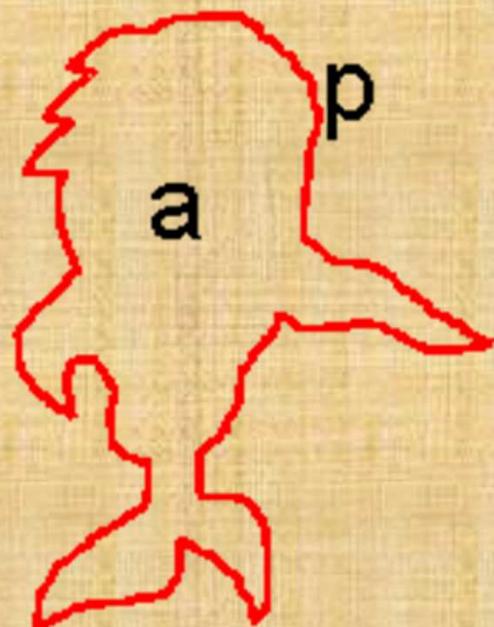
Geometría Simple

Se trata de un descriptor **muy simple**, pero **muy eficaz** en determinados contextos

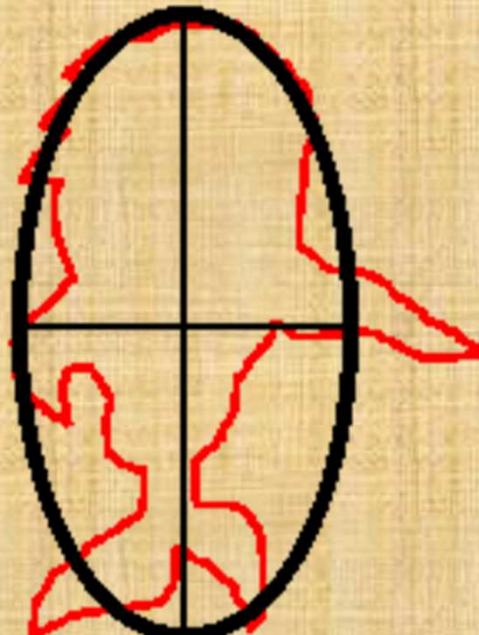
Forma
Local

Área (a)

Compacidad (p^2/a)



Elongación



Momento

Considerar una imagen binaria con valor 1 para los puntos del contorno y valor 0 para el resto

**Forma
Local**

El momento de orden ij viene dado por

$$M_{ij} = \sum_{x,y} f(x,y)x^i y^j$$

Los momentos **son invariantes a translaciones, rotaciones y escala** y pueden considerarse normalizados respecto al momento central M_{00} y a la varianza de la distribución respecto al baricentro

Son muy utilizados en la descripción de formas sencillas



Transformada de Fourier

La transformada 2D de Fourier de una función $f(x,y)$ definida en $\{0, \dots, x-1\} \times \{0, \dots, y-1\}$ viene dada por

**Forma
Local**

$$F(m,n) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x,y) e^{-2\pi j \left(\frac{mx}{X} + \frac{ny}{Y} \right)}$$

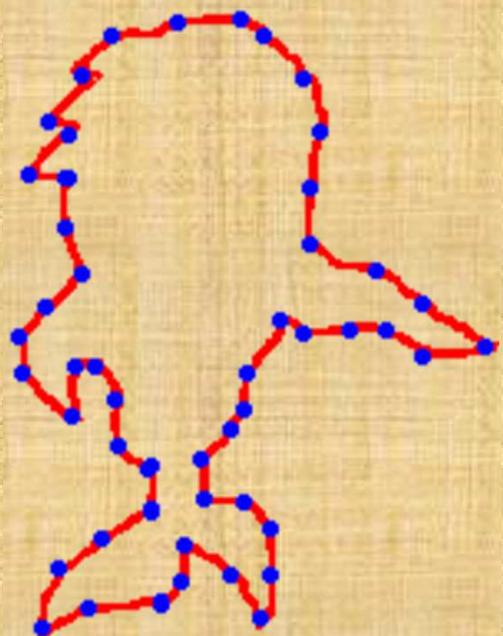
Su uso no resulta sencillo pues no existe una relación inmediata entre las componentes de la frecuencia y su significado en la forma del objeto



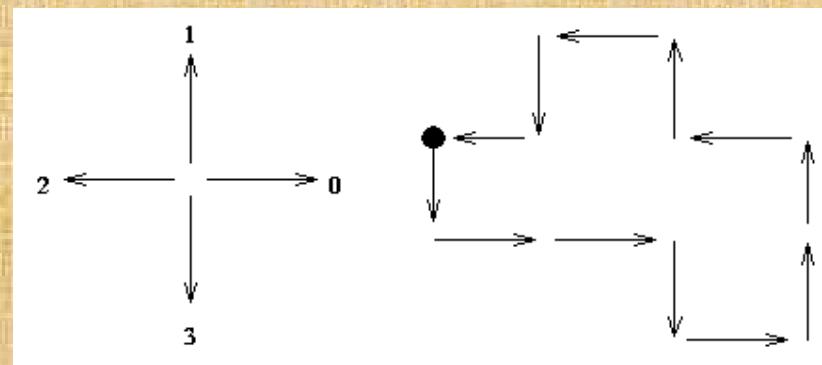
Códigos de Cadena

Se subdivide la curva en **segmentos** y se caracterizar cada segmento por su **longitud** y su **orientación**

**Forma
Local**



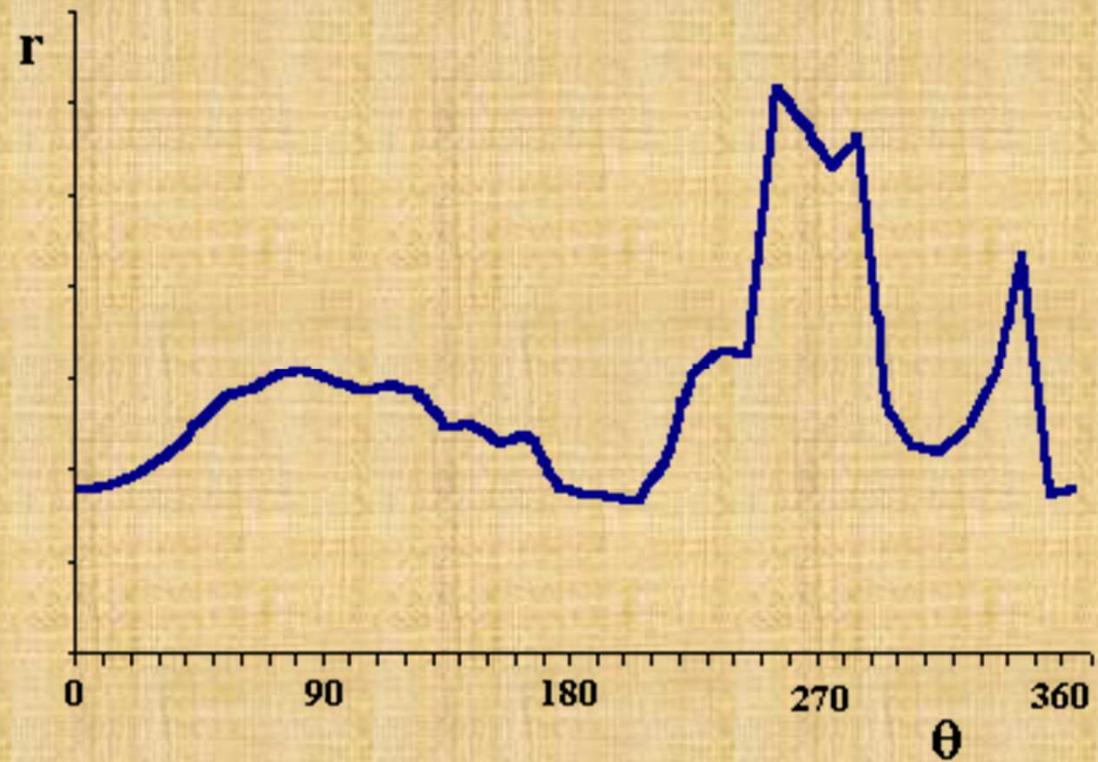
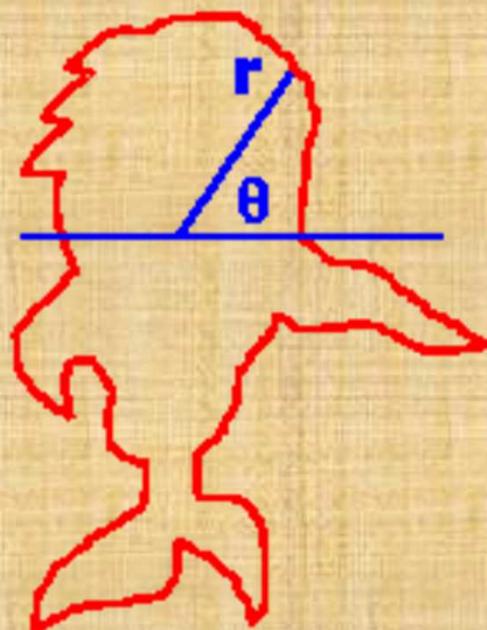
La forma del objeto queda descrita por la secuencia $(l_1, \theta_1) (l_2, \theta_2) \dots (l_n, \theta_n)$



Coordenadas Polares

Se describe la forma por la **función Radio frente a Ángulo** calculada desde el baricentro del objeto

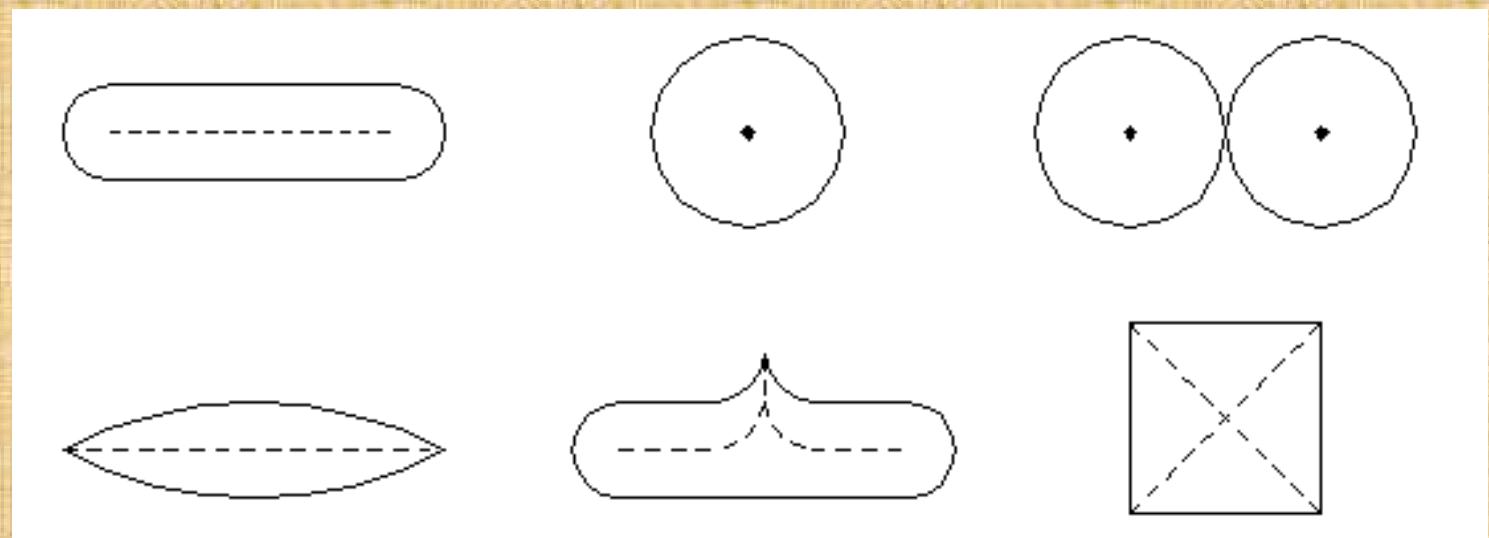
Forma Local



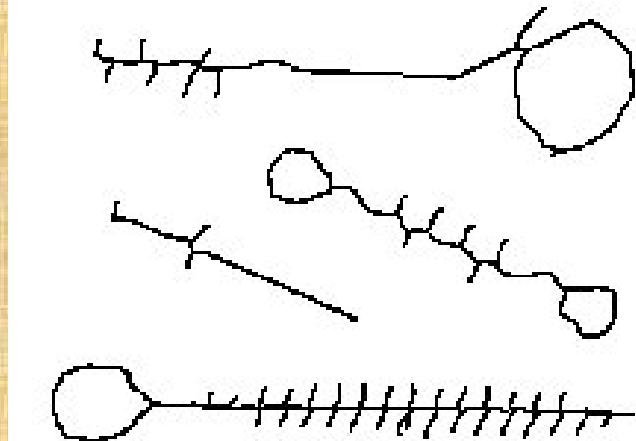
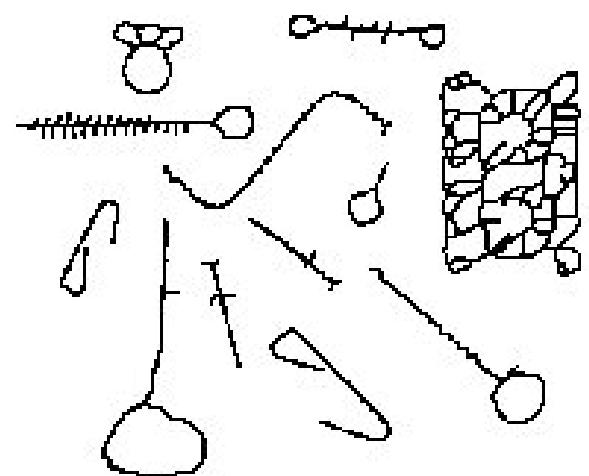
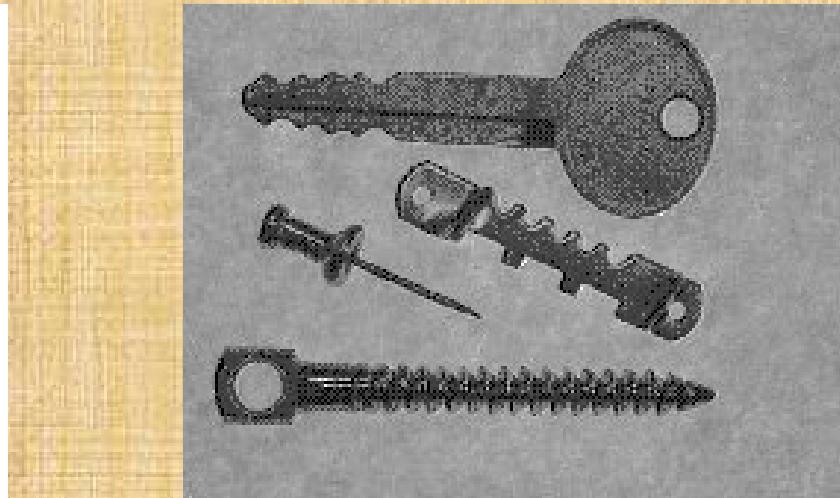
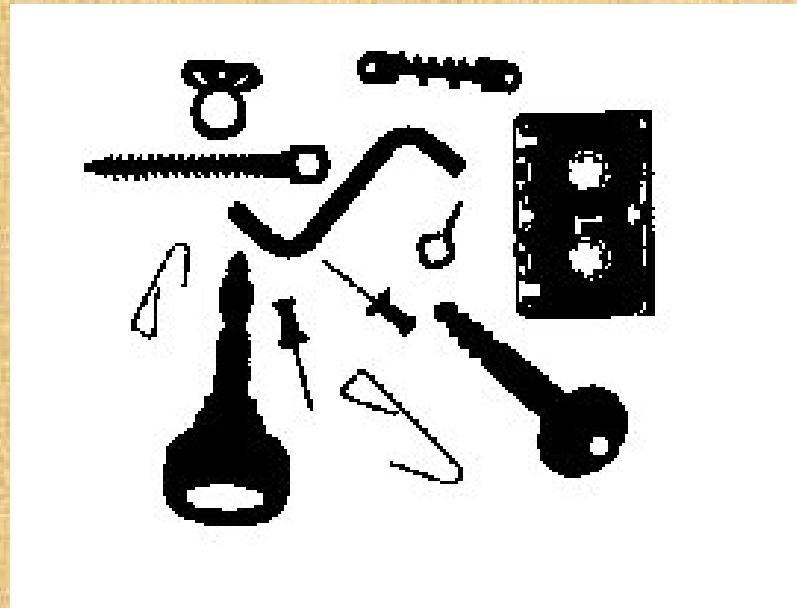
Esqueleto

El **esqueleto** es una representación mediante un grafo de la forma del objeto

Para calcularlo se realizan **erosiones desde los puntos del contorno**; cuando dos direcciones diferentes de erosión coinciden, ese punto pertenece al esqueleto



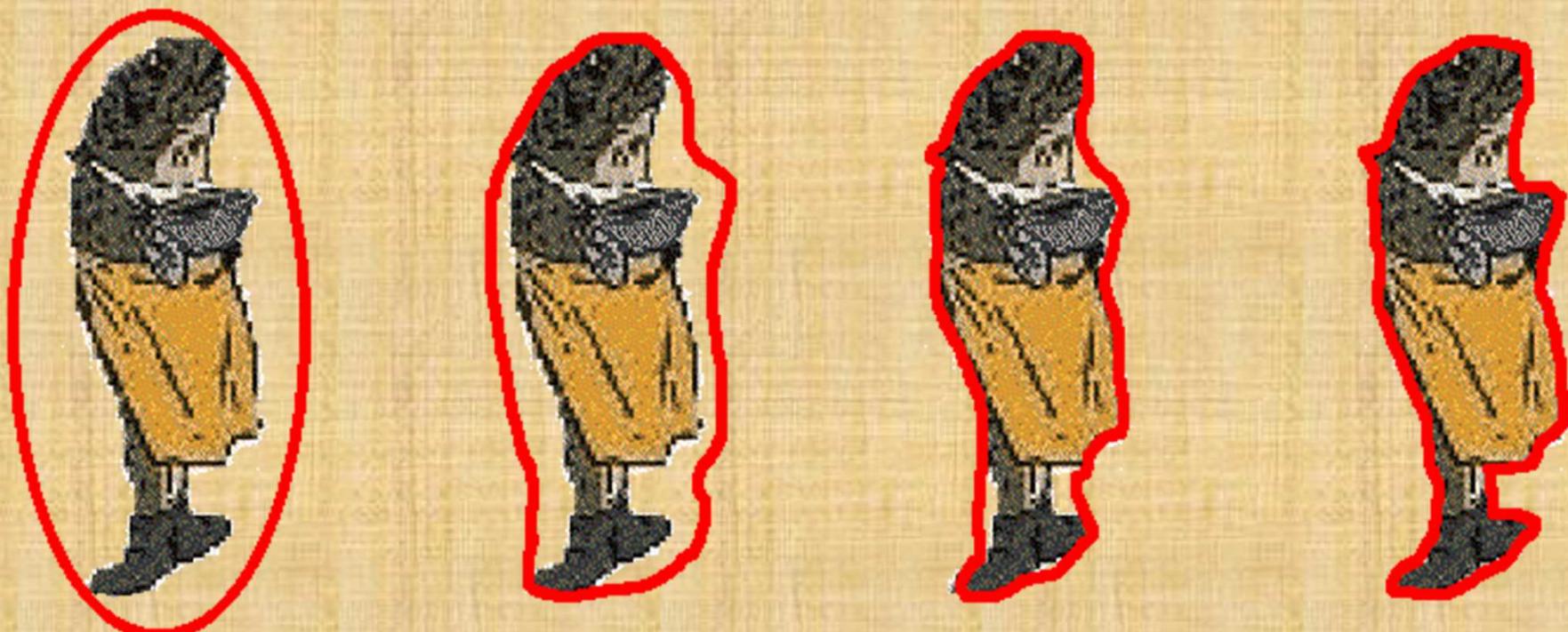
Ejemplos Esqueleto



Contornos Activos

La técnica de los contornos activos permite localizar formas en imágenes mediante la **evolución de una curva cerrada** predefinida sobre la imagen de ejes

Forma
Global



Contornos Activos

El proceso de deformación elástica corresponde a la **minimización de un funcional** formado por dos términos de energía; el primero relacionado con la **forma del contorno activo** y el segundo relacionado con su **posición respecto a los ejes** presentes en la imagen

Forma Global

Una vez minimizado el funcional se pueden utilizar los **valores de sus dos términos en medidas de similaridad**



Correlación de Ejes

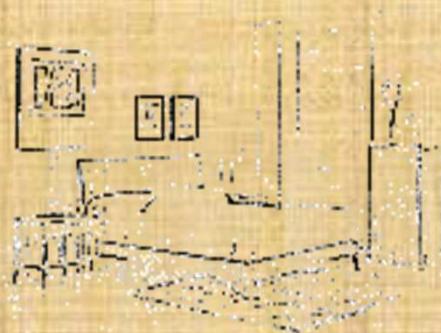
Se representa el contenido de forma de las imágenes mediante las **imágenes de ejes**

Forma Global

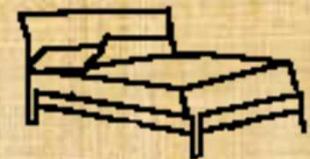
Base de datos
de imágenes



Imágenes
filtradas



Patrones



correlación

Correlación de Ejes

Requisitos para que la correlación de ejes sea eficaz:

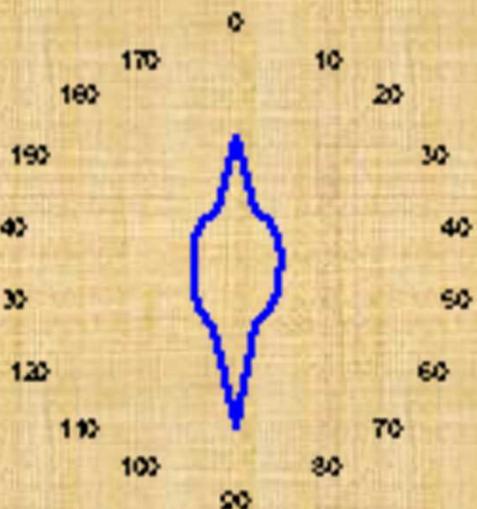
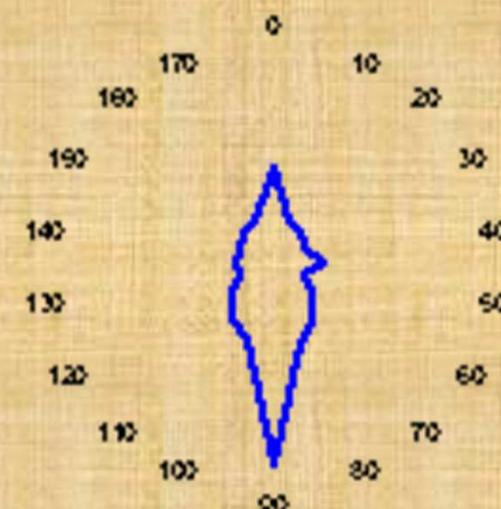
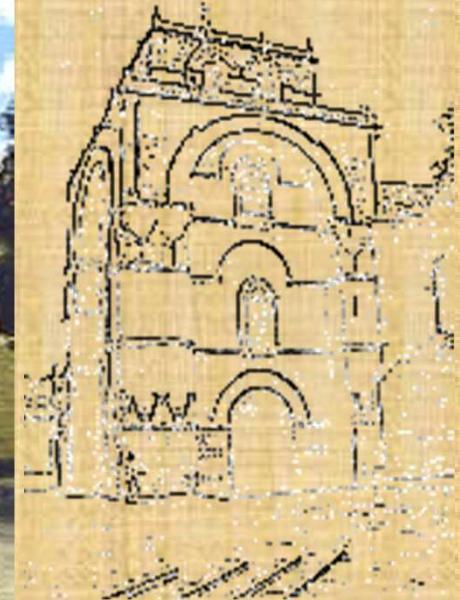
Los **ejes relevantes deben corresponder** efectivamente **a elementos estructurales** de la escena y no a detalles insignificantes

La correlación debe tener una **cierta elasticidad** para localizar objetos que no sean exactamente iguales a los patrones



Histograma de Orientaciones

Se representa el contenido de forma mediante su **distribución de ejes**



Recuperación mediante Forma



Distancia entre Formas

Existen evidencias experimentales que indican que **la percepción humana de la similaridad entre formas no sigue completamente los axiomas de la métrica**

A pesar de ello, la mayoría de los sistemas de recuperación usan **modelos de similaridad métricos**



Forma Local

Para formas **aproximadas mediante curvas poligonales** que usan propiedades locales como Av (ángulo del vértice), Dv (distancia al siguiente vértice), Xv e Yv (coordenadas del vértice)

la similaridad se calcula midiendo el **número de cambios necesarios** para transformar una curva en otra



Forma Local

Los sistemas que emplean **representaciones de forma basadas en propiedades** (momentos, coeficientes de Fourier, histogramas de orientación, etc) tratan la forma como **vectores de propiedades** y calculan la similaridad como **distancia en el espacio de propiedades**

Si la dimensión de los vectores de propiedades no es demasiado grande, los vectores se organizan en estructuras de índice que permiten facilitar el acceso a la base de datos

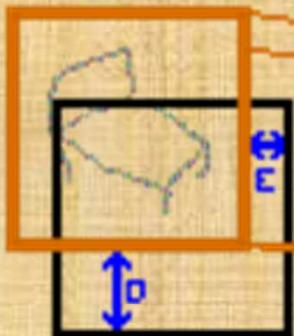
Estos modelos utilizan típicamente distancia Euclidea



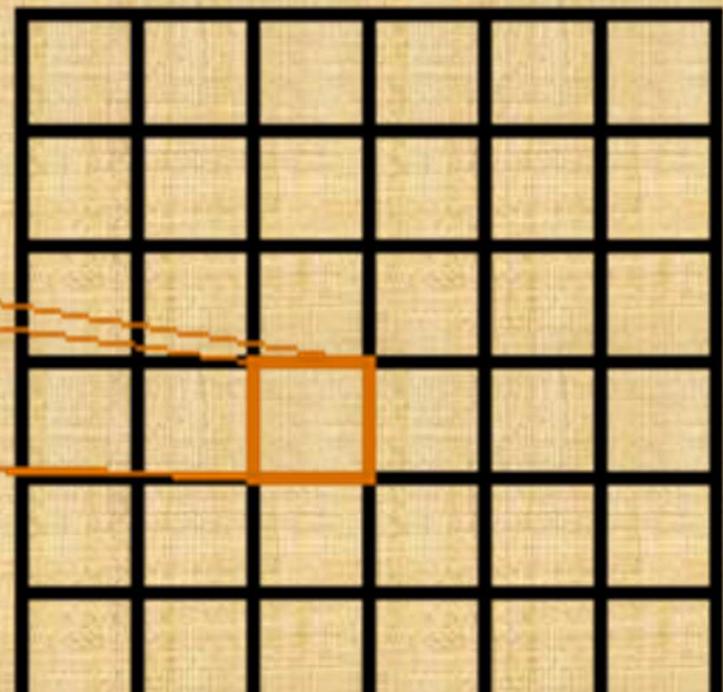
Correlación de Ejes

La medida de similaridad viene dada por el cálculo de la **correlación entre el patrón consulta y la imágenes de la base de datos**

$$C = \sum_a^M \sum_b^N C^{ab}$$



$$C^{ab} = \max C_{DE}^{ab}$$



M x N Bloques

Modelado de la Textura



¿Qué es una Textura?



Texturas Artificiales

Primitivas idénticas

* * * * * * * * *
* * * * * * * * *
* * * * * * * * *
* * * * * * * * *
* * * * * * * * *
* * * * * * * * *
* * * * * * * * *

** * * * * * * *
** * * * * * * *
** * * * * * * *
** * * * * * * *
** * * * * * * *
** * * * * * * *
** * * * * * * *

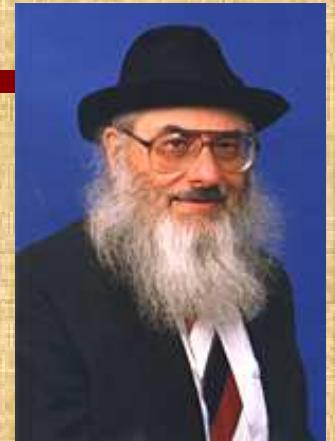
#

Relaciones espaciales idénticas



Coocurrencia Matemática

Caracteriza la textura mediante la **dependencia espacial** que aparece **en la distribución de los niveles de gris** en la imagen



$$P_{xy}(i_1, i_2) = \frac{\# \left\{ [(x_1, y_1), (x_2, y_2)] \mid \begin{array}{l} |x_1 - x_2| = x \quad |y_1 - y_2| = y \\ I(x_1, y_1) = i_1 \quad I(x_2, y_2) = i_2 \end{array} \right\}}{\# I}$$



**Histograma
de Orden 2**

La matriz de coocurrencia matemática no es necesariamente simétrica

Ejemplo Coocurrencia

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

$$P_{0^\circ,1} = \begin{vmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$$

$$P_{135^\circ,1} = \begin{vmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{vmatrix}$$



Coocurrencia Matemática

La textura se caracteriza mediante **descriptores calculados sobre las matrices de coocurrencia** matemática

Energía

$$\sum_{i_1, i_2} P_{xy}^2(i_1, i_2)$$

Entropía

$$\sum_{i_1, i_2} P_{xy}(i_1, i_2) \log P_{xy}(i_1, i_2)$$

Máximo

$$\max P_{xy}(i_1, i_2)$$

Contraste

$$\sum_{i_1, i_2} (i_1 - i_2)^k P_{xy}^{-1}(i_1, i_2)$$

Correlación

$$\sum_{i_1, i_2} i_1 i_2 P_{xy}^2(i_1, i_2)$$

Clusterización

$$\sum_{i_1, i_2} (i_1 + i_2 - 2\mu)^k P_{xy}(i_1, i_2)$$



Autocorrelación

Correlación de la imagen consigo misma desplazada

Si el tamaño del **patrón de textura es grande**,
la función de autocorrelación varía **lentamente**
con la distancia

Si el tamaño del **patrón de textura es pequeño**,
la función de autocorrelación varía **rápidamente**
con la distancia



Dimensión Fractal

La dimensión fractal se mide, por ejemplo, mediante el coeficiente de Hurst

Coeficiente de Hurst: Pendiente de la mayor diferencia de nivel de gris en píxeles a lo largo de un línea en escala logarítmica frente a la distancia

Muy útil para texturas isotrópicas



Modelos Autoregresivos

La textura se modela mediante los **coeficientes** del modelo

$$v(p) = \mu + \sum_{q \in W} \theta(q)v(p+q) + \epsilon(p)$$

p es un píxel de la imagen;

W es una ventana centrada en p;

μ es la media de la imagen;

θ caracteriza la dependencia de p con sus vecinos;

ϵ es un ruido gaussiano de media nula y varianza σ^2

θ **caracteriza la direccionalidad** de la textura

σ **caracteriza la granularidad** de la textura

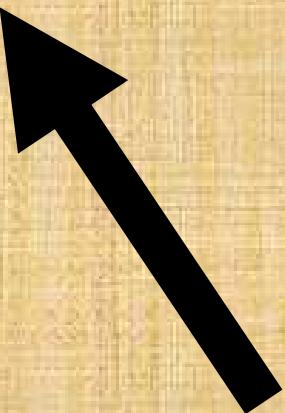


Recuperación mediante Textura



Distancia entre Texturas

Cuando se trabaja con texturas, lo usual es realizar el **cálculo de distancia Euclídea en el espacio de propiedades utilizado**



n-dimensional,
con **n** **usualmente bastante alto**

Modelado de las Relaciones Espaciales



Cadenas 2D

Una cadena 2D es una **representación simbólica de la disposición** de los objetos

Cada símbolo de la cadena se obtiene proyectando la posición de cada región etiquetada sobre ejes cartesianos



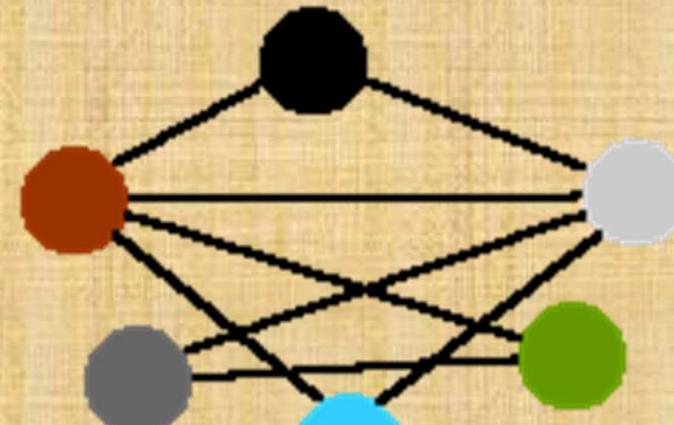
Hor: **PG<R**

Ver: **PR<G**

Grafos de Adyacencias



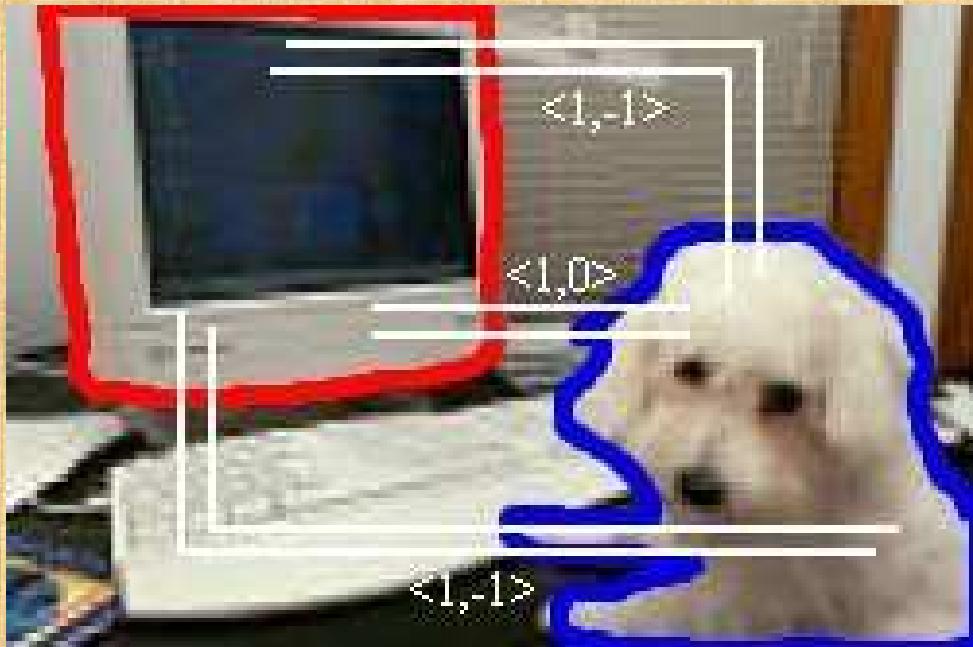
segmentación



descripción
de relaciones

Walkthrough

Un Walkthrough es una **clase de equivalencia** sobre el **camino** entre cada punto del **objeto A** y cada punto del **objeto B**



El camino se forma en dos pasos ordenados: desplazamiento horizontal y desplazamiento vertical

Walkthrough

La relación espacial entre A y B
se describe a través de la **matriz**

$$S(A,B) = \begin{bmatrix} w_{-1,1}(A,B) & w_{0,1}(A,B) & w_{1,1}(A,B) \\ w_{-1,0}(A,B) & w_{0,0}(A,B) & w_{1,0}(A,B) \\ w_{-1,-1}(A,B) & w_{0,-1}(A,B) & w_{1,-1}(A,B) \end{bmatrix}$$

Donde $w_{i,j}(A,B)$ representa el **número de parejas de puntos relacionadas por un camino** equivalente $\langle i,j \rangle$



Recuperación mediante Relaciones Espaciales



Cadenas 2D

La similaridad se establece **comparando las cadenas** que representan a las imágenes

Dos imágenes son similares si la cadena que define a una de ellas **está incluida** en la cadena que define a la otra



Grafos de Adyacencias

imagen $\stackrel{\text{def}}{=}$ $\langle E, a, w \rangle$

E = conjunto de relaciones espaciales

$a : E \longrightarrow A$

$w : E \times E \longrightarrow W$

La comparación entre modelos de grafos de una especificación de consulta $\langle Q, a_q, w_q \rangle$ y una descripción de imagen $\langle D, a_d, w_d \rangle$ **involucra una asociación de entidades** en la consulta con un conjunto de entidades en la imagen llamada interpretación



Walkthrough

La similaridad de **relaciones especificadas mediante matrices de walktrought** se obtiene **comparando los coeficientes** de las matrices



Otras Dimensiones en la Información Multimedia



Información Multimedia

Imagen tridimensional



Vídeo



Imagen tridimensional en movimiento

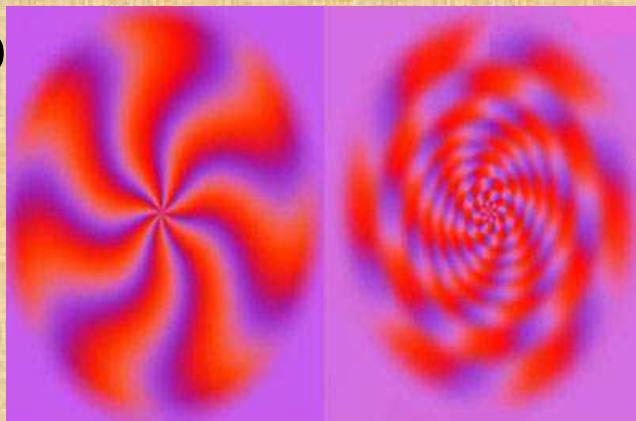
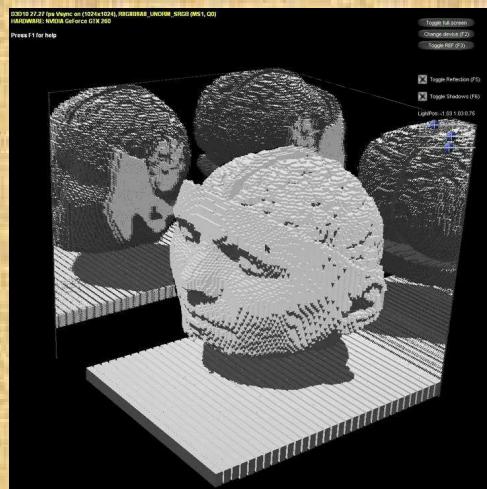


Imagen Tridimensional

Una **imagen tridimensional** se puede definir como un valor particular de la **distribución de probabilidad conjunta** de un conjunto de **N₁xN₂xN₃ variables aleatorias** ordenadas espacialmente sobre una rejilla de N₁xN₂xN₃ posiciones

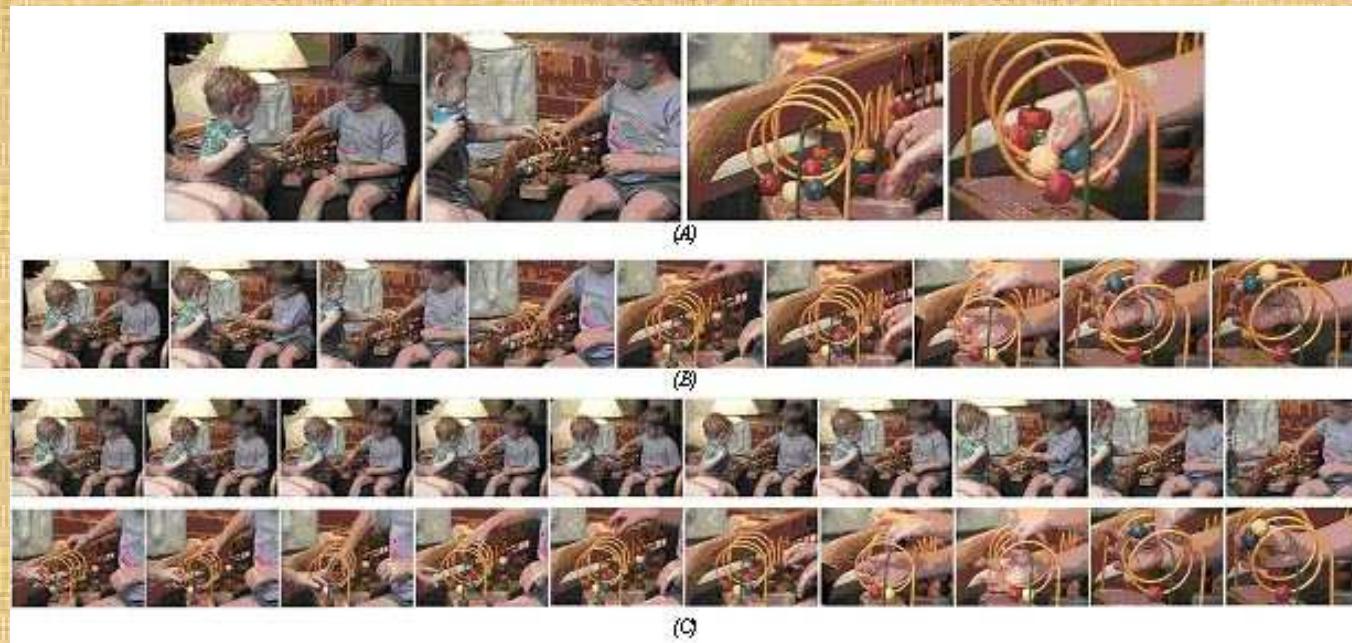


VOXEL (Volume Element)



Vídeo

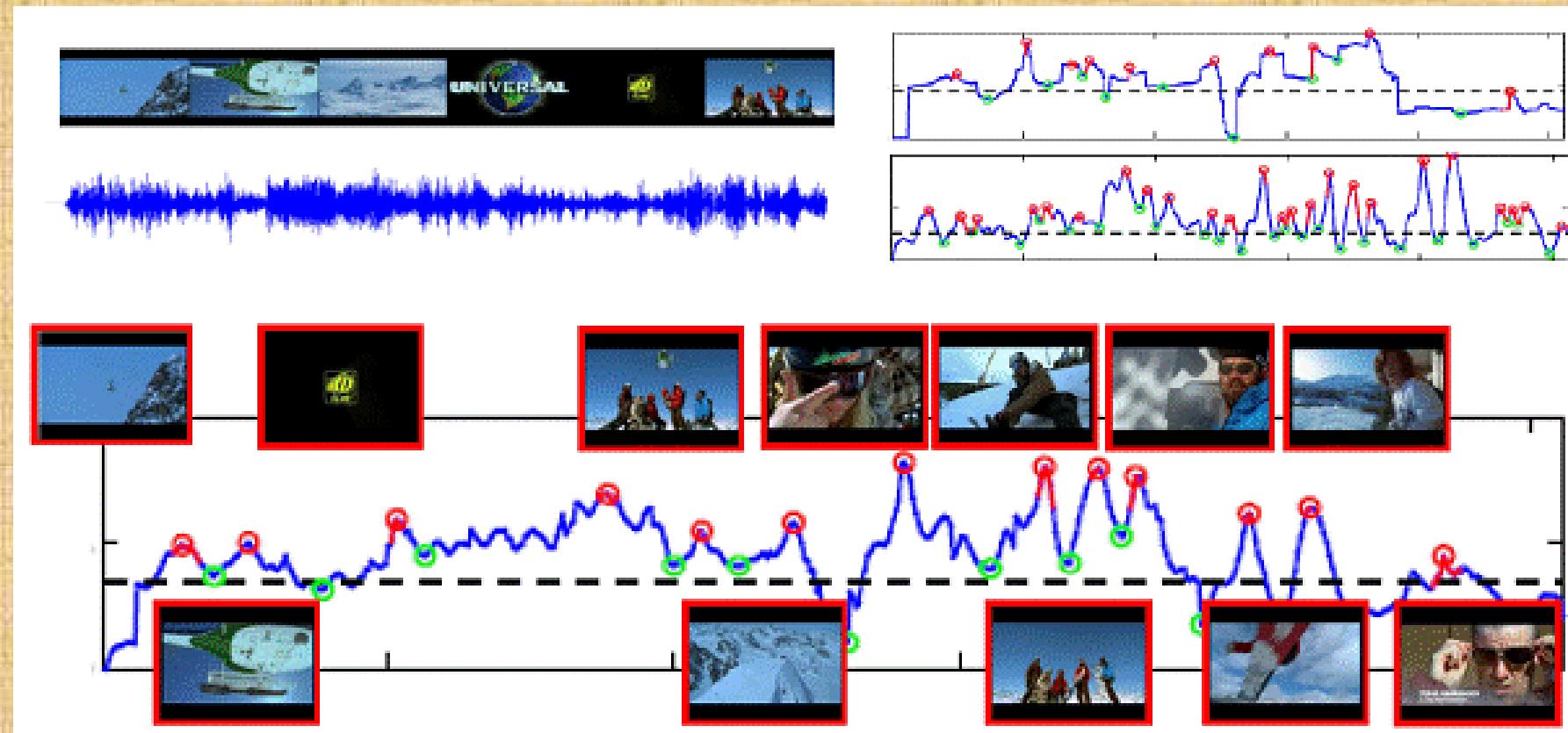
Un **vídeo** puede considerarse como una **sucesión de imágenes**, denominadas **frames**



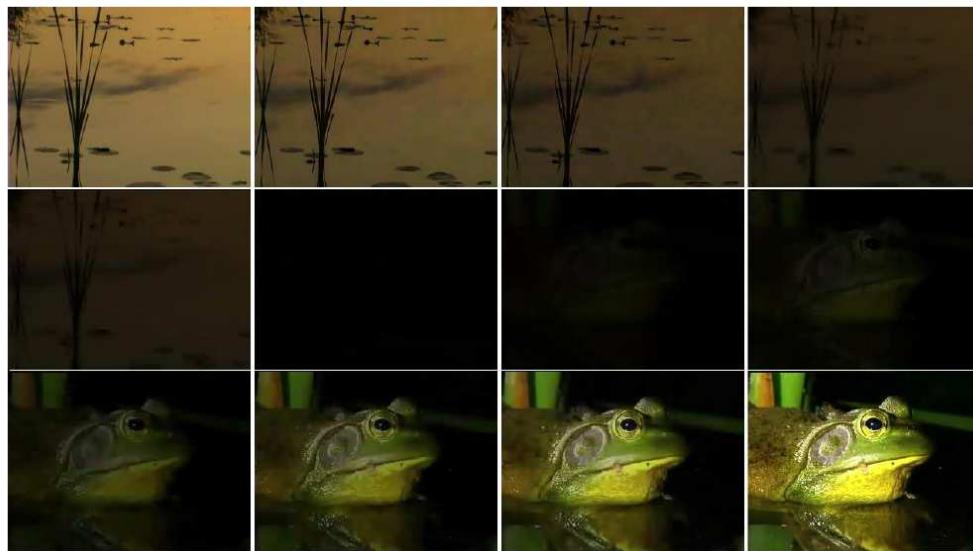
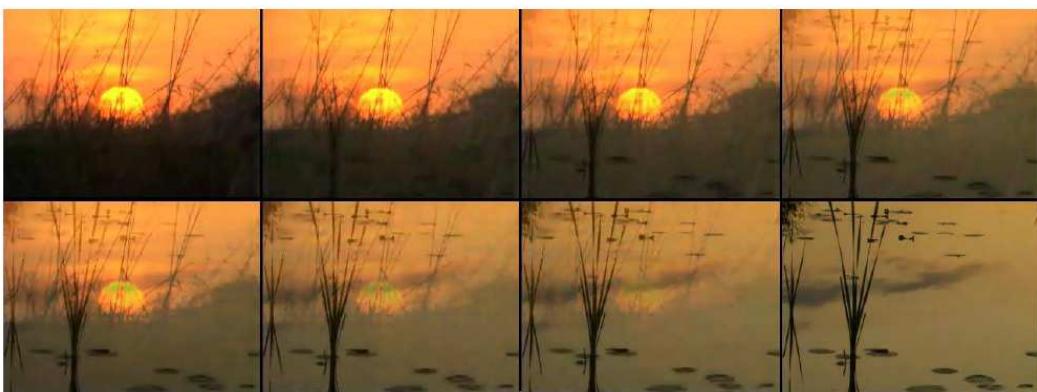
Keyframes: frames representativos de las diferentes escenas que forman un vídeo

Cambios de Escena

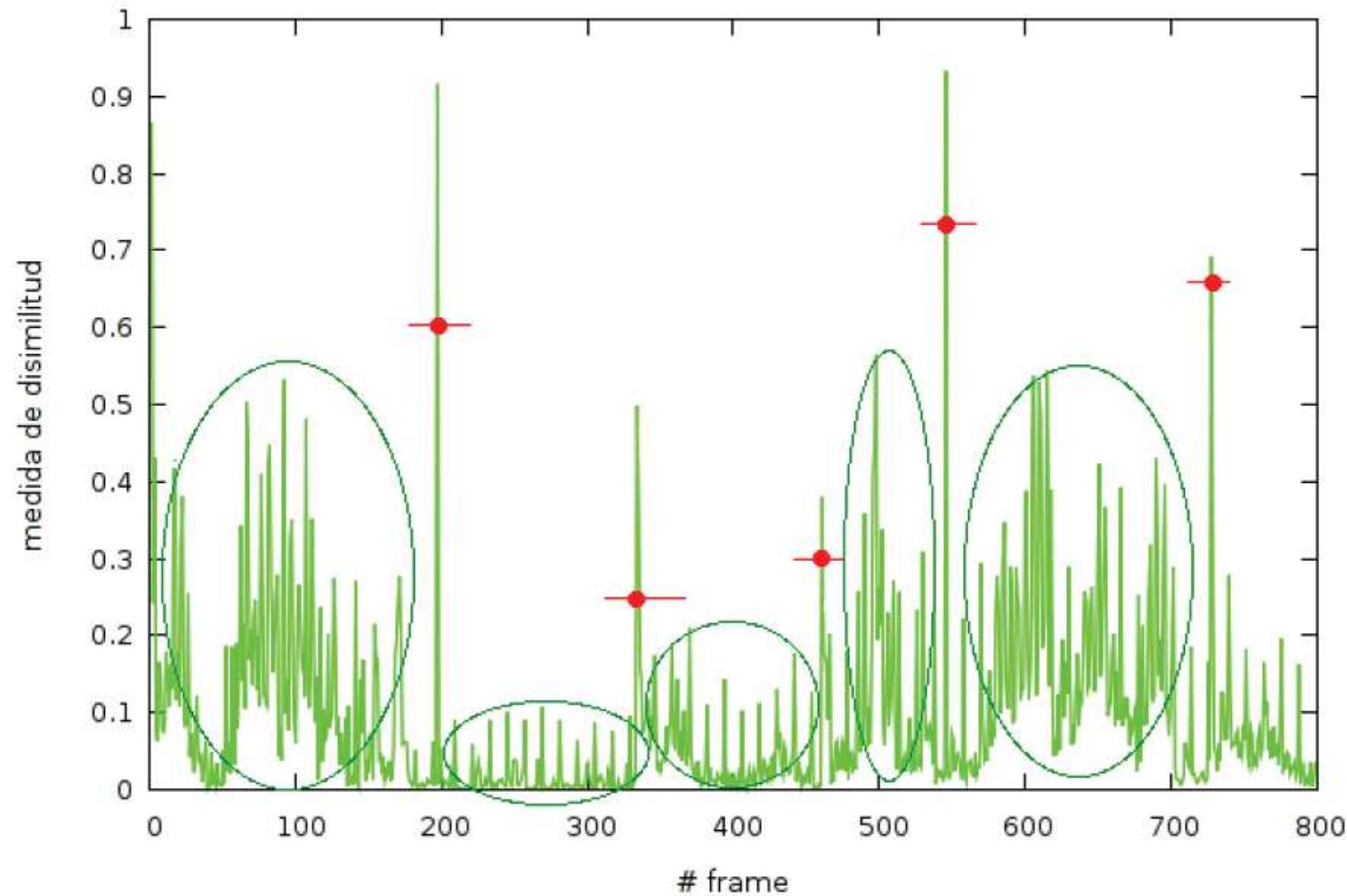
Un **vídeo** se puede **segmentar** localizando los **cambios de escena** mediante la evolución temporal de la medición de alguna propiedad de las imágenes que forman el vídeo



Tipos de Cambios de Escena



Tipos de Cambios de Escena



Medidas de Disimilaridad

Pixel a Pixel

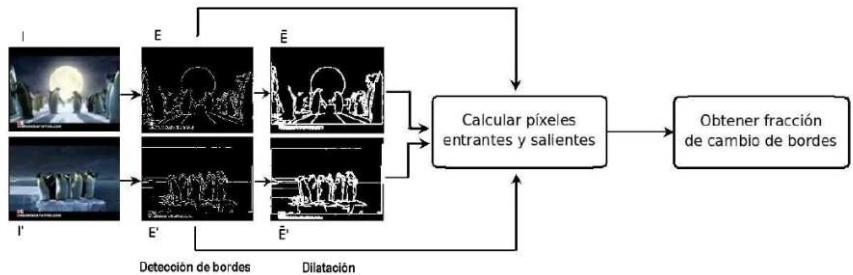


Disimilaridad basada en Bordes

Disimilaridad basada en Correlación



Distancia entre Histogramas



S.R.I.B.C. de Vídeo

**S.R.I.B.C.
de Vídeo**



S.R.I.B.C. de Imágenes
sobre los **keyframes**
de las secuencias del vídeo



Bibliografía

Bibliografía del Tema



Libros Castellano

Recuperación de Información. Un Enfoque Práctico y Multidisciplinar

Fidel Cacheda, Juan Manuel Fernández, Juan Francisco Huete; Ed. Ra-Ma; ISBN: 978-84-9964-112-6; 2011

Recuperación de Imágenes en Bases de Datos a partir del Color y la Forma

José Manuel Fuentes García; Tesis Doctoral; Universidad de Jaén; 1999



Libros Inglés

Visual Information Retrieval

Alberto del Bimbo; Morgan Kaufmann Publishers, Inc.; ISBN:
1-55860-624-6; 1999



Artículos

Ricerca per Contenuto di Immagini e Video

A. del Bimbo, E. Vicario, C. Colombo, P. Pala, J. Assfalg, S. Berreti y M. Bertini; Visual Information Processing Lab.; Dipartement. Sistemi e Informatica. Universitá di Firenze.

Similarity Search

A. del Bimbo, E. Vicario, C. Colombo, P. Pala, J. Assfalg, S. Berreti y M. Bertini; Visual Information Processing Lab.; Dipartement. Sistemi e Informatica. Universitá di Firenze.

