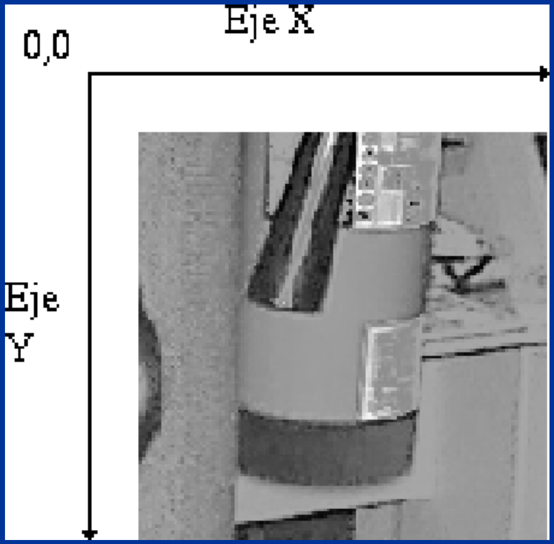
Visión de Alto Nivel

Dr. Luis Enrique Sucar

INAOE

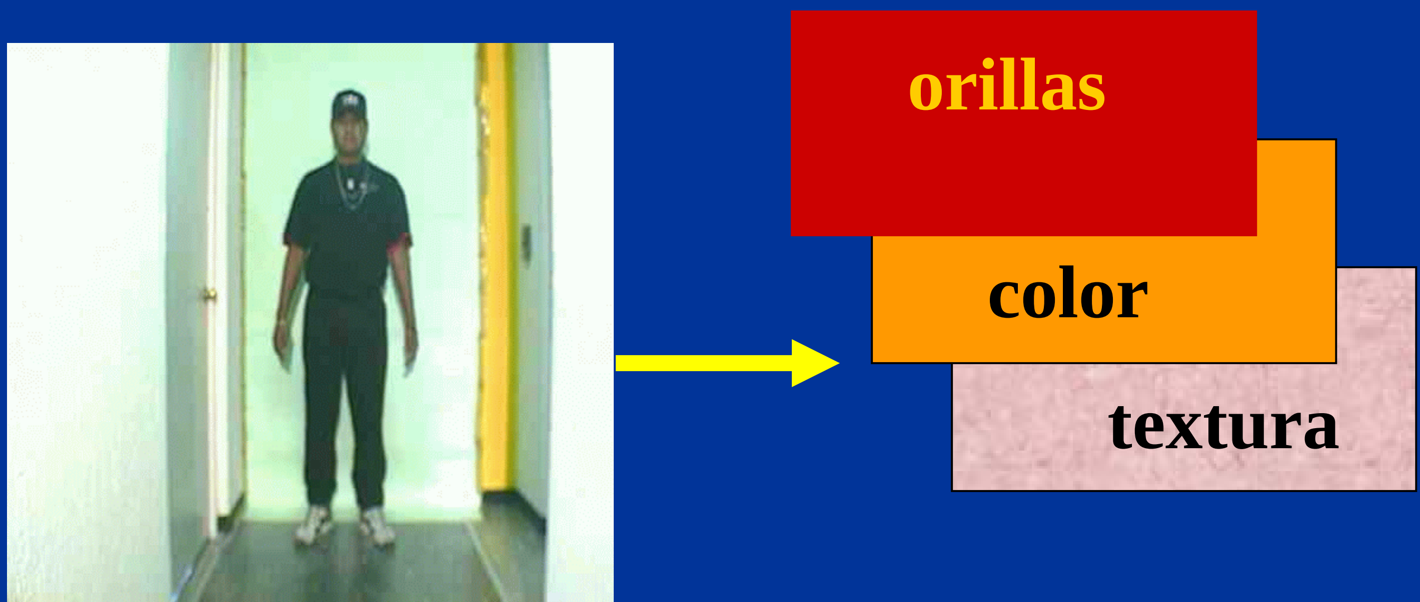
esucar@inaoep.mx
ccc.inaoep.mx/~esucar



Sesión 2
Bajo Nivel

Visión de Bajo Nivel

Extracción de Características



Orillas

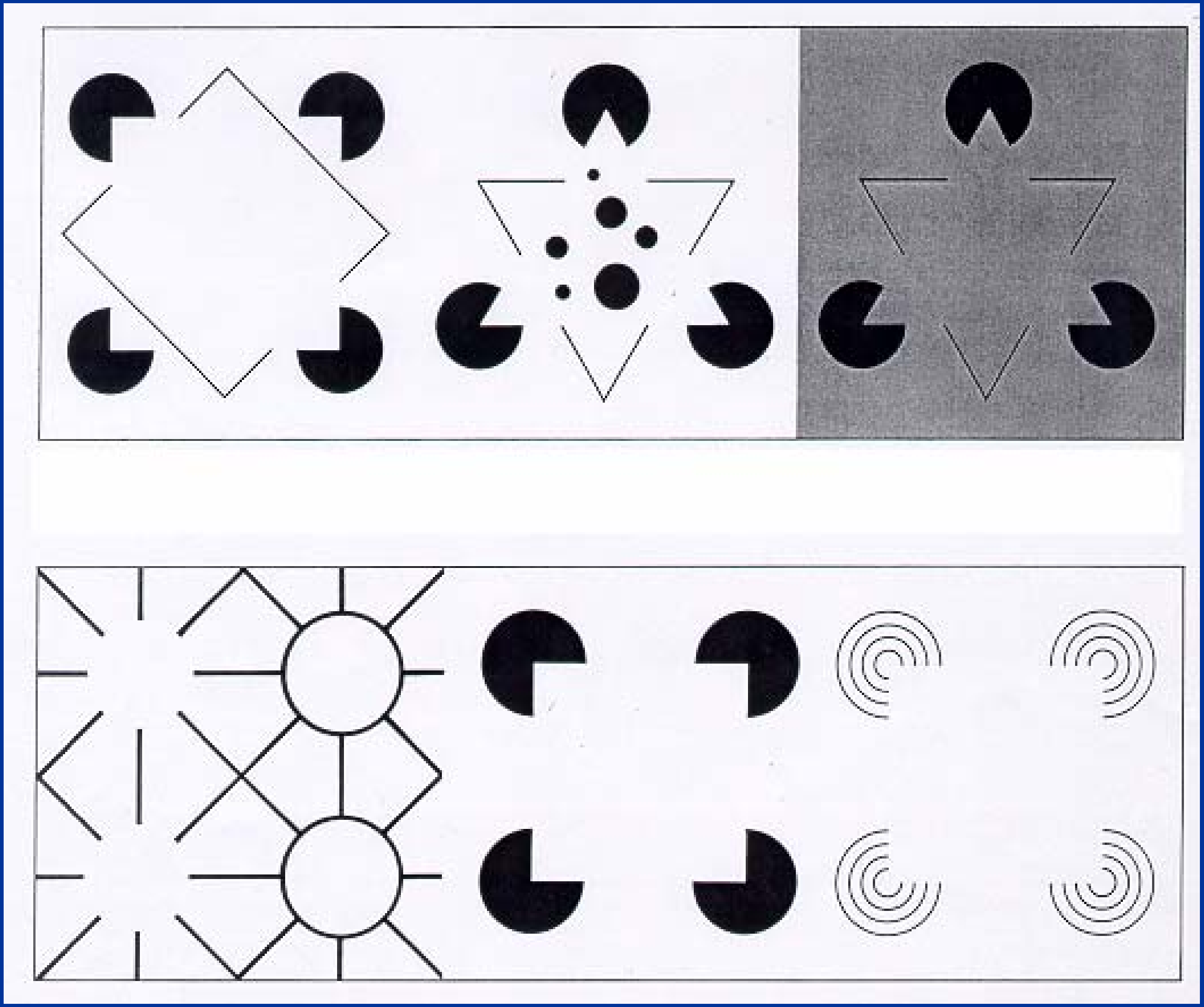
- Las variaciones de intensidad, *orillas* o *bordes* son muy importantes en visión.
- Muchas veces podemos reconocer un objeto sólo en base a su *silueta*
- Las orillas son una de las características más importantes y útiles para etapas posterior de visión como la segmentación y el reconocimiento de objetos.

Ejemplo de reconocimiento en base sólo a orillas o silueta

Contornos “subjetivos”

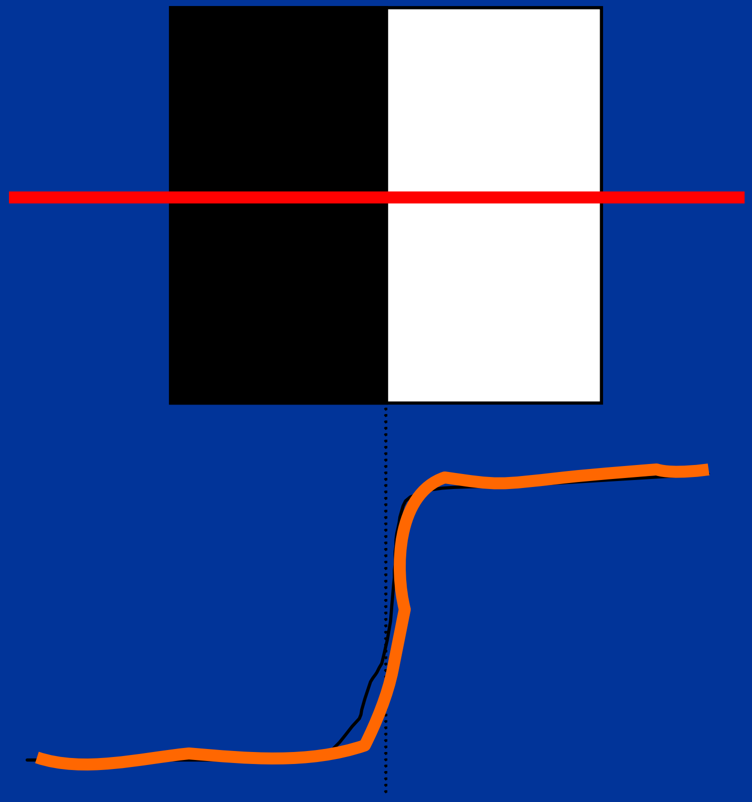
- De estudios en animales, se sabe que los sistemas de visión biológicos tienen “celdas” especializadas para detectar orillas
- Un ejemplo de esto, es que “vemos” contornos aún donde no los hay, completando bordes ocluidos o implícitos, como en los contornos subjetivos

Contornos subjetivos de Kanizsa



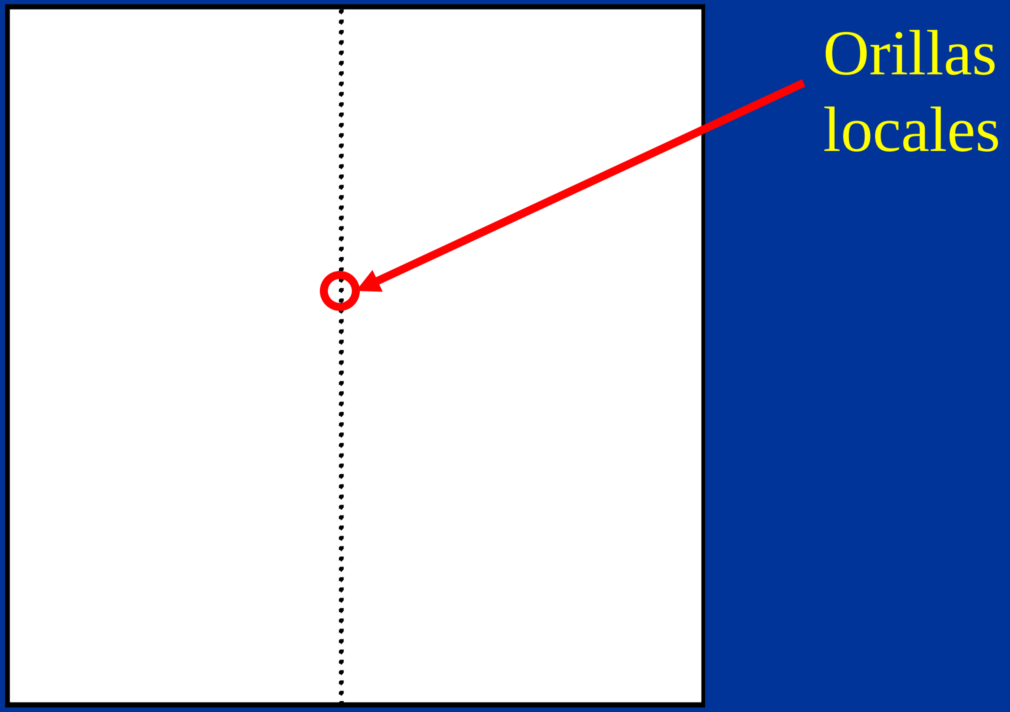
¿Qué es una orilla?

- Parte de la imagen en que hay un cambio *brusco* o discontinuidad en la intensidad de la imagen- derivada “alta”



Orillas locales

- Las orillas de una imagen normalmente se detectan como pequeños segmentos o secciones de un borde que se integran en etapas posteriores



Técnicas

- Técnicas de detección de orillas:
  - operadores de gradiente (primera derivada),
  - operadores de segunda derivada,
  - múltiples respuestas a diferentes orientaciones.

- Técnicas de post-procesamiento:
  - operador Canny
  - relajación.

Operadores de gradiente

- Se basan en diferenciar la imagen, es decir, encontrar el gradiente:

$$\Delta f = (df/dx,\ df/dy)$$

- Magnitud del gradiente:

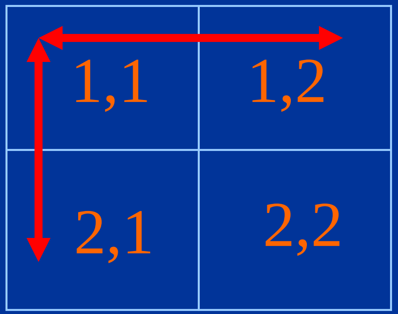
$$|\Delta f| = [\ (df/dx)^2 + (df/dy)^2\ ]^{1/2}$$

Aproximación al gradiente

- Se puede aproximar el gradiente tomando la diferencia de valores contiguos en la imagen.
- Para una sección de 2 x 2:

$$df/dx = I_{1,2} - I_{1,1}$$

$$df/dy = I_{2,1} - I_{1,1}$$

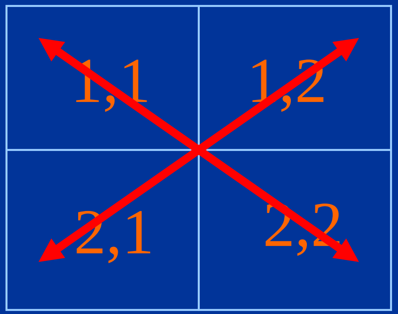


Aproximación al gradiente

- Otra opción es considerar las diferencias cruzadas:

$$df/dx = I_{1,1} - I_{2,2}$$

$$df/dy = I_{1,2} - I_{2,1}$$



Aproximación al gradiente

- Podemos también considerar una sección de 3 x 3 y aproximar el gradiente de la siguiente forma:

$$df/dx = (I_{1,3} + I_{2,3} + I_{3,3}) - (I_{1,1} + I_{2,1} + I_{3,1})$$

$$df/dy = (I_{3,1} + I_{3,2} + I_{3,3}) - (I_{1,1} + I_{1,2} + I_{1,3})$$

Las aproximaciones anteriores se pueden implementar como filtros espaciales (máscaras)

Operadores como el de <u>Roberts, Prewitt y Sobel</u>, se implementan con dos máscaras: una para *dx* y otra para *dy*

Operadores de Roberts

Corresponden a las diferencias cruzadas de 2 x 2

1	0
0	-1

0	1
-1	0

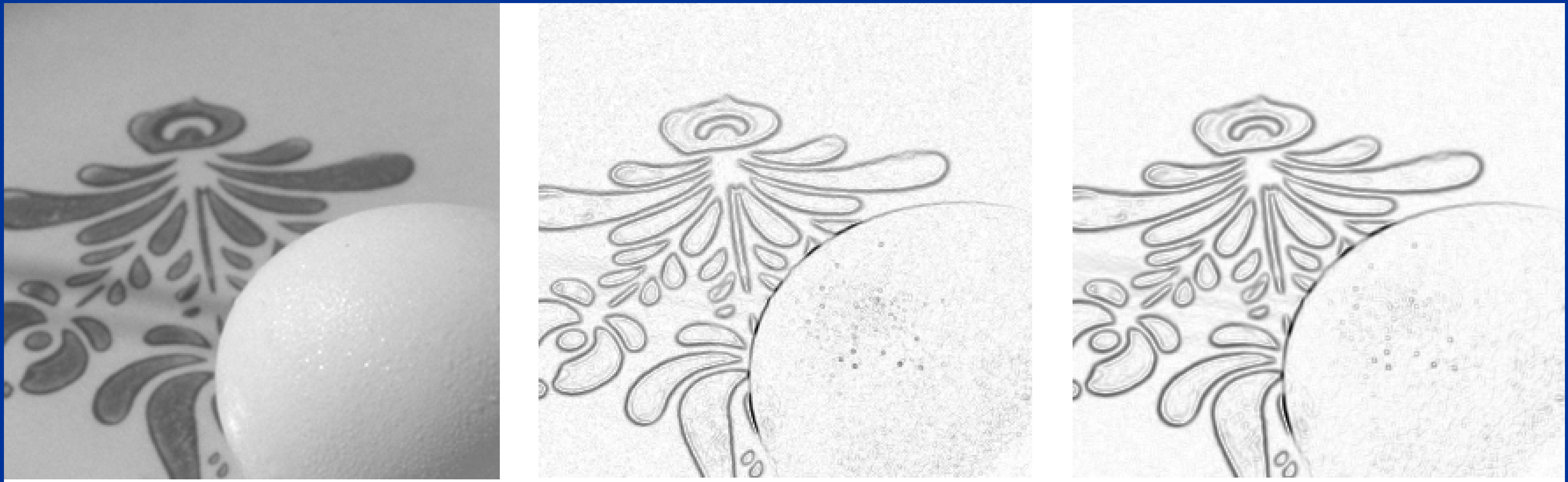
Operadores de Prewitt

Corresponden a las diferencias en secciones de 3 x 3

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

Ejemplos con los operadores de Roberts y Prewitt



(a) (b) (c)

Operadores de Sobel

- Los detectores de orillas como *Roberts* y Prewitt tienden a amplificar el ruido en la imagen
- Los operadores de *Sobel* reducen este efecto, al combinar en la misma máscara, la diferenciación con suavizamiento.

Operadores de Sobel

- El detector de Sobel se puede ver como la combinación de un filtro de suavizamiento con un operador de gradiente:

$$Sobel = D\,G^T$$

- Donde:

  D = [-1 0 1]

  G = [1 2 1]

Operadores de Sobel

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

Ejemplos con operadores de Sobel

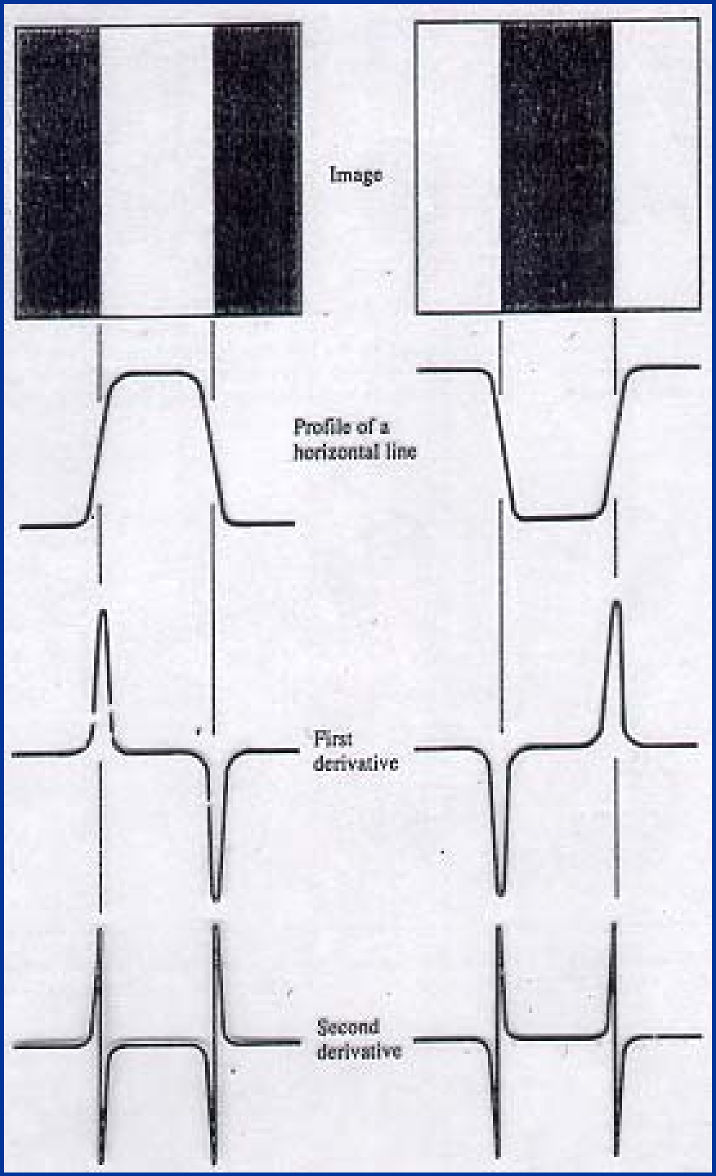
(a) (b)

(c) (d)

Operadores de segunda derivada

- Una ventaja de usar operadores de segunda derivada es que se puede estimar mejor la localización de la orilla - donde la segunda derivada cruza cero
- Ejemplos de estos detectores de orillas son:
  - Laplaciano
  - Laplaciano de una gaussiana (LOG)

Cruces por cero



Laplaciano

- Laplaciano de una función de 2 variables:

$$\Delta^2 f = (d^2f/dx^2,\ d^2f/dy^2)$$

- El cual se puede aproximar en forma discreta como:

$$\Delta^2 f = 4 * I_{2,2} - I_{1,2} - I_{2,1} - I_{2,3} - I_{3,2}$$

Máscara para el operador Laplaciano

0	-1	0
-1	4	-1
0	-1	0

Laplaciano de una Gaussiana

- En forma análoga al operador de Sobel, combina el efecto de una suavizamiento gaussiano con el Laplaciano en una sola máscara.
- El Laplaciano de una Gaussiana (LOG) es:

$$\Delta^2 G = (d^2G/dx^2,\ d^2G/dy^2)$$

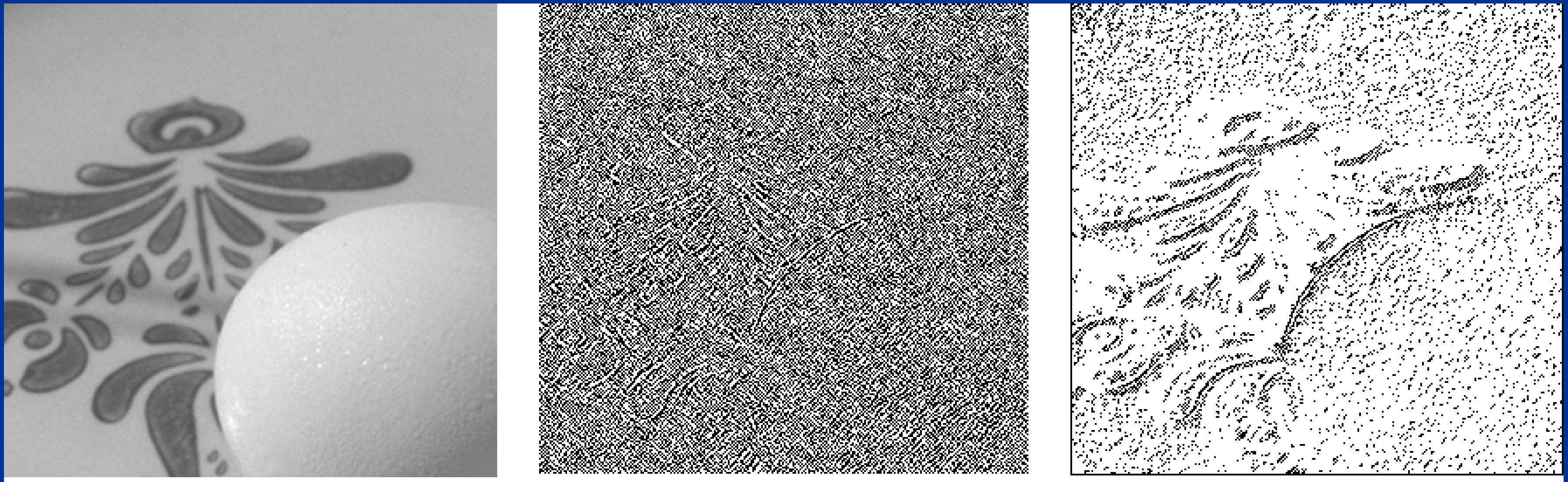
- El cual también se puede aproximar con una máscara.

Máscara para el operador Laplaciano de una Gaussiana

1	-2	1
-2	4	-2
1	-2	1

Ejemplos con operador LOG

Original | LOG (3x3) | Supresión dobles

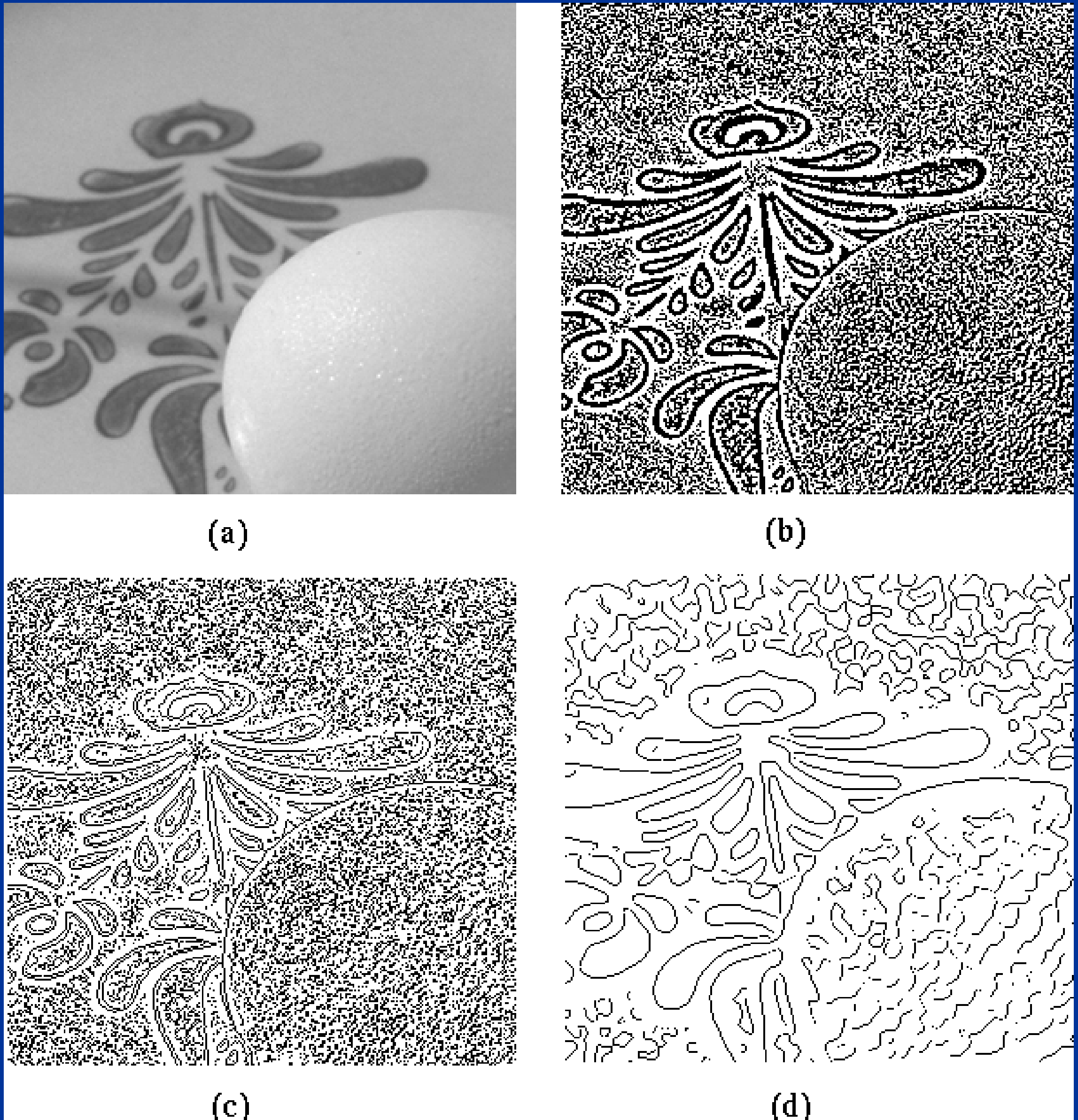


(a) (b) (c)

Otros ejemplos con operador LOG:

implementación como diferencia de 2 gaussianas

original



0.5, 0.8

supresión dobles

2.5, 4.0 y supresión

Operadores direccionales

- En general es útil conocer no sólo la magnitud de las orillas sino también su dirección:

$$\phi f = \tan^{-1} \frac{df/dy}{df/dx}$$

- Esto se puede obtener con los operadores de de Prewitt y Sobel, así como con otros operadores direccionales más sofisticados

Operadores de Kirsch

- Detectan la máxima respuesta en direcciones espaciadas 45°, es decir en orientaciones de 0, 45, 90 y 135 grados - 4 máscaras
- Se pueden definir a diferentes tamaños: 2x2, 3x3, 5x5

Máscaras para operadores de Kirsch de 3x3

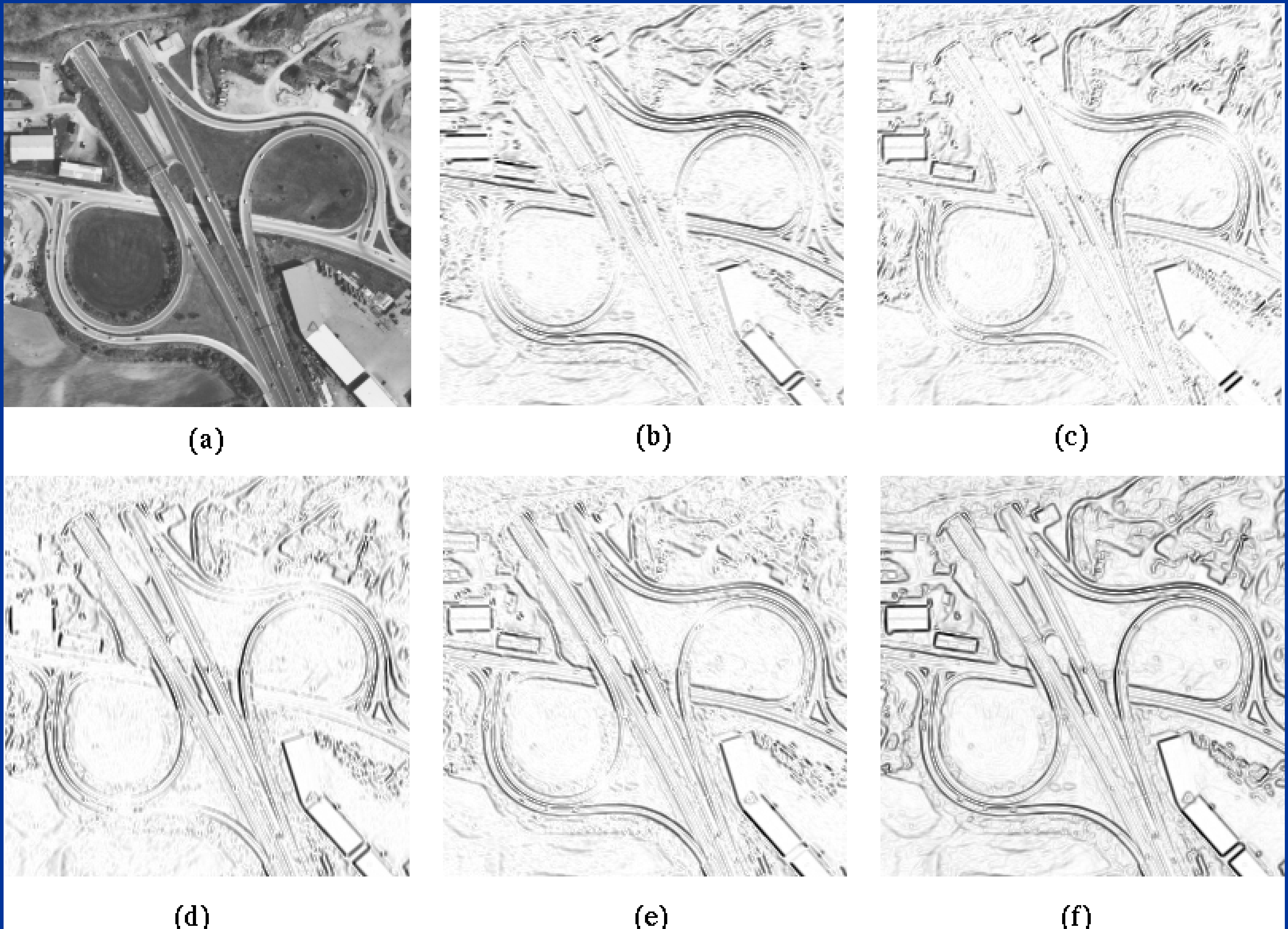
-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

-1	-1	0
-1	0	1
0	1	1

0	1	1
-1	0	1
-1	-1	0

Ejemplo de aplicación de operadores de Kirsch



(a) (b) (c)

(d) (e) (f)

Color

- El color es importante porque:
  - ayuda a la extracción de características,
  - apoya los niveles superiores como segmentación y reconocimiento.
- El ojo humano distingue miles de colores y en cambio sólo aprox. 20 niveles de gris - el color es importante en el reconocimiento visual.

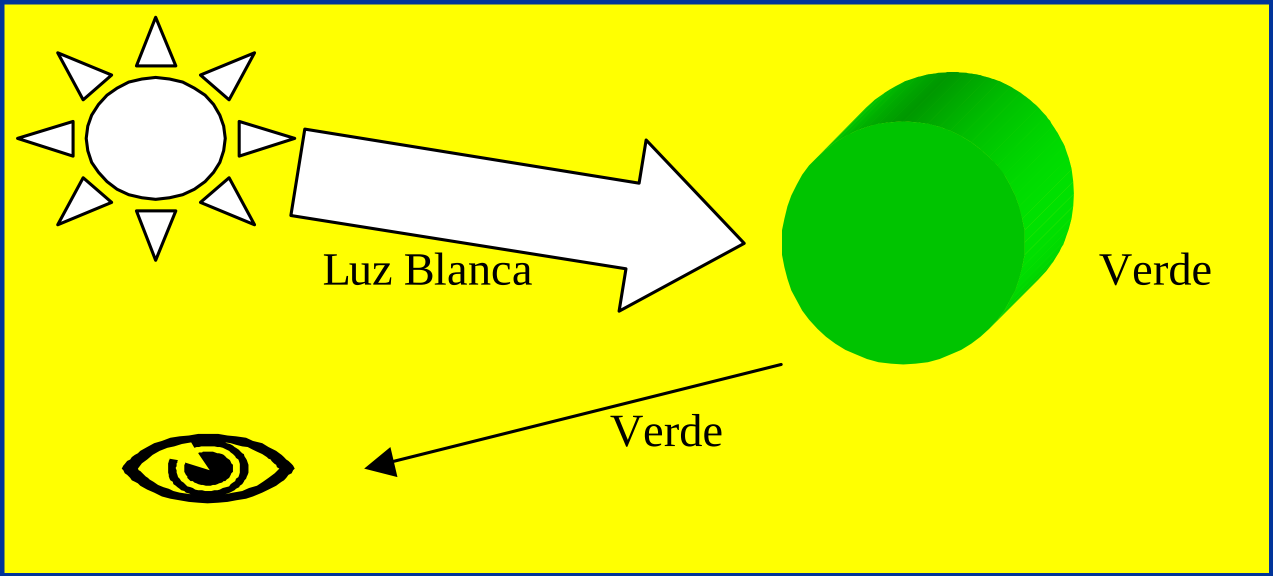
Espectro

- El color tiene que ver con la longitud de onda dentro de la luz visible del espectro electromagnético

color:	violeta azul verde amarillo naranja rojo	
longitud de onda (nm)	400	700

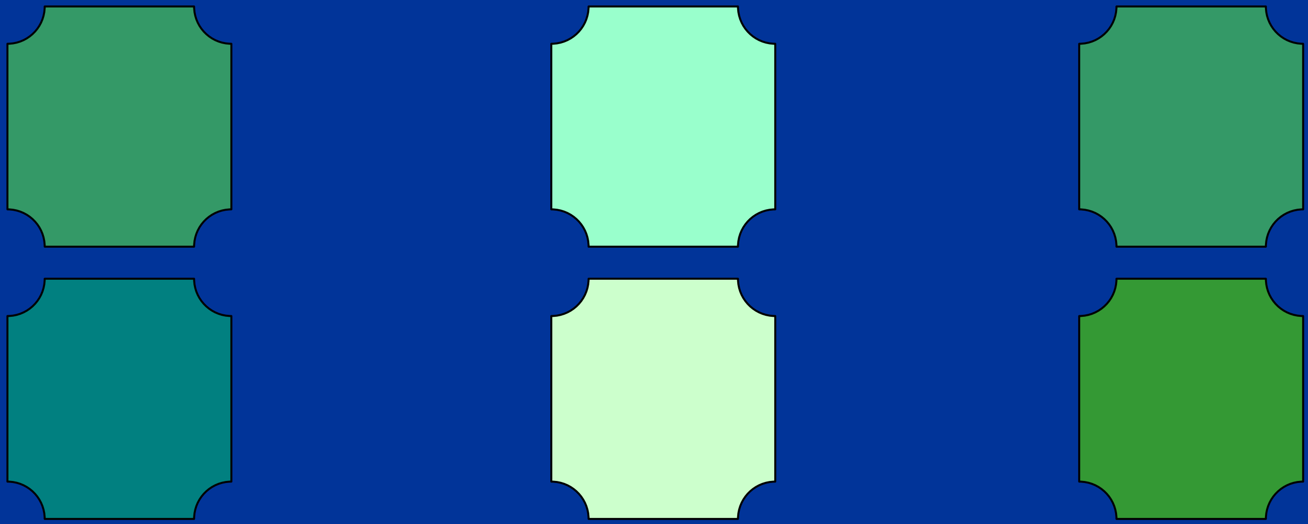
Percepción del Color

- Un objeto se ve de cierto color si refleja las longitudes de onda de dicho color (por ejemplo, verde: 500-570 nm) y absorbe el resto, al ser iluminado por luz "blanca"



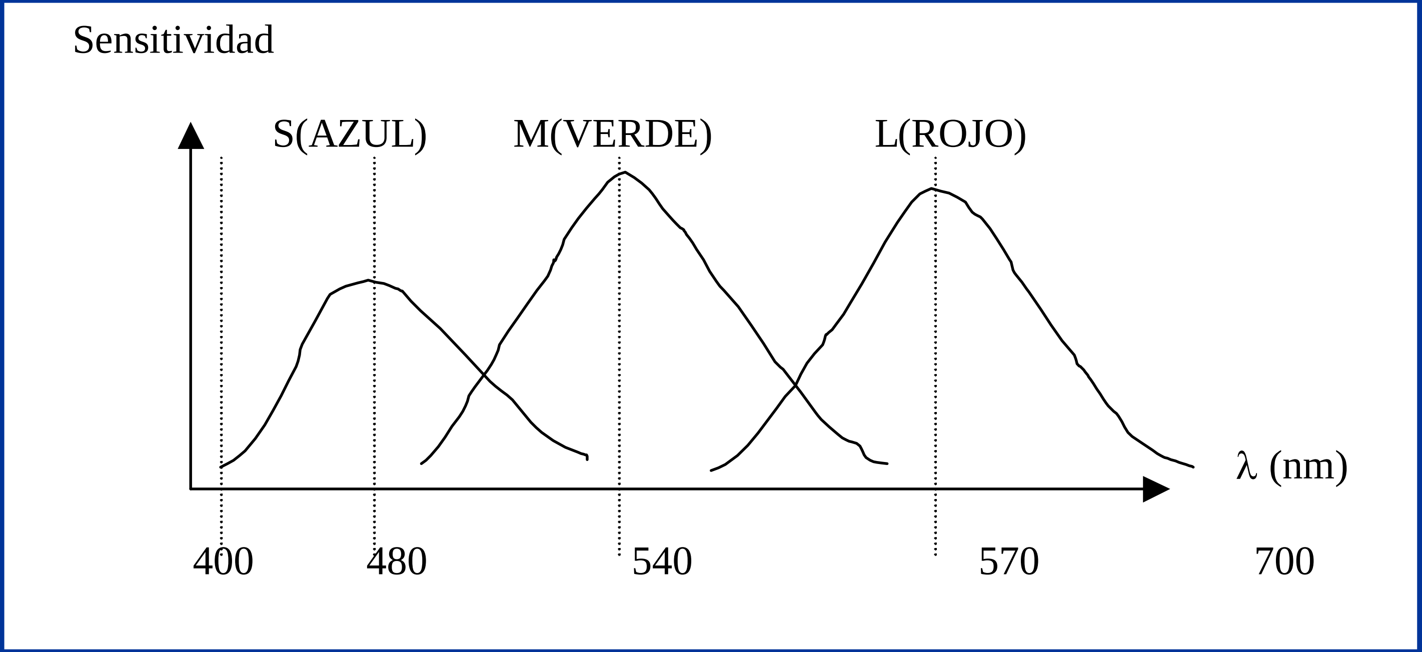
Atributos básicos del color

- Croma o longitud de onda dominante (Hue)
- Pureza o saturación
- Brillantez o intensidad



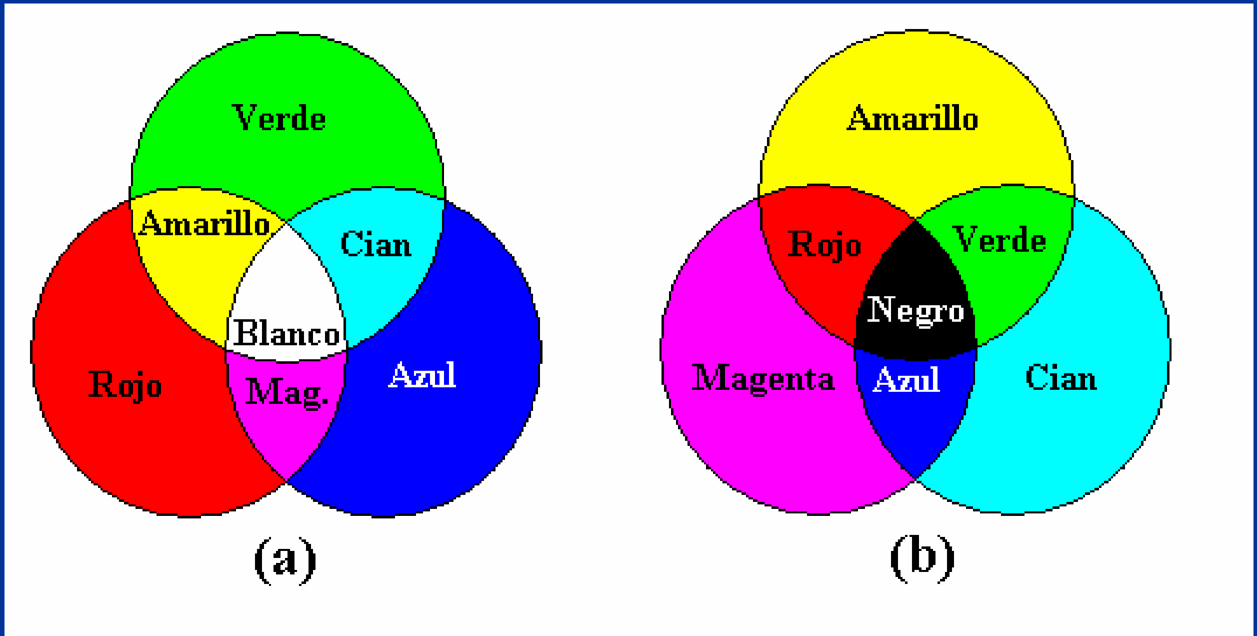
Percepción humana del color

- Percibimos el color mediante sensores (conos) que son de diferentes tipos - mayor sensibilidad a diferentes longitudes de onda: α (azul), β (verde), γ (rojo)

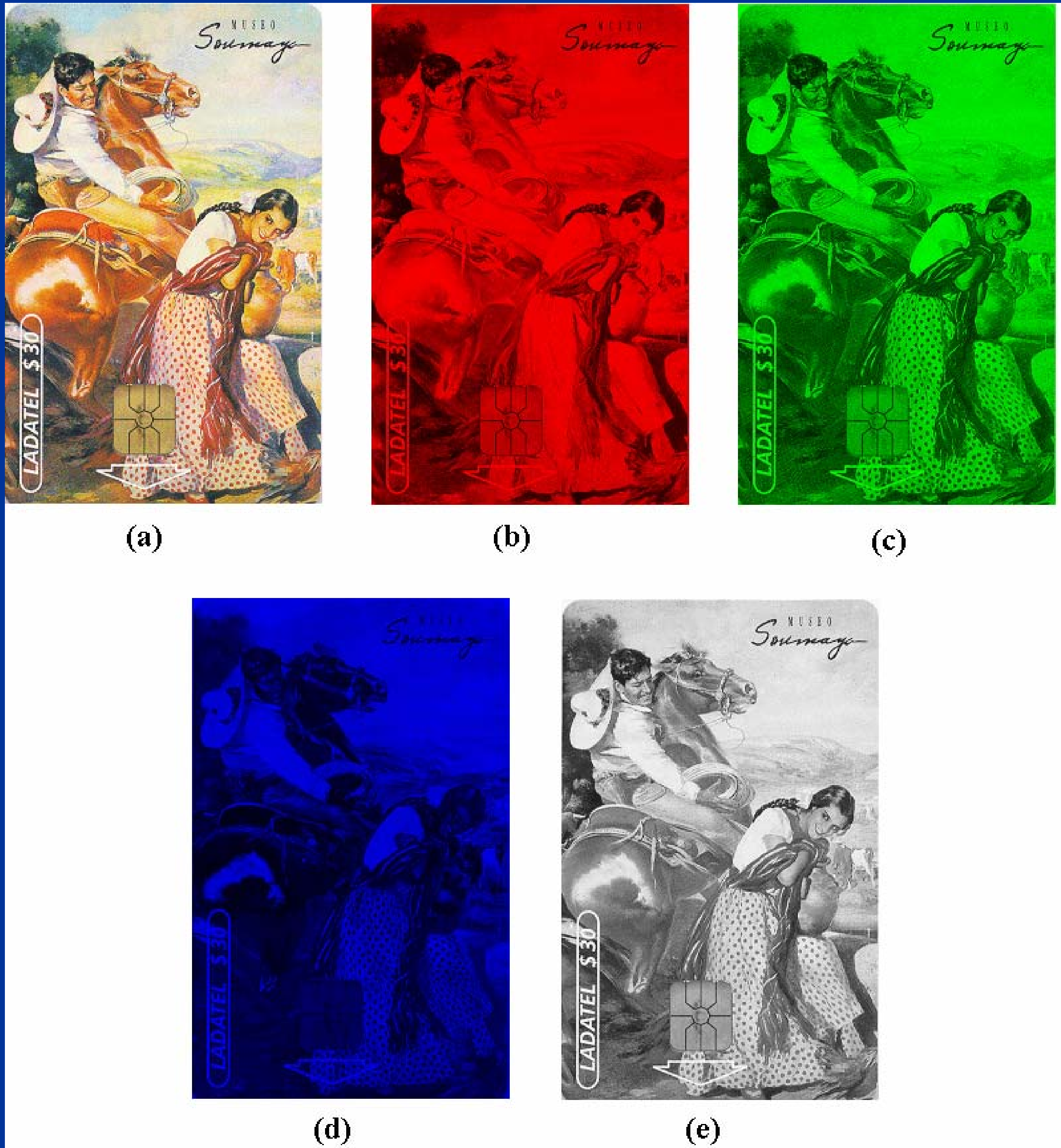


Combinación de colores

- La identificación del color se hace mediante la combinación de los 3 tipos de sensores - combinación de colores “primarios” (RGB)
- Combinación aditiva - luz (a) y substractiva - pigmentos (b)

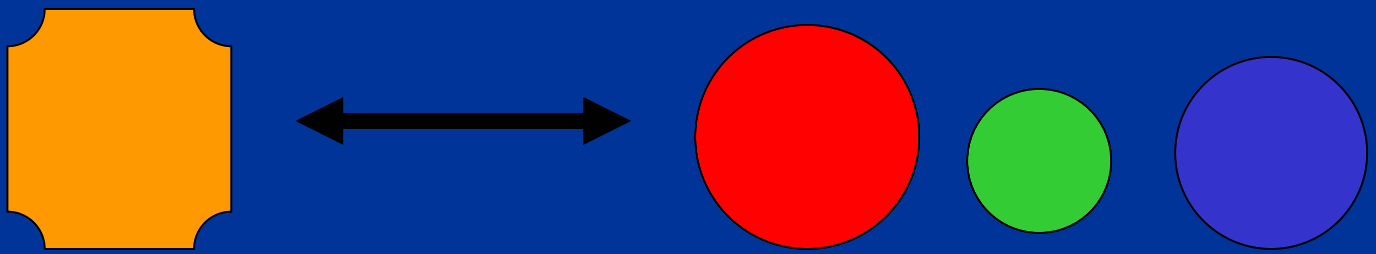


Componentes de una imagen a color



Sistema CIE

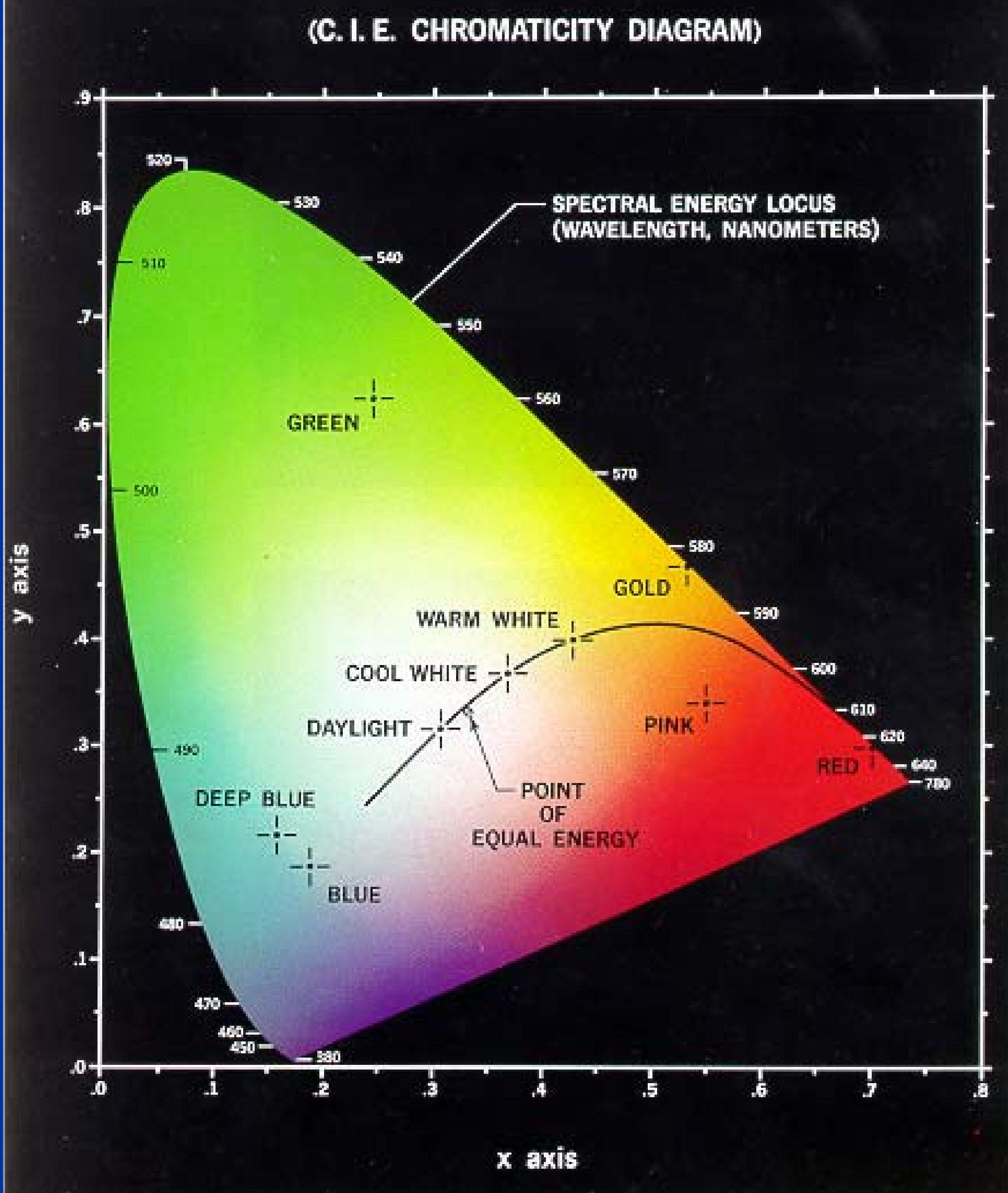
- Generación de colores a partir de combinación de los 3 primarios
- “*Color matching*” - persona compara luz de cierto color vs. combinación de luces en colores primarios (psico-fisiológico)



- Se realizó para todos los colores del espectro visible
- Ciertos colores no se lograban igualar!

Diagrama Cromático CIE

azul - 435.8 nm
verde - 541.6 nm
rojo - 700 nm



Coordenadas Cromáticas

- Se normalizan los valores de R,G,B de forma de que sumen 1:

$$r = R/(R + G + B)$$

$$g = G/(R + G + B)$$

$$b = B/(R + G + B)$$

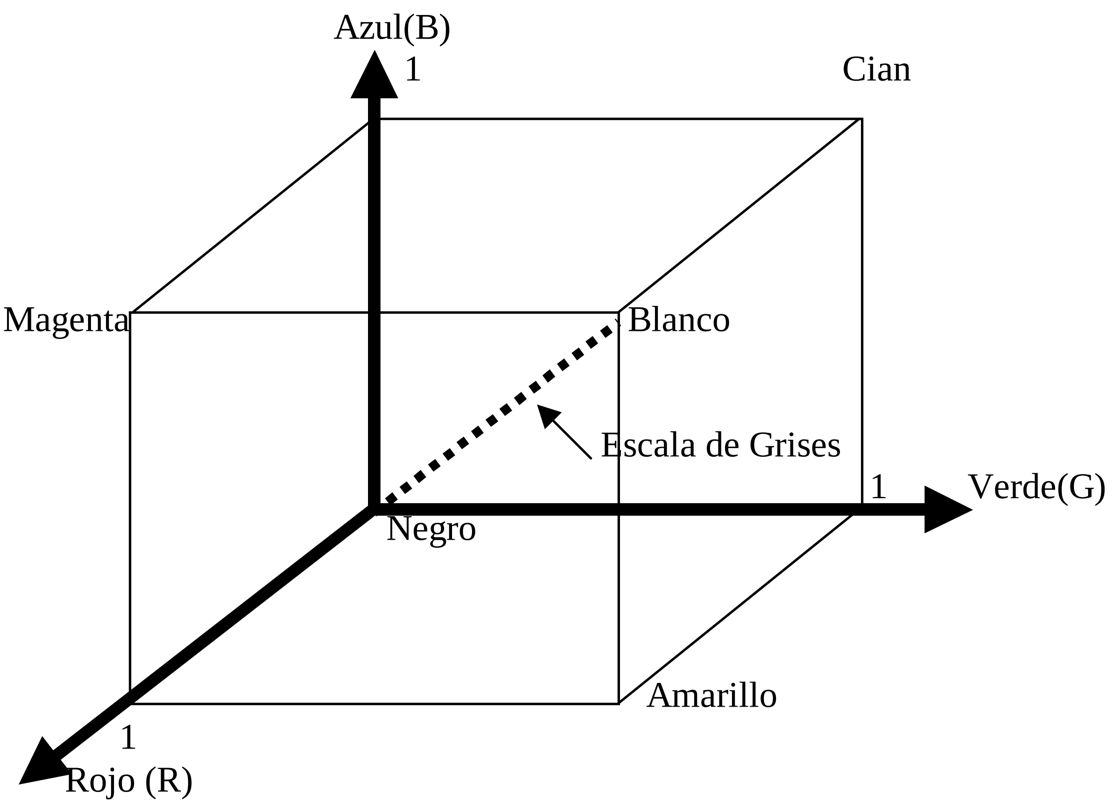
- De esta forma el espacio de colores lo podemos representar en dos dimensiones (por ejemplo, r - g)

Modelos de Color

- Diferentes formas de representar el color:
  - modelos sensoriales - orientados a los equipos
    - RGB, CMY, YIQ
  - modelos perceptuales - se asemejan a la percepción humana y se orientan al procesamiento de imágenes y visión
    - HSV, HLS, HSI

Modelo RGB

- Se basa en los componentes primarios: RGB
- Se puede representar como un “cubo” con un primario en cada eje



Modelo CMY

- Se basa en los componentes secundarios: Cian, Magenta y Amarillo (Y)
- Se puede obtener del RGB (normalizado) como:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- En la práctica se agrega el negro (CMYK) para facilitar la impresión de negro

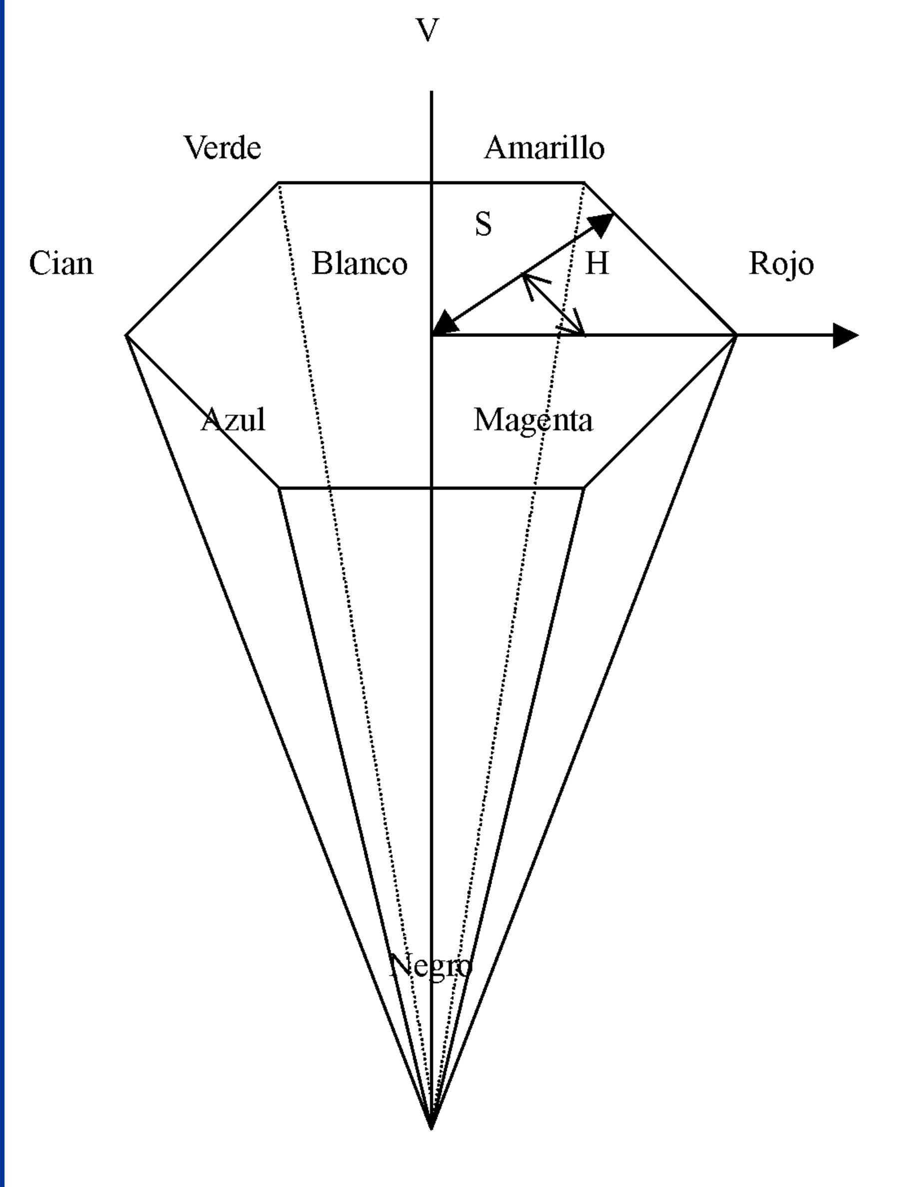
Modelo YIQ

- Separa la información de intensidad (Y) de la información de color (I,Q)
- Se obtiene mediante la siguiente transformación a partir del RGB:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.231 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Modelo HSV

- Modelo perceptual en que el color se representa en base a croma (H), saturación e intensidad o valor (V)
- Se obtiene "deformando" el cubo RGB de forma que se convierte en una pirámide hexagonal invertida



Conversión RGB a HSV

$V = M$

Si $m=M$ $S=0$ sino $S=(M-m)/M$

Si $m=B$ $H=120\,(G-m)/(R+G-2m)$

Si $m=R$ $H=120\,(B-m)/(B+G-2m)$

Si $m=G$ $H=120\,(R-m)/(R+B-2m)$

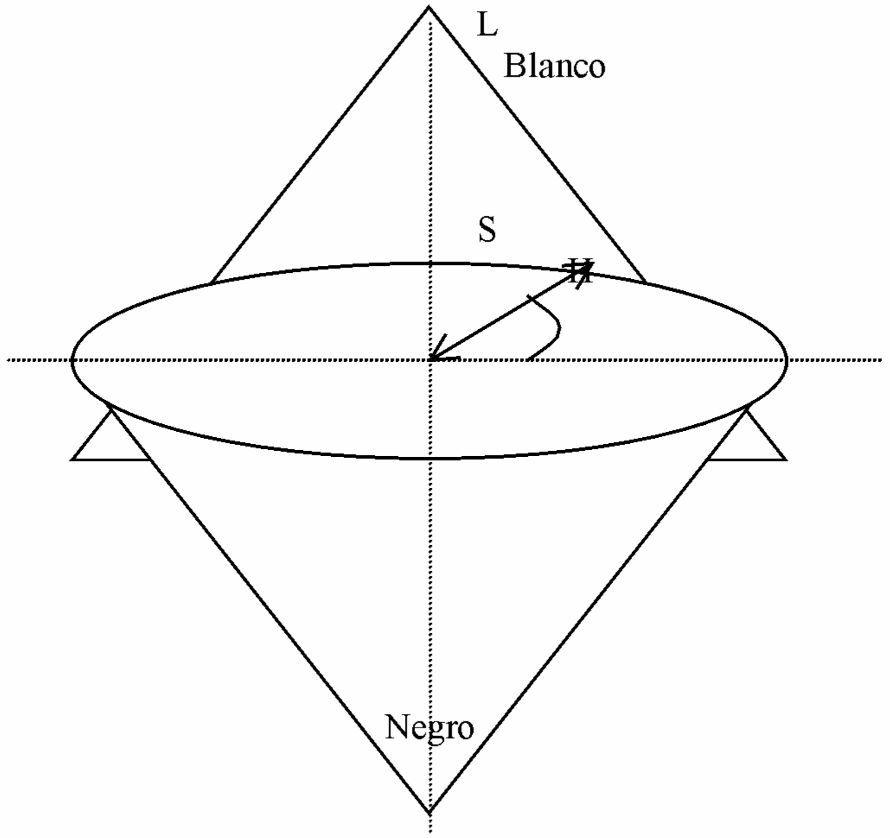
$m = \min(R,G,B)$

$M = \max(R,G,B)$

$V, S: [0, 1]$ $H: [0\ 360]$

Modelo HLS

- Modelo perceptual: *Hue, Level, Saturation*
- Se basa en coordenadas polares en 3-D, obteniéndose un espacio en forma de 2 conos unidos en la base



Conversión RGB a HLS

$L = (M + m)/2$

$S = (M + m)/(M - m)$, si $L < 0.5$

$S = (M - m)/(2 - M - m)$, si $L > 0.5$

Si $m=B$ $\quad$ $H=120\ (G-m)/(R+G-2m)$

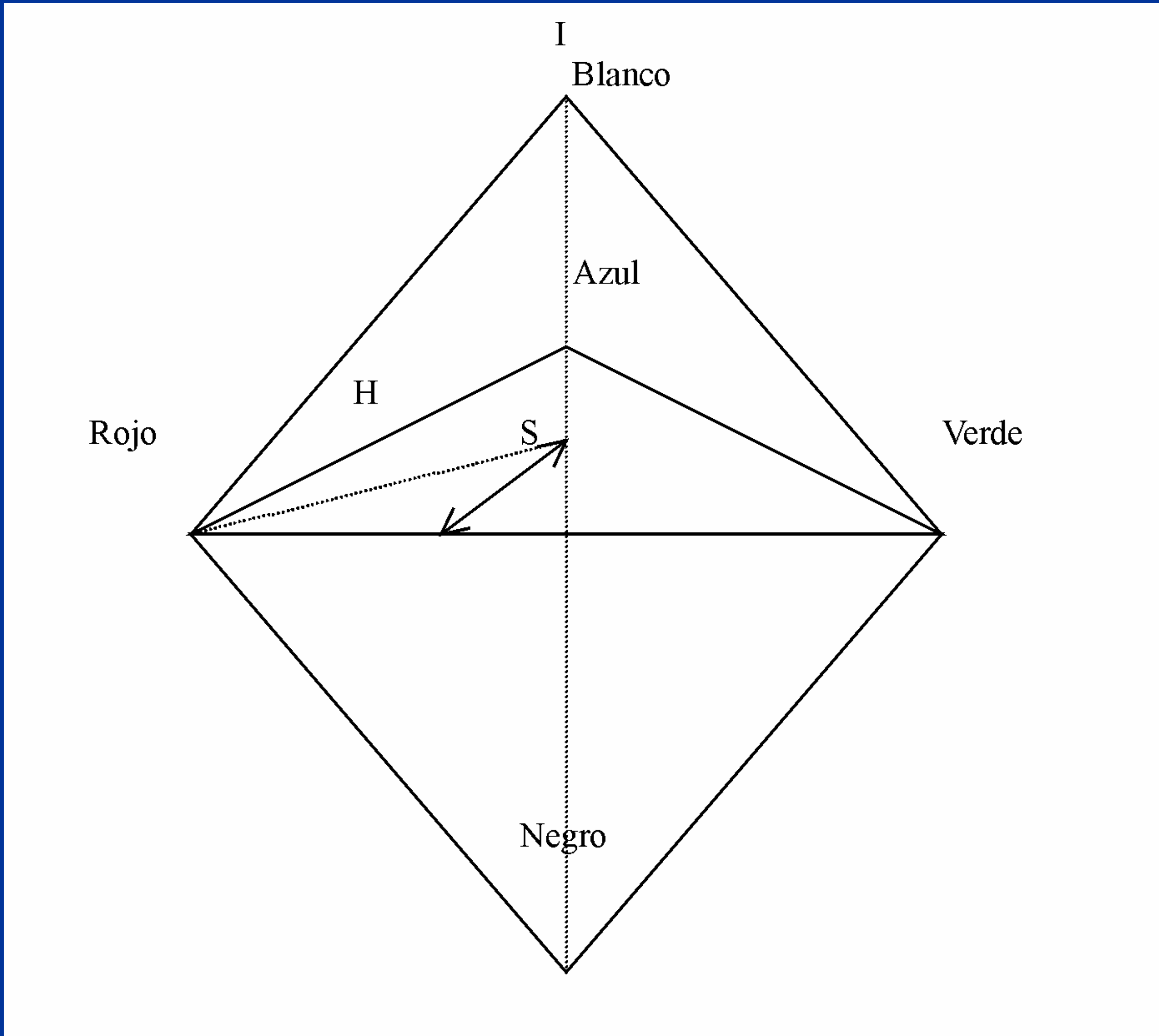
Si $m=R$ $\quad$ $H=120\ (B-m)/(B+G-2m)$

Si $m=G$ $\quad$ $H=120\ (R-m)/(R+B-2m)$

$m = \min(R,G,B)$, $\quad$ $M = \max(R,G,B)$

Modelo HSI

- Transformación del espacio RGB al espacio perceptual
- Forma de 2 pirámides triangulares unidas en su base



Conversión RGB a HSI

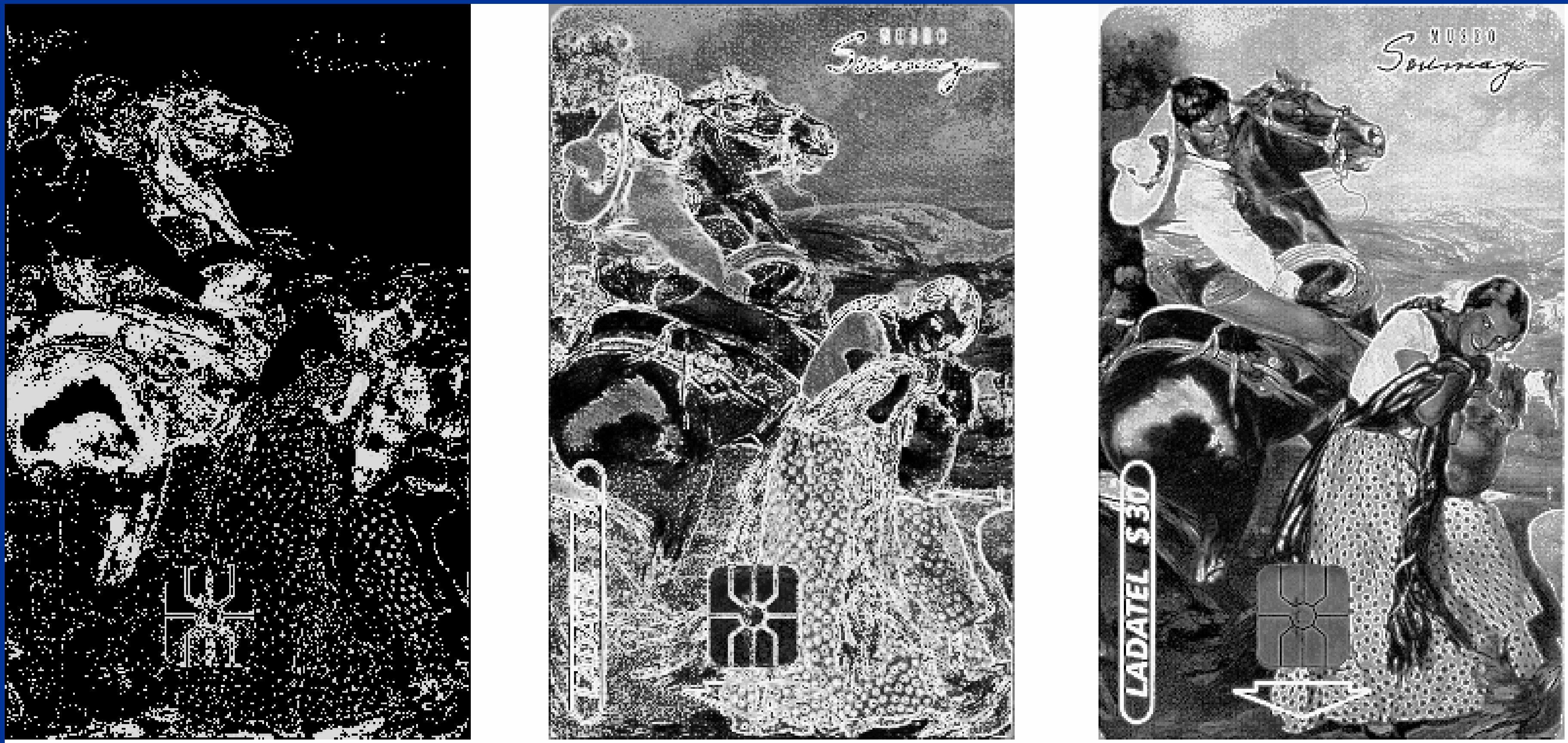
$$I = 1/3\ (R + G + B)$$

$$S = 1 - 3\,m / (R + G + B)$$

$$H = \cos^{-1}\left[\frac{\tfrac{1}{2}\,[(R - G) + (R - B)]}{[\,(R-G)^2 + (R-B)\,(G-B)]^{1/2}}\right]$$

$$m = \min(R,G,B)$$

Ejemplo de Imagen en HSI



(a) (b) (c)

Textura

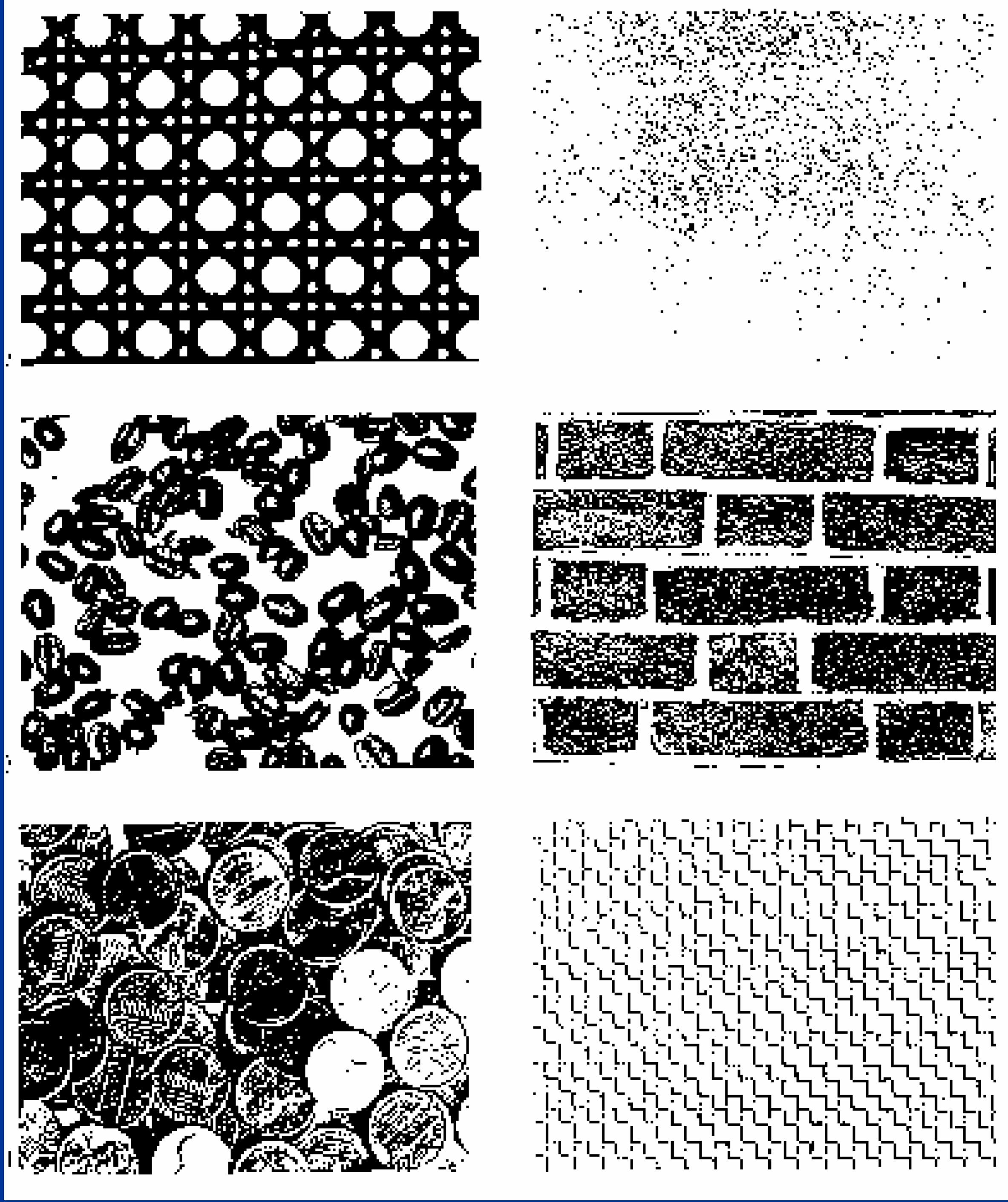
- Muchos objetos no son uniformes, tienen cierta “textura”

Textura:

“compuesta de pequeños elementos indistinguibles y entrelazados”

- La textura depende de la resolución
- La información de textura se puede usar para segmentación, reconocimiento de objetos y obtención de forma

Ejemplos de texturas



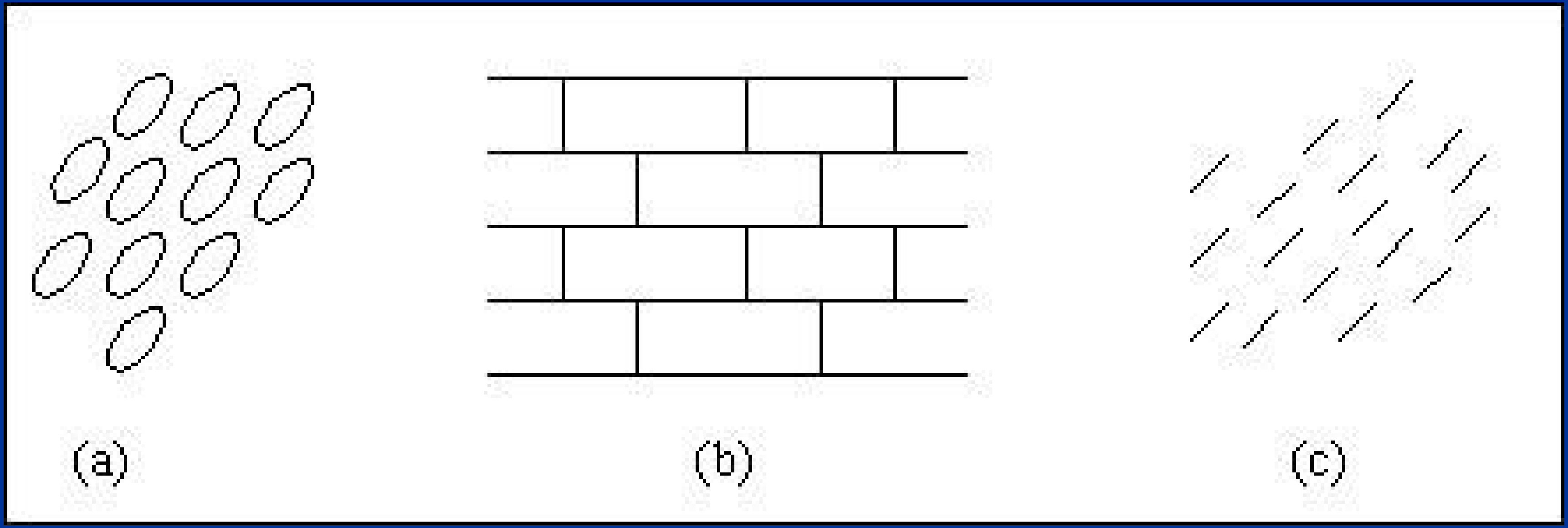
Descripción de texturas

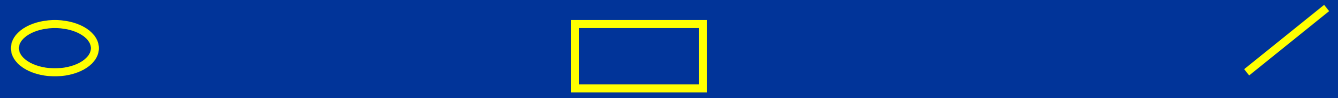
- Existen diferentes formas de describir una textura:
  - Modelos estructurales
  - Modelos estadísticos
  - Modelos espectrales

Texels

- Elementos básicos o primitivas de textura
- *Texel*: "primitiva visual con propiedades invariantes que ocurre a diferentes posiciones, deformaciones y orientaciones en un área"
- Propiedades invariantes: *forma, tamaño, nivel de gris, color.*

Ejemplos de *texels*





Tipos de texturas

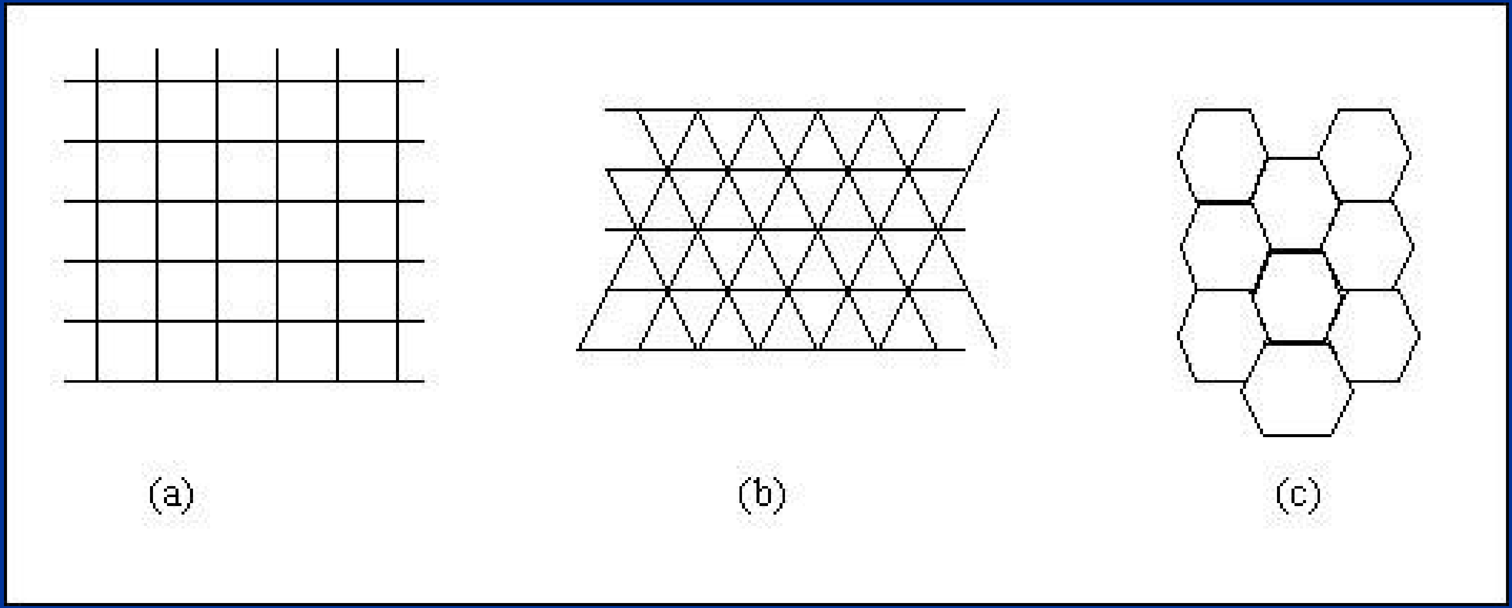
- Jerárquicas (ladrillos) vs. fractales (arbustos)
- Bidimensionales vs. tridimensionales
- Regulares vs. estadísticas

- Dependiendo del tipo de textura se aplican diferentes modelos para su descripción

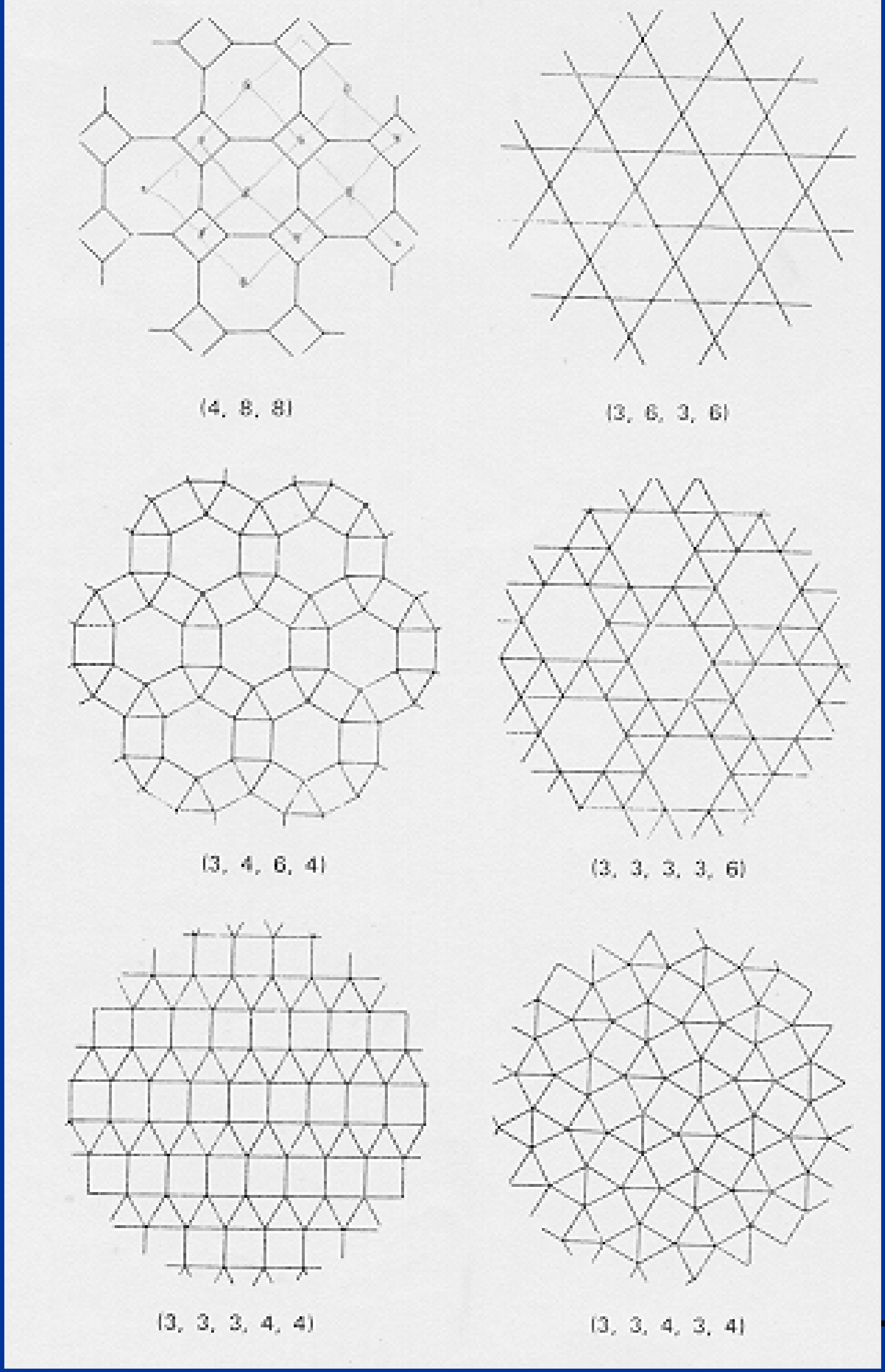
Modelos Estructurales

- Se aplican a texturas altamente regulares
- Se pueden describir en base a formas básicas que se repiten uniformemente
- Texturas básicas:
  – regulares
  – semi-regulares

Texturas Regulares



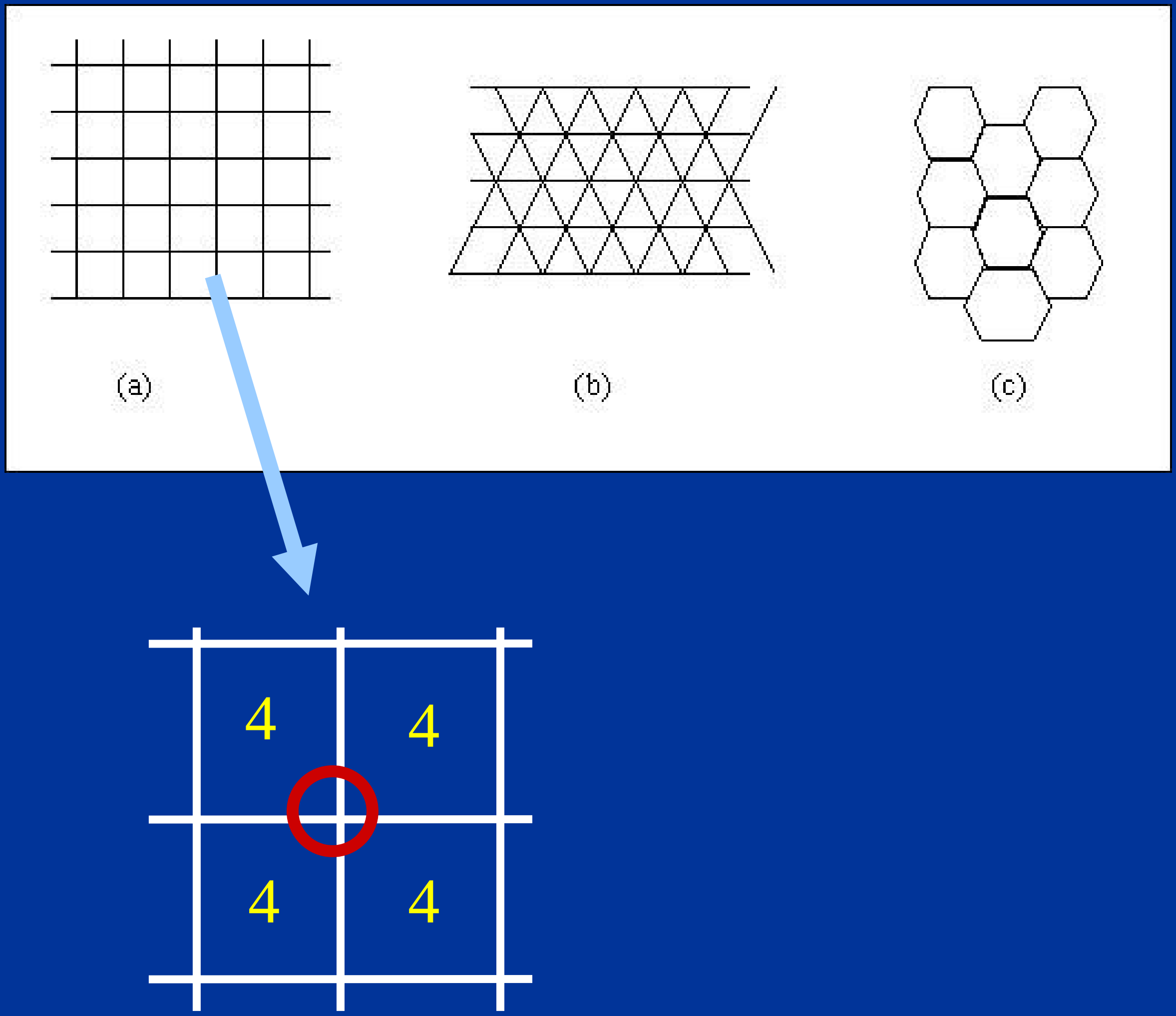
Texturas Semi-regulares



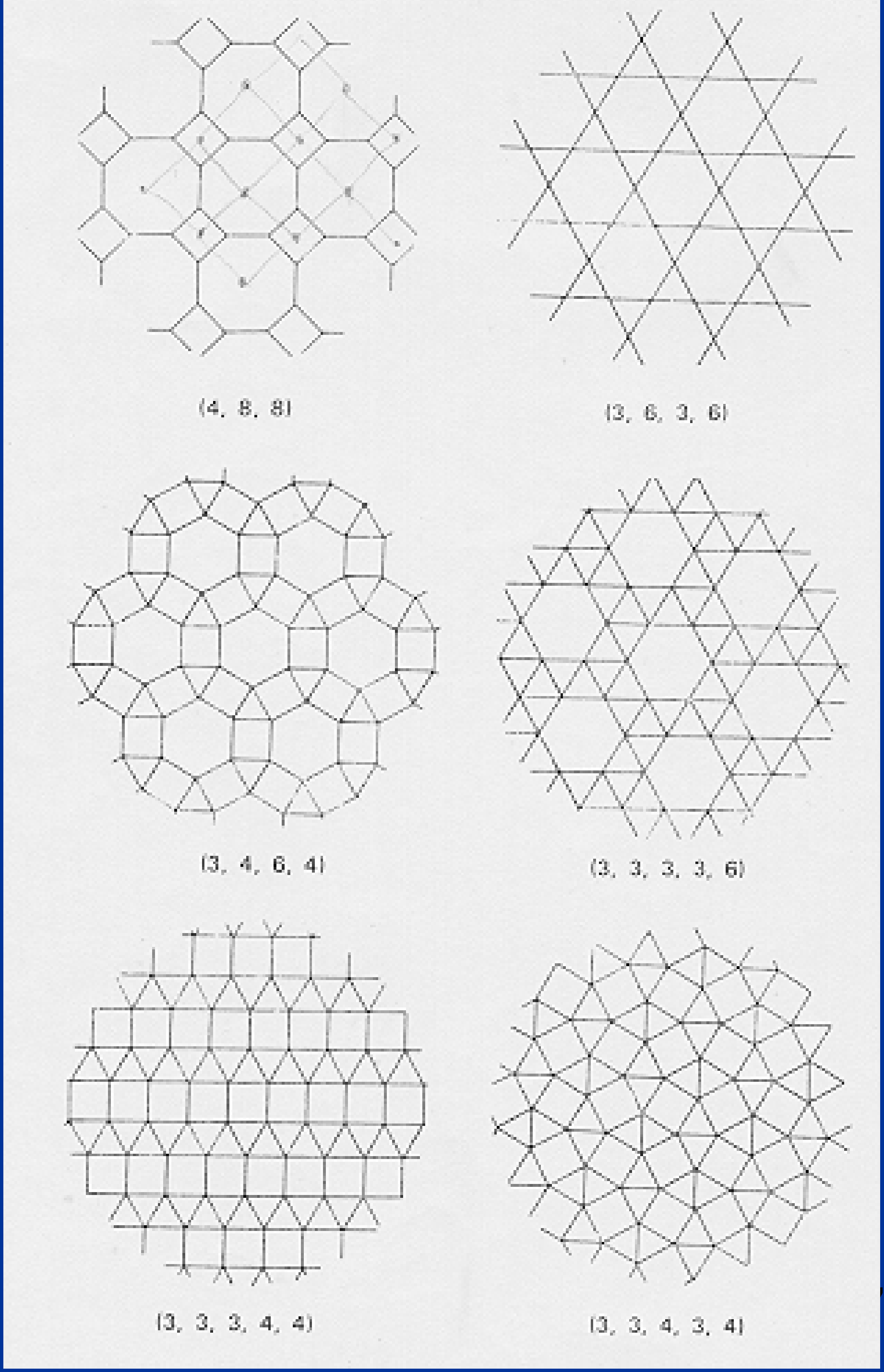
Descripción

- Código: secuencia de números donde c/u corresponde a el número de lados de los polígonos adyacentes a un vértice
- Ejemplos:
  - regular cuadrada: (4,4,4,4)
  - regular hexagonal: (6,6,6)
  - semi-regular triangular-hexagonal (3,6,3,6)

Códigos de texturas



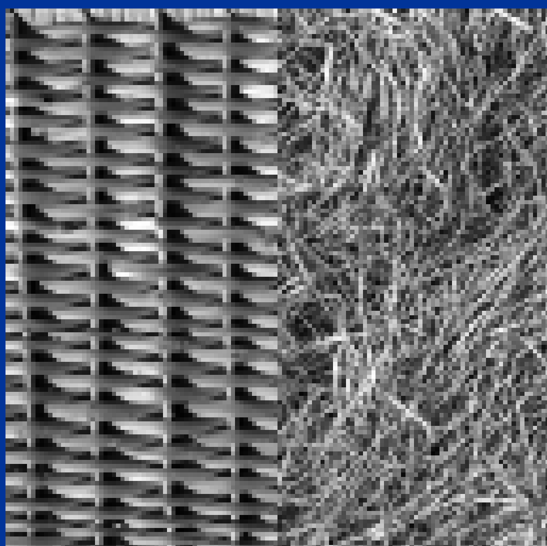
Códigos de texturas

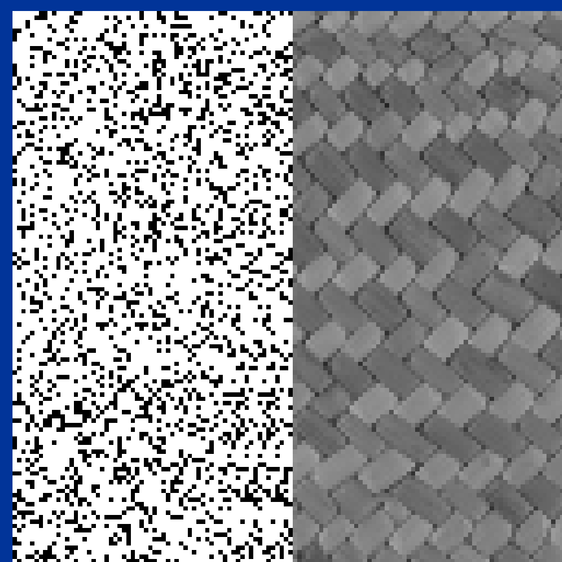


Modelos Estadísticos

- Texturas no-regulares
- Su descripción se basa en parámetros estadísticos como:
  - momentos
  - energía en el dominio espacial
  - matrices de dependencia espacial
  - transformada de Fourier (modelos espectrales)

Ejemplos de texturas no regulares

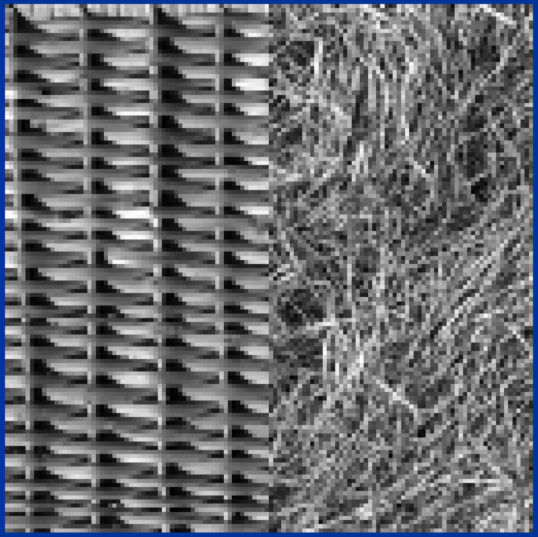


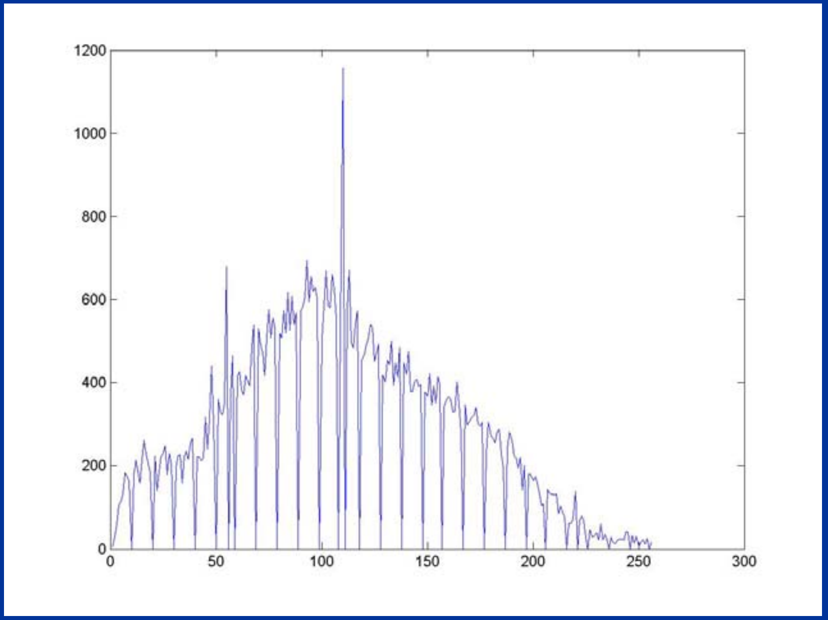


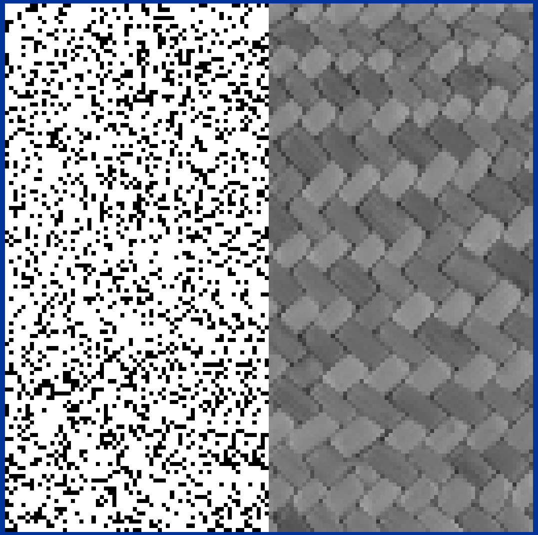
Momentos

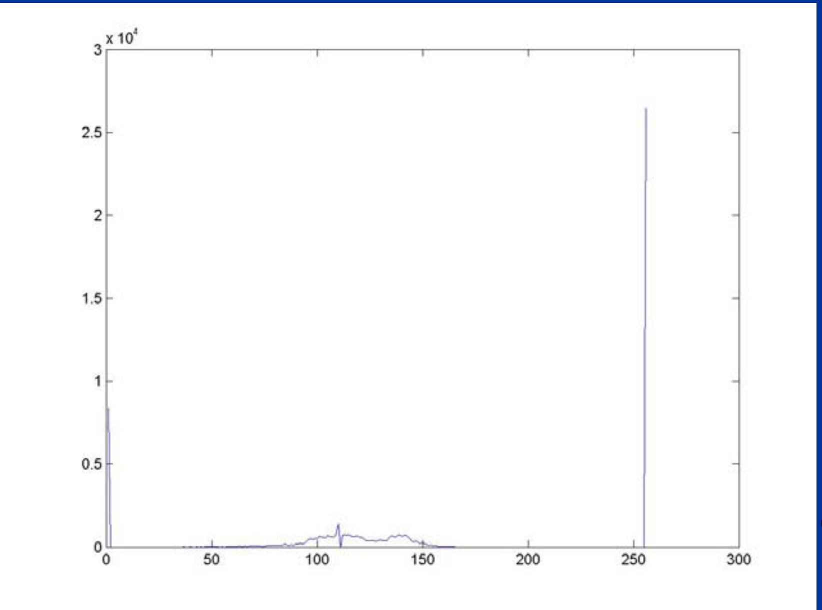
- Un forma de caracterizar una textura es obteniendo el histograma de niveles de gris
- Del histograma se pueden calcular ciertos parámetros que lo caracterizan - una posibilidad es calcular sus *momentos*

Ejemplos de Histogramas









Momentos

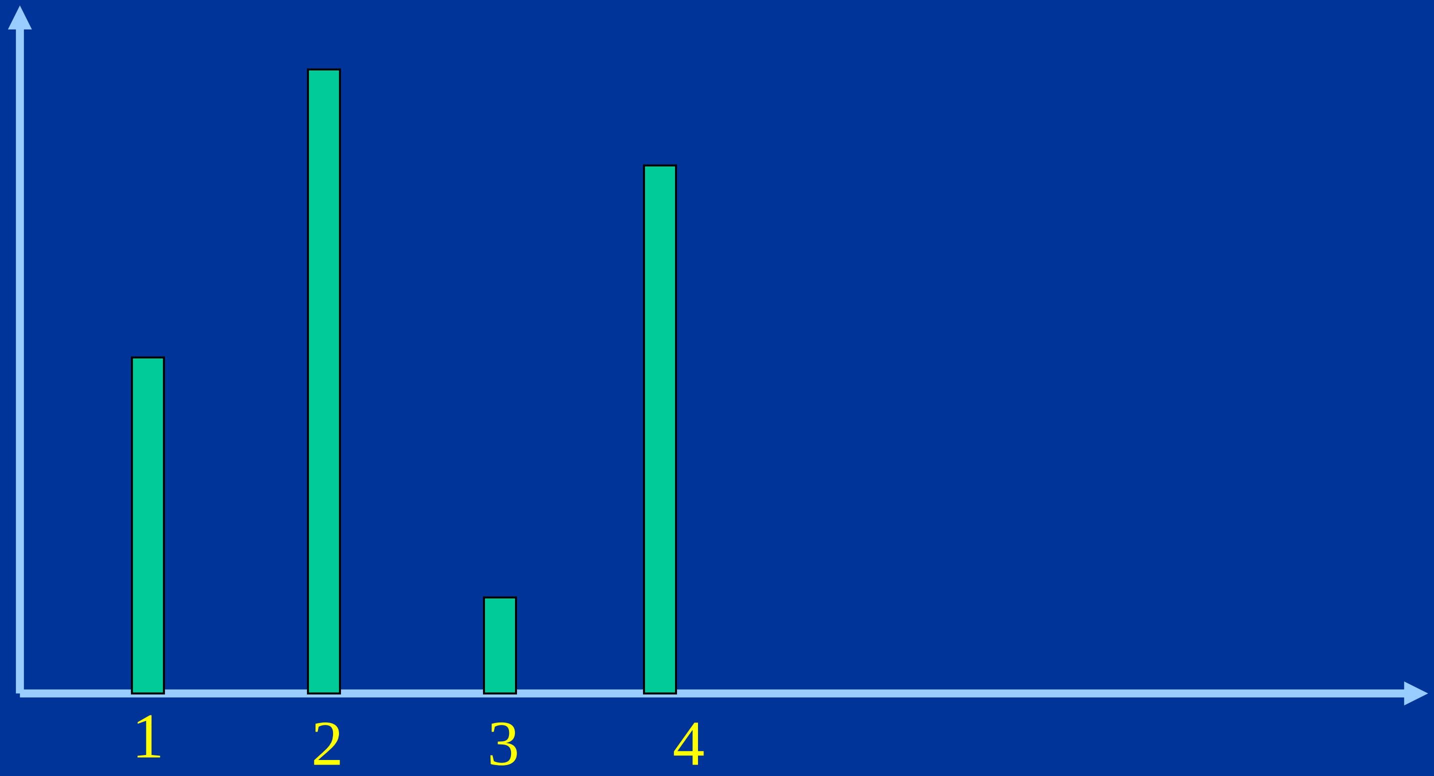
- primero (promedio): $m(z) = \sum z_i \, P(z_i)$
- segundo (desviación): $\sigma^2(z) = \sum (z_i - m)^2 \, P(z_i)$
- tercero (desplazamiento): $u_3(z) = \sum (z_i - m)^3 \, P(z_i)$
- cuarto (uniformidad): $u_4(z) = \sum (z_i - m)^4 \, P(z_i)$
- "n-ésimo": $u_n(z) = \sum (z_i - m)^n \, P(z_i)$

Medida de uniformidad

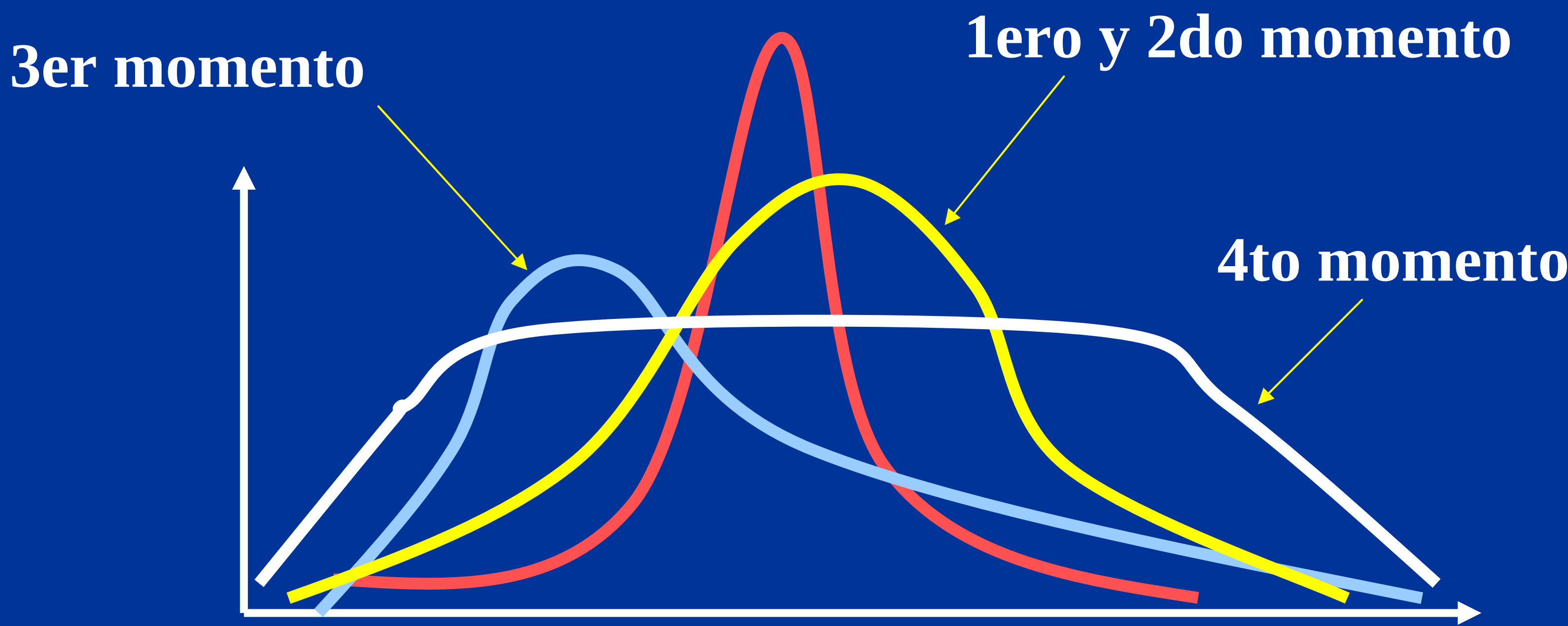
$$R = 1 - [\, 1 / (\, 1 + \sigma^2(z) \,) \,]$$

- $R = 0 \rightarrow$ uniforme
- $R \rightarrow 1 \rightarrow$ alta varianza

Histograma



Momentos



Vector de Características

- Los “n” momentos se pueden agrupar en un vector de características (*feature vector*):

$$V = (v1, v2, \ldots, vn)$$

- Este vector da una representación compacta de la textura correspondiente

Otras formas de obtener características:

- Energía en el dominio espacial
- Matrices de dependencia espacial
- Modelos Espectrales

Energía en el dominio espacial

- Se hace una transformación de la imagen mediante diferentes máscaras
- Se calculan características de cada pixel que dependen de la textura en una región local
- Se clasifica cada pixel de acuerdo a su tipo de textura

Energía en el dominio espacial

- Procedimiento:
  - Ecualización por histograma
  - Convolución con 12 máscaras diferentes (funciones base)
  - Obtención del promedio absoluto de una ventana de 15 x 15
  - Clasificación de cada pixel de acuerdo al vecino más cercano

Matrices de dependencia espacial

- Consiste en obtener unas matrices intermedias a partir de la imagen
- Se obtienen características (similares a los momentos) de dichas matrices
- Estas características se utilizan para clasificar a las texturas

Matrices de dependencia espacial

- Procedimiento:
  - Obtener las matrices *S(d,t)* para diferentes distancias (*d*) y orientaciones (*t*)
  - Cada elemento *s(i,j)* de las matriz es el número de veces que un pixel de valor *i* tiene una relación (*d,t*) con otro pixel de valor *j*
  - Normalizar cada matriz *S*

Matrices de dependencia espacial

– Obtener características de cada matriz:

- energía: $\Sigma_i\Sigma_j \; P^2_{ij}$
- entropía: $\Sigma_i\Sigma_j \; P_{ij} \log P_{ij}$
- correlación: $\Sigma_i\Sigma_j \; (1\text{-}mx)(1\text{-}my) \; P_{ij}$
- inercia: $\Sigma_i\Sigma_j \; (i - j)^2 \; P_{ij}$
- homogeneidad local: $\Sigma_i\Sigma_j \; (1/[1+ (i - j)^2]) \; P_{ij}$

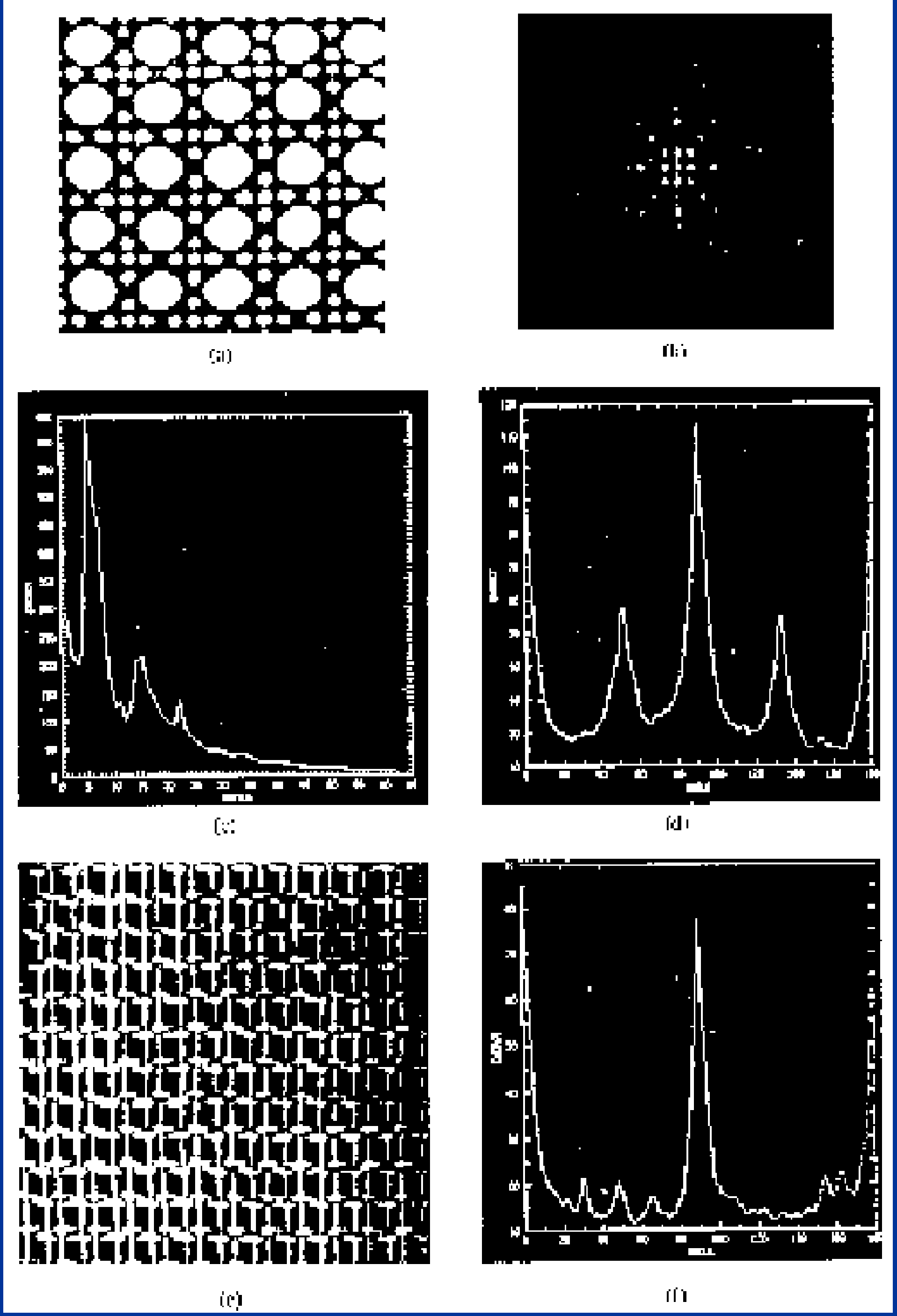
Modelos Espectrales

- Muchas texturas presentan patrones periódicos - por ello la transformada de Fourier es adecuada para describirlas
- Estos modelos consisten en obtener la transformada en frecuencia de la imagen y a partir de esta obtener ciertas características
- Dichas características sirven de base para la clasificación

Modelos Espectrales

- Características:
  - magnitud de "picos" prominentes en frecuencia
  - localización de los "picos"
  - aplicar técnicas estadísticas a partes a-periódicas
- Estas características son más fáciles de obtener del espectro en coordenadas polares

Ejemplos de espectros



Referencias

- Sucar & Gómez: Cap. 3, 4, 5
- Forsyth & Ponce: Cap. 6, 8, 9