

Visión de Alto Nivel

Dr. Luis Enrique Sucar

INAOE

esucar@inaoep.mx

ccc.inaoep.mx/~esucar



Sesión 4
Características

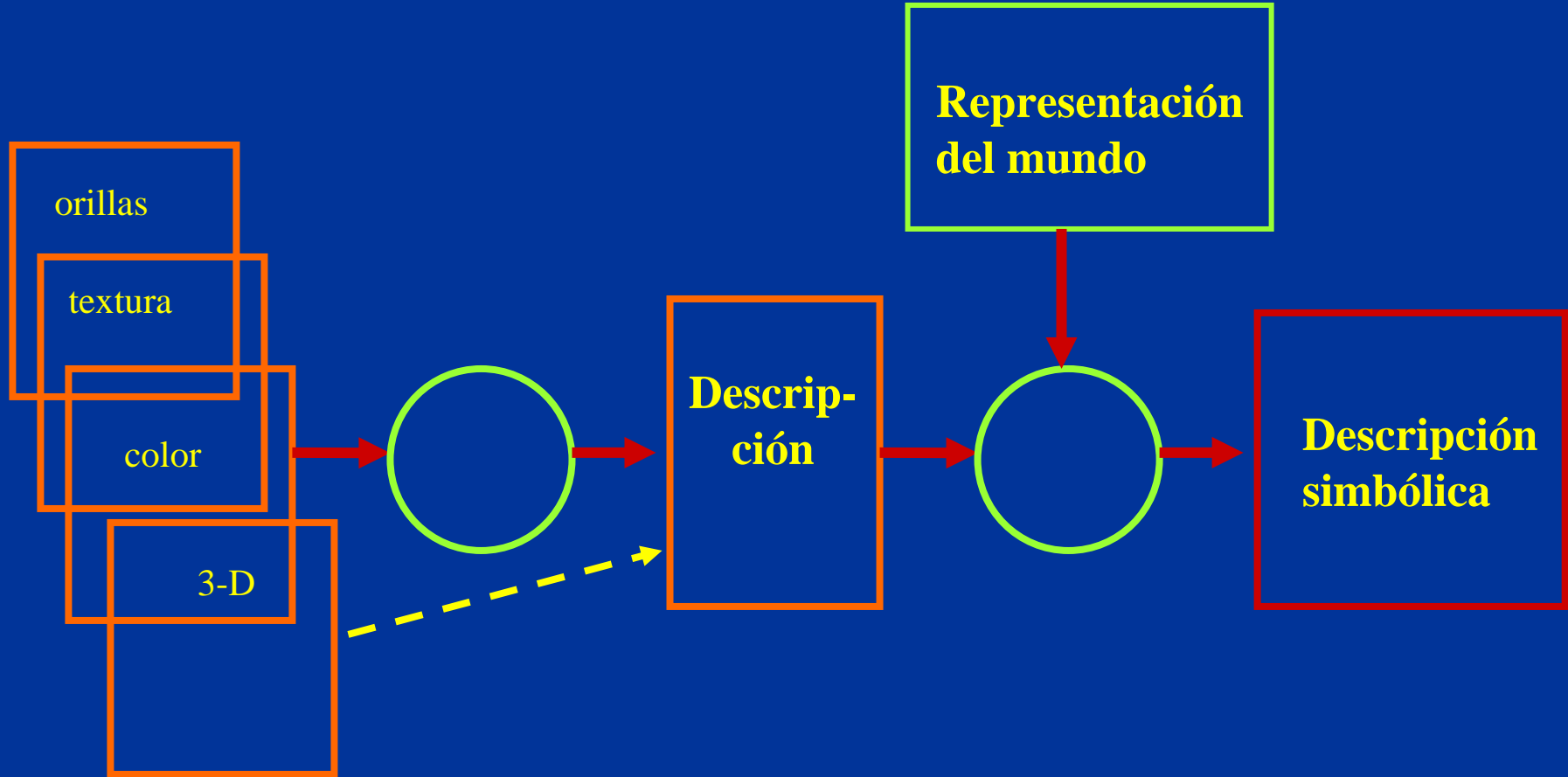
Visión de Alto Nivel

- Obtiene una interpretación consistente de las características obtenidas en visión de nivel bajo e intermedio
- Se basa en utilizar conocimiento de los objetos del dominio de interés
- En base al conocimiento y las características se realiza el *reconocimiento*

Aspectos básicos

- Descripción – caracterización de la información obtenida de la imagen
- Representación - forma de modelar el mundo, en particular los objetos de interés para el sistema
- Reconocimiento - como el modelo y la descripción de la imagen(es) son utilizadas para identificar los objetos

Reconocimiento



Representación

- “un sistema formal para hacer explícitas ciertas propiedades, incluyendo una especificación de como usarlas”
- Propiedades importantes:
 - genérica: aplica a diferentes clases de objetos
 - invariante: rotación, traslación y escala
 - robusta: tolerante a ruido y datos incompletos
 - eficiente: en espacio y tiempo

Descripción

- Para poder reconocer los objetos es necesario *describirlos* de alguna manera; para posteriormente utilizar dicha descripción para su reconocimiento
- Dicha descripción resume las características obtenidas a partir de la imagen en los procesos de visión de nivel bajo e intermedio

Descripción

- Existente básicamente 3 tipos de descripciones (que se pueden combinar):
 - Modelos geométricos en 2-D o 3-D
 - Características globales
 - Características locales

Modelos Geométricos

- Describen la “forma” del objeto
- Pueden ser en 2 o 3 dimensiones
- Se basan en el contorno (superficie) o en la región (volumen)
- Se busca que sean invariantes ante cambios de escala, rotación, traslación; y robustas ante ruido y cambios de iluminación
- Puede utilizarse un sólo modelo “genérico” o varios modelos “representativos”

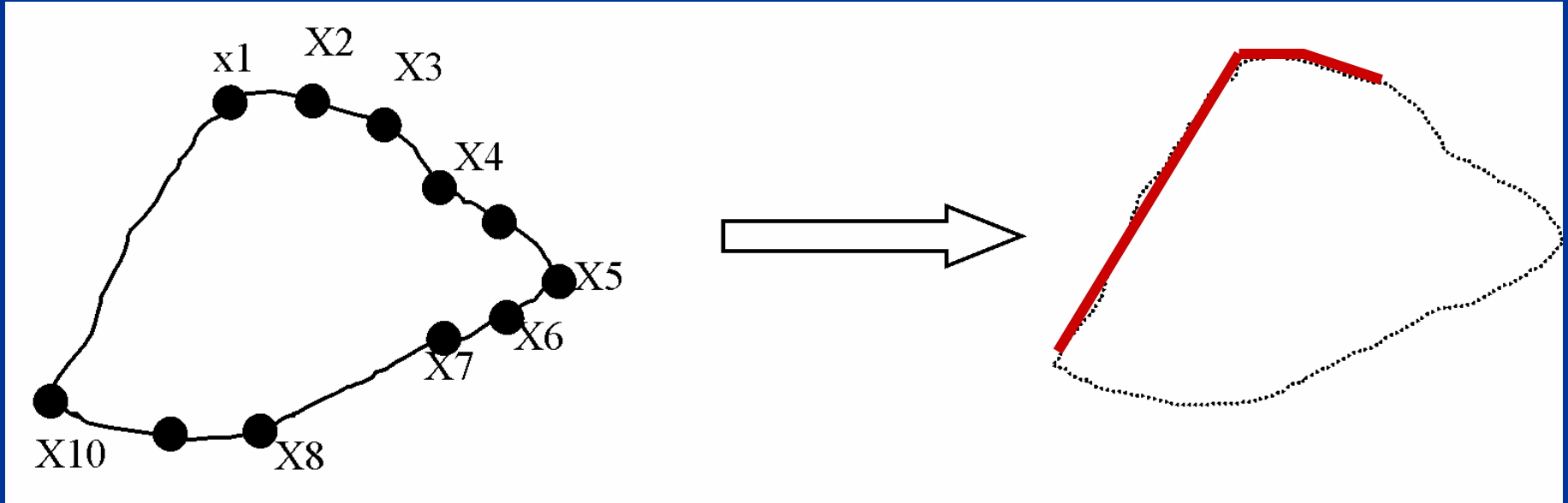
Modelos en 2 dimensiones

- En base a Contornos
 - polilíneas
 - códigos de cadena
 - descriptores de Fourier
 - secciones cónicas
- En base a Regiones:
 - arreglos de ocupación espacial
 - eje "Y"
 - árboles cuaternarios
 - esqueletos

Polilíneas

- Se representa el contorno en base a la concatenación de segmentos de línea

$$X_1, X_2, \dots, X_N, X_1 \quad X_i = (x_i, y_i)$$



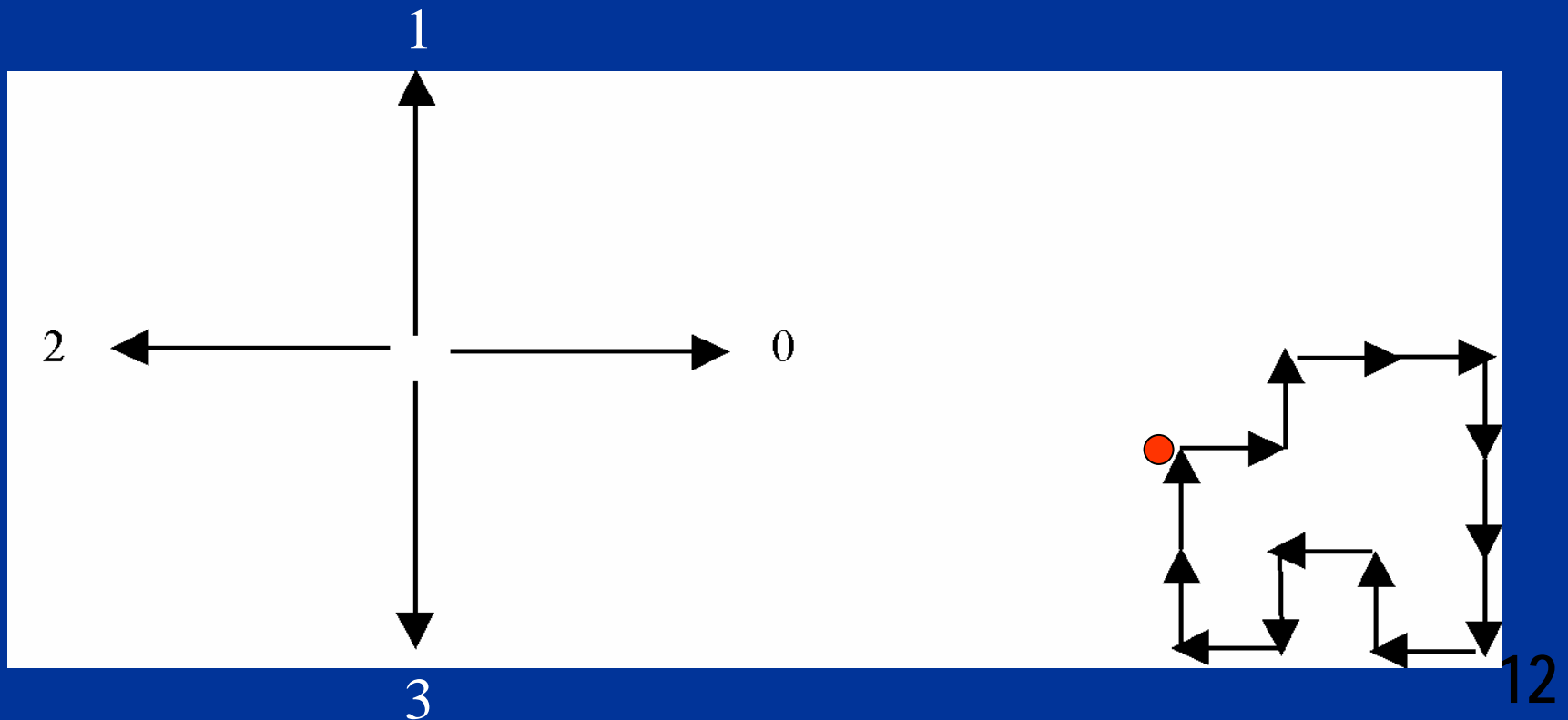
Polilíneas

- Para encontrar los puntos hay, básicamente, 2 estrategias:
 - agrupamiento → seguir la orilla hasta encontrar un cambio “brusco” de orientación
 - división → iniciar con “pocos” segmentos y dividir hasta que la distancia a la recta sea menor a un umbral

Códigos de Cadena

- Segmentos de línea en una retícula con un número de orientaciones limitadas (4 u 8)

Ejemplo: 0,1,0,0,3,3,3,2,1,2,3,2,1,1



Códigos de Cadena - derivada

- Se obtiene otro código si se toma la derivada (diferencias entre segmentos)

$$d = (X_i - X_{i-1}) \text{ MOD } N$$

- Ejemplo:

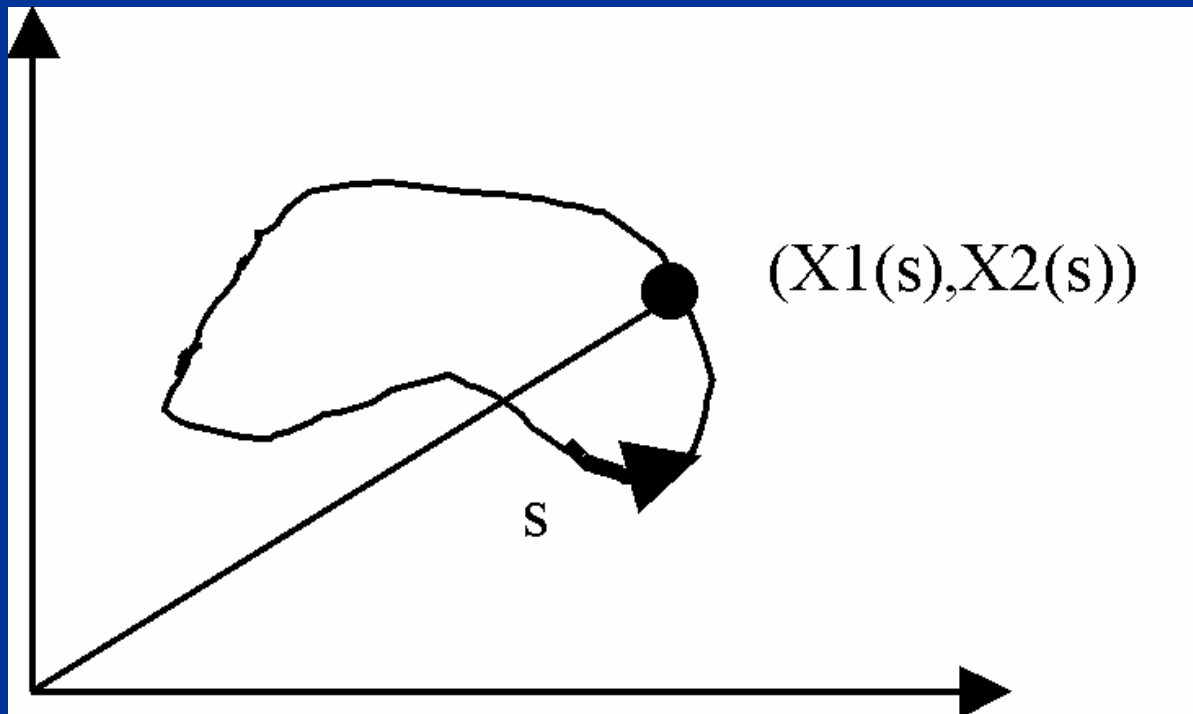
Código: 0,1,0,0,3,3,3,2,1,2,3,2,1,1

Derivada: 0,1,3,0,3,0,0,3,3,1,1,3,3,0

- Ventajas:
 - representación compacta
 - independiente de posición (invariante traslación)
 - derivada invariante a rotación

Descriptores de Fourier

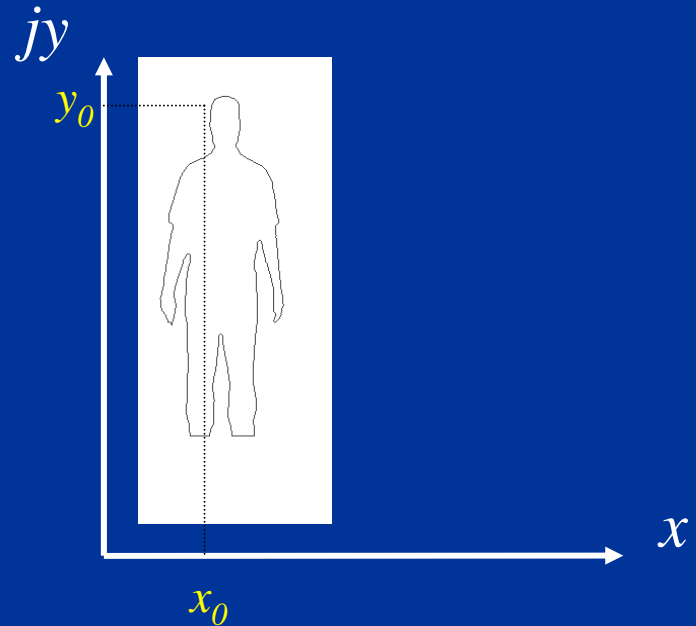
- Expansión en series de Fourier de la curva
- Ventajas:
 - Buena parametrización con pocos términos
 - Invariante a traslación y rotación



Descriptores de Fourier

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp[-j2\pi uk / N]$$

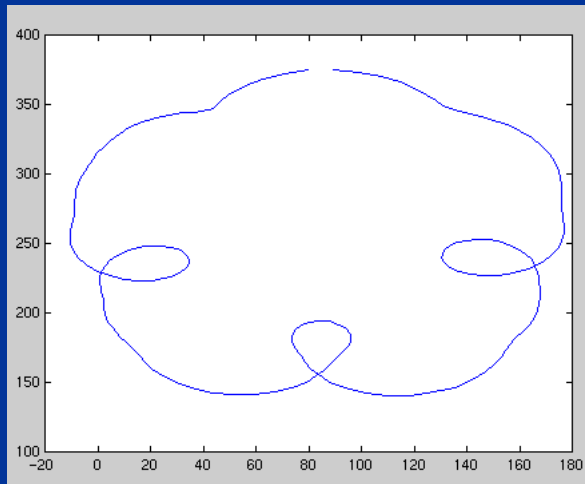
para $u = 0, 1, 2, \dots, N - 1$



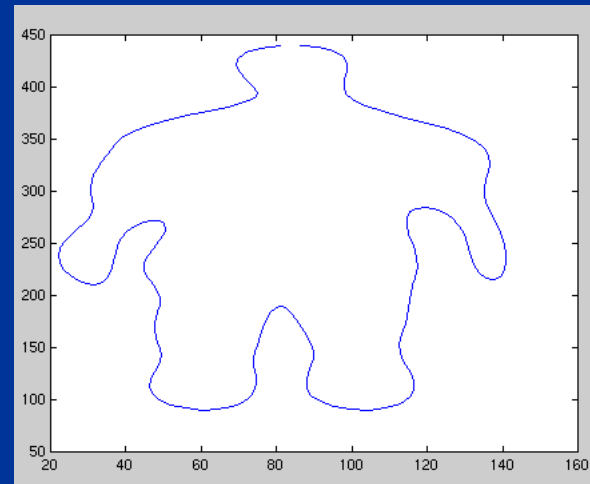
$$s(k) \exp[-jp] = [x(k) \cos(p) + y(k) \sin(p), -x(k) j \sin(p) + y(k) j \cos(p)]$$

donde $p = 2\pi uk / N$

Efecto del número de términos

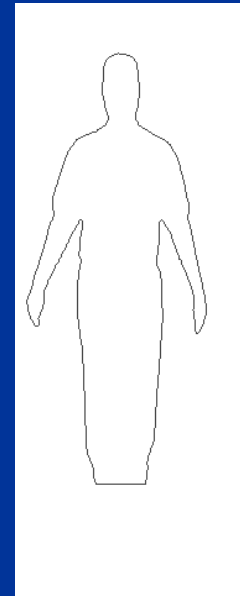
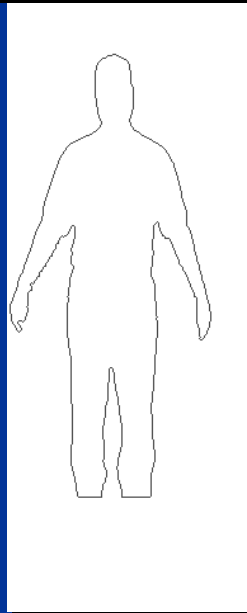


20 – bajas
frecuencias

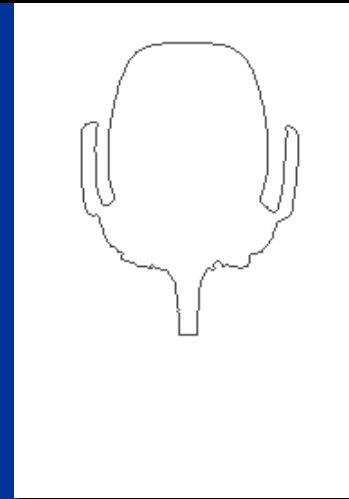
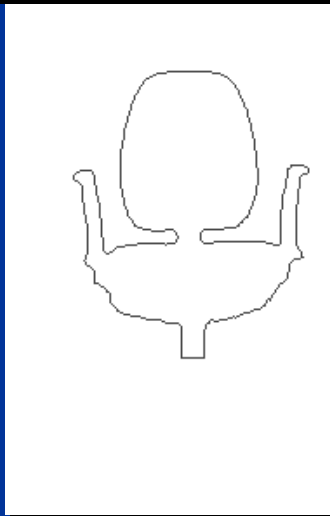


40 – bajas y altas
frecuencias

Ejemplo – humano

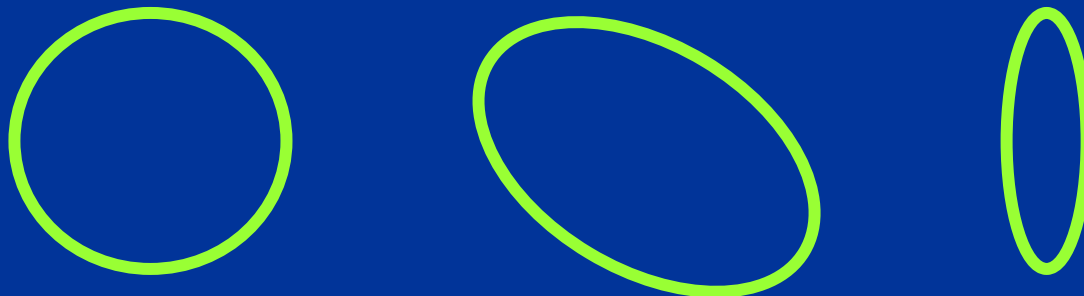


Ejemplo - silla



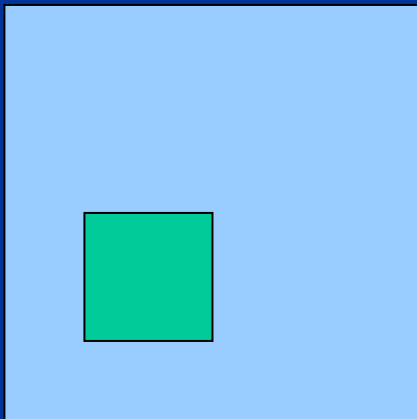
Secciones cónicas

- Representación de objetos simples con pocos parámetros:
 - círculos (3)
 - elipses (5)
 - cónicas genéricas (6)



Regiones

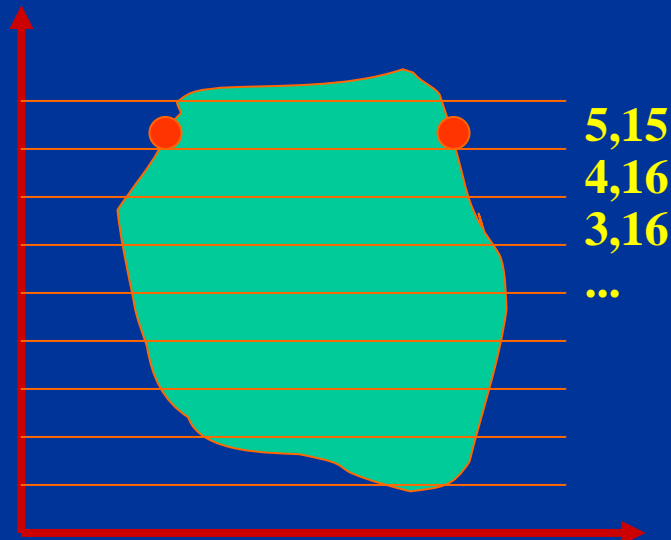
- Arreglos de ocupación espacial
 - predicado de pertenencia $p(x,y)$: 1 si el pixel pertenece a la región y 0 si no pertenece



	1	1			
	1	1			

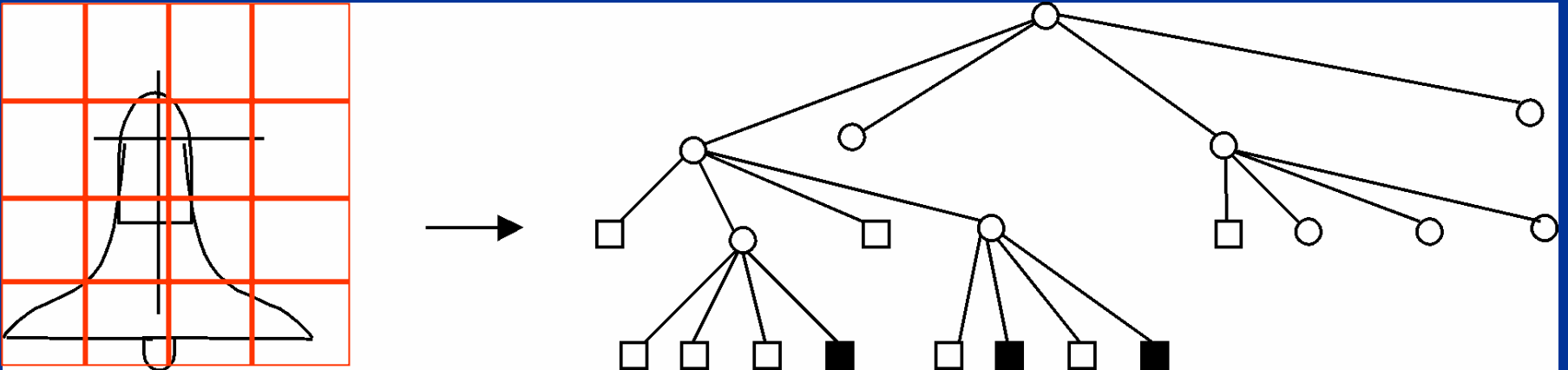
Eje-Y

- Se codifica la región en un conjunto de listas por renglón - cada lista tiene las coordenadas en X donde la línea entra/sale de la región



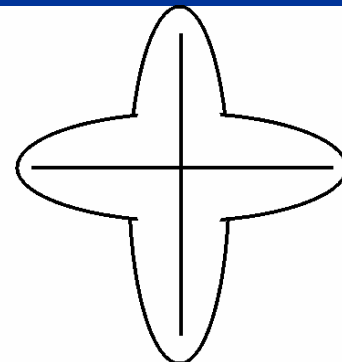
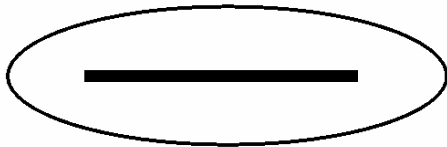
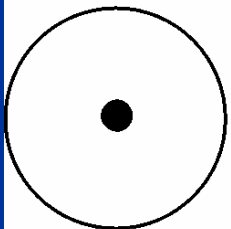
Árboles Cuaternarios

- Se representa la región como la unión de nodos de un *quadtree* a varios niveles:
 - negro: pertenece
 - gris (círculo): pertenece parcialmente
 - blanco: no pertenece

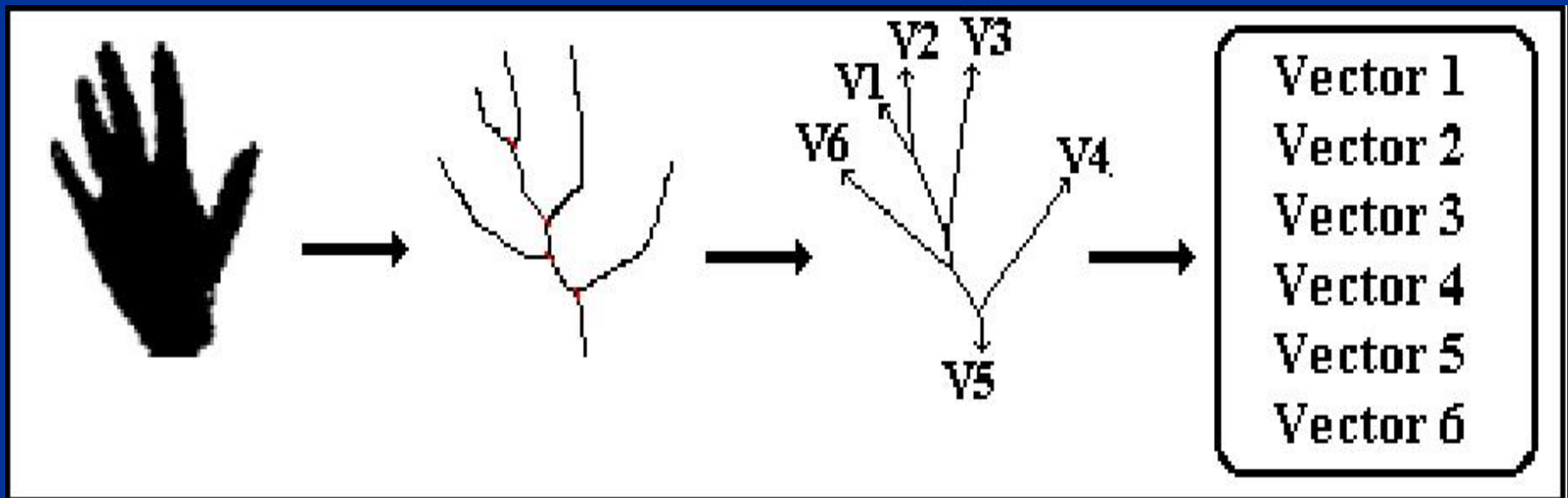


Esqueletos

- Se obtiene el “esqueleto” de la región, que corresponde a los “ejes” centrales de cada componente (diagrama de *Voronoi*)
- Adecuada para regiones *delgadas*
- Obtención del esqueleto:
 - transformada de eje medio (MAT)
 - técnicas de morfología matemática



Ejemplo de esqueleto - representación de la mano



Descriptores globales

- Propiedades simples de una forma (bidimensional):
 - área
 - excentricidad = cuerda mayor / cuerda menor
 - número de Euler = # regiones - # hoyos
 - compactes = $\text{perimetro}^2 / \text{área}$
 - firmas - proyecciones de la región sobre diferentes ejes
 - momentos - de diferente orden calculados respecto a un punto (centro de masa)
 - número de forma - basados en códigos de cadena

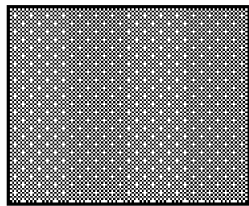
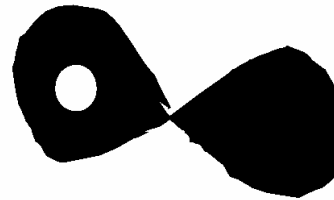
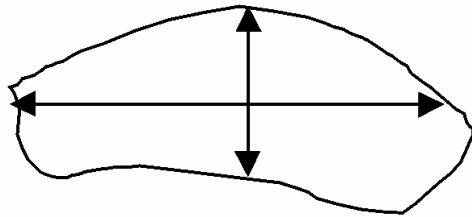
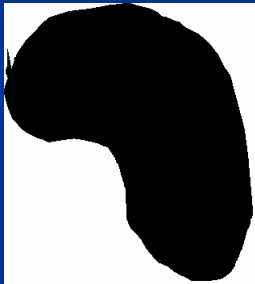
Descriptores globales

área

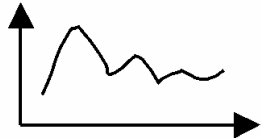
eccentricidad

número
de Euler

compactez



firmas



Modelos en 3 Dimensiones

- Representación de objetos en 3-D independiente del punto de vista
- Se consideran objetos sólidos
- Dos enfoques:
 - en base a su superficie
 - en base a una representación volumétrica

Tipos de Modelos

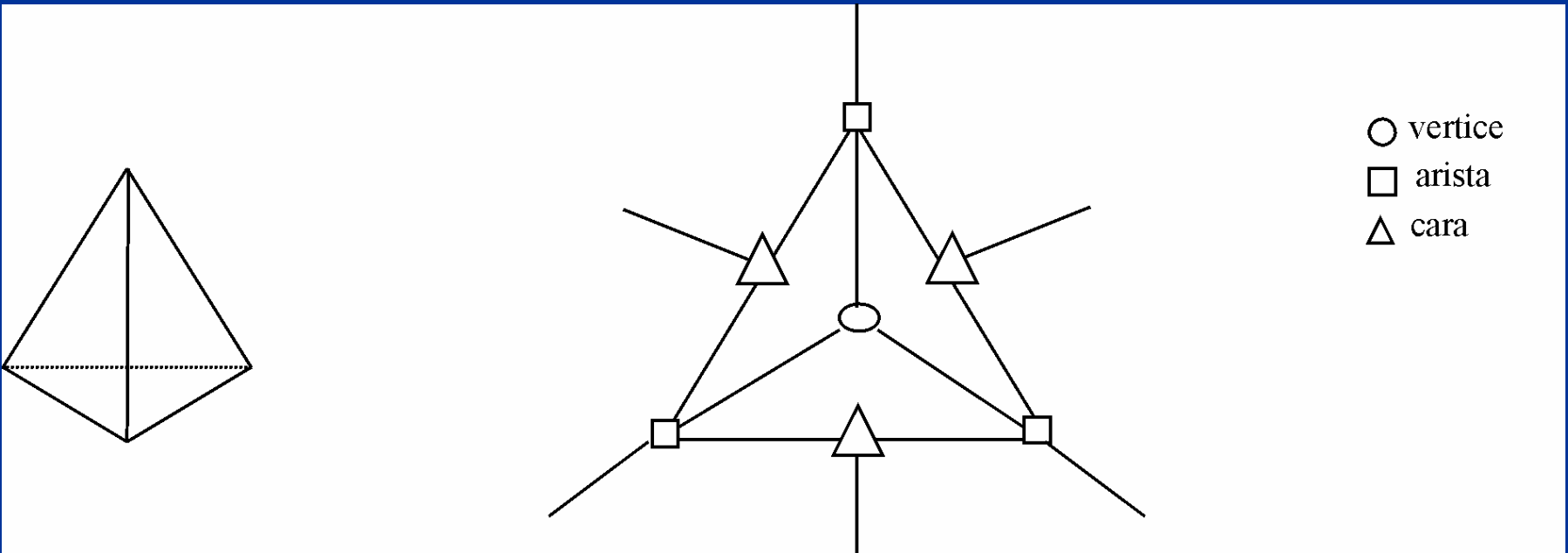
- Poliedros planos
- Cilindros generalizados
- Geometría sólida constructiva

Poliedros Planos

- Representación en base a superficie
- Se aproxima mediante poliedros planos
- Se representa mediante un grafo
- Nodos:
 - caras
 - aristas
 - vértices
- Ligas: relaciones entre nodos

Poliedros Planos

- Ejemplo - tetraedro

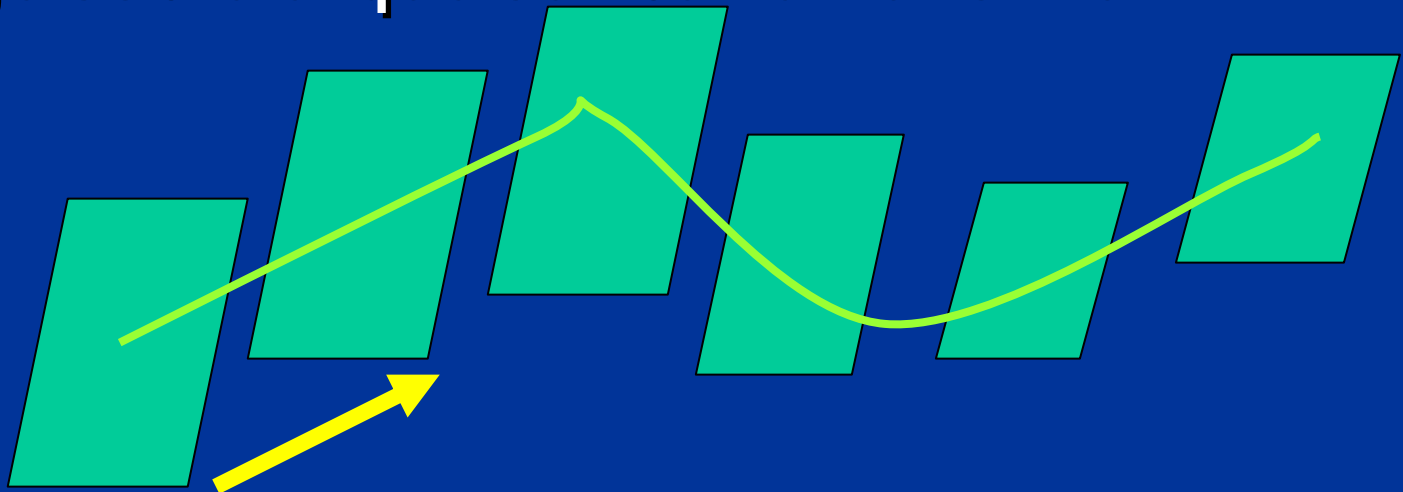


Poliedros Planos

- Reconocimiento - combinación de información de esquinas, vértices y regiones
- Aplicación - objetos que se puedan aproximar por poliedros planos:
 - manufactura,
 - ambientes de oficina,
 - etc.

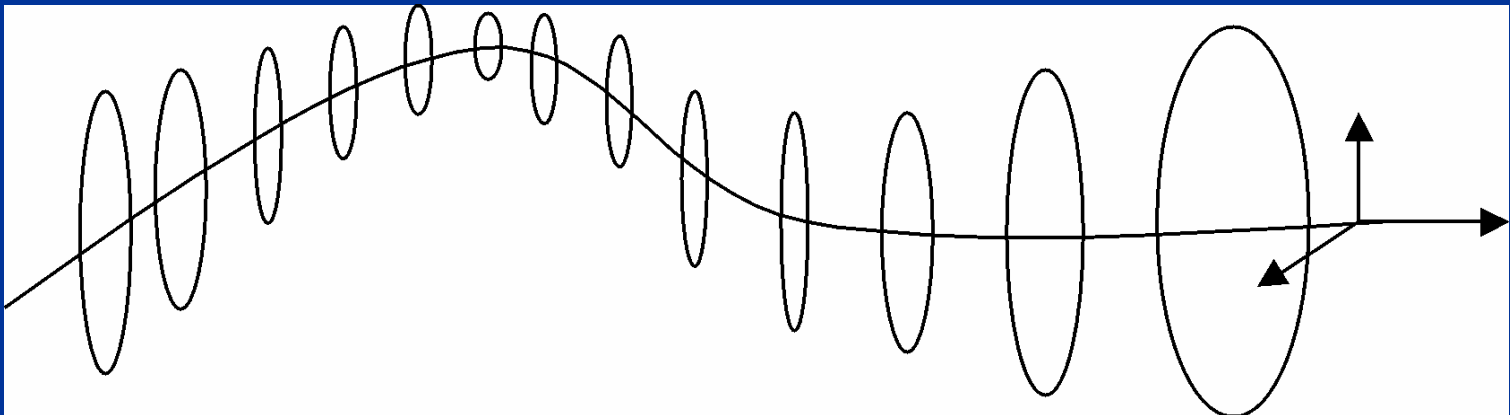
Cilindros Generalizados

- Representación que generaliza un cilindro variando la forma de la superficie y del eje
- Dos elementos básicos:
 - forma bidimensional (que puede variar)
 - eje sobre el que se “barre” la forma



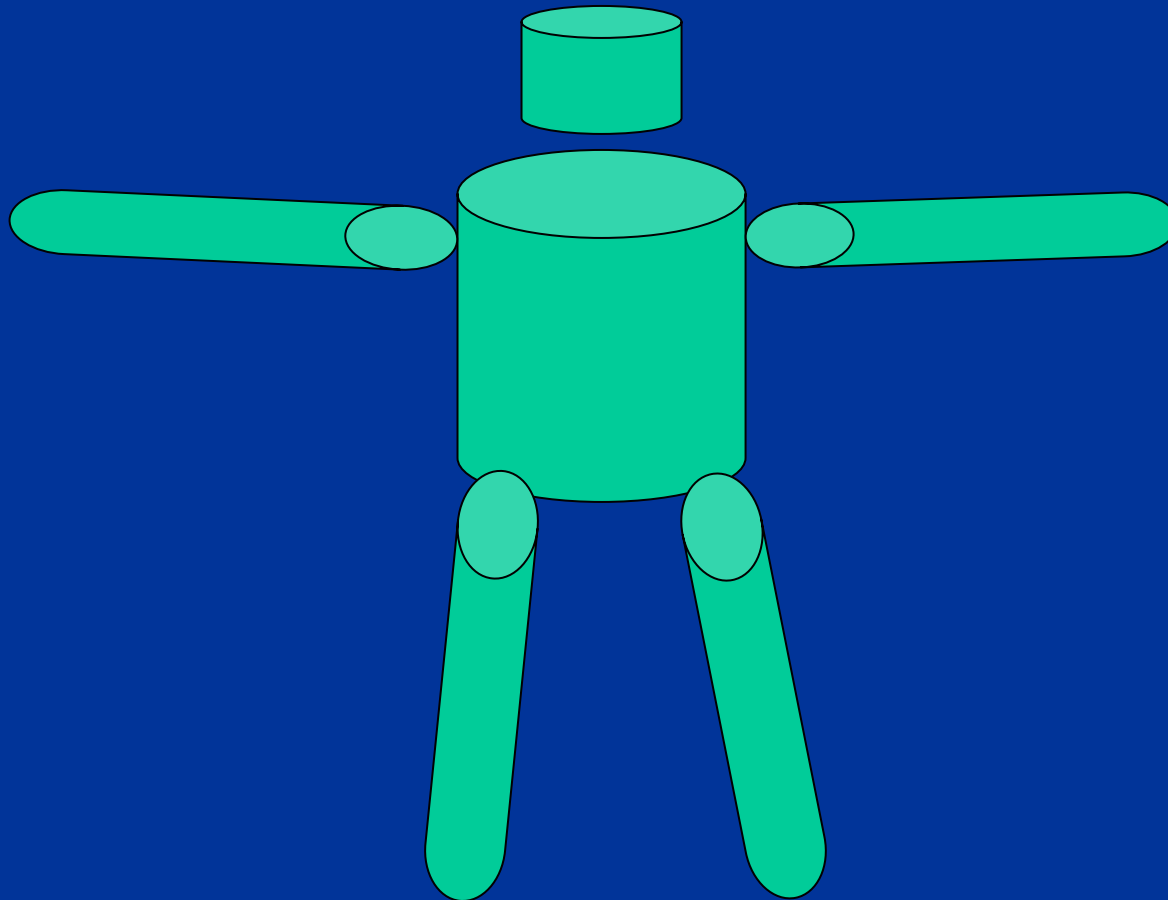
Cilindros Generalizados

- Se especifica con dos funciones:
 - una para el eje
 - otra para la superficie
- Se puede extender considerando varios cilindros generalizados, que se combinan en una estructura jerárquica
- Útil para objetos artificiales y naturales



Cilindros Generalizados

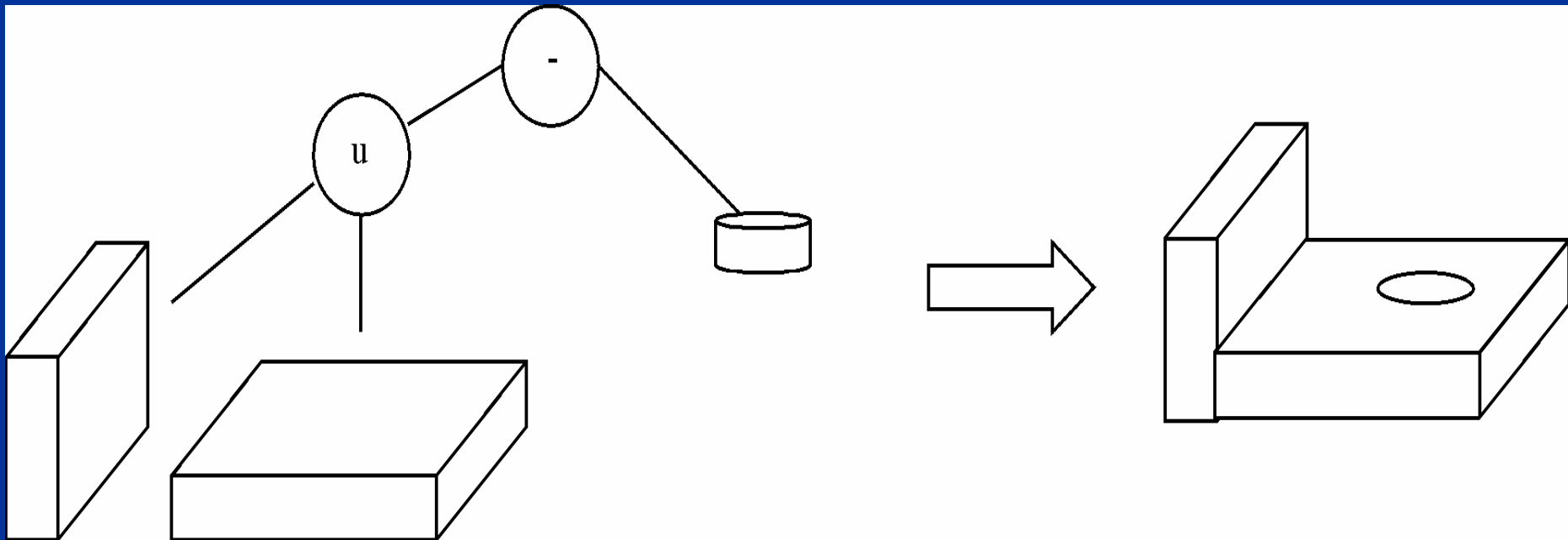
- Ejemplo de estructura jerárquica



Geometría Sólida Constructiva

- Composición de objetos sencillos (esferas, prismas, ...) para formar objetos complejos
- Se combinan mediante operaciones lógicas:
 - unión
 - intersección
 - diferencia
- Apropriada para objetos artificiales como partes para manufactura

Geometría Sólida Constructiva



Propiedades globales

- Al igual que para objetos en 2-D, se pueden definir parámetros que describen globalmente objetos en 3-D, como:
 - volumen
 - centroide
 - momento de inercia
 - producto de inercia

Propiedades Globales

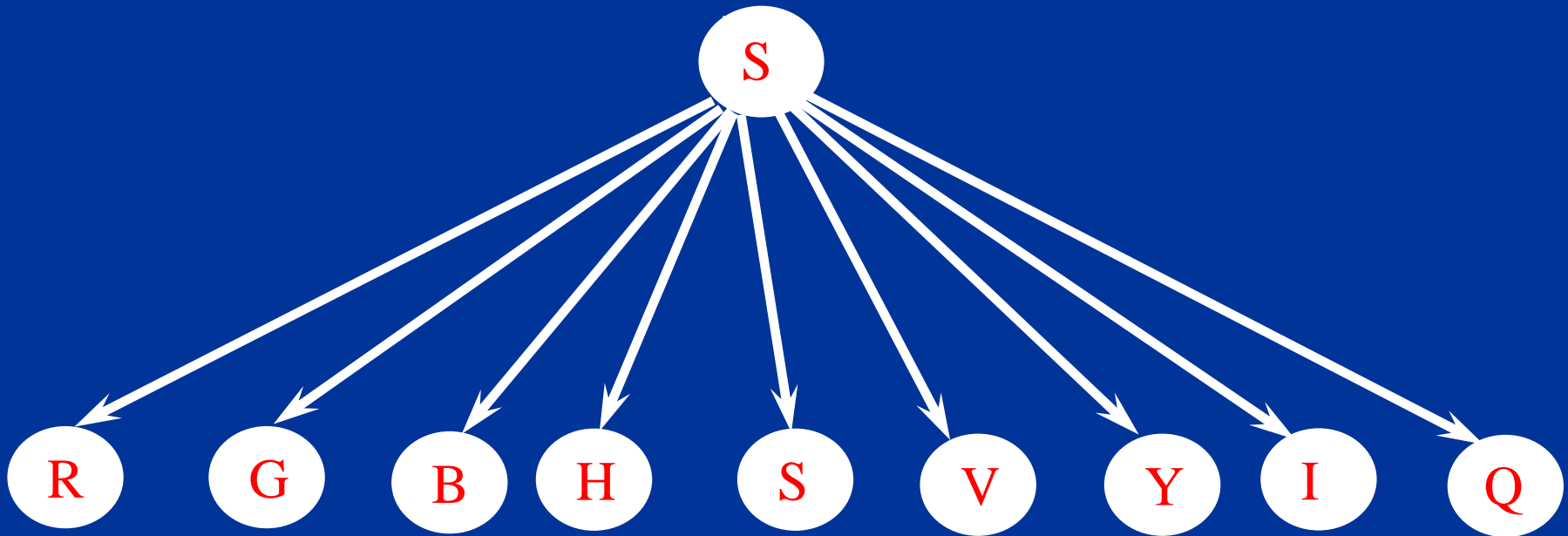
- Caracterizan la “región” correspondiente al objeto, normalmente después de un proceso de segmentación (aunque en ocasiones se pueden aplicar a toda la imagen)
- Se consideran 3 tipos de características:
 - Color
 - Textura
 - Forma

Color

- Tratan de caracterizar la distribución de color de un objeto o clase de objetos
- Se consideran diferentes modelos de color: RGB, HSI, YIQ, Luv, ... o una combinación de estos
- Se puede utilizar directamente el histograma de cada componente, o una representación compacta de este (media, varianza, momentos, ...)

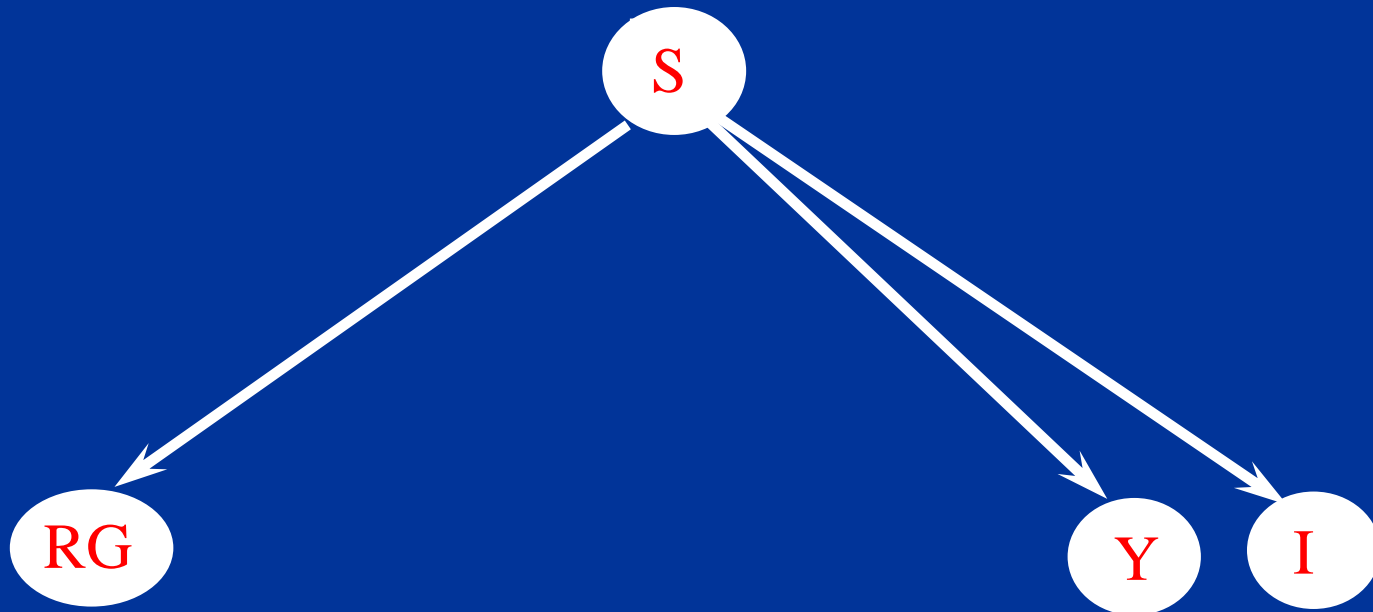
Ejemplo: clasificación de piel en base a color

- 3 modelos de color:
RGB, HSV, YIQ



Selección de atributos (mejora estructural)

Precisión: modelo completo (9 atributos) 94%
modelo final (3 atributos) 98%



Textura

- Describen la clase de textura del objeto
- Se busca caracterizar a la textura con pocos parámetros, utilizando alguna de las técnicas como momentos del histograma, matrices de dependencia espacial, ...
- Otra forma de caracterizar una textura es mediante la respuesta a un conjunto de filtros a diferentes escalas, como los filtros de Gabor

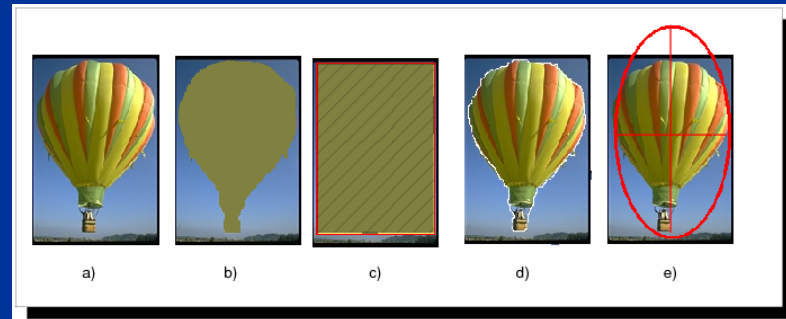
Forma

- Caracterizan la forma geométrica del objeto (región), normalmente en 2-D
- Para ello se pueden utilizar las descripciones o características globales que hemos visto para los modelos geométricos
- Estas no aplican cuando se caracteriza a la imagen en forma global o cuándo se utilizan segmentaciones en *rejilla*

Ejemplo: caracterización de regiones para etiquetado automático

Forma:

- a) Área
- b) Área convexa
- c) Perímetro
- d) Eje mayor y menor
- e) Solidez
- f) Excentricidad

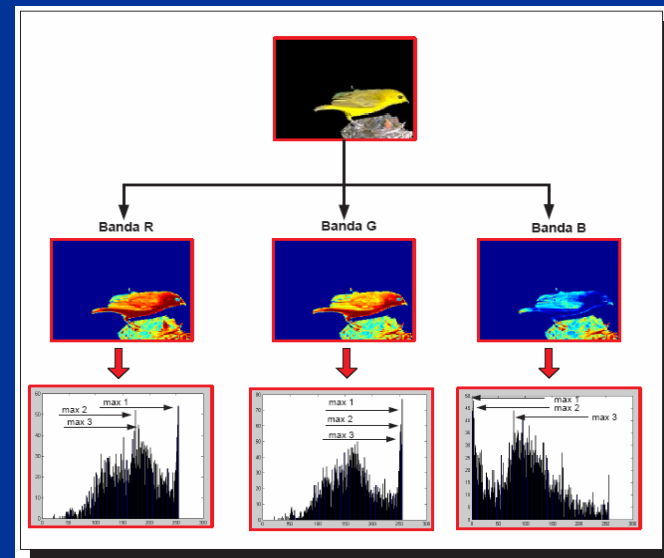


Color:

- a) Valores máximos en RGB
- b) Promedio de luminosidad
- c) Desviación estándar

Textura:

- a) Filtros de Gabor :0,45,90 y 135



Ejemplo: caracterización de regiones para etiquetado automático

temp
Menu_Regiones

EXTRACCION DE CARACTERISTICAS

Nombre del Archivo:

No. Regiones:

Region:





Imagen Original



Imagen Segmentada

Características de Color

	Primer Color	Segundo Color	Tercer Color
Maximos_R :	256	255	254
Maximos_G :	256	254	253
Maximos_B :	4	2	5

Promedios_G globales Desviacion Standar

Promedio_R	28.4869	Desviacion_R :	8.2347
Promedio_G	26.8433	Desviacion_G :	9.1381
Promedio_B	6.8041	Desviacion_B :	8.548

Características de Textura

	Media	Varianza
Orientacion 45 grados	1.0227	18.9225
90 grados	0.65442	1.8807
135 grados	0.78826	4.0376
0 grados	14.6867	905099.1162

Características de Forma

Area:	2047	Areaconvexa	2561
Perimetro	51.0521	Solidez	0.7993
Eje_Mayor	72.9302	Eccentricidad	0.85608
Eje_Menor	37.6937		

Etiquetas 1

- birds
- branch
- nest

Etiquetas 2

- city
- mountain
- sky
- sun
- water
- clouds
- tree

Tomar Etiquetas 2

Características locales

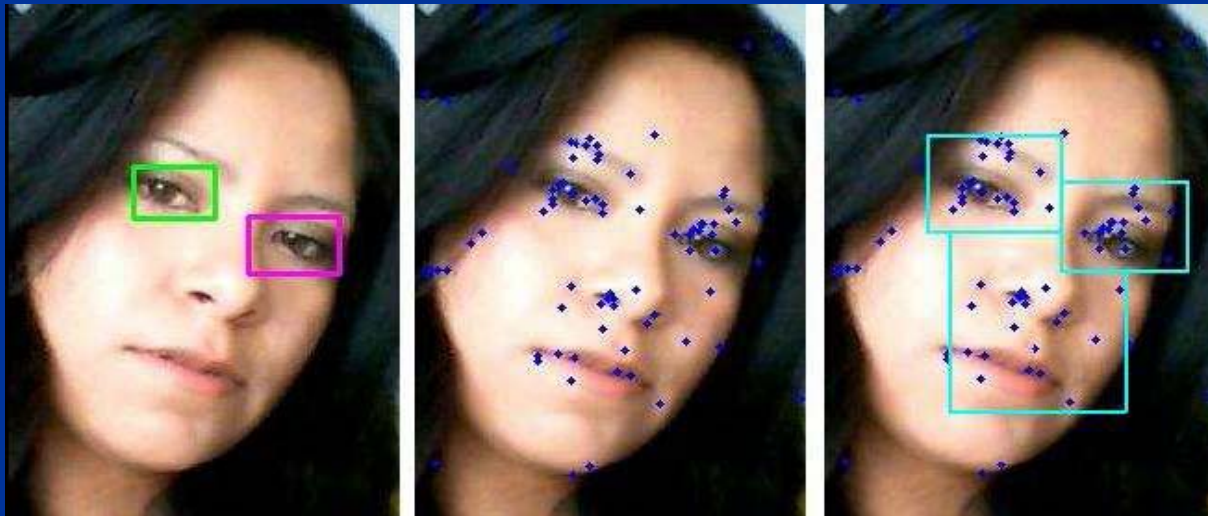
- Dado que la segmentación es un problema muy complejo aún no completamente resuelto, una alternativa es evitar la segmentación
- Entonces, para reconocer un objeto, se utilizan ciertas características locales distintivas
- Se busca que dichas características sean robustas ante cambios de escala, rotación y traslación; así como ante cambios de iluminación

Características locales

- Existen una gran variedad de características locales que se pueden utilizar:
 - Pequeños parches o *templates*
 - Orillas, esquinas (*Harris*)
 - Filtros sencillos (*Haar*)
 - Características invariantes (*SIFT*)

Ejemplo: características *SIFT*

Uso de características invariantes (SIFT)



Reconocimiento de rostros en base a características SIFT

Reconocimiento basado en SIFT



¿Qué características?

- La descripción o características a utilizar dependen de:
 - La técnica de reconocimiento – por ejemplo, los modelos geométricos son apropiados para técnicas de alineamiento; mientras que las características globales/locales son apropiadas para las técnicas basadas en reconocimiento de patrones
 - El dominio de aplicación – objetos de manufactura (geométricas), caras (locales), paisajes (globales)

¿Qué características?

- Se pueden combinar diferentes tipos de características, por ejemplo una caracterización geométrica con propiedades globales
- Se pueden utilizar técnicas de aprendizaje computacional – selección de atributos

Referencias

- Sucar & Gómez: Caps. 10
- Forsyth & Ponce: Caps. 18, 24
- M. Martínez, L.E. Sucar, "Learning an optimal naive bayes classifier", Inter. Conference on Pattern Recognition (ICPR), pp. 1236-1239, 2006.
- R. Díaz de Leon, L.E. Sucar, Human Silhouette Recognition with Fourier Descriptors. Inter. Conference on Pattern Recognition, pp. 713-716, 2000
- H. Marin-Castro, L.E. Sucar, E. Morales, "Automatic image annotation using a semi-supervised ensemble of classifiers", Lecture Notes in Computer Science 4756, Springer-Verlag, pp. 487-495, 2007.