

## Capítulo 3. Técnicas de representación de formas.

Los humanos tendemos a percibir las escenas tal y como son, compuestas de objetos individuales que se identifican mejor por medio de sus formas. La forma es una de las características más importantes de bajo nivel para la percepción humana.

Una característica es un atributo utilizado para hacer decisiones respecto a objetos en la imagen. Algunos atributos son naturales y se definen mediante la apariencia visual de la imagen, otros son artificiales, los cuales son el resultado de operaciones realizadas a la imagen.

Las figuras las podemos representar verbalmente o en imágenes. No existe una metodología aceptada para la descripción de una forma. Las estrategias matemáticas hacen representaciones efectivas de las formas, para algunas de estas estrategias no se puede utilizar la graficación por computadora, aunque es posible encontrar representaciones comunes, para la mayoría de las estrategias de descripciones de forma [41].

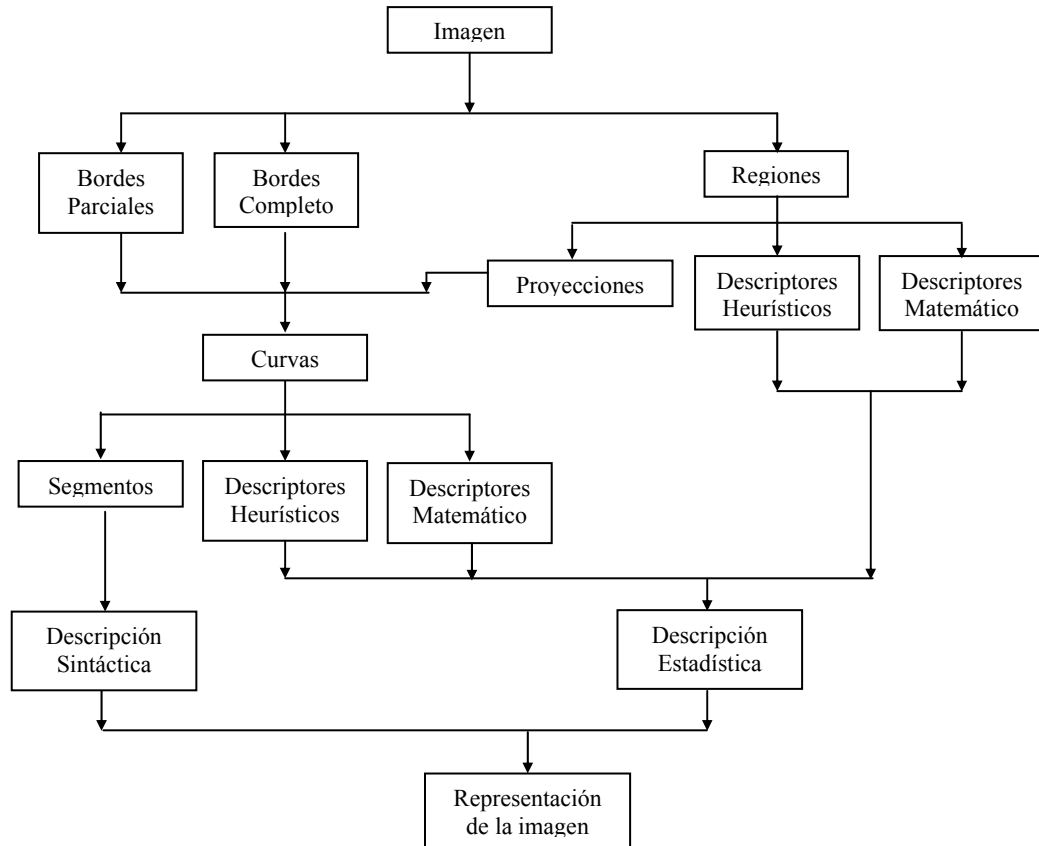
### 3.1 Clasificación de métodos de representación y descripción de formas

Los métodos de descripción de formas se pueden clasificar en ( ver Figura 3.1):

- Representación de entrada: La descripción del objeto se puede basar en los límites ó en datos más complejos de la región completa.
- Descripciones locales y globales: Los descriptores globales son aquellos que se pueden utilizar si se conoce completamente el objeto a analizar. Los descriptores locales describen propiedades del objeto usando información parcial que se conoce del objeto.
- Técnicas heurísticas y matemáticas: En técnicas matemáticas se tiene por ejemplo la transformada de Fourier; en los métodos heurísticos se puede mencionar la elongación.
- Descripción estadística.

La forma es una característica visual muy importante de una imagen. Existen muchas técnicas de representación y descripción de formas. La representación se puede basar en características ya sea partiendo solamente del contorno o del contorno además del contenido interno. Se han diseñado varias características como incluyendo: firma de forma, firma de histograma, invariantes de forma, momentos, curvatura, contexto de forma, matriz de forma, características de espectrales, etc. Dichas características se evalúan por cuanto precisión obtienen en recuperar formas similares desde un repositorio de imágenes. La recuperación efectiva requiere un descriptor de forma que sea capaz de encontrar similitudes perceptuales entre imágenes, que estén rotadas, trasladadas, escaladas.

También es importante la precisión con que pueda encontrar similitud entre formas distorsionadas, que para el ojo humano sean parecidas. A este tipo de requerimientos se le conoce como robustez.

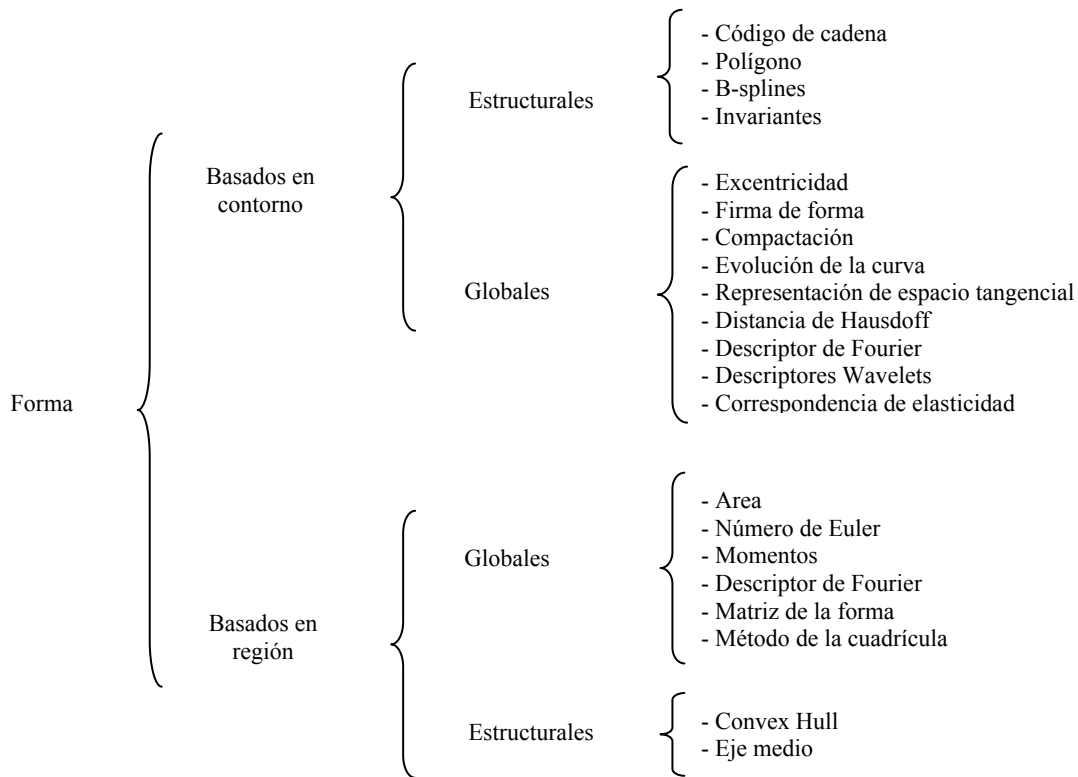


**Figura 3.1: Métodos de análisis de imágenes [41].**

La representación y descripción de formas también puede clasificarse en dos grandes clases de métodos. Métodos basados en contorno y métodos basados en región. En éste trabajo nos centraremos en los métodos basados en contorno. La clasificación está basada según si las características de la forma son extraídas solamente del contorno o son extraídas de toda la región de la forma. Cada clase a su vez se divide en estrategias estructurales y estrategias globales. Estas subclases se basan si la forma es representada como un todo o es representada por segmentos o secciones [84].

Las técnicas de contorno de forma solamente toman información de los bordes de la forma. Se dividen en dos métodos: el *método continuo (global)* y el *método discreto (estructural)*. Los métodos globales no dividen la forma en sub-partes, sino que un vector de características se basa en los bordes para describir la forma, la medida de similitud es la distancia entre los vectores de características adquiridos. Los métodos estructurales, segmentan los bordes usando un criterio en

particular. La representación final es regularmente una cadena o un grafo, la medida de similitud se obtiene comparando cualquiera de estas dos representaciones.



**Figura 3.2: Clasificación de las técnicas de descripción y representación de formas.**

### 3.1.1 Métodos Globales

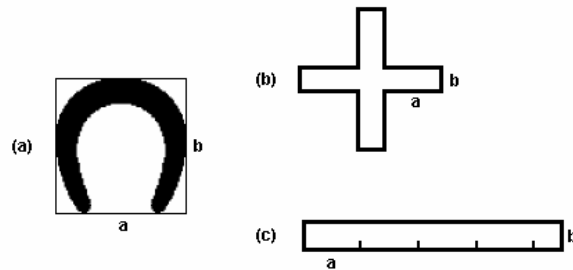
Las técnicas de representación global del contorno de una forma, comúnmente calculan un vector de características multidimensional con información del contorno de la forma. La correspondencia entre formas se lleva a cabo calculando distancias Euclidianas o distancias de bloques. Las técnicas globales de contorno de forma, trabajan con todo el contorno de la forma para como su representación. En la descripción de la forma, siempre hay cambios entre la seguridad y eficiencia, ésta debe ser tan precisa como sea posible, además de ser compacta para simplificar la indexación y la recuperación [40-42].

#### 3.1.1.1 Descriptores de formas simples

Los descriptores globales simples son el *área*, la *circularidad* ( $\text{perímetro}^2/\text{área}$ ), *excentricidad* (que se refiere a la longitud del eje mayor y/o menor), orientación del mayor eje y la intensidad de

la curvatura [40]. Regularmente estos descriptores, solo pueden descartar formas con diferencias muy notables.

La excentricidad en la forma (a) de la Figura 3.3, es cercana a 1 ( $a = b$ ) por lo que no describe correctamente la forma, ya que es una forma alargada. En este caso la circularidad es un mejor descriptor. Las dos formas (b) y (c) en la Figura 3.3, tienen la misma circularidad ( $a = 2b$ ), sin embargo, hay muchas formas diferentes. En este caso la excentricidad es un mejor descriptor.



**Figura 3.3: Ejemplo de (b) excentricidad y (c) circularidad de la forma (a).**

### 3.1.1.2 Similitud de forma basada en correspondencia

Esta similitud trabaja en el dominio de espacio, determina la similitud de formas midiendo la correspondencia punto a punto de la forma, la correspondencia se conduce a un espacio 2D. Se puede utilizar la distancia Hausdorff para medir la similitud entre las formas de los objetos [41].

Dadas dos formas representadas por dos conjuntos de puntos:  $A = \{a_1, a_2, \dots, a_p\}$  y  $B = \{b_1, b_2, \dots, b_p\}$ , la Distancia Hausdorff entre A y B, es definida como:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (3.1)$$

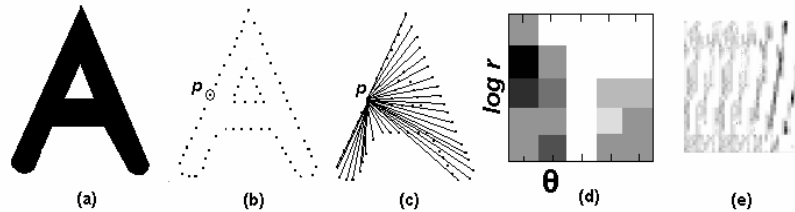
donde 
$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (3.2)$$

donde  $\|a - b\|$  es la normal subyacente en los puntos de A y B, que es usualmente la Distancia Euclidiana. Sin embargo esta medida de distancia es muy sensitiva al ruido. Un punto en A que este lejos de alguno en B, ocasiona que la distancia  $h(A, B)$  sea muy larga.

La ventaja de usar la distancia Hausdorff es que la forma puede ser correspondida parcialmente, sin embargo, esta distancia no es invariante a rotación, translación y escalamiento [42].

### 3.1.1.3 Contexto de imagen.

La correspondencia usando el contexto de la forma es una mejora del método de distancia de Hausdorff [43], el cual extrae una característica global, para cada punto correspondiente. Para extraer el contexto de la forma en un punto  $P$ , se encuentran todos los vectores de  $P$  para todos los puntos del borde.

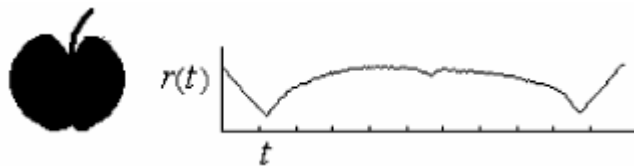


**Figura 3.4:** (a) Objeto original; (b) Contorno de (a); (c) vectores desde un punto  $P$ ; (d) mapa de histograma; (e) contexto de la imagen.

En la Figura 3.4 (c), la longitud de  $r$  y la orientación  $\theta$  de los vectores se contabilizan para crear un mapa de histograma, que es usado para representar el punto  $P$  como se observa en Figura 3.4 (d), en donde el eje  $\theta$  representa el ángulo con respecto a cada vector en (c) y el eje  $\log r$ , representa el logaritmo del radio  $r$ . El histograma de cada punto se nivela y se concatena para formar el contexto de la forma (Figura 3.4 (e)). Estos vectores están en el espacio logarítmico polar. La similitud entre dos formas se realiza mediante la correspondencia de los dos mapas de contexto de las formas, basado en matrices [43].

### 3.1.1.4 Firmas de forma.

Representa una forma con una función unidimensional derivada de los puntos del borde. Existen muchas firmas de este tipo como: distancia del centroide  $r(t)$  en Figura 3.5, coordenadas complejas, ángulo tangente, ángulo acumulativo, curvatura, y área, entre otros [44].



**Figura 3.5:** Gráfica de la distancia del centroide hacia los puntos del contorno

Debido a que los objetos del mundo son tridimensionales, al observar un objeto desde diferentes posiciones u orientaciones en el espacio, se pueden formar diferentes proyecciones en dos

dimensiones, con lo cual sería ideal tener un descriptor capaz de detectar estos cambios para diseñar descriptores invariantes a proyecciones.

Las firmas de las formas se normalizan para que sean invariantes a escalamiento y a rotación [45].

### 3.1.1.5 Momentos

Los momentos se pueden utilizar para reducir las dimensiones de la representación de los bordes. Asumiendo que el borde de la forma ha sido representado por medio de una firma de forma  $z(i)$ , el  $r$ -ésimo momento  $m_r$  y el momento central  $\mu_r$  se pueden calcular por [46]:

$$m_r = \frac{1}{N} \sum_{i=1}^N [z(i)]^r \quad \text{y} \quad \mu_r = \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^r \quad (3.3) \text{ y } (3.4)$$

Siendo  $N$  el número de puntos del borde de la forma. Los momentos normalizados son:

$$\bar{m}_r = m_r / (\mu_2)^{r/2} \quad \text{y} \quad \bar{\mu}_r = \mu_r / (\mu_2)^{r/2} \quad (3.5) \text{ y } (3.6)$$

Las ecuaciones 3.5 y 3.6 son invariantes a translación, rotación y escalamiento.

Aunque los momentos son fáciles de implementar, son difíciles de asociar a interpretaciones físicas.

### 3.1.1.6 Correspondencia por elasticidad.

Fue propuesto por Bimbo y Pala [47], se genera una plantilla deformada, partiendo de la suma de la plantilla original  $\tau(s)$  y el esfuerzo de la deformación  $\theta(s)$ :

$$\varphi(s) = \tau(s) + \theta(s), \quad (3.7)$$

donde  $\tau = (\tau_x, \tau_y)$  es un *spline* de segundo orden y  $\theta = \theta_x, \theta_y$  es la deformación. La similitud entre la forma original de la plantilla y la forma del objeto en la imagen se calcula minimizando la función:

$$F = S + B + M = \alpha \int_0^1 \left[ \left( \frac{d\theta_x}{ds} \right)^2 + \left( \frac{d\theta_y}{ds} \right)^2 \right] ds + \beta \int_0^1 \left[ \left( \frac{d^2\theta_x}{ds^2} \right)^2 + \left( \frac{d^2\theta_y}{ds^2} \right)^2 \right] ds + \int_0^1 I_E(\varphi(s)) ds \quad (3.8)$$

$IE$  es el objeto de la imagen,  $S$  es la fuerza de estiramiento y  $B$  la fuerza de encorvar y  $M$  es el grado de trasladar la plantilla deformada y el objeto en la imagen. Debido a que estos tres parámetros no son suficientes para medir la similitud entre formas, se obtienen los parámetros de complejidad de la forma  $N$ , (número de ceros de la función de curvatura asociada con las plantillas del contorno) y la correlación  $C$  (entre la función de curvatura asociada con la plantilla, asociada con la deformada). Por lo que se obtienen cinco parámetros ( $S$ ,  $B$ ,  $M$ ,  $N$ ,  $C$ ) los cuales son clasificados por una red neuronal de retropropagación. La descripción de la forma no es invariante a rotación; el criterio para deformar la plantilla no está definido con lo que puede ser arbitrario [47].

### 3.1.1.7 Métodos estocásticos.

Entre estos métodos se encuentra la autoregresión (AR), basado en el modelo estocástico de una función unidimensional  $f$  como las obtenidas en la firma de forma [48-50]. Un modelo lineal autoregresivo define un valor de función como una combinación lineal de un cierto número de valores que le preceden. Cada valor de la función, tiene una correlación con respecto a los valores de la función previa y puede ser predecida a través de un número  $M$  de observaciones de valores de funciones previas.

El modelo de autoregresión es un simple pronosticador definido por:

$$f_t = \alpha + \sum_{j=1}^m \theta_j f_{t-j} + \sqrt{\beta} \omega_t \quad (3.9)$$

donde  $\theta_j$  es el modelo de coeficientes de autoregresión,  $m$  es el modelo de orden, que dice cuantos valores de funciones precedentes usa el modelo.  $\sqrt{\beta} \omega_t$  es el error del término actual o residuo, el cual refleja la precisión de la predicción.  $\alpha$  es proporcional al término medio de los valores de las funciones. Los parámetros  $\alpha$  y  $\beta$  no son invariantes a escala, pero el factor  $\alpha/\sqrt{\beta}$  que refleja el

radio de la señal de ruido, se considera como invariante. El vector de características que se utiliza

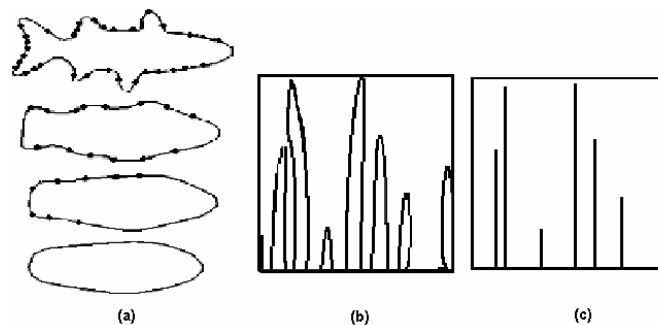
como descriptor de formas esta definido por:  $\left[ \theta_1, \dots, \theta_m, \alpha/\sqrt{\beta} \right]^T$ . La desventaja de este método es

que en caso de bordes complejos, no es suficiente un pequeño número de parámetros de autoregresión para una adecuada representación. Además la selección de  $m$  es complicada, se define empíricamente y el significado físico asociado con el valor  $\theta_j$  no es muy claro [48-50].

### 3.1.1.8 Método de escala de espacio

Este método crea la representación de la forma, rastreando los puntos de inflexión en el borde de la forma, que es filtrada previamente con el filtro Gaussiano de paso bajo, con variable ( $\sigma$ ). Mientras esta variable incrementa, se eliminan puntos de inflexión insignificantes y la forma se va suavizando. Los puntos de inflexión que permanecen, son características significantes del objeto (Figura 3.6 (b)) [85].

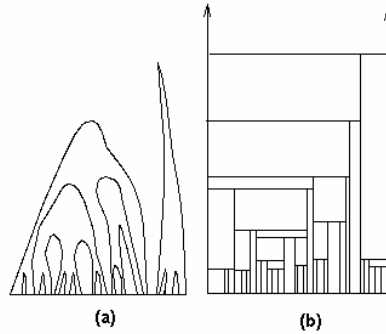
Básicamente el procedimiento de éste método es procesar cada límite de una imagen CSS y entonces encontrar el punto máximo de los contornos CSS que son usados como descriptores de formas para comparar objetos. La imagen CSS es una organización multi-escala de los puntos de inflexión del contorno donde se desarrolla. La curvatura es una medida local de que tan rápido, un contorno plano está dando vuelta. Las coordenadas de esos puntos junto con el *aspect ratio* de la imagen (total de columnas y total de renglones), la excentricidad, la circularidad y el nombre de la imagen original, constituyen características que representan al objeto. Cuando se comparan dos imágenes, se encuentra el cambio horizontal óptimo de los máximos en una de las imágenes CSS, indicando la traslapación posible más eficiente con los máximos de las otras imágenes CSS. El costo de comparación se define como la suma en parejas de distancias entre los pares correspondientes de máximos [85].



**Figura 3.6:** (a) Evolución del contorno de la forma, cuando incrementa ( $\sigma$ ).  $\sigma = 1,7,12,14$ , bordes con puntos de inflexión; (b) mapa de curvatura, imagen CSS; (c) picos de la curvatura [51].

Mokhtarian y Mackworth [52] extendieron este método a la escala del espacio de la curvatura (Curvatura Scale Space CSS). Los picos, en lugar de interpretarse como primitivas de alto nivel, se utilizan para la correspondencia entre dos curvas a analizar. Para cada correspondencia individual, la ramificación del contorno tiene que ser escalada y acomodada de tal manera que sea invariante a traslación y a escalamiento.





**Figura 3.7: Escala de espacio, (a) variación del número y localizaciones de los puntos de segmentación de la curva; (b) representación de la curva en un árbol de intervalos.**

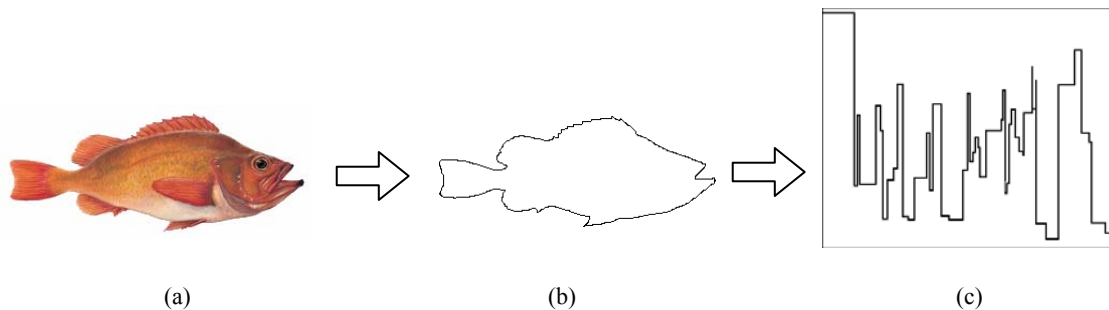
Las ventajas que presenta este método son [52]:

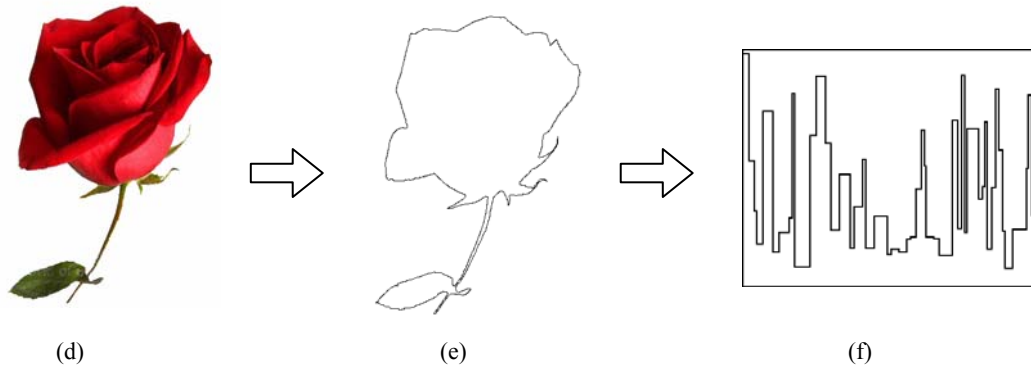
- Robusto con respecto al ruido, escala y cambios de orientación.
- Retiene información local sobre la forma de la imagen de entrada. Cada convexidad de la forma de la imagen tiene su propio contorno correspondiente en la imagen CSS.
- Es rápido. Al procesar una consulta, lo que se hace es procesar la imagen CSS de entrada y extraer sus máximos. Entonces el sistema compara el vector de características de todas las imágenes en la base de datos.

Sistemas como SQUID utilizan este método para la representación y comparación de formas, obteniendo resultados satisfactorios.

### 3.1.1.9 Representación de espacio tangencial

En este método, una curva  $C$  se representa por el grafo que forma una función escalón, los pasos en el eje  $X$  representan la longitud normalizada del arco de cada segmento  $C$ , el eje  $Y$  representa el ángulo de giro entre dos segmentos consecutivos en  $C$ . Se recorre la curva en dirección de las manecillas del reloj y se asigna a cada línea de segmento máximo de la descomposición de la curva, un escalón en el espacio tangencial [53].





**Figura 3.8: Representación de espacio tangencial. La dirección del ángulo de giro entre dos segmentos consecutivos (e) y (f) se determina con el eje Y. La longitud de cada línea normalizada del segmento con respecto a la longitud de la curva es representada por el eje X.**

Esta función también es conocida como representación de espacio tangente (Tangent Space Representation). La representación de la tangente es invariante a rotación, escalamiento y a traslación.

### 3.1.1.10 Evolución discreta de curvas.

Esta técnica consiste en reducir el conjunto de vértices de un polígono a un subconjunto de vértices, que contiene información relevante del polígono original. Se asigna una medida de relevancia a cada vértice, para que el vértice menos importante sea removido. Cuando un vértice remueve, sus vértices vecinos deben ser conectados. Éste proceso se repite hasta que se obtenga una forma simplificada deseada. Las sustituciones de vértices se dan por la medida de relevancia  $K$ , siendo  $\beta(s_1, s_2)$  el ángulo del vértice común de los segmentos  $s_1, s_2$ ,  $l$  es la longitud normalizada con respecto a la longitud total de la curva poligonal  $C$ . Mientras más pequeño es el valor de  $K(s_1, s_2)$ , menos es la contribución a la forma de la curva, del arco  $s_1 \cup s_2$ . Para detener el proceso de eliminación de vértices (evolución de la curva), es necesario especificar un parámetro que defina el número de iteraciones [54].

$$K(S_1, S_2) = \frac{\beta(S_1, S_2)l(S_1)l(S_2)}{l(S_1) + l(S_2)} \quad (3.10)$$

Propiedades y ventajas de la técnica de evolución de la curva:

- Siempre induce a un polígono convexo. Debido a que el algoritmo elimina un vértice en cada paso. Si se ejecutan el número máximo de iteraciones, obtenido cuando se tienen tres vértices, se forma un triángulo que es convexo. Aunque el estado convexo se puede obtener antes.

- Es invariante a rotación, escalamiento, reflexión y translación. Los valores de la función de costo se calculan en el espacio de tangente, el cual es un toro. Invariante a escalamiento, ya que la longitud de la curva es normalizada. Invariante a reflexión ya que la función de costo mínimo, no depende de la dirección ni del orden decreciente de los segmentos. Invariante a translación porque el espacio de tangente, sólo contiene información sobre posiciones relativas (al primer vértice) de los vértices.
- Reduce la complejidad de las curvas. La complejidad de una forma poligonal, se define como la suma de los ángulos entre segmentos adyacentes. En cada ciclo se reduce la complejidad ya que se va eliminando un ángulo.
- Se preservan las posiciones de las características importantes. Los puntos irrelevantes se van eliminando, por lo que los puntos relevantes no varían su posición, esta propiedad, facilita encontrar características importantes en un nivel más bajo de la jerarquía de la evolución de la curva.
- Estable a deformaciones de ruido. El ruido es el primer elemento que se elimina, ya que sólo se conserva información importante de la curva.
- Puede determinar un segmento recto. Debido a su propiedad de eliminación de ruido, por lo que líneas de segmentos pueden ser detectadas.
- Nivel de abstracción universal. La curva final, tiene un estado en donde los puntos obtenidos son suficientes para el reconocimiento de formas, no importa que tan distorsionado esté el contorno inicial.

### **3.1.1.11 Transformaciones espectrales**

Los descriptores espectrales superan el problema de la sensibilidad al ruido y variaciones en el borde, analizando la forma en el dominio espectral. Los descriptores espectrales como el descriptor de Fourier (FD) y el descriptor Wavelet (WD), se derivan de transformaciones espectrales en firmas de forma unidimensionales.

Las Wavelets son funciones matemáticas que cortan datos a diferentes componentes de frecuencia, estudiando cada componente con una solución que se empareja con su escala.

Recientemente, muchos trabajos proponen el uso de los descriptores Wavelets para la descripción de la forma. Tienen las ventajas sobre métodos de Fourier en analizar situaciones físicas donde la señal contiene discontinuidades y puntos agudos [55].

Este tipo de descriptores es multi-resolución, trabaja tanto en el espacio espectral como en el espacial, el incremento de la resolución espacial, sacrificará la resolución de la frecuencia. El procedimiento de las transformadas Wavelets es adoptar una función prototipo llamada un

anализador Wavelet. Se realiza un análisis temporal con una versión contractada de alta frecuencia del prototipo Wavelet, mientras que el análisis de la frecuencia es ejecutada por una versión dilatada de baja frecuencia del mismo Wavelet [55, 56].

Los descriptores de Fourier solo trabajan con curvas cerradas. Rui et al [57] proponen una medida de distancia para clasificar la similitud usando coeficientes de la transformada de Fourier. Esta distancia es la suma ponderada de la varianza de la magnitud de las razones y la varianza de la diferencia entre los coeficientes de dos series de Fourier. Los coeficientes de Fourier se derivan de bordes de la forma reconstruidos por Fourier, en vez de los bordes originales.

Las ventajas que presentan el descriptor de Fourier sobre otros descriptores de forma son: fáciles de implementar, cada descriptor tiene un significado físico específico, simple para hacer normalización, que captura características globales y locales. Estos descriptores resuelven la habilidad de discriminación de los descriptores globales, además de resolver problemas de sensibilidad al ruido y de dificultad en representar la normalización de la firma de la forma.

Suponiendo que C es una curva cerrada y viajando en contra de las manecillas del reloj, se obtiene la función compleja  $z(t)$ , donde  $t$  es la variable del tiempo.

$$z(t) = \sum_n T_n e^{int} \quad (3.11)$$

Los coeficientes  $T_n$  son llamados descriptores de Fourier, de la curva C. Es más común considerar la distancia de la curva  $s$  en comparación con el tiempo.

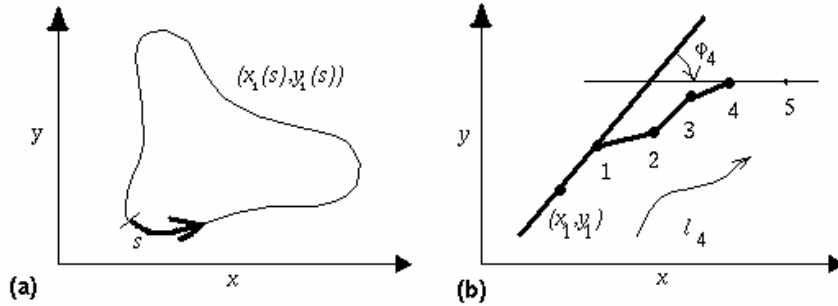
$t = 2\pi s / L$  donde L es la longitud de la curva. Los descriptores de Fourier  $T_n$  están dados por:

$$T_n = \frac{1}{L} \int_0^L z(s) e^{-i(2\pi/L)ns} ds \quad (3.12)$$

Un borde cerrado puede ser representado como una función del ángulo tangente contra la distancia entre los puntos del borde desde los cuales fueron determinados los ángulos.

Definiendo así las ecuaciones:  $a(l_k) = \varphi_k + u_k$  y  $u_k = 2\pi l_k / L$ .

El conjunto de descriptores está dado por:  $S_n = \frac{1}{2\pi} \int_0^{2\pi} a(u) e^{-inu} du$  (3.14)



**Figura 3.9: Descripción de bordes por Fourier: (a) Descriptores  $T_n$ , (b) descriptores  $S_n$**

La mayoría de los trabajos basados en descriptores de Fourier, se enfocan a reconocimiento de caracteres y clasificación de objetos. Las coordenadas complejas y la función de ángulo acumulado son utilizadas en este tipo de enfoques para derivar los descriptores de Fourier. Zhang y Lu [58, 59] concluyeron que para formas generales, la función de distancia hacia el centroide es la firma de forma, más deseable para derivar el descriptor de Fourier.

### 3.1.2 Métodos estructurados.

Estos métodos parten las formas en segmentos de bordes llamados primitivas, estos métodos difieren entre sí, en la selección de primitivas y su organización para la representación de la forma. Los métodos más comunes de descomposición de borde están basados en aproximaciones poligonales, descomposición de curvatura y ajustes de curva.

El resultado es codificado dentro de una cadena de la forma general,  $S = S_1, S_2, \dots, S_n$  donde  $S_i$  puede ser un elemento de un código de cadena, el lado de un polígono, un arco, una tira, etc.  $S_i$  puede contener un número de atributos como, longitud, promedio de la curvatura, máxima curvatura, esfuerzo de deformación, orientación, etc. La cadena se puede usar directamente para la descripción o puede ser usada como entrada para un analizador de más alto nivel. Los beneficios de los métodos estructurales, son su capacidad de manejar el problema de oclusión en una escena y permitir una correspondencia parcial, sin embargo tiene varios inconvenientes, como a continuación se explican en cada técnica [86].

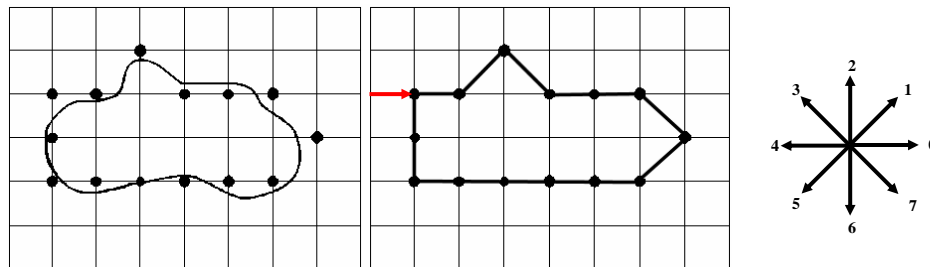
#### 3.1.2.1 Representación de código de cadena.

Describe un objeto por una secuencia de segmentos de línea de un tamaño unitario con una orientación dada. El método fue introducido por Freeman en 1961 [60], en donde una curva arbitraria se representa por una secuencia de pequeños vectores de longitud unitaria y un conjunto

limitado de posibles direcciones. En la implementación se sobrepone una rejilla en una imagen que contiene el borde de un objeto, los puntos del borde son aproximados al punto más cercano de la rejilla, al final se obtiene una imagen aproximada. Seleccionando un punto de inicio, se genera un código de cadena usando un código de dirección 4 o de dirección 8. Si el código de cadena es usado para la correspondencia de formas, debe ser independiente del primer píxel del borde en la secuencia. Una forma para normalizar el código de cadena es encontrar el píxel en la secuencia de bordes que tenga el entero menor, si la descripción de la cadena es interpretada como una base de 4 números, entonces ese píxel será el punto de inicio, para hacer la correspondencia. Posteriormente se representará el borde con las diferencias las direcciones sucesivas en el código de cadena, en lugar de representarlo por direcciones relativas. Con lo que se puede calcular restando cada elemento del código de cadena del previo y tomar el valor del módulo  $n$  (que es la conectividad) [60].

Permutando cíclicamente, se obtiene el número más pequeño, resultando en un código de cadena invariante a rotación. Éste código de cadena normalizado, es llamado *número de forma* y no es invariante a escalamiento.

El código de cadena es de dimensiones grandes y sensitivo a ruido, comúnmente se usa como entrada para un análisis de más alto nivel, por ejemplo para encontrar la curvatura de la forma.



**Figura 3.10:** Borde obtenido sobreponiendo una rejilla, cuyo código de conectividad 8 es: 0, 1, 7, 0, 0, 7, 5, 4, 4, 4, 4, 4, 2, 2 y su punto de inicio es el más a la izquierda.

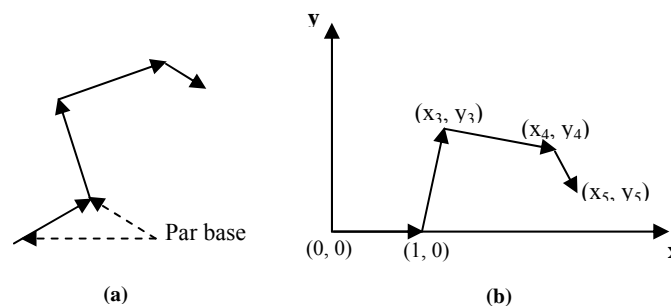
### 3.1.2.2 Descomposición del polígono

En trabajos como [61, 62] el borde de la forma se parte en segmentos de línea para obtener una aproximación de polígono. Los vértices de los polígonos son utilizados como primitivas. Las características para cada primitiva, se expresa como una cadena de cuatro elementos: ángulos internos, distancia del siguiente vértice, y sus coordenadas  $x$  y  $y$ .

Este tipo de característica no es invariante a traslación, a escalamiento o a rotación. La diferencia de similitud entre las dos formas será la distancia de las dos cadenas de características. Para tener una mejor eficiencia y robustez, sólo se escogen un número determinado de vértices para

cada forma. Se genera una colección de características pertenecientes a todos los modelos en la base de datos. Para el índice de característica, éstas características se almacenan en un árbol binario. Para la correspondencia entre las formas, primero de una consulta de forma, se obtienen los datos de características y estos datos se buscan en el árbol de índices, si se encuentra un modelo similar a las características, se retornan todas las imágenes que estén asociadas con el modelo recuperado. Después la similitud entre el modelo recuperado y la forma, se basa en la distancia entre las dos cadenas de primitivas [62].

Mehrotra y Gary en [63], representan una forma como una cadena de vectores. En una forma, se detectan una serie de puntos para obtener una aproximación poligonal del borde de la forma. Si esta forma tenía  $n$  puntos, se escogían un par de puntos para formar vectores base, el cual después se normaliza como un vector unitario a lo largo del eje  $x$  y los demás puntos se transforman a este sistema de coordenadas (ver Figura 3.11.). La forma se representa por el conjunto de coordenadas  $(x_1, y_1), \dots, (x_n, y_n)$  donde  $(x_i, y_i)$  son las coordenadas normalizadas del  $i$ -ésimo punto.



**Figura 3.11: Vectores cadena: (a) Descripción original del borde (b) Descripción normalizada del borde [62].**

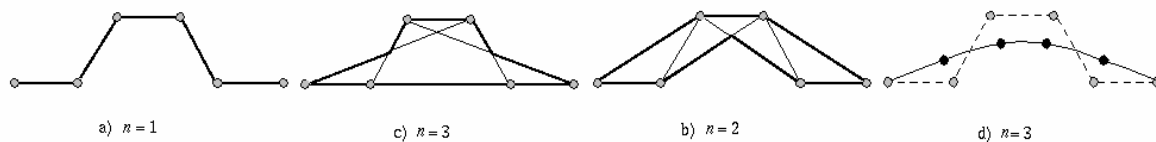
Además se añaden 4 parámetros al conjunto de coordenadas para crear el vector final de características:  $S$  que representa la escala,  $T_x$  y  $T_y$  que representan la traslación y  $\theta$  que representa el ángulo del vector base. La similitud entre dos características es obtenida por la distancia Euclidiana. Los vectores de características se organizan en un árbol  $B$  de dimensión  $k$ . En una consulta, se busca el índice de las características de la forma a consultar y se retorna la lista de las formas que tienen características similares [63].

### 3.1.2.3 Representación B-spline

Es una representación de curvas usando un conjunto de interpolaciones polinomiales para obtener curvas suavizadas. Las formas de estas curvas polinomiales, están relacionadas con su polígono base, como una cadena de vértices dando una representación poligonal de la curva. Los B-spline de tercer orden son los más comunes porque éste es el orden más bajo, el cual incluye

cambios de curvatura. Este método tiene buenas propiedades de representación y son fáciles de calcular:

- Cambian su forma menos que la de su polígono y no oscilan entre otros puntos de representación. Una curva de *Spline* siempre está posicionada dentro de un polígono convexo de  $n+1$  lados, para un *B-spline*, es de  $n$ -ésimo orden.
- Si un vértice del polígono base cambia su posición, ocurre un cambio de la curva *spline*, pero solamente en la vecindad de dicho vértice.
- Los métodos para la correspondencia de regiones representadas por *splines*, se basan en la búsqueda directa de los datos de la imagen original.



**Figura 3.12: Splines de orden n. (a), (b), (c), polígonos convexos  $n+1$ , para un B-spline de  $n$ -ésimo orden; (d) splin de 3er orden.**

Cada parte de una curva B-spline cúbica es un polinomio de tercer orden, lo que significa que su primera y segunda derivadas son continuas. Los B-splines están dados por:

$$x(s) = \sum_{i=0}^{n+1} v_i B_i(s) \quad (3.15)$$

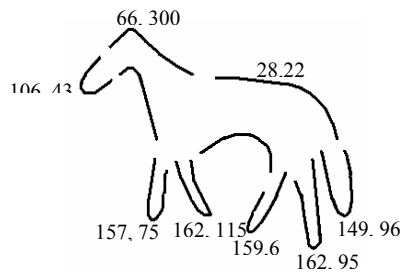
### 3.1.2.4 Descomposición de curva suavizada

En trabajos como [64], la curvatura de los puntos de intersección 0 de un borde suavizado por un filtro Gaussiano, se utiliza para obtener primitivas de comparación llamadas “tokens”. La característica de cada *token* es su curvatura máxima y su orientación. La similitud entre dos tokens es medido por el peso de la distancia Euclidiana. Este método no es invariante a rotación ya que está basado en la orientación de la curva, se utiliza un árbol-M para indexar los tokens dentro de la base de datos de características. Al dar una consulta de forma, se obtienen los N tokens de la forma a consultar y se encuentran los tokens similares, atravesando el árbol índice N veces. El conjunto de tokens recuperado que tienen el mismo identificador, son formas similares potenciales. Posteriormente se realiza la correspondencia entre modelos, lo cual requiere  $O(MN)$  operaciones (M y N son los tokens de dos formas a corresponder.) La eficiencia de la correspondencia depende del número de tokens para cada forma.



### 3.1.2.5 Método de espacio de escala.

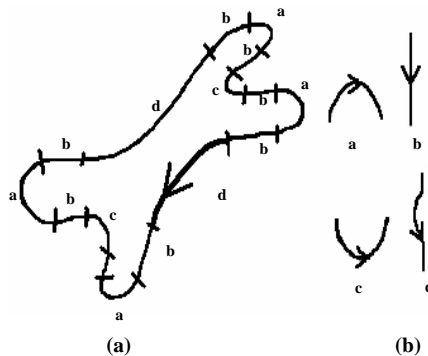
Trabajos como el de Dudek y Tsotsos [65] analizan la forma en el espacio de escala, utilizando un esquema de correspondencia modelo a modelo, las primitivas de las formas, se obtienen de las técnicas de ajuste de curvatura. Un segmento que es descriptor consiste de la longitud del segmento, de la posición ordinaria y el valor de ajuste de la curvatura, extraído de cada primitiva. Se crea una cadena de un descriptor de segmento para describir la forma. Así, para dos formas A y B representadas con sus descriptores de cadenas  $A = (S_1^A, S_2^A, \dots, S_N^A)$  y  $B = (S_1^B, S_2^B, \dots, S_N^B)$ , se usa una correspondencia modelo a modelo, usando programación dinámica para encontrar el grado de similitud entre las dos formas. Para que el proceso sea más robusto, las características de la forma se trasladan al espacio de escala de curvatura, donde las formas puedan ser correspondidas en diferentes escalas. Los descriptores no son invariantes a escala debido a la inclusión de la longitud en los segmentos descriptores [65].



**Figura 3.13:** División en “tokens”, los números corresponden a la curvatura y a la orientación del token, respectivamente [65].

### 3.1.2.6 Análisis sintáctico

En estos métodos la forma es representada por un conjunto predefinido de primitivas, llamado “libro de código” y las primitivas son llamadas “palabras de código” [66], como se muestra en la Figura 3.14.



**Figura 3.14:** (a) Descripción estructural de la forma de un cromosoma; (b) palabras de código, que son parte del libro de códigos. El cromosoma puede ser representado como una cadena

$S = dbabcbabdb \text{ } abcbab$  [46].

La correspondencia entre las formas, se hace encontrando el número mínimo de operaciones para modificar y convertirla en otra forma. Existe otro método que es el de representar una gramática de cadena, donde las primitivas son interpretadas como un alfabeto de alguna gramática, que es un conjunto de reglas de sintaxis que definen la generación de sentencias formadas por los símbolos del alfabeto.

El conjunto de sentencias generadas por una gramática  $G$  es llamado lenguaje  $L(G)$ . Las sentencias son cadenas de símbolos, que representan patrones y los lenguajes son los patrones de clases. De una sentencia representando una forma desconocida, se tiene que decidir en que lenguaje, dicha forma representa una sentencia válida [46].

El análisis sintáctico de formas, intenta simular la naturaleza de los sistemas de visión humana, lo cual no es práctico porque no siempre es posible inferir un patrón de gramática que pueda generar solamente patrones válidos.

La diferencia entre los métodos sintácticos y los métodos como el *código de cadena*, *descomposición de curva*, *aproximación del polígono* y *espacio de escala*, es que las primitivas son predefinidas y no adaptadas a la base de datos de imágenes, la aproximación predefinida es conveniente si se conocen en general los perfiles del borde de la forma. Sin embargo, el alfabeto será dependiente de la aplicación [66].

### 3.1.2.7 Invariantes de forma

Pueden ser vistos como una aproximación estructural, porque también representan formas basadas en primitivas de los bordes. Las técnicas que utilizan invariantes, pretenden representar propiedades de las configuraciones de los bordes, los cuales permanecen sin cambio, con una serie de transformaciones apropiadas.

La teoría de invariantes se basa en una serie de transformaciones que puede componerse e invertirse.

Se le llama invariante respecto al número de características utilizadas para definirlo.

Entre los invariantes más comunes, se encuentran [66]:

1. Invariantes *geométricos*, tales como el cruce de proporción, la longitud de la proporción, la distancia de la proporción, ángulo y área triángulo, invariantes de puntos coplanares.
2. Invariantes *algebraicos*, tales como determinantes, valores propios, trazo.
3. Invariantes *diferenciales*, como la curvatura, torsión y la curvatura Gaussiana.

Los invariantes geométricos y los algebraicos, son recomendables para situaciones donde los bordes se pueden representar con líneas rectas o curvas algebraicas, para los casos en que esto no es posible, se pueden utilizar los invariantes diferenciales.

Existen problemas al usar invariantes para representar una forma, comúnmente son derivados de transformaciones geométricas de la forma, son muy sensitivos al ruido y a errores en los bordes. Para la mayoría de los métodos invariantes, la parte difícil es la correspondencia, que generalmente toma técnicas de partes y relaciones entre ellas y requieren un subgrafo de correspondencia del orden NP-completo.

En trabajos recientes como el de Kliot y Rivlin [67], se propone el uso de firmas de invariantes para la descripción de forma. Los invariantes geométricos como la longitud, ángulo, área de la proporción, intersección de la proporción y longitud de la proporción, se utilizan para derivar firmas de invariantes para cada curva del borde (llamadas firmas multivaluadas). Esta firma se asigna a una matriz la cual es utilizada para hacer la correspondencia entre dos curvas de dos formas. Para incrementar su eficiencia, se construye la firma de histograma y se utiliza para el proceso de correspondencia inicial, después de esto, las curvas están sujetas a la matriz de correspondencia.

Otros trabajos como el realizado por Squire y Caelli [68], utilizan firma de invarianza para la descripción de formas. La firma de invarianza es una función de densidad de probabilidad, derivada de los bordes representados por un conjunto de variables de curvas algebraicas. Las firmas son invariantes a rotación, escala y traslación. Después se construye un histograma de los cálculos de la firma de invarianza, para hacer la descripción y correspondencia de la forma. El histograma después se pasa a una red neuronal, para su clasificación.

### **3.1.3 Ventajas y Desventajas de los descriptores globales y estructurales.**

Los descriptores globales no son tan precisos como descriptores de formas, por lo que necesitan ser combinados con otros descriptores de formas. Los métodos de correspondencia por elasticidad y de Wavelets son complejos para implementarlos. Los métodos de autoregresión requieren operaciones de matrices, además de que es difícil asociar los descriptores de autoregresión con su significado físico. La transformada de Fourier es fácil de implementar, su descriptor, aparte de ser compacto su correspondencia es simple. El descriptor de momento del borde es similar al descriptor de Fourier y es fácil de adquirir, aunque solo los primeros descriptores de momento de más bajo orden tienen interpretación física.

Los descriptores estructurales presentan un inconveniente, la generación de primitivas y características. Debido a que no hay una definición formal para un objeto o forma, es desconocido el número de primitivas que se requieren para cada forma. La efectividad de aplicar estos métodos depende del conocimiento previo que tengamos de los bordes de la forma. Este tipo de métodos trabaja para casos particulares y no tanto generales, ya que es imposible conocer los tipos de

primitivas en una base de datos genérica de formas. Otro inconveniente de los métodos estructurales es su complejidad computacional, especialmente la correspondencia, que es muy compleja. A diferencia de los métodos globales, la correspondencia en los métodos estructurales, utilizan medidas de similitud no métricas, es por eso que se pueden hacer correspondencias parciales, utilizando un subgrafo de correspondencia.

Además pueden fallar al capturar características globales de la forma, las cuales son importantes para la representación de la forma. Debido a que su representación estructural no conserva la estructura topológica de un objeto, variaciones del borde del objeto pueden causar cambios significativos.

### **3.1.4 Métodos seleccionados**

Después de analizar varias técnicas de representación de forma para este trabajo de tesis, se decidió utilizar los siguientes métodos globales, en base a las características y ventajas que presentan, el método de *evolución discreta de la curva* (Sección 3.1.1.10) en combinación con el método de *representación de espacio tangencial* (3.1.1.9). Se eligieron la combinación de dos métodos globales ya que estos métodos son fáciles de implementar y trabajan mejor combinándolos con otro método. A diferencia de los demás métodos, no ocupan mucho espacio para su creación y por lo tanto se procesan rápidamente, por ejemplo no ocupan un libro de palabras como los métodos sintácticos, ni son tan complejos como las transformaciones espectrales, son más precisos que las firmas de forma o como el código de cadena, el cual tiene restricciones de invarianzas.

El método de evolución discreta de la curva presenta características que lo llevan a ser un método sencillo, rápido y eficaz de implementar, además de tener las ventajas de ser invariante a rotación, escala y traslación, otra ventaja es que reduce la complejidad de las formas de los objetos, eliminando información irrelevante como se muestra en la sección 3.1.1.10. Se optó además, por combinar éste método de evolución de curva con el método de representación de espacio tangencial, éste último facilita la comparación de formas, representándolas de una manera fácil y obteniendo buenos resultados.

La técnica de representación de espacio tangencial también presenta invarianzas a rotación, escalamiento y traslación, además su ejecución es de forma rápida y da como resultado una gráfica sencilla 2D escalonada como se vio anteriormente en la Sección 3.1.1.9.

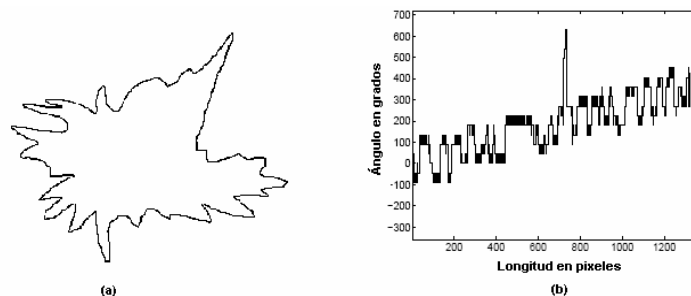
## 3.2 Algunos trabajos utilizando descriptores de formas

En [69], se basan en la forma, obtienen características como: partes rugosas, forma ovalada o circular, etc. Analizando la forma de un contorno distinguen entre tumores malignos y benignos. El método que proponen es obtener un índice de especulación, un índice de convexidad y una medida para la dimensión fractal, partiendo del cálculo de la función de ángulo de giro (Turning Angle Function). Un contorno de dos dimensiones, se convierte a una dimensión, obteniendo la distancia de cada punto del contorno hacia el centroide del contorno. Este vector calculado, se utiliza para derivar las dimensiones fractales y representar la complejidad del contorno.

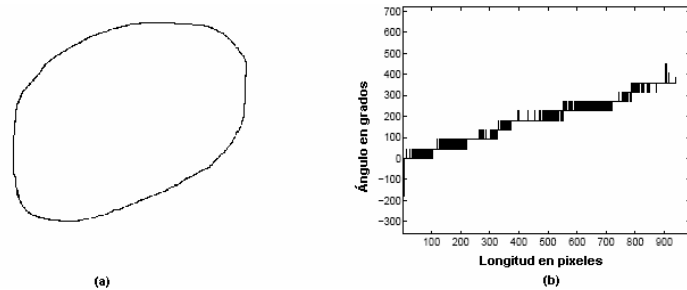
La función de ángulo de giro del contorno  $s$ , de una imagen dada  $T_s(s(n))$ , definida como el ángulo, obtenido en dirección de las manecillas del reloj, entre la tangente del contorno en  $s(n)$  y en referencia con el eje  $x$ , se expresa como una función de contorno con índice  $n$ . La función de ángulo de giro también es conocida como función tangente. En las Figuras 3.15 3.16, se muestran las funciones de ángulo de giro, para los contornos de tumores benignos y malignos, respectivamente[69].

Para el contorno convexo, como se muestra en Figura 3.16, la función de ángulo de giro es en general una función incremental (Figura 3.16 (b)). Para un contorno con secciones cóncavas y convexas, la función de ángulo de giro, se empieza a decrementar al iniciar la sección cóncava y se mantiene así, hasta que la dirección de la tangente cambia de dirección en el inicio de una sección convexa. En Figura 3.15 una espícula en el contorno se relaciona a una sección delimitada por un par de ángulos sucesivos eliminados.

Otras características de complejidad que se pueden obtener son, la densidad, la concavidad fraccional, índice de especulación, factores del descriptor de Fourier, dimensión fractal, momentos, estadísticas de la longitud de cuerda, y transformadas Wavelets, las cuales se han utilizado para reconocer patrones en el área del cáncer de mama [69].



**Figura 3.15: (a) Tumor maligno con un contorno espiculado, con segmentos cóncavos. (b) Función ángulo de giro del contorno (a) [69].**



**Figura 3.16: (a) Tumor benigno con un contorno convexo y suavizado (b) Función ángulo de giro del contorno (a) [69].**

Para evaluar su funcionamiento, se basaron en términos de eficiencia para la clasificación de tumores.

Resultados y experimentos que realizaron:

**Tabla 3.1: Resultados de pruebas.**

Casos probados	111
Tumores Benignos	65
Tumores Malignos	46

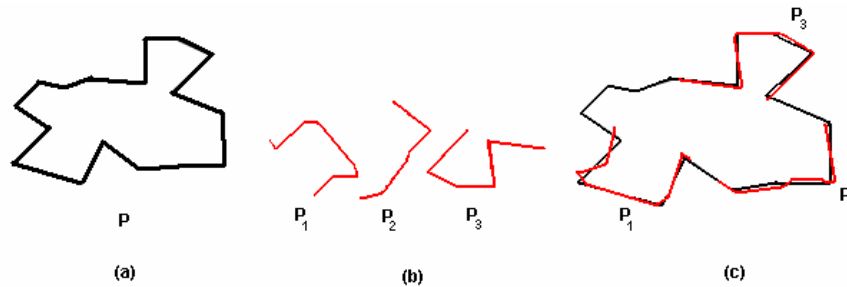
**Tabla 3.2: Eficiencia del método.**

Característica	Eficiencia
Índice de espicularidad	0.92
Índice de convexidad	0.93
Dimensiones fractales	0.91

Este trabajo propone métodos partiendo de la función de ángulo de giro (índice de especularidad, dimensiones fractales e índice de convexidad) para obtener características de una forma de contorno enfocadas a la clasificación de tumores.

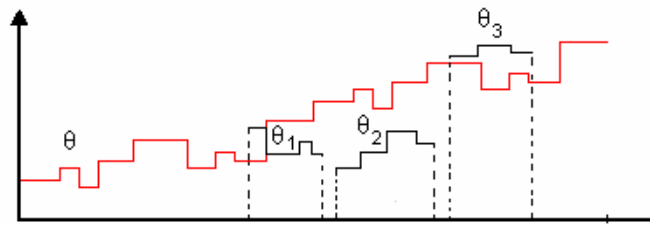
Otros trabajos ayudan a encontrar similitud entre formas, aún cuando una porción de la forma no es visible, o está distorsionada [70]. En este enfoque tiene la ventaja de que realiza menos trabajo haciendo una correspondencia parcial, en vez de hacer una global como comúnmente se realiza. Permite que se busquen ciertas partes de una figura dada. Las figuras polilíneas Figura 3.17 (b), que son figuras de una línea continua, compuesta de uno o más segmentos de líneas, pueden ser piezas del contorno de un objeto incompleto extraído de una imagen. Ellos se basan en la representación de la función de ángulo de giro tanto de los polígonos, como de las figuras polilíneas. Dados un conjunto de formas y una serie de consultas con figuras polilíneas, recuperan las formas del conjunto que mejor corresponda a la consulta. En este método tanto la base de datos de formas como las partes de la consulta son seleccionadas por el usuario. Su conclusión fue que el método de

correspondencia parcial, es mucho mejor que la ejecución de correspondencia global, en categorías difíciles de formas [70].



**Figura 3.17: Correspondencia (c) de un conjunto ordenado de figuras polilíneas (b) y un polígono (a) [70]**

La correspondencia se hace a través de un conjunto ordenado  $\{P_1, P_2, \dots, P_k\}$  de  $k$  polilíneas contra un polígono  $P$ . Se calcula que tan cerca está una figura polilínea, de encajar en alguna parte del borde de  $P$ , para esto las figuras de polilíneas se rotan y ajustan a lo largo de todo el polígono  $P$ , de tal manera que esas partes del borde de  $P$ , sean cubiertas por las  $k$  figuras polilíneas. Los puntos de inicio de  $P_1, P_2, \dots, P_k$ , son correspondidos con los puntos de los bordes de  $P$  en sentido de las manecillas del reloj. La medida de similitud, se basa en la función de ángulo de giro obteniendo su distancia correspondiente a los puntos coincidentes del polígono y las  $k$  figuras polilíneas, la suma de las distancias que sea más cercana a 0 será la mejor coincidencia posible entre las  $k$  figuras polilíneas y el polígono [70].



**Figura 3.18: Distancias  $d(P_1, \dots, P_k; P)$  entre las polilíneas  $P_1, P_2, P_3$  y el polígono  $P$ , cambiando tanto vertical como horizontalmente las funciones de ángulo de giro  $\theta_1, \theta_2, \theta_3$ , a lo largo de  $\theta$ .**

El tiempo de ejecución de su algoritmo es de  $O(km^2n^2)$ , siendo  $m$  el número de vértices del polígono  $P$ , y  $n$  el total de vértices en las  $k$  figuras polilíneas.

Las consultas realizadas para la evaluación de este método, representan un conjunto de partes discontinuas del borde de una figura, el proceso de correspondencia evalúa que tan cercana esta esa

parte de encajar en alguna parte del contorno de la forma. Para una consulta se tienen que especificar las partes de la forma a buscar, las cuales son generadas automáticamente partiendo de una forma dada.

La colección de formas que utilizaron para la evaluación de la recuperación, fue la base de datos de siluetas de formas Core Experiment “CE-Shape-1” part B, que consiste de 1400 imágenes, con 70 clases de formas y 20 imágenes por cada clase. Probaron con 20 clases y obtuvieron los siguientes resultados [70].

**Tabla 3.3: Resultados obtenidos en [70]**

Clase	Promedio de porcentaje de correspondencia de la consulta en la clase especificada
1	55 %
2	50 %
3	37.5 %
4	37.5 %
5	50 %
6	52.5 %

Otro método para la indexación y recuperación de formas, es el método basado en rejilla [71]. La forma se mapea a una rejilla de tamaño fijo y es justificada a la esquina superior izquierda. Después se recorre esta rejilla de izquierda a derecha, asignando un “1” a cada celda de la reja que es parcial o totalmente cubierta por la forma y asignando un “0”, a la celda que no se cubre por la forma, obteniendo una secuencia de 0’s y 1’s. Esta secuencia puede ser utilizada para la representación de la forma, que solamente es robusta a translación. Se utiliza el eje mayor para hacer su representación invariante a translación y a escalamiento. El eje mayor de la forma es el segmento de línea mas largo, cuyos puntos extremos están más alejados el uno del otro, lo que representa un costo computacional muy alto. En la mayoría de las formas el eje mayor es único. Para encontrar el eje mayor AB, se inicia en un punto  $C_0$  del contorno, con el cual se calculan todas las distancias hacia los otros puntos del borde. La distancia mayor para  $C_0$ , se compara con las distancias mayores obtenidas para  $C_1, \dots, C_n$ .

Para obtener la diferencia entre dos formas  $d(A, B)$ , es necesario contar las celdas de la rejilla de figura A que contienen 1’s y que en la figura B en la misma posición de la celda no contenga un 1. También se calcula la excentricidad que se define como la longitud del radio del eje mayor y la longitud del eje menor. Dicho eje mayor debe ser perpendicular al eje menor, para formar así un rectángulo que encierre el borde de la forma.



La distancia  $d(A, B)$  entonces se puede obtener comparando los números binarios de todas las formas, con los números binarios de la forma a consultar. La mejor correspondencia es la mínima distancia a la forma a consultar [71].

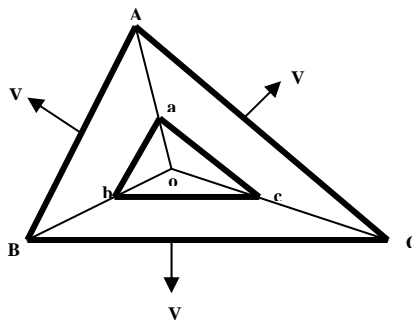
Dalong Li y S. Simske en [72] proponen un algoritmo basándose en la distribución de la distancia de la proporción. Siguiendo el algoritmo, se mide la transformación entre las formas a comparar. La desviación estándar de éste algoritmo se utiliza para representar la distribución. Éste algoritmo es invariante a rotación, escalamiento y a rotación.

Dicho algoritmo se puede resumir en los siguientes pasos:

1. Se obtiene el contorno de la forma con  $N$  puntos.
2. Se calcula el centroide de los  $N$  puntos.
3. Se obtienen los vectores (rv1, rv2).
4. Se calculan las distancias proporcionales.
5. Se calculan las desviaciones estándar y se elige la mínima.
6. Se compara la distancia mínima para decidir si las dos formas son similares.

Para obtener los vectores en el paso 3, se saca la longitud de la distancia que existe desde el centroide de la forma hacia el borde de la misma. Se eligen intervalos para obtener el vector  $\{L_0, L_1, \dots, L_i, \dots, L_{n-1}\}$ , donde  $L_i$  es el  $i$ -ésimo radio desde el centroide de la forma. Se toman intervalos en dirección de las manecillas del reloj, iniciando en la dirección del eje de las  $x$ . Para que esta característica sea invariante a escalamiento, todas las longitudes de los radios se normalizan dividiéndolas entre la longitud mayor.

Dos formas se consideran similares si las longitudes entre sus vectores, difieren por un valor trivial [72].



**Figura 3.19: Similitud geométrica, para obtener longitudes de los radios  $\{oa / oA, ob / oB, oc / oC\}$ , con desviación estándar de 0.**

Para la Figura 3.19 es necesario rotar los vectores de distancia antes de hacer una comparación general. Para un vector dado  $\{a, b, c\}$ , existen  $2*N$  vectores posibles diferentes

$\{b, c, a\}; \{c, a, b\}; \{c, b, a\}; \{b, a, c\}; \{a, c, b\}$ . Estas comparaciones entre vectores son más factibles cuando dos figuras son diferentes; manteniendo el centroide uniforme, se puede encontrar la desviación estándar. Si sus radios no son proporcionales, entonces se puede formar una nueva forma. La idea es tener a la desviación estándar como una medida de transformación para encontrar la similitud entre las formas.

Para la evaluación utilizaron las formas de la base de datos de SQUID, con 1100 figuras de pescados, que aún no han sido clasificados. Si se tienen 20 figuras ejemplo, entonces se necesitarán  $20 \times 1100$  comparaciones, las cuales necesitarían realizarse manualmente ya que la base de datos aun no está clasificada [72].

Dada una consulta  $Q$ , la precisión de sus pruebas, está basada en:  $Pr\ ecisión = C / R$ , donde  $R$  es el conjunto de respuestas recuperadas y  $C$ , es la respuesta correcta en  $R$ .

**Tabla 3.4: Resultados obtenidos en [72]**

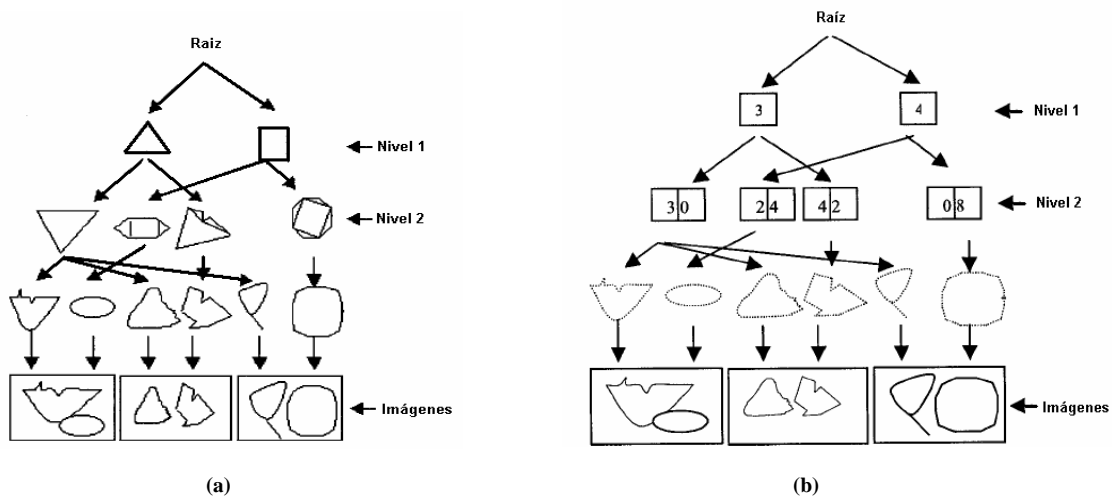
Precisión	Tamaño del conjunto de imágenes
0.4-0.8	5
0.2-0.95	110

Tao Shen [73], propone un esqueleto para una recuperación rápida basada en formas. Esta estructura construye una jerarquía de aproximaciones de formas de objetos. Múltiples formas en un nivel más bajo pueden ser mapeadas a una sola forma de un nivel más alto, la jerarquía sirve para particionar la base de datos a varios niveles de granularidad. Dada una consulta, se busca solo en las rutas involucradas de la jerarquía, “podando” una porción de la base de datos. Su método se basa en el mapeo de ángulos, que transforma una forma de un nivel a otra de nivel más alto. Reemplaza algunos de los bordes de una forma, por un número menor de lados en el borde, basándose en los ángulos entre las líneas del contorno, lo cual reduce la complejidad de la forma original. Para esto proponen dos estructuras jerárquicas que facilitan la recuperación, una estructura es la de particionamiento jerárquico en representación de formas (Hierarchical Partitioning on Shape Representation - HPSR), que usa la información de la representación de formas como llave de indexación; la otra estructura es la de particionamiento jerárquico en vectores de ángulos (Hierarchical Partitioning on Angle Vector - HPAV) que captura la información del ángulo en la representación de la forma [73].

Cada forma de objeto se representa como segmentos de líneas. En esta estructura, cada una de las formas en los niveles altos sirve como una aproximación para una forma en un nivel más bajo. Éste método reduce la secuencia de segmentos conectados y con ello la dimensión de la forma. En los niveles más altos de la estructura hay menos particiones y existen más formas sin cambios

compartiendo la misma aproximación, en los niveles más bajos hay más particiones y menos número de formas comparten la misma representación.

En una consulta, se tiene un menor costo al comparar dos formas con un número menor de vértices, que al comparar dos figuras con una gran cantidad de vértices. Para esto se utiliza el método de Mapeo de ángulo, para mapear una representación de forma de dimensiones grandes en una representación de menor dimensión. En base a que los ángulos agudos y los segmentos de líneas más largas, representan información más importante sobre la forma, se eliminan los ángulos mayores a  $90^\circ$  y por consecuencia también los segmentos que tendrían menor relevancia. Lo que definiría un ciclo hasta que la condición de nivel de representación en la jerarquía se cumpla, definiendo este número como  $N$ . Para almacenar los datos referentes a los ángulos de las formas, utilizan la estructura HPAV, que contiene información únicamente de los ángulos [73].



**Figura 3.20:** Árboles de indexación de las estructuras (a) HPSR y (b) HPAV, en la base de datos, para 3 imágenes [73].

Para la correspondencia, utilizaron la función de ángulo de giro, buscando la menor distancia entre las figuras. Sus pruebas consistieron de 5000 imágenes, cada una de las cuales con un promedio de 15 objetos y las consultas se realizaron con 15 objetos.

Estos son los resultados que obtuvieron:

**Tabla 3.5:** Resultados obtenidos en [73]

Estructura	Porcentaje de eficiencia en su búsqueda de similitud de formas
HPSR	77 %
HPAV	93 %

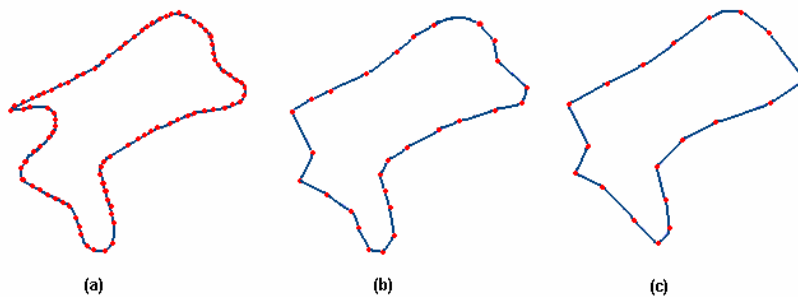
Muchas aplicaciones que abarcan comparaciones de formas, se enfocan a la medicina. Trabajos como [74] implementan la recuperación basada en formas de imágenes de rayos x sobre vértebras. Representando un reto ya que éstas se parecen.

Implementaron la técnica de evolución de la curva de un polígono, para reducir el número de puntos, obteniendo sólo los puntos significativos que caracterizan el contorno de la forma. Las coordenadas  $x$  y  $y$  del borde se utilizan para calcular la función de ángulo de giro, normalizando la longitud de la curvatura. Posteriormente utilizando descriptores de Fourier se calcula la medida de similitud.

La función de curvatura sirvió para analizar la curva del contorno y determinar qué puntos son relevantes al comparar segmentos. Utilizando el método de evolución de la curva se remueven pequeñas variaciones que existen y las características menos significativas para representarlas en la curva de la función de ángulo de giro. Después con los descriptores de Fourier se mide la similitud, ya que presentan invarianza con respecto al cambio del punto de inicio de la curva del polígono.

Desarrollaron una variante para el algoritmo de evolución de la curva, para preservar las características importantes del contorno de una forma y eliminar las características insignificantes, como lo son los segmentos de líneas.

Un alto valor de relevancia, significa que el vértice tiene una contribución mayor para la curva. En cada iteración, el vértice con menor medida relevancia, es eliminado y se crea un segmento uniendo los vértices adyacentes [74].



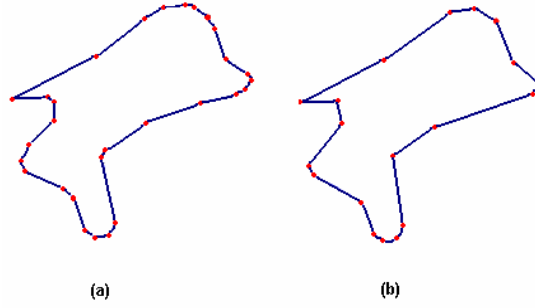
**Figura 3.21 (a) Figura original con 172 puntos, (b) reducción a 30 puntos, (c) reducción a 20 puntos, utilizando la ecuación 3.16**

$$K(S_1, S_2) = \frac{\beta(S_1, S_2)l(S_1)l(S_2)}{l(S_1) + l(S_2)} \quad (3.16)$$

En la Figura 3.21 se muestra que algunos puntos críticos son removidos por la Ecuación 3.16, para preservar estos puntos propusieron una nueva medida, definida en la Ecuación 3.17. Esta

ecuación remueve segmentos pequeños y líneas rectas, manteniendo características significativas, puntos convexos y cóncavos, eliminando vértices que tienen longitud corta y/o sus ángulos de giro se acercan a 180°.

$$K(S_1, S_2) = \frac{|\beta(S_1, S_2) - 180|l(S_1)l(S_2)|}{l(S_1) + l(S_2)} \quad (3.17)$$



**Figura 3.22 (a) Polígono reducido a 30 puntos y (b) 20 puntos, usando Ecuación 3.17.**

A partir de estos polígonos, se calcula la función de ángulo de giro normalizando el eje de las  $x$ . El punto de inicio no es invariante al cambio, por lo que utilizaron la propiedad de fuerza del espectro [74]. Definida por:

$$A_n = \sqrt{a_n^2 + b_n^2}, \quad F_{jk} = j * \tan^{-1}(b_k / a_k) - k * \tan^{-1}(b_j / a_j) \quad (3.18)$$

donde  $a_n$  y  $b_n$  son los coeficientes para cada componente de frecuencia:

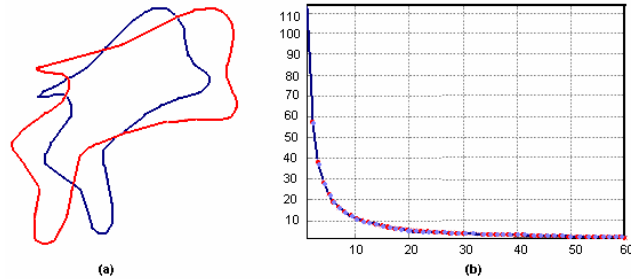
$$a_n = -\frac{1}{n\pi} \sum_{k=1}^m \theta_k \sin \frac{2\pi n \lambda_k}{L}, \quad b_n = -\frac{1}{n\pi} \sum_{k=1}^m \theta_k \cos \frac{2\pi n \lambda_k}{L} \quad (3.19) \text{ y } (3.20)$$

El proceso de correspondencia calcula los descriptores de Fourier definidos como:

$$\Theta(l) = \mu_0 + \sum_{n=1}^{\infty} (a_n \cos nl + b_n \sin nl) \quad (3.21)$$

Finalmente el proceso de correspondencia se realiza mediante la sumas de las diferencias de los descriptores de Fourier:  $D_A[c, c'] = \sqrt{\left(\sum |A_k - A'_k|^2\right)}$  para la amplitud.

La similitud entre las dos curvas esta dada por  $S_A[c, c'] = 100 - D_A$ . Un resultado de 100, significa que las curvas son idénticas y una resultado de 0, significa que las curvas son muy poco parecidas [74].



**Figura 3.23 (a) Forma original en color azul, forma más larga escalada 1.2 veces y girada 20°, en rojo. (b) Descriptores de Fourier de las funciones de ángulo de giro en (a) [74].**

**Tabla 3.6: Medidas de similitud con alteraciones [74].**

Cambios de la forma original	Medida de Similitud
20° de rotación	99.9999
Escalamiento del 1.2	99.9999
Punto de inicio: 10	100.0
Combinación de las tres anteriores	99.9999
Un punto trasladado a coordenada (20,20)	98.3516

Para su evaluación seleccionaron 20 formas aleatorias, las formas que eran similares, las agruparon por categoría. Evaluaron la ejecución, examinando el número de imágenes recuperadas para un mismo grupo. Seleccionando una a la vez, se compara contra el resto de las formas en la base de datos. Reportaron un resultado preciso, en donde dos de sus 20 imágenes reportaron mayores aciertos y los demás se agrupaban por números cercanos.

Zhang y G. Lu, en [58] utilizan firmas de forma a través de los descriptores de Fourier. En su trabajo, las formas pueden ser descritas como curvas cerradas. Su base de datos esta formada de siluetas de objetos y/o de esquemas que dibuja un usuario. Trabajaron con 4 firmas para las formas, la distancia hacia el centroide, coordenadas complejas, la curvatura y la función acumulativa de ángulo. Estos datos después son utilizados por los descriptores de Fourier.

Teniendo que las coordenadas del borde de una forma están definidas como:  $(x(t), y(t)), t = 0, 1, \dots, L - 1$ .

Para obtener las distancias hacia el centroide  $r(t)$ , que es invariante a traslación, la distancia de los puntos del borde al centroide  $(x_c, y_c)$  que es el promedio de las coordenadas de la forma, se define por:

$$x_c = \frac{1}{L} \sum_{t=0}^{L-1} x(t), y_c = \frac{1}{L} \sum_{t=0}^{L-1} y(t), \quad (3.22)$$

Teniendo que

$$r(t) = \sqrt{([x(t) - x_c]^2 + [y(t) - y_c]^2)} \quad (3.23)$$

Las coordenadas complejas se obtienen restando el centroide a cada coordenada compleja:

$$z(t) = [x(t) - x_c] + i[y(t) - y_c] \quad (3.24)$$

La curvatura que es invariante a traslación y a rotación, representa la segunda derivada del borde y la primera derivada de la tangente del borde [58]. Su función está definida por la diferencia entre ángulos sucesivos del borde calculados en  $w$  y se representa por:

$$K(t) = \theta(t) - \theta(t - 1) \quad (3.25)$$

$$\text{donde } \theta(t) = \arctan \frac{y(t) - y(t - w)}{x(t) - x(t - w)} \quad (3.26)$$

La función angular acumulativa, es la cantidad de ángulos de giro entre la posición de inicio  $z(0)$  y posición  $z(t)$  del borde de la forma:

$$\psi(t) = \varphi\left(\frac{L}{2\pi}t\right) - t \text{ donde } \varphi(t) = [\theta(t) - \theta(0)] \bmod(2\pi). \quad (3.27), (3.28)$$

Posteriormente los descriptores de Fourier, utilizan estas firmas para analizar la forma. Los coeficientes de Fourier describen la forma en el dominio de frecuencia. Se obtienen un conjunto de

coeficientes para capturar las características generales de una forma [58]. Para una firma de alguna forma dada, normalizada a N puntos, la transformada discreta de Fourier esta dada por:

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \exp\left(\frac{-j2\pi nt}{N}\right), n = 0, 1, \dots, N-1, \quad (3.29)$$

Los coeficientes de Fourier están dados por  $u_n, n = 0, 1, \dots, N-1$ , denotados como  $FD_n$ . Para indexar las formas utilizando éstos coeficientes, utilizan los vectores de características:

Para los descriptores de coordenadas complejas:

$$f = \left[ \frac{|FD_2|}{|FD_1|}, \frac{|FD_3|}{|FD_1|}, \dots, \frac{|FD_{N-1}|}{|FD_1|} \right] \quad (3.30)$$

Para los descriptores de distancia al centroide y la curvatura:

$$f = \left[ \frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \dots, \frac{|FD_{N/2}|}{|FD_0|} \right] \quad (3.31)$$

Para los descriptores de función angular acumulativa:

$$f = [ |FD_0|, |FD_1|, \dots, |FD_{N/2}| ] \quad (3.32)$$

Para medir la similitud entre dos formas se tienen las diferencias entre los descriptores de características indexados  $f_m = [f_m^1, f_m^2, \dots, f_m^N]$  y los datos de la forma, indexados por su respectivo descriptor de Fourier  $f_d = [f_d^1, f_d^2, \dots, f_d^N]$  [58].

La distancia Euclidiana entre los vectores se obtiene:

$$d = \sqrt{\sum_{i=0}^{N_c} |f_m^i - f_d^i|^2} \quad (3.33)$$

donde  $N_c$  es el número necesario de ciclos para indexar la forma.



Para sus experimentos, utilizaron 70 formas poligonales y 25 formas de botella. Se crearon para las 95 formas, 4 diferentes imágenes con distorsiones de diferentes parámetros, y para las 70 formas se generó una forma rotada. En total se probó con 640 imágenes.

La precisión de su método está basada en el radio, que es la proporción entre el número de recuperaciones relevantes  $r$  y el total de formas  $n$ ,  $P = \frac{r}{n}$ . El número de llamadas al sistema está dado por el número total de formas relevantes recuperadas  $r$ , entre el total de formas relevantes en la base de datos  $m$ .  $R = \frac{r}{m}$ . Usaron 16 formas para hacer pruebas de consultas. El mejor resultado se obtuvo con el descriptor de distancia al centroide siguiéndole el de coordenadas complejas [58].

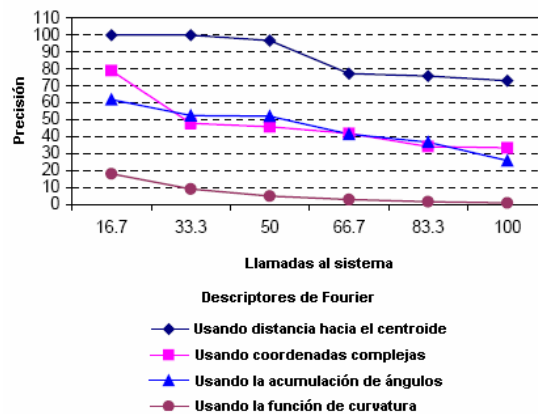


Figura 3.24 Eficiencia de pruebas para el sistema basado en formas de imágenes de rayos x sobre vértebras [58].

### 3.3 Conclusiones

La comparación de formas es un aspecto difícil, ya que su representación es más compleja que el color y la textura. Los métodos de descripción de formas se clasifican como métodos basados en contorno y métodos basados en región. Dentro de cada tipo existe otra subclasificación de métodos globales y estructurales, los métodos globales por sí solos no son tan precisos por lo que se tienen que combinar con otros descriptores de formas para obtener un mejor resultado, entre estos tenemos la autoregresión, la transformada de Fourier, momentos, etc. Los métodos estructurales, tienen gran capacidad para manejar oclusión en imágenes, permitiendo una correspondencia parcial, aunque sólo trabajan para aplicaciones específicas y no generales, además tienen una complejidad computacional muy alta. Debido a esto último, los métodos globales presentan más ventajas que los métodos estructurales.

En éste capítulo también se analizaron algunos trabajos que utilizan descriptores de formas, implementando diferentes técnicas en diferentes campos de aplicación como medicina, etc.

Como conclusión de este capítulo, después del estudio de varias técnicas para la representación de forma, como se menciona anteriormente, en éste trabajo se decidió utilizar los métodos globales: de *evolución discreta de la curva* en combinación con el método de *representación de espacio tangencial*.

En algunos trabajos analizados, que también se estudiaron en este capítulo, hacen uso de éstas técnicas y obtienen buenos resultados. Ambos métodos globales son fáciles de implementar.