

Capítulo 3

REPRESENTACIÓN Y CARACTERIZACIÓN DE OBJETOS

En este capítulo se estudia el caso particular de aplicación de la Red Preservadora de la Topología aplicada al espacio bidimensional, de modo que se realiza una representación de los objetos de la escena mediante un gas neuronal. En caso de que se desee obtener una representación del objeto, independientemente del tiempo requerido, se empleará una Neural Gas, que preserva mejor la topología de un espacio de entrada. Si, por contra, existen restricciones temporales o espaciales, se hará uso de una Growing Neural Gas.

Posteriormente, se estudia su aplicación al análisis de imágenes, realizando simplificaciones del grafo (red) obtenido para un mejor manejo de la información. Asimismo, se presentan diversas características que pueden ser extraídas del grafo original.

A continuación, se estudia la aplicación de la Red Preservadora de la Topología a la síntesis de imágenes, aprovechando la capacidad de compresión de la información que poseen los modelos auto-organizativos.

Finalmente, se presenta un ejemplo de aplicación de la adaptación de un gas neuronal a la forma de los objetos, empleándolo para la clasificación de estos a partir de sus contornos.

3.1 REPRESENTACIÓN DE UN OBJETO 2D CON UN GAS NEURONAL

La capacidad de preservación de la topología que poseen los gases neuronales va a ser empleada, particularmente en este trabajo, para la representación de objetos bidimensionales. Identificando los puntos de la imagen que pertenecen a los objetos, se posibilita la adaptación de la red a este subespacio de entrada, obteniendo una triangulación de Delaunay inducida por el objeto.

Sea un objeto $O = [A_G, A_V]$ que está definido por una apariencia geométrica A_G y una apariencia visual A_V . La apariencia geométrica A_G viene dada por unos parámetros morfológicos (deformaciones locales) y por unos parámetros posicionales (traslación, rotación y escalado):

$$A_G = [G_M, G_P] \quad (3.1)$$

La apariencia visual A_V está establecida por un conjunto de características del objeto como pueden ser el color, la textura, el brillo,...

En (Escolano, 1997) se indica que dados un dominio soporte $S \subseteq R^2$, una función de intensidad de imagen $I(x, y) \in R$ tal que $I : S \rightarrow [0, I_{max}]$, y un objeto O , su *campo de potencial normalizado* $\psi_T(x, y) = f_T(I(x, y))$ es la transformación $\psi_T : S \rightarrow [0, 1]$ que asocia a cada punto $(x, y) \in S$ el grado de cumplimiento de la propiedad visual T del objeto O por su intensidad asociada $I(x, y)$.

Considerando:

- El espacio de las señales de entrada como el conjunto de puntos de la imagen:

$$V = S$$

$$\xi = (x, y) \in S$$

- La función de densidad de probabilidad en función del campo de potencial normalizado obtenido para cada uno de los puntos de la imagen:

$$p(\xi) = p(x, y) = \psi_T(x, y)$$

se realiza el aprendizaje de un gas neuronal, ya sea una Neural Gas o una Growing Neural Gas, siguiendo el algoritmo general o cualquiera de sus variantes expuestas en el Capítulo 2. De modo que, realizando este proceso, se obtiene la red neuronal que preserva la topología del objeto O a partir de una cierta propiedad T del mismo. Es decir, a partir de la apariencia visual A_V del objeto se consigue una aproximación a su apariencia geométrica A_G .

De ahora en adelante, se denominará *Grafo Preservador de la Topología* al grafo no dirigido $GPT = \langle A, C \rangle$, definido por un conjunto de vértices (neuronas) A y un conjunto de aristas C que los conectan, que preserva la topología de un objeto a partir del campo de potencial normalizado considerado.

En la Figura 3.1 se muestra una descripción global del sistema de obtención del GPT de un objeto a partir de una escena, donde se observa que se pueden obtener GPT para diferentes propiedades T de los objetos sin modificar el algoritmo de aprendizaje de los gases neuronales, ya que obteniendo $\psi_T(x, y)$ distintas, se provocan adaptaciones de la red para cada uno de estos campos de potencial normalizados (Figura 3.2).

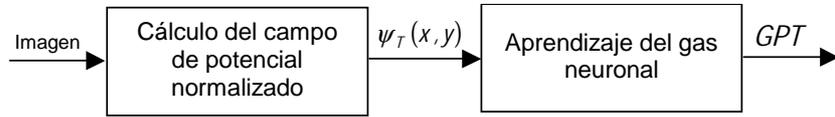


Figura 3.1. Descripción global del sistema de obtención del Grafo Preservador de la Topología de un objeto.

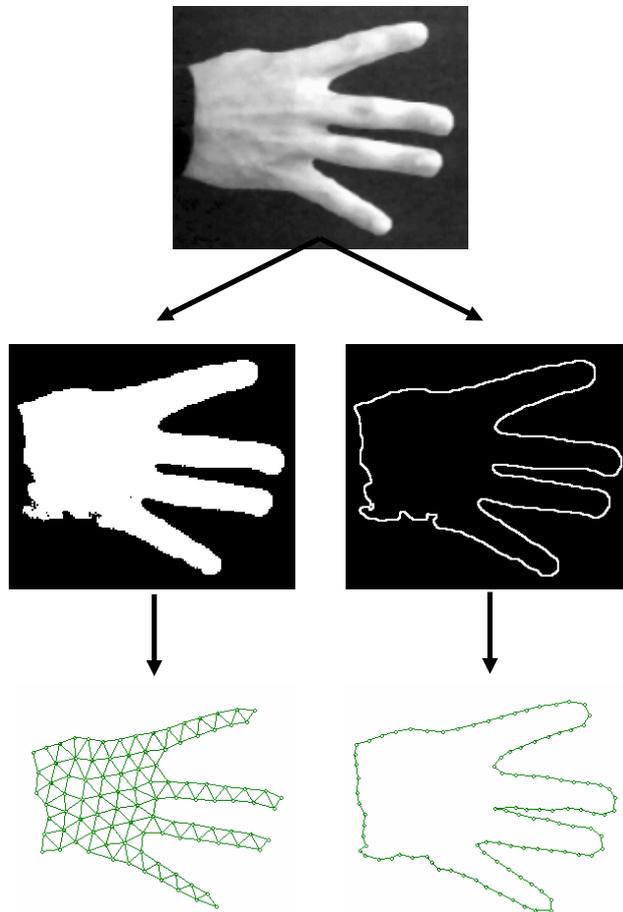


Figura 3.2. Diferentes adaptaciones del gas neuronal a un mismo objeto.

3.2 APLICACIÓN AL ANÁLISIS DE IMÁGENES

Tal y como se expresa en (González y Woods, 1996) la representación de una región u objeto admite dos posibilidades: en términos de sus características externas (contorno) o en términos de sus características internas (superficie). Generalmente, se elige una representación externa cuando el objetivo principal se centra en las características de la forma y una representación interna cuando el principal interés se centra en propiedades como el color, la textura,...

Una vez se extrae de la escena el objeto, el siguiente paso es el de descripción o caracterización del mismo. Las características extraídas deben ser, en la medida de lo posible, invariantes a modificaciones de tamaño, traslación y rotación; de modo que se pueda llevar a cabo una correcta clasificación o reconocimiento de los mismos al realizar una comparación con modelos almacenados en una base de conocimiento.

A continuación, se indica el empleo del aprendizaje de los *GPT* para extraer y caracterizar tanto contornos como superficies, así como posibles características que se pueden obtener del mismo que permitan mejorar reconocimientos posteriores.

3.2.1 EXTRACCIÓN Y CARACTERIZACIÓN DE CONTORNOS

Como se ha expresado anteriormente, una forma de representar los objetos es mediante su contorno. El contorno de un objeto puede ser obtenido mediante un *GPT* eligiendo de forma adecuada la propiedad visual T para la que se obtendrá el campo de potencial normalizado ψ_T de la escena.

Tomando $T = \nabla$ como la función gradiente, el campo de potencial normalizado ψ_{∇} indica, para cada uno de los puntos de la imagen I , si pertenece o no al contorno del objeto (Figura 3.3b).

A partir de este campo de potencial, se extrae el Grafo Preservador de la Topología GPT_{∇} que se adapta al contorno del objeto (Figura 3.3d). En ocasiones, cuando el contorno del objeto posee zonas con una curvatura importante, se crean aristas entre neuronas que no se encuentran próximas en el contorno (Figura

3.3c). Este problema es sencillo de resolver, eliminando del GPT_{∇} obtenido todos los ciclos indeseados que se produzcan.

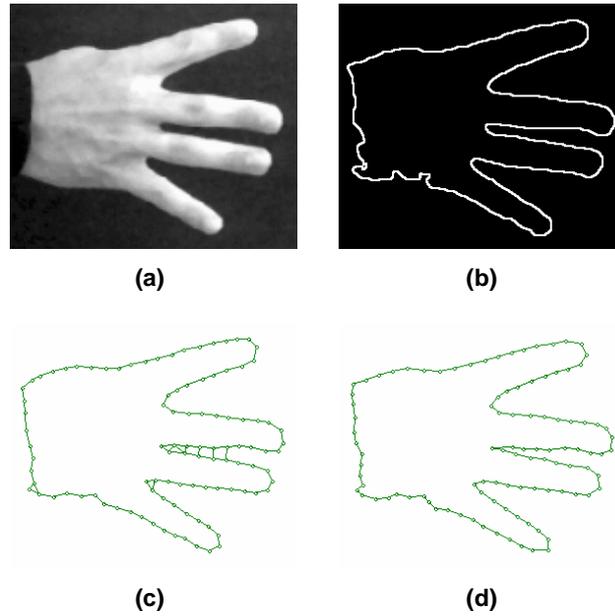


Figura 3.3. Extracción del GPT del contorno de un objeto.

El modelo es capaz de extraer, asimismo, los diversos contornos que pueda tener un objeto (Figura 3.4), debido a la capacidad de división que poseen los gases neuronales.

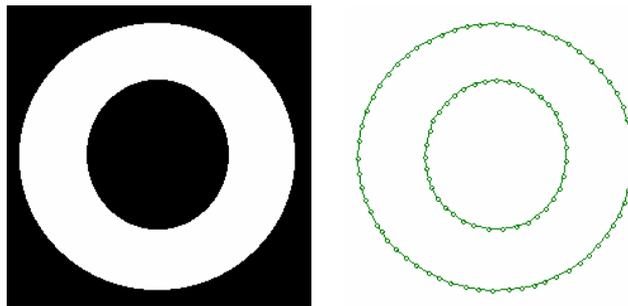


Figura 3.4. Obtención de los diversos contornos de un objeto.

Aunque, este GPT_{∇} podría ser utilizado como característica del objeto, es usual en la literatura encontrar diferentes características que son extraídas de los contornos para facilitar clasificaciones o reconocimientos posteriores (Loncaric, 1998) (Sonka et al., 1998) (González y Woods, 1996). Entre éstas se pueden encontrar algunos descriptores simples como su longitud, su diámetro, códigos de cadena, curvatura, firmas, momentos,...

DIÁMETRO

Dado un GPT_{∇} del contorno del objeto O , se obtiene su diámetro como la distancia máxima existente entre dos vértices (neuronas) cualesquiera del mismo:

$$\text{diámetro}(GPT_{\nabla}) = \max_{i,j \in A} (\|w_i - w_j\|) \quad (3.2)$$

Esta medida es invariante a traslación y rotación, aunque no lo es al escalado.

A partir del diámetro, otra característica que se puede obtener es la orientación del eje mayor del contorno, que es la línea que conecta ambos vértices.

LONGITUD

Dado un GPT_{∇} del contorno del objeto O , se obtiene su longitud como la suma de la longitud de todas las aristas que forman parte del mismo:

$$\text{long}(GPT_{\nabla}) = \sum_{\forall (i,j) \in C} \|w_i - w_j\| \quad (3.3)$$

Esta característica es invariante a traslación y rotación, pero no lo es al escalado de los objetos. Con el objetivo de hacer invariante esta medida también al escalado se ha transformado la expresión anterior, normalizando la longitud del contorno con respecto a su diámetro:

$$long_norm(GPT_{\nabla}) = \frac{long(GPT_{\nabla})}{diámetro(GPT_{\nabla})} \quad (3.4)$$

En la Tabla 3.1 se presentan los diámetros y longitudes obtenidos a partir del GPT_{∇} calculados para algunos de los objetos que están siendo cuestión de estudio. Para comprobar la efectividad de la longitud normalizada como medida invariante de la forma de los objetos, se han realizado tres instancias, a distintos tamaños, de cada uno de los espacios de entrada.

Objeto	Longitud	Diámetro	Longitud normalizada
Cuadrado	319 / 639 / 1280	112 / 225 / 452	2.83 / 2.84 / 2.83
Círculo	274 / 551 / 1100	89 / 179 / 358	3.06 / 3.06 / 3.07
Mano	660 / 1320 / 2647	136 / 272 / 546	4.85 / 4.85 / 4.84

Tabla 3.1. Longitud, diámetro y longitud normalizada de los contornos.

CURVATURA

Dado un GPT_{∇} del contorno del objeto O , la curvatura viene dada por el ritmo de variación de su pendiente. Como se observa en la Figura 3.5 la caracterización mediante curvatura del contorno de los objetos es bastante inestable, debido a que el GPT_{∇} realiza una aproximación mediante rectas del contorno; sin embargo, esto puede ser resuelto mediante alguna técnica de suavizado.

Para poder utilizar la curvatura como característica clasificadora, se debe tomar un mismo punto de partida en todos los casos. Para ello, por ejemplo, se puede comenzar el cálculo de la curvatura en el punto (neurona) que se encuentra más alejado del centro de masas del grafo.

El análisis de las curvaturas en los diferentes puntos del contorno permite dos modos de simplificación del mismo:

- detección de las esquinas del objeto y conexión de éstas,
- detección de las aristas o segmentos rectos del objeto.

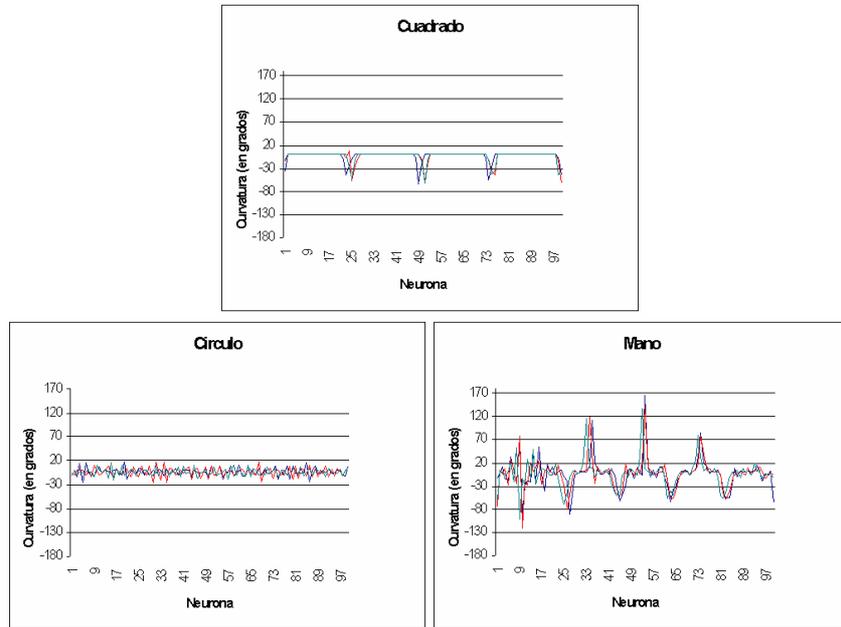


Figura 3.5. Curvaturas de los contornos.

OBTENCIÓN DE ESQUINAS

Las esquinas son extraídas a partir de la función de curvatura del objeto, detectando aquellos nodos donde la curvatura supera un cierto umbral. Estas neuronas determinan las esquinas del objeto y, mediante su conexión, se realiza una simplificación del GPT_{∇} (Figura 3.6).

OBTENCIÓN DE ARISTAS

La detección de aristas se consigue estudiando la curvatura acumulada entre neuronas sucesivas, de modo que cuando se supera una cierta curvatura umbral, la primera y última neurona de dicha sucesión determinan los extremos de la arista. Esta simplificación es dependiente del punto de inicio, por lo que se ha empleado la neurona con mayor curvatura como partida en la búsqueda de aristas.

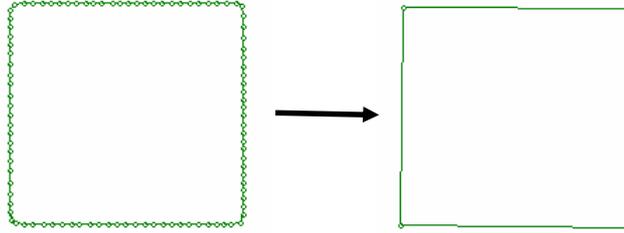


Figura 3.6. Simplificación del GPT_{∇} mediante la detección de esquinas.

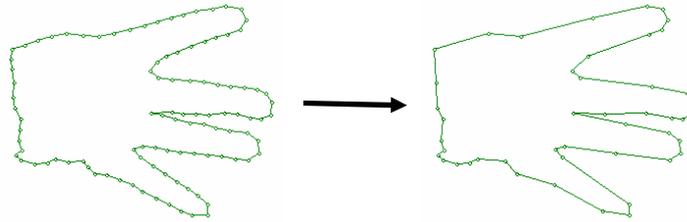


Figura 3.7. Simplificación del GPT mediante la obtención de aristas.

FIRMA

Dado un GPT_{∇} del contorno del objeto O , su firma puede ser calculada como la distancia de cada uno de sus vértices al centro del grafo. Al igual que en el caso del cálculo de la curvatura, se debe comenzar desde un mismo punto de partida. Si, además la distancia es dividida entre la máxima distancia de un nodo del GPT_{∇} al centroide, se obtiene una medida invariante a traslación, rotación y escalado (Figura 3.8).

CÓDIGOS DE CADENA

Los códigos de cadena se emplean para representar un contorno mediante una sucesión de segmentos de dirección especificada. Usualmente, esta representación se basa en segmentos de conectividad cuatro u ocho. En el caso de la Figura 3.9 se ha empleado conectividad ocho. Esta característica también depende del punto de partida, aunque existen diversos criterios para normalizar la cadena resultante (González y Woods, 1996).

MOMENTOS

Los momentos son uno de los métodos más utilizados para caracterizar contornos. Son calculados a partir de una representación unidimensional de los contornos, como pueden ser las curvaturas o las firmas.

Si se considera V como la función unidimensional elegida, k el número de nodos del GPT_{∇} , se obtiene el momento n -ésimo de V como:

$$\mu_n(V) = \frac{\sum_{i=1}^k (v_i - m)^n}{k} \quad (3.5)$$

donde

$$m = \frac{\sum_{i=1}^k v_i}{k} \quad (3.6)$$

En la Tabla 3.2 se representan los momentos de los contornos calculados a partir de las firmas no normalizadas de los objetos.

Objeto	m	μ_2	μ_3	μ_4
Cuadrado	92.78 / 92.90	103.61 / 105.49	525.07 / 524.27	21101 / 22000
Círculo	86.79 / 86.79	4,73 / 2.76	-5.20 / 0.82	50.39 / 15.46
Mano	90.04 / 87.62	1319.42 / 1394.14	-10277 / -9561	4048945 / 4575068

Tabla 3.2. Momentos de los contornos.

3.2.2 EXTRACCIÓN Y CARACTERIZACIÓN DE SUPERFICIES

La extracción del Grafo Preservador de la Topología GPT de la superficie de un objeto se puede obtener tomando como propiedad T para construir el campo de potencial cualquier

característica interna del objeto, como puede ser su textura, su nivel de gris, su color, su saturación,...; así como combinaciones de éstas.

Como resultado del aprendizaje del gas neuronal de este campo de potencial, se obtiene un GPT que realiza una esqueletización del objeto deseado (Figura 3.10).

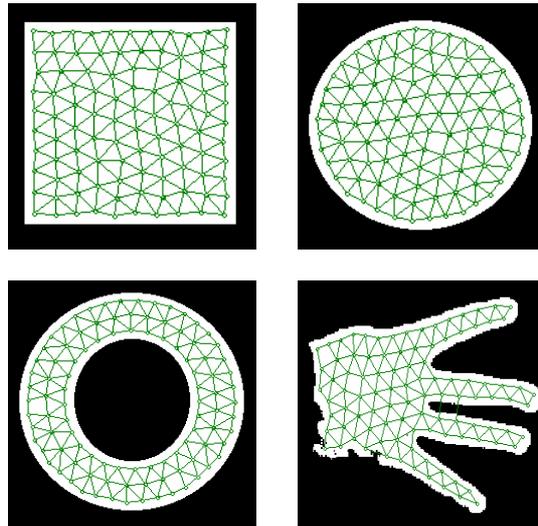


Figura 3.10. Grafos Preservadores de la Topología de diversos objetos bidimensionales.

Del mismo modo que el modelo es capaz de representar los diversos contornos de un objeto, es posible caracterizar las diversas partes de un objeto, o varios objetos presentes en la escena que posean los mismos valores para la propiedad visual T (Figura 3.11).

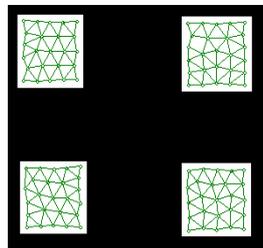


Figura 3.11. Representación de varios objetos.

Si bien el Grafo Preservador de la Topología GPT caracteriza el objeto, el empleo de éste para el reconocimiento o clasificación de los objetos viene limitado por restricciones temporales. Esto es debido a que la comparación entre grafos o cliques es un problema que se ha demostrado que es NP, pero se desconoce si es P o NP-completo (Fortin, 1996). Es por esto, que, a continuación se introducen diversos subgrafos del GPT que reducen el tiempo de comparación de los mismos al tener un menor número de aristas.

Asimismo, al igual que se ha realizado con los contornos, se presentan diversas características que pueden ser obtenidas del GPT del objeto. Hay bastante trabajo realizado en este sentido, que puede ser aprovechado, dado que la representación establecida por el grafo está muy relacionada con estructuras de geometría computacional.

SUBGRAFOS DEL GPT

Tal y como se ha expresado en el punto 2.3.1, el GPT establece una triangulación de Delaunay inducida de los vértices del grafo. En (Fortune, 1997) (O'Rourke y Toussaint, 1997) se establece la siguiente relación entre grafos:

$$NNG \subseteq MST \subseteq RNG \subseteq GG \subseteq DT$$

donde NNG es el grafo de vecinos más cercanos, MST es el árbol de expansión mínimo, RNG es el grafo de vecindad relativa, GG es el grafo Gabriel y DT es la triangulación de Delaunay (Figura 3.12).

GRAFO GABRIEL

El Grafo Gabriel GG es un subgrafo de la triangulación de Delaunay establecida por el Grafo Preservador de la Topología GPT que:

- posee su mismo conjunto de vértices A ,
- posee aquellas aristas $(p,q) \in C$ que cumplen que

$$B\left(\frac{p+q}{2}, \frac{\delta(p,q)}{2}\right) \cap \mathcal{C} = \emptyset \quad (3.7)$$

donde $B(x,r)$ es un círculo de centro x y radio r , y

$$\delta(p,q) = \sum_{i=1}^2 |w_{pi} - w_{qi}|$$

esto es, los puntos p y q definen el diámetro de un círculo que no es cruzado por ninguna otra arista.

En (Mou y Yeung, 1994) se presenta una red neuronal auto-organizativa, denominada Red Gabriel, que modifica durante el aprendizaje las conexiones entre las neuronas de modo que se obtenga el grafo Gabriel de las mismas. Sin embargo, los ejemplos presentados no son muy representativos y el coste computacional es elevado.

GRAFO DE VECINDAD RELATIVA

El grafo de vecindad relativa RNG es un subgrafo de la triangulación de Delaunay establecida por el Grafo Preservador de la Topología GPT que:

- posee su mismo conjunto de vértices A ,
- posee aquellas aristas $(p,q) \in \mathcal{C}$ que cumplen que $\Lambda_{p,q} \cap \mathcal{C} = \emptyset$, donde $\Lambda_{p,q} = B(p, \delta(p,q)) \cap B(q, \delta(p,q))$ es una luna o intersección entre dos círculos; esto es, no existe ningún punto Z que esté más cercano a p que q , y no existe ningún punto Z que esté más cercano a q que p .

ÁRBOL DE EXPANSIÓN MÍNIMO

El árbol de expansión mínimo MST es un subgrafo de la triangulación de Delaunay establecida por el Grafo Preservador de la Topología GPT que:

- posee su mismo conjunto de vértices A ,

- posee aquellas aristas $(p,q) \in \mathcal{C}$ que forman un árbol de longitud total mínima, de entre todos los posibles árboles que se pueden formar a partir del grafo GPT .

GRAFO DE VECINOS MÁS CERCANOS

El grafo de vecinos más cercanos NNG es un subgrafo de la triangulación de Delaunay establecida por el Grafo Preservador de la Topología GPT que:

- posee su mismo conjunto de vértices A ,
- posee aquellas aristas $(p,q) \in \mathcal{C}$ que cumplen que $B(p, \delta(p,q)) \cap \mathcal{C} = \emptyset$.

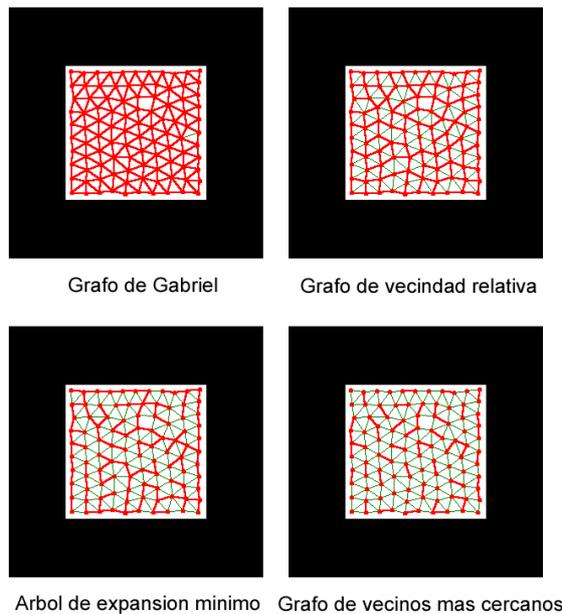


Figura 3.12. Subgrafos del GPT .

CARACTERIZACIÓN DE LA TOPOLOGÍA DEL GPT

Dentro de este apartado se incluyen todas aquellas medidas que pueden ser obtenidas de la propia topología del Grafo Preservador de la Topología. Esta información aporta datos sobre las

características morfológicas de los objetos. Entre ellas, se pueden obtener:

- Número de aristas: un mayor número de aristas sugiere una forma del objeto más compacta, ya que existe una mayor capacidad de conexión (Tabla 3.3).
- Histograma de vecindad: obtención del número de neuronas que poseen cada una de las vecindades. En la Figura 3.13 se observa que las figuras más compactas poseen unas vecindades más altas. Por otro lado, si el objeto posee esquinas (cuadrado) o zonas muy estrechas (los dedos de la mano) se generan neuronas con baja vecindad.
- Medidas estadísticas de las vecindades: datos como la vecindad media, la mediana, la desviación típica,... aporta información similar a la del histograma de vecindades (Tabla 3.4).
- Medidas estadísticas de las longitudes de las aristas: la presencia en un *GPT* de aristas de menor longitud sugiere la presencia de más aristas y, por tanto, una forma compacta (Tabla 3.3).
- Número y tipos de polígonos: aunque el *GPT* debería realizar una triangulación del espacio de entrada, al estar inducida por el objeto, provoca que algunas de las aristas de la triangulación de Delaunay desaparezcan y, por tanto, se presenten líneas, así como otros polígonos.

Objeto	Nº de aristas	Longitud Media	Mediana	Desviación típica
Cuadrado	258 / 258	17.67 / 17.63	17.46 / 17.49	1.88 / 1.91
Círculo	256 / 257	16.71 / 16.77	16.76 / 17.03	2.02 / 2.06
Anillo	226 / 227	17.17 / 17.17	17.00 / 17.12	2.06 / 2.09
Mano	228 / 228	19.07 / 19.06	18.87 / 18.97	2.25 / 2.26

Tabla 3.3. Información extraída del número y longitud de las aristas.

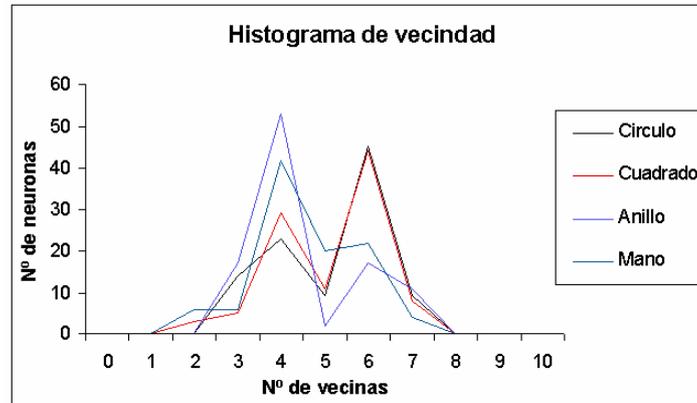


Figura 3.13. Histograma de vecindad.

Objeto	Nº de vecindades	Vecindad media	Mediana	Desviación típica
Cuadrado	516 / 516	5.16 / 5.16	6 / 6	1.20 / 1.22
Círculo	512 / 514	5.12 / 5.14	6 / 6	1.25 / 1.30
Anillo	454 / 452	4.54 / 4.52	4 / 4	1.27 / 1.26
Mano	456 / 456	4.56 / 4.56	4 / 4	1.18 / 1.17

Tabla 3.4. Información extraída de las vecindades.

MOMENTOS

Al igual que en el caso de los contornos, el cálculo de los momentos de una superficie es uno de los métodos más utilizados para su caracterización.

Dado un GPT se define su momento de orden $(p+q)$ como:

$$m_{pq} = \sum_{\forall (x,y) \in A} x^p y^q \quad (3.8)$$

para $p, q = 0, 1, 2, \dots$

Sus momentos centrales pueden ser expresados como:

$$\mu_{pq} = \sum_{\forall (x,y) \in A} (x - \bar{x})^p (y - \bar{y})^q \quad (3.9)$$

donde

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (3.10)$$

Finalmente, los momentos centrales normalizados se definen como:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (3.11)$$

donde

$$\gamma = \frac{p+q}{2} + 1 \quad (3.12)$$

para $p+q = 2, 3, \dots$

A partir de estos momentos, se pueden obtener un conjunto de siete momentos invariantes (Hu, 1962):

$$\phi_1 = \eta_{20} + \eta_{02} \quad (3.13)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (3.14)$$

$$\phi_3 = (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3.15)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (3.16)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (3.17)$$

$$\begin{aligned} \phi_6 = & (\eta_{20} - \eta_{02}) + [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (3.18)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (3.19)$$

Estos momentos obtenidos a partir de los *GPT* son válidos para la clasificación y reconocimiento de los objetos (Tabla 3.5).

	Cuadrado	Círculo	Anillo	Mano
η_{20}	23.18 / 23.44	20.71 / 20.89	38.13 / 37.95	45.72 / 45.45
η_{02}	21.70 / 21.32	18.35 / 18.64	36.24 / 36.33	23.92 / 23.98
η_{11}	-0.12 / 0.13	0.22 / -0.26	0.01 / -0.05	1.23 / 1.58
η_{30}	-0.32 / -0.45	-0.75 / -0.15	-0.81 / 3.60	21.21 / 18.62
η_{12}	-0.08 / -1.42	0.11 / 0.32	1.08 / 3.12	47.06 / 46.21
η_{21}	-0.80 / 0.30	0.32 / 0.66	-0.34 / -0.42	-15.29 / -15.41
η_{03}	0.94 / 1.02	1.15 / -0.30	-4.12 / -7.59	7.96 / 9.93
ϕ_1	44.88 / 44.76	39.07 / 39.53	74.37 / 74.29	69.63 / 69.43
ϕ_2	2.26 / 4.56	5.77 / 5.35	3.59 / 2.61	481.66 / 471.06
ϕ_3	2.42 / 25.86	4.64 / 3.44	32.32 / 246.48	27807 / 26042
ϕ_4	0.18 / 5.24	2.57 / 0.15	19.98 / 109.40	4714.47 / 4232.51
ϕ_5	-0.22 / 11.45	-4.00 / 0.04	337.05 / 2087.62	-31364862 / -28128226
ϕ_6	0.23 / 2.36	-4.98 / -0.28	-37.62 / -20.84	97994 / 87341
ϕ_7	0.09 / -39.30	5.15 / -0.13	358.62 / 1523.76	-32844276 / -25550254

Tabla 3.5. Momentos del *GPT*.

EXTRACCIÓN DEL CONTORNO DEL GPT

Dado un GPT de un objeto, se define su contorno como el conjunto de aristas que pertenecen únicamente a un polígono. Resultados de la aplicación de esta operación a varios objetos se presentan en la Figura 3.14.

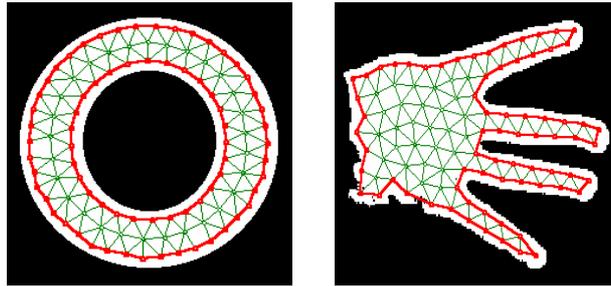


Figura 3.14. Contorno del GPT .

ESQUELETIZACIÓN (EJE MEDIO)

Tal y como se ha explicado anteriormente el GPT establece una esqueletización del objeto. Sin embargo, la representación del objeto mediante un esqueleto más ampliamente utilizada es la obtenida a partir de la transformada del eje medio, mucho más simple que el resultado obtenido en este trabajo. Este esqueleto viene definido por aquellos centros de círculos inscritos en el objeto que no están incluidos completamente en otro círculo. El resultado final presenta una figura estilizada del objeto.

Dado el gran número de trabajos que emplean la transformada del eje medio, se ha planteado la obtención del eje medio del objeto a partir de su GPT . Para ello, se calcula el diagrama de Voronoi de las aristas que forman el contorno del grafo (Figura 3.15). Mediante la unión de aquellos puntos que son frontera para tres o más regiones de Voronoi se obtiene el esqueleto del objeto (Figura 3.16).

De igual modo, se puede operar con el GPT_{∇} obtenido a partir del contorno del objeto.

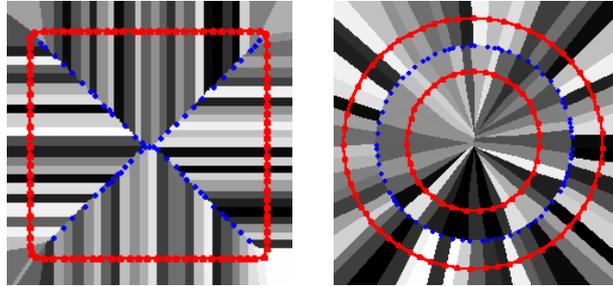
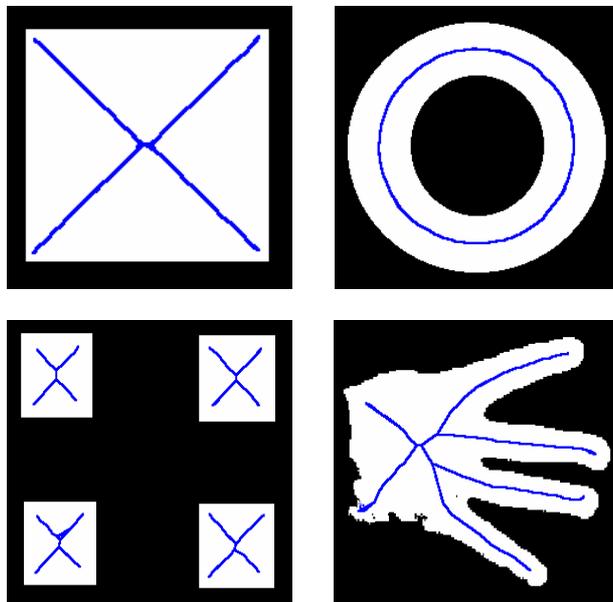
Figura 3.15. Diagrama de Voronoi del contorno del *GPT*.

Figura 3.16. Esqueletos.

3.3 APLICACIÓN A LA SÍNTESIS DE IMÁGENES

La capacidad de compresión de información de las redes neuronales auto-organizativas se ve claramente demostrada en este trabajo. La representación de un objeto bidimensional mediante un *GPT* extrae la información topológica más importante para realizar el análisis de la forma del objeto.

En este punto, se plantea el problema inverso: ¿cómo generar la forma del objeto a partir de su *GPT*? Es decir, cómo aplicar este trabajo a la síntesis de objetos.

Para ello, se procede siguiendo dos alternativas, dependiendo de la información disponible en el momento de realizar la síntesis:

- Si en el momento de la adaptación del *GPT* se guarda información asociada a cada neurona, relativa a la distancia r al punto del contorno del objeto más cercano a la misma; se procede a dibujar para cada una de las neuronas un círculo con radio r .
- Si únicamente se dispone del *GPT*, sin información adicional, se procede a dibujar círculos con centro en cada una de las neuronas y con un radio r que está en función del conjunto de aristas que salen de cada una de ellas:

$$r_i = \frac{\sum_{\forall j \in N_i} \|w_i - w_j\|^2}{\text{card}(N_i)} \quad (3.20)$$

Asimismo, si se guarda información de la intensidad o color del punto de la imagen que corresponde a cada una de las neuronas, se puede dotar a cada uno de los círculos de dicho tono, de modo que la síntesis posea la misma apariencia visual que el objeto real (Figura 3.17). Para mejorar la calidad de la síntesis, eliminando el efecto del dibujo de múltiples círculos, se realiza un filtrado (mediana, erosión / dilatación,...).

Por supuesto, existen técnicas más complejas de sintetización de objetos y de coloreado de los mismos, que podrían ser empleados; las cuales se han dejado para trabajos posteriores.

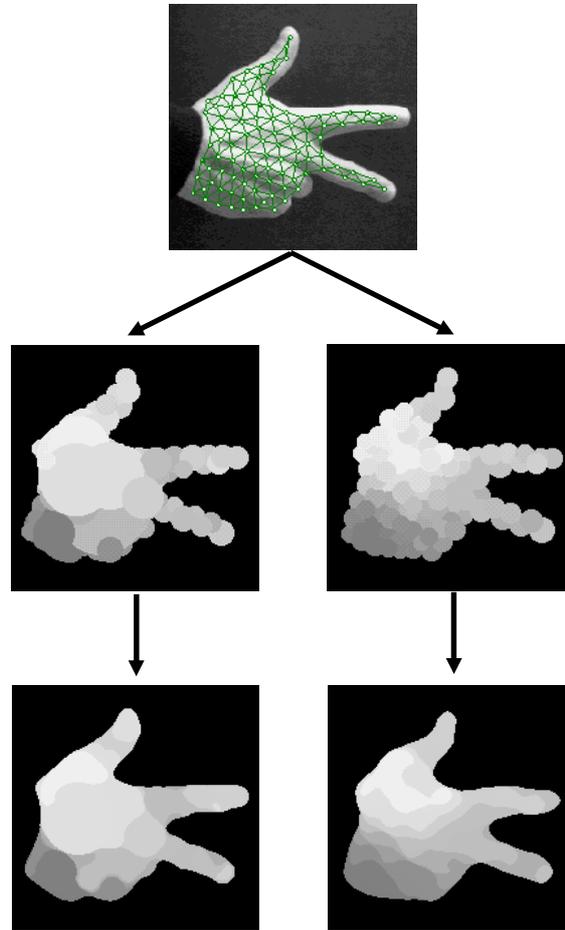


Figura 3.17. Síntesis a partir de un *GPT*.

3.4 EJEMPLO DE APLICACIÓN: CLASIFICACIÓN DE OBJETOS

En este punto, se aplica el modelo a la clasificación de objetos según su estructura externa. Para ello, se realiza la adaptación de una Growing Neural Gas al contorno de los objetos, obteniendo una caracterización de su forma.

Se ha empleado una base de datos de 1100 imágenes de contornos de peces utilizada en trabajos relacionados con la

búsqueda de similitud entre formas (Mokhtarian et al., 1996) (Veltkamp y Hagedoorn, 1999). Cada uno de los ficheros contiene un objeto representado por un número variable (entre 400 y 1600) de coordenadas x e y .

3.4.1 DESCRIPCIÓN GLOBAL DEL SISTEMA

El procedimiento a seguir es similar al presentado al inicio de este capítulo. En primer lugar, se debe extraer el campo de potencial normalizado ψ_{∇} , es decir, la probabilidad de cada uno de los puntos de pertenecer al contorno del objeto. Sin embargo, este paso no es necesario en esta aplicación, ya que la base de datos provee ya los contornos.

A continuación se procede a extraer el grafo preservador de la topología del contorno del objeto, mediante la adaptación del gas neuronal, en este caso una Growing Neural Gas.

Con esta estructura se puede actuar de dos modos: empleando el propio grafo preservador de la topología como entrada al clasificador, o extrayendo un vector de características a partir del mismo (Figura 3.18). En este apartado, se abordarán ambas posibilidades.

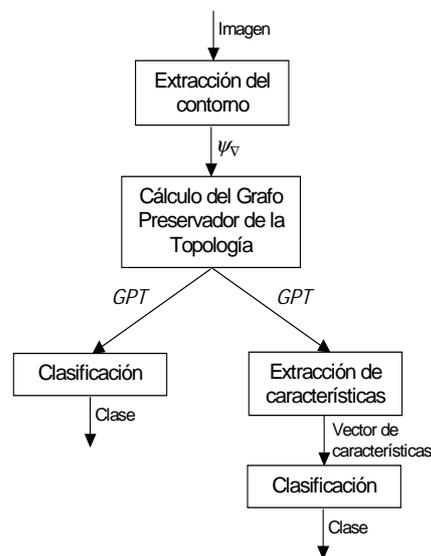


Figura 3.18. Sistema de clasificación de formas.

La clasificación se va a realizar, igualmente, con una Growing Neural Gas, pero en este caso con su empleo usual como clasificadora de patrones.

3.4.2 OBTENCIÓN DEL *GPT* DEL CONTORNO

A partir del campo de potencial normalizado se realiza el aprendizaje de la Growing Neural Gas, siguiendo el algoritmo del punto 2.2.2 con $\varepsilon_1 = 0.1$, $\varepsilon_2 = 0.01$, $\lambda = 10000$, $\alpha = 0$, $\beta = 0$. En este caso en particular, se educa la red con la condición de que finalice la adaptación cuando posea 100 neuronas.

Debido a la complejidad de algunos de los contornos se producen aristas indeseadas, que deben ser eliminadas para obtener una representación adecuada (Figura 3.19). Para ello, se localiza un ciclo que posea tantas aristas como neuronas posee la red. Se ha observado que esta operación es muy costosa en determinados casos, por lo que se está buscando algún tipo de solución, alguna de las cuales se presenta en el capítulo de “Líneas de continuación”.

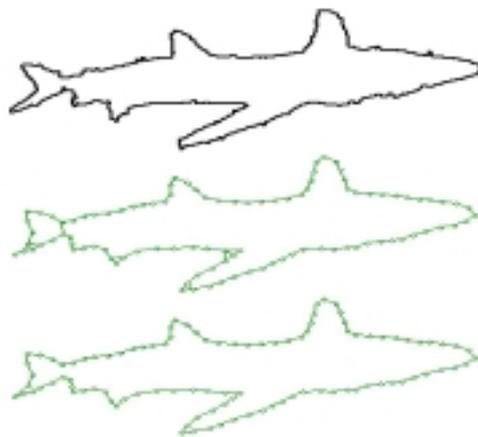


Figura 3.19. Obtención y simplificación del grafo del contorno.

3.4.3 CLASIFICACIÓN DE CONTORNOS A PARTIR DEL *GPT*

En este caso, la red neuronal clasificadora, una Growing Neural Gas, tiene como entrada un grafo preservador de la topología, que es, asimismo, otra Growing Neural Gas. Esta es una situación que no se suele dar en la literatura, en la que la entrada a una red neuronal sea otra red neuronal. Por ello, se debería continuar con esta línea de trabajo para extraer posibles conclusiones que llevaran al desarrollo de nuevos modelos.

El grafo obtenido en el paso previo es normalizado a traslación, orientación y escalado, del siguiente modo:

- se obtiene su centro de masas situándolo en el origen de coordenadas,
- se localiza la neurona más alejada al centroide, rotando todo el grafo para situarla en la horizontal,
- se inscribe el contorno en un círculo de radio estándar

Se realiza el aprendizaje de la red clasificadora, que posee 200 neuronas, empleando los mismos parámetros que para la extracción del contorno, con los 1100 *GPT* normalizados obtenidos de los respectivos contornos; de modo que, finalmente, cada una de las neuronas posee como vector de referencia el contorno determinado por un *GPT*.

Una vez completado el aprendizaje de la red, se realiza un ciclo de etiquetado, asociando cada uno de los patrones de entrada con la neurona cuyo contorno es más similar.

En la Figura 3.20 se muestra el resultado de la clasificación de diversos peces según este método: en la parte superior se muestra los *GPT* representantes de algunas neuronas, asociando a éste diversos peces para los que dicha neurona es la más cercana.

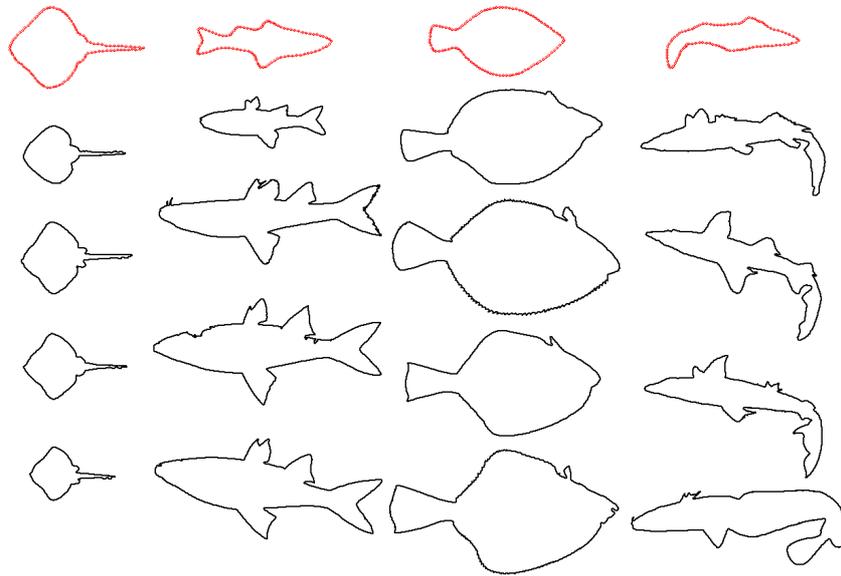


Figura 3.20. Clasificación de los objetos a partir del *GPT*.

3.4.4 CLASIFICACIÓN DE CONTORNOS A PARTIR DE CARACTERÍSTICAS EXTRAÍDAS DEL *GPT*

La segunda alternativa en el empleo del *GPT* como caracterizador del contorno de los objetos, es la extracción de un vector de características a partir del mismo, que reduzca la dimensionalidad de los patrones de entrada y mejore la caracterización. En este caso, se han probado dos de los métodos presentados previamente: los análisis de firmas normalizadas y de curvaturas. En ninguno de estos casos es necesario realizar un proceso de transformación del *GPT* original para normalizarlo.

El tamaño de estos vectores de características es igual al número de neuronas del *GPT*. En ambos casos, se observa que tras realizar el aprendizaje de la red clasificadora, se obtienen como vectores de referencia, funciones suavizadas de las firmas y curvaturas (Figuras 3.21 y 3.22). En el primer caso no es un problema, pero sí que se han encontrado dificultades en el caso del análisis de curvaturas ya que las funciones son mucho más

fluctuantes entre valores positivos y negativos. Esto produce que el suavizado lleve la mayor parte de los valores de las curvaturas a cero. Sin embargo, los valores significativos (picos o valles) de la función permiten realizar una correcta clasificación de los peces.

3.5 CONCLUSIONES

En este capítulo se ha demostrado la capacidad de representación de objetos bidimensionales, mediante un gas neuronal. Estableciendo una función de transformación adecuada, el modelo es capaz de adaptarse tanto a las características internas como externas del objeto, obteniendo lo que se ha denominado Grafo Preservador de la Topología.

El modelo, por su propio método de adaptación, es capaz de dividirse de modo que pueda extraer los múltiples contornos que puede tener un objeto, así como caracterizar las diferentes partes en las que pueda estar fragmentado.

Con el objetivo de facilitar y acelerar tareas de más alto nivel, como clasificaciones y reconocimientos, se extraen características de este grafo, como momentos, esqueletos, firmas,...

No sólo se obtienen resultados prometedores en el análisis de objetos, sino que su aplicación a la síntesis de objetos ha sido también satisfactoria. La triangulación de Delaunay establecida por el modelo, permite una compresión de la información de la morfología, a partir de la cual se consigue sintetizar correctamente el objeto.

Para finalizar, se ha mostrado la capacidad de representación de la forma de los objetos que posee el modelo. Para ello, se ha extraído el contorno de los objetos mediante un *GPT*, empleándolo para la clasificación de los mismos. Se han presentado dos posibles alternativas: la primera de ellas novedosa, en cuanto a que el vector de características y el clasificador son redes neuronales auto-organizativas del mismo tipo; la segunda de ellas, usual, en la que se extraen características del contorno, construyendo un vector que sirve de entrada a la red clasificadora.

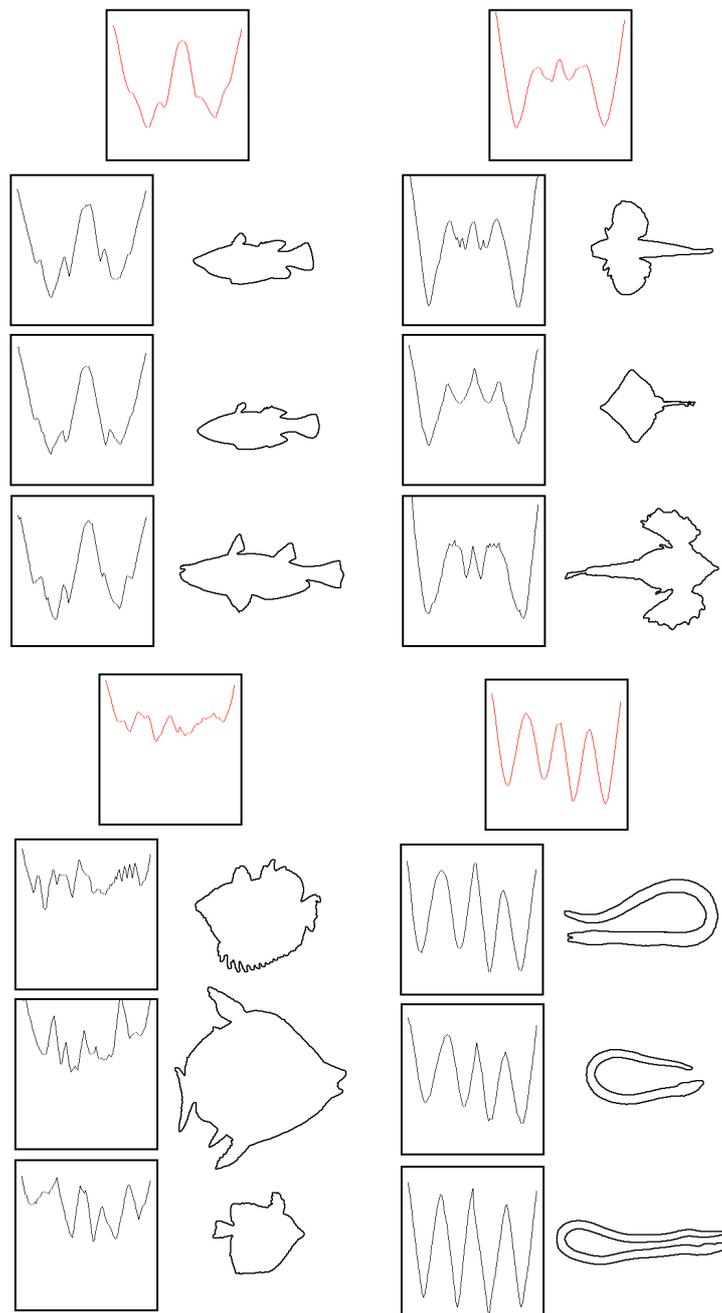


Figura 3.21. Clasificación de los objetos según su firma.

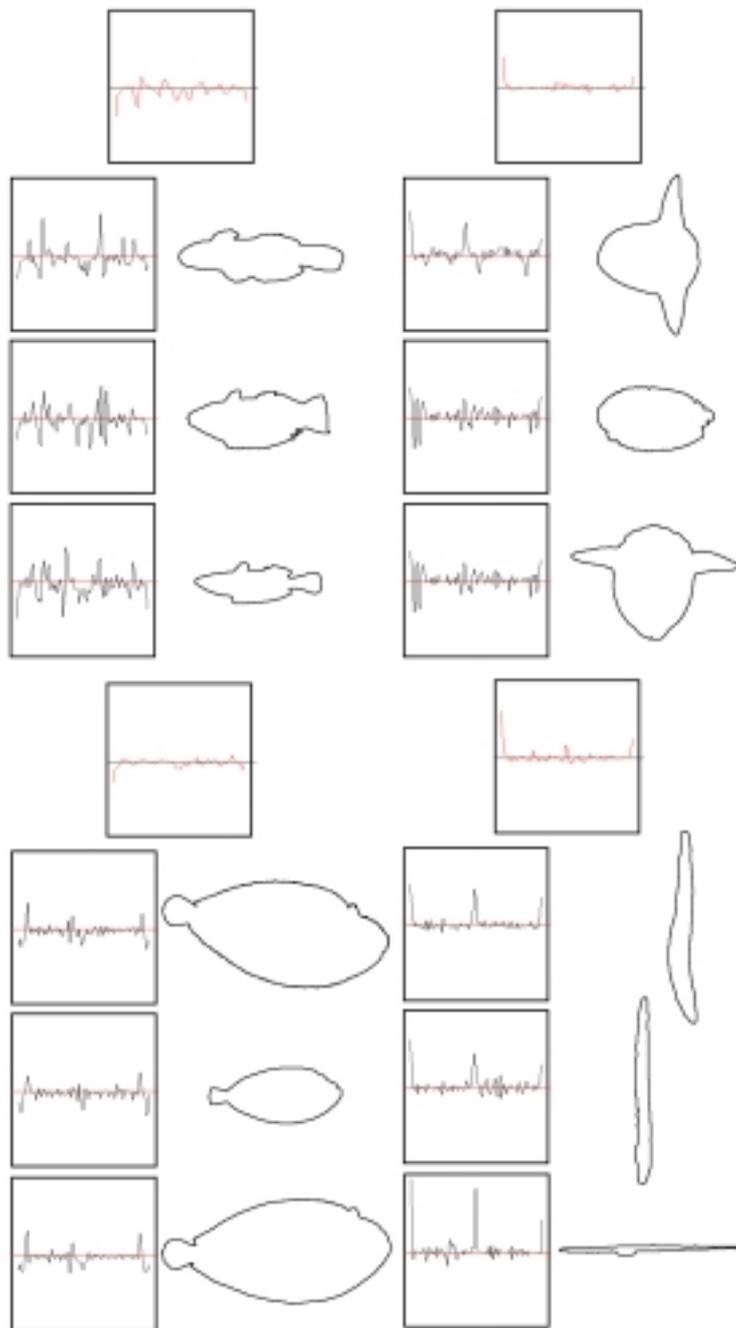


Figura 3.22. Clasificación de los objetos según su curvatura.

