

UNIVERSIDAD DE CANTABRIA

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS
DE CAMINOS, CANALES Y PUERTOS

TESIS DOCTORAL

Desarrollo de una metodología para el estudio de la
morfología de playas basado en mapas auto-organizativos de
imágenes digitales.

Presentada por: **OMAR QUETZALCÓATL GUTIÉRREZ GUTIÉRREZ**

Dirigida por: **ERNESTO MAURICIO GONZÁLEZ RODRÍGUEZ**
RAÚL MEDINA SANTAMARÍA

Santander, julio 2011

Aplicación de Mapas Auto-Organizativos para identificación y clasificación de imágenes

3.1. Introducción

Los mapas auto-organizativos (SOMs), desarrollados por *Kohonen* (2000), forman parte del grupo de redes neuronales competitivas no supervisadas y están específicamente indicadas para trabajar con espacios vectoriales de altas dimensiones. Las redes neuronales competitivas tienen la habilidad de detectar automáticamente grupos homogéneos a partir de un conjunto de datos disponible, mientras que al ser no supervisadas implica que no se presupone un patrón de salida y se permite que la red evolucione libremente. Los SOMs, como los algoritmos de tipo cluster, tienen la habilidad de automáticamente detectar grupos, que cumplan con la característica de que las diferencias internas de los miembros de cada grupo son minimizadas, mientras que las diferencias entre grupos se maximizan. Difieren de los métodos tradicionales de *clusters* en dos características (*Hewitson y Crane, 2002*):

- Durante la formación de los grupos, no solo el grupo *ganador* es modificado sino que los grupos vecinos también.
- El resultado de los SOMs se puede representar en un mapa bidimensional fácil de comprender

Una de las principales fortalezas de los SOMs es la capacidad de encontrar patrones complejos de bajas dimensiones ocultos en el conjunto de variables de entrada que suelen ser de una alta dimensionalidad por medio de un proceso de aprendizaje no supervisado. Los SOMs están compuestas de un mapa bi-dimensional de nodos en los que los vectores multidimensionales se proyectan de forma no lineal (*Somervuo y Kohonen, 1999*).

La proyección se realiza por un proceso comparativo, en busca de la representación más fiel de aquellas dimensiones de las variables de entrada entre las que la variación es más importante (*Cavazos et al., 2002*). A cada nodo del mapa le corresponde un vector de pesos o prototipo con el que cada set de datos es comparado por medio de una métrica (usualmente la distancia euclidiana) para encontrar el nodo ganador (BMU en inglés) en el mapa. El proceso comparativo es recursivo y a cada iteración la BMU y sus vecinos más cercanos son modificados, mientras que los más alejados permanecen inhibidos. Este proceso continua hasta que el sistema converge.

Los SOMs comparados con otras técnicas de clasificación *i.e.* *k*-medias, resulta costosa computacionalmente debido a las restricciones topológicas impuestas, es decir, mientras en otras técnicas los centroides son libres para moverse en la malla, los SOMs restringen su movimiento. Sin embargo, tienen algunas ventajas, ya que permiten estudiar las transiciones o posiciones relativas entre los distintos prototipos. Descripciones más detalladas de la técnica se pueden encontrar en *Gutiérrez et al. (2004)*, *Hewitson y Crane (2002)* y *Molinier et al. (2005)*. Las ventajas de los SOMs han sido aprovechadas en áreas como la meteorología (*Hewitson y Crane, 2002; Cavazos, 2000*) y oceanografía (*Schizas et al., 1994; Hsieh y Tang, 1998; Camus, 2009*).

3.2. Estado del arte

Los SOMs han despertado gran interés en los últimos años y se han aplicado con éxito en diversos campos *e.g.*, biología (*Palakal et al.*, 1995), medicina (*Joutsiniemi et al.*, 1995), química (*Chen y Gasteiger*, 1997) y desde luego la geofísica (*e.g. Schizas et al.*, 1994; *Hsieh y Tang*, 1998; *Hewitson y Crane*, 2002), incluso han sido utilizados como mecanismos de clasificación *Malmgren y Winter* (1999), *Cavazos* (2000) y *Ambroise et al.* (2000).

En la clasificación de imágenes destaca **PicSOM** (*Laaksonen et al.*, 2000) por su robustez. **PicSOM** se encuentra disponible a través de la red en el sitio de internet de la Helsinki University of Technology (<http://www.cis.hut.fi/picsom>). Este sistema ha sido utilizado en variadas aplicaciones prácticas, en la visión digital (*Laaksonen et al.*, 2000; *Brandt et al.*, 2002; *Koskela et al.*, 2001), en la identificación de estructuras hechas por el hombre en imágenes de satélite (*Molinier et al.*, 2005), identificación de objetos en imágenes (*Laaksonen et al.*, 2005a,b), entre otros. Otra aplicación interesante de clasificación de imágenes digitales por medio de SOMs es **aiSOMPic** (*Chesnut*, 2004), disponible en la red (<http://www.brains-n-brawn.com/default.aspx?vDir=aisompic>), donde incluso una vez entrenado el mapa se puede proponer un prototipo sencillo de ciertas características de color y el mapa selecciona las imágenes que cumplen con las mismas.

3.3. Planteamiento teórico

3.3.1. Descripción de los SOMs

Los SOMs están compuestos por (ver figura 3.1)

1. Una base de datos de entrada n -dimensional \vec{x}_n , donde n es el número de datos.

2. Un conjunto de grupos o unidades representados por las neuronas de salida $y_{i,j}$, localizados en una malla regular 2D o mapa, para propósitos visuales, de dimensiones (I, J) .
3. Cada neurona $y_{i,j}$ tiene asociado un vector de pesos $\beta_{i,j} = (\beta_{1,1}, \dots, \beta_{I,J})$ (el valor original de los pesos se asigna de forma aleatoria) asociado a la posición central de la neurona en el espacio n -dimensional, y
4. una métrica, usualmente la distancia euclidiana, por medio de la cual los vectores prototipos $\beta_{i,j}$ son modificados según los datos \vec{x}_n .

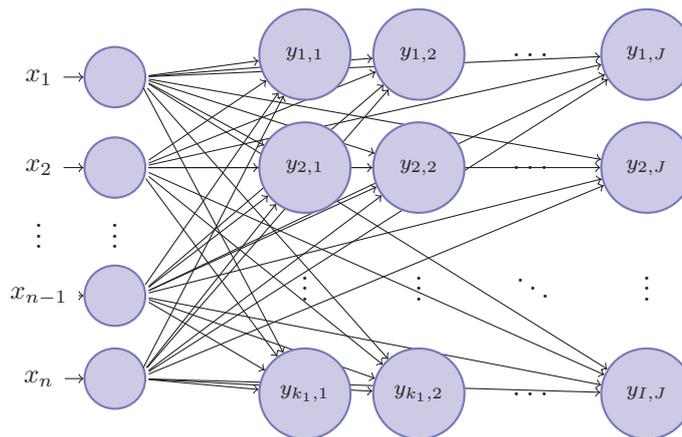


Figura 3.1. Esquema típico de un Mapa Auto-Organizativo.

Los elementos de los SOMs pueden ser comparados con aquellos necesarios para realizar una clasificación óptima de una base de datos cualquiera por una persona. Por ejemplo, en el caso de clasificar una serie de documentos, estos son la base de datos (\vec{x}_n), un archivador hace las funciones de los SOMs, los separadores de este equivalen a las unidades ($y_{i,j}$) del mapa, la métrica al criterio de clasificación de los documentos, digamos para este caso *palabras clave*, y los prototipos que definen las características de cada separador (ver figura 3.2). El proceso clasificar los documentos se traduce en el entrenamiento del mapa que se explica a continuación.

3.3 Planteamiento teórico

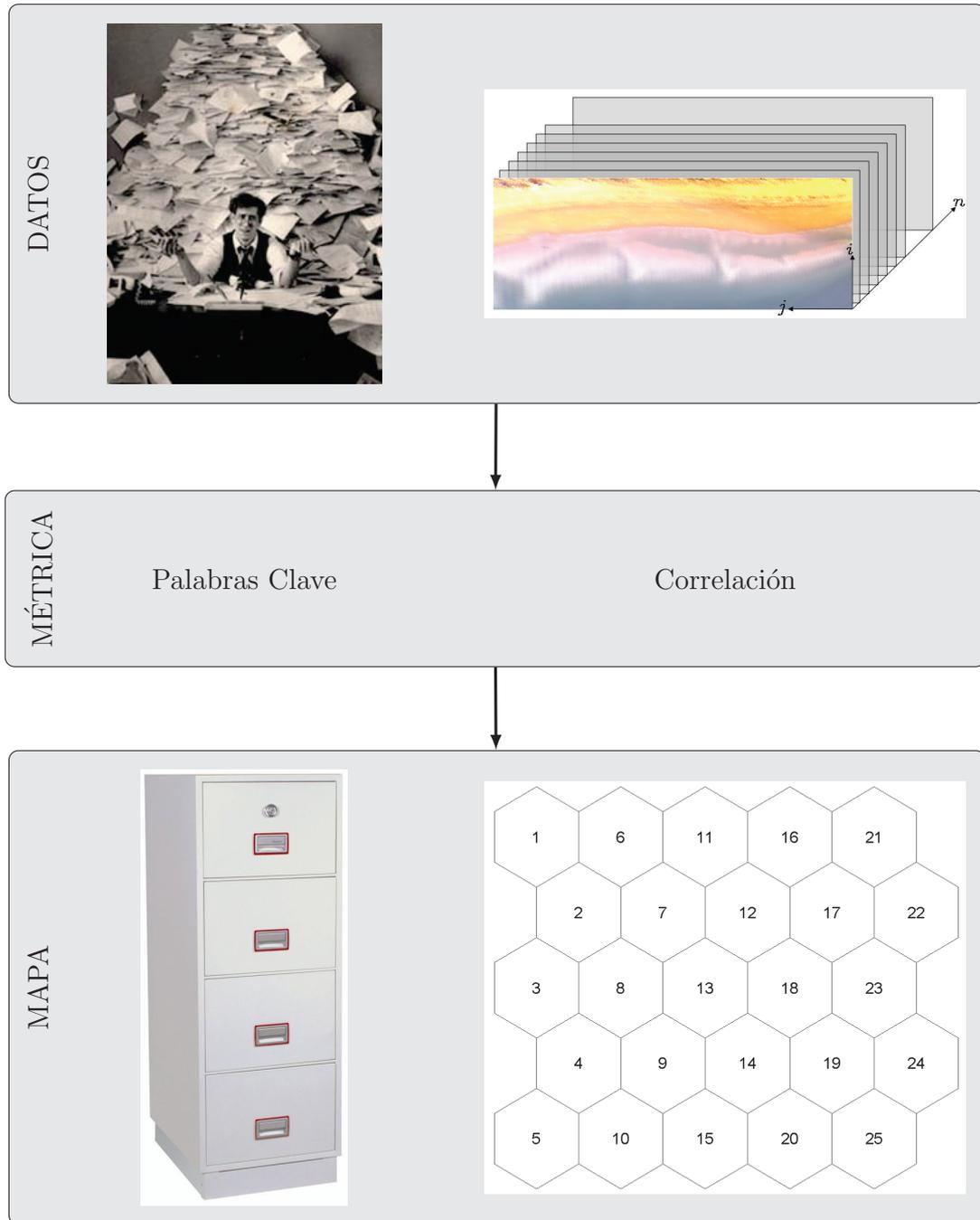


Figura 3.2. Comparación entre el procedimiento de clasificación de documentos en un archivero y el entrenamiento de una colección de IMDP en un SOM.

3.3.2. Entrenamiento de un mapa

Para conseguir que el mapa represente las variables del espacio n -dimensional se requiere de un entrenamiento. El proceso de entrenamiento consiste en la adaptación iterativa de los valores de los vectores prototipos, de forma que cada prototipo final sea representativo de un grupo de datos. Este procedimiento es similar al realizado por otras técnicas como k -medias (*Gutiérrez et al.*, 2004), sin embargo, lo que hace distinta a los SOMs de otras técnicas de clasificación es que incluye un mecanismo auto-organizativo de vecindad de tal forma que los grupos vecinos en el mapa tienen características similares en el espacio n -dimensional o real.

La implementación del algoritmo se realiza en ciclos de forma iterativa. En cada ciclo, para cada uno de los vectores de datos x_i , $i = 1, \dots, n$ se determina el prototipo más cercano $y_{w(i)}$, como aquel que minimiza la distancia al vector de datos:

$$\|x_i - y_{w(i)}\| = \min_k \|x_i - y_k\|, k = 1, \dots, m. \quad (3.1)$$

Una vez determinadas las distancias en cada ciclo, se recalculan los prototipos basándose en los valores de los centroides de la unidad correspondiente y vecinas:

$$y_j = \frac{\sum_{i=1}^a h(\|r_j - r_{w(i)}\|)x_i}{\sum_{i=1}^a h(\|r_j - r_{w(i)}\|)}, j = 1, \dots, m, \quad (3.2)$$

donde la función $h(r)$, es un núcleo de vecindad que mide las distancias de los grupos en el mapa $2D$ y determina la tasa de cambio de un prototipo con base en los grupos vecinos, usualmente es una función gaussiana,

$$h(r) = \exp\left(\frac{-r}{s(t)}\right), \quad (3.3)$$

donde $s(t)$ es el radio de vecindad que decrece en el tiempo, suavizando las restricciones topológicas. Una descripción más detallada de distintas implementaciones se puede encontrar en *Kaski y Oja* (1999).

3.4. Aplicación en la clasificación de imágenes digitales

3.4.1. Introducción

La diferencia fundamental en el uso de SOMs en la clasificación de imágenes, comparado con variables geofísicas consiste en que, las últimas tienen un sentido físico, en función de lo cual se realiza la clasificación. Las imágenes digitales, por el contrario, son un conjunto de valores de intensidad de píxel que no tienen una relación física con el contenido de la imagen *e.g.* un píxel con intensidad igual a 1, blanco, puede estar asociado a cualquier objeto dentro de la imagen, sin tener que guardar ninguna relación con otro píxel cuya intensidad sea idéntica. Al ojo humano es muy simple identificar patrones y similitudes entre las imágenes, para los SOMs cada imagen no es más que un conjunto de píxeles y valores que no necesariamente están relacionados con el contenido. Aun con estas limitaciones, los SOMs han sido aplicadas con éxito en diferentes problemas prácticos de visión digital (*Laaksonen et al.*, 2000; *Brandt et al.*, 2002; *Koskela et al.*, 2001), de identificación de estructuras construidas por el hombre en imágenes de satélite (*Molinier et al.*, 2005) y objetos en imágenes digitales (*Laaksonen et al.*, 2005a,b).

Por lo tanto el mayor esfuerzo realizado en la adaptación de los SOMs para clasificar imágenes digitales se ha enfocado en el aspecto de determinar, a partir de las imágenes, los parámetros más adecuados para la clasificación. Otro de los aspectos a considerar en el funcionamiento de los SOMs es la métrica a utilizar. Los SOMs como se explicó previamente utilizan la distancia euclidiana como métrica, sin embargo *Di Gesù y Starovoitov* (1999) discutieron las limitaciones de usar esta distancia como otras medidas de mayor orden de complejidad en la comparación de imágenes digitales. Los autores propusieron diferentes métricas más apropiadas con ordenes de complejidad mayores, sin embargo no determinaron una métrica óptima.

A continuación se describen los parámetros que han sido utilizados por otros autores para clasificar imágenes digitales, así como de las métricas propuestas para cuantificar la similitud entre imágenes.

3.4.2. Características de los SOMs en la clasificación de imágenes

3.4.2.1. Parámetros a clasificar

El sistema **PicSOM** tiene implementados SOMs estructurados en forma árbol y usa como métrica la distancia euclidiana. La metodología empleada por **PicSOM** divide cada imagen en 5 zonas, y calcula las siguientes variables para cada zona:

- **Color promedio:** Obtenido al calcular el promedio de los valores de rojo (**R**), verde (**G**) y azul (**B**) en cada una de las 5 zonas de la imagen. El resultado es un mapa a entrenar con un vector de datos x_n^r con $r = 3 \times 5 = 15$ dimensiones que describe el color de la imagen y proporciona información de la distribución del color en la imagen.
- **Momentos espectrales del color:** Considerados por primera vez por *Stricker y Orengo* (1995). Los momentos de color de la imagen se calculaban tratando cada canal de color, en cada una de las zonas, como una distribución de probabilidad independiente. Los primeros tres momentos (media, varianza y el sesgo) se calculaban para cada canal de color, resultando un vector de características con $r = 3 \times 3 \times 5 = 45$ dimensiones.
- **Textura:** La textura, definida para un cluster de 9 píxeles es la probabilidad de que la intensidad de alguno de los píxeles que rodean el cluster sea mayor que la del píxel central. Como medida de intensidad se consideraban los valores Y de la representación **YIQ** de la imagen (Los valores Y se pueden obtener al resolver el

sistema lineal definido en el apartado 4.1 utilizando los valores de colores **RGB** de la imagen). El resultado es un mapa a entrenar con un vector de $r = 8 \times 5 = 40$ dimensiones.

- **Forma del histograma:** Esta característica esta basada en el histograma en las 8 direcciones principales de los ejes de la imagen. Con un histograma por cada una de las secciones se tiene un vector de $r = 8 \times 5 = 40$ dimensiones. Esta propiedad describe la distribución de colores en cada una de las 8 principales direcciones de la imagen (columna inicial, columna final, fila inicial, fila final, fila central horizontal, fila central vertical y dos diagonales principales) y con esto representa el contenido de la imagen por medio de un estadístico de bajo nivel.
- **Forma de la Transformada de Fourier (FFT):** Esta característica consiste en determinar la transformada de Fourier de la imagen binaria (en este caso, la imagen no se divide en zonas). El tamaño de las imágenes era normalizado a 512×512 píxeles previo al cálculo de la FFT. El resultado es filtrado con un pasa-bajos y diezmada en un factor de 32, de tal forma que el resultado es un vector de $r = 128$ dimensiones (*Brandt et al., 2002*).

Estas variables, determinadas a partir de las imágenes digitales, no tienen un sentido físico, sin embargo nos proporcionan información de la forma en que los colores se distribuyen en las misma. A diferencia de **PicSOM**, en **aiSOMPic** en esta aplicación solamente se usaron características basadas en el color y la textura. En el caso del color realizaron pruebas con diferentes esquemas de representación de los colores de la imagen (**RGB**, **HSL**, **YUV**, y **CIE-L*a*b***), encontrando los mejores resultados con el formato HSL. En el caso de la textura, ésta se determinó con un algoritmo simple que calcula la diferencia de intensidad entre un píxel y sus vecinos. Si la diferencia es grande, la textura es alta y la imagen es ruidosa. Las mediciones se realizaron de dos maneras: la primera en forma de un histograma de 16-bins que genera un vector de datos 16-dimensional y la segunda, separando la imagen en 9 secciones, por lo que, el vector de datos es 9-dimensional. Esta aplicación es capaz de, conocida una imagen,

buscar imágenes semejantes en una colección. El programa tiene la suficiente robustez como para encontrar la imagen adecuada aún cuando la imagen de entrada es un simple boceto.

3.4.2.2. Métricas

Tanto **PicSOM** como **aiSOMPic** utilizan como métrica la distancia euclidiana. Sin embargo, existen otras métricas más adecuadas para la comparación de imágenes. *Di Gesù y Starovoitov* (1999), propusieron 4 métricas de alto grado de complejidad: la distancia Hausdorff, distancia local promedio, distancia global y distancia simétrica y realizaron una comparación de los resultados con los valores obtenidos usando métricas de bajo orden de complejidad como la correlación y la distancia euclidiana. Desafortunadamente, ya que las comparaciones realizadas por estos autores son cualitativas, no determinan cual es la métrica más apropiada en la comparación, sin embargo sugieren la distancia local promedio como la más adecuada.

A continuación se describe brevemente las características de estas métricas. Las formulaciones consideran dos imágenes digitales con las mismas características $A = \{A_{i,j}\}$ y $B = \{B_{i,j}\}$ como conjuntos de píxeles definidos por las coordenadas espaciales (i, j) y un conjunto de niveles de intensidad $\{0, 1, \dots, G\}$, de tal forma que un píxel cualquiera es $A_{i,j} = (i, j, a_{i,j})$ donde $a_{i,j}$ es la intensidad de la imagen en las coordenadas (i, j) .

1. *Distancia de Hausdorff (Hausdorff-based distance)*

Esta medida fue especialmente diseñada para calcular la máxima distancia entre píxeles de las imágenes A y B

$$HG(\mathbf{A}, \mathbf{B}) = \max_{i,j} \{d_{city}(A_{ij}, \mathbf{B}), d_{city}(B_{ij}, \mathbf{A})\}, \quad (3.4)$$

donde d_{city} es la función distancia entre los píxeles de A_{ij} a la imagen B y desde

B_{ij} a A

$$d_{city}(A_{i,j}, \mathbf{B}) = \min_{B_{lm} \in \mathbf{B}} \{A_{ij}, B_{lm}\}, \quad (3.5)$$

$$d_{city}(B_{i,j}, \mathbf{A}) = \min_{A_{lm} \in \mathbf{A}} \{B_{ij}, A_{lm}\}. \quad (3.6)$$

La función distancia utilizada es,

$$d_{city}(A_{ij}, B_{lm}) = \frac{|i-l| + |j-m|}{N} + \frac{|a_{ij} - b_{lm}|}{G}, \quad (3.7)$$

El orden de complejidad de este calculo es $O(N^4)$.

2. Distancia Local (Local distance based function)

Esta distancia es la raíz cuadrada del promedio de las distancias locales medias,

$$AD(\mathbf{A}, \mathbf{B}) = \sqrt{\frac{1}{2(N-2W)} \sum_{i,j=W}^{N-W} \sqrt{d_{city}^2(A_{ij}, \mathbf{B}_{W_{ij}}) + d_{city}^2(B_{ij}, \mathbf{A}_{W_{ij}})}, \quad (3.8)$$

donde W indica una sub-imagen limitada por un cuadrado digital de tamaño $2W+1$ centrado en la posición (i, j) , mientras que d_{city} esta definida en 3.7. Usando $(N-2W)^2$ píxeles centrales la complejidad del calculo es $O((N-2W)^2(2W)^2)$.

3. Distancia Global (Global feature based distance)

Esta distancia combina características espaciales y de intensidad de los píxeles,

$$GD(\mathbf{A}, \mathbf{B}) = \frac{2}{N^2(N^2-1)} \sum_{A_{ij} \in \mathbf{A}} \sum_{B_{lm} \in \mathbf{B}} \delta(A_{ij}, B_{lm}), \quad (3.9)$$

$$\delta(A_{ij}, B_{lm}) = \alpha d_{city}((i, j), (l, m)) + \beta \mathbf{GR}(a_{ij}, b_{lm}), \quad (3.10)$$

$$\mathbf{GR}(a_{ij}, b_{lm}) = \frac{|a_{ij} - b_{lm}|}{\max(a_{ij}, b_{lm})}, \quad (3.11)$$

$$(3.12)$$

donde $\alpha = \beta = 0.5$ es el peso de los componentes de δ . El orden de esta medida

es $O(N^4)$. Sin embargo, GD puede ser aproximada como el valor medio de la misma función calculada para sub-imágenes $A_{st} \subset A$ y $B_{st} \subset B$ con dimensiones $n \times n$,

$$GD(\mathbf{A}, \mathbf{B}) = \frac{1}{n^2} \sum_{s=0}^{\frac{N}{n}} \sum_{t=0}^{\frac{N}{n}} GD(\mathbf{A}_{st}, \mathbf{B}_{st}), \quad (3.13)$$

los índices (s, t) definen la posición del píxel superior-izquierda $(i, j) = (s \times n, t \times n)$ en las imágenes A y B , de tal forma que la complejidad del calculo se reduce a $O(n^4)$ para cada $GD(A_{st}, B_{st})$ de n^2 pares y en total de $O(N^2 n^2)$.

4. Distancia Simétrica (Symmetry based distance)

Esta medida compara las características simétricas de las imágenes con respecto a varios ejes. Esta medida se basa en la *Discrete Symmetry Transform*, T , (*Di Gesù y Valenti*, 1996). Esta transformada proporciona un conjunto de momentos axiales K calculados en un disco digital C_R de radio R centrado en el píxel (i, j) de la imagen,

$$T_k^A(i, j) = \sum_{(i+r, j+s) \in C_R} \left| r \sin\left(\frac{k\pi}{N}\right) - s \cos\left(\frac{k\pi}{N}\right) \right| a_{i+r, j+s}, \quad (3.14)$$

con $k = 0, 1, \dots, k, \dots, K - 1$, donde K depende on del circulo C_R . La imagen A esta representada por un conjunto de K características,

$$\mathbf{T}(\mathbf{A}) = (T_0^A, T_1^A, \dots, T_k^A, \dots, T_{K-1}^A).$$

así la distancia simétrica SD es calculada en función de $T(A)$ y $T(B)$,

$$SD(\mathbf{T}(\mathbf{A}), \mathbf{T}(\mathbf{B})) = \frac{1}{N^2 K} \sum_{i, j=0}^N \sum_{k=0}^K GR(T_k^A(i, j), T_k^B(i, j)), \quad (3.15)$$

donde GR es calculado como 3.11. La complejidad es $O(N^2 KC)$ donde C es el numero de puntos en el disco C_R .

5. Mediciones de bajo orden de complejidad

Las medidas consideradas son la *Distancia Euclidiana (ED)* y la *Correlación cruzada (CO)*, ambas de complejidad $O(N^2)$.

$$ED(\mathbf{A}, \mathbf{B}) = \frac{1}{NG} \sqrt{\sum_{i,j=0}^N (a_{ij} - b_{ij})^2}, \quad (3.16)$$

$$CO(\mathbf{A}, \mathbf{B}) = 1 - \frac{\sum_{i,j=0}^N a_{ij} b_{ij}}{\sqrt{(\sum_{i,j=0}^N a_{ij}^2) (\sum_{i,j=0}^N b_{ij}^2)}}. \quad (3.17)$$

3.5. Implementación de los SOMs

La Helsinki University of Technology en 2002 desarrolló una serie de herramientas para calcular SOMs en Matlab y las agrupó en un *toolbox*. En esta *toolbox* se agrupa el conjunto de funciones necesarias para aplicar la técnica. El *toolbox* incluye las funciones necesarias para preparar los datos, entrenar mapas y visualizar resultados. En el apéndice B se describen las principales características del funcionamiento del *toolbox*. A continuación, para ilustrar el funcionamiento de los SOMs, se presenta una serie de ejemplos en los que se realizan la clasificación de bases de datos reales y ficticias.

3.5.1. Ejemplos

- Ejemplo 1.

La base de datos *Iris.dat* es una base de datos conocida y utilizada como ejemplo por *Vesanto et al.* (2000), para mostrar las características de los SOMs. Este conjunto de datos consiste de mediciones de longitud (L) y ancho (W) de las hojas del cáliz (*Sepal*) y de los pétalos (*Petal*) de 150 flores el género *Iris*: 50 de *Iris setosa* (se), 50 de *Iris versicolor* (ve) y 50 de *Iris virginica* (vi), por tanto en \vec{x}_n , $n = 150$ y cada vector tiene dimensiones $r = 4$, dadas por las variables L ,

W , *Sepal* y *Petal*.

La base de datos se lee y normaliza. El entrenamiento se realiza en un solo paso y con inicio aleatorio en un mapa de 16×4 unidades. Los resultados del entrenamiento se muestran a continuación en la figura 3.3. El panel a) muestra la matriz de distancia entre las unidades (*U-matrix*), en esta se observa que las cuatro primeras filas forman un *cluster* muy homogéneo. Las etiquetas asociadas (panel f) indican que el cluster corresponde a la subespecie Setosa. La parte inferior de la *U-matrix* incluye las otras dos subespecies; no se observa una separación entre ambos, sin embargo, en las etiquetas se nota que ocupan un lugar fijo en el cluster.

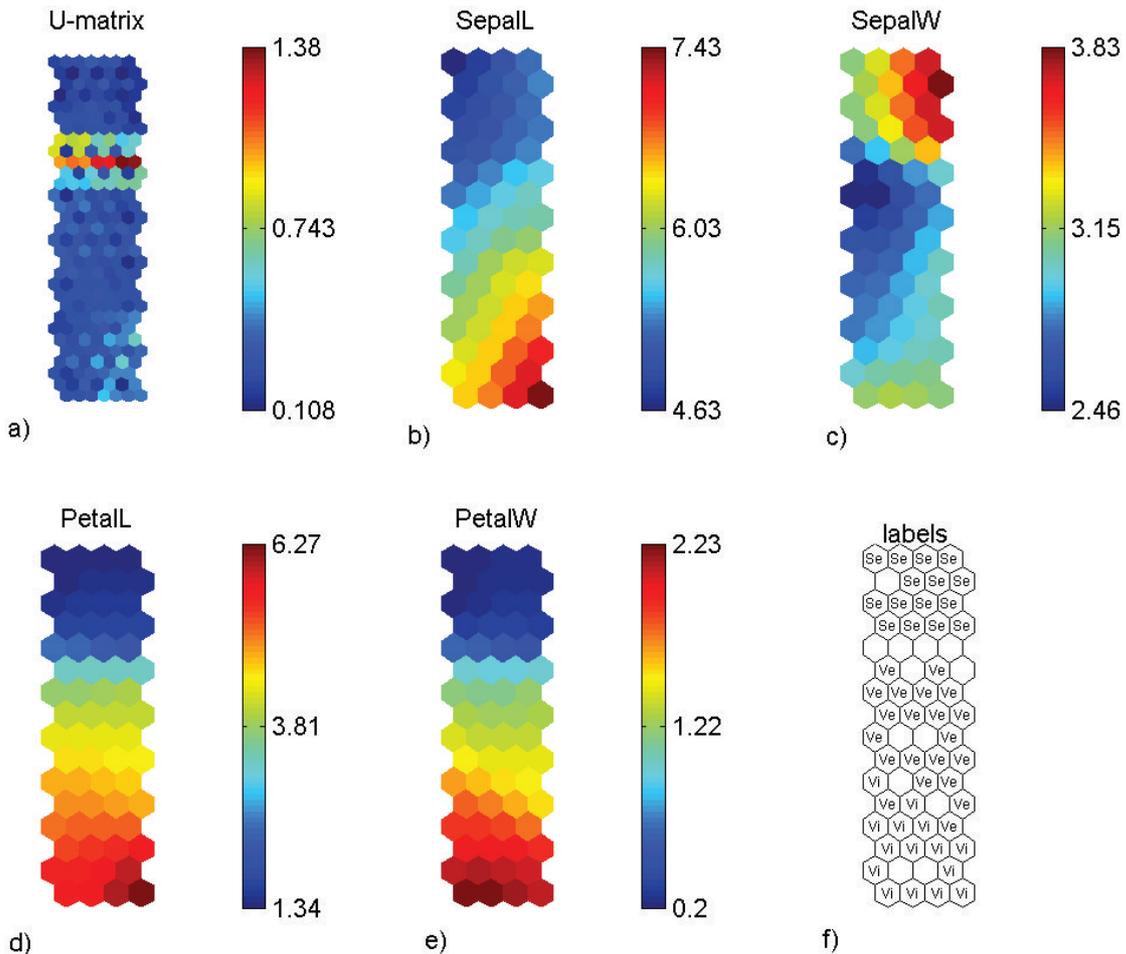


Figura 3.3. Visualización de los SOMs de *Iris* data. a) Mapa de distancias entre unidades; b), c), d) y e) muestran los mapas de distancia entre las variables longitud L y ancho W del cáliz *Sepal* y pétalo *Petal*. El gráfico f) muestra un mapa con las etiquetas de cada unidad.

3.5 Implementación de los SOMs

En los paneles de las variables (b, c, d y e) se observan las relaciones entre las variables de los prototipos. La ventaja de este tipo de presentación de resultados, implica poder crear relaciones entre variables, por ejemplo, se puede observar que la subespecie Setosa, ubicado en la parte superior del mapa de etiquetas (panel f), tiene pétalos pequeños (panel d y e) y hojas del cáliz cortas (panel b) y anchas (panel c). El panel a) no muestra separación entre las subespecies virginica y versicolor, sin embargo, el mapa de etiquetas (panel f) muestra que ocupan una posición distinta dentro del mapa: *I. virginica* en la parte inferior y *I. versicolor* en la parte central. Así mismo, en los paneles d) y f) se puede observar que la longitud y el ancho del pétalo en cada especie es ligeramente distinto.

Estos planos son muy convenientes cuando existen muchas variables a comparar. Cuando son, solamente unas pocas las variables que interesan los diagramas de la dispersión son mas eficientes, puesto que muestran como se agrupan las variables en el espacio de los SOMs. La figura 3.4 muestra el espacio SOM, y la posición que ocupa cada subespecie. La figura 3.5 muestra para cada unidad un histograma con el valor de las cuatro variables correspondientes. El color en la figura indica la subespecie a la que pertenecen.

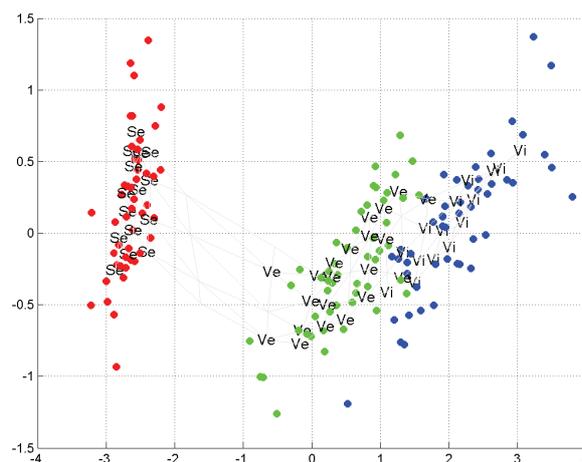


Figura 3.4. Proyección del espacio de los SOMs. Las tres subespecies se muestran de distinto color. La malla se ha proyectado en el mismo subespacio.

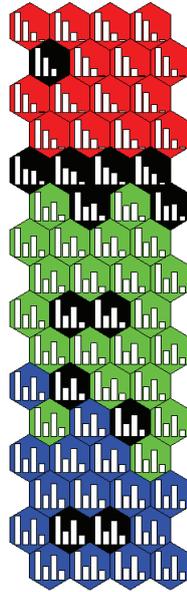


Figura 3.5. El mapa entrenado, donde los colores indican la localización de las subespecies, mientras que los histogramas los valores en cada unidad.

■ Ejemplo 2.

Para realizar este ejercicio se generó una base de datos con $n = 1000$ mediciones de tres variables: Rojo, Verde y Azul ($r = 3$). El valor de cada variable se eligió de forma aleatoria de entre los valores de intensidad 0 y 255, de tal forma que la combinación de estas tres variables, permiten la existencia en el espacio de la muestra de 8 colores distintos, rojo (R), verde (G), azul (B), blanco (W), amarillo (Y), negro (K), azul celeste (C) y rosa (M). A cada uno de las combinaciones en la muestra se le asignó una etiqueta correspondiente a cada color.

Esta base de datos se entrenó en un solo paso y con inicio aleatorio en un mapa de 20×20 . La figura 3.6 muestra los resultados obtenidos. El panel a) que muestra la distancia entre las unidades, indica que existen ocho grupos bien diferenciados de resultados. Los paneles b), c) y d) indican la distribución en el mapa de valores para las variables Rojo, Verde y Azul respectivamente. De estas gráficas se observa que la técnica fue capaz de encontrar y separar los 8 colores que fueron incluidos en la base de datos. La combinación de los valores en los paneles b), c) y d) forman el color correspondiente a esa posición en el mapa. En la figura 3.7 se

muestran el mapa con las etiquetas que identifican a cada grupo con una posición. Por ejemplo, en la zona central de los paneles b), c) y d) se encuentran los valores iguales a cero, lo que corresponde al color negro, indicado en el mapa de etiquetas con la **K**. Lo mismo ocurre en la zona central superior donde los paneles b), c) y d) tienen valores iguales a 255, que representan el color blanco (indicado con **W**).

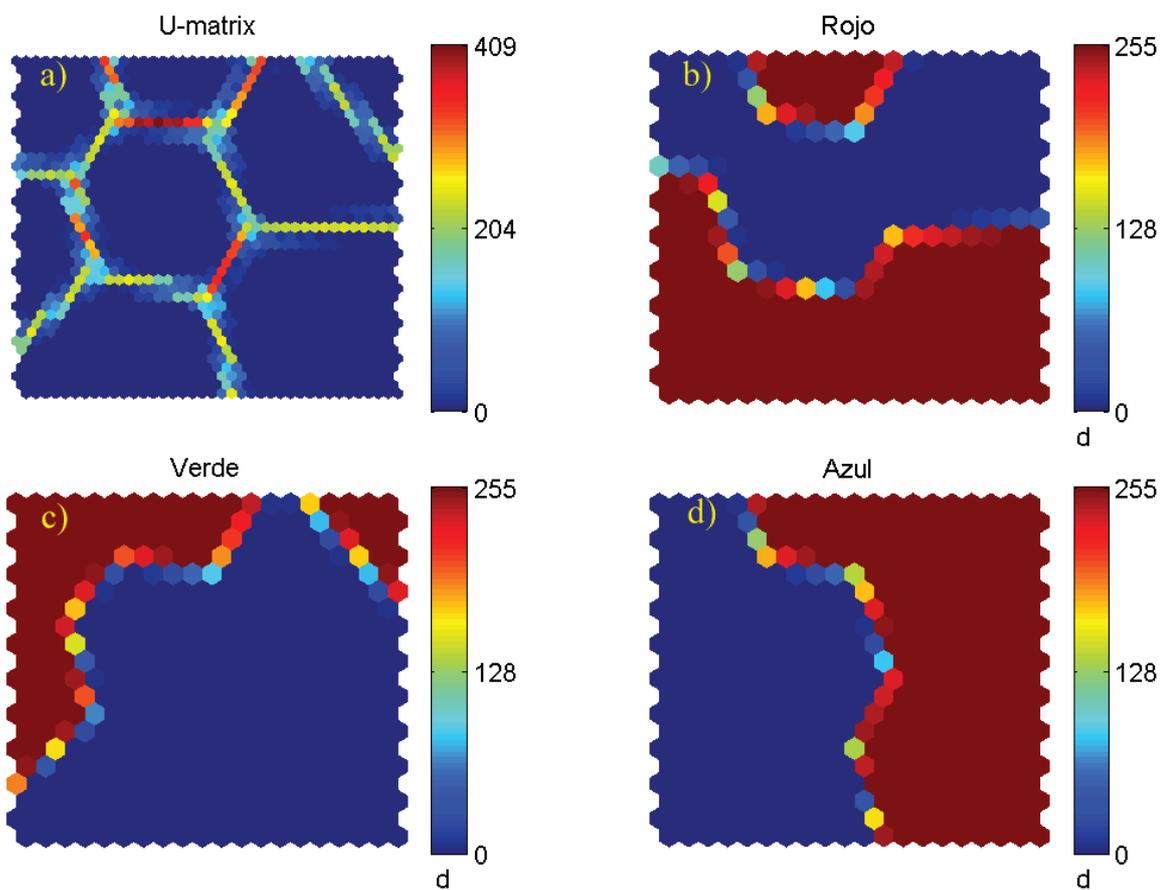


Figura 3.6. Visualización del mapa entrenado del ejercicio 2. El panel a) muestra las distancias entre las unidades. Los paneles b), c) y d) muestran las relaciones entre las variables y su disposición en el mapa.

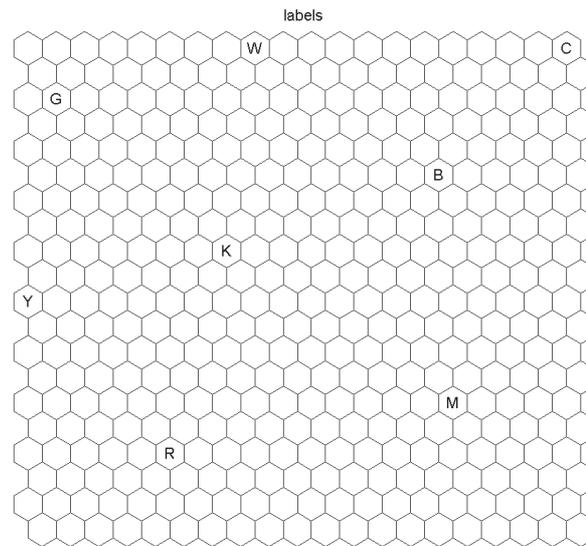


Figura 3.7. En el panel se muestran las etiquetas que identifican a cada grupo: *G* verde, *W* blanco, *C* azul celeste, *B* azul, *K* negro, *Y* amarillo, *R* rojo y *M* rosa .

■ Ejemplo 3.

En el ejercicio 2, se presentó una base de datos con 8 clases distintas que la técnica fue capaz de encontrar y clasificar adecuadamente. Sin embargo, en una muestra real de datos, no se conoce cuantas clases existen o cual es su distribución, por esta razón se generó una nueva base de datos, con la misma estructura, y valores aleatorios entre 0 y 255, de tal forma que en una muestra existen todos los colores del espectro visible. En este caso la base de datos incluye 5000 muestreos. El entrenamiento de esta base de datos nuevamente se realizó en un solo paso y con un inicio aleatorio.

La figura 3.8 muestra al igual que en el caso anterior las distancias (panel a) y la distribución (paneles b, c y d) entre unidades. En el panel a) aún cuando se observa la formación de grupos, la distancia entre ellos no es lo suficiente para dejarlos claramente en evidencia. Sin embargo, en los paneles b), c) y d), se muestra el gradiente de intensidad de cada color entre las unidades del mapa. Es posible ubicar los tonos más oscuros en la esquina superior izquierda del mapa, y los más claros en la esquina inferior derecha. Los tonos rojizos en la esquina inferior izquierda, los azules al centro en la parte superior y los verdes también

al centro pero en la parte inferior del mapa.

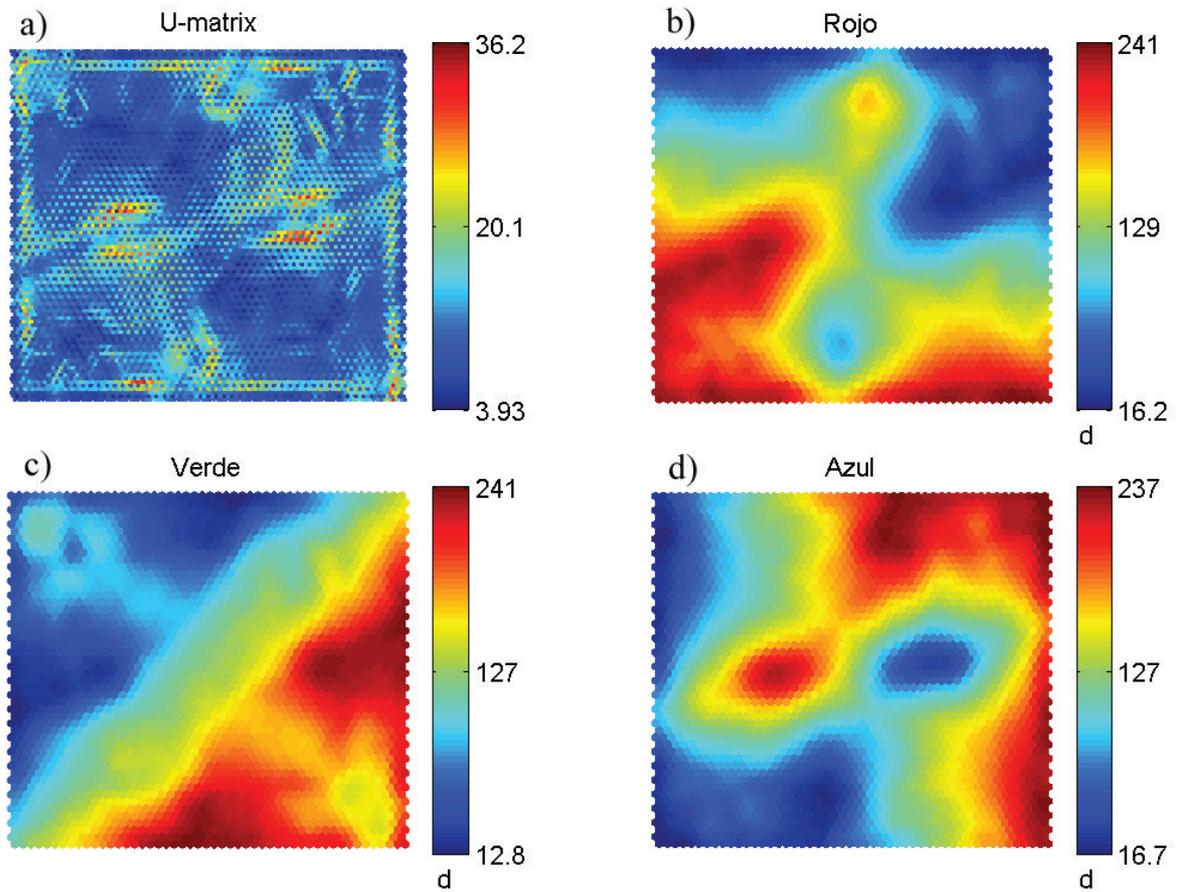


Figura 3.8. Visualización del mapa del ejercicio 3. El panel a) muestra las distancias entre las unidades. Los paneles b), c) y d) muestran las relaciones entre las variables y su disposición en el mapa.

3.6. Conclusiones

En este capítulo se profundizó en los fundamentos teóricos de los SOMs y en su viabilidad para ser usados como técnica de clasificación de IMDP. Se encontró que los SOMs son redes neuronales competitivas no supervisadas y son capaces de trabajar con espacios vectoriales de altas dimensiones. Tienen la habilidad de detectar, a partir de los datos, grupos con características homogéneas, de tal forma que las diferencias internas son mínimas, mientras que las diferencias entre los grupos se maximizan.

La aplicación de los SOMs, a diferencia de otras técnicas de clasificación, es costosa computacionalmente por las restricciones topológicas. Sin embargo, presenta ciertas ventajas como por ejemplo, que los resultados se presentan en un mapa bidimensional en el que se proyectan los vectores multidimensionales. Esta propiedad permite estudiar transiciones o posiciones relativas entre prototipos.

Los SOMs están compuestos por una base de datos de entrada multidimensional, un mapa formado por neuronas o unidades que agrupan conjuntos de vectores organizados en una malla regular de bidimensional; cada neurona esta asociada a un vector de pesos que representa las características del conjunto agrupado, y una métrica por medio de la cual los vectores prototipos son modificados según los datos de entrada. El entrenamiento del mapa se realiza de forma iterativa, asignando cada vector de datos a la unidad o neurona que mejor la representa por medio del vector prototipo. A cada paso del entrenamiento los vectores prototipos son recalculados en función de los vectores recién agrupados.

Los SOMs se han aplicado con éxito en diversos campos (*e.g.* Palakal *et al.*, 1995; Joutsiniemi *et al.*, 1995; Chen y Gasteiger, 1997), incluida la geofísica (*e.g.* Schizas *et al.*, 1994; Hsieh y Tang, 1998; Hewitson y Crane, 2002; Camus, 2009), incluso han sido utilizados como mecanismos de clasificación Malmgren y Winter (1999), Cavazos (2000) y Ambroise *et al.* (2000). En lo que respecta a su aplicación a imágenes digitales destaca **PicSOM** (Laaksonen *et al.*, 2000), que ha sido aplicado en problemas de visión digital, identificación de estructuras en imágenes de satélite, identificación de objetos en imágenes digitales, etc. La diferencia fundamental en el uso de SOMs en la clasificación de imágenes, comparado con variables geofísicas consiste en que, las últimas tienen un sentido físico, en función de lo cual se realiza la clasificación. Las imágenes digitales, por el contrario, son un conjunto de valores de intensidad de píxel que no tienen una relación física con el contenido de la imagen. Por lo tanto el mayor esfuerzo de adaptación de los SOMs para clasificar imágenes digitales se ha enfocado en el aspecto de determinar los parámetros más adecuados a calcular a partir de las imágenes, así como la métrica a utilizar.

3.6 Conclusiones

Finalmente en este capítulo se realizaron ejemplos sencillos de clasificación de bases de datos simples para comprender el funcionamiento de la técnica.

