



Instituto Politécnico Nacional
“La Técnica al Servicio de la Patria”

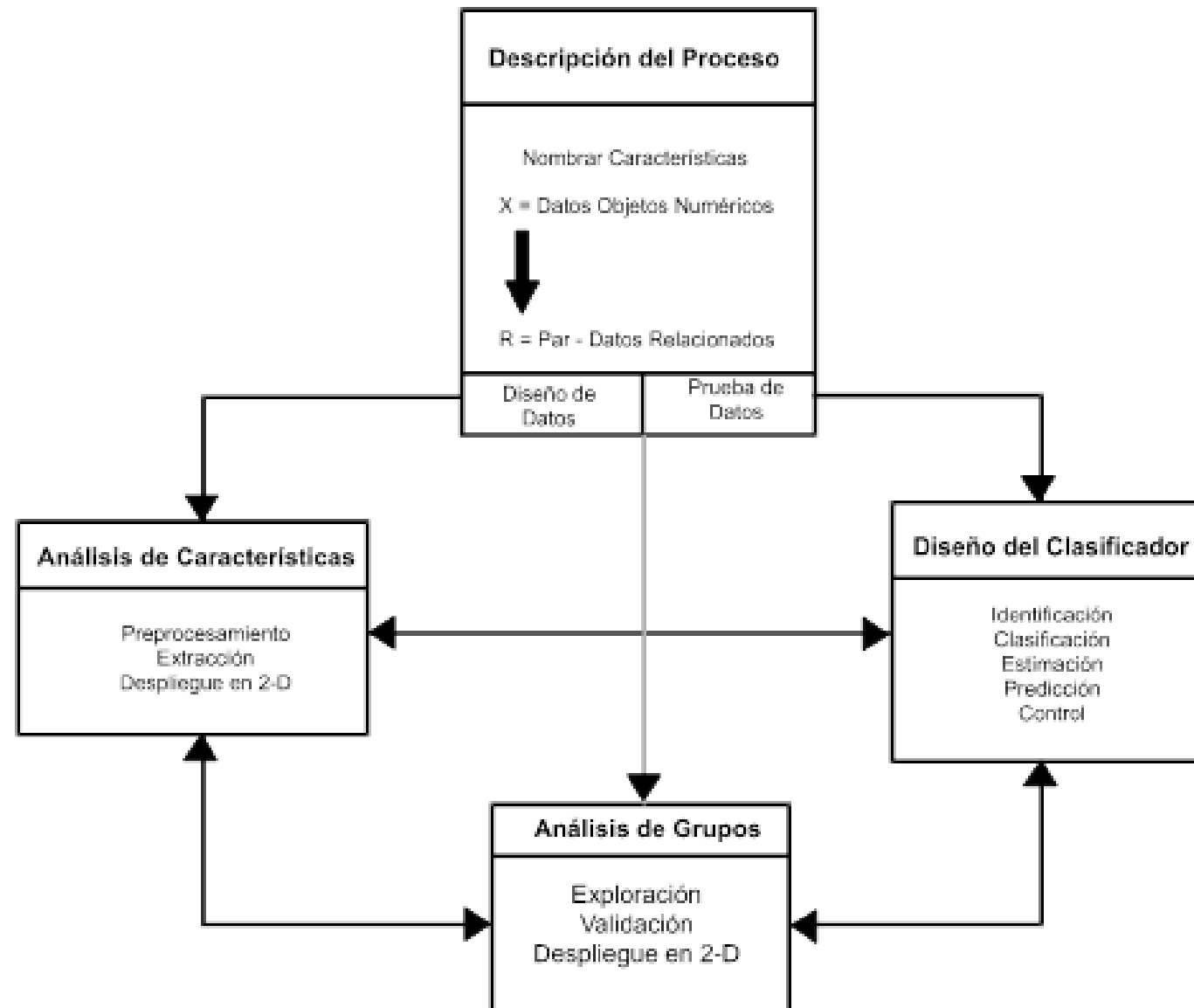


Microtecnología y
Sistemas Embebidos

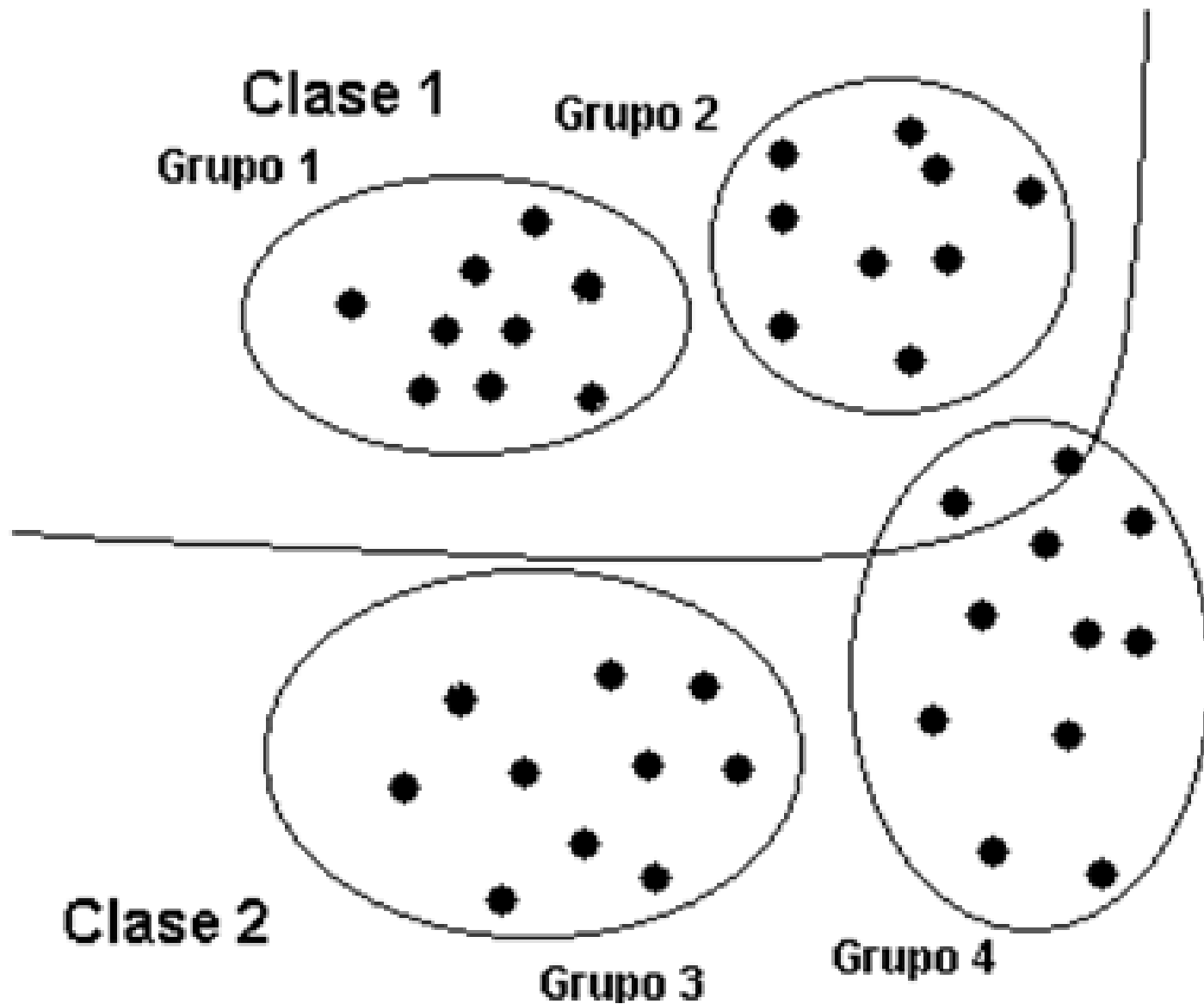
II MINI WORKSHOP ON HIGH PERFORMANCE COMPUTING IN
SCIENCE AND ENGINEERING

CLUSTERING DIFUSO

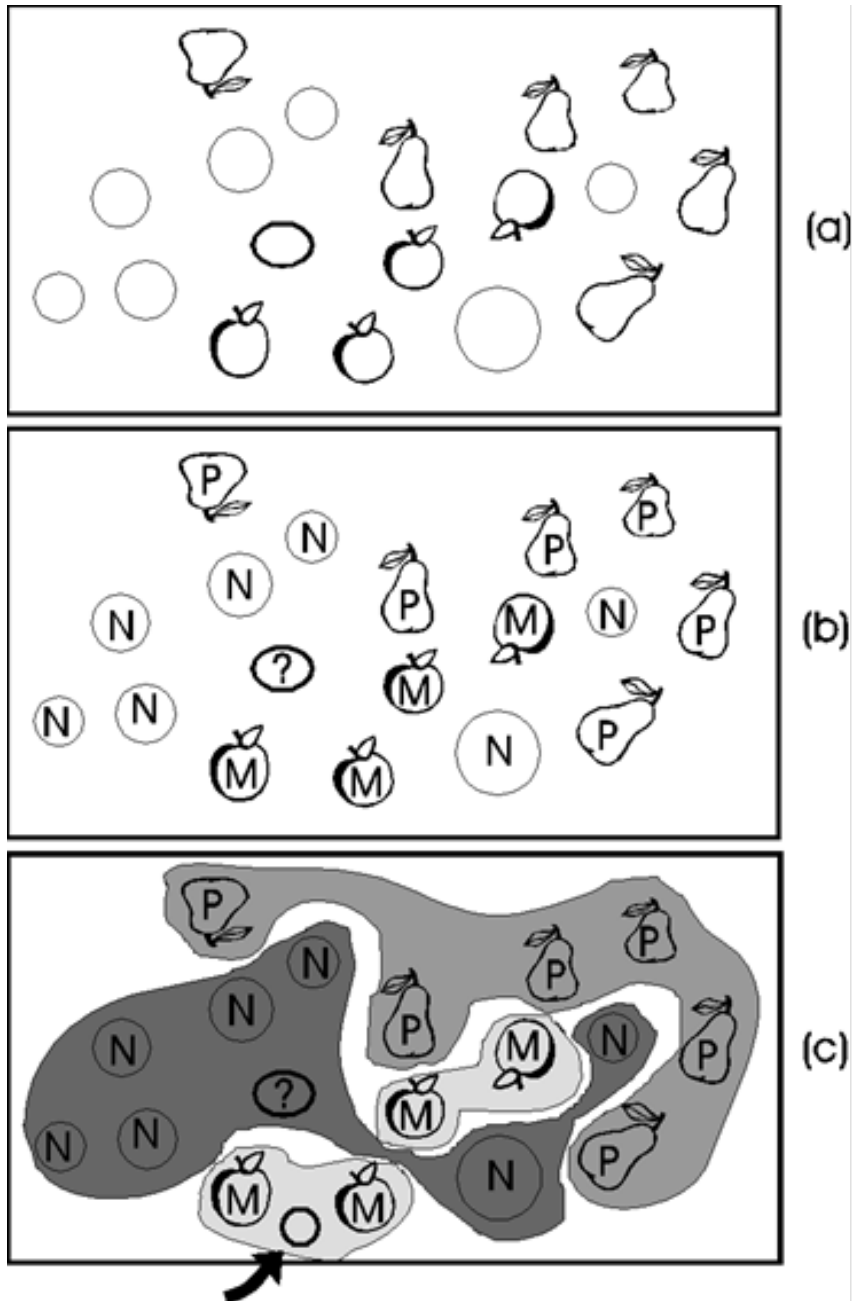
Instituto Politécnico Nacional
Centro de Investigación en Computación
Dr. Herón Molina Lozano



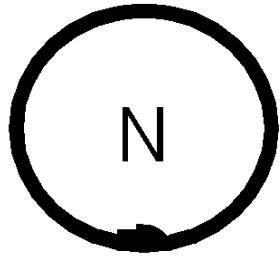
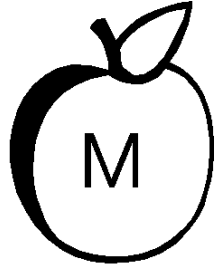
ELEMENTOS DE UN SISTEMA DE RECONOCIMIENTO DE PATRONES NUMÉRICO TÍPICO



AGRUPAMIENTO Y CLASIFICACIÓN



Descripción del
proceso:
a) descripción de
los objetos;
b) análisis
numérico;
c) análisis de
características.



P → 2
D → 1
T → 0

1

2



1

1.4

0

Selección

Selección de las características de textura, para separar y clasificar a las naranjas las manzanas y las peras.

Técnicas de agrupamiento

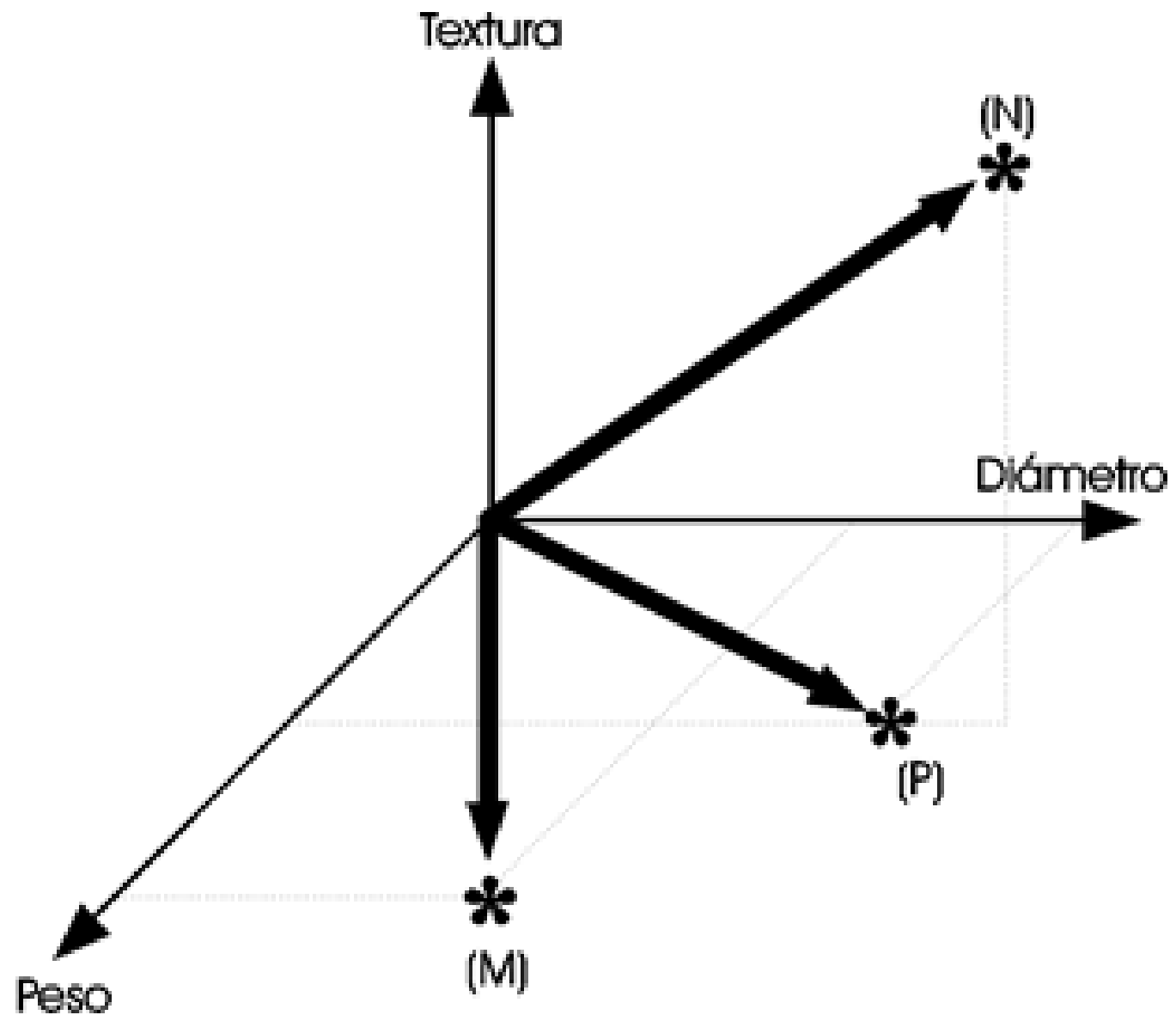
- Incompleta o heurística.
- Determinística.
- Certeras con traslape.
- Probabilísticas.
- Posibilísticas.
- Jerárquica de grupos.
- Por funciones objetivo.
- Estimación de agrupamientos.

Extracción de características

Básicamente cualquier función que realice la siguiente transformación

$$R^p \rightarrow R^q \text{ donde } p \geq q,$$

se le llamará extractor de características. Es decir, se tiene una imagen Y de X a través de f_E , $Y = f_E [X]$. La idea básica es realizar una compresión de la información que nos sea útil y eliminar la información que sea redundante.

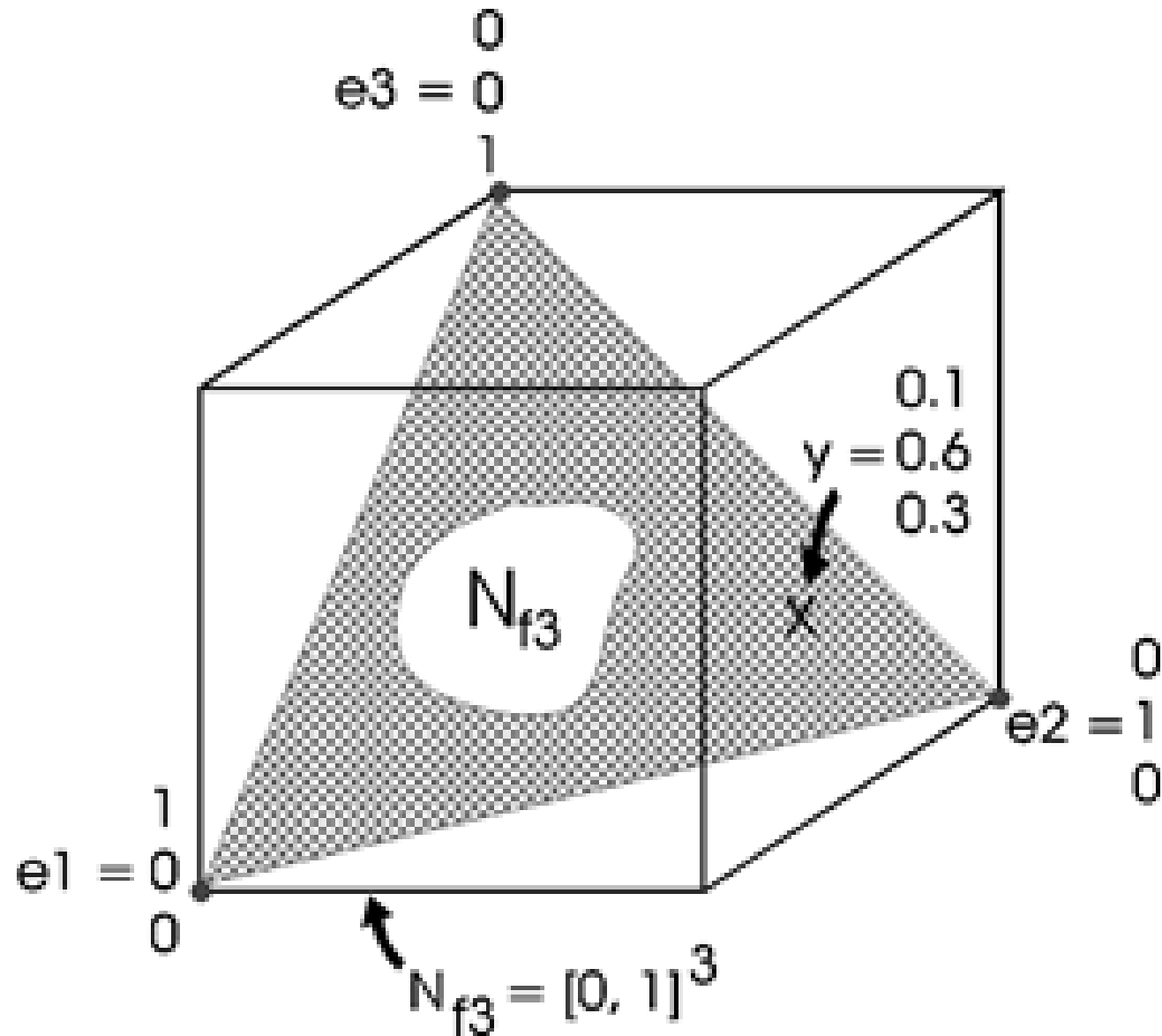


DATO OBJETO-
NUMÉRICO PARA EL
RECONOCIMIENTO
DE PATRONES

Dato objeto-numérico para el reconocimiento de patrones

- Es necesarios definir algunos objetos para llevar a cabo la descripción de los dato-objeto que ayudaran en la descripción de la tarea de agrupamiento.
- - (n) Objetos $\mathbf{O} = \{o_1, o_2, \dots, o_n\}$
- - Dato-Objeto $\mathbf{X} = \{x_1, x_2, \dots, x_n\} \subset \mathbf{R}^p$
- - $x_k \in \mathbf{R}^p = \phi(\mathbf{O}_k) = (x_{k1}, x_{k2}, \dots, x_{kp})^T,$
- - donde x_{kj} = k-ésima “observación” de la característica j y p es el número de características que definen (se seleccionaron) para cada objeto.

VECTORES- ETIQUETA PARA TRES CLASES



Espacios para la solución de agrupamiento

Se define una matriz universal de datos que contienen las características que describen a los objetos de dos formas, dura M_{cn} y difusa M_{fcn} . De aquí, se define una sub-matriz U , para el caso duro y difuso.

A continuación, se muestran dos de estas matrices U :

$U \in M_{cn}$	$U \in M_{fcn}$
<i>DURO</i>	<i>DIFUSO</i>
\updownarrow	\updownarrow
$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$	$\begin{vmatrix} 1 & .07 & 0 & .44 \\ 0 & .91 & 0 & .06 \\ 0 & .02 & 1 & .50 \end{vmatrix}$

“Tomar un segundo vistazo en la 3ª columna”

Vectores etiqueta

Para caracterizar los espacios de solución para los agrupadores y el diseño de clasificadores nos auxiliaremos de tres grupos de vectores etiqueta. Sea c el número de grupos, $2 < c < n - 1$, y sea:

$$N_{fcu} = \{ y \in R^c \mid y_k \in [0, 1] \forall k \} = \text{etiquetas sin restricción},$$

$$N_{fc} = \{ y \in N_{fcu} \mid \sum y_k = 1 \} = \text{etiquetas (restringidas)},$$

$$N_c = \{ y \in N_{fc} \mid y_k \in \{0, 1\} \forall k \} = \text{etiquetas duras para clase } c.$$

Vectores etiqueta

N_c es la base canónica (vector unidad) del espacio Euclidiano c ; N_{fc} , es una pieza de un hiperplano, es el caso convexo; y N_{fcu} es el hipercubo unitario en \mathbf{R}^c . De manera usual $N_c \subset N_{fc} \subset N_{fcu}$. Para $c = 3$, N_3 es el conjunto de los tres vértices de la porción triangular de la porción triangular de un hiperplano: $N_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. N_{f3} es el triángulo en si mismo. El vector de etiqueta típico.

Con N_{fc} se define un clasificador en \mathbf{R}^p como una función $\underline{\mathbf{D}}$ imaginada en N_{fc} . Esto es, los clasificadores son un tipo especial de vector de campo, que se puede denotar como $\underline{\mathbf{D}}$: $\mathbf{R}^p \rightarrow N_{fc} \subset \mathbf{R}^c$. Así, el valor de $\underline{\mathbf{D}}$ para cualquier $x \in \mathbf{R}^p$ es $\mathbf{y} = \underline{\mathbf{D}}(x)$, el vector etiqueta para x en N_{fc} . Con esta notación, se puede decir que $\underline{\mathbf{D}}$ es un clasificador certero (o duro, o convencional) si y solo si la imagen de \mathbf{R}^p bajo $\underline{\mathbf{D}}$ es:

$N_c (\underline{\mathbf{D}} [\mathbf{R}^p] = N_c)$; de otra forma el clasificador es difuso.

Análisis de grupos

Dado un conjunto de datos no etiquetados, el problema del agrupamiento en \mathbf{X} es asignar las etiquetas objeto (duras o difusas o probabilísticas) que identifiquen “subgrupos naturales” en \mathbf{X} . Algunas veces este problema es llamado aprendizaje no supervisado (o auto organización).

Sea (c) , $2 < c < n - 1$ y $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$.

Son vectores característicos no etiquetados $\in \mathbf{R}^p$. Cada grupo c son particiones de \mathbf{X} , caracterizados como conjuntos de (c_n) por medio de valores $\{u_{ik}\}$, es decir, u_{ik} $i = 1, \dots, c$, $k = 1, \dots, n$ es el grado de pertenencia que asigna el método de agrupamiento (agrupador) a cada dato-objeto por cada grupo, y que satisface alguna de las siguientes condiciones:

$$0 \leq u_{ik} \leq 1 \quad \forall i, k$$

$$0 < \sum u_{ik} < n \quad \forall i$$

$$\sum u_{ik} = 1 \quad \forall k$$

Análisis de grupos

Los valores $\{u_{ik}\}$ son los elementos de una matriz de $(c \times n)$ $U = [u_{ik}]$, se define como;

$$\begin{aligned} M_{fcnu} &= \{ U \in \mathbf{R}^{cn} \mid u_{ik} \forall i, k \} \\ M_{fcn} &= \{ U \in M_{fcnu} \mid u_{ik} \in [0, 1] \forall i, k \} \\ M_{cn} &= \{ U \in M_{fcn} \mid u_{ik} = 0 \text{ ó } 1 \forall i, k \}. \end{aligned}$$

Las ecs. definen, respectivamente conjuntos difusos no restringidos, conjuntos difusos restringidos o probabilísticos, y particiones-c certeras de \mathbf{X} . Se puede representar los algoritmos de agrupamiento como $A: \mathbf{X} \rightarrow M_{fcnu}$. Cada columna de U en M_{fcnu} (M_{fcn} , o M_{cn}) es un vector etiqueta de N_{fcnu} (N_{fc} , o N_c). La razón por la cual estas matrices son llamadas particiones sigue de la interpretación de u_{ik} como la membresía de \mathbf{X}_k en el i -ésimo subconjunto particionado (grupo) de \mathbf{X} .

Partición de grupos

La aplicación de cualquier técnica de agrupamiento de datos conduce necesariamente a una partición del espacio al cual pertenecen estos datos, en distintos grupos según sus propiedades. En efecto, un conjunto de datos n los cuales pertenecen a un espacio total de datos $\mathbf{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^p$, es dividido según sus propiedades en un determinado número de grupos; la agrupación de los mismos a su vez forma subclases o familias, denotadas por $c \in \{2, \dots, n - 1\}$. Así, a cada dato se le puede asignar una *etiqueta*, siendo esta etiqueta la familia a la que pertenece dicho dato. Por lo tanto, el agrupamiento se establece como una herramienta para particionar una estructura \mathbf{X} de datos.

Partición de grupos

Una partición generalmente se define como una matriz \mathbf{U} de dimensión $c \times n$; donde los elementos de esta matriz están definidas como $\{u_1, u_2, \dots, u_c\}$ subgrupos de cada dato. Por simplificación $u_i(x_k) = u_{ik}$, donde $x_k \in \mathbf{X}$, ($k = 1, \dots, n$; $i = 1, \dots, c$) y u_{ik} es el grado de membresía que se le asigna a cada elemento x_k .

Así, se define la matriz \mathbf{U} de particiones como:

$$U = \begin{array}{c} \begin{array}{ccc} & x_1 & x_k & x_n \\ \begin{array}{c} 1 \\ i \Rightarrow \\ c \end{array} & \begin{array}{c} \Downarrow \\ u_{ik} \\ \Downarrow \\ \text{clases} \end{array} & \Rightarrow \end{array} \Rightarrow \text{objetos}$$

Normas de partición

El conjunto de *partición dura* se define como:

$$N_{\text{hcn}} = \{ \mathbf{U} \in \{0,1\}^{\text{cn}} \mid \sum_{i=1}^c u_{ik} = 1, k = 1, \dots, n, \sum_{k=1}^n u_{ik} > 0, i = 1, \dots, c \},$$

El conjunto de *particiones difusas*:

$$N_{\text{fcn}} = \{ \mathbf{U} \in [0,1]^{\text{cn}} \mid \sum_{i=1}^c u_{ik} = 1, k = 1, \dots, n, \sum_{k=1}^n u_{ik} > 0, i = 1, \dots, c \},$$

El conjunto de *particiones difusas ruidosa*:

$$N_{\text{ncn}} = \{ \mathbf{U} \in [0,1]^{\text{cn}} \mid 0 < \sum_{i=1}^c u_{ik} \leq 1, k = 1, \dots, n, \sum_{k=1}^n u_{ik} > 0, i = 1, \dots, c \},$$

Y el conjunto de *particiones posibilísticas*:

$$N_{\text{pcn}} = \{ \mathbf{U} \in [0,1]^{\text{cn}} \mid \sum_{k=1}^n u_{ik} > 0, k = 1, \dots, n \},$$

Hay que notar que $N_{\text{hcn}} \subset N_{\text{fcn}} \subset N_{\text{ncn}} \subset N_{\text{pcn}} \subset [0, 1]^{\text{cn}}$.

Modelo de agrupamiento difuso

Una de las características principales que presentan los modelos de agrupamiento *fuzzy c-means* y *possibilistic c-means* es que se pueden obtener modelos para distribuciones de datos en forma de cúmulos, inclusive si estos llamados cúmulos o hiper-esferas tienen cierto grado de traslape. A continuación, se presentarán estos dos modelos, poniendo especial énfasis en los fundamentos matemáticos en que se sustentan, así como sus algoritmos de implementación computacional.

Modelo de agrupamiento difuso

El modelo FCM se apoya en la minimización de una función objetivo definida como:

$$J_{\text{FCM}}(\mathbf{U}, \mathbf{V}; \mathbf{X}) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m \|x_k - v_i\|^2$$

y

$$\sum_{i=1}^c u_{ik} = 1, k=1, \dots, n,$$

donde \mathbf{X} es un conjunto de datos, $\mathbf{U} \in \mathbf{N}_{\text{fcn}}$ y $\mathbf{V} = \{v_1, \dots, v_c\} \subset \mathbb{R}^p$ es un conjunto de puntos prototipo (centros característicos de cada grupo).

Modelo de agrupamiento posibilístico

En el caso del modelo de agrupamiento PCM la condición anterior es eliminada, lo que lleva a una partición posibilística en vez de una partición difusa. Para evitar una solución trivial $u_{ik} = 0, i = 1, \dots, c, k = 1, \dots, n$, Krishnapuram y Keller, añadieron un término para considerar datos cuyos grados de membresía sean muy bajos para ser considerados en la función definida en el caso de agrupamiento difuso, por lo que el modelo PCM se desarrolla debido a la minimización de la función objetivo definida como:

$$J_{\text{PCM}}(\mathbf{U}, \mathbf{V}; \mathbf{X}) = \sum_{k=1}^n \sum_{i=1}^c \left(u_{ik}^m \|x_k - v_i\|^2 + \eta_i (1 - u_{ik})^m \right),$$

Modelos de agrupamiento posibilístico

donde $\mathbf{U} \in \mathbf{N}_{\text{pcn}}$, v_i siendo el centro inicial para cada grupo, y los parámetros de distancia $\eta_1, \dots, \eta_c \in \mathbb{R}^+ \setminus \{0\}$ son especificados por el usuario. Krishnapuram y Keller recomiendan varios métodos para seleccionar a η_i , por ejemplo:

$$\eta_i = K \frac{\sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2}{\sum_{k=1}^n u_{ik}^m},$$

donde $K \in \mathbb{R}^+ \setminus \{0\}$ típicamente es 1; o, también:

$$\eta_i = \frac{\sum_{k=1}^n (u_{ik})^{\geq \alpha} \|x_k - v_i\|^2}{\sum_{k=1}^n (u_{ik})^{\geq \alpha}},$$

donde $\alpha \in (0, 1]$ es una partición de corte α certero definida como:

$$(u_{ik})^{\geq \alpha} = \begin{cases} 0, & \text{si } u_{ik} < \alpha \\ 1 & \text{si } u_{ik} \geq \alpha \end{cases},$$

Algoritmos de optimización alternativa

La mayoría de los modelos de agrupamiento son optimizados por un algoritmo de "optimización alternativa" (OA). A continuación, se discutirá el caso de los puntos prototipos. Los prototipos son inicializados como $\mathbf{V}^{(0)}$. En cada paso de optimización t la partición $U^{(t)}$ y los prototipos $\mathbf{V}^{(t)}$ son optimizados. El algoritmo se termina cuando un máximo número de iteraciones se ha alcanzado $t = t_{\max}$, o cuando las aproximaciones sucesivas del prototipo se han estabilizado, $\|\mathbf{V}^{(t)} - \mathbf{V}^{(t-1)}\|_{\text{error}} < \varepsilon_V$, donde $\varepsilon_V > 0$ y $\|\cdot\|_{\text{error}}$ es una norma apropiada. Un algoritmo similar OA' se obtiene si se cambia U y \mathbf{V} en OA. En OA', la partición $U^{(0)}$ es inicializada aleatoriamente y OA' termina si $\|U^{(t)} - U^{(t-1)}\|_{\text{error}} < \varepsilon_U$, donde $\varepsilon_U > 0$. OA requiere menos comparaciones y almacenamiento que OA'.

Algoritmo de OA-FCM

Para la solución de OA los modelos de agrupamiento FCM y PCM, las ecuaciones para actualizar U y V son las condiciones necesarias de extremo local de sus respectivas funciones objetivo. Más específicamente, cuando $\|\cdot\|$ es una norma de producto interno $\|x\|_A^2 = x^T A x$, para OA-FCM, cuando $\|x_k - v_j\|_A > 0 \quad \forall j, k$:

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{\|x_k - v_i\|_A}{\|x_k - v_j\|_A} \right)^{\frac{2}{m-1}}, \quad i = 1, \dots, c, \quad k = 1, \dots, n$$

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m}, \quad i = 1, \dots, c$$

Algoritmo de OA-PCM

Para la OA-PCM se tiene que $\eta_i > 0 \forall i$,

$$u_{ik} = \frac{1}{1 + \left(\frac{\|x_k - v_i\|}{\sqrt{\eta_i}} \right)^{\frac{2}{m-1}}}, \quad i = 1, \dots, c, \quad k = 1, \dots, n$$

OA-PCM actualiza los prototipos de la misma forma como lo hace OA-FCM.

Las secuencias generadas $\{(U^{(1)}, V^{(1)}), (U^{(2)}, V^{(2)}), \dots, (U^{(*)}, V^{(*)})\}$ generadas por OA-FCM para $m > 1$ fijas siempre convergen (o contienen una secuencia que lo hace) a un mínimo local de J_{FCM} o J_{PCM} , respectivamente. Se puede encontrar un número razonable de iteraciones para OA-FCM para $t_{\max} \approx 100$.

Algoritmo Fuzzy C-Means (FCM)

Dado el grupo de datos \mathbf{Z} , seleccionar el número de grupos $2 < c < N-1$, el exponente de ponderación $m > 1$, la tolerancia para terminar $\varepsilon > 0$ y una matriz de norma inducida \mathbf{A} . Inicializar la matriz de partición en forma aleatoria. Tal que $\mathbf{U}^{(0)} \in M_{fc}$.

Repetir para $l = 1, 2, \dots$

Paso 1. Calcular los grupos prototipos (media):

$$v_i^{(l)} = \frac{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m z_k}{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m}, \quad 1 \leq i \leq c.$$

Paso 2. Calcular las distancias:

$$D_{ikA}^2 = \left(z_k - v_i^{(l)} \right)^T A \left(z_k - v_i^{(l)} \right), \quad 1 \leq i \leq c, 1 \leq k \leq N.$$

Paso 3: Actualizar la matriz de partición:

Si $D_{ikA} > 0$ para $1 \leq i \leq c, 1 \leq k \leq N$,

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{m-1}}},$$

de otra forma:

$$\mu_{ik}^{(l)} = 0 \text{ si } D_{ikA} = 0, \text{ y } \mu_{ik}^{(l)} \in [0, 1] \text{ con } \sum_{i=1}^c \mu_{ik}^{(l)} = 1.$$

Hasta que $\| \mathbf{U}^{(l)} - \mathbf{U}^{(l-1)} \| < \varepsilon$.

Algoritmo Possibilistic C-Means (PCM)

Dado el conjunto \mathbf{Z} y una buena partición inicial $\mathbf{U}^{(0)} \in M_{pc}$, seleccionar el número de grupos $1 < c < N$, el exponente de ponderación $m > 1$, la tolerancia de terminación $\varepsilon > 0$ y la matriz de norma inducida \mathbf{A} . Estimar η_i utilizado:

$$\eta_i = \frac{\sum_{k=1}^N (\mu_{ik})^m D_{ikA}^2}{\sum_{k=1}^N (\mu_{ik})^m}.$$

Repetir para $l = 1, 2, \dots$

Paso 1: Calcular los grupos prototipo (medios):

$$v_i^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m z_k}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m}, \quad 1 \leq i \leq c.$$

Paso 2: Calcular las distancias:

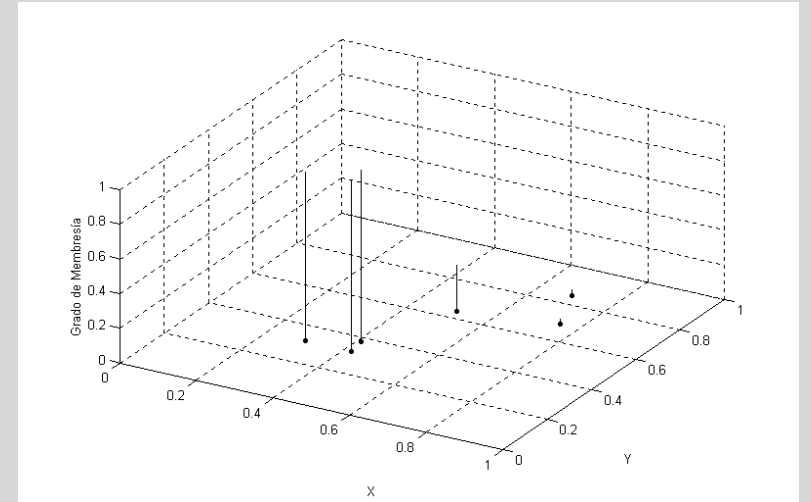
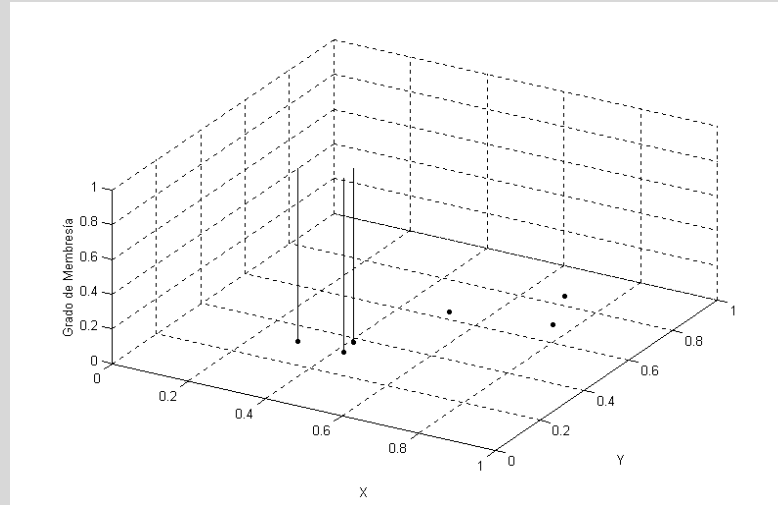
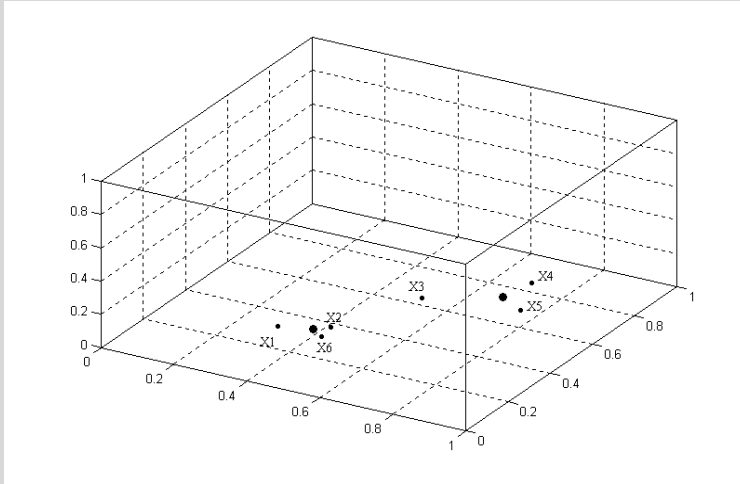
$$D_{ikA}^2 = (z_k - v_i^{(l)})^T A (z_k - v_i^{(l)}), \quad 1 \leq i \leq c, 1 \leq k \leq N.$$

Paso 3: Actualizar la matriz de partición:

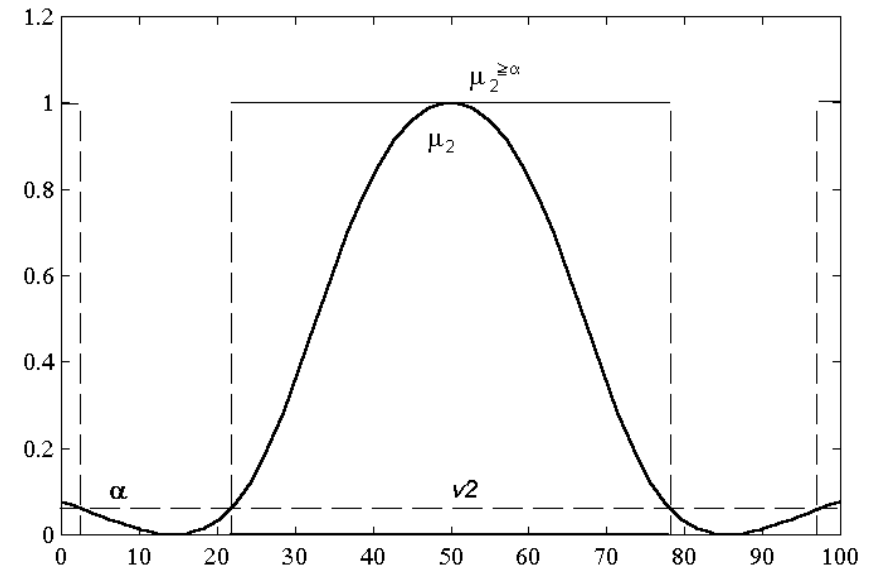
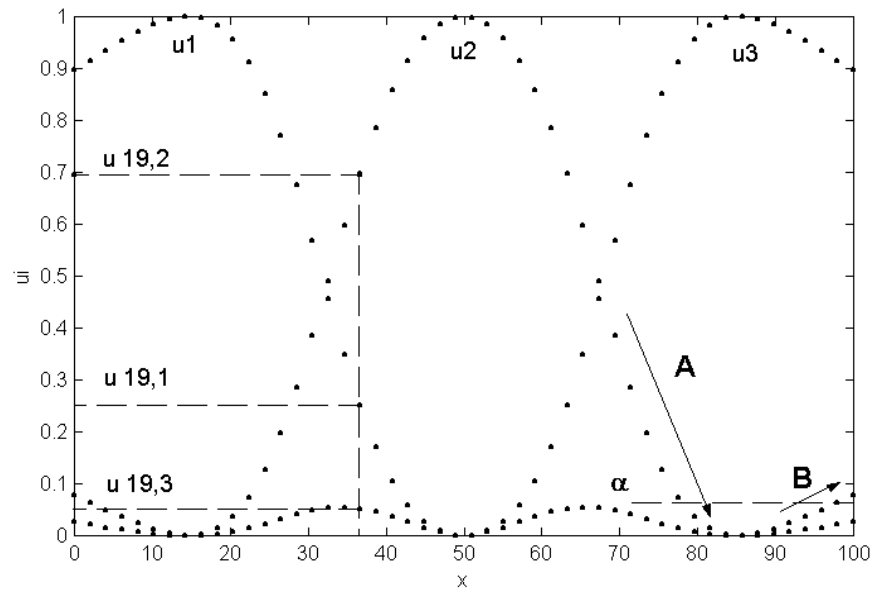
$$\mu_{ik}^{(l)} = \frac{1}{1 + \left(\frac{D_{ikA}}{\eta_i} \right)^{2/(m-1)}}, \quad 1 \leq i \leq c, 1 \leq k \leq N,$$

hasta que $\| \mathbf{U}^{(l)} - \mathbf{U}^{(l-1)} \| < \varepsilon$.

Ejemplos de partición de datos por medio de métodos de agrupamiento duro y difuso



Funciones de membresía $u_i(x_k)$ obtenidas por medio de OA-FCM



Modelos de estimación alternativa de grupos

Otro método analizado es el denominado de *estimación alternativa de grupos* EAG (alternating cluster estimation-ACE), propuesto por J. A. Runkler y J. C. Bezdek. Este método se diferencia del FCM-AO y el PCM-AO en que no se utiliza la minimización de una función objetivo. Así, este método se basa en una arquitectura de iteraciones alternativas, pero las funciones de membresía y las funciones prototipo pueden ser seleccionadas por el usuario. Virtualmente cualquier modelo de agrupamiento puede ser realizado como un ejemplo del modelado por estimación alternativa de grupos. El algoritmo utilizado en este método es el siguiente:

1. Seleccionar las ecuaciones que en este caso definen a la función prototipo, la cual puede ser alguna de las utilizadas por la metodología AO (funciones Gaussianas, campana generalizada, etc.)
2. Seleccionar los parámetros de control, es decir, $t_{\text{máx}}$ (número de iteraciones máxima), $\| \cdot \|_{\text{error}}$ (norma apropiada), ε_V (razón de error, es decir la que define a U^*)
3. Inicializar los centros de las funciones prototipo
4. Continuar mientras no se alcance el número de iteraciones máxima ($t \leq t_{\text{máx}}$) o $\| V^{(t)} - V^{(t-1)} \| < \varepsilon_V$
5. Calcular la nueva partición $U^{(t)}$ ($V^{(t-1)}, X, \dots$)
6. Calcular el nuevo prototipo $V^{(t)}$ ($V^{(t-1)}, X, \dots$)

Modelos de estimación alternativa de grupos

Una función prototipo puede ser definida por un hipercono, cuya ecuación está dada por:

$$\mu_i(x) = \begin{cases} 1 - \left(\frac{\|x - v_i\|}{r_i} \right)^\alpha, & \text{para } \|x - v_i\| \leq r_i \quad \alpha > 0 \\ 0, & \text{de otra forma} \end{cases}$$

Donde v_i y r_i son el centro y el radio del cono, respectivamente, α es una constante la cual permite modificar el grado de pertenencia de x , de tal forma que si $\alpha < 1$, el cono se hace más angosto. Normalmente $\alpha = 1$.

Una ventaja que se tiene al utilizar esta función es que permite definirla únicamente a partir de su centro y un radio inicial el cual no es necesario cambiar. Una forma conveniente de definir el radio r_i , es proporcionada por la siguiente ecuación:

$$r_i = \frac{\sqrt{\sum_{l=1}^P s_l^2}}{c + 1}$$

Donde $s_l = l_{\text{mín}} - l_{\text{máx}}$, $l = 1, \dots, P$, y $l \in \mathfrak{R}^P$, siendo \mathfrak{R}^P el espacio que describe a los datos a ser agrupados.

Algoritmo de los conos danzantes (DC)

Dado el grupo de datos \mathbf{Z} , seleccionar el número de grupos $1 < c < N$, los centros iniciales $\mathbf{V}^{(l)}$, tolerancia de terminación $\varepsilon > 0$ y una matriz norma inducida \mathbf{A} .

Paso 1: Calcular $s_1 = \text{abs}(\max(\mathbf{Z}_x) - \min(\mathbf{Z}_x))$ y $s_2 = \text{abs}(\max(\mathbf{Z}_y) - \min(\mathbf{Z}_y))$.

Paso 2: Calcular el radio de los hiperconos:

$$r_i = \frac{\sqrt{\sum_{l=1}^p s_l^2}}{c+1}.$$

Paso 3: Calcular las funciones de membresía para cada uno de los hiperconos:

$$\mu_i = \begin{cases} 1 - \frac{D_{ikA}^2}{r_i} & \text{para } D_{ikA}^2 \leq r_i \\ 0 & \text{de otra forma} \end{cases}.$$

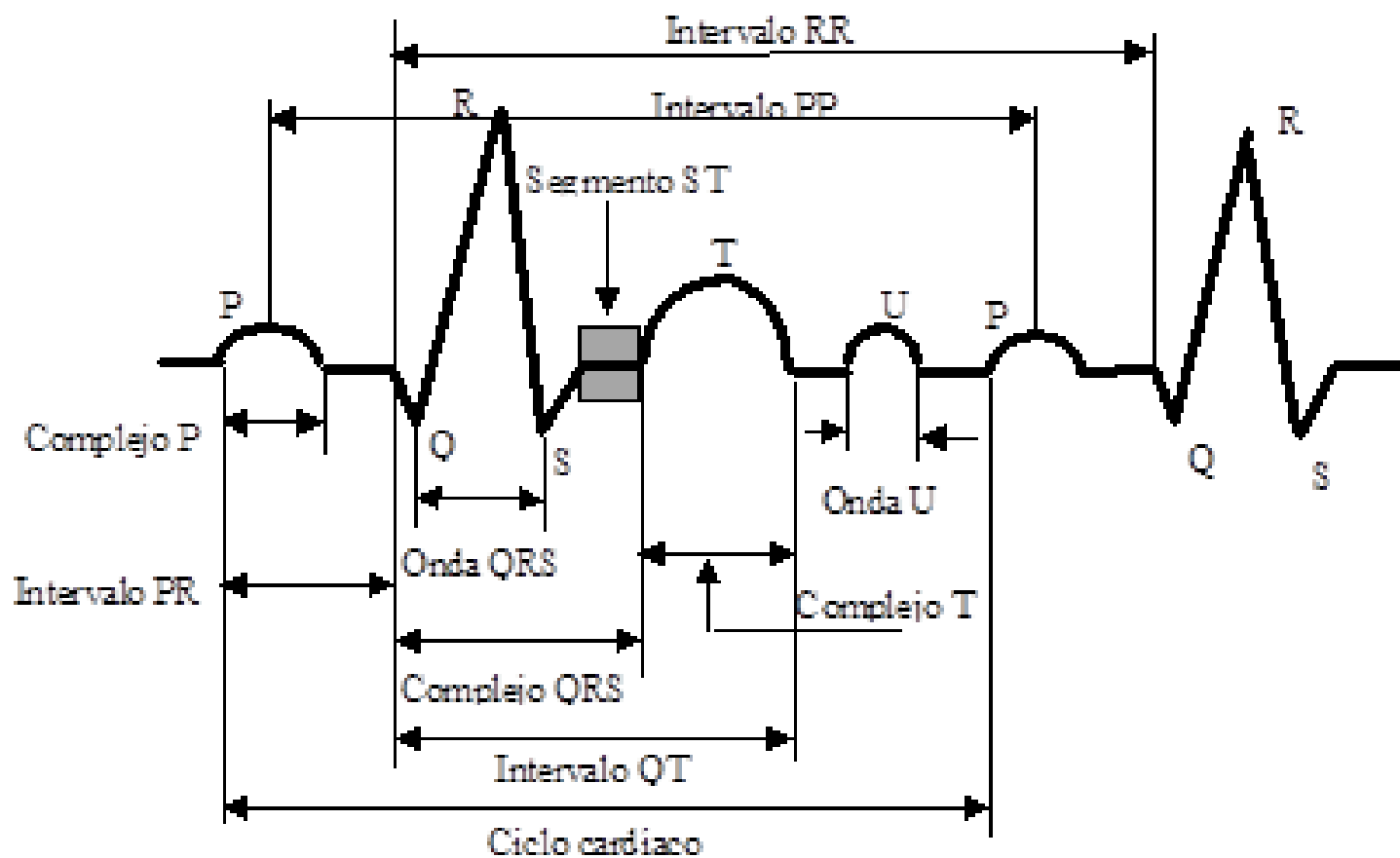
Paso 4. Calcular las distancias:

$$D_{ikA}^2 = \left(\mathbf{z}_k - \mathbf{v}_i^{(l)} \right)^T \mathbf{A} \left(\mathbf{z}_k - \mathbf{v}_i^{(l)} \right), \quad 1 \leq i \leq c, 1 \leq k \leq N.$$

Paso 1: Calcular los grupos prototipo (medios):

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m \mathbf{z}_k}{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m}, \quad 1 \leq i \leq c.$$

hasta que $\| \mathbf{V}^{(l)} - \mathbf{V}^{(l-1)} \| < \varepsilon$.



RECONOCIMIENTO DE ARRITMIAS VENTRICULARES

Tipos de arritmias

1) Originadas en las *aurículas*:

- a) Arritmia sinusal
- b) Taquicardia sinusal
- c) Bradicardia sinusal
- d) Síndrome del seno enfermo
- e) Contracción supraventricular o contracción prematura
- f) Taquicardia supraventricular, taquicardia auricular parasistólica
- g) Aleteo auricular
- h) Fibrilación auricular
- i) Síndrome de Wolf-Parkinson-White

2) Originadas en los *ventrículos*:

- a) Taquicardia ventricular
- b) Fibrilación ventricular
- c) Complejos ventriculares prematuras (PVC)

Arritmias PVC

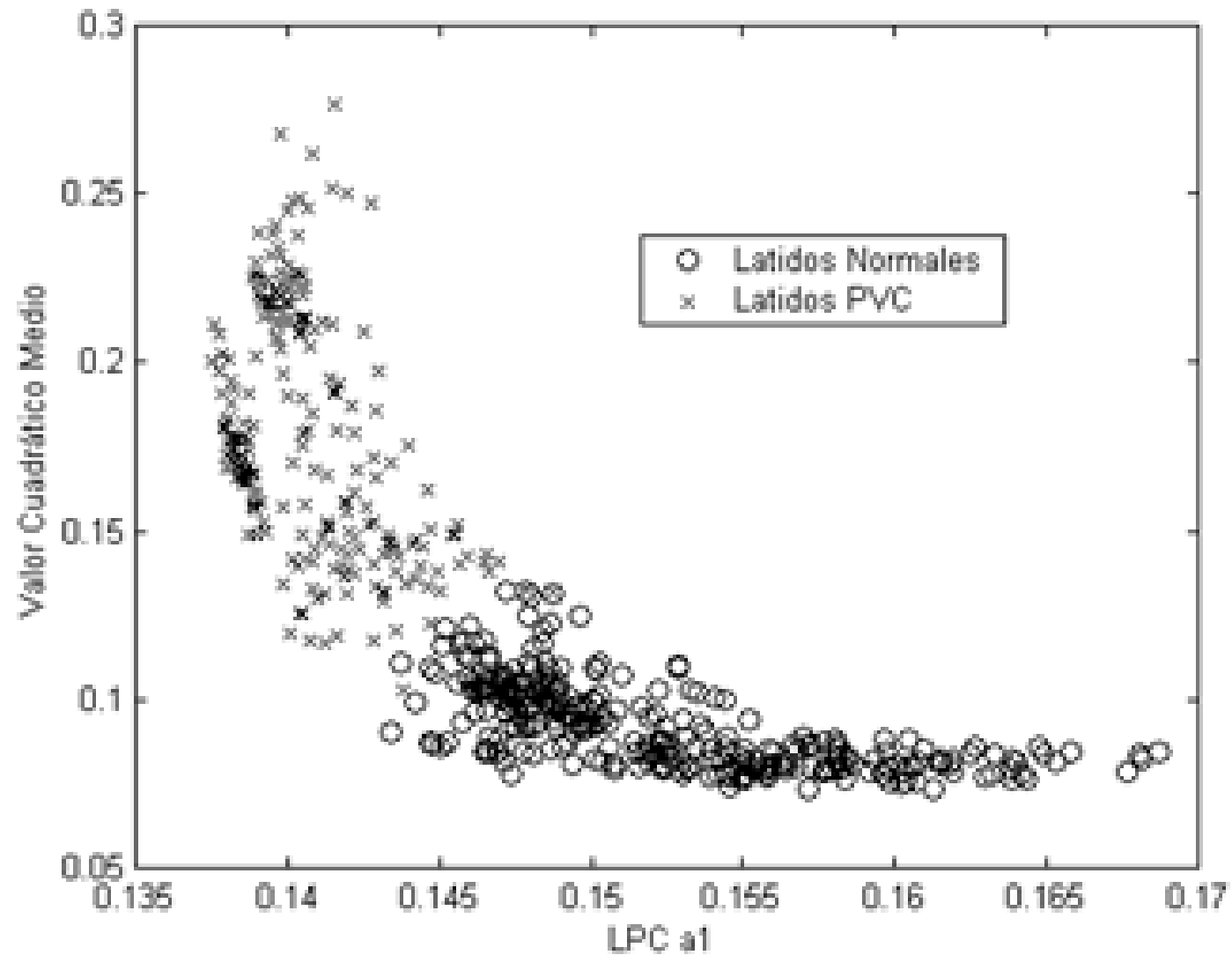
Los *complejos ventriculares prematuros* (PVC) pueden surgir de un foco ectópico situado en cualquier parte del miocardio ventricular o por reentrada de un impulso (*ventricular o supraventricular*) a una zona de tejido ventricular. La característica principal que se manifiesta en el ECG es la presencia de un *complejo QRS* que difiere notablemente de los otros QRS del registro por su forma y su tamaño. Los complejos ventriculares son anchos y de aspecto anormal, con una duración superior 0.12 seg. El segmento ST y la onda T suelen estar desplazados en sentido opuesto al de la deflexión principal de QRS. A veces, un complejo PVC se produce entre dos impulsos sinusales sin modificar el ritmo sinusal, lo que se conoce como complejo PVC interpolado.

Extracción de rasgos característicos de las arritmias PVC

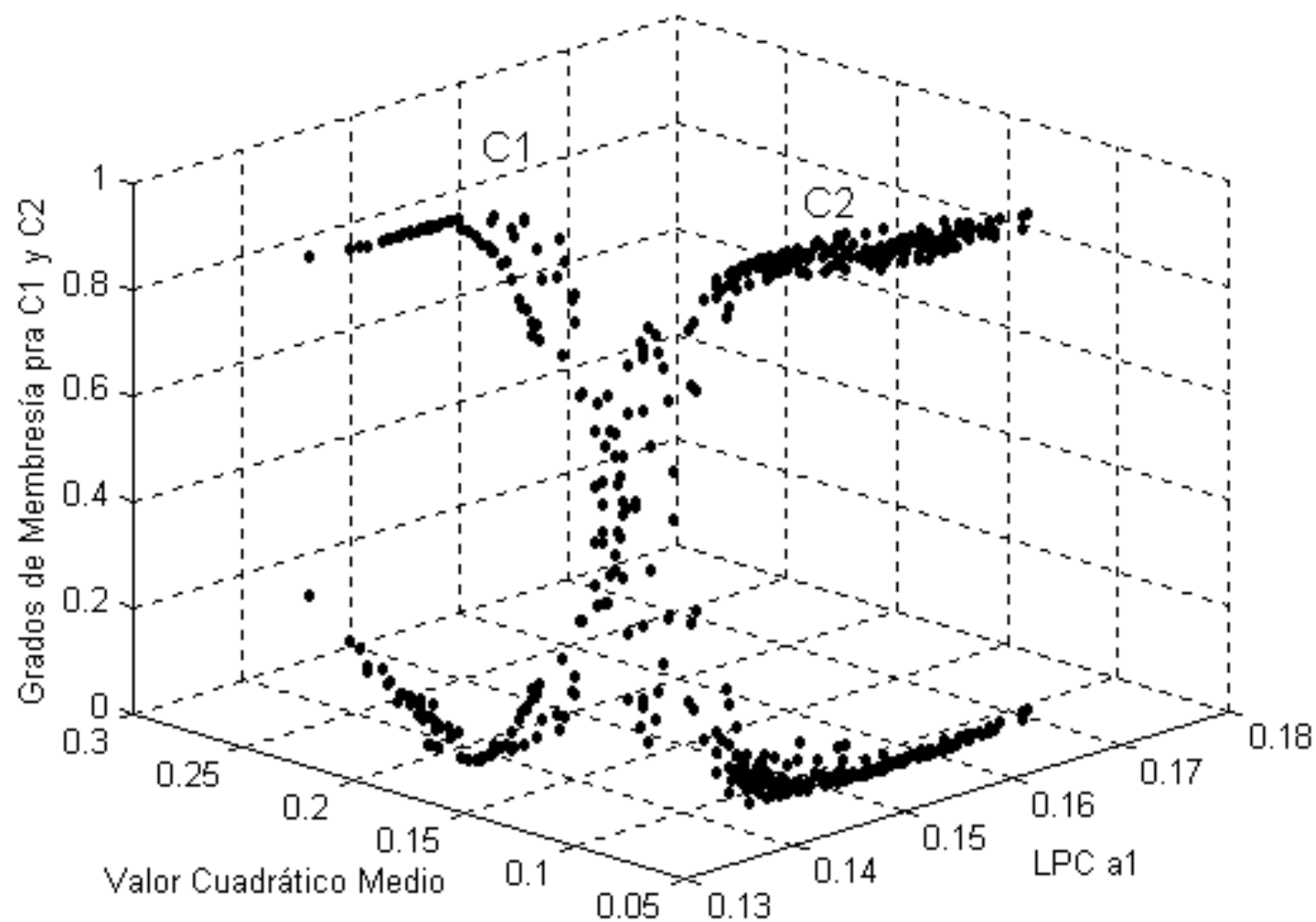
A continuación, se presenta la aplicación de los métodos FCM y PCM a un grupo de datos que fue generado a partir de un grupo de *señales electrocardiográficas* (ECG). Este grupo de señales que se utilizó está contenido en una base de datos del MIT, y en este caso se tomaron señales electrocardiográficas con *señales normales* y con *arritmia PVC* de siete pacientes. A cada señal se le aplicó un *preprocesamiento* que, básicamente, consistió en aplicar un *filtro pasa altas* con una frecuencia de corte de 1 Hz, seguido por un filtro pasa altas con corte a 20 Hz, así como un filtro de rechazo de banda a 60 Hz. Posteriormente, se realizó el centrado de cada segmento QRS de la señal electrocardiográfica, utilizando la media aritmética para cada segmento.

Extracción de rasgos característicos de las arritmias PVC

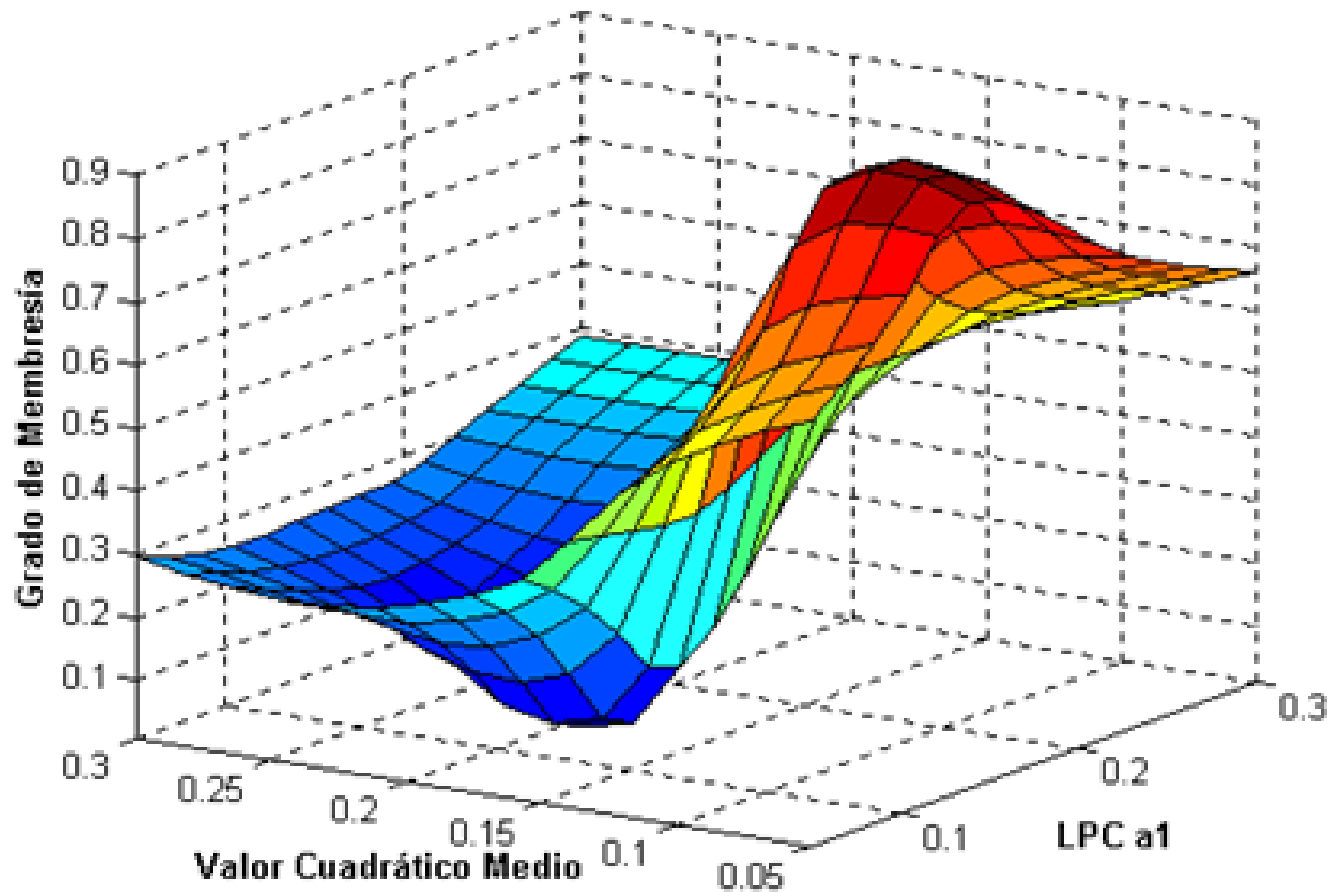
A continuación, se escalaron las señales entre los valores -1 y $+1$; y se aplicó una *ventana de Hamming*, la cual permitió reducir discontinuidades y resaltar la señal de la onda R del electrocardiograma, después, para minimizar la cantidad de datos que permitieran representar cada señal ECG se aplicó el método de cálculo de coeficientes de predicción lineal, encontrándose que se necesitan únicamente dos coeficientes de predicción lineal, definidos como a_1 y a_2 , para dicha representación. Con estos dos coeficientes para cada señal, se puede representar de una manera apropiada un latido de la señal electrocardiográfica. Para obtener los coeficientes, se utilizó el método de *Burg* de máxima entropía; de manera adicional también se calculó el valor cuadrático medio de cada complejo QRS. Por lo tanto, de cada señal electrocardiográfica se extrajeron tres características: los coeficientes de predicción lineal a_1 y a_2 , y el valor cuadrático medio. Así, se obtuvo una distribución de datos graficando el coeficiente LPC a_1 contra el valor cuadrático medio, obteniéndose una distribución adecuada para realizar la clasificación de arritmias tipo PVC.



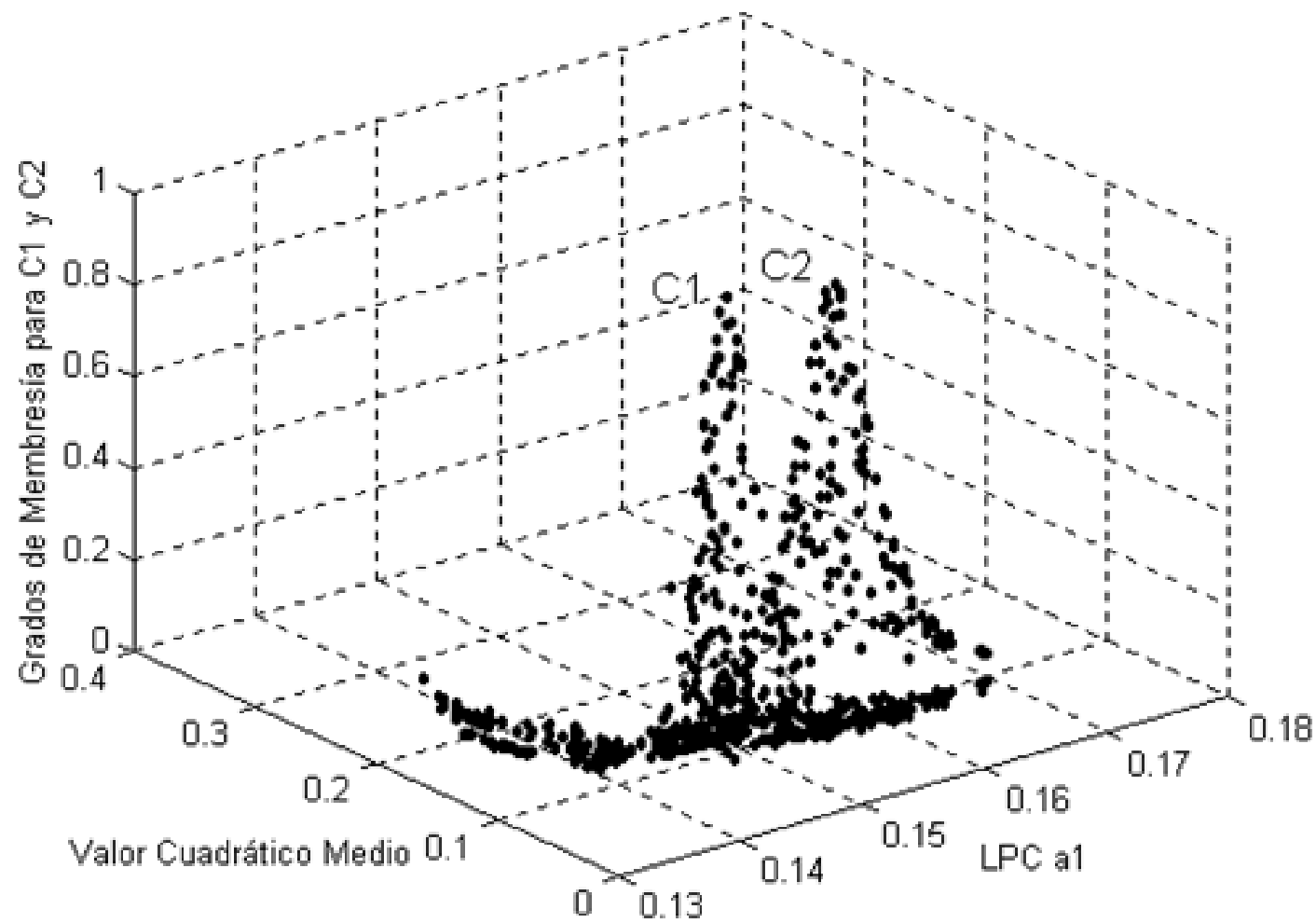
DISTRIBUCIÓN DE LOS
COEFICIENTES LPC
OBTENIDOS PARA
LATIDOS NORMALES Y
PVC DE DIFERENTES
SEÑALES ECG PARA EL
VALOR CUADRÁTICO
MEDIO CONTRA A1.



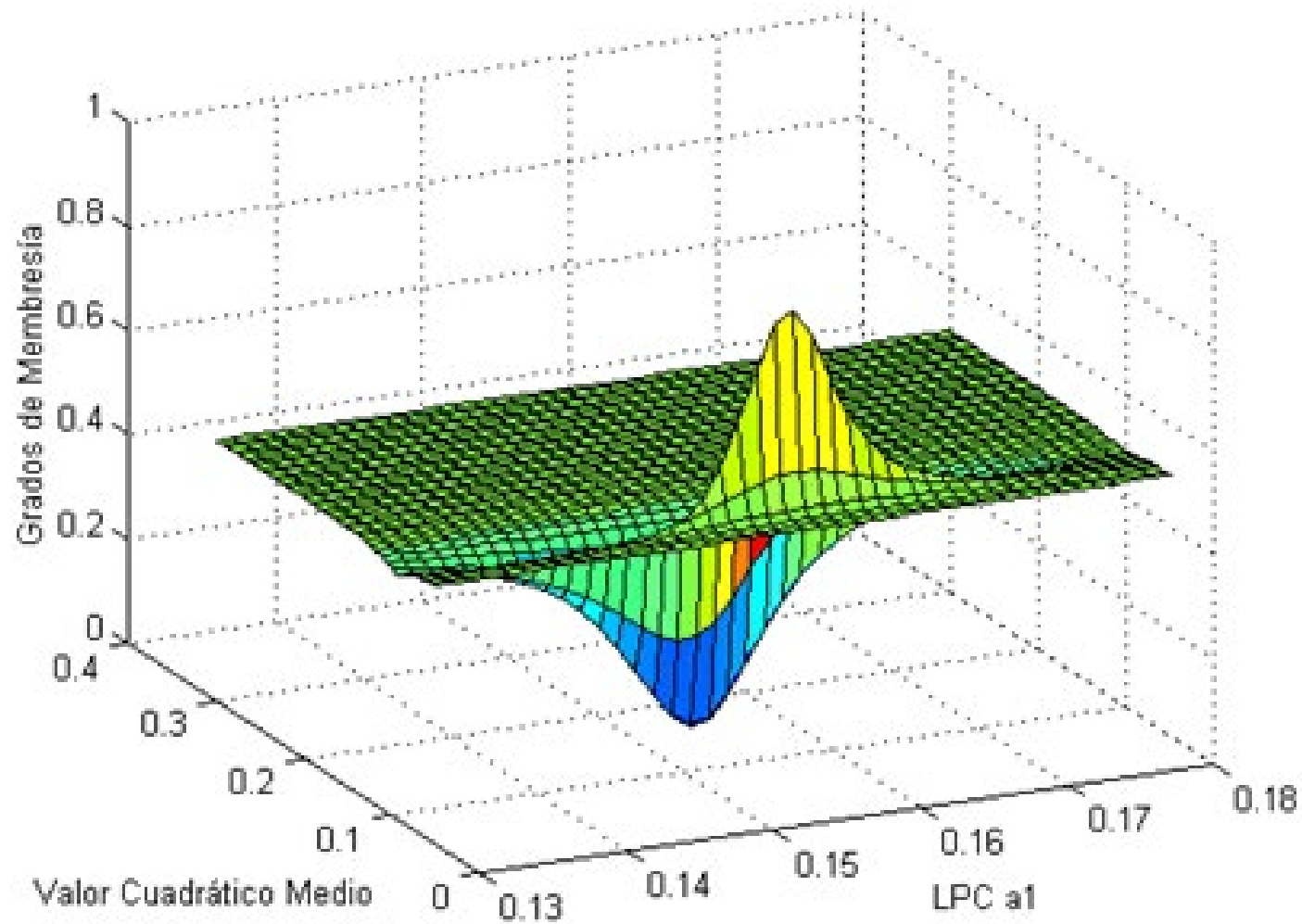
Grados de
membresía obtenidos
después de aplicar
FCM al grupo de
datos con arritmia
PVC y datos
normales sin
etiquetar.



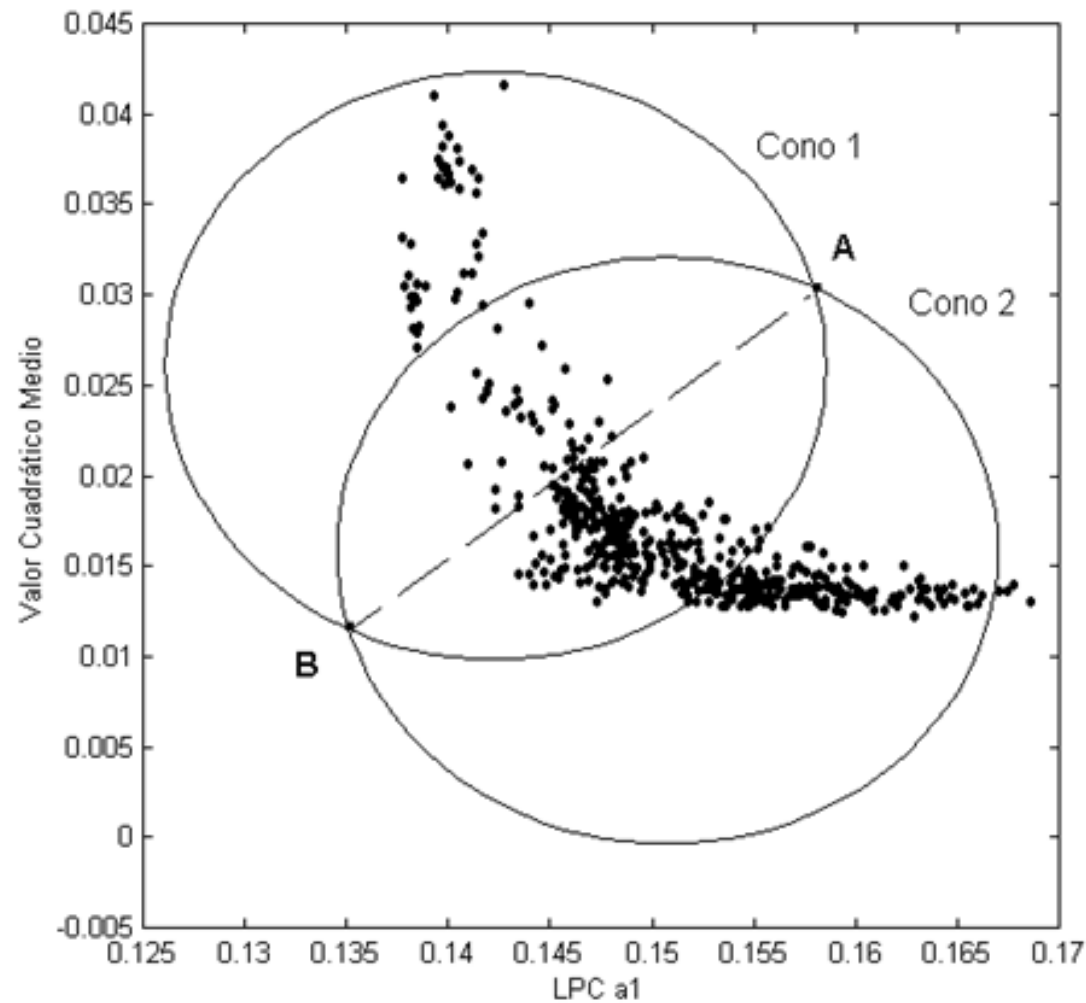
Superficie generalizada
de datos obtenidos
después de aplicar FCM
y aplicando el algoritmo
que permite aprovechar
una sola función para
separar el grupo de datos
normales y PVC.



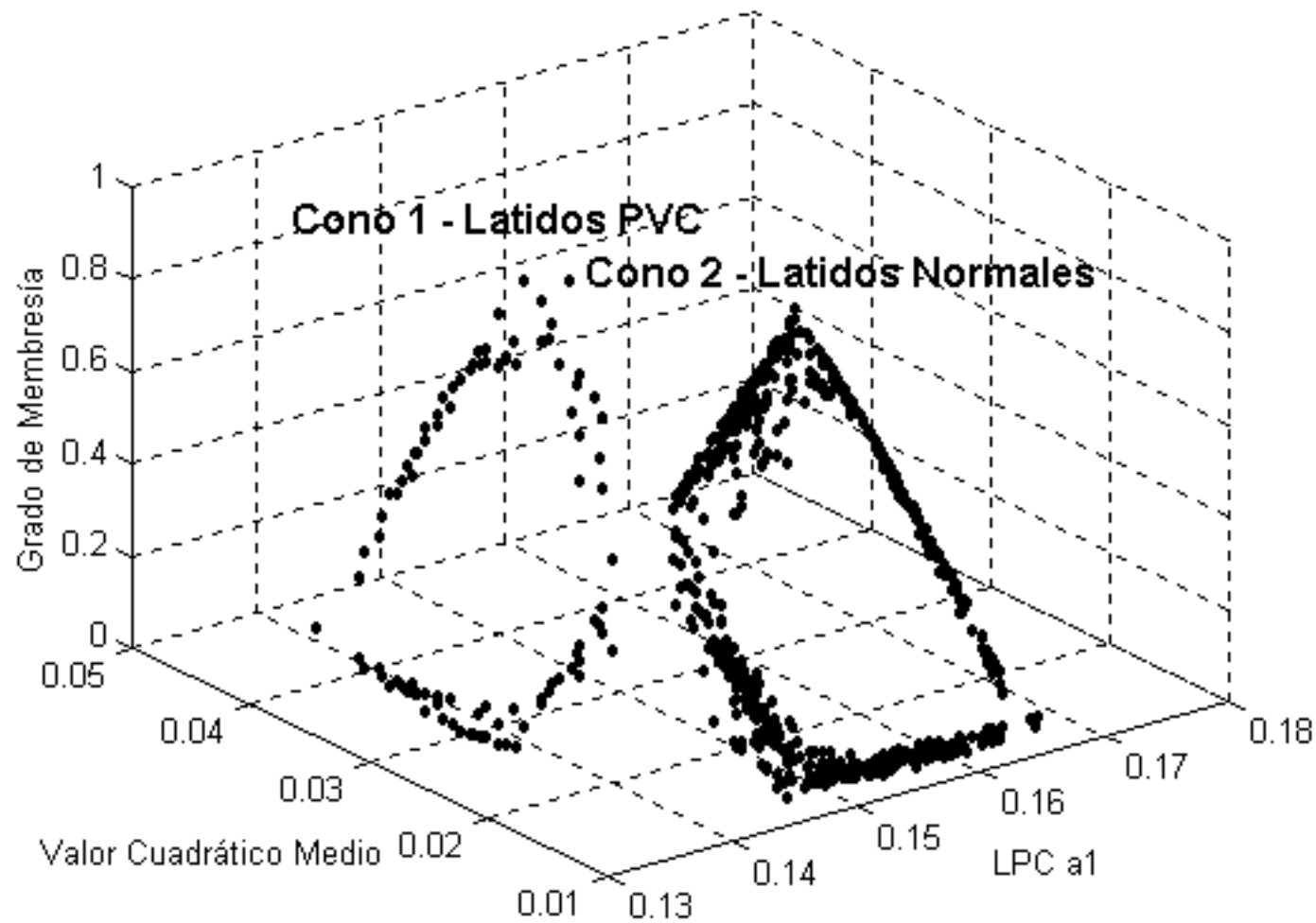
Grados de membresía
obtenidos después de
aplicar PCM al grupo de
datos con arritmia PVC
y datos normales sin
etiquetar.



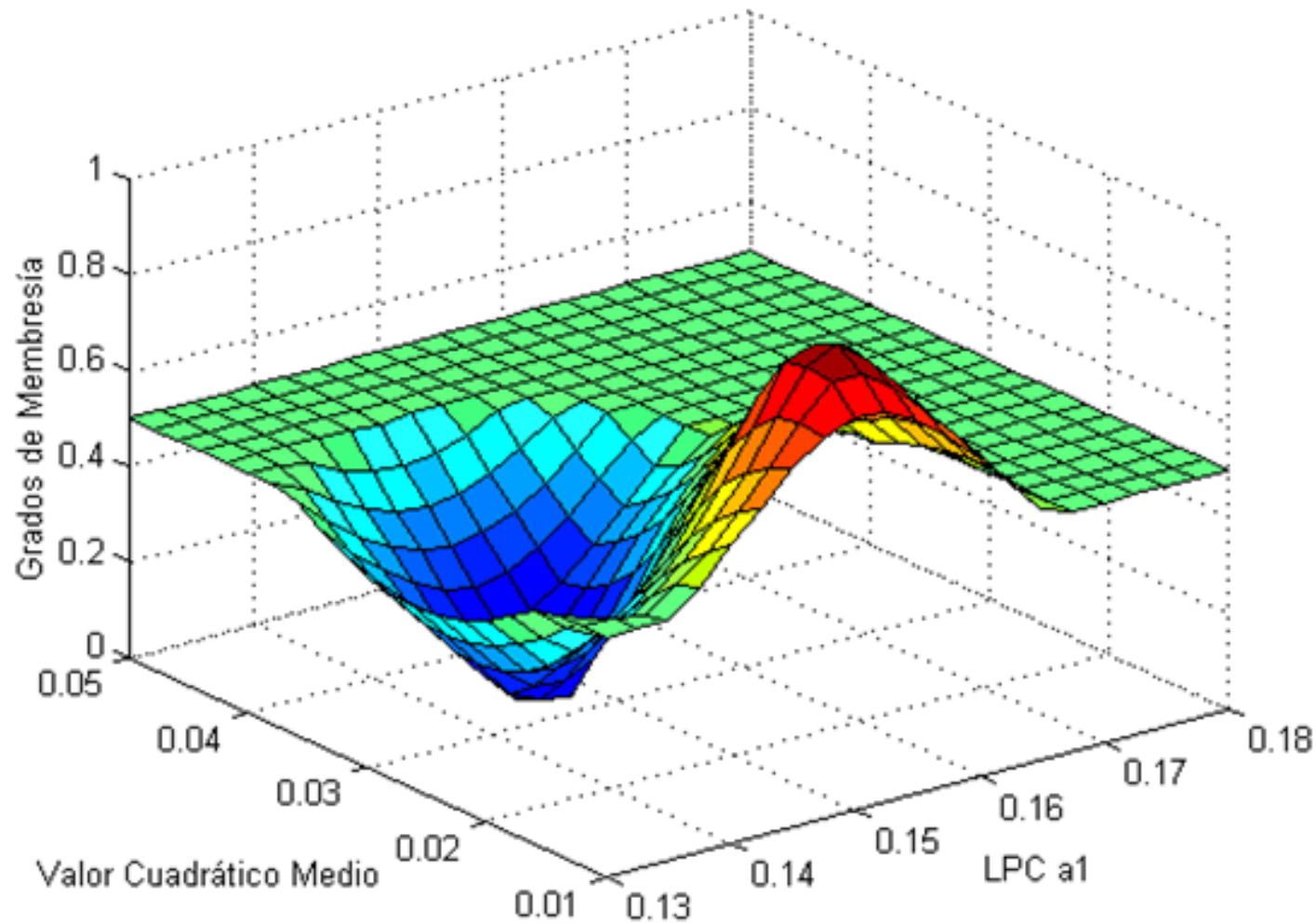
Superficie
generalizada para los
grupos de datos PVC
y normales generada
por el método de
PCM-AO.



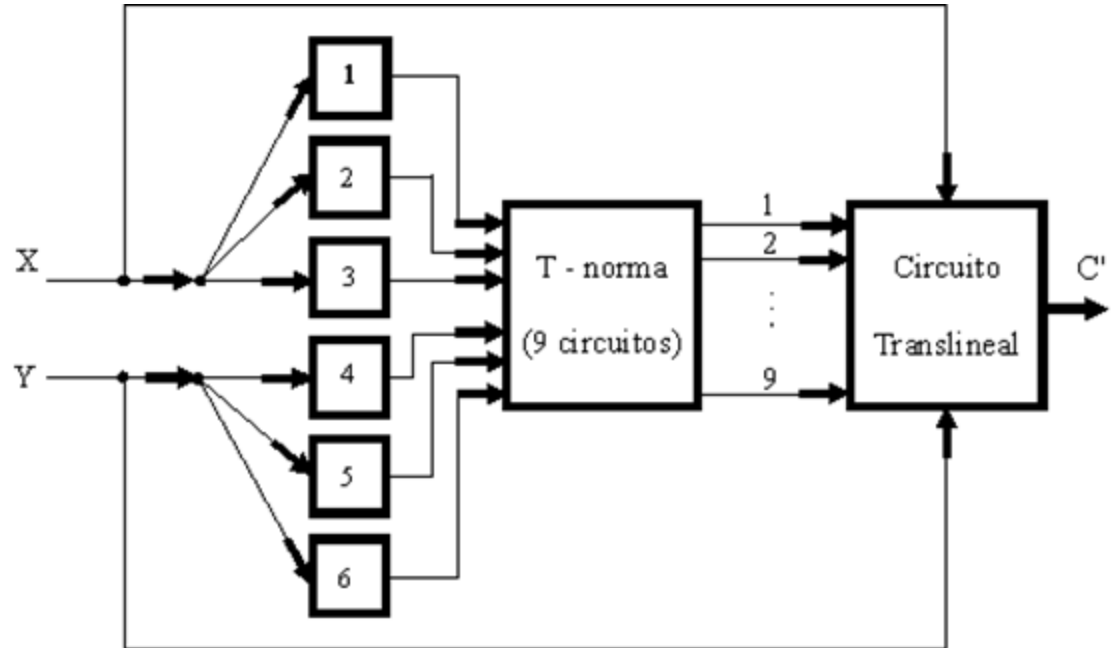
Aplicación del algoritmo EGA basado en conos danzantes al grupo de datos Normales y arritmia PVC. La línea AB separa los dos conjuntos de datos (por cuestiones de escala no se observan círculos perfectos, sino elipses.)



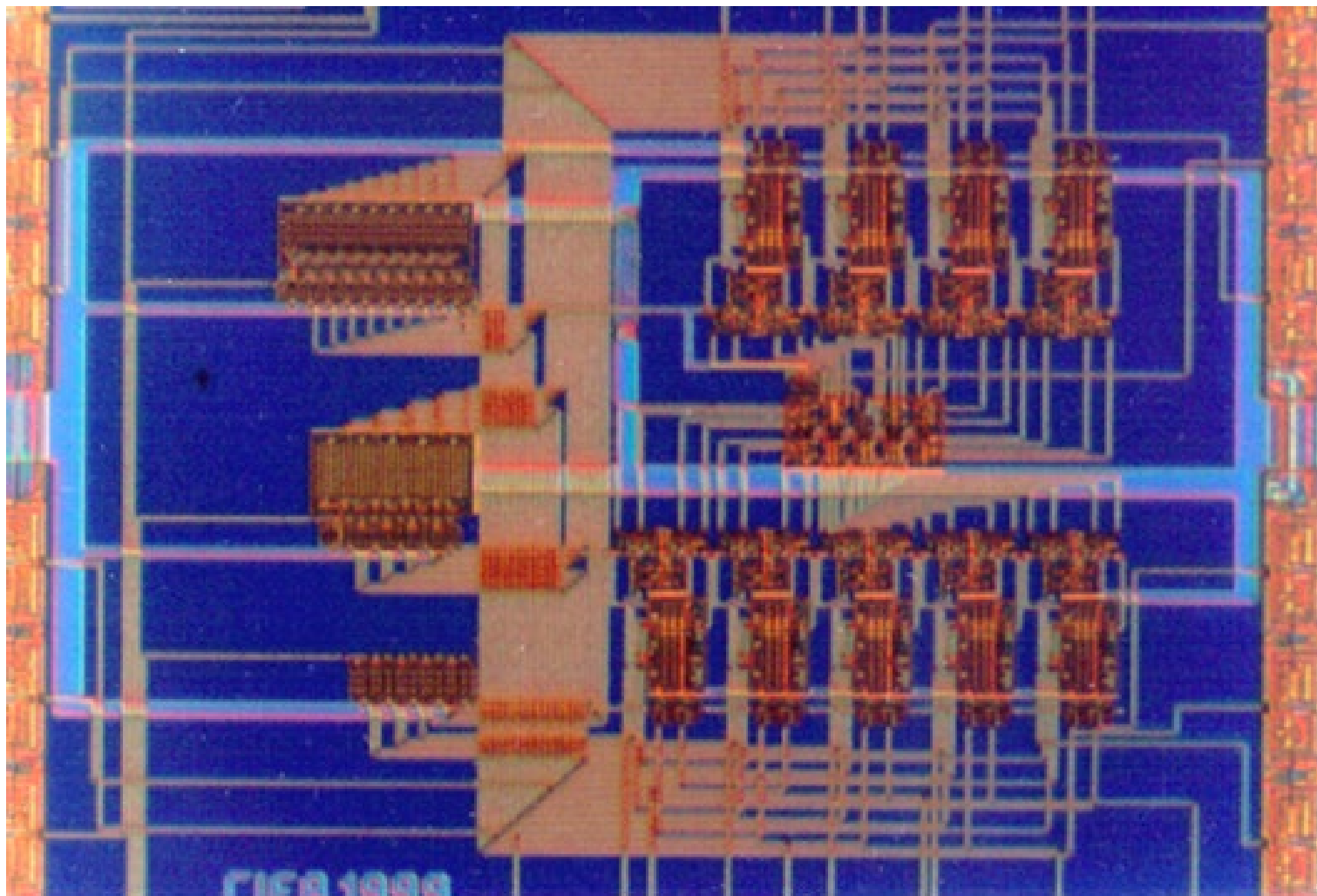
Grados de membresía obtenidos para cada conjunto de datos. El cono 1, contiene los grados de membresía del grupo de datos PVC, y el cono 2 contiene los grados de membresía de los datos normales.



Superficie
generalizada de datos
normales y PVC
utilizando el método
EAG con conos
danzantes.

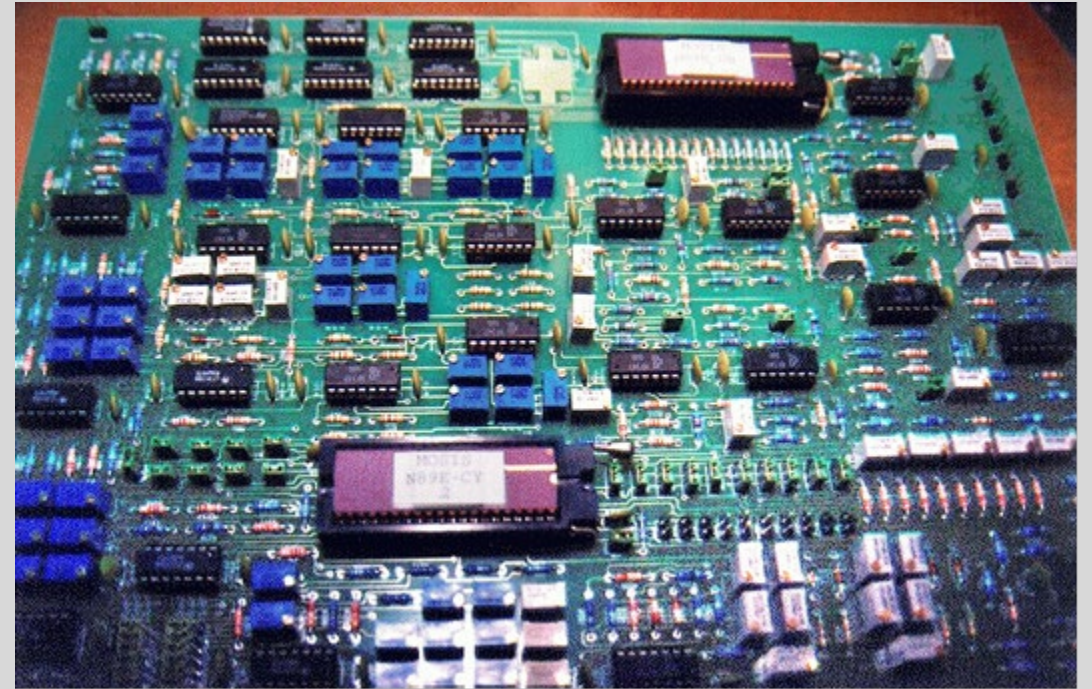
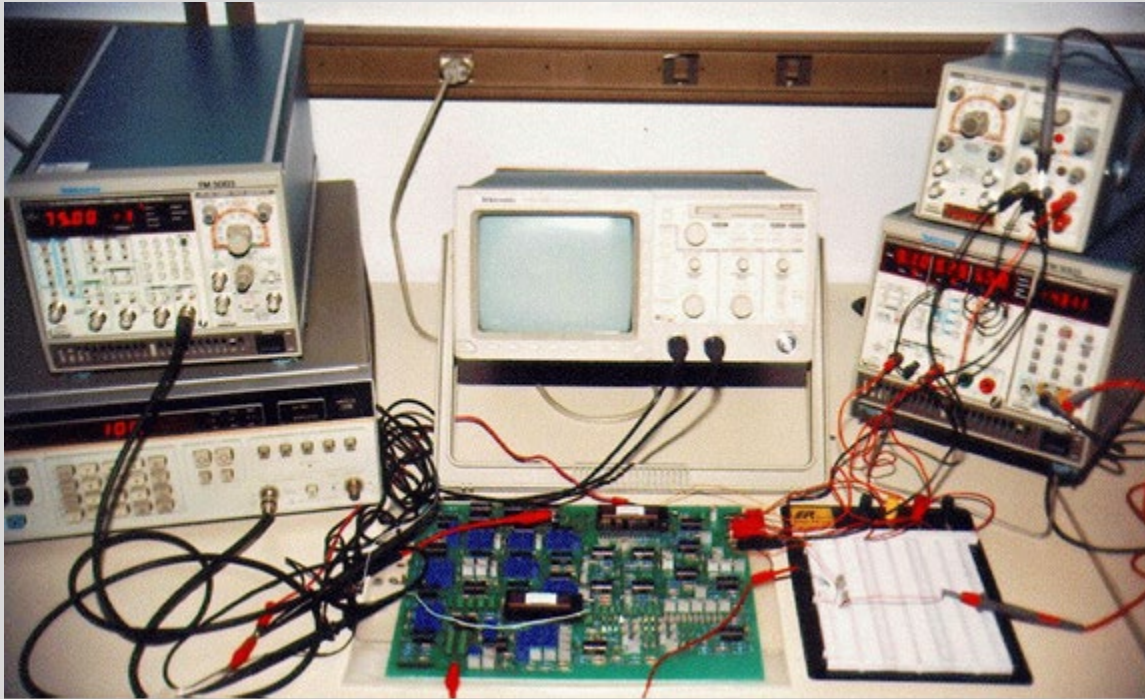


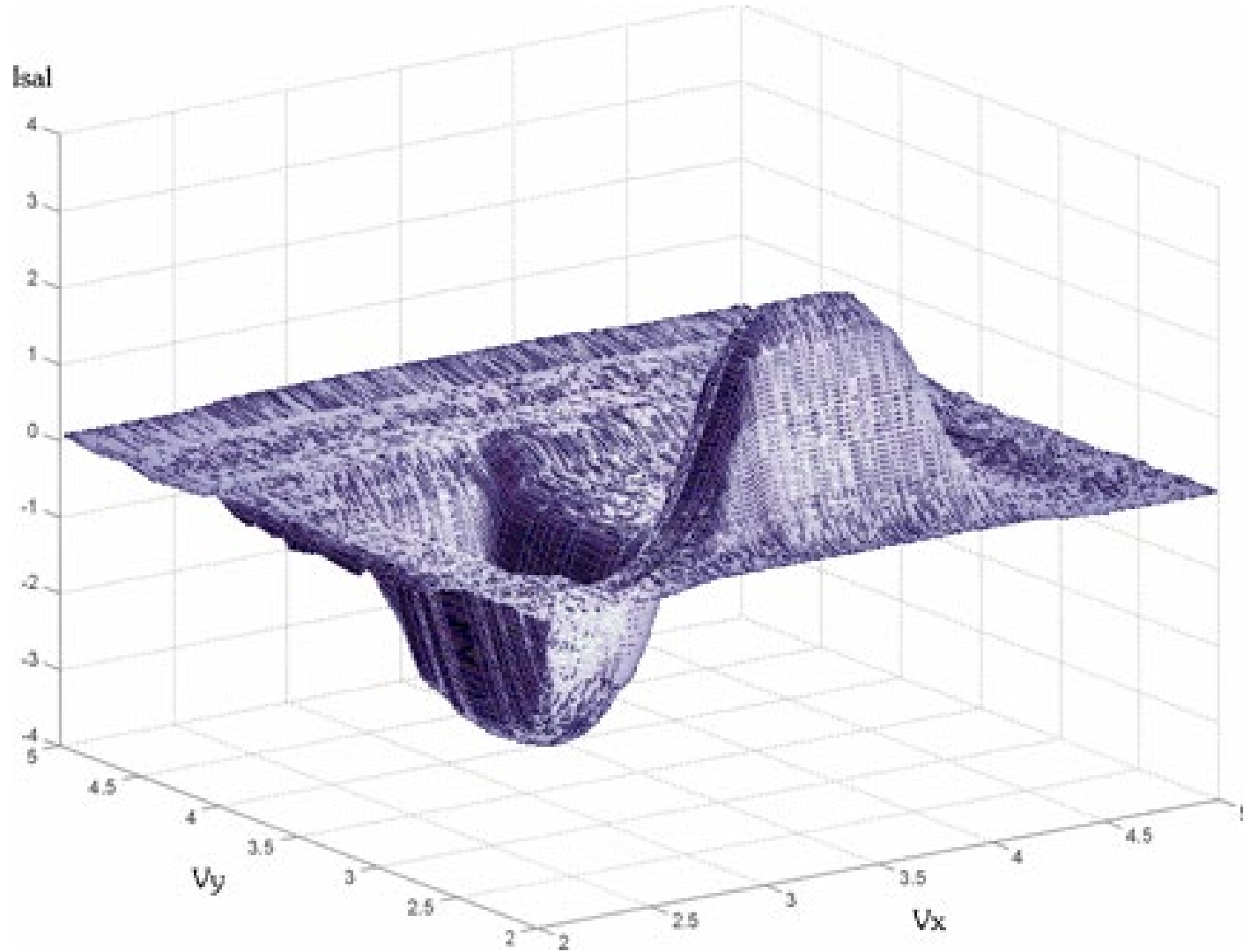
Implementación
en hardware de
un sistema difuso



Microfotografía
del circuito
integrado
fabricado

Sistema difuso implementado en hardware





Superficie
discriminatori
a obtenida con
el sistema
ANFIS, para
el ejemplo de
clasificación
de arritmias
tipo PVC

Resultados obtenidos

- A continuación, se muestran dos tablas en donde se presentan los resultados de la clasificación del sistema ANFIS implementado, así como una evaluación del desempeño del sistema utilizado como clasificador. Cabe aclarar que estas dos últimas tablas se construyeron al realizarse mediciones en forma conjunta con el responsable de la etapa correspondiente a las capas 1 y 2 de ANFIS, por lo que se repiten aquí estos resultados.
-
- Para la evaluación del sistema ANFIS, se plantearon tres escenarios: 1) Aplicando 500 valores de entrada que correspondan con los que se hizo el entrenamiento del sistema en Matlab; 2) Aplicar 500 valores de entrada correspondientes a datos que se utilizaron como comprobación en Matlab, una vez entrenado el sistema; 3) Aplicar otros 500 datos, tomados de los siete pacientes, no considerados en los casos anteriores.
- Las mediciones de los tres *escenarios* se resumen en la tabla, en donde se muestran los resultados obtenidos para cada tipo de latido (normal o PVC). Los resultados de la siguiente tabla, se obtuvieron considerando un umbral positivo de aproximadamente equivalente al 4 % del valor máximo positivo de salida.

Resultados de la clasificación del sistema de desarrollo

Casos de Prueba	Normal ND / NV	PVC VD / VN	No Clasificados NNC / VNC
Escenario 1 (N = 500, V = 250)	485 / 7	241 / 4	8 / 5
Escenario 2 (N = 500, V = 250)	484 / 6	240 / 6	10 / 4
Escenario 3 (N = 500, V = 250)	480 / 8	289 / 6	12 / 5

Desempeño del sistema de desarrollo para clasificación de señales ECG

Casos de Prueba	% ET	% S	% DN	% DV
Escenario 1	3.2	96.4	96.2	93.6
Escenario 2	3.47	96.0	95.6	93.6
Escenario 3	3.88	96.33	94.8	93.67
Evaluación Global	3.52	96.25	95.53	93.63



Instituto Politécnico Nacional
“La Técnica al Servicio de la Patria”



Microtecnología y
Sistemas Embebidos

¡Gracias por todas sus atenciones!

◦ Dr. Herón Molina Lozano

◦ hmolina@ipn.mx