



Clasificación del artículo:
REFLEXIÓN

# RECONOCIMIENTO DE PRIMITIVAS 3D, USANDO AUTOCORRELACIÓN Y ANFIS

**Emil Javier Guerra Monterroza***

## Resumen

El presente trabajo muestra una metodología para el reconocimiento y clasificación de cuatro, primitivas tridimensionales como son: cilindro, cubo, esfera y cono, las cuales están contenidas en imágenes de 128X128 píxeles. La metodología se fundamenta en la aplicación de un algoritmo de autocorrelación, para la extracción de características en las imágenes, y el proceso de clasificación se realiza mediante ANFIS, que es un híbrido entre las redes neuronales artificiales y la lógica difusa. En la fase de operación, el sistema deber ser capaz de diferenciar las primitivas mencionadas, en forma independiente del ángulo de adquisición de la imagen. El sistema descrito, por ejemplo puede ser acoplado a un brazo robótico, para realizar tareas de clasificación de objetos según su forma.

*Palabras clave*
Primitivas 3D, visión artificial, correlación, NeuroFuzzy.

*Summary*
The present work shows a rnethodology for recognition and classification of four (4), primitive three-dimensional such as: cylinder, cube, sphere and cone, which are contained in images of 128X128 pixels. The methodology is based on the application of autocorrelation algorithm, for the extraction of characteristics in images, and the classification process is made by means of ANFIS, that is a hybrid between the artificial neuronal networks and the fuzzy logic. In the phase of operation, the system must be able to differentiate the primitive ones mentioned, in an independent way of the angle of acquisition of the image. The described system, for example, can be connected to a robotic arm stops to perform tasks of classification of objects according to its form.

*Keywords*
Primitive 3D, artificial vision, correlation, NeuroFuzzy

* Ingeniero en Control Electrónico e Ins-trumentación de las Unidades Tecnológicas de Santander, miembro del Grupo de Investigaciones en Control Inteligente, de las mismas Unidades. Correo electrónico: eguerra_inpesa@yahoo.es

## 1. Introducción

Día a día, los sistemas de control basados en visión artificial tienen mayor aplicación en el campo industrial, siendo un reto interesante el reconocimiento de estructuras físicas (en forma independiente su ángulo de adquisición), para la generación de acciones de control. Aquí se describe una metodología que utiliza un algoritmo para autocorrelación para la extracción de características en las imágenes. En la primera fase esas características extraídas se utilizan para construir patrones para entrenar el ANFIS que actúa como clasificador. En la fase de operación las características son entregadas al clasificador y éste responde con la categoría a la que pertenece la primitiva evaluada.

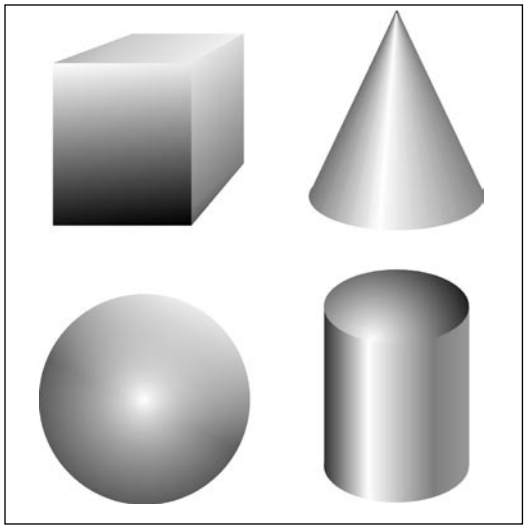
En la figura 1 se encuentran las cuatro primitivas consideradas que fueron escogidas para su identificación y clasificación, porque éstas constituyen la base de estructuras más complejas.

Las primitivas se encuentran contenidas en imágenes de 128x128 píxeles, que pueden proceder de una cámara o ser generadas en un PC.

Las imágenes contienen una gran cantidad de información, situación que imposibilita su utilización para entrenar directamente el ANFIS, por ello deben ser tratadas previamente mediante el algoritmo de autocorrelación que, en esencia, transforma cada imagen en una pareja de datos, con el propósito de reducir o eliminar la información redundante que puedan contener.

Terminado el paso anterior el híbrido es configurado y entrenado para reconocer las primitivas citadas. Finalmente, una vez la termina el entrenamiento, el sistema debe ser evaluado con datos no mostrados en la fase de entrenamiento, que pueden ser imágenes contaminadas con ruido, para valorar su capacidad de respuesta ante este tipo de información.

**Figura 1. Primitivas 3D**



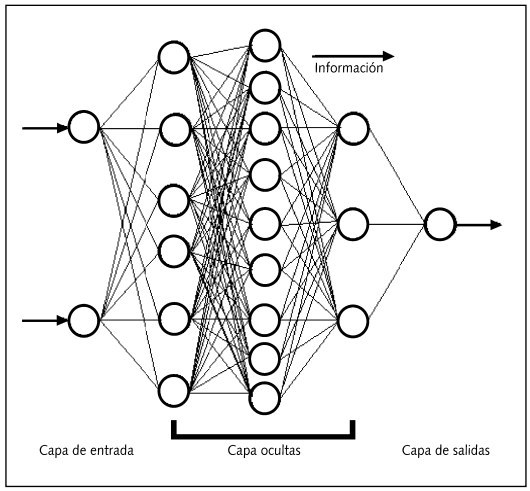
## 2. Redes neuronales artificiales, lógica difusa y ANFIS

Las Redes Neuronales Artificiales (RNA) son un conjunto de elementos simples, fuertemente interconectados y de procesamiento paralelo. Su mecanismo para procesar la información se encuentra inspirado en la naturaleza cerebral. Las RNA hacen parte de la inteligencia artificial, donde ocupan un lugar privilegiado, gracias a su amplia aplicación en múltiples áreas.

Las principales características de las RNA son:

1. Capacidad de aprendizaje, es decir, pueden aprender a realizar tareas a partir de datos representativos del problema.

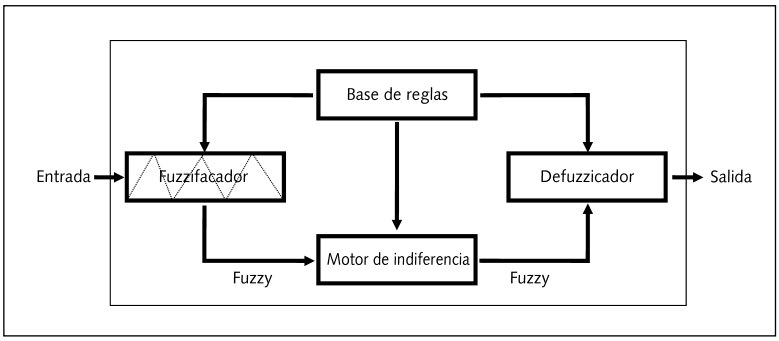
**Figura 2. Red neuronal artificial**




2. Son capaces de organizar la información adquirida durante el aprendizaje, entre sus elementos de proceso.
3. Pueden responder apropiadamente ante datos nunca antes vistos (generalización).
4. Además brindan resultados satisfactorios cuando se les suministra información ruidosa, e incluso cuando la red se encuentra parcialmente destruida.

La red BackPropagation (BP) es una de las una de las más utilizadas por su capacidad de aprendizaje y generalización, características necesarias para determinar las relaciones que existen entre un conjunto de datos de entrada y salida y así poder realizar una buena aproximación. El aprendizaje consiste en la presentación iterada de los datos y la modificación sistemática de los pesos de la red, según el algoritmo de retropropagación del error BP. Los datos utilizados para el entrenamiento deben ser previamente escalados a valores comprendidos entre: 0 y 1 o –1 y 1.

**Figura 3. Sistema de inferencia difuso**



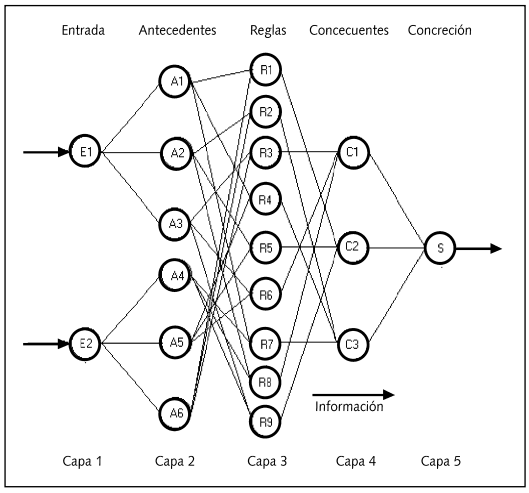

La Lógica Difusa (LD), es la rama de las matemáticas que se encarga del procesamiento de información ambigua, que puede expresarse en forma de reglas lingüísticas, como por ejemplo: **si** temperatura **es** agradable **entonces** velocidad_ventilador **es** baja.

La LD aparece como respuesta a las dificultades presentadas por la lógica tradicional binaria, en la que no existen estados intermedios. Mediante LD es posible establecer el grado de pertenencia de un elemento a uno o varios de los conjuntos con que se encuentra relacionado. El grado de pertenencia puede tomar valores dentro del rango continuo [0,1], incluyendo los extremos y es determinado mediante las Funciones de Pertenencia (FP).

Los sistemas difusos se caracterizan por:

1. Fácil incorporación del conocimiento, ello se debe a que estos sistemas se basan en reglas lingüísticas y no en constructores matemáticos, por lo tanto los modelos difusos pueden ser modificados simplemente creando, eliminando o modificando reglas lingüísticas.
2. Explicación de resultados, mediante un análisis de las reglas lingüísticas que participan en la generación de un resultado; es posible establecer cómo se llegó a éste.

**Figura 4. ANFIS para dos entradas**




Existen diferentes formas de combinar las redes neuronales con la LD, pero una de las más exitosas en el modelado de sistemas continuos no-lineales es el Adaptive Neuro Fuzzy Inference System (ANFIS) que conserva la estructura de una red neuronal, siendo sus conexiones de tipo difuso. El ANFIS es multientrada con única salida, está constituido por cinco capas de neuronas. El aprendizaje se realiza mediante BackPropagation o BackPropagation modificado.

Las características más importantes del ANFIS son:

1. Mayor capacidad de aprendizaje y en menos ciclos de entrenamiento, comparado con las redes neuronales.
2. Capacidad de extraer y refinar reglas de control, a partir de los datos de entrenamiento.
3. Capacidad de refinar de sus funciones de pertenencia, se hace mediante el ajuste de las curvas de las funciones de pertenencia a los datos de entrenamiento.
4. Fácil interpretación del conocimiento adquirido durante el entrenamiento.
5. Al igual que en LD es posible determinar cómo se llegó a un resultado.

## 3. Algoritmo de correlación

El algoritmo de correlación típico (ecuación 1) se utiliza para determinar el grado de similitud entre dos series de datos; éste brinda un resultado que varia entre -1 y 1, representado -1 una similitud inversa, el cero ninguna similitud y el 1, una completa similitud.

$$R = \frac{\sum_{1}^{N}\left[(X(i)-X_m)(Y(i-d)-Y_m)\right]}{\sqrt{\sum_{1}^{n}\left[(X(i)-X_m)\right]^2}\sqrt{\sum_{1}^{n}\left[(Y(i-d)-Y_m)\right]^2}} \quad (1)$$

En la ecuación anterior Xm y Ym, corresponden con la media de cada serie y d, corresponde con la ventana de tiempo y tiene la desventaja que la correlación no es absoluta. Para superar la dificultad anterior se plantea una modificación de la ecuación anterior en la ecuación 2.

$$R = \frac{\sum_{1}^{N}\left[(X(i)-X_m)(Y(i-d)-Y_m)\right]}{\sqrt{\sum_{1}^{N}\left[(X(i)-X_m)\right]^2}\sqrt{\sum_{1}^{N}\left[(Y(i-d)-Y_m)\right]^2}}\frac{\sum_{1}^{N}|Y(i)|}{\sum_{1}^{N}|X(i)|} \quad (2)$$

## 4. Metodología

A continuación se describen los pasos necesarios para construir el sistema de reconocimiento de primitivas 3D.

1. Se debe disponer de un conjunto de patrones representativos de la tarea que, en este caso, corresponde a unas 600, imágenes de 128x128 píxeles, por cada primitiva. Los patrones son para entrenar y evaluar el ANFIS, y estos deben contener suficiente información acerca de variaciones en el tamaño, ubicación

y perspectiva en cada una de las cuatro primitivas. Los patrones fueron divididos en dos grupos con el 90 y 10%, para entrenamiento y prueba respectivamente.

2. Inicialmente, a los patrones de entrenamiento se les debe pasar a escala de grises, luego eliminar la información redundante mediante la ecuación 2. El resultado de este paso es una reducción drástica en las dimensiones de los patrones de entrenamiento, sin perder la información que describe la tarea.
3. Una vez procesados los patrones, se configura el ANFIS para su entrenamiento, que para las condiciones actuales tiene dos neurona de entradas y una de salida.
4. El ANFIS entrenado debe ser evaluado con patrones no presentados en la fase de entrenamiento, que pueden ser: imágenes en perspectivas diferentes a las utilizadas en el entrenamiento o ruidosas. El objetivo de este paso es probar que la red responde en forma adecuada ante este tipo de información.

## 5. Resultados

La figura 5, muestra un cubo en escala de grises. Las gráficas 1 y 2 muestran la autocorrelación para el cubo en la dirección X e Y, respectivamente.

Los resultados del procesamiento para las otras tres primitivas fueron semejantes a los presentados para el cubo.

La gráfica 3 muestra el resultado de la clasificación mediante el ANFIS. En ésta los datos en azul representan la salida ideal, los datos en verde son los de entrenamiento y los datos en rojo corresponden a los patrones de prueba que nunca fueron mostrados en la fase de entrenamiento.

El sistema tiene un mal desempeño con datos ruidosos, seguramente porque se redujo mucho las dimensiones de los patrones y con ello se pierde información valiosa, que permitiría responder adecuadamente.

**Figura 4. ANFIS para dos entradas**

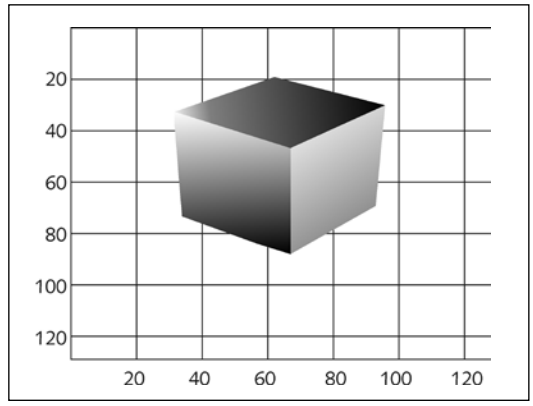



**Figura 4. ANFIS para dos entradas**

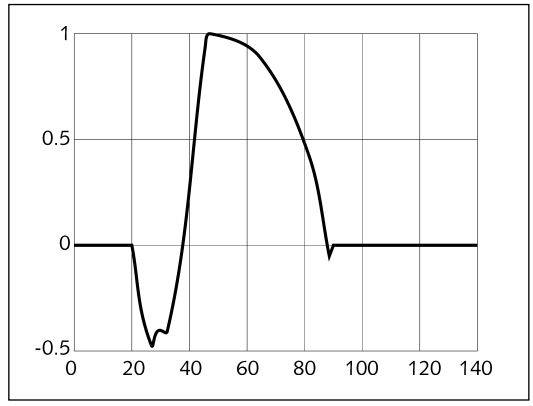



**Figura 4. ANFIS para dos entradas**

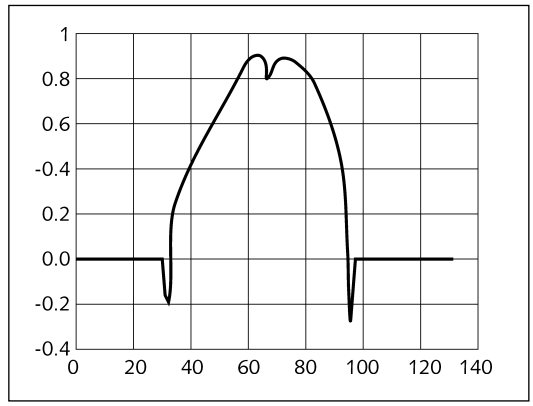
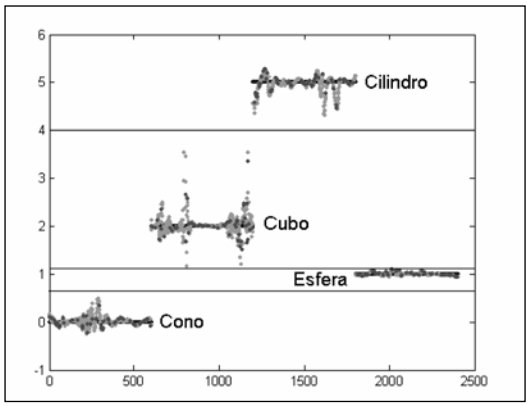


**Figura 4. ANFIS para dos entradas**

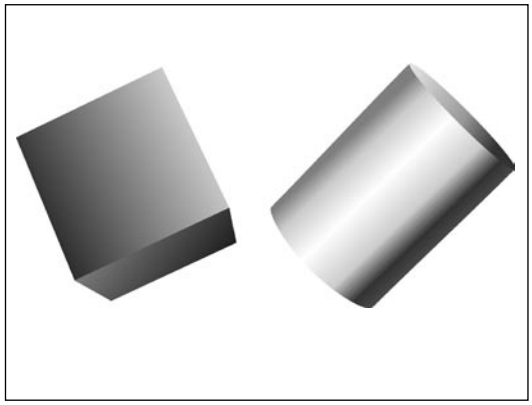



El sistema fue probado con datos contaminados con ruido, pero su desempeño fue pobre, porque sólo responde adecuadamente hasta un nivel de ruido del 5%.

## 6. Conclusiones

1. El sistema es capaz de reconocer imágenes en perspectivas difíciles, con un alto grado de acierto.
2. El ANFIS demostró ser útil también en tareas de clasificación, ya que tradicionalmente se emplea como aproximador universal.
3. El sistema tiene un mal desempeño con datos ruidosos, seguramente porque se redujo mucho las dimensiones de los patrones y con ello se pierde información valiosa, que permitiría responder adecuadamente.
4. Las dificultades presentadas respecto al ruido pueden ser evitadas mediante una buena iluminación y sistema de adquisición.
5. Se recomienda emplear una técnica para reducir la dimensionalidad del problema que incluya una mayor cantidad de información de las primitivas.

**Figura 4. ANFIS para dos entradas**



## Referencias bibliográficas

[1] Torres, Emmanuel y Guerra Emil. (Noviembre de 2002). *Reconocimiento de objetos 3D proyectados en un plano.* VII Simposio de tratamiento de señales imágenes y visión artificial. Bucaramanga, Colombia.

[2] González, Rafael y Woods. Richard. (1996). *Tratamiento digital de imágenes.* S.d.: Addison Wesley.

[3] Bruno, N. (1999). *Sistemas de inferencia difusos basados en redes neuroadaptativas.* Murcia: Universidad de Murcia.

[4] J. S. Roger. (1993). *ANFIS: Adaptive-Network-Based Fuzzy Inference System.* Berkeley: Universidad de California.

[5] Freeman, J. A. Skapura D. M. (1993). *Redes neuronales algoritmos, aplicaciones y técnicas de programación*. S.d.: Addison Wesley.