

Reconocimiento visual de múltiples objetos sobre una superficie con redes neuronales y su clasificación mediante un brazo robot

El objetivo del trabajo es la identificación de un número indeterminado de diferentes objetos sobre una superficie dispuestos en cualquier orden.

Tras la identificación de esos objetos, un brazo robot es el encargado de su clasificación.

Jesús Manuel Besada Juez

Nació en Madrid en 1976. Ingeniero Industrial del ICAI, Universidad Politécnica de Comillas. 1999. Investigador adscrito al Instituto de Investigación Tecnológica y alumno de doctorado de ICAI. Sus áreas de investigación son los sistemas expertos, redes neuronales, lógica difusa, sistemas multiagentes, visión artificial y robótica.



Miguel Ángel Sanz Bobi

Trabaja en IIT desde 1987. Profesor en la E.T.S. de Ingenieros Industriales de ICAI, Universidad Politécnica de Comillas. Doctor en 1997. Ha dirigido tesis doctorales. Ha desarrollado aplicaciones industriales en el campo de la fabricación y mantenimiento y se centra en expertos y aplicaciones de sistemas de Inteligencia Artificial para la supervisión y diagnóstico de procesos industriales complejos tales como centrales de generación eléctrica y máquinas herramienta.



Para lograr alcanzar este objetivo, se ha de capturar la imagen de la superficie sobre la que deben de encontrarse los objetos. Se procederá al procesamiento y segmentación de la misma con objeto de determinar los bordes de los objetos presentes. Previamente una red neuronal basada en una arquitectura perceptrón multicapa ha sido entrenada en la identificación de todos los objetos posibles. El entrenamiento se ha realizado utilizando múltiples imágenes de los objetos en diferentes posiciones. Cada objeto ha sido caracterizado por sus descriptores cartesianos de Fourier aplicados a sus bordes, éstos han sido las entradas de la red neuronal. Con la red entrenada y detectados los bordes de los objetos de la superficie, éstos son identificados y de acuerdo a sus características geométricas, una pinza plana de brazo robot es capaz de coger el objeto por la parte más conveniente y moverlo para su clasificación al sitio adecuado. Los resultados obtenidos han sido satisfactorios tanto en la fiabilidad de la identificación como en su clasificación.

Introducción

El reconocimiento y clasificación de objetos en procesos industriales es una tarea que realizada por una persona consume tiempo, es monótona, la cadencia de clasificación puede depender del estado de ánimo, puede causar problemas de salud, etc. Todo esto hace que desde un

tiempo acá, estas tareas industriales tiendan a dejarse en manos de máquinas o procesos automatizados. Sin embargo la capacidad humana de reacción ante cualquier eventualidad dentro del proceso de identificación y clasificación y su correspondiente corrección, es difícil de automatizar. Es por ello que la Inteligencia Artificial (Russell, 1996) en su afán de emular las habilidades humanas, es un buen campo donde desarrollar aplicaciones industriales del tipo descrito.

Dentro del mundo de la Inteligencia Artificial son múltiples las referencias que han abordado el problema del reconocimiento de objetos y su clasificación en un proceso industrial. Este artículo pretende sumarse a ellas presentando otra nueva versión simple, fiable y eficiente. El procedimiento utilizado es independiente del tipo de objeto a clasificar y de su forma. Su contenido será descrito en los diferentes apartados de este artículo.

Objetivos y técnicas empleadas

El objetivo del trabajo es la identificación de un número indeterminado de diferentes objetos sobre una superficie dispuestos en cualquier orden. Tras la identificación de esos objetos, un brazo robot es el encargado de su clasificación.

Para lograr este objetivo se han empleado diferentes técnicas que se pueden catalogar en tres áreas diferentes; determinación

de las características de los objetos, identificación de objetos mediante redes neuronales y definición de la captura de objetos y trayectoria a seguir por un brazo robot. Seguidamente se describirán brevemente los métodos utilizados en estas tres áreas.

Determinación de las características de los objetos:

- Captura de la imagen de la superficie con las piezas u objetos a clasificar.
- Tratamiento preliminar en que se realiza el preprocesado y segmentación de la imagen.
- Cálculo de contornos. A partir de la imagen en binario y segmentada se procede a la obtención de los bordes de los objetos para su posterior análisis.
- Caracterización del objeto. Se ha de encontrar una manera para que de forma muy concisa, representativa y unívoca se pueda identificar un objeto. Aquí se ha optado por elegir los descriptores cartesianos de Fourier (Jähne, 1991).

Identificación de objetos mediante redes neuronales:

- El tipo de arquitectura de red neuronal elegida es un perceptrón multicapa (Muller. B., Reinhardt. J., Strickland. M.T. 1995, Haykin. Simon 1994, Hilera. José R., Martínez. Víctor J. 1995).
- Se eligió un conjunto de entrenamiento compuesto por imágenes de cada objeto a identificar. Estas imágenes contemplaban múltiples posiciones del objeto.

• Las entradas de la red eran los descriptores cartesianos de Fourier mencionados antes.

• Se procedió al entrenamiento de la red neuronal.

• Se realizó un intensivo análisis de validación de las imágenes aprendidas por la red.

Definición de la captura de objetos y trayectoria a seguir por un brazo robot:

• Comunicación con el robot para enviar y recibir los datos necesarios.

• Orientación espacial. Una vez identificado el objeto se debe proceder a su captura. Para ello se ha de conocer su centro de masa con los descriptores cartesianos de Fourier y su orientación con los momentos centrales (Jähne, 1991).

• Trayectoria del brazo robot. Una vez conocida la posición y orientación del objeto el robot lo cogerá y dejará en un lugar según el tipo de objeto.

Procedimiento de caracterización de objetos para su identificación y clasificación

El objetivo de este procedimiento es obtener una definición clara y unívoca del objeto que se está caracterizando mediante unos parámetros. Estos parámetros van a basarse en los

descriptores cartesianos de Fourier, tal y como se ha mencionado en el apartado 2 de este artículo. Para llegar a obtener estos parámetros para un objeto particular partiendo de la imagen de una superficie donde puede haber varios objetos, se ha de seguir un proceso que se describe en la figura 1.

El proceso empieza por la captura de la imagen a tratar que se almacena en la memoria de la tarjeta digitalizadora en formato 512*512 pixels. La imagen capturada es de 256 tonalidades de grises. Seguidamente se realiza un

tratamiento previo o preliminar que consiste en reducir la resolución de la imagen original a 128*128 pixels sin una pérdida sustancial de calidad de imagen y ganando velocidad de procesado. Este tratamiento previo incluye el cálculo del umbral óptimo de detección que permite separar el/los objeto/s del fondo de la imagen, es decir, permite la segmentación de la imagen en sus diferentes objetos. Este es un parámetro crítico que condicionará los resultados posteriores siendo muy sensible a las variaciones de iluminación, por ello se ha optado por ofrecer la posibilidad de cam-

biar el valor óptimo estimado si se cree conveniente.

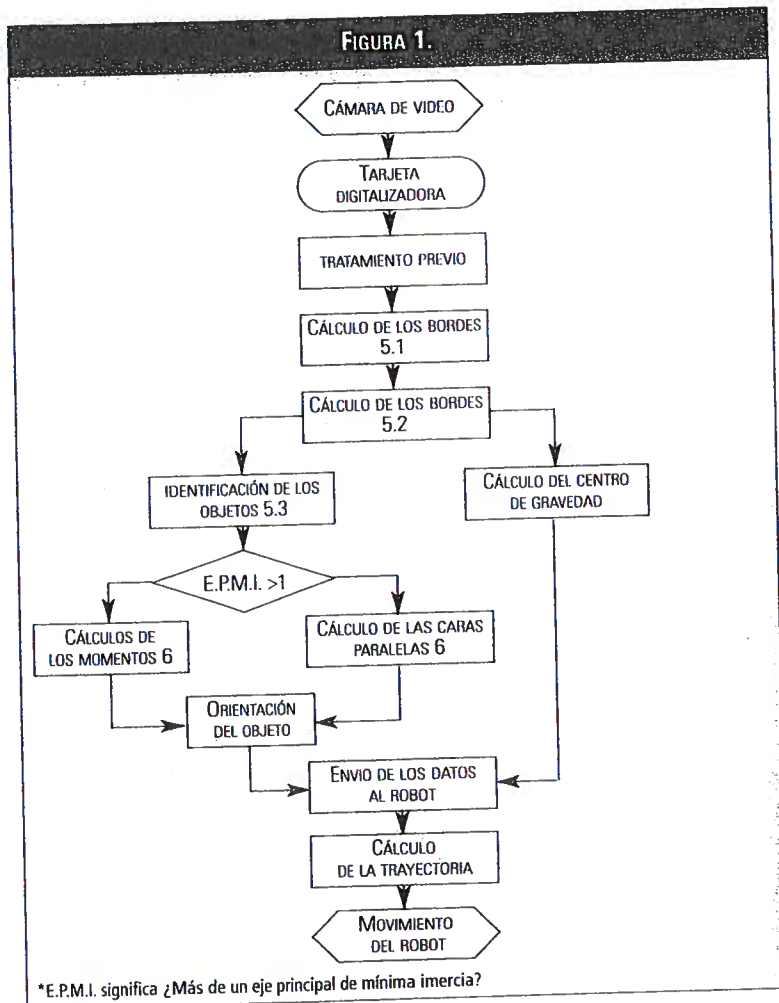
Con la imagen segmentada se está en disposición de calcular los bordes de los objetos de la imagen. Los bordes permitirán calcular los descriptores cartesianos de Fourier para caracterizar a los objetos y éstos a su vez suministran el centro de gravedad del objeto.

Una vez obtenidos estos parámetros se pretende utilizarlos como estímulos en una red neuronal para identificar cada objeto. Identificado el objeto, éste ha de ser cogido por el brazo robot. Se usarán los momentos centrales de segundo orden en objetos poco regulares (sin caras paralelas y con un solo eje principal de mínima inercia) y un algoritmo desarrollado para objetos con caras paralelas o con más de un eje principal de mínima inercia. Con esta información el robot puede proceder a la captura del objeto conociendo quien es, su posición, su orientación, y su verticalidad u horizontalidad.

En las siguientes secciones del artículo se comentarán los medios utilizados así como el fundamento del algoritmo de reconocimiento de bordes, identificación de los objetos y manipulación del brazo robot.

Instrumentación y medios utilizados

Se listan a continuación los medios utilizados en la implantación del proceso descrito en la figura 1.



- Ordenador compatible de bajas prestaciones encargado de realizar todos los cálculos y controlar a los restantes dispositivos empleados.

- El lenguaje de desarrollo utilizado ha sido C.

- Cámara de vídeo blanco y negro.

- Monitor auxiliar: Muestra la imagen capturada y los sucesivos tratamientos a la imagen.

- Tarjeta digitalizadora inserta en el PC y encargada de procesar la imagen, almacenar y realizar diferentes funciones a través de librerías en C.

- Robot industrial Staubli con lenguaje de programación V+.

- Objetos a identificar.

- Mesa: Es la superficie plana donde se colocan los distintos objetos. El color de la superficie de la mesa debe tener suficiente contraste con respecto a los ob-

jetos para poder distinguirlo. En la mesa existen también unos ejes que servirán de referencia espacial.

Principios sobre los que se basa la identificación de objetos

Se describe en sucesivos apartados aquellas partes del proceso de identificación más relevante. Éstas son la detección de los bordes de los objetos, la caracterización de los objetos mediante descriptores Cartesianos de Fourier y la identificación propiamente dicha por una red neuronal.

Detección de bordes

A partir de una matriz de puntos (pixels) obtenida de la imagen a estudiar, el algoritmo de detección de bordes utilizado

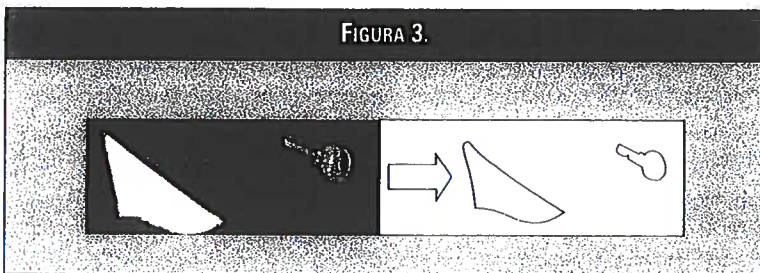
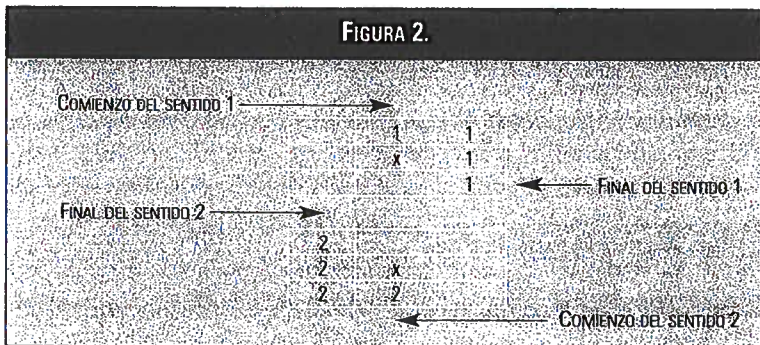
parte del punto (0,0) que es la esquina superior izquierda de la imagen. A partir de él se recorre la matriz (la imagen) barriendo por filas, así en la fila i se empezará por la columna 0 hasta la n y luego se pasará a la fila $i+1$. El objetivo de este recorrido es encontrar puntos con intensidad luminosa por encima de un umbral calculado, capaz de reconocer qué es fondo y qué es objeto. El primer punto que se detecte como perteneciente a un objeto representa siempre el punto más cercano a la referencia (0,0) origen. A partir de este punto se analizan los nueve puntos vecinos. La continuidad del borde sugiere que siempre se habrá de encontrar un punto vecino salvo que el primer punto sea único.

Se definen dos sentidos a la hora de mirar los puntos vecinos, de tal manera que siempre se empieza con el sentido 1 y salvo que no encuentre un punto vecino pasa al sentido 2 donde permanecerá hasta que no encuentre punto vecino pasando otra vez al uno. Los citados sentidos se muestran en la figura 2.

Siendo X el punto considerado como contorno y los nueve puntos vecinos como posibles nuevos puntos del contorno. Una vez definidas las reglas de búsqueda de puntos vecinos y con objeto de evitar profundizar en el objeto, se introduce una nueva condición que consiste en comprobar si todos los puntos vecinos son 1. Si lo son y según el sentido actual se ha de cambiar de sentido y las reglas que se aplican son:



En esta fotografía se muestra la mayor parte de los medios enumerados anteriormente



- Si sentido=1 \diamond El nuevo punto es misma fila y $P_{columna-1}$ y sentido=2
- Si sentido=2 \diamond El nuevo punto es misma fila y $P_{columna+1}$ y sentido=1

De esta forma se asegura el permanecer en la superficie del objeto y obtener el contorno correcto.

Además de lo anterior una variable de alarma (Watchdog) avisa si el algoritmo ha dado más de un número determinado de pasos suficientemente grande como para obtener cualquier contorno. El algoritmo también desprecia todos los supuestos objetos que no superen un mínimo número de puntos.

En la figura 3 se muestra el resultado de segmentar una imagen en sus objetos.

Caracterización de los objetos mediante descriptores cartesianos de Fourier

Para reconocer un objeto es necesario caracterizarlo de la mejor manera posible de forma unívoca con independencia de la posición del mismo. La manera utilizada para su caracterización ha sido mediante el empleo de unos parámetros que actúan de "descriptores" del mismo. Las propiedades que se han exigido a estos descriptores son las que se enumeran a continuación (FU. K. S., González. R.C., Lee. C.S. G., 1988):

- Inmunidad frente al ruido. Los descriptores no deben verse afectados por pequeñas perturbaciones de los contornos de los objetos.
- Invarianza al escalado. Se ha de cumplir que un mismo ob-

jeto de distinto tamaño no dé lugar a descriptores diferentes.

- Invarianza a traslaciones. La posición del objeto no debería influir en su identificación.
- Invarianza a rotaciones.
- Cálculo rápido de los descriptores.

De entre todos los posibles descriptores analizados, se ha decidido utilizar los descriptores Cartesianos de Fourier ya que sólo necesitan el contorno del objeto con el consiguiente ahorro de tiempo de cálculo, y poseen buenas propiedades de invarianza.

El principio de cálculo de los descriptores Cartesianos de Fourier se basa en considerar el contorno descrito por dos vectores circulares de tamaño M "x" e "y" de tal manera que pueden componerse en un solo vector complejo tal que " $e = x + y$ " sea una descripción completa del contorno. Los coeficientes de la transformada discreta de Fourier sobre este vector complejo son:

$$\hat{i}_n = \frac{1}{M} \sum_{m=0}^{M-1} i_m \exp\left(-\frac{2\pi i m n}{M}\right)$$

El primer coeficiente de la transformada de Fourier es el centro de masas del objeto como se observa simplificando la fórmula:

$$\hat{i}_0 = \frac{1}{M} \sum_{m=0}^{M-1} x_m + i \frac{1}{M} \sum_{m=0}^{M-1} y_m$$

El cuál no es útil como descriptor pero servirá para localizar la pieza.

Todos los descriptores de Fourier son invariantes a las translaciones y son fácilmente normalizables para hacerlos invariantes a escalado y rotaciones. En la figura 4 se observa la representación de los descriptores cartesianos de Fourier de varios objetos.

• Parámetros de un paralelepípedo y una pirámide sin normalizar:

En el paralelepípedo (figura 4) se observa que todos sus parámetros impares son nulos debido a la simetría de la figura. A partir del parámetro octavo los valores de los descriptores son excesivamente pequeños y por tanto son poco significativos.

En la pirámide (figura 5) debido a que no es simétrica respecto a sus dos ejes principales los parámetros impares son tan importantes como los pares.

De igual manera si viéramos los parámetros de una esfera todos sus parámetros excepto el primero serían despreciables debido a que tiene infinitos ejes de simetría.

Se considera que los descriptores usados permitirán identificar los objetos de manera unívoca.

Identificación de objetos con una R.N.A.

La identificación de los objetos presentes en la superficie que se está viendo se realiza tras la obtención de los descriptores Cartesianos de Fourier de los objetos y el paso de los mismos a través de una red neuronal.

El conjunto de entrenamiento utilizado para la identificación de los objetos posibles está constituido de 32 ensayos de cada objeto en posición ho-

rizontal con poco más de 1° entre rotación a la par que se realizan traslaciones del objeto para intentar abarcar las posiciones más significativas y también de 10 rotaciones para el mismo objeto en vertical.

La arquitectura de red neuronal elegida fue un perceptrón multicapa con función de activación tangente hiperbólica. El número de entradas es 5, correspondientes a los cuatro primeros pares y al primero impar, debido a la simetría de la mayoría de las piezas utilizadas. El número de salidas es 5, cuatro de ellas corresponden a un código de objeto reconocido y la restante indica si el objeto está en vertical u horizontal.

A través de un análisis pormenorizado de errores cometidos se ha obtenido que el mejor número de capas ocultas es dos con 12 y 8 neuronas respectivamente.

FIGURA 4 Y 5.

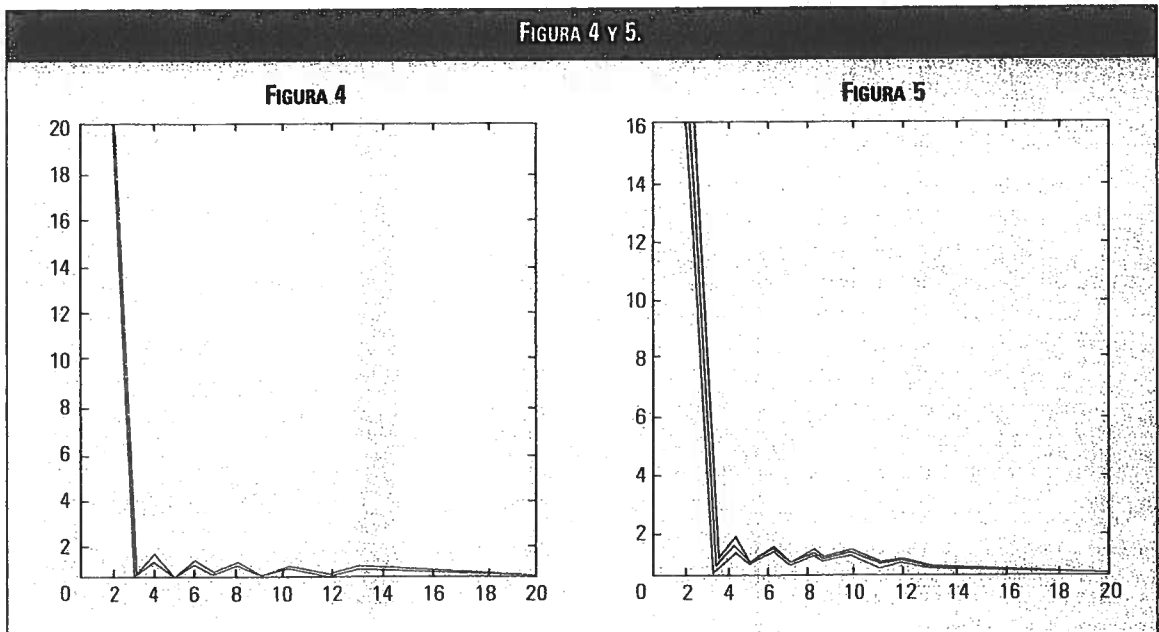
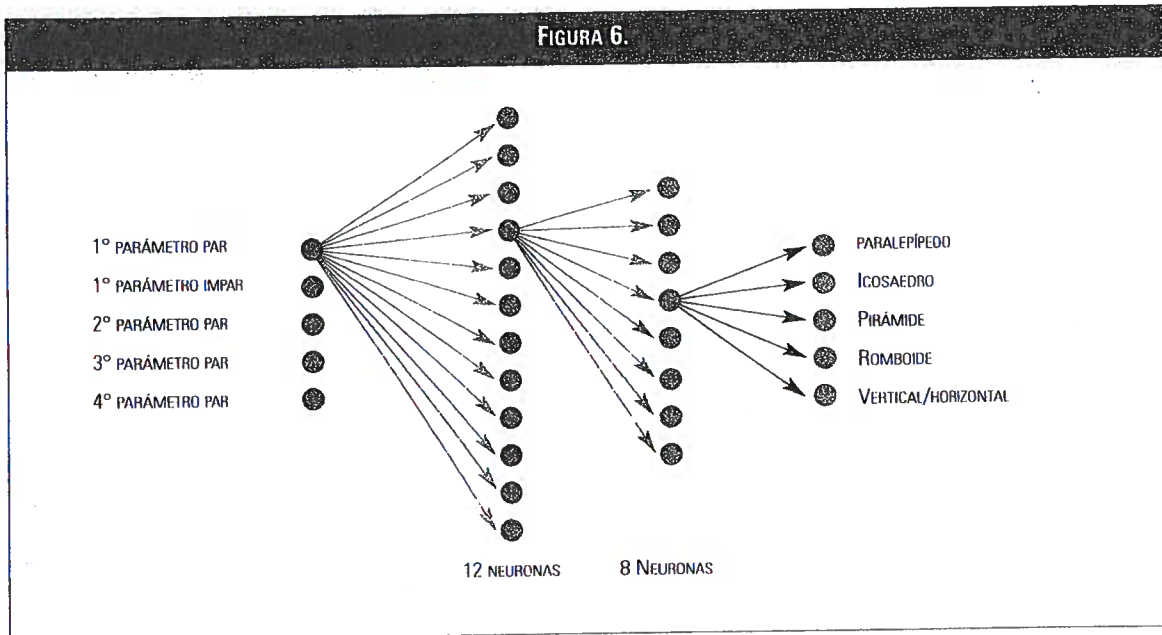


FIGURA 6.



a) La arquitectura de la red usada con función de activación tangente hiperbólica ha sido la presentada en la figura 6:

b) Validación de la red:

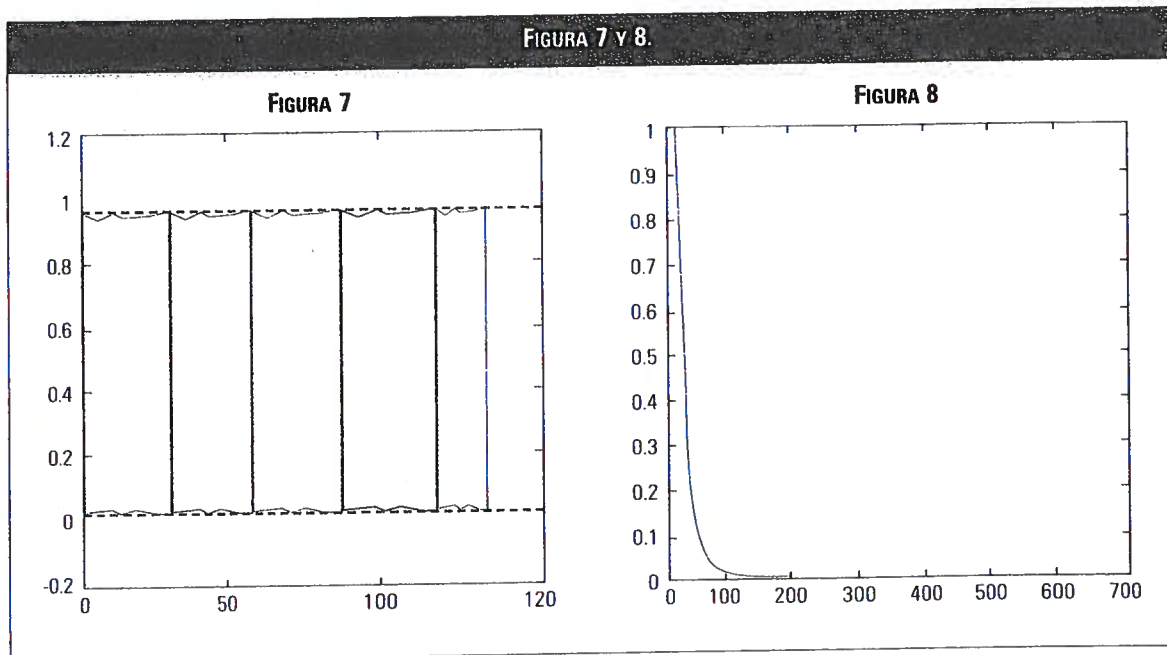
El error cometido en el entre-

namiento con la estructura óptima y usando un método cuasi-Newton fue de $1e-5$ con 707 pasos, el error en cada iteración está representado en la figura 8.

La respuesta de la red al conjunto de entrenamiento forma

como era de esperar conjuntos disjuntos representando los cuatro primeros conjuntos cada una de las piezas y los dos últimos la verticalidad u horizontalidad (figura 7), obteniéndose igual respuesta ante un conjunto de test previamente seleccionado.

FIGURA 7 Y 8.

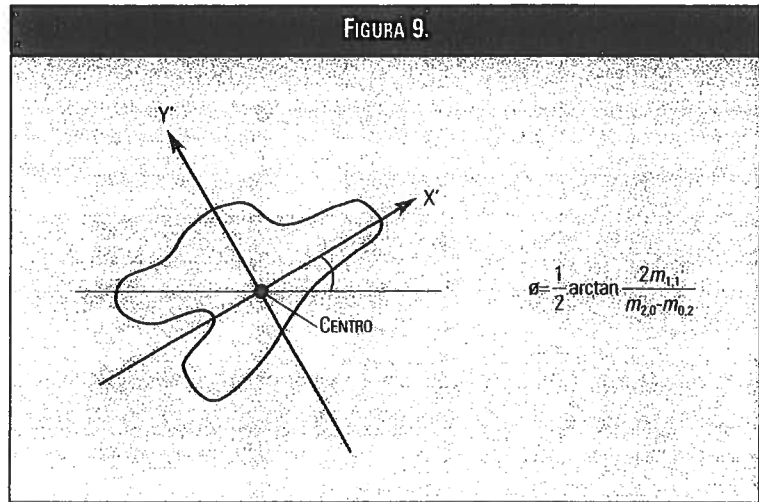


Recogida del objeto y su clasificación por un brazo robot

Con la respuesta de la red neuronal se está en disposición de proceder a la captura del objeto identificado. La orientación del objeto se obtiene o por un algoritmo de caras paralelas o del cálculo de los momentos de segundo orden del objeto, según la regularidad del objeto identificado. El fundamento de ambos algoritmos se describe a continuación.

a) Orientación por los momentos centrales (Jähne, 1991):

De los momentos centrales de orden 2 se obtiene el ángulo que forma el eje principal de mínima inercia con respecto al eje X según está definido en la figura 9.

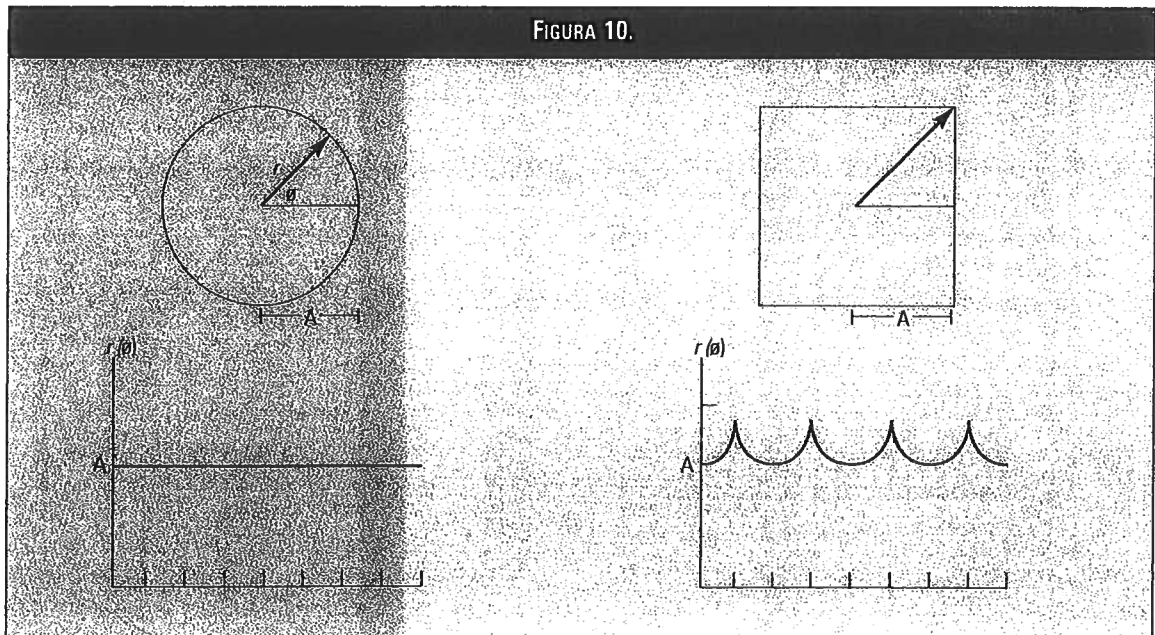


b) Estudio de las caras paralelas por el análisis de la signatura del contorno (FU. K. S., González. R.C., Lee. C.S. G., 1988):

Conocida la respuesta de la red y por tanto la figura a estudiar, la detección de caras paralelas resulta trivial con un análisis de los máximos y mínimos de la signatura del contor-

no, ya que, por ejemplo, en el cuadrado de la figura 10 los máximos representan los vértices y los mínimos el punto medio de una cara, calculando la posición de dos mínimos consecutivos impares obtenemos una cara paralela.

Con la orientación calculada por alguno de los dos métodos previamente comentados se en-



vían estos datos al robot, el cual está en un bucle cerrado a la espera de nueva información. El primer parámetro que espera el robot es el número de objetos en la zona de trabajo seguidamente de los datos de cada pieza. Con todos los datos enviados al robot, éste procede a la clasificación de los objetos o a cualquier proceso que se quiera someter a cada pieza.

En la zona de trabajo hay unos ejes ortogonales de referencia que sirven como referencia relativa para los objetos realizando una transformada de los ejes absolutos del robot a los relativos, éstos además nos permitirán independizar la distancia de la cámara a los objetos ya que conociendo la distancia entre los centros de gravedad de los ejes en la dirección de uno de los ejes y calculando el centro de gravedad de los ejes por Fourier obtenemos el factor Pixel/milímetros.

Para cada tipo de objeto se definirá una posición que hará

de contenedor al cual una vez capturada la pieza se llevará el objeto a dicha posición con trayectorias rectas y aproximaciones suaves al nivel del suelo a velocidades del 20% por ser este momento el más peligroso para la seguridad del robot, como se observa en la figura 2, el robot agrupa las piezas fuera de la zona de visión (fondo negro) y dentro de su campo de acción (la mesa que representa la zona de trabajo).

Conclusiones

El sistema descrito en el artículo identifica y clasifica sin error seis tipos de objetos en una superficie en apenas cinco segundos. El sistema resulta flexible y puede ampliarse a otro número y tipo de objetos muy fácilmente. El uso de los descriptores cartesianos de Fourier para caracterizar los objetos se ha visto que es un método eficaz y ello unido al uso de una

red neuronal lo hace fiable y robusto.

La precisión conseguida en la localización del objeto ha sido de ± 3 mm en la posición y casi 1° en la orientación. El uso de otra cámara de vídeo para simular la visión 3D permitiría obtener una herramienta tan precisa como podría serlo una persona y mucho más rápido que ésta a la hora de clasificar múltiples figuras desordenadas sobre una superficie, obteniendo una herramienta muy útil para la industria en algo tan usado como es la clasificación visual de objetos.

Agradecimientos

Agradecemos a Francisco Nieto su colaboración tanto en orientaciones técnicas, material usado y tiempo dedicado, así como al departamento de mecánica de la ETSI por las facilidades prestadas para el uso del robot. \otimes

Referencias

- Jähne, B. (1991): *Digital Image processing, Concepts, Algorithms and Scientific Applications*. Springer-Verlag.
- Müller, B., Reinhardt, J., Strickland, M.J. (1995): *Neural Networks, An introduction*. Springer.
- Haykin, Simon. (1994): *Neural Networks, A Comprehensive Foundation*. IEEE PRESS.
- Hillera, José R., Martínez, Víctor J. (1995): *Redes Neuronales Artificiales, fundamentos, modelos y aplicaciones*. Rama serie paradigma.
- FU, K. S., González, R.C., Lee, C.S. G. (1988): *Robótica, Control, detección, visión e inteligencia*. Mc Graw Hill.
- Gasca González, Mariano. (1991): *Cálculo numérico I*. UNED.
- STAUBLI S.C.A. (1998): *Manual de instrucciones V-P-I*. STAUBLI.
- STAUBLI V-2 (1998): *Instruction Manual*. STAUBLI.
- Russell, S., Norvig, P. (1996): *Inteligencia Artificial. Un enfoque moderno*. Prentice Hall.

Palabras clave: Redes neuronales; reconocimiento de patrones; visión artificial; robot; descriptores cartesianos de Fourier; reconocimiento de bordes.