

Extracción de las características fundamentales en el control de la trayectoria de un robot móvil

¹José Juan Hernández Mora, ²Erick Juárez Bernabé, Hernández Cortés Roberto, José Federico Ramírez Cruz

División de Estudios de Posgrado e Investigación
Departamento de Sistemas y Computación
Instituto Tecnológico de Apizaco
Avenida Instituto Tecnológico s/n 90300 Apizaco, Tlaxcala
¹jhhmora@yahoo.com.mx, ²itapizaco_inf@hotmail.com
Paper received on 12/08/08, accepted on 10/09/08.

Resumen. El proyecto se basa en la implementación de robots móviles en base a juguetes, obteniendo así tecnologías económicas, para este fin se diseñó una interfaz infrarroja entre la computadora y el robot, que permite controlarlo; además la retroalimentación se realiza con una webcam que obtiene una imagen de una escena en un ambiente controlado en iluminación y fondo, posteriormente el sistema extrae las características de objetos con algoritmos de Visión por Computadora, que nos permiten reconocerlos, clasificarlos y obtener su posición, para generar la trayectoria del robot móvil hacia un objeto dado como objetivo, cabe mencionar que en el espacio hay obstáculos que impiden el libre trayecto.

Palabras clave: robots móviles, visión por computadora, redes neuronales, identificación de objetos.

1 Introducción

Al construir un robot autónomo móvil pretendemos que sea capaz de desenvolverse en un entorno real, sin modificaciones que ayuden al desarrollo de la tarea a llevar a cabo. Esto conlleva el problema de que el conocimiento que tenemos del entorno es parcial: Las imágenes obtenidas varían en base al tipo de iluminación, el entorno suele ser cambiante y las tareas a llevar a cabo no están completamente especificadas.

Este proyecto tiene la finalidad de implementar robots móviles económicos basándose en la tecnología de los juguetes y el uso de computadoras personales equipadas con "webcams". De esta manera se desarrollan algoritmos propios para el reconocimiento del robot móvil y el control de trayectorias y que tenga como retroalimentación un sistema de procesamiento digital de imágenes de bajo costo

computacional. Así tendremos tecnología de robots móviles económica, que nos permita desarrollar proyectos más avanzados en robótica móvil.

2 Visión por Computadora

2.1 Obtención de Imágenes

Se refiere al procesamiento de datos de cualquier tipo que use el espectro electromagnético que produce la imagen. Una imagen implica múltiples lecturas dispuestas en una rejilla o matriz de dos dimensiones.

Las cámaras de vídeo tienen como ventajas:

- Brindan mucha información en cada imagen.
- Son el sensor ideal para detectar colores, texturas y para reconocer de objetos.

Por otro lado existen desventajas como:

- Alto coste computacional. Debido a esto, los robots suelen tener varios procesadores, algunos de ellos asignados exclusivamente al procesado de imágenes.
- El precio es elevado. Un sistema de visión artificial, suele constar de una o dos cámaras de vídeo, un procesador dedicado a cálculos y una plataforma.
- Las imágenes dependen de la iluminación, aunque ya se están explorando técnicas que disminuyen este problema [3].
- No es el sensor adecuado para medir profundidad, aunque se pueden hacer mapas de profundidad a partir de dos cámaras de vídeo [4] [5].

El primer paso en el desarrollo de esta investigación, fue la adquisición de imágenes del mundo real, es decir, utilizando la cámara Web se capturo la imagen del Robot (Megabyte) en posiciones diferentes con diferentes objetos, obteniendo 22 cuadros por segundo. La figura No. 1 muestra los módulos que componen este trabajo, cada uno de estos se explica más adelante.

2.2 Sistema de visión

El propósito de emplear un sistema de visión es el de permitir obtener una completa obtención de datos para los diferentes procesos. La cámara Web utilizada proporcionara la información sobre las coordenadas espaciales del móvil y obstáculos dentro de ambiente que serán procesadas por los algoritmos.

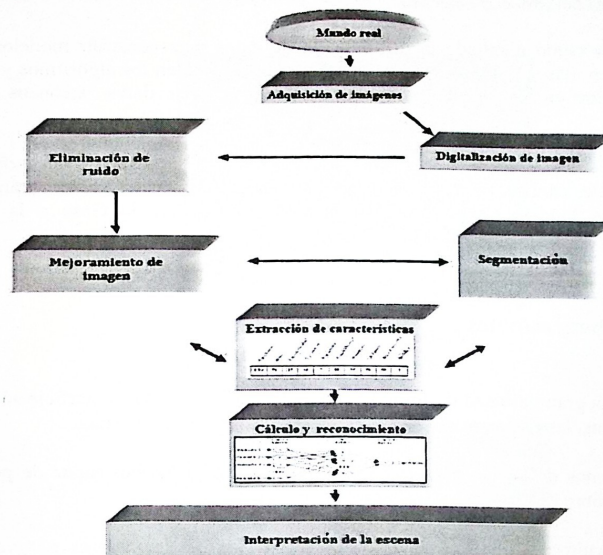


Fig. 1 Diagrama descriptor de la metodología propuesta.

2.3 Selección de características

La extracción o selección de características relevantes es una fase primordial para el reconocimiento de patrones y el aprendizaje automático. La gran mayoría de los procesos en el mundo real se representan en función a un gran número de variables o características que se supone forma parte del mismo proceso sin embargo esta representación a menudo se sujeta meramente a un pequeño número de dichas variables.

Así podemos definir a la selección o extracción de características como:

“La obtención de un conjunto de valores numéricos que representen al objeto”

[1]

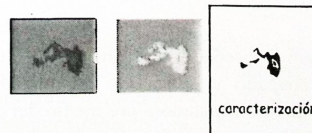


Fig 2 Caracterización de una imagen

Estos valores nos darán la pauta para determinar la función de transformación de entrada /salida del proceso analizado.

Dándonos como resultado la disminución de datos para construir modelos de categorización más sencillos y así lograr una mayor rapidez en los algoritmos y una mejor precisión en sus resultados debido a la supresión de datos erróneos (ver figura No. 2).

Los conjuntos de datos a los que se les dará tratamiento serán señales bidimensionales (imágenes recibidas de una webcam), a las cuales se extraerán sus características fundamentales que nos servirán para definir la clase a la que pertenecen.

3 Robots móviles

Existe una gran variedad de formas de desplazarse sobre una superficie sólida, entre los robots, las más comunes son las ruedas, las cadenas y las patas.

Así podemos definir a un robot móvil como: "Aquel robot capaz de poder desplazarse sobre una superficie."

Los vehículos de ruedas son, por lo general, los más populares por varios motivos prácticos, son más sencillos y más fáciles de construir, relativamente. Tanto los robots basados en cadenas como en patas se pueden considerar más complicados y pesados, generalmente, que los robots de ruedas para una misma carga útil. A esto podemos añadir el que se pueden transformar vehículos de ruedas de radio control para usarlos como bases de robots.

3.1. Prototipo MEGABYTE

Megabyte es un juguete ordinario que se comunica vía infrarroja con un control remoto, ver figura No.3.

4 Subsistema de control

El subsistema de control diseñado para este trabajo, es la interfaz que permite controlar el encendido/apagado y el sentido de giro de los motores del robot (MEGABYTE) a través de la computadora por medio del puerto paralelo (ver figura No.4), logrando así cumplir con tareas sencillas de navegación, de tal forma que la computadora manda la señal activa los servomotores del robot a través de las secuencias infrarrojas que el perro robot reconoce. Se encuentra diseñada y

acondicionada para permitir la elección de puerto, calibración del voltaje de armadura de cada motor para lograr la misma velocidad de trabajo en cada uno [17].

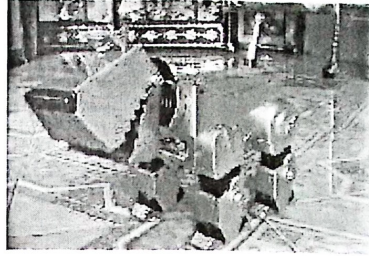


Fig. 3 Robot Móvil MEGABYTE

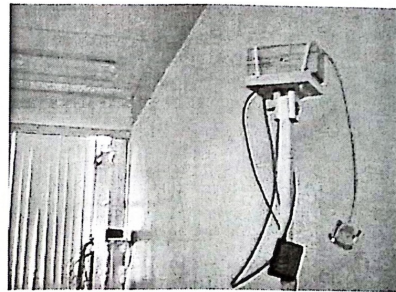


Fig. 4 Subsistema de control del robot móvil

5 Sistemas de Visión por Computadora

El sistema de visión está diseñado de tal manera que recibe las imágenes del mundo real por medio de una cámara Web para ser procesadas. Se hace notar que en este trabajo se tiene un ambiente controlado de iluminación y fondo para facilitar el uso de algoritmos de Visión por Computadora.

La primera fase es la del proyecto consiste en obtener la imagen, para posteriormente extraer las características más relevantes y así generar una trayectoria desde la posición del robot hasta el objeto de interés, esto se hace con la cámara web y se obtienen 22 imágenes por segundo, esto se hace con un programa en visual Basic ver 6.0 como se muestra en la figura No. 5.

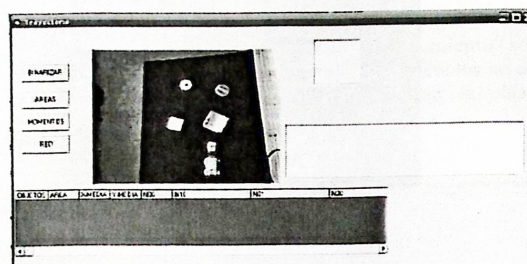


Fig. 5 Imagen obtenida del robot y varios objetos

El segundo paso es de pasar la imagen adquirida por un filtro pasa bajo de orden 7x7 (1), para disminuir el ruido y pasarla del plano de color al plano de grises. El resultado se ve en la figura No. 7.

. Matriz de 7*7 para el proceso de suavizado

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \cdot \frac{1}{49} \quad (1)$$

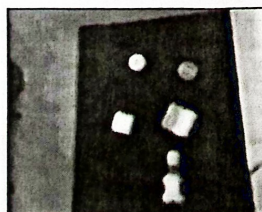


Fig. 7 Proceso de suavizado de la imagen

El tercer paso es binarizar la imagen de la figura No. 7 para esto se aplica el algoritmo de Otsu para obtener el umbral óptimo, que esta dado por la fórmula (16) para obtener el dicho umbral [9]. La imagen resultante se ve en la figura No. 8:

$$\delta^2(t) = \frac{m_1(t)p_2(t) - m_2(t)p_1(t)}{p_1(t)p_2(t)} \quad (2)$$

$$T = \arg \left\{ \max_t \left| \delta^2(t) \right| \right\} \quad (3)$$

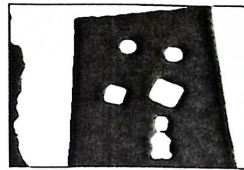


Fig. 8 Proceso de binarización de la imagen.

Como se puede apreciar los diferentes objetos así como el robot se distinguen perfectamente del fondo, cabe destacar que el robot aun no es tratado como un solo objeto, para solucionar este problema la imagen de la figura No. 8 es pasada por un algoritmo de dilatación para darnos como resultado la imagen de la figura No. 9.

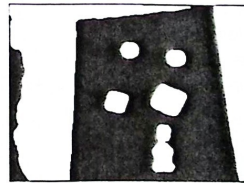


Fig. 9 Proceso de dilatación de la imagen

De esta manera cualquier objeto que se encuentre muy cerca de otro será tratado como uno solo, esto no afecta a la trayectoria generada pues los espacios entre los objetos no sobrepasan el perímetro del robot por lo que no importa que se unan porque de todos modos el móvil no puede pasar entre estos objetos.

Para el siguiente paso se debe localizar las diferentes áreas de los objetos con los que se trabajaran en el siguiente proceso, aquí se implemento un algoritmo de detección de áreas que se basa en asignarle un valor a partir del 100 a cada región localizada y por medio de la vecindad-8 se hacen varias pasadas hasta obtener un valor uniforme para cada área como se muestra en la imagen de la figura No. 10.

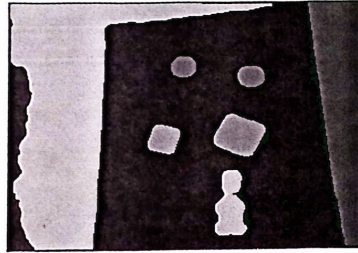


Fig. 10 Proceso de detecciones de cinco áreas

Para la siguiente fase la imagen resultante de la figura No. 10 es utilizada por el algoritmo que calcula los momentos geométricos de cada una de las áreas de los cuales, los momentos invariantes son los más importantes de ellos, pues estos serán utilizados por la red neuronal para poder identificar que objetos es el perro y cual son los obstáculos para así poder trazar la trayectoria a partir del robot como punto de inicio.

Los momentos utilizados en este trabajo, son propiedades numéricas que se pueden obtener de una determinada imagen, teniendo en cuenta todos los píxeles de la imagen y no solo los bordes. La aplicación más útil de los momentos son las funciones que podemos extraer de ellos que son invariantes a transformaciones geométricas tales como: traslación, escalado y rotación [8].

Se clasifican en:

- Momentos Simples
- Momentos Centrales
- Momentos Centrales Normalizados

Los momentos centrales se usan para reconocer una imagen independiente de su situación en un eje de coordenadas. Para calcular los momentos centrales se usa:

$$\mu_{pq} = \iint (x - X)^p (y - Y)^q f(x, y) \quad (4)$$

$$MC_{pq} = \sum \sum (x - X)^p (y - Y)^q f(x, y) \quad (5)$$

Los momentos simples de orden 0, suman todos los píxeles cuyo valor es uno, es decir los que son distintos de blanco, por lo tanto calculan el área. Estos momentos están dados por:

$$M(0,0) = \sum_x \sum_y f(x,y) \quad (6)$$

Los momentos de orden simple 1 se usan principalmente para hallar el centro de masas de una figura y están dados por las siguientes formulas:

$$M(1,0) = \sum_x \sum_y xf(x,y) \quad (7)$$

$$M(0,1) = \sum_x \sum_y yf(x,y) \quad (8)$$

Los momentos de orden 2 es en donde comienza el análisis de las imágenes mediante el reconocimiento de formas y están dados por:

$$U(p,q) = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (9)$$

Los últimos momentos de orden 3 sirven para calcular los momentos invariantes y como su nombre lo indican estos se mantienen invariantes ante rotaciones, traslaciones y cambios de escalas de una figura concreta. Están dados por:

$$I_1 = N(2,0) + N(0,2) \quad (10)$$

$$I_2 = (N(2,0) - N(0,2))^2 + 4(N(1,1))^2 \quad (11)$$

$$I_3 = (N(3,0) - 3N(1,2))^2 + (3N(2,1) - N(0,3))^2 \quad (12)$$

$$I_4 = (N(3,0) + N(1,2))^2 + (N(2,1) + N(0,3))^2 \quad (13)$$

$$I_5 = (N(3,0) - 3N(1,2))(N(3,0) + N(1,2)) \left[(N(3,0) + N(1,2))^2 - 3(N(2,1) + N(0,3))^2 \right] \\ + (3N(2,1) - N(0,3))(N(2,1) + N(0,3)) \left[3(N(3,0) + N(1,2))^2 - (N(2,1) + N(0,3))^2 \right] \quad (14)$$

$$I_6 = (N(2,0) - N(0,2)) \left[(N(0,3) + N(1,2))^2 - (N(2,1) + N(0,3))^2 \right] \\ + 4N(1,1)(N(3,0) + N(1,2))(N(2,1) + N(0,3)) \quad (15)$$

$$I_7 = (3N(2,1) - N(0,3))(N(3,0) + N(1,2)) \left[(N(3,0) + N(1,2))^2 - 3(N(2,1) + N(0,3))^2 \right] \\ + (3N(1,2) - N(3,0))(N(2,1) + N(0,3)) \left[3(N(3,0) + N(1,2))^2 - (N(2,1) + N(0,3))^2 \right] \quad (16)$$

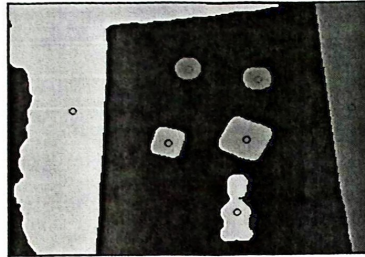


Fig. 11 Proceso de obtenciones de los momentos y centros de cada área.

Como se puede ver los centros de gravedad de cada área son visualizados para mayor detalle por la interfaz (ver Figura No.11). Los datos obtenidos de este proceso además de los ya mencionados son:

Las áreas localizadas (objetos), el área de estos objetos (AREA), los puntos centrales de los mismos (XMEDIA, YMEDIA), los momentos de orden 0, 1, 2 (N00, N10, N01, N20, N02, N11, N30, N12, N21, N03). (Ver Tablas No. 1, No. 2 y No. 3)

Tabla 1 Momentos de orden 0, 1, 2

OBJETOS	AREA	XMEDIA	YMEDIA	N00	N10	N01	N20
1	19294	56	105	1	0.326864100399596	-0.320653259729466	0.3797002
2	7771	301	101	1	-2.27216517896541	2.26643995835697	5.70823
3	403	160	64	1	-4.7825950485893	4.79067670547588	22.94617
4	443	221	74	1	-6.99276071364107	6.98957223219823	48.96668
5	1355	211	131	1	-2.17999964012191	2.16478250723935	4.823309

Tabla 2 Momentos de orden 0, 1, 2

OBJETOS	N02	N11	N30	N12	N21
1	0.177497946614701	-0.151755091454798	0.329790726394546	5.48835508099358E-02	
2	5.15341073302422	-5.11349640571115	-15.3193980080195	-11.5490129021438	
3	23.0393327955963	-22.9142966214927	-110.440353936226	-110.201728978099	
4	48.9532226915806	-48.8762286686811	-343.362836123113	-342.277018202433	
5	4.7802230361787	-4.71553056194769	-10.8238328583719	-10.4053746286268	

Tabla 3 Momentos de orden 0, 1, 2

OBJETOS	N21	N03
1	-0.140623819875167	-6.89797030759645E-02
2	12.7827265260829	11.7545602691891
3	109.946714921267	111.210284250341
4	342.234435151734	343.488591842651
5	10.4251919906912	10.7552002443784

Así como los momentos de orden 3 (Ver tablas No. 4 y No.5).

Tabla 4 Momentos invariantes

OBJETOS	h1	h2	h3	h4
1	0.557198239021237	0.133004219773537	0.369004868934346	0.191907936376316
2	10.8616481091141	104.899214568735	3203.86538589441	1323.98994611081
3	45.9855118866565	2100.26863624367	242320.18502991	97593.3470353638
4	57.919907672142	9555.54309673143	2344213.64309162	940318.079589898
5	9.60353208698135	86.9467703273068	2188.44378306489	899.288265757526

Tabla 5 Momentos invariantes

OBJETOS	h4	h5	h6	h7
1	0.191907936376316	3.06135646610163E-02	6.99807619351696E-02	1.16496934378215E-02
2	1323.98994611081	1583433.33394834	13551.3473499759	33219.4996705557
3	97593.3470353638	9459743164.96483	4472574.84506162	-15586.3264367589
4	940318.079589898	881180666556.485	91918400.7469101	-3119583.74108887
5	899.288265757526	780177.257445064	6481.30932700191	-251.56858231721

Así bien una vez obtenidos los siete momentos invariantes procedemos a pasarlos como parámetros a una red neuronal, en este caso se seleccionó el modelo backpropagation [16] para poder reconocer cual de todos los objetos es el robot, punto de llegada y obstáculos en general. La figura No. 12 detalla la arquitectura de la red a utilizada.

Las entradas a la red son un vector que contiene la información de los algoritmos de procesamiento digital de imágenes y cálculos geométricos (momentos) los cuales sirven para identificar los objetos de la imagen. La red se entrena con diferentes objetos además del objeto destino y el robot móvil con 2000 épocas.

Una vez clasificados a los objetos por la red neuronal se procede a distinguir los objetos (ver figura No. 13), para este caso marcamos de amarillo al móvil y al objeto destino. Y así trazar la trayectoria que el robot va a seguir hasta el objeto destino. La salida deseada para el reconocimiento de objetos, Megabyte y otros objetos se utilizó en un rango de 0 a 0.4 clasificándolo como un objeto extraño, de 0.2 a 0.7 como robot Megabyte y de 0.5 a 1 como otro objeto extraño. Para estas

pruebas se utilizaron 30 imágenes. Las características de cada muestra con las cuales se entrenó la red neuronal artificial fueron los momentos invariantes ideales para cada objeto. Los resultados obtenidos en estas pruebas fueron del 91 % de reconocimiento.

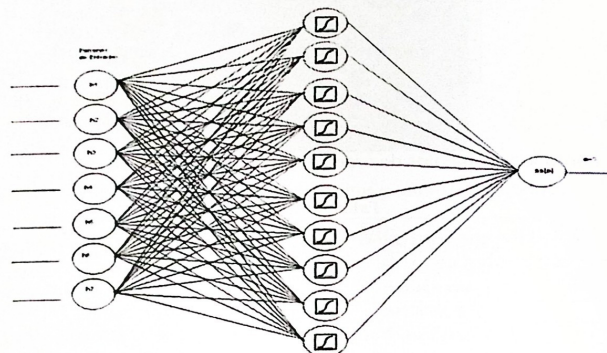


Fig. 12 Arquitectura de la red neuronal implementada

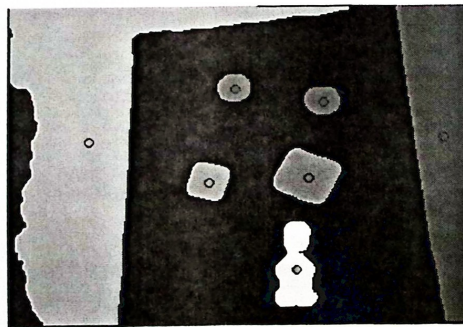


Fig. 13 Objeto origen (amarillo) y objeto destino (rojo).

6 Conclusiones

Una de las partes importantes en la construcción de un robot móvil más allá de la alta tecnología en hardware y software son los principios básicos para que este logre cumplir con las tareas asignadas.

En este trabajo se logró con una webcam, de baja resolución (1.3MPíxeles) identificar perfectamente el robot móvil, el objetivo y demás objetos; con técnicas

de procesamiento digital de imágenes, como el algoritmo de Otsu, algoritmos de dilatación y suavizado de la imagen, segmentación de objetos por medio de algoritmos de detección de áreas y extracción de características las áreas encontradas con momentos invariantes y la clasificación e identificación con una red neuronal Backpropagation de tres capas y 40 neuronas ocultas, para finalmente poder trazar una trayectoria.

También se obtuvo el control del robot móvil con una interfaz sencilla a través del puerto paralelo y el control del robot original que emite la señal infrarroja.

Referencias

1. Campoy Pascual, "Extracción de Características", Dpto. Automática, Ingeniería Electrónica e Informática Industrial. Universidad Politécnica de Madrid. 2008
2. Béjar Javier, "Apuntes de Aprendizaje Automático", Inteligencia Artificial, Departamento de Lenguajes y Sistemas Informáticos. Ingeniería en Informática. 2º cuatrimestre 2005/2006
3. Umbaugh F., Scoot E. C., "Computer Vision and image Processing". Englewood Cliffs, Prentice-Hall, 1998.
4. Dean, T., Allen, J., Aloimonis Y., "Artificial intelligence: Theory and Practice". Benjamin/Cummings Publishing, 1995.
5. Horn B. K. P., "Robot Vision". MIT Press, 1986
6. Cañas Plaza J. Maria "Visión 3D en robots móviles" III Jornadas de robótica URJC, 17 octubre 2006
7. Marrón Romera Marta "Sensores para posicionamiento de robots móviles" Departamento de Electrónica. Universidad de Alcalá.
8. Hernández Cortés Roberto. "Control de trayectorias de un robot móvil a través de una cámara web", División de Estudios de Posgrado e Investigación. Instituto Tecnológico de Apizaco.
9. Begoña Acha y Carmen Serrano, "Tratamiento digital de imágenes", 5º Telecomunicación. Universidad de Sevilla 2007/2008
10. Vega J, "Visión por computadora". Facultad de Ingeniería UACH 1998, P 1,2
11. Parker, J.R., "Algorithms for Image Processing and computer Vision", Wiley Computer Publishing, 997.
12. Ripley, B.D "Pattern recognition and Neural Networks", Cambridge University Press, 1996.
13. Carl G Looney. "Patter recognition Using Neural Networks", Oxford University Press, 1997.
14. "Técnicas de programación", cap. 32. ed, McGraw Hill.
15. "Revista Pc World", Octubre 94. Sección de programación.
16. Rojas Raul, "Neural Networks A Systematic Introduction" Springer, 1996
17. Contreras Badillo Edgar Alan. "Interfaz para el control de un robot móvil" Facultad de Ciencias Básicas Ingeniería y Tecnología de la Universidad Tecnológica de Tlaxcala