

# SISTEMA DE VISION ARTIFICIAL PARA RECONOCIMIENTO DE OBJETOS.

**Ing. Melvin Andrés González Pino\***  
**Ing. José Gabriel Ojeda Rodríguez\*\***

Universidad de Pamplona  
Grupo de Investigación: Automatización y Control A&C.  
Pamplona, Colombia  
\*Ingmelvin@unipamplona.edu.co  
\*\*Ingjgabriel@unipamplona.edu.co

**Abstract.** Sistema de Visión Artificial desarrollado bajo MS Visual Basic que permite el reconocimiento de objetos para que un robot ubique y manipule objetos de forma autónoma.

**Keywords:** Reconocimiento, Visual, Robótica, Píxeles

## 1. FUNCIONAMIENTO Sistema De Reconocimiento Y Localización De Objetos.

Es un sistema experimental ya desarrollado y probado, el cual tiene como objetivo básico el darle la posibilidad al robot de localizar e identificar en cuanto a forma un objeto utilizando una imagen de video proporcionada por una cámara ubicada en una posición estratégica, la cámara estará ubicada sobre el brazo apuntando hacia abajo a una altura que le permita captar toda el área necesaria para que el robot realice una tarea específica, este sistema trabaja bajo condiciones especiales que serán predeterminadas por el operador del brazo; una de estas condiciones es que el objeto a localizar y reconocer debe tener un color diferente al de la superficie de trabajo y del robot, ya que el computador identificara el objeto por

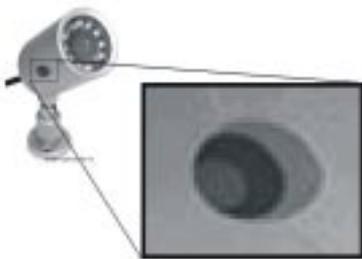
su color específico, el tamaño del objeto no debe ser muy pequeño ya que esto dificulta y retarda el proceso de localización, deben tener superficies opacas, la iluminación debe ser difusa, la geometría del objeto debe permitir su manipulación por las pinzas del brazo, por ejemplo una pieza con forma de pirámide no podría ser manipulada. El proceso de localización es tal vez el mas sencillo y trabaja de la siguiente manera: primero que todo el brazo debe moverse de manera que no se interponga entre la cámara de video y el área de búsqueda, luego el programa en el computador captura una imagen de la cámara de video la cual podemos acceder como un grafico cualquiera, a continuación empezamos a rastrear el objeto mediante una rejilla o cuadrícula formada por una serie de puntos (figura 1), es decir se revisa el color de cada

punto de izquierda a derecha hasta detectar el objeto, la distancia entre los puntos debe ser menor al ancho o el alto del objeto para que este sea interceptado por lo menos por un uno de estos puntos, en el momento que el color de uno de estos puntos coincida con el rango de colores buscado se habrá encontrado el objeto.



**Fig. 1. Muestreo de una imagen mediante malla de puntos o píxeles.**

Se habla de rango de colores ya que se presenta el fenómeno de antialias, que consiste básicamente en asignarle un color a un píxel, este color es el promedio de los colores presente en el área de imagen correspondiente a este píxel. En la figura 2 podemos observar como se compone una imagen a partir de píxeles.



**Fig. 2. Composición de imagen a partir de píxeles**

Se recomienda utilizar los colores primarios para las piezas, son mas fáciles de identificar debido que el programa utiliza el código de colores RGB ( red green blue ) usando 24 bits para especificar un color, por ejemplo el código 255,0,0 correspondería al color rojo, el 0,0,0 seria el negro, el 255,255,255 serial el color blanco, obteniendo una gran gama de colores combinado los tres bytes del código RGB (figura 3), mas de 16 millones de colores, aunque usando esta resolución de color se puede diferenciar entre tonos de colores muy parecidos se recomienda utilizar rangos de colores muy grandes y separados debido a que las imágenes pueden verse afectadas por factores externos como por ejemplo la iluminación.

Después de encontrar el objeto utilizando la cuadrícula de puntos, se procede a registrar la forma, para esto empezamos en el píxel donde se encontró el color determinado, luego nos desplazamos al siguiente píxel a la izquierda es decir en el eje x (aunque podría ser en cualquier dirección preferiblemente ortogonal), repitiendo esta operación hasta encontrar el color característico de la superficie de trabajo, regresamos al ultimo píxel que presento un color distinto al de la superficie de trabajo no el color del objeto, ya que muy probablemente que se presentara un efecto de antialias en este punto, ya ubicado en este punto se procede a revisar los píxeles vecinos empezando desde el píxel de la derecha en sentido antihorario, hasta encontrar el color de la superficie de

trabajo. Avanzando hasta el último píxel que presente un color distinto al de la superficie de trabajo, repitiendo el proceso pero esta vez empezando desde el píxel anterior, de esta manera se traza el contorno del objeto en sentido horario hasta regresar a la posición inicial ( figura 3).



**Fig. 3. Trazo del contorno de un objeto en una imagen.**

Después de determinar el contorno de la pieza se procede a encontrar su centro geométrico esto lo hacemos hallando la media de las componentes X y Y de las coordenadas que conforman su perímetro luego procesando la información proporcionada por la imagen de video capturada obtenemos una serie de características especiales (figura 4) que en teoría no dependan de la posición y del ángulo del objeto como

- Longitud del perímetro en píxeles: Es el número de píxeles que forman el contorno del objeto y se haya haciendo el seguimiento del borde hasta llegar a la posición inicial.



**Fig. 4. Características especiales de la imagen.**

- Distancia mínima en píxeles: Es la distancia mínima de un punto del contorno al centro geométrico ( podría ser también al centro de gravedad ) del objeto llamado también vector de distancia mínima.
- Distancia máxima en píxeles: Es la distancia máxima de un punto del contorno al centro geométrico (podría ser también el centro de gravedad) del objeto llamado también vector de distancia máxima.
- Área : Es el área ocupada por el objeto en píxeles cuadrados, esta se haya haciendo un barrido píxel por píxel del rectángulo que lo contiene y revisando si el color del píxel corresponde a un color diferente al fondo de trabajo.
- Distancia promedio al centro geométrico.
- Distancia promedio al centro de gravedad.
- Distancia entre el centro geométrico y el centro de gravedad.

- Relación entre la distancia máxima y la mínima.
- Distancia entre los puntos de distancia mínima y distancia máxima.
- Angulo mínimo entre los vectores de distancia mínima y máxima .
- Numero de agujeros.
- Mapa polar:

Consiste en una grafica de la distancia en función del ángulo alrededor del contorno, muy importante para el reconocimiento de objetos con la misma forma exclusiva pero de distintos tamaños, el reconocimiento se hace reescalando el patrón de comparación utilizando como dato la distancia máxima ( no la mínima ya que esta puede variar en grandes porcentajes dependiendo de factores como el ángulo del objeto entre otros ) de tal forma que coincidan luego se desplaza la grafica polar del objeto analizado hasta que coincidan los puntos de distancia máxima a continuación se halla el área ( valor absoluto ) de diferencia entre el patrón y la grafica analizada, y si esta diferencia no es muy grande se puede afirmar que las formas coinciden. Cabe destacar que el eje X de la grafica debe ser reescalado para que coincida con el del patrón de comparación, ya que en la practica el perímetro de un una imagen correspondiente a un objeto puede variar alrededor de un 20% en promedio o incluso mas al variar el ángulo de orientación. Aunque existen casos especiales en donde un objeto puede presentar varios vectores de

distancia máxima en este caso se repite el procedimiento anterior para cada uno de estos puntos haciendo coincidir cada uno de ellos con el punto de distancia máxima del patrón de comparación, este caso es común en objetos con formas con simetría radial o axial como por ejemplo un cuadrado presenta cuatro vectores de distancia máxima, otro caso especial es el de un circulo, para reconocer esta forma en particular solo habría que constatar de que la distancia mínima, máxima, y la promedio son muy parecidas, ahora como se diferenciaría un disco de un anillo de igual radio exterior, teniendo en cuenta que la distancias mínima, máxima y media coinciden en gran manera, en este caso se recurriría a comparar las áreas, ya que por lógica el disco tendría mas píxeles cuadrados de área que el anillo.

- Diámetro: Es el diámetro mínimo para que una circunferencia pueda contener la figura del objeto.
- Rectángulo de contención mínima: Consiste en hallar el ancho y el alto de un rectángulo mínimo que pueda contener la figura del objeto. En este caso también es importante la relación ancho alto.
- Determinar si el centro geométrico y de gravedad se encuentran dentro o fuera del contorno exterior.
- Color : Es el color promedio del objeto, aclarando que se debe contar con una buena iluminación, es decir ni muy baja ni muy alta, esta debe ser difusa o ambiental,

de color blanca y los objetos a identificar deben tener cierto grado de opacidad para evitar la reflexión de la luz. Los objetos pueden tener texturas preferiblemente finas, con la restricción que contengan colores similares al del fondo de trabajo, ya que esto podría producir un reconocimiento erróneo del borde.

Entre otras características, aunque en la realidad practica no se pueden tomar las mismas medidas de ciertas características del mismo objeto, ya que estas varían en un cierto porcentaje, la variación de este porcentaje es distinta para cada características, incluso se presenta una variación considerable de cada una de las características en dos tramas de video del mismo objeto sin hacerle modificación de iluminación, desplazamiento o rotación, esto debido al ruido eléctrico que se filtra en la señal de video, aunque se podrían aplicar filtros de imagen para disminuir este ruido. Esto sumado a la variación de las características por razones como ángulo de orientación que modifica características tales como perímetro y área hasta mas o menos un 15%, otra variación de estas características se presenta dependiendo de la distancia del objeto al eje focal de la cámara o al centro de la imagen de video capturada, ya que entre mayor sea esta distancia el objeto tiende a verse un poco mas pequeño debido al fenómeno llamado ojo de pez que es común en las cámaras de video, este problema aunque no es grave

se puede solucionar alejando la cámara del el plano de trabajo y adaptándole un lente de ampliación. Por estas razones se debe asignar un porcentaje de tolerancia para cada una de las características en el proceso de reconocimiento de objetos, en la figura 6 podemos apreciar como afecta el valor de tolerancia al proceso.

Aunque todas las características explicadas anteriormente nos permiten hacer un reconocimiento de un objeto , en la practica solo utilizamos las siguientes:

- Perímetro.
- Vector de distancia mínima.
- Vector de distancia máxima.
- Área.
- Distancia promedio al centro geométrico.

Ya que son suficientes para reconocer una gran cantidad de objetos con gran variedad de formas y tamaños. Si fuera necesario una mayor confiabilidad en la tarea de reconocimiento de objetos se podrían seleccionar las características mas adecuadas dependiendo de los tipos de objetos con que se trabajarían.

Con base en practicas realizadas con una tarjeta importadora de video y cámara CCD a color de mediana resolución con campo de visión ángulo 60 grados, situada perpendicularmente sobre un fondo gris oscuro opaco a una altura de 70 cms y utilizando varios objetos con diferentes características como tamaño, forma, color y haciendo todo tipo de variaciones en su posición y ángulo de orientación, se obtuvieron una gran cantidad de datos que nos permite obtener las variaciones

máximas para cada una de las características:

- perímetro: hasta un +/- 15% del promedio
- Vector de distancia mínima: hasta +/- 2 píxeles del promedio

- Vector de distancia máxima: hasta +/- 2 píxeles del promedio
- área: hasta un +/- 15% del promedio

- Distancia promedio al centro geométrico: hasta +/- 2 píxeles del promedio

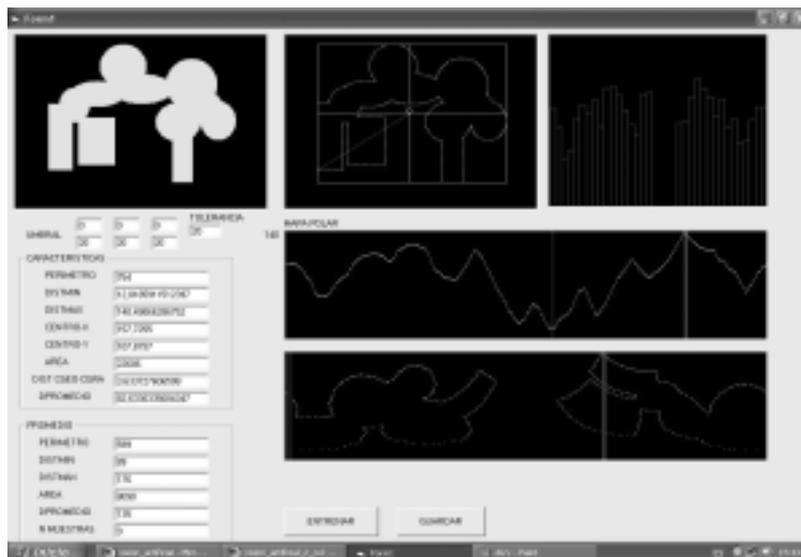
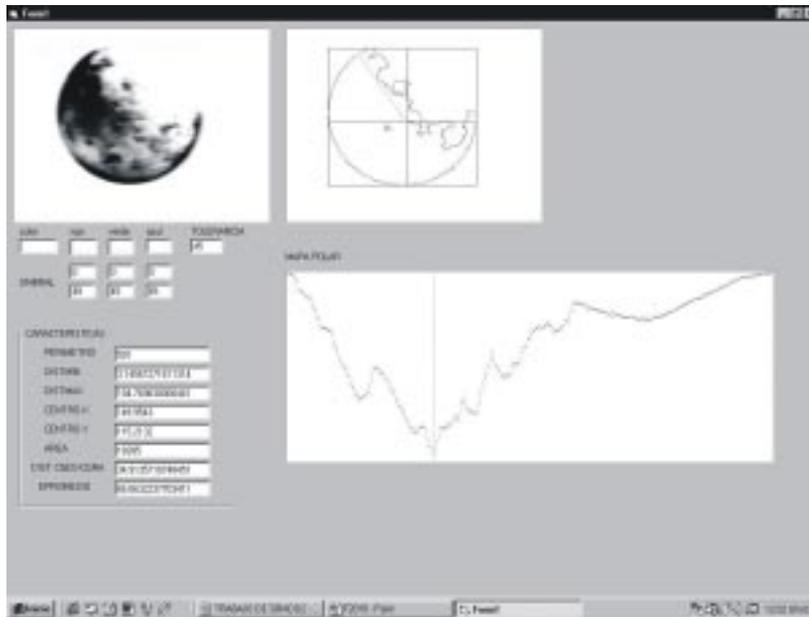


Fig. 5. Programa de reconocimiento

Estos valores pueden cambiar modificando la distancia de la cámara. Es muy importante minimizar las sombras ya que pueden producir en gran medida una variación en las características anteriores, como el perímetro y el área. Especialmente sobre un fondo claro.

Actualmente se está probando una técnica que hemos denominado comparación de formas por promedio de distancia por región angular, el cual tiene como objetivo acelerar el proceso de comparación de mapas polares sobre todo cuando los objetos tienen formas complejas, varios tamaños, presentan varios vectores de distancia mínima y máxima, y los números de patrones de comparación son grandes. Debido a que la comparación con un patrón implicaría varios desplazamientos del mapa polar en el eje X y posteriormente hallar el área de diferencia de los dos patrones.

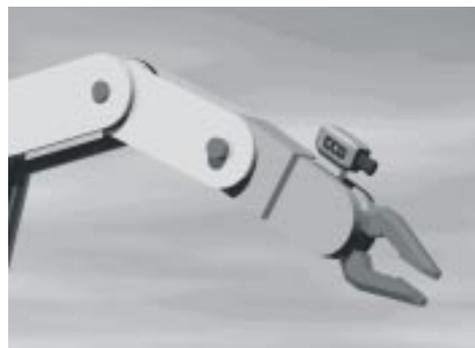
La técnica consiste básicamente en encontrar un ángulo especial único entre uno de los vectores de distancia máxima y uno de los vectores de distancia mínima se debe tomar un rango de tolerancia para detectar la distancia mínima y la máxima así como un ángulo mínimo permitido entre vectores vecinos. Para generar un patrón de 36 valores donde cada valor corresponde a un promedio de distancia al centro geométrico del conjunto de puntos pertenecientes al contorno que se encuentran dentro de una región angular de 10 grados, es decir el contorno del objeto se parte como un pastel en 36 partes iguales y sacamos el promedio de distancia de los puntos dentro de cada parte. Aclarando que el

ángulo 0 grados se fija utilizando la localización de los vectores de distancia mínima y máxima mencionados anteriormente, esto aceleraría en gran medida el proceso de comparación ya que solo sería necesario ubicar el valor máximo de promedios en el arreglo de 36 posiciones, reescalarlo, hacer la resta elemento a elemento y constatar si la diferencia está dentro del valor máximo de tolerancia.

En el desarrollo de este proyecto se trabajó con una tarjeta ATI ALL-IN-WONDER 128 que es una tarjeta importadora de video de 32 Mb, aceleradora de gráficos, importadora de video y sintonizador de televisión.

Esta tarjeta tiene la posibilidad de ajustar características como color, brillo, contraste y tinte del video importado así como filtros especiales para la reducción del ruido producido por la estática o interferencia externa.

Otra posibilidad en este proyecto es montar la cámara en un lugar estratégico del brazo de tal manera que el robot primero ubique la cámara para capturar la imagen, procesarla, localizar el objeto y posteriormente manipularlo.



**Fig. 6. Cámara de video montada de manera estratégica.**

## Sistema De Detección De Movimiento.

Este sistema experimental trabaja de una forma similar al sistema de reconocimiento de formas, pero tiene como objetivo detectar el cambio de posición de un objeto sobre un fondo inmóvil, localizando el centro geométrico de objeto en movimiento. Básicamente consiste en comparar dos imágenes de video consecutivas e identificar la regiones que hayan cambiado de alguna forma, al iniciar el programa este captura una imagen. Esta es rastreada en su totalidad con una cuadrícula de NxN píxeles para acelerar el proceso, el tamaño de la cuadrícula determina el área mínima de detección de movimiento e influye en el tiempo de procesamiento de cada imagen, en este caso no se pretende identificar el objeto en movimiento sino su ubicación, por esta razón podemos utilizar baja resolución en el proceso de rastreo, es decir las casillas de la cuadrícula serán mas grandes, luego de configurar la cuadrícula de rastreo se procede a leer el color de los píxeles correspondientes al centro de cada casilla, estos datos son almacenados en un arreglo bidimensional, el cual será comparado uno a uno con los datos obtenidos de la siguiente imagen usando el algoritmo de seguimiento de borde se determinara el contorno del objeto en movimiento y su centro geométrico, en la comparación de los datos de la matriz actual con la anterior es conveniente permitir un valor de tolerancia debido a que dos imágenes consecutivas idéntica siempre presentan diferencias en el color de cada uno de sus píxeles.

La matriz bidimensional que guarda los datos de los colores de los píxeles se ira actualizando con cada imagen que sea tomada por la cámara de video. Después de determinar la ubicación del objeto en movimiento se puede realizar cualquier tipo de tarea con la información obtenida.

Este sistema abre las puertas a muchas posibilidades de uso y aplicaciones como por ejemplo:

- Localización de objetos en entornos no predefinidos: Por ejemplo se podría aplicar el algoritmo de identificación de objetos (siempre y cuando este en movimiento) comparando los píxeles del contorno con los de la imagen anterior, esto solucionaría el problema del sistema de reconocimiento de formas sobre un fondo no uniforme o texturizado.
- Predicciones de trayectorias: Ya es conocido el trabajo realizado en el exterior por unos estudiantes que construyeron un brazo mecánico que tenia la capacidad de atrapar objetos lanzados hacia el, apoyado por dos cámaras de video ubicadas ortogonalmente dentro de un cuarto, estas cámaras proporcionaban las imágenes de donde se hallaba la posición del objeto lanzado, con la desventajas de las paredes debían ser de un color uniforme distinto al del objeto con este sistema se podría solucionar este problema. se podrían predecir trayectorias de un objeto hallando su velocidad, dirección y posición ya sea bidimensional o tridimensional en el espacio aplicando la leyes que rijan

su movimiento, como por ejemplo en lanzamiento parabólico se aplicarían la ley de la gravedad y la resistencia del aire, para determinar su posición en un instante de tiempo cualquiera.

- Sistemas de vigilancia continua: Sería un gran aporte en el desarrollo de sistemas de seguridad relacionados con la vigilancia, un ejemplo sería usar una cámara de video montada sobre una plataforma móvil la cual le permita cubrir una gran zona a su alrededor, esta cámara tomaría unas cuantas imágenes de una zona para buscar algún tipo de movimiento en ella, en caso de no encontrarlo la cámara sería orientada a otra zona y así sucesivamente hasta cubrir toda la zona en vigilancia, en caso de detectar algún movimiento el sistema activaría alguna alarma o otros dispositivos de seguridad. Cabe aclarar que sistema no realizaría el proceso de detección mientras la cámara se mueva.

Una condición para este sistema sería que el entorno de vigilancia debe ser estático y no perturbado por fuentes de luz externas. Otra condición tiene que ver con el nivel de luminosidad, debe

existir un nivel mínimo pero tampoco muy grande preferiblemente que produjera sombras, una posibilidad sería utilizar cámaras infrarrojas para vigilancia nocturna, en caso de un área de vigilancia exterior este no sería afectado por el cambio de luminosidad producido por el día y la noche ya que este cambio es gradual y es aceptado por el valor de tolerancia predeterminado. En teoría este sistema puede ser burlado con movimientos demasiado lentos para realizarlos en la práctica alrededor de un píxel por cada 5 minutos, aunque esto depende de la tolerancia usada y de las imágenes por segundo procesadas.

## CONCLUSION

Este sistema de visión artificial se ha probado con resultados muy satisfactorios, logrando reconocer una gran cantidad de objetos de diferentes formas, tamaños y colores, con un alto porcentaje de confiabilidad. El programa de reconocimiento puede actualizar una base de datos con una gran cantidad de objetos, otra fase de este proyecto es la integración del sistema de reconocimiento de objetos con un robot manipulador, permitiéndole a este la ubicación y manipulación de objetos de forma autónoma.

## REFERENCIAS BIBLIOGRÁFICAS:

1. J. Lladós, E. Martí, Graph-edit algorithms for hand-drawn graphical document recognition and their automatic introduction in CAD systems, ( accepted for publication in Machine Graphics & Vision journal, special issue on Graph transformations in Pattern Recognition and CAD, 1999).
2. D. Ponsa, A. F. Solé, A. López, C. Cañero, P. Radeva, J. Vitrià, Regularized EM, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 3-10, 1999.
3. F. Lumbreras, R. Baldrich, M. Vanrell, J. Serrat, J.J. Villanueva, Multiresolution colour texture representations for tile classification, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 145-152, 1999.
4. D. Guillamet, J. Vitrià, Using Eigenspace analysis of color distributions for object recognition, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 219-226, 1999.
5. A. Pujol, F. Lumbreras, X. Varona, J. J. Villanueva, Template matching through invariant eigenspace projection, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 227-234, 1999.
6. A.F. Solé, A. López, C. Cañero, P. Radeva, J. Saludes, Crease enhancement diffusion, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 279-286, 1999.
7. J. M. Sánchez, X. Binefa, Automatic digital TV commercial recognition, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 313-320, 1999.
8. A. López, R. Toledo, J. Serrat, J. Villanueva, Extraction of vessel centerlines from 2D coronary angiographies, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 489-496, 1999.
9. A. López, W. Niessen, J. Serrat, K. Nicolay, B. Ter Haar Romeny, J. Villanueva, M. Viergever, New improvements in the multiscale analysis of trabecular bone patterns, Proceedings of the VIII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Bilbao, 1: 497-504, 1999.
10. X. Binefa, X. Roca, J. Vitrià. A Contrast Approach to Depth from Focus. VII Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Barcelona, Abril 1997.
11. Javier Varona and Juan Jose Villanueva. NeuroFilters: Neural Networks for image Processing. Vision Systems: New image Processing Techniques and Applications Algorithms, Methods, and Components. Proceedings of the SPIE. Munich, June 1997.