

# Sistema de detección de esquinas en imágenes digitales en tonos de gris, basado en redes neuronales artificiales

## Corner detection system in gray-scale digital images based on artificial neural network

Hugo de Jesús Mesa Yepes<sup>1</sup>, M Sc (c); Jhon William Branch Bedoya<sup>2</sup>, Ph D; Cesar López Amaya<sup>1</sup>, Ing.  
1. Grupo de Investigación en Sistemas Inteligentes, Institución Universitaria de Envigado - IUE, 2. Escuela de Ingeniería de Sistemas, Universidad Nacional de Colombia - Sede Medellín  
hjmesa@gmail.co, jwbranch@unalmed.edu.co, clopezam@une.com.co

Recibido para revisión: 26 de Septiembre de 2008, Aceptado: 28 de Noviembre de 2008, Versión final: 15 de Diciembre de 2008

**Resumen**—La detección de esquinas, es un proceso fundamental en muchas aplicaciones de visión artificial, tales como el reconocimiento de patrones, la identificación de objetos, tareas de planificación, calibración de cámaras, sistemas de navegación automática, entre otras. Las técnicas usadas actualmente para la detección de esquinas, presentan algún tipo de limitación, de las cuales las principales son: susceptibilidad al ruido, detección de esquinas falsas, imposibilidad para encontrar algunos puntos de esquina, alto costo computacional en consumo de tiempo y recursos de máquina, no se identifica el tipo de esquina. En el presente artículo, se presenta un nuevo modelo de detección de esquinas basado en propiedades geométricas, mediante la aplicación de una red neuronal, el cual identifica la esquina y el tipo de la misma. Los datos obtenidos, pueden ser aplicados en la reconstrucción tridimensional de objetos.

**Palabras Clave**—Visión artificial, Puntos de interés, Esquina, Reconstrucción 3D, Red neuronal artificial.

**Abstract**—The detection of corners, is a fundamental process in many machine vision applications, such as pattern recognition, objects identification, planning, camera calibration, automatic navigation systems, among others. The techniques currently used to detect corners, have some kind of limitation, of which the main ones are: susceptibility to noise, detecting false corners, unable to find some corner, high computational cost consumption of time and machine resources, not identification of the type of corner. This article presents a new model for detecting corners based on geometric properties, using a neural network, which identifies the corner and type. The obtained data can be applied in the reconstruction of three-dimensional objects.

**Keywords**—Artificial vision, Points of interest, Corner, 3D reconstruction, Neural network.

### I. INTRODUCCIÓN

El ser humano percibe al mundo que lo rodea esencialmente a través de su sentido de la vista. Más del 99% de la información que procesa el cerebro humano es captada a través de sus ojos [1]. Es por ello, que la idea de aumentar y mejorar la percepción del mundo a partir de la visión por computador, juega un papel importante en el desarrollo científico y tecnológico del quehacer humano.

La visión por computador es una rama de la inteligencia artificial, que tiene como propósito la descripción explícita y significativa de las características de los objetos físicos, captados en una imagen digital. Este objetivo se logra mediante la combinación de dos elementos: 1) un tratamiento de bajo nivel de una imagen, como es el reconocimiento de bordes y esquinas, remover ruidos, mejorar el contraste, entre otros., 2) un procesamiento de alto nivel, como es el caso de la segmentación de una imagen, reconocimiento de texturas, reconocimiento de objetos, unión de imágenes y detección de movimiento. Un sistema de visión no sólo se enfoca al problema de la extracción de información, sino también el como debe ser extraída, representada y empleada [2], [3].

Uno de los procesos fundamentales en visión por computador, es la detección y extracción de las características de los objetos presentes en la escena. Los bordes y las esquinas son considerados como una de las características básicas [4], [5], [6], [7], [8] [9] [10], [11] [12]. Los bordes y las esquinas sirven como elementos de referencia, llamados puntos de control o puntos prominentes, para procesos de alto nivel. Dichos puntos de control son útiles en los siguientes procesos: 1) La unión de

imágenes; algoritmos de correspondencia entre dos o más imágenes, 2) reconocimiento de objetos; por ejemplo, la detección de las esquinas de un poliedro, 3) calibración de cámaras; encontrar los parámetros de las deformaciones proyectivas y factores de escala provocados por la arquitectura de la cámara, y 4) reconstrucción tridimensional; a través de un conjunto de imágenes, información de contexto y parámetros de calibración de la cámara, se puede lograr una reconstrucción tridimensional de la escena observada.

Debido a que pequeños errores en las mediciones en el plano de la imagen provocan, en general, errores significativos en su interpretación tridimensional, la ubicación precisa de la esquina impacta directamente en la calidad de dicha interpretación, ya que se afectan a los procesos de más alto nivel; por ejemplo a los algoritmos de calibración y reconstrucción tridimensional. En este sentido, la exactitud en la ubicación de las esquinas es un factor fundamental en el desarrollo de un sistema de visión por computador.

El presente artículo plantea un nuevo método de detección de esquinas múltiples basado en redes neuronales artificiales, el cual permite determinar la ubicación del punto donde se encuentra la esquina a un nivel de píxel y el tipo de esquina detectado. El detector opera sobre una región cuadrada de la imagen definida por el usuario. Dicho detector tiene las propiedades siguientes:

1. Emplea redes neuronales artificiales (RNA) para la detección del punto y tipo de esquina. Se evaluaron los modelos de RNA "MLP, RBF y Kohonen", encontrando una mejor respuesta en las MLP y RBF.
2. Los datos usados como entrada en la RNA, son los valores de intensidad de color de los tonos de gris de los puntos de entornos cuadrados de la imagen. Se utilizaron entornos de 3x3, 5x5, 7x7 y 9x9, encantándose un mejor rendimiento en los datos de 5x5.

## II. MODELO DE DETECTOR DE ESQUINAS BASADO EN REDES NEURONALES ARTIFICIALES

Los detectores actuales buscan determinar la ubicación de los puntos de esquina, además de identificar algunas propiedades importantes tales como el ángulo de la esquina, la orientación y su relación con otros puntos de interés.

No se ha abordado la identificación del tipo de esquina (L, Y, K, T, X y Vértices), dato que es de importancia en el proceso de reconstrucción tridimensional y la planificación de tareas, debido a que la morfología de la esquina guarda una íntima relación con el tipo de objeto que la genera. La identificación del tipo de esquina puede ser usada, junto con los demás datos relacionados con la geometría, para realizar procesos de identificación de objetos en diferentes escenas y ampliar las técnicas de reconocimiento de patrones y objetos.

Otro de los factores que presentan dificultades en el uso de

los detectores, es que fueron desarrollados para operar bajo determinadas condiciones de adquisición de la imagen, lo cual implica el mantener estas condiciones en las imágenes usadas en el análisis, o la intervención de expertos en procesos de calibración del sistema y establecimiento de parámetros de operación bajo las nuevas condiciones. Este aspecto los hace poco eficientes al ser utilizados en condiciones en las cuales se presente una alta variabilidad en los procesos de adquisición de imágenes.

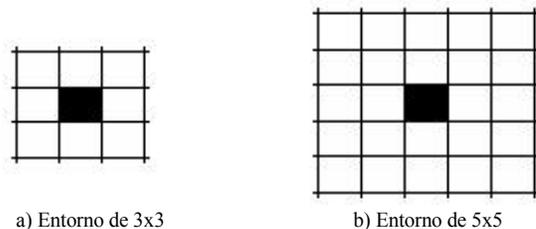
En la literatura revisada, se plantea la alternativa de uso de las redes neuronales artificiales como una opción al proceso de tratamiento de imágenes y reconocimiento de patrones [13], [14], [15]. El uso de las redes neuronales ha mostrado su aplicabilidad en proceso de reconocimiento de objetos [16], [17], detección de contornos [18], mejoramiento de imágenes [19], [20] y en otra serie de aplicaciones relacionadas con los procesos de visión artificial y procesamiento digital de imágenes [21].

Una alternativa para mejorar la detección de esquinas, es el empleo de técnicas basadas en redes neuronales, las cuales permiten reducir la influencia del ruido y el uso de parámetros, logrando con ello diseñar un método automático, capaz de realizar los procesos de detección bajo diferentes condiciones de adquisición y operación.

### A. Datos de entrada para los modelos de redes neuronales propuestos para el detector

Los datos con los cuales se alimentará la red neuronal para la implementación del detector de esquinas, dependen del color de los píxeles y de su entorno, un píxel no define una esquina por sí solo, sino, por su relación con los demás píxeles que se encuentran en su entorno cercano, la principal dificultad radica en determinar qué significa cercano, ya que este puede ser el conjunto de los 8-vecinos del píxel o puede requerirse una cantidad mayor de píxeles alrededor para poder determinar la propiedad ser o no un punto de esquina. Los datos analizados en el modelo corresponden a entornos de 3x3, 5x5, 7x7 y 9x9, para los cuales el punto de esquina se ubica en el punto central en cada caso.

En la Figura 1, se presentan los entornos propuestos para el modelo de datos de entrada, en éste se propone analizar la información capturada en entornos de 3x3, 5x5, 7x7 y 9x9, los cuales se usarán para conformar los conjuntos de datos.



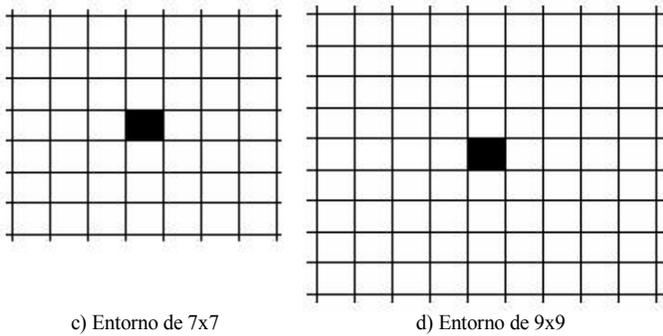


Figura 1. Entornos de captura de datos de entrada para el detector propuesto

Para cada uno de estos entornos, se evalúan tres modelos diferentes de datos, los cuales dependen de la información capturada en el entorno, el tratamiento previo realizado y la cantidad de parámetros extraídos. Las características de los modelos de datos usados para la determinación del detector son:

- Sólo se incluye la información correspondiente a la intensidad de tonos de gris.  $I(x, y)$
- Incluir tonos de gris y bordes, siendo estos bordes calculados previamente sobre toda la imagen, mediante el algoritmo de Canny [22].
- Incluir tonos de gris y gradientes (Los bordes se encuentran inmersos en los gradientes)  $I_x, I_y, I_{xx}, I_{yy}, I_{xy}$

El gradiente de primer orden  $I_x, I_y$ , en la dirección  $k$ , se determina mediante la siguiente ecuación:

$$I_k = I_{k-1} - I_k$$

Donde  $k$  representa la dirección  $x$  o  $y$

Los gradientes de segundo orden  $I_{xx}, I_{yy}, I_{xy}$  se obtienen aplicando el cálculo del gradiente de primer orden, en la dirección adecuada, a la imagen resultante al aplicar el gradiente de primer orden una vez.

Esta Información es utilizada para conformar un vector de características  $X$  que se aplicará a los modelos de red neuronal artificial plantados. El tamaño de dicho vector de características depende del modelo de datos usado, por tanto, se tendrán vectores de características de tamaño 9, 18, 27 y 36 para el caso de modelos de datos de 3x3, este tamaño depende no sólo del modelo usado, sino de la combinación de datos usados.

Las combinaciones de datos usadas son las siguientes:

Vector de características:

- [X1] Intensidades: Un dato por pixel, solo se incluye el tono de gris del pixel analizado.
- [X2] Bordes: Un dato por pixel, 1 si es un pixel de borde, 0 en caso contrario.

- [X3] d1: Un dato por pixel, el cual corresponde a la primera derivada en la dirección  $x$ .
- [X4] d2: Un dato por pixel, el cual corresponde a la primera derivada en la dirección  $y$ .
- [X5] d3: Un dato por pixel, el cual corresponde a la segunda derivada en la dirección  $x$ .
- [X6] d4: Un dato por pixel, el cual corresponde a la segunda derivada en la dirección  $y$ .
- [X7] d5: Un dato por pixel, correspondiente a la segunda derivada en las dirección  $x$  e  $y$ .
- [X8]  $I + B$ : Se incluyen los datos de intensidad y los de bordes, se tienen en total dos datos por pixel analizado.
- [X9]  $I + d1$ : Se incluyen los datos de intensidad y la primera derivada en dirección  $x$  ( $I_x$ ), se tienen un total de dos datos por pixel analizado.
- [X10]  $I + d2$ : Se incluyen los datos de intensidad y la primera derivada en dirección  $y$  ( $I_y$ ), se tienen un total de dos datos por pixel analizado.
- [X11]  $I+B+d1$ : Se incluyen los datos de intensidad, bordes y la primera derivada en dirección  $x$  ( $I_x$ ), se tienen un total de tres datos por pixel analizado.
- [X12]  $I+B+d2$ : Se incluyen los datos de intensidad, bordes y la primera derivada en dirección  $y$  ( $I_y$ ), se tienen un total de tres datos por pixel analizado.
- [X13]  $I+B+d1+d2$ : Se incluyen los datos de intensidad, bordes y la primera derivada en dirección  $x$  ( $I_x$ ), y en dirección  $y$  ( $I_y$ ), se tienen un total de cuatro datos por pixel analizado.

De esta forma, los vectores de características usados en las pruebas tienen las siguientes dimensiones.

Tabla 1. Número de datos en los vectores de características usados en las pruebas

VECTOR	DIMENSIÓN DE LA VENTANA DE DATOS			
	3X3	5X5	7X7	9X9
[X1]	9	25	49	81
[X2]	9	25	49	81
[X3]	9	25	49	81
[X4]	9	25	49	81
[X5]	9	25	49	81
[X6]	9	25	49	81
[X7]	9	25	49	81
[X8]	18	50	98	162
[X9]	18	50	98	162
[X10]	18	50	98	162
[X11]	27	75	147	243
[X12]	27	75	147	243
[X13]	36	100	196	324

**B. Modelos de red neuronal usados para la implementación del detector**

El detector propuesto se probó con tres modelos de red neuronal artificial, el modelo del perceptron multicapa (MLP),

redes de base radial (RBF) y red de Kohonen, para cada modelo se probaron diferentes topologías y se analizaron los resultados obtenidos, determinando el modelo que mejores características presenta para la implementación del detector.

**Perceptron Multicapa (MLP):** Se implementaron redes MLP para cada uno de los vectores de entrada definidos, con las características mostradas en la tabla 2.

Se construyeron, entrenaron y evaluaron un total de 192 MLP, después de entrenar las redes, se evaluó su rendimiento encontrando que las mejores respuestas se presentan para:

- MLP de 7x7 con 49 nodos en la capa de entrada, correspondientes a los datos de primera y segunda derivada.
- Una capa oculta.
- Los tiempos de respuesta son del orden de 120 ms
- Son muy sensibles al ruido
- No presentan un buen comportamiento en imágenes jpg
- La implementación presenta un alto costo computacional debido al cálculo de los valores de primera y segunda derivada.

Tabla 2: Topología de los MLP probados

RED - MPL	1	2	3	4	5	6
Capas ocultas	1	1	1	1	1	1
Nodos capa de entrada	49	49	49	49	49	49
Nodos capa oculta 1	35	32	29	27	30	22
Nodos Capa de Salida	7	7	7	7	7	7
Función de activación	Neta	Sigm	Tanh	Neta	Sigmo	Tanh

**Redes de base radial (RBF):** Se implementaron redes RBF para cada uno de los vectores de entrada definidos, con las siguientes características:

- Número de clases: 3, 5, 7 y 9
- Nodos de la capa de entrada igual a la cantidad de datos de los vectores de prueba
- Selección de centros mediante k-medias con inicialización de centros ideales.
- Cálculo de centros de clase con la distancia media a los dos más cercanos al centro de clase.

Se implementaron y evaluaron un total de 96 RBF, encontrando la mejor respuesta para:

- RBF de 7x7 con 49 nodos en la capa de entrada, correspondientes a los datos de primera y segunda derivada.
- Los tiempos de respuesta son del orden de 80 ms
- Mejor respuesta al ruido que para el caso de los MLP
- Mejora el comportamiento en imágenes jpg, respecto a los MLP, pero sigue siendo un factor crítico

- Presenta un alto costo computacional debido al cálculo de los valores de primera y segunda derivada, pero la complejidad algorítmica es menor que para el caso de los MLP
- La mejor respuesta se encontró para el caso de nueve clases, (Seis tipos de esquina, plano variable, borde y plano estable)

**Redes de Kohonen:** Se implementaron redes RBF para cada uno de los vectores de entrada definidos, con las siguientes características:

- Nodos de la capa de entrada igual a la cantidad de datos de los vectores de prueba
- Número de clases: Siete, una para cada tipo de esquina (L, T, Y, K, X, Vértice, No esquina)

Se implementaron y evaluaron un total de 13 redes de Kohonen, pero para ninguna de las topologías probadas, se encontró una respuesta adecuada, en el mejor de los casos se logró una eficiencia de reconocimiento de un 60%. Por esta razón, se descarta este modelo.

### III. EVALUACIÓN DE LOS MODELOS DE RED NEURONAL IMPLEMENTADOS

El programa desarrollado se encuentra orientado a la evaluación de detectores de esquinas, dentro del sistema se encuentran implementados los detectores de Plessey [23], Susan [24] y Sojka [25], los cuales se encuentran optimizados mediante rutinas de bajo nivel desarrolladas en Matlab.

El sistema realiza el proceso de detección de esquinas en imágenes digitales, mediante la implementación de los algoritmos de Plessey, Susan y Sojka, reportando la ubicación estos, mediante la ubicación de píxeles de diferentes colores para cada detector, en la posición del punto de esquina identificado; adicionalmente, se reporta las coordenadas de cada uno de los puntos de esquina y los tiempos empleados en la detección.

Estos datos son comparados con los resultados obtenidos por el detecto de esquinas desarrollado mediante el modelo de la MLP y RBF para datos de entornos de 7x7 usando los datos de primera y segunda derivada.

En todos los casos la detección por medio de la RNA fue de menor calidad, encontrándose que ese modelo emplea más tiempo que los otros tres algoritmos usados en la comparación, adicionalmente, la eficiencia del algoritmo basado en redes MLP y RBF fueron mucho menor que la detectada por los algoritmos de comparación, pero el modelo desarrollado presenta una cualidad por encima de los otros modelos, en éste se detecta el tipo de esquina, lo cual es fundamental en procesos de reconstrucción tridimensional, ya que esta información permite realizar con más facilidad los procesos de apareamiento de puntos en la imagen de prueba usada.

En la Figura 2 se presenta la evaluación general de los resultados de comparación entre los modelos de Plessey, Susan y Sojka, con el modelo de red neuronal RBF, MLP y Kohonen desarrolladas.

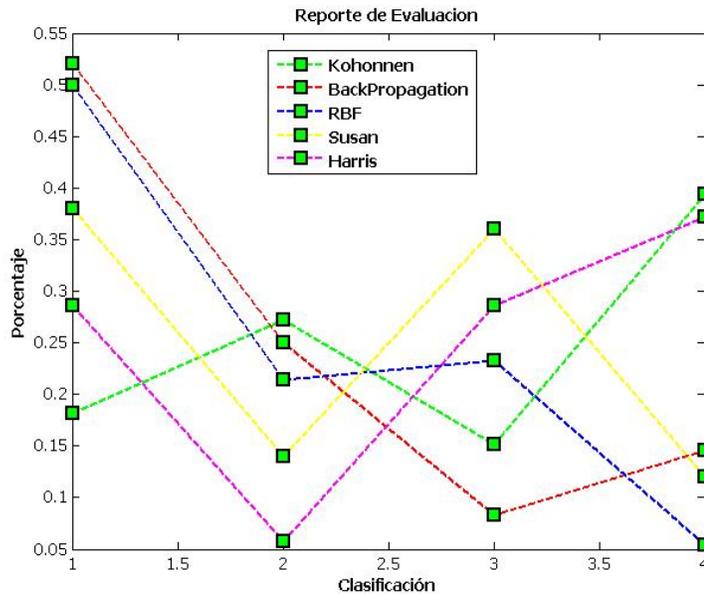


Figura 2. Comparación de los detectores de Harris, Susan, MLP, Kohonen y RBF.

#### IV. CONCLUSIONES

La detección de puntos de esquina mediante redes neuronales artificiales es una alternativa viable, la cual puede entregar resultados de mayor precisión que muchos de los detectores actuales.

La principal dificultad al desarrollar un detector basado en redes neuronales artificiales, radica en la necesidad de probar múltiples modelos de redes, conjuntos de datos y técnicas de entrenamiento. En el proceso de construcción del modelo se deben probar y evaluar múltiples alternativas.

La arquitectura de redes neuronales artificiales que presenta una mejor respuesta al caso de detección de esquinas es la MLP y RBF, pero deberán explorarse otro tipo de arquitecturas y evaluar en forma sistemática las respuestas, para poder identificar la arquitectura con una mejor respuesta.

Los detectores de esquina usados como comparación (Susan, Sojka y Harris), presentan una respuesta más rápida y eficiente en los procesos de detección de esquinas que los modelos propuestos, esto es debido a:

- Los métodos tradicionales se encuentran implementados mediante librerías de bajo nivel, las cuales se encuentran optimizadas y operan más rápido

- No se realiza un proceso de clasificación de la esquina, solo se limita a identificar puntos en los cuales se encuentran.
- El modelo propuesto no se enfoca solo en la detección, adicionalmente se busca realizar un proceso de clasificación, lo cual es fundamental en los procesos de reconstrucción 3D.

Para realizar procesos de reconstrucción 3D, no solo basta la información sobre la nube de puntos pertenecientes al objeto, además se puede realizar procesos de reconstrucción con mayor eficiencia conociendo puntos característicos como las esquinas y la información sobre la morfología de los mismos.

#### REFERENCIAS

- [1] Russ, J. C., 1995. The image processing handbook CRC Press, Inc., second edition.
- [2] Faugeras, O., 1999. Three-Dimensional Computer Vision: a geometric Viewpoint. MIT Press, Cambridge, Massachusetts, third printing.
- [3] Faugeras, O., and Luong, Q.T., 2001. The Geometry of Multiple Images, The MIT Press, Cambridge Massachusetts, London, England.
- [4] Deriche, R., and Giraudon, G., 1990. Accurate corner detection: An analytical study. In Proc. 3rd Int. Conf. on Computer Vision, pp. 66 - 70.
- [5] Deriche, R.; Giraudon, G., 1990. Accurate corner detection: an analytical study Computer Vision. Proceedings, Third International Conference. pp.66 - 70
- [6] Deriche, R., and Giraudon, G., 1993. A Computational Approach for Corner and Vertex Detection. International Journal of Computer Vision. 10(2), pp. 101-124. Kluwer Academic Publishers.
- [7] Dreschler, L., and Nagel, H. H., 1981. Volumetric Model and 3D Trajectory of a Moving Car Derived from Monocular TV-frame Sequence of a Street Scene. Computer Vision, Graphics and Image Processing, 20(3), pp. 199 - 228.
- [8] Dreschler, L., and Nagel, H. H., 1982. On the Selection of Critical Points and Local Curvature Extrema of Region Boundaries for Interframe Matching. In Proc. Of the International conference on Pattern Recognition. pp. 542-544.
- [9] Hernández, B. y Olague, G., 2001. Un detector sub-pixel paramétrico de esquina múltiples. En memorias del VI Taller Iberoamericano de Reconocimiento de Patrones. México D.F. pp. 37-49.
- [10] Rohr, K., 1992. Recognizing Corners by Fitting Parametric Models. International Journal of Computer Vision. 9(3), pp. 213-230. Kluwer Academic Press.
- [11] Rohr, K., 1992. Modelling and identification of characteristic intensity variations. Image and Vision Computing, 10(2): pp. 66 - 76.
- [12] Rosin, P., 1996. Augmenting Corner Descriptors. Graphical Models and Image Processing, 58(3), pp. 286-294.
- [13] Dias, P.G.T.; Kassim, A.A. y Srinivasan, V., 1995. A neural network based corner detection method Neural Networks. Proceedings., IEEE International Conference on Volume 4, No. 27, pp. 2116 - 2120.
- [14] Wan-Ching C.; Rockett, P., 1997. Bayesian labelling of corners using a grey-level corner image model Image Processing. Proceedings., International Conference on Vol. 1, pp. 687 - 690.
- [15] Harvey A. Cohen Craig McKinnon, and J. You, Neural-Fuzzy Feature Detectors, DICTA-97, Auckland, N.Z., Dec 10-12, pp 479-484.
- [16] Chen K. and Wang D.L., 1999. Image segmentation based on a dynamically coupled neural oscillator network. Proceedings of IJCNN-99.
- [17] Liu X. and Wang D.L., 1999. Range image segmentation using a LEGION network. IEEE Transactions on Neural Networks, vol. 10, pp. 564-573.
- [18] Chen K. and Wang D.L., 2002. A dynamically coupled neural oscillator network for image segmentation. Neural Networks, vol. 15, pp. 423-439.
- [19] Chen Y., Sheng M., and He Y., 1993. A method of pattern recognition based upon synthetic technology of fuzzy logic and neural network tencon '93. Proceedings. Computer, Communication, Control and Power Engineering.

- IEEE Region 10 Conference on Vol.2 Issue 0, pp. 815 - 818
- [20] Kwanghoon S., 1996. Recognition Of Partially Occluded Target Objects Image Processing. Proceedings., International Conference on Volume 3, pp 595 – 598.
- [21] Egmont-Petersen, M., De Ridder, D. and Handels. H., 2002. Image processing with neural networks - a review. Pattern Recognition, Vol. 35, No. 10, pp. 2279-2301.
- [22] Canny, J., 1986. A computational approach to edge detection. En: IEEE Pattern Analysis and Machine Intelligence. Vol. 8, No. 6, pp. 679-698.
- [23] Harris C. and Stephens M., 1988. A Combined Corner and Edge Detector. Proc. Alvey Vision Conf., Univ. Manchester, pp. 147-151.
- [24] Smith S.M. and Brady M., 1997. SUSAN - A New Approach to Low Level Image Processing. International Journal of Computer Vision, Vol. 23(1), pp. 45-78.
- [25] Sojka, E., 2004. A New Algorithm for Detecting Corners in Digital Images. 2004.12.18, <http://www.cs.vsb.cz/sojka/cordet/presentation.html>