Journal of Engineering and Sustainable Development
www.jeasd.org
Vol. 20, No.06, November 2016
ISSN 2520-0917

A NOVEL EDGE DETECTION METHOD USING K-MEANS CLUSTERING

*Dr. Walaa Mohammed Khalaf [1], Dr. Kadhum Al-Majdi [1], Noor Hashim Hamed [2]

1) Assist Prof., Computer & Software Engineering Department, University of Almustansiriayah/ Baghdad, Iraq.
2) Assist Lect., Building and Construction Department, University of Technology/Baghdad, Iraq.

Abstract: In this paper, a new approach is introduced to detect the edges of any kind of gray scale images by using k-means clustering algorithm, where three novel features are proposed by taking the advantages of the similarity of image pixel with its eight surrounding neighbors through feeding these features as attributes to the clustering system. This method of edge detection does not use neither any smoothing filter nor threshold values. The experimental results show that an acceptable detection of the edges is done.

Keywords: *edge detection, clustering, image processing.*

طريقة جديدة للكشف عن الحافات باستخدام تجميع k-means

الخلاصة: في هذا البحث تم تقديم طريقة جديدة للكشف عن الحافات في اي نوع من انواع الصور ذات التدرج الرمادي بأعتماد خوارزمية تجميع (ك) من المعدلات، حيث تم افتراض ثلاثة ميزات جديدة بالاستفادة من التشابه ما بين البيكسل و متجاوراتها الثمانية المحيطة بها و من ثم تزويد تلك الميزات الخاصة بكل بيكسل كصفات الى نظام التجميع. هذه الطريقة الخاصة بأكتشاف الحافات للصور لا تستخدم اي مرشح تجانسي ولا اي قيمة للعتبة. النتائج التجريبية اظهرت بأن كشف مقبول للحواف قد تم من خلال استخدام هذه الطريقة.

1. Introduction

Edges are defined typically as the local changes of intensity in an image, these changes are classified as geometric events like object boundary and surfaces, and non-geometric events like specularity, shadows, and inter-reflect.

Edge detection is an important component of many vision systems, including object detectors which have remained a critical task in computer vision like pattern recognition, segmentation, and active contour. Therefore; edge detection has been applied in many fields like, optical character recognition [1], automatic target recognition [2], and medical image applications [3].

**Corresponding Author* walaakhalaf@yahoo.com

Many operators have been introduced in edge detection, and the most famous and applicable are Sobel and Prewitt operators [4, 5]. First derivatives and second derivatives are used to detect edges, which are called gradient and Laplacien respectively. Gradient based edge detection methods have been used two-linear filters for processing horizontal edges and vertical edges individually for approximating the first-order derivative of each pixel value in the image. Laplacian is considered more sensitive to noise because it uses second derivatives. To get rid this problem, Canny edge detector and LoG algorithm were introduced, which merge smoothing functions with zero crossing detection technique [6, 7]. LoG algorithm used Gaussian and Laplacian differential operator for smoothing the image and detect the edge respectively.

Most edge detection algorithms have been introduced and developed based on the computation of intensity differences of image. In this article a new proposed heuristic approach is introduced, which tries to detect edges of an image most easily. The key feature of the proposed approach which is differentiated from others is the use of K-means clustering algorithm to get the edges of the image, as well as the use of novel artificial features for the image as a feature domain of the K-means clustering algorithm.

This paper is organized as follows. The proposed method and novel artificial features is described in Section 2. The use of k-means clustering algorithm for edge detection is presented in Section 3. While experimental results are introduced in Section 4. Conclusions are presented in Section 5.

2. Proposed Method

This section includes a rich description of the new three proposed features of the image pixel. After that, these 3 features are fed to the K-means algorithm for clustering.

2.1. Pixel Features

In this paper the work will be done with a 3x3 matrix of image pixels to get good information about the central pixel comparing it with its surrounding neighbors.

Each pixel[i,j] (p_1), has 8 neighbors, [i-1,j-1] (p_9), [i-1,j] (p_2),…, [i+1,j+1] (p_5), as shown in "Fig .1".

Pixel[i-1,j-1] p_9	Pixel[i-1,j] p_2	Pixel[i-1,j+1] p_3
Pixel[i, j-1] p_8	Pixel[i,j] $\mathbf{p_1}$	Pixel[i,j+1] p_4
Pixel[i+1,j-1] p_7	Pixel[i+1,j] p_6	Pixel[i+1,j+1] p_5

Figure 1: pixel definition

This 3x3 neighborhood matrix will be used to define the three features for each pixel. Using these attributes via k-means clustering algorithm to detect edge and nonedge pixels for the tested image. The new algorithm will be applied for a gray scale images which has a color intensity f(i,j) for each pixel[i,j] ranging from 0-255.

Feature 1

Some statistical properties like standard deviation gives meaningful information about pixel, which measure the distance between the data point and the average of the these data points, small value signify small variation in grayness, while large value signify big variation in grayness [8]. Therefore a pixel p_1 with high standard deviation value measured with respect to its 8 neighbors will be candidate as an edge pixel.

$$\mu = \frac{1}{9}\sum_{k=1}^{9} p_k \quad (1)$$

$$\sigma = \sqrt{[\frac{1}{9}\sum_{k=1}^{9}(p_k - \mu)^2]} \quad (2)$$

Where μ represents mean value (average), and σ represents standard deviation value.

In this article a new value is defined, which measures the differences not with respect to the mean as in standard deviation but will measure the differences with respect to central pixel p_1.

$$\sigma_{new} = \sqrt{[\frac{1}{8}\sum_{k=2}^{9}(p_1 - p_k)^2]} \quad (3)$$

If some pixels' values are close to the central pixel (p_1) value, then the calculated variable in equation 3 (σ_{new}) will be small, therefore this pixel (p_1) will not be candidate as edge pixel; if some pixels' values are far away from the central pixel (p_1) value, then the calculated variable in equation 3 (σ_{new}) will be large, therefore this pixel (p_1) will be candidate as edge pixel. If all data values are equal (have same value), then the calculated variable in equation 3 (σ_{new}) will equal to zero.

Feature 2

Mean value represents the average of all nine pixels in 3x3 mask, as in equation 1, but in this article a new parameter is invented which measures the difference between the average of the surrounding pixels (the 8 neighbors) and the central pixel (p_1), as in equation 4

$$d_{new} = \left|\left[\frac{1}{8}\sum_{k=2}^{9} p_k\right] - p_1\right| \quad (4)$$

If the average value of the surrounding pixels is near the value of the central pixel (p_1), d_{new} is small and this pixel will be candidate as nonedge pixel, while if the average value of the surrounding pixels is far away from the central pixel (p_1), d_{new} is large and this pixel will be candidate as edge pixel.

Feature 3

Creating four elements vector called $\boldsymbol{D}$, contains the absolute differences between the central pixel p_1 and the four main directions neighbors pixels (north (p_2), east (p_4), south (p_6), and west (p_8)), $\boldsymbol{D} = [\ |p_1 - p_2|, |p_1 - p_4|, |p_1 - p_6|, |p_1 - p_8|\ ]$.
Then calculating the norm value (L_2_norm) of this vector as follows:

$$\boldsymbol{D}_{\mathbf{norm}} = \|\boldsymbol{D}\|_2 = \sqrt[2]{\sum_{k=1}^{4} D_k^2} \qquad (5)$$

If $\boldsymbol{D}_{\mathbf{norm}}$ is small, this pixel will be candidate as nonedge pixel, while if $\boldsymbol{D}_{\mathbf{norm}}$ is large, this pixel will be candidate as edge pixel.

3. K-Means Clustering

After calculating these three features for each pixel, a vector of length $\boldsymbol{s}$ ($\boldsymbol{s} = m$xn) and width 3 will be creating as follows:

	Feature 1	**Feature 2**	**Feature 3**
$Pixel_1$	P_{11}	P_{12}	P_{13}
$Pixel_2$	P_{21}	P_{22}	P_{23}
….	…	…	…
$Pixel_s$	P_{s1}	P_{s2}	P_{s3}

Sending these $\boldsymbol{s}$ vectors to K-means clustering algorithm, by considering $\boldsymbol{k} = 2$, cluster 1 will be defined for edge pixels while cluster 2 for nonedge pixels.

Clustering is a series of operations for dividing data points group into smaller groups called clusters, in this article the number of clusters will be two, first one will be defined for edge pixels while the second one for nonedge pixels. In general, there are $\boldsymbol{s}$ data points ($\boldsymbol{s}$ equals the multiplication result of m and n) $Pixel_i$, i=1,2,...$\boldsymbol{s}$ that have to be separated into k (k = 2 in the proposed method) clusters.

In this article, the main target is to allocate a class for each data point in this vector. K-means clustering method is used to find the positions λ_i, i = 1, 2 of each cluster which minimize the square of the *distance* (d) between data points and the cluster [9]. K-means clustering method will solve the following:

$$\arg min_c \sum_{i=1}^{k} \sum_{x \in c_i} d(x, \lambda_i)^2 = \arg min_c \sum_{i=1}^{k} \sum_{x \in c_i} \|x - \lambda_i\|_2^2 \qquad (6)$$

where c_i represents the set of data points which belong to cluster i. K-means clustering method uses the Euclidean distance for measuring the distance which is defined as the following $d(x,\lambda_i) = \|x - \lambda_i\|_2$. Equation 6 can't be considered as trivial (in fact it is NP-hard), therefore; K-means algorithm only desires to find the global minimum solution. Fig.2 explains the algorithm of K-means clustering, while Fig.3 explains the whole proposed edge detection algorithm.

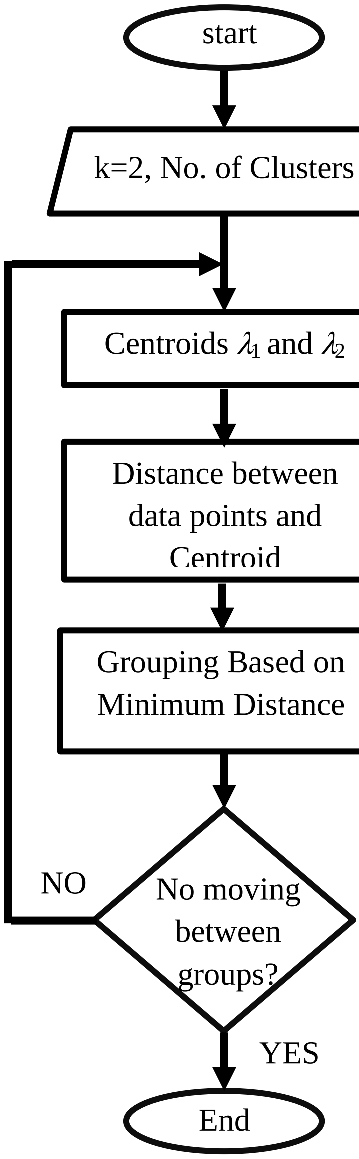


Figure 2. K means clustering algorithm

4. Experimental Results

The performance of the proposed method is tested on some test images which are shown in "Fig .4". These images have various types of edges and different gradients of edge complexity**.**The results are seems good and acceptable as well as tried to find most true edges in that images as shown in 'Fig .5". For the purpose of comparison, standard edge detection methods like Canny and Sobel are also applied on the same test images, by using standard library of image processing toolbox of matlab, the results are shown in "Fig. 6" and "Fig. 7".

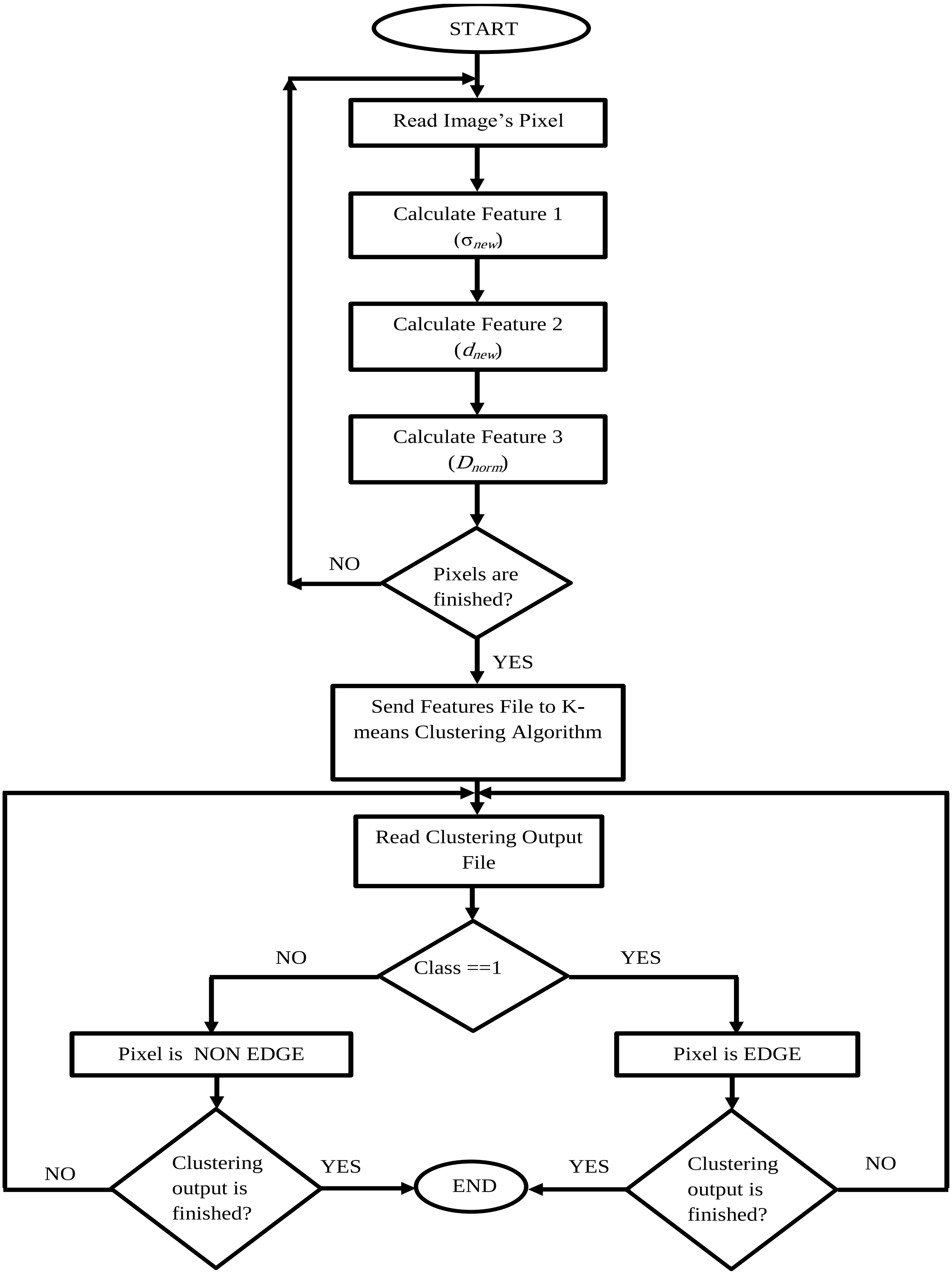
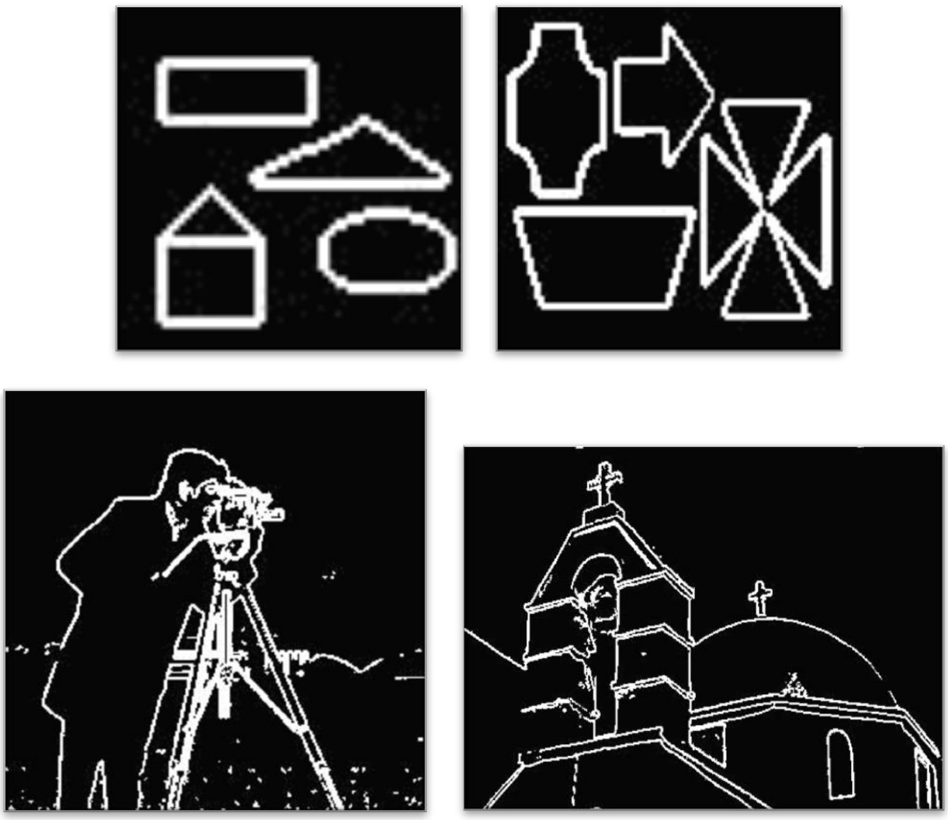


Figure 3. Proposed edge detection algorithm.

Figure 4. Test Images



\

Figure 5. Results of proposed method

The advantage of the proposed method that it does not use any smoothing filter for the preprocessing stage as well as does not use any threshold value as always have been used in standard edge detection methods.

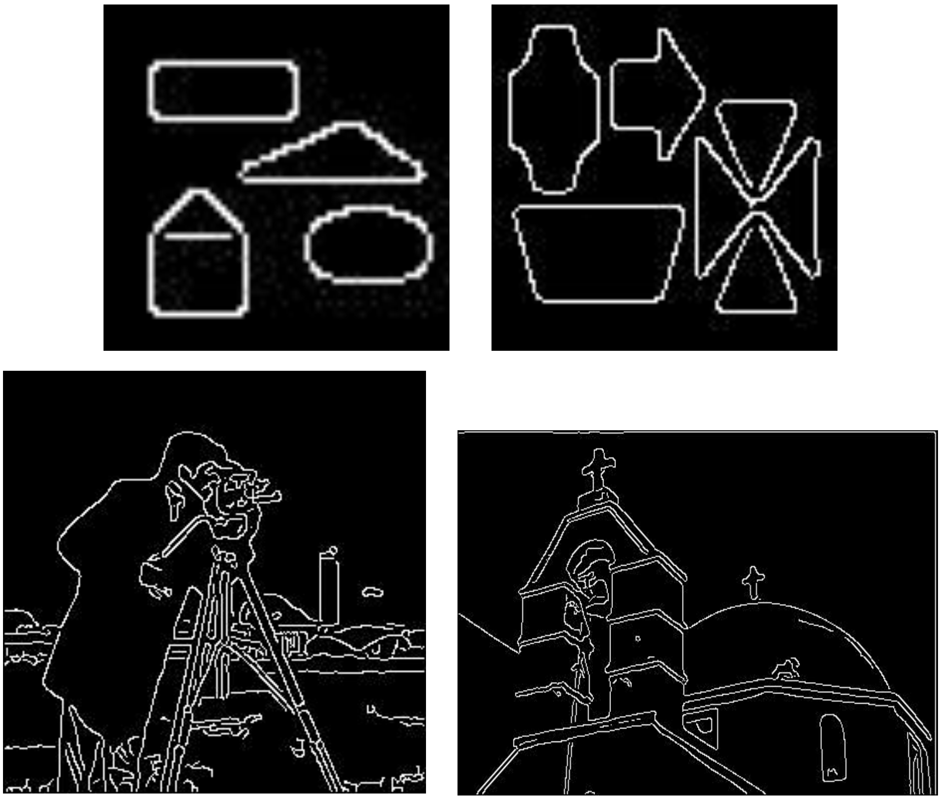


Figure 6. Results of Canny method

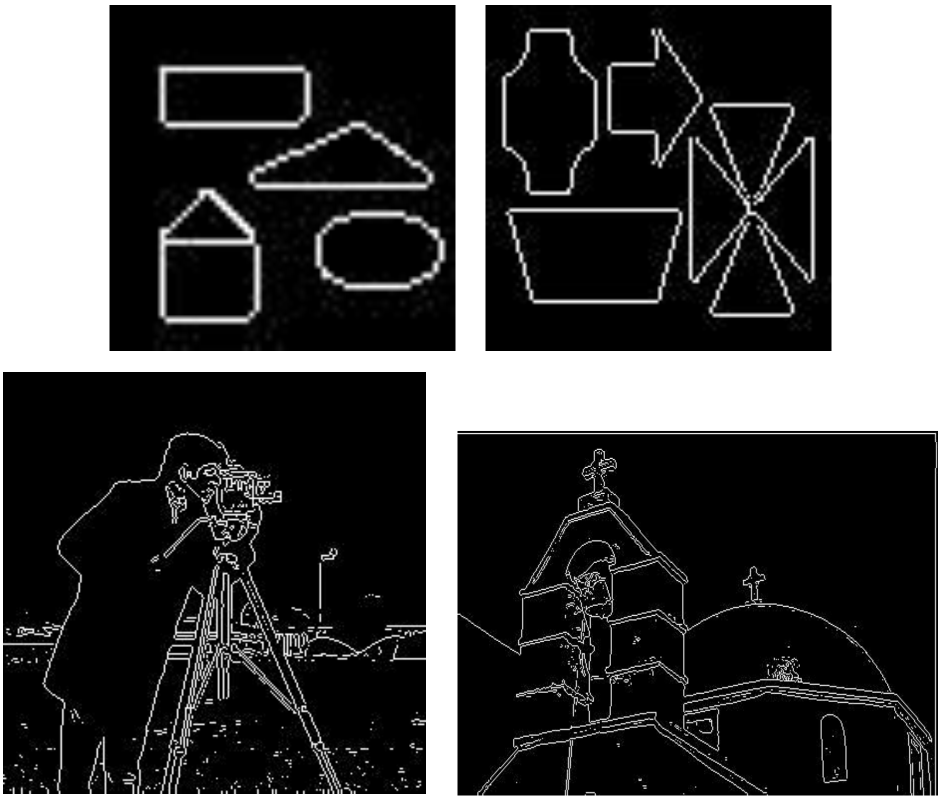


Figure 7. Results of Sobel method

5. Conclusions

Most of the applied methods in edge detection field start with a smoothing filter for removing the noise and making edges more blurred, and also are based on calculating the gradient and trying different threshold values to get better results.

In this article a new attributes are proposed which then applied to a clustering system for taking the decision on each pixel to be labeled as edge or nonedge pixel. These three new proposed attributes are depending on the relationship between each pixel and its eight neighbors, from these attributes a good indication is given for the clustering algorithm to take the decision for defining each pixel as edge or nonedge pixel.

The algorithm is applied for different types of gray scale images and acceptable results are gotten, so it can be considered as a good and easy tool in pattern recognition systems for expert robotic system, and other fields.

References

1. J. Lzaro, J. L. Martn, J. Arias, A. Astarloa, C. Cuadrado, (2010)." *Neuro semantic thresholding using OCR software for high precision OCR applications*". Image Vision Comput., 28, pp. 571–578,
2. G. C. Anagnostopoulos, (2009),"*SVM-based target recognition from synthetic aperture radar images using target region outline descriptors*". Nonlinear Analysis-Theor. Meth. App., 71, pp. 2934–2939 .
3. M. T. Doelken, H. Stefan, E. Pauli, A. Stadlbauer, T. Struffert, T. Engelhorn, G. Richter, O. Ganslandt, A. Doerfler, T. Hammen, (2008). "*1H-MRS profile in MRI positive versus MRI negative patients with temporal lobe epilepsy*". Seizure, **17**, pp. 490–497.
4. M. Basu, (1994)"*A Gaussian derivative model for edge enhancement.*", Pattern Recognition, 27, pp. 1451-1461 .
5. C. Kang, and W. Wang, (2007)."*A novel edge detection method based on the maximizing objective function.*", Pattern Recognition, 40, pp. 609-618
6. J. F. Canny, "*Computational approach to edge detection,* 1986". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, pp. 679{698,.
7. L. S. Davis., "*A Survey of Edge Detection Techniques.* 1995," Computer Graphics Image Processing. vol. 4, no. 3, pp. 243–270, Mar..
8. D. Ganguly, S. Mukherjee, K. Mitra, and P. Mukherjee, 2009 "*A novel approach for edge detection of images*". in Proceedings of the International Conference on Computer and Automation Engineering (ICCAE'09), pp. 49-53,.
9. D. MacKay, *2003"An Example Inference Task: Clustering Algorithms". Cambridge University Pres. pp. 284–292.*