

Isolating Healthy Bananas from Unhealthy Ones Based on Feature Extraction and Clustering Method Using Neural Network

Meysam Siyah Mansoori (Corresponding author)

Department of Computer Science and Electrical Engineering, Islamic Azad University

Majlesi Branch, Isfahan, Iran

E-mail: Meysam_bme@yahoo.com

Hamidreza Fardad

Department of Computer Science and Electrical Engineering, Islamic Azad University

Majlesi Branch, Isfahan, Iran

E-mail: Hfardad@iaumajlesi.ac.ir

Reza Enteshari

Department of Computer Science and Electrical Engineering, Islamic Azad University

Majlesi Branch, Isfahan, Iran

E-mail: Renteshari@iaumajlesi.ac.ir

Yaser Siah Mansouri

Agro Technology Department, College of Abureyhan, University of Tehran

Tehran, Iran

E-mail: ymansouri@ut.ac.ir

Abstract

Due to the high sorting speed required during inspection and classification in packing lines, most of the current automatic systems, based on machine vision, are used. Fruit industries are not excluded about this fact. In this paper a method is proposed for detection healthy bananas and defective one. Our algorithm has 4 steps.

First, we eliminated background using segmentation methods such as FCM, HCM, Kmeans. Then we extracted the boundaries of a sample banana using edge detection approach. After that, feature from surface of a sample was extracted. Finally, by using a neural network, healthy bananas and defective one was detected.

Keywords: Banana, Clustering Method, Feature extraction, Neural network

1. Introduction

Agriculture products are being more demanding in market today. In order to increase its productivity, speed and eliminating human errors due to hard conditions, automation to produce these products will be very helpful(S. A. R. A. B. M.M.Mokji, 2006).

Using image processing technology and machine vision in recent decade had much progress in minor and major industries. At the present, human operator does detecting defections in factory. As intelligence level of human operator is decreased gradually, by difficult conditions of working environment such as air pollution, sound pollution. It is probable that some surface defection is hidden from human operator's eye(M.S.Mansoori, H.tajik, G.Mohhamadi, M.Pashna 2008), so developing an image processing system seems necessary.

Segmentation is frequently used to separate significant regions from background. Hence, accurate image segmentation is the primary component in many image-processing applications.

In order to segment an object many algorithms has been developed such as kmeans (P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. 2002),Hard C Means (HCM)(J.C Dunn. 1983),Fuzzy C Means (FCM) (J.C. Bezdek and S. Pal. 1991), Neural Gas (NG) (B. Hammer, M. Strickert, and T. Villmann. 2005) and Maximum Entropy(ME)(S. Pakhomov. 2002).There is a wide range of feature extraction and analysis techniques that are used with different

criteria for feature extraction: statistical methods (grey level co-occurrence matrix (Sonka M, 1998), semivariogram analysis (O. S. Junior, A. C. Silva and Z. Abdelouah. 2009)); filter techniques (energy filters (A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, 1987), Gabor filters (S. E. Grigorescu, N. Petkov, and P. Kruizinga, 2002)); or the most recent techniques based on wavelet decomposition (S. Arivazhagan and L. Ganesan. 2003).

2. Method

2.1 Image acquisitions

This section describes how corresponding digital images are acquired for the conducted experiments. The imaging system consisted of a digital camera with silicon CCD detector connected to a lens assembly. The camera was attached to a computer with MATLAB software capable of capturing image.

In order to use computer vision to evaluate objects by proposed algorithm, it is important to create a uniform light field around the object. To achieve this and additionally to eliminate any effect of environmental light, a cubic box of $0.80 \times 0.80 \times 0.80$ m with white reflecting walls was constructed.

The bananas were rotated around the longitudinal axis by the rollers. The rollers and the bottom of the frame were colored in blue in order to simplify the separation of the banana from its background. Above the set-up, a digital camera was placed and directed to the frame. This camera acquires images taken in the red-green-blue color space (640×480 pixels) of the banana at 30 frames per second.

2.2 Proceeding of image analysis

After image acquisition, images are given to MATLAB software as an input image. Then this image, which is color image, is transformed to gray level image by the following formula,

$$I = (R + G + B) / 3 \quad (1)$$

In this formula R is the red component of a sample image and G is the green component and finally B is the blue component. The used method in this article has 4 steps.

- 1) Background elimination using clustering method
- 2) Finding the edge of a sample using SUSAN edge detector
- 3) Feature extraction using texture analysis
- 4) Applying neural network as a classifier.

2.3 Background Elimination

Among of the clustering method such as K-Means (O. S. Junior, A. C. Silva and Z. Abdelouah. 2009), Hard C-Means (J.C Dunn. 1983), Neural Gas (B. Hammer, M. Strickert, and T. Villmann. 2005), Maximum Entropy (S. Pakhomov. 2002) and Fuzzy C-Means (J.C. Bezdek and S. Pal. 1991), we used FCM algorithm for background elimination. The Fuzzy C-Mean (FCM) clustering algorithm was first introduced by Dunn and later extended by Bezdek (J.C. Bezdek and S. Pal. 1991). In this work, the FCM algorithm has been applied to perform the classification. This algorithm is an iterative clustering method and presents some advantages with respect to other classifiers.

The fuzzy c-means (FCM) algorithm is one of the most widely used methods in fuzzy clustering. It is based on the concept of fuzzy c-partition, summarized as follows,

Let $X = \{x_1, \dots, x_n\}$ be a set of given data, where each data point $x_k (k = 1, \dots, n)$ is a vector in R^p , U_{cn} be a set of real $c \times n$ matrices and c be an integer, $2 \leq c < n$.

Then the fuzzy c-partition and space for X is the set

$$M_{fcn} = \{U \in U_{cn} : u_{ik} \in [0, 1], \sum_{i=1}^c u_{ik} = 1, 0 < \sum_{k=1}^n u_{ik} < n\} \quad (2)$$

where u_{ik} is a membership value of x_k in cluster $i (i = 1, \dots, c)$.

The aim of the FCM algorithm is to find an optimal fuzzy c-partition and corresponding prototypes minimizing the objective function

$$J_m(U, V; X) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m \|x_k - v_i\|^2 \quad (3)$$

In the above equation $V = (v_1, v_2, \dots, v_c)$ is a matrix of unknown cluster centers (prototypes) $v_i \in R^p$, $\|\dots\|$ is Euclidean norm, and the weighting exponent m in $[1, \infty]$ is a constant that influences the membership values.

To minimize criterion J_m , under the fuzzy constraints defined on (3), the FCM algorithm is defined as an alternating minimization algorithm, as follows. Choose a value for c, m and ε as a small positive constant; then generate randomly a fuzzy c -partition U^0 and set iteration number $t=0$. A two-step iterative process works as follows, given the membership value $U_{ik}^{(t)}$, the cluster centers $V_i^{(t)} (i=1, \dots, c)$ are calculated by

$$V_i^{(t)} = \frac{\sum_{k=1}^n (u_{ik}^{(t)})^m x_k}{\sum_{k=1}^n (u_{ik}^{(t)})^m} \quad (4)$$

given the new clusters $V_i^{(t)}$, update membership values u_{ik}^t ;

$$u_{ik}^{(t+1)} = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i^{(t)}\|^2}{\|x_k - v_j^{(t)}\|^2} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (5)$$

The process stops when $|U^{(t+1)} - U^t| < \varepsilon$ or predefined number of iteration is reached.

2.4 Finding the Edge of Banana

The *Smith* edge detector (S.M. Smith and J.M. Brady, 1997) is different from the other detectors in nature. Each pixel in an image is used as the center of a small circular mask. The grayscale values of all the pixels within this circular mask are compared with that of the center pixel (the *nucleus*). All pixels with *similar* brightness to that of the nucleus are assumed to be part of the same structure in the image.

Figure 1 shows the masks with pixels of *similar* brightness to the nucleus colored black, and pixels with *different* brightness colored white. Smith calls the black image area the *Univalue Segment Assimilating Nucleus* (USAN)

In SUSAN algorithm, the mask (see figure 2) is placed at each point of the image and, for each point, the brightness of each pixel within the mask is compared with that of the nucleus (the center point), originally a simple equation determined this comparison,

$$c(\vec{r}, \vec{r}_0) = \begin{cases} 1 & \Rightarrow \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq t \\ 0 & \Rightarrow \text{if } |I(\vec{r}) - I(\vec{r}_0)| > t \end{cases} \quad (6)$$

Where $\vec{r}_0$ is the position of the nucleus in the two dimensional image, $\vec{r}$ is the position of any other point within the mask, $I(\vec{r})$ is the brightness of any pixel, t is the brightness difference threshold and C is the output of the comparison. This comparison is done for each pixel within the mask and a running total, n , of outputs (C) is made,

$$n(\vec{r}_0) = \sum_{\vec{r}} c(\vec{r}, \vec{r}_0) \quad (7)$$

This total n is just the number of pixels in the USAN, i.e. it gives the USAN's area.

For an edge to be detected, a local minimum in n must be found,

$$\frac{d_n(\vec{r}_0)}{d_{\vec{r}_0}} = 0 \quad (8)$$

That is, differentiating with respect to both components of $\vec{r}_0$ must give zero. (n) must also be below the geometric threshold. Thus, using equation 7 and 8,

$$\frac{d}{d_{r_0}} \int_{\text{over}} \int_{\text{mask}} e^{-\left(\frac{I(\vec{r})-I(\vec{r}_0)}{t}\right)^6} dxdy = 0 \quad (9)$$

Where $dxdy$ is the elemental area within the mask. (The summation has been replaced by a double integral.) This gives on differentiation,

$$\int_{\text{over}} \int_{\text{mask}} B(\vec{r}, \vec{r}_0) \left(\frac{d_I}{d_{r_0}}(\vec{r}) - \frac{d_I}{d_{r_0}}(\vec{r}_0) \right) dxdy = 0 \quad (10)$$

With B given by

$$B(\vec{r}, \vec{r}_0) = \left(\frac{I(\vec{r}) - I(\vec{r}_0)}{t} \right)^5 - e^{-\left(\frac{I(\vec{r}) - I(\vec{r}_0)}{t}\right)^6} \quad (11)$$

this function B has the property of picking out the narrow bands that lie on boundaries between the univalve regions; it is zero everywhere apart from fairly sharp peaks centered at $I(\vec{r}) = I(\vec{r}_0) \pm t$.

2.5 Feature Extraction

For extracting feature, we used properties of a Gray-Level Co-occurrence Matrix (GLCM) (Sonka M, 1998). GLCM is the matrix that defined by calculating how often a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j (Figure 3).

We explain some of these features (Haralick RM, Shapiro LG. 1993) as follows:

Entropy: a term more commonly found in thermodynamics or statistical mechanics. Entropy is a measure of the level of disorder in a system. Image of highly homogeneous scenes have high associated entropy, while in homogenous scenes poses a low entropy measure. The GLCM entropy is obtained with the following expression.

$$H = 1 - \frac{1}{Nc \cdot \ln(nc)} \sum_i \sum_j M(i, j) \cdot \ln(M(i, j)) \cdot 1_{M(i, j)} \quad (12)$$

Where $1_{M(i, j)} = 0$ when $M(i, j) = 0$ and 1 otherwise.

Contrast: it is a measure of the difference between intensity values of the neighboring pixels. It will favor contributions from pixels located away from the diagonal of $M(i, j)$

$$C = \frac{1}{Nc(L-1)^2} \sum_k^{L-1} k^2 \sum_{|i-j|=k} M(i, j) \quad (13)$$

Where Nc are the number of occurrences and L is the number of gray level variations.

Homogeneity: this gives a measure of how uniformly a given region is structured, with respect to its gray level variations.

$$Ho = \frac{1}{Nc^2} \sum_i \sum_j M(i, j)^2 \quad (14)$$

Local homogeneity: this measure provides the homogeneity of the image using a weight factor which gives small value for none-homogenous images when $i \neq j$.

$$G = \frac{1}{Nc} \sum_i \sum_j \frac{M(i, j)}{1 + (i - j)^2} \quad (15)$$

Directivity: this measure provides bigger value when two pixel regions with same gray values are separated by a translation.

$$D = \frac{1}{Nc} \sum_i \sum_j M(i, j)^2 \quad (16)$$

Uniformity: this a measure of the uniformity of each gray level

$$U = \frac{1}{Nc^2} \sum_i M(i, i)^2 \quad (17)$$

Moments: moments express common statistical information, such as the variance for that corresponds to second moment. this descriptor increase when the majority of the value of $M(i, j)$ are not on diagonal.

$$Mom_k = \sum_i \sum_j (i - j)^k M(i, j) \quad (18)$$

Inverse moments: this produce the opposite effect compared to previous descriptor

$$Mom_k^{-1} = \sum_i \sum_j \frac{M(i, j)}{(i - j)^k} \quad (19)$$

Maximum probability: considering the GLCM as an approximation of joint probability density between pixels, this operator extracts the most probable difference between pixel gray scale values

$$MAX = \max(M(i, j)) \quad (20)$$

Correlation: this is measure of gray scale linear dependencies between pixels at the specified positions relative to each other.

$$S = \frac{1}{Nc\sigma_x\sigma_y} \left| \sum_{ii} \sum_j (i - m_x)(j - m_y)M(i, j) \right| \quad (21)$$

2.6 Neural Network Classifier

The perceptron is the simplest form of a Neural Network used for the classification of a special type of patterns said to be linearly separable (i.e., patterns that lie on opposite sides of a hyperplane) (Hykin, Simon). The algorithm used to adjust the free parameters of this Neural Network first appeared in a learning procedure developed by Rosenblatt (Rosenblatt, F., 1958)(Rosenblatt, F., 1962) for his perceptron brain model. Multilayer perceptrons have been applied successfully to solve some difficult and diverse problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm. Neural Networks can have several layers. In multilayer feedforward networks the direction of signals is from input to output, there is no feedback in the layers. The diagram below (Figure 4) shows a 3-layered feedforward network (Callinan, Tim., 2003).

The input layer had 10 neurons for the first ANN classifier, because of, 24 neurons for the second, and 2 neurons for the third one using wavelet transform. The number of classes desired determined the output layers. In our study, the 2 neurons of the output layer correspond to the healthy bananas and unhealthy bananas. In practice, the number of neurons in the hidden layer varies according to the specific recognition task and is determined by the complexity and amount of training data available.

3. Results and conclusion

A sample banana image was captured (Figure 5) and then it was converted to gray level image (Figure 6). After that FCM algorithm was done to eliminate object from its background (Figure 7).

In addition to higher speed, the reason of using FCM is that the distance function versus iteration started at the lower value and with the less iteration it will be converged (Figure 8).

SUSAN edge detector uses no spatial derivatives at all. The *Smith* corner detector does not require any smoothing and so there is no degradation in localization accuracy due to smoothing.

For evaluating SUSAN algorithm, in comparison to some common edge detectors such as Canny (Albovik, 2000), Perwit (Albovik, 2000), Sobel(Albovik, 2000) and Roberts(Albovik, 2000), we added a Gaussian noise with variable level of noise variance; the variances were between 0.001 and 0.0105. We tested our algorithm for the all of the images in the database. To demonstrate the performance of the proposed algorithms, a sample of five hundred bananas, 350 defective and 150 healthy, were analyzed for sorting purposes and final inference for acceptance or rejection. The results are shown in the following figure.

We used one network structure of 10-24-2 (i.e. eight neurons at the input layer, twenty four at the hidden layer and two at the output layers), and the two other structures respectively of eight-three-two and eight-four-two. In practice, the number of neurons in the hidden layer varies according to the specific recognition task and is determined by the complexity and amount of training data available. If too many neurons are used in the hidden layer, the network will tend to memorize the data instead of discovering the features. This will result in failing to classify new input data. . To generate an efficient network, different learning coefficients and momenta were selected for different layers. In

the present work, the normalized root-mean-square (RMS) error of the output layer was used as a criterion to select these parameters. The selected learning coefficients and momenta correspond to the deepest slope of the normalized RMS error. Using the hyperbolic tangent sigmoid as the neural transfer function, the input feature vectors were scaled to the range from -1 to +1 to fit into the dynamic range of this function. Before the training process was started, all the weights were initialized to small random numbers. This ensured that the classifier network was not saturated by large values of the weights. The results of using neural network are shown in the following tables (Table 1).

References

- A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent. (1987). "Minimum average correlation energy filters," *Appl. Opt.* 26, pp. 3633-3630, 1987.
- Albovik. (2000). "Handbook of Image and Video Processing", Academic Press, 2000.
- Albovik. (2000). "Handbook of Image and Video Processing", Academic Press, 2000
- B. Hammer, M. Strickert, and T. Villmann. (2005). "Supervised neural gas with general similarity measure", *Neural Processing Letters* 21:21-44.
- Callinan, Tim. (2003). "Artificial Neural Network identification and control of the inverted pendulum". 2003, 18-24.
- Haralick RM, Shapiro LG. (1993). "*Computer and Robot Vision*". Addison-Wesley Publishing Company, 1993: volume1, 453-507.
- Hykin, Simon. "Neural Networks a comprehensive foundation", Prentice Hall. 106-235.
- J.C Dunn. (1983). "A fuzzy relative of the ISODATA process and its use detecting compact well-separated clusters", *J. Cybernetics*, Vol. 8, pp. 32-57, 1983.
- J.C. Bezdek and S. Pal (eds.) (1991). "*Fuzzy Models for Pattern Recognition*", IEEE Press, 1991.
- M.S.Mansoori, H.tajik, G.Mohhamadi, M.Pashna. (2008). 'Edge Defect Detection in Ceramic Tile Based on Boundary Analysis Using Fuzzy Thresholding and Radon Transform.', proceeding of ISSPIT 2008,IEEE
- O. S. Junior, A. C. Silva and Z. Abdelouah. (2009). Personal Identification Based on Iris Texture Analysis Using Semivariogram and Correlogram Functions. *International Journal for Computacional Vision and Biomechanics*, (2009).
- P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. (2002). Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. In *Seventh European Conf. on Computer Vision*, pages 97-112, 2002
- Rosenblatt, F. (1958). "The Perceptron: A probabilistic model for information storage and organization in the brain." *Psychological Review* 65, 386-408.
- Rosenblatt, F. (1962). "*Principles of Neurodynamics*". Washington, DC: Spartan Books.
- S. A. R. A. B. M.M.Mokji. (2006). "Starfruit Grading on 2-Dimensional Color Map," Regional Postgraduate Conference on Engineering and Science, Johore, 2006..
- S. Arivazhagan and L. Ganesan. (2003). Texture Segmentation Using Wavelet Transform. *Pattern Recognition Letters*, 24(16):3197-3203, December 2003.
- S. E. Grigorescu, N. Petkov, and P. Kruizinga. (2002). "Comparison of texture features based on Gabor filters," *IEEE Trans. Image Processing*, vol. 11, pp. 1160-1167, Oct. 2002.
- S. Pakhomov. (2002). Semi-supervised maximum entropy based approach to acronym and abbreviation normalization in medical text. In *ACL 2002*, 2002.
- S.M. Smith and J.M. Brady. (1997). "SUSAN - a new approach to low level image processing", *Int. Journal of Computer Vision*, Vol.23(1), pp.45-78, May 1997
- Sonka M. (1998). "Image Processing, Analysis, and Machine Vision. Brooks and Cole Publishing", 1998.

Table 1. Average of 10 extracted features

Healthy Bananas	Unhealthy Bananas
0.5724	0.2523
0.2515	0.0524
0.0895	0.2147
0.0098	0.5412
0.8540	0.0254
0.4785	0.9854
0.1924	0.5194
0.5147	0.9874
0.3564	0.8475
0.2504	0.7415

Table 2. Results of net

	1 st neural classifier	2 nd neural classifier	3 rd neural classifier
Normal	91.3	92.7	95.7
Abnormal	92.1	90.2	93.8

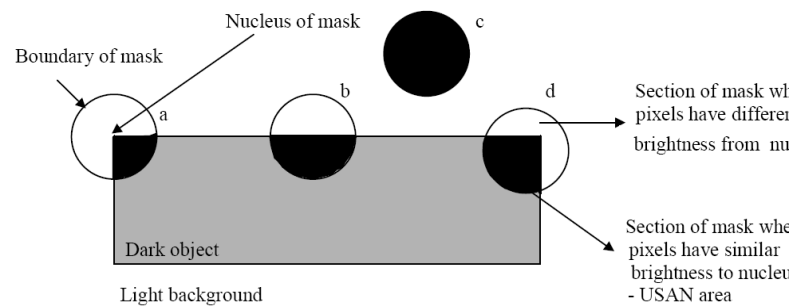


Figure 1. Four Smith edge finding masks at different positions in an image (S.M. Smith and J.M. Brady, 1997)

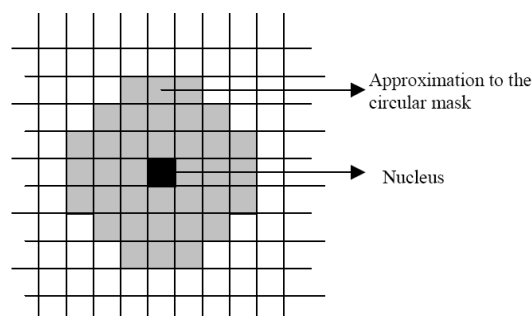


Figure 2. The Smith USAN (SUSAN) edge finding mask(S.M. Smith and J.M. Brady, 1997)

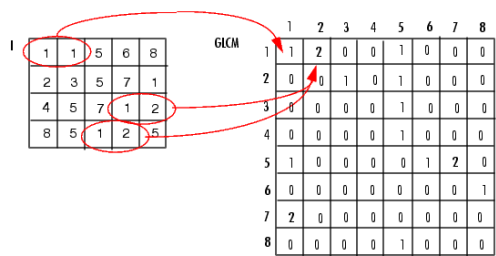


Figure 3. Calculating GLCM

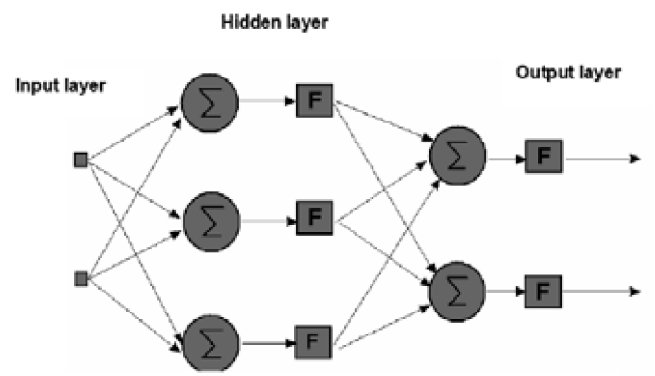


Figure 4. Diagram of a multi-layered perception



Figure 5. Input image



Figure 6. Gray image



Figure 7. Back ground elimination using FCM

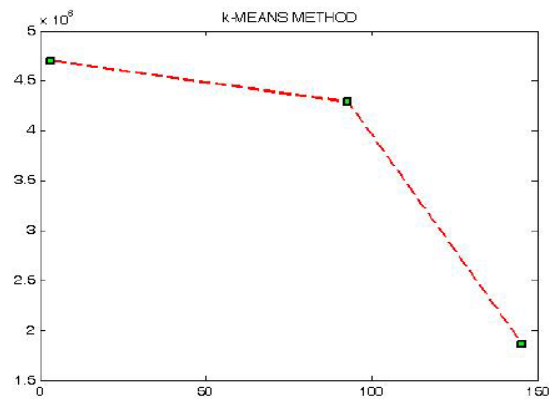


Figure 8.a. Distance function versus iteration for K-Means algorithm

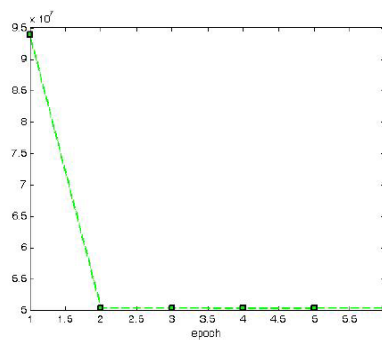


Figure 8.b. Distance function versus iteration for Hard C-Means algorithm

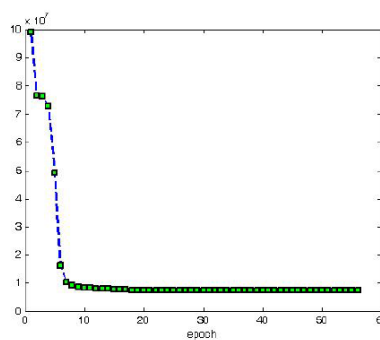


Figure 8.c. Distance function versus iteration for Neural-Gas algorithm

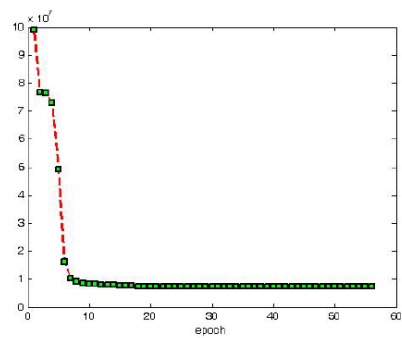


Figure 8.d. Distance function versus iteration for FCM algorithm.

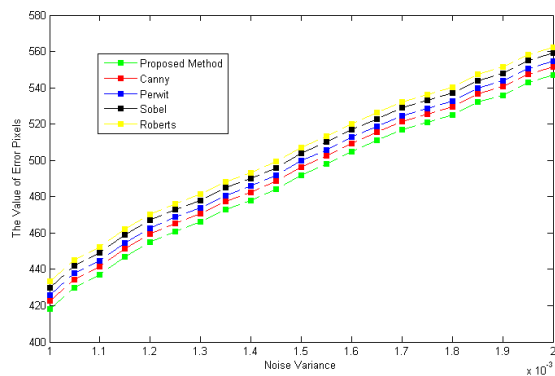


Figure 9. Evaluating edge detectors