JAST, Vol. 5, No. 3, pp 123–128

JAST Journal of Aerospace Science and Technology

Aircraft Visual Identification by Neural Networks

F. Saghafi[1], S. M. Khansari Zadeh[2], V. Etminan Bakhsh[3]

In the present paper, an efficient method for three dimensional aircraft pattern recognition is introduced. In this method, a set of simple area based features extracted from silhouette of aerial vehicles are used to recognize an aircraft type from its optical or infrared images taken by a CCD camera or a FLIR sensor. These images can be taken from any direction and distance relative to the flying aircraft. A multilayer perceptron neural network has been used for the purpose of aircraft classification. The network training has been carried out using a library of images generated by a 3D model of each aircraft. The neural network is successfully trained and used to recognize and classify arbitrary real aircraft images. The results show more than 90% accuracy in ideal conditions and very good robustness in the presence of noise.

INTRODUCTION

Accurate identification of unknown aircraft is very important in military operations. A machine vision (MV)-based system has been proposed by many authors for close proximity operations of aerospace vehicles [1-2] and for the navigation of UAVs [3]. Automated systems that quickly and accurately determine the identity of an aircraft could be a benefit in backing up electronic-signals identification methods. Also, autonomous recognition of aircraft is required in tracker systems especially those used in military engagement. Optical trackers are of interest because they are passive in nature and do not reveal the location of the imaging system. Therefore, the target is not aware of being tracked and cannot take any measures to evade the threat. Recent advances in FLIR infrared imaging technology has even made it possible to take images at night. Figure 1 shows examples of the type of a real image taken by a CCD camera or an IR sensor.

There are two approaches to automatic visual object recognition, local methods and global methods. Local methods use features such as critical points or high-resolution pursuit [4]. Systems using local features can perform well in the presence of noise, distortion, or partial occlusion because only one distinctive part needs to be recognized. However, the distinctive local features are vague to define for the aircraft recognition, and also, searching for them could be computationally intensive, which defeats our intention of designing a system which does not require sophisticated hardware.

Global methods include Fourier descriptors [5, 6], moments [7], and autoregressive models [8]. In contrast to local methods, global methods use total feature of aircraft, and as a result, they are suitable for aircraft recognition algorithms. Zaki et al [9] implemented an enhanced motion-based object tracker by using these global features. Dudani et al [7] used moment invariants for classification of airplane. In their work, six aircraft types are used and the images were taken from physical models. The training set was based on over 3000 images taken in a 140 by 90 degree sector. The testing set contained 132 images (22 images of each of the six classes) obtained at random viewing aspects. Reeves et al [10] have tried to improve the method by using Fourier Descriptors with the same data and obtained a best classification result of 93%. Alves et al [11] has classified ships using infrared silhouette with 70% accuracy. They have employed scale-invariant moments based on edge detection as features and a neural network for classification. The neural networks are also successfully used in other works such as Peterson [12] and Alsultanny [13].

This work addresses classification of aircraft by type from optical or infrared images assuming a single object presented in the image. The major question which is addressed in this work is whether it is pos-

1. *Associate Professor, Dept. of Aerospace Eng., Sharif Univ. of Tech., Tehran, Iran.*
2. *Ph.D. Candidate, EPFL Univ., Lausanne, Switzerland, Formerly M.Sc. Student, Dept. of Aerospace Eng., Sharif Univ. of Tech., Tehran, Iran.*
3. *M.Sc. Graduate, Dept. of Aerospace Eng., Sharif Univ. of Tech., Tehran, Iran*

sible to recognize aerial objects based on very simple features, but with a relatively bigger image library, and what will be the system performance and robustness in the presence of noise.

VISION-BASED RECOGNITION ALGORITHM

A new and simple area-based feature extraction method combined with a neural network approach is used for three dimensional aircraft identification. The trend so far has been to calculate or extract complicated feature sets from original images which are as descriptive of the original object as possible and has a small sized feature vector at the same time. This required the extracted features to be invariant to rotation or other possible changes in view.

However, if enough all direction views of the object are stored in the classification library, the feature set used for classification need not to account for these changes. For this purpose, a system has to be introduced with more built-in knowledge so that it can categorize the input patterns correctly even with minimal details or preprocessing. This eliminates the requirement for processing intensive feature extraction; thus, better real time implementation of the algorithm is possible.

The procedure of aircraft recognition from the captured images consist of three major steps:

1) Recognizing the aircraft pattern from the background and converting it into a binary image

2) Extracting feature vector from the obtained binary image

3) aircraft identification based on extracted feature vector using neural networks.

The implemented neural networks classify patterns with a very high generalization capability. A large library of patterns provides enough information for an efficient and acceptable system training.

Figure 1. Example of an aircraft image taken by an IR sensor (darker areas are less radiant).

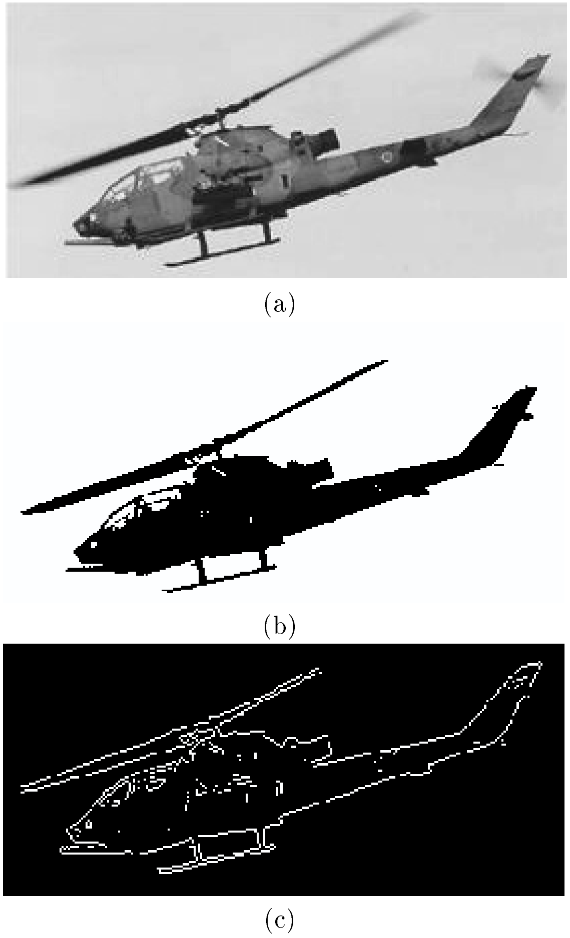

Figure 2. (a) Original image, (b) Area detection using Thresholding method, (c) Edge detection using Sobel method.

AIRCRAFT PATTERN RECOGNITION FROM THE IMAGE

One important factor for an accurate estimation of the aircraft relative orientation and position is to correctly recognize the aircraft pattern from the background. There are two common methods to accomplish this work, called the edge and thresholding techniques.

The first method is to detect the image edges, and then filter the detected aircraft pattern from the rest of the image. Basically, most edge detection techniques rely on the idea of computing a local derivative operator to obtain the intensity changes [9]. The intensity value of image at position x, y can be described by the function $f(x, y)$. Then, the image gradient $\nabla f(x, y)$ is defined as:

$$\nabla f(x,y) = \left[\begin{array}{cc} \partial f/\partial x & \partial f/\partial y \end{array}\right] = \left[\begin{array}{cc} f_x & f_y \end{array}\right] \quad (1)$$

There are many techniques that are applied for edge detection using local derivative operator such as Sobel method, Robert method, etc [14-16]. Figure 2-c demonstrates how a Sobel operator can detect the edges of an image.

The second method is to use some thresholds to recognize the aircraft pattern in an image. This method is based on dividing the image pixels to background and object portions by a threshold value. This threshold value is calculated using Otsu's method which chooses the threshold to minimize the intraclass variance of the black and white pixels [17]. Figure 2-b shows how a thresholding operator can recognize an aircraft pattern in an image.

In this paper, the thresholding method is used as our approach in recognizing the aircraft pattern for two reasons. First, the edge detection method is computationally more expensive than thresholding method. Second, using the aircraft pattern area for the relative orientation and position estimation demonstrates to be more robust in the presence of noise than the edge detection method.

As can be seen in Figure 2, the main problem with edge detection and thresholding method is that they cannot solely detect the desired object. Usually, there are a few other undesired objects that appear in the image. Furthermore, there might be noisy pixels which could affect the total computation. Reference [1] addresses some methods on filtering these noisy pixels and removing unwanted objects from the image. In the present work, it is assumed that the achieved image from the thresholding method leads to a single object.

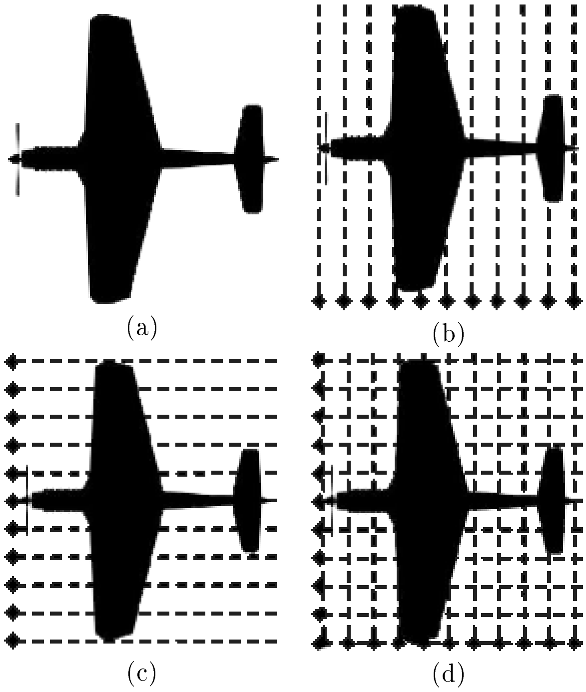

Figure 3. (a) Silhouette image of an aircraft, (b) vertical strip segmentation, (c) horizontal strip segmentation, (d) rectangular grid segmentation.

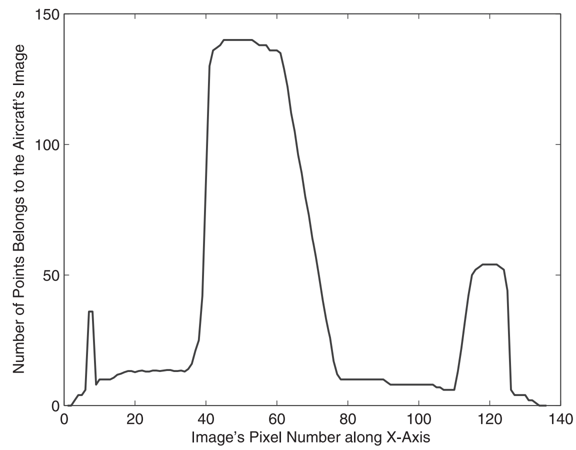

Figure 4. Horizontal area distribution of the aircraft shown in Figure 3-a.

Feature Vector

Four different feature vectors, all being area features, were tested. The idea is that if an image is turned into black and white in which the background is white and the object is black (see Figure 3), the distribution of the area of image over the background may reveal the type of an image (Figure 4). In order to do so, the images are divided to segments and the proportion of the image area to the background area in each segment is caculated. Putting these proportions together a feature vector can be created.

The segmentations used were vertical strips (Figure 3-b), horizontal strips (Figure 3-c), rectangular grid (Figure 3-d), and a combination of vertical and horizontal strips. Then the ratio of the object area to the total area of each segment is considered as an element of the feature vector. This vector has as many elements as the number of strips.

Generating Image Library

In order to train the implemented neural networks, it is necessary to provide an appropriate number of training data. As a result, 3D computer models of five different aircraft including Bell 206, C-130 Hercules, AH-1 Cobra, Su-25 and Mustang were used (Figure 5). Images are generated by camera rotating around the object in 3 mutually perpendicular planes by 1 degree steps (Figure 6). For each model, 1080 images were generated. Images have a resolution of 320 by 240. Then these images were converted into four aforementioned feature vectors to generate required training data for the implemented neural networks which will be discussed in the next section.

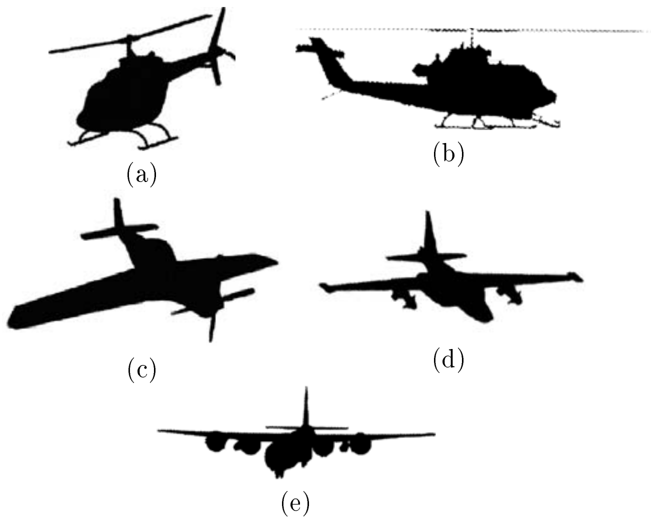

Figure 5. Aircraft used for recognition, (a) Bell 206, (b) AH1 (Cobra), (c) Mustang, (d) Su-25, (e) C-130.

Networks Architecture

A three layer back-propagation [18] neural network is used to produce a decision function with enough hidden units as shown in Figure 7.

The input layer contains the input nodes which interact with the outside environment. The input units are only buffer units that pass the signal without changing it. The hidden layer size is left to the appreciation of the user that will estimate the number of hidden nodes by his experience or by trial and error. This number must not be too large to avoid waste of computations and slow convergence process and not too low to make the network capable to absorb the set of training pixels. The output layer represents the number of nodes which is equal to the number of classes, each output node representing a class. Since in the present work, the goal is to determine the class of the aircraft in the image between five distinct flying vehicles, output layer consist of five neurons.

The activation function used in layers is sigmoid activation function. The output of sigmoid activation function (a) is calculated based on:

$$a = \frac{1}{1+e^{-n}} \tag{2}$$

where n is the net input of each neuron. If the weights and biases of a layer are defined as:

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix} \tag{3}$$

$$b = \left[\begin{array}{cccc} b_1 & b_2 & \cdots & b_S \end{array}\right]^T \tag{4}$$

then, the net input for the ith neuron can be calculated according to the following equation.

$$n_i = W(i, 1:R).b = (x_1 w_{i,1} + x_1 w_{i,1} + \cdots + x_R w_{i,R}) + b_i \tag{5}$$

The network is trained with Levenberg-Marquardt ([19] and [20]) back-propagation method to learn patterns.

As stated in the previous section, computer generated images were used to train the network; however, for the network performance evaluation, both computer generated and real images were employed.

SIMULATION RESULTS

In order to evaluate effects of feature vector on the ability of the networks to identify flying vehicles, performance resulted from use of each feature vector

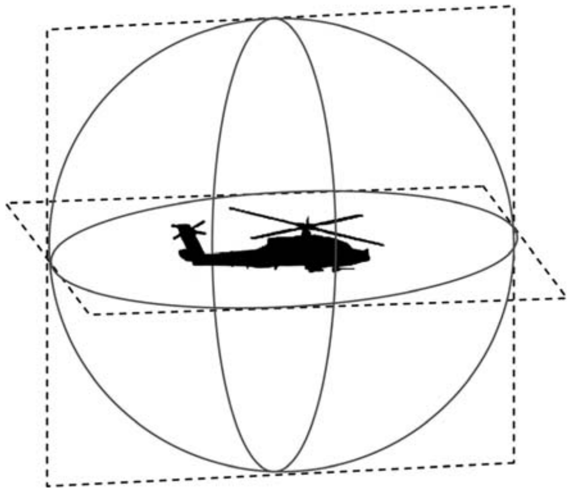

Figure 6. Generating computer images through camera rotation around the object.

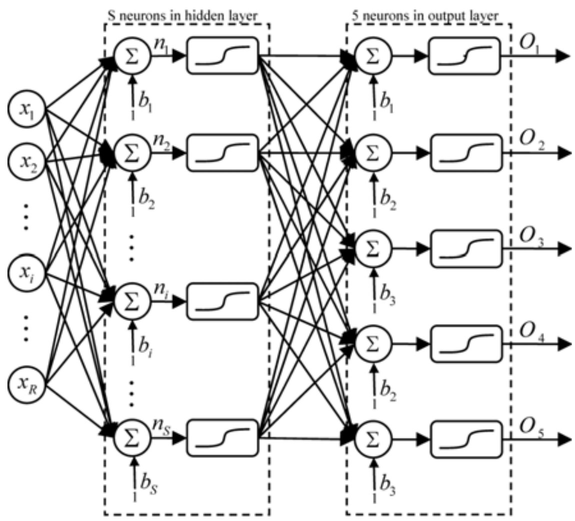

Figure 7. Back-propagation network architecture used as a decision function.

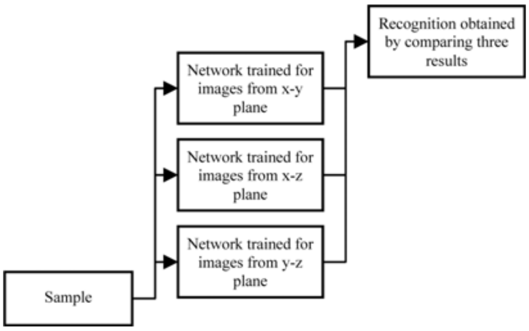

Figure 8. Network arrangement used for aircraft recognition.

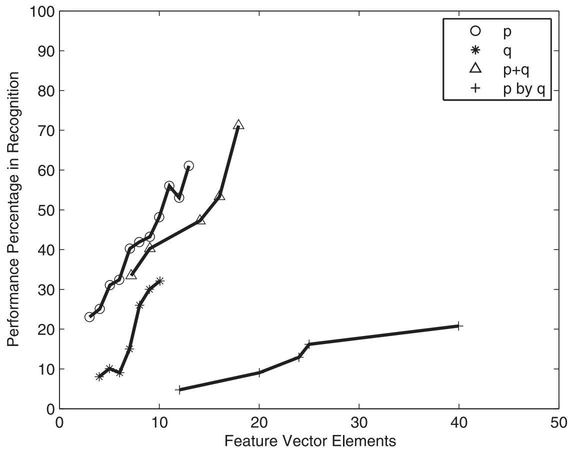

Figure 9. Comparing performance in recognition for different feature vector types.

including vertical strips, horizontal strips, rectangular grid, and combination of horizontal and vertical strips, was investigated. Figure 9 shows how each type of feature vector affects results. In this figure, "p" and "q" stands for number of vertical and horizontal strips respectively. "p by q" is rectangular grid and "p + q" is combination of vertical and horizontal strips. As can be seen for all types of feature vectors, by increasing the number of segments in the image, the performance of the networks in recognition are improved. Furthermore, among all used feature vectors, the vector that consists of a combination of vertical and horizontal strips is able to provide the highest performance with respect to its size. Subsequently, this feature vector was adopted for the rest of evaluations.

Furthermore, another achievement which is attained during the algorithm evaluations was that performance of the algorithm in identification can be improved by changing the network architecture. At first attempt, a single network with all training data was trained. The results were not satisfactory because the feature set utilized had no cue of how the camera was looking at the object and as such a classification of the different views of the aircraft was necessary to be somehow implemented *e.g.* by modifying feature vectors. Instead, three distinct networks were implemented, each of which was trained with a different set of images. These images were classified as images taken in $x-y$, $y-z$, and $x-z$ plane (Figure 8). Accordingly, each of the networks can decide on the type of the presented image. Outputs from these three networks were put together and whenever at least two of the three networks classified the image as corresponding to the same category, it was assumed as the output of the combined network. This highly improved the results. Table 1 shows how the accuracy in identification is increased using three networks architecture for "p + q" feature vector.

The networks architecture used in this paper are capable of recognizing every image made by computer from 3D models. Although not every possible view is tested one-by-one, all challenging situations involving perspective and rotation were tested. The networks were also evaluated by real images of aircraft at different flying attitudes. As for real images, there were some degradations in quality. These degradations are as a result of defects in the original image. The defects include noise, complex or dark background, multi-object occlusions. Because it is assumed that a single aircraft is in the vicinity, the occlusion and multi-object are not implemented. However, the multi object problem can be overcome by using object recognition algorithms and presenting these distinct objects to the neural network one-by-one [1]. The network could recognize the real images with a performance of better

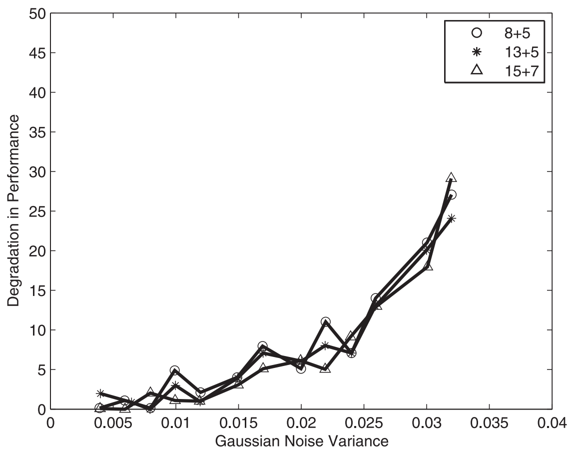

Figure 10. Effect of Gaussian noise on network performance.

Table 1. Effects of using a three network architecture in aircraft identification improvement.

Feature vector: p+q	Percentage of identification for a single network	Percentage of identification for three networks
18	67%	74%
19	73%	79%
20	81%	87%

than 90% when no noise is present. These images were arbitrarily selected from a library of real flying aircraft.

Gaussian noises were added to images to investigate the performance of the network in the presence of noise. This type of noise is selected because it is one of the hardest models of noises to remove with image restoration algorithms. Figure 10 shows how adding Gaussian noise of increasing variance can affect the performance of the network. This is shown for different sizes of feature vectors and shows.

CONCLUSION

In this paper a neural network algorithm was implemented, which is able to recognize five distinct aerial vehicles from captured images. The proposed neural network was trained with a large number of training data generated from different viewpoints. We realized that when a large number of training data are available, and when the proper classification approach is employed, recognition of images can be done with rather simple feature vectors. This significantly decreases the amount of calculation, and in the meantime, maintains the performance of aircraft recognition within an acceptable limit. Extension of this work could include estimating the detected aircraft's altitude and position from the captured images.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Sharif University of Technology for supporting the research presented in this paper.

REFERENCES

1. Saghafi F., Khansari Zadeh S.M., "Vision-Based Trajectory Tracking Controller for Autonomous Close Proximity Operations", *Proceeding of 29th IEEE Aerospace Conference, Big Sky*, (2008).
2. Philip N.K., Ananthasayanam M.R., "Relative Position and Orientation Estimation and Control Schemes for the Final Phase of an Autonomous Docking Mission of Spacecraft", *Acta Astronautica*, **52**, PP 511-522(2003).
3. Sinopoli B., Micheli M., Donato G., Koo T.J., "Vision Based Navigation for an Unmanned Aerial Vehicle", *Proceedings of the 2001 IEEE Conference on Robotics and Automation*, Seoul, (2001).
4. Jaggi S., Karl C., Mallat S., and Willsky A., "Silhouette Recognition Using High Resolution Pursuit", *Pattern Recognition*, **32**, PP 753-771(1999).
5. Wallace T.P., and Wintz P.A., "An Efficient 3D Aircraft Recognition Algorithm Using Fourier Descriptors", *Computer Graphics Image Processing*, **13**, PP 99-126(1980).
6. Richard C., and Hemani H., "Identification of Three-Dimensional Objects Using Fourier Descriptors of the Boundary Curve", *IEEE Transactions on Systems, Man, and Cybernetics, SMC-4*, **4**, PP 371-378(1974).
7. Dudani S. A., et. al., "Aircraft Identification by Moment Invariants", *IEEE Transactions on Computers*, **26(1)**, PP 39-46(1977).
8. Kashyap R. and Chellapa R., "Stochastic Models for Closed Boundary Analysis: Representation and Reconstruction", *IEEE Trans. on Inf. Theory*, **27**, PP 109-119(1981).
9. Zaki M., El Nahas M.Y. and Youssef M., "EMBOT: An Enhanced Motion-Based Object Tracker", *The Journal of Systems and Software*, **69**, PP 149-158(2004).
10. Reeves A., Prokop R., Andrews S. , and Kuhl F., "Three-Dimensional Shape Analysis Using Moments and Fourier Descriptors", *IEEE Trans. on PAMI*, **10(6)**, PP 937-943(1988).
11. Alves J., et. al., "Robust Recognition of Ship Types from an Infrared Silhouette", US Naval Postgraduate School, Monterey, CA, (2004).
12. Egmont-Peterson M., de Ridder D., and Handels H., "Image Processing With Neural Nerworks-A Review", *Pattern Recognition*, **35**, PP 2279-2301(2002).
13. Alsultanny Y. A., and Aqel Musbah M., "Pattern Recognition Using Multilayer Neural-Genetic Algorithm", *Neurocomputing*, **51**, PP 237-247(2003).
14. Canny J., "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8(6)**, PP 679-698(1986).
15. Lim J. S., *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, NJ, Prentice Hall, PP 478-488(1990).
16. Parker J. R., *Algorithms for Image Processing and Computer Vision*, John Wiley & Sons, Inc., PP 23-29(1997).
17. Otsu N., *A Threshold Selection Method from Gray-Level Histograms*, MIT Press, Cambridge, (1969).
18. Hagan M.T., Demuth H.B. and Beale M.H., *Neural Network Design*, Boston: PWS publishing company, .
19. Levenberg K., "A Method for the Solution of Certain Problems in Least Squares", *Quart. Appl. Math.*, **2**, PP 164-168(1944).
20. Marquardt D., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", *SIAM J. Appl. Math.*, **11**, PP 431-441(1963).