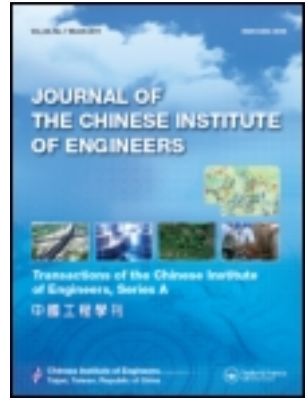


This article was downloaded by: [National Chiao Tung University 國立交通大學]

On: 28 April 2014, At: 01:17

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Journal of the Chinese Institute of Engineers

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tcie20>

### Applying fuzzy logic in the modified single-layer perceptron image segmentation network

Jyh-Yeong Chang<sup>a</sup> & Jia-Lin Chen<sup>a</sup>

<sup>a</sup> Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C.

Published online: 03 Mar 2011.

To cite this article: Jyh-Yeong Chang & Jia-Lin Chen (2000) Applying fuzzy logic in the modified single-layer perceptron image segmentation network, Journal of the Chinese Institute of Engineers, 23:2, 197-210, DOI: [10.1080/02533839.2000.9670538](https://doi.org/10.1080/02533839.2000.9670538)

To link to this article: <http://dx.doi.org/10.1080/02533839.2000.9670538>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>



# APPLYING FUZZY LOGIC IN THE MODIFIED SINGLE-LAYER PERCEPTRON IMAGE SEGMENTATION NETWORK

Jyh-Yeong Chang\* and Jia-Lin Chen  
Department of Electrical and Control Engineering  
National Chiao Tung University  
Hsinchu, Taiwan 300, R.O.C.

**Key Words:** Neural network, fuzzy logic, image segmentation, unsupervised learning.

## ABSTRACT

In this paper, we propose a fuzzy-logic-based modified single-layer perceptron (MSLP) image segmentation network for object extraction. We select a sigmoid gray level transfer function with the aid of the input image histogram and map the input gray levels into the interval  $[0,1]$ . Then we adopt the linear index of fuzziness of the output nodes as the error function of the image segmentation system to incorporate the learning capability of a neural network. Our scheme can successfully extract objects from the background. To further enhance the capability of the segmentation system, the proposed network is incorporated with fuzzy if-then rules to adaptively adjust the threshold of the activation function of the MSLP output neuron for best matching the local characteristics of the image. Fuzzy if-then rules involving the edge intensities and vertical positions of pixels are reasoned to determine the threshold adaptively. From the results of segmenting forward looking infrared (FLIR) images, better segmentation images have been obtained by incorporating fuzzy if-then rules with the MSLP segmentation technique. As demonstrated by this study, it is promising and worthy of study that incorporating human knowledge in terms of fuzzy rules into a designed numerical algorithm can further improve performance, not only in the segmentation problem we present.

## I. INTRODUCTION

Image segmentation plays a key role in image analysis and computer vision. From this perspective, a correct segmentation must be efficiently made between the object and the background. Many segmentation techniques have been proposed and developed. Some important and commonly used segmentation methods are relaxation labeling, Markov random field modeling, region growing and splitting, and global or local thresholding (Haralick and Shapiro, 1985).

Among them, the image thresholding technique is the most general concept for separating the pixels of an image into two regions, sometimes more than two regions, of the object and the background. The threshold is usually set at the deep and obvious valley of the histogram. However, the valley of an image histogram is usually not obvious and thus it is not easy to determine the threshold (Abutaleb, 1989; Lee and Chung, 1990; Sahoo *et al.*, 1988). Another obstacle is that these methods belong primarily to sequential processing and do not include the positional

---

\*Correspondence addressee

relationship information about a pixel and its neighbors. A neural-based segmentation approach (Banz and Gish, 1990; Ghosh and Pal, 1992; Ghosh *et al.*, 1991; Ghosh *et al.*, 1992; Rananath and Kuntimad, 1994), however, can easily include a pixel as well as its neighborhood information in the network structure because the inputs of a pixel and its neighborhood pixels are weighted summed and then processed. Moreover, fast computation can also be achieved since a neural network is highly parallel in structure, which is especially advantageous for a computationally involved task of image processing. These two facts suggest that neural models can be good architectural frameworks for image segmentation applications. In this application, a self-supervised multilayer perceptron (MLP) neural network (Ghosh *et al.*, 1992) and a self-organized mapping (SOM) based method (Ghosh and Pal, 1992) were proposed for object extraction. In the MLP-based segmentation method, they incorporated fuzziness measures judiciously into the neural model to avert the limitation of unknown desired outputs for the supervised MLP network. The SOM-based method explores Kohonen's concept of self-organized feature mapping for object extraction from an image. From our experience, the MLP-based scheme sometimes fails to extract part of a particular object and the SOM-based scheme frequently produces a segmented image with pepper-and-salt noise. Moreover, the SOM-based method is computationally intensive and its noise immunity is weak. To eliminate these defects, we propose an algorithm, named Modified Single-Layer Perceptron (MSLP), which has the characteristic of self-supervision. Inspired by the MLP approach (Ghosh *et al.*, 1992), the proposed MSLP network is a recursive single-layer perceptron structure between the input and output layers, and hence is structurally simpler than the MLP model. In the MSLP segmentation network, both the nodes of the input and output layers are equal to the dimensions of the image. The output layer nodes are connected from the corresponding input window mask, i.e., every pixel is associated with a mask. In this way, the segmentation network is structured in a fully parallel form so that whole image pixels can be segmented simultaneously. Hence, the image segmentation task can be performed very fast. It would seem that the linear relationship among a neighborhood connection of pixels is enough for the image segmentation task, without the necessity to resort to a much more complex nonlinear relationship structure of neighbor pixels provided by an MLP (Ghosh *et al.*, 1992). More seriously, a gradient-based error back-propagation learning algorithm used in a complex structure like MLP is slower in learning convergence and is more prone to get stuck on local minima than the proposed

simpler MSLP approach.

In our proposed MSLP segmentation network, the brightness of a pixel is first transformed into the interval  $[0,1]$  using a sigmoid gray level transfer function. To obtain a bi-tone image of the object and background, we take an index of fuzziness, the sum of each image pixel's bi-tone fuzziness, as the error function of the system. The delta learning rule is used to adjust the synapse weights so that the system error reduces gradually. When the error converges to a small value, the segmentation between the object and the background is completed. In the present model, we only need a *rough* estimation of the valley of the histogram to prevent the difficult problem of precise valley estimation encountered in thresholding based segmentation techniques.

Our proposed segmentation network does not lead to a precise enough segmentation for images with small regions of low-contrast in gray level. Such regions in the image are referred to as *ambiguous regions* hereafter. In these regions, the object pixels' gray levels are very close to that of the background. Because an image, especially in ambiguous regions, is full of uncertainty, vagueness, and imprecision, it is wise to describe and manipulate this kind of information to further improve accuracy during the course of image segmentation. Such ambiguous information is best managed with fuzzy logic (Keller, 1997; Zadeh, 1965; Zimmermann, 1996) because fuzzy sets have flexible representation ability in describing uncertainty, vagueness, and imprecision, and also have an intuitively easy thinking process. Fuzzy logic works well where the important decision making criteria can be expressed as knowledge in terms of linguistically stated rules. Therefore, if the fuzzy set concept and the knowledge of fuzzy inference are properly added into the MSLP segmentation algorithm to deal with the ambiguous regions of the image, a better segmentation result can be expected. This is because not only are the numerical data processed, but also are the uncertain and imprecise features existing in the image modeled and taken into account at the same time. As a result, incorporating fuzzy rules into the MSLP segmentation network can improve its accuracy greatly in a segmentation task.

To model the uncertainty and the vague characteristics existing in the input image, we will first calculate the edge intensities of the gray values of pixels over a  $3 \times 3$  region. Then, we introduce fuzzy logic rules, whose fuzzy sets describe the edge intensities and the vertical position (an observed feature) of pixels in the image. Fuzzy if-then rules are inferred to adaptively determine the threshold of the activation function of output neurons in the MSLP network. The inclusion of if-then rules into the MSLP network has led to a great improvement in

the segmentation correctness in the ambiguous regions.

The rest of this paper is organized as follows. Section II briefly reviews the conventional single-layer perceptron network and its learning rules. Section III introduces the structure and synapse weight updating rule of the present modified single-layer perceptron segmentation network for object extraction. The simulation results are also provided and compared to the existing methods. In Section IV, we incorporate fuzzy if-then rules into the proposed segmentation network to dynamically infer the threshold of the activation function of output neurons. The numerical simulation has shown that the MSLP net with such fuzzy if-then rule incorporation has enhanced its segmentation accuracy, especially for the ambiguous regions of the images. Concluding remarks are finally made in Section V.

## II. SINGLE-LAYER PERCEPTRON NET AND DELTA LEARNING RULE

Since our proposed segmentation system makes use of a modified version of a single-layer perceptron structure, it is instructive to review this net briefly. A schematic representation of a conventional Single-Layer Perceptron (SLP) neural network (Zurada, 1992) is given in Fig. 1. Such a net is composed of sets of nodes arranged in two layers. The input layer consists of  $I$  input neurons, and these neurons correspond to an  $I$ -dimensional input vector  $X=[x_1, x_2, \dots, x_I]$ . The output layer consists of  $J$  output neurons. The output neurons,  $y_j, j=1,2,\dots, J$ , are the network responses corresponding to an input vector  $X$ . Each output node in the output layer is activated in accordance with the weighted sum of input data and the activation function of the node. The  $j$ -th output node value  $y_j$  is

$$y_j = a(\text{net}_j) = a\left(\sum_{i=1}^I w_{ij}x_i\right), \quad (1)$$

where  $w_{ij}$  is the connection weight between the  $i$ -th node of the input layer and the  $j$ -th node of the output layer and  $a$  is the activation function. The activation function adopted in this paper is the sigmoid function, defined as

$$a(\text{net}) = \frac{1}{1 + \exp^{-\lambda(\text{net} - \theta)}}, \quad (2)$$

where  $\lambda$  controls the steepness of the function and  $\theta$  is the thresholding value.

Now we consider an SLP network trained by feeding training patterns sequentially. Let  $\mathbf{d}=[d_1, d_2, \dots, d_J]^T$  denote the desired output vector of a training pattern. The square error  $E$  for the training instance

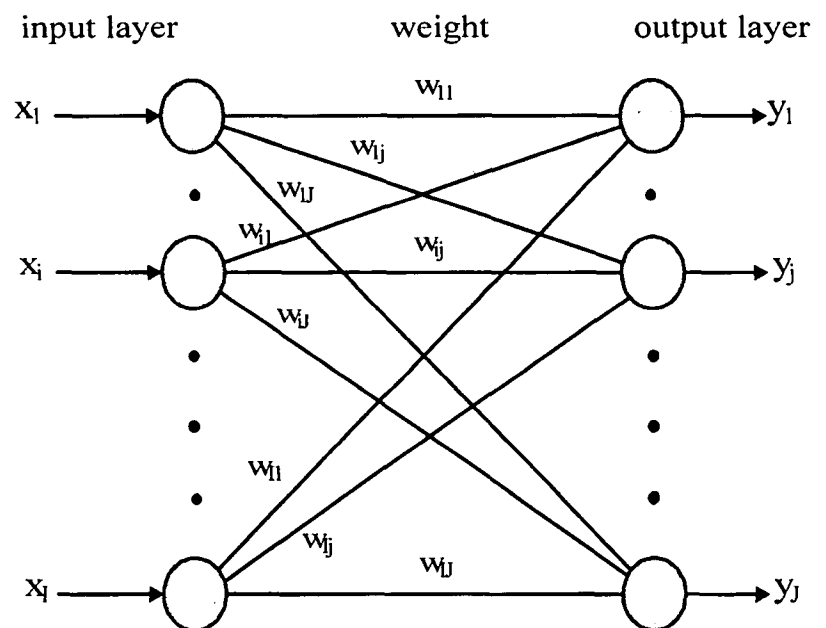


Fig. 1. The conventional single-layer perceptron structure.

is given by

$$E = \frac{1}{2} \sum_{j=1}^J (d_j - y_j)^2.$$

The synapse weight learning involves varying the weights in a manner to reduce the error  $E$  as much as possible. This *sequential mode* of training can be achieved by moving the weights in the direction of the negative gradient of  $E$  as given by

$$\begin{aligned} \Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}} \\ &= -\eta \frac{\partial E}{\partial (\text{net}_j)} \cdot \frac{\partial (\text{net}_j)}{\partial w_{ij}} \\ &= -\eta \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial (\text{net}_j)} \cdot \frac{\partial (\text{net}_j)}{\partial w_{ij}}. \end{aligned}$$

Thus,

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial y_j} a'(\text{net}_j) x_i, \quad (3)$$

where  $\eta$  is a positive learning constant. Because the sigmoid function is used as the activation function, we have

$$a'(\text{net}_j) = y_j(1 - y_j).$$

The delta learning rule then becomes

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial y_j} y_j(1 - y_j) x_i. \quad (4)$$

Note that the SLP net can also be trained using a *batch mode* instead of a sequential mode as described above. But these two training performances are almost the same (Zurada, 1992).

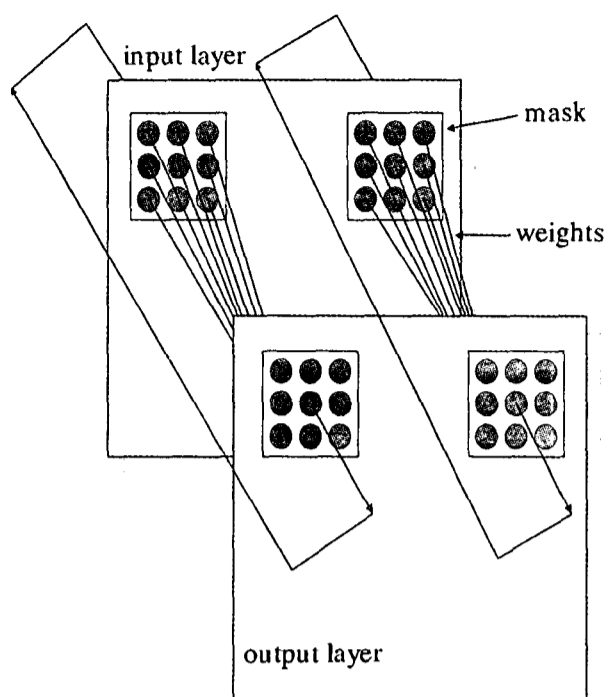


Fig. 2. The modified single-layer perceptron structure.

### III. THE MODIFIED SINGLE-LAYER PERCEPTRON SEGMENTATION NETWORK FOR OBJECT EXTRACTION

In this section, a segmentation network for object extraction is proposed. Since our present model modifies a conventional single-layer perceptron structure in several respects, hence it is referred to as a Modified Single-Layer Perceptron (MSLP) network.

#### 1. The MSLP Network Architecture

Figure 2 depicts the structure of the proposed MSLP network. If the dimensions of the image to be segmented are  $M \times N$ , then both the input and output neuron dimensions of the proposed MSLP net are also  $M \times N$ . In this network, each neuron in the output layer corresponds to a pixel in the image. Because the MSLP net is designed to segment the whole image iteratively, all the  $M \times N$  input neurons are fed back from the corresponding output neurons each time a new learning iteration is started. Instead of a sequence of training by feeding training patterns, i.e., pixels, in the conventional SLP net, our proposed net would feed back, in parallel, all the output neurons to the input neurons to perform a new learning iteration. This feedback iterative learning scheme will be further illustrated in the next section. Each output neuron is only partially, i.e., locally, connected to a set of input neurons. The rationale behind this partial connection in the MSLP net is due to the exploitation of the spatial filtering concept as described in the following.

In digital image processing, spatial domain methods use the concept of mask frequently (Chen *et al.*; Gonzalez and Woods, 1992). Generally

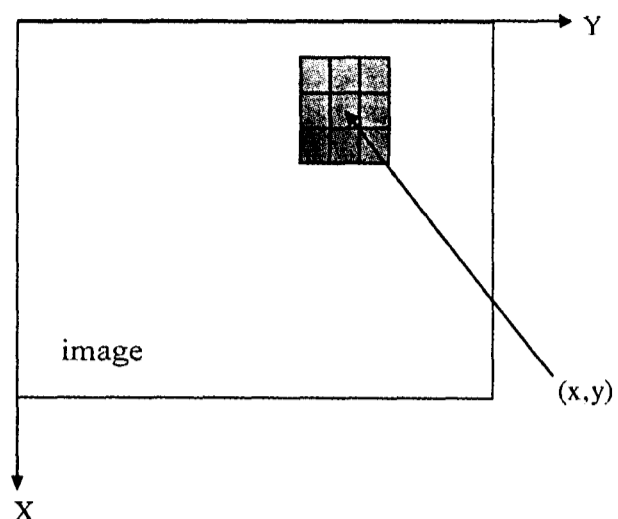


Fig. 3. A 3x3 window mask about a pixel.

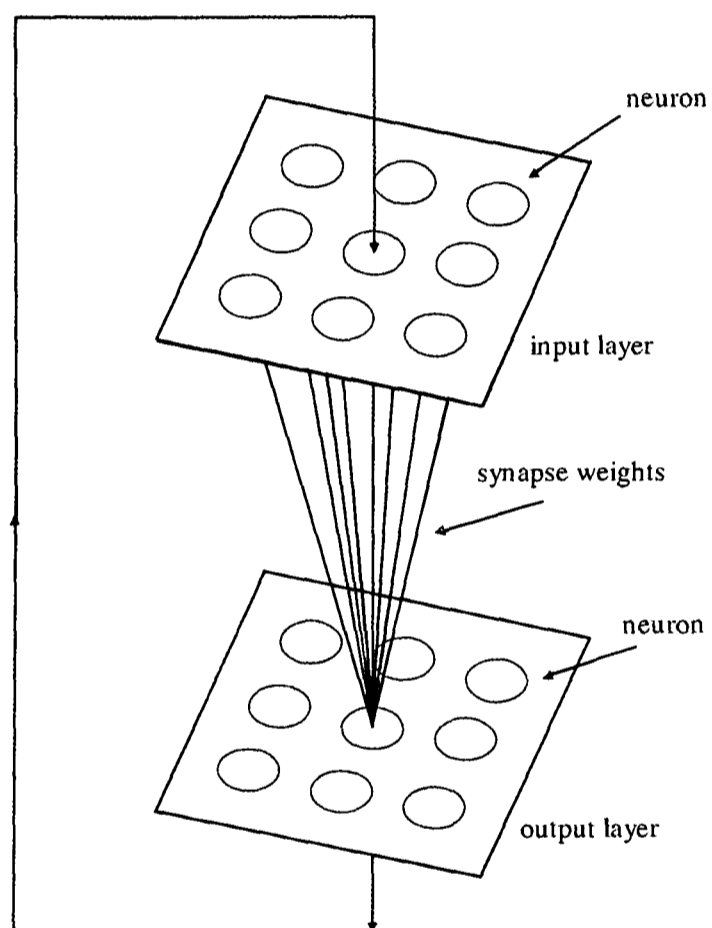


Fig. 4. An output neuron of the modified single-layer perceptron structure.

speaking, there exists a similarity among neighborhood pixels in an image. A pixel and its surrounding eight pixels in general have very close gray values. For example, if most of the eight pixels lean toward a white color, then this pixel will lean highly toward a white color too. For this reason, a 3x3 window mask, as shown in Fig. 3, will be used for each output neuron to confine its input set of neurons. Each neuron in the output layer is connected to a set of nine input neurons in a corresponding 3x3 window mask, as shown in Fig. 4. The center of the window mask is the input neuron fed back from the corresponding output neuron. Therefore,  $M \times N$  output neurons are associated with  $M \times N$  3x3 window masks, in which consecutive output neurons are also associated

with consecutive window masks whose centers are the input neurons, also consecutive. Output neurons on the boundaries of the image will have fewer than nine input neurons, which is specified by the window mask and also in the image. The output neurons on the boundaries are treated as special cases. Accordingly, each output neuron, except those neurons on the boundaries, has nine links from the input neurons of the corresponding  $3 \times 3$  window mask.

A slight difference from the conventional single-layer perceptron arises in terms of the initial weight setting of the network. It is known that the random setting of the initial weights may cause a pseudo noise effect (Ghosh *et al.*, 1992), thereby affecting the data structure of the original image, and possibly increasing the processing time. For this reason, all of the weights are set to unity in the beginning. Moreover, the threshold value of the activation function of the output neuron must also be selected. Since each neuron in the output layer has nine links, and each input neuron pixel value lies in  $[0,1]$ , we may choose  $9/2$ , i.e., the middle value of the total input range, as the threshold value.

## 2. Sigmoid Gray Level Transfer Function of the Input Image Pixels

The input to a neuron in the input layer is given as a real number in  $[0,1]$ . First the gray level of input image  $X$  is transferred in the following form

$$g: [l_{\min}, l_{\max}] \rightarrow [0,1],$$

where  $l_{\min}$  and  $l_{\max}$  are, respectively, the lowest and the highest gray levels in the input image. Through the transformation, the brighter pixel produces the larger value, and vice versa. To do this, it is common to use a linear gray level transfer function  $g_l$ , defined by

$$g_l(x_{mn}) = \frac{x_{mn} - l_{\min}}{l_{\max} - l_{\min}}.$$

The transferred value is proportional to the gray level of the pixel. Ghosh and Pal (1992) and Ghosh *et al.* (1992) use the above linear gray level transfer function for image segmentation. This paper instead made use of the sigmoid function  $g_s$ ,

$$g_s(x_{mn}) = \frac{1}{1 + \exp^{-(x_{mn} - t)/t_o}}, \quad (5)$$

as the gray level transfer function, with the threshold  $t$ , chosen with the aid of the input image histogram as described below. Parameter  $t_o$ , considered approximately as the slope of  $g_s$ , can be chosen as a value

proportional to the standard deviation of the image histogram and 1-2 multiples of the standard deviation is generally appropriate, from our experience. In this study, we set  $t_o$  equal to 50.

The sigmoid gray level transfer function above, Eq. (5), has lower tangent values at both ends and a much higher slope around the threshold. This steep region sharpens the difference between the transformed gray levels around the neighborhood of the threshold, i.e., the boundary between the objects and the background. The pixels' sharpened difference caused by this nonlinear transform reduced the work load in the following segmentation task performed using the MSLP net, and hence will speed up the convergence rate. For a bright-background and dark-object image, the first valley from the right side (the brightest side) of the histogram can be used as a suitable threshold value for distinguishing the background and the objects. Consequently, this valley position can also be used as the threshold  $t$  of the sigmoid gray level transfer function. In the case of a bright-object and dark-background image, the above procedure can still be applied, except that the threshold of the sigmoid gray level transfer function is chosen as the first valley from the left side of the histogram. Note that only a rough estimation of the valley of the histogram is required, since determining the precise location of the valley is relatively difficult. Via the continuing learning steps of the MSLP net, the neighborhood connection provides local information efficiently and leads to a correct segmentation of each pixel, either belonging to the object or the background. Therefore, each segmented pixel can best match the given image, even though the selection of the valley is not precise.

## 3. The Learning Procedures of the MSLP Segmentation Network

We attempt to segment image  $X$  of dimensions  $M \times N$  by the modified single-layer perceptron network. Note that the single-layer perceptron based learning algorithm is conventionally a supervised learning scheme which needs desired outputs for training. However, the desired output is unavailable in the case of object extraction by segmentation. To solve this problem, we may utilize the output nodes' linear index of fuzziness (Ghosh *et al.*, 1992) as the error function  $E$  in the delta learning rule of the system. Before giving the definition of this index, we have to define  $\hat{X}$ , the nearest bi-tone version of  $X$ , as given by

$$\mu_{\hat{X}}(x_{mn}) = 0, \text{ if } \mu_X(x_{mn}) < 0.5,$$

$$\mu_{\hat{X}}(x_{mn}) = 1, \text{ if } \mu_X(x_{mn}) \geq 0.5,$$

where  $\mu_X(x_{mn})$  represents the membership value of brightness at the  $(m,n)$ -th pixel  $x_{mn}$ . The linear index of fuzziness  $v_l$  for the  $M \times N$  image is defined as

$$v_l = \frac{2}{MN} \sum_m \sum_n \left| \mu_X(x_{mn}) - \mu_{\tilde{X}}(x_{mn}) \right|$$

$$= \frac{2}{MN} \sum_m \sum_n \min(\mu_X(x_{mn}), 1 - \mu_X(x_{mn})). \quad (6)$$

This index measures the fuzzy difference between  $X$  and  $\tilde{X}$ , that reflects the average amount of ambiguity involved in deciding whether a pixel belongs to a specific set, for instance, object or background. In the proposed MSLP net, the values of the output neurons represent the degree of brightness of the corresponding pixel in the image: the larger the value is, the brighter it is, and vice versa. Therefore, this measure of fuzziness of the output neurons can be considered as the error of the MSLP net because this error reflects the deviation from the well-segmented state of the network. Ideally, as the index converges to zero, the output neurons are divided into two subsets, zero and one, and thus segmentation is completed. Consequently, the linear index of fuzziness of the output neurons is the key in changing the proposed MSLP segmentation network from a supervised learning mode to a self-supervisory process (Ghosh *et al.*, 1992).

Next, the learning operation of the system is considered. The initial values of all neurons in the input layer are determined using sigmoidally transformed values, Eq. (5), of image pixels. The input net value to any neuron in the output layer is the weighted sum of the input neurons in the corresponding  $3 \times 3$  window mask. The activation function of Eq. (2) is then applied to get the outputs of the neurons in the output layer and the linear index of the fuzziness  $v_l$ , Eq. (6), of the output layer is calculated. If  $v_l$  is greater than a preassigned small positive value, then the learning iteration continues and the synapse weights are updated according to the delta learning rule, i.e.,

$$\Delta w_i^{(j)} = -\eta \frac{\partial E}{\partial y_j} y_j (1 - y_j) x_i^{(j)}, \quad i=1, \dots, 9,$$

$$j=1, \dots, MN, \quad (7)$$

where  $E$  denotes the linear index of the fuzziness  $v_l$  of the output layer;  $y_j$ ,  $j=1, 2, \dots, MN$ , represents the output of the  $j$ -th output neuron, and  $x_i^{(j)}$  represents the  $i$ -th input neuron in the mask corresponding to the  $j$ -th output neuron. A more compact form of Eq. (6) for the proposed segmentation net becomes

$$E = \frac{2}{MN} \sum_j [\min(y_j, 1 - y_j)]. \quad (8)$$

It follows from Eq. (8) that

$$-\frac{\partial E}{\partial y_j} = \begin{cases} -\frac{2}{MN}, & 0 \leq y_j < 0.5 \\ \frac{2}{MN}, & 0.5 \leq y_j \leq 1. \end{cases} \quad (9)$$

Thus the delta synapse weight  $\Delta w_i^{(j)}$  to be learned in each iteration becomes

$$\Delta w_i^{(j)} = \begin{cases} -\eta \frac{2}{MN} y_j (1 - y_j) x_i^{(j)}, & 0 \leq y_j < 0.5 \\ \eta \frac{2}{MN} y_j (1 - y_j) x_i^{(j)}, & 0.5 \leq y_j \leq 1. \end{cases} \quad (10)$$

After the synapse weights have been changed by Eq. (10), all of the outputs of the neurons in the output layer are then fed back to the input layer as a new set of input data to start a new learning iteration. The learning step above will reduce the system's bi-tone fuzziness gradually, which is equivalent to classifying the objects from the background better. Such a learning iteration will not stop until the network stabilizes, i.e.,  $v_l$  converges to a negligible value. After arriving at this situation, the output neurons in the output layer become either 0 or 1 or very close to 0 or 1. Neurons with values very close to 0 constitute one group and those neurons with values very close to 1 constitute the other group. This iterative learning procedure will make the synapse weights of the MSLP net properly adjusted and thus the image segmentation will be sharp.

Note that the mechanism of our proposed system is quite similar to the manner of self-organization and feature mapping in the eyes. Self-organizing is a basic feature of the human visual system, which can learn the structure of a given data set gradually, without any a priori information about that data. The local information, through the neighboring pixel's connection to the pixel, is considered in the MSLP structure instead of just the pixel gray level being taken into account. Hence, the system can improve its segmentation accuracy. To summarize, the algorithmic procedures of the MSLP segmentation network are described in the following.

Filter input image  $X$  using a lowpass filter to reduce high frequency noise, if necessary; the filtered image is denoted as image  $Y$ .

- (1) For an image with a bright background, find the valley from the right side of the histogram of image  $Y$ ; Otherwise, for an image with a dark background, find the valley from the left side of the histogram instead.
- (2) Transform the gray levels of image  $Y$  into the interval  $[0, 1]$  using the sigmoid function with the threshold value equal to the valley found above.

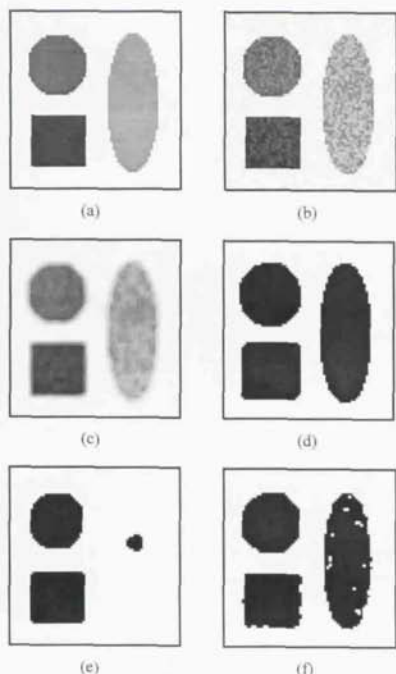


Fig. 5. (a) The synthetic image. (b) The noisy image. (c) The  $3 \times 3$  lowpass filtered image. (d) Processed by our method. (e) Processed by the MLP method. (f) Processed by the SOM method.

- (3) Employ the MSLP to segment image  $Y$  until the linear index of fuzziness  $v_i$  of the output nodes converges to a small enough value.

#### 4. Simulation Results

In order to check the effectiveness of the proposed technique, a synthetic image and a real world image are provided in this section for computer simulation. Simulations were run using a  $3 \times 3$  mask as the neighborhood region and  $9/2$  as the threshold value of the activation function, i.e., half of the number of corresponding input nodes. To avoid the pseudo noise effect, all of the initial weights were set to unity. The learning constant  $\eta$  was set as 0.01. For comparison, same image data were applied to the MLP and SOM approaches. The structure of the MLP used in the numerical simulations is the same as that used in the work of Ghosh *et al.* (1992) and the SOM ap-

proach was conducted under an optimal threshold ( $T$ ) selection (Ghosh and Pal, 1992) for best segmentation results.

Figure 5(a) is a synthetic image of dimensions  $72 \times 72$  containing three objects of different gray levels 50, 100, and 150. Random noise is also added to the image to test the noise immunity of this model. Fig. 5(b) is a noisy image with a signal-to-noise ratio (SNR), defined in the mean-square sense (Abutaleb, 1989), of 22.2 dB. To filter out high frequency noise, we applied a  $3 \times 3$ , spatial lowpass filter to the input image and obtained Fig. 5(c). Fig. 5(d) shows the result by the MSLP algorithm applied to Fig. 5(c), whose segmentation quality outperformed the other two methods. Fig. 5(e) indicates the resulting image using multilayer perceptron method (Ghosh *et al.*, 1992). The MLP-based algorithm can extract two objects successfully, but it extracts only a very small portion of the third object. The result of the SOM-based approach (Ghosh and Pal, 1992) at the optimal threshold  $T$  of 0.2 is shown in Fig. 5(f). From this figure, we can see that the segmented result at the edge of the square object is not accurate enough and a lot of holes are observable in the ellipse object. When we processed the same image with different gray levels of objects and SNRs, the best performance still consistently went to the MSLP approach.

In the following, a real world image, Fig. 6(a), of an aircraft model taken by a CCD camera, was segmented. The image has dimensions of  $67 \times 100$ . Fig. 6(b) is a noisy image with an SNR of 23.6 dB. The lowpass filtered image is shown in Fig. 6(c). The result using the proposed MSLP method is displayed in Fig. 6(d). Fig. 6(e), produced by the multilayer perceptron method, caused part of the object to disappear. Fig. 6(f) is the result produced by the SOM-based method with the optimal threshold  $T = 0.1$ , and from this figure, the weak noise immunity of the method is observed. The segmentation image quality of the MSLP method is much better than the other two approaches.

The segmentation accuracy rate and CPU processing time required on an HP model/712 workstation of the proposed, MLP-based, and SOM-based methods for these two patterns are summarized in Table 1. As Table 1 indicates, the MSLP not only produces higher segmentation accuracy but also requires shorter CPU training time than the other two methods. Aside from the difficulty in finding the precise valley from the histogram, the pixel's neighborhood connection of the MSLP net would change the output nodes toward a region having a similar statistical property. From the comparison above, we can conclude that our method outperformed the others in terms of both classification accuracy and computational time.



**Table 1. Comparison of three methods (MSLP, MLP, and SOM) in the segmentation accuracy rate and CPU processing time.**

Pattern	MSLP		MLP		SOM	
	CPU Time(s)	Accuracy Rate	CPU Time(s)	Accuracy Rate	CPU Time(s)	Accuracy Rate
Three Objects	8.31	95.87%	9.34	81.54%	307.75	93.77%
Aircraft	8.22	—	9.47	—	338.25	—

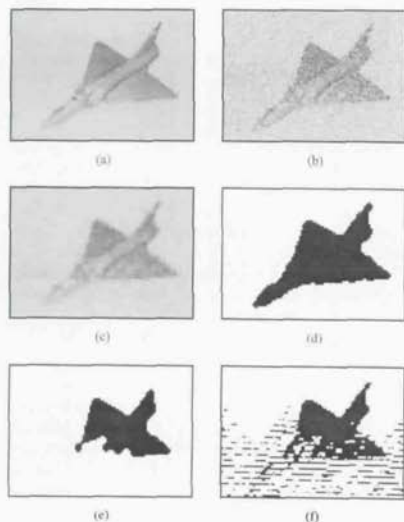


Fig. 6. (a) The real world image. (b) The noisy image. (c) The  $3 \times 3$  low-pass filtered image. (d) Processed by our method. (e) Processed by the MLP method. (f) Processed by the SOM method.

#### IV. INCORPORATING FUZZY IF-THEN RULES WITH THE MODIFIED SINGLE-LAYER PERCEPTRON NETWORK

In the previous section, the MSLP method provided satisfactory segmentation results on the simulation examples. The objects in these examples were somewhat self-compact in shape and significantly different in gray levels from the background. There are, however, still many ill-conditioned images that are not easily segmented. Such segmentation-difficult images usually have low-contrast ambiguous regions and the pixels of objects in these regions will frequently be identified as background and vice versa. These ambiguous regions, which will usually be categorized as background by conventional approaches,

can be easily identified as objects by human beings. Common sense, or a priori, knowledge of humans demonstrates a crucial capability in image segmentation, whereas numerical based algorithms through computing machines do not have. Furthermore, it is well known that human pattern recognition adopts more or less a set of linguistically descriptive knowledge rules because human reasoning and decision making are approximate and based on a larger and higher scale of viewpoint. Linguistic description usually is given in fuzzy terms, not only because they are the most common form of representation of human knowledge but also because our knowledge about many aspects is fuzzy. These facts suggest that research toward incorporating fuzzy logic structures into an image segmentation algorithm is not only interesting but also worthwhile. In this perspective, the proposed segmentation algorithm framework could be positively impacted by the incorporation of a fuzzy logic structure. Namely, if the proposed MSLP system is embedded with human-like visual recognition knowledge, in the form of fuzzy if-then rules to tackle the uncertainty and ambiguity existent in the low-contrast regions of an image, the quality of the segmentation solution could be further improved.

The MSLP network was used to segment FLIR ship images, Figs. 8(a), 10(a), and 11(a) (see Section IV-2), under a sea background, an essential task required in infrared searching and tracking systems (Bhanu and Holben, 1990) and also required in our on-going project as well. These ship images are difficult to segment because they have many low-contrast ambiguous regions, especially in the mast of the ship. The segmentation results of this kind of images by the MSLP network were not satisfactory in segmenting the mast from the background (see Figs. 8(b), 10(b), and 11(b)). The ship pixels in the ambiguous regions of the image are prone to converge to the wrong category of background. From our experiments, we observed that the threshold setting (4.5 as noted before) of the activation function of the output neurons in the MSLP network is an important factor, because it plays the vital role in separating the pixels into object or background. In the iteration

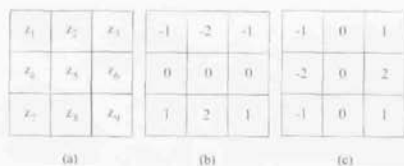


Fig. 7. (a) A 3x3 image region. (b) The mask used to compute  $G_x$  at the center point of the 3x3 region. (c) The mask used to compute  $G_y$  at that point. These masks are often referred to as Sobel operators.

of the network, once the learning process misclassifies a pixel initially, the iterative learning procedure can hardly recover it from the wrong category. This is especially true for pixels in the ambiguous regions. Therefore, if the threshold of the activation function of the output neuron can be adaptively adjusted (not necessarily set to 4.5 for all output neurons) according to the local characteristics of the corresponding image pixels, more effective segmentation results can be expected. Fuzzy sets are known to be very powerful in describing uncertain, imprecise, and vague characteristics and in representing human-like knowledge existing in a set, and likewise in an image. Fuzzy logic will be employed in the present MSLP segmentation net to manipulate this imprecise and uncertain information existing in the image so that the threshold of output neurons can adapt to the local characteristics of pixels. Such a learning threshold in each output neuron through fuzzy inference will lead to a better match to the local characteristics of the nine pixels in the window mask being proposed in the corresponding MSLP net. As a result, better segmented results can be obtained because the output neuron's threshold of the MSLP net is adapted to interpret the uncertainty and ambiguity in the image data using fuzzy rules.

### 1. Fuzzy Reasoning in Adjusting the Threshold of Output Neurons

This subsection will introduce a scheme that incorporates fuzzy if-then rules into the MSLP algorithm. Hence it will include human-like intelligence in helping the treatment of uncertainty and ambiguity encountered in the segmentation task. Two features which include the edge intensity and the vertical position of a pixel are defined as fuzzy sets to characterize the uncertain and vague variations existing in the ambiguous regions. Fuzzy if-then rules are then exploited to tune the threshold of each output neuron to make the MSLP net more adaptable to the local characteristics of the image data.

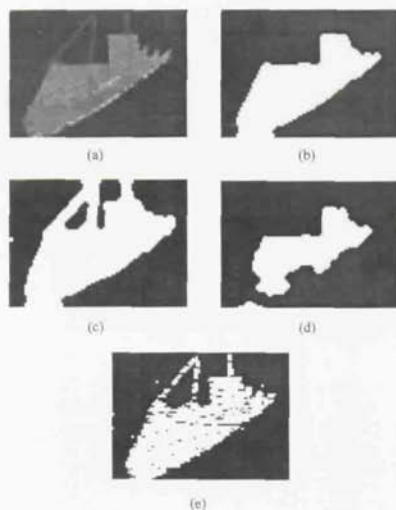


Fig. 8. (a) The first FLIR image. (b) The result of using only the MSLP method. (c) Processed by the MSLP method with fuzzy rules. (d) Processed by the MLP method. (e) Processed by the SOM method.

The first feature is the edge intensity measured by Sobel operators (Gonzalez and Woods, 1992). An edge is the boundary between two regions with relatively distinct gray levels. To detect an edge, using the gradient for image differentiation is the basic concept. The gradient  $\nabla f$  of an image  $f(x,y)$  at location  $(x,y)$  is the vector

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}. \quad (11)$$

Common practice is to approximate the gradient with absolute values using

$$\nabla f = |G_x| + |G_y|. \quad (12)$$

It follows from Eq. (11) that computation of the gradient of an image is based on obtaining the partial derivatives  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$  at every pixel location. The derivative may be implemented using the Sobel operator, as shown in Figs. 7(b) and 7(c), because it has the advantage of providing both a differentiating and a smoothing effect (Gonzalez and Woods, 1992). From Fig. 7, derivatives based on Sobel operator masks are

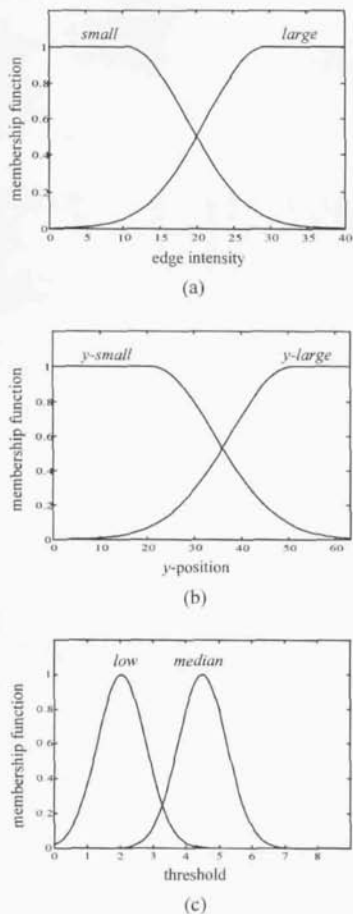


Fig. 9. (a) The membership functions of the edge intensity of Fig. 8(a). (b) The membership functions of the y-position of Fig. 8(a). (c) The membership functions of the threshold.

$$G_5 = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_6 = (z_5 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

where  $z_i$ 's are the gray levels of the pixels in the operator mask. The edge intensity at the center of a mask is computed by Eqs. (12) and (13). Naturally, the boundary of the mast is composed of edge pixels, which have a larger edge intensity. The second

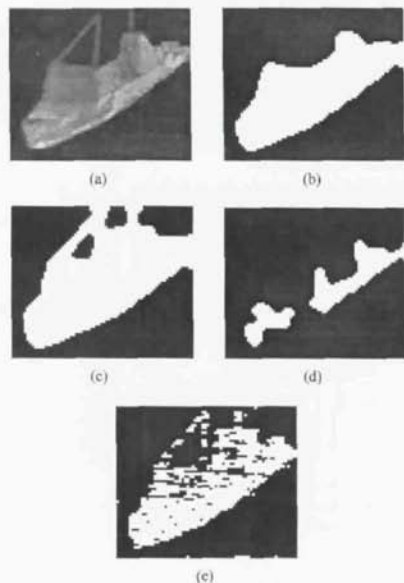


Fig. 10. (a) The second FLIR image. (b) The result of using only the MSLP method. (c) Processed by the MSLP method with fuzzy rules. (d) Processed by the MLP method. (e) Processed by the SOM method.

feature we used is the vertical position, i.e., y-position, of a pixel in the image. If a pixel is in the upper part of an image, then we say that the y-position of the pixel is *y-large*. This feature is acquired by observing the FLIR ship images of Figs. 8(a), 10(a), and 11(a) because all of these images contain ambiguous regions of mast at the upper portion of the images. Hence, we use y-position as the second feature to construct the fuzzy rules.

The threshold of the output node should be properly tuned so that the upper and ambiguous pixels, mostly referring to the mast, can be segmented as the ship. Since the mast pixels are in the region of dark background, to segment the mast as the ship, the threshold of the output neural nodes in this region must be reduced to make the pixels around the mast match the fact of dark surroundings. While, for the remainder of the image, the threshold of the output nodes can still be kept at the previously successful range, i.e., around 4.5. We try to construct fuzzy if-then rules to adapt the threshold of the output neurons in a manner meeting these observations.

Two fuzzy sets, *small* and *large*, were used for

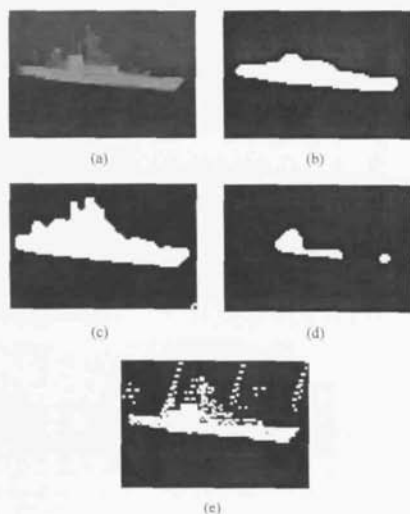


Fig. 11. (a) The third FLIR image. (b) The result of using only the MSLP method. (c) Processed by the MSLP method with fuzzy rules. (d) Processed by the MLP method. (e) Processed by the SOM method.

constructing the edge intensity. As to the attribute of the  $y$ -position, two linguistic labels composing the *universe of discourse* are  $y$ -small and  $y$ -large. Since these two features were used and each feature contains two labels, the following four rules can be obtained:

- Rule 1: IF the edge intensity of a pixel is *large* and its  $y$ -position is  $y$ -large,  
THEN the threshold of the pixel is *low*.
- Rule 2: IF the edge intensity of a pixel is *large* and its  $y$ -position is  $y$ -small,  
THEN the threshold of the pixel is *median*.
- Rule 3: IF the edge intensity of a pixel is *small* and its  $y$ -position is  $y$ -large,  
THEN the threshold of the pixel is *median*.
- Rule 4: IF the edge intensity of a pixel is *small* and its  $y$ -position is  $y$ -small,  
THEN the threshold of the pixel is *median*.

The consequence of the above four rules, *low* and

*median*, are the fuzzy sets used to represent the threshold. Only the consequent part of Rule 1 is *low*, the rest of the consequent parts are *median*. Rule 1 defines a rule concerning the upper and ambiguous part of the image. The mast, which has large edge intensity and large  $y$ -position, is in the region defined by Rule 1. This rule tends to adjust the neural threshold lower in this region to reflect the fact of a dark background, and the remaining three rules maintain the threshold at its original median range of 4.5.

Next we discuss how to determine the membership functions of the above linguistic labels. All of the membership functions involved in the rules above were chosen to be Gaussian functions, including a clipped version. For the edge intensity membership function *small*, its crossover point, membership degree equal to 0.5, was set to the value that the edge intensities for most, more than 80%, of the mast pixels are observable. The edge intensities of pixels smaller than that of the crossover value were collected and then the mean and variance were estimated. This mean and variance were used as the parameters of clipped Gaussian membership function *small* (Chi and Yan, 1995). Membership function *large* is chosen to be the symmetrical clipped Gaussian function of *small* using the crossover point as the reflection point. The crossover point of the  $y$ -position membership functions,  $y$ -small and  $y$ -large, was chosen at the position of the deck. The corner points starting full degree of these two membership functions were set at the positions  $\pm 15$  pixels with respect to the crossover point. For the membership functions of the consequent part, it is appropriate to set the center of the membership function *median* to be 4.5, the threshold assigned before. As to the *low* membership function, we first found pixels, called pro-rule-1 pixels, which fire Rule 1 with a firing strength larger than a preset value, 0.7 in our experiments. Note that each of these pixels corresponds to a  $3 \times 3$  window mask. Each of the nine pixel values in the mask of a pro-rule-1 pixel was first transformed by the sigmoid gray level transfer function, Eq. (5), and then added. This value is the threshold suitable for this pixel. The mean and variance of the threshold of these pro-rule-1 pixels were calculated. The center of the *low* membership function will be set to this mean. Both fuzzy sets, *low* and *median*, used this variance as their common variance in defining their respective membership functions.

For each image pixel, we used max-min composition and product implication in reasoning the above four rules and then the center of area (COA) method was employed (Zimmermann, 1996) for defuzzification. The inferred result, which has been adapted to the local characteristics of the output neuron pixel, was then used as the threshold value of each

**Table 2. The CPU processing time comparison of the MSLP with/without fuzzy rules, MLP, and SOM.**

Pattern	MSLP		MLP	SOM
	CPU Time(s)		CPU Time(s)	CPU Time(s)
	MSLP	MSLP+Rule		
Fig. 8(a)	8.15	12.28	10.10	332.86
Fig. 10(a)	7.51	10.69	7.78	256.46
Fig. 11(a)	8.03	12.11	10.01	329.54

output neuron pixel. And finally, the learning procedures, described in Section III, of the MSLP segmentation net were then followed to segment the image iteratively.

## 2. Simulation Results

In this subsection, the segmentation results of a modified single-layer perceptron network with or without incorporating fuzzy if-then rules were compared. With the threshold of the activation function of every pixel being selected by fuzzy rules, the accuracy of the segmentation has been greatly increased as illustrated by the three examples below. The results of the MLP- and SOM-based methods are also provided for comparison.

Using only the modified single-layer perceptron network with the threshold of every output neuron equal to 4.5, we obtained the segmented result, Fig. 8(b), of the input image Fig. 8(a). Based on the procedure mentioned in the previous subsection, the membership functions related to the above four if-then rules were obtained. The fuzzy sets concerning the edge intensity, the vertical position, and the threshold are shown in Figs. 9(a), 9(b), and 9(c), respectively. Applying the MSLP network to Fig. 8(a) with the threshold of output neurons being inferred by these fuzzy rules, produced the segmented image of Fig. 8(c). Comparing Figs. 8(b) and 8(c), we can see that Fig. 8(c) is much better than Fig. 8(b), especially in the mast region of the ship. Fig. 8(d) is the segmented result of Fig. 8(a) using the MLP-based algorithm. In this figure, the MLP-based approach produced an unsatisfactory result since several parts of this ship disappear. Applying the SOM-based method to Fig. 8(a), with the optimal threshold  $T=0.1$ , led to the resultant image of Fig. 8(e). As can be seen in this figure, even though the mast of the ship is segmented correctly, there are still many pixels which are wrongly segmented and this makes the ship full of holes and discontinuity.

Along similar lines, the MSLP networks with and without if-then rules were applied to Figs. 10(a) and 11(a). The resultant images are shown in Figs.

10(b) and 10(c), and Figs. 11(b) and 11(c), respectively. From these figures, a consistent improvement in the segmentation accuracy of the MSLP with fuzzy rules has also been obtained. A simulation study of the MLP- and SOM-based methods has also been done on Figs. 10(a) and 11(a). Figs. 10(d) and 10(e) are the results of applying the MLP- and SOM-based methods to Fig. 10(a), respectively. Applying the MLP- and SOM-based methods to Fig. 11(a) produced Figs. 11(d) and 11(e), respectively. The optimal thresholds  $T$  of the SOM method used for Figs. 10(a) and 11(a) were set to 0.1 and 0.15, respectively. We can see that the ship disappears partially in Figs. 10(d) and 11(d) again. Observing the results of Figs. 10(e) and 11(e), although there are many detailed parts being segmented correctly, we can observe that a compact ship image still cannot be obtained by the SOM-based method.

The amounts of CPU processing time required on an HP model/712 workstation for the proposed MSLP with/without fuzzy rules, MLP, and SOM methods for these three FLIR images are summarized in Table 2. As Table 2 indicates, the processing time needed for the SOM approach is significantly longer than ours and the MLP approach. Although the CPU processing time required for the MSLP with fuzzy rules is slightly longer than that for the MLP model, the segmentation results of the MSLP model, with or without fuzzy rules inclusion, have overridden the other two schemes.

## V. CONCLUSION

A self-supervised modified single-layer perceptron neural network for image segmentation is presented in this paper. The proposed segmentation network structure is simple and can be computed very fast. Our proposed model achieves very good segmentation accuracy and eliminates the problems of partial and/or noisy segmentation that arises in some recent approaches. Next, fuzzy if-then rules are integrated with an MSLP network, in which the threshold of the activation function of the output neurons is adaptively determined from fuzzy rule inference on

localized information of the image. With such localized tuning of this parameter, the threshold of output neurons can be best matched to the local information content of the image, and the segmentation network becomes much more accurate and robust. FLIR image segmentation examples demonstrate that incorporating such a fuzzy methodology with the MSLP network greatly enhances the results over the original MSLP numerical algorithm itself. This encouraging result suggests that integrating fuzzy logic into a designed numerical-based algorithm in some way can lead to a performance escalation. This is because the concepts and experience of human beings, which are modeled in fuzzy rules, and numerical algorithms can work together and cooperatively to solve a problem. Through the inclusion of fuzzy rules, the numerical-based algorithm is enriched with human-like knowledge as well as human information processing ability in tackling the imprecision, uncertainty, and ambiguity in the data. In this way, we can thus mimic human decision making and information processing ability along the execution lines of the numerical-based algorithm. The idea of incorporating fuzzy logic structure into a numerical algorithm is promising and is worthy of further investigation in other application areas, in addition to the image segmentation problem we have presented.

### NOMENCLATURE

$a(\bullet)$	activation function of a node.
$d_j$	desired value of the $j$ -th output node.
$E(\bullet)$	error function.
$g_s$	sigmoid gray level transfer function.
$t$	threshold of the sigmoid gray level transfer function.
$t_o$	approximate slope of $g_s$ .
$w_{ij}$	synapse weight between the $i$ -th input node and the $j$ -th output node.
$w_i^{(j)}$	synapse weight between the $i$ -th input node (of the $j$ -th output node) and the $j$ -th output node of image $X$ .
$\hat{X}$	bi-tone version of image $X$ .
$x_i$	$i$ -th input node.
$x_i^{(j)}$	$i$ -th input of the $j$ -th output node of image $X$ .
$x_{mn}$	gray level of the $(m,n)$ -th pixel.
$y_j$	actual value of the $j$ -th output node.
$\eta$	learning constant.
$\theta$	threshold value of a sigmoid activation function.
$\mu_X(x_{mn})$	membership value of image $X$ at the $(m,n)$ -th pixel $x_{mn}$ .
$v_l$	linear index of fuzziness.
$\nabla f$	gradient of image $f$ .
$\Delta w_{ij}$	the increment of synapse weight $w_{ij}$ .

$\Delta w_i^{(j)}$  the increment of synapse weight  $w_i^{(j)}$ .

### REFERENCES

1. Abutaleb, A.S., 1989, "Automatic thresholding of gray-level picture using two-dimensional entropy," *Computer Vision, Graphics, and Image Processing*, Vol. 47, pp. 22-32.
2. Bhanu, B., and Holben, R.D., 1990, "Model-based segmentation of FLIR images," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 26, pp. 2-11.
3. Blanz, W.E., and Gish, S.L., 1990, "A Connectionist classifier architecture applied to image segmentation," *Proceedings, 10th International Conference Pattern Recognition*, Atlantic City, NJ., Vol. 2, pp. 272-277.
4. Chen, J.L., Chang, J.Y., and Shieh, K.L., "2-D discrete signal interpolation and its image resampling application using fuzzy rule-based inference," *Fuzzy Sets and Systems*, to be published.
5. Chi, Z., and Yan, H., 1995, "Image segmentation using fuzzy rules derived from K-means clusters," *Journal of Electronic Imaging*, Vol. 4, pp. 199-206.
6. Ghosh, A., and Pal, S.K., 1992, "Neural network, self-organization and object extraction," *Pattern Recognition Letters*, Vol. 13, pp. 387-397.
7. Ghosh, A., Pal, N.R., and Pal, S.K., 1991, "Image segmentation using a neural network," *Biological Cybernetics*, Vol. 66, pp. 151-158.
8. Ghosh, A., Pal, N.R., and Pal, S.K., 1992, "Self-organization for object extraction using a multi-layer neural network and fuzziness measures," *IEEE Transactions on Fuzzy Systems*, Vol. 1, pp. 54-68.
9. Gonzalez, R.C., and Woods, R.E., 1992, *Digital Image Processing*, Addison Wesley, Reading, MA.
10. Haralick, R.M., and Shapiro, L.G., 1985, "Survey: image segmentation techniques," *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 100-132.
11. Keller, J.M., 1997, "Fuzzy set theory in computer vision: a prospectus," *Fuzzy Sets and Systems*, Vol. 90 pp. 177-182.
12. Lee, S.U., and Chung, S.Y., 1990, "A comparative performance study of several global thresholding techniques for segmentation," *Computer Vision, Graphics, and Image Processing*, Vol. 52, pp. 171-190.
13. Rananath, H.S., and Kuntimad, G., 1994, "Image segmentation using pulse coupled neural networks," *Proceedings, 1994 IEEE International*

- Conference on Neural Networks*, Orlando, FL., Vol. 2, pp. 1285-1290.
14. Sahoo, P.K., Soltani, S., and Wong, A.K.C., 1988, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing*, Vol. 41, pp. 233-260.
15. Zadeh, L.A., 1965, "Fuzzy sets," *Information and Control*, Vol. 8, pp. 338-353.
16. Zimmermann, H.J., 1996, *Fuzzy Set Theory and Its Applications*, Kluwer Academic, Norwell, MA.
17. Zurada, J.M., 1992, *Introduction to Artificial Neural Systems*, West, St. Paul.

Discussions of this paper may appear in the discussion section of a future issue. All discussions should be submitted to the Editor-in-Chief.

*Manuscript Received: Nov. 05, 1998*

*Revision Received: Apr. 07, 1999*

*and Accepted: July 30, 1999*

## 應用模糊邏輯於改良式單層認知影像分割網路

張志永 陳嘉麟

國立交通大學電機與控制工程學系

### 摘要

針對物體擷取，本文提出一個以模糊邏輯為基礎的改良式單層認知影像分割網路。我們選取一個S型灰階轉換函數，並配合輸入影像的灰階統計圖，將該影像圖素的灰階值映設到 $[0,1]$ 區間；接著我們採用輸出神經元之輸出值的線性模糊指標，作為影像分割系統的誤差函數，以發揮神經網路的學習能力。所提出的系統能成功地將物體擷取出來。為了更進一步提昇分割網路的能力，本系統再加入模糊若則法則，以適應性調整輸出神經元活化函數的門檻值，期能達到對影像局部特性做最佳的配合；我們依據圖素的邊界強度與垂直位置的模糊若則法則，適應地推論此門檻值。經由比較紅外線影像的模擬結果，所提出之改良式單層認知分割網路，在加入模糊若則法則後，可得到更佳的分割結果。此篇研究結果顯示，將人類知識以模糊法則的方式融入所設計之數值演算法中，對結果有進一步的提昇，且應可利用於影像分割以外的問題。

關鍵詞：神經網路，模糊邏輯，影像分割，非指導性學習。