OBJECT RECOGNITION OF MONOCHROMATIC IMAGES IN THE FREQUENCY DOMAIN

Risanuri Hidayat*)

ABSTRACT

   A large image usually consists of several smaller objects. People can recognize the objects automatically. The objects can be differented because they have different patterns. The aim of this research is for computer to recognize an object in image.

   The objects which will be recognized are transformed to the frequency domain, so spectrum frequencies are obtained for patterns, and these spectra are used as input. The objects are sampled by 4x4, 8x8, and 16x16 pixels. The object recognition uses an Artificial Neural Network method with step function.

   From this research, it is found that pattern recognition by spectrum frequency inputs is resistant to changes in position, such as rotation, translation, and reflection. Thus the same object is still recognized well, even though it has different position, or different place. Our experiment results show that pattern recognition in the frequency domain is more resistant to changes in position changing than pattern recognition in the spatial domain.

BACKGROUND

   Image processing is developing very quickly in the last two decades. One interesting aspect of this field is in image recognition. For example, in aerial-photos, one tries to recognize a forest, city, water, etc.

   Object recognition in the spatial domain uses pixel intensity as inputs. This kind of recognition requires a match in the objects position. If the object's position is changed, the object becomes unrecognizeable.

   Inputs in this research are in the frequency domain, i.e pixels are transformed to the frequency domain, so that they produce frequency spectra. These frequency spectra are used as inputs for the object recognition. The images are not only limited to aerospace images, but also in other fields, such as medical images, satellite images, and so on.

F OURIER TRANSFORMATION

   Images are transformed to the frequency domain by Fourier Transformation [Wintz,1987], according to the equation :

$$F(u,v) = \frac{1}{N}\sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x,y).e^{\frac{-2\pi j(ux+vy)}{N}} \qquad 1)$$

Note :
F(u,v) = frequency spectrum of an image at coordinate (u,v)
N = width of image
f(x,y) = pixel intensity of image at coordinate (x,y)

   Inverse Fourier Transformation transforms the spectrum frequency back to the spatial domain, that is, pixel intensity, based on the equation :

$$f(x,y) = \frac{1}{N}\sum_{v=0}^{N-1} \sum_{u=0}^{N-1} f(u,v).e^{\frac{-2\pi j(ux+vy)}{N}} \qquad 2)$$

   Objects that are similar lie in a certain pattern region, and objects that are different lie in the other patterns region [Duda,1973]. This property will be used in this research to recognize the objects.

   Before attempting to recognize, the computer system is given samples of the object as a reference and trained to recognize these. An Object will be recognized as a certain sample if it has a maximum similarity. The maximum value of correlation indicates that the object is similar to the sample [Schalkoff,1992].

ARTIFICIAL NEURAL NETWORK

   Below is the equation of a one layer artificial neural network with R inputs and S outputs, meanwhile Fig. 1. ilustrates about the network.

$$\begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_S \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1R} \\ w_{21} & w_{22} & \cdots & w_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S1} & w_{SR} & \cdots & w_{SR} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_R \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{pmatrix} \qquad 3)$$

*) Ir. Risanuri Hidayat, Lecture Electrical Engineering Dept., Fakultas Teknik, UGM
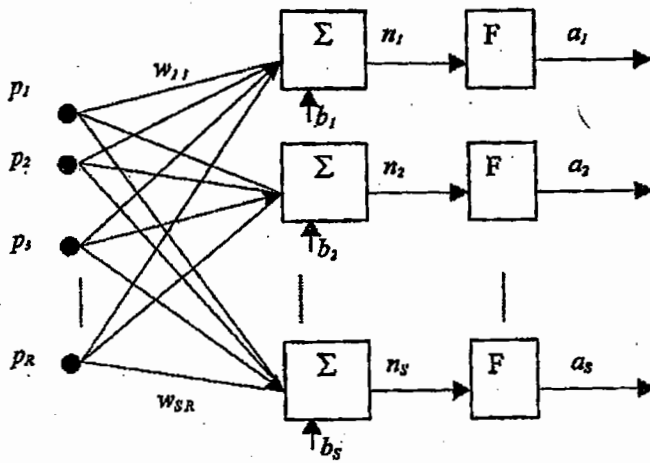
Figure 1. One layer Artificial Neural Network, which has R inputs and S outputs

The inputs are pj , and the outputs are aj , where aj are functions of nj. There are three kinds of the function, these are, step function, linear function, sigmoid function. The figure shows about the function. Step function yields a by $a = F(n)$ equation.
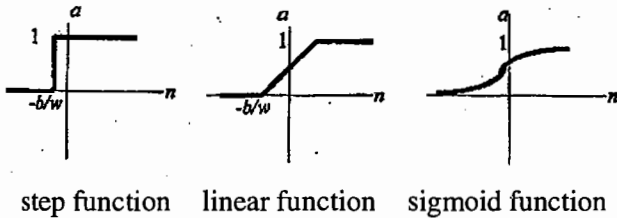

step function   linear function   sigmoid function

Figure 2. Functions of nj

The Training Process

The training process is the process to determine wij weights, so the wieghts are suitable between inputs and outputs. These are steps of the training process,
1. The values Wij weights are determined randomly, and maximum error is determined as well.
2. The weights above are applied for the $a = U(n = w*p + b)$
3. Output values are compared to the desired outputs t, and errors are used to modify the weights. according to the equations,

$$\Delta w = (t - a)p^T \qquad\qquad 4)$$
$$\Delta b = (t - a) \qquad\qquad 5)$$

$$w_{new} = w_{old} + \Delta w \qquad\qquad 6)$$
$$b_{new} = b_{old} + \Delta b \qquad\qquad 7)$$

where $e = (t-a)$ is the error. If $e <=$ maximum error, then the recursive process is halted. If $e >$ maximum error then the process goes back to the step 2.

METHOD OF THE RESEARCH

The research was done in the four steps. In the first step, the images data were collected. Second, the program of Fourier Transformation to transform the object into the frequency domain and to recognize the object was developed. The third step was the training step, and then followed the examination. The end of the research were analysed. The analysis included variations in position for the same object, and for different objects that had the same pattern because they lie on the different coordinates in the image. The analysis also compared object recognition between the frequency domain and the spatial domain.

RESULT OF THE RESEARCH

Two images were used in this research, namely, Cilacap and UGM. These images are shown on Figure 3.

The images were split in to block of 4x4, 8x8, and 16x16 pixels for recognition. The pixels represent objects, such as sea, land, grass, trees, etc. Figures 4. illustrate examples of 4x4, 8x8, 16x16 pixels, and in the frequencies domain.

From the research, it is visible that object recognition in the frequency domain is better than that in the spatial domain. The recognition validity in this research is measured by the equation :

$$\eta = \frac{\sum v}{A} x100\% \qquad\qquad 8)$$

with η as the percentage of successful recognition by the system. It is determined by comparing the number of objects that can be recognized successfully ($\sum v$) and the total number of the objects (A). The table below shows the results of the percentage of succesful recognitions, in 4x4, 8x8, and 16x16 pixels objects, both in the frequency domain and in the spatial domain.
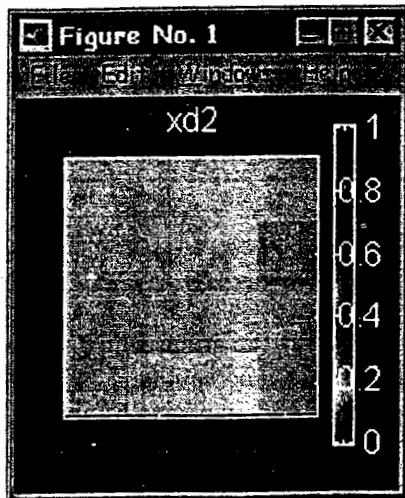
trees

grass
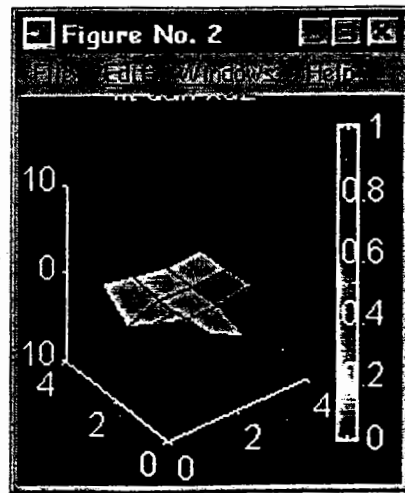
(a) UGM Central Office

land   rivers   seas
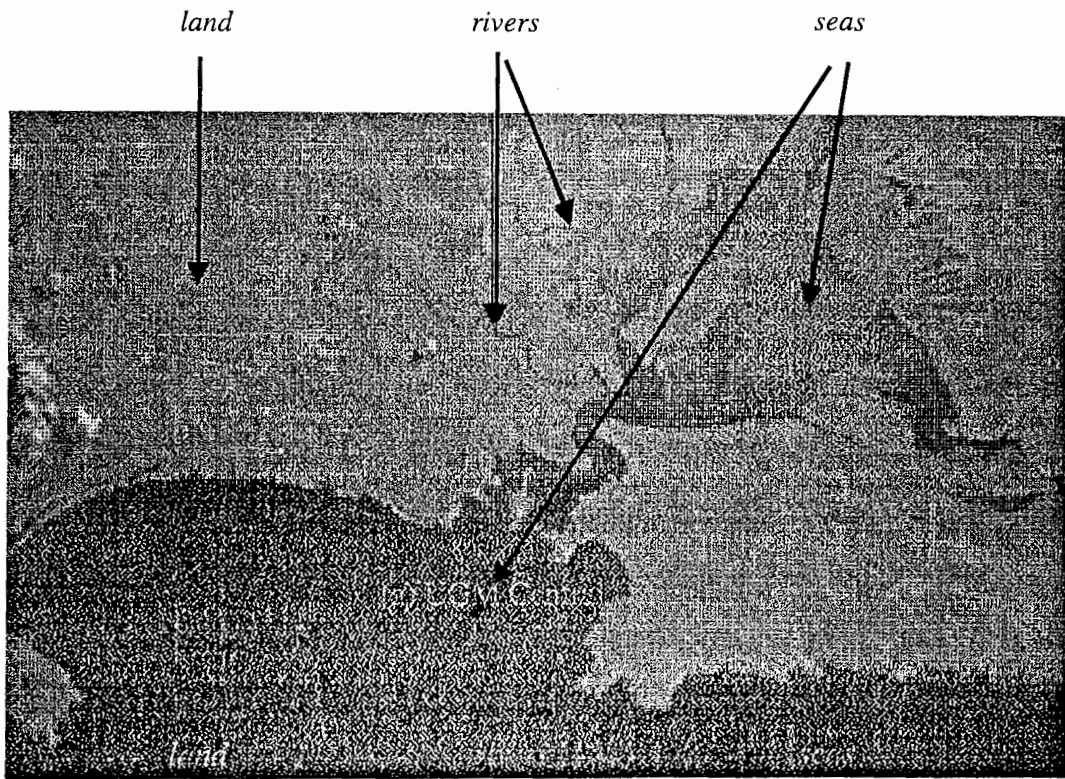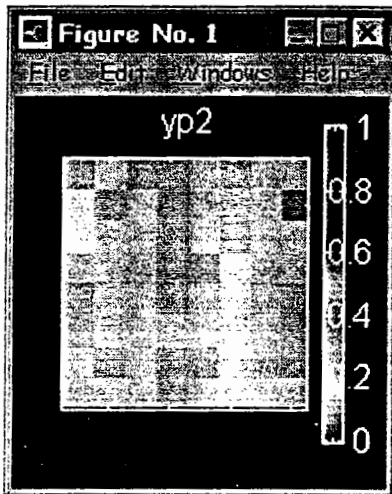
(b) Cilacap map

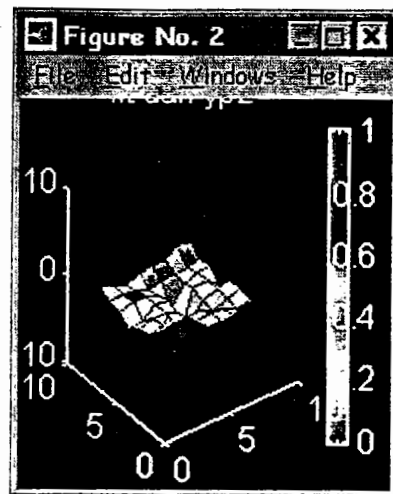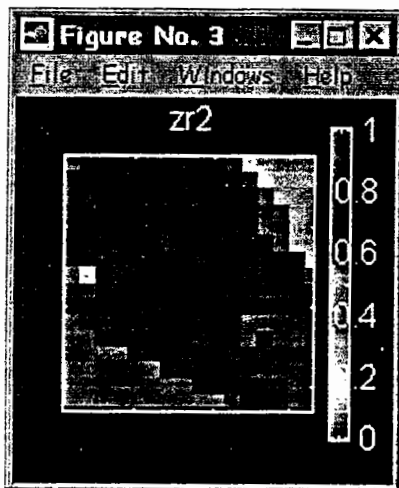Figure 3. Two images which are used in the research

(a) 4x4 pixels (land)
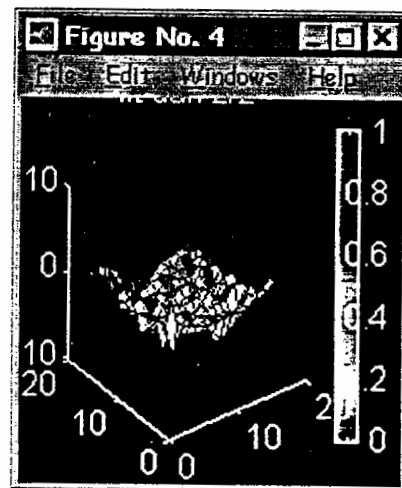
(b) 4x4 frequency domain (land)

(c) 8x8 pixels (trees)

(d) 8x8 frequency domain (trees)

(e) 16x16 pixels (grass)

(f) 16x16 frequency domain (grass)

Figure 4. Examples of 4x4, 8x8, and 16x16 pixel blocks and their frequency domain representations

Table 1. The percentage of successful recognition on the same sample

	4x4 %	8x8 %	16x16 %
Frequency domain	70,83	68,75	75,00
Spatial domain	50,00	31,25	41,67
Differences	20,83	37,50	33,33

Table 2. The percentage of successful object recognition on the different sample

	4x4 %	8x8 %	16x16 %
Frequency domain	50,00	33,33	60,00
Spatial domain	41,67	25,00	30,67
Differences	8,33	8,33	30,00

Below are the graphs which show the relationships between the percentage of successful object recognition and pixel size of object.

Below are the graph which show the relationships between the percentage of successful object recognitions and pixel size of object.
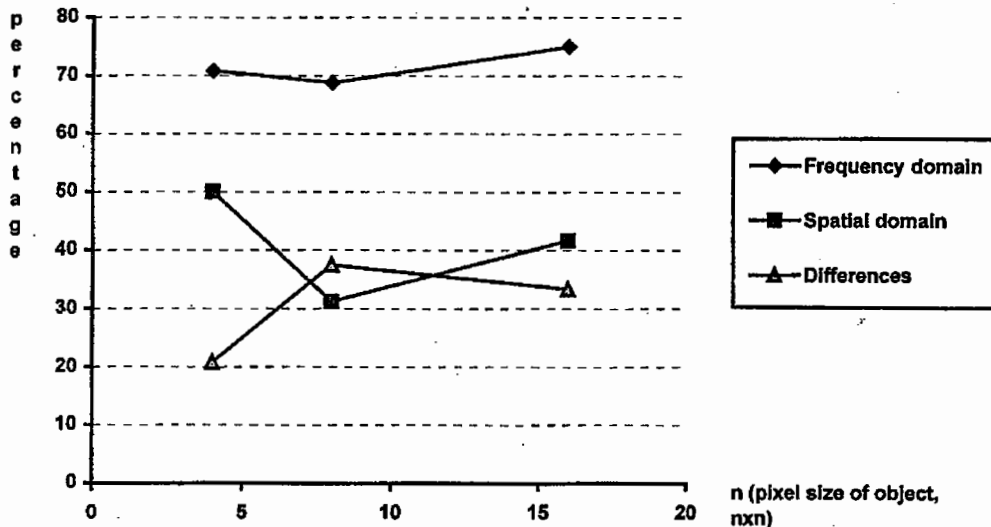

Figure 5. Graphs of the relationship between the percentage of successful recognition and size of objects each for the same sample.
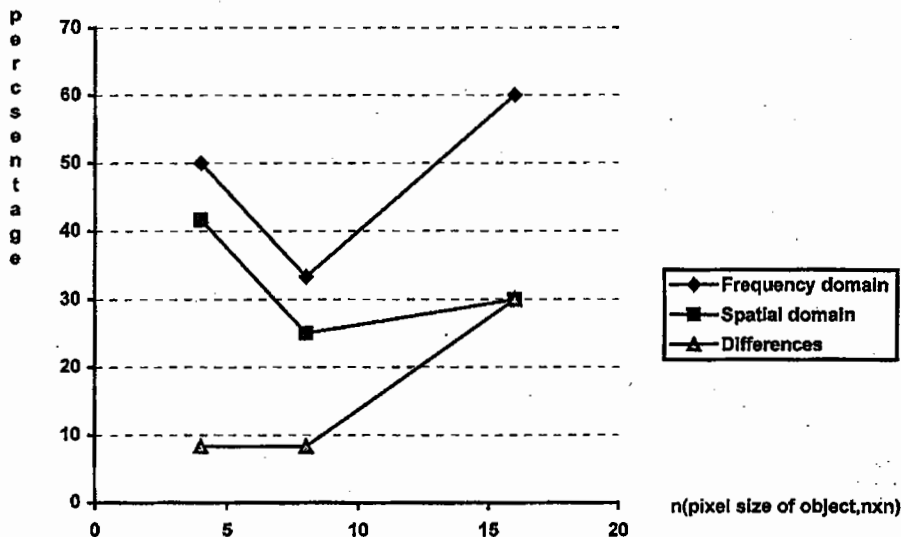

Figure 6. Graphs of the relationship between percentage of successful reconition and size of objects for the different sample

Based on the above graphs, we can comment that :

1) Regardless of the different samples, object recognition in the frequency domain is better than object recognition in the spatial domain.

2) For all samples, if the size of objects is increased then the success of object recognition in the frequency domain tends to increase, but the success of object recognition in the spatial domain tends to decrease.

3) Object recognition on the same samples is better than object recognition on the different samples, and on the different samples, object recognition in the frequency domain is still better.

CONCLUSION

1. Pattern recognition in the frequency domain is better than pattern reconition in the spatial domain.

2. If the size of objects is larger, the objects in the frequency domain will be recognized better, but the objects in the spatial domain will be recognized even worse.

3. Both in the frequency domain and spatial domain, the recognition of the different object samples is worse than that of the same samples, but the recognition is still better in the frequency domain.

REFERENCES

Demuth H, Beale Mark, 1994, *Neural Network TOOLBOX, For Use with MATLAB*, The MATHWORKS Inc., Massachusetts.

Duda Richard O.,1973, *Pattern Classification and Scene Analisys*, John Wiley & Sons, New York

Fu K.S., Gonzalez R.C., 1987, *Robotics, control, sensing, vision and intelligence*, McGraw-Hill, New York

Schalkoff Robert, 1992, *Pattern Recognition : Statistical, Structural and Neural Approaches*, John Wiley & Sons, New York

Wintz Paul, Gonzalez R.C., 1987, *Digital Image Processing*, Addison-Wesley, California.