Application of Fourier Descriptors and Neural Network to Shape Recognition

Leonid Tolstoy *leon_tol@yahoo.com*
Advisor: Dr. Hamed Parsiani *parsiani@ece.uprm.edu*
University of Puerto Rico, Mayagüez Campus
Mayagüez, Puerto Rico 00681-5000

Abstract

The paper represents an application of Neural Network and Fourier Descriptors to the problem of shape recognition of the 2-D binary image. The recognition is based on the features extracted from the Fourier Transformation of the shape of the original image. The shape of the image is described by small number of descriptors (Fourier Descriptors). This description of the shape makes picture translation-, shift-, and rotation-independent. The program in MatLab to find Fourier Descriptors and to apply them as input to feed-forward backpropagation Neural Network was developed. Several different variants of the Neural Network organization are tested and compared from the point of view of best matching to target.

1. Introduction

During the project, algorithms for boundary tracing of the shape of the binary image and for obtaining of Fourier features (Fourier Descriptors) were implemented in MatLab. The features obtained were used as an input to two-layers feed-forward backpropagation Neural Network. Neural Network was trained then to recognize several images and their rotated copies. Series of tests to find an optimal structure of network were accomplished, and their results are presented in this paper.

1.1 Background

The recognition is based on the contour of the pattern. There are many approaches to description and classification of contours of the pattern: statistical methods, based on method on moments, curve signatures, circular autoregressive model, etc. In this paper the shape of the pattern is described by coordinates of it contour in Cartesian (complex) coordinate system, and then the Discrete Fourier Transformation is applied to this data, to achieve the Fourier Descriptors (FD) of the shape. The advantage of FD is that it is possible to restore an original shape of the pattern, and they are easily computed because they are based on the well-developed theory of Fourier transformation.

2. Fourier Descriptors

Let us suppose that we have closed contour in a complex X-Y plane. A point moving around the contour generates a sequence of coordinates (x(m), y(m)), where m=1,2,…,N, and N is a number of points in the boundary. Since boundary is a closed curve,

$$x(N)=x(1),\ y(N)=y(1) \qquad (1)$$

We can represent each coordinate pair as a complex number:

$$a(n) = x(n) + i * y(n) \qquad (2)$$

The Discrete Fourier transformation, u(n), for this coordinate sequence a(n) is defined as follow:

$$u(n) = \sum_{k=0}^{N-1} a(k) e^{\frac{i2pkn}{N}} \qquad (3)$$

Inverse Fourier transform is given by:

$$a(k) = \frac{1}{N} \sum_{n=0}^{N-1} u(n) e^{\frac{-i2pkn}{N}} \qquad (4)$$

$$0 \le n \le N-1$$

This is a transformation from spatial to frequency domain. It is known that:

a) scaling the axes in spatial domain causes an inverse scaling in frequency domain;
b) rotating the image in the spatial domain causes Fourier representation to be rotated by the same angle;
c) shift in the spatial domain causes a liner shift in the phase component in the frequency domain;

To make Fourier Descriptors rotation- and shift- invariant, we have to use only absolute values of coefficients u(k), and in order to make them scale invariant, we normalize them by dividing each one by the first value.

Fourier Descriptors s(n) are computed as follows:

$$s(n) = \frac{|u(n)|}{|u(1)|} \qquad (5)$$

The most important Fourier Descriptors are the ones describing the lowest frequencies. To describe the shape, we are going to use only the first 10 descriptors.

An approximation of original shape could be obtained by making inverse Fourier transformation of these descriptors. If only the first M descriptors are used, then approximation of the shape is given by eq. (6).

$$a'(n) = \sum_{k=0}^{M-1} u(k) e^{\frac{i2pkn}{N}}, M < N \qquad (6)$$

Figures.1-3 show an example of this operation.

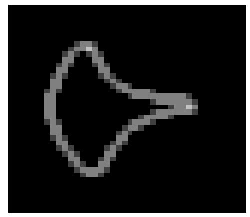


Figure 1 Original binary image

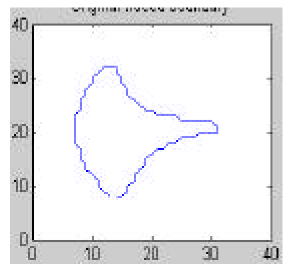


Figure 2 Traced Boundary of the object

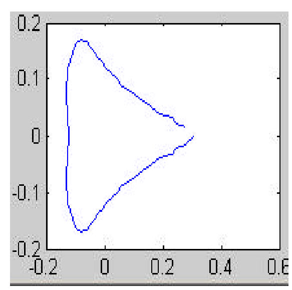


Figure 3 Shape restored from Fourier Descriptors

3. Edge Tracing

This research deals with binary images containing different objects. Object is represented by its shape. To get the Fourier Descriptors of a given shape, we first have to calculate the coordinates of all of the object's boundary points.

To do this, simple boundary tracing algorithm described in [Ramesh

95] was implemented as a MatLab function. This algorithm works for all objects with sizes greater than one pixel point.

a) Find a starting point S of the object, using scan.
b) Let the current point in the boundary be denoted by C. Set C=S and let the 4-neighbor to the west of S be B.
c) Let the eight 8-neighbors of C starting with B in clockwise order be N1, N2, …,N8. Find Ni, for the first i that is in S.
d) Set C=Ni and B=N_{i-1} .
e) Repeat steps c and d until C=S.

Fig.2 shows the result of applying this algorithm to the object from Fig.1

4. Neural Network

4.1 Organization

The recognition system was implemented using Neural Network. To recognize test pattern, two layers feed-forward backpropagation Neural Network was implemented in MatLab as a function. Fig.4 shows the structure of this network. To introduce Fourier Descriptors to the first layer of the network, 10 inputs are used. Neurons in the first layer use logsig transfer function. Initial weights and biases are generated randomly, and number of neurons in this layer changes from 1 to 10 during experiment, to find optimal structure.

Second layer includes only one neuron. Its transfer function is purelin, and output of this layer used to determine the pattern used on the input.

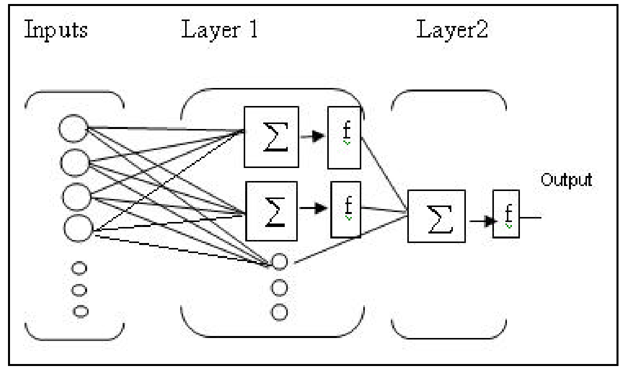


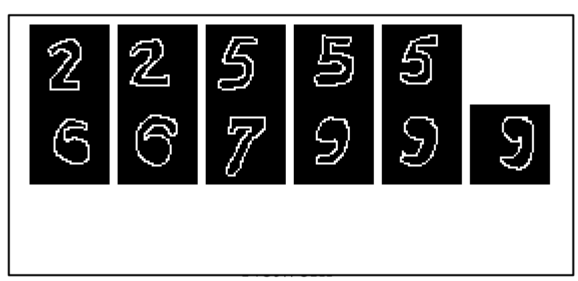
Figure 4 Neural Network

4.2 Training

Neural Network was trained using supervised learning to recognize patterns correctly. Ten FDs per image are calculated and assigned to a set A_i. Therefore, Ai, $\forall$ i=1, 2,3,4, each contain 10 FD's corresponding to the same image with rotations of 0, 90, 180, and 270 degrees, respectively.

In this project, three different images each with their corresponding sets A_i , B_i, and C_i were assigned to the inputs of Neural Network. Target values TA=1, TB=2, and TC=3 were assigned, for each respective image. Network was trained to minimize mean square error MSE, the average squared error between the network outputs and the targets Ti, $\forall$ i=A,B,C. Maximum number of training epochs was set to 700, and performance goal was 10^{-6} .

Figure 5 shows examples of images used in this work. Several tests carried out, each used a different set of three of these images.



From the previous results of using FD in recognition problems, it was reported [Man90] that there are difficulties of distinguishing between “2” and “5” , “6” and “9” due to their rotational invariance properties. The network was trained to recognize different forms of 5 from different forms of 2, and 6 from 9.

The number of neurons in the first layer was incrementally changed from 1 to 10, to find the optimal structure. Figure 6 shows a typical number of Neural Network epochs until convergence vs. number of neurons in the first layer. It can be observed from Fig. 6 that best results were achieved when number of neurons in the first layer were in the range from 4 to 7. In these cases, network converges faster and produces lesser error.

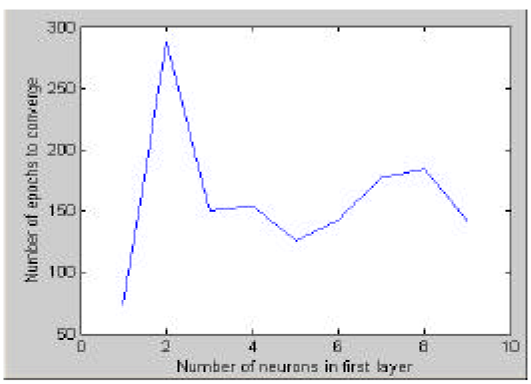


Figure 6 Number of epochs to converge vs Number of neurons

6. Conclusions

It was shown that FD contains a set of useful features, which can be used in shape recognition. The advantages of the FD are that they can be easily computed, and have shift, rotation, and scale invariant capabilities. The 2-layers feed-forward backpropagation Neural Network was implemented and trained to recognize different classes of shapes with or without rotation. It was shown that it can generally distinguish shapes even when their FDs are close due to rotation or shift invariance. Different structures of Neural Networks were tested and it was found that the optimum number of neurons in the first layer is in the range from 4 to 7.

Reference

[Czajka01] . Pacut, A.; Czajka, A. Recognition of human signatures Neural Networks, 2001. Proceedings. IJCNN '01. International Joint Conference on , Volume: 2 , 2001 Page(s): 1560 -1564 vol.2

[Ghorbel92] F.Ghorbel. “A complete and stable set of invariant Fourier descriptors for a random planar shape”. Image Processing and its Applications, 1992., International Conference on , 1992. Page(s): 278 -281

[Man90] Gary M.T.Man, Joe C.H.Poon. “An Enhanced Approach to Character Recognition by Fourier Descriptor”. Singapore ICCS/ISITA '92. 'Communications on the Move' , 1990
Page(s): 558 -562 vol.2

[Nghia00] Stanislaw Osowski, Do Dinh Nghia, “Neural Network for classification of 2-D patterns”. Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on , Volume: 3 , 2000 .Page(s): 1568 -1571 vol.3

[Ramesh 95] Ramesh Jain, Rangachar Kasturi, Brian G.Schunck. “Machine Vision”. McGraw-Hill 1995