A Radial Basis Function Neural Network to Recognize Handwritten Numerals with normalized moment features from skeletons

N. Venkateswara Rao¹
G. Rama Mohan Babu³
¹Dept. of Computer Science & Engineering
²Dept. of Information Technology
RVR & JC College of Engineering,
Guntur, INDIA.
{vraonaramala, rmbatram}@gmail.com.

Dr. B. Raveendra Babu²
²Dept. of Computer Science & Engineering
VNR Vignana Jyothi Institute of Engineering and Technology,
Hyderabad, INDIA.
rbhogapathi@yahoo.com.

Abstract—Handwritten numeral character recognition has been an intensive research in the field of artificial intelligence since many decades. This paper proposes a radial basis function neural network model for recognizing handwritten numerals. The geometric shape of handwritten numerals is described by computing a feature vector based on the skeleton of the images. The normalized central moment features are extracted from the skeleton of the images. Classification is performed with these normalized moment features by a radial basis function neural network. The novelty of this approach is that the normalized moment features from the skeletons gives good recognition rate than the contour images and thinned images with radial basis function neural network. The performance of the proposed work is computed from the error rate. Results of this proposed method on MNIST handwritten numeral database is reported.

Key words—Normalized moments, Skeleton, Character recognition, Feature extraction, Radial basis function.

I. INTRODUCTION

The character reorganization task can be generally separated into two categories: the recognition of machine printed characters and the recognition of hand written characters. Machine printed characters are uniform in size, position and width for any given font. Where as in handwritten, the characters are not in uniform, because they can be written in many different sizes, width and styles by different writers or by the same writer. Therefore, the recognition of handwritten characters is a much complex task than recognition of machine printed characters.

The recognition of the handwritten digits is one of the most active areas of research in computer science because of the high variability of writing styles. This system is used in many areas like Postal Identification Number (PIN) code, Bank cheque processing and Tax form verification and automation of historical documents, etc. Many researchers have proposed different methods of unconstrained handwritten characters recognition and are available in the literature. Some systems got higher accuracy and some got errors because of similar shape characters. In this approach, we proposed a normalized moment feature extraction method from the skeletons of the images to improve the recognition rate by using a radial basis function neural network.

An important approach for representing the structural shape of a region is to reduce it to a graph. The graph reduction may be accomplished by obtaining the skeleton of the shape region. A skeleton or Medial Axial Transformation (MAT) based feature extraction method was developed, which has produced an excellent performance for recognition. Moments of image skeletons provide efficient local descriptors and have been used extensively in image recognition applications. The advantage of moment features is its ability to provide invariant measures of shape images. Artificial Neural Networks are used in character recognition systems because of its distributive and massive parallel computation nature. Learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data. Cover (1965) proposed pattern-classification problem cast in a high-dimensional space is more likely to be linearly separable than in a low-dimensional space. Radial-basis functions were introduced in the solution of the real multivariate interpolation problem by Powell (1985). Broomhead and Lowe (1988) were the first to exploit the use of radial-basis functions in the design of neural networks.

II. RELATED WORK

In Brown et al. [1], a recognition system for the unconstrained hand printed numerals was proposed. This system used topological, geometrical and local measurements to identify the character or to reject the character as unrecognizable. This system yielded a recognition rate of 97% with a substitution error rate of 0.3% and a rejection rate of 2.7%. Suen et al. [2], proposed the recognition of hand written digits. Le Cun et al. [3] achieved good results by using convolution neural networks, which were specifically designed to deal with the variability of two dimensional shapes. For the recognition of handwritten numerals, the recognition rate with this method could be as high as 99.18% on the MNIST database. Mitchell and Gillies [4] used the tools of mathematical morphology to extract cavity features as the starting input for their specialized digit recognizers. A classification system was implemented by a symbolic model matching process. Simard et al [5], expanded the training set of the MNIST dataset by adding a new form of distorted data,
and the convolution neural networks were better suited for classification purposes. The recognition rate was achieved at 99.60%. Shi et al. [6] proposed a handwritten digit recognition system using the gradient and curvature of the gray character image in order to improve the accuracy of handwritten numeral recognition. The experiments were conducted on IPTP CDROM1, NIST SD3, and SD7 databases. The recognition rates could reach from 98.25% to 99.49%. Teow and Loe [7] proposed a handwritten digit recognition system based on a biological vision model. The features were extracted by the model, which could linearly separate over a large training set (MNIST). The high recognition rate was reported, where the error rate was 0.59%. Decoste and Scholkopf [8] proposed a handwritten digit recognition system where the prior knowledge about invariance of a classification problem was incorporated into the training procedure. Support Vector Machines (SVMs) were used as classifiers. The system achieved a low error rate of 0.56% when using this procedure with the MNIST dataset.

In the statistic feature domain, Hu [9] introduced the use of moment invariants as features for pattern recognition. In Krzyzak et al. [10], features were firstly extracted from the contours of numerals: 15 complex Fourier descriptors were extracted from the outer contours and simple topological features were extracted from the inner contours. These features were directly presented as the input of a three-layer ANN for recognition. In recent years, wavelet transform has been an emerging tool for feature extraction. In Chen, Bui and Krzyzak’s paper [11], a multi wavelet ortho normal shell expansion was used on the contour of the character to get several resolution levels and their averages. Finally, the shell coefficients were used as the features input into a feed-forward neural network to recognize handwritten numerals. Tao et al. [12] investigated the utility of several emerging techniques to extract features. The central projection transformation was applied to describe the shape of the characters. In Lee’s paper [13], Kirsch masks were adopted for extracting four directional local feature sets and one global feature set. In Yang et al. [14], high-order B-splines were used to calculate the curvature of the contours of handwritten numerals. Oliveira et al. [15] proposed a specific concavity, contour-based feature sets for the recognition and verification of handwritten numeral strings.

Nearest Neighbor classifier (NN) was well described in a book written by Duda et al. [16]. It uses a predefined distance to measure the similarity between a feature vector of the testing sample and a feature vector set for the class. The distance function can be Euclidean or Hamming distance. The polynomial discriminant classifier [17] assigns a pattern to a class with the maximum discriminant value. Tree classifiers are used to reduce the complexity in prototype matching. There are many well-known tree classifiers, such as CART [18] and C4.5 [19]. Ho [20] used the C.5 decision tree and reported good results on recognition problems. Hidden Markov Model (HMM) [21] consists of a set of states and the transition probabilities between consecutive states. Artificial Neural Networks (ANN), due to its useful properties such as: highly parallel mechanism, excellent fault tolerance, adaptation, and self-learning, has become increasingly developed and successfully used in character recognition [22-27].

III. METHODOLOGY

The proposed method was divided into three major steps: the first step is concerned with the generation of skeletons for the input images. The skeleton is very useful here because it provides a simple and compressed representation of the images that preserves many of the topological and size characteristics of the original images. The second step deals with how to use normalized central moments of the skeleton images to extract the features. Moments are the projections of the image function into a polynomial basis. The third step is the classification of the handwritten numerals from the features with Radial Basis Function neural networks which gives good recognition rate. A radial basis function neural network is a type of artificial neural network for application to problems of supervised learning.

A. Skeletonization

In our proposed system we had taken 1000 samples from MNIST data set and performed skeletonization which gives better normalized moment features. Skeletonization is a process of reducing the foreground regions in a binary image to a skeleton that largely preserves the extent and connectivity of the original image region while throwing away most of the original foreground pixels of the image.

The skeleton $S(A)$ of a set $A$ is defined as

$$S(A) = \bigcup_{k=1}^{\infty} S_k(A)$$

(1)

In terms of erosions and openings the skeleton of $A$ is defined by:

$$S_k(A) = (A \ominus E_k) - (A \ominus E_k) \ominus B$$

(2)

Where $B$ is the structuring element and $(A \ominus E_k)$ indicates $k$ successive erosions of $A$:

$$(A \ominus E_k) = (A \ominus E) \ominus (A \ominus E) \ominus \cdots \ominus E$$

(3)

$k$ times, and $k$ is the last iterative step before A erodes to an empty set, i.e.,

$$k = \max \{i | (A \ominus E_i) = \emptyset \}$$

(4)

So the skeleton $S(A)$ can be obtained as the union of skeleton subsets $S_k(A)$.

B. Feature Extraction

Features are the representation of the image pixels into an equivalent reduced representation. In this system, geometrical moments which are invariant under scaling, rotation and translation are computed. They are widely used in the process
of pattern recognition problems because of their discrimination power and robustness. The geometric shape of a handwritten digit is described by computing a feature vector based on moment invariants.

The two-dimensional moment of order \((p + q)\) of a density distribution function \(f(x, y)\) is defined as

\[
\mu_{pq} = \int x^p y^q f(x, y) \, dx \, dy
\]  

(5)

The moments that have the property of translation invariance are called central moments and are defined as

\[
\mu_{pq}^c = \int (x - \bar{x})^p (y - \bar{y})^q f(x, y) \, dx \, dy
\]  

(6)

Where \(x = \frac{\sum x}{N}\), \(\bar{y} = \frac{\sum y}{N}\) are the coordinates of the centroid of the image.

The normalized central moments, which are invariant to the scale of the image, can be defined as

\[
\eta_{pq} = \frac{\mu_{pq}^c}{\mu_{00}^c} \quad \text{Where} \quad \gamma = \frac{\mu_{00}^c + 1}{2}
\]  

The geometric moments, \(\phi_1\) to \(\phi_7\) with respect to translation, rotation and scaling invariants are defined as

\[
\phi_1 = \eta_{20} + \eta_{02}
\]  

(7)

\[
\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}
\]  

(8)

\[
\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2
\]  

(9)

\[
\phi_4 = (\eta_{20} + \eta_{02})^2 + (\eta_{11} + \eta_{11})^2
\]  

(10)

There are seven invariant moments from \(\phi_1\) to \(\phi_7\) to represent the image feature. We observed good recognition rate from the skeletons of the handwritten digits with first four moments. So we calculated the features of the images using these four moments. The extracted features from moment functions are used in training and classification is performed using neural network classifier. Results of the classification using these features are compared in terms of accuracy with respect to thinned images and contour images.

C. Radial Basis function Neural Network Model for classification

A radial basis function (RBF) is a real valued function whose values depend only on the distance from the origin, so that \(\Phi(x) = \Phi(||x||)\). It can also depend on the distance from some other point \(c\), called a center, so that \(\Phi(x, c) = \Phi(||x-c||)\). Any function \(\Phi\) that satisfies the property \(\Phi(x) = \Phi(||x||)\) is a radial function. The norm normally used is the Euclidian norm. A function is radial basis if its output depends on the distance of the input from a given stored vector. In a RBF network one hidden layer uses neurons with RBF activation functions describing local receptors. Then output node is used to combine linearly the outputs of the hidden layer neurons.

Radial basis function (RBF) networks typically have three layers: an input layer, a hidden layer and an output layer. The input layer takes four feature values computed from \(\phi_1\) to \(\phi_4\) for all the samples of the handwritten digits. In this layer the radial basis neurons calculates the vector distance between its weighted vector and the input vector multiplied by the bias.

The hidden layer size contains the number of samples taken for classification. In this layer it uses the pure linear transfer function which calculates the weighted input and the net input. The output layer contains the number of classes in the input samples shown in Fig. 1. The output \(\Phi = \mathbb{R}^n \rightarrow \mathbb{R}\) of the network is thus,

\[
\Phi(x) = \sum_{i=1}^{N} a_i \exp(\beta ||x-c_i||^2)
\]  

(11)

Where \(N\) is the number of neurons in the hidden layer, \(c_i\) is the center vector for neuron \(i\), and \(a_i\) are the weights of the linear output neuron. In the basic form all inputs neurons are connected to each hidden neuron. The norm is generally taken to be the Euclidean distance and the basis function is taken to be Gaussian

\[
\rho(||x - c_i||) = \exp(-\beta ||x-c||^2)
\]  

(12)

The Gaussian basis functions are local in the sense that

\[
\lim_{||x - c_i|| \to \infty} \rho(||x - c_i||) = 0
\]  

(13)

i.e. changing parameters of one neuron has only a small effect for input values that are far away from the center of that neuron.

RBF networks are universal approximators on a compact subset of \(\mathbb{R}^n\). This means that a RBF network with enough hidden neurons can approximate any continuous function with arbitrary precision. The weights \(a_i\), \(c_i\) and \(\beta\) are determined in a manner that optimizes the fit between \(\Phi\) and the data.

IV. RESULTS AND DISCUSSIONS

All experiments reported in this paper are conducted on handwritten digit samples taken from the MNIST database to evaluate the performance of the proposed scheme. There are 60,000 and 10,000 training and test samples in the MNIST database for handwritten English numerals. In our experiments we have taken randomly 100 samples from each class, a total of 1,000 samples to evaluate the performance of our method. We tested these normalized moment features on contour images, thinned images, and skeleton images. The recognition rate for contour images is 63.4%, for thinned images is 71.5%, and for skeleton images is 95.4%. The error rate is calculated

Fig. 1. Architecture of Radial Basis Function Neural Network
for all samples collected from the MNIST digit dataset. The error rate is the difference between the output generated by the radial basis net and the actual output given to the network. We observed an error rate of 36.6%, 28.5%, and 4.6% for contour images, thinned images and skeleton images respectively. Skeleton images gives low error rate, which is better than the contour images and thinned images as shown in Fig. 2. The normalized moment features from skeleton images gives good recognition rate with radial basis function neural network. For each digit we taken 100 samples and the number of samples correctly classified for each digit e detailed results of the recognition of each digit are shown in Table 1. Some of the digit samples collected from the MNIST numeral database is shown in Fig. 3.

![Graph showing error rate for digits 0 to 9](image)

**Fig. 2. Error Rate of each digit**

<table>
<thead>
<tr>
<th>Digit</th>
<th>Contour Images</th>
<th>Thinned Images</th>
<th>Skeleton Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>89</td>
<td>89</td>
<td>96</td>
</tr>
<tr>
<td>1</td>
<td>92</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>70</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>55</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>62</td>
<td>95</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>61</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
<td>72</td>
<td>96</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>64</td>
<td>97</td>
</tr>
<tr>
<td>8</td>
<td>68</td>
<td>75</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>71</td>
<td>99</td>
</tr>
</tbody>
</table>

**All** | **63.4%** | **71.5%** | **95.4%**

### Table 1. Experimental Results of Each Digit

<table>
<thead>
<tr>
<th>Digit</th>
<th>Contour Images</th>
<th>Thinned Images</th>
<th>Skeleton Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>89</td>
<td>89</td>
<td>96</td>
</tr>
<tr>
<td>1</td>
<td>92</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>70</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>55</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>62</td>
<td>95</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>61</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
<td>72</td>
<td>96</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>64</td>
<td>97</td>
</tr>
<tr>
<td>8</td>
<td>68</td>
<td>75</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>71</td>
<td>99</td>
</tr>
</tbody>
</table>

| All   | **63.4%** | **71.5%** | **95.4%** |

Some of the digit samples collected from the MNIST numeral database is shown in Fig. 3.
V. CONCLUSION

Handwritten numeral recognition is presented in this paper. The handwritten numeral recognition process is performed using skeletons of the images by computing normalized central moments. From the skeleton shape features we got better recognition rate than the contour and thinned images. From seven invariant moments we used only four moment features. For the handwritten numerals the first four moments given good recognition rate so we considered four features only. The radial basis function neural network is used on skeleton image moment features to improve the recognition results of two similar shaped digits for classifying unconstrained offline handwritten numerals. We observed a recognition rate of 95.4% and an error rate of 4.6% on subset of MNIST handwritten numeral digit data.

REFERENCES