ZONE BASED RELATIVE DENSITY FEATURE EXTRACTION ALGORITHM FOR UNCONSTRAINED HANDWRITTEN NUMERAL RECOGNITION

K.N. SARAVANAN, 1,2 Dr. R. ANITHA

1 Assistant Professor, Christ University, Bangalore, Karnataka, India
2 Prof. & Director, Muthayammal Engg. College, Rasipuram, Tamilnadu, India
E-mail: 1 saravanan.kn@christuniversity.in, 2 aniraniraj@rediffmail.com

ABSTRACT

The recognition of handwritten digit recognition has been a challenging problem among the researchers for few decades. This paper proposes a relative density feature extraction algorithm for recognizing unconstrained single connected handwritten numerals independent of the languages. The proposed method consists of four phases, namely, image enhancement (dilation), representation (zone based), feature extraction (relative density) and recognition (minimum distance classifier). The handwritten numerals must be enhanced with dilation, in order to connect the broken digits. After enhancement, the dilated binary images can be represented as a mid-point aspect ratio class interval values. The minimum distance classifier technique has been used to recognize the given numerals. The method yielded a satisfactory recognition rate of 92.85%, 99.28%, 98.95%, 98.72%, 99.29%, and 99.48% for Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya Handwritten Numerals respectively.

Keywords: Handwritten digit recognition, Feature extraction, Zoning, Minimum distance classifier, K-NN classifier.

1. INTRODUCTION

A numeral recognition system has a variety of commercial and practical applications in postal automation, bank cheque processing, automatic data entry, reading aid for the blinds, vehicle number plate recognition etc. The challenge of building a numeral recognition system which can match the human competence provides a strong motivation for research in this field. The recognition of handwritten numerals, characters, symbols and multivariants by computer has been a topic of intense research for almost five decades. A variety of algorithms, combination of features and feature sets have been adopted to try to capture essential information from handwritten numerals. Using different feature sets, and recognition algorithms researchers have obtained accuracy rate up to 99.10% ±0.40 for numerals of their respective languages, but the feature sets and recognition algorithms may not work for other languages because of variant in shapes, orientation etc.[1] The aim of the proposed research work is to recognize the numerals for language independence. The organization of the paper is as follows. In section 2 a brief review of earlier work and in section 3 database collections and pre-processing of handwritten digits. In section 4 representations of the handwritten digits and in section 5 and 6 zone based feature identification followed by feature selection for zone based representation has been described respectively. The proposed methodology has been described in section 7. The experimental results are reported in section 8 and the conclusion of the work is given in section 9.

2. A BRIEF REVIEW OF EARLIER WORK

The strategy used for pattern recognition, especially for handwritten digit recognition can be broadly classified into three categories, namely, Statistical approach, Syntactical approach, Hybrid approach. In statistical approach, a pattern is represented as a vector of an ordered, fixed length list of numeric features. For example, the outer densities of pixels for each of the direction are computed in four directions, namely, top, bottom, left and right. In structural or syntactical approach, a pattern is represented as a set of shapes of an unordered, variable length list of geometric features of mixed type. For example, shapes include end point, fork point and cross point. In hybrid approach, these two approaches are combined for representation of numerals and utilizing them for classification of unknown digits. For example, the Euler number was used as a pre-classification step and statistical features were used for classification. In
the Expert2 method is a time consuming process, it was recognized rate of 86.05% was achieved. In the second to the features extracted from these branches along with called branches. The pattern is then classified according to the features extracted from their skeletons and the other makes use of their contours, and has achieved recognition rate 84.85%. TUAN A. MAI and CHING Y. SUEN, 1990 [5], the computer generated feature is the primary feature used for the coarse classification of numerals, and the secondary features originates from human perception. This study has 16985 samples, 8500 were used in training, the remaining 8485 used in testing, from a training set 407 features were established, recognition result based on inference method and structural method 96.4%, 79.9% respectively. This method works for numerals of English language and thinning algorithms is a mandatory one. The database they have collected from U.S. Postal Services at different locations in the United States. J. T Lin and R.M. Inigo [6] used Back Propagation Neural Network with different structures and images are normalized to size 19 * 19. A lucid scope has been given in the paper as, increasing training numerals the recognition rate would proportionally increase. CHING Y. SUEN, CHRISTINE NADAL et.[7], four algorithms developed independently, three standard sets of data were prepared, each containing 2000 digits (200 of each class). Two of these sets, labelled A and B, are used for training and the third labelled T is used for testing. In the first method, the skeleton of a character pattern is decomposed into parts called branches. The pattern is then classified according to the features extracted from these branches along with the relations which exist among the parts. The recognized rate of 86.05% was achieved. In the second method, the primitives selected to represent the handwritten numerals are line segments, convex polygons and loops. The recognized rate 93.10% was achieved with this method. It has been mentioned that the Expert2 method is a time consuming process, it was considered inadvisable and unnecessary to use the method as the sole classifier. In the third method, the primary features are endpoint, fork point and cross point by the computer and the secondary features are obtained based on the human perception, the recognition rate of 92.95% was achieved. In the fourth method, features are extracted from the contours of the digits, a tree classifier is used and the recognition rate of 93.90% was achieved. Jun Cao, M. Shridhar, F. Kimura, M. Ahmadi [8], each image is divided into rows and columns, called zones. In each zone, a local histogram of the chain code is calculated. The feature vector is composed of these local histograms. They concluded that the neural net classifier outperforms the statistical classifier when the feature vector size is small. Both classifiers are effective in recognizing handwritten numerals with very low error rates and low rejection rates. Heutte L, Moreau J.V., Plessis B, Plagnaud J.L., Lecourtier Y [9], used four feature extraction families, first, concavity measurements, second, horizontal and vertical projections, third, polygonization of both the internal and the external numeral boundaries and fourth feature includes top, bottom, left and right extrema of the numeral. The system was tested on 100,000 digits extracted from the NIST database. 50,000 digits were used for the learning stage and 50,000 digits for the recognition rate. The recognition rate has been 98.05%. Yi Lu, Steven Schlosser, Michael Janezko [10], used Fourier descriptor to represent the handwritten digits and proposed five Fourier descriptor representations namely, SB, GCT, ERIM1, ERIM2 and DFT and have achieved 99.8%, 98.6%, 99.2%, 98.8% and 99.5% recognition rate respectively. They have also mentioned, a Fourier Descriptor based digit classifier can serve as a pre-classifier in a digit recognition system. Takahiko KAWATA [11], used four feature extraction families, first, concavity measurements, second, horizontal and vertical projections, third, polygonization of both the internal and the external numeral boundaries and fourth feature includes top, bottom, left and right extrema of the numeral. The recognition rate has been 98.05%. Feng Pan, Mike Keane [12], proposed a new set of aspect invariant moments, which are suitable for neural networks. Their experimental results have proved high recognition rate of 98.73% and low substitution rate of 1.06%. M.H. Shirali-Shahreza e.t.al [13], designed 32 segment bar mask for shadow coding Arabic Numerals irrespective of size and translation and obtained recognition accuracy of 97.8%. Jianming Hu and Hong Yan [14] proposed structural method for describing both printed and handwritten characters. The additional features considered are direction points (D) to characterize the curve changes in horizontal and vertical directions and the bend points (B) are used to detect the curvature changes of a curve in one direction. They have used
In this work, the MNIST standard database comprising of English numerals has been used for both training and testing. It has separate training set and testing set data. Total digits of 58631 used for training purpose and 9621 handwritten samples were used for testing purpose. To validate the numerals available in the database, enhancement and size limits in terms of tolerance limits have been applied. In the enhancement procedure, gaps between the pixels have been bridged using dilation process. One of the advantages of the morphological dilation process over the low pass filtering method is that the morphological method resulted directly in a binary image. Low pass filtering, on the other hand, started with a binary image and produced a gray scale image, which would require a pass with a thresholding function to convert it back to binary form. Statistical tolerance limits has also been applied to obtain the numerals with a standard height, width, and minimum and maximum aspect ratio. The lower tolerance limit can be obtained by applying the formula, \( \text{Mean} - K \times S \) and the upper tolerance limit = \( \text{Mean} + K \times S \), where Mean refers to average height of a particular numeral, \( K=3.291 \) (sample size > 1000 and confidence limit =99.9%) and S is the standard deviation. From the experiment conducted, the tolerance values obtained for validation of the numerals is tabulated in Table 1. Figure 1 shows the sample data set for training which have been obtained after dilation and tolerance limits procedures.
In addition to the MNIST database, we have collected Assamese, Devanagari, Manipuri, Malayalam and Oriya handwritten numerals as the primary database. Total number digits collected for Assamese language is 21745, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Devanagari language is 21314, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Manipuri language is 19573, 11700 digits for training set and 7873 digits for testing set. Total digits collected for Malayalam language is 20874, 14400 digits for training set and 6474 digits for testing set. Total digits collected for Oriya language is 21146, 15300 digits for training set and 5846 digits for testing set.

4. REPRESENTATION OF THE HANDWRITTEN DIGITS

Representation plays a vital role in handwritten recognition system. The selection of good representation without affecting the shape of the numerals leads to a better recognition rate. The ratio of Height / Width varies from 0.33 to 5.40 for handwritten numerals due to variance of the numerals of the same class. Based on the aspect ratio midpoint value of a numeral, the number of zones can be obtained along x and y axis. Table 2 shows the class interval, midpoint value and number of zones along x axis and y axis.

![Sample Data Set](image)

Figure 1 shows sample data set for training which was obtained after dilation and tolerance limits procedure.

In addition to the MNIST database, we have collected Assamese, Devanagari, Manipuri, Malayalam and Oriya handwritten numerals as the primary database. Total number digits collected for Assamese language is 21745, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Devanagari language is 21314, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Manipuri language is 19573, 11700 digits for training set and 7873 digits for testing set. Total digits collected for Malayalam language is 20874, 14400 digits for training set and 6474 digits for testing set. Total digits collected for Oriya language is 21146, 15300 digits for training set and 5846 digits for testing set.

5. ZONE BASED FEATURE IDENTIFICATION

The methodology followed for extraction of features based on the above representation scheme plays a vital role in the recognition rate of a numeral. Theoretically,
there are ‘M’ zones in the set A and ‘N’ zones in the set B. Then the set of all ordered pairs (a, b) where a ∈ A and b ∈ B, is called the Cartesian product and is denoted by A × B. Since a relation from A to B is precisely a subset of A × B, the set of all relations from A to B is precisely the set of all subsets of A × B. Therefore, the number of relations from A to B is equal to the number of subsets of A × B. Since, number of elements in A is M zones and number of elements in B is N zones, we have A × B = M × N zones. Therefore, A × B has \(2^{MN}\) number of subsets [25]. Let A = \{1, 2, 3…M\} and B = \{1, 2, 3…N\} and let the relation R from A to B be defined as relational density of the zone (feature), i.e., Relational density of the zone = total number of white (black) pixels in the given area / (total number of white pixels + total number of black pixels). Therefore, \(2^{MN}\), relational density zone can be obtained from the given M * N zones. [1]

Table 3 Depiction Of Three Row Zones Along Y Axis And Four Column Zones Along X Axis

<table>
<thead>
<tr>
<th></th>
<th>Z1</th>
<th>Z2</th>
<th>Z3</th>
<th>Z4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z5</td>
<td>Z6</td>
<td>Z7</td>
<td>Z8</td>
<td></td>
</tr>
<tr>
<td>Z9</td>
<td>Z10</td>
<td>Z11</td>
<td>Z12</td>
<td></td>
</tr>
</tbody>
</table>

6. FEATURE SELECTION FOR ZONE BASED REPRESENTATION

From feature identification procedure, the number of features we could extract is exponential, so, we have narrowed down the number of features based on four primitive descriptors, namely, relational density along x axis, along y axis, along the major diagonal and along the minor diagonal. Hence, the numbers of features become polynomial W. For example, the number of features generated using midpoint value method for the zone size (53, 10) is 83145.

7. PROPOSED METHODOLOGY

The proposed method uses relative density of the pixels as the main feature in the classification process. Initially, enhancement is carried out for connecting broken numerals using dilation method, the number of zones selection has been done using midpoint class interval values. After selecting the number of zones along x and y axis, standard block size namely (16 *16) or (8*8) of the zone has been assigned and the image is re-sized for feature extraction using interpolation method. This procedure will be done for all similar images of same zone size along both axes. The relative densities of the zone are computed for 1*1, 1*2, 1*3……up to M * N. The mean relative density of the pixels for each group (cluster) are found using hierarchical clustering algorithm and finally stored in the feature vector library. To classify the numeral, we have used minimum distance classifier and the nearest feature vector is estimated. The Euclidean distance between the feature vector and the mean feature vector is determined and assigned the numeral class to the nearest mean vector.

Algorithm for Training Numerals

Input: Binary Numeral Image from the database
Output: Store Feature Vector in the Library
Step 1: Pre-process the image to connect the broken digits (dilation)
Step 2: Perform labelling connected component algorithm to crop the image.
Step 3: Obtain the number of zones along x and y axis based on the midpoint aspect ratio interval
Step 4: Resize the image
Step 5: Compute the relative density of the pixels for 1≤i≤M and 1≤j≤N
Step 6: Repeat the steps 1 to 5 to group the images using Hierarchical Clustering Algorithm and store the mean relative density feature vector in the library.

End of the Training Numerals Algorithm.

Algorithm for Testing a Numerical

Input: Isolated Binary Numeral Image
Output: Recognition of the numeral
Step 1: Pre-process the image to connect the broken digits.
Step 2: Perform labelling connected component algorithm to crop the image.
Step 3: Obtain number of zones along x and y axis based the midpoint aspect ratio interval from the global database
Step 4: Resize the image
Step 5: Compute the relative density of the pixels for 1≤i≤M and 1≤j≤N
Step 6: Calculate the Euclidean distance between the input feature vector and the mean feature vectors and assign the class to the nearest mean vector

End of the Testing a Numerical Algorithm.

On the basis of the proposed algorithm, the flow of the system is provided in the following figures and tables.
The grouped images based on midpoint aspect ratio class interval

In the Experiment I, we have achieved 92.85% as recognition rate for the Latin numerals. In the Experiment II, 99.28% recognition rate has been achieved in Assamese language. In the Experiment III, 98.95% recognition rate has been achieved in Devanagari script. In the Experiment IV, 98.72% recognition rate has been obtained in Manipuri language. In the Experiment V, 99.29% has been obtained for Malayalam language. In the Experiment VI, 99.48% has been obtained for Oriya language.

Table 4 Sample Confusion Matrix For Latin Numerals (Experiment I)

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Accuracy %</th>
<th>#recognized Data</th>
<th>#Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>952</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>99.06</td>
<td>952</td>
<td>961</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1101</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>97.86</td>
<td>1101</td>
<td>1125</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>0</td>
<td>904</td>
<td>32</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>11</td>
<td>13</td>
<td>0</td>
<td>92.15</td>
<td>904</td>
<td>981</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>888</td>
<td>0</td>
<td>29</td>
<td>1</td>
<td>7</td>
<td>33</td>
<td>8</td>
<td>90.70</td>
<td>888</td>
<td>979</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>857</td>
<td>3</td>
<td>11</td>
<td>3</td>
<td>6</td>
<td>46</td>
<td>91.17</td>
<td>857</td>
<td>940</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>48</td>
<td>5</td>
<td>706</td>
<td>12</td>
<td>1</td>
<td>24</td>
<td>2</td>
<td>87.59</td>
<td>706</td>
<td>806</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>886</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>97.46</td>
<td>886</td>
<td>909</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>926</td>
<td>4</td>
<td>41</td>
<td>92.32</td>
<td>926</td>
<td>1003</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>9</td>
<td>8</td>
<td>27</td>
<td>9</td>
<td>33</td>
<td>7</td>
<td>8</td>
<td>818</td>
<td>16</td>
<td>85.92</td>
<td>818</td>
<td>952</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>16</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>7</td>
<td>910</td>
<td>94.30</td>
<td>910</td>
<td>965</td>
</tr>
</tbody>
</table>

Mean 92.85

To prove the eminence of the proposed algorithm a comparative analysis with the popular and best algorithms has been attempted. The table 8 shows the comparison of proposed with existing algorithms for Latin Numerals.

8. EXPERIMENTAL RESULTS

For the purpose of experimentation, Handwritten Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya numerals have been used.
Table 5 Comparison Of Proposed Method With The Existing Methods For Latin Numerals

<table>
<thead>
<tr>
<th>References</th>
<th>No. of samples in the Data set</th>
<th>Feature Extraction Method</th>
<th>Classifier</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>9621</td>
<td>Selective Subset of Relative Density Feature Extraction</td>
<td>Minimum Distance Classifier</td>
<td>93.02%</td>
</tr>
<tr>
<td>[2]</td>
<td>300</td>
<td>Structural features</td>
<td>Decision tree</td>
<td>96%</td>
</tr>
<tr>
<td>[3]</td>
<td>Not available</td>
<td>Contour based on the direction and curvature for a character contour</td>
<td>Basic Multi-layered neural network with three kinds of inputs</td>
<td>98.3%, 98.8%, 99.1%</td>
</tr>
<tr>
<td>[4]</td>
<td>2000</td>
<td>Structural features from their skeletons, other makes use of their contours</td>
<td>A tree classifier</td>
<td>84.85%</td>
</tr>
<tr>
<td>[6]</td>
<td>500</td>
<td>Not available</td>
<td>Back Propogation Neural network</td>
<td>70%</td>
</tr>
<tr>
<td>[8]</td>
<td>11000</td>
<td>Local histogram of the chain codes</td>
<td>Statistical and neural network</td>
<td>96.02%</td>
</tr>
<tr>
<td>[11]</td>
<td>10000</td>
<td>concavity measurements, horizontal and vertical projections, polygonization of both the internal and the external numeral boundaries and top, bottom, left and right extrema of the numeral</td>
<td></td>
<td>98.05%</td>
</tr>
<tr>
<td>[12]</td>
<td>8000</td>
<td>Aspect invariant, moment order</td>
<td>Neural networks</td>
<td>98.73%</td>
</tr>
<tr>
<td>[13]</td>
<td>2600</td>
<td>Shadow coding</td>
<td>Probabilistic Neural Network</td>
<td>97.8%</td>
</tr>
<tr>
<td>[14]</td>
<td>10000</td>
<td>Primitive coding and global description</td>
<td>Structural method</td>
<td>99.7%</td>
</tr>
<tr>
<td>[15]</td>
<td>17000</td>
<td>Projections and contour histograms</td>
<td>Knn</td>
<td>99.54%</td>
</tr>
<tr>
<td>[17]</td>
<td>195000</td>
<td>Foreground features namely, transitions from background and foreground, Background features based on concavity information</td>
<td>HMM</td>
<td>97.9%</td>
</tr>
<tr>
<td>[19]</td>
<td>Not available</td>
<td>Statistical and structural features</td>
<td>Neural network</td>
<td>95.7%</td>
</tr>
<tr>
<td>[20]</td>
<td>100</td>
<td>Linear discriminant analysis</td>
<td>Improved LDA and Bhattacharyya Distance</td>
<td>96.93%</td>
</tr>
<tr>
<td>[22]</td>
<td>18467</td>
<td>Holes, vertical up and down cavities and endpoints, horizontal left, right cavities, horizontal left and right end points</td>
<td>Knn</td>
<td>85.26%</td>
</tr>
<tr>
<td>K. N. Saravanan &amp; Dr. R. Anitha</td>
<td>9621</td>
<td>Relative density of pixels</td>
<td>Minimum Distance Classifier</td>
<td>92.8567%</td>
</tr>
</tbody>
</table>

9. CONCLUSION

In this paper we have proposed a relative density feature extraction algorithm for the recognition of single connected component numerals. Minimum distance classifier has been used for classification. The recognition rate of 92.8567%, 99.28%, 98.95%, 98.72%, 99.29%, 99.48% has been achieved for Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya numerals respectively. The table shows the previous work cited has better performance than the proposed method for Latin Numerals. The reason could be the dataset used and the size of the database. Our future work aims to decrease the number of features and combine other feature set to improve the recognition rate and apply the proposed algorithm to other language numerals and check the robust nature of the algorithm.

10. LIMITATIONS AND FUTURE ENHANCEMENT

The maximum number of features generated using midpoint value method for the zone size (53, 10) is 83145 and (37, 10) is 41665, though it is polynomial, but it occupies more space to store the floating point values. Our future work aims to decrease the number of features which are generated using the proposed algorithm to reduce the space and apply the proposed algorithm to other language numerals.

11. ACKNOWLEDGEMENTS

The authors are grateful to Dr. Nangegowda, Prof. Chandrasekhar, Dr. Fr. Jossy, Dr. Nachamai M. and...
REFERENCES


