

Handwritten Text Recognition Using Information Energy And Neural Network

Kanchan Keisham¹, Sunanda Dixit²

¹Dayananda Sagar College of Engineering, M. Tech, (CSE), ISE Department, Bangalore, India

²Dayananda Sagar College of Engineering, Associate Professor, ISE Dept, Bangalore, India

Abstract: In handwritten character recognition one of the most challenging task is segmentation. This is mainly due to different challenges like skewness of textlines, overlapping characters, connected components etc. This paper proposes a character recognition method of handwritten English documents. The textlines are segmented based on information energy that is calculated for every pixel in the scanned document and the characters are recognized using artificial neural network (ANN).

Keywords: line segmentation, word segmentation, character segmentation, information energy, neural networks, recognition etc

1. Introduction

Text recognition plays an important role in document image processing. The segmentation of machine printed documents are performed more effectively than handwritten documents. This is mainly because in machine printed documents the text are straight, uniform alignment, uniform spacing between the characters and stable alignment etc. However in handwritten documents many challenges occur due to the following reasons: 1) Multi-orientations 2) skewness of the textlines 3) overlapping of characters 4) connected components etc. The segmentation methods for machine printed and handwritten documents are different. Some of the methods for machine printed are as follows. [4] Xujun Ping uses an approach to segment printed text using Markov Random fields (MRF). This method separates noise, printed and handwritten text. [5] Shetty uses an approach to segment and label by using Conditional random fields (CRF). One advantage of this method over MRF is that it does not assume the conditional independence of the data. Handwritten character recognition can be either offline or online.

In offline character recognition the text are written in paper or document are automatically converted into letter codes which are then used within the computer. In online character recognition it involves the automatic conversion of text which are written on special digitizer or PDA. Offline character recognition is more difficult since different people have different writing styles. Offline handwritten character recognition involves segmentation of textlines, words and characters which are then recognized. Many methods have been proposed for text, line and word segmentation. Studies have shown that text and word segmentation have achieved very high accuracy.

This paper proposes an English document handwritten character recognition using information energy that are used to formulate an energy map. The energy map is then used to construct seams to separate the text lines [1]. The seam passes through minimum energy levels. The characters are then segmented and recognized using feed forward artificial neural network [5].

The rest of the paper is organized as follows. Section 2 gives related work. The proposed method is explained in Section 3. Section 4 gives the experimental result and the conclusion is given in section 5.

The general diagram for types of character recognition is given below:

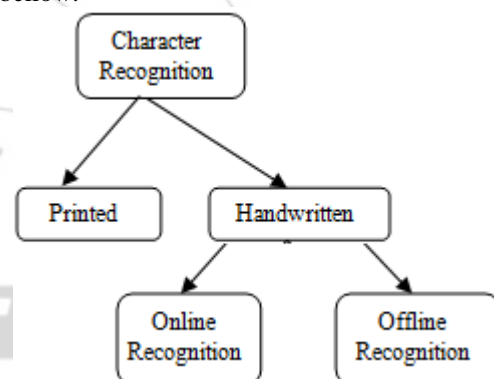


Figure 1: Block diagram for character recognition

2. Related Work

This section gives a summary of works that has been done for line, word and character segmentation and recognition. [2] performs textline extraction using Hough transform. This method also involves postprocessing in order to extract the lines which the hough transform fails to do so. [3] uses Minimum spanning tree [MST] to segment the lines. In this method the characters of the same textline will be present on the same subtree. [6] proposes a method based on Active Contour (Snakes). In this method the snakes are adapted over the ridges of the textline. The ridges pass through the middle of the textlines. [7] uses a MLP based classifier in order to recognize the segmented words. It is a holistic approach as it recognizes the word from its overall shape. [8] uses a method known as Transformation based error driven learning. This algorithm generates a set of rules and scores are evaluated. This method applies the rules to different sentences repeatedly until the scores remain unchanged. [9] uses pivot based search algorithm to recognize the words. This method selects some words from the training set as pivot words. The remaining words are then compared with the pivot words and similarity is computed.

[10] uses Convex Hull to determine the initial segmentation points and SVM classifier is used to classify the segmented characters.[11] uses multiple Hidden Markov Model(HMM).Multiple HMM are used since different characters can be written in different writing styles.[12] uses horizontal,vertical and radial histogram to segment and recognize the characters.

3. Proposed Method

Our proposed method computes line segmentation based on information energy of the pixels[1] and segmented characters are recognized based on feed forward Artificial Neural Network(ANN). Our schematic diagram for the proposed method is given as follows:

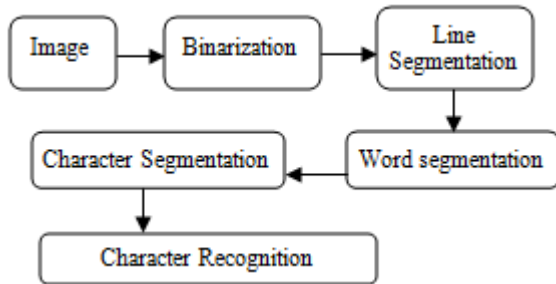


Figure 2: Block diagram for proposed method

In a document image every pixel has certain level of information present in it. The blank spaces between the text lines represent low information pixels. If high energy pixels are removed it leads to information loss. The idea is to generate an energy map based on the information content of each pixel.From the computed energy map the location of the text lines can be found out.

The energy value of each pixel is computed as follows[]:

$$E(i, m) = \left| \frac{I(i, m+1) - I(i, m-1)}{2} \right| + \left| \frac{I(i+1, m) - I(i-1, m)}{2} \right|$$

The energy map between two lines is given by $E_a = E(K, :) \in R^{d \times e}$ where $K = \{i, \dots, m\}$ (1)

Which gives the set of co ordinates between the start and end of energy block map. Here $d = m - i + 1$

Now based on the energy map a seam is generated by dynamic programming [13]. The image block is transverse and the minimum energy N_b is computed.

The energy N_b is computed backwards.Now the optimal seam is finally generated which segments the text line. After the lines are extracted and skew and slant are corrected the next step is word segmentation. This paper uses vertical histogram to identify the segmentation points. Sometimes the gaps between the characters may be larger than the gaps between words. To segment the

words in such cases K-means clustering is used[14] with $k=2$.

First vertical histogram is constructed by using the following equation:

$$S_{p_i} = \begin{cases} 1 & (V_h(k) = 0 \ \&\& \ V_h(k+1) > 0) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Here $S_{p_i}(k)$ refers to the starting point of all the words. The end point of the words is given as follows:

$$St_{p_i}(k) = \begin{cases} 1 & (V_h(k+1) = 0 \ \&\& \ V_h(k) > 0) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Based on the start and end points the words are segmented.

Now vertical histogram of the words are computed to find the initial segmentation points. After the characters are segmented its features are extracted using diagonal based feature extraction method and feed forward artificial neural network (ANN) is used for classification and recognition.

Our detailed process is given in the fig below(fig 2).

The diagonal based feature extraction is given below:

- i. Each character of size 90*90 pixel is taken.
- ii. Divide each character into 54 zones of 10*10 pixel
- iii. Each zone has 19 diagonals.Now features are extracted by moving along the diagonals.
- iv. Thus 19 subfeatures are extracted from each zone.
- v. These 19 subfeatures are averaged to form a single feature and it is then placed in the corresponding zone.
- vi. In case of empty diagonals the feature corresponds to zero
- vii. Thus 54 features are extracted for each character.Also (9+6) features are also extracted for rows and columns zones.Thus totally (54+16) features are extracted.

Feed forward Artificial Neural Network(ANN) is then trained for classification and recognition. The neural network has two hidden layers of 30-100-100-30 to perform classification. The output layer gives the identified character. The feature vector is denoted by V where $V = (d_1, d_2, \dots, d_n)$ where d denotes feature and n is the number of zones for each character. Here no. of neurons in the outer layer is given by the total no of characters. Input layer neurons by the feature vector 'n' length. However trial and error method is used to determine the hidden layer neurons. It is as shown in the fig. The network training parameters are:

- Input node:35
- Hidden node:100
- Output node:62(52 alphabets,0-9 numerals)
- Training algorithm: Gradient propogation with adaptive learning rate
- Perform function: Sum square error
- Training goal achieved: 0.000001
- Training epochs: 5000
- Training momentum constant:0.9

- Output node:62(52 alphabets,0-9 numerals)
- Training algorithm: Gradient propogation with adaptive learning rate
- Perform function: Sum square error
- Training goal achieved: 0.000001
- Training epochs: 5000
- Training momentum constant:0.95

Our detailed design is given below:

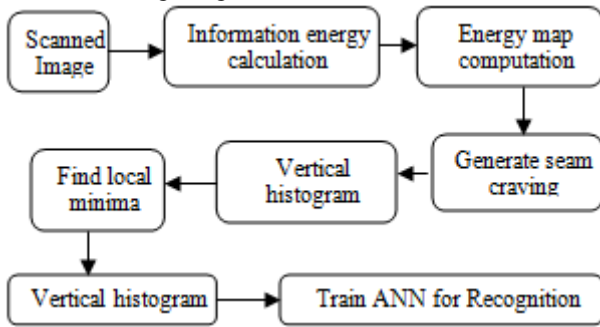


Figure 3: Detailed design of the proposed method

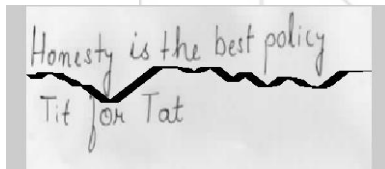
4. Experimental Results

The neural network is trained with 52 English characters capital and small. It is also trained using numbers ranging from 0-9. The input consist of handwritten sentences which in first segmented with respect to lines, words and characters which is as shown in the fig. Finally it is then recognized using the trained neural network. The segmentation accuracy achieved can be tabulated below as:

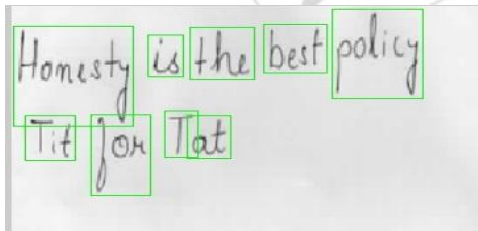
Table 1: Segmentation Accuracy

| Segmentation | Accuracy |
|---------------------------|----------|
| 1) Line segmentation | 95% |
| 2) Word segmentation | 94% |
| 3) Character segmentation | 94% |

The output is as shown below:



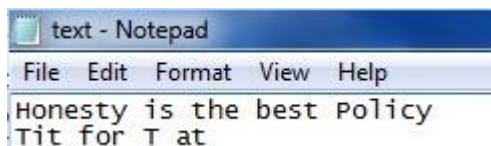
(a) Line segmentation



(b) Word segmentation



(c) Character segmentation



(d) Recognition using neural network

5. Conclusion

Using the above proposed method the line segmentation achieves an accuracy of 95%. The word segmentation achieves an accuracy of almost 94%. Character segmentation has an accuracy of about 94%. The recognition has an accuracy of almost 92%. The proposed method can also be further improved to work on other languages as well as increase the accuracy.

References

- [1] C.A. Boiangiu "Handwritten Documents Text Line Segmentation based On Information Energy, INT J COMPUT COMMUN, ISSN 1841-9836
- [2] L.Likformann-Sulem "A Hough based algorithm for extracting text lines in handwritten documents", International Conference, 95
- [3] Fei Yin,"Handwritten Text line extraction based on minimum spanning tree clustering "International Conference Of wavelet Analysis and Pattern Recognition,2-4 Nov,07.
- [4] Xujun Peng," Markov Random Field Based Text Identification from Annotated Machine Printed Documents" 2009 10th International Conference on Document Analysis and Recognition.
- [5] Neeta Nain Subhash Panwar,"Handwritten Text Recognition System Based on Neural Network".
- [6] Bukhari S.S "Script Independent Handwritten text line segmentation using Active contour", ICDAR, 09.
- [7] Ankush Acharyya," Handwritten word recognition using MLP based classifier" IJCSI, 2013.
- [8] David D.Palmer," A trainable rule-based algorithm for word segmentation", ACL'98 Proceedings of the 35th annual ruling of the association for computational Linguistics and eight conference.
- [9] Laszlo Czuni." Pivot-based search for words spotting in achieve documents", ICEPAF, 15th Aug, 14.
- [10] PP Roy," Muti oriented and Multi sized touching character segmentation using Dynamic Programming", ICDAR, 09.
- [11] Rajib Lochan Das," HMM based Offline Handwritten Writer Independent English Character Recognition using Global and Local Feature Extraction",.
- [12] E.Kavallieratou," Handwritten Character Recognition based on Structural Characteristics".
- [13] M. Baechler and R. Ingold. Multi Resolution Layout Analysis of Medieval Manuscripts Using Dynamic MLP. In International Conference on Document Analysis and Recognition, pages 1185{1189, 2011.
- [14] Kumar, M.; Jindal, M.K.; Sharma, R.K.; , "k-nearest neighbor base offline handwritten Gurmukhi character recognition," Image Information Processing (ICIIP), 2011 International Conference on , vol., no., pp.1-4, 3-5 Nov. 2011