A Review: English Online Handwriting Recognition

Pratishruti Saxena
School of Engineering and Technology
Jaipur National University, Jaipur, Rajasthan, India
pratishrutsaxena8@gmail.com

ABSTRACT

Online handwriting recognition is the trajectory of the pen which is recorded during writing. Handwritten script deal with pen tip where tip is trace by pen positions (up or down). Representation of written output is provided by trajectory as well with compact & complete manner. Handwriting Recognition of is very crucial task to perform with least error. It is use in real time system like PDA, digitizer etc. In this paper we describe the detail study of experiments & their technologies which used in online handwriting recognition.

Keywords: Handwriting; Handwriting recognition; Preprocessing; Segmentation; Feature extraction; Classification; Convolution neural net

I. INTRODUCTION

Recognizing handwritten character written by human sometimes difficult to recognize by other human but still humans are performing far better in comparison to machines in many areas, so making an interface which is capable of recognizing characters written by human yet requires intensive research. Handwriting recognition classified into two types as off-line and on-line handwriting recognition methods. In off-line recognition, the writing is usually captured optically by a scanner and complete writing is available as an image. But, in the on-line handwriting recognition words are generally written on a pressure sensitive surface from which real time information, such as the order of the stroke made by the writer is obtained and preserved. On-line handwriting recognition is shown as superior in comparison to off-line handwriting recognition as temporal information will be available with the on-line handwriting recognition[13].

Handwriting recognition has a wide variety of applications. They range from address recognition for mail sorting to the use of the pen for filling in electronic forms. Depending on the application, the recognition system is either based on scanned image data (off-line recognition) or on pen-trajectory data which is recorded during the writing process (on-line recognition).

(a) Offline Character Recognition
(b) Online Character Recognition

On-line versus off-line recognition. According to the different nature of the data processed, on-line and off-line recognition are each confronted with various specific problems. An advantage of the on-line data is that spatially overlapping characters do not pose a segmentation problem as they are still separable by the time difference of the writing.

On the other hand the temporal information can also complicate the recognition. For example the letter E can be written in one to four strokes and various stroke orders and still produce the same static image [4].

Handwritten character and numeral recognition has a great potential in data and word processing for instance, automated postal address and PIN code reading, data acquisition in bank checks, processing of archived institutional records etc. Several studies have been carried out on recognition of characters in the languages like English, Chinese, Japanese and Arabic. Some studies are reported on the recognition of other languages like Tamil, Telugu, Oriya, Kannada, Panjabi, and Guajarati[12]

A typical handwriting recognition system have following steps:
1. Data acquisition
2. Pre-processing
3. Segmentation
4. Feature extraction
5. Classifier
6. Post-processing

DATAACQUISITION

PRE-PROCESSING

SEGMENTATION

FEATURE EXTRACTION

CLASSIFICATION

POST PROCESSING

Fig. 1

(a) Offline Character Recognition
(b) Online Character Recognition
The issue of classifying preprocessed versus raw data has broad relevance to machine learning, and merits further discussion. Using handcrafted features often yields superior results, and in some cases can render classification essentially trivial. However, there are three points to consider in favor of raw data. Firstly, designing an effective pre-processor requires considerable time and expertise. Secondly, hand-coded features tend to be more task specific. For example, features designed for English handwriting could not be applied to languages with substantially different alphabets, such as Arabic or Chinese. In contrast, a system trained directly on pen movements could be applied to any alphabet. Thirdly, using raw data allows feature extraction to be built into the classifier, and the whole system to be trained together. For example, convolutional neural networks, in which a globally trained hierarchy of network layers is used to extract progressively higher-level features, have proved effective at classifying raw images, such as objects in cluttered scenes or isolated handwritten characters [7].

II. METHODOLOGY FOR ONLINE HANDWRITING RECOGNITION SYSTEM

In this section of the paper discussed about stages of the existing methodologies of the online handwriting recognition system.

A. DATA ACQUISITION

Online handwritten recognition requires a transducer that captures writing as it is written. The most common of these devices is the electronic tablet or digitizer. These devices use a pen that is digital in nature. Data collection is the first phase in online handwriting recognition that collects the sequence of coordinate points of the moving pen. A typical pen includes two actions, namely, PenDown and PenUp [6]. The connected parts of the pen trace between PenDown and PenUp is called a stroke. These pen traces are sampled at constant rate, therefore these pen traces are evenly distributed in time and not in space. The common names of electronic tablet or digitizer are personal digital assistant (PDA), cross pad (or pen tablet) and tablet PC. The appearances of personal digital assistant, cross pad and tablet PC [1].

B. PREPROCESSING

Preprocessing phase in handwriting recognition is applied to remove noise or distortions present in input text due to hardware and software limitations vis-à-vis smooth handwriting. These noise or distortions include irregular size of text, missing points during pen movement collections, jitter present in text, left or right bend in handwriting and uneven distances of points from neighboring position[1]. Preprocessing involves series of operations performed to enhance it suitable for segmentation [14].

Steps which follows by pre-processing:
1. Noise removal.
2. Binirization
3. Slant Angle Correction
4. Resize

The main objective of pre-processing is to denoise and enhance the quality of scanned digit image.

C. SEGMENTATION

Generally, document is processed in hierarchical way. At first level lines are segmented using row histogram. From each row, words are extracted using column histogram and finally characters are extracted from words. Accuracy of final result is highly depends on accuracy of segmentation [14]. A key problem of handwriting recognition, handwriting as well as cursive writing, is character segmentation (separation). While extreme in cursive writing where several characters can be made with one stroke, this problem remains significant with handwriting because the characters can consist of one or more strokes, and it is often not clear which strokes should be grouped together [8].

We can define segmentation in following types:
1. Line segmentation.
2. Word segmentation.
3. Character segmentation.

D. FEATURE EXTRACTION

Feature extraction is the process of collecting distinguishable information of an object or a group of objects so that on the basis of this information we can classify objects with different features [13]. Feature extraction is the heart of any pattern recognition application [4].

Feature extraction techniques are:
1. Principal Component Analysis (PCA),
2. Linear Discriminant Analysis (LDA),
3. Independent Component Analysis (ICA),
4. Chain Code (CC),
5. Scale Invariant Feature Extraction (SIFT),
6. Zoning,
7. Gradient based features [13].

Feature extraction based on three types of feature:

1.) Statistical Feature

These features are derived from statistical distribution points. They provide high speed and low complexity and take care of style variation. Zoning, characteristic loci, crossing and distance are the main statistical features.

2.) Structural Features

Structural features are based on topological and geometrical properties of character, such as, aspect ratio, cross point, loops, branch points, strokes and thire directions, inflection between two points, horizontal curve on top or bottom, etc.
3.) Global transformation and series expansion
A continuous signal generally contains more information than needs to be represented for the purpose of classification. This may be true for discrete approximations of continuous signals as well. One way to represent a signal is by a linear combination of a series of simpler well-defined functions.

E. CLASSIFICATION AND RECOGNITION
The Classifiers compare the input feature with stored pattern and find out the best matching class for input[14]. Classification phase is the decision making part of the recognition system.

Handwriting recognition is the task of transforming a language represented in its spatial form of graphical marks into its symbolic representation. Symbolic representation refers to the digital representation of characters like in the case of 8-bit ASCII character set. Handwriting has been a basic form of communication and still a good way of expressing one’s ideas. In relation to this fact, handwriting recognition systems are useful and are used to realize ideal applications of computers such as pen computing[1].

There are four classification methods:
1. Template Matching
2. Statistence Techniques
3. Structural Techniques
4. Neural Networks

Some techniques or Classifiers which use mostly:-
1. Artificial Neural network
2. Euclidian distance
3. Chess board distance.
4. Taxicab Metric[14].

F. POST-PROCESSING
Post-processing stage is the last stage of the proposed recognition system. It prints the corresponding recognized characters in the structured text form. System results usually contain errors because of character classification and segmentation problems. For the correction of recognition errors, CR systems apply contextual post-processing techniques. The two most common post-processing techniques for error correction are dictionary lookup and statistical approach. The advantage of statistical approach over dictionary-based methods is computational time and memory utilization. The simplest way of incorporating the context information is the utilization of adidiction for correcting the minor mistakes.

III. HANDWRITING RECOGNITION APPROACHES
There are numerous methods for handwriting recognition. In general, most of them are applicable both to online and offline recognition approaches. The main difference between the two is the set of features that is being recognized.

Various approaches which used for online handwriting systems as discussed below:

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Author</th>
<th>Method</th>
<th>Classifier</th>
<th>Approach</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Andreas Kosmala, Joerg Rottland, Gerhard Rigoll[5]</td>
<td>Context Dependent Models</td>
<td>HMM</td>
<td>Trigraph-Based System</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>Han Shu[6]</td>
<td>Feature Extraction Methods</td>
<td>HMM</td>
<td>Baseline Bnn</td>
<td>86.4 (height feature)</td>
</tr>
<tr>
<td>4</td>
<td>Alex Graves, Santiago Fern &amp; Andez, Marcus Lwicke[9][7]</td>
<td>Global Feature Vector Matching</td>
<td>Recurrent Neural Networks</td>
<td>CTC Network</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cursive writing 86.7</td>
</tr>
</tbody>
</table>

THE k-NEAREST-NEIGHBOR (k-NN) rule is a popular non-parametric recognition method, where a posteriori probability is estimated from the frequency of nearest neighbours of the unknown pattern. This approach has been reported to give rather good results, although it is noted that it requires a considerable computational power during the classification process.

HIDDEN MARKOV MODEL (HMM) is another popular way of solving the problem. The model is a doubly stochastic process: it includes an unobservable process (hence “hidden”), but can be observed through another stochastic process that produces the sequence of observations. HMMs have been proven to be one of the most powerful tools for modeling speech and later on a wide variety of other real-world signals. These probabilistic models offer many desirable properties for modeling characters or words. One of the most important properties is the existence of efficient algorithms to train the models automatically without any need of labeling pre-segmented data. HMMs have been extensively applied to handwritten word recognition, including combinations with other approaches, such as stochastic grammars and neural networks.

SUPPORT VECTOR MACHINE (SVM): SVM is based on the statistical learning theory and quadratic programming optimization. An SVM is basically a binary classifier and multiple SVMs can be combined to form a system for multi-class classification. In the recent years, SVM has received increasing attention in the community of machine learning due to its excellent generalization performance [2].

IV. COMPARATIVE ANALYSIS OF ONLINE ENGLISH HANDWRITING RECOGNITION
In above table we show some previous work related to online English handwriting recognition. The work perform by researcher is good in many cases but there is always a scope to perform better. The brief review of character recognition which is done by researcher in past is given:


This system has some unique features that are rarely found in other HMM-based character recognition systems, such as:

1) Option between discrete, continuous, or hybrid modeling of HMM probability density distributions.
2) Large vocabulary recognition based on either printed or cursive word or complete sentence input.
3) Optimized HMM topology with an unusually large number of HMM states.
4) Use of multiple label streams for coding of handwritten information.

Emphasis in this paper is the comparison between continuous and discrete density HMMs, since this is still an open question in handwriting recognition, and is crucial for the future development of the system. However, in order to give a complete description of the basic system architecture, some of the above mentioned issues are also addressed in the next sections. The surprising result of our investigation was the fact that discrete density models led to better results than continuous models, although this is generally not the case for HMM-based speech recognition systems. With the optimized system, a 70% word recognition rate was obtained for a challenging large vocabulary, writer-independent sentence input task.

The maximum recognition rate of the discrete model is 99.35% compared to a maximum recognition rate of the continuous model with one mixture of 97.42%. An improvement of the recognition rate of the continuous model by increasing the number of mixtures is not possible. Andreas Kosmala, Joerg Rottlaad, Gerhard Rigoll[5] in this paper introduce context dependent Hidden Markov Models for cursive, unconstrained handwriting recognition with large vocabularies. Since context dependent models were successfully introduced to speech recognition, it seems obvious, that the use of trigrams could also lead to improved line handwriting recognition systems.

The selective approach represents a smart and simple method for parameter reduction, avoiding the problem of ‘unseen trigrams’ (trigrams with no samples in the training set). The disadvantage is, that this approach does not consider possible similarities between the context of two trigrams with the same center grapheme. The number of training samples of each of two trigrams with similar context and the same center grapheme may fall both below the desired threshold and would be left as monographs. While more generalized context would lead to an increased number of training samples for each trigram and could thus keep trainability for both trigrams. On the other hand, state clustering ensures a more generalized context, while the problem of clustering together seen with unseen trigrams remains unsolved.

In analogy to triphones in speech recognition, trigrams are context dependent sub-word units representing a single written character in its left and right context. The tests were conducted on a writer dependent system with three different environments and two different vocabulary sizes (1000 words and 30000 words). The results we obtained with the trigram-based system compared to the monograph system are very encouraging: A mean relative error reduction of 46% for the 1000 word handwriting recognition system and a mean relative error reduction of 37% for the same system with the 30000 word vocabulary.

We believe that this represents one of the first systematic investigations of the influence of context dependent models and parameter reduction methods for a difficult large vocabulary handwriting recognition task.

Han Shu[6] completed his thesis on online handwriting recognition using HMM. In this thesis, define complete deep study of HMM for online handwriting recognition with every parameter and mainly focus on HMM with various study.

In this thesis, four separate feature experiments were performed. With all four sets of new global-information bearing features, the system obtained a word error rate of 9.1%, a 34% reduction in error from the performance of the baseline system of 13.8%. Among them, the space feature and the substroke features were most effective. Each of them reduced error approximately 15%. These features are height with 86.4% accuracy, space with 88.7% accuracy, hat stroke with 89.3% accuracy & substroke with 99.9% accuracy obtained by the system.

Alex Graves, Santiago Fernándezand Marcus Liwicki[7] describes a system capable of directly transcribing raw online handwriting data. The system consists of an advanced recurrent neural network with an output layer designed for sequence labelling, combined with a probabilistic language model. In experiments on a unconstrained online database, we record excellent results using either raw or preprocessed data, well outperforming a state-of-the-art HMM based system in both cases.

In this paper, author apply a recurrent neural network (RNN) to online handwriting recognition. The RNN architecture is bidirectional Long Short-Term Memory, chosen for its ability to process data with long time dependencies. The RNN uses the recently introduced connectionist temporal classification output layer, which was specifically designed for labelling unsegmented sequence data.

An algorithm is introduced for applying grammatical constraints to the network outputs, there by providing word
level transcriptions. Experiments are carried out on the IAM online database which contains forms of unconstrained English text acquired from a whiteboard. The performance of the RNN system using both raw and preprocessed input data is compared to that of an HMM based system using preprocessed data only. To the best of our knowledge, this is the first timewhole sentences of unconstrained handwriting have been directly transcribed from raw online data.

In this paper combined a BLSTM CTC network with a probabilistic language model & applied this system to an online handwriting database and obtained results that substantially improve on a state-of-the-art HMM based system. We have also shown that the network’s performance with raw sensor inputs is comparable to that with sophisticated preprocessing.

The character error rate for the CTC network with the preprocessed inputs was 11.5 0.05%. with a dictionary and a language model this translates into a mean word error rate of 20.4%, which is a relative error reduction of 42.5% compared to the HMM. Without the language model, the error reduction was 26.8%. With the raw input data CTC achieved a character error rate of 13.9 0.1%, and word error rates that were close to those recorded with the preprocessed data, particularly when the language model was present.


Both systems use a sliding window technique which avoids any segmentation before recognition. The recognizer is a hybrid system containing a neural network and a hidden Markov model. New normalization and feature extraction techniques for the off-line recognition are presented, including a connectionist approach for non-linear core height estimation. Results for uppercase, cursive and mixed case word recognition are reported.

For our final set of recognizers we chose to use the connectionist normalisation with the feature set 2. The on-linetrajectories were converted into off-line bitmaps of 60 pixelsheight. The line width was chosen randomly between 2 and4 pixels. The neural networks used for character probability estimation have TDNN architectures with 4 computational layers with about 30,000 independent parameters.

We implemented a traditional method based on histograms. This algorithmic method uses a linear regression to calculate raw first estimate of the skew angle. For fine tuning projection histograms of the image are calculated for several angles around the horizontal direction. The entropy for each histogram (taken to be a probability distribution) is calculated and the angle corresponding to the minimal entropy chosen as the orientation of the word. An estimate of the core height is obtained by thresholding the histogram at its mean height.

This approach proved to be quite sensitive to the presence of many ascenders or descenders, resulting in errors of the core height estimate. Also short words can be problematic. An optical evaluation of one hundred words, randomly chosen from misrecognized test data, showed about one third of the recognition errors being due to inconsistent skew or core height estimation. By this get 98.3% accuracy for upper case & 86.7% accuracy for cursive handwriting.

P.M. Lallican, C. Viard-Gaudin, S. Knerr[11] propose a new approach for recovering the time order of the off-line writing signal. Starting from an over-segmentation of the off-line handwriting into regular and singular parts, the time ordering of these parts and recognition of the word are performed simultaneously.

This approach, termed “OrdRec”, is based on a graph description of the handwriting signal and a recognition process using Hidden Markov Models (HMM). A complete omni-scriptorisolated word recognition system has been developed. Using a dynamic lexicon and models for upper and lower case characters, our system can process binary and gray value word images of any writing style (script, cursive, or mixed). Using a dual handwriting data base which features both the on-line and the off-line signal for each of the 30 000 words written by about 700 scriptors, we have shown experimentally that such an off-line recognition system, using the recovered time order information, can achieve recognition performances close to those of an on-line recognition system.

In this paper, propose a methodology, termed “OrdRec”, for the reconstruction of the temporal order of the off-line handwriting signal which is based on the simultaneous time ordering and recognition of the signal at the word level. “OrdRec” uses (i) a graph based optimization process which generates candidates for the time ordering and (ii) Hidden Markov Models for the word recognition. Thereby, the decisions as to the time ordering of the writing signal are made globally within the word context instead of locally. It also shows that the “OrdRec” approach often succeeds at recovering the true time order of the handwriting signals, even in cases where purely local analysis does not work.

For instance, the system which has been trained using the true on-line ordering available in the IRONOFF database and the best recovered ordering (Ord. on&off) achieves a 93% recognition rate with N1=N2=6, instead of the 90.2% without “OrdRec”, thereby coming close to the recognition rate of the on-line system which achieves 94.5%. Increasing the number of “OrdRec” candidates beyond 40 or 50 does not bring any improvement. For comparison, our best on-line recognition system which uses a data representation based on a resampled sequence of on-
line points provides a recognition rate of 96% on the same data set. The same system obtains 97.5% when using a “RefRec” approach, i.e., several candidates for the reference lines are used and the best candidate is again found by the recognition process.

In order to evaluate the quality of the recovered stroke order, we have conducted a comparison of the word likelihoods for the true word model computed by both the on-line recognition system and the “OrdRec” system. This comparison shows that for about 80% of the samples the best “OrdRec” candidate obtains a likelihood which is close, equal or larger than the likelihood of the on-line system.

Therefore, we conclude that for about 80% of the samples of the test set the “OrdRec” approach recovers the true (or close to true) time order of the handwriting signal. Note that approximately 15–20% of the test samples are correctly recognized despite an unsatisfactory restoration of the time order.

Muhammad Faisal Zafar, DzulkifliMohamad, Razib M. Othman[9] Instead of doing such lengthy preprocessing, author presents a simple approach to extract the useful character information. This work evaluates the use of the counter-propagation neural network (CPN) and presents feature extraction mechanism in full detail to work with on-line handwriting recognition. The obtained recognition rates were 60% to 94% using the CPN for different sets of character samples. This paper also describes a performance study in which a recognition mechanism with multiple thresholds is evaluated for counter-propagation architecture. Theresults indicate that the application of multiple thresholds has significant effect on recognition mechanism. The method is applicable for off-line character recognition as well. The technique is tested for upper-case English alphabets for a number of different styles from different peoples.

Hecht-Nielsen proposed CPN as an alternate function approximator which can be developed on the available input output data. The first counter-propagation network consisted of a bi-directional mapping between the input and output layers. In essence, while data is presented to the input layer to generate a classification pattern on the output layer, the output layer in turn would accept an additional input vector and generate an output classification. The network got its name from this counter-posing flow of information through its structure.

Most developers use a uniflow variant of this formal representation of counterpropagation. In other words, there is only one feed-forward path from input layer to output layer. The forward-only counterpropagation network architecture, consists of threelabs: an input layer (layer 1) containing n fan out units that multiply the input signals x1, x2, …, xn, and m units that supply the correct output signal values y1, y2, …, ym to the output layer), a middle layer (layer 2 or Kohonen layer) with N processing elements that have output signals z1, z2, …, zN, and a final layer (layer 3) within processing elements having output signals y1’, y2’, …, ym’. The outputs of layer 3 represent approximations to the components y1, y2, …, ym of y = f(x).

From the results it can be concluded that CPN is a good promise in terms of recognition capability which has not been put on trial in the field of handwriting recognition. More over CPN is more economical than convergence of other NN architectures e.g., back-propagation where the training time can take long time. The experiments provided the authors an opportunity to explore this pattern recognition methodology. The exercise provided a theoretical base for further investigations and impetus for development work in this discipline. The obtained results motivate the continuity of the system development considering a preprocessing mechanism including normalization and slant removal.

VI. CONCLUSION

The character recognition methods have developed amazingly in the last decade. A variety of techniques have emerged, influenced by developments in related fields such as image recognition and face recognition. In this paper we provide review of various techniques used in offline handwritten character recognition. These techniques provide better accuracy by use of different classifier. This review provides information about different classifier used in character recognition techniques. This comprehensive discussion will provide insight into the concepts involved, and perhaps provoke further advances in the area. The promise for the future is significantly higher performance for almost every character recognition technology area.

REFERENCES


Aiquanyuan, Gang Bai, Lijing Jiao, Yajie Liu, "Offline handwritten english character recognition based on convolution neural network, 2012, 10th IAPR International workshop on document analysis system."


Rakesh Kumar Mandal, N. R. Manna, "Handwritten english character recognition using rowwise segmentation techniques," ISDMISC 2011 By IJCA.

Huiqin Lin, Wennuan Ou, Tanglin Zhu, "The research of algorithm for handwritten character recognition in correcting assignment system.",

D. K. Patel, T. Som, Manoj Kumar Singh, "Multiresolution technique to handwritten english character recognition using learning rule and Euclidean distance metric.",

Huihung Zhao, Dejian Zhou and Zhaohua Wu, "SMT product character recognition based on BP neural network," 2010 Sixth international conference on natural computation (ICNC 2010).
