Improving the Segmentation of Cursive Handwritten Words using Ligature Detection and Neural Validation

Chun Ki Cheng and Michael Blumenstein

School of Information Technology, Griffith University-Gold Coast Campus

PMB 50, Gold Coast Mail Centre, QLD 9726, Australia

M.Blumenstein@griffith.edu.au

ABSTRACT

This paper describes an enhanced neural network-based segmentation technique for improving the segmentation process in cursive handwriting recognition. The technique has two major steps: First, Prospective Segmentation Points (PSPs) are found by an Enhanced Heuristic Segmenter (EHS). EHS locates possible segmentation points based on the ligature and global characteristics of the handwriting. The second part of the technique fuses confidence values obtained from the left and centre character recognition outputs in addition to segmentation point validation output. A recently proposed feature extraction technique (Modified Direction Feature) was chosen for representing segmentation points and characters to enhance the overall segmentation process. The proposed neural network-based segmentation technique is implemented and tested on a benchmark database providing encouraging results.

Keywords : Cursive Handwriting Segmentation, Neural Networks

1. Introduction

The nature of cursive handwriting [1] is often such that it produces challenging tasks for an automated recognition process. Active research still continues in order to attain a satisfactory solution for recognizing off-line cursive handwriting. The motivating factors include commercial applications and scientific progress in an age-old artificial intelligence problem.

Offline Handwriting Recognition refers to recognising handwriting scripts that have been optically scanned and stored in digital format. One of the major problems in recognizing unconstrained cursive words is the process of segmentation [2, 3]. Segmentation in this context refers to the process of isolating a word image into separate characters. It is an important step in unconstrained handwriting recognition. Poor segmentation contributes heavily to recognition errors. Some systems use the method of over-segmentation to dissect the word at many intervals into character components. Following initial over-segmentation, various techniques may be used to correctly assemble the primitives using contextual processing to recognise entire words. The removal of incorrect segmentation points from over-segmented words is still a difficult problem. A solution to this problem would guarantee a higher success rate for handwritten word recognition. A number of segmentation techniques have been proposed in the literature [4-6].

In this research, further enhancements to the heuristic segmenter of a neural-based segmentation technique are proposed. The current technique analyses the surroundings of every suspicious segmentation point found by the heuristic segmenter. It incorporates a rule-based fusion component to combine three neural network-based confidence values for verification of correct and incorrect segmentation points [7].

To enhance the current segmentation technique, an Enhanced Heuristic Segmenter (EHS) is developed for obtaining better inputs for the neural validation process. In addition, a recently proposed feature extraction technique [8] is used for performing Segmentation Point Validation (SPV) [9] and recognizing left and centre characters (LC and CC respectively) associated with each segmentation point. Finally, a method for generating segmentation paths between cursive characters is described for the purpose of enhancing character extraction.

The remainder of the paper is broken down into 4 sections. Section 2 describes enhancements to the neural confidence-based segmentation technique; Section 3 provides experimental results, followed by a discussion in Section 4. Finally, conclusions are drawn in Section 5.

2. Enhanced Segmentation Technique

In previous work [9], a feature-based heuristic segmenter (FHS) was used to perform over-segmentation of handwritten words. Following this, an SPV technique was used to validate the Segmentation Area (SA) of each PSP. Further work [7] included LC and CC validation in the segmentation process. For SPV, a Density Feature (DF) extraction technique was used to divide the SA into small windows of equal size and analyse the number of black pixels in each window [9]. For LC and CC recognition, the Transition Feature (TF) extraction technique [10] was used for processing the characters. In this section, some enhancements to the previous heuristic segmenter are presented. The enhancements include a new segmenter, EHS and the use of a recently proposed feature extraction technique (MDF) for SA, LC and CC representation.

2.1. Modified Direction Feature (MDF)

Recent work has shown that the Modified Direction Feature (MDF) enhances the character recognition process and outperforms TF [8]. This work demonstrated the superiority of MDF for describing patterns based on their contour or boundary. This prompted an investigation to determine the feasibility of employing MDF for SPV, LC and CC recognition to enhance the overall segmentation process. Figure 1 provides an overview of the entire segmentation technique.

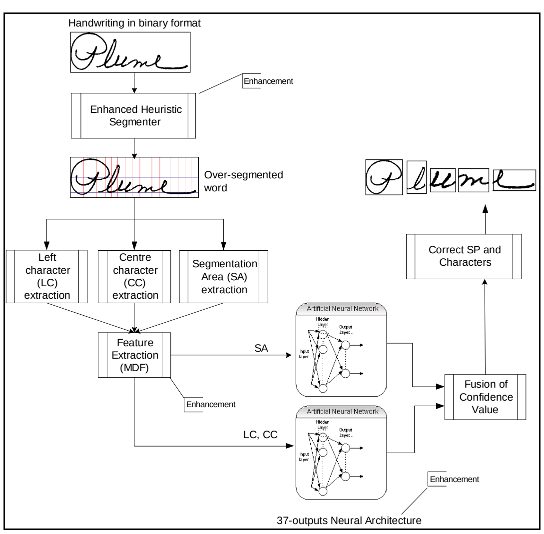

Figure 1 : Overview of the Neural Confidence based Segmentation Technique

Feature vector creation in MDF [8] is based on the location of transitions from background to foreground pixels in the vertical and horizontal directions of a binary image. When a transition is located, two values are stored: the Location of the Transition (LT) and the Direction Transition (DT). An LT is calculated by taking the ratio between the location of the pixel where a transition occurs and the distance across the image in a particular direction [10]. The DT value at a particular location is also stored. The DT is calculated by examining the stroke direction of an object's boundary as defined in Blumenstein *et al.* [11]. Finally, a vector comprising the [LT, DT] values in each of the four possible traversal directions is created.

2.2. Enhanced Heuristic Segmenter (EHS)

The segmentation point assignment of the previous segmenter FHS [7] is based on various handwriting characteristics such as boundary, lower and upper contour minima, vertical density histogram, holes and confidence assignment. The weakness of FHS is its ability of locating segmentation points under overlapped strokes. Although the test words were pre-processed by slant correction, in some words, overlapping characters still existed. As such, EHS was developed with an additional feature, ligature detection. A ligature is a small stroke that is used to connect joined characters.

The algorithm first calculates the position of the upper and lower baselines of the word. The baselines are used to separate the main body of the word with the overlapped strokes. Following this, a modified vertical histogram that represents the distance between the first and last black pixels in each column is generated based on the main body. Finally, the modified vertical histogram is normalized based on the average stroke width. Regions with values less than the average stroke width are identified as ligatures.

2.3. Neural confidence calculation

2.3.1. Segmentation Point Validation (SPV)

As mentioned above, following heuristic segmentation it is necessary to discard "incorrect" segmentation points while preserving the "correct" points in a cursive word. This is achieved by calculating a number of confidences for each Prospective Segmentation Point (PSP) generated by the heuristic segmenter. Three neural networks are used for this step. Firstly, a neural network is trained with features extracted from SAs originally located by the heuristic algorithm. The neural network verifies whether each particular area is or is not characteristic of a segmentation point [9]. If an area is positively identified as a segmentation point, the network outputs a high confidence (> 0.5). Otherwise the network will output a confidence close to 0.1. In this research, the MDF extraction technique was used to describe the segmentation area.

2.3.2. Left and centre character classification

Two additional neural networks trained with handwritten characters (upper case and lower case) are required to confirm the first neural network's output. Each network is presented with areas immediately centred on/adjacent to each segmentation point. Fusion of character and segmentation point confidences using MDF and a single network of 37 outputs is detailed in [12].

2.4. Segmentation Path Detection (SPD)

The first step of extracting characters using SPD is to measure the ascenders and descenders in a word image. Ascenders and descenders are strokes that extend above or below the middle zone or main body of a handwriting sample. Next, the main body of the image is equally divided up into 4 sections, namely sections 1, 2, 3, and 4 (See Figure 2). Based on the x-coordinate of a segmentation point, SPD performs backward traversal. Once a foreground (black) pixel is encountered, the system checks whether the location of the black pixel is below section 1. The line at the bottom of section 1, in Figure 2, is called the "best-fit" line.

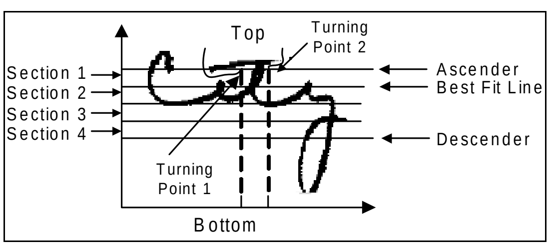

Figure 2 : Word sample sections and segmentation path generation

The "best-fit" line is used as a threshold position, which informs the algorithm whether or not an alternate extraction path should be detected. If the encountered black pixel is below the "best-fit" line, then this pixel, along with all the continuous foreground pixels, are ignored. However, if this black pixel exists on or above the "best-fit" line, this is considered to be the starting point of an overlapping stroke. This pixel is called the "turning point". Commencing from this turning point, a path directed around the overlapping stroke is explored. The algorithm attempts to investigate the right hand side of the turning point. If it is possible to reach the top row of the image, then the extraction path is found. Otherwise, if the traversal to the right hand side is blocked, then it will go back to the turning point, and traverse towards the left hand side. As shown in Figure 3, both left-hand and right-hand segmentation paths of the character "t" are detected. Once an extraction path is detected, all pixel coordinates are stored for the purpose of character extraction.

3. Experimental Results

3.1. Handwriting Database

For the experiments detailed in this section, patterns for SPV and character recognition were obtained from handwritten words contained in the CEDAR benchmark database [13]. Specifically, segmentation points and characters for training and testing were obtained from "/train/cities/BD" and "/test/cities/BD" directories respectively.

3.2. Neural Classifier Configuration

The classifiers chosen for the task of SPV and character classification was a feed-forward Multi-layered Perceptron (MLP) trained with the resilient backpropagation (BP) algorithm. For experimental purposes, the architectures were modified varying the number of inputs, outputs and hidden units. The number of inputs to each network was associated with the size of the feature vector for each image. Various vector dimensions were investigated. The most successful vector configurations were of size 80 for SPV and 120 for character classification (using MDF).

3.3. Segmentation Results

Table 1 shows the segmentation result comparing the current technique [7] and the enhanced technique using the same set of the handwriting database with 317 test words.

Table 1: Segmentation Results

	Segmentation Error Rates		
	Over-segmented [%]	Missed [%]	Bad [%]
Enhanced	8.73	0.1	8.63
Current	7.08	2.33	10.86

4. Discussion of Segmentation Technique

As may be seen from Table 1, the "over-segmentation" and "bad errors" were reasonably low, the "missed" error, which is only 0.1%, is a very promising result. The segmentation technique was successful at discarding bad segmentation points as well as recovering "missed" segmentation points by adding them at large gaps between points in words based on the average character width. One of the reasons, the 37-output neural classifier performed well is that characters were extracted by the aforementioned segmentation path detection technique. "Clean" characters were easier for the 37-output neural classifier to recognize. Although the over-segmentation rate was not very low, it is possible to recover this at a later post-processing stage. It may be noted that it is quite difficult to compare the results obtained here with those in the literature as different data sets were used. Hence, the results in this research can mainly be compared with the previous neural-based segmentation technique [7] that uses the same data set.

5. Conclusions and Future Work

This paper described an enhanced neural confidence-based segmentation technique for cursive handwritten words. The improvements focused on an enhanced feature-based heuristic segmentation algorithm and the use of an MDF extraction technique for SPV, left character and centre character classification. Encouraging results were obtained that can increase the overall segmentation performance. The use of the SPD technique for character extraction proved to be successful, and has the potential for enhancing the process of left and centre character extraction.

In future work, the above mentioned enhanced segmentation technique will be tested on a larger dataset and MDF will be optimized for improving SPV. Finally, SPD will be improved to include direction information.

References

[1] R. G. Casey, "Moment Normalization of Handprinted Characters", *IBM Journal of Research Development*, pp. 548-557, 1970.

[2] R. G. Casey and E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, pp. 690-706, 1996.

[3] Y. Lu and M. Shridhar, "Character Segmentation in Handwritten Words – An Overview", *Pattern Recognition*, Vol. 29, pp. 77-96, 1996.

[4] B. Yanikoglu and P. A. Sandon, "Segmentation of Off-Line Cursive Handwriting using Linear Programming", *Pattern Recognition*, Vol. 31, pp. 1825-1833, 1998.

[5] G. Nicchiotti and C. Scagliola, "A Simple and Effective Cursive Word Segmentation Method", *Proceedings of the 7th International Workshop on the Frontiers of Handwriting Recognition (IWFHR-7)*, pp. 499-504, 2000.

[6] X. Xiao and G. Leedham, "Knowledge-based Cursive Script Segmentation", *Pattern Recognition Letters*, Vol. 21, pp. 945-954, 2000.

[7] M. Blumenstein, and B. K. Verma, "Analysis of Segmentation Performance on the CEDAR Benchmark Database", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR '01)*, pp. 1142-46, 2001.

[8] M. Blumenstein, X. Y. Liu and B. Verma, "A Modified Direction Feature for Cursive Character Recognition", *Proceedings of the International Joint Conference on Neural Networks (IJCNN '04)*, Budapest, Hungary, (CD-ROM Proceedings), 2004.

[9] M. Blumenstein and B. Verma, "A New Segmentation Algorithm for Handwritten Word Recognition", *Proceedings of the International Joint Conference on Neural Networks (IJCNN '99)*, Washington D.C., pp. 878-882, 1999.

[10] P. D. Gader, M. Mohamed and J-H. Chiang, "Handwritten Word Recognition with Character and Inter-Character Neural Networks", *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, Vol. 27, pp. 158-164, 1997.

[11] M. Blumenstein, B. Verma and H. Basli, "A novel feature extraction technique for the recognition of segmented handwritten characters", *Proceedings of the Seventh International Conference on Document Analysis and Recognition, (ICDAR '03)*, Edinburgh, Scotland, pp. 137-141, 2003.

[12] C. K. Cheng, X. Y. Liu, M. Blumenstein, and V. Muthukkumarasamy, "Enhancing Neural Confidence-Based Segmentation For Cursive Handwriting Recognition", *5th International Conference on Simulated Evolution And Learning (SEAL '04)*, Busan, Korea, SWA-8, CD-ROM Proceedings, 2004.

[13] J. J. Hull, "A Database for Handwritten Text Recognition", *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol. 16, pp. 550-554, 1994.