

Implicit Vs Explicit based Script Segmentation and Recognition: A Performance Comparison on Benchmark Database

Amjad Rehman, Dzulkipli Mohamad and Ghazali Sulong

Department of Computer Graphics and Multimedia
Faculty of Computer Science and Information System
University Technology Malaysia

Abstract

This paper compares implicit and explicit based segmentation techniques for off-line cursive handwriting recognition. Firstly, limitations in conventional implicit based approaches are handled by imposing sequence of heuristic rules to locate prospective segmentation points. However, fine segmentation delayed until character recognition. Additionally, for character hypothesis generation, forward and backward strategies were adopted, whereas character verification was lexicon based. Secondly, in explicit based approach, two ANN were employed. One for segmentation point validation forwarded by heuristic segmenter based on character shape analysis and other for recognition. Finally, for character classification, hybrid statistical features were extracted from segmented characters. Techniques are tested and compared on IAM benchmark database with consistent platform.

Keywords: ANN validation, character recognition, explicit segmentation, implicit segmentation, Hybrid features, Features extraction.

1 Introduction

Despite achievements in off-line cursive handwriting segmentation and recognition of characters has been reported in the literature [1-7]. It is still an open issue to have an objective evaluation of the performance comparison in realistic circumstances with respect to accuracy and computational complexity [8, 9].

Researchers have acknowledged the importance of segmentation in the handwriting recognition process and consider it the most crucial step [10]. In fact correct recognition often depends on a correct segmentation [11]. This is precisely why more innovative and accurate methods need to be explored and compare. Accurate segmentation techniques are still in demand to be used as important components in the overall handwriting recognition process [12]. Broadly, analytical segmentation approaches are divided into two categories: implicit segmentation and explicit based segmentation [18]. Implicit based segmentation or internal segmentation, in which the system searches the image for components that match classes in its alphabet also termed as recognition based segmentation [19,23,24]. Whereas, explicit based segmentation approach or external segmentation, in which segments are identified based on "character-like" properties [1-4,25,26]. This process of cutting up the image into meaningful components is given a special name, "dissection".

This research presents and compares two analytical techniques for segmentation and recognition of cursive handwritten words. Experiments were conducted using IAM benchmark database [13]. Initially, in implicit based approach, heuristic rule-based segmentation algorithm is used to find prospective segmentation points in cursive handwritten words [14]. Finally, recognition based segmentation verifies the segmentation points. To recognize segments as valid characters, hybrid features are extracted from each segment [15].

Secondly, in explicit based approach, shape based analysis is performed to locate ligatures for character segmentation. Two neural networks are further used. First, ANN is used to validate segmentation points following heuristic segmentation [16]. Second ANN is used to classify segmented characters using hybrid statistical features [15].

The remainder of the paper is organized into sections. Section 2 briefly describes proposed techniques along with algorithms, section 3 presents implementation and classification results of each technique, analysis and comparison of results follows in section 4 and a conclusion is drawn in section 5.

2 Proposed techniques

Preprocessing, feature extraction scheme and the classifier are same for the both analytical approaches. However, first segmentation approach is not integrated with any intelligent technique at segmentation stage. The comparison is made on the basis of accuracy and computational complexity.

2.1 Preprocessing

Following digitization, handwritten word images are threshold to filter noise [17]. The both techniques follow vertical cut strategy therefore to avoid shadow of one character to the neighboring, slant correction is applied. Additionally, to

accommodate large variability of the handwriting stroke width that found in IAM database, thinning operation is performed.

2.2 Segmentation algorithms

In cursive handwriting ligatures are ideal segmentation points. Therefore, both segmentation algorithms are applied to detect ligatures between letters. However strategy of both algorithms is different that is one is implicit based and other is explicit based.

2.2.1 Implicit based segmentation approach

Generally, in this approach segmentation and recognition of characters are achieved at the same time. Segmentation is only a byproduct of the recognition process and therefore is a simple process yet provides all tentative segments and let the recognizer to decide best segmentation hypothesis. However, there is a tradeoff in selecting total number of segments for a word. Less number of segments is the base of efficient computation but on downfall side characters written wide can not be covered in the hypothesis. Whereas, large number of segments, generates more slices that again have two main shortcomings. First, it is computationally expensive since it increase number of character hypothesis and all hypotheses generated must be evaluated. This is a very important issue that has been ignored very often in the literature [28]. Second and more severe number of junks increase significantly that is additional burden on character recognizer in modeling junks [27]. Additionally, parts of characters are recognized as valid character, which is commonly known as class overlapping problem [19]. In this regard, Britto et al. [20, 21] have observed some loss in recognition performance caused by combining segmentation with recognition.

To overcome problems mentioned above, our implicit based segmentation approach is a two stage process. First, stage is heuristic rule based segmentation that provides all tentative segmentation points to be finalized by the recognizer in second stage. The goal of heuristic segmenter is not only to over-segment words but also an attempt to remove incorrect segmentation points which are tradeoff of over-segmentation. Segments were joined or separated and possible valid unions were found based on pre-defined sequence of heuristic rules. Consequently, it reduced number of character hypothesis and therefore lessened burden of classifier in second stage. Additionally, it increases character recognition efficiency and might also increase chances of correct word recognition. The first stage of the implicit based segmentation is as under.

Over-segmentation: Following preprocessing, image is over segmented heuristically at horizontal distance 'x' where $x = h/18$, h is the height of the word image and 18 is evaluated empirically (Fig. 1b).

Loop determination: For each vertical line, count its crossing the foreground pixels. If count is more than 1, extract vertical segment line (Fig. 1c).

Character boundaries: If successive vertical lines are at distance 'x', accept first and last vertical lines, excluding in between (Fig. 1d).

Junk detection and removal: If consecutive vertical segment lines are at distance less than or equal to 'x', set their mean as letter boundaries. (Fig. 1e)

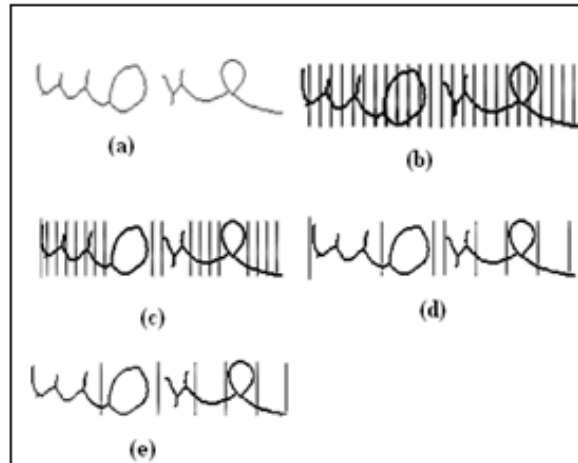


Fig 1: A sequence of processing results of stage 1 (implicit based approach)

The heuristic rule-based approach performed well in most of the cases; however, few characters were over-segmented as shown in figure 2. Nevertheless, over-segmented characters were not segmented into more than two pieces. Hence our character hypothesis consists of maximum two fragments, that not only reduced burden of classifier and but was also base of efficient computation. It is mention worthy, that miss-segmentation occurred in case of touched characters which is out of scope of this paper.

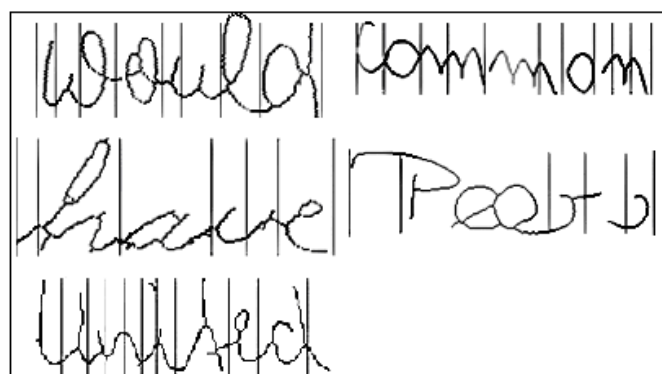


Fig 2: Failure results of heuristic rule based approach

2.2.2 Explicit based segmentation approach

In this approach, following preprocessing (figure 2b, 2c), number of foreground pixels are counted vertically in each column to detect holes and concavities. Only those columns are set as candidate segmentation columns (CSC) for which count was 0 or 1. Consequently, we faced over-segmentation problem that emerged from two sources. Firstly, ligatures/union of characters put forward many CSC (see figure 2d). Secondly, characters without hole/concavity vertically such as m, n, u, v, w etc. Therefore problem of over-segmentation was handled in two phases. In first phase, to avoid problem of over-segmentation in ligatures, an average of those CSC was taken which are at distance less than threshold to merge them into one segment column. Threshold is the minimum horizontal distance between successive CSC that can not accommodate a character and was set to value 4 experimentally. In second phase one more heuristic was added to locate and handle over-segmented characters nonetheless over-segmentation was minimum.

Each segmentation point (SP) was given more confidence by checking its crossing of foreground pixels and nearest neighbors. If histogram for current SP was zero then highest confidence was allocated without any further processing, otherwise presence of hole/concavities in its nearest neighbors are investigated. The current SP that did not have hole/concavity/highest confidence segment line in its nearest neighbor was probably over-segmented character. To trounce this problem, density features of over-segmented character were extracted and are fed to ANN to validate correctness of each segmentation point. The correct ones are retained whereas incorrect are extracted. Detailed discussion of training set creation and testing is available in [16]. Stepwise results are demonstrated in fig. 3(a, b, c, d, e).

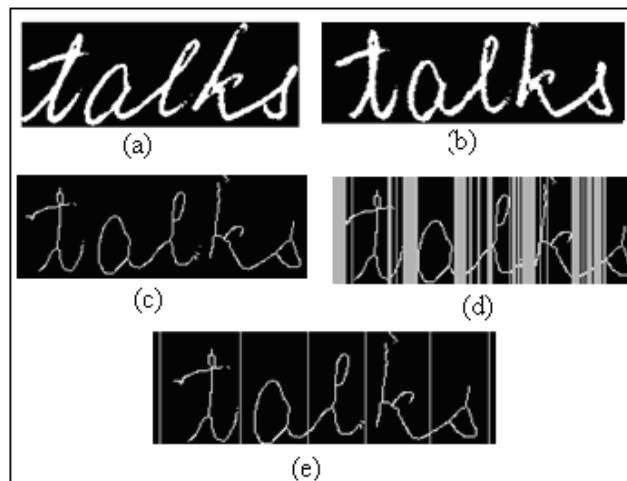


Fig 3: (a) The original IAM word image sample which is cursive and slanted (b) resulting image after slant correction (c) resulting image after thinning (d) All CSC over-segmenting the image (e) Fine segmented word after thresholding

Proposed segmentation algorithm is explained as under

Let say an image denoted by P where $P \in \{0,1\}$

$$p_{i,j} \in P \begin{cases} i = 1, 2, \dots, h \\ j = 1, 2, \dots, w \end{cases}$$

Where h, w is height and width of P respectively.

i. Calculate sum of foreground pixels for each column.

$$a_j = \sum_{i=1}^h p_{i,j}, j = 1, 2, \dots, w .$$

ii. Detection of candidate segment columns (CSC)

$$CSC = \{csc \in a_j \mid a_j \leq 1, j = 1, 2, \dots, w\}$$

iii. Thresholding, first initialize $m, n, sum = 0$

for $k = 1..length(CSC) - 1$

if $(csc_{k+1} - csc_k \leq threshold)$ then

$$sum = sum + csc_k$$

$$n = n + 1$$

else

if $n > 0$ then

$$SP_m = round(sum / n)$$

else

$$SP_m = csc_k$$

$$m = m + 1$$

$$sum = 0, n = 0$$

endfor

iv. Segmentation Point (SP) validation except first and last

$$validate_SP_1 = SP_1$$

$$k = 2, m = 2$$

while $(m < length(SP))$

$save_segment_point = true$

$jump = false$

[checking current segment point]

if $histogram(SP_m) > 0$ then

[Checking hole/concavities in nearest neighbor of current SP]

if $transition_feautres(SP_{m-1}, SP_m) = 1$ OR $transition_feautres(SP_m, SP_{m+1}) = 1$

extract $transition_feautres(SP_{m-1}, SP_{m+1})$ then feed to ANN validation.

if ANN output is “incorrect” then

$save_segment_point = false$

else

$$m = m + 1$$

$$jump = true$$

end

```

    end
  end
  [save valid segment point]
  if save_segment_point then
     $validate\_SP_k = SP_m$ 
     $k = k + 1$ 
    if jump then
       $m = m + 1$ 
    end
  end
end while
[save last segment point]

```

$validate_SP_k = SP_m$

2.3 Hybrid feature extraction from segmented characters and ANN training

For classification of segmented characters, two types of statistical features: transition features & profile projection features are combined in hybrid fashion and extracted from each segment. To allow varying sizes of segments/characters no resizing is applied. The transition features computes number of transition from background to foreground and vice versa in each row and column. While profile projection features divides the character image into two equal halves and calculates distances of upper and lower projections from the middle line. NIST SD 19 is used to train ANN based on hybrid features extracted from both lower and upper case characters. Detailed discussion for feature extraction and ANN training/testing can be found in [15]. Following number of experiments optimal structure of ANN was finalized, that is presented in table 1. Furthermore, the whole platform was kept constant for character recognition performance, segmented from each segmentation approach.

Table 1: ANN optimal structure for character classification

| Input units | Hidden units | Output units | Momentum/ learning rate | Training data |
|-------------|--------------|--------------|-------------------------|---------------|
| 174 | 89 | 53 | 0.05 | 30,000 |

3 Classification results

Following ANN training for segmented characters recognition in previous section it was combined with both segmentation approaches separately to classify the segmented characters. For meaningful experiments and comparisons 350 test words were taken from IAM benchmark database [13] and a lexicon was developed.

3.1 Implementation and Experimentation of implicit based segmentation approach

Following heuristic rule based segmentation as detailed in section 2.2.1, fine segmentation was finalized based on recognition. To handle ornaments of the characters, a special junk class is introduced and ANN is trained with true characters and junks [15]. Finally, in second stage, hybrid statistical feature are extracted from each segment, normalized and are fed to ANN for their classification since fine segmentation was based on character recognition. As mentioned earlier, over-segmented characters were fragmented into maximum two halves, therefore, our character recognition system does not join more than two fragments. The joining of character's fragments is lexicon based recognition for which both forward - backward strategies are adopted.

3.2 Implementation and Experimentation of explicit based segmentation approach

Following intelligent segmentation of cursive handwritten words described in section 2.2.2. For each cursive word, a character matrix was extracted between successive dissections, normalized and are fed to a trained ANN for character classification [15].

3.3 Word recognition

Finally, word likelihood is lexicon based for each analytical approach. In order to reduce lexicon and to make recognition process effective, a simple guess strategy is adopted to recognize word from lexicon based on sequential character matching. Accordingly, based on recognition and matching of first few recognized segmented characters, lexicon was reduced continually until one target word was located. This not only reduced matching process but also increased accuracy. Additionally, all characters recognition and their matching were not desired.

Table 2: Word recognition results based on implicit Vs explicit segmentation approaches

| Segmentation technique | No. of IAM words for testing | Classification rate (%) |
|-----------------------------|------------------------------|-------------------------|
| Implicit based segmentation | 350 | 79.23 |
| Explicit Based segmentation | 350 | 80.91 |

4 Analysis and Comparison of results

It is very interesting that segmentation based approaches produce almost similar results despite of mistakes of each approach were different.

Initially, implicit segmentation based recognition defeated class overlapping problem and reduction of character hypothesis/ junks that not only reduced burden of classifier but also increased character recognition accuracy in the state of the art. On the other hand, segments of few characters such as m, n, w etc had less discriminative nature and were classified as true character as compared to other segments that do not have such competition. Therefore at this stage we took lexicon help, to join the pieces of such characters.

Explicit based segmentation integrated with neural validation of segmentation points remained very successful as it rejected incorrect segmentation points and facilitated further character recognition process. Yet some shortcomings were observed during experimentation. On the downfall, touching and broken characters could not be segmented and also created problems for the succeeding characters recognition, however touching and broken character's segmentation was out of scope for this research. Finally, as character boundaries are detected and segmented vertically, horizontally overlapped characters also created problems that were solved by applying the heuristic segmentation in the core-zone and characters were detected by tracing connected components [22].

4.1 Comparison of segmentation techniques

It is hard to compare both segmentation techniques as first is implicit segmentation based approach whereas, second is explicit segmentation based approach. Nevertheless, these two techniques can be compared for recognition of

cursive handwritten words. However, it must be stressed once again that both segmentation approaches use same preprocessing and post-processing treatment.

It can be seen from table 2 that character recognition results of implicit based segmentation approach is slightly less than explicit based segmentation approach. Whereas explicit segmentation approach employed two ANN, one for fine segmentation and the other for segmented character recognition. That not only increased complexity and processing time but also was laborious due to manual separation of correct and incorrect segmentation points during ANN training phase. On the other hand, implicit segmentation approach made use of one ANN for fine segmentation and recognition. Hence, it is mention worthy that against less processing time and computational efforts, character recognition results are comparable to those obtained from explicit segmentation based approach. For a large real life handwriting recognition system, it may be considered a big advantage. Finally, the experiments were taken on small data set and therefore for larger data set explicit approach may be more successful from accuracy point of view.

5 Conclusion and Future Plan

Two analytical segmentation based approaches for cursive handwriting recognition are presented and compared on benchmark database from accuracy and computational complexity points of view. Both techniques produced very encouraging results with interesting findings. Explicit segmentation based approach was computationally complex but yield slightly better results than less complex implicit based segmentation approach.

This research is still on going and more intelligent features are required to be explored to enhance accuracy and overcome problems in segmentation and recognition due to touched and broken characters in cursive handwriting.

6 Open Problem

In off-line cursive handwriting recognition domain character segmentation, feature extraction of segmented characters and their classification is still an open problem. Considerable work has been carried out by research community but still the results are far from maturity. Additionally, it is required to compare segmentation and recognition results of various approaches on benchmark database.

ACKNOWLEDGEMENTS

This research work is fully supported by Ministry of Science and Technology Innovation (MOSTI) Malaysia. Authors would like to thank RMC UTM staff for their support.

References

- [1] Verma, B. "A Contour Character Extraction Approach in Conjunction with a Neural Confidence Fusion Technique for the Segmentation of Handwriting Recognition. Proceeding of the 9th International Conference on Neural Information Processing, Vol. 5, (2002), pp. 2459-2463.
- [2] Verma, B. "A Contour Code Feature Based Segmentation for Handwriting Recognition", In Proceedings of 7th International Conference on Document Analysis and Recognition (ICDAR'03), (2003), pp. 1203-1207.
- [3] Cheng, C. K., Liu, X., Y., Blumenstein, M., and Muthukkumarasamy, V. "Enhancing Neural Confidence-Based Segmentation for Cursive Handwriting Recognition", Proceedings of 5th International Conference on Simulated Evolution and Learning, Busan, Korea, SWA-8, (2004).
- [4] Hamamura, T., Akagi, T., Irie, B. "An Analytic Word Recognition Algorithm Using a Posterior Probability". In Proceedings of International Conference on Document Analysis and Recognition. Vol. 02, (2007), pp. 669-673.
- [5] Blumenstein, M., Liu, X., Verma. B. "An investigation of the modified direction feature for cursive character recognition", Pattern Recognition, Vol. 40, (2007), pp. 376-388.
- [6] Araki, N.; Okuzaki, M.; Konishi, Y.; Ishigaki, H. "A Statistical Approach for Handwritten Character Recognition Using Bayesian". Proceeding of 3rd International Conference on Filter Innovative Computing Information and Control, (2008), pp. 194-198.
- [7] Blumenstein, M., Verma, B., Basli, H. "A novel feature extraction technique for the recognition of segmented handwritten characters". In Proceedings of 7th International Conference on Document Analysis and Recognition, (2003), pp. 137-141.
- [8] Gatos, B., Antonacopoulos, A., Stamatopoulos, N. "Handwriting Segmentation Context". Proceedings of the International Conference on Document Analysis and Recognition (ICDAR'2007), pp.1284-1288.
- [9] Borji, A., Hamidi, M. "Optical character recognition motivated by primate visual system". Neural Network World; Vol. 17, No. 5, (2007), pp. 431.
- [10] Verma and Blumenstein (2008). Fusion of Segmentation Strategies for Off-Line Cursive Handwriting Recognition, chapter 1, page 2
- [11] Bortolozzi, F., Souza, A., Britto Jr., Luiz S. Oliveira and Morita, M. Recent Advances in Handwriting Recognition. Document Analysis, Editors: Umapada Pal, Swapan K. Parui, Bidyut B. Chaudhuri, (2005), pp. 1-30.

- [12] Lorigo, L.M. and Govindaraju, V. "Offline Arabic handwriting recognition: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 5, (2006), pp. 712-724.
- [13] Marti, U., and Bunke, H. "The IAM database: An English Sentence Database for Off-line Handwriting Recognition". *International Journal of Document Analysis and Recognition*, Vol. 15, (2002), pp. 65-90.
- [14] Rehman, A., Kurniawan, F and Dzulkifli, M. "Off-line Cursive Handwriting Segmentation, A Heuristic Rule-based Approach". *Journal of Institute of Mathematics and Computer Science (Computer Science Series)*, Vol. 19, No. 2, (2008), pp.135-140.
- [15] Rehman, A., Kurniawan, F., and Dzulkifli, M. Off-line cursive character recognition based on hybrid statistical features. *Proceedings of International Graduate Conference on Engineering and Science, 2nd RPCES conference*, (2008).
- [16] Rehman, A. and Dzulkifli, M. A Simple Segmentation Approach for Unconstrained Cursive Handwritten Words in Conjunction with the Neural Network. *International Journal of Image Processing*. Vol 2, No. 3, (2008), pp. 29-35.
- [17] Otsu, N. "A Threshold Selection Method from Gray level Histograms", *IEEE Trans. on Systems, Man and Cybernetics* Vol. 9, No. 1, (1979), pp. 62-66.
- [18] Yanikoglu, B and Sandon, P.A "Segmentation of off-line cursive handwriting using linear programming", *Pattern Recognition*, Vol. 31, (1998), pp. 1825-1833.
- [19] Cavalin, P., R., Britto, A., S., Bortolozzi, F., Sabourin, R., and Oliveira, L., (2006). An Implicit Segmentation based Method for Recognition of Handwritten Strings of Characters. *Proceedings of ACM symposium on applied computing*, 836-840.
- [20] Britto, A. S., Sabourin, R., Bortolozzi, F., Suen, C. Y.: An enhanced HMM topology in an LBA framework for the recognition of handwritten numeral strings, *Proceedings of the International Conference on Advances in Pattern Recognition*, Vol 1, (2001), pp. 105-114
- [21] Britto, A. S., Sabourin, R., Bortolozzi, F., Suen, C. Y.: A two-stage HMM-based systems for recognizing handwritten numeral strings. *Proceedings of the International Conference on Document Analysis and Recognition*, (2001), pp. 396-400.
- [22] Dzulkifli, M., Rehman, A., and Kurniawan, F. A New Approach for Segmenting Difficult Cursive Handwritten Words from Benchmark Database. *Proceedings of 4th International Conference on Information & Communication Technology and Systems (ICTS)*, Vol.1,(2008), pp.17-21.

- [23] Gillies, M.: "Cursive word recognition using hidden Markov models". In Proc. Fifth U.S. Postal Service Advanced Technology Conference, pp. (1992), pp. 557-562.
- [24] Cho, W., Lee, S. W., Kim, J. H. "Modelling and recognition of cursive words with hidden Markov models". Pattern Recognition, Vol. 28, No. 12, (1995), pp. 1941-1953.
- [25] El-Yacoubi, A., Gilloux, M., Sabourin, R., Suen, C. Y. "An HMM-based approach for off-line unconstrained handwritten word modelling and recognition". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 8, (1999), pp. 752-760.
- [26] Arica, N., Yarman-Vural, F. T. "Optical character recognition for cursive handwriting". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 24, No. 6 , (2002), pp.801-813.
- [27] Sayre, K., M. "Machine Recognition of Handwritten Words: A Project Report". Pattern Recognition, Vol. 5, (1973), pp.213-228.
- [28] Oliveira, L.S., Britto, A.S. and Sabourin, R. "A Synthetic Database to Assess Segmentation Algorithms". Proceedings of Eight International Conference on Document Analysis and Recognition (ICDAR'2005), Vol. 1, pp.207-211.