

A feedback-based Approach for Segmenting Handwritten Legal Amounts on Bank Cheques

Jun Zhou, Ching Y. Suen and Ke Liu
Centre for Pattern Recognition and Machine Intelligence
Concordia University
Montreal, Quebec H3G 1M8, Canada
{junzhou, suen, keliu}@cenparmi.concordia.ca

Abstract

The proposed feedback-based approach is implemented in two steps. In the first step, segmentation is done according to the structural features between the connected components in the legal amounts. In the second step, a feedback process is introduced to re-segment the parts that could not be identified in the first step. Then a multiple neural network classifier is used to verify the re-segmentation result. The confidence value produced by the classifier is used to determine the best segmentation points. This approach is tested on a new CENPARMI database and the result indicates that the correct segmentation rate increased by 13.4% from the previous approach.

1. Introduction

Automatic processing of bank cheques has been studied extensively in the past decade. Topics include the recognition of legal amount, courtesy amount and date [1, 3, 5, 11, 12]. Among them, automatic processing of legal amount is a challenging task because of the high variability of characters, words and writing styles. Before the recognition of legal amount, reliable sentence-to-word segmentation is a difficult task. Currently, most approaches focus on the identification of physical gaps between words. Besides that, some researchers have proposed another method which incorporates the writing style of author in terms of spacing [4].

A typical sentence-to-word segmentation approach may encompass several steps from finding the connected components, computing and sorting the distances between components, to selecting a threshold to determine the inter-word and inter-character gaps.

During this procedure, computation of the distances is quite important. The distances can be calculated according to bounding box, Euclidean distance, minimum run-length, convex hull and their combinations [6, 8]. Most approaches assume that the distances between words are bigger than those between characters. Sometimes, other information, such as transition from a string of lower case characters to upper case characters, may be utilized to determine the segmentation points. However, bank cheque users may not obey the rules when writing cheques. The distances between words and characters may not be far apart. Normally, the distances between words in a line vary considerably because the users often write big characters at the beginning of a line and later on find that the space is not large enough at the end. As usual, a transition between upper and lower characters may not exist. In such cases, the selection of a threshold to distinguish inter-word and inter-character gaps may become a troublesome task. Even though a good threshold can be reached after the training of the segmentation algorithm, the rates of under-segmentation and over-segmentation may still be very high. Here, under-segmentation and over-segmentation mean the presence of fewer or more words than are actually present. And the adjustment of the threshold will cause a drop of one rate while greatly increasing the other.

In this paper, a feedback-based approach is proposed to improve the performance of the segmentation. In the CENPARMI cheque processing system, a KNN classifier based on global features [3] and an HMM-MLP hybrid model classifier [10] have been implemented to recognize the legal amounts after preprocessing and segmentation. The performance of both recognizers depends heavily on the result from the sentence-to-word segmentation stage. In our approach, segmentation is divided into two steps. In the first step, a structural

feature based segmentation module is called to calculate the Euclidean distance between the connected components in a sample and to make an initial segmentation. In the second step, feedback information on the length of the segments is collected from the result of the first step. By comparing the information with the global information of the legal amount sample, re-segmentation is called to split some selected segments. Several segmentation points of the selected segments are produced. A multiple neural network classifier is then introduced to generate confidence values that can be used by the feedback system to select the best segmentation points. The structure of the approach is shown in Figure 1.

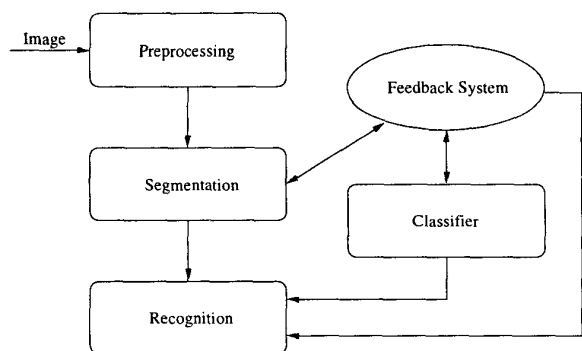


Figure 1. A block diagram for feedback-based segmentation approach

2. Feedback from first segmentation step

The first segmentation step (SEG1) adopts the structural features to calculate the Euclidean distance between connected components and produces word segments. The performance of SEG1 is then analyzed by a feedback system. It finds those segments that are not completely segmented, and re-segment them with a different threshold based on the feedback from the first calling (SEG2).

2.1. First segmentation step

A sample of legal amount is shown in Figure 2(a). SEG1 first represents a binary image as a list of contours of the connected components, then the components that are recognized as line or dash and punctuations are removed from further consideration. In this step, several classifiers have been designed and trained for the recognition and removal [2]. For example, a

Bayesian classifier based on simple features (i.e. aspect ratio and average vertical density) is designed and trained to recognize and remove the lines that people have the tendency to write at the beginning or end of the legal amount. Then the system groups some connected components together, for example, those fully overlapped components such as the small dot over “i” or the vertical stroke of the letter “T”. This process is also useful for those words composed of broken strokes due to some weakness of the binarization process. The components that are very “close” to each other are also grouped together. After the grouping step, the system computes the distance from one group to the next. The distances are sorted by size and the largest inter-word gap is found. All those gaps that are larger than a given threshold ($t1$) of the largest candidates are classified as inter-word gaps. Another threshold ($t2$) is also defined in terms of pixels as the smallest value of inter-word gaps. Figure 2(b) shows the result of first segmentation step of the legal amount in Figure 2(a).

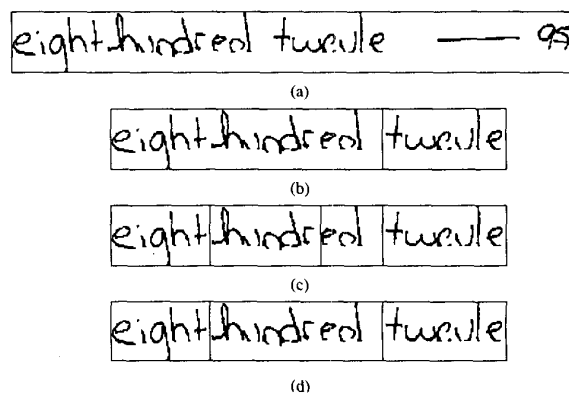


Figure 2. A legal amount sample and segmentation. (a) is the original image; (b) is the result after first segmentation step; (c) is the result after re-segmentation; (d) is the final segmentation result.

2.2. Feedback and re-segmentation

As we can see from Figure 2(b), when the distances between words vary greatly, it is very difficult to select a sound threshold to distinguish inter-word gaps and inter-character gaps. Adjustment of a well-trained threshold may incur an increase in under-segmentation or over-segmentation. However, it is possible that a threshold be selected to control either under-segmentation or over-segmentation to a very low

level. Then we can find out the under-segmented or over-segmented parts with the help of global information and classifier. In our approach, we trained the t_1 and t_2 to certain values so that the amount of over-segmentation is minimized. Then we try to solve the problem of under-segmentation according to the feedback of the first segmentation step.

By analyzing the 32 words in the lexicon of our cheque processing system, we find that the only longest word and next five longest words are composed of 9 characters and 8 characters respectively. The average length of all the words in lexicon is 5.66 characters. Thus the ratios of the length of longest and second longest words to the average length of all words in lexicon are 1.59 and 1.41 respectively. If we define $t_3 = 1.5$, which is the middle value of 1.59 and 1.41, then t_3 can be considered as a theoretical value to identify under-segmented segments. Once the stated property of a segment exceeds this value, under-segmentation may occur. A test of SEG1 on CENPARMI database which contains 408 legal amounts shows that in 85.4% cases, the longest segments are not successfully separated. Thus, among all the segments identified by the first segmentation step, the longest segment is most probably under-segmented part. In our system, the segments are sorted according to their lengths after SEG1. Feedback information on the length of each segment is collected to calculate the average length of each segment. The algorithm to select a probably under-segmented segment is:

$$t_4 = \frac{\max\{l_i | i = 1, \dots, n\}}{\frac{1}{n} \sum_{i=1}^n (l_i)} \quad (1)$$

where l_i is the length of the i th segment, n is the number of segments. If $t_4 > t_3$, the longest segment has to be re-segmented. If there is only one segment and it is longer than a certain threshold, it will be automatically sent to SEG2. There are some special cases that we only get two segments after the first step, and the lengths of the segments are similar. Then there is great possibility that both of them are under-segmented. Thus, both of them need to be re-segmented. Figure 2(c) is the result after re-segmentation of Figure 2(a).

The thresholds t_1 and t_2 are carefully selected in the SEG2. They are smaller than that of the SEG1 so that the under-segmented segments can be separated in SEG2. Here, we intend to increase the over-segmentation rate. Once a segment is re-segmented and several segments are given out after SEG2, a classifier will be used to select the best combination of them.

3. Feedback from classifier

The purpose of introducing a classifier after re-segmentation is for feedback system to acquire useful information in the form of confidence values. We are not trying to get final recognition result at this stage. However, the classifier should be effective in providing useful confidence value in a fairly short time.

The method used in the classifier is whole word recognition with classification by multiple neural networks trained by back propagation. The classifier is composed of three simpler neural networks. The confidence value is obtained from the sum of the outputs from three neural networks. The classifier is trained on 5317 words from 32 classes. The recognition result on a CENPARMI test set of 2514 words is 86.4% [9].

When incorporating the classifier after re-segmentation, not all the segments that are sent to SEG2 need verification from the classifier because some segments remain intact. For these cases, we directly send them to the recognition module. Only those that are split in SEG2 are verified. As mentioned above, it is very likely that some words have been over-segmented. Thus, combination needs to be done. However we do not want to try all the possible combinations by the classifier because the computation complexity would be too high. Hence, we adopt a simplified algorithm for the combination as illustrated in Figure 3.

Algorithm: combination(x_1, \dots, x_n)

```

S ← ∅
for i ← 1 to n
  for j ← i to n
    Try all combinations from  $x_i$  to  $x_j$ 
    Get confidence value  $c_{ij}$  from classifier
  end for
   $c_k = \max(c_{ij}, \dots, c_{in})$ .  $k \in [j, n]$ 
   $S = x_k \cup S$ 
   $i = k$ 
end for
return(S)

```

Figure 3. Algorithm for combination.

In fact, the incorporation of classifiers can be made from the beginning of the segmentation module. However, there may exist too many segments after the first segmentation, which may greatly increase the computation complexity. A direct consequence is that the processing time of the system may be multiplied according to the number of segments. In our system,

each time the classifier is called, the segments need to be combined come from only one segment after SEG1. Thus, the number of segments is fairly small compared with the number coming from direct segmentation of the original legal amount image. It makes the implementation of our combination method practical and effective. Figure 2(d) shows the final segmentation result of Figure 2(a).

4. Experiments and results

The proposed approach was tested on a new CENPARMI database that contains 389 English legal amount samples from bank cheques. Comparing with the previous segmentation approach implemented by Guillevic and Suen in the CENPARMI cheque processing system[3], the correct segmentation rate has been increased for 13.4%. The experiment result is shown in Table 1. It is difficult to compare our result with those provided by other researchers due to the use of different databases. One research was reported by Paquet and Lecourtier. They tested their word separation algorithm on a restricted database involving 10 writers and acquired a successful segmentation in 50% of the images [7].

Table 1. Performance on English test set (389 samples)

	Correct	Under	Over	Noise
<i>Guillevic et al.</i>	58.6%	34.9%	5.7%	0.8%
<i>Zhou et al.</i>	72.0%	16.9%	10.3%	0.8%

The “Correct” word segmentation indicates that all the words in the amount are isolated successfully. The samples in the “Under” and “Over” categories are under-segmented or over-segmented. Some useless parts such as small dashes and printed characters may affect the recognition result if they are not removed successfully and are grouped with the words in the samples. We also consider these cases as under-segmentation. The samples in the “noise” category come from the errors in the binarization step and the item extraction module. A typical example is that the background is not completely removed and is connected with some words. It will prevent the segmentation approach from finding and calculating the gaps between the connected components. Some samples of these four categories are shown in Figure 4.

Comparing with the previous approach, the under-

segmentation problem has been suppressed considerably. But under-segmentation and over-segmentation still remain as the major sources of errors. When words are connected together or connected with lines, printed characters and noises, it is very difficult to isolate each word correctly. This may be a reason why the literature on word segmentation is very limited.

5. Conclusion

We have presented the feedback-based word segmentation approach for separating legal amounts in bank cheques. This approach has been tested on a new CENPARMI database obtained from bank cheques. The approach has improved the correct rate of segmentation by 13.4% comparing with the previous approach in the CENPARMI cheque processing system.

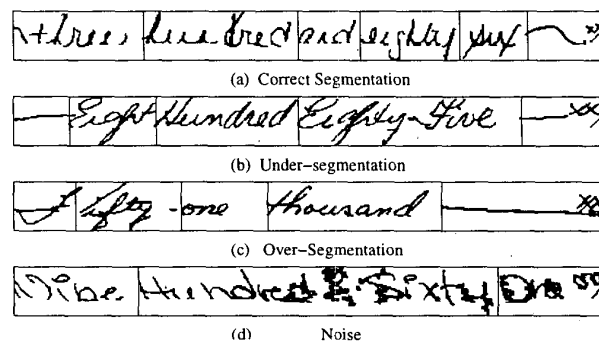


Figure 4. Word segmentation results

Acknowledgement

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

References

- [1] G. Dzuba, A. Filatov, D. Gershuny, I. Kil, and V. Nikitin. Check amount recognition based on the cross validation of courtesy and legal amount. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(4):639-655, June 1997.
- [2] D. Guillevic. *Unconstrained handwriting recognition applied to the recognition of bank cheques*. Ph. D Thesis, Concordia University, Montreal, Canada, 1995.
- [3] D. Guillevic and C. Y. Suen. Recognition of legal amounts on bank cheques. *Pattern Analysis and Applications*, 1(1):28-41, 1998.

- [4] G. Kim, V. Govindaraju, and S. N. Srihari. An architecture for handwritten text recognition systems. *International Journal on Document Analysis and Recognition*, 2(1):37-44, 1999.
- [5] K. Liu, C. Y. Suen, and C. Nadal. Automatic extraction of items from cheque image for payment recognition. In *Proc. of the International Conference on Pattern Recognition*, pages 798-802, Vienna, Austria, August 1996.
- [6] U. Mahadevan and R. C. Nagabhushanam. Gap matrices for word separation in handwritten lines. In *Proc. of Third International Conference on Document Analysis and Recognition*, pages 124-127, Montreal, Canada, August 1995.
- [7] T. Paquet and Y. Lecourtier. Handwriting recognition: application on bank cheques. In *Proc. of the International Conference on Document Analysis and Recognition*, pages 749-757, Saint Malo, France, September 1991.
- [8] G. Seni and E. Cohen. External word segmentation of off-line handwritten text lines. *Pattern Recognition*, 27(1):41-52, January 1994.
- [9] N. W. Strathy. Handwriting recognition for cheque processing. In *Proc. of the Second International Conference on Multimodal Interface*, volume 3, pages 47-50, Hong Kong, China, 1999.
- [10] C. Y. Suen, J. Kim, K. Kim, Q. Xu, and L. Lam. Handwriting recognition - the last frontiers. In *Proc. 15th International Conference on Pattern Recognition*, pages 1-10, Barcelona, Spain, September 2000.
- [11] C. Y. Suen, L. Lam, D. Guillivic, N. W. Strathy, M. Cheriet, J. N. Said, and R. Fan. Bank check processing system. *International Journal of Imaging Systems and Technology*, 7(4):392-403, 1996.
- [12] C. Y. Suen, K. Liu, and N. W. Strathy. Sorting and recognizing cheques and financial documents. In *Proc. of the Third International Association for Pattern Recognition Workshop on Document Analysis Systems*, pages 1-18, Nagano, Japan, November 1998.