Water Reservoir Based Approach for Touching Numeral Segmentation

U. Pal[1,2], A. Belaïd[1] and C. Choisy[1]
[1]Group READ, LORIA, Campus Scientifique B. P. 239
54506 Vandoeuvre-Les-Nancy, France
(pal, abelaid, choisy)@loria.fr
[2]Computer Vision and Pattern Recognition Unit
Indian Statistical Institute, 203 B. T. Road, Calcutta-35, India
Email: (umapada)@isical.ac.in

Abstract

This paper deals with a new scheme for automatic segmentation of unconstrained handwritten connected numerals. The scheme is mainly based on features obtained from a new concept based on water reservoir. A reservoir is a metaphor to illustrate the region where numerals touch. Reservoir is obtained by considering accumulation of water poured from the top or from the bottom of the numerals. At first, considering reservoir location and size, touching positions (top, middle and bottom) are decided. Next, analyzing the reservoir boundary, touching position and topological features of the touching pattern, the best cutting point is determined. Finally, combined with morphological structural features the cutting path for segmentation is generated.

1. Introduction

The segmentation of connected handwritten digits is the main bottleneck in the handwritten recognition system. In general, there are two types of segmentation schemes: Recognition-free-segmentation and Recognition-based-segmentation. In recognition-free-segmentation, a numeral string can be divided into segments by rules without recognition. In recognition-based-segmentation, candidate segmentation points are verified with recognizer. Here a recognition-free-segmentation scheme is proposed.

In the past years many algorithms for the segmentation of connected numerals have been proposed [1,2,3,5-10]. One class of approaches use contour features of the component for segmentation. Analyzing the contour of a touching pattern valley and mountain points are derived. Next, cutting path is decided to segment the touching pattern by joining valley and mountain points. In general contour based methods do not provide more accurate results. The situations where contour based method may fail are given in the reference [1]. Some researchers use profile features for touching numeral segmentation [2]. Profile based methods fail when the handwritings are strongly skewed or overlapped. Another class of approaches is based on thinning [1,3]. In this approach, thinning of foreground and/or background pixels of a connected pattern are processed. The end point and fork points obtained by thinning are used for cutting point extraction of the touching pattern. Thinning based method is time consuming and generate protrusions. These protrusions sometimes generate wrong results in the segmentation.

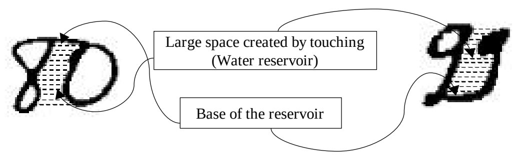

Figure 1. Examples of touching numeral and Space created by the touching.

In this paper, a more direct approach for classification and segmentation of numerals is proposed without using any thinning or normalization. The classification part of the scheme first detect whether a component of numeral(s) is isolated or touching. If it is touching, segmentation scheme is applied on it. When two numeral touch each other, they create large space between the numerals. This large space is very important for segmentation. The touching points of the connected numerals lie on the boundary of this large space (see Fig.1). Using a simple concept based on water reservoir this space is encountered for segmentation. If water is poured from top (bottom) of a connected numeral then water will be stored in this large space. This water stored area is called "water reservoir". For illustration see Fig.1. At first, the positions and sizes of the reservoirs are analyzed and a reservoir is detected where touching is made. Considering the type (top or bottom reservoir) and analyzing *base* of this reservoir the touching position (top, middle or bottom touching) is decided. Now noting touching position and analyzing the profile of the reservoir the initial feature points for segmentation are determined. By considering close loops, reservoir heights and distance from center of the component the initial feature points are ranked and the best feature point (highest rank point) is noted. Finally, based on the best feature point, close loop positions and morphological structure of touching region the cutting path is generated. The flow-chart of the scheme is shown in Fig.2.

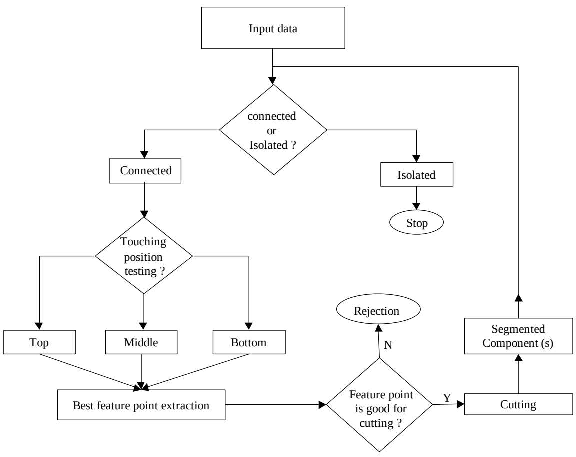

Figure 2. Block diagram of the scheme.

2. Water reservoir principle

The water reservoir principle is as follows. If water is poured from top and bottom of the numeral, the cavity regions of the numerals where water will be stored are considered as reservoirs. For illustration see Fig.3. Here by top (bottom) reservoirs we mean the reservoirs obtained when water is poured from top (bottom). (Here, water pouring from bottom we mean the water pouring from top after rotating the component by 180°). We analyze these top and bottom reservoirs for the segmentation.

All reservoirs obtained in this way may not be considered for further processing. Those reservoirs whose heights are greater than a threshold T_1 are considered for future processing. The value of T_1 is 1/8 of the numeral height (The threshold value is obtained from the experiment).

3. Feature selection

Since different individuals can have various writing styles, the features to be chosen should be independent of writing styles of the individuals. To achieve writing independent features we consider here reservoir based features.

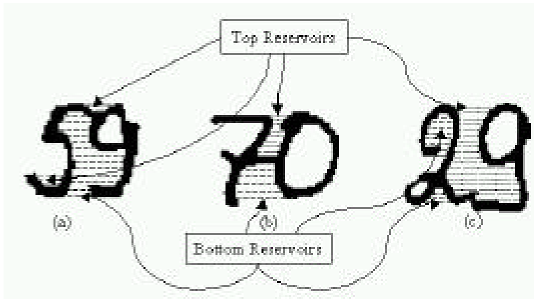

Figure 3. Reservoirs obtained from water flow from top and bottom are shown for (a) Top (b) middle and (c) bottom touching numerals. Top Reservoirs are marked by dots and bottom reservoirs are marked by small line segments.

The important reservoir features considered in the scheme are: (i) number of reservoirs (ii) positions of reservoirs with respect to bounding box of the touching pattern (iii) size and shape of the reservoirs (iv) center of gravity of the reservoirs (v) relative positions of the reservoirs. Except the reservoir features, some structural and topological features are also used. Close loop features were the main topological features in our scheme. Number of close loop, positions with respect to the bounding box of the component, center of gravity and the ratio of close loop height to component height are considered here. In the structural feature we consider morphological structural pattern of the touching region.

4. Numeral classification

This stage of the scheme classifies an input numeral string into isolated or touched digit group. Generally in the earlier studies, aspect ratio (height width ratio) of the component is used for the classification of isolated or connected numerals [2]. If two or more numerals are connected, then the width of the connected component should be larger than its height. Although, by and large, it may true in printed text, it is not true in handwritten cases because of different writing style of different individual. For example, from the experiment it is noticed that in some handwriting, the digits "2" and "1" are connected and the width of the touching component is not greater that the isolated numeral “4”. Complexity (in terms of curvature) is also used for isolated/touching numeral separation. The idea of using complexity is as follows. When two or more numerals are connected, then the connected pattern should be more complex than isolated numeral. This is not true in general. For example, if two zeros touch each other to make a touching pattern, the complexity of this touching pattern is more or less equal to the numeral “8”.

In principle, when two numerals get connected one of the two followings happens in most of the cases: (1) two numerals create a large reservoir (for example see the Fig.1). (2) the number of reservoirs (obtained from both top and bottom) in a connected numeral will be greater than that of a isolated numeral. Based on the number of reservoirs, their size and positions, and number of close loop and their location, the isolated digit and touching digit groups are identified.

For classification, at first, number of close loop is tested. If there are two or more close loops and the positions of two close loops are side by side the component is connected. Else, the number of reservoirs is used. If the number of reservoir is greater than 3 the component is touching. Else reservoir features are used. To get a more precise idea of reservoir features consider the examples of numerals shown in Fig.3. In the first and third touching numerals of this figure, along with other reservoirs we get a big reservoir whose height is nearly equal to the height of the component. This big reservoir identifies a touching numeral from isolated numeral. In the second touching component (Fig.3(b)), two reservoirs are top-bottom fashion which is another property of touching numeral separation.

The advantage of this method is that it is size independent and there is no need any normalization of the component. To evaluate the performance of the method a data set of 3190 components was extracted from French bank checks. The proposed method has 98.75% accuracy for separating isolated digits and touched strings. To make a comparison, result obtained by Kim et al.[4] is also provides in Table 1.

Table 1: Comparison of classification accuracy.

Method	Data size	Accuracy
Our approach	3190 [2934 (Iso) + 256 (Conn)]	98.75% [99.12% (Iso) 94.56% (Conn)]
Kim et al. [4]	2500 [2000 (Iso) + 500 (Conn)}	96.5% [97.9% (Iso) 91.0% (Conn)]

Iso = Isolated. Conn= Connected.

5. Touching digit segmentation

For the segmentation of touching pattern at first, the touching position is found. The touching position has been classified into three regions: top, middle and bottom. Next, based on the touching position, reservoir position, topological features (number and position of close loops) of the component etc., the feature points for segmentation are extracted. Finally, considering loops, structural features and reservoir features the path to segment the touching pattern is constructed.

Touching position detection: Let BB be the bounding box area of a touching component. BB is divided horizontally in three regions. The top region (h_t) is 25% of BB. Middle region (h_m) and bottom region (h_b) are 50% and 25% of BB, respectively. Similarly, the vertical division of BB in three regions is done. The left (v_l), middle (v_m) and right (v_r) regions are 25%, 50% and 25% of BB, respectively. For illustration see Fig.4. (Based on experiment these values are decided). At first, the largest reservoir of the component whose center of gravity lies in v_m region is found. This reservoir is called as the *best reservoir* for touching. The *base-line* (lowermost row of the reservoir) of the best reservoir is then detected. The best reservoir and its base-line are shown in Fig.4. If the base line of the best reservoir of a touching numeral lies in h_t, h_m, h_b regions then the touching numeral is top, middle, bottom touching, respectively.

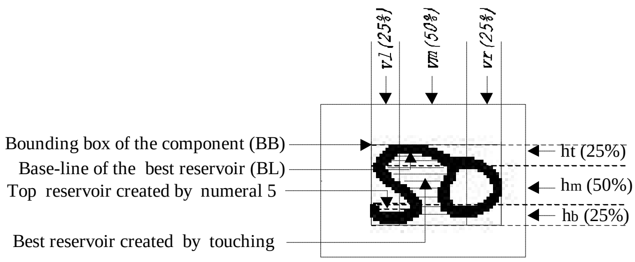

Figure 4. Feature detection approach.

Feature point extraction: For feature points extraction the touching position is noted. If the touching position is top then all reservoirs whose base-line lies in the h_t region are considered for feature extraction. Similarly, if the touching position is bottom (middle) then all reservoirs whose base-line lies in the h_b (h_m) region are considered for feature extraction. The leftmost and rightmost points of the base-line of considered reservoirs are the feature points. These points are initial feature points. If the distance (L) between the leftmost and rightmost points of a base-line is less than 2R (R is the length of most frequent occurring horizontal black run of a touching component) then instead of two feature points (leftmost and rightmost points of base line) we consider the midpoint of leftmost and rightmost points as an initial feature point The initial feature points for three different touching components are shown in Fig.5. For touching numerals shown in Fig.5(a), only one reservoir is considered for feature extraction, because the base lines of other reservoirs do not line in the region h_t. Now from initial feature points the best feature point (which gets maximum confidence value) is chosen for segmentation. To compute confidence value (CV) following features are considered. (1) Euclidean distance of feature points from the center of gravity of the touching component. Let there are F initial feature points and the Euclidean distance of these points from center of gravity (CG) of the component are d_1, d_2 …….d_F. Then the confidence value of a feature point with distance d_i is 1/K, where K = $d_i/(d_1+d_2+……d_F)$. (2) Distance from the CG of the close loop. Confidence value detection procedure for this feature is similar to above feature. (3) Height of the reservoir. The main idea of this feature is as follows. Those points which come from the bigger reservoir should get more confidence value for this feature. Let F initial points are obtained from p reservoirs of heights R_1, R_2, …. R_P. The confidence value of a point obtained from reservoir R_i is R_i/S, where S = $R_1+R_2+…. +R_P$. The best feature points for the three components shown in Fig.5 are shown by A.

Other existing methods [1,2,3,6,8] consider whole touching pattern for feature points extraction. Instead of whole pattern a portion of the touching component is considered here for feature points extraction. Which is an important characteristics of the proposed method.

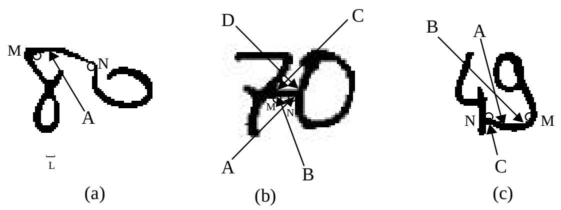

Figure 5. Initial feature points (A, B, C, D) and node points (N, M) are shown in (a) Top (b) middle and (c) bottom touching numerals. A is the best feature point. L is length of base-line.

Determination of segmentation path: Both the straight line and curve segmentation are done in the scheme. The segmentation algorithms are different for two different cases of touching numerals. Case 1: components having two side by side and touching close loops. Case 2: other touching components.

For the first case, segmentation is done through the middle of their common touching area. From the common touching portion a starting point is detected. (detection of starting point is given below). From this point a pointer is moved upwards though the middle of the touching area until it reaches to a reservoir (or boundary of the component). The path moved by the pointer is noted. Similarly, the pointer is moved downwards from the starting point and the path is noted. The total path moved by the pointer is the segmentation path. An example of this type touching numeral and its segmentation result are shown in the first column of Fig.6.

The starting point detection is done as follows. The center of gravity points of two close loops are noted and the line segment obtained by joining these two points is examined. The portion of this line segment which lies in the common touching area of the close loops are detected. The middle point of this portion is the starting point.

For the second case, from the best feature point the boundary of the reservoir (the reservoir from which the best feature point is obtained) is traced pixel-wise in clock-wise direction to find a *node point*. During tracing, the length of vertical black run at each tracing point is computed. Here, by run at each tracing point means the vertical black run which contains the point. The boundary point where this run length is greater than 3R/2 is considered as node point (R is described earlier). Let this node point be N. Similarly, boundary of the reservoir is traced from the best feature point in anti-clock-wise direction to find another node point (M). See Fig.5 where node points obtained in the touching numerals are shown. As our best feature point detection is based on the base-line of water reservoir, the best feature point may not always obtained in the exact position where cutting should be made. This situation is tackled by the node points and obtained good segmentation path.

For the top and bottom touching components vertical segmentation is done. From the two node points best node point is chosen and the cutting is done vertically at the best node point. To choose best node point we note positions of the node points. If one of the node points lies in v_m region (v_m region is shown in Fig.4), the node within v_m region is the best node. If both the node points lie in the v_m region then the node which is nearer to the biggest close loop in the component is the best node. The node nearer to the middle of the touching component is the best node point if no close loop is obtained in the component. If no node point lies in the v_m region we segment the component at the middle of the two node points. See the components shown in 3rd and 5th columns of Fig.6. Here, segmentation is done at the middle of the two node points (node points are shown by small circles).

For the middle touching components the cutting method is different. From the best node we try to associate the other initial features (see Fig.5 where initial points are shown) which are obtained from the opposite type reservoir to best node reservoir and choose a point (best associated point) from these initial points. The line obtained by joining this pair is the segmentation path. If the best node point is obtained from top reservoir then other initial points which are obtained from bottom reservoirs will be considered to find the best association point. The best association is measured by the distance of these points and the location of the reservoirs. For an example, consider the component shown in Fig. 5(b). Here, the best feature point (A) and best node point (N) are same. Since the best node point comes from bottom reservoir, so other initial feature points (C and D) obtained from top reservoirs are considered to find best associate point of N. Here, D is the best associate point of A. The line obtained by joining the points A and D is the segmentation path of this component.

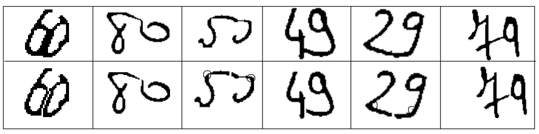

Figure 6. Different types of touching components (shown in 1st row) and their segmentation results (shown in 2nd row).

6. Results and Discussion

The evaluation of the segmentation scheme was done on 978 images of numeral string of French bank check courtesy amount. The results were verified manually and observed that 94.34% of the connected numerals were correctly segmented.

Table 2. Performance of various segmentation approaches.

Method	Source [size] of data set	Accuracy	Rejection rate
Cheng & Wang [1]	UN[150]	96%	7.8 %
Lu et al (HRR) [3]	NIST [823]	97%	28.6%
Lu et al (LRR) [3]	NIST [823]	92.5%	4.7%
Chi et. al (HRR) [5]	UN [UN]	95.1%	32.7%
Chi et. al (LRR) [5]	UN [UN]	89.2%	2.8%
Congedo et al [6]	CEDAR [UN]	91%	0%
Oliveira et al. [8]	BC [900]	95.24%	0%
Shi et al [9]	USPS [2579]	85.7%	0%
B. Zhao et al [10]	UN [172]	87.2%	0%
Our approach	FC [978]	94.34%	3.16%

Here, BC = Brazilian bank check, LRR=Low Rejection rate HRR= High rejection rate, UN= Unknown

The rejection rate of our segmentation system was 3.16%. The main features for rejection were: (1) the widths of one of the segmented part is very small compare to width of the other part (2) the length of cutting path is very long compare to the height of the touching pattern (3) no best reservoir is obtained in a touching component. At present, no confusion stage is implemented in our segmentation scheme. In future we plan to implement it. Fig.6 shows some numeral strings and their segmentation results obtained by the proposed approach. From our experiment we noticed that most of the errors came from those components where touching numerals have some common portion. Errors also obtained from the double touching components.

To get an idea about the performances of some earlier pieces of work and the proposed work, a performance table (Table 2) is given. The drawback of the proposed method is that it will fail if there is a break point on the contour used as the boundary of a reservoir. In that case reservoir cannot get properly and hence miss-segmentation occurs. We observed that such cases occur rarely.

References

[1] Y. K. Chen and J. F. Wang, "Segmentation of handwritten connected numeral string using background and foreground analysis", Proc. 15th ICPR, pp. 598-601, 2000.

[2] H.Fujisawa, Y.Nakano and K.Kurino, "Segmentation methods for character recognition from segmentation to document structure analysis", Proceeding of the IEEE, vol.80, pp.1079 - 1092, 1992.

[3] Z. Lu, Z. Chi, W. Siu and P. Shi, "A background-thinning based approach for separating and recognizing connected handwritten digit strings", Pattern Recognition, vol.32, pp. 921-933, 1999.

[4] K. Kim, J. H. Kim and C. Y. Suen, "Recognition of Unconstrained Handwritten Numeral Strings by Composite Segmentation method", Proc. 15th ICPR, pp. 594-597, 2000.

[5] Z. Chi, M. Suters and H. Yan, "Separation of single and double touching handwritten numeral strings", Optical Engineering, vol. 34, pp.1159-1165, 1995.

[6] G. Congedo, G. Dimauro, S. Impedovo and G. Pirlo, "Segmentation of numeric string", Proc. 3rd ICDAR, pp. 1038-1041, 1995.

[7] S. Knerr, V. Anisimov, O. Baret, N. Gorski, D. Price and J. C. Jimon, "The A2iA intercheque system: Courtesy amount and legal amount recognition for French checks", Automatic Bank Check Processing, editors, S. Impedovo, P.S.P. Wang and H. Bunke, World Scientific, pp. 43-86, 1997.

[8] L. S. Oliveria, E. Lethelier, F. Bortolozzi and R. Sabourin, "A new approach to segment handwritten digits", Proc. 7th IWFHR, pp. 577-582, 2000.

[9] Z. Shi, S. N. Srihari, Y. Shin and V. Ramanaprasad, "A system for segmentation and recognition of totally unconstrained handwritten numeral string", Proc. 4th ICDAR, pp. 455-458, 1997.

[10] B. Zhao, H. Su, S. Xia, "A new method for segmenting unconstrained handwritten numeral string", Proc. 4th ICDAR, pp. 524-527, 1997.