RESULTS FROM A PERFORMANCE EVALUATION OF HANDWRITTEN ADDRESS RECOGNITION SYSTEMS FOR THE UNITED STATES POSTAL SERVICE

D'Amato, D.P.; Kuebert, E.J.; Lawson, A.

Published in:
EPRINTS-BOOK-TITLE

IMPROVEMENT IN HANDWRITTEN NUMERAL STRING RECOGNITION BY SLANT NORMALIZATION AND CONTEXTUAL INFORMATION

ALCEU DE S. BRITTO JR.

Pontifícia Universidade Católica do Paraná (PUC-PR), R. Imaculada Conceição, 1155
Curitiba (PR) 80215-901 – Brazil
Universidade Estadual de Ponta Grossa (UEPG), Praça Santos Andrade S/N, Centro,
Ponta Grossa (PR) 84100-000 – Brazil

E-mail: alceu@cenparmi.concordia.ca

ROBERT SABOURIN

École de Technologie Supérieure (ETS), 1100 Rue Notre Dame Ouest
Montreal (QC) H3C 1K3 - Canada
Centre for Pattern Recognition and Machine Intelligence (CENPARMI), 1455 de
Maisonneuve Blvd. West, Suite GM 606 - Montreal (QC) H3G 1M8 - Canada

E-mail: sabourin@gpa.etsmtl.ca

EDOUARD LETHELIER AND FLAVIO BORTOLOZZI

Pontifícia Universidade Católica do Paraná (PUC-PR), R. Imaculada Conceição, 1155
Curitiba (PR) 80215-901 - Brazil

E-mail: {edouard, fborto}@ppgia.pucpr.br

CHING Y. SUEN

Centre for Pattern Recognition and Machine Intelligence (CENPARMI), 1455 de
Maisonneuve Blvd. West, Suite GM 606 - Montreal (QC) H3G 1M8 - Canada

E-mail: suen@cenparmi.concordia.ca

This work describes a way of enhancing handwritten numeral string recognition by considering slant normalization and contextual information to train an implicit segmentation-based system. A word slant normalization method is modified in order to improve the results for handwritten numeral strings. We assume that each connected component (CC) in the string has its own slant. The slant and contour length of each CC are used for obtaining the mean slant of the string. Both the original and modified methods are evaluated by means of some interesting analyses on the NIST SD19 database. These analyses show (a) the positive impact of slant correction on the number of overlapping numerals in strings, and (b) the difference in normalizing isolated numerals based on the slant estimated from their own images and the slant estimated from their original string images. Slant normalization and contextual information regarding string slant and digit size variations within the string are used to train numeral HMMs. Preliminary string recognition results, produced by a system under construction, are shown.

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 323-332

1 Introduction

Slant normalization plays an important role in different methods of recognizing numeral strings. The general objective is to reduce script variability. However, we can also identify some specific objectives, which are dependent on the recognition approach.

For instance, in implicit segmentation-based methods, usually a vertical strip is used to scan the word image for feature extraction. The slant may be corrected in order to minimize the problem when slanted characters overlap adjacent characters, interfering in the columns of pixels extracted from them.

This paper focuses on the use of slant normalization in this context. Let us consider the system in Figure 1(b). In this example, the numeral strings are slant-normalized in order to reduce the overlap between adjacent numerals. To this end, a slope (θ_2) is estimated from the whole numeral string image. The string recognition is done by matching numeral HMMs against the normalized string by means of dynamic programming. The construction of these models is presented in Figure 1(a). They are trained from isolated numerals, which are also slant-normalized. However, a different slope (θ_1) is used, which is estimated from the numeral image. Although the same slant normalization method may be used to estimate θ_1 and θ_2 , they may present a significant difference.

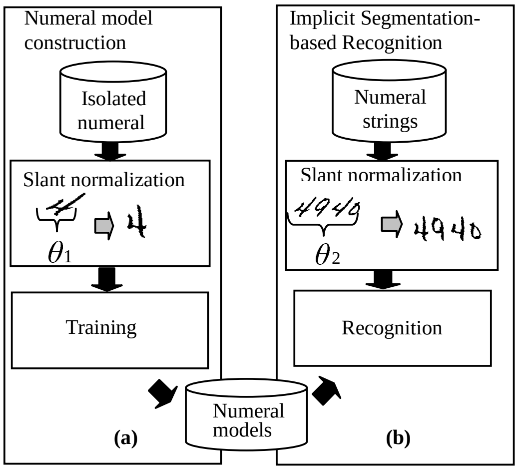


Figure 1. Slant normalization and a hypothetical implicit segmentation-based system

At this point, we can formulate the following questions:

1. What is the real impact of slant normalization on the number of overlapping numerals in the strings?
2. How different are θ_1 and θ_2 ?

3. If the difference between $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ is really significant, what is the real contribution to string recognition of considering some contextual information during slant normalization, *i.e*. correcting the slant of the training samples using the slope estimated from their original strings ($\boldsymbol{\theta_2}$)?

To answer these questions we have first adapted a word slant normalization method to normalize numeral strings. After that, we have performed some interesting analyses using well-segmented numeral strings extracted from NIST SD19 database. These analyses have also been used to compare the original and the modified slant normalization algorithms.

This study investigates the real contribution to string recognition of considering contextual information on the slant normalization of numerals used for training the HMMs, and also for dealing with intra-string size variations. Preliminary recognition results are given using an implicit segmentation-based system currently under construction.

This paper is organized into 6 sections. Section 2 presents the adaptation of an approach to word slant normalization technique to correct the slant of numeral strings. Section 3 shows some interesting analyses performed on the NIST SD19 database. In Section 4, experimental results on both isolated numerals and numeral strings are shown. A discussion is presented in Section 5, and the conclusion in Section 6.

2 Slant normalization method

Many methods for word slant normalization can be found in the literature [1-4,7]. However, the majority is strongly based on heuristics. An exception is the KSC method [4], where the chain code of the word contour is used to estimate slant. The horizontal chain elements (code $\boldsymbol{n_0}$) are not considered, and the other elements are divided into slant chain segments: $\boldsymbol{n_1}$, $\boldsymbol{n_2}$ and $\boldsymbol{n_3}$, respectively 45°, 90° and 135°. The average orientation of these segments is given by:

$$\theta = \tan^{-1}\left(\frac{n_1 + n_2 + n_3}{n_1 - n_3}\right) \tag{1}$$

The slant is estimated from the whole word, which is unsatisfactory for numeral strings, since they have individual components (digits or fragments), each of which has its own slant. However, an independent correction of each component is also not viable, since this may produce distortions when broken digits are present in the numeral string.

We propose a modified KSC method (MKSC), which consists of estimating the mean slant based on the average of the slants of each string component. The string slant is obtained through a weighted average, where the slant of each string component is weighted by its contour length. Let $\boldsymbol{N}$ be the number of components (digits or parts of digits) in a numeral string; the mean slant can be estimated by the following expression:

$$\bar{\theta} = \frac{\sum_{i=1}^{N} (def - \theta_i) \times w_i}{\sum_{i=1}^{N} w_i} \qquad (2)$$

where, ***def*** is the normal slant (90^o),

$$\theta_i = \tan^{-1}\left(\frac{n_{i1} + n_{i2} + n_{i3}}{n_{i1} - n_{i3}}\right) \qquad (3)$$

and $\boldsymbol{w_i}$ is the chain code length of the $\boldsymbol{i^{th}}$ connected component. The objective of using a weighted average is to avoid distortions, which can be caused by small fragments in the string.

3 Analyses on the NIST SD19 database

In these analyses we have used a total of 44,256 well-segmented numeral strings from the *hsf_0, hsf_1, hsf_2* and *hsf_3* series of the NIST SD19 database, and 197,784 isolated digits extracted from them. To create this data set we implemented a process based on the digit classifier proposed in [8]. A handwritten numeral string is considered a "well-segmented string" when its components are recognized as isolated digits. Moreover, the recognition result of this string, from its digits, must correspond to that labeled by NIST. In this data set the isolated digits have a link to their original strings.

3.1 Impact of slant normalization on the number of overlapping numerals

One objective of this analysis is to estimate the impact of correcting the string slant on the number of overlaps between adjacent numerals. Both KSC and MKSC methods are used in this analysis.

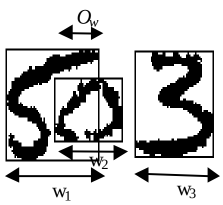


Figure 2. Overlap estimation

The overlap between adjacent numerals is estimated from the overlap between adjacent bounding boxes. For example, given the overlapping numerals in Figure 2, the algorithm calculates for each bounding box involved its overlapping percentage ($\boldsymbol{Op_i}$), as:

$$Op_i = O_w / w_i * 100 \qquad (4)$$

In this example the overlap (O_w) represents 47.6% of the width of the first bounding box (numeral 5), and 65.2% of the width of the second one (numeral 0). Figure 3 shows the impact of slant correction on the number of overlapping numerals in the 44,256 numeral strings analyzed. Before correcting the string slant, there were 9,425 numeral strings with overlapping numerals, which represents 21.29% of all analyzed strings. After correcting the string slant using the KSC method, the number of strings with overlapping numerals was reduced to 4,767, which represents 10.77% of all analyzed strings.

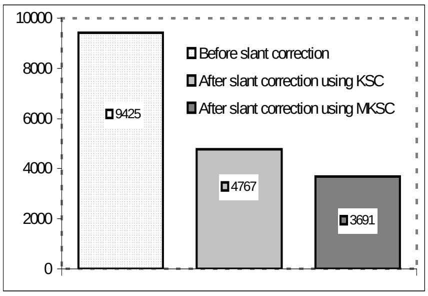


Figure 3. Number of overlapping numerals before and after slant correction using the MKSC method

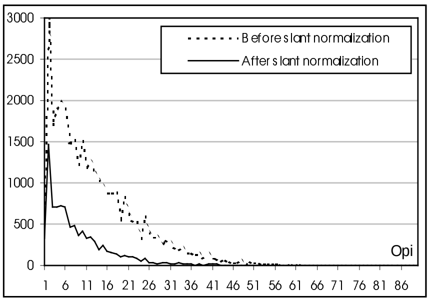


Figure 4. Distribution of overlapping percentages (***Op***) before and after slant correction using the MKSC method

The number of overlapping numerals was reduced by 49.42%. On the other hand, using the MKSC method, the reduction was more significant. After correcting the string slant using this method, the number of strings with overlapping numerals was reduced to 3,691, which represents 8.34% of all analyzed strings. The number of overlapping numerals was thus reduced by 60.83%. Figure 4 shows the ***Op*** distribution before and after slant correction using the MKSC method.

3.2 Slant normalization with and without contextual information

The KSC and MKSC methods are also used to show the difference between correcting the slant of isolated numerals extracted from strings using two techniques: slant correction with contextual information, where the slant is estimated from the original string; and slant correction without contextual information, where the slant is estimated for each isolated numeral in particular, without taking into account the origin. The objective of this analysis is to evaluate the difference between both techniques in terms of the slope estimated. The isolated numerals extracted from well-segmented strings are used in this analysis (197,784 samples), since they have a link to their original strings. For each numeral image $\Delta\theta$ is calculated, as $\Delta\theta = |\theta_1 - \theta_2|$, where θ_1 is the slant of a numeral, estimated from its image; θ_2 is the slant of a numeral, estimated from its original string image.

In addition, we have calculated the $\Delta\boldsymbol{\theta}$ mean ($\mu_{\Delta\theta}$) and the dispersion ($\vartheta_{\Delta\theta}$) for each numeral class. The results are shown in Table 1 for both slant correction methods. We can see that the algorithms for slant correction show similar results. However, the MKSC method provides a smaller dispersion than that obtained by using the KSC method for all the numeral classes.

Table 1. $\Delta\boldsymbol{\theta}$ mean and dispersion using KSC and MKSC methods

Class	KSC method		MKSC method		Class	KSC method		MKSC method	
	$\mu_{\Delta\theta}$	$\vartheta_{\Delta\theta}$	$\mu_{\Delta\theta}$	$\vartheta_{\Delta\theta}$		$\mu_{\Delta\theta}$	$\vartheta_{\Delta\theta}$	$\mu_{\Delta\theta}$	$\vartheta_{\Delta\theta}$
0	4.76	4.82	5.05	4.20	5	6.85	5.75	6.51	5.18
1	6.30	6.89	6.94	6.35	6	5.45	5.89	5.93	4.90
2	6.08	5.51	5.95	4.76	7	7.11	7.11	7.40	6.10
3	6.40	5.51	6.15	4.94	8	4.73	4.55	4.92	3.98
4	4.74	5.79	5.18	4.74	9	4.23	5.05	4.70	4.12
Global						**5.66**	**5.83**	**5.88**	**5.07**

Using the MKSC method the mean variation becomes 5.88^o with a dispersion of 5.07^o, with the largest mean variation occurring in numeral class seven (7.40^o).

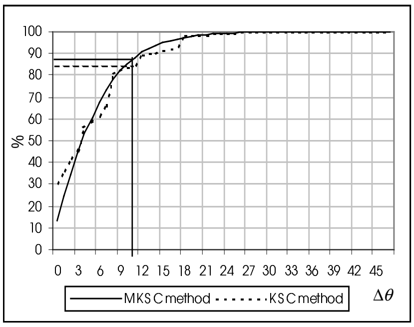


Figure 5. Cumulative distribution of $\Delta\boldsymbol{\theta}$

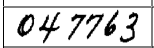
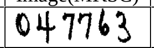
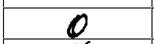
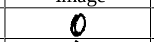
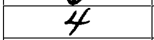
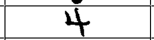
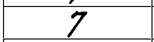
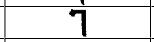
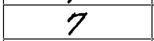
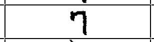
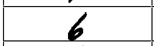
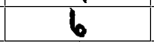
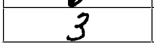
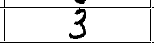
Orig. Image	Slant-corrected Image(MKSC)	$\boldsymbol{\theta_2}$	
		-35.6^o	
Orig. Image	Slant-corrected Image	$\boldsymbol{\theta_1}$	$\Delta\boldsymbol{\theta}$
		-18.4^o	**17.2^o**
		-44.9^o	9.3^o
		-44.9^o	9.3^o
		-44.9^o	9.3^o
		-44.9^o	9.3^o
		-14.0^o	**21.6^o**

Table 2. Example of numeral string with significant $\Delta\boldsymbol{\theta}$ (17.2^o and 21.6^o)

Figure 5 shows that there is a significant number of cases where $\Delta\boldsymbol{\theta}$ is greater than 10.95^o , i.e. ($\mu_{\Delta\theta} + \vartheta_{\Delta\theta}$). This number represents 11.74% of all the analyzed images (23,223 numeral images) using the MKSC method, and 15.23% of all the analyzed images (30,313 numeral images) using the KSC method. We can see that the modified method shows better results, since the slant estimated from the string is closer to that estimated from each isolated numeral. However, even using the MKSC method the difference is still significant. In Table 2 we can see an example, where $\Delta\theta$>$(\mu_{\Delta\theta}$+$2\times\vartheta_{\Delta\theta})$ based on the use of MKSC method.

4 Experimental results

In order to demonstrate the real contribution on string recognition by considering slant normalization and contextual information, an implicit segmentation-based system has been used. This system, currently under construction, matches numeral HMMs against the string using the Level Building Algorithm [5,6]. To this end, the numeral string is scanned from left-to-right, while local and global features are extracted from each column. The local features are based on transitions from background to foreground pixels and *vice versa*. For each transition the following features are calculated: its circular mean direction and the corresponding variance; the relative position, taking into account the top of the bounding box; and the information whether the transition belongs to the outer or inner contour. The global features are based on horizontal projection (HP) of black pixels for each column, and the derivative of HP between adjacent columns. The structure of the numeral HMMs was experimentally defined. The best results were obtained using the Bakis model with 5 states. A codebook with 128 entries was used.

During the experiments, we done a comparison in terms of isolated digit and numeral string recognition of considering no slant normalization, and slant normalization with and without contextual information (CI). In the experiment based on slant normalization without contextual information, each isolated numeral used for training the numeral HMMs is slant corrected using the slope estimated from its own image ($\boldsymbol{\theta_1}$). On the other hand, when contextual information is used, the slope for each isolated numeral is estimated from its original string ($\boldsymbol{\theta_2}$).

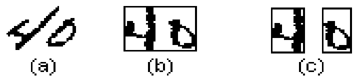


Figure 6. a) Original string; b) String bounding box after slant normalization; c) Training samples linked to their orginal strings and the bounding box used for feature extraction.

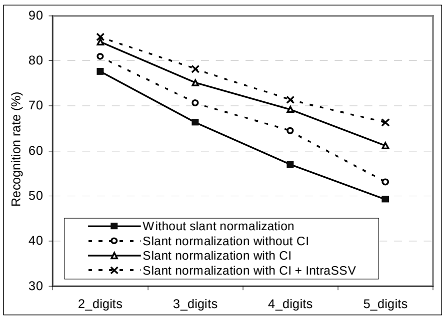
In addition, based on the same idea for providing similar experimental conditions to system training and testing, we have again experimented the use of contextual information for training the numeral HMMs. Here, the objective consists of taking into account the intra-string size variation (IntraSSV), *i.e.*, height variation of the digits in the strings. For this purpose, we have extracted features from the isolated numerals used for training the numeral HMMs, taking into account the bounding box height of their slant-normalized original strings (see Figure 6). To deal with the inter-string size variation, *i.e.*, the size variation between different strings, we have considered the use of size-invariant features.

The isolated digits used in the experiments come from the data set extracted from the well-segmented numeral strings. 50,000 samples are used for training, 10,000 for validation and 10,000 for testing. In all the experiments zero-rejection level is used.

Table 3. Preliminary recognition rates for handwritten isolated numerals

Experiments	Validation (%)	Test (%)
Without slant normalization	93.1	86.2
Slant normalization without CI	95.9	93.0
Slant normalization with CI	95.8	92.6
Slant normalization with CI + IntraSSV	94.2	91.1

Table 3 provides our preliminary recognition results for isolated digits. We can observe that slant normalization based on contextual information does not contribute to the recognition of isolated numerals. Moreover, the experiment based on IntraSSV loses in terms of numeral recognition. It is possible since the use of this additional contextual information increases the numeral variability.



Experiments	2_dig (%)	3_dig (%)	4_dig (%)	5_dig (%)
Without slant normalization	77.63	66.37	57.01	49.30
Slant normalization without CI	80.97	70.64	64.47	53.10
Slant normalization with CI	84.21	75.17	69.25	61.18
Slant normalization with CI + IntraSSV	85.32	78.19	71.34	66.32

Figure 7. Recognition rates by string class (2,3,4 and 5 digit strings)

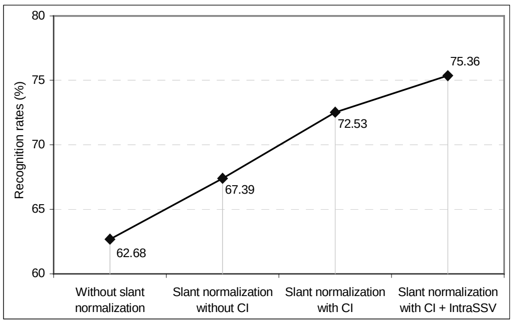


Figure 8. Global string recognition rates

On the other hand, the experiments using numeral strings show some significant improvement in recognition performance. In these experiments 9,416 numeral strings have been used. They have been extracted from the *hsf-7* series of the NIST SD19 database and distributed into 4 classes: 2_digit (2,370), 3_digit (2385), 4_digit (2,345) and 5_digit strings (2,316). These strings show different problems, such as touching, overlapping and fragmentation. We can see the recognition rates by class in Figure 7, while the global recognition rates are shown in Figure 8.

5 Discussion

The first analysis on the NIST database shows that slant normalization brings a significant reduction in the number of overlaps between adjacent numerals in strings. The MKSC method achieves a more significant reduction (60.83%) than that provided by the KSC method (49.42%). This answers our first question and shows us that slant normalization is really helpful in suppressing overlapping problems.

The second analysis shows that there is a representative part of the analyzed data, referring to those cases where $\Delta\boldsymbol{\theta} > (\mu_{\Delta\theta} + \vartheta_{\Delta\theta})$, in which the slants estimated with and without contextual information present a significant difference. In other words, the slant estimated from the isolated digit ($\boldsymbol{\theta_1}$) and that estimated from its string ($\boldsymbol{\theta_2}$) differ by more than 10.95^{o} in 11.74% of 197,784 analyzed digits using the MKSC method. This answers our second question, and at the same time justifies an investigation of the real contribution to string recognition of considering this contextual information in the slant normalization of the numeral training samples. Moreover, we observe that using MKSC we reduce the number of cases where $\Delta\boldsymbol{\theta}$ is greater than ($\mu_{\Delta\theta} + \vartheta_{\Delta\theta}$). This means that we can approximate the slant estimated from the string to the slant estimated from each isolated numeral in the string.

Finally we have the answer to the third question. Even with a system under construction, where the feature extraction method needs to be improved, we have observed interesting results. The slant normalization without contextual information has brought an improvement of 4.71% to the global string recognition rate. On the other hand, with slant normalization based on contextual information, we achieve an improvement of 9.85%.

In addition, the experiment based on intra-string size variation brought a further improvement of 2.83% to the global string recognition rate, even with a small loss in terms of numeral recognition rate. This shows again that the use of contextual information to provide the same conditions during training and testing is a promising strategy in implicit segmentation-based systems.

6 Conclusion and future work

This work focuses on the use of slant normalization and contextual information in an implicit segmentation-based system. We have adapted a word slant normalization method in order to improve the results for handwritten numeral strings. The original and modified methods are evaluated by means of some interesting analyses on the NIST SD19 database. These analyses include the impact of slant correction on the number of overlapping numerals in strings, and the difference in terms of slant estimated between two slant correction techniques of isolated numerals extracted from strings.

Some preliminary recognition results using a system under construction are shown. In the experiments, we evaluated the use of contextual information on the slant normalization of numerals used for training the system, and also to deal with the intra-string size variation. The results have been very encouraging.

In summary, these experiments have shown that the same experimental conditions should be used during system training and testing. The use of contextual information has contributed to this. Our next step aims at improving the feature extraction method and the numeral HMMs, taking into account as much contextual information as possible.

7 Acknowledgements

The authors wish to thank the Pontifícia Universidade Católica do Paraná (PUC-PR, Brazil), the Universidade Estadual de Ponta Grossa (UEPG, Brazil), the École de Technologie Supérieure (ETS, Canada) and the Centre for Pattern Recognition and Machine Intelligence (CENPARMI, Canada), which have supported this work.

References

1. Bozinovic R.M. and Srihari S., Off-line Cursive Script Word Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **11**(1989) pp. 68-83.
2. Brown M.K. and Ganapathy S., Preprocessing Techniques for Cursive Word Recognition, *Pattern Recognition* **16**(1983) pp.447-458.
3. El-Yacoubi A., Gilloux M., Sabourin R. and Suen C.Y., An HMM-Based Approach for Off-Line Unconstrained Handwritten Word Modeling and Recognition, *Pattern Analysis and Machine Intelligence*, **21**(1999) pp. 752-760.
4. Kimura F., Shridhar M. and Chen Z., Improvements of a Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words, *Proceedings of the 2nd ICDAR* (1993), pp. 18-22.
5. Procter S., Illingworth J. and Elms A.J., The recognition of handwritten digit strings of unknown length using hidden Markov models, *Proceedings of the 14th I CPR* (1998), pp. 1515-1517.
6. Rabiner L. and Juang B.H., Fundamentals of Speech Recognition. Prentice Hall, New Jersey (1993).
7. Simoncini L. and Kovacs V., A system for Reading USA Census 90 Handwritten Fields, *Proceedings of the 3rd ICDAR* (1995), pp. 86-91.
8. Suen C.Y., Liu K. and Strathy N.W., Sorting and Recognizing Cheques and Financial Documents, *Proceedings of the 3rd IAPR - Workshop on Document Analysis Systems* (1998), pp. 1-18.