Word Slant Estimation using Non-Horizontal Character Parts and Core-Region Information

A. Papandreou and B. Gatos
Computational Intelligence Laboratory,
Institute of Informatics and Telecommunications,
National Research Center "Demokritos",
153 10 Athens, Greece
{alexpap,bgat}@iit.demokritos.gr

Abstract — **In this paper, we propose a new technique for estimating the word slant that is based on analyzing the non-horizontal parts of the characters. We calculate the orientation, the height of the bounding box that includes all non-horizontal parts as well as their location related to the word. We estimate the non-horizontal parts orientation weighted according to the height of the corresponding bounding box. An additional weight is applied if the fragment is outside of the core-region of the word, which indicates that this fragment is probably one of the strokes that should be, by definition, vertical to text orientation. Extensive experimental results prove the efficiency of the proposed method.**

Keywords - word slant estimation; core-region information; document image pre-processing;

I. INTRODUCTION

A crucial preprocessing technique is the removal of the unwanted slant variations. By the term “character slant” we refer to the angle in degrees clockwise from vertical at which the characters are drawn. Character slant estimation can be very helpful in handwritten text processing while it is also a salient feature of handwriting and it is considered to be an important writer-specific feature. Though it is controversial if correcting the slant will help in writer identification, estimating it definitely provides additional information.

There are several techniques used in order to estimate slant so at a next step we are able to correct it most of the times uniformly by shear transform. The average slant angle can be estimated by averaging angles of near-vertical strokes[1]-[2], by analyzing projection histograms [3]-[4] or by using statistics of chain code contours [5]-[7].

Referring to near vertical stroke techniques, according to Bozinovich and Srihari [1], for a given word all horizontal lines which contain at least one run of length greater than a parameter M (depending on the width of the letters) are removed. Additionally, all horizontal strips of small height are also removed. By deleting these horizontal lines, only orthogonal window parts of the image remain in the text. For each letter the parts that remain are those that have relatively small horizontal slant. For each of these parts, the angle between the line indicated by the centers of gravity for its upper and lower halves of the page is estimated and their mean value is the overall text’s character slant. In the approach of Kim and Govindaraju [2], vertical and near-vertical lines are extracted by tracing chain code components using a pair of one dimensional filters. Coordinates of the start and end points of each vertical line extracted provide the slant angle. Global slant angle is the average of all the angles of the lines, weighted by their length direction since the longer line gives more accurate angle than the shorter one.

Analyzing projection histograms, Vinciarelli and Luettin [3] have proposed a deslanting technique based on the hypothesis that the word has no slant when the number of columns containing a continuous stroke is maximum. On the other hand, Kavallieratou et al. [4] have proposed a slant removal algorithm based on the use of the vertical projection profile of word images and the Wigner-Ville distribution (WVD). The word image is artificially slanted and for each of the extracted word images, the vertical histogram as well as the WVD of these histograms is calculated. The curve of maximum intensity of the WVDs corresponds to the histogram with the most intense alternations and as a result to the dominant words slant.

Techniques based on statistics of chain code contours can estimate average slant of a handwritten word by using the 4-directional chain code histogram of border pixels according to Kimura et al. [5]. This method although tends to underestimate the slant when its absolute value is close or greater than 45^0. To solve this problem, Ding et al. [6] proposed a slant detection method using an 8-directional chain code. Chain code methods of 12 and 16 directions are also examined in [6] but the experimental results show that these methods tend to overestimate the slant.

Additionally a slant estimation method for handwritten characters by means of Zernike moments has been proposed by Ballesteros et al. [8] which is based on the average inclination of the Zernike reconstructed images for low moments. It is claimed that this method improves the slant estimation accuracy in comparison with the chain code based methods. Finally Nagendraprasad et el. [9] proposed a method for slant correction in handwritten numerals which performs a binary search on the angular slant values and is based on the observation that if a numeral is transformed through a series of slanted positions it attains its minimum width when it is near-vertical.

978-0-7695-4661-2/12 $26.00 © 2012 IEEE
DOI 10.1109/DAS.2012.97

Our technique of estimating the slant is motivated by [1] and depends on the orientation of the non-horizontal parts of the characters, the height of their bounding box and their location related to the word. We estimate the boxes that include all the non-horizontal fragments of the word, we estimate their orientation and we weight it according to the height of the box. An additional weight is applied if the fragment is outside of the core-region of the word, which indicates that this fragment is probably one of the strokes that should be, by definition, vertical to text orientation. In that way, we calculate the dominant slant of the word and we get far better results compared to other state of the art algorithms.

II. Proposed Methodology

The proposed world slant estimation method consists of three steps. Firstly, we find the core-region of the word with an Otsu based thresholding technique [2]. Afterwards, we remove the near-horizontal parts of the word by extracting the corresponding line that includes them and we put the remaining parts (fragments) into boxes (as in [1]). Finally we estimate the average slant by measuring the orientation of the remaining fragments averaging them weighted depending on the height of the box that bounds them and their position related to the word's core.

The first step of our method is to find the core-region of the word using the method presented in [2]. The core-region of the word includes the main parts of the characters while it excludes the parts that are lower than the lower baseline or higher than the upper baseline. Most of the fragments that are outside the core are strokes of the characters that should, by definition, be vertical to the word's direction when the word is properly deslanted (see Fig.1).

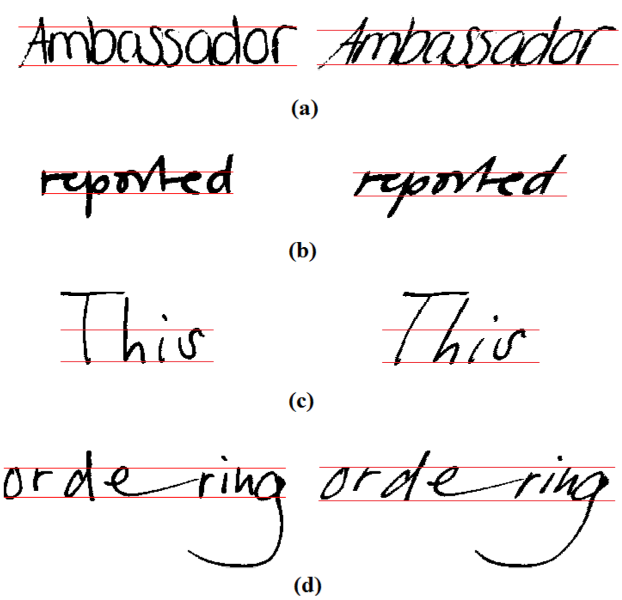

Figure 1. Example of the core-region of handwritten words with and without slant.

In the next step we exclude all the horizontal lines which contain at least one near-horizontal stroke by detecting black runs of length that exceed a parameter M defined as follows.

$$M = 3 \times dBR \tag{1}$$

when dBR is the dominant black run frequency and imprints the width of the letters.

Additionally, we remove all the horizontal strips of small height that might have appeared between the zones that we excluded in order to smoothen the area [1]. By deleting these horizontal lines, only orthogonal window parts of the picture remain in the text (see Fig.2).

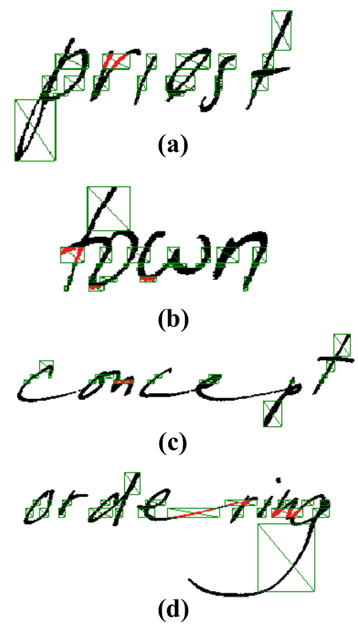

Figure 2. Boxes that bound the remaining fragments. Highlighted with red colour the fragments with misleading orientation.

In the last step we calculate the average slant using the information that we gathered in the first two steps. As we observe in Fig.1a,b most of the strokes outside the core-region of a deslanted word are near-vertical strokes so their orientation should have greater contribution in the total average slant when they are detected in a slanted word. Furthermore, most of the non-vertical fragments of the strokes outside the core-region are horizontal or near-horizontal (see Fig. 1c,d) so they have already been extracted in the previous step of our algorithm. In that way misleading fragments have not significant contribution. Taking that into consideration, we weight the remaining fragments depending on whether their bonding box has its upper or lower limits higher or lower from the upper or lower baseline respectively. For the average slant calculation the fragment's orientation is weighted by c_i which is defined as follows.

$$c_i = \begin{cases} 2 & \text{if} \quad ul_i < UL_i OR\, dl_i > DL_i \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

$$\forall\ i \in (0, b]$$

where ul_i and dl_i are the vertical distance of each box's upper and lower limits respectively from the top of the document, $UL_i$and DL_i are the vertical distance of the core-region's upper and lower baseline respectively from the top of the document and b is the number of boxes.

Here we should point out that the better the core-region is estimated the bigger its contribution to the improvement of the results of slant estimation is. It is true that in this way the improvement of our method depends on the core-region estimation but on the other hand even if the core-region is miscalculated, we would have a result as if there was no core-region estimation (see Fig. 3a,b) (all the fragments are weighted so it is as if no fragment was), or as if there was partial core-region detection (see Fig. 3c,d) resulting in partial improvement from Bozinovich method [1] (only part of the misguiding fragments will be weighted as much as the correct ones). It is highly unlike that a core-region detection algorithm will fail that much to its estimations that will include the vertical strokes that are above the upper line or lower than the baseline (since it is the sparsest area), and at the same time exclude the authentic core-region strokes, (which are in the thickest area), something that would be misleading. As much as words in capitals is concerned (see Fig. 3e,f), since the core-region includes the whole word all the boxes will be weighted uniformly.

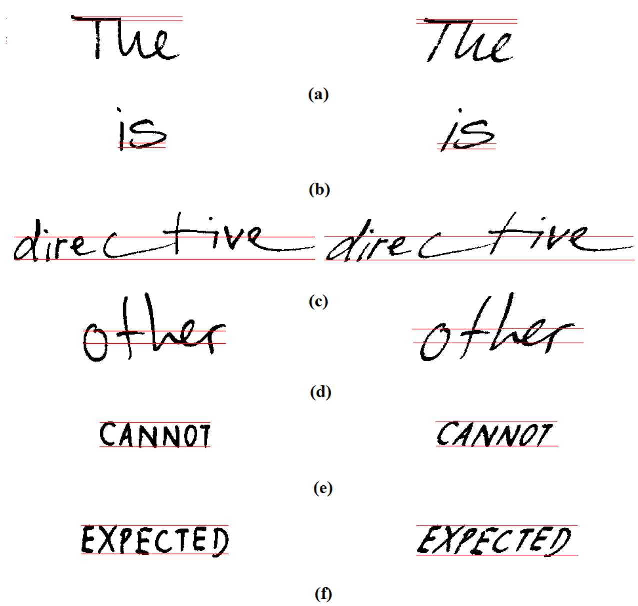

Figure 3. Example of false estimations of the core-region of handwritten words (a-d) and words in capital letters (e,f) from IAM Database [10]

Moreover the proposed methodology takes into account the fact that larger fragments should have larger impact on the outcome of the algorithm. We also observed that the strokes that by definition should be oriented vertically and their orientation points the slant are longer and bounded in bigger boxes compared to the rest of the fragments. On the other hand the strokes that are opposite-than-the-slant-oriented and misguide our algorithm are smaller (see Fig.2).

We also use the height of the bounding box as an extra weight, h_i, in order to make our technique more robust in greater angle scales. h_i is defined as follows:

$$h_i = 0.3 \times (dl_i - ul_i) \quad \forall\, \mathrm{i} \in (0, b] \tag{3}$$

In order to estimate the average slant of a word S_{AV} we combine both weights and take advantage of all the information that we gain from the core-region in the first step and the height of the boxes (see Fig.4) as follows:

$$S_{AV} = \frac{\sum_{i=1}^{\mathrm{b}} \{\mathrm{s_i} \times \mathrm{h_i} \times \mathrm{c_i}\}}{\sum_{i=1}^{\mathrm{b}} \{\mathrm{h_i} \times \mathrm{c_i}\}} \tag{4}$$

where $\mathrm{s_i}$ is the orientation of the fragment in the *i*-th box calculated by the inclination of the line connecting the centers of gravity for the box's upper and lower halves.

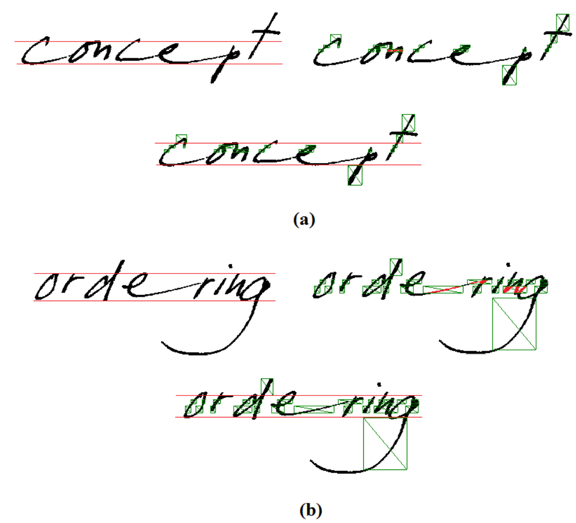

Figure 4. Combination of the information of the first two steps.

As we observe in Fig.4 the boxes that bound the largest fragments and are outside the core region contain the strokes that by definition should be vertical to the word's orientation. Our algorithm considers these kind of strokes as dominants.

III. Experimental Results

In Fig.5 we are demonstrating the performance of our technique against the classic algorithm of Bozinovich [1] while in Fig. 6 we have some more examples with results of our method in words with great slant. As it can be observed our approach outperforms method of [1].

In order to further test the proposed methodology we binarized (based on [11]) 540 words (from 9 different documents) from the IAM database [10] that seemed to have almost zero slant, 126 of which (from 3 different documents) were written fully in capital letters. We considered them as a ground-truth and we slanted them using the shear transform in 19 different angles from -45° to 45° with 5° step. In that way, we formed a test-set of 10260 words in which we applied 3 different algorithms: the classic Bozinovich algorithm [1], an enhanced Bozinovich algorithm that uses as weight the height of each box and the proposed methodology. We should point out here that there are lots of words in the dataset which consist of only few letters (1 to 3 letters) and numbers.

Furthermore, we used 159 printed handwritten-like words that we obtained from a list [12] using Bradley Hand ITC fonts, 10 of which were written fully in capital letters. We took them as a ground-truth and we slanted them using the shear transform in 19 different angles from -45° to 45° with 5° step. In that way we formed a test-set of 3021 words in which we applied the same 3 algorithms. In that way we can be more positive about the ground-truth that we use and we also test the algorithms in uniformly slanted words. In either way both of our test-sets provide all kind of advantages and disadvantages, same for all three algorithms so we can be positive about our results.

A. Full Test-Set

The results presented in Table I are from 540 handwritten words and 10260 different estimations for each of the algorithms. The metrics that are presented are the average error deviation in degrees and the percentage of the exactly correct estimation of the algorithms according to the ground-truth.

TABLE I. RESULTS FROM HANDWRITTEN WORDS

Different Slant Estimation Techniques	***Error Deviation in Degrees***	***Percent of Correct Estimation***
Bozinovich [1]	7,48	8,53%
Bozinovich + Weighted Boxes	**6,33**	**11,20%**
Proposed Methodology	**6,00**	**12,07%**

It can be observed that our method outperforms the classic slant estimation method [1], and we also clearly demonstrate the contribution of each of the weights that we use in our method. This is done by comparing the "Bozinovich + Weighted Boxes" method that uses as weights only the height of each box and our methodology that also integrates information about the location of its fragment.

Furthermore the results presented in Table II are from 159 printed handwritten-like words and 3021 different estimations for each of the algorithms.

TABLE II. RESULTS FROM PRINTED HANDWRITTEN-LIKE WORDS

Different Slant Estimation Techniques	***Error Deviation in Degrees***	***Percent of Correct Estimation***
Bozinovich [1]	9,62	6,75%
Bozinovich + Weighted Boxes	**9,34**	**7,31%**
Proposed Methodology	**8,36**	**8,17%**

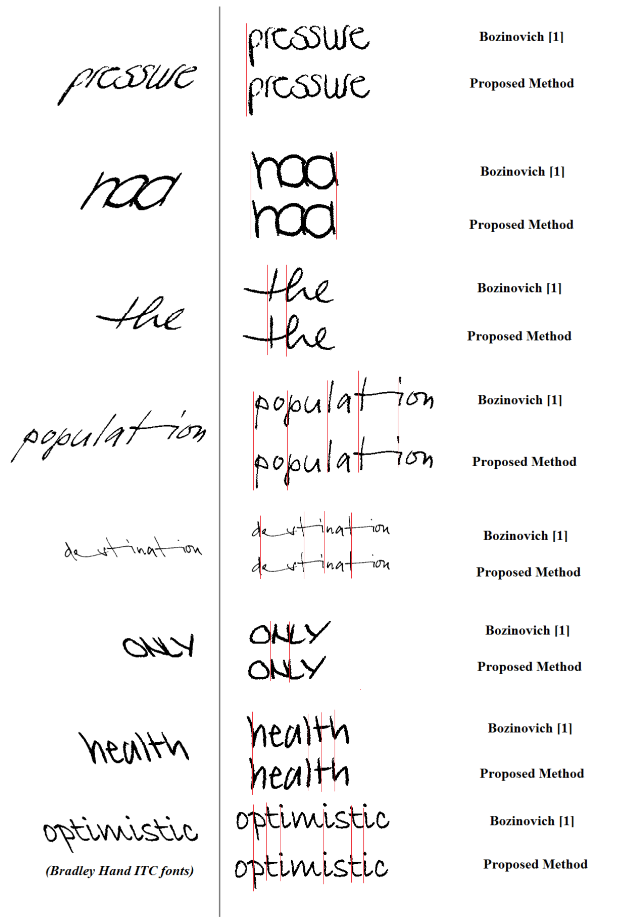

Figure 5. Examples in order to compare the proposed and the Bozinovich [1] method.

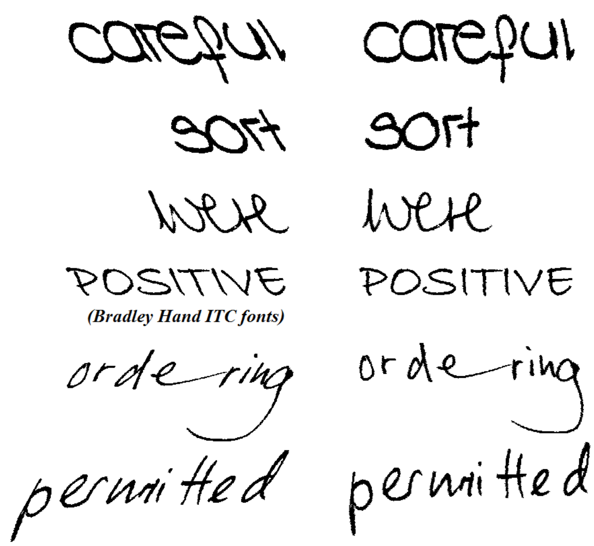

Figure 6. Deslant results using the proposed method.

As it can be observed in the results of the printed handwriten -like words our algorithm performs better than [1] while the contribution of each inovative weight is demonstrated.

B. *Test in words with the core-region corectly estimated*

In order to detect the dependency of our method from the core-region estimation algorithm we selected the 352 word-images in which the core-region detection algorithm was accurate in most of the slant angles and made a test supposing that we have a very good core estimation.

The results presented in Table III are from 352 handwritten words and 6688 different estimations for each of the algorithms.

TABLE III. RESULTS FROM HANDWRITTEN WORDS WITH CORRECT ESTIMATED CORE-REGION

Different Slant Estimation Techniques	***Error Deviation in Degrees***	***Percent of Correct Estimation***
Bozinovich [1]	5,73	10,59%
Bozinovich + Weighted Boxes	**4,43**	**13,89%**
Proposed Methodology	**4,08**	**15,61%**

As it is demonstrated when the core-region is correctly estimated our algorithm is improving its results and it becomes even better than the other algorithms.

We performed the same test with printed handwritten-like words and the results presented in Table IV are from 58 words and 1102 different estimations for each of the algorithms.

TABLE IV. RESULTS FROM PRINTED HANDWRITTEN-LIKE WORDS WITH CORRECT ESTIMATED CORE-REGION

Different Slant Estimation Techniques	***Error Deviation in Degrees***	***Percent of Correct Estimation***
Bozinovich [1]	6,66	10,07%
Bozinovich + Weighted Boxes	**5,65**	**11,70%**
Proposed Methodology	**5,10**	**12,89%**

As we can observe also in that case our algorithm behaves better when the core-region is correctly estimated and in all cases it behaves better than the state of the art algorithm [1]. Also, we can see that the average error deviation is again in general, for all 3 algorithms, higher than it was in the handwritten words and that is due to the nature of the specific fonts.

IV. CONCLUSION

In this paper we proposed a word slant estimation algorithm that was motivated by Bozinovich method [1]. After extensive testing it is proved to outperform the state of the art method [1]. We made two rational assumptions and we weighted the orientation of each non - horizontal detected fragment accordingly. The height of the box that bounds each fragment is crucial information and should be taken into consideration. While also the information whether the fragment is part of the core-region or not is crucial too. The strokes outside the core-region of the word are mostly near-vertical strokes and by definition should by vertical to the word's orientation. The contribution of each innovative element is demonstrated through the experiments in numbers but also by examples in Fig.5 and 6. Through the experiments we demonstrated also that our method depends on the core-region estimation algorithm but in any case our technique proves to be more efficient.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n^o 215064 (project IMPACT).

REFERENCES

[1] A. Bozinovich, A.Srihari, "Off-Line Cursive Script Word Recognition", *Transactions on Pattern Analysis and Machine Intelligence*, Vol II, No 1, pp. 69-82, 1989.

[2] G. Kim, V. Govindaraju, "A Lexicon Driven Approach to Handwritten Word Recognition for Real-Time Applications", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 19, Issue 4, pp. 366-379, 1997.

[3] A. Vinciarelli, J. Juergen, "A new normalization technique for cursive handwritten words", Pattern Recognition Letters, Vol. 22, Num. 9, pp. 1043-1050, 2001.

[4] E. Kavallieratou, N. Fakotakis, G. Kokkinakis, "Slant estimation algorithm for OCR systems", *Pattern Recognition*, Vol. 34, pp. 2515-2522, 2001.

[5] F.Kimura, M.Shridhar, Z.Chen, "Improvements of a Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words", *2nd International Conference on Document Analysis and Recognition (ICDAR 1993)*, pp. 18-22, Oct. 1993.

[6] Y. Ding , F. Kimura , Y. Miyake, "Slant estimation for handwritten words by directionally refined chain code", *7th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2000)*, pp. 53-62, 2000.

[7] Y. Ding, W. Ohyama, F. Kimura, M.Shridhar, "Local Slant Estimation for Handwritten English Words", *9th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2004)*, pp. 328-333, 2004.

[8] J. Ballesteros, C.M. Travieso, J.B. Alonso, M.A. Ferrer, "Slant estimation of handwritten characters by means of Zernike moments", *Electronics Letters*, Vol. 41, Issue 20, pp. 1110-1112, 2005.

[9] M.V. Nagendraprasad, A. Gupta, V. Feliberti, "Correction of Slant in Handwritten Numerals for Automated Document Processing", *Engineering Applications of Artificial Intelligence*, Vol. 8, Num. 4, pp. 469-472, 1995.

[10] www.iam.unibe.ch/fki/databases/iam-handwriting-database

[11] B. Gatos, I. Pratikakis and S. J. Perantonis, "Adaptive Degraded Document Image Binarization", Pattern Recognition, Vol. 39, pp. 317-327, 2006.

[12] http://www.psychpage.com/ learning/library/assess/feelings.html