

A Coarse to Fine Skew Estimation Technique for Handwritten Words

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— The estimation and correction of handwritten word skew is a difficult and challenging task since it has to be independent of the variations due to handwriting style and writing conditions. In this paper, a coarse-to-fine technique that integrates core-region information is presented. At first, a rough estimation and correction of the skew is accomplished by cutting vertically the word in two overlapping parts, detecting the center of mass in each part and calculating the inclination of the line that connects the two centers of mass. Afterwards, the core-region of the word is detected, the word is cut again in two overlapping parts and the centers of mass are calculated disregarding all the information outside the core-region (ascenders and descenders). The inclination of the line that connects the updated centers of mass corresponds to a finer estimation of the word skew. After correcting the detected skew the last step of core-region detection and skew correction is repeated iteratively in order to reach a finer word skew estimation that will contribute to a successful handwritten word recognition system. Extensive testing based on various test-sets has demonstrated that the proposed method outperforms the state-of-the-art algorithms concerning word skew estimation while it is more robust in variations of the writing style.

Keywords — *Handwritten Words; Skew Estimation; Normalization*

I. INTRODUCTION

One of the most important steps required for handwritten word recognition is that of detecting and correcting the word skew. Word skew is defined as the deviation of the baseline of the word from the horizontal direction. Some degree of word skew is unavoidable and it is introduced by handwriting styles of different writers or different writing conditions. The task of skew correction relates to the normalization of the handwritten word in terms of skew in order to reduce the word shape variability. By normalizing each handwritten word, the subsequent tasks of feature extraction and recognition are significantly facilitated. Previous research has indicated that the effect of word skew highly impact on the success of a handwriting recognition system [1].

In the literature, several methods have been proposed for handwritten word skew. Madhvanath et al. [2] have used the image contour to detect the word's minima and determine the baseline as the regression line through those minima. The inclination of the regression line is regarded to be the word skew. Morita et al. [3] have proposed a method based on mathematical morphology to obtain a pseudo-convex hull image. At a next step, the minima are detected on the pseudo-convex image, a reference line is fit through those points and the inclination of this line is considered to be the word skew.

The primary challenge in these two methods is the rejection of spurious minima. Furthermore, the regression-based methods do not work well on words of short length because of the lack of a sufficient number of minima points. Other approaches for handwritten word skew detection are based on density distribution. In [4], several histograms are computed for different vertical projections. Then, the entropy is calculated for each one of them. The histogram with the lowest entropy determines the word skew angle. In [5], Kavallieratou et al. calculate the Wigner-Ville distribution for several horizontal projection histograms. The word skew angle is selected by the Wigner-Ville distribution having the maximal intensity. The main problem for these projection based methods is the high computational cost since an image has to be rotated for each angle. Jian-xiong Dong et al. [1] have proposed to maximize a global measure which is defined by the Radon transform and calculate its gradient in order to estimate the word skew. Furthermore, Blumenstein et al. [6] identify the skew by detecting the center of mass in each half (right and left) of a word image. The skew angle is estimated by hypothesizing a line between the two centers and by measuring its angle with the x-axis.

In this paper we present a novel coarse-to-fine handwritten word skew estimation method that is fast, has low computational cost, is accurate and doesn't depend on the minima of the word. In that way, the proposed technique provides an accurate estimation of the word skew, even for words with spurious minima and small length. It is based on the detection of the core-region and the calculation of the centers of mass of the right and left part of the word image. At first, a rough estimation and correction of the skew is made by cutting vertically the word in two equal overlapping parts, detecting the corresponding center of mass in each part and calculating the inclination of the line that connects them. This step is essential in order to have accurate core-region detection. Afterwards, we take advantage of the information inside the core-region in order to iteratively make a finer estimation of the skew.

II. PROPOSED METHODOLOGY

The proposed coarse-to-fine method is based on two steps as it is demonstrated in the flowchart of Fig.1. At first, (a) the word image is divided vertically into two equal overlapping parts, (b) the corresponding center of mass of the foreground pixels is detected in each part and (c) the inclination of the line that connects the two centers of mass is calculated. This leads to a first rough estimation of the skew angle and the corresponding correction is applied in the handwritten word image. Afterwards, at the next step, the core-region of the word

is detected and the word is divided again vertically in two equal overlapping parts. The corresponding centers of mass are calculated, though this time only the information inside the core-region is taken under consideration. The inclination of the line that connects the updated centers of mass is calculated and a new finer skew correction is made in the word image. The second step is repeated iteratively till the finest skew estimation is accomplished.

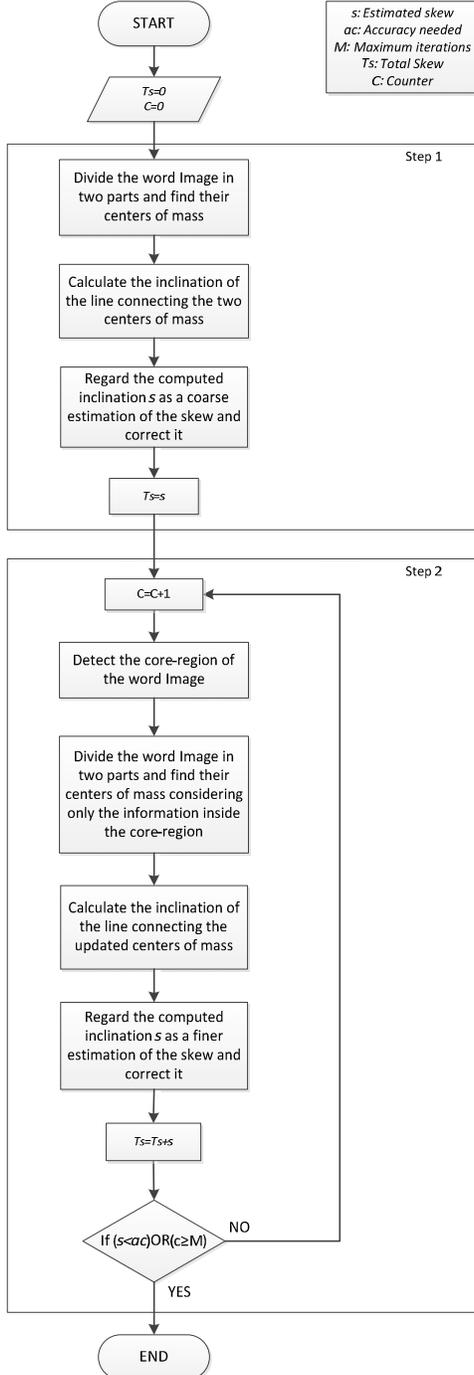


Fig. 1. Flowchart of the proposed coarse-to-fine methodology.

A. Step 1: Coarse Estimation of the Word Skew

Let $I(x, y)$ be the binary handwritten word image $I_x \times I_y$ having 1s for foreground pixels and 0s for background pixels (see Fig.2).

$$I(x, y) = \begin{cases} 1 & 0 \leq x < I_x \\ 0 & 0 \leq y < I_y \end{cases} \quad (1)$$

In the first step, the word image is divided in two equal overlapping parts. In order to define those parts the handwritten word image is divided horizontally in three equal stripes as it is demonstrated in Fig.2. The first two thirds of the image compose the left part with dimensions $I_x^l \times I_y$ and the rightmost two thirds form the right part with dimensions $I_x^r \times I_y$.

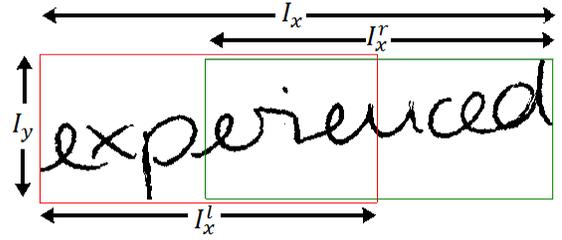


Fig. 2. The word is divided in two equal horizontal overlapping parts with length I_x^l and I_x^r respectively for the left and the right part.

Afterwards, the corresponding centers of mass (x_{cm}^l, y_{cm}^l) and (x_{cm}^r, y_{cm}^r) of the overlapping parts are calculated according to the following equations:

$$x_{cm}^l = \frac{\sum_{x=0}^{x=2I_x/3} \sum_{y=0}^{y=I_y} (x - I_x/6) * f(x, y)}{\sum_{x=0}^{x=2I_x/3} \sum_{y=0}^{y=I_y} f(x, y)} \quad (2)$$

$$y_{cm}^l = \frac{\sum_{x=0}^{x=2I_x/3} \sum_{y=0}^{y=I_y} y * f(x, y)}{\sum_{x=0}^{x=2I_x/3} \sum_{y=0}^{y=I_y} f(x, y)} \quad (3)$$

$$x_{cm}^r = \frac{\sum_{x=I_x/3}^{x=I_x} \sum_{y=0}^{y=I_y} (x + I_x/6) * f(x, y)}{\sum_{x=I_x/3}^{x=I_x} \sum_{y=0}^{y=I_y} f(x, y)} \quad (4)$$

$$y_{cm}^r = \frac{\sum_{x=I_x/3}^{x=I_x} \sum_{y=0}^{y=I_y} y * f(x, y)}{\sum_{x=I_x/3}^{x=I_x} \sum_{y=0}^{y=I_y} f(x, y)} \quad (5)$$

Due to the fact that the two parts are overlapping horizontally, the calculation of x_{cm}^l and x_{cm}^r is done as if the word image's width was the sum of the width of the two parts.

So, the left center of mass shifts by half the overlapping distance to the left and the right center of mass shifts by the other half of the overlapping distance to the right.

In order to have a coarse estimation of the skew angle of the word image, the inclination of the line connecting the two reference points (see Fig. 3) is calculated using eq. 6.

$$s = \tan^{-1} \left(\frac{y_{cm}^r - y_{cm}^l}{x_{cm}^r - x_{cm}^l} \right) \quad (6)$$

The word image is then rotated to $-s$ in order to correct the course estimated word skew.

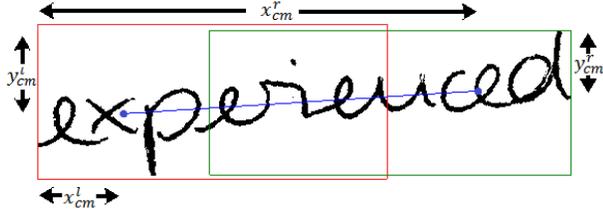


Fig. 3. The line connecting the two centers of mass. Its inclination corresponds to the coarse estimation of the word image skew angle.

Unlike the method presented in [6], in the proposed method, the word image is not divided vertically in the middle but it is separated in two equal overlapping parts. In that way, the overlapping parts are providing continuity in the information of the left and the right side. Also, the proposed algorithm, due to the overlapping parts that introduces, calculates the centers of mass in a suitable way. With the use of overlapping parts and the corresponding centers of mass, better results are achieved even for words of few letters and limited length, where most of the state-of-the-art methods fail.

B. Step 2: Iterative Finer Estimations of the Word Skew

At the next step, the overlapping parts are calculated again for the corrected word image and the core-region of the word is detected. In that way, the main part of the word image (inside the core-region) is distinguished from the ascenders and descenders (see Fig.4).

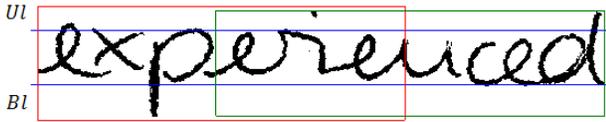


Fig. 4. The overlapping parts are calculated for the corrected image and the core-region of the handwritten word is detected.

In order to detect the core-region of the word image, the algorithm of reinforced projection profiles described in [7] is used. In that way the upper and lower baselines of the word are detected.

At this point the centers of mass of the two overlapping parts are calculated again, though only the information inside the core-region is taken under consideration. The updated coordinates of the centers of mass (x_{cm}^l, y_{cm}^l) and (x_{cm}^r, y_{cm}^r) are given by eq. 7-10:

$$x_{cm}^l = \frac{\sum_{x=0}^{x=2l_x^l/3} \sum_{y=Ul}^{y=Bl} (x - l_x^l/6) * f(x, y)}{\sum_{x=0}^{x=2l_x^l/3} \sum_{y=Ul}^{y=Bl} f(x, y)} \quad (7)$$

$$y_{cm}^l = \frac{\sum_{x=0}^{x=2l_x^l/3} \sum_{y=Ul}^{y=Bl} y * f(x, y)}{\sum_{x=0}^{x=2l_x^l/3} \sum_{y=Ul}^{y=Bl} f(x, y)} \quad (8)$$

$$x_{cm}^r = \frac{\sum_{x=l_x^l/3}^{x=l_x^r} \sum_{y=Ul}^{y=Bl} (x + l_x^l/6) * f(x, y)}{\sum_{x=l_x^l/3}^{x=l_x^r} \sum_{y=Ul}^{y=Bl} f(x, y)} \quad (9)$$

$$y_{cm}^r = \frac{\sum_{x=l_x^l/3}^{x=l_x^r} \sum_{y=Ul}^{y=Bl} y * f(x, y)}{\sum_{x=l_x^l/3}^{x=l_x^r} \sum_{y=Ul}^{y=Bl} f(x, y)} \quad (10)$$

where l_x^l is the length of the corrected word image and Ul and Bl are the upper and lower baseline respectively, detected from the core-region detection algorithm.

After defining the updated centers of mass (see Fig. 5) the inclination of the line that connects them is calculated by eq. 11 and in that way a finer estimation of the word image skew angle is achieved.

$$s' = \tan^{-1} \left(\frac{y_{cm}^r - y_{cm}^l}{x_{cm}^r - x_{cm}^l} \right) \quad (11)$$

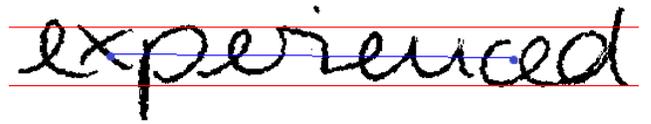


Fig. 5. The line that connects the updated centers of mass. Its inclination corresponds in a finer estimation of the word image skew angle.

After correcting the skew detected, the second step is repeated iteratively until a finest skew estimation is reached. When our last estimation of the word image skew angle is smaller than a defined accuracy ac , the algorithm stops and the total skew is defined as the sum of all previous calculated values as follows:

$$T_s = \sum_i s_i \quad (12)$$

where i is the number of skew angle estimations needed to achieve accuracy ac and s_i is the estimated word skew in the i -th calculation.

By excluding the information outside the core-region of the word and by updating the centers of mass considering only remaining information, a finer estimation of the skew angle is achieved. This is due to the fact that the ascenders and descenders of a word have a major contribution in the vertical average deviation of the foreground pixels and they affect erroneously the skew estimation.

Also, it should be pointed out that the proposed method is based on the core-region detection algorithm which performs better after each iteration, as the skew is minimized. Consequently after a few iterations the proposed method will reach to the finest skew estimation and the line connecting the centers of mass will be horizontal.

III. EXPERIMENTAL RESULTS

In order to test the proposed methodology two datasets were created. Dataset A was based on real data and dataset B on synthetic data, while both datasets are available upon request.

For the construction of dataset A, we used words obtained from 9 different documents (100 words) from the IAM database [8] that were visually checked to have no skew. Those 9 documents were picked so that the proposed method would be tested on various handwriting styles such as cursive handwriting, round and elongated letters, thick and thin strokes. All words were first binarized using the adaptive technique of [9]. At a next step, they were rotated in 11 different angles from -5° to 5° with 1° step. In that way we formed a dataset of 1100 real words for which the ground-truth was known.

In order to form dataset B, we obtained instances (100 words) from a random list of feeling words found in [10] and Bradley Hand ITC fonts was used to construct a synthetic printed handwritten-like dataset. Again, the word images were rotated in 11 different angles from -5° to 5° with 1° step and formed a dataset of 1100 synthetic words with known skew ground truth. Dataset B compared to dataset A contains words with more accurate ground truth while representative samples of datasets A and B together with the skew estimation results using the proposed methodology are given in Fig.6.

For the experiments conducted the accuracy demanded from the proposed algorithm was 0.1° while the maximum possible iterations were set to be 4. In order to make a comparison with current handwritten word skew estimation state-of-the-art techniques, we implemented (i) Blumenstein et al. [6] algorithm, (ii) Safabakhsh and Khadivi [11] minimization of the bounding box technique, (iii) the classical Projection Profiles of Postl et al. [12] and (iv) the Papandreou and Gatos [13] enhanced Projection Profile. Also, we enhanced Blumenstein et al. [6] algorithm with the proposed overlapping parts in order to measure the influence of the innovative elements that were introduced in the proposed technique. It should also be mentioned that neither regression algorithms nor Hough transform could perform successively with that little information of a word and therefore they were not implemented. For all these methods, the average word image skew error estimation was calculated in degrees. The results are presented in Table I.

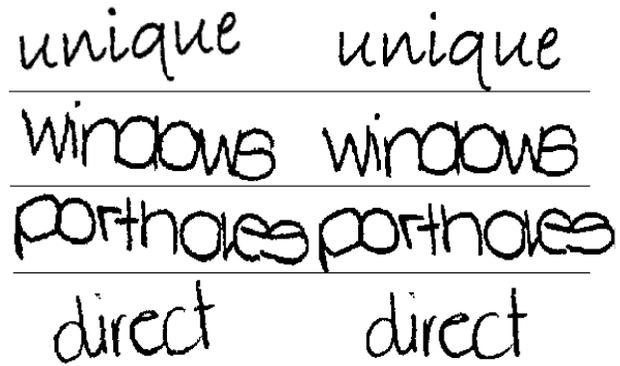


Fig. 6. Representative samples of datasets A and B with skew (left) and corrected with the proposed algorithm (right).

TABLE I. AVERAGE ERROR DEVIATION IN WORD IMAGE SKEW ESTIMATION

Different Word Skew Estimation Techniques	Error Deviation in Degrees		
	Dataset A	Dataset B	Dataset A&B
Blumenstein et al. [6]	0,993	0,666	0,829
Safabakhsh and Khadivi [11]	1,795	2,100	1,947
Postl et al. [12]	1,410	0,536	0,973
Papandreou and Gatos [13]	1,145	0,479	0,812
Blumenstein et al. [6] with overlapping parts	0,863	0,575	0,719
Proposed Technique	0,580	0,415	0,497

As it is demonstrated in Table I the proposed method outperforms the above mentioned state-of-the-art algorithms in both datasets A and B. Furthermore, it is easily observed that most of the algorithms perform better in dataset B, that has synthetic data, than in dataset A that has real words due to the lack of variations in the handwriting style. All the algorithms, except the one proposed, tested in real words have an average error deviation that almost exceeds 1° . We can state that the proposed algorithm can better handle the variations in the handwriting styles.

Also, it is profound from the results demonstrated in Table I that the overlapping parts along with the corresponding centers of mass that are described in this paper is a more efficient way to divide the word image than the one proposed in [6]. Furthermore, the contribution of the core-region information is shown from the performance of the proposed algorithm against the enhanced algorithm of Blumenstein et al. [6].

In Fig. 7 there is a representative example of the iterative results of the proposed method for a word image.

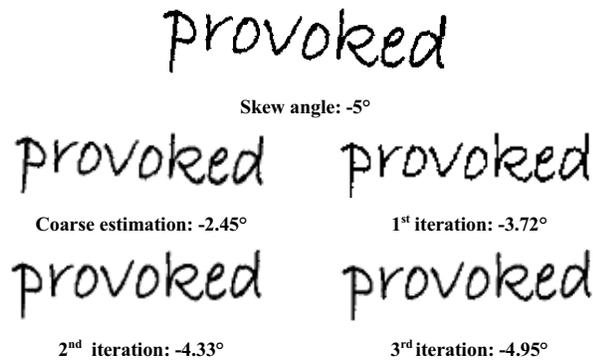


Fig. 7. Representative example of the iterative results of the proposed method for a word image

IV. CONCLUSION

In this paper, a novel skew estimation method for handwritten word images is presented. It's a coarse-to-fine technique based on two steps in which the word image is divided vertically into two equal overlapping parts, the corresponding center of mass of the foreground pixels is detected in each part and the inclination of the line that connects them is calculated. After this first rough estimation of the skew, the corresponding correction is applied in the handwritten word image and the core-region of the word is detected. The word is divided again in two equal overlapping parts and the centers of mass are updated taking under consideration only information inside the core-region. The inclination of the line that connects the updated centers of mass is calculated and a new finer skew correction is made in the word image. After correcting the detected skew, the second step is repeated iteratively till the finest skew estimation is done.

The proposed coarse-to-fine method uses a new way to divide the word image in overlapping parts which provide continuity in the information of the left and the right side and introduces the suitable way to calculate their centers of mass. Also, it integrates the core-region information of the word in the algorithm and in this way, the ascenders and descenders which have a major contribution in the vertical average deviation of the foreground pixels are excluded from the algorithm's calculations providing a finer estimation of the skew. Repeating the correction of the word image and the update of the centers of mass iteratively, consequently the proposed method will reach to the finest skew estimation.

Extensive testing based on various test-sets has demonstrated that the proposed method outperforms the state-of-the-art algorithms concerning word skew estimation. Moreover, it is demonstrated that this coarse-to-fine technique is more robust in variations of the writing style while the

contribution of each of the innovative elements proposed is experimentally proved. The proposed algorithm is not computationally expensive, it is more robust and accurate and it can perform well also with words of small length.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n^o 600707 - tranScriptorium.

REFERENCES

- [1] J. Dong, D. Ponson, A. Krzyżak and C. Y. Suen, "Cursive word skew/slant corrections based on Radon transform," Proc. of 8th International Conference on Document Analysis and Recognition, pp.478-483, 2005.
- [2] S. Madhvanath, G. Kim and V. Govindaraju, "Chaincode contour processing for handwritten word recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 9, pp. 928-932, 1999.
- [3] M. Morita, J. Facon, F. Bortolozzi, S. Ganes and R. Sabourin, "Mathematical morphology and weighted least squares to correct handwriting baseline skew," Proc. of 5th International Conference on Document Analysis and Recognition, pp. 430-433, 1999.
- [4] M. Cote, E. Lecolinet, M. Cheriet and C. Suen, "Automatic reading of cursive scripts using a reading model and perceptual concepts," International Journal on Document Analysis and Recognition, vol. 1, no. 1, pp. 3-17, 1998.
- [5] E. Kavallieratou, N. Fakotakis and G. Kokkinakis, "New algorithms for skewing correction and slant removal on word-level," Proc. IEEE 6th International Conference on Electronics, Circuits and Systems, pp. 1159-1162, 1999.
- [6] M. Blumenstein, C. K. Cheng and X. Y. Liu, "New Preprocessing Techniques for Handwritten Word Recognition", Proc. of 2nd IASTED International Conference on Visualization, Imaging and Image Processing, pp.332-336, 2002.
- [7] A. Papandreou and B. Gatos, "Slant estimation and core-region detection for handwritten Latin words", Pattern Recognition Letters, Available online 29 August 2012, ISSN 0167-8655, 10.1016/j.patrec.2012.08.005.
- [8] www.iam.unibe.ch/fki/databases/iam-handwriting-database, last accessed on 27 of June 2012
- [9] B. Gatos, I. Pratikakis and S. J. Perantonis, "Adaptive Degraded Document Image Binarization", Pattern Recognition, Vol. 39, pp. 317-327, 2006.
- [10] <http://www.psychpage.com/learning/library/assess/feelings.html>, last accessed on 27 of June 2012
- [11] R. Safabakhsh and S. Khadivi, "Document Skew Detection Using Minimum-Area Bounding Rectangle", Proc. of International Conference on Information Technology: Coding and Computing, pp. 253 - 258, 2000.
- [12] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents", Proc. of 8th International Conference on Pattern Recognition, pp. 687-689, 1986.
- [13] A. Papandreou and B. Gatos, "A Novel Skew Detection Technique Based on Vertical Projections", Proc. of 11th International Conference on Document Analysis and Recognition, pp. 384-388, 2011.