

Boise State University

ScholarWorks

---

Electrical and Computer Engineering Faculty  
Publications and Presentations

Department of Electrical and Computer  
Engineering

---

12-1-2019

## Normalization of Unconstrained Handwritten Words in Terms of Slope and Slant Correction

Suman Kumar Bera  
*Jadavpur University*

Akash Chakrabarti  
*Jadavpur University*

Sagnik Lahiri  
*Jadavpur University*

Elisa H. Barney Smith  
*Boise State University*

Ram Sarkar  
*Jadavpur University*

---

### Publication Information

Bera, Suman Kumar; Chakrabarti, Akash; Lahiri, Sagnik; Barney Smith, Elisa H.; and Sarkar, Ram. (2019). "Normalization of Unconstrained Handwritten Words in Terms of Slope and Slant Correction". *Pattern Recognition Letters*, 128, 488-495. <https://dx.doi.org/10.1016/j.patrec.2019.10.025>

© 2019, Elsevier. Licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 license. The final, definitive version of this document can be found online at *Pattern Recognition Letters*, doi: [10.1016/j.patrec.2019.10.025](https://dx.doi.org/10.1016/j.patrec.2019.10.025)

# *Normalization of unconstrained handwritten words in terms of Slope and Slant Correction*

Suman Kumar Bera <sup>1</sup>, Akash Chakrabarti <sup>1</sup>, Sagnik Lahiri <sup>1</sup>, Elisa H. Barney Smith <sup>2</sup> and Ram Sarkar <sup>1</sup>

<sup>1</sup> Computer Science and Engineering Department,  
Jadavpur University, Kolkata, India

<sup>2</sup> Electrical and Computer Engineering Department,  
Boise State University, Boise, USA

**Abstract**— In offline handwritten text slope (or skew) and slant are inevitably introduced, but to varying degrees depending on several factors, such as the writing style, speed and mood of the writers. Therefore slope and slant detection in offline handwritten text and their subsequent correction have become the critical pre-processing steps for document analysis and retrieval systems to neutralize the variability of writing styles and to improve the performance of word and character recognition systems. In this paper, we present new methods that use two novel core-region detection techniques to estimate both the slope and slant angles of offline handwritten word images. Also we prepare multilingual datasets comprised of both real and synthetic handwritten word images, along with ground truth information related to the slope and slant of each word, to address the lack of standard datasets for this research. These datasets of Bangla, Devanagari and English words are made publicly available. Extensive experimental results prove the efficacy of the proposed methods compared to contemporary state-of-the-art methods. Moreover, the methods are robust, efficient, and easily implementable.

**Keywords**—*slope; slant; core-region;*

## I. INTRODUCTION

As the volume of information worldwide has risen, digitalization of paper documents has become a priority for conversion of existing handwritten text documents into electronic ones for easy archiving and maintenance. In handwritten text slope and slant are inevitably introduced. Slope is the angle between the horizontal and the direction of the line along which the word is aligned. Slant is the angle between the vertical direction and the dominant direction of the vertical strokes of a word. Moreover, during the scanning process additional slope may be introduced due to physical misalignment of the scanned document on the scanning surface. This is also an issue if the document is camera captured. It may be noted that most character recognition algorithms based on the texture features are highly sensitive to the presence of slope and slant in the word images if they use character segmentation. Even holistic approaches which refrain from segmentation suffer from variability of character shape; their treatment of lexicon words as distinct pattern classes has traditionally limited their application to recognition scenarios. Even with small lexicons, the intertwining of successive words and touching lines may also impede the efficiency of the entire system. Therefore, detection as well as correction of slope and slant in cursive handwritten words are indispensable pre-processing steps which helps to reduce their effect on succeeding stages like feature extraction during character or word recognition and thus improves the accuracy of the systems. To solve this problem, many

researchers have proposed algorithms for slope and slant correction of word images, but most of these methods are very time consuming, and bear certain limitations such as working only for a particular script. To address this problem, in this paper we propose a novel algorithm that is not only efficient and fast, but also works well for word images written in different scripts.

## II. RELATED WORK

Slope and slant detection in document images can be performed either at the page, text-line or word level. In this paper we focus on the slope and slant estimation of offline handwritten text at the word level. The major approaches found in the literature can be classified into three broad categories of methods, those based on the Hough Transformation, Projection Profiles and evaluating the Core-region of the given word.

The Hough transformation [1] is one of the most popular techniques for correcting the slope angle of word images because of its robustness and simplicity, but it becomes computationally expensive when the size of the Hough space or number of pixels in a word image increase. Recently, Limonova et al.[2] use a Fast Hough Transform to determine the characters' slant angle in Russian passport recognition by analyzing the vertical strokes which are extracted with the help of the x-derivative of the text-line. A method called Progressive Probabilistic Hough Transform (PPHT) is proposed by Boudraa et al. [3], to identify the most prominent lines in a scanned document from which the global slope angle is estimated. Many methods as reported in [4–6] attempt to reduce the size of the Hough space, but still the computational complexity remains high.

The methods based on projection profile make use of projection features of the word image to determine the slope or slant angle [7–10]. The horizontal projection is applied to the rotated word image to find the slope angle and vertical projection is applied to sheared images to find the slant angle of the given word image. Kavallieratou et al. [8, 9] emphasize the Winger-Ville distribution for both horizontal and vertical projection profiles to determine the slope and slant angles respectively. Pastor et al. [10] focus on the observation that the column distribution of the vertical projection profile yields the maximum variance for a non-skewed word. Therefore, these methods are highly sensitive to sparse pixel distribution as well as the ascenders and/or descenders of the word images.

Methods belonging to the third category try to estimate the core-region of a given word, which is basically the area between the upper and lower baselines of a word image [11], and then determine a best line along that core-region. Accurate estimation of the core-region actually helps to determine the inclination of a word image, and thus the slope and slant angles, even more precisely. The best fit line along the core-region is approximated by linear regression [12] or maximum Eigen-vectors of the covariance matrix [13]. Bera et al. [14] propose a hybrid method where they first fit an ellipse over a word image, and then the major and minor axes are used to determine the slope and slant angles simultaneously. Most recently Kar et al. [15] try to optimize the frequency function over the horizontal and vertical projection profiles to determine the slope and slant angles of curvy word images.

Some other significant works also address this problem. Alaei et al. [16] use the piece-wise painting algorithm (PPA) to detect slope angle. Gatos et al. [17] estimate the minimum entropy of the sheared image to find the overall slant angle of a text word. Haji et al. [18] concentrate on splitting an entire word in two vertical slices and then joining their centers of gravity to estimate the slope angle of a text-word. Manish et al.[19] use the chain code method for slant estimation of handwritten words that overcomes the problems of underestimate and overestimate while using 4/12/16 directional methods. Most recently, Huang et al. [20] use a residual neural network (ResNet) to correct the skewedness of Chinese characters, but it is highly affected by an improper and inadequate training set.

This brief literature review reveals that the third category which uses the core-region is the most effective approach for estimating slope and slant angles of a given word. Hence, in the present work, we have emphasized the core-region based approach in designing the techniques for slope and slant correction of the handwritten word images.

### III. RESEARCH GAP

It is observed that some of the existing methods are script dependent, which means a method developed for Matra based scripts (e.g. Bangla or Devanagari) may not be applicable for non-Matra based scripts (e.g. Roman) or vice-versa. A Matra is defined as the distinctive horizontal line running along the top of the characters that links them to make an individual word. Furthermore, it may be noted that most of the slant estimation algorithms [8–11] are developed by conceptualizing words without any slope. But how could the handwritten words be perfectly horizontal along the x-axis when the word is written by human beings? This implies that these methods may fail to estimate accurate slant angles for such types of oblique words. Keeping this fact in mind, we have proposed a novel algorithm which first corrects any slope present in the word image, and then subsequently corrects its slant. Our algorithm has the potential to work on handwritten word images written in any script to find the obliquity of the words most efficiently.

The contributions of our methods can be summarized as follows:

1. Designing two new methods for the normalization of slope and slant angles of offline handwritten word images using a novel core-region detection technique for both skewed and non-skewed words.
2. Preparation of four datasets along with corresponding ground truth information - among these three are benchmark datasets having 1500 handwritten words each written in Bangla, Devanagari and Roman, and the remaining one is a synthetic dataset having 180,500 words taken from the IAM dataset [21]. These datasets are made publicly available for the research community [22].

This paper is organized into seven sections, of which the introduction part has already been discussed in section 1. Section 2 narrates the several existing works related to slope and slant correction and their relative merits and demerits. The present section 3 conveys the motivation behind this approach and a brief overview

of our contributions. In the following section 4, we present the proposed work detailing processes for the estimation and correction of skewed and slanted words. Section 5 describes the processes of data collection (isolated word samples) and preparation of Ground Truth (GT) information. Experimental results comprising the quantitative evaluation process and comparison with some state-of-the-art slope and slant estimation methods are presented in section 6. Finally, the epilogue including future plans is reported in section 7.

#### IV. PROPOSED METHODOLOGY

Our proposal consists of three main steps: pre-processing, slope correction and slant correction as shown in Figure 1. Each step further consists of several sub-steps which are detailed next. In the first step, the word image is converted to binary form and then is skeletonized to retain the best component shape as well as to reduce the computational complexity. The second step assesses the inclination of the word image. Finding the inclination of a skewed word from its core-region is always a challenging task, even though it is quite simple to look for the top and bottom profiles of that word. Our algorithm increases the density of the busy zone to make a denser word. Finally, slope corrected word is checked to determine whether it is vertically aligned or not. To do that, the algorithm operates only on the longest strokes which are more significant for slant detection. This simple but effective adjustment reduces random shearing of word images as well as fragmenting of strokes.

### A. Pre-Processing

The first step of the proposed method is to binarize the input word image. The RGB image  $I_r$  is converted to gray scale, then binarized using the well-known Otsu's method [23] to produce

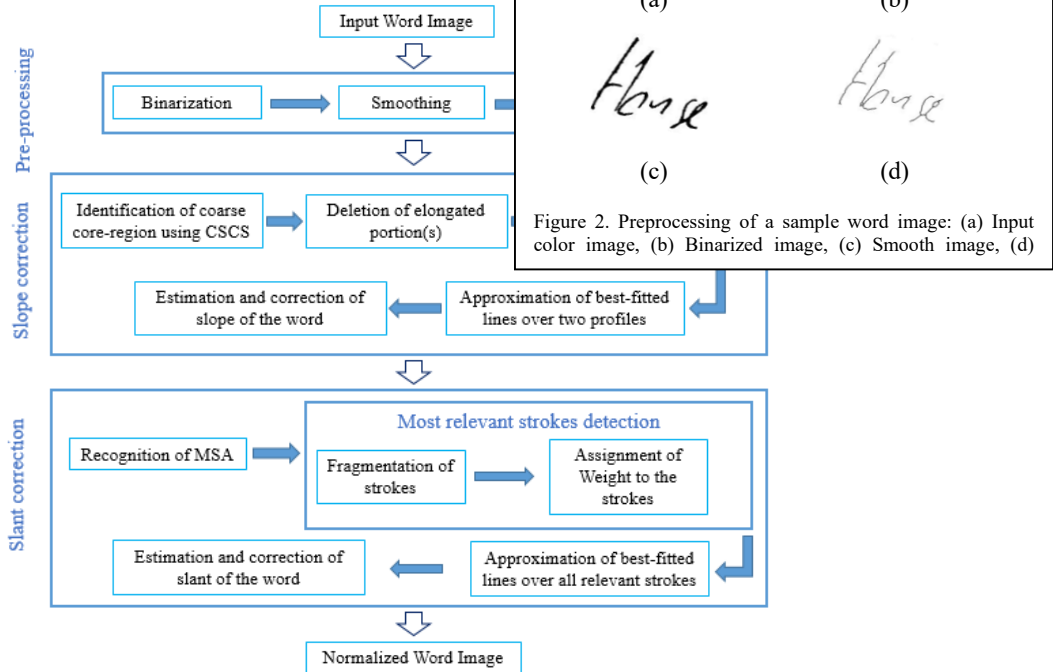


Figure 2. Preprocessing of a sample word image: (a) Input color image, (b) Binarized image, (c) Smooth image, (d)

the binary word image  $I_b$  having 1s for white (background) pixels and 0s for black (foreground) pixels. The shrink and swell filters are then applied to smooth the contour of each component. Finally, the smoothed image is cropped to its minimum bounding box and skeletonized (let 1s) to reduce computational complexity as well as to get a more concise representation prior to further applications such as feature estimation and stroke detection. The skeleton is a one-pixel thick representation showing the centerlines of the text. The outcomes of these preprocessing steps are shown in Figure 2.

### B. Slope Correction

Our slope correction method uses a two-pass approach, consisting of coarse core-region detection and statistical modelling of the best fitting lines on two end profiles of the word image. The horizontal histogram profile which is often used for core-region detection may lead to loss of relevant information. Instead, we gradually increase the density of the busy zones in a word image to identify the core-region more accurately.

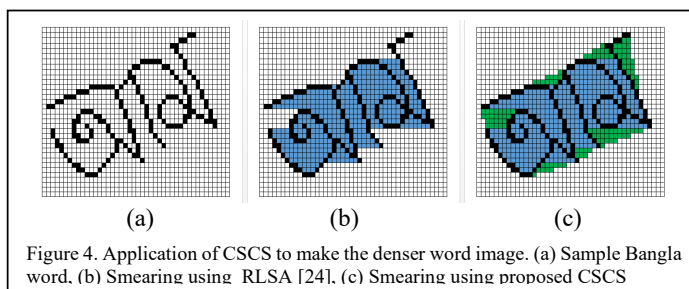
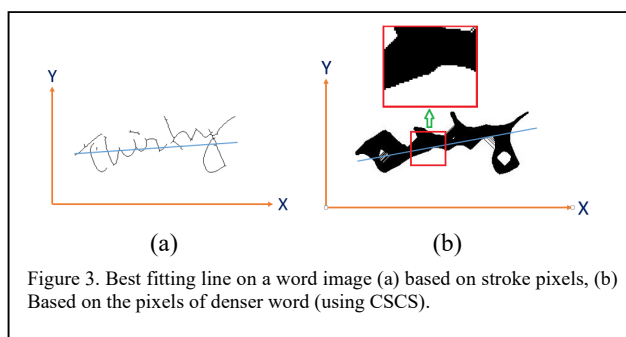
### i. Core region detection

As we look into the orientation of a word by fitting a line using linear regression, we also need to consider the presence of ascenders and descenders, which are present in most words. The sparse areas outside the core-region may mislead the regression procedure of slope detection when the text pixels appear intrinsically nonlinear due to the non-uniform

distribution of ascenders and descenders. Hence, the detection of the core-region and removal of these elongated parts within word images are the crucial stages of slope detection. The core-region detection algorithms based on the horizontal histogram tend to remove most of the text pixels appearing on the top and bottom portions of a word image and naturally result in the loss of relevant data. To address this problem, we approximate the coarse core-region by smearing the components to their neighboring components. We have introduced a simple smearing algorithm, called Component Smearing with Concentric Squares (CSCS), to increase the density of a skeletonized word's busy zone.

As we look into the orientation of word by fitting a line using linear regression, we also need to consider the presence of ascenders and descenders, which are obvious parts in most of words. It is because, the sparse areas outside the core region may mislead the regression procedure of slope detection when the text pixels appear intrinsically nonlinear due to the non-uniform distribution of ascenders and descenders. Hence, the detection of core region and removing these elongated parts within word images appear as a crucial stage of slope detection. While considering skewed words, the core region detection algorithms based on the horizontal histogram, tends to remove most of the text pixels appearing on the top and bottom portions of word image and naturally results in the loss of relevant data. To address this problem, we may approximate the coarse core region by smearing the components to its neighboring components.

In general, the busy zone corresponds to the core-region of the word image. Here, concentric squares are hypothetically fitted surrounding each text pixel by increasing  $S_h$  pixels on its both sides.  $S_h$  varies from 1 to a threshold  $Th_1$ . All the text pixels within the squares are then joined to that pixel by a black run using the Digital Differential Analyzer (DDA) algorithm [25]. Particularly on skewed words CSCS produces better statistical estimates about the orientation of the word because it avoids the staircasing problem on the contour of the strokes commonly provided by conventional smearing algorithms. The application of CSCS to find a best regression line is shown in Figure 3. The threshold  $Th_1$  for the CSCS is set to the average character width, which is calculated in terms of the vertical histogram profile after a rough correction of the skewed image. The steps to



implement the CSCS are described in Algorithm 1 and the output of this algorithm on a word image is depicted in Figure 4.

ii. Finding best fitting lines for Slope Angle Estimation

Our slope angle detection method is a two-pass approach that individually finds different slope angles for the sets of text pixels in two different profiles. Finally, a mean angle is calculated from this set of angles which is used as the estimated slope angle of the word image. The linear regression is used here to find the best fit lines and their corresponding slope angles on the top profile and the bottom profile of the denser word image  $I_d(x; y)$ . The procedure of coarse core-region  $I_c(x; y)$  detection is given in Algorithm 2.

**Algorithm 1** Procedure of CSCS

**Pre-requisite:**

$I_s(x, y)$  : The skeletonized word image

$Th_1$  : Average character width

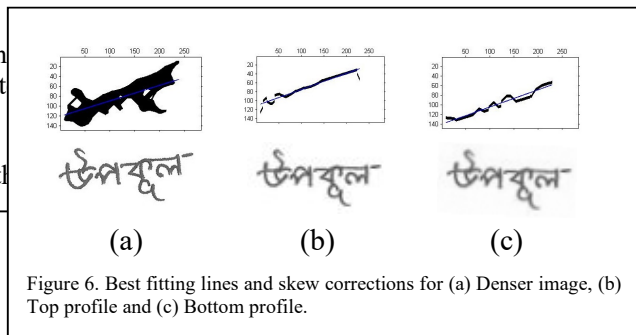
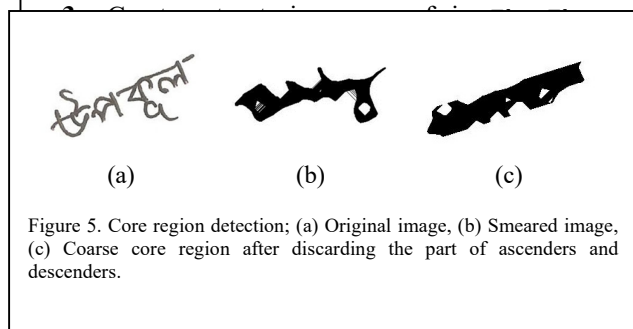
**Ensure:**

$I_c(x, y)$  : The denser word image

**Procedure CSCS**

2. Make a duplicate word image of skeletonized word

$$I_c(x, y) = I_s(x, y)$$



In the next step, we extract the top and bottom profiles of the core image  $I_c(x; y)$ . The top profile is extracted by scanning each column of the word image. The scanning starting from the topmost row for each column continues until a black run of length five pixels is obtained or the bottommost row of the word image is reached. Thus, the collection of the top five pixels for each column produces top profile. A similar procedure is applied for extracting the bottom profile, except this time the scanning starts from the bottommost row and goes progressively upwards. Linear regression is again applied on each of these two sets of profile pixels to get the best fit lines and their corresponding slope angles  $\theta_1$  and  $\theta_2$  respectively. These intermediate steps are illustrated with an example in Figure 6. Note that, the line  $L$  in Fig. 6(a) is slightly affected by the presence of ascenders and descenders, Fig. 6(b) and Fig. 6(c) show that the angles  $\theta_1$  and  $\theta_2$  are more accurate estimates of slope angles when those unwanted portions are eliminated. The mean  $\theta$  of these two angles is calculated which corresponds to the actual inclination of skewed word. Finally, the slope of the word image is corrected by rotating the word image about its centroid by the negative of the slope angle  $-\theta$ .



**Algorithm 2** Procedure of coarse core-region detection

**Pre-requisite:**

$I_d(x, y)$  : The dense word image

**Ensure:**

$I_c(x, y)$  : The image with coarse core-region

**Procedure CSCS**

1. Make a duplicate word image of skeletonized word

$$I_c(x, y) = I_d(x, y)$$

Let  $(x_i, y_j)$  denote the coordinate of text pixel  $P_{ij}$  of  $I_c$

2. Fit a regression line  $L$  on  $I_d(x, y)$  by a model function

$$y = \alpha x + \beta$$

where,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (X')^{-1}Y$$

$$X = [x_1, x_2, \dots, x_n], Y = [y_1, y_2, \dots, y_n]$$

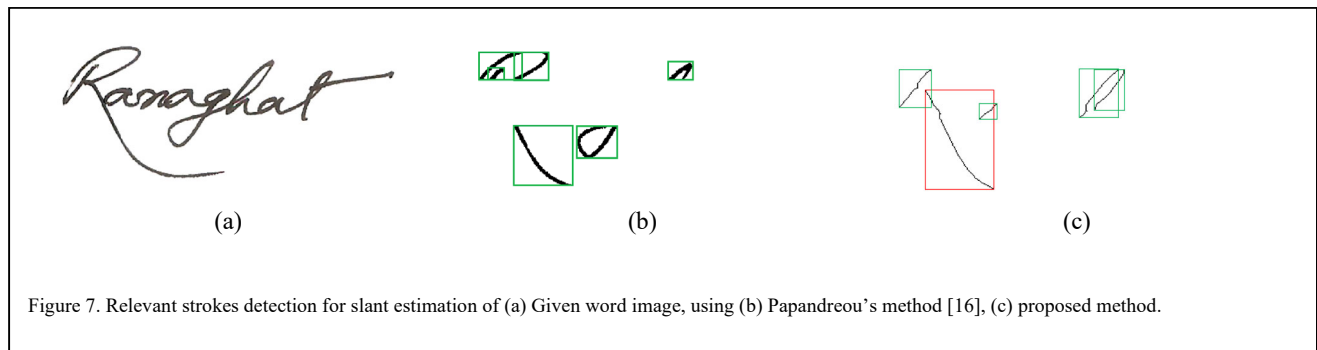
3. Calculate perpendicular distance  $D_{ij}$  from  $P_{ij}$  to  $L$
4. Calculate  $\mu_1$  and  $\sigma_1$ , the mean and standard deviation of all  $D_{ij}$
5. Discard the pixels  $P_{ij}$  from  $I_c(x, y)$ , if  $D_{ij} > Th$  where,  
Th =  $\mu_1 + \sigma_1$ , i.e.
6. Scan  $I_s(x, y)$  from top to bottom and do the followings

$$I_c(x_i, y_i) = 1 \forall d_i > \mu_1 + \sigma_1$$

*C. Slant correction*

Our slant correction method is based on identifying the relevant strokes in the word image and determining the average slant of all such strokes as shown in Figure 7. The entire procedure is subdivided into three parts.

*i. Recognition of most significant area (MSA)*



Let  $I_{ds}(x, y)$  be the de-skewed skeletonized word image obtained from the previous step and let,  $I_p$  and  $I_q$  be the height and width of  $I_{ds}$  respectively. In a word image, the MSA always corresponds to its busy zone or

core-region, because it contains most of the relevant information and the portions outside the busy zone are mostly multi-oriented small strokes which are less significant for slant estimation. So, busy zone detection plays an important role here. The core-region is the area between the upper and lower base lines. In general the core-region of a word image is estimated by a statistical measurement over the horizontal projection profile. This may not be a suitable technique for a word where the pixel distribution over the ascenders and descenders is similar to that of the core-region. Thus, we introduce a simple horizontal black run profile  $R(x)$  to emphasize the length of the horizontal black runs as well as their frequency in a line.  $R(x)$  is defined as:

$$R(x) = \max(L(x, i) * [C(x)]^2) \sum_{i=1}^{C(x)} L(x, i) \quad (8)$$

where  $L(x; i)$  is the run length of the  $i^{th}$  black run and  $C(x)$  is the number of black runs in the  $x^{th}$  line. Figure 8 shows the quantitative comparison between the traditional horizontal profile and that obtained by our technique. The upper baseline  $UB$  and the lower baseline  $LB$  are defined as the longest run of  $R(x)$  where,  $R(x) > \mu_2$  and  $\mu_2$  is the mean of  $R(x)$ ,

$$\arg \max_i \left( \sum_{x=S_z}^{E_z} (R(x) > \mu_2) \right) \quad \{UB, LB\} = \{S_z, E_z\} \quad \text{where, } z = \quad (9)$$

### ii. Removal of Insignificant strokes

This section identifies the most relevant strokes for slant angle estimation. It is axiomatic that horizontal or near horizontal strokes do not have any significant contribution towards estimation of the slant angle as long as we do not consider word images which are slanted through large angles. Therefore, it is beneficial to remove the near horizontal strokes or the near horizontal portions of any strokes. To do that, we propose another simple and effective algorithm based on the skeletonized features of word images. When there is a patch of three or more consecutive text pixels in a row of skeletonized word  $I_{ds}$ , then all the consecutive pixels starting from the 3rd pixel are made background, as shown in Figure 9. This simple algorithm retains the original shape of a stroke, as it only removes the horizontal stroke or the horizontal portion of any stroke, without removing the entire row. Finally, in this step the small connected components (CCs) or the components having  $C_{height} = C_{width} \leq 1:2$  are deleted, as these components are less significant in the slant estimation. From Figure 7, it is evident that the stroke shapes are better preserved in our proposed method than in the most recent approach proposed by Papandreou et al. [11].

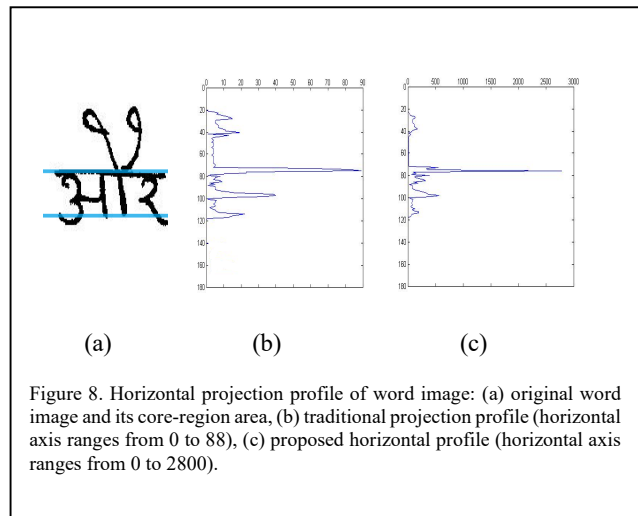


Figure 8. Horizontal projection profile of word image: (a) original word image and its core-region area, (b) traditional projection profile (horizontal axis ranges from 0 to 88), (c) proposed horizontal profile (horizontal axis ranges from 0 to 2800).

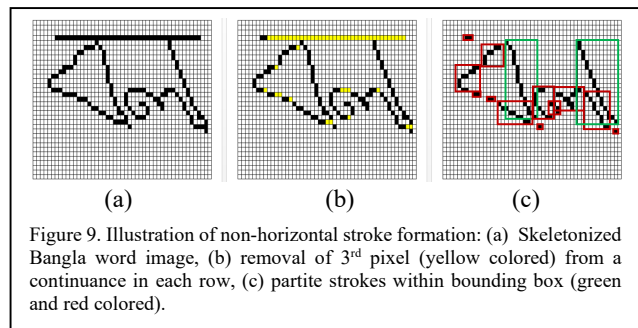


Figure 9. Illustration of non-horizontal stroke formation: (a) Skeletonized Bangla word image, (b) removal of 3<sup>rd</sup> pixel (yellow colored) from a continuance in each row, (c) partite strokes within bounding box (green and red colored).

### iii. Slant Angle Detection and Correction

It is evident that the longest linear (vertical) parts of the characters inside the core-region have greater significance than the parts outside the core-region to overall estimation of the slant angle of a word. The weighting  $W$  of each such stroke  $S_i$  is given by the expressions:

$$W(S_i) = ([H(S_i)]^2 * P(S_i)) \quad (10)$$

$$P(S_i) = \begin{cases} 0, & \text{if } S_i \text{ is outside of core region} \\ \frac{H(CR) - L_n}{H(CR)}, & \text{otherwise} \end{cases} \quad (11)$$

where,  $H(S_i)$  and  $P(S_i)$  are the height and percentage of height of the  $i^{\text{th}}$  stroke belonging to the core-region respectively.  $H(CR)$  is the height of the core-region of the word image.

At the next step, the strokes having  $W(S_i) > Th_2$  are considered as the most significant strokes for overall slant estimation, where,

$$Th_2 = \mu_3 + \sigma_3$$

and  $\mu_3, \sigma_3$  are estimated over  $W(S_i)$ . For all such extracted strokes, the best lines are fitted. Finally, some of the extracted strokes are again discarded if they are fewer in number as compared to the number of strokes of opposite orientation. The mean of angles for the remaining strokes is considered as the overall slant angle  $\phi$  for the word image. The slope corrected word is finally slant corrected using a shear transformation at angle  $-\phi$ . The intermediate steps of the overall slant correction procedure are depicted in Figure 10.

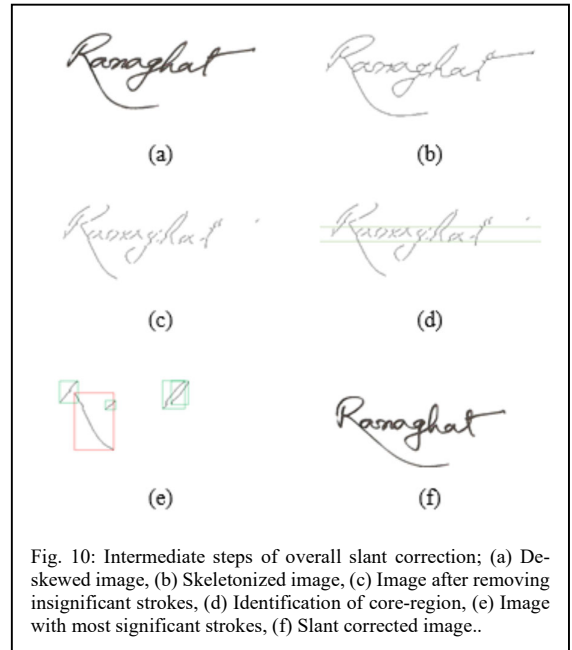


Fig. 10: Intermediate steps of overall slant correction; (a) De-skewed image, (b) Skeletonized image, (c) Image after removing insignificant strokes, (d) Identification of core-region, (e) Image with most significant strokes, (f) Slant corrected image..

## V. DATASET AND GROUND TRUTH (GT) PREPARATION

In any document analysis and recognition system, the dataset always plays a vital role. Though many methods have been developed for slope and slant correction at the word level, in most of the cases researchers have not made their datasets with corresponding GT information publicly available for future research. To address this problem, we have prepared four different datasets [22], among which

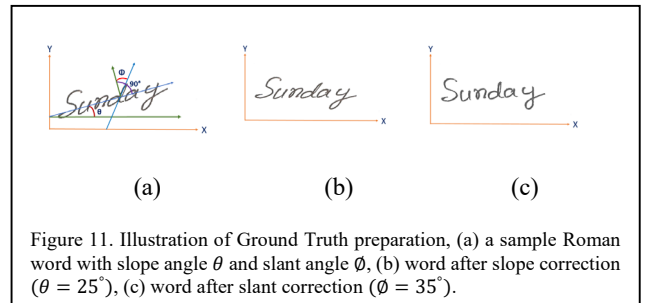


Figure 11. Illustration of Ground Truth preparation, (a) a sample Roman word with slope angle  $\theta$  and slant angle  $\phi$ , (b) word after slope correction ( $\theta = 25^\circ$ ), (c) word after slant correction ( $\phi = 35^\circ$ ).

three are in-house datasets namely Dataset A, Dataset B and Dataset C written in Bangla, Devanagari and

Roman scripts respectively and a synthetic dataset, named Dataset D, prepared using word images taken from the IAM dataset [21].

To prepare the in-house datasets, we use handwritten words in three different scripts Bangla, Devanagari and Roman. The words are written by several writers from different age groups, starting from school-going children to elderly people. As different people have different writing styles, such variations make our datasets more realistic as well as more challenging to demonstrate the robustness of our algorithm. Nevertheless, these Matra based and non-Matra based languages have significantly different characteristics e.g. the English words can be written in uppercase or lowercase letters whereas, Bangla and Devanagari have compound characters, modifiers, ascenders, and descenders apart from basic characters. The compound characters are usually more complex than the constituent basic characters. The text written in three scripts contains abundant variations with enough challenge to motivate the researchers making a script (or language)-invariant slope and/or slant correction method. To collect the word samples, we made 30 equal sized blocks on an A4-size white sheet and writers were requested to write words naturally inside the blocks. Then the A4 sheets were scanned at a 300 dpi resolution and saved as RGB images. The individual words were then cropped programmatically. Finally, we selected 500 words for each of the three scripts to prepare 1500 benchmark text words. It has been noticed that even though words are written within boxes, most of the word samples contain natural skew as well as slant.

The qualitative comparison among different algorithms is only possible if the proper GT values are available. The GT images (i.e. de-skewed and then de-slanted) along with slope and slant angles corresponding to each dataset have been prepared in a semi-automated way. This is because the estimation of actual slope and slant angles is only possible by visual perception of human beings because the traditional methods fail in some cases. For slope estimation, we use the work introduced by Bera et al. [14] and the failure cases are handled by the well-known software **IrfanView444**. After suitably deskewing, slant angles associated with the word images are estimated by using another recent work suggested by Papandreou et al. [11] and in a similar way the remaining cases where the outputs are not as desirable, are handled by shearing the image. For better understanding, the GT preparation is pictorially depicted in Figure 11. On analysis of 1500 handwritten words, it is noticed that the slope angles varied from  $0^\circ$  to  $25^\circ$ , whereas the slant angles varied from  $0^\circ$  to  $40^\circ$  in most cases.

Dataset D is completely constructed from the synthetic data collected from the IAM dataset which is publicly available. In order to prepare the Dataset D, we selected 500 words from a large set of IAM. The chosen words were visually checked to be free of skew as well as slant and be more than one character in length. The selected words are then binarized and cropped to contain all the text pixels and reduce extra manipulation of background portions. Then the cropped words are slanted at 19 different angles from  $-45^\circ$  to  $45^\circ$  with a step of  $5^\circ$  by using the shear transform. In this way we generated 9,500 synthetic words with known GT. Finally, each of the slanted words are skewed at 19 different angles from  $-45^\circ$  to  $45^\circ$  with  $5^\circ$  steps by rotating the word images about their centroids. This procedure generated 180,500 synthetic words for which the GT values are known.

## VI. EXPERIMENT AND ANALYSIS

We have implemented our algorithm using MATLAB2017 software on a PC with 16 GB RAM and Intel(R) Core(TM) i5-7500 CPU @ 3.40GHz 3.41 GHz processor. We have also implemented 14 well-known techniques, 7 for slope correction [5–7, 12–15] and 7 for slant correction [2, 9–11, 14, 15, 19] on the same platform to compare our techniques with state-of-the-art slope and slant correction techniques. For all these algorithms, the estimated slope and slant angles of each word are compared with the corresponding GT angles in terms of absolute error ( $E$ ) in degrees. Lower error implies more accuracy. We then calculated the average value ( $\mu_E$ ) of these absolute errors. Then the average ( $\mu^-$ ) and variance ( $S^2$ ) values of all  $\mu_E$  are calculated.

The proposed algorithm takes all such raw words images (naturally skewed and slanted) as inputs and subsequently corrects the slope and slant angles to normalize it. For the three datasets A, B and C, Table I proves that our methods outperform the state-of-the-art slope and slant estimation methods. The efficacy of our algorithm is pictorially displayed in Figure 12 for three representative word images, where the outcomes of our proposed methods are much closer to the ground truth word images than the other state-of-the-art methods.

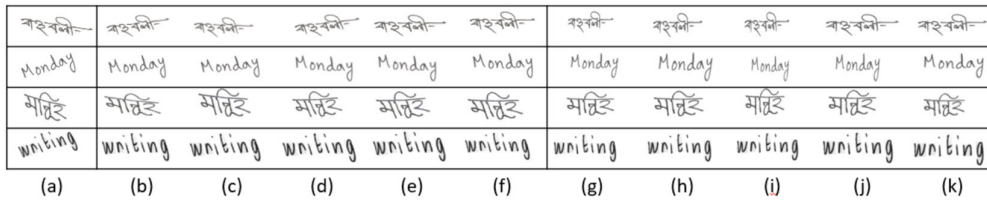


Figure 12. A sample word of IAM database and its different forms prepared in the present work to create synthetic database D where,  $\theta$  and  $\phi$  correspond to slope and slant angles respectively.

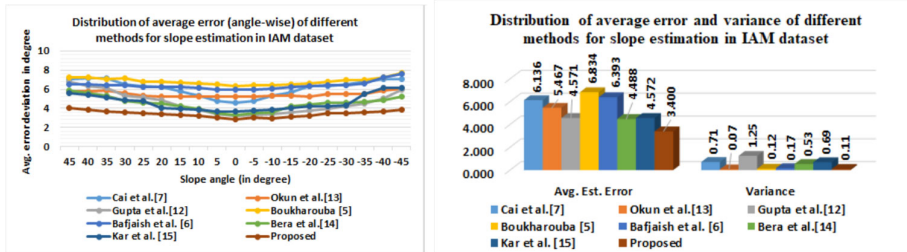


Figure 13. Distribution of (a) Average error  $\mu_E$ (angle-wise), (b) Average error  $\mu^-$  and Variance  $S^2$  of different methods for slope estimation in IAM dataset

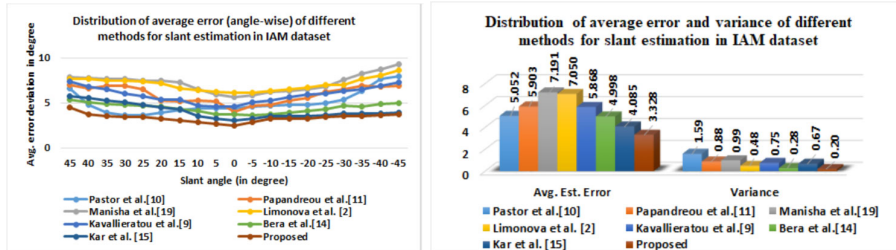


Figure 14. Distribution of (a) Average error  $\mu_E$ (angle-wise), (b) Average error  $\mu^-$  and Variance  $S^2$  of different methods for slant estimation in IAM dataset

The comparison results on the IAM dataset, as shown in Figure 13 and Figure 14, demonstrate the qualitative assessment of state-of-the-art slope and slant estimation methods along with the proposed methods. In both plots, the lines and bars with deep brown color correspond to the accuracy of our proposed method. It is observed that the brown lines in both cases lie below (mostly) all the other lines, and the brown bars are shorter than the others. These reveal that our methods outperform the state-of-the-art methods consistently at any range of curvature. We have also measured the time requirement of each algorithm. The average execution time (in seconds) of all these 14 methods along with our methods are given in Table 1. We can see that not only does our method produce accurate results on the words written in different scripts, but it is also fastest among all the 14 comparative algorithms.

We have evaluated our methods on the words written in three different scripts, having different characteristics. More specifically, two of these are Matra based namely Bangla and Devanagari, whereas the other one (Roman) is non-Matra based. Even though Bangla and Devanagari both are Matra based scripts, the number of characters in each alphabet is different. This difference would be wider if we consider the compound characters present in each alphabet. The nature of strokes, the ascenders and descenders are some other features which make each alphabet set unique. On the other hand, the Roman script has the lower case and upper case character sets. The three scripts are thus different in many aspects and still our technique shows reasonably satisfactory results on all three.

The experiments show, although the Hough transform based methods and projection profile based methods provide satisfactory results for Matra-based word images, they fail to produce desired results especially for images from non Matra-based scripts where the pixel distribution of outside core-region looks like that of the core-region. It may also be concluded that the slant estimation method, proposed by Papandreou et al. [11], is highly affected by the near vertical elongated strokes outside the core-region of the word. However, our proposed methods yield better estimation in all types of cursive word images because as in human perception,

TABLE I. AVERAGE ERROR DEVIATION (IN DEGREES) AND AVERAGE RUNNING TIME (IN SECONDS) OF PROPOSED METHODS AND STATE-OF-THE-ART METHODS

	Methods	Dataset A (Bangla)	Dataset B (Devanagari)	Dataset C (Roman)	$\mu$	S2	Avg. Running Time (s)
Slope Estimation	Cai et al.[7]	5.113	4.23	5.472	4.938	0.409	2.25
	Okun et al.[13]	4.418	4.964	4.668	4.683	0.075	1.84
	Boukharouba[5]	4.258	4.011	3.868	4.046	0.039	3.4
	Bafjaish et al.[6]	4.96	4.602	5.022	4.861	0.051	3.83
	Gupta et al.[12]	5.118	4.554	4.987	4.886	0.087	2.69
	Bera et al.[14]	4.102	4.556	4.652	4.437	0.086	1.99
	Kar et al.[15]	3.942	4.028	4.251	4.074	<b>0.025</b>	2.24
	<b>Proposed</b>	<b>3.824</b>	<b>3.744</b>	<b>3.46</b>	<b>3.676</b>	0.037	<b>0.87</b>
Slant Estimation	Pastor et al.[10]	4.356	5.248	4.602	4.735	0.212	1.96
	Manisha et al.[19]	4.882	6.312	5.864	5.686	0.535	3.34
	Limonova et al.[2]	4.212	5.642	6.022	5.292	0.911	4.4
	Papandreou et al.[11]	3.447	4.95	4.642	4.346	0.63	1.21
	Kavallieratou et al.[9]	4.986	5.352	4.756	5.031	0.09	3.74
	Bera et al.[14]	3.876	4.404	4.025	4.102	0.074	2.54
	Kar et al.[15]	3.86	4.201	3.491	3.851	0.126	1.48
	<b>Proposed</b>	<b>3.528</b>	<b>3.248</b>	<b>3.014</b>	<b>3.263</b>	<b>0.066</b>	<b>0.22</b>

it highlights the use of the top and bottom profiles in slope estimation rather than focusing only on the accurate core-region and it emphasizes the near vertical strokes (as a whole) throughout the core-region for slant estimation.

## VII. CONCLUSION

A simple and well formalized approach for slope and slant correction of handwritten text words is presented in this paper. The novelty of our proposed methods are that they contribute in three areas: slope and slant angle estimation, along with a unique core-region detection. This is done by first empirically separating the top and bottom profiles from the word's coarse core-region, which are then used to determine the existing overall skewedness of given text word. Afterwards, a sensible horizontal black run profile is applied on the

de-skewed word to mark out the actual core-region to sieve the longest near vertical strokes throughout that extent. Those extracted near vertical strokes eventually help to fix up the inclination of words. The experimental results with corresponding discussion reveal that the proposed methods are quite efficient as well as effective to normalize the unconstrained cursive handwritten text words by adjusting the slope and slant angles. The values of slope angle and slant angle obtained by our proposed methods are much closer to the GT (actual) values, thus we can say that this algorithm would certainly make the OCRing of the text words much easier.



## References

- [1] R. O. Duda, P. E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Communications of the ACM* 15 (1) (1972) 11–15.
- [2] E. Limonova, P. Bezmaternykh, D. Nikolaev, V. Arlazarov, Slant rectification in Russian passport ocr system using fast Hough transform, in: *Ninth International Conference on Machine Vision (ICMV 2016)*, Vol. 10341, International Society for Optics and Photonics, 2017, p. 103410P.
- [3] O. Boudraa, W. K. Hidouci, D. Michelucci, An improved skew angle detection and correction technique for historical scanned documents using morphological skeleton and progressive probabilistic hough transform, in: *Electrical Engineering-Boumerdes (ICEE-B), 5th International Conference*, IEEE, 2017, pp. 1–6.
- [4] U. Pal, B. B. Chaudhuri, An improved document skew angle estimation technique, *Pattern Recognition Letters* 17 (8) (1996) 899–904.
- [5] A. Boukharouba, A new algorithm for skew correction and baseline detection based on the randomized Hough transform, *Journal of King Saud University-Computer and Information Sciences* 29 (1) (2017) 29–38.
- [6] S. S. Bafjaish, M. S. Azmi, M. N. Al-Mhiqani, A. R. Radzid, H. Mahdin, Skew detection and correction of mushaf al-quran script using Hough transform, *International Journal of Advanced Computer Science and Applications* 9 (8) (2018) 402–409.
- [7] J. Cai, Z. Liu, Offline unconstrained handwritten word recognition, *International Journal of Pattern Recognition and Artificial Intelligence* 14 (3) (1998) 259280.
- [8] E. Kavallieratou, N. Fakotakis, G. Kokkinakis, Skew angle estimation for printed and handwritten documents using the wignerville distribution, *Image and Vision Computing* 20 (11) (2002) 813–824.
- [9] E. Kavallieratou, L. Likforman-Sulem, N. Vasilopoulos, Slant removal technique for historical document images, *Journal of Imaging* 4 (6) (2018) 80.
- [10] M. Pastor, A. Toselli, E. Vidal, Projection profile based algorithm for slant removal, in: *International Conference Image Analysis and Recognition*, Springer, Berlin, Heidelberg, 2004, pp. 183–190.
- [11] A. Papandreou, B. Gatos, Slant estimation and core-region detection for handwritten Latin words, *Pattern Recognition Letters* 35 (6) (2014) 16–22.
- [12] J. D. Gupta, B. Chanda, An efficient slope and slant correction technique for offline handwritten text word, in: *Proceedings of the Fourth International Conference of Emerging Applications of Information Technology (EAIT)*, 2014, pp. 204–208.
- [13] O. Okun, M. Pietikinen, J. Sauvola, Document skew estimation without angle range restriction, *International Journal on Document Analysis and Recognition* 2 (2) (1999) 132–144.
- [14] S. K. Bera, R. Kar, S. Saha, A. Chakrabarty, S. Lahiri, S. Malakar, R. Sarkar, A one-pass approach for slope and slant estimation of tri-script handwritten words, *Journal of Intelligent Systems*.
- [15] R. Kar, S. Saha, S. K. Bera, E. Kavallieratou, V. Bhateja, R. Sarkar, Novel approaches towards slope and slant correction for tri-script handwritten word images, *The Imaging Science Journal* 67 (3) (2019) 159–170.
- [16] A. Alaei, U. Pal, P. Nagabhushan, F. Kimura, A painting based technique for skew estimation of scanned documents, in: *proceedings of the Document Analysis and Recognition (ICDAR), International Conference*, IEEE, 2011, pp. 299–303.
- [17] B. Gatos, I. Pratikakis, A. L. Kesidis, S. J. Perantonis, Efficient offline cursive handwriting word recognition, in: *Tenth international workshop on frontiers in handwriting recognition*, 2006.
- [18] S. A. B. Haji, A. James, S. Chandran, A novel segmentation and skew correction approach for handwritten malayalam documents, *Procedia Technology* 24 (2016) 1341–1348.
- [19] C. N. Manisha, Y. S. Krishna, E. S. Reddy, Slant correction for offline handwritten telugu isolated characters and cursive words, *International Journal of Applied Engineering Research* 11 (4) (2016) 2755–2760.
- [20] Z. Huang, Q. Zhang, Skew correction of handwritten Chinese character based on resnet, in: *International Conference on High Performance Big Data and Intelligent Systems (HPBDIS)*, IEEE, 2019, pp. 223–227.
- [21] U. V. Marti, H. Bunke, The iam-database: an English sentence database for offline handwriting recognition, *International Journal on Document Analysis and Recognition* 5 (1) (2002) 39–46.
- [22] S. Bera, A. Chakrabarti, S. Lahiri, R. Sarkar, E. H. Barney Smith, *Handwriting-word-Database*, <https://scholarworks.boisestate.edu/saipl/2/> (2019).
- [23] N. Otsu, A threshold selection method from gray-level histograms, *IEEE transactions on systems, man, and cybernetics* 9 (1) (1979) 62–66.
- [24] P. P. Roy, U. Pal, J. Llads, Morphology based handwritten line segmentation using foreground and background information, in: *International Conference on Frontiers in Handwriting Recognition*, 2008, pp. 241–246.
- [25] P. G. McCrea, P. W. Baker, On digital differential analyzer (DDA) circle generation for computer graphics, *IEEE Transactions on Computers* 100 (11) (1975) 1109–1110.