

A Comparison Paper on Skew Detection of Scanned Document Images Based on Horizontal and Vertical Projection Profile Analysis

Bishakha Jain¹, Mrinaljit Borah²

¹Sikkim Manipal Institute of Technology, Majhitar, Sikkim.

²Jorhat Engineering College, Garmur, Jorhat, Assam.

Abstract- Skew detection has been an important part of the document recognition system. A lot of techniques already exist and has currently been developing for detection of skew of scanned document images. This paper describes the skew detection and correction of scanned document images written in Assamese language using the horizontal and vertical projection profile analysis and brings out the differences after implementation of both the techniques.

Index Terms- Skew Detection, Skew Correction, Projection-profile analysis.

I. INTRODUCTION

One of the most important steps of offline character recognition system is skew detection and correction which has to be used in scanned documents as a pre-processing stage in almost all document analysis and recognition systems. The skew of the scanned document image specifies the deviation of its text lines from the horizontal axis. The skew of the document image can be a global (all document's blocks have the same orientation), multiple (document's blocks have a different orientation) or nonuniform (multiple orientation in a text line) [1]. A variety of skew detection and correction techniques are available. Skew estimation approaches are classified into four main categories according to the basic approach they adopt [2]. It includes Hough transform [3], projection profile [4-7], nearest neighbor clustering [9] and interline cross correlation [10].

The traditional projection profile approach was initially proposed by Postl [4] and is based on horizontal projection profile. According to this approach, a series of horizontal projection profiles are calculated at a range of angles. The profile with maximum variation refers the best alignment to the text lines. At this stage, projection angle is the actual skew angle of the skewed document. In order to reduce high computational costs, several variations of this basic method have been proposed. Baird [5] proposes a technique for selecting the points to be projected: for each connected component the midpoint of the bottom side of the bounding box is projected. The objective function is to compute the sum of the squares of the profiles. Ciardiello [6] projected selected sub-region (one with high density of black pixels per row) of the document image; the function is to maximize the mean square deviation of the profile. Ishitani [7] uses a profile which is defined in a different style. A cluster of parallel lines on the image is selected and the bins of

the profile store the number of black/white transitions along the lines. These methods are limited to estimate skew angle within $\pm 10^\circ$ to 15° [8].

1.1. Introduction to Assamese script

The **Assamese script** is a writing system of the Assamese language. It used to be the script of choice in the Brahmaputra valley for Sanskrit as well as other languages such as Bodo (now Devanagari), Khasi (now Roman), Mising (now Roman) etc. The current form of the script has seen continuous development from the 5th-century Umachal/Nagajari-kanikargaon rock inscriptions written in an eastern variety of the Gupta script, adopting significant traits from the Siddham Script along the way. The present standard is identical to the Bengali script *except for three letters*. The alphabet of the modern Assamese script consists of 11 vowels and 40 consonants. These characters are called as basic characters. Writing style in Assamese is from left to right and the concept of upper/lower case is absent in this script. It can be seen that most of the characters of Assamese have a horizontal line (Matra) at the upper part. From a statistical analysis we notice that the probability that a Assamese word will have horizontal line is 0.994. In Assamese script a vowel following a consonant takes a modified shape. Depending on the vowel, its modified shape is placed at the left, right, both left and right, or bottom of the consonant. These modified shapes are called modified characters. A consonant or a vowel following a consonant sometimes takes a compound orthographic shape, which is called as compound character. Compound characters can be combinations of two consonants as well as a consonant and a vowel. Compounding of three or four characters also exists in Assamese. There are about 280 compound characters in Assamese.

II. METHODOLOGY

The projection profile can be used as a suitable feature for skew detection. We need to create a feature to describe which one is more peaked for comparing peaks of projection profiles. So employing a criterion function provides a numerical description of the peaks. The projection profile analysis process is as follows:

1. Rotate the binary input image to different angles and at any angle do "a" and "b".
 - a. Obtain the projection profile.
 - b. Calculate criterion function.

2. Skew estimation: obtain the angle corresponding to the maximum value of criterion function.

The horizontal projection profile is based on the histogram of black pixels along horizontal scan-lines. For a script with horizontal text lines, the horizontal projection profile will have peaks at text line positions and troughs at positions in between successive text lines [9]. This concludes to the fact that any noise and warp will ruin those peaks and troughs of the horizontal projection histogram and the efficiency of this technique. On the contrary a vertical projection method is robust to noise and warp of the image. The sum of squares of the projection profile elements as the value of the criterion function. This method also works well for the languages where most of their letters include at least one vertical line, such as languages with Latin alphabets. The pseudo codes of both the algorithms are as follows:

2.1. Vertical Projection profile Analysis Algorithm:

1. Read the image data into a matrix and convert it to grayscale.
2. This grayscale image is changed to black background and white writing on comparison each pixels with 0.34
3. Searches for the first column with a white pixel, i.e., with a written pixel.
4. The entire image column-wise is stored in a variable (Skew_input).
5. Each element of the input image matrix is added column-wise to get the number of white pixels per column and is stored in a variable Sum_col.
6. Sum of the squares of each Sum_col gives the value of energy function for the skew angle.
7. Input Image is rotated by angle "rot_angle" and steps 5 and 6 are repeated for this angle to obtain the value of energy function.
8. Input Image is rotated by angle "(-)rot_angle" and steps 5 and 6 are repeated for this angle to obtain the value of energy function.
9. $rot_angle = rot_angle - 1$.
10. Repeat steps 7, 8 & 9 till $rot_angle \neq 0$.
11. Find the angle for which the value of Energy function is maximum.
12. This angle gives the skew angle.
13. To display as output the values of energy function for each angle is displayed along with the bar graph for the column values for the skew angle and the corrected image segment.

2.2. Horizontal Projection profile Analysis Algorithm:

1. Read the image data into a matrix and convert it to grayscale.
2. This grayscale image is changed to black background and white writing on comparison each pixels with 0.34
3. Searches for the first column with a white pixel, i.e., with a written pixel.
4. One-Fourth of the image row-wise is stored in a variable (Skew_input).

5. Each element of the input image matrix is added row-wise to get the number of white pixels per column and is stored in a variable Sum_row.
6. Sum of the squares of each Sum_row gives the value of energy function for the skew angle.
7. Input Image is rotated by angle "rot_angle" and steps 5 and 6 are repeated for this angle to obtain the value of energy function.
8. Input Image is rotated by angle "(-)rot_angle" and steps 5 and 6 are repeated for this angle to obtain the value of energy function.
9. $rot_angle = rot_angle - 1$
10. Repeat steps 7, 8 & 9 till $rot_angle \neq 0$
11. Find the angle for which the value of Energy function is maximum.
12. This angle gives the skew angle.
13. To display as output the values of energy function for each angle is displayed along with the bar graph for the row values for the skew angle and the corrected image segment.

III. RESULTS AND DIFFERENCES

The implementation of both the algorithms brings out the following differences:

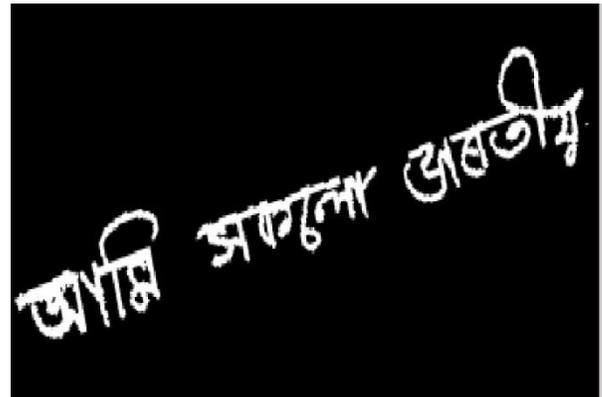


Fig1: Input Image

1. In the case of vertical projection profile analysis technique, the written line is scanned vertically and then aligned to find the maximum value of the objective function. But in the case of horizontal projection profile analysis technique the written line is scanned horizontally and then aligned to find the maximum value of the objective function.

2. The result obtained for Vertical Projection Profile Analysis Algorithm is not uniformly increasing showing that a little noise can create skew detection errors in this case.

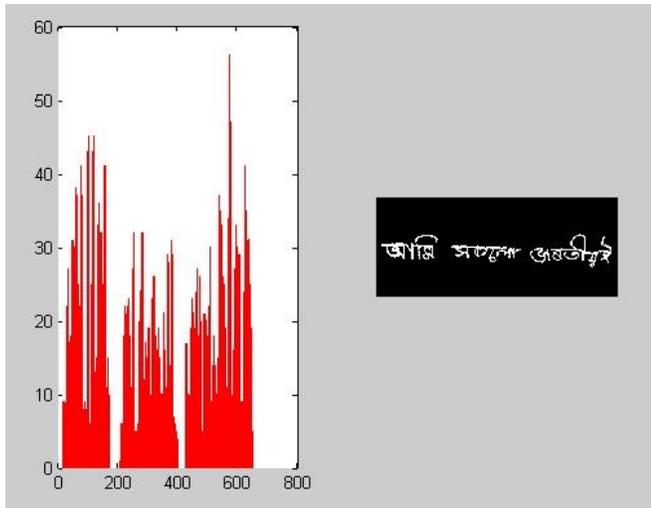


Fig2: Corrected Image Segment

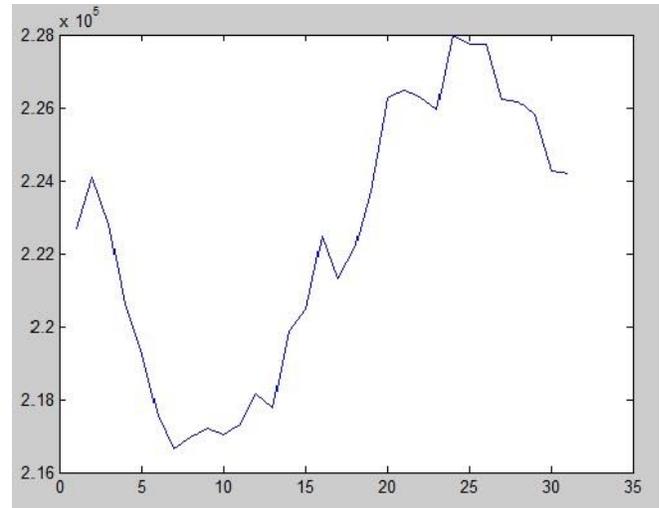


Fig3: Bar graph for column values of the skew angle

The angle of skew from this algorithm for the input image is -8 degrees.

Whereas the results obtained for horizontal projection profile analysis algorithm is uniformly increasing showing that it can give accurate results for a little noisy image.

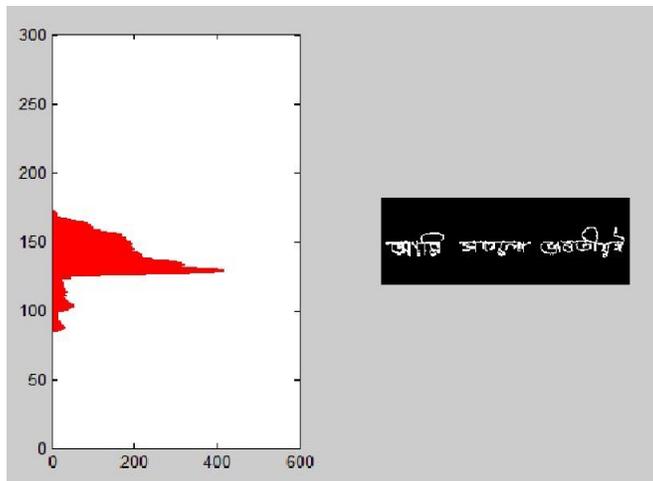


Fig4: Corrected Image Segment

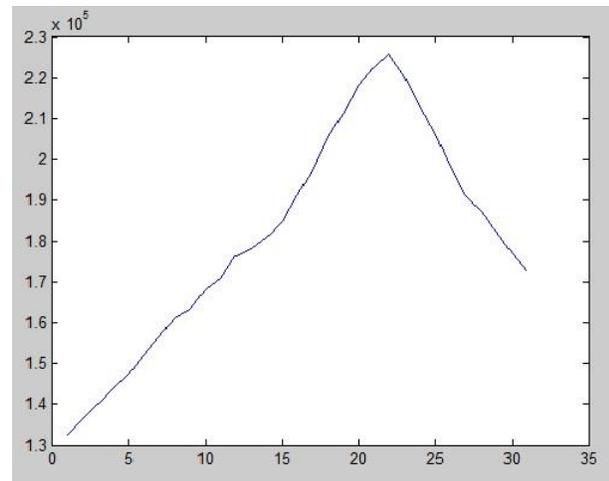


Fig5: Bar graph for column values of the skew angle

The angle of skew from this algorithm for the input image is -6 degrees.

3. Due to the non-uniform variation of the objective function in the vertical projection profile analysis technique, for accuracy in result we have to process the complete image to get better results. Whereas the uniform variation in the horizontal projection profile analysis technique helps to get better results with processing of even small portion of the written image.

4. The time complexity of the algorithm designed for vertical projection profile analysis technique is higher than the algorithm for horizontal projection profile analysis technique.

The differences are presented in tabular form as follows:

Vertical Projection Profile Analysis	Horizontal Projection Profile Analysis
<i>Working Principle:</i>	
1. The image is scanned vertically column-wise for number of white pixels for each column.	1. The image is scanned horizontally row-wise for number of white pixels for each row.
<i>Effect of Noise:</i>	
2. The resultant distribution of values of the objective function for various angles, contains multiple peaks	2. The resultant distribution of values of the objective function for various angles, contains a single peak value.

allowing a small noise at the peaks to produce errors.	Thus effect of noise is reduced greatly.
<i>Input Image data required:</i>	
3. To reduce effect of noise and thus improve accuracy the complete image is required to be processed.	3. As the effect of noise on the result is negligible so processing a section of the image can also yield results with high accuracy.
<i>Time Complexity:</i>	
4. The time complexity of the algorithm designed is high.	4. The time complexity of the algorithm designed is less than vertical projection profile analysis.
<i>Angle obtained:</i>	
5. The angle obtained through experimental results from this algorithm is -8 degrees	5. The angle obtained through experimental results from this algorithm is -6 degrees.

Fig6:Differences in tabular form

IV. SIMILARITIES

Apart from the differences observed there even exists some similarities between both the techniques of Projection profile analysis. Some of them are stated below:

1. Projection profile based approaches are computationally expensive as different projections are calculated at different angles in a particular range.
2. Projection profile methods are limited to estimate skew angle within $\pm 10^\circ$ to 15° .
3. The accuracy depends upon the angular resolution of the projection profile.
4. Projection profile based approaches cannot deal with noisy documents and broken characters.

V. CONCLUSIONS

In this work, we presented two algorithms for skew detection & correction. Both the algorithms are based on projection profile analysis, one based on vertical & other on horizontal projection profile. The subjective & objective evaluation of both these algorithms was executed and the results

of both the techniques were compared. The algorithm was implemented on input images of Assamese language. The skew for the test images were estimated and corrected and the performance was observed to be improved in the horizontal projection profile from that in the vertical projection profile. It was also observed that the horizontal profile technique could be used for skew correction with images with some noise. However the algorithm could only estimate skew if the angle is less than $\pm 15^\circ$.

REFERENCES

- [1] O. Okun, M. Pietikäinen, and J. Sauvola, "Document skew estimation without angle range restriction," International Journal on Document Analysis and Recognition, vol. 2, pp. 132-144, 1999.
- [2] K. Jung, K. I. Kim and A. K. Jain, "Text information extraction in images and video: a survey," Pattern Recogn vol. 37 issue 5, pp. 977- 997, 2004
- [3] S. N. Srihari and V. Govindaraju, "Analysis of textual images using the Hough transform," Mach Vis A, pp. 141-153, 1989.
- [4] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," Proc. 8th international conference on pattern recognition, pp. 687-689, 1986.
- [5] H. S. Baird, "The skew angle of printed documents," Proc. SPSE 40th symposium hybrid imaging systems, Rochester, NY, pp 739-743M, 1987.
- [6] G. Ciardiello, G. Scafuro, M. T. Degrandi, M. R. Spada and M. P. Roccotelli, "An experimental system for office document handling and text recognition," Proc. 9th international conference on pattern recognition, pp. 739-743, 1988.
- [7] Y. Ishitani, "Document skew detection based on local region complexity," Proc. 2nd international conference on document analysis and recognition, Tsukuba, Japan, pp. 49-52, 1993.
- [8] J. Sadri, M. Cheriet, "A new approach for skew correction of documents based on particle swarm optimization," Proc. 10th international conference on document analysis and recognition, ICDAR '09, pp. 1066-1070, 2009.
- [9] L. Gorman, "The document spectrum for page layout analysis," IEEE Trans Pattern Anal Mach Intell vol. 15 issue 11, pp. 1162-1173, 1993.
- [10] T. Akiyama and N. Hagita, "Automatic entry system for printed documents," Publisher Elsevier Science, Pattern Recognition vol. 23 , issue 11, pp.1141 - 1154, 1990

AUTHORS

First Author – Bishakha Jain, Sikkim Manipal Institute of Technology, Majhitara, Sikkim.
Second Author – Mrinaljit Borah, Jorhat Engineering College, Garmur, Jorhat, Assam.