International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.9, (2016), pp.139-146
http://dx.doi.org/10.14257/ijsip.2016.9.9.13

Skew Detection and Correction of Gurmukhi Words from Natural Scene Images

Balwinder Singh[1*] and Raman Maini[2]

[1]Assistant Professor (Computer Science),Yadavindra College of Engineering, Punjabi University, Guru Kashi Campus,Talwandi Sabo, Punjab (India)
[2]Professor, Department of Computer Engineering, Punjabi University,Patiala,Punjab (India)
[1]balwinder_cs@pbi.ac.in, [2]research_raman@yahoo.com

Abstract

Natural scene images are more susceptible to skew deformation as compared to document text which makes skew correction an indispensable step in scene text extraction. This work evaluates Murthy's Devanagari scene word slant correction method [Signal, Image and Video Processing, 7(6), 2012] on Gurmukhi scene images. The method makes use of headline feature of Devanagari which also exist in Gurmukhi script. The headline of Gurmukhi word is found by perceiving farthest located salient points as its end-points and skew angle of headline is calculated from its slope. Gurmukhi word image is de-skewed using skew angle of identified headline. The method has been tested on 100 self-captured good quality Gurmukhi and 117 publically available Devanagari scene words with average accuracy of 62.8% and 72.2% respectively. The method has been found to be working well on few samples of defective scene words, provided actual end-points of headline are preserved. It has been observed that Murthy's method is very simple to implement, does not require any pre-processing and give good results in wide variety of practical situations. However, this method does not work well for single character words with vowel above headline and words with identified headline parallel to horizontal axis.

Keywords: *Scene text extraction, skew correction, Gurmukhi, Devanagari*

1. Introduction

Skew correction is a pre-processing step to improve the performance of optical character recognizer. Skew [1] is deviation of baseline/headline of text word from horizontal axis as shown in Figure 1. Document text usually suffer from skew angle due to misalignment of an image during scanning/copying. Skew correction becomes even more crucial in scene text extraction (STE), where chances of dislocation are high due to multiple issues like non-stationary camera, 3-D space and distant focus *etc.*

*Corresponding Author

ISSN: 2005-4254 IJSIP

Figure 1. Gurmukhi Script Words Before and After Skew Correction

STE is an active area of research and a lot of work [2-7] has been done in past few years for various scripts including Indian. The method proposed for recognition

(a) (b) (c)

(d) (e) (f)

Figure 2. Natural Scene Images Containing (a)-(c) Gurmukhi Words and (d)-(f) Devanagari and Roman Words

of document text may not work well for scene text due to various challenges present in scene images [5-6] such as complex background, uneven lighting, and unpredictable text location, variety of fonts and sizes, perspective distortions *etc.*, So, methods for STE need to be investigated from a different viewpoint. Some example natural scene images containing Gurmukhi, Devanagari and Roman words are shown in Figure 2 (a)-(f).

This work replicates and carry out performance analysis of Murthy *et. al.*, [7] slant (skew) correction method of Devanagari for Gurmukhi scene words. Section 2 presents the overview of Gurmukhi script and details of Gurmukhi and Devanagari Scene text dataset used in our experiments. Section 3 discusses the details of skew correction method as emulated from Murthy. Simulation results are presented in Section 4 and Section 5 concludes the paper and give scope for future improvement.

2. Dataset Collection

2.1. Overview of Gurmukhi Script

Gurmukhi script is used to write Punjabi language in Punjab, a northern state of India. Gurmukhi script was standardized in 16th century by Guru Angad Dev, second Guru of Sikhs which is the youngest religion of India. The translated meaning of Punjabi word ‘Gurmukhi’ in English is ‘from the mouth of the spiritual Teacher’. Gurmukhi script consists of 35 basic consonants, 6 special consonants, 12 vowels, 3

half characters and 10 numerals (Figure 3). There is no concept of upper or lowercase characters and characters are normally aligned below the line of writing.

All Gurmukhi characters have a full/partial horizontal line (except numerals) at the top, called headline (Figure 1 and 3). Gurmukhi words are formed by joining common headline of each character to generate a bigger headline. A Gurmukhi script word can be partitioned into three distinct horizontal zones namely upper, middle and lower zone as shown in Figure 3(f). Gurmukhi script is structurally similar to Devanagari due to presence of similar headline in both scripts. Headline is missing in numerals of both Gurmukhi and Devanagari script.

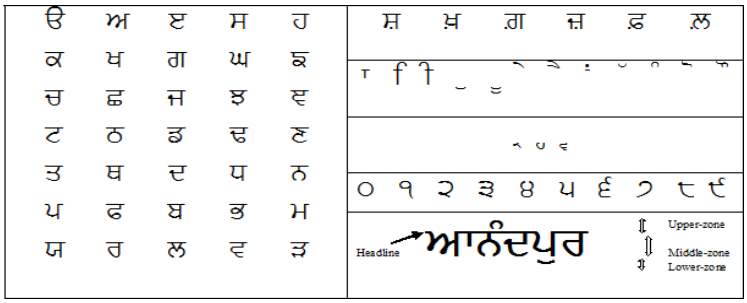


Figure 3. Gurmukhi Script (Clockwise from Top-Left), (a) Basic Consonants (b) Special Consonants (c) Vowels (d) Conjunct (Half) Characters (e) Numerals (f) Zones of Gurmukhi Word

2.2. Natural Scene Words Dataset

Due to unavailability of any standard Gurmukhi scene image dataset, colored images containing printed Gurmukhi words have been captured (jpeg format) in daylight under variable lighting conditions using Canon PowerShot A2400 Digital camera (16MP). Some of these original images are shown in Figure 2. The images contain mostly good quality Gurmukhi text of different sizes, styles and colors captured from road signboards, display boards, shop names and banners etc. Few special case images of hand-printed, noisy, low contrast, embossed and worn out text are also included in the dataset to check effectiveness of method in most practical situations. Further word image dataset is prepared by manually segmenting the Gurmukhi words with ImageJ software [9] and out of this 100 word images are used for experimentation. Some of the segmented Gurmukhi words used in our experiments are shown in Figure 4(a). Images containing Gurmukhi numerals are excluded from the scope of present study due to absence of headline in numerals.

(a)

(b)

Figure 4. Scene Word Samples of (a) Gurmukhi and (b) Devanagari [7]

As Murthy *et.al.,* [7] did not report any performance results for their skew correction technique, the replicated method has been also tested on Devanagari scene word dataset [1] released by them (jpeg & png format) for sake of result comparison. This publically available dataset consists of 117 Devanagari (both printed and hand-printed) scene words images captured from street names, product advertisement, shop names, caution signs etc. Few word images from Murthy's Devanagari scene words dataset are shown in Figure 4(b). This dataset do not contain any image with Devanagari numerals.

3. Methodology

Murthy *et.al.,* [7] proposed slant correction technique for pre-detected Devanagari scene word images. In literature, the approach to correct deviation of baseline/ headline [1] is traditionally known as correction of skew rather than slant. In present work, word 'skew' has been used as per usual convention in contrast to word 'slant' used by Murthy. The idea is to take advantage of unique headline property of Gurmukhi and Devanagari script words. As top of each character/word in both scripts is covered by headline (excluding numerals), firstly location of headline is identified and its skew angle with horizontal axis is detected. Then, skew of headline is corrected using traditional rotation transformation, which results in normalization of word image. The present work has been implemented on pre-segmented Gurmukhi words in MATLAB using Murthy's algorithm [7], whose step-by-step description is given below:

Step 1: In upper-half of the input word image salient points are detected using [8] as shown in Figure 5(b), as headline is presumed to be present in this upper region of word image.

Step 2: Euclidean distances and intermediate angles between all salient point pairs identified in step 1 are calculated.

Step 3: The pair with longest distance and intermediate angle less than 5˚ is perceived as end points of the header line as shown in Figure 5(c).

Step 4: The skew angle of identified headline is calculated and image is de-skewed only in case of significant skew (threshold >1˚) by using following rotation transformation as shown in Figure 5(d).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where θ, (x, y), and (x', y') represent skew angle of identified headline, pixel values on original, and skew corrected image respectively.

[1] Available at https://sites.google.com/site/dsiwiit/

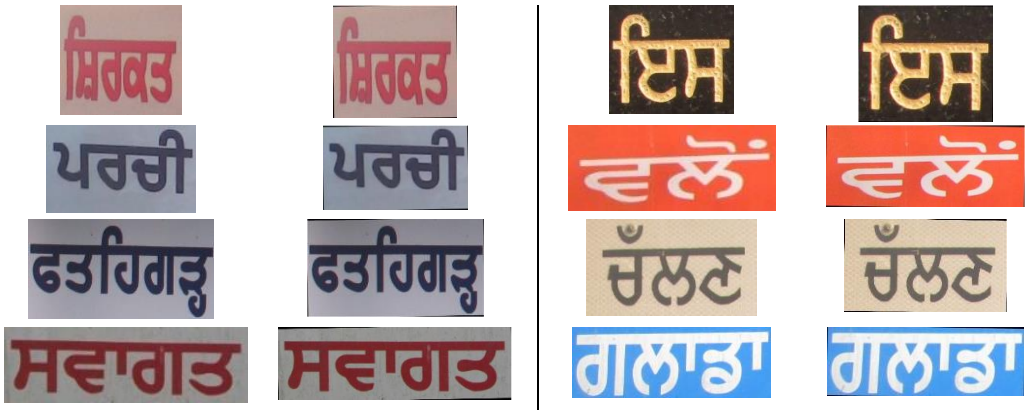
(a) (b) (c) (d)

Figure 5. Skew Correction of Gurmukhi and Devanagari Scene Words. a) Original Word Image b) Salient Points in Upper-Half c) Identified Headline d) De-Skewed Image

Simulation results of skew correction of Gurmukhi and Devanagari words are shown in Figure 5. Salient points are detected in upper-half of word image (marked red in Figure 5b). As the end points of headline of a Gurmukhi word are always at farthest distance, the pair of salient points with longest distance and intermediate angle less than 5˚ is earmarked as end points of the identified headline (drawn red in Figure 5c). The intermediate angle between end points of headline is kept small (<5˚) to avoid selecting a salient point (as headline end point) laying on a vowel above headline. The slope and accordingly skew angle of identified headline is calculated. Word image is de-skewed by using skew angle of identified headline (Figure 5d).

4. Results and Discussion

Primarily, the simulation of Murthy's Devanagari skew correction method [7] has been carried out to investigate its effectiveness for Gurmukhi scene words. As the Murthy's method did not make available any quantitative results of Devanagari skew correction, the experimental evaluation has been performed on both self-captured Gurmukhi and publically available Devanagari scene words to compare their results. The method has been tested on 100 Gurmukhi and 117 Devanagari scene words dataset (as discussed in section 2.2) separately. Figure 6-8 depicts the outcome of the skew correction experiments and results are listed in Table 1.



(a) (b) (c) (d)

Figure 6. Correctly Skewed Source (a),(c) and Resultant Images (b),(d)

Figure 6, shows side-by-side comparison of few test and resultant Gurmukhi word images obtained after experimentation. The results of skew correction are estimated visually, as human skew estimation has been found to be competitive. As noticed

from visual comparison of words before and after skew correction, the method is able to correct skew in Gurmukhi words of various size and text/background colour. The method amplify size of image after skew correction and introduces black colour background in skewed words along image border as can be noticed in Figure 6.

The results of the method on both Gurmukhi and Devanagari scene words are summarized in Table 1. The method has been found to be possessing an accuracy rate of 62.8% and 72.2% for Gurmukhi and Devanagari scene words respectively.

Table 1. Result of Skew Correction

Dataset	Total Words	Skewed Words	Corrected	%age Accuracy
Gurmukhi	100	70	44	62.8
Devanagari [7]	117	90	65	72.2

Accuracy is measured as the percentage of correctly skewed words to the total number of skewed words. A word image has been categorized as correctly skewed if de-skewed image shows a visible positive change (partial or full) in comparison to original source image when examined visually. The decision of skew correction results has been made by visual observation only and no evaluation metrics are used to arrive at the results.

To check effectiveness of technique in practical situations, method has been also tested on few special case images generally found in real situations. Figure 7 shows side-by-side comparison of special case source and de-skewed images. The method has been found to

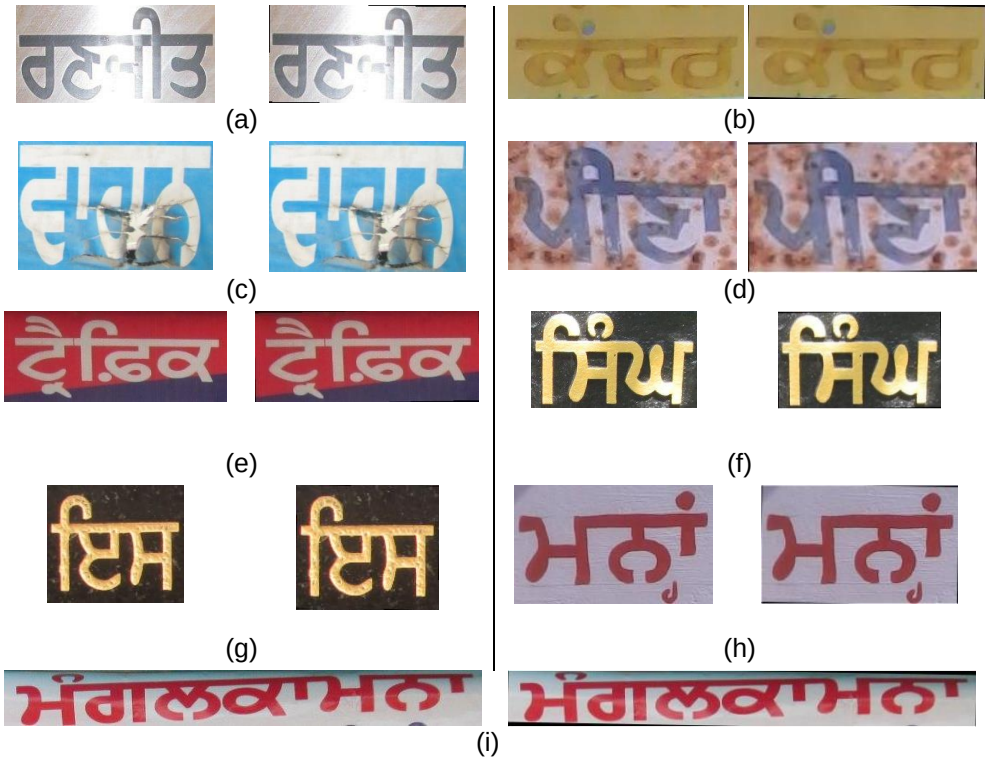


Figure 7. Results on Special Case Source and De-Skewed Word Images (a) Varying Illumination (b) Low Contrast (c) Distorted (d) Noisy (e) Multi-Colour (f) Embossed (g) Engraved (h) Hand-Printed (i) Curled Text

be working well on images with varying illumination and low contrast. The technique is also able to correct skew successfully even in distorted, noisy and curled images

provided actual end-points of headline are preserved. The method also works quite well for multi-coloured, engraved, embossed and hand-printed words. These results demonstrate that method can be applied well in a wide variety of practical situations.

(a) (b)

Figure 8. Failure Case Source and De-Skewed Images (a) Single Character Word with Vowel (b) Identified Parallel Headline

However, there are few failure cases also when the method is not able to correct the skew of Gurmukhi scene words properly. Figure 8(a), shows the failure instance in case of single character word with vowel above the headline, as headline is incorrectly identified being based on longest length. Figure 8(b), shows another failure instance when identified headline is almost parallel to horizontal axis which results in failure to detect the presence of skew. Sometimes method also do not work in low resolution images due to failure of Awrangjeb method [8] to detect salient points.

The beauty of Murthy's method is that it is very simple to implement and no pre-processing operation like enhancement or binarization is required. The technique is able to correct skew both in clockwise as well as anti-clockwise direction. Overall accuracy of method turns out to be 62.8% and 72.2% for Gurmukhi and Devanagari scene words respectively as calculated by visual estimation. It has been observed that performance of method is excellent on good quality words. The method also works on different type of Gurmukhi words found in most practical situations provided actual end-points of headline are preserved. Few failure cases are reported for less occurring single-character scene words and incorrect assessment of skew resulting from identified parallel headline.

5. Conclusion

In this paper, an attempt has been made to evaluate the performance of Murthy's skew correction method on Gurmukhi words dataset extracted from self-captured images of road signboards, display boards, shop names and banners *etc.* The method works well on good quality word images as well as on few samples of defective words found in most practical situations provided end-points of headline are preserved. Method is very simple to implement, robust to outdoor illumination conditions and does not need any pre-processing or enhancement operation. The method gives an average accuracy of 62.8% despite wide variation in text size and colour of words in dataset. However, the method do not work correctly in case of single character words and in case identified headline is almost parallel to horizontal axis.

Not much work has been done towards skew correction of scene images of Indian scripts in general and Gurmukhi in particular. This work would be helpful for selecting a skew correction method for Gurmukhi as well as other headline based Indian scripts at word level and results obtained may serve as baseline results for future evaluations of similar studies. In future, algorithm will be modified to incorporate slant correction module along with solution to flaws unveiled in present method.

Acknowledgment

The first author of paper would like to acknowledge financial assistance from the Punjabi University, Patiala in the form of study leave.

References

[1] P. SlavõÂk and V. Govindaraju, "Equivalence of different methods for slant and skew corrections in word recognition applications", IEEE Transaction of Pattern Analysis and Machine Intelligence, vol. 23, no. 3, **(2001)**, pp. 323–326.
[2] K. Jung, K. I. Kim and A. K. Jain, "Text information extraction in images and video: a survey", Pattern Recognition, vol. 37, no. 5, **(2004)**, pp. 977–997.
[3] H. Zhang, K. Zhao, Y. Z. Song and J. Guo, "Text extraction from natural scene image: A survey", Neurocomputing, vol. 122, **(2013)**, pp. 310–323.
[4] Q. Ye and D. Doermann, "Text Detection and Recognition in Imagery: A Survey", IEEE Transaction of Pattern Analysis and Machine Intelligence, vol. 37, no. 7, **(2014)**, pp. 1480-1500.
[5] J. Liang, D. Doermann and H. Li, "Camera-based analysis of text and documents: a survey", International Journal of Document Analysis and Recognition, vol. 7, no. 2, **(2005)**, pp. 84–104.
[6] S. Uchida, "Text Localization and Recognition in Images and Video", in Handbook of Document Image Processing and Recognition, Edited D. Doermann, and K. Tombre, Springer, London, **(2014)**, pp. 843-883.
[7] O. V. R. Murthy, S. Roy, V. Narang, M. Hanmandlu and S. Gupta, "An approach to divide pre-detected Devanagari words from the scene images into characters", Signal, Image and Video Processing, vol. 7, no. 6, **(2012)**, pp. 1071–1082.
[8] M. Awrangjeb, G. Lu, C. S. Fraser and M. Ravanbakhsh, "A fast corner detector based on the chord-to-point distance accumulation technique", Proceedings of the Digital Image Computing: Techniques and Applications, Melbourne, Australia, **(2009)** December 1-3.
[9] M. D. Abràmoff, P. J. Magalhães and S. J. Ram, "Image Processing with Image", Biophotonics International, vol. 11, no. 7, **(2004)**, pp. 36-42.

Authors

Balwinder Singh, received B.E. and M.Tech. degree from Punjab Technical University, Jalandhar and Punjabi University, Patiala respectively. He is currently pursuing Ph.D. from Department of Computer Engineering, Punjabi University, Patiala. He is working as Assistant Professor in Computer Science at Yadavindra College of Engineering, Punjabi University, Guru Kashi Campus, Talwandi Sabo. His current area of research is Image Processing, Document analysis and Pattern Recognition.

Raman Maini, received B.Tech. and M.Tech. in (Computer Science & Engineering) from PTU, Jalandhar and PAU, Ludhiana respectively. He got Merit certificate in his M.Tech at PAU. He received his Ph.D. from Punjabi University Patiala. He is working as a Professor in Department of Computer Engineering, Punjabi University, Patiala. He is an editorial board members and a reviewer of various National and International journals of repute. He is a life member of ISTE, IETE and Punjab Academy of sciences. His current area of research is Image Processing, Computer Networks, Software Engineering and Data mining. He has more than 40 research publications to his credit in National and International journals/Conferences of repute.

