Research Survey on Skew Detection of Devanagari Script

Trupti Jundale
Dept. of Computer Science
Solapur University
Solapur, Maharashtra

Ravindra Hegadi
Dept. of Computer Application
Solapur University
Solapur, Maharashtra

ABSTRACT
Skew detection is one of the important part of any document image processing or character recognition system. The successful skew detection and correction may leads to success of document image processing or character recognition system. The skew is may be of scanned document image or of handwritten data. This paper gives the survey of various skew detection techniques used for Devanagari script as well as other scripts. The methods which are used mostly for skew detection like Hough transform, Fourier transform, projection profile, principle component analysis, nearest-neighbor clustering and cross-correlation are described in this paper.

Keywords
Devanagari script, pre-processing, skew detection, skew correction

1. INTRODUCTION
In the recent years, the usage of digital based document as compared to paper based document is increased at a brisk rate. With modern appearance and extensive application of multimedia equipments, there is a growing demand to produce a paperless environment, hence document image processing and Optical Character Recognition (OCR) is playing a key role in conversion of the paper based environment to paperless electronic environment. Handwriting recognition has been one of the challenging research areas in the field of image processing and pattern recognition in the recent years. To make handwriting recognition successful it is required to detect and correct the skew of document text. Majority of writer fails to write in straight line and write in a particular angle from horizontal line. This angle from horizontal line is called the skew or slope. Moreover, writing style of script may vary from person to person and also may differ within the same person, so there is presence of skew in the handwritten data. Skew refers to the text which neither parallel nor at right angles to a specified or implied line. Skew is a distortion that is often introduced during scanning or copying of a document and it is unavoidable. Skew angle is the angle that text lines deviate from the x-axis. Generally skew varies in types as global skew (when all text lines have the same orientation), multiple skew (when some text lines have a different orientation than the others) and non-uniform skew (when the orientation changes within a text line).

Devanagari is one the mostly used writing system. India's national/official language, Hindi uses Devanagari script. Many other languages like Marathi, Sanskrit, Nepali, Kashmiri, Bhojpuri, Maithili, Dogri, Bodo uses Devanagari script. As the writing system of the Devanagari script is different than other scripts, it is difficult to detect skew of Devanagari script as compared to other scripts.

2. LITERATURE SURVEY
In the literature survey there is lot of research available for skew detection of scanned document image but less work is available for skew detection of handwritten text/data. Also the methods which are mostly used for skew detection are using Hough transform, principle component analysis, cross correlation, projection profile, nearest-neighbor clustering, Fourier transform. [1] proposed a method for skew detection and correction of handwritten Devanagari script using Hough transform and bounding box technique is used for extraction of words for skew detection. A morphological and projection profile analysis is used in [2] for skew detection of fabric images. In this, projection profile is computed at each angle and the angle that maximizes a criterion function is regarded as the skew angle. A fast and robust skew detection algorithm for gray-scale images is presented in [6].They use MCCSD (modified cross-correlation skew detection) algorithm for horizontal and vertical cross-correlation simultaneously to deal with vertically laid-out text, which is commonly used in Chinese or Japanese documents. Fourier transformation and angle of elevation theory is used to detect skew of document image in [3] which contains input of different languages such as English, Hindi, Panjabi (uses Devanagari script) and picture images. The nearest neigbhor chains (NNCs) are extracted from the adjacent nearest neigbhor (NN) pairs, in which the slopes of the NNCs with a largest possible number of components are computed to give the skew angle of document image in [7].

Apart from these techniques [4] uses linear regression technique for calculating skew of handwritten Devanagari script word and line. This paper illustrates the axes-parallel rectangle of binary image by finding minimum row and minimum column pixels and then applies linear regression technique on axes-parallel rectangle to detect skew angle. [5] uses central moments to detect the skew of Urdu document images. For thinning free pre-processing, they use random polygon over the text in document and then skew angle is calculated using central moments and the centroid of the document image. [8] calculates the Wigner-Ville distribution (WVD) of the projections taken at various angles of the Devanagari script word which are extracted from document image and are fitted in a standard frame. It takes the angle as a maximum intensity after applying WVD for skew correction. A robust technique for multi-skew angle detection of Indian documents containing Devanagari and Bangla script is described in [9]. The technique is used for skew detection of each text line containing in the document and they used

properties of digital straight line and clusters for detection of skew angle of each text line.

3. DEVANAGARI SCRIPT

Devanagari Script is a part of Brahmic family which is belonging from Indo-Aryan languages. It is written from left to right. Unlike Latin script, concept of upper/lower case is absent in Devanagari script. It consists of 14 vowels and 33 consonants. Generally every word written in most of the Devanagari Script will have a header line on group of characters, called as 'Shirorekha' and this is considered as one word [1]. Vowels that can be written as separate characters or by using diacritic marks on below, upper, before or after consonants are called modifiers. In Devanagari script, two or three consonants can be written as a single character, which is known as compound character. Fig.1 shows different features of Devanagari script.

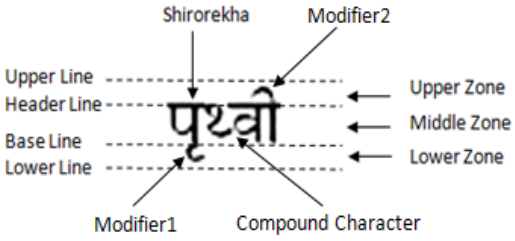
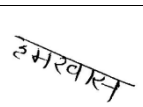
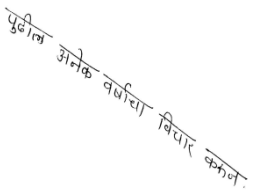
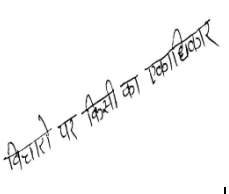
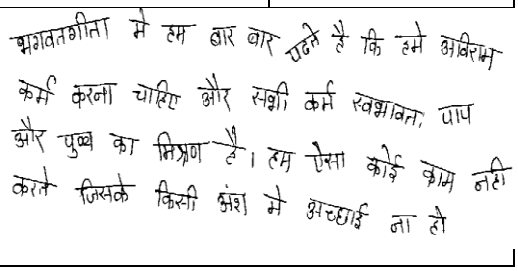


Fig. 1 Devanagari Script Word

The main characters of word are written in middle zone. Upper zone and lower zone are for modifiers and Shirorekha is drawn at header line. In Fig.1 two characters are combined and form a new shape of single character is a compound character.

The table 1 contains the samples of handwritten skew of Devanagari script contains word and line of single skew and text of non-uniform(multiple) skew.

Table 1 Samples of Skew of word, line and text

Word		
Line		
Text		

4. PROCESS OF SKEW DETECTION

The processes of skew detection consist of input acquisition, pre-processing, segmentation, skew detection and skew correction steps. Figure 2 shows the flow diagram of skew detection system.

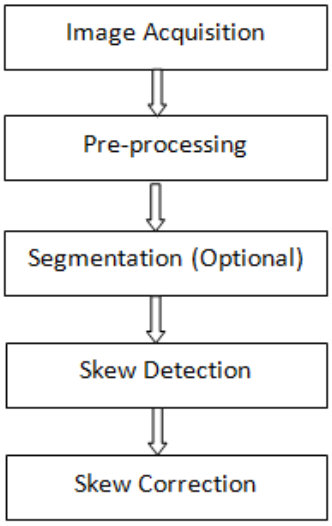


Fig. 2 Flow diagram of skew detection system

4.1 Input Acquisition

The input to the skew detection system is a scanned image of any picture supporting format like .jpg, .png, .bmp etc. The image is acquired through a scanner, digital camera or any other suitable digital input device.

4.2 Pre-processing:

Pre-processing is the major step in any recognition system. Pre-processing aims to produce clean document images that are easy for the next recognition systems to operate accurately. The pre-processing is a series of operation performed on the scanned input images. The scanned image is pre-processed for noise removal. The processes which get involved in pre-processing are illustrated below:

4.2.1 Binarization

Document image binarization (thresholding) refers to the conversion of a gray-scale image into a binary image (black and white image). Binarization separates the foreground (black) and background information (White). Thresholding (Binarization) methods can be classified into two categories as, global and local thresholding. Global methods apply one threshold value to the entire image. Local or adaptive thresholding methods apply different threshold values to different regions of the image.

4.2.2 Noise Removal

Scanned documents often contain noise that arises due to printer, scanner, print quality, age of the document, etc. Therefore, it is necessary to filter this noise before we process the image. Noise reduction improves the quality of the document. Two main approaches for noise reduction are, filtering or masks (which uses spatial domain and frequency domain filters for various operations like smoothing, sharpening, contrast stretching etc) and morphological operations (which uses operations like dilation, erosion, opening, closing etc).

4.2.3 Normalization

Normalization is the process of converting a random sized image into a standard size. It is applied to obtain characters of uniform size, slant and rotation. It includes basic techniques like, scaling, translation, and rotation etc.

4.3 Segmentation

After pre-processing, the next step towards angular skew removal is may be segmentation because segmentation step is may not use all skew detection systems. This step is used only when the system is used for detecting non-uniform skew of text. If system wants to detect and correct skew of words or lines that time segmentation is used. Only successful segmentation of individual words or lines or non-uniform data can ensure that the skew removal is performed. Segmentation is one the most important process that decides the success of character recognition technique. Segmentation is performed at different levels like line segmentation, word segmentation, and character segmentation. It is used to decompose an image of a sequence of characters into sub images of individual symbols by segmenting lines and words. In line segmentation the input image of the document is segmented into lines of text. This is achieved by reading the horizontal projection of pixels in the document. If the projection becomes zero this means that a line of text is finished and subsequently the line of text is segmented out of the image. After lines are segmented from text, words are segmented from lines by vertical projection profile. For each column of the line the number of black pixels is counted and the columns with zero black pixels are used as delimiters for word separation.

4.4 Skew Detection

Skew angle detection is considered as a significant part of any Optical Character Recognition and document analysis system because correct skew angle has a direct effect on the next recognition system. The largest classes of methods for skew detection are based on projection profile analysis, Hough transform, nearest-neighbor clustering, Fourier transformation, principle component analysis, cross correlation, histogram analysis, binary moments and other methods. Some of mostly used skew detection techniques are described here.

4.4.1 Hough Transform

Hough transform is a feature extraction technique used in digital image processing, image analysis and computer vision. The purpose of technique is to find lines, curves or any other parametric curves. It was introduces in 1962 by Paul Hough. The simplest case of Hough transform is the linear transform for detecting straight lines. The slope-intercept model of straight line is,

$$y = mx + c$$

where *m* is the slope and *c* is the y intercept. This straight line equation can be written in the form of parameter space is;

$$\rho = x\cos\theta + y\sin\theta$$

where ρ is the distance of line from origin, and θ is the angle of ρ with respect to x axis. The line in the image space is just a point in the parameter space. Hough transform uses two dimensional arrays called an accumulator array for detecting the existence of lines in image space where each row and column corresponds to θ and ρ values respectively. Peak is the strong point in the accumulator array which represents straight line in the image space. Once peak points are detected we can find endpoints of line segments corresponding to peak values. Hough transform method gives very accuracy rate as compared to any other methods but it is relatively slow and computationally expensive. Also, method is slow in presence of noise.

4.4.2 Projection Profile

Projection profile is one of the popular method used for skew detection in image processing. Projection Profile can be a horizontal projection profile or vertical projection profile. The total numbers of rows/columns in an image are stored in equal numbers of locations in one-dimensional array. The horizontal/vertical projection profile is a histogram of the number of black pixels along horizontal/vertical scan lines. In projection profiles, where a histogram is created at each possible angle and a 'cost function' is applied to this histogram. The skew angle is the angle at which this cost function is maximized. The disadvantages of this method are; when document contains diagrams, graphs etc., it fails to calculate skew angle. This method is not very efficient as projection profile of all possible angles need to be calculated. Methods are limited to estimate skew angle within ±10° to 15° and the accuracy of skew detection depends upon the angular resolution of the projection profile. It is very sensitive to noise.

4.4.3 Fourier transform

The Fourier transform of an image divides the image into different components, one for each frequency. Displaying the properties of each component is usually done in the form of an image - the spectrum, and interpreting this image/spectrum takes a bit of getting used to. The position of a value in a spectrum does not correspond to any position in the original image. Instead the position indicates the sine wave frequency that the value is providing data on. Each element in the Fourier transform represents the magnitude and phase of a basis function. We can reconstruct an image from its transform if we take each Fourier component, multiply it by its corresponding basis function, and finally sum them all. Fourier components come in pairs. For a real valued signal, the Fourier component at a frequency ω is $F(\omega)$, which is the complex conjugate of the Fourier component, $F(-\omega)$, at $-\omega$. The Fourier transform a 2D image having spatial domain $f(x,y)$ of size M*N is given by the following Equation, which calculates the frequency spectrum of the spatial co-ordinate function $f(u,v)$.

$$f(u,v) = \frac{1}{MN}\sum_{x=0}^{M}\sum_{y=0}^{N} f(x,y)\, e^{-2j\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

For large images, this method is computationally expensive since 2D Fourier transform of each pixel in the document has to be computed. Also very often for a document image, the largest density direction of Fourier space may be different than the true skew direction. These are the disadvantages of Fourier transform

4.4.4 Cross Correlation

Cross correlation matches a template or a pattern in a document. It measures the degree of similarity between an image and template. 2D cross correlation function is defined as,

$$c(u,v) = \frac{\sum_x \sum_y T(x,y)\, I(x-u, y-v)}{\sqrt{\left[\sum_x \sum_y I^2(x-u, y-v)\right]}}$$

Where T is a (X*Y) template and I is an (U*V). To find the degree of similarity between a template and an image, the cross correlation must be computed over all possible translations. This method calculate the document skew by finding the amount of vertical shifts needed to maximize the cross-correlation between pairs of narrow vertical columns of the documents.

4.4.5 Nearest Neighbor Clustering

Nearest neighbor method is based on connected components present in the documents. The first nearest-neighbors of all connected components are found and the histogram of the direction vectors for all nearest-neighbors is obtained. By using histogram peak, the skew angle can be found. Usually in the nearest neighbor method connected components are determined and using Euclidean distance nearest neighbor of each component is calculated. Then angles are accumulated in a histogram which is normally calculated between centroids of nearest neighbor components. The prevailing peak in the histogram represents the skew angle. The main advantage of this method is, it is not limited to any range of skew angle. Accuracy of this method decreases significantly, when noise and subparts of characters are present, is the disadvantage of nearest neighbour method. Also the process of nearest neighboring method is heavily depends on the quality of the binarization process output.

4.5 Skew Correction

After the skew angle of the input has been detected, it must be rotated in order to correct the skew. Various methods are used for skew correction are; direct method, indirect method and others like contour-oriented, projection based etc. The direct method uses rotation transformation in which corresponding pixels in the input image will be transformed to new location by using equation (1)

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

Where (x, y) are the co-ordinates of skew detected word and (x', y') are the co-ordinates for skew correction of word. The opposite of direct method is the indirect method. For a pixel (x', y') in the output image, the indirect method finds corresponding pixel in the input image and assigns a value of (x', y') to (x, y) using equation (2).

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

In general the skew is rotated with θ angle. If angle of skew is positive then it is corrected by rotating at negative angle and if angle of skew is negative then it is corrected by rotating at positive angle.

5. CONCLUSION

This paper presents the research survey of skew detection of Devanagari script. The literature review of existing methods is described and also the techniques which are mostly used for skew detection are described in the paper. The review concludes that most of skew detection techniques are based on Hough transform, Fourier transform, projection profile, principle component analysis, cross correlation etc. Also many skew detection techniques are prepared for document skew detection as compared to skew detection of handwritten or printed data.

6. REFERENCES

[1] Trupti A. Jundale, Ravindra S. Hegadi, "Skew Detection and Correction of Devanagari Script Using Hough Transform", International Conference on Advanced Computing Technologies and Applications, Procedia of Computer Science, Journal of Elsevier, ISSN: 1877-0509, March 2015, in press.

[2] Liu, Zhoufeng, Jie Huang, Chunlei Li, "Skew detection of fabric images based on edge detection and projection profile analysis", Foundations of Intelligent Systems, Springer Berlin Heidelberg, Vol. 122, pp 483-488, 2012.

[3] Ruby Singh, Ramandeep Kaur, "Improved Skew Detection and Correction Approach Using Discrete Fourier Algorithm", International Journal of Soft Computing and Engineering, Volume 3, Issue 4, pp 5-7, September 2013.

[4] Trupti A. Jundale, Ravindra S. Hegadi, "Skew detection of Devanagari script using axes-parallel rectangle and linear regression", International Conference on Innovations in Information Embedded and Communication Systems, IEEE, March 2015, in press

[5] R. J. Ramteke, Imran Khan Pathan, S.C. Mehrotra, "Skew angle estimation of Urdu document images: A moment based approach", International Journal of Machine Learning and Computing, Vol. 1, No. 1, pp. 7-12, April 2011.

[6] Ming Chen, Xiaoqing Ding, "A robust skew detection algorithm for grayscale document image", Document Analysis and Recognition, IEEE, pp. 617-620, Sep 1999.

[7] Yue Lu, Chew Lim Tan, "A nearest-neighbor chain based approach to skew estimation in document images", Pattern Recognition Letters, Elsevier, pp.2315-2323, March 2003.

[8] Rajiv Kapoor, Dipak Bagai, T.S. Kamal, "Skew angle detection of a cursive handwritten Devanagari script character image", J. Indian Inst. Sci., Vol. 82, pp. 161-175, 2002.

[9] Pal U., Mitra M., Chaudhari B. B., "Multi-skew detection of Indian script documents", Document Analysis and Recognition, IEEE, pp. 292-296, Sep 2001.