A Range Free Skew Detection Technique for Digitized Gurmukhi Script Documents

G S Lehal[1], Renu Dhir[2]

[1]*Department of Computer Science and Engineering, Punjabi University, Patiala, India, 147002*

gslehal@mailcity.com

[2]*Department of Computer Science and Engineering, Regional Engineering College, Jalandhar, India.*

renudhir@mailcity.com

Abstract

In this paper, a range free skew detection technique for machine printed Gurmukhi documents has been presented. This approach can easily be extended to other Indian language scripts such as Devnagri and Bakngla. Most characters in these scripts have horizontal lines at the top called headlines. The characters forming a word are joined at top by headlines, so that the word appears as one single component with headline. The ratio of pixel density above and below the headline of any word in Gurmukhi script is always less than 1. These inherent characteristics of the script have been employed and a new algorithm based on projection profile method has been devised. By inspecting horizontal and vertical projections at different angles in range [0°, 90°], the skew angle of the document in range [-180°, 180°] can be determined. Thus this approach is not limited to any range of skew angle and skewness in any document with orientation portrait or landscape and placed at any angle can easily be detected and removed.

Keywords : *Projection profile; Gurmukhi script; Skew detection*

1. Introduction

Skewness refers to the tilt in the bitmapped image of the scanned paper for Optical character Recognition (OCR). It is usually caused if the document is not well aligned on the scanner, thus yielding a skewed (rotated) digital image. Most of the OCR algorithms are sensitive to the orientation (or skew) of the input document image making it necessary to develop algorithms to perform skew detection and correction automatically. Several attempts to detect the skew angle in a document image have been reported. Most of the popular methods use projection profile [Postl, 1986] and its variation [Baird, 1987][Akiyama and Hagita,, 1990][Pavildas and Zhou, 1991]. In these approaches, the image is projected along a number of axes and the orientation-dependent histograms are computed. The skew angle is determined by the direction that maximizes an alignment criterion. Hough transform based approach to find the skew angle by selecting the peak value in the Hough plane as the detected skew angle has also been used for skew detection [Srihari and Govindaraju, 1989][Hinds et al, 1990][Le et al, 1994][Yu and Jain, 1996]. The running cost of the above two approaches is $O(n\theta)$, where n is the number of foreground pixels and θ is the number of different values of angles. Another approach uses k-NN (nearest neighbour) clustering [Hashizume et al, 1986][O' Gorman, 1993] of the connected components. This approach has a relatively high accuracy but a large computational cost, independent of the detection range, $O(n^2)$, where n is the number of connected components.

The problem of skew angle detection of documents containing major Indian language script forms has not received as much attention. Chaudhary and Pal [Chaudhary and Pal, 1997] have developed a technique for Indian language scripts such as Bangla and Devnagri which exploits the inherent characteristics of the script to determine the skew angle. The idea is to detect the head lines automatically and estimate the skew angles of these head lines. The method works well for detecting skew angles in range (-45° to 45°). Lehal and Madan [Lehal and Madan, 1998] have used the physical properties of Gurmukhi script to device a range free skew detection scheme for Gurmikhi script, which detects skewness in range -180° to 180°. A limitation of their algorithm is that it may not give correct results if graphic images or tables are present along with the text.

In this paper a technique for range free skew detection in machine printed Gurmukhi script has been suggested. The projection profile method together with the physical structure of Gurmukhi script have been combined to design the new skew detection technique, which can detect skewness at any angle. This approach can easily be adapted to other major Indian language scripts such as Devnagri and Bangla which have the concept of head lines.

2. Properties of Gurmukhi script

The word 'Gurmukhi' literally means from the mouth of the Guru. Gurmukhi script is used primarily for the Punjabi language. It is spoken by 84 million native speakers and is the world's 14th most widely spoken language. The Gurmukhi script alphabet consists of 40 consonants and 12 vowels (Fig 1). The writing style is from left to right. The concept of upper/lower case is absent. From Fig.1 it can be noted that most of the characters have a horizontal line at the upper part. The characters of words are connected mostly by this line called head line. A word in Gurmukhi script can be partitioned into two horizontal zones. The upper zone denotes the region above the head line, while the lower zone represents the area below the head line (Fig 2). The major part of the characters is located in the lower zone.

Consonants

u	a	e	s	h
c	k	g	G	L
C	x	j	J	M
t	T	D	Q	N
V	W	d	Y	n
p	f	b	B	m
y	r	l	v	R
S	z	K	F	Z

Vowels

A	i	I	U	<
-	O	E	>	
:	&	^		

Fig 1 : Character set of Gurmukhi script

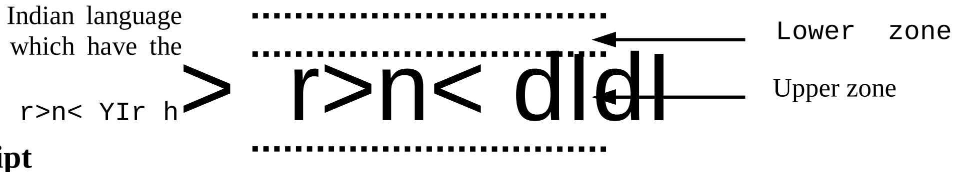

Fig 2 : Upper and lower zones of a word in Gurmukhi script

3. Proposed Skew Detection Algorithm

The skew angle is determined by calculating horizontal and vertical projections at different angles at fixed interval in range [0°, 90°]. A horizontal projection maps a two-dimensional bitmap into one-dimensional by recording the number of foreground pixels in each row of the bitmap in a one-dimensional array, where an entry in the array corresponds to the y coordinate of the row in the bitmap. Similarly the vertical projection keeps track of foreground pixel in each column (Fig. 3(a)). Under such projections, for an image with no skew, headlines appear as distinct peaks while gaps between successive text rows will be represented by valleys (Fig. 3(b)). Our task is to determine the angle at which the highest peaks and deepest valleys in the projections are present. Both the horizontal and vertical profiles are simultaneously examined for peaks and valleys. According to our empirical study, 10 peaks and 10 valleys at each angle are identified, assuming there are at least 10 lines of text, and a measure of difference of sum of heights of peaks and valleys was made and the angle at which the difference was maximum was recorded. It was observed that if the document contained large strips of rows or columns of foreground and background pixels, then the highest peaks and valleys occurred in consecutive rows or columns and thus highest peaks and valleys were detected in skewed image also, resulting in wrong skew angle estimation. To overcome this problem, the bitmapped image is partitioned into ten equal sized horizontal and vertical zones and highest peaks and valleys determined for projections in each zone. The angle at which the difference of the sum of heights of peaks and valleys is maximum is identified as the skew angle.

To decrease the computational cost, first the course skew angle is calculated by taking the angle interval 3°. Once the course skew angle is found, the accurate skew angle θ is determined by looking in the range [θ - 3°, θ + 3°] at intervals of 0.25°. The image is then rotated over -θ, where θ is the skew angle Since the skew angle is checked only in the range [0°-90°] and the image can be skewed at any angle in the range [-180°, 180°], the rotated image may need another additional rotation by 90°, -90° or 180°, depending on the range of skewness of the image (Table 1). After first rotation the bitmap image will be aligned along x or y-axis. If the rotated image is skewed at 90° or -90°, then the highest peaks and valleys would be present in vertical projection else they will be reported in horizontal projection. The physical characteristics of the Gurmukhi script are then used to determine the skew angle of the image after first rotation. To determine the skew angle of the image aligned with y-axis, if the foreground pixel density on the left side of headlines is greater than pixel density on right side for text rows then the image is skewed at -90 else it is skewed at 90° (Fig. 4 and 5). Similarly for the image aligned with x-axis, if the foreground pixel density above the headlines is lesser than pixel density below then the image is straight else it is upside down (Fig. 6). In the end the image is rotated by the second rotation angle to completely remove any skew present in the image.

It can be observed from above discussion, there are three main steps for skew detection and correction. A detailed description of these three steps follows.

Step 1. Course skew angle estimation.

for $\theta = 0°$ to $90°$ step $3°$
{
Rotate image by $\theta°$
Divide the image into 10 equal sized vertical and horizontal zones and compute vertical and horizontal projections in each zone.
Find the maximum and minimum value of projections in each zone and store in arrays maxv[10], minv[10], maxh[10], minh[10]
$d1 = \sum(maxv[i] - minv[i])$
$d2 = \sum(maxh[i] - minh[i])$
$d = max(d1, d2)$
if d = d1 then orient = VERT else orient = HORIZONT
}
Find the angle θ_c at which d is max and record θ_c as the course skew angle.

Step 2. Accurate skew angle detection and rotation of image

*Repeat all the substeps of step 1, by checking for skew angle in range [$\theta_c°$–3°, $\theta_c°$+3°] with step size 0.25° and determine the accurate skew angle θ_A in the range **[0°, 90°]**.*
Rotate the image by $\theta_A°$.

Step 3. Second rotation of image to remove skewness in the range [-180°, 180°]

If orient=VERT then for each maxv[i] determine the ratio of foreground pixel density to left and right of the column corresponding to maxv[i], where i ranges from 1 to 10. The pixel density is found by moving along a direction and summing the foreground pixels in each column until a column is encountered in which sufficient background pixels are found. If for majority of maxv[i], the ratio is lesser than 1 then rotate the image by -90° else rotate the image by +90°.

If orient=HORIZONT then for each maxh[i] determine the ratio of foreground pixel density to above and below the row corresponding to maxh[i], where i ranges from 1 to 10. If for majority of maxh[i], the ratio is more than 1 then rotate the image by +180°.

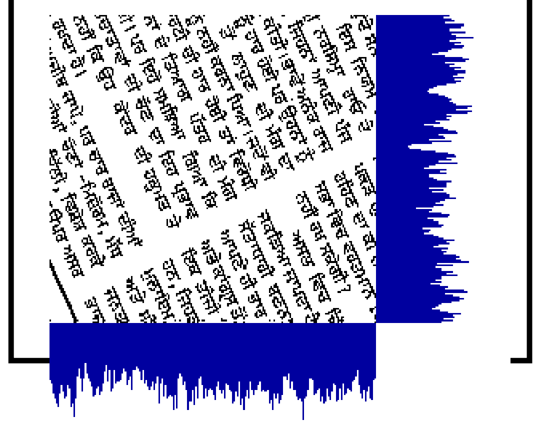

Fig 3(a) : Horizontal and vertical projections of image 1

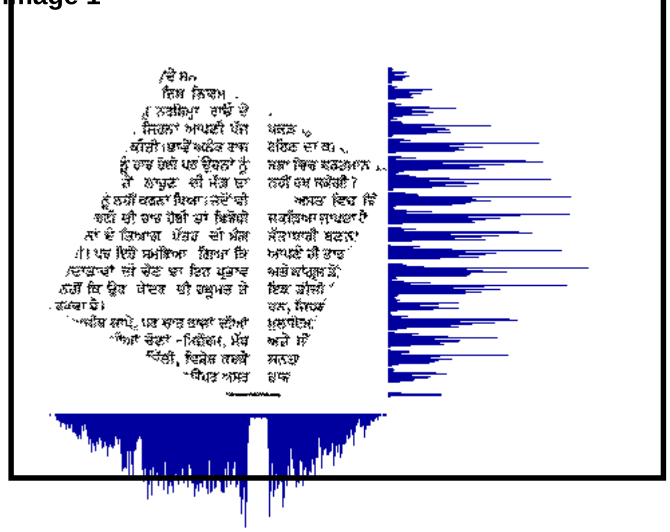

Fig 3(b) : Horizontal and vertical projections of image 1 after first rotation

Skew Angle Range	Alignment axis	Second Rotation Angle
[-180°, -90°]	y-axis	90°
[-90°, 0°]	x-axis	0
[0°, 90°]	y-axis	-90°
[90°, 180°]	x-axis	180°

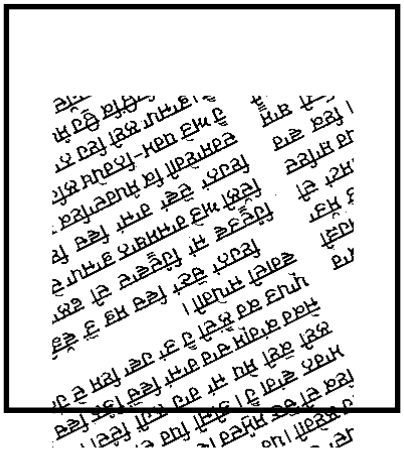

Table 1: Angle for additional rotation

Fig. 4(c) : Image 2 after second rotation

Fig 4(a) : Image 2

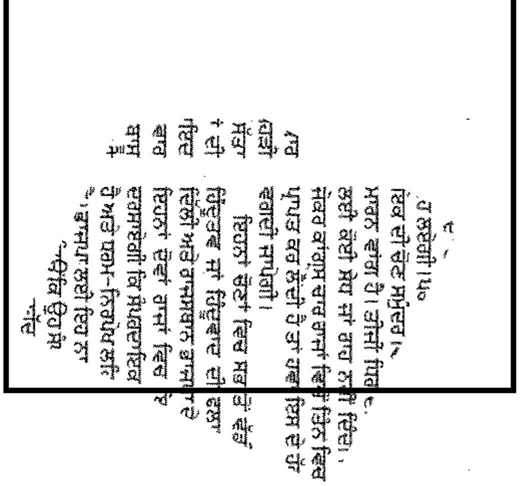

Fig. 4(b) : Image 2 after first rotation

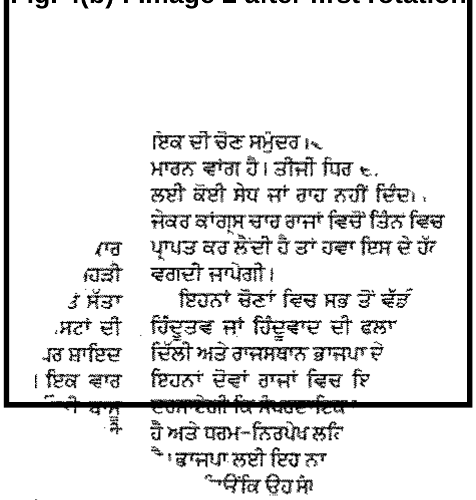

Fig 5(a) : Image 3

Fig 5(b) : Image 3 after first rotation

Fig. 5(c) : Image 3 after second rotation

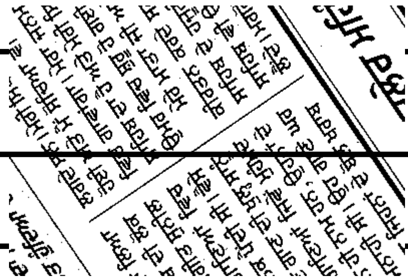

Fig 6(a): Image 4

Fig 6(b): Image 4 after first rotation

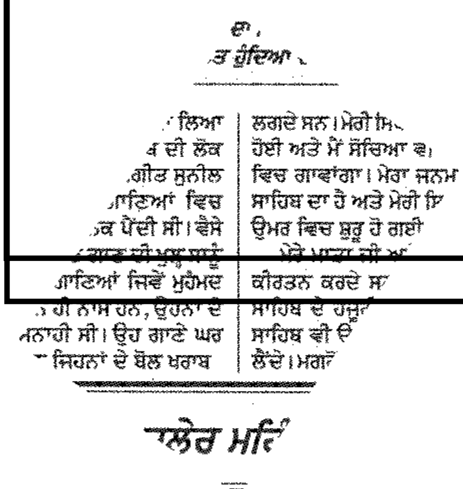

Fig 6(c): Image 4 after second rotation

4. Experimental Results

The skew detection algorithm proposed in this paper has been applied to a large number of document images. These images were obtained by scanning books, journals, magazines and newspapers using an HP ScanJet P5 scanner. The skew detection and correction program was executed on 133 MHz Pentium PC. For illustration purpose we have taken four scanned images (Figs. 3-6), skewed at different angles in each of the four quadrants. Table 2 gives information about the input image sizes, scanning resolution, estimated skew angle and processing time. The algorithm works well on all these document images which have different characteristics, noise levels and resolutions.

Image No.	Image Size	Skew Angle (°)	Resolution	Time (Sec)
1	204x184	-64.25	75 dpi	2.4
2	300x351	-153.5	150 dpi	3.9
3	340x299	54.0	75 dpi	5.0
4	408x367	124.5	150 dpi	5.2

Table 2: Characteristics of the images

5. Conclusions

A new technique for detection of skewness in machine printed Gurmukhi script, which takes advantage of the structure of characters of the Gurmukhi script to detect skew angle in the range -180° to +180° has been presented. One major advantage of the new algorithm is that now any document with orientation portrait or landcsape can automatically be made straight and fed to the character recognition algorithm. Previously on an A4 sized scanner an A4 sized page with landscape could not be as such scanned and recognized. The scanned image had to be first manually rotated at right angle and then skewness if any was removed and then it was passed for recognition. Another application has been in the usage of Automatic Document Feeder (ADF) where multiple pages of documents are fed for recognition. Now the documents could be of mixed orientations and placed at any angles, including upside down, and without any human intervention they can all be made straight. The algorithm has been found working satisfactorily for a large variety of documents tilted at all possible angles.

References

[Akiyama and Hagita, 1990]
Akiyama T and Hagita N. Automated entry system for printed documents, Pattern recognition, Vol. 23, pp. 1141-1154, 1990.

[Baird, 1987]
Baird HS. The skew angle of printed documents, Proc. SPIE 40th Ann. Conf. Symp. Hybrid Imaging Syst., pp. 21-24, 1987.

[Chaudhary and Pal, 1997]
Chaudhary BB and Pal U. Skew angle detection of digitized Indian script documents, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, pp. 182-186, 1997.

[Hashizume et al, 1986]
Hashizume A, Yeh PS and Rosenfeld A. A method of detecting the orientation of aligned components, Pattern Recognition Letters , Vol. 4, pp.125-132, 1986.

[Hinds et al, 1990]
Hinds S, Fisher J and D'Amato D. A document skew detection method using run-length encoding and the Hough transform, Proc. 10th Conf. Pattern Recognition, Atlantic City, pp. 464-468, 1990.

[Le et al, 1994]
Le DX, Thoma G and Weschler H. Automated page orientation and skew angle detection for binary document images, Pattern Recognition, Vol. 27, pp.1325-1344, 1994.

[Lehal and Madan, 1998]
Lehal G S and Madan S. A New Approach to Skew Detection and Correction of Machine Printed Gurmukhi Script, Proc. Int. Conf : KBCS 98, Mumbai, pp. 215-223, 1998.

[O'Gorman, 1993]
O'Gorman L. The document spectrum for structural page layout analysis, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 15, pp. 1162-1173, 1993.

[Pavildas and Zhou, 1991]
Pavildas T and Zhou J. Page segmentation by white streams, Proc. 1st Int. Conf. Document Anal. Recognition, Paris, pp. 945-953, 1991.

[Postl, 1986]
Postl W. Detection of linear oblique structures and skew scan in digitized documents, Proc. 6th Int. Conf. Pattern Recognition, Paris, pp. 687-689, 1986.

[Srihari and Govindaraju, 1989]
Srihari SN and Govindaraju V. Analysis of textual images using the Hough transform, Mach. Vis. Appl., Vol. 2, pp. 141-153, 1989.

[Yu and Jain, 1996]
Yu B and Jain K. A robust and fast skew detection algorithm for generic documents, Pattern Recognition, Vol. 29, pp. 1599-1629, 1996.