World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:8, 2007

Skew Detection Technique for Binary Document Images based on Hough Transform

Manjunath Aradhya V N*, Hemantha Kumar G, and Shivakumara P

Abstract**—**Document image processing has become an increasingly important technology in the automation of office documentation tasks. During document scanning, skew is inevitably introduced into the incoming document image. Since the algorithm for layout analysis and character recognition are generally very sensitive to the page skew. Hence, skew detection and correction in document images are the critical steps before layout analysis. In this paper, a novel skew detection method is presented for binary document images. The method considered the some selected characters of the text which may be subjected to thinning and Hough transform to estimate skew angle accurately. Several experiments have been conducted on various types of documents such as documents containing English Documents, Journals, Text-Book, Different Languages and Document with different fonts, Documents with different resolutions, to reveal the robustness of the proposed method. The experimental results revealed that the proposed method is accurate compared to the results of well-known existing methods.

Keywords**—**Optical Character Recognition, Skew angle, Thinning, Hough transform, Document processing

I. Introduction

THE study of Optical Character Recognition (OCR) by machine has attracted more attention recently for its application in many areas including bank cheque processing, postal ZIP codes and address recognition, office automation, document and text recognition and reading aid for the blind.

When the document is scanned, skew is inevitably introduced into the image due to various factors. Most of the popular algorithms for character recognition and layout analysis are however very sensitive to the distortion of document images. And skew may cause serious problem for document analysis. Furthermore, it is very difficult, if not totally impossible, to prevent skew distortion produced in scanning procedure. It is therefore preferable to detect and correct the skew document image at the preprocessing stage in order to avoid the disturbance of skew to the further processing.

Several attempts have been made for skew detection. And the methods can be mainly categorized into five groups. The one based on Hough transform, Cross Correlation, Projection profile, Fourier transformation and *K* nearest neighbor *(K-NN)* clustering.

Hough transform has been used by [13] for skew detection. The basic method consists of mapping points in the Cartesian space (x, y) to sinusoidal curves in (ρ, θ) space via the transformation ρ=x cos θ + ysin θ.

Each time a sinusoidal curve intersects another particular value of ρ and θ, the likelihood increases that a line corresponding to that (ρ, θ) coordinates value is present in the original image. An accumulator array is used to count the number of intersect the various ρ and θ values. The skew is then determined by the θ values corresponding to the highest number of counts in the accumulator array. In [7] use bottom pixels of the candidate objects within a selected region for Hough transformation. The hierarchical Hough transformation technique is also adopted in another paper [15]. The main idea of the above methods is to reduce the amount of Input data, but their computational complexities are still very high. In [9] they proposed an improved method to overcome the drawback of the method [7].

The cross-correlation method proposed by [14] is based on the correlation between two vertical lines in a document image. Since the pixels in the two parallel lines are translated due to skew, correlation matrix can be produced, it is defined as

$$R(x_o, s) = \sum_y L_1(x_o, y) L_2(x_o + d, y + s) \quad (1)$$

(here, L1(x, y) and L2(x, y) denote the two parallel vertical lines, respectively, d is the space between L1 and L2and s is the maximum translation . though this technique can result in accuracy, it is time consuming to calculate the correlation matrix and its projection profile. Furthermore, in certain situations parameter d should be changed and backtracked, which may increase the computing cost.

The horizontal (vertical) projection profile [6] is a histogram of the number of black pixels along horizontal (vertical) scan lines. For a script with horizontal text lines the horizontal projection profile will have peaks at text lines positions and troughs at positions in between successive text lines. To determine the skew of a document, the projection profile is computed at a number of angles and for each angle a measure of difference peak and trough height made. The maximum difference corresponds to the best alignment with the text line direction which, determine the skew angle. In [1] they have described an approach where the document is partitioned into vertical strips. The horizontal projection profiles are calculated for each strip and from the correlation of the profiles of the neighboring strips, the skew angle is determined. Although the proposed method is computationally inexpensive, it works well if the document is skewed within $\pm 10^0$.

Method proposed by [10] is based on vertical projection profile of horizontal strips which works well if the skew angle is small.

The method proposed by [11] belongs to Fourier Transformation approach. According to their method, the

Manuscript received on 8th April 2006.

Manjunath Aradhya V N is with the Department of Studies in Computer Science, University of Mysore, Mysore, INDIA. Ph: +91-9886896108, *Email: mukesh_mysore@rediffmail.com

Hemantha Kumar G is with the Department of Studies in Computer Science, University of Mysore, Mysore, INDIA.

Shivakumara P is with Dept. of Computer Science, National University of Singapore, Singapore.

skew angle of a document image corresponding to the direction where the density of Fourier space becomes the largest. However its computing complexity is very high.

In [5] proposed nearest neighbor clustering to skew detection. He found all the connected components in the documents and for each component computed the direction of its nearest neighbor. A histogram of the direction angle are computed, the peak of which indicates the document skew angle. In [8] the K nearest centers of the successive connected components is selected to calculate the vector directions between random pairs. The histogram peak corresponding to the skew angle of the whole document image will be generated afterwards. This approach achieves high accuracy, yet with high computing complexity O (N^2) (here N is the number of connected component).

Method by [16] proposed a nearest neighbor chain based approach to skew estimation in document images. Size restriction is introduced to the detection of nearest neighbor pairs. Then the chains with a largest possible number of nearest neighbor pairs are selected and their slopes are computed to give the skew angle of document image.

Method by [3] proposed skew detection and correction in document images based on straight–line fitting. The bottom center of the bounding box of a connected component is regarded as an eigen-point. According to the relations between the successive eigen-points in every text line, the eigen-points laid on the baseline are extracted as sample points. Then these samples are adopted by the least squares method to calculate the baseline direction.

In order to reduce the computing cost and to gain high accuracy of the above Hough transform based approaches, a fast skew angle detection method is proposed in this paper. The proposed method considers some of the selected character present in the document. The selected component is blocked and thinning is performed to the region. The obtained thinned points of the document region are then applied to Hough transform to estimate skew angle accurately. Experiments prove that through the proposed approaches the speed can be improved and the higher accuracy can be achieved.

The organization of the paper is as follows. We present proposed methodologies and their algorithm in section 2. Experimental results are reported in section 3. We give a comparative study of the proposed method with the well-known existing methods in section 4. Discussion based on experimental results is given in section 5. Conclusion is given at the end.

II. Proposed Methodlogy

This section presents the proposed methodology that is based on thinning and Hough transform (HT) to determine the skew angle accurately. The method has two stages. In the first stage, selected characters from the document image are blocked and thinning is performed over the blocked region. In the second stage, the thinned coordinates are fed to Hough transform (HT) to estimate the skew angle accurately. The block diagram of the proposed methodology is given in Fig. 1.

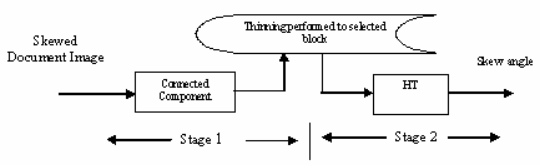


Fig. 1 block diagram of the proposed model

All of the connected components in a document image are detected by connected component analysis algorithm. For a component c_i, its centroid is represented by (x_{ci}, y_{ci}), the upper – left and bottom right coordinates of the rectangles enclosing the component is represented by (x_{li}, y_{ti}) and (x_{ri}, y_{bi}) respectively, and its height and width is represented by using h_{ci} and w_{ci} respectively. The average height (AH) of the bounding box is found and only the components with bounding box height less than AH are considered. The uppercase characters, Numerals and characters like b,d,f,g,h,j,k,l,p,q,t,y, do not participate in our algorithm as their heights are more than Average height. Also we have debarred those components whose box height is very small so that dots of the character like i and j, punctuation marks like full stop, comma, hyphen etc., are removed. For example, see Fig. 3.

Block the selected component present in the document image as shown in the Fig. 4. Apply thinning algorithm to each block present in the document image. Here we used [17] thinning algorithm to thin the component block. The unique feature that distinguishes the thinning system is that it thins symbols to their central lines. This means that the shape of the symbols is preserved. It also means that the method is rotation invariant. The system has 20 rules in its inference engine. These rules are applied simultaneously to each pixel in the image. Therefore, the system has the advantages of symmetrical thinning and speed. The system is very efficient in preserving the topology of symbols and letters written in any language. Fig.6 shows the result of thinning over the blocked component.

Hence to improve the accuracy and reduce the computing time, the parallel straight lines present in the Fig. 7, which is marked with circle is removed. Thus obtained points from the Fig. 8 are then subjected to Hough transform to estimate the skew angle accurately. Detailed Hough transform is explained below.

A. Hough Transform

Hough Transform technique is an approach used for fitting lines and curves. This approach is preferred when the objective is to find lines or curves formed by groups of individual points on an image plane. The method involves a transformation from an image plane to a parameter space.

Consider the case in which lines are the objects of interest. The line is expressed as $\rho = X\cos\theta + Y\sin\theta$. There are two line parameters namely, the distance (ρ) and the angle (θ) which defines transformation space. Each coordinate (x, y) of ON pixel in the image plane is mapped onto the locations in the transformed plane for all possible straight lines. This is depicted in Fig. 12 and 13. For all possible values of ρ and θ the transformations intersect at the same point on the

transformed plane when multiple points are collinear. Therefore, the point (ρ, θ), which has the greatest accumulation of mapped points, indicates lines with these parameters. In practice, due to discretization error and noise, points mapped will not be exactly collinear. Thus the points do not map on to exactly the same location on the transformed plane. For connected lines or positions of lines, computations can be reduced greatly by considering not all (ρ, θ) points but only those (ρ, θ) points that are in one orientation as indicated by the angle. The HT has a limitation that it does not give the coordinates of end points of the line, and further the long lines are favored over short lines. For these reasons, spatial domain method is considered to be often faster and more effective for skew detection and estimation.

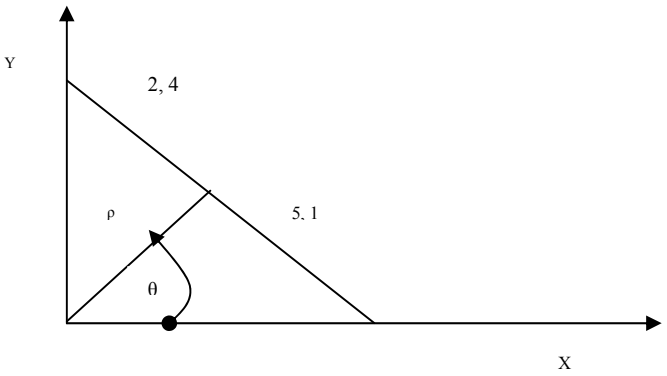


Fig. 12 X, Y plane for line detection

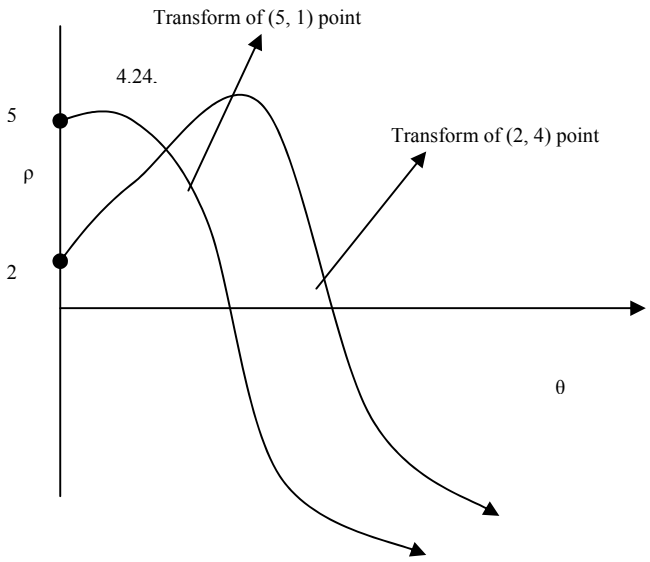


Fig.13 Sinusoidal Curves for Corresponding Points in Fig. 12

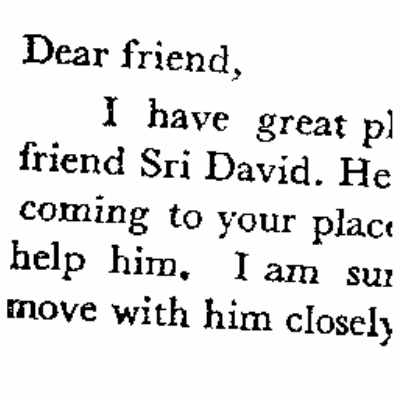


Fig. 2 English skewed text

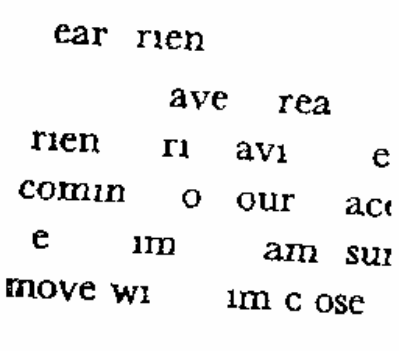


Fig. 3 Components selected from Fig. 2

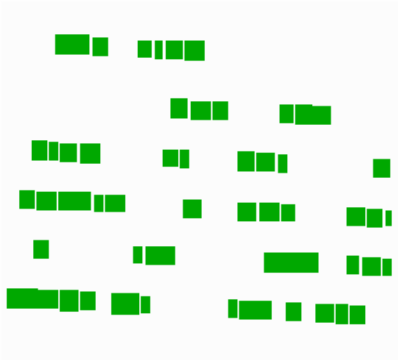


Fig. 4 Selected components are blocked

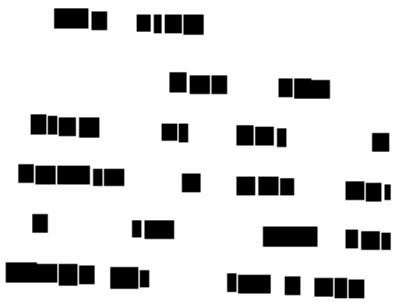


Fig. 5 Converted to Monochrome

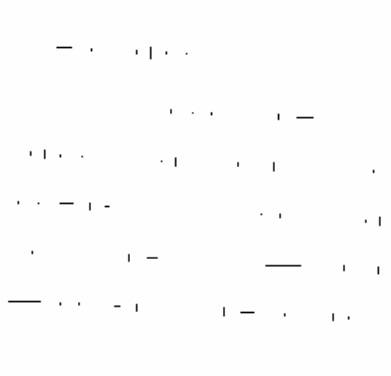


Fig. 6 Thinned result performed over the block

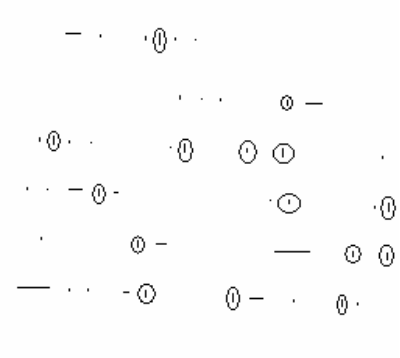


Fig. 7 Parallel Straight line points are circled

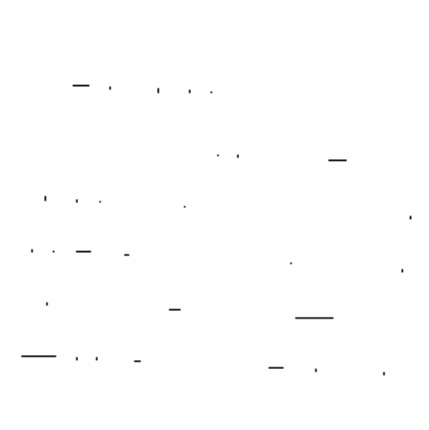


Fig. 8 Final thinned points that are subjected to HT

The detailed algorithm of the proposed method is shown below:

Algorithm SKEW

Step1: Find connected components in the document image and compute average bounding height (AH).

Step2: Select those connected components whose height is less than AH and remove very small connected components so that the dots of the character i,j, punctuation marks like full stop, comma, hyphen etc. are deleted.

Step3: Block the selected component present in the document.

Step4: Perform thinning operation over the selected block region.

Step5: Remove the parallel straight lines using prespecified threshold.

Step6: The obtained points are then subjected to Hough transform to estimate skew angle accurately.

Step7: Stop.

III. EXPERIMENTAL RESULTS

This section presents the results of the experiments conducted to study the performance of the proposed method. The method has been implemented in the C language on a Pentium IV 1.4 GHZ. We have considered different skewed documents from different sources like journals, textbooks, newspapers and the like. For experimentation purpose more than 200 documents are considered, samples of text documents are shown in Fig. 9. Obtained Mean Skew Angle (M), Standard Deviation (SD) and Mean Computing Time (CT) taken by the proposed methodology for these documents are reported in Table 1. To further establish the suitability of our method for document analysis and understanding purpose, documents with noise, documents with different resolutions, different textbook documents and the like are also considered.

TABLE I
ESTIMATED SKEW ANGLE, STANDARD DEVIATION AND MEAN TIME TAKEN BY THE PROPOSED METHOD

True angle	Proposed Method		
	Mean	Standard deviation	Mean Time (Secs)
3	2.89	0.346	1.78
5	5.08	0.211	1.68
10	9.86	0.487	1.77
15	15.12	0.11	1.70
20	20.11	0.231	1.80
30	30.0	0.147	1.71
40	39.89	0.365	1.72
45	44.98	0.220	1.74

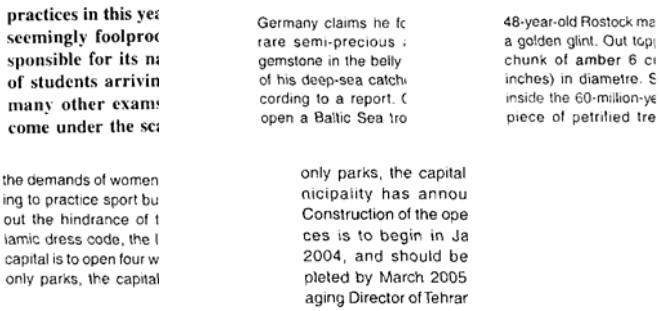


Fig. 9 Different samples scanned text document images

IV. COMPARATIVE STUDY

A comparative study with certain existing methods is carried out to establish the superiority of our method in terms of accuracy and efficiency. The scanned text document image , shown in Fig. 9 (a) to Fig. 9 (e) with different skew angles say 3, 5, 10, 15, 20, 30, 40, and 45 degrees is actually considered as an input to the proposed method as well as to the existing methods. The mean, standard deviation and computing time obtained using the proposed method and the other methods are reported in Table 2 (a) and Table 2 (b).

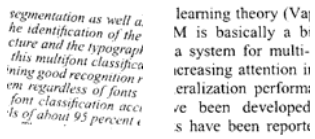
It is observed from Table 2 (a) and Table (b) that the skew angle estimation done by proposed methodology is better than the existing methods Akiyama and Hagita (1990), Pavlidis and Zhau (1992), Hashizume et al, (1986), Srihari and Govindaraju (1989), Le et al, (1994) and Pal and Chaudhuri (1996), Yan, (1993), Lu and Tan, (2003) and Cao Y et al (2003) with respect to mean (M). It is also observed that the proposed method is precise too with respect to the standard deviation that varies from 0.1 to 0.4. The variations in standard deviation (SD) of the other methods namely Akiyama and Hagita (1990) is (0.99 to 2.86), Pavlidis and Zhau (1992) is (0.76 to 3.25), Hashizume et al, (1986) is (0.5 to 1.5), Srihari and Govindaraju (1989) is (0.3 to 1.45), Le et al, (1994) is (0.35 to 0.85), Pal and Chaudhuri (1996) is (0.3 to 1.42), Yan, (1993) is (0.63-1.04), Lu and Tan (2003) is (051-0.95), and Cao Y et al (2003) is (0.2-0.76), which reveals that the proposed method is precise.

The above discussion revealed that the proposed method is precise and efficient compared to the existing methods.

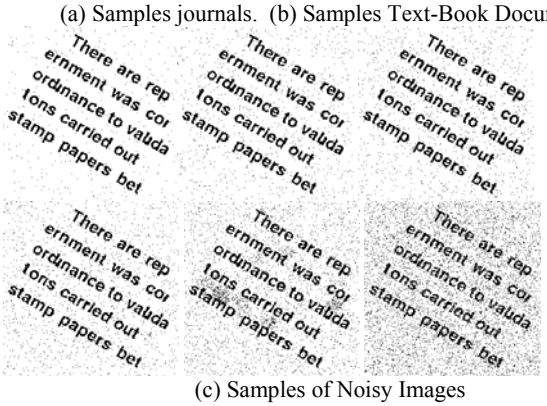
V. DISCUSSION

The simplicity, generality, superiority and applicability of the proposed method considering special cases comprising of different document images namely language documents, journals, document with different resolutions, synthetic text documents, newspaper cuttings, and noisy documents, is discussed in this section.

Experiments are conducted on different types of documents as shown in Fig. 10 (a) to Fig. 10 (c) and the computed skew angles using the proposed methods as well as the existing methods are given in Table 3.



(a) Samples journals. (b) Samples Text-Book Documents



(c) Samples of Noisy Images

Fig. 10 Sample images of Different document images

An experiment has also been conducted on different resolution documents and the performance of the proposed method does not degrade even for different resolution document. The computed skew angle using the proposed method and other methods is given in Table 4.

However the proposed method fails if the noise density (Salt and Pepper) increases by 0.05. To illustrate this we have conducted experiments and based on experimental results we have tabulated the values in Table 5 and the graphical representation is shown in Fig. 11. From Table 5 and Fig. 11, it is noticed that the proposed method works for the noise documents up to 0.05 level densities but not beyond.

TABLE V
PERFORMANCE OF THE PROPOSED METHOD WITH DIFFERENT NOISE DENSITIES

Noise Density	Known Angle	Computed Angle	Performance (%)	Remarks
0.01	30	29.9	99.66	Algorithm Works
0.02	30	30.12	99.60	
0.03	30	29.94	99.80	
0.04	30	29.85	99.50	
0.05	30	30.16	99.45	
0.06	30	34.64	83.33	Performance Degrades
0.07	30	21.69	73.01	

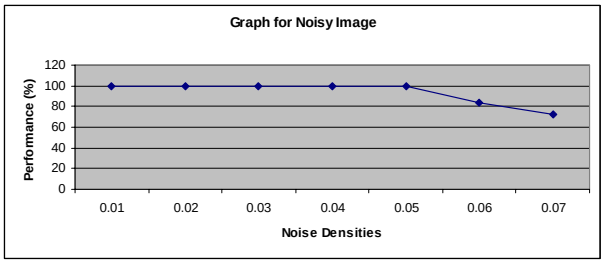


Fig 11 showing the graph for noisy images (30^0 documents)

VI. CONCLUSION

In summary, an efficient, novel and accurate methodology to estimate skew angle is presented in this paper. The proposed methods work based on thinning and Hough transform. Maher and Ward thinning algorithm is used to thin the blocked component. The proposed method is fast compared to other HT based methods. We have shown that the proposed method is superior with respect to accuracy, computational time and suitability. However, all the methods including the proposed method fail for document images containing text with picture. The authors are working with documents containing text with picture to make the proposed algorithm more generic and robust.

REFERENCES

[1] Akiyama T and Hagita N, Automated entry system for printed documents, Pattern Recognition, Vol. 23, No. 11, 1990, pp 1141-1158.
[2] Baird H.S, The Skew Angle of Printed Documents, Proceedings of Conference Society of Photographic Scientists and Engineers, Rocherster, New York, 1987, pp 14-21.
[3] Cao Yang, Shuhua Wang, Li Heng., Skew detection and correction in document images based on straight-line fitting, Pattern Recognition Letters, 24, pp 1871-1879, 2003.
[4] Gonzales R.C and Woods R.E, Digital Image Processing, 2nd ed., Pearson Education Asia, 2002.
[5] Hashizume A Yeh P S and Cosenfeld A, A Method of Detecting the Orientation of Aligned Components, Pattern Recognition Letters, Vol. 4, April 1986, pp 125-132.
[6] Hou H.S., Digital Document Processing, Wisely New York, 1983.
[7] Le D S, Thoma G R and Wechsler H, Automatic page orientation and skew angle detection for binary document images. Pattern Recognition 27, 1994, pp 1325-1344.
[8] O'Gorman L, The document spectrum for page layout analysis, IEEE Transactions on Pattern Analysis and machine Intelligence, No 15, vol 11, 1993, pp. 1162-1173.

[9] Pal U and Chaudhuri B. B, An Improved document skew angle estimation technique, Pattern Recognition Letters, Vol. 17, 1996, pp 899-904.
[10] Pavlidis T and Zhou J, Page segmentation by white streams, Proceedings of first International Conference on Document Analysis and Recognition (ICDAR), France, September 30, October 2, 1991, pp 945-953.
[11] Postl W, Detection of linear oblique structures and skew scan in digitized documents. Proceedings 8th International Conference on Pattern Recognition, 1986, pp. 687-689.
[12] Postl W, Detection of linear oblique structures and skew scan in digitized documents. Proceedings 8th International Conference on Pattern Recognition, 1986, pp. 687-689.
[13] Srihari SN and Govindaraju V, Analysis of textual images using the Hough Transform, Machine Vision and Applications, vol 2, 1989, pp. 141-153.
[14] Yan, H. Skew correction of document images using interline cross-correlation, Computer Vision, Graphics, and Image Processing 55, 1993, pp 538-543.
[15] Yu, B., Jain, A.K., A robust and fast skew detection algorithm for generic documents, Pattern Recognition 29 (10), pp 1599-1629, 1996.
[16] Yue Lu and Chew Lim Tan, A nearest neighbor chain based approach to skew estimation in document images, Pattern Recognition Letters 24, 2003, pp 2315-2323.
[17] M. Ahmed and R. Ward, (2002), Rotation Invariant Rule-Based Thinning Algorithm for Character Recognition, IEEE. Trans. Pattern Analysis and Machine Intelligence, vol. 24, No. 12, December 2002.

TABLE II (A)
COMPUTED MEAN, STANDARD DEVIATION AND COMPUTING TIME USING PROPOSED METHOD AND THE OTHER METHODS TO MEASURE ACCURACY, CONSISTENCY AND EFFICIENCY

True angle in degrees	Akiyama & Hagita (1990)			Pavlidis & Zhau (1992)			Hashizume et al (1986)			Srihari and Govindaraju (1989)			Le et al. (1994)		
	M	SD	CT (s)	M	SD	CT (s)	M	SD	CT (s)	M	SD	CT (s)	M	SD	CT (s)
3	7.3	2.86	1.16	3.15	0.761	1.28	3.54	0.545	2.2	3.2	0.3	13.66	3.150	0.365	2.32
5	8.04	2.12	1.16	5.28	1.543	1.28	4.8	0.649	2.15	5.59	0.548	14.06	5.35	0.426	2.31
10	12.19	1.761	1.16	11.2	1.949	1.28	8.05	1.246	1.83	12.88	1.456	13.13	10.125	0.416	2.21
15	15.94	1.44	1.16	14.76	2.135	1.28	15.82	1.14	1.95	15.72	0.826	11.19	15.18	0.496	2.42
20	22.6	1.562	1.16	19.14	3.13	1.28	20.78	0.901	1.96	18.9	0.749	12.04	20.17	0.398	2.10
30	32.7	1.769	1.16	27.5	2.256	1.28	27.8	1.469	2.03	30.02	0.994	11.36	30.5	0.856	2.61
40	40.2	2.02	1.16	36.2	2.866	1.28	43.19	1.558	2.01	40.9	0.618	11.17	40.21	0.358	2.68
45	45.9	0.99	1.16	46.16	3.257	1.28	45.95	0.6449	2.03	45.3	0.462	10.5	45.15	0.926	2.59

TABLE II (B)
CONTINUATION TO TABLE II (A)

True angles in degree	Pal and Chaudhari (1996)			Yan (1993)			Lu and Tan (2003)			Cao Y et al (2003)			Proposed Method		
	M	SD	CT (s)	**M**	**SD**	**CT** (s)	**M**	**SD**	**CT** (s)	M	SD	CT(s)	M	SD	CT(s)
3	3.128	0.348	2.12	3.439	0.9488	2.49	3.868	0.819	2.28	3.21	0.51	1.89	2.89	0.346	1.78
5	5.68	0.451	1.98	5.0955	1.0401	2.48	5.7124	0.958	2.31	5.52	0.68	1.78	5.08	0.211	1.68
10	10.17	0.425	2.38	10.191	0.8201	2.492	10.7363	0.798	2.22	9.86	0.54	1.90	9.86	0.487	1.77
15	15.72	0.51	2.12	15.358	0.936	2.501	15.5330	0.517	2.29	15.02	0.32	1.88	15.12	0.11	1.70
20	20.11	0.722	2.38	20.439	0.8376	2.685	20.97	0.801	2.58	19.68	0.76	1.85	20.11	0.231	1.80
30	30.32	1.42	2.18	30.847	0.9717	2.48	30.174	0.521	2.58	30.54	0.46	1.92	30.0	0.147	1.71
40	40.42	0.719	2.17	40.382	0.846	2.369	39.474	0.664	2.40	40.12	0.26	1.90	39.89	0.365	1.72
45	45.2	0.86	2.32	44.422	0.6399	2.589	45.927	0.530	2.69	44.56	0.66	1.91	44.98	0.220	1.74

TABLE III
COMPUTED MEAN SKEW ANGLE FOR DIFFERENT DOCUMENTS WITH 10^0 TRUE SKEW ANGLES

Cases	Existing methods									
	Akiyama & Hagita (1990)	Pavlidis & Zhau (1992)	Hashizume et al (1986)	Srihari and Govindaraju (1989)	Le et al. (1994)	Pal and Chaudhari (1996)	Yan (1993)	Lu and Tan (2003)	Cao Y et al (2003)	Proposed Method
Different Language	39.11	43	24	11	11.5	11.5	7.63	10.963	11.02	10.54
Noise	27	31	0	11	9.5	9.5	11.921	9.236	8.98	9.84
Text Book	29	25	12.1	10.5	11	10.5	10.326	10.459	10.21	10.21
Journal	3.58	2.7	35	12	10.5	11	11.021	10.679	10.24	10.23
Text with Picture	71	81	0	45	1.7	38	16.796	14.719	12.42	13.56

TABLE IV
COMPUTED SKEW ANGLES USING PROPOSED AND EXISTING METHODS FOR DIFFERENT DPIS OF 23^0

Resolution	Existing methods									
	Akiyama & Hagita (1990)	Pavlidis & Zhau (1992)	Hashizume et al (1986)	Srihari and Govindaraju (1989)	Le et al. (1994)	Pal and Chaudhari (1996)	Yan (1993)	Lu and Tan (2003)	Cao Y et al (2003)	Proposed method
75	22.12	23	90	18.5	23	23.5	26.75	23.96	22.86	23.01
100	46	45	21	17.5	22.5	23.5	24.21	23.02	23.01	23.12
150	29	38	21	4.5	23	23	21.65	22.89	22.69	22.98
300	41	44	19	22.5	23	23	26.08	23.86	23.14	22.88
400	75	90	0	48	22.5	23.5	24.06	23.59	23.58	23.41