Engineering Letters, 14:1, EL_14_1_22 (Advance online publication: 12 February 2007)

Skew Estimation Technique for Binary Document Images based on Thinning and Moments

Manjunath Aradhya V N, Hemantha Kumar G, Shivakumara P

***Abstract* -When a document is fed to a scanner either mechanically or by a human operator for digitization, it suffers from some degrees of skew or tilt. Skew angle detection is an important component of any Optical Character Recognition (OCR) and Document Analysis System. In this paper we present a novel skew estimation technique for document images. The method uses Boundary Growing (BG) technique, Thinning, and Moments for skew angle detection. Experimental results are presented and compared with results on several other skew detection methods.**
***Keywords:* Optical Character Recognition, Skew angle, BG, Thinning, Moments, 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 for the blind.

Over the past two decades, many people have attempted to solve the problem of skewed documents using different methods. Applying Hough Transform to a set of points to determine straight lines in the image has been a popular approach. [Srihari and Govindaraju, 1989] use the Hough transform on all black pixels, while Hinds et al, 1990 reduce the data with the use of horizontal and vertical run length computations. [Le et al, 1994] apply this transform only to the bottom pixels of a connected component. [Baird, 1987] uses connected component analysis and creates a projection profile using a single point to represent each connected component. Postl, 1986 briefly discusses method which maximizes a given function, called the premium, with respect to the angle by simulating scan lines. [Akiyama and Hagita, 1990] describes a method that is extremely fast, where the document is segmented into vertical partitions, and projection profiles are created for each partition. The skew angle is then calculated by determining the average shift in zero crossings between partitions. [Hashizume et al, 1986] propose a method that computes the nearest neighbor of each character and creates a histogram of these values to detect the skew angle. [Liu et al, 1992] use a similar technique to [Hashizume et al, 1986] in which they detect and remove the characters with ascenders or descenders and then calculate the inclination between adjacent base points of connected components. [O'Gorman, 1993] discusses skew correction in the framework of his 'docstrum' analysis. Here, the skew angle is determined by finding the nearest neighbors for a connected component and determining the most common angle between neighboring connected components. [Smith, 1995] discusses an algorithm that groups the connected components into lines, moving from left to right across the document, and uses a continually in the correct line. Method by Cao Y et al (2003) 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 this paper, we describe a new method for the detection of skew angle based on BG, Thinning, and Moments. The organization of the paper is as follows. Proposed methodologies are given in section 2. Experimental results are reported in section 3. A comparative study of the proposed method with the well-known existing methods is presented in section 4. Discussion based on experimental results is given in section 5. Conclusion is reached in section 6.

II. Proposed Methodolgy

This section presents the proposed methodology that is based on boundary growing and moments to determine the skew angle. The method has two stages. In the first stage, the text line form by the document is extracted using Boundary Growing method (BGM). Resultant text line extracted from BGM is then applied to Thinning. In the next stage obtained thinned line is then passed onto moments to estimate skew angle. The block diagram of the proposed methodology is given in Fig. 1. In the following subsections, each stage is explained in detail.

A. Boundary Growing Method (BGM)

The method begins with fixing the boundaries (Fig. 5) for each character in the skewed text document. The boundary for each character is fixed by scanning the document column-wise and the centroid of bounding box of a character is computed when the character is found. The boundary of a character is expanded row-wise as well as column-wise until it reaches the pixel of the neighbor-character. The procedure of centroid computation of bounding box of a character and boundary expansion continues till the end of the text line. The end of the text line is decided based on the space between the characters and the words. The height (H) of each boundary of characters in the text line is computed. This procedure is repeated for all the text lines present in the document.

Manuscript received on 26th 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.

B. Thinning

Thinning is the process of reducing thickness of each line of pattern to just a single pixel. Thinning is usually used as the first step in applications such as optical character recognition to improve the recognition rate. Here a new improved 4-step thinning algorithm is proposed. Proposed algorithm is iterative; at each iteration proposed algorithm deletes every point that lies on the outer boundaries of the symbol, as long as the width of the symbol is more than one pixel wide. Region points are assumed as 1 and background points to have value 0. The proposed method is improved version of 2-step method (Rafeael and Richard, (2002)). Fig 2 shows the 8-neighborhood of the pixel. In 2-step method (Rafeael and Richard, (2002)) 8-neighborhood of a pixel is used to thin a character image. The proposed method uses 4-step to thin a given image. First, the number of nonzero neighbors of p0 is calculated as N (p0) = p1+ p2 + p3 + p4 + p5 + p6 + p7 + p8. Transition say, T (p0) is calculated as the number of 0-1 transition in the order sequence p1, p2, p3, p4, p5, p6, p7, p8 respectively. Finding number of nonzero neighbors and 0-1 transition using 8-neighborhood, the proposed method uses 4 steps to complete iteration. In every iteration the proposed method deletes on all 4-direction i.e., p2, p4, p6, p8. The iterations are repeated until no further changes occur. The 4-step of the proposed method is shown in the following algorithm.

p1	p2	p3
p8	p0	p4
p7	p6	p5

Fig 2 showing 8-neighborhood of the pixel

Algorithm for thinning
Input: Any Symbols
Output: Thinned Image
Method:

Step 1: If 2 < =N (p0) < = 8
If T (p0) = 1
If p2 = 0
If p3 = 0
If the above condition is satisfied, make p (0) =0.

Step 2: If 2 < =N (p0) < =8
If T (p0) = 1
If p4 = 0
If p5 = 0
If the above condition is satisfied, make p (0) =0.

Step 3: If 2 < =N (p0) < = 8
If T (p0) =1
If p6 = 0
If p7 = 0
If the above condition is satisfied, make p (0) =0.

Step 4: If 2 < =N (p0) < = 8
If T (p0) =1
If p8 = 0
If p1 = 0
If the above condition is satisfied, make p (0) =0.

Step 5: Repeat the above procedure for all the region points.
Step 6: Stop.

The proposed thinning algorithm is also compared with [Maher and Ward, 2002 and Gonzalez and Woods, 2002] methods. The proposed method preserves the shape of the given image even after rotation. Fig 3(a) to 3 (d) shows the results obtained from the proposed method and existing methods. From Fig 3 it is clear that the proposed method preserves shape of the symbol and preserves the topology even after rotation.

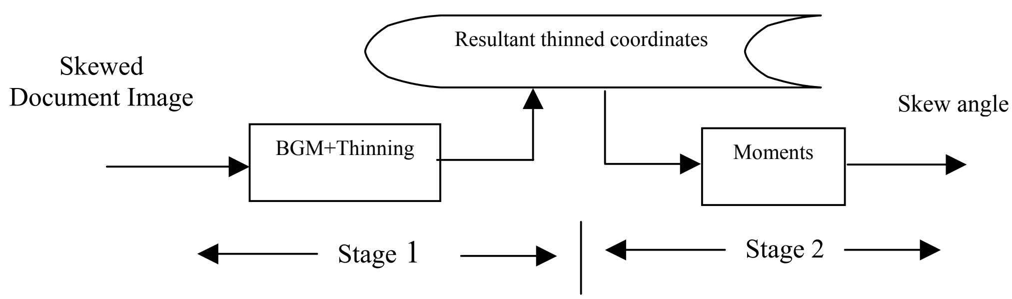


Fig. 1 Block diagram of the proposed model for skew angle estimation

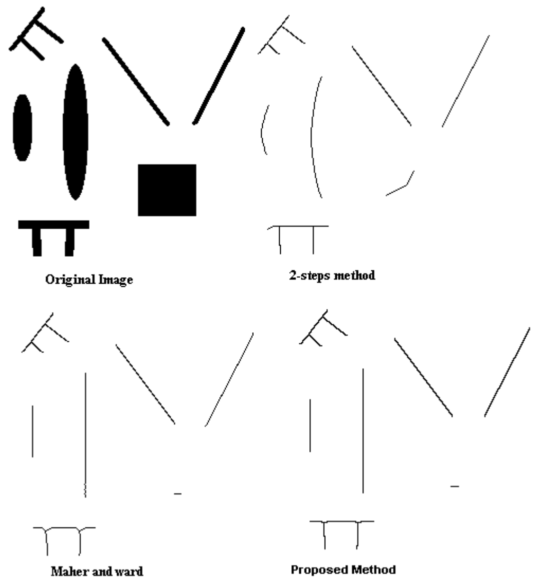


Fig 3 (a)

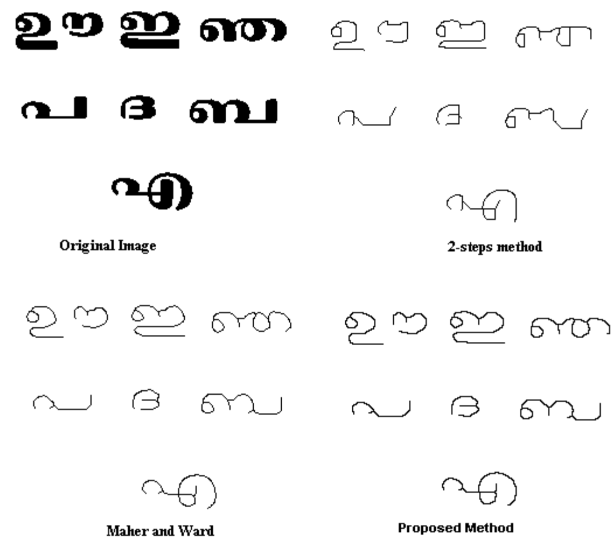


Fig 3(b)

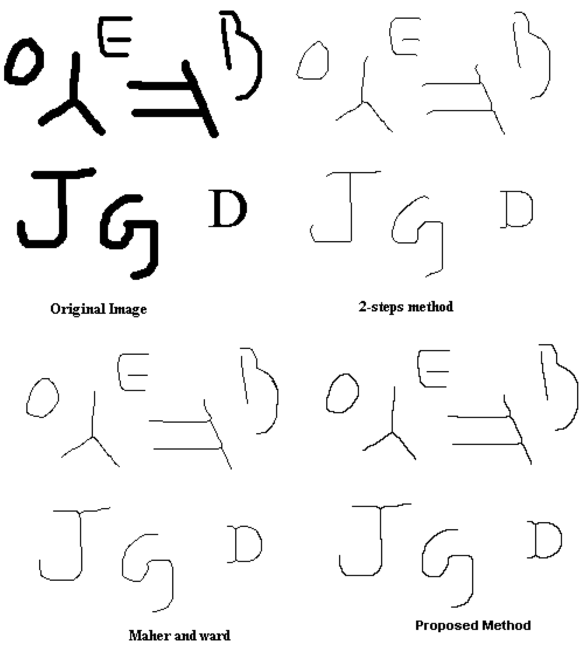


Fig 3 (c)

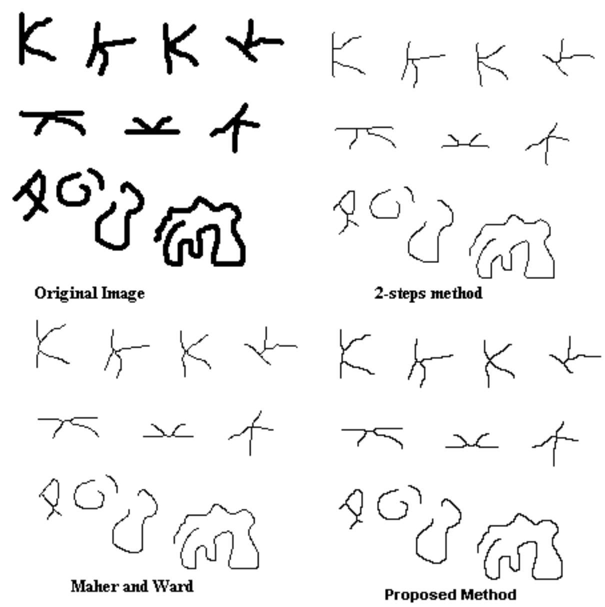


Fig 3(d)

Fig. 3 showing the resultant thinned image for different images

C. Moments for finding the Inclination

In this section, we present moments based method for computing skew angle using extracted coordinates. The Moments are computed using equation 1, where f(x, y) is pixel values in the 2D image, x and y are the coordinates of the pixels, (p + q) defines the order of the moments. For continuous 2D image, equation of the moments is given in equation 1. For p, q=0, 1 …. For discrete 2D image the equation 1 becomes equation 2. Further, for the binary image we require equation 3.

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y)\,dxdy \quad \text{……………………. (1)}$$

$$m_{pq} = \sum_{0}^{n} \sum_{0}^{n} x^p y^q f(x,y) \quad \text{……………………….. (2)}$$

$$m_{pq} = \sum_{0}^{n} \sum_{0}^{n} x^p y^q \ \ldots\ldots\ldots\ldots \quad (3)$$

Where n * n is the size of the image.

From equation 1 we have obtained central moments and is given in equation 4 by computing centroid (ref. equation 3). Equation 5 for discrete image, equation 6 for binary image since f(x, y) = 1 in this work.

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y)\,dxdy \ \ldots\ldots \quad (4)$$

$$\mu_{pq} = \sum_{0}^{n} \sum_{0}^{n} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \ \ldots\ldots \quad (5)$$

$$\mu_{pq} = \sum_{0}^{n} \sum_{0}^{n} (x-\bar{x})^p (y-\bar{y})^q \ \ldots\ldots \quad (6)$$

$$m_{10} = \sum_{0}^{n} \sum_{0}^{n} x^1 y^0 \ \ldots\ldots \quad (7)$$

$$m_{101} = \sum_{0}^{n} \sum_{0}^{n} x^0 y^1 \ \ldots\ldots \quad (8)$$

$$m_{00} = \sum_{0}^{n} \sum_{0}^{n} x^0 y^0 \ \ldots\ldots \quad (9)$$

Using equation 7, 8 and 9 we compute the equation 10 and 11.

$$\bar{x} = \frac{m_{10}}{m_{00}} \ \ldots\ldots \quad (10)$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \ \ldots\ldots \quad (11)$$

$$\mu_{11} = \sum_{0}^{n} \sum_{0}^{n} (x-\bar{x})^1 (y-\bar{y})^1 = m_{11} - \frac{m_{10} m_{01}}{m_{00}}$$

$$\mu_{11} = m_{11} - \bar{x} m_{01}$$

$$\mu_{11} = m_{11} - \bar{y} m_{10} \ \ldots\ldots \quad (12)$$

$$\mu_{20} = \sum_{0}^{n} \sum_{0}^{n} (x-\bar{x})^2 (y-\bar{y})^0 = m_{20} - \frac{2m_{10}^{\ 2}}{m_{00}} + \frac{m^2{}_{10}}{m_{00}}$$

$$\mu_{20} = m_{20} - \frac{m_{10}^{\ 2}}{m_{00}}$$

$$\mu_{20} = m_{20} - \bar{x} m_{10} \ \ldots\ldots \quad (13)$$

$$\mu_{02} = \sum_{0}^{n} \sum_{0}^{n} (x-\bar{x})^0 (y-\bar{y})^2 = m_{02} - \frac{m_{01}^{\ 2}}{m_{00}}$$

$$\mu_{02} = m_{02} - \bar{y} m_{01} \ \ldots\ldots \quad (14)$$

The theta (Skew angle) is computed using equation 12, 13 and 14 and it results in equation 15.

$$\theta = \frac{1}{2} \tan^{-1}\left(\frac{2\mu_{11}}{(\mu_{20} - \mu_{02})} \right) \ \ldots\ldots \quad (15)$$

Equation 15 gives the θ for the set of points.

The skew angle is estimated using equation 15. Finally the average skew angle is computed for four or five text line using the equation 16.

Average skew angle

$$= \frac{\sum_{i=1}^{p} \theta_i}{p} \ \ldots\ldots \quad (16)$$

Where θ is the skew angle each text line and p indicates that the number of text lines.

Fig. 4(a) - (e) show the working principle of the proposed method. Fig. 4(a) skewed document, Fig. 4(b) components selected from Fig. 4(a), dots, comma, and small connected components are removed. Fig. 4(c) shows the BGM applied and corresponding text line fetched and converted to monochrome (see Fig.4 (d)). Fig.4 (e) shows the thinning obtained using proposed thinning algorithm. The obtained resultant thinned image coordinates are then passed to Moments to find the inclination.

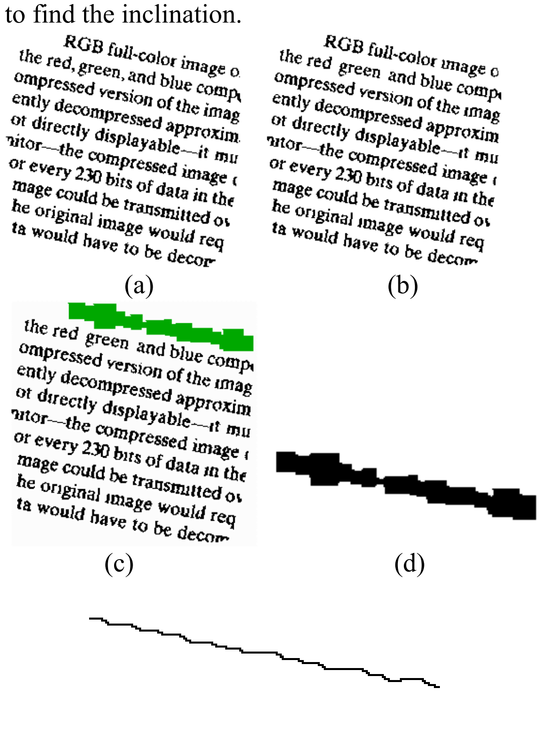


Fig. 4 showing working procedure of the proposed method

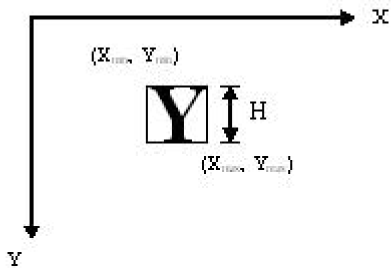


Fig. 5 Boundary fixed for character

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. 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 as shown in Fig 6.

Table 1 Estimated Skew angle, SD, and Mean Time taken by the proposed method

True Angles	Proposed Method		
	Mean	**SD**	**Time**
3	3.101	0.446	1.23
5	5.21	0.460	1.32
10	9.89	0.326	1.22
15	14.81	0.39	1.20
20	20.03	0.37	1.18
30	30.23	0.41	1.02
40	40.48	0.126	1.11
45	44.80	0.30	1.036

IV. COMPARITVE 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 images with different skew angles say 3, 5, 10, 15, 20, 30, 45 degrees is actually considered as an input to the proposed method as well as existing methods. The M, SD and CT obtained using the proposed method and other methods are listed in Table 1 for English documents.

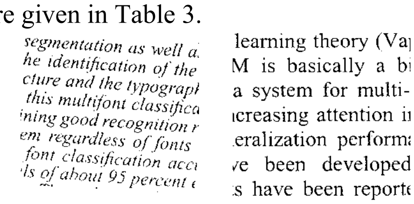
It is observed from Table 2(a) and Table 2 (b) that the skew angle estimation obtained by proposed method gives better results compared to the existing methods in terms of accuracy and standard deviation.

From Table 2(a) and Table 2 (b), it is also observed that the proposed method is consistent 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 Chaudhari, 1996] is (0.3 to 1.42), [Yan H, 1993] is (0.6 to 1.0), [Lu and Tan, 2003] is (0.5 to 0.9), and Cao Y et al (2003) is (0.2-0.76) which reveals that the proposed method is consistent.

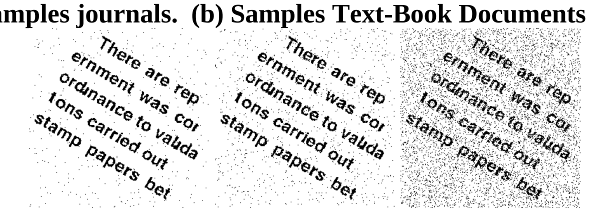
V. DISCUSSION

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. 6 (a) to Fig. 6 (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. 6 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. 7. From Table 5 and Fig. 7, it is noticed that the proposed method works for the noise documents up to 0.05 level densities but not beyond.

Table 5 Performance of the proposed method with different noise densities

Noise Density	Known Angle	Computed Angle	Performance (%)	Remarks
0.01	30	29.89	99.63	Algorithm Works
0.02	30	30.06	99.80	
0.03	30	30.12	99.60	
0.04	30	29.931	99.66	
0.05	30	29.84	99.46	
0.06	30	20.45	68.16	Performance Degrades
0.07	30	18.56	61.86	

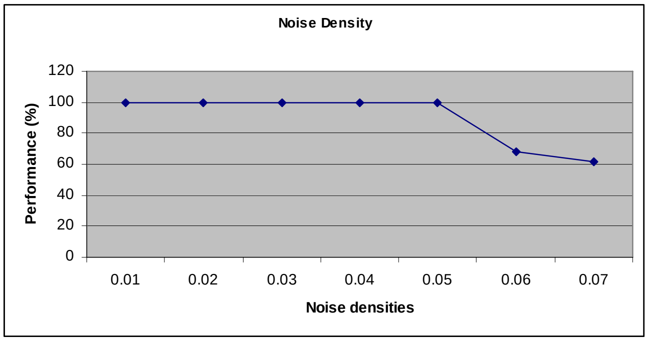


Fig 7 showing the graph for noisy images (30^0 documents)

VI. CONCLSUION

In summary, an efficient, novel and simple methodology to estimate skew angle is presented in this paper. The proposed methods work based on boundary growing, thinning and moments. Here a novel thinning algorithm is described and the proposed thinning algorithm preserves the shape of the symbol even after rotation. Also 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

Akiyama T and Hagita N, Automated entry system for printed documents, Pattern Recognition, Vol. 23, No. 11, 1990, pp 1141-1158.

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.

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.

Gonzalez Rafeael and Woods Richard E, Digital image Processing, Second Edition, Person Education, 2002.

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.

Hinds, S. C, Fisher J L and D'Amato D P, A Document Skew detection method using run-length encoding and the Hough transform. In Proc. International conference on Pattern Recognition, Volume I, 1990, pp 464-468.

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.

Liu j, Lee C.M, Shu R.B, An efficient method for the skew normalization of a document image. 11th International conference on Pattern Recognition, 1992, pp. 152-155.

M. Ahmed and R. Ward, A Rotation Invariant Rule-Based Thinning Algorithm for Character Recognition, *IEEE. Trans. Pattern Analysis and Machine Intelligence*, vol. 24, No. 12, December 2002.

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.

Pal U and Chaudhuri B. B, An Improved document skew angle estimation technique, Pattern Recognition Letters, Vol. 17, 1996, pp 899-904.

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.

Postl W, Detection of linear oblique structures and skew scan in digitized documents. Proceedings 8th International Conference on Pattern Recognition, 1986, pp. 687-689.

Smith R, A simple and efficient skew detection algorithm via text row accumulation, Proceedings 3rd International Conference on Document Analysis and Recognition, vol 2, 1995, pp. 1145-1148.

Srihari SN and Govindaraju V, Analysis of textual images using the Hough Transform, Machine Vision and Applications, vol 2, 1989, pp. 141-153.

Yan, H. Skew correction of document images using interline cross-correlation, Computer Vision, Graphics, and Image Processing 55, 1993, pp 538-543.

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.

Table 2 (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 2 (b) Continuation to Table 2 (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	3.101	0.446	1.23
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	5.21	0.460	1.32
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	9.89	0.326	1.22
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	14.81	0.39	1.20
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	20.03	0.37	1.18
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	30.23	0.41	1.02
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	40.48	0.126	1.11
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	44.80	0.30	1.036

Table 3 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	10.22
Noise	27	31	0	11	9.5	9.5	11.921	9.236	8.98	10.123
Text Book	29	25	12.1	10.5	11	10.5	10.326	10.459	10.21	10.002
Journal	3.58	2.7	35	12	10.5	11	11.021	10.679	10.24	10.312
Text with Picture	71	81	0	45	1.7	38	16.796	14.719	12.42	13.63

Table 4 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	22.96
150	29	38	21	4.5	23	23	21.65	22.89	22.69	23.24
300	41	44	19	22.5	23	23	26.08	23.86	23.14	23.51
400	75	90	0	48	22.5	23.5	24.06	23.59	23.58	23.02