A Hybrid Approach for Skew Detection and Correction in the Multi-script Scanned Document

M. Ramanan[1*]

[1]*Department of Computer Science, Faculty of Applied Science, Trincomalee Campus, Eastern University, Sri Lanka.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Original Research Article

ABSTRACT

Skew detection and correction of a scanned document is a very important step in Optical Character Recognition because skew of scanned document is reducing the accuracy of text line approach for skew detection and correction to calculate the skew angle on multi-script scanned document using Radon transform, Hough transform, Harries corner, Wiener filter and smearing algorithm. In this paper, a proposed approach is compared existing skew detection and correction techniques for printed documents having different scripts: English, Tamil, Sinhala and mixed-script. A proposed hybrid method is tested on 160 documents. The overall testing results is 90.62% for skew detection and correction.

1. INTRODUCTION

Skew detection and correction are important steps of document layout analysis and Optical Character Recognition (OCR) because skew of scanned document is reducing the accuracy of text line and character segmentation. The skew angle of a scanned document is defined as the

_____

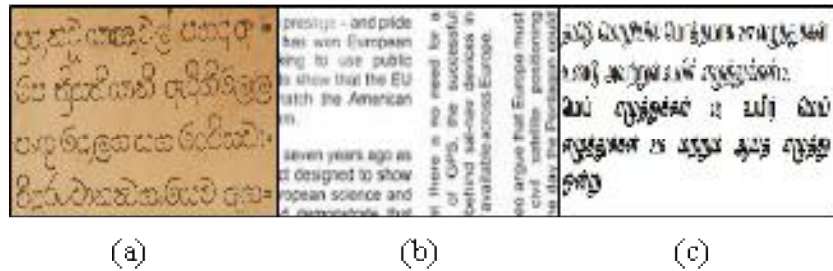*Corresponding author: E-mail: mramanan1981@gmail.com*

Fig. 1. skew on scanned document

variation of the dominant orientation of the text lines from the y-axis. When a document is fed to the scanner either mechanically or by a human intervention, a few degrees of skew is unavoidable.

The skew angle in a scanned document is divided into two types: Clock wise skew angle and Anti-clock wise skew angle.

There are three types of skew in the scanned document basis on number of skew angle and orientation: (1) Single Skew on scanned document, (2) multiple skew on scanned document, and (3) multiple skew in single text line

A. Single Skew on scanned document

In this skew, the document is skewed to single angle. Fig. 1(a) shows the sample single skew document

B. Multiple skew on scanned document

The document can have many sections; each may be skewed to different angles. Fig. 1(b) shows the sample multiple skew documents

C. Multiple skew in single text line

When documents contain several orientations in the single line is called as multiple skew in single text line. In multiple skew in single text line, there are several orientations in each line of the document. Fig. 1(c) shows the sample multiple skew in single text lines

Many techniques of skew detection can process text only, and good quality document images and single skew document successfully. But it is challenging problem to process documents with non-text document which consists of tables, charts, graphics, figures or different font sizes.

This study proposes a hybrid approach for skew detection and correction techniques using Radon transform, Hough transform, Harris corner, smearing technique and Wiener filter, and focus only single skew on scanned document; following this introductory section, the rest of the paper is organised as follows. Section 2 summarize the related work in skew detection and correction technique, whereas section 3 describes the background of Hough transform, radon transform and Harris corner detection method. The proposed methodology is presented in Section 4. Section 5 briefly describes the constructed dataset. Experimental setup and testing results are presented in section 6. Finally, section 7 concludes this paper with future extension.

2. RELATED WORK

In [1], Salem Saleh Bafjaish, et al. proposed Skew Detection and Correction of Mushaf Al-Quran Script using Hough transform. The proposed method involved in converting to grayscale image, binary image, foreground image detection, Hough transform to detect lines, calculate skew angle and rotate image. The technique using Hough transform lines detection for calculating the skew angulations. It works for different version of Mushaf AlQuran image pages which has skewed text zones. Moreover, it can detect and correct the skew angle in the range between 20 degrees.

In [2], Abdelhak Boukharouba proposed a new algorithm for skew correction and baseline detection based on the randomized hough transform. The proposed algorithm involved lower edge determination and segmentation, small edge filtering, skew angle detection, baseline estimation with skew correction and baseline estimation without skew correction. Authors have collected a variety of printed Arabic documents consisting of magazines, books and

newspaper. This algorithm can also contribute significantly for text line extraction from skewed document images for many languages.

In [3], Rubani and Jyoti Rani proposed the Skew Detection and Correction in Text Document Image using Projection Profile Technique. The proposed method works in two phases: calculate the skewed angle from the input text document, and correct the sleekness in the given text document on line by line basis. The proposed algorithm involved the binarization, line extraction, Convert the extracted lines into independent block using RLSA algorithm, calculate the corner coordinate points, and calculate the skew angle. The proposed system is tested on 30 text documents which are written in various languages: Punjabi, Hindi and English. The overall recognition rate is around 97%.

In [4], Ramandeep Kaur et al. proposed a hybrid approach to detect and correct a skew in scanned document images using fast fourier transform and nearest neighbor algorithm. The proposed method for Skew Detection and correction involved preprocessing, apply Faster Fourier Transform, apply nearest neighbor, calculate the skewed angle. The overall recognition rate of proposed method is 96%.

In [5], MSLB. Subrahmanyam et al. proposed a new algorithm for Skew Detection of Telugu Language Document based on Principle-axis Farthest Pairs Quadrilateral (PFPQ). The proposed system consists of preprocessing, derivation of PFPQ, parameters estimation on PFPQ, painting and directional smearing, final estimation. The proposed technique is independent on font size, line spacing between text lines etc.

There are many methods: Hough transform [6]-[8], Discrete Fourier Algorithm [9], Fast Fourier transform [10,11], Fast frequency domain [12], Nearest neighbor [4], Interline cross-correlation [13], Projection Profile [14], Principal Component Analysis [15,16], Radon transform [17], Hough transform and Harris corners [18], Radon transform and Wiener filter [19], and straight-line fitting [19] for skew estimation and correction.

3. BACKGROUND

A. Hough Transform

Standard Hough Transform (SHT) uses the parametric representation of a line:

$$rho = x*cos(theta) + y*sin(theta) \qquad (i)$$

The variable rho is the distance from the origin to the line along a vector perpendicular to the line. Theta is the angle of the perpendicular projection from the origin to the line measured in degrees clockwise from the positive x-axis. The range of theta is −90°≤θ<90°. The angle of the line itself is θ+90°, also measured clockwise with respect to the positive x-axis.
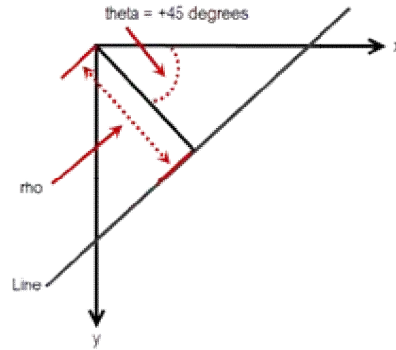
Fig. 2. Parametric representation of Standard Hough Transform (SHT)

The SHT is a parameter space matrix whose rows and columns correspond to rho and theta values respectively. The elements in the SHT represent accumulator cells. Initially, the value in each cell is zero. Then, for every non-background point in the image, rho is calculated for every theta. rho is rounded off to the nearest allowed row in SHT. That accumulator cell is incremented. At the end of this procedure, a value of Q in SHT(r,c) means that Q points in the xy-plane lie on the line specified by theta(c) and rho(r). Peak values in the SHT represent potential lines in the input image.

B. Radon Transform

The Radon transform of an image is the sum of the Radon transforms of each individual pixel. The algorithm first divides pixels in the image into four subpixels and projects each subpixel separately, as shown in the following figure.

Each sub pixel's contribution is proportionally splits into the two nearest bins, according to the distance between the projected location and the bin centers. If the sub pixel projection hits the center point of a bin, the bin on the axes gets the full value of the sub pixel, or one-fourth the value of the pixel. If the sub pixel projection hits the border between two bins, the sub pixel value is split evenly between the bins.

C. Harris Corner Detection Method

The Harris corner detection method [20] avoids the explicit computation of the eigenvalues of the sum of squared differences matrix by solving for the following corner metric matrix, R:

$$R = AB - C^2 - k(A + B)^2)$$ (ii)

Minimum Eigenvalue Method is more computationally expensive than the Harris corner detection algorithm because it directly calculates the eigenvalues of the sum of the squared difference matrix, M.
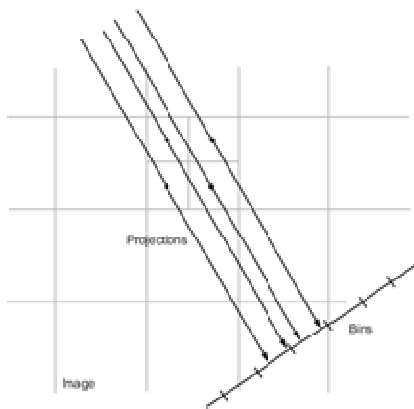
The sum of the squared difference matrix, M, is defined as follows: $M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$

The previous equation is based on the following values:

$$A = (I_x)^2 \otimes w$$
$$B = (I_y)^2 \otimes w$$
$$C = (I_x I_y)^2 \otimes w$$

Where I_x and I_y are the gradients of the input image, I, in the x and y direction, respectively. The $\otimes$ symbol denotes a convolution operation.

Use the Coefficients for separable smoothing filter parameter to define a vector of filter coefficients. The block multiplies this vector of coefficients by its transpose to create a matrix of filter coefficients, w.

Fig. 3. Parametric representation of random transform

The block calculates the smaller eigenvalue of the sum of the squared difference matrix. This minimum eigenvalue corresponds to the corner metric matrix.

The variable k corresponds to the sensitivity factor. You can specify its value using the Sensitivity factor (0<k<0.25) parameter. The smaller the value of k, the more likely it is that the algorithm can detect sharp corners.

Use the Coefficients for separable smoothing filter parameter to define a vector of filter coefficients. The block multiplies this vector of coefficients by its transpose to create a matrix of filter coefficients, w.

4. METHODOLOGY

In this section, the proposed skew detection and correction technique is described in details. The block diagram of skew detection and correction technique consists of various steps as shown in Fig. 4.

The following steps are applied to perform skew detection and correction:

Step 1: Input the printed scanned documents typed in single or multiple scripts.

Step 2: Convert the RGB image into Gray Scale Image.

Step 3: Convert the grayscale image into a Binary image- Binarization is the process of converting a gray scale image into binary image using Otsu's method.

Step 4: Detection of bounding box - Each binary image is enclosed in a tight fit rectangular boundary. The portion of the image outside this boundary is discarded using projection techniques.

Step 5: Divide an image into three blocks of equal size name as Block 1 (B1), Block 2 (B2) and Block 3 (B3) steps (see Fig.4).

Step 6: **Apply Radon transform** is the sum of each pixel of the edge map. The Radon function computes projections of Block 1 matrix along specified directions. The Radon transformation is applied to determine the largest line in the Block 1 (B1) and identifies most visible line on the Block 1 to estimate the angle of skewness (ӨB1).

Apply Hough transform: Hough transform is applied to either vertical or horizontal Block2

(B2). After that calculate the angle of skewness (ƟB2).

Apply Radon transform with smearing and Wiener filter

• Apply Wiener filter to the Block 3.
• Apply smearing algorithm
• The Radon transformation is applied to determine the largest line in the Block 3(B3) and identifies most visible line on the Block 1 to estimate the angle of skewness (ƟB3).

Step 7:   Calculate skew angle
if (Ɵ B1== Ɵ B2== Ɵ B3)
skewangle= Ɵ B1;
elseif (Ɵ B1== Ɵ B2)

skewangle= Ɵ B2;
elseif (Ɵ B2== Ɵ B3)
skewangle= Ɵ B3;
elseif (Ɵ B1== Ɵ B3)
skew angle= Ɵ B1;
else
(Ɵi, Ɵj)=minimum (|ƟB1- ƟB2|, |ƟB2- ƟB3|, |ƟB1- ƟB3|)
where i, j=B1,B2,B3;
skew angle=average of (Ɵi, Ɵj)
end

Step 8:   Skew correction - Is to rotate the original image with the calculated skew angle in the opposite direction.
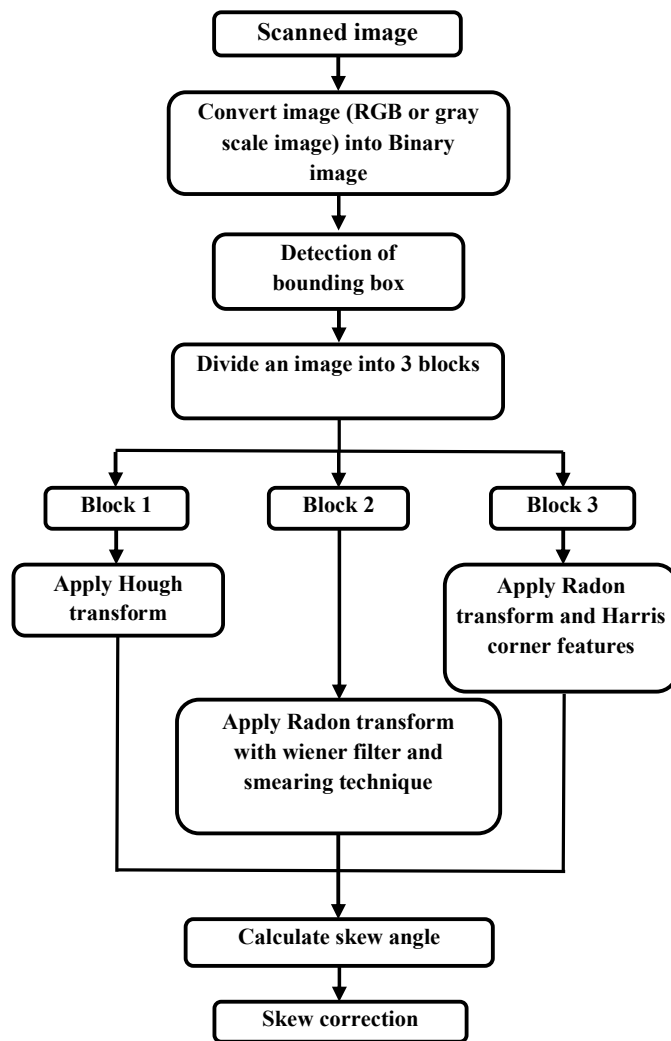
Step 9:   End of program

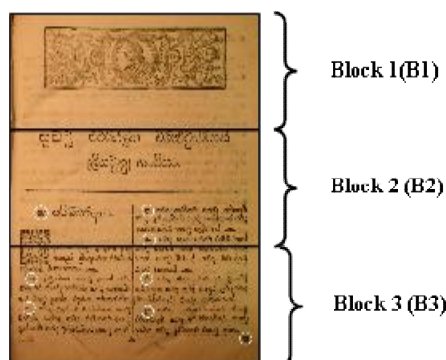**Fig. 4. Block diagram of skew detection and correction technique**

Fig. 5. Block segmentaion on scanned image

5. DATA SET

The Created database for the printed scanned documents consists of 160 samples collected from different old books, magazines, newspapers and pamphlets. This dataset include documents having different scripts: English, Tamil, Sinhala and mixed-script (see Fig. 6). This dataset contains 4 different script consisting 40 images per script. The constructed dataset consists of multi-columns, text and non-text blocks, broken characters, different font size and formatting.

6. RESULTS AND DISCUSSION

The recognition rate is evaluated using the following criteria:

$$\text{Recognition rate} = \frac{\#\, correctly\, detected\, skew\, documents}{\#\, documents}$$

In Tables 1 and 2, presents the recognition rate evaluated on created dataset along with the testing times (in seconds). The best results in the proposed method were obtained by using the Hough transform, Radon transform, Harris corner, Wiener filter, and smearing technique.

The range of skew angle considered here is -20 to +20 degrees.

Table 1. A comparison text only for printed scanned document

Method	Script							
	English		Tamil		Sinhala		Mixed-script	
	Rate (%)	Time (s)	rate (%)	Time (s)	Rate (%)	Time (s)	Rate (%)	Time (s)
Shafii's method [21]	75	1.85	75	1.90	85	2.01	70	2.32
Radon method with Harries corner [18]	80	4.20	80	3.57	90	3.28	80	3.85
Radon method with smearing and Wiener filter [19]	90	2.65	85	2.32	90	2.67	85	2.75
Proposed method	95	2.12	90	2.45	95	2.08	85	2.53

Table 2. A comparison text and non-text for printed scanned document

Method	Script							
	English		Tamil		Sinhala		Mixed-script	
	rate (%)	time (s)	rate (%)	time (s)	rate (%)	time (s)	rate (%)	time (s)
Shafii's method [21]	70	2.05	75	2.10	75	2.35	75	2.25
Radon method with Harries corner [18]	85	4.32	80	4.10	85	3.96	80	4.35
Radon method with smearing and Wiener filter [19]	85	3.95	85	3.35	85	3.24	80	3.60
Proposed method	95	3.70	90	3.30	90	3.10	85	3.23

Fig. 6. Sample of various script document (a) Tamil (b) sinhala (c) English (d) Mixed

The proposed approach is evaluated on the entire dataset (i.e, 160 images) by comparing with the Shafii's method, Radon method with Harries corner, and Radon method with smearing and Wiener filter. Shafii's method yields a recognition rate of 75%, Radon method with Harries corner yields 82.5%, Radon method with smearing and Wiener filter yields 85.62%, whereas my proposed approach yields 90.62%

7. CONCLUSION

In this paper, a method for skew angle detection and correction is proposed. The estimation of skew angle is performed using Radon transform, Hough transform, Harris corner, smearing technique and Wiener filter. The proposed technique is experimentally evaluated on two types of dataset that are based on text only, and text and non-text of printed various script documents. The skew detection method yields a recognition rate of 91.25% for text only of printed various script documents and yields a recognition rate of 90% for text and non-text of printed various script documents. The overall recognition rate is around 90.62%. The recognition rates confirm that the proposed technique is effective for preprocessing step of Optical Character Recognition on multi-script printed document.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Salem Saleh Bafjaish, Mohd Sanusi Azmi, Hairulnizam Mahdin. Skew detection and correction of Mushaf Al-Quran script using hough transform. International Journal of Advanced Computer Science and Applications. 2018;9(8):402-409.
2. Abdelhak Boukharouba. A new algorithm for skew correction and baseline detection based on the randomized hough transform. Journal of King Saud University–Computer and Information Sciences. 2017; 29-38.
3. Rubani Jyoti Rani. Skew detection and correction in text document image using projection profile technique. International Journal of Computer Sciences and Engineering. 2018;6(7):986-990.
4. Ramandeep Kaur, Seema, Sunil Kumar. A hybrid approach to detect and correct a skew in scanned document images using fast Fourier transform and nearest neighbor algorithm. International Journal of Advances in Electronics and Computer Science. 2016;3(5):1-6.
5. Subrahmanyam MSLB, Vijaya Kumar V, Eswara Reddy B. A new algorithm for skew detection of Telugu language document based on principle-axis farthest pairs quadrilateral (PFPQ). International Journal of Image, Graphics and Signal Processing. 2018;47-58.
6. Papandreou A, Gatos B. A course to fine skew estimation technique for handwritten documents. International conference on document analysis and recognition (IEEE); 2013.
7. Aditya B, Kumar A, Yadav BM. Skew Detection in handwritten documents. International Journal of Computer Applications. 2013;69:17-18.
8. Jipeng T, Kumar HG, Chethan HK. 'A skew correction for Chinese character using hough transform. International Journal of Advanced Computer Science and Applications. 2011;45-48.
9. Singh R, Kaur R. 'Improved skew detection and correction approach using discrete fourier algorithm. International Journal of Soft Computing and Engineering. 2013; 3(4):55-57.
10. Kaur L, Jindal S. Skew detection technique for various scripts. International Journal of Scientific & Engineering Research. 2011;2(9):1-3.
11. Rao KR, Kim DN, Hwang JJ. Fast Fourier Transform - Algorithms and Applications, Springer, Dordre ht; 2010.
12. Mehta S, Walia E, Dutta M. Fast frequency domain method to detect skew in a document image. Seventh International Conference on Graphic and Image Processing. 2015;9817.
13. Yan H. Skew correction of document images using interline cross-correlation. CVGIP Graphical Models and Image Processing. 1993;55(6):538-543.
14. Li ST, Shen QH, Sun J. Skew detection using wavelet decomposition and projection profile analysis. Patter Recognition Letters. 2007;28(5):555-562.
15. Basavanna M, Gornale SS. Skew detection and skew correction in scanned document image using principal component analysis. International Journal of Scientific & Engineering Research. 2015;6(1):1414-1417.

16. Steinherz T, Intrator N, Rivlin E. "Skew detection via principal components analysis. Proceedings of the International Conference on Document Analysis and Recognition. 1999;153-156,

17. Aithal K, Rajesh G, Acharya D, Siddalingaswary P. A fast and novel skew estimation approach using radon transform. International Journal of Computer Information Systems and Industrial Management Applications. 2013;5:337-344.

18. Ahmed Gari, Ghizlane Khaissidi, Mostafa Mrabti, Driss Chenouni, Mounim El Yacoubi. Skew detection and correction based on Hough transform and Harris corners, 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS); 2017.

19. Ramanan M, Ramanan A, Charles EYA. A preprocessing method for printed Tamil documents: Skew correction and textual classification. IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS). 2015;495-500.

20. Harris C, Stephens M. A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference. 1988;147-151.

21. Shafii M. Optical character recognition of printed persian arabic documents. Electronic Thesis and Dissertations; 2014.