# LANGUAGE INDEPENDENT ROBUST SKEW DETECTION AND CORRECTION TECHNIQUE FOR DOCUMENT IMAGES 

N. Neha<br>Department of Electronics and Communication The National Institute Of Engineering, Mysore, India, nehanie2010@gmail.com

Follow this and additional works at: https://www.interscience.in/ijess
Part of the Electrical and Electronics Commons

## Recommended Citation

Neha, N. (2012) "LANGUAGE INDEPENDENT ROBUST SKEW DETECTION AND CORRECTION TECHNIQUE FOR DOCUMENT IMAGES," International Journal of Electronics Signals and Systems: Vol. 2 : Iss. 2 , Article 9.

DOI: 10.47893/IJESS.2012.1077
Available at: https://www.interscience.in/ijess/vol2/iss2/9

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Electronics Signals and Systems by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

# LANGUAGE INDEPENDENT ROBUST SKEW DETECTION AND CORRECTION TECHNIQUE FOR DOCUMENT IMAGES 

Neha.N<br>Department of Electronics and Communication The National Institute Of Engineering, Mysore, India<br>e-mail: nehanie2010@gmail.com


#### Abstract

Document image processing is an increasingly important technology essential in all optical character recognition (OCR) systems and for automation of various office documents. A document originally has zero-skew (tilt), but when a page is scanned or photo copied, skew may be introduced due to various factors and is practically unavoidable. Presence even a small amount of skew $\left(0.5^{0}\right)$ will have detrimental effects on document analysis as it has a direct effect on the reliability and efficiency of segmentation, recognition and feature extraction stages. Therefore removal of skew is of paramount importance in the field of document analysis and OCR and is the first step to be accomplished. This paper presents a novel technique for skew detection and correction which is both language and content independent. The proposed technique is based on the maximum density of the foreground pixels and their orientation in the document image. Unlike other conventional algorithms which work only for machine printed textual documents scripted in English, this technique works well for all kinds of document images (machine printed, hand written, complex, noisy and simple). The technique presented here is tested with 150 different document image samples and is found to provide results with an accuracy of $0.1^{0}$


Keywordsg; skew, morphological, projection, radial line

## I. INTRODUCTION

All optical character recognition (OCR) systems need the input in the form digital document. The digital image of the document may be skewed (rotated) because of how it placed on the platen when scanned or because of the malfunction of the document feeder. The skew is inevitably introduced in the document images especially when digitizing huge document bulks. Document images can be of various types. Simple textual document images are the ones that consist of only text without any graphic or image embedded with it. Complex document images on the other hand consist of text along with images/graphics like barcodes, stamps, signatures, figures, block diagrams etc. Presence of skew in document images will dramatically degrade the feature extraction efficiency of OCR systems, thus making Skew correction a crucial requirement.

In the past, this problem has been addressed by several researchers by considering the inherent properties of the textual content in the document image like font, base line, border, statistical features of connected components. But these algorithms work well only for simple document images scripted in a particular language. The other algorithms include Fourier transform method, hough transform method, straight line fitting method, histogram based projection profile method, cross correlation based technique. However, these approaches are designed mainly for machine printed documents and fail for hand written, complex and noisy documents. Also some of these approaches are capable of correcting small skew angles within $\pm 45^{\circ}$, They fail to handle the document images having a skew greater than $45^{\circ}$. Moreover, some of them entail high computational cost, especially in the case where the Hough and Fourier transform is used. The main disadvantage of

## these algorithms is that their accuracy is limited to $0.5^{0}$

The technique presented in this paper nullifies almost all the limitations of the existing techniques and is made suitable to handle wide variety of documents that are produced in day to day life with an accuracy of $0.1^{0}$.

Paper outline. Section II presents the key observations which form the basis of the technique presented in this paper. The details of the proposed technique are described in section. III. Section IV provides experimental results. Section V and VI provide the comparison and advantages respectively

## II. KEY OBSERVATIONS

The core of the presented technique is an isophote driven from how human brain determines the presence of skew. The contents (text \& other graphics) in the document image will have the same orientation as that of the document. i.e., the contents present in a document image having zero skew will be properly oriented with respect to its vertical and horizontal positions. For tilted document images may it be simple or complex, the contents present in it will be at some angle with respect to the horizontal depending upon the tilt. Humans usually determine document orientation by collecting some regularly aligned symbols, mostly text. A symbol line is defined as a group of regularly aligned symbols that are adjacent, relatively close to each other.


Figure 1 A skewed complex document image
From figure-1 we observe that, the relative distances between characters in same lines are generally smaller than the distances between text lines which give rise to more number of foreground pixels in the same line. Also the edges of various kinds of graphics that may be present in the document image along with the text, are distributed over a particular area and their contribution to the density of foreground pixels in each textual line is less. This suggests that it is the textual part of the document that contributes majorly for finding the document tilt.

The characteristics of the document image observed here are neither dependant on language in which it is not scripted nor the complexity of the document like whether the document is hand written, machine printed, simple or complex. This paper presents a novel technique for skew detection and correction based on the above observations thus making it language independent and efficient than the previous works.

## III. PROPOSED TECHNIQUE

The proposed skew detection and correction technique consists of three main stages:

1. Preprocessing and obtaining the binary version of the input document image.
2. Detecting the amount of skew present in the document image.
3. Skew Correction
A. Stage-1: Pre-processing and obtaining the binarized version of the scanned document image.

Pre-processing is the first step in any kind of document image processing. In pre-processing stage, the input document image is processed so as to obtain an enhanced binary version of it. The pre-processing stage consists of three steps:
In Step-1, the input document image is checked if it is a color image or a gray scale image. In case of a color image, it is converted into a gray-scale image by eliminating the hue and saturation information while retaining the only the luminance. In step-2 the gray scale image obtained is converted to binary such that
the foreground text and the edges of various graphics that may be present in the image are replaced by white pixels and image background is replaced by black pixels. Finally in Step-3, the binary image obtained is enhanced by morphological thickening of the foreground pixels. Morphological thickening thickens all the connected foreground pixels thus suppressing the noise( salt and pepper). This makes the skew detection process more accurate. If the input document image is noise free then step 3 can be avoided and thereby reduce the computational cost.

## B Stage-2: Detection of the amount of skew present in the binarized version of the input document image.

On obtaining the enhanced binary version of the document image from stage-1, the amount of skew is found by finding the orientation of the line along which we obtain the maximum density of the foreground pixels which is nothing but the orientation of the input document image with respect to the horizontal. In order to achieve this, a reference line oriented along $-90^{\circ}$ is positioned exactly at the centre of the binary image such that its centre coincides with the centre of the binary image as shown in figure 2. Next we find the projection of the image intensity onto this orientation of the reference line. The lines perpendicular to the reference line in figure 2 are imaginary lines showing the projections taken at a distance of one pixel width. Projection is nothing but the summation of intensities of the pixels, perpendicular to the radial line direction. The number of projections taken is decided by the number of pixels on the diagonal of the input document image. The projection values so obtained are stored along the column of a predefined 2-D matrix.

Further the orientation of the reference line is changed by rotating it in the anticlockwise direction by $0.1^{\circ}$ i.e., from $-90^{\circ}$ to $-89.9^{\circ}$ and the projection of the image intensity is taken onto the new orientation of the reference line.


Figure 2: Image showing the projections taken onto -90 ${ }^{\mathbf{0}}$ orientation of the reference line.

The process of rotating the reference line counterclockwise in steps of 0.1 is continued till $+89.9^{\circ}$ and projection is taken for every orientation of the imaginary radial line. Doing this will result in a complete 2-D matrix, in which each column contains the projected values for a particular orientation of the imaginary radial line. Maximum projection value will be obtained along the direction where the foreground pixel density is maximum.

In order to obtain the skew of the document image, first the orientation of the imaginary radial that corresponds to maximum projection value is extracted from the 2D matrix . To the angle so found $90^{\circ}$ is either added or subtracted depending upon whether the obtained angle is negative or positive respectively. This is done as projection is always taken perpendicular to the reference line. Doing this will give us the actual skew angle of the input document image.

## C. Stage-3: Skew correction

If the skew angle is determined using the procedure described above is not zero, erect input document image by rotating it back by the skew angle detected in stage-2. The rotation is carried out the by using the following formula

$$
\begin{aligned}
& X=x \cos \theta+y \sin \theta \\
& Y=y \sin \theta-x \cos \theta
\end{aligned}
$$

This transfers a point ( $\mathrm{x}, \mathrm{y}$ ) from original image to a new point ( $\mathrm{X}, \mathrm{Y}$ ) in the same coordinate system. Here $\theta$ is the skew angle between the document orientation and the x -axis along horizontal direction. For transferring the original image to a de-skewed image, we need to adjust the size of the output buffer and shift the $(\mathrm{X}, \mathrm{Y})$ coordinates in a usual way.

## IV. EXPERIMENTAL RESULTS

The algorithm presented here is tested for 150 document image samples which include simple, complex, noisy and handwritten. The technique is proved to provide the result with accuracy of $0.1^{0 .}$ An accuracy of $0.1^{0}$ indicates that even a skew of $0.1^{0}$ present in the document image can be detected and corrected. Finally a GUI is created for the ease to the end user. Figure 3 shows a few tested results.

## V. COMPARISION

Table 1 provides the comparison of the technique presented in this paper with few other popular methods.

Table 1

| SI | Method | Skew <br> range <br> in <br> degrees | Accuracy <br> in <br> degrees | Type of <br> document <br> that can <br> be <br> handled |
| :---: | :---: | :---: | :---: | :---: |
| 1. | Fourier <br> transform | $\pm 45^{\circ}$ | 2.0 | Simple <br> textual, <br> complex |
| 2. | Straight <br> line fitting | $\pm 90^{\circ}$ | 0.5 | Simple <br> textual <br> (hindi, <br> Sanskrit, <br> bangla) |
| 3. | Histogram <br> based <br> projection <br> profile | $\pm 15^{\circ}$ | 0.7 | Simple <br> textual |
| 4. | Cross <br> correlation | $\pm 45^{\circ}$ | 1.0 | Simple <br> textual |
|  | ( <br> Proposed | $\mathbf{\pm 9 0 ^ { \circ }}$ | 0.1 <br> (can be <br> reduced <br> further) | All kind <br> of <br> document <br> images |
|  | Pechnique |  |  |  |

## VI. ADVANTAGES AND APPLICATIONS

The skew detection and correction technique presented in this paper has a number of advantages over the previous works which can be listed as follows:

1. It is language independent.
2. It works well with both simple textual as well as complex document images with an accuracy of $0.1^{0}$. As accuracy is dependent on the incremental angle used in the algorithm, it can further be reduced at cost of a little increased execution time. Thus the technique presented provides flexibility in accuracy.
3. Range of skew angle that can be detected and corrected is $\pm 90^{\circ}$.
4. It works well with noisy images as well

(
a
)

(
b
Figure 3: (a) Result obtained for the image of a cheque leaf having a skew of $43^{\circ}$. (b) Result obtained for the noisy document image having a skew of $65^{0}$

Skew angle estimation and correction procedure presented, has a wide range of direct and indirect applications in the field of document image processing and character recognition systems. It is the first crucial step (pre-requisite) to be carried out in any character recognition system such as Bank cheque processing. Postal zip code and address recognition, Automatic license plate recognition etc.

## AKNOWLEDGEMENT

The satisfaction and euphoria accompanying the successful completion of this is incomplete without the mention of those people who have been a source of great inspiration and help. Firstly I thank god for his blessings. I extend my sincere thanks to Dr.

Rohini Nagapadma. Professor and head, department of electronics and communication N.I.E .I also express my gratitude to my parents for their support and encouragement throughout this effort

## REFERENCES

[1] C. Sun and D. Si, Skew and Slant Correction for Document Images Using Gradient Direction, In: Proceedings of the 4th International Conference on Document Analysis and Recognition, Ulm, Germany, 1997, pp.142-146.
[2] Huiye Ma, Zhenwei Yu, "An enhanced skew angle estimation technique for binary document images".
[3] S. Chen, R.M. Haralick, and I.T. Phillips, Automatic Text Skew Estimation in Document Images,In: Proceedings of the 3rd International Conference on Document Analysis and Recognition,Montreal, Canada, 1995, pp. 1153-1156.
[4] Abhisek Gattani, Maitrayee Mukerji, Hareish Gur, "A fast multifunctional approach for document image analysis ", 2003 IEEE.
[5] Adnan Amin, Stephen Fischer, Tony Parkinson, Ricky Shiu, "Fast algorithm for skew detection".
[6] Hua Shen, Xiaoou Tang, "Generic Sign Board Detection in Images".
[7] A. Bagdanov and J. Kanai, Projection Profile Based Skew Estimation Algorithm for JBIG Compressed Images, In: Proceedings of the 4th International Conference on Document Analysis and Recognition, Ulm, Germany, 1997, pp. 401-405.
[8] L. O'Gorman, The Document Spectrum for Page Layout Analysis, IEEE Trans. on Pattern Analysis and MachineIntelligence, Vol. 15, No. 11, 1993, pp. 1162-1173.
[9] B. Yu and A. Jain, A Robust and Fast Skew Detection Algorithm for Generic Documents, Pattern Recognition, Vol. 29, No. 10, 1996, pp. 1599-1629.
[10] S.C. Hinds, J.L. Fisher, and D.P. D'Amato, "A Document Skew Detection Method Using Runlength Encoding and the Hough Transform," Proceedings, 10th International Conference on Pattern Recognition, pp. 464-468, 1990.
[11] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing"
[12] C.Sun and D. Si, "Skew and Slant Correction for Document Images using Gradient Direction", Proc. of Forth International Conference on Document Analysis and Recognition, pp. 142-146, August 1997.
[13] R. Safabakhsh and S. Khadivi, "Document Skew Detection Using Minimum-Area Bounding Rectangle" Proc. of The International Conference on Information Technology:Coding and Computing, pp. 253-258, March 2000
[14] H. Ma and Z. Yu, "An enhanced Skew Angle Estimation Technique for Binary Document Images", Proc. of Fifth International Conference on Document Analysis and Recognition, September 1999

