Various Skew Correction Techniques: Kannada Handwritten Script

Malini. M, Hemanth. K.S

School of CSA, REVA University, India

Corresponding author: Malini. M, Email: r21pca02@reva.edu.in

In a handwritten text document, before identifying the character it is essential to pre-process. During the pre-processing stage, identifying the skewed angle and correcting the text document is a significant step. A few effective algorithms are used in later stages to extract the features and identify the characters. In this paper, Fourier Transform, Hough Transform, and Horizontal Projection Profile Techniques are described in steps. These are implemented forskew detection, and correcting techniquesin a Kannada handwritten document. Comparison analysis is also undertaken in this study, to get unambiguousness about the suitable algorithm to use on different documents for effectual results.

Keywords: Skew Detection, Skew Correction, Handwritten Document, Hough Transform, Fourier Transform, Projection Profile

2023. In Satyasai Jagannath Nanda & Rajendra Prasad Yadav (eds.), *Data Science and Intelligent Computing Techniques*, 633–642. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.56155/978-81-955020-2-8-57

1 Introduction

Kannada Handwritten text recognition is a challenging task, due to various reasons such as different styles of handwriting, different font sizes, spacing between letters and characters, overlapping, and multi-touch of nearby characters. When the document is scanned unnecessary noises [1] in artifacts are captured, errors in segmentations, and complexity in computing with several deep learning techniques. Along with all these and more, there exists a skew identification and correction [2] method which becomes very much essential even before the recognition of handwritten Kannada characters[3]. If undone, then the final output will surely be an improper result.

The process of digitalization has various steps that include in sequence as scanning the handwritten document - removing the unwanted noises that get created during scanning -converting to binary format - skew detection and correction – segmentation - and recognizing characters with accuracy and efficiency.

The complete steps specified starts with pre-processing techniques [4] that require error minimization to acquire accurate result. To achieve this, the following various techniques are specified to find the skewed angle and correct the tilted document. components, incorporating the applicable criteria that follow.

2 Related Works

Enhanced Supervised Learning Distance (ESLD) and R_Clustering algorithms [5] are used to segment and calculate skewness in the handwritten document. According to the algorithm, after preprocessing is completed, the document is divided into several segments, first, classify the connected components, second, identify the average width and height of each, and next, group using the R_Clustering Algorithm. Further, calculate the gap using ESLD and extract the words. Now skew angle is estimated based on the gap and deskewed the word. Here, care is taken not to consider overlapped words.

Various skew techniques that include the Hough Transform, Projection histogram, Method of Least Squares, and Word Centroid Least Squares [6] were analyzed in the English handwritten paper.According to their study, the Hough transform produces more accurate skewed results than the rest of the skew techniques. The projection histogram delivers the least accurate outcome.The document image while scanned might get skewed, then 4 skew correction methods viz., Hough Transform, cross-correlation, K-Nearest Neighbour, and Fast Fourier Transformation were experimented with various digitalized documents [7] and found that a Fast Fourier transform would provide more correct results with less time than the rest.

The skew correction technique is applied to the ancient Kannada document. Since it is a historical document, it needs a morphological as well as clustering technique [8] for identifying skew angles and correcting them. The technique that was proposed, according to their experiments, the method provided a good result. Another method involves determining the coordinates from the starting letter to the ending letter. The slope is computed and converted from radians to the desired angle[9]. According to the standard dataset images were chosen, the result would give both positive and negative skews in both left and right directions.

The skewed angle in the printed document can also be identified using deep learning techniques [10]. The machine is first allowed to learn the features of different languages and later feature testing is applied. Gradually the skewed angle is identified by classifying the rotated image document from class 0 to class 360.

3 Various Techniques to Identify Skew Angle and Correct

The skewed angle in the document refers to the deviation in angle from the horizontal line. Some of the techniques to identify deviated angle and correct this deviated angled document are specified as follows:

3.1 Fourier Transform:

The name Fourier was from the scientist Jean Baptiste Joseph Fourier. This technique is a significant method that is used to simplify the given image in sinusoidal waves, which are periodic signals to the weighted summation of different frequency components. If the input is in terms of (x, y) coordinates, then the output is transformed into a frequency series (see Fig. 1). Each wave can be represented in a specific frequency and phase[18]. All waves put together form a summation of discrete frequency components without loss of information.

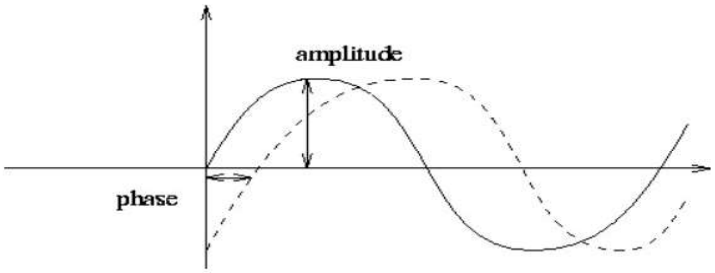

Figure 1: Amplitude and phase wave at particular frequency[19].

$$f(x) = A\sin(2\pi ux + \emptyset) \quad (1)$$

where A is the amplitude of the sinusoid, $\emptyset$ is a phase, u is the frequency of the sinusoid, and 2π is the period of the sinusoid.
The sinusoidal wave is represented as the sum of several sinusoids of different frequencies with amplitudes and phases as shown (see Fig. 2 and 3):

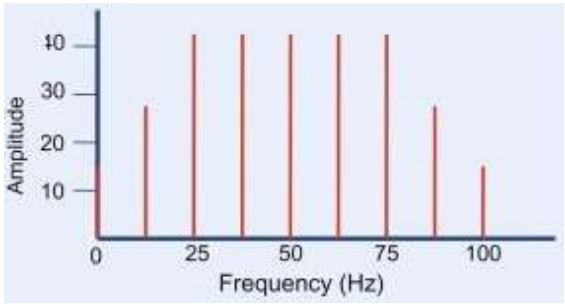

Figure 2: Amplitude Spectrum [20].

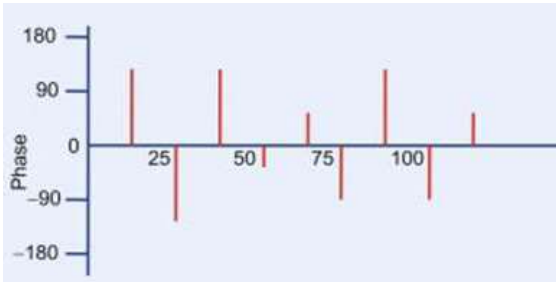

Figure 3: Phase Spectrum [20].

This transformation is important, as certain operations can be performed and then reconverted to the original image. These transformations can identify the minute detailed feature, which otherwise could not be identified. Moreover, it is used for edge detection, smoothing, and denoising on various applications.

Fourier Transform represents a signal f(x) to the sum of amplitudes and phases of different frequencies.

$$F(x) \xrightarrow[\text{Transform}]{\text{Fourier}} F(u) \qquad (2)$$

Equation of Fourier Transform

$$F(u) = \int_{-\infty}^{\infty} F(x) e^{-j.2\pi.u.x}\, dx \qquad (3)$$

, which is a continuous integral component and the variables used in the equation, represent such as x: space, u: frequency, $e^{i\theta} = cos\theta + isin\theta$, i=$\sqrt{-1}$, j: direction, F(x): original signal

$$F(u) = \text{Real } F(u) + i \text{ Imaginary } F(u) \qquad (4)$$

Fourier Transform is a complex form and represents frequencies of both positive and negative components.

The equation for Discrete Fourier Transform

$$F(u_k) = \frac{1}{N}\sum_{n=0}^{N-1} F(n)\, e^{\frac{-j\,2\pi kn}{N}} \qquad (5)$$

which is a discrete component and the variables used in the equation represent such as u is frequency, evaluating n values of N samples in kth frequency and in j direction. It is the summation of k frequencies from n=0 ton=N-1.The steps involved in Discrete Fourier Transform is as follows:

1. The scanned image is converted to grayscale and undergoes a noise-reducing process.
2. The grayscale and the denoised image are binarized and the threshold value is calculated.
3. The center of the image is shifted at the origin of the cartesian system by multiplying it with the phase shift term.
4. The power spectrum is computed using an absolute square of the Discrete Fourier coefficients.

Magnitude Spectrum=sqrt(Real(F(x, y))^2 + I(F(x, y))^2 (6)

where Real(F(x, y)) is real components of F(x, y)
I(F(x, y)) is imaginary components of F(x, y)

5. Calculate the frequency in the power spectrum. This corresponds to the angle of the image.

theta=atan(y/x) (7)

where theta is the skew angle.
x, y are frequency coordinates of the peakmagnitude.

6. Apply the rotation algorithm based on the angle found in step 5.

Figure 4: Original Scanned Document

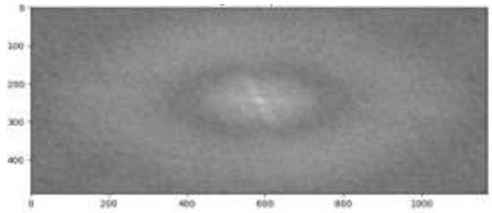

Figure 5: Computed frequency spectrum angle.

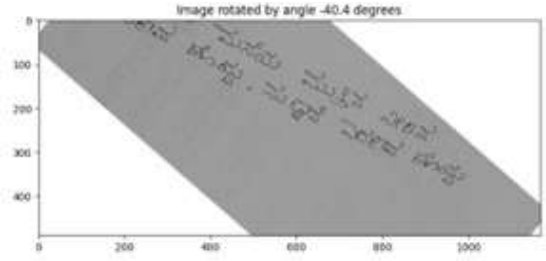

Figure 6: Rotated by calculated angle.

The input document is scanned, converted to grayscale, and removed noise (see Fig. 4). Then binarization process is undergone and the threshold value is calculated. The frequency spectrum is calculated (see Fig. 5). Based on the angle obtained, a rotation algorithm is used to correct the skewed document. The above document was skewed to -40.4 degrees (see Fig. 6).

The Fourier transform will be better suited for detecting global skew i.e., when the entire document is rotated by the same angle. It might not be as effective in detecting local skew, where different parts of the document have different rotation angles.

3.2 Hough Transform

The algorithm was designed by Paul. V. C. Hough. It is used mainly for edge-linking in image processing.
The following steps are used to identify skewness [11, 12, 13,14, 15, 17] and correct the document.

1. To detect the edge, a few edge-detecting algorithms are used such as Sobel Operator, Prewitz Operator, Canny Edge Detector, Robertz Operator, and Scharr Operator. Using any of these mask algorithms, boundaries are identified. If there are discontinuous boundaries, then the Hough Transform algorithm is used.
2. The parameter(slope and y-intercept) space is used to indicate the possible shapes in the image.
3. For every boundary pixel in the image, a line is drawn to indicate likely shapes that could occur at that point.
4. An array is generated that specifies the likely shape in the parameter plane of the image. This array is used to compute the threshold value to find the probable shape of the given image. Only the values that are greater than the threshold value are selected.
5. These are then extracted to represent the original image shape.

In this algorithm, a line equation is used to join the points. Here while doing so, the input image which is in the spatial plane is converted to Hough space. This Hough space is otherwise called parameter space that is in the form of a c-m (intercept-slope) plane.
The line equation,

$$y_j = mx_j + c \text{ or } c = -x_j m + y_j \,. \tag{8}$$

Further, to detect lines drawn in the image

$$\rho = x\cos\theta + y\sin\theta, \tag{9}$$

where ρ is the length of a point to the line detected, (x, y) is the coordinate of a point in an image, and θ is the angle between the line and the x-axis.

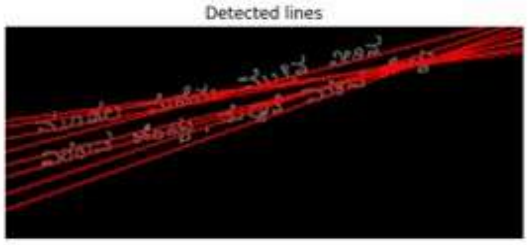

Figure 7: Lines drawn for the edge detected.

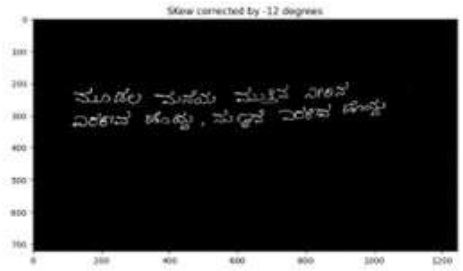

Figure 8: The line peak angle is found and rotated accordingly.

The input document is scanned, converted to grayscale, and removed noise. Then binarization process is undergone and the threshold value is calculated. Here edge is detected using the Canny edge detector, and the Hough transform is used to draw the line (see Fig. 7). The above document was skewed to -12 degrees (see Fig. 8).
It can handle noisy images effectively, as it works based on a voting scheme. Even if some edge points are corrupted by noise, the dominant lines can still be detected accurately.Detects lines of any orientation that need not be necessarily horizontal or vertical. However, with its computational complexity, it can be a concern in large-scale image analysis.

3.3 Horizontal Projection Profile

In Horizontal Projection Profile (HPP), it is the sum of positive numbers along the x-axis and y-axis. In binarized images, the image that is first converted to grayscale is the pixel values from the range of 0 to 255 according to the pixel intensities or brightness values. Later identified the threshold value of these pixels. Further, replace the values with binarized one that is either 0 to 1 depending on the less than threshold value and greater than threshold value respectively.

The pixel values which are 1, positive values are summed up either the x-axis or y-axis. The pixel values are summed up in row-wise is horizontal and the values summed up in column-wise is a vertical projection profile.

The steps to follow to identify skewness and rotate the handwritten document are as follows:

1. The gray-scaled image is converted to a binary image.
2. The binary image is used to compute the horizontal projection profile.
3. $$P(i) = \sum[j = (0 \text{ to } n - 1)]\ I(i, j) \quad (10)$$

where P(i) is the projection profile for row i, n is the total columns in the image, I(i, j) is the pixel value at row i and column j.

4. Calculate the mean and standard deviation of the values in the projection profile array.
5. Compute the skewness of the horizontal projection profile using the formula:

$$\text{skewness} = \left(\frac{1}{N}\right) * \sum \left[{(x_i - mean)}/{\text{std\_dev}}\right] \text{^}3 \quad (11)$$

where N is the total rows in the image, xi is the value of the horizontal projection profile at row i, mean is the mean value of the horizontal projection profile, and std_dev is the standard deviation of the horizontal projection profile.

6. The value can be either positive or negative depending upon the slantness.
7. Applying a rotation transform to correct the skew in the document.

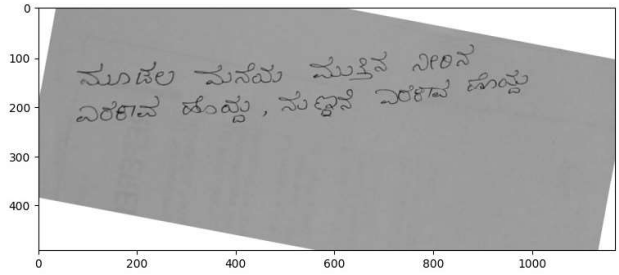

Figure 9: The input document was rotated to an angle of -10.53

In this process, the input document (see Fig. 4) is converted to a binary image. The horizontal projection profile is calculated on the binary image. The mean and the standard deviation of the array are computed. Based on these values, the skew angle is determined [12,16, 17]. Rotation transformation is applied to the document so as to align it horizontally (see Fig. 9).

Though HPP can be used in an efficient manner for well-structured documents, care has to be taken while implementing it in certain documents. Those documents that have more noise, small fonts, lower / higher resolutions, and skewed text, cannot provide a satisfactory result. As more noises can lead to false peaks that can affect the succeeding outcomes, small fonts might not generate peaks, lower resolutions can result in loss of fine details, high resolutions can result in increased computations, and skewed text might require extra steps of preprocessing.

4 Observation Analysis

In the Fourier Transform, directionality is lacking, as it has no parameters to identify the orientation of directions and it is suitable only for the periodic continuous wave and not for the non-periodic wave or signal. As the document image is transformed to a frequency plane from a spatial (x,y) plane, that can lead to difficulty in analyzing the result because of transformations.For Kannada handwritten documents, this algorithm is suitable for reducing noises, pattern enhancements, compressions, and text extractions.

In Hough Transform, the limitation is that, if the document is vertically inclined, then it will have an infinity slope. Moreover, every time the output line must be converted into polar coordinates. It is difficult to accurately detect overlapping shapes. But from the output perspective, that is arrived, the Hough transform can be applied to the handwritten Kannada document to identify the skewness and correct the angle. Few refinements need to the applied to the algorithm when in case it is applied to handwritten documents. This is because of the font styles, size, and shape written in the artifact.Also, those documents that require the identification of patterns, and shapes can get expected results from the algorithm.

The Horizontal Projection Profile can be difficult to use if the text is of a small size. Because the basic concept of HPP is the sum of black pixels in the row. This summation will not create a peak that in turn couldn't be used to segment the line or the word. If the document contains irregular text layouts, the profile may not accurately represent the dominant orientation of the text. Moreover, if there are large gaps between columns can also lead to incorrect skew detection. The Kannada documents that are required for segmenting paragraphs, lines , detecting skew angle is suitable to use this HPP algorithm.

5 Conclusion

In this study, various skew detection and correction algorithms were used along with different skewed handwritten documents. Every algorithm has its strengths and limitations. These algorithms cannot provide a perfect solution for a particular handwritten document, as every algorithm has its expertise in a different arena. But when applied in a particular way, the algorithms will give the desired result. The Fourier Transform when applied to find frequency-based analysis such as filtering, compression, and so on, gives the finest result. In the same manner, when the Hough Transform is applied in computer vision tasks like object recognition, detecting lines, circles, and other shapes will obviously provide the best solution. When the Horizontal Projection Profile is applied to analyze or extract information from an image based on horizontal or vertical pixel values, will give the required answer. In future work, an algorithm that could minimize the loss of features in an image and align the document with the best accuracy needs to be worked on.

References

[1] Muhammad Sarfraz, Zeehasham Rasheed(2008). Skew Estimation and Correction of Text using Bounding Box. Fifth International Conference on Computer Graphics, Imaging and Visualization, DOI 10.1109/CGIV.2008.10, 978-0-7695-3359-9/08.

[2] Liang Junjuan, Tan Guoxin (2008). An Efficient Algorithm for Skew-Correction of Document Image Based on Cyclostyle Matching. International Conference on Computer Science and Software Engineering, DOI 10.1109/CSSE.2008.443, 978-0-7695-3336-0/08.

[3] M. Sarfraz, S.A. Mahmoud and Z. Rasheed (2007). On Skew Estimation and Correction of Text. Computer Graphics, Imaging and Visualisation (CGIV 007) 0-7695-2928-3/07.

[4] AdityaVighnesh, B., Kumar, A., & Manikanta Yadav, B. (2013). Skew Detection in Handwritten Documents. International Journal of Computer Applications, 69(11), 17-18.

[5] Shakunthala B S, Dr. C S Pillai (2021). Enhanced Text Line Segmentation And Skew Estimation For Handwritten Kannada Document. Journal of Theoretical and Applied Information Technology, 15th January 2021. Vol.99. No 1, ISSN: 1992-8645.

[6] Wesley Chin, Man Harvey, Andrew Jennings(1997). Skew Detection In Handwritten Scripts. IEEE TENCON - Speech and Image Technologies for Computing and Telecommunications.

[7] A.Sakila, Dr. S.Vijayarani (2017). Skew Detection and Correction in the Document Image. International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Issue 8, August 2017. DOI:10.15680/IJIRSET.2016.0608274.

[8] Ravi P, Naveena C, Sharathkumar Y H (2020). Skew Detection and Correction System for Ancient Kannada Document. International Journal of Computational Linguistics Research Volume 11 Number 4 December 2020 128-134.

[9] Banumathi K L, Dr. Jagadeesh Chandra A P (2018). A New Algorithm for Skew Detection and Correction of Historical Kannada Documents. International Journal of Research in Advent Technology. Special Issue, June 2018, E-ISSN: 2321-9637, International Conference on Multimedia Processing, Communication, and Information Technology.

[10] Shaheera Saba Mohd Naseem Akhter, Priti P Rege(2020). Improving Skew Detection and Correction in Different Document Images using a Deep Learning Approach. IEEE – 49239. 11th ICCCNT 2020. July 1-3, 2020 - IIT – Kharagpur.

[11] Bafjaish, Salem & Azmi, Mohd & Al-Mhiqani, Mohammed & Radzid, Amirul Ramzani & Mahdin, Hairulnizam. (2018). Skew Detection and Correction of Mushaf Al-Quran Script using Hough Transform. International Journal of Advanced Computer Science and Applications.

[12] Daniel Rosner, Costin-Anton Boiangiu, Mihai Zaharescu, Ion Bucur (2014). Image Skew Detection: A Comprehensive Study. Conference Paper.

[13] Chinmay Chinara, Nishant Nath, Subhajeet Mishra, Sangram K. Sahoo, Farida A. Ali (2018). A Novel Approach to Skew-Detection and Correction of English Alphabets for OCR. IEEE Student Conference on Research and Development.

[14] Chandan Singh, Nitin Bhatia, Amandeep Kaur (2008). Hough transform based fast skew detection and accurate skew correction methods. Doi:10.1016/j.patcog.2008.06.002, Pattern Recognition 41 (2008) 3528 – 3546.

[15] Amjad Rehman & Tanzila Saba (2011). Document Skew Estimation And Correction: Analysis Of Techniques, Common Problems And Possible Solutions. Applied Artificial Intelligence, 25:769–787, 2011, Copyright # 2011 Taylor & Francis Group, LLC, DOI: 10.1080/08839514.2011.607009.

[16] Sneha L. Bagadkar, Dr. L. G. Malik(2014). Review on Extraction Techniques for Images, Text lines and Keywords from Document Image. IEEE .

[17] Barun Biswas, Dr. Ujjwal Bhattacharya, Bidyut B Chaudhuri (2023).Document Image Skew Detection and Correction: A Survey. IJIRT (International Journal Of Innovative Research In Technology) 158871, Volume 9 Issue 10, ISSN: 2349-6002.

[18] The University of Edinburgh, School of Informatics, Computer Vision IT412, "The Discrete Fourier Transform",https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT4/node2.html. [Accessed June 5, 2023]

[19] Robyn Owens, Computer Vision IT412, Department of Computer Science, Fourier transform theory: https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT4/node2.html. [Accessed June 5, 2023]

[20] Enwenode Onajite, Chapter 2 - Understanding Seismic Wave Propagation, Seismic Data Analysis Techniques in Hydrocarbon Exploration, Elsevier, 2014, Pages 17-32, ISBN 9780124200234, https://doi.org/10.1016/B978-0-12-420023-4.00002-2. (https://www.sciencedirect.com/science/article/pii/B9780124200234000022)