

Segmentation of Connected Components and Overlapping Lines in Gurumukhi Handwritten Documents

Snehdeep

Student, M.tech, Final Year,
Yadavindra College of Engineering,
Talwandi Sabo, Punjab, India

Manoj Kumar

Assistant Professor,
Yadavindra College of Engineering,
Talwandi Sabo, Punjab, India

ABSTRACT

Text line segmentation is extremely important phase of OCR. Overlapped lines, skewed lines and connected components make the problem of line segmentation more complicated in Gurumukhi handwritten documents. The existence of these problems in handwritten text documents declines the performance of OCR system. In this paper, we present a technique to solve these problems. The proposed algorithm is based on mid-point detection. The algorithm deals with these problems and gives effective results 90% in case of overlapped lines and 94% accurate results for segmentation of connected components between neighboring lines. This paper also provides a review on major problems in line segmentation that decreases the accuracy of recognition system. The proposed method has achieved 93.05% accuracy in text line segmentation.

Keywords

Line Segmentation, Gurumukhi, Skew Lines, Overlapped Lines, Connected components.

1. INTRODUCTION

Segmentation is a pre-processing phase of optical character recognition. Segmentation of offline handwritten documents is a crucial phase in OCR process. OCR is a technique to encode the offline handwritten as well as printed documents. Before recognition, segmentation is performed on the scanned handwritten documents. Results of OCR mostly depend upon effective line segmentation. Incorrect segmentation of lines affects the recognition. There are different scripts available to write different languages. Different properties of languages and variations in writing styles of different writers may complicate the process of segmentation. Many researchers have developed different methods to solve different problems for Gurumukhi handwritten text documents and provide efficient results. In this paper, we propose an algorithm to solve overlapping lines and connected component problems in handwritten documents. Simpel Jindal et al. [1], have proposed an algorithm based on the concept of vertical stripes and Text blocks for segmenting lines in Gurumukhi handwritten manuscripts. Firstly, they found the optimal or average text block size form the division of whole document, and then compare the average text block with others large and small text blocks. The algorithm works efficiently for segmenting heavily overlapping lines and multiple skew documents. Namisha Modi, et al. [2], presented a technique based on average line height to segment lines in Gurumukhi handwritten scripts. This technique works efficiently in overlapping line segmentation with 75.78% accuracy. Naresh Kumar Garg, et al. [3], has proposed a technique based on cut text method (CTM).The proposed technique uses the concept of difference between header and

base lines for segmenting handwritten Hindi text documents. M.K.Jindal, et al. [4], have proposed an algorithm based upon horizontal projection profile and average line height to segment Gurumukhi printed scripts and obtain 99% accuracy in Printed documents. Naresh Kumar Garg, et al. [5], presented an algorithm based upon header line and base line detection and gives accurate results for uniform and non-uniform skewed lines. Ashu Kumar, et al. [6], has proposed an algorithm based upon piecewise projection profile for segmenting handwritten Gurumukhi text. Saiprakash Palakollu, et al. [7], has proposed the same concept of average line height and horizontal projection profile. Row height is calculated by replacing the rows containing pixels with black pixels. Nallapareddy Priyanka, et al. [8], have proposed a technique based upon run length smearing technique and solve the problem of overlapping and touching component in different Indian scripts.

This paper proposes an algorithm for line segmentation. The algorithm is based on mid-point detection concept and is implemented in two steps. In the first step, gaps between consecutive lines are detected and find the mid-value of these gaps to segment the lines. In the second step, connected components and overlapped lines are detected in the given text document. These text lines are segmented by applying mid-detection approach.

2. PROPERTIES OF GURUMUKHI SCRIPT

As Devnagari script is used to write different Indian language like Hindi, Sanskrit. Similarly, Gurumukhi script is used to write Punjabi language. Gurumukhi script is very popularly used language of Punjab state. Gurumukhi language is the invention of Sikh Gurus. Religious persons use Gurumukhi script to write old manuscripts. This script is 12th widely used in the world. Gurumukhi script uses the writing style from top to bottom and left to right. There is no concept of case sensitive letters in Gurumukhi script. Gurumukhi script composed of forty one consonants, twelve vowels and three half characters. The characters of Gurumukhi scripts are connected with the help of header line. Many characters of Gurumukhi script have similar features. A single word of Gurumukhi is composed of three different zones: upper zone, middle zone and lower zone. In Gurumukhi script, text line is partitioned in different portions. Upper line connects the upper modifiers or upper zone characters with middle zone. Lower line connects the lower modifiers to the middle zone characters. Header line connects the upper parts of different character bodies and the header line is detected with the help of maximum black pixels. Base line connects the lower parts of different character bodies and the base line is detected with the help of minimum black pixels.

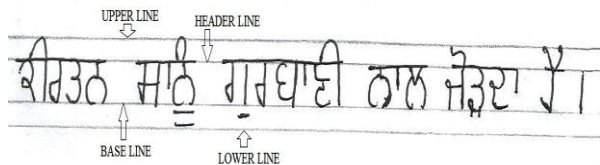


Fig. 1: Different Sections of Text Line

2.1 Major Problems in Line Segmentation

Line segmentation of a handwritten or printed document is one of the major challenges in optical character recognition. There are various problems in the segmentation of handwritten documents, for example, structural properties of the script, varying writing styles of different persons and uneven spaces between consecutive lines. In Gurumukhi handwritten documents, line segmentation is a complex task because of irregularities in geometrical properties such as line height, width, and distance in between line.

2.1.1 Overlapped lines

The problem of overlapping arises in both printed and handwritten documents. Overlapping here means upper and lower modifiers and characters of one line may cover the modifiers or characters of next line. In handwritten documents, problem of overlapping lines becomes more complex due to presence of unequal spaces between neighboring lines.

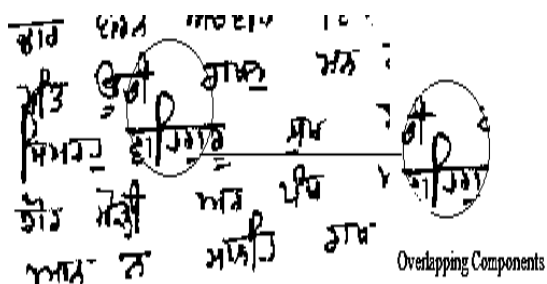


Fig. 2: Text Containing Overlapping Lines

2.1.2 Connected Components

Connected components implies when lower modifier of one character connects with upper modifiers of next line. The problem of connected components arises due to handwriting styles of different writers.

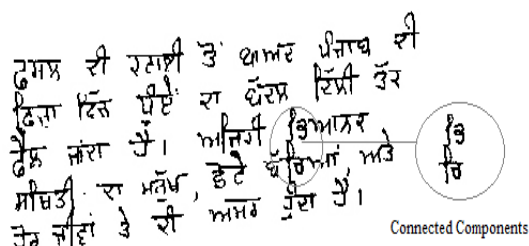


Fig. 3: Text Containing Connected Components

2.1.3 Skewed Lines

Skewness present in text lines due to hasty writing of some authors. Skew means slanting position of text line. Skewness

may be positive, negative and non-uniform. The skewed text makes the document analysis and segmentation more complex. Skewness may be positive, negative and non-uniform.

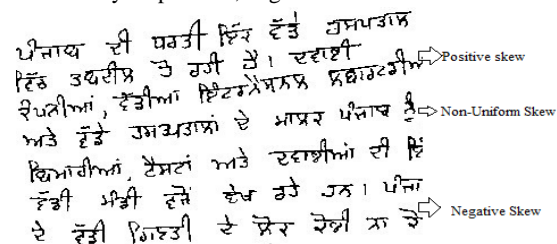


Fig. 4: Skewed Text Lines Document

3. LINE SEGMENTATION

Line segmentation is a technique to extract lines from a scanned document. Line segmentation techniques basically categorized into three different categories: Hough Transform [10], Projection profiles [4, 6] and Smearing [8]. A lot of research work carried out by various authors on Gurumukhi handwritten documents to extract lines. In this paper, novel approach is used to extract lines from Gurumukhi handwritten documents and try to find the solutions of connected components between lines, skewed lines and overlapped lines. The proposed algorithm is based on Mid-Point Detection concept to segment handwritten text document. Segmentation of lines is done by analysis the gap between lines. This algorithm effectively deals with skewed text, connected components and overlapped lines.

The following assumptions made about the data:

1. The average line height of the text line is between 25 to 30 pixels.
2. The maximum height of text line is 30 pixels.
3. The maximum gap between two consecutive lines is 5 pixels.
4. If the difference between two consecutive mid-values is more than 60 pixels, it must have overlapped text lines or connected components.
5. If the gap between two consecutive lines is less than 5 pixels, then it must have upper or lower modifiers.

3.1 Proposed Methodology

In this paper, a novel technique for text line segmentation is presented. The approach is based on the assumption that two consecutive lines must have some white spaces between them. The following steps are systematically performed on text document to extract lines.

3.1.1 Collect Database

Number of handwritten text documents is collected from different writers. The documents are assumed to bear some kind skew in the header line and having the some sort of problems such as connected components and overlapped lines.

3.1.2 Pre-Processing

The document image is further scanned at 200dpi. Further, the gray scale image converted into binary format. After binarization, Proper value of threshold must be specified.

3.1.3 Division into Stripes

The text document image is divided into number of non-overlapping stripes to detect the header line in skewed text

document. Horizontal projections are used to detect gaps between the consecutive lines. The mid-values of these gaps are calculated to segment non-overlapping lines.

3.1.4 Identify Overlapped Lines and Connected Components

The proposed algorithm works on the assumption that average height of text line is 30 pixels. If the difference between consecutive mid-values is more than 70 pixels, then the text must have overlapped lines or connected components.

3.1.5 Segmentation of Text Lines

After detecting the overlapped lines and connected components in text document, these text lines are segmented using Mid-Detection approach.

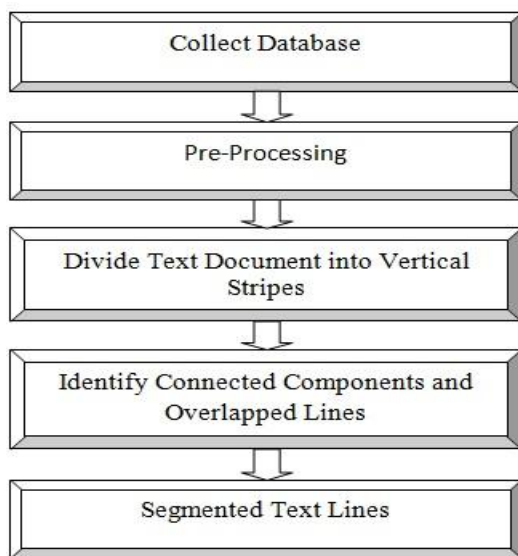


Fig. 5: Block Diagram of Proposed method

3.2 Algorithm

Step 1: Divide the document into vertical stripes by assuming the strip size equal to 100 pixels.

Step 2: Using the concept of Horizontal Projections, find the white spaces between the consecutive lines.

Step 3: Calculate the mid-points of these white spaces to segment the consecutive non-overlapping lines.

Step 4: Identify the overlapped lines and connected components by analyzing the difference between two mid-points, which is explained in further sub steps:

Step 4.1: By taking the assumption that average line height is of 30 pixels.

Step 4.2: If the difference between two mid-points is greater than 70 pixels, then it is assumed that document contains some overlapped lines and connected components.

Step 5: Segment the overlapped lines and connected components by calculating the mid-value from previous and next mid-point value.

Step 6: Draw horizontal black lines to represent the segmentation.

Step 7: Repeat same process for each strip.

4. RESULTS

The proposed algorithm has been implemented on many handwritten Gurumukhi text documents. Most of the

Gurumukhi handwritten documents contain skewed, overlapping lines and connected components. Using this algorithm, these problems are sorted out up to high extent. Some results of proposed algorithm shown below:

4.1 Segmentation of simple lines

These type of documents taken is containing simple lines that mean document comprise non-uniform skew with reasonable spaces between lines.

ਗੁਰਮੁਖੀ ਦਾ ਮੁੱਖ ਉਦੇਸ਼ ਖੜੂ ਦੀ ਖ਼ਾਖ਼ੀ ਤੋਂ ਜੀਨ ਨਿਰਧਾਰਤ
ਦੀ ਖਤਰੇ ਵੀਰ ਵਰੀਦਿਆ ਜੀਵ ਤੇ ਜੀਨ ਉਦੇਸ਼ ਤੋਂ ਘੱਟ
ਸਾਹਮਣੇ ਹੋਣਾ ਹੈ, ਜਿਹਦੇ ਦੁਨਿਆਵੀ ਦੇਸ਼-ਭਾਗਾਂ ਵਿੱਚ ਹੀ ਉਥਰ
ਦੇ ਨਹੀਂ ਹੋ ਸਕੇ। ਗੁਰਮੁਖੀ ਅਤੇ ਅਰਥਸ਼ਾਸਤਰ ਵਿੱਚ ਇਸ ਯੱਥ ਤੋਂ
ਵੱਡਾ ਅੰਤਰ ਹੈ। ਗੁਰਮੁਖੀ ਦਾ ਵਿਸ਼ਵੀਕਰਣ ਸਮਾਜਵਾਦ ਵਾਂਗ
ਜਿਹਦੇ ਪਰਾਭਵ ਤੌਰ 'ਤੇ ਜੀਨਾਂ ਨਹੀਂ ਹੈ। ਗੁਰਮੁਖੀ ਪਾਠਸ਼ਾਲਾ
ਅੱਖਾਂ ਦੀ ਆਰਾਮ ਅਤੇ ਸਹੀਦੇ ਦੇਣ ਵਾਲੇ ਕਾਰਜਕ ਵਰਤਾਵਾਂ ਹੈ।
ਇਹ ਦੋਹਾ ਵੀਆਂ ਵੱਖ-ਵੱਖ ਸੋਝਾਂ ਨੂੰ ਪੂਰਾ ਰੱਖਣ ਦਾ ਸਿੱਕਾਂਤ
ਹੈ। ਇਹ ਦਾ ਮਤਲਬ ਸੀਏ ਦੇ ਵੀ ਸੋਝਾਂ ਨੂੰ ਖੁਲ੍ਹੇ ਰੱਖਣਾ ਹੈ,
ਜਿਸ ਵਲੀ ਮੱਥਾਂ ਨੇ ਜਿਹਦੇ ਵਰੀਦਿਆ ਆਪਣਾ ਇਹ ਨਿਰੰਜਨ ਨਾਮ
ਜੋੜੀ ਹੋਵੇ।

Fig. 6: Input Simple Text Line Document

ਗੁਰਮੁਖੀ ਦਾ ਮੁੱਖ ਉਦੇਸ਼ ਖੜੂ ਦੀ ਖ਼ਾਖ਼ੀ ਤੋਂ ਜੀਨ ਨਿਰਧਾਰਤ
ਦੀ ਖਤਰੇ ਵੀਰ ਵਰੀਦਿਆ ਜੀਵ ਤੇ ਜੀਨ ਉਦੇਸ਼ ਤੋਂ ਘੱਟ
ਸਾਹਮਣੇ ਹੋਣਾ ਹੈ, ਜਿਹਦੇ ਦੁਨਿਆਵੀ ਦੇਸ਼-ਭਾਗਾਂ ਵਿੱਚ ਹੀ ਉਥਰ
ਦੇ ਨਹੀਂ ਹੋ ਸਕੇ। ਗੁਰਮੁਖੀ ਅਤੇ ਅਰਥਸ਼ਾਸਤਰ ਵਿੱਚ ਇਸ ਯੱਥ ਤੋਂ
ਵੱਡਾ ਅੰਤਰ ਹੈ। ਗੁਰਮੁਖੀ ਦਾ ਵਿਸ਼ਵੀਕਰਣ ਸਮਾਜਵਾਦ ਵਾਂਗ
ਜਿਹਦੇ ਪਰਾਭਵ ਤੌਰ 'ਤੇ ਜੀਨਾਂ ਨਹੀਂ ਹੈ। ਗੁਰਮੁਖੀ ਪਾਠਸ਼ਾਲਾ
ਅੱਖਾਂ ਦੀ ਆਰਾਮ ਅਤੇ ਸਹੀਦੇ ਦੇਣ ਵਾਲੇ ਕਾਰਜਕ ਵਰਤਾਵਾਂ ਹੈ।
ਇਹ ਦੋਹਾ ਵੀਆਂ ਵੱਖ-ਵੱਖ ਸੋਝਾਂ ਨੂੰ ਪੂਰਾ ਰੱਖਣ ਦਾ ਸਿੱਕਾਂਤ
ਹੈ। ਇਹ ਦਾ ਮਤਲਬ ਸੀਏ ਦੇ ਵੀ ਸੋਝਾਂ ਨੂੰ ਖੁਲ੍ਹੇ ਰੱਖਣਾ ਹੈ,
ਜਿਸ ਵਲੀ ਮੱਥਾਂ ਨੇ ਜਿਹਦੇ ਵਰੀਦਿਆ ਆਪਣਾ ਇਹ ਨਿਰੰਜਨ ਨਾਮ
ਜੋੜੀ ਹੋਵੇ।

Fig. 7: Resultant Segmented Simple Text Line Document

4.2 Segmentation Connected Components

Connected lines are having upper and lower modifiers of two consecutive lines touches with each other i.e. modifiers of two consecutive lines are linked with each other. Because of this, segmentation of document makes difficult.

ਦੁਖਪਾਲੀ ਨਗਰ ਤੋਂ ਇੱਕ ਆਈ ਜਿਸਦਾ ਨਾਮ ਸ਼ਾਇਬ ਵੇਰ ਸੀ
ਸੇਵਾ ਰਹੀ ਹੁੰਦੀ ਸੀ। ਇਸ ਆਈ ਵੀਆਂ ਚਾਰ ਖੜੂਵੀਆਂ
ਖੁੱਲ੍ਹੇ ਨਗਰ ਵਿੱਚ ਹੀ ਵਿਆਹੁੰਦੀ ਸਨ। ਇਹ ਦੋਹਾ ਮਾਤਾ
ਸ਼ਾਇਬ ਵੇਰ ਇੱਕ ਖੜੂਵੀ ਨੂੰ ਸੈਂਦੇ ਬਾਬਾ ਜੀ ਪਾਸ
ਹਜ਼ਾਰ ਹੋਈ ਅਤੇ ਬੇਨਤੀ ਕੀਤੀ ਕਿ ਅਗਲਾ ਇਸ ਮੇਰੀ
ਖੜੂਵੀ ਦਾ ਨਾਮ ਹਰ ਵੇਰ ਹੈ। ਹੁਣ ਸੁਰੱਖੇ ਵੀ ਇਸਨੂੰ
ਪਸੰਦ ਨਹੀਂ ਰਹੇ, ਮੇਰੇ ਵੇਰ। ਵੇਰੀ ਬੱਚਾ ਤੇ ਜਾਣੇ

Fig. 8: Non- uniform Skew Text Lines with Connected Components

ਦਿਖਾਈ ਨਗਰ ਤੇ ਇੱਕ ਆਈ ਜਿਸਦਾ ਨਾਮ ਸ਼ਾਇਦ ਦੋਰ ਸੀ
ਮੇਰਾ ਰਹੀ ਜੁੱਤੀ ਸੀ। ਇਸ ਆਈ ਦੀਆਂ ਚਾਰ ਖੜਕੀਆਂ
ਭੁੱਲੇ ਨਗਰ ਵਿੱਚ ਹੀ ਵਿਆਹੀਆਂ ਸਨ। ਇਹ ਵਾਰ ਸਾਡਾ
ਸਾਹਿਬ ਦੋਰ ਵਿੱਚ ਲੜਕੀ ਨੂੰ ਸੈਂ ਦੇ ਬਾਬਾ ਜੀ ਪਾਸ
ਪਹੁੰਚ ਗਈ ਅਤੇ ਚੇਨੜੀ ਵੀੜੀ ਵਿੱਚ ਮਹਾਗਜ਼ ਵਿਸ਼ ਮੇਰੀ
ਲੜਕੀ ਦਾ ਨਾਮ ਹਰ ਦੋਰ ਤੇ। ਤੁਝ ਸੁੱਚੇ ਵੀ ਇਸਨੂੰ
ਪਸੰਦ ਨਹੀਂ ਰਹੇ ਮੇਰਾ ਦੋਰ। ਦੋਈ ਕੱਚਾ ਤੇ ਜਾਣੇ

Fig. 9: Segmented Text Lines Containing Connected Components

4.3 Segmentation of overlapped lines

Overlapped lines present in both printed and handwritten documents. Lower zone and upper zone or lower zone and middle zone characters are being the part of neighboring line, thus cause the problem of overlapping lines.

ਦੁਨੀਆ ਦੀ ਨਗਰ ਨਾਲ ਦੋਖਿਆ ਜਾਵੇ, ਤਾਂ ਇਹ ਇੱਕ ਚੜ੍ਹਤ
ਜੀ ਸਖਤ ਸਿਮਿਤਗਤ ਸੀ ਪਰ ਆਪ ਜੀ ਤੇ ਦੁਨੀਆਂ ਦੀ
ਜ਼ੋਰ-ਸ਼ਾਨ ਦੀ ਜ਼ਬਾ ਹੀ ਪੁਰਾਣੇ ਸੀਤੇ ਕਿਸਾ ਤਰੀਕੇ ਖੁਰਾਈ ਬਚਨ
ਤੇ ਫੁੱਲ ਚੜ੍ਹਦੇ। ਆਪਣੇ ਜੱਥੀ ਹੋਈ ਦੀ ਇੱਟ-ਚਿੱਟ ਰਹੇ
ਉਸਦਾ ਸਾਗ ਮਲਥਾ ਟਰੱਕਾਂ ਅਤੇ ਟਰੈਕਟਰ-ਕੁਸ਼ੀਆਂ ਵਿੱਚ
ਜੁਰੇ ਆਪਣੇ ਪਿਆਰੇ ਖੁਰਾ ਦੀ ਪੀੜਤਰ ਹੜ੍ਹੀ ਵਿੱਚ
ਸੈਂ ਆਏ।

Fig. 10: Document Containing Overlapping Skewed Text Lines

ਦੁਨੀਆ ਦੀ ਨਗਰ ਨਾਲ ਦੋਖਿਆ ਜਾਵੇ, ਤਾਂ ਇਹ ਇੱਕ ਚੜ੍ਹਤ
ਜੀ ਸਖਤ ਸਿਮਿਤਗਤ ਸੀ ਪਰ ਆਪ ਜੀ ਤੇ ਦੁਨੀਆਂ ਦੀ
ਜ਼ੋਰ-ਸ਼ਾਨ ਦੀ ਜ਼ਬਾ ਹੀ ਪੁਰਾਣੇ ਸੀਤੇ ਕਿਸਾ ਤਰੀਕੇ ਖੁਰਾਈ ਬਚਨ
ਤੇ ਫੁੱਲ ਚੜ੍ਹਦੇ। ਆਪਣੇ ਜੱਥੀ ਹੋਈ ਦੀ ਇੱਟ-ਚਿੱਟ ਰਹੇ
ਉਸਦਾ ਸਾਗ ਮਲਥਾ ਟਰੱਕਾਂ ਅਤੇ ਟਰੈਕਟਰ-ਕੁਸ਼ੀਆਂ ਵਿੱਚ
ਜੁਰੇ ਆਪਣੇ ਪਿਆਰੇ ਖੁਰਾ ਦੀ ਪੀੜਤਰ ਹੜ੍ਹੀ ਵਿੱਚ
ਸੈਂ ਆਏ।

Fig. 11: Segmented Overlapped Text Lines

Table 1. Results of Text Line Segmentation

Type of Sample Images	Total No. of Lines	Correctly segmented Lines	Accuracy
Simple Text Documents	100	98	98%
Text Documents having Connected Components	250	235	94%
Text Document having Overlapped Lines	240	216	90%
Total	590	549	93.05%

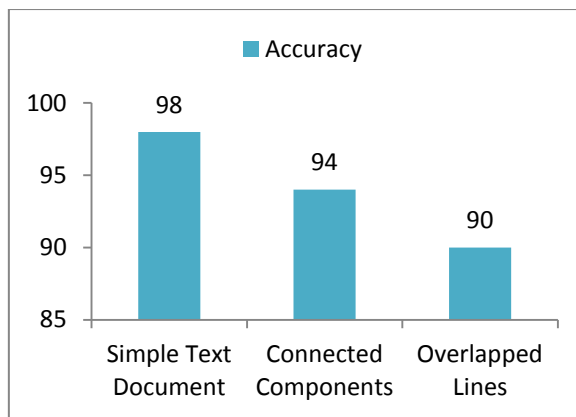


Fig. 12: Segmentation Results

Table 2. Comparison of Proposed method with Existing Techniques

Ref.	Technique used	Type of input	Accuracy
Namisha et. al [2]	Average height Detection	Handwritten Gurumukhi Text Line Detection of overlapped and skew Text Lines	75.78%
M. Kumar et. al [4]	Horizontal Projection Profile	Overlapped Text Lines in Printed Gurumukhi Scripts	99%
Naresh et. al [5]	Header and Base line Detection	Simple Hindi text	91.5%
Proposed Method	Detection of Mid Point from Existing Gap between Lines	Gurumukhi Handwritten Scripts, Connected components, overlapping and Skew Lines	93.05%

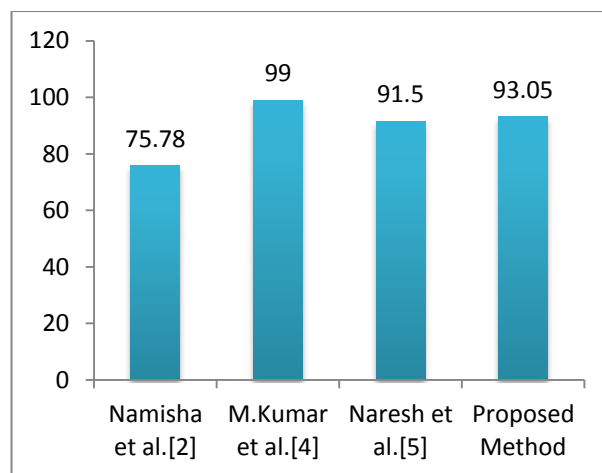


Fig. 13: Comparison analysis

5. CONCLUSION AND FUTURE WORK

Segmentation of Gurumukhi handwritten text is very complicated task due to skewed, overlapped and connected components present between lines. The complexity arises in handwritten documents because of writing style and structural properties of language. In this paper, we have used the proposed algorithm based on mid-point detection to extract lines from Gurumukhi handwritten documents. The proposed algorithm is used to segment skewed lines, overlapped lines and connected components between the neighboring lines. This technique provides effective results for text line segmentation. This algorithm distinguishes the concept of segmenting connected components from overlapped lines.

In future work, we are trying to develop same algorithm for Devnagari handwritten scripts. Although, we received efficient result using this technique yet some problem remains there i.e. during the segmentation of overlapping lines, some of touching characters are getting degrade from their actual features or existence. This may create difficulty during recognition. Further work can be extended to solve this problem.

6. REFERENCES

- [1] Simpel Jindal, Gurpreet Singh Lehal, "Line Segmentation of Gurumukhi Manuscripts", December 2012.
- [2] Namisha Modi, Khushneet Jindal, "Text line detection and segmentation in Handwritten Gurumukhi Scripts", International Journal of Advanced Research in Computer Science and Software Engineering, vol.3, Issue 5, PP:1075-1080, May, 2013.
- [3] Naresh Kumar Garg, Lakhwinder Kaur, M.K.Jindal, "A New Method for Line Segmentation of Handwritten Hindi Text", IEEE, 2010.
- [4] M.K. Jindal, R.K. Sharma, G.S.Lehal, "Segmentation of Horizontally Overlapping lines in Printed Gurmukhi Script", IEEE, 2006.
- [5] Naresh Kumar Garg, Lakhwinder Kaur, M.K. Jindal, "Segmentation of Handwritten Hindi Text", International Journal of Computer Applications, Vol.1, No.4, 2010.
- [6] Ashu Kumar, Simpel Rani Jindal, "Segmentation of Handwritten Gurumukhi text into Lines", International Conference on recent advances and Future Trends in Information Technology (iRAFIT2012), pp. 13-17,2012.
- [7] Saiprakash Palakollu, Renu Dhir, Rajneesh Rani, "Handwritten Hindi Text Segmentation Techniques for Lines and Characters", World congress on Engineering and Computer Science, Vol. 1, 2012.
- [8] Nallapareddy Priyanka, Srikanta Pal, Ranju Mandal, "Segmentation of handwritten Gurumukhi Text into Lines", IJCA, 2010.
- [9] U.Pal, Sagarika Datta, "Segmentation of Bangla Unconstrained Handwritten Text",IEEE, 2003.
- [10] G. Louloudis, B.Gatos, I. Pratikakis,C. Halatsis, "A Block-Based Hough Transform Mapping for Text Line Detection in Handwritten Documents", Proceedings of the tenth International Workshop on Frontiers in Handwriting Recognition, La Baule,Oct.2006.
- [11] Rajiv Kumar, Amardeep Singh, "Detection and Segmentation of Lines and Words in Gurumukhi Handwritten Text", IEEE,2010.
- [12] M.K. Jindal, R.K. Sharma, G.S.Lehal, "Segmentation of Horizontally Overlapping lines in Printed Indian Script", International Journal of Computational Intelligence Research, Vol.3, No.4, pp.277-286, 2007.
- [13] Munish Kumar, M.K.Jindal, R.K. Sharma, "Classification of Characters and Grading Writers in Offline Handwritten Gurmukhi Script", IEEE, 2011.
- [14] G.S.Lehal, Chandan Singh, Ritu Lehal, "A Shape Based Post Processor for Gurumukhi OCR", IEEE, 2001.
- [15] G S Lehal, Chandan Singh, "A Gurumukhi Script Recognition System", IEEE, 2000.