International Journal of Scientific & Engineering Research, Volume 8, Issue 2, February-2017
ISSN 2229-5518

Feature Extraction for Marathi Compound Character Using Edge Map

Mrs.Snehal S.Golait, Dr.Latesh Malik, Prof.A.Thomas

1Research Scholar ,Department of Computer Science and Engineering, G.H.Raisoni College of Engineering,Nagpur,
2 Professor, Department of Computer Science and Engineering, G.H.Raisoni College of Engineering,Nagpur,
3 Head of Department, Department of Computer Science and Engineering, G. H. Raisoni College of Engineering, Nagpur

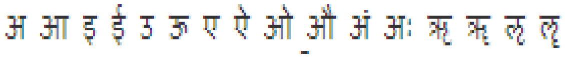
Abstract— Feature extraction is one of the important phase of handwritten Script Identification. It involves measuring those features of the input pattern are relevant to classification. This paper provides a feature extraction for Marathi Compound Character using syntactical methods. Aside from the similarity of Character researcher founds difficulty to find features of characters. The recognition is carried out using structural and statistical feature extraction and multistage classification scheme. The Proposed system used horizontal projection method for line segmentation, vertical projection method for word segmentation and minutiae detection algorithm is used for compound character segmentation. An input image is preprocessed and segmented. In this paper proposed a novel technique for feature extraction of segmented character called edge map. Edge map is a structural feature extraction technique for extracting the features of segmented character. With this technique able to get near about 95% accuracy in recognition process.

Index Terms— Segmentation, Minutiae detection , Binarization, edge map, feature extraction, wavelet Decomposition , Erosion.

u

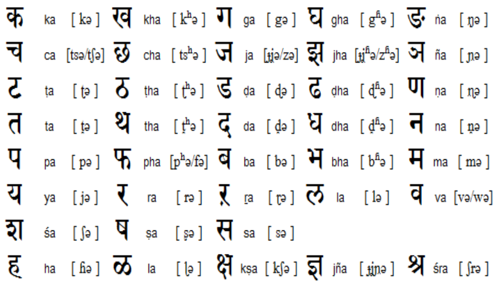
1 INTRODUCTION

Handwritten character recognition is an important field of Optical Character Recognition (OCR). The objective of OCR is automatic reading of optically sensed document text materials to translate human readable characters to machine understandable codes. OCR is popular for its various application potentials in banks, library automation post-offices and defense organizations. Handwritten character recognition aims at converting handwritten characters in images into text that can be stored, edited or can be converted into speech. This field of research finds applications in various areas that aim in automation so as to reduce the human efforts like postal automation bank automation form filling etc. Handwritten character recognition for Indian scripts is quite a challenging task due to several reasons. One of the Indian Script is Devnagari Script. Devnagari is third most widely used script, used for Indian languages such as Hindi, Sanskrit, Nepali and Marathi, and is used by more than 500 million people. Unconstrained Devnagari writing is more complex than English language due to the possible variations in the shape, number and direction of the constituent strokes. Devnagari script has 50 characters which can be written as individual symbols in a word. Devnagari Character recognition is complicated process due to presence of multiple conjuncts, loops, lower and upper modifiers and the number of disconnected and multistroke characters, in a word where all characters are connected through Shirorekha. OCR is further complicated by compound characters that makes character separation and identification is very difficult. Marathi script is one of the script of Devanagri Script. Marathi belongs to the group of Indo-Aryan languages which are one of the part of Indo-European languages, all of which can be traced back to a common root. Among the Indo-Aryan languages, Among the Indo-Aryan languages, Marathi is the southern-most language. All of the Indo-Aryan languages originated from Sanskrit. Prakrit languages simpler in structure and originated from Sanskrit. Some of the Prakrit languges were Saurseni, Maharashtri and Magadhi. Marathi is said to be a descendent of Maharashtri which was the Prakrit spoken by people residing in the region of Maharashtra. The script currently used in Marathi is called 'Balbodh' which is a modified version of Devnagari script. Earlier, another script called 'Modi' used till the time of the Peshwas. This script was introduced by Hemadpanta, a minister in the court of the Yadava kings of Devgiri. This script looked more like today's draviDian scripts and offered the advantage of greater writing speed because the letters could be joined together. Today only the Devanagari script is used which is easier to read. Marathi script derived from Devanagari, is an official language of Maharashtra. It is the 4th most spoken language in India and 15th most spoken language in the world[1]. Marathi script consists of 16 vowels and 36 consonants making 52 alphabets[1]. Marathi is written from left to right. It has no lower and upper case characters. Each character has a horizontal line at the top called as the header line. The header line joints the characters in a word. The vowels, consonants and modifiers in Marathi language shown in fig 1, 2 and 3.



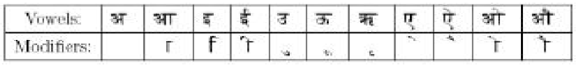
अ आ इ ई उ ऊ ए ऐ ओ औ अं अः ऋ ॠ ऌ ॡ

Fig 1. Vowels in Marathi Script



क	ka	[kə]	ख	kha	[kʰə]	ग	ga	[gə]	घ	gha	[gʱə]	ङ	ṅa	[ŋə]
च	ca	[tsə/tʃə]	छ	cha	[tsʰə]	ज	ja	[ʤə/zə]	झ	jha	[ʤʱə/zʱə]	ञ	ña	[ɲə]
ट	ṭa	[ʈə]	ठ	ṭha	[ʈʰə]	ड	ḍa	[ɖə]	ढ	ḍha	[ɖʱə]	ण	ṇa	[ɳə]
त	ta	[t̪ə]	थ	tha	[t̪ʰə]	द	da	[d̪ə]	ध	dha	[d̪ʱə]	न	na	[n̪ə]
प	pa	[pə]	फ	pha	[pʰə/fə]	ब	ba	[bə]	भ	bha	[bʱə]	म	ma	[mə]
य	ya	[jə]	र	ra	[rə]	ऱ	ṟa	[ɽə]	ल	la	[lə]	व	va	[və/wə]
श	śa	[ʃə]	ष	ṣa	[ʂə]	स	sa	[sə]						
ह	ha	[ɦə]	ळ	ḷa	[ɭə]	क्ष	kṣa	[kʃə]	ज्ञ	jña	[ʤɲə]	श्र	śra	[ʃrə]

Fig 2. Consonants In Marathi Script



Vowels:	अ	आ	इ	ई	उ	ऊ	ऋ	ए	ऐ	ओ	औ
Modifiers:		ा	ि	ी	ु	ू	ृ	े	ै	ो	ौ

Fig 3. Modifiers In Marathi Script

Marathi Script have a character set with a large number of characters in it. The shape of the characters is complex and they may have modifiers, present above, below or in line with the character. The modifiers and the vowels that change their shape when they jointed to the consonants. Moreover, some character pairs are almost similar to each other that make them quite difficult to classify. Another reason is presence of compound characters in some scripts like Bangla, Devanagari etc. where two or more consonants are joint together to form a special character as shown in fig 4.

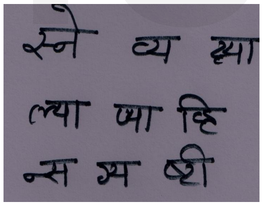


Fig 4. Samples of Marathi Compound Characters

Further handwritten characters shows a large variation in the basic shapes of the characters due to the pen ink, pen width, accuracy of acquisition device, the stroke size and location in the character. Physical and mental situation of the writer affects the writing style and in turn the recognition accuracy to a considerable extent. Research for Indian offline character recognition first started with the recognition of printed characters and then extended to the recognition of handwritten numbers. Some research is also devoted towards segmentation of touching characters, recognition of hand- written compound characters and words in various Indian scripts.

OCR work started on printed Devanagari script in early 1970s. Veena Bansal and R.M.Sinha worked on printed Devanagari text. First system for hand-written numeral recognition of Devanagari script was proposed by R. Bajaj , P. M. Patil and T.R. Sontakke also presented an algorithm for handwritten Devanagari numeral recognition which used concept of scaling, rotation and translation invariant. U. Pal proposed a system for off-line handwritten character recognition of De vanagari using directional information as features. A technique for accuracy improvement of Devanagari character recognition system was proposed by U. Pal using two features based on directional and curvature information in the characters and applied to the classifiers support vector machines and modified quadratic discriminant function. A comparative study of various features and classifiers used for handwritten Devanagari character recognition was proposede by U. Pal. A multi-feature, multi-classifier scheme for handwritten Devanagari Marathi characters is proposed by S. Shelke and S. Apte [1]. Work on handwritten Bangla compound characters is carried out by U. Pal.Recently, some piece of work are found on handwritten Marathi compound characters. In this paper, proposed algorithm for segmentation of Marathi compound character and applying syntactical feature extraction technique on it.

2 PROPOSED APPROACH

The proposed system consist of following stages of OCR which includes preprocessing steps and recognition step. The preprocessing steps Shown in Fig 5.

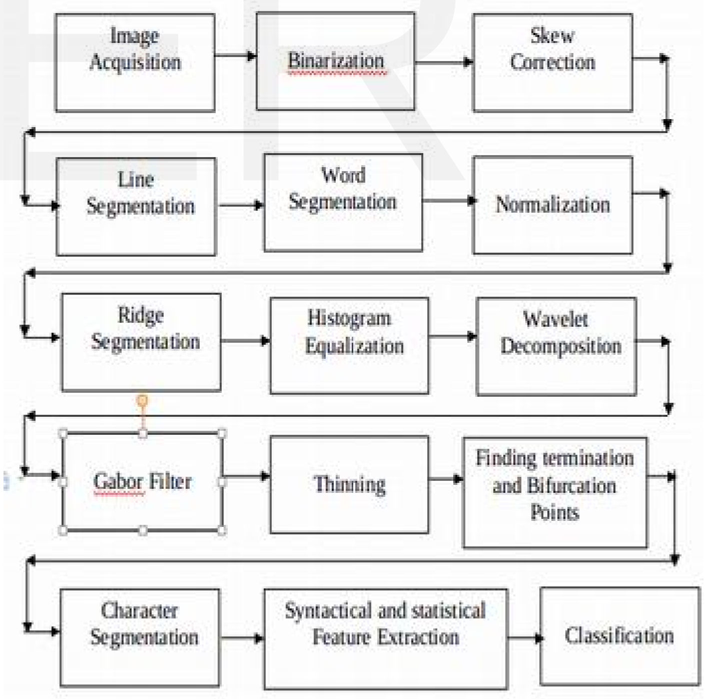


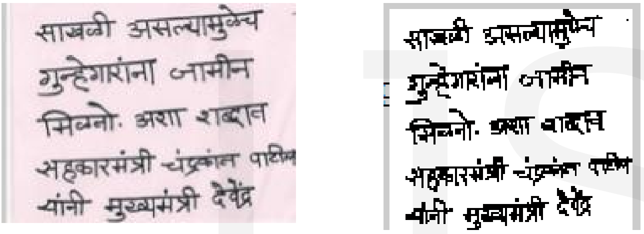
Fig 5. Flow for Compound Character Recognition

Image Acquisition

Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever processes need to occur afterward. Here we can use hardware as Scanner. Performing image acquisition in image processing is always the first step in the workflow sequence. The Input Image that we get is completely unprocessed. One of the ultimate goals of this process is to have a source of input that operates within such controlled and measured guidelines that the same image can reproduced under the same conditions so anomalous factors are easier to locate and eliminate.

Binarization

Binarization means digitization of image. Binarization of an Image is shown in fig 6. Binarize an Image based on the threshold value. Thresholding is an image processing technique for converting a color or gray scale image to a binary image based upon a threshold value. If a pixel in the image has an intensity value less than the threshold value, set the corresponding pixel in the resultant image to black. Otherwise, if the pixel intensity value is greater than or equal to the threshold intensity, the resulting pixel is set to white. Thus used an image with only 2 colors, black (0) and white (255) .

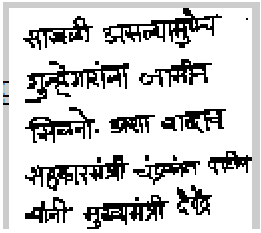


a) Input Image b) Binarized Image

Fig 6. Binarization of Input Image

Skew Correction

When a document is fed to the scanner either mechanically or by a human operator, a few degrees of skew is unavoidable. Skew angle is the angle that the text lines in the digital image makes with the horizontal direction. Image in the Fig 7 a) is with skew. Skew estimation and correction are important preprocessing steps of document layout analysis and OCR approaches. One of the popular skew estimation techniques is based on projection profile of the documents. The horizontal or vertical projection profile is a histogram of the number of black pixels along horizontal or vertical scan lines. For a script with horizontal text lines, the horizontal projection profile will have peaks at text line positions and thorough at positions in between successive text lines. To determine the skew in document, the projection profile is used at a number of angles and for each angle, a measure of difference of peak and through height is made. The maximum amount of difference corresponds to the best alignment with the text line direction which, in turn, determines the skew angle.

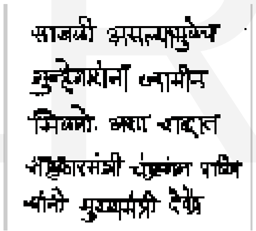


(a)

For Skew Correction use a very straightforward approach . For a detected skew angle rotate the document image using the following formula. In MATLAB function imdeskew() is used to deskew the Input Image. The Image with removal of Skew is shown in fig 7b).

$$X = x \cos \theta + y \sin \theta$$

$$Y = y \sin \theta - x \cos \theta$$



(b)

Fig 7: a) Input Image with Skew b) Deskew Image

Line Segmentation

Segmentation process is segmenting the text documents into lines, also called as line segmentation is shown in fig 8. First we need to calculate header lines and base lines for the Line segmentation. Header lines are those with maximum number of black pixels and base lines are rows with minimum number of black pixels. Finding header line is a challenge task because of skew in header line. Now a day's most of the researchers are detecting the header line by finding the row with maximum pixel density, but it cannot work for skew variable text. This method gives good results for uniform and non-uniform skewed lines.

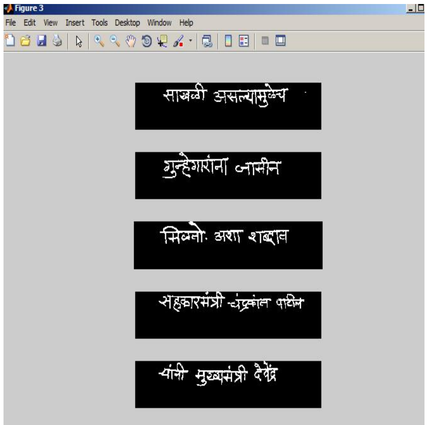
Determining Location of Text Line

1) For each scan line the proposed system will check all the pixels on that scan line.

2) If for particular pixel intensity value is 1,then system will store that scan line number.

3)For the stored scan line position the system will check subsequent scan lines till a scan line containing no black pixels is obtained.

4) Then the dimension of the text line will be found from stored scan line positions.



साखळी असल्यामुळेच

गुन्हेगारांना जामीन

मिळतो. अशा शब्दात

सहकारमंत्री चंद्रकांत पाटील

यांनी मुख्यमंत्री देवेंद्र

Fig 8. Line Segmentation Word Segmentation

Word Segmentation

Word segmentation is easier task as compared to line segmentation and character segmentation. Space between two words is generally more than two or three pixels. Words segmentation is done by the projection based method. For word segmentation use the following algorithm. The result of word segmentation is shown in fig 9.

Determining Location of Word in Text Line

1) For each vertical scan line all horizontal pixels will be examined.
2) If any pixel having intensity value 1 is found, note its position.
3) Subsequent scan lines are checked till we get a scan line with all pixels having intensity value 0. Position of that scan line will be noted.
4) Position 1 and 2 noted in above two steps determine the location and boundary of the word to be segmented.

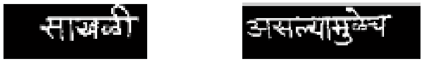


Fig 9. Word Segmentation

Normalization

Normalization is one of basic step for pre-processing factors of character recognition. Normally, in normalization the character image is linearly mapped on to a standard plane by interpolation/extrapolation. The position and size of character is controlled such that the width and length of normalized plane are filled. By linear mapping, the character shape is not only deformed but also the aspect ratio changes.

Ridge Segmentation

The need of ridge segmentation is for finding the break points in character. In mathematics the ridges of a smooth function of two variables are a set of curves whose points are local maxima of the function in one dimension. These notation captures the intuition of geographical ridges. For a function of N variables, its ridges are a set of curves whose points are local maxima in N-1dimensions. Correspondingly, the notion of valleys for a function can be defined by replacing the condition of a local maximum with the condition of a local minimum. The union of valley sets and ridge sets, together with a related set of points called the connector set form a connected set of curves that partition, intersect, or meet at the critical points of the function.

Histogram Equalization

Apply the Histogram equalization for adjusting image intensities to enhance contrast. In Matlabhisteq() function is used to enhanced the image.

Wavelet Decomposition

The wavelet decomposition is merely done for the lossless reduction of the size of an image[7]. DWT decomposes the signal into mutually orthogonal set of wavelets. DWT decomposition is used for finding the pixels at horizontal, vertical and in diagonal direction. In CWT, the wavelets are not orthogonal and the data obtained by this transform are highly correlated. The four values of the decomposition are shown in fig 10.

CA : Accurate Value
CD : Dimension value
CV : Vertical Value
CH : Horizontal Value

The four divisions in the image was the same in all level of the decomposition. These four divisions of the decomposition levels have their own characteristics and preserved the values of the image. The values are

LL: Smoothing of original image
LH : Preserves edge at horizontal side
HL : Preserves edge at Vertical side
HH : Preserves edge at diagonal side

The original image is decomposed into many levels for that used the dwt function in the MATLAB. The db1 function is used in MATLAB for wavelet transforms and the four level decomposition has been carried out for the characters.

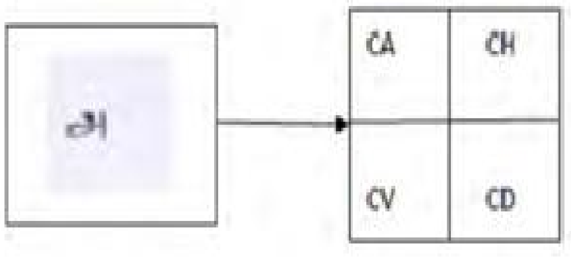


Fig 10. 4 Level Wavelet Decomposition

Gabor Filter

In the literature, there are several image noise removal methods which are applied in spatial, transform, or time-frequency domains. In the spatial domain a small mask is convolved with the image. The mask can be mean, Gaussian or an average filter. In the transform domain, first the image is translated, then is multiplied by a low pass filter and at the end perform the inverse transformation to enhanced the image. In the transform domain, noise in the grey levels of an image contributes heavily to the high frequency components and the most of the image energy is concentrated in the low frequency components. Although, applying a low pass filter to a noisy image in the transform domain reduces the noise, as well as it could eliminate some high frequency components that are not related to noise and weaken sharp transitions like edges. Furthermore, the transforms which perform on the whole image, do not show any spatial information where the frequency components come from. Therefore, noise reduction by low pass filtering in such domains does not preserve the local information of the image. Time-frequency transforms combine time-domain and frequency-domain analysis and allow obtaining a revealing picture of the temporal localization of the signal's spectral components. Due to this problem we consider the Gabor filter as a noise reduction technique.

Thinning

The next step is to thin the processed binary image using the morphological thinning operation. The thinning algorithm removes pixels from ridges until the ridges are one pixel wide[13]. The thinning of an image I by a structuring element J = (J_1, J_2) is given by

$$\mathrm{Thin}(I,J)= I - (I \otimes J)$$

Where, the subtraction is the logical subtraction defined by:

$$X – Y = X \cap \text{Not } Y$$

Finding termination and bifurcation point using Hit or Miss Transform

After thinning apply hit or Miss transform to find the termination and bifurcation points[13]. Results of this algorithm is shown in fig 11(a) and (b). Termination is those pixels in an image which have only one neighbour in 3X3 neighbourhoods. The terminations are given by applying Hit or Miss Transform on I by J as follows:

$$M1 = (I \otimes J)$$

Where, I is the thinned image and J is the sequence of structuring element pairs (J_1, J_2)

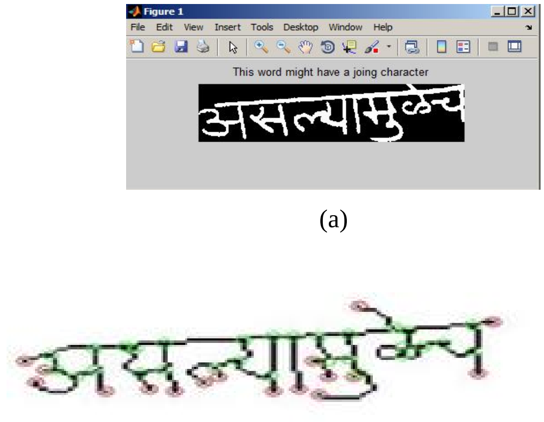
$$I \otimes J = (I \Theta J_1) \cap (I^c \Theta J_2)$$

Ridge bifurcations are those pixels in an image which have only three neighbours in 3X3 neighbourhoods and these neighbours are not adjacent to each other. The minutiae image M2 containing ridge terminations is given by

$$M2 = (I \otimes J)$$

Where, I is the thinned image and J is the sequence of structuring element pairs (J_1, J_2)

$$I \otimes J = (I \Theta J_1) \cap (I^c \Theta J_2)$$



This word might have a joing character

(a)

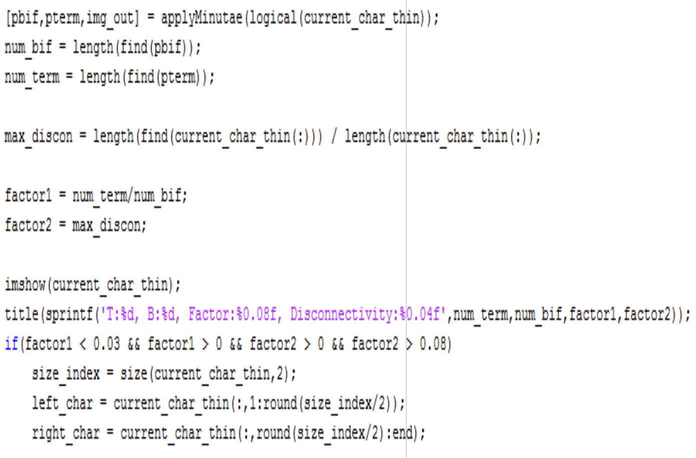
(b)

Fig 11. a) Word having Compound Characters

b) Result of getting termination and bifurcation points

Character Segmentation

Character Segmentation for isolated character is very easy task as compared to compound character separation. For character segmentation need to find number of bifurcation points and number of termination points. After that need to find maximum disconnectivity. The code is shown in fig 12. Results for compound character segmentation is shown in fig 13. a) and b)



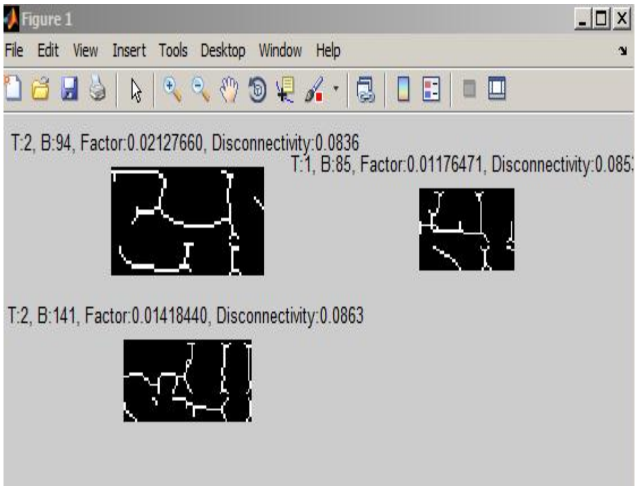
```
[pbif,pterm,img_out] = applyMinutae(logical(current_char_thin));
num_bif = length(find(pbif));
num_term = length(find(pterm));

max_discon = length(find(current_char_thin(:))) / length(current_char_thin(:));

factor1 = num_term/num_bif;
factor2 = max_discon;

imshow(current_char_thin);
title(sprintf('T:%d, B:%d, Factor:%0.08f, Disconnectivity:%0.04f',num_term,num_bif,factor1,factor2));
if(factor1 < 0.03 && factor1 > 0 && factor2 > 0 && factor2 > 0.08)
    size_index = size(current_char_thin,2);
    left_char = current_char_thin(:,1:round(size_index/2));
    right_char = current_char_thin(:,round(size_index/2):end);
```

Fig 12: Code for Character Segmentaion.

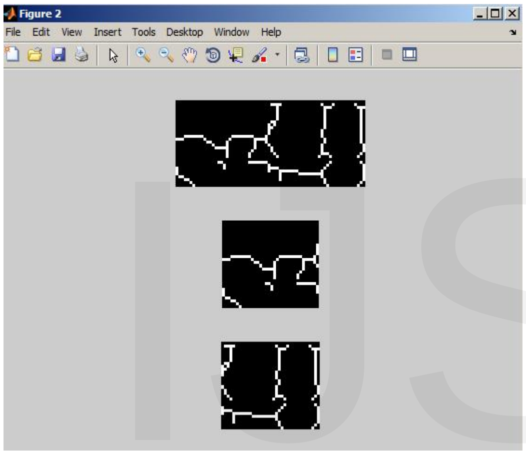


T:2, B:94, Factor:0.02127660, Disconnectivity:0.0836

T:1, B:85, Factor:0.01176471, Disconnectivity:0.085

T:2, B:141, Factor:0.01418440, Disconnectivity:0.0863

(a)



(b)

Fig 13. a) Character segmentation

b) Compound character separation

Feature Extraction

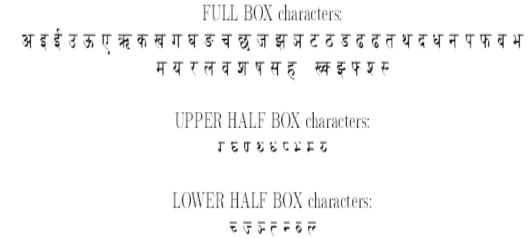
The main aim of feature extraction is to make improvement in the accuracy and speed of the classifies for the pattern recognition. The extraction of the features of the characters is done in such a way that the complete portion of binary image covered and there is a distinct property associated with the each position. Feature extraction method categories into three types.
1. Structural 2. Statistical 3. Hybrid

Structural Features

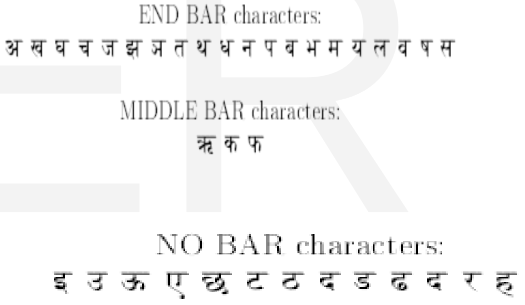
Characters can be represented by structural features with high tolerance to distortions and style variations. This type of representation may encode some knowledge about the structure of the object or may provide some knowledge as to what sort of components make up that object. Structural features are based on topological and geometrical properties of the character, such as aspect ratio, end points, loops, dots, branch points, junctions, strokes and their directions, inflection between two points, horizontal curves at bottom or top, etc.

The pre-classification is done using a two stage classification based upon the structural features. The first stage employs classification using global features like presence of vertical line in the character, its position in the character and the presence of holes. These features can be termed as global features. The detection of global features is followed by the detection of the local features such as presence of end points.

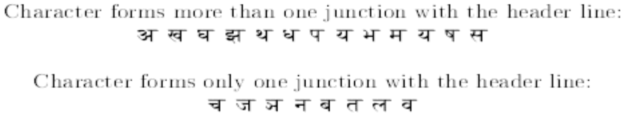
Before extracting features of Marathi Compound Character need to see some Characteristics of Compound Characters. These are some structural properties of character obtained by Visual Inspection Shown in fig 14.


FULL BOX characters:
अ इ ई उ ऊ ए ऋ क ख ग घ ङ च छ ज झ ञ ट ठ ड ढ ढ त थ द ध न प फ ब भ
म य र ल व श ष स ह
UPPER HALF BOX characters:
LOWER HALF BOX characters:

a) **Classes of core characters based on the Coverage of core strip**


END BAR characters:
अ ख घ च ज झ ञ त थ ध न प ब भ म य ल व ष स
MIDDLE BAR characters:
ऋ क फ
NO BAR characters:
इ उ ऊ ए छ ट ठ द ड ढ द र ह

b) **Classes of FULL BOX characters based on the presence and Position of vertical bar**


Character forms more than one junction with the header line:
अ ख घ झ थ ध प य भ म य ष स
Character forms only one junction with the header line:
च ज ञ न ब त ल व

c)**classes of END BAR characters based on joining pattern of the character with header line**

Fig 14. Features of Character

Edge is one of the important feature of extracting Marathi characters. Edges are important features in an image since they represent significant local intensity changes. They provide important clues to identify the handwritten character. Canny Edge detection is used in pre-processing stage for feature extraction.Canny Edge Detector is an optimal edge detection method that detects edges with noise suppressed at same time.

Canny Edge Detection Algorithm

1. Smoothing: Blurring of the image to remove noise.
2. Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.
3. Non-maximum suppression: Only local maxima should be marked as edges.
4. Double thresholding: Potential edges are determined by thresholding.
5. Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

After detecting edge, used edge map[12] for feature extraction of character. Results of feature extraction using edge map is shown in fig 15. Gets the pixels and check the neighboring pixels, and if an edge occurs on the pixels, then the probability of edge increases. This map of the probability of the edges is called as edge map.

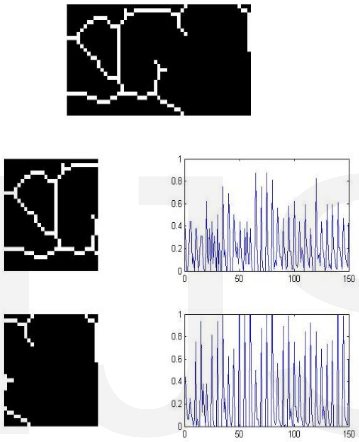


Fig 15: Result of Edge Map

3 Observations

The proposed system is implemented in MATLAB2013 , for this taken the samples of handwriting from 400 peoples. As per the earlier work no such standard database is available in literature. For making the database scan the image of character and crop. After worked on proposed work get the observations shown in table 1.

Segmentation	Recognition Accuracy
Line segmentation	100%
Word segmentation	100%
Compound Character segmentation with Minutiae Detection algorithm	95%
Feature extraction using Edge map	94%

Table 1. Recognition Rate

4 Conclusion

Worked on new novel segmentation algorithm and new structural feature extraction method for recognizing handwritten Marathi compound character. The proposed minutiae detection algorithm gave 95% accuracy for segmentation. By using structural feature edge map getting 94 % recognition accuracy for character recognition.

Acknowledgment

This paper is made possible through the help and valuable support of my guide Dr.Latesh Malik. Finally, I sincerely thanks to my parents and God.

References

[1]Sushama Shelke, Shaila Apte, “A Multistage Handwritten Marathi Compound Character Recognition Scheme using Neural Networks and Wavelet Features ” International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 4, No. 1, March 2011.

[2] Miss Vandana M. Ladwani, Dr.latesh Malik, “Novel Approach to Segmentation of Handwritten Devnagari Word”, 978-0-7695-4246-1/10 $26.00 © 2010 IEEE DOI 10.1109/ICETET.2010.143.

[3]U.K.S. Jayarathna, G.E.M.D.C. Bandara,“ A Junction Based Segmentation Algorithm for Offline Handwritten Connected Character Segmentation”, International Conference on Computational Intelligence for Modeling Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06).

[4] Ambadas B. Shinde, yogesh H. Dandawate, “ Shirorekha Extraction in Character Segmentation for Printed Devanagri Text In Document Image Processing”, 2014 Annual IEEE India Conference (INDICON).

[5]Roli Bansal, Priti Sehgal & Punam Bedi,” Effective Morphological Extraction of True Fingerprint Minutiae based on the Hit or Miss Transform”, International Journal of Biometrics and Bioinformatics (IJBB), Volume (4) : Issue (2).

[6] Dhaval Salvi, Jun Zhou, Jarrell Waggoner, and Song Wang,“ Handwritten Text Segmentation using Average Longest Path Algorithm”, 978-1-4673-50542-95/132/$31.00 ©20132 IEEE.

[7] R.G. Casey et.al. “A Survey of Methods and Strategies in Character Segmentation”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18,pp 690-706, 1996.

[8] Veena Bansal and R.M.K. Sinha. “Segmentation of touching and Fused Devnagari characters, ". Pattern recognition, vol. 35: 875-893, 2002.

[9] Ms. Aarti Desai, Dr. Latesh Malik,“ A Modified Approach to Thinning of Devanagri Characters”, 978-1-4244-8679-3/11/$26.00 ©2011 IEEE.

[10] K.B.M.R. Batuwita, G.E.M.D.C. Bandara,“ Meaningful Segmentation of Offline Individual Handwritten Numeric Characters”, 2006 IEEE International Conference on Fuzzy Systems ,Vancouver, BC, Canada July 16-21, 2006.

[11]Dr.Amitabh Wahi, Mr.Sundaramurthy.S,Poovizhi.P,”Recognition of Handwritten Tamil Characters using wavelet”, International Journal of Computer Science & Engineering Technology (IJCSET).

[12] N.N.Khalsa, Parag.P.Gudadhe, Dr. V. T. Ingole,” Advance Image Classification System”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3210 - 3214

[13]Mrs.Snehal Golait,Dr.Latesh Malik, ″Handwritten Marathi Compound Character Segmentation using Minutiae Detection Algorithm”, International Conference on Recent Trends in Computer Science and Engineering (ICRTCSE 2016), Published by Elsevier B.V.