ENHANCEMENT AND SEGMENTATION OF HISTORICAL RECORDS

Soumya A[1] and G Hemantha Kumar[2]

[1]Dept. of Computer Science & Engg,
R V College of Engineering, Bangalore, India
`soumyaa@rvce.edu.in`
[2]Dept. of Studies in Computer Science,
University of Mysore, Mysore, India
`ghk.2007@yahoo.com`

ABSTRACT

Document Analysis and Recognition (DAR) aims to extract automatically the information in the document and also addresses to human comprehension. The automatic processing of degraded historical documents are applications of document image analysis field which is confronted with many difficulties due to the storage condition and the complexity of the script. The main interest of enhancement of historical documents is to remove undesirable statistics that appear in the background and highlight the foreground, so as to enable automatic recognition of documents with high accuracy. This paper addresses pre-processing and segmentation of ancient scripts, as an initial step to automate the task of an epigraphist in reading and deciphering inscriptions. Pre-processing involves, enhancement of degraded ancient document images which is achieved through four different Spatial filtering methods for smoothing or sharpening namely Median, Gaussian blur, Mean and Bilateral filter, with different mask sizes. This is followed by binarization of the enhanced image to highlight the foreground information, using Otsu thresholding algorithm. In the second phase Segmentation is carried out using Drop Fall and WaterReservoir approaches, to obtain sampled characters, which can be used in later stages of OCR. The system showed good results when tested on the nearly 150 samples of varying degraded epigraphic images and works well giving better enhanced output for, 4x4 mask size for Median filter, 2x2 mask size for Gaussian blur, 4x4 mask size for Mean and Bilateral filter. The system can effectively sample characters from enhanced images, giving a segmentation rate of 85%-90% for Drop Fall and 85%-90% for Water Reservoir techniques respectively.

KEYWORDS

Document Analysis, Preprocessing, Filters, Segmentation, Drop Fall Technique, Water Reservoir Technique

1. INTRODUCTION

A generic Optical Character Recognition (OCR) system comprises of different stages like preprocessing, segmentation, feature extraction and classification. Preprocessing is one of the most interesting and challenging topics in DAR. Preprocessing of document involves converting scanned images or photographed images of machine printed or handwritten text which may

include numbers, letters and symbols into system processable format. Segmentation is an important assignment of any OCR system and it separates the image text documents into lines,words and characters. Hence the accuracy of OCR system primarily depends on the segmentation algorithm been used.

Segmentation of handwritten text of Indian languages is challenging when compared with Latin based languages because of its structural complication and presence of compound characters. This complexity increases further if were to recognize text of ancient Indian or non-Indian epigraphical documents. The epigraphical records engraved on stones, rocks, pillars or on some other writing material are non-linear in their shapes and non-uniform in their sizes.

Raw image of an epigraph contains unwanted symbols or marks, noise embedded and text engraved with much skew.The spacing between characters and also between the lines and the skew could complicate the process of translating the scripts. Some touching lines as well as characters complicates the process of segmentation which is input for the recognition process in the later stages. Hence the input document image of epigraphs is to be preprocessed for removal of noise, skew detection and correction, followed by segmentation of characters [1].

Inspite of several positive works on OCR across the world, development of OCR tools in Indian languages is still a challenging task. Character segmentation plays an important role in character recognition since incorrectly segmented characters are susceptible to be recognized wrongly. Hence the proposed work focuses on preprocessing and segmentation of ancient handwritten documents. This is an initial step towards developing OCR for ancient scripts, which can be used by archaeologists and historians for digitization and further exploration of ancient records.

This paper is organized as follows: Section 2 elaborates the related works in the field. The system architecture is highlighted Section 3. The theory and related mathematical background of the approaches in current system is discussed in Section 4. Methodology is given in Section 5. Experimental results and performance analysis is covered in Section 6 and Section 7 provides conclusion.

2. RELATED WORK

Researchers have worked on many approaches for preprocessing and segmentation of various languages. In this section, some of the works are discussed.

The linear Unsharp Masking (USM) technique [2] is adopted to increase the pictorial presence of an image by highlighting its regularity contents to improve the edge and detailed information in it. Nevertheless this method is easy and gives good result for many applications. It has two limitations, one it is tremendously sensitive to noise and other one is that it increases high contrast areas much more than area that do not show high image dynamics. Therefore output image suffers from unkind overshoot objects. Adaptive Unsharp Masking hires an adaptive filter that controls the contribution of sharpening in the manner that contrast enhancement occurs in high dense areas and less or no image sharpening occurs in soft areas. This algorithm is better compared with several other methods available in linear unsharp masking filter technologies.

Binarization is the initial step for processing, with the fact of degradation of the source document, whichever global or local thresholding approaches are chosen. The Otsu thresholding algorithm

[3] using the histogram shape analysis, which is most widespread global binarization algorithm. The thresholding of Otsu yields a promising performance when the histogram has twin modal distribution. The global threshold is designated automatically by a discriminant standard.

Another method utilize image contrast defined as local image minimum and maximum when compared with the image gradient process, the image contrast derived by the local maximum and minimum process has a good property and it is more tolerant to the uneven illumination, document degradation such as smudge [4]. This method is superior when handling document images with difficult background variation. Finally, the ancient document image is binarized based on the local thresholds that are derived from the detected high contrast image pixels when the same is compared with previous method based on image contrast, the method uses the image contrast to recognize the text stroke boundary and it can be used to produce high accurate binarization results.

A common technique for scrubbing the degraded documents is modified iterative global threshold algorithm [5]. A best approach in the separation of object information from foreground is to compute a global threshold of intensity value based on which two clusters can be diverted. It is an iteration approach which can handle many degraded conditions. In each iteration the intermediate tones are shifted towards background there by providing efficient difference between foreground and background. It is mostly useful for the documents having non-uniform distribution of noises.

In order to make foreground inscriptions clearly visible from background, Histogram normalization is used. The image obtained still may suffer from uneven background intensity variation which in turn reduces the clarity of the foreground. Further processing of image is done to get an image with better foreground information. As the intensity of the foreground pixels differs from the intensity of background, this key factor is used to identify the foreground characters. The main criteria is to find a threshold value which causes the image components to lie in one of two levels L0, which is below the threshold value and L1, which is above the threshold value. The pixels which are above the threshold value represent the nodes of a graph. The nodes form the basis for representing the foreground. The nodes that are neighbors in the sampling grid are joined by an edge [6].

This technique is used to reduce the number of pixels in the image by a factor of 4 or 8, which in turn will decrease the number of pixels that has to be processed. The image is represented as a graph by associating each pixel to a vertex of a graph and connecting the pixels that are neighbors in the sampling grid by an edge. The gray value of the pixel is considered as an attribute of the vertex. Since the image is of finite size so also the graph. Pixels represent the finite regions and vertices represent the faces. The dual of this graph represents borders of the faces which are inter-pixel edges and vertices. Dual graph pyramids are constructed by adopting bottom-up approach. Each level of pyramid represents an adjacency graph where vertices correspond to regions and edges represent the relation between the regions [7].

The procedure of segmentation has huge importance in the handwritten script identification. Thus an abundant study of research outcome in related segmentation field was surveyed. The algorithm based on connected components [8], segmenting the document image into non-overlapping equi-width vertical zones.

A method is to decrease the noise level, which is existing in the distorted image is proposed in [9]. The distorted image will be first binarized using the threshold, which is determined by the Otsu's method. Next for each pixel p(x, y) estimate the horizontal and vertical run length count. If horizontal and vertical run length count is less than a specified threshold then it is assumed to be noise and will be eliminated.

Gaussian kernel is a linear operation; convolution is used to find the common area between the profile and the Gaussian kernel. The degree of shift in the Gaussian kernel during the convolution process linearly varies with horizontal profile information. So this can be used to represent randomness in the profile and provides a zero crossing smooth curve, when it is convolved with the profile, represented by 'C'. The peaks which are above zero are treated as the gaps between the lines. Based on this information, the line segmentation is performed [10].

Another method for segmentation is nearest neighbor algorithm [11] which is iterative in nature scans the character from the top left portion of the image. When it reaches a first black pixel, then the first symbol is identified through the connected component. If it is found to be the first character of the script, then it will be placed as the first character of the new line used for placing the character segmentation. The centroid of the character is calculated and stored in an array separation the x and y coordinators. The document is again scanned from left top to locate the next black pixel and hence the next character in the document. The centroid of the character also computed. The distance between the centroid is computed using the distance formula. If the distance is less than or equal to threshold value then the character is assume to lie on same line. Otherwise the character is consider being the part of the next line and transferred to the next line in the result part. This process is continued until all the characters are scanned and whole image has been traversed. At the end of the iterative algorithm the separated lines are obtained from the source. The individual character can also be obtained in this process itself.

Text line segmentation is necessary to detect all text regions in the document image. The algorithm based on multiple histogram projections using morphological operators to extract features of the image. Horizontal projection is performed on the text image, and then line segments are identified by the peaks in the horizontal projection. Threshold applied to divide the text image into segments. False lines are eliminated using other threshold. Vertical histogram projections are used for the line segments and decomposed into words using threshold and further decomposed to characters. This kind of approach provides best performance based on the experimental results such as Detection rate DR (98%) and Recognition Accuracy RA (98%) [12,13].

Contour tracing is a technique applied to digital images for extracting the boundary of any object. This kind of system applies one of the recent contour tracing algorithms to separate character by using the Theo Pavlidis's algorithm [15]. It works with 4-connected patterns. The width of the segmented components from this process is checked. If it is more than the criteria value of the average width of the components, it will be processed in the next stage. This means there are some touching characters that are not separated [14].

Tracing of Background Skeleton approach, is applied to separate some touching components to segment touching characters, background skeleton is processed by using the Zhang-Suen thinning algorithm. Then, contour tracing algorithm is applied to abstract the skeleton of the background. Subsequently, the characters in each line will be sorted by checking the column position in order

to determine the sequence of the characters. This is applied to practical data from ancient documents [15, 16].

Enhanced stroke filter have shifted the attention to the skeletons of potential strokes. In this manner, the task of distinguishing strokes from a complicated background is converted to the task of comparing the difference between skeletons of potential strokes and those of disturbing patches, both of which can be extracted from the resulting images of previous Stroke filters. Skeleton constraints such as length and width constraints can be introduced into the method to enhance stroke information [17].

For the segmentation of unconstrained handwritten connected numerals, Water Reservoir technique is used. A reservoir location and size, touching position (top, middle or bottom) is decided. Analyzing the reservoir boundary, touching position and topological structures of the touching pattern, the best cutting point and then the cutting path for segmentation is generated [18].

Segmented linked characters are critical preprocessing steps in character recognition applications. Old drop fall algorithm has proved to be an efficient segmenting method due to its simplicity and effectiveness. However it is subject to small convexes on the contour of characters. Xiujuan Wang, Kangfeng Zheng, and Jun Guo presented Innertial Drop fall algorithm and big drop fall algorithm to avoid this defect [19, 20, 21].

3. SYSTEM ARCHITECTURE

The system "Enhancement and Segmentation of Historical Records" designed, mainly consists of the subcomponents - Preprocessing and Segmentation as shown in Figure 1. The input to the system is ancient epigraphic documents of varying amount of degradation.

- **Image Enhancement :** This sub system enhances the quality of the ancient document images by reducing noise. This is carried out by providing four different filtering options of various filter sizes.

- **Binarization:** This sub-system converts RGB images to binary images using thresholding technique known as Otsu algorithm.

- **Segmentation:** This sub-system samples out characters from the ancient documents and is achieved through Drop Fall and Water reservoir techniques.
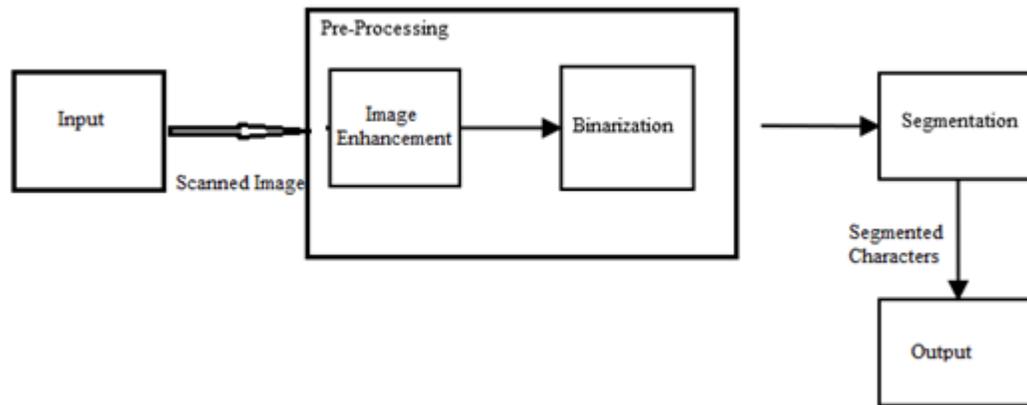
Figure 1. Enhancement and Segmentation System

4. REVIEW ON THE APPROACHES USED IN PROPOSED SYSTEM

The methods used in current system are described in this section:

Preprocessing stage of degraded ancient document images includes: enhancement and reduction in noise, which is achieved through different filtering methods for smoothing or sharpening namely Bilateral, Mean, Median, and Gaussian Blur Filters. These filters are provided with different mask sizes and parameter values. This is followed by binarization of the enhanced image to highlight the foreground information, using Otsu thresholding algorithm. Character segmentation is performed based on Drop Fall and Water Reservoir concept.

4.1 Mean Filter

The mean filter is a sliding-window longitudinal filter that exchanges the center value in the window with the average (mean) of all the pixel values in that window [22]. Let Sxy represent the set of coordinates in a rectangular sub image window of size m x n, centered at point (x, y). The arithmetic mean filtering process computes the average value of the corrupted image g(x, y) in that area defined by Sxy. The restored image value at any point (x, y) is merely the arithmetic mean calculated using the pixel in that region, indicated in Equation 1.

$$\widetilde{f}(x,y) = \frac{1}{m^n} \ \sum_{(s,t)\in S_{xy}} g(s,t) \tag{1}$$

This operation can be applied using a convolution mask in which all coefficients have value 1/mn. Mean Filters smoothes local variations in an image and as a result of blurring, noise is reduced.

4.2 Median Filter

In image processing, neighborhood averaging is the best method to perform the noise reduction, whereas the method can overturn isolated out of range noise, however the adverse effect is that it also distorts sudden changes such as sharp edges. The median filter can suppress the noise without damaging the sharp edges. In median filtering, all the pixel values are first sorted into numerical order and then replaced with the middle pixel value [22].

Let y be a pixel location and w a neighborhood centered on location (m, n) in the image, therefore median filter is given by Equation 2,

$$y [m, n] = median\{x[i ,j], (i , j) \text{ belongs to w}\} \tag{2}$$

Subsequently the pixel y [m, n] represents the location of the pixel y, m and n represents the x and y co-ordinates of pixel y. w represents the neighborhood pixels surrounding the pixel position at (m, n), (i, j) belongs to the same neighborhood centered on (m, n). Hence the median method will take the median of all the pixels within the range of (i, j) represented by x [i, j].

4.3 Gaussian blur Filter

A Gaussian blur or Gaussian smoothing involves blurring an image by Gaussian function and used to decrease image noise and image details. Gaussian smoothing is used as a preprocessing stage in computer vision algorithms in order to enhance image structures at different scales. The 2- dimension Gaussian function is given by Equation 3.

$$G(x, y) < -\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

where x is the distance from the origin (Horizontal axis), y is the distance from the origin (vertical axis), and σ is the standard deviation of the Gaussian distribution. The standard deviation σ of the Gaussian determines the amount of smoothing [22].

4.4 Bilateral Filter

Bilateral filter [23] is a non linear filter in spatial domain, which does averaging without smoothing the edges. The bilateral filter inputs a weighted sum of the pixels in a local neighborhood; the weights depend on both the spatial distance and the intensity distance. Essentially the bilateral filter has weights as a product of two Gaussian filter weights, one of which corresponds to average intensity in a spatial domain, and second weight corresponds to the intensity difference. Hence no smoothing occurs, when one of the weights is close to 0, which means the product becomes insignificant around the region where intensity changes swiftly, which represents usually the sharp edges. As a result, the bilateral filter preserves sharp edges [28]. Pixel location x, bilateral filter output is given in Equation 4

$$\hat{I} = \frac{1}{C}\sum_{y\in N(x)} e^{\frac{-||y-x||^2}{2\sigma_d^2}} - e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}} I(y) \tag{4}$$

There parameters controlling the fall-off of weights in spatial and intensity domains, respectively. And are inputs and output images respectively are spatial neighborhood of pixel I(x), and C can be given as

$$C = \sum_{y\in N(x)} e^{\frac{-||y-x||^2}{2\sigma_d^2}} - e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}} \tag{5}$$

4.5 Otsu's Method of Binarization

Binarization is the method of converting a grey scale image to a binary image by using threshold selection procedures to categorize the pixels of an image into either one of the two classes. Binarization of the image using Otsu method [23] is used to automatically accomplish histogram shape-based image thresholding or the decrease of a gray level image to a binary image. This algorithm adopts that the image as thresholded contains two classes of pixels, then calculates the

optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal.

In an Otsu's method weighted sum of variances of the two classes is given by:

$$\sigma_\omega^2(t) = \omega_1(t)\,\sigma_1^2(t) + \omega_2(t)\,\sigma_2^2(t) \tag{6}$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variances of these classes. The class probability is

$$\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t) = \omega_1(t)\omega_2(t)\,[\mu_1(t) - \mu_2(t)]^2 \tag{7}$$

The class probability $\omega_1(t)$ is calculated from the histogram as t:

$$\omega_1(t) = \sum_0^t p(i) \tag{8}$$

While the class mean $\mu_1(t)$ is:

$$\mu_1(t) = [\sum_0^t p(i)x(i)] \ / \ \omega_1 \tag{9}$$

where x(i) is $\omega_2(t)$ the value at the center of the ith histogram bin. Also can calculate μ_2 on the right-hand side of the histogram for bins greater than t.

4.6 Drop Fall Algorithm for Segmentation

Drop fall algorithm [24] with respect to the principle that an equally ideal cut between two touched characters can be created, if one has to role a hypothetical marble off the top of the first character and create the cut where the marble falls. The important things to be addressed for this implementation are where to drop the marble from because it is important if the algorithm starts at the wrong place. The marble can simply roll down the left side of the first digit or the right side of the second digit and, hence, it would be completely unsuccessful. The best approach to start drop falling process is possible to the point at which two characters are touched. In this process the pixels are scanned row by row until a black boundry pixel with adjacent black boundry pixel to the right of it is identified, where as the two pixels are separated by white space. This pixel is used as a point to start the drop fall as shown in Figure 2.
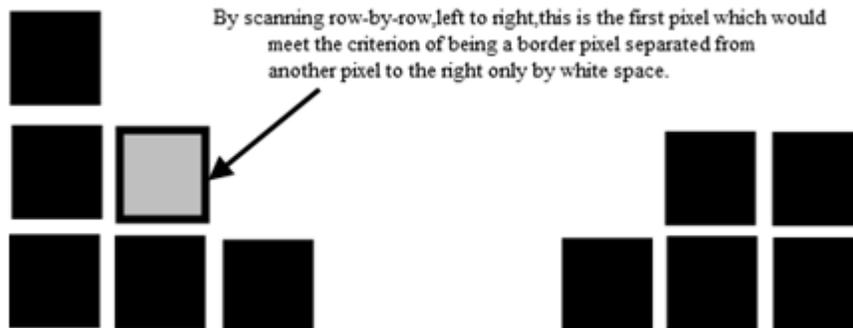

Figure 2. Identification of Initial Pixel Positions

The direction that the algorithm will move is according to the current pixel position and its surroundings as shown in Figure 3.
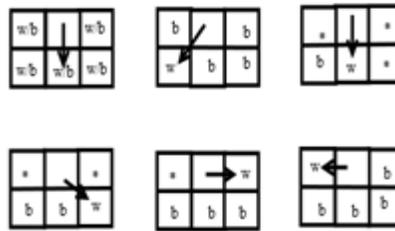

Figure. The neighborhood pixels of n$_0$

Figure. The principle of drop-fall algorithm

Figure 3. The Principle of Drop Fall algorithm

4.7 Water Reservoir Algorithm

The larger space generated by touching characters is analyzed with the help of water reservoir concept. The working principle of water reservoir method is illustrated in Figure 4. When water is poured from top (bottom) of a component, the regions of the component where water will be stored are considered as top (bottom) reservoir. Top (bottom) reservoir is the reservoir obtained when water is poured from top (bottom). The white spaces are found in the regions in the bounding box of the components where water can be stored. These regions are called water reservoirs. The reservoirs obtained in this procedure are not considered for further processing. Those reservoirs whose heights are greater than a threshold value T1 are considered for further processing.
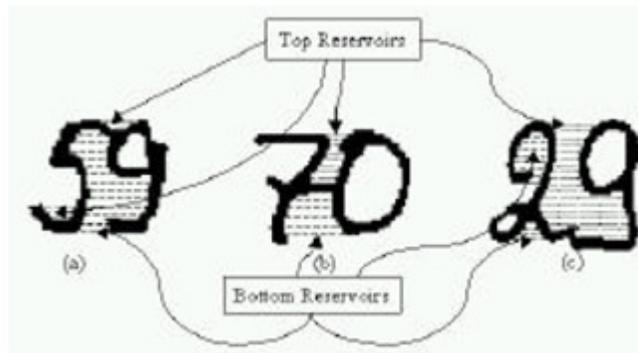

Figure 4. Reservoirs formed from water flow from top and bottom is shown for (a) top (b) middle and (c) bottom touching numerals. The top Reservoirs are marked by Dots and bottom reservoirs are marked by small line segments

When two characters touch each other, they create a space (reservoir) between the characters. This space is very important for segmentation because,

a) As cutting points are concentrated around the base of the reservoir, and hence, decreases the search area.

b) The cutting points lie on base of the reservoir.

c) The space attributes (center of gravity and height) aid to go near the best touching position.

If water is poured from top (bottom) of large space created by touching (Water reservoir) Base of the reservoir connected numeral then water will be stored in this large space. This water stored area is named "Water Reservoir" [25]. Figure 5 illustrates the same.

Figure 5. Examples of touching numeral and Space created by the touching

Figure 6 depicts the detection of touching position recognition. The largest reservoir of the component whose center of gravity lies in vm region is found. This reservoir is known as the best reservoir for touching. The base-line (lowermost row of the reservoir) of the best reservoir is then identified. The best reservoir and its base-line are shown and to find this touching position in the components, morphological thinning operation is applied to touching components for further processing. For feature points extraction the touching position is renowned. The leftmost and rightmost points of the base-line of considered reservoirs are the feature points. These points are initial feature points. With this initial feature points the best feature point (which gets maximum confidence value) is chosen for segmentation. To calculate confidence value (CV) following features are considered. Euclidean distance of feature points from the center of gravity of the touching component.
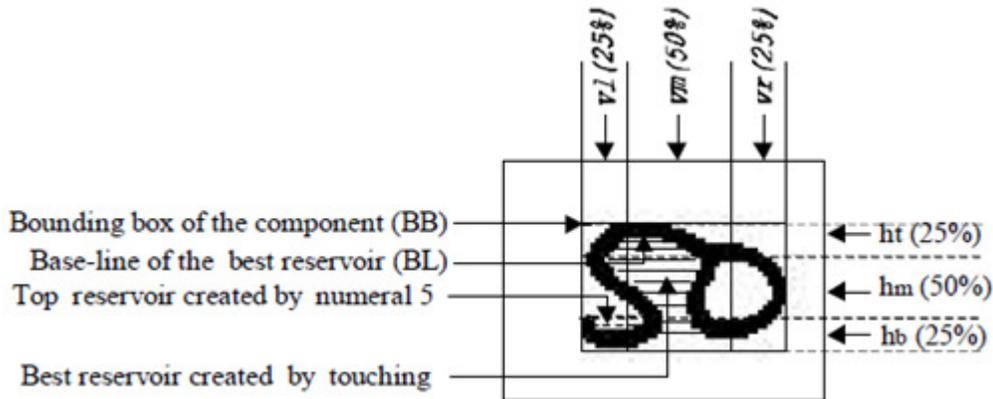

Figure 6. Feature Detection Approach.

5. PROPOSED SYSTEM AND METHODOLOGY

5.1 DFD of the Preprocessing and Segmentation System

The Data Flow Diagram(DFD) of the current system is shown in Figure 7. This work is carried out in two phases. In the first phase - Preprocessing, the degraded ancient document image is taken as input and it is converted to grayscale. Then the smoothing or sharpening filters, namely Median filter, Gaussian filter, Mean filter, Bilateral filter of different mask sizes are applied to reduce the amount of noise and thus enhances the image. Next, the enhanced image is binarized using Ostu algorithm to differentiate background and foreground of ancient document.

In the second phase Segmentation is carried out using Drop Fall algorithm and Water Reservoir algorithm, to obtain sampled characters, which can be used later stages of OCR.
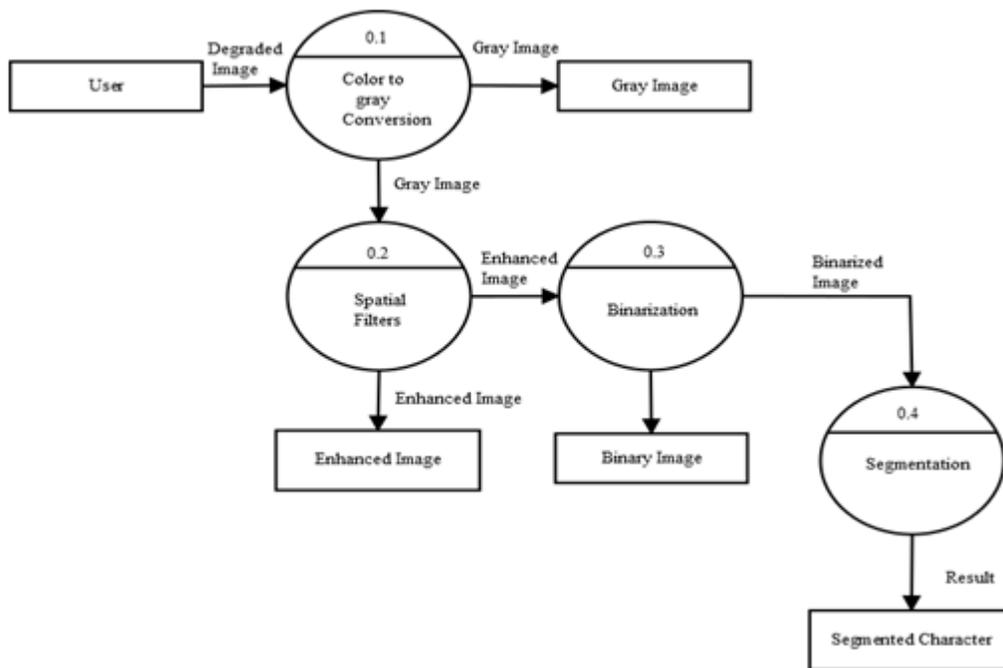

Figure 7. DFD of the Preprocessing and Segmentation Phases

5.2 Methodology

The functionality of the phases – Preprocessing and Segmentation in detail is covered in this section. The ancient input image is first converted to grayscale image. The image is enhanced and noise is reduced by applying four different filters with mask size of 2x2 and 4x4. Next, The enhanced image is converted to binary image. Lastly the characters in the document are sampled out using Drop Fall and Water Reservoir approaches.

5.2.1 Preprocessing

➢ **Image Enhancement**

- ▪ **Input:** Degraded Gray Scale image
- ▪ **Functionality:** Enhances epigraphic image of medium-level degradation using spatial filters namely Median, Gaussian blur, Mean, Bilateral filter on the input image.
- ▪ **Output:** The enhanced image with reduced noise.

- ▪ **Algorithm for Enhancement using Median filter**
  [Step 1]: Read the gray image.
  [Step 2]: Compute y [m, n] = median{x [i, j], (i, j) belongs to w}
  y be a pixel position, w represent a neighborhood centered around location (m, n) in the image.
  [**Step 3]:** Apply the Median filter designed over the entire input image to obtain enhanced image.

- ▪ **Algorithm for Enhancement using Gaussian blur filter**
  [Step 1]: Read the gray image.
  [Step 2]: Compute$e$ $G(x, y) < -\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$
  x is the distance from the origin (Horizontal axis), y is the distance from the origin (vertical axis), σ is the standard deviation of the Gaussian distribution.
  [Step 3]: Apply the Gaussian filter designed over the entire input image to obtain enhanced image.

- ▪ **Algorithm for Enhancement using Mean filter**
  [Step 1]: Read the gray image.
  [Step 2]: Compute $\hat{f}(x, y) < -\frac{1}{m^n}\sum_{(s,t)\epsilon S_{xy}} g(s,t)$
  Sxy represent the set of coordinates in a rectangular subimage window of size m X n, centered at point (x, y).
  [Step 3]: Apply the Mean Filter designed over the entire input image to obtain enhanced image.

- ▪ **Algorithm for Enhancement using Bilateral filter**
  [Step 1]: Read gray image.
  [Step 2]: Compute $C = \sum_{y\in N(x)} e^{\frac{-||y-x||^2}{2\sigma_d^2}} - e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}}$
  [Step 3]: Compute $\hat{I} = \frac{1}{C}\sum_{y\in N(x)} e^{\frac{-||y-x||^2}{2\sigma_d^2}} - e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}} \quad I(y)$
  [Step 4]: Apply the Bilateral Filter designed over the entire input image to obtain enhanced image.

➢ **Binarization**
  - ▪ **Input:** Enhanced Image
  - ▪ **Functionality:** The enhanced images are converted to binary image consisting of ones and zeroes.

- **Output:** Binarized image

- **Algorithm for Binarization**
  [Step 1]: Compute histogram and probabilities of each intensity level
  [Step 2]: Initialize initial $\omega_i(0)$ and $\mu_i(0)$
  [Step 3]: Compute through all possible thresholds t=1. Maximum intensity
      Revise ω_i and μ_i; Compute $\sigma_b^2(t)$
  [Step 4]: Desired threshold corresponds to the maximum $\sigma_b^2(t)$
  [Step 5]: Compute two maxima (and two corresponding thresholds). $\sigma_{b1}^2(t)$
  [Step 6]: Compute greater max and $\sigma_{b2}^2(t)$ is the greater or equal maximum
  [Step 7]: Compute required threshold ← threshold1+threshold2/2

5.2.2 Segmentation

The segmentation of the document image is carried out at the character level using Drop Fall and Water Reservoir Approaches.
- **Input:** Binary epigraph image
- **Functionality:** The binarized image is segmented to characters
- **Output:** Segmented characters of the input epigraph.

➢ **Drop fall algorithm**
Drop falling algorithm forms segmentation path by rolling in between two touching characters and displays the segmented characters.
  [Step 1]: Input the binary image
  [Step 2]: Find the Height and Width of the touched characters
  [Step 3]: Apply Breadth First Search (BFS) algorithm to find the touched characters
  [Step 4]: If found start the Drop fall, the drop falling algorithm it will always move downwards, crossways down-wards, to the right, or two the left.
  [Step 5]: Make the slice where marble parks. Thus Segmentation path for   connected components is found
➢ **Water Reservoir Algorithm**
  [Step 1]: Find the size of the characters to find touched characters.
  [Step 2]: The positions and sizes of the reservoirs are analyzed and a reservoir is detected where touching is made, the initial feature points for segmentation are noted.
  [Step 3]: The best feature points are noted from the initial feature points.
  [Step 4]: Based on touching position, close loop positions and morphological structure of touching region the cutting path is produced.

6. EXPERIMENTAL RESULTS, ANALYSIS AND DISCUSSION

6.1 Experimental Results

The system developed is tested on nearly 150 ancient epigraphic images and the results are found to be satisfactory. The sample experimental results are depicted in following figures. Figure 8 shows the input ancient historical record, which is analysed for different spatial filtering techniques.
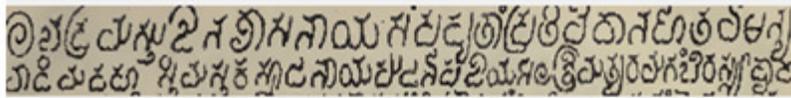
Figure 8. Input Image selected for Pre-processing

Figure 9 shows the results of color to gray scale conversion, when carried out on the image shown in Fig 8.
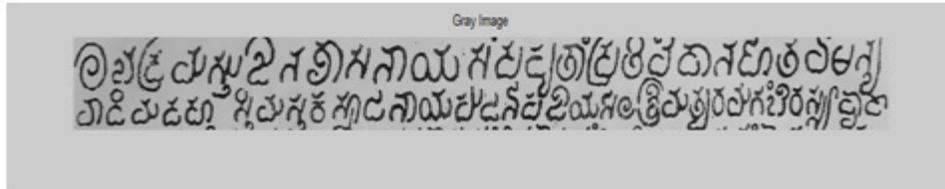

Figure 9. The result of Gray Scale Conversion

Figure 10(a) and 10(b) shows the results after Median filtering, for the mask size of 2x2 and 4x4 respectively. The median filter is an effective method that can suppress isolated noise without blurring sharp edges.
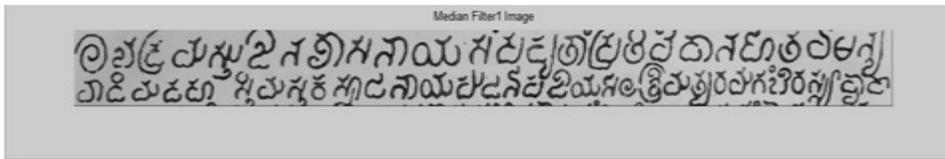

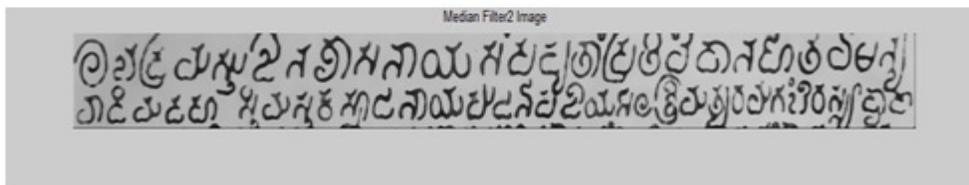Figure 10(a). The result of Median filtering for Mask size 2x2

Figure 10(b). The result of Median Filter for Mask size 4x4

Figure 11(a) and 11(b) shows the results of Gaussian blur filtering for the mask size of 2x2 and 4x4. The Gaussian blur method is used to blur the sharpen image so that a less edge highlighted image is produced.
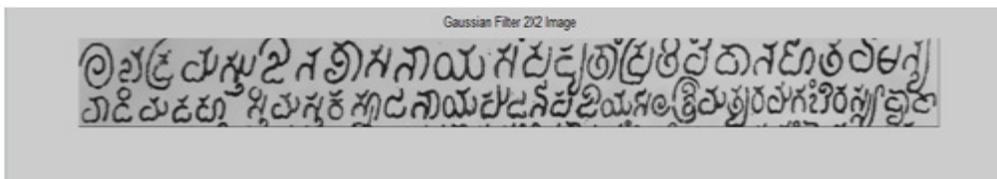

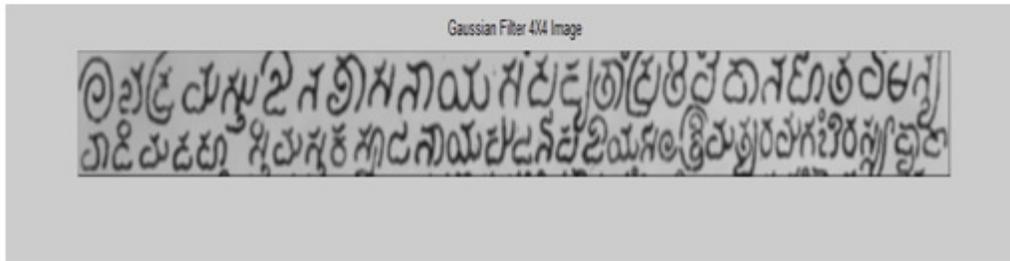Figure 11(a). The result of Gaussian Blur Filtering for Mask size 2x2

Figure 11(b). The result of Gaussian blur Filtering for Mask size 4x4

Figure 12 shows the results of Mean filtering for the mask size of 4x4. The Mean filter is a simple filter that replaces the center value in the window with the mean of all the pixel values in that window.
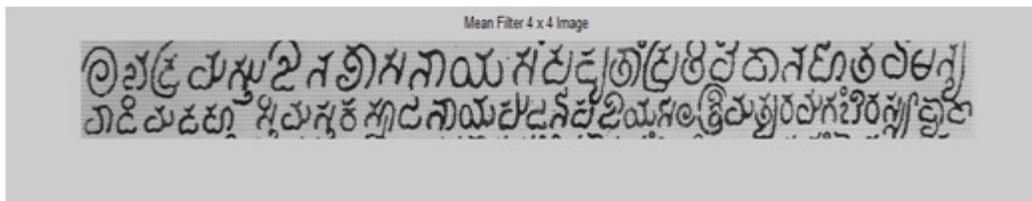

Figure 12. The result of Mean Filtering for Mask size 4x4

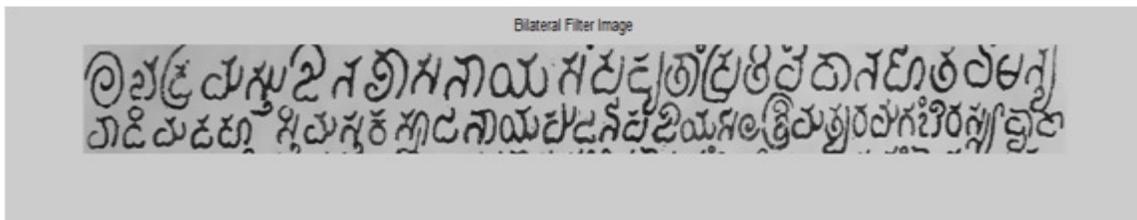Figure 13 shows the results of bilateral filtering.

Figure 13. The result of Bilateral Filtering

Figure 14 shows the result of binarization, in which the enhanced image is converted to binary image.
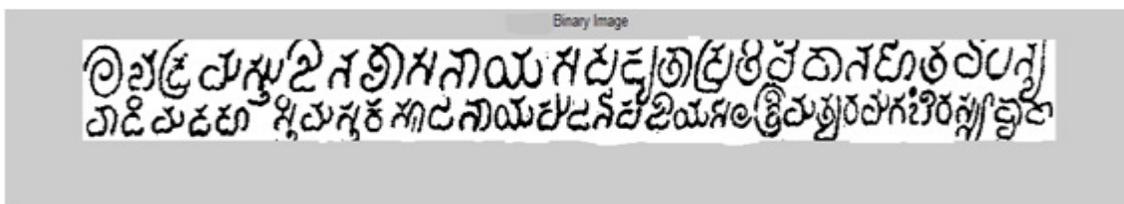

Figure 14. The results of Binarization

Figure 15 and Figure 16 represents the result of Segmentation of Characters using Drop Fall algorithm and Water Reservoir algorithm respectively.

Figure 15. The result of Segmented Characters using Drop Fall algorithm

Figure 16. The result of Segmented Characters using Water Reservoir algorithm

6.2 Performance Analysis

The dataset includes 150 samples of medium degraded images for preprocessing and segmentation. The performance of the 2 phases is discussed below:

6.2.1 Filtering Techniques

This system is tested on 150 samples of medium degraded images, using four spatial filtering techniques of varying mask sizes. The enhancement was found to be appreciable for the mask size 4x4 for Median filter when the mask size is high then the output image will appear clear with sharp edges. The mask size of 2x2 for Gaussian blurs results in blurred image. The mask size of 4x4 for Mean filter typically smoothens local variations in an image and noise is reduced as a result of blurring. Bilateral filter sharpens the edges. The smoothing Gaussian filter will result in good accuracy if the edge of the input image is very thick, where as in case of sharpening Bilateral filter gives better output for the medium degraded images.

6.2.2 Segmentation

The system showed good results when tested on 150 varying degraded images, giving segmentation rate of 85%-90% for Drop Fall algorithm, 85%-90% for Water Reservoir algorithm.
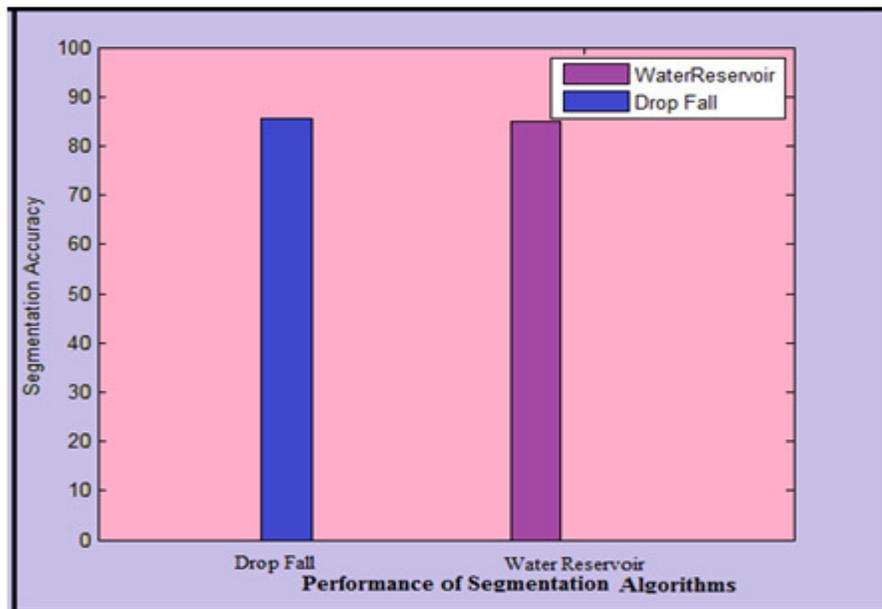

Figure 17 . Segmentation Rate of Drop fall and Water Reservoir techniques

7. CONCLUSION

The system showed good results when tested on the 150 samples of varying degraded epigraphs. It provides better enhanced output on applying filters of appropriate mask size - 4x4 mask size for Median filter, 2x2 mask size for Gaussian blur, 4x4 mask size for Mean and Bilateral filter. Segmentation is carried out using Drop Fall and Water Reservoir algorithms and system can efficiently segment characters from ancient document images. System segments the compound characters correctly when connectivity present. Few cases where in connectivity is absent, compound character is segmented separately. Segmentation rate of 85%-90% for Drop fall algorithm and 85%-90% for Water Reservoir algorithm is achieved.

REFERENCES

[1]   Tanzila Saba, Ghazali Sulong & Amjad Rehman, "A survey on Methods and Strategies on Touched Characters Segmentation", International Journal of Research and Reviews in Computer Science (IJRRCS), Vol.1, No.2, June 2010.

[2]   M. Trentacoste et al. "Unsharp Masking, Counter shading and Halos: Enhancements or Artifacts", 2012, The Euro graphics Association and Blackwell publishing Ltd. 2012.

[3]   Maya R. Gupta, National P. Jacobson, Eric K. Garcia, "Binarization and Image Preprocessing for searching Historical Documents", University of Washington, Seattle, Washington 98195, April 2006.

[4]   Bolan Su, Shijian Lu, Chew Lim Tan, "Binarization of Historical Document images Using the local Maximum and Minimum", pp.9-11, June 2010.

[5]   N. Venkata Rao, A.V.Srinivasa Rao, S Balaji and L. Pratap Reddy, "Cleaning of Ancient Document Images Using Modified Iterative Global Threshold" ,IJCSI international Journal of Computer Science Issues, Vol.8.Issue 6, No 2, pp. ISSN (online):1694-0814, November 2011.

[6]   Dr. B.P.Mallikarjunaswamy, Karunakara K, "Graph Based Approach for Background and Segmentation of the Image", Vol 02, pp.ISSN: 2230-8563, June 2011.

[7]   Mamatha H.R and Srikanta Murthy K, "Morphological Operations and Projection Profiles based Segmentation of handwritten Kannada Document", International Journal of Applied Information System (IJAIS)-ISSN: 2249-0868 Foundation of computer Science FCS, New York, USA Vol 4-N0.5, pp. 13-19, October 2012.

[8]   Das, Reddy, Govardhan, Saikrishna, "Segmentation of Overlapping text lines, Characters in printed Telugu text document images", International Journal of Engineering science and technology, Vol.2(11), pp.6606-6610, 2010.

[9]   Karthik S, Mamatha H.R, Srikanta Murth y K, "An Approach based on Run Length Count for denoising the Kannada Characters", Internatinal journal of Computer applications, Vol 50-No.18, pp.0975-8887, July 2012.

[10]   N. Anupama, Ch. Rupa & Prof. E. Sreenivasa Reddy, "Character Segmentation for Telugu Image Document using Multiple Histogram Projections", Global Journal of Computer Science and Technology Graphics & Vision, Vol 13, 2013.

[11]   Srinivasa Rao A.V, segmentation of ancient Telugu Text Documents, published online in MECS ,I.J. Image, "Graphics and Signal Processing", pp.8-14 Published Online July 2012.

[12]   Partha Pratim Roy, Umapada Pal, Josep Llados, Mathieu Delalandre, "Multi-oriented and Mutisized Touching Character Segmentation using Dynamic programming", 10th international conference on Document analysis and Recognition, DOI 10.1109/ICDAR.124 IEEE 2009.

[13]   Indu Sreedevi, Rishi Pandey, N. Jayanthi, Geetanjali Bhola, Santanu Chaudhury2, "NGFICA Based Digitization of Historic Inscription Images, ISRN Signal Processing", Vol 3, 2013.

[14]   Pavlidis, T. "Algorithms for Graphics and Image Processing", Computer Science Press, Rockville, Maryland, 2009.

[15]   Zhang, T. Y. and Suen, C. Y. "A Fast parallel algorithm for thinning digital patterns. Communication. ACM", Vol 3, No 27, pp 236-239, 2009.

[16]   B Gangamma,Srikanta Murthy K,Arun Vikas Singh, "Restoration of Degraded Historical Document Image", Journal of Emerging Trends in computing and Information Science, Vol 3, No 5, May 2012.

[17]   Xiaoqing Lu, Zhi Tang, Yan Liu, Liangcai. "Stroke –based character segmentation of low –quality images on ancient Chinese tablet", 12th International conference on document Analysis and recognition, 2013.

[18]   U. Pal, A. Belaid and Ch. Choisy "Touching numeral segmentation using water reservoir concept", Pattern Recognition Letters, Vol.24, pp. 262-272, 2003.

[19]   S.Khan."Character Segmentation Heurisrtics for Check Amount Verification "Master Thesis, Massachusetts Institute of Technology, 1998.

[20]   X.Wang, K. Zheng and J.Guo. "Inertial and Big Drop Fall Algorithm", International Journal of information Technology, vol.12, No.4, 2006.

[21] Parminder Singh and Harjinder Singh, " A Comparison of High Pass Spatial Filters using Measurements and Automation", Int. J. IJERT, Vol 1, May 2012.

[22] Mamatha H.R, Sonali Madireddi, Srikanta Murthy K, "Performance Analysis of various filters for De-noising of Handwritten Kannada Documents", International Journal of Computer Applications (0975 – 888) Volume 48– No.12, June 2012.

[23] Messaoud, I.B, et.al. "Region Based Local Binarization approach for Handwritten Ancient Documents, "in Int. Conf. Front. Handwrit recognit. Bari, pp.633-638, 2012.

[24] Srinivasa Rao A V, D R Sandeep, V B Sandeep,S Dhanam Jaya "Segmentation of Touching Hand written Telugu Characters by using Drop Fall Algorithm, International Journal of Computers & Technology, Volume 3 No. 2, OCT, 2012.

[25] Praveen kumar. C, Kiran. Y. C, " Kannada Handwritten Character Segmentation using Water Reservoir Method" , International Journal of Systems, Algorithms & Application 2012.