Improved Document Image Binarization using Hybrid Thresholding Method

Neha¹ Deepak²
¹M.Tech. Student ²Assistant Professor
¹,²Department of Electronics & Communication Engineering
¹,²DIET, KUK

Abstract—The paper presents a hybrid thresholding approach for binarization and improvement of despoiled documents. Historical documents contain information of great cultural and scientific value. But such documents are often degraded over time. Digitized degraded documents require focused processing to remove different kinds of noise and to improve readability. During the processing of binarization some problem occurs such as part of text remains unexplored. In the proposed system we will try to implement existing system using morphological operators and will improve the values of parameters like PSNR, F-Measure and NRM.

Key words: Documents, Binarization, thresholding, binary image

I. INTRODUCTION

The preprocessing step for document image analysis is Image Binarization. Image segmentation of the document image is performed for separating the foreground text from background using image binarization. Image binarization is the separation of the image into dual pixels, black as foreground and white as background. The main objective of image binarization is to separate the document foreground text and background. The outcome of image binarization is a Binary Image. It produces a binary image from a gray scale image. Binary image is a digital image that has only two set of pixels to represent it. The two pixel values that are used to represent the binary image are usually 0 and 1, 0 for black and I for white. The resultant of the image binarization is the enhanced image. It improves the quality of the document image.

Image binarization is the important preprocessing step in many document analysis techniques. It improves the further steps of document analysis and processing. It plays an important role for document processing techniques like OCR and document layout analysis. Binarization has emerged as an intense area of research in the past decade. The approach to image binarization uses thresholding. Thresholding is the simplest method of image segmentation. It converts the gray scale image into binary image. It chooses a threshold and separates the pixels by comparing them with the threshold value. It is used to separate the foreground text from the background in image binarization.

But to choose a threshold is very challenging task for degraded documents. To choose a threshold value for degraded document is an unsolved problem due to high variation between foreground and background. The degraded documents suffer from various types of degradations. The document gets degraded with the passage of time, seeping of the ink from the other side of the paper, background noise, shadow, uneven illumination, variation in illumination and contrast. In such cases to calculate an optimal threshold is a very challenging and demanding task.

The wrong estimation of the threshold value can reduce the efficiency of the document analysis system. In case of wrong estimation of threshold value there will be misinterpretation of the pixels as foreground or background. This will result in inefficient binarization and also affect the document processing techniques.

Binarization techniques are generally classified as Global and Local. Global thresholding technique uses a single optimal threshold to binarize the whole image. It compares the pixels with the single threshold and classifies them as foreground or background. This is a fast process. But, this technique fails for the degraded documents suffering from poor illumination and non-uniform color. Local thresholding technique uses different threshold for different pixels of the image. It chooses the threshold depending upon the neighboring pixels. But, this technique is inefficient for documents suffering from background noise.

Binarization techniques can be categorized depending upon the criteria used to select the threshold value.

1) Clustering Based Methods
The pixels are divided into two clusters one for foreground and other for background. Otsu method is a clustering based method that is based on image variance. The purpose of clustering based method is to divide the pixels in such a way that it maximizes the inter-class variance and minimizes the intra-class variance.

2) Entropy based methods
These methods use the frequency of occurrence of each grey level. They consider foreground and background of the image as different signal sources. The optimal threshold is the value when the sum of the foreground and background entropies maximum.

3) Local Adaptive Thresholding Methods
The algorithms calculate the threshold for each pixel depending upon the local statistics like illumination, variance, contrast and range. Niblack method chooses the threshold based on the local mean and standard deviation of the rectangular window. Bernsen used local contrast and proposed a local adaptive method. The threshold is chosen as the mean of the minimum and maximum value of the grey values.

4) Histogram based methods
Histogram of the image is used to decide the threshold value.

A. Various Binarization Techniques
Some of the Global and Local binarization techniques are discussed below

1) Otsu Method
Otsu Method is the clustering-based method for thresholding. It is the best Global thresholding method. It assumes that an image pixels are classified as foreground
and background. The histogram based thresholding is performed to convert the gray scale image into binary image. The threshold is chosen in such a way that the inter-class variance within the class is maximum and the intra-class variance between the classes is minimum. The weighted sum of variance of two classes is given as:

\[ \sigma^2(t) = p_1(t)\sigma^1_2(t) + p_2(t)\sigma^2_2(t) \] (1.1)

Otsu method is the best method for Global thresholding and thus performs efficiently for binarization. But, it gets affected by image contrast. And thus do not perform well for degraded images with uneven illumination and shadows. So, Otsu method has the limitation that it can’t be used for the degraded documents that do not have clear bi-modal histogram.

2) **Niblack Method**

Niblack Method is the Local thresholding method. It uses a rectangular window to estimate the pixel-wise threshold. The rectangular window is made to slide over the gray scale image. This method uses mean and standard deviation for the calculation of threshold.

The threshold using Niblack method is defined as:

\[ T = m + K x s \]

Where \( m \) is the mean and \( s \) is the standard deviation within the window. \( K \) is the constant. The quality of binarization is defined by \( K \) and the size of the sliding window.

As Niblack is a local thresholding method it gets affected by the local features of the image. Gaps between pixel values usually in blank areas also influence the Niblack method. The another limitation of Niblack method is that it is not suitable for document images affected by background noise.

3) **Sauvola Method**

It solves the problem of noise of Niblack method. It is local variance based method. This method also uses standard deviation like Niblack method. This method can be used for document images with uneven illumination, stains and documents with low contrast and high illumination.

The threshold using Sauvola Method is defined as:

\[ T = m \times [1 + (K(s) / R)] \] (1.2)

Where \( m \) is the mean and \( s \) is the standard deviation within the window. \( K \) and \( R \) are constants. The value of \( K \) is 0.5 and \( R \) is 128 are used.

The limitation of Sauvola method is that the text becomes thinner after binarization.

4) **Bernsen Method**

Bernsen is contrast based local adaptive method. Threshold in Bernsen method is defined as the mean of maximum and minimum intensity of the gray scale image. Depending upon this threshold the pixel is identified as foreground or background. In Bernsen method image contrast is calculated as

\[ C_{ij} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \] (1.3)

The pixel is classified as foreground if the value of \( C_{ij} \) is greater than the threshold and background otherwise. Bernsen method is not suitable for degraded document images having complex background.

II. LITERATURE REVIEW

S. vishnupriya et al. [1] has proposed that the extraction of text from badly degraded image document is a very challenging task due to the variation between the foreground and background text of various document images. The proposed system tackles this problem by the combination of several state-of-the-art binarization methodologies as well as on the efficient incorporation of the edge information of the gray scale source image. Given a degraded document image an adaptive contrast map is first constructed. The contrast map is then binarized and combined with Canny’s edge map to identify the text stroke edge pixels. The document text is further extracted using globally matched wavelet filters. Finally Fisher classifier is used for improving the result of image segmentation. The approach is simple, robust and applicable to images of typewritten text as well as hand written text or a mixture of both. Experimental result shows that the proposed method gives superior performance compared with other technique.

Bolan su et al. [3] Segmentation of text from badly degraded document images is a very challenging task due to the high inter/intra-variation between the document background and the foreground text of different document images. In this paper, we propose a novel document image binarization technique that addresses these issues by using adaptive image contrast. The adaptive image contrast is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation caused by different types of document degradations. In the proposed technique, an adaptive contrast map is first constructed for an input degraded document image. The contrast map is then binarized and combined with Canny's edge map to identify the text stroke edge pixels. The document text is further segmented by a local threshold that is estimated based on the intensities of detected text stroke edge pixels within a local window. The proposed method is simple, robust, and involves minimum parameter tuning. It has been tested on three public datasets that are used in the recent document image binarization contest (DIBCO) 2009 & 2011 and handwritten-DIBCO 2010 and achieves accuracies of 93.5%, 87.8%, and 92.03%, respectively, that are significantly higher than or close to that of the best-performing methods reported in the three contests. Experiments on the Bickley diary dataset that consists of several challenging bad quality document images also show the superior performance of our proposed method, compared with other techniques.

Debdoot Sheet et al. [11] proposes a novel modification of the brightness preserving dynamic histogram equalization technique to improve its brightness preserving and contrast enhancement abilities while reducing its computational complexity. The modified technique, called Brightness Preserving Dynamic Fuzzy Histogram Equalization (BPDHE), uses fuzzy statistics of digital images for their representation and processing. Representation and processing of images in the fuzzy domain enables the technique to handle the inexactness of gray level values in a better way, resulting in improved performance. Execution time is dependent on image size and nature of the histogram, however experimental results show it to be faster as compared to the techniques compared here. The performance analysis of the BPDHE along with that for BPDHE has been given for comparative evaluation.

Tien-Ying Kuo et al. [15] proposed a simple yet effective image binarization method to work with various kinds of images, such as images with the color bleeding...
characters, or with non-uniform background exposed to poor ambient light, and so forth. Our approach is based on a hybrid color quantization process, which adaptively takes the global and local image characteristics into account; therefore, it can deal with complex images. The experiment result shows that our proposed hybrid method outperforms the literature methods.

Nija Babu et al. [17] presents a hybrid thresholding approach for binarization and enhancement of degraded documents. Historical documents contain information of great cultural and scientific value. But such documents are frequently degraded over time. Digitized degraded documents require specialized processing to remove different kinds of noise and to improve readability. The approach for enhancing degraded documents uses a combination of two thresholding algorithms. First, iterative global thresholding is applied to the smoothed degraded image until the stopping criteria is reached. Then a threshold selection method from gray level histogram is used to binarize the image. The next step is detecting areas where noise still remains and applying iterative thresholding locally. A method to improve the quality of textual information in the document is also done as a post processing stage, thus making the approach efficient and better suited for character recognition applications.

OD Trier et al. [26] Presents a methodology for evaluation of low-level image analysis methods, using binarization (two-level thresholding) as an example. Binarization of scanned gray scale images is the first step in most document image analysis systems. Selection of an appropriate binarization method for an input image domain is a difficult problem. Typically, a human expert evaluates the binarized images according to his/her visual criteria. However, to conduct an objective evaluation, one needs to investigate how well the subsequent image analysis steps will perform on the binarized image. We call this approach goal-directed evaluation, and it can be used to evaluate other low-level image processing methods as well. Our evaluation of binarization methods is in the context of digit recognition, so we define the performance of the character recognition module as the objective measure. Eleven different locally adaptive binarization methods were evaluated, and Niblack's method gave the best performance.

III. PROBLEM FORMULATION

In the existing system, the author have used edge detection technique for detecting the edges of the old documents manuscripts, though the technique was new and also the outputs were improved from the existing technique in literature but not that much accurate. Canny edge detection has a limit to constrain only inside the edge that means a part of text remains unexplored. From literature survey, we come across a number of suggested ways for Digital Image Binarization but the implementation of these ways inspires us to look forward for Region based Segmentation. In the proposed system we will try to implement existing system using morphological operators and will improve the values of parameters like PSNR, F-Measure and NRM.

IV. PROPOSED WORK

A. RGB to Gray Image Conversion

RGB is a device-dependent color model. The fundamental reason for the RGB shading model is for the sensing, representation and presentation of pictures in electronic systems. To form a color with RGB, three light beams (one red, one green, and one blue) must be superimposed. Each of the three beams is known as a component of that color.

B. Fuzzy Histogram Equalization

Technique consists of following operational Stages:
- Fuzzy Histogram Computation.
- Partitioning of the Histogram.
- Dynamic Histogram Equalization of the Partitions.
- Normalization of the image brightness.

C. Morphological Operators

This is a simple technique in which input as well as output both have binarized image. But other procedures use gray scale image as input and binary image as output by using threshold function. Basic Operation used in morphology are Dilation, Erosion, Opening and Closing etc. Dilation and Erosion are two fundamental operations. Dilation enlarges foreground, shrinks background. Erosion shrinks foreground, enlarges background. Opening and Closing derived from basic Dilation and Erosion operations. Opening is the dual of shutting (closing) i.e. opening the foreground pixels with a specific structuring element is equivalent to closing the background pixels with the same element.

![Flowchart](image.png)
scans for the limit that minimizes the intra-class fluctuation, characterized as the weighted sum of variations of the two classes.

$$\sigma^2(t) = \beta_1(t)\sigma^2(1) + \beta_2(t)\sigma^2(0)$$  \hspace{1cm} (4.4)

Otsu method gives better performance when the numbers of pixels in each class are close to each other. The expansion of the first technique to multi-level thresholding is referred to as the Multi Otsu method.

E. Sauvola’s Thresholding

The strategy proposed by Sauvola’s et al. is local-variance based method. It is a change on the system proposed by Niblack, particularly when the background contains light composition, enormous variations, re-colored and badly degraded documents. It adapts the contribution of the local mean and standard deviation. When document is dirty or re-colored paper then threshold value is lowered. The threshold is calculated as follows:

$$T(x, y) = m(x, y) \star \left[ 1 + k \left( \frac{\sigma(x, y)}{\bar{R}} \right)^2 - 1 \right]$$  \hspace{1cm} (4.5)

Where value of m(x, y) and \(\sigma(x, y)\) same as in Niblack system. Here, \(m\) and \(\sigma\) denotes the mean and standard deviation of the entire window. Sauvola’s estimate value of \(k=0.5\) and \(R=128\) and \(k\) is a fixed value. It was conclude that the suggestion of \(R\) has a little impact on the quality while the estimations of \(k\) and window size influence fundamentally. The smaller the estimation of \(k\), the thicker is the binarized stroke, and the more cover exists between characters. A smaller window size will create thinner strokes. An ideal blend of \(k\) and the sliding window will deliver a good binary image.

V. EXPERIMENTAL RESULTS

A. Proposed Technique

Segmentation of badly degraded document images is done for discriminating a text from background images but it is a very challenging task. So, to make a robust document images many binarization techniques are used. But in existing binarization techniques thresholding and filtering is an unsolved problem. In the existing method, edge based segmentation can be done and Canny edge detector used. In our proposed technique, Image Binarization for degraded document images has been use Region based segmentation. Firstly, an RGB image covert into gray image then image filtering can be done on the basis of Wiener Filtering and Gaussian filter. Secondly, morphological operators use to discriminate foreground from background. Then Otsu and Sauvola’s thresholding did for better results. Finally, proposed method results compare with the method used in DIBCO 2011 dataset.

---

Fig. 2: (Erosion Process) [4]

Fig. 3: (Dilation process) [4]

The opening of X by Y is obtained by the erosion of X by Y, followed by dilation of the resulting image by Y:

$$X \circ Y = (X \ominus Y) \oplus Y$$  \hspace{1cm} (4.1)

The closing of X by Y is obtained by the dilation of X by Y, followed by erosion of the resulting structure by Y:

$$X \bullet Y = (X \oplus Y) \ominus Y$$  \hspace{1cm} (4.2)

In our existing method, Canny edge detector has been used after contrast mapping but drawback is that it can be based on edge segmentation which is not capable for outer edges segmentation. In our proposed method, morphological operators can be used in place of canny edge detector which is based on Region based segmentation and segment both inner as well as outer edges for preserve text stroke pixels.

D. Thresholding

Thresholding is the first preprocessing scheme that can be used before document analysis. We can use two thresholding procedures in our proposed method. Firstly, Otsu and then for further improvement Sauvola’s thresholding.

1) Otsu Thresholding

Otsu Thresholding is most successful global thresholding technique. The simplest way for implementation of thresholding is to select an intensity value as threshold level. The value which is below the threshold level is treated as 0(black) which is above the threshold level that select as 1(white). If we assume that Z is the global threshold of an given image \(f(x, y)\) and then \(g(x, y)\) is the threshold image that can be given by as follows:

$$g(x, y) = \{ \begin{array}{ll} 1 & \text{when } f(x, y) \geq Z, \\ 0 & \text{otherwise} \end{array} \}$$  \hspace{1cm} (4.3)

The strategy proposed by Otsu explained a clustering analysis built technique based with respect to picture variation. It automatically performs histogram shape-based picture thresholding for the decrease of a grey level picture to a binary picture. The calculation expect that the image for thresholding contains two classes of pixels (e.g., foreground area and background) and then calculates the optimum threshold differentiating those two classes so that their combined spread is minimal. It exhaustively
B. Parameter Used

For evaluation there are few parameters that can be use to check the Binarization performance like
- F-Measure,
- Peak Signal to Noise Ratio (PSNR),
- Distance Reciprocal Distortion (DRD) and
- Misclassification Penalty Metric (MPM).

C. Testing on Competition Dataset

In this part, we compare our proposed method result with another techniques result that used in DIBCO 2011 dataset. This method includes Otsu’s method, Sauvola method, Brensen’s method, Gatos et al.’s method, LMM, BE, LELO, SNUS, HOWE methods and also Bolan Su Adaptive Contrast Mapping method.

The DIBCO 2011 dataset includes eight degraded handwritten documents and eight degraded printed documents. Therefore, total 16 document images. Table 5.1 shows the evaluation results as follows for figure 5(A) Binarization results of the sample document image (PR 08).

![Fig. 5(A): Binarization results of the sample document image (PR 08) (a) Input Image (b) Proposed Method Image](image1)

![Fig. 5(B): Binarization results of the sample document image (HW 08) (a) Input Image (b) Proposed Method Image](image2)

![Fig. 5(C): Binarization results of the sample document image (PR 06) in DIBCO 2011 dataset produced by different methods. (a) Input Image. (b) OTSU (c) SAUV(d) NIBL (e) BERN (f) GATO (g) LMM (h) BE (i) LELO (j) SNUS (k) HOWE (l) Bolan Su. (m) Proposed)](image3)

<table>
<thead>
<tr>
<th>Methods</th>
<th>F-Measure (%)</th>
<th>PSNR</th>
<th>DRD</th>
<th>MPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTSU</td>
<td>82.22</td>
<td>15.77</td>
<td>8.72</td>
<td>15.64</td>
</tr>
<tr>
<td>SAUV</td>
<td>82.54</td>
<td>15.78</td>
<td>8.09</td>
<td>9.20</td>
</tr>
<tr>
<td>NIBL</td>
<td>68.52</td>
<td>12.76</td>
<td>28.31</td>
<td>26.38</td>
</tr>
<tr>
<td>BERN</td>
<td>47.28</td>
<td>7.92</td>
<td>82.28</td>
<td>136.54</td>
</tr>
<tr>
<td>GATO</td>
<td>82.11</td>
<td>16.04</td>
<td>5.42</td>
<td>7.13</td>
</tr>
<tr>
<td>LMM</td>
<td>85.56</td>
<td>16.75</td>
<td>6.02</td>
<td>6.42</td>
</tr>
<tr>
<td>BE</td>
<td>81.67</td>
<td>15.59</td>
<td>11.24</td>
<td>11.40</td>
</tr>
<tr>
<td>LELO</td>
<td>80.86</td>
<td>16.13</td>
<td>104.48</td>
<td>64.43</td>
</tr>
<tr>
<td>SNUS</td>
<td>85.2</td>
<td>17.16</td>
<td>15.66</td>
<td>9.07</td>
</tr>
<tr>
<td>HOWE</td>
<td>88.74</td>
<td>17.84</td>
<td>5.37</td>
<td>8.64</td>
</tr>
<tr>
<td>BOLAN</td>
<td>87.8</td>
<td>17.56</td>
<td>4.84</td>
<td>5.17</td>
</tr>
<tr>
<td>Proposed</td>
<td>91.5</td>
<td>20.89</td>
<td>3.37</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Table 5.1: (Evaluation Results of the Dataset Of DIBCO 2011)

As shown in Table 5.1 our proposed method results is best in terms of all above parameters. That means our proposed method maintains best visibility and provides good text stroke contours.

Above figures 5 (A), (B) and (C) shows three document images (PR 08, HW-08 and PR-06) from DIBCO 2011 dataset.

VI. Conclusion

There are so many parameters are used in our proposed method to check the ability to remove the different kinds of degradation in an input document images.
Our proposed technique makes the document images stable and noise free.

Region based segmentation gives better performance instead of edge based segmentation.

Our proposed techniques extract foreground from background by using morphological operators.

The proposed method is easy, more reliable and an efficient way. The proposed method makes use of morphological operators then Otsu and Sauvola’s thresholding. The output can compare with DIBCO 2011 dataset on the basis of PSNR, F-measure, DRD and MPM.

VII. FUTURE SCOPE

This work can be further enhanced by using various pictures to improve the quality of image by using combination of different algorithms and by using different upgraded techniques which can make the image much clear.

REFERENCES


