Efficient Document Image Binarization for Degraded Document Images using MDBUTMF and BiTA

Leena.L.R , Gayathri. S2

1 Leena. L.R,Author is currently pursuing M.Tech (Information Technology) in Vins Christian College of Engineering. e-mail: leenalr699@gmail.com.

Abstract:
It is a challenging task for segmenting the foreground text from the background of the document. Salt and pepper noise can corrupt the images in which the corrupted pixels will take either maximum or minimum grey level. In the proposed technique, Modified Decision Based Unsymmetric Trimmed Median Filter algorithm processes the corrupted images by detecting the impulse noise. Salt and pepper noise is removed by the MDBUTMF algorithm. Global contrast enhancement is required to reveal hidden details in dark and bright regions. In addition to enhancing regions with extremely high or low luminance, proposed BiTA also significantly stretches the contrast in mid tone regions. The adaptive contrast map for degraded document is then constructed and binarized. Adaptive image contrast is the combination of both local image contrast and gradient. It is then combined with Canny’s edge detection method to identify the text stroke edge pixels. The local threshold is estimated from the output of Canny’s edge detection method. Based on those intensities, the text is segmented from the background of the degraded document accurately. The proposed method is simple, efficient and has minimum parameter tuning.

Keywords— Adaptive image contrast, MDBUTMF, BiTA, degraded document image binarization, pixel classification.

I. INTRODUCTION

Segmentation of the foreground text from the background of the document is very difficult due to the inter intra variations between the foreground and background in the historical degraded documents. The thresholding of the degraded document image is an unsolved problem.

The handwritten documents show variations in stroke width, stroke brightness and document background. Historical documents are often affected by ink bleed issues [1]. These problems will induce errors during the thresholding and makes segmentation of the foreground text from the background a quiet difficult task. This method presents the binarization method that extends the previous maxima minima method. The proposed method is simple, efficient and includes minimum parameter tuning. It makes use of local image contrast and local image gradient. In local maximum minimum method the image contrast alone is used to construct the contrast map. In r the proposed system both local image contrast and local image gradient is used to construct the adaptive contrast map.

Compared with the image gradient local image contrast is more capable of detecting the high contrast image pixels. Document image binarization divides the document image in to two classes. They are (1) foreground text (2) document background. It is usually performed on the preprocessing stage of document analysis. Compared with the printed documents the historical documents have the variations in the foreground and the background. So, fast and accurate document image binarization method is required.

Document image binarization methods uses a variety of scanned machine printed and hand written documents. There is currently a substantial and growing interest in the field of document image understanding where several research groups are developing and designing systems to automatically process and extract relevant information from documents such as drawings, maps, magazines, newspapers, forms and mail envelopes. Approaches dealing with the document image binarization are
traditionally differentiated into global and local approaches.

Global method uses a single computed threshold value to classify image pixels. The pre and post processing normally consist of speckle noise removal, hole filling and in some cases the merging of broken characters using heuristic rules. For some poor quality document images with variable and inhomogeneous document background intensity like shadows, smear or smudge, complex background patterns and signal-dependent noise, a practical problem is that no thresholding algorithm could work well for all kind of document images. The aim of post processing is to remove the noise and false print to improve the image quality.

The rest of the paper is organized as follows. Section II first reviews existing technologies. Section III describes the proposed system and the major work. Section IV presents the experimental results that shows the superior performance of the system when coared with other systems.

**II. RELATED WORKS**

Shijian et al. [2] proposed document image binarization using background estimation and stroke edges. This is a method in which estimates the document background first through an iterative polynomial smoothing procedure. Different types of document degradation are then compensated by using the estimated document background surface.

The text stroke edge is further detected from the compensated document image by using L1-norm image gradient. The document text is segmented by a local threshold that is estimated based on the detected stroke edges. Text stroke edges is used to estimate the local threshold and accordingly overcomes the limitation of many existing adaptive thresholding methods. L1-norm image gradient is often more suitable for the edge detector and edge profile used for the text stroke detection and based on empirical observations.

Yi Huang et al. [3] distinguished ink bleed reduction from the existing techniques by its allowance for user interaction and our dual layer MRF formulation. Ink bleed characteristics, namely that ink bleed intensity is lighter than the foreground or that the grey scale intensity values of ink bleed, foreground, and background follow a parametric distribution.

Rajat gupta et al. [4] proposed the concept of matched wavelet filters to develop the globally matched wavelet filters specifically adapted for the text and non text region. These filters are used for detecting text regions in scene images and for segmentation of document images into text, picture and background.

M.Cheriet et al. [5] proposed a new approach that goes beyond segmenting only one bright object from an image in to an approach that recursively segments the brightest object at each recursion, leaving the darkest object in the given digitized image. In digitized images, the uniformity of objects plays a significant role in separating these objects from the background. Ostu’s method in thresholding grey scale images is efficient on the basis of uniformity measure between the two classes C0 and C1 that should be segmented.

Eric Saund et al. [6] designed and implemented a customized pixel level labeling tool, for document images. Document images are characterized by certain kinds of structure that lend themselves to the use of accelerated user interface commands so that large collections of like label pixels can be tagged together.

Oivind Due Trier et al [7] evaluated binarization methods requiring manual inspection of sub images and fine tuning of parameters were executed. This is because the binarization algorithms are intended to be party of an automatic digitizing system for maps, being able to process a large number of maps per time unit.

**III. PROPOSED SCHEME**

In the proposed system, when a degraded document is given as input noise reduction is performed by MDBUTMF [8] and then it is enhanced using bilateral tone adjustment. The enhanced image is then used for adaptive contrast map construction and then it is binarized and Canny’s edge detection is done. Then the foreground text is segmented from the background.

**A. Noise Removal**

Every pixel of image is checked for salt and pepper noise. If the selected window contains salt/pepper noise as processing pixel and neighboring pixel values contains all pixels that adds salt and pepper noise to the image:
In the matrix (1), 255 is the processing pixel. The mean of the processing window is found and processing pixel is replaced by the mean value.

\[
\begin{bmatrix}
0 & 255 & 0 \\
0 & (255) & 255 \\
255 & 0 & 255
\end{bmatrix}
\]

(1)

In the matrix (1), 255 is the processing pixel. The mean of the processing window is found and processing pixel is replaced by the mean value. If the selected window contains salt or pepper noise as processing pixel and neighboring pixel values contains some pixels that adds salt and pepper noise to the image:

\[
\begin{bmatrix}
78 & 90 & 0 \\
120 & (0) & 255 \\
97 & 255 & 73
\end{bmatrix}
\]

(2)

In matrix (2), the processing pixel is 0. Now eliminate the salt and pepper noise from the selected window. Here the median value is replaced by median value. If the selected window contains a noise free pixel as a processing pixel:

\[
\begin{bmatrix}
43 & 67 & 70 \\
55 & (90) & 79 \\
85 & 81 & 66
\end{bmatrix}
\]

(3)

In matrix (3), 90 is the processing pixel. Since it is a noise free pixel it does not require further processing.

### B. Image Enhancement

It improves the visibility of details in dark and bright regions using the bilateral gamma adjustment [9].

\[
G_a(L) = (G_d + G_b) / 2 
\]

(4)

\[
G_d(L) = L^{1/\gamma} 
\]

(5)

\[
G_b(L) = 1 - (1 - L)^{1/\gamma} 
\]

(6)

Where L is the input luminance and \(\gamma\) is a user specified variable that indicates the degree of enhancement. Note that \(G_d\) is a convex function for enhancing dark regions and \(G_b\) is a concave function for bright regions. The final function of gamma adjustment \(G_a\) is the average of \(G_d\) and \(G_b\). Substantially, \(G_a\) is an inverse S curve which reveals hidden details in dark and bright regions. Although \(G_a\) can enhance the contrast in dark and bright regions, it unavoidably reduces the contrast in mid-tone regions.

### C. Adaptive Contrast Image Construction

To overcome the over normalization problem the local image contrast is combined with the local image gradient and adaptive local image contrast is calculated in eqn (7)

\[
C_a(i,j) = \alpha C(i,j) + (1 - \alpha)(I_{max}(i,j) - I_{min}(i,j)) 
\]

(7)

Where \(C(i,j)\) denotes the local contrast. \((I_{max}(i,j) - I_{min}(i,j))\) denotes the image gradient.

### D. Canny’s Edge Detection

The purpose of the contrast image construction is to detect the stroke edge pixels of the document properly. The constructed contrast image has a clear bi-modal pattern where the adaptive image contrast computed at text stroke edges is obviously larger than that computed within the document background. It therefore detects the text stroke edge pixel candidate by using Otsu’s global thresholding method. For the contrast images a binary map by Otsu’s algorithm that extracts the stroke edge pixels properly. As the local image contrast and the local image gradient are evaluated by the difference between the maximum and minimum intensity in a local window, the pixels at both sides of the text stroke will be selected as the high contrast pixels.

The binary map can be further improved through the combination with the edges by Canny’s edge detector because Canny’s edge detector has a good localization property that it can mark the edges close to real edge locations in the detecting image. In addition, canny edge detector uses two adaptive
thresholds and is more tolerant to different imaging artifacts such as shading. It should be noted that Canny’s edge detector by itself often extracts a large amount of non-stroke edges without tuning the parameter manually. In the combined map, it keep only pixels that appear within both the high contrast image pixel map and canny edge map.

Algorithm 1: Edge Width Estimation

Require: The Document Image I and Binary Text Stroke Edge Image Edg
Step 1: Get the width and height of I
Step 2: for Each Row i = 1 to height in Edg do
Step 3: Scan from left to right to find edge pixels that meet the following criteria:
   a) Its label is 0 (background)
   b) The next pixel is labeled as 1(edge)
Step 4: Examine the intensities in I of those pixels selected in Step 3, and remove those pixels that have a lower intensity than the following pixel next to it in the same row of I
Step 5: Match the remaining adjacent pixels in the same row into pairs, and calculate the distance between the two pixels in pair.
Step 6: end for
Step 7: Construct a histogram of calculated distances.
Step 8: Use the frequently occurring distance as the estimated stroke edge width EW

E. Threshold Estimation And Segmentation

The text can then be extracted from the document background pixels once the high contrast stroke edge pixels are detected. Two characteristics can be observed from the document images: First, the text pixels are close to the detected text stroke edge pixels. Second, there is a distinct intensity difference between the high contrast stroke edge pixels and the surrounding background pixels.

The neighborhood window should be at least larger than the stroke width in order to contain stroke edge pixels. So the size of the neighborhood window W can be set based on the stroke width of the document image under study, EW, which can be estimated from the detected stroke as stated in Algorithm

Algorithm 2: Post Processing Procedure

Require: The Input Document Image I, Initial Binary Result B and Corresponding Binary Text Stroke Edge Image Edg
Ensure: The Final Binary Result Bf
Step 1: Find out all the connected components of the stroke edge pixels in Edg.
Step 2: Remove those pixels that will not connect with other pixels.
Step 3: for Each remaining edge pixels (i, j): do
Step 4: Get its neighborhood pairs: (i – 1, j) and (i + 1, j) and (i, j – 1) and (i, j + 1)
Step 5: if the pixels in the same pairs belong to the same class (both text and background) then
Step 6: Assign the pixel with lower intensity to foreground class (text), and the other to background class.
Step 7: end if
Step 8: end for
Step 9: Remove single pixel artifacts along the boundaries after the document thresholding
Step 10: Store the new binary result to Bf.

IV. EXPERIMENTAL RESULTS

It quantitatively compare the proposed method with other state of the art techniques on DIBCO 2009, H-DIBCO 2010 and DIBCO 2011 datasets. These methods include Otsu’s method (OTSU), Sauvola’s method (SAUV) [18], Niblack’s method (NIBL), Bernsen’s method (BERN), Gatos et al.’s method (GATO), and the previous methods (LMM [5], BE). The three datasets are composed of the same series of document images that suffer from several common document degradations such as smear and smudge. The DIBCO 2009 dataset contains ten testing images that consist of five degraded handwritten documents and five degraded printed documents. The DIBCO 2011 dataset contains eight degraded handwritten documents and eight degraded printed documents. In total, we have 36 degraded document images with ground truth.
The superior performance of the proposed method can be explained by several factors. First, the proposed method combines the local image contrast and the local image gradient that help to suppress the background variation and avoid the over-normalization of document images with less variation. Second, the combination with edge map helps to produce a precise text stroke edge map. Third, the proposed method makes use of the text stroke edges that help to extract the foreground text from the document background accurately.

### TABLE I - EVALUATION RESULTS

<table>
<thead>
<tr>
<th>methods</th>
<th>F - measure</th>
<th>PSNR</th>
<th>DRD</th>
<th>MPM</th>
<th>Rank Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSTU</td>
<td>82.22</td>
<td>15.77</td>
<td>8.72</td>
<td>15.64</td>
<td>412</td>
</tr>
<tr>
<td>SAVU</td>
<td>82.54</td>
<td>15.78</td>
<td>8.09</td>
<td>9.20</td>
<td>403</td>
</tr>
<tr>
<td>BERN</td>
<td>47.28</td>
<td>7.92</td>
<td>82.28</td>
<td>136.54</td>
<td>664</td>
</tr>
<tr>
<td>GATO</td>
<td>82.11</td>
<td>16.04</td>
<td>5.42</td>
<td>7.13</td>
<td>353</td>
</tr>
<tr>
<td>LMM</td>
<td>82.56</td>
<td>16.75</td>
<td>6.02</td>
<td>6.42</td>
<td>516</td>
</tr>
<tr>
<td>BE</td>
<td>81.67</td>
<td>15.59</td>
<td>11.24</td>
<td>11.40</td>
<td>376</td>
</tr>
<tr>
<td>SNUS</td>
<td>85.2</td>
<td>17.16</td>
<td>15.66</td>
<td>9.07</td>
<td>279</td>
</tr>
<tr>
<td>Howe</td>
<td>88.74</td>
<td>17.14</td>
<td>5.37</td>
<td>8.64</td>
<td>299</td>
</tr>
<tr>
<td>proposed</td>
<td>87.8</td>
<td>17.56</td>
<td>4.84</td>
<td>5.17</td>
<td>307</td>
</tr>
</tbody>
</table>

### V. CONCLUSION

This paper presents an adaptive image contrast based document image binarization technique that is tolerant to different types of document degradation such as uneven illumination and document smear. The proposed technique is simple and robust, only few parameters are involved. Moreover, it works for different kinds of degraded document images. The proposed technique makes use of the local image contrast that is evaluated based on the local maximum and minimum. The proposed method has been tested on the various datasets. Experiments show that the proposed method outperforms most reported document binarization methods in term of the F-measure, pseudo F-measure, PSNR, NRM, MPM and DRD.

In future work it can explore the proposed algorithm with noise reduction for of noise in degraded documents and image enhancement to improve accuracy. In the future enhancement, we have planned to improve the segmentation quality by improving the accuracy.

### REFERENCES


