Image Segmentation for Document Image Binarization: A Brief Review

Kirti S. Datir¹, Prof. J.V. Shinde²
datir.kirti0@gmail.com¹, jv.shinde@rediffmail.com²
M. E. Student, Dept. of Computer Enng. Late G.N. Sapkal COE, Nashik, Maharashtra, India¹
Assistant Professor, Dept. of Computer Enng. Late G.N. Sapkal COE, Nashik, Maharashtra, India²

Abstract

Binarization is nothing but to generate binary image from document image. Document image binarization has already under study from past many years, and many binarization algorithms have been proposed for degraded document images. Document image Binarization is very trendy to improve old handwritten and machine printed documents. Still to recover degraded document is very tedious job. Such document has the much scratched also presence of noise and degradation. There is a lot of scale to improve old and degraded documents. Image segmentation is method which used usually in image processing. Thresholding is an important pre-processing step for the degraded image to improve their quality. The between the foreground text and the background of different document images is a tricky task. New Binarization method using image segmentation is proposed for better results.

Keywords - Document image binarization, Color-to-gray image conversion, Thresholding, Historical document analysis.

1. Introduction

Historical documents are of great interest to scholars in the community sciences and humanities. The value of historical documents is very much improved for consultation by the general public and for research purposes through digitization, which involves acquisition, processing and dissemination of knowledge. A historical document is unique, i.e., it does not have multiple copies. It contains specific difficulties impeding access to its content, e.g., the presence of physical degradation caused by environmental conditions, dust, dirt, etc. Such phenomena continue to harm these precious objects, and thus there is an urgent need for a method of preserving and providing broader access to them.

Image binarization is one of the pre-processing shackles of document image enhancement required before different tasks of analysis and recognition (i.e., OCR). Document image binarizations, in particular those based on thresholding strategy, aim at finding an optimal threshold (gray-level) which separates the document image pixels into two classes, foreground and background.
Though document image binarization has been studied for many years, the thresholding of degraded document images is still an unsolved problem due to the high intra/inter variation between the text stroke and the document background across different document images.

As illustrated in example Fig. 1, the handwritten text within the degraded documents often shows a certain amount of variation in terms of the stroke width, stroke brightness, stroke connection, and document background. In addition, historical documents are often degraded by the bleed-through as illustrated in Fig. 1(a) and (c) where the ink of the other side seeps through to the front. In addition, historical documents are often degraded by different types of imaging artifacts as illustrated in Fig. 1(e). These different types of document degradations have a tendency to induce the document thresholding fault and make degraded document image binarization a big challenge to most state-of-the-art techniques.

2. Literature survey

Algorithms for converting color to gray scale

1. The lightness method averages the most well-known and least well-known colors: \((\max(R, G, B) + \min(R, G, B)) / 2\).[16]

2. The average method simply averages the values: 
\((R + G + B) / 3\).[16]

Average method is the most simple one. You just have to take the average of three colors. Since its an RGB image, so it means that you have add \(r\) with \(g\) with \(b\) and then divide it by 3 to get your desired grayscale image.

3. The luminosity method is a more complex version of the average method. It also averages the values, but it forms a weighted average to account for human observation. We’re more perceptive to green than other colors, so green is weighted most heavily. The formula for luminosity is \(0.21 R + 0.72 G + 0.07 B\).[16]

4. Decolorize

Goals of algorithm:

- **Contrast Magnitude**: The magnitude of the gray scale contrasts should visibly reveal the magnitude of the color contrasts.
- **Contrast Polarity**: The positive or negative polarity of gray level change in the gray scale contrasts should visibly correspond to the polarization of luminance change in the color contrasts.
- **Dynamic Range**: The dynamic range of the gray levels in the gray scale image should plainly accord with the dynamic range of luminance values in the color image.[12]

5. Color2Gray

The Color2Gray algorithm has three steps: first, convert a color image to a perceptually regular color space, then compute target differences in order to combine luminance and chrominance differences, and finally, use a least squares optimization to selectively modulate the source luminance differences in order to replicate changes in the source image’s chrominance.[13]

Algorithm of Binarization

1. **Otsu’s Binarization Method**

Otsu's method [1], is used clustering-based image thresholding, or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels foreground pixels and background pixels, it then calculates the optimum threshold sorting out the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal.

In Otsu’s method carefully search for the threshold that minimizes the intra-class variance (the variance within the class), defined as a weighted sum of variances of the two classes:

\[
\sigma^2_t(t) = \omega_1(t) \sigma^2_1(t) + \omega_2(t) \sigma^2_2(t)
\]
2. Adaptive and threshold-based binarization method

The method proposed by Sauvola et al. [17] is a local-variance-based method. It is an improvement on the method proposed by Niblack [18], especially when the background contains light texture, big variations, stained and badly and unevenly illuminated documents. It adapts the contribution of the standard deviation. For example, in the case of text on a dirty or stained paper, the threshold is lowered. The threshold is calculated as follows:

\[ T(i,j) = m(i,j) \times \left[ 1 + k \left( \frac{\sigma(i,j)}{R} - 1 \right) \right] \]

The typical values of \( k = 0.5 \) and \( R = 128 \) are suggested. Here, \( m \) and \( \sigma \) are again the mean and standard deviation of the entire pane, and \( k \) is a fixed value. It was found that the value of \( R \) has a very small effect on the significance while the values of \( k \) and window size affect it significantly. The smaller the value of \( k \), the thicker is the binarized stroke, and the more overlap exists between characters. A smaller window size will generate thinner strokes. An optimal combination of \( k \) and the down window will produce a good binary image.

3. Lu et al.’s algorithm [6]

This algorithm consists of the following three steps:
1. The document background plane is estimated using an iterative polynomial smoothing procedure, which makes it possible to compensate for the various levels of document degradation.
2. A histogram of local image variation of the contender text stroke edge pixels is calculated, and then the real text stroke edge pixels can be detected by using Otsu’s global thresholding method.
3. Text extraction based on the estimated stroke width is performed using a local thresholding technique.

3. Proposed System

In pre-processing of document image binarization is translation of gray scale image from RGB image or color image. Gray scale image is necessary for the elimination of noise, smoothing of background texture of degraded input document. Then on that gray scale image new image segmentation algorithm is applied to segment image into windows. In this method, each pixel in the image has its own threshold by calculating the statistical information of the greyscale values of its neighbourhood pixels. According to threshold value gray scale image can binarize. Some post-processing algorithm is applied on the binarize image. Foreground pixels that is separate from other foreground pixels are clean out. Post – Processing can also connect break edges due to degradation. So we can produce more clear binarize image.

![Fig 2.Block Diagram](image)

4. Conclusion

In this paper an overview of gray scale algorithms and binarization algorithms has been present and according to the result of existing algorithm conclude that there does not exist a universal algorithm for segmenting all types of images so, image segmentation algorithm is proposed for degraded document image binarization. Its require less parameter setting which makes it simple and robust. The Post-Processing technique is useful to generate clear document image due to connecting unclear edges. The proposed algorithm will takes little time to get a accurate result when small size structural operator is selected.

5. Acknowledgement

I am thankful to Prof J .V. Shinde Assistant professor in the Department of Computer Engineering in Late G.N. Sapkal College of Engineering, Nasik. For providing constant
guidance and encouragement for this research work.

6. References


