

A BINARIZATION TECHNIQUE FOR EXTRACTION OF DEVANAGARI TEXT FROM CAMERA BASED IMAGES

Rajesh K. Bawa¹ and Ganesh K. Sethi²

¹Department of Computer Science, Punjabi University, Patiala, India

²M.M. Modi College, Patiala, India

ABSTRACT

This paper presents a binarization method for camera based natural scene (NS) images based on edge analysis and morphological dilation. Image is converted to grey scale image and edge detection is carried out using canny edge detection. The edge image is dilated using morphological dilation and analyzed to remove edges corresponding to non-text regions. The image is binarized using mean and standard deviation of edge pixels. Post processing of resulting images is done to fill gaps and to smooth text strokes. The algorithm is tested on a variety of NS images captured using a digital camera under variable resolutions, lightening conditions having text of different fonts, styles and backgrounds. The results are compared with other standard techniques. The method is fast and works well for camera based natural scene images.

KEYWORDS

Text information extraction (TIE), Global Thresholding, Local Thresholding, Canny edge detection, Morphological dilation

1. INTRODUCTION

Extracting text from camera based natural scene images is very challenging problem due to variation in font size, style, complex backgrounds, shadows, reflections from background surface, and uneven lightening conditions. To extract the textual information from the images, the images is segmented into its constituent components and the text part is separated from nontext part. For segmenting the image, the first step applied is image thresholding or image binarization. This step is very crucial as the success of complete method depends upon it. A large number of techniques have been given in the literature. A survey of binarization techniques for nondestructive testing images and document images was given by Sezgin and Sankur [3] in which various techniques were divided into six categories: histogram shape, measurement space clustering, entropy, object attributes, spatial correlation, and local gray-level surface. No technique fits well for all type of images has every technique has its pros and cons. Binarization of camera based natural scene images for the purpose of text extraction is a challenging task due to complex backgrounds, uneven lighting conditions, shadows, different types of text (font, style,

orientations). So the algorithms of binarization which produce good results for document images sometimes fail completely for NS images.

Binarization techniques can be divided into two categories: Global and Local. Most of the binarization techniques in literature follow the architecture given in figure 1. Preprocessing is required to reduce noise so as to make binarization successful and post-processing of binarized image is done so fill gaps in character strokes and to remove non-text areas. A chronological listing of binarization techniques from literature is given in Table 1.

Method	Year	Category	Approach
Otsu [4]	1979	Global	Histogram Analysis to minimize within class variance
Iterative [5]		Global	Iterations till convergence
Minimum Error Thresholding [6]	1985	Global	Optimise Average Pixel Classification Error Rate
Niblacks [7]	1986	local	Mean and Standard deviation of sub-image
Sauvola [8]	2000	local	Mean and Standard deviation of sub-image with dynamic range of standard deviation
B. Gato et. al. [12]	2005	local	Sauvola method with nearest neighbour interpolation
T.Kasar [13]	2007	local	Edge Based Binarization
Extreme Value Based [9]	2010	Global	Gamma Transforms and Extreme Value Theory

Table 1 A chronological listing of binarization techniques

2. PROPOSED METHOD

The proposed methodology for binarization is based on edge detection and morphological dilation. The flowchart of the proposed methodology is presented in Figures 1 and 2.

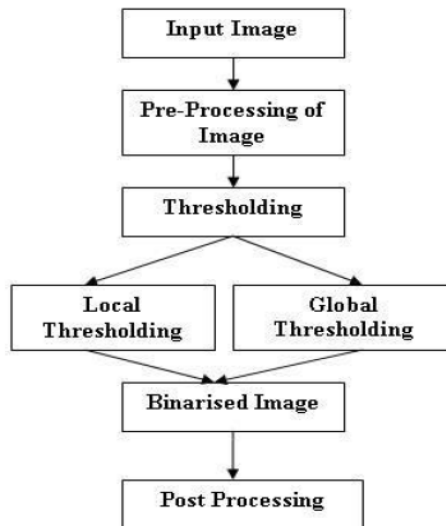


Figure 1: Flow diagram for Binarization of Image

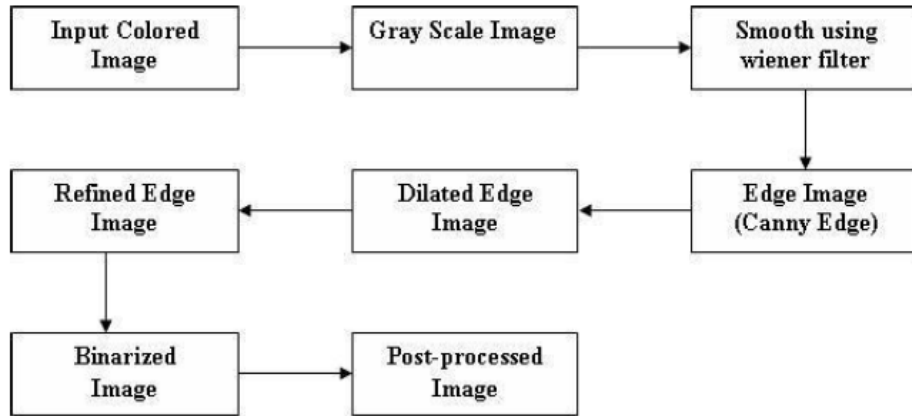


Figure 2: Proposed Method for Binarization of Images

2.1. Image preprocessing

The input colored image is first converted to a grey scale image. Since the images are captured under real conditions, these may contain noise of different types. To minimize the effect of noise, a pre-processing stage is required for smoothing of background texture and contrast enhancement between background and text areas. We applied a low-pass Wiener filter [14] to the grey scale image to meet the above goals. Original image captured by camera is shown in figure 3 and corresponding grey scale image after applying Wiener filter is shown in figure 4.



Figure 3: Original Colour Images Captured by digital camera



Figure 4: Grey Scale Image after preprocessing

2.2. Edge Detection

The text in images is mostly highlighted as compared to background objects. This results in difference in intensity levels between text regions and background regions in grey scale image. The change of intensity creates edges that can be detected using edge detection technique. The proposed method of binarization exploits the edges that are formed by text components in the image due to the difference in contrast with the background. The Canny edge detector [15] is used for detecting the edges. Canny edge detector works as follows:

1. Smooth the image with a two dimensional Gaussian.
2. Take the gradient of the image.
3. Non-maximal suppression.
4. Edge Thresholding.



Figure 5 Edge Image formed by Canny Edge Detection

2.3. Morphological Dilation

Morphology is a broad set of image processing operations that process images based on shapes. Morphological operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. By choosing the size and shape of the neighborhood, we can construct a morphological operation that is sensitive to specific shapes in the input image. Dilation is the most basic morphological operation. Dilation adds pixels to the boundaries of objects in an image. The number of pixels added in an image depends on the size and shape of the structuring element used to process the image. For the proposed method we applied horizontal dilation followed by a vertical dilation using a structuring element of size 1×3 and 3×1 respectively to connect the edges obtained in previous step. The text in natural scene images mostly appears in the center of image and is not connected to the image borders. Using this assumption we eliminated all the edges that are touching the image border.



Figure 6 Edge Image after morphological dilation and eliminating border connected edges

2.4. Edge Analysis

After applying morphological operations, the next step is to analyze the resulting edge image. For this first we assign the unique label to each edge component by performing 8 connected analyses and computed edge box for each component. Since all the edge components are not text, we used some heuristics to filter edge components:

- All the edge boxes having aspect ratio less than 0.2 and more than 15 are ignored as text normally fall within this aspect ratio.
- Next all edges having height more than half of the image height and width more than 90 percent of image width are ignored.
- Next all edge boxes having area less than 20 pixels and more than 20th part the image are ignored.

After filtering we analyzed the remaining edge boxes. It may happen that one box may contain one or more edge boxes into it. If it happens then we simply ignore the inner boxes and retain the outer box to reduce the computations. The resulting boxes are considered for the next step.

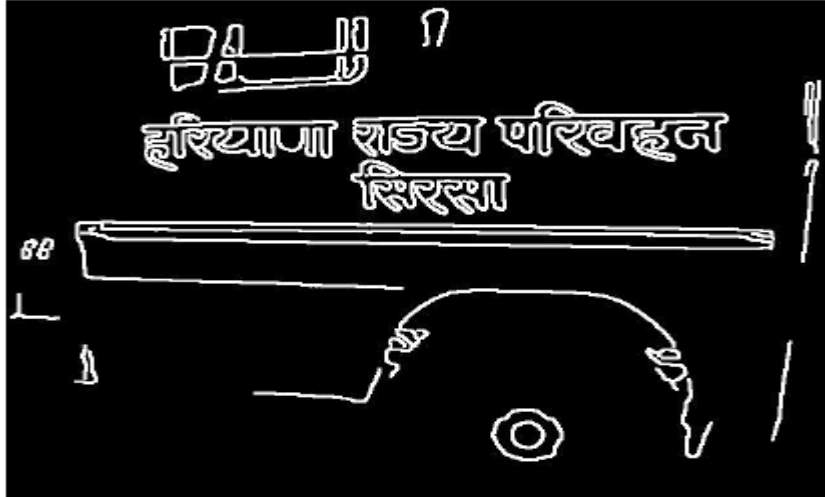


Figure 7: Edge Image after performing edge analysis

2.5. Binarization

For binarizing the image we considered only the filtered edge image obtained in previous step. In each edge box we calculated the mean and standard deviation of edge pixels by taking intensity from gray scale image. The threshold value for each edge box is calculated using the relation.

$$T(x, y) = m(x, y) + k *s(x, y)$$

where $m(x, y)$ and $s(x, y)$ are the mean intensity and standard deviation of the edge pixels in the edge box respectively and k is a parameter set at 0.2 for larger components and -0.1 for smaller components. The complete image is binarized against a white background. In some edge boxes we inverted the output using the logic given in [13] by using median of intensity of corner pixels.

2.6. Post Processing

Post processing is based on shrink and swell filter [12]. A shrink filter is used to reduce noise from the background. The resulting binary image is scanned and each foreground pixel is checked. If BGN is the number of background pixels in a sliding $n \times n$ window, which has the foreground pixel as the central pixel, then this pixel is changed to background if BGN is more than 'k' where k can be defined experimentally. A swell filter is used to fill possible breaks, gaps or holes in the foreground. The entire binary image is scanned and each background pixel is examined. If FGN is the number of foreground pixels in a sliding $n \times n$ window, which has the background pixel (x, y) as the central pixel, then this pixel is changed to foreground if FGN is more than 'm' where m can be defined experimentally.

3. RESULTS AND PERFORMANCE EVALUATION

For the purpose of performance evaluation we collected data set containing images of natural scene images under real conditions using a digital camera of upto 14 mega pixels. The images are captured at various resolutions, under different lightening conditions, from various sources. The results obtained by the proposed algorithm and those obtained by other standard techniques are given at the end of paper in Table 2.



Figure 8: Image after performing Binarization






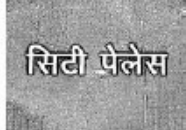










Figure 9: Binarized Image after performing post processing

4. CONCLUSIONS AND FUTURE SCOPE

In this paper we have analyzed different global and local methods of thresholding an image and then proposed a binarization technique based on edge detection and morphological dilation operation. As the results show proposed technique is better as compared to other methods and is capable of extracting text from a variety of natural scene images. After binarizing the image we have a set of connected components which has to be classified into text and non-text. So the future plan is to extract features from these components and classify into text and non-text and finally compare the results with other methods.

Table 2 : Comparison of Results of various techniques

Original	Otsu	Niblack	Sauvola	Proposed
				
				
				
				
				

REFERENCES

- [1] Jung K., Kim K. I. & Jain A. K., (2004) "Text Information Extraction in Images and Video : a Survey" Pattern Recognition Vol. 37, pp 977-997.
- [2] Sethi G. K. & Bawa R. K., (2011) "Text Information Extraction : In Context of Indian Languages", Proceedings of International conference on Control, Robotics and Cybernetics (ICRC 2011), Vol. 2, 2011, pp 221-226.
- [3] Sezgin M. & Sankur B., (2004) "Survey over image thresholding techniques and quantitative performance evaluation" Journal of Electronic Imaging, Vol. 13(1), pp 146-165.
- [4] Otsu N., (1979) "A threshold selection method from gray level histograms" IEEE Transactions on Systems, Man, and Cybernetics, SMC-9, pp 62-66.
- [5] Gonzalez R.C. & Woods R.E. (2009) "Digital Image Processing" Third Edition, pp 763-764.
- [6] Kittler J. & Illingworth J., (1986) "Minimum error thresholding" Pattern Recognition, Vol 19(1), pp 41-47.

- [7] Niblack W., (1986) "An Introduction to Image Processing" Prentice-Hall, Englewood Cliffs, NJ.
- [8] Sauvola J. & Pietikainen M., (2000) "Adaptive document image Binarization" Pattern Recognition, Vol 33(2), pp 225-236.
- [9] Fernando B. & Karaoglu S., (2010) "Extreme Value Theory Based Text Binarization In Documents and Natural Scenes" The 3rd International Conference on Machine Vision (ICMV 2010), pp 144-151.
- [10] Gatos B., Ntirogiannis K. and Pratikakis I., (2009) "DIBCO 2009 Document Image Binarization Contest" <http://users.iit.demokritos.gr/bgat/DIBCO2009>
- [11] ICDAR 2003 Robust reading and text locating competition image database <http://algoval.essex.ac.uk/icdar/Datasets.html>
- [12] Gatos B., Pratikakis I. and Perantonis S.J., (2005) "Text Detection in Indoor/Outdoor Scene Images" First International Workshop on Camera-based Document Analysis and Recognition (CBDAR'05), pp 127-132.
- [13] Kasar T., Kumar J. and Ramakrishnan A.G., (2007) "Font and Background Color Independent Text Binarisation" Camera Based Document Analysis and Recognition (workshop of ICDAR), pp. 3-8.
- [14] Jain A., (1989), "Fundamentals of Digital Image Processing" Prentice Hall, Englewood Cliffs.
- [15] Canny J., (1986) "A Computational Approach to Edge Detection" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, pp 679-698.

Authors

1. Rajesh K. Bawa is Professor and Head , Dept. of Computer Science, Punjabi University Patiala. He is MSc, PhD. His areas of research includes Parallel, Scientific Computing and Digital Image Processing.
2. Ganesh K Sethi is Assistant Professor in Dept. of Computer Science, MM Modi College Patiala. He is MCA (NET), MSc(NET). Currently he is doing research in the area of Digital Image Processing under the guidance of Rajesh K. Bawa