An Analysis of Image Binarization Techniques for Natural Scene Images

Dharam Veer Sharma
Assistant Professor
Computer Science Department,
Punjabi University Patiala

Sukhdev Singh
Assistant Professor
Mutani Mal Modi College Patiala

ABSTRACT
Text extraction from natural scene images is an emerging field in computer graphics. Extracted text contains important information that can be used for various purpose like vehicle number plate detection to identify the vehicle, to provide information of surrounding to visually impaired persons, preservation of information of historical documents etc. Binarization is a key process in text extraction process. It is challenging take in case of natural scene images due to uneven lighting conditions, complex background and unpredicted text size, color and layout. Three well known binarization techniques namely Otsu's, Niblack's and Sauvola's binarization techniques are test on natural scene images. We found that, Sauvola's algorithm can achieve better performance than Niblack's. In most of cases Sauvola and Niblack gave good results as compare to Otsu's method. Otsu binarization technique is good for uniform background. Window based Niblack's and Sauvola's methods are useful to find local threshold to binrize natural scene images.

Keywords
Binarization,Otsu, Sauvola; images thresholding, Niblack, local thresholding,Text extraction.

1. INTRODUCTION
With the increasing availability of digital imaging devices such as digital camera and mobile camera, a new field for text recognition from natural scene images is expanding. It has many applications like, the navigation systems for visually impaired persons [1]. This system may help visually impaired persons to walk freely by providing surrounding information. Such system may have GPS facility and a hand held Camera system to capture surrounding area to recognize text from it. Text extraction techniques can be helpful to extract text from historical documents [2] as a lot of vital information is preserved in historical documents. The physical condition of most of the documents is too bad that even Photostat of these papers is not possible. In such a case a camera based system is used to extract text from these camera images. In this way these vital information can be preserved for a long time. Similarly text extraction from natural scene is useful to overcome languages barrier. When the people from foreign lands visit India, they face difficulty in interacting with local residents and interpret messages written on notice boards, sign boards and direction boards. The written matter on display/sign boards provides information necessary for the needs and safety of people and may be written in unknown languages. The written matter can be names of streets, restaurant names, building names, company names, traffic directions, warning signs, etc. It is the language barrier for foreign origin persons who want to move across India for business, field works and/or pleasure. Text written in unknown language can be extracted from natural scene images captured by camera and converted into known language extraction from natural scene.

In all the cases text extraction techniques are used and binarization is one of the key phases of this process. Binarization is a process to convert a color or gray scale image into blank and white image. It is done to divide the image into foreground and background pixels so that text extraction can easily perform. In case of binarization of natural scene images there are lots of complexities like multicoloured background, different size of the text, different layout of text, similarity of text with other object in the background. These are the challenges to binarize a natural scene images. Generally, binarization algorithms are classified into two categories [3] based on discontinuity or similarity of grey values. Algorithms, using discontinuity, segment an image based on abrupt changes in grey level, whereas algorithms using similarity are based on threshold. In case of thresholding algorithms, the conversion is based on finding a threshold grey value and deciding whether a pixel having a particular grey value is to be converted to black or white. Usually within an image the pixels having grey value greater than the threshold is transformed to white and the pixels having grey value lesser than the threshold is transformed to black. It is further divided into two categories local threshold and global threshold. The simplest and oldest the method is to find a global threshold for the whole image and binarize the entire image using the single threshold value. In this technique, the local variations are actually suppressed or lost, though they may have important contribution towards the information content within it. On the other hand, local thresholding is very effective. In this method, a window is used around a pixel and threshold value is calculated for the window. In this case threshold value can be calculated for the entire region of the window, the binarization is done on pixel-by-pixel basis, where each pixel may have a calculated threshold value, or on region-by-region basis where all pixels in a region or window have same threshold value. In this paper we have taken Otsu's Global Binarization method, Niblack's window based binarization technique and Sauvola's window based binarization technique.

2. Challenges in Binarization of Natural Scene Images
Natural scene images have complex background usually not present in machine printed ones such as

- Shadow mark on Text Area in natural scene image
- Text background with different colors.
- Text Orientation and variable color of text
- Perspective Projection of Text on tiled surface
- Uneven lighting(Sharp bright spot due to sun light)
- perspective distortion
- foreground and background almost same in scene

Text extraction process faces above problems in binarization stage because of numerous degradations in natural scene.

3. Binarization Techniques

3.1 Otsu's Binarization Technique

It is a global binarization method. It assumes that the image to be threshold [4] contains two groups of pixels (e.g. foreground and background) and calculates the optimum threshold value to separating those two groups so that their intra-class variation is minimal.

Where intra-class variance is calculated as

$$\sigma^2_{\omega}(t)=\omega_1(t)\sigma^2_1(t)+\omega_2(t)\sigma^2_2(t)$$

Otsu's algorithm is not satisfactory performance in uneven illumination and is not real-time implementation.

3.2 Niblack's Binarization Technique

Niblack's algorithm [5] is a local thresholding method based on the calculation of the local mean and local standard deviation. In this method the averaged grey level of a neighbourhood of a given point is used to set a threshold for that point. The threshold is decided by the formula:

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

Where, m(x, y) and s(x, y) are the average of a local area and standard deviation values, respectively. Drawback of this method is a considerable sensitivity to window size and the persistence of background noise in the output image .The improved Niblack's method uses parameters k and R to reduce its sensitivity to noise.

$$T(x, y) = m(x, y) \cdot [1+k \cdot (1-s(x, y)/R)]$$

Where, k and R are empirical constants.

3.3 Sauvola's Binarization Technique

Sauvola's binarization technique [6] is window based, which calculates a local threshold for each image pixel at (x, y) by using the intensity of pixels within a small window W (x, y). Here we have taken the window of size N x N pixels with (x, y) as centre except at the edge pixels of the image frame. The threshold T (x, y) is computed using the following formula:

$$T(x, y) = M(1 + k(SD/R - 1))]$$

Where M is the mean of gray values in the considered window W (x, y), SD is the standard deviation of the gray levels ,R is the dynamic range of the variance(Standard Deviation), k is a constant (0 to1).

4. EVALUATION AND COMPARISION

Comparison of these binarization algorithms prove to be difficult task since there is no efficient way to compare the results. In literature several papers are found that use precision and recall analysis. But in case of natural scene image, the background is too complex that it is not easy to separate object from it. There is lot of text like objects like tree leaf, branches, window, etc. In current study we consider visual observation is a most efficient method to observe the effect of binarization techniques separating text from background. We have collected 1,027 images using mobile camera (5.0 mega pixel) of different location at Patiala City. These Natural scene images contain sign board, notice board, banners, and text written on wall and vehicles. Text present in these boards is in Gurmukhi Script. Images are captured at different locations and at time so that different lighting conditions are covered. Images are resized to cover come computation over head. Test images are of size 336 x 448 in jpeg format. The Window size is 15 x15 considered in case of Niblack's method and Sauvola's method. The value of parameter k is set as 0.5 and -0.2 in case of Sauvola's method and Niblack's method respectively. The following table shows original image with outputs of different methods.

Fig 1: Visual Analysis of Binarization Techniques: Otsu,Niblack,Sauvola on Natural Scene Images

Sno.	Original Image	Otsu's Binarization	Niblack's Binarization	Sauvola's Binarization
1.				

Sno.	Original Image	Otsu's Binarization	Niblack's Binarization	Sauvola's Binarization
2.				
3.				
4.				
5.				
6.				
7.				
8.				

5. CONCLUSIONS

This paper has presented a comparison of binarization techniques using Visual observation. Otsu's, Niblack's and Sauvola's techniques has been carried out on more than 1000 samples images. We drew several conclusions.

- The local Threshold methods gave better result as compare to global threshold methods as Niblack's and Sauvola's algorithms performed batter than Otsu's algorithm.
- The local Threshold methods of Sauvola and Niblack's are sensitive to window size. The window small size (smaller than 7x7) shows more noise where as Size greater than 15 x 15 take more computation time.
- These methods are also sensitive to value of k, best value for Sauvola's method is 0.5 and -0.2 for Niblack;s method. The value can be change to control in noise level in the image.
- Sauvola's method shows better performance in case of small size text surrounding by complex back ground.
- Otsu's Binarization method is efficiently binarize images where text has uniform background but it fails when text background complex.
- The images with uneven lighting condition such as sun sport or sharp reflection of sun light are not binarized by global Ostu's method. Such images may have uniform background even than small spot with high value of intensity create problem for global thresholding.
- Local threshold methods are good to deal images with uneven lighting and high intensity sport. Both Sauvola's and Niblack's methods give good results. But results of Sauvola are better as there is very less amount of noise.
- Both the local thresholding methods gave poor result where text in the scene is at distance or text size is very small.

6. REFERENCES

[1] Nobuo Ezaki, Marius Bulacu and Lambert Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", in the proceedings of 17th International Conference on Pattern Recognition, Cambridge ,UK, vol. 2, 683-686, 2004.

[2] D. E. Ventzas, N. Ntogas, "A Binarization Algorithm for Historical Manuscripts", in the proceedings of 12th WSEAS International Conference on Communications, Greece, pp.41-51, 2008.

[3] M. Sezgin, B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", Journal of Electronic Imaging,146-165, 2004.

[4] Wayne Niblack. An Introduction to Digital Image Processing. Prentice Hall, New Jersey, 1986.

[5] J.Sauvola,T,Seppanen, S.Haapakoski and M.Pietikainen, "Adaptive Document Binarization", in proceedings of 4th International Conference on Document Analysis and Recognition, Ulm, Germany, August 1997,147-152.

[6] J. Sauvola, M. Pietikainen, "Adaptive document image binarization", Journal of Pattern Recognition,225-236, 2000