Abstract

In this paper we propose a method of document image binarization that performs well on
grayscale images with complex backgrounds, maintains good text extraction abilities and retains the graphic features that might be present in the image. The proposed method employs a coarse thresholding step that uses a contrast feature for classification of pixels into foreground and background followed by Niblack thresholding for finer classification of the pixels. The proposed method was found to perform better or at-par with four other popular thresholding methods that it was compared against.

References

- B. Gatos, I. Pratikakis, and S. J. Perantonis, "Improved document image

- Bitbucket URL: http://bit.ly/1Hi9H1

**Index Terms**

Computer Science  
Image Processing

**Keywords**

Document Image Processing  
Binarization  
Niblack Thresholding