

A Weighted Hybrid Thresholding Approach for Text Binarization

S.T.Deepa

Computer Science dept. Shri Shankarlal
Sundarbai Shasun Jain College for Women
Chennai, India

S.P.Victor

Phd, Computer Science dept. St.Xavier's
College
Palayamkottai, India

ABSTRACT

Text extraction in real images taken in unconstrained environments remains surprisingly challenging in Computer Vision due to language characteristics, complex backgrounds and the text color. Extraction of text and caption from images and videos is important and in great demand for video retrieval, annotation, indexing and content analysis. In this paper we propose a weighted hybrid thresholding approach. It is demonstrated that the proposed method achieved reasonable accuracy of the text extraction for moderately difficult examples.

General Terms

Image processing, image segmentation.

Keywords

Thresholding, weight, hybrid approach

1. INTRODUCTION

In today world, video place a major role in the area of data mining. Superimposed text in images gives lot of information. Extraction of text in images is of greater demand.

A comprehensive strategy combining bottom-up and top-down mechanisms to detect textboxes is present in [1]. The bottom up part is based on Character segmentation and grouping. The top-down part is achieved with a statistical learning approach based on box descriptors.

A method is presented for adaptive document image binarization [2], where the page is considered as a collection of subcomponents such as text, background and picture. Using characteristics analysis, two new algorithms are applied to determine a local threshold for each pixel. An algorithm based on soft decision control is used for thresholding background and picture regions. An approach utilizing local mean and variance of gray values is applied to textual regions.

The historical manuscripts digital processing algorithm consists of 5 stages viz image acquisition, image preparation, denoising, thresholding and refinement [3]. The images are acquired by a digital camera with high resolution ratio. The images are converted to grayscale. A refinement procedure is based on erosion and dilation.

A binarization technique based on a probabilistic MRF model is presented in [4]. The learning of clique parameters from training data does not result in an energy function invariant to changes in the text properties. The binary image is estimated in a Bayesian framework. During the annealing process, for each pixel, the energy potential is calculated before and after flipping it.

A new text extraction method for natural scene camera based pictures, showing the relevance of the use of two clustering distances in the original color space instead of choosing different color spaces experimentally is present in [5]. As no transformation of color spaces needs to be applied, computational time is drastically reduced and no dynamic choice between several color spaces is useful reducing much processing time.

A binarization approach to handle a large variety of images from scanned flatbed images to images acquired by mobile phone cameras is present in [6]. The binarization is targeted at creating layers of binary images for processing by OCR engines. The layers are classified spatially and by intensity and color. First textual pixels are classified by a text operator. The text kernel is then segmented by intensity/color levels and layout analysis techniques to create regions of similar text. Adaptive binarization is applied to each region to obtain superior binary images.

A novel approach to accurately detect text regions including shopname in signboard images with complex background for mobile system applications is present in [7]. The method is based on the combination of text detection using edge profile and region segmentation using fuzzy c-means method. The edge profile and geometrical characteristics of each object contour are carefully examined to construct candidate text regions and classify the main text region. The fuzzy c-means algorithm is performed to segment and detected binarized text region.

A binarization method based on edge for video text images, especially for images with complex background or low contrast is present in [8]. The binarization method first detects the contour of the text and utilizes a local thresholding method to decide the inner side of the contour and then fills up the contour to form character that recognizable to OCR software.

An algorithm to localize artificial text in images and videos using a measure of accumulated gradients and morphological post processing to detect the text is present in [9]. The quality of the localised text is improved by robust multiple frame integration. The detection algorithm is applied to each frame of the sequence. The detected text rectangles are passed to a tracking step, which finds corresponding rectangles of the same text appearance. From the frames of an appearance a single enhanced image is generated and binarized.

A novel image binarization method that can deal with degradations such as shadows, non-uniform illumination, low contrast, large signal dependent noise, smear and strain. A pre-processing procedure based on morphological operations is first applied to suppress light/dark structures connected to image border. A novel binarization concept based on difference of gamma functions is present in [10]. Next generalised extreme value distribution is used to find proper threshold for

binarization with a significance level. The method emphasizes on region of interest and generates less noisy artifacts.

The pixel in the document image is presented as random variables in an MRF and new energy functions are introduced on these variables [11]. Each variable takes a foreground or background label and the quality of the binarization is determined by the value of the energy function. The energy function is minimized using an iterative graph cut scheme.

An objective evaluation methodology for document image binarization techniques that aims to reduce the human involvement in the ground truth construction and consecutive testing is present in [12]. A skeletonized ground truth image is produced by the user following a semi-automatic procedure. The estimated ground truth image can aid in evaluating the binarization result in terms of recall and precision as well as to further analyse the result by calculating broken and missing text, deformations and false alarms.

A sliding window method to binarize text from textual images with textured background is present in [13]. Suitable pre-processing techniques are applied first to increase the contrast of the image and blur the background noises due to textured background. Subsequently formed edge boxes are analysed to remove unwanted edges due to complex background and binarized by sliding window approach based character size uniformity check algorithm.

We propose a weighted hybrid thresholding approach.

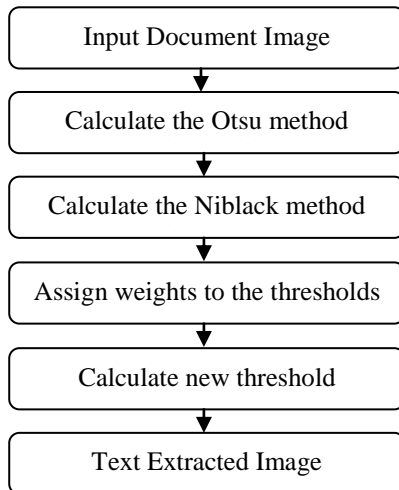


Figure 1. Flow Chart of the approach proposed

2.METHODOLOGY

A challenging problem for character recognition applications is text extraction from images, due to language characteristic, complex background and unknown text color. In this paper, an efficient approach for text extraction based on Otsu and Niblack is proposed. Text pixels having simple background is extracted very well by Otsu method and the edges are well identified by Niblack method. To achieve better text extraction, the advantages of both the techniques are combined to form a new threshold. Initially, threshold values are calculated from the state of art techniques Otsu and Niblack techniques. From the two threshold, a new threshold for text extraction is derived which is the average of the thresholds. The threshold values calculated from the Otsu and Niblack method are enhanced by different weight values before new threshold calculation. Experimental results of the proposed hybrid approach are well satisfied compared to Otsu and Niblack method.

Algorithm:

1. Read the text image
2. Calculate the threshold value Ω_1 by Otsu method
3. Calculate the threshold value Ω_2 by Niblack method
4. Assign weight ω_1 for Otsu and ω_2 for Niblack threshold value
5. Calculate the new threshold Ω by using the formula.

$$\Omega = \frac{\Omega_1\omega_1 + \Omega_2\omega_2}{2} \quad \text{where } \omega_1 + \omega_2 = 1$$

3.EXPERIMENTAL RESULTS

The performance of the proposed system is evaluated in this section. We first find the threshold Ω_1 using the Otsu method and the threshold Ω_2 using the Niblack method. Weight ω_1 for Otsu and ω_2 for Niblack threshold value is assigned. The new threshold is assigned to extract the text from images. Finally we get the text extracted image. The images are shown below

Input image



Figure 2: Original Image

Otsu Method



Figure 3. Text Binarization using Otsu method

Niblack Method



Figure 4. Text Binarization using Niblack method

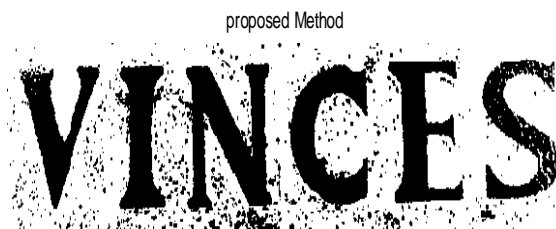


Figure 5. Text Binarization using Proposed method

4.CONCLUSION

This paper presents a novel approach on text extraction. The method uses Dual Tree complex wavelet transform, morphological dilation and logical operation. It is robust against various conditions such as shadows, degradations, non-uniform illuminations, highlights, specular reflections, different font style and size and low contrast images. The experiment result showed that the proposed method reasonably extract text regions with eliminating most non-text regions well.

5.REFERENCES

- [1] R.Minetto,N.Thome, M.Cord, J.Stolfi, F.Precioo, J.Guyumard and N.J.Leite, “Text Detection and recognition in Urban scenes”, IEEE International Conference on Computer Vision workshop, 2011.
- [2] J.Sauvola, T.Seppanen, S.Haapakoski and M.Pietikainen, “Adaptive Document Binarization”, Proceedings of the IEEE fourth International Conference on Document Analysis and recognition,IEEE, 1997, vol.1, pp 147 – 152.
- [3] Ntogas, Nikolas, Ventzas, Dimitrios, “A Binarization algorithm for historical manuscripts”, 12th WSEAS International Conference on Communications, Heraklion, Greece, July 23 – 25, 2008
- [4] Christian Wolf and David Doermann, “Binarization of low quality text using a Markov Random field model” Proceedings of IEEE 16th International Conference on Pattern Recognition, vol.3, pp. 160 – 163, 2002.
- [5] Celine Mancas, Thillou and Bernard Gosselin, “ Color Text Extraction from Camera based images-the impact of the choice of the clustering distance”, Proceedings of the IEEE 8th International Conference on Document Analysis and recognition, 2005, vol.1, pp 312 – 316.
- [6] Yaakov Navon, “Layer based binarization for Textual images”, 19th International conference on pattern recognition, Vol. 1-6, (2008), p. 2634-2638
- [7] J.Park, T.N.Dinh and G.Lee, “Binarization of text region based on fuzzy clustering and histogram distribution in signboards”, World Academy of Science, Engineering and Technology, 43, 2008
- [8] Z.Zhou, L.Li and C.L.Tan, “Edge based binarization for video text images”, 2010 International Conference on Pattern Recognition.
- [9] C.Wolf, J.M.Jolion and F.Chassaing, “Text localization, enhancement and binarization in multimedia documents”, Proceedings of IEEE 16th International Conference on Pattern Recognition, Vol.2, pp 1037 -1040.
- [10] B.Fernando, S.Karaoglu and A.Tremeau, “Extreme value theory based text binarization in documents and natural scenes”, 2010, The 3rd International Conference on machine Learning.
- [11] A.Mishra, K.Alahari and C.V.Jawahar, “An MRF model for binarization of Natural scene text”, Proceedings of IEEE International Conference on Document Analysis and Recognition
- [12] K.Ntirogiannis, B.Gatos and I.Pratikakis, “An objective evaluation methodology for document image binarization techniques”, The 8th IAPR workshop on Document analysis systems.
- [13] Chitrakala Gopalan and D.Manjula, “ Sliding window approach based text binarization from complex textual images”, International Journal on Computer Science and engineering, Vol 2, 2010, pp. 309 - 313