

A Phase Congruency Based Document Binarization

Hossein Ziaei Nafchi¹ and Hamidreza Rashidy Kanan²

¹ Department of Electrical, Computer and IT Engineering, Qazvin Branch,
Islamic Azad University, Qazvin, Iran
h.ziaei@qiau.ac.ir

² Department of Electrical Engineering, Bu-Ali Sina University, Hamedan, Iran
h.rashidykanan@basu.ac.ir

Abstract. In this paper, three new methods proposed for binarization of degraded documents and manuscripts. Phase congruency used to select regions of interest (ROI) of document's foreground. The main idea is to achieve an optimal recall measure (recall~1), while the precision value is at an acceptable level. Further processing should be performed to focus on the ROI. We also used a modified adaptive thresholding method in the next step. This method uses a global variance, a global mean and local means for thresholding. Finally, a new method called early exclusion criterion (EEC) proposed for document enhancement. The experimental results on the datasets introduced in DIBCO 2009, H-DIBCO 2010 and DIBCO 2011 shows that near optimal recall value (recall~0.99) obtained, while precision measure's value is acceptable.

Keywords: Degraded document binarization, Phase congruency, Adaptive thresholding, Early exclusion criterion.

1 Introduction

The purpose of document binarization is to convert input gray-scale documents into binary form. Usually, the latter form will be used in most document analysis systems as the first step. The performance of document binarization step highly impacts the subsequent steps such as page segmentation and optical character recognition. A number of historical and badly degraded documents can be found in libraries and archives. Usually, reading or processing these documents is not easy. For converting such injured documents, adaptive thresholding techniques are the best choices. Adaptive thresholding technique is a robust method to handle strong illumination changes. A global binarization method such as Otsu's method [1], usually fail in such an environment conditions. Global binarization methods try to find a threshold and use it for whole document image.

In this paper, three methods proposed for document processing. Phase congruency [2] is a well-known edge detector. It widely used in the machine vision literature. Phase congruency shows advantages against gradient-based edge detectors such as Sobel and Canny due to their sensitivity to variations in image illumination, blurring and magnification [2]. In this paper, phase congruency is used to select edges of

foreground. Then, a morphological grey-scale image reconstruction [3] used to fill the obtained edges. After conversion to binary form, it contains all the foreground information. This means that recall~1.

A number of adaptive thresholding methods have been proposed [4, 5, 6, 7, 8, 9, 10, 11] and [12]. Sauvola et al [4] proposed an adaptive thresholding method for image binarization. They used local mean and local variance to compute a threshold for each pixel in the input image. Shafiat el al [5] improved the speed of the Sauvola's approach by using two integral images. An adaptive thresholding approach has been proposed by Wellner [6]. In this approach, each pixel is set to 0 (black) if the pixel intensity value is smaller than 85% of average of the intensity values of some surrounding pixels in the x-axis. Bradley and Roth [7] improved this approach by including surrounding pixels in both x and y axes and speeded-up their method by utilizing integral image and achieved real-time thresholding. We improved this approach by interfering the global mean and global standard deviation.

In the face detection literature, Liu [13] proposed a method called early exclusion criterion (EEC) to exclude windows which cannot be faces at all. This criterion used as first step (early exclusion) to speed up overall face detection speed. We used a different form of this approach for document enhancement.

The rest of the paper is organized as follows: In section 2, we discussed the related works. In section 3, proposed phase congruency based ROI selection is introduced. Section 4 elaborates our proposed modified adaptive thresholding method. In section 5, the proposed early exclusion (EEC) method for document enhancement is introduced. Section 6 deals with the experimental results and Section 7 draws a conclusion.

2 Related Works

Many of the thresholding methods have been surveyed in [14, 15] and [16]. Sauvola's method uses local mean and local variances to compute a threshold $t(x,y)$ for pixel $g(x,y)$ in the grey-scale image g . Each threshold is computed by the following equation:

$$t(x, y) = m(x, y) \left[1 + k \left(\frac{\sigma(x, y)}{R} \right) - 1 \right]. \quad (1)$$

where $m(x, y)$ is the local mean, $\sigma(x, y)$ is local standard deviation, R is the maximum value of the standard deviation and k is a parameter which takes the positive values in the range of [0.2,0.5], [4]. The value of k is 128 for grey-scale images. If the contrast of surrounding pixels of $g(x, y)$ is high, then $\sigma(x, y) \cong R \cong 128$ and then $t(x, y) = m(x, y)$. If the value of $s(x, y)$ becomes low then $t(x, y)$ becomes less than $m(x, y)$, and as a result the dark side of background will be removed. This approach shows acceptable results even for the severely degraded documents. Sauvola's method is the modification of Niblack's method [8]. A problem with Sauvola's approach is its slow process time. Computing local mean and local variance for each pixel is a slow process. Shafiat et al [5] improves the sauvola's method by using two integral images. Remember that integral image widely used in the machine vision literature, after utilizing by Viola and Jones [17]. By utilizing the integral image, we can compute sum of rectangles in the constant time independent of rectangle size.

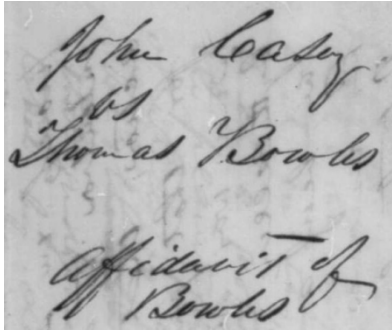
Bradley et al [7] used local mean of the surrounding pixels $S \times S$ for thresholding. $S \times S$ means that pixels of the both x and y axes used. They choose the surrounding pixels as $1/8$ of image length. A pixel is set to dark if its value is less than the product of a coefficient C and local mean. They chose the value of C as 0.85 like the Wellner [6].

Otsu's method [1] tries to choose a threshold to minimize the intraclass variance of the black and white pixels. Moghaddam et al [10] introduced AdOtsu, which is an adaptive and parameterless generalization of the Otsu's method. They used multiscale background estimation and a skeleton-based postprocessing to remove false positive sub-strokes. Moghaddam et al [11] proposed a multi-scale framework for adaptive binarization of degraded document images. This framework is based on the several binarizations on different scales and use of AdOtsu. Hedjam et al [12] proposed a spatially adaptive statistical method for image binarization. They used Sauvola's method to obtain an initial map and adapt a two-class maximum likelihood classifier to the pixels. The parameters of the class are computed locally from the grey-level distribution. Moghaddam et al [18] proposed a non-local patch means (NLPM) restoration and reconstruction method for preprocessing degraded document images. The image data is represented by a content-level descriptor based on patches. Then a modified genetic algorithm is used to correct the patched image based on the similar patches identified.

Finally, a number of hybrid methods used for document binarization. For example, Gatos et al [19] proposed a hybrid adaptive thresholding method based on combination of several methods. They also used edge information and enhancement step based on mathematical morphology operation.

3 Phase Congruency Based ROI Selection

A new procedure, based on phase congruency is proposed to perform document image binarization. Phase congruency is a robust method to detect edges and corners. Phase congruency's robustness to image variations stems from the multi-scale and multi-orientational approach to phase congruency calculation and from the fact that phase rather than magnitude information is considered for line or edge detection. We refer to reference [2] for more study about phase congruency. In this paper, phase congruency is used to detect edges of document's information. A number of parameters impact the phase congruency output. Specially, the number of $2D$ log-Gabor filters scales p and the number of orientations of $2D$ log-Gabor filters r should be set according to the application. As we see in Fig. 1 and Fig. 2 Phase congruency detects edges. As a result inner information of edges of document information will be lost. Therefore, we use grey-scale image reconstruction [3] to fill the inner parts of edges. Then a conversion to binary form must be performed. We choose a global threshold for this purpose. This global threshold can be either the mean of the filled image or even can be obtained threshold from Otsu global thresholding method. In our experiments, we used half of Otsu's global threshold. This threshold achieve near optimal recall measure while the precision value is acceptable for a ROI selector. The output of this step is the input for the proposed adaptive thresholding method. Fig. 1 and Fig. 2 shows above mentioned process.



a) A degraded document image (from DIBCO 2009 dataset)



b) The edge image obtained by using the phase congruency

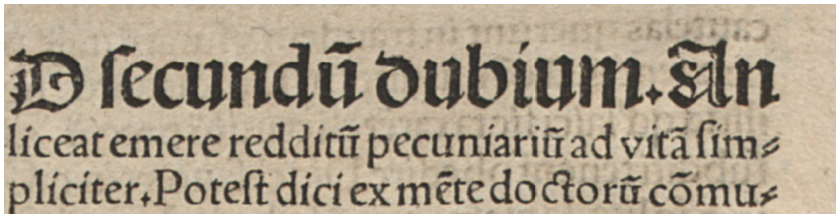


c) Filled image of (b)



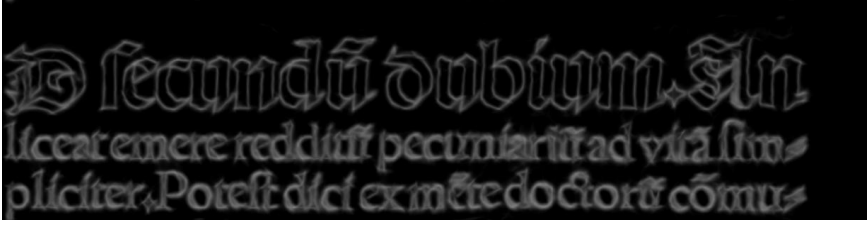
d) Binary conversion of (c).

Fig. 1. Phase congruency based ROI selection process. We used $p = 10$, $r = 10$ in our experiments. The phase congruency was able to reject a majority of the background pixels. ($Recall = 100$, $F - measure = 44.35$).

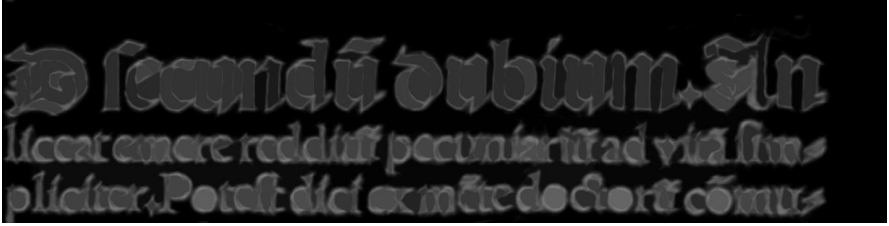


a) A degraded Input image (from DIBCO 2009 dataset)

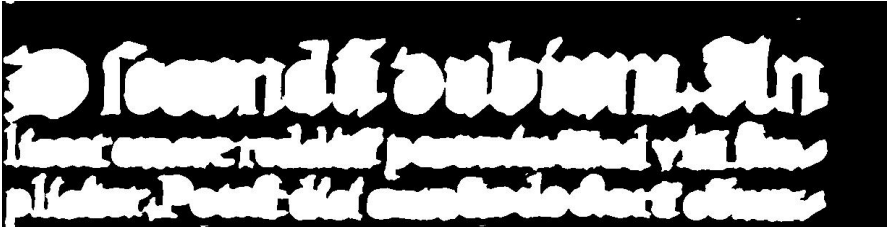
Fig. 2. Phase congruency based ROI selection process. ($Recall = 100$, $F - measure = 61.95$).



b) The edge image obtained by using the phase congruency



c) Filled image of (b)



d) Binary conversion of (c).

Fig. 2. (Continued)

4 The Proposed Adaptive Thresholding Method

While many of the adaptive thresholding methods use local mean and local variance to compute a threshold for each pixel, we use a different manner. The proposed method which is an improvement of the Bradley et al [7] approach is as follows. For each pixel, we compute the average of $S \times S$ surrounding pixels of that pixel and compare value of that pixel with product of obtained average with a coefficient. This coefficient is computed with the following equation:

$$C = 0.85 + \frac{|\mu + \sigma|}{1000} \quad (2)$$

where, μ and σ are the average and standard deviation of intensities of input image. A pixel is set to 0 (dark) if the value of that pixel is smaller than the product of S mean values and C , otherwise pixel is set to 1 (white). Wellner [6] and Bradley et al [7] choose the value of the C as 0.85 and number of S as 1/8 of image length. Furthermore, Bradley & Roth suggested that for different applications one can use a

different C . Suppose that we have an image with high intensity pixels, Bradley et al adaptive thresholding method may fail because pixel value usually becomes more than surrounding pixels and maybe set to 1 erroneously. In images with low intensity values the same scenario repeated in setting pixels with 0 (dark). The proposed method interfere the mean and standard deviation of input image in the coefficient to overcome this problem. If input image has dark values and also low variations, then proposed approach considers low intensity values as background in according to its $S \times S$ mean. This is because the coefficient C becomes less than 1. We also choose the number of S as $1/16$ of image length.

In this paper, the output of the phase congruency ROI is the input for adaptive thresholding. In the input image those pixels in the ROI which classified as background, replaced by the mean of the input image. Instead of replacing by mean value some methods replaced the background values by *NaN* value. This approach also can be used as a preprocessing step without replacing by the mean value. This approach changes the F-measure value to some extent. Adaptive thresholding method converts this input image to binary form. At the end of this step, many of unnecessary holes will be removed while approximately all the sub-strokes remains. Results can be found in section 6.

5 Early Exclusion Criterion Enhancement

An early exclusion criterion has been proposed by Liu [13] for face detection preprocessing. A completely different manner of this criterion is used in this paper for document enhancement. While this criterion used as the first step in the face detection applications, we use it as final step. The reason is its slower process time in comparison with adaptive thresholding used in section 4. The purpose is to remove those background pixels which already classified as foreground. The proposed document enhancement method is as follow.

After removing isolated pixels from previous step, a 5×5 sliding window slides across whole input image. Average intensity of the sliding window m is computed from integral image. Then m_h is computed, where m_h is the average intensity of those pixels which has higher value than m . A pixel is set to background if it satisfies two conditions: $in(i, j) \times k^* > m_h$ and $m_h \neq 0$, where $k^* > 1$. We set $k^* = 1.05$ in our experiments. We observed that some inner parts of large connected foregrounds may set to background erroneously. Therefore, we use a third condition in addition to above two mentioned conditions. The value of $bw(i, j)$ and at least one of its 3×3 mask should be unequal. This condition solves the large connected foreground problem. It's clear that only foreground pixels from previous steps are evaluated by EEC. EEC removes those pixels in which they cannot be foreground. The experimental results show that outstanding results achieved for an enhancement procedure.

6 Experimental Results

The proposed methods have been tested on the standard datasets provided in DIBCO'09 [16], H-DIBCO'10 [20] and DIBCO'11 [21]. The measures recall, precision and F-measure used in our experiments:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (5)$$

where, TP, FP, FN denote the true positive, false positive and false negative values, respectively. Table 1 provide the results of the proposed methods, Otsu [1], spatially adaptive [12], and AdOtsu methods [10]. To the best of our knowledge, this approach [10] has highest F-measure reported to date.

Table 1. Experimental results comparison between proposed methods, Otsu's method, spatially adaptive method and AdOtsu method for DIBCO'09 dataset

<i>Method</i>	<i>Measures</i>		
	Recall	Precision	F-measure
Phase congruency (1)	99.83	30.17	44.83
Phase congruency + Adaptive thresholding (2)	98.82	61.15	71.85
(1) + (2) + Early Exclusion Criterion	98.50	68.18	78.30
Otsu [1]	94.25	73.66	78.59
Spatially adaptive [12]	92.10	90.72	91.35
AdOtsu [10]	90.09	93.22	91.57

The proposed adaptive thresholding method improved the F-measure value from phase congruency by about 60% at the cost of 1% decrease of the recall value. The EEC improved F-measure by about 10% while reduction of recall value is only 0.3%. Recall and F-measure values for printed images in the DIBCO'09 dataset (P01-P05) are 98.21 and 89.30, respectively. That is 98.78 and 67.30 for handwritten images. Total running time to convert all 10 images in the DIBCO'09 dataset is about 31 seconds. The experiments performed on a Pentium 4, 3.4 GHz with 4 GB of RAM.

Although our proposed methods achieve smaller F-measure values, it should be noticed that a recall~99 achieved. An alternative method can further improve the performance by focusing on the ROI instead of dealing with all pixels in the input image. We also noticed that many of the methods evaluated in the DIBCO'09 contest [16] have higher F-measure value. We believe that a well combination of several methods can result in a near optimum F-measure. Each of these methods must have optimal or near-optimal recall value and acceptable precision value. For example, Gatos et al [19] combined several state-of-the-art methods to achieve high F-measure value. For this end, we proposed three mentioned methods.

To show the robustness of the proposed preprocessing step by using the phase congruency, results for the H-DIBCO'10 and DIBCO'11 are listed in Table 2. The experimental results of the DIBCO'10 and DIBCO'11 reports [20, 21] indicate that these datasets includes more problematic document images than DIBCO'09 dataset.

Table 2. Experimental results of preprocessing by using the proposed phase congruency based region of interest selection

<i>Dataset</i>	<i>Measures</i>		
	Recall	Precision	F-measure
H-DIBCO 2010	99.72	29.01	44.32
DIBCO 2011 (Handwritten)	99.09	32.25	48.21
DIBCO 2011 (Machine Print)	98.33	40.45	57.12

7 Conclusion

In this paper, three methods proposed to select regions of interest and binarize degraded document images. Phase congruency, image filling and Otsu's global threshold used to select regions of interest. Then, we used regions of interest information obtained from first method and a modified adaptive thresholding method to further improve the binarization result. Finally, an enhancement method called early exclusion criterion proposed and used for further document enhancement. Experimental results on the DIBCO'09, H-DIBCO'10 and DIBCO'11 datasets shows that near optimal recall value obtained, while precision value is acceptable. A subsequent method should be employed to achieve better results. In future work, we focus on the shape based document binarization based on phase congruency.

References

1. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man, and Cybernetics* 9(1), 62–66 (1979)
2. Kovese, P.: Image Features from Phase Congruency. *Videre: Journal of Computer Vision Research* 1(3) (1999)
3. Soille, P.: *Morphological Image Analysis, Principles and Applications*. Springer (2007)
4. Sauvola, J., Pietikainen, M.: Adaptive binary image binarization. *Pattern Recognition* 33(2), 225–236 (2000)
5. Shafiat, F., Keysers, D., Breuel, T.M.: Efficient Implementation of Local Adaptive Thresholding Techniques Using Integral Images. *Document Recognition and Retrieval XV* (2008)
6. Wellner, P.D.: Adaptive thresholding for the digitaldesk. Tech. Rep. EPC-110 (1993)
7. Bradley, D., Roth, G.: Adaptive thresholding using the integral image. *Journal of Graphic Tools* 12(2), 13–21 (2007)
8. Niblack, W.: *An Introduction to Image Processing*. Prentice-Hall Press (1986)

9. Gatos, B., Pratikakis, I., Perantonis, S.J.: Adaptive degraded document image binarization. *Pattern Recognition* 39(3), 317–327 (2006)
10. Moghaddam, R.F., Cheriet, M.: AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization. *Pattern Recognition* 45, 2419–2431 (2012)
11. Moghaddam, R.F., Cheriet, M.: A multi-scale framework for adaptive binarization of degraded document images. *Pattern Recognition* 43(6), 2186–2198 (2010)
12. Hedjam, R., Moghaddam, R.F., Cheriet, M.: A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images. *Pattern Recognition* 44(9), 2184–2196 (2011)
13. Liu, C.: A Bayesian discriminating features method for face detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 25(6), 725–740 (2003)
14. Trier, O., Taxt, T.: Evaluation of binarization methods for document images. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 17, 312–315 (1995)
15. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging* 13(1), 146–165 (2004)
16. Gatos, B., Ntirogiannis, K.: ICDAR 2009 document image binarization contest (DIBCO 2009). In: *International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1375–1382 (2009)
17. Viola, P., Jones, M.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
18. Moghaddam, R.F., Cheriet, M.: Beyond pixels and regions: A non-local patch means (NLPM) method for content-level restoration, enhancement, and reconstruction of degraded document images. *Pattern Recognition* 44(2), 730–743 (2011)
19. Gatos, B., Pratikakis, I., Perantonis, S.J.: Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information. In: *International Conference on Pattern Recognition*, pp. 1–4 (2008)
20. Pratikakis, I., Gatos, B., Ntirogiannis, K.: H-DIBCO 2010 – Handwritten Document Image Binarization Competition. In: *International Conference on Frontiers in Handwritten Recognition*, pp. 727–732 (2010)
21. Pratikakis, I., Gatos, B., Ntirogiannis, K.: ICDAR 2011 Document Image Binarization Contest (DIBCO 2011). In: *International Conference on Document Analysis and Recognition*, pp. 1506–1510 (2011)