THE IMAGE BINARIZATION PROBLEM REVISITED: PERSPECTIVES AND APPROACHES

Costin-Anton Boiangiu[1]
Ion Bucur[2]
Andrei Tigora[3]

ABSTRACT

Image document analysis is a complex process that involves several steps of processing. However, due to their sensitivity to errors, most of these are not applied on the original image; instead, they use a simplified black and white version of the original image, which offers a clear separation between foreground and background. Unfortunately, achieving the optimal separation is difficult to achieve, as no proposed algorithm has managed to offer a solution that is adequate for any type of input. This paper aims to present some of the more recent approaches in the field and compare their results with some of the classic algorithms.

Keywords: Binarization, Local Thresholding, Global Thresholding, Variable Window, Niblack, Otsu

1. INTRODUCTION

Over the past decades document creation and storage has slowly switched from physical to electronic support. This has greatly changed the way humans interact with their data: the search times have decreased as the documents can be easily accessed and investigates from virtually anywhere, the storage requirements no longer represent a problem and backing up information is handled automatically with little or no human intervention whatsoever.

Though it is great that newly created documents are no longer printed or typed on paper for storage, there are still millions more that were created before the advent of electronic storage. These are either rare documents such as manuscripts or early printings, or archives that store large amounts of information that would be ideal for automatic processing [14] [21].

Though content conversion has come a long way since its early days, processing old documents does pose a new set of challenges [3]. These are far from the newly printed

[1] Associate Professor PhD., Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenţei 313, Bucharest, 060042, Romania, costin.boiangiu@cs.pub.ro

[2] Associate Professor PhD., Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenţei 313, Bucharest, 060042, Romania, ion.bucur@cs.pub.ro

[3] Engineer, Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenţei 313, Bucharest, 060042, Romania, andrei.tigora@cti.pub.ro

documents that are part of everyday activities; instead, they show signs of deterioration caused by improper handling or storage, parasites attacks etc [16] [19] [21]. Furthermore, the digitization process also introduces its errors, due to the poorly calibrated scanning devices, resulting in non-uniform brightness and noise.

As documents themselves have a complicated structure, the analysis components should not have to the deal with the extra complexity previously presented [20]. The documents are instead simplified, cataloguing the pixels as either foreground or background through bitonal conversion or binarization. A lot of work has been dedicated to image binarization achieving various degrees of success, depending on the input source [13] [14] [16] [19] [21].

2. GENERAL PURPOSE (TRADITIONAL) BINARIZATION APPROACHES

Image binarization algorithms use a variety of approaches such as error dispersion in 1D/2D domain, cell-dithering based conversions, thresholds. Among them, thresholding methods have proven to be the more popular, whether they are performed globally or locally [13] [15] [17].

Figure 1. Test image “Lena” in True-Color and Grayscale form (wherever possible, for an algorithm the perceptual distances are measured in the color space, otherwise, the grayscale variant is used)

Threshold methods involve computing a threshold value that is used as classification criteria to differentiate foreground from background pixels. The threshold can be computed globally, for the entire image [15], or locally for individual or group of pixels [18]. The global approach is fast, but for most old images, treating them uniformly generates poor results, as different regions of the image display different features. Localized evaluation of the pixels is perhaps more reliable, but even so, it is sometimes difficult to predict how slight variations of the threshold value can affect the resulting image [13].

Figure 2. Black and White conversion using a global threshold at the middle of the [Black…White] range and a zoom on a significant area of the result

Figure 3. Black and White conversion using a global threshold computed using the Otsu approach and a zoom on a significant area of the result

There is another category of binary conversion algorithms that tend to process the image by allocating different maps or patterns for every image area. This category contains methods such as Halftone [7] [9] and Ordered Dithering [11]. The basic idea is that by "summation" of the Black and White values, the resulted image will tend to reproduce the same local color/grayscale density as the initial image but using only the colors present in the building maps/patterns [6] [11]. This group of algorithms usually generates irregular contours [10]. This can have a negative impact on character recognition, which relies on evaluating the geometric characteristics of the characters.

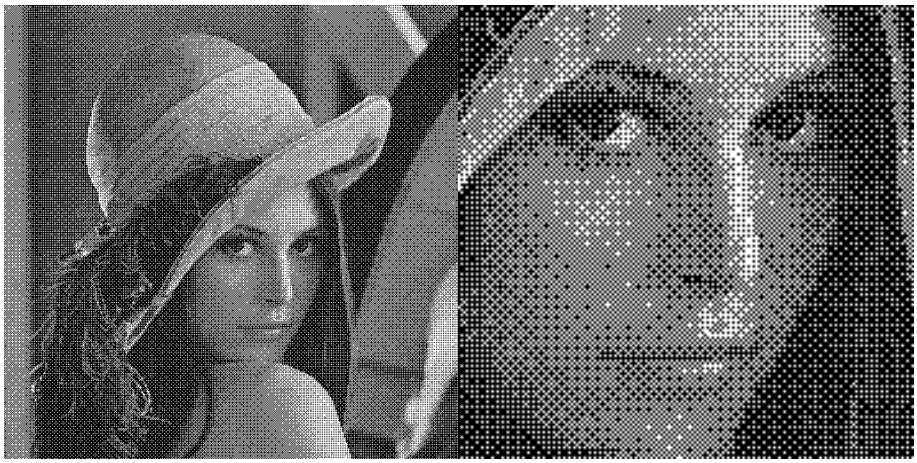


Figure 4. Black and White conversion using an Ordered Dithering algorithm and a zoom on a significant area of the result

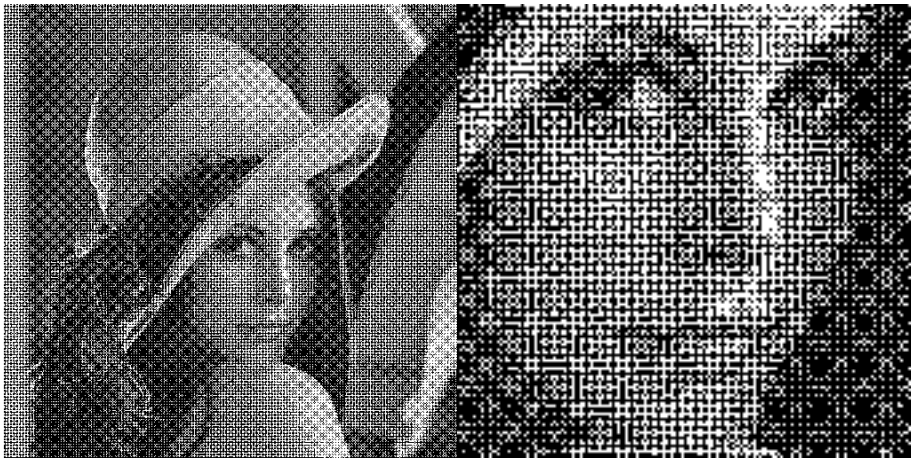


Figure 5. Black and White conversion using a Halftone Dithering algorithm and a zoom on a significant area of the result

Error dispersion techniques like Floyd-Steinberg [12] try to minimize the error representation across the page by approximating and propagating the resulted color difference at every pixel to his neighbors and rerun the approximation/propagation until all pixels are processed. The main hurdle for this approach is the image representation mechanism itself. Transitioning from dark foreground to white background is usually done smoothly, through several intermediate gray pixels. Though this is pleasant from a human perspective, it also creates regions of gray pixels that are difficult to catalogue and establish a clear contour.

Figure 6. Black and White conversion using a Floyd-Steinberg error diffusion algorithm and a zoom on a significant area of the result

As it can be easily observed from the figures above, there is no perfect conversion technique for all purposes. In fact, it cannot be since the Black and White image conversion is an operation with a significant degree of information loss.

3. SPECIFIC PURPOSE (MODERN) BINARIZATION APPROACHES

Most modern approaches concentrate on threshold based methods [13] [17]. Whereas global methods are largely considered inappropriate for most input types, as they require sharp images that clearly separate between foreground and background, the local or adaptive method still has a lot of potential.

Some of the first methods are trying to find thresholds globally (Otsu) [15] or locally (Niblack, Sauvola) [18]. Local thresholding tend to preserve better the image details and the objects contour, but also they are much more sensitive to the noise.

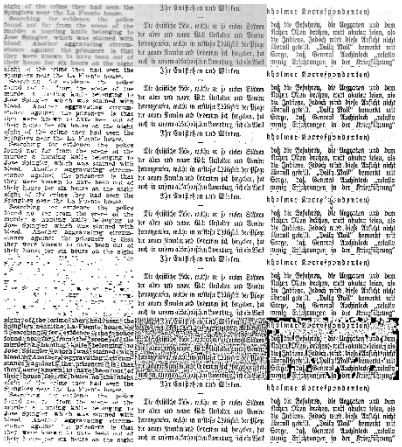
The presence of noise is a significant problem, as it disturbs the statistics for a particular image region [5]. In order to reduce the impact of noise, some authors suggest smoothing the image, even as a first preprocessing step, or as an intermediary enhancement step. This ensures that small patches of dark in the image, which usually represent noise, are less visible and therefore have a smaller impact on the result.

Other methods may choose to enhance the contrast, through various mechanisms, such as histogram stretching. The aim is to have a histogram as large as possible, which spreads the values over a larger interval. This is supposed to insure the contrast between foreground and background pixels that guarantees successful binarization.

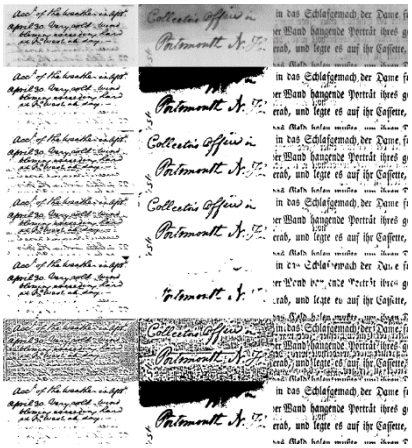
Many algorithms rely on histogram manipulation, reminiscent of Otsu's method. For document images, the typical histogram is bell shaped, having a unique global maximum.

However, this does not always happen, so the histogram may also be smoothed, usually through a triangle filter to remove non-relevant local minima and maxima.

Test set 1 was considered representative for documents containing both written and graphical material, various font sizes, uneven background and big-sized, non-uniform-spread noise specific to advanced document decay.



Test set 2 was considered representative for historical books containing Fraktur-style mixed with Antiqua-style fonts, with a lot of background noise, high-density text, blurry textual regions and presenting difficulties due to uneven exposure during document acquisition.



Test set 3 was considered representative for old documents written on thin paper with both manual handwritings and automatic printing machines

Illustrated Algorithm Order

In every test scenario:
Line 1: the original grayscale image;
Then results obtained on the original image with:
Line 2. Boiangiu Adaptive/Otsu approach;
Line 3. Boiangiu Adaptive approach;
Line 4. Lu and Tan's method;
Line 5. J. Fabrizio and B. Marcotegui's method;
Line 6. Niblack's method;
Line 7. Otsu's method.

Figure 7. Various test sets and binarization algorithms and scenarios

In [1] Boiangiu proposed a local thresholding technique but using a variable window neighborhood instead of a fixed one. There were two different proposals regarding the method used for threshold computation:

- A local thresholding based on the mean value (adaptive approach)
- A weighted local and global thresholding based on both mean value and global Otsu threshold (adaptive/Otsu approach)

The proposed methods has the advantage of being a completely parameter free approaches. A comparison against specific document binarization techniques [22] [23] [24] (in this case the parameters were tuned for their best results) shown that there are several advantages in binarization quality also. The results are presented in Figure 7.

4. CONCLUSIONS

Comparing the performances of different binarization algorithms based on scientific papers alone is difficult, as there is no predefined input set for evaluating an algorithms performance.

Even if several algorithms use the same material as input, the evaluation criteria are sometimes vague and are most often presented in comparison to other algorithms in terms of better or worse. An ideal solution would be to use a predefined set of images, coupled with their "ideal" binarized versions.

This would allow an objective evaluation of an algorithm's performances, in terms of true and false positives and negatives. On the other hand, the existence of such a set would determine researchers to highly "tune" their algorithms for those particular images, as it happens in the world of hardware benchmarking. Furthermore, the "ideal" binarization might not be that ideal after all, as it is not always clear which pixels represent foreground and which background.

While the end result is definitely important, the means of obtaining it is also relevant, and this aspect is often downplayed in many papers. Algorithms that show good results for a wide range of input images are often parameterized [22] [24], requiring human intervention for fine tuning. Setting these parameters requires some experience with the algorithm and its output, and obtaining the best out of an algorithm becomes more of an art than science. This could hardly be considered a "sin", but such an approach renders an algorithm unusable for automatic processing, at least in its original implementation. To compensate for this aspect, some authors propose dynamic mechanisms for setting the values, which can, to some extent, yield results that are almost just as good as the human fine tuned ones.

Another aspect that is rarely tackled is resource usage. Hardly any algorithm requires more memory than what is usually available for modern commodity computers. On the other side, execution time can spiral out of control. Whereas for Otsu's algorithm the image pixels are inspected once for histogram creation and once again for the actual image binarization

[15], applying local Otsu for 10 by 10 square windows results in 100 more inspections than the original approach. The processing can become even slower if the computations performed on the window are more complicated that basic pixel counting.

The time complexity is usually linear, directly proportional to the number of pixels in the image. However it can sometimes depend on the square of the image size, when variable size windows are used, and even when it does not, the proportionality constant can be really high. Still, for large scale document analysis projects, the time constraint is not a deal breaker; the processing involved is usually easy to parallelize so more hardware and efficient implementations should normally solve, at least partially, this problem.

5. ACKNOWLEDGEMENTS

The work presented in this paper was funded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the financial agreement POSDRU/89/1.5/S/62557.

REFERENCES

1. Costin-Anton Boiangiu, Alexandra Olteanu, Alexandru Stefanescu, Daniel Rosner, Nicolae Tapus, Mugurel Andreica "Local Thresholding Algorithm Based on Variable Window Size Statistics" Proceedings CSCS-18, The 18-th International Conference on Control Systems and Computer Science, May 24-27 2011, Bucharest, Romania, Volume 2, Pag. 647-652, Politehnica Press, ISSN: 2066-4451.
2. Costin-Anton Boiangiu; Olteanu, A.; Stefanescu, A. V.; Rosner, D. & Egner, A. I. (2010). Local Thresholding Image Binarization using Variable-Window Standard Deviation Response (2010), Annals of DAAAM for 2010 & Proceedings of the 21st International DAAAM Symposium, 20-23 October 2010, Zadar, Croatia, pag. 133-134, ISBN 978-3-901509-73-5, ISSN 1726-9687, Editor B. Katalinic, Published by DAAAM International, Vienna, Austria 2010.
3. Costin-Anton Boiangiu, Andrei Iulian Dvornic. "Methods of Bitonal Image Conversion for Modern and Classic Documents". WSEAS Transactions on Computers, Issue 7, Volume 7, pp. 1081 1090, July 2008, ISSN: 1109-2750.
4. Costin-Anton Boiangiu, Andrei Iulian Dvornic, Dan Cristian Cananau, „Binarization for Digitization Projects Using Hybrid Foreground-Reconstruction", Proceedings of the 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing", Cluj-Napoca, Editor Ioan Alfred Letia, August 27-29, pp.141-144, ISBN: 978-1-4244-5007-7, Library of Congress Number: 200906791, IEEE Catalog Number: CFP0909D-PRT, 2009.
5. Lawrence G. Roberts, Picture Coding Using Pseudo-Random Noise, MIT, S.M. thesis, 1961 online
6. L. Schuchman (December 1964). "Dither Signals and Their Effect on Quantization Noise" (abstract). IEEE Trans. Communications 12 (4): 162–165. doi:10.1109/TCOM.1964.1088973.
7. Ulichney, Robert A (1994). "Halftone Characterization in the Frequency Domain". Retrieved 2012-7-20.
8. 6-Bit vs. 8-Bit... PVA/MVA vs. TN+Film Are Things Changing? [1]
9. Crocker, Lee Daniel; Boulay, Paul & Morra, Mike (1991-06-20). "Digital Halftoning". Computer Lab and Reference Library. Retrieved 2007-09-10

10. Silva, Aristófanes Correia; Lucena, Paula Salgado & Figuerola, Wilfredo Blanco (2000-12-13). "Average Dithering". Image Based Artistic Dithering. Visgraf Lab. Retrieved 2007-09-10.
11. Ulichney, Robert A (1993). "The void-and-cluster method for dither array generation". Retrieved 2012-7-19.
12. R.W. Floyd, L. Steinberg, An adaptive algorithm for spatial grey scale. Proceedings of the Society of Information Display 17, 75–77 (1976).
13. Sezgin, M., and Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging (2004)
14. Halabi, Y. S., SA, Z., Hamdan, F., Haj Yousef, K.: Modeling Adaptive Degraded Document Image Binarization and Optical Character System. European Journal of Scientific Research (2009)
15. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics (1979)
16. Gatos, B., Pratikakis, I., Perantonis, S. J.: Adaptive degraded document image binarization. Pattern Recognition Journal, Elsevier Science Inc., New York, NY, USA (2006)
17. Zhang, J., Hu, J.: Image segmentation based on 2D Otsu method with histogram analysis. Proceedings of the International Conference on Computer Science and Software Engineering. IEEE Computer Society, Washington, DC, USA (2008)
18. Niblack, W.: An introduction to digital image processing. Prentice-Hall International edition (1986)
19. Gatos, B., Pratikakis, I., Perantonis, S. J.: An adaptive binarization technique for low quality historical documents. In Lecture Notes in Computer Science,of Document Anal-ysis Systems VI, Springer, Berlin (2004)
20. Kasar, T., Kumar, J., Ramakrishnan, A.G.: Font and background color independent text binarization. Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR). Bangalore, India (2007)
21. He, J., Do, Q.D.M., Downton, A.C., Kim, J.H.: A comparison of binarization methods for historical archive documents. Proceedings of the Eighth International Conference on Document Analysis and Recognition (ICDAR), IEEE Computer Society, Washington, DC, USA (2005)
22. Gatos, B., Pratikakis, I., Perantonis, S. J.: Dibco 2009: document image binarization contest. International Journal on Document Analysis and Recognition (IJDAR) (2010)
23. Gatos, B., Pratikakis, I., Perantonis, S. J.: Icdar 2009 document image binarization contest (dibco 2009). Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR), IEEE Computer Society, Washington, DC, USA (2009)
24. Fabrizio, J., Marcotegui, B., Cord, M.: Text segmentation in natural scenes using toggle-mapping. 16th IEEE International Conference on Image Processing (ICIP) (2009)
25. Bayer, Bryce (June 11-13 1973). "An optimum method for two-level rendition of continuous-tone pictures". IEEE International Conference on Communications 1: 11–15.