Different Thresholding Techniques in Image Processing: a Review

Radha Seelaboyina and Rajeev Vishwkarma

January 11, 2022

Different Thresholding Techniques in Image processing : A Review

Radha Seelaboyina
Research Scholar
radha.seelaboyina@gmail.com
Dr. Rajeev Vishwkarma
Professor
rajeev@mail.com
Dr.A.P.J. Abdul Kalam University Indore

Abstract

Document data is captured through optical scanning or digital video, resulting in a file of picture elements, or pixels, which serves as the raw input for document analysis. These pixels are samples of intensity values taken in a grid pattern throughout the document page, with intensity values ranging from OFF (0) to ON (1) for binary pictures, 0-255 for grayscale images, and 3 channels of 0-255 colour values for colour images. The initial stage in document analysis is to process this image so that it may be analysed further. Thresholding is used to convert a grayscale or colour image to a binary image, noise reduction is used to remove superfluous data. The goal of this paper is to summarize some thresholding technique for image processing.
Keywords:Thresholding,Global Thresholding,Binarizations,Adaptive Thresholding,Intensity Histogram.

1 Introduction

In this treatment of document processing, we deal with images containing text and graphics of binary information[5]. That is, these images contain a single foreground level that is the text and graphics of interest, and a single background level upon which the foreground contrasts. We will also call the foreground: objects, regions of interest, or components. The documents may also contain true gray scale (or color) information, such as in photographic figures; however, besides recognizing the presence of a gray-scale picture in a document, we leave the analysis of pictures to the more general fields of image analysis and machine vision. Though the information is binary, the data — in the form of pixels with intensity values — are not likely to have only two levels, but instead a range of intensities[4]. This may be due to nonuniform printing or nonuniform reflectance from the page, or a result of intensity transitions at the region edges that are located between foreground and background regions. The objective in binarization is to mark pixels that belong to true foreground regions with a single intensity (ON) and background regions with a different intensity (OFF). Figure 1 illustrates the results of binarizing a document image at different threshold values. The ON-values are shown in black in our figures, and the OFF-values are white[2].

For documents with a good contrast of components against a uniform background, binary scanners are available that combine digitization with thresholding to yield binary data. However, for the many documents that have a wide range of background and object intensities, this fixed threshold level often does not yield images with clear separation between the foreground components and background. For instance, when a document is printed on differently colored paper or when the foreground components are faded due to photocopying, or when different scanners have different light levels, the best threshold value will also be different. For these cases, there are two alternatives[3]. One is to empirically determine the best binarization setting on the scanner (most binary scanners provide this adjustment), and to do this each time an image is poorly binarized. The other alternative is to start with gray-scale images (having a range of intensities, usually from 0 to 255) from the digitization stage, then use methods for automatic threshold determination to better perform binarization. While the latter alternative requires more input data and processing, the advantage is that a good threshold level can be found automatically, ensuring consistently good images, and precluding the need for time- consuming manual adjustment and repeated digitization.

2 Thresholding

The following discussion presumes initial digitization to gray-scale images. If the pixel values of the components and those of the background are fairly consistent in their respective values over the entire image, then a single threshold value can be found for the image[7].This use of a single threshold for all image pixels is called global thresholding. Processing methods are described below that automatically determine the best global threshold value for differentImages. For many documents, however, a single global threshold value cannot be used even for a Single image due to non-uniformities within foreground and background regions[9]. For example, for a document containing white background areas as well as highlighted areas of a different background color, the best thresholds will change by area. For this type of image, different threshold values are required for different local areas; this is adaptive thresholding.

2.1 Global Thresholding

The most straightforward way to automatically select a global threshold is by use of a histogram of the pixel intensities in the image. The intensity histogram plots the number of pixels with values at each intensity level. See Figure1 for a histogram of a document image. For an image with well-differentiated foreground and background intensities, the histogram will have two distinct peaks[10]. The valley between these peaks can be found as the minimum between two maxima and the intensity value there is chosen as the threshold that best separates the two peaks.

2.1.1 Drawbacks

There are a number of drawbacks to global threshold selection based on the shape of the intensity distribution. The first is that images do not always contain well-differentiated foreground and background intensities due to poor contrast and noise. A second is that, especially for an image of sparse foreground components, such as for most graphics images, the peak representing the foreground will be much smaller than the peak of the background intensities. This often makes it difficult to find the valley between the two peaks. In addition, reliable peak and valley detection are separate problems unto themselves.

2.1.2 Solution

One way to improve this approach is to compile a histogram of pixel intensities that are weighted by the inverse of their edge strength values. Region pixels with low edge values will be weighted more highly than boundary and noise pixels with higher edge values, thus sharpening the histogram peaks due to these regions and facilitating threshold detection between them[12].
Second an analogous technique is to highly weight intensities of pixels with high edge values, then choose the threshold at the peak of this histogram, corresponding to the transition between regions. This requires peak detection of a single maximum, and this is often easier than valley detection between two peaks. This approach also reduces the problem of large size discrepancy between foreground and background region peaks because edge pixels are accumulated on the histogram instead of region pixels; the difference between a small and large size area is a linear quantity for edges versus a much larger squared quantity for regions.
A third method uses a Laplacian weighting. The Laplacian is the second derivative operator, which highly weights transitions from regions into edges (the first derivative highly weights edges). This will highly weight the border pixels of both foreground regions and their surrounding backgrounds, and because of this the histogram will have two peaks of similar area. Though these histogram shape techniques offer the advantage that peak and valley detection are intuitive, still peak detection is susceptible to error due to noise and poorly separated regions. Furthermore, when the foreground or background region consists of many narrow regions, such as for text, edge and Laplacian measurement may be poor due to very abrupt transitions (narrow edges) between foreground and background.

2.2 Multi Thresholding

A number of techniques determine classes by formal pattern recognition techniques that optimize some measure of separation. One approach is minimum error thresholding. Here, the foreground and background intensity distributions are modeled as normal (Gaussian or bell-shaped) probability density functions. For each intensity value (from 0 to 255, or a smaller range if the threshold is known to be limited to it), the means and variances are calculated for the foreground and background classes, and the threshold is chosen such that the misclassification error between the two classes is minimized[13]. This latter method is classified as a parametric technique because of the assumption that the gray-scale distribution can be modeled as a probability density function. This is a popular method for many computer vision applications, but some experiments indicate that documents do not adhere well to this model, and thus results with this method are poorer than non-parametric approaches . One non-parametric approach is Otsu's method. Calculations are first made of the ratio of between-class variance to within-class variance for each potential threshold value. The classes here are the foreground and background pixels and the purpose is to find the threshold that maximizes the variance of intensities between the two classes, and minimizes them within each class. This ratio is calculated for all potential threshold levels and the level at which the ratio is maximum is the chosen threshold. A similar approach to Otsu's employs an information theory measure, entropy, which is a measure of the information in the image expressed as the average number of bits required to represent the information. Here, the entropy for the two classes is calculated for each potential threshold, and the threshold where the sum of the two entropies is largest is chosen as the best threshold. Another thresholding approach is by moment preservation. This is less popular than the methods above, however, we have found it to be more effective in binarizing document images containing text. For this method, a threshold is chosen that best preserves moment statistics in the resulting binary image as compared with the initial gray-scale image. These moments are calculated from the intensity histogram — the first four moments are required for binarization. Many images have more than just two levels. For instance, magazines often employ boxes to highlight text where the background of the box has a different color than the white background of the page. In this case, the image has three levels: background, foreground text, and background of highlight box. To properly threshold an image of this type, multi-thresholding must be performed.

There are many fewer multi-thresholding methods than binarization methods. Most require that the number of levels is known. For the cases where the number of levels is not known beforehand, one method will determine the number of levels automatically and perform appropriate thresholding. This added level of flexibility may sometimes lead to unexpected results. For instance, a magazine cover with three intensity levels may be thresholded to four levels instead due to the presence of an address label that is thresholded at a separate level.

2.3 Adaptive Thresholding

A common way to perform adaptive thresholding is by analyzing gray-level intensities within local windows across the image to determine local thresholds . White and Rohrer describe an adaptive thresholding algorithm for separating characters from background. The threshold is continuously changed through the image by estimating the background level as a two-dimensional running-average of local pixel values taken for all pixels in the image[12]. Mitchell and Gillies describe a similar thresholding method where background white level normalization is first done by estimating the white level and subtracting this level from the raw image. Then, segmentation of characters is accomplished by applying a range of thresholds and selecting the resulting image with the least noise content. Noise content is measured as the sum of areas occupied by components that are smaller and thinner than empirically determined parameters. Looking back at the results of binarization for different thresholds in it can be seen that the best threshold selection yields the least visible noise[1]. The main problem with any adaptive binarization technique is the choice of window size. The chosen window size should be large enough to guarantee that a large enough number of background pixels are included to obtain a good estimate of average value, but not so large as to average over nonuniform background intensities. However, often the features in the image vary in size such that there are problems with fixed window size. To remedy this, domain dependent information can be used to check that the results of binarization give the expected features (a large blob of an ON-valued region is not expected in a page of smaller symbols, for instance). If the result is unexpected, then the window size can be modified and binarization applied again.

3 Choosing a Thresholding Method

Whether global or adaptive thresholding methods are used for binarization, one can never expect perfect results. Depending on the quality of the original, there may be gaps in lines, ragged edges on region boundaries, and extraneous pixel regions of ON and OFF values[11]. This fact that processing results will not be perfect is generally true with other document processing methods, and indeed image processing in general. The recommended procedure is to process as well as possible at each step of processing, but to defer decisions that don't have to be made until later steps to avoid making irreparable errors. In later steps there is more information as a result of processing to that point, and this provides greater context and higher level descriptions to aid in making correct decisions, and ultimately recognition. Deferment, when possible, is a principle appropriate for all stages of document analysis .A number of different thresholding methods have been presented in this section. It is the case that no single method is best for all image types and applications. For simpler problems where the image characteristics do not vary much within the image or across different images, then the simpler methods will suffice[10]. For more difficult problems of noise or varying image characteristics, more complex (and time-consuming) methods will usually be required. Commercial products vary in their thresholding capabilities. Today's scanners usually perform binarization with respect to a fixed threshold. More sophisticated document systems provide manual or automatic histogram-based techniques for global thresholding[8]. The most common use of adaptive thresholding is in special purpose systems used by banks to image checks. The best way to choose a method at this time is first by narrowing the choices by the method descriptions, then just experimenting with the different methods and examining their results. Because there is no "best" thresholding method, there is still room for future research here. One problem that requires more work is to identify thresholding methods or approaches that best work on documents with particular characteristics[6]. Many of the methods described above were not formulated in particular for documents, and their performance on them is not well known. Documents have characteristics, such as very thin lines that will favor one method above another.

4 Conclusion

In this paper we discussed different techniques of thresholding as well as how best to quantify the results of thresholding. For text, one way is to perform optical character recognition on the binarized results and measure the recognition rate for different thresholds.We also discussed problem that requires further work is that of multi-thresholding. While multi-thresholding capabilities have been claimed for some of the methods discussed above, not much dedicated work has been focused on this problem. For other reviews and more complete comparisons of thresholding methods on global and multi-thresholding techniques, and on adaptive techniques. We suggest just manually setting a threshold when the documents are similar and testing is performed beforehand. For automatic, global threshold determination, we have found that the moment-preserving method works well on documents. For adaptive thresholding, the method of is a good choice. This paper also gives background and comparison on these adaptive methods. For multi-thresholding, the method is appropriate if the number of thresholds is known, and the method if not.

References

[1] Bilal Bataineh, Siti Norul Huda Sheikh Abdullah, and Khairuddin Omar. An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. *Pattern Recognition Letters*, 32(14):1805–1813, 2011.

[2] D Gnanadurai and V Sadasivam. An efficient adaptive thresholding technique for wavelet based image denoising. *International Journal of Electronics and Communication Engineering*, 2(8):1703–1708, 2008.

[3] Maarten Jansen and Adhemar Bultheel. Empirical bayes approach to improve wavelet thresholding for image noise reduction. *Journal of the American Statistical Association*, 96(454):629–639, 2001.

[4] Lakhwinder Kaur, Savita Gupta, and RC Chauhan. Image denoising using wavelet thresholding. In *ICVGIP*, volume 2, pages 16–18, 2002.

[5] Nirpjeet Kaur and Rajpreet Kaur. A review on various methods of image thresholding. *International Journal on Computer Science and Engineering*, 3(10):3441, 2011.

[6] N Nacereddine, L Hamami, M Tridi, and N Oucief. Non-parametric histogram-based thresholding methods for weld defect detection in radiography. In *Transactions on Enformatika, Systems Sciences and Engineering*. Citeseer, 2005.

[7] Yasuo Nakagawa and Azriel Rosenfeld. Some experiments on variable thresholding. *Pattern recognition*, 11(3):191–204, 1979.

[8] N Ramesh, J-H Yoo, and IK Sethi. Thresholding based on histogram approximation. *IEE Proceedings-Vision, Image and Signal Processing*, 142(5):271–279, 1995.

[9] Payel Roy, Saurab Dutta, Nilanjan Dey, Goutami Dey, Sayan Chakraborty, and Ruben Ray. Adaptive thresholding: A comparative study. In *2014 International conference on control, Instrumentation, communication and Computational Technologies (ICCICCT)*, pages 1182–1186. IEEE, 2014.

[10] O Imocha Singh, Tejmani Sinam, O James, and T Romen Singh. Local contrast and mean thresholding in image binarization. *International Journal of Computer Applications*, 51(6), 2012.

[11] Bolan Su, Shijian Lu, and Chew Lim Tan. Combination of document image binarization techniques. In *2011 International Conference on Document Analysis and Recognition*, pages 22–26. IEEE, 2011.

[12] Pierre D Wellner. Adaptive thresholding for the digitaldesk. *Xerox, EPC1993-110*, pages 1–19, 1993.

[13] Shi Zhong and Vladimir Cherkassky. Image denoising using wavelet thresholding and model selection. In *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, volume 3, pages 262–265. IEEE, 2000.