Segmentation with image thresholding algorithms

Sylwia Sikorska

Cracow University of Technology, Institute of Applied Informatics,
e-mail: sysi@poczta.onet.pl.

The first step in image analysis and pattern recognition is image segmentation and it is one of the most difficult tasks in image processing. It determines the quality of the final result of analysis because it is very important and critical component. There are hundreds of segmentation techniques in literature. There is no single method which can be considered good for all sorts of images and conditions. In many applications of image processing, the gray levels of pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. Thresholding then becomes a simple but effective tool to separate objects from the background. To improve the segmentation results, a strategy consists in combining algorithms in order to obtain a robust segmentation by exploiting the advantages of one method to reduce the drawbacks of the second one. This paper provides a summary of approaches to image segmentation by thresholding available at the present and describes the properties of different kinds of methods and problems encountered. There will be also presented some advanced algorithms with their practical application.

Keywords and phrases: image analysis, segmentation, thresholding algorithm.

Introduction

The first step in image analysis and pattern recognition is image segmentation and it is one of the most difficult tasks in image processing. It determines the quality of the final result of analysis because it is very important and critical component of image analysis. Image segmentation is a process of dividing an image into different regions such that each region is homogeneous while not the union of any two adjacent regions.

A formal definition of image segmentation is as follows [1]:

If $P(\cdot)$ is a *homogeneity predicate* defined on groups of connected pixels, then segmentation is a partition of the set K into connected subsets or regions $\{S_1, S_2, S_3\}$ such that

$$\bigcup_{i=1}^{n} S_i = K \quad \text{with } S_i \cap S_j = \Phi \;\; (i \neq j) \tag{1}$$

The uniformity predicate $P(S_i)$ is true for all regions S_i, and $P(S_i \cup S_j)$ is false, when $i \neq j$ and sets S_i and S_j are neighbors.

According to the following [2]: *the image segmentation problem is basically one of psychophysical perception, and therefore not susceptible to a purely analytical solution.* Probably that is way, literally, there are hundreds of segmentation techniques in literature. There is no single method which can be considered good for all sorts of images and conditions. In many applications of image processing, the gray levels of pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. Thresholding then becomes a simple but effective tool to separate objects from the background. To improve the segmentation results, a strategy consists in combining algorithms in order to obtain a robust segmentation by exploiting the advantages of one method to reduce the drawbacks of the second one. This paper provides a summary of approaches to image segmentation by thresholding available at the present and describes the properties of different kinds of methods and problems encountered. There will be also presented some advanced algorithms with their practical application.

Definition of thresholding

Ref. [3] surveyed segmentation algorithms based on thresholding and attempted to evaluate the performance of some thresholding techniques using uniformity and shape measures. It categorized global thresholding techniques into two classes: point-dependent techniques (gray level histogram based) and region-dependent techniques (modified histogram or co-occurrence based). Discussion on probabilistic relaxation and several methods of multi-thresholding techniques was also given.

This technique is based upon a simple concept. A parameter θ called the brightness threshold is chosen and applied to the image $a[m,n]$ as follows:

if $a[m,n] \geq \theta$ $\quad a[m,n] = object = 1$
else $\quad a[m,n] = background = 0$

This version of the algorithm assumes that there are light objects on a dark background. For dark objects on a light background should be used:

if $a[m,n] < \theta$ $\quad a[m,n] = object = 1$
else $\quad a[m,n] = background = 0$

The output is the label "object" or "background" which, due to its dichotomous nature, can be represented as a Boolean variable "1" or "0". In principle, the test condition could be based upon some other property than simple brightness.

The output of the thresholding operation is a binary image whose one state will indicate the foreground objects, that is, printed text, a legend, a target, defective part of a material, etc., while the complementary state will correspond to the background. Depending on the application, the foreground can be represented by gray-level 0, that is, black as for text, and the background by the highest luminance for document paper, that is 255 in 8-bit images, or conversely the foreground by white and the background by black. Various factors, such as nonstationary and correlated noise, ambient illumination, busyness of gray levels within the object and its background, inadequate contrast, and object size not commensurate with the scene, complicate the thresholding operation. Finally, the lack of objective measures to assess the performance of various thresholding algorithms, and the difficulty of extensive testing in a task-oriented environment, have been other major handicaps.

The central question in thresholding is how to choose the threshold θ? While there is no universal procedure for threshold selection that is guaranteed to work on all images, there are a variety of options.

Groups of thresholding methods

Thresholding methods are categorized in six groups according to the information they are exploiting. These categories are [4]:

- histogram shape-based methods, where, for example, the peaks, valleys and curvatures of the smoothed histogram are analysed;
- clustering-based methods, where the gray-level samples are clustered in two parts as background and foreground (object), or alternately are modeled as a mixture of two Gaussians;
- entropy-based methods result in algorithms that use the entropy of the foreground and background regions, the cross-entropy between the original and binarized image, etc.;
- object attribute-based methods search a measure of similarity between the gray-level and the binarized images, such as fuzzy shape similarity, edge coincidence, etc.;
- the spatial methods use higher-order probability distribution and/or correlation between pixels;
- local methods adapt the threshold value on each pixel to the local image characteristics.

The histogram and the probability mass function (PMF) of the image are indicated, respectively, by $h(g)$ and by $p(g)$, $g = 0 \ldots G$, where G is the maximum luminance value in the image, typically 255 if 8-bit quantization is assumed. If the gray value range is not explicitly indicated as $[g\min, g\max]$, it will be assumed to extend from 0 to G. The cumulative probability function is defined as:

$$P(g) = \sum_{i=0}^{g} p(i) \tag{2}$$

It is assumed that the PMF is estimated from the histogram of the image by normalizing it to the total number of samples. In the context of document processing, the foreground becomes the set of pixels with luminance values less than T, while the background pixels have luminance value above this threshold. In NDT images, the foreground area may consists of darker (more absorbent, denser, etc.) regions or conversely of shinier regions, for example, hotter, more reflective, less dense, etc., regions. In the latter contexts, where the object appears brighter than the background, obviously the set of pixels with luminance greater than T will be defined as the foreground.

The foreground (object) and background PMFs are expressed as $p_f(g)$, $0 \leq g \leq T$, and $p_b(g)$, $T+1 \leq g \leq G$, respectively, where T is the threshold value. The foreground and background area probabilities are calculated as:

$$P_f(T) = P_f = \sum_{g=0}^{T} p(g) \tag{3}$$

$$P_b(T) = P_b = \sum_{g=T+1}^{G} p(g) \tag{4}$$

The Shannon entropy, parametrically dependent on the threshold value T for the foreground and background, is formulated respectively:

$$H_f(T) = -\sum_{g=0}^{T} p_f(g)\log p_f(g) \qquad (5)$$

$$H_b(T) = -\sum_{g=T+1}^{G} p_b(g)\log p_b(g) \qquad (6)$$

The sum of these two is expressed as $H(T) = H_f(T) + H_b(T)$. When the entropy is calculated over the input image distribution $p(g)$ (and not over the class distributions), then obviously it does not depend on the threshold T, and hence is expressed simply as H. For various other definitions of the entropy in the context of thresholding, with some abuse of notation, we use the same symbols of $H_f(T)$ and $H_b(T)$.

The fuzzy measures attributed to the background and foreground events, that is, the degree to which the gray level, g, belongs to the background and object, respectively, and are symbolized by $m_f(g)$ and $m_b(g)$. The mean and variance of the foreground and background as functions of the thresholding level T can be similarly denoted as:

$$m_f(T) = \sum_{g=0}^{T} gp(g) \quad \sigma_f^2(T) = \sum_{g=0}^{T}[g - m_f(T)]^2 p(g) \qquad (7)$$

$$m_b(T) = \sum_{g=T+1}^{G} gp(g) \quad \sigma_b^2(T) = \sum_{g=T+1}^{G}[g - m_b(T)]^2 p(g) \qquad (8)$$

This is referring to a specific thresholding method, which was programmed in the simulation analysis. For example, Shape–Sezan and Cluster–Otsu, refer, respectively, to the shape-based thresholding method introduced in a paper by Sezan and to the clustering-based thresholding method first proposed by Otsu.

Thresholding algorithms application

In many applications of image processing, the gray levels of pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. Thresholding then becomes a simple but effective tool to separate objects from the background. Examples of thresholding applications are document image analysis, where the goal is to extract printed characters, logos, graphical content, or musical scores: map processing, where lines, legends, and characters are to be found [5] scene processing, where a target is to be detected and quality inspection of materials [6, 7] where defective parts must be delineated. Other applications can be listed as follows: cell images [8] and knowledge representation [9] segmentation of various image modalities for nondestructive testing applications, such as ultrasonic images in [10], eddy current images [11], thermal images [12], x-ray computed tomography [13], endoscopic images [14], laser scanning confocal microscopy [13], extraction of edge field [15], etc.

Fixed threshold

One alternative is to use a threshold that is chosen independently of the image data. If it is known that one is dealing with very high-contrast images where the objects are very dark and the background is homogeneous and very light, then a constant threshold of 128 on a scale of 0 to 255 might be sufficiently accurate.

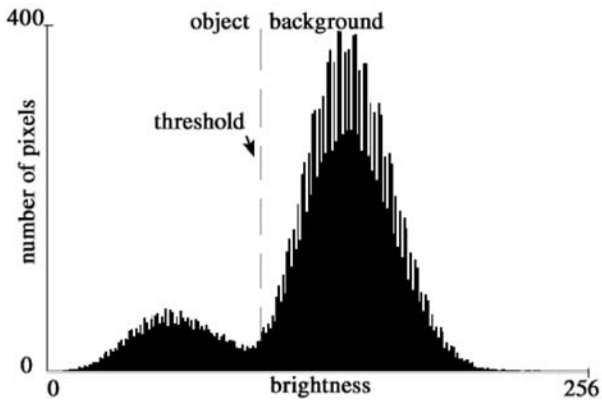
Histogram-derived thresholds

In most cases the threshold is chosen from the brightness histogram of the region or image that we wish to segment. An image prepared to segmentation by author and its associated brightness histogram are shown in Fig. 1.

A variety of techniques have been devised to automatically choose a threshold starting from the gray-value histogram. The most common one is presented below. This translates into a zero-phase smoothing algorithm given below where typical values for W are 3 or 5 [16]:

$$h_{smooth}[b] = \frac{1}{W}\sum_{w=-(W-1)/2}^{(W-1)/2} h_{raw}[b-w] \qquad (9)$$

a)



b)

Fig. 1. a) Image to be thresholded, b) brightness histogram of the image: pixels below the threshold are labeled as object pixels; those above the threshold are labeled as background pixels.

Triangle algorithm

A line is constructed between the maximum of the histogram at brightness b_{max} and the lowest value b_{min} = (p = 0)% in the image. The distance d between the line and the histogram is computed for all values of b from $b = b_{min}$ to $b = b_{max}$. The brightness value b_o where the distance between $f[b_o]$ and the line is maximal is the threshold value, that is, $\theta = b_o$. This technique is particularly effective when the object pixels produce a weak peak in the histogram. This technique due to Zack [17] is illustrated in Fig. 2.

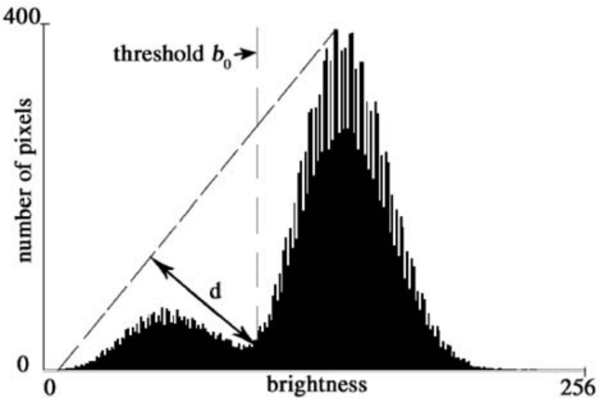


Fig. 2. The triangle algorithm basing on finding the value of b that gives the maximum distance d.

Thresholding does not have to be applied to entire images but can be used on a region by region basis. Chow and Kaneko [18] developed a variation in which the M x N image is divided into non-overlapping regions. In each region a threshold is calculated and the resulting threshold values are interpolated to form a thresholding surface for the entire image. The regions should be of reasonable size so that there are a sufficient number of pixels in each region to make an estimate of the histogram and the threshold. The utility of this procedure depends on the application at hand.

Conclusion

There are hundreds of segmentation techniques in literature. There is no single method which can be considered good for all sorts of images and conditions. In many applications of image processing, the gray levels of pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. Thresholding then becomes a simple but effective tool to separate objects from the background. In this study there was demonstrated taxonomy of thresholding algorithms based on the type of information used, and we assess their performance comparatively using a set of objective segmentation quality metrics. There was distinguished six categories, namely, thresholding algorithms based on the exploitation: histogram shape information, measurement space clustering, histogram entropy information, image attribute information, spatial information, and local characteristics. There was also presented some advanced algorithms with their practical application.

References

[1] Pal, S.K., et al. "A review on image segmentation techniques". *Pattern Recognition* 29 (1993): 1277–1294.
[2] Fu, K.S., and J.K. Mui. "A survey on image segmentation". *Pattern Recognition* 13 (1981): 3–16.
[3] Sahoo P.K., et al. "A survey of thresholding techniques". *Computer Vision Graphics Image Process* 41 (1988): 233–260.
[4] Sezgin, M., and rB. Sanku. "Survey over image thresholding techniques and quantitative performance evaluation". *Journal of Electronic Imaging* 13(1), 2004: 146–165.
[5] Trier, O.D., and A.K. Jain. "Goal-directed evaluation of binarization methods". *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (1995): 1191–1201.
[6] Sezgin, M., and R. Tasaltin. "A new dichotomization technique to multilevel thresholding devoted to inspection applications". *Pattern Recogn. Lett.* 21 (2000): 151–161.
[7] Sezgin, M., and B. Sankur. "Comparison of thresholding methods for non-destructive testing applications". IEEE ICIP'2001. *Intl. Conf. Image Process*. 2001: 764–767.
[8] Russ, J.C. "Automatic discrimination of features in grayscale images". *J. Microsc*. 148(3), 1987: 263–277.
[9] Bock, P., et al. "Gray-scale ALIAS". *IEEE Trans. Knowl. Data Eng*. 4 (1992): 109–122.
[10] Wu, L.U., M.A. Songde, and L.U. Hanqing. "An effective entropic thresholding for ultrasonic imaging". ICPR'98: *Intl. Conf. Patt. Recog*. 1998: 1522–1524.
[11] Moysan, J., G. Corneloup, and T. Sollier. "Adapting an ultrasonic image threshold method to eddy current images and defining a validation domain of the thresholding method". *NDT & E Intl*. 32 (1999): 79–84.
[12] Chang, J.S., et al. "New automatic multi-level thresholding technique for segmentation of thermal images". *Image Vis. Comput*. 15 (1997): 23–34.
[13] Oh, W., and B. Lindquist. "Image thresholding by indicator kriging". *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-21, 1999: 590–602.
[14] Srikanthan, T., and K.V. Asari. "Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images". *Med. Biol. Eng. Comput*. 39(1), 2001: 8–14.
[15] Venkatesh, S., and P.L. Rosin. "Dynamic threshold determination by local and global edge evaluation". CVGIP: *Graph. Models Image Process*. 57 (1995): 146–160.
[16] Young, I.T., J.J. Gerbrands, and L.J. van Vliet. *Fundamentals of Image Processing, An interactive image processing course*. Delft University of Technology, 2003.
[17] Zack, G.W., W.E. Rogers, and S.A. Latt. "Automatic Measurement of Sister Chromatid Exchange Frequency". *J Histochem Cytochem*. 25(7), 1977: 741–753.
[18] Chow, C.K., and T. Kaneko. "Automatic boundary detection of the left ventricle from cineangiograms". *Computers and Biomedical Research* 5 (1972): 388–410.