Schedae Informaticae Vol. 24 (2015): 21–29
doi: 10.4467/20838476SI.15.002.3024

Cross-Entropy Based Image Thresholding

Mateusz Malik, Przemysław Spurek[1], Jacek Tabor[2]
Faculty of Mathematics and Computer Science
Jagiellonian University, ul. Łojasiewicza 6, 30-348 Kraków, Poland
e-mail: *mateusz.malik@uj.edu.pl*, *przemyslaw.spurek@ii.uj.edu.pl*, *jacek.tabor@ii.uj.edu.pl*

Abstract. This paper presents a novel global thresholding algorithm for the binarization of documents and gray-scale images using Cross-Entropy Clustering. In the first step, a gray-level histogram is constructed, and the Gaussian densities are fitted. The thresholds are then determined as the cross-points of the Gaussian densities. This approach automatically detects the number of components (the upper limit of Gaussian densities is required).

Keywords: Cross-Entropy Clustering, thresholding, binarization, Otsu.

1. Introduction

Segmentation of images into homogeneous regions is an important part of the ongoing research in computer vision. The process of binarization is a common step during the various methods of image representation techniques [1] and plays a very important role in a large variety of tasks in pattern recognition, computer vision, image and video retrieval [2].

Thresholding techniques can be categorized into two classes: global and local. The global algorithms use a single threshold, while the local binarization algorithms compute a separate threshold for each pixel based on its neighborhood.

In this paper a new global thresholding algorithm will be presented. It is based on the Cross-Entropy Clustering (CEC) algorithm [3] which uses the classical Shannon Entropy Theory [4] and the Minimum Description Length Principle [5].

[1] The work of this author was supported by the Polish National Centre of Science Grant No. 2013/09/N/ST6/01178.
[2] The work of this author was supported by the Polish National Centre of Science Grant No. 2014/13/B/ST6/01792.

There are several global thresholding methods widely used in practice. The most successful is the Otsu method [6, 7] which takes the intensity of pixels in gray-level pictures into account, and divides them into two groups by maximizing the between-class variance. The Otsu method has the tendency to divide the data into two groups of similar sum of squares within the cluster, and therefore it does not cope well with the case of prevailing number of elements from the background compared to the foreground. In such a case the Otsu threshold has the tendency to place the barrier too far into the foreground (since the background is usually more concentrated and consists of more points). Consequently, after Otsu thresholding, some important details are lost, which can be of crucial importance in further processing.

Other traditional approaches use: valleys of the histogram [8], median of the histogram [9], entropy functions [10, 11], iterative methods [12] or same first three moments [13] for the choice of thresholding borders. They have similar limitations – all of them consider a fixed threshold value according to the gray-level histogram and cannot process images whose histograms are nearly unimodal, especially when the target region is much smaller and low-contrasted respective to the background area.

An effective and adaptive approach to background subtraction is to construct a statistical model which represents the probabilistic distribution of the pixel's color intensity. Kittler and Illingworth [14] present an algorithm based on fitting a mixture of Gaussian distributions. It therefore transforms the binarization problem into a minimum-error Gaussian density fitting problem. Similarly, in [12] the iterative thresholding based on two-class Gaussian mixture models is presented. It adapted a single Gaussian to represent the background model [15]. A more advanced application of GMM in image thresholding is presented in [16] where the authors construct an algorithm which automatically detects the number of components.

In this paper a thresholding method which uses Cross-Entropy Clustering instead of GMM is presented. The idea of using the CEC algorithm for thresholding problems was mentioned in [3]. This approach has several advantages over the classical GMM methods. In the case of two class binarization all possible thresholds can be verified, like in the Otsu algorithm, and consequently the method does not depend on initial conditions. For multiclass thresholding, only the upper bound of the number of components is required. The algorithm automatically reduces unnecessary clusters.

This paper is arranged as follows. In the next section the basic idea of CEC, and the possible adaptation of the algorithm to image thresholding, are presented. In the third section, the application and comparison of different methods is shown.

2. Cross-Entropy Clustering

In this section, CEC will be described by comparison to the well-known Expectation Maximization method which is used in the GMM approach. Let it be recalled that in general EM aims at finding $p_1, \ldots, p_k \geq 0$, $\sum_{i=1}^{k} p_i = 1$ and $f_1, \ldots, f_k$ Gaussian densities (where k is given beforehand) such that the convex combination

$$f := p_1 f_1 + \ldots p_k f_k$$

optimally approximates the scatter of the data $X = \{x_1, \ldots, x_n\}$ with respect to the MLE cost function

$$\mathrm{MLE}(f, X) := -\sum_{l=1}^{n} \ln(p_1 f_1(x_l) + \ldots + p_n f_n(x_l)). \tag{1}$$

The optimization in EM consists of the Expectation and Maximization steps. While the Expectation step is relatively simple, Maximization usually needs complicated numerical optimization.

On the other hand, the goal of CEC is to minimize another cost function, which is a small modification of the one given in (1) by substituting the sum with a maximum:

$$\mathrm{CEC}(f, X) := -\sum_{l=1}^{n} \ln(\max(p_1 f_1(x_l), \ldots, p_n f_n(x_l))). \tag{2}$$

However, instead of focusing on the density estimation as its first aim, CEC concerns the clustering.

To explain CEC applied to image histogram thresholding, let a set of possible gray-scale colors $X = \{0, \ldots, 255\}$ be considered. As the histogram, the function $h \colon X \to \mathbb{R}$ such that $h(x)$ is a number of occurrences of pixels with color $x \in X$, is understood. In the following calculations only the colors which exist in the image $\bar{X} = \{x \in X : h(x) \neq 0\}$ are considered. In such a case

$$\mathrm{mean}(\bar{X}; h) := \sum_{x \in \bar{X}} h(x) \cdot x, \quad \mathrm{var}(\bar{X}; h) := \sum_{x \in \bar{X}} h(x) \cdot (x - \mathrm{mean}(\bar{X}))^2$$

are used as estimators for mean and variance. Now the cost function, for the purpose of minimization, can be introduced:

$$E(X_1, \ldots, X_k, h) = \sum_{i=1}^{k} p_i \cdot \left(-\ln(p_i) + \frac{1}{2}\ln(2\pi e) + \frac{1}{2}\ln(\sigma_i^2) \right),$$

where $\sigma_i^2 = \mathrm{var}(X_i; h)$, $p_i = \dfrac{|X_i|}{|\bar{X}|}$ and $\bar{X} = X_1 \cup \ldots \cup X_k$.

Detailed description of Cross-Entropy Clustering may be found in [3].

Consequently, for the purpose of minimizing the cost function, the greedy algorithm can be applied, and all the possible thresholds in the histogram can be verified (similarly to the Otsu method). In general, it is not possible to do it using GMM, which is a fuzzy method. The algorithm for greedy CEC with one threshold is presented in Algorithm 1.

Single-level binarization is well suited for images with clear foreground–background relationships [17]. In some situations, images contain more then only two different types of elements. Consequently, multilevel thresholding is needed [18].

An example of multilevel thresholding is presented in Fig. 2. It is easy to see that the histogram (see Fig. 1) contains three segments which are connected with

Algorithm 1 (CEC greedy approach):

input
histogram $h\colon [0, 255] \to \mathbb{R}$
$E = \infty$, threshold $= 0$
for $i \in [0, \ldots, 255]$, $h(i) \neq 0$ **do**
$\sigma_1 = \mathrm{var}(1, \ldots, i; h)$, $\sigma_2 = \mathrm{var}(i + 1, \ldots, 255; h)$, $p_1 = \frac{i}{255}$, $p_2 = \frac{255-i}{255}$.
if $E(\{0, \ldots, i\}, \{i + 1, \ldots, 255\}, h) < E$ **then**
E = E({0,. . . ,i}, {i+1,. . . ,255 } ,h)
threshold $= \frac{i+(i+1)}{2}$
end if
end for

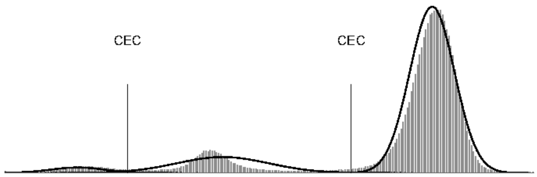


Figure 1. Multilevel binarization of the H04 image.

different parts of the image: the background, the gray letters in the title and the black lettering. Consequently, to extract all components, two thresholds are needed instead of one.

The greedy approach to the CEC algorithm can easily be extended onto multiple thresholds. Unfortunately, this approach is very time consuming in the case of a large number of components. In such a situation, the classical Cross-Entropy Algorithm can be used for extracting the components [3]. This approach reduces the number of components and finds the optimal number of thresholds.

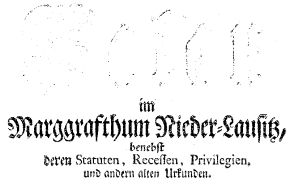
3. Experimental results

Table 1. Thresholds chosen by CEC, Otsu, GMM and Max Entropy.

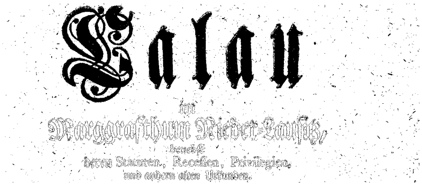
Method \ Image	H01	H02	H03	H04	H05	P01	P02	P03	P04	P05
CEC	170	185	171	179	204	140	151	172	185	130
Otsu	153	134	151	154	178	135	124	146	140	14
GMM	172	0	176	180	0	150	158	174	0	147
ME	165	165	154	91	116	138	152	178	154	14

In this section, the results of the CEC binarization are presented and compared with the Otsu, GMM and Maximum Entropy (ME) [11] methods. The algorithms

(a) original image



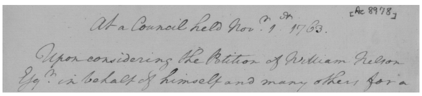
(b) lower section



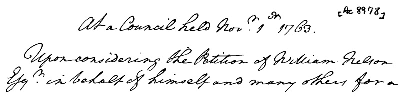
(c) middle section

(d) upper section

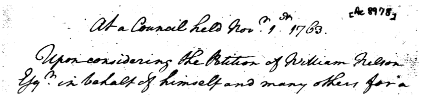
Figure 2. Multilevel binarization of the P03 image.



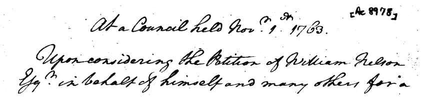
(a) Original



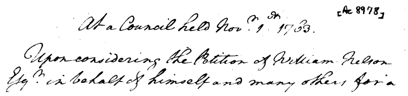
(b) Gold standard thresholding



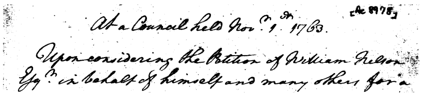
(c) CEC



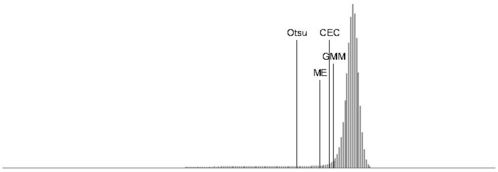
(d) Max Entropy



(e) Otsu



(f) GMM



(g) Histogram

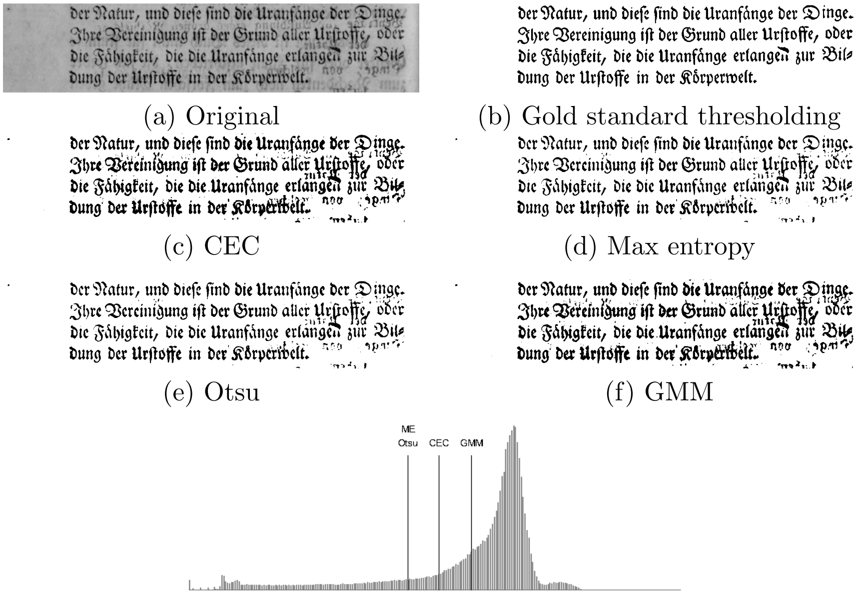
Figure 3. Handwritten text binarization comparison (H01).

Table 2. Comparison of the results according to precision, recall and MCC.

		CEC	Otsu	GMM	ME
H01	precission	0.7109	0.9263	0.6505	0.7880
	recall	0.9952	0.9006	0.9976	0.9868
	MCC	0.8286	0.9072	0.7898	0.8727
H02	precission	0.2662	0.7779	0.9990	0.4733
	recall	0.9922	0.9393	0.2816	0.9793
	MCC	0.4979	0.8513	0.5262	0.6721
H03	precission	0.5147	0.7178	0.4496	0.6911
	recall	0.9973	0.9752	0.9987	0.9804
	MCC	0.6790	0.8172	0.6244	0.8018
H04	precission	0.1765	0.2471	0.1740	0.8201
	recall	0.9995	0.9894	0.9996	0.7137
	MCC	0.3335	0.4297	0.3294	0.7479
H05	precission	0.1382	0.1622	1.0000	0.6969
	recall	0.9991	0.9617	0.0000	0.7653
	MCC	0.3223	0.3501	0.0051	0.7191
P01	precission	0.7765	0.8331	0.6408	0.7999
	recall	0.9840	0.9694	0.9963	0.9786
	MCC	0.8554	0.8839	0.7674	0.8678
P02	precission	0.8326	0.9686	0.7670	0.8250
	recall	0.9991	0.9644	0.9997	0.9993
	MCC	0.8876	0.9578	0.8400	0.8823
P03	precission	0.9250	0.9815	0.9146	0.8889
	recall	0.9827	0.9552	0.9843	0.9869
	MCC	0.9436	0.9618	0.9379	0.9230
P04	precission	0.4817	0.7235	1.0000	0.6606
	recall	0.9999	0.9594	0.0003	0.9870
	MCC	0.6490	0.8115	0.0165	0.7819
P05	precission	0.7212	0.8736	0.4822	0.9736
	recall	0.9824	0.9106	0.9990	0.9106
	MCC	0.8116	0.8729	0.6269	0.8729

are applied to real images and compared according to precision, recall and Matthews Correlation Coefficient [19]. Both precision and recall are in the range $[0, 1]$, where 1 indicates perfect partition. MCC is defined in the range $[-1, 1]$, where 1 indicates perfect partition, 0 – partition close to random and -1 means total disagreement between the received partition and the desired one. The results are compared with the gold standard thresholding as given in the DIBCO 2009 contest (http://users.iit.demokritos.gr/~bgat/DIBCO2009/).

Figures 3 and 4 show the results of Otsu, GMM, Maximum Entropy and CEC on images of handwriting [20]. The histograms printed along with the resulting images compare the thresholds chosen by the algorithms. The overall comparison of the thresholds for all images from DIBCO2009 contest can be found in Table 1 and comparison of results in Table 2.



(a) Original (b) Gold standard thresholding

(c) CEC (d) Max entropy

(e) Otsu (f) GMM

(g) Histogram

Figure 4. Printed text binarization comparison (P05).

As can be seen in the results, CEC always chooses a higher threshold than Otsu and usually higher or comparable to those chosen by ME. This means that more noise may be left in the resulting image, which causes more false positive classifications. On the other hand, more pixels are classified as foreground, therefore, there is a lower risk of false negative classifications. This means that CEC should have a higher recall than Otsu and ME, but lower precision. The thresholds given by CEC and GMM are very similar.

The Matthews Correlation Coefficient is a balanced measure, so it fits the domain well, as the foreground and background clusters are usually of very different sizes. According to MCC, CEC is better than GMM, but Maximum Entropy provides the best classifications of the four methods tested.

4. Summary

Cross-Entropy Clustering can be seen as a good thresholding algorithm, which attempts to match the results of the GMM methods while preserving the simplicity of the algorithm, like Otsu does. As CEC chooses the thresholds more aggressively than Otsu, it is a better option for those images in which every detail is important.

Consequently, CEC algorithm is a good tool which can by successfully used as a preprocessing step in more complicated image processing procedures.

The implementation of CEC algorithm as a plug–in for ImageJ is available on page http://ww2.ii.uj.edu.pl/~spurek/imageJ/CECMultilevelThresholding/CECMultilevel Thresholding.html.

5. References

[1] Gonzalez R.C., Woods R.E., *Digital image processing*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, United States of America, 2002.

[2] Huang D.S., *Systematic theory of neural networks for pattern recognition.* Publishing House of Electronic Industry of China, Beijing, 1996, 201.

[3] Tabor J., Spurek P., *Cross-entropy clustering.* Pattern Recognition, 2014, 47(9), pp. 3046–3059.

[4] Cover T.M., Thomas J.A., *Elements of information theory*, John Wiley & Sons, New York, United States of America, 2012.

[5] Grünwald P.D., *The minimum description length principle.* MIT Press, Cambridge, Massachusetts, United States of America, 2007.

[6] Otsu N., *A threshold selection method from gray-level histograms.* Automatica, 1975, 11(285–296), pp. 23–27.

[7] Liu Y., Srihari S.N., *Document image binarization based on texture features.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1997, 19(5), pp. 540–544.

[8] Prewitt J., Mendelsohn M.L., *The analysis of cell images.* Annals of the New York Academy of Sciences, 1966, 128(3), pp. 1035–1053.

[9] Doyle W., *Operations useful for similarity-invariant pattern recognition.* Journal of the ACM (JACM), 1962, 9(2), pp. 259–267.

[10] Johannsen G., Bille J., *A threshold selection method using information measures.* Proc. Sixth Int. Conf. Pattern Recognition, 1982, pp. 140–143.

[11] Kapur J.N., Sahoo P.K., Wong A.K.C., *A new method for gray-level picture thresholding using the entropy of the histogram.* Computer Vision, Graphics, and Image Processing, 1985, 29, pp. 273–285.

[12] Ridler T.W., Calvard S., *Picture thresholding using an iterative selection method.* IEEE transactions on Systems, Man and Cybernetics, 1978, 8(8), pp. 630–632.

[13] Tsai W.H., *Moment-preserving thresolding: A new approach.* Computer Vision, Graphics, and Image Processing, 1985, 29(3), pp. 377–393.

[14] Kittler J., Illingworth J., *On threshold selection using clustering criteria.* Systems, Man and Cybernetics, IEEE Transactions on, 1985, 5, pp. 652–655.

[15] Friedman N., Russell S., Image segmentation in video sequences: A probabilistic approach. In: *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., 1997, pp. 175–181.

[16] Huang Z.K., Chau K.W., *A new image thresholding method based on Gaussian mixture model.* Applied Mathematics and Computation, 2008, 205(2), pp. 899–907.

[17] Papamarkos N. B. G., *A new approach for multilevel threshold selection.* Graphical Models and Image Processing, 1994, 56, pp. 357–370.

[18] Mehmet S., Sankur B., *Survey over image thresholding techniques and quantitative performance evaluation.* Journal of Electronic Imaging, 2004, 13, pp. 146–165.

[19] Matthews B., *Comparison of the predicted and observed secondary structure of t4 phage lysozyme.* Biochimica et Biophysica Acta (BBA) – Protein Structure, 1975, 405, pp. 442–451.

[20] Gatos B., Ntirogiannis K., Pratikakis I., *ICDAR 2009 document image binarization contest (DIBCO 2009).* In: *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*, IEEE, 2009, pp. 1375–1382.