

Microscopic image segmentation with two-dimensional exponential entropy based on hybrid microcanonical annealing

A. NAKIB, H. OULHADJ and P. SIARRY

Université de Paris XII, Laboratoire Images, Signaux et Systèmes Intelligents
(LISSI, E. A. 3956), 61 avenue du Général De Gaulle
94010 Créteil, FRANCE

E-mail address: nakib@univ-paris12.fr

Abstract

Counting cells and following the evolution of the biological layers are important applications in microscopic imagery. In this paper, a microscopic image segmentation method with two-dimensional (2D) exponential entropy based on hybrid microcanonical annealing is proposed. The 2D maximum exponential entropy does not consider only the distribution of the gray-level information but also takes advantage of the spatial information using the 2D-histogram. The problem with that method is its time-consuming computation that is an obstacle in real time applications, for instance. We propose to combine the microcanonical annealing with the Nelder-Mead method, that was proved very efficient for non convex and combinatorial optimization. As the method is deterministic, the reproduction of the result is guaranteed, thus avoiding any randomization of the solution. The experiments on segmenting microscopic images proved that the proposed method can achieve a satisfactory segmentation with a low computation cost.

1. Introduction

Image segmentation is an important component in an artificial visual system; it consists in extracting objects from the background or separate different regions in an image. The segmentation problem has received a great deal of attention, thus any attempt to survey the literature would require too much space. However, a presentation of most segmentation methods may be found in [1-3].

The development of metaheuristic computing has been flourishing during the last decade. Many metaheuristics such as genetic algorithms [4], particle swarm optimization [5-6], simulated annealing [7] and others [2] have been applied to image segmentation problems. However, in most of these works the reproducibility of the solution was not studied. As the metaheuristic techniques are probabilistic, the reproduction of the solution cannot be guaranteed.

The initial motivation for applying microcanonical annealing (MA) [8] came from our interest in removing the stochastic character of the solution.

We also exploit the power of MA for searching in vast combinatorial state spaces for optimal thresholds. MA has some definite advantages over the canonical simulated annealing [9-10], since it does not require the generation of random numbers or the evaluation of transcendental functions, thus allowing much faster implementations.

In this study, the development of the hybrid algorithm aims at improving the performance of the segmentation

techniques in current practice. MA is hybridized with Nelder-Mead (NM) simplex method [11] in order to combine their advantages. NM method is very efficient for local search but its convergence is very sensitive to the starting point selected.

In this paper, we propose an automatic multilevel thresholding technique involving the optimization of the two-dimensional exponential entropy criterion based on hybrid microcanonical annealing for microscopic images. We first propose to extend the exponential entropy defined by Campbell [12] to the two-dimensional (2D) case and then to the multilevel thresholding case. The thresholding method based on the two-dimensional exponential entropy takes advantage of the gray-level spatial information. Afterwards, the 2D exponential entropy criterion is maximized through the hybrid NM-MA optimization algorithm.

This paper is outlined as follows: in the next section the computation of the two-dimensional histogram is presented. In section 3 the extension of the exponential entropy to the two-dimensional case is presented. In section 4 the optimization method based on the hybridization of the microcanonical annealing and the proposed segmentation algorithm are described. The results are discussed in section 5. Finally, we conclude in the last section.

2. Two-dimensional histogram

The two-dimensional (2D) histogram [1] of a given image is computed as follows: the average gray-level value of the neighbourhood of each pixel is calculated. Let $g(x, y)$ be the averaged image $f(x, y)$ using a window of size 3×3 .

In order to solve the frontier problem we disregard the top and bottom rows and the left and right columns. Then the 2D histogram is constructed using expression (1).

$$hist(i, j) = \frac{\text{Cardinal}(f(x, y) = i \text{ and } g(x, y) = j)}{\text{pixels number in the image}} \quad (1)$$

The joint probability is given by:

$$p_{ij} = hist(i, j) \quad (2)$$

where $i, j \in \{0, 1, 2, \dots, 255\}$.

The 2D histogram plane is represented in Fig.1: the first and the second quadrant denote the background and the objects respectively, the third and the fourth quadrant contain information about noise and edges alone, they are not considered. A threshold vector is (s, t) , where s , for $g(x, y)$, represents the threshold of the average gray-level of the pixel's neighbourhood and t , for $f(x, y)$, represents the threshold of the gray-level of the pixel. The quadrants containing background and objects (first and second) are

considered to be independent probability distributions; values in each case must be normalized in order to have a total probability equal to 1. In the case of image segmentation into N classes the *a posteriori* class probabilities are given by:

$$P_{m-1}[a_{n-1}, a_n] = \sum_{i=s_{n-1}}^{s_n-1} \sum_{j=t_{n-1}}^{t_n-1} p_{ij} \quad (3)$$

$$P_m[a_n, a_{n+1}] = \sum_{i=s_n}^{s_{n+1}-1} \sum_{j=t_n}^{t_{n+1}-1} p_{ij} \quad (4)$$

where $a_n \equiv (s_n, t_n)$, $n=2, \dots, N-1$, $m=2, \dots, N$ and N is classes number.

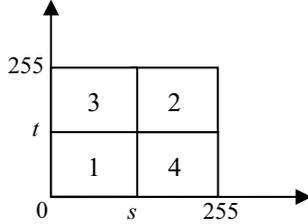


Figure 1: Two-dimensional histogram.

3. Two-dimensional exponential entropy criterion

We define the 2D exponential entropy by:

$$H = \left(\sum_i \sum_j p_{ij}^\alpha \right)^{1/(1-\alpha)} \quad (5)$$

where $\alpha \in \mathfrak{R}$ and $\alpha \neq 1$.

Thus the exponential entropies associated with the distributions of different image classes are defined below:

- The 2D entropy of the class $m-1$ can be computed through:

$$H^{(m-1)}[a_{n-1}, a_n] = \left(\sum_{i=s_{n-1}}^{s_n-1} \sum_{j=t_{n-1}}^{t_n-1} \left(\frac{p_{ij}}{P_m[a_{n-1}, a_n]} \right)^\alpha \right)^{1/(1-\alpha)} \quad (6)$$

- The 2D entropy of the class m can be computed through:

$$H^{(m)}[a_n, a_{n+1}] = \left(\sum_{i=s_n}^{s_{n+1}-1} \sum_{j=t_n}^{t_{n+1}-1} \left(\frac{p_{ij}}{P_m[a_n, a_{n+1}]} \right)^\alpha \right)^{1/(1-\alpha)} \quad (7)$$

For the convenience of illustration, two vectors $(s_0, t_0) = (1, 1)$ and $(s_N, t_N) = (255, 255)$ were added, where $t_0 < t_1 < t_2 < \dots < t_N$ and $s_0 < s_1 < s_2 < \dots < s_N$.

Then the total entropy is:

$$H^T[a_0, \dots, a_N] = \sum_{i=0}^{N-1} H^{(i+1)}[a_i, a_{i+1}] \quad (8)$$

According to the maximum entropy principle, the optimal vectors $(a_{1, \dots, N-1}^*) \equiv ((s_1^*, t_1^*), \dots, (s_{N-1}^*, t_{N-1}^*))$ should meet:

$$H^T(a_{1, \dots, N-1}^*) = \max \{ H^T(a_{1, \dots, N-1}) \} \quad (9)$$

where: $1 < s_1 < s_2 < \dots < s_{N-1} < 255$ and

$1 < t_1 < t_2 < \dots < t_{N-1} < 255$

In the case of one threshold ($N=2$) the computational complexity for determining the optimal vector (s^*, t^*) is $O(L^4)$. However, it is too time-consuming in the case of multilevel thresholding. For the n -thresholding problem, it requires $O(L^{2n+2})$. In this paper, we further present a microcanonical annealing algorithm for solving $\arg \max \{ H^T[(s_1, t_1), (s_2, t_2), \dots, (s_{N-1}, t_{N-1})] \}$ efficiently.

4. Hybrid microcanonical annealing

4.1. Microcanonical annealing

The microcanonical annealing is based on Creutz's method [9] known as microcanonical Monte Carlo simulation or the 'demon' algorithm. In its original form the demon algorithm does not aim at generating low energy states and hence is not directly useful for optimization. Optimization problems can usually be expressed in terms of a cost or energy function which is to be minimized over a space of possible solutions.

Microcanonical annealing based on Creutz's microcanonical algorithm can be stated as:

1. Choose an initial configuration S
2. Choose a demon energy $D > 0$
3. Repeat: until stopping conditions
 - a. Choose a new configuration S'
 - b. Let $\Delta E = E(S') - E(S)$
 - c. If $\Delta E \leq D$ accept the new configuration and update the demon, i. e. $S = S', D = D - \Delta E$
 - d. Else reject the new configuration
 - e. If quasi-equilibrium is reached, reduce the demon according to the schedule: $D = \varphi \cdot D$

Generation of a new configuration (step a) is performed as in the Metropolis algorithm. Any new configuration which would reduce the system energy is accepted. However, the energy lost by the system is given to an artificial variable called a 'demon'. Increases in system energy are only allowed if the demon can provide the necessary energy, which is then lost.

The acceptance function is deterministic and computationally simple. It replaces an exponential and the generation of a random number with a comparison and a subtraction. The sequence of states produced remains stochastic, but derives its randomness from the generating function. Here we anneal the demon value as Kirkpatrick *et al.* [9] and others [10] have annealed the temperature in simulated annealing.

4.2. Nelder-Mead method

It is a classical powerful local descent algorithm making no use of the objective function derivatives. A simplex is a geometrical figure consisting, in n -dimensions, of $(n+1)$ points x_0, \dots, x_n [8]. If any point of the simplex is taken as the origin, the n other points define vector directions that span the n -dimension vector space. If we randomly draw as initial starting point x_0 , then the other points are generated through the relation $x_i = x_0 + \lambda \cdot e_j$, where the e_j are n unit vectors, and λ is a constant (typically equal to one).

The simplex uses four elementary geometric transformations: reflection, contraction, expansion and multi-contraction. Through these operations the simplex can improve itself and come closer to a local optimum sequentially. The initial simplex moves, expands or contracts. To select the appropriate transformation, the method only uses the values of the function to be optimized at the vertices of the simplex considered. After each transformation, the current worst vertex is replaced by a better one. Trial moves are generated according to the following basic operations (where \bar{x} is the average value of the vector components x_1, \dots, x_n and δ, β, γ are constants):

$$\text{reflection: } x_r = (1 - \delta)\bar{x} - x_{n+1}$$

$$\text{expansion: } x_e = \gamma x_r + (1 - \gamma)\bar{x}$$

$$\text{contraction: } x_c = \beta x_{n+1} + (1 - \beta)\bar{x}$$

The algorithm starts by moving only the point of the simplex, where the objective function is “high” and another point image of the worst point is generated (reflection operation). If the new point is better than all the others, the simplex is expanded in this direction, otherwise, if it is at least better than the worst, the reflection is performed again with the new worst point. The algorithm performs a contraction step when the worst point is at least as good as the reflected point, in such a way that the simplex adapts itself to the function landscape and finally surrounds the optimum. If the worst point is better than the contracted point, the multi-contraction is performed. At each step, we check that the generated point is not outside the allowed reduced solution space.

4.3. Proposed segmentation algorithm

In this section, the proposed segmentation algorithm based on the optimization of the two-dimensional exponential entropy is presented. Figure 2 shows the different steps of the algorithm. The method exploits the hybrid microcanonical annealing to solve the segmentation problem expressed by (9). The hybridization of the microcanonical annealing consists of using the simplex method to find the local solution (thresholds). The algorithm does not require any special initialization. In order to get the contour of the regions, a contour detection based on the gradient operator is performed.

5. Results and discussion

In this section, we discuss the experimental results obtained using the proposed method. This discussion includes the choice of the optimal thresholds and the presentation of some microscopic images (bacteria, blood, cells, and vessels). Here, are presented only the results in the case of two classes’ segmentation.

Figure 3 shows the different original images used to illustrate our method. In table 1 the values of the parameters of NM-MA are summarized (these values were empirically chosen).

In table 2 the size of the different image tests is specified in term of the number of pixels. In order to quantify the performance of the optimization algorithm we define the speed gain ratio, that corresponds to the ratio of the

number of the possible solutions to the evaluation number of the objective function.

Results have been compared with those provided by the well known and robust Canny method (CM) [14]. The values in bold correspond to the optimal thresholds achieved through the NM-MA method.

1. Initialization,
 - 1.1. Define the minimal percent p of accepted transitions in the first energy level.
 - 1.2. Initialize D such that $p\%$ transformations tested can be accepted.
 - 1.3. Generate a random acceptable solution and compute its energy E .
 - 1.4. Choose the maximal number of tests in each energy level.
2. While the number of accepted solutions is non null.
 - 2.1. Find a local solution (configuration) by using the simplex method.**
 - 2.2. Repeat : until $Nbmaxtest$
 - 2.2.1. Generate a new configuration (S') by perturbing the obtained local configuration.
 - 2.2.2. Compute the associated (ΔE).
 - 2.2.3. If $\Delta E \leq D$ accept the new configuration and update the demon, i. e. $S = S', D = D - \Delta E$
 - 2.2.4. Compare the accepted configuration to the best state found since the beginning and save it if it is better.
 - 2.3. Reduce the demon according to the schedule: $D = \varphi.D$
3. Show the best configuration.

Figure 2: Microcanonical annealing for maximizing 2D exponential entropy.

In our first example, (figure 3(a)) the segmentation result obtained through NM-MA is shown in figure 4(a) and the result obtained via CM is shown in figure 5(a). By comparing these two images one can remark that only some of the bacteria were detected by CM, but NM-MA method detected all the bacteria. The number of regions corresponds to the number of bacteria, that is false in the case of CM.

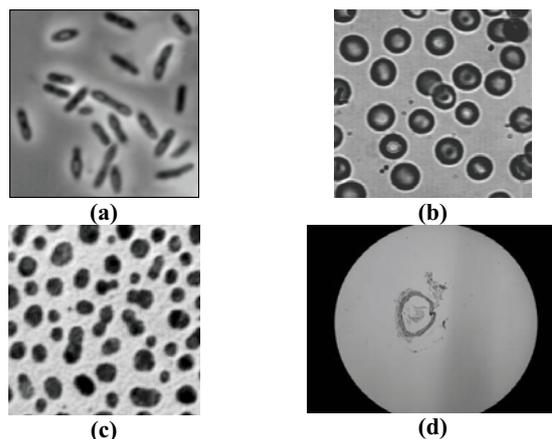


Figure 3. Original microscopic images. (a) bacteria, (b) blood, (c) cells, (d) vessel.

Table 1: Fitting of parameters of our algorithm.

Parameters	Value
$Nbmaxtest$	1000
p, φ, α	90%, 0.99, 0.5
δ, γ, β	1, 0.5, 0.5

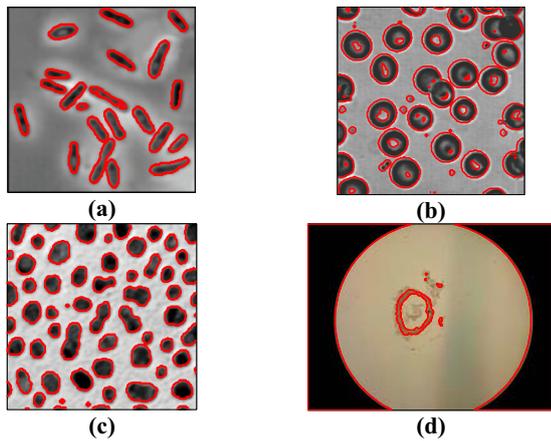


Figure 4. Segmentation results. (a) bacteria, (b) blood, (c) cells, (d) vessel.

Figure 3(b), figure 4(b) and figure 5(b) show the second image test, the segmentation result obtained through NM-MA and CM respectively. The result obtained via CM is inaccurate because of the double detection of each blood cell: the cell and another surrounding region. This is due to the gray-level variation around the cells.

Table 2: Experimental results for images in figure 3 provided by NM-MA.

Images in figure 3	Size (pixels)	Segmentation thresholds (S,T)	Speed gain factor
(a)	178x178	90, 98	32,76
(b)	272x265	197, 131	32,76
(c)	178x178	118, 141	21,84
(d)	2048x1536	171, 70	21,84

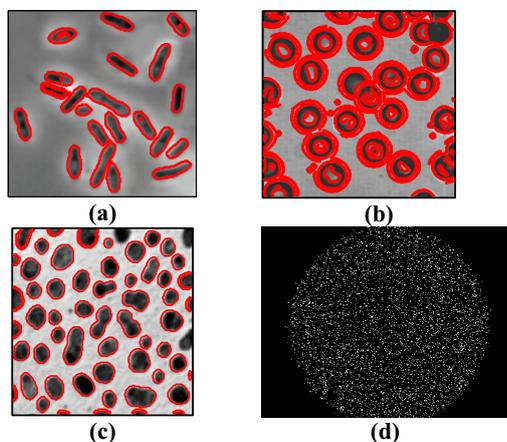


Figure 5. Results obtained through the Canny method. (a) bacteria, (b) blood, (c) cells, (d) vessel.

In the third example (figure 3(c)) the results achieved by both methods (figure 4(c) and figure 5(c)) are equivalent.

The last microscopic image is presented in figure 3(d). The goal is to detect the media of a vessel and, after applying the method, to follow the evolution of this layer. One can notice the different variations in the contrast on the original image, making CM unable to produce a good segmentation. On the contrary, the NM-MA method allows for efficient detection of the contours of the media. Some methods use CM to localize the cells [2], and if the localization is false, like in figure 5 (d), the segmentation algorithm will fail to converge.

6. Conclusion

In this paper, we proposed a new approach to find optimal thresholds, based on hybrid microcanonical optimization. In the first phase the two dimensional histogram was constructed using spatial information, local average gray value, to choose optimal thresholds. In the second phase, a new extension of the one-dimensional exponential entropy of degree α to the two-dimensional case and its generalization to multilevel segmentation were developed. In the third phase, microcanonical annealing (MA) is introduced to remove the stochastic feature of the solution. The algorithm convergence speed is improved by hybridizing MA with NM.

It is clearly seen from the experimental results that the presented method based on hybrid microcanonical optimization is more efficient than the classical Canny method in the case of vessel microscopic images. This approach allows for good microscopic image segmentation by using exponential two-dimensional entropy. It can also be useful for gray-level image segmentation.

References

- [1] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146-165, 2004.
- [2] T. Jiang, F. Yang, "An evolutionary Tabu search for cell image segmentation", *IEEE Trans. On Syst., Man and Cybernetics-Part B: Cybernetics*, vol. 32, no. 5, pp. 675-678, Oct., 2002.
- [3] A. Garrido, N. Pérez de la Blanca, "Applying deformable templates for cell image segmentation", *Pattern Recognition*, vol. 33, pp. 821-832, 2000.
- [4] W. Tao, J. Tian, J. Liu, "Image segmentation by three level thresholding based on maximum fuzzy entropy and genetic algorithm", *Pattern Recognition Letters*, vol. 24, pp. 3069-3078, 2004.
- [5] Y. Peng-Yeng, "Multilevel minimum cross entropy threshold selection based on particle swarm optimization, *Applied Mathematics and Computation*, vol. 40, pp. 596-608.
- [6] E. Zahara, S. S. Fan, D. Tsai, "Optimal multi-thresholding using a hybrid optimisation approach", *Pattern Recognition Letters*, 26, 1082-1095, 2004.
- [7] W. Synder, G. Bilbro, "Optimal thresholding: A new approach", *Pattern Recognition Letters*, vol. 11, pp. 803-810, 1990.
- [8] M. Creutz, "Microcanonical Monte Carlo simulation", *Phy. Rev. Letters*, vol. 50, no. 19, pp. 1411-1414, 1983.
- [9] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science*, vol. 220, pp. 671-680, 1983.
- [10] R. W. Eglese, "Simulated annealing: A tool for operational research", *Euro. J. of Op. Research*, vol. 46, pp. 271-281, 1990.
- [11] J.A. Nelder, R. Mead, "A simplex method for function minimization", *Computer Journal*, vol. 7, pp 308-313, 1965.
- [12] L. L. Campbell, "A coding theorem and Rény's entropy", *Inform. Contr.*, vol. 29, pp. 423-429, 1965.
- [13] J. Canny, "A Computational Approach to Edge Detection", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, No. 6, pp. 679-698, 1986.