A Study of The Firefly Meta-Heuristics for Multi-Threshold Image Segmentation

L. A. Lopes, G. A. Wachs-Lopes, H. Erdmann and P. S. Rodrigues
Signal Processing Group
Centro Universitário da FEI
São Bernardo do Campo, São Paulo, Brazil
lukercio@live.com, guilhermewachs@gmail.br, horste@fei.edu.br, psergio@fei.edu.br

Abstract

Thresholding-based image segmentation algorithms are usually developed for a specific set of images because the objective of these algorithms is strongly related to their applications. The binarization of the image is generally preferred over multi-segmentation, mainly because it's simple and easy to implement. However, in this paper we demonstrate that a scene separation with three threshold levels can be more effective and closer to a manually performed segmentation. Also, we show that similar results can be achieved through a firefly-based meta-heuristic. Finally, to compare the distance between a manual and an automatic segmentation through a similarity measure that is suggested and proven.

1. Introduction

Image segmentation is a task with applications in several areas related to digital image processing. It can be done by estimating the number of thresholds used to partition an image into regions of interest. The most simple thresholding technique is to divide the image in two regions, binarizing the search space.

Among the known techniques to define a threshold that splits an image in two clusters, there are those based on the probabilistic-distribution entropy for the color intensities. In [15], T. Pun wrote the first algorithm for image binarization based on the traditional Shannon entropy, assuming that the optimal threshold is the one that maximizes the additivity property for its entropy. Such property states that the total entropy for a whole physical system (represented by its probability distribution) can be calculated from the sum of entropies of its constituent subsystems (represented by their individual probability distributions).

Kapur et al. [9] maximized the upper threshold of the maximum entropy to obtain the optimal threshold, and Abutaleb [1] improved the method using bidimensional entropies. Furthermore, Li and Lee [11] and Pal [14] used the direct Kullback-Leibler divergence to define the optimal threshold. And some years before, Sahoo et al. [16] used the *Reiny*-entropy seeking the same objective. More details about these approaches can be found in [3], which presents a review of entropy-based methods for image segmentation.

Considering the restrictions of Shannon entropy, Albuquerque et al. [2] proposed an image segmentation method based on Tsallis non-extensive entropy [18], a new kind of entropy that is considered as a generalization of Shannon entropy through the inclusion of a real parameter q, called "non-extensive parameter". The work of Albuquerque showed promising results and a vast literature demonstrating the performance of this method against the Optimal Threshold Problem Although it is a new contribution to the field, this paper will not address the Tsallis entropy.

A logical extension of binarization is called multi-thresholding [20] [8], which considers multiple thresholds on the search space, leading to a larger number of regions in the process of segmentation.

However, since the optimal threshold calculation is a direct function of the thresholds quantity, the time required to search for the best combination between the thresholds tends to grow exponentially. Furthermore, the optimum quantity of thresholds is still a topic for discussion. Thus, the literature has proposed the use of meta-heuristics that may be efficient for the calculation of thresholds, one of them being the Firefly.

Recently, M. Horng [8] proposed an approach based on Minimum Cross-Entropy thresholding (MCET) for multilevel thresholding with the same objective function criterion as proposed by P. Yin [20]. The main conclusion of the work was that the Cross-Entropy based method, a linear time algorithm, obtained thresholds values very close to those found by equivalent brute-force algorithms. However, the results were inconclusive since their methodology to evaluate the experiment was subjective.

This article proposes an analysis of the Firefly meta-heuristics for multi-threshold-based image segmentation. We also present the use of a Golden Standard Image Base that allows us to compare the segmentation results of different algorithms in an objective manner.

2. The Proposed Methodology

The strategy used in this study is the comparison of the obtained results with exhaustive methods results, both manual and automatic. Although these methods have polynomial complexity in $O(n^{d+1})$ order, it is computationally expensive to calculate the results for $d \geq 3$.

One important issue is to define the the number of thresholds required to obtain a segmentation result as close as possible to that obtained manually. The answer seems subjective and dependent of cognitive factors that are outside the scope of this paper. Thus, the database used for comparison of the results consists of several images that were manually segmentated during psychophysical experiments that were defined and performed in [13]. Moreover, the results will be compared in two directions. First, we compare the results of the exhaustive segmentation with the respective manual one. Then, we compare the results of the manual segmentation with the ones obtained with the Firefly meta-heuristics, allowing us to draw a comparison between both methods used.

Although answering cognitive questions is not the purpose of this paper, the exhaustive search of the entire result space, allows us to observe the minimum amount of thresholds required to obtain the closest result to the manual segmentation. This lower limit can be used for any segmentation algorithms besides those cited in this paper.

The method used to compute the threshold-based multi segmentation with the Firefly meta-heuristics is shown in Figure 1.

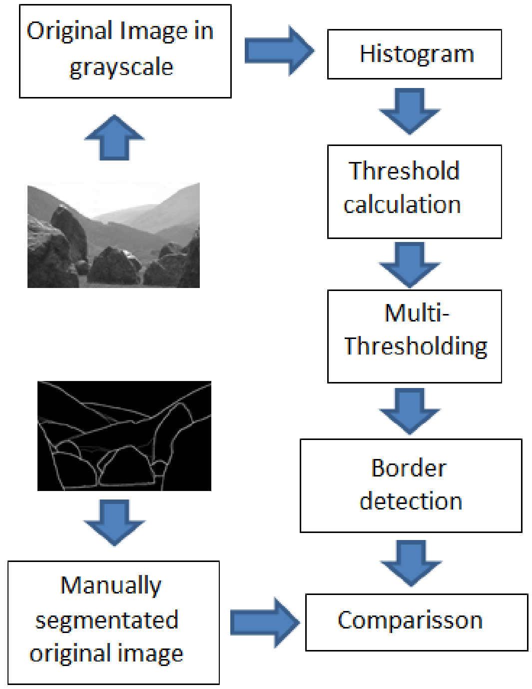

Figure 1: Proposed comparison methodology scheme. The left column shows an example of original image and its manual segmentation taken as a basis of comparison.

3. Firefly Meta-Heuristics (FF)

The Firefly (FF) algorithm was proposed by Xin-She Yang [19] and is a meta-heuristics inspired on the fireflies behavior, which are attracted among themselves according to their natural luminescence.

After their interactions, convergence is achieved through the creation of fireflies clusters, on which the brighter attract the less bright ones under certain restrictions, such as: (i) all fireflies are unisex so that one firefly will be attracted to any other fireflies regardless of their sex; (ii) attractiveness is proportional to their brightness, thus for any two flashing fireflies, the less bright one will move towards the brighter one, remembering that the glow declines as the distance between them increases; (iii) if, for any given firefly, there isn't any other brighter firefly, than it moves randomly.

The algorithm is designed to model a non-linear optimizer associating the thresholds to fireflies. The kernel depends on these variables, which are associated with the fireflies glow and can be modified according to be more appropriate to the data that is being manipulated. Then, the fireflies luminescences are updated iteratively under pre-established rules until the algorithm convergence to a global minimum.

The papers of Lukasik and Zak [12] and Yang [19] suggest that the FF overcome other meta-heuristics, such as the Ant Farm [4], Tabu search [5], PSO [10] and Genetic Algorithms [6]. Thus, the FF was presented as a computing-time efficient method to the Multilevel Thresholding Problem (MLTP) Recently, the work of[7] showed a computational time comparison of the FF against the Otsu method, demonstrating that the FF is more efficient when the evaluation function is modeled with the maximum inter-cluster variance. Other works, such as [8] and [20] also showed similar results when applied to the MLTP.

Specifically for the MLTP modeling, each firefly is a d-dimensional variable, where each dimension is a single threshold that partitions the histogram space. In the specific work of M. H. Horng and R. J. Liou [8], the goal was to minimize the objective function using the Cross-Entropy of the intensities histogram associated with each segmented image criteria.

The Algorithm 1 describes the FF, where a solution set of n initial fireflies is given on line 3. Each firefly f_i is a d-dimensional vector and x_k^i is the k-eth threshold of i-eth solution. More Details about The FF can be found in [8] and [19].

Algorithm 1 FF Algorithn to MLTP (adapted from [8])

1: Input: n : number of fireflies; d: dimention; γ: absorving coeficient; α: motion; β: attractness factor; MG: maximum number of generations;
2: Output: aceptable threshold set $f_i^* = \{x_1^i, x_2^i, \dots, x_d^i\}$
3: Define initial values: $t = 0$, $\alpha_0 = 1.0$
4: Randomly define the initial population $\{f_1, f_2, \dots, f_n\}$ where $f_i = \{x_1^i, x_2^i, \dots, x_d^i\}$ is the i-eth d-dimensional solution of the firefly
5: **while** $t \leq MG$ **do**
6:   **for** $i = 1 \to n$ **do**
7:     **for** $j = 1 \to n$ **do**
8:       Calculate the $r_{i,j}$ distance between the glows $Z(f_i)$ and $Z(f_j)$
9:     **end for**
10:   **end for**
11:   **for** $i = 1 \to n$ **do**
12:     Glow evaluation $Z(f_i)$
13:     **for** $j = 1 \to n$ **do**
14:       **if** $Z(f_i)$ is less bright than $Z(f_j)$ **then**
15:         Move the f_i firefly towards the f_j firefly, according to the following update rule:
16:         Randomly generate a new solution $\mu_t = \{x_1, x_2, \dots, x_d\}$
17:         $\alpha_t = \alpha\alpha_t$
18:         $\beta_0 = \beta \exp(-\gamma r_{i,j}^2)$
19:         **for** $k = 1 \to d$ **do**
20:           $x_k^i = (1 - \beta_0)x_k^i + \beta_0 x_k^j + \alpha_t\mu_t$
21:         **end for**
22:       **end if**
23:     **end for**
24:   **end for**
25:   Sort the fireflies according to their glow $Z(.)$
26:   Define the brightest firefly f_i^* as the current result
27:   $t = t + 1$
28: **end while**

4. Image Database

In this work, we made use of 300 images from the Berkeley University database [13]. Such images are composed of various natural scenes, wherein each was manually segmented. The task to segment an image into different cognitive regions is still an open problem. It is possible to highlight two main reasons for it to be considered a difficult task: (i) a good segmentation depends on the context of the scene, as well as from the point of view of the person who is analyzing it; and (ii) it is rare to find a database for formal comparison of the results. Generally, researchers present their results comparing just a few images, pointing out what they believe is correct. In these cases, probably the same technique will work only with other images that belong to the same class. Still, the question that remains unresolved is: " What is a correct segmentation?".

In the absence of an answer to the question, a reference is necessary that allows the comparison of several techniques under the same database or parametrization. Regarding this, the image database used here can be considered as an attempt to establish such reference.

The Figure 2 shows 10 examples of the pictures that belong to the database and the overlapping of 5 edge-maps derived from the manual segmentation, which denotes the high level of consistency between segmentations done by different persons. Additional details about this image database can be found on [13].

When overlapping the five edge-maps of the same image as in Figure 2, some edges do not match, thus the final intensity of each edge of the overlapped image is going to be higher if it overlaps more edges and less intense otherwise. In this article, we made use of 300 images as comparison base (gold standard) for our experiments.

Furthermore, the divergence of information in the absolute value between the automatically-obtained segmentations and the golden standard (manually-obtained segmentations) were also not considered as a segmentation-quality measure. So, the image database is used as a tool for comparison between the results of the two evaluated methods.

5. Similarity Measure

We defined a function to measure the similarity between the manual and the automatic segmentation. However, this is a difficult task and the problem is still unsolved. Sezgin and Sankur [17] proposed 5 quantitative criterias for measuring the luminance region and shaped 20 classical methods to measure the similarity between them. But the criteria they proposed was not based on a golden standard defined set of images, thus the method of comparison proposed in [17] can be used only as an intrinsic quality evaluation of the segmented areas: i.e, one output image segmented into uniformly molded regions can not be considered as close as expected to the manual segmentation.

On the other side, golden standard based measuring techniques are also difficult to propose when the system needs to detect several regions of the image at the same time, a common task in computer vision. Besides that, to compare corresponding edges brings difficulty to detect entire regions, as well as their location in space. Also, in the area of computer vision, is an important demand to be able to deduct

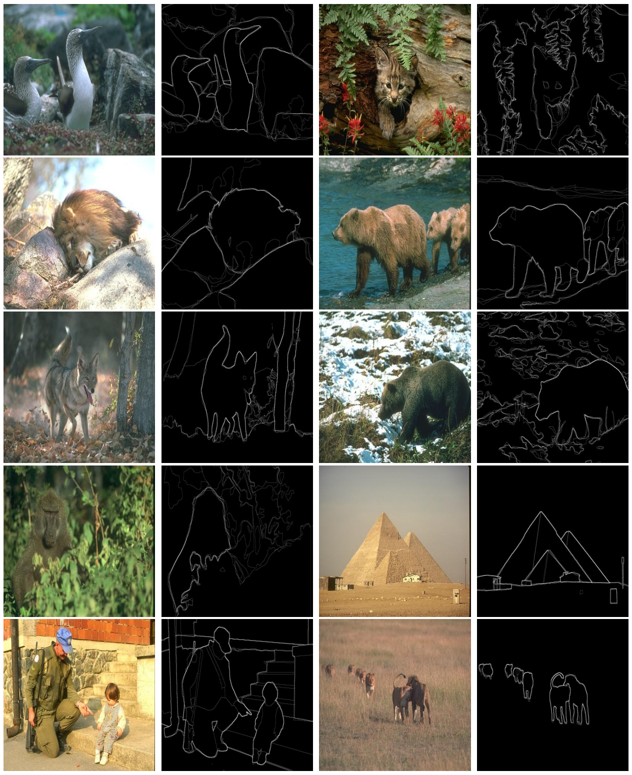

Figure 2: Segmented images used on the experiments.

regions that are interrelated.

Although it is possible to design an algorithm which tolerates localization errors, it is likely that detecting only the matching pixels and assuming all others are flaws or false positive and may provide a poor performance.

One can speculate from Figure 2 that the comparison between the edge-maps derived from the automatic and manual segmentations must tolerate localization errors as long as there are also divergences on the edges of the golden standard. Thus, the consideration of some differences can be useful in the final result as shown in [13].

On the other hand, from 2D edge-maps, such as the one we used, one can obtain two types of information: geometrical dispersion and intensity dispersion. The geometric dispersion measures the size and the location of the edges; the intensity dispersion measures how common is that edge among all manual segmentations that were overlapped. Thus, the geometric dispersion between two edge-maps has its information measured in a quantitative manner, in the x and y dimensions, while the luminance dispersion can be represented by the z dimension.

The divergence of information between the two edge-maps of an $M \times N$ image in the x dimension is calculated by the Euclidean distance between the two maps (i.e. the M_x as vertical projection at the edge map for automatic segmentation and the H_x is the corresponding vertical projection for the manual one). So, in this article, we propose an evaluation function between the two edge-vertical-projection M_x and H_x of the x dimension presented in (1) to measure how far the automatically-obtained segmentation is from the manual one in this specific direction:

$$Sim_x(M_x|H_x) = \sqrt{\sum_M (M_x(i) - H_x(i))^2}, \qquad (1)$$

where M_x and H_x are the image edges projections in the x direction, manual and automatic respectively.

Similarly, the corresponding functions are proposed respectively for y (2) and z (3) directions:

$$Sim_y(M_y|H_y) = \sqrt{\sum_N (M_y(i) - H_y(i))^2} \qquad (2)$$

and

$$Sim_z(M_z|H_z) = \sqrt{\sum_L (M_z(i) - H_z(i))^2}, \qquad (3)$$

where N e L are the sizes of y e z distributions (for horizontal and luminance distribution, respectively). Note that N is the image resolution on the y dimension and L is the total of gray levels (i.e. 256).

Thus, in this study, we propose the following evaluation function to measure the similarity between two edge-maps:

$$Sim(M|H) = Sim_x + Sim_y + Sim_z \qquad (4)$$

6. Experiments and Discussion

The methodology shown in Figure 1 describes both scenarios used in this paper: i) the segmentation with 1, 2 and 3 thresholds found by an exhaustive search; and ii) the segmentation with 1, 2 and 3 thresholds obtained with the use of the FF meta-heuristic.

The main reason for using the exhaustive search was to guarantee that the whole solution space is explored in order to find the thresholds that provide the closest results to the golden standard for each image.

The authors of [19] and [8] presented multi-thresholding approaches based on the FF algorithm and made a comparison with the exhaustive strategy, where the FF's kernel was chosen as being the Cross-Entropy approach. This type of comparison is limited, since it is only a relative matching between the FF result and the one obtained with the entropic method (achieved in an exhaustive manner). So there is no way of knowing if there are other better solutions (threshold levels), since the search space was not entirely explored. Another limitation of the method presented in [19] and [8] is the similarity measure used, since they used the noise difference of each segmentation as a metric.

In this article the two limitations listed above were addressed in the following manner: we explored the entire solution space, for 1, 2 and 3 thresholds, ensuring that there was no better solution from the similarity measure point of view. And we also used the manually segmented image set presented in Section 4 as the basis for comparing the results of our experiments.

6.1. Exhaustive segmentation

As in Figure 2, we applied a threshold (1 level) for each image. Then, for each possible threshold, the image was segmentated. Then we applied a gradient-based edge detector which return de boundaries of the regions that were found. Next, the comparison between the newly obtained edge-map and the golden standard is given by Equation (4). If $T = \{t_1, t_2, \ldots, t_L\}$, where $L = 256$, then the optimal threshold $t_{opt} \in T$ is the one that minimizes Equation (4). These procedure was then repeated for 2 and 3 levels, remembering that the solutions space grows exponentially, since we need $|T|^2$ and $|T|^3$ tests for segmentating with 2 and 3 levels respectively.

Despite being an exhaustive strategy, the algorithm surely returns the optimal results, which can be used as a lower boundary for minimizing the Equation (4). This strategy is more appropriate than the noise minimization that was proposed in [19] and [8].

If $I = \{i_1, i_2, \ldots, i_{300}\}$ is the 300 image set, for each $i_j \in I$, we can associate an array $S_i = [s_{i1}, s_{i2}, s_{i3}]$, where s_{i1} is the value given by Equation (4) for the binarization of i_j with the optimal t_{opt}; s_{i2} is the following value for the multi-thresholding of i_j with the optimal thresholds $\{t_{opt1}, t_{opt2}\} \in T$; and finally, s_{i3} is the corresponding array with 3 thresholds $\{t_{opt1}, t_{opt2}, t_{opt3}\} \in T$.

For better visualization of the results, we created an $M_{300\times 3}$ matrix, where each M_{ij} ($1 \leq i \leq 300$ and $1 \leq j \leq 3$) element is the value of $s_{ij} \in S_i$ associated with the i image. Each i line of M was normalized into 3 intensity values $L \in \{0, 128, 255\}$, so that $M_{ij} = 0$ if $s_{ij} = maxS_i$; $M_{ij} = 255$, if $s_{ij} = minS_i$; and $M_{i,j} = 128$, if s_{ij} is the median of S_i. The Figure 3 shows M as one single image with dimensions 300×3 resized to 300×300 for better visualization.

Thus, for cell (i, j) of M on Figure 3, the brighter the pixel, the more the image segmentated with the j-eth threshold resembles the manually segmentated image. The darker the pixel, greater the difference between them.

6.2. Segmentation with the FF meta-heuristics FF

The experiments were repeated using the FF segmentation, except for the threshold calculation, that is done with the Algorithm 1

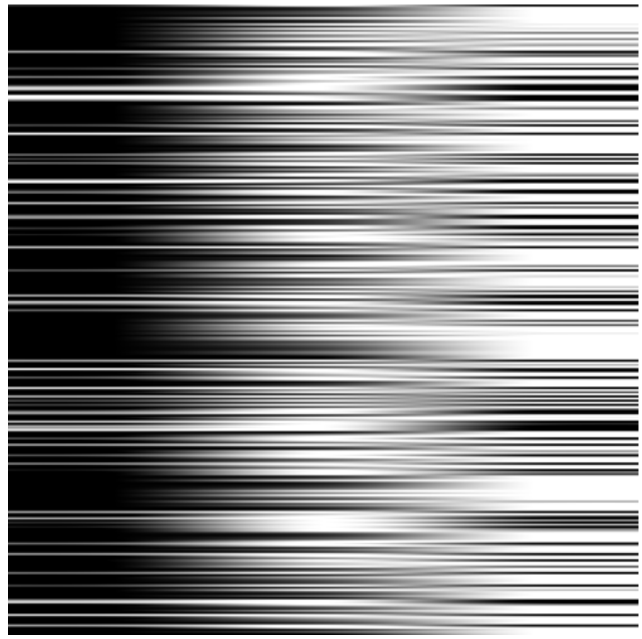

Figure 3: Exhaustive segmentation results. Each line represents one of the 300 images. The columns are the results of the segmentation with 1, 2 and 3 thresholds.

Just like the experiments with the exhaustive search method, we also created a $M_{300\times 3}$ matrix with the same properties as the previous one.

Comparing the Figures 3 and 4, it is possible to notice the similarity between them, indicating that the FF results are as good as the exhaustive method.

Another observation is that, due to the matrix normalization, the brighter the $M_{i,j}$ cell, the closer the j-eth threshold segmentation is to the manual segmentation. So, it is possible to conclude that this approximation gets higher as the j value increases. That is, if the goal of the threshold segmentation is to find the threshold set that results in a segmentation that is close to the manual one, then, to use 3 thresholds is more efficient than 2 which in turn is better than 1. However, one can speculate that beyond 3 thresholds, the results tend to get worse since this leads to the over-segmentation of the image. But this is a further investigation out of the current scope.

7. Conclusion

This paper presented the application of a meta-heuristic inspired by the fireflies behavior for multi-thresholding image segmentation. The proposed method's results were compared with the results of exhaustive search for 1, 2 and 3 thresholds through a manually segmented database. By searching all the solution's up to 3 thresholds space, we were able to establish a lower limit for the comparison with

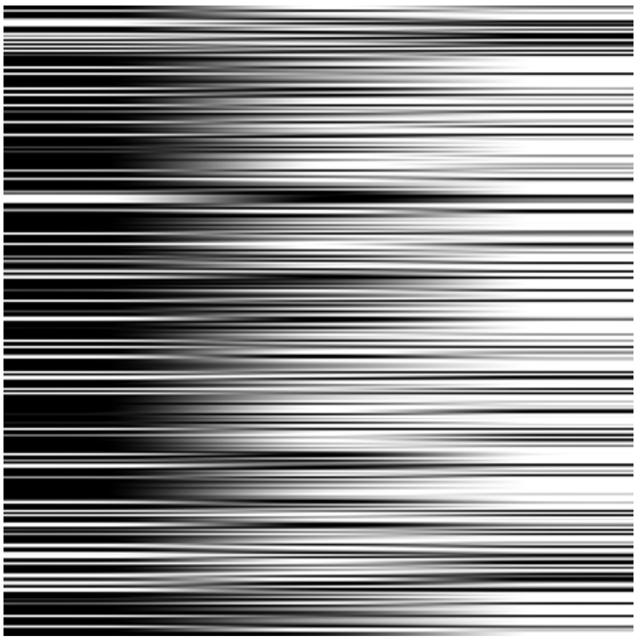

Figure 4: FF Meta-Heuristics segmentation results. Each line represents one of the 300 images from the database. The columns are the results of the segmentation with 1, 2 and 3 thresholds.

the manual segmentation results. This limit is useful for other algorithms or thresholds-based segmentation strategies.

The experiments indicate that the FF results are those close to the exhaustive search. Moreover, these results suggest that, for threshold-based segmentation, separating the image into four groups with three thresholds, provides better chances to reach the edges obtained with the manual segmentation as a final result than dividing into three groups. Furthermore, this last separation is still closer to the manual results than the separation in two groups, the so-called binarization.

Acknowledgments

We thank CNPq and CAPES for supporting FEI pos-graduation. We are also thankful for FAPESP for the direct support to the project (Case 2012/17439-2).

References

[1] A. S. Abutaleb. A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Graphics Image Process*, 47:22–32, 1989.

[2] M. P. Albuquerque, M. P. Albuquerque, I. Esquef, and A. Mello. Image thresholding using tsallis entropy. *Pattern Recognition Letters*, 25:1059–1065, 2004.

[3] C.-I. Chang, Y. Du, J. Wang, S.-M. Guo, and P. Thouin. Survey and comparative analysis of entropy and relative entropy thresholding techniques. *IEEE Proceedings, Vision, Image and Signal Processing*, 153(6):837–850, Dec. 2006.

[4] M. Dorigo. *Optimization, learning, and natural algorithms*. Ph.D. Thesis, Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy, 1992.

[5] F. Glover. Tabu search. *PART I, ORSA Journal of Computing 1*, pages 190–206, 1989.

[6] D. E. Goldberg. *Genetic Algorithms in search, Optimization, and Machine Learning*. Addison Wesley, Reading, MA, 1997.

[7] T. Hassanzadeh, H. Vojodi, and A. M. Eftekhari. An image segmentation approach based on maximum variance intra-cluster method and firefly algorithm. In IEEE, editor, *Seventh International Conference on Natural Computation*, pages 1844–1848, Shanghai, China, 2011.

[8] M. H. Horng and R. J. Liou. Multilevel minimum cross entropy threshold selection based on firefly algorithm. *Expert Systems with Applications*, 38:14805–14811, 2011.

[9] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Graphics Image Process*, 29:273–285, 1985.

[10] J. Kennedy and R. C. Goldberg. Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks*, volume IV, pages 1942–1948, 1997.

[11] C. H. Li and C. K. Lee. Minimum cross entropy thresholding. *Pattern Recognition*, 26:617–625, 1993.

[12] S. Lukasik and Zak. Firefly algorithm for continuous constrained optimization tasks. In *1st international conference on computational collective intelligence, semantic web*, 5-7 October 2009.

[13] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.

[14] N. R. Pal. On minimum cross entropy thresholding. *Pattern Recognition*, 26:575–580, 1996.

[15] T. Pun. Entropic thresholding: A new approach. *Comput. Graphics Image Process*, 16:210–239, 1981.

[16] P. Sahoo, S. Soltani, A. Wong, and Y. Chen. A survay of thresholding techniques. *Computer Vision Graphics Image Processing*, 41(1):233–260, 1988.

[17] M. Sezgin and B. Sankur. Survay ove image thresholding techniques and quantitative performance evaluation. *Journal of Eletronic Imaging*, 13(1):146–165, January 2004.

[18] C. Tsallis. Nonextensive statistics: Theoretical, experimental and computational evidences and connections. *Brazilian Journal of Physics*, 29(1), March 1999.

[19] X. S. Yang. Firefly algorithms for multimodal optimization. In *Stochastic algorithms: Fundation and applications, SAGA 2009. Lecture notes in computer science*, volume 5792, pages 169–178. 2009.

[20] P. Y. Yin. Multilevel minimum cross entropy threshold selection based on particle swarm optimization. *Applied Mathematics and Computation*, 184:503–513, 2007.