Magnetic Resonance Image Segmentation Based on Two-Dimensional Exponential Entropy and a Parameter Free PSO

Amir Nakib, Yann Cooren, Hamouche Oulhadj, and Patrick Siarry

Université de Paris XII, Laboratoire Images, Signaux et Systèmes Intelligents (LISSI, E. A. 3956), 61 avenue du Général De Gaulle
94010 Créteil, France
{nakib,cooren,oulhadj,siarry}@univ-paris12.fr

Abstract. In this paper, a magnetic resonance image (MRI) segmentation method based on two-dimensional exponential entropy (2DEE) and parameter free particle swarm optimization (PSO) is proposed. The 2DEE technique 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 this method is its time-consuming computation that is an obstacle in real time applications for instance. We propose to use a parameter free PSO algorithm called TRIBES, that was proved efficient for combinatorial and non convex optimization. The experiments on segmentation of MRI images proved that the proposed method can achieve a satisfactory segmentation with a low computation cost.

Keywords: image segmentation, two-dimensional exponential entropy, particle swarm optimization, tribes, parameter free.

1 Introduction

The increasing need for analyzing the brain magnetic resonance images (MRI) allowed to establish MRI segmentation as an important research field. For instance, in order to make easy the evaluation of the ventricular space evolution, a multilevel MRI segmentation is required. In this paper, we consider the problem of detecting the ventricular space from MRI of the brain. The image segmentation problem at hand is difficult because of the common occurrence of peri-ventricular lesions in MRI of even normal aging subjects, which locally alter the appearance of the white matter surrounding the ventricular space.

The segmentation problem has received a great deal of attention, thus any attempt to survey the literature would be too space-consuming. The most popular segmentation methods (tissue classification methods) may be found in [1] to [13]. The common class of parametric methods used in brain MRI segmentation is based on an expectation-maximization framework. This class of methods is based on the assumption that a mixture Gaussian distribution is assumed as a model for the voxel intensity probability distribution. However, in most cases, the distribution is far from being Gaussian. Many

N. Monmarché et al. (Eds.): EA 2007, LNCS 4926, pp. 50–61, 2008.

authors tried to overcome this problem by regularizing the misclassification error through spatially constraining the segmentation process with prior information from a probabilistic atlas [4]. However, the method becomes very sensitive to the correct alignment of the atlas with the image and too time consuming. Actually, doctors do not want to spend a lot of time waiting the result of segmentation, because of the large number of subjects. That induces the need for a fast segmentation algorithm.

Many authors have applied to brain MRI classical segmentation methods, details are given in [3], [6], and [7] to [13]. In order to overcome the problem of these methods, some post-classification methods were proposed [7].

The main contribution of the work we present here is a novel method for brain MRI segmentation based on an information measure, defined in [8], called exponential entropy (EE). The EE information measure solves the different problems related to the use of the classical Shannon entropy, pointed out in [9], i.e. Shannon's entropic description is not defined for distributions that include probabilities of 0. To avoid the problem of the spatial distribution, we defined a two-dimensional histogram, that takes into account the pixel spatial distribution. We also extended the EE to the two-dimensional and multilevel case.

As the computation complexity of the problem at hand exponentially increases with the increase of the number of classes, a fast optimization metaheuristic is needed to search for the optimal solution. Most metaheuristics have the drawback of having parameters which must be set by the user. According to the values given to these parameters, the algorithm is more or less efficient. However, there are many applications for which the user of the algorithm has no time to waste with parameter tuning. Practically, if the values of the objective function result from an experimental time costly process, it would be not possible to lead tests on the values of parameters, particularly in industrial applications. Tuning the parameters requires a minimum of experience about the used algorithm, so, it would be difficult and time consuming for a novice user to find the optimal set of parameters.

In this paper, we propose to use a parameter free PSO algorithm, called TRIBES, that does not need any parameter fitting [14]. Many authors tried to make the PSO algorithm free of parameters [15], [16] and [17]. But the first really parameter free algorithm, called TRIBES, was proposed by Clerc [14].

This paper is outlined as follows: in the next section, the computation of the two-dimensional histogram is presented. In section 3, definition of the exponential entropy is given and the extension of the exponential entropy to the two-dimensional case is presented. A quick description of the TRIBES parameter free Particle Swarm Optimization algorithm is given in section 4. The proposed segmentation algorithm is presented in section 5. Experimental results are discussed in section 6. Finally, we conclude in the last section.

2 Two-Dimensional Histogram

The two-dimensional (2D) histogram [18] of a given image is computed as follows. One calculates the average gray-level value of the neighborhood of each pixel. Let $w(x, y)$ be the averaged image of $f(x, y)$ using a window of size 3x3 defined by:

$$w(x,y)=\left\lfloor \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} f(x+i,y+j)\right\rfloor . \tag{1}$$

where $\lfloor x \rfloor$ denotes the integer part of the number x. 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 (2).

$$h(i,j)=\text{Cardinal}(f(x,y)=i \text{ and } w(x,y)=j)/\text{image size} . \tag{2}$$

The joint probability is given by:

$$p_{ij}=h(i,j) , \tag{3}$$

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

The 2D histogram plane is represented in figure 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 here. A threshold vector is (s, t), where s, for g(x, y), represents the threshold of the average gray-level of the pixel neighborhoods and t, for f(x, y), represents the threshold of the gray level of the pixel. The quadrants containing the background and the 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, a posteriori class probabilities are given by:

$$P_{m-1}\left[a_{n-1},a_n\right]=\sum_{i=s_{n-1}}^{s_n-1}\sum_{j=t_{n-1}}^{t_n-1} p_{ij} \quad .(4)$$

$$P_m\left[a_n,a_{n+1}\right]=\sum_{i=s_n}^{s_{n+1}-1}\sum_{j=t_n}^{t_{n+1}-1} p_{ij} . \tag{5}$$

where $a_n\equiv(s_n,t_n)$; n=1,…,N; m=2,…,N and N is the number of classes.

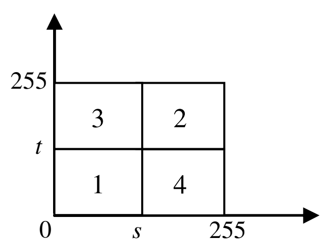

Fig. 1. Two-dimensional histogram plane, where s and t are the thresholds for w(x,y) and f(x,y), respectively

3 Two-Dimensional Exponential Entropy

We define the 2D exponential entropy (2DEE) by:

$$H_\alpha = \left(\sum_i \sum_j p_{ij}^\alpha \right)^{1/(1-\alpha)} , \tag{6}$$

where $\alpha \in \Re$ and $\alpha \neq 1$.
Thus the exponential entropies associated with different image classes' distributions are defined below:

- The 2DEE of the class m-1 can be computed through:

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

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

$$H_\alpha^{(m)}\left[a_n, a_{n+1}\right] = \left(\sum_{i=s_n}^{s_{n+1} - 1} \sum_{j=t_n}^{t_{n+1} - 1} \left(\frac{p_{ij}}{P_m\left[a_n, a_{n+1}\right]} \right)^\alpha \right)^{1/(1-\alpha)} . \tag{8}$$

For the convenience of illustration, two vectors $(s_0, t_0) = (0, 0)$ and $(s_N, t_N) = (255, 255)$ were added, where $t_0 < t_1 < t_2 < ... < t_N$ and $s_0 < s_1 < s_2 < ... < s_N$.
Then the total 2DEE is:

$$H_\alpha^T\left[a_0, ..., a_N\right] = \sum_{i=0}^{N-1} H_\alpha^{(i+1)}\left[a_i, a_{i+1}\right] \tag{9}$$

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

$$H_\alpha^T\left(a^*_{1,...,N-1}\right) = \max\left\{H_\alpha^T(a_{1,...,N-1})\right\} \tag{10}$$

where: $0 < s_1 < s_2 < .... < s_{N-1} < 255$ and $0 < t_1 < t_2 < .... < t_{N-1} < 255$.

In the case of one threshold (N=2) the computational complexity for determining the optimal vector (s^*, t^*) is $O(L^4)$, where L is the total number of gray-levels (usually

256). 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 parameter free PSO algorithm for solving $\arg\max\left\{H_{\alpha}^{T}\left[(s_1,t_1),(s_2,t_2),...,(s_{N-1},t_{N-1})\right]\right\}$ efficiently.

4 Parameter Free PSO Algorithm (TRIBES)

The Particle Swarm Optimization (PSO) is a population based stochastic technique developed by Kennedy and Eberhart (1995). PSO has similarities with the genetic algorithms: a population of potential solutions is used in the search. However there is no evolution operator in PSO. The technique starts with a random initialization of a swarm of particles in the search space. Each particle is modeled by its position in the search space and its velocity. At each time step, all particles adjust their positions and velocities, thus their trajectories, according to their best locations and the location of the best particle of the swarm, in the global version of the algorithm, or of the neighbors, in the local version. Here appears the social behavior of the particles. Indeed, each individual is influenced not only by its own experience but also by the experience of other particles.

TRIBES is an adaptive algorithm of which parameters change according to the swarm behavior. In TRIBES, the user only has to define the objective function and the stopping criterion. The method incorporates rules defining how the structure of the swarm must be modified and also how a given particle must behave, according to the information gradually collected during the optimization process.

However, it must be pointed out that TRIBES, like all competing optimization algorithms, cannot solve with certainty all the problems. Moreover, TRIBES is a stochastic algorithm, thus results given by the algorithm are probabilistic. The aim of TRIBES is to be an algorithm which is efficient enough in most cases and which permits to the users to gain time by avoiding the fitting of parameters.

4.1 Swarm's Structure and Communication

The swarm is structured in different "tribes" of variable size. The space search is simultaneously explored and all tribes exchange results in order to find the global optimum. The algorithm includes two different types of communication: intra-tribe communication and inter-tribes communication, more details about these types of communication are given in [14].

To set rules to modify the swarm's structure, quality qualifiers are defined for each particle and likewise for the tribes. These qualifiers allow defining two rules: removal of a particle and generation of particles. These structural adaptations are not done at all iterations. In practice, if NL is information links number at the moment of the last adaptation, the next adaptation will occur after $NL/2$ iterations. For more details see [14].

4.2 Swarm Evolution

The swarm is initialized by only one particle, that represents a single tribe. A second tribe is created if, at the first iteration, the initial particle does not improve its location. The same process is then applied for the next iterations. The size of the swarm increases until promising areas are found. In other words, the capacity of the swarm to explore increases, but the time between successive adaptations decreases. Then, the swarm has more and more chances to find a good solution between two adaptations. This can be seen as a strategy of displacement. Other implemented strategies are described below.

4.3 Strategies of Displacement

The second strategy to adapt the swarm to the found results is by selecting a different strategy of displacement of each particle according to its recent past. Then the algorithm chooses to call for the best strategy of displacement in order to move the particle to the best possible location, that can be reached.

TRIBES tries to overcome an important problem of metaheuristics: the fitting of parameters. TRIBES frees users of defining parameters by adapting the structure of the swarm and the strategies of displacement of the particles. The particles use their own history and the history of the swarm to decide the way of their move and the organization of the swarm in view of approaching as efficiently as possible the global optimum. Fig. 2 shows a summary of TRIBES process.

1. Initialization of a population of particles with random positions and velocities.
2. Evaluate the objective function for each particle and compute g.
For each individual i, p_i is **initialized** at X_i.
3. Repeat until the stopping criterion is met
- 3.1. Determination of status of all particles
- 3.2. Choice of the displacement strategies
- 3.3. Update the velocities and the positions of the particles.
- 3.4. Evaluate the objective function $H_\alpha^T\left[a_0,...,a_N\right]$ for each individual.
- 3.5. Compute the new p_i and g.
  If $n<NL$
  - Determination of tribes qualities
  - Swarm's adaptations
  - Computation of NL
  End if

4. Show the best solution.

Fig. 2. Principle of TRIBES, where g is the best location reached by the swarm, p_i is the best location for particle I, X_i the position vector of the particle i, NL is the number of information links at the last structure of the swarm, and n is the number of iterations since the last adaptation of the swarm.

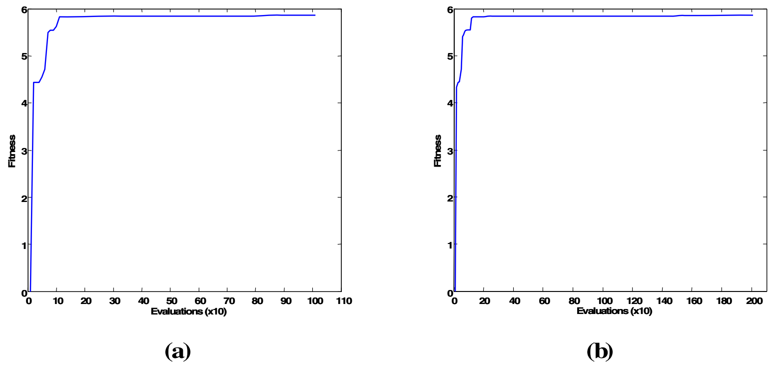

Fig. 3. Example of the evolution of the fitness function in logarithmic scale (for image of Fig. 4 (a)) for : (a) 2000 evaluations, (b) 1000 evaluations. The curves are the result of the averaging of 25 runs.

5 The Proposed Image Segmentation Algorithm

The proposed image segmentation algorithm is based on the maximization of the total 2DEE using TRIBES. The method exploits the particle swarm approach to solve the segmentation problem expressed by (10). The algorithm does not require any special initialization. The number of evaluations was used as stopping criterion. Looking at our experiments (Fig. 3), the value of the fitness function does not increase significantly after 1000 evaluations of the objective function, that explains our decision to fix the maximum number of evaluations of the objective function at 1000.

6 Experimental Results and Discussion

In this section, we discuss the selection of the optimal thresholds and the presentation of some MR images. The performances of the method are compared to those of five other methods, over the segmentation of a synthetic images. The results on MRI segmentation were compared to those provided by the 2D Shannon entropy (2DSE) method [11]. Here, are presented only the results in the case of four and five classes' segmentation.

The value of the optimal threshold depends on the 2DEE order (α). In order to find the optimal value (α^*), the well known uniformity criterion is used. This criterion is given by:

$$U_{(\alpha)} = \left(\frac{1-2N}{M}\right) \sum_{j=0}^{N} \sum_{i \in C_j} (f_i - \mu_i)^2 \Big/ (f_{\max} - f_{\min})^2 \ . \tag{11}$$

where N is the number of thresholds, C_j the jth class, M the number of pixels in the image, f_i the gray level of pixel i, μ_i the mean gray level of pixels in jth class, f_{max} and f_{min} the maximum and the minimum gray levels of pixels in the image, respectively. U has a positive value and lies between 0 and 1. When U is close to 1, the uniformity is very good and vice versa.

6.1 Comparison to Other Methods

We compared the performance of the proposed method to those of five other methods: EM algorithm based method (EM) [20], one method based on valley-emphasis (VE) [21], the well known Otsu method [9], the classical Kapur *et al.* method [9], and Sahoo *et al.* method based on 2D Tsallis entropy (TE) [22]. The comparison is based on synthetic images, noised with different degrees of noise (Fig. 3). To measure these performances, the misclassification error (*ME*) criterion was used [9]. *ME* is defined in terms of correlation of the images with human observation. *ME* is expressed by:

$$ME(\%) = \left(1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|}\right) \times 100 \tag{12}$$

where background and foreground are denoted by B_O and F_O for the original image, and by B_T and F_T for the thresholded image, respectively. In the best case of ideal thresholding, *ME* is equal to 0% and, in the worst case, *ME* value is 100%.

(A) (B) (C) (D)

Fig. 4. (A) Original synthetic image, (B) to (D) noised images

Table 1. Performance evaluation of the proposed method compared to competing methods

Test images of Fig. 4	Segmentation methods						
	Otsu	Kapur	EM	VE	TE	EE2D	
	ME(%)	*ME(%)*	*ME(%)*	*ME(%)*	*ME(%)*	*ME(%)*	α
Image B	0.45	5.61	8.68	0.34	0.64	**0.23**	0.4
Image C	0.88	4.50	12.50	0.63	1.12	**0.56**	0.4
Image D	12.22	4.97	28.87	11.59	12.90	**3.57**	0.6

The quantitative comparison of the results provided by our method and the five other methods, based on segmentation of synthetic images, is presented on table 1. As it can be seen, the proposed method provides better results than the other methods, only VE method provides a better performance in the case of image B.

Table 2. Experimental results for image in Fig. 4 (a)

Number of classes (*N*)	Time (s)	Speed gain factor
3	14.8	$106.\,10^{4}$
4	19.6	687.10^{8}
5	26.7	194.10^{16}

6.2 Examples of Results and Discussion

The obtained results through the application of our segmentation algorithm are illustrated with two brain MRI. Fig. 5 shows the original images and their multilevel classification (segmented) version when N=4 and 5. The results in the case of a sane subject are in Fig. 5 (c) and (e); those in the case of an atrophy pathology are shown in Fig. 5 (d) and (f). Our goal is to detect the different spaces and the white matter surrounding the ventricular space quickly. In order to quantify the performance of the optimization algorithm, we define the speed gain factor, that corresponds to the ratio of the number of the exhaustive search solutions to the evaluation number of the objective function.

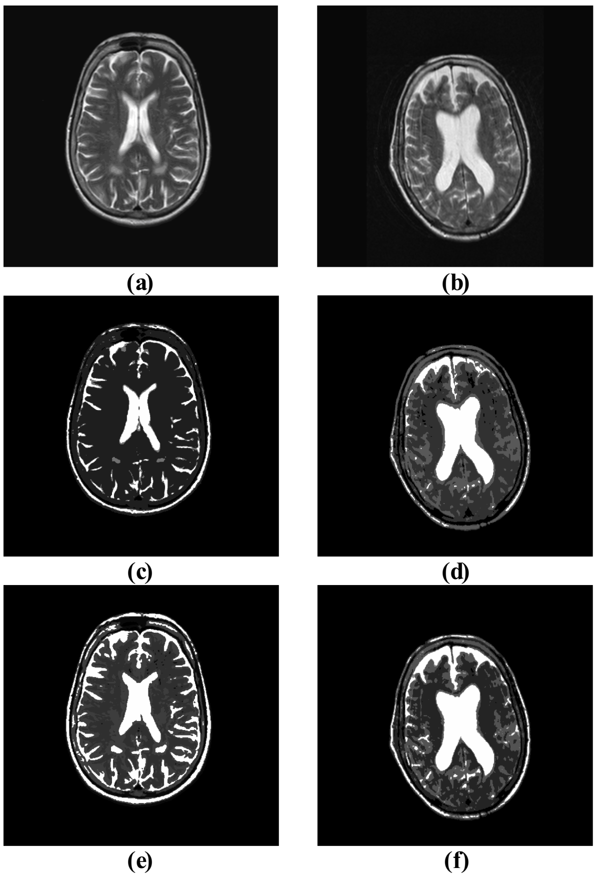

Fig. 5. Segmentation of sane and pathologic MRI. (a) Original image of a sane brain, (b) Original image of a pathologic brain, (c) 4 classes segmented image T=(30, 112, 134) with α=0.1, where T is the threshold vector, (d) 4 classes segmented image T=(49, 88, 180) with α=0.3, (e) 5 classes segmented image T=(10, 47, 61, 97) with α=0.4, (f) 5 classes segmented image T=(36, 82, 138, 153) with α= 0.2.

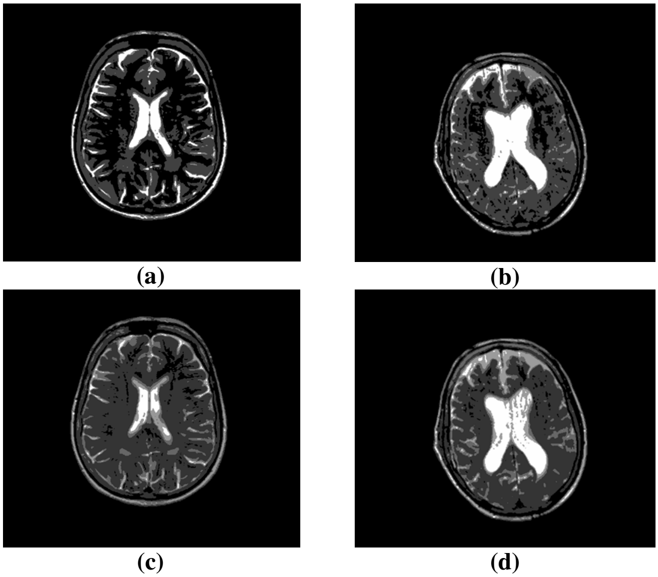

Fig. 6. 2DSE segmentation results. (a) 4 classes segmented image *T*(64, 128, 191), where *T* is the threshold vector, (b) 4 classes segmented image *T*(65, 131, 192), (c) 5 classes segmented image *T*(52, 102, 152, 203), (d) 5 classes segmented image *T*(52, 103, 155, 205).

The number of points for which the criterion function must be evaluated, in the case of an exhaustive search, is $\left(L!/\left((L+1-N)!(N-1)!\right)\right)^2$, where L is the total number of gray-levels (usually 256). For instance, when L=256 and N=2, the number of objective function evaluations is 65536 and, when L= 256 and N=3, it is 32640^2! [14]. Table 1 shows the experimental results obtained on the image of Fig. 5 (a). The speed gain factor and the time values show effectiveness of TRIBES algorithm and confirm that our method is fast compared to those in [1] to [7], where the result is obtained after more than 120s [7]. As it can be seen, in table 2, the speed gain factor increases by a factor higher than 10^4 when one class is added to the problem.

Fig. 6 shows the results obtained via the application of 2DSE. One notices that the results provided by our method are more homogeneous than those provided by 2DSE. This can be seen clearly, for instance, through the comparison of the detected white matter, between Fig. 5 (f) and Fig. 6 (d).

7 Conclusion

In this paper, we proposed a new fast approach to find the optimal thresholds, based on 2DEE to avoid the problems related to the use of Shannon entropy to segment images. We also proposed to use a parameter free PSO algorithm and our experiments proved that TRIBES can be used as a black box optimization tool to solve a segmentation problem. The use of TRIBES allows to avoid the parameter tuning step that requires a minimum of experience about the used algorithm.

It is clearly seen from the experimental results that the presented method is more efficient than the classical 2DSE and using TRIBES allows to obtain good results quickly. However, the use of the method to segment other kinds of images does not provide good segmentation results when the images are strongly noised. In the work in progress we use a multiobjective optimization based on parameter free PSO in order to add information to segment noised images.

Acknowledgements. The authors would like to thank Dr. Raphael Blanc, from Neurosurgery service of "Centre Hospitalier Universitaire de Mondor", Créteil, France, for supplying the MR images.

References

1. Drapaca, C.S., Cardenas, V., Studholme, C.: Segmentation of tissue boundary evolution from brain MR image sequences using multi-phase level sets. Computer Vision and Image Understanding 100, 312–329 (2005)
2. Qiao, Y., Hu, Q., Qian, G., Luo, S., Nowinski, W.L.: Thresholding based on variance and intensity contrast. Pattern Recognition 40, 596–608 (2007)
3. Warfield, S.K., Kaus, M., Jolesz, F.A., Kikinis, R.: Adaptive template moderate spatially varying statistical classification. Medical Image analysis 4, 43–55 (2000)
4. Murgasova, M., Dyet, L., Edwards, D., Rutherford, M., Hajnal, J.V., Rueckert, D.: Segmentation of Brain MRI in Young Children. In: 9th Int. Conf. MICCAI Copenhagen, pp. 687–694 (2006)
5. Song, Z., Tustison, N., Avants, B., Gee, J.C.: Integrated Graph Cuts for Brain MRI Segmentation. In: 9th Int. Conf. MICCAI, Copenhagen, Denmark, pp. 831–838 (2006)
6. Kamber, M., Shinghal, R., Collins, D.L., Francis, G.S., Evans, A.C.: Model-based segmentation of multiple sclerosis lesions in magnetic resonance brain images. IEEE Trans. on Med. Imaging 14, 442–453 (2000)
7. Cocosco, C.A., Zijdenbos, A.P., Evans, A.C.: A fully automatic and robust brain MRI Tissue Classification. Medical Image Analysis 7, 513–527 (2003)
8. Zografos, K., Nadarajah, S.: Survival Exponential Entropies. IEEE Trans. on Information Theory 51, 1239–1246 (2005)
9. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging 13, 146–165 (2004)
10. Tao, W., Tian, J., Liu, J.: Image segmentation by three level thresholding based on maximum fuzzy entropy and genetic algorithm. Pattern Recognition Letters 24, 3069–3078 (2004)
11. Peng-Yeng, Y.: Multilevel minimum cross entropy threshold selection based on particle swarm optimization. Applied Mathematics and Computation 184, 503–513 (2007)
12. Zahara, E., Fan, S.S., Tsai, D.: Optimal multi-thresholding using a hybrid optimisation approach. Pattern Recognition Letters 26, 1082–1095 (2004)
13. Synder, W., Bilbro, G.: Optimal thresholding: A new approach. Pattern Recognition Letters 11, 803–810 (1990)
14. Clerc, M.: TRIBES - Un exemple d'optimisation par essaim particulaire sans paramètres de contrôle. In: OEP 2003, Paris (2003)
15. Ye, X.F., Zhang, W.J., Yang, Z.L.: Adaptive Particle Swarm Optimization on Individual Level. Int. Conf. on Signal Processing (ICSP), Beijing, China, 1215–1218 (2002)

16. Zhang, W., Liu, Y., Clerc, M.: An adaptive PSO algorithm for real power optimization. In: APSCOM (Advances in Power System Control Operation and Management), S6: Application of Artificial Intelligence Technique (part I), Hong Kong, pp. 302–307 (2003)
17. Yasuda, K., Iwasaki., N.: Adaptive particle swarm optimization using velocity information of swarm. In: IEEE Conference on System, Man and Cybernetics, The Hague, Netherlands, pp. 3475–3481 (2004)
18. Nakib, A., Oulhadj, H., Siarry, P.: Microscopic image segmentation based on two-dimensional exponential entropy with hybrid microcanonical annealing. In: Proceedings of Int. Conf. IAPR- MVA2007, Tokyo, pp. 420–423 (2007)
19. Nakib, A., Oulhadj, H., Siarry, P.: Image histogram thresholding based on multiobjective optimization. Signal processing 87, 2516–2534 (2007)
20. Bazi, Y., Bruzzone, L., Melgani, F.: Image thresholding based on the EM algorithm and the generalized Gaussian distribution. Patter Recognition Journal 40, 619–634 (2007)
21. Ng, H.: Automatic thresholding for defect detection. Pattern Recognition Letters 27, 1644–1649 (2006)
22. Sahoo, K.P., Arora, G.: Image thresholding using two dimensional Tsallis-Havrda-Charvat entropy. Pattern Recognition Letters 27, 520–528 (2006)