Image Segmentation Using an Evolutionary Method Based on Allostatic Mechanisms

Valentín Osuna-Enciso, Virgilio Zúñiga, Diego Oliva, Erik Cuevas and Humberto Sossa

Abstract In image analysis, segmentation is considered one of the most important steps. Segmentation by searching threshold values assumes that objects in a digital image can be modeled through distinct gray level distributions. In this chapter it is proposed the use of a bio-inspired algorithm, called Allostatic Optimisation (AO), to solve the multi threshold segmentation problem. Our approach considers that an histogram can be approximated by a mixture of Cauchy functions, whose parameters are evolved by AO. The contributions of this chapter are on three fronts, by using: a Cauchy mixture to model the original histogram of digital images, the Hellinger distance as an objective function, and AO algorithm. In order to illustrate the proficiency and robustness of the proposed approach, it has been compared to the well-known Otsu method, over several standard benchmark images.

V. Osuna-Enciso (✉) · V. Zúñiga
Departamento de Ingenierías, CUTonalá, Universidad de Guadalajara,
Tonalá, Jalisco, Mexico
e-mail: Valentin.Osuna@cutonala.udg.mx

V. Zúñiga
e-mail: Virgilio.Zuniga@cutonala.udg.mx

D. Oliva
Departamento de Ciencias Computacionales, Tecnológico de Monterrey,
Campus Guadalajara, Guadalajara, Jalisco, Mexico
e-mail: Diego.Oliva@itesm.mx

E. Cuevas
Departamento de Ciencias Computacionales, CUCEI, Universidad de Guadalajara,
Guadalajara, Jalisco, Mexico
e-mail: Erik.Cuevas@cucei.udg.mx

H. Sossa
Instituto Politécnico Nacional-CIC, México D.F., Mexico
e-mail: HSossa@cic.ipn.mx

A.I. Awad and M. Hassaballah (eds.), *Image Feature Detectors and Descriptors*,
Studies in Computational Intelligence 630, DOI 10.1007/978-3-319-28854-3_10

1 Introduction

Image segmentation is considered as an important operation for meaningful analysis and interpretation of images acquired. In particular, image segmentation aims to group pixels within meaningful regions. Commonly, gray levels belonging to an object, are substantially different from those featuring by other objects or by the background. Segmentation is typically conducted considering two main criteria: similarity of image regions and discontinuity between adjacent disjoint regions [1, 2]. Among the segmentation approaches based on similarity, thresholding is considered the simplest technique [3, 4]. It involves the basic assumption that the objects and the background in the digital image have distinct gray level distributions. Under such assumption, the gray level histogram contains two or more distinct peaks and threshold values separating them that can be obtained. Therefore, segmentation is performed by assigning regions having gray levels below the threshold to the background, and assigning those regions having gray levels above the threshold to the objects, or vice versa. Segmentation by thresholding has been used in several areas where a correct separation of the objects in images is a vital step to perform fully automatic vision systems for detection and classification such as medical imaging [5–11], aviation [12], spacecraft imagery [13] and nondestructive tests [14], among many other applications. Several thresholding segmentation approaches have been reported in the literature [15–20], being the most popular the Otsu [21] method.

In statistics, the Gaussian distribution [22] is a standard modeling tool which satisfies the central limit theorem. A Gaussian distribution assumes that the probability of any occurring value falls off rapidly as it is moved further away from the central value. However, several problems, such as those that involve the presence of several outliers in the population, cannot be appropriately modeled under such assumption. Similar to the Gaussian distribution, the Cauchy distribution [23] is a symmetric bell-shaped density function but with a greater probability mass in the tails. This fact allows that the probabilities of points with large deviations from the central value, such as outliers, do not drop off as precipitously as it is obtained by the Gaussian distribution [24]. Although the Cauchy distribution possesses better modeling capabilities (in presence of outlier data) than other distributions, it presents serious difficulties in estimating its behavior parameters [25]. The capacity of the Cauchy distributions to model complex process has been demonstrated in several engineering applications such as impulsive noise cancellation [26], image denoising algorithms [27, 28], among others.

In this work, the segmentation algorithm is based on a parametric model which groups a mixture of several Cauchy density functions (Cauchy mixture, CM). CM involves the model selection, i.e., to determine the number of components in the mixture, and the estimation of the parameters of each component in the mixture that better adjust the statistical model. In general, computing the parameters of each Cauchy function is considered a difficult task, sensible to the initialization [29–31]

and full of possible singularities [32]. In order to calculate such parameters, several methods have been proposed in the literature [33–35], presenting most of them flaws such as high computational overhead and sub-optimal values as a result of getting trapped in a local minimum. In the proposed approach, the parameter estimation of the CM has been faced as an optimization problem.

On the other hand, an impressive growth in the field of biologically inspired evolutionary algorithms for search and optimization has emerged during the last decade. Several bio-inspired algorithms have been proposed in the literature. Some examples include methods such as the Evolutionary Algorithm (EA) proposed by Fogel et al. [36], De Jong [37], the Genetic Algorithms (GA) proposed by Holland [38], the Artificial Immune System proposed by De Castro and Von Zuben [39] the Particle Swarm Optimization (PSO) method proposed by Kennedy and Eberhart [40] and the Artificial Bee Colony (ABC) proposed by Karaboga [41].

The interesting and complex behavior of biological organs from the human body have fascinated and attracted the interest of researchers for many years. Biologists have studied these phenomena to model organ operations, and engineers applied these models as a framework for solving complex real-world problems. An important biological phenomenon is Allostasis which explains how the modifications of specialized organ conditions inside the body allow achieving stability when an unbalance health condition is presented. Therefore, if a body decompensation happens, according to the allostatic mechanisms, several set points compound by blood pressure, oxygen tension and others indexes are proved in order to get a stability state. Such set points are generated by using different specialized mechanisms.

In this chapter, a multi-thresholding segmentation algorithm based on a new evolutionary algorithm called Allostatic Optimization (AO) is presented. In the approach, the segmentation process is considered as an optimization problem by approximating the 1-D histogram of a given image in terms of a Cauchy mixture model, whose parameters are calculated through the AO algorithm. In AO, the searcher agents emulate different body conditions which interact to each other by using operators based on the biological principles of the allostasis mechanism. The proposed approach encodes the parameters of the CM as an individual. An objective function by using the Hellinger distance evaluates the matching quality between the CM candidate and the original histogram. Guided by the values of this objective function, the set of encoded candidate mixtures are evolved using the operators defined by AO so that they can fit into the original histogram. In order to illustrate the proficiency and robustness of the proposed approach, it has been compared to the well-known Otsu method. The comparison examines several standard benchmark images that are commonly considered within the literature. Experimental results show a high performance of the proposed method for searching appropriate threshold values in terms of accuracy and robustness.

2 Allostatic Optimization

2.1 Allostasis

Organ systems (OS) into the body are composed of organ groups working in coordinated ways in order to maintain vital functions [42]; the human being has eleven of such systems, and each one is responsible to perform several specialized tasks as can bee seen in Table 1.

Even though there are other forms to describe the body organization, in the explanation given next it is only considered the organization based on OS. Communication among cells belonging to different OS with the brain it is achieved by means of two systems: the nervous and the endocrine, who are responsible of the coordination among OS for regulation of each essential function inside the body. Once the brain detects some external or internal change (stress, pollution, social status, disease, etc.), it determines if the stability of the body is compromised, in whose case it uses those channels to communicate with the adequate OS, trying to cope with such perturbation in order to get again the stability of the body. Chemical messages from the endocrine system are sent through hormonal substances, which are in charge of triggering or inhibiting responses from several tissues through target cells (who usually belong to several OS), whereas the nervous systems mainly uses electrical messages, activated

Table 1 Organ systems in the body

No.	Organ system	Task(s)	Some elements
1	Integumentary	External protection, sensory receptors	Skin, hair
2	Skeletal	Internal protection of tissues and organs	Bones
3	Muscular	External and internal movements	Muscles
4	Circulatory	Carrying vital substances	Heart, blood vessels
5	Respiratory	Control and regulation of the breathing process	Lung, nose
6	Digestive	Turn substances in energy to cells	Stomach, intestines
7	Excretory	Disposal of wastes	Kidney
8	Lymphatic	Internal protection against toxins and substances	T-cells, B-cells
9	Reproductive	Sexual reproduction	Testicles, ovaries
10	Nervous	First communication center	Neurons, Nerves
11	Endocrine	Second communication center	Pituitary, hypothalamus

through neurotransmitters, also hormones, or a combination of them. Both chemical and electrical messengers are referred simply as mediators. One of the main theories that explains how the body achieves the stability of the body, or health, as well as the coordination with the different OS, it is called 'homeostasis', which means 'to maintain stability (of a system, organ, OS, condition, health, etc.) through constancy (of a determined mechanism's set point (SP))' [43–51]. Among the several types of homeostasis are counted: of glucose, intestinal, of immunologic resources, of lipids, of cholesterol, of zinc, of energy, pulmonary, of epidermis, of blood pressure, etc [42].

Let's consider the blood pressure of a person being sit: at such a moment, according to the homeostatic theory, one mechanism maintains the blood pressure into stability, by keeping up a SP inside of a narrow range (around of 10 beats/min in healthy people). As the hypothetical person is standing up, there is a difference between the actual SP and the required SP of blood pressure of a standing person, due simply to the gravity force; in other words, there is no stability of blood pressure. As the brain detects this instability, also starts sending signals (Mediators) to the related mechanism through the communication channels (Nervous and Endocrine systems), in order to activate the adequate response from the tissues involved (to increase the heart bumping, in this example changed to a new SP).

The problem being solved has only one liberty degree. In other words, it is required to find only one SP, and this must be inside a narrow margin. The homeostatic model has demonstrated its utility in medicine [43, 44]; however, in some cases that model is not enough to explain neither complex behaviors of OS in the body, nor even some disease patterns, and therefore medical prescriptions homeostasis-based tend to fail. By considering such a problem it was proposed an alternative model who is called allostasis, that means 'to maintain stability (of a system, organ, OS, condition, health, etc.) through change (of several mechanisms)' [51]; in this case, and different to the homeostasis model, are taken in account several SPs of mechanisms and the non-linear relationships among mediators, OS and the brain. A simple example of allostasis is shown in Fig. 1.

The fundamental difference between this model and the homeostasis is the existence of several SPs in several mechanisms involved in returning to stability. As can be seen from Fig. 1, some SPs are increased whereas other are decreased; moreover, there are relationships among mechanisms, mediators, other mechanisms, as well as OS, which are not fully understood by scientists working at medical areas [51]. Prior to the explanation of the proposed computational algorithm, it is important to consider both a standard vocabulary and considerations. For instance, we will consider that the communication groups of OS are Group A. In the computational algorithm we consider only three groups of mediators, even though that could exist more in natural allostasis; those groups are called Groups B1, B2 and B3, or simply, Group B. Both groups, A and B, are responsible of coordination between brain and mechanisms that directly changes the appropriated SPs that could cope with the perturbation found; also, we argue that those groups contain different versions of SPs

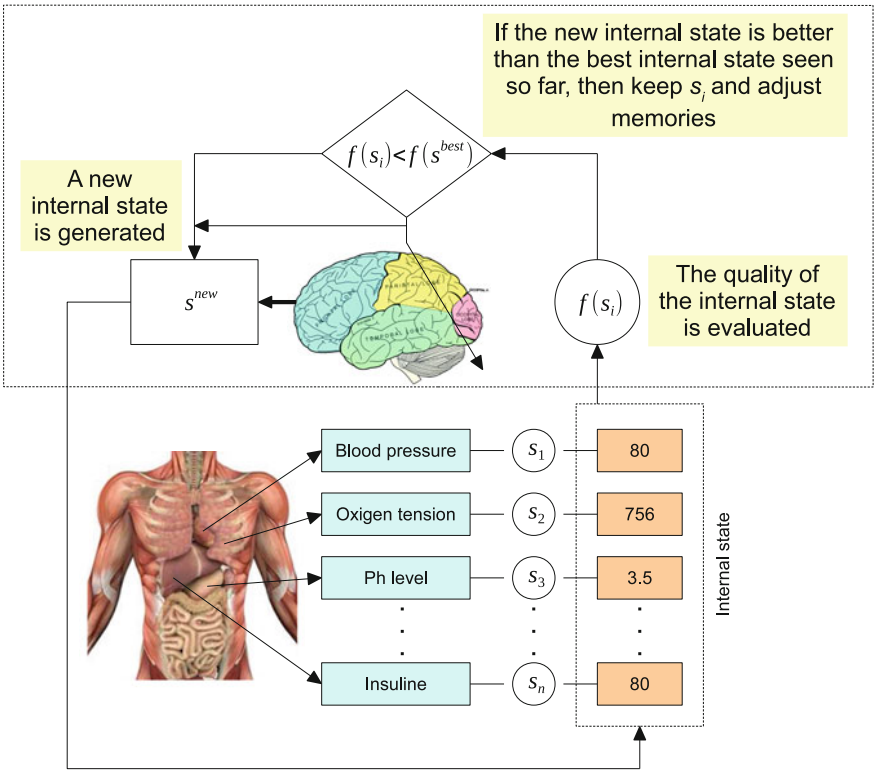

Fig. 1 Allostasis, a simple illustration scheme

(in hormonal as well as other kind of signaling means) historically used. In other words, we say that Groups A, B, and mechanisms, are different forms of SPs used in the past. By considering the aforementioned, in allostasis, the generation of new SPs is done by using the SPs historically used [49]. Whereas SPs of Group A do not suffer collective changes, those contained in Group B are constantly modified by collective operations. In order to generate new SPs, allostasis considers several procedures, being the main the so called 'combination' [52], which combines information of Groups A, B, and random variations. Once a new SP is generated, it is evaluated its capacity to reach a stable state, and whether the new SP improves the stability provided by the actual SP, a collective change is carried out over all elements contained in Group B [47, 48, 50].

2.2 Allostatic Optimization Algorithm

The computational approach of allostasis is called Allostatic Optimization (AO), which implements operations that resembles the interaction rules modeled by the mechanism of allostasis. In the algorithm, each *candidate solution* within the search space represents a *SP* vector, whereas the *fitness* value equals to a *degree of stability (or allostasis)* of each SP. The *population of candidate solutions* correspond to *stored SPs* in natural allostasis. The *optimum candidate solution* corresponds to the *whole stability*, or simply, *stability*. Finally, we consider a perturbation as a function of the difference between the allostasis and the stability; such a function it is called *err* in this chapter.

Also, the AO defines several operators, as the combination operation, which is considered the main operator and is applied over all individuals of the population. Other operators called collective (group operators: B1, B2 and B3) are also implemented in AO, and they affect only a group of elements. Following the biological model of allostasis, the AO approach divides the entire population in four different groups: Group A, group B1, group B2 and group B3. The elements of group A are only modified by the combination operator whereas the elements of groups B1, B2 and B3 are affected by the combination operator and other collective operations. Even though in the natural process each group could have different sizes, in the computational approach we consider the size of each group as one fourth of the entire population. Thus, the population size must be selected in such a way that it can be entirely divided by four (20, 40, etc.).

2.3 Description of the AO Algorithm

The AO algorithm starts by initializing the population randomly (candidate random solutions or SPs) and later, the evolutionary process acts as follows: The combination operator is applied to the first individual (SP) of the population, obtaining in such a way a new individual. Whether the new individual is better than the original one according to their allostasis (fitness value), the original individual is replaced by the new one whereas the groups B1, B2 and B3 are modified used collective operators. On the other hand, if the original individual is better than the new individual, no changes are executed to the population. An iteration is completed when the combination operator has been applied to the last individual. This procedure is applied until a termination criterion is met (i.e. the iteration number NI). Following the evolution process of AO, the following operators are employed:

1. Initialization.
2. Combination.
3. Collective B1.
4. Collective B2.
5. Collective B3.
6. Update the best element.

and the pseudo code of the proposal is

```
AO algorithm
1  Initialize S and determine its allostasis,
   divide S in groups A and B
2  while (criterion){
   for i=1 to Ns {
      Generate new individual snew
      by using the combination operator
      if f(snew)<f(si){
         if f(snew)<f(sbest){
            Calculate e and m
            Modify B1 by using collective operator B1
            Modify B2 by using collective operator B2
            Modify B3 by using collective operator B3
            sbest=snew; f(sbest)=f(snew);
         }
         si=snew; f(si)=f(snew);
      }
   }
}
```

Next, each operator is defined.

Initialization

In the first part, the algorithm initializes a population S of Ns set point vectors ($S = \{s_1, s_2, \ldots, s_{Ns}\}$), where each set point (SP) s_i is a D-dimensional vector containing parameter values to be optimized. Such values are randomly and uniformly distributed between the pre-specified lower initial parameter bound s_j^{low} and the upper initial parameter bound s_j^{high}:

$$s_{ij} = s_j^{low} + rand\,() \cdot \left(s_j^{high} - s_j^{low}\right); i = 1, \ldots, Ns;\ j = 1, \ldots, D \qquad (1)$$

with i and j being the individual and parameter indexes, respectively. Hence, s_{ij} is the jth parameter of the ith individual. After initialization of SPs, it is found the best individual from the population, e.g.

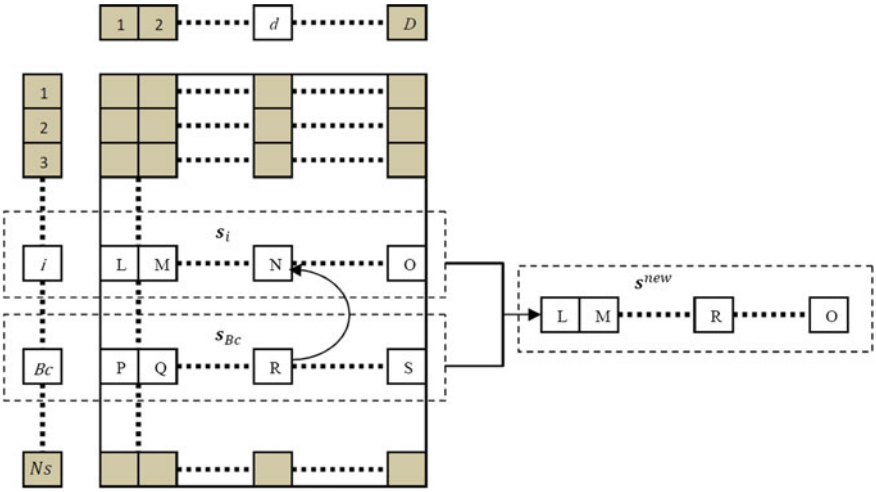

Fig. 2 The combination operator

$$s^{best} \in \left\{ S | f\left(s^{best}\right) = min(f\left(s_1\right), f\left(s_2\right), \ldots, f(s_{Ns})) \right\} \tag{2}$$

where $f(.)$ represents the cost function.

Combination

In allostasis, this operation combines each SP in the population, with information provided by other SPs. In this work, such effect is simulated by using a single operation of mutation, by replacing information of an original individual s_i with information extracted from other s_{Bc}, obtaining in such a way a new individual s^{new}, who combines information from both individuals. In order to implement this operator, two different integers are randomly generated, Bc inside the number of SPs $(1, \ldots, Ns)$ and d inside the dimension number $(1, \ldots, D)$. The combination takes place substituting the element s_{di} from the original s_i with the element s_{dBc} from the element s_{Bc}. Therefore, the only difference between s_i and s_{Bc} is the element in the position d. Figure 2 shows graphically the combination operation. Once the new individual s^{new} is generated by using the combination operator, it is compared whether such individual is better than the original one s_i and also the best found to far s^{best}, according to their fitness values. If s^{new} is better, the elements s_i as well as s^{best} are replaced by s^{new}, whereas the groups B1, B2 and B3 are modified used the collective operators. However, if s_i is better than s^{new}, no changes are executed to the population.

Collective B1

This operator modifies only the elements of group B1, namely SPs from $(Ns/4) + 1$ to $2{\cdot}(Ns/4)$. In the allostasis mechanism, SPs from Group B1 are substituted by similar versions of the average answer produced by the entire set of SPs. In the AO approach, the average answer $a = \{a_1, \ldots, a_D\}$ is computed as:

$$a_j = \left(\frac{1}{Ns}\right) \cdot \sum_{i=1}^{Ns} s_{ij};\ j = 1, \ldots, D \tag{3}$$

The modification, applied to each element, depends on the existent difference between s^{new} and s^{best}. Such relationship, defined as m, is calculated by using:

$$m = \psi \left[1.1 - \frac{1}{e^{\psi \cdot err}}\right] \tag{4}$$

where $\psi \in [0.01, 1.5]$ and $err = \left(\left(f\left(s^{new}\right) - f\left(s^{best}\right)\right) / \left(f\left(s^{new}\right) + f\left(s^{best}\right)\right)\right)$ Therefore, the SPs of group B1 are updated according to:

$$s_{g1,j} = a_j - m + 2{\cdot}m{\cdot}rand\,() \tag{5}$$

where $j \in \{1, \ldots, D\}$, $g1 \in \left\{\left(\frac{Ns}{4}\right) + 1, \left(\frac{Ns}{4}\right) + 2, \ldots, 2{\cdot}\left(\frac{Ns}{4}\right)\right\}$ and $rand(\,)$ is a number randomly generated between 0 and 1.

Collective B2

According to the allostasis mechanism, elements of group B2 are replaced by SPs randomly generated inside the average answer. Such effect is simulated modifying the elements of group B2 according to the following model:

$$s_{g2,j} = a_j{\cdot}rand\,() \tag{6}$$

where $g2 \in \left\{2{\cdot}\left(\frac{Ns}{4}\right) + 1, 2{\cdot}\left(\frac{Ns}{4}\right) + 2, \ldots, 3{\cdot}\left(\frac{Ns}{4}\right)\right\}$.

Collective B3

Following the allostasis model, SPs of Group B3 are replaced by those who have demonstrated to be successful when a similar decompensation has happened. Such a behavior is emulated producing perturbed versions of the best SP $s^{best} = \left\{s_1^{best}, s_2^{best}, \ldots, s_D^{best}\right\}$ found so-far. Thus, the elements of group B3 are modified by using:

$$s_{g3,j} = s_j^{best} - m + (2{\cdot}m{\cdot}rand\,()) \tag{7}$$

where $g3 \in \left\{3{\cdot}\left(\frac{Ns}{4}\right) + 1, 3{\cdot}\left(\frac{Ns}{4}\right) + 2, \ldots, Ns\right\}$.

Update the Best Element

In order to update de best SP s^{best} seen so-far, the best found individual from the current population $s^{best,k}$ is compared with the best individual $s^{best,k-1}$ of the last generation. If $s^{best,k}$ is better than $s^{best,k-1}$ according to their fitness values, s^{best} is updated with $s^{best,k}$, otherwise s^{best} remains without changes. Therefore, s^{best} stores the best historical SP found so-far.

3 Parametrical Model

3.1 Histogram Approximation by Using a CM

Let consider an image holding L gray levels $[0, \ldots, L-1]$ whose distribution is displayed within the histogram $h(g)$. In order to simplify the description, the histogram is normalized just as a probability distribution function, yielding:

$$h(g) = \frac{n_g}{N}, h(g) \geq 0, N = \sum_{g=0}^{L-1} n_g, \text{ and } \sum_{g=0}^{L-1} h(g) = 1, \tag{8}$$

where n_g specifies the number of pixels with gray level g, whereas N represents the total number of pixels contained in the image. The image histogram can thus be approximated by a CM of the form:

$$p(x) = \sum_{i=1}^{K} P_i \cdot p_i(x) = \sum_{i=1}^{K} P_i \left[\frac{{\gamma_i}^2}{(x-\rho_i)^2 + {\gamma_i}^2} \right] \tag{9}$$

where P_i is the a priori probability of class i, $p_i(x)$ is the probability distribution of gray-level random variable x in class i, ρ_i and γ_i are the location and the scale parameter of the ith Cauchy distribution and K is the number of classes contained in the image. In addition, the constraint $\sum_{i=1}^{K} P_i = 1$ must be satisfied.

In the proposed approach, the parameters $(P_i, \rho_i, \gamma_i, i = 1, \ldots, K)$ of the CM are encoded in an individual (a possible candidate solution). In order to correctly evaluate the matching quality between a candidate CM and the original histogram, the Hellinger distance E [53] is used. Such distance is defined as follows:

$$E = \sqrt{\sum_{j=0}^{L} \left[\sqrt{p(x_j)} - \sqrt{h(x_j)} \right]^2} \tag{10}$$

where $p(x_j)$ is the probability defined by the candidate CM, in the gray level point x_j whereas $h(x_j)$ represents its respective histogram value. Therefore, Eq. 10 is the objective function used by the AO algorithm to assess the quality of each individual.

Once obtained the best histogram approximation by a CM, the next step is to determine the optimal threshold values. At first, the location parameters are organized such as $\rho_1 < \rho_2 < \cdots < \rho_K$; then, the threshold values are calculated by estimating the overall probability error for two adjacent Cauchy functions:

$$E(T_i) = P_{i+1} \cdot E_1(T_i) + P_i \cdot E_2(T_i), \quad i = 1, 2, \ldots, K-1 \tag{11}$$

considering

$$E_1(T_i) = \int_{-\infty}^{T_i} p_{i+1}(x)dx, \quad \text{and} \quad E_2(T_i) = \int_{T_i}^{\infty} p_i(x)dx \tag{12}$$

$E_1(T_i)$ is the probability of mistakenly classifying the pixels in the $(i+1)$th class belonging to the ith class, while $E_2(T_h)$ is the probability of erroneously classifying the pixels in the ith class belonging to the $(i+1)$th class. P_i's are the a-priori probabilities within the combined probability density function, and T_i is the threshold value between the ith and the $(i+1)$th classes. The T_i value is chosen as to minimize the error $E(T_i)$. By differentiating $E(T_i)$ with respect to T_i and equating the result to zero, it is possible to use the following equation to define the optimum threshold value T_i:

$$AT_i^2 + BT_i + C = 0 \tag{13}$$

where

$$A = \gamma_i^2 - \gamma_{i+1}^2 \tag{14}$$

$$B = 2 \cdot (\rho_i \gamma_{i+1}^2 - \rho_{i+1} \gamma_i^2) \tag{15}$$

$$C = (\gamma_i \rho_{i+1})^2 - (\gamma_{i+1} \rho_i)^2 + 2 \cdot (\gamma_i \gamma_{i+1})^2 \cdot \ln\left(\frac{\gamma_{i+1} P_i}{\gamma_i P_{i+1}}\right) \tag{16}$$

From Eq. 13, it is only considered the solution whose value is positive and falls inside the valid interval. Figure 3 shows the determination process of threshold points, considering only two consecutive Cauchy functions.

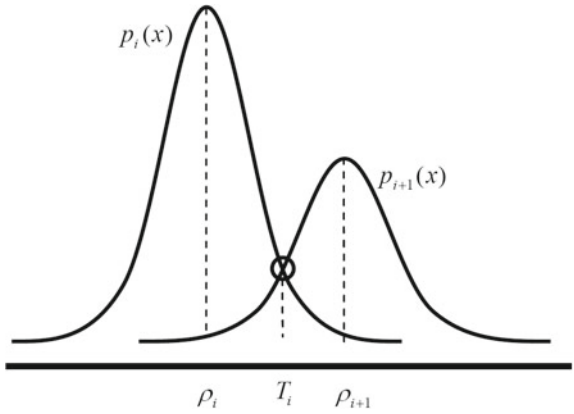

Fig. 3 Determination of the threshold points

3.2 Otsu's Method

This method is a nonparametric technique for thresholding segmentation proposed by Otsu [21] that employs the maximum variance value of the different classes as a criterion to segment the image. In this approach the image histogram $h(g)$ is divided in m threshold values $T = \{T_1, \ldots, T_{m-1}\}$, considering $T_0 = 0$ and $T_m = L$. Each i-partition of m is defined as:

$$C_i = \{g | g \in (1, \ldots L-1),\ T_{i-1} < g < T_i\}, \quad i = 1, \ldots, m \tag{17}$$

Such values are calculated as follows:

$$q_1 = \sum_{i=0}^{T_1} h(g_i), \quad \mu_1 = \sum_{i=0}^{T_1} \frac{h(g_i) \cdot i}{q_1}, \quad \sigma_1^2 = \sum_{i=0}^{T_1} \frac{(i-\mu_1)^2 \cdot h(g_i)}{q_1} \tag{18}$$

$$q_i = \sum_{T_i+1}^{T_{i+1}} h(g_i), \quad \mu_i = \sum_{T_i+1}^{T_{i+1}} \frac{h(g_i) \cdot i}{q_i}, \quad \sigma_i^2 = \sum_{T_i+1}^{T_{i+1}} \frac{(i-\mu_i)^2 \cdot h(g_i)}{q_i} \tag{19}$$

where $i = 1, \ldots, m-1$. Therefore, the variance for the K-class is computed following the model:

$$\sigma_{WC}^2 = \sum_{j=1}^{K} q_j \cdot \sigma_j^2 \tag{20}$$

In order to find the threshold values, Eq. 21 must be minimized:

$$\sigma_{WC}^2(T_1^*, \ldots, T_K^*) = \min_{0 \le T_1 \le \cdots \le T_K \le L-1} \sigma_{WC}^2(T_1^*, \ldots, T_K^*) \tag{21}$$

The Otsu method is considered the most popular segmentation algorithm with respectable results. Nevertheless, if the number of threshold values increases, the number of function evaluations increases. Such fact is considered its main drawback. Due to its wide popularity, the Otsu algorithm is used for comparing the performance of the approach proposed in this chapter.

4 Experimental Results

In the proposed approach, the parameters of the CM are encoded as an individual. An objective function by using the Hellinger distance evaluates the matching quality between the CM candidate (individual) and the original histogram. Guided by the values of this objective function, the set of encoded candidate mixtures are evolved using the operators defined by AO so that they can fit into the original histogram.

In this section, several experiments have been conducted considering different classes. Since each Cauchy function is defined by three parameters (P, ρ, γ), each individual l will have $3xK$ dimensions, if K different classes would be considered $[P_1^l, \rho_1^l, \gamma_1^l, \ldots, P_K^l, \rho_K^l, \gamma_K^l]$. Table 2 shows the general parameters utilized by AO. All the experiments are performed on a desktop computer with Intel® Core i7-2600 3.4 GHz, 8 GB of RAM and programmed in Matlab® 7.13.0.

Table 2 Parameters used in AO

Parameter	Value	Observations
L	256	Number of gray levels
N_p	90	Population size
N_{max}	200	Maximum number of iterations
x_l^{high}	$L-1$	Higher limits of candidate l
x_l^{low}	0	Lower limits of candidate l
K	[2, 7]	Number of classes to find
T	$K-1$	Number of thresholds to find
Ψ	0.03	Tuning parameter of AO

In order to demonstrate the performance of the proposed algorithm, several images extracted from the Berkeley and the All-IDB databases [54, 55] have been used. Figures 4, 5, 6, 7, 8 and 9 present the experimental results after applying the

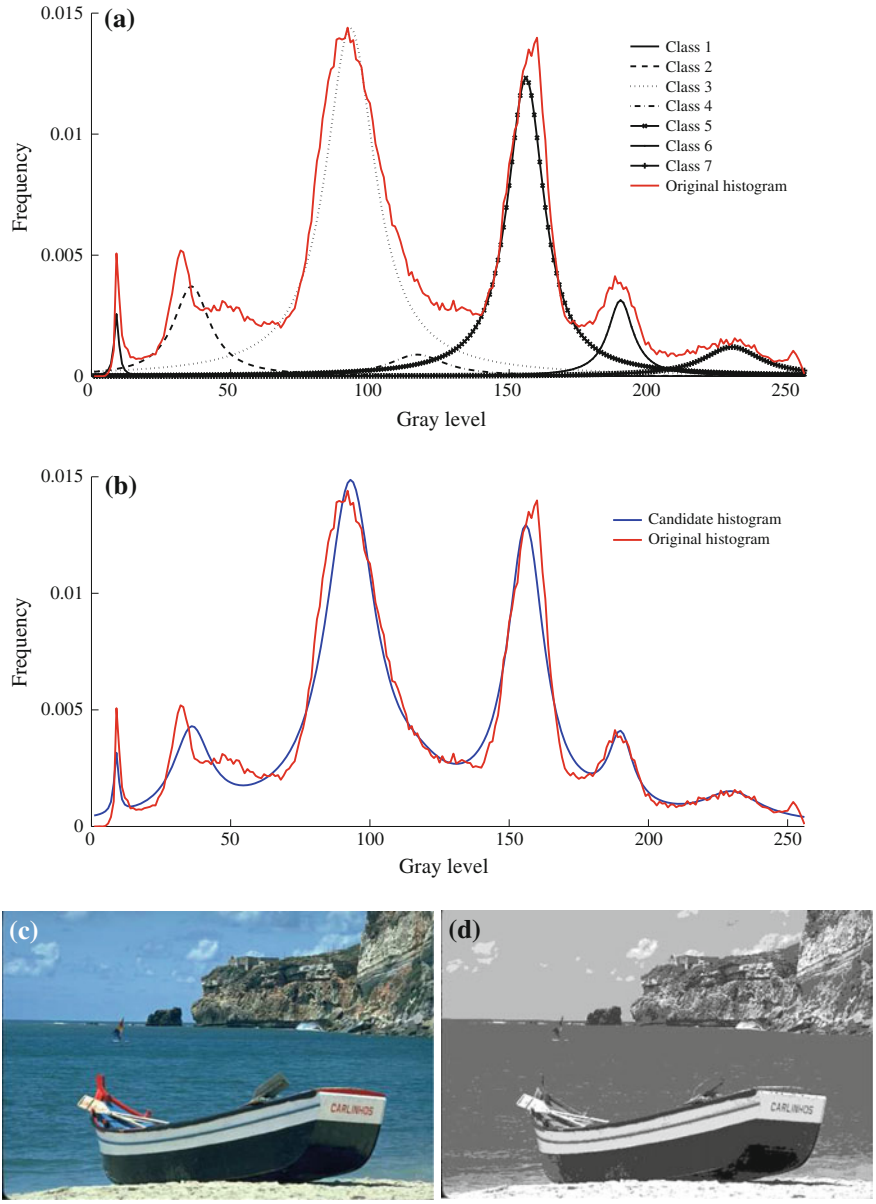

Fig. 4 Image 233, **a** class distribution with seven classes ($K = 7$), **b** approximation considering seven classes, **c** original image, **d** segmented image

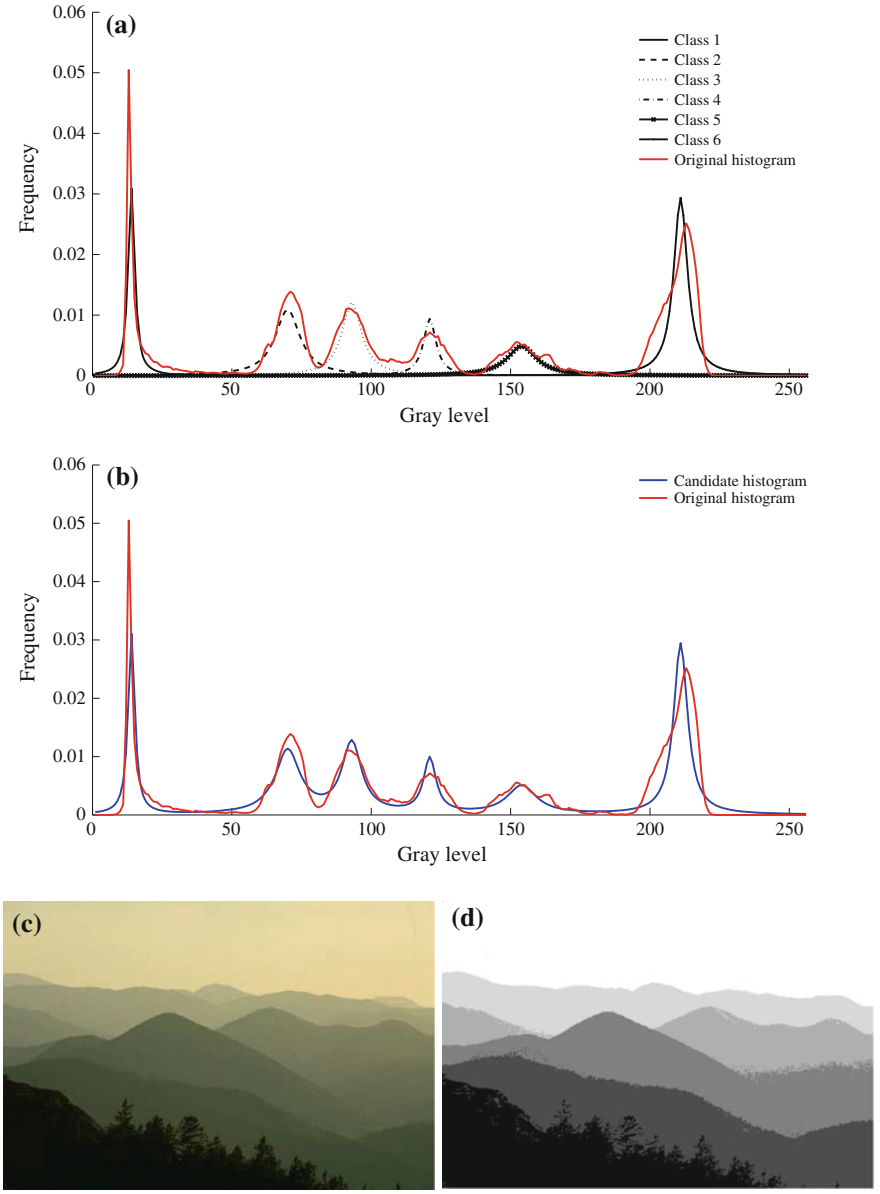

Fig. 5 Image Q24a, **a** class distribution with six classes ($K = 6$), **b** approximation considering six classes, **c** original image, **d** segmented image

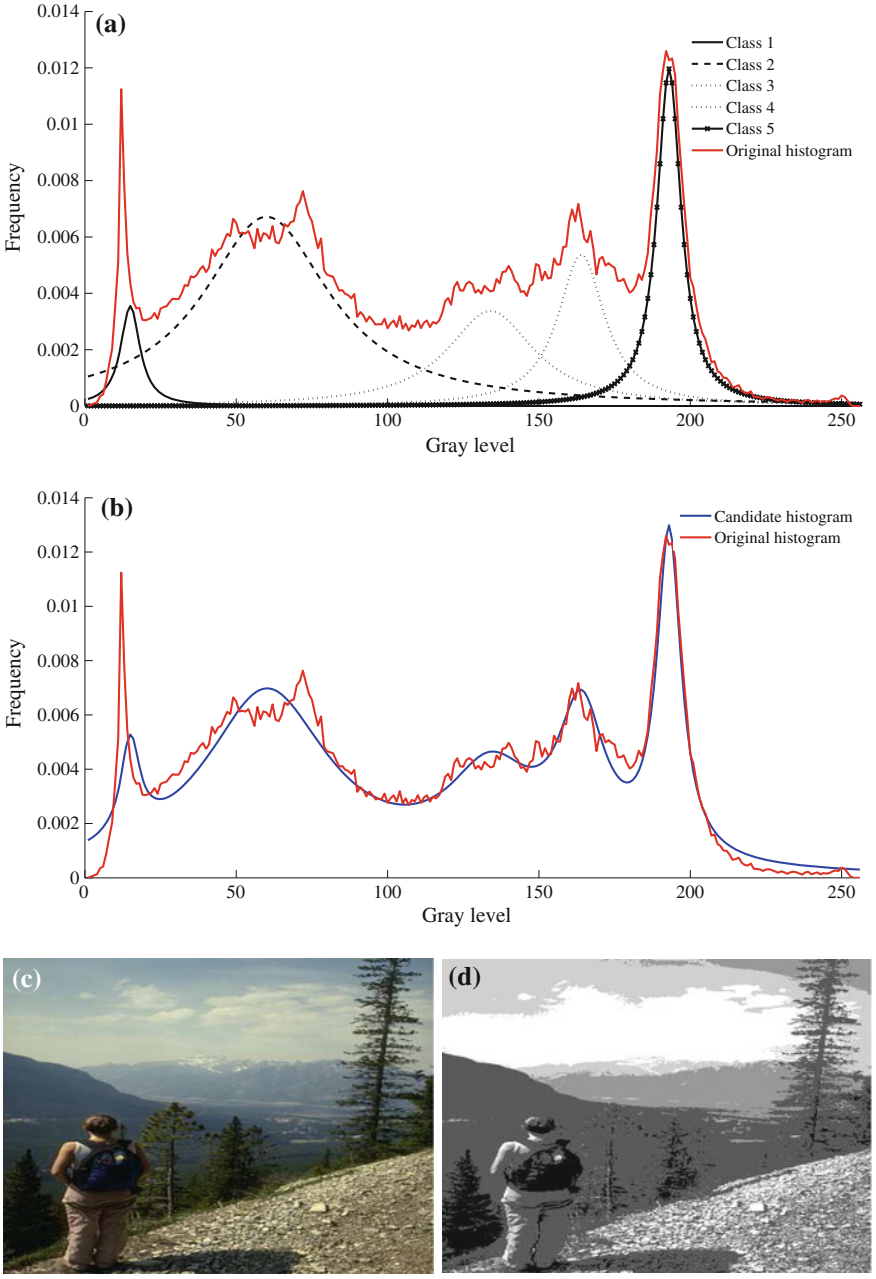

Fig. 6 Image Im001_1, **a** class distribution with five classes ($K = 5$), **b** approximation considering five classes, **c** original image, and **d** segmented image

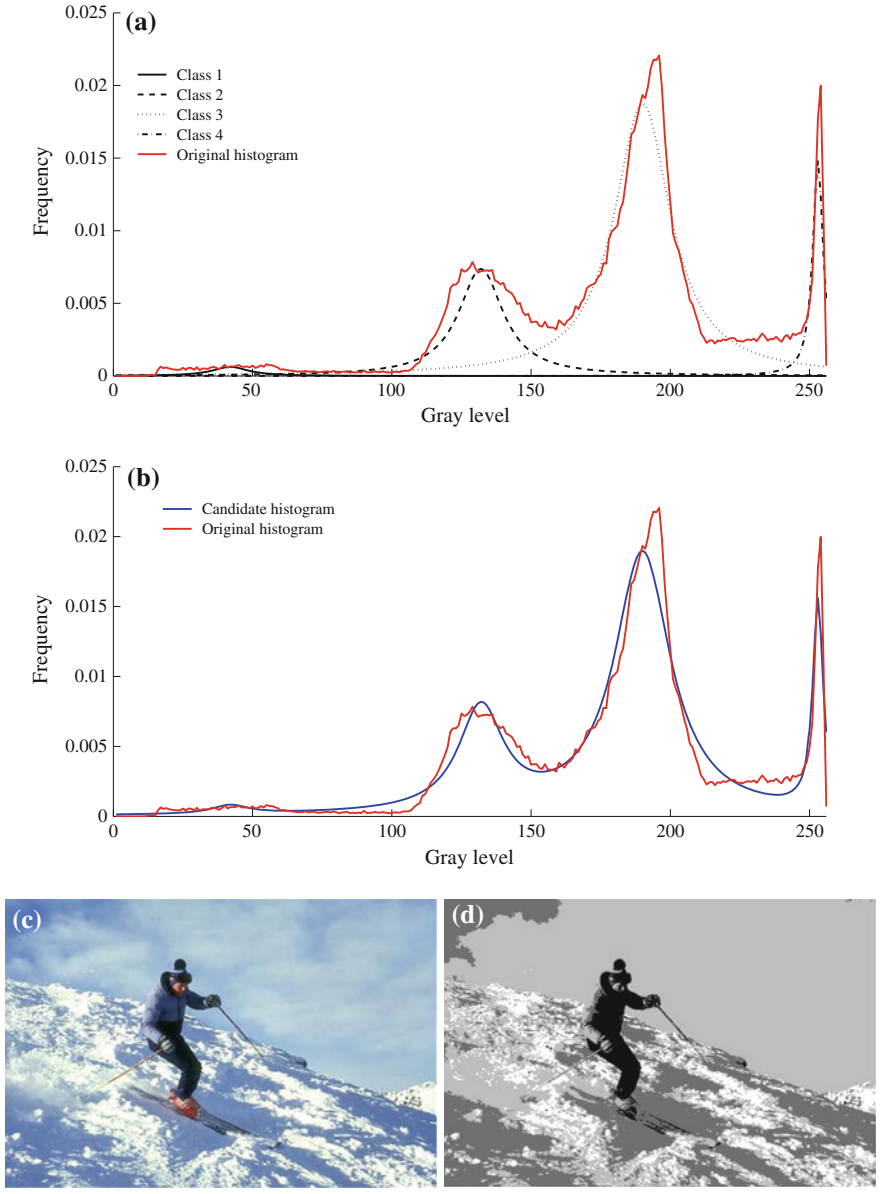

Fig. 7 Image Im002_1, **a** class distribution with five classes ($K = 5$), **b** approximation considering five classes, **c** original image, and **d** segmented image

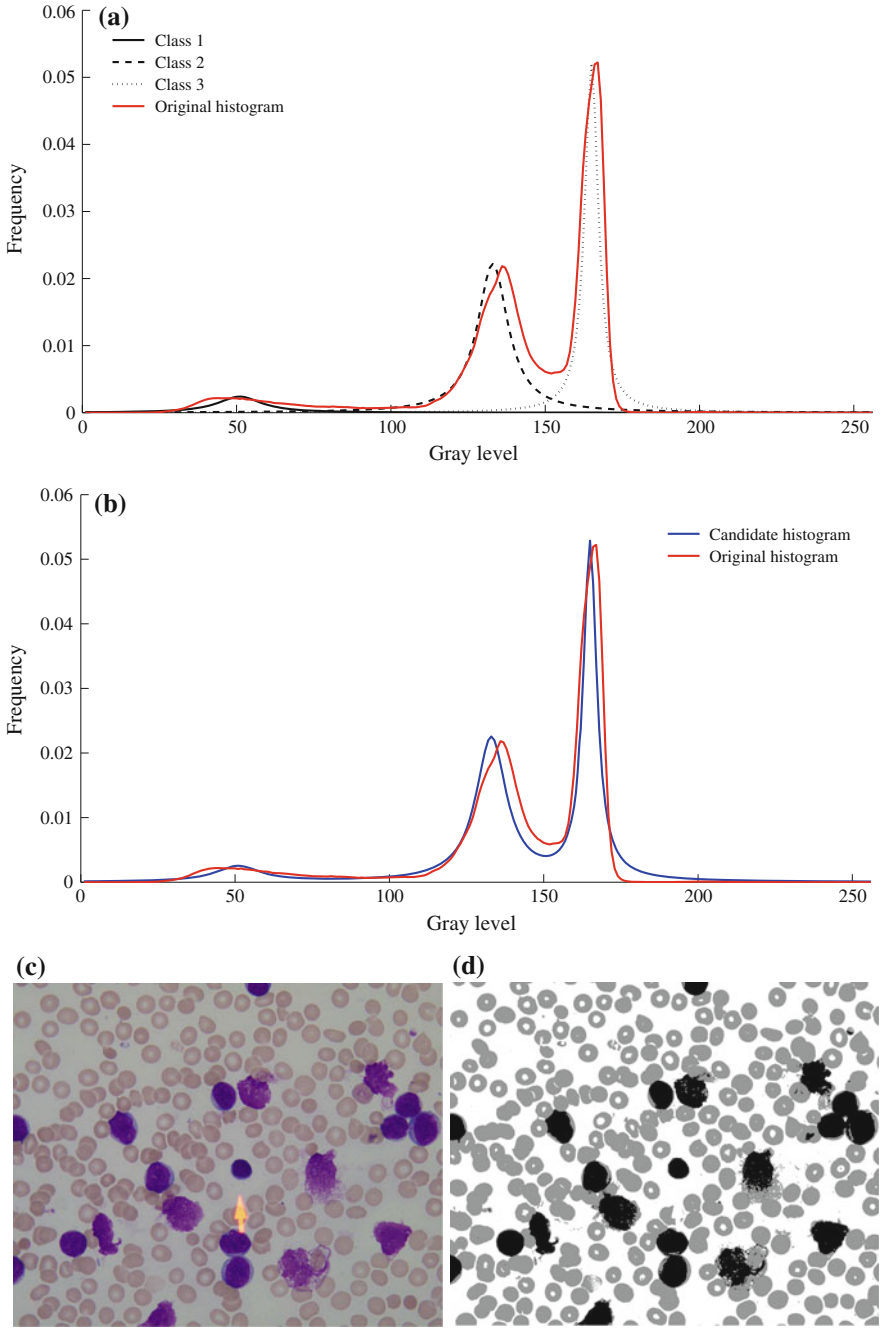

Fig. 8 Image 61060, **a** class distribution with three classes ($K = 3$), **b** approximation considering three classes, **c** original image, and **d** segmented image

Table 3 Experimental results obtained from the comparison between the Otsu and the AO methods

Image	Clases	AO $\mu(\sigma)$			Otsu $\mu(\sigma)$		
		T_1	T_2	T_3	T_1	T_2	T_3
233	2	98(1.09)	NA	NA	90(0)	NA	NA
Q24a	2	117(1.94)	NA	NA	125(0)	NA	NA
Im001_1	3	105(1.36)	154(0.54)	NA	97(0)	148(0)	NA
Im002_1	3	97(2.81)	154(0.63)	NA	96(0)	148(0)	NA
61060	4	84(1.18)	157(0.96)	241(1.89)	91(0)	162(0)	215(0)
253036	4	142(1.77)	189(3.84)	232(3.61)	138(0)	191(0)	222(0)

AO-based algorithm. In all figures, the approximation results over the original histogram are also shown.

In order to enhance the performance analysis, the proposed approach has been compared with the Otsu method [21]. Table 3 shows some results obtained by both methods, considering the mean and standard deviation of threshold values obtained by each algorithm when they have been executed 50 times for each image. The results have been presented in the format mean value μ (standard deviation, σ) whereas the elements that not correspond for a specific experiment are marked by NA (Figs. 4, 5, 6, 7).

Figure 10 shows two images proposed in [54] as segmentation benchmarks. Such problems consist in segmenting different cells, considering that their optimal results have been already obtained by a human expert (ground-truth). Under these conditions, it is possible to compare the segmentation results obtained by our approach and the Otsu method in terms of the optimal results. Figure 11 presents the results obtained by the Otsu method and the AO-based algorithm considering the benchmark images from [54]. A visual inspection of Fig. 11 demonstrates that the Otsu method presents more undesirable artifacts as a consequence of a poor segmentation.

In order to appropriately compare the results from Fig. 11, the Hausdorff distance in terms of the ground-truth has been used. Table 3 shows the averaged Hausdorff distances considering both images from Fig. 11. Considering the mean value μ of the Hausdorff distance, it is clear that the proposed method produces better results, than Otsu's method, as can be seen from Table 4.

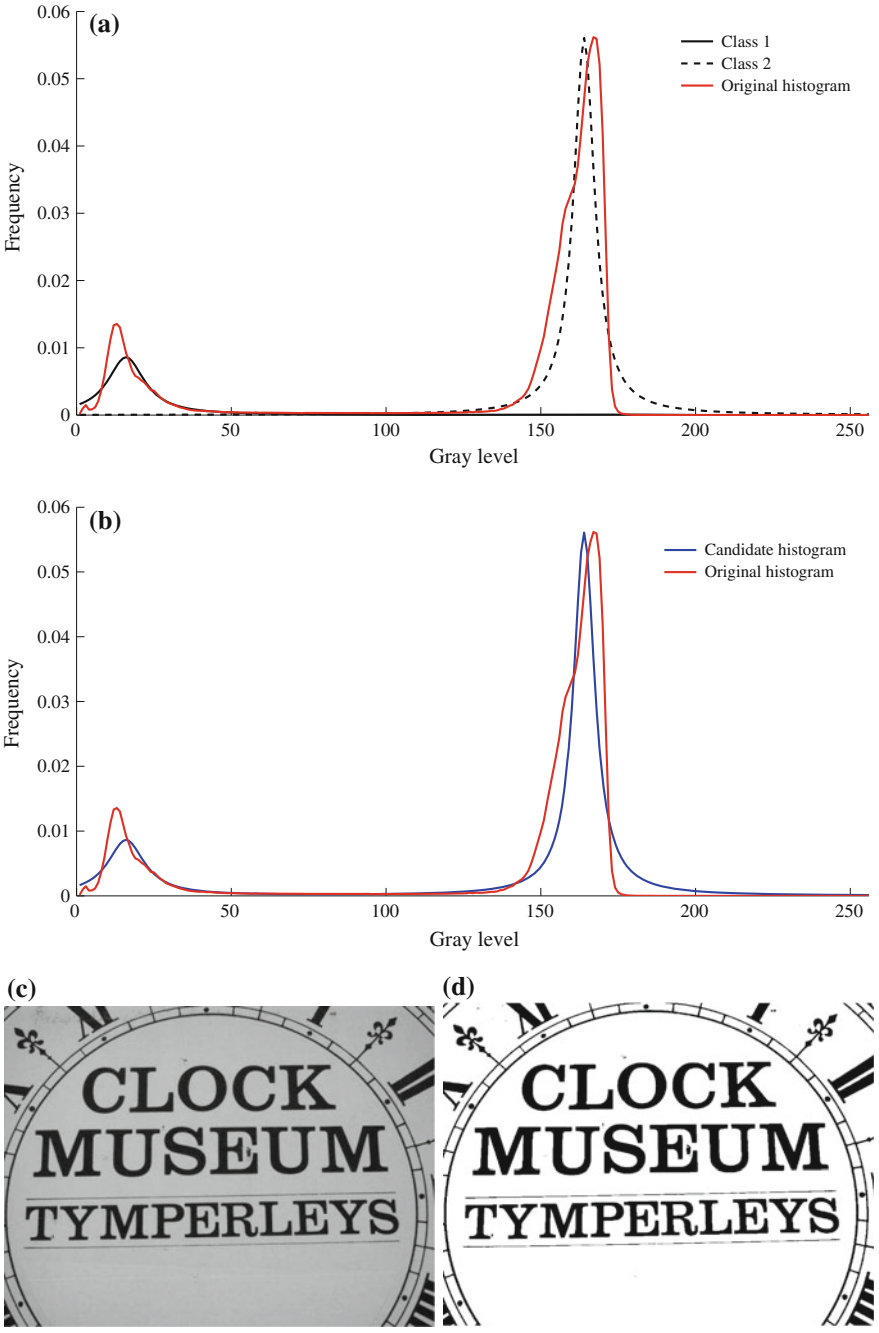

Fig. 9 Image 253036, **a** class distribution with two classes ($K = 2$), **b** approximation considering two classes, **c** original image, and **d** segmented image

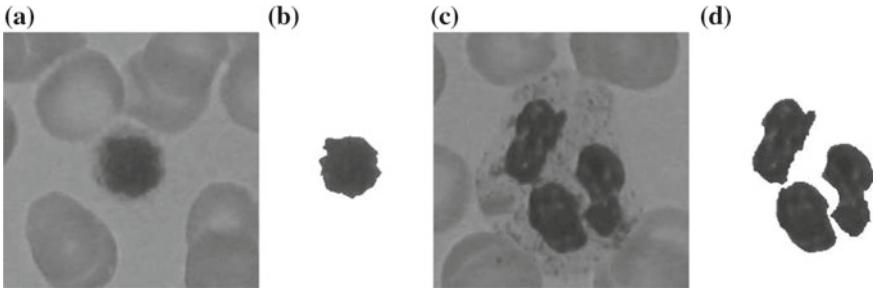

Fig. 10 Benchmark images for comparison with ground-truth. **a** Oiginal image with a single cell, **b** ground-truth, **c** original image with multiple cells and **d** ground-truth

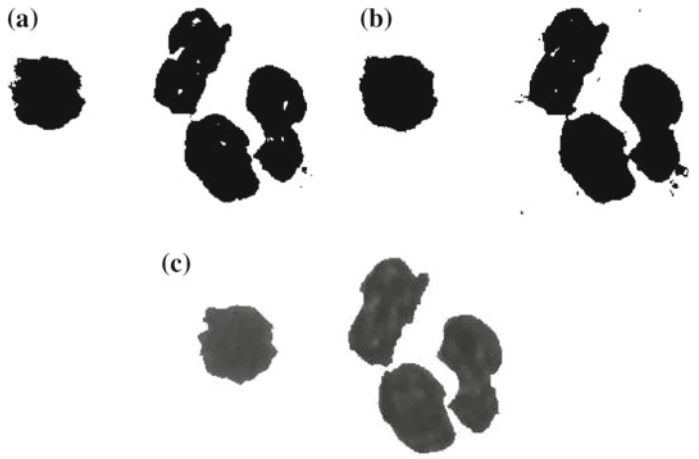

Fig. 11 Some results of the Otsu method and the AO-based algorithm. **a** AO-based. **b** Otsu. **c** Ground-truth images

Table 4 Hausdorff distance of AO against Otsu method

Method	Hausdorff distance	
	μ	σ
AO	2.1364	0.6535
Otsu	2.4655	2.9779×10^{-15}

5 Conclusions

In this chapter, a multi-thresholding segmentation algorithm based on a new evolutionary algorithm called Allostatic Optimization (AO) has been proposed. In the approach, the capacity of the Cauchy distribution to model complex problems (in presence of outliers) is exploited. Our approach assumes that the segmentation

process is considered as an optimization problem by approximating the 1-D histogram of a given image in terms of a Cauchy mixture (CM) model, whose parameters are calculated through the AO algorithm.

In AO, the searcher agents emulate different body conditions which interact to each other by using operators based on the biological principles of the allostasis mechanism. The proposed approach encodes the parameters of the CM as an individual. An objective function by using the Hellinger distance evaluates the matching quality between the CM candidate and the original histogram. Guided by the values of this objective function, the set of encoded candidate mixtures are evolved using the operators defined by AO so that they can fit into the original histogram.

References

1. Arora, S., Acharya, J., Verma, A., Panigrahi, P.K.: Multilevel thresholding for image segmentation through a fast statistical recursive algorithm. Pattern Recogn. Lett. **29**(2), 119–125 (2008)
2. Naz, S., Majeed, H., Irshad, H.: Image segmentation using fuzzy clustering: a survey. In: 2010 6th International Conference on Emerging Technologies, pp. 181–186. IEEE (2010)
3. Janev, M., Pekar, D., Jakovljevic, N., Delic, V.: Eigenvalues driven gaussian selection in continuous speech recognition using hmms with full covariance matrices. Appl. Intell. **33**(2), 107–116 (2010)
4. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. J. Electron. Imaging **13**(1), 146–168 (2004)
5. Beevi, S.Z., Sathik, M.M., Senthamaraikannan, K., Yasmin, J.H.J.: A robust fuzzy clustering technique with spatial neighborhood information for effective medical image segmentation: an efficient variants of fuzzy clustering technique with spatial information for effective noisy medical image segmentation. In: 2010 Second International Conference on Computing, Communication and Networking Technologies, pp. 1–8. IEEE (2010)
6. Chitsaz, M., Seng, W.C.: A multi-agent system approach for medical image segmentation. In: 2009 International Conference on Future Computer and Communication, pp. 408–411. IEEE (2009)
7. Halim, N., Mashor, M., Abdul Nasir, A., Mokhtar, N., Rosline, H.: Nucleus segmentation technique for acute Leukemia. In: 2011 IEEE 7th International Colloquium on Signal Processing and Its Applications, pp. 192–197. IEEE (2011)
8. Mohapatra, S., Patra, D.: Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images. In: 2010 International Conference on Systems in Medicine and Biology, pp. 49–54. IEEE (2010)
9. Mohapatra, S., Patra, D., Kumar, K.: Blood microscopic image segmentation using rough sets. In: 2011 International Conference on Image Information Processing, pp. 1–6. IEEE (2011)
10. Mohapatra, S., Samanta, S.S., Patra, D., Satpathi, S.: Fuzzy based blood image segmentation for automated leukemia detection. In: 2011 International Conference on Computers and Devices for Communication, pp. 1–5. IEEE (2011)
11. Nor Hazlyna, H., Mashor, M., Mokhtar, N., Aimi Salihah, A., Hassan, R., Raof, R., Osman, M.: Comparison of acute leukemia Image segmentation using HSI and RGB color space. In: 10th Internaional Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), pp. 749–752. IEEE (2010)
12. Yang, G., Chen, K., Zhou, M., Xu, Z., Chen, Y.: Study on statistics iterative thresholding segmentation based on aviation image. In: Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2007), vol. 2, pp. 187–188. IEEE (2007)

13. Li, X., Ramachandran, R., He, M., Rushing, J., Graves, S., Lyatsky, W., Germany, G.: Comparing different thresholding algorithms for segmenting auroras. In: International Conference on Information Technology: Coding And Computing 2004. Proceedings. ITCC 2004, vol. 2, pp. 594–601. IEEE (2004)
14. Li, Z., Liu, C., Liu, G., Cheng, Y., Yang, X., Zhao, C.: A novel statistical image thresholding method. AEU Int. J. Electron. Commun. **64**(12), 1137–1147 (2010)
15. Akay, B.: A study on particle swarm optimization and artificial bee colony algorithms for multilevel thresholding. Appl. Soft Comput. **13**(6), 3066–3091 (2013)
16. Cuevas, E., Zaldivar, D., Pérez-Cisneros, M.: A novel multi-threshold segmentation approach based on differential evolution optimization. Expert Syst. Appl. **37**(7), 5265–5271 (2010)
17. Hammouche, K., Diaf, M., Siarry, P.: A comparative study of various meta-heuristic techniques applied to the multilevel thresholding problem. Eng. Appl. Artif. Intell. **23**(5), 676–688 (2010)
18. Horng, M.H.: A multilevel image thresholding using the honey bee mating optimization. Appl. Math. Comput. **215**(9), 3302–3310 (2010)
19. Sathya, P., Kayalvizhi, R.: Modified bacterial foraging algorithm based multilevel thresholding for image segmentation. Eng. Appl. Artif. Intell. **24**(4), 595–615 (2011)
20. Sathya, P., Kayalvizhi, R.: Optimal multilevel thresholding using bacterial foraging algorithm. Expert Syst. Appl. **38**(12), 15549–15564 (2011)
21. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1)
22. Al-Hussaini, E., Ateya, S.: Parametric estimation under a class of multivariate distributions. Stat. Pap. **46**(3), 321–338 (2005)
23. Liu, T., Zhang, P., Dai, W.S., Xie, M.: An intermediate distribution between Gaussian and Cauchy distributions. Phys. A Stat. Mech. Appl. **391**(22), 5411–5421 (2012)
24. Ateya, S., Madhagi, E.: On multivariate truncated generalized cauchy distribution. Stat. Pap. **54**(3), 879–897 (2013)
25. Zhang, J.: A highly efficient l-estimator for the location parameter of the cauchy distribution. Comput. Stat. **25**(1), 97–105 (2010)
26. Pander, T., Przybya, T.: Impulsive noise cancelation with simplified Cauchy-based p-norm filter. Signal Process. **92**(9), 2187–2198 (2012)
27. Gao, Q., Lu, Y., Sun, D., Sun, Z.L., Zhang, D.: Directionlet-based denoising of SAR images using a Cauchy model. Signal Process. **93**(5), 1056–1063 (2013)
28. Guozhong, C., Xingzhao, L.: Cauchy pdf modelling and its application to SAR image despeckling. J. Syst. Eng. Electron. **19**(4), 717–721 (2008)
29. Kocsor, A., Tth, L.: Application of kernel-based feature space transformations and learning methods to phoneme classification. Appl. Intell. **21**(2), 129–142 (2004)
30. Olsson, R.K., Petersen, K.B., Lehn-Schiøler, T.: State-space models: from the EM algorithm to a gradient approach. Neural Comput. **19**(4), 1097–1111 (2007)
31. Park, H., Amari, S.I., Fukumizu, K.: Adaptive natural gradient learning algorithms for various stochastic models. Neural Netw. **13**(7), 755–764 (2000)
32. Park, H., Ozeki, T.: Singularity and slow convergence of the EM algorithm for gaussian mixtures. Neural Process. Lett. **29**(1), 45–59 (2009)
33. I Abdul-Moniem, Y.M.S.: Tl-moments and l-moments estimation for the generalized pareto distribution. Appl. Math. Sci. **3**(1)
34. Reeds, J.A.: Asymptotic number of roots of cauchy location likelihood equations. Ann. Stat. **13**
35. Barnett, V.D.: Order statistics estimators of the location of the cauchy distribution. J. Am. Stat. Assoc. **61**
36. Fogel, L.J., Owens, A.J., Walsh, M.J.: Artificial Intelligence through Simulated Evolution. Wiley, Chichester (1966)
37. De Jong, K.A.: An analysis of the behavior of a class of genetic adaptive systems. Ph.D. Thesis, Ann Arbor, MI, USA (1975)
38. Holland, J.H.: Adaptation in Natural and Artificial Systems. MIT Press, Cambridge (1992)

39. De Castro, L.N., Von Zuben, F.J.: Artificial immune systems: Part i-basic theory and applications. Universidade Estadual de Campinas, Dezembro de, Technical Report 210 (1999)
40. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks, vol. 4, pp. 1942–1948. IEEE (1995)
41. Karaboga, D.: An idea based on honey bee swarm for numerical optimization. Technical Report TR06, Engineering Faculty, Computer Engineering Department, Erciyes University (2005)
42. Encyclopedia of Human Body Systems. Julie McDowell, Greenwood (2011)
43. Cannon, W.: Bodily Changes in Pain, Hunger, Fear and Rage: An Account of Recent Researchers into the Function of Emotional Excitement. Appleton, New York (1929)
44. Cannon, W.: The Wisdom of the Body. Norton (1932)
45. Gross, C.G.: Claude bernard and the constancy of the internal environment. Neuroscientist **4**
46. Fletcher, J.M.: Homeostasis as an explanatory principle in psychology. Psychol. Rev. **49**(1)
47. McEwen, B.: Allostasis and allostatic load: implications for neuropsychopharmacology. Neuropsychopharmacology **22**(2)
48. McEwen B.S., Wingfield, J.C.: The concept of allostasis in biology and biomedicine. Horm. Behav. **43**(1)
49. Romero, L.M., Dickens, M.J., Cyr, N.E.: The reactive scope model a new model integrating homeostasis, allostasis, and stress. Horm. Behav. **55**(3), 375–389 (2009)
50. Schulkin, J.: Allostasis: a neural behavioral perspective. Horm. Behav. **43**(1)
51. Sterling, P.: Allostasis: a model of predictive regulation. Physiol. Behav. **106**(1)
52. Martinez-Lavin, M., Vargas, A.: Complex adaptive systems allostasis in fibromyalgia. Rheum. Dis. Clin. North Am. **35**(2), 285–98 (2009)
53. Karunamuni, R., Wu, J.: Minimum Hellinger distance estimation in a nonparametric mixture model. J. Stat. Plan. Infer. **139**(3), 1118–1133 (2009)
54. Labati, R.D., Piuri, V., Scotti, F.: All-IDB: The acute lymphoblastic leukemia image database for image processing. In: 2011 18th IEEE International Conference on Image Processing, pp. 2045–2048. IEEE (2011)
55. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of Eighth IEEE International Conference on Compututer Vision. ICCV 2001, vol. 2, pp. 416–423. IEEE Computer Society (2001)