31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

MRI Segmentation using Dialectical Optimization

Wellington P. dos Santos, Francisco M. de Assis, and Ricardo E. de Souza

***Abstract*— Biology, Psychology and Social Sciences are intrinsically connected to the very roots of the development of algorithms and methods in Computational Intelligence, as it is easily seen in approaches like genetic algorithms, evolutionary programming and particle swarm optimization. In this work we propose a new optimization method based on dialectics using fuzzy membership functions to model the influence of interactions between integrating poles in the status of each pole. Poles are the basic units composing dialectical systems. In order to validate our proposal we designed a segmentation method based on the optimization of k-means using dialectics for the segmentation of MR images. As a case study we used 181 MR synthetic multispectral images composed by proton density, T_1- and T_2-weighted synthetic brain images. Comparing our proposal to k-means, fuzzy c-means, and Kohonen's self-organized maps, concerning the quantization error, we proved that our method can improved results obtained using k-means.**

I. Introduction

The dialectical method is based on considering parts of reality (or *phenomena*) as dynamical *systems*. These systems are composed by several integrating *poles*. Each pole has a *potency* (as named by Aristotle, the philosopher) or *force*. These poles interact with each other, in a process called *pole struggle*, where the strongest poles become the dominant poles, while similar poles are fused, new poles are generated from the hardest conflicts between poles, and the weakest poles are absorbed or destroyed [1].

This paper proposes a class of algorithms based on a specific interpretation of the dialectics, namely the Objective Dialectical Method, to be used in optimization problems. In order to validate our proposal we designed a segmentation method based on the optimization of k-means using dialectics for the segmentation of magnetic resonance (MR) images. As a case study we used MR synthetic multispectral images composed by proton density, T_1- and T_2-weighted synthetic images of 181 slices with 1 mm, resolution of 1 mm^3, for a normal brain and a noiseless MR tomographic system without field inhomogeneities, amounting a total of 543 images, generated by the simulator BrainWeb [2]. Our principal target here is comparing our proposal with other non-supervised classifiers, namely k-means, classical fuzzy c-means, and Kohonen's self-organized maps, concerned with the quantization error.

This work is organized as follows: section II presents the proposal of the dialectical optimization method based on dialectics; section III shows the image fidelity expressions, the parameters of the non-supervised image classification methods, and the synthetic brain MR images used in this work; the quantitative and qualitative experimental results for image quantization are presented in section IV, where in section V some discussion on experimental results is performed, whilst conclusions are also presented.

II. Dialectical Optimization Method

The fundamental idea of the dialectical optimization method is to associate the objective function of the optimization problem to the social force of each pole: the adjustment of poles depends on the present hegemonic pole and the historical hegemonic pole. The hegemonic pole is the pole with the greatest social force among the set of the forces of all poles in a determined historical moment. The present hegemonic pole is the hegemonic pole of the present instant, whilst the historical hegemonic pole is the hegemonic pole of the historical period from the beginning of dialectical system to the actual instant. The search for possible solutions occurs in two intertwined phases: evolution and revolutionary crisis.

A. General Definition

To understand the dialectical optimization method it is important to present some important definitions and assumptions:

- Pole: It is the fundamental integrating unit of a dialectical system. Given the set of poles
  $$\Omega = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\},$$
  the *pole* i, as defined in [3], is associated to the vector of weights $\mathbf{w}_i = (w_{i,1}, w_{i,2}, \dots, w_{i,n})^T$, where $\mathbf{w}_i \in S$, m is the *number of poles* and n is the *dimensionality of the system*. In our proposal, poles also have the role of solution candidates. Objective function's and system's dimensionalities are the same, where $f : S \to \mathbb{R}$, $S \subseteq \mathbb{R}^n$ and $\Omega \subseteq S$.
- Social force: To each pole i is associated a *social force* equals to the value of the objective function f in the i-th pole, that is, the social force of the i-th pole is given by $f(\mathbf{w}_i)$.
- Hegemony: In the process of pole struggle, the k-th pole has the *hegemony* in instant t when:
  $$f(\mathbf{w}_k(t)) = f_C(t) = \max_{1 \leq j \leq m(t)} f(\mathbf{w}_j(t)), \qquad (1)$$

Wellington P. dos Santos is with Departamento de Engenharia Elétrica, Universidade Federal de Campina Grande, 58109-970, Campina Grande, Paraíba, Brazil, and Departamento de Sistemas e Computação, Escola Politécnica de Pernambuco, Universidade de Pernambuco, 50.720-001, Recife, Pernambuco, Brazil (e-mail: wps@dsc.upe.br).

Francisco M. de Assis is with Departamento de Engenharia Elétrica, Universidade Federal de Campina Grande, 58109-970, Campina Grande, Paraíba, Brazil (e-mail: fmarcos@dee.ufcg.edu.br).

Ricardo E. de Souza is with Departamento de Física, Universidade Federal de Pernambuco, 50670-901, Recife, Pernambuco, Brazil (e-mail: res@df.ufpe.br).

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

where $1 \leq k \leq m(t)$. The vector $\mathbf{w}_C(t) = \mathbf{w}_k(t)$ is called the *present hegemonic pole*, or *contemporary hegemonic pole*, where $f_C(t)$ is the *present hegemonic force*, or *contemporary hegemonic force*. The *historical hegemonic force* in instant t, $f_H(t)$, is given by:

$$f_H(t) = \max_{0 \leq t' \leq t} f_C(t'), \tag{2}$$

where $\mathbf{w}_H(t) = \mathbf{w}_C(t')$, for $f(\mathbf{w}_C(t')) = f_H(t)$ and $0 \leq t' \leq t$.

- Absolute antithesis: Given x for $a \leq x \leq b$, where $a, b \in \mathbb{R}$, the *opposite* of x is given by [4]:

$$\breve{x} = b - x + a. \tag{3}$$

This result has a geometrical interpretation: considering the translation $x' = x - (a+b)/2$, the translation opposite is $\breve{x}' = -x' = -x + (a+b)/2$; the opposite inversion is $\breve{x} = \breve{x}' + (a+b)/2 = -x + a + b = b - x + a$. Consequently, the opposite number is defined in relation with the mean of the interval. Assuming $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and $\mathbf{x} \in S \Rightarrow r_i \leq x_i \leq s_i$, $\forall i = 1, 2, \ldots, n$, where r_i and s_i are the inferior and superior boundaries of the i-th dimension of S, the associated *opposite vector*, $\breve{\mathbf{x}} = (\breve{x}_1, \breve{x}_2, \ldots, \breve{x}_n)^T$ has its coordinates calculated as follows [4]:

$$\breve{x}_i = s_i - x_i + r_i, \tag{4}$$

where $i = 1, 2, \ldots, n$. The *absolute antithesis vector* of pole $\mathbf{w}$ is defined by its opposite vector $\breve{\mathbf{w}}$. Rahnamayan *et al* affirm that, in evolutionary programming and particle swarm optimization, the presence of pairs of opposite vectors in the initial population typically accelerates the convergence of algorithms in 10% [4].

- Contradiction: The contradiction between poles $\mathbf{w}_p$ and $\mathbf{w}_q$ is given by:

$$\delta_{p,q} = d(\mathbf{w}_p, \mathbf{w}_q), \tag{5}$$

where $d : S^2 \rightarrow \mathbb{R}_+$ is a distance function, for $\delta_{p,q} = \delta_{q,p}$, $\forall p, q$. A typical distance function is the Euclidean distance function.

- Synthesis: According to the dialectical conception, the synthesis is the resolution of the contradiction between two poles, where one of them is thesis and the other is antithesis [1]. The poles $\mathbf{w}_u, \mathbf{w}_v \in S$ are the possible syntheses between poles $\mathbf{w}_p$ and $\mathbf{w}_q$, calculated as follows:

$$\mathbf{w}_u = g_1(\mathbf{w}_p, \mathbf{w}_q), \tag{6}$$

$$\mathbf{w}_v = g_2(\mathbf{w}_p, \mathbf{w}_q), \tag{7}$$

where $g : S^2 \rightarrow S$. A very intuitive approach strictly based on the dialectical conception is to consider that all syntheses inherit characteristics of theses and antitheses [1], as following:

$$w_{u,i} = \begin{cases} w_{p,i}, & i \bmod 2 = 0, \\ w_{q,i}, & i \bmod 2 = 1, \end{cases}, \tag{8}$$

$$w_{v,i} = \begin{cases} w_{p,i}, & i \bmod 2 = 1, \\ w_{q,i}, & i \bmod 2 = 0, \end{cases}, \tag{9}$$

for $i = 1, 2, \ldots, n$. Notice that there are similarities between this definition and models of inheritance in genetic algorithms. It is also possible to use different definitions of synthesis, using different inheritance criteria, in case we need to generate diversity in other manners.

B. Algorithm for Search and Optimization

First of all, it is necessary to set the *initial number of poles*, $m(0)$, integrating the dialectical system $\Omega(0)$, the *number of historical phases*, n_P, and the *duration of each historical phase*, n_H. The number of initial poles must be even, because a half population is randomly generated, and the other half is obtained by the generation of the respective opposite poles inside the domain of the functions. From the point of view of dialectics, here we have a set of initial poles composed by thesis-antithesis pairs of poles in antagonic contradiction, generating a more intense initial dynamics, once the pole struggle is more intense in this case [1]. Consequently:

$$w_{i,j}(0) = \begin{cases} U(r_j, s_j), & 1 \leq j \leq \frac{1}{2}m(0), \\ \breve{w}_{i',j}(0), & 1 + \frac{1}{2}m(0) \leq j \leq m(0), \end{cases} \tag{10}$$

for $i' = i - \frac{1}{2}m(0)$, $1 \leq i \leq m(0)$ and $1 \leq j \leq n$, where n is the dimensionality of the optimization problem, $U(r_j, s_j)$ is a random number uniformly distributed in the interval $[r_j, s_j]$, and $S = \bigcap_{j=1}^{n} [r_j, s_j]$, since $s_j > r_j$ and $s_j, r_j \in \mathbb{R}$.

While the maximum number of historical phases, n_P, is not reached, and the historical hegemonic force is not bigger than a given superior threshold of force (initial estimative of the maximum value of the objective function), $f_H(t) < f_{\sup}$ (criterion to estimate the maximum value of objective function is reached), the phases of evolution and revolutionary crisis are repeated, in this order.

Evolution: While the maximum number of iterations, n_H, is not reached, and $f_H(t) < f_{\sup}$, poles are adjusted according to the following expression:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \Delta\mathbf{w}_{C,i}(t) + \Delta\mathbf{w}_{H,i}(t), \tag{11}$$

for

$$\Delta\mathbf{w}_{C,i}(t) = \eta_0 (1 - \mu_{C,i}(t))^2 (\mathbf{w}_C(t) - \mathbf{w}_i(t)), \tag{12}$$

$$\Delta\mathbf{w}_{H,i}(t) = \eta_0 (1 - \mu_{H,i}(t))^2 (\mathbf{w}_H(t) - \mathbf{w}_i(t)), \tag{13}$$

where $0 < \eta_0 < 1$. The terms $\Delta\mathbf{w}_{C,i}(t)$ and $\Delta\mathbf{w}_{H,i}(t)$ are used to model the influence of present and historic hegemonies, in this order, above the i-th pole. The terms $\mu_{C,i}$ and $\mu_{C,i}$ are the *present membership* and the *historical membership*, respectively, defined as following, based on the membership functions of the classical version of *fuzzy* c-means non-supervised classifier [5]:

$$\mu_{C,i}(t) = \left(\sum_{j=1}^{m} \frac{|f(\mathbf{w}_i(t)) - f_C(t)|}{|f(\mathbf{w}_j(t)) - f_C(t)|} \right)^{-1}, \tag{14}$$

$$\mu_{H,i}(t) = \left(\sum_{j=1}^{m} \frac{|f(\mathbf{w}_i(t)) - f_H(t)|}{|f(\mathbf{w}_j(t)) - f_H(t)|} \right)^{-1}, \tag{15}$$

where $1 \leq i \leq m(t)$. Therefore, when $f(\mathbf{w}_i(t))$ is close to $f_C(t)$, the term $\mu_{C,i}(t)$ is close to 1, turning $\Delta\mathbf{w}_{C,i}(t)$ closer to 0 and, consequently, turning the influence of the present force correlation almost null. When $f(\mathbf{w}_i(t))$ is close to $f_H(t)$, the behavior is similar.

Revolutionary Crisis: In the stage of revolutionary crisis, the following steps are executed:

1) All contradictions $\delta_{i,j}$ are evaluated; the poles whose contradictions are smaller than the *minimum contradiction* $\delta_{\min}$ are fused, in a process defined as following

$$\delta_{i,j}(t) > \delta_{\min} \Rightarrow \mathbf{w}_i(t), \mathbf{w}_j(t) \in \Omega(t+1), \quad (16)$$

$$\delta_{i,j}(t) \leq \delta_{\min} \Rightarrow \mathbf{w}_i(t) \in \Omega(t+1). \quad (17)$$

$i \neq j$, $\forall i,j$ where $1 \leq i,j \leq m(t)$ and $\Omega(t+1)$ is the new set of poles.

2) From the contradictions evaluated previously, the greatest are selected and considered the *principal contradictions* of the dialectical system; poles involved in principal contradictions are considered thesis-antithesis pairs and, with their synthesis poles, they now belong to the new set of poles, that is:

$$\delta_{i,j}(t) = \delta_{\max} \Rightarrow \mathbf{w}_u(t), \mathbf{w}_v(t) \in \Omega(t+1), \quad (18)$$

where

$$\delta_{\max}(t) = \max_{1 \leq p,q \leq m(t)} \{\delta_{p,q}(t) : p \neq q\},$$

$$\mathbf{w}_u(t) = g_1(\mathbf{w}_i(t), \mathbf{w}_j(t)),$$

$$\mathbf{w}_v(t) = g_2(\mathbf{w}_i(t), \mathbf{w}_j(t)),$$

for $i \neq j$, $\forall i,j$ where $1 \leq i,j \leq m(t)$.

3) The *effect of crisis* is added, given the *maximum crisis*, $\chi_{\max}$, to all poles in the dialectical system $\Omega(t+1)$, generating the new set of poles, $\Omega(t+2)$, for $\mathbf{w}_k(t+2) \in \Omega(t+2)$, since

$$w_{k,i}(t+2) = w_{k,i}(t+1) + \chi_{\max} G(0,1), \quad (19)$$

for $1 \leq k \leq m(t+1)$ and $1 \leq i \leq n$, where $G(0,1)$ is a random number distributed according to the distribution of Gauss, with expected value 0 and variance 1.

4) If the stop criterion is not reached, a new set of poles is generated, as following:

$$\mathbf{w}_i(t+2) \in \Omega(t+2) \Rightarrow \breve{\mathbf{w}}_i(t+2) \in \Omega(t+2), \quad (20)$$

for $1 \leq i \leq m(t+2)$, where $m(t+2) = 2m(t+1)$. Consequently, the set of poles is enlarged by adding poles in antagonic antithesis to the others previously existing. Such a procedure models the dialectical property of a system carrying opposite forces as a seed of potential qualitative transformation in every historical changing.

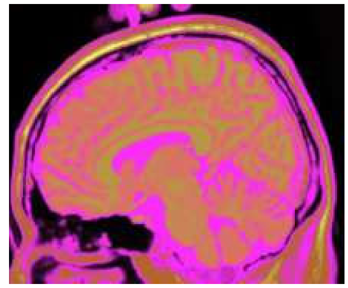

Fig. 1. R0-G1-B2 colored composition of PD-, T_1-, and T_2-weighted MR images of the 97th slice

III. Materials and Methods

A. MR Images

In this work we adopted the following case study: we used MR synthetic multispectral images composed by proton density, T_1- and T_2-weighted synthetic sagital images of 181 slices with 1 mm, resolution of 1 mm^3, for a normal brain and a noiseless MR tomographic system without field inhomogeneities, amounting a total of 543 images, generated by MR image simulator BrainWeb [6], [2]. All images were composed by R0-G1-B2 colored compositions, where bands 0, 1 and 2 are PD- (proton density), T_1- and T_2-weighted MR images. Figure 1 shows the image of the 97th slice.

B. Quantization Error

The quantization error is an indirect measure of the quantization distorsion. It is used to evaluate the problem of clustering the pixels of the image $f : S \to W^n$, with dimensions $n_H \times n_W$ and $n_B = n$ bands, in n_G groups (classes) with centroids $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n_G}\}$. Hence the clustering process is reduced to minimize the following function:

$$J_e = \sum_{i=1}^{n_G} \sum_{\forall f(\mathbf{u}) \in G_i} \frac{||f(\mathbf{u}) - \mathbf{v}_i||}{n_G n_{G,i}}, \quad (21)$$

where $\mathbf{u} \in S$ and $||f(\mathbf{u}) - \mathbf{v}_i||$ is the distance between the pixel *pixel* $f(\mathbf{u})$ of the image $f : S \to W^n$ and the centroid of the i-th group $\mathbf{v}_i$, whilst $n_{G,i}$ is the number of elements of f grouped in the i-th cluster and J_e is the quantization error [7], [8]. Hence the solution candidates are:

$$\mathbf{x} = (\mathbf{v}_1^T, \mathbf{v}_2^T, \ldots, \mathbf{v}_{n_G}^T)^T. \quad (22)$$

C. Non-Supervised Image Classification Methods

The synthetic multispectral images obtained by colored compositions R0-G1-B2 were classified using the following methods, also used to evaluate vector quantization performance:

1) *Kohonen self-organized map classifier (KO)*: 3 inputs, 13 outputs, maximum of 200 iterations, initial learning rate $\eta_0 = 0.1$, circular architecture, Gaussian function of distance;
2) *Fuzzy c-means classifier (CM)*: 3 inputs, 13 outputs, maximum of 200 iterations, initial learning rate $\eta_0 = 0.1$;

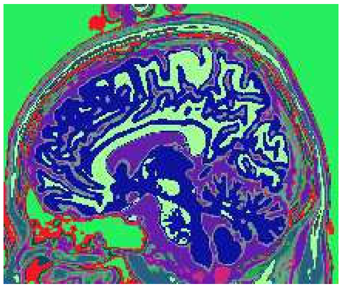

Fig. 2. Classification results of the 97th slice using KO method

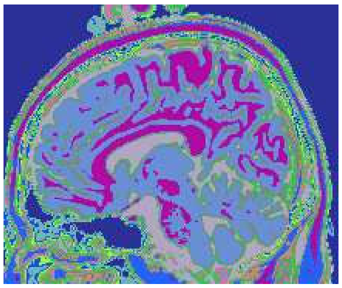

Fig. 3. Classification results of the 97th slice using CM method

3) *K-means classifier (KM)*: 3 inputs, 13 outputs, maximum of 200 iterations, initial learning rate $\eta_0 = 0.1$.
4) *Objective dialectical method (ODM)*: 20 initial poles, 10 historical phases of 20 iterations each phase (stop criteria), initial step $\eta_0 = 0.99$, minimum contradiction of 0.1, maximum contradiction of 0.9, maximum crisis of 0.9, threshold value of 0.01. After all historical phases, the training process was finished with only 2 poles. As can be seen, ODM is a bit more complex than other segmentation methods.

IV. Experimental Results

Figures 2, 3, 4 and 5 show segmentation results for the image of the 97th slice, figure 1, using methods KO, CM, KM and ODC, respectively, where each class is associated to a specific random color. The optimization process was performed according to the quantization error considering the 97th slice, once it presents all structures of interest. Table I shows the results of the quantization error of the segmentation methods for the 97th slice, whilst table II shows results (sample average and mean deviation) considering all 181 slices with 3 bands (DP, T_1 and T_2), for KO, CM, KM and ODM methods.

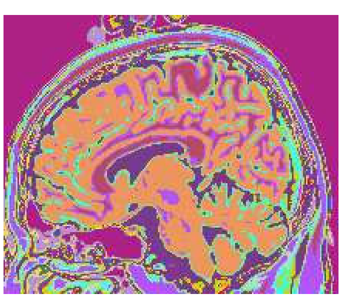

Fig. 4. Classification results of the 97th slice using KM method

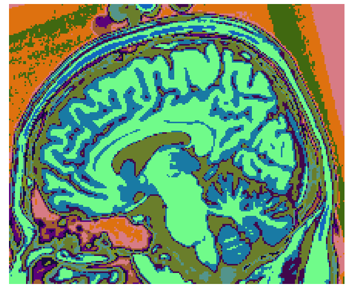

Fig. 5. Classification results of the 97th slice using ODM method

	KO	CM	KM	ODM
J_e	9.02	7.89	9.37	8.65

TABLE I

Measures of quantization error considering the 97th slice

V. Discussion and Conclusions

From the analysis of table I it is clearly seen that ODM improved the results of KM, since we got J_e of 8.65 against 9.37 for the 97th slice. This result is also a bit better than the result obtained by KO, but still worse than CM's, with quantization errors of 9.02 and 7.89, respectively. From table II we can see that ODM obtained a sensible reduction of J_e compared with KM, but the general results got with KO and CM were better. These results prove that the use of dialectical optimization can improve the performance of classical k-means according to a given cluster validity index.

References

[1] A. S. Vázquez. *Filosofia da Práxis*. Expressão Popular, São Paulo, 2007.

[2] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and Construction of a Realistic Digital Brain Phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.

[3] W. P. Santos, F. M. Assis, R. E. Souza, and P. B. Santos Filho. Evaluation of Alzheimer's Disease by Analysis of MR Images using Objective Dialectical Classifiers as an Alternative to ADC Maps. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vancouver, Canada, 2008. EMBS-IEEE.

[4] S. Rahnamayan, H. R. Tizhoosh, and M. M. A. Salama. A novel population initialization method for accelerating evolutionary algorithms. *Computers and Mathematics with Applications*, (53):1605–1614, 2007.

[5] C. Zhu and T. Jiang. Multicontext fuzzy clustering for separation of brain tissues in magnetic resonance images. *NeuroImage*, (18):685–696, 2003.

[6] R. K. S. Kwan, A. C. Evans, and G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, 1999.

[7] M. Omran, A. P. Engelbrecht, and A. Salman. Particle swarm optimization method for image clustering. *International Journal of Pattern Recognition and Artificial Intelligence*, 19(3):297–321, 2005.

[8] D. W. van der Merwe and A. P. Engelbrecht. Data clustering using particle swarm optimization. In *IEEE Congress on Evolutionary Computation*, Canberra, Australia, 2003. CIS-IEEE.

	KO	CM	KM	ODM
J_e	8.93 ± 0.25	8.27 ± 0.47	9.21 ± 0.29	9.17 ± 0.27

TABLE II

Measures of quantization error considering all slices