XL CILAMCE
IBERO-LATIN AMERICAN
CONGRESS ON
COMPUTATIONAL
METHODS IN
ENGINEERING

NOVEMBER
11-14, 2019
Praiamar Natal Hotel & Convention
Natal, RN-BRAZIL

GENETIC ALGORITHMS FOR MULTI-THRESHOLDING OPTIMIZATION FOR MRI SEGMENTATION

Tiago S. Ferreira
Luís Fillype da Silva
Luan A. Sousa
Ritta S. Abreu
Italo R. Feitosa
Marta de Oliveira Barreiros
tiago.s.f3@gmail.com
silvaluis_@outlook.com
luanfelcomp@gmail.com
rittaseixas@gmail.com
italo85199@gmail.com
marta-barreiros@hotmail.com
Dept. of Computer Engineering, State University of Maranhão
Cidade Universitária, R. Paulo VI - São Cristovão, 65055-000, MA, Brasil

Abstract. Image segmentation is an essential task in image processing that aims to simplify and/or change its representation for easier analysis. Medical images of the brain are extremely important in the detection and diagnosis of diseases, as well as relevance in research in the biological area. In this article, an approach for the segmentation of brain images is proposed, using the method based on genetic algorithms with multiple thresholds, in order to find thresholds that optimally separate gray matter, white matter and cerebrospinal fluid in cranial images. The most appropriate threshold parameters were identified in order to obtain the best possible solution for image segmentation. Magnetic resonance imaging of 10 patients was obtained from the National Center for Image Guided Therapy database. Next, a preprocessing was established to improve the input images in the segmentation algorithm. It was based on the application of multi-threshold image segmentation, which aims to maximize intra-class variation between object(s) and background, as well as the minimization of inter-class variation, i.e. between background pixels among pixels of objects, then obtaining the tissues of interest. The genetic algorithm proved to be an efficient method to obtain optimal values for the thresholds, so it is possible to perform segmentation of the image showing the brain tissues of interest.

Keywords: Image processing, Genetic Algorithm, Brain MRI

1 Introduction

Image segmentation has been an important and challenging issue in the field of image processing and it plays a critical role for most image analysis tasks such as object recognition, object-based image compression and content based indexing. Image segmentation has been varying different interpretation for different kinds of application towards content analysis and image understanding [3] . Brain MRI segmentation is often the first step in many clinical applications, commonly used for measuring and visualizing the brain's anatomical structures, for analyzing brain changes, for delineating pathological regions, and for surgical planning and image-guided interventions .

The accuracy of image segmentation stage would have a great impact on the effectiveness of subsequent stages of the image processing. Image segmentation problem has been studied by many researchers for several years; however, due to the characteristics of the images such as their different modal histograms, the problem of image segmentation is still an open research issue and so further investigation is needed. Providing ideal solutions for many practices applications, genetic algorithms have recently been used in order to find optimal solutions to the image problem segmentations. In general terms, a Genetic Algorithm (GA) is a computational model that simulates the biological evolutionary process of natural selection .

In this paper, genetic algorithm it's used with to convert thresholding problem into an optimization problem by finding threshold(s), in order to segment MRI and find brain structures of interest. The proposed method aims to maximize intra-class variation between object(s) and background, as well as the minimization of inter-class variation, i.e. between background pixels among pixels of objects[2].

2 Image Threshold Segmentation

A digital image can be viewed as two-dimensional matrix of discrete points called pixels, in a gray level image each element is an approximation of the gray level of the image at the point sampled to a value in the set 0, 1, ..., L-1. Pixels with values between 0 and L-1 will be displayed in different shades of gray, which will be darker as closest to zero their values are [1]. Many image processing problems use gray level images because of their simplicity over color images, for that reason, color images are converted to grayscale images before the processing algorithm is applied. The most adopted gray level is 256, i.e. each pixel has a value between 0 and 255.

Multi thresholding is a process that determine more than one threshold for a given image, in order to segment a grey level image into one background and several objects. The method works very well objects with complex background which is our case since we are analyzing the brain structure.

A histogram shows the graph of pixel intensity of image, it helps to select the threshold level to separate the image into groups [1].

However, for multi-threshold segmentation, the possible combinations the follow the calculation: $L \times (L-1) \times ... \times (L-t+1)$, where is the number of limits. It is obvious that The naive of trying all the thresholds is not computationally efficient or feasible as the value of t increases. In fact, a computational question of the naive approach the use of faster algorithms to find t thresholds.

In natural images, one can see that same objects have similar pixels and different objects pixels, this will be taken into consideration is the inter-class variance and intra-class variance; that is the variance in gray level between the pixels belong to one object and variance between pixels in different objects. Such variance represents higher values than the mean. Generally, the pixels that represent an object in the image tend to have the color level, having their minimum variance. Thus, if it is a greater variance, it is not an object. In the proposed algorithm, the calculation of the fitness based on variance in the color intervals. For example, with 2 thresholds, values 75 and 150, the variation of color levels between 0 and 75, 75 and 150, 150 and 255 is calculated. Taking the first interval, and its variance to the low, the probability of having an object there is great. The genetic algorithm searches for the best thresholds that minimize variance.

The computational complexity of naïve approach, and the existence of goodness measure in the case of multiple thresholds motivated the use of an efficient search algorithm.

3 Genetic Algorithms

Genetic algorithms(GA) are probabilistic optimization methods which are based on the principles of evolution. A genetic algorithm has commonly four stages: population initialization, evaluation of fitness, reproduction and termination.

The search is done in a way that mimics evolution, in the method a “population” of possible solutions is formed, and a set of new solutions are formed by “breeding” the best solutions from the population’s members, to form a new generation, the population evolve in many generations; till the best solution is returned.

Those proposed solutions are called a chromosome and each chromosome is built up from genes, the number of solutions is the population size. The representation a chromosome has need to suit the problem, bad representation may cause false positive solutions to the problem. To choose the best solution a function is used, the fitness function evaluates the population in order to select the best chromosomes, the way fitness its calculated changes with different representations of the chromosome.

In the reproduction part two of the chosen solutions, merge their genes to form new solutions, two good parents produce good children, creating in this way, an entirely new population, this part it’s also called crossover to overcome the problem of local minima, that the crossover may cause, the mutation is used, where a gene is selected randomly and its value is changed.

The iteration of the algorithm stops when a certain criterion is met, the most used stopping criteria is number of iterations. When a defined number of iterations is reached, the algorithm stops, and the best solutions are showed. The process is ilustred below in Fig.1:

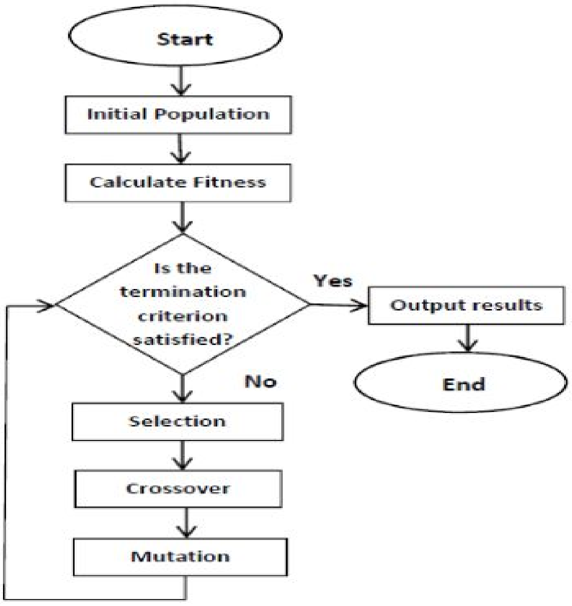

Figure 1. Fluxogram of GA [1]

4 The Proposed Algorithm

The algorithm used was inspired by the paper of et.al. Banimelhem [1], where the chromosomes are represented has a vector of bits. The Fig. 2 represents a chromosome's structure where each vector has $L*n$ bits in which L is the logarithm of the number of gray levels, and n is the number of thresholds used. Each value of L, it's a bit that represents a threshold value.

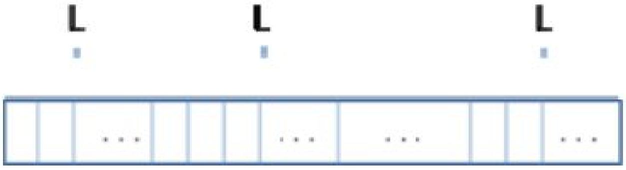

Figure 2. Chromosome structure [1]

First, is found the histogram of the image. The histogram gives information that will be used to evaluate the fitness function. Then, are generated randomly K initial chromosomes. After that, the genetic algorithm iterates for fixed number of iterations, and finally the best chromosome will be selected as the solution.

The pseudo code for the proposed algorithm as showed in Fig. 3.

Inputs: image Im, population size, crossover rate, mutation rate, number of iterations, number of thresholds

Outputs: segmented image

```
Im = Read Image
Hist = image hsitogram(Im)
Pop = initial K chromosomes
for i = 1 to number of iterations
    while (size(Pop) > size(NewPop))
    Ch1, Ch2 <- Select two chromosoms by rolletwheel
        if(rand < crossover rate) then
            ChNew1, ChNew1 <- crossover(ChNew1)
            if(rand < mutation rate) then
                ChNew1 <- mutate(ChNew1)
            if(rand < mutation rate) then
                ChNew2 <- mutate(ChNew2)
    if(fit(ChNew1))>max(fit(Ch1),fit(Ch2)) then
        NewPop.add(ChNew1)
    if(fit(ChNew2)>max(fit(Ch1),fit(Ch2)) then
        NewPop.add(ChNew2)
```

Figure 3. Genetic algorithm pseudo code [1]

In general, the pixel variance between pixels from different objects is large, while between pixels within the same object is small. For measure the fitness of a given threshold the fitness function is[2]:

$$F = S_{\text{Between objects}} / S_{\text{within objects}} \qquad (1)$$

In Eq. (1), we have the variance between objects divided by the variance within objects.

5 Data Base

To obtain the images for segmentation 3D Slicer was utilized, which is a software for medical image analysis and visualization. In the program the brain can be viewed in the three anatomical planes: coronal, transverse or horizontal and sagittal. The images used in this article are views from the horizontal plane of the brain. There are two types of images, from ten different patients: MRI scan acquired before surgery and MRI scan acquired during tumor resection.

6 Implementation and Results

The algorithm was implemented in GNU Octave. The initial population size is twenty chromosomes, number of iterations is 50, mutation and crossover rates are 0.1 and 0.8, respectively and three thresholds are used since the segmentation aims to find three tissues in MRI of our database. It was run on 2.27 Ghz core i3 processor

The algorithm showed ability segment image with high quality and was capable to segment them and find the anatomical structures of interest. The results shown below in Fig. 5 and Fig. 6 were obtained with the limits thresholds values: 37, 67, 127 in a patient image of the database, validating all the applied process, because for each limit a specific object was found, as shown in the Fig. 4, separately.

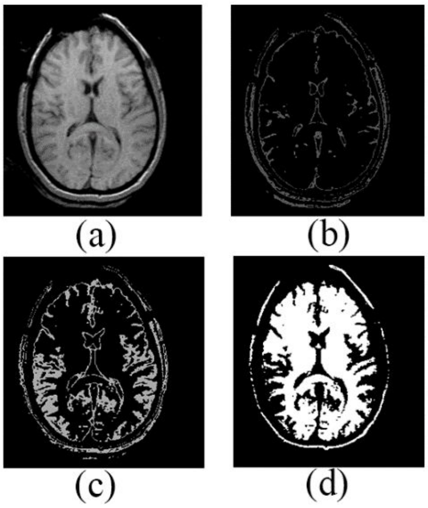

Figure 4. (a) Patient MRI during in tumor resection (b) cerebrospinal fluid (c) grey matter (d) white matter

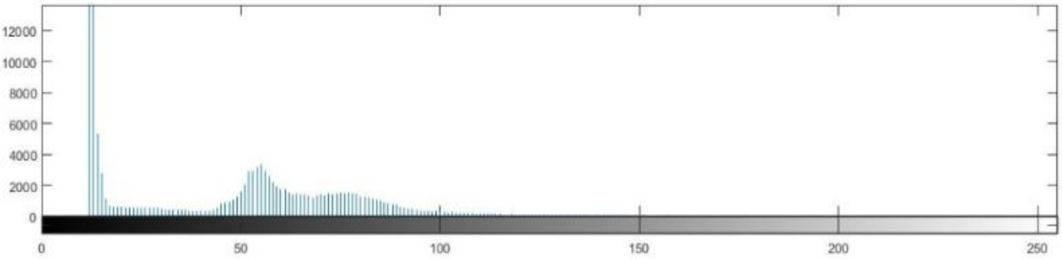

Figure 5: histogram of original image (a)

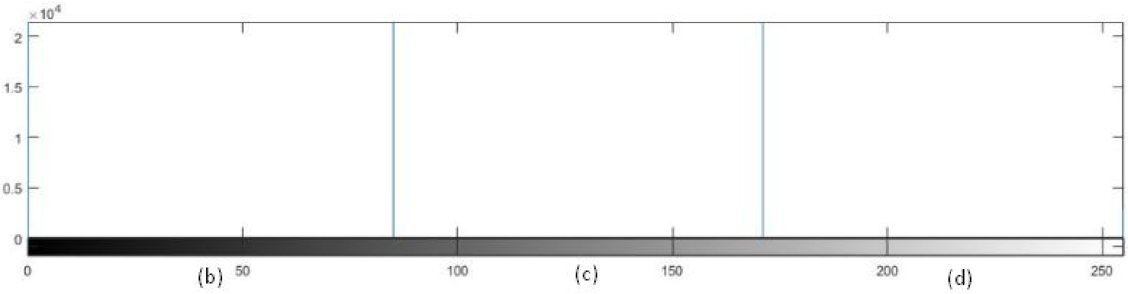

Figure 6. (b) histogram image and threshold at 37 levels(c) histogram image and threshold at 67 levels (d) histogram image and threshold at 127 levels

Below is a Table 1 containing the multi-threshold results for 10 patient images, as well as their execution time.

Table 1: Thresholds selected by the proposed algorithm

Patients	Thresholds			Time(seconds)
	T1	T2	T3	
P1_T2	36	70	104	37.45
P2_T1	39	97	209	32.71
P3_T2	50	93	120	31.40
P4_T2	33	63	93	30.40
P5_T2	45	166	205	33.25
P6_T2	72	131	193	42.85
P7_T1	59	112	158	35.54
P8_T2	55	89	120	35.17
P9_T2	43	95	146	35.46
P10_T2	50	80	126	31.70

7 Conclusion

The proposed algorithm showed how powerful the genetic algorithms are in the process of image segmentation, both in terms of processing and computational cost. It was able to segment the image into different tissues of the brain, showing how powerful the algorithms are; however, failed to segment high-contrast images to a satisfactory degree of separation between the parts, which is a known disadvantage of multi-threshold segmentation.

It can also be concluded that working with concepts of the image statistic, such as histogram, was quite advantageous and the algorithm won in terms of performance, since the problem in question ceased to be simply image processing and became statistical . In future work, different brain parameters and visualizations could be used to improve the result, as well as applying different preprocessing methods in the image prior to the segmentation step.

References

[1] BANIMELHEM, Omar; YAHYA, Yahya Ahmed. Multi-Thresholding Image Segmentation Using Genetic Algorithm. WorldComp, 2011.
[2] R. Gonzalez and R. Woods, "Digital image processing", third edition, Pearson Prentice Hall, 2008, pp. 738-761.
[3] Y.Rui, A.C. She, T.S. Huang, ―Modified Fourier descriptors for shape representation—a practical approach‖, In Proceedings of the First International Workshop on Image Databases and Multi Media Search, pp. 284-295, 1996.