Advances in Mechanical Engineering and its Applications (AMEA)
Vol. 2, No. 3, 2012, ISSN 2167-6380

www.worldsciencepublisher.org

2D Maximum Entropy Method for Image Thresholding Converge with Differential Evolution

Sushil Kumar, Millie Pant, A.K. Ray

Indian Institute of Technology Roorkee, India

Email: kumarsushiliitr@gmail.com, millidma@gmail.com, amiayakumarray@gmail.com

Abstract – Threshold segmentation is a simple and important method in grayscale image segmentation. Segmentation refers to the process of partitioning a digital image in order to locate different objects and regions of interest. Gray information of an image utilizes by Maximum Entropy threshold segmentation method.2D entropy computation can be increased by Differential Evolution method. This paper proposed a new method DE-, using entropy as the basis for measuring the optimum threshold value for the image.

Keywords – Entropy; Image Segmentation; Thresholding; Differential Evolution

.

1. Introduction

The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer [8, 9]. In a broader context, it implies digital processing of any two-dimensional data. A digital image is an array of real numbers represented by a finite number of bits. An image may be defined as a two-dimensional function, f(x,y), where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point. When x,y, and the intensity values of f are all finite, discrete quantities, then image is called a digital image [9].

Image segmentation is a basic technology in image processing and early vision, and is also an important part of most systems of image analysis and vision. At present, there are many thresholding segmentation methods, e.g. histogram thresholding, Ostu, maximal entropic thresholding and so on. Maximal entropic thresholding method doesn't need prior knowledge and can separate the images which have nonideal double-humped histograms, but it needs too much computation when evaluating the optimal thresholding [12].

The focus of the present study is thresholding technique which is perhaps simplest but effective segmenting methods. CT and MRI are grayscale images. Image segmentation is to separate pixel of the image into some non-intersecting regions [11]. Image segmentation is the base of image 3D reconstruction. It is a critical step in image processing. At present, there isn't a common segmentation algorithm. According to features of image and application purposes, different algorithms are used in application.

In the present study a simple and effective method based on Differential Evolution is proposed for image thresholding. In the proposed method called DE-based Entropy method, differential evolution is embedded in the 2D maximum entropy method to obtain the optimized threshold value.

The remaining of the paper is organized as follows; in the next section we give a brief description of Differential Evolution Algorithm; in section III, 2D Maximum Entropy method is described. The proposed DE-based algorithm is given in section IV. Results are discussed in section V, and the paper finally concludes with section VI.

2. Differential Evolution

DE, a kind of Evolutionary Algorithm (EA), was proposed by Storn and Price in [1] as a heuristic optimization method for minimizing nonlinear and non-differentiable continuous space functions having real-valued parameters. The most important characteristic of DE is that it uses the differences of randomly sampled pairs of object vectors to guide the mutation operation instead of using probability distribution functions as other EAs. The distribution of the differences between randomly sampled object vectors is determined by the distribution of these object vectors. Since the distribution of the object vectors is mainly determined by the corresponding objective function's topography, the biases where DE tries to optimize the problem match those of the function to be optimized. Consequently DE results as a robust and more generic global optimizer in comparison to other EAs [2].

According to Price [3], the main advantages of a DE include,

- Fast and simple for application and modification
- Effective global optimization capability
- Parallel processing nature
- Operates on floating point format with high precision
- Efficient algorithm without sorting or matrix multiplication
- Self-referential mutation operation
- Effective on integer, discrete and mixed parameter optimization
- Ability to handle non-differentiable, noisy, and/or time-dependent objective functions
- Operates on flat surfaces
- Ability to provide multiple solutions in a single run and effective in nonlinear constraint optimization problems with penalty functions

DE relies on the initial population generation, mutation, recombination and selection through repeated generations till the termination criteria is met. The fundamentals of DE are introduced accordingly in the sequel.

DE is a parallel direct search method using a population of N parameter vectors for each generation. At generation G, the populationP^G is composed of $X_i^G, i = 1, 2, \dots N$. The initial population P^{G0} can be chosen randomly under uniform probability distribution if there is nothing known about the problem to be optimized.

$$\boldsymbol{X}_i^G = \boldsymbol{X}_{i(L)} + rand_i[0,1].\left(\boldsymbol{X}_{i(H)} - \boldsymbol{X}_{i(L)}\right) \quad (1)$$

Where $\boldsymbol{X}_{i(L)}$ and $\boldsymbol{X}_{i(H)}$ are the lower and higher boundaries of dimensional vector $x_i = \{x_{j,i}\} = \{x_{1,i}, x_{2,i}, \dots, x_{d,i}\}^T$. If some *a priori* knowledge is available about the problem, the preliminary solution can be included to the initial population by adding normally distributed random deviations to the nominal solution.

The key characteristic of a DE is the way it generates trial parameter vectors throughout the generations. A weighted difference vector between two individuals is added to a third individual to form a new parameter vector. The newly generated vector will be evaluated by the objective function. The value of the corresponding objective function will be compared with a predetermined individual. If the newly generated parameter vector has lower objective function value, it will replace the predetermined parameter vector. The best parameter vector is evaluated for every generation in order to track the progress made throughout the minimization process. The random deviations of DE are generated by the search distance and direction information from the population. Correspondingly, this adaptive approach is associated with the normally fast convergence properties of a DE.

For each parent parameter vector, DE generates a candidate child vector based on the distance of two other parameter vectors. For each dimension j∈[1, d], this process is shown, as is referred to as scheme DE 1 by Storn and Price.

$$\boldsymbol{x}' = \boldsymbol{x}_{r3}^G + F.(\boldsymbol{x}_{r1}^G + x_{r2}^G) \quad (2)$$

Where, the random integers$r_1 \neq r_2 \neq r_3 \neq i$are used as indices to index the current parent object vector. As a result, the population size N must be greater than 3. F is a real constant positive scaling factor and normally $F \in$ (0, 1+). F controls the scale of the differential variation $(\boldsymbol{X}_{r1}^G - \boldsymbol{X}_{r2}^G)$[1], [3].

Selection of this newly generated vector is based on comparison with another DE1 control variable, the crossover constant $CR \in [0,1]$ to ensure the search diversity. Some of the newly generated vectors will be used as child vector for the next generation, others will remain unchanged. The process of creating new candidates is described in the pseudo code shown in Fig. 1, [1] [3], [4-6].

```
Mutation and recombination
For each individual j in the population
Generate 3 random integers, r1, r2, and r3 ∈(1,N) and r1 ≠ r2 ≠ r3 ≠ j
Generate a random integer irand∈ (1,N)
    For each parameter i
    If rand(0,1) <CR or i = irand
        x'_{i,j} = x_{i,r3} + F.(x_{i,r1} − x_{i,r2})
        Else
            x'_{i,j} = x_{i,j}
        End If
    End For
```

3. Two-Dimensional Maximum Entropy Thresholding Method

The maximum entropy principle states that, for a given amount of information, the probability distribution which best describes our knowledge is the one that maximizes the Shannon entropy subjected to the given evidence as constraints [13-14].

The two-dimensional entropy thresholding segmentation method separate an image with the thresh holding based on two-dimensional histogram which is formed of the gray level values of pixels of the image and the average gray level values of their neighborhoods [15]. For example, there is a gray-level image with L level, the size of which is × n , and its neighborhood with k × k size also has L level. Then, the total number of its pixel can be evaluated in terms of the equation: = m× n . The two-dimensional histogram of this image is calculated by the following formula:[10]

$$h_{ij} = p_{ij} = \frac{f_{ij}}{N}, \qquad 0 \leq i, j \leq L - 1 \quad (3)$$

Where,

- i is the gray level value of a pixel;
- j is the average gray level value of a pixel's neighborhood;
- f_ij is the number of the pixels, of which the gray level values are i and the average gray level values of their neighborhoods are j;
- $\left(p_{ij} \middle| i, j = 0,1,2, \dots L - 1\right)$ is the two-dimensional histogram about the gray level value of each pixel

and the average gray level value of its neighbourhood.

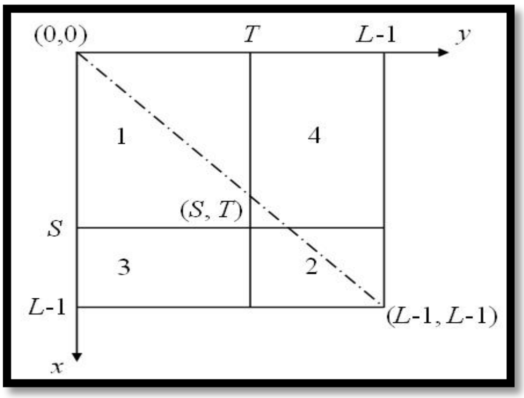

Figure 1. Two dimensional Histogram

In the figure, field-1 and field-2 along the diagonal represent the objects and the background of an image; field-3 and field-4 away from the diagonal represent the boundaries and noises in the image. Accordingly, we should find the optimal thresholding in field-1 and field-2 by using of the two-dimensional maximal entropic thresholding segmentation method in order to make the information representing the object and background maximal [10].

Suppose that field-1 and field-2 have different probability distribution. Then, the probability of field-1 and field-2 is normalized using their posterior probability, which is represented as pij. Suppose the thresholding is (s,t), then:

$$p_1 = \sum_{i=0}^{s}\sum_{j=0}^{t} p_{ij} \qquad p_2 = \sum_{i=s+1}^{L-1}\sum_{j=t+1}^{L-1} p_{ij} \tag{4}$$

According to Shannon's entropy function we can define two-dimensional discrete entropy as follows:

$$H = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p_{ij} \tag{5}$$

Then, the two-dimensional entropy of object and background of an image are appointed by following:

$$H_1 = -\sum_{i=0}^{s}\sum_{j=0}^{t} p_{ij}/p_1 \lg\frac{p_{ij}}{p_1} = lg\, p_1 + \frac{H_1}{p_1} \tag{6}$$

$$H_2 = -\sum_{i=s+1}^{L-1}\sum_{j=t+1}^{L-1} p_{ij}/p_2 \lg\frac{p_{ij}}{p_2} = lg\, p_2 + \frac{H_2}{p_2} \tag{7}$$

In the formula, $H1$ and $H2$ are calculated as follows:

$$H_1 = -\sum_{i=0}^{s}\sum_{j=0}^{t} p_{ij}\, logp_{ij}$$

$$H_2 = -\sum_{i=s+1}^{L-1}\sum_{j=t+1}^{L-1} p_{ij}\, logp_{ij}$$

Because field-3 and field-4 represent the information about boundaries and noises, they can be ignored and then the following equations are obtained:

$$p_2 = 1 - p_2, \quad H_2 = 1 - H_1 \tag{8}$$

Then, the following equation is deduced through equation (10) and (11):

$$H(2) = \lg(1 - p_1) + H - H_1/1 - p_1 \tag{9}$$

So, discriminant function about entropy can be defined as follows

$$\boldsymbol{\Psi(s,t) = H(A) + H(B) = \lg[p_1(1-p_1)] + \frac{H_1}{p_1} + \frac{H_2}{p_2}} \tag{10}$$

At last, select the optimal thresholding by following formula:

$$\Psi(s^*,t^*) = max\,\{\Psi(s,t)\} \tag{11}$$

4. Methodology

In the previous section, we saw that the objective function may be formulated as an optimization problem, with the help of 2D Entropy method. After constructing the objective function, DE is used to solve it and to obtain the optimized values. It is a simple and easy to implement process in which we first of all take an image which has not a clear valley in its histogram. With the help of 2D Entropy we construct the objective function and optimize it with the help of DE to obtain the threshold value of entropy. A step-by-step procedure is explained here:

Step 1: The function of (s,t) will be considered as an objective function for Differential Evolution. Size of the population initializing 0-255 a range of a gray image intensity values. $X_i(0)$will provide a fitness value of $f(X_i(0))$. Compute the personal fitness for each population member.

Step 2: Apply mutation operation on according (2) and compute fitness value for each mutant population member, compute the fitness value according these updated vectors.

Step 3: Update it with the best fitness value.

Step 4: Then apply selection procedure on the known fitness values. It will provide us a threshold value of the given image.

6. Results

The proposed method is validated on a set of three gray level images; Lena, Coins and Brain. Table shown a comparison between SOPSO [10] calculated entropy and DE based entropy.

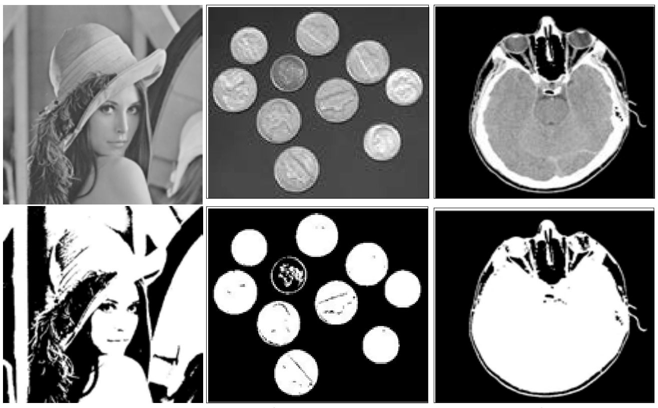

Figure 2. (a) Original Image of Lena (b) Segmented image of Lena (c) Original image of Coins (d) Segmented image of Coins (e) Original image of Brain MRI (f) Segmented image of Brain MRI

Table 1. Comparison between SOPSO and DE-Based Algorithm

Algorithm	Optimal Threshold	Maximal Entropy	Minimum Steps	Maximum Steps	OOT
SOPSO(Lena)	(126,113)	136383	11	83	16
DE-Based(Lena)	(124,115)	136475	8	70	13
SOPSO(coins)	(129,119)	116331	19	65	15
DE-Based(coins)	(124,113)	116331	10	48	9
SOPSO(MRI)	(128,111)	128972	21	95	19
DE-Based(MRI)	(123,112)	128972	16	72	12

6. Conclusion and Future Scope

2D maximum entropy method is well known for obtaining threshold value for the images, having a well-defined thresholding value with the help of global thresholding method. Embedding DE in 2D maximum entropy method further improves it which helps in obtaining a better result by providing optimum threshold value. The 2D maximum entropy method in its basic form however has a drawback that computational time is very high. In future we will try to extend this work by modifying the Differential Evolution method for multilevel images as well.

Acknowledgements

The first author acknowledges with thanks the financial support provided by MHRD.

References

[1] R. Storn and K. Price, "Differential Evolution – A simple and efficient adaptive scheme for global optimization over continuous spaces", *Technical Report TR-95-012,* March 1995, ftp.ICSI.Berkeley.edu/pub/techreports/1995/tr-95-012.ps.Z.

[2] Wong, K.P. and Dong, Z.Y, "Differential Evolution, an Alternative Approach to Evolutionary Algorithm". In K. Y. Lee and M. A. El-Sharkawi (Ed.) 2007, *Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems* (pp. 171-187) USA: John Wiley & Sons, Ltd.

[3] K. Price, "An introduction to differential evolution", Chapter 6, in D. Corne, M Dorigo and F. Glover, edt. *New Ideas in Optimization*, McGraw Hill, 1999, pp.79-108.

[4] K. Price, "Differential evolution: a fast and simple numerical optimizer". In M. Smith, M. Lee, J. Keller, and J. Yen (eds.) *1996 Biennial Conference of the North American Fuzzy Information Processing Society*, IEEE Press, New York, NY, pp. 524−527.

[5] K.V. Price, "Differential evolution vs. the functions of the 2nd ICEO", *Proc. IEEE International Conference on Evolutionary Computation*, 13-16 April 1997, pp. 153 – 157.

[6] R.L. Becerra and C.A. Coello Coello, "Culturizing differential evolution for constrained optimization", *Proc. The fifth Mexican International Conference in Computer Science (ENC'04)*, 2004, pp. 304 – 311.

[7] N. Otsu, "A threshold selection method from gray-level histogram," IEEE Transactions on System Man Cybernetics, 1979 pp. 62-66.

[8] Anil K. Jain, "Fundamentals Of Digital Image Processing", Prentice-Hall, 1989.

[9] R.C. Gonzalez, R.E. Woods, "Digital Image Processing", Prentice Hall, 2009.

[10] Liping Zheng, Quanke Pan, Guangyao Li, Jing Liang, "Improvement of Grayscale Image Segmentation Based On PSO Algorithm" proc. *Fourth International Conference on Computer Sciences and Convergence Information Technology,2009,* pp.442-446.

[11] Pal N R., Pal S K. "A review on image segmentation techniques", 1993, Pattern Recognition.

[12] Xiujuan Lei, Ali Fu, "2-D Maximum-entropy Thresholding Image Segmentation Method Based on Second-order Oscillating PSO", ICNC,2009,pp.161-165.

[13] A butaleb A S. Automatic thresholding of gray-level pictures using two-dimension entropy. Computer Vision Graphics and Image Processing, 1989, 47(1):22~32.

[14] Kennedy J, Eberhart R C. Particle swarm optimization. International Conference on Neural Net-works Prodeedings, 1995:1942~1948.

Vitae

Sushil Kumar was born in 1982. He obtained a B. Tech. degree in Computer Science and Engineering in Computer Science and Engineering department from Uttar Pradesh Technical University in 2006. Obtained his M. Tech degree in Computer Science and Engineering from Maulana Azad National Institute of Technology, Bhopal(India) in 2008.

He worked as a Trainee software engineer in Impetus InfoTech Pvt. Ltd during 2008.

He is working as a Research Scholar in Indian Institute of Technology Roorkee, Roorkee(India). His research interest includes Image Processing, Soft Computing.