Image Thresholding Using Differential Evolution

Shahryar Rahnamayan, Hamid R. Tizhoosh, Magdy M.A. Salama
Medical Instrument Analysis and Machine Intelligence Research Group
Faculty of Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
shahryar@pami.uwaterloo.ca, tizhoosh@uwaterloo.ca, msalama@hivolt1.uwaterloo.ca

Abstract - Image thresholding is a challenging task in image processing field. Many efforts have already been made to propose universal, robust methods to handle a wide range of images. This paper introduces a new optimization-based thresholding approach. The optimizer, Differential Evolution (DE) algorithm, minimizes dissimilarity between the input grey-level image and the bi-level (thresholded) image. The proposed approach is compared with a well-known thresholding method, Kittler algorithm, through subjective and objective assessments, and experimental results are provided.

Keywords: Image thresholding, Differential evolution, Optimization, Kittler, Objective assessment, Subjective assessment

1 Introduction

In many image processing applications, the crucial role of the image thresholding can be observed (e.g. medical image processing [1]). Numerous thresholding techniques have already been proposed [2]. However, almost all of them are application- or domain-oriented solutions suffering from lack of universality. Therefore, this research filed is still open to investigate and introduce new robust and universal techniques.

In this paper, a new thresholding technique is proposed which generates corresponding binary image by minimization of dissimilarity between input and output image. Hence, the input image itself is directly used to measure the quality of the threshoded image. Thus, this method can be introduced as a candidate for universal thresholding. On the other hand, by splitting the input image in several sub-images and assigning an optimal thresholding level for each sub-image, we improve our chance to generate better thresholded image compared to applying a global thresholding technique. For those images which suffer from some disturbing factors such as non-uniform illumination or reflectance effects, this local approach can perform better.

After comprehensive evaluation of more than 40 image thresholding techniques, Sezgin and Sankur [2] concluded that the Kittler [3] is the best overall performing method. For this reason, the proposed approach is compared with Kittler algorithm, using subjective and objective assessment methodologies. The results are promising.

Organization of this paper is as follows: In section 2, a brief review of differential evolution (DE) algorithm is given. DE is employed as an optimizer in the current study. The proposed approach is presented in section 3. Experimental verification is presented in section 4 and performance assessments in section 5. The work is concluded in section 6.

2 Review of Differential Evolution

Differential Evolution (DE) is a population-based, efficient, robust, and direct search method [4,5]. We selected DE since it offers fast convergence rate and capability of working directly with real numbers (threshoding levels).

Like other evolutionary algorithms, DE starts with an initial population vector, which is randomly generated when no preliminary knowledge is available. Let assume that $X_{i,G}(i = \{1, 2, ..., N_p\})$ are N_p N_v-dimensional parameter vectors of generation G (N_p is a constant number which represents the population size) [6]. In order to generate a new population of vectors, for each target vector in population three vectors are randomly selected, and weighted difference of two of them is added to the third one. For classical DE, the mutation, crossover, and selection have straightforward procedures as follows [6]:

Mutation - For each vector $V_{i,G}$ from generation G a mutant vector $V_{i,G}$ is defined by

$$V_{i,G} = X_{r1,G} + F(X_{r2,G} - X_{r3,G}), \tag{1}$$

where $i = \{1, 2, ..., N_p\}$ and r_1, r_2, and r_3 are mutually different random integer indices selected from $\{1, 2, ..., N_p\}$. Further, i, r_1, r_2, and r_3 are different such that $n \geq 4$. $F \in [0, 2]$ is a real constant which determines the amplification of the added differential variation of $(X_{r2,G} - X_{r3,G})$. Larger values for F result in higher diversity in the generated population and the lower values in faster convergence.

Crossover - DE utilizes crossover operation to increase diversity of the population. It defines following trial vector:

$$U_{i,G} = (U_{1i,G}, U_{2i,G}, ..., U_{N_v i,G}), \tag{2}$$

where $j = 1, 2, .., N_v$ and

$$U_{ji,G} = \begin{cases} V_{ji,G} & \text{if } rand_j(0,1) \leq Cr, \\ X_{ji,G} & \text{otherwise.} \end{cases} \tag{3}$$

$Cr \in (0,1)$ is predefined crossover constant; $rand_j(0,1)$ is $j^{\text{th}} \in [0,1]$ evaluation of uniform random generator. Most popular value for Cr is in the range of $(0.4, 1)$ [7].

Selection - Now it must be decided which vector ($U_{i,G}$ or $X_{i,G}$) should be a member of new generation, $G+1$. Vector with the higher fitness value is chosen.

There are other variations of DE [8]. In this work, the classical version of DE, shown in Fig. 1, has been utilized in all experiments.

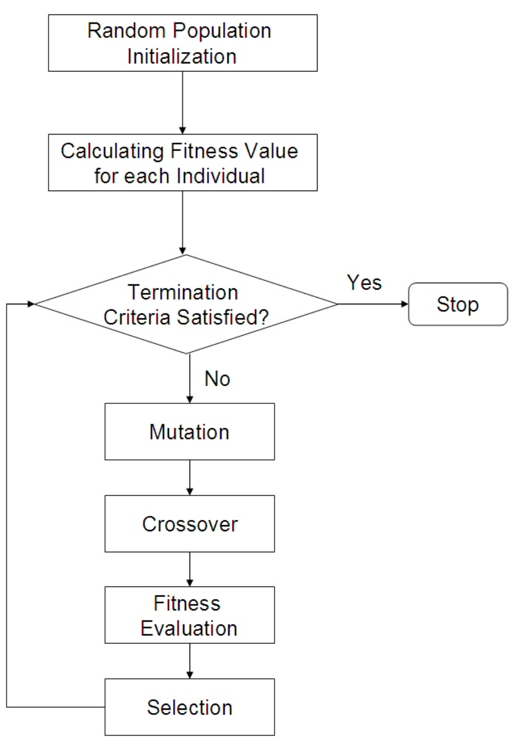


Figure 1: Classical differential evolution (DE) algorithm.

3 Proposed Approach

After a brief introduction to differential evolution algorithm, we are ready to concentrate on the proposed approach. Before starting to describe the new approach we need to define fitness function for our optimization algorithm.

Fitness function - The optimizer in our approach minimizes the dissimilarity between $M \times N$ input grey-level image, I (normalized in $[0,1]$), and the corresponding generated bi-level thresholded image, $B \in \{0,1\}$. The fitness function f can be defined as follows:

$$f = \sum_{i=1}^{M} \sum_{j=1}^{N} |I_{ij} - B_{ij}|. \tag{4}$$

Minimization of this fitness function means mapping darker pixels from grey-level image to the background and lighter pixels to the foreground in the threshoded image.

Individual steps of the proposed approach are as follows:

(1) Splitting input image I in n sub-images: $I_1, I_2, \cdots, I_n$.

(2) Assigning a threshold level for each sub-image: $T_1, T_2, \cdots, T_n$.

(3) Applying DE algorithm to find optimal values for $T_1, T_2, \cdots, T_n$ by minimizing the fitness function (see Eq. 4).

(4) Thresholding the sub-images by using optimized values from step (3).

(5) Assembling thresholded sub-images to obtain final threshoded image.

By splitting the input image in n sub-images and assigning an optimal thresholding level for each sub-image, we improve our chance to have better results compared to applying global thresholding methods. Table 1 shows details of proposed approach by a pseudo-code representation. In conducted experiments we heuristically set $n = 4$.

4 Experimental Verification

In order to investigate the effect of the proposed approach, 26 hard to threshold images were selected to build our test set; all of them are frequently used images in the image processing literature.

The following DE control parameters are set for all conducted experiments with no attempt to achieve optimal values for them.

Table 1: Proposed approach: Image thresholding using DE algorithm.

```
Begin
    Splitting Input Image, I, in n Sub-images: I_1, I_2, ..., I_n;
    Random Population Initialization;
/* Each individual in the population has n variables: T_1, T_2, ..., T_n (thresholding level of sub-images) */
    Calculate Fitness Value for each Individual in the Population;   /* see Eq. 4 */
      while (satisfying termination criteria)
            /* DE evolution steps (mutation, crossover, and selection) */
            Mutation;   /* see Eq. 1 */
            Crossover;   /* see Eq. 2 and Eq. 3 */
            Selection;   /* see Sec. 2 */
        Calculate Fitness Value for each Individual in the Current_Population;   /* see Eq. 4 */
      while end
    Thresholding Sub-images by Using Optimized Values for Thresholding Levels;
    Assembling Thresholded Sub-images to Obtain Final Threshoded Image;
End
```

- Population size, $N_p = 40$
- Differential amplification factor, $F = 0.9$
- Crossover probability constant, $Cr = 0.9$
- Strategy [8]: DE/rand/1/bin
- Maximum function calls, NFC_{MAX}=1000
- Number of sub-images, $n = 4$

Some sample results of applying the proposed approach are presented in Table 2. As shown, also corresponding ground-truth image (created manually) and the result of Kittler method are given for comparison. A sample fitness plot for one of the test images is given in Fig. 2.

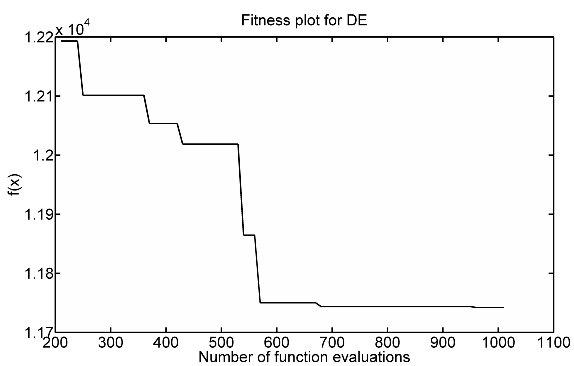


Figure 2: A sample fitness plot (minimization) for a test image.

In following section, the results of the proposed approach and Kittlers' have been compared using two completely different methods, namely, objective and subjective assessment.

5 Performance Assessment

A wide range of image quality measures have been proposed by image processing researchers [9–11]. In this section, results of Kittler and the proposed approach are compared by human judgment (subjective assessment) and also by reference based objective assessment. Reference or ground-truth images have been manually prepared to serve as gold/ideal thresholded image for each test image.

5.1 Objective Assessment

In order to compare the performance of different methods, an objective metric is required. To compare two binary images, Misclassification Error (ME) [2,12] can be a reasonable and straightforward measure to use. It calculates percentage of foreground pixels which assigned wrongly to background and vice versa:

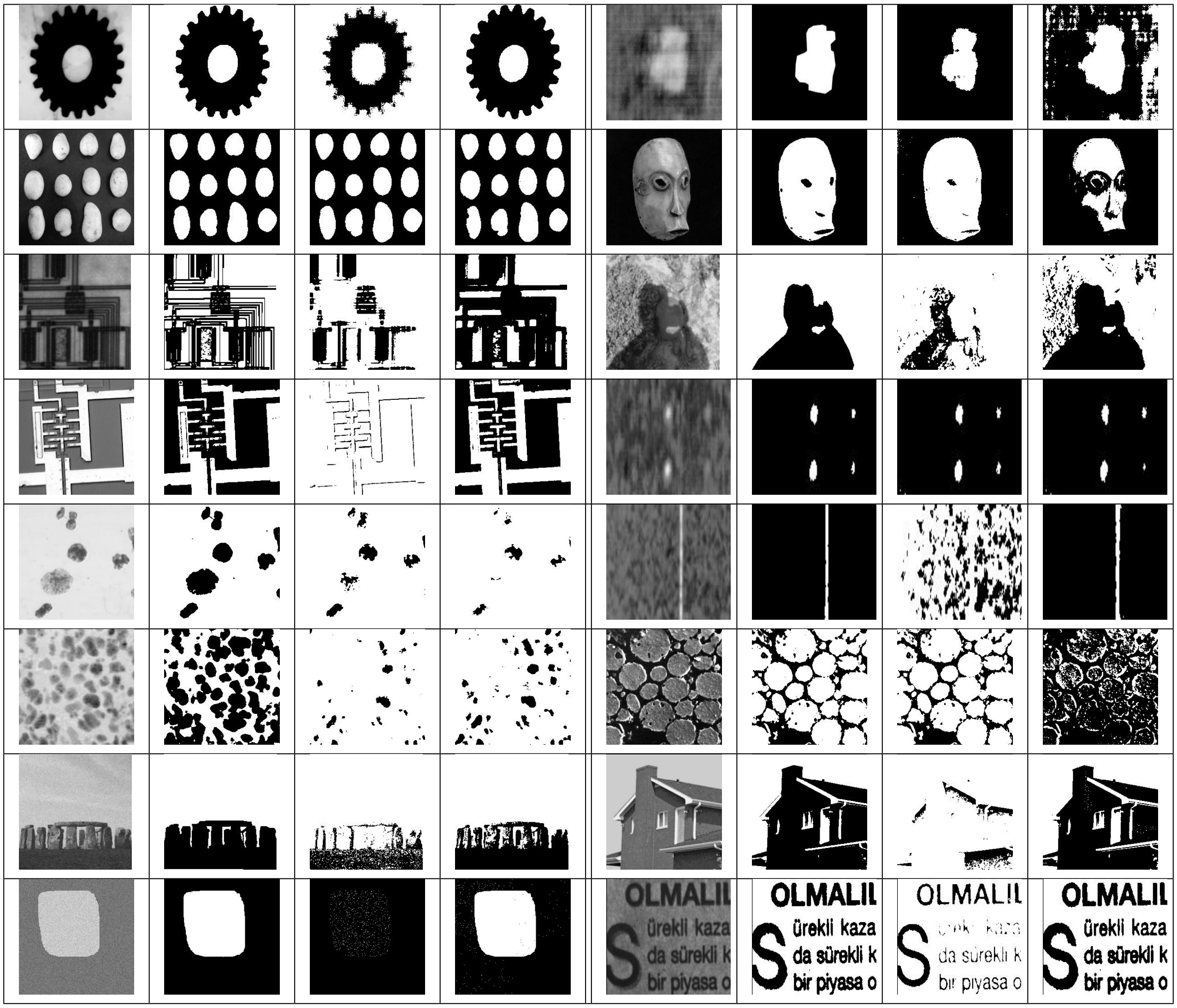
$$ME = \frac{|B_O \cap F_T| + |F_O \cap B_T|}{|B_O| + |F_O|}, \tag{5}$$

where B_O, F_O, B_T, and F_T are the background and foreground pixels of the ground-truth image and the background and foreground pixels of the test image, respectively. $|.|$ denotes the cardinality of the set.

By utilizing this error measure, similarity index, η, can be defined as follows:

Table 2: Some experimental results. Input image, corresponding manually created ground-truth (gold) image, result of Kittler method, and result of the proposed approach (DE) for some test images.

Image	Ground-truth	Kittler	DE	Image	Ground-truth	Kittler	DE



$$\eta = (1 - ME) \times 100\%. \tag{6}$$

The similarity index of 100% means complete matching of two images. Also, the overall similarity index $\bar{\eta}$ (or generalization index) for a set of images is defined as follows:

$$\bar{\eta} = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n} \eta_i, \tag{7}$$

where n_{test} is the number of test images.

Table 3 summarizes the results of objective assessment for 26 test images. As seen, the overall similarity index, $\bar{\eta}$, for the new approach (89.16%) is higher and the standard deviation (12.58%) is lower than Kittlers' $(81.53\%,\ 20.08\%)$. Also, the 95% confidential intervals (CI) for the proposed approach is more compact with higher boundaries $([84.07, 94.24])$ than Kittlers' $([73.42, 89.95])$.

The best result in each case has been indicated in bold. For 9 cases (34.61%) the Kittler shows better results; in 12 cases (46.15%) DE performs better; and for 5 cases (19.23%) the results are almost the same. Results with difference less than 1% are assumed to be the same.

Table 3: Results of objective assessment for 26 test images. The best result in each case has been highlighted in bold (for difference less than 1%, the results are assumed to be the same). η: similarity index, $\bar{\eta}$: overall similarity index (for 26 test images), σ: standard deviation, K: Kittler, DE: proposed method, 95% CI: 95% confidential intervals.

Image No.	η_K	η_{DE}	Image No.	η_K	η_{DE}
1	**98.07**	86.61	14	100.0	100.0
2	89.88	**97.85**	15	**92.83**	90.05
3	**99.44**	98.40	16	28.14	**99.25**
4	**98.32**	84.48	17	99.83	99.57
5	77.84	**79.85**	18	**97.38**	93.60
6	78.19	**91.18**	19	52.98	**57.78**
7	81.80	**93.87**	20	**97.16**	49.76
8	92.42	92.22	21	99.68	99.51
9	**81.11**	80.20	22	59.40	**78.68**
10	92.51	92.66	23	48.68	**97.97**
11	**93.47**	82.45	24	71.93	**95.37**
12	**91.53**	82.49	25	72.78	**99.68**
13	41.49	**96.14**	26	83.07	**98.56**
$\bar{\eta}_K$=81.53%, σ_K=20.08%					
$95\%\mathrm{CI}_K = [73.42, 89.95]$					
$\bar{\eta}_{\mathrm{DE}}$=**89.16**%, σ_{DE}=**12.58**%					
$95\%\mathrm{CI}_{\mathrm{DE}}$ = **[84.07, 94.24]**					

5.2 Subjective Assessment

In order to subjectively assess the two approaches (Kittler and DE), the results of thresholding for 26 test images were presented to 9 expert observers. They had three options to select from for each test image: Kittler (is better), DE (is better), and the (results are) same. The subjective evaluations by 9 observers are summarized in Table 4. As seen, according to evaluation of each observer and also in overall, the proposed approach shows a higher performance than Kittlers'.

For a better illustration, the subjective assessment results of 9 observers are represented as a chart in Fig.3.

Result analysis - The results of Kittler and the proposed approach for 26 test images were compared by two totaly different methodologies, subjective and objective assessments. In both cases, the proposed approach achieved a higher performance than Kittler method. Table 5 shows in how far objective and subjective evaluations differ. As seen, interestingly, the difference between objective and subjective assessments for all three cases is less than 3%.

Table 4: Results of subjective assessment by 9 expert observers. The best result in each case has been highlighted in bold. n_K: number of images (out of 26 test images) for which Kittler performs better, n_{DE}: number of images for which DE performs better, n_{S}: number of images for which Kittler and DE results are almost the same ($\bar{n}$ is the average value calculated among 9 observers).

Observer No.	n_K	n_{DE}	n_S
1	8 (30.76%)	**14** (53.84%)	4 (15.38%)
2	9 (34.61%)	**13** (50.00%)	4 (15.38%)
3	12 (46.15%)	**13** (50.00%)	1 (3.84%)
4	12 (46.15%)	**14** (53.84%)	0 (0%)
5	4 (15.38%)	**10** (38.46%)	12 (46.15%)
6	5 (19.23%)	**11** (42.30%)	10 (38.46%)
7	6 (23.07%)	**11** (42.30%)	9 (34.61%)
8	10 (38.46%)	**12** (46.15%)	4 (15.38%)
9	8 (30.76%)	**16** (61.53%)	2 (7.69%)
$\bar{n}_K$=31.61%, σ_K=11.01%			
$\bar{n}_{\mathrm{DE}}$=**48.71**%, σ_{DE}=**7.19**%			
$\bar{n}_S$=19.65%, σ_S=16.25%			

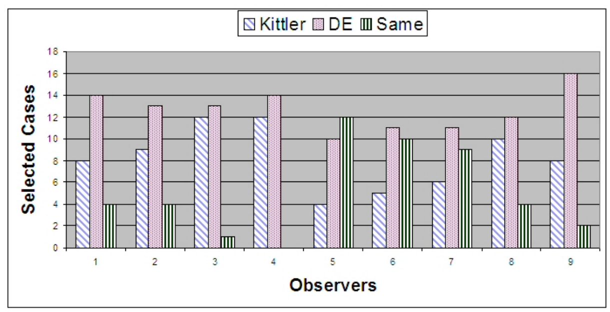


Figure 3: Chart representation of subjective assessments. Number of selected options (Kittler, DE, and Same) by each observer for 26 test images. For example, observer number 1 has selected: 8 times **Kittler**, 14 times **DE**, and 4 times **Same**.

6 Concluding Remarks

In this paper, an optimization-based image thresholding approach has been introduced. An differential evolution algorithm (DE) segmented the image into two classes by minimization of the dissimilarity between input grey-level image and binary (thresholded) image. The proposed approach was compared with a well-known method, Kittler method, through subjective and objective assessments. Both evaluations confirmed that the proposed approach is superior to Kittler algorithm.

The most important part of the proposed approach is the definition of the fitness function. As seen, DE,

Table 5: Results comparison of objective and subjective assessment. |difference|: difference between the overall result of subjective and objective assessments.

	Objective Assessment	Subjective Assessment	|difference|
Kittler	34.61%	31.61%	3%
DE	**46.15**%	**48.71**%	2.56%
Same	19.23%	19.65%	0.42%

as an optimizer, minimizes the dissimilarity between grey-level image and thresholded image. This dissimilarity is measured by pixel-by-pixel comparison of the binary and (in $[0, 1]$ normalized) grey-level images. So, it can directly affect the final result. Introducing a more accurate and also universal measure is our main direction for future work. The main drawback is that employing an evolutionary algorithm (DE) to threshold image shows a higher computational time. We are faced with this common disadvantage when we employ population-based optimization methods to solve any problem in science or engineering. In order to answer this crucial general problem, we are working to speed up convergence rate of differential evolution [13, 14].

Acknowledgements: The authors would like to thank Medical Instrument Analysis and Machine Intelligence (MIAMI) Research group members (University of Waterloo, Canada) for sharing their ideas and offering useful comments to improve the quality of this work. Special thanks go to Erik Jonasson (visiting scholar at the University of Waterloo, Canada) for conducting some experiments.

References

[1] S. Rahnamayan, H.R. Tizhoosh, M.M.A Salama, *Automated Snake Initialization for the Segmentation of the Prostate in Ultrasound Images*, Int. Conf. on Image Analysis and Recognition-ICIAR2005, Springer Lecture Notes in Computer Science Series (Springer LCNS), Toronto, Canada, pp. 930-937, 2005.

[2] Mehmet Sezgin and Bulent Sankur, *Survey over image thresholding techniques and quantative performance evaluation*, Journal of Electronic Imaging 13(1), pp. 146-165, Jan. 2004.

[3] M. I. Sezan, *A peak detection algorithm and its application to histogram-based image data reduction*, Graph. Models Image Process. 29, pp. 47-59, 1985.

[4] K. Price, R. M. Storn, J. A. Lampinen, *Differential Evolution : A Practical Approach to Global Optimization (Natural Computing Series)* Springer; 1st Edition, 2005, ISBN: 3540209506.

[5] K. Price, *An Introduction to differential evolution*, In: D. Corne, M. Dorigo, F. Glover (eds) *New Ideas in Optimization*, NcGraw-Hill, London (UK), pp. 79-108, 1999, ISBN:007-709506-5.

[6] Godfrey C. Onwubolu and B.V. Babu, *New optimization techniques in engineering*, Berlin ; New York : Springer, 2004.

[7] S. Das, A. Konar, U. Chakraborty, *Improved differential evolution algorithms for handling noisy optimization problems*, IEEE Congress on Evolutionary Computation Proceedings, Vol.2, pp. 1691-1698, 2005.

[8] R. Storn and K. Price, *Differential evolution- A simple and efficient heuristic for global optimization over continuous spaces*, Journal of Global Optimization 11, pp. 341-359, 1997.

[9] Stefan Winkler. *Vision models and quality metrics for image processing applications*, PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland, 2000.

[10] Z. Wang and A. C. Bovik, *A universal image quality index*, IEEE Signal Processing Letters, Vol. 9, No. 3, pp. 81-84, March 2002.

[11] J.B. Martens and L. Meesters, *Image dissimilarity*, Signal Processing, Vol. 70, pp. 1164-1175, Aug. 1997.

[12] W. A. Yasnoff, J. K. Mui, and W. Bacus, *Error measures for scene segmentation*, Pattern Recognition 9, pp. 217-231, 1977.

[13] S. Rahnamayan, H.R. Tizhoosh, M.M.A. Salama, *Opposition-Based Differential Evolution Algorithms*, IEEE Congress on Evolutionary Computation, to be held as a part of IEEE World Congress on Computational Intelligence, Vancouver, July 16-21, 2006.

[14] S. Rahnamayan, H.R.Tizhoosh, M.M.A. Salama, *Opposition-Based Differential Evolution for Optimization of Noisy Problems*, IEEE Congress on Evolutionary Computation, to be held as a part of IEEE World Congress on Computational Intelligence, Vancouver, July 16-21, 2006.