International Journal of Soft Computing 2 (4): 562-571, 2007

Fuzzy Clustering Based Parallel Cultural Algorithm

Jihane Alami, L. Benameur and A.El Imrani
Laboratoire Conception and Systèmes,
Faculty of Sciences, Rabat, Morocco

Abstract: This study presents a new method, which combines cultural algorithms and a fuzzy clustering technique to improve the performance of cultural algorithms in multimodal optimization. Based on the metaphor of social environment principles, this model promotes the formation and maintenance of populations and implements the concept of cultural exchanges among them. Computer simulations show good performance for several multimodal test functions.

Key words: Cultural algorithms, multimodal optimization, self-learning, fuzzy clustering

INTRODUCTION

Cultural Algorithms (CA's) are suitable for locating the global optimum of functions as they converge to a single solution of the search space. Real optimization problems, however, often lead to multimodal domains and so require the identification of multiple optima, either global or local. In this context, several methods have been presented in the literature including niching methods, methods based on immune system and techniques based on particle swarm. Most of these techniques depend on specific parameters which can reflect on the quality and on the number of detected optima.

For this purpose, a new approach called Fuzzy Clustering based Parallel Cultural Algorithm (FC-PaCA) is proposed in this study. This method extends CA's by promoting the formation of stable nations in the neighbourhood of optimal solutions, using a fuzzy clustering algorithm.

The use of fuzzy clustering model offers a natural way to deal with overlapping clusters and do not require prior information on data distribution. They present a good alternative to probabilistic methods when such information is not available.

Based on social environment where different nations can evolve and exchange some cultures among them, the proposed model promotes the formation and maintenance of nations and implements the concept of cultural exchanges among them. In the context of this approach, a nation represents a population and its belief space.

CULTURAL ALGORITHMS

Cultural Algorithms consist of a social population and a belief space (Reynolds, 1979) as shown in Fig. 1.

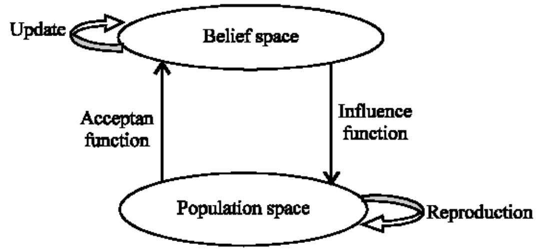

Fig. 1: Cultural algorithm framework

Selected individuals from the population space contribute to cultural knowledge by means of the acceptance function. The cultural knowledge resides in the belief space where it is stored and updated based on individual experiences and their successes or failures. In turn, the cultural knowledge controls the evolution of the population by means of an influence function. A Cultural Algorithm thereby provides a framework in which to accumulate and communicate knowledge so as to allow self-adaptation in both the population and the belief space (Angeline, 1995; Holland, 1992).

The main idea behind this approach is to preserve beliefs that are socially accepted and discard unacceptable beliefs. Therefore, if a CA is applied for global optimization, then acceptable beliefs can be seen as constraints that direct the population at the micro-evolutionary level. These constraints can influence directly the search process, leading to an efficient optimization process.

Some versions of cultural algorithms have been built, with different choices for the implementation of the micro and macro levels (Reynolds and Zannoni, 1994, 1997). In this study, the population space is supported by an Evolutionary programming system and the belief

Corresponding Author: Jihane Alami, Laboratoire Conception and Systèmes, Faculty of Sciences, Rabat, Morocco

represents both situational knowledge, which provides the exact point where the best individual of each generation was found; and normative knowledge, which stores intervals for the decision variables of the problem that correspond to the regions where good results were found (Chung, 1997). The shell is called, CAEP, Cultural Algorithm with Evolutionary Programming.

The basic idea of the CAEP is to “influence” the mutation operator (the only operator in evolutionary programming). In CAEP the belief space is initialized with the given problem domain and candidate solutions.

MULTIMODAL OPTIMIZATION

Several methods have been proposed for multimodal function optimization. These methods include the niching techniques which have been developed to maintain population diversity into genetic algorithm GA. They permit the GA to investigate multiple peaks in parallel and so prevent the GA from being trapped in local optima (Michalewicz, 1994).

Several niching techniques have been developed in the literature, including techniques of: Sharing (Goldberg and Richardson, 1987), sequential niching (Beasley *et al*., 1993), dynamic niching (Miller and Shaw, 1995), Clearing (Petrowsky, 1996) and Crowding (Mahfoud, 1995) etc.

The effectiveness of niching methods requires an a priori knowledge of the niche radius and the spatial disposition of the niches. These limitations can reflect on the number and the quality of the expected optima (Sareni and Krahenbuhl, 1998; Beasly *et al*., 1993; Petrowsky, 1997; Goldberg and Wang, 1997).

Other techniques have been recently developed using other concepts including artificial immune systems (Castro and Timmis, 2002) and techniques based on Particle Swarm Optimization PSO such Species-based Particle Swarm Optimization (SPSO) (Li, 2004) and Species Conservation Genetic Algorithm (SCGA) (Li *et al*., 2002).

Most of the multimodal optimization techniques cited above require an estimation of some parameters. For example, the effectiveness of niching methods depends on the fine-tuning of niche radius (Sareni and Krahenbuhl, 1998; Beasly *et al*., 1993). Also, techniques based on particle swarm optimization, such SPSO, introduce another parameter, which denotes the radius measured in Euclidean distance from the center of a species to its boundary and is used to classify particles into species (Li, 2004).

For this purpose, this study presents a new approach based on cultural algorithms and fuzzy clustering technique.

FUZZY CLUSTERING BASED PARALLEL CULTURAL ALGORITHM (FC-PACA)

Basic principles: The FC-PaCA is a new model inspired from social environment as depicted in Fig. 2. This figure presents the metaphor used by CA and FC-PaCA in relation with real world.

As shown in Fig. 2, cultural algorithms are based on the metaphor of real world. On the other hand, in real world, there are different nations naturally separated, which can evolve and exchange their cultures. Based on this metaphor, the FC-PaCA uses a fuzzy clustering technique in order to classify the whole population into a set of clusters. These clusters are characterized by their centers or prototypes, radius and cardinal. In the context of the proposed model, a cluster represents a nation, i.e., population and its culture and the center the nation’s elite which corresponds to the best individual and its culture in the nation. The elite of each nation represents an optimum.

The aim of this approach consists in the partition of the population into several nations, so that each of them is processed by a cultural algorithm and thus explores different regions of the search space. So that such an approach can be useful for multimodal optimization problems.

The principle of the proposed technique is based on a three-layer strategy (Fig. 3). The first layer is the basic Cultural Algorithm with Evolutionary Programming (CAEP). The output of this level represents the input for the second layer (FC). This layer is based on an unsupervised fuzzy clustering algorithm, which performs the partition of the individuals into a set of C clusters so that each of them corresponds to a nation which has a different aspect of culture. The final layer implements the principle of Spatial Separation (SS) to generate nations from the resulting cluster characteristics (centre and radius) provided by the previous layer. After separating the nations, a concept of cultural exchange is performed to neighbourhood ones.

In the context of optimization, the number and center of clusters represents the number of identified optima and the expected optima, respectively.

To validate the clustering results, several validity criteria were tested. These include the partition coefficient, the Xie and Beni index, the Fukuyama and Sugeno index, the separation/compactness index, the Bensaid index and the entropy criterion (Pal and Bezdek, 1995). Experimental results showed that, in this study, the entropy criterion (h) is the most reliable index and the

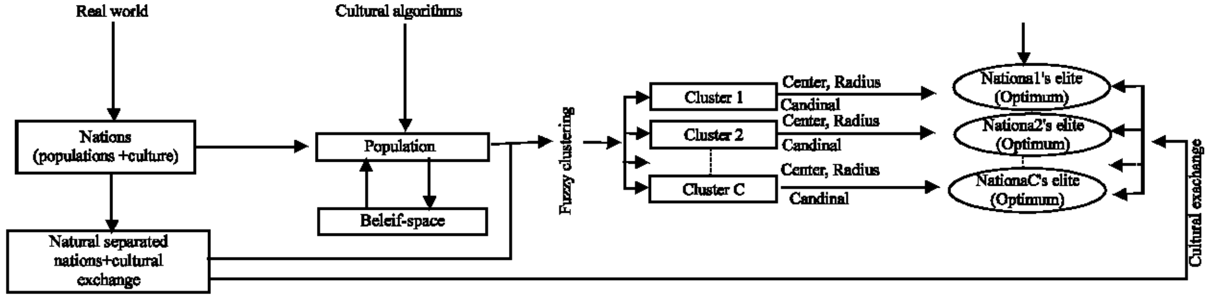

Fig. 2: Analogies between real world, CA and FC-PaCA

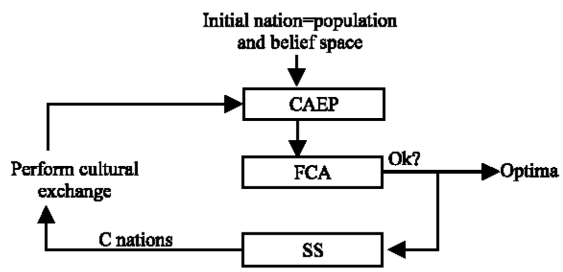

Fig. 3: Strategy of the proposed model

whole procedure of the proposed model is iterated until the entropy criterion h reaches a prefixed minimum value ($<10^{-3}$).

CAEP: This technique was selected for a number of reasons. First, EP supports a phenotypic description (real encoding) that provides flexibility in data structures and coding (Fogel, 1995). This, in turn, provides a natural representation of the information. Also, a CAEP shell is available for real-valued function optimization.

The CAEP parameters are those of the evolutionary programming, i.e., number of generations and population size. The population and belief space are initialized once only in the whole search process at the first cycle. In the intermediate cycles, the population will contain the sub-populations which interact with their own belief sub-spaces; both subspaces (population and belief) are built using features of each detected cluster.

Fuzzy clustering technique: In a totally unsupervised environment and in the absence of any priory knowledge, it is difficult to admit that objects are distributed according to a special probability density function or that clusters in presence are well separated. Fuzzy clustering models offer a natural way to deal with overlapping clusters and do not require prior information on data distribution. They present a good alternative to probabilistic methods when such information is not available.

Cluster analysis attempts to organize a set of unlabeled input data into a number of natural groups (or clusters) in such a way that elements within a same cluster are as similar as possible while the elements belonging to different clusters are highly dissimilar. Clustering algorithms proposed in literature can be divided into two main categories: crisp (or hard) clustering procedures where each data point belongs to only one cluster and fuzzy clustering techniques where every data point belongs to every cluster with a specific degree of membership (Duda and Hart, 1973). Fuzzy clustering can be considered as a generalization of hard clustering and presents the advantage of dealing efficiently with overlapping clusters.

The proposed algorithm, which uses a measure of similarity between individuals in order to separate them into different clusters, is constituted of two main phases. The first one is an unsupervised learning procedure, which yields an initial fuzzy C-partition of the individuals, by exploring them sequentially. The procedure starts by generating the first class around the first individual encountered. Then a new cluster is automatically created each time the current object presents a small similarity, i.e. less than a specified threshold S_{min}, to the entire already existing cluster centers.

To measure the similarity between two vectors x_i and x_j of R^p, the following expression is proposed

$$S(i,j)=1-\frac{d(x_i,x_j)}{\sqrt{p}} \tag{1}$$

Where $d(x_i, x_j)$ is a distance measure, calculated on the basis of the normalized values of x_i and x_j.

In Eq. 1 when x_j is replaced by a cluster prototype v_j, the relation is also interpreted as the membership degree of x_i to the jth cluster.

$$\mu_{ij} = 1 - \frac{d(x_i, v_j)}{\sqrt{p}} \tag{2}$$

At each iteration of the learning process, the relation (2) is used for calculating the membership degree of the current pattern to all C existing cluster (C is variable). The condition of creating a new cluster is

$$\max_{i \le j \le c}(\mu_{ij}) < S_{min} \tag{3}$$

When Eq. 3 does not occur, the centers of all existing clusters are updated according to:

$$v_j = \frac{\sum_{k=1}^{i} \mu_{kj} x_j}{\sum_{k=1}^{i} \mu_{kj}}; 1 \le j \le C \tag{4}$$

S_{min} represents the similarity threshold. It belongs to an interval whose limits s_1 and s_2 are derived from the data at hand using the relations:

$$S_1 = \min_{i \ne k}\{S(x_i, x_k)\} \text{ and } S_2 = \max_{i \ne k}\{S(x_i, x_k)\} \tag{5}$$

The second phase of this fuzzy clustering approach is an optimization procedure that ameliorates the learned partition, generated during the previous phase. It uses the principle of the well-know fuzzy C-means algorithm (Bezdeck, 1981).

This clustering algorithm is, of course, sensitive to the choice of the similarity threshold value S_{min}. This means that different choices of this parameter may lead to different results. To solve this problem, the partition entropy defined by Bezdeck (1981) is used as validity criterion.

$$h(U) = -\frac{1}{\log(C)} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{C} \mu_{ij} \log(\mu_{ij}) \tag{6}$$

The best solution U^* is the one that minimizes h. Once this solution is achieved, a defuzzification procedure is performed in order to affect definitely each individual to its natural group, i.e., the class for which it presents the maximum membership degree. This results in a final hard C-partition with C cluster centers v_i (1<i<C), which represent the expected optima. For each cycle, the parameters, calculated for each detected cluster C_i (1.i.C), are:

- Cardinal (N), i.e., The number of individuals,
- Center (v_i) : The mean vector of the cluster elements,
- Radius (r_i) of the cluster,

These parameters are of great use to follow the evolution process of niches. For example, v_i and r_i define the geometrical structure of the i^{th} subspace, whereas N represents the number of individuals to generate in each nation.

Spatial separation layer: There are two reasons why spatial should be desirable in evolutionary computation. One reason is that in nature the populations are actually divided in a number of subpopulations that can interact. Another reason is that separating a number of subpopulations allows an effective parallel implementation and is therefore interesting from the point of view of computational efficiency (Pelikon and Goldberg, 2001).

The main goal of this layer is to induce local geography in the population and force local competition within this structure. It involves the formation of nations using results of the clustering procedure. At each cycle, a nation is formed using the center and the radius of each detected cluster.

The cultural exchanges process: This study introduces a new strategy of cultural exchange which is performed between belief spaces. This strategy has an influence on both evolutionary levels. First, it allows injecting the cultural diversity within nations (UNESCO, 2001) and influencing the genetic evolution of individuals, via influence function, on the other hand. The cultural diversity represents one of the roots of the cultural development (UNESCO, 2001).

Cultural exchange occurs with neighbouring nations and the frequencies of exchange decrease during the evolution of the process as well as the nation are well separated in the space. The neighbourhood structure is defined using the Euclidean distance between centers of the clusters. Therefore, the frequency of exchanges is automatically adjusted and depends on the degree of neighbourhood.

The strategy used in this study is defined as follows:

- Define a neighborhood structure using the Euclidean distance between centers of the clusters.
- For each nation exchange the normative knowledge of neighboring nations.

In the context of optimization, the cultural exchanges process allows, in the first hand, individuals of the same nation to exploit knowledge of the neighbouring nation through merging and adding new beliefs to the set of actual belief space. In the other hand, it permits to improve the quality of the solutions found so far. These new beliefs constitute the new constraints that will direct the population through the mutation operator and therefore, speed up the convergence rate of the algorithm.

Time complexity: The time complexity of the evolutionary programming is in order O(n), where n is the population size. The proposed approach presents an additive time complexity induced by the fuzzy clustering process. This complexity corresponds to the computation of the C cluster's centers and the distance between each of the n data vectors and each center. Both steps can be done in O(np) time. Furthermore, the determination of each new partition during the iterating process requires the calculation of the membership degree of each data vector to each class. This step is done in O(npC) time. Since the procedure is iterated until the algorithm has converged, the time complexity of the fuzzy clustering algorithm is in order of O(npCt) where t is the number of iterations.

However, for many real world applications, the dominant cost is that of calculating the fitness value for each population member for each generation (Mahfoud, 1995) and the time complexity O(n) or O(npCt) can be small compared with the cost of evaluation of objective function.

RESULTS AND DISCUSSION

To illustrate the performance of the proposed approach, in this section the results obtained for a set of well known test functions are presented. The experiments were performed using the multimodal functions used in (Beasley *et al*., 1993; Li, 2004). Note that in the clustering process, each object is defined by its coordinate(s) and fitness.

Test functions: In this study, the test functions that have been used to test the performance of the proposed model are presented.

$$F_1(x) = \sin^6(5\pi x); F_2(x) = \exp\left(-2\log(2).\left(\frac{x-0.1}{0.8}\right)^2\right).$$

$$\sin^6(5\pi x); F_3(x) = \sin^6\left(5p\left(x^{3/4} - 0.05\right)\right);$$

$$;F_4(x) = \text{Exp}\left(-2\log(2)\left(\frac{x\text{-}0.08}{0.854}\right)^2\right)$$

$$\text{Sin}^6(5\pi(x^{0.75} - 0.05)); F_5(x,y) =$$

$$\frac{2186 - (x^2 + y - 11)^2 - (x + y^2 - 7)^2}{2186};$$

$$F_6(x,y) = x.\sin(4px) - y.\sin(4py + p)$$

$$+1; F_7(x) = \sum_{i=1}^{n}(x_i^2 - 10\cos(2px_i) + 10)$$

As shown in Fig. 4, F_1 has 5 evenly spaced maxima with a function value of 1.0.F_2 has 5 peaks decreasing exponentially in height, with only one peak as the global maximum. F_3 and F_4 are similar to F_1 and F_2 but the peaks are unevenly spaced. F_5 Himmelblau's function has two variables x and y, where $-6 \le x,y \le 6$. This function has 4 global maxima at approximately (3.58,-1.86), (3.0, 2.0), (-2.815, 3.125) and (-3.78, -3.28). F_6 defined above is the Multi function (Cartro and Timmis, 2002). This function has a several local optima solutions and a single global optimum all distributed non-uniformly. F_7 is the Rastrigin function, where $-5.12 \le x_i \le 5.12$, I = 1;......;30, has one global minimum (which is (0, 0) for dimension = 2) and many local minima (Li, 2004). F_7 with a dimension of 2, 3, 4, 5 and 6 were used to test the ability of the proposed approach in dealing with functions with numerous local minima and of higher dimensions.

Numerical results and discussions: Table 1 shows the evolution of the proposed approach to identify F_1's optima. The proposed model has been applied to a population of 80 individuals. For each cycle, the best partition is obtained for the minimum entropy (h). Also for each detected cluster, the process provides several characteristics, particularly center and radius. In the next cycles, the nations, i.e., subpopulations and their belief space are created using only the coordinates of the cluster's center and the radius. For this test function, each cluster's center is defined by its coordinate (x) and its fitness. Note that for each test function, C^* represent the numbers of the identified clusters.

In the first cycle, the analysis of the clusters centres (v_i) and radius (r_i) shows that the second cluster and the third one contain the same optima of coordinate x=0.3 and thus are overlapping. This is confirmed in the next cycle where the two clusters (0.304, 0.018) and (0.293, 0.174) are merged in a unique cluster (0.302, 0.019), so five clusters are identified and the entropy continues to decrease. In cycle 3, five clusters are detected and the optima are identified. When the algorithm has converged at cycle 4, the entropy is less than 10^{-3} approximately. At this stage, all individuals of the same nation are identical and have a similar culture; this yields a cluster radius close to zero. This is confirmed in Fig. 5 which shows the distribution of individual at each cycle.

Function 2: The proposed model has been applied with different population size and the best one is 80. The process converges at the fourth cycle when the entropy value is 10^{-4}. Results obtained for F_2 are represented in Fig. 6 which presents the distribution of individuals in the search space at different cycle of the process.

Figure 6 shows that on F_2, the proposed model was able to locate all maxima, that is, at the fourth cycle; individuals of the same subpopulation were able to converge to the corresponding maximum.

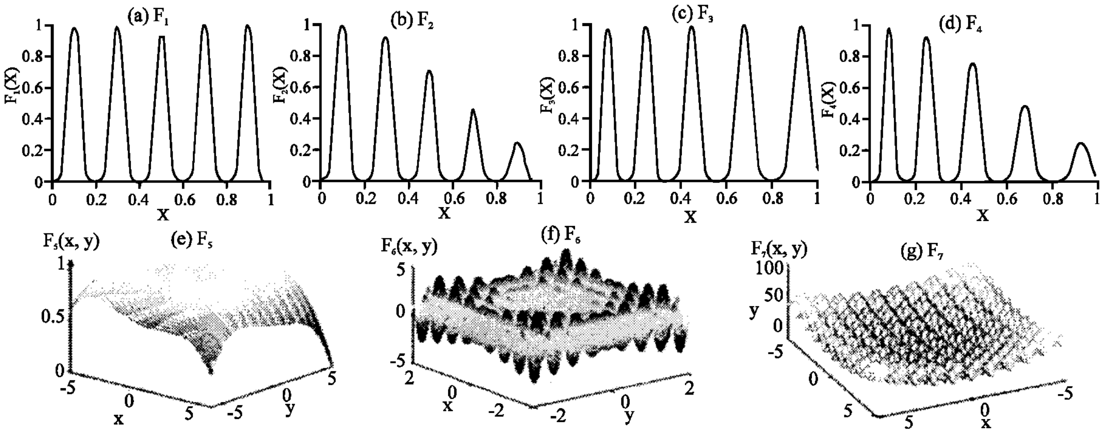

Fig. 4: Test functions

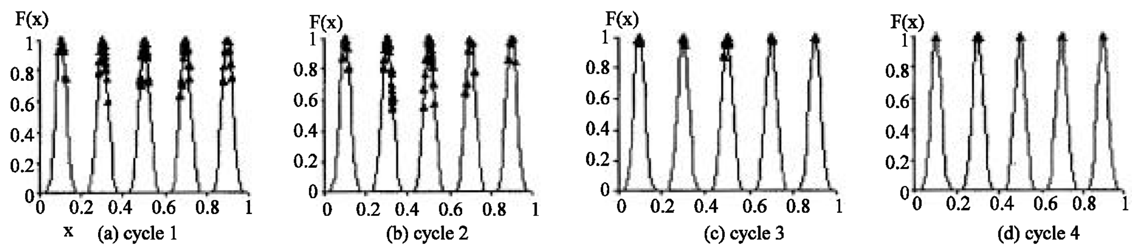

Fig. 5: The distribution of individuals in the search space at each cycle

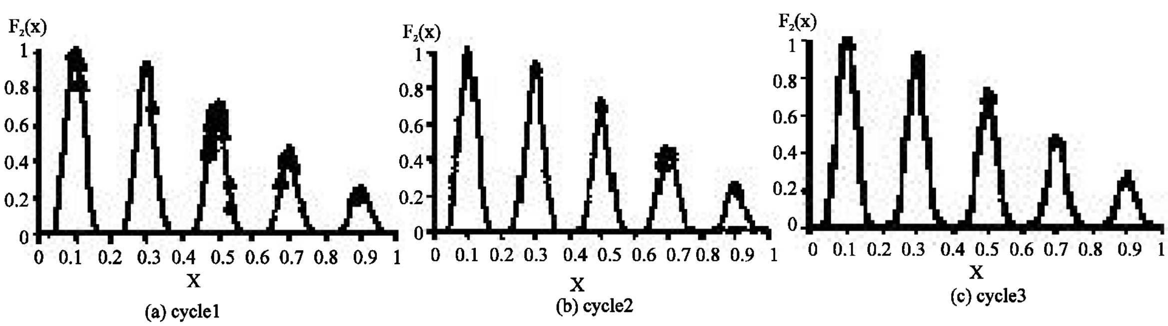

Fig. 6: Placement of individuals in search space

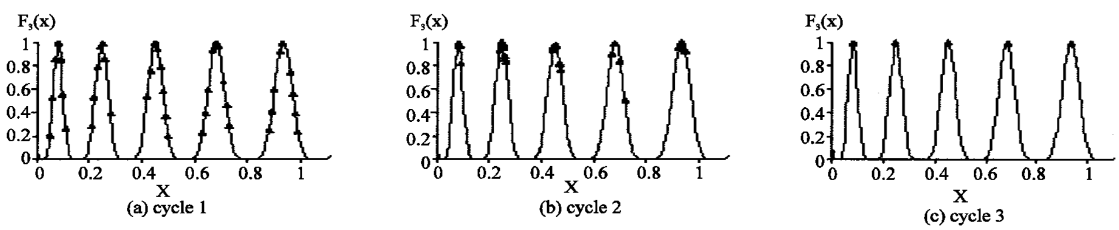

Fig. 7: The distribution of individuals in population at different cycles

Function 3: The method has been performed to the function F_3 using a population of 80 individuals. The entropy criterion value is 10^{-6} when the algorithm converges at fifth cycle. The distribution of individuals in the search space at different iteration is depicted in Fig. 7.

Function F: The proposed model has been applied to a population of 80 individuals. The model converges at the fourth cycle when the value of the entropy is 10^{-16} and locates all optima of the function F_4. Figure 8 shows the location of individuals in population at different cycles.

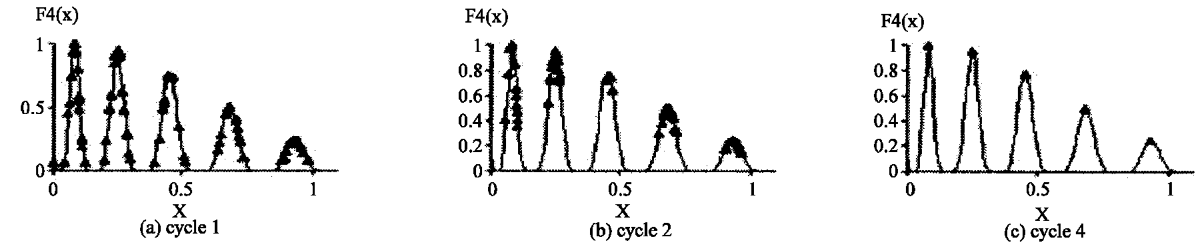

Fig. 8: The distribution of individuals in the search space at each cycle

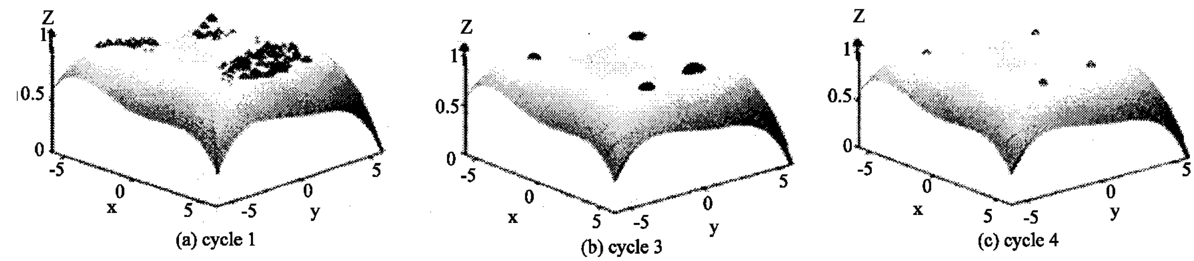

Fig. 9: Placement of individuals at different cycles

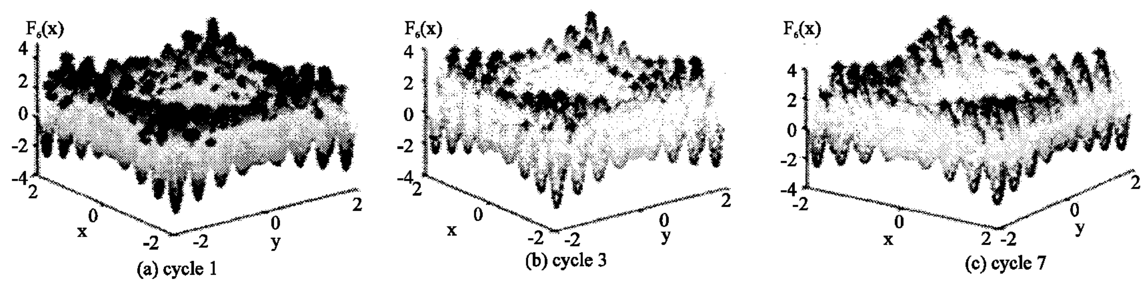

Fig. 10: The distribution of individuals in population at different cycles

Function F: For this test function, the population size used is 150. The optima of F_5 are more difficult to identify and locating them all is not immediate. Results obtained for this example are summarized in Fig. 9. When the algorithm converges at the fourth cycle, the value of the entropy is 10^6. All optima of the Himmelblau's function are identified.

Function F: This function is more difficult than the previous ones. The method has been performed to the function F_9 using a population of 500 individuals. The entropy criterion value is 3.10^{-4} when the algorithm converges at the iteration step 7. The number of detected cluster, i.e., optima is 83. The distribution of individuals in the search space at different iteration is depicted in Fig. 10.

Comparisons of results with other multimodal optimization techniques: In this study, the comparison between the results obtained by the proposed model and other techniques including fitness sharing, Sequential Niched Genetic Algorithm (SNGA) (Beasly *et al*., 1993), Species Conservation Genetic Algorithm (SCGA) (Li *et al*., 2002), Species Particle Swarm Optimization (SPSO) (Li, 2004) and a technique based on artificial immune network, opt-aiNet, (Castro and Thimmis, 2002) for the test functions used in this study, is presented.

Comparisons with sharing fitness: The efficiency of multimodal optimization methods is related to their capacity to maintain the Maximum of function's Peaks Maintained (MPM), to identify all solutions that are most close to the theoretical optima, i.e., Maximum Peaks Ratio (MPR) and to have the lowest Number of Fitness Evaluation (NFE). Table 2 displays the average values of the three performance criterion, obtained for the four functions F_4 and F_5, of the proposed model and sharing scheme during 10 runs.

Table 2 shows that the fitness sharing technique was capable to identify all F_4's and F_5's optima at each run. The implementation of the proposed approach to both test functions F_4 and F_5 displays that it was successfully able to locate all optima of these functions. This is confirmed by the average value of the maximum peaks maintained criterion obtained by the two methods.

In addition, the value of maximum peaks ratio criterion obtained using both techniques indicates that the quality of the optima identified using the proposed approach is better than that obtained by fitness sharing technique. The performance of the proposed approach is confirmed too by the number of function fitness evaluations required to converge.

Comparison with a technique based on artificial immune network (opt-aiNet): Function F_6 was used in order to compare the performance of the proposed method and opt-aiNet technique (Castro and Thimmis, 2002).

As reported in Castro and Thimmis (2002) the performance measurements are the ability of identification of the global optima and the number of detected optima.

Numerical results relative to F_6 show that both methods were able to identify the global optimum. However, the artificial immune network method located only 61 peaks at iteration step 451 (Castro and Thimmis, 2002) while the proposed model locates 83 peaks at the seventh iteration.

Comparison of results with SNGA, SCGA and SPSO: To allow comparisons between the proposed model and SNGA, SCGA and SPSO, this section presents the results obtained for functions F_5 and F_1. Table 3, for example, presents the average values of NFE and the success rate of the four methods over 30 runs.

As shown in Table 3, for F_1 and F_5, the proposed approach and SPSO were able to locate the required solutions with 100% success rate in all runs. In addition, the number of function fitness evaluations required by the proposed model to converge, is less than that required by SPSO, SCGA and SNGA (Beasly *et al*., 1993; Li, 2004; Li *et al*., 2002).

The performance of the proposed model is confirmed too by the results obtained for the function F_7 with dimension varying from 2-6. Table 4, depicts the average performances of the proposed method and SPSO, for the function F_7.

As shown in Table 4, the efficiency of FC-PaCA is confirmed by both the average NFE required to its convergence and the number of optima identified both global and local, as well as the dimension of the function is increased.

Table 1: Evolution of the entropy criterion, centres and radius of clusters throughout the cycling process on function F_1

C*=6		C*=5		C*=5		C*=5	
Centers	Radius	Centers	Radius	Centers	Radius	Centers	Radius
0.105	0.015	0.105	0.018	0.100	0.007	0.100	0.001
0.304	0.018	0.302	0.019	0.302	0.007	0.301	0.002
0.293	0.174	0.452	0.071	0.498	0.011	0.499	0.001
0.501	0.179	0.517	0.187	0.700	0.006	0.700	0.000
0.544	0.166	0.872	0.076	0.900	0.004	0.900	0.000
0.871	0.155						
H	0.448		0.381		0.05		0.0001

Table 2: Performance criteria obtained by sharing fitness and the proposed model for F_1, F_2, F_3 and F_4 functions

	Maximum peaks maintained		Maximum peaks ratio		Number of fitness evaluation	
Techniques	F_4	F_5	F_4	F_5	F_4	F_5
FC-PaCA	5	4	1	1	1400	2500
Fitness sharing	5	4	0.99	0.89	20000	12150

Table 3: Comparison of performance results on F_1 and F_5

		SNGA		SCGA		SPSO		FC-PaCA	
Function	Num. of optima	NFE	Success rate	NFE	Success rate	NFE	Succes rate	NFE	Succes rate
F_1	5	1900	99%	3310	100%	1383.33	100%	1120	100%
F_5	4	5500	76%	-	-	3155	100%	1800	100%

Table 4: Comparison of performance results on F_7

	SPSO		FC-PaCA		
Dimension	Num. of detected optima (Means ± stand dev)	Success rate	Num. of Evals. (Means ± stand dev)	Success rate	Num. of detected optima
2	3711.67±911.87	100%	2396.00±948.00	100%	30.5
3	9766.67 ±4433.86	100%	5275.00±2266.67	100%	28
4	36606.67 ±14662.38	33.3%	25560.00±15258.33	100%	26.6
5	44001.67 ±10859.84	26.7%	40100±1500	100%	31
6	50000.00 ±0.00	0%	40650±5325	94%	23.5

CONCLUSION

This study presents a new approach based on cultural algorithm and fuzzy clustering method, in order to improve the performance of cultural algorithm in the context of multimodal optimization. The model uses a fuzzy clustering technique for two main reasons. At the first hand, fuzzy clustering technique starts with an unsupervised learning procedure, which can be useful to identify the clusters (optima) without any prior information, so the number of the expected optima is automatically computed. At the other hand, it presents the advantage of dealing efficiently with overlapping clusters, which lead to the identification of other unidentified clusters and then other optima.

The proposed model has been reported to display better search performance than other multimodal technique; both in terms of the quality of the solutions found as well as identified new solutions. One reason for the improved search quality is that this approach lies in the partition of the population into several nations, each one of them being processed by a CAEP. Besides, it implements the cultural exchange concept which can be useful to mingling cultures, at the macro evolutionary level. This model allows the exploration of new regions of the search space and permits the location of non detected optima. Since the clusters radius is dynamically adjusted, fine local tuning is improved allowing successive amelioration of the solutions during the cycling process.

The comparison between results obtained by the proposed approach and other techniques confirms the efficiency of the proposed approach in dealing with a hard function with numerous local minima and with high dimension.

In conclusion, generally, for real problems, in which we do not have any information about either search space or the number of optima, the proposed approach can be efficiently applied. In fact, it does not require as much knowledge as other multimodal technique about the problem to solve.

REFERENCES

Angeline, P.A., 1995. Adaptive and Self-Adaptive Evolutionary Computation, in Communication Intelligence, Eds. Marimuthu Palaniswami, *et.al.*, IEEE. Press, New York, pp: 152-163.

Beasley, D., D.R. Bull and R.R. Martin, 1993. An overview of genetic algorithms: Part 2, Research topics. University Computing, 15: 170-181.

Beasley, D., D.R. Bull, R.R. Martin, 1993. A sequential niche technique for multimodal fction optimization. Evolutionary Computation, 1: 101-125.

Bezdeck, J.C., 1981. Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York.

Chung, C.J., 1997. Knowledge Based Approaches to Self Adaptation in Cultural Algorithms. PhD thesis, Wayne State University, Detroit, Michigan.

de Castro L.N. and J. Timmis, 2002. An artificial Immune Network for multimodal function optimization, the IEEE World Congress on Computational Intelligence.

Duda, R.O. and P.E. Hart, 1973. Pattern classification and scene analysis, Wiley, New York.

Fogel, D.B., 1995. A comparison of evolutionary programming and genetic algorithms on selected constrained optimization problems, Simulation, pp: 397-403.

Goldberg, D.E. and J. Richardson, 1987. Genetic algorithms with sharing for multimodal function optimization. In: Proc. 2nd Int. Conf. Genetic Algorithms, pp: 41-49.

Goldberg, D.E. and L. Wang, 1997. Adaptative Niching via Coevolutionary Sharing, In: Quagliarella *et al*. (Eds.). Genetic Algorithms in Engineering and Computer Science (John Wiley and Sons, Ltd.), pp: 21-38.

Holland, J.H., 1992. Adaptation in Natural and ArtificialSystems, (1st Edn.), Ann Arbor: The University of Michigan Press, 1975; First MIT Press Edition.

Li, J.P., M.E. Balazs, G. Parks and P.J. Clarkson, 2002. A Species Conserving Genetic Algorithm for Multimodal Function Optimization. Evolutionary Computation, 10: 207-234.

Li, X., 2004. Adaptively choosing neighbourhood bests using species in swarm optimizer for multimodal function optimization, in GECCO., pp: 105-116.

Mahfoud, S.W., 1995. Niching methods for Genetic Algorithms. PhD Thesis, University of Illinois.

Michalewicz, Z., 1994. Genetic Algorithms + Data Structures = Evolution Programs, Springer Verlag, New York.

Miller, B.L. and M.J. Shaw, 1995. Genetic Algorithms with Dynamic Niche Sharing for Multimodal Function Optimization. IlliGAL Report no. 95010.

Pal, N.R. and J.C. Bezdek, 1995. On Cluster Validity for Fuzzy C-means model, IEEE. Trans. Fuzzy Sys., 3: 370-379.

Pelikan, M. and D.E. Goldberg, 2001. Escaping hierarchical traps with competent genetic algorithms, In: Proc. the Genetic and Evolutionary Computation Conf., pp: 511-518.

Petrowsky, A., 1996. A clearing procedure as a niching method for Genetic Algorithms, In: Proc. IEEE. Int. Conf. Evolutionary Computation, Japan, Nagoya, pp: 798-803.

Petrowsky, A., 1997. A New selection operator dedicated to speciation, In: Proc. 7th Int. Conf. Genetic Algorithms, (Morgan Kaufmann, San Francisco).

Reynolds R.G. and E. Zannoni, 1994. Learning to Understand Software Using Cultural Algorithms, In: Proceeding of the Third Annual Conference on Evolutionary Programming. (Sebald and Fogel, Eds., World Scientific).

Reynolds, R.G., 1979. An Adaptive Computer Model of theEvolution of Agriculture for Hunter-gatherers in the Valley of Oaxaca, Mexico, Doctoral dissertation, University of Michigan, Annual Arbor.

Säreni, B. and L. Krähenbühl, 1998. Fitness sharing and niching methods revisited, IEEE. Trans. Evolutionary Computation, 2: 97-106.

UNESCO, 2001. Universal Declaration on cultural diversity, Adopted by the 31st session of the general conference of UNESCO, Paris article 2: From Cultural diversity to cultural pluralism.

UNESCO, 2001. Universal Declaration on cultural diversity, Adopted by the 31st session of the general conference of UNESCO, Paris article 3: Cultural diversity as a factor in development.

Zannoni E. and R.G. Reynolds, 1997. Extracting Design Knowledge from Genetic Programs Using Cultural Algorithms, Full version in J. Evolutionary Computation (MIT, Press).