



Murdoch
UNIVERSITY

MURDOCH RESEARCH REPOSITORY

<http://researchrepository.murdoch.edu.au/>

This is the author's final version of the work, as accepted for publication following peer review but without the publisher's layout or pagination.

The definitive version is available at

<http://dx.doi.org/10.1109/ICMLC.2006.258347>

Nguyen, D.H.M., Wong, K.P. and Chung, C. (2006) *Multi-modal search with convex bounding neighbourhood*. In: 2006 International Conference on Machine Learning and Cybernetics, 13 - 16 August 2006, Dalian, China.

<http://researchrepository.murdoch.edu.au/12040/>

Copyright © 2006 IEEE

Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

MULTI-MODAL SEARCH WITH CONVEX BOUNDING NEIGHBOURHOOD

D. H. M. NGUYEN¹, K. P. WONG², C.Y. CHUNG²

¹School of Engineering Science, Murdoch University.

²Computational Intelligence Applications Research Laboratory, Department of Electrical Engineering,
The Hong Kong Polytechnic University.

E-MAIL: h.nguyen@murdoch.edu.au, eekpwong@polyu.edu.hk, eecychun@polyu.edu.hk

Abstract:

This paper presents a new dynamic method of subpopulation in solving multi-modal search problems with evolutionary algorithms. The new method identifies the modes found at each generation and equalises the subpopulation sizes assigned to each mode. Modes are identified sequentially starting with the highest fitness mode. Mode membership is determined by successive grouping of fitness dominated convex bounding neighbours, starting from the fittest individual. This new dynamic modal subpopulation approach is able to find a representative sample of optima for multi-modal landscape with infinite number of global and local optima with uneven heights and non-uniform distribution. The algorithm also facilitates parallel implementation.

Keywords:

Multi-modal search; evolutionary computation; parallel algorithm; subpopulation techniques.

1. Introduction

General optimisation concerns the identification of sets of values for some variables that give rise to optima of some objectives defined over variables' domain. If only one objective is defined then it is a single objective problem, otherwise it is specified as a multiple objective problem. Single objective problems are further classified as uni-modal or multi-modal. In uni-modal problems the aim is to find the global optimum if it is unique, otherwise it is sufficient to find any one of the global optima. With multi-modal problems the goal is to search for all of the global and local optima if they are finite in numbers, otherwise a representative sample is sought.

Existing multi-modal search techniques comprise a range of deterministic and stochastic approaches. This paper presents a new method for employing Evolutionary Algorithms (EA) in multi-modal search. EA are a class of biologically inspired guided stochastic search algorithms. They operate by mimicking the process of evolution to

evolve a population of candidate solutions whose fitness is a measure of objective satisfaction. EA have been used in multi-modal search since its inception with the ecological concepts of niching and speciation. Essentially these concepts state that isolated subpopulations or species can evolve concurrently in geographically bounded areas or niches. Hence by associating niches with modes in the objective landscape multi-modal search with EA is possible.

Many techniques exist for applying niching, speciation, or some hybrid combination, in EA multi-modal search. In the niching approaches the idea is to explicitly or implicitly identify niches and limit subpopulation to the niches. Examples of this approach include fitness sharing [1], crowding [2] and clustering [3]. The speciation approaches maintain subpopulations by restricting mating between the subpopulations. Examples of this approach include tagging [4], co-evolution and island migration [5]. The approach of this paper belongs to the class of explicitly niching methods since it involves the complete identification of modes found by the population at the current generation. Since the method evaluates the modes at every generation it is a dynamical niching method.

In essence the new dynamic niching method operates by assigning individuals to modes. The members of the modes are then treated as a subpopulation which can be evolved in parallel by any standard EA. The algorithm for determining individual's membership of modes is described in Section 2. Section 3, describes the procedure for evaluating the performance of the new dynamic niching method which is compared with fitness sharing niching scheme. Experimental results are summarized and appraised together with discussion of the limitations of the new method. Finally, in Section 4 conclusions are drawn and directions for further investigation are given.

2. Modal Subpopulation Algorithm

The overall scheme for the new modal subpopulation consists of three stages. The first is a Crowd Clearing stage.

Next come the Neighbourhood Identification stage before the Modal Partition stage proper. Overall the scheme can be incorporated into an EA generation cycle to partition the population into subpopulations for possible parallel evolution implementation per Figure 1. Note that under this scheme the Evolve Subpopulation tasks operate independently of each other. Consequently, different EA and even non-evolutionary algorithms can be employed with each subpopulation. Furthermore the all parameters of each EA can be set independently of each other, thus it is possible to have different population size and final age.

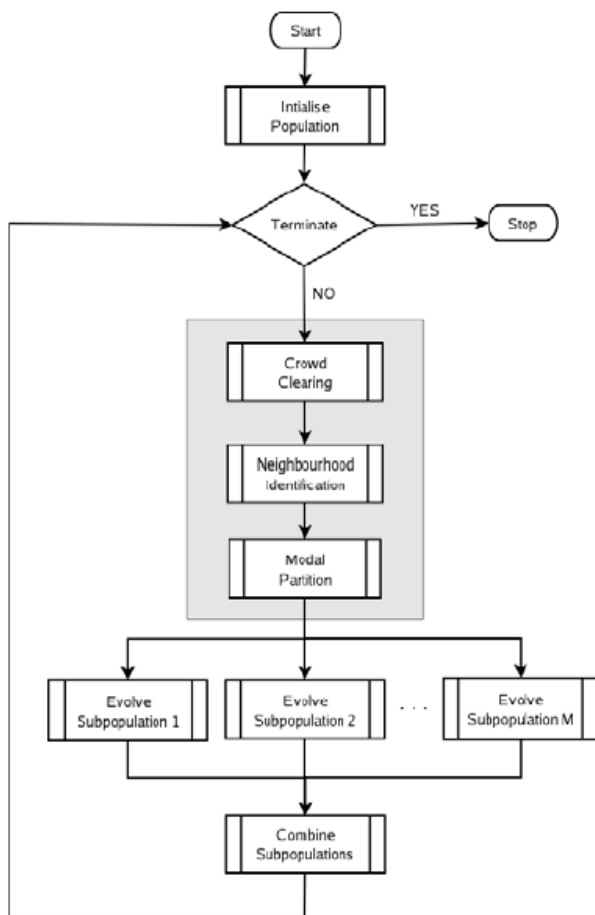


Figure 1. Incorporating new modal subpopulation scheme for EA

2.1. Crowd clearing stage

The idea with the Crowd Clearing stage is to provide a unique set of genome and corresponding fitness $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ from the given set of genomes \mathbf{X} , for the succeeding

Neighbourhood Identification stage to operate on. This is achieved by replacing the i^{th} set of crowded genomes ${}^i\mathbf{X} = \{{}^{i,j}\mathbf{x} : j = 1, 2, \dots, {}^iC\}$, where iC is the number crowded genomes in the set, with a new representative genome ${}^i\tilde{\mathbf{x}}$ and corresponding fitness ${}^i\tilde{y}$. The representative genome and its fitness are defined by (1) with K being the gene space dimension and f is the gene to fitness space mapping. Thus it can be seen that ${}^i\tilde{\mathbf{x}}$ is simply chosen as the centroid of ${}^i\mathbf{X}$ and ${}^i\tilde{y}$ is a new fitness evaluation if ${}^iC > 1$ otherwise the fitness and genome is unchanged.

$${}^i\tilde{\mathbf{x}} = \frac{1}{{}^iC} \begin{cases} {}^i\mathbf{x} & : {}^iC = 1 \\ \sum_{j=1}^{{}^iC} \begin{bmatrix} {}^{i,j}x_1 \\ {}^{i,j}x_2 \\ \vdots \\ {}^{i,j}x_K \end{bmatrix} & : {}^iC > 1 \end{cases} \quad (1)$$

$$\text{and } {}^i\tilde{y} = \begin{cases} {}^iy & : {}^iC = 1 \\ f({}^i\tilde{\mathbf{x}}) & : {}^iC > 1 \end{cases}$$

The crowded sets are selected in sequence as follows, first set ${}^1\mathbf{X}$ is found by randomly selecting a genome ${}^{1,1}\mathbf{x}$ from \mathbf{X} . The members of ${}^1\mathbf{X}$ are taken from \mathbf{X} if they satisfy equation (2) for $i = 1$. In (2) $\delta_x > 0$ is the genome sampling resolution hence the equation simply states that ${}^i\mathbf{X}$ is the set of genomes that are within and on the boundary of the cuboids centered on ${}^{1,1}\mathbf{x}$, with edge length δ_x . Once set ${}^1\mathbf{X}$ is found these genomes are removed from \mathbf{X} and set ${}^2\mathbf{X}$ is found similarly. This process is continued until no genomes are left in \mathbf{X} . Thus the set of unique genome and fitness is $\tilde{\mathbf{X}} = \{{}^i\tilde{\mathbf{x}} : i = 1, 2, \dots, \tilde{I}\}$ and $\tilde{\mathbf{Y}} = \{{}^i\tilde{y} : i = 1, 2, \dots, \tilde{I}\}$ where \tilde{I} is the number of crowded sets found.

$${}^i\mathbf{X} = \left\{ \mathbf{x} \in \mathbf{X} : k \in \mathbf{K}, \left| {}^{i,1}x_k - x_k \right| \leq \frac{\delta}{2} \right\} \quad \text{and } \mathbf{K} = \{1, 2, \dots, K\} \quad (2)$$

2.2. Neighbourhood identification stage

With unique and distinct set of genomes $\tilde{\mathbf{X}}$, a convex

bounding neighbourhood for each genome can be found. First the separation between all genome pairs are calculated and sorted from nearest to farthest. Then for the i^{th} genome ${}^i\tilde{\mathbf{x}}$, its nearest bounding facet is formed from its K nearest neighbours. If the facet vertices degenerate into lower subspace then only the \tilde{K} nearest linearly independent vertices ${}^{i,\tilde{K}}\tilde{\mathbf{X}} = [{}^{i,1}\tilde{\mathbf{x}}, {}^{i,2}\tilde{\mathbf{x}}, \dots, {}^{i,\tilde{K}}\tilde{\mathbf{x}}]$ are kept. The other $K - \tilde{K}$ vertices are randomly generated so that the resulting facet spans the gene space. This can be accomplished by requiring that the randomly generated vertices $\tilde{\mathbf{X}} = [{}^1\tilde{\mathbf{x}}, {}^2\tilde{\mathbf{x}}, \dots, {}^{K-\tilde{K}}\tilde{\mathbf{x}}]$ satisfy (3).

$$K = \text{rank} \left(\left[\begin{matrix} {}^{i,\tilde{K}}\tilde{\mathbf{X}} \\ \tilde{\mathbf{X}} \end{matrix} \right] \right) \quad (3)$$

The vertices in ${}^{i,\tilde{K}}\tilde{\mathbf{X}}$ are then collected as members of ${}^i\tilde{\mathbf{x}}$ neighbourhood, and also assigned as the frontier vertices of the neighbourhood. The facet is then used to eliminate genomes from $\tilde{\mathbf{X}}$ which cannot be members of a bounding convex neighbourhood. The convexity condition is ensured by eliminating genomes that are not on same side as ${}^i\tilde{\mathbf{x}}$, relative to the plane spanned by the facet's vertices. Equation (4) defines the genomes to be eliminated to satisfy the convexity condition. Note $\delta_x > 0$ is used to exclude genomes that, within sampling error, would lie on the plane.

$$\left\{ \mathbf{x} : \mathbf{x} \in \tilde{\mathbf{X}}, \mathbf{I} \cdot {}^i\mathbf{N} \cdot (\mathbf{x} - {}^i\mathbf{x}) \geq 1 - \delta_x \cdot \|\mathbf{I} \cdot {}^i\mathbf{N}\|_1 \right\} \quad (4)$$

where ${}^i\mathbf{N} = \left(\left[\begin{matrix} {}^{i,\tilde{K}}\tilde{\mathbf{X}} \\ \tilde{\mathbf{X}} \end{matrix} \right] - {}^i\mathbf{x} \cdot \mathbf{I} \right)^{-1}$

$$\left\{ \mathbf{x} : \mathbf{x} \in \tilde{\mathbf{X}}, \mathbf{v} = {}^i\mathbf{N} \cdot (\mathbf{x} - {}^i\mathbf{x}), v_j \geq \delta_x \cdot \sum_{k \in \mathbf{K}} |n_{j,k}|, j \in \mathbf{K} \right\} \quad (5)$$

where ${}^i\mathbf{N} = \left(\left[\begin{matrix} {}^{i,\tilde{K}}\tilde{\mathbf{X}} \\ \tilde{\mathbf{X}} \end{matrix} \right] - {}^i\mathbf{x} \cdot \mathbf{I} \right)^{-1}$ and $\mathbf{K} = \{1, 2, \dots, K\}$

Similarly, to meet the bounding condition genomes covered by the solid angle centered at ${}^i\tilde{\mathbf{x}}$ and extended by the facet vertices are eliminated. Equation (5) defines the genome to be eliminated to satisfy the bounding condition, again δ_x is used to account for the effect of discreet sampling. For the special case when the angle form a hyper-plane, that is, when ${}^i\tilde{\mathbf{x}}$ is coplanar with the facet, then the facet's plane and solid angle eliminations are identical. The zones where genomes are to be excluded as

defined by (4) and (5) are illustrated in Figure 2 for a two dimensional gene space.

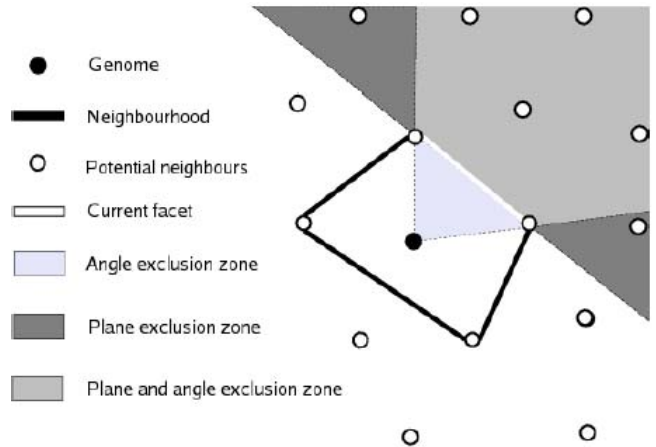


Figure 2. Facet non-neighbour genomes elimination in 2D

2.3. Modal partition stage

With the neighbourhood of all genomes given along with its fitness, it is possible to partition the population genomes into distinct modes. This is accomplished by recognising that a mode is defined by the property where by starting from the fittest genome(s) all members are connected to neighbours with equivalent or lower fitness. Employing this property the modes of the population are search for starting from the mode with the highest fitness. Once mode all members of the mode with highest fitness are found they are removed. The search is repeated for the mode with the second highest fitness. This is done until there are no genomes left. The algorithm for this search is illustrated in Figure 3 below. Note the fitness resolution constant $\delta_y > 0$ is used in the "Assign Peak(s) to Frontier & Members" block to account for the effect finite genome sampling on the fitness landscape. It is a tolerance allowing genomes whose fitness is slightly less than the highest to also be considered as a one of the peaks ${}^m\hat{\mathbf{X}}$ belonging to the m^{th} mode per (6).

$${}^m\hat{\mathbf{X}} = \left\{ \mathbf{x} \in \tilde{\mathbf{X}} : f({}^i\mathbf{x}) - y \geq \max\{\tilde{\mathbf{X}}\} - \delta_y \right\} \quad (6)$$

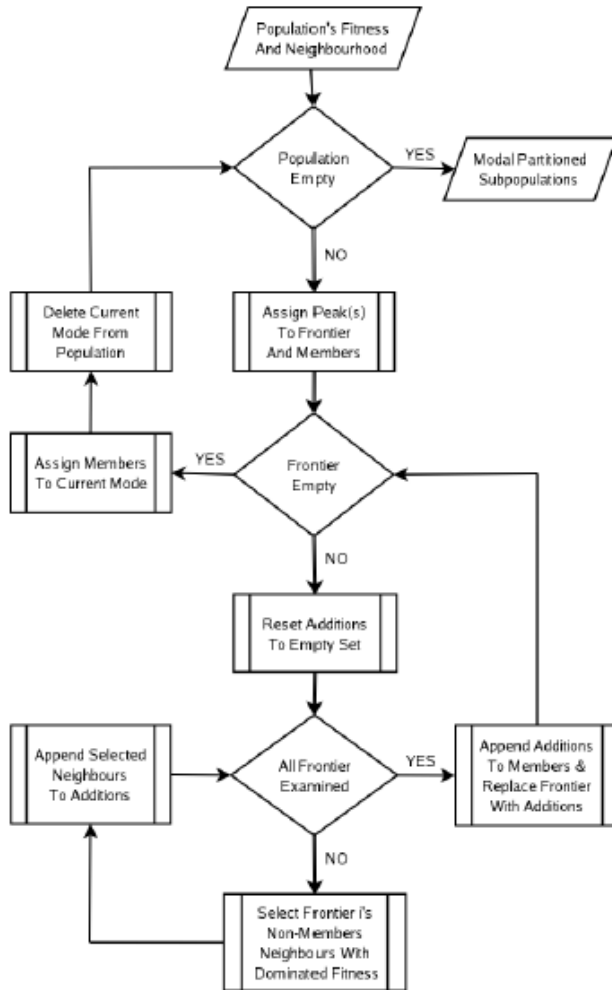


Figure 3. Population modal partitioning algorithm

3. Performance and Discussion

To evaluate the performance of the new algorithm it is compared to the standard fitness sharing method. Both methods employ a canonical evolutionary programming (EP) algorithm as the underlying EA. The EP parameters are set with population size of $N=100$. For one dimensional problems one trial of 100 generations performed. In two dimensions 20 trials of 20 generations are performed. Binary tournament selection is used with Gaussian mutation. The mutation deviation σ_x is set as the geometric mean of δ_x and a maximum value $\hat{\sigma}_x$ per (7). This value is chosen to be $\hat{\sigma}_x = R_x/3$ where R_x is the geometric mean of the gene ranges. The factor of one third ensures that as a maximum mutation can cover the

expected genome range with a probability approaching 50%. An elitist strategy is also implemented in the EP where by the fittest genome of the previous generation is replaces a randomly chosen genome in the new generation if it is not represented in the new generation. For the new approach this is done on a subpopulation level.

$$\sigma_x = \sqrt{\hat{\sigma}_x \cdot \delta_x} \quad (7)$$

The two parameters of the new algorithm, δ_x and δ_y is set per (8) where R_y is the expect range of fitness in the problem domain, and R_x and N is as above. Additionally the new approach also employs the modal partitioning algorithm after the final generation to select only the peaks of the final modes as solutions. From (8) it can be seen that in practical application δ_y is really the only parameter that need to be tune as there is no prior information on R_y . For the fitness sharing algorithm the sharing radius parameter ρ is set at the actual mode radius if the modes are uniform distributed, if not they are set at the arithmetic mean of the distance between modes that are found in bounded gene space. The gene space for the test problems is confined to be in the range from zero to four for all genes.

$$\delta_x = \frac{R_x}{N} \text{ and } \delta_y = \frac{R_y}{N} \quad (8)$$

3.1. Test functions

The new algorithm and the reference fitness sharing approach is tested on a family of functions based on the canonical form of (9). In (9) parameter α controls the distribution of modes with uniform distribution generated only with $\alpha=1$ and non-uniform distribution resulting for all $\alpha>1$. Parameter β controls the heights of the modes with equal height modes generated only with $\beta=1$ and unequal mode heights for all $\beta>1$. Also with (9) it can be observed that for one dimension, $K=1$, and with bounded domain there is generally a finite numbers of peaks or optima. In contrast for higher dimensions $K>1$ there are an infinite number of global and local optima although the numbers of global and local modes are still finite as in the $K=1$ case.

$$y = f(\mathbf{x}) = \frac{\sin(\|\alpha \cdot \mathbf{x}\|^\alpha)}{\beta \|\mathbf{x}\|} \text{ where } \alpha \geq 1 \text{ and } \beta \geq 1 \quad (9)$$

The distinction between peaks and modes is that a mode can be constituted as a single peak or a series of peaks of equal height that are directly connected. Thus a plateau and a ridge as well as an isolated peak would be considered as a mode. For the paper three cases of (9) is considered. In Case I, $K=1$, $\alpha=1$ and $\beta=2$, this case has unequal but uniformly distributed optima with one global optimum and two local optima. In Case II $K=1$, $\alpha=2$ and $\beta=1$, this case has equal height optima, hence there are no local optima, and the global optima are non-uniformly distributed. Finally, in Case III $K=2$, $\alpha=2$ and $\beta=2$, the optima are neither uniformly distributed nor of equal height. Furthermore the gene space is now two-dimensional and there are an infinite number of global and local optima.

3.2. Test results discussion

The solutions of the fitness sharing and new algorithm are illustrated in Figure 4 for Case I which has uniform mode distribution but unequal mode heights. In general it can be seen that the fitness sharing solutions is able to identify all the modes. However the mode's peaks are poorly defined as the EA selection pressure is too low. In contrast, by using modal partitioning the new algorithm is not only able to find the modes it is also able to pick the peaks. The results for Case II which has even mode heights but non-uniform peak distribution is illustrated in Figure 5. It can be seen that the fitness sharing approach, which presupposes a uniform distribution of modes, perform poorly with the narrower mode poorly represented. By comparison the new approach does not have any problems since it does not pre-suppose any mode geometry. In fact the algorithm allows the mode geometry to be discovered dynamically in the population at generation. Consequently, the distribution of modes and their relative heights has no effect as long as they are above the resolution limits of δ_x and δ_y .

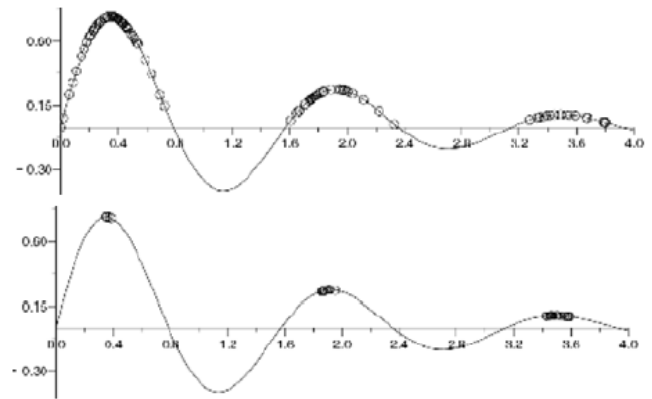


Figure 4. Case I, uniform mode locations and uneven mode heights

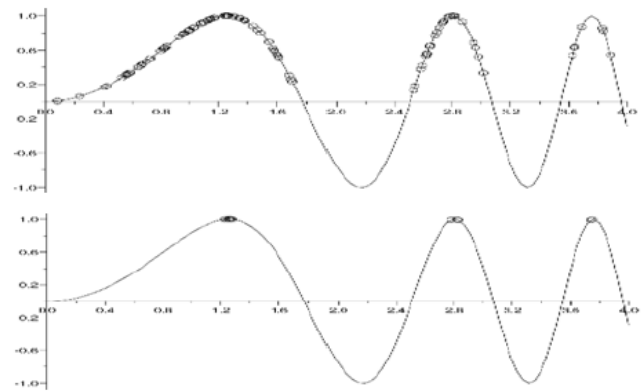


Figure 5. Case II, non-uniform mode locations and even mode heights

The solution for the two dimensional Case III, where the modes are of unequal heights, are non-uniformly spaced and are constituted of infinite numbers of local global optima, is illustrated in Figure 6 for the fitness sharing approach. From Figure 6 it can be seen that only the two highest of the four local modes could be by fitness sharing. Again the poor of the peak resolution is manifested. This is especially evident with the global mode, where the top half of the mode is selected as peaks of the mode. It is possible to increase the resolution by increasing the selection of the underlying EA however this would probably result in lost of the local modes.

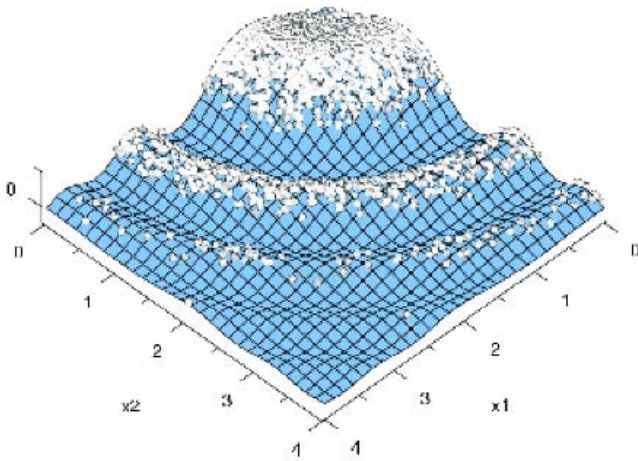


Figure 6. Case III, non-uniform mode locations & uneven mode heights – fitness sharing solutions.

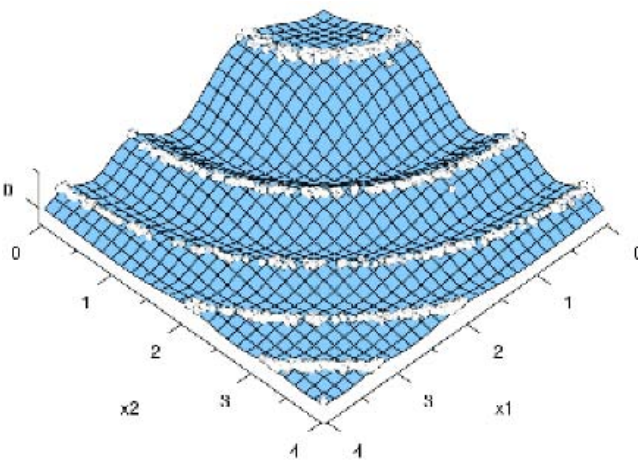


Figure 7. Case III non-uniform mode locations & uneven mode heights – new algorithm's solutions.

Case III solutions by the new approach, given in Figure 7, is significantly better than those of Figure 6. From Figure 7 it can be seen that all the modes, global and local, can be found with the peaks of each mode clearly defined. Furthermore the peaks are found to be evenly distributed there by providing a representative sampling of the infinite no of peaks constituting each mode.

4. Conclusions and Future Work

In this paper a new sub-populating algorithm for multi-modal search is presented. The new method uses convex bounding neighbourhood information to identify modes in the population and assignment of subpopulations to the modes found. The new approach performance is excellent on all problems tested. However, evaluation of convex bounding neighbourhood is very computationally intensive. Thus future research is needed to reduce the computation necessary to determine the neighbourhood. In conclusion the significance of this paper is that it introduces the concept of modal subpopulating and has demonstrated the effectiveness convex bounding neighbourhood information in multi-modal search results.

Acknowledgements

The work was supported by Research Grants Council, Hong Kong (PolyU 5214/03E) and The Hong Kong Polytechnic University.

References

- [1] Sareni B. and Krahenbuhl L., "Fitness sharing and niching methods revisited", IEEE Transactions on Evolutionary Computation, vol 2, no. 3, September 1998, pp. 97-106.
- [2] DeJong K. A., "An analysis of the behaviour of a class of genetic adaptive systems", Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, 1975.
- [3] Habagandi V. and Nicholaou M., "A hybrid approach to global optimisation using a clustering algorithm in a genetic search framework", Computers and Chemical Engineering, 1995.
- [4] Spears W. M., "Simple subpopulation schemes", Evolutionary Programming Conference, 1994, World Scientific
- [5] Martin W. N. Lienig J. and Cohom J. P. "Island (migration) models: Evolutionary algorithms based on punctuated equilibria", Handbook of Evolutionary Computation, Sect. C, Chapter 6, pp. C6:3:1-C6:3:16, 1997.