文章编号：1000-8152(2004)02-0302-03

多峰搜索的自适应遗传算法

刘洪杰，王秀峰
（南开大学 信息技术科学学院 自动化系，天津 300071）

摘要：对多峰函数问题提出了基于峰值转换和优育子群相结合的遗传搜索策略.主要是：通过变换函数将多峰问题中的所有峰变成"等高"峰，从而保证每个峰都有同等机会被找到；在种群中实施各种遗传操作及近亲排斥策略，以保证种群的多样性；将种群中适应值超过阈值的个体迁徙形成一个子群，在子群中实施"梯度操作"，对个体进行精细进化.该方法不仅可保证较快地找到所有峰，而且无需对多峰函数做峰的个数已知、峰均匀健分布等任何先验假设.最后与Spears的简单子群法进行了对比实验.

关键词：遗传算法；多峰搜索；梯度算子；聚类算子
中图分类号：TP13 **文献标识码**：A

Adaptive genetic algorithm for multi-peak searching

LIU Hong-jie, WANG Xiu-feng
(College of Information Technology and Science, Nankai University, Tianjin 300071, China)

Abstract: An adaptive multi-peak genetic searching strategy based on optimal subgroup migrating and the functional transformation is proposed. The main idea is, all peaks of multi-peak problems whose peaks are not equally high are transformed into those whose peaks are equally high by functional transformation so as to find all peaks in the same probability; Some genetic operators and near relative excluding strategy are executed in order to maintain the population diversity; The excellent individuals whose fitness are bigger than a threshold are migrated into a subpopulation; Gradient operator in subpopulation is applied to make the individuals evolved subtly. This searching strategy not only can ensure to find all peaks, but it doesnot need any pre-knowledge hypothesis, such as the number and distribution of the peaks. Finally, a comparison test with Spears's Simple Sub Population strategy is performed.

Key words: genetic algorithm; multi-peak searching; gradient operator; clustering operator

1 引言(Introduction)

遗传算法(genetic algorithm，简称GA)已经广泛应用于函数优化、任务调度、神经网络、系统辨识、机器学习等各个领域[1,2].在实际应用的过程中，人们开始注意到：通常的遗传算法虽然能够进行全局最优搜索，但是只能收敛到一个极值点.而在实际问题中，经常需要求出多个极值点，在这方面，传统的遗传算法则显出其不足.

Holland[3]，Goldberg和Richardson[4]通过引入小生境和物种理论，用分享和限制交配机制相结合的方法成功地实现了多峰搜索.但他们的方法是基于两个假设：一个是假设搜索空间中峰的个数已知，另一个假设是在搜索空间内所有峰均匀分布.这在实际问题中是很难满足的.另外，该算法在对每个体进行相似性衡量时，需花费相当大的计算量.

M. William[5]针对Goldberg的算法中的问题，提出了一种较为简单的基于分享机制的算法——简单子群法.该方法避免了个体之间的距离运算，对峰间距也没有作等距的约定，但需假设子群的个数多于峰的个数（这实际上也暗含了对峰数的假设）.这样一来，种群规模会随着子群数的增长呈指数级的增长.这种方法的成功很大程度上依赖于分组和子群的数目.

由于现行方法有诸多先验苛刻的假设，因此很难实际应用.另外，即使满足假设，由于它们都是采用分享机制，当峰的高低相差较大时，则常常会漏掉低峰.鉴于上述缺点，本文另辟新径，提出了一种转变策略，突出"峰"的特征，自动地识别"峰"与"非峰".只要是峰，不论高低，一律"平等"，使每个峰都具有相同的被选力度.进化过程采取了"优育子群迁

收稿日期：2002-10-28；收修改稿日期：2003-06-09.

徙”和近亲排斥策略，既保证了种群的多样性又确保了整体收敛.另外，本文还引入了梯度算子和聚类算子，提高了算法的搜索速度和精度.实验表明，本文提出的方法不需对多峰函数做任何的先验假设，且收敛速度快、精确度高，能够找到所有的峰值.

2　梯度算子(Gradient operator)

设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_m)$($m$ 是染色体的维数)为一个染色体，设 f 为本文研究的对象函数.任取染色体中的一个元素，例如第 j 个元素，在其上产生一个小随机扰动 $\Delta \boldsymbol{x}_j$，记 $\Delta \boldsymbol{x}_j = (0, 0, \cdots, \Delta x_j, \cdots, 0)$，则第 j 个分量上的近似梯度为

$$\Delta_j(\boldsymbol{x}) = \frac{f(\boldsymbol{x} + \Delta \boldsymbol{x}_j) - f(\boldsymbol{x})}{\Delta \boldsymbol{x}_j}.$$

记 $\Delta(\boldsymbol{x}) = (\Delta_1(x), \Delta_2(x), \cdots, \Delta_m(x))$ 为梯度向量，$|\Delta(\boldsymbol{x})| = \sqrt{\sum_{j=1}^{m} (\Delta_j(x))^2}$ 称为梯度向量的矩.经过梯度算子操作后产生的新个体为 $\boldsymbol{x}' = \boldsymbol{x} + \eta \frac{\Delta(\boldsymbol{x})}{|\Delta(\boldsymbol{x})|}$，其中 η 称为步长，是一个小的正数.

3　聚类算子(Clustering operator)

为了使峰值点附近的个体快速的向其最近的峰值点聚集，本文建立了聚类算子，使得每个峰值点附近的个体自动地向峰值点聚集，从而加快了进化的速度和精度.其基本思想如下：

设任意两个染色体 $\boldsymbol{x}^i = (x_{i1}, x_{i2}, \cdots, x_{im})$ 和 $\boldsymbol{x}^j = (x_{j1}, x_{j2}, \cdots, x_{jm})$，定义其距离 $d_{ij} = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2}$.在种群中选取适应值高于平均适应值一定比例的个体 $\boldsymbol{x}^i = (x_{i1}, x_{i2}, \cdots, x_{im})$，以 $\boldsymbol{x}^i$ 为中心产生两个子体：$\boldsymbol{x}'_i = \boldsymbol{x}^i + \boldsymbol{\varphi}_i$ 和 $\boldsymbol{x}''_i = \boldsymbol{x}^i - \boldsymbol{\varphi}_i$，然后淘汰其中适应值较小的一个.这里

$$\boldsymbol{\varphi}_i = \sum_{j=1}^{\mathrm{popnum}} w_{ij} \boldsymbol{x}^j \Big/ \sum_{j=1}^{\mathrm{popnum}} w_{ij},$$

$$w_{ij} = \frac{\mathrm{fitness}\ (j)}{e^{\beta d_{ij}}}.$$

fitness (j) 是染色体 $\boldsymbol{x}^j$ 的适应值；β 称为聚合系数，用来调整聚类强度，一般选取 $\beta = \min\left(\frac{1}{H_k - L_k}\right)$，$k = (1, \cdots, m)$(其中 H_k, L_k 分别是染色体第 k 个分量的最大值和最小值)；popnum 是种群的规模.这样新生成的子体就向离它最近的并且适应值大的个体靠拢，从而达到聚类的目的.

4　适应值函数(Fitness function)

现行的多模态遗传算法中，适应度函数强烈地依赖于函数值，试图采用峰附近的个体“均分”适应值的方法达到使不等高的峰具有相等的选择力度的目的.为此需假定“峰间距相等”等一些苛刻的条件.而这里构造的适应度函数只突出“峰”的特点，只区别“峰”与“非峰”，无论峰的高低，所有峰都具有相同的适应值，从而使所有峰都有同等的机会被找到.要构造这样的适应度函数，就要找出峰的共同点.峰具有以下特点：

1) 对于峰附近光滑的函数，峰值点的函数值比其附近点的函数值都高；

2) 越是接近峰，梯度就越接近于零.

针对以上特点，引入示性函数

$$h(\boldsymbol{x}) = e^{-|\Delta(x)|}. \tag{1}$$

由式(1)可见，越接近峰值，梯度越小，当梯度向量的矩 $|\Delta(\boldsymbol{x})| \to 0$ 时，$h(\boldsymbol{x}) \to 1$；梯度越大，$h(\boldsymbol{x})$ 的值越小，当 $|\Delta(\boldsymbol{x})| \to \infty$ 时，$h(\boldsymbol{x}) \to 0$.不难看出，$h(\boldsymbol{x})$ 可以很好地刻画了峰的上述特点且不区分峰的高低.但进一步分析不难发现，作为描述峰，还存在如下问题：

1) 对于峰附近不光滑的函数，$h(\boldsymbol{x})$ 不能正确反应峰的特征.

2) $h(\boldsymbol{x})$ 在谷底和平原处，与在峰处具有相同的函数值.

为了克服上述的缺点，引入惩罚项如下：

$$\theta(\boldsymbol{x}) = \sum_{j=1}^{m} \left[\begin{array}{l} \mathrm{sgn}\ (f(\boldsymbol{x}) - f(\boldsymbol{x} + \Delta \boldsymbol{x}_j)) + \\ \mathrm{sgn}\ (f(\boldsymbol{x}) - f(\boldsymbol{x} - \Delta \boldsymbol{x}_j)) \end{array} \right]. \tag{2}$$

其中：$\Delta \boldsymbol{x}_j$ 由第 2 小节所定义，sgn $(\cdot)$ 是符号函数.

当 $\boldsymbol{x}$ 处于峰值点时，$\theta(\boldsymbol{x}) = 2m$；

当 $\boldsymbol{x}$ 处于谷底点时，$\theta(\boldsymbol{x}) = -2m$；

当 $\boldsymbol{x}$ 处于平原点时，$\theta(\boldsymbol{x}) = 0$；

其他情况，$-2m < \theta(\boldsymbol{x}) < 2m$.

通过惩罚函数，可以明显地区分峰、谷和平原.综合式(1)、(2)，取评价函数为

$$g(\boldsymbol{x}) = h(\boldsymbol{x}) + \theta(\boldsymbol{x}) + 2m. \tag{3}$$

则 $g(\boldsymbol{x})$ 可以很明显地区分原函数 $f(\boldsymbol{x})$ 的峰值点与非峰值点，且在 $f(\boldsymbol{x})$ 的所有峰值点处，$g(\boldsymbol{x})$ 都取最大值 $4m + 1$.这是一个重要的先验知识，据此很容易确定个体的优劣程度.为了对变换函数有个直观的印象，图 1 和图 2 分别给出了仿真实验中第 2 个函数 $f_2(\boldsymbol{x})$ 和相应的 $g(\boldsymbol{x})$ 的函数图形.

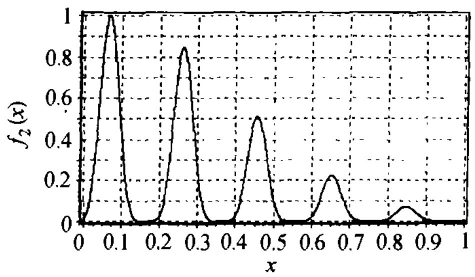

图 1 $f_2(x)$ 函数曲线

Fig. 1 Curve of $f_2(x)$

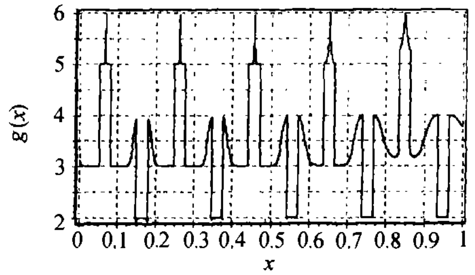

图 2 相应的 $g(x)$ 函数曲线

Fig. 2 Curve of $g(x)$ corresponding to $f_2(x)$

5 优育子群迁徙的遗传策略(Excellent individual sub-population migirating stratergy)

通过评价函数,将峰不等高的函数转换成对应的等高函数,这为找到所有峰创造了良好条件,但还不能保证收敛到所有的峰.必须有好的遗传策略.为此,本文提出了基于优育子群迁徙的遗传策略.其中心思想是:设定一个适应值的阈值,称为迁徙阈值.大于迁徙阈值的个体迁徙到一个特殊的子群——优育子群中.通过这种迁徙操作将种群分成母群和优育子群.初始时,优育子群中没有个体.母群采用通常的遗传操作进化,但是没有选择操作,所有的个体直接进入下一代,这时种群的规模不断增大.随着遗传进化尤其在聚类算子的作用下,个体自动向某个或某几个峰值点聚集,当某个个体适应值超过迁徙阈值时,此个体被迁徙到优育子群中,母群中剔除该个体及其"近亲",即与该个体距离非常接近(距离小于某个阈值)的个体.这时种群的规模会减小,当母群的规模小于初始种群规模时,补充新个体,但补充的新个体不能在优育子群个体的附近,直到母群的规模达到初始种群的规模.在母群进化的同时,优育子群也在进化:一是每当有个体进入优育子群时,首先检查优育子群中是否有其近亲,如果有,则淘汰其中适应值较小的个体;二是对优育子群中的个体用梯度算子进行精细地进化,并随时淘汰其适应值低的近亲个体.

从以上进化过程可以看出,一旦某个峰值点的邻域有个体进入,在聚类算子的作用下,该个体会很快地向峰值点靠拢,且当适应值超过了迁徙阈值时,则被迁徙到优育子群中,且进行进一步的梯度进化,进而该个体所处的峰值点将会被精确地找到并保存在优育子群中.由于采用了淘汰近亲的策略,极大地增加了种群的多样性,为找到所有峰值点创造了必要条件.因此当运行代数足够长时,每个峰值点处都会有响应个体进入优育子群,从而使该算法可以稳定而较精确地找到定义域内的所有峰值点.进化过程结束后,优育子群中的个体即是所有的峰值点,不需要人工的筛选和分析.详细的实现步骤参见文献[6].

6 仿真实验(Simulation test)

本文用文献[5]给出的3个函数,并用其给出的简单子群法与本文的方法进行了对比测试.3个函数分别为:

1) $f_1(x) = \sin^6(5.1\pi x + 0.5)$,极值点等距且峰等高;

2) $f_2(x) = f_1(x) \times e^{-4\ln 2\times(x-0.0667)^2/0.64}$,极值点等距但峰不等高;

3) $f_3(x) = x\sin^6(5.1\pi(x^2 - 0.57))$,极值点不等距且峰也不等高.

其中 $x \in [0\ 1]$.

遗传算法采用实数编码,并用比例交叉和随机变异操作[2],交叉和变异的概率分别为0.1和0.01.对上述3个函数,两种方法都运行了5次,得到的最好结果如表1所列.表中精确度定义为:实际峰值点对应的个体与计算得到的最优个体之差的平方和.

从测试结果可以看出,本文的方法具有收敛速度快、精度高、需要的种群规模小、且找到了所有的峰等优点.

表 1 测试的结果对比

Table 1 Comparision of test results

函数	指 标	优育子群法	简单子群法
$f_1(x)$	种群数	10	至少为40
	进化代数	50	200
	峰的个数(5)	5	5
	精确度	0.000012	0.004
$f_2(x)$	种群数	10	至少为100
	进化代数	150	400
	峰的个数(5)	5	4
	精确度	0.00004	0.0024
$f_3(x)$	种群数	10	至少为100
	进化代数	180	500
	峰的个数(5)	5	4
	精确度	0.000051	0.0046

(下转第310页)

cinematically redundant Manipulators [J]. *IEEE Trans on Robotics and Automation*, 1996, 12(4): 543 – 552.

[4] LEE K, CARROLL V C. Control algorithms for stabilizing under-actuated robots [J]. *J of Robotic Systems*, 1998, 15(12): 681 – 697.

[5] MUKHERJEE R, CHEN D. Control of free-flying under-actuated space manipulators to equilibrium manifolds [J]. *IEEE Trans on Robotics and Automation*, 1993, 9(5): 561 – 570.

[6] BERGERMAN M, XU Yangsheng. Optimal control of manipulators with any number of passive joints [J]. *J of Robotic Systems*, 1998, 15(13): 115 – 129.

[7] 何广平,陆震,王凤翔.被动冗余度空间机器人动力学控制[J].空间科学学报, 2001, 21(1): 73 – 80.
(HE Guangping, LU Zhen, WANG Fengxiang. Dynamical control of under actuated redundant space robots [J]. *Chinese J of Space Science*, 2001, 21(1): 73 – 80.)

[8] XU Yangsheng. The measure of dynamic coupling of space robot systems [C] // *IEEE Int Conf on Robotics and Automation*. California: IEEE Computer Society Press, 1993: 615 – 620.

[9] 高为炳.变结构控制的理论及设计方法[M].北京:科学出版社, 1996.
(GAO Weibing. *The Theory and Designing Methods of Variable Structure Control* [M]. Beijing: Science Press, 1996.)

作者简介:

何广平 (1972 —),男.副教授,博士.主要研究领域:机器人技术,仿生机械. E-mail: guangpinghe@sohu.com;

陆 震 (1942 —),男.教授,博士生导师.主要研究领域:机器人学,智能机械. E-mail: zhenluh@sohu.com;

王凤翔 (1943 —),男.研究员.主要研究领域:机器人技术,火箭发动机. E-mail: wangfengxiang@sohu.com.

(上接第304页)

7 结论(Conclution)

本文的方法不仅不需要对多峰函数做任何先验假设,而且在性能上比现有的方法都有很大的提高.

参考文献(References):

[1] MICHALEWICZ Z. *Genetic Algorithms + Data Structures = Evolution Programs* [M]. Berlin, Heidelberg, New York: Springer-Verlag, 1994.

[2] WANG Xiufeng, ELBULUK M E. The application of genetic algorithm with neural networks to the induction machines modeling [J]. *System Analysis Modeling Simulation*, 1998, 31: 93 – 105.

[3] HOLLAND J H. *Adaptation in Natural and Artificial System: An Introduction Analysis with Applications to Biology, Control and Artificial Intelligence* [M]. Michigan, USA: The University of Michigan Press, 1975.

[4] GOLDBERG D E, RICHARDSON J. Genetic algorithms with sharing for multimodel function optimization [C] // *Proc of the Second Int Conf on Genetic Algorithms*: July 28 – 31, 1987 at the Massachusetts Institute of Technology. Massachusetts, USA: The Massachusetts Institute of Technology Press, 1987: 41 – 49.

[5] WILLIAM M. Spears, simple subpopulation schemes [C] // *Proc of the Third Annual Conference on Evolutionary Programming*, Feb. 24 – 26, 1994 at San Diego, California, USA. Singapore: World Scientific, 1994: 296 – 307.

[6] 刘洪杰.遗传算法及其在金融预测与金融决策中的应用研究[D].天津:南开大学, 2002.
(LIU Hongjie. *Research of genetic algorithms and its application in financial forecast and financial decision* [D]. Tianjin: Nankai University, 2002.)

作者简介:

刘洪杰 (1975 —),男.博士.主要研究方向:遗传算法,复杂系统建模,金融预测与决策. E-mail: liuhongjie@x263.net;

王秀峰 (1942 —),男.南开大学控制理论与控制工程教授,博士生导师.1965年毕业于南开大学数学系.出版了8本著作,发表了80余篇论文.主要研究方向为:复杂系统建模,遗传算法,数据挖掘,智能信息处理等. E-mail: wangxf@nankai.edu.cn.