

基于多种编码的多群体遗传算法

张莉芬 黎明 周琳霞

(南昌航空工业学院测试与控制系, 南昌 330034)

摘要 为了有效地克服标准遗传算法(SGA)中的早熟收敛现象,提出了一种基于多种编码的多群体遗传算法。该方法是采用3个群体同时进行进化的策略。其中,第1个群体是采用浮点数编码方法,以使该群体具有较强的局部搜索能力;第2个群体是采用二进制编码方法,以使该群体具有较强的全局搜索能力;第3个群体为“精华种群”,用于保存算法在进化过程中产生的优秀个体。在进化过程中,还通过引入“移民”策略来交换3个群体中的优秀个体,以有效地增加群体的多样性。该算法不仅不易陷入局部收敛,还具有较强的跳出局部收敛的能力,且收敛速度较快。通过对一系列典型复杂多模函数进行的优化计算试验,结果证实了该方法的有效性和优越性。

关键词 遗传算法 早熟收敛 多种群进化 多种编码

中图分类号: TN919.81 TP301.6 **文献标识码**: A **文章编号**: 1006-8961(2002)09-0980-05

Multi-species Genetic Algorithms Based on Multi-encoding

ZHANG Li fen, LI Ming, ZHOU Lin-xia

(Department of Test & Control, Nanchang Institute of Aeronautical Technology, Nanchang 330034)

Abstract In this paper, a genetic algorithm using multi-species and multi-encoding method is proposed to combat premature convergence inherent in Standard Genetic Algorithms (SGA). It involves with three species evolved simultaneously. By using float encoding method and binary encoding method respectively, the first species has stronger local search ability and the second has stronger global search ability. The third species, which called "elitist species", aims to keep the elitist individuals in the evolution process. And at the same time, it evolves too, which will enhance the convergence speed and improves the performance of GA. And the migration strategy adopted in the proposed method which immigrates elitist individuals among the three species can keep the population diversity efficiently. This multi-species method can help genetic algorithms to escape from possible local entrapment and obtain good tradeoff between exploration ability and exploitation ability. The experimental results of this method on a series of classical complex multimodal functions have shown its efficiency and superiority.

Keywords Genetic algorithms, Premature convergence, Multi-species evolution, Multi-encoding

0 引言

遗传算法(Genetic Algorithms, GA)是由美国Michigan大学的John Holland教授提出的一种借鉴生物体自然选择和自然遗传机制的随机搜索算法,其主要特点不仅是采用群体搜索策略和可实现群体中个体之间的信息交换,且搜索过程不依赖于

梯度信息。它尤其适用于处理传统搜索方法难以解决的复杂和非线性问题,可广泛应用于组合优化、机器学习、自适应控制、规划设计、人工生命和图象处理等领域,是21世纪有关智能计算的关键技术之一。这里应该指出的是,遗传算法虽然可以实现兼顾全局搜索和局部搜索的所谓均衡搜索,并且在许多复杂问题的求解中也往往能得到满意的结果,但是该算法的全局优化收敛性的理论分析尚待解决。目前普遍认为,标准遗传算法并不保证全局最优收

基金项目:江西省跨世纪学科带头人培养计划项目(第三批);江西省测控研究中心开放基金项目(200104003)

收稿日期:2001-04-18;改回日期:2001-12-10

敛^[1],而且用遗传算法求解大规模复杂的实际问题时,常常发生早熟收敛(Premature Convergence)现象,即种群演化到一种非全局最优状态,也就是说,算法的进一步迭代已不能产生更优可行解^[2]。

为克服早熟收敛问题,许多研究者从编码、遗传算子、并行进化等各个方面对遗传算法进行了改进^[3~7]。如文献[3]中提出了一种浮点数编码遗传算法,该方法虽收敛速度快,且能得到较高的优化精度,但由于浮点数编码方法本身的特点,该方法的全局搜索能力不高;文献[4]提出了一种新型结构的两代竞争遗传算法,即其选择过程是一个由本代种群和经过交叉、变异算子作用过的染色体种群的共同竞争,再经过优胜劣汰的选择来生成新一代染色体种群的过程。大量实例表明,该算法具有搜索效率高、鲁棒性强的特点。文献[5]提出了一种遗传-灾变变异的方法,这种方法以一个远大于通常概率的突变概率来进行一次突变操作,并将突变后产生的新个体加入到中间群体,以使其能始终保持群体的多样性,并使群体脱离“早熟”。但这种为了提高遗传算法的性能,对变异率进行改变的方法,其效果并不明显^[2]。文献[6]中提出了一种自适应的并行遗传算法,该算法是通过将多群体并行进化和自适应参数调整相结合来对每个群体的变异率、交叉率进行自适应调整,以使得各子群体分别具有较强的全局搜索能力和较强的局部搜索能力。由于该方法克服了以往定常参数单群体进化的不足,并综合了不同特性种群进化的长处,因而使得早熟收敛问题得以缓解,同时又提高了搜索的范围和效率。但是该方法中,需要调整自适应控制参数,以使每个群体具有不同的全局搜索能力和局部搜索能力,而控制参数的调整,则需根据实际问题试探性地给出,其不恰当的设定将会在很大程度上影响算法的性能,且较多的控制参数的调整,还将使得问题变得复杂。文献[7]中提出的多群体遗传算法主要是通过对每个群体设置不同的变异率和交叉率,以使每个群体分别具有较强的全局搜索能力或较强的局部搜索能力,进而使得算法不易陷入早熟收敛。另外,为避免早熟收敛,改变变异率的作用并不大,而增大群体规模比改变变异率则更为有效,但群体规模的增大,又使得计算量增加,进而影响算法性能。另外,交叉率的变化,对避免早熟收敛的作用也不大^[2],所以文献[7]中的方法对增加种群多样性的效果并不很明显。

为了有效地克服遗传算法中的早熟收敛现象,

本文提出了一种基于多种编码的多群体遗传算法,该算法中的两个子群体分别采用浮点数编码方法和二进制编码方法进行局部搜索和全局搜索,试验证明,该方法进一步提高了遗传算法克服早熟收敛的能力。

1 基于多种编码的多群体遗传算法

1.1 多种编码

编码方法是设计遗传算法时的一个关键步骤,它在很大程度上决定了如何进行群体的遗传进化运算以及遗传进化运算的效率,而目前主要有二进制编码、格雷码编码、浮点数编码及符号编码等编码方法^[8]。其中,二进制编码方法是遗传算法中最常用的一种编码方法,它使用的编码符号集是由二进制符号0和1所组成的二值符号集{0,1}。例如,对于 $x \in [0, 1023]$,若用10bit长的二进制编码来表示该参数的话,则字符串0010101111就可表示一个个体,它所对应的参数值是 $x=175$ 。对于用二进制编码方法表示的个体,其变异操作有时虽然只是一个基因座的差异,而对应的参数值却相差较大,所以,该编码方法具有较强的全局搜索能力,但对于一些连续函数的优化问题等,由于遗传算法的随机特性,而使得其局部搜索能力较差,而且,对于一些多维、高精度优化问题,又由于个体串长过大,而使得问题无法计算,还易产生早熟收敛。通常浮点数编码方法使用的是决策变量的真实值,其个体的每个基因值用某一范围内的一个浮点数来表示,且其个体的编码长度等于其决策变量的个数。例如,若某一个优化问题含有5个变量,且每个变量都有其对应的上下限 $[U'_{\min}, U'_{\max}]$,则

5.80	6.90	3.50	3.80	5.00
------	------	------	------	------

就表示一个个体的基因型,其对应的表现型是: $[5.80, 6.90, 3.50, 3.80, 5.00]^T$,实验证明,该编码便于较大空间的遗传搜索。对于用浮点数编码方法表示的个体,因为其变异操作通常是采用高斯变异方法来对原个体附近的某个局部区域进行重点搜索。所以,该编码方法具有较强的局部搜索能力。

本文方法结合浮点数编码方法局部搜索能力强以及二进制编码方法全局搜索能力较强的特点,同时采用多种编码方法有效地保持了群体的多样性,进而进一步提高了遗传算法避免早熟收敛的能力。

1.2 多群体进化

多群体遗传算法是在简单遗传算法中,利用并行遗传算法的思想,将群体划分为一些子群体,而各子群体又按一定的模式分别进行独立进化,并在适当的时候,某一些子群体之间还可交换一些信息.这样不仅可以维持群体的多样性,而且能获得抑制早熟现象的效果^[8].

本文方法中,多群体进化结构如图1所示.

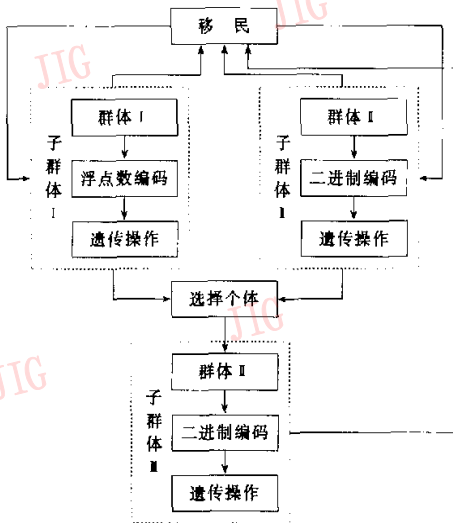


图1 基于多种编码的多群体遗传算法的结构

其中,将群体划分为以下3类群体同时进行进化:①子群体1中的个体采用浮点数编码方法进行进化搜索,使得群体中的进化寻优具有较强的局部搜索能力,其群体规模设为 N_1 ;②子群体2中的个体采用二进制编码方法进行进化搜索,因该方法具有较强的全局搜索能力,且不易陷入局部极值,其群体规模设为 N_2 ;③子群体3为精英种群,它开始并没有个体,它是由前两类子群体进化过程中选取的优秀个体组成的,其作用在于保存前两类子群体进化过程中的优秀个体,以便使优秀个体不遭受破坏,同时子群体3自身也在进化,其群体规模设为 $(N_1 + N_2)/2$.

在进化过程中,为了保持个体分布的多样性,采用了相应的“移民”策略,在子群体1、2之间进行“移民”,即相互交换优秀个体,这样可以吸取其他群体的优点和克服个体趋向,以便增加群体的多样性,其中,“移民”的数量设为 N_1 、 N_2 两者中较小数的一半.同时定期地从前两类子群体中选择优秀个体,以便对子群体3中较差个体进行更新.为加快收敛速度,还定期地用子群体3中最优个体来对群体1、2

中最差个体进行更新.该方法不仅有效地克服了进化过程中的早熟收敛现象,还增加了群体的多样性和加快了收敛速度,并进一步提高了算法性能.

1.3 算法结构

结合浮点数编码方法的强局部搜索能力以及二进制编码方法的强全局搜索能力,并结合多群体并行进化,提出了一种基于多种编码的多群体遗传算法.

该算法的结构可描述如下:

首先对子群体1、2分别进行初始化;然后采用浮点数编码方法和二进制编码方法进行局部搜索和全局搜索;

评价子群体1、2中各个体;

用子群体1、2中的优秀个体初始化子群体3,且子群体3采用二进制编码方法进行进行搜索;

while 不满足终止条件 do

{

对子群体1、2分别进行遗传操作;

用子群体1、2中优秀个体更新子群体3;

用子群体3中优秀个体更新子群体1、2;

子群体1、2之间进行“移民”操作;

}

该方法中,3个群体同时进行进化寻优,并将浮点数编码方法和二进制编码方法分别应用于不同子群体的局部及全局搜索,在进化过程中,通过对优秀个体进行“移民”,既增加了群体的多样性,也增加了找到全局最优解的可能.该方法不仅有效地增加了避免早熟收敛的可能性,也增加了跳出局部最优,趋向全局最优的可能性.

2 试验

因为数学函数优化问题具有普遍意义,且能较好地反映算法本身的实际效能,所以常常用来作为遗传算法的测试对象.这里对5个典型的函数进行优化,将本文提出的方法先与二进制编码遗传算法、格雷码遗传算法和浮点数编码遗传算法3种不同编码机制下的单群体遗传算法进行收敛性能的比较;再将本文方法与基于二进制编码的多群体遗传算法进行收敛性能的比较,其中,前3种单群体遗传算法中,群体规模设为80,本文方法中, N_1 、 N_2 分别设为30及50,用于性能比较的基于二进制编码的多群体遗传算法中,群体规模的设置与本文方法一致.在待测试的各种遗传算法中,均采用文献[4]中提出的两代竞争的选择方法.

用于算法收敛性能检验的标准为:(1)收敛率,

用以反映收敛的效率,即算法成功地找到全局最优解的次数;(2) 收敛代数及收敛时间,分别表示算法收敛时,种群进化代数的平均值及算法收敛所用时间的平均值,用以反映收敛的快慢。

2.1 试验函数

所选函数为如下 5 个典型的测试函数:

$$f_1 = \sum_{i=1}^3 x_i^2 \quad -5.12 \leq x_i \leq 5.12$$

该函数只有一个 (0,0) 为全局最小点,最小值为 0。

$$f_2 = 100 \cdot (x_1^2 - x_2)^2 + (1 - x_1)^2 \quad -2.048 \leq x_i \leq 2.048$$

该函数虽是单峰值函数,但为病态的,难以进行全局极小化,且只有一个全局极小点 $f(1,1)=0$ 。

$$f_3 = \sum_{i=1}^5 \text{int}(x_i) \quad -5.12 \leq x_i \leq 5.12$$

该函数为一不连续函数,对于 $x_i \in [-5.12, -5.0]$ 区域内的每一个点,它都取全局极小值 $f=-30$ 。

$$f_6 = 0.5 - \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1.0 + 0.001(x_1^2 + x_2^2)]^2} \quad 100 \leq x_i \leq 100$$

该函数有无数个局部极大点,其中只有一个 (0,0) 点为全局最大点,最大值为 1。由于此函数的最大峰周围有两圈脊,它们的取值分别为 0.990 284 和 0.962 776,因此优化过程很容易停滞在这些局部极大点。

$$f_8 = \sum_{i=1}^M \frac{x_i^2}{4000} - \prod_{i=1}^M \cos \frac{x_i}{\sqrt{i}} + 1 \quad -600 \leq x_i \leq 600$$

该函数为多维函数,只有当 $x_i=0$ 时,才取全局最小值 $f=0$,而当 $x_i = \pm k\pi \sqrt{i}$ 时,均为局部极小值。所以该函数因很容易陷入局部收敛而较难寻优到最优值,且维数越高,寻优越困难,本文分别对该函数的 10 维及 100 维函数进行寻优计算。

2.2 试验结果及分析

对于测试函数,其参数设置如表 1 所示。

表 1 测试函数的参数设置

函数	最优值	阈值	最大遗传代数	交叉率	变异率
f_1	0.0	0.001	100	0.8	0.1
f_2	0.0	0.001	100	0.8	0.1
f_3	-30	-30	100	0.8	0.1
f_6	1.0	0.999	200	0.8	0.1
$f_8(10 \text{ 维})$	0.0	0.001	5 000	0.8	1.0
$f_8(10 \text{ 维})$	0.0	0.000 1	5 000	0.8	1.0
$f_8(100 \text{ 维})$	0.0	0.001	40 000	0.8	1.0
$f_8(100 \text{ 维})$	0.0	0.000 1	40 000	0.8	1.0

其中,对每一目标函数设置一个阈值作为其收敛准则,当种群最大适应度大于该阈值时,就认为该算法收敛,则此次寻优结束,并进行下一次寻优计算;另外,对每一测试函数,设置遗传算法每次寻优的最大允许进化代数(例如 100 代)作为遗传算法不收敛的结束条件,若遗传算法在最大允许进化代数内,其得到的最优目标函数值超过给定的阈值,则认为算法成功地找到了全局最优解,否则,如没有达到预定的阈值,则认为算法阻滞于局部最优解。

对函数 f_1 、 f_2 、 f_3 和 f_6 ,分别用上述各种方法进行 100 次寻优计算,而对 10 维及 100 维的 f_8 函数,由于计算时间的关系,分别进行 20 次及 10 次寻优计算。表 2 为基于各种方法的遗传算法的性能比较。其中,基于二进制编码方法的多群体遗传算法是指各个子群体均采用二进制编码方法的多群体遗传算法,其种群的数目、种群规模、遗传算法的设计等都采用与本文方法相同的设置。表 2 中,收敛代数是指对同一目标函数重复做寻优试验(例如 100 次),算法收敛时而得到的进化代数平均值,收敛时间是指对同一目标函数重复做寻优试验,算法收敛时所需时间的平均值(以秒 s 单位)。试验中,所用计算机为联想奔月 2000,PⅢ 处理器、933MHz 主频、128M 内存。

表 2 基于各种方法的遗传算法的性能比较

函数	单群体									多群体					
	二进制编码方法			格雷码编码方法			浮点数编码方法			基于二进制编码方法			本文方法		
	收敛率 (%)	代数	时间 (s)	收敛率 (%)	代数	时间 (s)	收敛率 (%)	代数	时间 (s)	收敛率 (%)	代数	时间 (s)	收敛率 (%)	代数	时间 (s)
f_1	100	16.65	0.50	100	24.40	0.81	100	9.60	0.28	100	16.32	0.44	100	5.95	0.18
f_2	60	8.83	0.21	57	7.72	0.19	77	6.42	0.14	45	10.16	0.24	100	7.68	0.19
f_3	95	23.62	1.00	69	29.26	1.35	100	55.30	2.27	70	41.14	1.51	100	6.47	0.28
f_6	4	146.75	3.29	9	107.11	2.63	100	72.47	1.58	45	24.00	0.53	100	10.20	0.24
$f_8(10 \text{ 维})$	0	—	—	0	—	—	100	1 623.85	55.1	0	—	—	100	1 400.7	67.1
$f_8(10 \text{ 维})$	0	—	—	0	—	—	100	2 556.05	86.6	0	—	—	100	2 077.9	99.3
$f_8(100 \text{ 维})$	0	—	—	0	—	—	100	8 129.17	1 951.6	0	—	—	100	5 771.6	1 889.3
$f_8(100 \text{ 维})$	0	—	—	0	—	—	100	25 516.2	6 125.3	0	—	—	100	15 172.6	4 912.2

从上面的试验结果可以看出:

(1)在设定的最大遗传代数内,本文方法都能找到全局最优解,可见本文方法具有很强的跳出局部极值的能力,且不易陷入局部收敛,即使是对极易陷入局部最优的 f_1 函数,也能很快地达到最优.浮点数编码方法虽然对 f_1 函数的收敛性能很好,但对 f_2 函数却易陷入局部收敛.

(2)本文方法的收敛速度快,即使对于较难收敛的 f_1 函数,也能较快地达到全局收敛,且计算量明显少于上述的其他方法.

(3)本文方法中,由于对各子群体采用不同的编码方法,即当子群体1“移民”至子群体2时(或相反),要重新换一种编码方法,而新增加的子群体3还需同时进行遗传算法的操作,所以该方法运行的时间较长.但从表2可看出,虽然本文方法运行的时间较长,但由于该方法的收敛代数远少于其他方法,所以,与其他方法相比,本文方法的收敛时间较少.

3 讨 论

函数最优化技术在航空、航天等众多领域都有重要的应用价值,但对于现实中广泛存在的高维、多峰、非线性、非凸,甚至没有数学模型的函数最优化问题,传统技术常常无能为力,而遗传算法则提供了一条解决复杂优化问题的新途径,但标准遗传算法(SGA)对有的复杂多模函数仍存在寻优效率不高,且易出现早熟收敛,也不便于局部微调的问题.

为了克服遗传算法中的早熟收敛问题,本文将多群体并行进化与多种编码方法相结合,提出了一种基于多种编码的多群体遗传算法.该方法首先利用浮点数编码方法的强局部搜索能力以及二进制编码方法的强全局搜索能力,对两个群体分别同时进行进化,然后通过进化过程中进行“优秀个体的移民”来增加群体的多样性,以避免早熟收敛的产生.同时,引入第3个群体,用于保存进化过程中的优秀个体,且第3个群体本身也进行进化.该方法不仅具有较强的跳出局部收敛的能力,且收敛速度较快.通过对一系列典型函数的测试,也证明了该方法的有效性.

参 考 文 献

- 1 陈国良,王煦法,庄镇泉等.遗传算法及其应用[M].北京:人民邮电出版社,1999.
- 2 Leung Yee, Gao Yong, Xu Zong-Ben. Degree of population diversity — A perspective on premature convergence in genetic algorithms and its markov chain analysis[J]. IEEE Transactions On Neural Networks, 1997, 8(5): 1165~1176.
- 3 许海平,张彤,王子才等.浮点数编码遗传算法及其在电站机组组合优化中的应用[J].小型微型计算机系统, 1999, 20(8): 578~582.
- 4 于海斌,王浩波,徐心和.两代竞争遗传算法及其应用研究[J].信息与控制, 2000, 29(4): 309~314.
- 5 金希东,李治.遗传-突变算法及其在神经网络和控制系统中的应用[A].见:神经网络理论与应用研究[C].成都:西南交大出版社,1996.
- 6 周远晖,陆玉昌,石纯一.基于克服过早收敛的自适应并行遗传算法[J].清华大学学报(自然科学版), 1998, 38(3): 93~95.
- 7 叶在福,单渊达.基于多群体遗传算法的输电系统扩展规划[J].电力系统自动化, 2000, 24(5): 24~35.
- 8 周明,孙树栋.遗传算法原理及应用[M].北京:国防工业出版社,2000.



张莉芬 1977年生,1999年毕业于南昌航空工业学院电子工程系,现为南昌航空工业学院测试与控制系硕士研究生.主要研究方向为遗传算法、图象处理等.



黎明 教授,1985年获上海交通大学电子工程系学士学位,1990年和1997年先后获南京航空航天大学自动控制系及测试工程系硕士和博士学位.主要从事图象处理、模式识别、神经网络和遗传算法等方面的研究工作.已发表学术论文50余篇.



周琳琳 讲师,1994年毕业于南昌航空工业学院电子工程系,现为南昌航空工业学院测试与控制系硕士研究生.主要研究方向为神经网络、遗传算法、图象处理等.