

一种基于种群多样度的实数编码并行遗传算法

刘 胜,李高云,孙天英

(哈尔滨工程大学 自动化学院,黑龙江 哈尔滨 150001)

摘 要:为了改善遗传算法的收敛性能,提出了一种基于个体适应度的种群多样性度量函数,恰当地反映了遗传算法的进化阶段,预报了早熟收敛的趋势.设计了基于种群多样性函数的迁移算子和交叉算子,并对交叉、变异概率等进行了动态调整,构成了具有多层迁移特点的实数编码并行遗传算法.通过和其他优秀遗传算法对测试函数的验证比较,结果表明,该算法对于解决遗传算法中早熟、收敛速度慢等问题具有优越的性能.

关键词:遗传算法;种群多样性;迁移算子;实数编码

中图分类号: TP18 **文献标识码:** A **文章编号:** 1673-4785 (2008) 05-0423-06

A real coding parallel genetic algorithm based on diversity of population

LU Sheng, LI Gao-yun, SUN Tian-ying

(College of Automation, Harbin Engineering University, Harbin, 150001, China)

Abstract: In order to improve the convergence performance of genetic algorithms, a function measuring population diversity on the basis of the degree of individual adaptability was needed. This measuring function must reflect the evolutionary stage of the genetic algorithm and forecast premature trends appropriately. We designed a migration operator and crossover operator based on diversity of population functions, and made dynamic adjustments on crossover and mutation probabilities. This structured a parallel genetic algorithm with real coding and a multi-layer migration operator. Comparative experiments were made on benchmark functions. The results showed that this algorithm is obviously superior to other genetic algorithms in overcoming problems such as prematurity and slow convergence.

Keywords: genetic algorithm; diversity of population; migration operator; real coding

遗传算法是受生物进化理论启发的搜索算法,是由 Holland 在 19 世纪 70 年代提出来的^[1],其本身具备良好的并行设计结构,不易陷入局部最优,并且不依赖于梯度信息,因此特别适用于处理高度复杂的非线性问题,并在函数优化、参数辨识、模式识别、自动控制等许多领域得到广泛地应用.然而,遗传算法作为一种随机搜索方法,存在局部领域搜索不敏感,早熟收敛,寻优速度慢等诸多不足之处.为了改善性能,人们进行了大量的研究,并提出了许多解决的方案. J. Liening 博士通过大量的实验发现^[2]:过多以及过频繁的迁移会破坏子群体的多样性,致使多个搜索进程集中到相同的区域,不利于提高解的质量;过少的迁移以及迁移频率过低,使各子群体不

能充分利用其他子群体的信息,同样不利于提高解的质量.管宇等分析了迁移算子在整个算法收敛过程中的重要作用,提出一种基于模式定理的迁移策略 SMS,提高了算法的时间性能^[3]. Marin 等采用集中式方案^[4],从各进程在其子群体中执行 GA,并周期地将最好的部分结果发送给主进程,在网络上所作的实验表明,运行速度以接近线性加速比增加.但迁移时机的把握还需作进一步的研究. Braun 的方案^[5]是当 GA 在区域上产生早熟收敛时,才进行迁移. Krøger 等则采用了另一种途径,每当一个区域中发现一个改进的个体,就异步地执行迁移操作^[6]. 本文在以上基础上,提出了一种新的种群多样性度量函数,并基于此设计引入迁移算子的实数并行遗传算法,同时对实数交叉等方面也做了改进.算例结果表明,该算法具有良好的防止早熟特性,并且增强了全局最优解的搜索能力,能较快地搜索到全局最

收稿日期:2008-03-01.

基金项目:黑龙江省自然科学基金资助项目(A200419).

通信作者:李高云. E-mail: ligaojun@hrbeu.edu.cn

优解.

1 一种新的种群多样性度量

在遗传算法中,人们最关注的问题之一就是它的过早收敛.近几年的研究发现:种群只有在保持一定的多样性的基础上才能进化,过早收敛总是与种群中个体趋同、种群多样性 (population diversity, PD) 的迅速下降有密切的关系.遗传算法中多个遗传算子参数需要根据种群的 PD 进行自适应地调整.因此,如何合理、有效地刻画种群的多样性是遗传算法急需解决的问题.常用的测度方法是采用海明距离,但海明距离通常只适合二进制编码,从计算量的角度来讲,海明距离方法运算量又显得较大.因此,本文从个体表现性的角度出发,根据个体适应度来定义了一种可以应用到各类遗传算法,而且能恰当充分反映种群多样性的度量函数.

假设种群 pop 含有 M 个个体,分别为 P_1, P_2, \dots, P_M ,其所对应的适应度值为 f_1, f_2, \dots, f_M ,那么种群的平均适应度可记为

$$\bar{f} = \frac{1}{M} \sum_{i=1}^M f_i \quad (1)$$

则种群的多样性定义为

$$D = \frac{1}{M} \sum_{i=1}^M \left(\frac{f_i - \bar{f}}{f_i + \bar{f}} \right)^2 \quad (2)$$

由于在遗传算法设计适应度函数时,为了选择的方便且具有意义,一般情况下 $\forall i$ 满足 $f_i > 0$,那么就有 $\bar{f} > 0$,由此可以得到:

$$\frac{f_i - \bar{f}}{f_i + \bar{f}} = \frac{f_i - \bar{f}}{f_i + \bar{f}} \quad (3)$$

所以, $0 \leq \frac{f_i - \bar{f}}{f_i + \bar{f}} \leq 1$ 得到:

$$1 - \left(\frac{f_i - \bar{f}}{f_i + \bar{f}} \right)^2 \geq 0 \quad (4)$$

从而得到 $0 \leq D \leq 1$,种群多样性函数的曲线如图 1 所示.

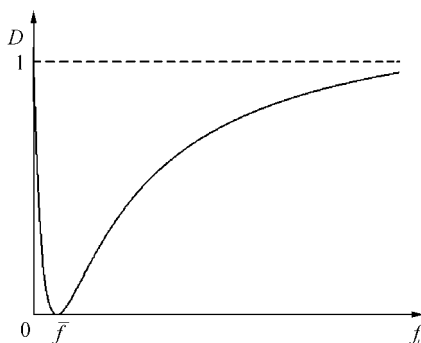


图 1 种群多样性函数

Fig 1 Diversity of population function

当 f_i 远离 \bar{f} 时,多样性函数值趋近于 1,而 f_i 接近于 0 时,其值同样接近于 1,且呈现出以 \bar{f} 为中心单

调形式,这恰好符合了种群多样度的各项要求.

本文为了检测种群多样性函数 D 值的性能,做了以下检验,在第 i 次随机产生 10 000 个 $[0, 1]$ 之间的随机数,然后用 1 替换其中的 100*i* 个随机数, $i \in [1, 100]$,分别计算种群多样性 D 的值,得到随机种群多样性曲线,如图 2 所示.

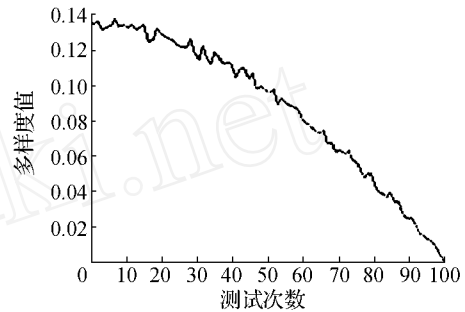


图 2 随机种群多样性

Fig 2 Diversity of random population function

测试曲线中, D 值逐渐较稳定的从 0.14 左右降为 0,表明本文提出的 D 函数对于种群的多样性具有极良好的反映性能.当 D 越接近于 1,种群的多样性就越高,越不易于陷入早熟;反之接近于 0 时,种群很可能陷入到局部最优解之中,这时就需要采取其他有效的方式来使遗传算法跳出该局部最优解.

2 基于种群多样度的迁移算子

生物界具有大规模的迁移习性,通过这种行为,往往使种群得到极大的发展,而遗传算法本质上是一种模拟生物进化的过程,本文从中得到启发,设计了迁移算子改进遗传算法.

然而,迁移时机很难把握.进一步分析遗传算法,由于选择算子的作用,加之人们对遗传算法收敛性的要求,较为理想的遗传算法的种群多样性呈现出总体缓慢递减,随后逐步稳定的一个过程.当遗传算法早期陷入到局部最优解时,解将在这一局部最优解集中,同时遗传算法的多样性会呈现出较快下降的态势.依据以上特点,本文采用监测种群多样性 D 函数的变化,从而适时的引入迁移算子,解决迁移时机很难把握的难题.

当种群多样性函数 D 低于某一预先设定的值 $(T) = \phi - \phi T / T_2$ 时,对种群 pop 进行整体迁移替换,即将种群中个体适应度最差的 $(100n\%)$,用相应的种群来替换更新,称此种操作为整体迁移算子. (T 为当前种群代数, T_2 为延拓代数,规定为进化种群代数的 1.5 倍, ϕ 为整体迁移因子,在下面的测试实验中取 ϕ 为 0.065, T_2 为 150, $n=3$).

整体迁移算子的引入实质是为了使种群保持一

定的多样度,调整种群的多样度按期望的梯度下降有利于遗传算法的全局收敛.整体迁移算子引入准则示意图如图 3 所示.

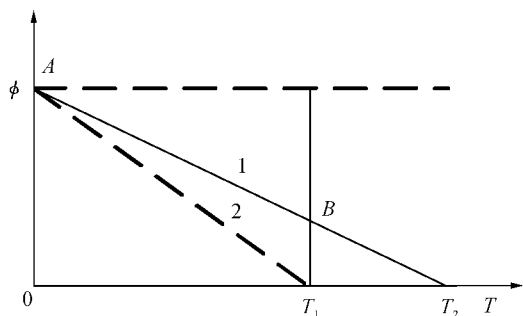


图 3 迁移算子引入准则示意图

Fig 3 Migration operator import rule

图 3 中, T 为进化代数, T_1 为种群总进化代数. 图中 AB 线以上 (即 1 区) 为期望的种群多样度变化域, 随着种群代数的进化呈现出线性递减规律; 若在进化的过程中多样度低于 AB 线 (即 2 区) 时, 遗传算法有可能陷入到局部最优解当中, 在此种情况下需要对种群中的最差个体进行随机替换. 遵循陷入早熟代数越早, 替换越多种群的原则, 本文采用正比于多样度函数下降斜率替换的策略.

上述设计方式的优点在于:

- 1) 由于种群多样度函数具有监测作用, 预报了遗传算法陷入早熟的可能性概率和时间;
- 2) 一旦陷入到早熟状态, 可以根据所处进化阶段的不同情况, 自适应地控制随机迁移算子引入的多寡;
- 3) 可以根据计算的要求, 动态地调整种群延拓代数的大小, 方便对选择压力进行自适应地调节.

3 并行遗传算法算子设计及其流程

3.1 基于种群多样度的交叉算子改进

在文献 [9] 中给出了关于交叉算子的改进算法, 但其计算的方案并未根据遗传算法的进化过程进行动态地调节. 而实际过程中, 当种群进化到一定阶段后, 就必须在所搜索到的个体中进行重点的搜索, 而不应该与算法初期搜索的方式一样. 由于种群多样度函数恰当地反映了遗传算法的进化阶段, 因此本文将多样度函数嵌入到交叉策略中, 对于文献 [9] 中的搜索方式进行改进.

2 个待交叉个体为 x_i, x_j , 判断其适应度的高低, 并记为 $f_{\text{high}}, f_{\text{low}}$, 相应的个体记为: x_h, x_l . 令交叉的搜索方向 $d = \frac{f_{\text{high}} - f_{\text{low}}}{f_{\text{high}}}$, 即一个个体沿着方向 d 从区域 f_{low} 向另一个区域 f_{high} 前进, 则实数交叉策略操作如

下:

$$\begin{cases} x_h^{t+1} = x_l^t + (1 - D) \cdot / x_h^t - x_l^t / \cdot d, \\ x_l^{t+1} = x_h^t + (1 + D) \cdot / x_h^t - x_l^t / \cdot d \end{cases} \quad (5)$$

式中: D 为种群多样度的大小. 相比于文献 [9], 该算法除了保持原有算法的优点外, 由于在遗传算法初期 D 较大, 而在后期会逐渐减小, 因此在算法初期扩大了搜索的范围, 而在算法后期加速了算法的收敛, 提高了算法的整体性能.

3.2 参数的自适应调整

一方面, 为了使遗传算法具有更强的鲁棒性、全局最优性和效率, 当算法提前收敛时, 加大 p_c 和 p_m , 在算法后期减小 p_c 和 p_m ; 另一方面, 为了保持优良个体, 不能对所有的个体采取同一个 p_c 和 p_m , 本文结合 Striniras 的自适应交叉和变异率, 对其进行改进设计 p_c 和 p_m , 如下 [8]:

$$p_c = \begin{cases} \frac{P_{c \max} - P_{c \min}}{1 + \exp(A(1 - \frac{2(f_{\text{avg}} - f_{\max})}{f_{\text{avg}} - f_{\min}}))} + P_{c \min}, & f_{\max} \leq f_{\text{avg}} \\ P_{c \min}, & f_{\max} > f_{\text{avg}} \end{cases} \quad (6)$$

$$p_m = \begin{cases} \frac{P_{m \max} - P_{m \min}}{1 + \exp(A(1 - \frac{2(f_{\text{avg}} - f_{\max})}{f_{\text{avg}} - f_{\min}}))} + P_{m \min}, & f_{\max} \leq f_{\text{avg}} \\ P_{m \min}, & f_{\max} > f_{\text{avg}} \end{cases} \quad (7)$$

式中: $A = 9.903438$, f_{\max}, f_{\min} 分别为交叉的 2 个个体的高、低适应度值, f_{avg} 为该群体的平均适应度, $P_{c \max}, P_{c \min}, P_{m \max}, P_{m \min}$ 为个人设定的交叉最大概率、交叉最小概率、变异最大概率、变异最小概率. 本文在下面的测试过程中本文取子群体 $P_{c \max}, P_{c \min}, P_{m \max}, P_{m \min}$ 为 1, 0, 0.01, 0, 取主群体 $P_{c \max}, P_{c \min}, P_{m \max}, P_{m \min}$ 为 0.7, 0, 0.0005, 0.

3.3 算法并行结构设计及其流程

算法采用并行结构, 共有 3 个子种群, 并由此产生一个主群体和一个融合种群. 本文提出“整体迁移算子 + 相互迁移”的思想, 即各个子群体本身实现由上文提出的整体迁移算子, 并在进化过程中, 子群体间也进行种群相互迁移. 其迁出、迁入的个数由各种群进化效果决定 [10-11], 但其迁移总数目保持不变, 具体如式 (8) 所示:

$$E = af_{\text{elist}} + bf + cf \quad (8)$$

式中: f_{elist} 为种群最优个体的适应度; f 为种群的平均适应度; f 为相邻 2 次个体迁移间种群平均适应度的变化, 表示种群的进化速度; a, b, c 为常数, 且 $a + b + c = 1$, 一般取 $a > b > c$.

迁移后的种群大小由式 (9) 决定.

pop size_i = $\frac{E_i}{\sum_{i=1}^3 E_i} \cdot \text{pop size}_r$ (9)

所设计的 MPGA 主要算法过程描述如下：

- 1)初始化子种群 P_1 、 P_2 、 P_3 ;
- 2)选择各个子种群中适应度较高的前 1/3 个体构成主进化群体 A ;
- 3)选择各个子种群中适应度较高的前 1/3 个体构成融合群体 B ;
- 4)从融合群体中选择,替换每个子群体中的最差的 1/3 个体;
- 5)检测各个子群体多样性,并相应调整各遗传算子,当种群多样性低于 (T) 时,实现自适应整体迁移算子;
- 6)检测主群体多样性,是否需要子群体间的互相迁移,若需要则进行互相迁移;
- 7)各个子群体、主群体进行选择、交叉、变异等遗传操作;
- 8)检测子群体最优个体是否优于主群体,倘若 是,则用最优个体替换掉主群体最差个体;
- 9)检验主群体是否满足设计计算指标,满足则 停止运算,否则转到 2)步。

程序流程图如图 4 所示。

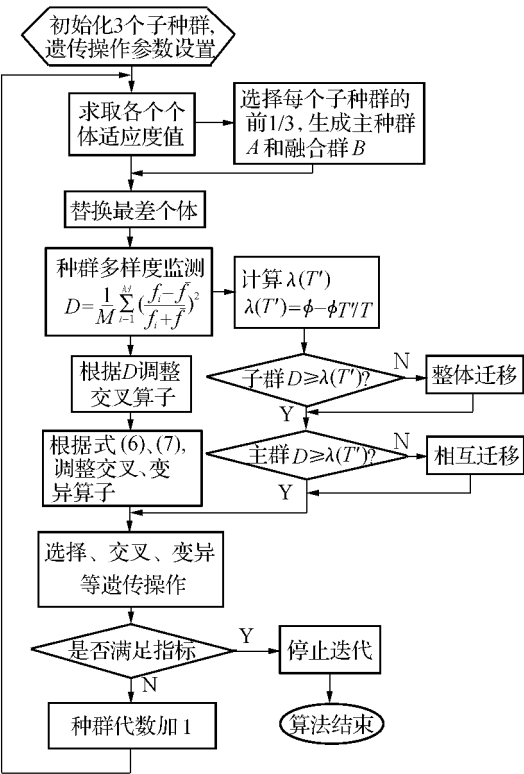


图 4 算法结构流程图

Fig 4 Flow chart of algorithm

4 算例分析

数值函数测试不需要专门的领域知识,且能很好地反映算法本身的实际效能,因此精选了 2 个测试函数,从不同的角度来检测本文设计算法的性能, 2 个函数如下分别如下:

$F_1 = 100(x_1^2 - x_2)^2 + (1 - x_1)^2 - 5.12x_1 + 5.12x_2$ (10)

$F_2 = 0.5 - \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{(1 + 0.001(x_1^2 + x_2^2))^2}$ (11)

当上述函数满足 $-5.12 \leq x_1 \leq 5.12$ 时, F_1 最小值位于 (0, 0) 点, F_2 的最大值位于 (0, 0) 点. F_2 是一个快速变化的多模态函数,有无数个局部极大值点,且全局最优点周围有一圈脊,取值均为 0.990 283,可以参考图 5 的三维模型,其越接近于原点,函数变化越剧烈,在最优点附近形成很密、很陡的振荡峰,可以很好考察函数的全局最优收敛性。

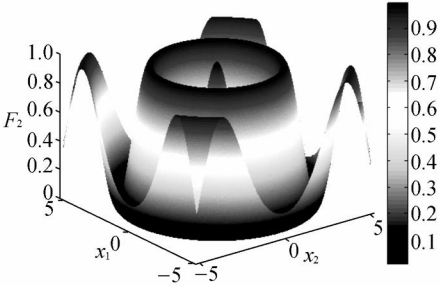


图 5 F_2 三维曲线图

Fig 5 F_2 three-dimensional curves

为了比较各个算法的具体性能,本文分别采用 5 种方法对以上 2 个函数进行测试,其中算法 1 为标准的遗传算法;算法 2 为文献 [7] 提出的算法;算法 3 为文献 [9] 提出的算法,算法 4 采用本文提出的不含有整体迁移算子的 MPGA,算法 5 为本文提出的 MPGA,MPGA 具体的参数设置如表 1 所示。

表 1 MPGA 参数设定

Table 1 Parameter setting of MPGA

MPGA 参数	
子群体的迁移频率 / 代	10
子群体迁移百分比 / %	20
子群体的规模	50
选择策略	最优保留
编码方式	实数编码
种群代数	100

其中整体迁移算子、自适应交叉、变异参数已经由前面几节给出,这里不再累述.本文运用以上算法,对每个测试函数都重复计算 100 次.

由于仿真曲线数目众多,下面给出针对 F_2 函数测试的其中一次仿真曲线图,如图 6 所示.其中,横坐标为迭代次数,纵坐标为算法进化精度.

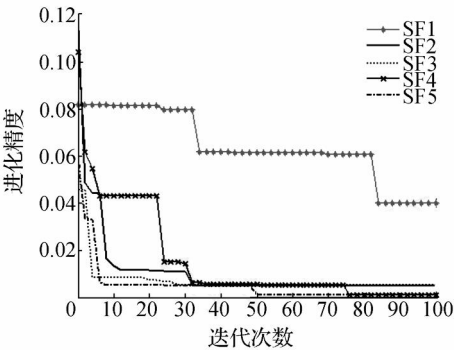


图 6 收敛精度曲线

Fig 6 Curves of convergence precision

从图 6 中可以清晰看到,在进化 30 代以后,收敛几乎处于了停滞状态,此时种群个体趋同,种群多样度值迅速下降,算法 1、2 和 3 在此次仿真中最终没能达到预期的收敛精度;本次仿真中算法 4 在 75

代左右达到了预期收敛精度,而算法 5 在 48 代左右就达到了预期收敛精度.表明 MPGA 算法(算法 5)具有较好的防止早熟、跳出局部最优解的能力.

针对以上各种算法具有随机性的因素,本文对以上算法做了统计性的测试,统计结果如表 2 所示,给出了对上述 5 种算法测试的统计性能比较.从表 2 中可看出,尽管函数具有多峰性、凹凸性和连续性,MPGA 找到全局最优解的概率还是较大的,比其他 4 种算法均有所改善. MPGA 达到精度的平均代数比文献中其他算法有所降低,特别是在处理非线性严重的 F_2 测试函数时,其性能更加的明显,这进一步说明了 MPGA 在处理高度非线性函数寻优时的优越性.与算法 4 相比较,由于其加入了整体迁移算子,因此对于早熟现象有更好的解决能力,在算法陷入早熟的时候,整体迁移算子给予了种群二次发展的机会,使其跳出了 F_2 测试函数的局部最优解,发挥了整体迁移算子对多样度调整的性能,但在运算时间的比较上,由于 MPGA 采用了较多的算子,相比于其他算法,其计算时间略微有所增加,但增加的幅度很小,相对于计算效率而言,这样的损耗是值得的(本算法实现环境为 P4 处理器,2.8G 主频,256M 内存).

表 2 各个函数性能测试

Table 2 Test of each function capability

函数	精度	收敛次数 达到精度平均代数					达到指定精度平均运行时间 /s				
		算法 1	算法 2	算法 3	算法 4	算法 5	算法 1	算法 2	算法 3	算法 4	算法 5
F_1	0.000 01	43/80	77/61	83/59	91/43	100/40	0.15	0.18	0.17	0.18	0.19
F_2	0.000 01	32/93	55/90	52/94	70/90	91/95	0.41	0.47	0.50	0.52	0.53

5 结束语

本文提出了一种新的种群多样度量函数,基于在适当的时机引入整体迁移算子,并利用多样度进行交叉方式的改进,采用二层迁移的并行结构体系,对于改善遗传算法的性能具有一定程度的提高,且可以很方便地在其他各类遗传算法中实现.

参考文献:

[1] HOLLAND J H. Adaptation in natural and artificial systems [M]. Cambridge, Massachusetts: MIT Press, 1992: 87-95.
[2] LIEBIG J. A parallel genetic algorithm for performance driven VLSI routing[J]. IEEE Trans on EC, 1997, 1(1): 29-39.
[3] 管宇,徐宝文. 基于模式迁移策略的并行遗传算法

[J]. 计算机学报, 2003, 5(3): 294-301.
GUAN Yu, XU Baowen. Parallel genetic algorithms with schema migration[J]. Chinese Journal of Computers, 2003, 5(3): 294-301.
[4] AKRA S. Experimental results in distributed genetic algorithms[C]// Applied Corporate Computing Monterrey, Mexico, 1994: 99-108.
[5] BRAUN H. On solving travelling salesman problems by genetic algorithms[M]. Berlin: Springer-Verlag, 1991: 129-133.
[6] KROGER B, SCHWENDERLING P, VORNBERGER O. Parallel genetic packing of rectangles[M]. Berlin: Springer-Verlag, 1991: 160-164.
[7] SRINIVAS M, PATNAIK L M. Adaptive probabilities of crossover and mutation in genetic algorithms [J]. IEEE Transactions on Systems, Man and Cybernetic, 1994, 24

(4): 655-677.

- [8] 陈小平, 于盛林. 实数遗传算法交叉策略的改进 [J]. 电子学报, 2003, 1(1): 71-74.

CHEN Xiaoping, YU Shenglin. Improvement on crossover strategy of real-valued genetic algorithm [J]. Acta Electronica Sinica, 2003, 1(1): 71-74.

- [9] 吴好阳, 朱长春, 刘建华. 自适应遗传算法改进种群早熟收敛 [J]. 西安交通大学学报, 1999, 33(11): 27-30.

WU Haoyang, ZHU Changchun, LI Jianhua. Adaptive genetic algorithm to improve group premature convergence [J]. Journal of Xi'an Jiaotong University, 1999, 33(11): 27-30.

- [10] 黄晓峰, 潘立登, 陈标华, 等. 用改进的实数编码遗传算法估计反应动力学参数 [J]. 高校化学工程学报, 1999, 13(1): 50-55.

HUANG Xiaofeng, PAN Lideng, CHEN Biahua, et al. Estimating reaction kinetics parameters with an improved realcoded genetic algorithm [J]. Journal of Chemical Engineering of Chinese University, 1999, 13(1): 50-55.

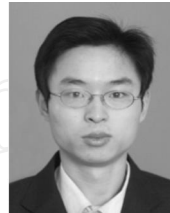
- [11] 李波, 邱枫. 基于单亲遗传算法的动态设备布局仿真研究 [J]. 智能系统学报, 2007, 2(1): 74-79.

LI Bo, QIU Feng. Simulation of the dynamic plant layout problem based on partheno genetic algorithm [J]. CAAI Transactions on Intelligent System, 2007, 2(1): 74-79.

作者简介:



刘胜, 男, 1957年生, 教授, 博士生导师, 曾获国防科学技术进步三等奖、中国船舶工业集团科技进步三等奖、黑龙江省科技进步二等奖. 主要研究方向为智能控制、鲁棒控制、船舶姿态控制.



李高云, 男, 1981年生, 博士研究生, 主要研究方向为智能控制、故障诊断与容错控制、船舶姿态控制.



孙天英, 男, 1981年生, 硕士研究生, 主要研究方向为智能控制、船舶姿态控制.

2009 IEEE机电一体化与自动化国际会议 The 2009 IEEE International Conference on Mechatronics and Automation (ICMA 2009)

The 2009 IEEE International Conference on Mechatronics and Automation (ICMA 2009) will take place in Changchun, Jilin, China from August 9 to August 12, 2009. The objective of ICMA 2009 is to provide a forum for researchers, educators, engineers, and government officials involved in the general areas of mechatronics, robotics, automation and sensors to disseminate their latest research results and exchange views on the future research directions of these fields. The topics of interest include, but not limited to the following:

- Intelligent mechatronics, robotics, biomimetics, automation and control systems
- opto-electronic elements and Materials, laser technology and laser processing
- Elements, structures, mechanisms, and applications of micro and nano systems
- Teleoperation, telerobotics, haptics, and teleoperated semi-autonomous systems
- Sensor design, multi-sensor data fusion algorithms and wireless sensor networks
- Biomedical and rehabilitation engineering, prosthetics and artificial organs
- Control system modeling and simulation techniques and methodologies
- AI, intelligent control, neuro-control, fuzzy control and their applications
- Industrial automation, process control, manufacturing process and automation

All papers must be submitted in PDF format prepared strictly following the IEEE PDF Requirements for Creating PDF Documents for IEEE Xplore®. The standard number of pages is 6 and the maximum page limit is 8 pages with extra payment for the two extra pages. See detailed instructions in the conference web site. All paper accepted by IEEE ICMA 2009 will be indexed by EI and included in IEEE Xplore®.

Website: <http://www.ieee-icma.org/>