

采用改进的粒子群算法训练 CNNE 模型

杨春玲,王健来,朱 敏

(哈尔滨工业大学 电气工程及自动化学院,黑龙江 哈尔滨 150001)

摘 要:提出用人工智能算法——粒子群优化算法(PSO)对 CNNE 模型进行训练,并针对标准粒子群算法易限于局部极小点的局限性,采用了一种带有梯度加速的粒子群算法,通过引入梯度信息来影响粒子速度的更新.为防止陷入局部最优,在群体最优信息陷入停滞时,对部分粒子进行重新初始化,从而保持群体的活性,减小群体陷入局优的可能性.采用粒子群算法训练的 CNNE 模型较原来的分布式最速下降法而言,在保证精度的前提下,提高了算法的收敛速度,解决了发射率的在线实时测量问题.

关键词:CNNE 模型;粒子群;梯度

中图分类号:TP183 **文献标识码:**A **文章编号:**1673-4785(2007)03-0067-06

Using the improved particle swarm optimization to train the CNNE model

YANG Chun-ling, WANG Jian-lai, ZHU Min

(School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China)

Abstract: An artificial intelligence algorithm—particle swarm optimization (PSO)—is proposed to train the CNNE model. To overcome a limitation of standard particle swarm optimization, which can easily be restricted at a local optimum point, a particle swarm optimization (PSO) algorithm with gradient acceleration was created that adds gradient information to periodically force updating of the velocities of particles. When the optimum information for the swarm is stagnant, some particles in the population are initialized again to reduce the possibility of becoming trapping in a local optimum. Compared to the step steepest descent algorithm, training the CNNE model with the particle swarm optimization algorithm can improve the speed of convergence of the algorithm while keeping its precision, which solves the challenge of real-time measurement of emissivity.

Keywords: CNNE model; particle swarm optimization; gradient

光谱发射率是目标红外辐射特性的一个重要参量.对其测量可以了解材料的红外辐射特性,这对于材料科学的发展,从而进一步促进航天航空、国防科技事业的进步有着重要意义.

在文献[1-2]中,建立了激活函数可调(tunable activation function)的组合神经网络发射率模型(简称 TAF-CNNE 模型)用于目标发射率的辨识,并用分布式最速下降法对神经网络模型进行了训练.仿真结果表明该方法搜索精度高,但由于有梯度运算,结构比较复杂,迭代次数较大,且对初值变化

比较敏感,导致收敛速度较缓慢,不能满足发射率在线实时测量问题.因此引入一种新颖的人工智能算法——粒子群优化算法用来对 CNNE 模型进行训练,在保证精度的前提下,大大提高了收敛速度.

粒子群优化算法(particle swarm optimization, PSO)是由 Kennedy 和 Eberhart 于 1995 年提出的一种基于群体智能(swarm intelligence)的进化计算技术,其基本思想源于对鸟群捕食等群体行为的研究^[3-4].目前,PSO 已广泛应用于函数优化、神经网络训练、模式分类、模糊系统控制以及其他的应用领域.近几年的研究和实践表明,PSO 在多维空间函数寻优、动态目标寻优等方面有着收敛速度快、解质量高、鲁棒性好等优点,特别适合工程应用^[5].

收稿日期:2006-10-13.

基金项目:国家自然科学基金资助项目(60377037);黑龙江省自然科学基金资助项目(F2004-13).

PSO 算法在搜索的初期收敛速度很快,但在后期却易于陷入局部极优点,尤其当解空间为非凸集时,这是 PSO 算法的主要缺点.针对这一缺点,本文在用 PSO 算法搜索解空间的同时加入了梯度信息,从而帮助算法摆脱易于陷入局部最优点的束缚,同时又保持搜索速度快的特性,并经过仿真测试验证了该算法的优越性.

1 粒子群算法基本原理

使用 PSO 方法解决优化问题,即把问题的潜在解定义为搜索空间中的一个“粒子”,用适应度值来评价这个粒子的好坏,适应度值是由目标函数唯一确定的.粒子群优化算法依靠个体间的信息交换来达到整个群体的共同演化.

首先,粒子群初始化为可行解空间的一组初始值^[5],然后通过迭代找到最优解.在每一次迭代中,粒子通过跟踪 2 个“极值”来更新自己:一个是该粒子经历过的最好位置(有最好的适应度值),即该粒子本身所找到的最优解,这个解叫做个体极值,用 P_{best} 来表示;另一个是群体中所有粒子经历过的最好位置(有最好的适应度值),即整个种群目前所找到的最优解,这个解叫做全局极值,用 G_{best} 来表示.

假设在一个 D 维的目标搜索空间中,有 m 个粒子组成一个群体,其中第 i 个粒子($i = 1, 2, \dots, m$)的位置为 X_i ,表示为 $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T$.因为每个粒子的位置 $P_{best_i} = (p_{best_{i1}}, p_{best_{i2}}, \dots, p_{best_{iD}})^T$,就是一个潜在解,将 X_i 代入目标函数就可以计算出其适应度值,根据适应度值的大小衡量其优劣.设该粒子经历过的最好位置 $P_{best_i} = (p_{best_{i1}}, p_{best_{i2}}, \dots, p_{best_{iD}})^T$,整个群体所有粒子经历过的最好位置为 $G_{best} = (g_{best_1}, g_{best_2}, \dots, g_{best_D})^T$,粒子 i 的速度表示为 $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})^T$,那么粒子 i 根据下面的规律更新自己的速度和位置:

$$v_{id}^{k+1} = v_{id}^k + c_1 \text{rand}_1^k (p_{best_{id}}^k - x_{id}^k) + c_2 \text{rand}_2^k (g_{best_d}^k - x_{id}^k), \quad (1)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}. \quad (2)$$

式中: $i = 1, 2, \dots, m$; $d = 1, 2, \dots, D$; k 为当前迭代次数; $\text{rand}_1^k, \text{rand}_2^k$ 是 $[0, 1]$ 区间均匀分布的随机数; v_{id}^k 是粒子 i 在第 k 次迭代中第 d 维的速度; $p_{best_{id}}^k$ 是粒子 i 在第 k 次迭代中第 d 维的位置; $p_{best_{id}}^k$ 是粒子 i 在第 k 次迭代中第 d 维的个体极值点的位置(即坐标); $g_{best_d}^k$ 是整个种群在前 k 次迭代中第 d 维的全局极值点的位置. c_1, c_2 是加速系数(或称学习因子),分别调节向全局最好粒子和个体最好粒子方向飞行的最大步长,若太小,则粒子可能远离目标区域,若太大,则会导致突然向目标区域飞去,或飞过目标区域,合适的 c_1, c_2 可以加快收敛且不易陷入局部最优.称为惯性因子,较大适于对解空间进

行大范围搜索,较小适于进行小范围挖掘.对全局搜索,通常较好的方法是在前期有较高的探索能力以得到合适的种子,而在后期有较高的开发能力以加快收敛速度.因此,可将 w 设为随时间或者迭代次数的增加而减小.一般可用下式来确定: $w =$

$$\frac{\max - \min}{\text{最大迭代次数}} \times \text{迭代次数}, \text{式中: } \max, \min \text{ 分别为开始时和结束时的权重,一般取 } \max = 0.9, \min = 0.4.$$

注意:粒子在第 d ($1 \leq d \leq D$) 维的位置变化范围为 $[- X_{\max_{id}}, X_{\max_{id}}]$,速度变化范围 $[- V_{\max_{id}}, V_{\max_{id}}]$,迭代中若位置和速度超过边界范围,则取边界值.设定较大的 $V_{\max_{id}}$ 可以保证粒子种群的全局搜索能力, $V_{\max_{id}}$ 较小则粒子种群的局部搜索能力加强.通常 $V_{\max_{id}} = k X_{\max_{id}}, 0.1 \leq k \leq 1.0$.

Maurice Clerc^[6]和 F. V. D. Bergh^[7]对于 PSO 的收敛性和稳定性作了初步分析,并给出一些保证 PSO 算法收敛的参数条件.

PSO 算法流程图如图 1 所示,具体步骤如下:

- 1) 设定种群规模,给定求解精度和最大迭代次数;
- 2) 随机初始化粒子种群:初始化种群中所有粒子的速度和位置(即可行解);
- 3) 使用根据优化问题目标定义的适应度函数对所有粒子进行评价;
- 4) 对每个粒子 i ,将其适应度与其经历过的最好位置 P_{best_i} 的适应度作比较,如果较好,则将 X_i 作为当前的最好位置 P_{best_i} ;
- 5) 对每个粒子 i ,将其适应度与其经历过的最好位置 G_{best} 的适应度作比较,如果较好,则将 X_i 作为当前所有粒子的最好位置 G_{best} ;
- 6) 根据式(1)和式(2)更新粒子的速度和位置;
- 7) 检查是否满足迭代终止条件,若是,则终止迭代,求得最优值;若否,转至 3),直到满足算法的迭代终止条件为止.迭代终止条件根据具体问题设定,一般选为是否达到预定最大迭代次数或(和)粒子群目前为止搜索到的最优位置是否满足预定最小适应阈值.

2 CNNE 模型的粒子群算法实现

参照文献[1]可知对 CNNE 模型训练的最终结果是得到一组最佳系数 W, B_1, V, B_2 ,即待优化问题的一组最优解.而在 PSO 算法中,把粒子看作是待优化问题的潜在解,所以在用粒子群算法实现 CNNE 模型的时候,每个粒子是由系数 W, B_1, V, B_2 构成的.

在 CNNE 模型中,取隐层节点数为 6 ($K = 6$).可以确定权系数 W 为 1×6 维向量 $[w_1, w_2, \dots, w_6]$,隐层偏差 B_1 为 1×6 维向量 $[b_{11}, b_{12}, \dots, b_{16}]$,

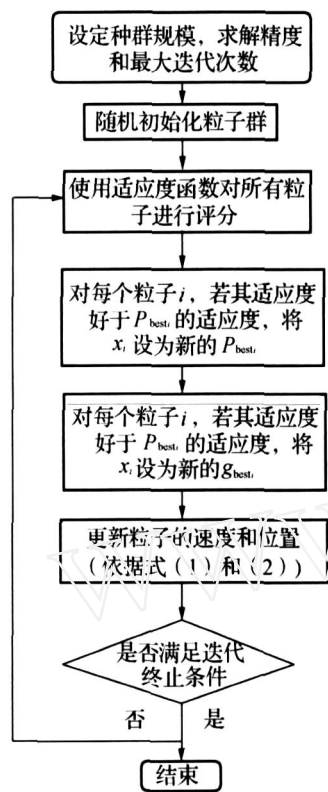


图 1 PSO 的算法流程

Fig. 1 Flow chart of PSO algorithm

权系数 V 为 6×1 维向量 $[v_1, v_2, \dots, v_6]^T$, 输出层偏差 B_2 为 1×1 一维向量 $[b_2]$, 所以每个粒子 (particle) 的解空间为 $6 + 6 + 6 + 1 = 19$ 维, 其排列形式如式 (3) 所示:

$$\begin{bmatrix} W^T \\ B_1^T \\ V \\ B_2 \end{bmatrix}_{19 \times 1} \quad (3)$$

在 PSO 优化算法中, 粒子数的增加意味着更快的求解速度, 但也意味着更大的计算量. 这里选粒子数 (particle) 为 20. 那么 20 个粒子的解向量可以用一个维的矩阵来表示, 其通用形式如式 (4) 所示:

$$\begin{bmatrix} W_1^T & W_2^T & \dots & W_{20}^T \\ B_{11}^T & B_{12}^T & \dots & B_{120}^T \\ V_1 & V_2 & \dots & V_{20} \\ B_{21} & B_{22} & \dots & B_{220} \end{bmatrix}_{19 \times 20} \quad (4)$$

在 PSO 算法中, 用适应度函数来判断每个粒子的好坏, 一般取适应度函数为误差函数. 常用的误差函数 J 有 $J_1 = \int_0^T |e(t)| dt$, $J_2 = \int_0^T e^2(t) dt$, $J_3 = \int_0^T te^2(t) dt$ 和 $J_4 = \int_0^T t|e(t)| dt$. 将 $\int_0^T e^2(t) dt$ 离散化得 $E = \frac{1}{2} \sum_{p=1}^L (d^p - \hat{y}^p)^2$, 这是在神经网络中常用的均方误差公式. 在 CNNE 模型中, 设 $d_{i,j} = \frac{1}{T_i} - \frac{1}{T_j}$

为样本数据, 若波长数为 n (这里 $n = 15$), 则输入样本总数为 $P = C_n^2 = C_{15}^2 = 105$ 个, 依据 $E = \frac{1}{2} \sum_{p=1}^L (d^p - \hat{y}^p)^2$, 则误差函数可以取 $E = \frac{1}{2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (d_{i,j} - \hat{y}_{i,j})^2$, 总共为 105 个数据的误差和.

3 仿真结果分析与改进

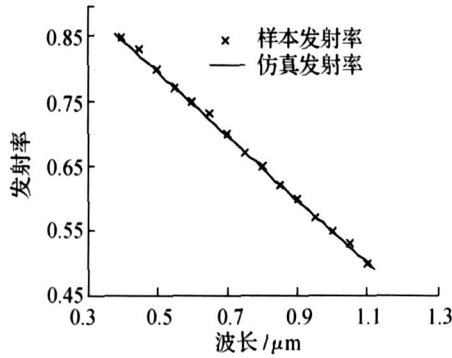
用 Matlab 对程序进行仿真, 样本数据选择表 1 中 A、B、C 3 种不同变化率的发射率模型. 图 2 是假定的这 3 种发射率模型与仿真发射率模型的比较图; 图 3(a)、图 3(b) 分别是采用分布式最速下降法和 PSO 算法训练的发射率模型 A 的误差函数曲线.

表 1 CNNE 模型的发射率样本数据
Table 1 Emissivity sample data of the CNNE model

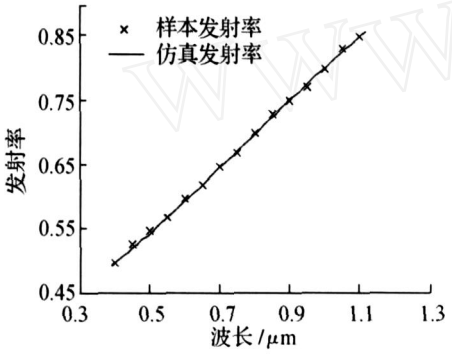
通道	波长 / μm	发射率		
		模型 A	模型 B	模型 C
1	0.40	0.85	0.50	0.85
2	0.45	0.83	0.53	0.83
3	0.50	0.80	0.55	0.80
4	0.55	0.77	0.57	0.77
5	0.60	0.75	0.60	0.75
6	0.65	0.73	0.62	0.73
7	0.70	0.70	0.65	0.70
8	0.75	0.67	0.67	0.69
9	0.80	0.65	0.70	0.70
10	0.85	0.62	0.73	0.73
11	0.90	0.60	0.75	0.75
12	0.95	0.57	0.77	0.77
13	1.00	0.55	0.80	0.80
14	1.05	0.53	0.83	0.83
15	1.10	0.50	0.85	0.85

观察图 2 可以看出, 对于线性变化的样本发射率模型, 即样本发射率模型 A 和样本发射率模型 B, 采用 PSO 算法训练出的发射率能够较好的逼近

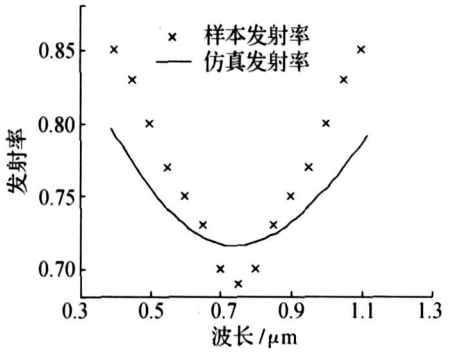
样本发射率曲线;但是对于非线性变化的(如正弦)样本发射率模型,逼近程度则不好.



(a) 模型 A

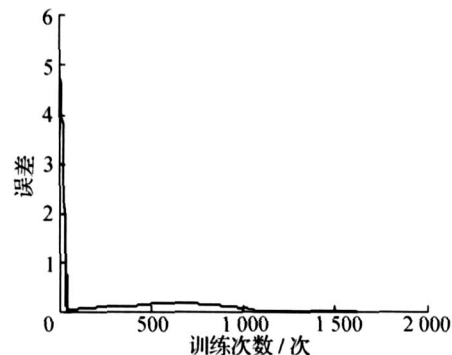


(b) 模型 B

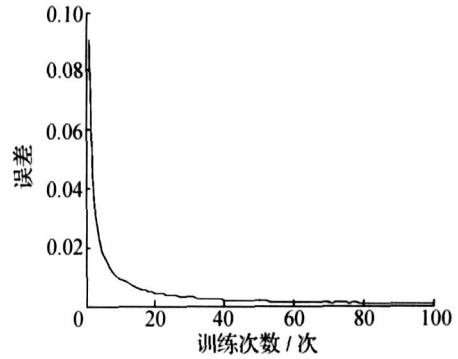


(c) 模型 C

2 采用 PSO 算法训练的仿真发射率与样本发射率的比较图
Fig 2 Comparison of the simulated emissivities and the sample ones for training the CNNE model with PSO



(a) 采用分布式最速下降法训练的发射率模型 A 的误差函数曲线



b) 采用 PSO 算法训练的发射率模型 A 的误差函数曲线
3 采用不同算法训练的发射率模型 A 的误差函数曲线
Fig 3 Comparison of the error function curve of model A for training the CNNE model with different algorithms

通过观察图 3,可以比较这 2 种算法收敛的快速性. 显而易见, PSO 算法的收敛性要明显好于分布式最速下降法的收敛性. 在实际应用中,为了实现发射率的实时测量,在保证精度的前提下,必须寻求具有非常快收敛速度的训练方法,而分布式最速下降法由于结构比较复杂,且对初值变化特别敏感,这些都延缓了收敛速度,因而不能满足实际需求. 相比较而言, PSO 算法结构比较简单,对初始值变化不甚敏感(初始的位置和速度是随机的),一般在 80 ~ 200 次以内就会达到收敛,更适合于实际应用.

造成对非线性变化样本发射率模型逼近程度不好的原因是:分析式(1)和式(2)不难发现,粒子们在搜索时,总是追逐当前全局最优点和自己迄今搜索到的个体最优点,因此粒子们的速度很快降到接近于 0,导致粒子们陷入局部极小而无法摆脱,称这种现象为粒子群的趋同性.

这种趋同性限制了粒子的搜索范围. 要想扩大搜索范围,就要增加粒子群的粒子数,或者减弱粒子对整个粒子群当前搜索到的全局最优点的追逐. 增加粒子数将导致算法计算复杂度增高,而减弱粒子对全局最优点的追逐又存在算法不易收敛的缺点.

针对以上问题,对标准 PSO 算法进行改进,采用一种带有梯度加速的 PSO 算法^[8]. 与其他进化算法一样,标准 PSO 算法利用种群进行随机搜索,没有考虑具体问题的特性,不使用梯度信息,而梯度信息往往包含目标函数的一些重要信息.

带有梯度加速的 PSO 算法主要步骤与标准 PSO 算法步骤基本相同,只需把步骤 6 稍作修正,改进措施如下:在每次粒子进行速度和位置更新时,每个粒子以概率 p 按式(1)和式(2)更新,并以概率 $1 - p$ 按梯度信息更新,在负梯度方向进行一次直线搜索来确定移动步长. 梯度信息的加入使粒子的移动更具针对性和效率. 为防止陷入局部最优,在群体最优信息陷入停滞时,对部分粒子进行重新初始化,

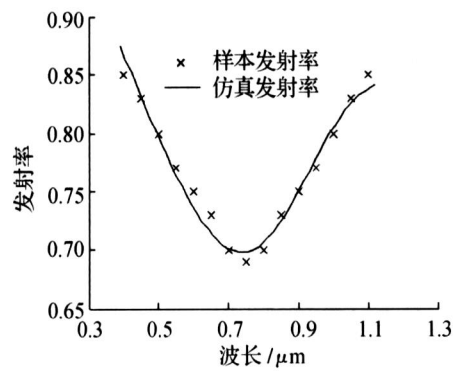
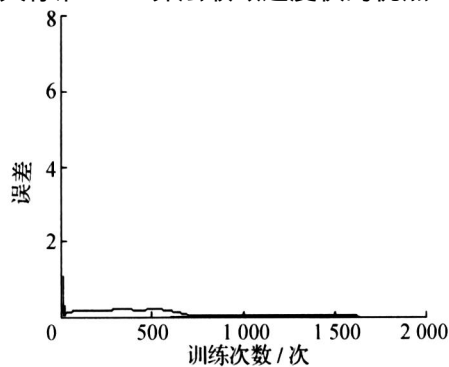


图 4 采用改进后的 PSO 算法训练的仿真发射率与样本发射率模型 C 的比较图

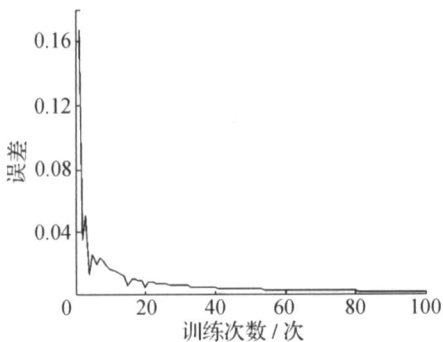
Fig. 4 Comprison of the simulated emissivities and the sample ones of model C for training the CNNE model with advanced PSO algorithm

从而保持群体的活性,减小群体陷入局优的可能性.应用改进的 PSO 算法对样本发射率模型 C 重新进行仿真,结果如图 4 所示.

通过比较图 4 和图 2 (c) 可以看出:改进后的 PSO 算法对于非线性变化的样本发射率模型的逼近程度有了大幅度的提高,而且从图 5 可以看出,采用改进后的 PSO 算法对 CNNE 模型进行训练并没有丧失标准 PSO 算法收敛速度快的优点.



(a)采用分布式最速下降法训练的发射率模型 C 的误差函数曲线



(b)采用改进后的 PSO 算法训练的发射率模型 C 的误差函数曲线

图 5 采用不同算法训练的发射率模型 C 的误差函数曲线

g. 5 Comprison of the error function curve of model C or training the CNNE model with different algorithms

4 结束语

本文针对分布式最速下降法结构复杂,算法对初值变化敏感,收敛速度较缓慢等缺点^[1],将人工智能算法——粒子群优化算法引入到神经网络模型 CNNE 中,并通过改进标准 PSO 算法,加入梯度信息,有效地增强了粒子群优化算法跳出局部最优解的能力,使算法既摆脱了后期易陷入局部最优点的束缚,又保持了其前期搜索速度快的优点,从而在保证精度的前提下,使 CNNE 模型迅速收敛,解决了发射率的在线实时测量问题.通过仿真验证,该算法较分布式最速下降法在训练速度上有了明显的提高,一般在 100 次以内就会达到收敛.

参考文献:

[1] YANG Chunling, YU Yong, ZHAO Dongyang, et al. Study on modeling of multispectral emissivity and optimization algorithm [J]. IEEE Trans on Neural Networks, 2006,17(1): 238 - 242.

[2] YANG Chunling, ZHAO Dongyang, DAI Jingming. A new method for constructing spectral emissivity models for measuring the real temperature of targets[J]. IEEE Trans on Instrumentation and Measurement, 2005,54(6): 2549 - 2553.

[3] EBERHART R C, KENNEDY J. A new optimizer using particle swarm theory[A]. Proceedings of Sixth International Symposium Micro Machine and Human Science [C]. Nagoya, Japan, 1995.

[4] EBERHART R C, KENNEDY J. Swarrn intelligence [M]. San Francisco: Mornan Kaufmann Publishers, 2001.

[5] KENNEDY J, EBERHART R C. Particle swarm optimization[A]. Proceedings of IEEE International Conference on Neutral Networks[C]. Perth, Australia, 1995.

[6] CLERE M, KENNEDY J. The particle swarm: explosion, stability and convergence in a multidimensional complex space[J]. IEEE Trans on Evolutionary Computation, 2002,6(1): 58 - 73.

[7] BERGH F V D. An analysis of particle swarm optimizer [D]. South Africa: University of Pretoria, 2002

[8] 王俊伟,汪定伟. 一种带有梯度加速的粒子群算法[J]. 控制与决策,2004,19(11): 1298 - 1230.

WANG Junwei, WANG Dingwei. Particle swarm optimization algorithm with gradient acceleration[J]. Control and Decision, 2004,19(11): 1298 - 1230.

作者简介:



杨春玲,女,1965年生,教授,博士生导师,IEEE会员;东北地区高校电子技术研究会理事;主要研究方向为神经网络建模及优化算法、目标红外辐射特性参数的测量、红外图像处理、微弱信号检测及FPGA硬件实现技术等.承担省部级科研项目6项.在国内外发表论文40余篇,被SCI检索5篇,主编教材4部.

E-mail: yangcl1@hit.edu.cn.



王睐来,男,1982年生,硕士研究生,主要研究方向为基于FPGA/CPLD的数字系统设计、人工神经网络建模等.



朱敏,男,1981年生,硕士研究生主要研究方向为红外检测.

IEEE 计算智能世界大会

IEEE World Congress on Computational Intelligence (WCCI 2008)

The 2008 IEEE World Congress on Computational Intelligence (WCCI 2008) will be held at the Hong Kong Convention and Exhibition Centre during June 1 ~ 6, 2008. WCCI 2008 is composed of the 2008 International Joint Conference on Neural Networks (IJCNN 2008), the 2008 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008) and the 2008 IEEE Congress on Evolutionary Computation (CEC 2008). The congress will feature world-renowned plenary speakers, regular technical sessions, state-of-the-art special sessions, interactive poster sessions, moderated panel discussions, informative pre-conference tutorials, themed post-conference workshops, and entertaining social functions. Researchers are invited to contribute high-quality papers to WCCI 2008. All papers are to be submitted electronically through the Congress website by December 1, 2007. All contributed papers will be reviewed by experts in the fields based on the criteria of originality, significance, quality, and clarity. For general inquiries, please contact General Chair Jun Wang at jwang@mae.cuhk.edu.hk. For program inquiries, contact IJCNN2008 Program Chair Derong Liu at dliu@ece.uic.edu, FUZZ-IEEE2008 Program Chair Gary Feng at megfeng@cityu.edu.hk, or CEC2008 Program Chair Zbigniew Michalewicz at zbyszek@cs.adelaide.edu.au. For more details, please visit the conference web site at <http://www.wcci2008.org/>.