

文章编号: 1001-0920(2009)03-0347-05

基于改进量子遗传算法的过程神经网络训练

李欣^{1a,2}, 程春田^{1b}, 曾筠^{1b}

(1. 大连理工大学 a. 电子与信息工程学院, b. 水电与水信息研究所, 辽宁 大连 116024; 2. 大庆石油学院 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要: 针对过程神经网络由于模型参数较多 BP 算法不易收敛的问题, 提出一种基于量子位 Bloch 坐标的量子遗传算法. 将该算法融合于过程神经网络的训练, 按权值参数的个数确定量子染色体上的基因数并完成种群编码, 通过新的量子旋转门完成个体的更新. 算法中的每条染色体携带 3 条基因链, 因此可扩展对解空间的遍历性, 加速优化进程. 以两组二维三角函数的模式分类问题为例, 仿真结果表明该方法不仅收敛速度快, 而且寻优能力强.

关键词: 过程神经网络; 量子遗传算法; 学习算法

中图分类号: TP183

文献标识码: A

Training of process neural networks based on improved quantum genetic algorithm

LI Xin^{1a,2}, CHENG Chun-tian^{1b}, ZENG Yun^{1b}

(1a. School of Electronics and Information Engineering, 1b. Institute of Hydropower System and Hydroinformatics, Dalian University of Technology, Dalian 116024, China; 2. School of Computer and Information Technology, Daqing Petroleum Institute, Daqing 163318, China. Correspondent: LI Xin, E-mail: lixin_dq@163.com)

Abstract: Aiming at the problem that it is difficult for BP algorithm to converge because of more parameters in training of process neural networks based on orthogonal basis expansion, a solution on the basis of an improved quantum genetic algorithm is proposed in the paper. An improved quantum genetic algorithm based on Bloch coordinates of qubits is proposed, which is integrated into the training of process neural networks. The number of genes on a chromosome is determined by the number of weight parameters and population coding is completed. Individuals in the population are updated by new quantum rotation gate. In this method, each chromosome carries three chains of genes, so can extend ergodicity for solution space and accelerate optimization process. Taking the pattern classification of two groups of two-dimensional trigonometric functions as an example, the simulation results show that the method has not only fast convergence but also good optimization ability.

Key words: Process neural networks; Quantum genetic algorithm; Learning algorithm

1 引言

针对许多实际应用系统的输入是依赖于时间的连续函数, 一些控制信号的输出既依赖于输入函数的空间聚合, 又与时间的累计效应密切相关的问题, 文献[1, 2]提出一种过程神经网络模型(PNN). 过程神经元的结构与传统 MP 模型的结构相似, 由加权、聚合和激励运算三部分组成; 与传统神经元不同之处在于: 过程神经元的输入和权值可以是时变的, 其聚合运算既有对空间的多输入聚合, 也有对时间过程的积累. 对于过程神经网络的训练问题, 文

献[3-5]提出了基于函数正交基展开的学习算法. 其基本思想是将网络输入函数和链接权函数用某种正交基函数展开, 利用基函数的正交性, 将网络的积分聚合运算简化为展开项系数间的加、乘运算. 然而, 当展开式项数较多时, 需要优化的参数数目比普通神经网络大得多, 这将使优化曲面变得极不规则, 存在许多局部极小值, 最终导致 BP 算法不易收敛.

量子遗传算法(QGA)是一种基于量子计算原理的概率优化算法, 它以量子计算的一些概念和理论为基础, 用量子位编码表示染色体, 用量子门更新

收稿日期: 2008-01-30; 修回日期: 2008-04-25.

基金项目: 国家自然科学基金项目(50479055, 50679011).

作者简介: 李欣(1979—), 男, 吉林蛟河人, 博士生, 从事神经网络、智能算法的研究; 程春田(1965—), 男, 湖北孝感人, 教授, 博士生导师, 从事神经网络、智能控制等研究.

染色体并完成进化搜索. 该算法具有种群规模小、收敛速度快、全局搜索能力强等优点^[68]. 由于使用了量子计算机, 使 QGA 具有良好的鲁棒性和全局搜索能力, 目前已在人工智能领域取得了不少成功的应用. 然而, 目前的 QGA 在编码机制、进化策略等方面还存在某些缺陷.

本文对此作了改进, 提出一种基于量子位 Bloch 坐标的量子遗传算法 (BQGA); 将改进的 BQGA 应用于 PNN 的训练, 提出一种基于 BQGA 的 PNN 训练算法. 仿真实验表明, BQGA 的优化效率明显高于 QGA, 它比 BP 算法具有更快的收敛速度和更强的寻优能力.

2 过程神经网络模型

2.1 过程神经元

过程神经元是传统神经元在时域上的扩展, 传统神经元可看成过程神经元的特例^[1,21]. 单个过程神经元的输入输出关系为

$$Y = \int_0^T \sum_{i=1}^n w_i(t) x_i(t) dt. \quad (1)$$

2.2 过程神经网络模型

过程神经网络是由若干个过程神经元按一定的拓扑结构组成的网络. 设输入层有 n 个单元, 中间层(隐层)有 m 个单元, 输出层为线性关系. 过程式输入与输出之间的关系可描述为

$$y = \sum_{j=1}^m v_j f \left(\sum_{i=1}^n \int_0^T w_{ij}(t) x_i(t) dt - j \right). \quad (2)$$

其中: $w_{ij}(t)$ 为输入层与隐层的连接权函数, v_j 为隐层过程神经元到输出层的连接权值, j 为隐层的输出阈值, $[0, T]$ 为时间采样区间.

3 量子遗传算法

3.1 QGA 存在的问题

目前的量子遗传算法主要存在以下问题:

1) 其编码机制是通过测量量子位的状态获得二进制解, 这是一个概率操作过程, 具有很大的随机性和盲目性, 因此在种群进化的同时, 个体将不可避免地产生退化.

2) 二进制编码虽然适合于某些组合优化, 但是对于连续优化, 由于需要频繁的解码操作, 无疑加大了计算量.

3) 对于量子旋转门的转角方向, 几乎都是基于查询表, 由于涉及到多路条件判断, 影响了算法的效率; 对于转角的大小, 现有方法对全部种群一视同仁, 没有考虑各染色体之间的差异.

4) 量子遗传算法中的量子态, 是在实域 Hilbert 空间中平面单位圆周上的描述, 只有一个可变量, 并不是通常的 Bloch 球面描述, 量子特性被削

弱.

综上所述, 如何对量子染色体进行编码, 如何确定量子态的描述, 如何完成量子态的更新, 是目前没有得到较好解决的问题, 也是限制量子遗传算法效率的主要问题.

3.2 BQGA 染色体的编码

在量子计算中, 最小的信息单位用量子位表示. 量子位又称量子比特, 一个量子比特的状态可表示为

$$|\phi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi} \sin(\theta/2)|1\rangle. \quad (3)$$

其中: $\cos(\theta/2)$ 和 $e^{i\varphi} \sin(\theta/2)$ 是复数, $|\cos(\theta/2)|^2$ 和 $|e^{i\varphi} \sin(\theta/2)|^2$ 分别表示量子位处于 $|0\rangle$ 或 $|1\rangle$ 的概率, 且满足归一化条件

$$|\cos(\theta/2)|^2 + |e^{i\varphi} \sin(\theta/2)|^2 = 1. \quad (4)$$

满足式(4)的一对复数 $\cos(\theta/2)$ 和 $e^{i\varphi} \sin(\theta/2)$ 称为一个量子比特相应状态的概率幅. 因此, 量子比特也可用概率幅表示为 $[\cos(\theta/2), e^{i\varphi} \sin(\theta/2)]^T$. 其中 θ 和 φ 定义了 Bloch 球面上一点 P , 如图 1 所示.

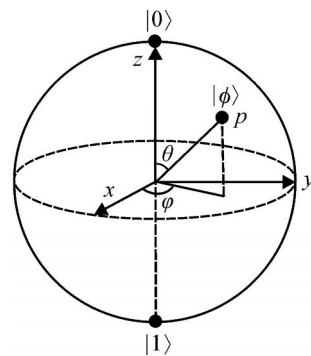


图1 量子比特的 Bloch 球面表示

由图 1 知, 任何量子位都与 Bloch 球面上一点相对应, 因此量子位可用 Bloch 坐标表示为

$$|\phi\rangle = [\cos \theta \sin \theta \sin \theta \cos \theta]^T.$$

在 BQGA 中, 直接采用量子位的 Bloch 坐标作为编码. 设 p_i 为种群中第 i 条染色体, BQGA 的编码方案如下:

$$p_i = \begin{bmatrix} \cos \theta_{i1} \sin \theta_{i1} & \dots & \cos \theta_{in} \sin \theta_{in} \\ \sin \theta_{i1} \sin \theta_{i1} & \dots & \sin \theta_{in} \sin \theta_{in} \\ \cos \theta_{i1} & \dots & \cos \theta_{in} \end{bmatrix}. \quad (5)$$

其中: $\theta_{ij} = 2 \times \text{rnd}$, $\varphi_{ij} = \text{rnd}$, rnd 为 $(0, 1)$ 之间的随机数; $i = 1, 2, \dots, m$, m 是种群规模; $j = 1, 2, \dots, n$, n 是量子位数.

在 BQEA 中, 将量子位的 3 个坐标看作 3 个并列的基因, 每条染色体包含 3 条并列的基因链, 分别称为 X 链、 Y 链和 Z 链. 每条基因链代表一个优化解, 每条染色体同时代表搜索空间中的 3 个优化解

$$p_{ix} = (\cos \theta_{i1} \sin \theta_{i1}, \dots, \cos \theta_{in} \sin \theta_{in}),$$

$$p_{iy} = (\sin \theta_{i1} \sin \theta_{i2}, \dots, \sin \theta_{in} \sin \theta_{in}),$$

$$p_{iz} = (\cos \theta_{i1}, \cos \theta_{i2}, \dots, \cos \theta_{in}).$$

分别称 p_{ix}, p_{iy} 和 p_{iz} 为 X 解、 Y 解和 Z 解。

设待优化问题的全局最优解由解空间映射到单位空间 $I^n = [-1, 1]^n$, 其形式为 $P = (x_1, x_2, \dots, x_n)$. 不失一般性, 令 $x_i \geq 0, i = \arccos(x_i), i = \theta/2, i = 1, 2, \dots, n$.

在 Bloch 球面上取点 $Q_z(\cos \theta_i \sin \theta_i, \sin \theta_i \sin \theta_i, \cos \theta_i)$, 则 Q_z 的 z 坐标等于 x_i . 过 Q_z 作垂直于 z 轴的圆周 C_z , 因 Bloch 球面上点的 z 坐标与 θ 无关, 故圆周 C_z 上所有点的 z 坐标都等于 x_i . 由 C_z 的取法知其半径为 $\sin \theta_i$. 同理, 存在半径为 $\sin \theta_i$ 垂直于 x 轴的圆周 C_x 和垂直于 y 轴的圆周 C_y , 其上所有点的 x 坐标和 y 坐标也都等于 x_i . 因此, 全局最优解中每个分量 $x_i, i = 1, 2, \dots, n$, 在 Bloch 球面上存在 3 个相同半径的圆周 C_x, C_y, C_z , 3 个圆周上的任意一点表示的量子比特都与 x_i 相对应, 如图 2 所示.

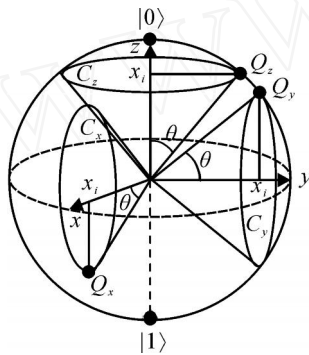


图 2 x_i 对应的量子位在 Bloch 球面上的位置

综上所述, 这种基于量子位 Bloch 坐标的编码方案具有以下优点: 1) 能避免对量子位测量生成二进制编码带来的随机性; 2) 当用于连续空间优化时, 可避免频繁的二进制数解码过程; 3) 能扩展全局最优解的数量, 提高获得全局最优解的概率.

3.3 解空间变换

群体中的每条染色体包含 n 个量子比特的 $3n$ 个 Bloch 球面坐标, 利用线性变换, 可将这 $3n$ 个坐标由 n 维单位空间 $I^n = [-1, 1]^n$ 映射到优化问题的解空间, 每个坐标对应解空间中的一个优化变量. 记第 i 条染色体 p_i 上第 j 个量子位的 Bloch 坐标为 $[x_{ij}, y_{ij}, z_{ij}]^T$, 则相应的解空间变换公式为

$$X_{ix}^j = \frac{1}{2} [b_j(1 + x_{ij}) + a_j(1 - x_{ij})], \quad (6)$$

$$X_{iy}^j = \frac{1}{2} [b_j(1 + y_{ij}) + a_j(1 - y_{ij})], \quad (7)$$

$$X_{iz}^j = \frac{1}{2} [b_j(1 + z_{ij}) + a_j(1 - z_{ij})]. \quad (8)$$

其中: $i = 1, 2, \dots, m; j = 1, 2, \dots, n$. 每条染色体对

应于优化问题的 3 个近似解.

3.4 BQGA 染色体的进化

本文提出如下量子旋转门更新 BQGA 染色体中量子位的相位:

$$U = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos(\theta/2) & \sin(\theta/2) \\ 0 & 0 & -\sin(\theta/2) & \cos(\theta/2) \end{bmatrix}. \quad (9)$$

根据

$$U \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos(\theta + \alpha) & \sin(\theta + \alpha) \\ \sin(\theta + \alpha) & \cos(\theta + \alpha) \end{bmatrix},$$

可知 U 的作用是使量子位的相位旋转 α 和 β .

关于 α 和 β 方向的确定, 文献[6]的作法是构造一个查询表, 列出各种可能情况, 作为辅助决策的工具, 但由于涉及到多路条件判断, 影响了算法的效率. 关于 α 和 β 的方向问题, 本文有如下结论:

令 $q_{0j}(x_{0j}, y_{0j}, z_{0j})$ 是当代最优染色体中第 j 量子位的 Bloch 坐标, $q_{ij}(x_{ij}, y_{ij}, z_{ij})$ 是当前种群第 i 条染色体中第 j 量子位的 Bloch 坐标. 其中: $i = 1, 2, \dots, m; j = 1, 2, \dots, n$. 记 $A = x_{0j}y_{ij} - x_{ij}y_{0j}, B = z_{0j} - z_{ij}$.

1) 方向的选取规则: 当 $A \neq 0$ 时, 方向为 $\text{sgn}(\alpha) = -\text{sgn}(A)$; 当 $A = 0$ 时, 方向取正负均可.

2) 方向的选取规则: 当 $B \neq 0$ 时, $\text{sgn}(\beta) = -\text{sgn}(B)$; 当 $B = 0$ 时, 方向取正负均可.

α 和 β 的大小通常需要根据具体问题来确定. 取值过小, 会使优化过程缓慢而降低效率; 取值过大, 会使算法越过全局最优解或陷入早熟收敛. 在 BQEA 中, 转角的大小可在区间 $(0.005, 0.05)$ 中选取.

4 基于 BQGA 的 PNN 训练算法

4.1 PNN 输入和权函数的正交基展开

在式(2)所示的 PNN 模型中, 设系统输入函数空间为 $(C[0, T])^n$, 输出为反映系统阶段累计效应的一组有序离散值, $b_1(t), b_2(t), \dots, b_n(t)$ 为 $C[0, T]$ 中一组标准正交基函数, 则输入函数和网络权函数在该组基函数下的展开式为

$$x_i(t) = \sum_{l=1}^L a_{il} b_l(t), \quad i = 1, 2, \dots, n; \quad (10)$$

$$w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t), \quad j = 1, 2, \dots, m. \quad (11)$$

其中: L 为满足展开精度要求的自然数, a_{il} 和 $w_{ij}^{(l)}$ 分

别为 $x_i(t)$ 和 $w_{ij}(t)$ 相对于 $b_l(t)$ 的展开式系数.

将式(10)和(11)代入式(2),则网络输入输出关系可表述为

$$y_k = \sum_{j=1}^m v_{jk} f \left(\sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} b_l(t) \right) \times \left(\sum_{s=1}^L a_{ks} b_s(t) \right) dt - j, \quad (12)$$

由基函数的正交性,式(12)可简化为

$$y_k = \sum_{j=1}^m v_{jk} f \left(\sum_{i=1}^n \sum_{l=1}^L a_{il} w_{ij}^{(l)} - j \right), \quad k = 1, 2, \dots, p. \quad (13)$$

给定学习样本集,通过设计适当的学习算法,即可确定出网络的各层权参数.

4.2 基于BQGA的PNN训练算法

BQGA是一种全局优化算法,而PNN训练可看作优化求解以权值为自变量的输出误差函数的全局极小值.因此,实现二者的融合可借助于量子计算机制的高效性,达到网络训练快速收敛的目的.由基函数的确定性易知, $w_{ij}(t)$ 与一组基函数展开式的系数 $w_{ij}^{(l)}$ ($l = 1, 2, \dots, L$) 存在一一对应关系,即可由一组 $w_{ij}^{(l)}$ 唯一确定,因此只需优化求解 $w_{ij}^{(l)}$ 即可.令 $w_{ij}^{(l)}$, j 和 v_{jk} 的取值区间为 $[a, b]$, 权值总数 $N = (nL + p + 1)m$. 一种优化求解 $w_{ij}^{(l)}$, j 和 v_{jk} 的BQGA结构描述如图3所示.

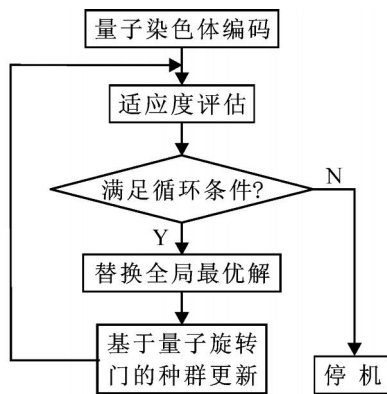


图3 基于量子位概率幅编码的量子遗传算法

具体实施方案可描述如下:

Step1: 种群初始化.按式(5)随机产生 m 条染色体组成初始群体,设定转角步长 θ , $(0.005, 0.05)$,最大进化代数 \max_gen ,限定误差 \max_E ,置当前代数 $t = 0$.

Step2: 解空间变换.将每条染色体代表的近似解,由单位空间 $[-1, 1]^N$ 映射到权值优化问题的解空间 $[a, b]^N$.

Step3: 计算适应度.每条染色体对应3组网络权值,其适应度函数可结合式(2)过程神经网络的输出误差函数来构造.考虑到适应度的实际意义,可

取误差函数的倒数^[9,10],并按适应度函数计算各染色体的适应值.记当代最优解为 \bar{X}_0 ,对应染色体 \tilde{p}_0 ,到目前为止得到的最优解为 X_0 ,对应染色体 p_0 .若 $\text{fit}(\bar{X}_0) > \text{fit}(X_0)$,则 $p_0 = \tilde{p}_0$.

Step4: 对于种群中每条染色体,以 p_0 中各量子位为目标,用式(9)定义的量子旋转门更新其量子位.

Step5: 若当前误差 $E > \max_E$,且 $t < \max_gen$,则 $t = t + 1$,返回Step2;否则,记录优化结果,停机.

5 训练实例

为验证BQGA-PNN算法的有效性,采用基于沃尔什正交基函数展开的PNN,实现如下两组二维三角函数的模式映射,并用BQGA实现PNN权参数的训练.

第1组

$$\begin{cases} (\sin(2(x - 0.1)), \cos(2(x - 0.1)), 0.5), \\ (\sin(2(x - 0.15)), \cos(2(x - 0.15)), 0.5), \\ (\sin(2(x - 0.2)), \cos(2(x - 0.2)), 0.5). \end{cases}$$

第2组

$$\begin{cases} (1.5\sin(2.2(x - 0.1)), \\ 0.5\cos(1.8(x - 0.1)), 1), \\ (1.5\sin(2.2(x - 0.15)), \\ 0.5\cos(1.8(x - 0.15)), 1), \\ (1.5\sin(2.2(x - 0.2)), \\ 0.5\cos(1.8(x - 0.2)), 1). \end{cases}$$

为体现BQGA的高效性,将本文算法与文献[3]的正交基BP算法、文献[6]的QGA算法进行对比.3种算法采用相同的网络参数,网络结构及算法参数设置如下:

输入节点 $n = 2$,隐层节点 $m = 15$,输出节点 $p = 1$,序列长度 $L = 20$,误差精度为0.05,学习速度为0.5,惯性系数为0.1,最大迭代次数为1000,种群规模为30,量子位数为630,转角初值为0.01,各权值和阈值的取值区间为 $[-10, 10]$.

文献[6]中QGA采用二进制编码,每个权值用15位二进制数表示,左起第1位为符号位,2~5位为整数部分,6~15位为小数部分.

分别用BQGA、QGA和正交基BP算法进行10次仿真,记录各自的收敛次数、迭代步数及训练误差等评价指标,对比结果如表1,图4和图5所示.

实验中基函数取20个,网络权参数高达630个,属于高维优化问题.正交基BP算法由于自身的缺陷很容易陷入局部极小,从而导致收敛次数较少;BQGA算法使用了基于量子位Bloch坐标的三链基因编码机制,因此有较强的全局寻优能力,其收敛次

表 1 3 种算法的优化性能对比

算 法	收敛次数	最少步数	平均步数	最大误差	平均误差	平均时间/s
BQGA-PNN	10	398	629.2	0.050	0.050	40.71
QGA-PNN	9	509	719.2	0.063	0.051	199.96
BP-PNN	6	673	869.9	0.089	0.059	41.83

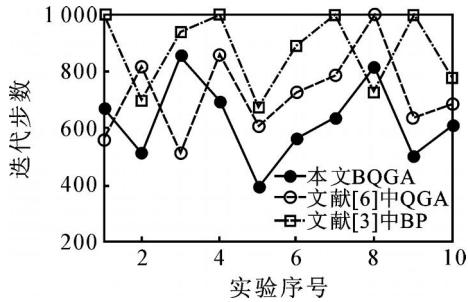


图 4 3 种算法的迭代步数对比

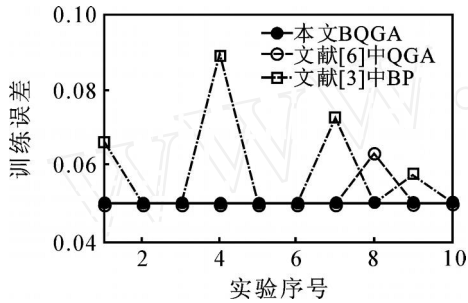


图 5 3 种算法的训练误差对比

数、平均迭代步数和平均训练误差等指标,均优于 QGA 算法和正交基 BP 算法。值得指出的是,尽管文献[6]的 QGA 有较强的寻优能力(9 次收敛),但由于使用基于量子位测量的二进制编码,需要进行频繁的解码操作,使平均训练时间显著延长,从而降低了优化效率。

6 结 论

对于输入为过程而输出为实值向量的过程神经网络的训练问题,本文提出一种基于 BQGA 的 PNN 训练方法。该方法融合了过程神经网络能处理过程式输入,便于对时变系统建模;量子遗传算法能进行全局搜索,并且收敛速度快。仿真实验表明,该算法与文献[6]的 QGA 相比,优化效率有明显的改善;与文献[3]的正交基 BP 算法相比,不仅收敛速度快,而且寻优能力强。该方法为 PNN 训练提供了一条新的途径。

参考文献(References)

[1] 何新贵,梁久祯. 过程神经网络的若干理论问题[J]. 中国工程科学, 2000, 12(2): 40-44.
(He X G, Liang J Z. Some theoretical issues on procedure neural networks [J]. Engineering Science, 2002, 12(2): 40-44.)

[2] 何新贵,梁久祯,许少华. 过程神经网络的训练与应用[J]. 中国工程科学, 2001, 3(4): 31-35.
(He X G, Liang J Z, Xu S H. Learning for process neural networks and its applications [J]. Engineering Science, 2001, 3(4): 31-35.)

[3] 许少华,何新贵. 基于函数正交基展开的过程神经网络学习算法[J]. 计算机学报, 2004, 27(5): 645-650.
(Xu S H, He X G. Learning algorithms of process neural networks based on orthogonal function basis expansion[J]. Chinese J of Computers, 2004, 27(5): 645-650.)

[4] 许少华,何新贵,李盼池. 一类用于连续过程逼近的过程神经网络及其应用[J]. 信息与控制, 2004, 33(1): 116-119.
(Xu S H, He X G, Li P C. A process neural network for continuous process approximation and its application [J]. Information and Control, 2004, 33(1): 116-119.)

[5] 许少华,何新贵,尚福华. 基于基函数展开的双隐层过程神经网络及其应用[J]. 控制与决策, 2004, 19(1): 36-40.
(Xu S H, He X G, Shang F H. Research and application of process neural network with two hidden layer based on expansion of basis function[J]. Control and Decision, 2004, 19(1): 36-40.)

[6] Han K H, Kim J H. Genetic quantum algorithm and its application to combinatorial optimization problem [C]. Proc of the 2000 Congress on Evolutionary Computation. New York: IEEE Press, 2000, 7: 1354-1360.

[7] 张葛祥,李娜,金炜东. 一种新量子遗传算法及其應用[J]. 电子学报, 2004, 32(3): 476-479.
(Zhang G X, Li N, Jin W D. A novel quantum genetic algorithm and its application [J]. Acta Electronica Sinica, 2004, 32(3): 476-479.)

[8] Yang J A, Li B, Zhuang Z Q. Multi-universe parallel quantum genetic algorithm its application to blind source separation [C]. Proc of IEEE Int Conf on Neural Networks & Signal Processing. New York: IEEE Press, 2003, 12: 393-398.

[9] 王小平,曹立明. 遗传算法——理论应用与软件实现[M]. 西安:西安交通大学出版社, 2002.
(Wang X P, Cao L M. Genetic algorithm: Theory, application and software implementation [M]. Xi 'an: Xi 'an Jiaotong University Press, 2002.)

[10] 何新贵,许少华. 过程神经网络[M]. 北京:科学出版社, 2007.
(He X G, Xu S H. Process neural networks [M]. Beijing: Science Press, 2007.)