

自适应变异粒子群算法优化 BP 神经网络在音乐分类中的应用*

彭建喜, 喻晓

(佛山职业技术学院 电子信息系, 广东 佛山 528137)

摘要: 自适应变异粒子群算法具备了基本粒子群算法和遗传算法优点, 用此算法寻找 BP 网络较好的网络权值和阈值, 使得 BP 网络的全局误差最小化, 不仅可以克服基本 BP 算法收敛速度慢和易陷入局部极值的局限, 而且模型的精度高。仿真实验结果表明, 本算法与传统的分类方法相比, 具有更高的正确率, 验证了自适应变异粒子群算法优化 BP 神经网络是一种有效的分类方法。

关键词: 音乐分类; 神经网络; 粒子群优化算法; 交叉算子

中图分类号: TH124

文献标识码: A

文章编号: 1674-7720(2012)01-0064-03

Music category based on adaptive mutation particle swarm optimization BP neural network

Peng Jianxi, Yu Xiao

(Department of Electronic Engineering, Foshan Polytechnic, Foshan 528137, China)

Abstract: Adaptive mutation particle swarm optimization algorithm has the advantages of PSO and GM algorithm. Using the algorithm to find a better network weights and threshold can not only overcome the slow convergence of basic and easy to fall into local minimum limitations, but also has a high accuracy of the model. The simulation results show that the algorithm has higher accuracy than traditional classification methods, and verified adaptive mutation particle swarm optimization is an effective classification to optimize BP neural network.

Key words: music category; neural network; particle swarm optimization(PSO); hybrid genes

音频分类识别技术是利用计算机对音频信号进行分析, 从而实现声音自动理解的一门学科。音频分类识别技术已经成为信息科学中一个十分活跃的研究领域。作为一门交叉学科, 它正逐步成为信息技术中人机交互的关键技术。音频分类识别是一个复杂的非线性过程, 本质上是一个模式识别的问题, 很多研究者在这个领域做了大量的工作, 但均存在需要改进之处。如基于规则的音频分类方法难于满足复杂的、多特征的音乐分类应用^[1]; 模式匹配法计算量大, 分类精度低^[2]; 隐马尔可夫模型(HMM)算法分类决策能力差, 需要语音识别的先验统计知识等缺陷^[3]。

人工神经网络是一种分布式并行信息处理系统, 它的自适应、自组织及自学习等特征使其特别适合于音频

识别中的分类问题, 为解决音频分类识别这样一个复杂的模式分类问题提供了新的途径。众多神经网络中, BP 神经网络因其结构简单、算法成熟并能精确寻优而被广泛地应用于音频分类识别领域。但是, 由于 BP 算法是一种基于梯度下降的算法, 因而不可避免地存在收敛速度慢、易陷入局部极小点等缺点。近年来, 人们围绕如何加速传统 BP 网络的收敛速度及尽量避免陷入局部最优解等问题做了大量的工作, 并提出了许多改进方案^[4-5]。

本文借鉴了遗传算法中的变异思想, 提出了基于遗传交叉算子的自适应变异粒子群算法, 它具备了基本粒子群算法和遗传算法优点, 使粒子能够跳出先前搜索到的局部最优解, 在更大的空间中开展搜索, 同时保持了种群多样性, 提高算法寻找到更优解的可能性。因此可

* 基金项目: 国家自然科学基金(50865003)

技术与方法

Technique and Method

以充分利用 BP 神经网络与自适应变异粒子群算法的优点,把两者结合起来,优势互补,在全局搜索的同时利用梯度法加速寻优,从而能改善整个网络的学习性能和收敛性能,最终提高整个系统的识别率。

1 BP 神经网络

误差反向传播神经网络(简称 BP 神经网络),它是一种单向传播的多层前向网络,网络中除了输入输出节点外,还有一层或多层的隐含层节点,且同层节点间没有任何耦合。输入信号从输入层节点依次传过各隐含层节点,然后传到输出层节点,每一层节点的输出只影响下一层节点的输出。

图 1 所示为一个两层的 BP 神经网络结构,该网络只有一个隐含层。

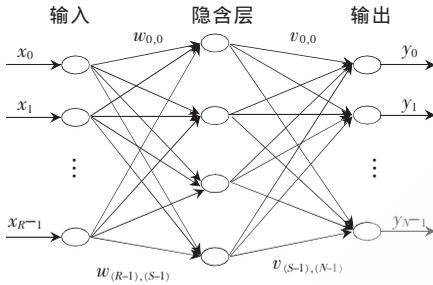


图 1 BP 神经网络的拓扑结构图

$$W_{R \times S} = \begin{bmatrix} W_{0,0} & W_{0,1} & \cdots & W_{0,S-1} \\ W_{1,0} & W_{1,1} & \cdots & W_{1,S-1} \\ \vdots & \vdots & \vdots & \vdots \\ W_{R-1,0} & W_{R-1,1} & \cdots & W_{R-1,S-1} \end{bmatrix} \quad (1)$$

$$V_{S \times N} = \begin{bmatrix} V_{0,0} & V_{0,1} & \cdots & V_{0,S-1} \\ V_{1,0} & V_{1,1} & \cdots & V_{1,S-1} \\ \vdots & \vdots & \vdots & \vdots \\ V_{R-1,0} & V_{R-1,1} & \cdots & V_{R-1,S-1} \end{bmatrix} \quad (2)$$

其中矩阵 W 是输入层与隐含层的连接权矩阵,矩阵 V 是隐含层与输出层的连接权矩阵。

假设隐含层、输出层的神经元阈值分别为:

$$\alpha = \{\alpha_0, \cdots, \alpha_{S-1}\} \quad (3)$$

$$\beta = \{\beta_0, \cdots, \beta_{S-1}\} \quad (4)$$

则该网络各层的输入输出可以表示为:

输入矢量为:

$$X = \{X_0, X_1, \cdots, X_{R-1}\} \quad (5)$$

隐含层的网络输入为:

$$\text{net}_j = \sum_{i=0}^{R-1} X_i W_{i,j} - \alpha_j \quad j=0, 1, \cdots, (S-1) \quad (6)$$

隐含层的网络输出为:

$$z_j = g(\text{net}_j) \quad j=0, 1, \cdots, (S-1) \quad (7)$$

输出层的网络输出为:

$$y_k = g\left(\sum_{j=0}^{S-1} Z_j V_{j,k} - \beta_k\right) \quad k=0, 1, \cdots, (N-1) \quad (8)$$

输出矢量:

$$Y = \{y_0, y_1, \cdots, y_{N-1}\} \quad (9)$$

对应的期望输出值为:

$$O = \{o_0, o_1, \cdots, o_{N-1}\} \quad (10)$$

2 自适应变异粒子群算法

粒子群优化算法 PSO (Particle Swarm Optimization) 是由 Kennedy 和 Eberhart 于 1995 年提出的一种基于群体智能理论的全局优化方法,通过群体中粒子在解空间追踪最优的例子进行搜索^[6]。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、非劣解质量高、鲁棒性好等优点,特别适合工程应用。但同时 PSO 也存在早熟收敛、搜索精度不高、后期迭代效率不高的缺点。

粒子的速度和位置更新公式如下:

$$v_i^{k+1} = \omega v_i^k + c_1 \times r_1 (pbest_i^k - x_i^k) + c_2 \times r_2 (gbest^k - x_i^k) \quad (11)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (12)$$

本文借鉴了遗传算法中的变异思想,提出了基于遗传交叉算子的自适应变异粒子群算法。主要方法是通过采用交叉算子产生出代表新的解集的种群,其交叉方法如下:每一次迭代中,取排序后适应度好的前半粒子直接进入下一代,后半粒子放入粒子选择池中两两配对,产生一个随机交叉位置进行遗传选择和交叉操作,生成和父代同样数目的子代后再和父代做比较,选择适应度好的一半进入下一代,以保持种群的粒子数目不变。这样通过交叉既可以增加粒子多样性,跳出局部最优,还可以加快收敛速度^[7]。

设 a, b 分别表示被选择进行遗传选择和交叉操作的两个父代个体的指针,则该操作的具体计算公式如下:

$$x_i^{k+1} = r_1 x_i^k + (1.0 - r_1) x_b^k \quad (13)$$

$$x_b^{k+1} = r_1 x_b^k + (1.0 - r_1) x_i^k \quad (14)$$

$$v_i^{k+1} = \frac{v_i^k + v_b^k}{\|v_i^k + v_b^k\|} \|v_i^k\| \quad (15)$$

$$v_b^{k+1} = \frac{v_b^k + v_i^k}{\|v_b^k + v_i^k\|} \|v_b^k\| \quad (16)$$

经过上述计算,由父代粒子形成的超立方体中随机产生了两个新的位置,其中在速率的交叉处将两个父代个体的速率之和的长度进行了规格化。因此,只有粒子的方向受到影响,数量却不会改变。

3 自适应变异粒子群算法优化 BP 神经网络

自适应变异粒子群算法优化 BP 神经网络分为 BP 神经网络结构确定、自适应变异粒子群算法优化和 BP 神经网络预测三部分。其中 BP 神经网络结构确定部分是根据应用的输入输出参数个数确定 BP 神经网络结构,进而确定粒子群算法个体的长度。自适应变异粒子群算法优化 BP 神经网络的权值和阈值,种群中的每个个体都包含了一个网络所有权值和阈值,个体通过适度函数计算个体适应度值,然后更新个体极值和群体极值来寻找到最优适应值对应个体。BP 神经网络预测用自

技术与方法

Technique and Method

自适应变异粒子群算法得到的最优个体对网络初始权值和阈值赋值,网络经训练后预测结果。

自适应变异粒子群算法优化 BP 神经网络算法流程图如图 2 所示。

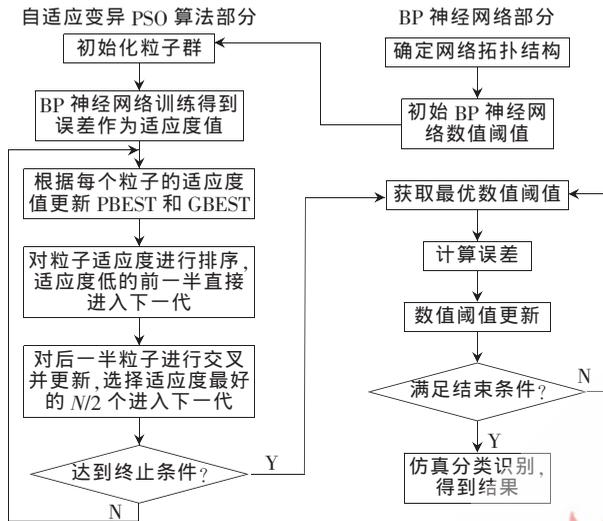


图2 算法流程图

4 实验仿真与结果讨论

4.1 网络训练与识别

本文选取了民歌、古筝、摇滚和流行四类不同音乐,每段音乐都用倒谱法提取 500 组 24 维语音特征信号,共有 2 000 组语音特征信号。由于语音特征输入信号有 24 维,待分类的语音信号共有 4 类,所以 BP 神经网络的结构为 24-25-4,即输入层有 24 个节点,隐含层有 25 个节点,输出层有 4 个节点。

从语音特征信号数据中随机选择 1 500 组数据作为训练数据用来训练网络。根据语音特征信号的特性并参考经典 PSO 参数集,本文设置算法的基本参数如下:(1)粒子规模数 $n=30$;(2)粒子维数 $D=729$;(3)最大速度 $v_{\max}=1$;(4)最大迭代次数设为 100 次;(5)终止条件,循环达到终止迭代次数或最优适应值连续迭代 50 次,计算结果差值小于 0.000 5;(6)粒子群节点适应度函数使用 BP 算法的 MSE (Mean Squared Error) 定义。

用 PSO 得出的最优解确定 BP 网络的权值和阈值,从 2 000 组语音特征信号中随机选择 1 500 组数据作为训练数据用来训练网络,500 组数据作为测试数据用来测试网络的分类能力。将这个分类号与输入自带的分类号进行比较,相等则识别正确;反之,识别错误。最后将识别正确的个数与所有待识别数作比值即可得到最终的识别率。

4.2 结果与分析

用训练好的 BP 神经网络分类语音特征信号测试数据,图 3 所示是 BP 神经网络分类误差,图 4 所示是预测结果。

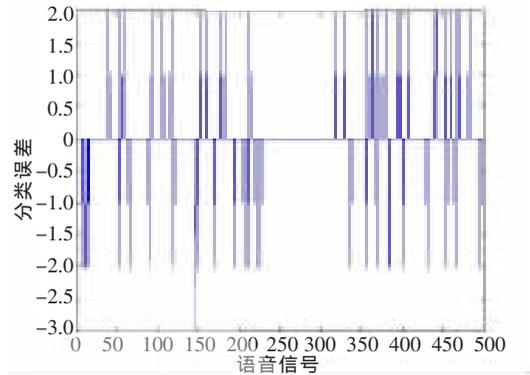


图3 BP神经网络分类误差

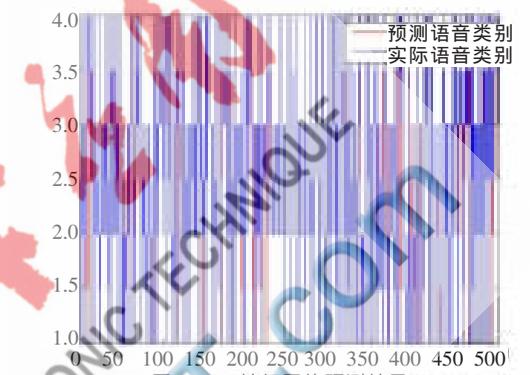


图4 BP神经网络预测结果

为了验证自适应变异粒子群算法优化 BP 神经网络基于遗传交叉算子的改进粒子群算法优化 BP 神经网络 (HPSOBPNN) 的有效性,同时用其他模型 (GABPNN、PSOBPNN、BPNN、模式匹配法、HMM) 进行了实验,结果如表 1 所示。从表 1 可知,HPSOBPNN 分类正确率比其他模型方法有了明显的提高。

表 1 各类模型分类正确率/%

语音识别类别	各类模型分类正确率/%					
	HPSOBPNN	GABPNN	PSOBPNN	BPNN	模式匹配法	HMM
民歌	92.68	88.71	90.17	80.62	77.40	76.589
古筝	100	100	100	100	96.00	95
摇滚	90.87	83.33	86.42	81.43	78.17	77.358 5
流行	94.79	93.23	93.15	92.24	88.55	87.628

音频的自动分类,尤其是语音和音乐的分类,作为提取音频内容语义和结构的重要手段之一,其研究也日益地引起关注。本文采用倒谱系数法提取音乐特征,利用自适应变异粒子群算法优化 BP 神经网络对音乐类型进行分类,与其他方法相比,其分类正确率有了明显的提高。仿真实验结果表明,该方法分类性能较好,具有一定的现实意义与参考价值。

参考文献

- [1] Tian Lan, Lu Xiaoshan, Bai Shuzhong. Speaker-independent speech recognition based on a fast algorithm [J]. Control and Decision, 2002, 17(1): 65-68.
- [2] DOWNIE T R, SILVERMAN B W. The discrete multiple

- wavelet transform and thresholding methods [J]. IEEE Trans on Signal Processing, 1998, 46(9):2558-256.
- [3] TZANETAKIS G, COOK P. Musical genre classification of audio signals [J]. IEEE Trans. on Speech and Audio Processing, 2002, 10(5):293-302.
- [4] MANIEZZO V. Genetic evolution of the topology and weight distribution of neural networks [J]. IEEE Trans on Neural Networks, 1994, 5(1):39-53.
- [5] HAGAN M T, MENHAJ M B. Training feed forward networks with the marquardt algorithm [J]. IEEE Trans on Neural Networks, 1994, 5(6):989-993.
- [6] EBERHART R C, KENNEDY J. A new optimizer using particle swarm theory [C]. Proc. of the 6th Int'l Symp. on Micro Machine and Human Science. Nagoya, Japan:[s. n.], 1995.
- [7] CLERC M. The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization [C]. Proc 1999 Congress on Evolutionary Computation Washington, DC: [s. n.], 1999:1951-1957.

(收稿日期:2011-07-14)

作者简介:

彭建喜,男,1975年生,高级工程师,主要研究方向:BP神经网络优化。

