

基于 PCA 和神经网络的故障诊断技术

汪 蔚¹, 王荣杰², 胡 清²

(1. 广东技术师范学院自动化学院, 广州 510665; 2. 广东工业大学信息工程学院, 广州 510006)

摘要: 提出一种基于 PCA 和神经网络的故障诊断/识别方法, 利用主元分析法提取故障样本集的主元, 实现故障样本的最优压缩, 简化故障诊断中神经网络分类器的结构, 提高神经网络的分类速度和测试精度。仿真结果表明, 该方法可以有效减少输入层神经元个数, 提高神经网络模型的学习效率和诊断的准确性, 具有良好的故障识别能力。

关键词: 主元分析; 神经网络; 故障诊断

Fault Diagnosis Technology Based on PCA and Neural Network

WANG Wei¹, WANG Rong-jie², HU Qing²

(1. School of Automation, Guangdong Polytechnic Normal University, Guangzhou 510665;

2. School of Information Engineering, Guangdong University of Technology, Guangzhou 510006)

【Abstract】 A new method of fault diagnosis distinguishing based on PCA-neural network is raised. It uses PCA theory to extract the main element from the fault sample data, realizes optimum compression of fault sample data, simplifies structure of neural network classifier in fault diagnosis, and enhances classification speed and precision. The results of power electronic circuit experiment show that the method can decrease the number of the network input nerve cells effectively, and enhance study efficiency and diagnosis accuracy. The way has very good fault distinguishing ability and vast prospect.

【Key words】 Primary Cell Analysis(PCA); neural network; fault diagnosis

1 概述

近年来, 电力电子变流技术得到了迅猛的发展, 经过变流技术处理的电能在整个国民经济的耗电量中所占比例越来越大。在发达国家, 约 75% 的电能经过电力电子技术变换后使用, 预计在 21 世纪将发展到 95% 以上。因此, 对电力电子电路进行故障检测与诊断/识别显得日趋重要^[1-4]。

故障诊断在本质上可归为一类模式识别问题, 即依据相关征兆对故障模式进行分类。特征提取是故障诊断的第 1 步, 不论用计算机还是由人去识别首先都要分析各种特征的有效性并选出最有代表性的特征。特征选取的好坏直接影响到识别的效果, 即分类器的设计与性能。

另一方面, 模式识别问题往往基于一定数量的样本, 即利用从样本中提取的特征将样本划分成相应的模式类别。对原始样本或特征进行降维处理, 从而得到最能反映分类本质的特征是十分必要的, 而特征提取正是模式降维处理的有力工具。另外, 从分类性能角度来看, 降维也是有益处的, 因为利用特征提取对原始特征作适当的降维处理有利于减小分类器的误识率。多元统计理论中的主元分析法可以提取样本集的主元, 从而降低样本的维数, 甚至可以实现样本的最优压缩。文献[5]介绍了混合神经网络电力电子电路故障诊断方法, 该方法非常适用于具有较强非线性的电力电子电路。文献[6]介绍了采用模式识别技术的电力电子电路故障诊断方法, 该方法通过对故障模式矢量的逻辑预处理, 简化了故障模型。本文根据多元统计理论PCA原理^[7], 提出了基于PCA特征提取方法和神经网络相结合的电力电子整流装置故障诊断识别方法, 并将其应用于实现三相桥式可控整流电路中的晶闸管断路故障诊断实验中。该方法可简化故障诊断中神经

网络分类器的结构, 提高神经网络的分类速度和测试精度, 具有很好的故障识别能力, 且适用于其他电路的故障诊断/识别。

2 基于 PCA 的故障特征提取

根据 PCA 原理, 基于 PCA 的电力电子整流装置故障信号特征提取算法如下:

(1) 设故障信号的采样点为 r 个, 将其排列成一维列向量, 可以表示成下式:

$$\mathbf{x} = (s_1, s_2, \dots, s_r)^T \quad (1)$$

用于训练的故障信号样本总数为 N , 每个故障样本均为 r 维, 将其按行抽取成对应的列向量为 x_1, x_2, \dots, x_N 。

(2) 计算全体训练样本的平均值向量

$$\mathbf{m} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (2)$$

(3) 计算协方差矩阵

$$\begin{aligned} C_x &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T = \frac{1}{\sqrt{N}} \frac{1}{\sqrt{N}} \sum_{i=1}^N (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T \\ &= \mathbf{X}\mathbf{X}^T \end{aligned} \quad (3)$$

其中,

$$\mathbf{X} = \frac{1}{\sqrt{N}} [(\mathbf{x}_1 - \mathbf{m}), (\mathbf{x}_2 - \mathbf{m}), \dots, (\mathbf{x}_N - \mathbf{m})]$$

(4) 求构造矩阵 $C_x = \mathbf{X}\mathbf{X}^T$ 的特征值 λ_i 及相应的正交归一化特征矢量 \mathbf{u}_i :

1) 一般情况, 原始故障信号的维数远比故障训练样本的

作者简介: 汪 蔚(1962 -), 女, 讲师、硕士, 主研方向: 智能控制; 王荣杰, 硕士研究生; 胡 清, 副教授、博士研究生

收稿日期: 2007-04-10 **E-mail:** gduthq@163.com

大, 即 $r \gg N$ 时, 为了减小计算量, 利用 SVD 解方法求解 $C_x = X^T X$ 的 λ_i 和 v_i ;

2) 当 $r < N$, 可直接求解 $C_x = X X^T$ 的特征值和特征向量。

(5) 根据前面选取的特征值所对应的正交归一化特征矢量; 当 $r \gg N$ 时,

$$u_i = \frac{1}{\sqrt{\lambda_i}} X v_i \quad i=1, 2, \dots, M \quad (4)$$

则有

$$A = (u_1, u_2, \dots, u_M) \quad (5)$$

(6) 将每种故障信号训练样本的矢量向步骤(5)中的特征矢量 A 张成的子空间投影, 即

$$y_i = A^T x_i \quad i=1, 2, \dots, N \quad (6)$$

这样就可提取故障信号的 M 维特征矢量。

3 PCA-神经网络的故障诊断

神经网络是一个天然的分器, 具有作为分类器众多的优点, 根据神经网络的原理和 PCA 的故障特征提取算法提出了基于 PCA-神经网络的电力电子整流装置故障诊断识别方法, 此方法充分结合两者的优点。基于 PCA-神经网络故障诊断识别方法具体的实现步骤如下:

(1) 故障原始特征样本集的形成。

(2) 根据前面的故障特征提取算法计算故障特征矢量, 把每种故障样本向前 M 大的特征值所对应的正交归一化特征矢量张成的子空间投影, 提取每种故障信号 M 维特征矢量。

(3) 根据步骤(2)构造用于诊断电力电子整流装置故障的 BP 神经网络, 并进行神经网络的训练。

(4) 故障训练样本或测试样本经 PCA 提取 M 维特征向量, 作为已训练神经网络的输入, 神经网络输出为对应故障类型, 从而实现故障诊断和故障元的定位。

4 实验验证

故障诊断识别实例采用三相可控整流电力电子电路, 如图 1 所示。

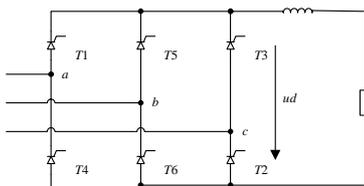


图 1 三相可控整流电路

本文将通过 $a=30^\circ$ 和 $a=0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$ 三相桥式可控整流电路中晶闸管断路故障诊断实验验证 PCA 的故障特征提取方法和 PCA-神经网络的电力电子整流装置故障诊断识别方法。以 $a=30^\circ$ 时三相桥式可控整流电路为例, 根据 PCA 特征提取的算法对训练样本进行特征提取, 求解 $C_x = X X^T$ 的 22 个特征值如下:

$$\lambda = [1.1917e+06 \quad 1.0925e+06 \quad 7.1296e+05 \quad 2.9087e+05 \quad 2.1448e+05 \quad 1.2921e+05 \quad 40802 \quad 22752 \quad 11806 \quad 7866.5 \quad 5297.8 \quad 1591.2 \quad 888.42 \quad 1.0235e-04 \quad 8.929e-05 \quad 4.3098e-05 \quad 1.6232e-05 \quad 1.0334e-05 \quad 7.5528e-06 \quad 4.9884e-06 \quad 3.4876e-06 \quad 3.7362e-026]$$

从矩阵 C_x 中选取与这 4 个特征值相对应的 4 个标准特征向量构成主成分计算矩阵。表 1 为无故障和 $T_1, T_3, T_1 T_4, T_2 T_4, T_5 T_6$ 发生断路故障时的训练样本经 PCA 方法提取 4 维特征向量组成的新训练样本。通过 PCA 方法提取的特征向量

如表 1 所示。

表 1 部分故障特征向量

故障类型	经 PCA 提取的故障特征向量			
无故障	-55.329 3	0.822 6	-391.548 2	677.419 9
T_1	-321.957 4	157.638 0	-135.737 8	508.286 0
T_3	-37.318 7	-331.639 3	-409.006 4	363.252 2
$T_1 T_4$	11.534 5	37.409 4	91.214 9	325.957 9
$T_2 T_4$	-175.714 1	-258.336 2	-198.207 4	206.600 7
$T_5 T_6$	225.764 6	509.582 1	-294.288 7	150.315 0

新训练样本作为神经网络的输入, 构造 $n_i = 4, n_o = 6, n_H = 9$ 的网络, 神经网络的训练误差曲线如图 2 所示。

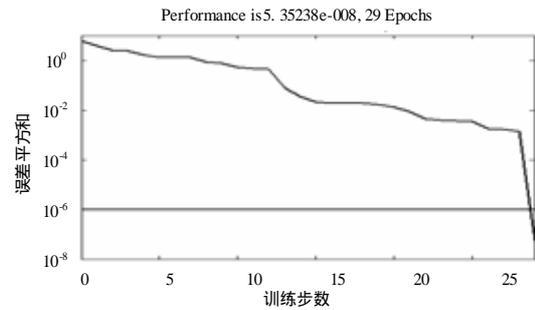


图 2 训练误差曲线

同样利用主元分析法提取三相桥式可控整流电路 $a=0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$ 时晶闸管断路故障时训练样本的特征向量, 提取 6 维故障特征向量为 BP 神经网络的新训练样本, 构造 $n_i = 6, n_o = 9, n_H = 11$ 的神经网络分类器。图 3 为神经网络训练误差曲线, 图 4 为部分训练样本和测试样本的实验诊断结果。训练样本和测试样本的诊断精度和在各个不同 α 精度分布的情况如表 2 所示。

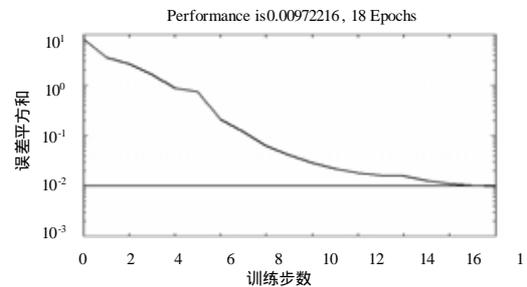


图 3 神经网络训练误差曲线

$a/(^\circ)$	诊断样本	故障元	神经网络实际输出				诊断结果
			C_9		C_8		
			C_6	C_5	C_4	C_3	
30	P_{23}	无故障	0.00	0.00	-0.01	0.01	无故障
60	P_{55}	$T_1 T_3$	-0.01	0.01	0.02	1.02	$T_1 T_3$
90	P_{70}	T_3	-0.03	1.00	0.42	0.95	T_3
120	P_{110}	$T_5 T_6$	0.79	1.03	-0.06	-0.14	$T_5 T_6$
0	P_{t22}	$T_5 T_6$	0.71	1.03	0.19	-0.07	$T_5 T_6$
30	P_{t150}	$T_1 T_3$	0.38	0.61	0.13	0.41	误诊
60	P_{t165}	$T_1 T_3$	0.10	0.16	-0.00	0.93	$T_1 T_3$
90	P_{t188}	$T_3 T_5$	-0.01	0.98	0.04	1.04	$T_3 T_5$
			0.55	0.96	-0.07	-0.07	

图 4 部分故障样本诊断结果

(下转第 188 页)