

文章编号: 1006-4710(2012)02-0204-06

# 基于非线性主元分析的水轮机调节系统 传感器故障诊断

刘明华, 南海鹏, 余向阳

(西安理工大学 水利水电学院, 陕西 西安 710048)

**摘要:** 利用现场的运行数据, 将基于输入训练神经网络的非线性主元分析(PCA)方法应用到水轮机调节系统传感器故障诊断中, 讨论了基于输入训练神经网络的非线性主元分析实现方法, 建立了输入训练神经网络和反向传播网络, 实现了对实测数据的重构, 讨论了利用平方预测误差(SPE)进行故障检测和识别的方法, 并用现场实测数据对该方法进行了仿真。仿真结果表明, 该方法有效且实用。

**关键词:** 传感器; 故障诊断; 非线性主元分析; 输入训练神经网络; 水轮机调节系统

**中图分类号:** TM312, TP277      **文献标志码:** A

## Fault Diagnosis for Sensors of Hydro Turbine Regulation System Based on Nonlinear Component Analysis

LIU Minghua, NAN Haipeng, YU Xiangyang

(Faculty of Water Resources and Hydroelectric Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** A nonlinear principal component analysis methodology based on input-training neural network is proposed and applied to sensors diagnosis of hydro turbine regulating process, which is completed by establishing an input-training neural network and a backpropagation network to reconstruct sensors value. The scheme of fault detection and fault identification is discussed via the application of the squared prediction error (SPE). Simulating results prove that this method is practically feasible with high fault recognizing rate and application value.

**Key words:** sensor; fault diagnosis; PCA; input-training neural network; regulation of hydro turbine

将主元分析(PCA)应用到故障诊断, 是基于多元统计分析的方法, 根据系统运行过程的实测参量, 经过线性变换和降维得到主元参量, 再将主元参量经过逆变换得到实测参量的估计值, 根据实测参量与其估计值之间的残差是否超出某一阈值来进行故障诊断。这种方法利用系统运行过程积累的历史数据和实时数据进行故障诊断, 适合于系统模型较难建立的系统。笔者将这一方法应用到水轮机调节系统, 利用调节系统运行过程中各传感器统计值判断各传感器值之间的相关性是否遭到破坏, 从而进一步判断是否有故障发生。

传统的主元分析方法是基于线性变换的方法,

而水轮机调节系统各传感器参数之间的关系既有线性的(如静特性曲线), 也有非线性的(如中间接力器位移与机组频率的关系)。用线性的方法来表征非线性相关性, 不能反映各传感器之间的本质关系, 会忽略非线性信息, 从而产生误判, 而采用非线性PCA的故障诊断方法具有更强的实用性。

非线性主元分析故障诊断的方法主要有两类。一类是主元曲线法, 是Hastie和Stuetzle于1989年提出的<sup>[1]</sup>, 该方法通过将数据投影到主元曲线或主元曲面上, 得到非线性主元, 有效地提取了数据中的非线性信息, 但是它不能对新样本进行数据重构, 因此不能直接应用到故障诊断中。

**收稿日期:** 2012-03-01

**基金项目:** 教育部博士点专项基金资助项目(20096118110012)。

**作者简介:** 刘明华(1976-), 男, 陕西咸阳人, 博士生, 研究方向为水轮机调节系统的故障诊断。E-mail: l\_mh11@126.com。南海鹏(1963-), 男, 陕西乾县人, 教授, 博导, 研究方向为水轮发电机组在线监测和故障诊断。E-mail: hxnhp@163.com。

另一类是基于自组织神经网络的非线性 PCA 方法,是 Kramer 于 1991 年提出的<sup>[2]</sup>,网络分为输入层、映射层、主元层、反映射层和输出层共 5 层,输入层到主元层的神经网络构造数据的非线性得分向量,主元层到输出层的神经网络则形成数据的重构向量,该方法存在神经网络结构复杂、难以训练且易产生训练过适度等缺点。Dong D 在 1993 年提出基于神经网络和主元曲线的方法,构造两个神经网络,分别实现实测数据到主元得分以及主元得分到重构数据的映射,并通过构造主元曲线获得训练样本,但主元曲线构造较困难,且适用的范围受到限制<sup>[3]</sup>。Tan S 和 Mavrouniotis 于 1995 年提出了一种输入训练神经网络<sup>[4]</sup>,输入对应主元,输出对应实测数据,通过后向传递算法不断调整输入和权值直到收敛,训练结束后同时得到主元得分向量和重构数据的映射,并证明了在单输入情况下与主元曲线的方法是等价的。

本文将输入训练神经网络非线性 PCA 方法应用于水轮机调节系统的故障诊断,选择了实测参量,确定了神经网络的结构,提出了故障检测和识别的方法,并分析了不同工况下故障诊断的方法,运用现场运行数据对该方法进行了仿真,仿真结果表明该方法有效。

## 1 输入训练神经网络的非线性 PCA 故障诊断方法

### 1.1 非线性 PCA 故障诊断原理

PCA 是多元统计分析的一种降维方法,通过线性组合和正交变换,将多元数据投影到维数较低的主元空间,降低数据的相关性和冗余,以便对数据进行分析。

设观测数据样本集  $\mathbf{X} = [X^1, X^2, \dots, X^m] \in R^{n \times m}$ ,其中  $n$  为样本数, $m$  为测量变量个数,如果测量变量间存在冗余,理论上可以用  $l(l < m)$  个变量表示数据所包含的信息,PCA 的方法是将  $\mathbf{X}$  作如下分解:

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} \quad (1)$$

$$\hat{\mathbf{X}} = \mathbf{TP}^T = \sum_{i=1}^l t_i \mathbf{p}_i^T \quad (2)$$

$$\mathbf{E} = \mathbf{T}_e \mathbf{P}_e^T = \sum_{i=l+1}^m t_i \mathbf{p}_i^T \quad (3)$$

式中, $\hat{\mathbf{X}}$  为  $\mathbf{X}$  的重构值,表示了样本集  $\mathbf{X}$  的主要信息, $\mathbf{E}$  为残差矩阵,即建模误差, $\mathbf{E}$  被认为是由于观测和干扰引起的。 $l(l < m)$  为主元个数, $\mathbf{T}$  和  $\mathbf{P}$  分别

表示主元得分向量和负载矩阵, $\mathbf{X}$  阵的分解方式确保  $[\mathbf{P}, \mathbf{P}_e]$  正交、 $[\mathbf{T}, \mathbf{T}_e]$  正交,主元投影将原来的  $m$  个变量减少到  $l$  个,实现了数据的降维。

非线性 PCA 是线性 PCA 的扩展,但  $\hat{\mathbf{X}}$  与  $\mathbf{T}$  不再是线性变换关系,而是非线性的映射关系:

$$\hat{\mathbf{X}} = \mathbf{F}(\mathbf{T}_N) \quad (4)$$

映射函数  $\mathbf{F}(\cdot)$  相当于线性 PCA 的矩阵  $\mathbf{P}$ ,函数  $\mathbf{F}(\cdot)$  的确定是非线性 PCA 要解决的主要问题。

另一方面,非线性 PCA 故障诊断还要构造从  $\mathbf{X}$  到  $\mathbf{T}_N$  的映射关系,此时:

$$\mathbf{T}_N = \mathbf{G}(\mathbf{X}) \quad (5)$$

$$\hat{\mathbf{X}} = \mathbf{F}(\mathbf{T}_N) = \mathbf{F}(\mathbf{G}(\mathbf{X})) \quad (6)$$

可见在建立以上两个映射后,可以得到任一实测值  $\mathbf{X}(k)$  的重构值  $\hat{\mathbf{X}}(k)$ ,实测值  $\mathbf{X}(k)$  与其重构值  $\hat{\mathbf{X}}(k)$  形成残差  $\mathbf{e}(k)$ ,故障诊断的原理就是在某一给定的阈值  $\alpha$  下对残差  $\mathbf{e}(k)$  的统计假设检验。

非线性 PCA 故障诊断中,核心问题就是要解决对实测值的重构,一般多采用神经网络的方法<sup>[2-4]</sup>,其中自组织神经网络的方法,是构造一个五层网络,直接建立  $\mathbf{X}$  与  $\hat{\mathbf{X}}$  的关系;而输入训练神经网络的方法则只建立从主元得分向量  $\mathbf{T}_N$  到重构值  $\hat{\mathbf{X}}$  的映射  $\mathbf{F}(\cdot)$ ,实测值  $\mathbf{X}$  的主元得分向量  $\mathbf{T}_N$  则由输入训练神经网络的训练过程获得。

### 1.2 输入训练神经网络的非线性 PCA 的故障诊断方法

为了简化自组织神经网络的结构,输入训练神经网络只建立得分向量  $\mathbf{T}_N$  与重构值  $\hat{\mathbf{X}}$  的映射函数  $\mathbf{F}(\cdot)$ ,网络的输入为得分向量  $\mathbf{T}_N$ ,输出为重构值  $\hat{\mathbf{X}}$ 。网络采用含一个隐含层的三层结构,只要隐含层节点数足够多,它可以拟合任意的非线性函数,隐含层的传输函数采用 tangent sigmoid 函数,输入输出层采用线性传输函数,网络的结构如图 1 所示。网络训练时,输出节点值采用实测值,输入节点值未知,可赋一初值,训练过程中输入向量和网络权值同时调整,以使每一组输入向量所得到网络的输出  $\hat{\mathbf{X}}$  与实测值  $\mathbf{X}$  的差值最小,训练结束后不仅得到网络参数即映射函数  $\mathbf{F}(\cdot)$ ,同时可得到每组实测值所对应的输入值,即对应的主元得分向量。对不同的输入初值,将得到不同的网络参数及输入向量,所以要进行多次训练以使网络获得尽可能精确的输出。网络采用反向传播最速下降算法,学习算法如下<sup>[5]</sup>。

$$\Delta t_i = \sum_{j=1}^d W_{1ji} d_j \quad (7)$$

$$\Delta W_{1ji} = t_i \delta_j \quad (8)$$

$$\Delta W_{2kj} = \eta \cdot f \left( \sum_{i=1}^l W_{1ji} t_i + b_{1j} \right) (X_k - \hat{X}_k) \quad (9)$$

$$\delta_j = \eta \cdot f' \left( \sum_{i=1}^l W_{1ji} t_i + b_{1j} \right) \cdot \sum_{k=1}^m W_{2kj} (X_k - \hat{X}_k) \quad (10)$$

式中,  $l, d, m$  分别是输入层、隐含层、输出层的节点数目,  $W_1, W_2$  是隐含层及输出层的权值向量,  $b_1$  是隐含层的阈值,  $f$  为传输函数,  $X_k$  是第  $k$  个实测值,  $\hat{X}_k$  是其相应的重构值。

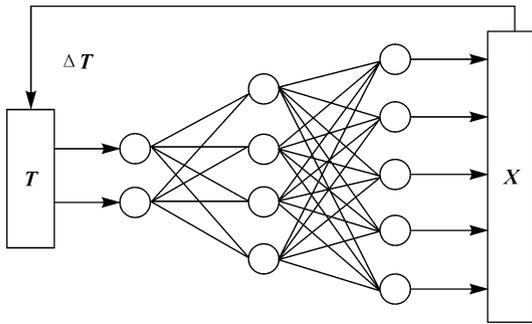


图1 输入训练神经网络的结构  
Fig. 1 Input-training neural network

实际训练输入神经网络时,将线性 PCA 和非线性 PCA 结合起来,输入神经网络的输出层不直接采用实测数据,而是将实测数据进行线性 PCA 变换,将数据用一组新的正交变量重新定义,去除实测数据的线性相关性。这样做的理由是避免实测数据中线性相关的结构增加输入训练神经网络的非线性主元个数,从而使训练的网络产生“过适度”。这里,对观测数据进行线性 PCA 变换不同于典型的线性 PCA 变换,典型的线性 PCA 变换只保留代表实测数据显著的主元方向,这里则要增加主元个数,应尽可能多地保留数据信息,防止非线性信息的遗失。区别非线性信息与噪声没有明显的界限,一般认为主元的贡献率超过 97% 以上将代表大多数的非线性信息<sup>[5]</sup>。输入训练神经网络建立的非线性主元与实测数据重构值的映射关系为:

$$X = F'(T_N) \cdot P^T + E = \hat{X} + E \quad (11)$$

式中,  $F'(\cdot)$  为输入网络建立的非线性映射,  $T_N$  为非线性主元得分,  $P$  为进行线性 PCA 变化得到的主元负载矩阵。

应用网络时,对一组新的实测数据,网络的输入是未知的,需在网络参数不变的情况下,应用非线性优化的算法计算相应的得分向量  $T_N$ ,这一过程将产

生时间消耗,不适于进行在线的故障诊断。对  $T_N$  的求取,可利用另一个反向传播神经网络,建立实测值  $X$  和  $T_N$  之间的映射关系,网络训练时,输出值采用输入神经网络训练过程所得的  $T_N$ ,网络的输入不直接采用实测值,而采用实测值经线性 PCA 变换后得到的线性得分向量  $T$ ,这样可简化过程,无需再处理相互结构关系,映射关系为:

$$T = G(X \cdot P) \quad (12)$$

式中,  $G$  为反向传播神经网络所建立的线性主元得分与非线性主元得分的映射关系。

经过以上过程,对一个给定的新的实测值  $X$ ,可以得到它的重构值  $\hat{X}$ ,  $X$  与  $\hat{X}$  形成残差向量  $e$ ,通过对  $e$  的统计分析可以对故障进行检测和识别。

### 1.3 故障检测与识别

多变量分析的数据检验方法多是基于 Hotelling  $T^2$  统计方法,  $T^2$  统计方法针对的是一系列独立的符合正态分布的随机变量,而 PCA 则被应用到相关变量的分析,因而对故障的检测采用 Q 统计方法,主要采用平方预测误差 SPE (Squared Prediction Error) 来检测传感器的相关性<sup>[6]</sup>。对一个新的实测值  $X_{new}$ ,在获得它的重构值  $\hat{X}_{new}$  后,构造其 SPE 值。

$$SPE = e^T e = (X_{new} - \hat{X}_{new})(X_{new} - \hat{X}_{new})^T \quad (13)$$

$$SPE < Q_\alpha \quad (14)$$

$Q_\alpha$  为阈值,当 SPE 大于  $Q_\alpha$  时,可判断有故障发生。

Q 统计的方法可以对故障进行检测,但却不能对故障进行识别,故障识别有基于贡献图的方法和基于向量重构的方法。贡献图的方法根据各实测参量对 SPE 的贡献率来对故障进行识别<sup>[7]</sup>,此方法在有故障传感器的情况下可能导致对其它传感器不正确的估计,易引起误诊。Dunia 等提出基于故障子空间的向量重构方法<sup>[8]</sup>,该方法的思想是假设任何一个传感器都是可能的故障源,用基于 PCA 的模型重构假定有故障的传感器信号,其它传感器信号仍保留为原输入变量,通过检验重构前后的 SPE 值来确定故障传感器。

在传感器系统发生故障之后,基于非线性 PCA 的重构模型,应用  $k$  时刻以前的测量数据,可以得到各传感器  $k$  时刻的重构值:

$$\hat{X}(k) = [\hat{X}_1(k), \hat{X}_2(k), \dots, \hat{X}_m(k)]$$

在  $k$  时刻传感器的实际测量值中第  $j$  个传感器在  $k$  时刻的实测数据  $X_j(k)$  用 PCA 重构信号  $\hat{X}_j(k)$  代替,重新定义实际测量信号为:

$$\mathbf{X}(k) = [X_1(k), X_2(k), \dots, \hat{X}_j(k), \dots, X_m(k)]$$

得到:

$$SPE'_j(k) = \sum_{i=1}^m e_i^2(k) \quad (15)$$

式中,  $SPE'_j(k)$  表示重构了第  $j$  个变量之后的  $SPE(k)$  值, 因此, 一旦有故障的传感器测量值被重构,  $SPE'_j(k)$  值会由于故障被重构排除, 而处于故障阈值以内。相反, 若重构的不是故障变量,  $SPE'_j(k)$  值则仍会受故障影响, 而处于故障阈值之外。总之, 当故障被检测出来, 分别计算重构后各传感器的  $SPE'_j(k)$  值, 若满足  $SPE'_j(k) < Q_\alpha$  则可以确定第  $j$  个传感器是有故障传感器, 从而实现传感器故障识别<sup>[9]</sup>。

## 2 水轮机调节系统故障诊断

### 2.1 水轮机调节系统故障诊断的实测参量的选择

水轮机调节系统包括调速器、引水管道、水轮机、发电机和负载, 图 2 为典型的水轮机调节系统框图, 图中  $x$  为机组转速,  $m_g$  为负荷扰动力矩,  $m$  为水轮机输出转矩,  $c$  为给定转速。

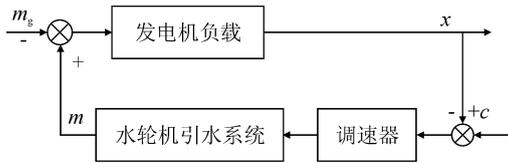


图 2 水轮机调节系统框图

Fig. 2 Hydro turbine regulation system

调节对象总的开环传递函数为:

$$\frac{X(S)}{Y(S)} = \frac{e_y + (e_y e_{qh} - e_h e_{qy}) T_w S}{e_{qh} T'_a T_w S^2 + [(e_{qh} e_n + e_n e_{qx}) T_w + T'_a] S + e_n}$$

式中,  $X$  为机组频率相对值,  $Y$  为导叶接力器相对开度,  $e_y$ 、 $e_h$ 、 $e_n$  分别为水轮机转矩  $m$  对接力器行程、水头、转速的传递系数,  $e_{qy}$ 、 $e_{qh}$ 、 $e_{qx}$  分别为水轮机流量  $q$  对接力器行程、水头、转速的传递系数。

水轮机的转矩可以用机组的有功功率  $w$  反映。可见在系统运行正常时, 水头  $H$ 、流量  $q$ 、有功功率  $w$ 、转速  $x$ 、接力器行程  $y$ 、中间接力器位移  $h$  呈一定的相关关系, 除  $x$  与  $y$  呈近似线性关系外, 其它参量之间呈明显的非线性关系。系统中, 这些参量都由相应的传感器进行检测, 在系统发生故障时, 这些传感器的相关性会遭到破坏, 可以据此对各传感器进行故障诊断。

采用输入训练神经网络的非线性 PCA 方法对水轮机调节系统传感器进行故障诊断, 输入神经

网络的输出对应实测向量, 网络的输入为主元得分向量, 即用输入神经网络构造各实测参量的相关关系, 实测参量的选择至关重要, 实测参量的数量应足够多, 以表征系统的运行状态, 同时各实测参量间应存在相关关系。这里选择上述 6 个传感器值构成实测向量  $\mathbf{X} = [y, f, h, w, q, H]$ , 利用它们的历史数据训练输入神经网络和反向传播网络。对一个实时的观测向量  $\mathbf{X}(k)$ , 利用传感器之间的相关关系对其进行重构, 可得到重构值向量  $\hat{\mathbf{X}}(k)$  和残差  $\mathbf{e}(k)$ , 通过分析  $SPE$  值对传感器的故障进行检测和识别。

在故障诊断中, 为了消除量纲的影响, 需首先对实测向量进行标准化的处理<sup>[6]</sup>。

$$X_i^* = \frac{X_i - E(X)}{\sqrt{\text{Var}(X_i)}} \quad i = 1, 2, \dots, n \quad (16)$$

水轮机调节系统是一动态过程, 在线诊断时, 用  $k-1$  时刻的实测值  $\mathbf{X}(k-1)$  进行重构得预报值  $\hat{\mathbf{X}}(k)$ , 利用  $\hat{\mathbf{X}}(k)$  与  $k$  时刻的实测值  $\mathbf{X}(k)$  形成的残差向量  $\mathbf{e}(k)$  来对故障进行判别。

### 2.2 多工况的故障诊断方法

水轮机调节系统是一个变工况的动态过程, 如所带负荷的调整, 在此过渡过程中, 各测量值都有较大的变化, 故障诊断有可能产生误判, 特别是在甩负荷的大波动过程。另一方面, 调节对象在不同的工况点, 各传递系数及时间常数都有较大变化, 如水头  $H$ 、负载的变化对  $e_h$ 、 $e_{qh}$ 、 $e_n$ 、 $T_w$ 、 $T'_a$  的影响, 这些实际上改变了各传感器之间的相关关系, 之前建立的输入训练神经网络 PCA 映射关系已不能用于故障诊断。

在实际的故障诊断中, 需在稳定的工况下进行故障检测和识别, 为此系统要对实测数据进行分类, 区分过渡过程和稳定状态, 同时要在不同的负荷状态、水头等工况下训练多个输入神经网络, 即建立多个 PCA 数据重构网络。对于一组实测值  $\mathbf{X}$ , 首先判断系统是否处于过渡过程, 只有系统达到稳定调节状态后, 才可用观测值进行故障判别, 其次, 再根据负荷、水头等参数确定用哪一组输入 PCA 神经网络进行故障诊断<sup>[10]</sup>。

### 2.3 系统仿真

根据现场实测的数据进行仿真, 在机组额定参数附近, 稳定小波工况下共获取 300 组数据, 采样周期为 60 s, 前 150 组数据用于对网络进行训练, 后 150 组数据用于对故障检测和识别, 部分观测数据见表 1。对数据进行标准化处理, 然后进行线性 PCA

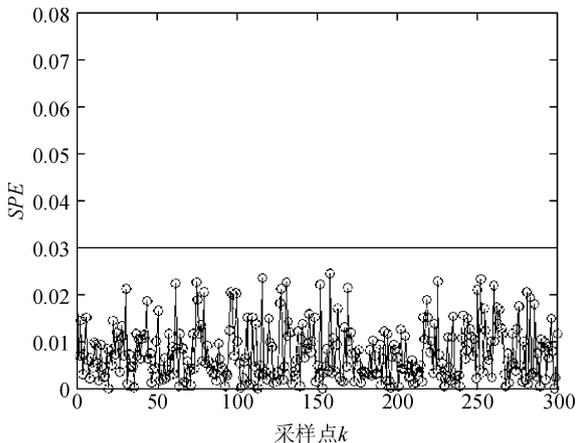
变换,得到线性主元负载矩阵  $\mathbf{P}$  和主元得分向量  $\mathbf{T}$ 。

表1 水轮机调节系统部分运行数据

Tab.1 Partail data of hydro turbine regulation system

样本	$y/\text{mm}$	$f/\text{Hz}$	$h$	$w/\text{MW}$	$q/(\text{m}^3/\text{s})$	$H/\text{m}$
1	231.5	50.01	-55	47.68	26.05	184.72
2	230.6	50.04	-30	47.59	26.08	184.72
3	230.4	49.95	-66	48.04	26.02	184.72
4	229.8	49.98	-70	48.02	25.98	184.72
5	230.2	50.02	+15	47.97	25.96	184.72
⋮	⋮	⋮	⋮	⋮	⋮	⋮
296	231.0	50.04	-50	47.96	25.99	185.03
297	231.7	50.01	-35	48.03	25.94	185.03
298	230.8	49.97	-25	48.10	26.00	185.03
299	230.3	49.95	+63	48.06	26.01	185.03
300	229.8	49.98	-58	48.03	26.07	185.03

用标准化的数据训练输入神经网络,网络的输出为6个实测量  $y, f, h, w, q, H$ , 输入为主元得分向

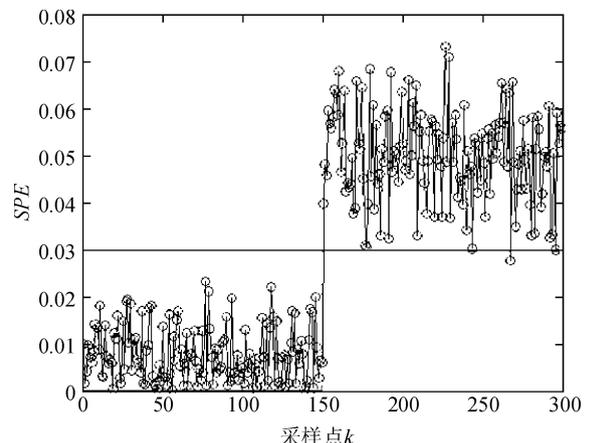


(a) 无故障发生

量  $\mathbf{T}_N$ 。根据经验,主元个数取4,隐层数为1,隐层节点数为11,网络的结构如图1所示。训练算法采用反向传播最速下降算法,为提高训练速度,引入动量因子  $\eta$ ,根据经验,输入层初值可取线性 PCA 变换得到的  $\mathbf{T}^{[10]}$ 。训练开始,输入层和网络权值、阈值同时调整,训练结束后得到非线性主元得分向量  $\mathbf{T}_N$  和输入神经网络的权值。

训练第二个反向传播神经网络,采用 BP 网络,建立实测数据  $\mathbf{X}$  到非线性主元得分向量  $\mathbf{T}_N$  的映射  $\mathbf{G}(\cdot)$ ,输入层为实测数据经线性 PCA 变换得到的主元得分向量  $\mathbf{T}$ ,输出层为输入神经网络所得的非线性主元得分  $\mathbf{T}_N$ 。

用训练所得的两个网络可以对  $k$  时刻的实测数据  $\mathbf{X}(k)$  进行重构和故障诊断。对实测的后150组数据进行故障模拟,将所观测的主接力器行程传感器值  $y$  加一个增量  $\Delta y$ ,模拟位移传感器恒偏差故障。将实测数据和模拟了故障的数据分别输入反向传播网络和输入训练神经网络,得到数据的重构值  $\hat{\mathbf{X}}$  和  $SPE$  值,绘制两组数据的  $SPE$  曲线(如图3所示)。



(b) 有故障发生

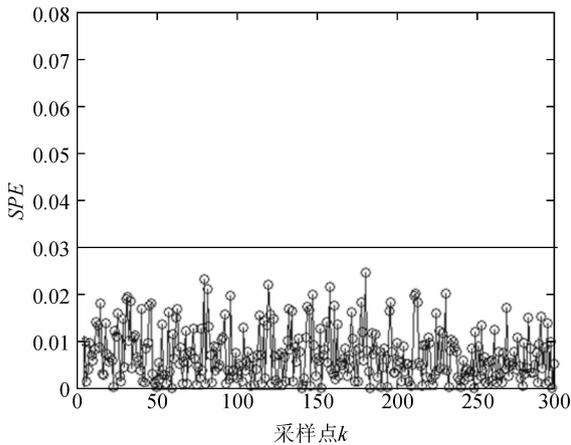
图3 故障检测  $SPE$  曲线

Fig.3  $SPE$  curves of fault detection

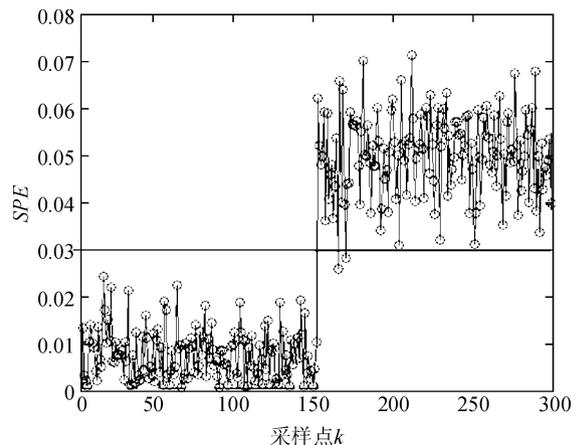
图中直线为故障发生的  $SPE$  阈值,图3(a)中所有的点都落在阈值以下,说明没有故障发生,图3(b)从第151个模拟故障点开始,曲线有跳变,越过阈值,说明有故障发生,可见,可以通过对重构向量  $SPE$  值的判断而较明显地检测出故障。

对故障进行识别,仍然使用上述的150组模拟故障数据,对任一组实测向量  $\mathbf{X}(k)$ ,得到其重构向量  $\hat{\mathbf{X}}(k)$ ,用  $\hat{\mathbf{X}}(k)$  的第  $j$  个分量  $\hat{X}_j(k)$  代替  $k$  时刻实测值的  $X_j(k)$  分量,得到有别于  $\mathbf{X}(k)$  的另一个  $k$  时

刻的实测向量  $\mathbf{X}'(k)$ ,根据式(13),利用  $\hat{\mathbf{X}}(k)$  与  $\mathbf{X}'(k)$  构造  $SPE$  值,绘制相应的  $SPE$  曲线,如图4所示。从图4(a)可以看出,当导叶开度行程传感器实测值  $y$  被重构后,其  $SPE$  点全部落在阈值线以下,说明导叶行程传感器值  $y$  重构后故障被消除了,可以判定导叶行程传感器为故障传感器。其余传感器实测值被重构后,  $SPE$  点从150点以后依然落在阈值线以上,类似于图4(b)所示曲线,说明重构的不是故障传感器。



(a) 故障传感器被重构时



(b) 无故障传感器被重构时

图4 故障识别 SPE 曲线

Fig. 4 SPE curves of fault identification

### 3 结论

用输入训练神经网络建立非线性主元和实测值的映射关系,用前馈网络建立其逆映射,根据这两个映射关系得到了实测向量  $X$  的重构值  $\hat{X}$ ,通过分析它们平方预测误差 SPE 进行故障检测和识别。

将这种方法应用到水轮机调节系统的故障诊断,将水头  $H$ 、流量  $q$ 、有功功率  $w$ 、转速  $x$ 、接力器行程  $y$ 、中间接力器位移  $h$  作为实测向量,选择适当的网络结构和训练算法,用现场运行数据对网络进行训练并模拟故障诊断的过程,仿真结果证明该方法有效。

### 参考文献:

- [1] Hastie T, Stuetzle W. Principal curves [J]. Am Stat Asso, 1989, 84(406):502-517.
- [2] Kramer M A. Nonlinear principal component analysis using auto-associative neural networks [J]. AIChE, 1991, 37(4):313-328.
- [3] Dong D, McAvoy T J. Nonlinear principal component analysis-based on principal curves and neural networks [J]. Computers and Chemical Engineering, 1996, 20(1):65-78.
- [4] Tan S, Mavrouniotis M L. Reducing data dimensionality through optimizing neural network inputs [J]. AIChE, 1995, 41(6):1471-1480.
- [5] Jia F, Martin E B, Morris A J. Non-linear principal component analysis with applications to process fault detection [J]. International Journal of Systems Science, 2000, 31

(11): 1473-1487.

- [6] 李尔国,俞金寿. PCA 在过程故障检测与诊断中的应用 [J]. 华东理工大学学报, 2001, 27(10):572-576.  
Li Erguo, Yu Jinshou. Fault detection and diagnosis based on principal component analysis [J]. Journal of East China University of Science and Technology, 2001, 27(10):572-576.
- [7] Nomikos P, MacGregor J F. Multivariate SPC charts for monitoring batch processes [J]. Technometrics, 1995, 37(2):41-59.
- [8] Dunia R, Qin S J, Edgar T F, et al. Identification of faulty sensors using principal component analysis [J]. American Institute of Chemical Engineers Journal, 1996, 42(10):2797-2812.
- [9] 陈楚瑶,朱大奇. 神经网络主元分析的传感器故障诊断方法 [J]. 系统工程与电子技术, 2010, 32(7):1549-1552.  
Chen Chuyao, Zhu Daqi. Sensor fault diagnosis method based on neural network principal component analysis [J]. Systems Engineering and Electronic, 2010, 32(7):1549-1552.
- [10] 牛征,刘吉臻,牛玉广. 动态多主元模型故障检测方法在变工况过程中的应用 [J]. 动力工程, 2005, 25(4):554-558.  
Niu Zheng, Liu Jizhen, Niu Yuguang. Fault detection under varying load conditions based on dynamic multi-principal component models [J]. Journal of Power Engineering, 2005, 25(4):554-558.

(责任编辑 王卫勋)