

基于 PCA 的管道缺陷导波信号特征优化方法

许桢英 罗来齐 王 匀 俞慧芳 刘 欢

(江苏大学机械工程学院 镇江 212013)

摘要: 针对超声导波管道缺陷检测中存在的识别率低、鲁棒性差等问题,应用了主成分分析对管道缺陷回波信号进行特征优化。首先,通过对超声导波缺陷回波信号进行处理,提取了信号在时域和时频域内的特征参数,构成联合特征向量。然后使用主成分分析法(principal component analysis, PCA)对联合特征向量进行降维处理,通过提取累计贡献率达到 89% 的主成分得到融合特征。最后用 BP 神经网络对融合特征进行训练和识别。这种方法可以有效的识别管道缺陷,与联合特征向量相比具有更高的识别率。

关键词: 超声导波;管道检测;缺陷识别;主成分分析;特征优化;BP 神经网络

中图分类号: TP2 **文献标识码:** A **国家标准学科分类代码:** 510.4030

Feature optimization method of pipe defect guided signals based on PCA

Xu Zhenying Luo Laiqi Wang Yun Yu Huifang Liu Huan

(School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: In order to solve low recognition rate and poor robustness in pipe defect recognition based on ultrasonic guided wave, principal component analysis (PCA) is used to optimize the feature of pipe defect echo signals. First, a few characteristic parameters in domain of time and time-frequency were extracted by means of dealing with the ultrasonic guided echo signals of the pipe defect to construct a multi-feature vector. The multi-feature vector dimension is then reduced using principal component analysis. The fusion feature is generated by extracting the principal component whose cumulative contribution rate is about 89%. Finally, BP neural network is used to train and recognize fusion feature. This method can effectively recognize the pipe defect, and has higher recognition rates than that of the multi-feature vector.

Keywords: ultrasonic guided wave; pipe defection; defect recognition; PCA; feature optimization; BP neural network

1 引言

管道超声导波检测因为其一次扫描即可测量整个长度范围内外表面所有缺陷、能实现远距离缺陷的快速定位、可以实现在役检测等优点在管道检测中具有良好的应用前景。国内外对这一领域做了广泛的研究^[1-2]。管道的定位和定性是管道检测研究的热点。缺陷的定性分析需要提取一些表征缺陷类型的特征参数,为了全面的反映缺陷信息,一般同时采用多种特征参数,而多个特征参数之间的相关性往往会造成分析的困难。同时特征参数过多又会给模式识别计算方面带来负担。因此选用合适的特征提取方法和分类器对缺陷的正确识别有很大帮助。

PCA 是基于二阶统计特性的一种多元统计方法,在信息压缩和消除数据相关性方面有显著地效果,有效用于人脸识别、故障检测等特征提取^[3-7]。神经网络技术因为综合

分析数据的能力在模式识别方面得到了广泛的应用。文中选用 BP 神经网络作为分类器,对 PCA 融合后的管道缺陷的特征参数进行数据融合,实现了多类缺陷的智能识别,有效地提高了缺陷的识别率。

2 主成分分析(PCA)

设管道缺陷信号模式向量矩阵为 $\boldsymbol{X}_{(m \times n)}$, 首先为了消除原始数据各维数据之间量纲不同和数值差异过大带来的影响,对原始数据做标准化处理:

$$\bar{\boldsymbol{X}}_j = \frac{1}{m} \sum_{i=1}^m \boldsymbol{X}_{ij} \quad (1)$$

$$s_j = \sqrt{\frac{1}{m} \sum_{i=1}^m (\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_j)^2} \quad (2)$$

$$z_{ij} = \frac{\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_j}{s_j} \in \boldsymbol{Z} \quad (3)$$

式中: \bar{x}_j 为样本均值, s_j 为样本标准差, $i = 1, 2, \dots, m; j = 1, 2, \dots, n$, z_{ij} 为数据标准化后的数据。并建立 Z 的相关系数矩阵 $\mathbf{T} = (t_{ij})_{n \times n}$, 式中:

$$t_{ij} = \frac{\sum_{k=1}^m (Z_{ki} - \bar{Z}_i)(Z_{kj} - \bar{Z}_j)}{\sqrt{\sum_{k=1}^m (Z_{ki} - \bar{Z}_i)^2 \sum_{k=1}^m (Z_{kj} - \bar{Z}_j)^2}} \quad (4)$$

求解矩阵 \mathbf{T} 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ 及相应的单位特征向量, $\mathbf{a}_1 = [a_{11}, \dots, a_{m1}]$, \dots , $\mathbf{a}_n = [a_{1n}, \dots, a_{mn}]$ 。得到新的特征向量 $\mathbf{F}_j = a_{1j}\mathbf{x}_1 + \dots + a_{mj}\mathbf{x}_m, j = 1, \dots, n$ 。所有的特征向量组成一个 n 维正交空间。特征向量对应的特征值越大, 它在重构时的贡献也越大, 特征值越小的特征向量在重构时的贡献就越小。根据特征值的累积贡献率来确定主成分, 其累积贡献率为:

$$\varphi(j) = \frac{\sum_{k=1}^j \lambda_k}{\sum_{k=1}^n \lambda_k} \quad (5)$$

一般选取 $\varphi(j)$ 在 85%~95% 的整数 $j (j < n)$, 就确定了前 j 个主成分来表征 85%~95% 原始数据信息, 在保证信息完整的情况下, 达到数据降维。

3 缺陷回波信号特征优化

3.1 缺陷回波信号特征分析

超声导波在管道传播过程中遇到声阻抗突变的结构, 会产生反射回波和模态转换, 由于入射角度的不同以及反射点存在空间分布区域差异, 导波在遇到不同形状的缺陷时, 其反射回波在时域、时-频域存在差异。采用人工智能方法进行缺陷的识别, 要首先提取能反映不同缺陷类别的特征。

时域分析: 研究发现, 在相同强度激励信号下, 缺陷反射回波的幅值随缺陷截面积的增大而呈线性增高, 因此诸如幅值、能量、包络面积等时域有量纲特征指标不能用来表征缺陷的类型。缺陷回波信号在时域的差异主要体现在时域无量纲特征参数, 本文选用峰度系数、偏度系数、离散系数和形状系数^[8] 4 个时域无量纲特征参数。峰度系数用来描述缺陷信号的陡峭度, 偏度系数用来描述缺陷信号幅值统计分布的对称度, 离散系数用来描述缺陷信号单位均值上的离散程度, 形状系数为缺陷信号的轮廓长度与轮廓线包围面积的比值。

时频域分析: 由于不同缺陷对导波的影响, 使得缺陷回波信号在各频段内的能量分布产生差异。小波包分析的每一层分解都同时对低频部分和高频部分进行分解, 因此信号可以获得更高的时频分辨率, 从小波包分解结果提取频段能量作为信号特征可以有效反映缺陷回波信号在各频带内能量分布的差异^[9-10]。

对不同缺陷的回波信号进行傅里叶变换, 发现信号的频率主要集中在 200 kHz 以下, 本研究的信号的采样频率为 2.5 MHz, 根据奈奎斯特采样定理可得出采集有效信号

的最高频率为 1.25 MHz。根据小波包分解理论, 研究选用 db9 小波基对缺陷回波信号进行 7 层小波包分解, 可将信号分解为 128 个频段, 那么前 20 个频段的信号将包括信号频率为 0.195 kHz 的范围。因此取前 20 个频段的信号作为研究对象, 就可以包括原信号时频域中的绝大多数有用信息。计算第 7 层前 20 个频段信号的能量与第 7 层总能量的比值, 得到一个 20 维的特征向量 $\{x_1, x_2, \dots, x_{20}\}$, 并将其作为缺陷回波信号的时频域特征值, 其中 x_k 可由式 (6) 计算获得, 即

$$x_k = x_{7,j} / x_7 \quad (6)$$

式中: $j = 0, 1, \dots, 19; k = 1, 2, \dots, 20$, X_k 为第 k 个频段能量与第 7 层总能量比值, $x_{7,j}$ 为第 7 层上第 j 个频段的能量; x_7 为第 7 层上各个节点的总能量。

分别提取腐蚀、孔洞、裂纹 3 种缺陷回波信号, 每种缺陷测取 30 组数据进行计算, 提取其缺陷特征参数。

图 1 给出了管道 3 种缺陷的 4 个时域特征参数, 它们是峰度系数(c1)、离散系数(c2)、偏度系数(c3)和形状系数(c4), 其中横坐标为样本序号, 从 1 开始每 30 个样本属于一种缺陷; 纵坐标为幅值。由图 1 可知, 这些特征参数对缺陷信号的表征各不相同, 有的参数对缺陷反映较敏感, 有些则比较稳定。

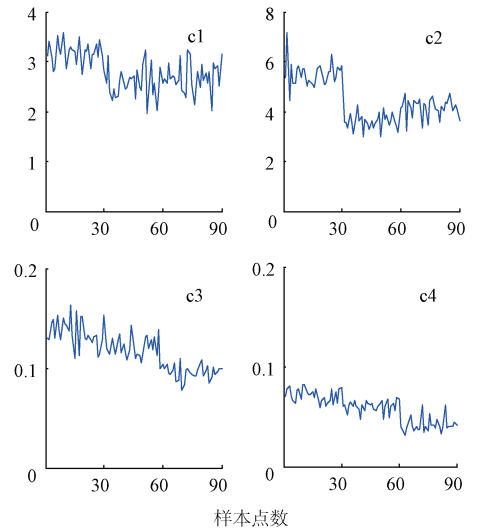


图 1 缺陷回波信号时域特征参数

图 2 给出了腐蚀(d1)、孔洞(d2)和裂纹(d3) 3 种缺陷信号的前 20 个频段归一化能量分布。从图 2 可以看出, 3 种缺陷信号前 20 个频段内的能量分布有较大差别。

4 个时域特征参数和 20 维小波包能量谱组成 24 维联合特征参数矩阵。

3.2 基于 PCA 的特征提取

PCA 用少数的几个综合特征来代表众多的原始特征, 综合特征由原始特征线性组合构成, 相互之间正交, 消除了特征之间的相关性, 同时最大限度地反映原始特征集的有

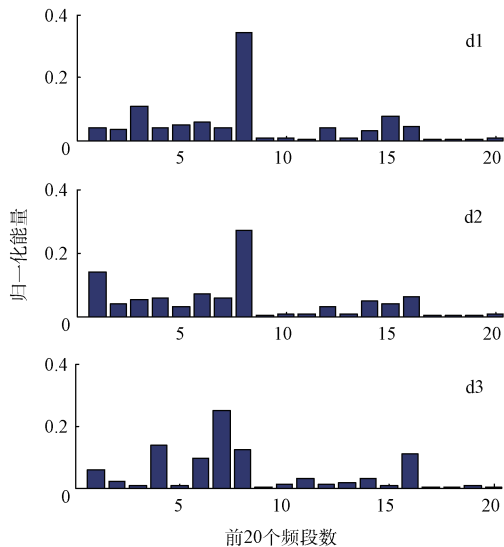


图 2 三种缺陷回波信号小波包能量谱

用信息。管道缺陷类型识别需要获取对缺陷反映敏感性强的特征来进行样本分类,在小波包能量谱中,不同频段的能量特征值之间存在一定的关联性,使得特征向量集存在一定程度上的信息重叠,同时主观选择很可能包括与输出相关性很小的输入向量,增加了神经网络陷入局部极小点的可能。因此,在上述特征分析基础上来降维并提取对缺陷敏感的特征尤为重要。采用 PCA 方法来提取反映缺陷类型的主要特征参数。

分别选取腐蚀(d1)、孔洞(d2)和裂纹(d3)3种缺陷信号各30组进行分析。图3和图4分别是3种缺陷信号的时域波形和频域图,由图3和图4可以看出,从时域和频域不能直接将3类管道缺陷进行分类。基于前述分析,采用特征参数的低维主成分来表征管道缺陷类型,首先提取所有信号的特征参数组成联合特征向量,在预处理后,求解其相关系数矩阵的全部特征值及对应的特征向量。然后可以得到联合特征参数的主分量。由式(5)计算,前4个主成分

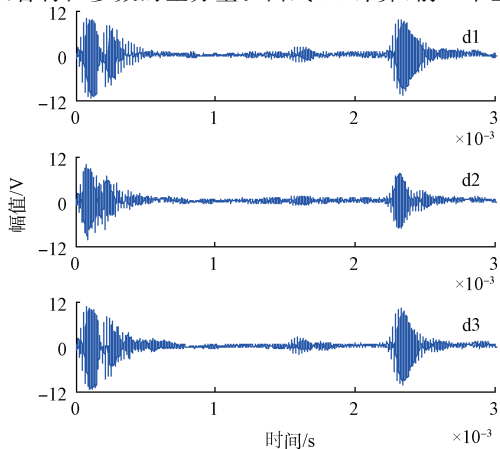


图 3 3种缺陷管道回波信号

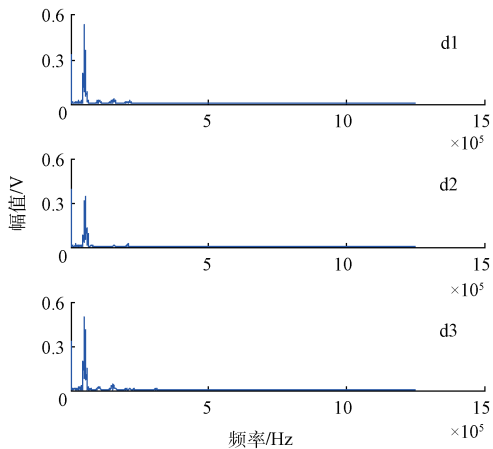


图 4 回波信号频谱

的累积方差贡献率达到 89%,因此选用前 4 个主成分来组成融合特征向量。图 5 为利用 PCA 方法提取的前 4 个特征参数主成分。由图可以看出,前两个特征参数主成分 p1 和 p2 能较好地表征不同缺陷的区别特征。而其他特征参数主成分反映的是局部变化特征,反映不同缺陷特征的能量相对较弱。

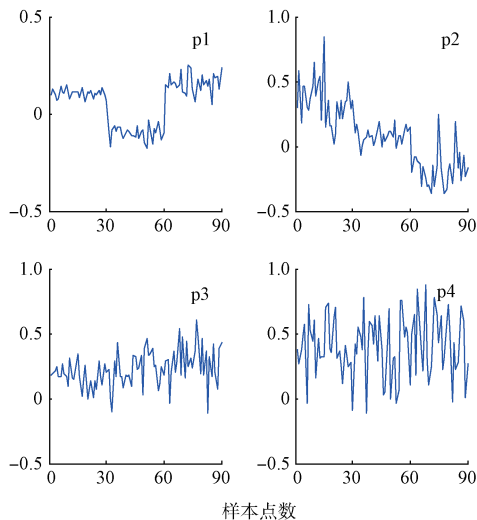


图 5 缺陷信号前 4 个主成分特征

4 管道缺陷识别

4.1 实验

如图 6 所示,实验系统包括管道、压电传感器阵列、任意函数发生器、功率放大器、前置放大器、示波器及计算机,本实验中,被测试的管道为 3 根碳钢无缝管,每根长 3 m,钢管外直径为 76 mm,壁厚为 4 mm,材料为 Q235。在管道 2 m 处制作孔洞、裂纹、腐蚀 3 种人工缺陷。压电传感器为 PZT-5 材质压电陶瓷片,将 20 片压电传感器构成的阵列用于激励管道 L(0,2) 模态导波,8 片压电传感器并

联构成的阵列用于采集激励信号和回波信号。激励信号采用 10 个周期的汉宁窗调制正弦波,采样频率为 2.5 MHz。



图 6 超声导波管道检测实验系统

激励信号的中心频率为 90 kHz,通过调节激励信号的峰峰值,对于 3 种缺陷类型、每种缺陷的 3 种尺寸,于缺陷回波最大处重复采集 10 次,得到 30 个孔洞、30 个裂纹、30 个腐蚀共 90 个缺陷样本。缺陷的类型、形状等数据如表 1 所示。

表 1 缺陷信息

类型	形状	尺寸/mm
孔洞	圆形	$\phi 2/6/10$
裂纹	线形	截面积比 4/6.8/10%
腐蚀	长方形	15 * 15/15 * 30/30 * 30

4.2 结果与分析

对于采集到的 90 组缺陷信号,可以获得 90×24 维联合特征矩阵,采用前述的 PCA 方法,提取前 4 个主成分特征,如图 5 所示。4 个主成分特征构成一个 90×4 的融合特征矩阵,每种缺陷的前 20 组数据用于训练,后 10 组数据用于测试。用神经网络分别对联合特征矩阵和融合特征矩阵进行训练和识别,其中网络传递函数选择 log-sigmoid 函数,训练次数设为 1000 次,性能参数设为 0.0001,对于联合特征矩阵训练了 280 步即达到了性能参数的要求,而对于融合特征矩阵训练了 64 步即达到了性能参数的要求。

为了提高识别率的精度,两次测试均训练 5 次神经网络,并对测试样本分别进行识别,得到 PCA 处理前后的平均识别率如表 2 所示。

表 2 PCA 处理前后管道缺陷识别率

缺陷类型	测试样本个数	联合特征平均识别率	融合特征平均识别率
腐蚀	10	80%	82%
裂纹	10	76%	82%
孔洞	10	86%	90%

识别结果表明,主成分分析可以有效地消除各变量之间的冗余,有效地减少了特征的维数,减少了后续训练和识别的时间,提高了识别的准确率。

5 结 论

应用超声导波技术对带有人工缺陷管道进行检测,并对缺陷信号在时域和时频域的特征进行了分析研究,利用 PCA 对原始特征参数进行融合,提取原始特征参数的主成分表示。采用的主成分分析法,既能降低维数减少计算的复杂性,又能有效表征缺陷类型。

利用 BP 神经网络作为分类器,对原始特征和融合特征进行识别,实现了 3 种缺陷的识别。使用 PCA 特征融合能有效地提高缺陷的识别率。有助于管道质量的评定。

参考文献

- [1] RATASSEPP M, FLETCHER S, LOWE M J S. Scattering of the fundamental torsional mode at an axial crack in a pipe[J]. The Journal of the Acoustical Society of America, 2010, 127(2): 730-740.
- [2] 姜银方,井然,严有琪,等. 基于 L(0,2)模式导波的缺陷反射信号的实验研究[J]. 机械科学与技术,2014(04):551-554.
- [3] 刘永斌,何清波,孔凡让,等. 基于 PCA 和 SVM 的内燃机故障诊断[J]. 振动. 测试与诊断,2012(02):250-255,342.
- [4] 叶永伟,刘志浩,黄利群. 基于 PCA 的汽车涂装线设备信号特征提取[J]. 仪器仪表学报,2011(10): 2363-2370.
- [5] 常玉清,王姝,王福利,等. 基于多 PCA 模型的过程监测方法[J]. 仪器仪表学报,2014(04):901-908.
- [6] 胡敏,程天梅,王晓华. 融合全局和局部特征的人脸识别[J]. 电子测量与仪器学报,2013(09):817-822.
- [7] 王培良,夏春江. 基于 PCA-PDBNs 的故障检测与自主学习辨识[J]. 仪器仪表学报,2015(05):1147-1154.
- [8] 胡文刚,刚铁. 基于超声信号和图像融合的焊缝缺陷识别[J]. 焊接学报,2013(04):53-56,115-116.
- [9] 孙洁娣,靳世久. 基于小波包能量及高阶谱的特征提取方法[J]. 天津大学学报,2010(06):562-566.
- [10] 丁硕,常晓恒,巫庆辉,等. 基于 Elman 神经网络的传感器故障诊断研究[J]. 国外电子测量技术,2014(04):72-75.

作者简介

许桢英,博导,教授,主要研究方向为光、电无损检测理论与技术、机器视觉理论及检测技术等。

Email: xuzhenying@ujs.edu.cn

罗来齐,硕士研究生,主要研究方向为管道检测,模式识别。