

压缩感知域智能天然气管道泄漏孔径识别^{*}

孙洁娣¹, 乔艳雷¹, 温江涛²

- (1. 燕山大学信息科学与工程学院 秦皇岛 066004;
2. 燕山大学 河北省测试计量技术及仪器重点实验室 秦皇岛 066004)

摘要:天然气管道泄漏监测正在进入大数据时代,针对传统方法存在的采集数据冗余、特征提取及识别受主观因素影响较大等问题,结合压缩感知与深度学习理论,提出一种在变换域进行泄漏信号的压缩采集、在压缩感知域进行自适应特征提取及识别的智能天然气管道泄漏孔径识别方法。通过随机高斯矩阵获取压缩采集数据,并通过深度学习挖掘测量信号中隐藏的泄漏孔径信息,经稀疏滤波实现特征的自动筛选,最后研究了 softmax 回归实现孔径的高精度分类识别。实验结果表明,该方法实现了监测数据的压缩,对压缩感知域采集信号的识别性能明显优于传统方法。

关键词:管道泄漏孔径识别;压缩感知采集;深度学习;稀疏滤波;softmax 回归

中图分类号: TH865 TE88 **文献标识码:** A **国家标准学科分类代码:** 460.40

Pipeline leak aperture identification based on compressed sensing

Sun Jiedi¹, Qiao Yanlei¹, Wen Jiangtao²

- (1. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China; 2. Key Laboratory of Measurement Technology and Instrumentation of HeBei Province, Yanshan University, Qinhuangdao 066004, China)

Abstract: Natural gas pipeline leak monitoring is entering the age of big data. Aiming at the problems of traditional methods, such as redundant data, subjective feature extraction and identification, an intelligent pipeline leak aperture identification method is proposed combined compressed sensing (CS) and deep learning theory, which can achieve compressed sampling, adaptive feature extraction and recognition. The random Gaussian matrix is used to acquire the compressed acquisition data, and the aperture information contained in measured samples in CS domain is analyzed by deep learning. The sparse filtering is applied to realize the automatic feature selection. Finally, the high precision classification and recognition of the aperture is obtained by softmax regression. Experimental results show that this method realizes the compression of the monitoring data, and the identification performance for data of compressed sensing domain is better than traditional methods.

Keywords: pipeline leak aperture recognition; compressed sensing samples; deep learning; sparse filtering; softmax regression

0 引言

随着经济发展及环保要求的提高,我国的天然气管网已进入高速发展阶段。输气管网规模增大以及运营时间增加的同时,因材料缺陷、腐蚀等原因而产生泄漏的危险逐渐增加,已成为威胁输气管网健康运营的重要因素。由于天然气管道泄漏会造成巨大的经济损失、严重的人员损失及不良的环境后果,因此进行天然气管道泄漏实

时检测意义重大^[1]。

目前天然气管道实时泄漏检测主要是基于硬件检测和软件分析的方法、通过长时间在线监测的方式实现。常见的管道泄漏监测系统一般包含数据采集、数据通信、数据处理与存储3个部分,依据 Nyquist 采样定理,由计算机控制的实时高速采集系统进行全天候的管道状态数据采集,由此产生的大量监测数据经通信通道汇聚到处理与存储系统,以实现泄漏信息获取和历史数据存储。随着监控网络规模的扩大和监控点不断增加,采集的管

道状态数据越来越多,形成海量数据,使得泄漏监测进入了“大数据”时代^[2]。油气管道泄漏本质上也是一种故障,随着大量监测数据的产生,如何从海量数据中挖掘信息,实现自动、高效、准确的识别管道的故障状况,成为管道大数据智能检测面临的新问题。

目前研究成果中,各种管道泄漏检测系统多采用“高速采样—海量数据—宽带通信—少量有用信息”的采集与处理模式,相对于所获取的有用信息量,采集的数据过度冗余,使得基于传统时频域及变换域的特征提取及识别算法复杂、且易受混杂的干扰及噪声影响,并且由于数据量巨大,采集信号后期处理会对硬件在计算能力、存储容量与传输带宽等多方面提出过高要求。天然气管道泄漏孔径与其危险程度密切相关,为了获取泄漏孔径信息,目前的研究成果通常采用“特征提取+构造分类器”的方法,即采用传统高速实时采集系统获取时域信号之后,在时域、频域及变换域研究复杂的信号处理方法^[3-4]、基于先验知识提取出信号的特征。为了全面反映管道的监测状态,通常特征维数较高,其中掺杂的无用特征,可能会对诊断结果产生较大影响,因此还需依赖专家的诊断经验及模式识别相关原理,如主成分分析法、特征判别式分析等降维方法^[5-6],对特征进行进一步的主观选择,选出更为敏感的泄漏信号特征,最后通过构造分类器,利用神经网络、支撑向量机等^[7-8]实现泄漏孔径的识别。

近年来发展的压缩感知(compressed sensing, CS)^[9-10]以及深度学习^[11-12]理论提出了全新的信息感知及挖掘方式,将对信号的采样变为对信息的采样,为天然气管道泄漏信息采集与分析提供了新的思路。本文所提方法正是在此新思路下,针对传统时域采样处理系统存在的问题,提出在变换空间实现远低于奈奎斯特定理要求的数据压缩采集;结合深度学习理论,提出构建以稀疏滤波为基础的深层次学习模型,通过贪婪逐层训练算法,直接对压缩感知域信号进行自适应特征提取,并通过构造回归分类器实现泄漏孔径的高精度分类识别。所提方法对建立新型的管道泄漏采集及处理结构、降低硬件复杂度及成本、减少采集数据量等具有重要意义。

1 基于压缩感知的泄漏数据采集

Candes E. J. 等人^[8]基于信号稀疏性提出的压缩感知理论指出,如果被测信号 $\mathbf{X} \in \mathbf{R}^{A \times 1}$ 本身是稀疏的或在某个变换域 Ψ 上的变换系数 Θ 是稀疏的,就可以用一个与变换基 Ψ 不相关的观测矩阵 $\Phi \in \mathbf{R}^{B \times A}$ ($B \ll A$) 实现线性投影,得到观测向量 $\mathbf{Y} \in \mathbf{R}^{B \times 1}$,然后利用优化求解方法从观测集合 \mathbf{Y} 精确或高概率地重构原始信号 \mathbf{X} ,其观测模型如式所示:

$$\mathbf{Y} = \Phi\mathbf{X} = \Phi\Psi\Theta \quad (1)$$

压缩感知实现了信号从 A 维到 B 维的降维投影,从而实现数据的压缩。压缩感知作为一种全新的信号获取与处理的理论框架,将传统信号采样与数据压缩合并进行,不再受 Nyquist 定理的采样率限制,只需少量的采样测量值即可获取原始信号的全部信息,实现时域信号到压缩感知域信号的直接获取。压缩感知理论主要包括信号的稀疏表示、观测矩阵的设计以及信号的重构 3 个方面^[13-14],其中观测矩阵的作用就是实现对原始信号的降维投影,从而实现数据的压缩。定义压缩率(compression ratio, CR)为:

$$CR = \frac{A - B}{A} \times 100\% \quad (2)$$

CR 用来衡量泄漏信号的压缩度,CR 越大,泄漏信号的压缩率越高,测得的压缩感知域数据越少,其中 A 为原始泄漏信号的样本长度, B 为测量矩阵观测后获得的信号长度,因此不同维度的观测矩阵即可实现不同程度的压缩采集。

压缩感知测量过程中,测量矩阵 Φ 必须服从“有限等距性质(restricted isometric property, RIP)”^[15],即对于任意具有严格 K -项稀疏矢量,矩阵都能保证:

$$1 - \delta \leq \frac{\|\Phi\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} \leq 1 + \delta \quad (3)$$

式中: δ 为 RIP 常数, $0 < \delta < 1$ 。

目前常用的测量矩阵包括随机高斯矩阵、随机伯努利矩阵、随机次高斯矩阵等,Candes E. J. 等人^[16]证明了随机高斯矩阵几乎与任意稀疏信号都不相关,能以较大概率满足约束等距性条件,具有良好的普适性,因此在压缩感知研究中被广泛使用。

管道泄漏信号具有稀疏特性,因此本文提出利用随机高斯矩阵实现变换域投影,并研究以不同的压缩率将高维泄漏信号投影到低维空间上,在保留泄漏信号大部分重要信息基础上实现数据量的大幅度压缩。

2 基于稀疏滤波的自适应特征提取

传统泄漏信号分析方法多针对时域信号进行主观特征提取,通过分析间接计算所得特征而非测量信号来实现识别。本文压缩采集方法获取的是泄漏信号的压缩感知域测量信号,少量信号即含有极大量信息,直接分析该信号实现孔径识别是更加有效的泄漏孔径分类方法。

深度学习通过学习复杂的非线性关系,挖掘隐藏在数据内部的重要信息,放大信息中那些能够区别不同原始数据的部分,并抑制不相关部分,从而改变传统方法中对先验知识、人工和经验的严重依赖,使得泄漏诊断实现真正的智能化、自动化^[17]。泄漏诊断通常包括特征学习过程,稀疏自编码、堆叠自编码等传统的无监督学习方法需要对输入数据进行建模估计,因而受参数设置影响较

大,灵活性较差^[18]。稀疏滤波学习方法由 Ngiam J 等人^[19]提出,其核心思想就是避免对数据分布的显式建模,通过直接优化特征分布的稀疏性从而得到更好的特征表达,具有调节参数少、能处理高维度数据的优点^[20]。本文正是基于该算法的优点,提出将压缩感知域获取的采集信号通过稀疏滤波学习,自适应的获取其中最能代表泄漏孔径特征的信息,相比传统特征提取方法,此方法从携带大量泄漏信息的“原始”采集信号中学习得到的特征,更具有代表性,有利于进一步的准确分类识别。

在泄漏孔径特征提取过程中,设每类信号样本表示为 $\mathbf{x}^{(i)}$, 则所有样本构成的样本集为:

$$\mathbf{X} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots, \mathbf{x}^{(N)}\} \quad (4)$$

式中: N 为样本总数。

假设待识别的泄漏信号由 P 类特征构成,引入特征分布矩阵,矩阵的每一行是一个特征,每一列是一个样本,则特征分布矩阵可表示为:

$$\mathbf{F} = \begin{pmatrix} f_1^{(1)} & f_1^{(2)} & \dots & f_1^{(N)} \\ f_2^{(1)} & f_2^{(2)} & \dots & f_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ f_p^{(1)} & f_p^{(2)} & \dots & f_p^{(N)} \end{pmatrix} \quad (5)$$

式中: $f_j^{(i)}$ 表示第 i 个样本的第 j 个特征的激活值。

对于天然气管道泄漏智能诊断系统来说,提取到好的泄漏孔径特征至关重要(即获取好的特征分布),需要满足 3 个条件。

- 1) 每个样本的特征应该是稀疏的,即特征分布矩阵的每一列中只有少数非零项,其余的都是零值。
- 2) 样本间的特征应该是稀疏的,即特征分布矩阵的每一行只有少数非零项,其他的都是零值。
- 3) 特征的分布应是均匀的,即每种特征应该具有相似的统计特性,即特征的高分散性。

稀疏滤波能满足上述要求,它包括输入层和输出层。泄漏信号训练样本由输入层进入,经过该方法处理后即可自适应地得到训练特征,其结构简图如图 1 所示。

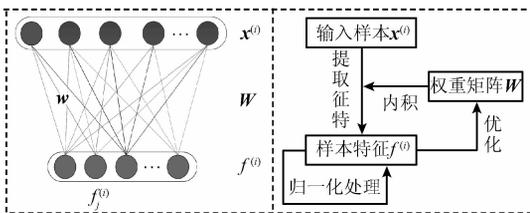


图 1 稀疏滤波结构简图

Fig. 1 Schematic diagram of sparse filtering

稀疏滤波方法通过优化目标函数实现由输入到特征的映射,从而提取到所需的特征。该方法只需要调整提

取特征数量一个参数即可实现对前述 3 个条件的同时优化。

设训练样本集为 $\{\mathbf{x}^{(i)}\}_{i=1}^M$, $\mathbf{x}^{(i)} \in \mathbf{R}^{N \times 1}$ 为训练集中的一个样本, M 表示该样本集中的样本个数,样本 $\mathbf{x}^{(i)}$ 在经权重矩阵 \mathbf{W} (特征提取函数) 处理后得到样本线性特征 $f_j^{(i)}$:

$$f_j^{(i)} = \sqrt{\varepsilon + (\mathbf{w}_j^T \mathbf{x}^{(i)})^2} \approx |\mathbf{w}_j^T \mathbf{x}^{(i)}| \quad (6)$$

式中: $\varepsilon = 10^{-8}$ 为误差项。

对特征分布矩阵按式(7)、(8)分别进行行、列归一化处理。

$$\tilde{f}_j = \frac{f_j}{\|f_j\|_2} \quad (7)$$

$$\tilde{f}^{(i)} = \frac{\tilde{f}^{(i)}}{\|\tilde{f}^{(i)}\|_2} \quad (8)$$

特征分布矩阵归一化使每个特征为相等的激活值,并使其落于二范数的单位球面上。归一化后,如果只有一个特征分量增大,那么奇异特征分量值将会减小,反之亦然。

设定提取特征数量以及目标函数如式(9),通过有限拟牛顿法(limited-memory BFGS, L-BFGS)等算法优化目标函数直至收敛,对特征分布进行优化:

$$\underset{\mathbf{W}}{\text{minimize}} \sum_{i=1}^M \|\tilde{f}^{(i)}\|_1 = \sum_{i=1}^M \left\| \frac{\tilde{f}^{(i)}}{\|\tilde{f}^{(i)}\|_2} \right\|_1 \quad (9)$$

深度学习神经网络正是通过稀疏滤波构建深层次模型实现对泄漏信号的特征提取。它可以使用逐层贪婪算法来训练,先用稀疏滤波训练得到一个单层的归一化特征,然后用得到的特征当作第 2 层的输入去训练第 2 层,其他层依次类推,随着层数的增多,所提取到的特征越具有典型的表达性。

3 基于回归分类的泄漏孔径识别

在深度学习神经网络中,softmax 回归通常作为分类器应用于深度神经网络的最后一层^[21]。本文在自适应提取出所需特征的基础上,采用 softmax 回归最终实现对泄漏孔径的高精度分类识别。

设有训练样本集 $\{\mathbf{x}^{(i)}\}_{i=1}^M$, $\mathbf{x}^{(i)} \in \mathbf{R}^{N \times 1}$ 和标签集 $y \in \{1, 2, 3, \dots, K\}$, 对于每个输入样本 $\mathbf{x}^{(i)}$, 利用 softmax 回归预测出其对应标签集中每个标签时的概率 $p(y = j | \mathbf{x}^{(i)})$, 即估算出 $\mathbf{x}^{(i)}$ 的每一种分类结果所对应的概率。首先确定 softmax 回归模型,根据预测模型会针对每个输入样本 $\mathbf{x}^{(i)}$ 对应的每个标签 $y = 1, 2, 3, \dots, K$ 得出一个 K 维的预测向量(向量元素的和为 1)来表示 K 个估计的概率值,模型如下:

$$\mathbf{h}_\theta(\mathbf{x}^{(i)}) = \begin{bmatrix} p(y=1|\mathbf{x}^{(i)};\boldsymbol{\theta}) \\ p(y=2|\mathbf{x}^{(i)};\boldsymbol{\theta}) \\ \vdots \\ p(y=K|\mathbf{x}^{(i)};\boldsymbol{\theta}) \end{bmatrix} = \frac{1}{\sum_{j=1}^K e^{\theta_j^T \mathbf{x}^{(i)}}} \begin{bmatrix} e^{\theta_1^T \mathbf{x}^{(i)}} \\ e^{\theta_2^T \mathbf{x}^{(i)}} \\ \vdots \\ e^{\theta_K^T \mathbf{x}^{(i)}} \end{bmatrix} \quad (10)$$

式中: $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_k]^T$ 为一个矩阵, 矩阵的每一行可以看作是一个类别所对应分类器的模型参数, 总共 K 行。

$1/\sum_{j=1}^K e^{\theta_j^T \mathbf{x}^{(i)}}$ 可对概率分布进行归一化, 使得所有概率之和为 1, 则输入样本 $\mathbf{x}^{(i)}$ 分类为类别 j 的概率为:

$$p(y=j|\mathbf{x}^{(i)};\boldsymbol{\theta}) = \frac{e^{\theta_j^T \mathbf{x}^{(i)}}}{\sum_{j=1}^K e^{\theta_j^T \mathbf{x}^{(i)}}} \quad (11)$$

输出的最大值概率值则属于相应的类, 例如: P_1 最大, 则说明输入样本 $\mathbf{x}^{(i)}$ 属于这一类。

在上述模型的基础上, 通过代价函数 $J(\boldsymbol{\theta})$ (见式(12)) 对其进行训练, 并对代价函数 $J(\boldsymbol{\theta})$ 求导, 如式(13), 使用迭代优化算法(如梯度下降法或 L-BFGS) 对代价函数进行最小化并保证收敛到全局最优解:

$$J(\boldsymbol{\theta}) = -\frac{1}{M} \left[\sum_{i=1}^M \sum_{j=1}^K \mathbf{1}\{y=j\} \log \frac{e^{\theta_j^T \mathbf{x}^{(i)}}}{\sum_{i=1}^K e^{\theta_i^T \mathbf{x}^{(i)}}} \right] + \frac{\lambda}{2} \sum_{j=1}^K \sum_{i=1}^N \theta_{ji}^2 \quad (12)$$

$$\nabla_{\theta_j} J(\boldsymbol{\theta}) = -\frac{1}{M} \sum_{i=1}^M [\mathbf{x}^{(i)} (\mathbf{1}\{y=j\} - p(y=j|\mathbf{x}^{(i)};\boldsymbol{\theta}))] + \lambda \theta_j \quad (13)$$

权重衰减项 λ 是 softmax 回归中的重要参数, 它可以解决 softmax 回归的参数冗余所带来的参数值过大问题。当 λ 为任意大于 0 的值时, 代价函数 $J(\boldsymbol{\theta})$ 是一个严格的凸函数, 这时 softmax 回归模型确保有唯一解^[22], 在此基础上, 梯度下降法和 L-BFGS 等算法可以保证收敛到全局最优, 实现一个可用的 softmax 回归模型。

本文所述方法中, softmax 回归分类器的输入样本即为利用稀疏滤波对压缩采集信号提取到的样本特征 $f^{(i)}$, 与训练样本为 $\mathbf{x}^{(i)}$ 时一样, 通过对样本特征添加标签的有监督方式, 间接实现了对泄漏孔径的高精度分类识别。

4 本文所提方法

综上所述, 本文提出的压缩感知域智能天然气管道泄漏孔径识别方法主要处理过程为首先基于压缩感知理论在变换域选择观测矩阵, 实现对泄漏信号压缩采集; 其次在深度学习理论基础, 采用稀疏滤波方法, 构建深层次模型对压缩采集信号通过逐层贪婪的算法直接进行自适应特征提取; 最后将提取到的特征加入构建的 softmax

回归分类器中, 实现对泄漏孔径的高准确度分类识别。

5 实验结果及分析

5.1 实验情况概述及泄漏识别结果

实验使用的泄漏信号由分布于不同位置的传感器在不同压强下采集得到, 泄漏信号由 3 种模拟泄漏孔产生: 3、4、5 mm, 即故障类型为 3 类。实验中对不同条件下的大量泄漏信号进行了处理, 以验证本文方法, 并对重要参数对结果的影响进行了研究, 后文将逐一介绍。以泄漏样本长度 4 096 点为例, 根据压缩采集原理, 采用 819 行 4 096 列的随机高斯矩阵即可对 4 096 点的泄漏信号实现压缩采集, 获取 819 点的压缩采集信号, CR 为 80%; 将多组采集信号输入稀疏滤波器进行自适应特征提取, 稀疏滤波器的输出特征数 N_{out} 为 600; 将获得的特征输入 softmax 分类器进行泄漏孔径分类(权重衰减项 $\lambda = 1 \times 10^{-8}$), 进行多次实验后得到的平均诊断准确率为 99.499%。

按照类似过程对大量信号进行处理, 通过对实验结果的分析, 本研究组发现以下因素对于最终的识别结果具有较大影响, 包括压缩采集的 CR、稀疏滤波的输出参数个数 N_{out} 、回归分类器的权重衰减因子 λ 。下面对各因素的影响进行分析, 为了排除随机性影响以下均为进行 10 次实验的结果。

5.2 压缩率的影响

本文采用随机高斯矩阵作为观测矩阵, 根据式(2)可知压缩率与高斯矩阵的行列数有直接关系, 压缩率越大则获得的测量信号数量越少, 为了更好地研究压缩率对后续识别结果的影响, 本文研究了 CR 变化范围在 50% ~ 85% 时的识别准确率及所用时间, 其他步骤的参数如下设置: 稀疏滤波输出特征数为 600, softmax 回归的权重衰减 $\lambda = 1 \times 10^{-8}$, 从样本集中取出 3/4 的样本作为训练样本, 剩余的样本作为测试样本。图 2 所示为不同 CR 情况下, 每种情况各进行 10 次实验并取平均值的结

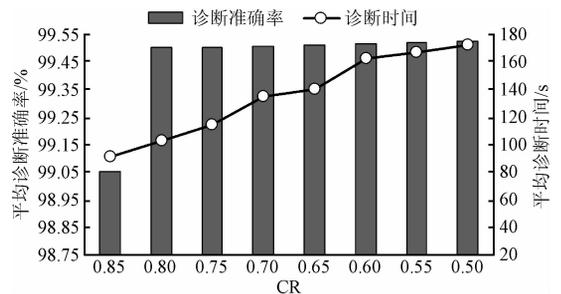


图2 CR 对诊断准确率的影响

Fig. 2 Effect of compression ratio on diagnostic accuracy

从图2中可以看出在 $CR = 85\%$ 的时候完成诊断用时最少,虽然平均准确率略低于其他情况,但仍可以实现准确分类;随着 CR 的减小,平均诊断准确率变化不明显,但耗时明显增加,表明当压缩率达到一定时,压缩采集的信号已包含所有泄漏信息,即使采集再多的数据,对提高诊断准确率贡献也不大,反而增加了识别时间,这也说明了本文提出的压缩采集方法可以实现利用很少的观测数据就能实现高准确度的分类识别。压缩率过高时泄漏孔径信息会有所损失,此时识别准确率会下降。因此实际中需综合考虑时间成本与诊断准确率的要求。本文取 $CR = 80\%$ 作为较理想的参数进行后续处理。

5.3 稀疏滤波输出特征数量的影响

压缩采集的数据作为稀疏滤波的输入,通过深度学习得到 N_{out} 个自适应特征输出。 N_{out} 的大小直接影响了稀疏滤波对特征分布矩阵及相关处理的结果,并直接影响后续的分类结果,因此此处对稀疏滤波阶段的输出特征数 N_{out} 进行研究。采用 80% 的压缩率,softmax 回归权重衰减 $\lambda = 1 \times 10^{-8}$ 的情况下,输出特征数 N_{out} 在 $400 \sim 700$ 的取值时,图3所示为各进行10次实验的平均识别准确率及诊断时间。

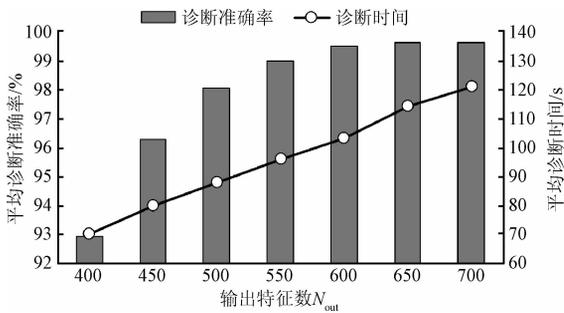


图3 输出特征数 N_{out} 对诊断结果的影响

Fig. 3 Effect of output features number on diagnostic accuracy

由图3中可看出随着输出特征数 N_{out} 的增加故障诊断的准确率也随之提高,当 $N_{\text{out}} = 600$ 时准确率为 99.499% ,接近 100% ,而后趋于平缓;但是随着 N_{out} 的增大诊断时间明显增加,当输出特征数大于 600 后,诊断准确率无明显变化,但诊断时间增加明显,因此综合考虑诊断及诊断准确率的影响, $N_{\text{out}} = 600$ 是较好的选择。

以上分析中稀疏滤波的输入是压缩率为 80% 情况下的压缩采集数据,通过实验选择输出特征数量。在实际应用中,也需综合考虑准确率及识别耗费时间选择稀疏滤波输出特征数。

5.4 softmax 回归权重衰减项的影响

在 $CR = 80\%$,稀疏滤波输出特征数为 600 的情况下,自适应地提取特征并输入 softmax 回归分类器进行识别,权重衰减项 λ 对回归分类器影响较大,下面对 λ 的影响进行研究,图4所示为不同 λ 下的分类识别结果。

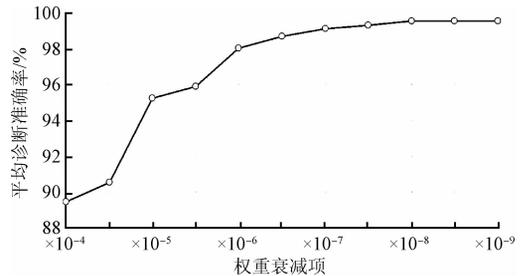


图4 权重衰减项对孔径识别准确率的影响

Fig. 4 Effect of weight decay item on the diagnostic accuracy

从图4中可以看出当 λ 较大时,识别准确率较低,这是由于 λ 过大会使 softmax 回归在分类过程中得到的特征参数大多趋近于 0 ,这样会严重影响系统的诊断准确率;随着 λ 的减小,诊断准确率逐渐增加,直到权重衰减项 $\lambda = 1 \times 10^{-8}$ 时可达 99.499% ,而后趋于平缓,因此在本文实验中,权重衰减因子 $\lambda = 1 \times 10^{-8}$ 可以得到较好的识别结果。

6 本文方法与常用方法的比较

6.1 与 BPNN 及 SVM 识别方法的比较

管道泄漏故障诊断中较为常用的识别方法是传统时域信号采集基础上的特征提取,并通过反向传播神经网络(back-propagation neural network, BPNN)或支持向量机(support vector machine, SVM)对特征进行识别。本文所提出的基于压缩采集与深度学习的智能天然气管道泄漏孔径识别方法,与常用方法相比,从占用存储空间、诊断精度、运算复杂度上来说都有较大优势。为了说明其优势,下面将本文方法与采用 BPNN 识别以及基于小波包分解特征与 SVM 相结合的识别方法进行比较,图5所示为 10 次独立实验的结果。

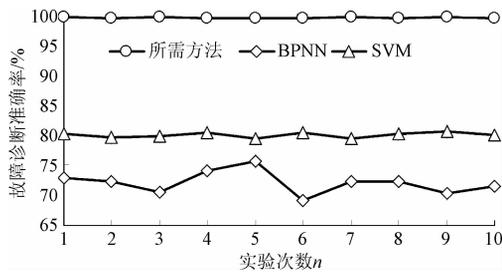


图5 本文方法与BPNN及SVM算法的识别结果比较

Fig. 5 Comparison of the proposed method with BPNN and SVM algorithm

从图5中可以看出,若采用BPNN识别,10次实验的诊断准确度集中在69.080%~74.004%,整体偏低,且单次识别准确度变化较大。基于“时域信号特征提取+SVM”识别的方法,10次实验的诊断准确率在79.372%~80.556%,变化幅度相比BPNN趋于平稳,诊断准确率也较BPNN高;而本文所提方法的10次实验诊断准确率范围为99.317%~99.681%,且每次识别准确率相差很小。计算10次实验的平均诊断准确率以及相应的标准差,如表1所示。

表1 十次实验的识别准确率统计

Table 1 Identification accuracy statistics of 10 experiments

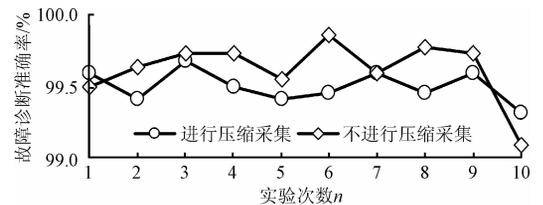
诊断方法	平均诊断准确率/%	准确率的标准差
所提方法	99.499	0.105
BPNN	72.106	1.802
特征提取+SVM	80.048	0.382

由图5及表1可以看到,BPNN方法识别准确度及稳定性都较差;基于SVM方法的诊断平均准确率为80.048%,标准差为0.382,好于BPNN方法,差于本文方法,说明基于SVM方法的识别结果较稳定,但是准确度还有待提高。相比之下,本文方法通过压缩采集、稀疏滤波自适应特征提取及识别,在建立深层次学习模型的基础上,能够快速高效地实现高精度的识别,较BPNN及SVM算法无论从单次识别准确率以及多次实验识别准确率波动方面,都具有明显优势。

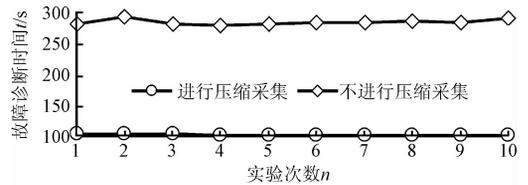
6.2 压缩采集在本文方法中的重要性

通过以上实验及分析,可以明显看到本文方法的诊断性能远远优于BPNN和基于小波包分解特征与SVM相结合的算法。在这一新的信息采集模式中,基于CS的压缩采集所起作用及进行压缩采集较没有进行压缩采样差别有待研究。以某泄漏信号为例,分别将未进行压缩采集与进行了压缩采集处理的数据输入

稀疏滤波网络进行特征提取及识别,图6所示为10次实验结果。



(a) 诊断准确率比较
(a) Comparison of diagnosis accuracy



(b) 诊断时间比较
(b) Comparison of diagnosis time

图6 信号进行压缩采样与不进行压缩采样的诊断效果比较

Fig. 6 Comparison of compressed sampling and un-compressed sampling

如图6(a)所示,进行压缩采集处理后泄漏孔径的平均诊断准确率和没有进行压缩采集处理的平均诊断准确率都可达到99.5%左右,说明压缩采集方法获取了绝大部分泄漏信息。从图6(b)中可以看到,压缩采集处理在大幅度降低泄漏信号数据量的同时(从每个样本的数据量为4096降低到了819,即 $CR=80\%$),也减少了诊断时间(从283s降低到103s),说明所提方法对泄漏孔径信息的提取是十分有效的;在泄漏检测过程中进行压缩采集处理也是非常必要的,它不仅能够大幅度减少数据量,也能够用更少的诊断时间实现高准确度的识别。

7 结论

针对天然气管道泄漏监测正在向大数据分析发展的趋势,结合压缩感知理论,实现了变换域的泄漏信号压缩采集;并在分析传统智能泄漏识别方法局限性的基础上,提出了一种通过深度学习获取压缩采集信号特征并识别的智能天然气管道泄漏孔径识别方法。本文方法中,泄漏信号经过压缩采集后,在少量观测数据中保留了原信号的大部分信息;基于深度学习的稀疏滤波网络能够直接从压缩感知域采集信号中自适应获取典型特征并通过softmax回归实现泄漏孔径的准确识别。实验结果表明所提方法的性能与传统方法相比更为优越。

参考文献

- [1] 孙洁娣, 肖启阳, 温江涛, 等. 局域均值分解分析的管道泄漏孔径识别及定位[J]. 仪器仪表学报, 2014, 35(12):2835-2842.
SUN J D, XIAO Q Y, WEN J T, et al. Pipeline leak aperture classification and location based on local mean decomposition analysis [J]. Chinese Journal of Scientific Instrument, 2014, 35(12): 2835-2842.
- [2] 王学伟, 苏丹, 袁洪芳, 等. 小波包多级树模型管道泄漏信号压缩感知方法[J]. 仪器仪表学报, 2014, 35(3):520-526.
WANG X W, SU D, YUAN H F, et al. Pipeline leakage signal compressed sensing based on wavelet packet hierarchical tree model [J]. Chinese Journal of Scientific Instrument, 2014, 35(3): 520-526.
- [3] 何炜琨, 郭双双, 王晓亮, 等. 基于奇异值分解的风电场杂波微动特征提取[J]. 电子测量与仪器学报, 2017, 31(4): 588-595.
HE W K, GUO SH SH, WANG X L, et al. Micro-motion features extraction of wind farm echoes based on singular value decomposition [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(4): 588-595.
- [4] 王军阵, 王建斌, 张轩硕. 超声导波管道缺陷检测的小波阈值去噪法[J]. 国外电子测量技术, 2010, 29(8):33-35.
WANG J ZH, WANG J B, ZHANG X SH. Ultrasonic guided wave pipeline defect detection using wavelet threshold selection method analysis [J]. Foreign Electronic Measurement Technology, 2010, 29(8):33-35.
- [5] YOU D, GAO X, KATAYAMA S. WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM [J]. IEEE Transactions on Industrial Electronics, 2015, 62(1):628-636.
- [6] 许桢英, 罗来齐, 王匀, 等. 基于PCA的管道缺陷导波信号特征优化方法[J]. 电子测量技术, 2016, 39(4):160-163.
XU ZH Y, LUO L Q, WANG J, et al. Feature optimization method of pipe defect guided signals based on PCA [J]. Electronic Measurement Technology, 2016, 39(4): 160-163.
- [7] SUN J, XIAO Q, WEN J, et al. Natural gas pipeline small leakage feature extraction and recognition based on LMD envelope spectrum entropy and SVM[J]. Measurement, 2014, 55(9):434-443.
- [8] 焦敬品, 李勇强, 吴斌, 等. 基于BP神经网络的管道泄漏声信号识别方法研究[J]. 仪器仪表学报, 2016, 37(11):2588-2596.
JIAO J P, LI Y Q, WU B, et al. Research on acoustic signal recognition method for pipeline leakage with BP neural network [J]. Chinese Journal of Scientific Instrument, 2016, 37(11):2588-2596.
- [9] CANDES E J. Compressive sampling [J]. Marta Sanz Solé, 2006, 17(2): 1433-1452.
- [10] SCHMIDHUBER J. Deep learning in neural networks: An overview [J]. Neural Networks, 2014, 61(10): 85-117.
- [11] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786):504.
- [12] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [13] 沈跃, 丁灵卫, 吴翔轩, 等. 回溯自适应匹配追踪电能质量信号重构方法[J]. 电子测量与仪器学报, 2017, 31(5):731-738.
SHEN Y, DING L W, WU X X, et al. Power quality signal reconstruction of backtracking-based self-adaptive matching pursuit method [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(5): 731-738.
- [14] RENNA F, CALDERBANK R, CARIN L, et al. Reconstruction of signals drawn from a gaussian mixture via noisy compressive measurements[J]. IEEE Transactions on Signal Processing, 2013, 62(9): 2265-2277.
- [15] TILLMANN A M, PFETSCH M E. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing [J]. IEEE Transactions on Information Theory, 2014, 60(2):1248-1259.
- [16] CANDES E J, TAO T. Decoding by linear programming [J]. IEEE Transactions on Information Theory, 2005, 51(12):4203-4215.
- [17] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21):49-56.
LEI Y G, JIA F, ZHOU X, et al. A deep learning-based method for machinery health monitoring with big data [J]. Journal of Mechanical Engineering, 2015, 51(21):49-56.
- [18] HINTON G E. A practical guide to training restricted boltzmann machines [C]. Momentum, 2012, 9(1): 599-

619.

- [19] NGIAM J, PANG W K, CHEN Z, et al. Sparse filtering [C]. IEEE International Conference on Neural Information Processing Systems, 2011:1125-1133.
- [20] RAJA K B, RAGHAVENDRA R, VEMURI V K, et al. Smartphone based visible iris recognition using deep sparse filtering [J]. Pattern Recognition Letters, 2015, 57(C):33-42.
- [21] LEI Y, JIA F, LIN J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5):3137-3147.
- [22] LIU W, LIU H, TAO D, et al. Multiview Hessian regularized logistic regression for action recognition[J]. Signal Processing, 2015, 110(5):101-107.

作者简介



孙洁娣, 1998 年于河北师范大学获得学士学位, 2001 年于燕山大学获得硕士学位, 2008 年于天津大学获得博士学位, 现为燕山大学副教授, 主要研究方向为智能故障诊断、压缩感知、深度学习理论及应用、复杂振动信号处理、管道泄漏监测及预警。

E-mail: wjtsjd@163.com

Sun Jiedi received her B. Sc degree from Hebei Normal University in 1998, M. Sc. degree from Yanshan University in 2001 and Ph. D. degree from Tianjin University in 2008. She is currently an associate professor in Yanshan University. Her main research interests include intelligent diagnosis, compressed sensing, deep learning and various applications, complex vibration signals processing and pipeline leak detection and pre-warning.



温江涛(通讯作者), 分别在 1998 年和 2003 年于燕山大学分别获得学士学位和硕士学位, 2009 年于清华大学获得博士学位, 现为燕山大学副教授, 主要研究方向为故障诊断、复杂振动信号处理及模式识别。

E-mail: wens2002@163.com

Wen Jiangtao (Corresponding author) received his B. Sc. and M. Sc. degrees both from Yanshan University in 1998 and 2003, respectively, and received his Ph. D. degree from Tsinghua University in 2009. He is currently an associate professor in Yanshan University. His main research interests include intelligent diagnosis, complex vibration signals processing and pattern recognition.