

学校代码: 10126

学号: 31656052

分类号: _____

编号: _____



内蒙古大学
INNER MONGOLIA UNIVERSITY

INNER MONGOLIA UNIVERSITY

硕士学位论文

MASTER DISSERTATION

(专业学位)

基于支持向量机的供水管道泄漏检测算法研究

学 院: 电子信息工程学院

专 业: 电子与通信工程

研究方向: 智能信息检测与处理

姓 名: 青春

指导教师: 刘洋 教授

学校代码: 10126
分类号: _____

学号: 31656052
编号: _____

论文题目

基于支持向量机的供水管道泄漏 检测算法研究

学 院: 电子信息工程学院

专 业: 电子与通信工程

研究方向: 智能信息检测与处理

姓 名: 青春

指导教师: 刘洋

2018 年 6 月 10 日

原创性声明

本人声明：所呈交的学位论文是本人在导师的指导下进行的研究工作及取得的研究成果。除本文已经注明引用的内容外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得内蒙古大学及其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：青春 指导教师签名：刘洋
日 期：2018.6.13 日 期：2018.6.13

在学期间研究成果使用承诺书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，即：内蒙古大学有权将学位论文的全部内容或部分保留并向国家有关机构、部门送交学位论文的复印件和磁盘，允许编入有关数据库进行检索，也可以采用影印、缩印或其他复制手段保存、汇编学位论文。为保护学院和导师的知识产权，作者在学期间取得的研究成果属于内蒙古大学。作者今后使用涉及在学期间主要研究内容或研究成果，须征得内蒙古大学就读期间导师的同意；若用于发表论文，版权单位必须署名为内蒙古大学方可投稿或公开发表。

学位论文作者签名：青春 指导教师签名：刘洋
日 期：2018.6.13 日 期：2018.6.13

基于支持向量机的供水管道泄漏检测研究

摘要

每年由于管道腐蚀和地质下沉等自然因素和城市发展建设的人为因素，城市供水管道不可避免的发生泄漏，如果能够及时的发现管道漏损及泄漏状况，将节约大量的水资源。因此供水管道泄漏检测在整个监测供水管网的系统的占据重要地位。管道泄漏检测的方法主要分基于声信号和非声信号两种，而在基于非声信号的主流方法中包括流量平衡法、红外热成像法、模型法等。由于受到管道埋设及要求精度高等限制，目前主流检测方法为基于声信号的检测。早期采用听音法，完全依赖工人的经验，而后来的研究者主要研究的内容有：1) 通过分析管道的振动频谱谱峰的位置可判断管道是否发生泄漏，然而此方法与泄漏信号频谱图相似的窄带噪声无法辨别。2) 将信号分解理论与人工神经网络等机器学习算法结合训练数据，用于判别泄漏信号，而此方法高度依赖数据量。

对此本文提出了专门针对小样本处理的基于机器学习的泄漏识别模型，具体内容如下：

(1) 分析了管道泄漏的主要形式及在发生泄漏时在泄漏口处引起管道振动的激励源，通过仿真分析了有无泄漏时的频谱差异以及管道材质对信号时域和频域的影响。

(2) 采用哈夫曼编码对管道信号在数据级做无损融合，有利于数据的传输。并对信号存在管道背景噪声情况提出 LMS 自适应抵消算法，对信号做滤波。

(3) 基于泄漏信号和非泄漏信号在特征上的差异性，对信号做经验模态分解后

提取固有模态函数的平均功率谱密度特征、原始信号近似熵特征以及代表信号的主成分。仿真分析基于信号单个特征在不同信噪比下的泄漏识别效果。再组合多个特征组成特征矩阵，用二分类支持向量机训练数据特征矩阵，将用已知信号测试训练好的网络，通过优化网络提高准确率。最后用优化好的网络去检测实际供水管道判别泄漏。

关键词：供水管道；特征提取；支持向量机；经验模态分解；近似熵；主成分分析

WATER SUPPLY PIPE LEAKAGE DETECTION BASED ON ACOUSTIC SIGNAL CHARACTERISTICS

ABSTRACT

Every year due to natural factors such as pipeline corrosion and geological subsidence and human factors in urban development and construction, urban water supply pipelines will inevitably leak. If pipeline leakage and leakage conditions can be discovered in time, a large amount of water resources will be saved. Therefore, water supply pipeline leak detection plays an important role in the entire system of monitoring the water supply network. Pipeline leakage detection methods are mainly based on two kinds of acoustic signals and non-acoustic signals, and mainstream methods based on non-acoustic signals include flow balance method, infrared thermal imaging method, and model method. Due to restrictions imposed by pipelines and high precision requirements, the current mainstream detection method is based on the detection of acoustic signals. Earlier adopting the listening method completely relied on the experience of the workers. The main research contents of the later researchers were: 1) The analysis of the position of the peak of the vibration spectrum spectrum of the pipeline can be used to determine whether the pipeline is leaking. However, this method works against the spectrum of the leakage signal. Similar narrowband noise cannot be discerned. 2) Combine training data with signal decomposition theory and machine learning algorithms such as artificial neural networks to discriminate leakage signals. This method is highly dependent on the amount of data.

This paper proposes a machine-learning-based leak identification model specifically for small sample processing. The specific content includes:

(1) Analyze the main form of pipeline leakage and the excitation source that causes pipeline vibration at the leakage outlet in the event of a leak. Analyze the frequency spectrum difference in the presence or absence of leakage and the effect of pipe material on the time domain and frequency domain of the signal.

(2) The use of Huffman coding for lossless fusion of pipeline signals at the data level is conducive to data transmission. The LMS adaptive cancellation algorithm is proposed to filter the background noise of the pipeline, and the signal is filtered.

(3) Based on the difference in characteristics of the leakage signal and the non-leakage signal, after the empirical mode decomposition of the signal, the average power spectral density characteristic of the intrinsic modal function, the approximate entropy characteristic of the original signal, and the principal component of the representative signal are extracted. Simulation analysis is based on the leakage recognition effect of single signal features at different signal-to-noise ratios. Then combine multiple feature composition feature matrices, use the binary support vector machine to train the data feature matrix, test the trained network with known signals, and improve the accuracy by optimizing the network. Finally, the optimized network is used to detect the actual water supply pipeline discriminating leakage.

KEYWORDS: water supply pipeline; feature extraction; Support Vector Machine; Empirical Mode Decomposition; Approximate Entropy; Principal Component Analysis

目 录

第一章 绪论.....	1
1.1 供水管道漏损现状及泄漏检测研究意义.....	1
1.2 供水管道泄漏检测研究现状.....	1
1.3 论文主要研究内容.....	3
第二章 供水管道泄漏信号分析.....	4
2.1 管道泄漏的形式.....	4
2.2 泄漏声信号的种类.....	5
2.3 管道泄漏段管内流场.....	6
2.4 漏水信号时频特性分析.....	6
2.5 小结.....	8
第三章 管道信号的数据融合及噪声自适应抑制.....	10
3.1 数据融合在管道泄漏检测中的应用.....	10
3.1.1 供水管道数据融合的意义.....	10
3.1.2 数据融合的分类.....	10
3.1.3 哈夫曼编码原理.....	11
3.1.4 基于哈夫曼编码的供水管道数据融合.....	13
3.2 自适应噪声抵消系统对泄漏信号的滤波.....	14
3.2.1 自适应噪声抵消技术的原理.....	14
3.2.2 LMS 自适应信号滤波.....	15
3.2.3 应用 LMS 滤波算法去除管道泄漏信号噪声.....	16
3.3 小结.....	18
第四章 基于支持向量机的管道泄漏检测方法.....	19
4.1 基于 EMD 平均功率谱密度的泄漏检测.....	19
4.1.1 经验模态分解算法原理.....	19
4.1.2 基于经验模态分解法的信号分解.....	21

4.1.3 基于 EMD 功率谱密度的泄漏检测实验.....	23
4.2 基于近似熵特征的泄漏检测.....	25
4.2.1 近似熵算法原理.....	25
4.2.2 基于近似熵的泄漏检测实验.....	26
4.3 基于主成分分析法的泄漏检测.....	27
4.3.1 主成分分析算法原理.....	27
4.3.2 基于主成分分析法的泄漏检测实验.....	29
4.4 基于信号多特征的泄漏检测系统构造.....	31
4.5 支持向量机理论.....	32
4.5.1 支持向量机分类原理.....	32
4.5.2 核函数理论与核函数选择.....	35
4.5.3 核函数参数选择.....	36
4.6 系统仿真结果分析.....	37
4.7 小结.....	42
第五章 总结与展望.....	43
5.1 总结.....	43
5.2 展望.....	43
参考文献.....	44
致谢.....	47

第一章 绪论

1.1 供水管道漏损现状及泄漏检测研究意义

由我们国家的人均水资源的占有率来看,我国是一个极具缺少水资源的国家世界,而人均占有率只有世界平均值的四分之一^[1]。然而平均每年水资源的漏损量达到 30%以上,在发达国家每年的水资源的漏损量控制在 20%以下。近年来国家城镇化的趋势下城市每年水的供应量也大大地提高,水资源的污染、社会水资源的分配不均匀也直接的导致了水的稀缺程度增加,显然成为了国家倡导的可持续发展道路的一种阻碍^[2]。在一些老城区,供水管道常年失修,铺设时间较远的管道长时间的腐蚀并造成爆管,而这些管道早已不符合国家规定的标准。随着城市基础设施建设的不断发展,为适应城市变化的步伐,供水方式也在不断地改进,一个城市不同在不同发展时期采用的管道材质、管道结构,导致供水网络结构错综复杂,承载压力标准各不一,增加了管道爆裂的概率。

供水管道泄漏浪费的是宝贵的水资源,自然导致了净水的成本。现如今节约用水是国家提倡并且成为了全民共识,根据各个国家对供水漏损率检测得到如表 1.1 所示的漏损状况。发达国家的管道漏损导致水的损失量远远低于发展中国家。

表 1.1 各个国家和地区的供水损失量

Fig1.1 Water loss rate in various countries and regions

发达国家	8-20%
发展中国家	20-40%

我国对供水管道漏损控制的起步相比于发达国家晚,供水损失量为 20-40%。经常性的因爆管或断裂导致停水,很大程度的影响了人们的生活质量。所以漏损控制问题得到国家相当的重视,相继出台了相关政策,推动城市老旧管网的改造,降低供水损失量,保证城市居民的供水完全,并鼓励建立供水管网检测系统。

1.2 供水管道泄漏检测研究现状

目前国内外较普遍认可并应用在供水管道的泄漏检测模型依据声学分为基于声信号的检

测和基于非声信号的检测。而基于非声信号的检测主要包括:

(1) 流量平衡法^[3]: 当供水管道处于正常运行状态下时, 管道上下游的流速和流量保持稳定, 两端的差值也会保持在一定的范围之内。而当管道发生泄漏时, 管道下游的流量明显减少, 导致两端的相差数值将明显增大, 当超过原先设定好的阈值时自动触发泄漏报警, 从而判别管道泄漏的发生。

(2) 红外热成像法^[4]: 此方法为通过采集红外热成像仪接受埋地管道表面产生而向外辐射的红外线, 采用相应的技术转化为热力场的分布图, 通过观测温度的差距来确定管道泄漏状态。红外热成像技术可以分为主动监测和被动监测, 其中被动监测为通过光敏元件接受管道自身发射的红外线来分析管道是否处于正常状态, 而主动监测为对管道进行局部加热, 使得加热部位失去热平衡, 当管道的加热部位热场分布图不均匀时可判别为管道发生泄漏。

(3) 实时模型法^[5]: 该方法是通过由计算机建立多个方程组实时的模拟管道, 并且使这模型与实际中的供水管道同步运行。定时或者实时的选取管道的运行状态并与模型相比较分析, 当计算得到的值与模型相差超过拟定的门限时触发泄漏报警。而在模型中取得的状态量包括管道中流体的流量、压力等实时数据。模型中采用的状态方程包括流体状态方程、能量平衡状态方程、动量平衡状态方程等, 这需要在管道的各个位置安装控制系统来实时取得流量和压力, 所以用于检测的点越多越好。

在实际的管道泄漏检测中最重要的是及时发现泄漏故障并做维护。而基于非声信号的泄漏检测在实际应用中由于受到检测区域的土质和地面的影响存在管道检测耗时的问题。又比如当管道内的水流波动幅度较大时流量平衡法容易导致误报。一般情况下供水管道是深埋于地下的, 如果通过红外热成像法来检测的话, 被动检测时很难感应到管道的温度分布, 而主动检测也很难通过厚厚的土层对管道做局部加热。在复杂的供水管网上通过实时模型法建立模型难度极大, 并且随着采集到的数据量的增多导致运算量极具变大。所以目前由于受到各种条件的限制及算法发展的限制应用于供水管道泄漏的主流算法为基于声音信号的泄漏识别。国内外对依赖声音信号分析管道漏损方面做了很多的研究工作^[6-9]。

在早期基于声信号的泄漏检测中检漏人员是通过使用听音棒等设备来监听流水声, 而这很大程度上依赖于工人的经验判断。Wan 等人(1993)对管道信号做自适应滤波处理, 通过查看滤波后的频谱波形图来判断管道的泄漏状态^[10]。李光海等人(2002)基于声发射特征, 通过采用功率谱密度与频率之间存在的泄漏信号特有的的幂指数关系来判别泄漏, 然而随着泄漏形式和管道材质的不同泄漏信号的频谱也会发生相应的变化, 所以很难获取各种管道在不同流体

状态下的泄漏频谱先验知识^[11]。唐秀家(1997)认为泄漏信号的产生是流体湍流与管道相互作用的结果,基于此采用了使用信号的幅度和平均值作为特征结合神经网络辨识模型的检测方法,对泄漏与非泄漏状态做区别,但是当官内外固定干扰声源时,将会导致误判^[12]。基于此,为了区别官内外固定干扰声源和泄漏信号,Ai 等人(2006)引用隐马尔可夫模型(HMM)和线性预测编码倒谱系数(LPCC)作为特征来检测判别信号^[13]。Toshitaka 等人(2007)对管道信号提取功率谱密度特征和 AR 模型的破坏参数特征并用 SVM 来判别信号的来源是属于泄漏管道还是非泄漏管道^[14]。文玉梅等人(2014)对管道信号提取了谱宽参数特征和近似熵特征结合 SVM 来判断管道的泄漏与非泄漏^[15]。

1.3 论文主要研究内容

在本论文中针对目前基于声音信号处理的供水管道泄漏检测的主要发展方向,结合实际的地下管道复杂的情况,重点分析了泄漏信号的产生条件,泄漏信号的特征提取与泄漏检测,以及考虑影响信号分析的多方面的因素。

- 1) 结合目前国内外对于供水管网泄漏的实际情况阐述了研究管道泄漏辨识的意义和目的以及对供水管道泄漏的研究进展。
- 2) 介绍管道泄漏的形式,泄漏导致产生的各种复杂声音以及在泄漏处管道内部水流的流场,分析管道有无泄漏时时频特性,管道材质等因素影响泄漏声的时域、频域特性,以及噪声信号种类和来源。
- 3) 分析数据融合技术在检测管道泄漏方面的应用并提出管道信号的融合方法。采用自适应噪声抵消算法滤除管道背景噪声,在一定程度上提高了漏水信号的信噪比。
- 4) 从泄漏信号的产生模式分析泄漏声信号所独有的特征信息,分别对泄漏信号、管道背景噪声、管外固定噪声(例如阀门噪声,工地施工噪声耦合到管道上的振动信号)等信号做特征提取,并做基于单个特征的泄漏检测识别诊断。
- 5) 基于支持向量机的管道泄漏辨识。详细阐述了支持向量机的概念、分类原理、主要参数及特点。分析在不同的核函数和网络参数,通过寻优算法将检测模型优化到最适合处理漏水信号的状态,并且分析模型在不同信噪比环境下检测判别准确率。

第二章 供水管道泄漏信号分析

由于各种原因导致管道破裂,从而导致管内流水改变原来的运动状态从漏口向外喷射,形成泄漏声信号的同时将引起管道振动。正是对泄漏声信号的处理被广泛应用到实际的泄漏检测中。然而了解供水管道泄漏声的产生机理是检测的前提。国内外研究者通过在不同采集环境下对不同管道材质、管道尺寸、管道埋设条件、泄漏形式做了大量的仿真分析,得到了泄漏信号的时域和频域特征。基于此本章主要分析了管道的泄漏形式、泄漏信号的种类以及管道泄漏段的管内流场,并通过仿真分析了有无泄漏时的时频特征以及不同材质对泄漏信号的时域和频域特性的影响,为后续的处理做铺垫。

2.1 管道泄漏的形式

在上世纪七十年代及以前,国内供水管道采用的主流材质为钢管,混凝土管和铸铁管等。然而到了 80 年代,随着国家对基础设施建设的投入导致原有的管道材料不再满足需求,塑料管、复合型管道在供水管道中广泛应用。地下供水管网铺设错综复杂,而这些管道在地下由于季节变化、地质变化、管内外压力和人为因素,管道将不可避免的引发泄漏,其主要形式有^[16-18]:

- 1) 管道断裂泄漏:在管道铺设施工时未能充分考虑到地面对管道的压力而导致管道受力不均匀,管道局部受力过大将造成管身被压断或铸铁管破裂,从而发生大面积泄漏,无法运输自来水。
- 2) 管道间接口及焊接口泄漏:地下供水管道是将多个管道连接而成的,特别是在管道拐弯处必然存在接口。用水泥和麻绳密封的接口部分未拧紧,石棉水泥封口处理不严实,在外力冲击下造成接口处损坏。钢管焊接时,焊接处未焊透、管道接口未对正、多个管道汇聚于一点都容易造成焊接处开裂,从而引发泄漏。
- 3) 管道小孔泄漏:管道深埋地下多年后难免会受到地下环境的影响造成局部腐蚀,土壤的酸碱性 and 杂散电流要比空气对管道的腐蚀速度要快,这容易导致管道局部穿孔而造成小孔泄漏。

2.2 泄漏声信号的种类

由于供水管道内存在高压水流，管内和管外压力差很大，而随着供水管道的老化、腐蚀等影响地下管道发生泄漏时，管内高压水向外喷射，而这喷射的水与周围的土层和埋设介质相互作用会引发各类不同频率的信号，而将此类信号大致可以分为以下三种：

- 1) 漏口处泄漏声：当发生泄漏时管道中正常运行的水流发生变化，形成流体紊乱，并且跟管道相互作用导致管道振动，因为实际中采用的输送管一般情况下采用弹性管，所以容易引发共振。而这类信号随着管道铺设路径传播，在一定的传播距离内通过对管道的暴露点监听得到，比如在阀门和消防栓的位置^[14]。
- 2) 水头撞击声：这类声音表示的是喷出的水跟与管道一起埋设用来保护管道的介质发生碰撞产生的。而此类信号的频率成分较低，保持在 20~300 Hz 之间，传播方向为地下土壤向地面传播，所以可以选择用听漏仪测听^[19]。
- 3) 介质摩擦声：这是喷出的水流跟周围土壤相互作用而产生的信号，当流出的水流达到一定程度后会形成泥浆，导致周围土壤中的小石子来回移动。而此类信号的频率成分照样很低，还是在 20~300 Hz 之间^[19]。

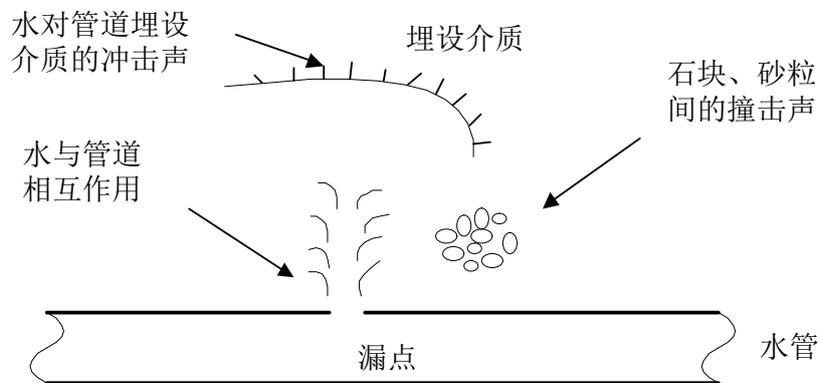


图 2.1 供水管道泄漏处主要声源

Fig2.1 Main source of leakage in water supply pipelines

在实际的应用检测时，用来采集的传感器大多数为吸附在管壁上，从上面的分析可以得知介质摩擦声和水头撞击声的传播介质为地面，很少通过管壁传播，而相反漏口处泄漏声主要是通过管壁传播，所以放置在管壁上的传感器采集得到信号为这类的泄漏信号，而且此类信号有更能表达泄漏的特征的部分。基于此本文论主要以此类信号作为研究对象，而该漏口处泄漏声是喷射出的流体与管道共同作用的结果，所以分析该类信号的特征时应该分析因为泄

漏导致引发的管内流场的变化

2.3 管道泄漏段管内流场

在供水管道泄漏检测研究中基于声信号的方法普遍处于热门的地位，而很多研究者对泄漏声信号的怎样产生做了很多的研究。唐秀家(1997)对一个长输管道建立了管道振动模型后发现当发生泄漏时在泄漏口处流体状态方程发生改变形成湍流，而向外喷射的流体使管内流体发生紊乱并且跟管道发生共振现象^[12]。而 Thompson 等人(2001)进一步研究后提出泄漏信号中不仅包括湍流声，还夹带着一些少部分的气穴声^[20]。

当管道出现漏口时必然会有一部分水流向管外以一定的速度喷射出，且喷射的速度较高，这会导致在漏口处水流分离，而靠近漏口一侧的水流会改变原来的流水轨道。在泄漏口的位置漏口与管壁形成一个直角锐缘处，这会使流体分离时在分离处存在一个小型的低压区域，低压区的压力变化慢慢向管道两侧扩散，当扩散到一定远距离时压力差保持在稳定状态。又因为水的饱和蒸汽压力相比于空气分离压力低很多，所以当发生泄漏时局部空气分离压力会低于水的蒸汽压力，这时管道内水中包含的气核会分离出来形成空泡^[21]。而形成的空泡一部分随着水流沿着管道下游传播，随着传播距离的增大管内压力重新回到原来的状态，这时空泡会发生爆破引起气穴声，而另一部分直接随着水流在泄漏口处流向管道外侧。

在泄漏口处分别有左右两个直角锐缘处，当往外喷射的流体速度较高时在这两个锐缘处分别产生高强度的湍流。湍流状态也会沿着管道两侧传播，尤其是下游区域更加明显，而随着距离的增大流体又从湍流状态变为层流状态^[22]。因此当管道发生泄漏时靠近泄漏口一侧的管内流体的运行状态从层流变为湍流，从而引发湍流声，而此类声音耦合到管道上引起管道局部振动并向管道两侧传播^[23]。同时向外喷射的高速水流与管道泄漏口相互摩擦也将导致管道振动。通过上面的分析可以发现，泄漏导致管道振动的激励源包括以下三个部分：空泡声、管内流体湍流声以及高速喷射的水流与泄漏口相互摩擦的作用。

2.4 漏水信号时频特性分析

对不同的泄漏环境和不同的管道环境下通过大量的仿真实验分析泄漏信号的时域和频域特性，采用高采样率环境下的采集卡和振动传感器对管道信号进行采集，并参考了多个文献发现金属管道的泄漏信号的主要频段集中在 1KHz~2KHz 之间，所以选择 5000Hz 的采样率对

管道振动信号及声信号采集。此时可分析的频率范围为[0~2500Hz]。本论文采用的采集卡为NI公司的MCC BTH-1208LS数据采集卡，与之匹配的是CT1010加速度传感器，此传感器为在机械领域广泛应用的，是具有高灵敏度的传感器。采集卡用来将数据保存到上位机端，用于分析信号。图2.2为采集简易的硬件系统。

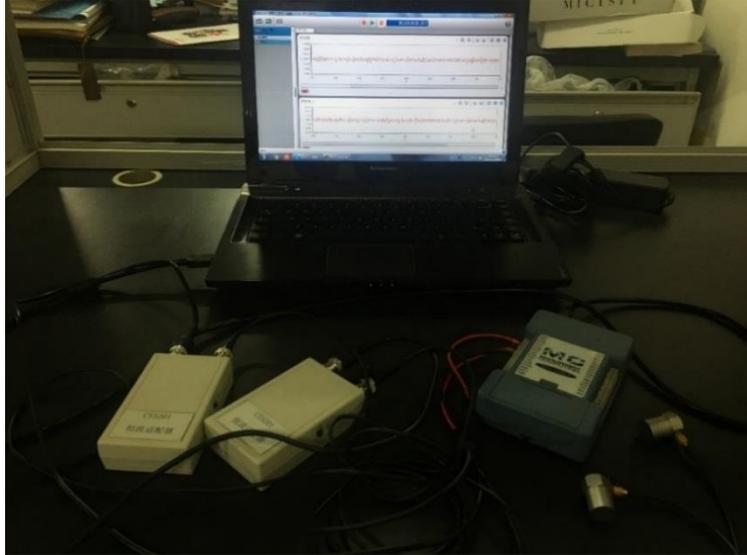


图 2.2 漏水信号采集系统

Fig2.2 Leakage signal acquisition system

以水龙头作为泄漏声源，调节流量，将传感器放置在距离水龙头 20cm 远处位置，将采集得到的管道振动信号认为泄漏导致产生的管道振动信号。首先调节流量采集了泄漏信号，图 2.3-2.4 分别是背景噪声和管道泄漏时的时域和频域图，通过比较可以看出：

- 1) 管道背景噪声能量不集中，频率几乎均匀分布在整个频带，而不是集中在某个频段内；
- 2) 泄漏信号的主要频率成分还是保持不变，主要频率成分集中在 1.6kHz~ 1.8kHz 之间，以及少部分的低频频段。

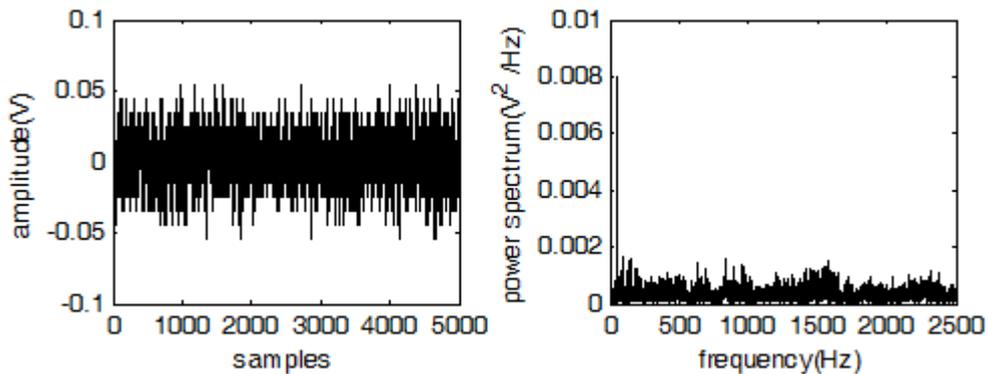


图 2.3 管道背景信号时域波形和功率谱

Fig2.3 Waveform and power spectrum of pipe background signal

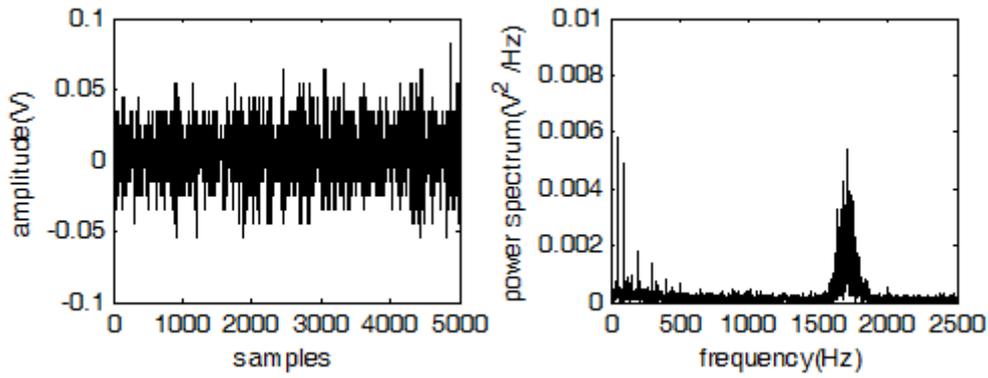


图 2.4 管道泄漏信号时域波形和功率谱

Fig.2.4 Waveform and power spectrum of pipe leakage signal

下面分析 PVC 管道泄漏时的频谱，分析管道振动引起的泄漏信号，得到的泄漏信号时域和频域波形如图 2.5 所示。

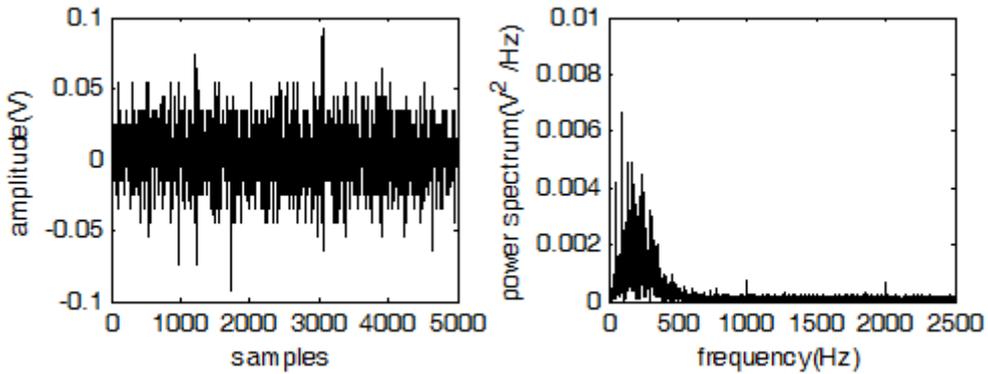


图 2.5 PVC 管泄漏信号时域波形和功率谱

Fig.2.5 Waveform and power spectrum of PVC pipe leakage signal

通过在实验环境中仿真分析自来水 PVC 管道的频率分布可以发现 PVC 管道上的泄漏信号频率成分主要集中在 400Hz 以下，而时域波形没有很明显的差距。将金属管道和 PVC 管道的泄漏信号的频谱比较可得出，两个信号的主要频率不在一个频段上，金属管道的泄漏信号主要频率成分较高于 PVC 管道的泄漏信号频谱。

2.5 小结

介绍了日常供水管道破裂导致泄漏的几种形式，分析了管道泄漏段的管内流场，认为空泡声、湍流声、在高压下泄漏口管道与水流的摩擦声引起管道振动的信号为泄漏处的激励源。通过仿真实验结论得知当管道没有发生泄漏时管道背景噪声的频谱分布在整个频带上，发现

了管道材质的变化对管道泄漏信号时域和频域之间的影响，PVC 管道的泄漏信号主要频率部分较金属管道要低。

第三章 管道信号的数据融合及噪声自适应抑制

3.1 数据融合在管道泄漏检测中的应用

3.1.1 供水管道数据融合的意义

在供水管道管壁上安装的单个传感器所获得的信息往往是片面的或结果达不到需求，而多个传感器之间收到的信息又会出现冗余，会导致传输数据是网络堵塞，每个传感器都具有不同的特征，所以基于这些问题多传感器数据融合技术应运而生，国内外很多的研究者都对多传感器数据融合具有不同的解释，大致上可以概括为：充分利用不同时间或同一时间上的、确定的或模糊的、相关的或互补的，也可能是相互矛盾的多传感器信息采用算法技术，在一定准则下加以分析、剔除冗余、增加判断、估计和检测的准确性和可靠性，消除单一传感器的局限性，使系统的性能更加稳定。

3.1.2 数据融合的分类

无线传感器的数据融合算法依据融合的层次大致可分为一下三种^[24]：数据级融合、特征级融合和决策级融合。

(1) 数据级融合：将各个传感器（包括同类或者异类传感器）采集到的不同时间段或同一时间段采集到的管道信号数据进行融合处理，为最低层次的融合。这种融合的主要特点是尽可能保证原始数据信息的前提下减少网络的传输数据量。其融合流程如图 3.1 所示：

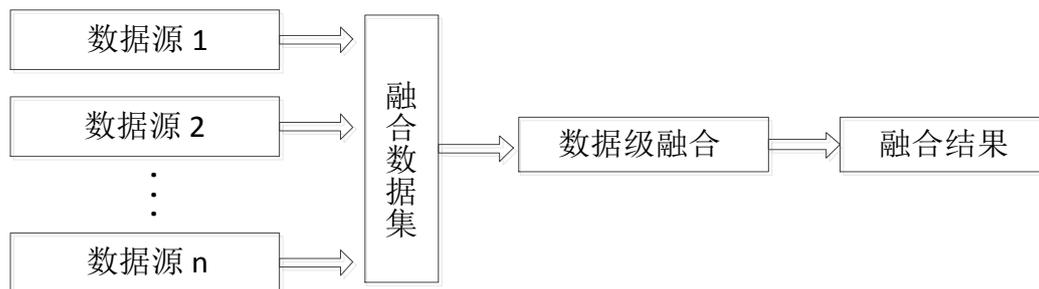


图 3.1 数据级融合流程图

Fig3.1 Data level fusion flowchart

(2) 特征级融合：这是在中间层次的融合，将对每个传感器采集到的原始管道信号进行特征提取后，对提取的特征做二次特征提取将冗余的特征剔除，进行分类和综合。提取的特征必

须能够充分表达原始信号部分。在这个层次的融合主要特点是实现信息压缩,实时处理信号。其融合流程如图 3.2 所示。

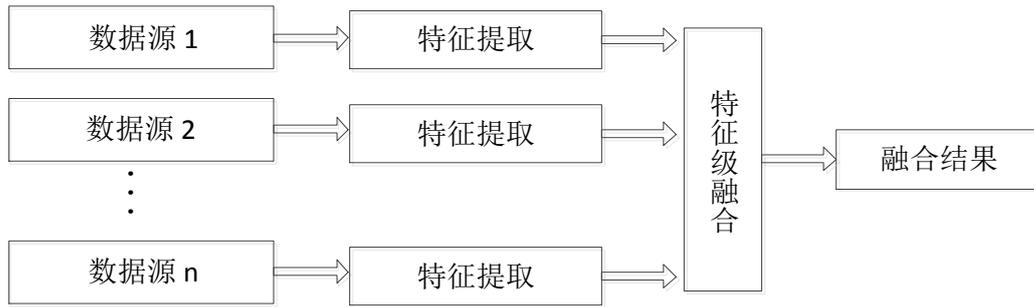


图 3.2 特征级融合流程图

Fig3.2 Feature level fusion flowchart

(3) 决策级融合: 这是处于最高层次的融合, 对每个传感器采集到的管道信号单个做泄漏判别, 对判别后得到的结果做融合。对每个判别结果做决策、控制, 得到最终的判断和检测结果。其融合流程如图 3.3 所示。

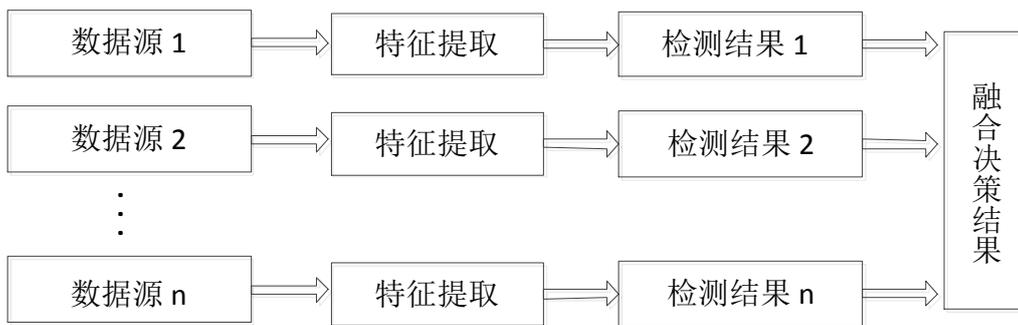


图 3.3 决策级融合流程图

Fig3.3 Decision-level fusion flowchart

无线传感器网络的数据融合技术如果从融合前后数据的信息的缺损程度划分, 可以分为有损融合和无损融合:

(1) 有损融合: 有损失融合使用的方法为剔除一些细节信息和降低信号的冗余部分, 从而达到减少数据量存储的目的, 间接的节约了存储空间和节点的能量损耗。但是信息损失部分不可以是数据处理的必要部分。

(2) 无损融合: 相比与有损融合, 所有的信息均被保留, 没有任何损失。是将多个信号打包成一组数据信号, 但不改变组内所携带的所有信息内容。一般情况下是剔除一组数据中的冗余信息, 但是数据整体减少的程度会受到其嫡值的上限。

3.1.3 哈夫曼编码原理

Huffman (1952)提出的哈夫曼编码 (Huffman Coding) 是一种编码算法。算法原理是根据

信号中各种数值点的值出现的概率进行编码运算，其中出现概率越高的数据点的值的码字越短，而出现概率越小的数据点的值，其对应的码字长度越长，从而减少整体码字的长度^[25]。

哈夫曼编码的实现步骤如下：

- (1)首先统计信号中各数据点的值出现的概率；
- (2)将信号中各个数据点的值出现的概率按照从大到小的顺序排列；
- (3)将两个概率最小的数据点的值进行相加，得到相加后的概率；将这个新的概率值与剩余的概率比较并进行步骤(2)和(3)的计算，直到最后的概率值为 1.0；
- (4)在计算过程中概率大的数据点的值标为 1，较少的则标为 0；
- (5)记录由每个数据点的值的概率到达 1.0 处的路径上的 1 和 0；
- (6)对每个数据点的值，从右到左记录 1、0 序列，至此得到了每个数据点的哈夫曼码。

编码实现过程通过举例来说明，设在一组数据中包含 A, B, C, D 和 E 五个数据值，出现的概率分别为 0.40、0.18、0.15、0.15 和 0.12。

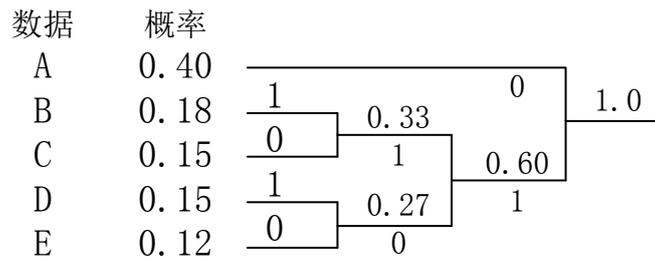


图 3.4 哈夫曼编码流程图

Fig3.4 Huffman coding flowchart

图 3.4 中显示的数据点的值 Huffman 码如表 3-1 所示，数据值 A 出现的概率最大，相应的代码最短，而其它出现概率较小的数据值的 Huffman 码相对较长。这样可以使数据的码长整体变少，达到压缩的目的。

表 3-1 数据的 Huffman 表

Table3-1 Huffman table of data

信源数据	码字	码长
A	0	1
B	111	3
C	110	3
D	101	3
E	100	3

3.1.4 基于哈夫曼编码的供水管道数据融合

本实验中结合属于无损融合领域的哈夫曼编码技术和数据融合的低层融合技术——数据级融合对管道振动信号做分析。在实际采集时发现管道振动信号时一个微弱的信号，一般时域幅值在-0.2v 到 0.2v 之间，应用哈夫曼编码数据融合时先将数据变成无符号 8 位二进制的数，即区间在[0~255]的十进制数。如果将原始数据矩阵做编码预处理的话将变成零数据，所以通过变换将数据投影到[0~255]上。变换公式为：

$$B = (A - A_{\min}) \times \frac{255}{(A_{\max} - A_{\min})} \quad (3-1)$$

式中 A 为原始数据矩阵，可以是多个传感器信号在同一个时间点采集到的信号或单个传感器在不同时间点采集到的信号。 A_{\min} 和 A_{\max} 为矩阵内元素的最小值和最大值，B 为变换后的数据矩阵。矩阵 B 为哈夫曼编码处理信号，对数据做压缩，传到上位机端，在上位机做反变换解码得到数据编码前的数据 B。

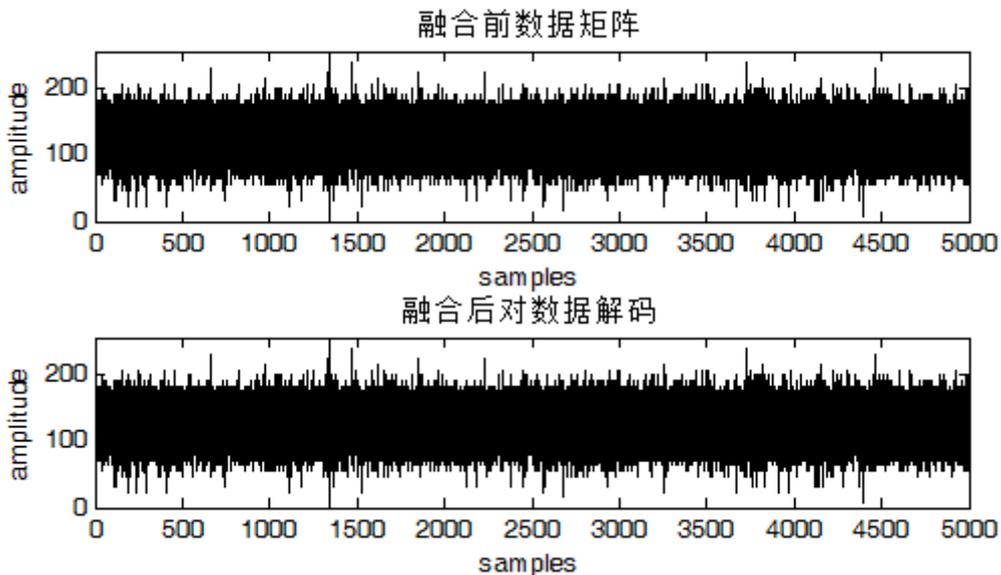


图 3.5 融合前和融合后解码的数据图

Fig3.5 Pre-encoding and decoded data diagram

在本次仿真中，将采用 50 组信号，每组有 5000 个数据点的原始数据矩阵，投影到[0~255]区间上的 50*5000 uint8 型矩阵，哈夫曼编码后成为 1*102620 uint8 型数据矩阵，通过仿真分析得到压缩比为 2.3290。图 3.5 为数据融合前和融合后解码信号的数据矩阵，可以看到数据一样，即数据融合为无损融合。当对管道信号提取的特征过多或者有冗余信息时可以选择采用特征级融合，对特征数据做二次特征提取。对依靠管道的多种参数做泄漏判别时，例如官

道上采用加速度传感器采集振动信号、温度传感器采集局部温度信号、压力传感器采集管道压力信号，依赖各个参数做泄漏判别后，按照对传感器和算法的信任度分配概率，依据概率分配做融合进而做最终的判断，即为决策级融合。

3.2 自适应噪声抵消系统对泄漏信号的滤波

在实际的管道信号采集时或多或少会混入噪声，包括高斯分布的随机噪声，突发的强干扰噪声以及管道背景流水声，这些噪声与泄漏信号混合重叠在一起，很大程度上影响着信号的有效分析，所以对漏水信号做滤波或者去除噪声是很有必要的。在本节首先介绍自适应噪声抵消技术的原理及 LMS 自适应滤波算法原理，并对含有噪声的管道泄漏信号做滤波处理，证明此算法可用于对泄漏信号做滤波。

3.2.1 自适应噪声抵消技术的原理

自适应噪声抵消系统是一种能够改变自身权值而达到最优状态的滤波器，是通过自适应的更新滤波器系数来实现的算法。自适应噪声抵消系统一般是由两部分组成，一个为滤波器的内部结构，另一个为调节更新滤波器系数的自适应算法。此系统不需要知道输入信号的先验知识，有计算量小和能实时处理信号的特点^[26]。图 3.6 为系统原理：

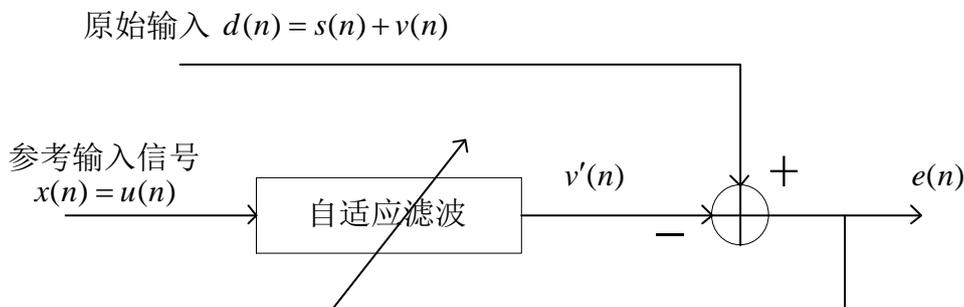


图 3.6 自适应噪声抵消原理图

Fig3.6 Principle diagram of adaptive noise offset

系统的输入包括一个原始输入和一个参考输入。一般情况下原始输入 $d(n)$ 为有用信号 $s(n)$ 和噪声信号 $v(n)$ 的叠加，参考输入是与有用信号 $s(n)$ 不相关，而与噪声信号 $v(n)$ 具有相关性的噪声 $u(n)$ 。如图 3.6 所示，对参考噪声 $u(n)$ 做滤波可以得到自适应滤波器的输出 $v'(n)$ ，而整个系统的输出为：

$$e(n) = s(n) + v(n) - v'(n) \quad (3-2)$$

即输出结果为有用信号的相似信号。对式 (3-2) 两边平方得:

$$e^2(n) = s^2(n) + [v(n) - v'(n)]^2 + 2s(n)[v(n) - v'(n)] \quad (3-3)$$

对式 (3-3) 两边求得期望值, 并简化得:

$$E[e^2(n)] = E[s^2(n)] + E\{[v(n) - v'(n)]\}^2 \quad (3-4)$$

滤波器系数时有用信号的功率 $E[s^2(n)]$ 无关, 则不会随着系数的更新而改变。所以要使

$E[e^2(n)]$ 最小, 应该使 $E[(v(n) - v'(n))^2]$ 最小, 由式 (3-2) 可得:

$$e(n) - s(n) = v(n) - v'(n) \quad (3-5)$$

由式 (3-5) 可得, 当 $E[(v(n) - v'(n))^2]$ 最小时相应的 $E[(e(n) - s(n))^2]$ 也最小。就是当有用信号 $s(n)$ 的均方差跟整个自适应噪声抵消系统的输出信号 $e(n)$ 的均方差越接近时最小。换句话说在理想的状态是 $v'(n) = v(n)$, 则 $e(n) = s(n)$ 。就是自适应滤波器自动更新自身的脉冲, 将 $u(n)$ 通过调节变成 $v(n)$, 将原始输入 $d(n)$ 中的 $v(n)$ 滤除掉, 使得输出信号 $e(n)$ 中只有有用信号 $s(n)$, 而噪声部分完全被抵消掉。

3.2.2 LMS 自适应信号滤波

在上一小节中提到自适应噪声抵消系统的剩下的另一个组成部分为自适应滤波算法, 用来调节信号的滤波器系数达到误差最小化。这里采用最小均方误差准则, 对信号做滤波处理。最小均方误差(LMS)算法是将滤波器系统的总输出信号与参考输入信号, 即期望输出信号, 之间的均方误差最小^[27]。最小均方误差算法是一种较为简单并且便捷的梯度最速下降算法。假设自适应滤波器的误差为:

$$e(n) = d(n) - y(n) = d(n) - x^T(n)w(n) = d(n) - w^T(n)x(n) \quad (3-6)$$

式中 $x(n)$ 为输入信号矢量。 $y(n) = x^T(n)w(n) = w^T(n)x(n)$ 为自适应滤波器的输出信号。而现在找到一种方法, 即将最速下降法中理想状态下的梯度真值用均方误差 $MSE = E[e^2(n)]$ 的梯度值的估计结果来代替。用最小均方误差算法来计算梯度估计的算法是将用以误差信号的每一次迭代的瞬时平方值代替它的均方值, 并以此来估算误差函数 $\xi = E[e^2(n)]$ 的梯度 $\nabla(n)$ 为:

$$\hat{\nabla}(n) = \left[\frac{\partial e^2(n)}{\partial w_0(n)} \frac{\partial e^2(n)}{\partial w_1(n)} \cdots \frac{\partial e^2(n)}{\partial w_m(n)} \right]^T \quad (3-7)$$

若写成矢量形式，有：

$$\hat{\nabla}(n) = \frac{\partial e^2(n)}{\partial w(n)} \quad (3-8)$$

将式 (3-2) 和式 (3-6) 带入式 (3-8) 可得：

$$\hat{\nabla}(n) = 2e(n) \frac{\partial e(n)}{\partial w(n)} = -2e(n)x(n) \quad (3-9)$$

最速下降法中的梯度值有：

$$w(n+1) = w(n) + \mu[-\hat{\nabla}(n)] = w(n) + 2\mu e(n)x(n) \quad (3-10)$$

式中 $W(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T$ 表示为在 n 时刻的权系数矢量，参考输入信号的矢量定义为 $X(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ ， N 为自适应滤波器长度， $y(n)$ 为期望输出， $d(n)$ 为主输入信号， $e(n)$ 为整个抵消系统的输出值，即误差信号。 n 为时间序列， μ 为收敛因子。最小均方误差算法中收敛因子符合 $0 < \mu < \lambda_{\max}$ 时才保证收敛，其中 λ_{\max} 是对输入原始信号自相关矩阵求得的最大特征值。概括计算步骤为^[28]：

- (1) 设置初始条件： $w(0) = 0$ 。
- (2) 对每个时间点 $n = 1, 2, \dots$ 都做迭代运算。

$$\text{滤波器输出: } y(n) = w^T(n)x(n)$$

$$\text{误差信号: } e(n) = d(n) - y(n)$$

$$\text{权向量调整: } w(n+1) = w(n) + 2\mu e(n)x(n)$$

式中 μ 为收敛的步长因子。为了保证自适应处理过程的稳定性， μ 必须符合 $0 < \mu < NPin$ ，其中 $Pin = E[X^2(n)]$ 为输入信号功率。

3.2.3 应用 LMS 滤波算法去除管道泄漏信号噪声

在实验环境下夜间采集管道泄漏信号，将环境噪声和管道背景噪声降低到最低，几乎接近理想状态下采集到的泄漏信号作为自适应噪声抵消系统的期望输出信号。图 3.7 为采集到的

期望信号：在噪声环境下采集管道泄漏作为待滤波信号，输入到自适应噪声抵消系统中，滤波器层数设为 20 层。图 3.8 为待滤波信号，设置信噪比为 2dB。滤波输出信号为图 3.9 所示。

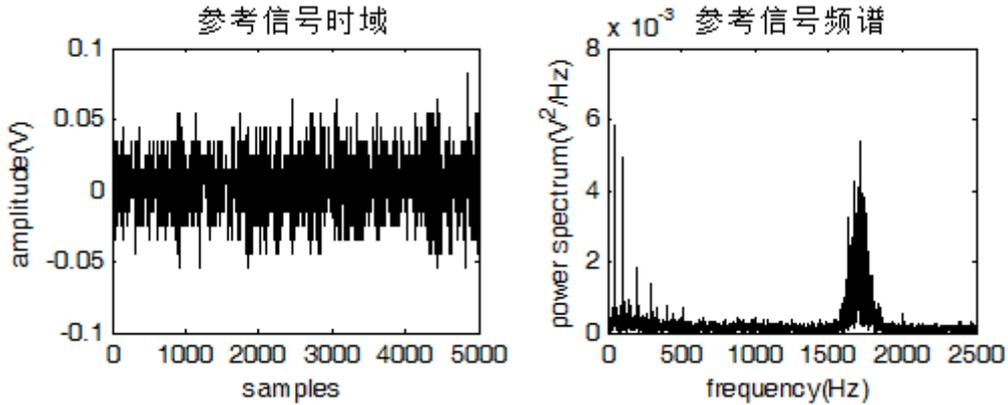


图 3.7 纯净参考信号时域和频域图

Fig3.7 Waveform and power spectra of the pure reference signal

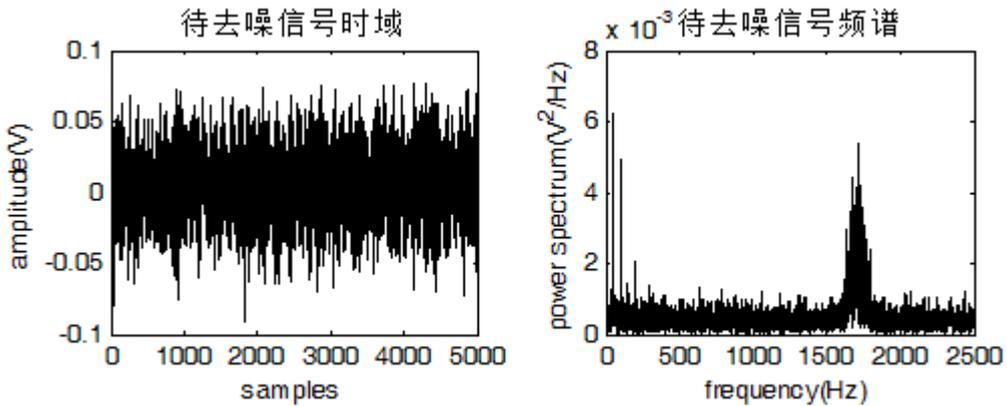


图 3.8 待去噪信号时域和频域图

Fig3.8 Waveform and power spectra of the undenoised signal

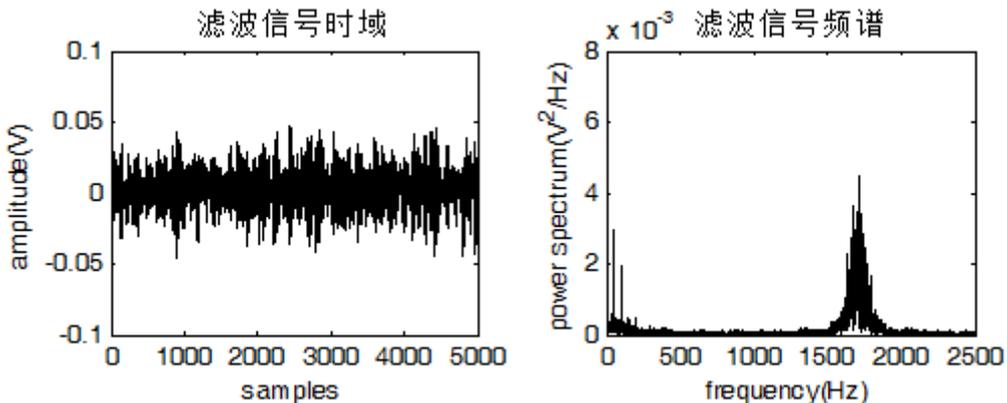


图 3.9 滤波后信号的时域和频域图

Fig3.9 Waveform and power spectra of the after filtered signal

通过仿真分析可以看到背景噪声被抵消掉很大部分，达到了滤波效果，表明采用 LMS 误

差准则对管道泄漏信号做自适应滤波处理可以很好的滤除背景干扰噪声的影响。

3.3 小结

在本章首先介绍了数据融合的基本概念和详细分类内容以及哈夫曼编码的算法原理。通过仿真处理对管道数据用哈夫曼编码技术做无损融合，在数据处理前又做反变换得到原始数据。接下来介绍了 LMS 自适应噪声抵消算法的原理以及算法步骤，最后对管道信号做处理，有效抑制了管道的背景宽带噪声，得到较为纯净的泄漏信号。

第四章 基于支持向量机的管道泄漏检测方法

在第二章中分析了管道泄漏形式、漏水信号的产生原理以及泄漏导致管道振动的漏水信号的时域和频域特性。并在第三章提出了基于无损融合的在数据级融合技术对管道信号做融合，利用 LMS 自适应噪声抵消系统减少采集信号中的背景噪声的影响。为了获取供水管道泄漏信号的各种特征，特别是辨别泄漏信号与突发干扰，背景噪声的差异性，从多方面获取信号的全面特征是非常必要的。本章首先从信号的时域和频域、自相似角度、主要成分对管道振动信号进行分析提取了三种能够识别泄漏的信号特征：固有模态函数的平均功率谱密度，近似熵，信号主成分特征并对管道漏水信号做检测，通过仿真分析了各个检测算法的特点。最后提出了一种基于提取信号的多特征并结合支持向量机等辨识模型的泄漏检测算法。

4.1 基于 EMD 平均功率谱密度的泄漏检测

4.1.1 经验模态分解算法原理

经验模态分解 (Empirical Mode Decomposition, EMD) 算法是目前在工程领域中被广泛应用的对信号做自适应的时频分解算法，非常适合于处理供水管道振动信号的分解处理^[29]。而现在先普及一下瞬时频率和固有模态函数 (Intrinsic Mode Function, IMF) 的概念^[30]。在很多情况下信号的频率随着时间的变化而变化，所以如果通过傅里叶变换来获得瞬时频率则达不到很好的效果，而 Hilbert 变换方法能够较好地得到信号的瞬时频率。现在对任意时间序列 $x(t)$ ，得到 Hilbert 变换

$$y(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (4-1)$$

构造解析函数：

$$z(t) = x(t) + iy(t) = a(t)e^{i\Phi(t)} \quad (4-2)$$

其中相位函数：

$$\Phi(t) = \arctan \frac{y(t)}{x(t)} \quad (4-3)$$

相位函数的导数为瞬时频率：

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/808023131037006045>