



图8 研究总体技术方案

3 主要研究内容

在声学智能感知领域，主要针对声源分离算法、声成像算法及基于多维信息融合的声学故障辨识方法等方面，主要通过解决多声源耦合问题及多维信息融合提升声学故障辨识的精度；在超声智能感知领域，主要对非接触超声损伤特征提取方法、高精度超声定位方法、基于概率的超声成像算法等方面进行布局，从而实现高精度的超声缺陷识别与定位；在边缘节点软硬件设计方面，主要针对神经网络加速单元和信号处理加速单元设计、集成编译开发与运行时管理软件设计等方面进行布局，实现算法加速及多速率、多任务算法模型的在线评测与快速应用部署。在云平台与边云协同部署策略方面，主要在面向多模数据融合的云平台架构设计与面向分布式与集成式混合管理的边云协同部

署策略方面进行专利布局，实现计算实时性强、模型迭代高效、模型更新及时的边协同系统。

3.1 工业多模态智能感知认知基础理论与系统框架研究

3.1.1 工业中传统感知技术收集的信息难以满足AI的决策的质量需求

新型感知声成像和超声成像技术可以非接触、快速无损探测目标对象的故障及位置，因此声和超声在工业诊断中具有广阔的前景。开展声和超声感知的多模态信息融合技术研究，可在工业诊断中实现产品或零部件的快速、无损检测和定位。

研究以质量为基础开展工业多模态智能感知认知基础理论研究。将对声和超声的多模态采集分析机理、多模态信息融合机理及深度学习故障辨识理论等进行研究。

3.1.2 工业多模态智能感知认知系统框架

工业多模感知与边云协同系统总

体框架如图9所示。系统主题框架分为三层：

(1) 感知层：包含声像智能感知系统和超声智能系统和超声智能感知系统。其中声像智能感知系统实现声学信息的感知、采集、声源分离、定位、成像和信号传输。而超声智能感知系统实现超声信息的感知、采集、损伤特征提取、成像及信号传输等。

(2) 边缘节点层：以国产应用处理器及FPGA实现边缘节点，以FPGA实现神经网络及信号处理的加速，并构建集成编译开发与运行时管理软件。在边缘节点层实现声学及超声的故障检出、辨识与定位功能等。

(3) 云平台层：工业多模感知融合云平台，云平台包含工业IaaS层、工业PaaS层及工业SaaS层，提供边云协同功能，实现模型的管理、训练、部署、调度及更新迭代等。研究将应用在空调设备运转噪音检测，实现多区域多厂区部署。如图9

3.2 基于声像智能感知的故障辨识与定位技术研究及感知系统研制（如图10）

本文研究基于声像智能感知的故障辨识与定位技术，采用深度卷积、声达时延差、波束成形等算法实现声源分离、定位、成像、采用声像对抗和迁移学习等建立小样本模型，采用自组织神经网络对新样本进行在线学习以提升模型辨识率，采用基于记忆回放机制的增量学习模型实现新故障自动标注识别，采用基于时序/空间/时频多维信息融合推理模型的故障检出、辨识和故障声定位，并研制声像智能感知系统，达到产品或零部件故障检测行业应用的故障检出率、辨识准确率及故障定位精度等要求^[8-11]。

3.2.1 工业声像智能感知系统研制

(1) 声像智能感知系统设计(如图11)

声像智能感知系统实现声学信息的感知、采集、处理、计算和传输,以及软件/模型远程更新等功能。其硬件主要包括:声学阵列、声学高速采集单元、同步采集控制单元、主处理单元、存储单元、运算加速单元、声学高速传输单元等;其软件主要执行声源分离、声源定位、声成像等声音数据运算以及感知系统软件/模型更新任务。该系统拟达到采样速率48KHz、采样精度32bits、采样通道24路等指标,其关键技术在于声学阵列、同步采集控制,高速采集与传输等。如图12

(2) 声学阵列及同步采集技术研究(如图13)

为实现从检测对象声源与环境噪声的时频混合声中提取目标声源信号,本文研究声学阵列技术。拟采用24个声学传感器构成声学阵列,以实现声像智能感知系统可分离故障声源6个以上。研究基于控制变量法研究阵列孔径、阵元数量、阵元空间分布、声学传感器性能等参数对阵列性能的影响特性,以获得声学阵列的优化设计模型,从而研制频率响应、空间分辨、混叠抑制、阵列结构等最优的声学阵列,满足声源分离、声源定位、阵列安装等条件。

声学采集单元之间的时钟频率差异,造成采样间隔存在误差,且此误差会随时间累积。若不及时消除累积误差,易导致声音数据失去时间同步性,从而影响数据处理结果。研究通过研究声学阵列同步采集技术,可消除声学采集单元之间的采样累积误差,确保声音数据的时间同步性小于25us。



图9 系统总体框架图

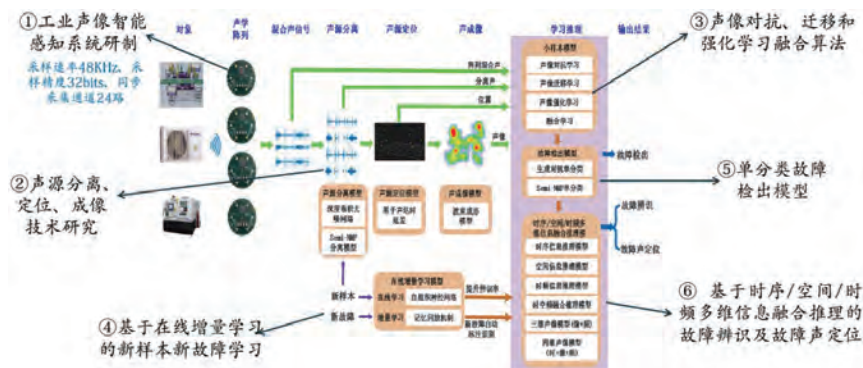


图10 基于声像智能感知的故障辨识与定位技术示意图

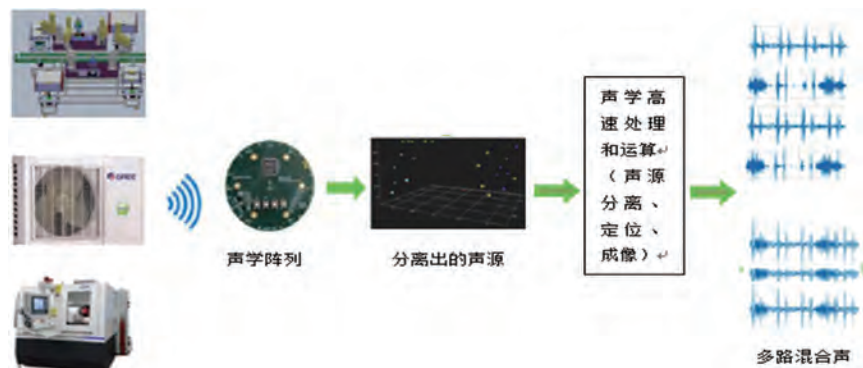


图11 声像智能感知系统总体设计示意图

(3) 声学高速采集与传输(如图14)

本文研究声学高速采集与传输技术。为满足采样速率48KHz、采样精度32bits、采样通道24路等系统性能,声学高速采集单元与声学高速传输单元之间的通信带宽须不低于3.5Mbps,声学高速传输单元与边缘节点之间的

通信带宽不低于500Mbps。研究采用高性能DSP处理器+高精度音频AD实时采集声学信息,并通过高速总线向FPGA处理器传输,每通道传输带宽可达6.25Mbps。FPGA处理器+ARM处理器传输,每通道传输带宽可达6.25Mbps。FPAG处理器+ARM处理

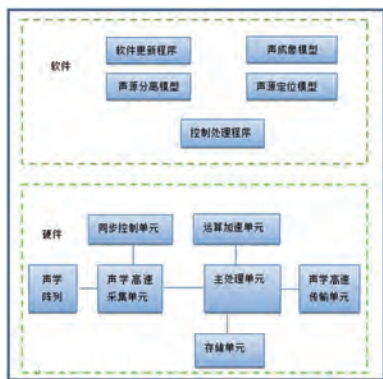


图12 声像智能感知系统框图

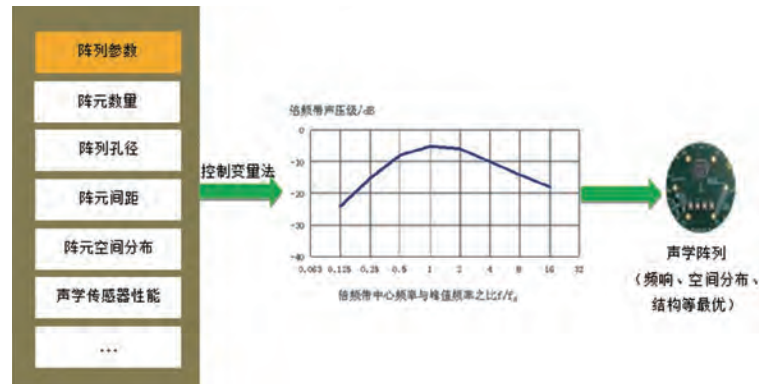


图13 声学阵列优化设计研究示意图

器完成声学信息处理和运算，并将声学信息和运算结果通过工业级以太网向边缘节点传输，传输带宽可达1000Mbps。

3.2.2 声源分离、定位、成像技术研究

声学阵列拾取的是检测对象正常声、故障声及环境声等形成额混合声，不便目标声源提取和定位。为此，本文研究声源分离技术，以实现从单通道混合声中提取目标声源，从而利用多个通道提取的目标声源对其进行定位。针对混合声信号中各声源时频域基本完全重叠的情况，研究采用基于深度卷积去噪的单通道声源分离技术；针对混合声信号中各声源频域仅部分重叠的情况，研究采用基于Semi-NMF（半非负矩阵分解）的单通道声源分离技术。

(1) 基于深度卷积去噪的单通道声源分离技术（如图15）

针对混合声信号中声源时频域基本完全重叠的情况，本文研究基于深度卷积去噪的单通道声源分离技术：为各声源分别学习分离模型，实现分离模型只需学习其所对应的声源的特征，以获得更好的分离效果。研究构建如图16所示的声源 i ($i=1, 2, \dots, n$) 分离网络：混合信号先经STFT（短时傅里叶变换）计算网络得到时频谱和

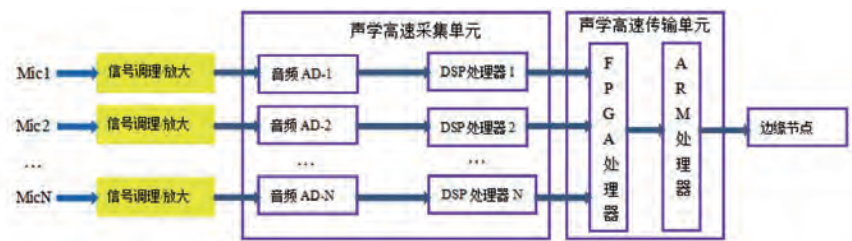


图14 声学高速采集与传输结构示意图

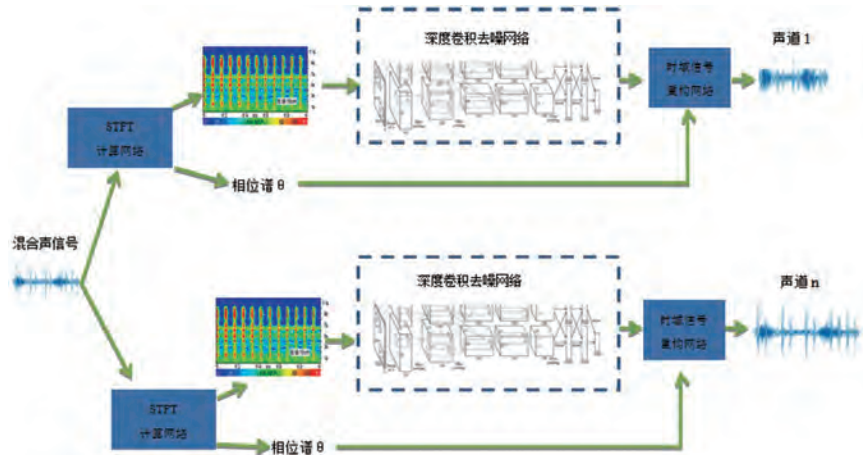


图15 基于深度卷积去噪的单通道声源分离技术示意图

相位谱，由深度卷积去噪网络从混合信号的时频谱中提取声源 i 的时频信息，把此时频信息和相位谱输入时域信号重构网络运算得到干净的声源 i 。

(2) 基于Semi-NMF的单通道声源分离技术^[12]

针对混合声信号中各声源频域仅部分重叠的情况，本文研究基于Semi-NMF（半非负矩阵分解）的单通道声源分离技术。如图11所示，声

源分离流程包括：单通道混合声音信号 S 由若干独立的声音信号 S^1, S^2, \dots, S^n 混合而成，计算 S 的傅里叶变换结果 F ，根据 F 计算频谱 X ；对 X 进行半非负矩阵分解，得到特征矩阵 W 和系数矩阵 H ；根据 X 得到声音信号 S^1, S^2, \dots, S^n 各自的初始估计频谱 X^1, X^2, \dots, X^n ；对其进行半非负矩阵分解，得到对应的特征矩阵 W^1, W^2, \dots, W^n 和系数矩阵 H^1, H^2, \dots, H^n ；根据傅里叶变

换结果F以及频谱 X^1, X^2, \dots, X^n , 获得声音信号 S^1, S^2, \dots, S^n 各自的傅里叶变换结果 F^1, F^2, \dots, F^n 进行傅里叶逆变换, 得到声音信号 S^1, S^2, \dots, S^n 实现声源分离。

(3) 多声源定位技术

为了声源成像及故障声源精准定位, 本文研究声源定位技术, 实现高精度声源定位。实现高精度声源定位。声源传播至各声学传感器的时间差, 暗含其在传感器坐标系中的位置信息, 因此本文研究基于声达时间差的声源定位网络。采用基于广义互相关函数的时延估计算法, 获得声源到各声学传感器的时间差, 再将此时间差输入声源定位网络, 以计算出声源位置。为了便捷的训练声源定位网络, 本文研究思路如下: 建立声学传感器坐标系后, 模拟若干声源坐标, 计算声源传播至各声学传感器的时间差, 从而利用这些声达时间样本训练声源定位网络。经测算, 研究理论上可以实现声源定位精度约0.5cm。如图17

(4) 基于分离声源的波束成形声成像技术

声成像是基于声学阵列的测量技术, 通过测量空间内声波到达各阵元的信号的时间差异, 确定声源的位置并测量声源的幅值, 以图像的方式显示声源在空间的分布。由于声音信号在发声面中各个点的声压不同, 故声像能反映整个发声面的声压分布情况, 反映声源的空间信息。因此, 与其他声学测量模态信号相比, 利用声像进行故障检测更有优势。如图18

波束成形是声成像的主要方法, 通过麦克风阵元的响应向量, 计算声压在目标平面的分布。我们使用延时求和算法求解波束成形, 计算框架图



图16 基于Semi-NMF单通道声源分离技术示意图

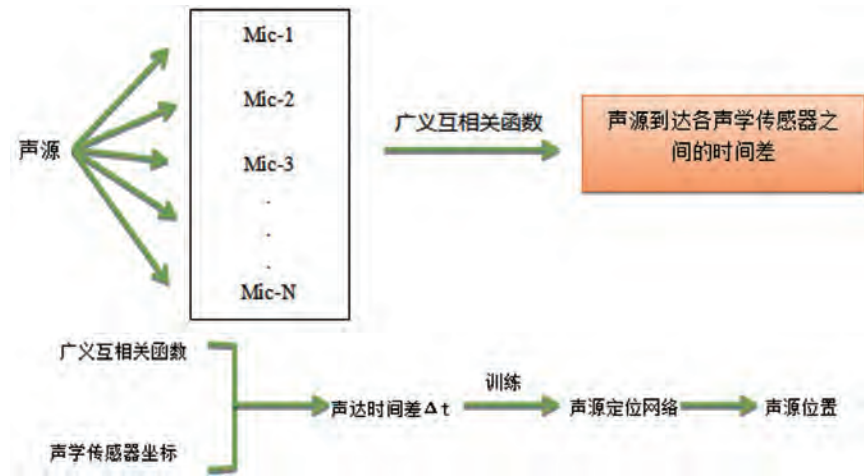


图17 基于声达时间差的声源定位技术示意图

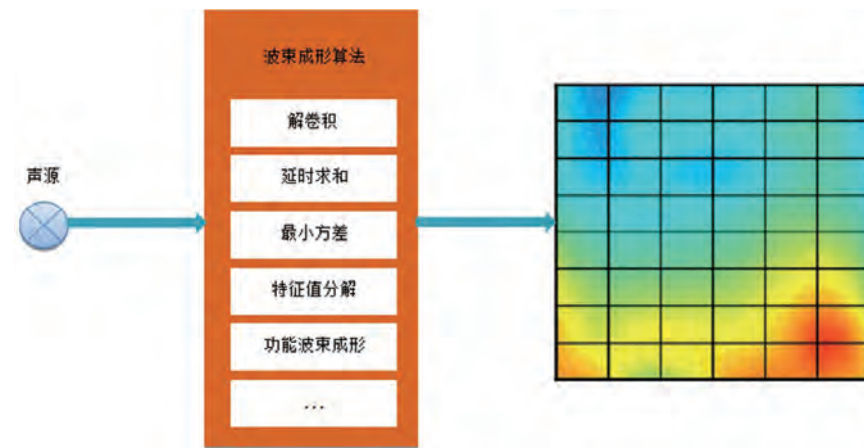


图18 声成像技术计算框架图

参考图18.假设有m个阵元阵列, 每个阵列间距为d, 第一个阵列的权重为 $W_i(i \in \{0, m\})$ 。声源发射入射波入射到阵元平面, 根据公式(1)(2)(3), 第n时刻时m个阵元所收到的波形可表示为:

$$X(n)=[x(n-\delta_1), x(n-\delta_2), \dots, x(n-\delta_m)] \quad (1)$$

$$\delta_i = \frac{roc(\theta) * p_i}{c} \quad (2)$$

$$roc(\theta)=[-\sin(\theta)-\sin(\theta)\cos(\theta)] \quad (3)$$

其中 θ 为入射角度, δ_i 为第i个阵

元接受的波束时延, $roc(\theta)$ 以旋转矩阵为基的入射角向量, c 为声速, P_i 为第i个阵元的坐标位置。将m个阵元按照波束成形算法最后得到输出波束y:

$$y(n) = \sum_{i=1}^m W_i x(n) \quad (4)$$

根据公式(4)最后输出波束值为带权重的m个阵元相应的累加和。

相比于其他波束成形解算方法例如解卷积、最小方差、特征分解和功能波束成形算法, 延时求和算法具有易于求解, 计算效率高的优势。

3.2.3 声像对抗, 迁移和强化学习融合算法

在利用声学感知信息进行故障辨识的过程中, 一些故障声的声像样本数量少, 使用少量声像样本进行故障辨识, 会导致声像图的特征学习不充分, 学习效果不佳。本研究使用声像对抗、迁移和强化学习融合算法研究故障声像的小样本建模问题, 使用少量的样本信息生产更多标注信息的感知样本数据。

(1) 声像对抗学习算法 (如图19)

研究利用生成对抗网络 (GAN) 学习故障声的声像图, 生成与故障声像分布相同的声像图样本, 达到扩大故障声像图样本数量的目的。生成器和判别器构成了GAN的主要结构, 生成器输入时噪声变量, 输出是生成样本; 判别器用来判断输入的样本是真实还是伪造。判别器、生成器损失函数在公式(5)(6)中:

$$\max_D V(D, G) = E_{x \sim p_{data}} [\log(D(x))] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \quad (5)$$

$$\min_G V(D, G) = E_{z \sim p(z)} [\log(1 - D(G(z)))] \quad (6)$$

其中D, G分别表示生成器和判别器, x表示真实样本, z表示噪声变量。完成判别模型的优化后, 对生成的模型进行优化, 最后通过生成模型生成与真实分布相同的故障声的声像图样本。

相比于其他小样本建模方法, 使用GAN对故障声的声像图进行小样本建模, 无需对声像图原始分布进行预先假设, 可以自动学习真实声像样本的数据分布。

(2) 声像迁移学习算法

本文研究迁移学习对小样本进行多任务迁移的方法, 使用生成模型学习故障的声像图特征信息, 并将声像

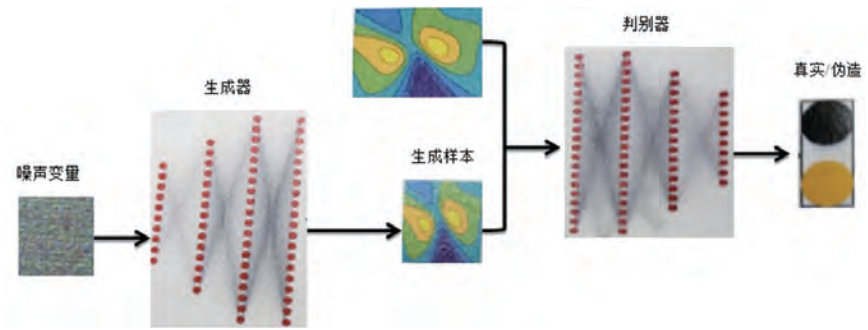


图19 小样本生成对抗网络计算框架

图特征知识迁移到相似故障类别的声像图生成学习任务中, 实现小样本多任务迁移学习。声像迁移学习计算框架如图20。

声像迁移学习的算法流程: 第1类故障声像图经过生成模型1学习图像特征, 得到第1类故障声像生成样本; 将模型1的知识迁移到生成模型2中, 得到第2类故障声像生成样本; 最后迁移生成第N类故障声像样本。

相比于其他小样本建模方法, 使用迁移学习模型对故障声的声像图进行小样本建模, 将学习到的模型知识迁移到多任务中, 提升学习模型的泛化性。

(3) 声像强化学习算法

针对数据样本多而标注样本少的情况, 本文研究强化学习对小样本进行标注学习的方法。设计训练样本分发器将少量声像样本数据分发给判别网络

(处理能够识别的数据)。其中环境空间包括了近似度分析模块, 实现与前样本的近似性对比, 输出近似系数; 手动标注模块, 输出标注系数, 实现专家打标签功能。此外, 训练样本针对不确定数据输出计算延迟系数。将延迟系数、标签分数、标注系数、近似系数一起输入到政策学习网络, 输出标签、需要专家标注和标注延迟的分数信息, 传给智能体进行决策。智能体选择获得奖励最高的行为反馈给环境空间。声像强化学习计算框架如图21。

相比于其他小样本建模方法, 使用强化学习模型对故障声的声像图进行小样本建模, 将资源的延迟、判别网络、政策学习网络三者相结合, 可在样本数量多、标签少的情况下应用。

(4) 基于声像对抗迁移的融合学习算法

本文研究融合学习对小样本进行

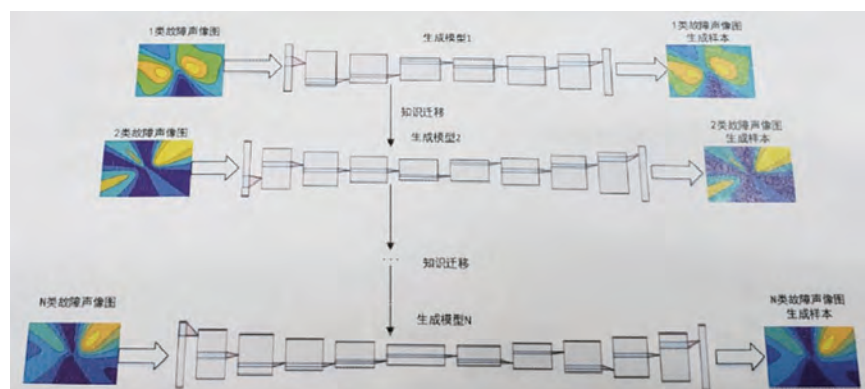


图20 声像迁移学习计算框架