



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

声人工结构的智能化设计与应用

作者姓名: 孙雪聪

指导教师: 杨军 研究员 中国科学院声学研究所

贾晗 研究员 中国科学院声学研究所

学位类别: 工学博士

学科专业: 信号与信息处理

培养单位: 中国科学院声学研究所

2022 年 6 月

Intelligent Design and Application of the Acoustic Metamaterials

A dissertation submitted to the
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in Signal and Information Processing
By
Sun Xuecong
Supervisor: Professor Yang Jun
Professor Jia Han

Institute of Acoustics, Chinese Academy of Sciences

June, 2022

中国科学院大学

学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日 期：

中国科学院大学

学位论文授权使用声明

本人完全了解并同意遵守中国科学院大学有关保存和使用学位论文的规定，即中国科学院大学有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘要

声人工结构是一种人工设计的复合声学结构，这种复合结构的性质主要取决于其结构的尺寸、形状以及排列方式，因而具备很多天然材料所不具备的新奇物理特性。在过去的二十年中，人们利用声人工结构实现了负折射、各向异性等奇异的声学现象，因而构造了具有声聚焦、声隐身等多功能的声学器件，为有效地操纵和调制声波提供了新的技术思路。然而，在对具有特定功能需求的声学系统进行设计时，声学结构的逆设计一直是一个十分具有挑战性的问题。一方面，我们需要解决如何对复杂声学结构进行快速精准化设计的问题，即声学结构的智能化设计问题；另一方面，我们也在设计过程中考虑与后端信号处理算法的适配性，使整个声学系统的性能最优化，即声学结构的智能化应用问题。在此背景下，本文针对在声学结构的智能化设计与应用中所存在的技术问题展开研究，具体内容如下：

1. 针对复杂声学结构参数多，难以通过人工手动设计的问题，提出了一种基于遗传算法的声学结构自动化设计方法。该方法不仅可以与层状严格耦合波分析等解析类算法相结合，用于提升吸声体在目标频段的吸声性能；还可以与有限元等数值类算法相结合，对声学隐身器件的隐身效果进行优化。通过仿真与实验相结合的方式证明了所提出方法的通用性和有效性。
2. 针对优化类算法计算成本高等问题，将深度学习模型和集总参数模型相结合，提出了一种快速高效的声学结构逆设计方法。通过对声学结构进行解析化建模，正向生成、筛选了足量的样本数据，并将其用于深度学习模型的训练和测试。基于提出的模型实现了二阶亥姆霍兹共鸣器单元的自动化按需设计，并对所设计的单元进行组合完成了低频宽带隔声器件的设计工作。
3. 针对声学结构在实际应用中的多样化需求，将深度学习与其他算法相结合，进一步拓展了深度学习在声学结构设计领域的应用。针对声学结构设计过程中普遍存在的多解现象，将所提出的模型与降维类算法相结合，实现了声学结构的批量化设计与多维度筛选；针对优化类算法对初值敏感、鲁棒性较差的问题，利用所提出的模型为优化类算法提供良好的初始条件，有效避免了算法过早地陷入局部最优点，显著地提升了最终的优化效果。最后我们将所提出的深度学习

模型在集总参数模型的框架下向其他声学结构进行了迁移和推广，进一步展现了所提出的方法的灵活性和通用性。

4. 针对多通道声传感系统通道多、能耗高、不易操控等问题，提出了一种基于声学结构的单通道声学传感系统，实现了三维空间中多个声源的方位辨识与内容识别。在对声学结构的设计过程中充分考虑了与信号处理算法的联动性和相互作用，基于物理层的声学滤波机制实现了对来自不同方向声波的差异性编码，并结合字典学习算法，以较低的复杂度实现了对声音方位信息和音频内容信息的重建。与传统的多通道声学传感系统相比，基于声学结构的单通道系统具有系统简单、便于操控、功耗小等优点，为设计新型轻量化振动传感器件与系统开辟了新的道路，具有广阔的应用前景。

总之，本文系统地对声人工结构的智能化设计与应用展开了深入的研究，在对多种复杂声学结构进行理论分析和仿真计算的基础上，针对不同声学结构的物理特性和应用需求提出了与之适配的智能化设计策略，设计了新型声学系统装置原型，实现了新奇、丰富的声学现象和功能。研究成果使声人工结构设计突破以经验和大量“试错”为前提的设计困境，极大地节约了人力成本和时间成本，为多种类型的声学器件和声学系统设计提供了理论与技术支持，进一步推动了声人工结构的实用化和产业化发展。

关键词：声人工结构，优化算法，深度学习，单通道声学传感系统

Abstract

Acoustic metamaterials are artificial composite acoustic structures, whose properties mainly depend on their geometric structures and arrangement. Therefore, acoustic metamaterials have many novel physical properties that natural materials do not have. In the past two decades, people have used acoustic metamaterials to realize a lot of novel physical phenomena, such as negative refraction and anisotropic materials. And many acoustic devices, including acoustic lens and cloaks, have been constructed based on acoustic metamaterials, which provided a new material platform for the manipulation of sound propagation. However, the development of accurate and computationally efficient design and optimization approaches for the acoustic metamaterials is still in the early stages. On the one hand, we need to design the complex acoustic structures efficiently and accurately to meet the requirements, that is, the intelligent design of acoustic metamaterials. On the other hand, we also need to consider the adaptability of the structures with the signal processing algorithm during the design process to optimize the performance of the whole acoustic system, that is, the intelligent application of acoustic metamaterials. This paper is devoted to explore the technical problems existing in the intelligent design and application of acoustic metamaterials. The main contents of this paper include following aspects:

1. An optimization strategy based on the genetic algorithm is proposed to solve the problem that it is difficult to design complex acoustic structures manually. The proposed method can not only be combined with analytical algorithms, such as the layered rigorous coupled-wave analysis, to improve the sound absorption performance of the absorber in the target frequency range, but also be combined with numerical algorithms, such as finite element method, to optimize the stealth effect of the acoustic cloak. The performance superiority of the proposed method has been demonstrated through simulations and experiments.
2. In view of the high complexity of the optimization algorithms, we propose an efficient, flexible and universal acoustic structure design method to solve the inverse

problem of the acoustic structures based on the deep learning and lumped-parameter technique (LPT). In our method, we analyze acoustic structures by LPT and develop connections between geometric parameters and equivalent electrical parameters. Then, we generate the datasets to train and test the deep learning model. To evaluate the effectiveness of the proposed model, a two-order Helmholtz resonator is designed to realize acoustic insulation at specific frequencies. Moreover, we design a composite structure with 9 two-order Helmholtz resonators using the proposed method to realize low-frequency broadband sound insulation.

3. Some extended applications of the deep learning model in the field of acoustic structure design are discussed. Considering that the corresponding relation between target acoustic functionality and acoustic structures is usually not a deterministic one-to-one mapping, we combine the proposed model and the principal component analysis to search for multiple solutions of the inverse design problem and make further selections based on different properties. Moreover, the proposed model can provide a good initial condition for the optimization algorithm to avoid fall into local optimum and improve the optimization effect. Finally, considering that various acoustic structures can be analyzed by the LPT exactly in the low frequency range, the proposed approach has a strong versatility and scalability, which can be further extended to other acoustic structures.

4. We develop a metamaterial-based single-microphone listening system (MSLS), which can localize and separate multiple sound signals from an overlapping signal in three-dimensional(3D) space. A 3D metamaterial enclosure (ME) is used to provide monaural cues to the inversion task by coding sound signals as a function of the source direction, which is designed from the point of view of adaptation of the structures with the signal processing algorithm. During signal processing, a joint algorithm of variable sparsity principal component analysis and orthogonal matching pursuit (VSPCA-OMP) is used to solve the multi- source listening problem, which has the advantages of low computational complexity and good real-time performance. A lot of experiment results promise a wide range of potential applications for our proposed system, such as intelligent scene monitoring and robot audition.

In a word, this paper systematically makes an in-depth study on the intelligent

design and application of acoustic metamaterials. Corresponding theoretical analysis and simulation calculations have been performed. According to the different physical characteristics and application requirements, several automatic design methods are proposed. The prototypes of new metamaterial devices are designed to realize novel and rich acoustic phenomena and functions. The proposed methods break through the current predicament of the structure design overly relying on the experience, which greatly save the labor cost and time cost. This study provides an important theoretical and technical basis for the design of various types of acoustic devices and acoustic systems, which further promotes the practical and industrial development of acoustic metamaterials.

Keywords: Acoustic metamaterials, Optimization algorithm, Deep learning, Single-sensor acoustic sensing system

目 录

第 1 章 绪论 ······	1
1.1 研究背景与意义 ······	1
1.2 研究现状及分析 ······	2
1.2.1 声人工结构概述 ······	2
1.2.2 声人工结构的自动化设计 ······	4
1.2.3 研究难点及关键科学问题 ······	7
1.3 研究内容与结构安排 ······	8
第 2 章 基于遗传算法的声人工结构优化 ······	11
2.1 引言 ······	11
2.2 理论基础 ······	11
2.2.1 遗传算法简介 ······	11
2.2.2 COMSOL Multiphysics with MATLAB 简介 ······	16
2.3 宽频吸声体的形状优化 ······	17
2.3.1 层状严格耦合波分析 ······	17
2.3.2 吸声体分层占空比优化 ······	20
2.3.3 吸声体表面形状优化 ······	21
2.4 声学隐身器件的分层优化 ······	23
2.4.1 五模材料的变换声学理论 ······	24
2.4.2 声学隐身器件的分层设计 ······	27
2.4.3 最小化散射声能量 ······	28
2.4.4 最小化反射声能量 ······	32
2.5 本章小结 ······	34
第 3 章 基于深度学习的声人工结构逆设计 ······	35
3.1 引言 ······	35
3.2 物理模型 ······	35
3.2.1 集总参数模型 ······	37
3.2.2 传输矩阵模型 ······	38
3.2.3 有限元模型 ······	41
3.3 逆设计问题的分析 ······	42
3.3.1 复杂性 ······	42
3.3.2 非唯一性 ······	44

3.4 数据生成	46
3.5 深度学习模型的训练及测试	51
3.6 二阶亥姆霍兹共鸣器单元的按需设计	53
3.6.1 设计流程	53
3.6.2 与优化类算法的性能对比	55
3.6.3 实验结果	57
3.7 基于多个二阶亥姆霍兹共鸣器的低频宽带隔声装置设计	60
3.7.1 耦合效应	60
3.7.2 设计方法	62
3.7.3 隔声效果	64
3.8 本章小结	66
第 4 章 深度学习辅助的多样化声人工结构设计	67
4.1 引言	67
4.2 二阶亥姆霍兹共鸣器的批量化设计	67
4.2.1 主成分分析	68
4.2.2 计算案例	69
4.3 二阶亥姆霍兹共鸣器的优化设计	72
4.3.1 最大化目标频点处的 STL	73
4.3.2 最小化共振峰间隔	75
4.4 深度学习模型用于其他声学结构的逆设计	76
4.5 本章小结	78
第 5 章 基于声学结构的智能化声学传感系统	81
5.1 引言	81
5.2 物理模型	82
5.2.1 声学滤波器	82
5.2.2 半球状外壳	84
5.3 压缩感知算法	87
5.3.1 字典学习	88
5.3.2 变稀疏度主成分分析	91
5.3.3 稀疏恢复	92
5.4 实验测试	95
5.4.1 实验装置	95
5.4.2 多声源辨识	95
5.4.3 实时追踪	104
5.5 本章小结	107

第 6 章 总结与展望	109
6.1 工作总结	109
6.2 前景与展望	110
参考文献	113
致谢	121
作者简历及攻读学位期间发表的学术论文与研究成果	123

图形列表

1.1 具有不同等效参数的典型声人工结构模型 ······	3
1.2 基于拓扑优化的双负声学超材料 ······	6
1.3 结合遗传算法和模拟退火算法的贝塞尔散射体优化 ······	6
2.1 GA 流程示意图 ······	12
2.2 具有(a)平面, (b)三角形, (c)矩形和(d)T 形表面的多孔材料照片, 及其对应的横截面 (e) 平面, (f) 三角形, (g) 矩形和(h)T 形。 ······	17
2.3 基于 LRCWA 方法计算非平整界面的吸声体的吸声系数 ······	18
2.4 多孔材料等效声学参数测量装置示意图 ······	18
2.5 某多孔材料的 (a) 等效折射率和 (b) 等效密度 ······	19
2.6 矩形和三角形多孔材料的结构示意图及吸声系数 ······	19
2.7 优化后的吸声体的吸声系数 ······	21
2.8 具有不同表面形状的吸声体的结构示意图: (a) 三角形; (b) 高斯形; (c) 平面 ······	22
2.9 具有不同表面形状的吸声体的吸声曲线 ······	22
2.10 二维五模材料微结构和能带: (a) 二维五模材料微结构示意图; (b) 二维五模材料能带。 ······	25
2.11 五模材料坐标变换示意图: (a) 虚拟空间 Ω ; (b) 物理空间 ω 。 ······	26
2.12 五模材料隐身衣分层示意图和材料参数分布 ······	29
2.13 GA 进化曲线 ······	29
2.14 优化前后总声场对比: (a) 优化前; (b) 优化后。 ······	31
2.15 优化前后散射声场对比: (a) 优化前; (b) 优化后。 ······	31
2.16 优化前后隐身效果对比 ······	31
2.17 优化前后散射声场对比: (a) 优化前; (b) 优化后。 ······	33
2.18 优化目标为最小化反射声能量时隐身效果对比 ······	33
3.1 THR 的几何结构示意图 ······	36
3.2 THR 作为管道旁支结构的示意图 ······	36
3.3 THR 作为管道旁支结构时的等效电路图 ······	37
3.4 基于 TMM 对 THR 单元进行建模 ······	38
3.5 基于 LPT, TMM 和 FEM 的 THR 单元的 STL 的计算结果对比 ······	41
3.6 THR 单元的共振频率对不同几何参数的灵敏度 ······	43
3.7 期望的声学特性与声学结构之间的映射关系 ······	44

3.8 满足同一设计目标的不同声学结构	45
3.9 样本生成流程图	49
3.10 生成的样本数据	50
3.11 数据集分布	50
3.12 Loss 下降曲线	51
3.13 测试集样本的测试结果	52
3.14 基于深度学习的 THR 单元按需设计流程图	53
3.15 设计目标为在 150 Hz 和 250 Hz 实现 STL 大于 10dB 的设计结果 ..	54
3.16 深度学习与 GA 的设计效率对比	56
3.17 针对三个不同的优化目标的优化结果对比	57
3.18 THR 单元的样品照片及其几何参数	57
3.19 STL 频谱的实验和仿真结果对比	58
3.20 降噪前后时域波形对比：(a) 降噪前；(b) 降噪后。	59
3.21 降噪前后语谱图对比：(a) 降噪前；(b) 降噪后。	59
3.22 多个 THR 单元串联排列示意图	60
3.23 多个 THR 单元间的耦合效应	61
3.24 低频宽带隔声装置的设计过程	64
3.25 宽频隔声装置结构示意图	65
3.26 宽频隔声装置的 STL 频谱	65
4.1 PCA 空间中 β_0 附近的结构的 AERF 分布	70
4.2 在 250 Hz 处的 STL 大于 10 dB 的基础上最大化 150 Hz 处的 STL ...	71
4.3 在 150 Hz 处的 STL 大于 10 dB 的基础上最大化 250 Hz 处的 STL ...	72
4.4 基于两种初始条件最大化 150 Hz 和 250 Hz 处的 STL	74
4.5 基于两种初始条件最大化 300 Hz 和 450 Hz 处的 STL	75
4.6 基于两种初始条件最小化两个共振峰到 400 Hz 的距离	76
4.7 具有两个一阶短管的 THR 单元的逆设计：(a) 结构示意图及几何参数； (b) 结构的 STL 频谱。	77
4.8 具有两个二阶短管的 THR 单元的逆设计：(a) 结构示意图及几何参数； (b) 结构的 STL 频谱。	77
5.1 二阶声学滤波器结构示意图及腔体体积对结构频响的影响	83
5.2 穿孔板孔隙率对二阶声学滤波器频响的影响	83
5.3 穿孔板孔的位置分布对二阶声学滤波器频响的影响	84
5.4 半球状外壳的结构示意图	84
5.5 半球状外壳不同方向的频响	85
5.6 半球状外壳不同方向频响的相关度	86

5.7 单通道定位与分离系统的算法流程示意图	88
5.8 PCA 变换后不同方向频响的相关度	90
5.9 字典矩阵 A 和稀疏向量 s 的对应关系	93
5.10 实验测量: (a) 实验装置照片; (b) 半球状外壳样品照片; (c) 声源位置示意图。	94
5.11 测量装置示意图	94
5.12 街道场景下语料的语谱图和功率谱: (a) 警车的鸣笛声; (b) 倒车的提示音; (c) 消防车的鸣笛声; (d) 轿车的鸣笛声; (e) 消防车的鸣笛声; (f) 自行车的铃声。	97
5.13 街道场景下听音测试结果	98
5.14 动物园场景下语料的语谱图和功率谱: (a) 猫的叫声; (b) 蝉鸣声; (c) 蟋蟀的叫声; (d) 蜜蜂的叫声; (e) 青蛙的叫声; (f) 鸭子的叫声。 ..	99
5.15 动物园场景下听音测试结果	100
5.16 音乐厅场景下语料的语谱图和功率谱: (a) 吉他; (b) 编钟; (c) 萨克斯; (d) 鼓; (e) 小提琴; (f) 中国民乐。	101
5.17 音乐厅场景下听音测试结果	102
5.18 家庭场景下语料的语谱图和功率谱: (a) 做饭的声音; (b) 抽水马桶的声音; (c) 吹风机吹头发的声音; (d) 消声; (e) 撕纸的声音; (f) 东西碎了的声音。	103
5.19 家庭场景下听音测试结果	104
5.20 对话场景下语料的语谱图和功率谱: (a)-(c) 三位不同男性讲话的声音; (d)-(f) 三位不同女性讲话的声音。	105
5.21 对话场景下听音测试结果	106
5.22 基于 Speech Commands 数据库的测试结果	106
5.23 针对俯仰方向的听音测试及结果	107
5.24 对移动声源进行实时追踪	107

表格列表

2.1 优化后高斯形状表面的系数	23
2.2 几种典型的映射关系	28
2.3 优化前后隐身衣每一层的半径	30
3.1 几何参数和等效电学参数的转换关系	38
3.2 样本分布	47
3.3 3个THR单元的几何参数及共振频率	61
3.4 低频宽带隔声装置的THR单元的几何参数及共振频率	66
5.1 不同场景下听音测试结果	96

第1章 绪论

1.1 研究背景与意义

在日常生活中，我们无时无刻不在接收声音，也在发出声音。因此，声音是人类最早研究的物理现象之一，声学也是经典物理学中历史最悠久，并且当前仍处在前沿地位的物理学分支学科。很长时间内人们对声学的探索都是基于经验的规律性探索，直到17世纪初，伽利略对单摆周期和物体振动的研究正式开启了将声学视为一种自然科学的系统性研究。随后，无数杰出的物理学家和数学家投身于研究物体的振动和声的产生原理。1877年，瑞利对19世纪以前两三百年大量的声学研究成果进行了系统地总结，出版的《声学原理》一书至今仍是声学研究领域的经典著作[1, 2]。现如今，随着电子技术的发展，人们开始可以通过制造电子仪器随时随地地发射和接收声波，对声学的研究也从可听声频段逐渐扩展至超声频段和次声频段。

声波一方面具有很强的穿透力，可以绕过障碍继续传播；另一方面由于媒介的粘滞性等原因又会伴有不可避免的衰落和损耗。因此，如何对声波进行有效地操控和调制一直是研究者们密切关注的话题。无论是在自然界还是工业界，材料都被普遍用于调控各类波动现象。例如，我们可以用放大镜实现光的聚焦，鲸类动物可以通过头部的瓜状体来对声信号进行放大等。近年来，在传统结构学与材料学发展的基础上，人们开始对声人工结构展开研究。通过将传统的声学材料以一定的几何结构、空间阵列进行组合，可以实现具有特殊物理参数的声人工结构，从而对声源辐射特性、声波传播路径、声能量吸收效率等进行调控。与传统材料相比，声人工结构的最大优势在于其可以利用亚波长尺寸的结构实现材料物理特性的调节以及对声波的调控。在此基础上，研究人员也基于声人工结构构造了声学弯管[3, 4]、声学隐身毯[5, 6]、声学超表面[7]等具有实际应用价值的声学器件。

然而，在构造具有特定功能需求的声学器件和声学系统的过程中，声人工结构的逆设计依然一个十分具有挑战性的问题。所谓逆设计，指的是通过对结构的几何参数进行设计使该结构具有某种期望的物理性质或物理响应。传统的声学结构设计主要针对的是形状较为规则、参数量比较少的简单结构，设计者可

以基于一些经典的物理模型以及设计经验和直觉来确定基本的模型框架，再手动调整结构的几何参数使结构的物理性质与预期目标相吻合。然而，这个过程是全程需要人工干预的，并且高度依赖于设计者的专业知识和经验，设计效率很低，很难应用于设计不规则的复杂结构。因此，对于较为复杂的声学结构，我们通常会借助时域有限差分（Finite-Difference Time-Domain, FDTD）和有限元仿真（Finite Element Method, FEM）等数值方法对其物理性质进行正向求解，并与各类优化算法相结合，在参数空间进行自动化寻优。与传统的人工设计所得到的声学结构相比，优化后的结构往往会展现出更新奇的物理现象，拥有更极致的声学性能。但是，由于优化方法的本质还是通过多次的正向计算来逐渐逼近优化目标，因此计算成本非常高；且大部分优化算法本身的参数也需要调节，具有对初值很敏感、容易陷入局部最优等问题。

近年来，机器学习技术不断发展，在除计算机领域外的生物学，材料科学，量子物理等领域都展现了其独特的优势 [8–12]。机器学习可以通过数据驱动的方式从大量的数据中自动发现和学习有用的信息，具备自主学习、联想储存、高速寻优等优点。采用机器学习模型对声人工结构进行设计可以略过传统设计中复杂的建模过程和正向求解过程，将基于经验的直观设计转变成基于大数据的自动化设计。并且前期的仿真和实验数据都可以视为是一次性投入的计算成本，模型一旦训练完成即可满足后续多个结构的自主设计，具有简单高效等优点。

本文在前人工作的基础上对声人工结构的智能化设计与应用展开了理论与实验研究。首先，针对不同的应用场景和需求，提出了适用于不同复杂声学结构的智能化设计方法，并结合具体的设计案例对所提出的设计方法进行了深入的分析，证明了方法的有效性。同时也对声人工结构的智能化应用进行了探索，基于所设计的声学结构提出了一种新型的声学传感系统，在仅使用一个传声器的基础上实现了三维空间中多个声源的辨识。

1.2 研究现状及分析

1.2.1 声人工结构概述

声人工结构是一种人工设计的复合声学结构，这种复合结构的性质主要取决于结构的尺寸、形状以及排列方式，因而具备很多天然材料所不具备的新奇物理特性，如图 1.1 所示 [13]。Liu 等 [14] 利用置于环氧树脂基体中的包覆铅球

的硅橡胶构造了三组元立方声子晶体结构，实现了负等效质量，对应图 1.1 的第二象限；Lee 等 [15] 在系统中同时引入单极谐振的亥姆霍兹腔与偶极谐振的薄膜结构，在入射波频率小于单极共振频率时，系统的等效质量密度和等效模量同时呈现出负值，成功构建了一种“双负”材料，对应图 1.1 的第三象限；Fan 等 [16] 通过实验证明，含有一维亥姆霍兹共振腔的声学超材料模型在共振频率处可以产生负的等效弹性模量，对应图 1.1 的第四象限；Park 等 [17] 通过构建等效密度近零的声人工结构，实现了倏逝波的放大，对应图 1.1 的纵轴；Cselyuszka 等 [18] 基于亥姆霍兹共鸣器实现了等效模量接近零的声人工结构，对应图 1.1 的横轴；Dubois 等 [19] 基于声学狄拉克锥首次实现了阻抗匹配的双零折射率结构，实现点声源的指向性辐射与准直，对应图 1.1 的原点。

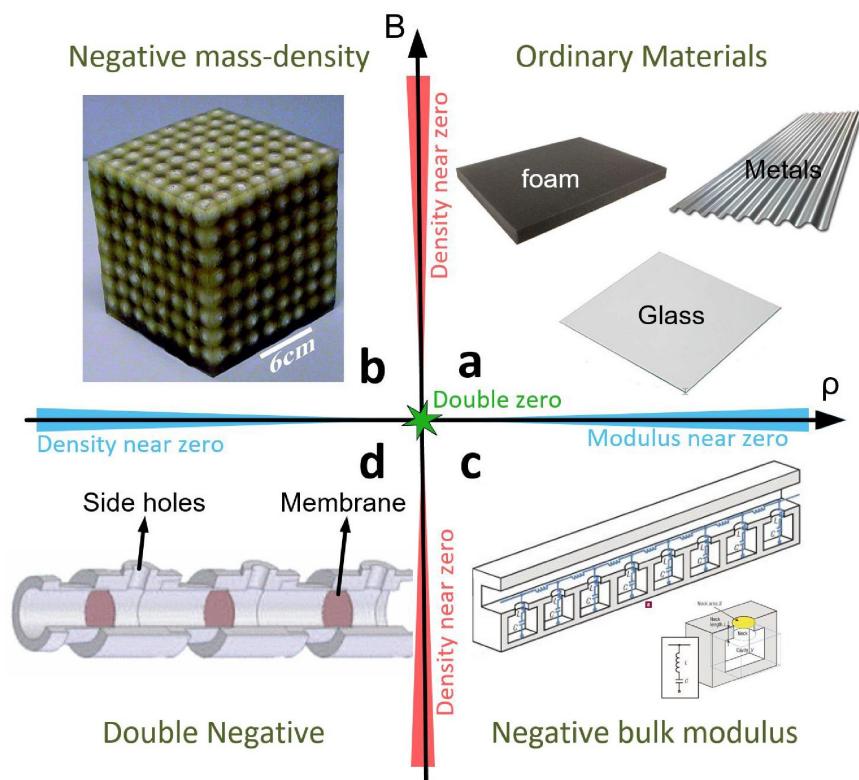


图 1.1 具有不同等效参数的典型声人工结构模型

Figure 1.1 Acoustic metamaterials with different equivalent parameters

与传统材料相比，声人工结构的最大优势在于其可以利用亚波长尺寸的结构实现材料物理特性的调节以及对声波的调控。因此，除了上述一系列对声人工结构基础理论的探索，研究人员还基于声人工结构提出了一系列具有实际应用价值的声学器件：Mei 等 [20] 提出一种“暗声学”薄膜超材料结构，通过在弹性薄膜上镶嵌一些非对称的金属片，在低频范围内实现了多个高吸声系数的吸收

峰；Li 等 [21] 采用亥姆霍兹共鸣器阵列设计了一种传输型超表面结构，可以实现自弯曲的声学波束；Melde 等 [22] 利用 3 D 打印技术制作了一种可以实现声学全息成像的超表面结构，这种结构克服了传统声学换能器阵列结构复杂的缺点，可以实现灵活的声波波前调控；Zhu 等 [23] 利用声学超表面结构将传统施德罗扩散体的厚度由波长的 $1/2$ 减小至 $1/20$ ，消除了传统施德罗扩散体在器件尺寸和几何形状方面的制约；Memoli 等 [24] 提出了一种离散化、可压缩编码的声学超表面结构，可以实现微小物体的声学悬浮；Liu 等人将各向异性的概念引入到三维折叠空间的系统中，提出了一种打开低频、宽频声子带隙的方法，并设计了既能隔声降噪，又能通风透气的三维声学超构笼子 [25]。

近十年来，研究人员发现经过精心设计的声人工结构不仅可以实现声场的任意调控 [26, 27]，还可以在空域和频域中实现声波的传输编码，为发展兼顾高精度、轻量化的新型声学传感器件提供了新的思路 [28–32]。Rogers 等 [33] 将随机散射掩模与衍射理论结合，实现了单传感器声源方位角辨识。Kruizinga 等 [34] 利用随机掩模对三维空间的波场编码，首次实现了超声频段的单传感器三维压缩成像。Xie 等 [35] 提出一种包含多个亥姆霍兹共鸣器阵列的声源辨识装置，该装置通过多个亥姆霍兹共鸣器实现了不同方向入射声信号的传输编码，实现了二维空间内的声源定向与分离。Jiang 等 [36] 提出了一种空间折叠机构，并基于该结构搭建了声学相机装置原型，实现对空间平面多声源的辨识与成像。

1.2.2 声人工结构的自动化设计

无论是基于声人工结构的声场调控，还是在空域和频域中实现声波的传输编码，结构设计均是其中的核心问题。目前，声学结构的设计首先需要基于经典的物理模型和实践中积累的经验与直觉来确定模型的框架。例如，改变声子晶体的点阵排列、周期尺寸可以实现对能带的调节；共振腔的单极子共振可以对结构等效模量进行调节；膜结构和穿孔板等偶极子振动能够对结构等效密度进行调节等等。确定结构模型后，在波动声学理论的基础上对模型进行简化，然后采用集总参数模型 (Lumped-parameter Technique, LPT)、传输矩阵模型 (Transmission Matrix Method, TMM)、耦合模式理论等方法对结构响应进行求解。虽然这些经典物理模型为声学结构的设计提供了指导方针，但是要确定结构精确的几何参数并获得期望的声学响应并不是一件容易的事情，尤其当结构的几何形状和空间排列都比较复杂的时候。这种情况下，我们通常借用 FDTD, FEM 等数值仿

真方法进行求解，在明确边界条件的基础上对时域和空间结构进行离散化处理，通过建立足够精细的网格计算给定结构的声学特性。

对于未被设计并优化过的声学结构，在基于上述方法得到的声学特性往往无法满足需求，因此我们需要对结构的几何参数进行调整，通过反复计算、对比和筛选最终获得具有期望响应的声学结构。这样的设计过程高度依赖于研究人员的工作经验，对计算资源和时间成本也造成极大的消耗。因此人们尝试使用优化类的算法在参数空间中自动化地寻找具有期望响应的声学结构。常用的优化算法包括拓扑优化、伴随方法、水平集方法等基于梯度的优化方法和遗传算法（Genetic Algorithm, GA）、粒子群算法（Particle Swarm optimization, PSO）等启发式优化方法 [37–41]。如图 1.2 所示为 Dong 等人基于拓扑优化得到的双负声学超材料的结构示意图 [39]；如图 1.3 为 Lu 等人结合 GA 和模拟退火算法对散射体形状和数量进行优化，利用 20 个不规则非均匀分布的离散散射体实现内部物体的定向声隐藏图 [40]。采用优化算法获得的物理结构的几何外形更加复杂，实现的声学响应也更加新奇。但是，优化过程依赖于大量的仿真计算和反复的参数迭代。并且每一次优化都是独立的，即使是对同一个声学结构，期望的声学响应发生改变优化过程也需要被重新执行，这会带来很大的时间成本和计算负担。

科学的研究的范式经历了实验科学、理论归纳以及仿真模拟的三个阶段，目前大数据和人工智能的结合被称为是“科学的第四范式”和“第四次工业革命”[42]。近年来，以人工神经网络为代表的机器学习在包括计算机视觉 [43]、自然语言处理 [10]、语音识别 [44]、知识图谱 [12] 等计算机科学及工程领域取得了突破性的进展，也以惊人的速度在材料科学 [45–47]，计算化学 [48]，凝聚态物理 [49, 50] 等其他学科领域展示了其独特的优势。此外，机器学习方法已经成为光学和电磁学结构设计的一个全新手段 [51–55]：例如建立深度学习网络，通过纳米结构的远场响应即可快速预测复杂亚波长光学结构的几何形状 [56]；利用全连接网络不仅实现了多层纳米粒子的散射场估计，还可以针对特定的散射特性进行纳米结构的逆向设计 [57]；通过神经网络学习，在可见光场下获得性能优越的隐形纳米粒子的逆向设计 [58]；向深度学习模型输入三波段的吸收谱，即可自动计算超表面结构分布 [59]。与传统方法相比，机器学习可以通过数据驱动的方式从大量的数据中自动发现和学习有用的信息，具备自主学习、联想储存、高速寻优等优点。声学和光学具有相似的波动方程，都可采用结构对波传播路径进行

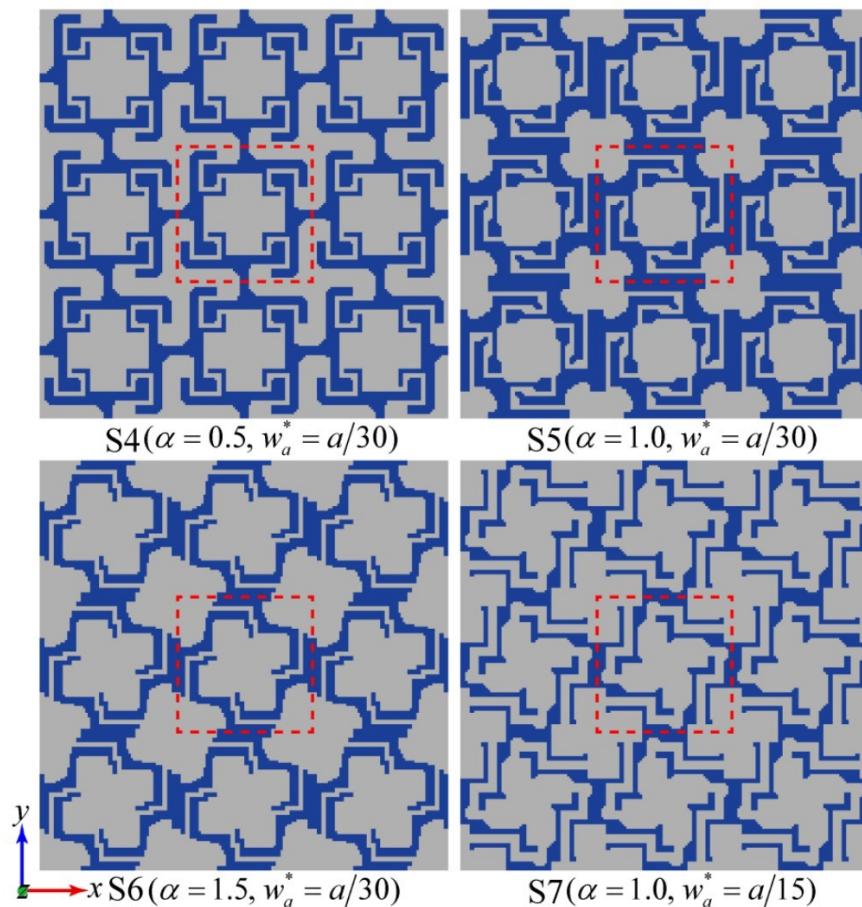


图 1.2 基于拓扑优化的双负声学超材料

Figure 1.2 Systematic design and realization of double-negative acoustic metamaterials by topology optimization

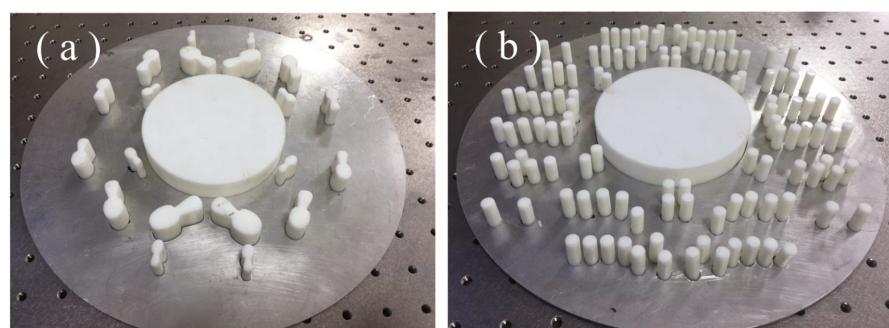


图 1.3 结合遗传算法和模拟退火算法的贝塞尔散射体优化

Figure 1.3 The optimization of Bézier scatterers based on the combination of the genetic algorithm and the simulated annealing

调控；两者传统的结构设计也采用相同的思路：确定模型→参数空间扫描→优化结构参数→筛选目标结构。因此，和光学结构设计类似，采用机器学习模型设计声学结构也可以略过传统设计的复杂流程，将基于经验的直观设计转变成基于数据模型的自动优化设计 [60–64]。设计同类的声学结构只需要一个机器学习模型，前期的仿真和实验数据提供一次数据投入即可满足后续多个结构的自主设计。

1.2.3 研究难点及关键科学问题

随着信息化时代的到来，声学系统也日趋智能化，在智能家居，视频会议，智能机器人，可穿戴设备等领域都有着广泛的应用。声学结构作为声学系统前端的重要组成部分也被赋予了更多的需求，声学结构的设计直接影响着声学系统的整体性能。一方面，我们需要解决如何对复杂声学结构进行快速精准化设计的问题，即声学结构的智能化设计的问题；另一方面，我们也在设计过程中考虑与后端信号处理算法的适配性，使整个声学系统的性能最优化，即声学结构的智能化应用的问题。

在声学结构的智能化设计方面，目前主要的技术难点在于如何在实现设计过程自动化的同时尽量降低计算成本。智能化声学系统中所使用的声学结构往往是十分复杂且不规则的，参数多、设计难度大。传统的人工设计方法需要设计者依据自身的专业知识和设计经验手动地对参数进行调节，且调节过程中还要兼顾参数之间的耦合关系，调节难度大，往往很难得到理想的设计效果。因此，声学结构的设计过程应提高设计效率，实现智能化设计。

在声学结构的智能化应用方面，目前主要的技术难点在于如何将前端的声学结构和后端的信号处理算法进行联合设计。在现有的声学系统设计中，声学结构的贡献大多是在前端传感部分，而声信号的处理工作是由后端算法完成的，两者通常是独立进行设计的。事实上，声学结构对声波在物理层面上的调制作用可以很大程度上减轻后端信号处理算法的工作量，从而提升整个声学系统的性能。因此，在对声学系统中的声学结构进行设计时，应充分考虑与后端算法的适配性。

1.3 研究内容与结构安排

本文针对声学结构的智能化设计与应用中所存在的技术问题，在对多种复杂声学结构进行理论分析和仿真计算的基础上，针对声学结构的物理特性和设计需求提出了与之适配的声学结构自动化设计策略。然而，单一的声学结构所实现的功能是有限的，想要实现智能化应用往往需要配合后端的信号处理算法构建智能化声学系统。因此，我们对前端声学结构和后端信号处理算法的联动性和相互作用展开讨论，在充分考虑二者之间的适配性的基础上，提出了一种基于声学结构的新型智能化传感系统，将前端声学结构的滤波作用与后端压缩感知算法进行了有机地融合，在仅使用一个传声器的基础上实现了三维空间中多个声源方位和音频内容的辨识。具体内容如下：

第1章介绍了论文工作的研究背景及意义，对声人工结构的国内外研究进展与发展现状进行了介绍和总结，并着重对声人工结构的自动化设计方法及其在传感领域的应用研究做了详细的介绍，最后对本文的主要内容和结构安排进行了说明。

第2章对GA的原理进行了基本的介绍，并结合具体的声学结构提出了基于GA的声学结构优化方法并展开讨论。首先将GA与层状严格耦合波分析(Layered Rigorous Coupled-wave Analysis, LRCWA)相结合，对基于多孔材料的吸声体的分层占空比和表面形状进行了优化，显著提升了吸声体在目标频段的吸声性能；然后又基于Norris教授提出的五模材料变换声学理论，对基于五模材料的隐身器件的分层方式进行了优化，并对优化前后隐身衣的隐身性能进行了对比分析，进一步证明了所提出方法的有效性。

第3章将深度学习模型和集总参数模型相结合，提出了一种高效且通用的声学结构逆设计方法，并基于该方法在给定几何参数范围和工作频段的基础上，完成了对二阶亥姆霍兹共鸣器(Two-order Helmholtz Resonator, THR)的几何参数的自动化设计，以实现特定的声学特性。最后，基于上述物理模型和设计方法，完成了低频宽带隔声器件的设计和性能验证。

第4章中将深度神经网络与其他算法相结合，探索了深度学习在声学结构设计领域的拓展性应用。深度学习模型可以与降维类算法结合，实现声学结构的批量化设计；同时，在解决结构优化问题上，深度学习模型还可以为优化类算法提供良好的初始条件，提高优化效率，在一定程度上避免其陷入局部最优解，提

升最终的优化效果；最后我们将所提出的深度学习模型在 LPT 框架下向其他声学结构进行迁移，证明了该设计方法的灵活性和普适性。

第 5 章提出了一种基于声学结构的单通道智能化声学传感系统。该系统在原理上借鉴了人类的听觉系统，并在对声学结构与算法的设计过程中充分考虑了二者的联动性和相互作用，在只使用一个传声器的基础上实现了三维空间内多个声源的辨识。

第 6 章对全文进行了总结，并对本研究目前存在的问题和下一步工作计划进行了讨论和展望。

第2章 基于遗传算法的声人工结构优化

2.1 引言

通常情况下，智能声学系统由于功能较为复杂，所使用的声学结构也是复杂且不规则的。而传统的声学结构设计方法主要针对的是形状较为规则且参数量比较少的简单结构，设计者可以基于一些经典的物理模型和自身的专业知识、设计经验来确定大体的结构框架，再手动调整参数来实现期望的声学响应。但是这个过程高度依赖于研究人员的专业知识和设计经验，手动调整参数设计效率低，很难用于设计不规则的复杂结构。因此，对于较为复杂的结构，我们通常会借助时域有限差分（Finite-Difference Time-Domain, FDTD），有限元仿真（Finite Element Method, FEM）等数值仿真方法对结构的物理特性进行正向求解，并与各种类型的优化算法相结合，在参数空间进行自动化搜索。本章节中，我们将以较为经典的遗传算法（Genetic Algorithm, GA）为例，结合具体的设计案例，实现复杂声人工结构的优化设计。

基于优化类算法对声学结构进行设计的本质是通过不断迭代的正向计算来逐步逼近期望的结果，在具体实施的过程中，我们可以根据所采用的正向计算方法的不同把优化类方法分为两类：当所设计的声学结构相对简单时，结构可以被解析化地建模，从而使得我们可以基于公式对声学结构的物理性质进行正向计算。在 2.3 章节中，我们将 GA 与层状严格耦合波分析（Layered Rigorous Coupled-wave Analysis, LRCWA）算法相结合对吸声棉的形状进行优化，以提升其吸声效果；然而，当我们所设计的声学结构较为复杂时，我们很难通过解析的公式对其进行正向计算，这时我们往往会求助于 COMSOL Multiphysics 等有限元仿真工具，同时利用软件内置的 MATLAB 接口对声学结构进行优化。在 2.4 章节中，我们将该方法应用于声学隐身器件的分层优化，以提升其隐身效果。

2.2 理论基础

2.2.1 遗传算法简介

优胜劣汰，适者生存是自然界的进化规律，GA 就是在这一规律基础上抽象而成的一种近似随机的搜索方法 [65, 66]。传统优化算法往往需要对目标函数进

行求导，对目标函数有连续性要求。GA 则可以解决未知形式的函数优化问题，采用概率性寻优的方式，自适应地确定每一步的优化方向。GA 最初是借鉴了进化生物学中的一些现象而发展起来的，这些现象包括自然选择 (Selection)、交叉 (Crossover) 以及变异 (Mutation) 等。GA 是通过随机生成一组目标函数作为父代样本，经过自然选择确定保留以及淘汰的个体，随后再进行杂交以及变异获得子代样本。与自然进化进程类似，子代个体比父代个体获得更好的适应性。由于 GA 具有上述优点，已被广泛地应用于声学结构设计领域解决实际的优化问题，它也是现代智能计算中的关键技术 [67]。

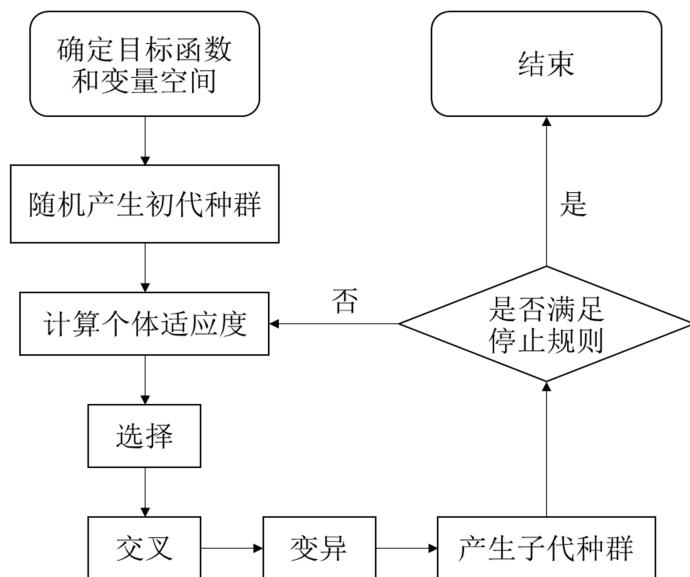


图 2.1 GA 流程示意图

Figure 2.1 Flow chart of GA

在 GA 的具体实现上，我们可以调用一些开源的数据库，也可以自行进行编写。本章中所使用的 GA 均是在 MATLAB 平台上基于英国 Sheffield 大学推出的遗传算法工具箱实现的 [68]，算法流程如图 2.1 所示，其核心步骤如下：

1. 染色体的编码表示

GA 首先需要将待搜索的参数通过编码的方式形成若干个个体，然后借助个体间自然选择的进化过程来完成后续的参数搜索。在进化过程中，每个个体均可视为优化问题的一个备选解，个体的性状对应优化问题所搜索的变量，每一个变

量以基因的形式存在于染色体上。因此，使用 GA 来解决优化问题，一个重要的环节就是对备选解进行合理地描述，即对变量编码。

GA 的经典编码形式是二进制编码，对于含有 n 个变量的优化问题，每个变量用 d 维比特串编码，这就要求将有限的连续搜索空间映射到离散规划，即满足 $\phi : \mathbb{R}^n \rightarrow (0, 1)^{n \times d}$ 。标准的二进制编码可以把个体 $\phi : \mathbb{R}^n \rightarrow (0, 1)^{n \times d} x = (x_1, \dots, x_j, \dots, x_n), x_j \in \mathbb{R}$ ，转换为二进制个体 $b = (b_1, \dots, b_j, \dots, b_n), b_j = e_{(j-1)d+1} \dots e_{jd}, e_l \in \{0, 1\}$ 。在这样的映射条件下，搜索的离散空间成网格状，参数 x_j 最大可获得的精度 $\Delta x_j (j = 1, 2, \dots, n)$ 为：

$$\Delta x_j = \frac{x_{j,\max} - x_{j,\min}}{2^d - 1}. \quad \dots (2.1)$$

其中， $x_{j,\max}$ 和 $x_{j,\min}$ 分别为变量 x_j 的最大值和最小值。

二进制编码一个最明显的不足是会造成汉明悬崖（Hamming Cliff），即相邻的参数值用比特串表示时相互远离。例如，对于十进制数 7 和 8，相应的 4 位比特串为 0111 和 1000，汉明距离为 4；假设 7 为全局最优适应度，当前求得的最优适应度为 8，若使算法能从 8 搜索到 7 则需要改变 4 个比特位。格雷编码可以有效解决二进制编码的汉明悬崖问题，确保了相邻数值的汉明距离均为 1。二进制编码与格雷编码满足如下转换关系：

$$\begin{aligned} g_1 &= e_1 \\ g_l &= e_{l-1}\bar{e}_l + \bar{e}_{l-1}e_l. \end{aligned} \quad \dots (2.2)$$

其中， e_l 为二进制编码的第 l 位， g_l 为格雷编码的第 l 位。 \bar{e}_l 表示非 e_l ，+ 表示逻辑或，乘表示逻辑与。

2. 初始种群

确定了编码方式后，我们需要在可行域上产生一个初始种群。当我们对该优化问题没有先验知识时，通常会随机产生初始种群；而如果我们掌握了一定的先验知识，我们可以在初始种群中加入一定量的精英个体，用于加速种群的进化。本章中所介绍的优化案例中的初始种群均为随机产生，4.2 章节中则介绍了带有精英个体的优化案例。此外，初始种群的大小也直接关系到优化效率和优化结果，种群含有的个体越多，多样性越丰富，也就增强了最优解被搜索到的概率，从而减少了迭代次数，但会因此增加每一代的计算量；相应的，若种群较小，每

一代的计算量减小了，不仅迭代次数会增多，陷入局部最优的风险也升高了。

3. 适应度函数

自然进化的原则是适者生存，即适应度高性状优良的个体有更大概率存活下去并参与繁殖下一代。相应的，在GA中，我们引入适应度函数来描述个体的优劣程度。适应度越高表明该个体与设计目标越吻合，则存活几率越大，反之则将面临被淘汰的危险。GA的重要一环就是对适应度函数进行设计，若适应度函数选取不当，有可能会在进化过程中出现欺骗。例如，对于一些具有高适应度的个体，他们较强的竞争力可能会影响进化的进程，使得种群中的个体差异性逐渐缩小，导致优化收敛到局部最优解。

适应度函数通常是单值的，且能够有效地反映出种群个体的优劣程度。在声学结构优化问题中，适应度函数通常是通过对声学结构的物理特性进行正向计算得到的。考虑到在每一次迭代的过程中，我们需要对种群中每一个个体计算其适应度，因此适应度函数的计算复杂度越低，优化效率越高。

4. 选择

选择算子体现了进化论模型中适者生存的理念，选择的目的是获得性能更加优良的下一代，在进化算法中分两个步骤来完成：

(a) 父代繁殖子代：父代个体的染色体经过交叉，保留下适应度较高的优良个体；之后变异算子会有很大可能作用在“劣势”基因，提高该个体的适应度和生存能力。

(b) 新种群的生成：每一次进化结束后会生成新的子代个体，组成最终的新种群，作为下次进化的附带个体。新种群也可以包含原先的父代个体，保证优良的个体可以持续存活。

选择算子以选择压力为特征，选择压力指的是选择算子单独反复作用使得最优解占据整个种群的速度。我们可以对不同的问题选取不同的选择算子，常见的选择算子包括轮盘赌选择，随机竞争选择，最佳17保留选择和均匀排序选择等。其中，轮盘赌选择是最为经典的一种选择方法，其基本思想是保证每个个体被选中的概率与其适应度大小成正比。本章所涉及的优化问题的选择算子均为轮盘赌选择。

5. 交叉

在自然进化的过程中，交配行为使两个同源染色体互相交换部分基因，繁殖

出下一代。相对应的，GA 中引入了交叉算子。在交叉运算前还需要为被选择的个体进行两两配对，随机配对是目前较为常用的配对策略。随后再通过控制交叉点的位置和交叉方式来完成整个交叉操作。本章所采用的是最基本的单点交叉，此外还有多点交叉，启发式交叉等。

6. 变异

同样是借鉴了生物遗传和进化过程中的变异现象，GA 中的变异操作是以较小概率改变个体编码串的某一位或某几位的值，如将二进制编码中“0”和1互换，形成新的基因。变异操作本身是随机的，但可以有效避免由于选择和交叉运算而造成的某些信息（基因）的丢失，使进化更加完备。选取合适的遗传算子可以提高算法的搜索能力，交叉算子掌控全局搜索，而变异算子则会影响算法局部搜索能力。且上述两个算子可以使群体一直保有种群多样性，避免在局部最优解处提前收敛，出现早熟等现象。

GA 中常常包含很多预设参数，合理选择预设参数有助于提高算法的优化性能，主要的预设参数包括：

1. 交叉概率 P_c

影响交叉操作的频繁程度，过小会降低解空间的搜索能力，使种群的进化速度缓慢；过大则会破坏群体进化的优良模式，使进化变得盲目而随机。一般建议 P_c 的取值范围是 0.4~0.99。

2. 变异概率 P_m

影响变异操作的频繁程度，过小会降低种群产生新个体的能力，在进化过程中容易收敛到局部极值；过大则会使种群过于多样，进化模式遭到破坏。所以在求解过程中常根据进化进程改变 P_m 的大小，例如在进化初期需要大范围搜索最优解，所以应选用较大的 P_m ；在进化后期需要加速收敛，一般推荐使用较小的 P_m 。变异概率 P_m 一般在 0.0001~0.1 范围内取值。

3. 群体规模 N

关系到算法的收敛性和计算效率，群体过小会导致算法收敛到局部最优解，规模过大则会带来很大的计算量，影响优化效率。 N 的取值为 10~200 最佳。

很多实际中遇到的优化问题常含有多个优化目标，对这些目标的优化往往是相互制约的。多目标优化问题可以有多种解决策略，这里主要介绍较为常用的权重系数法和并列选择法：

1. 权重系数法

对于一个多目标优化问题，若给每一个子目标函数 $f_i(x) (i = 1, 2, \dots, n)$ 赋予权重 $w_i (i = 1, 2, \dots, n)$ ，则多目标优化问题的总评价函数为 $u = \sum_{i=1}^n w_i f_i(x)$ 。

2. 并列选择法

并列选择法为每一个子目标函数都分配均等大小的子群体，每一个子群体在进化过程中都是相互独立的，再对每一个子群体的选择结果进行整合，使他们之间可以相互交叉，并完成变异等操作，从而产生新一代子群体。

2.2.2 COMSOL Multiphysics with MATLAB 简介

当我们所要设计的声学结构较为复杂时，我们常常需要借助有限元的方法对其进行仿真计算，而 COMSOL Multiphysics 便是较为常用的一款商业有限元仿真软件。尽管 COMSOL Multiphysics 在仿真方面功能很强大，但是由于软件内部无法输入编程语言，因此如果我们想对软件中的物理模型进行优化，就需要调用软件与 MATLAB 的接口 LiveLink for MATLAB。

由于 GA 每一代都会产生新的个体，他们都对应了一组新的结构参数。因此我们首先需要根据该结构参数在 COMSOL Multiphysics 中进行建模，这一过程就需要基于 LiveLink for MATLAB 接口来实现。在这一步中，我们常常需要通过 for 或 while 语句执行嵌套的循环，通过 if 或 switch 语句进行条件式建模，或者使用 try 和 catch 来处理异常等；随后，我们需要借助 COMSOL Multiphysics 对上述结构进行仿真计算，得到该结构的吸声系数、透射谱等物理特性，并将其通过 LiveLink for MATLAB 接口传递回 MATLAB，并作为可以参与后续计算的变量。最后，我们可以基于所得到的物理特性来计算该个体的适应度，并参与下一轮进化。

值得注意的是，LiveLink for MATLAB 接口的大部分功能是基于 java 实现的，但是这并不需要使用者有很强的 java 基础。一个较为快速地掌握 API 的调用方法是从 COMSOL Desktop 将模型保存成 M 文件，在 M 文件中寻找相关的命令语句。默认情况下，M 文件包含模型的完整命令历史。如果希望只包括当前模型中有效的命令，可以在保存 M 文件之前进行压缩模型历史的操作。

2.3 宽频吸声体的形状优化

多孔材料内部充满狭小的空隙，声音在其中传播引起空气分子的振动，通过狭小空隙中空气的粘滞效应可以大幅度将声能转化为热能，达到高效、宽频吸声的目的。多孔材料具备吸声效果优异、生产成本低等优点，是厅堂设计、轨道交通以及制造业中常用的吸声材料。例如，我们可以将不同形状的多孔材料作为吸声体安装在墙体上，以调节房间的混响效果等。与平面多孔材料相比，具有周期波纹结构的多孔材料工作频段更宽，往往具有更好的吸声效果。而吸声体的形状会直接影响吸声体的吸声性能，因此，在实际应用中，我们需要根据具体需求对吸声体的形状进行优化设计，以实现我们期望的吸声效果。本章节将 LRCWA 与 GA 相结合，对基于多孔材料的吸声体的分层占空比和表面形状进行了优化，显著提升了吸声体在目标频段的吸声性能。与传统的设计方法相比，所提出的优化方法具有简单高效的特点，并且可以拓展至其他种类的吸声体设计。

2.3.1 层状严格耦合波分析

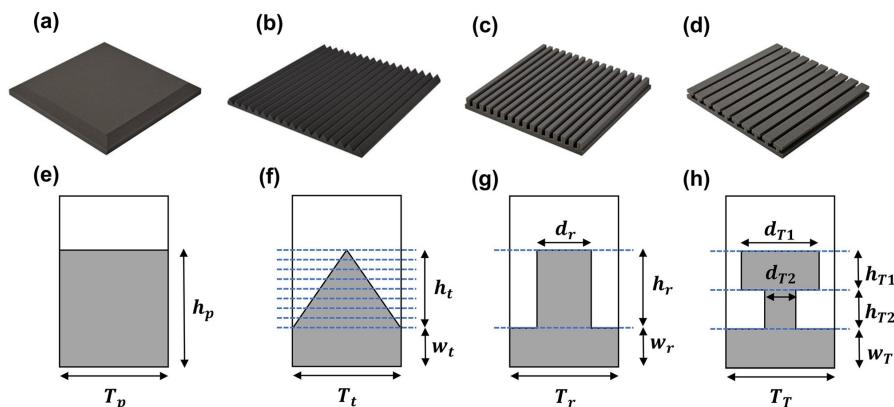


图 2.2 具有 (a) 平面, (b) 三角形, (c) 矩形和 (d)T 形表面的多孔材料照片, 及其对应的横截面 (e) 平面, (f) 三角形, (g) 矩形和 (h)T 形。

Figure 2.2 Photos of the sound-absorbing sponges with surfaces of (a) plane, (b) triangle, (c) rectangle and (d) T shape. Cross sections of the sound absorbing sponges with surfaces of (e) plane, (f) triangle, (g) rectangle and (h) T shape.

如图 2.2 所示为常见的几种不同的周期非平整界面多孔材料的照片及其对应的横截面示意图。目前，对于这种周期非平整界面多孔材料吸声性能的研究主要以仿真计算和实验测试为主。常用的仿真计算方法包括边界元、FEM 以及 FDTD 等方法，这些方法大多涉及几何建模和网格划分，且三维计算量大，对计算资源

要求较高。而实验测试通常需要制备大量不同材料参数、不同几何外形和不同结构参数的样件，进行大量重复的测量工作，然后对所有测试数据进行归纳统计并总结规律。这样的实验测试流程操作繁琐、耗时耗力，也不利于工程应用中的快速优化设计。在对吸声体进行优化设计时，需要进行大量的吸声系数计算，因此计算的准确度和复杂度是至关重要的。针对这一需求，Yang 等提出了具有较低复杂度和较高精确度的 LRCWA 方法用于周期非平整界面的吸声体的吸声系数计算 [69]，算法流程如图 2.3 所示，具体的算法细节在此不再赘述。

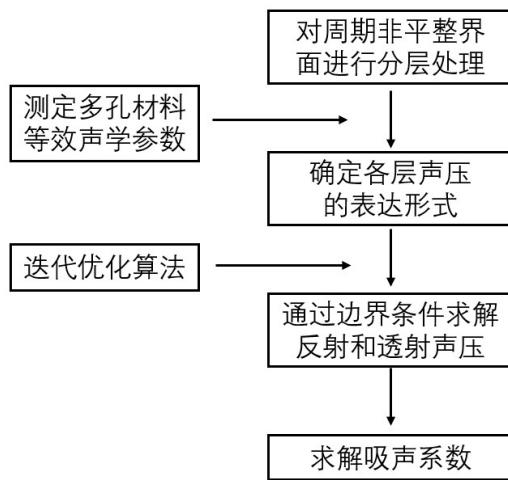


图 2.3 基于 LRCWA 方法计算非平整界面的吸声体的吸声系数

Figure 2.3 Absorption performance of nonplanar periodic structures solved by LRCWA

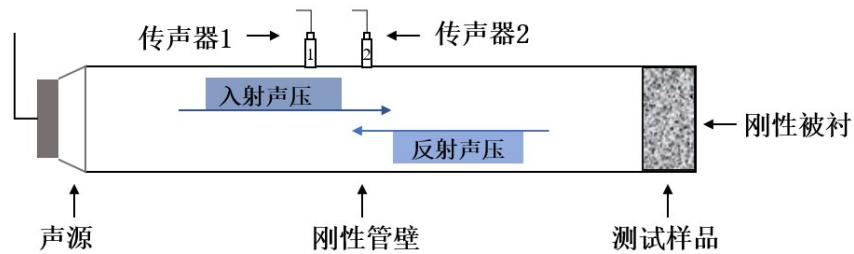


图 2.4 多孔材料等效声学参数测量装置示意图

Figure 2.4 Schematic diagram of the measuring system

LRCWA 算法需要已知基底多孔材料的等效折射率和等效密度，因此我们首先通过如图 2.4 所示的二传声器阻抗管系统测得了厚度分别为 2 cm 和 4.8 cm 的某多孔材料的反射系数，再通过数值计算得到了该多孔材料的等效折射率和等效密度，分别如图 2.5(a) 和图 2.5(b) 所示。

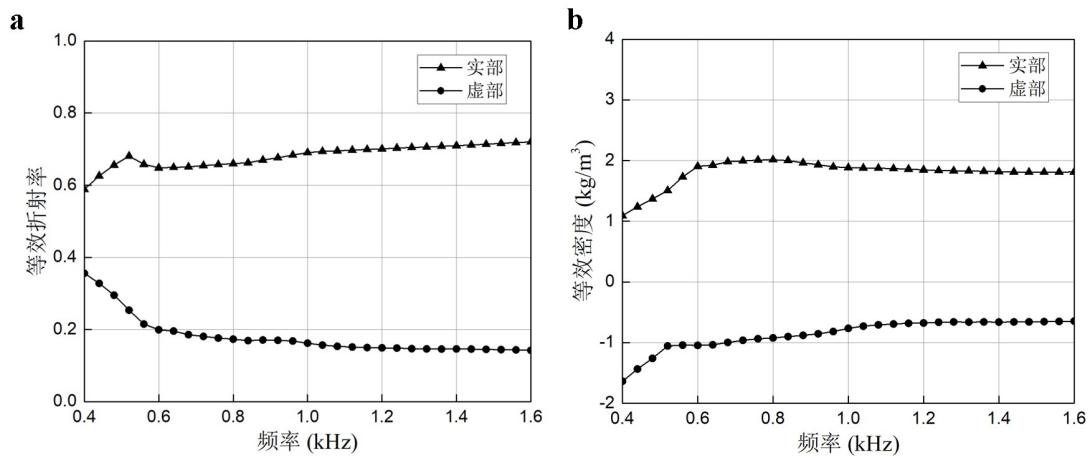


图 2.5 某多孔材料的 (a) 等效折射率和 (b) 等效密度

Figure 2.5 Real and imaginary parts of (a) the effective refractive index and (b) the effective mass density of the absorbing sponge

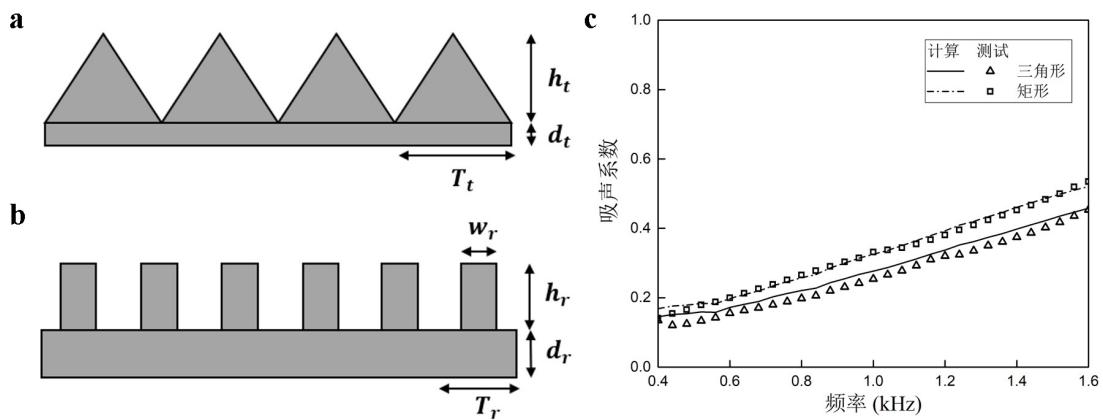


图 2.6 矩形和三角形多孔材料的结构示意图及吸声系数

Figure 2.6 Diagrams and absorption coefficients of the absorbers with triangular surface and rectangle surface

为了证明 LRCWA 方法用于计算周期非平整界面多孔材料吸声系数的有效性，我们分别对较为常见的三角形周期调制的吸声棉和矩形周期调制的吸声棉的吸声系数进行了计算，并与实验结果进行了对比。如图 2.6(a) 所示为三角形周期调制的吸声棉的结构示意图，该结构的周期 $T_t = 5 \text{ cm}$ ，基底厚度 $d_t = 1 \text{ cm}$ ，三角形尖端厚度 $h_t = 3.7 \text{ cm}$ ，其吸声系数的数值计算结果如图 2.6(c) 中实线所示，其吸声系数的实验测试结果如图 2.6(c) 中的三角形散点所示；如图 2.6(b) 所示为矩形周期调制的吸声棉的结构示意图，该结构的周期 $T_r = 3.3 \text{ cm}$ ，基底厚度 $d_r = 1.9 \text{ cm}$ ，矩形尖端厚度 $h_r = 2.8 \text{ cm}$ ，矩形尖端宽度 $w_r = 1.5 \text{ cm}$ ，其吸声系数的数值计算结果如图 2.6(c) 中虚线所示，其吸声系数的实验测试结果如图 2.6(c) 中的矩形散点所示。通过观察图 2.6(c) 中的实验和仿真结果可知，基于 LRCWA 方法计算得到的吸声系数与实验测试结果十分接近，证明了 LRCWA 方法的有效性。而与 FEM、FDTD 等传统的数值计算方法相比，LRCWA 方法计算效率更高，更适用于宽频吸声体的优化任务。

2.3.2 吸声体分层占空比优化

LRCWA 方法本质上是将周期非平整界面的吸声体分成许多薄层，使得每一层可以近似为矩形周期调制。这时，每一层的声传播都可以等同为单层周期结构中的声传播进行处理。然后，利用边界连续性条件确定层间耦合关系，建立整个结构的声学耦合方程，以此求解周期非平整界面吸声体的反射系数、透射系数和吸声系数。因此，吸声体每一层的占空比大小都会直接影响吸声体的整体吸声性能。

这里，我们将吸声体均匀分成了 30 层，并基于 GA 对每一层的占空比进行了优化，以提高 2.9 kHz 附近的吸声性能。吸声体每一层的占空比 $x_i (i = 1, 2, \dots, 30)$ 可以视为 0~1 之间的待优化变量，对应 12 位的二进制编码。种群采用随机初始化，规模为 100。经过 50 次迭代后算法收敛，此时的最优个体的结构示意图如图 2.7(a) 中的灰色插图所示。我们分别基于 LRCWA 方法和 FEM 方法对最优个体的吸声曲线进行了计算，分别对应图 2.7(a) 中的黑色曲线和红色曲线。在图 2.7(b) 中，我们还将最优个体（绿色实线）的吸声性能与较为常见的平板结构（黑色实线），矩形结构（红色实线），三角形结构（蓝色实线）和椭圆结构（紫色实线）进行了对比。通过观察上述计算结果可知，优化后的吸声体在 2.9 kHz 附近的吸声系数在 0.9 附近，显著高于上述常规结构在此频带附近的吸声系数。

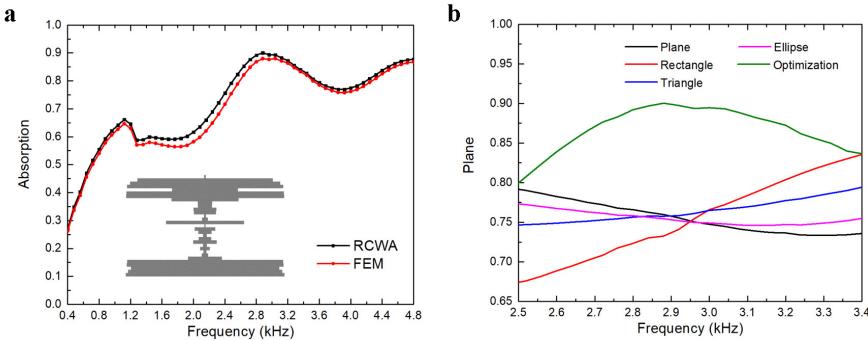


图 2.7 优化后的吸声体的吸声系数

Figure 2.7 Absorption coefficients of the optimized structure

2.3.3 吸声体表面形状优化

在上一小节中，我们并没有对吸声体的几何形状做任何的假设，而是采用了最朴实的思路直接对吸声体每一层的占空比进行了优化。因此，最终得到的吸声体的形状十分不规则，若不加以约束很难对其进行加工。在这一小节，我们希望能够兼顾结构的吸声性能和可加工性，直接对吸声体的表面形状进行优化，希望能得到表面连续、形状较为规则的吸声体，并在此基础上最大化 60 Hz 至 400 Hz 频段内吸声体的最低吸声系数。同时，为了凸显表面形状对吸声性能的影响，我们在优化过程中始终保持吸声棉的总厚度为 1 m，周期 $T = 30$ cm。

首先，我们针对较为常见的三角形表面吸声体进行了优化设计，待调节参数为吸声体的空气背腔 t_1 和基底厚度 t_2 ，调节范围均为 0~50 cm。考虑到这里的参数个数只有 2 个，因此可以直接基于 LRCWA 算法在上述范围内进行网格搜索，最终得到的最优化参数为 $t_1 = 25.0$ cm 和 $t_2 = 46.8$ cm，最优结构如图 2.8(a) 所示，该结构与消声室中的尖劈结构十分相似。该结构对应的吸声曲线为图 2.9 中的红色曲线和三角形散点，分别对应基于 LRCWA 和 FEM 的计算结果。

一个更不失一般性的建模方式是将吸声体与空气接触的表面在二维情况下视为一条曲线并对其进行高斯展开，这里我们将其建模为 5 阶高斯函数的叠加：

$$y = \sum_{i=1}^5 a_i e^{-(x-b_i)^2/2c_i^2} \quad (-2.5 \ll x \ll 2.5). \quad \dots (2.3)$$

其中， a_i ， b_i 和 c_i ($i = 1, 2, \dots, 5$) 为待优化系数，共 15 个。值得注意的是，由于我们需要控制吸声结构的整体厚度不变，因此在优化过程中我们还需要结合背腔的厚度对 y 进行缩放。随后，我们便基于 GA 对这 15 个系数及背腔厚度 t_1 进行

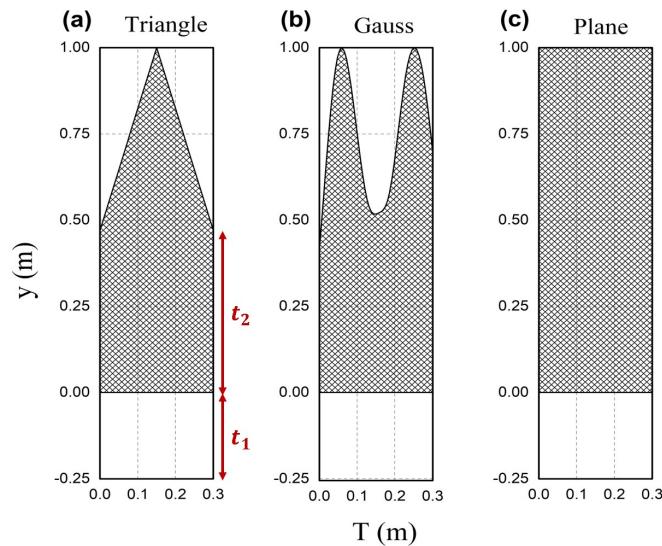


图 2.8 具有不同表面形状的吸声体的结构示意图: (a) 三角形; (b) 高斯形; (c) 平面

Figure 2.8 Optimized structures with (a) triangular surface, (b) surface formed by the super-position of Gaussian functions and (c) the planar structure for comparison.

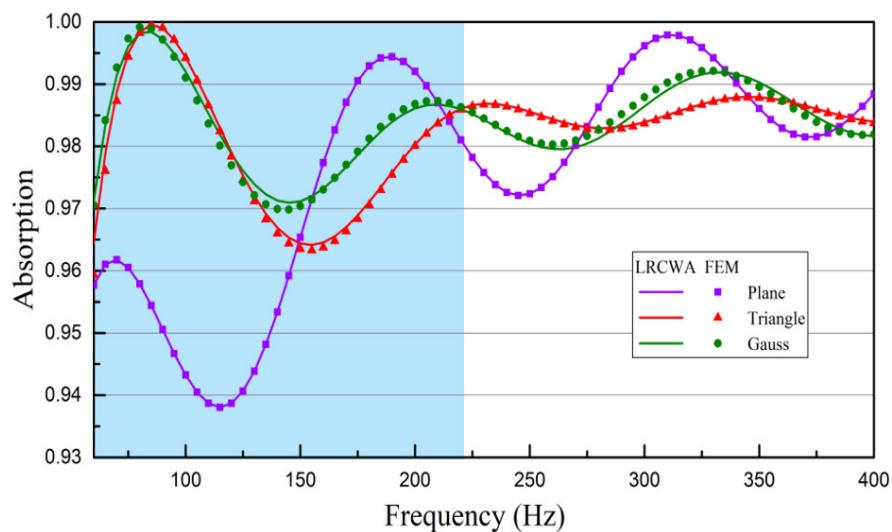


图 2.9 具有不同表面形状的吸声体的吸声曲线

Figure 2.9 Absorption coefficients of the optimized structures with different surface

了寻优，调节范围为 $0 < a_i < 1$, $-2 < b_i < 2$, $0 < c_i < 1$, $0 \text{ cm} < t_1 < 50 \text{ cm}$ 。最终得到的 15 个系数的最优值如表 2.1 所示，最优的背腔厚度为 $t_1 = 25.5 \text{ cm}$ 。该结构的示意图如图 2.8(b) 所示，对应的吸声曲线如图 2.9 中的绿色曲线和圆形散点所示，分别对应基于 LRCWA 方法和 FEM 的计算结果。

表 2.1 优化后高斯形状表面的系数

Table 2.1 Gaussian coefficients after optimization

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
a_i	0.3168	0.7706	0.0703	0.1242	0.4786
b_i	-1.2453	1.6993	0.1294	-0.4320	-1.7344
c_i	0.7335	0.9426	0.3073	0.5543	0.6673

为了方便对比，我们还计算了如图 2.8(c) 所示的具有平整表面的矩形吸声体的吸声曲线，背腔厚度为 $t_1 = 25.0 \text{ cm}$ 。计算结果如图 2.9 中的紫色曲线和方形散点所示，分别对应基于 LRCWA 方法和 FEM 的计算结果。通过对图 2.9 中三种不同形状吸声体的吸声曲线，我们发现具有非平整表面的吸声体的低频吸声性能要显著好于具有平整表面的吸声体（对应图 2.9 的蓝色阴影区域）。此外，与较为常用的三角形尖劈结构相比，优化后的具有高斯形状的吸声结构在 150 Hz 以下频段的吸声性能大体不变，而在 150 Hz~200 Hz 频段的吸声性能得到了显著的提升。

2.4 声学隐身器件的分层优化

1995 年，美国犹他大学的 Milton 教授和 Cherkaev 教授首先提出五模材料 (Pentamode Material) 这一概念，该材料在形变时仅能承受主轴方向的应力，与流体的“流动”颇为相似，因此受到了研究者的广泛关注 [70]。2008 年，美国罗格斯大学的 Norris 教授提出了基于五模材料的变换声学理论，完成了利用五模材料对声波进行调控的理论分析，也激起了广大科研工作者对五模材料及其在声波控制领域中的研究热情 [71]。2012 年，Gokhale 等在 Norris 所提出的五模材料变换理论基础上，针对环形隐身衣，研究了理想五模材料属性与变换映射之间的

联系，并给出常密度分布、常径向模量分布、常切向模量分布、级数密度分布、级数径向模量分布、级数切向模量分布以及最小各项异性模量分布等几种环形斗篷对应的映射函数 [72]。此后，五模材料在水下隐身 [6, 73]、梯度透镜 [74]、负折射成像 [75] 及水声复路通信 [76] 等领域均取得了进展。

原则上，基于五模材料的声学隐身衣应具有连续的物理参数。然而，受制备工艺等客观条件的限制，很难制备出密度和模量连续变化的五模材料。作为一种简化手段，实际制备时必须对隐身衣进行分层，层内的密度与模量是均匀分布的。然而，由于分层人为引入了交界面，对隐身衣性能必然会造成影响，因此研究分层对隐身衣性能的影响具有重要的理论及实际意义。Cai 等 [77] 研究了变换函数对分层惯性隐身衣隐身性能的影响。Scandrett 等 [78] 对五模材料球形隐身衣的密度和模量进行优化，以提升隐身性能。Zhang 等 [79] 研究了圆柱状五模材料隐身衣层数、层厚分布等分层因素对隐身衣性能的影响。然而，以上工作所涉及的分层方式均为均匀分层，而对于非均匀分层的情况，目前讨论并不多见。由于非均匀分层突破了分层厚度的限制，对隐身衣性能的影响多了一个自由度，因此有必要对其进行深入的研究，以便寻找最优的分层优化方案。

在本小节中，我们分析了分层方式对隐身效果的影响，提出了一种基于 GA 的五模材料分层优化策略，并借助 2.2.2 中所介绍的 COMSOL Multiphysics with MATLAB 对基于五模材料的声学隐身器件的隐身性能进行了优化。仿真结果证明了该策略对窄带探测信号和宽带探测信号均十分有效。所提出的分层优化策略具有易于实现，通用性强等优点，为其他类型的超材料器件设计提供了可行的技术途径。

2.4.1 五模材料的变换声学理论

二维空间中五模材料的弹性张量可以表示为：

$$\mathbf{C} = \mathbf{S} \otimes \mathbf{S}. \quad \dots (2.4)$$

其中， \mathbf{S} 是五模材料的特征应力张量，在主轴坐标系下有较为简单的对角矩阵形式：

$$\mathbf{S} = \begin{pmatrix} s_{11} & 0 \\ 0 & s_{11} \end{pmatrix} = \begin{pmatrix} \sqrt{K_x} & 0 \\ 0 & \gamma \sqrt{K_x} \end{pmatrix}. \quad \dots (2.5)$$

联立公式 (2.4) 和公式 (2.11)，可以得到弹性矩阵 \mathbf{C} 在主轴坐标系下有如下形式：

$$C = \begin{pmatrix} K_x & \gamma\sqrt{K_x K_y} & 0 \\ \gamma\sqrt{K_x K_y} & K_y & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad \dots (2.6)$$

根据系数 γ 的正负可以进一步定义正五模材料 ($\gamma = 1$) 和负五模材料 ($\gamma = -1$)，本文主要讨论正五模材料，因此如无特别说明，下文五模材料均指正五模材料。公式 (2.6) 中的 K_x 和 K_y 分别为五模材料 x 主轴和 y 主轴的压缩模量，在这两个主方向声波的相速度分别为：

$$c_x = \sqrt{\frac{K_x}{\rho}}, c_y = \sqrt{\frac{K_y}{\rho}}. \quad \dots (2.7)$$

根据以上分析，改变五模材料的模量或者密度，可以实现对声速的有效调节，这是五模材料声波调控的基础。

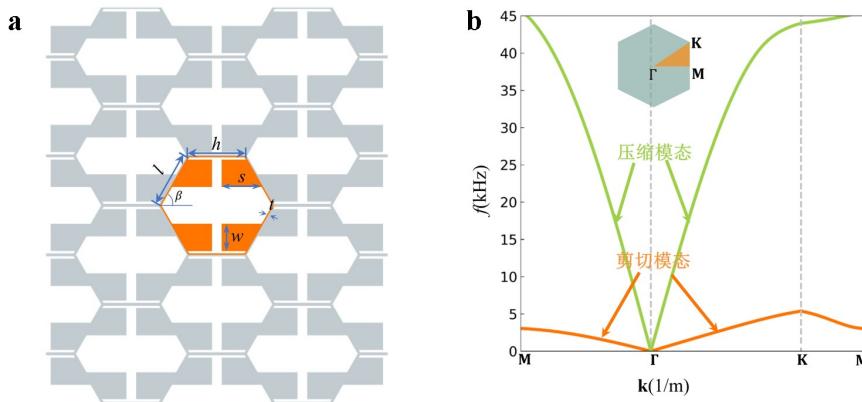


图 2.10 二维五模材料微结构和能带：(a) 二维五模材料微结构示意图；(b) 二维五模材料能带。

Figure 2.10 The unit cell and band diagram of pentamode materials: (a) Schematic diagram of the pentamode material unit cell; (b) Band diagram of the pentamode materials.

二维五模材料的微结构呈六边形，如图 2.10(a) 所示，其中橘黄色区域是一个完整的五模材料单元示意图。五模材料单元的基材一般采用硬质金属，如铝合金、钛合金等，六边形内部的空白填充空气。图 2.10(a) 中的五模材料微结构基材为铝合金，杨氏模量为 $E_s = 69$ GPa，泊松比为 $\sigma = 0.33$ ，密度为 $\rho_s = 2700$ kg/m³，单元的几何参数为 $l = h = 5$ mm, $w = 2.38$ mm, $s = 3.50$ mm, $t = 0.25$ mm，因此单元的等效密度为 $\rho_{eff} = 1263.48$ kg/m³。图 2.10(a) 中的五模材料的频散曲线图 2.10(b) 所示，其中橘黄色曲线代表结构的剪切模态，绿色曲线代表结构的压

缩模态。根据图 2.10(b) 中的能带可以计算出微结构压缩模态的等效声速分别为 $c_x = 805.34 \text{ m/s}$, $c_y = 1115.21 \text{ m/s}$, 剪切模态的等效声速分别为 $c_{sx} = 86.34 \text{ m/s}$, $c_{sy} = 73.35 \text{ m/s}$ 。因此, 通过计算可以得到该单元的等效弹性张量近似为:

$$C_{PM} = \begin{bmatrix} 0.698 & 0.504 & 0 \\ 0.504 & 0.364 & 0 \\ 0 & 0 & 0.0035 \end{bmatrix} B. \quad \dots (2.8)$$

其中, $B = 2.25 \text{ GPa}$ 是水的体模量。结合图 2.10(b) 中的频散曲线和公式 (2.8), 不难发现五模材料的剪切模量远小于压缩模量, 因此在很宽频率范围内结构中只存在压缩模态, 表现出近似流体的声学特性 [80]。通过以上分析可知, 五模材料微结构的等效参数能够很好地描述五模材料微结构的性能, 因此在设计五模材料器件的过程中, 常常先研究等效参数描述下五模材料器件的性能。

2008 年, Norris 教授基于空间坐标变换提出了基于五模材料的变换声学理论, 为五模材料的器件设计提供了理论基础。该理论建立了物理空间与虚拟空间的坐标变换关系, 通过对隐身衣材料进行参数设计, 来操控声波绕过隐藏区域最终实现声隐身。

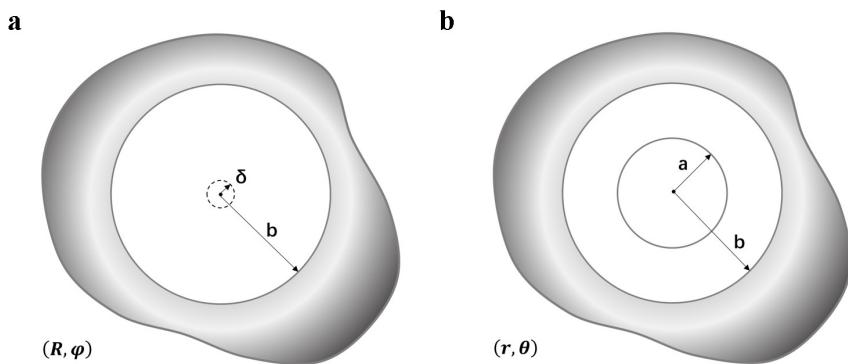


图 2.11 五模材料坐标变换示意图: (a) 虚拟空间 Ω ; (b) 物理空间 ω 。

Figure 2.11 Diagram of coordinate transformation of pentamode materials: (a) Virtual space Ω ; (b) Physical space ω .

图 2.11(a) 为二维环形五模材料的虚拟空间, 图 2.11(b) 为其物理空间, 灰色的背景区域均为相同的背景介质。虚拟空间和物理空间之间可以建立如下映射

关系：

$$R = \begin{cases} r & (r \geq b) \\ f(r) & (a < r < b) \\ \delta & (r = a) \end{cases}, \quad \varphi = \theta . \quad \dots (2.9)$$

由于虚拟空间和物理空间具有不同的拓扑属性，为了避免出现奇点，在虚拟空间的中心处挖半径为 δ 的小洞，映射的边界条件应满足 $b = f(b), \delta = f(a)$ 。映射梯度 \mathbf{F} 在极坐标系下可以表示为：

$$\mathbf{F} = \begin{pmatrix} \frac{1}{f'(r)} & 0 \\ 0 & \frac{r}{f(r)} \end{pmatrix}, \quad J = \det \mathbf{F} = \frac{r}{f'(r)f(r)}. \quad \dots (2.10)$$

由于 \mathbf{F} 是对称张量，可令特征应力张量满足：

$$\mathbf{S} = J^{-1}\mathbf{F}. \quad \dots (2.11)$$

假定整个虚拟空间和物理空间中隐身衣以外的区域均分布着密度为 ρ_0 模量为 K_0 的声学介质，隐身衣密度为 ρ ，径向模量为 K_r ，切向模量为 K_θ 。由于声波在两个空间遵从相同的物理定律，波动方程具有相同的形式，结合坐标变换法则可以计算出隐身衣材料参数的表达式：

$$\begin{aligned} \rho &= \rho_0 \frac{f(r)f'(r)}{r} \\ K_r &= K_0 \frac{f(r)}{rf'(r)} \quad K_\theta = K_0 \frac{rf'(r)}{f(r)} \end{aligned} . \quad \dots (2.12)$$

由式(2.12)可以看出，隐身衣的密度和模量分布取决于映射关系 $f(r)$ 。表2.2给出了常密度映射、常模量映射和线性映射3种具有代表性的坐标变换映射方式以及对应的参数分布。表中的密度 ρ 、径向模量 K_r 、切向模量 K_θ 均为与背景介质的密度 ρ_0 和模量 K_0 的相对值。

2.4.2 声学隐身器件的分层设计

表2.2给出了不同的映射方式对应的参数分布，对于设计五模材料隐身衣具有很好的指导意义。在利用表2.2所给出的参数设计隐身衣器件时，需要先对其进行分层，本文仅考虑标准环形隐身衣的分层优化问题。对某一层而言，层内的密度与模量视为均匀分布。以起始半径和终止半径的算术均值视为该层的参数半径，然后将其代入表2.2的映射关系中，则能够计算出该层对应的模量和密度。图2.12(a)为将二维五模材料隐身衣分为5层后的结构示意图，其中隐藏空间的

半径 $r_0 = 0.1$ m, 隐身衣的外径 $r_5 = 0.5$ m; 由内到外每一层的终止半径分别为 r_1 、 r_2 、 r_3 、 r_4 和 r_5 。2.12(b) (d) 分别为表 2.2 中 3 种映射关系下隐身衣的密度、切向模量和径向模量随半径变化的对比图。由上述 3 幅图可知, 无论是在常密度、常模量还是线性映射下, 五模材料隐身衣的材料参数都是随半径连续变化的, 且材料参数和半径之间并不都是简单的线性关系, 这意味着隐身衣的分层方式会对其最终的隐身效果造成一定的影响, 传统的均匀分层方式并不是最优的选择。因此, 为了得到最佳的隐身效果, 本章将以表 2.2 中的线性映射为研究对象, 对基于五模材料的声学隐身衣的分层方式展开研究。

表 2.2 几种典型的映射关系

Table 2.2 Mapping relations

映射类型	映射表达式	材料参数分布
常密度映射	$f(r) = \sqrt{Ar^2 + B}$ $A = \frac{b^2 - \delta^2}{b^2 - a^2}, \quad B = -\frac{b^2(a^2 - \delta^2)}{b^2 - a^2}$	$\rho = A, \quad K_r = \frac{f^2(r)}{Ar^2}, \quad K_\theta = \frac{1}{K_r}$
常模量映射	$f(r) = b\left(\frac{r}{b}\right)^A, \quad A = \frac{\ln(b/\delta)}{\ln(b/a)}$	$\rho = A\left(\frac{b}{r}\right)^{2-2A}, \quad K_r = \frac{1}{A}, \quad K_\theta = A$
线性映射	$f(r) = Ar + B$ $A = \frac{b-\delta}{b-a}, \quad B = -\frac{b(a-\delta)}{b-a}$	$\rho = A \frac{f(r)}{r}, \quad K_r = \frac{f(r)}{Ar}, \quad K_\theta = \frac{1}{K_r}$

2.4.3 最小化散射声能量

首先选择最小化隐身衣的散射声能量 E_s 作为优化问题的目标, 最小化隐身衣的散射声能量可以用于抵御声场中任何位置的探测。基于 COMSOL Multiphysics with MATLAB 进行算法优化, 计算不同分层方式对应的散射声能量, 并在此基础上构造适应度函数。算法对该优化问题展现出较好的鲁棒性, 并随迭代次数的增加逐渐收敛。图 2.13 为目标频率为 10 kHz 时算法的进化曲线, 在优化计算的过程中, 迭代超过 60 次时已经出现了最优解, 此后便是优势逐渐扩大的过程。

优化前后隐身衣的分层方式如表 2.3 第一列和第二列所示, 优化前为均匀分层, 而优化后的隐身衣每一层的厚度不再均匀。隐身衣优化前后的总声场和散射

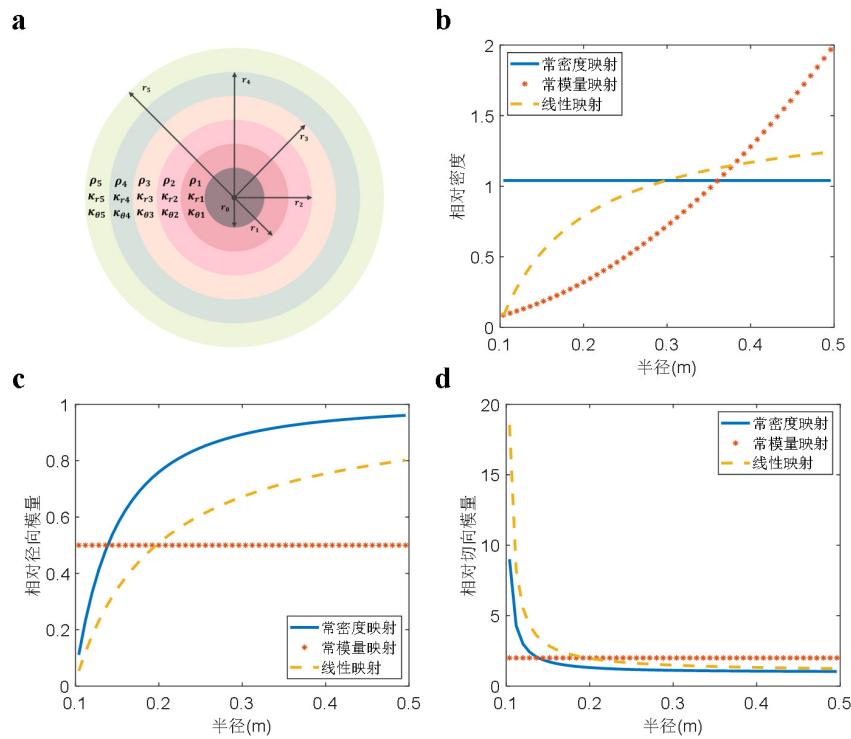


图 2.12 五模材料隐身衣分层示意图和材料参数分布

Figure 2.12 The diagram and material parameter distributions of pentamode materials

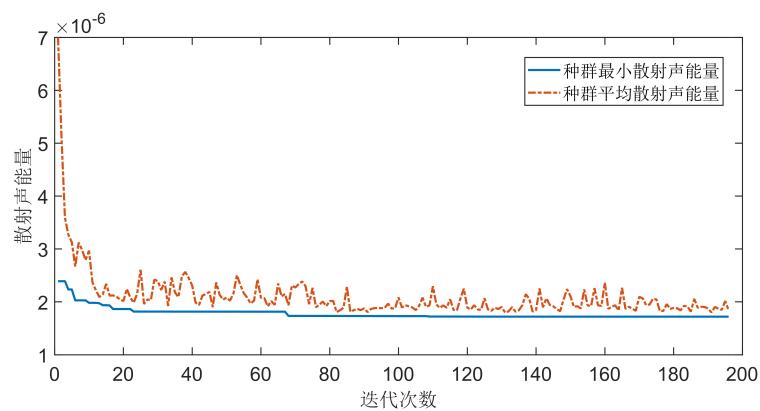


图 2.13 GA 进化曲线

Figure 2.13 Evolution curve of the GA

声场可以由 COMSOL Multiphysics 进行数值仿真得到，数值模拟的区域是一个边长为 3 m 的正方形区域，背景介质为水。入射波为从仿真区域左侧入射的 10 kHz 的平面波，总的入射声能量 E_{in} 可以通过入射声场对入射边界积分求得。为了模拟理想状态无限大流体介质中平面声波的传播特性，正方形区域的 4 条边均设置为“平面波辐射”边界条件。隐身衣的总散射声能量 E_s 可以通过散射声场对正方形区域的 4 条边界积分求得。图 2.14(a) 和 (b) 分别为优化前后总声场的仿真结果，图 2.15(a) 和 (b) 分别为优化前后散射声场的仿真结果。通过对比这 4 幅图可以较为直观地看出，优化前由于存在比较严重的散射效应，造成总声场的波阵面发生变形。而优化后散射声场被明显削弱，总声场的波阵面更加连续，与平面波声场更接近，隐身效果更好。

表 2.3 优化前后隐身衣每一层的半径

Table 2.3 The radius of each layer of the pentamode material cloak before and after optimization

优化前	最小化 10 kHz 散射声能量		最小化 9 kHz、10 kHz、11 kHz 散射声能量	
	能量	能量	能量	能量
r_1	0.180 m	0.166 m	0.158 m	0.161 m
r_2	0.260 m	0.247 m	0.240 m	0.245 m
r_3	0.340 m	0.344 m	0.329 m	0.336 m
r_4	0.420 m	0.428 m	0.418 m	0.417 m

为了定量地展现优化前后的性能提升，常用散射截面 σ_s 来描述隐身衣的隐身效果：

$$\sigma_s = E_s/E_{in} \quad \dots (2.13)$$

其中， E_{in} 为入射声能量， E_s 为散射声能量，散射截面 σ_s 正比与散射声能量 E_s ，散射截面越小，隐身效果越好。图 2.16 展示了优化前后隐身衣散射截面随频率的变化曲线，与未优化的时均匀分层的隐身衣（蓝色实线）相比，优化后隐身衣的散射截面（红色点划线）在以 10 kHz 为中心频率一定宽度的频带内整体均有所

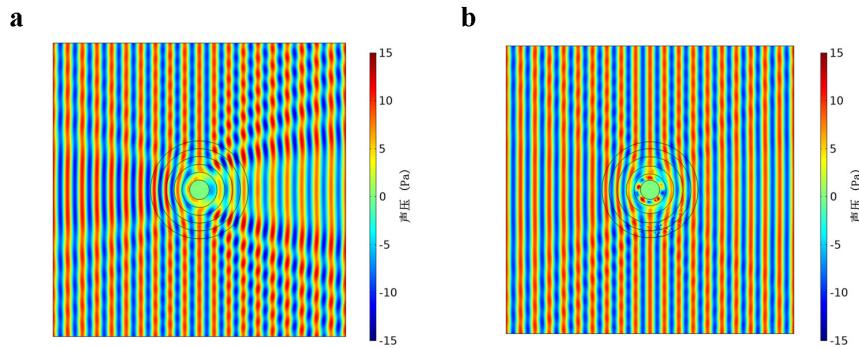


图 2.14 优化前后总声场对比: (a) 优化前; (b) 优化后。

Figure 2.14 Comparison of the total acoustics fields (a) before and (b) after optimization

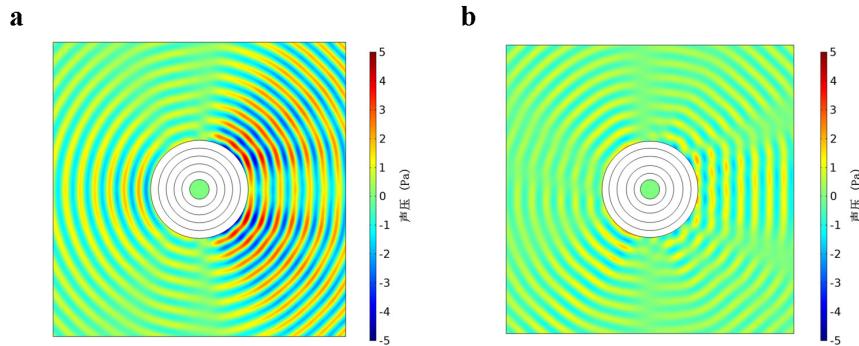


图 2.15 优化前后散射声场对比: (a) 优化前; (b) 优化后。

Figure 2.15 Comparison of the scattered acoustics fields (a) before and (b) after optimization

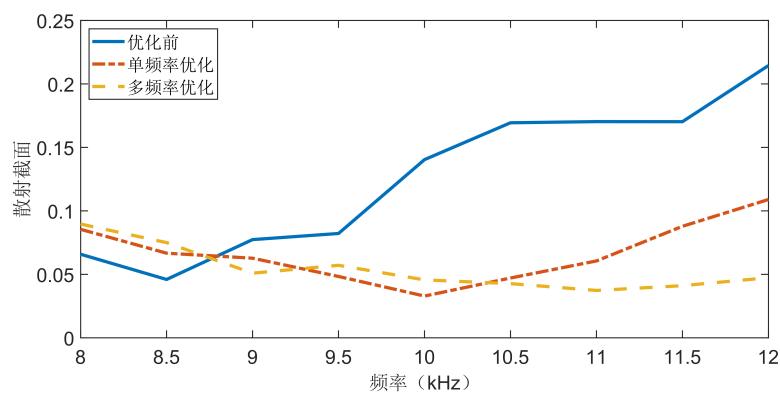


图 2.16 优化前后隐身效果对比

Figure 2.16 Comparison of the stealth effects before and after optimization

降低。其中，目标频率 10 kHz 处的散射声能量较优化前降低了 77%。由此可见，基于单目标频点的 GA 优化不仅可以使目标频点的散射声压大幅度降低，以目标频点为中心的一定带宽内的散射声压也会随之被降低。

为了拓宽优化频带的宽度，躲避宽频声呐信号的探测，也可以对多个频点的散射声场同时进行最小化。这里选取目标频率点为 9 kHz、10 kHz 和 11 kHz，优化后的隐身衣每一层的半径如表 2.3 第三列所示，优化后隐身衣散射截面随频率的变化曲线如图 2.16 中黄色虚线所示。与图 2.16 中红色点划线对比可知，虽然单目标优化可以令 $f = 10 \text{ kHz}$ 处的散射声能量降得很低，但偏离 10 kHz 时散射声压会稍有提升；若对 9 kHz、10 kHz、11 kHz 三个频率点同时进行优化，在更加宽频的范围内的散射声能量均被降低了，且散射声能量的变化趋势也更加平缓。

2.4.4 最小化反射声能量

现阶段使用的声呐探测仪多为自发自收型，考虑实际需求只需要最小化回波能量，即反射声能量，就能躲避自发自收型水下声呐信号探测，获得很好的隐身效果。

假定探测信号为由仿真区域左方入射的 5 kHz 的平面波，优化目标为最小化反射声能量，即最小化左半边区域的散射声能量。整个优化过程与数值仿真方法与 2.4.3 节完全相同，其中隐身衣的反射声能量可以通过散射声场对左半边区域的边界积分求得。图 2.17 为优化前后散射声场分布对比，可以看到优化前散射声场能量有很大一部分集中在左侧回波部分，而对分层方式进行优化可以使散射声场能量几乎全部集中在透射区域，从而大大降低回波强度，隐身性能得到大大加强。

与 2.4.3 节中的散射截面类似，这里定义反射截面来定量地描述反射声场的强度：

$$\sigma_r = E_r / E_{in} \quad \dots (2.14)$$

其中， E_r 为反射声能量，反射截面 σ_r 正比于反射声能量 E_r ，反射截面越小，在入射端进行探测时越不易探测到，隐身效果越好。

图 2.18 为优化前后反射截面和散射截面随频率的变化曲线的对比，实线代表优化前的仿真结果，虚线为优化后的仿真结果。由图 2.18 可知，在目标频率附近的反射声能量得到了明显的抑制，反射能量降低了 90% 以上。值得注意的

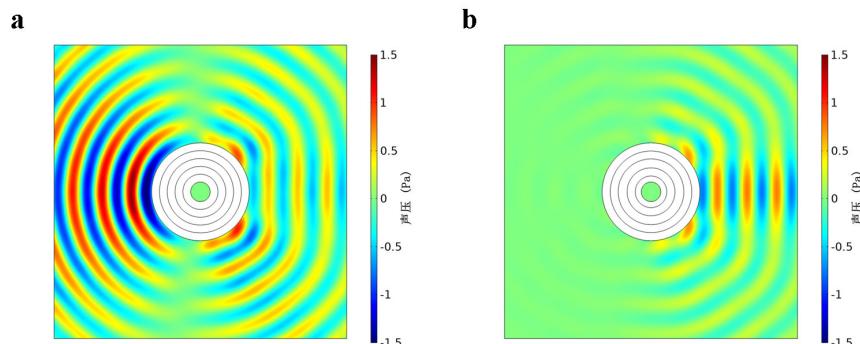


图 2.17 优化前后散射声场对比：(a) 优化前；(b) 优化后。

Figure 2.17 Comparison of the scattered acoustics fields (a) before and (b) after optimization

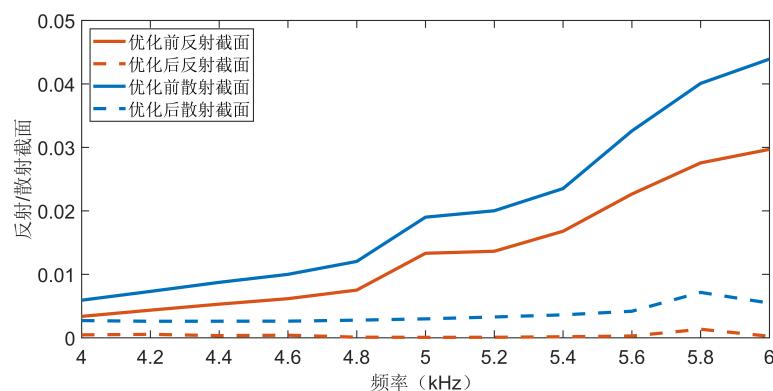


图 2.18 优化目标为最小化反射声能量时隐身效果对比

Figure 2.18 Comparison of stealth effect for minimizing reflected acoustic energy

是，压低反射区域声能量的同时并不会造成透射区域声能量的增高，整个背景声场的散射声能量也均比优化前有所降低。

2.5 本章小结

GA 作为一种经典的启发式优化算法，被广泛应用于解决各类的优化问题。本章对 GA 的原理进行了基本的介绍，并将 GA 应用到声学结构的优化问题的求解，并给出了两个优化案例。首先，我们将 LRCWA 方法与 GA 相结合，对基于多孔材料的吸声体的分层占空比和表面形状进行了优化，显著提升了吸声体在目标频段的吸声性能；而后基于 Norris 教授提出的五模材料变换声学理论，提出了基于 GA 的五模材料隐身衣分层优化策略。仿真结果表明优化前的均匀分层并不能得到最理想的隐身效果，优化后散射声能量和反射声能量都有大幅下降，对于窄带和宽带探测信号都具有更好的隐身效果。

由于声人工结构设计是一个十分复杂的问题，很难得到确切的解析解，而在对数值解的探索方面尚有很多问题亟待解决。鉴于 GA 在声人工结构优化问题上取得了较好的效果，且具有原理简单、易于实现、通用性强等优点，在之后的研究中也可以向其他声学结构进行推广，进一步推动声人工结构从理论模型迈向实际应用。

第3章 基于深度学习的声人工结构逆设计

3.1 引言

在声人工结构的实际应用中，我们常常需要根据期望的物理性能对单元的几何结构进行设计，这种设计是一种典型的逆向设计问题。传统的设计路线可以分为两部分：首先，我们可以借助一些经典的物理模型和从之前的实践中积累的经验和直觉来确定模型的框架。例如我们可以基于亥姆霍兹共鸣器结构来实现吸隔声，可以通过迷宫结构构造超表面来调控相位实现准直等。尽管这些物理模型为我们的设计提供了重要的指导方针，但最终确定结构的几何参数来实现期望的物理特性并非易事，尤其是当结构的几何参数较多且相互存在耦合时。这时就进入到第二个步骤，我们需要利用有限元法（Finite Element Method, FEM）等数值模拟方法不断微调这些几何参数，并结合优化算法进行大量的正向计算，不断迭代，来逐步接近我们的设计目标。在第2章中，我们以遗传算法（Genetic Algorithm, GA）为例介绍了基于优化类算法的声学结构设计方法。这一类方法虽然也可以实现声学结构的自动化设计，但本质上还是在进行正向设计，会带来较高的计算成本，设计效率并不高。因此，为了能够从根本上解决这个具有挑战的逆设计问题，我们将深度学习模型和集总参数模型相结合，提出了一种高效且具有一定通用性的声学结构逆设计方法，并基于该方法在给定几何参数范围和工作频段的基础上完成了对二阶亥姆霍兹共鸣器（Two-order Helmholtz Resonator, THR）几何参数的自动化设计，以实现特定的声传输损失（sound transmission loss, STL）。

3.2 物理模型

如图3.1所示为THR单元的几何结构示意图，其中包含8个几何参数：第一阶单元的细管半径 a_1 ，细管长度 l_1 ，腔体半径 r_1 ，腔体深度 h_1 ；以及第二阶单元的细管半径 a_2 ，细管长度 l_2 ，腔体半径 r_2 ，腔体深度 h_2 。腔体深度和腔体半径共同决定了腔体体积，腔体体积直接对应结构的声容。为了简化分析，我们固定腔体的半径 $r_1 = r_2 = 5\text{ cm}$ ，因此一个THR结构的几何自由度为6，待设计的几何参数可以被表示为 $\mathbf{g} = [a_1, l_1, h_1, a_2, l_2, h_2]$ 。与传统的一阶亥姆霍兹共鸣器

相比，THR 结构拥有更丰富的共振模式。在结构大小相同的情况下，THR 可以在低频共振频率基本不改变的基础上，比传统的一阶亥姆霍兹共鸣器多出额外的高频共振峰 [81, 82]。

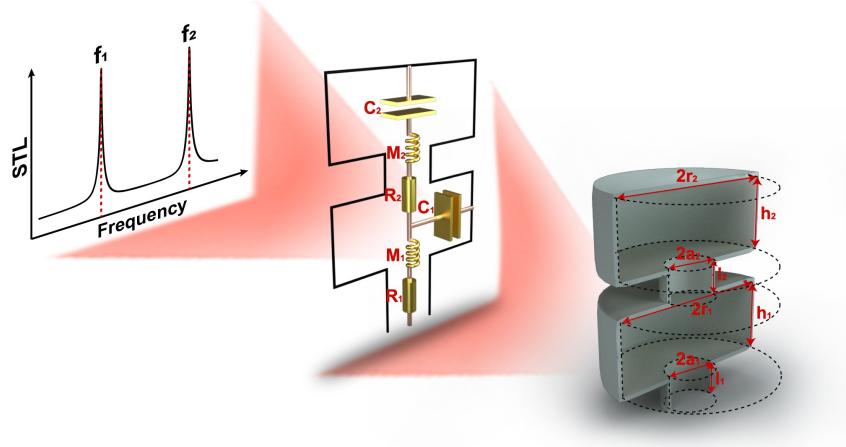


图 3.1 THR 的几何结构示意图

Figure 3.1 Schematic view of the THR.

在实际应用中，亥姆霍兹共鸣器常作为管道的旁支结构用于管路的隔声降噪。如图 3.2 所示为 THR 作为管道旁支结构的结构示意图，我们通常可以基于集总参数模型（Lumped-parameter Technique, LPT），传输矩阵模型（Transmission Matrix Method, TMM）和有限元模型（Finite Element Method, FEM）等物理模型对其物理性质进行正向计算。

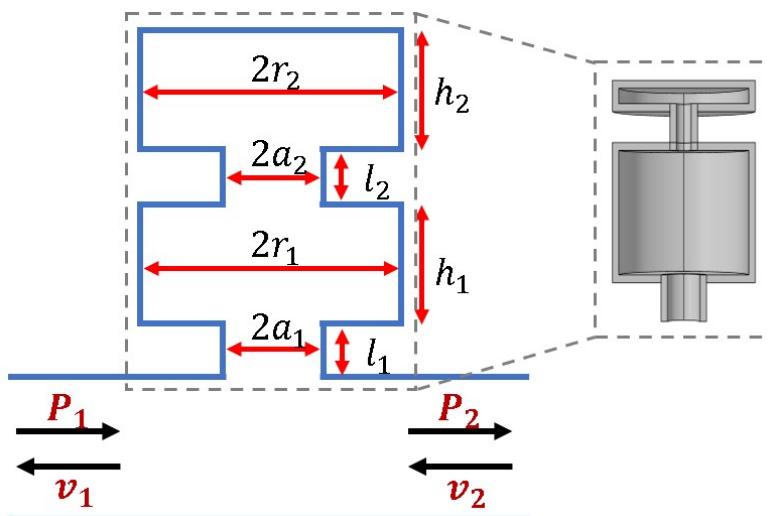


图 3.2 THR 作为管道旁支结构的示意图

Figure 3.2 Schematic view of THR as the side branch of a tube.

3.2.1 集总参数模型

图 3.3 为图 3.2 所示结构的等效电路图。这里我们将短管等效为声阻 R 和声质量 M ，将腔体等效为声容 C 。因此，在 LPT 模型下 THR 可以由等效电学参数 $e = [R_1, M_1, C_1, R_2, M_2, C_2]$ 进行描述，其中， R_1, M_1 和 C_1 分别为 THR 第一阶单元的等效声阻，等效声质量和等效声容； R_2, M_2 和 C_2 分别为 THR 第二阶单元的等效声阻，等效声质量和等效声容。几何参数 g 和等效电学参数 e 之间的转换关系如表 3.1 所示。其中， ρ_0 和 c_0 分别为空气的密度和声速； η 为空气的粘滞系数； $\delta_i = \frac{8a_i}{3\pi} \left(2 - \beta_i \frac{a_i}{r_i}\right)$ 为第 i 阶短管的末端修正，修正因子 $\beta_1 = 0.75$ ， $\beta_2 = 1.05$ 。利用 LPT 可以求得该结构的声阻抗 Z_{THR} 为：

$$Z_{THR} = R_1 \sqrt{\omega} + j\omega M_1 + \frac{1}{j\omega C_1 + \frac{1}{R_2 \sqrt{\omega} + j\omega M_2 + 1/j\omega C_2}}. \quad \dots (3.1)$$

其中， $\omega = 2\pi f$ 为角频率。此时带有旁支结构的管路的 STL 随频率的变化 $t(\omega)$ 可以被表示为：

$$t(\omega) = 10 \log \frac{X_b^2 + \left(\frac{Z_0}{2S} + R_b\right)^2}{R_b^2 + X_b^2}. \quad \dots (3.2)$$

其中， R_b 和 X_b 分别为声阻抗 Z_{THR} 的实部和虚部； $Z_0 = \rho_0 c_0$ 为空气的特性阻抗； S 为管道的横截面积。通过分析公式 (3.2) 可知，当声阻抗 Z_{THR} 的虚部 $X_b = 0$ 时，声传输损失为极大值，降噪效果最好，此时对应的频率即为 THR 的共振频率。一个 THR 单元通常有两个共振频率，分别对应 STL 频谱中的两个共振峰。

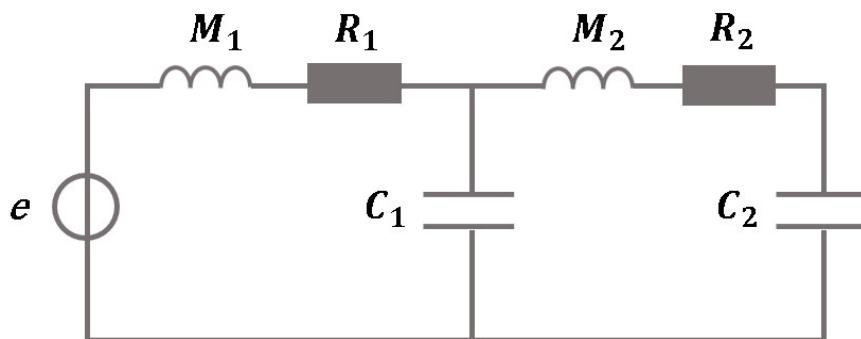


图 3.3 THR 作为管道旁支结构时的等效电路图

Figure 3.3 Equivalent circuit diagram of THR as the side branch of a tube.

表 3.1 几何参数和等效电学参数的转换关系

Table 3.1 Relationships between geometric parameters and equivalent electrical parameters

几何参数 → 等效电学参数	等效电学参数 → 几何参数
$R_i = l_i \sqrt{2\eta\rho_0}/\pi a_i^3$	$\frac{\rho_0 R_i}{\sqrt{2\eta\rho_0}} a_i^2 - \left(\frac{8\rho_0 \beta_i}{3r_i \pi^2} + M_i \right) a_i + \frac{16\rho_0}{3\pi^2} = 0$
$M_i = \rho_0 (l_i + \delta_i)/\pi a_i^2$	$l_i = \frac{3\pi R_i a_i}{\sqrt{2\eta\rho_0}}$
$C_i = V_i/\rho_0 c_0^2$	$V_i = C_i \rho_0 c_0^2$

3.2.2 传输矩阵模型

在 LPT 模型中，我们将腔体等效为声容而忽略了它的声质量，这种近似在有些情况下会引入较大的误差。TMM 模型通过将声学结构的不同部分等效为传输矩阵的方式对声学结构进行建模，如图 3.4 所示，与 LPT 相比基于 TMM 对 THR 单元的 STL 进行计算往往更加准确。

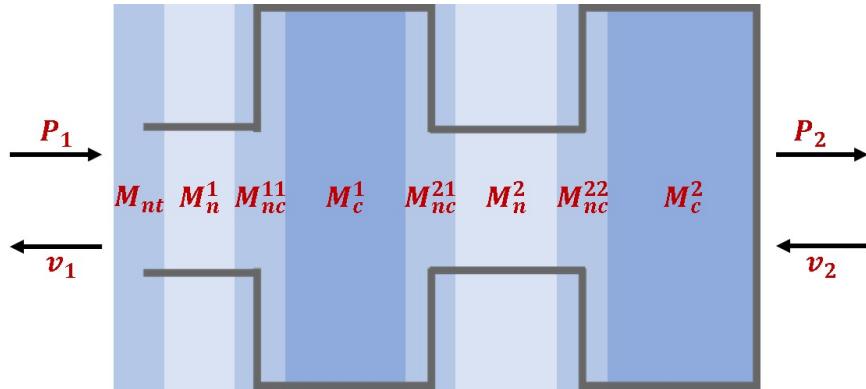


图 3.4 基于 TMM 对 THR 单元进行建模

Figure 3.4 TMM model of the THR.

基于 TMM 计算图 3.2 所示结构的 STL 的计算流程如下：

1. 计算一阶细管与流道连接处的传递矩阵 M_{nt} ：

$$M_{nt} = \begin{bmatrix} 1 & jz_n^1 k \Delta l_{nt} \\ 0 & 1 \end{bmatrix}. \quad \dots (3.3)$$

其中， $z_n^1 = \frac{\rho_0 c_0}{\pi a_1^2} \left(1 + \frac{\delta(1-j)}{\sqrt{2}a_1} \left(1 - \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为一阶细管的等效声阻抗， $\delta =$

$\sqrt{\frac{2\eta}{\rho_0\omega}}$ 为边界层厚度, $\eta = 1.8e^{-5}$ Pa · s 为空气粘滞系数, $\gamma = 1.4$ 为空气的比热比, $Pr = 0.702$ 为普拉克常数; 一阶细管与流道连接处的修正长度 Δl_{nt} 为 $\Delta l_{nt} = 0.82 \left[1 - 0.235 \frac{a_1}{r_t} - 1.32 \left(\frac{a_1}{r_t} \right)^2 + 1.54 \left(\frac{a_1}{r_t} \right)^3 - 0.86 \left(\frac{a_1}{r_t} \right)^4 \right] a_1$, $2r_t$ 为流道的边长; $k = \omega/c$ 为波矢, $\omega = 2\pi f$ 为角频率, $\rho_0 = 1.21\text{kg/m}^3$ 和 $c_0 = 343\text{m/s}$ 分别为空气的密度和声速。

2. 计算一阶细管的传递矩阵 \mathbf{M}_n^1 :

$$\mathbf{M}_n^1 = \begin{bmatrix} \cos(k_n^1 l_1) & j z_n^1 \sin(k_n^1 l_1) \\ \frac{j \sin(k_n^1 l_1)}{z_n^1} & \cos(k_n^1 l_1) \end{bmatrix}. \quad \dots (3.4)$$

其中, $k_n^1 = \frac{\omega}{c_0} \left(1 + \frac{\delta(1-j)}{\sqrt{2}a_1} \left(1 + \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为一阶细管的等效波矢;

3. 计算一阶细管和一阶腔体连接处的传递矩阵 \mathbf{M}_{nc}^{11} :

$$\mathbf{M}_{nc}^{11} = \begin{bmatrix} 1 & j z_n^1 k \Delta l_{11} \\ 0 & 1 \end{bmatrix}. \quad \dots (3.5)$$

其中, $\Delta l_{11} = 0.82 \left[1 - 1.35 \frac{a_1}{r_1} + 0.31 \left(\frac{a_1}{r_1} \right)^3 \right] a_1$ 为一阶细管和一阶腔体连接处的修正长度。

4. 计算一阶腔体的传递矩阵 \mathbf{M}_c^1 :

$$\mathbf{M}_c^1 = \begin{bmatrix} \cos(k_c^1 h_1) & j z_c^1 \sin(k_c^1 h_1) \\ \frac{j \sin(k_c^1 h_1)}{z_c^1} & \cos(k_c^1 h_1) \end{bmatrix}. \quad \dots (3.6)$$

其中, $k_c^1 = \frac{\omega}{c_0} \left(1 + \frac{\delta(1-j)}{\sqrt{2}r_1} \left(1 + \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为一阶腔体的等效波矢; $z_c^1 = \frac{\rho_0 c_0}{\pi r_1^2}$ $\left(1 + \frac{\delta(1-j)}{\sqrt{2}r_1} \left(1 - \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为一阶腔体的等效阻抗。

5. 计算一阶腔体和二阶细管连接处的传递矩阵 \mathbf{M}_{nc}^{21} :

$$\mathbf{M}_{nc}^{21} = \begin{bmatrix} 1 & j Z_n k \Delta l_{21} \\ 0 & 1 \end{bmatrix}. \quad \dots (3.7)$$

其中, $Z_n = \frac{\rho_0 c_0}{\pi a_2^2} \left(1 + \frac{\delta(1-j)}{\sqrt{2}a_2} \left(1 - \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为二阶细管的等效声阻抗, $\Delta l_{21} = 0.82 \left[1 - 1.35 \frac{a_2}{r_1} + 0.31 \left(\frac{a_2}{r_1} \right)^3 \right] a_2$ 为一阶腔体和二阶细管连接处的修正长度。

6. 计算二阶细管的传递矩阵 \mathbf{M}_n^2 :

$$\mathbf{M}_n^2 = \begin{bmatrix} \cos(k_n^2 l_2) & j z_n^1 \sin(k_n^2 l_2) \\ \frac{j \sin(k_n^2 l_2)}{z_n^2} & \cos(k_n^2 l_2) \end{bmatrix}. \quad \dots (3.8)$$

其中 $k_n^2 = \frac{\omega}{c_0} \left(1 + \frac{\delta(1-j)}{\sqrt{2}a_2} \left(1 + \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为二阶细管的等效波矢;

7. 计算二阶细管和二阶腔体连接处的传递矩阵 \mathbf{M}_{nc}^{22} :

$$\mathbf{M}_{nc}^{22} = \begin{bmatrix} 1 & jz_n^2 k \Delta l_{22} \\ 0 & 1 \end{bmatrix}. \quad \dots (3.9)$$

其中, $\Delta l_{22} = 0.82 \left[1 - 1.35 \frac{a_2}{r_2} + 0.31 \left(\frac{a_2}{r_2} \right)^3 \right] a_2$ 为二阶细管和二阶腔体连接处的修正长度:

8. 计算二阶腔体的传递矩阵 \mathbf{M}_c^2 :

$$\mathbf{M}_c^2 = \begin{bmatrix} \cos(k_c^2 h_2) & j z_c^2 \sin(k_c^2 h_2) \\ \frac{j \sin(k_c^2 h_2)}{z_c^2} & \cos(k_c^2 h_2) \end{bmatrix}. \quad \dots (3.10)$$

其中, $k_c^2 = \frac{\omega}{c_0} \left(1 + \frac{\delta(1-j)}{\sqrt{2}r_2} \left(1 + \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为二阶腔体的等效波矢; $z_c^2 = \frac{\rho_0 c_0}{\pi r_2^2} \left(1 + \frac{\delta(1-j)}{\sqrt{2}r_2} \left(1 - \frac{\gamma-1}{\sqrt{Pr}} \right) \right)$ 为二阶腔体的等效阻抗;

9. 计算 THR 单元的等效传递矩阵 \mathbf{M}_{total} :

$$\mathbf{M}_{total} = \mathbf{M}_{nt} \mathbf{M}_n^1 \mathbf{M}_{nc}^{11} \mathbf{M}_c^1 \mathbf{M}_{nc}^{21} \mathbf{M}_n^2 \mathbf{M}_{nc}^{22} \mathbf{M}_c^2. \quad \dots (3.11)$$

10. 计算 THR 单元作为流道的旁支结构时的等效传递矩阵 \mathbf{T} :

$$\mathbf{T} = \begin{bmatrix} 1 & 0 \\ 1/Z_{THR} & 1 \end{bmatrix}. \quad \dots (3.12)$$

其中, $Z_{THR} = \mathbf{M}_{total}(1, 1)/\mathbf{M}_{total}(2, 1)$ 为 THR 的等效阻抗, $\mathbf{M}_{total}(i, j)$ 代表矩阵 \mathbf{M}_{total} 第 i 行第 j 列的元素;

11. 对于频率为 ω 的入射波, 计算 THR 单元作为流道的旁支结构时的 STL 大小 $t(\omega)$

$$t(\omega) = 20 \lg \frac{|T_{11} + T_{12}/Z_t + T_{21}Z_t + T_{22}|}{2}. \quad \dots (3.13)$$

其中, T_{ij} 为等效传递矩阵 \mathbf{T} 的第 i 行 j 列的元素 ($i = 1, 2, j = 1, 2$); $Z_t = \frac{\rho_0 c_0}{S}$ 为流道的平面波声阻抗, $S = 4r_t^2$ 为流道的横截面积; 当 THR 单元的几何参数

固定时, $t(\omega)$ 是频率 ω 的函数, 改变频率 ω 的大小即可计算出该 THR 的 STL 频谱。

3.2.3 有限元模型

除了上面所介绍的 LPT 模型和 TMM 模型, 另一个较为通用的计算工具是借助商业仿真软件 COMSOL Multiphysics 对声学结构进行 FEM 仿真, 计算得到结构的 STL 频谱。值得注意的是, 在实际应用中我们常常需要同时使用多种方法对声学结构进行仿真计算, 因此在使用 COMSOL Multiphysics 进行仿真时我们需要保证软件调用的密度、声速等物理参数与解析方法中一致。同时, 考虑到细管中的粘滞常常较大, 因此仿真中 THR 单元的一阶短管和二阶短管区域的物理场设置应为热粘性声学, 该部分的网格也应尽可能的细。

为了进一步验证三种计算方法的等效性, 我们针对大量的具有不同共振频率的 THR 单元基于上述三种方法进行了仿真计算, 所得到的 STL 频谱一致性都很好。图 3.5 中展示了三个不同的 THR 单元的计算结果对比, 相应几何参数也记录在每幅图下方的表格中。因此, 考虑到基于 FEM 的计算方法耗时较高, 在实际应用中我们为了提高设计效率, 在设计初期往往会选择 LPT 模型、TMM 模型等解析的方法对结构的几何参数进行设计, 然后再使用 COMSOL Multiphysics 对设计结果进行验证。

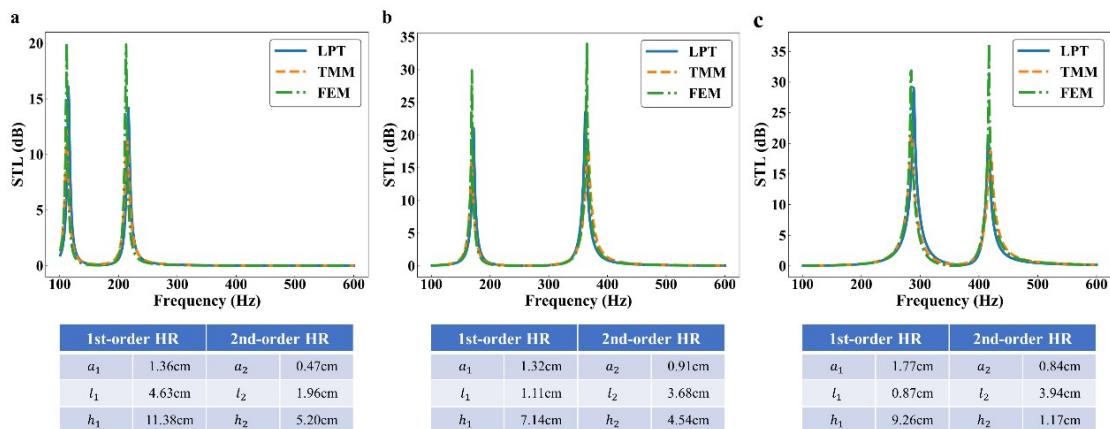


图 3.5 基于 LPT, TMM 和 FEM 的 THR 单元的 STL 的计算结果对比

Figure 3.5 STL spectrum of the THR calculated by LPT, TMM and FEM.

3.3 逆设计问题的分析

在上一小节中，我们介绍了三种较为常用的 THR 单元的正向计算模型，使得我们可以在给定结构的几何参数的基础上较为容易的计算出结构的物理特性。然而，实际应用中我们常常需要构造具有指定物理特性的结构，如在 150 Hz 和 250 Hz 实现 10 dB 以上的降噪等。这种逆设计问题往往是十分具有挑战性的，无法通过解析地方式进行求解。在本小节中，我们将着重对逆设计问题的复杂性和非唯一性展开分析。

3.3.1 复杂性

由于 THR 单元在共振频率处降噪效果最好，因此我们常常希望结构的共振频率与背景环境中的噪声频率相一致。结合 (3.1)式和 (3.2)式可知，当声阻抗 Z_{THR} 的虚部 $X_b = 0$ 时，STL 为极大值点，此时对应的频率即为 THR 的共振频率。因此，令 $X_b = 0$ 可以得到一个关于共振频率 ω_0 的六次方程：

$$M_1 C_1^2 M_2^2 \omega_0^6 + M_1 C_1^2 R_2^2 \omega_0^5 - C_1 M_2 \left(2 M_1 \frac{c_1+c_2}{c_2} + M_2 \right) \omega_0^4 - C_1 R_2^2 \omega_0^3 + \left[M_1 \left(\frac{c_1}{c_2} \right)^2 + 2 (M_1 + M_2) \frac{c_1}{c_2} + M_1 + M_2 \right] \omega_0^2 - \frac{c_1+c_2}{c_2^2} = 0. \quad \dots (3.14)$$

Abel–Ruffini 定理表明，五次及更高次的代数方程不存在一般的代数解法，这为我们的求解造成了很大的困难 [83]。值得注意的是，我们不仅需要关注结构的共振频率 ω_0 ，还需要对结构的声阻进行调节，以确保结构在共振频率 ω_0 处的 STL 足够高以满足实际应用中的隔声需求。而结构声阻与 STL 之间的关系又是十分复杂的，因此我们很难通过解析的手段对 THR 单元进行逆设计

此外，THR 单元的几何参数 $[a_1, l_1, h_1, a_2, l_2, h_2]$ 之间存在耦合，很难对其进行独立的调节，以实现我们期望的共振频率。为了定量地分析 THR 单元的共振频率对不同几何参数的敏感度，我们定义共振频率的平均误差（Average Error of the Resonant Frequency, AERF）为：

$$\text{AERF}(\mathbf{g}) = \frac{|f_1(\mathbf{g}) - f_1^{\text{target}}| + |f_2(\mathbf{g}) - f_2^{\text{target}}|}{2}. \quad \dots (3.15)$$

其中， $f_1(\mathbf{g})$ 和 $f_2(\mathbf{g})$ 分别为几何参数为 \mathbf{g} 的 THR 单元的一阶共振频率和二阶共振频率； f_1^{target} 和 f_2^{target} 分别为一阶目标共振频率和二阶目标共振频率。随后，我们对一个共振频率在 150 Hz 和 250 Hz 的 THR 单元（即 $f_1^{\text{target}} = 150 \text{ Hz}$,

$f_2^{\text{target}} = 250 \text{ Hz}$) 的 6 个几何参数 $[a_1, l_1, h_1, a_2, l_2, h_2]$ 分别进行了一定范围内的扰动，并计算了扰动后的结构 \tilde{g} 的共振频率 $f_1(\tilde{g})$ 和 $f_2(\tilde{g})$ ，最后基于公式 (3.15) 计算出 AERF，如图 3.6 (a) 所示。其中横坐标为归一化的几何参数，即扰动后的几何参数与原始的几何参数的比值，扰动范围为原始的几何参数的 $\pm 10\%$ ；纵坐标为扰动后结构的 AERF。由图 3.6 (a) 可知，上述这 6 个几何参数 $[a_1, l_1, h_1, a_2, l_2, h_2]$ 均会对 THR 单元的共振频率产生影响，对其中任何一个几何参数进行独立调节都会导致 THR 单元的共振频率发生不同程度的偏移；并且斜率越大意味着改变该参数对结构的共振频率的影响越大，即共振频率对该参数越敏感。因此，短管的半径 a_1 和 a_2 及一阶腔体的深度 h_1 对 THR 单元的共振频率影响较大。

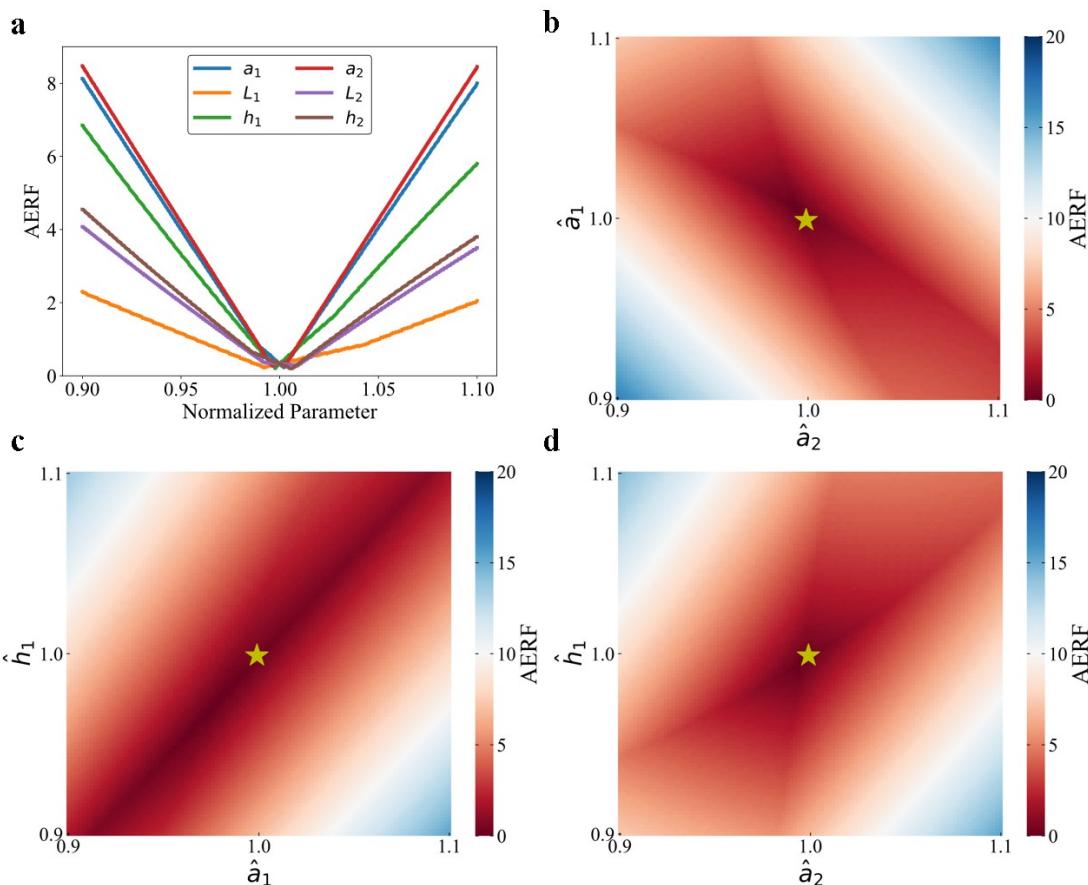


图 3.6 THR 单元的共振频率对不同几何参数的灵敏度

Figure 3.6 The sensitivity of the resonant frequency on different geometric parameters of the THR.

为了进一步展现参数间的耦合效应，我们分别对共振频率影响较大的 a_1 , a_2 和 h_1 两两组合进行联合调节，调节范围仍然为 $\pm 10\%$ ，同时保持其他 4 个几何参数不变。如图 3.6 (b)~(d) 所示为扰动后结构的 AERF 分布，横坐标和纵坐标

为归一化后的几何参数，颜色代表 AERF，黄色星星代表共振频率为 150 Hz 和 250 Hz 的原始 THR 单元。通过观察上述三幅图可以发现，黄色星星附近的暗红色区域所对应的 THR 的共振频率依然为 150 Hz 和 250 Hz，也就是说不同几何参数对结构共振频率的影响有时会叠加，有时会相互抵消。因此，当我们同时对多个几何参数同时进行调节时，结构的共振频率的变化是很难预测的。

因此，THR 单元的共振频率对几何参数是非常敏感的，我们在对 THR 单元进行设计时很难通过手动地调节每个几何参数来使结构的两个共振频率同时满足我们的要求。如果使用遍历搜索的方法需要对 6 个几何参数同时进行遍历，且步长需要取的比较小，这会使计算成本急剧增加。

3.3.2 非唯一性

对于绝大多数声学结构而言，期望的物理响应和几何参数之间的对应关系往往不是一对一的：一方面，受限于结构本身的物理机制，有一些较为极端的性质是无法被物理实现的，此时该逆设计问题是无解的；另一方面，对于一些期望的物理响应，往往存在多个几何结构可以与之对应，此时逆设计问题就是一个多解问题。这里我们主要以 THR 单元的为例，对逆设计问题的非唯一性进行说明。



图 3.7 期望的声学特性与声学结构之间的映射关系

Figure 3.7 Corresponding relation between target acoustic functionality and acoustic structures.

在 LPT 模型下，该逆设计问题可以被拆解成如图 3.7 所示的三个级联的映射关系，即：期望的物理性质 → STL 频谱 → 等效电学参数 → 几何参数。首先，对

于期望的物理响应，存在不止一条 STL 频谱符合设计需求，即映射①是非唯一的；其次，由表 3.1 可知，对于给定的一组等效电学参数，往往可以反解出不止一组几何参数，即映射③也是非唯一的。

图 3.8 给出了一个具体的设计案例，这里我们的设计目标是希望 THR 单元在 150 Hz 和 250 Hz 处的 STL 大于 10 dB。图 3.8(a) 和 (c) 展示了两条不同的 STL 频谱，这两条 STL 频谱在 150 Hz 和 250 Hz 均存在两个高度大于 10 dB 的共振峰，对应图 3.7 中映射①的非唯一性；上述两条 STL 频谱对应的等效电学参数记录在图片右上角的表格中，对应图 3.7 中映射②，该映射是一对一映射；每一组等效电学参数又对应两组不同的几何参数，分别记录在图 3.8(b) 和 (d) 的表格中，对应图 3.7 中映射③的非唯一性。上述 4 个 THR 单元的物理性质均满足我们的设计需求，并且它们之间的几何结构的差异性是非常大的，如果我们将期望的物理性质作为模型的输入，结构的几何参数做为模型的输出，由于输出的非唯一性训练过程中模型将面临多个收敛方向，从而为训练造成极大地困难。

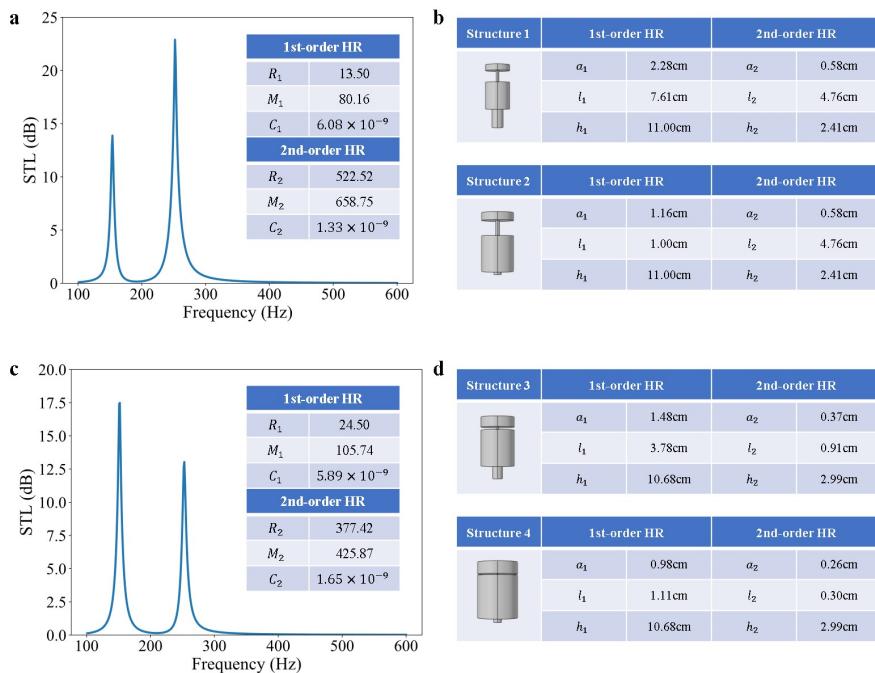


图 3.8 满足同一设计目标的不同声学结构

Figure 3.8 Different acoustic structures corresponding to the same target acoustic functionality..

深度学习模型所拟合的是模型输入和输出之间唯一的映射关系，然而声学结构的逆设计是一个非唯一问题，会给模型训练造成极大地困扰，往往会导致模

型难以收敛，或收敛后模型的性能并不理想。针对这一问题，一种解决思路是引入双向网络模型 [53, 56]，以确保模型在训练过程中可以顺利地收敛。然而，双向网络模型包含一个正向预测模型和一个逆向设计模型，模型规模较大，待训练参数较多，因此需要大量的训练样本参与模型的训练，训练难度也较大。

如果我们希望使用更小的模型来实现声学结构的逆设计，就需要采用一些手段来削弱这种非唯一性。本文所提出的深度学习模型为 5 层全连接模型，同时为了确保输入输出间映射关系的唯一性，模型的输入为期望的 STL 频谱，模型的输出为结构的等效电学参数，即模型学习的是图 3.7 中的映射 ②。这个做法不仅会为我们很好的规避掉映射 ① 和映射 ③ 的非唯一性，帮助模型更好的收敛，还可以使所得到的模型具有更好的普适性，可以在 LPT 框架下向其他声学结构进行迁移（详见 4.5 章节）。

值得注意的是，对于某些物理响应，并不存在与之对应的声学结构，此时逆设计问题就是一个无解的问题。其中，一部分问题可以通过分析结构背后的物理机制对其无解性进行判断，例如通过分析 THR 单元的阻抗 Z_{THR} 的虚部 X_b 的零点可知，并不存在一个 THR 单元的 STL 频谱具有三个及以上的共振峰。然而，更多的逆设计问题是很难通过理论分析来判断其是否有解的，例如是否存在一个 THR 单元其 STL 频谱的两个共振峰距离小于 20Hz。对于这一类问题我们若想判断其是否有解，一个较为可行的解决思路是借助优化算法来进行试探性求解（详见 4.3.2 章节）。此外，若将这些较为极端的物理响应直接作为深度学习模型的输入，模型的输出往往也是不理想的，模型往往只能预测出物理响应与设计目标尽可能接近的声学结构。

3.4 数据生成

为了搭建符合我们要求的深度学习模型，我们首先需要生成足够数量的数据用于模型的训练、验证和测试。正如 3.2 章节所提到的，深度学习模型的输入是 STL 频谱，输出为等效电学参数，因此所生成的样本的 feature 为 THR 单元从 101 Hz 至 600 Hz 的 STL 频谱，长度为 500 维，样本的 label 为 6 维的等效电学参数 $[R_1, M_1, C_1, R_2, M_2, C_2]$ 。同时，考虑到不同电学参数间具有数量级的差异性，这里需对等效电学参数进行归一化处理。为了保证结构的物理可实现性，样本的电学参数是通过随机产生几何参数并基于表 3.1 中的公式计算

表 3.2 样本分布

Table 3.2 Data distribution of datasets.

	共振频率范围	样本个数
NO.1	$100 < f_1 \leq 150$	$150 < f_2 \leq 200$ 5000
NO.2		$200 < f_2 \leq 250$ 5000
NO.3		$250 < f_2 \leq 300$ 5000
NO.4		$300 < f_2 \leq 350$ 5000
NO.5		$350 < f_2 \leq 400$ 5000
NO.6		$400 < f_2 \leq 450$ 5000
NO.7		$450 < f_2 \leq 500$ 5000
NO.8		$500 < f_2 \leq 550$ 5000
NO.9		$550 < f_2 \leq 600$ 5000
NO.10	$150 < f_1 \leq 200$	$200 < f_2 \leq 250$ 5000
NO.11		$250 < f_2 \leq 300$ 5000
NO.12		$300 < f_2 \leq 350$ 5000
NO.13		$350 < f_2 \leq 400$ 5000
NO.14		$400 < f_2 \leq 450$ 5000
NO.15		$450 < f_2 \leq 500$ 5000
NO.16		$500 < f_2 \leq 550$ 5000
NO.17		$550 < f_2 \leq 600$ 5000
NO.18	$200 < f_1 \leq 250$	$250 < f_2 \leq 300$ 5000
NO.19		$300 < f_2 \leq 350$ 5000
NO.20		$350 < f_2 \leq 400$ 5000
NO.21		$400 < f_2 \leq 450$ 5000
NO.22		$450 < f_2 \leq 500$ 5000
NO.23		$500 < f_2 \leq 550$ 5000
NO.24		$550 < f_2 \leq 600$ 5000
NO.25	$250 < f_1 \leq 300$	$300 < f_2 \leq 350$ 0
NO.26		$350 < f_2 \leq 400$ 5000
NO.27		$400 < f_2 \leq 450$ 5000
NO.28		$450 < f_2 \leq 500$ 5000
NO.29		$500 < f_2 \leq 550$ 5000
NO.30		$550 < f_2 \leq 600$ 5000
NO.31	$300 < f_1 \leq 350$	$350 < f_2 \leq 400$ 0
NO.32		$400 < f_2 \leq 450$ 5000
NO.33		$450 < f_2 \leq 500$ 5000
NO.34		$500 < f_2 \leq 550$ 5000
NO.35		$550 < f_2 \leq 600$ 5000
NO.36	$350 < f_1 \leq 400$	$400 < f_2 \leq 450$ 0
NO.37		$450 < f_2 \leq 500$ 0
NO.38		$500 < f_2 \leq 550$ 5000
NO.39		$550 < f_2 \leq 600$ 5000

得到的，对应的 STL 频谱是基于上述电学参数通过 LPT 模型正向计算得到的。考虑到结构的可加工性，这里限定结构的几何参数范围为 $0.1 \text{ cm} < a_i < 2.5 \text{ cm}$, $0.1 \text{ cm} < l_i < 5 \text{ cm}$, $0.1 \text{ cm} < h_i < 12.7 \text{ cm}$ ($i = 1, 2$)；结构的等效电学参数的范围为 $1 < R_i < 170$, $1 < M_i < 300$, $7 \times 10^{-10} < C_i < 7 \times 10^{-9}$ ($i = 1, 2$)。为了使模型能更多地学习具有良好的隔声性能的 THR 结构的特征，我们在生成样本时只保留了共振频率处 STL 大于 10 dB 的样本。同时，为了使样本的分布尽可能均匀，在生成样本的同时需要引入一定的筛选机制，使模型能更全面地学习到整个目标频带的特征信息。具体的筛选方法是我们需要以样本的两个共振频率 f_1 和 f_2 所在频段作为分组标准对样本进行分组，且 1 阶共振频率 f_1 和 2 阶共振频率 f_2 一般不会处于同一个频段。如表 3.2 所示为样本的分组情况，第一组中的样本的 1 阶共振频率 f_1 和 2 阶共振频率 f_2 分别满足 $100 \text{ Hz} < f_1 \leq 150 \text{ Hz}$ 和 $150 \text{ Hz} < f_2 \leq 200 \text{ Hz}$ ；第二组中的样本的 1 阶共振频率 f_1 和 2 阶共振频率 f_2 分别满足 $100 \text{ Hz} < f_1 \leq 150 \text{ Hz}$ 和 $200 \text{ Hz} < f_1 \leq 250 \text{ Hz}$ …以此类推，最后一组中的样本的 1 阶共振频率 f_1 和 2 阶共振频率 f_2 分别满足 $500 \text{ Hz} < f_1 \leq 550 \text{ Hz}$ 和 $550 \text{ Hz} < f_2 \leq 600 \text{ Hz}$ 。按上述方式将样本分为 N 组，且初始状态每组中样本个数均为 0 个，当组内样本个数达到 5000 个时将不再保存该组的样本。

产生样本的流程示意图如图 3.9 所示，具体包括以下步骤：

1. 固定腔体半径 $r_1=r_2=5 \text{ cm}$ ，在上述给定的几何参数范围内随机产生一组几何参数 $\mathbf{g} = [a_1, l_1, h_1, a_2, l_2, h_2]$ ；
2. 将上述几何参数 \mathbf{g} 代入表 3.1 中，计算出对应的等效电学参数 $\mathbf{e} = [R_1, M_1, C_1, R_2, M_2, C_2]$ ；
3. 判断上述等效电学参数 \mathbf{e} 是否均在给定的等效电学参数范围内，若在的话则进入步骤 4，否则返回步骤 1。
4. 基于公式 (3.1) 和公式 (3.2) 计算得到 STL 频谱 $\mathbf{t} = [t_{f=100}, t_{f=101}, t_{f=102}, \dots, t_{f=600}]$ 。
5. 计算 STL 频谱 \mathbf{t} 的两个峰值所对应的频点（即共振频率），若共振频率处的 STL 大于 10 dB，则进入步骤 6，否则返回步骤 1。
6. 按照共振频率所在的区间对该样本进行分组，若该组样本数小于 5000，则保存该样本并进入步骤 7，否则返回步骤 1。
7. 判断是否每一组样本数均达到了 5000，若是则停止计算，否则返回步骤

1。

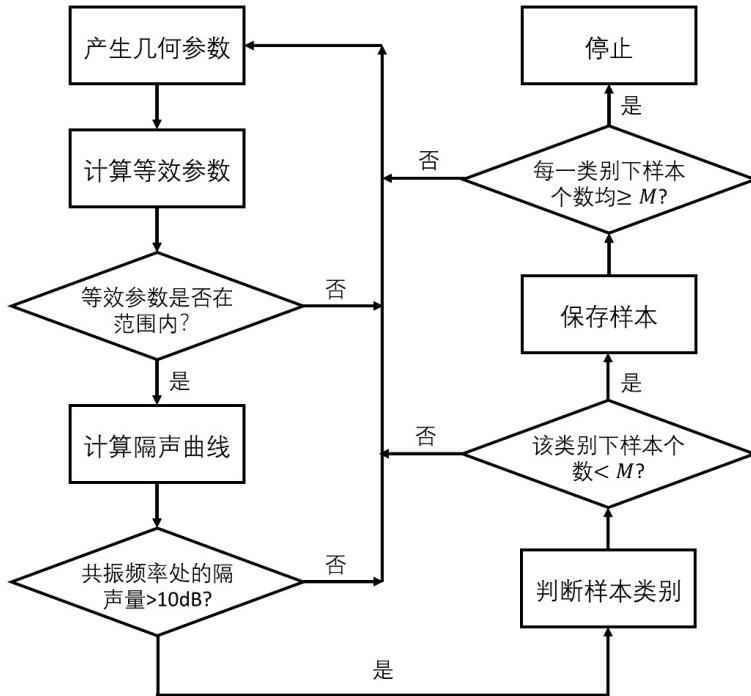


图 3.9 样本生成流程图

Figure 3.9 Concept schematic of the data generation.

如图 3.10 所示为产生的一个样本示例，其中 STL 频谱第一个峰值对应 1 阶共振频率 $f_1=128$ Hz，第二个峰值对应 2 阶共振频率 $f_2=247$ Hz。按照之前约定的分组标准，该样本对应的组为 NO.2，因此需要保存该样本的几何参数 \mathbf{g} ，等效电学参数 \mathbf{e} 和 STL 频谱 \mathbf{t} 至该组下。

基于上述方法，我们共得到了 195000 个样本数据，样本的共振频率覆盖了 101 Hz~600 Hz，且在共振频率处的 STL 均大于 10 dB。每个频率区间中的样本共振频率处的平均 STL 分布如图 3.11 所示，横轴和纵轴分别代表每个频率区间的中心频率，颜色的深浅代表着共振频率处 STL 的大小。由图 3.11 可知，当结构的共振频率较低时，对应的 STL 数值也一般较小，即低频的共振峰一般较低，反之高频的共振峰一般较高。但这只代表了大多数情况，还是存在少数结构低频共振峰是高于高频共振峰的。最后，这些数据被分为了训练集，验证集和测试集，分别占比 70%，20% 和 10%，用于模型的训练，验证和测试。

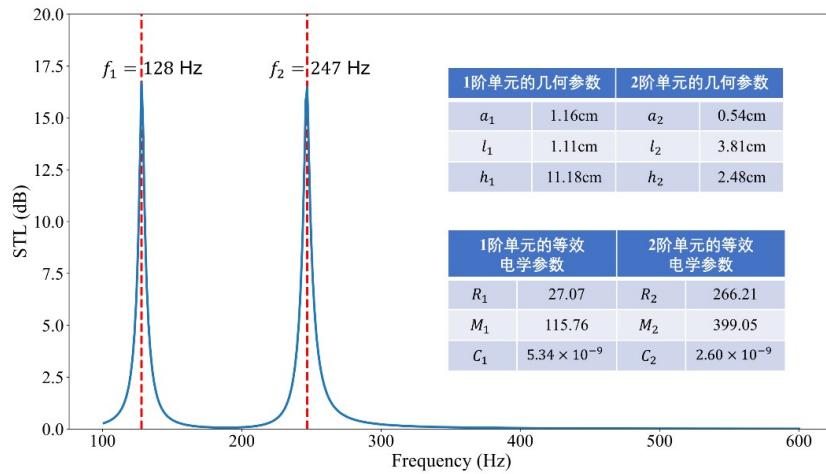


图 3.10 生成的样本数据

Figure 3.10 One of the THR samples.

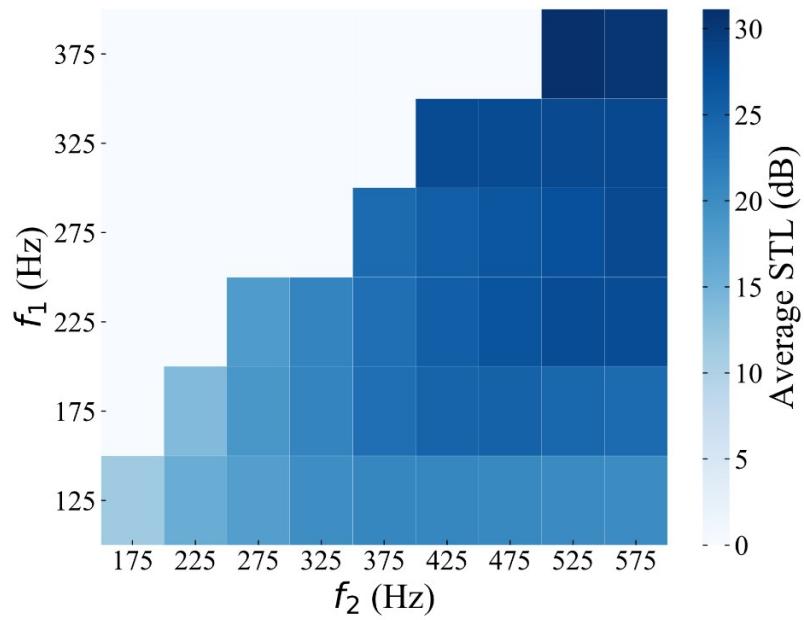


图 3.11 数据集分布

Figure 3.11 Distribution of samples.

3.5 深度学习模型的训练及测试

我们基于 pytorch 框架搭建了一个 5 层全连接神经网络，网络的输入为 101 Hz~600 Hz 间隔为 1 Hz 的 STL 频谱，输出为 6 维的等效电学参数，隐含层的神经元个数分别为 450, 250 和 220。模型训练时所使用的损失函数为均方误差函数 (Mean Square Error, MSE)，优化算法为 Adam 算法 [84]，学习率为 0.001，批大小 (Batch Size) 为 256。同时，为了缩短训练时间，防止过拟合，我们在训练过程中采用了批标准化 (Batch Normalization) [85] 和 Dropout[86] 等策略。经过约 200 个 epoch 的训练后模型收敛，训练过程中的 Loss 下降曲线如图 3.12 所示。

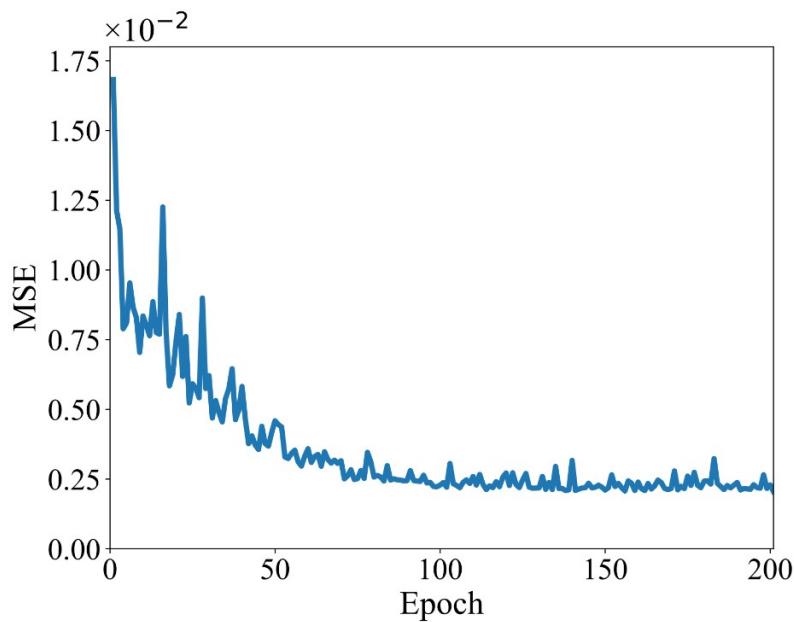


图 3.12 Loss 下降曲线

Figure 3.12 Learning curve of the DNN model.

最后，我们用测试集的样本对该模型进行了测试。我们将测试样本的 STL 频谱作为期望频谱并输入模型，得到模型输出的一组预测的等效电学参数，再根据表 3.1 中的公式反解出 THR 单元的几何参数，最后基于 TMM 方法和 FEM 方法计算该 THR 单元真实的 STL 频谱。图 3.13 中展示了某 6 个测试样本的测试结果，每幅图中蓝色的实线代表期望的 STL 频谱，黄色的虚线代表模型预测的 THR 单元基于 TMM 方法计算得到的 STL 频谱，绿色的点线代表模型预测的 THR 单元基于 FEM 方法计算得到的 STL 频谱，每个测试样本对应的 MSE 值也标记在了图片的正上方。由图 3.13 可知，这六个样本的真实 STL 频谱均与期望

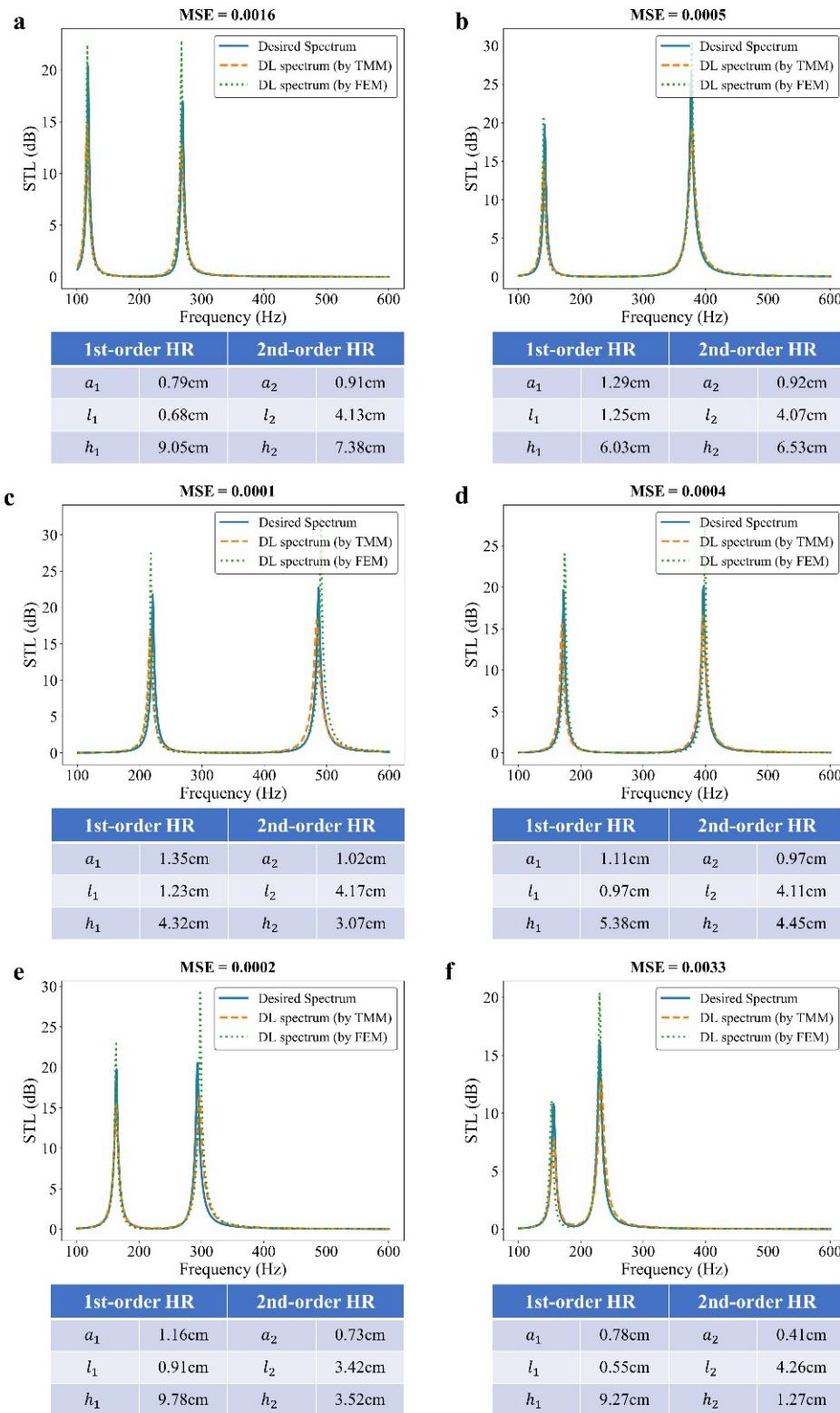


图 3.13 测试集样本的测试结果

Figure 3.13 Prediction results of the test samples.

值吻合的很好。

为了定量地评价模型在整个测试集上的效果, 我们引入了平均均方误差(Average Mean Square Error) \overline{MSE} 这一指标:

$$\overline{MSE} = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{n} \sum_{j=1}^n (y_j - \tilde{y}_j)^2 \right). \quad \dots (3.16)$$

其中, $\mathbf{y} = [y_1, y_2, \dots, y_n]$ 和 $\tilde{\mathbf{y}} = [\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n]$ 分别为模型的期望输出和实际输出; N 为测试集中的样本个数。最终我们所提出的模型在测试集上的 \overline{MSE} 为 0.0029, 参考图 3.13(f) 中 MSE 为 0.0033 的测试结果, 我们认为这个误差是在可接受的范围内的, 所得到的模型已经具备了足够的泛化性能。

3.6 二阶亥姆霍兹共鸣器单元的按需设计

3.6.1 设计流程

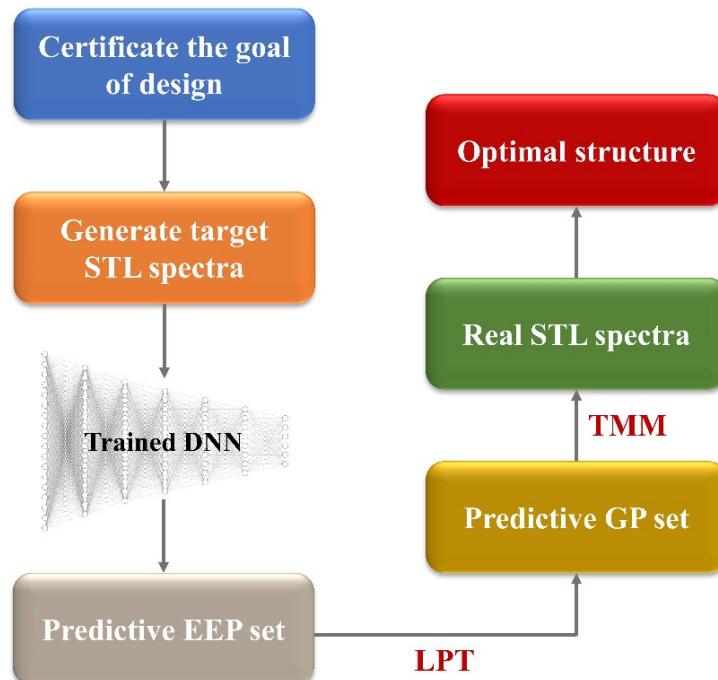


图 3.14 基于深度学习的 THR 单元按需设计流程图

Figure 3.14 Concept schematic of the design process.

实际应用中我们常常需要根据具体的噪声频率设计滤波器来实现降噪, 而不是期望的 STL 频谱, 并且满足我们的设计需求的 STL 频谱也是不唯一的。这

里我们以噪声频率为 150 Hz 和 250 Hz 为例对基于深度学习模型的 THR 单元的按需设计方法进行说明，我们的设计目标是希望结构在 150 Hz 和 250 Hz 处的 STL 大于 10dB，设计流程图如图 3.14 所示。具体设计方法如下：

1. 产生 N 条形状不尽相同的 STL 频谱，要求频谱在 150 Hz 和 250 Hz 有两个共振峰，且共振峰的高度大于 10 dB；
2. 将上述频谱输入到深度学习模型中，得到 N 组等效电学参数；
3. 基于得到的 N 组等效电学参数计算得到对应的 N 组几何参数；
4. 将这 N 组几何参数依次代入传递矩阵公式，基于 TMM 计算对应的 STL 频谱并判断是否满足设计需求。若出现满足设计需求的结果，则保留对应的几何参数作为最终的设计结果并终止计算；若 N 组几何参数均不满足需求，则增大 N 值并回到步骤 1.

其中，步骤 1 中所述的产生 STL 频谱的方式有很多种选择。比如，我们可以基于一阶亥姆霍兹共鸣器产生一条只包含一个共振峰的 STL 曲线并对其进行归一化作为标准曲线；然后，将上述标准曲线根据设计需求分别进行两次平移和缩放并叠加，得到一条包含两个峰的曲线，且该曲线峰的位置、高度和宽度均符合设计目标，由此生成一条符合要求的 STL 频谱。同时，由于深度学习模型的推理过程是并行的，步骤 2 中一次性输入 N 条 STL 频谱并不会带来额外的计算负担。

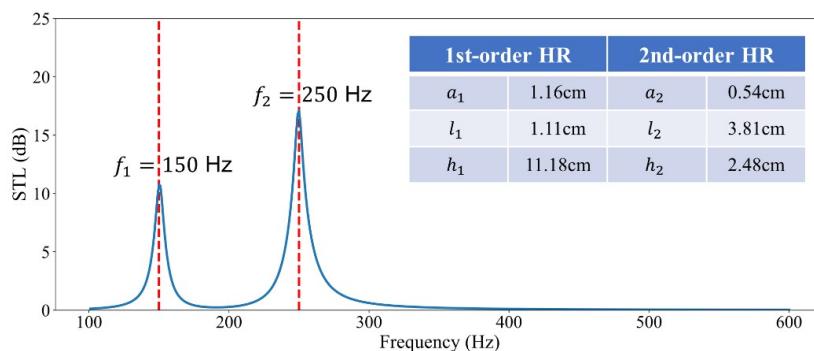


图 3.15 设计目标为在 150 Hz 和 250 Hz 实现 STL 大于 10dB 的设计结果

Figure 3.15 Results of the inverse design whose STL is more than 10 dB at 150 Hz and 250 Hz.

图 3.15 为最终的设计结果，所得到的 STL 频谱在 150 Hz 和 250 Hz 均存在一个大于 10 dB 的共振峰，符合我们的设计需求，该 THR 单元对应的几何参数

展示在图 3.15 右上角表格中。

3.6.2 与优化类算法的性能对比

为了进一步展现我们所提出的基于深度学习（Deep Learning, DL）的设计方法在计算效率上的优越性，在本小节我们将其与第 2 章所介绍的基于 GA 的设计方法进行了对比，从计算效率和设计效果两方面展开讨论。

在上一小节我们详细介绍了基于 DL 的声学结构设计方法，大致分为 4 个步骤。其中步骤 2 是 DNN 的正向推理过程，计算效率很高；而与步骤 4 相比，步骤 1 和步骤 3 的计算成本可以忽略不计，因此 DL 方法的计算复杂度主要由步骤 4 决定，即根据指定的几何参数基于 TMM 正向计算其 STL 频谱。而使用 GA 对结构进行设计时，算法的主要的计算复杂度也是由计算每个个体对应的 STL 频谱贡献的，与步骤 4 的计算内容一致。因此，我们可以通过比较两种方法所需的正向计算次数来对两种方法的计算效率进行对比。在使用 GA 对结构进行优化设计时，假定种群规模为 n ，进化 m 代后出现符合要求的个体，这一过程大约需要进行 $2nm$ 次正向计算， $2nm$ 通常是 10^3 量级的；而基于 DL 进行结构设计时，正向计算的次数基本处于 10^1 到 10^2 量级，因此具有明显的速度上的优势。

为了进一步对这两种方法的计算效率进行对比，我们分别基于这两种方法对在目标隔声频率处 STL 大于 10 dB 的 THR 单元进行了设计。所使用的 GA 是基于开源工具箱 Geatpy 中的 soea_SEGA 所实现的，初始种群规模为 100，所采用的适应度函数 F 为：

$$F = \begin{cases} 1000 & \text{if } t(f_1^{\text{target}}), t(f_2^{\text{target}}) > 10 \text{ dB} \\ t(f_1^{\text{target}}) + t(f_2^{\text{target}}) - \lambda & \text{else} \end{cases} \dots (3.17)$$

其中， $t(f_1^{\text{target}})$ 和 $t(f_2^{\text{target}})$ 分别为结构在目标频率 f_1^{target} 和 f_2^{target} 处的 STL 值； λ 为用于辅助进化的惩罚因子：

$$\lambda = \begin{cases} 50 + |f_1 - f_1^{\text{target}}| + |f_2 - f_2^{\text{target}}| & \text{if } 101 \leq f_1, f_2 \leq 600 \\ 1000 & \text{else} \end{cases} \dots (3.18)$$

其中， f_1 和 f_2 分别为结构的一阶共振频率和二阶共振频率。 λ 的引入可以在很大程度上引导结构的共振频率向 f_1^{target} 和 f_2^{target} 方向进化，从而更快地达到期望的设计目标。我们使用上述适应度函数对 THR 单元的 6 个待定几何参数

$[a_1, l_1, h_1, a_2, l_2, h_2]$ 进行寻优，考虑到结构的可加工性，几何参数调节范围被设置为 $0.1 \text{ cm} < a_i < 2.5 \text{ cm}$, $0.1 \text{ cm} < l_i < 5 \text{ cm}$, $0.1 \text{ cm} < h_i < 25.5 \text{ cm}$ ($i = 1, 2$)。在整个优化过程中，当种群中出现了适应度为 1000 的个体时便停止计算。

这里我们选取了三组目标隔声频率，并且为了不失一般性，我们针对每组设计目标都分别基于 DL 和 GA 进行了 20 次求解并记录了每次求解所用的时间，如图 3.16 所示。第 1 组目标频率为 $f_1^{\text{target}} = 150 \text{ Hz}$ 和 $f_2^{\text{target}} = 250 \text{ Hz}$ ，基于 DL 的 20 次设计的平均耗时为 8.7s，基于 GA 的 20 次设计的平均耗时为 27.5s。第 2 组目标频率为 $f_1^{\text{target}} = 250 \text{ Hz}$ 和 $f_2^{\text{target}} = 400 \text{ Hz}$ ，基于 DL 的 20 次设计的平均耗时为 2.5s，基于 GA 的 20 次设计的平均耗时为 43.1s；第 3 组目标频率为 $f_1^{\text{target}} = 300 \text{ Hz}$ 和 $f_2^{\text{target}} = 450 \text{ Hz}$ ，基于 DL 的 20 次设计的平均耗时为 4.4s，基于 GA 的 20 次设计的平均耗时为 55.3s。我们从上述三个对比实验的最终计算结果中各选取了一组 THR 单元，并基于 TMM 方法计算了结构对应的 STL 频谱，如图 3.17 所示。可以看到，通过这两种方式设计得到的 THR 单元在性能方面并无明显的差异，都与预期目标吻合的很好。因此，与传统的优化类算法相比，所提出的基于深度学习的设计方法具有更高的计算效率，模型一旦训练完成，我们可以在很短的时间内设计出满足需求的 THR 单元，这为我们实现低频宽带隔声装置的设计提供了很大的便利。

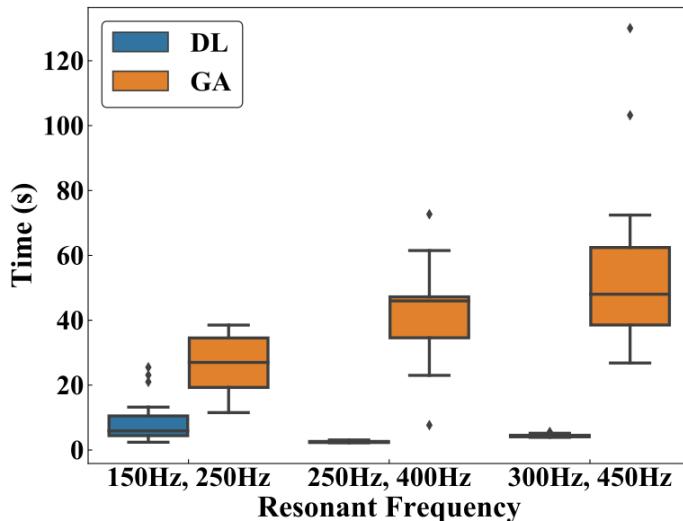


图 3.16 深度学习与 GA 的设计效率对比

Figure 3.16 Comparison of the computational efficiency of deep learning and genetic algorithm

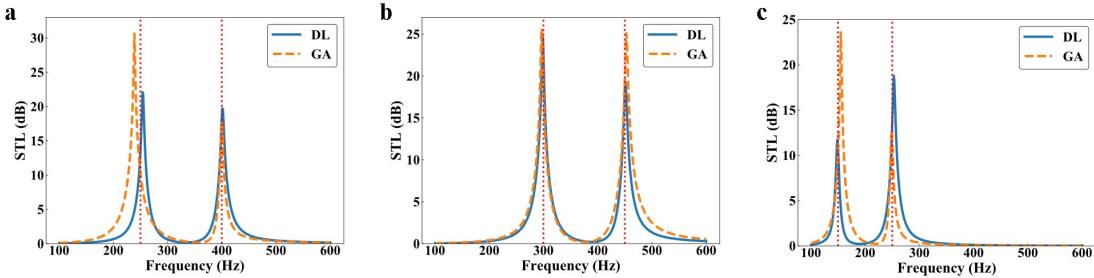


图 3.17 针对三个不同的优化目标的优化结果对比

Figure 3.17 Comparison of the optimization results for three different optimization objectives

3.6.3 实验结果

为了进一步证明我们所提出的基于深度学习的逆设计方法的有效性，我们基于此方法设计了两个 THR 单元用于对 150 Hz, 200 Hz, 250 Hz 和 300 Hz 的噪声进行降噪，并按照深度学习模型所预测的几何参数制备了样品，在驻波管中进行了测试。样品是基于 3D 打印技术制备的，所使用的材料为光敏树脂，厚度为 5 mm，密度为 1.13 g/cm³。且该光敏树脂的阻抗远大于空气，因此可被视为刚性边界。这两个 THR 单元的样品照片及其几何参数如图 3.18 所示，左侧的 THR 单元的共振频率为 150 Hz 和 250 Hz，几何参数记录在左侧橙色的表格中；右侧的 THR 单元的共振频率为 200 Hz 和 300 Hz，几何参数记录在右侧的绿色表格中。

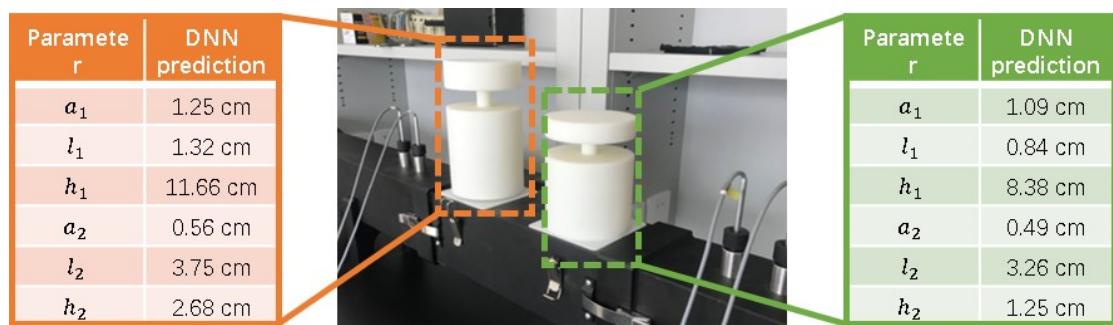


图 3.18 THR 单元的样品照片及其几何参数

Figure 3.18 Photo and GPs of the THR.

测量所使用的方形驻波管的边长为 10 cm，上述两个 THR 单元作为方形驻波管的旁支结构放置在同一侧，连接处用橡皮泥进行密封。首先，我们基于双负载法对结构的 STL 频谱进行了测量，4 个 1/4 英寸的传声器 (Brüel & Kjær, type-4187) 用于接收信号，实验中所使用的两种边界条件分别是硬边界和吸声海绵。我们将实验所测得的 STL 频谱与仿真结果进行了对比，如图 3.19 所示：图中蓝

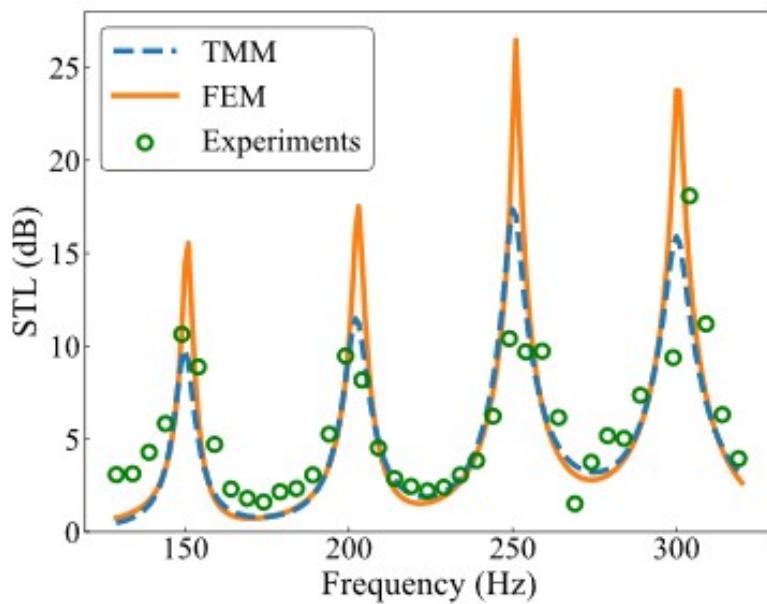


图 3.19 STL 频谱的实验和仿真结果对比

Figure 3.19 The theoretical,simulated and experimental STL spectrum of the THR.

色虚线为基于 TMM 的仿真结果，黄色实线为基于 FEM 的仿真结果，绿色圆圈为实验结果。结果表明，在 150 Hz, 200 Hz, 250 Hz, 300 Hz 这四个目标频率，所设计的结构均实现了 10 dB 以上的降噪效果，与预期目标相吻合。

为了进一步展现该结构的精准降噪性能，我们将一段纯净的语音信号与上述四个频率的噪声相混合，通过方管左端口的扬声器入射至方管内。方管的右端口放置了吸声棉，并在中心处放置了一个 1/2 英寸的传声器 (Brüel & Kjær, type-4189) 用于接收信号。若 THR 单元未作为方管的旁支结构，右端口接收所接受的信号为降噪前的语音信号，其时域波形图如图 3.20(a) 所示，语谱图如图 3.21(a) 所示。当 THR 单元作为方管的旁支结构放置时，结构会过滤掉 150 Hz, 200 Hz, 250 Hz, 300 Hz 这四个目标频率的噪声，因此右端口所接受的信号为降噪后的语音信号，其时域波形图如图 3.20(b) 所示，语谱图如图 3.21(b) 所示。由图 3.20 和图 3.21 可知，所设计的两个 THR 单元在显著降低了线谱噪声的能量的同时，并未对语音信号本身产生过多的影响，实现了精准降噪。

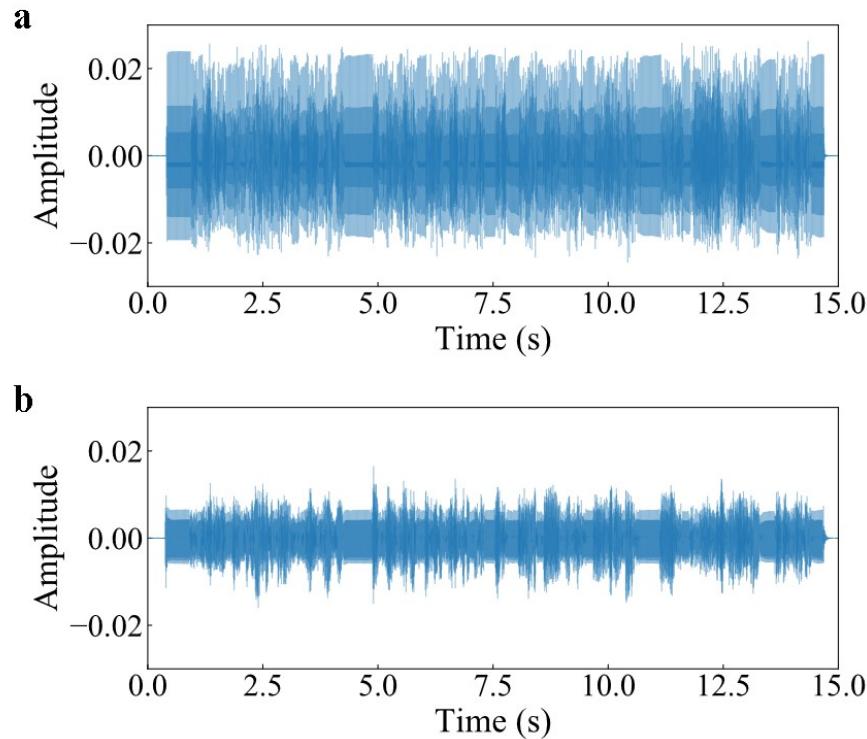


图 3.20 降噪前后时域波形对比：(a) 降噪前；(b) 降噪后。

Figure 3.20 The time-domain waveforms of the test signal (a) before and (b) after filtering.

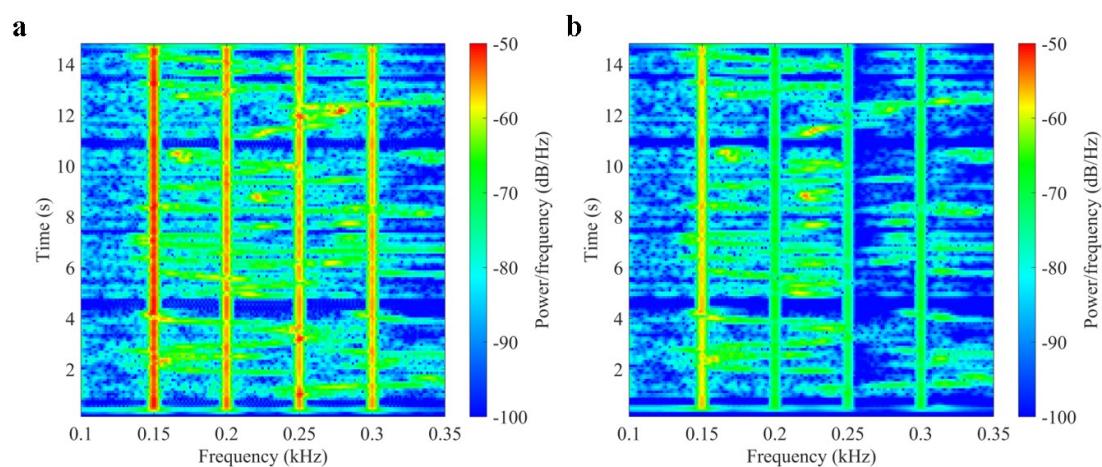


图 3.21 降噪前后语谱图对比：(a) 降噪前；(b) 降噪后。

Figure 3.21 Spectrograms of the test signal (a) before and (b) after filtering.

3.7 基于多个二阶亥姆霍兹共鸣器的低频宽带隔声装置设计

虽然与一阶亥姆霍兹共鸣器相比，THR 在不增加额外体积的基础上又引入了额外的高阶共振频率，但是单独的一个 THR 依然只能针对两个共振峰附近频带内的噪声进行降噪。因此，对于宽带噪声我们常常需要对多个 THR 进行组合，以实现低频宽带隔声。在本章中，我们基于所提出的深度学习模型对多个 THR 设计并组合，完成了低频宽带隔声器件的设计工作。所设计的低频宽带隔声器件可以作为管道的旁支结构，在不影响管道中流体流速的基础上对低频宽带的管路噪声进行降噪。此项工作进一步证明了深度学习模型用于声学结构设计的优越性，也为实现低频宽带隔声提供了新的技术路径。

3.7.1 耦合效应

多个 THR 单元通常如图 3.22 所示作为管道的旁支结构依次串联排列，其中 THR 单元间的间隔为 L 。为了使隔声装置更加紧凑，相邻的单元间的距离一般比较小，所以单元间会存在一定的耦合效应。

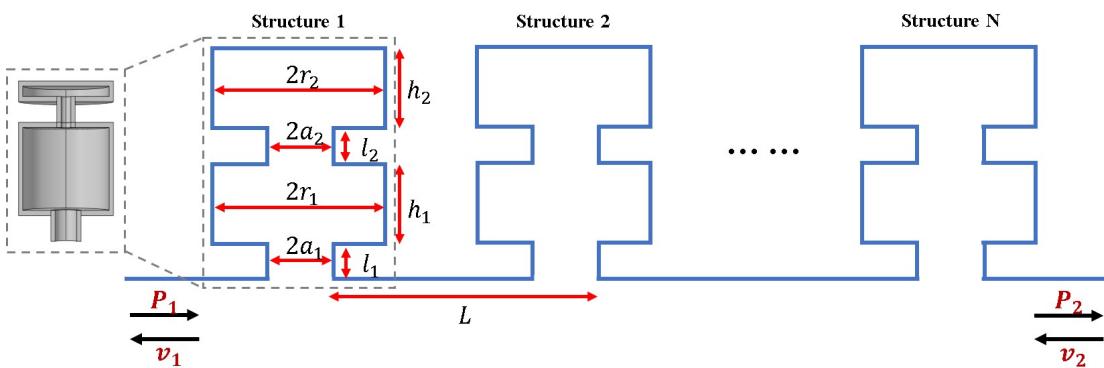


图 3.22 多个 THR 单元串联排列示意图

Figure 3.22 Schematic view of several THRs in series.

为了更加形象地展现单元间的这种耦合效应，下面以三个 THR 单元的不同组合方式为例展开分析。为了方便叙述，这三个 THR 单元分别被编号为 1 号、2 号和 3 号，对应的几何参数如表 3.3 所示。同时，我们也借助 FEM 对这三个 THR 单元单独作为管道旁支结构时的 STL 频谱进行了计算，如图 3.23(a) 所示。这里的黄色虚线为 1 号单元的 STL 频谱，其共振峰分别出现在 251 Hz 和 430 Hz；绿色的点线为 2 号单元的 STL 频谱，其共振峰分别出现在 245 Hz 和 454 Hz；红色点划线为 3 号单元的 STL 频谱，其共振峰分别出现在 275 Hz 和 504 Hz。这三个

单元单独作为管道旁支结构时的共振频率也记录在表3.3的最后两行。

表3.3 3个THR单元的几何参数及共振频率

Table 3.3 Geometric parameters and resonance frequencies of the three THRs.

	Structure 1	Structure 2	Structure 3
a_1/cm	1.38	1.89	1.98
l_1/cm	1.12	0.53	0.13
h_1/cm	6.85	10.82	8.58
a_2/cm	0.96	1.12	0.83
l_2/cm	3.88	2.35	3.59
h_2/cm	1.88	3.89	1.87
f_1/Hz	251	245	275
f_2/Hz	430	454	504

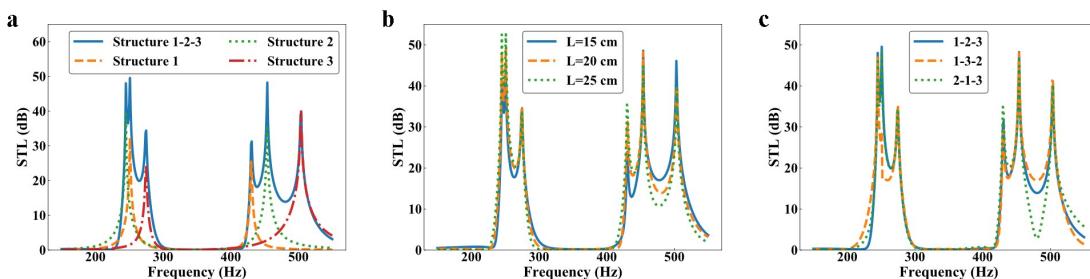


图3.23 多个THR单元间的耦合效应

Figure 3.23 Coupling effect between the THRs.

然后，我们将这三个THR单元以图3.22的形式，从左到右按照1号结构、2号结构和3号结构的顺序依次排列，间距为20 cm，基于FEM计算了该组合结构的STL频谱，如图3.23(a)中蓝色实线所示。我们可以看到，该组合结构的STL具有6个共振峰，分别出现在245 Hz, 251 Hz, 275 Hz, 431 Hz, 454 Hz, 504 Hz，高度均在30 dB以上。与每个单元单独作为旁支结构相比，组合后结构的共振频率基本和每个单元的共振频率吻合，这说明在这种排列方式下耦合效应基本不会造成共振频率的偏移。同时，由于耦合效应的存在，组合结构共振频率附近的STL有了较为明显的提升，其最终的隔声效果并不是每个单元隔声量的简单叠加。

其次，我们改变了组合结构单元之间的间距L，基于FEM计算了不同单元间距的组合结构的STL频谱，如图3.23(b)所示。其中，蓝色实线为间距L=15 cm

时组合结构的 STL 频谱，黄色虚线为间距 $L=20\text{ cm}$ 时组合结构的 STL 频谱，绿色点线为间距 $L=25\text{ cm}$ 时组合结构的 STL 频谱。观察图 3.23(b) 的三条曲线可知，改变单元之间的间距依然不会造成共振频率的明显偏移，但是会对共振频率附近的 STL 造成一定的影响。

最后，我们还探究了单元顺序对组合结构 STL 频谱的影响。我们改变了三个单元的相对位置，同时固定单元间距 $L=20\text{ cm}$ 不变，基于 FEM 计算了不同顺序下组合结构的 STL 频谱，如图 3.23(c) 所示。其中，蓝色实线为将三个单元按照 1 号、2 号和 3 号的顺序依次排列后的组合结构（以下简称“1-2-3”组合）的 STL 频谱；黄色虚线为将三个单元按照 1 号、3 号和 2 号的顺序依次排列后的组合结构（以下简称“1-3-2”组合）的 STL 频谱；绿色点线为将三个单元按照 2 号、1 号和 3 号的顺序依次排列后的组合结构（以下简称“2-1-3”组合）的 STL 频谱。对比“1-3-2”组合的 STL 频谱和另外两条频谱，可以看到“1-3-2”组合的 STL 频谱在 251 Hz 处的共振峰的高度明显低于另外两条频谱。该共振峰为 1 号单元的一阶共振峰，且与 2 号单元的一阶共振峰在频率上较为接近。而在“1-3-2”组合中，1 号单元和 2 号单元间放置了一个 3 号单元，导致两个单元相距较远，因此这两个共振峰间的耦合效应被削弱了，致使“1-3-2”组合在 251 Hz 处的共振峰的高度更接近于 1 号单元原本的一阶共振峰的高度，并没有因为耦合效应的存在有明显的提升。类似的效应也存在于“2-1-3”组合第五个共振峰和第六个共振峰的中间区域，在这个频带内“2-1-3”组合的 STL 频谱较另外两条频谱相比出现了很明显低谷。这其实是因为组合中 2 号单元和 3 号单元距离较远，因此 2 号单元的二阶共振峰和 3 号单元的二阶共振峰间的耦合效应被削弱了，导致组合结构在该频率区间的 STL 较低。

综上所述，由于耦合效应的存在，对组合结构中的单元进行独立设计并不是最优的选择。我们应该在对组合结构设计的过程中充分考虑单元间的耦合效应，并对其加以利用。例如，可以将共振频率相近的单元尽可能放置在相邻的位置，以达到提升组合结构的隔声效果的目的。

3.7.2 设计方法

由上一章节的分析可知，在设计宽带隔声装置的过程中，如果忽视单元间的耦合效应生成一组共振频率有特定间隔分布的结构并将他们组合，最终的设计结果并不理想，不仅隔声频带较窄，还可能会出现一些隔声效果很差的频点。因

此，我们在设计时应该充分考虑并利用单元间的耦合效应。这里我们基于贪婪算法的思想提出了一种宽带隔声装置的设计方法，设计流程如图3.24所示，核心思想如下：

1. 将有隔声需求的频带划分为低频区和高频区，同时确定一个阈值 T ，我们的目标是利用 THR 单元的一阶共振峰使低频区的 STL 均大于 T ，同时利用 THR 单元的二阶共振峰使高频区的 STL 均大于 T ；
2. 找到两个区域内 STL 小于 T 的最低的频点 $[f_{10}, f_{20}]$ ，基于深度学习模型生成 N 个共振频率分别在 $[f_{10} - \delta_1, f_{10} + \delta_1]$ 和 $[f_{20} - \delta_2, f_{20} + \delta_2]$ 范围内的 THR 单元，其中 δ_1 和 δ_2 为事先确定的搜索范围；
3. 将上述 N 个结构和已有结构进行组合，计算组合结构的 STL，并根据一定的标准计算每个组合结构的得分 S ，选择得分最高的结构；
4. 判断两个区域内的各个频点处的 STL 是否都大于 T ，若是则结束计算，否则回到步骤 2。

其中，步骤 3 所述的打分标准可以根据具体的应用需求自行进行设计，且该标准会很大程度上影响整个装置的设计过程及最终设计结果，这里给出 2 个范例：

标准 1：计算组合结构两个区域内 STL 小于 T 的最低的频点 f_{11} 和 f_{21} ，组合结构的得分可以被表示为 $S = f_{11} + f_{21}$ 。此标准可以使每一步迭代的步长尽可能的大，以减少所用 THR 单元的个数。

标准 2：计算组合结构 STL 频谱中低于 T 的频点个数 t_{num} ，组合结构的得分可以被表示为 $S = 1 / t_{num}$ 。此标准意在确保每一个结构都能使得更多的频点达到既定的隔声量标准。

由于低频的共振峰往往较为尖锐，高度也较低，在设计过程中低频部分的设计进度往往是落后于高频部分的，可能会出现高频区的所有频点的隔声量都达标了，但是低频区仍有频带隔声量不达标。为了避免这种情况的发生，我们可以在划分区域时适当增大高频区的带宽，减少低频区的带宽；在制定打分标准时也可以对低频区域和高频区域分开进行打分，并赋予低频部分更高的权重，使得在挑选结构时优先考虑该结构对低频区隔声量的提升效果。

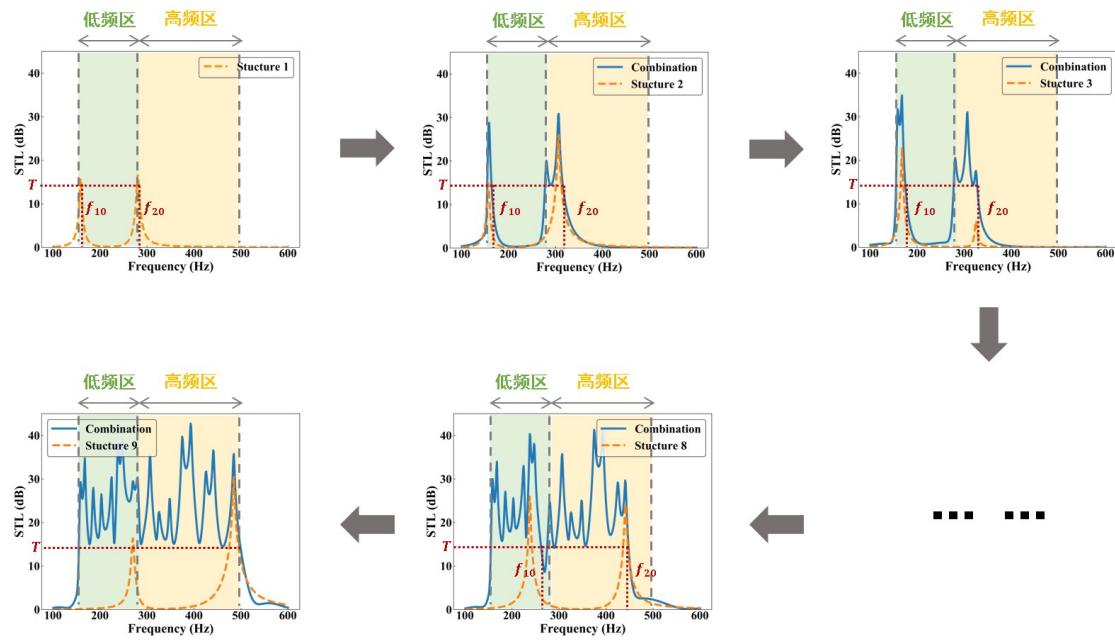


图 3.24 低频宽带隔声装置的设计过程

Figure 3.24 Design process of the low-frequency broadband sound insulation device

3.7.3 隔声效果

基于 3.7.2 中所述的设计方法，我们设计了如图 3.25 所示的低频宽带隔声装置，共包含了 9 个 THR 单元，每个单元的间隔为 20 cm，各个 THR 单元对应的几何参数和共振频率如表 3.4 所示。组合后的宽频隔声装置的 STL 频谱的仿真计算结果如图 3.26 所示，蓝色虚线为基于 FEM 的计算结果，黄色实线为基于 TMM 算法的计算结果。由于存在较强的耦合效应，基于 FEM 的结果和基于 TMM 的结果之间存在一定的差异，且这种差异性在高频更加明显。但是，二者所呈现出的大体趋势是一致的，且 TMM 算法的计算复杂度更低，因此在设计宽频器件的过程中使用 TMM 算法可以大大提升设计效率。

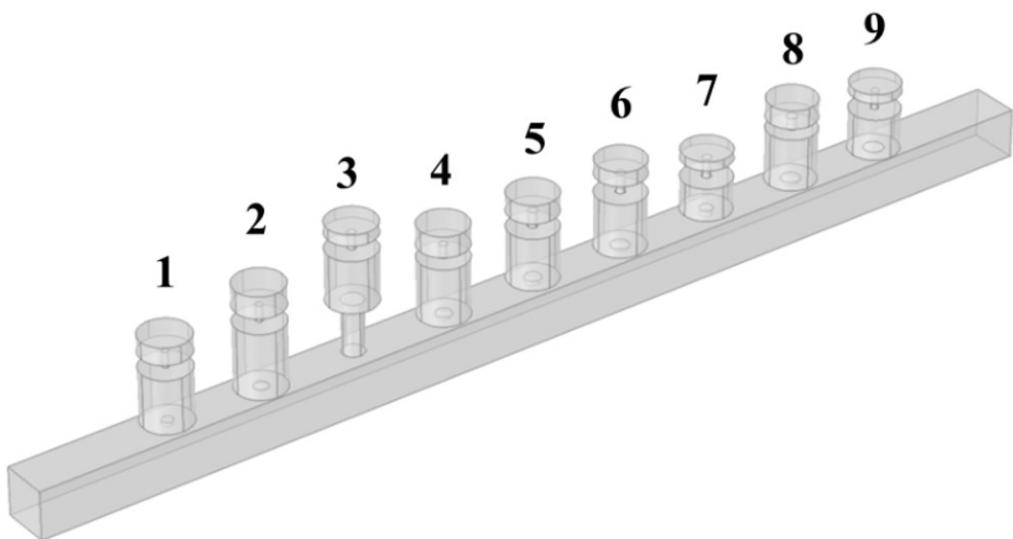


图 3.25 宽频隔声装置结构示意图

Figure 3.25 Design process of the low-frequency broadband sound insulation device

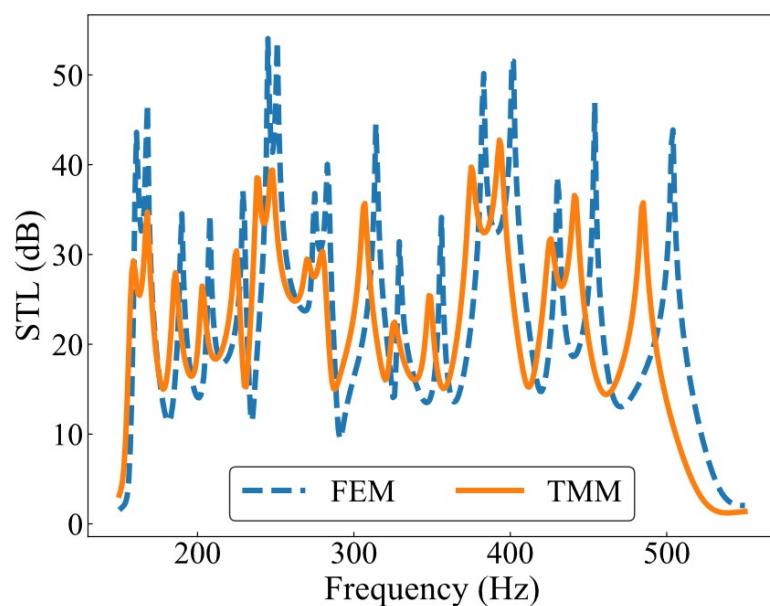


图 3.26 宽频隔声装置的 STL 频谱

Figure 3.26 STL spectrum of the low-frequency broadband sound insulation device.

表 3.4 低频宽带隔声装置的 THR 单元的几何参数及共振频率**Table 3.4 Schematic view of the low-frequency broadband sound insulation device.**

结构编号	1	2	3	4	5	6	7	8	9
a_1/cm	1.13	1.46	2.23	1.61	1.72	1.83	1.38	1.89	1.98
l_1/cm	0.77	0.24	10.75	0.54	0.81	0.46	1.12	0.53	0.13
h_1/cm	11.08	13.57	10.68	11.93	10.82	11.27	6.85	10.82	8.58
a_2/cm	0.69	0.73	0.85	0.83	0.97	0.98	0.96	1.12	0.83
l_2/cm	3.13	3.25	3.19	2.54	3.22	3.84	3.88	2.35	3.59
h_2/cm	3.55	4.48	2.30	4.31	3.84	3.00	1.88	3.89	1.87
f_1/Hz	159	157	168	186	203	225	248	238	270
f_2/Hz	280	307	325	349	375	393	425	441	485

3.8 本章小结

本章将深度学习与 LPT 模型相结合，提出了一种从需求出发的声学结构逆设计方法，并以 THR 结构为例结合仿真与实验的结果证明了所提方法的有效性。首先对声学结构的正向计算方法进行了介绍，同时指出虽然基于给定的几何参数计算其声学性质的正向计算过程是较为容易的，但从期望的声学性质出发对几何参数进行逆设计是十分具有挑战性的，并着重讨论了逆设计的复杂性和非唯一性。因此，借助深度学习模型对逆设计问题进行求解是兼具学术意义和应用价值的。随后，我们详细介绍了用于模型训练的数据的产生方法、模型结构及训练过程，并给出了模型在测试集上的测试结果。此外，我们还结合具体案例对基于所提出的模型进行 THR 结构的逆设计过程进行了详细的阐述，并与优化类算法进行对比证明了所提出方法的优越性。最后，我们基于所提出的模型完成了低频宽带隔声装置的设计，用于解决实际应用中常见的低频宽带隔声需求。所提出的方法能精准高效地对声学结构进行逆设计，为推动复杂声学结构走向实际应用提供有力的理论基础和技术支持。

第4章 深度学习辅助的多样化声人工结构设计

4.1 引言

在第3章中，我们基于深度神经网络建立了期望的声学响应（STL频谱）和等效电学参数之间的映射关系，实现了声学结构的自动化精准逆设计，大大提升了声学结构的设计效率。考虑到单一的深度神经网络能力较为有限，仍有很多声学结构设计领域的问题无法被解决。因此，本章将深度神经网络与其他算法相结合，探索了深度学习在声学结构设计领域的拓展性应用，为声学结构的批量化设计和优化设计提出了具体的解决方案。由于我们所提出的模型拟合的是期望的物理响应和等效电学参数之间的映射关系，因此该模型不仅可以用于设计THR单元，还可以在LPT框架下向其他声学结构进行迁移。

4.2 二阶亥姆霍兹共鸣器的批量化设计

在3.3.2章节中，我们分析了声学结构的物理性质和几何参数之间非唯一的对应关系，考虑到深度神经网络只能拟合“一对一”的映射关系，我们又采取了一系列手段弱化了期望的物理响应和结构的几何参数和之间的“一对多”效应，从而确保模型的收敛。然而，虽然我们通过这种方式完成了模型的训练，但是所构建的模型对于一个输入通常也只会输出一组可以被物理实现的几何参数，因此无法高效地预测出满足我们设计需求的所有结构。且这些满足我们设计需求的结构虽然具有相似的物理性质，但是他们的几何参数可能相差很大，因此可以看做是一些散落在几何参数空间内的点，很难通过暴力搜索的方式快速地将他们找到。因此，在这一章节我们将所提出的深度学习模型和降维类算法相结合，并引入一种经典的降维算法主成分分析（Principal Component Analysis, PCA）算法对声学结构的几何参数进行变换，在变换后的PCA空间，具有相似物理性质的声学结构往往聚集在一起。因此，我们只需要通过深度学习模型得到满足期望的一个声学结构，并基于PCA算法求解出该结构在PCA空间的坐标系数，然后以该结构为中心在PCA空间内进行小范围的搜索，即可快速获得大量具有期望的物理性质的声学结构，从而实现THR单元的批量化设计。

4.2.1 主成分分析

伴随着信息化进程的不断推进，大规模数字化技术的应用在极大地便利了人们的生产生活的同时，也带来了数据规模的急速膨胀。虽然这些数据的背后往往潜藏着大量可被利用的信息，但是这些数据往往是多特征、强耦合的高维数据，为后续的分析和处理造成了困难。因此，人们希望能够选取合适的数学模型对原始数据的最本质特性进行挖掘，提取数据之间的关联结构，整合形成新的数据特征，这一类算法被称之为降维类算法。

作为一种经典的降维类算法，PCA 算法可以将原来样本数据以方差最大化为原则投影至一个新的空间（即 PCA 空间），这一投影过程可以视为一次线性变换 [87]。因此，我们可以通过 PCA 算法将声学结构 6 维的几何参数特征无损地线性变换至 6 维的 PCA 空间中。变换后的 PCA 空间中的点可以通过 6 个正交向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_6$ 来描述，这 6 个正交向量也被称为主成分。因此，对于声学结构 β 我们可以在几何参数空间和 PCA 空间之间建立如下的对应关系：

$$\mathbf{g}^\beta = \mathbf{v}_0 + p_1^\beta \mathbf{v}_1 + p_2^\beta \mathbf{v}_2 + p_3^\beta \mathbf{v}_3 + p_4^\beta \mathbf{v}_4 + p_5^\beta \mathbf{v}_5 + p_6^\beta \mathbf{v}_6 \quad \dots (4.1)$$

其中， $\mathbf{g}^\beta = [a_1^\beta, l_1^\beta, h_1^\beta, a_2^\beta, l_2^\beta, h_2^\beta]^T$ 为声学结构 β 的几何参数的向量表示； \mathbf{v}_0 为用于去中心化的常数向量； $p_1^\beta, p_2^\beta, \dots, p_6^\beta$ 为结构 β 在 PCA 空间内的坐标系数。从几何参数空间到 PCA 空间的变换过程在数学上可以被建模为投影矩阵 \mathbf{B} ，矩阵中每一行即为主成分向量，且按照重要程度从上至下排列。PCA 的具体的计算过程如下：

1. 在数据集中收集 N 个具有相似物理性质的声学结构，将其对应的几何参数向量排列得到训练集矩阵 $\mathbf{X} = [\mathbf{g}^1, \mathbf{g}^2, \dots, \mathbf{g}^N] \in \mathbb{R}^{6 \times N}$
2. 计算矩阵 \mathbf{X} 的每一行的均值 $\boldsymbol{\mu}$ 和标准差 $\boldsymbol{\sigma}$ ：

$$\begin{aligned} \boldsymbol{\mu} &= [\mu_1, \mu_2, \dots, \mu_6]^T, \\ \mu_i &= \frac{1}{N} \sum_{j=1}^N x_{ij}, \\ \boldsymbol{\sigma} &= [\sigma_1, \sigma_2, \dots, \sigma_6]^T, \\ \sigma_i &= \sqrt{\frac{1}{N} \sum_{j=1}^N (x_{ij} - \mu_i)^2} \end{aligned} \quad \dots (4.2)$$

其中， x_{ij} 为矩阵 \mathbf{X} 第 i 行第 j 列的元素； μ_i 为均值向量 $\boldsymbol{\mu}$ 的第 i 个分量 ($i = 1, 2, \dots, 6$)； σ_i 为标准差向量 $\boldsymbol{\sigma}$ 的第 i 个分量 ($i = 1, 2, \dots, 6$)；

3. 对矩阵 \mathbf{X} 进行标准化，得到标准化的训练集矩阵 $\bar{\mathbf{X}} \in \mathbb{R}^{6 \times N}$ ，该矩阵的第 i 行 j 列的元素 \bar{x}_{ij} 为：

$$\bar{x}_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i} \quad \dots (4.3)$$

4. 计算矩阵 $\bar{\mathbf{X}}$ 的协方差矩阵 $\mathbf{C} = \frac{1}{N} \bar{\mathbf{X}} \bar{\mathbf{X}}^T \in \mathbb{R}^{6 \times 6}$ ；

5. 计算协方差矩阵 \mathbf{C} 的特征值和特征向量，并按特征值从大到小将相应的特征向量进行排列，得到 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_6 \in \mathbb{R}^{6 \times 1}$ ；

6. 将上述特征向量依次转置为行向量组成投影矩阵 $\mathbf{B} = [\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_6^T]^T \in \mathbb{R}^{6 \times 6}$

7. 基于上述投影矩阵 \mathbf{B} 即可实现从几何参数空间到 PCA 空间的变换，得到结构 β 在 PCA 空间内的坐标系数 \mathbf{p}^β ：

$$\mathbf{p}^\beta = \mathbf{B} \bar{\mathbf{g}}^\beta \quad \dots (4.4)$$

其中， $\bar{\mathbf{g}}^\beta \in \mathbb{R}^{6 \times 1}$ 为标准化后的结构 β 的几何参数； $\mathbf{p}^\beta = [p_1^\beta, p_2^\beta, \dots, p_6^\beta]^T \in \mathbb{R}^{6 \times 1}$ 为结构 β 在 PCA 空间内的坐标系数的向量形式。

由于具有相似物理性质的声学结构的几何参数可能会很不一样，因此他们在几何参数空间中往往相距较远，但这些结构的坐标会在 PCA 空间中聚集。因此，我们只需要通过深度学习模型得到满足期望的一个声学结构，并基于上述算法求解出该结构在 PCA 空间的坐标系数，然后以该结构为中心在 PCA 空间内进行小范围的搜索，即可快速获得大量具有期望的物理性质的声学结构。

4.2.2 计算案例

在 3.6.1 章节，我们针对具体的设计目标基于所提出的深度学习模型完成了对 THR 单元结构参数的预测，所得到的结构的 STL 在 150 Hz 和 250 Hz 均大于 10dB，如 3.15 所示。且通过 3.3.2 章节的分析可知，该逆设计问题实为一个“一对多”问题，即存在多个几何结构不尽相同的 THR，他们的 STL 均满足在 150 Hz 和 250 Hz 均大于 10 dB 这个设计需求。这里我们将以这个设计需求为例，结合 PCA 算法，快速地找到多个满足该设计需求的 THR 结构。

为了方便叙述，我们将 3.6.1 章节中深度学习所预测出的结构命名为 β_0 （即图 3.15 中所展示的结构），该结构的几何参数记作 \mathbf{g}^{β_0} 。考虑到我们的设计需求要求结构的共振频率处在 150 Hz 和 250 Hz 附近，因此我们选择表 3.2 中 No.2，

No.3, No.10 和 No.11 四组样本组成训练集矩阵 \mathbf{X} , 然后按照 4.1.1 中所介绍的 PCA 算法流程计算得到投影矩阵 \mathbf{B} , 最终得到结构 β_0 在 PCA 空间的坐标系数 $\mathbf{p}^{\beta_0} = [p_1^{\beta_0}, p_2^{\beta_0}, \dots, p_6^{\beta_0}]$ 。而对于 PCA 空间中任意点 $\mathbf{p}^i = [p_1^i, p_2^i, \dots, p_6^i]$, 我们也可以通过 PCA 的逆变换求解出对应结构的几何参数 $\mathbf{g}^i = [a_1^i, l_1^i, h_1^i, a_2^i, l_2^i, h_2^i]$ 。为了方便我们更好地筛选符合设计需求的结构, 这里我们沿用公式(3.15)中所定义的 AERF 来定量地描述这些结构的性能, 在本案例中 $f_1^{\text{target}}=150 \text{ Hz}$, $f_2^{\text{target}}=250 \text{ Hz}$ 。一个结构的 AERF 越小, 代表该结构越符合我们的预期。

为了可视化变换后的样本分布, 缩小搜索范围, 这里我们在 \mathbf{p}^{β_0} 附近沿着方差最大的两个主成分 \mathbf{v}_1 和 \mathbf{v}_2 的方向进行了遍历搜索, 搜索的结果如图 4.1 所示。该平面内的任意一点 $[p_1^i, p_2^i]$ 都对应一个 THR 单元结构 β_i , 由(4.1)式可知, 该结构的几何参数 $\mathbf{g}^{\beta_i} = [a_1^{\beta_i}, l_1^{\beta_i}, h_1^{\beta_i}, a_2^{\beta_i}, l_2^{\beta_i}, h_2^{\beta_i}]$ 可以被表示为 $\mathbf{g}^{\beta_i} = \mathbf{v}_0 + p_1^{\beta_i} \mathbf{v}_1 + p_2^{\beta_i} \mathbf{v}_2 + p_3^{\beta_i} \mathbf{v}_3 + p_4^{\beta_i} \mathbf{v}_4 + p_5^{\beta_i} \mathbf{v}_5 + p_6^{\beta_i} \mathbf{v}_6$ 。图 4.1 中黄色星星所在的位置对应的结构即为结构 β_0 , 而其周围暗红色区域内的点所对应的结构均在 150 Hz 和 250 Hz 拥有较好的隔声性能, 均为符合我们设计需求的 THR 结构, 且这些结构的几何参数往往具有较大的差异性。

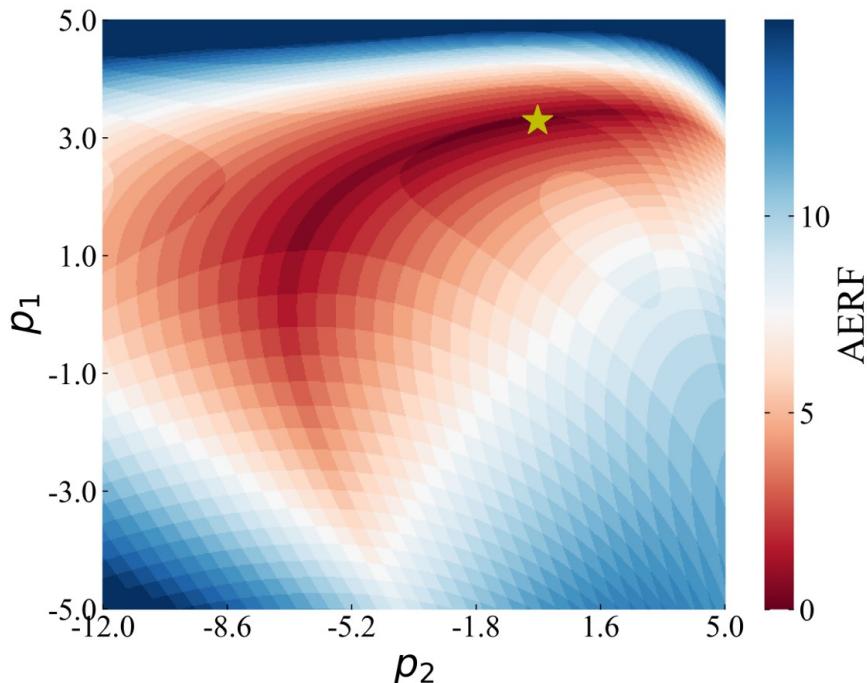


图 4.1 PCA 空间中 β_0 附近的结构的 AERF 分布

Figure 4.1 AERF varies with the two most important principal components p_1 and p_2 around β_0 in PCA space.

通过与 PCA 算法相结合，我们完成了声学结构的批量化设计，以较小的计算成本得到了一组符合我们预期的声学结构。在实际应用中，我们还可以对所得到的结构基于其他方面的特性进行二次筛选。例如，我们希望该结构在 150 Hz 处的 STL 能尽可能的大，同时在 250 Hz 处的 STL 仍然不小于 10 dB。针对这一筛选标准，我们定义筛选指标 H_1 为：

$$H_1 = \begin{cases} t(f_1^{\text{target}}, g) & \text{if } t(f_1^{\text{target}}, g), t(f_2^{\text{target}}, g) > 10 \text{ dB} \\ 0 & \text{else} \end{cases} \dots (4.5)$$

其中， $t(f_1^{\text{target}}, g)$ 和 $t(f_2^{\text{target}}, g)$ 分别为几何参数为 g 的 THR 结构在 f_1^{target} 和 f_2^{target} 频点处的 STL 值。我们可以在 PCA 空间中计算 β_0 附近结构的 H_1 ， H_1 越高越符合我们的设计要求。计算结果如图 4.2(a) 所示，颜色代表对应结构 H_1 的大小，我们可以在搜索范围内选择 H_1 最大的点（黄色星星）所对应的结构 β_1 ，该结构的 STL 频谱和几何参数如图 4.2(b) 所示。由计算结果可知，与深度学习模型直接预测出的结构 β_0 相比，结构 β_1 在保证 250 Hz 处的 STL 大于 10 dB 的基础上，在 150 Hz 处的 STL 提升至了 14.26 dB。

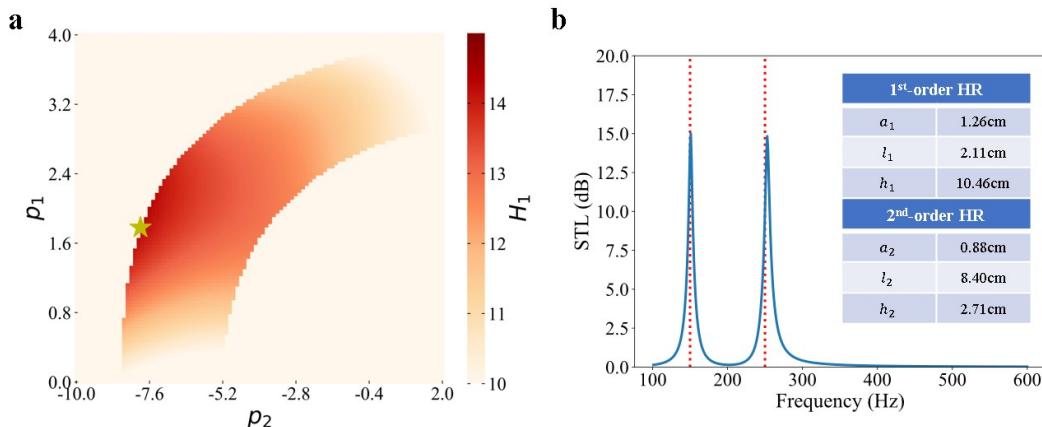


图 4.2 在 250 Hz 处的 STL 大于 10 dB 的基础上最大化 150 Hz 处的 STL

Figure 4.2 Searching for the THR whose STL in 150 Hz can be as high as possible and STL in 250 Hz can higher than 10 dB.

同理，我们也可以定义评价指标 H_2 用于计算并寻找在 250 Hz 处 STL 尽可能大的结构：

$$H_2 = \begin{cases} t(f_2^{\text{target}}, g) & \text{if } t(f_1^{\text{target}}, g), t(f_2^{\text{target}}, g) > 10 \text{ dB} \\ 0 & \text{else} \end{cases} \dots (4.6)$$

计算所得到 H_2 分布如图 4.3(a) 所示, 黄色星星所对应的结构为搜索区域内 250 Hz 处 STL 最大的结构 β_2 , 该结构的几何参数和 STL 频谱如图 4.3(b) 所示。由计算结果可知, 与深度学习模型直接预测出的结构 β_0 相比, 结构 β_2 在保证 250 Hz 处的 STL 大于 10 dB 的基础上, 在 250 Hz 处的 STL 提升至了 17.31 dB。

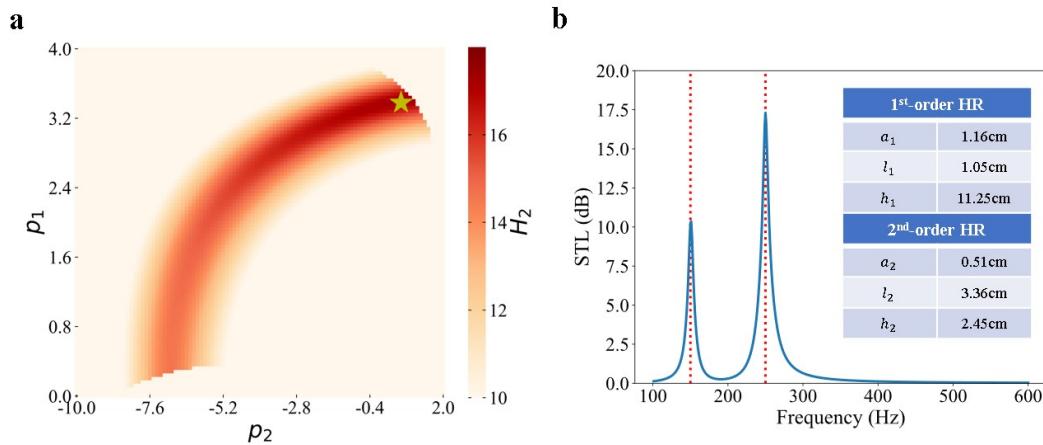


图 4.3 在 150 Hz 处的 STL 大于 10 dB 的基础上最大化 250 Hz 处的 STL

Figure 4.3 Searching for the THR whose STL in 250 Hz can be as high as possible and STL in 150 Hz can higher than 10 dB.

综上所述, 我们通过将深度学习模型与以 PCA 为代表的降维类算法相结合, 实现了声学结构的批量化设计, 能够在较短时间内找到多个具有期望物理性质的 THR 结构。同时, 我们也可以基于其他的评价标准对上述结构进行二次筛选, 以满足其他的应用需求。

4.3 二阶亥姆霍兹共鸣器的优化设计

在实际应用中, 我们经常会需要处理一些优化类问题, 如最大化某些频点处的隔声量等。然而, 深度学习模型往往只能用于预测某种符合要求的结构参数, 但并不能确保该结构性能的最优性, 因此并不适合用来独立解决这一类最优化问题。对于这类优化问题, 较为常见的一种解决手段是借助优化类算法进行求解。在第 2 章中, 我们介绍了基于 GA 对声人工结构进行优化设计的思路及相关案例。同时也指出, 诸如 GA、粒子群算法等启发式优化算法往往对初值是十分敏感的, 优秀的初始种群不仅能最大限度地避免算法过早地陷入局部最优, 还能加快进化速度降低计算成本。然而, 在大多数的应用场景下, 我们是无法获得一

个优秀的初始种群的，因此很多时候我们会选择对种群进行随机初始化，并且增大种群的规模，来避免进化陷入局部最优，使优化结果更加稳定。但这就会使计算成本大幅度增加，因此设计效率很低。在第4章中我们借助深度学习模型实现了指定共振频率的THR设计，虽然我们无法确保所预测的THR在指定的共振频率处拥有最大化的隔声效果，但对于“最大化指定频率隔声量”这一最优化问题而言，深度学习模型所给出的预测结果无疑是一个很好的初值。因此，本章节将深度学习模型所预测的声学结构作为一种先验知识，将其作为GA算法初始种群的精英个体参与后续的进化，并与不包含精英个体的情况进行了对比。仿真结果表明，深度学习模型通过为GA提供较好的初始条件，加快了进化速度，提升了最终的优化结果。

4.3.1 最大化目标频点处的STL

关于GA的基本原理在论文中的第2章已经进行了详细的介绍，这里不在赘述。在第4章中，我们基于深度学习模型实现了对指定共振频率的THR的几何参数的精准预测。这里我们将深度学习模型和GA相结合，希望能在确保目标频点处的STL不小于10 dB的基础上最大化这两个频点处的隔声量。这一优化目标可以被表示为：

$$\begin{aligned} \min_{\mathbf{g}} J = & \min_{\mathbf{g}} - \left[t(f_1^{\text{target}}, \mathbf{g}) + t(f_2^{\text{target}}, \mathbf{g}) \right] \\ \text{s.t. } & t(f_1^{\text{target}}, \mathbf{g}) > 10 \text{ dB} \text{ and } t(f_2^{\text{target}}, \mathbf{g}) > 10 \text{ dB} \end{aligned} \quad \dots (4.7)$$

其中， f_1^{target} 和 f_2^{target} 为优化的目标频率； $\mathbf{g} = [a_1, l_1, h_1, a_2, l_2, h_2]$ 为THR的几何参数； $t(f_1^{\text{target}}, \mathbf{g})$ 和 $t(f_2^{\text{target}}, \mathbf{g})$ 分别代表该THR在频率 f_1^{target} 和 f_2^{target} 处的STL值。

算法的初始种群包含50个个体，采用二进制编码方式，几何参数的优化范围与第4章相同。为了更好地展现深度学习模型对算法性能的提升效果，这里我们选取了两种初始种群的初始化条件分别进行实验：第一种情况考虑没有深度学习模型的辅助的情况，我们就无法获得关于该优化问题的先验知识，因此我们只能选择随机初始化这50个个体；第二种情况考虑有深度学习模型的辅助的情况，我们基于深度学习模型产生了5个共振频率在150 Hz和250 Hz并且隔声量不小于10 dB的样本作为种群的精英个体，其余45个个体进行随机初始化。我们基于上述两种初始条件利用GA对THR的几何参数进行优化，其他算法参数

保持一致。值得注意的是，除了初始种群，目标函数的选取同样也会对最终的优化结果产生很大的影响。为了更好的达成设计目标，这里我们定义优化的目标函数 F 为：

$$F = \begin{cases} -\left[t(f_1^{\text{target}}, \mathbf{g}) + t(f_2^{\text{target}}, \mathbf{g})\right] & \text{if } t(f_1^{\text{target}}, \mathbf{g}), t(f_2^{\text{target}}, \mathbf{g}) > 10dB \\ \lambda - \left[t(f_1^{\text{target}}, \mathbf{gp}) + t(f_2^{\text{target}}, \mathbf{gp})\right] & \text{else} \end{cases} \dots (4.8)$$

其中， λ 为惩罚因子，用于促进结构的共振频率向 f_1^{target} 和 f_2^{target} 进化，可以被进一步地表示成

$$\lambda = \begin{cases} 10 + |f_1 - f_1^{\text{target}}| + |f_2 - f_2^{\text{target}}|, & \text{if } 101 \leq f_1, f_2 \leq 600 \\ 1000, & \text{otherwise} \end{cases} \dots (4.9)$$

其中， f_1 和 f_2 分别为 THR 真实的一阶共振频率和二阶共振频率。

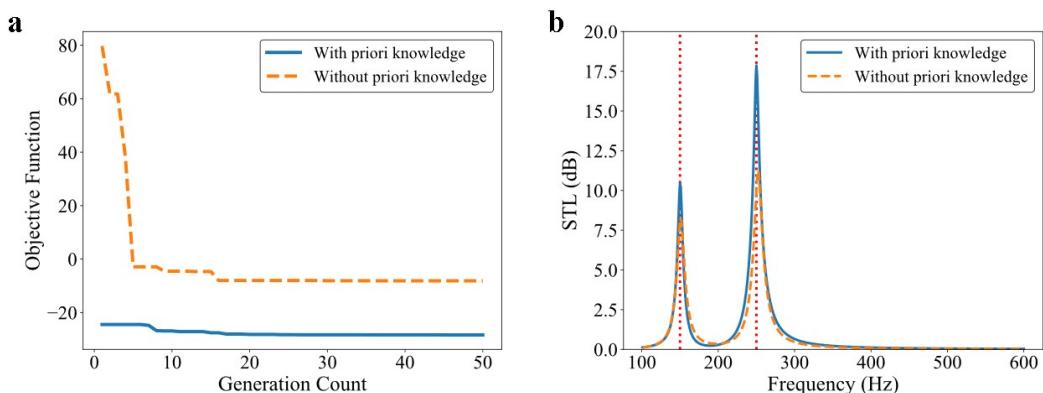


图 4.4 基于两种初始条件最大化 150 Hz 和 250 Hz 处的 STL

Figure 4.4 The optimization effect through the GA to maximize the average value of the STL at 150 Hz and 250 Hz with and without prior knowledge.

图 4.4 为将目标频点设为 150 Hz 和 250 Hz 时算法的优化结果对比，图 4.5 为将目标频点设为 300 Hz 和 450 Hz 时算法的优化结果对比。图 4.4(a) 和图 4.5(a) 为两种初始条件下进化曲线的对比，可以看出加入精英个体可以起到引导种群进化方向，加速算法的收敛的效果。图 4.4(b) 和图 4.5(b) 为经过 50 代进化后种群的最优个体的 STL 频谱对比。由于精英个体的加入可以在一定程度上避免算法过早地陷入局部最优，因此也会起到提升最优个体性能的作用。值得注意的是，由于深度学习模型一旦训练完成正向推理过程是十分迅速的，因此产生精英个体的计算成本非常的小，因此算法总体的优化效率依然得到了提升。

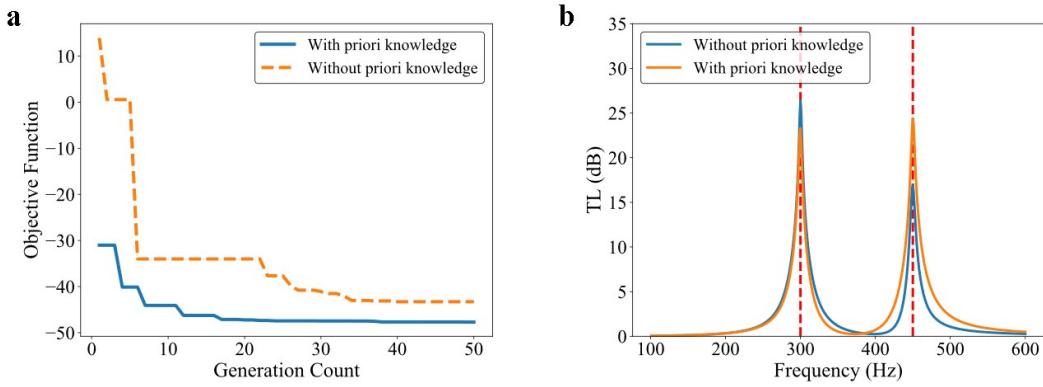


图 4.5 基于两种初始条件最大化 300 Hz 和 450 Hz 处的 STL

Figure 4.5 The optimization effect through the GA to maximize the average value of the STL at 3000 Hz and 450 Hz with and without prior knowledge.

若不引入惩罚因子 λ , 即令算法的目标函数为 $F' = -[t(f_1^{\text{target}}, \mathbf{g}) + t(f_2^{\text{target}}, \mathbf{g})]$, 不引入精英个体的优化算法失败率会非常的高。这是由于在算法的进化初期, 大多数结构的共振频率都是偏离 f_1^{target} 和 f_2^{target} 很远的, 对于这些结构而言计算他们的 F' 所得到的结果是很相近的, 因此很难对其适应度进行区分。而此时种群中若不存在优势个体, 算法的进化方向很可能会出现偏差, 最终导致算法过早地收敛至不理想的局部最优解。因此, 当目标函数设计的不够合理时, 精英个体的存在会对整个进化过程产生更为决定性的作用。

4.3.2 最小化共振峰间隔

在实际应用中, 当噪声集中在某个频带时, 我们总是希望结构的两个共振峰都能处于该频带内, 以便于我们用更少的结构来实现指定频带内的降噪。然而, 由于 THR 的一阶单元和二阶单元之间存在耦合, STL 频谱中的两个共振峰的距离很难靠的很近, 因此通过人工设计是很难让两个共振峰同时处于一个较窄的频带内的。这里我们针对这一问题展开了优化, 目的是使 THR 的两个共振峰都尽可能地靠近 400 Hz。这里的目标函数 F 可以表示为:

$$F = \begin{cases} |f_1 - f^{\text{target}}| + |f_2 - f^{\text{target}}| & \text{if } 101 \leq f_1, f_2 \leq 600 \\ 1000 & \text{otherwise} \end{cases} \quad \dots (4.10)$$

其中, f_1 和 f_2 分别为 THR 的一阶共振频率和二阶共振频率, f^{target} 为目标频率, 这里设置为 400 Hz。

与 4.3.1 章节类似, 我们依然在两种初始条件下进行了对比实验。这里的精

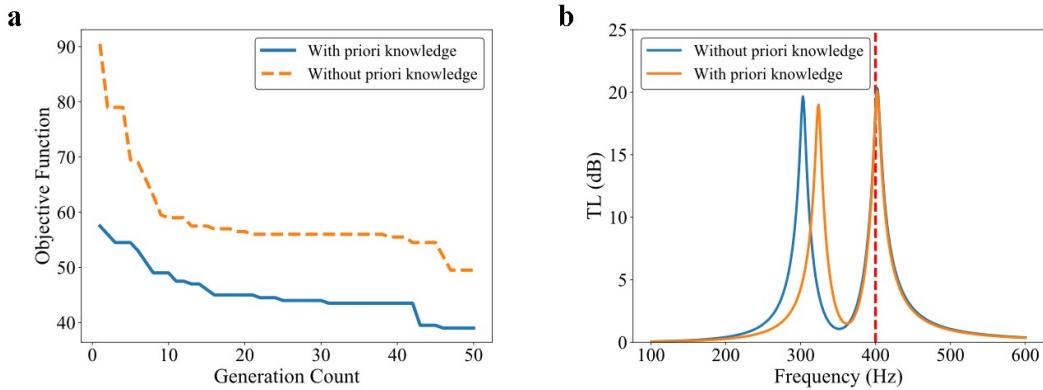


图 4.6 基于两种初始条件最小化两个共振峰到 400 Hz 的距离

Figure 4.6 The optimization effect through the GA to minimize the distance from the resonant frequencies to 400 Hz with and without prior knowledge.

英个体为 5 个共振峰在 400 Hz 附近的样本，均由深度学习模型生成。如图 4.6 所示为基于两种初始条件的优化结果对比，不含精英个体情况下得到的最优个体的共振频率分别为 303 Hz 和 402 Hz，加入精英个体后最优个体的共振频率分别为 324 Hz 和 402 Hz，共振峰到 400 Hz 的距离缩短了 21 Hz，优化效果得到了提升。

4.4 深度学习模型用于其他声学结构的逆设计

在之前的章节中，我们将所提出的深度学习模型用于 THR 结构的逆设计，满足了实际应用中的多种设计需求。值得注意的是，所提出的模型直接预测的是声学结构的等效电学参数，并没有对声学结构本身做任何限制，而 THR 只是该组电学参数的一种物理实现。受限于声学结构和等效电学参数之间的映射关系，模型预测的等效电学参数可能会出现无法被某一种特定的声学结构实现的情况，这时我们可以在 LPT 的框架下向其他的声学结构进行迁移。由表 3.1 可知，给定等效电学参数 R_i 和 M_i 的条件下，若下列方程无解，则我们无法在 THR 的物理框架下找到对应的几何参数。

$$\frac{\rho_0 R_i}{\sqrt{2\eta\rho_0}} a_i^2 - \left(\frac{8\rho_0\beta_i}{3r_i\pi^2} + M_i \right) a_i + \frac{16\rho_0}{3\pi^2} = 0. \quad \dots (4.11)$$

此外，即使该方程有解，我们所最终求解的几何参数也可能是比较极端的，很难通过常规的加工手段进行样品制备。因此，对于这一类等效电学参数，我们可以尝试使用其他声学结构对其进行物理实现。在本章中，我们将以双管亥姆霍兹共

鸣器为例对其进行说明，从而进一步展示所提出的深度学习模型的普适性和可迁移性。

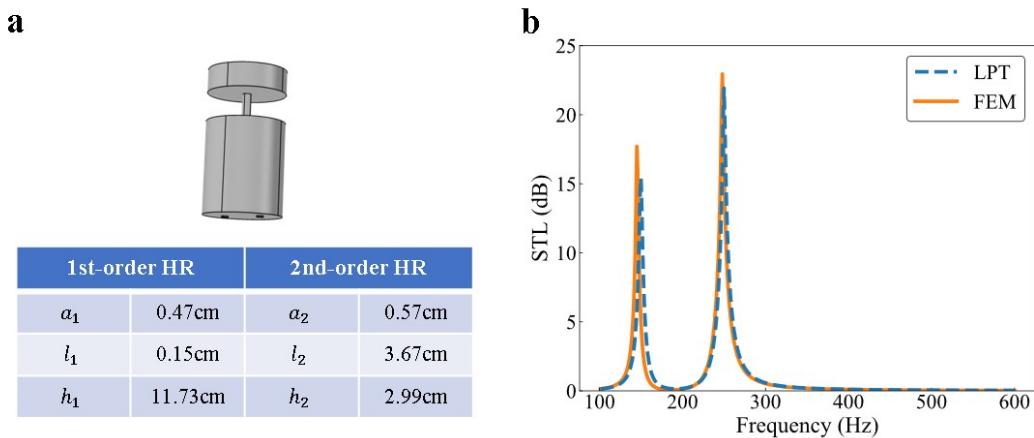


图 4.7 具有两个一阶短管的 THR 单元的逆设计：(a) 结构示意图及几何参数；(b) 结构的 STL 频谱。

Figure 4.7 The structure with double necks in the first order element designed by the DNN model: (a) The schematic view and GPs of the structure; (b) STL spectrum of the structure.

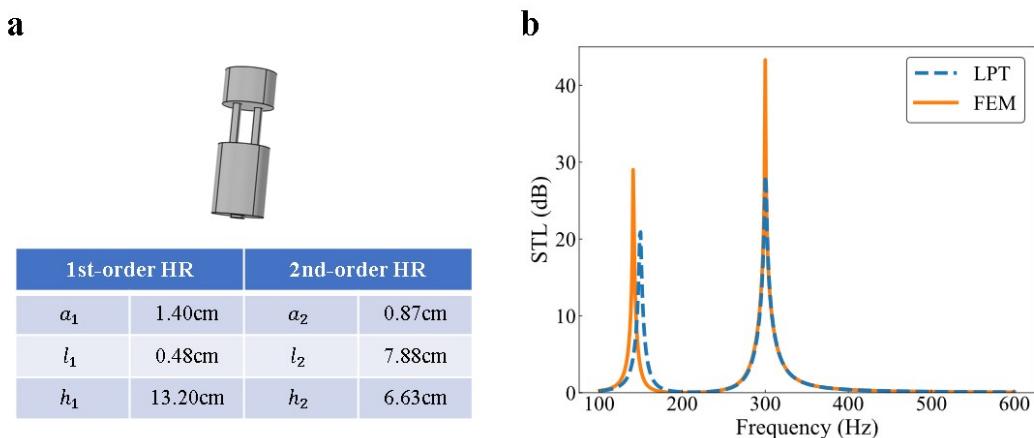


图 4.8 具有两个二阶短管的 THR 单元的逆设计：(a) 结构示意图及几何参数；(b) 结构的 STL 频谱。

Figure 4.8 The structure with double necks in the second order element designed by the DNN model: (a) The schematic view and GPs of the structure; (b) STL spectrum of the structure.

当我们的目标共振频率为 150 Hz 和 250 Hz 时，模型某一次输出的等效电学参数为 $e = [14.97, 79.76, 6.47 \times 10^{-9}, 409.50, 537.41, 1.65 \times 10^{-9}]$ 。当我们

将 $R_1 = 14.97$ 和 $M_1 = 79.76$ 代入后会发现方程 (4.11) 无实根。此时，我们可以令 THR 的一阶单元具有两个相同尺寸的短管，这两个短管之间是并联关系，因此每个短管的声阻 $R_1^* = 2R_1 = 29.94$ ，声质量 $M_1^* = 2M_1 = 159.52$ 。然后，我们将 R_1^* 和 M_1^* 代入方程 (4.11) 求解出一阶短管的 a_1 和 l_1 ，再基于 $C_1 = 6.47 \times 10^{-9}$, $R_2 = 409.50$, $M_2 = 537.41$, $C_2 = 1.65 \times 10^{-9}$ 计算结构的 h_1 , a_2 , l_2 和 h_2 。最终的双管亥姆霍兹共鸣器的结构示意图及对应的几何参数如图 4.7(a) 所示，该结构的 STL 频谱如图 4.7(b) 所示，黄色实线为基于 TMM 的计算结果，蓝色虚线为基于 LPT 的计算结果。由图 4.7 可知，该结构在 150 Hz 和 250 Hz 处均实现了 10 dB 以上的隔声，且对应的几何参数具备可加工性，与预期目标相吻合。

与第一个例子类似，我们也可以令 THR 的二阶单元具有两个相同尺寸的短管。当我们的目标共振频率为 150 Hz 和 300 Hz 时，模型的某一次输出为 $e = [3.67, 51.16, 7.28 \times 10^{-9}, 124.42, 233.17, 3.66 \times 10^{-9}]$ 。当我们令 $R_2 = 124.42$ 和 $M_2 = 233.17$ 代入后会发现方程 (4.11) 无实根。此时，我们可以令 THR 的二阶单元具有两个相同尺寸的短管，每个短管的声阻 $R_2^* = 2R_2 = 248.84$ ，声质量 $M_2^* = 2M_2 = 466.34$ 。然后，我们将 R_2^* 和 M_2^* 代入方程 (4.11) 求解出二阶短管的 a_2 和 l_2 ，再基于 $R_1 = 3.67$, $M_1 = 51.16$, $C_1 = 7.28 \times 10^{-9}$, $C_2 = 3.66 \times 10^{-9}$ 计算结构的 a_1 , l_1 , h_1 和 h_2 。最终的双管亥姆霍兹共鸣器的结构示意图及对应的几何参数如图 4.8(a) 所示，该结构的 STL 频谱如图 4.8(b) 所示，黄色实线为基于 TMM 的计算结果，蓝色虚线为基于 LPT 的计算结果。由图 4.8 可知，该结构在 150 Hz 和 350 Hz 处均实现了 10 dB 以上的隔声，且对应的几何参数具备可加工性，与预期目标吻合。

因此，与直接学习某种声学结构的几何参数和物理响应之间的映射相比，学习等效电学参数和物理响应之间的映射可以使模型更具灵活性和通用性，在实际应用中可以满足更为多样化的需求。

4.5 本章小结

本章在前述研究的基础上，探索了深度学习在声学结构设计领域的拓展性应用。首先，针对声学结构设计过程中普遍存在的多解现象，我们将训练好的深度学习模型与降维类算法结合，实现了声学结构的批量化设计，解决了实际应用

中相同的物理响应对应多种不同结构的设计难题；然后，针对实际应用中的声学结构优化问题，将所提出的模型与智能优化类算法相结合，通过为遗传算法提供较好的初始条件从而提高优化效率，并在一定程度上避免算法过早地陷入局部最优，提升了最终的优化效果；最后，我们将所提出的深度学习模型在 LPT 框架下向其他声学结构进行了推广，进一步说明了所提出的基于深度学习的声学结构设计方法具有很强的灵活性和通用性，并不会局限于某一种特定的声学结构。

第5章 基于声学结构的智能化声学传感系统

5.1 引言

近些年来，随着科学技术的飞速发展，智能设备越来越广泛地应用在我们的日常生活、工作、工业生产、测量采集、医疗康复、救灾救援等诸多领域。与人类类似，智能设备主要借助图像和声音来完成和外界的信息交互。而与图像信息相比，声音信息有着许多明显的优势：首先，声音信息有着较强的传播能力，一台设备可以在很大的范围内采集到所需要的声音信息；其次，声音信号有着较强的穿透能力，可以在更恶劣的环境中传播（如大雾天、昏暗的房间、水中、固体中等）；再次，声音信息有着较强的衍射能力，在传感器和声源信号之间有物体阻隔的情况下，声音信息也可以进行传播。作为智能语音处理模块的重要组成部分，声源定位与识别技术在众多领域有着广泛的需求和应用。现阶段，声源定位与识别主要依靠大型传声器阵列实现[88]。传声器阵列需要使用多个传声器（通常十个以上），利用声音到达各个传声器的时间差进行适当的阵列配置，以增强特定方向的信号，实现声源定位与识别[89]。这种声阵列不仅体积较大，不便于安装和操控，同时多通道采集到的信号的信息量也很大，后端的信号处理过程对计算能力与功耗要求很高。因此，基于多通道的定位与分离系统在无人机、小型机器人等对尺寸和功耗要求较高的场景下会面临应用受限的问题。

人们通过对高级哺乳动物特别是人类的观察发现，它们仅依靠两只耳朵甚至一只耳朵就能完成高精度的声定位和分离任务。在生物学与声学研究人员深入研究后发现，这是由于高级哺乳动物的耳朵并不像常规的传声器一样是直接裸露在空气中的，声波在到达耳膜之前会受到头部、肩膀、外耳和耳道等一系列物理结构的散射作用，由于这些结构是不规则的，因此最终收到的来自空间各个方向的信号也就具有了差异性，高级哺乳动物的大脑可以基于这种差异性完成后续的定位和分离任务的。上述散射过程被建模为头部相关传递函数（Head Related Transfer Functions, HRTF）[90, 91]，这种散射机制本质上可以视为一种物理层的差异性编码，这与声学滤波器的原理十分类似。近年来，研究人员开始尝试将精心设计的声人工结构应用于对声波的空域和频域编码，并在此基础上构建了轻量化的新型声学传感系统[35, 36, 92]。然而，现有的研究大多只关注了二

维平面内的声源的辨识问题，且所使用的重建算法往往需要借助信号的相位信息，因此信号采集过程需要借助额外的参考传声器，算法复杂度较高。

针对上述问题，本章受高级哺乳动物听觉机制的启发，为单通道传声器设计了一个空间高度不对称的半球状外壳，并在压缩感知的框架下提出了一种低复杂度的字典学习算法，构建了基于声人工结构的单通道传感系统，实现了三维空间内多个声源的精准辨识。

5.2 物理模型

由于传统的单通道传声器在空间中是全指向的，因此各个方向的声波对传声器而言是没有差异性的。如果我们希望能够搭建一个单通道定位分离系统，就势必需要将这种与来波方向有关的差异性引入我们的前端接收装置。在本章中，我们受声学滤波器的滤波机制的启发，为单通道传声器设计了一款具有高度空间各向异性的半球状外壳，用于对来自不同方向的声波进行物理层的差异性编码。

5.2.1 声学滤波器

如图 5.1(a) 所示为一个二阶声学滤波器的结构示意图，该结构由两个腔体和三层穿孔板构成，其频响 $T(\omega)$ 可以被建模为：

$$T(\omega) = M_1(\omega)N_1(\omega)M_2(\omega)N_2(\omega)M_3(\omega). \quad \dots (5.1)$$

其中， $M_1(\omega)$ 、 $M_2(\omega)$ 和 $M_3(\omega)$ 分别为三层穿孔板的频率响应； $N_1(\omega)$ 和 $N_2(\omega)$ 分别为两个腔体的频率响应。因此，腔体体积，穿孔板孔隙率及孔的位置分布均会影响声学滤波器模块的频响 $T(\omega)$ 。

首先，我们探究了腔体体积对结构频响的影响。这里我们固定了腔体的边长 $a = 0.1$ m，通过调节腔体的深度 d_1 和 d_2 来改变腔体的体积。同时，我们也固定了每一层穿孔板的孔隙率及孔分布，图 5.1(b) 中展示了上层、中层和底层穿孔板的结构示意图。我们分别基于 COMSOL Multiphysics 计算了三个具有不同腔体体积的二阶声学滤波器的频响曲线，如图 5.1(c) 所示。通过观察这三条曲线可知，改变腔体的体积会同时影响二阶声学滤波器的低频响应和高频响应。

然后，我们探究了孔隙率对结构频响的影响。这里我们固定了腔体的边长

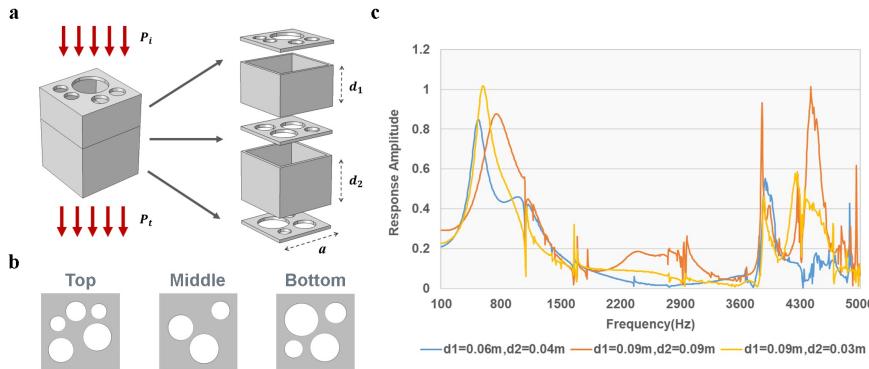


图 5.1 二阶声学滤波器结构示意图及腔体体积对结构频响的影响

Figure 5.1 Schematic view of the two-order acoustic filter and frequency responses of the two-order acoustic filter with different volume cavities.

$a=0.1\text{m}$, 一阶腔体深度 $d_1=0.05\text{ m}$, 二阶腔体深度 $d_2=0.08\text{ m}$ 。同时, 我们也固定了每一层穿孔板的位置分布, 只改变孔的半径以调节穿孔板的孔隙率。最终, 我们构造了三个具有不同孔隙率的二阶声学滤波器, 对应的每一层的穿孔板结构示意图如图 5.3(a) 所示。我们分别基于 COMSOL Multiphysics 计算了上述三个二阶声学滤波器的频响曲线, 如图 5.3(b) 所示。通过观察这三条曲线可知, 改变穿孔板的孔隙率也会同时影响二阶声学滤波器的低频响应和高频响应。

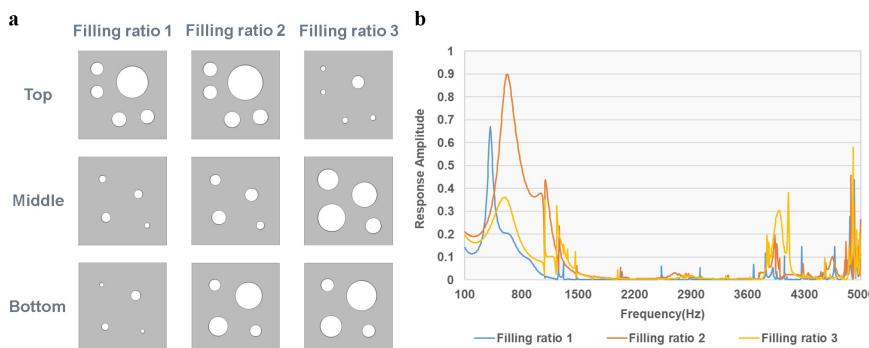


图 5.2 穿孔板孔隙率对二阶声学滤波器频响的影响

Figure 5.2 Frequency responses of the two-order acoustic filter with different filling ratios.

最后, 我们探究了孔的位置分布对结构频响的影响。这里我们固定了腔体的边长 $a = 0.1\text{ m}$, 一阶腔体深度 $d_1 = 0.05\text{ m}$, 二阶腔体深度 $d_2=0.08\text{ m}$, 顶层穿孔板的孔隙率 $w_1=0.32$, 中层穿孔板的孔隙率 $w_2=0.25$, 底层穿孔板的孔隙率 $w_3=0.4$, 只改变每一层穿孔板孔的位置分布。最终, 我们构造了三个具有不同孔的位置分布的二阶声学滤波器, 对应的每一层的穿孔板的结构示意图如图 5.3(a)

所示。我们分别基于 COMSOL Multiphysics 计算了上述三个二阶声学滤波器的频响曲线，如图 5.3(b) 所示。通过观察这三条曲线可知，改变穿孔板孔的位置分布只会影响二阶声学滤波器的高频响应。

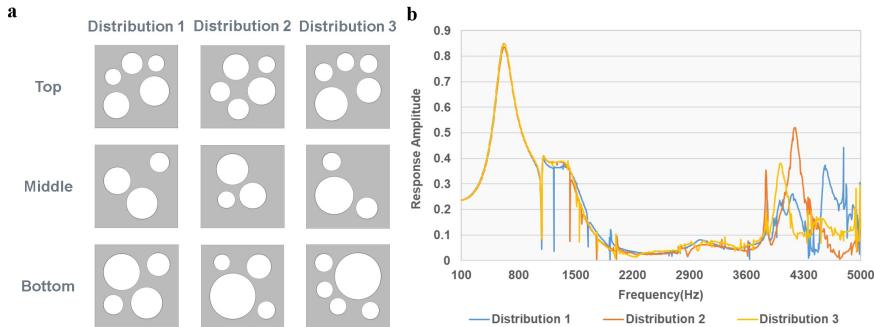


图 5.3 穿孔板孔的位置分布对二阶声学滤波器频响的影响

Figure 5.3 Frequency responses of the two-order acoustic filter with different distribution of holes.

因此，通过改变腔体体积，穿孔板孔隙率及孔的位置分布就可以构造出在宽频范围内具有很强差异性的二阶声学滤波器。

5.2.2 半球状外壳

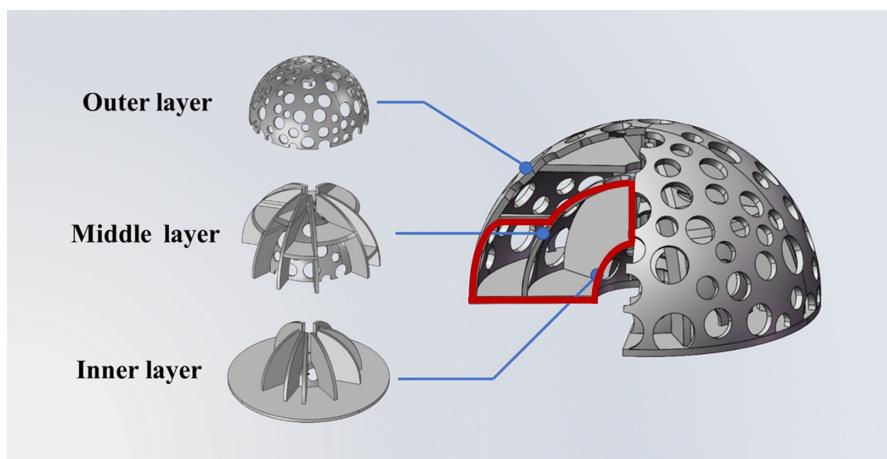


图 5.4 半球状外壳的结构示意图

Figure 5.4 Schematic view of the hemispherical shell.

如图 5.4 所示为我们所设计的半球状外壳，该外壳可以视为是由多个不同尺寸、不同朝向的二阶声学滤波器组合而成的，红线所指示的区域就是其中一个二阶声学滤波器单元。该外壳由三层同心的半球壳嵌套而成，由外到内的半径分别为 0.24m, 0.168m 和 0.072m；每层半球壳的表面均设置了多个半径不等的圆

孔，圆孔之间无重叠；球壳与球壳之间放置了若干块横向和纵向的隔板，将这部分区域划分成了许多个腔体。球心位置放置了一个单通道传声器，用于接收被球壳调制过的声音信号，底部铺有吸声棉用于削弱反射声。

在构建外壳的过程中，每一层球壳上都随机布放大小不一的圆孔，横向和纵向的隔板也是随机地插入到球壳之间的，因此朝向不同方向的声学滤波器模块具有不同的腔体体积，穿孔板孔隙率和孔的位置分布。由于这些声学滤波器结构的差异性，使得整个半球状外壳对不同方向的频响也具有了差异性，从而打破了单通道传声器的全指向性。

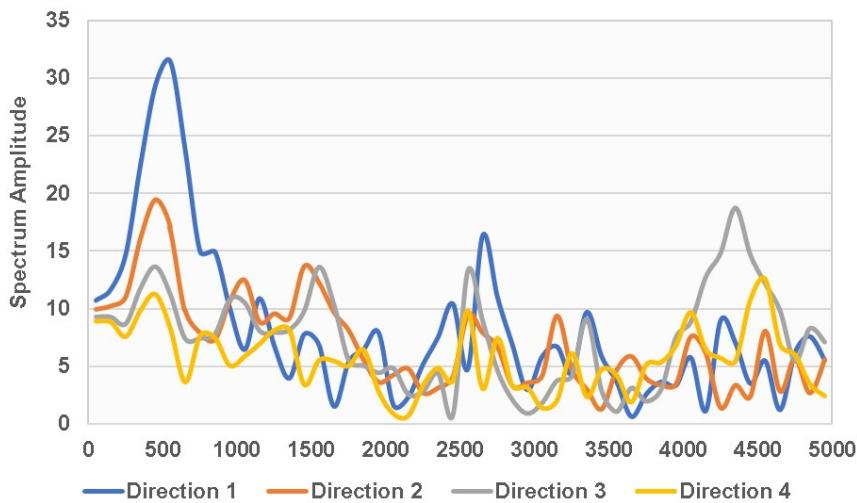


图 5.5 半球状外壳不同方向的频响

Figure 5.5 Frequency responses of the hemispherical shell in four different directions.

然后，我们基于 COMSOL Multiphysics 对图 5.4 所示的半球状外壳的频响进行了仿真计算。我们从不同方向发射了 1 Hz 5 kHz 的扫频信号，步长为 1Hz，并记录了球心处的声压幅值用于计算频响。图 5.5 所示为其中某 4 个方向的结构频响。仿真结果表明，半球状外壳结构在不同方向的频响均具有明显的差异性，这种差异性在中高频段尤为显著；而受限于结构本身的尺寸，这种与方向有关的差异性会在低频被削弱。由于生活中绝大多数的声信号均为宽频信号，因此对于这类信号而言结构的低频性能并不是必须的，我们可以基于被结构调制后的信号的高频差异性来完成后续的定位分离工作。而对于能量只集中于低频的特殊声信号，我们可以考虑加大结构的尺寸并采用更多层的球壳嵌套结构，以使其能在低频有效地工作。

为了定量地描述半球状外壳对不同方向频响的差异性，我们定义第 i 个方向

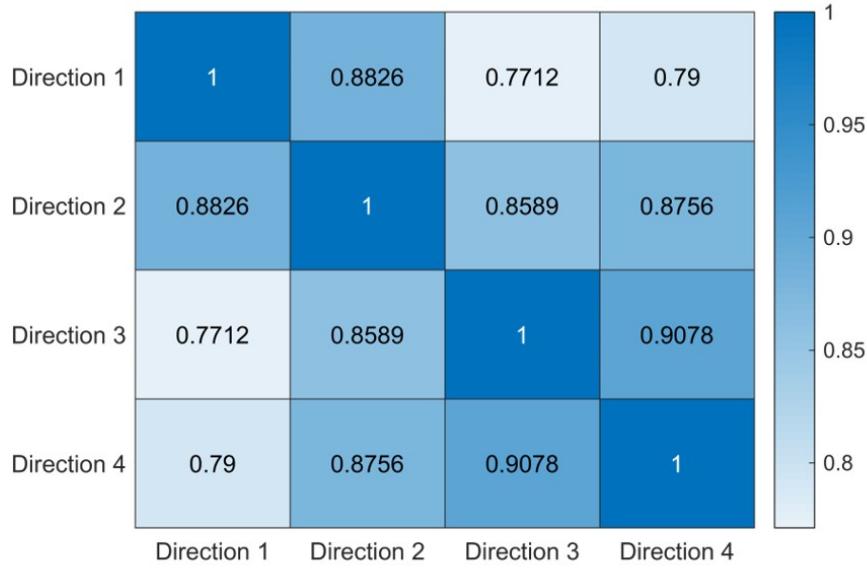


图 5.6 半球状外壳不同方向频响的相关度

Figure 5.6 Frequency response coherences of the hemispherical shell between 4 different directions.

和第 j 个方向频响之间的相关度 μ_{ij} 为：

$$\mu_{ij} = \langle h_i(\omega), h_j(\omega) \rangle / |h_i(\omega)| |h_j(\omega)|. \quad \dots (5.2)$$

其中， $h_i(\omega)$ 和 $h_j(\omega)$ 分别为第 i 个方向和第 j 个方向的频响， $\langle \cdot \rangle$ 代表内积算子， $|\cdot|$ 代表对向量取模。 μ_{ij} 可以用于衡量 $h_i(\omega)$ 和 $h_j(\omega)$ 之间的相似度， μ_{ij} 越接近于 1，则代表这两个方向的频响的差异性越小；反之， μ_{ij} 越接近于 0，则代表这两个方向的频响差异性越大。因此，对于一个裸露在空气中的单通道传声器，不同方向的频响之间 μ_{ij} 会非常接近于 1，而引入半球状外壳后不同方向频响的相干性会被降低。如图 5.6 所示为上述四个方向频响的相关度，可以看到，由于半球状外壳的调制作用，不同方向频响的相关度均有所降低。然而，对于某些相距较近的方向，如方向 3 和方向 4，他们之间的频响相关度依然较高。因此，在 5.3 章节中，我们将引入变稀疏度主成分分析（Variable Sparsity Principal Component Analysis, VSPCA）通过信号处理地方式进一步降低不同方向频响的相关度，起到降低冗余放大差异的效果。

半球状外壳的引入使得系统在信号采集阶段通过物理层编码的方式引入了与方向有关的差异性。声波从声源发出后在自由空间内传播，然后进入半球状外

壳被其调制，最终被球心处的传声器所接收。当空间中同时有 K 个声源发声时，信号的采集过程的数学模型可以表示为：

$$y(\omega) = \sum_{i=1}^K s_i(\omega) \circ h_i(\omega). \quad \dots (5.3)$$

其中， $s_i(\omega)$ 为第 i 个方向的声信号频谱； $h_i(\omega)$ 为第 i 个方向半球状外壳的频响； \circ 为 Hadamard 乘积算子； $y(\omega)$ 为传声器接收的混合信号的频谱，在低噪声的环境下可以近似地看做 K 个受到外壳调制后的信号的叠加。

5.3 压缩感知算法

在上一章节中，我们为单通道传声器设计了一个空间高度各向异性的半球状外壳，从而使得不同方向的声波被传声器接收时具有了与来波方向有关的差异性。然而，由公式 (5.3) 可知，传声器最终输出的是来自不同方向的声音叠加后的信号，我们还需要通过信号处理的方式基于上述差异性对其进行分离，从而确定每一个信号的方位及内容。在这一章节中，我们将在压缩感知的框架下提出一种变稀疏度主成分分析 (Variable Sparsity Principal Component Analysis, VSPCA) 和正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 的联合算法，用于实现单通道系统的定位与分离。

在压缩感知的框架下，我们的系统可以被建模为：

$$\mathbf{y} = \mathbf{A}\mathbf{s}. \quad \dots (5.4)$$

其中， \mathbf{y} 为观测向量，对应于传声器的接收信号； \mathbf{A} 为字典矩阵，需要基于一定的先验知识提前训练得到； \mathbf{s} 为稀疏表达，与真实空间内声音的方位与内容相关，在实际应用中需要基于观测向量 \mathbf{y} 和字典矩阵 \mathbf{A} 计算得到。

单通道定位与分离系统的工作流程图如图 5.7 所示，左侧为信号收集过程，右侧为信号处理过程。我们首先需要在搭建系统时对字典 \mathbf{A} 进行训练，然后在实际应用中通过传声器实际接收的信号计算出观测向量 \mathbf{y} ，最后基于压缩感知中的重建算法对稀疏表达 \mathbf{s} 进行求解（即稀疏恢复），从而实现空间中多个声信号的定位与分离。

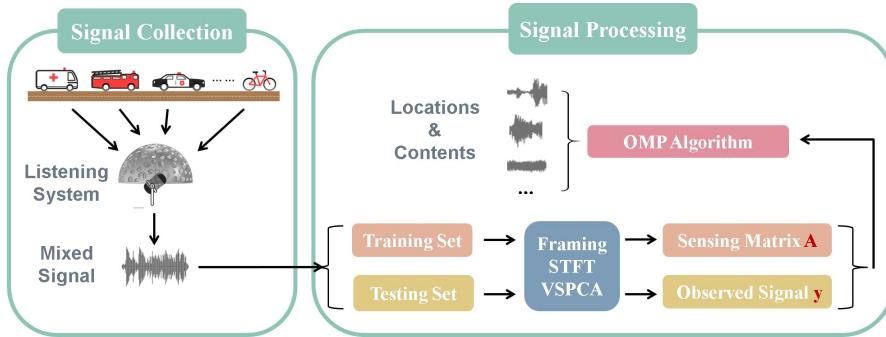


图 5.7 单通道定位与分离系统的算法流程示意图

Figure 5.7 Schematic of data collection and processing of the metamaterial-based single-microphone listening system.

5.3.1 字典学习

公式(5.4)中的字典矩 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_Q]$ 是一个 P 行 Q 列的矩阵，矩阵 \mathbf{A} 中每一列都与传声器所接收到的空间中某个方向发出的某一种声音信号相对应。也就是说，字典矩阵 \mathbf{A} 中同时包含了半球状外壳的频响信息和空间中可能出现的声源信息。因此，在定位与分离系统的搭建阶段，字典 \mathbf{A} 需要基于一定的先验知识进行预训练。

首先，我们需要收集空间中可能出现的声源信息，并将其整合成训练集矩阵 \mathbf{X} ，具体步骤如下：

1. 将三维空间划分为 N 个方向，编号为 $1 \sim N$ ，每个方向都是声音可能出现的方向；
2. 收集空间中可能出现的声音样本，将其整合成一个语料库，语料库中包含 M 种不同的声音信号，编号为 $1 \sim M$ ；
3. 在所有方向放置扬声器，依次播放语料库中的所有声音信号。每一段信号在空间中自由传播至半球状外壳的表面后被其调制，最后被球心的传声器拾取，传声器联合后端的采集模块最终输出训练信号，空间中第 n 个方向的第 m 个音频信号的训练信号记作 \mathbf{x}_{nm} ；
4. 将训练信号 \mathbf{x}_{nm} 分为 Q 帧，并对每一帧进行加窗和短时傅里叶变换，得

到训练信号每一帧的频谱，最终组成该训练信号的训练集矩阵 $\mathbf{X}_{nm} \in \mathbb{R}^{P \times Q}$ ：

$$\mathbf{X}_{nm} = \begin{bmatrix} x_{nm}^1(\omega_1) & x_{nm}^2(\omega_1) & \dots & x_{nm}^Q(\omega_1) \\ x_{nm}^1(\omega_2) & x_{nm}^2(\omega_2) & \dots & x_{nm}^Q(\omega_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_{nm}^1(\omega_P) & x_{nm}^2(\omega_P) & \dots & x_{nm}^Q(\omega_P) \end{bmatrix}. \quad (5.5)$$

其中， $x_{nm}^q(\omega_p)$ 为训练信号 \mathbf{x}_{nm} 第 q 帧的频谱在 ω_p 频率处的幅值 ($p = 1, 2, \dots, P; q = 1, 2, \dots, Q$)， P 为短时傅里叶变换的总频点数， Q 为训练信号分帧的总帧数；

5. 将所有训练信号的训练集矩阵进行合并，得到全部训练信号的训练集矩阵 $\mathbf{X} \in \mathbb{R}^{P \times QMN}$ ：

$$\mathbf{X} = \left[\mathbf{X}_{11} \ \dots \ \mathbf{X}_{1M} \ \mathbf{X}_{21} \ \dots \ \mathbf{X}_{2M} \ \dots \ \mathbf{X}_{N1} \ \dots \ \mathbf{X}_{NM} \right]. \quad (5.6)$$

然而，正如 5.2.2 小节所讨论的，虽然我们通过引入半球状外壳实现了对来自不同方向声波的差异性编码，但是对于一些相邻的方向频响依然会存在较高的相似度。因此，如果我们直接基于矩阵 \mathbf{X} 来构造字典 \mathbf{A} ，就会导致字典中条目间相关性较高。在这种情况下，为了确保稀疏恢复的精度，我们只能使用计算复杂度较高的重建算法对稀疏表达 \mathbf{s} 进行求解，这会带来较高的计算成本，导致定位与分离系统的实时性降低。另一种解决思路是通过降低字典中条目间的相关性来降低稀疏恢复的难度，从而使得我们可以基于复杂度较低的重建算法完成对稀疏表达 \mathbf{s} 的求解。因此，我们基于传统的 PCA 算法对训练集进行了变换用于构造字典 \mathbf{A} ，并提出了 VSPCA 算法对观测向量 \mathbf{y} 进行变换（见 5.3.2 章节），通过降低字典中条目间的相关度降低了稀疏恢复的难度。字典 \mathbf{A} 的计算过程如下：

1. 计算矩阵 \mathbf{X} 每一行的均值 $\boldsymbol{\mu}$ 和标准差 $\boldsymbol{\sigma}$ ；

$$\begin{aligned} \boldsymbol{\mu} &= [\mu_1, \mu_2, \dots, \mu_P]^T, \\ \mu_i &= \frac{1}{L} \sum_{j=1}^L x_{ij}, \\ \boldsymbol{\sigma} &= [\sigma_1, \sigma_2, \dots, \sigma_P]^T, \\ \sigma_i &= \sqrt{\frac{1}{L} \sum_{j=1}^L (x_{ij} - \mu_i)^2} \end{aligned} \quad (5.7)$$

其中， x_{ij} 为矩阵 \mathbf{X} 第 i 行第 j 列的元素； $L = Q \times M \times N$ 为训练集样本总数，即矩阵 \mathbf{X} 的列数； μ_i 为均值向量 $\boldsymbol{\mu}$ 的第 i 个分量； σ_i 为标准差向量 $\boldsymbol{\sigma}$ 的第 i 个分量；

2. 对矩阵 \mathbf{X} 进行标准化，得到标准化的训练集矩阵 $\bar{\mathbf{X}} \in \mathbb{R}^{P \times QMN}$ ，该矩阵的第 i 行 j 列的元素 \bar{x}_{ij} 为：

$$\bar{x}_{ij} = \frac{x_{ij} - \mu}{\sigma_i}. \quad (5.8)$$

3. 使用 PCA 算法计算训练集 $\bar{\mathbf{X}}$ 的投影矩阵 \mathbf{B} ，并保留前 \tilde{P} 个主成分，因此 $\mathbf{B} \in \mathbb{R}^{\tilde{P} \times P}$, $\tilde{P} < P$ ；

4. 使用投影矩阵 \mathbf{B} 对标准化后的训练集矩阵 $\bar{\mathbf{X}}$ 进行投影，得到字典 $\mathbf{A} \in \mathbb{R}^{\tilde{P} \times QMN}$ ：

$$\mathbf{A} = \mathbf{B}\bar{\mathbf{X}}. \quad (5.9)$$

为了证明 PCA 算法在降低字典条目间相关度任务上的有效性，我们使用该算法对图 5.5 中展示的 4 个方向的频响做了变换，变换后频响间的相关度如图 5.8 所示。与图 5.6 中的结果相比，变换后不同方向间的频响相关性被显著降低了。根据公式 (5.3)，传声器最终所接收的信号频谱为信号的原始频谱与半球状外壳在来波方向的频响的 Hadamard 乘积。因此，PCA 变换可以通过降低不同方向频响的相关度来显著降低字典条目间相关度。

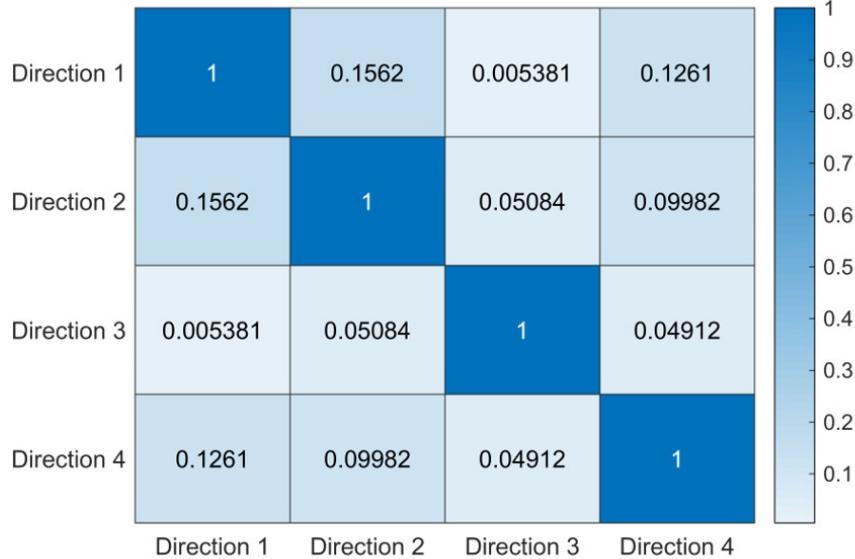


图 5.8 PCA 变换后不同方向频响的相关度

Figure 5.8 Frequency response coherences of the hemispherical shell between 4 different directions after PCA.

5.3.2 变稀疏度主成分分析

传统的PCA算法可以将原始数据以方差最大化为原则投影至一个新的空间，因此变换后样本间的冗余将会被显著降低，差异性将被增强。然而，传统的PCA算法通常作用于稀疏度为1的样本，例如，在使用PCA算法提取人脸的主成分信息时，同一图像区域往往只存在一张人脸，不会出现几张脸叠加在一起的情况。然而，对于我们所提出的定位与分离系统而言，我们的训练样本为独立的某一条语音信号（此时样本的稀疏度为1），测试样本则可能是 K 个不同的语音信号的叠加（此时样本的稀疏度为 K ）。因此，我们需要对传统的PCA算法进行改进，通过将稀疏度的概念引入到算法中，使得算法可以作用于任意稀疏度的样本。因此，在本小节中，我们在传统的PCA算法的基础上提出了VSPCA算法用于对稀疏度不为1的测试信号进行投影变换。变换后的测试信号可以表示为字典 A 中 K 个条目的叠加，可以通过压缩感知框架下的重建算法对其稀疏系数进行求解。

当空间中同时有 K 个声源同时发声时，测试信号的稀疏度为 K 。对于稀疏度为 K 的信号，我们需要使用VSPCA算法对其进行变换，具体步骤如下：

1. 对采集到的测试信号进行分帧、加窗和短时傅里叶变换处理，得到信号的每一帧的频谱 $\mathbf{y}^{\tilde{q}} = [y^{\tilde{q}}(\omega_1), y^{\tilde{q}}(\omega_2), \dots, y^{\tilde{q}}(\omega_P)]^T$ ($\tilde{q} = 1, 2, \dots, \tilde{Q}$)，将上述频谱向量作为矩阵的列向量构建信号的测试集矩阵 $\mathbf{Y} \in \mathbb{R}^{P \times \tilde{Q}}$ ：

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}^1 & \mathbf{y}^2 & \dots & \mathbf{y}^{\tilde{Q}} \end{bmatrix}. \quad (5.10)$$

其中， $y^{\tilde{q}}(\omega_p)$ 为第 \tilde{q} 帧测试信号的频谱在 ω_p 频率处的幅值； P 为短时傅里叶变换的频点数， \tilde{Q} 为测试信号分帧的总帧数。

2. 基于训练样本的标准差 σ 和均值 μ 对测试集矩阵 \mathbf{Y} 标准化，得到标准化后的测试集矩阵 $\bar{\mathbf{Y}} \in \mathbb{R}^{P \times \tilde{Q}}$ ，标准化的过程与训练信号标准化类似。标准化后的测试集矩阵 $\bar{\mathbf{Y}}$ 第 s 行 t 列的元素 \bar{y}_{st} 为：

$$\bar{y}_{st} = \frac{y_{st} - K\mu_s}{\sigma_s}. \quad (5.11)$$

其中， y_{st} 为矩阵 \mathbf{Y} 第 s 行 t 列的元素； μ_s 为均值 μ 的第 s 个分量， σ_s 为标准差 σ 的第 S 个分量， μ 和 σ 均由公式(5.7)计算得到；

3. 使用基于训练样本计算得到投影矩阵 \mathbf{B} 对标准化后的矩阵 $\bar{\mathbf{Y}}$ 进行投影变

换，得到变换后的测试集矩阵 $\mathbf{G} \in \mathbb{R}^{\tilde{P} \times \tilde{Q}}$ ：

$$\mathbf{G} = \mathbf{B}\bar{\mathbf{Y}} = [\mathbf{g}^1, \mathbf{g}^2, \dots, \mathbf{g}^{\tilde{Q}}]. \quad (5.12)$$

其中， $\mathbf{g}^{\tilde{q}}$ 为矩阵 \mathbf{G} 的第 \tilde{q} 列，对应于测试信号的第 \tilde{q} 帧 ($\tilde{q} = 1, 2, \dots, \tilde{Q}$)； \tilde{P} 为矩阵 \mathbf{G} 的行数，对应于 PCA 变换保留的主成分个数；

5.3.3 稀疏恢复

在实际应用中，我们希望所提出的系统能够对空间中的声音进行监测，并实时地输出声音的方位与内容信息。当空间中同时有 K 个声源同时发声时，传声器所接收到的是一路单通道的混合信号，首先我们会使用 5.3.2 章节所述的 VSPCA 算法对该信号进行变换，得到变换后的测试集矩阵 \mathbf{G} ；然后，我们将基于预训练的字典矩阵 \mathbf{A} 和测试集矩阵 \mathbf{G} 对测试信号的稀疏表达进行重建；最后，基于重建结果得到 K 个声源的方位信息和音频内容信息。

由公式 (5.4)、(5.9) 和 (5.12) 可知，对于测试集矩阵 \mathbf{G} 的第 \tilde{q} 列 $\mathbf{g}^{\tilde{q}}$ ，应满足：

$$\mathbf{g}^{\tilde{q}} = \mathbf{A}\mathbf{s}^{\tilde{q}}. \quad (5.13)$$

其中，向量 $\mathbf{s}^{\tilde{q}}$ 是一个稀疏向量，为该测试信号的稀疏表达。因此，求解向量 $\mathbf{s}^{\tilde{q}}$ 的问题可以在压缩感知的框架下被转化为最小化向量 $\mathbf{s}^{\tilde{q}}$ 的 0 范数的优化问题，即 L_0 问题：

$$\begin{aligned} & \min_{\mathbf{s}^{\tilde{q}}} \|\mathbf{s}^{\tilde{q}}\|_0, \\ & \text{s.t. } \|\mathbf{A}\mathbf{s}^{\tilde{q}} - \mathbf{g}^{\tilde{q}}\|_2 \leq \varepsilon \end{aligned} \quad (5.14)$$

有很多经典的稀疏恢复算法可以用于对上述优化问题进行求解，这里我们采用的是正交匹配追踪算法 (Orthogonal Matching Pursuit, OMP) [93]。与用于求解 L_1 问题的稀疏恢复算法相比，OMP 算法更加简洁高效，可以使定位与分离系统具有更好的实时性 [94]。

基于 OMP 算法，我们可以得到重建后的向量 $\mathbf{s}^{\tilde{q}}$ ，该向量是长度为 $Q * N * M$ 的稀疏向量，理想情况下该向量中只有 K 个元素的值接近 1，其余元素均为接近 0 的小量。如图 5.9 所示，向量 $\mathbf{s}^{\tilde{q}}$ 和字典 \mathbf{A} 每一列均具有对应关系。若向量 $\mathbf{s}^{\tilde{q}}$ 中某一元素接近 1，则表示字典 \mathbf{A} 中与该元素对应的条目存在在向量 $\mathbf{g}^{\tilde{q}}$ 中，而该条目的下标 n 和 m 即对应该信号的方位索引和音频索引。通过这种方式，我们便可以基于稀疏向量 $\mathbf{s}^{\tilde{q}}$ 计算出 K 组 $[n, m]$ ，从而推算出空间中 K 个声源的方

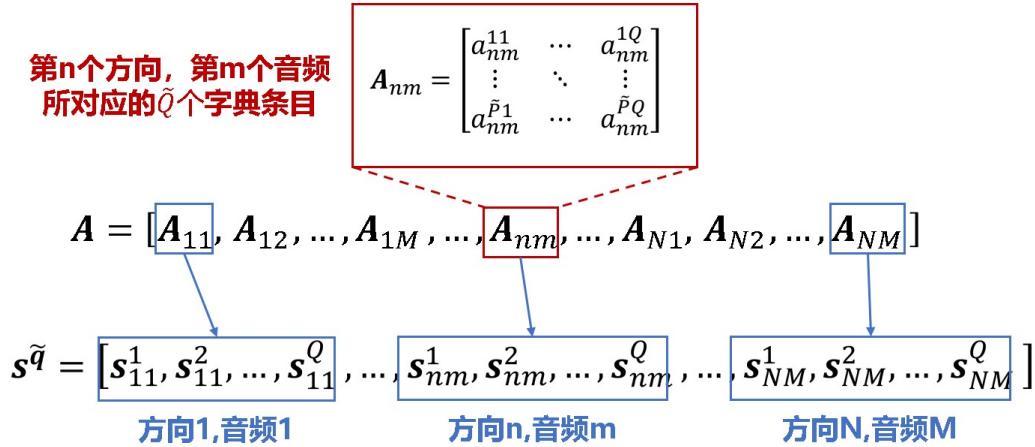


图 5.9 字典矩阵 A 和稀疏向量 s 的对应关系

Figure 5.9 Corresponding relationship between A and s .

位和音频内容。考虑到实际应用中会存在噪声等干扰因素，因此某一帧的计算结果可能会出现误差，并不可靠。因此，我们可以对测试集矩阵 G 的每一列都重复上述操作得到 \tilde{Q} 个重建结果，然后对这个结果进行统计分析，最终确定 K 个声源的方位和音频内容。上述重建过程的具体步骤如下：

1. 基于测试集矩阵 G 的第 \tilde{q} 列和字典 A ，通过 OMP 算法求解对应的稀疏向量 $s^{\tilde{q}}$ ，并统计该向量前 K 大的元素所对应的位置索引 $d_1^{\tilde{q}}, d_2^{\tilde{q}}, \dots, d_K^{\tilde{q}}$ ，计算对应的方向编号 $n_k^{\tilde{q}}$ 和音频内容标号 $m_k^{\tilde{q}}$ ：

$$n_k^{\tilde{q}} = \left\lceil \frac{d_k^{\tilde{q}} / (M \times Q)}{\right\rceil$$

$$m_k^{\tilde{q}} = \left\lceil \frac{d_k^{\tilde{q}} \bmod (M \times Q)}{Q} \right\rceil. \quad (5.15)$$

$$k = 1, 2, \dots, K$$

其中， $\lceil \cdot \rceil$ 为向上取整， \bmod 为取余数，最终得到 K 组结果 $[n_k^{\tilde{q}}, m_k^{\tilde{q}}] (k = 1, 2, \dots, K)$ 。

2. 针对测试集矩阵 G 的每一列重复步骤 1 中的操作，最终得到 $\tilde{Q} * K$ 组结果 $[n_k^{\tilde{q}}, m_k^{\tilde{q}}] (k = 1, 2, \dots, K; \tilde{q} = 1, 2, \dots, \tilde{Q})$ ，对上述结果进行统计分析，保留出现频率最高的 K 组结果；
3. 根据步骤 3 中得到的 K 组结果，结合 5.3.1 章节中空间方位的划分方式和所用的语料库，即可获取 K 个声源的具体方位和音频内容。

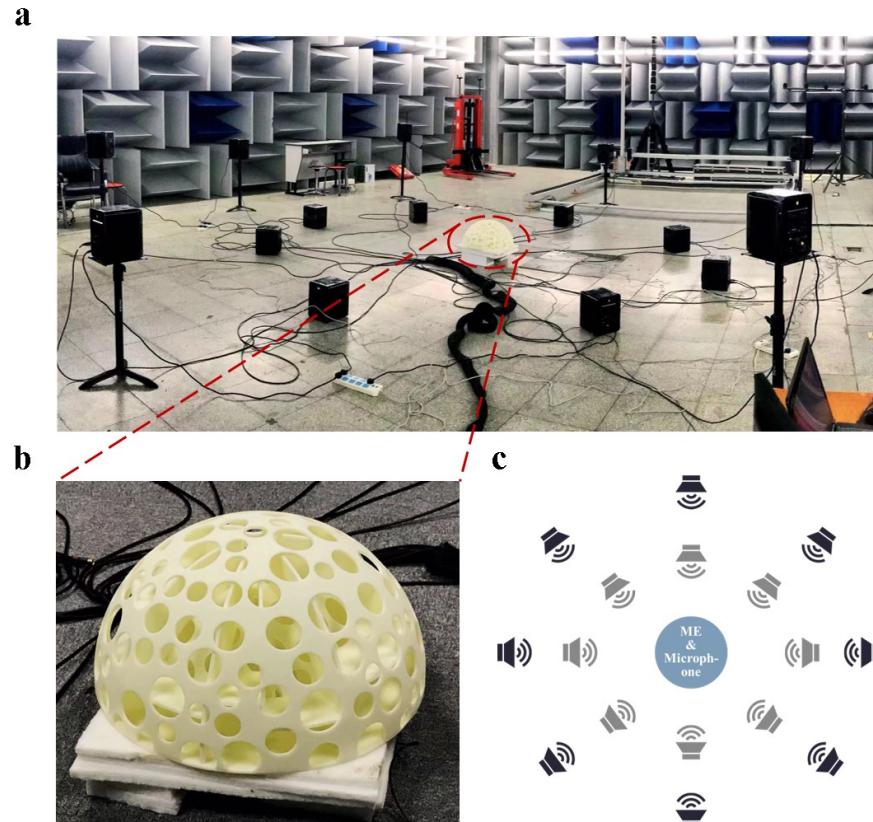


图 5.10 实验测量：(a) 实验装置照片；(b) 半球状外壳样品照片；(c) 声源位置示意图。

Figure 5.10 Experiment measurement : (a) Photo of the experimental setup in the semi-anechoic room. (b) Enlarged photo of the hemispherical shell. (c) Schematic of the distribution of sound sources.

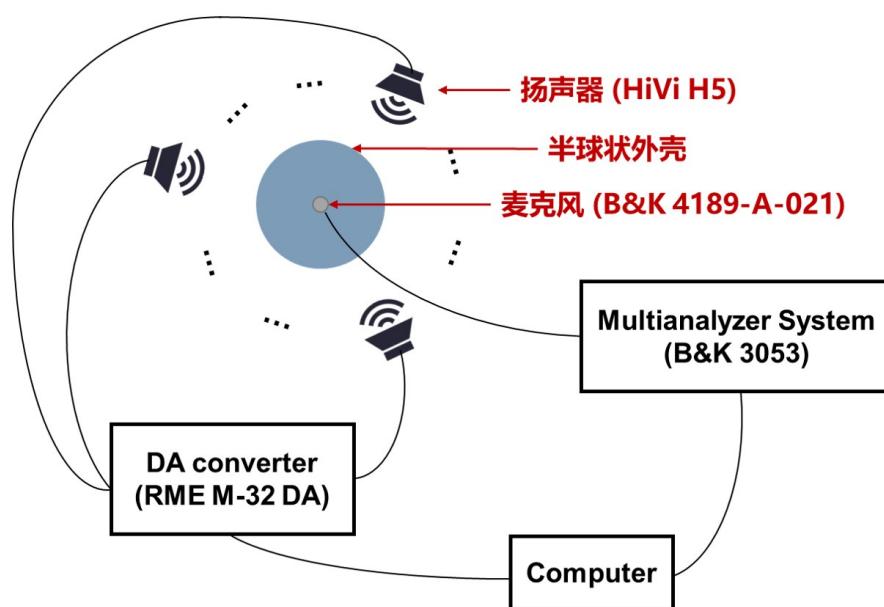


图 5.11 测量装置示意图

Figure 5.11 Schematic of the system layout.

5.4 实验测试

5.4.1 实验装置

为了进一步对所提出的定位与分离系统的性能进行验证，我们在半消声室中对其进行了听音测试，如图 5.10(a) 所示。定位与分离系统的信号采集端由半球状外壳和单通道传声器组成，半球状外壳放置在房间中心的地面上，底部铺有吸声棉，该样品的照片如图 5.10(b) 所示。该外壳由三层同心穿孔半球壳嵌套而成，球壳的半径由外到内分别为 24 cm, 16.8 cm 和 7.2 cm，球壳的厚度由外到内分别为 1 cm, 0.7 cm 和 0.3 cm。球壳表面随机开孔，孔的半径范围为 0.3 cm ~ 3 cm，相邻孔之间的距离大于 0.5 cm，穿孔球壳的孔隙率较高有利于减少声波的耗散，提高接收信号的信噪比。半球状外壳基于 3D 打印技术制备，打印精度为 0.06 mm。所使用的基材为光敏树脂 (Lasty-KS)，密度为 1.13 g/cm³。该树脂的阻抗远大于空气阻抗，因此可以视为刚性边界。

如图 5.11 所示为实验的测量装置示意图，半球状外壳周围布放了 16 个扬声器 (HiVi H5) 用于模拟声源发声，其位置的俯视示意图如图 5.10(c) 所示。这些扬声器被放置成两个同心圆环，内圈半径为 2.5 m，外圈半径为 4.5 m，圆心即为系统的信号采集装置。为了评估系统在俯仰维度上的定位分离能力，内外圈扬声器距离地面的高度并不相同。内圈扬声器直接放置在地面上，而外圈扬声器则被放置在了扬声器支架上，支架的高度从 90 cm 到 100 cm 不等。这 16 个扬声器通过 DA 转换器 (RME M-32 DA) 与计算机连接，可以独立地发射给定的声信号。半球状外壳的球心处放置了一个 1/2 英寸的全指向传声器 (B&K 4189-A-021) 用于接收测试信号，传声器接收的信号可以通过 Pulse (B&K 3053) 录制后进行离线处理，也可以通过声卡 (Audient-iD14) 传输到计算机中进行实时处理。

5.4.2 多声源辨识

为了测试系统的定位与分离性能，我们通过计算机随机选择 k 个扬声器发声，发声内容也是从语料库中随机挑选的。这 k 个声音在传播过程中相互混合，最终被传声器所接收。经过如图 5.7 所示的信号处理后，系统最终输出每个声源的方位和音频内容信息。为了定量地描述系统定位与分离性能，我们定义重建准确率 $\alpha = n/k$ ，其中 n 为方位和音频内容均被成功重建的声源个数， k 为空间中同时发声的声源个数，也就是重建算法的稀疏度。我们针对 5 个常见的应用场景

表 5.1 不同场景下听音测试结果**Table 5.1 The results of listening tests in different scenarios.**

	街道	动物园	音乐厅	家庭	对话
k=1	100.00%	100.00%	100.00%	100.00%	100.00%
k=2	99.50%	97.00%	96.50%	96.50%	97.50%
k=3	93.67%	91.67%	93.33%	91.00%	91.67%
k=4	81.00%	86.75%	77.75%	81.00%	81.00%
k=5	68%	77.80%	68.60%	77.60%	69.40%

进行了听音测试，不同场景下都选择了 6 种常见的声音作为测试的语料，每种场景下考察了 k 从 1 至 5 的情况并分别进行了 100 次随机实验，最后对重建结果进行了统计分析，测试结果如表 5.1 所示。具体细节如下：

1. 街道：

在街道场景下，我们选择的 6 种声音信号分别为救护车的鸣笛声，警车的鸣笛声，消防车的鸣笛声，轿车的鸣笛声，自行车的铃声和倒车的提示音。这些声音信号均为宽带信号，他们的语谱图和功率谱如图 5.12 所示。针对 k 从 1 至 5 分别进行了 100 次随机实验，实验结果如图 5.13(a) 所示：横坐标为稀疏度 k ，代表空间中同时发声的声源个数；左侧的纵坐标为柱状图的纵坐标，代表实验发生的次数，柱状图的颜色代表实验的成功率；右侧的纵坐标为折线图的纵坐标，代表该稀疏度下 100 次听音测试的平均成功率。例如：当 $k=2$ 时，空间中有两个声源同时发声；在 100 次实验中有 99 次实验两个声源的位置信息和音频内容信息均被成功重建，因此这 99 次实验的成功率 α 为 100%，柱状图为蓝色；而有一次实验只成功重建出一个声源的位置信息和音频内容信息，因此这次实验的成功率为 50%，柱状图为绿色；因此，这 100 次实验的平均成功率为 $(1*99+0.5*1)/100*100\% = 99.5\%$ 。测试所用的 6 个语料在不同稀疏度的条件下的重建成功率如图 5.13(b) 所示，横坐标为稀疏度 k ，代表空间中同时发声的声源个数；纵坐标为重建成功率，等于该语料被成功重建的次数与该语料出现总次数的比值；不同颜色的柱状图对应不同的语料。通过观察图 5.13(a) 和 (b) 可知，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统在街道场景下的重建成功率逐渐下降，分别为 100%，99.5%，93.67%，81% 和 68%。当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80%

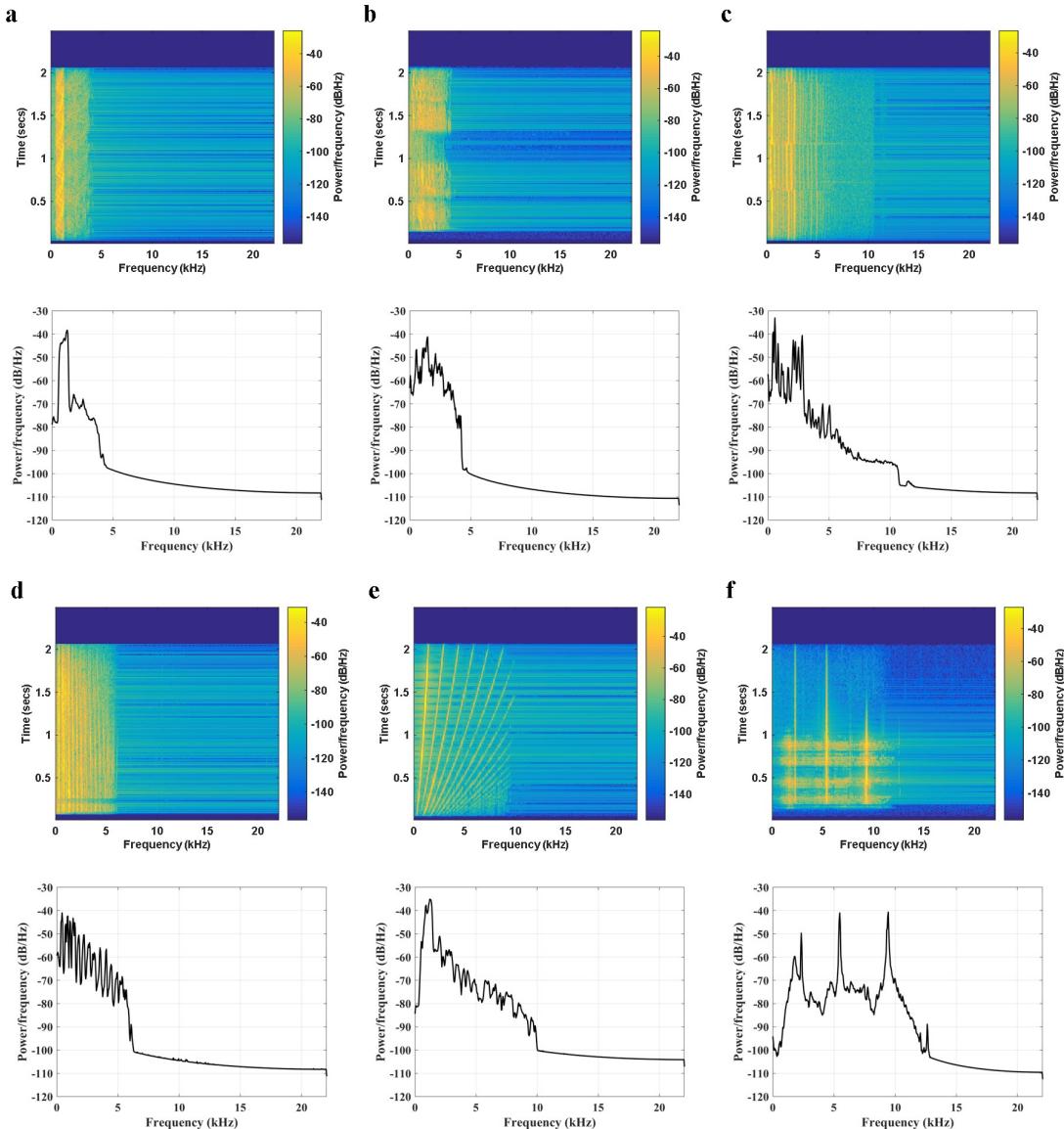


图 5.12 街道场景下语料的语谱图和功率谱: (a) 警车的鸣笛声; (b) 倒车的提示音; (c) 消防车的鸣笛声; (d) 轿车的鸣笛声; (e) 消防车的鸣笛声; (f) 自行车的铃声。

Figure 5.12 Spectrogram (top) and power spectral density (bottom) of six types of signals often heard in the street, including (a) police car, (b) backing car, (c) ambulance, (d) car whistle, (e) fire engine and (f) bicycle bell.

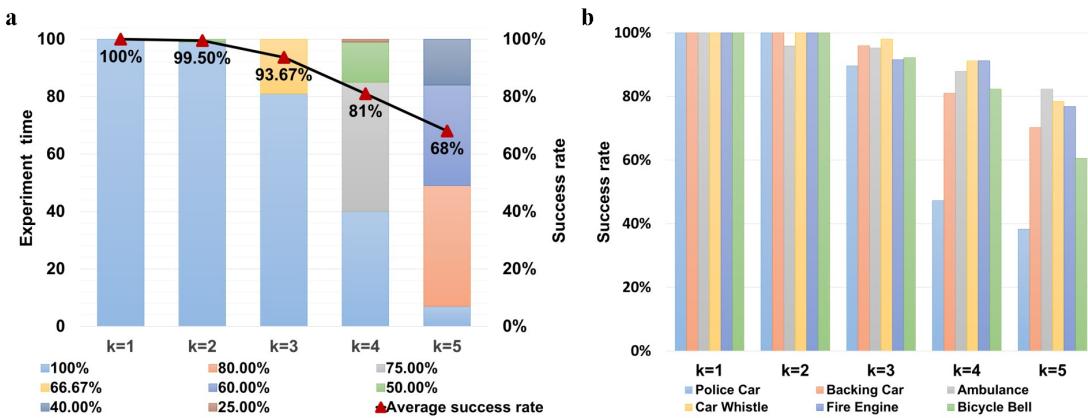


图 5.13 街道场景下听音测试结果

Figure 5.13 The results of the listening tests in the scenario of the street organized by (a) the number of activated sources and (b) the audio contents.

以上。这一测试过程的相关视频可以从 <https://doi.org/10.1002/advs.201902271> 处获取。

2. 动物园：

在动物园场景下，我们选择的 6 种声音信号分别为猫的叫声，蝉鸣声，蟋蟀的叫声，蜜蜂的叫声，青蛙的叫声和鸭子的叫声。这些声音信号均为宽带信号，他们的语谱图和功率谱如图 5.14 所示。针对 k 从 1 至 5 分别进行了 100 次随机实验，实验结果如图 5.15 所示，这两幅图的绘制方式与图 5.13 相似，在此不再赘述。通过观察图 5.15 可知，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统在动物园场景下的重建成功率逐渐下降，分别为 100%，97%，91.67%，86.75% 和 77.8%。当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80% 以上。

3. 音乐厅：

在音乐厅场景下，我们选择的 6 种声音信号分别为吉他，编钟，萨克斯，鼓，小提琴和中国民乐。这些声音信号均为宽带信号，他们的语谱图和功率谱如图 5.16 所示。针对 k 从 1 至 5 分别进行了 100 次随机实验，实验结果如图 5.17 所示，这两幅图的绘制方式与图 5.13 相似，在此不再赘述。通过观察图 5.17 可知，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统在音乐厅场景下的重建成功率逐渐下降，分别为 100%，97.5%，93.67%，80.25%

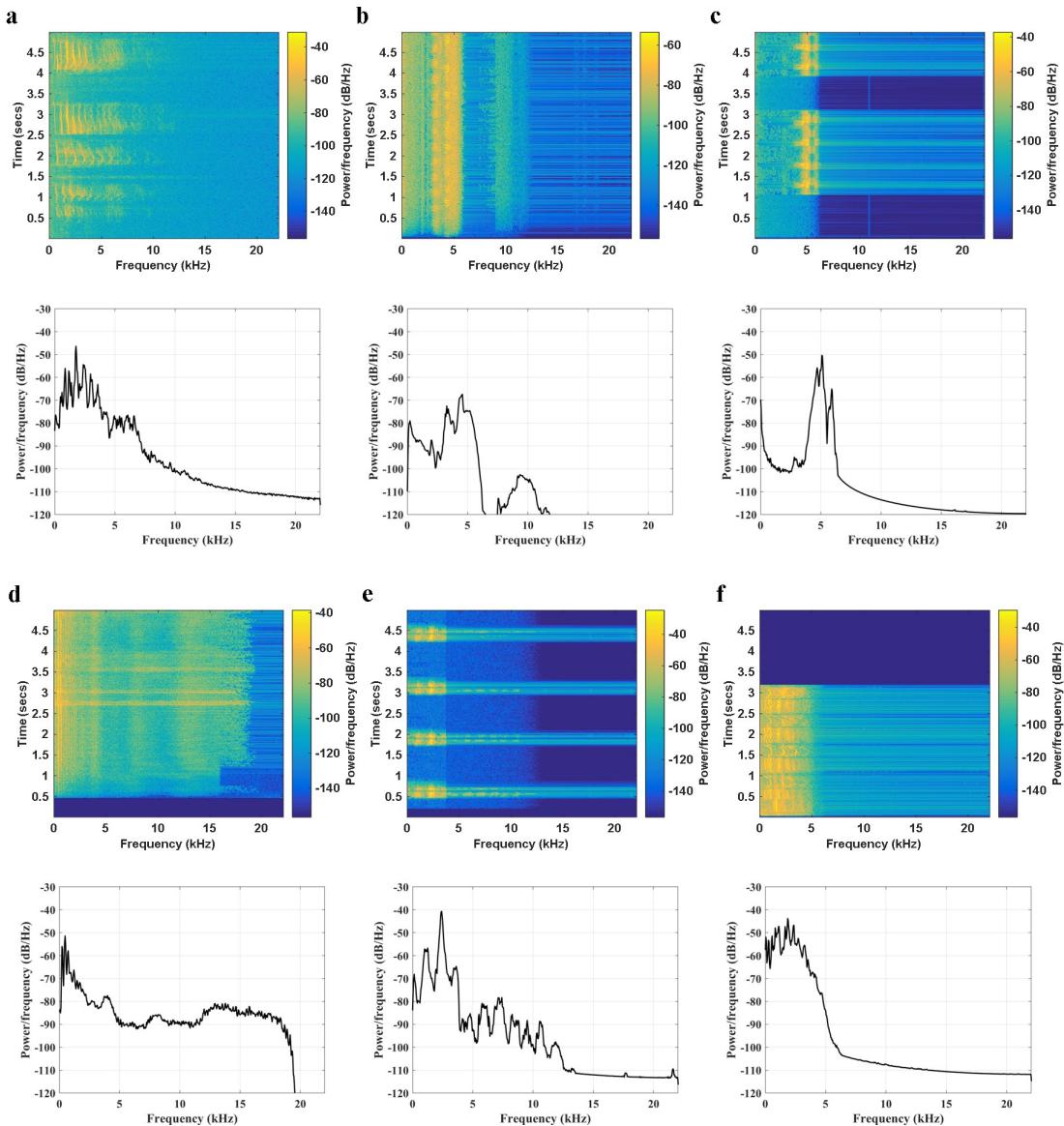


图 5.14 动物园场景下语料的语谱图和功率谱; (a) 猫的叫声; (b) 蝉鸣声; (c) 蟋蟀的叫声;
(d) 蜜蜂的叫声; (e) 青蛙的叫声; (f) 鸭子的叫声。

Figure 5.14 Spectrogram (top) and power spectral density (bottom) of six types of signals often heard in the animal farm, including (a) cat, (b) cicada, (c) cricket, (d) bee, (e) frog and (f) duck.

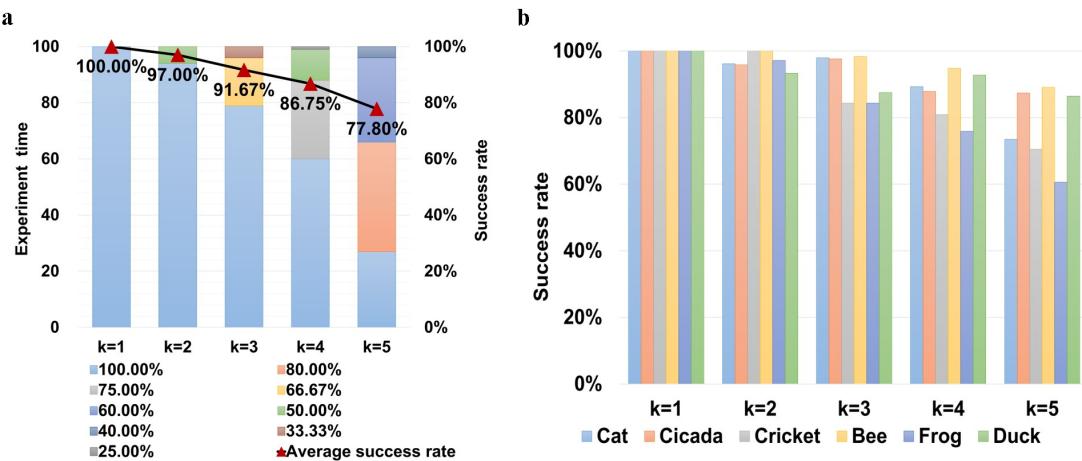


图 5.15 动物园场景下听音测试结果

Figure 5.15 The results of the listening tests in the scenario of an animal farm organized by (a) the number of activated sources and (b) the audio contents.

和 69.6%。当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80% 以上。

4. 家庭：

在家庭场景下，我们选择的 6 种声音信号分别为做饭的声音，抽水马桶的声音，吹风机吹头发的声音，笑声，撕纸的声音和东西碎了的声音。这些声音信号均为宽带信号，他们的语谱图和功率谱如图 5.18 所示。针对 k 从 1 至 5 分别进行了 100 次随机实验，实验结果如图 5.19 所示，这两幅图的绘制方式与图 5.13 相似，在此不再赘述。通过观察图 5.19 可知，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统在家庭场景下的重建成功率逐渐下降，分别为 100%，97%，92.67%，81.75% 和 78.4%。当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80% 以上。

5. 对话：

在对话场景下，我们选择了 3 段不同男性讲话的声音和 3 段不同女性讲话的声音作为测试的语料。这些声音信号均为宽带信号，它们的语谱图和功率谱如图 5.20 所示。针对 k 从 1 至 5 分别进行了 100 次随机实验，实验结果如图 5.21 所示，这两幅图的绘制方式与图 5.13 相似，在此不再赘述。通过观察图 5.21 可知，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系

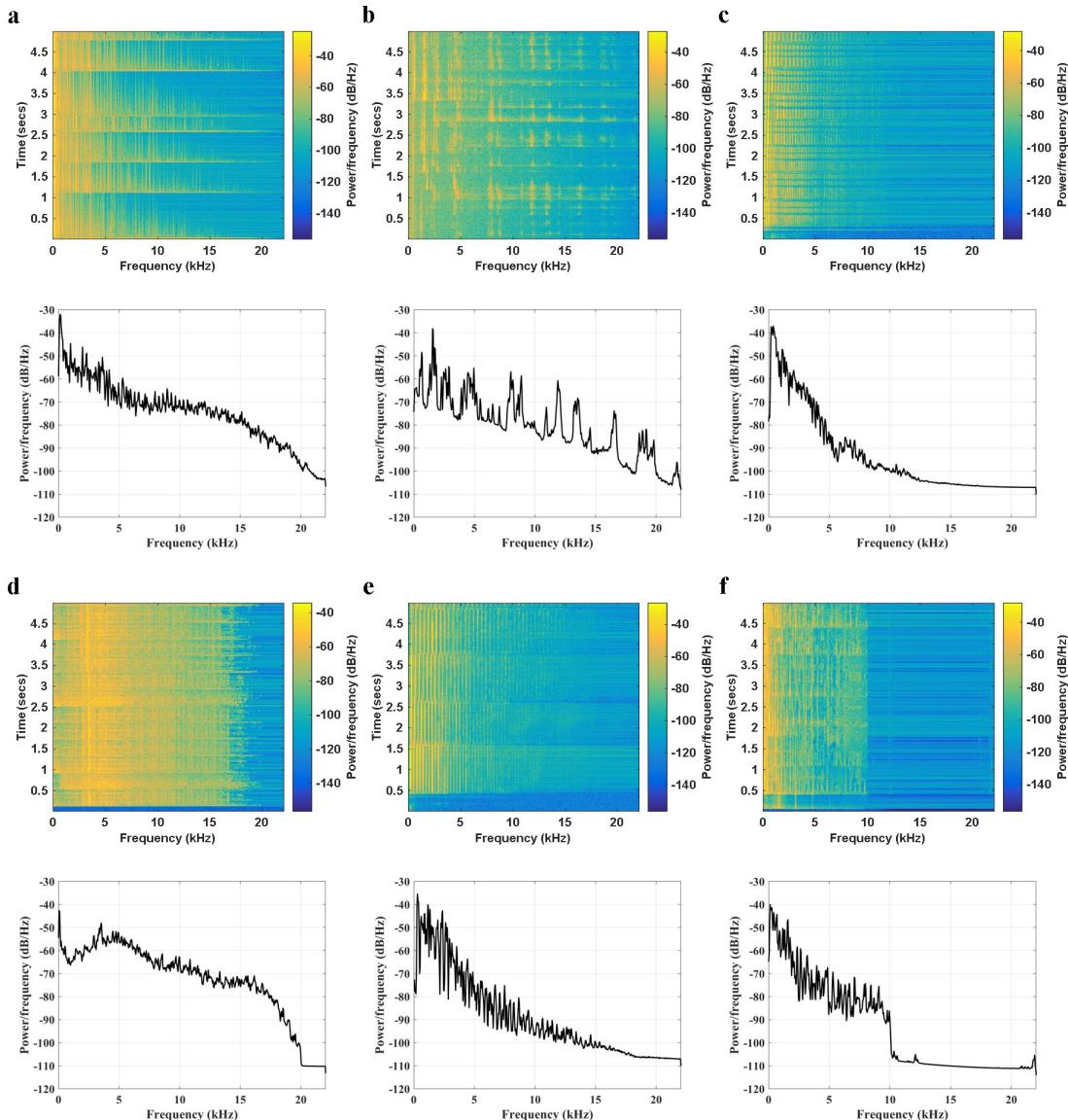


图 5.16 音乐厅场景下语料的语谱图和功率谱: (a) 吉他; (b) 编钟; (c) 萨克斯; (d) 鼓; (e) 小提琴; (f) 中国民乐。

Figure 5.16 Spectrogram (top) and power spectral density (bottom) of six types of signals often heard in the concert, including (a) guitar, (b) chime, (c) saxophone, (d)drum, (e) violin and (f) Chinese folk music.

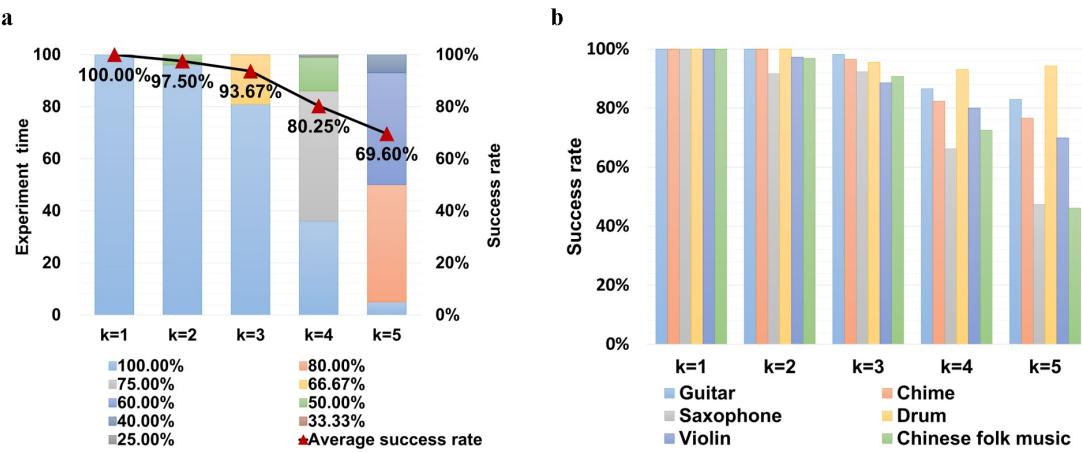


图 5.17 音乐厅场景下听音测试结果

Figure 5.17 The results of the listening tests in the scenario of a concert organized by (a) the number of activated sources and (b) the audio contents.

统在家庭场景下的重建成功率逐渐下降，分别为 100%，97.5%，91.67%，81% 和 69.4%。当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80% 以上。

此外，为了测试系统在规模更大的语料库上的性能，我们使用谷歌开源的 Speech Commands 命令词数据库对系统进行了听音测试 [95]。该数据库中包含了 30 个诸如“yes”，“no”等常用的命令词，测试方法与上述的听音测试相同，这里不再赘述。测试结果如图 5.22 所示，当空间中同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统在家庭场景下的重建成功率逐渐下降，分别为 100%，96.25%，84.67%，79.83% 和 73.7%。当空间中同时发声的声源个数不多于 2 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 80% 以上。从整体来看，本次测试结果略差于之前的结果，这是由于语料库规模增大所导致的。

最后，我们还对系统在俯仰维度上的定位分离能力做了针对性测试。如图 5.23(a) 所示，我们将半球状外壳垂直布放在一张桌子上，并在其周围的不同的方位布放了多个扬声器作为声源，所使用的测试语料为图 5.12 所示的街道场景下的 6 种常见信号。测试结果如图 5.23(b) 所示，当同时发声的声源个数从 1 增加至 5 时，所提出的单通道定位与分离系统的重建成功率逐渐下降，分别为 100%，99%，95.67%，83% 和 69%。测试结果进一步证明了该系统在俯仰维度上

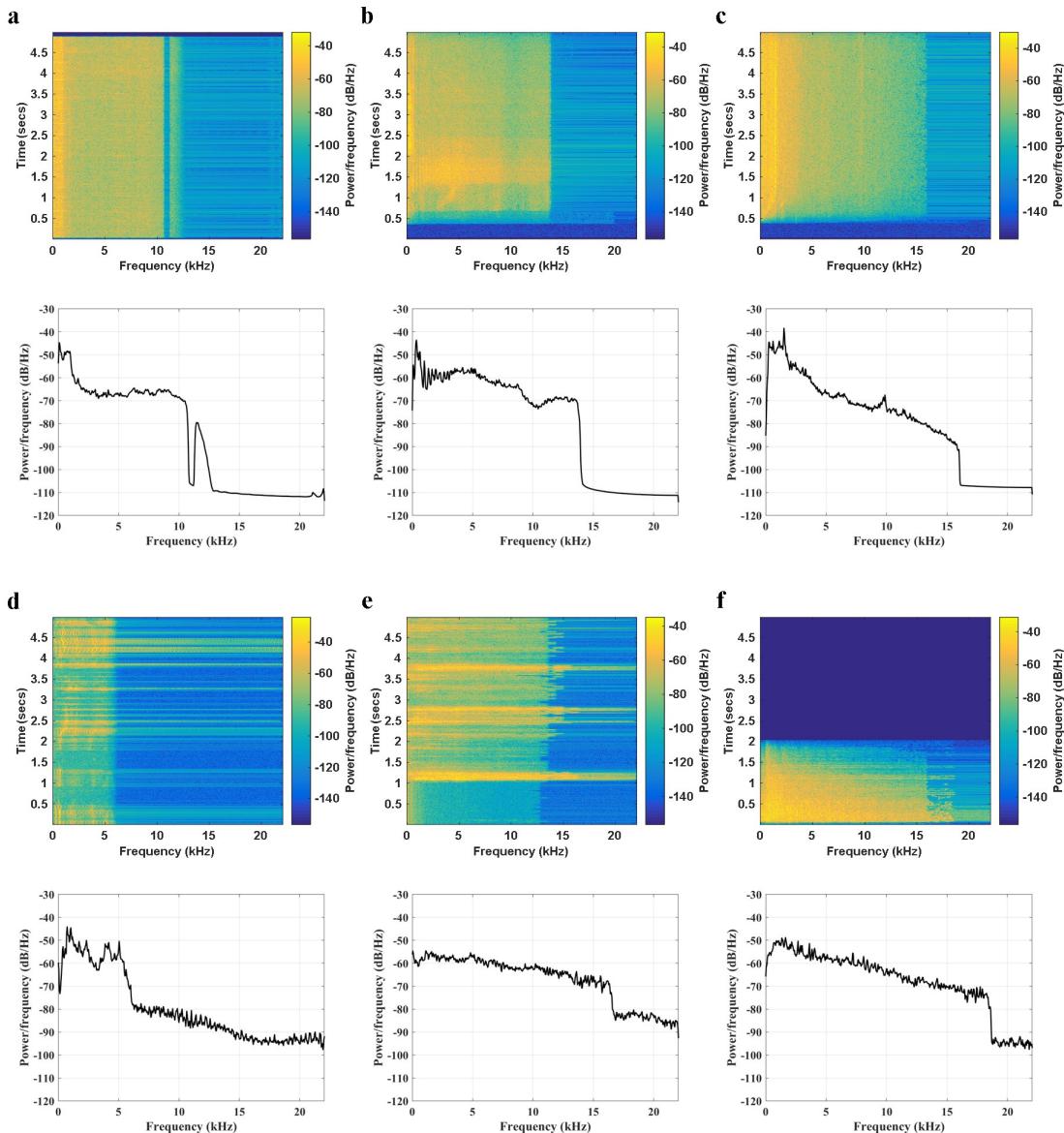


图 5.18 家庭场景下语料的语谱图和功率谱: (a) 做饭的声音; (b) 抽水马桶的声音; (c) 吹风机吹头发的声音; (d) 消声; (e) 撕纸的声音; (f) 东西碎了的声音。

Figure 5.18 Spectrogram (top) and power spectral density (bottom) of six types of signals often heard in the home, including (a) cooking, (b) flushing the toilet, (c) drying hair, (d) laughter, (e) tearing up paper and (f) something breaking.

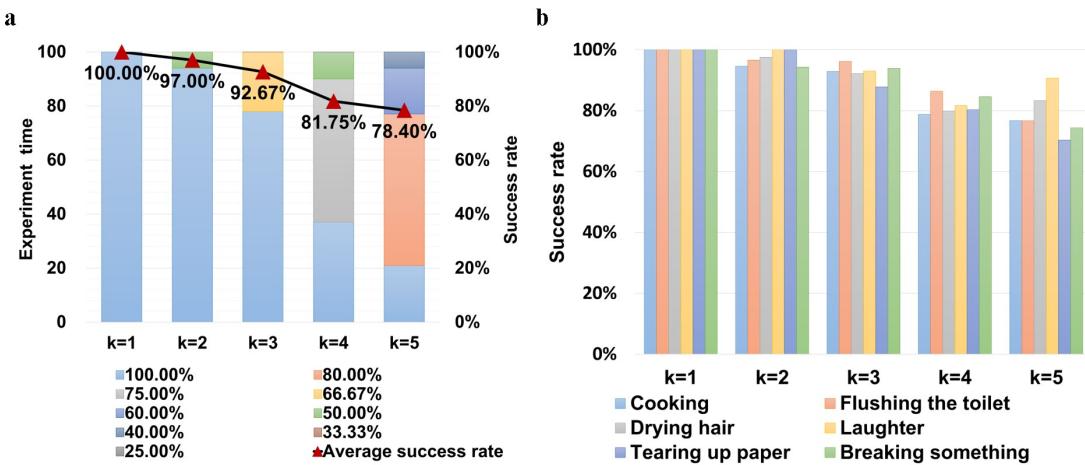


图 5.19 家庭场景下听音测试结果

Figure 5.19 The results of the listening tests in the scenario of home organized by (a) the number of activated sources and (b) the audio contents.

依然具有较强的空间辨识能力。

5.4.3 实时追踪

在 5.3.3 章节中，我们对所使用的稀疏恢复算法进行了介绍。其中较为核心的 OMP 算法属于贪婪类的算法，计算复杂度相对较低，这使得我们的系统具有较好的实时性。因此，所提出的系统不仅可以对声源进行定位和分离，还可以对多个移动的声源进行实时追踪。为了评估系统的实时追踪性能，我们依然在半消声室中做了相关测试。在测试过程中，我们人为地使扬声器在系统附近移动并不断播放图 5.12 中的汽车鸣笛声，以此来模拟在街道场景下行驶的汽车。同时，所提出的系统对环境内的声音进行监测，并实时输出车辆的行驶轨迹。

如图 5.24(a) 所示为使用 1 个随机移动扬声器播放救护车声音时，系统输出的声源轨迹；如图 5.24(b) 所示为使用 2 个随机移动扬声器分别播放消防车的声音和倒车的声音时，系统输出的声源轨迹。上述测试结果表明，我们的系统不仅可以在 1s 内完成对声源种类的识别，还可以实时地追踪到声源的方位形成移动轨迹。此外，在测试过程中人为地移动扬声器会带来不可避免的噪声，这些噪声也被系统监测到了。然而，在并没有为系统增加额外的降噪模块的条件下，系统依然表现出较为精准的辨识能力。这是由于压缩感知算法本身是具有一定的抗噪能力的，当噪声与目标信号相比能量较低时，算法在重建时会将其分解到多个字典条目对应的系数上，但是每个系数都将是一个小量，在后续处理中会被忽

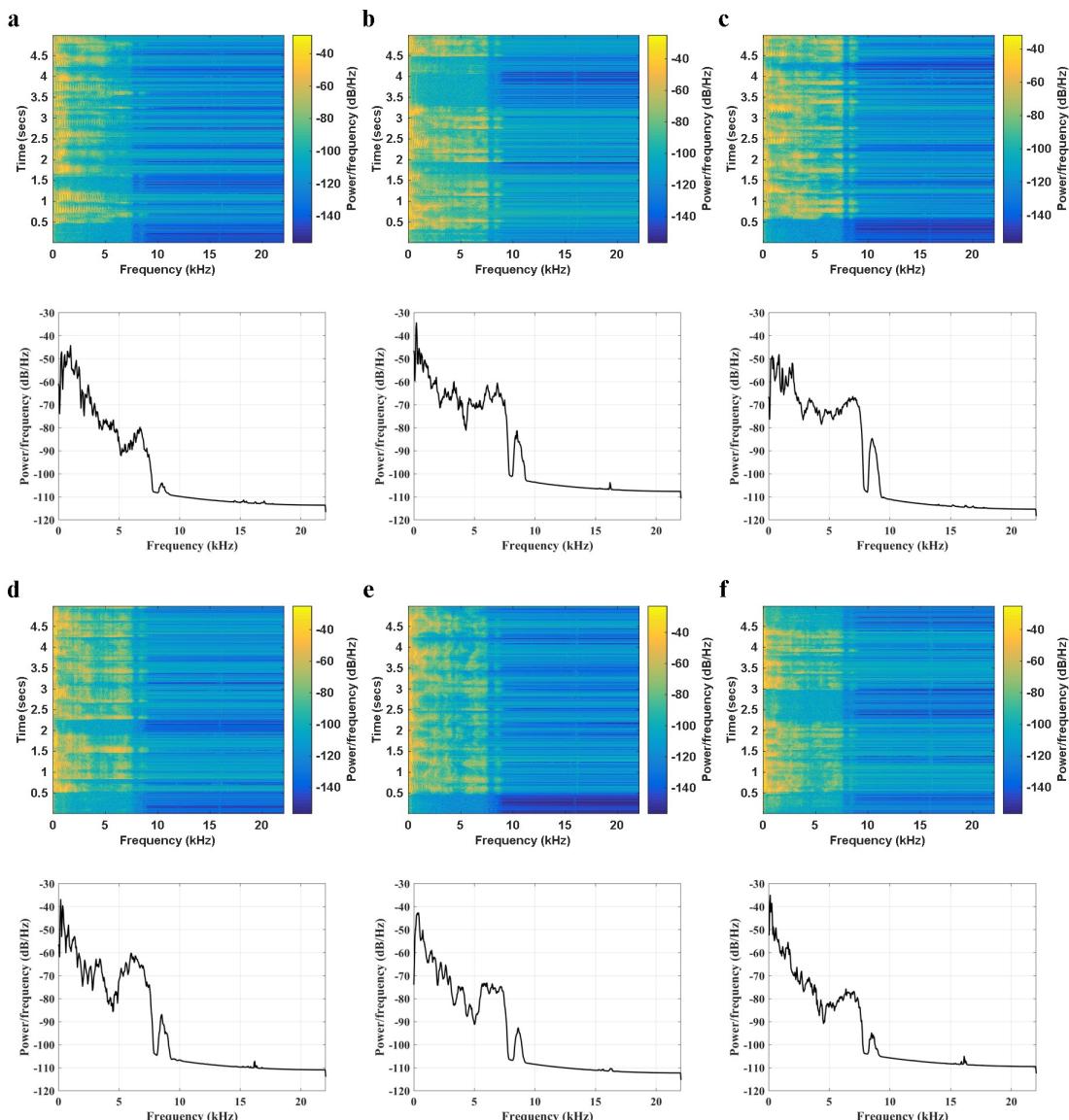


图 5.20 对话场景下语料的语谱图和功率谱: (a)-(c) 三位不同男性讲话的声音; (d)-(f) 三位不同女性讲话的声音。

Figure 5.20 Spectrogram (top) and power spectral density (bottom) of signals often heard in the speech, including (a)-(c) female speech and (d)-(f) male speech.

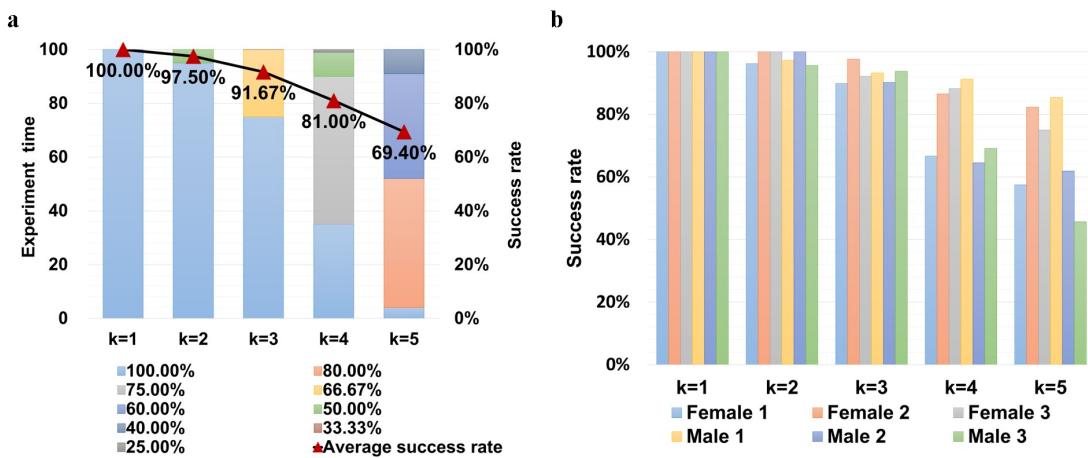


图 5.21 对话场景下听音测试结果

Figure 5.21 The results of the listening tests in the scenario of a speech organized by (a) the number of activated sources and (b) the audio contents.

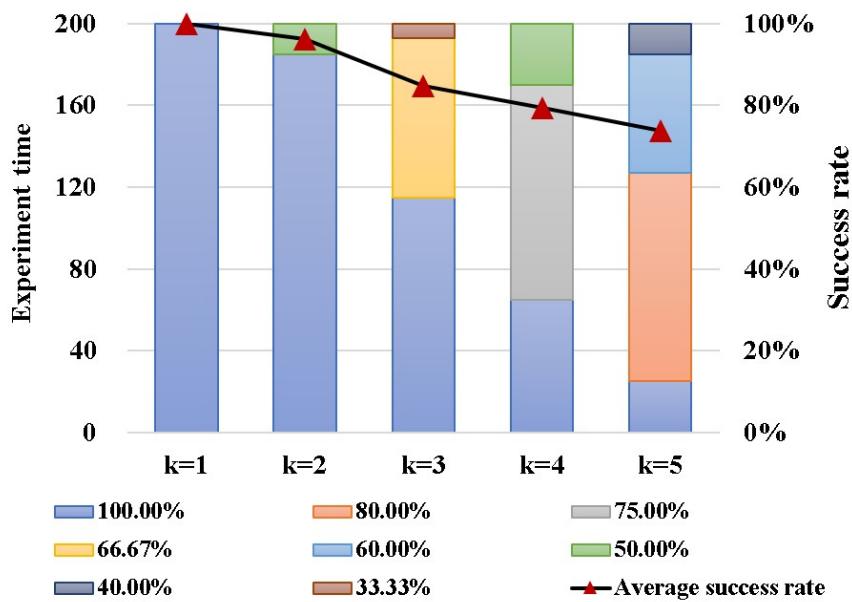


图 5.22 基于 Speech Commands 数据库的测试结果

Figure 5.22 The results of the listening tests based on the Speech Commands dataset.

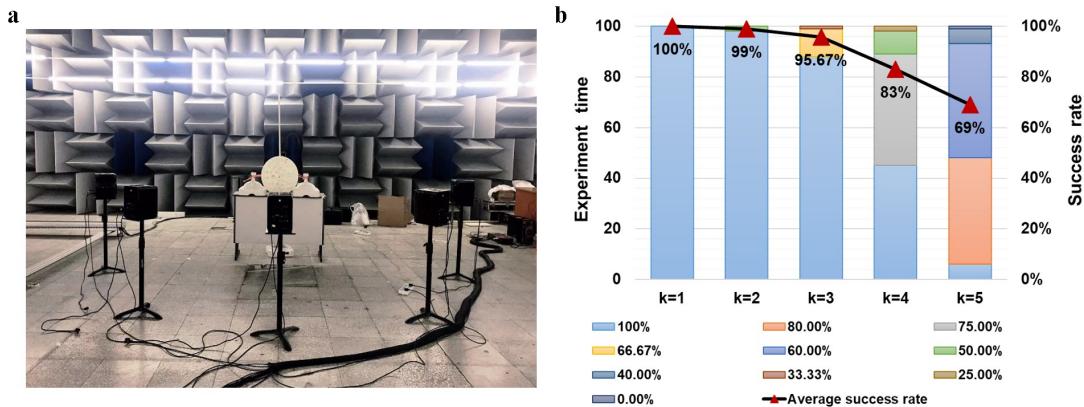


图 5.23 针对俯仰方向的听音测试及结果

Figure 5.23 Pitching direction Experiment and results.

略。上述实验的视频可以从 <https://doi.org/10.1002/advs.201902271> 处获取。

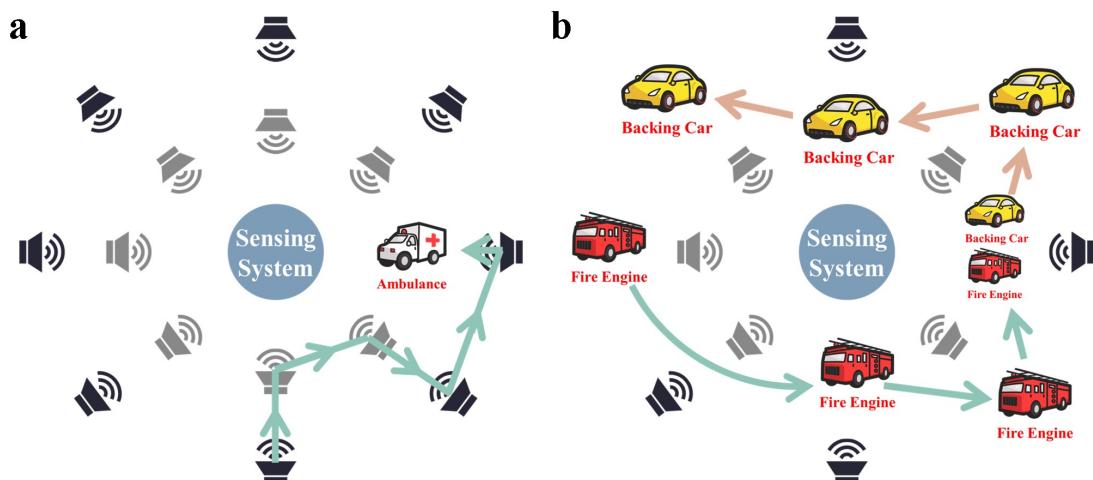


图 5.24 对移动声源进行实时追踪

Figure 5.24 Listening tests of source identification and tracking.

5.5 本章小结

在本章中，我们提出了一种基于声人工结构的智能化声学传感系统。该系统在原理上借鉴了人的听觉系统，在只使用一个单通道传声器的基础上在三维空间内实现了多个声源的实时定位与分离。首先我们基于声学滤波器的工作原理设计了一个空间上高度各向异性的半球状外壳，该外壳打破了单通道传声器的全指向性，在物理层实现了对来自不同方向声波的差异性编码；随后，我们提出了具有较低计算复杂度的 VSPCA-OMP 算法，基于上述差异性对声音的方位信

息和音频内容信息进行了重建。最后，我们在多个常见的声学场景下对该系统进行了听音测试。实验结果表明，当空间中同时发声的声源个数不多于 3 个时，分离和定位的准确率维持在 90% 以上；当空间中同时发声的声源个数不多于 4 个时，分离和定位的准确率维持在 80% 以上。并且，该系统可以在 1s 内输出定位和分离的结果，具有较好的实时性，因此也可用于对多个移动声源的实时追踪。

第6章 总结与展望

6.1 工作总结

本文对声学结构的智能化设计与应用展开了理论研究和实验验证，主要包含了声学结构的自动化设计方法和在声学系统中的智能化应用两大部分。在声学结构自动化设计方面，本文分别提出了基于遗传算法（Genetic Algorithm, GA）的声学结构优化方法和基于深度学习的声学结构逆设计方法；在声学结构的智能化应用方面，本文对声学结构与信号处理算法进行了联合设计，提出了一种新型的声学传感系统，实现了三维空间中单通道多声源的定位与分离。本论文的主要内容包括：

1. 提出了一种较为通用的基于 GA 的声学结构优化策略，并针对多种声学结构和优化目标对所提出的优化策略展开了分析和讨论。对于几何结构相对简单的声学结构，可以将 GA 与解析类的正向计算模型相结合，通过多次地正向迭代对声学结构进行优化。作为具体的案例，我们将 GA 与层状严格耦合波分析（Layered Rigorous Coupled-wave Analysis, LRCWA）相结合对吸声棉的形状进行优化，显著提升了目标频段的吸声效果。当我们所设计的声学结构较为复杂时，通常需要将 GA 与有限元仿真软件相结合对复杂结构进行优化。作为具体的案例，我们将 GA 与 COMSOL Multiphysics 相结合，利用软件内置的 MATLAB 接口实现了对基于五模材料的声学隐身器件的分层优化，显著提升了其隐身效果。
2. 将深度学习模型与集总参数模型相结合，实现了声学结构的快速精准化逆设计。首先，以吸隔声领域较为经典的二阶亥姆霍兹共鸣器（Two-order Helmholtz Resonator, THR）为例，介绍并对比了多种正向计算方法，生成并筛选了用于模型训练的样本数据；然后，对声学结构逆设计问题的复杂性和非唯一性展开讨论，在集总参数模型下基于深度神经网络建立了期望的声学响应和等效电学参数之间的映射关系，实现了 THR 单元的按需设计；最后，针对实际应用中低频宽带的隔声需求，基于所提出的深度学习模型对多个 THR 单元进行组合设计，完成了低频宽带隔声器件的构造。
3. 探索了深度学习辅助的多样化声人工结构设计，进一步拓展了深度学习在声人工结构设计领域的应用。考虑到声学结构逆设计问题的非唯一性，将深度

学习模型与降维类算法相结合，可以快速生成多个具有期望声学性质的声学结构；针对启发类优化算法对初值敏感这一问题，深度学习模型还可以为优化类算法提供良好的初始条件，提高优化效率，在一定程度上避免进化过早地陷入局部最优解；最后，将所提出的深度神经网络在集总参数模型的框架下向其他声学结构进行迁移，进一步展现了该设计方法的通用性。

4. 在三维空间内提出了一种基于声人工结构的智能化声学传感系统。充分考虑了实际应用中声学结构与信号处理算法的适配性，设计了一个空间上高度各向异性的半球状外壳，该外壳可以基于物理层的声学滤波机制实现对来自不同方向声波的差异性编码，从而打破了单通道传声器的全指向性；在压缩感知的框架下，提出了一种与所设计的外壳结构相适应的 VSPCA-OMP 算法，以较低的复杂度实现了对声音的方位信息和音频内容信息的重建，并通过多场景的听音测试对该系统的定位分离精度及实时性进行了验证。

6.2 前景与展望

本文对声学结构的智能化设计与优化展开了系统地研究，对声学结构在智能系统中的应用进行了探索，构建了物理模型、理论方法与器件原型，形成了一条完整的技术路线。虽然总结了一定的规律，取得了一定的成果，但由于研究经验、研究时间与实验条件等限制，以及个人精力等诸多方面的因素，本文中的一些研究方法与研究结论仍有进一步探索的空间，总结如下：

1. 复杂声学结构的多参数优化问题。实际应用中的声学结构设计往往包含多个待优化参数，这些参数之间往往存在一定的物理关联。而本文所提出的优化算法对所有的待优化变量均进行了无差别编码，因此优化过程中并没有考虑到变量之间的耦合关系，在对具有多个参数的复杂模型进行优化时效果往往并不理想。而对于一个具有实际功能的声学器件，往往是由很多个单元构成的，这些参数之间并不是无差别的，因此我们可以通过前期的理论分析将参数之间的耦合关系抽象成数学模型，将优化目标进行分解，并基于分布式优化的思想完成对复杂声学结构的优化。

2. 基于生成类深度学习模型解决声学结构的逆设计问题。与深度学习在其他领域的应用相比，基于深度学习对声学结构进行设计最大的不同之处在于模型所建立的映射关系是一对多的，即对于同一个声学特性往往存在多个与之对

应的声学结构。如果不做任何的处理，这种非唯一性会给模型训练带来很大的困难，因此本文所采用的策略是通过调整模型的输入输出来削弱这种非唯一性。而深度学习中的生成模型可以用于建立变量间的条件概率分布，因而可以对观测变量和目标变量之间更复杂的关系进行建模。因此，我们可以尝试采用高斯混合模型（Gaussian Mixture Model, GMM），生成式对抗网络（Generative Adversarial Networks, GAN）等生成模型对这种非唯一性问题进行求解。

3. 基于声学结构的智能传感系统的小型化问题。本文在第5章中提出了基于声学结构的单通道多声源辨识系统的装置原型，所设计的半球状外壳的物理机制决定了系统的工作频段很大程度上受限于外壳的尺寸大小。因此，在现有的物理模型下对系统进行小型化改造会对系统的低频性能造成影响。作者认为，一个可行的解决思路是在设计外壳的过程中引入薄膜结构，利用其在低频对声波进行调制，以实现智能传感系统的小型化。

4. 基于声学结构的智能传感系统的源依赖问题和环境依赖问题。本文在第5章所提出的系统是基于字典学习算法对信号进行后处理的，该算法需要预先掌握关于使用环境和目标声源的先验知识，并以此对字典矩阵进行训练。这种对声源和环境的依赖性很大程度上限制了该系统的应用及推广。作者认为，在接下来的研究中可以尝试使用深度学习模型消除上述依赖因素，使该系统可以在任意环境下对陌生信号进行辨识。

参考文献

- [1] RAYLEIGH J W S B. The theory of sound: volume 2[M]. Macmillan, 1896.
- [2] 杨庆余, 周荣生. 瑞利勋爵对现代声学基础理论的开创性贡献[J]. 徐州师范大学学报(自然科学版), 2001(03):36-39.
- [3] LU W, JIA H, BI Y, et al. Design and demonstration of an acoustic right-angle bend[J/OL]. The Journal of the Acoustical Society of America, 2017, 142(1):84-89[2021-04-01]. <https://asa.scitation.org/doi/full/10.1121/1.4990952>.
- [4] SUN Z, JIA H, CHEN Y, et al. Design of an underwater acoustic bend by pentamode metafluid [J/OL]. The Journal of the Acoustical Society of America, 2018, 143(2):1029-1034[2021-04-01]. <https://asa.scitation.org/doi/full/10.1121/1.5024351>.
- [5] BI Y, JIA H, SUN Z, et al. Experimental demonstration of three-dimensional broadband underwater acoustic carpet cloak[J/OL]. Applied Physics Letters, 2018, 112(22):223502[2021-09-26]. <https://doi.org/10.1063/1.5026199>.
- [6] SUN Z, SUN X, JIA H, et al. Quasi-isotropic underwater acoustic carpet cloak based on latticed pentamode metafluid[J/OL]. Applied Physics Letters, 2019, 114(9):094101[2021-09-26]. <https://doi.org/10.1063/1.5085568>.
- [7] ZHAO H, GUAN X, JIA H, et al. Design and verification of ultrathin Schroeder diffuser based on membrane-type resonators[J/OL]. Applied Physics Express, 2021, 14(9):094005. <https://doi.org/10.35848/1882-0786/ac1c42>.
- [8] HINTON G, DENG L, YU D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups[J/OL]. IEEE Signal Processing Magazine, 2012, 29(6):82-97. DOI: [10.1109/MSP.2012.2205597](https://doi.org/10.1109/MSP.2012.2205597).
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks[J/OL]. Commun. ACM, 2017, 60(6):84-90. <https://doi.org/10.1145/3065386>.
- [10] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation[J/OL]. 2014[2022-03-13]. <https://arxiv.org/abs/1406.1078v3>. DOI: [10.48550/arXiv.1406.1078](https://doi.org/10.48550/arXiv.1406.1078).
- [11] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J/OL]. Nature, 2016, 529(7587):484-489. <https://doi.org/10.1038/nature16961>.
- [12] SOCHER R, CHEN D, MANNING C D, et al. Reasoning With Neural Tensor Networks for Knowledge Base Completion[C/OL]//Advances in Neural Information Processing Systems:

- volume 26. Curran Associates, Inc., 2013[2022-03-13]. <https://proceedings.neurips.cc/paper/2013/hash/b337e84de8752b27eda3a12363109e80-Abstract.html>.
- [13] ZANGENEH-NEJAD F, FLEURY R. Active times for acoustic metamaterials[J/OL]. Reviews in Physics, 2019, 4:100031[2022-02-17]. <https://linkinghub.elsevier.com/retrieve/pii/S2405428318300649>. DOI: [10.1016/j.revip.2019.100031](https://doi.org/10.1016/j.revip.2019.100031).
- [14] LIU Z, ZHANG X, MAO Y, et al. Locally resonant sonic materials[J]. science, 2000, 289(5485):1734-1736.
- [15] LEE S H, PARK C M, SEO Y M, et al. Composite acoustic medium with simultaneously negative density and modulus[J]. Physical review letters, 2010, 104(5):054301.
- [16] FANG N, XI D, XU J, et al. Ultrasonic metamaterials with negative modulus[J]. Nature materials, 2006, 5(6):452-456.
- [17] PARK C M, PARK J J, LEE S H, et al. Amplification of acoustic evanescent waves using metamaterial slabs[J]. Physical review letters, 2011, 107(19):194301.
- [18] CSELYUSZKA N, SEČUJSKI M, BENGIN V C. Compressibility-near-zero acoustic metamaterial[J]. Physics Letters A, 2014, 378(16-17):1153-1156.
- [19] DUBOIS M, SHI C, ZHU X, et al. Observation of acoustic dirac-like cone and double zero refractive index[J]. Nature communications, 2017, 8(1):1-6.
- [20] MEI J, MA G, YANG M, et al. Dark acoustic metamaterials as super absorbers for low-frequency sound[J]. Nature communications, 2012, 3(1):1-7.
- [21] LI Y, JIANG X, LIANG B, et al. Metascreen-based acoustic passive phased array[J]. Physical Review Applied, 2015, 4(2):024003.
- [22] MELDE K, MARK A G, QIU T, et al. Holograms for acoustics[J]. Nature, 2016, 537(7621):518-522.
- [23] ZHU Y, FAN X, LIANG B, et al. Ultrathin acoustic metasurface-based schroeder diffuser[J]. Physical Review X, 2017, 7(2):021034.
- [24] MEMOLI G, CALEAP M, ASAOKAWA M, et al. Metamaterial bricks and quantization of meta-surfaces[J]. Nature communications, 2017, 8:14608.
- [25] LIU C, SHI J, ZHAO W, et al. Three-Dimensional Soundproof Acoustic Metacage[J/OL]. Physical Review Letters, 2021, 127(8):084301[2021-12-23]. <https://link.aps.org/doi/10.1103/PhysRevLett.127.084301>.
- [26] YANG Y, JIA H, BI Y, et al. Experimental Demonstration of an Acoustic Asymmetric Diffraction Grating Based on Passive Parity-Time-Symmetric Medium[J/OL]. Phys. Rev. Applied, 2019, 12(3):034040. <https://link.aps.org/doi/10.1103/PhysRevApplied.12.034040>.
- [27] YANG Y, JIA H, WANG S, et al. Diffraction control in a non-Hermitian acoustic grating

- [J/OL]. Applied Physics Letters, 2020, 116(21):213501[2021-09-26]. <https://doi.org/10.1063/5.0004104>.
- [28] SUN X, JIA H, ZHANG Z, et al. Sound Localization and Separation in 3D Space Using a Single Microphone with a Metamaterial Enclosure[J/OL]. Advanced Science, 2020, 7(3): 1902271[2021-09-26]. <https://doi.org/10.1002/advs.201902271>.
- [29] WRIGHT L G, ONODERA T, STEIN M M, et al. Deep physical neural networks trained with backpropagation[J/OL]. Nature, 2022, 601(7894):549-555[2022-02-22]. <https://www.nature.com/articles/s41586-021-04223-6>. DOI: [10.1038/s41586-021-04223-6](https://doi.org/10.1038/s41586-021-04223-6).
- [30] HUGHES T W, WILLIAMSON I A D, MINKOV M, et al. Wave physics as an analog recurrent neural network[J/OL]. Science Advances, 2019, 5(12):eaay6946. <https://www.science.org/doi/abs/10.1126/sciadv.aay6946>.
- [31] MA C, KIM S, FANG N X. Far-field acoustic subwavelength imaging and edge detection based on spatial filtering and wave vector conversion[J/OL]. Nature Communications, 2019, 10(1):204. <https://doi.org/10.1038/s41467-018-08221-7>.
- [32] ZANGENEH-NEJAD F, SOUNAS D L, ALÙ A, et al. Analogue computing with metamaterials[J/OL]. Nature Reviews Materials, 2021, 6(3):207-225. <https://doi.org/10.1038/s41578-020-00243-2>.
- [33] ROGERS J S, ROHDE C A, GUILD M D, et al. Demonstration of acoustic source localization in air using single pixel compressive imaging[J]. Journal of Applied Physics, 2017, 122(21): 214901.
- [34] KRUIZINGA P, VAN DER MEULEN P, FEDJAJEVS A, et al. Compressive 3d ultrasound imaging using a single sensor[J]. Science advances, 2017, 3(12):e1701423.
- [35] XIE Y, TSAI T H, KONNEKER A, et al. Single-sensor multispeaker listening with acoustic metamaterials[J]. Proceedings of the National Academy of Sciences, 2015, 112(34):10595-10598.
- [36] JIANG T, HE Q, PENG Z K. Proposal for the realization of a single-detector acoustic camera using a space-coiling anisotropic metamaterial[J]. Physical Review Applied, 2019, 11(3): 034013.
- [37] D'ALESSANDRO L, BAHR B, DANIEL L, et al. Shape optimization of solid-air porous phononic crystal slabs with widest full 3D bandgap for in-plane acoustic waves[J/OL]. Journal of Computational Physics, 2017, 344:465-484. <https://www.sciencedirect.com/science/article/pii/S0021999117303935>. DOI: <https://doi.org/10.1016/j.jcp.2017.05.018>.
- [38] KOOK J, JENSEN J S. Topology optimization of periodic microstructures for enhanced loss factor using acoustic-structure interaction[J/OL]. International Journal of Solids

- and Structures, 2017, 122-123:59-68. <https://www.sciencedirect.com/science/article/pii/S0020768317302573>. DOI: <https://doi.org/10.1016/j.ijsolstr.2017.06.001>.
- [39] DONG H W, ZHAO S D, WEI P, et al. Systematic design and realization of double-negative acoustic metamaterials by topology optimization[J/OL]. Acta Materialia, 2019, 172:102-120. <https://www.sciencedirect.com/science/article/pii/S1359645419302447>. DOI: <https://doi.org/10.1016/j.actamat.2019.04.042>.
- [40] LU Z, SANCHIS L, WEN J, et al. Acoustic cloak based on Bézier scatterers[J/OL]. Scientific Reports, 2018, 8(1):12924. <https://doi.org/10.1038/s41598-018-30888-7>.
- [41] YANG M, CHEN S, FU C, et al. Optimal sound-absorbing structures[J/OL]. Mater. Horiz., 2017, 4(4):673-680. <http://dx.doi.org/10.1039/C7MH00129K>.
- [42] AGRAWAL A, CHOUDHARY A. Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science[J/OL]. APL Materials, 2016, 4 (5):053208. <https://aip.scitation.org/doi/abs/10.1063/1.4946894>.
- [43] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks[C/OL]//Advances in Neural Information Processing Systems: volume 25. Curran Associates, Inc., 2012[2022-03-13]. <https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>.
- [44] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups[J/OL]. IEEE Signal Processing Magazine, 2012, 29(6):82-97. DOI: <10.1109/MSP.2012.2205597>.
- [45] SANCHEZ-LENGELING B, ASPURU-GUZIK A. Inverse molecular design using machine learning: Generative models for matter engineering[J/OL]. Science, 2018, 361(6400):360-365. <https://www.science.org/doi/abs/10.1126/science.aat2663>.
- [46] MAO Y, HE Q, ZHAO X. Designing complex architected materials with generative adversarial networks[J/OL]. Science Advances, 2020, 6(17):eaaz4169. <https://www.science.org/doi/abs/10.1126/sciadv.aaz4169>.
- [47] BESSA M A, GLOWACKI P, HOULDER M. Bayesian Machine Learning in Metamaterial Design: Fragile Becomes Supercompressible[J/OL]. Advanced Materials, 2019, 31(48):1904845. <https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201904845>. DOI: <https://doi.org/10.1002/adma.201904845>.
- [48] GOH G B, HODAS N O, VISHNU A. Deep learning for computational chemistry[J/OL]. Journal of Computational Chemistry, 2017, 38(16):1291-1307. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.24764>. DOI: <https://doi.org/10.1002/jcc.24764>.
- [49] CARRASQUILLA J, MELKO R G. Machine learning phases of matter[J/OL]. Nature Physics, 2017, 13(5):431-434. <https://doi.org/10.1038/nphys4035>.

- [50] LONG Y, REN J, CHEN H. Unsupervised manifold clustering of topological phononics [J/OL]. Phys. Rev. Lett., 2020, 124:185501. <https://link.aps.org/doi/10.1103/PhysRevLett.124.185501>.
- [51] MA W, LIU Z, KUDYSHEV Z A, et al. Deep learning for the design of photonic structures [J/OL]. Nature Photonics, 2021, 15(2):77-90. <https://doi.org/10.1038/s41566-020-0685-y>.
- [52] MA W, CHENG F, XU Y, et al. Probabilistic Representation and Inverse Design of Metamaterials Based on a Deep Generative Model with Semi-Supervised Learning Strategy[J/OL]. Advanced Materials, 2019, 31(35):1901111. <https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201901111>. DOI: <https://doi.org/10.1002/adma.201901111>.
- [53] LIU D, TAN Y, KHORAM E, et al. Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures[J/OL]. ACS Photonics, 2018, 5(4):1365-1369[2021-08-18]. <https://pubs.acs.org/doi/10.1021/acsphotonics.7b01377>.
- [54] KIARASHINEJAD Y, ABDOLLAHRAMEZANI S, ADIBI A. Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures[J/OL]. npj Computational Materials, 2020, 6(1):12. <https://doi.org/10.1038/s41524-020-0276-y>.
- [55] ZHU R, QIU T, WANG J, et al. Phase-to-pattern inverse design paradigm for fast realization of functional metasurfaces via transfer learning[J/OL]. Nature Communications, 2021, 12(1): 2974. <https://doi.org/10.1038/s41467-021-23087-y>.
- [56] MALKIEL I, MREJEN M, NAGLER A, et al. Plasmonic nanostructure design and characterization via Deep Learning[J/OL]. Light: Science & Applications, 2018, 7(1):60[2021-09-26]. <https://www.nature.com/articles/s41377-018-0060-7>. DOI: [10.1038/s41377-018-0060-7](https://doi.org/10.1038/s41377-018-0060-7).
- [57] Peurifoy John, Shen Yichen, Jing Li, et al. Nanophotonic particle simulation and inverse design using artificial neural networks[J/OL]. Science Advances, 2018, 4(6):eaar4206[2021-09-27]. <https://doi.org/10.1126/sciadv.aar4206>.
- [58] SHEVERDIN A, MONTICONE F, VALAGIANNOPoulos C. Photonic Inverse Design with Neural Networks: The Case of Invisibility in the Visible[J/OL]. Physical Review Applied, 2020, 14(2):024054[2021-01-18]. <https://link.aps.org/doi/10.1103/PhysRevApplied.14.024054>.
- [59] QIU T, SHI X, WANG J, et al. Deep Learning: A Rapid and Efficient Route to Automatic Metasurface Design[J/OL]. Advanced Science, 2019, 6(12):1900128. <https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.201900128>. DOI: <https://doi.org/10.1002/advs.201900128>.
- [60] LUO Y T, LI P Q, LI D T, et al. Probability-Density-Based Deep Learning Paradigm for the Fuzzy Design of Functional Metastructures[J/OL]. Research, 2020, 2020:8757403. <https://doi.org/10.34133/2020/8757403>.
- [61] AHMED W W, FARHAT M, ZHANG X, et al. Deterministic and probabilistic deep learning

- models for inverse design of broadband acoustic cloak[J/OL]. Phys. Rev. Research, 2021, 3(1):013142. <https://link.aps.org/doi/10.1103/PhysRevResearch.3.013142>.
- [62] GURBUZ C, KRONOWETTER F, DIETZ C, et al. Generative adversarial networks for the design of acoustic metamaterials[J/OL]. The Journal of the Acoustical Society of America, 2021, 149(2):1162-1174. <https://doi.org/10.1121/10.0003501>.
- [63] ZHANG H, WANG Y, ZHAO H, et al. Accelerated topological design of metaporous materials of broadband sound absorption performance by generative adversarial networks[J/OL]. Materials & Design, 2021, 207:109855. <https://www.sciencedirect.com/science/article/pii/S0264127521004081>. DOI: <https://doi.org/10.1016/j.matdes.2021.109855>.
- [64] BACIGALUPO A, GNECCO G, LEPIDI M, et al. Machine-Learning Techniques for the Optimal Design of Acoustic Metamaterials[J/OL]. Journal of Optimization Theory and Applications, 2020, 187(3):630-653. <https://doi.org/10.1007/s10957-019-01614-8>.
- [65] TANG K, MAN K, KWONG S, et al. Genetic algorithms and their applications[J/OL]. IEEE Signal Processing Magazine, 1996, 13(6):22-37. DOI: [10.1109/79.543973](https://doi.org/10.1109/79.543973).
- [66] DEB K, AGRAWAL S, PRATAP A, et al. A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II[C]//International Conference on Parallel Problem Solving from Nature. 2000.
- [67] WANG T, LI S, NUTT S R. Optimal design of acoustical sandwich panels with a genetic algorithm[J/OL]. Applied Acoustics, 2009, 70(3):416-425. <https://www.sciencedirect.com/science/article/pii/S0003682X08001382>. DOI: <https://doi.org/10.1016/j.apacoust.2008.06.003>.
- [68] 雷英杰, 张善文. MATLAB 遗传算法工具箱及应用[M]. MATLAB 遗传算法工具箱及应用, 2014.
- [69] YANG Y, JIA H, CAO H, et al. Absorption performance of nonplanar periodic structures solved by layered rigorous coupled-wave analysis[J/OL]. Mechanical Systems and Signal Processing, 2022, 168:108702[2022-02-19]. <https://www.sciencedirect.com/science/article/pii/S0888327021010232>. DOI: [10.1016/j.ymssp.2021.108702](https://doi.org/10.1016/j.ymssp.2021.108702).
- [70] MILTON G W, CHERKAEV A V. Which Elasticity Tensors are Realizable?[J]. Journal of Engineering Materials & Technology, 1995, 117(4):483-493.
- [71] NORRIS A N. Acoustic cloaking theory[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2008.
- [72] SCANDRETT C L, BOISVERT J E, HOWARTH T R. Acoustic cloaking using layered pentamode materials[J]. The Journal of the Acoustical Society of America, 2010, 127(5):2856.
- [73] YI C, ZHENG M, LIU X, et al. Broadband solid cloak for underwater acoustics[J]. Phys.rev.b, 2017, 95(18):180104.1-180104.5.

- [74] Su, Xiaoshi, Cushing, et al. Broadband focusing of underwater sound using a transparent pentamode lens[J]. Journal of the Acoustical Society of America, 2017.
- [75] HLADKY-HENNION A C, VASSEUR J O, HAW G, et al. Negative refraction of acoustic waves using a foam-like metallic structure[J]. Applied Physics Letters, 2013, 102(14):-.
- [76] SUN Z, SHI Y, SUN X, et al. Underwater Acoustic Multiplexing Communication by Pentamode Metasurface[J]. 2020.
- [77] CAI L, WEN J H, YU D L, et al. Design of the Coordinate Transformation Function for Cylindrical Acoustic Cloaks with a Quantity of Discrete Layers[J/OL]. Chinese Physics Letters, 2014, 31(9):094303. <https://doi.org/10.1088/0256-307x/31/9/094303>.
- [78] SCANDRETT C L, BOISVERT J E, HOWARTH T R. Broadband optimization of a pentamode-layered spherical acoustic waveguide[J]. Wave Motion, 2011, 48(6):505-514.
- [79] 张向东, 陈虹, 王磊, 等. 圆柱形分层五模材料声学隐身衣的理论与数值分析[J/OL]. 物理学报, 2015, 64(13):134303. <https://wulixb.iphy.ac.cn/cn/article/id/64604>.
- [80] YI C, XIAONING L, PING X, et al. Pentamode material for underwater acoustic wave control [J/OL]. Advances in Mechanics, 2016, 46(1):201609. <https://lxjz.cstam.org.cn/en/article/id/145825>.
- [81] LONG H, CHENG Y, LIU X. Reconfigurable sound anomalous absorptions in transparent waveguide with modularized multi-order Helmholtz resonator[J/OL]. Scientific Reports, 2018, 8(1):15678. <https://doi.org/10.1038/s41598-018-34117-z>.
- [82] LONG H, LIU C, SHAO C, et al. Tunable and broadband asymmetric sound absorptions with coupling of acoustic bright and dark modes[J/OL]. Journal of Sound and Vibration, 2020, 479:115371. <https://www.sciencedirect.com/science/article/pii/S0022460X20302029>. DOI: <https://doi.org/10.1016/j.jsv.2020.115371>.
- [83] BERNSTEIN D N. The number of roots of a system of equations[J/OL]. Functional Analysis and Its Applications, 1975, 9(3):183-185. <https://doi.org/10.1007/BF01075595>.
- [84] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [85] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]//ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37. JMLR.org, 2015: 448-456.
- [86] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. J. Mach. Learn. Res., 2014, 15(1):1929-1958.
- [87] PEARSON K. LIII. On lines and planes of closest fit to systems of points in space[J/OL]. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1901, 2(11): 559-572. <https://doi.org/10.1080/14786440109462720>.

- [88] LIAQUAT M U, MUNAWAR H S, RAHMAN A, et al. Localization of Sound Sources: A Systematic Review[J/OL]. Energies, 2021, 14(13). <https://www.mdpi.com/1996-1073/14/13/3910>. DOI: [10.3390/en14133910](https://doi.org/10.3390/en14133910).
- [89] VAN VEEN B, BUCKLEY K. Beamforming: a versatile approach to spatial filtering[J/OL]. IEEE ASSP Magazine, 1988, 5(2):4-24. DOI: [10.1109/53.665](https://doi.org/10.1109/53.665).
- [90] Pinna-based spectral cues for sound localization in cat[J/OL]. Hearing Research, 1992, 58(2):132-152[2022-03-14]. <https://sciencedirect.53yu.com/science/article/abs/pii/0378595592901235>. DOI: [10.1016/0378-5955\(92\)90123-5](https://doi.org/10.1016/0378-5955(92)90123-5).
- [91] Aytekin M, Grassi E, Sahota M, et al. The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation[J/OL]. Acoustical Society of America Journal, 2004, 116(6):3594-3605. DOI: [10.1121/1.1811412](https://doi.org/10.1121/1.1811412).
- [92] SAXENA A, NG A Y. Learning sound location from a single microphone[C/OL]//2009 IEEE International Conference on Robotics and Automation. 2009: 1737-1742. DOI: [10.1109/ROBOT.2009.5152861](https://doi.org/10.1109/ROBOT.2009.5152861).
- [93] PATI Y, REZAIIFAR R, KRISHNAPRASAD P. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition[C/OL]//Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. 1993: 40-44 vol.1. DOI: [10.1109/ACSSC.1993.342465](https://doi.org/10.1109/ACSSC.1993.342465).
- [94] TROPP J A, GILBERT A C. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit[J/OL]. IEEE Transactions on Information Theory, 2007, 53(12):4655-4666. DOI: [10.1109/TIT.2007.909108](https://doi.org/10.1109/TIT.2007.909108).
- [95] WARDEN P. Speech Commands: A Dataset for Limited-Vocabulary Speech Recognition [J/OL]. arXiv:1804.03209 [cs], 2018[2022-03-14]. <http://arxiv.org/abs/1804.03209>.

致 谢

纵使不舍，也终究到了要写下致谢二字之时。提笔之时，我漫长的学生时代也终于走到了结尾。求学二十余载，无数个被孤独和无力感侵袭的时刻，承蒙诸位良师益友不弃，才有了跌跌撞撞行至今日的我。

经师易遇，人师难遭。由衷地感谢我的博士生导师杨军老师。杨老师不仅是我学术上的导师，更是我生活中的引路人。杨老师经常会在工作的间隙出现在 411 与大家谈天说地，兴致勃勃地与我们分享他周遭的趣事，用他乐观开朗的生活态度感染着每一个人。热爱可抵漫长岁月，惟愿自己也能如杨老师般永远热爱这个世界。

感谢我的执行导师贾晗老师。回首这五年，有数不清的各执一词、不肯让步的争吵，也有无数个心照不宣、并肩战斗的瞬间。感谢贾老师包容我所有不切实际的想法，也感谢他一直以来的信任和支持，给了我失败后仍能重新出发的勇气。

感谢我的父母。你们是默默站在我身后无条件支持我所有决定的家人，也是我在这个世界上最早相识的朋友。儿时你们总是护在我身前为我披荆斩棘，未来请让我为你们遮风挡雨。

感谢我读博路上所有的战友们。感谢 meta-team 的毕亚峰师兄，孙兆永师兄，陆文佳师姐，杨玉真师姐，王真师姐，吴吉恩师兄，赵晗师兄，王苏豪师兄，张鹏，赵准，石玉，匡秋林，周萍，杨星航，喻宇驰，杨芸翰，施全权，林媛燕，张艺檬。私以为 meta-team 是整个声学所最有爱的课题组，欢声笑语总能冲散一切的苦闷与烦恼，与你们并肩战斗是我的荣幸。感谢美丽可爱温柔大方不仅瘦还很高的胖今，不知道为什么我们总是有说不完的话，发不完的神经，我们要一直做没有血缘关系的亲姐妹。感谢 3-141 寝室的禹蒲，伊伊，邢博，怡宁，媛爸，晓晓，翟，婉雪，雨薇，祝你们一直都是怎么也吃不胖的小仙女。感谢 dsp411 的所有同学们，夏天没有空调的闷热，网吧服务器的轰鸣，头顶突然就熄灭的灯泡... 这些恼人的小插曲不知道会不会成为我们今后不愿忘却的回忆。感谢美食博主诗洁，攒局小能手包哥，美貌与财富并存的焦焦，在另一座城市默默关心我的大泽，还有所有北京小团体的兄弟姐妹们，是你们让这座城市不再冰冷，也因为你

们我从未想过离开。谢谢，我的朋友们，感谢你们来到我的世界。少年人的肩上应有望不尽的清风明月和草长莺飞，愿我们都能拥有一个淋漓尽致的人生。

最后也想感谢一下自己。感谢你五年前的勇敢，也感谢你一次次地战胜心中的迷茫、委屈和软弱，在最后关头没有落荒而逃。这里也要感谢素未谋面的周奇墨老师，毛东老师，大壹老师，梁文道老师，窦文涛老师，老蒋，玫瑰叔...长夜难眠，幸得你们的陪伴。曾经的我对于结果总是执着而热切的，想要风光的学位，想要瞬间的博学，想要意气风发地走向所有人。而当我们拉长时间的维度，知道一件事为什么做不成可能比做成这件事更有意义。我们要耐得住不断延长的时间线，付出足够多的努力，命运才肯将我们想要的一切一点一点地递送到我们手中。就像森林中总有一座迷路了才能找到的城堡，旖旎的风光也总在与我们不期而遇的路上。

小时候很喜欢看武侠小说，幻想自己有朝一日也能成为一个一呼百应的大英雄，长大后发现似乎所有的英雄主义都源于心底的那份对抗现实的天真与浪漫。在象牙塔里关了这么久，终究还是要出来闯荡一番。更好的故事总是尚未发生，我也早已开始期待下一段旅程。奔走红尘，莫忘曾经是书生。

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历：

孙雪聪，女，汉族，1995年6月出生于河北省秦皇岛市。2013年9月至2017年6月，就读于南京大学物理学院声学专业，获理学学士学位；2017年9月至2022年6月，就读于中国科学院声学研究所噪声与振动重点实验室信号与信息处理专业，攻读工学博士学位。

期刊论文：

1. SUN X, JIA H, ZHANG Z, YANG Y, SUN Z, YANG J. Sound Localization and Separation in 3D Space Using a Single Microphone with a Metamaterial Enclosure[J]. *Advanced Science*, 2020, 7(3): 1902271. (SCI 1 区, IF = 16.8, 封面文章) .
2. SUN X, JIA H, YANG Y, ZHAO H, BI Y, SUN Z, YANG J. Acoustic Structure Inverse Design and Optimization Using Deep Learning[J]. *Mechanical Systems and Signal Processing*, 2022. (SCI 1 区, IF=6.82, 审稿中)
3. SUN X, YANG Y, JIA H, YANG J. Physics-aware Training for the Physical Machine Learning Model Building[J]. *The Innovation*, 2022. (SCI, 审稿中)
4. 孙雪聪, 孙兆永, 贾晗, 等. 基于遗传算法的五模材料分层优化 [J]. 应用声学, 2022. (中文核心, 已接收)
5. 孙雪聪, 贾晗, 杨玉真, 等. 基于深度学习的低频宽带隔声器件设计 [J]. 应用声学, 2022. (中文核心, 已接收)
6. YANG Y, JIA H, CAO H, SUN X, ZHAO H, BI Y, YANG J. Absorption performance of nonplanar periodic structures solved by layered rigorous coupled-wave analysis[J]. *Mechanical Systems and Signal Processing*, 2022, 168: 108702. (SCI 1 区, IF = 6.82)
7. SUN Z, SUN X, JIA H, BI Y, YANG J. Quasi-isotropic underwater acoustic carpet cloak based on latticed pentamode metafluid[J]. *Applied Physics Letters*, 2019, 114(9): 094101. (SCI 2 区, IF = 3.79)

8. SUN Z, SHI Y, SUN X, JIA H, JIN Z, DENG K, YANG J. Underwater Acoustic Multiplexing Communication by Pentamode Metasurface[J]. *Journal of Physics D: Applied Physics*, 2021, 54(20):205303. (SCI 2 区, IF = 3.21)
9. 赵晗, 贾晗, 孙雪聪, 等. 基于薄膜编码超表面的宽频超薄声散射体 [J]. 应用声学, 2022. (中文核心, 已接收)

学术会议:

1. 孙雪聪, 孙兆永, 毕亚峰, 等. 基于遗传算法的五模材料分层优化 [C]//第六届装备振动与噪声控制青年论坛. 2018
2. 孙雪聪. 基于声学超材料的单通道多声源定位与分离系统 [C]//中国声学学会声频技术交流峰会. 2019.

已授权专利:

1. 孙雪聪, 贾晗, 杨军. 基于声学超材料的单通道多声源定位与分离装置及方法 [P]. 北京市: CN110703200B, 2021-07-30.
2. 孙雪聪, 贾晗, 杨玉真, 毕亚峰, 杨军. 一种基于深度学习的二阶亥姆霍兹共鸣器设计方法 [P]. 北京市: CN112926159B, 2022-02-22.

获奖情况:

2020 年获博士研究生国家奖学金、中国科学院大学三好学生; 2021 年获中国科学院大学三好学生标兵; 2022 年获中科院声学所第三届科普作品创意大赛三等奖。