



中国科学院大学

University of Chinese Academy of Sciences

## 博士学位论文

远距离激光相干语音信号侦测技术研究

作者姓名: 吕 韬

指导教师: 郭 劲 研究员

中国科学院长春光学精密机械与物理研究所

学位类别: 工学博士

学科专业: 光学工程

培养单位: 中国科学院长春光学精密机械与物理研究所

2019 年 6 月

**Research on the Remote Laser Coherent Speech Signal**

**Detection Technology**

**A dissertation submitted to  
University of Chinese Academy of Sciences  
in partial fulfillment of the requirement  
for the degree of  
Doctor of Engineering  
in Optical Engineering**

**By**

**Tao LV**

**Supervisor: Professor Jin Guo**

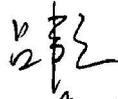
**Changchun Institute of Optics, Fine Mechanics and Physics,  
Chinese Academy of Sciences**

**June 2019**

## 中国科学院大学

### 研究生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。

作者签名：   
日期： 2019.05.31

## 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有关保存和使用学位论文的规定，即中国科学院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迟期后适用本声明。

作者签名：   
日期： 2019.05.31

导师签名：   
日期： 2019.05.31

## 摘要

与传统的语音侦测技术相比,激光语音侦测可在远距离、非接触的情况下侦测语音,具有隐蔽性强、操作便利等优点,因此被广泛应用于军事、安全、监控等领域。常见的激光语音侦测技术(例如光反射调制测量技术、散斑图像测量技术)存在环境适用性差、测量设备装调复杂等缺点,不能满足激光语音侦测的实际应用需求。激光相干多普勒测振技术具有作用距离远、精度高、非接触、空间分辨率高以及动态范围宽等优点,目前已经成为激光语音侦测领域的主要侦测手段。国外相关系统已实现 300 米的语音侦测,国内也开展了一系列相关研究,然而,与国外相比,无论是作用距离(相同激光功率下)抑或是获取语音的质量仍存在一定差距。国内要实现远距离、高质量语音侦测,无论是光路结构设计、测量目标的选择以及语音信号处理等方面都需要进行更深入的研究。因此,本文主要围绕上述三个方面,采用理论分析、数值仿真和实验研究相结合的方法,对远距离激光相干语音信号侦测技术进行了深入的研究。本文主要内容如下:

(1) 介绍了激光语音侦测技术在国内外的发展现状,分析并比较了目前常见的激光语音侦测方法的特点,提出了将激光相干多普勒测振技术应用于激光语音侦测系统中,回顾了激光相干多普勒测振技术的发展历程并介绍了国内外的研究现状,总结了国内发展激光相干语音侦测技术需要解决的技术难点,这为后续深入研究提供指导。

(2) 阐述了基于激光相干多普勒测振技术的语音信号侦测原理。通过理论计算给出了激光相干语音侦测系统的载噪比、探测分辨率和探测概率等评估系统性能指标的数学表达式。此外,通过理论推导和数值分析,研究了像差、激光器谱线展宽、探测方式以及发射光串扰等光学因素对激光相干语音侦测系统性能的影响。最后,在全光纤激光相干语音侦测系统光路结构基础上,针对光纤环形器隔离度不够的问题,设计并搭建了部分光纤光路,并根据理论研究结果,确定了系统的光学元器件。

(3) 激光相干语音侦测的本质是检测由声压引起的物体振动,因此被测物体选择的情况直接影响了侦测语音的质量。本文根据 STL 理论和 SSC 模型建立了不同材质及形状的平板在声场的振动模型模拟物体在声场的振动幅度,并通过

实验验证了分析的准确性。在激光语音侦测过程中,声源附近的物体相较于光波长而言是粗糙的。因此,本文采用数值仿真与实验相结合的方法讨论了不同粗糙表面对激光相干语音侦测系统载噪比的影响。当粗糙面相对于语音侦测系统存在除光轴方向之外的运动时,会给语音侦测过程引入散斑噪声。本文根据散斑特性讨论了散斑噪声对侦测语音的影响,并进行了实验验证,此外,根据理论计算得出了散斑噪声的功率谱密度。以上研究内容可为侦测过程中选择合适的测量目标提供指导。

(4) 系统地介绍了几种常见的解调方法,针对激光侦听系统需要同时保证信号处理的精度与实时性的要求,采用了正交解调与反正切鉴相结合的解调算法重构语音信号。由于环境噪声、物体粗糙面等因素,重构语音会受到噪声的干扰,因此需要降噪处理。根据噪声在相干侦测语音中的表现形式可分为背景噪声和冲击噪声。针对背景噪声,本文采用最优对数谱估计算法(Optimal Log-Spectral, OM-LSA)与相位补偿相结合的方法降噪;针对冲击噪声,本文首先使用峰度值比较法(kurtosis ratio)确定噪声在语音中的位置,再使用线性预测法估计出被冲击噪声污染的语音数据。此外,由于被测物体对于声场的振动响应会导致相干侦测语音带宽较窄,因此本文使用基于半波整流的语音带宽扩展算法扩展相干侦测语音带宽。

(5) 开展了室内 30m、70m、120m 距离,振动测量目标为文件夹(塑料制)、饮料瓶(纸制)以及电脑显示屏的语音侦测实验,实验结果表明自研系统可有效获取远距离(至少 120m)可理解语音信号。在充分证明自研侦测系统性能的基础上,分别进行了双模监控实验和远距离说话人识别实验,这两个实验结果分别表明了自研语音侦测系统在安防监控领域和生物识别领域应用的可行性和潜力。

综上,本论文通过对远距离激光相干语音信号侦测关键技术的理论与实验研究,逐步解决了各项关键难点,为后续实现更远距离更高质量的语音侦测奠定基础。

**关键词:** 激光语音侦测, 激光相干多普勒测振, 粗糙表面, 散斑噪声, 语音降噪, 带宽扩展

## Abstract

Compared with traditional speech detection technology, laser speech detection can detect speech signal in long-distance and non-contact situations, and has the advantages of strong concealment and convenient operation. Therefore, it is widely used in military, security, surveillance and other fields. Common laser speech detection technologies (such as light reflection modulation measurement technology and speckle image measurement technology) have shortcomings such as poor environmental applicability and complicated adjustment of measurement equipment, which cannot meet the practical application requirements of laser speech detection. Laser coherent Doppler vibration measurement technology has the advantages of long distance, high precision, non-contact, high spatial resolution and wide dynamic range. It has become the main detection method in the field of laser speech detection. Foreign related systems have achieved 300 meters of speech detection, meanwhile a series of related researches have been carried out in China. However, compared with foreign countries, there is still a certain gap in terms of the distance of operation (under the same laser power) and the quality of speech signal acquisition. In order to achieve long-distance, high-quality speech detection in China, more in-depth researches are needed in terms of optical path structure design, measurement target selection, and speech signal processing. Therefore, this paper focuses on the above three aspects, using theoretical analysis, numerical simulation and experimental research combined method to carry out in-depth research on long-range laser coherent speech signal detection technology. The main contents of this paper are as follows:

(1) This paper introduces the development status of laser speech detection technology at home and abroad. Characteristics of common laser speech detection methods are analyzed and compared. Additionally, the method about applying the laser coherent Doppler vibration measurement technology into laser speech detection system is proposed. The development history of laser coherent Doppler vibration measurement technology is reviewed and the research status at home and abroad is introduced. Besides, the technical difficulties that need to be solved in the development of laser coherent speech detection technology in China are summarized, which provides guidance for further in-depth research.

(2) The principle of speech signal detection based on laser coherent Doppler

vibration measurement technology is expounded. Mathematical expressions for assessment criteria which are used to evaluate the performance of the system such as the carrier-to-noise ratio, minimum resolution and detection probability of the laser coherent speech detection system are given by theoretical calculation. In addition, through theoretical derivation and numerical analysis, the effects of optical factors such as aberration, laser line width, detection mode and emission crosstalk on the performance of laser coherent speech detection system are studied. Finally, based on the optical path structure of the all-fiber laser coherent voice detection system, partial-fiber structures are designed and built to solve the problem of insufficient isolation of the fiber circulator. In addition, the optical components of the system are determined based on the results of the theoretical research.

(3) The essence of laser coherent speech detection is to detect the vibration of an object caused by sound pressure. Therefore, the selection of the object to be measured directly affects the quality of the detected speech signal. According to the STL theory and the SSC model, the vibration model of the flat plates with different materials and shapes in the sound field is used to simulate the vibration amplitude of the object in the sound field, and the accuracy of the analysis is verified by experiments. During laser speech detection, objects near the sound source are rough compared to the wavelength of the light. Therefore, this paper discusses the influence of different rough surfaces on the carrier-to-noise ratio of laser coherent speech detection system using the combination of numerical simulation and experiments. When the rough surface has motion other than the optical axis direction relative to the speech detection system, speckle noise will be introduced to the speech detection process. In this paper, the influence of speckle noise on the detected speech is discussed based on the speckle characteristics, and the experimental verification is carried out. In addition, the power spectral density of speckle noise is calculated theoretically. The above research can provide guidance for selecting suitable measurement targets during the detection process.

(4) Several common demodulation methods are introduced systematically. Aiming at the requirement that the laser interception system needs to ensure the accuracy and real-time during the signal processing at the same time, the demodulation algorithm combined with orthogonal demodulation and arctangent detection is used to reconstruct the speech signal. Due to environmental noise, rough surfaces of objects, etc., reconstructed speech is subject to noise interference.

Therefore, the noise reduction processing is required. According to the representation of noise in coherently detected speech signal, the noise can be divided into background noise and impact noise. For background noise, this paper uses the method combined with Optimal Logarithmic Spectrum estimation algorithm (OM-LSA) and phase compensation to reduce noise; For the impact noise, this paper first uses the kurtosis ratio to determine the position of the noise in the speech signal, and then uses the linear prediction method to estimate the speech data contaminated by the impact noise. In addition, since the vibration response of the measured object to the sound field will cause the bandwidth of the coherent detection speech to be narrow, this paper uses the half-wave rectification-based speech bandwidth extension algorithm to extend the coherent detection speech bandwidth.

(5) Speech detection experiments were conducted and the indoor distances are 30m, 70m, 120m, while the vibration measurement targets are folder (plastic), beverage bottles (paper) and computer display. The experimental results show that the self-developed system can effectively acquire long-distance (at least 120m) understandable speech signals. On the basis of fully demonstrating the performance of the self-research detection system, the dual-mode monitoring experiment and the long-distance speaker recognition experiment were carried out respectively. The results of the two kinds of experiments show the feasibility and potential of the self-developed speech detection system in the field of security monitoring and biometric identification.

In summary, this paper has gradually solved various key difficulties through the theoretical and experimental research on the key technologies of long-range laser coherent speech signal detection, and laid the foundation for the subsequent realization of longer-distance and higher-quality speech detection.

**Key Words:** Laser Speech Detection, Rough Surface, Speckle Noise, Speech Noise Reduction, Bandwidth Expansion, Laser Coherent Doppler Vibration Measurement

## 目 录

摘 要.....	I
Abstract.....	III
<b>第 1 章 绪论 .....</b>	<b>1</b>
1.1 激光语音侦测研究背景、意义及国内外现状.....	1
1.2 激光语音侦测技术综述.....	4
1.2.1 光反射调制测量技术.....	4
1.2.2 散斑图像测量技术.....	5
1.2.3 激光相干多普勒测振技术.....	6
1.3 激光相干多普勒测振技术研究进展及国内外现状.....	7
1.3.1 激光相干多普勒测振技术研究进展.....	7
1.3.2 国内外研究现状.....	12
1.4 论文研究思路和结构安排.....	15
<b>第 2 章 激光相干语音信号侦测理论研究和光路结构设计 .....</b>	<b>19</b>
2.1 激光相干语音侦测原理.....	19
2.1.1 激光相干语音侦测实现原理.....	19
2.1.2 激光相干多普勒测振原理.....	20
2.2 激光相干语音侦测性能分析.....	22
2.2.1 探测分辨率.....	22
2.2.2 探测概率.....	24
2.2.3 系统载噪比.....	28
2.3 激光相干语音侦测性能光学影响因素分析.....	31
2.3.1 像差对激光相干语音侦测系统的影响.....	31
2.3.2 激光器谱线展宽对激光相干语音侦测系统的影响.....	34
2.3.3 探测器探测方式对激光相干语音侦测系统的影响.....	38
2.3.4 发射光串扰对激光相干语音侦测系统的影响.....	40
2.4 激光相干语音侦测系统光路结构设计.....	42
2.4.1 全光纤激光相干语音侦测系统光路结构.....	42
2.4.2 部分光纤激光相干语音侦测系统光路结构.....	44
2.4.3 空间光与光纤高效耦合.....	44

2.4.4	全光纤结构与部分光纤结构性能比较实验.....	47
2.5	本章小结.....	50
<b>第3章</b>	<b>振动测量目标的选择 .....</b>	<b>51</b>
3.1	物体表面在声场的振动特性.....	51
3.1.1	物体表面在声场的振动模型.....	51
3.1.2	仿真分析与讨论.....	54
3.1.3	实验分析与讨论.....	57
3.2	物体表面粗糙度对激光相干语音侦测系统的影响.....	59
3.2.1	理论分析粗糙面对载噪比的影响.....	60
3.2.2	粗糙面对载噪比的影响的数值仿真.....	62
3.2.3	实验研究粗糙面对载噪比的影响.....	66
3.3	散斑噪声对激光相干语音侦测系统的影响.....	68
3.3.1	散斑特性.....	70
3.3.2	动态散斑特性.....	74
3.3.3	散斑噪声对激光相干语音侦测系统的影响.....	80
3.4	本章小结.....	85
<b>第4章</b>	<b>激光相干语音侦测系统信号处理 .....</b>	<b>87</b>
4.1	语音质量评价指标.....	87
4.1.1	主观评价.....	87
4.1.2	客观评价.....	88
4.1.3	本文采用的评价标准.....	90
4.2	相干侦测语音信号重构.....	90
4.2.1	传统载波解调技术.....	90
4.2.2	正交解调与反正切鉴相结合的解调算法.....	91
4.2.3	重构算法实验分析.....	93
4.3	相干侦测语音信号降噪 .....	94
4.3.1	背景噪声抑制算法研究.....	96
4.3.2	冲击噪声抑制算法研究.....	104
4.4	语音信号带宽扩展 .....	109
4.4.1	传统语音信号带宽扩展算法.....	109
4.4.2	基于半波整流的语音带宽扩展算法.....	111
4.5	本章小结.....	113

---

<b>第 5 章 激光相干语音侦测实验研究 .....</b>	<b>115</b>
5.1 语音侦测实验研究.....	115
5.2 双模监控实验研究.....	121
5.2.1 视频检测.....	121
5.2.2 双模监控系统工作方式.....	123
5.2.3 双模监控实验结果及分析.....	125
5.3 远距离说话人识别实验研究.....	127
5.3.1 训练集与测试集建立.....	127
5.3.2 语音特征提取.....	128
5.3.3 基于 GMM-UBM 的说话人识别模型.....	130
5.3.4 远距离说话人识别实验结果及分析.....	132
5.4 本章小结.....	134
<b>第 6 章 总结与展望 .....</b>	<b>135</b>
6.1 论文工作总结.....	135
6.2 研究展望.....	136
<b>参考文献.....</b>	<b>139</b>
<b>作者攻读学位期间发表的学术论文与研究成果.....</b>	<b>151</b>
<b>指导教师及作者简介 .....</b>	<b>153</b>
<b>致谢.....</b>	<b>155</b>

## 第 1 章 绪论

### 1.1 激光语音侦测研究背景、意义及国内外现状

由于语音信号是目前人类相互交流的最重要的方式,是信息传递的一个重要载体,因此语音侦测技术被广泛应用于安防监控、公共安全、军事等领域,例如入侵检测<sup>[1]</sup>,银行、地铁、机场及电梯等公共场所的异常检测<sup>[2~3]</sup>,反恐,缉毒等方面。传统的语音侦测技术基本都使用声音传感器获取语音信号。一般的声音传感器的工作原理是:由于其内置的电容式驻极体话筒的薄膜容易受到声压的影响而发生振动,这种振动会导致其电容产生变化,进而产生电压,通过检测声音传感器电压的变化便可获取语音信号。然而,由于声压在空气中传播时会快速下降,因此传统的语音侦测技术工作距离十分有限,往往需要近距离(距离声源一般为米级)放置侦听器材。若侦测区域被高度戒备,侦测人员便很难进入侦测区域,放置的侦听器材也极易被发现,此时传统的语音侦测方法便很难获取有效的语音信号。跟一般的声音传感器相比,抛物面麦克风虽然可以被用于侦测较远距离语音信号,但是其对环境噪声敏感<sup>[7]</sup>,并且作用距离仍然十分有限。激光语音侦测技术是一种较为新颖的语音侦测技术,相较于传统的语音侦测技术而言,激光语音侦测技术获取的语音信号是通过测量声源附近物体的振动信号所获得的,因此对环境噪声的敏感性较低,仅会受靠近声源处噪声的影响。另一方面,由于激光的特性,使得激光侦听技术可以在远距离,非接触的情况下侦测语音信号,无需近距离放置侦听器材,并且激光光源可以在红外范围内选择,这保证了激光语音侦测技术的隐蔽性。

随着激光器和激光测量技术的发展,激光语音侦测技术已经进入实用阶段。早在 1985 年,Modern Electronics and Creative Computing 期刊的特约编辑 FORREST MIMS 便对激光语音侦测系统的光源、接收器、光学结构及如何防范激光侦听等方面进行了详细的介绍<sup>[4]</sup>。海湾战争期间,美国便使用激光侦听设备侦听伊拉克高级将领车内的谈话<sup>[5]</sup>。此外,在 1997 年就有报道称英国警方使用激光侦听设备打击犯罪<sup>[6]</sup>。激光侦听设备除了广泛应用于军事、公共安全等领域,另一个典型的应用领域为安防监控。Weihong Li、Yufu Qu、Tao Wang 以及 Zhigang Zhu 等人<sup>[7~11]</sup>将德国 Polytec 公司生产的激光相干多普勒测振仪(Laser Doppler

vibrometer, LDV) 改造成激光侦听系统, 并结合 PTZ (pan/tilt/zoom) 相机与红外相机, 搭建了远距离多模监视系统, 其最远侦听距离可达 300 米(激光器功率为毫瓦级别)。图 1.1 为该系统的示意图, 图 1.2 为该系统距离目标 420 英尺时获取的视频与语音信息。除了这些实际应用之外, Rui Li、Tao Wang 等人<sup>[12]</sup>讨论了物体振动特性对激光语音侦测的影响。Rui Li、Nicholas Madampoulos 等人<sup>[13]</sup>研究了激光器相干长度对远距离激光语音侦测的影响。除此之外, Avargel<sup>[14]</sup>将激光侦听系统用于语音增强。Deng<sup>[15]</sup>利用激光侦听系统获取的语音信号进行远距离说话人识别等。

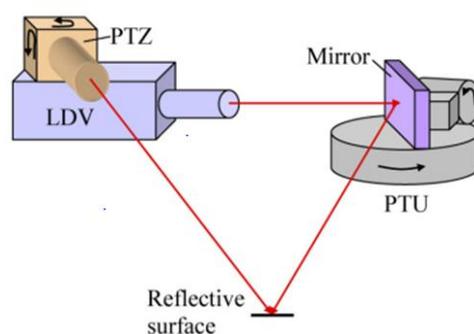


图 1.1 多模监控系统示意图

Figure 1.1 Schematic setup of the multimodal model detection system

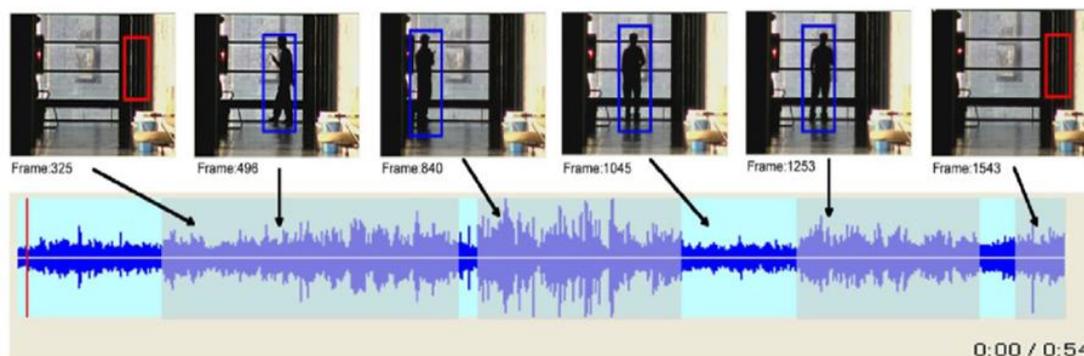


图 1.2 在 420 英尺的走廊中获取的视频与语音信息

Figure 1.2 Audio-video integration in a 420-foot corridor

与国外相比, 国内虽然对于激光语音侦测技术的研究起步较晚, 但是目前已经取得了一定的研究成果。例如, 2003 年, 华中科大的罗海俊和朱晓<sup>[16]</sup>对激光语音侦测技术的基本原理进行了讨论, 并进行了相关实验, 实验结果验证了激光语音侦测技术的可行性。在同一年, 国防科大的康宏向<sup>[17]</sup>针对激光侦听过程中遇

到的大气噪声和环境噪声的问题,采用激光光束强度调制、线性相位 FIR 滤波器以及时域延长相结合的方法来降低大气噪声和环境噪声的干扰。2007 年,基于 DSP 平台,华中科大的田玉敏<sup>[18]</sup>使用自适应变步长最小均方算法对激光侦测语音进行降噪。2008 年,国防科大的张超凡<sup>[19]</sup>利用“猫眼”效应对传统的激光语音侦测技术进行了改进,实现了激光语音侦测系统中发射系统与接收系统的相互结合以及侦测目标是否瞄准的自动判断。2009 年,上海光机所的尚建华、贺岩等人<sup>[20]</sup>设计并搭建全光纤激光多普勒测振仪以获取远距离语音信号。实验结果验证了该系统可获取远距离语音信号(理论可达 40 米)。2011 年,北交大的张勤、张建勇<sup>[21]</sup>等人基于 PGC 解调技术,设计并搭建了激光多普勒语音检测系统。实验结果表明,其系统有效工作距离可达 20 米。2012 年,湛江师范的黄贞和吴林富<sup>[22]</sup>研制了一套便携式远距离激光侦听系统,该系统工作距离可达 50.7 米。2014 年,长春理工的刘艳阳、母一宁等人<sup>[23~24]</sup>基于四象限探测器搭建了一套激光侦听系统。同年,中科院半导体所的张宇泽、李丽艳等人<sup>[25]</sup>研究了目标材料特性对激光语音侦测的影响。2015 年公安部第一研究所报道了一种利用散斑图像实现激光侦听的技术<sup>[26]</sup>。2017 年,南京航天航空大学的张展<sup>[27]</sup>利用激光语音侦测系统进行耳语音识别。同年,解放军信息工程大学的张文璐<sup>[28]</sup>对激光语音侦测过程中遇到的噪声按照其特性进行了分类,并针对不同类型的噪声,提出了相对应的降噪方案。同年,合肥工业的丁宁<sup>[29]</sup>使用四象限探测器研制了一套声音探测系统。2018 年,电子科大汪啸<sup>[30]</sup>利用激光反射调制原理,设计并搭建了一套收发分离的激光振动测量仪,该系统可实现 200 米处的语音侦测(激光器功率为 2W)。2019 年,中科院电子所的白涛、吴谨等人<sup>[31]</sup>使用深度循环神经网络(deep recurrent neural network, DRNN)对激光侦测语音进行降噪。除了上述科研院所对激光侦听技术进行研究之外,一些企业也对该技术表现出极大的兴趣,例如,杭州爱莱达公司斥巨资研制出多款声音侦听激光雷达(如图 1.3 所示为 BAT-200 型激光侦听雷达)。



图 1.3 BAT-200 型激光侦听雷达

Figure 1.3 BAT-200 Laser listening radar

虽然国内关于激光语音侦测方面取得了相当程度的进步，然而与国外相关系统相比，在同等激光器功率条件下，无论是作用距离、获取语音的清晰度还是系统的稳定性尚存在一定差距。国内要实现远距离、高质量语音侦测存在一些关键性难点，这些难点主要包括：激光语音侦测系统的系统设计、测量目标的选择以及语音信号处理等。因此，本论文作者基于上述考虑，通过对激光语音侦测关键技术进行理论和实验研究，逐步解决各项关键难点，以实现远距离、高质量的语音侦测。

## 1.2 激光语音侦测技术综述

目前而言，常见的激光语音侦测方法可分为：光反射调制测量技术、散斑图像测量技术以及激光相干多普勒测振技术。下文介绍这三种常见的激光语音侦测方法。

### 1.2.1 光反射调制测量技术

使用光反射调制测量技术侦测语音的过程可概述为：假设激光以一定角度聚焦在物体表面，并在反射光方向上放置光电探测器。若此时物体附近的声源发出声音，由于声压的作用会使得物体行受迫振动，从而使得反射光偏离预先的方向，最终导致光电探测器接收到的光强发生变化。因此，可以根据光强的变化量测量出物体的振动位移，再根据振动位移重构出语音信号。

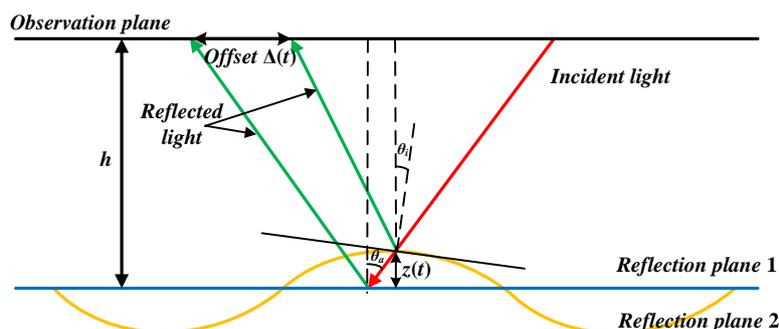


图 1.4 光反射调制原理图

Figure 1.4 Scheme of the light reflection modulation

如图 1.4 所示，反射面 1 为物体未产生振动，反射面 2 为物体产生振动， $h$  为观察平面距离反射面 1 的距离。假设未产生振动前，入射光与法线的夹角为  $\theta_a$ ，当物体开始振动时，振动幅度为  $z(t)$ ，并且反射面法线偏移了  $\theta_i$ 。此时，根据光的反射定律可知，在观察平面上，由于物体振动而引起的反射光偏移量  $\Delta(t)$  为：

$$\Delta(t) = [h + z(t)] \tan(\theta_a) - [h - z(t)] \tan(\theta_a - \theta_i) \quad (1-1)$$

一般而言，物体的振动幅度非常小，因此  $\theta_i$  要远小于  $\theta_a$ ，因此公式(1-1)可见简化为

$$\Delta(t) = 2z(t) \tan(\theta_a) \quad (1-2)$$

光反射调制测量技术具有结构简单、实现容易、技术较为成熟、研制成本低等优点，因此激光语音侦测系统大多采用该技术<sup>[4,16~19,22~24,27~30]</sup>。然而，在实际应用中，该技术具有两个较大的缺陷：第一，被测物体必须具有较强的反射率才能保证反射光足够强，这严重限制了激光语音侦测系统的应用场景。第二，激光发射器与探测器需放在特定的位置以保证反射光方向正对探测器，这给装调造成了困难。虽然国防科大的张超凡<sup>[19]</sup>利用猫眼效应对其进行了改进，将发射端与接收端组合在一起，这大大降低了装调难度，然而这个改进技术仍不成熟，还需要进一步研究。

### 1.2.2 散斑图像测量技术

散斑图像测量法是一种新颖的激光语音侦测技术<sup>[26]</sup>。如图 1.5 所示，将激光光束直接照射在被测物体上，由于被测物体表面的粗糙性使得回波光斑呈现明暗相间的散斑图样。当被测物体发生横向、轴向及斜向运动时，由于照射光斑与被

测物体之间的相对位置发生变化,从而导致回波光斑的散斑图样也发生改变。但是,当满足一定条件时(该条件详见第 3.3.2 节),可以使散斑图样的变化方式为仅改变其空间位置,不改变其大小和形状。因此通过光学成像技术,实时追踪回波散斑图样的时域轨迹,可获得被测物体的振动特征,从而重构出语音信号。然而该方法测量精度不高,且正处于研究阶段,不具备实际应用的条件。

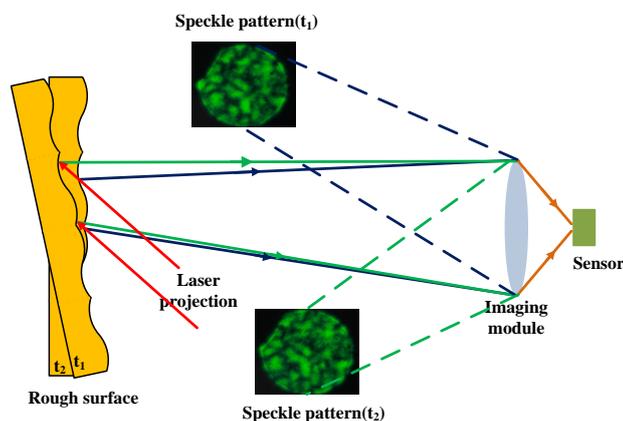


图 1.5 散斑图像测量法原理图

Figure 1.5 Scheme of the speckle image measurement

### 1.2.3 激光相干多普勒测振技术

使用激光相干多普勒测振技术侦测语音的过程可概述为:假设一束激光入射在物体表面,此时若物体受到的声压的作用而产生振动,这时由于物体的振动会导致回波光程发生变化,从而给回波光引入多普勒频率。由于光学频率太高,无法直接从回波光提取多普勒频率,而相干探测技术则可有效的提取回波光的多普勒频率。因此,使用相干探测技术提取回波光的多普勒频率便可测量出物体的振动特性,根据振动特性便可成功重构出语音信号。具体关于相干多普勒测振技术在激光语音侦测中的应用的介绍将在第二章进行详述。

激光相干多普勒测振技术相较于另外两种激光语音侦测技术而言,其测量的工作原理确保了检测的准确性,并且其探测灵敏度可比直接探测的方式高 7~8 个数量级<sup>[32]</sup>,这表明测量目标不仅可以选选择反射率较强的物体,还可以选选择常见的漫反射体作为测量目标。综合几种激光语音侦测技术的特点不难发现,激光相干多普勒测振技术更适合于激光语音侦测。因此,本文以激光相干多普勒测振技术作为测量方法,进行激光语音侦测的研究。

## 1.3 激光相干多普勒测振技术研究进展及国内外现状

### 1.3.1 激光相干多普勒测振技术研究进展

自从 Y. Yeh 和 H. Z. Cummins<sup>[32]</sup>于 1964 年发表了第一篇利用激光相干多普勒测量技术测量流速的论文以后,该技术就立即受到科学界及工业界的广泛关注。经过 50 余年不断的发展与完善,该技术逐渐走向成熟,其应用范围从最初的流体测速逐步扩展到固体测速和振动检测。与传统的振动检测方法(例如加速度仪)相比,激光相干多普勒测振技术具有作用距离远、精度高、非接触、空间分辨率高以及动态范围宽等优点,被广泛应用于军事<sup>[33~35]</sup>、安防<sup>[7~11]</sup>、工业<sup>[36~37]</sup>及生物医学<sup>[38~40]</sup>等领域,各国对该技术进行了大量的研究,并根据不同的应用需求,研制了不同结构的激光相干多普勒测振仪。

#### (1) 理论研究

激光相干多普勒测振技术是以激光的多普勒效应为基础,利用相干探测技术测量由被测物体振动而引起的回波光的多普勒频移,进而检测出物体的振动位移和速度,因此研究相干探测理论对于激光相干多普勒测振有着重要的意义。A.E. Siegma 等人<sup>[41]</sup>利用天线理论研究了失配角对相干探测性能的影响。Fink<sup>[42]</sup>根据信号光和本征光的光场分布函数以及探测器的形状和大小推导出来相干探测信噪比的一般方程。Steven C. Cohen 等人<sup>[43]</sup>研究了信号光与本振光波前失配、光斑尺寸以及光场分布对相干探测的影响。其研究结果表明,相干探测信噪比对于波前失配最敏感,并随着失配角的增加而下降;而相比于平面波而言,当回波波形为 Airy 型时,对于波前失配较不敏感。HongZhou Dong、RuoFu Yang 等人<sup>[44]</sup>从频域角度对相干探测进行了研究。其研究结果指出相干探测信号的电流幅值与探测器的量子效率函数的光谱值成正比,并且该值所对应的频率由光束入射角度决定,并且量子效率函数的光谱具有低通性,其截止频率由探测器的光敏面直径决定。该研究结论不仅适用于平面波,还可扩展到任意波。Holmes、Rask 等人<sup>[45]</sup>分析了探测器的饱和效应对于相干探测的影响,并根据光电二极管电流和光功率的响应曲线,制定了最佳本征光功率和最大可能信噪比的公式,并定义了有效的外差响应度。Jiang、Luu 等人<sup>[46]</sup>研究了弱本征光下的相干探测,提出来任意功率下漫反射目标或者镜面目标的相干检测理论。Salem、Rolland 等人<sup>[47]</sup>研究了部分相干光下相干探测的性能。并在假设光电探测器表面相干混合了两个部分相干的

高斯谢尔模光束时, 得出了外差效率的解析表达式。类似的, Chengqiang Li<sup>[48]</sup>研究了湍流环境下部分相干光的相干探测性能。Painchaud、Poulin 等人<sup>[49]</sup>研究了平衡探测对相干探测的影响。Gallion<sup>[50]</sup>及 Chengqiang Li<sup>[51]</sup>等人研究了激光器相位噪声对相干探测的影响。Chambers<sup>[52]</sup>利用赛德尔像差理论分析了像差对于相干探测的影响。Qi Hu<sup>[53]</sup>利用实验和仿真证明了球差会严重影响相干探测的性能。郭力仁<sup>[54]</sup>研究了声光移频器对基于微多普勒效应的探测的影响。

在实际振动测量中, 由于被测物体表面的粗糙性会给激光相干多普勒测振系统引入散斑噪声, 从而影响测量的准确性。针对这个问题, 进行了许多研究。Rahimi、Li 等人<sup>[55]</sup>搭建了一个试验台用来研究不同测量距离和速度下的散斑噪声。O'Malley、Woods 等人<sup>[56]</sup>进行了一系列的实验量化了被测物体表面粗糙度、激光聚焦以及测量间隔对于激光相干多普勒振动测量的影响。Rothberg<sup>[57]</sup>数值仿真了散斑噪声对于激光相干多普勒振动测量的影响。Pfister T、Fischer A 等人<sup>[58]</sup>研究了由于散斑噪声而引起的激光相干多普勒测量的极限不确定性。Rothberg、Halliwell 等人<sup>[59]</sup>讨论了散斑对多普勒信号的相位和幅度调制。Martin<sup>[60]</sup>研究了散斑造成的测量不确定性, 并通过实验量化了激光相干多普勒测振系统对于这种不确定性的灵敏度。Aranchuk、Sabatier 等人<sup>[34]</sup>研究了扫描式激光相干多普勒测振仪中散斑噪声的起源, 并提出了几种散斑消除方法。J. Vass 等人<sup>[61]</sup>通过峰度值比较法确定散斑噪声出现在测量信号中的位置。Ridley<sup>[62-63]</sup>, Hill<sup>[64]</sup>, Jiang<sup>[65]</sup> 以及 Dräbenstedt<sup>[66]</sup>等人通过理论或实验研究了由于散斑噪声而引起的噪声基底。

## (2) 系统结构研究

自从 20 世纪 80 年代第一台激光相干多普勒测振仪问世以来, 随着应用场景的日益多样化、复杂化, 对其系统结构的研究也不断深入。早期的激光相干多普勒测振仪基本是单点“轴向”振动测量仪, 只能用来测量沿激光视线方向上的单点振动分量。这类测振仪大多采用马赫-曾德尔干涉仪结构 (详见图 2.3)。如图 1.6 所示为德国 Polytec 公司生产的 OFV-5000 型激光相干多普勒测振仪, 该设备就是典型的单点“轴向”振动测量仪。



图 1.6 OFV-5000 型激光相干多普勒测振仪

Figure 1.6 OFV-5000 Laser Doppler vibrometer

由于单点“轴向”振动测量仪只能测量单点振动信息，为了扩大测量区域，弥补单点测量的不足，先后出现了扫描激光相干多普勒振动测量仪和多点激光相干多普勒振动测量仪。如图 1.7.a 所示，扫描激光相干多普勒振动测量仪是在单点“轴向”振动测量仪的基础上配备了扫描系统（一般由两个正交镜组成）。这些扫描激光相干多普勒振动测量仪通过改变扫描系统的偏转角度可快速精确的改变测量点的位置，从而实现面振动测量。此外，扫描激光相干多普勒振动测量仪还可以配备摄像机，将被测区域清晰的显示在电脑上，从而增强人机互动。如图 1.7.b 所示为德国 OptoMET 公司生产的扫描激光相干多普勒测振仪，该设备就是典型的扫描激光相干多普勒振动测量仪。除了使用机械装置进行扫描之外，Maru<sup>[67~68]</sup>等人使用频率可调的激光器和衍射光栅设计了可调谐激光器的垂轴扫描多普勒测量系统。该系统可通过改变激光器的频率改变激光通过衍射光栅时的偏转角，从而实现扫描测量。

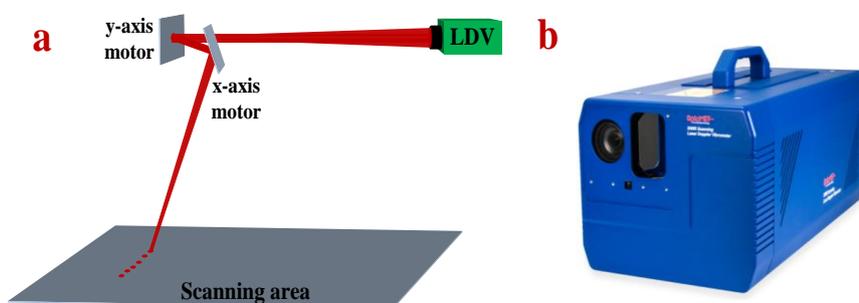


图 1.7 扫描式激光相干多普勒测振仪。(a)示意图。(b)实物图

Figure 1.7 Scanning Laser Doppler vibrometer. (a)Schematic diagram. (b)Physical map.

除了扫描激光相干多普勒振动测量仪，多点激光相干多普勒振动测量仪也可实现面振动测量。这类振动测量仪使用例如布拉格元件、多分束器及全息光学器件等光学器件将发射光由一束分为多束，由于这些发射光打在被测物体表面上的空间位置不同，因此可同时测量被测物体不同点的振动，从而实现面振动测量。然而，这类振动测量仪光路结构较复杂，很难大规模推广。图 1.8.a 为典型的多点激光相干多普勒振动测量仪结构图<sup>[65]</sup>，图 1.8.b 为德国 Polytec 公司生产的 MPV-800 型多点激光相干多普勒测振仪。

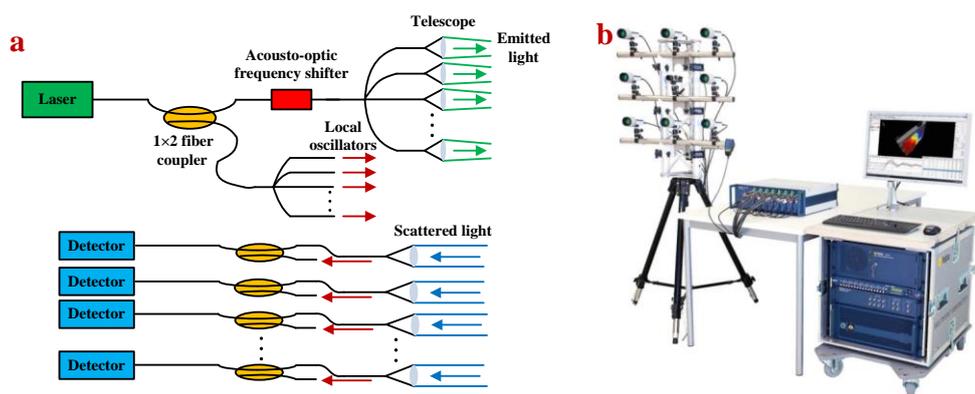


图 1.8 多点式激光相干多普勒测振仪 (a) 示意图。(b) MPV-800 型激光相干多普勒测振

Figure 1.8 Multi-point Laser Doppler vibrometer. (a) Schematic diagram.(b) MPV-800 Laser Doppler vibrometer.

无论是单点“轴向”结构的激光相干多普勒振动测量仪，还是扫描激光相干多普勒测振仪和多点激光相干多普勒测振仪都只能检测激光光轴方向上的振动，然而，有些物体不只沿着某个特定的方向振动，为了可以测量该点的全部振动特性，可使用 3D 激光相干多普勒测振仪。如图 1.9 所示，3D 激光相干多普勒测振仪是将三个各自独立的单点“轴向”振动测量仪组合在一起，通过同时测量被测物体三个分量上的振动特性从而实现 3D 振动测量。



图 1.9 PSV-500 型 3D 激光相干多普勒测振仪

Figure 1.9 PSV-500 3D Laser Doppler vibrometer

除此之外，自混频结构的激光相干多普勒测振仪凭借其低成本、结构紧凑等优点也受到广泛的关注。如图 1.10 为自混频激光相干多普勒测振仪的结构图<sup>[69]</sup>。与其他结构的测振仪相比，自混频测振仪除激光器和聚焦透镜之外，不需要附加任何光学器件，这是因为回波光与本征光在腔内发生混频。但是目前该技术尚未成熟，仍处于实验研究阶段。

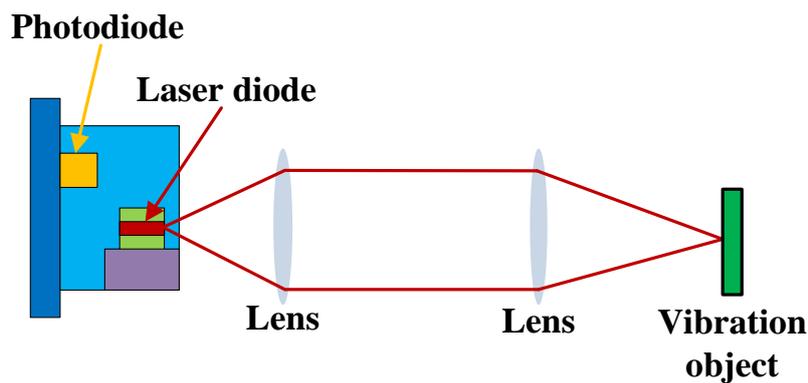


图 1.10 自混频激光相干多普勒测振仪结构图

Figure 1.10 Schematic of the self-mixing Laser Doppler vibrometer

随着激光器光源技术、光学器件技术及相关技术的不断发展,激光相干多普勒测振仪的结构也越来越多样。除了上述结构之外,还有全光纤声光结构<sup>[70]</sup>、双频自混频结构<sup>[71]</sup>、基于双频环形激光器的结构<sup>[72]</sup>、分集接收结构<sup>[73]</sup>等。

### 1.3.2 国内外研究现状

自 1983 年第一台激光相干多普勒测振仪问世以来,经过几十年的发展,激光相干多普勒振动测量技术已经相对成熟,世界许多国家都已经有了成熟的激光相干多普勒振动测量设备,例如德国、丹麦、日本、新加坡、英国等。如图 1.11 所示,为德国 Polytec 公司生产的 RoboVib 型结构测振站。该测振站将 3D 扫描式激光相干多普勒测振仪与工业机器人相结合,从而可完成从复杂部件到完整机器设备的一切测量工作,其中,采用的 3D 扫描式激光相干多普勒测振仪的频率测量范围为 DC-25MHz,振动速度测量范围为  $0.01\mu\text{m/s}$ -30m/s。



图 1.11 RoboVib 型结构测振站

Figure 1.11 The RoboVib Structural Test Station

图 1.12 为德国 OptoMET 公司生产的 Nova 系列激光相干多普勒测振仪,该产品的振动位移最佳分辨率可达  $2\text{pm}$ ,振动速度的最佳分辨率可达  $2.5\text{nm/s}$ 。该产品将光学结构和信号处理系统合二为一,这大大提高了产品的便携性。并且,该设备采用触摸屏和旋钮操作相结合,使得设备使用便捷。此外该设备可现场更换多款镜头,从而可满足多种不同的应用。



图 1.12 Nova 系列激光相干多普勒测振仪

Figure 1.12 Nova series Laser Doppler vibrometer

图 1.13 为新加坡舜宇仪器公司生产的 SLV-S02 型扫描激光相干多普勒振动测量仪，通过配备的两个扫描镜，可实现 X 轴和 Y 轴的扫描测振。



图 1.13 SLV-S02 型激光相干多普勒测振仪

Figure 1.13 SLV-S02 Laser Doppler vibrometer

图 1.14 为日本 ONOSOKK 公司生产的 LV-3800 型 3D 激光相干多普勒测振系统，该款产品的频率测量范围为 0.3-100KHz，可实现对振动物体进行三维振动测量。同时，该系统配备了 LV-0108 型定位相机，可将测量部分准确清晰的显示在电脑上。



图 1.14 LV-3800 型三维激光相干多普勒测振系统

Figure 1.14 LV-3800 3D Laser Doppler Vibration Measurement System

国内对于激光相干多普勒振动测量技术的实际应用也进行了很多研究，其中较具有代表性的是上海光机所的尚建华、赵曙光、贺岩等人<sup>[74]</sup>搭建的全光纤激光相干多普勒测振仪，该设备的振动位移分辨率可达 2.5pm，振动速度分辨率可达 96.9nm/s。其系统结构详见图 2.15。此外，西北工业大学的 Wending Zhang、Wei Gao、Ligang Huang 等人<sup>[70]</sup>基于光纤声光结构搭建了激光相干多普勒测振系统。图 1.15 为该系统的结构图，该系统的最小可测量振幅和分辨率在振动频率为数十赫兹时为 6pm,振动频率为数百至上千赫兹时为 1pm。

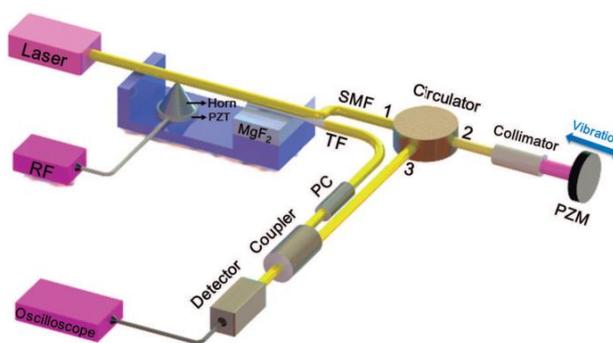


图 1.15 基于光纤声光调制器的激光相干多普勒测振系统结构图

Figure 1.15 Structure diagram of the laser Doppler vibration measurement system based on fiber optic acousto-optic modulator

除此之外，中北大学<sup>[75]</sup>，西安电子科大<sup>[76-77]</sup>，长春理工<sup>[78]</sup>，长光所<sup>[79]</sup>，国防科大<sup>[72,80]</sup>，浙大<sup>[81-83]</sup>，天大<sup>[84]</sup>等科研院所也对该技术进行了研究。除了上述

科研院所对激光相干多普勒测振技术进行研究之外,国内一些企业也对该技术表现出极大的兴趣,并成功研制出相关产品。如图 1.16 所示,为舜宇光学科技公司生产的 3DLVS-F 型三维激光相干多普勒测振仪,该设备可同时测量物体 X、Y、Z 三个方向上的振动,并且其光学头采用光纤结构,使得结构更加紧凑。



图 1.16 3DLVS-F 型三维激光相干多普勒测振仪

Figure 1.16 3DLVS-F 3D Laser Doppler vibrometer

#### 1.4 论文研究思路和结构安排

由前文所述研究现状可知,国内在激光语音侦测领域和激光相干多普勒测振领域已经获得了一定的进展,然而要实现远距离、高质量的激光语音侦测(针对采用激光相干多普勒测振技术作为测量方法的系统而言)仍有以下问题需要解决及完善:

1) 激光语音侦测系统设计。目前,根据已发表的文献可知,国外实现激光语音侦测的主要方法是直接利用商用激光相干多普勒测振仪侦测语音。但是,这些商用测振仪的波长多属于可见光波段,这大大降低了侦测的隐蔽性。除此之外,商用激光测振仪为了保证低频振动测量的准确性,大多采用空间光结构,这导致光学系统大且重,便携性差。国内虽然也研制了数套基于激光相干多普勒测量技术的激光语音侦测设备,然而,其性能与国外类似产品相比仍有一定差距。

2) 测量目标选择。由于激光语音侦测是通过检测声源附近物体的振动以获取语音信号,因此测量物体选择的好坏直接影响侦测语音的质量。然而,不同物体对于声场的振动响应、表面粗糙度以及与激光语音侦测系统的相对运动各不相

同,选择何种物体作为测量目标目前尚没有完善的理论研究。

3) 语音重构及增强。由于激光语音侦测系统直接输出的信号为混频信号,因此需要解调算法重构出语音信号。考虑实际应用需求,解调算法需要满足解调精度高,实时,抗干扰能力强等特点。此外,系统直接重构获取的语音存在噪声干扰、语音带宽较窄等问题,因此需要对其进行语音增强。

本论文主要研究内容将围绕上述三个问题展开,逐步解决各项问题的关键难点,为后续实现更远距离、更高质量的语音侦测奠定基础。本论文具体结构如下:

第一章:阐述了本文的研究意义并介绍了国内外激光语音侦测的现状,比较了目前几种常见的激光语音侦测技术,概述了激光相干多普勒测振技术的研究进展,对国内外激光相干多普勒测振技术的现状进行了介绍,最后给出了论文的结构安排。

第二章:阐述了基于激光相干多普勒测振技术的激光语音侦测的原理,并给出了可描述激光相干语音侦测系统性能指标的数学表达式,这些评价指标分别为载噪比、探测分辨率和探测概率。之后理论分析了像差、激光器线宽、探测方式以及发射光串扰等光学因素对激光相干语音侦测系统性能的影响。最后在参考了全光纤激光相干语音侦测系统光路结构的基础上,设计了部分光纤光路结构,并根据理论研究结果确定了系统的光学元器件。

第三章:基于 STL 理论和 SSC 模型建立了不同材质及形状的平板在声场的振动模型,并通过相关实验进行验证。之后采用仿真与实验结合的方式分析了物体表面粗糙度对激光相干语音侦测的影响。最后,根据散斑特性,研究了由于相对运动而产生的动态散斑噪声对激光相干语音侦测系统的影响,并进行实验验证,除此之外还通过理论计算得出了散斑噪声的功率谱密度。本章研究内容可为激光相干语音侦测中选择合适的测量目标提供指导。

第四章:介绍了常见的几种解调算法,针对激光侦听系统需要兼顾解调速度与精度的要求,采用正交解调与反正切鉴相结合的解调算法重构语音信号。并针对系统直接重构获取的语音存在噪声干扰、语音带宽较窄等问题,使用适合相干侦测语音的降噪算法抑制噪声,同时使用带宽扩展算法扩展语音带宽,进一步增强语音质量。

第五章:在不同距离,不同测量物体的情况下,进行了激光相干语音侦测实验,用以证明自研系统可获取远距离可理解语音信号。此外,分别进行了双模监

控实验和远距离说话人识别实验,用以表明自研系统在安防监控领域和生物识别领域应用的可行性和潜力。

第六章:总结了全文工作,给出了主要结论,并指出了工作的不足,最后,对激光相干语音侦测技术的下一步工作进行了展望。



## 第2章 激光相干语音信号侦测理论研究和光路结构设计

本章首先阐述了基于激光相干多普勒测振技术的语音侦测原理。其次，详细推导了激光相干语音侦测系统的载噪比、探测分辨率及探测概率的数学表达式，这些参数可以评估系统性能。然后详细分析了相干语音侦测过程中的光学影响因素，这些影响因素包括光学结构设计及装调不当而引入的像差、激光器固有的线宽、探测方式以及发射光的串扰。最后，结合激光侦听应用需求，并参考了全光纤激光相干语音侦测系统结构，设计了部分光纤结构，并根据理论研究结果，确定了自研激光相干语音侦测系统的光学元器件参数。

### 2.1 激光相干语音侦测原理

#### 2.1.1 激光相干语音侦测实现原理

研究表明：语音信号可以引起周围任何物体表面的振动，然而这种振动幅度通常较小，只有微米或者是纳米级<sup>[10]</sup>。但是激光相干多普勒测振技术具有远距离、非接触、精确检测物体振动的能力。因此正如图 2.1 所示，激光相干语音侦测指的是采用激光相干多普勒测振技术检测声源附近物体表面的振动信号从而获取语音信号，这种方法类似于传统录音设备通过检测麦克风膜的振动来获取声音信号。所以，本文所述的激光相干语音侦测技术可等效为指激光相干多普勒测振技术。

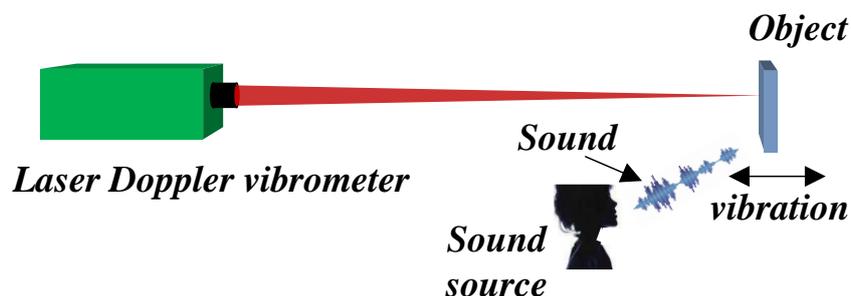


图 2.1 激光相干语音侦测示意图

Figure 2.1 The schematic diagram of laser coherent speech acquisition

### 2.1.2 激光相干多普勒测振原理

#### (1) 激光多普勒效应

激光多普勒效应是由奥地利科学家多普勒在 1842 年首先发现的，如图 2.2 所示，即当光源与探测器之间存在相对运动时，探测器探测到的光频率与光源处发射的光频率之间会存在一个多普勒频率差  $\Delta f$ ，多普勒频率差  $\Delta f$  可以表示为：

$$\Delta f = \frac{(\vec{v}_o + \vec{v}_d) \cdot \vec{e}}{\lambda} = \frac{\vec{V} \cdot \vec{e}}{\lambda} \quad (2-1)$$

公式 (2-1) 中， $v_o$  为光源的运动速度， $v_d$  为探测器的运动速度， $V=v_o+v_d$  为光源和探测器的相对速度， $\lambda$  为光波长。

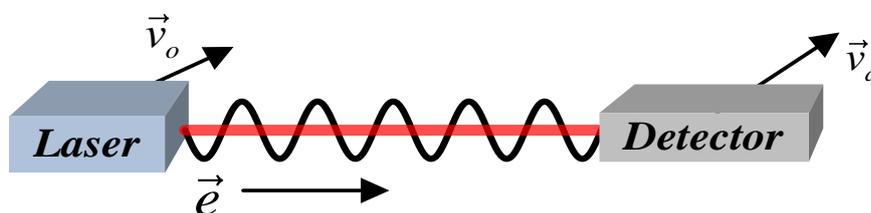


图 2.2 多普勒效应示意图

Figure 2.2 The schematic diagram of doppler effect

#### (2) 基于多普勒效应的激光相干多普勒振动测量原理

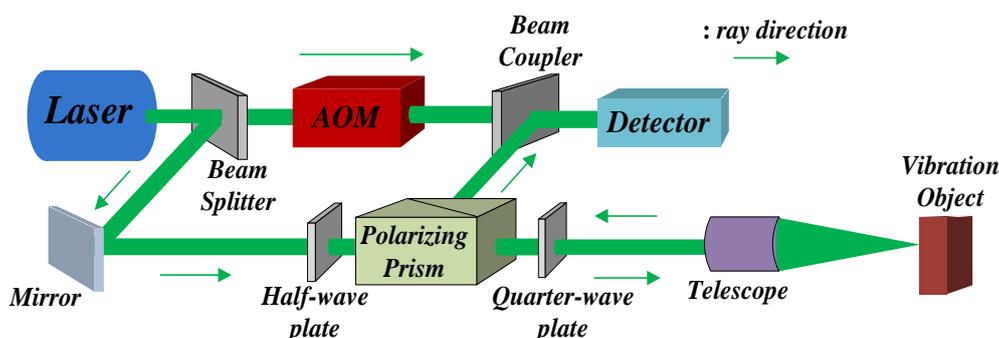


图 2.3 常见的激光相干多普勒测振仪结构图

Figure 2.3 Scheme of the Laser Doppler vibrometer

观察公式 (2-1) 可发现，通过检测激光的多普勒频移  $\Delta f$ ，可以获取物体的瞬时运动速度。激光相干多普勒测振也正是基于这个原理，即由于物体的振动，会使得物体的反向散射光中包含有多普勒频移，通过检测反向散射光的多普勒频移，便可获取物体振动的速度和位移。但是由于激光的光学频率太高导致探测器

不能响应,因此需采用相干探测的方式来提取反向散射光的多普勒频移。相干探测与直接探测相比,需要额外引入一束本征光与接收到的回波光进行混频,通过检测混频信号的频率与相位的变化来获取多普勒频移。一般而言,如图2.3所示,激光相干多普勒测振仪多采用改进型的马赫-曾德尔干涉仪结构<sup>[8]</sup>。

在激光相干多普勒测振仪中,激光器发射出的相干光束经过分束器分为本征光与测量光。测量光经过半波片、偏振棱镜、四分之一波片和聚焦光学系统聚焦在振动物体上。由于物体振动,会使得从振动物体散射的回波光携带多普勒频移,回波光经由聚焦光学系统接收后,再依次通过四分之一波片和偏振棱镜,最后在合束器处与本征光做混频,此时混频信号的光场可由公式(2-2)表示。另外,图2.3中,半波片、偏振棱镜、四分之一波片用于隔离测量光和回波光,并确保回波光的偏振方向和本征光相同。声光调制器(AOM, Acoustic-Optic modulator)则是让本征光产生一个固定的频移 $\omega_{if}$ ,以此来辨别物体振动的方向,当 $\omega_{if}=0$ 时,称为零差测量,当 $\omega_{if}\neq 0$ 时,称为外差测量。

$$E(x, y, t) = E_{Lo}(x, y) \exp(i[\omega_0 + \omega_{if}]t) + E_s(x, y) \exp(i[(\omega_0 + \omega_d)t - \omega_0 \tau_d]) \quad (2-2)$$

公式(2-2)中,  $E_{Lo}(x, y)$ 为本征光复振幅,  $\omega_0$ 为激光频率,  $E_s(x, y)$ 为回波光复振幅,  $\tau_d=2L/c$ 为延迟时间,其中  $L$ 为探测距离,  $c$ 为光速。 $\omega_d$ 为多普勒频移,由公式(2-1)可得,多普勒频移 $\omega_d$ 和物体振动速度之间的关系为:

$$\omega_d = 2\pi \frac{2V_r}{\lambda} \cos \theta \quad (2-3)$$

公式(2-3)中,  $V_r$ 为物体振动的速度,  $\theta$ 为振动速度与光轴的夹角。

光混频信号通过光电探测器输出的光电流为

$$i(t) = \iint \beta(x, y) [E_{Lo}(x, y)]^2 dx dy + \iint \beta(x, y) [E_s(x, y)]^2 dx dy + 2\text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_s^*(x, y) \exp[i(\omega_d + \omega_{if})t] \right\} dx dy \right) \quad (2-4)$$

公式(2-4)中,  $\beta(x, y) = e\eta(x, y)/h\nu$ 表示探测器响应度,  $\eta(x, y)$ 为量子效率,  $e$ 为电子电荷量,  $h\nu$ 为光子能量。\*代表复共轭。以下几个公式表示公式(2-4)中的各个部分:

$$i_{Lo} = \iint \beta(x, y) [E_{Lo}(x, y)]^2 dx dy \quad (2-5)$$

$$i_s = \iint \beta(x, y) [E_s(x, y)]^2 dx dy \quad (2-6)$$

$$i_{IF}(t) = 2 \operatorname{Re} \left( \iint \beta(x, y) \{ E_{Lo}(x, y) E_s^*(x, y) \exp [ i (\omega_d + \omega_{if}) t ] \} dx dy \right) \quad (2-7)$$

公式(2-5)和公式(2-6)分别为本征光和回波光在探测器的输出, 这两项与时间无关, 可称为直流电流。公式(2-7)与时间相关称为中频电流。一般而言, 中频电流是系统想获得的有用信号, 获得中频电流的一个常见方法是将探测器输出信号通过以  $\omega_{if}$  为中心频率的带通滤波器。

从公式(2-4)可以看出, 光电流信号随本征光和回波光的多普勒差频  $\omega_d$  成比例变化。通过解调这种成比例的变化便可获得多普勒差频  $\omega_d$ , 根据获得的多普勒差频  $\omega_d$  就可以得到目标振动速度和位移。

## 2.2 激光相干语音侦测性能分析

### 2.2.1 探测分辨率

如 2.1.1 节所述, 激光相干语音侦测的本质是检测声源附近物体的振动, 所以探测分辨率是评判语音侦测系统性能优劣以及能否实现语音侦测的一项重要参数。因此, 研究激光相干语音侦测系统的探测分辨率对于语音侦测系统的设计与搭建具有重要的指导意义。

假设振动目标做简谐振动, 振动幅度为  $A_0$ , 振动频率为  $f_0$ , 其振动位移  $s(t)$  可表示为:

$$s(t) = A_0 \cos(2\pi f_0 t) \quad (2-8)$$

并假设目标振动速度与光轴平行, 此时根据公式 (2-3) 可得, 由物体振动引起的多普勒频移  $f_d(t)$  和相位调制  $\varphi_d(t)$  分别为:

$$f_d(t) = \frac{2}{\lambda} \frac{d(s(t))}{dt} = \frac{4\pi f_0 A_0 \sin(2\pi f_0 t)}{\lambda} \quad (2-9)$$

$$\varphi_d(t) = \int_0^t 2\pi f_d(t) dt = \int_0^t \frac{4\pi}{\lambda} \frac{d(s(t))}{dt} dt = \frac{4\pi s(t)}{\lambda} = \frac{4\pi}{\lambda} A_0 \cos(2\pi f_0 t) \quad (2-10)$$

对光电探测器输出的电流在频率等于  $\omega_{if}$  做傅立叶变换可得:

$$\begin{aligned} F(i) &= \operatorname{Re} \left( \iint \beta(x, y) \{ E_{Lo}(x, y) E_s^*(x, y) \} dx dy \right) \exp(j\varphi_d) + n \\ &= c \exp(j\varphi_d) + n \end{aligned} \quad (2-11)$$

公式 (2-11) 中,  $c$  代表中频信号幅值,  $n$  代表噪声。与中频  $\omega_{if}$  相比, 多普勒相移  $\varphi_d(t)$  的变化速度要慢的多, 因此在这里多普勒相移  $\varphi_d$  与时间  $t$  的关系可以

忽略<sup>[65]</sup>。对于激光相干多普勒测振而言，散粒噪声是主要的噪声源，所以噪声  $n$  在复平面中为二维高斯随机分布，其概率密度函数为  $\exp[-(r^2+i^2)/(2n_{rms}^2)]/(2\pi n_{rms}^2)^{[65]}$ 。 $n_{rms}=\sqrt{ei_{Lo}B}$ 为噪声  $n$  的均方根值，其中  $B$  为带宽。公式 (2-11) 的相量图如图 2.4 所示

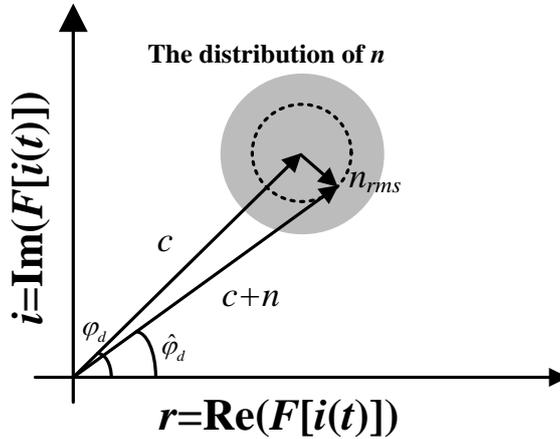


图 2.4 中频信号与噪声的相量图

Figure 2.4 Phasor diagram of carrier and noise.

如图 2.4 所示，一般而言中频功率远大于噪声功率，然而由于噪声的影响，会使得多普勒相移  $\varphi_d$  产生相位偏移  $\Delta\varphi_d$ ，相位偏移  $\Delta\varphi_d$  的标准差为：

$$\Delta\varphi_{d\_rms} = \arctan\left(\frac{n_{rms}}{|c|}\right) \approx \sqrt{\frac{n_{rms}^2}{|c|^2}} \quad (2-12)$$

由于激光相干多普勒测振系统载噪比的定义为中频电流处的功率与背景噪声在中频处功率的比值，因此，系统的载噪比 CNR 可表示为：

$$CNR = \frac{|c|^2}{n_{rms}^2} \quad (2-13)$$

考虑到低于和高于载波频率的两个不相关噪声分量总是会有助于噪声调制，因此相位偏移  $\Delta\varphi_d$  需要在公式 (2-12) 上引入因子  $\sqrt{2}$ <sup>[36]</sup>，此时再根据公式 (2-13) 可得，相位偏移  $\Delta\varphi_d$  可以表示为：

$$\Delta\varphi_d = \sqrt{\frac{2}{CNR}} \quad (2-14)$$

值得注意的是,在公式(2-13)中系统载噪比  $CNR$  是在带宽为  $B$  的情况下获得的,一般而言定义系统的位移分辨率和速度分辨率往往将带宽设为  $1\text{Hz}$ ,即将光电探测器输出的信号通过带宽为  $1\text{Hz}$  的带通滤波器。根据公式(2-14), (2-9)和(2-8)可得振幅误差  $\Delta s$  和速度误差  $\Delta v$  分别为:

$$\Delta s = \frac{\lambda}{4\pi} \Delta \varphi_d = \frac{\lambda}{4\pi} \sqrt{\frac{2}{CNR}} \quad (2-15)$$

$$\Delta v = \frac{\lambda}{2} f_{d\_max} \Delta \varphi_d = \pi f_0 A_0 \lambda \sqrt{\frac{2}{CNR}} \quad (2-16)$$

观察公式(2-15)和(2-16)可以发现,振幅误差  $\Delta s$  和速度误差  $\Delta v$  均与系统载噪比  $CNR$  的根号成反比,这表明为了提高激光相干语音侦测系统的探测分辨率,提高系统的载噪比是一条有效的途径。同时还可以看出,振幅误差  $\Delta s$  和速度误差  $\Delta v$  与波长  $\lambda$  成正比,波长  $\lambda$  越小,探测分辨率越高。因此为了提高系统的探测分辨率,在条件允许的情况下,激光器波长应选择较短的波长。另外,速度误差  $\Delta v$  会随着振动频率的增加而变大。

### 2.2.2 探测概率

正如前几节所述,激光语音相干侦测系统的侦测过程是通过探测多普勒相位调制信号  $\varphi_d(t)$  来获得语音信号。因此能否准确地探测到相位调制信号  $\varphi_d(t)$  是衡量系统性能的一个重要指标。由于在实际探测中,系统不可避免的会受到噪声的干扰,因此,本节主要分析了在带噪环境下,激光相干语音侦测系统正确探测到多普勒相位调制信号  $\varphi_d(t)$  的概率。

由公式(2-4)可得,光电探测器输出的电流包含了直流分量和中频分量。针对激光相干侦听而言,直流分量是没有价值的,而中频分量则包含了系统想要获得的多普勒相位调制  $\varphi_d(t)$ 。因此,为了滤除直流分量,同时滤除中频带外的噪声,会将探测器输出信号通过一个以频率  $\omega_{if}$  为中心的带通滤波器。一般而言,载波频率  $\omega_{if}$  有几十至上百  $\text{MHz}$  (由声光调制器调制频率决定),而带通滤波器的带宽  $B$  一般为几十至上百  $\text{KHz}$ ,因此,可将通过了带通滤波器的中频电流信号  $i_{IF}(t)$  称为窄带信号。与此同时,噪声  $n(t)$  也通过了以  $\omega_{if}$  为中心频率,频率宽带为  $B$  的带通滤波器,此时噪声  $n(t)$  的功率谱密度  $P_n(\omega)$  仅在频率  $\omega_{if}-1/2B < \omega < \omega_{if}+1/2B$  范

围内存在,通常把这种类型的噪声称为窄带噪声。根据公式(2-7)可将带噪的中频电流信号  $i_{IF}(t)$  表示为<sup>[85]</sup>:

$$\begin{aligned} i_{IF}(t) &= a_{if} \cos[\omega_{if}t + \varphi_d(t)] + n(t) \\ &= a_{if} \cos[\omega_{if}t + \varphi_d(t)] + a_n(t) \cos(\omega_{if}t + \theta_n(t)) \\ &= a_{if} (\cos\varphi_d(t) \cos(\omega_{if}t) - \sin\varphi_d(t) \sin(\omega_{if}t)) + n_R(t) \cos(\omega_{if}t) - n_I(t) \sin(\omega_{if}t) \quad (2-17) \\ &= [a_{if} \cos\varphi_d(t) + n_R(t)] \cos(\omega_{if}t) - [a_{if} \sin\varphi_d(t) + n_I(t)] \sin(\omega_{if}t) \\ &= i_{IF\_R}(t) \cos(\omega_{if}t) - i_{IF\_I}(t) \sin(\omega_{if}t) \end{aligned}$$

公式(2-17)中,  $a_{if}$  为中频电流幅值,  $a_n(t)$  和  $\theta_n(t)$  分别为噪声  $n(t)$  的随机幅度和随机相位,  $n_R(t)$  和  $n_I(t)$  为正交分量。

$$a_{if} = 2 \operatorname{Re} \left( \iint \beta(x, y) \{ E_{Lo}(x, y) E_s^*(x, y) \} dx dy \right) \quad (2-18)$$

$$n_R(t) = a_n(t) \cos\theta_n(t), \quad n_I(t) = a_n(t) \sin\theta_n(t) \quad (2-19)$$

$$i_{IF\_R}(t) = a_{if} \cos\varphi_d(t) + n_R(t) \quad (2-20)$$

$$i_{IF\_I}(t) = a_{if} \sin\varphi_d(t) + n_I(t) \quad (2-21)$$

根据公式(2-20)至(2-21)可知,从带噪中频电流  $i_{IF}(t)$  中探测得到的包络  $a_{if\_n}(t)$  和多普勒相移  $\dot{\varphi}_d(t)$  分别为

$$a_{if\_n}(t) = \sqrt{[i_{IF\_I}^2(t) + i_{IF\_R}^2(t)]} \quad (2-22)$$

$$\dot{\varphi}_d(t) = \arctan \frac{i_{IF\_I}(t)}{i_{IF\_R}(t)} \quad (2-23)$$

根据公式(2-22)和公式(2-23),公式(2-20)和(2-21)可以重新表示为

$$i_{IF\_R}(t) = a_{if\_n}(t) \cos\dot{\varphi}_d(t) \quad (2-24)$$

$$i_{IF\_I}(t) = a_{if\_n}(t) \sin\dot{\varphi}_d(t) \quad (2-25)$$

为了获得探测概率,首先需要获取噪声  $n(t)$  的正交分量  $n_R(t)$  和  $n_I(t)$  的统计特性。根据前几节叙述可知,散粒噪声是系统的主要噪声源,在中低频率下,散粒噪声可被认为是高斯白噪声,因此可认为窄带噪声  $n(t)$  是零均值的平稳的高斯噪声。由式(2-19)可得,当  $t=0$  时,  $n(t) = n_R(t)$ , 根据一维雅可比变换可得  $p(n_R) = p(n=n_R)$ 。类似地,当  $t=3\pi / (2\omega_{if})$ , 可得  $p(n_I) = p(n=n_I)$ 。此时可得  $E[n(t)] =$

$E[n_R(t)] = E[n_I(t)] = 0$ 。由此说明噪声  $n(t)$  的两个正交分量  $n_R(t)$  和  $n_I(t)$  的均值与噪声  $n(t)$  相同。

由(2-17)可得，噪声  $n(t)$  的自相关函数为：

$$\begin{aligned} r_n(\tau) &= E[n(t)n(t+\tau)] \\ &= r_{n_R}(\tau) \cos(\omega_{if}t) \cos(\omega_{if}(t+\tau)) - r_{n_R n_I}(\tau) \cos(\omega_{if}t) \sin(\omega_{if}(t+\tau)) \\ &\quad + r_{n_I}(\tau) \sin(\omega_{if}t) \sin(\omega_{if}(t+\tau)) - r_{n_R n_I}(\tau) \sin(\omega_{if}t) \cos(\omega_{if}(t+\tau)) \end{aligned} \quad (2-26)$$

因为窄带噪声  $n(t)$  是平稳的高斯噪声，所以可得

$$\sigma_n^2 = r_n(\tau) \Big|_{\tau=0} = \sigma_{n_R}^2 \Big|_{t=0, \tau=0} = \sigma_{n_I}^2 \Big|_{t=-\frac{\pi}{2\omega_{if}}, \tau=0} \quad (2-27)$$

根据上述分析可知，噪声  $n(t)$  的正交分量  $n_R(t)$  和  $n_I(t)$  为均值是 0，方差为  $\sigma_n^2$  的高斯噪声，并且相互统计独立。所以，对于给定多普勒相位  $\varphi_d$  的  $i_{IF\_R}$  和  $i_{IF\_I}$  信号来说，他们是相互独立的高斯随机过程。他们的条件均值和方差分别为：

$$E[i_{IF\_R}(t) | \varphi_d] = a_{if} \cos \varphi_d \quad \text{Var}[i_{IF\_R}(t) | \varphi_d] = \sigma_n^2 \quad (2-28)$$

$$E[i_{IF\_I}(t) | \varphi_d] = a_{if} \sin \varphi_d \quad \text{Var}[i_{IF\_I}(t) | \varphi_d] = \sigma_n^2 \quad (2-29)$$

此时  $i_{IF\_R}$  和  $i_{IF\_I}$  的联合条件密度函数为

$$p(i_{IF\_R}, i_{IF\_I} | \varphi_d) = \frac{1}{2\pi\sigma_n^2} \exp \left[ -\frac{(i_{IF\_R} - a_{if} \cos \varphi_d)^2 + (i_{IF\_I} - a_{if} \sin \varphi_d)^2}{2\sigma_n^2} \right] \quad (2-30)$$

根据公式(2-24)和(2-25)，和雅可比变换可得在给定多普勒相位  $\varphi_d$  的情况下，探测得到的包络  $a_{if\_n}(t)$  和多普勒相移  $\dot{\varphi}_d(t)$  的联合条件概率密度函数为

$$p(a_{if\_n}, \dot{\varphi}_d | \varphi_d) = \frac{a_{if\_n}}{2\pi\sigma_n^2} \exp \left[ -\frac{a_{if\_n}^2 + a_{if}^2 - 2a_{if\_n} a_{if} \cos(\dot{\varphi}_d - \varphi_d)}{2\sigma_n^2} \right] \quad (2-31)$$

根据公式(2-31)，可得到探测到的多普勒相移  $\dot{\varphi}_d(t)$  的条件概率密度函数为

$$\begin{aligned}
 p(\dot{\varphi}_d|\varphi_d) &= \int_0^\infty p(a_{if\_n}, \dot{\varphi}_d|\varphi_d) da_{if\_n} \\
 &= \int_0^\infty \frac{a_{if\_n}}{2\pi\sigma_n^2} \exp\left[-\frac{a_{if\_n}^2 + a_{if}^2 - 2a_{if\_n}a_{if} \cos(\dot{\varphi}_d - \varphi_d)}{2\sigma_n^2}\right] da_{if\_n} \\
 &= \frac{1}{2\pi} \exp\left[-\frac{a_{if}^2}{2\sigma_n^2}\right] + \frac{a_{if} \cos(\dot{\varphi}_d - \varphi_d)}{\sqrt{2\pi\sigma_n^2}} \exp\left[-\frac{a_{if}^2 \sin^2(\dot{\varphi}_d - \varphi_d)}{2\sigma_n^2}\right] \\
 &\quad \times \int_{-\frac{a_{if} \cos(\dot{\varphi}_d - \varphi_d)}{\sigma_n}}^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2}\right) d\mu
 \end{aligned} \tag{2-32}$$

根据载噪比定义可得，系统的载噪比为

$$CNR = \frac{1}{2} \frac{a_{if}^2}{\sigma_n^2} \tag{2-33}$$

则探测到的多普勒相移 $\dot{\varphi}_d(t)$ 的条件概率密度函数可表示为

$$\begin{aligned}
 p(\dot{\varphi}_d|\varphi_d) &= \frac{1}{2\pi} \exp[-CNR] + \frac{\sqrt{CNR} \cos(\dot{\varphi}_d - \varphi_d)}{\sqrt{\pi}} \exp[-CNR \sin^2(\dot{\varphi}_d - \varphi_d)] \\
 &\quad \times \int_{-\sqrt{\frac{CNR}{2}} \cos(\dot{\varphi}_d - \varphi_d)}^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2}\right) d\mu
 \end{aligned} \tag{2-34}$$

将探测概率定义为探测得到的多普勒相移 $\dot{\varphi}_d$ 与真实的多普勒相移 $\varphi_d$ 的差值不超过 $\Delta\varphi$ 的概率。则根据公式(2-34)可得探测概率 $P(|\dot{\varphi}_d - \varphi_d| \leq \Delta\varphi)$ 为

$$P(|\dot{\varphi}_d - \varphi_d| \leq \Delta\varphi) = \int_{-\Delta\varphi}^{\Delta\varphi} p(\dot{\varphi}_d|\varphi_d) d(\dot{\varphi}_d - \varphi_d) \tag{2-35}$$

图 2.5 为在不同载噪比下的探测概率曲线( $\Delta\varphi=0.01\pi$ )。从图 2.5 中可以发现，探测概率和载噪比有很大的关系，载噪比 $CNR$ 越大，探测概率越大。

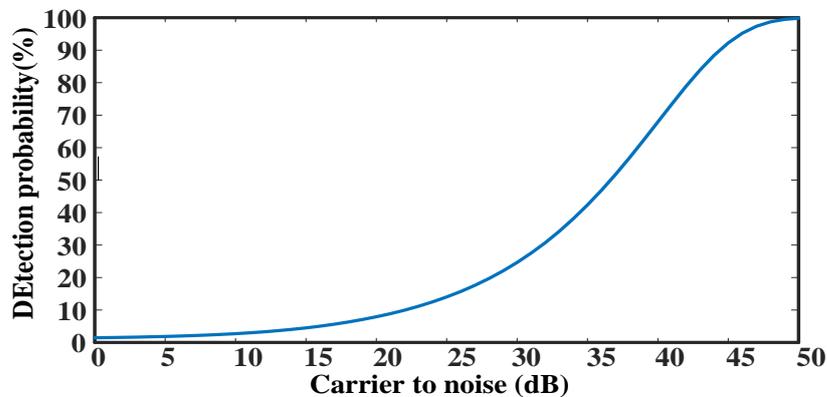


图 2.5 探测概率随载噪比的变化

Figure 2.5 The detection probability as a function of the CNR.

### 2.2.3 系统载噪比

从 2.2.1 节和 2.2.2 节可以看出，激光语音相干侦测系统的性能与系统载噪比 CNR 息息相关，系统载噪比 CNR 越高，语音侦测性能越好。因此研究系统的载噪比对激光相干语音侦测系统的设计与搭建具有重要的指导意义。对于激光相干语音侦测系统（也就是激光相干测振系统）而言，载波信号指的是以  $\omega_{if}$  为中心频率的中频电流（如公式(2-7)所示）。假设负载电阻为  $R_L$ ，则有效中频功率为：

$$P_{IF} = \overline{(i_{IF})^2} R_L = i_{\Delta\omega}^2 R_L = 2 \left\{ \iint \beta(x, y) [E_{Lo}(x, y) E_s^*(x, y)] dx dy \right\}^2 R_L \quad (2-36)$$

假设本征光和回波光为平面波，则他们的复振幅可以表示为

$$E_{Lo}(x, y) = E_{Lo} \exp(i\vec{k}_L \cdot \vec{r}) \quad (2-37)$$

$$E_s(x, y) = E_s \exp(i\vec{k}_s \cdot \vec{r}) \quad (2-38)$$

公式(2-37)和(2-38)中,  $E_{Lo}$  和  $E_s$  为常数,  $\vec{r}$  为在探测器光敏面的空间矢量,  $\vec{k}_s$  和  $\vec{k}_L$  分别为回波光和在本征光的传播矢量, 他们的传播方向如图 2.6 所示。

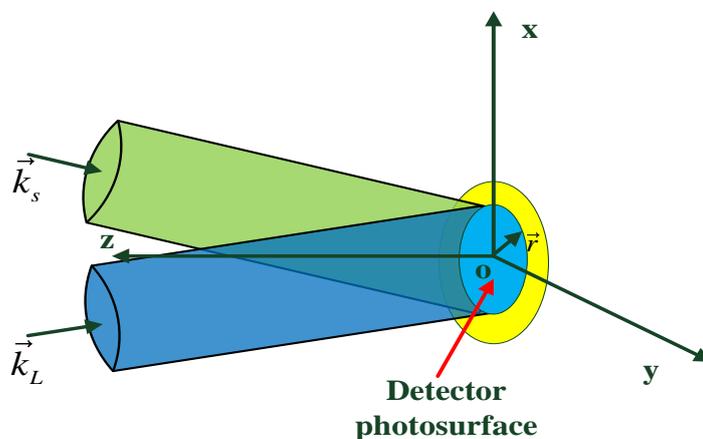


图 2.6 相对坐标系

Figure 2.6 Relative coordinates system.

为了方便起见，假设  $\vec{k}_s$  在  $oxz$  面内，并且与  $x$  轴的夹角为  $\theta_x$ ， $\vec{k}_L$  与  $y$  轴的夹角为  $\alpha_y$ ，与  $x$  轴的夹角为  $\alpha_x$ ，则公式(2-36)可被改写为<sup>[44]</sup>：

$$\begin{aligned}
 P_{IF} &= 2E_{Lo}^2 E_s^2 R_L \left\{ \iint \beta(x, y) \left[ \exp\left(i(k_L \cos \alpha_x - k_s \cos \theta_x)x + i(k_L \cos \alpha_y)y\right) \right] dx dy \right\}^2 \\
 &= 2E_{Lo}^2 E_s^2 R_L \left\{ \iint \beta(x, y) \left[ \exp\left(i(2\pi f_x x + 2\pi f_y y)\right) \right] dx dy \right\}^2
 \end{aligned} \quad (2-39)$$

公式(2-39)中  $f_x$  和  $f_y$  为空间频率。

$$f_x = \frac{\cos \alpha_x}{\lambda} - \frac{\cos \theta_x}{\lambda_s} \approx \frac{\cos \alpha_x - \cos \theta_x}{\lambda} \quad (2-40) \quad f_y = \frac{\cos \alpha_y}{\lambda} \quad (2-41)$$

假设探测器响应度  $\beta(x, y)$  为常数  $\beta_0$ ，并且探测器的光敏面面积为直径  $D$  的圆，则公式 (2-39) 可以表示为：

$$\begin{aligned}
 P_{IF} &= 2E_{Lo}^2 E_s^2 \beta_0^2 R_L \left\{ \int_0^{D/2} \int_0^{2\pi} \exp\left(i2\pi\rho(f_x \cos \theta + f_y \sin \theta)\right) \rho d\theta d\rho \right\}^2 \\
 &= 2E_{Lo}^2 E_s^2 \beta_0^2 R_L \left\{ \frac{DJ_1\left(\pi D\sqrt{f_x^2 + f_y^2}\right)}{2\sqrt{f_x^2 + f_y^2}} \right\}^2
 \end{aligned} \quad (2-42)$$

公式 (2-42) 中， $J_1(\cdot)$  表示为一阶贝塞尔函数。

对于激光相干语音侦测系统而言，系统主要的噪声源是散粒噪声，因此噪声的平均功率为：

$$P_N = 2eBR_L \langle i(t) \rangle \quad (2-43)$$

公式 (2-43) 中， $B$  代表带宽。根据式 (2-4)，公式 (2-43) 可表示为：

$$\begin{aligned}
 P_N &= 2eBR_L \beta_0 \left( \iint E_{Lo}(x, y) E_{Lo}^*(x, y) dx dy + \iint E_s(x, y) E_s^*(x, y) dx dy \right) \\
 &= 2eBR_L \beta_0 \frac{\pi D^2}{4} (E_{Lo}^2 + E_s^2)
 \end{aligned} \quad (2-44)$$

根据公式 (2-42) 和公式 (2-44)，可得系统的载噪比为  $CNR$  为：

$$CNR = \frac{P_{IF}}{P_N} = \pi D^2 \frac{E_{Lo}^2 E_s^2 \beta_0}{eB(E_{Lo}^2 + E_s^2)} \left\{ \frac{J_1\left(\pi D\sqrt{f_x^2 + f_y^2}\right)}{\pi D\sqrt{f_x^2 + f_y^2}} \right\}^2 \quad (2-45)$$

一般而言，本征光光强要远大于回波光光强，所以公式 (2-45) 可简化为：

$$CNR = \frac{P_{IF}}{P_N} = \pi D^2 \frac{E_s^2 \beta_0}{eB} \left\{ \frac{J_1 \left( \pi D \sqrt{f_x^2 + f_y^2} \right)}{\pi D \sqrt{f_x^2 + f_y^2}} \right\}^2 \quad (2-46)$$

由公式 (2-46) 可知, 系统的载噪比 CNR 主要取决于:

1) 回波光光强: 系统的载噪比与回波光的光强成正比。

2) 本征光与回波光的夹角, 以及他们与探测器之间的夹角: 如图 2.7 所示, 回波光光强一定时, 系统的载噪比最大值在  $f_x^2 + f_y^2 = 0$  即  $f_x = 0, f_y = 0$  处, 即回波光和本征光垂直入射在探测器时 ( $\theta_x = \alpha_y = \alpha_x = \pi/2$ ), 此时信噪比如式 (2-47) 所示; 当空间频率变大时, 即回波光与本征光之间夹角变大时, 载噪比也随之降低。

$$CNR_{\max} = \lim_{f_x \rightarrow 0, f_y \rightarrow 0} \pi D^2 \frac{E_s^2 \beta_0}{eB} \left\{ \frac{J_1 \left( \pi D \sqrt{f_x^2 + f_y^2} \right)}{\pi D \sqrt{f_x^2 + f_y^2}} \right\}^2 = \frac{\pi D^2 E_s^2 \beta_0}{4 eB} \quad (2-47)$$

3) 探测器带宽和响应度。

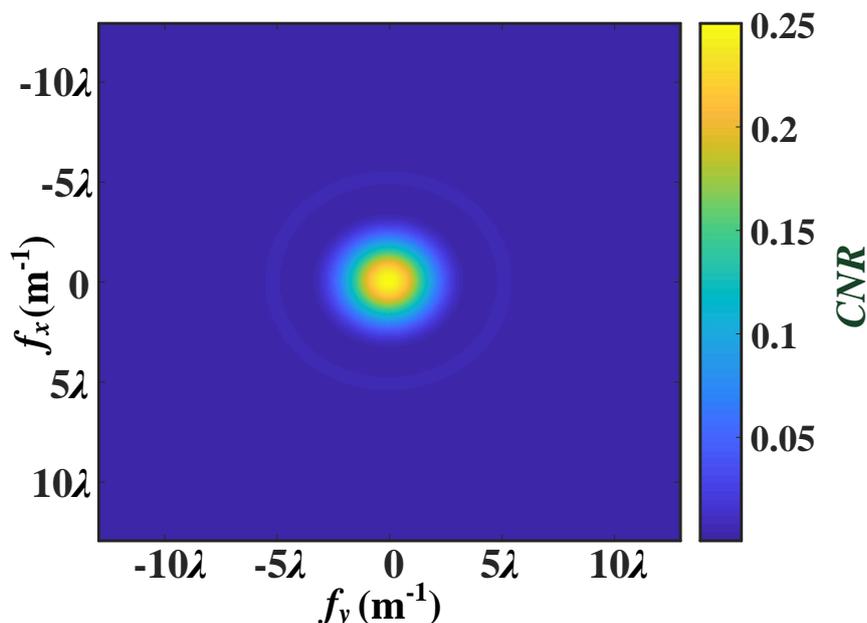


图 2.7 载噪比和空间频率的关系

Figure 2.7 Relationship between carrier-to-noise ratio and spatial frequency.

因此, 在实际系统搭建时, 为了提高系统的载噪比, 可适当选择功率较大的激光器作为光源; 也可以通过增加接收系统的口径来增加回波能量; 并且必须保

证本征光和回波光传播方向相互匹配且均垂直入射在探测器面上;另外还可以选择响应度好的光电探测器。

## 2.3 激光相干语音侦测性能光学影响因素分析

### 2.3.1 像差对激光相干语音侦测系统的影响

光学系统是激光相干语音侦测系统中激光传输的路径,光学系统性能的优劣直接决定了相干语音侦测系统的侦测性能。像差指的是实际光学系统所成的像与光学系统近轴区所成的像之间的差异,像差往往会损害光学系统的性能。但是在光学系统实际设计与搭建过程中,会因为装调不当、设计失误等原因,使得系统不可避免的存在像差。因此,分析像差对激光相干语音侦测系统的影响对于提高侦测性能具有重要的指导意义。由2.2节分析可知,激光语音侦测系统的性能可通过系统载噪比 CNR 衡量,因此分析像差对语音相干侦测系统的影响时,仅需分析像差对载噪比的影响。

像差的本征是由于光学系统不同位置处的光程对理想情况的偏差,从而导致光束产生波前误差  $\Delta W$ 。由于泽尼克多项式和光学检测中观测到的像差多项式的形式相互一致,并且在单位圆内部是连续正交的,因此在本节中,它被用来拟合光束的波前特性。值得注意的是,当存在空气扰动时,很难用泽尼克多项式拟合光束的波前,但是考虑到语音相干侦测的应用场合中空气扰动较小,因此暂不考虑空气扰动。在单位圆内 ( $\rho \leq 1$ ),泽尼克多项式定义为<sup>[86-87]</sup>:

$$U_{nm} = R_n^{n-2m}(\rho) \begin{cases} \sin \\ \cos \end{cases} (n-2m)\theta \quad (2-48)$$

公式(2-48)中,当  $n-2m > 0$  时取 sin 函数,当  $n-2m < 0$  时取 cos 函数。其中  $R_n^{n-2m}$  表示为:

$$R_n^{n-2m}(\rho) = \sum_{s=0}^m (-1)^s \frac{(n-s)!}{s!(m-s)!(n-m-s)!} \rho^{n-2s} \quad (2-49)$$

表(2.1)为初级像差的泽尼克多项式:

表 2.1 初级像差的泽尼克多项式

Table 2.1 Zernike Polynomial for Primary Aberrations

n	m	Meaning	Zernike polynomial
0	0	Piston	1
1	0	Y tilt	$\rho \sin \theta$
1	1	X tilt	$\rho \cos \theta$
2	1	Defocus	$2\rho^2 - 1$
2	0	Y primary astigmatism	$\rho^2 \sin 2\theta$
2	2	X primary astigmatism	$\rho^2 \cos 2\theta$
3	1	Y primary coma	$(3\rho^3 - 2\rho) \sin \theta$
3	2	X primary coma	$(3\rho^3 - 2\rho) \cos \theta$
3	0	Y trefoil	$\rho^3 \sin 3\theta$
3	3	X trefoil	$\rho^3 \cos 3\theta$
4	2	Primary spherical	$6\rho^4 - 6\rho^2 + 1$

由像差引起的波前误差  $\Delta W$  可以用泽尼克多项式表示，即

$$\Delta W(\rho, \theta) = \sum_{n=0}^k \sum_{m=0}^n A_{nm} U_{nm} \quad (2-50)$$

公式 (2-50) 中， $A_{nm}$  为泽尼克多项式项的系数。

由像差引起的波前误差  $\Delta W$  大小通常是由均方根 (rms)  $\sigma$  表示，即

$$\begin{aligned} \sigma^2 &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 [\Delta W(\rho, \theta) - \overline{\Delta W}]^2 \rho d\rho d\theta \\ &= \left[ \frac{1}{\pi} \int_0^{2\pi} \int_0^1 \Delta W^2(\rho, \theta) \rho d\rho d\theta \right] - \frac{1}{\pi^2} \left[ \int_0^{2\pi} \int_0^1 \Delta W(\rho, \theta) \rho d\rho d\theta \right]^2 \end{aligned} \quad (2-51)$$

又因为波前误差  $\Delta W$  可由泽尼克多项式表示，因此，根据公式 (2-50)，可将公式 (2-51) 简化为：

$$\sigma^2 = \sum_{n=0}^k \sum_{m=0}^n \frac{\varepsilon_{nm}}{2(n+1)} A_{nm}^2 \quad (2-52)$$

公式 (2-52) 中，当  $2m=n$  时， $\varepsilon_{nm}=2$ ，当  $2m \neq n$  时， $\varepsilon_{nm}=1$ 。

由 2.1.2 节所述，激光语音相干侦听系统的探测原理是探测回波光和本征光之间的混频信号。现假设本征光是无像差畸变的平面波（可继续用公式 2-37 表示），回波光则发生了像差畸变，因此，此时回波光可表示为：

$$E_s(x, y) = E_s \exp(i\vec{k}_s \cdot \vec{r}) \exp(i2\pi\Delta W) \quad (2-53)$$

根据公式 (2-42) 可得，此时中频功率为  $P_{IF}$  为：

$$P_{IF} = 2E_{Lo}^2 E_s^2 \beta_0^2 R_L \left\{ \frac{DJ_1 \left( \pi D \sqrt{f_x^2 + f_y^2} \right)}{2\sqrt{f_x^2 + f_y^2}} \right\}^2 \frac{1}{\pi^2} \left\{ \int_0^1 \int_0^{2\pi} \exp(i2\pi\Delta W(\rho, \theta)) \rho d\rho d\theta \right\}^2 \quad (2-54)$$

根据公式 (2-44) 和 (2-54) 可得, 当存在像差时, 系统载噪比 CNR 为

$$CNR = \frac{P_{IF}}{P_N} = \pi D^2 \frac{E_s^2 \beta_0}{eB} \left\{ \frac{J_1 \left( \pi D \sqrt{f_x^2 + f_y^2} \right)}{\pi D \sqrt{f_x^2 + f_y^2}} \right\}^2 \frac{1}{\pi^2} \left\{ \int_0^1 \int_0^{2\pi} \exp(i2\pi\Delta W(\rho, \theta)) \rho d\rho d\theta \right\}^2 \quad (2-55)$$

图 2.8 为不同类型的像差对激光相干语音侦测系统性能的影响。从图中可以发现, 当存在像差时, 系统载噪比 CNR 会下降, 并且均方根  $\sigma$  越大, 系统载噪比 CNR 越低。但是值得注意的是, 当像差较小时, 即均方根  $\sigma < 0.1\lambda$ , 系统对像差噪声相对来说不太敏感。然而, 当像差较大时, 相较于其他类型的像差, 倾斜、像散以及慧差对系统影响更大。

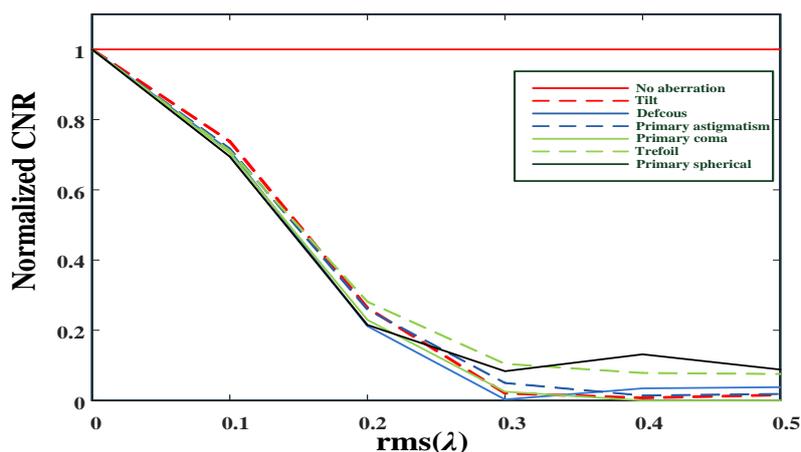
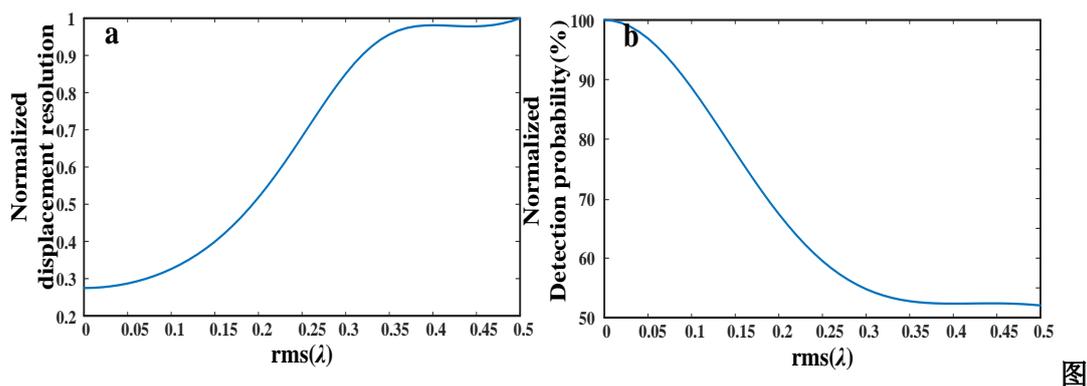


图 2.8 初级像差下, 系统载噪比随均方根的变化

Figure 2.8 The CNR as a function of the rms for the primary aberrations.

图 2.9.a 和图 2.9.b 分别为在存在像差的情况下, 系统的位移分辨率(归一化)以及探测概率 (归一化,  $\Delta\phi=0.01\pi$ )。从图中也可以发现, 系统性能随着像差的变大而变差。



2.9 (a) 系统位移分辨率随像差均方根的变化。(b) 系统探测概率随像差均方根的变化

**Figure 2.9 (a) The displacement resolution as a function of the rms for the primary aberrations. (b) The detection probability as a function of the rms for the primary aberrations.**

以上分析结果表明像差对系统性能影响很大,在实际系统设计与搭建过程中,需要重点优化光学系统的像差。

### 2.3.2 激光器谱线展宽对激光相干语音侦测系统的影响

激光产生的本质是受激辐射,但是在受激辐射的过程中,会伴随着自发辐射及各种随机因素,使得输出的激光光场在振幅和相位上存在随机波动。由于激光光场振幅波动相较于相位波动而言变化缓慢,并且波动幅度较小,因此可以认为激光光场振幅近似为常量。此时,激光器输出激光可表示为:

$$E(t) = A \exp\{i[\omega_0 t + \phi(t)]\} \quad (2-56)$$

公式(2-56)中,  $A$  为输出光场振幅,  $\omega_0$  为角频率,  $\phi(t)$  为相位噪声,正是由于相位噪声  $\phi(t)$  会使得激光器输出激光存在谱线展宽。此时,在不考虑输出激光在传输过程中产生的随机相位,回波光场  $E_s$  和本征光场  $E_{Lo}$  在探测器的光敏面处可分别表示为:

$$E_s(x, y, t) = \alpha A \exp\{i[\omega_0(t - \tau_d) + \phi(t - \tau_d) + \omega_d t]\} \exp(i\vec{k}_s \cdot \vec{r}) \quad (2-57)$$

$$E_{Lo}(x, y, t) = A \exp\{i[\omega_0 t + \phi(t)]\} \exp(i\vec{k}_{Lo} \cdot \vec{r}) \quad (2-58)$$

公式(2-57)中,  $\alpha$  为回波光场与本征光场振幅之间的比值,一般而言,  $\alpha \ll 1$ 。由于本节只讨论激光器的谱线展宽(即相位噪声  $\phi(t)$ )对相干语音侦测系统的影

响,因此假设本征场  $E_{Lo}$  与回波光场  $E_s$  垂直射入光敏面,并且相互重合,此时在探测器光敏面处光场可表示为:

$$\begin{aligned} E(x, y, t) &= E_{Lo}(x, y, t) + E_s(x, y, t) \\ &= A \exp\{i[\omega_0 t + \phi(t)]\} + \alpha A \exp\{i[\omega_0(t - \tau_d) + \phi(t - \tau_d) + \omega_d t]\} \end{aligned} \quad (2-59)$$

公式(2-59)中,  $\tau_d = 2L/C$  表示延迟时间,  $L$  表示系统与测量目标之间的距离,  $C$  表示光速。

为了计算方便,假设光敏面面积为1,探测器响应度  $\beta(x, y) = 1$ , 振幅  $A = 1$ , 此时探测器输出的中频电流信号  $i_{IF}(t)$  为

$$\begin{aligned} i_{IF}(t) &= \alpha \exp\left(i\left[\omega_0 \tau_d + (\phi(t) - \phi(t - \tau_d)) - \omega_d t\right]\right) \\ &\quad + \alpha \exp\left(-i\left[\omega_0 \tau_d + (\phi(t) - \phi(t - \tau_d)) - \omega_d t\right]\right) \end{aligned} \quad (2-60)$$

由2.2节所知,激光相干语音侦测系统的性能可用载噪比 CNR 衡量,因此研究激光器谱线宽度对系统的影响,可以近似等效为研究谱线宽度对载噪比的影响。研究对载噪比的影响,首先应获得中频电流  $i_{IF}(t)$  的功率,根据维纳-辛钦定理可得:宽平稳随机过程的功率谱密度是其自相关函数的傅立叶变换。因此可先通过求取中频电流信号  $i_{IF}(t)$  的自相关函数来获取功率谱密度,再根据  $i_{IF}(t)$  的功率谱密度求得  $i_{IF}(t)$  的功率。中频电流信号  $i_{IF}(t)$  的自相关函数  $\gamma_i(\tau)$  为:

$$\begin{aligned} \gamma_i(\tau) &= \langle (i_{IF}(t) i_{IF}(t - \tau)) \rangle \\ &= \alpha^2 \langle \exp(i\omega_d \tau) \rangle \langle \exp(i[\phi(t) - \phi(t - \tau) + \phi(t - \tau - \tau_d) - \phi(t - \tau_d)]) \rangle \\ &\quad + \alpha^2 \langle \exp(-i\omega_d \tau) \rangle \langle \exp(-i[\phi(t) - \phi(t - \tau) + \phi(t - \tau - \tau_d) - \phi(t - \tau_d)]) \rangle \\ &\quad + \alpha^2 \langle \exp(-i2\omega_d \tau) \rangle \langle \exp(i[\phi(t) + \phi(t - \tau) - \phi(t - \tau - \tau_d) - \phi(t - \tau_d)]) \rangle \\ &\quad \times \langle \exp(i[2\omega_0 \tau_d - \omega_d \tau]) \rangle + \alpha^2 \langle \exp(i2\omega_d \tau) \rangle \langle \exp(-i[2\omega_0 \tau_d - \omega_d \tau]) \rangle \\ &\quad \times \langle \exp(-i[\phi(t) + \phi(t - \tau) - \phi(t - \tau - \tau_d) - \phi(t - \tau_d)]) \rangle \end{aligned} \quad (2-61)$$

在公式(2-61)中,  $\exp(i2\omega_d \tau)$  在时间上的均值可表示为

$$\langle \exp(\pm i2\omega_d \tau) \rangle = \int_{-\infty}^{\infty} \exp(\pm i2\omega_d t) dt = \int_{-\infty}^{\infty} [\cos(\pm 2\omega_d t) + i \sin(\pm 2\omega_d t)] dt = 0 \quad (2-62)$$

并记  $\Delta\phi(t,\tau)=\phi(t)-\phi(t-\tau)$  为相位抖动, 即随机相位  $\phi$  在时间  $t-\tau$  和  $t$  之间的随机变化。相位抖动  $\Delta\phi(t,\tau)$  通常被认为是零均值的平稳高斯随机过程。并且存在如下关系式<sup>[51]</sup>:

$$\langle \exp(\pm i\Delta\phi(t,\tau)) \rangle = \exp\left[-\frac{1}{2}\langle \Delta\phi^2(t,\tau) \rangle\right] \quad (2-63)$$

为了避免相位噪声  $\phi(t)$  在时间轴上发生重叠, 因此当延迟时间  $\tau \leq \tau_d$  时, 记  $\Delta\phi_1(t,\tau)=\phi(t)-\phi(t-\tau)$ ,  $\Delta\phi_2(t-\tau_d,\tau)=\phi(t-\tau_d)-\phi(t-\tau_d-\tau)$ 。当延迟  $\tau > \tau_d$  时, 记  $\Delta\phi_1(t,\tau_d)=\phi(t)-\phi(t-\tau_d)$ ,  $\Delta\phi_2(t-\tau,\tau_d)=\phi(t-\tau)-\phi(t-\tau_d-\tau)$ 。

因此根据式 (2-62) 和 (2-63), 可将中频电流信号  $i_{IF}(t)$  的自相关函数  $\gamma_i(\tau)$  简化为:

$$\gamma_i(\tau) = \alpha^2 \cos(\omega_d \tau) \begin{cases} \exp\left[-\frac{1}{2}\langle \Delta\phi^2(t,\tau) \rangle\right] \exp\left[-\frac{1}{2}\langle \Delta\phi^2(t-\tau_d,\tau) \rangle\right] & \tau \leq \tau_d \\ \exp\left[-\frac{1}{2}\langle \Delta\phi^2(t,\tau_d) \rangle\right] \exp\left[-\frac{1}{2}\langle \Delta\phi^2(t-\tau,\tau_d) \rangle\right] & \tau > \tau_d \end{cases} \quad (2-64)$$

根据文献[50~51]可知, 若激光器的光源谱线是 Lorentz 型的, 则

$$\langle \Delta\phi^2(t,\tau) \rangle = 2\Delta\omega_L \tau \quad (2-65)$$

公式 (2-65) 中,  $2\Delta\omega_L$  为 Lorentz 型激光器在谱线上的半峰全宽 (FWHM), 并且对应的激光器的相干时间为  $\tau_c = \pi/\Delta\omega_L$ 。

根据公式 (2-64) 和维纳-辛钦定理可得, 中频电流信号  $i_{IF}(t)$  的功率谱密度  $P_{IF}(\omega)$  为:

$$\begin{aligned} P_{IF}(\omega) &= 2 \int_0^{\infty} \gamma_i(\tau) \exp(-j\omega\tau) d\tau \\ &= 2\alpha^2 \int_0^{\tau_d} \cos(\omega_d \tau) \exp(-2\Delta\omega_L \tau) \exp(-j\omega\tau) d\tau \\ &\quad + 2\alpha^2 \int_{\tau_d}^{\infty} \cos(\omega_d \tau) \exp(-2\Delta\omega_L \tau_d) \exp(-j\omega\tau) d\tau \end{aligned} \quad (2-66)$$

考虑实际的物理意义, 可将公式 (2-66) 化简为<sup>[51]</sup>:

$$\begin{aligned} P_{IF}(\omega) &= \alpha^2 \exp(-2\Delta\omega_L \tau_d) \delta(\omega - \omega_d) \\ &\quad + \alpha^2 \frac{\Delta\omega_L \exp(-2\Delta\omega_L \tau_d)}{\Delta\omega_L^2 + (\omega - \omega_d)^2} \left[ \exp(2\Delta\omega_L \tau_d) - \frac{\sin(\omega\tau_d - \omega_d\tau_d)}{(\omega - \omega_d)} - \cos(\omega\tau_d - \omega_d\tau_d) \right] \end{aligned} \quad (2-67)$$

由第 2.1 节可知，激光相干语音侦测实质是通过获取多普勒频移  $\omega_d$  来侦测语音信号，因此定义中频功率  $P_{IF}$  为频谱以  $\omega_d$  为中心，带宽为  $B$  之间的功率，如公式 (2-68) 所示：

$$P_{IF} = \int_{\omega_d - B/2}^{\omega_d + B/2} P_{IF}(\omega) d\omega \quad (2-68)$$

根据上几节所述，在激光相干语音侦测系统中，散粒噪声是系统的主要噪声源，在本节中其功率可表示为：

$$P_n = 2ei_{Lo}B = 2eB \quad (2-69)$$

根据公式 (2-68) 和 (2-69) 可得系统载噪比  $CNR$  为：

$$CNR = \frac{P_{IF}}{P_n} = \frac{\int_{\omega_d - B/2}^{\omega_d + B/2} P_{IF}(\omega) d\omega}{2eB} \quad (2-70)$$

激光器谱线展开对系统载噪比的影响可根据公式(2-70)进行数值仿真分析。在仿真过程中设带宽  $B \rightarrow 0$ 。如图 2.10 所示为在延迟时间  $\tau_d$  分别等于 66.67ns, 333.3ns, 666.7ns, 1000.0ns 和 1333.3ns 时，归一化的系统载噪比在不同激光器谱线宽度下的分布。分析图 2.10 可以发现，系统载噪比随着激光器谱线宽度的增加而减小。同时也可以发现，在激光器谱线宽度相同的情况下，延迟时间  $\tau_d$  越大，系统载噪比越低。从图中可以发现，由于语音探测的距离一般为百米级，因此，系统最好选择谱线展宽小于 0.1MHz 的激光器作为光源。

图 2.11.a 和图 2.11.b 分别为在存在激光器谱线展宽的情况下，系统的位移分辨率（归一化， $\tau_d=666.7ns$ ）以及探测概率（归一化， $\Delta\varphi=0.01\pi$ ， $\tau_d=666.7ns$ ）。从图中也可以发现，系统性能随着激光器谱线宽度的变大而变差。

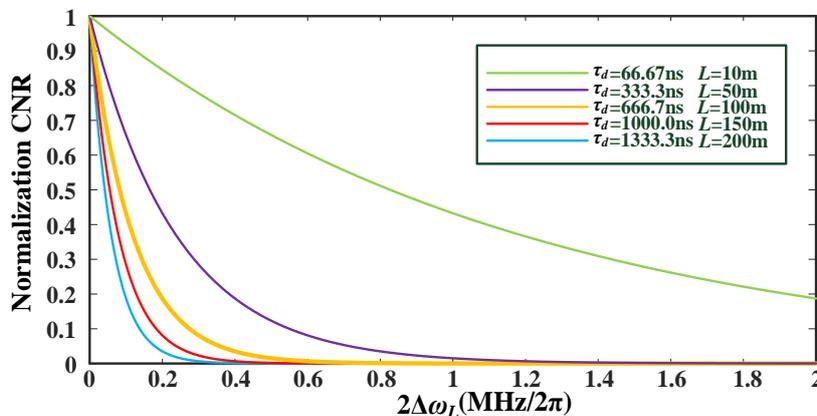


图 2.10 系统载噪比随激光器谱线宽带的变化

Figure 2.10 The CNR as a function of the laser linewidth.

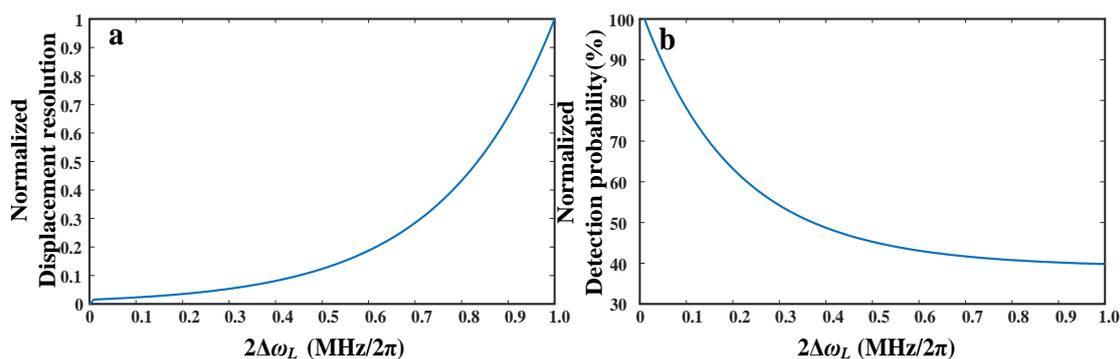


图 2.11 (a) 系统位移分辨率随线宽的变化。(b) 系统探测概率随线宽的变化

Figure 2.11 (a) The displacement resolution as a function of the laser linewidth.

(b) The detection probability as a function of the laser linewidth.

### 2.3.3 探测器探测方式对激光相干语音侦测系统的影响

激光相干语音侦测系统中，除了散粒噪声之外，本征过剩强度噪声也是主要的噪声源，该噪声是由于起伏的本征光引起的。消除此类噪声的方法中，平衡探测技术是一种非常有效的方法。

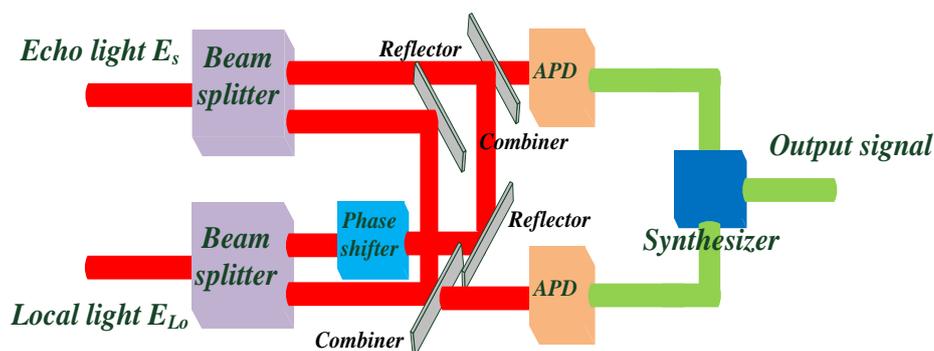


图 2.12 平衡探测器原理

Figure 2.12 The principle of balanced detector.

如图 2.12 所示为平衡探测器原理<sup>[49,88]</sup>。回波光  $E_s$  和本征光  $E_{Lo}$  分别经过光束分束器等分为两路。其中一路回波光和本征光在合束器里混合后入射到一个探测器里。另一路回波光与经过  $180^\circ$  相移的本征光在合束器里混合后入射到另一个探

测器里。两路探测器输出的信号经过合成器后输出差分中频信号。假设两个探测器的响应度分别为  $\beta_1$  和  $\beta_2$ ，光敏面的面积为 1，并且本征光和回波光都垂直入射在光敏面上，此时两个探测器输出的电流为：

$$i_1(t) = \beta_1 E_{Lo}^2 + \beta_1 E_s^2 + 2\beta_1 E_{Lo} E_s \cos(\Delta\omega t) \quad (2-71)$$

$$i_2(t) = \beta_2 E_{Lo}^2 + \beta_2 E_s^2 - 2\beta_2 E_{Lo} E_s \cos(\Delta\omega t) \quad (2-72)$$

根据[88]可知，过剩强度噪声功率的均方值为

$$P_{N\_exc} = 2eB\gamma R_L I_{dc}^2 \quad (2-73)$$

公式 (2-73) 中， $\gamma$  表示过剩强度噪声功率随本征光功率变化的系数， $I_{dc}$  为输出电流的直流项。根据公式 (2-71)、公式 (2-72) 和公式 (2-73) 可得系统的过剩强度噪声功率为：

$$P_{N\_exc} = 2eB\gamma R_L E_{Lo}^4 (\beta_2 - \beta_1)^2 \quad (2-74)$$

根据公式 (2-42)，(2-71)，(2-72) 和 (2-74) 可得，两路探测器输出的信号经过合成器后输出差分中频信号功率和噪声功率分别为：

$$P_{IF} = 2E_{Lo}^2 E_s^2 \beta_1^2 R_L + 2E_{Lo}^2 E_s^2 \beta_2^2 R_L \quad (2-75)$$

$$\begin{aligned} P_N &= P_{N\_exc} + |P_{N1\_shot} + P_{N2\_shot}| \\ &= 2eB\gamma R_L E_{Lo}^4 (\beta_1 - \beta_2)^2 + 2eBR_L E_{Lo}^2 |\beta_1 + \beta_2| \end{aligned} \quad (2-76)$$

公式 (2-76) 中，前一项表示平衡探测时的过剩强度噪声功率，后一项表示散粒噪声功率。此时，根据公式 (2-75) 和 (2-76) 可知，平衡探测器输出的载噪比 CNR 为：

$$CNR = \frac{P_{IF}}{P_N} = \frac{E_s^2 (\beta_1^2 + \beta_2^2)}{eB |\beta_1 + \beta_2| + eB\gamma E_{Lo}^2 (\beta_1 - \beta_2)^2} \quad (2-77)$$

当单点探测时，系统的载噪比 CNR 为：

$$CNR = \frac{E_s^2 \beta}{eB\gamma E_{Lo}^2 + eB} \quad (2-78)$$

比较公式 (2-77) 和 (2-78) 可以发现，平衡探测器输出信号的载噪比大于单点探测器输出信号的载噪比，并且如果选择两个完全相同的探测器，即  $\beta_1 = \beta_2$  时，系统的过剩强度噪声趋向于零。

### 2.3.4 发射光串扰对激光相干语音侦测系统的影响

为了最大程度的收集回波光光强以及装调的便利性,同时为了减小系统体积、节省系统成本,激光语音相干侦测系统通常使用收发一体的聚焦系统用于将激光器能量发射到探测区域上以及接收回波光能量。使用收发一体的光学镜头组意味着发射光和回波光共用一个镜头,然而,如果系统隔离度不够,会导致发射光串扰到回波光传输方向。同时,其他光学系统镜头的反射也会引起这种串扰。此时,在探测器光敏面处的光场分布由公式(2-2)改写为:

$$E(x, y, t) = E_{Lo}(x, y) \exp[i(\omega_0 + \omega_f)t] + E_{ct}(x, y) \exp(i\omega_0 t) + E_s(x, y) \exp[i[(\omega_0 + \omega_d)t - \omega_0 \tau_d]] \quad (2-79)$$

公式(2-79)中,  $E_{ct}(x, y)$  为串扰光复振幅。

此时,光电探测器输出的光电流由公式(2-4)改写为:

$$\begin{aligned} i(t) = & \iint \beta(x, y) [E_{Lo}(x, y)]^2 dx dy + \iint \beta(x, y) [E_s(x, y)]^2 dx dy + \iint \beta(x, y) [E_{ct}(x, y)]^2 dx dy \\ & + 2\text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_s^*(x, y) \exp[i(\omega_d + \omega_f)t] \right\} dx dy \right) \\ & + 2\text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_{ct}^*(x, y) \exp[i\omega_f t] \right\} dx dy \right) \\ & + 2\text{Re} \left( \iint \beta(x, y) \left\{ E_s(x, y) E_{ct}^*(x, y) \exp[i\omega_d t] \right\} dx dy \right) \end{aligned} \quad (2-80)$$

公式(2-80)中,串扰光的振幅  $E_{ct}(x, y)$  一般要远远小于本征光  $E_{Lo}(x, y)$  的振幅,所以串扰光和回波光之间的混频可以忽略。此时探测器的中频电流由公式(2-7)改写为:

$$\begin{aligned} i_{IF}(t) = & 2\text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_s^*(x, y) \exp[i(\omega_d + \omega_f)t] \right\} dx dy \right) \\ & + 2\text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_{ct}^*(x, y) \exp[i\omega_f t] \right\} dx dy \right) \end{aligned} \quad (2-81)$$

在不考虑噪声的情况下,对光电探测器输出的电流在频率等于  $\omega_{if}$  做傅立叶变换可得

$$\begin{aligned} F(i) = & \text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_s^*(x, y) \right\} dx dy \right) \exp(j\phi_d) \\ & + \text{Re} \left( \iint \beta(x, y) \left\{ E_{Lo}(x, y) E_{ct}^*(x, y) \right\} dx dy \right) \\ = & c \exp(j\phi_d) + d \end{aligned} \quad (2-82)$$

公式(2-82)中,  $c$  代表本征光与回波光混频信号的幅值,  $d$  代表本征光和串扰光混频信号的幅值。公式(2-82)的相量图如图 2.13 所示

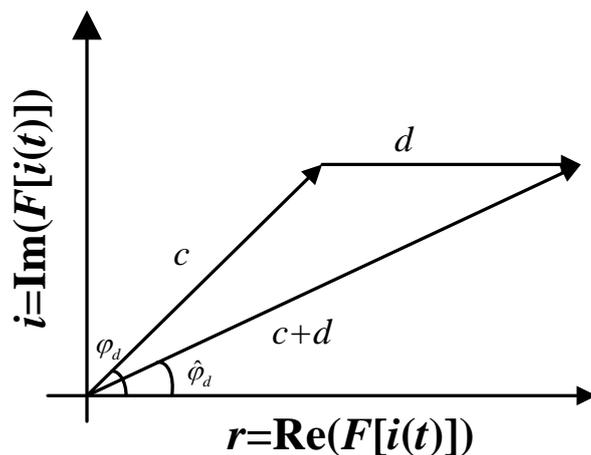


图 2.13 存在串扰的中频信号相量图

Figure 2.13 Phasor diagram of carrier with crosstalk.

如图 2.13 所示，此时观察到的多普勒频移  $\hat{\varphi}_d$  为：

$$\hat{\varphi}_d = \arctan \left( \frac{\sin \varphi_d}{\cos \varphi_d + d/c} \right) \quad (2-83)$$

由于串扰产生的相位偏移  $\Delta\varphi_d$  为：

$$\Delta\varphi_d = \varphi_d - \hat{\varphi}_d = \varphi_d - \arctan \left( \frac{\sin \varphi_d}{\cos \varphi_d + d/c} \right) \quad (2-84)$$

观察公式 (2-84) 和图 2.14 可以发现相位偏移  $\Delta\varphi_d$  随着  $d/c$  的增大而增大，这意味着，在回波光强度一定的情况下，串扰光强越强，探测得到的多普勒频移  $\hat{\varphi}_d$  与真实的多普勒频移  $\varphi_d$  之间的相位偏移  $\Delta\varphi_d$  也越大。以上结论说明，在系统设计与搭建时，要保证系统的隔离度。

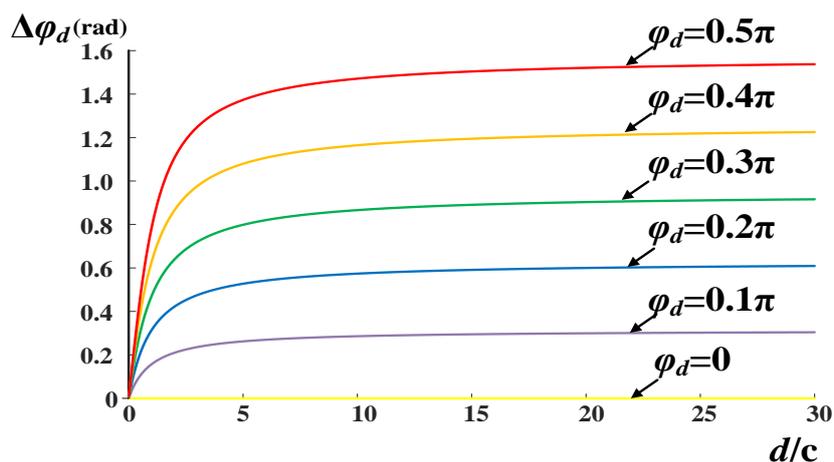


图 2.14 相位偏移随幅度比值的变化

Figure 2.14 The phase shift as a function of the ratio of the amplitudes.

## 2.4 激光相干语音侦测系统光路结构设计

激光相干语音侦测系统的光路结构与常见的激光相干多普勒测振仪类似(如图 2.3 所示),一般都采用改进型的马赫-曾德尔干涉仪结构。因此,使用高性能的商用激光相干多普勒测振仪也可以实现远距离的语音获取。例如 Weihong Li<sup>[7~11]</sup>等人采用德国 Polytec 公司生产的激光相干多普勒检测仪检测声源附近易振物体表面的振动以获取语音信号,并结合监视相机搭建了多模监视系统。Yekutiel Avargel 等人<sup>[14]</sup>使用商用激光相干多普勒检测仪检测发言人咽喉振动以获取语音信号。然而,目前商业激光相干多普勒测振仪的波长多属于可见光波段,这大大减弱了侦测的隐蔽性。除此之外,商用激光测振仪为了保证低频振动测量的准确性,大多采用空间光结构,这导致光学系统体积大并且重(例如, Polytec OFV 505 系统尺寸为 120 mm×80 mm×345 mm,重量为 3.4 kg)。以上这些弊端使得商用激光相干多普勒测振仪不合适直接用于激光相干语音侦测。本文在商用激光多普勒测振仪光路结构的基础上,采用光纤结构传输系统内部光束,这是由于光纤结构具有体积更小,重量更轻,结构更坚固的优点,除此之外,还大大降低了系统搭建难度。

### 2.4.1 全光纤激光相干语音侦测系统光路结构

图 2.15 为尚建华<sup>[20,74]</sup>以及 Rui Li<sup>[13]</sup>等人搭建的全光纤激光相干语音侦测系统的光路结构,尚建华使用该系统成功获取 25m 处的语音信号(振动物体为玻璃)。

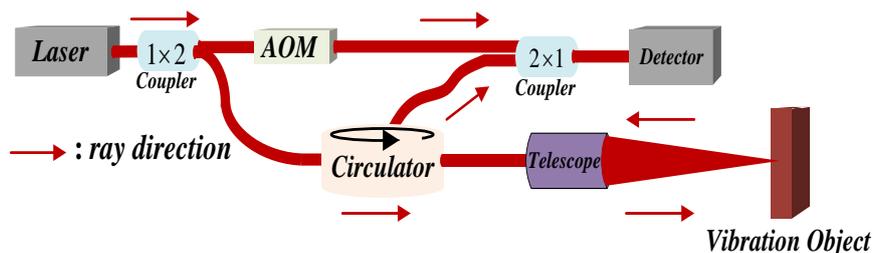


图 2.15 全光纤激光相干语音侦测结构图

Figure 2.15 Scheme of the all-fiber Laser Doppler vibrometer.

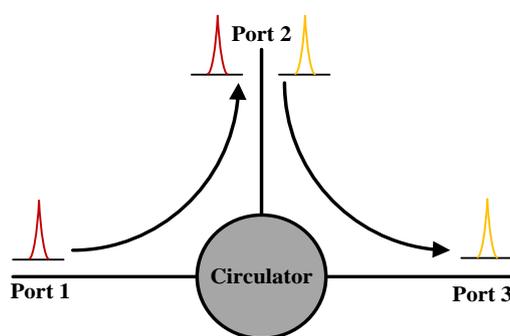


图 2.16 环形器示意图

Figure 2.16 Scheme of the Circulator.

观察图 2.15 可以发现，全光纤结构的相干语音侦测系统采用光纤环形器隔离发射光与回波光。图 2.16 为普通三端口环形器的示意图，其隔离原理可简单概述为从端口 1 进入的光传输路线与从端口 2 进入的光传输路线不在同一个空间区域内，从而避免了发射光与回波光之间的串扰。但是在实际应用中，由于环路器端口面存在反射，会导致约 3% 的发射光<sup>[89]</sup>串扰进回波光的传输路径中（不同类型的环形器串扰比例不同），即端口 1 至端口 2 传输的激光串扰至端口 2 到端口 3 的通道中。由 2.3.4 节可知，串扰光强会对激光相干语音侦测系统的性能造成影响，并且串扰光强与回波光强的比值  $d/c$  越大，对系统的性能影响也就越大。当探测距离较近时，此时由于回波光较强，串扰对系统性能的影响可以忽略。但是当探测距离为百米级时，结合实际情况和雷达方程（公式 2-85）可大致推算出回波能量要远小于串扰光能量，此时，串扰对系统的性能影响很大。

$$P_s = \frac{P_T}{4\pi R^2} \frac{\sigma_c}{4\pi R^2} \frac{\pi D^2}{4} \eta_{atm} \eta_{sys} \quad (2-85)$$

式 (2-85) 中,  $P_s$  为回波光功率,  $P_T$  为发射光功率,  $R$  为探测距离,  $D$  为接收孔径,  $\sigma_c$  为目标的雷达散射截面,  $\eta_{atm}$  为大气传输系数,  $\eta_{sys}$  为光学系统的传输系数。

### 2.4.2 部分光纤激光相干语音侦测系统光路结构

由上一节可知, 全光纤激光相干语音侦测系统由于存在串扰问题, 不适合用于远距离 (百米级) 语音侦测。为了解决串扰问题, 本文采用波片联合偏振棱镜的方式来隔离发射光与回波光, 这种方式不仅可以保证系统的隔离度、减小前几个反射面产生的后向反射噪声 (通过适当调整波片与棱镜之间的角度), 还可以保证回波光与本征光偏振方向相同。由于发射光与回波光的隔离过程是在自由空间中传输, 而其余过程, 例如混频、分束及合束都在单模保偏光纤内完成, 因此本文设计的光路结构可称为部分光纤结构。图 2.17 为本文采用的激光相干语音侦测系统光路结构图, 其工作原理已在 2.1.2 节中介绍, 在此不再赘述。

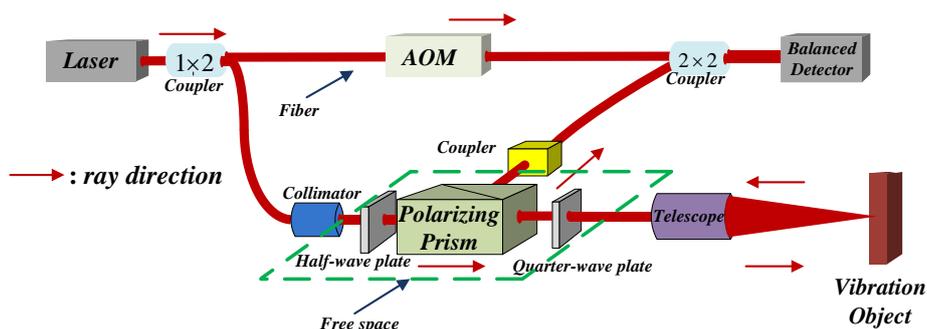


图 2.17 部分光纤激光相干语音侦测系统结构图

Figure 2.17 Scheme of the partial-fiber Laser Doppler vibrometer.

### 2.4.3 空间光与光纤高效耦合

正如 2.4.2 节所述, 在本文中激光相干语音侦测系统采用部分光纤的结构, 因此能否高效的将空间光耦合进单模保偏光纤对于语音侦测系统而言至关重要。空间光耦合进单模光纤的过程可以看作是空间光经过聚焦透镜后在焦面上产生的衍射模场与单模光纤的模场相互匹配的过程, 耦合效率  $\eta_c$  可表示为:

$$\eta_c = \frac{\left| \iint U_1^*(r) U_2(r) r dr d\theta \right|^2}{\iint U_1(r) U_1^*(r) r dr d\theta \iint U_2(r) U_2^*(r) r dr d\theta} \quad (2-86)$$

公式(2-86)中,  $U_1(r)$  为单模光纤端面上的电磁场分布,  $U_2(r)$  为空间光经过耦合聚焦透镜后, 在光纤端面上的电磁场分布。

由于单模光纤只能传输 LP<sub>01</sub> 模, 所以单模光纤端面上的电磁场分布  $U_1(r)$  可以近似为:

$$U_1(r) = \sqrt{\frac{2}{\pi}} \frac{1}{r_0} \exp\left[-\left(\frac{r}{r_0}\right)^2\right] \quad (2-87)$$

公式(2-87)中,  $r$  为到模场中心的距离,  $r_0$  为模场半径。一般而言:  $\lambda=632.8\text{nm}$ ,  $r_0=2.20\pm 0.25\mu\text{m}$ ;  $\lambda=1310\text{nm}$ ,  $r_0=4.60\pm 0.25\mu\text{m}$ ;  $\lambda=1550\text{nm}$ ,  $r_0=5.25\pm 0.5\mu\text{m}$ 。

假设空间光的波前为平面波(实际情况中, 由于回波光经过了长距离的传输, 可以将其等效为平面波), 假设耦合聚焦透镜的孔径为  $D_a$ 、焦距为  $f_a$ , 并在不考虑像差的情况下, 经过聚焦透镜后的空间光在光纤端面上的电磁场分布  $U_2(r)$  可表示为:

$$U_2(r) = \frac{\pi}{i\lambda f_a} \left(\frac{D_a}{2}\right)^2 \exp\left(i\left(kf_a + \frac{kr^2}{2f_a}\right)\right) \left(\frac{2J_1\left(\frac{\pi D_a r}{\lambda f_a}\right)}{\frac{\pi D_a r}{\lambda f_a}}\right) \quad (2-88)$$

根据公式(2-87)和公式(2-88)可得, 耦合效率  $\eta_c$  为

$$\eta_c = \frac{4 \left| \int_0^\infty \exp\left(-\left(\frac{r}{r_0}\right)^2\right) \exp\left(i\left(kf_a + \frac{kr^2}{2f_a}\right)\right) J_1\left(\frac{\pi D_a r}{\lambda f_a}\right) dr \right|^2}{r_0^2 \int_0^\infty \frac{(J_1(\pi D_a r / \lambda f_a))^2}{r} dr} \quad (2-89)$$

根据公式(2-89)可计算出耦合效率, 但是该式较为复杂, 不适合工程应用。根据 Parseval 定理可得, 入瞳面处计算得到的耦合效率和在光纤端面上计算得到的耦合效率一致, 在入瞳面处耦合效率  $\eta_{c,e}$  可表示为:

$$\eta_{c,e} = \frac{\left| \iint U_{1,e}^*(r) U_{2,e}(r) r dr d\theta \right|^2}{\iint U_{1,e}(r) U_{1,e}^*(r) r dr d\theta \iint U_{2,e}(r) U_{2,e}^*(r) r dr d\theta} \quad (2-90)$$

公式(2-90)中,  $U_{1,e}(r)$  为入瞳处单模光纤端面上的电磁场分布,  $U_{2,e}(r)$  为入瞳处空间光经过耦合聚焦透镜后的电磁场分布。

在入瞳面处, 归一化的光纤端面电磁场分布  $U_{1,e}(r)$  可表示为:

$$U_{1,e}(r) = \frac{kr_0}{\sqrt{2\pi}f_a} \exp\left[-\left(\frac{kr_0}{2f_a}r\right)^2\right] \quad (2-91)$$

空间光在入瞳处为平面波，即  $U_{2,e}(r) = 1$ ，根据公式 (2-90) 与公式 (2-91)，可得在入瞳面处计算得到的耦合效率  $\eta_{c,e}$  为：

$$\eta_{c,e} = \frac{2\left[1 - \exp(-a^2)\right]^2}{a^2} \quad (2-92)$$

公式 (2-92) 中,  $a = \pi r_0 D_a / (2\lambda f_0)$  为耦合参数。公式 (2-92) 求得的耦合效率与公式 (2-89) 求得的耦合效率是一致的，但是公式 (2-92) 求得的耦合效率表达形式更简单，更适合工程应用。

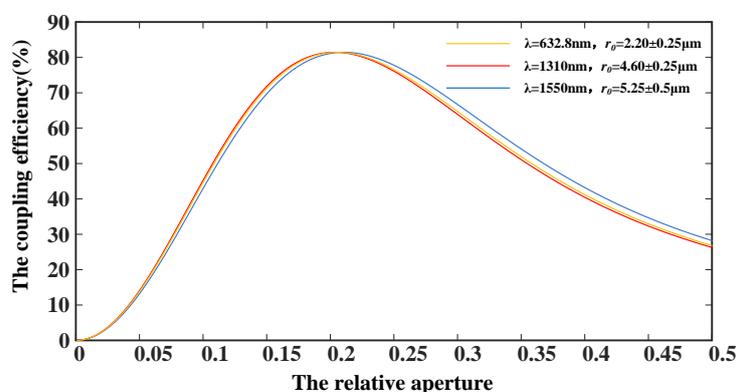


图 2.18 耦合效率与相对孔径的关系图

Figure 2.18 The coupling efficiency as a function of relative aperture.

图 2.18 为不同波长下，耦合效率与相对孔径  $D_a/f_0$  之间的关系。从图中可以发现：随着相对孔径  $D_a/f_0$  的增加，耦合效率呈现先增后减的趋势。这是由于当相对孔径  $D_a/f_0$  较小时，入射光斑能量受到衍射的影响而弥散，这会使得入射光斑面积与光纤端面面积失配，从而导致耦合效率变低。随着相对孔径  $D_a/f_0$  的增加，衍射影响减弱，耦合效率也逐步提高。但是当相对孔径  $D_a/f_0$  增加到一定值以后，会造成入射光斑的入射角过大，与光纤的数值孔径失配，又导致耦合效率降低。同时还发现为了获得最佳耦合效率 81.45%，不同波长配套的相对孔径略有差别。对于 1550nm 的光波而言，需要配套相对孔径  $D_a/f_0 = 0.211$  的透镜；对于 1310nm 的光波而言，需要配套相对孔径  $D_a/f_0 = 0.203$  的透镜；对于 623.8nm 的光波而言，则需要配套相对孔径  $D_a/f_0 = 0.205$  的透镜。

## 2.4.4 全光纤结构与部分光纤结构性能比较实验

## (1) 性能比较实验

为了验证部分光纤结构性能在远距离探测的情况下要优于全光纤结构,进行相关实验验证。图 2.19.a 为实验原理图。本次实验选择常见的矿泉水瓶作为测量目标,其位于激光相干语音侦测系统正前方 70m 处。声源位于矿泉水瓶左前方约 45cm 处,为了可以直观的比较测量结果,声源发出的声音为频率 500Hz 的单音信号,其音量约为 50.1 分贝(手机自带分贝仪检测)。在实验中,依次搭建部分光纤结构(如图 2.19.b)与全光纤结构(如图 2.19.c)的激光相干语音侦测系统测量矿泉水瓶的振动从而获得声源处产生的单音信号。为了排除其他因素对实验的干扰,在系统搭建过程中,全光纤结构与部分光纤结构的系统组成器件除环形器与偏振棱镜组之外全部相同。并在实验过程中,全光纤结构与部分光纤结构的系统所在位置相同,激光聚焦位置也相同,后续信号解调算法及算法中的参数也都相同,解调算法详见第 4 章。

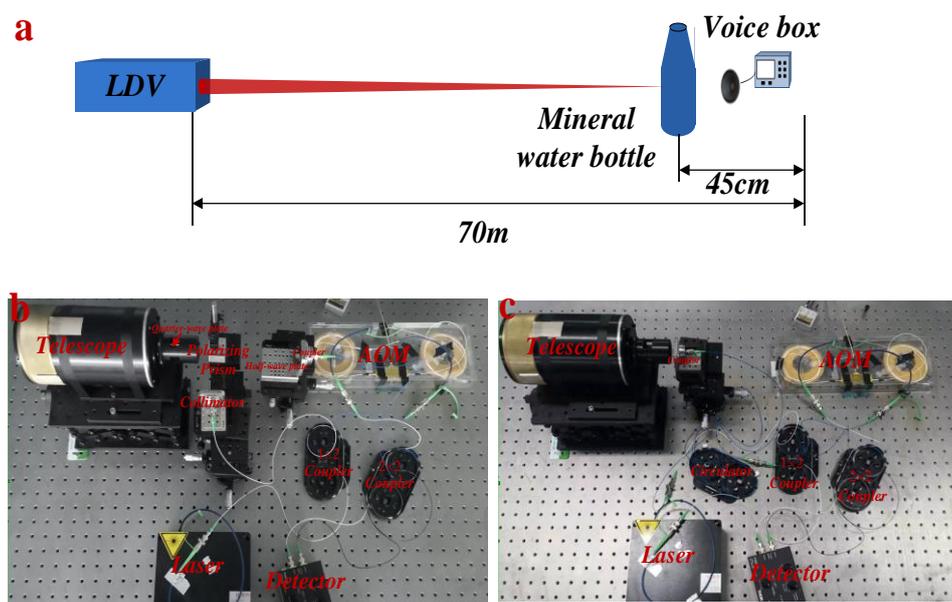


图 2.19 (a) 实验框图。(b) 部分光纤结构。(c) 全光纤结构

Figure 2.19 (a) Experimental block diagram. (b) Partial-fiber structure.

(c) All-fiber structure.

## (2) 实验系统关键器件介绍

根据 2.3 节研究结果可知，像差、激光器谱线展宽、过剩强度噪声等因素会严重影响激光相干语音侦测系统侦测性能，但是选择合适的光学元器件可以降低或消除上述因素对相干语音侦测系统的影响，因此，为了实现相干语音侦测系统性能最优化，根据理论分析结果并结合实际情况，本文搭建激光相干语音侦测系统（包括全光纤结构和部分光纤结构）时使用的关键光学元器件为：

1) 激光器：根据 2.3.2 节分析可知，激光器的谱线展宽会导致系统载噪比降低从而影响系统性能。为了减少激光器线宽的影响，并结合实际应用场景（侦测距离一般为百米级，从仿真图 2.10 可以发现，此时激光相干语音侦测系统最好选择谱线展宽小于 0.1MHz 的激光器作为光源。），同时为了保证侦测的隐蔽性，本文以输出波长为 1550nm，谱线宽度小于 10kHz，输出功率为 20mW 的单纵模连续波激光器作为系统的光源。

2) 收发一体的聚焦系统：据 2.3.1 节分析可知，光学系统的像差是影响激光相干语音侦测系统的关键因素之一。观察仿真图 2.8 和图 2.9 可以发现，语音侦测系统的性能随光学系统的像差的增加而降低，但是值得注意的是，当光学系统的像差较小时（均方根  $\sigma < 0.1\lambda$ ），其对系统的影响不明显。因此，为了减少像差对系统的干扰，本论文采用自主设计的小像差收发一体的聚焦系统来发射激光和接收回波光。该系统采用伽利略望远镜结构，采用这一结构的原因主要为：1) 将发射光束扩束后再聚焦，可让聚焦光斑能量更集中，使得回波能量更强；2) 可通过调节透镜间的间隔，实现不同距离的聚焦。为了有效消除光学系统球差，该聚焦系统由前后两组透镜组组成，前组负透镜口径为 10mm，后组口径为 100mm，两组透镜之间的间隔可调，从而可将激光聚焦在不同位置。表 2.2 为系统在不同距离下的像差。

表 2.2 不同距离下激光聚焦系统的像差

Table 2.2 The RMS at different distance

Distance/m	10	20	50	100	200
RMS/ $\lambda$	0.0776	0.0194	0.0159	0.0256	0.0311

3) 探测器：根据 2.3.3 节分析可知，与单点探测相比，平衡探测可以有效降

低本征过剩强度噪声，提高系统的载噪比。因此本文选择美国 Thorlabs 公司生产的 PDB570C 型光纤耦合自平衡探测器作为光电探测器。

4) 声光调制器：正如 2.1.2 节所述，声光调制器可让本征光产生稳定的频差从而实现物体振动方向的探测，并且还可以避免低频噪声的影响。本文使用的声光调制器为英国 Gooch&Housego 公司研制的 1550nm 光纤耦合式声光调制器，其频率移动量为 40MHz。

### (3) 实验结果分析

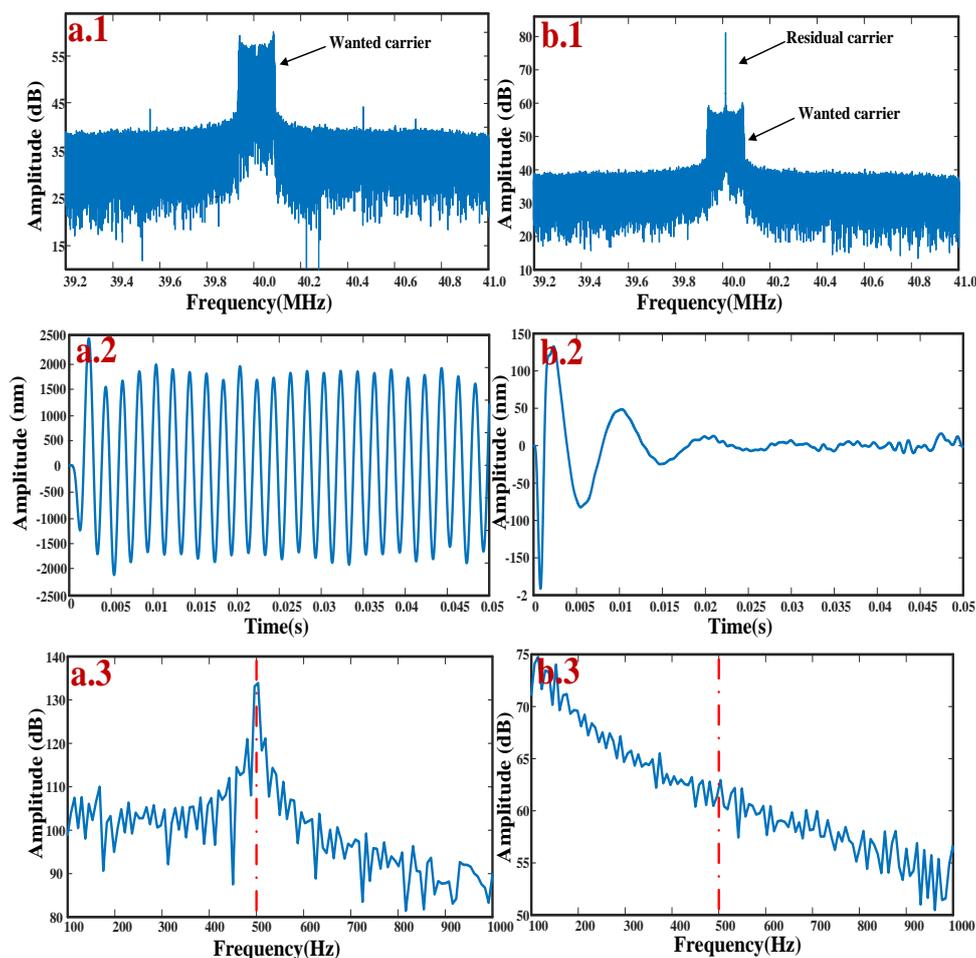


图 2.20 (a) 部分光纤结构实验结果。(b) 全光纤结构实验结果

Figure 2.20 (a) Experimental results of the partial-fiber system. (b) Experimental results of the all-fiber system.

图 2.20.a 与图 2.20.b 分别为使用部分光纤结构与全光纤结构的激光相干语音侦测系统的测量结果图。图中，图 2.20.a.1 与图 2.20.b.1 为探测器输出的载波信号，图 2.20.a.2 和图 2.20.b.2 为从载波信号中解调出的语音信号，其语音解调算

法详见第四章。图 2.20.a.3 和图 2.20.b.3 为语音信号的频谱图。比较图 2.20.a 与图 2.20.b 可以发现，全光纤结构由于隔离度不够，存在串扰，该串扰会严重影响系统获取语音的能力。而部分光纤结构则具有良好的隔离能力可以有效抑制串扰，其探测的语音信号与声源发出的语音信号相同。因此，部分光纤结构性能在远距离探测的情况下要优于全光纤结构。

## 2.5 本章小结

本章内容主要分四个方面：

(1) 介绍了激光相干语音侦测的原理，即采用激光相干测振技术获取由声压引起的物体的振动从而获得语音信号。

(2) 通过理论分析，给出了系统分辨率、探测概率以及系统载噪比的表达式，通过这些参数可以评估激光相干语音侦测系统的性能。

(3) 分析了影响激光相干语音侦测系统性能的光学因素：像差、激光器光谱展宽、探测器探测方式以及发射光串扰。理论与仿真结果表明：1.像差是影响系统性能的一个重要因素，像差越大，系统性能也就越差。在像差较小时（即均方根  $\sigma < 0.1\lambda$ ），系统对像差噪声相对不敏感。然而，当像差较大时，相较于其他类型的像差，倾斜、像散以及慧差对系统影响更大。2.激光器光谱展宽也是影响系统性能的一个重要因素，系统性能随着激光器线宽的增大而减小。由于语音探测的距离一般为百米级，从仿真结果可以发现，系统最好选择谱线展宽小于 0.1MHz 的激光器作为光源。3.相比较于单点探测方式而言，平衡探测可以有效消除过剩本征强度噪声，提高系统的载噪比。4.串扰光强对系统性能也有较大的影响，系统性能会随串扰光强的增强而降低。

(4) 在传统的全光纤激光相干语音侦测系统结构基础上，设计了部分光纤结构，并根据理论分析结果，为系统选择了合适的光学元器件。

## 第3章 振动测量目标的选择

由上一章的分析可知,激光相干语音信号侦测实质就是检测由声源驱动的对象表面的振动,因此选择合适的对象作为测量目标是实现侦测系统获取高质量语音信号的关键。本章通过理论与实验相结合的方法讨论了对象特性对语音相干侦测的影响:首先研究了不同材质的对象表面在声场驱动下的振动特性,然后研究了对象表面粗糙度对语音相干侦测载噪比的影响,最后分析了当被测对象与语音侦测系统之间存在相对运动时产生的散斑噪声对语音相干侦测的影响。以上研究可为激光相干语音侦测系统选择合适的测量目标提供指导。

### 3.1 对象表面在声场的振动特性

激光相干语音信号侦测技术类似于传统麦克风通过检查麦克风膜的振动以获取语言信号,也是通过检测声源周围对象表面的振动以达到语音侦测的目的。因此对象表面对声场的振动响应会直接影响侦测系统获取的语音信号质量。本节通过理论仿真与实验验证的方式研究了常见材质的对象表面在声场的振动特性。

#### 3.1.1 对象表面在声场的振动模型

如图 3.1 所示,为声波和对象之间的相互作用情况。假设对象的结构为由弹性固体介质组成的平板,声源产生的声波为单频信号,对象表面位于  $z=0$  的平面上,厚度为  $h$ ,坐标原点为对象的一角,声源位于  $z$  轴上并且距离对象表面的垂直距离为  $z-h$ ,此时,对象表面的入射声波声压  $P_{in}$  可表示为:

$$P_{in}(x, y, h) = \frac{A_a}{R_0} \exp \left[ ik_a \left( \sqrt{x^2 + y^2} \sin \theta_a - (z - h) \cos \theta_a \right) \right] \quad (3-1)$$

公式 (3-1) 中,  $A_a$  为声压的幅值,  $k_a=2\pi f_a/c_a$  为波数,  $f_a$  为声场频率,  $c_a$  为在空气中的声速,  $\theta_a$  为入射声波与对象表面法向量的夹角,  $R_0=\sqrt{x^2 + y^2 + (z - h)^2}$  为声源和对象表面之间的距离。

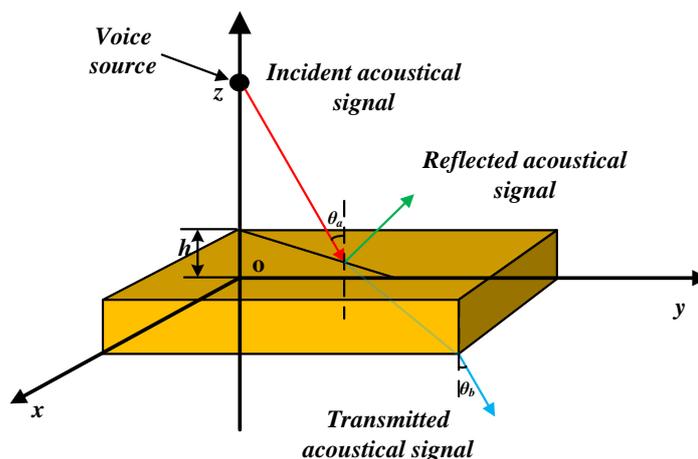


图 3.1 声波与物体之间的相互作用

Figure 3.1 Interaction between acoustical signal and object.

与入射光波类似，入射声波作用在物体表面时，会产生反射声波和透射声波，假设反射率为  $V$ ，透射率为  $T$ ，此时反射声压  $P_r$  和透射声压  $P_t$  可分别表示为：

$$P_r(x, y, h) = \frac{A_a V}{R_0} \exp \left[ ik_a \left( \sqrt{x^2 + y^2} \sin \theta_a + (z - h) \cos \theta_a \right) \right] \quad (3-2)$$

$$P_t(x, y) = \frac{A_a T}{R_0} \exp \left[ ik_a \left( \sqrt{x^2 + y^2} \sin \theta_a - z \cos \theta_b \right) \right] \quad (3-3)$$

公式 (3-3) 中， $\theta_b$  为入射声波与物体表面法向量的夹角。

根据公式 (3-2) 与公式 (3-1) 可得物体表面的总声压  $P_a$  可表示为：

$$\begin{aligned} P_a(x, y, h) &= P_{in}(x, y) + P_r(x, y) \\ &= \frac{A_a}{R_0} \exp \left[ ik_a \left( \sqrt{x^2 + y^2} \sin \theta_a - (z - h) \cos \theta_a \right) \right] \\ &\quad + \frac{A_a V}{R_0} \exp \left[ ik_a \left( \sqrt{x^2 + y^2} \sin \theta_a + (z - h) \cos \theta_a \right) \right] \end{aligned} \quad (3-4)$$

在  $z$  轴方向上，物体表面受到声压激励进行受迫振动，由于振动形式为频率等于  $f_a$  的简谐振动，因此物体表面沿  $z$  轴的振动速度可以写为：

$$v_a(x, y, h) = j2\pi f_a S_a(x, y, h) \quad (3-5)$$

公式 (3-5) 中， $v_a(x, y, h)$  与  $S_a(x, y, h)$  分别为物体表面沿  $z$  轴的振动速度幅度与振动幅度。假设物体表面振动速度满足连续性条件，则沿  $z$  轴的振动速度幅度  $v_a(x, y, h)$  与声压  $P_a$  之间的关系可表示为<sup>[12,25,90]</sup>：

$$v_a(x, y, h) = \frac{1}{-j2\pi f_a \rho_0} \frac{dP_a(x, y, h)}{dz} \quad (3-6)$$

公式(3-6)中,  $\rho_0$  为空气的密度。根据公式(3-4)、公式(3-5)和公式(3-6), 并考虑实际物理含义可得振动幅度  $S_a(x, y, h)$  为:

$$S_a(x, y, h) = \frac{\cos \theta_a}{2\pi f_a \rho_0 c_a} \frac{A_a}{R_0} (1-V) \quad (3-7)$$

物体表面在声场中的振动幅度可通过公式(3-7)表示, 但是由于声压反射率  $V$  的表达式较为复杂, 很难直接获取。但是根据能量守恒定律可得反射率  $V$  和透射率  $T$  存在如下关系

$$V^2 + T^2 = 1 \quad (3-8)$$

同时很多研究表明, 单层平板的声场传输损耗(sound transmission loss, STL)  $L$  与透射率  $T$  之间的关系可表示为<sup>[12,25,91-94]</sup>:

$$L = -20 \lg T \quad (3-9)$$

因此, 通过求解声场传输损耗  $L$  来代替直接求解反射率  $V$ , 从而简化计算物体表面在声场中的振动幅度的难度。在本文中, 声场传输损耗  $L$  通过 Sewell-Sharp-Cremer(SSC)模型<sup>[93]</sup>求解, 该模型被广泛应用于评估声场在单层平板中的传输损耗。在 SSC 模型中, 当声场频率  $f_a$  有不同的取值范围时, 声场传输损耗  $L$  也有不同的解析表达式。根据声场频率  $f_a$  与相干频率 (coincidence frequency)  $f_c$  的比值关系, 通常可分为三段:  $f_a < f_c/2$ ,  $f_c/2 \leq f_a \leq f_c$ , 和  $f_a > f_c$ , 其中  $f_c$  可表示为:

$$f_c = \frac{c_a^2}{2\pi} \sqrt{\frac{12\mu_a(1-\nu_a^2)}{Eh^3}} \quad (3-10)$$

公式(3-10)中,  $\mu_a = \rho h$  为单位面积的质量, 即面密度,  $\rho$  为材料的密度,  $E$  为杨氏模量,  $\nu_a$  为泊松比。

当声场频率  $f_a$  处于第一段频率范围时, 即  $f_a < f_c/2$ , 声场传输损耗  $L$  的表达式为:

$$L(\text{dB}) = -10 \lg \left\{ \frac{\ln(k_a \sqrt{A}) + 0.16 - U(\Lambda) + \frac{1}{4\pi A k_m^2}}{\left[ \left( \frac{\mu_a \pi f_a}{\rho_0 c_a} \right) \left( 1 - \frac{f_a^2}{f_c^2} \right) \right]^2} \right\} \quad (3-11)$$

公式 (3-11) 中,  $k_a=2\pi f_a/c_m$  为在材料中的波数,  $c_m$  为在材料中的声速,  $A$  为平板的面积(平板的面积不能太小),  $\Lambda$  为平板长与宽的比值,  $U(\Lambda)$  为对不规则的平板的形状修正系数, 修正系数  $U(\Lambda)$  可根据经验公式表示为:

$$U(\Lambda) = -0.0000311\Lambda^5 + 0.000941\Lambda^4 - 0.0107\Lambda^3 + 0.0526\Lambda^2 - 0.0407\Lambda - 0.00534 \quad (3-12)$$

当声场频率  $f_a$  处于第三段频率范围时, 即  $f_a > f_c$ , 声场传输损耗  $L$  的表达式为:

$$L(dB) = 20\lg\left(\frac{\mu_a \pi f_a}{\rho_0 c_a}\right) + 10\lg\left(\frac{2\eta_a f_a}{f_c}\right) - 5 \quad (3-13)$$

公式 (3-13) 中,  $\eta_a$  为平板的阻尼损耗系数。

当声场频率  $f_a$  处于第二段频率范围时, 即  $f_c/2 \leq f_a \leq f_c$ , 声场传输损耗  $L$  的表达式为传输损耗  $L$  在声场频率  $f_a=f_c/2$  处与  $f_c$  处之间的线性插值。

根据被测物体的相干频率  $f_c$ , 计算在不同声场频率下的声波的损耗  $L$ , 进而获得声压的反射率  $V$ , 最终求得声波作用下的物体表面的振动幅度。

### 3.1.2 仿真分析与讨论

根据上一节分析可知, 物体表面在声场中的振动幅度不仅仅和声压有关, 还和物体本身的参数: 形状、厚度、密度、杨氏模量、泊松比以及阻尼损耗系数密切相关。因此, 研究物体表面的振动幅度首先需了解被测物体的特性。激光相干语音侦测的应用场景通常在起居室、会议室、实验室或者是公共场所等地, 因此被侦测对象周围的物体大多是生活中常见的物体, 例如: 玻璃杯、纸杯、文件夹、纸盒、铁盒、铝罐等物品。按照材质可大致将这些物品分为五类, 即: 铁、铝、纸(纤维)、塑料以及玻璃。因此, 本章主要讨论由这五种材料组成的平板在声场中的振动幅度。表 3.1 为这五种材料的参数。

表 3.1 常见材料的参数

Table 3.1 Common material parameters

Parameters Materials	Iron	Aluminum	Plastic(PVC)	Glass	Paper (fiber)
Density $\rho$ (kg/m <sup>3</sup> )	7850	2700	1450	2500	900
Poisson ratio $\nu_a$	0.25	0.33	0.38	0.25	0.3
Young's modulus $E$ (Gpa)	206	70	3.14	55	1.5

Damping loss factor $\eta_a$	0.002	0.008	0.04	0.002	0.0177
Speed $c_m$ (m/s)	5200	5000	2388	5639	2591

需要注意的是，表 3.1 给出的只是大致值，由于所含成分的不同，实际物体的参数会与表 3.1 给出的数值有一定的差异。

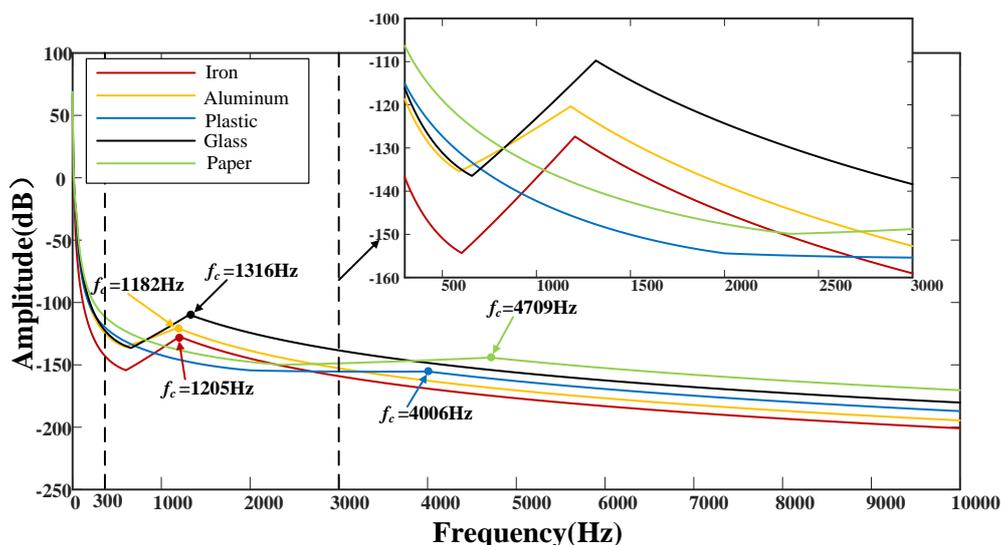


图 3.2 五种材料的振动幅度

Figure 3.2 Vibration amplitude of 5 materials.

图 3.2 为由上述五种材料组成的平板在声场中的振动幅度。在图 3.2 中，假设这五种材料组成的平板的厚度  $h$  均为 10mm，平板的面积  $A$  均为  $0.5 \times 0.5 \text{m}^2$ ，平板长与宽之间的比值  $\Lambda$  均为 1，声压的幅值与距离的比值  $A_d/R_0$  均为 1，入射声波与物体表面法向量的夹角  $\theta_a$  均为 0。观察图 3.2 可以发现：1) 铁、铝和玻璃的相干频率  $f_c$  在人说话的频率范围内 (300-3000Hz)，而纸和塑料的相干频率  $f_c$  则在人说话的频率范围外，这意味着在人说话的频率内，由铁、铝和玻璃这三种材料组成的平板的振动幅度随频率的变化存在较大波动，而由纸和塑料组成的平板的振动幅度则随频率的增加而单调递减。2) 在人说话的低频范围内 (<1000Hz)，纸组成的平板的振动幅度最大，其次由是塑料组成的平板，而在中高频范围内 (>1000Hz) 这两种材料振动幅度则最小，这意味选择塑料和纸制材料作为探测目标，可能会因为中高频部分的语言信号容易被过强的低频响应所淹没，从而造成获得的语音信号的辨识度不够。3) 铁组成的平板在人说话的低频

范围内 (<1000Hz) 振动幅度最小, 由于人声主要能量集中在低频部分, 因此选择铁制物体作为探测目标, 可能会由于物体的振动幅度太小, 导致探测得到的语音信号质量不高。4) 玻璃和铝材料组成的平板则在低频和中高频均有良好的振动响应, 但是由于这两种材料的振动幅度随频率的变化存在较大波动, 这可能会使得探测所得的语音信号的音量发生急剧变化。

从图 3.2 可以发现, 每种材料在声场的振动响应各有优劣, 但是激光相干语音侦测的目的是获得可理解的语音信号, 由于语音频率范围内各个频带的强度是不同的, 强度最高的语音频带处于 300~500Hz, 对于一般通讯而言, 去掉 2000Hz 以下的频率会难以理解语音的意思, 但去掉 2000Hz 以上的频率不会影响对语音的理解<sup>[105]</sup>, 因此从可理解这个角度而言, 纸和塑料这两种材料更适合作为语音侦测的目标物。

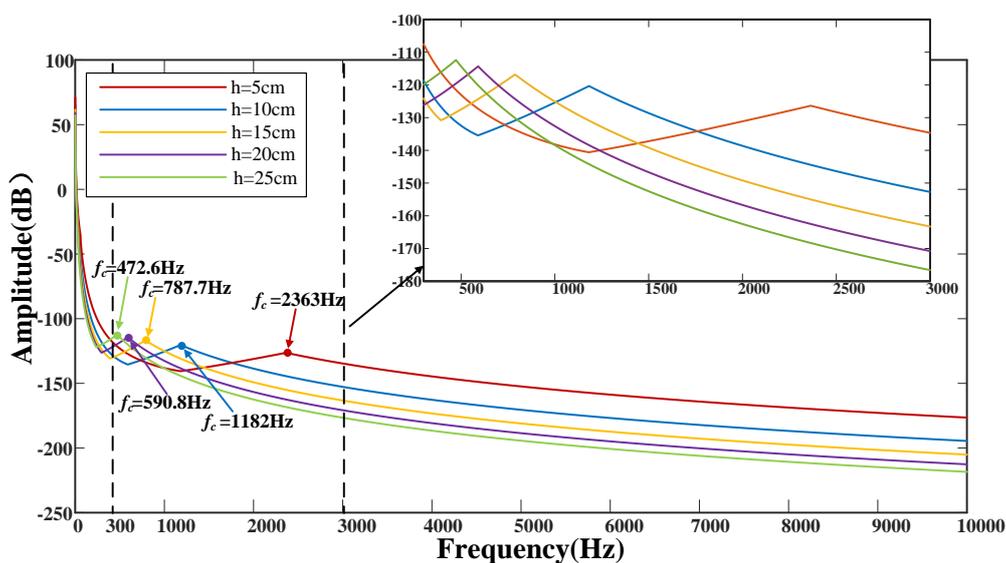


图 3.3 不同厚度下, 铝板的振动幅度

Figure 3.3 Vibration amplitude of aluminum plates in different thicknesses.

由 3.1 节分析可知, 物体的振动幅度除了与材料的性质有关, 还和自身的形状以及厚度有关系。由于 3.1 节的理论研究都是基于平板物体进行研究的, 因此在本章中, 只讨论物体厚度的影响, 不讨论物体形状的影响。图 3.3 为平板的厚度对振动幅度的影响。在图 3.3 中, 选择了铝制平板作为代表进行研究, 除了厚度  $h$  参数改变外, 其余参数都与图 3.2 中假设一致。观察图 3.3 可以发现, 随着铝板厚度的增加, 其相干频率  $f_c$  也随之降低, 同时, 平板的振动幅值也整体降低, 但是值得注意的是, 不同厚度物体的振幅曲线具有相同的形状。这表明物体的振

动幅度曲线与厚度无关，仅与物体的材料特性有关，但是物体的振动幅度会随着厚度增加而整体下降。因此，本课题采用物体的构成材料作为选择合适测量目标的重要评判依据。

### 3.1.3 实验分析与讨论

本章的前几节通过理论分析和数值仿真研究了不同材料组成的平板在声场中的振动特性。本节将通过实验实测不同材质的平板在声场的振动特性。如图 3.4.a 所示为实验原理图，本实验系统可分为两部分：测振部分与被测部分。本实验的测振部分为德国 Polytec 公司生产的 PSV-500 型激光相干多普勒测振仪（如图 3.4.b 所示），该测振仪的测量频率范围为 0~50KHz，速度分辨率为  $0.04\sim 0.5\mu\text{m s}^{-1}\sqrt{\text{Hz}}$ ，位移分辨率约为 0.3nm，该性能足以满足振动测量实验的要求。被测部分主要由信号发生器、音箱以及不同材质组成的平板（如图 3.4.c 所示）组成。音箱作为声源，信号发生器用来产生频率范围为 300-3000Hz，频率间隔为 100Hz 的正弦信号以驱动音箱发出单音信号（不同频率的正弦信号其电压峰值都相同以确保每个单音信号的强度相同），不同材质组成的平板则作为振动测量目标。因为本实验主要的目标是为了检测物体的振动特性，因此为了确保回波光能量足够强，在实验中，激光相干测振仪与被测目标之间的距离为 1m。同时，在本次实验中，各材质平板的尺寸如表 3.2 所示。

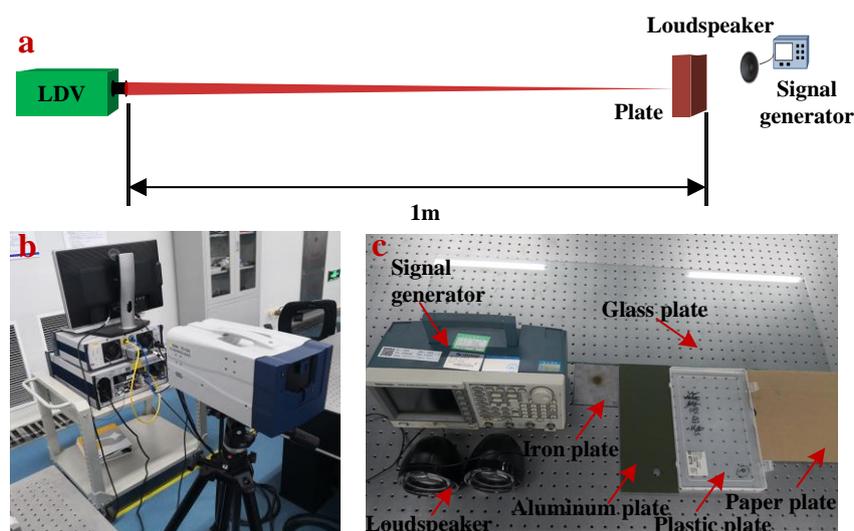


图 3.4 (a) 实验原理图。 (b) Polytec 测振仪。 (c) 被测部分

Figure 3.4 (a) Experimental setup. (b) The Polytec LDV. (c) Test component.

表 3.2 平板的尺寸

Table 3.2 Size of plates

<b>Parameters Materials</b>	Iron	Aluminum	Plastic(PVC)	Glass	Paper (fiber)
Area $A(\text{cm} \times \text{cm})$	10.0 × 10.0	28.7 × 11.2	26.8 × 18.5	95.2 × 29.3	24 × 21
Thickness $h(\text{mm})$	3.5	4.0	2.0	2.2	1.0

图 3.5a~3.5e 分别为铁（钢）制平板、铝制平板、塑料制平板、玻璃制平板以及纸制平板在声场中的振动幅度图。从图中可以发现：1）不同材质的平板在声场中的实测振幅与理论振幅有所不同，这可能是由于实际材料的参数与表 3.1 所列的参数存在差异，以及音箱的非线性效应（音箱的频率响应在语音范围内不均匀）导致每个频率的单音的声压略有区别，和被测物体的边界效应等诸多因素导致的，但是实测振幅的变化趋势与理论振幅的变化趋势是一致的。2）实测结果表明，在入射声压相同的情况下，纸制平板、塑料制平板和玻璃制平板在低频的振动幅度比其余两种材质的平板都要大，但是这三种平板对中高频的振动响应要比另外两种材质的平板差很多。3）相比纸制、塑料制和玻璃制材料，铝制平板与铁（钢）制平板对于频率变化的敏感度要低，它们在各个频率声场的振动幅度相差不大，但是这两种材质平板的整体振动幅度要小于另外三种材质。总体而言，由于激光相干语音侦测的目的是获取可理解语音，虽然纸制与塑料制材质在中高频振动响应较差，但是由于去除中高频信息对于语音的理解几乎没有影响，因此纸制与塑料制材料更适合作为语音获取的目标，这一结论与理论分析结果相一致。该研究结果为激光相干语音侦测系统选择合适的振动物体提供了有效的指导。

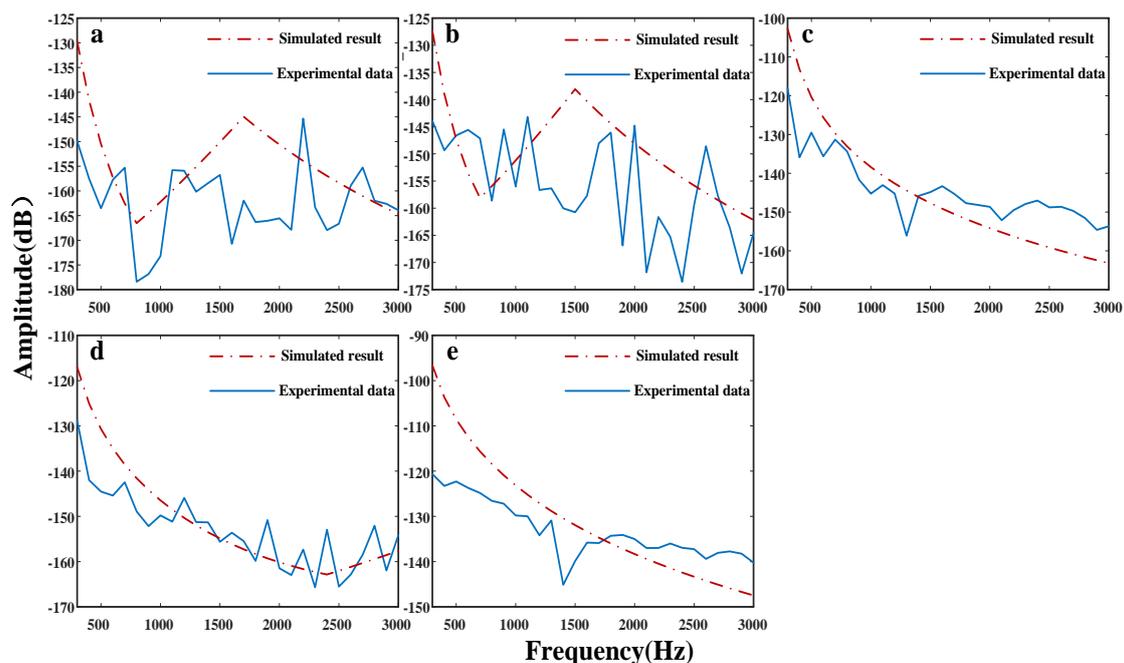


图 3.5 五种材料的振动幅度。(a) 铁(钢)板的振动幅度。(b) 铝板的振动幅度。(c) 塑料板的振动幅度。(d) 玻璃板的振动幅度。(e) 纸板的振动幅度

Figure 3.5 Vibration amplitude of 5 materials. (a) Vibration amplitude of iron (steel) plate. (b) Vibration amplitude of Aluminum plate. (c) Vibration amplitude of plastic plate. (d) Vibration amplitude of glass plate. (e) Vibration amplitude of paper plate.

### 3.2 物体表面粗糙度对激光相干语音侦测系统的影响

在实际的激光相干语音侦测过程中, 由于被测物体大多数属于漫反射物体, 其粗糙的表面一方面会使得激光发射散射不利于回波光的接收, 另一方面因为其凹凸不平的表面造成回波光波前产生畸变, 破坏其与本征光的相位匹配, 从而损害激光相干语音侦测系统的性能。因此研究物体表面粗糙度对相干语音侦测的影响, 有助于提出抑制或者改善物体表面粗糙度对激光相干语音侦测系统性能影响的方案。如 2.2 节所述, 激光语音侦测系统的性能可用系统载噪比 CNR 衡量, 因此在分析表面粗糙度对语音相干侦测系统的影响时, 只要分析表面粗糙度对系统载噪比的影响即可。

## 3.2.1 理论分析粗糙面对载噪比的影响

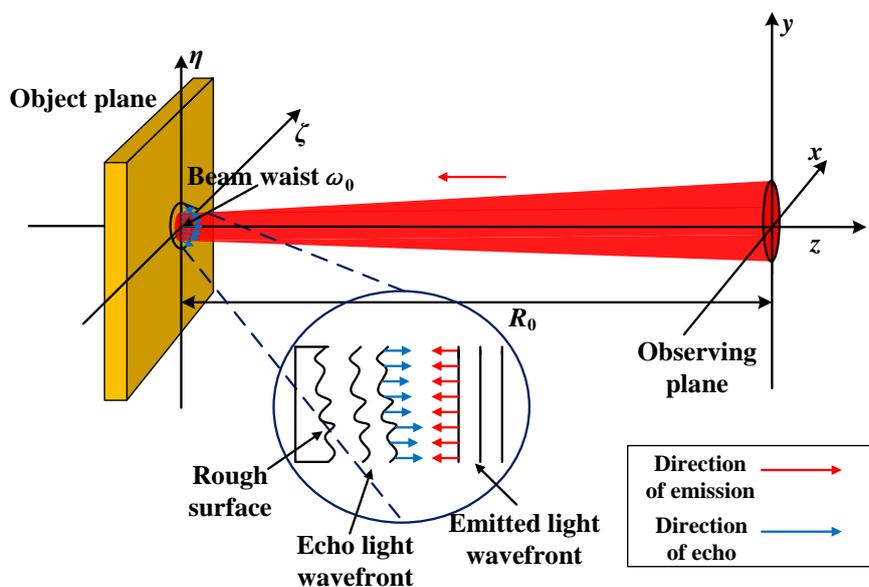


图 3.6 激光聚焦在粗糙表面示意图

Figure 3.6 Schematic diagram of laser focusing on a rough surface.

正如 2.1 节所述，对于激光相干语音侦测系统而言，其侦测的过程可简述为携带由物体振动而引起的多普勒频移的回波光与本征光在光电探测器的光敏面上发生混频，其混频信号通过探测器输出为中频电信号，通过解调中频电信号得到多普勒频移，根据多普勒频移即可获得物体的振动信息。回波光为发射光照射在物体表面散射回来的光。如图 3.6 所示，由于物体表面凹凸不平，回波光会发生相位延迟，其相位延迟  $\Delta\varphi_{de}$  可表示为：

$$\Delta\varphi_{de} = \frac{4\pi}{\lambda} Z(x, y) \quad (3-14)$$

式 (3-14) 中， $Z(x,y)$  为起伏的目标表面。此时回波光的复振幅  $E_s(x,y)$  根据惠更斯-菲涅耳定律<sup>[95]</sup>表示为：

$$E_s(x, y) = \iint E_o(\zeta, \eta) \exp(i\Delta\varphi_{de}) h(x-\zeta, y-\eta) d\zeta d\eta \quad (3-15)$$

式 (3-15) 中， $E_o(\zeta, \eta)$  为入射到粗糙表面上的复振幅。由于激光器发射的光束为高斯光束，其应用的模式大多为  $TEM_{00}$  模，并且高斯光束经过聚焦透镜聚焦后仍为高斯光束，在实际探测中，往往将发射激光聚焦在粗糙表面处，因此假设经过透镜后的束腰半径为  $\omega_0$ ，并且束腰位于粗糙表面上，则入射到粗糙表面上的复振幅  $E_o(\zeta, \eta)$  可表示为：

$$E_o(\zeta, \eta) = E_o \exp\left(-\frac{\zeta^2 + \eta^2}{\omega_0^2}\right) \quad (3-16)$$

$h(x-\zeta, y-\eta)$ 为传播函数, 可表示为:

$$h(x-\zeta, y-\eta) = \frac{1}{j\lambda r} \exp(ikr) \quad (3-17)$$

式(3-17)中,  $r$ 为观察面上的点与目标面上点之间的距离。因为在实际探测中, 目标物与观察面之间满足远场条件, 因此 $r$ 可以表示为:

$$r = R_0 + \frac{x^2 + y^2}{2R_0} - \frac{x\zeta + y\eta}{R_0} \quad (3-18)$$

式(3-18)中,  $R_0$ 为观察面与目标面之间的距离。

根据公式(3-16)至公式(3-18), 回波光的复振幅 $E_s(x, y)$ 可表示为:

$$E_s(x, y) = \frac{1}{i\lambda R_0} \exp(ikR_0) \exp\left[i\frac{k}{2R_0}(x^2 + y^2)\right] \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E_o \exp\left(-\frac{\zeta^2 + \eta^2}{\omega_0^2}\right) \times \exp(i\Delta\phi_{de}) \exp\left(-ik\frac{x\zeta + y\eta}{R_0}\right) d\zeta d\eta \quad (3-19)$$

通过2.2节可知, 粗糙面对激光相干语音侦测系统的影响可通过其对系统载噪比的影响进行衡量。假设本征光为无波前畸变的理想平面波, 其振幅 $E_{Lo}(x, y)=1$ , 并假设回波光与本征光的传播方向相同, 均垂直入射在探测器表面, 则根据公式(2-45)可得系统载噪比为:

$$CNR = \frac{\beta_0}{eB\Delta S} \left( \iint E_s(x, y) dx dy \right)^2 \quad (3-20)$$

式(3-20)中,  $\Delta S$ 为探测器面积。

从公式(3-14)和公式(3-19)可以发现, 粗糙表面的高低起伏会引起回波光的空间相位发生畸变, 导致回波光与本征光混频时混频信号的相位发生变化。若混频信号的相位差为 $\pi$ 的整数倍时, 会使得中频光电流信号为0, 从而使系统的载噪比下降, 影响系统性能。

### 3.2.2 粗糙面对载噪比的影响的数值仿真

#### (1) 粗糙面仿真

为了分析粗糙表面对载噪比的影响，首先必须仿真生成一个粗糙表面，该表面既要与实际粗糙面相似，又必须可以用数学表达式描述。由于大多数物体表面起伏分布满足均值为零的高斯随机分布，因此，本文采用蒙特卡罗方法仿真粗糙表面<sup>[96~98]</sup>，其仿真过程可以简述为先对粗糙面的功率谱密度进行滤波，然后将其通过快速逆傅立叶变换从而得到物体粗糙面的高低起伏函数。因此在仿真生成粗糙表面前，应先获得粗糙表面的功率谱密度。

假设目标表面起伏函数  $Z(x,y)$  的均方根为  $\delta_r$ ，横向相关长度为  $l_{rx}$ ，纵向相关长度为  $l_{ry}$ ，由于起伏函数  $Z(x,y)$  为均值为零的高斯随机分布，因此其相关函数  $G(R_x,R_y)$  为：

$$G(R_x, R_y) = \delta_r^2 \exp\left(-\frac{R_x^2}{l_{rx}} - \frac{R_y^2}{l_{ry}}\right) \quad (3-21)$$

式 (3-21) 中， $R_x$  为粗糙面两点在横向的距离， $R_y$  为粗糙面两点在纵向的距离。通过维纳-辛钦定理可得粗糙面的功率谱密度  $P_{ro}(k_x,k_y)$  为：

$$P_{ro}(k_x, k_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(R_x, R_y) \exp[i(k_x R_x + k_y R_y)] = \delta_r^2 \frac{l_{rx} l_{ry}}{4\pi} \exp\left(-\frac{k_x^2 l_{rx}^2 + k_y^2 l_{ry}^2}{4}\right) \quad (3-22)$$

在蒙特卡罗仿真中，粗糙表面被认为是由大量谐波相互累加而形成的，每一个谐波相互独立，且都服从高斯分布，其方差正比于对应谐波的功率谱密度。因此假设仿真生成的粗糙表面尺寸为  $L_x \times L_y$ ，由  $M \times N$  个点组成，目标表面起伏函数  $Z(x,y)$  可表示为：

$$Z(x, y) = \frac{1}{L_x L_y} \sum_{m_k=-M/2+1}^{M/2} \sum_{n_k=-N/2+1}^{N/2} F(k_{m_k}, k_{n_k}) \exp[i(k_{m_k} x + k_{n_k} y)] \quad (3-23)$$

式(3-23)中， $x=m\Delta x=mL_x/M$ ， $y=n\Delta y=nL_y/N$ ，其中  $m=-M/2+1 \dots M/2$ ， $n=-N/2+1 \dots N/2$ 。并且  $k_{mk}=2\pi m_k/L_x$ ， $k_{nk}=2\pi n_k/L_y$ 。

$$F(k_{m_k}, k_{n_k}) = 2\pi \left[ L_x L_y P_{ro}(k_{m_k}, k_{n_k}) \right]^{1/2} \begin{cases} \frac{[N(0,1) + iN(0,1)]}{\sqrt{2}}, & m_k \neq 0, M/2 \cup n_k \neq 0, N/2 \\ N(0,1), & \text{others} \end{cases} \quad (3-24)$$

为了使目标表面起伏函数  $Z(x,y)$  为实数，傅立叶系数  $F(k_{mk}, k_{nk})$  必须满足：

$$F(k_{m_k}, k_{n_k}) = F^*(-k_{m_k}, -k_{n_k}), \quad F(k_{m_k}, -k_{n_k}) = F^*(-k_{m_k}, k_{n_k}) \quad (3-25)$$

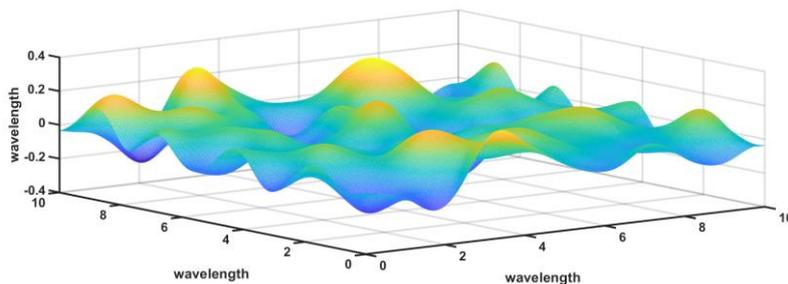


图 3.7  $\delta_r=0.1\lambda$ ,  $l_{rx}=l_{ry}=\lambda$

Figure 3.7  $\delta_r=0.1\lambda$ ,  $l_{rx}=l_{ry}=\lambda$ .

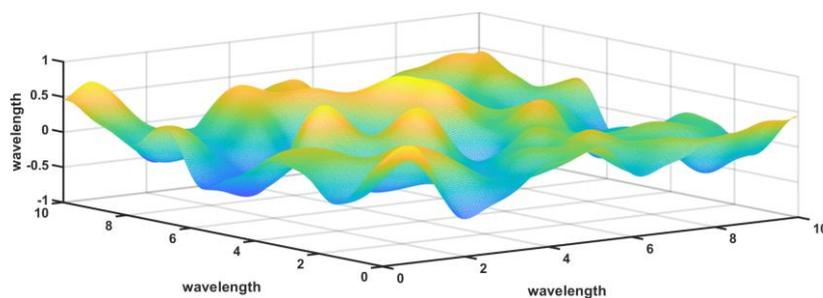


图 3.8  $\delta_r=0.3\lambda$ ,  $l_{rx}=l_{ry}=\lambda$

Figure 3.8  $\delta_r=0.3\lambda$ ,  $l_{rx}=l_{ry}=\lambda$ .

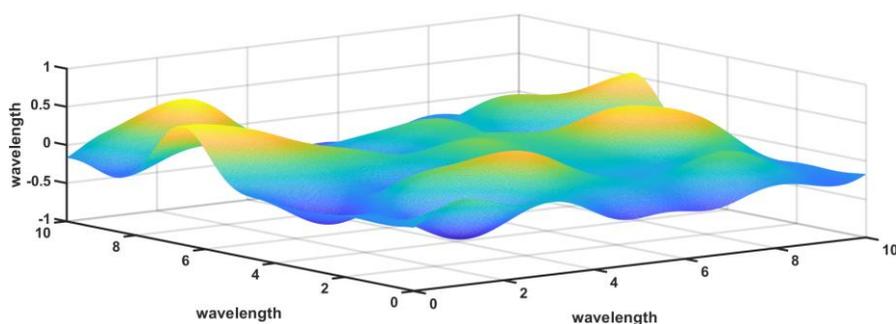


图 3.9  $\delta_r=0.3\lambda$ ,  $l_{rx}=l_{ry}=1.5\lambda$

Figure 3.9  $\delta_r=0.3\lambda$ ,  $l_{rx}=l_{ry}=1.5\lambda$ .

图 3.7, 图 3.8 与图 3.9 为不同均方根和相关长度下, 仿真生成的粗糙表面。比较图 3.7 与图 3.8 可以发现, 在相关长度一定的情况下, 物体表面粗糙度随均

方根的增加而增加，比较图 3.8 与图 3.9 可以发现，在均方根一定的情况下，物体表面粗糙度随着相关长度的增大而减小。这说明均方根越大，相关长度越小，则物体表面越粗糙，仿真结果与实际情况相对应。在现实中，孙鹤、郑小兵等人<sup>[99~100]</sup>以均方根和相关长度作为评判标准，将物体表面分为微粗糙面、弱粗糙面、强粗糙面以及极粗糙面，具体如表 3.3 所示。

表 3.3 表面粗糙程度的定性划分

Table 3.3 The qualitative division of the degree of surface roughness

	Micro-rough surface	small-rough surface	Strong-rough surface	Tremendous-rough surface
<b>Parameter</b>	$\delta_r < 0.1\lambda, l_r \gg \lambda$	$0.1\lambda < \delta_r < 0.3\lambda, l_r \gg \lambda$	$0.3\lambda < \delta_r < \lambda, l_r \approx \lambda$	$\delta_r > \lambda, l_r < \lambda$

(2) 粗糙面对载噪比影响的数值仿真

为了精确仿真粗糙面对载噪比的影响，先将回波光的复振幅  $E_s(x,y)$  离散化。假设将粗糙面均匀划分为  $M_r \times N_r$  块相同的面元，横向与纵向上的间隔分别为  $\Delta x_r$  和  $\Delta y_r$ ，将探测器的光敏面划分为  $M_d \times N_d$  块相同的面元，横向与纵向上的间隔分别为  $\Delta x_d$  和  $\Delta y_d$ ，此时，探测器任一区域的回波光复振幅  $E_s(x,y)$  为：

$$E_s(x,y) = E_s(m_d \Delta x_d, n_d \Delta y_d) = \frac{1}{i\lambda R_0} \exp(ikR_0) \exp\left[i \frac{k}{2R_0} (m_d^2 \Delta x_d^2 + n_d^2 \Delta y_d^2)\right] \times \sum_{m_r=-M_r/2+1}^{M_r/2} \sum_{n_r=-N_r/2+1}^{N_r/2} E_0 \exp\left(-\frac{m_r^2 \Delta x_r^2 + n_r^2 \Delta y_r^2}{\omega_0^2}\right) \exp[i\Delta\phi_{de}(m_r \Delta x_r, n_r \Delta y_r)] \times \exp\left(-ik \frac{m_d \Delta x_d m_r \Delta x_r + n_d \Delta y_d n_r \Delta y_r}{R_0}\right) \Delta x_r \Delta y_r \quad (3-26)$$

根据公式 (3-20) 与公式 (3-26) 可得系统载噪比为：

$$CNR = \frac{\beta_0}{eB\Delta S} \left( \sum_{m_d=-M_d/2+1}^{M_d/2} \sum_{n_d=-N_d/2+1}^{N_d/2} E_s(m_d \Delta x_d, n_d \Delta y_d) \Delta x_d \Delta y_d \right)^2 \quad (3-27)$$

根据公式 (3-27) 就可以计算出当被测物体表面为粗糙表面时系统的载噪比。

图 3.10a 和图 3.10b 分别为载噪比与粗糙表面相关长度以及均方根的关系。为了消除随机误差，提高仿真的准确性，在仿真程序计算过程中，每个载噪比都是由其对应的相关长度或者是均方根运算 1000 次后的和的均值。由于计算机性能的限制，在图 3.10a 与图 3.10b 中，均假设光波长  $\lambda=1\text{mm}$ ，粗糙面划分为

1024×1024 块区域,横向与纵向上的间隔  $\Delta x_r$  和  $\Delta y_r$  均为  $0.01\lambda$ ,探测距离为  $0.01\text{m}$ ,探测器面积为  $0.01\pi\text{m}^2$ ,并将其划分约为  $102\times 102$  块区域,  $(\Delta x_r\Delta y_r)^2\Delta S\beta_0/eB$  为 1, 高斯光束束腰半径  $\omega_0$  等于  $0.35\text{mm}$ , 入射到粗糙表面上的复振幅  $E_o$  等于 1。

在图 3.10a 中,均方根  $\delta_r=0.3\lambda$ ,观察图 3.10a 可以发现,在相关长度  $(l_{rx},l_{ry})<1.2\lambda$  时,系统载噪比随相关长度的增加而增加,但是增长速度随相关长度的增加而下降。这是由于相关长度是衡量粗糙表面上具有相关性的两点之间间隔的指标,在相关长度范围内的点所反射的回波相位是相关的,所对应的光程差也是一致的,因此不会造成回波光相位畸变.但是当粗糙表面的点的间隔大于相关长度时,此时,这些点彼此之间相互独立,因此反射的回波相位也彼此独立,这就容易导致中频光电流相互抵消,从而影响系统的载噪比。当  $(l_{rx},l_{ry}) >1.2\lambda$  时,由于此时目标表面已经接近光滑,系统载噪比基本稳定在 0.2 左右。

在图 3.10b 中,相关长度  $(l_{rx},l_{ry})= 0.3\lambda$ ,观察图 3.10b 可以发现,在均方根  $\delta_r<3\lambda$  时,系统载噪比随均方根  $\delta_r$  的增加而急剧下降。这是因为随着粗糙表面均方根值的变大,导致粗糙面表面起伏越剧烈,这就更容易使回波光与本征光的相位差值产生  $\pi$  的倍数,使得中频光电流相互抵消,最终影响系统的载噪比。当均方根  $\delta_r>3\lambda$  时,系统载噪比在 0.0015 附近处做随机波动,这可能是由于表面的随机起伏引起的。这说明当粗糙表面均方根到达一定程度以后,系统载噪比与均方根无明显关系。

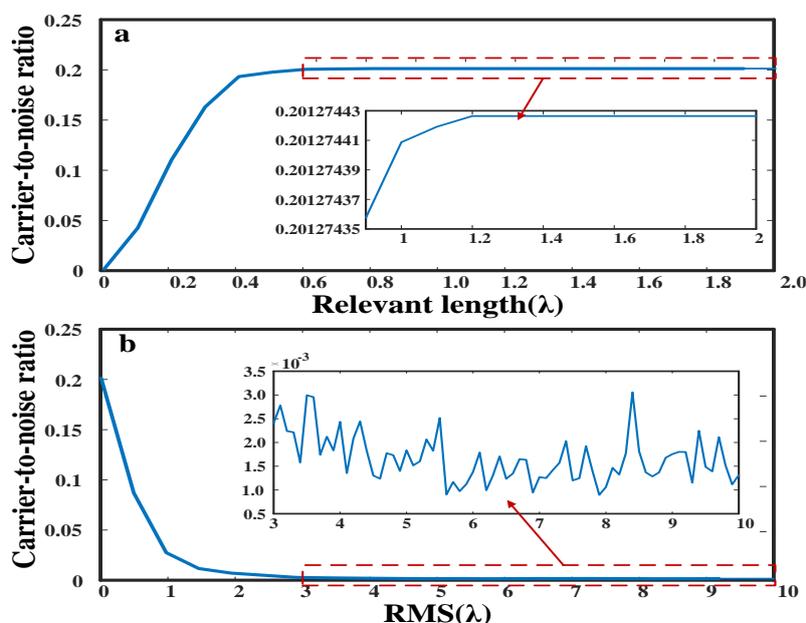


图 3.10 载噪比与表面粗糙度的关系。(a)载噪比和相关长度关系。(b)载噪比和均方根关系  
Figure 3.10 Relationship between carrier-to-noise ratio and surface roughness. (a) Relationship between carrier-to-noise ratio and relevant length.(b) Relationship between carrier-to-noise ratio and RMS.

图 3.11 为在粗糙表面相关长度以及均方根均为变量的情况下, 载噪比的变化情况。此处的参数与图 3.10 相同。从图 3.11 可以看出, 相较于相关长度( $l_{rx}, l_{ry}$ ), 系统载噪比对于均方根  $\delta_r$  的变化更敏感。但是总体而言, 系统载噪比随物体表面粗糙度的增加而减小, 而当表面粗糙度到达一定程度以后, 系统载噪比与表面粗糙度无明显关系, 只在某一定值处做随机波动。

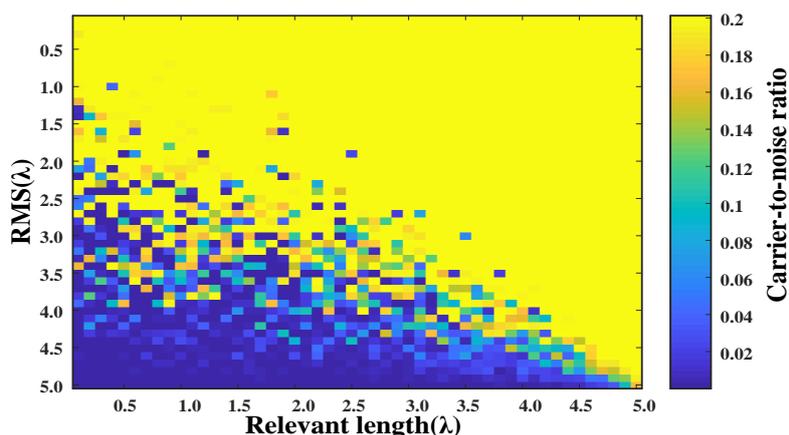


图 3.11 载噪比与相关长度以及均方根值的关系

Figure 3.11 Relationship between carrier-to-noise ratio and relevant length and RMS.

### 3.2.3 实验研究粗糙面对载噪比的影响

理论与仿真结果表明, 由于粗糙表面凹凸程度不同, 会使得回波光光程不同, 进而导致回波光产生波前畸变, 影响与本征光的混频效率, 最终降低激光相干语音侦测系统载噪比。语音侦测系统的载噪比随着表面粗糙度的增加而下降, 但是当表面粗糙度到达一定值以后, 系统的载噪比则与表面粗糙度无明显关系, 只在某定值附近做随机波动。为了验证表面粗糙度对于系统载噪比影响的分析结果的正确性, 本节进行了不同粗糙表面的相干探测实验, 如图 3.12.a 为实验设计图。本实验由探测部分以及被测目标两部分组成。

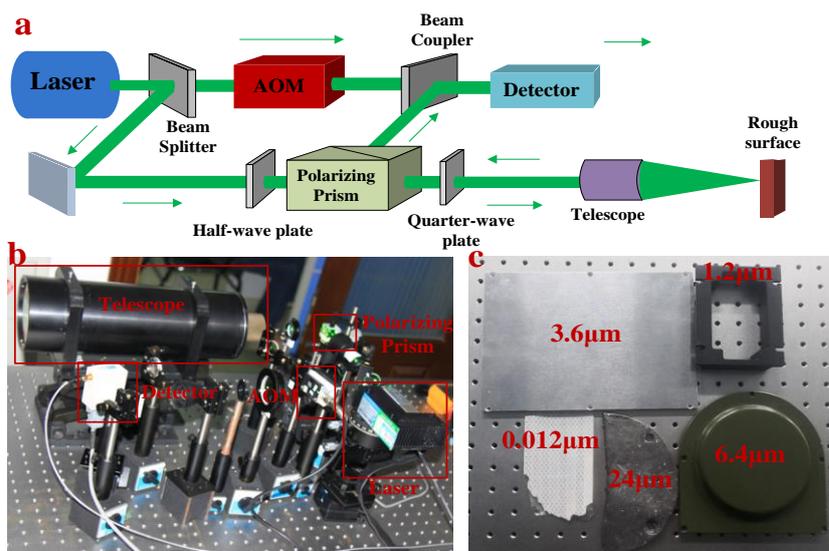


图 3.12 (a)实验原理图。(b)相干多普勒测振仪。(c) 被测部分

Figure 3.12 (a) Experimental setup. (b) The Coherent Doppler vibrometer. (c) Test component.

正如 2.1.1 节所述,在本文中激光相干语音侦测技术就是激光相干测振技术,实验研究粗糙表面对于激光相干语音侦测系统的载噪比的影响可等效为研究粗糙表面对于相干测振系统载噪比的影响,因此本文基于改进型马赫-曾德尔干涉仪结构搭建了一套的常见的激光相干测振系统作为实验的探测部分,如图 3.12.b 所示。在该系统中,激光器为法国 Oxxius 公司生产的窄线宽激光器,波长为 532nm,波长稳定性为 1pm,线宽小于 1MHz,功率可调,最大输出功率为 100mw。声光移频器为英国 Gooch&Housego 公司研制的,其移频为 80MHz。光电探测器为日本滨松(HAMAMATSU)公司生产的 C5658 型光电探测器,其探测带宽为 1GHz。为了获取载波强度以及噪声强度,采用美国 NI 公司生产的 PXIe-5162 高速采集卡将光电探测器输出的信号转换为数字信号,并存入计算机中,方便进一步计算分析。由于载波频率为 80MHz,因此为了保证采样得到的数字信号可以完整的保留光电探测器输出的信息,使用的采集卡采样率为 500MHz。在本实验中,载波强度定义为光电探测器输出的信号在中心频率为 80MHz,带宽为 1KHz 处频谱的强度,噪声强度则定义为当激光相干测振仪前无目标时,光电探测器输出的信号在中心频率为 80MHz,带宽为 1KHz 处频谱的强度。由于受实验条件的限制,本实验的被测目标由轮廓算术平均偏差 Ra 约为 0.012 $\mu\text{m}$ 、1.2 $\mu\text{m}$ 、3.6 $\mu\text{m}$ 、6.4 $\mu\text{m}$  以及 24 $\mu\text{m}$  这五个粗糙物块组成(如图 3.12.c 所示)。其中轮廓算术平均偏

差指的是在取样长度内被测轮廓偏离绝对值的算数均值,该参数用于描述表面粗糙度。

在实验中,每块测量物体与相干探测系统之间的距离都相同,激光发射光功率都为 20mw,本征光功率都为 5mw。同时为了消除实验过程中的随机误差,在每块被测物体上的不同位置处采集 20 个点,系统的载噪比为 20 个采样点的均值。

图 3.13 为不同粗糙度下系统的载噪比。观察图 3.13 可以发现,系统的载噪比随着表面粗糙度的增加而下降,但是当表面粗糙度到达一定值以后 ( $Ra > 1.2\mu\text{m}$ ),系统的载噪比则在 20.5dB 附近做随机波动。比较图 3.13 与仿真图 3.10.b,可以发现,实验中载噪比随表面粗糙度的变化趋势和仿真结果十分接近,这表明,当表面粗糙度超过  $3\lambda$  时,再纠结物体的表面粗糙度就没有意义了。该研究结果为激光相干语音侦测系统选择合适的振动表面提供了有效的指导。

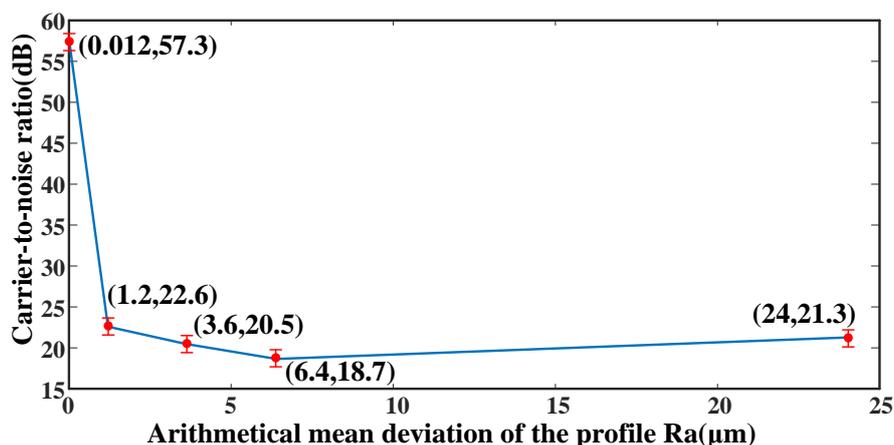


图 3.13 载噪比与表面粗糙度的关系

Figure 3.13 Relationship between carrier-to-noise ratio and roughness.

### 3.3 散斑噪声对激光相干语音侦测系统的影响

根据 3.2 节可知,当相干激光光束入射在粗糙表面时,粗糙表面会引起回波光的波前相位发生畸变,回波光的波前相位畸变不但导致与本征光的混频效率降低,从而降低系统的载噪比,还会使得回波光彼此之间发生建设性和破坏性的干涉,从而导致回波光斑呈现为亮斑和暗斑随机分布的形状。如图 3.14 所示,这些亮斑和暗斑随机分布的回波光斑具有颗粒状斑点外观,因此称之为散斑<sup>[10]</sup>。

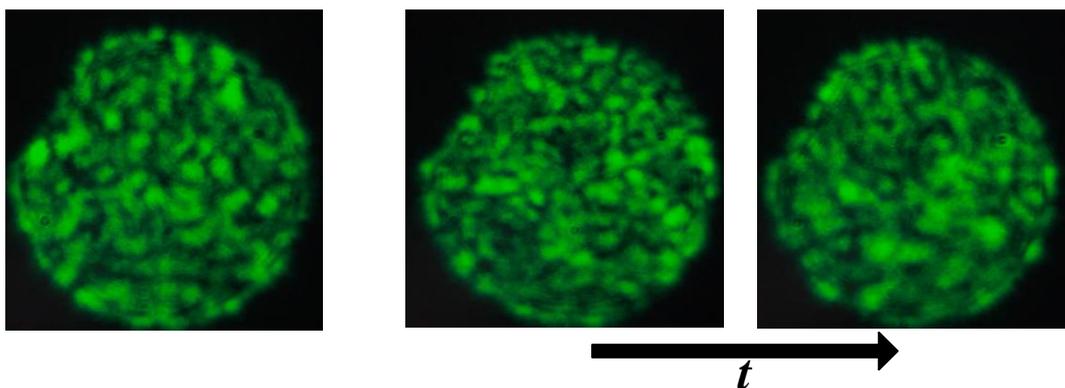


图 3.14 散斑

图 3.15 动态散斑

Figure 3.14 Speckles.

Figure 3.15 Dynamic speckles.

在实际应用中，由于环境干扰以及人为等因素，可能会导致被测物体与激光相干语音侦测系统之间存在除光轴方向外的相对运动。这种相对运动会改变出射光斑在粗糙表面处的位置，从而使得回波散斑模式发生变化（这种情况被称之为动态散斑，如图 3.15 所示）。由于散斑的模式随时间发生变化，因此回波光的振幅  $E_s$  也随时间的变化而发生变化，与此同时，由于粗糙表面而引入的相位畸变  $\Delta\varphi_{de}$  也会随时间的变化而发生变化。此时，中频电流公式 (2-7) 改写为：

$$i_{IF}(t) = 2\text{Re}\left(\iint \beta(x, y) \left\{ E_{Lo}(x, y) E_s^*(x, y, t) \exp\left[i(\omega_d + \omega_{if})t\right] \exp(i\Delta\varphi_{de}(x, y, t)) \right\} dx dy\right) \\ = I_{res}(t) \cos\left[\left(\omega_d + \omega_{if}\right)t + \Phi_{res}(t)\right] \quad (3-28)$$

公式 (3-28) 中， $I_{res}(t)$  和  $\Phi_{res}(t)$  分别为由于散斑模式的改变而引起的随时间变化的中频电流的幅度与相位。由于语音信号是通过解调中频电流  $i_{IF}(t)$  获得的，中频电流  $i_{IF}(t)$  的质量直接影响侦测过程。根据公式 (3-28) 可知，当散斑模式随时间变化而改变时，会导致中频电流  $i_{IF}(t)$  产生幅度调制  $I_{res}(t)$  和相位调制  $\Phi_{res}(t)$ ，这使得侦测的不确定性大大增加。将这种由于散斑模式的改变而引入的噪声称之为散斑噪声。特别强调的是，当散斑模式不随时间变化时，此时散斑对侦测系统的影响仅为降低系统的载噪比，这点已在 3.2 节进行了详细的研究。本节主要研究的是动态散斑引起的噪声对相干语音侦测的影响，这项研究有助于提出抑制或者改善动态散斑对激光相干语音侦测系统性能的影响的方案。

### 3.3.1 散斑特性

正如上节所述，散斑模式是由光学粗糙表面的相干光散射形成的。由于粗糙表面的微观特性是未知的，因此散斑模式的形成一般可以被认为是随机过程，所以采用统计特性描述散斑的特性。

#### (1) 散斑强度与相位的概率密度函数

散斑可简单被认为是许多基本散射单元相加的结果，因此，观察面上任一点散斑的复振幅  $E_s$  和相位  $\Phi$  可以用  $P$  个基本散射单元的相加表示<sup>[60]</sup>

$$E_s \exp(j\Phi) = \sum_{p=1}^P e_p \exp(j\varphi_p) \quad (3-29)$$

公式 (3-29) 中， $e_p$  和  $\varphi_p$  分别为第  $p$  个基本散射单元的振幅和相位。如图 3.16 所示，公式 (3-29) 中所示的总和可表示为在复平面下，各个散射单元随机累加的过程。

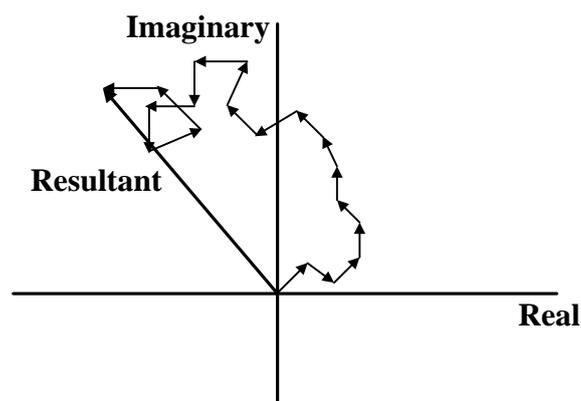


图 3.16 复平面中随机的相量累加

Figure 3.16 Random phasor accumulation in complex plane.

假设基本散射单元的振幅都为  $e$ ，此时散斑的复振幅  $E_s$  可表示为：

$$\begin{aligned} E_s &= e \left\{ \left[ \sum_{p=1}^P \cos \varphi_p \right]^2 + e^2 \left[ \sum_{p=1}^P \sin \varphi_p \right]^2 \right\}^{1/2} \\ &= e \left[ \sum_{p=1}^P \sum_{m=1}^P \cos \varphi_p \cos \varphi_m + \sum_{p=1}^P \sum_{m=1}^P \sin \varphi_p \sin \varphi_m \right]^{1/2} \\ &= e \left[ \sum_{p=1}^P \sum_{m=1}^P \cos(\varphi_p - \varphi_m) \right]^{1/2} \end{aligned} \quad (3-30)$$

相位  $\Phi$  可表示为:

$$\Phi = \arctan \left( \frac{\sum_{p=1}^P \sin \varphi_p}{\sum_{p=1}^P \cos \varphi_p} \right) \quad (3-31)$$

粗糙表面使得散射光的相位  $\varphi_p$  在区间  $(-\pi, \pi)$  上均匀分布, 因此, 散斑的相位  $\Phi$  在区间  $(-\pi, \pi)$  上也服从均匀分布, 因此方程 (3-29) 的实部与虚部的均值分别为:

$$\langle \operatorname{Re}[E_s \exp(j\Phi)] \rangle = e \sum_{p=1}^P \langle \cos \varphi_p \rangle = 0 \quad (3-32)$$

$$\langle \operatorname{Im}[E_s \exp(j\Phi)] \rangle = e \sum_{p=1}^P \langle \sin \varphi_p \rangle = 0 \quad (3-33)$$

同时, 方程 (3-29) 的实部与虚部的方差分别为:

$$\langle (\operatorname{Re}[E_s \exp(j\Phi)] - \langle \operatorname{Re}[E_s \exp(j\Phi)] \rangle)^2 \rangle = \langle (\operatorname{Re}[E_s \exp(j\Phi)])^2 \rangle = e^2 \sum_{p=1}^P \sum_{m=1}^P \langle \cos \varphi_p \cos \varphi_m \rangle = P \frac{e^2}{2} \quad (3-34)$$

$$\langle (\operatorname{Im}[E_s \exp(j\Phi)] - \langle \operatorname{Im}[E_s \exp(j\Phi)] \rangle)^2 \rangle = \langle (\operatorname{Im}[E_s \exp(j\Phi)])^2 \rangle = e^2 \sum_{p=1}^P \sum_{m=1}^P \langle \sin \varphi_p \sin \varphi_m \rangle = P \frac{e^2}{2} \quad (3-35)$$

正如公式 (3-29) 所述, 散斑的幅度与相位由大量随机相量累加而成。根据中心极限定理可得, 如果随机变量是由具有有限均值和方差的大量独立随机变量相加而成的, 那么它的分布倾向于高斯分布。从公式 (3-32) 至公式 (3-35) 可以发现, 散斑场的实部和虚部的方差相同, 均值都为零, 因此, 散斑场满足中心极限定理的条件。由此可得, 散斑幅度  $E_s$  与相位  $\Phi$  的联合密度函数为:

$$p(E_s, \Phi) = \frac{E_s}{\pi \langle E_s^2 \rangle} \exp \left[ -\frac{E_s^2}{\langle E_s^2 \rangle} \right] \quad (3-36)$$

又因为散斑的强度  $I_s = E_s^2$ , 根据密度函数转换公式可得散斑强度  $I_s$  与相位  $\Phi$  的联合密度函数为:

$$p(I_s, \Phi) = \frac{1}{2\pi \langle I_s \rangle} \exp \left[ -\frac{I_s}{\langle I_s \rangle} \right] \quad (3-37)$$

利用求边缘概率密度函数的方法, 可得散斑强度  $I_s$  的密度函数为:

$$p(I_s) = \int_{-\pi}^{\pi} p(I_s, \Phi) d\Phi = \begin{cases} \frac{1}{\langle I_s \rangle} \exp \left[ -\frac{I_s}{\langle I_s \rangle} \right], & I_s \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (3-38)$$

散斑相位  $\Phi$  的密度函数为:

$$p(\Phi) = \int_0^{\infty} p(I_s, \Phi) dI_0 = \begin{cases} \frac{1}{2\pi}, & I_0 \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (3-39)$$

观察公式 (3-38) 与公式 (3-39) 可知,  $p(I_0, \Phi) = p(I_0)p(\Phi)$ , 由此可知散斑强度的概率密度函数与散斑相位的概率密度函数相互独立。

## (2) 散斑尺寸

观察图 3.14 与图 3.15 可以发现散斑图案的空间结构是亮点与暗点之间的随机分布, 很难用常规的参数描述散斑图案, 因此采用散斑尺寸描述散斑的空间结构。除此之外, 散斑尺寸还可以反映在空间分布中散斑强度起伏变化的快慢。平均散斑尺寸由散斑强度的空间自相关函数定义, 该定义方法的实质是通过空间中两点之间的统计关系来量化散斑的大小。

散斑强度的自相关函数  $\Gamma_I(X_1, X_2)$  为

$$\Gamma_I(X_1, X_2) = \langle I_s(X_1)I_s(X_2) \rangle \quad (3-40)$$

公式 (3-40) 中,  $X=(x,y)$ 。假设散斑场的实部与虚部相互独立, 则可用圆域复高斯统计变量描述散斑场。根据复高斯统计矩定理可知, 散斑强度的自相关函数  $\Gamma_I(X_1, X_2)$  可从公式 (3-40) 表示为:

$$\begin{aligned} \Gamma_I(X_1, X_2) &= \langle I_s(X_1) \rangle \langle I_s(X_2) \rangle + \left| \langle E_s(X_1)E_s^*(X_2) \rangle \right|^2 \\ &= \langle I_s(X_1) \rangle \langle I_s(X_2) \rangle + \left| \Gamma_E(X_1, X_2) \right|^2 \\ &= \langle I_s(X_1) \rangle \langle I_s(X_2) \rangle + \Gamma_{\Delta I}(X_1, X_2) \end{aligned} \quad (3-41)$$

公式 (3-41) 中,  $\Gamma_E(X_1, X_2)$  为散斑场的自相关函数,  $\Gamma_{\Delta I}(X_1, X_2)$  为散斑强度波动的自相关函数。根据公式 (3-41) 可知, 由于散斑的强度  $\langle I_s \rangle$  是定值, 不随空间坐标的改变而改变, 因此只要计算散斑场的自相关函数  $\Gamma_E(X_1, X_2)$  或散斑强度波动的自相关函数  $\Gamma_{\Delta I}(X_1, X_2)$  即可得出散斑尺寸。由于散斑场  $E_s(X)$  可由公式 (3-19) 表示, 则远场条件下, 散斑场的自相关函数  $\Gamma_E(X_1, X_2)$  可表示为:

$$\begin{aligned} \Gamma_E(X_1, X_2) &= \frac{1}{\lambda^2 R_1 R_2} \exp[ik(R_1 - R_2)] \exp\left[i\frac{k}{2}\left(\frac{x_1^2 + y_1^2}{R_1} - \frac{x_2^2 + y_2^2}{R_2}\right)\right] \iint E_o(\zeta_1, \eta_1) E_o^*(\zeta_2, \eta_2) \\ &\quad \times \left\langle \exp\left\{i\left[\Delta\varphi_{de}(\zeta_1, \eta_1) - \Delta\varphi_{de}(\zeta_2, \eta_2)\right]\right\} \right\rangle \exp\left[-ik\left(\frac{x_1\zeta_1 + y_1\eta_1}{R_1} - \frac{x_2\zeta_2 + y_2\eta_2}{R_2}\right)\right] ds_1 ds_2 \end{aligned} \quad (3-42)$$

假设被测目标表面足够粗糙, 则

$$\left\langle \exp \left\{ i \left[ \Delta \varphi_{de} (X_1) - \Delta \varphi_{de} (X_2) \right] \right\} \right\rangle \approx \Delta S \delta (X_1 - X_2) \quad (3-43)$$

公式 (3-43) 中,  $\Delta S$  为面积,  $\delta(X)$  为  $\delta$  函数。将公式 (3-43) 代入公式 (3-42)

可得

$$\begin{aligned} \Gamma_E (X_1, X_2) = & \frac{\Delta S}{\lambda^2 R_1 R_2} \exp [ik (R_1 - R_2)] \exp \left[ i \frac{k}{2} \left( \frac{x_1^2 + y_1^2}{R_1} - \frac{x_2^2 + y_2^2}{R_2} \right) \right] \iint |E_o (\zeta, \eta)|^2 \\ & \times \exp \left[ -ik \left( \frac{x_1 \zeta + y_1 \eta}{R_1} - \frac{x_2 \zeta + y_2 \eta}{R_2} \right) \right] ds_1 ds_2 \end{aligned} \quad (3-44)$$

由于探测孔径远远小于探测距离, 因此  $R_1 \approx R_2 = R_0$ , 并用  $x = x_1 - x_2, y = y_1 - y_2$ ,

此时, 根据公式 (3-44), 可得散斑强度波动的自相关函数  $\Gamma_{\Delta I}(X)$  为:

$$\Gamma_{\Delta I} (X) = |\Gamma_E (X)|^2 = \frac{\Delta S^2}{\lambda^4 R_0^4} \left| \iint |E_o (\zeta, \eta)|^2 \exp \left[ -ik \left( \frac{\zeta x + \eta y}{R_0} \right) \right] ds_1 ds_2 \right|^2 \quad (3-45)$$

归一化的散斑强度波动的自相关函数  $\Upsilon_{\Delta I}(X)$  为:

$$\Upsilon_{\Delta I} (X) = \frac{\left| \iint |E_o (\zeta, \eta)|^2 \exp \left[ -ik \left( \frac{\zeta x + \eta y}{R_0} \right) \right] ds_1 ds_2 \right|^2}{\left| \iint |E_o (\zeta, \eta)|^2 ds_1 ds_2 \right|^2} \quad (3-46)$$

根据公式 (3-16) 可知, 照射在粗糙表面处的光束  $E_o(X)$  为高斯光束, 并且束腰正位于粗糙表面上, 因此将公式 (3-16) 代入公式 (3-46) 可得归一化的散斑强度波动的自相关函数  $\Upsilon_{\Delta I}(X)$  为:

$$\Upsilon_{\Delta I} (X) = \exp \left( -\frac{\pi^2 \omega_0^2}{\lambda^2 R_0^2} |X|^2 \right) \quad (3-47)$$

散斑的平均宽度  $\langle \sigma_0 \rangle$  为归一化的散斑强度波动自相关函数在  $\exp(-1)$  处, 具体可以表示为:

$$\langle \sigma_0 \rangle = \frac{\lambda R_0}{\pi \omega_0} \quad (3-48)$$

正如公式 (3-48) 所示, 散斑的平均宽度  $\langle \sigma_0 \rangle$  和激光波长  $\lambda$ 、探测距离  $R_0$  成正比, 与束腰半径  $\omega_0$  成反比。需要注意的是, 一些文献将散斑的平均宽度  $\langle \sigma_0 \rangle$  定义为

归一化的散斑强度波动自相关函数在  $\exp(-2)$  处<sup>[60]</sup>, 即  $\langle \sigma \rangle = \sqrt{2} \langle \sigma_0 \rangle$ 。在本文中, 则采用公式 (3-48) 表示散斑的平均宽度。

### 3.3.2 动态散斑特性

#### (1) 动态散斑的时空相关函数

3.3.1 节主要介绍的是静态散斑的特性, 当散斑模式不随时间变化时, 只会影响回波光与本征光的混频效率, 这点在 3.2 节进行了详细的研究。当被测目标与激光相干语音侦测系统之间存在除光轴方向外的相对运动时会引起散斑模式随时间的改变而改变, 从而给测量引入散斑噪声, 最终导致侦测的不确定性。这类随时间变化而变化的散斑称之为动态散斑。要研究散斑噪声对激光相干语音侦测系统的影响, 首先应分析动态散斑的特性。动态散斑的特性可通过散斑强度起伏的相关函数描述, 该函数量化了存在时间间隔的空间中两点之间散斑强度的统计关系。实际情况中, 由于外力很难使得被测物体发生转动, 因此本文主要考虑相对运动为平动时动态散斑的特性。

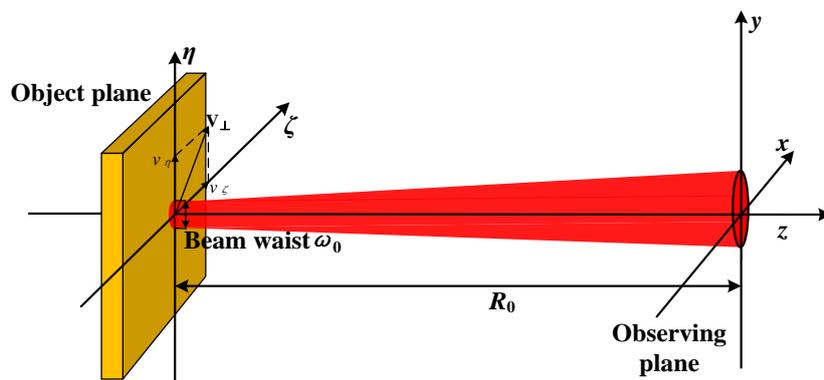


图 3.17 动态散斑坐标系示意图

Figure 3.17 Schematic of coordinate system for time varying speckles.

如图 3.17 所示, 假设粗糙表面的相对运动速度为  $\mathbf{V}_{\perp} = (v_{\eta}, v_{\zeta})$ 。在实际应用中, 往往将发射激光聚焦在目标表面上, 因此假设发射激光经过聚焦系统后的束腰位于粗糙表面上, 其束腰半径为  $\omega_0$ , 粗糙表面所在平面与侦测系统所在平面之间的距离为  $R_0$ 。此时, 根据菲涅耳衍射定律可知, 散斑场  $E_s(x, y)$  为:

$$E_s(x, y, t) = \frac{1}{i\lambda R_0} \exp(ikR_0) \int_{-\infty}^{+\infty} \int E_o \exp\left(-\frac{\zeta^2 + \eta^2}{\omega_0^2}\right) \exp\left[i\Delta\varphi_{de}(v_\zeta t + v_\eta t)\right] \times \exp\left[-ik\left[\frac{(x-\zeta)^2 + (y-\eta)^2}{2R_0}\right]\right] d\zeta d\eta \quad (3-49)$$

根据公式(3-41)可知, 散斑强度起伏的相关函数  $\Gamma_{\Delta I}(\mathbf{X}_1, \mathbf{X}_2; t_1, t_2)$  为

$$\Gamma_{\Delta I}(\vec{X}_1, t_1; \vec{X}_2, t_2) = \left\langle \left| E_s(\vec{X}_1, t_1) E_s^*(\vec{X}_2, t_2) \right|^2 \right\rangle = \left| \Gamma_E(\vec{X}_1, t_1; \vec{X}_2, t_2) \right|^2 \quad (3-50)$$

观察公式(3-50)可得, 散斑强度起伏的自相关函数  $\Gamma_{\Delta I}(\mathbf{X}_1, \mathbf{X}_2; t_1, t_2)$  可通过求取散斑场的自相关函数  $\Gamma_E(\mathbf{X}_1, \mathbf{X}_2; t_1, t_2)$  获得。将公式(3-49)代入公式(3-50)可得, 散斑场的相关函数  $\Gamma_E(\mathbf{X}_1, \mathbf{X}_2; t_1, t_2)$  为:

$$\Gamma_E(\vec{X}_1, t_1; \vec{X}_2, t_2) = \frac{1}{\lambda^2 R_1 R_2} \exp[ik(R_1 - R_2)] \iint E_o(\vec{\xi}_1) E_o^*(\vec{\xi}_2) \times \left\langle \exp\left\{i\left[\Delta\varphi_{de}(\vec{\xi}_1 - \vec{V}_\perp t_1) - \Delta\varphi_{de}(\vec{\xi}_2 - \vec{V}_\perp t_2)\right]\right\} \right\rangle \exp\left[-ik\left[\frac{(\vec{X}_1 - \vec{\xi}_1)^2}{2R_1} - \frac{(\vec{X}_2 - \vec{\xi}_2)^2}{2R_2}\right]\right] ds_1 ds_2 \quad (3-51)$$

公式(3-51)中,  $\mathbf{X}=(x, y)$ ,  $\boldsymbol{\xi}=(\zeta, \eta)$ 。根据公式(3-43)可知表面足够粗糙时

$$\left\langle \exp\left\{i\left[\Delta\varphi_{de}(\vec{\xi}_1 - \vec{V}_\perp t_1) - \Delta\varphi_{de}(\vec{\xi}_2 - \vec{V}_\perp t_2)\right]\right\} \right\rangle \approx \Delta S \delta\left[\vec{\xi}_1 - \vec{\xi}_2 - \vec{V}_\perp(t_1 - t_2)\right] \quad (3-52)$$

将公式(3-52)代入公式(3-51), 并令  $\Delta t = t_1 - t_2$ ,  $\vec{\xi}_1 = \vec{\xi} + 1/2 \vec{V}_\perp \Delta t$ ,  $\vec{\xi}_2 = \vec{\xi} - 1/2 \vec{V}_\perp \Delta t$ ,  $\Delta \mathbf{X} = \mathbf{X}_1 - \mathbf{X}_2$ , 并且由于探测孔径要远远小于探测距离, 因此  $R_1 \approx R_2 = R_0$ , 公式(3-51)可化简为

$$\Gamma_E(\Delta \mathbf{X}, \Delta t) = \frac{\Delta S}{\lambda^2 R_0^2} \exp\left[\frac{ik}{2R_0} \vec{V}_\perp \Delta t (\vec{X}_1 + \vec{X}_2)\right] \iint E_o\left(\vec{\xi} + \frac{1}{2} \vec{V}_\perp \Delta t\right) E_o^*\left(\vec{\xi} - \frac{1}{2} \vec{V}_\perp \Delta t\right) \times \exp\left[-ik\left[\frac{\vec{\xi}(\Delta \mathbf{X} - \vec{V}_\perp \Delta t)}{R_0}\right]\right] ds_1 ds_2 \quad (3-53)$$

其归一化的自相关函数  $\Upsilon_{\Delta I}(\Delta \mathbf{X}, \Delta t)$  为:

$$\Upsilon_{\Delta I}(\Delta \mathbf{X}, \Delta t) = \frac{\Gamma_{\Delta I}(\Delta \mathbf{X}, \Delta t)}{\Gamma_{\Delta I}(0, 0)} = \frac{\left| \iint E_o\left(\vec{\xi} + \frac{1}{2} \vec{V}_\perp \Delta t\right) E_o^*\left(\vec{\xi} - \frac{1}{2} \vec{V}_\perp \Delta t\right) \exp\left[-ik\frac{\vec{\xi}(\Delta \mathbf{X} - \vec{V}_\perp \Delta t)}{R_0}\right] ds_1 ds_2 \right|^2}{\left| \iint E_o(\vec{\xi}) E_o^*(\vec{\xi}) ds_1 ds_2 \right|^2} \quad (3-54)$$

正如图 3.17 所示,  $E_o(\mathbf{X})$  为照射在粗糙表面处的高斯光束, 并且束腰正位于粗糙表面上, 因此发射光束  $E_o(\mathbf{X})$  可用公式 (3-16) 表示。将公式 (3-16) 代入 (3-54) 可得:

$$\Upsilon_{\Delta}(\vec{X}, \Delta t) = \exp\left(-\frac{|\vec{V}_{\perp}|^2 \Delta t^2}{\omega_0^2}\right) \exp\left(-\frac{\pi^2 \omega_0^2}{\lambda^2 R_0^2} |\vec{X} - \vec{V}_{\perp} \Delta t|^2\right) \quad (3-55)$$

当  $\Delta t=0$  时, 此时公式 (3-55) 变为散斑光强起伏的空间相关函数, 该函数可以描述散斑大小, 这点已在 3.3.1 节中做了详细的介绍。在本章节中, 相较于散斑的空间相关函数, 更关心散斑光强随时间变化的强度起伏。由于在实际探测中, 探测距离远大于接收孔径  $\mathbf{X}$ , 并且可以把探测标定点选择在  $\mathbf{X}=0$  的中心点附近。因此, 此时归一化的散斑强度的起伏的相关函数为<sup>[102]</sup>

$$\Upsilon_{\Delta}(\Delta t) = \exp\left(-\frac{|\vec{V}_{\perp}|^2 \Delta t^2}{\omega_0^2}\right) \exp\left(-\frac{\pi^2 \omega_0^2}{\lambda^2 R_0^2} |\vec{V}_{\perp}|^2 \Delta t^2\right) = \exp\left(-\frac{\Delta t^2}{\tau_c^2}\right) \quad (3-56)$$

公式 (3-56) 中,  $\tau_c$  表示为散斑强度起伏的相关时间, 该参数描述了散斑的相关持续时间, 其可以表示为

$$\tau_c = \frac{1}{|\vec{V}_{\perp}|} \left( \frac{1}{\omega_0^2} + \frac{\pi^2 \omega_0^2}{\lambda^2 R_0^2} \right)^{-\frac{1}{2}} = \frac{1}{|\vec{V}_{\perp}|} \left( \frac{1}{\omega_0^2} + \frac{1}{\langle \sigma_0 \rangle^2} \right)^{-\frac{1}{2}} \quad (3-57)$$

观察公式 (3-57) 可以发现, 在给定束腰半径  $\omega_0$  的情况下, 相关时间  $\tau_c$  与相对运动的速度  $\mathbf{V}_{\perp}$  成反比。这说明, 被测物体与激光相干语音侦测系统之间的相对速度越快, 相关长度  $\tau_c$  越短, 相对应的散斑光强起伏越剧烈。

该结论可被实验验证。图 3.18.a 为实验的设计图。如图 3.18.b 所示, 在实验中, 采用的激光器为法国 Oxxius 公司生产的窄线宽激光器, 波长为 532nm。光电探测器为日本滨松(HAMAMATSU)公司生产的 C5658 型光电探测器, 其探测带宽为 1GHz。如图 3.18.c 所示, 被测物体为常见的纸盒, 为了模拟被测物体与侦测系统之间存在相对运动, 将纸盒放在与光轴垂直的电动导轨上。在实验过程中, 将发射激光聚焦在纸盒上。散斑强度随时间的起伏可通过对除去其均值的光电探测器输出信号作自相关获得。

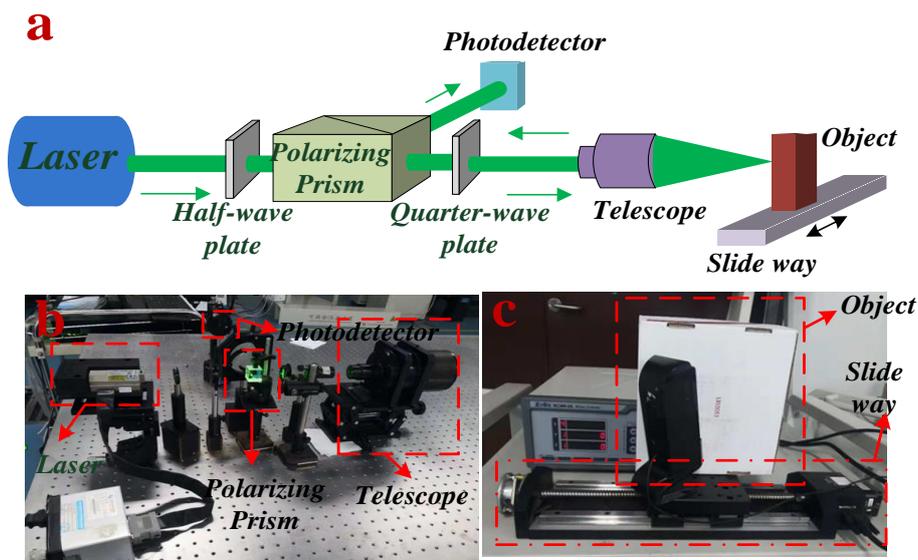


图 3.18 (a) 探测散斑强度的实验原理图。(b) 实验系统。(c) 目标

Figure 3.18 (a) Experimental setup for detecting speckle intensity. (b) Prototype of the experimental system. (c) Test component.

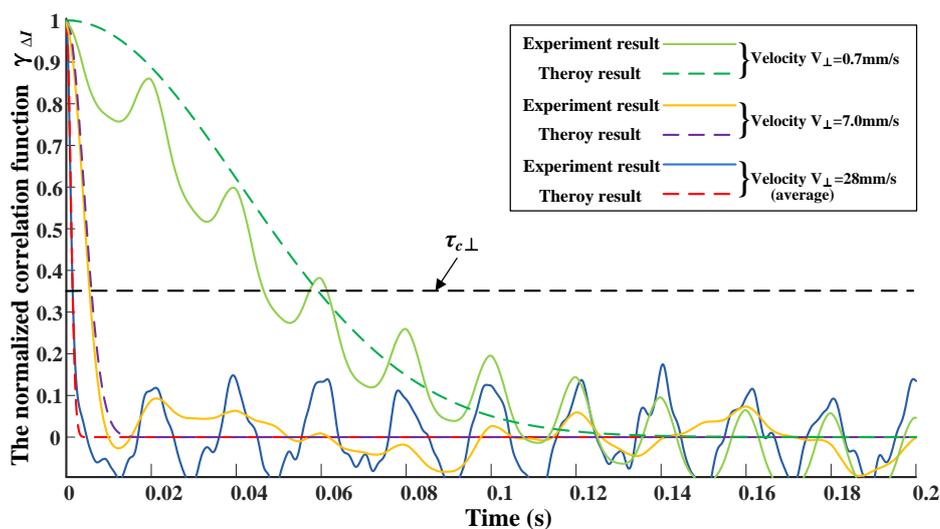


图 3.19 不同速度下归一化的散斑强度抖动的自相关函数  $\gamma_{\Delta I}$

Figure 3.19 The normalized correlation function of the intensity fluctuation  $\gamma_{\Delta I}$

under different velocity.

图 3.19 为不同速度下，散斑强度起伏的自相关函数。观察图 3.19，可以发现：

- 1) 实验值与理论计算值相一致。
- 2) 相关长度  $\tau_c$  随速度的增加而减小，这说明相对速度越快，散斑光强的起伏越剧烈。

(2) 动态散斑运动方式

动态散斑具有两种运动方式<sup>[60,102~104]</sup>用于描述其散斑图案随时间的变化。一种运动方式称为散斑平移(speckle translation), 如图 3.20 所示, 这种运动方式的特点是散斑图案在空间中平移, 但它的大小和形状不发生改变。另一种运动方式称为沸腾(boiling), 如图 3.21 所示, 这种运动方式的特点是散斑图案的形状和大小随着时间的变化而发生变化, 这种变化是散斑图案彼此之间独立的变化, 但是统计参数在时间上是静止的。

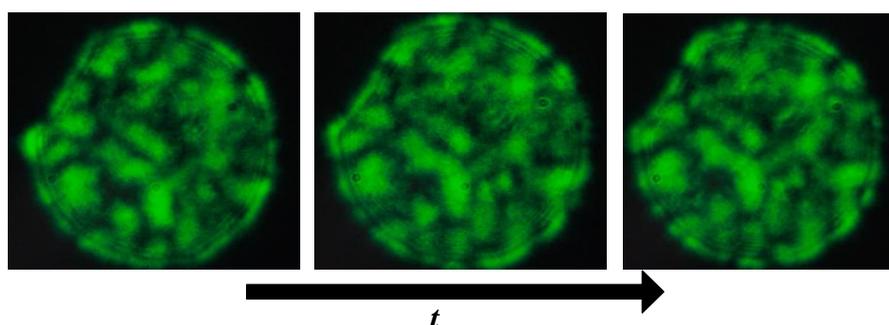


图 3.20 散斑平移

Figure 3.20 Speckle translation.

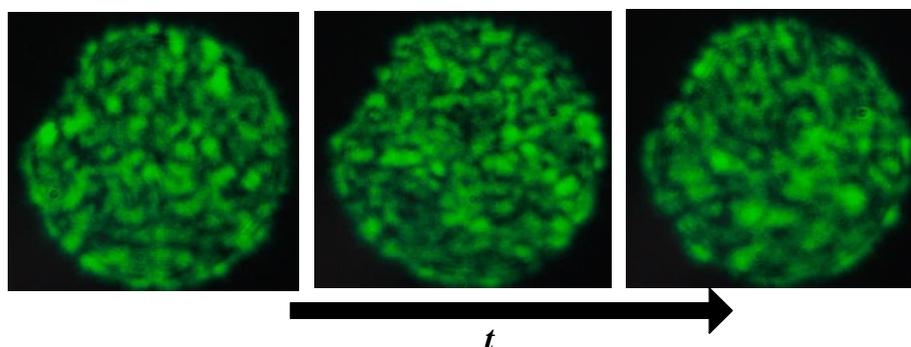


图 3.21 散斑沸腾

Figure 3.21 Speckle boiling.

散斑平移时, 归一化的散斑强度波动的相关函数的峰值具有空间延迟和时间延迟。假设时间延迟为  $\Delta t$ , 则根据公式 (3-55) 可知, 只有当空间延迟  $\mathbf{X} = \mathbf{V}_\perp \Delta t$  时, 散斑强度波动的相关函数最大。因此, 散斑平移量  $\mathbf{X}_T$  可以用被测物体与侦测系统之间除光轴方向外的相对位移  $\mathbf{X} = \mathbf{V}_\perp \Delta t$  表示。

散斑沸腾时，散斑的形状和大小随时间发生改变。散斑沸腾会使得归一化的散斑强度波动的相关函数的幅值降低。设空间延迟  $\mathbf{X} = \mathbf{V}_\perp \Delta t$ ，此时，归一化的散斑强度波动的相关函数为：

$$\Upsilon_{\Delta t}(\vec{X}, \Delta t) = \exp\left(-\frac{|\vec{V}_\perp|^2 \Delta t^2}{\omega_0^2}\right) \quad (3-58)$$

公式 (3-58) 的幅值下降到  $\exp(-1)$  时，相关函数去相关，此时相对位移  $|\mathbf{X}| = \omega_0$ 。

当被测物体与侦测系统存在相对运动时，会引起散斑运动。通常而言，散斑的运动方式同时包括散斑平移和散斑沸腾这两种运动方式。根据公式 (3-57) 可知，在相关时间内  $\tau_c$ ，被测物体与侦测系统之间的最大相对位移  $a_c$  为：

$$a_c = |\vec{V}_\perp| \tau_c = \left(\frac{1}{\omega_0^2} + \frac{1}{\langle \sigma_0 \rangle^2}\right)^{-\frac{1}{2}} = \omega_0 \left(\frac{1}{1 + \omega_0^2 / \langle \sigma_0 \rangle^2}\right)^{\frac{1}{2}} = \omega_0 \left(\frac{1}{1 + T^2}\right)^{\frac{1}{2}} \quad (3-59)$$

公式 (3-59) 可表示为，在散斑开始沸腾时散斑的平移量。其中  $T$  表示为光斑大小  $\omega_0$  与散斑大小  $\langle \sigma_0 \rangle$  的比值。将公式 (3-59) 代入公式 (3-56) 可得

$$\Upsilon_{\Delta t}(X, \Delta t) = \exp\left[-\left(\frac{1}{1 + T^2}\right)\right] \quad (3-60)$$

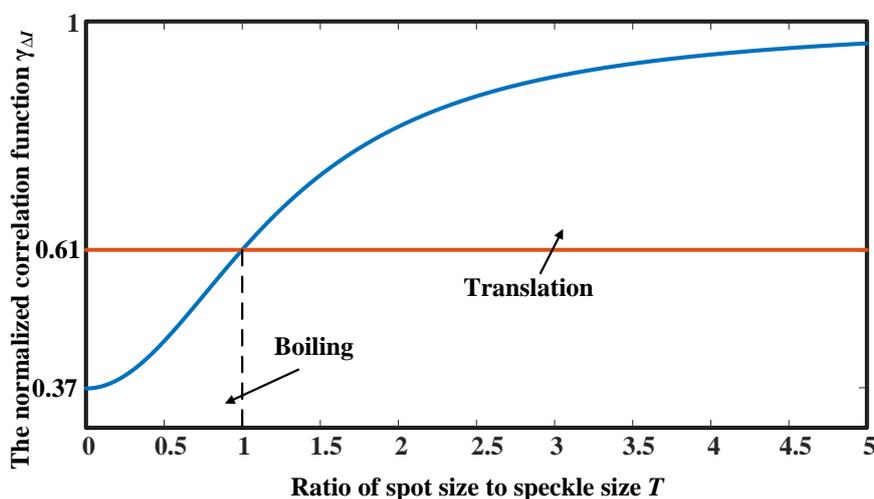


图 3.22 不同比值下的散斑强度抖动的自相关函数  $\Upsilon_{\Delta I}$

Figure 3.22 The normalized correlation function of the intensity fluctuation  $\Upsilon_{\Delta I}$  under different ratio.

图 3.22 为不同光斑尺寸与散斑尺寸的比值下，散斑强度起伏的自相关函数的幅

值。由图 3.22 可知, 当  $T=0$  时, 散斑的运动方式只有沸腾。当  $0<T<1$  时, 散斑的运动方式以沸腾为主导。当  $T=1$  时, 散斑的运动方式中沸腾与平移占的比重相同。当  $T>1$  时, 散斑的运动方式以平移为主导。

### 3.3.3 散斑噪声对激光相干语音侦测系统的影响

#### (1) 散斑噪声功率谱密度

正如公式 (3-28) 所示, 由于散斑噪声会导致中频电流  $i_{IF}(t)$  产生幅度调制  $I_{res}(t)$  和相位调制  $\Phi_{res}(t)$ , 通常而言, 这两种调制同时发生。由于中频电流  $i_{IF}(t)$  是载波信号, 为了最终获得语音信息, 需要对其进行进一步解调运算以获得多普勒频移, 再根据多普勒频移从而获得语音信号, 具体解调方法会在第四章详述。但是由于存在上述影响, 使得解调获得多普勒频率过程中引入散斑噪声。由于本章节只考虑散斑噪声对语音侦测系统的影响, 因此假定中频电流  $i_{IF}(t)$  得到了精确的解调, 即不考虑散粒噪声等因素对侦测系统的影响, 此时得到的多普勒频率  $\bar{f}_d$  可表示为:

$$\bar{f}_d = \frac{\omega_d}{2\pi} + \frac{1}{2\pi} \dot{\Phi}_{res}(t) \quad (3-61)$$

根据公式 (3-61) 所示, 散斑噪声会增加侦测的不确定性。评估散斑噪声对系统的影响的一个简单方法是检测散斑噪声的功率谱。根据维纳-辛钦定理可知: 宽平稳随机过程的功率谱密度与其自相关函数互为傅立叶变换。因此可通过计算散斑相位导数  $d\Phi_{res}(t)/dt$  的自相关函数获得散斑噪声的功率谱密度。

由 3.3.1 节所知, 散斑可以认为是一个高斯过程, 并假设散斑强度的实部和虚部分别为  $X_{sp}$  和  $Y_{sp}$ , 因此, 散斑的相位可以通过虚部  $Y_{sp}$  与实部  $X_{sp}$  之间比值的反正切计算获得。因此散斑相位导数  $d\Phi_{res}(t)/dt$  的自相关函数可表示为:

$$\begin{aligned} \langle \dot{\Phi}_{res}(0) \dot{\Phi}_{res}(\tau) \rangle &= \left\langle \left( \frac{X_{sp}(0) \dot{Y}_{sp}(0) - Y_{sp}(0) \dot{X}_{sp}(0)}{X_{sp}^2(0) + Y_{sp}^2(0)} \right) \left( \frac{X_{sp}(\tau) \dot{Y}_{sp}(\tau) - Y_{sp}(\tau) \dot{X}_{sp}(\tau)}{X_{sp}^2(\tau) + Y_{sp}^2(\tau)} \right) \right\rangle \\ &= \left[ \frac{g\ddot{g} + h\ddot{h} - \dot{g}^2 - \dot{h}^2}{2q} + \frac{(\dot{g}h - \dot{h}g)^2}{q^2} \right] \ln(1-q) + \frac{b^2 + 2b(\dot{g}h - \dot{h}g) + (\dot{g}h - \dot{h}g)^2}{(1-q)} / q \quad (3-62) \end{aligned}$$

公式 (3-62) 中

$$\begin{aligned}
 b &= \langle X_{sp}(0) \dot{Y}_{sp}(0) \rangle = \langle X_{sp}(\tau) \dot{Y}_{sp}(\tau) \rangle \\
 &= -\langle \dot{X}_{sp}(0) Y_{sp}(0) \rangle = -\langle \dot{X}_{sp}(\tau) Y_{sp}(\tau) \rangle \\
 g &= \langle X_{sp}(0) X_{sp}(\tau) \rangle = \langle Y_{sp}(0) Y_{sp}(\tau) \rangle \\
 h &= \langle X_{sp}(0) Y_{sp}(\tau) \rangle = -\langle X_{sp}(\tau) Y_{sp}(0) \rangle \\
 q &= (g^2 + h^2) / \langle X_{sp}^2 \rangle
 \end{aligned} \tag{3-63}$$

当散斑噪声的功率谱对称时, 公式 (3-62) 可简化为<sup>[62~63]</sup>:

$$\langle \dot{\Phi}_{res}(0) \dot{\Phi}_{res}(\tau) \rangle = \frac{1}{2} \left( \frac{\rho \ddot{\rho} - \dot{\rho}^2}{\rho^2} \right) \ln(1 - \rho^2) \tag{3-64}$$

公式 (3-64) 中,  $\rho$  为归一化的散斑强度实部的相关函数。根据公式 (3-64) 可知, 零频处散斑的功率谱密度  $S_{sp}(0)$  可表示为

$$S_{sp}(0) = 2 \int_0^\infty \left( \frac{\rho \ddot{\rho} - \dot{\rho}^2}{\rho^2} \right) \ln(1 - \rho^2) dt = 2 \int_0^\infty \frac{\dot{\rho}^2}{1 - \rho^2} dt \tag{3-65}$$

因为归一化的散斑强度实部的相关函数  $\rho$  可用公式 (3-56) 表示, 因此将公式 (3-56) 代入公式 (3-65) 可得, 零频处散斑的功率谱密度  $S_{sp}(0)$  为:

$$S_{sp}(0) = 2 \int_0^\infty \frac{\dot{\rho}^2}{1 - \rho^2} dt = 8 \int_0^\infty \frac{\exp\left(-2 \frac{t^2}{\tau_c^2}\right)}{\tau_c^4 \left(1 - \exp\left(-2 \frac{t^2}{\tau_c^2}\right)\right)} dt \approx \frac{3.27}{\tau_c} \tag{3-66}$$

经过仔细的测量由目标与激光多普勒测振仪之间横向运动而产生的散斑噪声的基底后, Dräbenstedt 发现, 在频率低于截止频率  $f_c = V_\perp / (2\omega_0)$  时, 散斑噪声的频谱密度恒定, 而超出截止频率时, 散斑噪声的频谱密度以  $1/f_c$  的速度下降<sup>[66]</sup>。因此, 根据散斑噪声在零频处的功率谱密度, 散斑噪声在全频域的功率谱密度  $S_{sp}(f)$  可表示为:

$$S_{sp}(f) = \begin{cases} \frac{3.27}{\tau_c} & f \leq f_c \\ \frac{3.27 f_c^2}{f^2} \frac{1}{\tau_c} & f > f_c \end{cases} \tag{3-67}$$

公式(3-67)的计算结果与 Dräbenstedt 的实测结果一致<sup>[66]</sup>。观察公式(3-67)可以发现,散斑噪声的功率谱密度随相关时间  $\tau_c$  的增加而减小。根据公式(3-57)可知,相关时间  $\tau_c$  与相对速度  $V_{\perp}$  成反比,并随着束腰半径  $\omega_0$  与散斑尺寸  $\langle\sigma_0\rangle$  的增加而增加。由于本文假设被测物体位于聚焦透镜焦点上,因此根据高斯光束在透镜组传播公式可得,束腰半径  $\omega_0$  与探测距离  $R_0$  的关系为:

$$\omega_0 = \frac{\lambda R_0}{\pi \omega'_0} \tag{3-68}$$

公式(3-68),  $\omega'_0$  为经过透镜组前,高斯光束的束腰半径。根据公式(3-68)与公式(3-48)可知,散斑尺寸  $\langle\sigma_0\rangle$  与探测距离  $R_0$  无关

$$\langle\sigma_0\rangle = \omega'_0 \tag{3-69}$$

将公式(3-69)与公式(3-68)代入公式(3-57)可得,相关时间  $\tau_c$  与相对速度  $V_{\perp}$  以及探测距离  $R_0$  的关系为

$$\tau_c = \frac{1}{|V_{\perp}|} \left( \frac{1}{\omega_0^2} + \frac{1}{\langle\sigma_0\rangle^2} \right)^{\frac{1}{2}} = \frac{1}{|V_{\perp}|} \left( \frac{\pi^2 \omega_0'^2}{\lambda^2 R_0^2} + \frac{1}{\omega_0'^2} \right)^{\frac{1}{2}} \tag{3-70}$$

观察公式(3-70)可以发现,散斑噪声相关时间  $\tau_c$  取决于相对速度  $V_{\perp}$ , 探测距离  $R_0$  以及出射前高斯光束的束腰半径  $\omega'_0$ 。因为本章节只讨论被测物体特性对相干语音侦测的影响,而出射前的高斯光束束腰半径  $\omega'_0$  由系统本身决定,因此只讨论相对速度  $V_{\perp}$  和探测距离  $R_0$  对散斑噪声的影响。

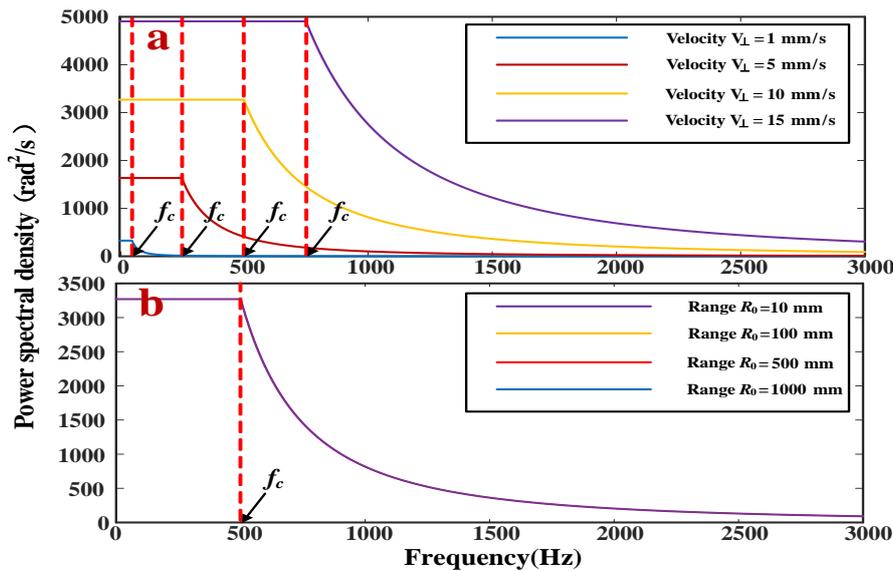


图 3.23 (a) 散斑噪声功率谱密度和相对速度关系。(b) 散斑噪声功率谱密度和距离关系  
 Figure 3.23 (a) Relationship between speckle noise spectral power density and relative velocity. (b) Relationship between speckle noise spectral power density and distance.

图 3.23a 和图 3.23b 分别为散斑噪声的功率谱密度  $S_{sp}(f)$  与相对速度  $V_{\perp}$  以及探测距离  $R_0$  的关系。在图 3.23a 中, 探测距离  $R_0=50\text{m}$ , 初始高斯光束的束腰半径  $\omega_0'$  为  $0.01\text{mm}$ 。观察图 3.23a 可以发现, 功率谱密度  $S_{sp}(f)$  随相对速度  $V_{\perp}$  的增加而增加。在图 3.23b, 相对速度  $V_{\perp}=10\text{mm/s}$ , 初始高斯光束的束腰半径  $\omega_0'$  为  $0.01\text{mm}$ 。观察图 3.23b 可以发现, 功率谱密度  $S_{sp}(f)$  对探测距离  $R_0$  的变化不敏感。

## (2) 散斑噪声对侦测语音的干扰

正如图 3.23 所示, 散斑噪声主要集中在中低频区域。然而, 人说话的频率范围大约在  $300\sim 3000\text{Hz}$ , 因此, 激光相干语音侦测系统侦测获得的语音信号会受到散斑噪声的干扰。具体的干扰形式表现为: 观察公式 (3-61) 可以发现, 由于散斑噪声的干扰, 会使得侦测系统获取的多普勒相位在  $-\pi$  与  $\pi$  之间的变化不连续, 这些不连续的变化会导致系统侦测输出的信号中出现尖刺。这种尖刺变化速度快, 变化幅度大, 听起来好像是音量很大、音调很高的‘吱吱’声, 这对侦测语音的质量影响很大。

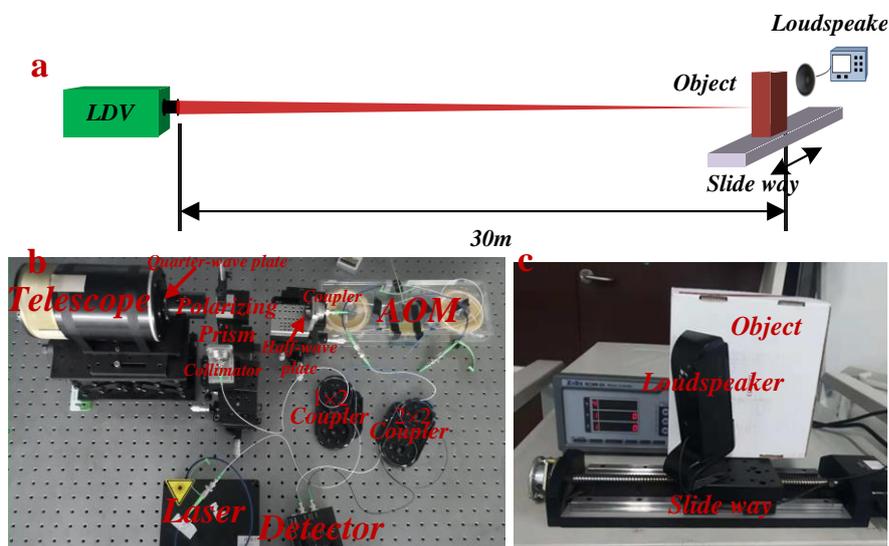


图 3.24 (a) 验证散斑噪声影响的实验原理图。(b) 激光相干语音侦测系统。(c) 被测目标

Figure 3.24 (a) Experimental setup for verifying the effect of speckles noise. (b) Prototype of the LDV system. (c) Target.

为验证散斑噪声对侦测语音影响分析的准确性,进行了实验验证。如图 3.24.a 所示,为实验原理图。实验可简单分为语音侦测系统和被测目标两个部分。实验中,如图 3.24.b 所示,采用自研的激光相干语音侦测系统作为语音侦测系统(激光相干语音侦测系统详见第一章,语音重构算法详见第四章)。如图 3.24.c 所示,被测目标部分由音箱和纸盒构成。音箱作为声源产生声音,产生的声音驱动纸盒发生振动。

实验时,为了可以产生散斑噪声,将音箱和纸盒放在与光轴垂直的电动导轨上。同时为了更好的分析散斑噪声对侦测语音的影响,并考虑到纸盒在低频时有良好的振动响应,因此让声源产生频率为 500Hz 的单音信号。另外,为了保证系统载噪比足够强,探测距离为 30m。

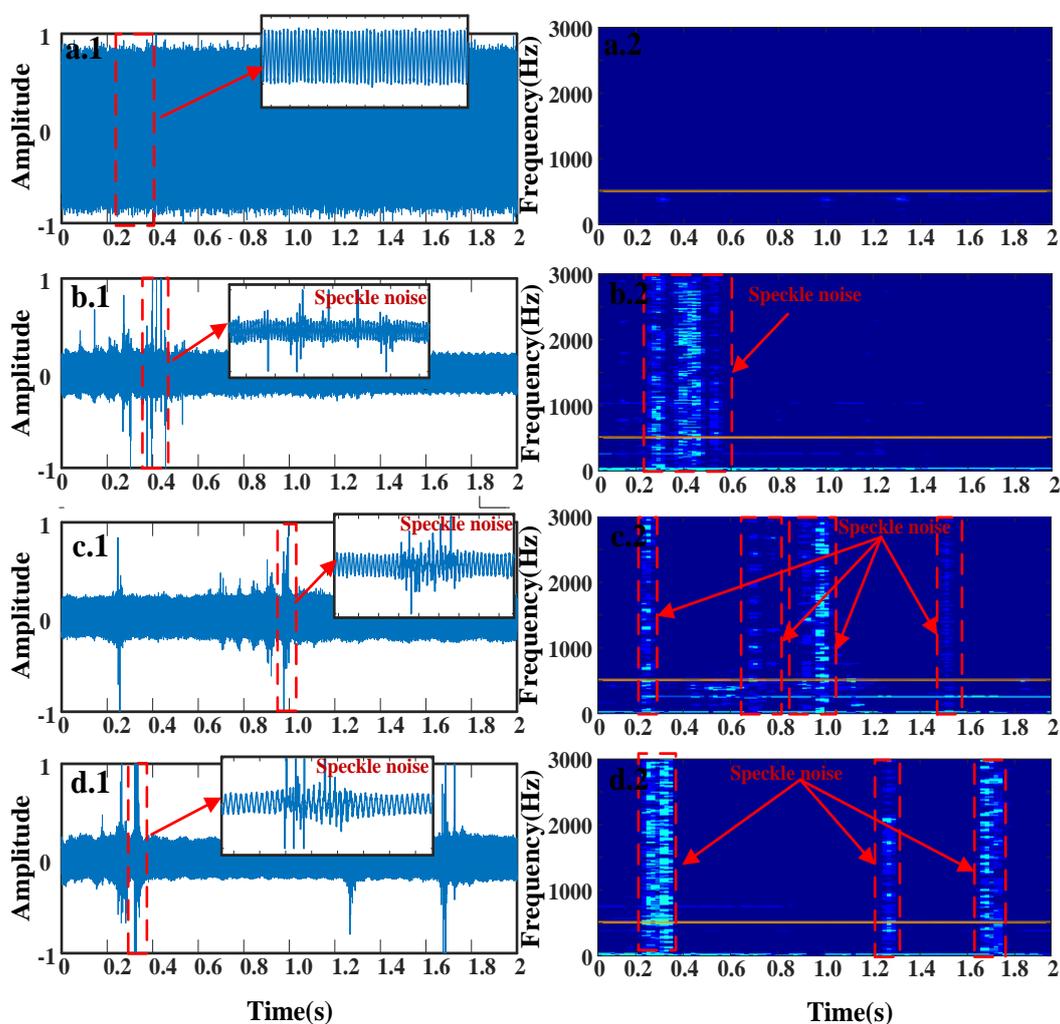


图 3.25 激光相干语音侦测系统在不同相对速度下获取的单音信号的波形图与语谱图。(a) 相对速度为 0。(b) 相对速度为 2.33mm/s。(c) 相对速度为 4.66mm/s。(d) 相对速度为 7.00mm/s。

Figure 3.25 The Speech waveforms and spectrograms of the LDV-captured single tone with different relative velocity. (a) velocity is zero. (b) Relative velocity is 2.33mm/s. (c) velocity is 4.66mm/s. (d) Relative velocity is 7.00mm/s.

图 3.25 为不同相对速度下,激光相干语音侦测系统获取的语音信号的波形图与语谱图。观察图 3.25.a 可以发现,当相对速度  $V_{\perp}=0$ ,即没有受到散斑噪声干扰时,相干语音侦测系统获取的语音信号与声源发出的信号完全一致。但是,如图 3.25.b 至 3.25.d 所示,随着散斑噪声的引入,系统获取的语音信号开始出现尖刺。从时频域上看,这种尖刺在声音上表现为能量强,频率高,听起来就类似于音量很大,音调很高的‘吱吱’声,非常影响听觉体验。另外,尖刺出现在侦测语音的位置以及次数是随机的。观察 3.25.b 至 3.25.d 还可以发现,随着相对速度  $V_{\perp}$  的增加,散斑噪声对侦测语音的影响似乎更大。

理论及实验结果表明,散斑噪声会使得相干语音侦测系统获取的语音信号出现尖刺,这些尖刺会严重影响侦测语音的质量。同时,随着相对速度的增加,散斑噪声也越强,对相干语音侦测系统的干扰也越大。由此可得,为了降低散斑噪声对激光相干语音侦测系统的影响,可以选择稳定的,除光轴外无相对运动的物体作为测量目标。

### 3.4 本章小结

本章内容主要分三个方面:

(1) 首先基于声场传输损耗和 SSC 模型建立了不同材质的平板在声场中的振动模型。然后,基于该振动模型,数值仿真了不同材质在声场下的振动响应。最后,通过实验验证了振动模型的准确性。理论和实验表明:1).纸制平板、塑料制平板对低频声场下有较好的振动响应,但是随着声场频率的增加,振动幅度快速下降。2).金属制平板则在不同频率的声场下的振动幅度相差不大,但是整体振动幅度(特别是低频部分)要小于非金属制材料。3).玻璃制平板则表现较为均衡。总体而言,从获取可理解语音角度而言,纸制、塑料制与玻璃制材料更适合作为语音获取的目标。

(2) 首先以远场衍射理论为基础建立了从粗糙表面反射的回波光的光场表达式。然后,根据回波光的光场表达式,数值分析了表面粗糙度对激光相干语音侦测系统载噪比的影响。最后通过实验验证了表面粗糙度对于系统载噪比的影响。

研究表明：系统的载噪比随着表面粗糙度的增加而下降，但是当表面粗糙度到达一定值以后（ $R_a > 3\lambda$ ），系统的载噪比则在某一值附近做随机波动，此时再纠结物体的表面粗糙度就没有意义了。

（3）首先理论推导了描述散斑特性的参数：散斑的概率密度、散斑尺寸、散斑时空相关函数以及散斑运动方式。然后详细计算了散斑噪声的功率谱密度。最后阐述了散斑噪声对激光相干语音侦测系统的影响，并通过实验进行了验证。理论和实验表明：散斑噪声会使得相干语音侦测系统获取的语音信号出现尖刺，这些尖刺听起来类似音量很大，音调很高的‘吱吱’声，非常影响语音质量。同时，理论和实验表明，散斑噪声与被测物体和侦测系统之间除光轴外的相对运动有很大关系，相对运动速度越快，散斑噪声越强。因此为了减少散斑噪声，最好选择除光轴外无相对运动的物体作为测量目标。

由此可得，一个合适的测量目标应该是表面光滑，与侦测系统除光轴外无相对运动，对声场具有的良好振动响应的物体。但是，需要注意的是，如果备选物体表面粗糙度都大于光波长时，表面粗糙度就不能作为选择测量目标的有效依据了。

## 第4章 激光相干语音侦测系统信号处理

对于激光相干语音侦测系统而言,除了光学系统的优劣以及振动测量目标选择的好坏之外,信号处理性能的好坏也是决定系统能否获得高质量语音信号的关键因素。语音侦测系统的光学结构以及参数一旦确定,振动测量目标也选择完毕,此时,获取的语音信号质量主要取决于信号处理系统。因此,选择合适的信号处理算法对于实现远距离、高质量的语音侦测至关重要。本章首先针对激光相干语音侦测的应用需求,采用正交解调与反正切鉴相结合的解调算法提取多普勒频移,重构语音信号。为了提高语音质量,采用适合于相干侦测语音的降噪技术抑制语音中存在的噪声。最后,采用语音带宽扩展算法扩展相干侦测语音的带宽,提高语音的可懂度和自然性。

### 4.1 语音质量评价指标

为了衡量激光相干语音侦测系统获得的语音信号质量,需要对侦测语音进行质量评价。由于语音质量好坏是人耳的主观感受,因此语音质量评价不仅仅是一个信号处理过程,还和心理学、生理学等学科紧密结合。大体而言,语音质量评价可分为主观评价和客观评价。

#### 4.1.1 主观评价

主观评价指的是通过人的主观感受判定语音质量。主观评价法可以直接反映人对语音质量的感受,但是为了测试的客观性,需要大量的人力和时间,并且评价的重复性与稳定性较差。语谱图和平均意见得分(Mean Opinion Score, MOS)是两种使用比较普遍的主观评价方法。

##### (1) 语谱图

语谱图是一种可以描述语音频率与时间变化关系的图形。在语谱图中,横轴为时间轴,纵轴为频率轴,图上的条纹表示在特定时刻下的语音频率成分,频率成分的强弱用灰度值或者色调表示。语谱图反映了语音信号频谱的动态变化,并包含丰富的语音特性,因此也被称为可视语音<sup>[105]</sup>。

## (2) MOS

MOS<sup>[106]</sup>指得是由多名试听者按照确定的评分标准对测试的语音质量评分,取所有分数的平均分作为测试语音质量的得分。表 4.1 为 MOS 评分标准。

表 4.1 MOS 评分标准

Table 4.1 The score standard of MOS

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

### 4.1.2 客观评价

客观评价是通过测试语音与标准语音之间差异的大小作为评判标准来评价测试语音的质量。客观评价法不受到人主观性的干扰,重复性和鲁棒性强,并且不需要大量时间和人力,易于实现。目前,使用比较普遍的客观评价方法有:分段信噪比、坂仓-斋田(Itakura-Saito)谱失真度<sup>[107]</sup>、语音质量评价(Perceptual Evaluation of Speech Quality, PESQ)<sup>[108-109]</sup>。

#### (1) 分段信噪比

信噪比是一种可以有效评价信号失真程度的方法,但是一般的信噪比只能提供整体信号失真度,然而语音信号是时变信号,不同帧之间语音能量不同,因此为了更准确的描述语音信号的失真程度,可以采用分段信噪比。分段信噪比是指计算每一帧测试语音的信噪比,然后取所有帧信噪比的均值作为测试语音质量的得分,其定义为:

$$SNR_{seg} = \frac{1}{L_{seg}} \sum_{l=1}^{L_{seg}} SNR(l) \quad (4-1)$$

公式(4-1)中,  $L_{seg}$  表示帧数,  $SNR(l)$ 表示第  $l$  帧的信噪比, 具体可表示为:

$$SNR(l) = 10 \log_{10} \frac{\sum_n s_l^2(n)}{\sum_n [s_l(n) - s_{l-n}(n)]^2} \quad (4-2)$$

公式(4-2)中,  $s_l(n)$ 表示为第  $l$  帧的标准语音,  $s_{l-n}(n)$ 表示为第  $l$  帧的测试语音。

### (2) 坂仓-斋田谱失真度

坂仓-斋田谱失真度是通过衡量标准语音和测试语音两者的线性预测参数之间的差异来量化测试语音的质量。线性参数之间的差异用坂仓距离  $d_{IS}$  来定义, 其可以表示为:

$$d_{IS}(s, s_n) = \frac{\sigma_{s_n}^2}{\sigma_s^2} \frac{\lambda_{s_n} R_{s_n} \lambda_{s_n}^T}{\lambda_s R_s \lambda_s^T} - \ln \left| \frac{\sigma_{s_n}^2}{\sigma_s^2} \right| - 1 \quad (4-3)$$

公式(4-3)中,  $\sigma_{s_n}$  和  $\sigma_s$  分别表示为标准语音  $s_n$  和测试语音  $s$  的声道幅度谱增益,  $\lambda_{s_n}$  和  $\lambda_s$  分别表示为标准语音  $s_n$  和测试语音  $s$  的线性预测系数,  $R_{s_n}$  和  $R_s$  分别表示为标准语音  $s_n$  和测试语音  $s$  的自相关矩阵。坂仓距离  $d_{IS}$  越小, 说明测试语音的失真度也越小, 语音质量也越高。

### (3) PESQ

PESQ 是 2001 年, 由国际电信联盟标准化组织 (International Telecommunications Union Telecommunication Standardization Sector, ITU-T) 提出的一种评价语音质量的方法。与其他的客观评价方法相比, PESQ 算法基于听觉和认知模型将语音频率、响度等物理特性与人类感知特性通过相互对应, 从而来模拟主观评价, 因此该方法是目前与主观评价相关度最高(相关度 0.97)的算法, 除此之外适用性广泛。PESQ 是通过计算标准语音与测试语音两者之间响度的误差来衡量测试语音的质量, 其公式可以表示为:

$$PESQ = a_0 + a_1 D_{ind} + a_2 A_{ind} \quad (4-4)$$

公式(4-4)中,  $a_0=4.5$ 、 $a_1=-0.1$ 、 $a_2=-0.0309$  为平衡参数,  $D_{ind}$  为对称平均扰动值,  $A_{ind}$  为非对称平均扰动值。PESQ 得分范围为 0.5~4.5 之间, 其得分越高, 说明测试语音的失真度越小, 语音质量也越高。

### 4.1.3 本文采用的评价标准

由于本论文的目的是为了获得可理解、听觉体验好的语音信号，语音的可懂度及听觉体验是一个较为主观的感受，因此，本文主要以主观评价的方式评价侦测语音质量。主观评价中包括 MOS 和语谱图。在 MOS 评价过程中，试听者由作者所在的部门中随机寻找的 10 名成员组成。

## 4.2 相干侦测语音信号重构

根据第二章可知，激光相干语音侦测的过程是通过检测物体的振动信息（位移或者速度）从而获得语音信号，因此重构语音的过程就是获得物体振动信息的过程。由于光电探测器输出的信号为载波信号，为了得到物体的振动信息，需要对载波信号进行解调。

### 4.2.1 传统载波解调技术

#### (1) 模拟速度解调

模拟速度解调是一种比较经典的解调方法，曾经被几乎所有的商用激光相干测振仪采用。模拟速度解调的核心思想是通过将瞬时多普勒频率转换为与速度成比例的电压值以获得物体的振动速度。模拟速度解调器的具体实现方案有：锁相环(phase locked loop, PLL)，脉冲密度或延迟线解调器等<sup>[110]</sup>，这些实现方案都有各自的优势和不足。器件的质量是模拟速度解调器中最重要的一环，器件的好坏直接影响了解调的精度，准确度和灵敏度。如图 4.1 所示为实现模拟解调的一般方案。

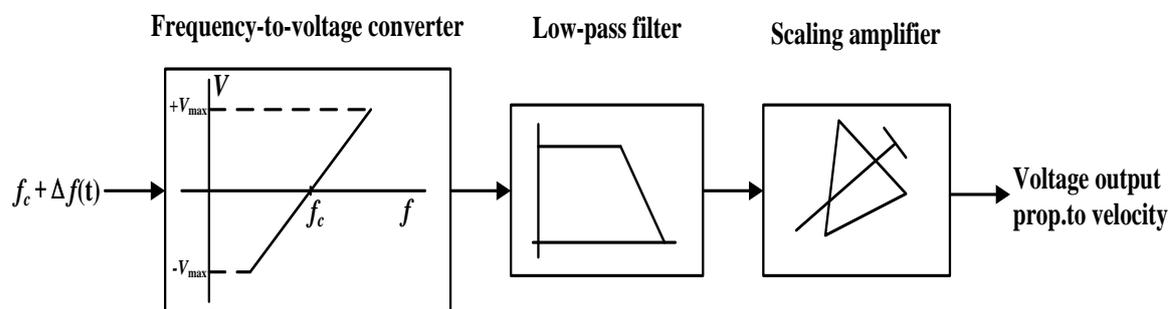


图 4.1 模拟速度解码器简化方案

Figure 4.1 Simplified scheme of an analog velocity decoder.

如图 4.1 所示，输入信号为中心频率为  $f_c$  的调频信号。线性频率-电压转换器

(frequency-to-voltage)产生与多普勒频移  $\Delta f(t)$ 成正比的电压,通过检测解调器输出的电压以获得瞬时多普勒频率,从而获得振动速度及位移。一般而言,模拟速度解调器输出电压前,需要通过低通滤波器来抑制寄生的高频成分,并且输出的电压可通过可调放大器进行适当的缩放和校准。设计精良的模拟速度解调器具有工作频率范围大,性能优异,实时响应等优点。但是,解调精度很大程度上取决于信号处理链中的每一个组件,并且容易发生热漂移和老化。因此,不能满足激光相干语音侦测的场景应用。

## (2) 条纹计数位移解调

条纹计数位移解调法是一个经典的从调制的多普勒信号中恢复位移信息的解调方法。在探测器中,观察到一个周期波形通过时,意味着多普勒相位  $\varphi_d(t)$  经过了一个完整的周期或者是  $2\pi$  的增量。根据公式(2-10)可得,当多普勒相位  $\varphi_d(t)$  移动  $2\pi$  时对应于振动位移  $s(t)$  移动了  $\lambda/2$ ,因此通过简单的周期或零交叉计数可实现对于较大位移的振动的精确测量,分辨率分别为  $\lambda/2$  (周期计数)或  $\lambda/4$  (零交叉计数)。此外,通过使用脉冲内插值的方法可提高测量分辨率,基于该方法,解调系统可实现  $2\text{nm}$  的分辨率,但是,由于技术原因,只能在相对低振动速度的情况下使用。总体而言,条纹计数法只能应用于低振动频率、大振动幅度的测量场景中,远远达不到激光相干语音侦测的应用需求。

### 4.2.2 正交解调与反正切鉴相结合的解调算法

反正切鉴相算法具有解调精度高、鲁棒性好等优点,因此采用反正切鉴相算法计算相位角是静态干涉仪和其他输出信号为正交格式的位移传感器的标准解调算法<sup>[110]</sup>。只要激光相干语音侦测系统可以产生正交格式的基带信号并且具有足够的计算力,同样可以采用反正切鉴相算法获得多普勒频移。目前,数字信号处理硬件的发展,可以确保激光相干语音侦测系统具备足够的计算能力,由此可得,反正切鉴相算法实现的关键是获得正交基带信号。

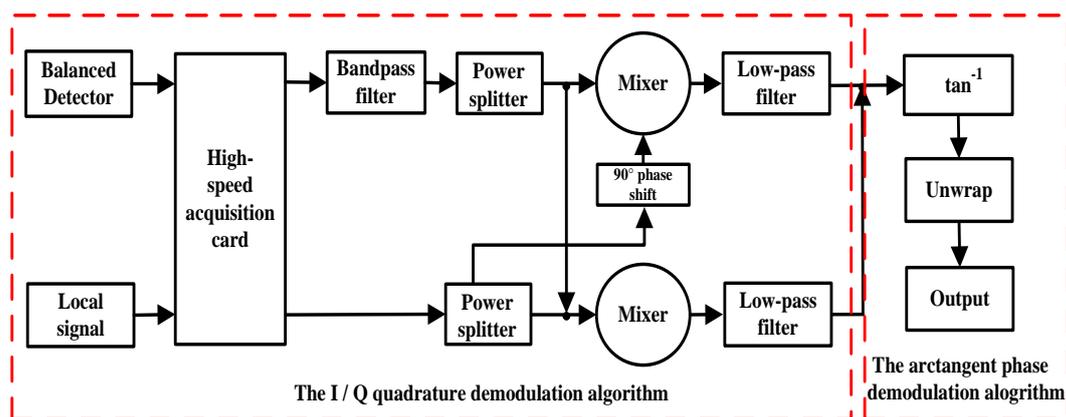


图 4.2 信号处理流程图

Figure 4.2 The diagram of the signal processing unit.

正交解调是获得正交基带信号的有效算法之一，该算法物理思想简单，实现方便，抗干扰性强。因此，本文使用正交解调算法获取正交基带信号，再对正交基带信号使用反正切鉴相算法以解调出多普勒相位信号。如图 4.2 所示为正交解调与反正切鉴相算法相结合的信号处理流程图。由公式(2-7)可知，光电探测器输出的信号为以频率  $\omega_{if}$  为中心的载波信号。不考虑输出电流的空间位置，只考虑时间，则公式(2-7)可以表示为：

$$i_{IF}(t) = c \cos(\omega_{if}t + \omega_d t) = c \cos\left[\omega_{if}t + \frac{4\pi}{\lambda} s(t)\right] = c \cos[\omega_{if}t + \varphi_d(t)] \quad (4-5)$$

公式(4-5)中， $c$  为电流幅度，可以表示为：

$$c = 2 \operatorname{Re}\left(\iint \beta(x, y) \{E_{Lo}(x, y) E_s^*(x, y)\} dx dy\right) \quad (4-6)$$

同时，以频率  $\omega_{if}$  为中心，生成本地振荡信号  $i_{loc}(t)$

$$i_{loc}(t) = \cos(\omega_{if}t) \quad (4-7)$$

将探测器输出信号  $i_{IF}(t)$  与本地振荡信号  $i_{loc}(t)$  相乘，并将相乘所得的结果通过低通滤波器即可得到基带信号  $u_I(t)$ 。

$$\begin{aligned} u_I(t) &= LPF[i_{IF}(t)i_{loc}(t)] = LPF\left[c \cos[\omega_{if}t + \varphi_d(t)] \cos(\omega_{if}t)\right] \\ &= LPF\left[\frac{c}{2}(\cos[\varphi_d(t)] + \cos[2\omega_{if}t + \varphi_d(t)])\right] \\ &= \frac{c}{2} \cos[\varphi_d(t)] \end{aligned} \quad (4-8)$$

公式(4-8)中,  $LPF$  代表低通滤波器。类似的, 与基带信号  $u_I(t)$  相互正交的基带信号  $u_Q(t)$ , 可通过将探测器输出信号  $i_{IF}(t)$  与经过  $\pi/2$  移相后本地振荡信号  $i'_{loc}(t)$  相乘, 并将相乘所得的结果通过低通滤波器得到。

$$\begin{aligned} u_Q(t) &= LPF [i_{IF}(t) i'_{loc}(t)] = LPF \left[ c \cos [\omega_{if} t + \varphi_d(t)] \cos \left( \omega_{if} t + \frac{\pi}{2} \right) \right] \\ &= LPF \left[ \frac{c}{2} (\sin [\varphi_d(t)] + \sin [2\omega_{if} t + \varphi_d(t)]) \right] \\ &= \frac{c}{2} \sin [\varphi_d(t)] \end{aligned} \quad (4-9)$$

根据正交解调得到的正交基带信号  $u_I(t)$  与  $u_Q(t)$ , 物体的振动位移  $s(t)$  可通过反正切鉴相算法获得。

$$s(t) = \frac{\lambda}{4\pi} \varphi_d(t) = \frac{\lambda}{4\pi} \arctan \left( \frac{u_Q}{u_I} \right) + \frac{\lambda}{4} m \quad (4-10)$$

直接使用反正切鉴相算法进行解调会引起相位跃变, 因此需要使用相位解缠算法对相位跃变进行修正。根据振动位移  $s(t)$  便可获得语音信号。

#### 4.2.3 重构算法实验分析

在 4.2.2 节中对使用正交解调与反正切鉴相结合的解调算法重构语音的过程进行了分析, 为了更加直观和深入的分析该算法对相干侦测系统侦测语音的影响, 进行了语音相干侦测实验。

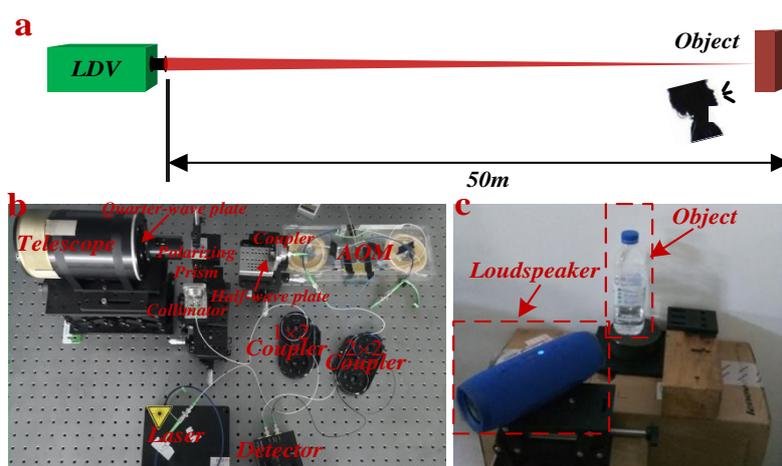


图 4.3 (a) 语音重构实验框图。 (b) 激光相干语音侦测系统。 (c) 被测目标

Figure 4.3 (a) Speech reconstruction experiment block diagram. (b) Prototype of the LDV system. (c) Target.

实验框图如图 4.3 所示。音箱播放事先准备好的纯净语音，发出的声波驱动间隔为 45cm 处的振动目标矿泉水瓶发生振动，振动目标的振动由距离约为 50m 的自研激光相干语音侦测系统（激光相干语音侦测系统详见第一章）检测，使用正交解调与反正切鉴相结合的解调算法重构语音。

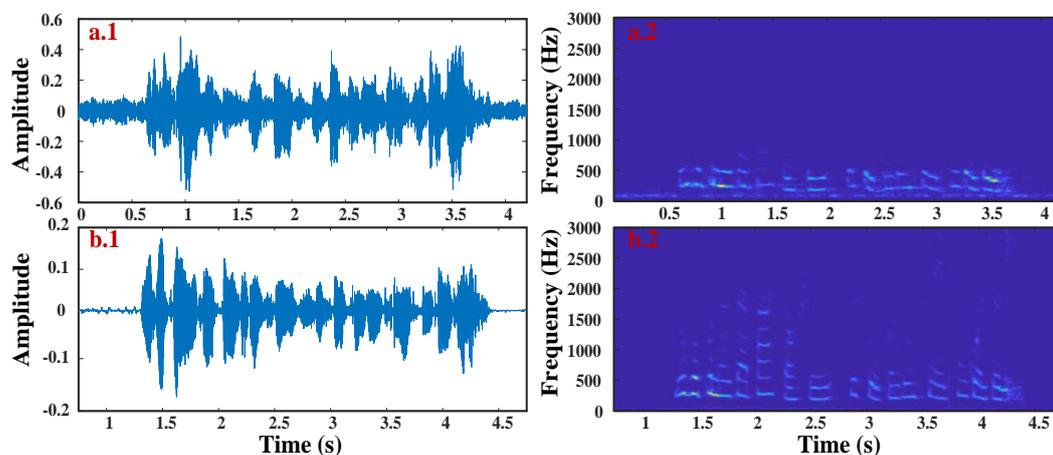


图 4.4 (a) 重构语音波形图与语谱图。 (b) 纯净语音波形图与语谱图

Figure 4.4 (a) The waveform and spectrogram of the reconstructed speech. (b) The waveform and spectrogram of the clean speech.

图 4.4.a 与图 4.4.b 分别为重构语音的波形图与语谱图以及纯净语音的波形图与语谱图。比较图 4.4.a 与图 4.4.b 可以发现，重构语音无论是语谱图还是波形图都和纯净语音十分相似，并且重构语音的 MOS 得分为 3.1，这说明正交解调与反正切鉴相结合的解调算法可以有效重构出可理解语音信号。但是，值得注意的是重构语音的频率范围在 1000Hz 以下，高频部分丢失，这是由于物体在声场高频振动不明显引起的。另外，从波形图和语谱图上可以观察到重构语音受到噪声的干扰，这对听觉体验造成了影响。

### 4.3 相干侦测语音信号降噪

观察图 4.4 可以发现，使用正交解调与反正切鉴相结合的解调算法可成功重构出语音信号，但是由于系统噪声、环境噪声以及物体粗糙面上固有的散斑问题，会使得重构出的相干侦测语音信号存在噪声，这些噪声会严重损害听众的听觉体验，虽然可以通过带通滤波等方式滤除人声频域范围外的噪声，然而，面对频域范围内的噪声便无能为力，因此需要对相干侦测获得的语音进行降噪处理。然

而，由于激光相干语音侦测技术的特殊性，使得相干侦测获得的语音中的噪声特性不同于传统语音信号的噪声，直接使用现有的语音降噪算法不能很好的满足应用需求。因此需要对相干侦测语音的噪声特点进行研究，以提出适用于相干侦测得到的语音的降噪方法。通过对大量的语音侦测实验（实验细节详见第5章）中获取的相干侦测语音进行分析，发现噪声在相干侦测语音的表现形式可大致分为背景噪声和冲击噪声两大类。

### (1) 背景噪声

造成背景噪声的因素有很多，例如相干侦测系统的散粒噪声、环境噪声以及信号处理的量化噪声等。图4.5为典型的相干语音侦测中的背景噪声。观察图4.5.a可以发现，整个相干侦测语音时间段内都存在背景噪声。观察图4.5.b可以发现背景噪声的频谱分布在整个语音频率范围内(300-3000Hz)，并且噪声能量分布均匀，但是小于语音能量。

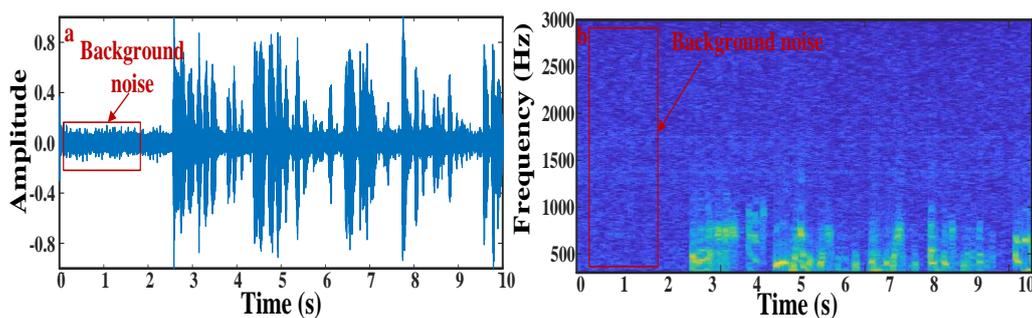


图 4.5 (a) 含噪语音波形图。 (b) 含噪语音语谱图

Figure 4.5 (a) The waveform of the noisy speech. (b) The spectrogram of the noisy speech.

### (2) 冲击噪声

冲击噪声主要是由于多普勒相位变化的不连续性引起的，而造成多普勒相位不连续变化的一个主要原因是动态散斑噪声（动态散斑对语音侦测的影响详见第3.3节）。除此之外，突发环境干扰，例如突然出现的喊叫声、鸣笛声以及撞击声也会引起冲击噪声。图4.6为典型的相干语音侦测中的冲击噪声。观察图4.6可以发现，冲击噪声具有能量强、频域分布广、持续时间短以及出现位置随机的特点。冲击噪声在语音信号中的表现为尖锐、刺耳、音调高、音量大的‘吱吱’声。

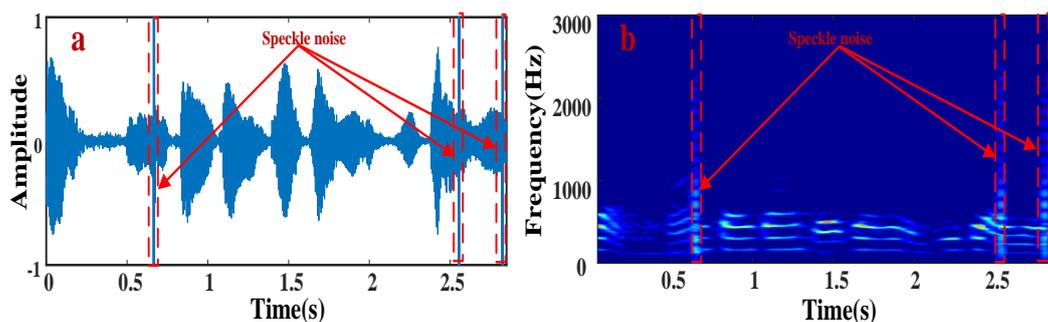


图 4.6 (a) 含噪语音波形图。 (b) 含噪语音语谱图

Figure 4.6 (a) The waveform of the noisy speech. (b) The spectrogram of the noisy speech.

由于背景噪声和冲击噪声这两种噪声差别很大，因此针对这两种噪声类型，本文采用不同的降噪方法。

#### 4.3.1 背景噪声抑制算法研究

由上文可知，侦测语音中存在的背景噪声能量分布均匀，可认为是高斯白噪声，这也是大多数语音降噪算法所要消除的噪声。现有的语音降噪算法根据其处理语音方式的不同可大致分为参数法<sup>[111~112]</sup>、子空间法<sup>[113~114]</sup>、小波法<sup>[115~116]</sup>、听觉遮蔽法<sup>[117~118]</sup>、短时谱估计法<sup>[119~124]</sup>等方法。参数法是通过估计语音生成模型来抑制噪声。但是在实际应用中，很难建立起准确的语音生成模型，因此局限性较大。子空间法则是利用语音信号的稀疏性将带噪语音分为语音子空间和噪声子空间，进而将噪声剔除。但是其计算量较大，实时性不强。小波法则是利用不同尺度下，语音信号和噪声的小波系数不同的性质来抑制噪声。但是小波法在分解尺度选择、阈值选取等方面存在诸多难点。听觉遮蔽则是利用人耳的听觉遮蔽特性掩盖噪声，但是该方法一般作为辅助算法。与上述其他方法相比，由于短时谱估计法既不需要从带噪语音中估计语音生成模型，也不需要构建训练集，并且计算复杂度较低，应用广泛，针对不同类型的背景噪声均有较好的表现。因此，本文采用基于短时谱估计的语音降噪方法抑制背景噪声。

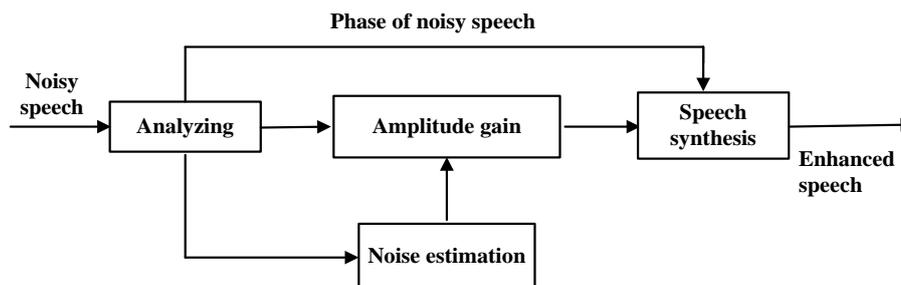


图 4.7 短时谱估计算法框图

**Figure 4.7 The block diagram of Short-Time Fourier Analysis-Modification-Synthesis algorithm.**

语音信号是一个时变、非稳态信号。但是，由于语音是由声带产生的激励信号通过声道形成的，而人的声道肌肉运动是缓慢的，因此语音信号在短时间内可认为是稳态、时不变的。因此，将语音在短时间内的稳态特性贯穿在语音降噪过程中，也就构成了短时谱估计语音增强技术。短时谱估计的语音增强算法框架如图 4.7 所示，首先对带噪语音信号进行短时傅里叶变换获得带噪语音频谱的幅度和相位。然后利用相关算法从带噪语音的幅度中估计出纯净语音频谱的幅度。最后，基于估计得到的纯净语音频谱的幅度和带噪语音的相位，利用短时傅立叶逆变换得到降噪语音。

传统的短时谱估计算法主要包括谱减法<sup>[119]</sup>、维纳滤波法<sup>[120]</sup>以及谱幅度最小均方误差估计法<sup>[121]</sup>。谱减法和维纳滤波法是短时谱估计中较早使用的方法。但是这两种算法很容易造成语音失真和引入音乐噪声。谱幅度最小均方误差估计法具有良好的语音降噪和削弱音乐噪声的效果，但是该方法是基于可以准确估计先验信噪比的前提下，如果噪声估计偏差较大时，则会进一步加剧噪声偏差。除了这些算法外，还有一些改进算法，例如 Martin 于 1994 年提出的基于最小统计量的增强方法<sup>[122]</sup>，并且他在 2001 年对该方法进行了改进，提出了更平滑的最小统计量噪声估计法<sup>[123]</sup>。但是，该算法对于非平稳噪声效果欠佳。最优对数谱估计算法(Optimal Log-Spectral, OM-LSA)<sup>[124]</sup>相较于其他短时谱估计的语音增强算法而言，具有噪声估计误差小，对非平稳噪声跟踪快的优点，可被认为是目前单声道语音增强算法中最优的方法。但是，该算法只对带噪语音的幅度进行降噪，而忽略了相位对于语音降噪的作用。本文采用 OM-LSA 算法估计纯净语音幅度，同时，对带噪语音的相位进行补偿，以实现抑制相干侦测语音中的背景噪声。

### (1) 最优对数谱增益函数估计

由图 4.5 可知, 背景噪声在整个语音频谱内均匀分布并和语音不相关, 因此可认为是加性高斯白噪声。假设  $x(t)$  和  $n(t)$  分别表示为纯净语音和噪声, 则带噪相干侦测语音可表示为

$$y(t) = x(t) + n(t) \quad (4-11)$$

通过窗函数对带噪相干侦测语音进行分帧处理, 并使用短时傅立叶变换 (short-time Fourier transform, STFT) 对每帧信号进行分析。此时, 带噪相干侦测语音在第  $l$  帧, 第  $k$  的频率点的频域表达式为:

$$Y(l, k) = X(l, k) + N(l, k) \quad (4-12)$$

假设  $H_0(l, k)$  和  $H_1(l, k)$  分别表示语音存在和语音不存在, 并假设语音和噪声的频域为复高斯分布, 则在探测到相干侦测语音情况下语音存在和不存在的概率密度函数分别为:

$$\begin{aligned} p(Y(l, k) | H_0(l, k)) &= \frac{1}{\pi \lambda_n(l, k)} \exp \left\{ -\frac{|Y(l, k)|^2}{\lambda_n(l, k)} \right\} \\ p(Y(l, k) | H_1(l, k)) &= \frac{1}{\pi (\lambda_x(l, k) + \lambda_n(l, k))} \exp \left\{ -\frac{|Y(l, k)|^2}{\lambda_x(l, k) + \lambda_n(l, k)} \right\} \end{aligned} \quad (4-13)$$

公式(4-13)中,  $\lambda_n(l, k)$  和  $\lambda_x(l, k)$  分别表示为纯净语音和噪声的方差。根据贝叶斯定理可知, 语音存在的概率  $p(l, k) = P(H_1(l, k) | Y(l, k))$  可以表示为:

$$p(l, k) = \left\{ 1 + \frac{q(l, k)}{1 - q(l, k)} (1 + \xi(l, k)) \exp(-\nu(l, k)) \right\}^{-1} \quad (4-14)$$

公式(4-14)中,  $q(l, k) = P(H_0(l, k))$  为语音不存在的先验概率 (Speech absence probability, SAP),  $\xi(l, k) = \lambda_x(l, k) / \lambda_n(l, k)$  为先验信噪比,  $\nu(l, k)$  可表示为:

$$\nu(l, k) = \frac{\gamma(l, k) \xi(l, k)}{1 + \xi(l, k)} \quad (4-15)$$

公式(4-15)中,  $\gamma(l, k) = |Y(l, k)|^2 / \lambda_n(l, k)$  为后验信噪比。

假设纯净语音频谱的幅度为  $A$ ,  $A'$  为经过最优对数谱估计得到的幅度, 假设频谱统计独立, 则

$$A' = \exp \left\{ E \left[ \log A(l, k) | Y(l, k) \right] \right\} = G(l, k) | Y(l, k) | \quad (4-16)$$

公式(4-16)中,  $G(l,k)$ 表示最优对数谱增益函数。根据统计模型可知:

$$E[\log A(l,k)|Y(l,k)] = E[\log A(l,k)|Y(l,k), H_1(l,k)]p(l,k) + E[\log A(l,k)|Y(l,k), H_0(l,k)](1-p(l,k)) \quad (4-17)$$

因此, 根据公式(4-17)可知, 当语音不存在时, 公式(4-16)可表示为:

$$A' = \exp\left\{E[\log A(l,k)|Y(l,k), H_0(l,k)]\right\} = G_{\min}|Y(l,k)| \quad (4-18)$$

公式(4-18)中,  $G_{\min}$ 为下限阈值。当语音存在时, 公式(4-16)可表示为:

$$A' = \exp\left\{E[\log A(l,k)|Y(l,k), H_1(l,k)]\right\} = G_{H_1}(l,k)|Y(l,k)| \quad (4-19)$$

根据[121]可知, 公式(4-19)中, 增益函数  $G_{H_1}$ 可表示为

$$G_{H_1}(l,k) = \frac{\xi(l,k)}{1+\xi(l,k)} \exp\left(\frac{1}{2} \int_{v(l,k)}^{\infty} \frac{e^{-t}}{t} dt\right) \quad (4-20)$$

将公式(4-18)和公式(4-19)代入公式(4-16)可得, 最优对数谱增益函数  $G(l,k)$ 可表示为:

$$G(l,k) = \left\{G_{H_1}(l,k)\right\}^{p(l,k)} \cdot G_{\min}^{1-p(l,k)} \quad (4-21)$$

观察公式(4-20)可知, 先验信噪比和后验信噪比是建立最优对数谱增益函数的关键参数。根据[121]可知, 先验信噪比可表示为

$$\xi'(l,k) = \alpha G_{H_1}^2(l-1,k) \gamma(l-1,k) + (1-\alpha) \max\{\gamma(l,k) - 1, 0\} \quad (4-22)$$

公式(4-22)中,  $\alpha$ 表示常数。后验信噪比是带噪相干侦测语音能量与噪声方差的比值即  $\gamma(l,k) = |Y(l,k)|^2 / \lambda_n(l,k)$ 。根据[124~125]可知, 语音信号中的噪声方差可估计为:

$$\lambda'_n(l,k) = \lambda'_n(l-1,k) \alpha_n(l,k) + (1-\alpha_n(l,k)) |Y(l-1,k)|^2 \quad (4-23)$$

公式(4-23)中, 参数  $\alpha_n$ 表示为:

$$\alpha_n(l,k) = \alpha_d + (1-\alpha_d) p(l,k) \quad (4-24)$$

公式(4-24)中,  $\alpha_d$ 表示常数。

## (2) 语音存在概率估计

观察公式(4-21)至(4-24)可知, 最优对数谱增益函数的建立、先验信噪比和后验信噪比的估计都是建立在语音存在概率估计的基础上进行的。本文利用 Israel

Cohen 在 2002 年提出的一种先验语音不存在概率估计算法来对语音存在概率进行估计<sup>[124]</sup>。首先基于先验信噪比, 并充分考虑连续语音相邻帧之间的强相关性推导出三个参数, 然后对这三个参数进行软判决从而计算出先验语音不存在概率, 最后, 根据计算获得的先验语音不存在概率, 估计出语音存在概率。

假设  $\zeta(l,k)$  为具有时间常数  $\beta$  的先验信噪比的递归均值:

$$\zeta(l,k) = \beta\zeta(l-1,k) + (1-\beta)\xi(l-1,k) \quad (4-25)$$

在频域上, 考虑频率点之间的相关性, 通过不同尺寸的窗函数求取先验信噪比递归均值  $\zeta(l,k)$  的局部均值  $\zeta_{\text{local}}(l,k)$  和全局均值  $\zeta_{\text{global}}(l,k)$ 。并定义参数  $P_{\text{local}}(l,k)$  和参数  $P_{\text{global}}(l,k)$  分别表示为局部均值和全局均值与第  $l$  帧, 第  $k$  的频率点的关系。

$$P_i(l,k) = \begin{cases} 0, & \zeta_i(l,k) \leq \zeta_{\min} \\ \frac{\log(\zeta_i(l,k)/\zeta_{\min})}{\log(\zeta_{\max}/\zeta_{\min})}, & \zeta_{\min} < \zeta_i(l,k) < \zeta_{\max} \\ 1, & \zeta_i(l,k) \geq \zeta_{\max} \end{cases} \quad (4-26)$$

公式(4-26)中,  $i$  表示'local'或者'global'。  $\zeta_{\min}$  和  $\zeta_{\max}$  为常数, 用来保留弱语音, 同时衰减噪声。为了进一步抑制噪声帧中的噪声分量, 除了定义参数  $P_{\text{local}}(l,k)$  和参数  $P_{\text{global}}(l,k)$  外, 根据相邻帧语音能量, 又定义了参数  $P_{\text{frame}}(l,k)$ 。

$$P_{\text{frame}}(l,k) = \begin{cases} 0, & \zeta_{\text{frame}}(l) \leq \zeta_{\min} \\ 1, & \zeta_{\min} < \zeta_{\text{frame}}(l-1) < \zeta_{\text{frame}}(l) \\ \mu(l), & \zeta_{\min} < \zeta_{\text{frame}}(l) < \zeta_{\text{frame}}(l-1) \end{cases} \quad (4-27)$$

公式(4-27)中,  $\zeta_{\text{frame}}$  表示频域的均值。

$$\zeta_{\text{frame}}(l) = \frac{1}{N/2+1} \sum_{k=1}^{N/2+1} \{\zeta(l,k)\} \quad (4-28)$$

$\mu(l)$  表示语音与噪声之间的软过渡, 具体表示为:

$$\mu(l) = \begin{cases} 0, & \zeta_{\text{frame}}(l) \leq \zeta_{\min} \zeta_{\text{peak}}(l) \\ \frac{\log(\zeta_{\text{frame}}(l) \zeta_{\min} / \zeta_{\text{peak}}(l))}{\log(\zeta_{\max} / \zeta_{\min})}, & \zeta_{\min} \zeta_{\text{peak}}(l) < \zeta_{\text{frame}}(l) < \zeta_{\max} \zeta_{\text{peak}}(l) \\ 1, & \zeta_{\text{frame}}(l) \geq \zeta_{\max} \zeta_{\text{peak}}(l) \end{cases} \quad (4-29)$$

公式(4-29)中,  $\zeta_{\text{peak}}$  为  $\zeta_{\text{frame}}$  的峰值。

此时, 根据参数  $P_{\text{local}}(l,k)$ 、 $P_{\text{global}}(l,k)$  以及  $P_{\text{frame}}(l,k)$ , 可得语音不存在的先验概率  $q'(l,k)$  为:

$$q'(l,k) = 1 - P_{\text{local}}(l,k)P_{\text{global}}(l,k)P_{\text{frame}}(l,k) \quad (4-30)$$

根据公式(4-30), 便可获得语音存在概率。

### (3) 相位修正

由于人耳对相位变化不敏感, 因此, 传统的短时谱估计算法只对带噪语音幅度进行增强, 而忽略了对带噪语音相位的处理。然而, 越来越多的研究和实验结果表明相位改善对于语音信号质量的提高有重要的作用。Wójcicki<sup>[126]</sup>、张文璐<sup>[28]</sup>等人通过实验发现在不改变语音频谱幅值的情况下, 仅对语音相位进行修正, 可以对语音信号进行增强。因此, 本文在不改变幅值的情况下, 采用语音信号短时频谱的共轭对称性对相位进行修正, 以达到语音增强的目的。该方法由 Wójcicki 于 2008 年提出, 是一种较为有效的语音相位修正方法<sup>[126]</sup>。

由于语音信号是实信号, 因此它的短时傅立叶频谱是共轭对称的, 即  $Y(l,k) = Y^*(l,N-k)$ 。因此, 通过改变共轭对称向量  $Y(l,k)$  与  $Y^*(l,N-k)$  在矢量图上角度, 来实现相位修正。相位修正过程如图 4.8 所示, 首先在带噪语音的短时傅立叶频谱上添加一个频率补偿函数  $\Lambda(l,k)$

$$Y_{\Lambda}(l,k) = Y(l,k) + \Lambda(l,k) \quad (4-31)$$

公式(4-31)中, 频率补偿函数可表示为

$$\Lambda(l,k) = \alpha_{\Lambda} \begin{cases} \lambda_n, & 0 \leq k/N < 0.5 \\ -\lambda_n, & 0.5 \leq k/N \leq 1 \end{cases} \quad (4-32)$$

公式(4-32)中,  $\alpha_{\Lambda}$  为补偿因子。从公式(4-32)中可得, 噪声越强, 相位修正值也就越大。此时, 通过反正切函数便可获得修正的相位:

$$\angle Y_{\Lambda}(l,k) = \arctan \left( \frac{\text{Im}\{Y_{\Lambda}(l,k)\}}{\text{Re}\{Y_{\Lambda}(l,k)\}} \right) \quad (4-33)$$

观察图 4.8.a 可以发现, 当噪声能量小于信号幅度时, 经过修正后得到的信号的实值部分与原始信号接近, 但是稍小于原始信号。观察图 4.8.b 可以发现, 当噪声能量接近或者大于信号幅度时, 修正后的信号与原始信号相比经过了较大

的削弱。因此，相位修正的实质是按照噪声与信号的比值，对带噪信号进行不同程度的削弱，以达到降噪的目的。

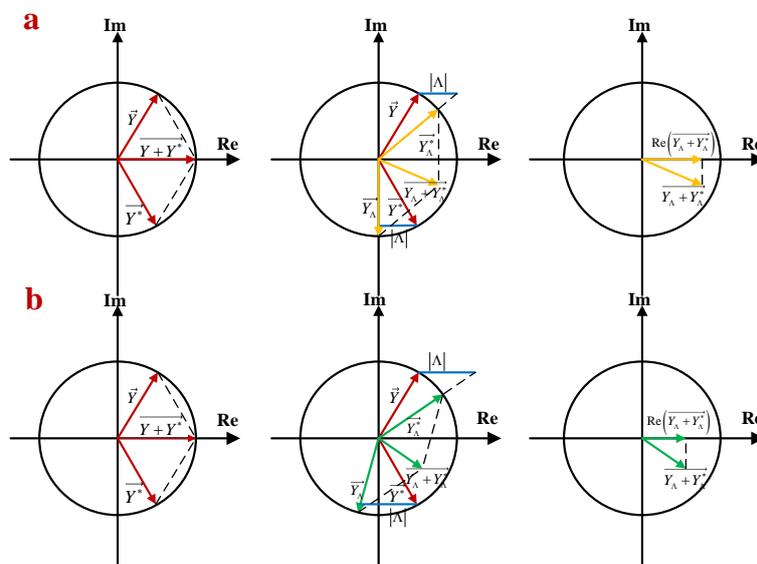


图 4.8 相位修正原理图

Figure 4.8 The schematic of Phase correction.

#### (4) 背景噪声抑制算法

本节采用的背景噪声抑制算法流程图如图 4.9 所示。采用相位修正算法对带噪语音的相位进行修正，同时，采用最优对数谱估计算法从带噪语音中估计纯净语音的幅度，将修正的相位与估计的幅度相结合便可获得降噪后的语音频谱  $X'(l, k)$

$$X'(l, k) = G(l, k) |Y(l, k)| \exp(j\angle Y_\Lambda(l, k)) \quad (4-34)$$

再对其做逆傅立叶变换便可获得降噪语音。

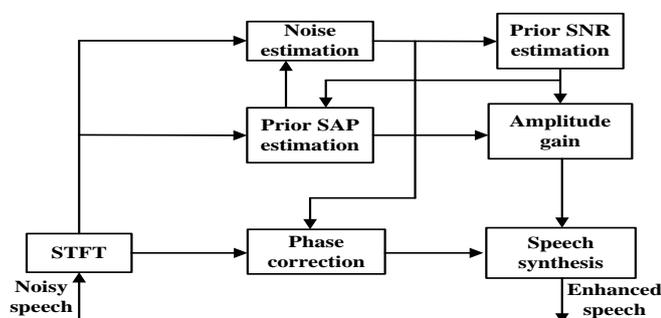


图 4.9 背景噪声抑制算法流程图

Figure 4.9 The schematic of background noise suppression algorithm.

### (5) 背景噪声抑制算法实验分析

采用实测相干侦测语音信号对背景噪声抑制算法性能进行验证,其验证结果如图 4.10 所示。图 4.10.a 为原始相干侦测语音的语谱图和波形图。图 4.10.b 为采用传统短时谱估计算法(维纳滤波法)降噪后的相干侦测语音的语谱图和波形图。图 4.10.c 为采用本文使用的降噪算法降噪后的相干侦测语音的语谱图和波形图。图 4.10.d 为纯净语音的语谱图和波形图。比较图 4.10.a、图 4.10.b 以及图 4.10.c 可以发现,本文使用的语音降噪算法与传统短时谱估计算法都可以有效抑制相干侦测语音中的背景噪声。然而,比较图 4.10.b、图 4.10.c 以及图 4.10.d 可以发现,传统短时谱估计算法造成了一定的语音失真,而本文使用的降噪算法表现则要优于传统的短时谱估计算法。另外使用 MOS 给语音评分。原始相干侦测语音 MOS 得分为 3.0。使用传统短时谱估计算法降噪后的语音 MOS 得分为 3.1。使用本节提出的语音降噪算法降噪后的语音 MOS 得分为 3.3。比较语音之间的 MOS 得分也可得出和图 4.10 相同的结论,即本文采用的降噪算法可以有效抑制背景噪声,提高语音质量,并且性能要优于传统的短时谱估计法。

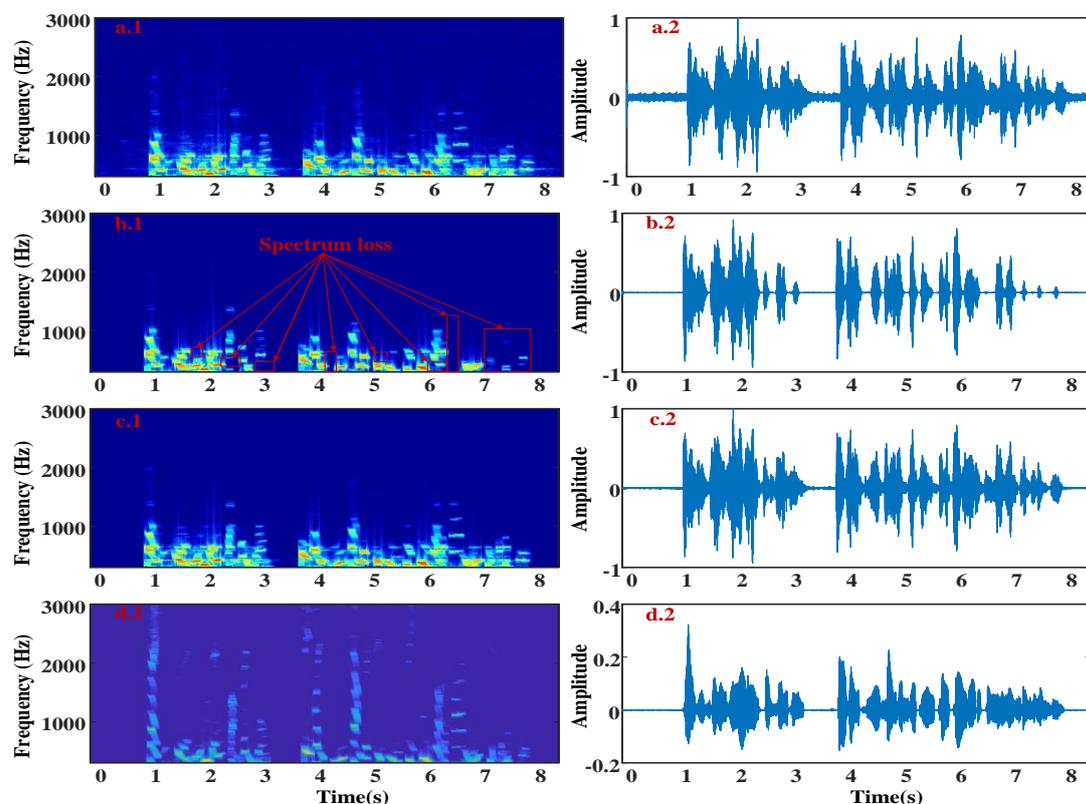


图 4.10 (a) 原始相干侦测语音波形图与语谱图。(b) 经过维纳滤波后语音波形图与语谱图。  
(c) 经过本节算法增强后语音波形图与语谱图。(d) 纯净语音波形图与语谱图

Figure 4.10 (a) The waveform and spectrogram of the raw speech. (b) The waveform and

**spectrogram of the speech enhanced by Wiener filtering. (c) The waveform and spectrogram of the speech enhanced by the algorithm proposed in this paper. (d) The waveform and spectrogram of the clean speech.**

### 4.3.2 冲击噪声抑制算法研究

相干侦测语音中的冲击噪声是一种较为特殊的噪声，它主要是由动态散斑噪声引起的，此外环境突发的强干扰也会引入这种形式的噪声。目前为止，已经提出了一些消除动态散斑噪声的方法，然而这些方法主要应用于激光相干测振方面。例如，J.Vass 用峰度值比率(kurtosis ratio)作为衡量标准，从振动位移数据段中选择未被动态散斑干扰的数据段作为振动测量结果<sup>[61]</sup>。然而，如果将此方法应用于激光相干语音侦测中，则会造成噪声语音数据段的丢失，影响语音理解。V.Aranchuk 使用非线性滤波器、系数缩放以及小波变化等方法滤除物体振动速度中的动态散斑噪声<sup>[34]</sup>。但是，如果将这些方法应用于激光相干语音侦测中，可能会造成语音畸变。本文在冲击噪声特性和语音信号处理方法的基础上，提出了一种消除算法以抑制相干侦测语音中的冲击噪声。

#### (1) 冲击噪声检测

由前文可知，冲击噪声具有随机出现的特性。因此为了消除冲击噪声，必须先确定噪声在相干侦测语音中的位置。观察图 3.25 和图 4.6 可以发现，冲击噪声的幅值与信号的幅值分布不同，并与信号幅值的均值存在极端差值。另一方面，峰度值可用于衡量随机实数概率分布的峰态。峰度值越大，意味着存在更多的极端值。因此，本文将峰度值作为判别指标，以检测冲击噪声在相干侦测语音中的位置。冲击噪声检测算法流程图如图 4.11 所示。

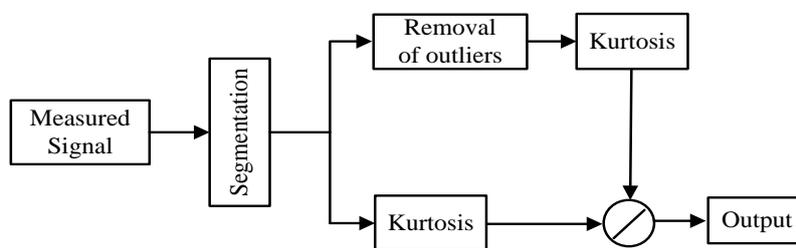


图 4.11 冲击噪声检测算法流程图

Figure 4.11 The block diagram of the algorithm for shock noise location.

1) 将相干侦测语音信号加窗分帧。由于统计指标的置信度随着样本数的增加而增加, 因此, 理论上每帧长度应该最大化。但是, 过长的帧长会影响时间分辨率, 从而不能准确检测出冲击噪声的位置。因此, 为了获得足够的帧长以保证可靠的峰度值估计和准确的时间分辨率, 本文采用高重叠分帧。

2) 对第  $l$  帧原始信号  $y^l$  进行阈值处理 (去除远离均值的数据点), 从而获得经过修整的信号  $y_t^l$ , 该过程可用公式(4-35)表示。

$$y_t^l = (thr_{low} < y^l[n] < thr_{high}) \quad (4-35)$$

公式(4-35)中,  $thr_{low}$  和  $thr_{high}$  分别为下限和上限。

3) 分别计算第  $l$  帧原始信号  $y^l$  和经过修整的信号  $y_t^l$  的峰度值。公式(4-36)为峰度值计算公式。

$$\beta_2(y) = \frac{\mu_4}{(\mu_2)^2} = \frac{E\{(x - \mu_y)^4\}}{\sigma_y^4} \quad (4-36)$$

公式(4-36)中,  $\mu_y$  和  $\sigma_y$  分别表示均值和标准差,  $E\{.\}$  表示为取均值。

4) 如公式(4-37)所示, 将第  $l$  帧原始信号  $y^l$  的峰度值  $\beta_2(y^l)$  与经过修整的信号  $y_t^l$  的峰度值  $\beta_2(y_t^l)$  相除, 以获得第  $l$  帧的峰度值比率  $KR^l$ 。若这一帧存在冲击噪声, 则原始信号的峰度值  $\beta_2(y^l)$  要远大于修整信号的峰度值  $\beta_2(y_t^l)$ 。因此, 若峰度值比率  $KR^l$  大于阈值  $thr_K$  时, 则可认为冲击噪声位于这帧数据中。

$$KR^l = \frac{\beta_2(y^l)}{\beta_2(y_t^l)} \quad (4-37)$$

图 4.12 为冲击噪声检测结果图。观察图 4.12 可以发现, 本文采用的冲击噪声检测算法可以准确检测出冲击噪声在相干语音信号中的位置, 这为冲击噪声消除奠定了基础

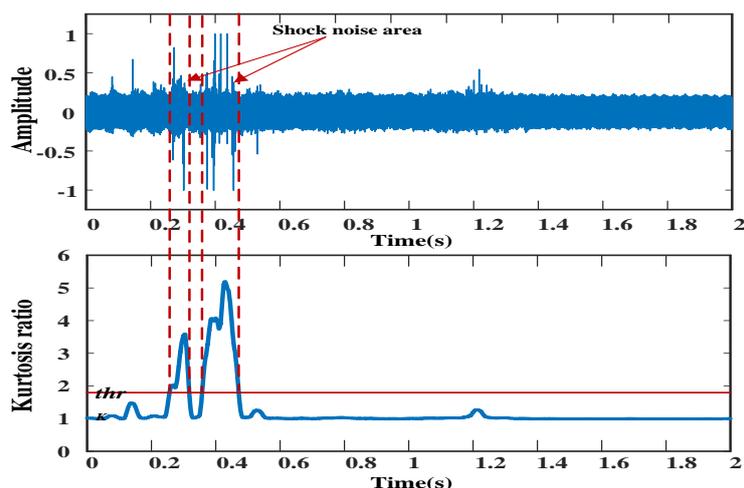


图 4.12 确认冲击噪声的位置

Figure 4.12 Confirmation of the shock noise area.

## (2) 冲击噪声消除

语音信号相邻点之间存在相关性，当前语音点可用过去的语音点线性拟合逼近，这个过程被为线性预测<sup>[127]</sup>。需要注意的是，相邻语音点之间的相关性随着间隔时间的增加而减弱，但是持续时间是冲击噪声的一个显著特征。因此，本文采用线性预测算法，从未被冲击噪声干扰的邻近语音片段中估计出受到干扰的原始语音，将估计出的原始语音代替受到干扰的语音片段，以达到消除冲击噪声的效果。

线性预测过程可用公式(4-38)表示

$$y(n) = \sum_{i=1}^p a_i y(n-i) + e(n) = y'(n) + e(n) \quad (4-38)$$

公式(4-38)中， $y'(n)$ 为  $y(n)$ 的估计值。 $a_i$ 为线性预测系数(Linear Prediction Coefficient, LPC)。 $e(n)$ 为线性预测误差。观察公式(4-38)可发现，一旦获得线性预测系数  $a_i$ ，便可估计出被冲击噪声干扰的原始语音。线性预测系数  $a_i$ 可通过最小均方误差准则求得。均方误差表示为

$$E_y = \frac{1}{n} \sum_n e^2(n) = \frac{1}{n} \sum_n [y(n) - y'(n)]^2 = \frac{1}{n} \sum_n \left[ y(n) - \sum_{i=1}^p a_i y(n-i) \right]^2 \quad (4-39)$$

为了得到最小均方误差，线性预测系数  $a_i$ 对均方误差的偏导为 0，即

$$\frac{\partial E_y}{\partial a_j} = \sum_n y(n) y(n-j) - \sum_{i=1}^p a_i \sum_n y(n-i) y(n-j) = 0 \quad (4-40)$$

公式(4-40)中,  $j=1,2\dots p; i=1,2\dots p$ 。定义

$$\phi(i, j) = \sum_n y(n-i)y(n-j) \quad (4-41)$$

则公式(4-40)可化简为:

$$\sum_{i=1}^p a_i \phi(i, j) = \phi(0, j) \quad (4-42)$$

语音信号  $y(n)$  的自相关函数可以表示为:

$$r(j) = \sum_{n=0}^{N-1} y(n)y(n-j) \quad (4-43)$$

公式(4-43)中,  $N$  为信号长度。根据公式(4-43), 公式(4-41)可表示为:

$$\phi(i, j) = r(i-j) \quad (4-44)$$

由于  $r(j)$  为偶函数, 因此公式(4-44)可表示为:

$$\phi(i, j) = r(|i-j|) \quad (4-45)$$

根据公式(4-45), 公式(4-42)可表示为:

$$\sum_{i=1}^p a_i r(|i-j|) = r(j) \Rightarrow \begin{bmatrix} r(0) & r(1) & r(2) & \cdots & r(p-1) \\ r(1) & r(0) & r(1) & \cdots & r(p-2) \\ r(2) & r(1) & r(0) & \cdots & r(p-3) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ r(p-1) & r(p-2) & r(p-3) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \\ r(3) \\ \vdots \\ r(p) \end{bmatrix} \quad (4-46)$$

因此, 通过求解方程(4-46)便可求得线性预测系数  $a_i$ 。值得注意的是, 方程(4-46)中的矩阵被称为 Toeplitz 矩阵, 方程(4-46)也被称为 Yule-Walker 方程。求解这类方程组不需要进行大量的运算, 可通过递归方法求解, Levinson-Durbin 算法是最常用的求解算法, 由于 Levinson-Durbin 算法是一个非常成熟的算法, 因此本文不再赘述, 具体解调过程可详见[106]。

本文采用的冲击噪声消除算法具体步骤为:

- 1) 通过冲击噪声检测算法确定冲击噪声在语音中的位置。
- 2) 计算冲击噪声存在处前后相邻帧的语音存在概率, 根据语音存在概率选择有效数据段  $y_e$ 。其中, 有效数据段的选择原则为: a)若前后相邻的语音存在概率都大于 0.5 或者前一帧语音存在概率大于 0.5, 而后一帧语音存在概率小于 0.5, 则选择前一帧作为有效数据段  $y_e$ 。b)若后一帧语音存在概率大于 0.5 而前一帧语

音存在概率小于 0.5，则选择后一帧作为有效数据段  $y_e$ 。c)若前后帧语音存在概率都小于 0.5，则认为冲击噪声位于噪声帧，可直接将冲击噪声所在处的数据用前后相临帧的均值代替。在实际应用中，可根据实际情况合理调整阈值。

3) 使用线性预测算法，从有效数据段中估计出原始语音  $y_{clean}$ 。然后用估计出的原始语音  $y_{clean}$  代替被冲击噪声干扰的语音  $y_{noise}$ ，从而起到消除冲击噪声的作用。在本文中，线性预测系数  $a_i$  的个数  $p=256$ 。

$$y_{clean} = \sum_{i=1}^{256} a_i y_e \quad (4-47)$$

### (3) 冲击噪声消除算法实验分析

采用实测相干侦测语音信号对冲击噪声消除算法性能进行验证，其验证结果如图 4.13 所示。图 4.13.a 为原始相干侦测语音的语谱图和波形图。图 4.13.b 为采用冲击噪声消除算法降噪后的相干侦测语音的语谱图和波形图。图 4.13.c 为纯净语音的语谱图和波形图。比较图 4.13.a、图 4.13.b 以及图 4.13.c 可以发现，本文使用的冲击噪声消除算法可以有效抑制相干侦测语音中的冲击噪声。并且使用 MOS 给语音评分。原始相干侦测语音 MOS 得分为 3.0。使用冲击噪声消除算法降噪后的语音 MOS 得分为 3.5。比较原始语音和增强语音之间的 MOS 得分也可得出和图 4.13 相同的结论，即本文采用的冲击噪声消除算法可以有效抑制冲击噪声，提高语音质量。

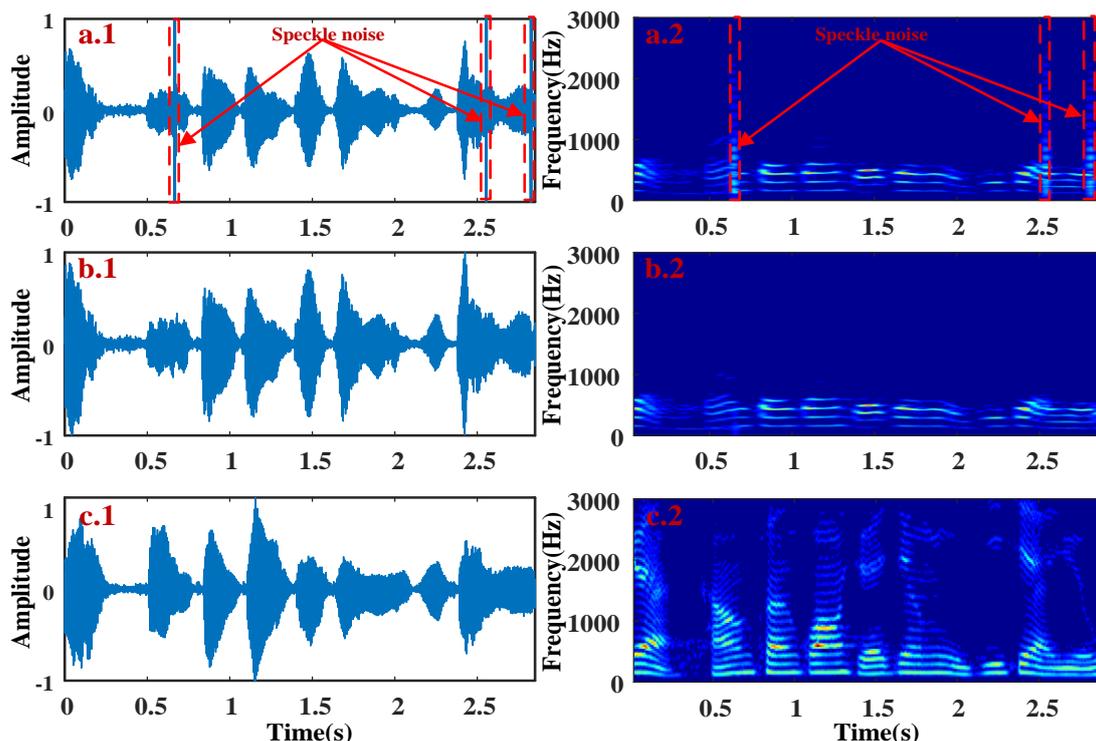


图 4.13 (a) 原始相干侦测语音波形图与语谱图。(b) 经过本节算法增强后语音波形图与语谱图。(c) 纯净语音波形图与语谱图

Figure 4.13 (a) The waveform and spectrogram of the raw speech. (b) The waveform and spectrogram of the speech enhanced by the algorithm proposed in this paper. (c) The waveform and spectrogram of the clean speech.

#### 4.4 语音信号带宽扩展

根据 3.1 节所述, 常见的非金属材料在高频处振动响应较差, 这就导致了侦测语音过程中, 声音高频部分的丢失, 这虽然不会对语音理解造成过大的影响, 但是在一定程度上会降低语音的可懂度和自然性。对于激光相干语音侦测系统而言, 获得宽带语音信号的一个切实有效的解决方案是在侦测到的语音信号上人为的加入一些高频丢失成分, 这个过程称之为带宽扩展(Bandwidth Extension, BWE) [128~132]。该方案环境适应性好, 并且不需要对侦测过程进行限制。

##### 4.4.1 传统语音信号带宽扩展算法

如图 4.14 所示为传统的带宽扩展算法方案。正如图 4.14 所示, 传统的语音带宽扩展基本算法可分为三个步骤<sup>[129~132]</sup>: 首先, 根据窄带语音(即物体在声场中的振动位移)信号  $s(t)$  计算出窄带语音的激励源  $e_n(t)$ , 再根据窄带语音激励源

$e_n(t)$ 估计出宽带语音信号的激励源  $e_b(t)$ 。同步的, 根据窄带语音信号  $s(t)$ 估计出频谱包络  $a_b(t)$ 。最后, 将宽带信号激励源  $e_b(t)$ 和频谱包络  $a_b(t)$ 相互合成, 从而获得宽带语音信号  $s_b(t)$ 。其中宽带激励源估计和频谱包络估计是带宽扩展算法中最重要的两个步骤。

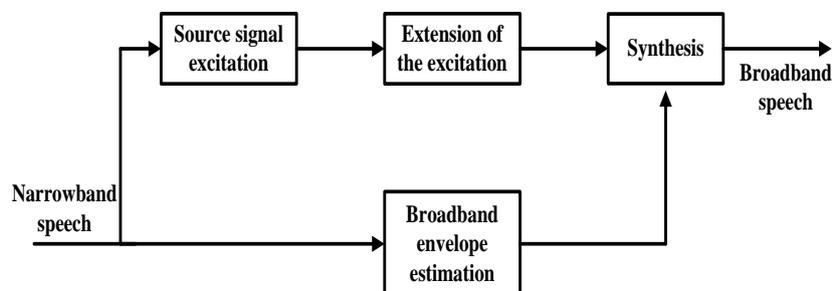


图 4.14 传统语音带宽扩展算法

Figure 4.14 Traditional speech bandwidth extension algorithm.

### (1) 宽带激励源估计

语音激励源包含语音频谱精细结构信息, 因此, 能否准确地从输入的窄带语音信号中估计出宽带语音信号的激励源, 直接决定了语音带宽扩展过程的成败。宽带激励源估计算法有很多, 大致可以分为: 直接信号产生、非线性失真、频谱移动等。

直接信号产生: 直接信号产生是指直接产生宽带激励源  $e_b(t)$ , 是最简单的宽带激励源估计算法。但是该算法非常依赖参数估计, 例如: 浊音度、基频频率以及增益。

非线性失真: 公式(4-48)所示, 非线性失真则是让窄带语音激励源  $e_n(t)$ 通过非线性函数  $g_{n-b}(\cdot)$ 从而估计出宽带激励源  $e_b(t)$ 。但是非线性函数输出结果很难预测, 输入信号参数发生微小变化, 就会极大影响输出结果, 并且很难找到最优的非线性函数  $g(\cdot)$ , 因此, 目前较少应用这种方法。

$$e_b(t) = g_{n-b}(e_n(t)) \quad (4-48)$$

频谱移动: 频谱移动主要包括谱折叠和谱平移这两种方法。谱折叠的思想是将低频部分通过镜像折叠的方式映射到高频部分。谱平移的思想则是将低频部分平移到高频部分。这两种方法由于可以很好的保留之前窄带信号, 是目前使用最广泛的方法。但是这两种方法的主要问题是估计出的宽带激励源  $e_b(t)$ 的高频谱波

与窄带信号中的真实谐波不一致,这会引入一些人为噪音,使得展宽后的语音自然度大大降低,但是与谱折叠相比,由于谱平移经过了两次滤波,这使得这种不自然的现象大幅减弱。

## (2) 频谱包络估计

相较于激励源而言,频谱包络包含有更多的信息,人耳对频谱包络也更敏感,因此频谱包络估计比激励源估计更为重要。频谱包络估计算法有很多,大致可以分为:码本映射、线性映射、统计映射等。

码本映射:码本映射的算法思想是首先设计两本码本,这两本码本分别对应于窄带语音信号与宽带语音信号,这两本码本之间一一对应并且同时训练。然后提取窄带语音的特征,从窄带语音码本中选择最相似的码字,通过窄带语音码本中的码字,可以直接从宽带码本中找寻相对应的码字,进而估计出宽带频谱包络  $a_b(t)$ 。码本映射算法频谱估算准确,计算量适中,可以保证实时性,是目前最为流行的频谱包络估计算法。但是该方法的主要缺点是没有考虑到语音信号之间时间的相关性,不能实现最优分类和搜索,并且需要较大的码本存储空间。

线性映射:如公式(4-49)所示,线性映射的算法思想是通过转移矩阵  $A_{n-b}$  将窄带语音信号的特征矢量  $x_n$  与宽带语音信号的特征矢量  $x_b$  相互联系起来。该方法计算简单,算法实现容易,但是由于估计模型太过简单,很难反映出窄带语音信号与宽带语音信号之间真实的联系。

$$x_b = A_{n-b}x_n \quad (4-49)$$

统计映射:统计映射的算法思想是通过分类的方式将窄带语音特征与宽带语音特征分为不同的类别,在包络估计时,计算每帧窄带语音特征矢量转移到各个类别时的概率,将每个类别的概率进行加权求和便可得到该帧宽带语音的特征矢量。该方法估计准确,鲁棒性好,目前受到广泛关注,但是由于模型较为复杂,实时性较差。

### 4.4.2 基于半波整流的语音带宽扩展算法

以上分析的传统语音带宽扩展算法都是基于源-滤波器模型的,这类算法虽然可以取得较好的效果,并在实际应用中得到了验证,但是这些算法需要大量的训练样本和先验知识。传统语音带宽扩展系统可通过各种语料库获得训练样本和先验知识,但是针对激光相干语音获取而言,目前还没有类似的语料库,因此,

不同于传统的语音带宽扩展模型, 本文采用半波整流来扩展语音带宽<sup>[133]</sup>。该算法通过对窄带语音信号中的最高倍频程(octave)进行半波整流以产生高频分量, 产生的高频分量通过增益系数进行适当的缩放加入到延迟输入信号中以实现语音带宽扩展, 这不仅仅可以有效扩展语音带宽, 并且不需要任何先验知识。

### (1) 语音带宽扩展算法流程

图 4.15 为本文采用的语音带宽扩展算法结构框图。在图 4.15 中, 滤波器 1(Filter 1)对窄带语音进行滤波操作以提取语音的最高倍频程。假设窄带语音的最高频率成分为  $f_h$ , 并假设信号的采样率  $f_s > 4f_h$ , 这可以通过改变采样率等预处理步骤完成。由于窄带语音中的最高倍频程在  $f_h/2$  与  $f_h$  之间, 因此滤波器的通带范围为  $f_h/2$  至  $f_h$ 。将窄带语音通过滤波器 1 从而获得最高倍频程。在图 4.15 中, 非线性装置(NLD)是将最高倍频程转换为高频频谱。由于半波整流器输出的频谱由原始输入频谱及其所有的谐波构成, 因此, 本文采用半波整流器作为非线性装置。在非线性输出的频谱中, 真正有用的为频率范围在  $f_h-2f_h$  之间的高频分量, 但是当有两个或者多个频率分量组成的信号通过非线性装置时, 输出频谱不仅仅包含高频分量, 而且还包含一些杂波分量, 这些杂波分量可能是由输入频谱之间的和差值构成的。这种现象称为互调失真(Intermodulation distortion, IMD)。为了消除由于互调失真而引入的杂波分量, 将非线性装置输出的信号通过通带范围为  $f_h$  至  $2f_h$  之间滤波器 2(Filter 2)以获得高频分量。最后, 将生成的高频分量添加到延时的输入信号中, 从而实现宽带扩展。其中延迟时间为高频分量生成所花费的时间。

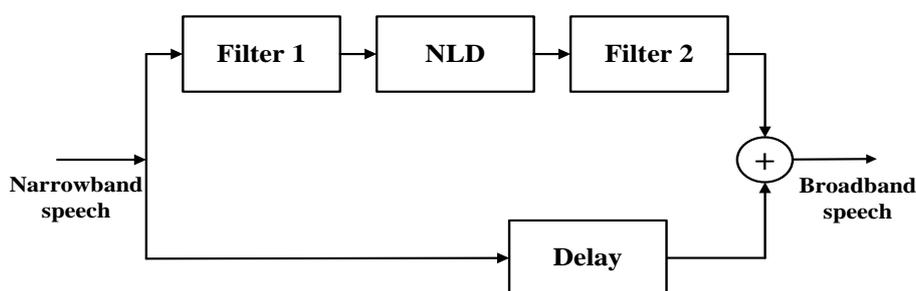


图 4.15 语音带宽扩展结构框图

Figure 4.15 The diagram of speech bandwidth extension.

## (2) 语音带宽扩展实验分析

采用实测相干侦测语音信号对语音带宽扩展算法有效性进行验证,其验证结果如图 4.16 所示。图 4.16.a 为原始相干侦测语音的频谱图。图 4.16.b 为扩展后的相干侦测语音的频谱图。比较图 4.16.a、图 4.16.b 可以发现,本文使用的语音带宽扩展算法可以有效实现窄带相干侦测语音的带宽扩展。并使用 MOS 给语音评分。原始相干侦测语音 MOS 得分为 3.1。带宽扩展后的语音 MOS 得分为 3.3。以上结果表明:本文采用的语音带宽扩展算法可以扩展窄带相干侦测语音的带宽,提升语音质量。

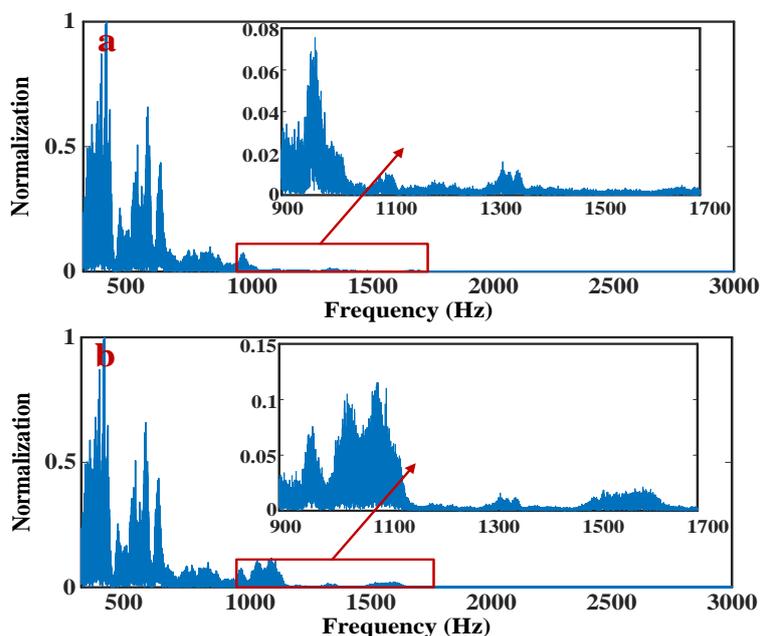


图 4.16 (a) 原始相干侦测语音频谱图。(b) 经过本节算法带宽扩展后语音频谱图。

Figure 4.16 (a) The spectrogram of the raw speech. (b) The spectrogram of the speech enhanced by the algorithm proposed in this paper.

## 4.5 本章小结

本章内容主要分三个方面:

(1) 为了重构语音信号,本文针对激光相干语音侦测系统需要同时保证信号处理的精度与实时性的特点,采用正交解调与反正切鉴相相结合的算法从载波信号中提取多普勒频移,进而重构语音信号。

(2) 为了降低相干侦测语音中的噪声,首先分析了相干侦测语音信号的噪声特点,根据噪声的表现形式可大致分为背景噪声和冲击噪声,对不同的噪声采

取不同的降噪方法。针对背景噪声，采用最优对数谱估计算法与相位补偿算法相结合的方式抑制背景噪声，提高语音质量。针对冲击噪声，本文通过峰度值比率法先确定冲击噪声的位置，然后利用线性估计算法估计出被冲击噪声干扰的原始语音段，从而消除冲击噪声，提高语音质量。

(3) 由于相干侦测语音带宽较窄，为了提高相干侦测语音的可懂度和自然性，本文采用基于半波整流的语音带宽扩展算法扩展窄带相干侦测语音的带宽，以提高语音的自然度和可懂度。

## 第 5 章 激光相干语音侦测实验研究

在前几个章节中已经完成了激光相干语音侦测系统光路结构与信号处理算法的设计和研究,并且分析讨论了被测物体特性对相干语音侦测的影响,这可以为语音侦测系统选择合适的测量物体提供指导。为了验证上述研究的准确性以及自研激光相干语音侦测系统的侦测性能,本章节进行了激光相干语音侦测实验。然后,在充分证明自研系统可有效侦测语音的基础上,又分别进行了双模监控实验和远距离说话人识别实验,通过这两个实验以表明自研系统在安防监控领域和生物识别领域应用的可行性和潜力。

### 5.1 语音侦测实验研究

为了验证自研激光相干语音侦测系统的语音侦测性能,在长春光机所对抗新楼走廊内进行了语音侦测实验,实验距离分别为 30m、70m 和 120m。实验装置示意图如图 5.1.a 所示,实验可分为检测部分与被测部分。如图 5.1.b 所示,检测部分由自研的激光相干语音侦测系统组成,其光路结构与关键器件详见第二章。使用高速采集卡对侦测系统中的光电探测器输出信号进行采集。如图 5.1.c 所示,本系统采用的数据采集卡为美国 NI 公司的 PXIe-5162,其最高采样率可达 5GS/s,采样分辨率为 10-bit。由于本系统的载波频率为 40MHz,为了保证采集信号不失真,采用的采样率为 312.5MHz。在计算机上使用 LabVIEW 作为开发环境,驱动采集卡将采集的数据实时上传到计算机中,同时采用正交解调与反正切鉴相结合的解调算法将采集数据实时重构为语音信号,并将这些语音信号保存以做进一步处理(降噪、带宽扩展等)。

本实验的被测部分由振动物体和声源两部分构成。为了验证自研激光相干语音侦测系统可获取不同类型人的语音,因此,使用录音设备在近距离、安静的环境下录制了 19 人(10 男 9 女,年龄分布在 27-65 之间)的语音作为本实验的声源。实验过程中,使用普通的便携式音箱依次播放事先录制的不同类型人的语音,播放的声波引起附近振动物体的振动,通过检测振动物体的振动便可获取语音信号。为了研究自研激光相干语音侦测系统的环境适用性,本实验中使用 3 种室内常见的物品:文件夹(塑料制)、饮料瓶(纸制)以及电脑显示屏作为振动物体

(如图 5.1.d 所示)。实验中，音箱播放语音音量为 50~60dB 之间，音箱与振动物体的距离约为 45 厘米。

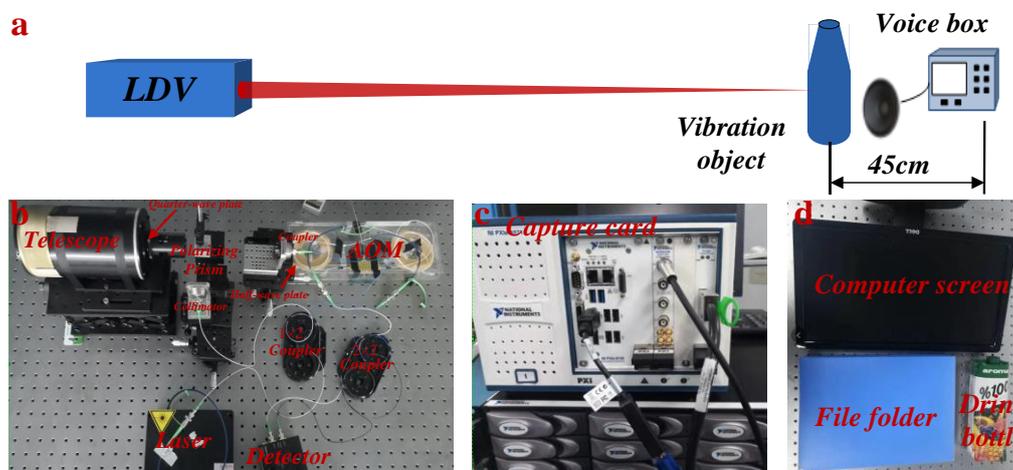


图 5.1 (a) 实验示意图。(b) 激光相干语音侦测系统。(c) 采集卡。(d) 被测目标

Figure 5.1 (a) Experimental diagram. (b) Prototype of the LDV system.(c)Capture card. (d) Target.

图 5.2 为侦测距离为 30m 时，自研激光相干语音侦测系统获取的一组实验数据，将其作为语音侦测实验结果的代表。图中 a、b、c 分别表示振动物体为文件夹、饮料瓶以及电脑显示屏时获取的语音信号、以及相对应的经过后续增强后的语音信号（降噪、带宽扩展等），d 代表纯净语音。图中数字 1 和 3 分别表示激光相干侦测语音及相对应的经过后续增强处理后语音信号的语谱图，数字 2 和 4 分别表示激光相干侦测语音及相对应的经过后续增强处理后语音信号的波形图。

观察图 5.2 可以发现，振动物体无论是文件夹、饮料瓶还是电脑显示屏，激光相干语音侦测系统均可获得可理解语音信号，这表明自研激光相干语音侦测系统具备获取远距离语音信号的能力，并且环境适用性好。然而相干侦测系统侦测的语音信号会不同程度的受到噪声的干扰，这严重损害了侦测语音的质量。但是使用本文提出的语音处理方法可明显抑制噪声，提高语音质量。此外，观察语谱图可以发现，从不同的振动物体中获取的语音信号的频谱成分互不相同，并且与纯净语音相比都存在高频频谱损失。这是由于不同种类的物体在声场的振动响应不同，以及所选材料在高频处振动响应差所引起的。在本次实验中，饮料瓶（纸制）相对于其他两种目标而言对于声场高频的振动响应最差，这一结果符合本文 3.1 节的分析结果。

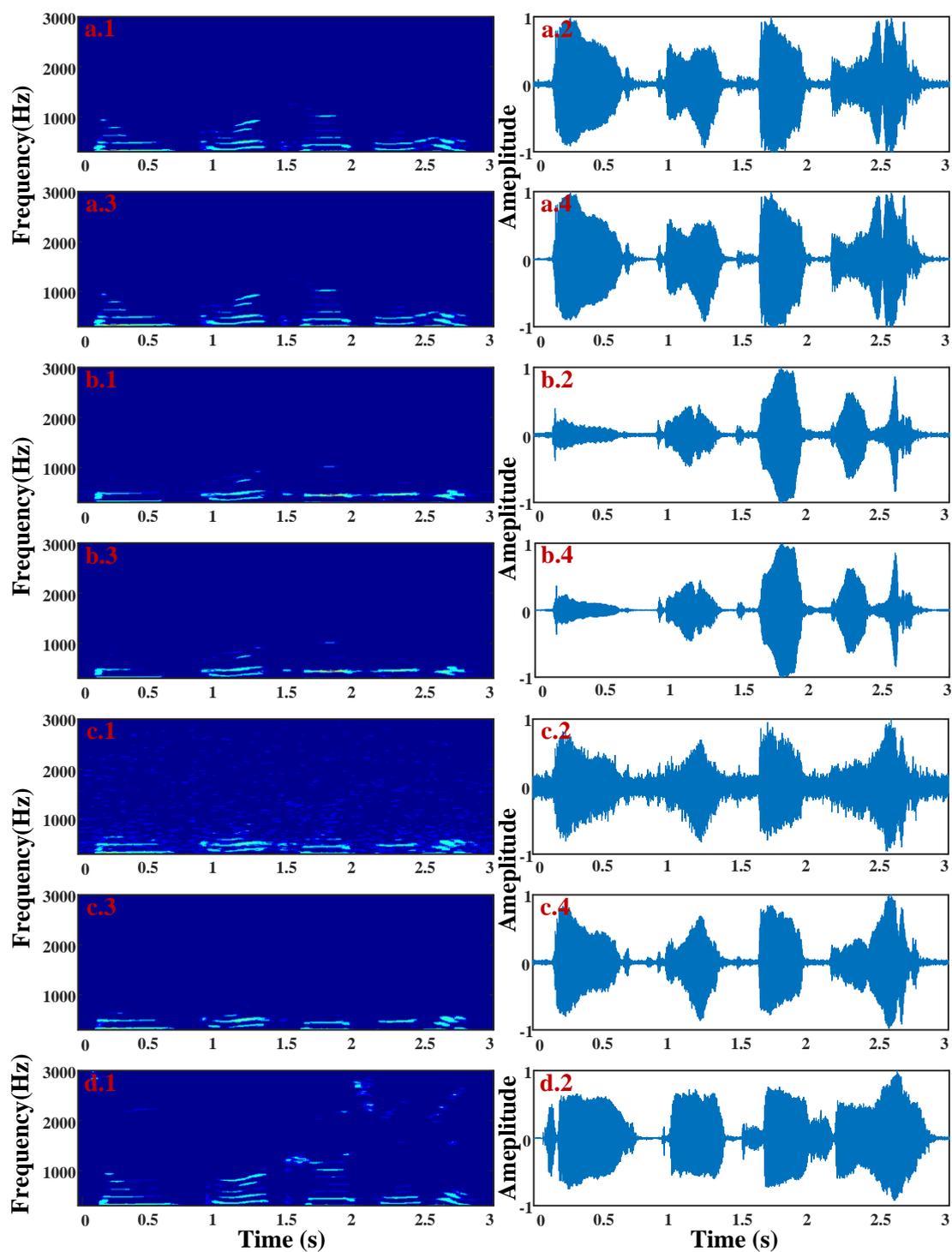


图 5.2 30m 侦测语音的语谱图和波形图

Figure 5.2 The spectrogram and waveform of the speech detected at 30m.

图 5.3 和图 5.4 为侦测距离分别为 70m 以及 120m 时，自研激光相干语音侦测系统获取的一组实验数据，将其作为语音侦测实验结果的代表。图 5.3 和图 5.4 中标记的字母和数字含义与图 5.2 中完全相同。

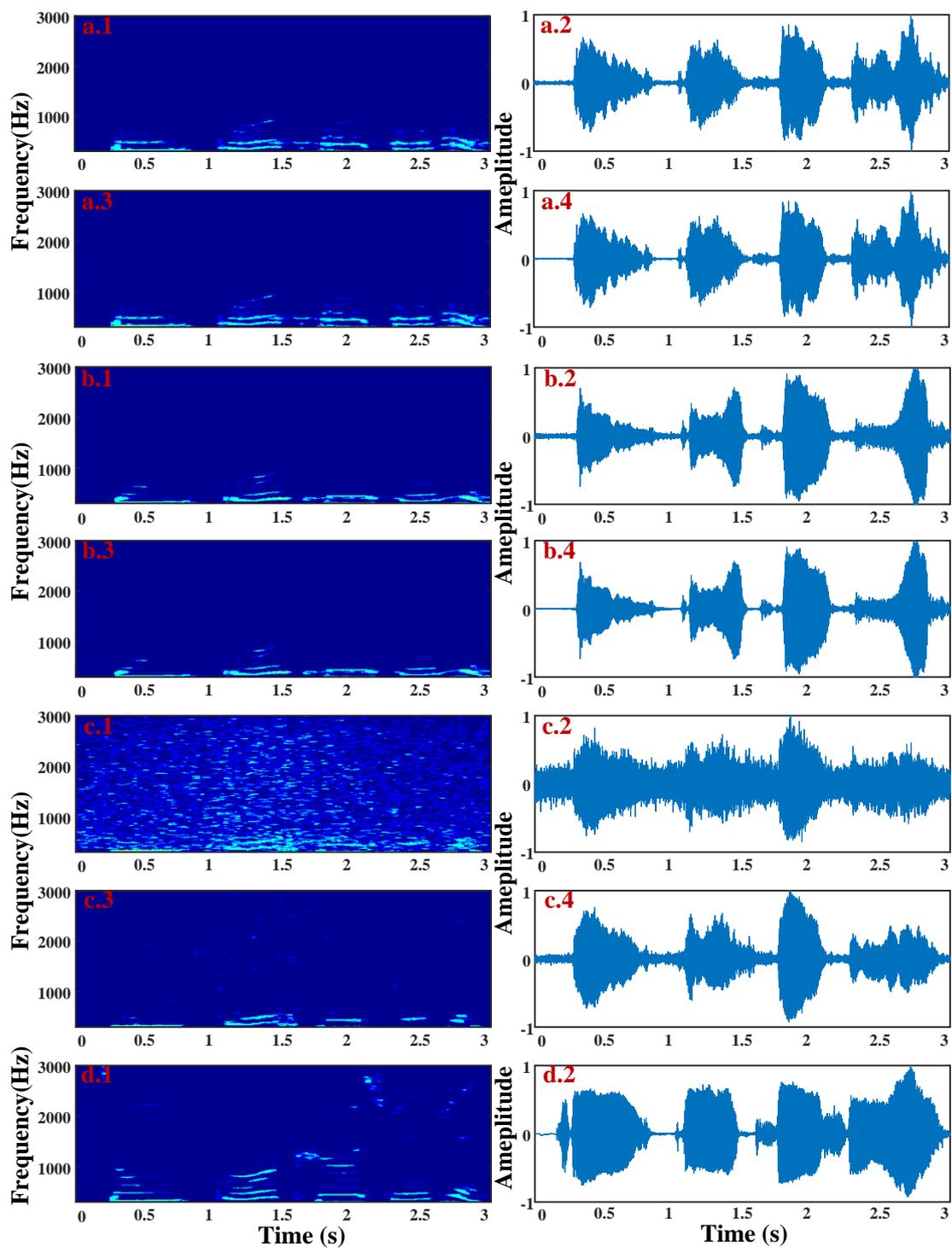


图 5.3 70m 侦测语音的语谱图和波形图

Figure 5.3 The spectrogram and waveform of the speech detected at 70m.

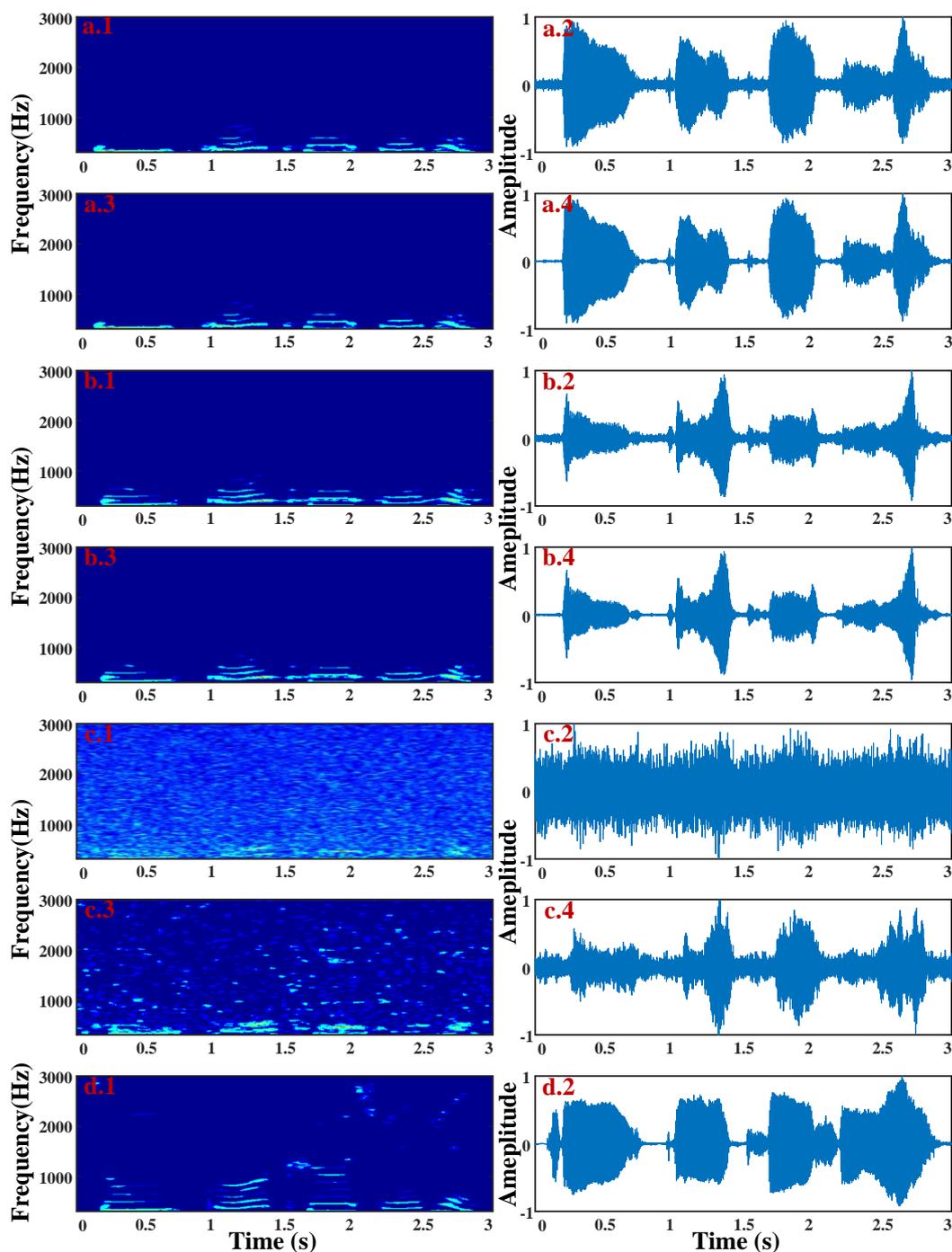


图 5.4 120m 侦测语音的语谱图和波形图

Figure 5.4 The spectrogram and waveform of the speech detected at 120m.

观察图 5.2、图 5.3 以及图 5.4 可以明显的发现，相干侦测语音的背景噪声随着侦测距离的增加而增加。这是因为回波光能量与侦测距离的四次方成反比（根据公式 2-85 可知），因此距离增加会导致回波光能量减弱，而回波光能量减弱则会使系统载噪比下降，最终影响相干语音侦测系统性能。但是，通过比较图 5.4

的侦测语音的语谱图和纯净语音的语谱图可以发现，其频谱有效成分相同（中低频率成分），这说明在距离 120m 处，系统仍可获取可理解语音信号。并且，使用本文提出的语音处理方法可明显抑制噪声，提高语音质量。另外，比较图 5.2、图 5.3 以及图 5.4 可以发现一个值得注意现象：不同振动物体获取的语音信号质量存在差异，这种差异随着侦测距离的增加而愈加明显。在本实验中，在侦测距离为 120m 时，振动物体为电脑显示屏时，获取的语音信号质量较差，振动物体为文件夹（塑料制）和饮料瓶（纸制）时，获取的语音信号质量较好。这一结果也说明选择合适的振动目标对于激光相干语音侦测的重要性。

表 5.1 MOS 平均得分

Table 5.1 The mean score of MOS

Vibration target	Distance (m)	Laser coherent detected speech (MOS)	Enhancement speech (MOS)
File Folder	30	3.3	3.6
	70	3.3	3.2
	120	3.0	3.3
Drink bottle	30	3.0	3.2
	70	3.1	3.1
	120	2.8	3.0
Computer screen	30	3.0	3.3
	70	2.5	2.9
	120	1.5	2.5

表 5.1 为自研激光相干语音侦测系统获取的语音信号以及相对应的增强信号的 MOS 得分的均值。观察表 5.1 可以得出与上述分析一致的结论，即：1) 自研的激光相干语音侦测系统可获取远距离（至少 120m）可理解语音信号。2) 系统直接获取的语音信号会受到噪声的干扰，使用本文采用的语音处理算法可明显抑制噪声，提高语音质量。3) 系统获取的语音信号质量随着侦测距离的增加而降低。4) 不同振动物体获取的语音信号质量不同。在本实验中，振动物体为电脑显示屏时，获取的语音信号质量明显逊色于振动物体为文件夹和饮料瓶时获取的

语音信号。

## 5.2 双模监控实验研究

监控系统作为安防系统中使用最多的系统之一，被广泛安装在各个场所，例如银行、地铁站、学校等。目前而言，几乎所有的监控系统都基于视频信息，而其他的可作为补充信息的信息形式（例如语音）尚未获得充分挖掘。造成这一结果的主要原因是因为声波在空气中传播时会快速衰减，这使得传统的语音传感器作用距离很短（米级），而监控系统的工作距离一般为十至百米级。由前文可知，自研的激光相干语音侦测系统可获取远距离可理解语音信号，与此同时相机可获取远距离视频信息。因此，在本实验中，结合自研激光相干语音侦测系统和监控相机搭建了双模监控系统来实现远距离视频和语音检测。由于本文已经对激光相干语音侦测过程进行了详细的阐述，因此不再对语音检测过程进行赘述。

### 5.2.1 视频检测

视频检测最主要的目的是在监控相机中获取的视频信号中识别出人体目标。此外，根据3.1节和上一个实验可知，相干侦测语音质量和振动目标的选择息息相关，因此视频检测的另一个重要目的为：识别人周围的物体，从而为相干语音侦测选择合适的振动目标提供指导。通过人工方式可实现视频检测，但是长时间的监控容易使人懈怠。随着计算机性能的飞速发展，深度学习在目标识别和检测方面得到了广泛的应用，其识别的准确率已经逼近人类，并在某些特定领域已经超越人类。YOLO(You Only Look Once)是深度学习成功应用在目标检测方面的一个典型代表<sup>[134~135]</sup>，该算法检测速度快，识别准确率高，因此在本实验中，使用该算法辅助监控人员来检测视频。

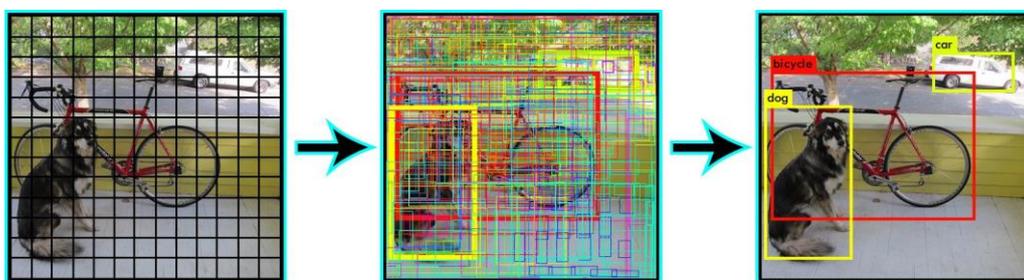


图 5.5 YOLO 算法识别过程

Figure 5.5 The identification process of YOLO algorithm.

YOLO 是一种先进的实时物体检测算法，该算法最初是由 Redmon J 等人在 2015 年提出的，在本实验中，该算法被用来辅助监控人员检测人体目标，以及人周围的物体。虽然 YOLO 算法属于卷积神经网络(Convolutional Neural Network, CNN)，但是与其他 R-CNN 系列的目标检测算法相比，YOLO 算法使用一个 CNN 网络便可实现图片的端对端(End-to-End)检测。因此 YOLO 算法的检测速度非常快，比 R-CNN 快 1000 倍以上，比 Fast R-CNN 快 100 倍。标准 YOLO 每秒可以检测 45 张图像，快速 YOLO 检测速度达到每秒 155 张图像<sup>[134~135]</sup>。如图 5.5 所示，YOLO 检测过程可概述为：首先将被检测图片划分为  $M \times M$  个网格。然后使用 CNN 网络预测每个网格的置信度(Confidence)，其计算公式可表示为：

$$Conf(Object) = Pr(Object) \times IOU \quad (5-1)$$

公式(5-1)中，若网格中包含物体，则  $Pr(Object)=1$ ，反之  $Pr(Object)=0$ 。IOU 为预测边界框和实际边界框的重叠率，其可表示为：

$$IOU = \frac{area(B_{Truth} \cap B_{pred})}{area(B_{Truth} \cup B_{pred})} \quad (5-2)$$

公式(5-2)中， $B_{Truth}$  和  $B_{Pred}$  分别代表实际边界框和预测边界框。

再将置信度乘以网格包含物体的预测类别，便可获得综合得分：

$$Score = Pr(Class_i | Object) Pr(Object) \times IOU \quad (5-3)$$

得到综合得分后，设置阈值，滤除低分项，并进行非极大值抑制处理和最大值操作，便可以确定被检查图片中，物体的类别和位置。

$$(i, pred) = \max_{i, pred} [Pr(Class_i | Object) Pr(Object) \times IOU] \quad (5-4)$$

公式(5-4)中， $i$  表示预测物体的类别， $pred$  表示预测物体的位置。

为了验证 YOLO 算法可有效辅助监控人员检测人体目标，以及人周围的物体，进行了一个针对性实验。本次实验的实验环境为：Intel Core i7-7700、NVIDIA GTX1070、32G 内存。实验的第一步是构建训练数据集。本文中的数据集是自制数据集，数据主要来源是互联网上可下载的图片。这些图片包含人以及现实生活中常见的并且可作为振动目标的物体，例如塑料袋、纸杯、矿泉水瓶等。为了增加训练数据量并提高模型的泛化能力，本文采用左右翻转来扩充数据集。然后基

于自制数据集,使用通过 ImageNet 1000-class 竞赛数据集预先训练好的 YOLOv2 模型进行训练。由于 YOLO 算法是一个应用广泛,并且较为成熟的算法,很多文献都对其进行了详细的描述<sup>[136-140]</sup>,因此,本文对 YOLO 算法的网络结构和训练过程不再进行赘述。最后,使用在实验室实景拍摄的视频以及互联网上的视频作为测试集,测试 YOLO 算法在视频检测中的性能。

表 5.2 准确率、召回率

Table 5.2 The precision rate, recall rate

	Precision	Recall
Human	99.89%	96.49%
Objects around people	89.17%	92.15%

YOLO 算法的准确率,召回率如表 5.2 所示。实验结果表明:本次实验使用的 YOLO 算法能够精确的检测出视频内人体目标和人周围的物体。并且其检测速度可满足实时检测,因此 YOLO 算法性能足以满足辅助监控人员检测视频的应用需求。

### 5.2.2 双模监控系统工作方式

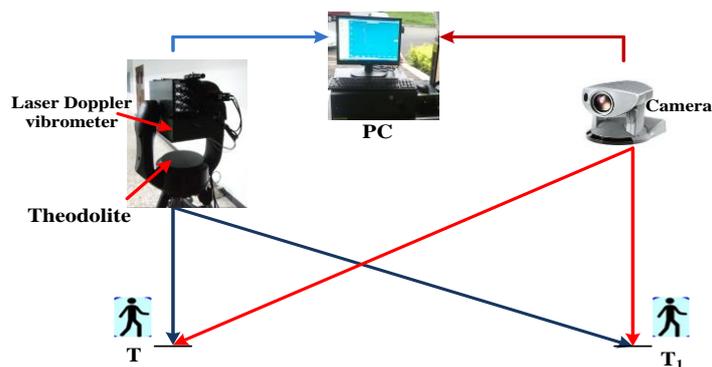


图 5.6 双模监控系统原理图

Figure 5.6 Schematic setup of the double model detection system.

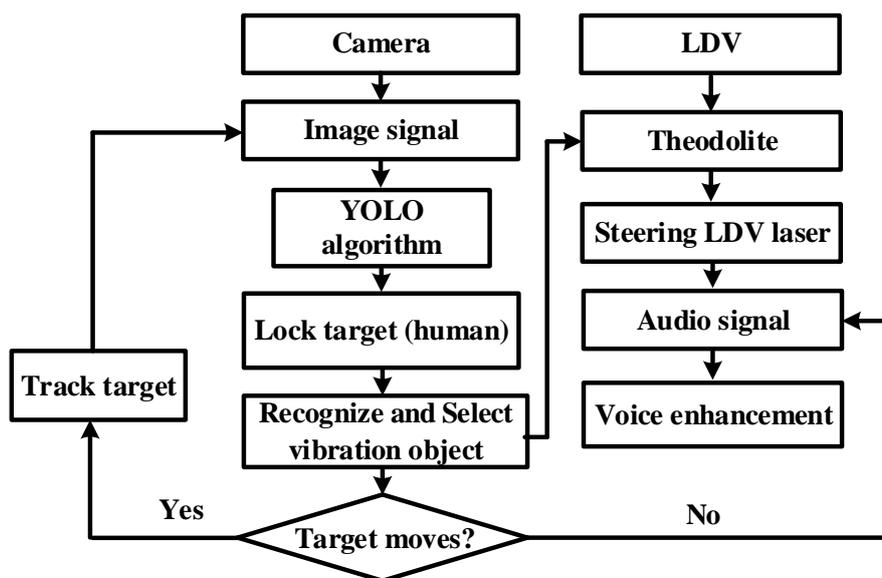


图 5.7 系统流程图

Figure 5.7 The system implementation flowchart.

如图 5.6 所示为双模监控系统的原理框图。该系统由安装在经纬仪上自研激光相干语音侦测系统、自带云台的监控相机和个人电脑组成。双模监控系统工作流程图如图 5.7 所示：首先使用监控相机获取监控区域的视频信息。然后，根据获取的视频信息，使用 YOLO 算法辅助监控人员检测监控区域是否出现被监控目标（人）。当监控区域出现被监控目标时，系统锁定被监控目标并聚焦监控相机以获得被监控目标的清晰图像。理想的图像应该包括被监控目标及其周围的物体。此时，利用 YOLO 算法识别被监控目标周围的物体，这可以为自研激光相干语音侦测系统选择合适振动目标提供指导。最后，通过经纬仪控制激光相干语音侦测系统的激光光束指向振动目标从而获取被监控目标的语音信号。如果被监控目标发生移动，带有云台的监控相机会跟踪目标，并帮助激光相干语音侦测系统重新选择合适的振动目标。该系统经纬仪的平移范围为 $-130^{\circ}$ 至 $130^{\circ}$ ，倾斜范围为 $-60^{\circ}$ 至 $70^{\circ}$ 。旋转分辨率为 $12.36''$ 。

## 5.2.3 双模监控实验结果及分析

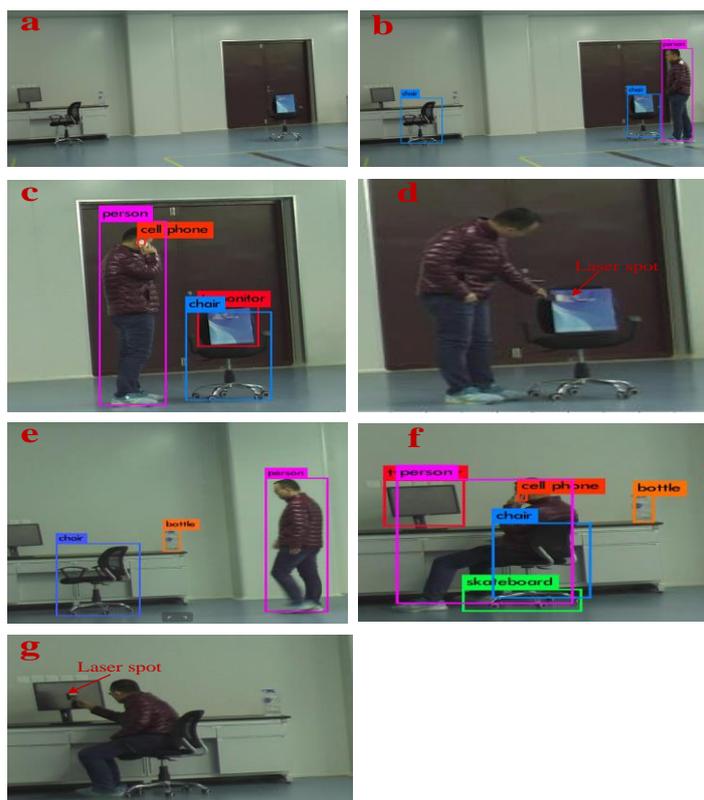


图 5.8 双模监控实验结果

Figure 5.8 Experiment results of double model detection.

为了验证双模监控的可行性，在实验室中进行了初步实验。实验结果如图 5.8 所示。如图 5.8.a 所示，由于实验室中所有的物体都自然的被放置，因此可以被认为是一个“非合作”的环境。当监控区域出现被监控目标时，系统自动识别并锁定被监控对象（图 5.8.b）。如图 5.8.c 所示，一旦锁定被监控目标，此时聚焦监控相机以获得被监控目标的清晰图像。根据获得的清晰图像，识别图像中的物体以便找到被监控对象附近的最佳振动目标。最后，如图 5.8.d 所示，控制激光相干语音侦测系统指向合适的振动目标（纸袋）以获取被监控目标的语音信号。当被监控目标开始移动时，系统会跟踪目标（图 5.8.e）。然后根据获得的图像，重新为激光相干语音侦测系统找寻合适的振动目标（图 5.8.f）。最后，如图 5.8.g 所示，将激光相干语音侦测系统重新指向合适的振动目标（电脑显示屏）以获取被监控目标的语音信号。

使用语谱图/波形图比较法和 MOS 评分法评估了双模监视实验中获取的语音信号质量。图 5.9 为振动目标为纸袋时，激光相干语音侦测系统获取的语音信

号的波形图和语谱图，及其对应的纯净语音信号的波形图与语谱图。图 5.10 为振动目标为电脑显示屏时，激光相干语音侦测系统获取的语音信号的波形图和语谱图，及其对应的纯净语音信号的波形图与语谱图。需要注意的是，本实验中获取的语音信号都经过后续的语音降噪、带宽扩展等处理，这部分过程可详见第 4 章。

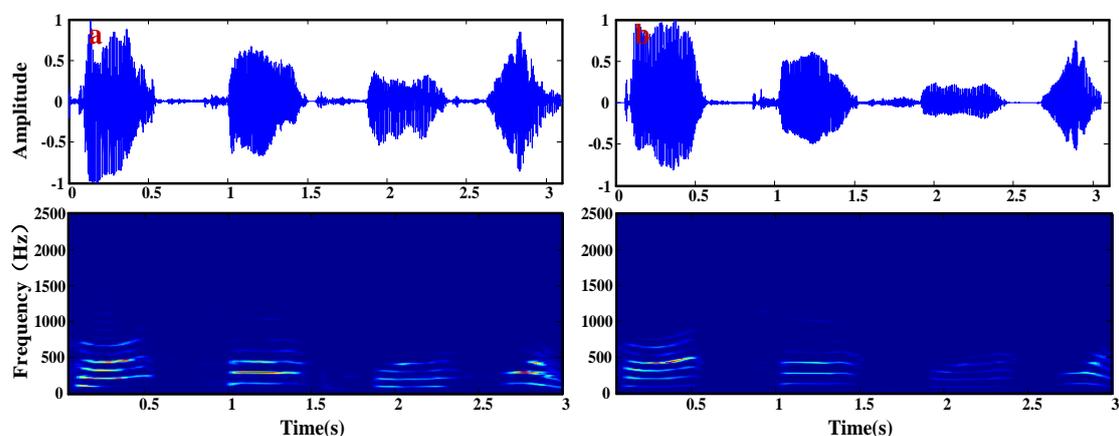


图 5.9 (a) 相干侦测语音。(b) 纯净语音

Figure 5.9 (a) Laser coherent detected speech signals. (b) Clean speech signals.

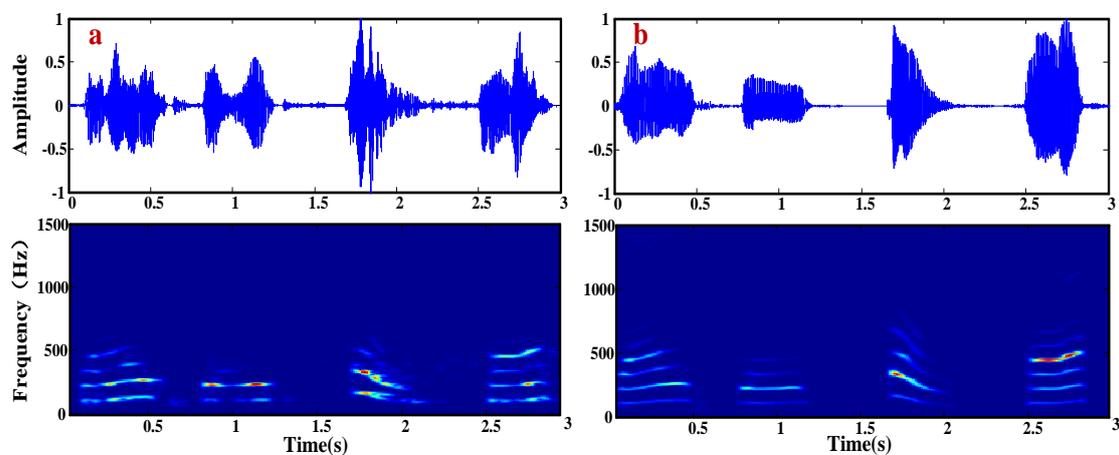


图 5.10 (a) 相干侦测语音。(b) 纯净语音

Figure 5.10 (a) Laser coherent detected speech signals. (b) Clean speech signals.

观察图 5.9 和图 5.10 可以发现，激光相干语音侦测系统获取的语音信号与纯净语音信号十分相似。并且，振动目标为纸袋时，相干侦测语音的 MOS 得分为 3.6。振动目标为电脑显示屏时，相干侦测语音的 MOS 得分为 3.1。语音侦测结果表明在双模监视实验中，可有效获取被监控对象的语音信号。

上述实验验证了双模监控的可行性，为今后双模监控系统的产品化研制奠定了基础。

### 5.3 远距离说话人识别实验研究

说话人识别技术是指从获取的语音信号中识别说话人的身份,是生物识别技术中的一种。由于该技术使用方便、非交互性等优点,已经成为生物识别领域的研究热点。说话人识别需要相对较高质量的语音,然而受限于传统语音采集系统的作用距离,使得说话人识别的应用场景主要集中在短距离条件下。由前文可知,自研激光相干语音侦测系统可获取远距离可理解语音信号。因此,在本实验中,基于获取的远距离可理解相干侦测语音,进行了远距离说话人识别实验。

#### 5.3.1 训练集与测试集建立

说话人识别可简单分为训练和识别两个阶段。在训练阶段时,根据训练集所提供的语音训练说话人识别模型。在识别阶段时,基于训练好的说话人识别模型,识别测试集中每个语音的说话人身份。由此可得,建立训练集与测试集是实现说话人识别的第一步。

在本实验中,使用录音设备在近距离、实验室环境下录制了 50 人的语音片段(包括 28 个男性和 22 个女性,年龄分布在 25-30 岁之间),每个语音片段的采样频率为 16Khz,采样时间为 1 分钟,单声道。由于合适的振动目标对于高频声音的振动响应较差,这会使得相干侦测语音的中高频频率成分丢失。因此,为了提高说话人识别的准确性,将录制的 50 人语音片段通过了一定频率范围的带通滤波器以作为训练集。

由于本实验的目的是验证远距离说话人识别的可行性,因此测试集由自研激光相干语音侦测系统获取的语音信号构成。测试集的语音获取实验过程与 5.1 的语音侦测实验基本相同。在本次实验中,声源由录制的 50 人的语音片段(未经过带通滤波器)构成。同时为了保证实验的普遍性,振动目标由塑料袋、矿泉水瓶和电脑显示屏构成。声源和振动目标之间的间隔约为 45cm,激光相干语音侦测系统与振动目标之间的距离约为 50m。值得注意的是,为了提高相干侦测语音质量,已对系统直接获取的语音进行语音降噪处理(详见第 4.3 节)。

图 5.11 为测试集语音的一组实验数据代表。图中 a、b、c 分别表示为振动物体为塑料袋、矿泉水瓶以及电脑显示屏时获取的语音信号,d 代表纯净语音(未通过带通滤波器)。图中数字 1 表示语音信号的波形图,数字 2 表示语音信号的语谱图。

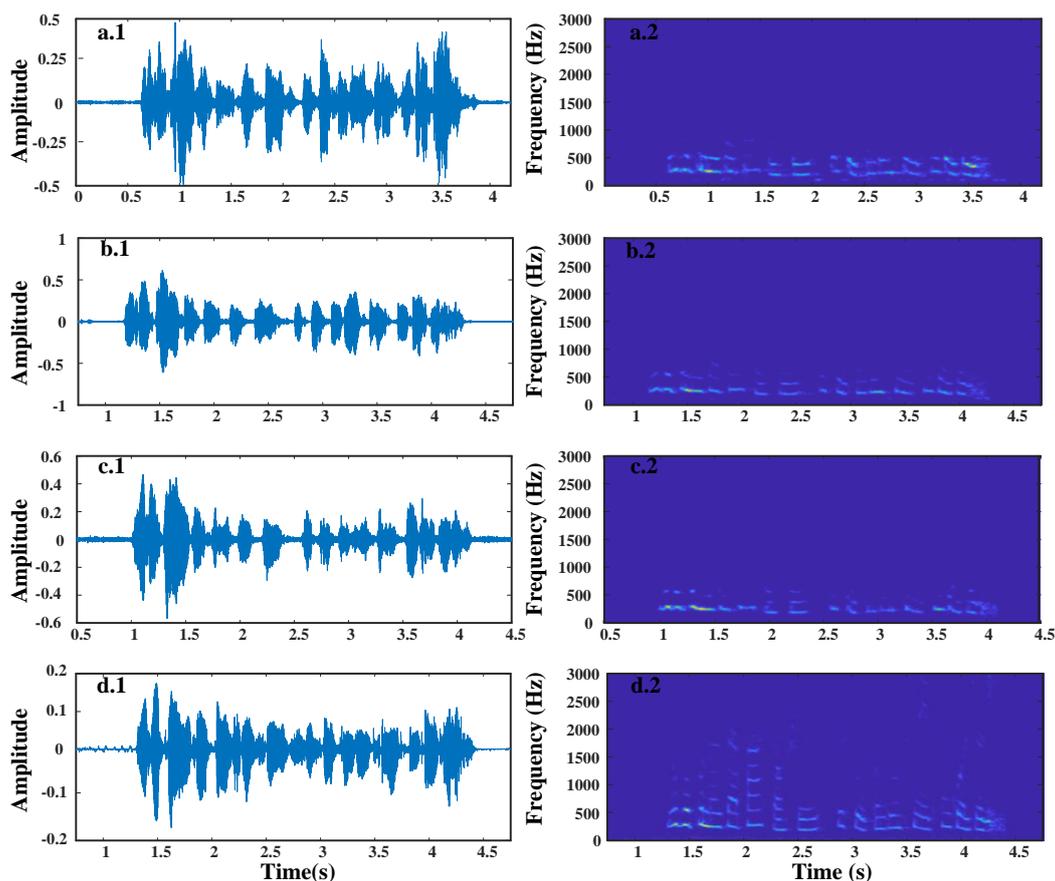


图 5.11 语音波形图与语谱图

Figure 5.11 Speech spectrograms and waveforms.

### 5.3.2 语音特征提取

语音特征提取指的是从语音信号中提取出可以反映说话人特点的基本特征，是说话人识别过程中的一个关键部分。由[15]与[141]可知，理想的特征应该满足以下特点：1) 相同说话人间的特征差异尽可能小，不同说话人间的特征差异尽可能大。2) 鲁棒性好。3) 计算量小，容易被获取，并且特征之间保持较好的独立性。4) 空间、时间的影响较小。5) 模仿难度大。Mel 频率倒谱系数(Mel Frequency Cepstrum Coefficient, MFCC)是目前普遍采用的说话人特征系数之一，该系数不仅包含很多与说话人相关的信息，并且还具有较强的鲁棒性。因此，在本实验中，选择 MFCC 作为语音特征系数。

MFCC 特征提取可大致分为 5 个阶段<sup>[106]</sup>：

1) 预处理。在该阶段中，主要对语音进行预加重、分帧以及加窗。经由这个阶段后，输入语音  $x(n)$  变为  $x_i(n)$ ，其中  $i$  表示分帧后的第  $i$  帧。

2) 短时傅立叶变换。在该阶段中, 使用傅立叶变换将每帧数据从时域转换为频域:

$$X(i, k) = FFT[x_i(m)] \quad (5-5)$$

3) 谱线能量计算。在该阶段中, 计算每帧频谱的能量:

$$E(i, k) = [X(i, k)]^2 \quad (5-6)$$

4) Mel 滤波器滤波能量计算。在该阶段中, 计算获得的频谱能量通过 Mel 滤波器组后的能量。该过程可以表示为将每帧频谱能量  $E(i, k)$  的第  $k$  条谱线与 Mel 滤波器组  $H_m(k)$  相乘, 最后再相加

$$S(i, m) = \sum_{k=0}^{N-1} E(i, k)H_m(k), \quad 0 \leq m < M \quad (5-7)$$

公式 (5-7) 中, Mel 滤波器组  $H_m(k)$  可表示为:

$$H_m(k) = \begin{cases} 0 & k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)} & f(m-1) \leq k \leq f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)} & f(m) \leq k \leq f(m+1) \\ 0 & k > f(m+1) \end{cases} \quad (5-8)$$

$$0 \leq m < M$$

公式 (5-8) 中,  $M$  表示为 Mel 滤波器组中滤波器的个数。  $f(m)$  为每个滤波器的中心频率, 可表示为:

$$f(m) = \left( \frac{N}{f_s} \right) F_{mel}^{-1} \left( F_{mel}(f_l) + m \frac{F_{mel}(f_h) - F_{mel}(f_l)}{M+1} \right) \quad (5-9)$$

公式 (5-9) 中,  $N$  表示每帧数据长度,  $f_s$  表示采样率,  $f_l$  表示最低频率,  $f_h$  表示最高频率。函数  $F_{mel}^{-1}$  可表示为

$$F_{mel}^{-1}(b) = 700(e^{b/1125} - 1) \quad (5-10)$$

5) DCT 倒谱计算。在该阶段中, 对获得的 Mel 滤波器滤波能量  $S(i, m)$  求取对数后, 对其进行离散余弦变换(Discrete Cosine Transform, DCT), 便可获得 MFCC 特征系数。

$$mfcc(i, n) = \sqrt{\frac{2}{M}} \sum_{m=0}^{M-1} \log[S(i, m)] \cos\left(\frac{\pi n(2m-1)}{2M}\right) \quad (5-11)$$

公式(5-11)中,  $n$  表示第  $n$  条谱线。

### 5.3.3 基于 GMM-UBM 的说话人识别模型

高斯混合模型(Gaussian Mixture Model, GMM)是目前实行与文本无关的说话人识别的主要办法<sup>[142-146]</sup>, 其识别过程可概述为: 首先依据训练集给每个说话人创建一个 GMM 模型, 然后计算测试语音在不同 GMM 模型下的概率, 最后通过比较测试语音在不同 GMM 模型中的概率来实现对说话人的判断。然而, 该模型对于语音的质量要求较高, 并且当训练集和测试集的语音长度不足时, 其识别性能不佳。通用背景模型(Universal Background Model, UBM)是一个通过对不同说话人的语音训练而得到的大型高阶 GMM 模型, 一般而言, 一个 UBM 模型可以在相当程度上反映一个大众水平的与说话人无关的特征分布。因此, 在 UBM 的基础上, 使用少量的训练集便可训练出可以充分反映每个说话人特性的 GMM 模型, 从而可以有效解决训练集和测试集不足的问题。本实验使用 GMM-UBM 模型来识别说话人。

GMM-UBM 模型识别过程可概述为:

- 1) 首先根据训练集中的语音特征系数构建 UBM 模型, 该模型可以等效于  $M$  个高斯概率密度函数的加权和:

$$\begin{aligned} P(x_t | \lambda_{UBM}) &= \sum_{i=1}^M \omega_i p_i(x_t) \\ &= \sum_{i=1}^M \omega_i \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(x - \mu_i)'(\Sigma_i)^{-1}(x - \mu_i)\right\} \end{aligned} \quad (5-12)$$

公式(5-12)中,  $p_i$  表示高斯概率密度函数,  $x_t$  表示第  $t$  帧的特征向量,  $D$  表示向量的维度,  $\mu_i$  表示样本的均值,  $\Sigma_i$  表示协方差矩阵,  $M$  为单个高斯分布的数量,  $\omega_i$  为权重因子, 并且满足:

$$\sum_{i=1}^M \omega_i = 1 \quad (5-13)$$

观察公式(5-12)可知, 一旦获得 UBM 模型的参数  $\lambda_{UBM} = \{\omega_i, \mu_i, \Sigma_i\}$  便可确定得 UBM 模型。最大似然估计(Maximum Likelihood, ML)是估计参数  $\lambda_{UBM}$  的一个有效的方法。假设训练集为  $X = \{x_t, t=1, 2, \dots, T\}$ , 则似然函数可表示为:

$$L(\lambda_{UBM}|X) = P(X|\lambda_{UBM}) = \prod_{t=1}^T P(x_t|\lambda_{UBM}) \quad (5-14)$$

通过求取似然函数  $L(\lambda_{UBM}|X)$  的最大值, 便可估计出参数  $\lambda_{UBM}$ 。

$$\hat{\lambda}_{UBM} = \arg \max L(\lambda_{UBM}|X) \quad (5-15)$$

求取似然函数  $L(\lambda_{UBM}|X)$  的最大值可采用最大期望估计算法 (Expectation Maximization Algorithm, EM)。EM 算法的实质是一种迭代算法, 在不断的迭代过程中, 根据 ML 优化参数  $\lambda_{UBM}$  的估计。在每次迭代中, 都必须满足:

$$L(\lambda_{UBM}^{k+1}|X) \geq L(\lambda_{UBM}^k|X) \quad (5-16)$$

公式 (5-16) 中,  $k$  表示迭代次数。实现 EM 算法时, 首先需要定义辅助函数  $Q$ :

$$Q(\lambda_{UBM}, \hat{\lambda}_{UBM}) = \sum_{k=1}^K \lg [L(\lambda_{UBM}|X)] P(k|X, \hat{\lambda}_{UBM}) = \sum_{k=1}^K \lg [P(X|\lambda_{UBM})] P(k|X, \hat{\lambda}_{UBM}) \quad (5-17)$$

EM 算法实现过程可分为求期望 (Expectation) 和最大化 (Maximization) 两个步骤。在求期望时, 计算辅助函数  $Q$ 。在最大化时, 则根据辅助函数  $Q$  求令  $\{\omega_i, \mu_i, \Sigma_i\}$  偏导数为 0 时的值。通过不断重复这两个步骤便可估计出参数  $\lambda_{UBM}$ 。

2) 然后在训练获得的 UBM 模型的基础上, 根据训练集中每一个说话人的特征参数, 估计出每个说话人的 GMM 模型。假设第  $l$  位说话人的特征参数为  $O = (O_1, O_2, \dots, O_T)$ , 则该说话人特征参数与 UBM 模型的匹配似然度为:

$$P(m|o_t) = \frac{\omega_m p_m(o_t|\mu_m, \Sigma_m)}{\sum_{j=1}^M \omega_j p_j(o_t|\mu_j, \Sigma_j)} \quad (5-18)$$

根据公式 (5-18), 重新计算第  $l$  位说话人的权值、均值和方差:

$$n_m = \sum_{t=1}^T P(m|o_t) \quad (5-19)$$

$$E_m(O) = \frac{1}{n_m} \sum_{t=1}^T P(m|o_t) o_t \quad (5-20)$$

$$E_m(O^2) = \frac{1}{n_m} \sum_{t=1}^T P(m|o_t) o_t^2 \quad (5-21)$$

利用更新的参数, 计算第  $l$  位说话人的 GMM 模型:

$$\hat{\omega}_m = [\alpha_m n_m / T + (1 - \alpha_m) \omega_m] / \gamma \quad (5-22)$$

$$\hat{\mu}_m = \alpha_m E_m(O) + (1 - \alpha_m) \mu_m \quad (5-23)$$

$$\hat{\Sigma}_m^2 = \alpha_m E_m(O^2) + (1 - \alpha_m)(\Sigma_m^2 + \mu_m^2) - \hat{\mu}_m^2 \quad (5-24)$$

公式 (5-22) 中,  $\gamma$  为调节系数, 以保证

$$\sum_{i=1}^M \hat{\omega}_i = 1 \quad (5-25)$$

公式 (5-22) 至公式 (5-24) 中,  $\alpha_m$  为自适应调节参数, 可表示为:

$$\alpha_m = \frac{n_m}{n_m + \tau} \quad (5-26)$$

公式 (5-26) 中,  $\tau$  为说话人的 GMM 模型与 UBM 模型之间的关联度, 通常可取 8~20 之间。

3) 在获取 UBM 模型和说话人的 GMM 模型的基础下, 计算测试语音在 UBM 模型和特定的说话人 GMM 的对数似然比, 根据获得似然比的大小识别测试语音的说话人身份。似然比越大, 说明越接近该说话人。

$$score = \log P(C | \lambda_{GMM}^l) - \log P(C | \lambda_{UBM}) \quad (5-27)$$

公式 (5-27) 中,  $C = (C_1, C_2, \dots, C_T)$  表示测试语音的特征参数,  $\lambda_{GMM}^l$  表示第  $l$  位说话人的 GMM 模型

#### 5.3.4 远距离说话人识别实验结果及分析

为了验证远距离说话人识别的可行性, 进行了初步实验。实验过程如流程图 5.12 所示。整个实验可分为训练和识别两个阶段。在训练阶段: 首先建立训练集语料库。然后在训练集语音中提取语音特征系数。最后依据训练集中的语音特征系数构建 UBM 模型, 并在 UBM 模型基础上, 生成训练集中每一个说话人的 GMM 模型, 以此用来识别语音的说话对象。在识别阶段: 首先将训练集语音作为声源, 使用自研激光相干语音侦测系统获取远距离声源语音, 将获得的相干侦测语音其作为测试集。然后提取测试集语音中的语音特征系数。最后将测试集中的语音特征系数放入已经训练好的 GMM-UBM 模型以识别远距离说话人。有关训练集以及测试集的建立、语音特征系数的提取以及 GMM-UBM 模型的构建已经在 5.3.1 至 5.3.3 节进行详述, 在此不再赘述。本实验中, GMM-UBM 模型的维度为 1024 维, 并且使用的语音特征系数 MFCC 经过了高阶差分、倒谱均值减 (Cepstral Mean Subtraction, CMS) 等参数优化, 其维度为 39 维<sup>[141~142]</sup>。

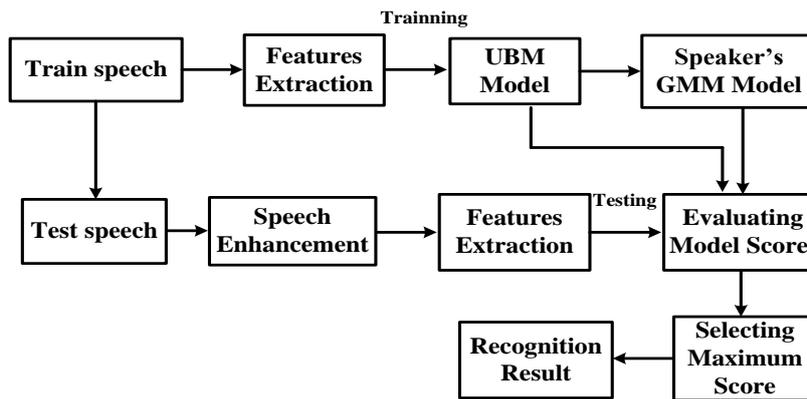


图 5.12 远距离说话人实验流程图

Figure 5.12 Flowchart for remote speaker recognition.

由于等错误率 (Equal Error Rate, EER) 评价指标具有简单明了的特点, 因此经常被用来评价识别结果。一般而言, EER 越低, 识别结果越好。检测误差权衡 (Detection Error Tradeoff, DET) 曲线也是评价说话人识别系统性能诸多评估方法中使用最为广泛的方法之一。该曲线的横纵坐标分别对应错误接受率和错误拒绝率在对数上的值。普遍而言, DET 曲线越靠近坐标轴, 其识别性能越好<sup>[142]</sup>。本实验中, 采用 EER 和 DET 曲线评估识别结果。表 5.3 为振动目标分别为塑料袋、矿泉水瓶和电脑显示屏时获得的相干侦测语音的 EER。图 5.13 为相对应的 DET 曲线图。另外, 在表 5.3 的最后一行中加入了用手机获取的语音的 EER 以提供比较, 手机与声源之间的间隔约为 15cm。

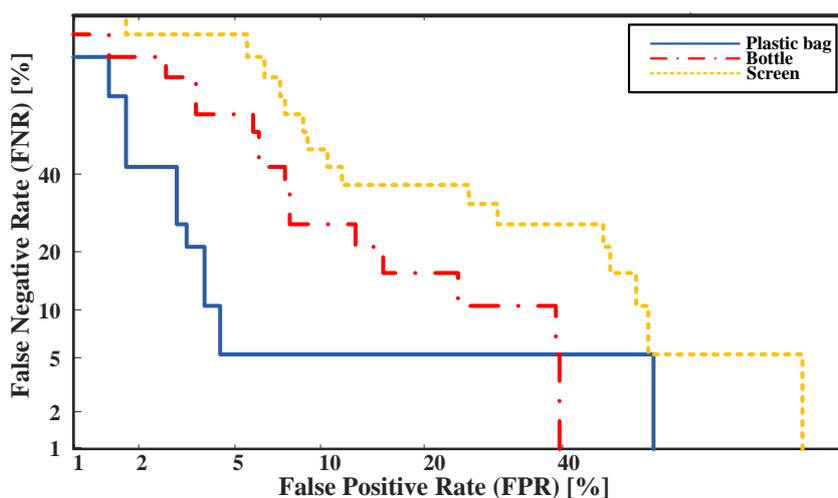


图 5.13 说话人识别 DET 曲线

Figure 5.13 Speaker Recognition DET Curves.

表 5.3 说话人识别等错误率

Table 5.3 Speaker Recognition EER

Vibration target	Speaker Recognition EER
Plastic bag	5.2630%
Mineral water bottle	15.7895%
Computer screen	29.8246%
Mobile phone	1.4620%

实验结果表明：振动目标分别为塑料袋、矿泉水瓶以及电脑显示屏时侦测获取语音的说话人识别 EER 分别为 5.2630%，15.7895%和 29.8246%，其均值 EER 为 16.9590%。手机录制语音的说话人识别 EER 为 1.4620%。观察上述实验结果可以发现：虽然相干侦测语音的说话人识别结果要差于手机录制语音的说话人识别结果，但是也显示了将激光相干语音侦测系统应用于远距离说话人识别领域的巨大潜力。此外，造成上述实验结果的一个主要原因是由于与手机麦克风振动膜相比，所选用的振动目标对于声音的振动响应较差。该结论可通过比较振动目标分别为塑料袋和电脑显示屏的 EER 以及比较相干侦测语音和其相对应的声源得到验证。因此，这也进一步表明选择一个合适的振动目标对于激光相干语音侦测的重要性。

#### 5.4 本章小结

本章内容主要分三个方面：

- (1)进行了激光相干语音侦测实验以获取远距离语音信号。实验结果表明：自研激光相干语音侦测系统可获取远距离（至少 120m）可理解语音信号。
- (2) 结合自研激光相干语音侦测系统和监控相机初步搭建了双模监控系统以实现远距离视频与语音双模检测。实验结果验证了双模监控的可行性，这为今后双模监控系统的产品化研制奠定了基础。
- (3) 在自研激光相干语音侦测系统获取的可理解语音信号的基础上，进行了远距离说话人识别实验。实验结果表明了远距离激光相干语音侦测技术在远距离说话人识别领域的巨大潜力。

## 第6章 总结与展望

### 6.1 论文工作总结

本文主要研究以激光相干多普勒测振技术为侦测手段的激光相干语音侦测系统的侦测原理、光路结构、测量目标选择以及信号处理等几方面的内容，并最终完成了激光相干语音侦测系统实验样机1台。主要研究内容如下：

1、分析比较了目前常见的激光语音侦测方法，提出了将激光相干多普勒测振技术应用于激光语音侦测系统中，简要回顾了激光相干多普勒测振技术的发展历程并介绍了国内外的研究现状，提出国内发展激光相干语音侦测技术需要解决的技术难点，为后续深入研究提供指导。

2、阐述了基于激光相干多普勒测振技术的语音信号侦测原理。详细推导了像差、激光器谱线展宽、探测器探测方式以及发射光串扰等光学因素对激光相干语音侦测系统的影响，并根据研究结果，确定了系统的光学元器件。参考了全光纤激光相干语音侦测系统的光路结构，并针对光纤环形器隔离度不够的问题，设计并搭建了部分光纤结构。

3、根据 STL 理论和 SSC 模型，建立了不同材质及形状的平板在声场的振动模型模拟物体在声场的振动幅度，并通过实验验证了分析的准确性。由于物体表面相对于光波而言是粗糙的，因此仿真了物体粗糙面对语音侦测的影响，并通过实验验证。当粗糙面相对于语音侦测系统存在除光轴方向之外的运动时，会给语音侦测过程引入散斑噪声。本文根据散斑特性讨论了散斑噪声对侦测语音的影响，并通过实验验证，此外，根据理论计算得出了散斑噪声的功率谱密度。

4、为了满足激光语音侦测系统对于语音重构算法精度与速度的要求，使用正交解调与反正切鉴相结合的解调算法重构语音信号。由于系统直接获得的语音信号会受到噪声的干扰，并且激光相干侦测语音的噪声与传统语音噪声特性不同，因此，详细分析了激光相干侦测语音中的噪声特性，针对不同类型的噪声，采用不同的降噪算法。例如针对背景噪声，采用 OM-LSA 与相位补偿相结合的方法降噪。针对冲击噪声，则先用峰度值比率确定噪声位置，再使用线性估计去除噪声。针对激光相干侦测语音普遍存在语音带宽窄的特点，采用基于半波整流的语音带宽扩展算法扩展相干侦测语音带宽，增强语音质量。

5、通过室内 30m、70m、120m 激光语音侦测实验验证了自研激光相干语音侦测系统的性能，随后，又分别进行了双模监控实验和远距离说话人识别实验，这两个实验结果分别表明了自研语音侦测系统在安防监控领域和生物识别领域应用的可行性和潜力。

本文的主要创新性内容包括：

1、在全光纤光路结构的基础上，提出了部分光纤结构，在保留光纤结构诸多优点的基础上，部分光纤结构解决了发射光串扰的问题，并通过实验得到了验证。

2、研究了散斑噪声对激光相干语音侦测的影响，并基于动态散斑特性，理论推导出了散斑噪声的功率谱密度与被测物体运动速度之间的关系。

3、研究了激光相干侦测语音的噪声特性，并根据不同类型的噪声，采用相对应的语音降噪方法消除噪声。并且，将半波整流技术应用于激光相干侦测语音的带宽扩展中以进一步提高激光相干侦测语音的质量。

4、将激光相干语音侦测技术分别与视频检测技术和说话人识别技术相结合，实现了远距离监控目标的图像和语音监控以及远距离说话人识别，这两项研究结果表明了激光相干语音侦测技术应用于安防监控领域和生物识别领域的可行性与潜力。

## 6.2 研究展望

激光相干语音侦测产品是光机电高度集成的精密设备，其实现需要涉及光学、电子、信号处理以及机械等诸多领域。本课题由于实验条件、时间、成本等限制，对自研激光相干语音侦测系统的产品化仍需要进行改进和进一步的研究，以下是对今后工作的设想及计划内容：

1、目前，信号处理算法都是在上位机上完成的，为了今后实现激光相干语音侦测系统的小型化和操作的便捷性、实时性，在后续工作中，需要设计基于 DSP 或 FPGA 的嵌入式信号处理电路，并尝试使用性能更强的语音降噪及增强算法，例如基于深度学习的语音降噪算法和带宽扩展算法，提高激光相干侦测语音的语音质量。

2、论文中所分析的物体粗糙面及散斑对于相干语音侦测的影响时都是基于被测物体位于光束焦点处的假设下进行的，未考虑离焦的情况，然而，在实际侦

测中,人工调焦很难做到完全聚焦,并且随着侦测距离的增加,人工调焦的难度也增加,因此,在未来可以分析聚焦效果对于激光相干语音侦测系统的影响,以及设计一套自动聚焦系统。

3、由于实际场地限制,未能进行测试距离大于 120m 的语音侦测实验,未能确定自研远距离语音侦测设备的最远侦测距离,从而不能针对最远侦测距离这一上限进行后续的研究工作。

4、在前期实验过程中,受到环境等因素的限制,只进行了室内语音侦测实验,对于室外环境下,能否成功侦测远距离语音信号以及侦测语音的质量如何仍是未知数,因此,后续工作中应该尽快开展室外语音侦测实验。此外,本文中,未考虑被测目标存在大平动速度时的情况(例如,车在行驶时,侦测车内人的语音)以及存在空气扰动的情况,因此,在后续研究中需要进行这方面的研究。



## 参考文献

- [1] Zieger C, Brutti A, Svaizer P. Acoustic Based Surveillance System for Intrusion Detection.[C]. Sixth IEEE International Conference on Advanced Video & Signal Based Surveillance. IEEE Computer Society, 2009:314-319.
- [2] Clavel C, Ehrette T. Fear-type emotion recognition and abnormal events detection for an audio-based surveillance system[C]. RISK ANALYSIS 2008. IEEE, 2008:1306-1309.
- [3] Radhakrishnan R, Divakaran A, Smaragdis A. Audio analysis for surveillance applications[C]. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005. IEEE, 2005:158-161.
- [4] Forrest M. Surreptitious interception of conversations with lasers[J]. OPTICS NEWS, 1985, 11(11):6-12.
- [5] 曾道红. 激光窃听[J]. 科学与文化, 2005(4):19-20.
- [6] 久期. 英国将采用先进监听监视跟踪技术[J]. 计算机与网络, 1997(4):71-73.
- [7] Qu Y, Wang T, Zhu Z. Remote audio/video acquisition for human signature detection[C]. IEEE Computer Society Conference on Computer Vision & Pattern Recognition Workshops. IEEE, 2009:66-71.
- [8] Li W, Liu M, Zhu Z, et al. LDV Remote Voice Acquisition and Enhancement[C]. International Conference on Pattern Recognition. IEEE, 2006: 262-265.
- [9] Wang A T, Zhu Z, Divakaran A. Long Range Audio and Audio-Visual Event Detection Using a Laser Doppler Vibrometer[J]. Evolutionary and Bio-Inspired Computation: Theory and Applications IV, 2010(7704):1-6.
- [10] Qu Y, Wang T, Zhu Z. Remote audio/video acquisition for human signature detection[C]. Cvpr'09 Biometrics, 2009: 66-71.
- [11] Zhu Z, Li W, Wolberg G. Integrating LDV Audio and IR Video for Remote Multimodal Surveillance[C]. IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2005:10-18.
- [12] Li R, Wang T, Zhu Z, et al. Vibration Characteristics of Various Surfaces Using an LDV for Long-Range Voice Acquisition[J]. IEEE Sensors Journal, 2011, 11(6):1415-1422.
- [13] Rui Li, Nicholas Madam Poulos, Zhigang Zhu, Liangping Xie. Performance comparison of an all-fiber-based laser Doppler vibrometer for remote acoustical

- signal detection using short and long coherence length lasers[J]. Appl. Opt.2012,51(21):5011-5018.
- [14] Avargel Y, Cohen I. Speech measurements using a laser Doppler vibrometer sensor: Application to speech enhancement[C]. Hands-free Speech Communication & Microphone Arrays. IEEE, 2011:109-114.
- [15] Deng Y. Long Range Standoff Speaker Identification Using Laser Doppler Vibrometer[C]. IEEE International Conference on Biometrics Theory. IEEE, 2016:1-6.
- [16] 罗海俊,朱晓. 激光窃听技术的研究[J].激光与光电子学进展.2003, 40(12):53-56.
- [17] 康宏向. 激光侦听信号中大气噪声与环境噪声的降噪研究[D]. 国防科学技术大学, 2003.
- [18] 田玉敏, 陈海清, 曾文锋. 基于改进算法的语音增强在激光侦测系统中的应用[J]. 光电子·激光, 2007, 18(12):1489-1491.
- [19] 张超凡. 激光窃听技术的改进与实现[J]. 激光与红外, 2008, 38(2):145-148.
- [20] Shang J , He Y , Liu D , et al. Laser Doppler vibrometer for real-time speech-signal acquirement[J]. Chinese Optics Letters, 2009, 7(8):732-733.
- [21] 张勤, 张建勇, 曾华林, et al. 基于 PGC 解调的激光多普勒语音检测系统[J]. 红外与激光工程, 2011, 40(6):1115-1118.
- [22] 黄贞, 吴林富. 便携式激光远程语音监听装置设计[J]. 激光与光电子学进展, 2012, 49(12):98-101.
- [23] 刘艳阳, 母一宁, 李野, et al. 基于四象限探测器的激光监听系统[J]. 压电与声光, 2014(4):675-678.
- [24] 母一宁, 刘艳阳, 李平, et al. 大气环境下的四象限激光监听方法与试验[J]. 半导体光电, 2014, 35(3): 523-526.
- [25] 张宇泽, 李丽艳, 曾华林, et al. 目标物材料特性对激光语音获取的影响研究[J]. 应用光学, 2014, 35(5):922-926.
- [26] 心愉. 基于散斑图像的激光侦听技术[J]. 中国安防, 2015(11):89-95.
- [27] 张展. 耳语音信号处理研究及其在激光侦听中的应用[D]. 南京航空航天大学, 2017.
- [28] 张文璐. 激光侦测系统中的语音增强技术研究[D]. 解放军信息工程大学,

2017.

- [29] 丁宁. 基于四象限测探测器的声音探测方法研究[D].合肥工业大学,2017.
- [30] 汪啸. 远距离激光振动检测技术的研究[D].电子科技大学,2018.
- [31] 白涛,吴谨,李明磊,万磊,李丹阳.DRNN 在激光多普勒测振仪测声系统中的应用[J].激光技术,2019,43(01):109-114.
- [32] Y. Yeh, H.Z.Cummins. Localized fluid flow measurements with an He-Ne laser spectrometer [J], Appl, Phys.lett, 1964, 176(4):176-178.
- [33] Lal, Amit. Multi-beam laser Doppler vibrometer for landmine detection[J]. Optical Engineering, 2006, 45(10),104302:1-10.
- [34] Aranchuk V , Lal A K , Hess C F , et al. Speckle noise in a continuously scanning multibeam laser Doppler vibrometer for acoustic landmine detection[J]. Proceedings of SPIE - The International Society for Optical Engineering, 2006, 6217(621716):1-10.
- [35] Abe T , Sugimoto T . Distinguishing Buried Objects in Extremely Shallow Underground by Frequency Response Using Scanning Laser Doppler Vibrometer[J]. Japanese Journal of Applied Physics, 2010, 49(7),07HC15:1-4.
- [36] Osten W. Optical inspection of microsystems[M]. Crc Press, 2006.
- [37] Cristalli C , Paone N , R.M. Rodríguez. Mechanical fault detection of electric motors by laser vibrometer and accelerometer measurements[J]. Mechanical Systems & Signal Processing, 2006, 20(6):1350-1361.
- [38] Cosoli G , Casacanditella L , Tomasini E P , et al. The non-contact measure of the heart rate variability by laser Doppler vibrometry: Comparison with electrocardiography[J]. Measurement Science and Technology, 2016, 27(6):1-8.
- [39] Melis M D, Grigioni M, Morbiducci U, et al. Optical monitoring of the heart beat[C]. Biomedicine, 2005: 181-190.
- [40] Kranjec J , Begu S , GerAk G , et al. Non-contact heart rate and heart rate variability measurements: A review[J]. Biomedical Signal Processing and Control, 2014, 13:102-112.
- [41] Siegman, A. E . The antenna properties of optical heterodyne receivers[J]. Proceedings of the IEEE, 1966, 54(10):1350-1356.
- [42] Fink D . Coherent Detection Signal-to-Noise.[J]. Appl Opt, 1975, 14(3):689-690.
- [43] Cohen S C . Heterodyne detection: Phase front alignment, beam spot size, and detector uniformity[J]. Applied Optics, 1975, 14(8):1953-1959.

- [44] Dong H , Yang R , Yang C , et al. Analysis of heterodyne detection in the frequency domain[J]. Applied Optics, 2015, 54(33):9883-9888.
- [45] Holmes J F , Rask B J . Optimum optical local-oscillator power levels for coherent detection with photodiodes.[J]. Applied Optics, 1995, 34(6):927-933.
- [46] Jiang L A , Luu J X . Heterodyne detection with a weak local oscillator[J]. Applied Optics, 2008, 47(10):1486-1503.
- [47] Salem M , Rolland J P . Heterodyne efficiency of a detection system for partially coherent beams[J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2010, 27(5):1111-1119.
- [48] Chengqiang L , Tingfeng W , Heyong Z , et al. The performance of heterodyne detection system for partially coherent beams in turbulent atmosphere[J]. Optics Communications, 2015, 356:620-627.
- [49] Painchaud Y , Poulin M , Morin M , et al. Performance of balanced detection in a coherent receiver[J]. Optics Express, 2009, 17(5):3659-3672.
- [50] Gallion P , Debarge G . Quantum phase noise and field correlation in single frequency semiconductor laser systems[J]. Quantum Electronics IEEE Journal of, 1984, 20(4):343-349.
- [51] 李成强, 王挺峰, 张合勇, et al. 激光光源线宽对外差探测性能的影响[J]. 物理学报, 2016, 65(8):174-181.
- [52] Chambers D . Modeling of heterodyne efficiency for coherent laser radar in the presence of aberrations.[J]. Optics Express, 1997, 1(3):60-67.
- [53] Hu Q , Rodrigo P J , Iversen T F Q , et al. Investigation of spherical aberration effects on coherent lidar performance[J]. Optics Express, 2013, 21(22):25670-25676.
- [54] 郭力仁,胡以华,李政,徐世龙.声光移频器对微多普勒效应探测的影响研究[J]. 光学学报,2015,35(02):208-215.
- [55] Rahimi S , Li Z , Dollevoet R . Measuring with laser Doppler vibrometer on moving frame (LDVMF)[J]. 2014:274-286.
- [56] O'Malley P , Woods T , Vignola J , et al. Mapping LDV Noise As A Function Of Standoff Distance, Beam Focus, And Surface Roughness[J]. AIP Conference Proceedings, 2010:291-297.
- [57] Rothberg S . Numerical simulation of speckle noise in laser vibrometry.[J]. Appl Opt, 2006, 45(19):4523-4533.
- [58] Pfister T , Fischer A , Czarske, Jürgen. Cramér–Rao lower bound of laser

- Doppler measurements at moving rough surfaces[J]. *Measurement Science and Technology*, 2011, 22(5):055301-055310.
- [59] Rothberg S J , Halliwell N A . Laser vibrometry on solid surfaces: The effects of laser speckle[J]. *Journal of Laser Applications*, 1994, 6(1):38-54.
- [60] Peter Martin. Uncertainty due to speckle noise in laser vibrometry[D]. Loughborough University, 2010.
- [61] Vass J , R Šmíd, Randall R B , et al. Avoidance of speckle noise in laser vibrometry by the use of kurtosis ratio: Application to mechanical fault diagnostics[J]. *Mechanical Systems & Signal Processing*, 2008, 22(3):647-671.
- [62] Ridley K D , Jakeman E . FM demodulation in the presence of multiplicative and additive noise[J]. *Inverse Problems*, 1999, 15(4):989-1002.
- [63] Ridley K D , Jakeman E . Signal-to-noise analysis of FM demodulation in the presence of multiplicative and additive noise[J]. *Signal Processing*, 2000, 80(9):1895-1907.
- [64] Hill C A , Harris M , Ridley K D , et al. Lidar frequency modulation vibrometry in the presence of speckle[J]. *Applied Optics*, 2003, 42(6):1091-1100.
- [65] Jiang L A , Albota M A , Haupt R W , et al. Laser vibrometry from a moving ground vehicle[J]. *Applied Optics*, 2011, 50(15):2263-2273.
- [66] Dräbenstedt A. Quantification of displacement and velocity noise in vibrometer measurements on transversely moving or rotating surfaces[C]. *Optical Measurement Systems for Industrial Inspection V*. 2007,661632(6616):1-10.
- [67] Maru K , Watanabe K . Non-mechanical scanning laser Doppler velocimetry with sensitivity to direction of transverse velocity component using optical serrodyne frequency shifting[J]. *Optics Communications*, 2014, 319:80-84.
- [68] Maru K , Hata T . Directional discrimination for fiber-optic non-mechanical scanning laser Doppler velocimeter using single transmission path[J]. *Optik - International Journal for Light and Electron Optics*, 2014, 125(20):6312-6314.
- [69] Scalise L , Paone N . Laser Doppler vibrometry based on self-mixing effect[J]. *Optics and Lasers in Engineering*, 2002, 38(3-4):173-184.
- [70] Zhang W , Gao W , Huang L , et al. Optical heterodyne micro-vibration measurement based on all-fiber acousto-optic frequency shifter[J]. *Optics Express*, 2015, 23(13):17576-17583.
- [71] Cheng C H , Lin L C , Lin F Y . Self-mixing dual-frequency laser Doppler velocimeter[J]. *Optics Express*, 2014, 22(3):3600-3610.

- [72] 张澍, 李玉, 卢广锋. 基于双频环形激光器的激光多普勒测振系统研究[J]. 光学学报, 2016, 36(3):111-116.
- [73] Dräbenstedt, Alexander. Diversity Combining in Laser Doppler Vibrometry for Improved Signal Reliability[C]. Intl Conference on Vibration Measurements by Laser & Noncontact Techniques. American Institute of Physics, 2014:263-273.
- [74] JianhuaShang, Shuguang Zhao, YanHe, WeibiaoChen, and NingJia. Experimental study on minimum resolvable velocity for heterodyne laser Doppler vibrometry[J]. Chin. Opt. Lett. 2011,9(8):35-37.
- [75] 刘丹, 郑宾, 郭华玲, 刘辉, 刘乃强. 基于外差干涉的微振动测量技术研究[J]. 应用光学, 2014, 35(05):858-861.
- [76] 李艳辉, 吴振森, 武颖丽. 基于激光多普勒法测量微小阻尼振动[J]. 中国激光, 2010, 37(01):231-235.
- [77] 刘琼. 激光振动目标识别技术研究[D]. 西安电子科技大学, 2008.
- [78] 李丽艳. 用于多普勒干涉测振的光学系统研究[D]. 长春理工大学, 2010.
- [79] 秦文强. 基于 FPGA 的激光多普勒测振的研究[D]. 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2018.
- [80] 周健. 用于车载自主导航激光多普勒测速仪的初步研究[D]. 国防科学技术大学, 2010.
- [81] 高宗梅. 激光多普勒测振技术用于西瓜成熟度检测的研究[D]. 浙江大学, 2016.
- [82] 黄浩钧. 激光多普勒干涉的双通道振动测量系统[D]. 浙江大学, 2017.
- [83] 张文. 梨质地的激光多普勒测振无损检测方法[D]. 浙江大学, 2016.
- [84] 李秀明. 用于弹丸速度测量的激光多普勒测速技术研究[D]. 天津大学, 2014.
- [85] 赵树杰, 赵建勋. 信号检测与估计理论[M]. 电子工业出版社, 2013.
- [86] N. Roddier. Atmospheric wavefront simulation using Zernike polynomials[J]. Optical Engineering. 1990, 29: 1174-1180.
- [87] Yang Q, Tan L, Ma J. Influence of aberrations of receiver antenna on heterodyne efficiency in free-space optical communications based on heterodyne detection[C]. International ICST Conference on Communications and NETWORKING. 2012:776-779.
- [88] 杨彦玲, 李彦超, 高龙, et al. 相干激光雷达平衡外差探测方法的数值仿真

- [J]. 红外与激光工程, 2011,40(10):1918-1922.
- [89] Karlsson C J , Fredrik Å. A. Olsson, Letalick D , et al. All-Fiber Multifunction Continuous-Wave Coherent Laser Radar at 1.55  $\mu\text{m}$  for Range, Speed, Vibration, and Wind Measurements[J]. Applied Optics, 2000, 39(21):3716-3726.
- [90] L. M. Brechovskich and O. A. Godin. Acoustics of Layered Media[M]. New York: Springer, 1990.
- [91] Tadeu A , J.M.P. ANTÓNIO. ACOUSTIC INSULATION OF SINGLE PANEL WALLS PROVIDED BY ANALYTICAL EXPRESSIONS VERSUS THE MASS LAW[J]. Journal of Sound and Vibration, 2002, 257(3):457-475.
- [92] António Tadeu, Julieta António, Mateus D . Sound insulation provided by single and double panel walls—a comparison of analytical solutions versus experimental results[J]. Applied Acoustics, 2004, 65(1):15-29.
- [93] Callister J R , George A R , Freeman G E . AN EMPIRICAL SCHEME TO PREDICT THE SOUND TRANSMISSION LOSS OF SINGLE-THICKNESS PANELS[J]. Journal of Sound and Vibration, 1999, 222(1):145-151.
- [94] Osipov A , Mees P , Vermeir G . Low-frequency airborne sound transmission through single partitions in buildings[J]. Applied Acoustics, 1997, 52(3-4):273-288.
- [95] Goodman J W . Introduction to Fourier Optics[M]. McGraw-Hill, 1968.
- [96] L Tsang, J A Kong, K Hau Ding et al. Scattering of Electromagnetic Waves, Numerical Simulations[M]. New York: Wiley Interscience, 2001.
- [97] 李甜田. 目标粗糙表面对回波外差探测的影响[D].西安电子科技大学,2013.
- [98] 党文佳. 粗糙目标对光外差探测信号的退相干效应[D].西安电子科技大学,2015.
- [99] 郑小兵. 随机表面光散射的数值分析[D]. 中科院安徽光机所,1998.
- [100] 孙鹤. 随机粗糙面的光散射特性的理论研究[D].黑龙江大学,2011.
- [101] Rothberg S J , Baker J R , Halliwell N A . Laser vibrometry: Pseudo-vibrations[J]. Journal of Sound & Vibration, 1989, 135(3):516-522.
- [102] T. Iwai, N. Takai, and T. Asakura. Dynamic statistical properties of laser speckle produced by a moving diffuse object under illumination of a Gaussian beam [J]. J. Opt. Soc. Am. 1982,4(72):460-467.
- [103] T. Asakura and N. Takai “Dynamic laser speckles and their application to velocity measurements of the diffuse object [J]. Appl. Phys.1981,25:179 -194.
- [104] N. Takai, T. Iwai, and T. Asakura. An effect of curvature of rotating diffuse

objects on the dynamics of speckle produced in the diffraction field [J]. Appl. Phys B. 1981, 26:185-192.

[105] 宋知用. MATLAB 在语音信号分析与合成中的应用[M]. 北京航空航天大学出版社, 2013.

[106] MolinaGarcía, Mariano, Alonso J I . Evolution of mean opinion Score due to handover events in IEEE 802.11 WLAN networks[C]. IEEE Radio & Wireless Symposium. IEEE, 2009,63(5):207-210.

[107] Gu L , Harris J G E , Shrivastav R , et al. Disordered Speech Evaluation Using Objective Quality Measures.[J]. 2005(1): 321-324.

[108] Rix A W, Beerends J G, Hollier M P, et al. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs[J]. IEEE Proceedings, 2001(2):749-752.

[109] Wenli Zhang. Algorithm Research and Implementation of PESQ Speech Quality Assessment System [D]. Dalian University of Technology, 2007.

[110] Bauer M , Ritter F , Siegmund G . High-precision laser vibrometers based on digital Doppler signal processing[J]. Proc Spie, 2002, 4827(4827):50-61.

[111] So S , Paliwal K K . Suppressing the influence of additive noise on the Kalman gain for low residual noise speech enhancement[J]. Speech Communication, 2011, 53(3):355-378.

[112] Yan Q , Vaseghi S , Zavarehei E , et al. Kalman tracking of linear predictor and harmonic noise models for noisy speech enhancement[J]. Computer Speech & Language, 2008, 22(1):69-83.

[113] Ephraim Y , Van Trees H L . A signal subspace approach for speech enhancement[C]. Icassp. IEEE Computer Society, 1993.251-256.

[114] Hu Y , Loizou P C . A generalized subspace approach for enhancing speech corrupted by colored noise[J]. IEEE Transactions on Speech and Audio Processing, 2003, 11(4):334-341.

[115] Johnson M T , Yuan X , Ren Y . Speech signal enhancement through adaptive wavelet thresholding[J]. Speech Communication, 2007, 49(2):123-133.

[116] Ghanbari Y , Karami-Mollaei M R . A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets[J]. Speech Communication, 2006, 48(8):927-940.

[117] Azirani A A , Jeannes R L B , Faucon G . Optimizing speech enhancement by

- exploiting masking properties of the human ear[C]. International Conference on Acoustics. IEEE, 1995,1: 800-803.
- [118] Gunawan T S , Ambikairajah E , Epps J . Perceptual speech enhancement exploiting temporal masking properties of human auditory system[J]. Speech Communication, 2010, 52(5):381-393.
- [119] Boll S . Suppression of acoustic noise in speech using spectral subtraction[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1979, 27(2):113-120.
- [120] Lim J S , Oppenheim A V . Enhancement and bandwidth compression of noisy speech[J]. Proceedings of the IEEE, 1979, 67(12):1586-1604.
- [121] EPHRAIM. Speech enhancement using a minimum mean square error short-time spectral amplitude estimator[J]. IEEE Trans, Acoust. Speech Signal Process. 1984, 32(6):1109-1121.
- [122] Rainer Martin. Spectral subtraction based on minimum statistics[C]. Proc.Euro.Sigal Processing Conf. 1994:1182-1185.
- [123] Martin R . Noise power spectral density estimation based on optimal smoothing and minimum statistics[J]. IEEE Trans Speech & Audio Processing, 2001, 9(5):504-512.
- [124] Israel Cohen. Optimal Speech Enhancement under Signal Presence Uncertainty Using Log-Spectral Amplitude Estimator [J]. IEEE SIGNAL PROCESSING LETTERS, 2002, 9(4):113-116.
- [125] Cohen I . Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recursive Averaging[J]. IEEE Transactions on Speech and Audio Processing, 2003, 11(5):466-475.
- [126] Wojcicki K , Milacic M , Stark A , et al. Exploiting Conjugate Symmetry of the Short-Time Fourier Spectrum for Speech Enhancement[J]. IEEE Signal Processing Letters, 2008, 15:461-464.
- [127] Markel J D , Jr A H G . Linear Prediction of Speech[J]. Communication & Cybernetics. 1976,53(3):266-267.
- [128] Jax P , Vary P . On artificial bandwidth extension of telephone speech[J]. Signal Processing, 2003, 83(8):1707-1719.
- [129] Iser B , Schmidt G , Minker W . Bandwidth extension of speech signals[J]. Springer Berlin, 2008, 13(2):45-52.
- [130] Peng R , Zheng C , Li X . Bandwidth extension for speech acquired by laser Doppler vibrometer with an auxiliary microphone[C]. International Conference on

Information. IEEE, 2015:1-4.

- [131] 隋经纬. 人工语音带宽扩展算法研究[D].大连理工大学,2014.
- [132] 王媛媛. 语音带宽扩展算法研究[D].大连理工大学,2009.
- [133] Somesh Ganesh. Audio Bandwidth Extension[C].
- [134] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-time Object Detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:779-788.
- [135] Redmon J , Farhadi A . YOLOv3: An Incremental Improvement[J]. arXiv, 2018,1804.02767:1-6.
- [136] 罗聪. 基于人脸检测 YOLO 算法的专用型卷积神经网络推理加速器的研究与设计[D].华南理工大学,2018.
- [137] 王超,付子昂.基于 YOLO v2 模型的交通标识检测算法[J].计算机应用,2018,38(S2):276-278.
- [138] Yang W , Zhang J , Wang H , et al. A vehicle real-time detection algorithm based on YOLOv2 framework[C]. Real-time Image & Video Processing. 2018(10670):1-8.
- [139] Jo K U , Im J H , Kim J , et al. A real-time multi-class multi-object tracker using YOLOv2[J]. ICSIPA, 2017:507-511.
- [140] Shafiee M J , Chywl B , Li F , et al. Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video[J]. Journal of Computational Vision and Imaging Systems, 2017.
- [141] 倪雷. 基于改进 Mel 特征提取算法的说话人识别研究[D].重庆邮电大学,2017.
- [142] 张黛. 基于 GMM-UBM 的稳健的说话人识别技术研究[D].华中科技大学,2016.
- [143] Reynolds D A, Quatieri T F, Dunn R B. Speaker verification using adapted Gaussian mixture models[J]. Digital signal processing, 2000, 10(1-3): 19-41.
- [144] Chao Y H , Tsai W H , Wang H M . Improving GMM-UBM speaker verification using discriminative feedback adaptation[J]. Computer Speech and Language, 2009, 23(3):376-388.
- [145] Chowdhury M F R , Selouani S A , O'Shaughnessy D . Distributed automatic text-independent speaker identification using GMM-UBM speaker models[C].

Conference on Electrical & Computer Engineering. IEEE, 2009: 372-375.

[146] Rakhmanenko I A , Meshcheryakov R V . Identification Features Analysis in Speech Data Using Gmm-Ubm Speaker Verification System[J]. Tr Spiiran, 2017, 3(52):32–50.



## 作者攻读学位期间发表的学术论文与研究成果

### 一作及通讯发表（或正式接受）的学术论文：

1. Tao LV, Xiyu HAN, Shisong WU, et. al. The effect of speckles noise on the Laser Doppler Vibrometry for remote speech detection[J]. *Optics Communications*, 2019,440. (第一作者, 已发表)
2. Tao LV, He-yong Zhang, Chun-hui YAN, et. al. Double mode surveillance system based on remote audio/video signal acquisition[J]. *Applied Acoustic*, 2017,129. (第一作者, 已发表)
3. Tao LV, GUO Jin, ZHANG He-Yong, et. al. Acquirement and enhancement of remote speech signals[J]. *OPTOELECTRONICS LETTERS*, 2017,13(4). (第一作者, 已发表)
4. 吕韬, 张合勇, 郭劲, et. al. 远距离语音的激光相干获取及增强[J]. *光学精密工程*, 2017, 25(03). (第一作者, 已发表)
5. Xiyu Han, Tao LV, Shisong Wu, et. al. A remote human activity detection system based on partial-fiber LDV and PZT camera[J]. *Optics and Laser Technology*, 2018,111. (通讯作者, 已发表)
6. Shuping Peng, Tao LV, Xiyu Han, et. al. Remote Speaker Recognition based on the Enhanced LDV-captured Speech[J]. *Applied Acoustic*, 2018,143. (通讯作者, 已发表)

### 合作发表（或正式接受）的学术论文：

7. Xiyu Han, Tao LV, XIANGYU SONG, et. al. An Adaptive Two-Scale Image Fusion of Visible and Infrared Images[J]. *IEEE Access*, 2019,07. (第二作者, 已发表)
8. He-yong Zhang, Tao LV, Chun-hui YAN, et. al. The novel role of arctangent phase algorithm and voice enhancement techniques in laser hearing[J]. *Applied Acoustic*, 2017,126. (第二作者, 已发表)
9. Shi-song Wu, Tao LV, Xi-yu Han, et. al. Remote audio signals detection using a partial-fiber laser Doppler vibrometer[J]. *Applied Acoustic*, 2017,130. (第二作者, 已发表)
10. Shi-song Wu, Tao LV, Yuan-Yang Li, et. al. Multi-beam interference competition

- in heterodyne detection[J]. *OPTOELECTRONICS LETTERS* (第二作者, 已录用)
11. Shisong WU, Yuanyang LI, Tao LV, et. al. All-fiber pulsed laser Doppler vibrometer development based on time-domain chopping techniques[J]. *CHINESE OPTICS LETTERS*, 2019,17(5). (第三作者, 已录用)
  12. Chun-Hui Yan, Ting-Feng Wang, Yuan-Yang Li, et. al. Investigating the displacement resolution of laser heterodyne detection system[J]. *Optics Communications*, 2019,435. (第四作者, 已发表)
  13. 晏春回, 王挺峰, 张合勇, et. al. 近距离激光外差探测光学极限位移分辨率[J]. *物理学报*, 2017,66(23).(第四作者, 已发表)
  14. 晏春回, 王挺峰, 张合勇, et. al. 激光语音检测系统反正切补偿算法[J]. *红外与激光工程*, 2017,46(09). (第四作者, 已发表)
  15. Chun-Hui Yan, Ting-Feng Wang, Yuan-Yang Li, et. al. An improved arctangent algorithm based on phase-locked loop for heterodyne detection system[J]. *Chinese Physics B*, 2019,28(03). (第四作者, 已发表)
  16. 吴世松, 张合勇, 王挺峰, et. al. 基于远程脉冲激光外差探测的匹配滤波过程研究[J]. *光学学报*, 2018,38(01). (第六作者, 已发表)

### 外审学术论文:

1. Two-scale Image Fusion Algorithm of Visible and Infrared Images. *Applied Optics* (Under review)
2. Speckles Intensity fluctuation and their impact on the LDV measurement. *JOURNAL OF RUSSIAN LASER RESEARCH* (Minor revision)
3. Influence of laser linewidth on the performance of the Laser Doppler Vibrometry in the presence of aberration. *Optics Communications* (Under review)

### 申请或已获得的专利:

1. 一种信号处理电路, 已获得, 专利号: 201611248378.9
2. 一种目标身份识别系统, 已受理, 申请号: 2016112470365
3. 一种探测系统, 已受理, 申请号: 2016112483666

## 指导教师及作者简介

### 指导教师简介

郭劲(1964-): 男, 吉林长春人, 工学博士, 研究员, 博士生导师, 目前为中国科学院长春光机所光电对抗部主任, 参加省部级课题及国防科研重点工程项目多项, 发表学术论文 50 余篇。获中科院科技进步一等奖一项, 获科工局科技进步二等奖一项, 获国务院政府特殊津贴一次。主要研究方向: 光电设备总体技术、激光应用技术。

### 作者简介

吕韬(1991-): 男, 浙江绍兴人, 汉族, 中国共产党预备党员, 2010~2014 年就读于浙江工业大学光信息科学与技术专业, 获学士学位; 2014 年 9 月就读于中科院长春光机所光学工程专业, 攻读硕士学位; 2016 年 9 月转为博士研究生。主要研究方向: 激光相干多普勒振动测量。



## 致谢

韶华易逝, 光阴冉冉, 转眼之间已经临近博士毕业。细细回想读博的这段时光里, 期间虽然遇到了不少令人望而却步的挫折与坎坷, 他们的艰难让我颤抖, 想要仓皇而逃, 可是, 在我导师、家人、朋友、同学及实验室同仁们的指导、帮助及鼓励下, 我才能鼓起勇气, 直面困难, 克服困难。如果没有他们, 我不可能会取得今天的成绩, 因此, 在论文的最后我想要向那些在我博士期间帮助我、支持我、鼓励我的人表示最真诚的谢意。

在众多需要感谢的人中, 首先要感谢的是我的导师郭劲研究员。郭老师渊博的学识、严谨的治学态度、以及敏锐的学术洞察力让我折服, 在我攻读博士学位的这段时光中, 郭老师自始至终都给予我悉心的指导, 本论文正是在郭老师的教导下, 才得以完成。除此之外, 郭老师身上勤奋、正直、严谨、乐观的特性无声的影响着我, 教会我如何做人做事。能够成为郭老师的学生是我莫大的荣幸, 在此, 我向郭老师表达我最真诚的谢意!

特别感谢张合勇副研究员和李远样博士, 我研究工作的相当于一部分内容都是在他们的指导下完成的, 毫不夸张的说, 没有他们的悉心指导与帮助, 我不可能顺利完成博士论文。能遇到亦师亦友的张合勇副研究员和李远样博士, 是我的荣幸。特别感谢冯群工程师在生活上给予我的照顾, 让我独处异乡还感受到家乡的温暖!

特别感谢李成强博士在光学相干探测原理方面提供极大的指导、特别感谢王灿近博士在电路调试及程序编写方面给予的极大的帮助。

衷心感激本部门王挺峰研究员、王锐副研究员、刘立生博士、王卫兵硕士, 正是在他们悉心的指导下, 我才能顺利完成本论文的研究工作。

感谢实验室同学杨国庆博士、王雅男硕士、李充硕士、晏春回、于湛、张翔、张绍军、吴世松、俞航航、宋纪坤、张绍、张玉莹、陈鸿凯, 他们无论是在学业还是生活上都给予了我极大的帮助, 非常感谢他们的陪伴, 正因为有他们的存在, 给稍显枯燥的科研生活增添了许多欢乐。

感谢韩玺钰, 她的出现让我的科研生涯中增添了些许色彩, 正因为有她默默的支持, 我才能顺利的完成博士论文。

最后要感谢我的父母和外婆，他们在背后默默的支持与付出，我才能心无旁骛的完成博士学业。感谢我的长辈们给我的支持和帮助。

要感谢的人实在太多，但是由于篇幅所限，只能就此搁笔。在今后的学习生活中，我唯有加倍努力，才能报答你们的恩情！