

WILEY

衍射、傅里叶光学 及成像

Diffraction, Fourier Optics and Imaging

[美] 奥坎 K. 埃尔索伊 (Okan K. Ersoy) 著

• 蒋晓瑜 闫兴鹏 等译



机械工业出版社
CHINA MACHINE PRESS

衍射、傅里叶光学及成像

Diffraction, Fourier Optics and Imaging

[美] 奥坎 K. 埃尔索伊 (Okan K. Ersoy) 著
蒋晓瑜 闫兴鹏 等译



机械工业出版社

衍射现象是光学及电磁学中的主要现象,成像是现代科学技术领域中应用广泛,而傅里叶分析方法作为一种有效工具,在衍射及成像中均具有广泛应用。本书将衍射、成像与傅里叶分析结合起来,系统而全面地介绍了光学中的衍射现象、成像系统和成像方法,以及傅里叶分析方法及其在多种衍射和成像领域中的应用。本书针对有一定理论基础和实践经验的读者,他们对物理光学有一定了解,并熟练掌握了傅里叶分析方法。本书可用作高等院校有关专业的研究生教材,也可作为相关领域研究人员或工程师的参考书。

Copyright © 2007 John Wiley & Sons, Inc. All rights reserved

All Rights Reserved. This translation published under license. Authorized translation from the English language edition, entitled Diffraction, Fourier Optics and Imaging, ISBN 9780471238164, by OKAN K. ERSOY, Published by John Wiley & Sons. No part of this book may be reproduced in any form without the written permission of the original copyrights holder.

本书中文简体字版由 Wiley 授权机械工业出版社独家出版。未经出版者书面允许,本书的任何部分不得以任何方式复制或抄袭。版权所有,翻印必究。

北京市版权局著作权合同登记 图字:01-2014-2699

图书在版编目(CIP)数据

衍射、傅里叶光学及成像/(美)埃尔索伊(Ersoy, O.)著;蒋晓瑜等译.—北京:机械工业出版社,2015.11

书名原文:Diffraction, Fourier Optics and Imaging

ISBN 978-7-111-52389-5

I. ①衍… II. ①埃…②蒋… III. ①衍射—研究②傅里叶光学—研究
IV. ①O436.1②O438.2

中国版本图书馆CIP数据核字(2015)第301113号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:吉玲 责任编辑:吉玲 熊海丽 刘丽敏

版式设计:霍永明 责任校对:张晓蓉

封面设计:张静 责任印制:乔宇

北京玥实印刷有限公司印刷

2018年1月第1版第2次印刷

184mm×260mm·16.5印张·404千字

标准书号:ISBN 978-7-111-52389-5

定价:65.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线:010-88361066

机工官网:www.cmpbook.com

读者购书热线:010-68326294

机工官博:weibo.com/cmp1952

010-88379203

金书网:www.golden-book.com

封面无防伪标均为盗版

教育服务网:www.cmpedu.com

译者序

人类社会已经由工业时代进入信息时代，信息正在以系统的方式改变整个社会，信息的获取、处理及分析的作用越来越重要，而成像是实现信息获取与处理的一种最主要的方式之一，被广泛应用于工业、军事、医学、科学研究等各行各业，深刻地影响着人类的日常生活与社会发展。

本书以傅里叶分析作为桥梁，将光学衍射现象和成像技术连接起来，系统地介绍了与成像有关的相关衍射光学现象和傅里叶分析方法。本书以原作者 30 多年的教学、科研为基础，包含了作者对于衍射与成像的独特理解，并融入了原作者在相关方面的原创性理论和应用实例。对于那些具有一定的光学基础，并渴望学习成像技术及其分析方法的读者来说，本书是实现快速入门的理想选择。本书注重理论与实际的结合，具有如下特点。

首先，对衍射现象的介绍深入浅出。对于波动传播和标量衍射理论的介绍，可以使读者快速掌握与衍射相关的物理基础理论；而对于衍射近似方法（菲涅耳衍射、夫琅禾费衍射）以及逆向波动理论的介绍及相关光学器件的介绍，可以让读者对衍射现象的认识与理解进一步加深。

其次，对成像技术的介绍细致全面。针对透镜成像，介绍了几何光学下的薄透镜与厚透镜成像原理，以傅里叶分析方法为工具，分别介绍了均匀介质与非均匀介质、相干与非相干条件下的波动传播方法及透镜成像特性，并介绍了相关成像技术的应用实例，如相衬显微成像与扫描共焦显微成像、电子束刻蚀、定向耦合器等。

此外，本书还对与衍射和成像相关的一些应用领域做了深入浅出的介绍。这些领域涵盖了全息及其编码，衍射光学器件及其编码，计算成像技术如合成孔径雷达及计算机断层扫描、阵列波导光栅等，并对孔径尺寸小于入射光波长时的衍射分析方法，如光束传播法、时域有限差分法和傅里叶模式法做了介绍。

本书的第 1 ~ 第 10 章等由裴闯、闫兴鹏、程长青、蒋晓瑜翻译，第 11 ~ 第 20 章与附录等由赵锴、严志强、赵哲、丛彬翻译，并由蒋晓瑜完成校正工作，蒋晓瑜与闫兴鹏共同完成全文定稿。

感谢 Okan K. Ersoy 教授对本书中文版出版给予的大力支持。感谢机械工业出版社吉玲编辑在本书翻译过程中给予的帮助与合作。由于译者水平有限，不妥和疏误之处在所难免，欢迎广大读者批评指正。

译者

前言

衍射与成像是许多现代科学领域中的核心课题之一，而傅里叶分析和合成技术又是贯穿这一课题的统一主线。例如，通过对傅里叶方法的研究与发展而衍生出了诸多现代成像技术。

本书具有显著的原创性，这源于作者长达 30 多年的教学、研究与项目开发经历。回顾本书涉及的内容在近年来取得的进展，并且随着技术的进步这些内容显得越来越重要，这本身就是一件令人快慰的经历。这些内容涉及的主题为数众多，且包含诸多学科。

即使傅里叶理论是最核心的，其也需要许多其他的主题来进行补充，例如，线性系统理论、优化理论、数值方法、成像理论以及信号与图像处理理论；并且，理论也需要与问题的具体实施和材料的制备相结合。因此很难对这一领域进行简单的描述。技术的日益进步使得该领域至关重要，从而需要开设一些课程以涵盖该科学与技术领域的主要议题；同时，也需要帮助学生理解这些课程的重要性，使他们为现代技术的学习做好准备。

本书可作为大学高年级和研究生阶段涉及领域的相关课程的教科书。以本书涵盖的议题为基础，可以设计开设 1/4 学期或者半学期的课程。

本书共包含 20 章和 3 个附录。前三章可认为是基础知识的导论性讨论。

第 1 章给出了衍射、傅里叶光学和成像的简要介绍，并给出了一些现代技术中正在出现的技术实例。

第 2 章综述了本书其他部分所需要的线性系统理论和变换理论。具体描述了连续空间的傅里叶变换、实傅里叶变换及其性质，并包含一些例题。其他涉及的主题包含在附录中：附录 A 介绍了脉冲函数，附录 B 介绍了线性矢量空间，附录 C 介绍了离散时间傅里叶变换、离散傅里叶变换和快速傅里叶变换 (FFT)。

第 3 章为波传播的基础知识。先给出了波的一般描述，并介绍了各种类型的波。接下来主要介绍了电磁波及其性质，并特别强调了平面波。

第 4 章为标量衍射理论的基础知识。介绍了亥姆霍兹方程、平面波角谱、衍射的菲涅耳-基尔霍夫和瑞利-索末菲理论。这些理论将波的传播描述为与傅里叶变换密切相关的线性积分变换。

第 5 章讨论了菲涅耳近似和夫琅禾费近似，其允许衍射可用傅里叶变换的方式进行描述。作为这些近似的特殊应用领域，还介绍了具有诸多用途的衍射光栅。

衍射通常是采用前向波动传播的方式来进行讨论的，本书第 6 章反其道而行之，引入了逆向波动理论来讨论衍射现象，这在一些特定类型的成像方法和光学元器件设计的迭代优化算法中有重要应用。本章中，着重介绍了逆向菲涅耳衍射、逆向夫琅禾费衍射和逆向角谱表述。



前述章节介绍的方法主要对靠近波动传播方向 z 轴的波动成立。换言之, 这些方法的结果仅在波动方向与 z 轴夹角很小的时候精确。而且当距离衍射面很近的时候, 菲涅耳衍射和夫琅禾费衍射也不再有效。在第7章中, 采用了一种新方法使这些问题得到了很大程度的解决, 这种新方法被称为近场-远场近似(NFFA)方法。该方法主要包含两方面内容: 第一个是包含泰勒级数展开中二次方以上的项, 第二个是推导了一种在输出面上确定半规则采样位置的方程, 从而使FFT仍能被用来计算波动传播。因此, NFFA方法的计算速度快, 并在宽角度衍射、近场和远场波动计算等应用中仍然适用。

第8章讨论当衍射孔径远大于波长时适用的几何光学。透镜的设计经常采用几何光学来进行, 本章给出了薄透镜和厚透镜的光线方程及其传播方程, 并给出了几何光学理论与波动光学理论的关系。

透镜成像是最典型的成像方式。第9章和第10章主要介绍均匀媒质中的透镜成像, 并将此成像过程转化为线性系统。第9章讨论了利用二维傅里叶变换描述的相干光成像。同时还讨论了相衬显微成像与扫描共焦显微成像这两种重要应用, 以此体现本章介绍的理论是如何在实际中应用的。

第10章是第9章的延续, 主要介绍了准单色波的透镜成像。阐述了相干成像和非相干成像, 详细讨论了包括希尔伯特变换和解析信号在内的一些理论基础。同时还介绍了光学畸变及利用泽尔尼克多项式进行的畸变估计。

前述章节介绍的均为成像理论。第11章介绍具体应用中的问题。实际应用中有诸多的方法, 本章主要介绍其中的两种, 分别是照相薄膜和胶片, 以及用于衍射光学的电子束刻蚀。

在第9章和第10章中, 传播介质被假设为均匀媒质(具有常数折射率)。第12章讨论非均匀媒质中的波传播问题, 此时波传播的数值求解变得更加困难。本章给出了非均匀媒质中的广义亥姆霍兹方程和傍轴波动方程。同时引入光束传播法(BPM), 这是一种用来求解非均匀媒质中波动传播的强有力的数值方法。本章还给出这种方法的应用实例: 一种允许光从一个波导转移到另一个波导的定向耦合器。

第13章讨论了一种最重要的三维成像技术——全息术。本章描述了用来分析全息成像、放大率、畸变等的最基本的全息方法。

第14章主要介绍了衍射光学器件(DOE)、新型成像方法和亚波长尺度的衍射现象。同时还讨论与信号或图像处理和迭代优化技术相关的内容。这些技术对前述章节也具有非常重要的意义, 特别是当光学图像被进一步进行数字处理的时候。

接下来的两章讨论了衍射光学, 包括利用数字计算机制作全息图(常称为DOE), 以及采用物理的方法即采用记录系统生成DOE。在一些约束条件下进行DOE设计及实现, 涉及对入射波的幅值和相位进行编码, 这是一种从通信工程中引入的思路。这种方法非常多, 第15章首先介绍了罗曼法, 这是历史上第一种此类编码方法。接下来介绍了在许多成像中(如三维成像)非常有用的两种方法, 一种是单幅图像全息术, 通过对幅值和相位的采样以及非线性编码, 其可以仅产生所需要的图像, 而将高阶衍射图像抑制掉。本章最后一节介绍了DOE用作平面透镜的二元光学器件——二元菲涅耳波带片。

第16章是第15章的延续。本章介绍了DOE编码的新方法及其改进。第14章中讨



论过的凸集投影法 (POCS) 在本章中被用来进行 DOE 编码。此外还讨论了一些其他方法, 包括实施更为简单的虚拟全息术, 利用 POCS 来进行子全息图优化的交叉迭代技术 (IIT), 利用 FFT 的频率抽取特性对 IIT 进行改进的 ODIFIIT, 以及可得到更高精度的混合罗曼-ODIFIIT 法。

第 17 章和第 18 章主要介绍计算成像技术的。第 17 章介绍了第一种计算成像技术——合成孔径雷达 (SAR)。在某些方面, 原始 SAR 图像与 DOE 图像类似, 也涉及光学和衍射原理, 例如菲涅耳近似的利用, 只有经过进一步处理, 或者更确切的说法是解码, 才能重构地面地形的图像。利用 SAR 得到的图像在地球遥感中非常有用。

第 18 章涵盖了计算成像的第二部分——计算机断层扫描 (CT)。CT 的理论基础是与傅里叶变换具有“近亲”关系的拉冬变换。投影切片理论给出了如何利用投影的一维傅里叶变换来在二维傅里叶变换平面上产生图像谱的切片。CT 是高度数值化的, 在本章的剩余部分中将给出它的一些图像重构的算法。

光学傅里叶变换技术在光通信和光网络中已变得非常重要。第 19 章介绍了用于密集波分复用 (DWDM) 的阵列波导光栅 (AWG)。AWG 也被叫作相控阵 (PHASAR), 它是一种采用阵列波导制成的成像器件。波导尺寸是不同的, 且差别为中心波长的整数 m 倍, 这样可以从相邻波导间得到大的相位差。整数 m 相当大, 例如可达到 30, 这对高分辨能力的相控阵器件是合理的, 意味着波长的微小变化也可以在输出面上被分辨出来。这就是采用波导而不是采用自由空间的原因。然而, 由于这些波长之间不会发生干涉, 通过波导在输出平面上生成不同波长的点的像, 利用的正是衍射, 这与 DOE 类似, 是一种被采样的器件。并且这些点的像以一定的间隔重复, 这限制了用于成像的波长数。接下来本章还讨论了一种解决上述问题的方法——不规则采样零交叉法 (MISZC)。MISZC 方法源于第 15 章所讨论的单图全息术。

当孔径尺寸比入射光的波长小时, 标量衍射理论给出的结果将不再精确。此时需要对麦克斯韦方程组进行数值求解, 目前已经有了一些求解方法, 如有限差分、傅里叶模式分析和有限元法。前两种方法在第 20 章中进行了讨论。首先采用克尼二氏法, 把第 12.4 节讨论的傍轴 BPM 写成有限差分形式, 然后讨论利用帕德近似得到的宽角度 BPM。最后一节重点介绍了时域有限差分法和傅里叶模式法。

作者在世界各地的诸多同事、秘书、朋友和学生为本书底稿的准备做出了有益帮助。我特别要感谢他们无论在何种境况下, 都给予我持续的激励。我也非常幸运能够与 John Wiley & Sons 在此书的出版项目中一起合作。他们对我付出了惊人的耐心。没有这些耐心, 我将完不成该出版项目。特别感谢 George Telecki 编辑在整个项目进行中的耐心和支持。

目 录

译者序

前言

第1章 衍射、傅里叶光学与成像	1
1.1 本章引论	1
1.2 日益重要的应用举例	2
1.2.1 密集波分复用 (DWDM) /解复用	2
1.2.2 光学与微波 DWDM 系统	2
1.2.3 衍射与亚波长光学元件	2
1.2.4 纳米尺度衍射器件与严格衍射理论	3
1.2.5 现代成像技术	3
第2章 线性系统与变换	4
2.1 本章引论	4
2.2 线性系统与平移不变性	4
2.3 连续空间傅里叶变换	6
2.4 傅里叶变换的存在条件	7
2.5 傅里叶变换的性质	7
2.6 实傅里叶变换	11
2.7 幅度谱和相位谱	13
2.8 汉克尔变换	13
第3章 波传播基本定理	16
3.1 本章引论	16
3.2 波	16
3.3 电磁波	20
3.4 相量表示法	21
3.5 非电介质中的波动方程	21
3.6 非电介质中的波动方程的相量表示	23
3.7 平面电磁波	24
第4章 标量衍射理论	27
4.1 本章引论	27
4.2 亥姆霍兹方程	27
4.3 平面波角谱	29
4.4 平面波角谱的快速傅里叶变换 (FFT)	31
4.5 基尔霍夫衍射理论	35
4.5.1 基尔霍夫衍射理论	35

4.5.2 菲涅耳-基尔霍夫衍射公式	36
4.6 瑞利-索末菲衍射理论	37
4.6.1 基尔霍夫近似	38
4.6.2 第二瑞利-索末菲公式	38
4.7 第一瑞利-索末菲衍射积分的另一种推导	38
4.8 非单色波的瑞利-索末菲衍射积分	39
第5章 菲涅耳与夫琅禾费近似	40
5.1 本章引论	40
5.2 菲涅耳衍射	40
5.3 菲涅耳衍射的快速傅里叶变换的实现	45
5.4 傍轴波动方程	46
5.5 夫琅禾费衍射	47
5.6 衍射光栅	48
5.7 正弦幅值光栅的夫琅禾费衍射	49
5.8 正弦幅值光栅的菲涅耳衍射	50
5.9 正弦相位光栅的夫琅禾费衍射	51
5.10 狭缝衍射光栅	52
第6章 逆衍射	53
6.1 本章引论	53
6.2 菲涅耳和夫琅禾费近似的逆衍射问题	53
6.3 角谱表述下的逆衍射	53
6.4 分析	54
第7章 宽角度下标量衍射理论的近场和远场近似	57
7.1 本章引论	57
7.2 菲涅耳和夫琅禾费近似的回顾	58
7.3 径向近似	59
7.4 高阶改进与分析	60
7.5 逆衍射与迭代优化	61
7.6 数值算例	62
7.7 更高精度近似	67
7.8 小结	68
第8章 几何光学	69
8.1 本章引论	69



8.2 光线的传播	69	11.5 漂白	113
8.3 光线方程	71	11.6 衍射光学、二元光学和数字光学	113
8.4 程函方程	72	11.7 电子束刻蚀	114
8.5 局部空间频率和光线	73	11.7.1 衍射光学元件的应用	115
8.6 子午面光线的矩阵表示	75	第12章 光在非均匀介质中传播	116
8.7 厚透镜	79	12.1 本章引论	116
8.8 光学系统的人瞳和出瞳	80	12.2 非均匀介质的亥姆霍兹方程	116
第9章 傅里叶变换和相干光学系统		12.3 非均匀介质的傍轴波动方程	117
成像	82	12.4 光束传播法 (BPM)	117
9.1 本章引论	82	12.4.1 波在折射率为 \bar{n} 的均匀介质中传播	118
9.2 薄透镜的相变	82	12.4.2 虚透镜效应	118
9.3 透镜的傅里叶变换	83	12.5 波在定向耦合器中的传播	119
9.3.1 波场入射在透镜上	83	12.5.1 耦合模理论概述	119
9.3.2 初始波场在透镜左侧	83	12.5.2 耦合模理论和光束传播法 (BPM) 的计算对比	120
9.3.3 初始波场在透镜右侧	84	第13章 全息	122
9.4 基于二维线性滤波的成像	85	13.1 本章引论	122
9.4.1 有限透镜孔径的影响	86	13.2 相干波前记录	122
9.5 相衬显微术	86	13.2.1 利思-厄帕尼斯克斯 (Leith-Upatnieks) 全息图	123
9.6 共焦扫描显微术	88	13.3 全息图的类型	124
9.6.1 图像形成	88	13.3.1 菲涅耳和夫琅禾费全息图	124
9.7 复杂光学系统的算子代数	90	13.3.2 像面全息图和傅里叶全息图	125
第10章 准单色波成像	94	13.3.3 体全息图	125
10.1 本章引论	94	13.3.4 模压全息图	126
10.2 希尔伯特变换	94	13.4 用计算机模拟全息图的再现	126
10.3 解析信号	96	13.5 全息成像的放大和分析	127
10.4 非单色波场的解析信号表示	99	13.6 像差	129
10.5 准单色波、相干和非相干波	99	第14章 切趾，超分辨率和缺失信息恢复	131
10.6 一般成像系统中的衍射效应	100	14.1 本章引论	131
10.7 准单色波场成像	100	14.2 切趾	132
10.7.1 干涉成像	101	14.2.1 离散时间窗	132
10.7.2 非相干成像	102	14.3 两点分辨率和信号恢复	134
10.8 衍射受限成像系统的频率响应	102	14.4 收缩	135
10.8.1 相干成像系统	102	14.4.1 收缩映射定理	136
10.8.2 非相干成像系统	102	14.5 用于信号恢复的收缩迭代方法	137
10.9 光学传递函数的计算机计算	105	14.6 约束迭代反卷积	138
10.10 像差	106	14.7 投影方法	140
10.10.1 泽尔尼克 (Zernike) 多项式	107	14.8 凸集投影方法 (POCS)	141
第11章 基于波调制的光学器件	110	14.9 盖师贝格-帕普里斯算法	
11.1 本章引论	110		
11.2 照相胶片和干板	110		
11.3 胶片的透光率	111		
11.4 调制传递函数	112		



(GP 算法)	142	16.4 交叉迭代技术 (IIT)	179
14.10 其他的凸集投影算法	142	16.4.1 IIT 算法实验	180
14.11 从相位恢复信号	143	16.5 最优频率抽取交叉迭代技术	
14.12 用离散傅里叶变换从离散相位		(ODIFIIT)	181
函数中重构信号	144	16.5.1 ODIFIIT 法的实验	184
14.13 广义投影	145	16.6 复合的罗曼-ODIFIIT 法	185
14.14 从幅值恢复信号	146	16.6.1 LM-ODIFIIT 法的计算机实验	186
14.14.1 陷阱和隧道	147	第 17 章 计算机成像技术 I：合成	
14.15 用最小二乘法和广义逆法恢复		孔径雷达	189
图像	148	17.1 本章简介	189
14.16 通过奇异值分解计算 H^+	148	17.2 合成孔径雷达	189
14.17 最速下降法	150	17.3 测距分辨率	190
14.18 共轭梯度法	151	17.4 脉冲波形的选择	191
第 15 章 衍射光学 I	153	17.5 匹配滤波器	192
15.1 本章引论	153	17.6 运用匹配滤波进行脉冲压缩	193
15.2 罗曼法	154	17.7 方位角分辨率	195
15.3 罗曼法中的近似	155	17.8 简化的 SAR 成像理论	196
15.4 常幅值罗曼法	156	17.9 用菲涅耳近似进行图像重构	198
15.5 量化罗曼法	156	17.10 数字图像重构算法	199
15.6 罗曼法的计算机仿真	157	第 18 章 计算机成像技术 II：图像	
15.7 基于硬限幅的傅里叶法	159	投影重构	201
15.8 一种 3-D 点图像重构的简单算法	160	18.1 本章引论	201
15.8.1 实验	161	18.2 Radon 变换	201
15.9 快速加权零交叉算法	162	18.3 投影切片定理	202
15.9.1 离轴平面参考波	164	18.4 逆 Radon 变换	203
15.9.2 实验	164	18.5 Radon 变换的性质	204
15.10 单图全息术	164	18.6 利用投影重构信号	205
15.10.1 成像分析	166	18.7 傅里叶重构法	205
15.10.2 实验	168	18.8 滤波-反投影算法	206
15.11 菲涅耳波带片	169	第 19 章 密集波分复用	209
第 16 章 衍射光学 II	170	19.1 本章引论	209
16.1 本章引论	170	19.2 阵列波导光栅	209
16.2 虚拟全息	170	19.3 不规则采样的零交叉法	
16.2.1 相位的确定	170	(MISZC)	212
16.2.2 孔径效应	172	19.3.1 修正项的计算方法	213
16.2.3 成像分析	172	19.3.2 推广 MISZC 到三维结构	214
16.2.4 信息容量, 分辨率, 带宽及		19.4 对 MISZC 的分析	214
冗余	174	19.4.1 色散分析	216
16.2.5 体积效应	175	19.4.2 有限尺寸孔径	217
16.2.6 波长变化和/或构建及重构中全息		19.5 计算实验	217
图尺寸的变化引起的失真	176	19.5.1 点源孔径	218
16.2.7 实验	176	19.5.2 多信道	219
16.3 用于二元 DOE 设计的 POCS 方法	178		



19.5.3 有限尺寸孔径	219	附录 A 脉冲函数	235
19.5.4 生成负相位的方法	220	附录 B 线性矢量空间	237
19.5.5 误差容限	220	B.1 本章引论	237
19.5.6 3-D 仿真	221	B.2 矢量空间的性质	239
19.5.7 相位量化	221	B.3 内积矢量空间	240
19.6 实现过程中的问题	222	B.4 希尔伯特空间	242
第 20 章 严格衍射理论的数值方法	224	附录 C 离散时间傅里叶变换、离散傅里	
20.1 本章引论	224	叶变换与快速傅里叶变换	243
20.2 基于有限差分法的 BPM	224	C.1 离散时间傅里叶变换	243
20.3 宽角度 BPM	226	C.2 离散时间傅里叶变换和傅里叶	
20.4 有限差分	228	变换之间的联系	244
20.5 时域有限差分法	228	C.3 离散傅里叶变换	245
20.6 计算机实验	231	C.4 快速傅里叶变换	246
20.7 傅里叶模式法	232	参考文献	247
附录	235		

第 1 章 衍射、傅里叶光学与成像

1.1 本章引论

当光场通过“障碍物”时，光场的表现不能简单地用一些光线来描述。例如，当一束平面波穿过某个孔径时，一些光波偏离了原来的传播方向，传播的光场在尺寸和形状上都不同于最初透过该孔径的光场分布。这种光偏离直线传播的现象被称之为衍射。

波的传播过程包含衍射。衍射发生于任何一种波的传播过程中，例如，电磁波、声波、无线电波、超声波以及海浪等。本书主要研究的是电磁波，研究结果也可以直接应用于其他类型的波。

由于成像系统的光学分辨率取决于衍射效应，因此在过去的光学设计中，衍射通常被认为是一种干扰。模拟全息技术（在 20 世纪 40 年代得到演示、20 世纪 60 年代得到实用）、合成孔径雷达（20 世纪 60 年代）、计算全息与相息图，以及更具一般意义的衍射光学元件（20 世纪 60 年代后期）的大力发展，标志着基于衍射的光学器件发展的开端。近年来，衍射与折射光学器件的结合，例如，采用衍射光学校正的折射透镜，显示出光学设计达到了一个新的阶段。

傅里叶光学研究包括连续空间与离散空间的傅里叶变换相关的光学课题及其应用。因此，标量衍射理论是傅里叶光学的一部分。在傅里叶光学的其他重要课题中，我们分析论证了透镜的傅里叶变换和成像特征、光学成像系统的频率分析、空间滤波与光信息处理、模拟与计算全息、衍射光学器件（Diffractive Optical Element, DOE）的分析与设计以及最新的成像技术。

现代衍射、成像理论，特别是依赖于傅里叶分析与合成技术的相关理论，已成为理解、分析与合成现代成像、光通信与网络、微/纳设备与系统必不可少的工具。一些典型的应用包括断层摄影术、核磁共振成像、合成孔径雷达（Synthetic Aperture Radar, SAR）、干涉合成孔径雷达、共焦显微镜，另外包括应用在通信与网络设备中的器件，例如，光纤与集成电路中的定向耦合器、超短光脉冲分析技术、计算全息技术、模拟全息、衍射光学器件、光栅、波带片、光学与微波相控阵以及利用电磁波的无线系统。

微纳光学在许多方向上处于飞速发展并产生新成果的阶段。在该领域中，衍射变得越来越重要。当感兴趣的波长与衍射源的尺寸相当甚至小于衍射源的尺寸时，所有的波动现象都受衍射的制约。显然，技术的发展会越来越注重尺寸更加微小的设备和系统。由于系统的复杂度增加，采用衍射来检测和分析这些系统将成为一种主要的方法。

在更先进的计算机技术中，在极端高速的情况下，采用传统的基于电子线路的时钟与同步通信方法的设计将难以继续发展。逐渐增加的复杂度使采用光互联的需求进一步增加，当前这种情况已经在具有非常复杂的集成电路的计算机工业中出现了，看来只有光学技术才能满足日渐复杂的微处理器需求的时代已经来临。原因很简单，相对电子传输，光子不需要受



传统电路中铜线的限制。

一些微电子实验仪器设备，例如，扫描电子显微镜、反应离子刻蚀设备等，对于现代微纳制造与测试将变得越来越重要。而采用衍射的方法测试这些技术获得的器件是一种非常有应用。与其他光学与数字信号处理设备相结合，使相机拍摄图像数字化并进一步进行图像处理变成了现实，从而能够处理不同种类的复杂微纳系统产生的衍射图像。同时，这种技术在实现衍射光学器件方面也最有竞争力。

综上所述，傅里叶及相关变换与衍射技术在科学技术的诸多不同领域中都具有广泛的应用，特别是在成像、通信与网络等相关领域中。由于系统经常被视为线性系统，线性系统理论在分析中变得十分重要。可利用卷积技术，采用傅里叶变换来进行系统分析。

1.2 日益重要的应用举例

随着上述技术的成熟与系统尺寸的缩小，这些技术产生了很多日益重要的应用。下面详细介绍近期出现的一些特例。

1.2.1 密集波分复用（DWDM）/解复用

多频谱通信、网络与计算等现代技术已经被越来越多地光学化。例如，在通信与联网的新近进展中，由于对频道（波长）需求的增长，密集波分复用（Dense Wavelength Division Multiplexing, DWDM）/解复用技术变得更加重要。

DWDM 技术为解决通信与网络的容量与灵活性问题提供了一个新的方向。该技术提供了非常大的传输容量和新的网络体系结构。DWDM 系统的主要组件包括波长多路复用器和解复用器。商业上可用的光学器件是基于光纤和微光学技术的。对于集成光学多路复用器（解复用器）的研究越来越多地集中在基于光栅与基于相控阵（Phased-array, PHASAR）（也是所谓的阵列波导光栅）的器件上。它们都是成像设备，即将输入波导的场分布用色散的方法成像在一个输出波导阵列上。在基于光栅的器件中，一个垂直蚀刻的反射光栅为解复用提供聚焦与色散特性；在基于相控阵的器件中，这种特性由波导阵列提供，根据所需要的成像与色散特性选择合适的波导阵列长度。由于相控阵器件采用传统的波导技术实现，不需要基于光栅器件中的垂直蚀刻步骤，因此其具有更好的鲁棒性和制造中的容错性。这些设备在很大程度上都是基于衍射的。

1.2.2 光学与微波 DWDM 系统

对于光学与无线微波相控阵密集波分复用系统的技术研究也正在迅速发展，由微波阵列天线提供复用/解复用在无线通信中变得越来越普遍。光学器件的应用，则能够产生更大的带宽，更小的尺寸和重量，以及对电磁干扰和串扰的高免疫性。

1.2.3 衍射与亚波长光学元件

传统的光学器件如透镜、反射镜及棱镜等都是基于光的反射与折射特性的。相反地，采用浮雕形式存在的衍射光学器件，是基于光的衍射原理工作的。这种器件在许多应用中变得越来越重要。



20 世纪 70 年代集成电路的革命性发展，导致光波长级别的衍射器件得以实际应用。光学与电子束光刻技术的发展，允许在抗蚀材料上生成高分辨率的复杂条纹。干法刻蚀技术使得通过具有精细线条纹与尖锐边缘的表面浮雕来控制相位成为可能。金刚石切削机器与激光写入技术的进展为构造高精度的衍射光学器件提供了新的途径。

近年来，商业化的基于晶片的纳米加工技术，为一种新的光学器件的产生提供了可能——亚波长光学器件（Subwavelength Optical Element, SOE）。由于具有比光的波长小很多的结构，光与具有如此精细表面结构的器件的相互作用的物理过程，产生了一种新的光处理函数的重组编排（分布）。与现有的许多技术相比，这种新的分布具有更高的密度、更好的鲁棒性、更高层次的整合性，能够从根本上改变现有的光学系统设计。

1.2.4 纳米尺度衍射器件与严格衍射理论

这种器件物理学上需严格地应用麦克斯韦方程的边界条件来描述光与该结构的相互作用。例如，在电信使用的 980 ~ 1800nm 的光波上，要求这种器件结构尺寸在几十到几百纳米的级别。在这种尺度的低端，将会发现单电子或者量子效应。在许多实际应用中，亚波长结构作为纳米衍射光栅，其与入射光产生互作用可以利用严格的衍射光栅理论和前面提到的麦克斯韦边界条件来建模。

尽管这种光学效应近年来已有研究，但是这种光学器件昂贵的工业制造尚难以实现。在实验环境中构建亚波长光栅结构需要高能技术，例如，电子束（E-beam）平版印刷术。E-beam 设备现在能够产生直径 5 ~ 10nm 的点，因此，能够曝光线宽在 0.1 μ m 或 100nm 的条纹。

“全光系统”的出现很大程度上依赖于新技术。这些新的技术将系统从数据率、带宽、反应时间、信息损失、成本以及协议（带有光电转换的光学系统协议）中解放出来。另外，新的技术，如基于微机械系统（Microelectro Mechanical System, MEMS）的微镜，允许在光路外进行光学开关控制，由此电子与光学参数才能够独立地进行调整以实现全局最优。针对这些器件衍射的研究对于新系统、新技术的发展也将同样重要。

1.2.5 现代成像技术

如果一个成像系统的源场（光场、电磁场）具有空间相干的特性，则该源场被称为相干场，同时可以被描述为复振幅的空间分布。例如，全息就是一种相干成像技术。如果源场不具有空间相干性，则被称为非相干场，可以被描述为实数亮度值的空间分布。激光与微波通常代表相干成像的源场，因此，傅里叶变换与衍射对于理解成像非常重要。太阳光通常代表非相干光源，非相干成像同样可以用傅里叶技术进行分析。

许多现代计算机成像技术在图像重构过程中严重依赖傅里叶变换及相关的计算算法，如合成孔径雷达、投影重构图像（包括 CT、磁共振成像、共焦显微镜、共焦扫描显微术等），都属于此类技术。

第 2 章 线性系统与变换

2.1 本章引论

衍射与成像过程经常被建模为**线性系统**。首先，一个系统是一个输入-输出的映射。也就是说，给系统一个输入，系统会生成一个输出。例如，在一个衍射和成像问题中，输入和输出分别是一个在输入平面的波和另一个与距输入平面有一定距离的对应的衍射波。

光学系统与通信系统是非常相似的。两种系统的首要目的是采集和处理信息。通信系统处理的语言是一维信号（1-D），而图像是二维信号（2-D）。一维信号是典型的时间信号，而二维信号是典型的空信号。例如，一个采用了激光光束的光学系统具有空间相干性。这样，信号可以被描述为 2-D 或者 3-D 复值幅度场。空间相干性对于观察衍射是十分必要的。普通照明（例如白昼）信号不具有空间相干性，因此普通照明信号被描述为 2-D 空间的实值的亮度。

对线性的时间不变与空间不变的通信与光学系统，经常采用傅里叶变换来进行频率分析。非线性光学元件如照相胶片与非线性电子元件如二极管，也具有相似的输入—输出特性。

在两种系统中，傅里叶技术同样可用于系统的合成。一个典型的例子是二维滤波。理论上，光学匹配滤波器、光学图像处理技术与通信和信号处理系统中的匹配滤波器、图像处理技术是相似的。

本章讨论了与衍射、光学成像有关的线性系统理论和傅里叶变换理论，以及相关领域。本章包含 8 个小节。在 2.2 节中强调了线性系统的特性，特别是卷积与平移不变性。在 2.3 节中介绍了 1-D 傅里叶变换和连续空间傅里叶变换（在本书的其他部分中简称为傅里叶变换（FT））。2.4 节给出了傅里叶变换存在的条件。在 2.5 节，对傅里叶变换的特性进行了总结。

到目前为止，讨论的傅里叶变换都具有复指数变换核。我们可以把傅里叶变换定义为一个利用余弦和正弦核函数的实变换，实傅里叶变换有时更加具有实用性。在 2.6 节讨论了 1-D 实傅里叶变换。在 2.7 节定义了 1-D 傅里叶变换的振幅谱和相位谱。

2-D 信号经常具有圆对称特性，特别是在光学与波的传播应用中。在这种情况下，傅里叶变换变成了柱面坐标下的汉克尔变换。在本章的 2.8 节中，将讨论汉克尔变换。

2.2 线性系统与平移不变性

线性系统允许将复杂信号分解成许多基本信号的组合，这些基本信号称为**基元信号**。在傅里叶分析中，基元信号或函数是正弦函数。

在一个线性系统中，一个给定的输入对应一个唯一的输出。但是，多个不同的输入可以



对应相同的输出。因此，这种对应关系可以是一对一，也可以是多对一。

图 2.1 给出了一个二维系统，其中， $u(x, y)$ 是输入信号， $g(x, y)$ 是输出信号，数学上，系统可以被表示为

$$g(x, y) = O[u(x, y)] \quad (2.2-1)$$

在连续空间， $O[\cdot]$ 是一个算符，将输入映射成输出。

在离散空间，点 (x, y) 被采样为 $[\Delta x \cdot m, \Delta y \cdot n]$ ，其中

Δx 和 Δy 是在两个方向上的采样间隔。 $[\Delta x \cdot m, \Delta y \cdot n]$ 可以被简单地表示为 $[m, n]$ ，系统可以写为

$$g[m, n] = O[u[m, n]] \quad (2.2-2)$$

接下来，考虑连续空间的情况。对于一个线性系统，如果以两个输入 $u_1(x, y)$ 和 $u_2(x, y)$ 的任意线性组合作为输入，系统将把它们各自的输出 $g_1(x, y)$ 和 $g_2(x, y)$ 同样进行线性组合输出。该特性称为叠加原理，写为

$$O[a_1 u_1(x, y) + a_2 u_2(x, y)] = a_1 O[u_1(x, y)] + a_2 O[u_2(x, y)] \quad (2.2-3)$$

其中， a_1 和 a_2 是标量。在离散线性系统中，上面的 (x, y) 被 (m, n) 代替。

假设在 (x_1, y_1) 处的输入为 δ 函数 $\delta(x_1, y_1)$ （参考附录 A 中关于 δ 函数的讨论），在 (x, y) 处的输出定义为

$$h(x, y; x_1, y_1) = O[\delta(x - x_1, y - y_1)] \quad (2.2-4)$$

$h(x, y; x_1, y_1)$ 被称为系统的脉冲响应（点扩散函数）。

δ 函数的筛选性质允许任意的输入 $u(x, y)$ 表示为

$$u(x, y) = \iint_{-\infty}^{+\infty} u(x_1, y_1) \delta(x - x_1, y - y_1) dx_1 dy_1 \quad (2.2-5)$$

此时，输出函数可以表示为

$$\begin{aligned} g(x, y) &= O[u(x, y)] \\ &= \iint_{-\infty}^{+\infty} u(x_1, y_1) O[\delta(x - x_1, y - y_1)] dx_1 dy_1 \\ &= \iint_{-\infty}^{+\infty} u(x_1, y_1) h(x, y; x_1, y_1) dx_1 dy_1 \end{aligned} \quad (2.2-6)$$

这个结果被称为“叠加积分”。物理上， δ 函数对应于一个点源。叠加积分告诉我们，我们要描述系统，就需要知道系统在感兴趣平面上对所有点源的响应。

如果一个系统输入的平移会产生相同的输出平移，则称之为空间不变或者是平移不变的线性成像系统，对于在 origin 的一个点源，一个平移不变系统的输出可以写为

$$h(x, y; 0, 0) = O[\delta(x, y)] \quad (2.2-7)$$

如果输入平移为 $\delta(-x_1, -y_1)$ ，该平移不变系统的输出一定是 $h(x - x_1, y - y_1; 0, 0)$ ，这种现象经常被简写为

$$h(x, y; x_1, y_1) = h(x - x_1, y - y_1) \quad (2.2-8)$$

此时，叠加积分变成

$$g(x, y) = \iint_{-\infty}^{+\infty} u(x_1, y_1) h(x - x_1, y - y_1) dx_1 dy_1 \quad (2.2-9)$$

采用变量代换，上式可以表示为

$$g(x, y) = \iint_{-\infty}^{+\infty} h(x_1, y_1) u(x - x_1, y - y_1) dx_1 dy_1 \quad (2.2-10)$$



这和 $h(x, y)$ 与 $u(x, y)$ 卷积生成 $g(x, y)$ 是一致的。其可以被写成

$$g(x, y) = h(x, y) * u(x, y) \quad (2.2-11)$$

这个结果的重要意义在于，一个线性平移不变 (Linear Shift-invariant, LSI) 系统可以通过卷积进行运算。这样，卷积定理可以用于表示系统的输入输出关系

$$G(f_x, f_y) = H(f_x, f_y) = U(f_x, f_y) \quad (2.2-12)$$

其中， $G(f_x, f_y)$ ， $H(f_x, f_y)$ 和 $U(f_x, f_y)$ 分别是 $g(x, y)$ ， $h(x, y)$ 和 $u(x, y)$ 的傅里叶变换。傅里叶变换将在下节中进行讨论。 $H(f_x, f_y)$ 由下式给出

$$H(f_x, f_y) = \iint_{-\infty}^{+\infty} h(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (2.2-13)$$

上式被称为系统的传递函数。

在离散空间系统中，叠加积分变成了叠加求和，表示为

$$g[m, n] = \sum_{m_1=-\infty}^{\infty} \sum_{n_1=-\infty}^{\infty} u[m_1, n_1] h[m - m_1, n - n_1] \quad (2.2-14)$$

在离散空间线性平移不变系统中，卷积积分变成了卷积求和，即

$$g[m, n] = \sum_{m_1=-\infty}^{\infty} \sum_{n_1=-\infty}^{\infty} u[m_1, n_1] h[m - m_1, n - n_1] \quad (2.2-15)$$

上式同样可以表示为

$$g[m, n] = \sum_{m_1=-\infty}^{\infty} \sum_{n_1=-\infty}^{\infty} h[m_1, n_1] u[m - m_1, n - n_1] \quad (2.2-16)$$

离散空间线性平移不变系统的传递函数是脉冲响应的离散空间傅里叶变换，表示为

$$H(f_x, f_y) = \sum_{m_1=-\infty}^{\infty} \sum_{n_1=-\infty}^{\infty} h[m_1, n_1] e^{-j2\pi(f_x m_1 \Delta x + f_y n_1 \Delta y)} \quad (2.2-17)$$

其中，卷积定理同样在公式 (2.2-12) 中给出。

2.3 连续空间傅里叶变换

线性性质的特征是允许复杂信号被分解成为许多基本信号的组合，这些基本信号经常被称为基元信号，在傅里叶分析中，基元信号或函数是正弦函数。

对于信号 $u(t)$ ， $-\infty \leq t \leq \infty$ ，其 1-D 傅里叶变换定义为

$$U_c(f) = \int_{-\infty}^{+\infty} u(t) e^{-j2\pi f t} dt \quad (2.3-1)$$

傅里叶逆变换是从基元函数 $e^{j2\pi f t}$ 的角度来表示 $u(t)$ 的，定义为

$$u(t) = \int_{-\infty}^{+\infty} U_c(f) e^{j2\pi f t} df \quad (2.3-2)$$

公式 (2.3-1) 被称为分析方程，而公式 (2.3-2) 被称为合成方程。

多维 (Multidimensional, MD) 傅里叶变换可变成一组可分离的一维变换，因为其傅里叶变换核在每个方向上是可分离的。例如，二维傅里叶变换核 $b(x, y, f_x, f_y)$ 可以表示为

$$b(x, y, f_x, f_y) = b_1(x, f_x) b_2(y, f_y) \quad (2.3-3)$$

$b_i(\cdot, \cdot)$ 中的 i 为 1 或者 2，是一维傅里叶变换核。

信号 $u(x, y)$ ， $-\infty \leq x, y \leq \infty$ ，其二维傅里叶变换定义为



$$U(f_x, f_y) = \iint_{-\infty}^{+\infty} u(x, y) e^{-j2\pi(xf_x + yf_y)} dx dy \quad (2.3-4)$$

其中, f_x 和 f_y 分别对应于 x 与 y 方向上的空间频率。

二维傅里叶逆变换定义为

$$u(x, y) = \iint_{-\infty}^{+\infty} U(f_x, f_y) e^{j2\pi(xf_x + yf_y)} df_x df_y \quad (2.3-5)$$

2.4 傅里叶变换的存在条件

傅里叶变换存在的充分条件总结如下:

- A. 在无穷空间内, 信号必须是绝对可积的;
- B. 在任意有限的子空间内, 信号必须只有有限个间断点和有限个极大值点和极小值点。
- C. 信号必须没有无穷大(小)的间断点。

如果需要, 这三个条件中的任意一个条件都可以减弱。例如, 通常用二维狄拉克 δ 函数表示强而窄的脉冲, 其表示式如下:

$$\delta(x, y) = \lim_{N \rightarrow \infty} N^2 e^{-N^2 \pi(x^2 + y^2)} \quad (2.4-1)$$

这个函数不满足条件 C。下面另外两个函数不满足条件 A:

$$\begin{aligned} u(x, y) &= 1 \\ u(x, y) &= \cos 2\pi f_x x \end{aligned} \quad (2.4-2)$$

对于这些函数, 其傅里叶变换仍然可以用上述非协作广义函数如 δ 函数、以极限方式来定义, 这通常被称为广义傅里叶变换。

例 2.1 根据 2D δ 函数定义, 计算其傅里叶变换。

解: 令

$$u(x, y) = N^2 e^{-N^2 \pi(x^2 + y^2)}$$

这样

$$\delta(x, y) = \lim_{N \rightarrow \infty} u(x, y)$$

我们可以写成

$$U(f_x, f_y) = e^{-\pi(f_x^2 + f_y^2)/N^2}$$

定义 $\Delta(f_x, f_y)$ 为在 $N \rightarrow \infty$ 极限上的傅里叶变换, 得到

$$\Delta(f_x, f_y) = \lim_{N \rightarrow \infty} U(f_x, f_y) = 1$$

2.5 傅里叶变换的性质

2-D 傅里叶变换的性质是一维傅里叶变换的扩展。需要定义在二维空间的奇函数和偶函数。

如果一个信号 $u(x, y)$ 满足

$$u(x, y) = u(-x, -y) \quad (2.5-1)$$

那么这个函数为偶函数。如果 $u(x, y)$ 满足

$$u(x, y) = -u(-x, -y) \quad (2.5-2)$$



那么这个函数是奇函数。

如上的定义显示其具有二重对称性，可以将其扩展到二维空间的四重对称。

下面列出傅里叶变换的性质。

性质 1：线性性质

如果 $g(x, y) = au_1(x, y) + bu_2(x, y)$ ，那么

$$G(f_x, f_y) = aU_1(f_x, f_y) + bU_2(f_x, f_y) \quad (2.5-3)$$

性质 2：卷积定理

如果 $g(x, y) = u_1(x, y) * u_2(x, y)$ ，那么

$$G(f_x, f_y) = U_1(f_x, f_y) U_2(f_x, f_y) \quad (2.5-4)$$

性质 3：相关性质

相关与卷积是相似的， $u_1(x, y)$ 与 $u_2(x, y)$ 的相关定义为

$$g(x, y) = \iint_{-\infty}^{\infty} u_1(x_1, y_1) u_2(x + x_1, y + y_1) dx_1 dy_1 \quad (2.5-5)$$

如果 $g(x, y) = u_1(x, y) \circ u_2(x, y)$ ，其中 \circ 代表 2-D 相关，那么

$$U(f_x, f_y) = U_1(f_x, f_y) U_2^*(f_x, f_y) \quad (2.5-6)$$

性质 4：调制性质

如果 $g(x, y) = u_1(x, y) u_2(x, y)$ ，那么

$$G(f_x, f_y) = F_1(f_x, f_y) * F_2(f_x, f_y) \quad (2.5-7)$$

性质 5：可分离性质

如果 $g(x, y) = u_1(x) u_2(y)$ ，那么

$$G(f_x, f_y) = U_1(f_x) U_2(f_y) \quad (2.5-8)$$

性质 6：平移性质

如果 $g(x, y) = u_1(x - x_0, y - y_0)$ ，那么

$$G(f_x, f_y) = e^{-j2\pi(f_x x_0 + f_y y_0)} \cdot U(f_x, f_y) \quad (2.5-9)$$

性质 7：频移性质

如果 $g(x, y) = e^{j2\pi(f_{x0}x + f_{y0}y)} \cdot u(x, y)$ ，那么

$$G(f_x, f_y) = U(f_x - f_{x0}, f_y - f_{y0}) \quad (2.5-10)$$

性质 8：空间域微分运算

如果 $g(x, y) = (\partial^k / \partial x^k) (\partial^l / \partial y^l) u(x, y)$ ，那么

$$G(f_x, f_y) = (2\pi j f_x)^k (2\pi j f_y)^l U(f_x, f_y) \quad (2.5-11)$$

性质 9：频域微分运算

如果 $g(x, y) = (-j2\pi x)^k (-j2\pi y)^l u(x, y)$ ，那么

$$G(f_x, f_y) = \frac{\partial^k}{\partial f_x^k} \frac{\partial^l}{\partial f_y^l} U(f_x, f_y) \quad (2.5-12)$$

性质 10：帕塞瓦尔 (Parseval) 定理

$$\iint_{-\infty}^{+\infty} u(x, y) g^*(x, y) dx dy = \iint_{-\infty}^{+\infty} U(f_x, f_y) G^*(f_x, f_y) df_x df_y \quad (2.5-13)$$

性质 11：对实函数 $u(x, y)$

$$U(f_x, f_y) = U^*(-f_x, -f_y) \quad (2.5-14)$$



性质 12: 对实偶函数 $u(x, y)$

$U(f_x, f_y)$ 同样是实偶数

性质 13: 对实奇函数 $u(x, y)$

$U(f_x, f_y)$ 是虚奇数

性质 14: 空间域拉普拉斯算子

如果 $g(x, y) = \left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \right) u(x, y)$, 那么

$$G(f_x, f_y) = -4\pi^2(f_x^2 + f_y^2)U(f_x, f_y) \quad (2.5-15)$$

性质 15: 频域拉普拉斯算子

如果 $g(x, y) = -4\pi^2(x^2 + y^2)u(x, y)$, 那么

$$G(f_x, f_y) = \left(\frac{\partial^2}{\partial f_x^2} + \frac{\partial^2}{\partial f_y^2} \right) U(f_x, f_y) \quad (2.5-16)$$

性质 16: 信号的二次方

如果 $g(x, y) = |u(x, y)|^2$, 那么

$$G(f_x, f_y) = U(f_x, f_y) * U^*(f_x, f_y) \quad (2.5-17)$$

性质 17: 频谱的二次方

如果 $g(x, y) = u(x, y) * u^*(x, y)$, 那么

$$G(f_x, f_y) = |U(f_x, f_y)|^2 \quad (2.5-18)$$

性质 18: 坐标轴旋转

如果 $g(x, y) = u(\pm x, \pm y)$, 那么

$$G(f_x, f_y) = U(\pm f_x, \pm f_y) \quad (2.5-19)$$

傅里叶变换的重要性质总结如表 2-1 所示。

表 2-1 傅里叶变换的性质 (其中, a 、 b 、 f_{x0} 和 f_{y0} 为非零实常数, k 和 l 是非负整数)

性 质	$g(x, y)$	$G(f_x, f_y)$
线性	$au_1(x, y) + bu_2(x, y)$	$aU_1(f_x, f_y) + bU_2(f_x, f_y)$
卷积性	$u_1(x, y) * u_2(x, y)$	$U_1(f_x, f_y)U_2(f_x, f_y)$
相关性	$u_1(x, y) \cdot u_2(x, y)$	$U_1(f_x, f_y)U_2^*(f_x, f_y)$
调制性	$u_1(x, y)u_2(x, y)$	$U_1(f_x, f_y) * U_2(f_x, f_y)$
可分离性	$u_1(x)u_2(y)$	$U_1(f_x)U_2(f_y)$
平移性	$u(x - x_0, y - y_0)$	$e^{-j2\pi(f_{x0}x + f_{y0}y)} \cdot U(f_x, f_y)$
频移性	$g(x, y) = e^{j2\pi(f_{x0}x + f_{y0}y)} \cdot u(x, y)$	$G(f_x, f_y) = U(f_x - f_{x0}, f_y - f_{y0})$
空间域微分运算	$\frac{\partial^k}{\partial x^k} \frac{\partial^l}{\partial y^l} u(x, y)$	$(2\pi j f_x)^k (2\pi j f_y)^l U(f_x, f_y)$
频域微分运算	$(-j2\pi x)^k (-j2\pi y)^l u(x, y)$	$\frac{\partial^k}{\partial f_x^k} \frac{\partial^l}{\partial f_y^l} U(f_x, f_y)$
空间域拉普拉斯算子	$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u(x, y)$	$-4\pi^2(f_x^2 + f_y^2)U(f_x, f_y)$
频域拉普拉斯算子	$-4\pi^2(x^2 + y^2)u(x, y)$	$\left(\frac{\partial^2}{\partial f_x^2} + \frac{\partial^2}{\partial f_y^2} \right) U(f_x, f_y)$
信号的二次方	$ u(x, y) ^2$	$U(f_x, f_y) * U^*(f_x, f_y)$
频谱的二次方	$x(x, y) * u^*(x, y)$	$ U(f_x, f_y) ^2$



(续)

性 质	$g(x, y)$	$G(f_x, f_y)$
坐标轴旋转	$u(\pm x, \pm y)$	$U(\pm f_x, \pm f_y)$
帕塞瓦尔定理	$\iint_{-\infty}^{+\infty} v(x, y) g^*(x, y) dx dy = \iint_{-\infty}^{+\infty} U(f_x, f_y) G^*(f_x, f_y) df_x df_y$	
实函数 $u(x, y)$		$U(f_x, f_y) = U^*(-f_x, -f_y)$
实偶函数 $u(x, y)$		$U(f_x, f_y)$ 是实偶函数
实奇函数 $u(x, y)$		$U(f_x, f_y)$ 是虚奇函数

例 2.2 计算如下函数 $g(x)$ 的 1-D 傅里叶变换, 表示成 $u(x, y)$ 傅里叶变换的函数形式。

$$g(x) = \int_{-\infty}^{+\infty} u(x, y) dy$$

解: 分别设 $g(x)$ 和 $\delta(x)$ 为 2-D 函数 $g(x) \cdot 1$ 和 $\delta(x) \cdot 1$, $g(x)$ 可以被表示为

$$\begin{aligned} g(x) \cdot 1 &= u(x, y) * [\delta(x) \cdot 1] \\ &= \iint_{-\infty}^{+\infty} u(x_1, y_1) \delta(x_1 - x) dx_1 dy_1 \\ &= \int_{-\infty}^{+\infty} u(x, y) dy \end{aligned} \quad (2.5-20)$$

计算式 (2.5-20) 两边的傅里叶变换, 根据卷积定理和傅里叶变换性质 5 (可分离函数), 得到

$$G(f_x) \delta(f_y) = U(f_x, f_y) \delta(f_y)$$

或者

$$G(f_x) = U(f_x, 0)$$

例 2.3 计算 $g(x, y)$ 的傅里叶变换

$$g(x, y) = u(a_1 x + b_1 y + c_1, a_2 x + b_2 y + c_2)$$

表示成 $u(x, y)$ 傅里叶变换的函数形式。

解:

$$G(f_x, f_y) = \iint_{-\infty}^{+\infty} u(a_1 x + b_1 y + c_1, a_2 x + b_2 y + c_2) e^{-j2\pi(f_1 x + f_2 y)} dx dy$$

令 $x_1 = a_1 x + b_1 y + c_1$, 得到

$$G(f_x, f_y) = \iint_{-\infty}^{+\infty} u \left[x_1, a_2 \left(\frac{x_1 - b_1 y - c_1}{a_1} \right) + b_2 y + c_2 \right] e^{-j2\pi \left[f_x \left(\frac{x_1 - b_1 y - c_1}{a_1} + f_y y \right) \right]} \frac{dx dy}{|a_1|}$$

令 $y_1 = a_2 \left(\frac{x_1 - b_1 y - c_1}{a_1} \right) + b_2 y + c_2$, 得到

$$G(f_x, f_y) = \frac{1}{|D|} \iint_{-\infty}^{+\infty} u(x_1, y_1) e^{-j2\pi(x_1 f'_x + y_1 f'_y + T_1 f_x + T_2 f_y)} dx_1 dy_1 \quad (2.5-21)$$

其中

$$\begin{aligned} D &= a_2 b_1 - a_1 b_2 \\ f'_x &= \frac{1}{D} (-a_1 f_x + a_2 f_y) \end{aligned}$$



$$f'_y = \frac{1}{D}(b_1 f_x - a_1 f_y)$$

$$T_1 = \frac{b_1 c_2 - b_2 c_1}{D}$$

$$T_2 = \frac{a_2 c_1 - a_1 c_2}{D}$$

式 (2.5-21) 等同于

$$G(f_x, f_y) = \frac{1}{|D|} e^{j2\pi(T_1 f_x + T_2 f_y)} U(f'_x, f'_y)$$

因此得到, 如果输入信号是比例的、平移的和偏移的, 那么其变换也是比例的、偏移的, 同时发生相位平移, 而不是空间平移。

例 2.4 计算如下“0-1”函数的傅里叶变换:

$$u(x, y) = \begin{cases} 1 & h_1(x) < y < h_2(x) \\ 0 & \text{其他} \end{cases}$$

其中, $h_1(x)$ 和 $h_2(x)$ 是给定的 x 的单值函数。计算 $U(f_x, f_y)$ 和 $U(f_x, 0)$ 。

解: 图 2.2 给出了 $u(x, y)$ 的定义。它的傅里叶变换为

$$\begin{aligned} U(f_x, f_y) &= \iint_{-\infty}^{+\infty} u(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \\ &= \int_{-\infty}^{+\infty} e^{-j2\pi f_x x} dx \int_{h_1(x)}^{h_2(x)} e^{-j2\pi f_y y} dy \end{aligned}$$

当 $f_y = 0$ 时

$$\begin{aligned} U(f_x, 0) &= \int_{-\infty}^{+\infty} [h_2(x) - h_1(x)] e^{-j2\pi f_x x} dx \\ &= H_2(f_x) - H_1(f_x) \end{aligned}$$

因此, $U(f_x, 0)$ 是函数 $h_2(x)$ 和 $h_1(x)$ 的 1-D 傅里叶变换的差值。

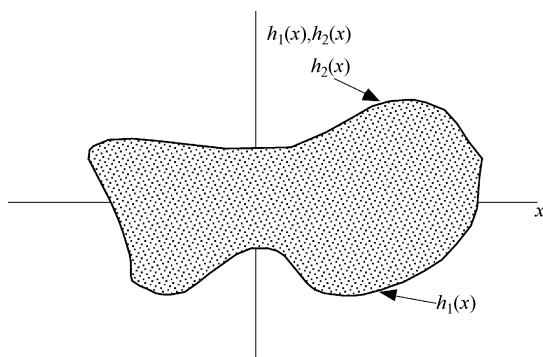


图 2.2 一个典型的 0-1 函数

2.6 实傅里叶变换

有些情况下, 采用实正弦函数和余弦函数为基函数来表示傅里叶变换会更加有效。这种表示方法被称为**实傅里叶变换** (Real Fourier Transform, RFT)。直到 1994 年 (Ersoy, 1994), 傅里叶变换的讨论才变成复数傅里叶变换 (CRF)。例如, 在本书 9.4 节中, 非单色波场可以用 RFT 进行更加有效的推导。在本节和下一节中, 将只讨论 1-D 的 RFT。

信号 $u(x)$ 的 RFT 定义为

$$U(f) = 2w(f) \int_{-\infty}^{+\infty} u(t) \cos(2\pi ft + \theta(f)) dx \quad (2.6-1)$$

其中



$$w(f) = \begin{cases} 1 & f \neq 0 \\ 1/2 & f = 0 \end{cases} \quad (2.6-2)$$

$$\theta(f) = \begin{cases} 0 & f \geq 0 \\ \pi/2 & f < 0 \end{cases} \quad (2.6-3)$$

RFT 逆变换定义为

$$u(t) = \int_{-\infty}^{+\infty} U(f) \cos(2\pi ft + \theta(f)) df \quad (2.6-4)$$

可以看到, 当 $f=0$ 时, $\cos(2\pi ft + \theta(f))$ 等于 $\cos(2\pi ft)$, 而当 $f<0$ 时, 其等价于 $\sin(2\pi |f|t)$ 。这是采用一个积分来涵盖正弦和余弦基本信号的“技巧”。为了达到这个目的, 函数中使用了负频率。有趣的是, 由于式 (2.6-4) 是实数, 因此对于变量 f 而言, $\sqrt{2} \cos(2\pi ft + \theta(f))$ 是正交的。因此,

$$2 \int_0^{+\infty} \cos(2\pi ft + \theta(f)) \cos(2\pi f\tau + \theta(f)) df = \delta(t - \tau) \quad (2.6-5)$$

其中, $\delta(\cdot)$ 是狄拉克脉冲函数。

图 2.3 绘出了 $t=0.25$ 和 $-10 < f < 10$ 时, 基函数 $\sqrt{2} \cos(2\pi ft + \theta(f))$ 的变化情况。

当 $f=0$ 时, 式 (2.6-1) 和式 (2.6-4) 可以被写为

$$U_1(f) = 2w(f) \int_{-\infty}^{+\infty} u(t) \cos(2\pi ft) dt \quad (2.6-6)$$

$$U_0(f) = 2 \int_{-\infty}^{+\infty} u(t) \sin(2\pi ft) dt \quad (2.6-7)$$

及

$$u(t) = \int_0^{\infty} [U_1(f) \cos(2\pi ft) + U_0(f) \sin(2\pi ft)] df \quad (2.6-8)$$

其中, 当 $f=0$ 时, $U(f)$ 等于 $U_1(f)$, 而当 $f<0$ 时, $U(f)$ 等于 $U_0(|f|)$ 。 $U_1(f)$ 和 $U_0(f)$ 被分别称为余弦和正弦部分。式 (2.6-1)、式 (2.6-6) 和式 (2.6-7) 被称为分解方程, 而式 (2.6-4) 和式 (2.6-8) 对应于合成方程。

当 $f \geq 0$ 时, CFT 和 RFT 之间的关系可以表示为

$$\begin{aligned} U_c(0) &= U(0) \quad f=0 \\ \begin{pmatrix} U_c(f) \\ U_c(-f) \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & -j \\ 1 & j \end{pmatrix} \begin{pmatrix} U_1(f) \\ U_0(f) \end{pmatrix} \quad f>0 \end{aligned} \quad (2.6-9)$$

式 (2.6-9) 表明了 $U_1(f)$ 和 $U_0(f)$ 分别是偶函数与奇函数。

当 $f \geq 0$ 时, 式 (2.6-9) 的逆向形式为

$$\begin{aligned} U_1(0) &= U_c(0) \quad f=0 \\ \begin{pmatrix} U_1(f) \\ U_0(f) \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ j & -j \end{pmatrix} \begin{pmatrix} U_c(f) \\ U_c(-f) \end{pmatrix} \quad f>0 \end{aligned} \quad (2.6-10)$$

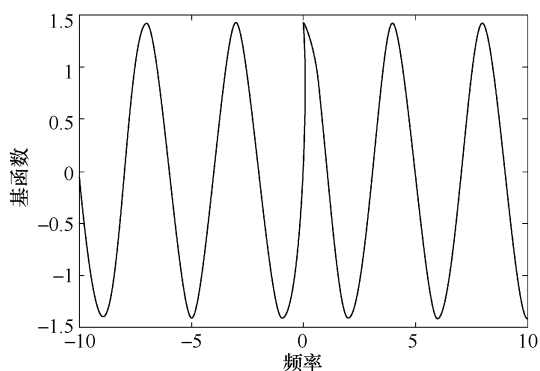


图 2.3 当 $t=0.25$ 和 $-10 < f < 10$ 时, 基函数 $\sqrt{2} \cos(2\pi ft + \theta(f))$ 的变化情况



式 (2.6-9) 和式 (2.6-10) 用来将一种表现形式转换为另一种表现形式。当 $x(t)$ 是实数时, $U_1(f)$ 和 $U_0(f)$ 同样是实数。这样, 式 (2.6-9) 表示 $U_c(f)$ 和 $U_c(-f)$ 是彼此的复共轭。

2.7 幅度谱和相位谱

$U_c(f)$ 可以表示为

$$U_c(f) = U_a(f) e^{j\phi(f)} \quad (2.7-1)$$

其中, 信号 $x(t)$ 的幅度谱 $U_a(f)$ 和相位谱 $\phi(f)$ 分别定义为

$$U_a(f) = |U_c(f)| = \frac{1}{2} [|U_1(f)|^2 + |U_0(f)|^2]^{1/2} \quad (2.7-2)$$

$$\phi(f) = \arctan \left\{ \frac{\text{Im}[U_c(f)]}{\text{Re}[U_c(f)]} \right\} \quad (2.7-3)$$

$U_a(f)$ 是一个偶函数。对于实数信号, $\phi(f)$ 是一个奇函数, 可以表示为

$$\phi(f) = \arctan [-U_0(f)/U_1(f)] \quad (2.7-4)$$

对于式 (2.3-2) 中的逆 CFT, 可以写成幅度谱和相位谱的形式

$$u(t) = \int_{-\infty}^{+\infty} U_a(f) e^{j[2\pi ft + \phi(f)]} df \quad (2.7-5)$$

当 $u(t)$ 是实数时, 式 (2.3-2) 可以简写成:

$$u(t) = 2 \int_0^{+\infty} U_a(f) \cos[2\pi ft + \phi(f)] df \quad (2.7-6)$$

因为 $U_c(f) = U_c^*(-f)$, 在公式

$$j \int_{-\infty}^{+\infty} U_a(f) \sin[2\pi ft + \phi(f)] df$$

中的被积函数是一个奇函数, 则积分值为 0。

式 (2.7-6) 将在 9.4 节中用来表示非单色光波场。

2.8 汉克尔变换

具有径向对称特征的函数在极坐标空间更加易于表示。这种情况是经常存在的, 例如, 光学中的透镜、孔径光阑等器件都是圆形的。

首先考虑极坐标系的傅里叶变换。图 2.4 表示了直角坐标和极坐标。从直角坐标到极坐标的变换表示为

$$\begin{aligned} r &= (x^2 + y^2)^{1/2} \\ \theta &= \arctan \left(\frac{y}{x} \right) \\ \rho &= (f_x^2 + f_y^2)^{1/2} \\ \phi &= \arctan \left(\frac{f_y}{f_x} \right) \end{aligned} \quad (2.8-1)$$

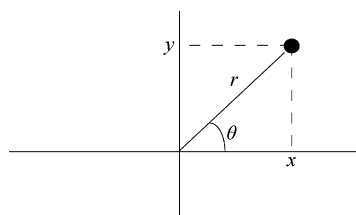


图 2.4 直角坐标和极坐标

函数 $u(x, y)$ 的傅里叶变换 (FT) 定义为

$$U(f_x, f_y) = \iint_{-\infty}^{+\infty} u(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (2.8-2)$$



$f(x, y)$ 在极坐标系中表示为 $f(r, \theta)$, $F(f_x, f_y)$ 在极坐标系中表示为 $F(\rho, \phi)$, 因此,

$$U(\rho, \phi) = \int_0^{2\pi} d\theta \int_0^{+\infty} u(r, \theta) e^{-j2\pi\rho(\cos\theta\cos\phi + \sin\theta\sin\phi)} r dr \quad (2.8-3)$$

当 $u(x, y)$ 是圆对称形时, 可以表示为

$$u(r, \theta) = u(r) \quad (2.8-4)$$

这样, 式 (2.8-3) 变成

$$U(\rho, \phi) = \int_0^{+\infty} u(r) r dr \int_0^{2\pi} e^{-j2\pi\rho r \cos(\theta - \phi)} d\theta \quad (2.8-5)$$

因此, 第一类零阶贝塞尔函数为

$$J_0(t) = \frac{1}{2\pi} \int_0^{2\pi} e^{-j t \cos(\theta - \phi)} d\theta \quad (2.8-6)$$

可以看出, 对于 ϕ 的所有值, $J_0(t)$ 都是相同的。将该等式代入式 (2.8-5), 并加入额外因子 2π , 得到

$$U(\rho, \phi) = U(\rho) = 2\pi \int_0^{+\infty} u(r) J_0(2\pi\rho r) r dr \quad (2.8-7)$$

因此, 一个圆对称 2-D 函数的傅里叶变换是自身圆对称的, 由式 (2.8-7) 表示。这种关系被称为零阶汉克尔变换, 或者傅里叶-贝塞尔变换。

零阶逆汉克尔变换可以用同样的积分形式给出

$$u(r) = 2\pi \int_0^{+\infty} U(\rho) J_0(2\pi\rho r) \rho d\rho \quad (2.8-8)$$

例 2.5 推导式 (2.8-5)。

解: x, y 和 f_x, f_y 表示为

$$\begin{aligned} x &= r \cos \theta & y &= r \sin \theta \\ f_x &= \rho \cos \phi & f_y &= \rho \sin \phi \end{aligned}$$

式 (2.8-2) 变成

$$U(\rho, \phi) = \int_0^{2\pi} d\theta \int_0^{+\infty} u(r, \theta) e^{-j2\pi\rho(\cos\theta\cos\phi + \sin\theta\sin\phi)} J dr \quad (2.8-9)$$

其中 J 为雅克比行列式, 为

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \quad (2.8-10)$$

其中 $|\cdot|$ 代表行列式。将 $J=r$ 代入式 (2.8-9), 得到结果。

例 2.6 (a) 给出如下圆柱函数 $\text{cyl}(r)$ 的汉克尔变换。

$$\text{cyl}(r) = \begin{cases} 1 & 0 \leq r < \frac{1}{2} \\ \frac{1}{2} & r = \frac{1}{2} \\ 0 & r > \frac{1}{2} \end{cases} \quad (2.8-11)$$



(b) 给出 $\text{cyl}(r/D)$ 的汉克尔变换。

解: (a) 令

$$u(r) = \text{cyl}(r)$$

这样

$$\begin{aligned} U(\rho) &= 2\pi \int_0^\infty u(r) J_0(2\pi r\rho) r dr \\ &= 2\pi \int_0^{1/2} J_0(2\pi r\rho) r dr \end{aligned}$$

第一类一阶贝塞尔函数由 [Erdelyi] 给出

$$J_1(\pi\rho) = 4\pi\rho \int_0^{1/2} J_0(2\pi r\rho) r dr \quad (2.8-12)$$

于是

$$U(\rho) = \frac{J_1(\pi\rho)}{2\rho}$$

宽边帽函数 $\text{somb}(r)$ 定义为

$$\text{somb}(r) = \frac{2J_1(\pi r)}{\pi r} \quad (2.8-13)$$

$U(\rho)$ 表示为由 $\text{somb}(r)$ 来表示

$$U(\rho) = \frac{\pi}{4} \text{somb}(\rho)$$

(b) 由于

$$U'(\rho) = 2\pi \int_0^{D/2} J_0(2\pi r\rho) r dr$$

令 $r' = r/D$, 于是 $dr = Ddr'$, 得到:

$$\begin{aligned} U'(\rho) &= 2\pi D^2 \int_0^{1/2} r' J_0(2\pi r D \rho) dr' \\ &= D^2 U(\rho) = \frac{D^2 \pi}{4} \text{somb}(D\rho) \end{aligned}$$

第 3 章 波传播基本定理

3.1 本章引论

在本章及第 4 章中，波在一般意义上被认为是存在于 3-D 空间的。然而，在某些应用中（例如集成光学），光被认为在一个 2-D 的平面上传播。例如，在第 19 章中的密集波分复用（DWDM）中，由于空间变量 y 被从方程中去掉，因此方程变成了更加简单的二维方程。于是，在以下 3-D 空间中讨论的结果可以容易地降维到 2-D 空间。

对于电磁波（EM），它由时变电场 $\boldsymbol{E}(\boldsymbol{r},t)$ 产生时变的磁场 $\boldsymbol{H}(\boldsymbol{r},t)$ 而生成，并在自由空间或空气等非波导介质、光纤等波导介质、陆地表面与电离层之间的各种介质中传播。在本章中，将主要考虑无边界介质的情况。

图 3.1a 所示的是一个天线源产生球面波的情况。当距离发射源很远时，波前上所有点都近似地具有同样的特性，球面波看起来像平面波，如图 3.1b 所示。另一个例子是指向 z 轴方向的一个电偶极子，在原点处以角频率 ω 振荡。它产生的电磁场是非常复杂的，但是距离原点很远时，该电磁波也可以视为平面波。一个完美的平面波在物理上是不存在的，但是它对各种波的分析建模则非常有用。

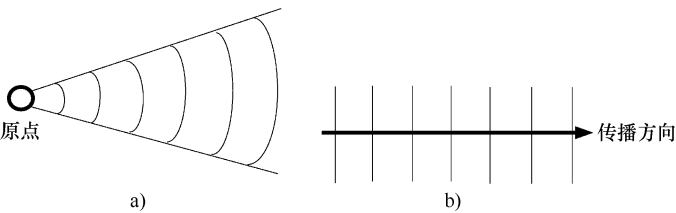


图 3.1 一个天线源产生球面波的情况

a) 由一个点源产生的球面波 b) 在传播方向上具有一致特性的平面波

但是距离原点很远时，该电磁波也可以视为平面波。一个完美的平面波在物理上是不存在的，但是它对各种波的分析建模则非常有用。

波在介质中传播，对于光波来说，光学特性由折射率 n 来表示，其定义为自由空间光传播速度与光在介质中传播速度的比值。如果 n 是常数，那么该介质是均匀的，否则该介质是非均匀的。在本章中，假设介质是均匀的。

本章包含七个部分。3.2 节介绍了波是如何产生的及波的一些基本特性。在 3.3 节中，对电磁波的基本特性及基尔霍夫方程进行了介绍。在 3.4 节中，对相量表示进行了回顾。在 3.5 节中，介绍了波动方程、在无源介质中的波动方程、基于波数与方向余弦的平面波解。3.6 节介绍了非电介质中波动方程的相量表示。平面波是电磁波的基本组成，在 3.7 节中对平面波及其偏振特性进行了详细介绍。

3.2 波

自然界有很多种波动形式，如电磁波、声波、水波及脑电波等。波可以被认为是一种以固定的速度和形式从一点到另一点传播的振动。

用 $u(x,t)$ 表示一个在均匀介质中沿 x 方向传播的一维波。如果其速度为 v ，且向右传



播, 则 $u(x, t)$ 满足

$$u(x, t) = u(x - vt, t) \quad (3.2-1)$$

相反, 若波向左传播, 则 $u(x, t)$ 满足

$$u(x, t) = u(x + vt, t) \quad (3.2-2)$$

假设波是向左传播的, s 定义为

$$s = x + vt \quad (3.2-3)$$

这样, 可以做以下计算:

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial s} \\ \frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial s^2} \\ \frac{\partial u}{\partial t} &= \frac{\partial s}{\partial t} \frac{\partial u}{\partial s} = v \frac{\partial u}{\partial s} \\ \frac{\partial^2 u}{\partial t^2} &= v^2 \frac{\partial^2 u}{\partial s^2} \end{aligned} \quad (3.2-4)$$

因此

$$\frac{\partial^2 u(x, t)}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 u(x, t)}{\partial t^2} \quad (3.2-5)$$

这个方程被称为非色散波动方程。上述方程的一个特解, 同时也是所有其他波动方程的解, 是一个简单的调和解

$$\Psi(x, t) = A \cos(k(x + vt)) = A \cos(kx + \omega t) = A \cos\left(2\pi\left(\frac{x}{\lambda} + ft\right)\right) \quad (3.2-6)$$

式中, $\omega = 2\pi f$; $k = 2\pi/\lambda$; λ 是波长; f 是时间频率; $1/\lambda$ 是空间频率; ω 和 k 是对应的角频率。

k 也被称为波数。式 (3.2-6) 给出的 $\Psi(x, t)$, 也被称之为 1-D 平面波。

为了数学上的简便性, 式 (3.2-6) 可以写成

$$u(x, t) = \text{Re}[A e^{j(kx + \omega t)}] \quad (3.2-7)$$

或者简化为

$$u(x, t) = A e^{j(kx + \omega t)} \quad (3.2-8)$$

需要再次强调的是波数 k 是 x 方向上的弧度空间频率, 也可以写成

$$k = 2\pi f_x \quad (3.2-9)$$

式中, f_x 是 x 方向上以周期/单位长度为单位的空间频率。

将式 (3.2-8) 代入式 (3.2-5), 得到

$$v = \frac{\omega}{k} \quad (3.2-10)$$

式中, v 为相位速度。

对于光波而言, 相位速度可以高于真空中的光速。

更一般地, 除式 (3.2-5) 之外, 式 (3.2-8) 给出的一维波可以满足另一个对 x 和 t 的偏微分方程。一个具有简单调和解的一般波动方程可以写为

$$f\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial t}\right)u(x, t) = 0 \quad (3.2-11)$$



其中, f 是 $\frac{\partial}{\partial x}$ 和 $\frac{\partial}{\partial t}$ 的多项式函数, 例如

$$\frac{\partial^3 u}{\partial x^3} + \alpha \frac{\partial^3 u}{\partial x^2 \partial t} - \beta \frac{\partial u}{\partial t} = 0 \quad (3.2-12)$$

利用式 (3.2-8) 获得上述简单单调和解, 用 $\frac{\partial}{\partial t} \rightarrow j\omega$, 式 (3.2-12) 变成

$$k^3 - \alpha k^2 \omega - \beta \omega = 0 \quad (3.2-13)$$

这个关于 ω 和 k 的方程被称为色散方程。在这种情况下, 相位速度 v 表示为

$$v = \frac{\omega}{k} = \frac{k^2}{\alpha k^2 + \beta} \quad (3.2-14)$$

不同于式 (3.2-10), 一个具有相位速度的波, 且其相位速度是波数和波长的函数, 这样的波被称为色散波。这样, 式 (3.2-14) 即为色散波动方程。

下面是一些波的例子:

A. 机械波, 如在可压缩流体中的纵向声波, 其波动方程为

$$\frac{\partial^2 u(x, t)}{\partial x^2} = \frac{\rho}{K_c} \frac{\partial^2 u(x, t)}{\partial t^2} \quad (3.2-15)$$

式中, ρ 是流体密度; K_c 是可压缩性参数。

因此其相位速度为

$$v = \sqrt{\frac{K_c}{\rho}} \quad (3.2-16)$$

B. 在稳态条件下的热扩散, 根据波动方程

$$\frac{K_h}{s} \frac{\partial^2 u(x, t)}{\partial x^2} = \frac{\partial u(x, t)}{\partial t} \quad (3.2-17)$$

式中, K_h 是传热系数; s 是单位体积的热容; $u(x, t)$ 是局部温度。

C. 量子力学中的薛定谔波动方程为

$$-\frac{h^2}{4\pi^2 m} \frac{\partial^2 u(x, t)}{\partial x^2} + V(x)u(x, t) = -j \frac{h}{2\pi} \frac{\partial u(x, t)}{\partial t} \quad (3.2-18)$$

式中, m 是质量; $V(x)$ 是粒子的势能; h 是普朗克常数; $u(x, t)$ 是粒子的态。

色散关系可以表示为

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} + V(x) \quad (3.2-19)$$

其中, \hbar 等于 $h/2\pi$ 。

D. 随时间变化的成像可以看成是一个三维函数 $g(x, y, t)$, 它的三维傅里叶变换表示可以写成

$$g(x, y, t) = \iiint_{-\infty}^{+\infty} G(f_x, f_y, f) e^{j2\pi(f_x x + f_y y + ft)} df_x df_y df \quad (3.2-20)$$

其中

$$\begin{aligned} G(f_x, f_y, f) &= \iiint_{-\infty}^{+\infty} g(x, y, t) e^{-j2\pi(f_x x + f_y y + ft)} dx dy dt \\ &= A(f_x, f_y, f) e^{j\Theta(f_x, f_y, f)} \end{aligned} \quad (3.2-21)$$



对于实数 $g(x, y, t)$, $G(f_x, f_y, f) = G^*(-f_x, -f_y, -f)$, 所以式 (3.2-21) 可以写成

$$g(x, y, t) = 2 \int_0^{+\infty} \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(f_x, f_y, f) \cos(2\pi(f_x x + f_y y + ft)) + \Theta(f_x, f_y, f) df_x df_y \right] df \quad (3.2-22)$$

$A(f_x, f_y, f) \cos(2\pi(f_x x + f_y y + ft)) + \Theta(f_x, f_y, f)$ 是一个二维的平面波, 其幅值为 $A(f_x, f_y, f)$, 相位为 $\Theta(f_x, f_y, f)$, 空间频率为 f_x 和 f_y , 时间频率为 f , 这样, 波矢量 \mathbf{k} 表示为

$$\mathbf{k} = 2\pi(f_x \mathbf{e}_x + f_y \mathbf{e}_y) = k_x \mathbf{e}_x + k_y \mathbf{e}_y \quad (3.2-23)$$

式中, \mathbf{e}_x 和 \mathbf{e}_y 分别是 x 方向和 y 方向上的单位向量; k_x 和 k_y 分别是 x 方向和 y 方向的弧度空间频率; f_x 和 f_y 是对应的空间频率。

二维平面波可以写成

$$u(\mathbf{r}, t) = A(\mathbf{k}, \omega) \cos(\mathbf{k} \cdot \mathbf{r} + \omega t + \Theta(\mathbf{k}, \omega)) \quad (3.2-24)$$

其中

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y \quad (3.2-25)$$

如果假设 $u(\mathbf{r}, t)$ 以速度 v 在 \mathbf{k} 方向上运动, $u(\mathbf{r}, t)$ 可以作为二维非色散波动方程的一个解

$$\nabla^2 u(\mathbf{r}, t) = \frac{1}{v^2} \frac{\partial^2 u(\mathbf{r}, t)}{\partial t^2} \quad (3.2-26)$$

其中

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (3.2-27)$$

综上所述, 随时间变化的成像可以表示为一系列二维平面波的叠加。这个结论可以被推广到任意的多维信号, 其中的一个维度被设定为时间。

式 (3.2-26) 是式 (3.2-5) 在两个空间轴 x 和 y 上的推广, 如果用同样的波的特性考虑三维空间轴 x 、 y 和 z , 式 (3.2-23) 和式 (3.2-25) 保持不变, 其中 \mathbf{r} 和 ∇^2 定义如下:

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z \quad (3.2-28)$$

其中, \mathbf{e}_z 是增加的 z 方向上的单位向量, 并且

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (3.2-29)$$

运算符 ∇^2 为拉普拉斯算子。式 (3.2-24) 中的三维波矢量为

$$\mathbf{k} = 2\pi(f_x \mathbf{e}_x + f_y \mathbf{e}_y + f_z \mathbf{e}_z) = k_x \mathbf{e}_x + k_y \mathbf{e}_y + k_z \mathbf{e}_z \quad (3.2-30)$$

式中, k_z 是增加的 z 方向上的弧度空间频率; f_z 是对应的空间频率。

使 $u(\mathbf{r}, t)$ 为常数的面称为波前, 它沿着 \mathbf{k} 方向以速度 v 运动, 在 3.5 小节中, 将会深入地讨论波前理论。

接下来, 除非在有特别说明的情况下, 后面将假设介质是线性、均匀的及各向同性的。同样假设采用十进制单位。

例 3.1 考虑一个向右传播的波 $u(x_1, t_1) = u(x_2, t_2)$, 说明 x_1 和 x_2 需满足何种关系。

解: $u(x_1, t_1)$ 可以被设定为以速度 v 向右运动, 令 $\Delta t = t_2 - t_1$, 当

$$x_2 - x_1 = v\Delta t$$

时, $u(x_1, t_1)$ 和 $u(x_2, t_2)$ 将是相同的。



3.3 电磁波

电磁波是三维波，具有三个空间维度和一个时间维度，有 \mathbf{D} 、 \mathbf{B} 、 \mathbf{H} 和 \mathbf{E} 四个空间坐标矢量，是 x 、 y 、 z 和 t 的函数，即 \mathbf{D} 、 \mathbf{B} 、 \mathbf{H} 和 \mathbf{E} 在空间上具有方向性。 \mathbf{D} 是电位移矢量，单位为 C/m^2 ； \mathbf{E} 是电场强度，单位为 V/m ； \mathbf{B} 是磁感应强度，单位为 Wb/m^2 ； \mathbf{H} 是磁场强度，单位为 A/m ； ρ 是电荷密度，单位为 C/m^3 ； \mathbf{J} 是电流密度，单位为 A/m^2 。

电磁波满足麦克斯韦方程，在公制单位下，其形式为

$$\nabla \cdot \mathbf{D} = \rho \quad (3.3-1)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (3.3-2)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad (3.3-3)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (3.3-4)$$

矢量 \mathbf{A} 包含成分 A_x 、 A_y 和 A_z ，其散度和旋度为

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \quad (3.3-5)$$

$$\begin{aligned} \nabla \times \mathbf{A} &= \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \\ &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \mathbf{e}_x + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \mathbf{e}_y + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \mathbf{e}_z \end{aligned} \quad (3.3-6)$$

另外还有

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (3.3-7)$$

$$\mathbf{B} = \mu \mathbf{H} \quad (3.3-8)$$

式中， ε 是介电常数； μ 是磁导率。

在自由空间或者真空，它们的值为

$$\varepsilon_0 = \frac{10^{-9}}{36\pi} \text{F/m} \quad (\text{注 } \varepsilon_0 = 8.854188 \times 10^{-12} \text{F/m})$$

$$\mu_0 = 4\pi \times 10^{-7} \text{H/m}$$

在电介质中， ε 和 μ 的值分别都比 ε_0 和 μ_0 大。在这样的介质中，我们同样考虑电偶极矩密度 \mathbf{P} （单位为 C/m^2 ），其与电场强度 \mathbf{E} 的关系为

$$\mathbf{P} = \chi \varepsilon_0 \mathbf{E} \quad (3.3-9)$$

式中， χ 称为电极化率，电极化率可以用来评价介质中电偶极子沿电场排列的能力。

同样，有

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 (1 + \chi) \mathbf{E} = \varepsilon \mathbf{E} \quad (3.3-10)$$

其中

$$\varepsilon = \varepsilon_0 (1 + \chi) \quad (3.3-11)$$



在均匀的各向同性介质中，空间电荷 ρ 和电流密度 \mathbf{J} 为零，麦克斯韦方程变为

$$\nabla \cdot \mathbf{D} = \varepsilon \nabla \cdot \mathbf{E} = 0 \quad (3.3-12)$$

$$\nabla \cdot \mathbf{B} = \mu \nabla \cdot \mathbf{H} = 0 \quad (3.3-13)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} \quad (3.3-14)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t} \quad (3.3-15)$$

3.4 相量表示法

我们考虑的电场和磁场，通常用随时间变化的正弦函数来表示

$$u(\mathbf{r}, t) = A(\mathbf{r}) \cos(\mathbf{k} \cdot \mathbf{r} \pm \omega t) \quad (3.4-1)$$

可以将 $u(\mathbf{r}, t)$ 表示为

$$u(\mathbf{r}, t) = \operatorname{Re}[A'(\mathbf{r}) e^{j\omega t}] \quad (3.4-2)$$

其中

$$A'(\mathbf{r}) = A(\mathbf{r}) e^{j\mathbf{k} \cdot \mathbf{r}} \quad (3.4-3)$$

式中， $A'(\mathbf{r})$ 为 $u(\mathbf{r}, t)$ 的相量表示，与时间无关。

我们发现：

$$\frac{d}{dt} A'(\mathbf{r}) e^{j\omega t} = j\omega A'(\mathbf{r}) e^{j\omega t} \quad (3.4-4)$$

$$\int A'(\mathbf{r}) e^{j\omega t} dt = \frac{1}{j\omega} A'(\mathbf{r}) e^{j\omega t} \quad (3.4-5)$$

这样，微分和积分分别等价于 $A'(\mathbf{r})$ 与 $j\omega$ 和 $\frac{1}{j\omega}$ 相乘。

电场和磁场的相量表示为

$$\mathbf{E}(\mathbf{r}, t) = \operatorname{Re}[\tilde{\mathbf{E}}(\mathbf{r}) e^{j\omega t}] \quad (3.4-6)$$

$$\mathbf{H}(\mathbf{r}, t) = \operatorname{Re}[\tilde{\mathbf{H}}(\mathbf{r}) e^{j\omega t}] \quad (3.4-7)$$

式中， $\tilde{\mathbf{E}}(\mathbf{r})$ 和 $\tilde{\mathbf{H}}(\mathbf{r})$ 是相量。

对于 \mathbf{D} 、 \mathbf{B} 和 \mathbf{J} ，其对应的相量采用同样的方法来定义。

以 $\tilde{\mathbf{E}}(\mathbf{r})$ 和 $\tilde{\mathbf{H}}(\mathbf{r})$ 来表示麦克斯韦方程为

$$\nabla \cdot \tilde{\mathbf{D}} = \rho \quad (3.4-8)$$

$$\nabla \cdot \tilde{\mathbf{B}} = 0 \quad (3.4-9)$$

$$\nabla \times \tilde{\mathbf{H}} = j\omega \varepsilon \tilde{\mathbf{E}} + \tilde{\mathbf{J}} \quad (3.4-10)$$

$$\nabla \times \tilde{\mathbf{E}} = -j\omega \tilde{\mathbf{B}} \quad (3.4-11)$$

3.5 非电介质中的波动方程

对式 (3.3-3) 两边取旋度，得到



$$\nabla \times \nabla \times \mathbf{H} = \varepsilon \frac{\partial}{\partial t} \nabla \times \mathbf{E} = -\varepsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (3.5-1)$$

$\nabla \times \nabla \times \mathbf{H}$ 可以被展开为

$$\nabla \times \nabla \times \mathbf{H} = \nabla(\nabla \cdot \mathbf{H}) - \nabla \cdot \nabla \mathbf{H} \quad (3.5-2)$$

式中, $\nabla \mathbf{H}$ 是 \mathbf{H} 的梯度。

这意味着, $\nabla H_i (i=x, y, z)$ 为

$$\nabla H_i = \frac{\partial H_i}{\partial x} \mathbf{e}_x + \frac{\partial H_i}{\partial y} \mathbf{e}_y + \frac{\partial H_i}{\partial z} \mathbf{e}_z \quad (3.5-3)$$

$\nabla \cdot \nabla H_i = \nabla^2 H_i$ 由下式给出

$$\nabla^2 H_i = \frac{\partial^2 H_i}{\partial x^2} \mathbf{e}_x + \frac{\partial^2 H_i}{\partial y^2} \mathbf{e}_y + \frac{\partial^2 H_i}{\partial z^2} \mathbf{e}_z \quad (3.5-4)$$

这样, $\nabla \cdot \nabla \mathbf{H}$ 沿三个方向的分量为 $\nabla^2 H_i (i=x, y, z)$ 的矢量。

可以看到, $\nabla(\nabla \cdot \mathbf{H})$ 是 $\nabla \cdot \mathbf{H}$ 的梯度矢量, 等于零。这样,

$$\nabla \times \nabla \times \mathbf{H} = -\nabla \cdot \nabla \mathbf{H} = -\nabla^2 \mathbf{H} \quad (3.5-5)$$

将该结果代入式 (3.5-1), 得到

$$\nabla^2 \mathbf{H} = \varepsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (3.5-6)$$

同样地, 可以得到

$$\nabla^2 \mathbf{E} = \varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (3.5-7)$$

式 (3.5-6) 和式 (3.5-7) 分别称为 \mathbf{E} 和 \mathbf{H} 的齐次波动方程。

综上所述, 电场和磁场矢量的每一个分量满足非色散波动方程, 其相量速度 v 等于 $1/\sqrt{\varepsilon\mu}$, 在自由空间

$$c = v = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \cong 3 \times 10^8 \text{ m/s} \quad (3.5-8)$$

下面我们考虑这样一个场 $u(\mathbf{r}, t)$, 它满足

$$\nabla^2 u(\mathbf{r}, t) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) u(\mathbf{r}, t) = \frac{1}{c^2} \frac{\partial^2 u(\mathbf{r}, t)}{\partial t^2} \quad (3.5-9)$$

这个方程的一个三维的平面波解为

$$u(\mathbf{r}, t) = A(\mathbf{k}, \omega) \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \quad (3.5-10)$$

其中的波矢量 \mathbf{k} 给定为

$$\mathbf{k} = k_x \mathbf{e}_x + k_y \mathbf{e}_y + k_z \mathbf{e}_z \quad (3.5-11)$$

相位速度 v 与 \mathbf{k} 和 ω 的关系如下:

$$v = \frac{\omega}{|\mathbf{k}|} \quad (3.5-12)$$

波前定义为

$$\mathbf{k} \cdot \mathbf{r} - \omega t = \text{常量} \quad (3.5-13)$$

这是一个平面, 其法向为矢量 \mathbf{k} 的方向。当 ωt 改变时, 这个平面发生变化, 波沿着 \mathbf{k} 方向以速度 c 传播。需要指出 $\cos(\mathbf{k} \cdot \mathbf{r} - \omega t)$ 也可以被选择为 $\cos(\mathbf{k} \cdot \mathbf{r} + \omega t)$, 这时波的传播方



向为 $-\mathbf{k}$ 。

波矢量 \mathbf{k} 的分量可以写成

$$k_i = \frac{2\pi}{\lambda} \alpha_i = k \alpha_i \quad i = x, y, z \quad (3.5-14)$$

式中, $k = |\mathbf{k}|$; α_i 是第 i 个方向的方向余弦。

如果用 f_i 表示沿着第 i 个方向的空间频率, 其等于 $k_i/2\pi$, 这样方向余弦可以写为

$$\alpha_x = \lambda f_x \quad (3.5-15)$$

$$\alpha_y = \lambda f_y \quad (3.5-16)$$

$$\alpha_z = \sqrt{1 - \lambda^2 f_x^2 - \lambda^2 f_y^2} \quad (3.5-17)$$

且

$$\alpha_x^2 + \alpha_y^2 + \alpha_z^2 = 1 \quad (3.5-18)$$

3.6 非电介质中的波动方程的相量表示

首先, 计算式 (3.4-11) 的旋度

$$\nabla \times \nabla \times \tilde{\mathbf{E}} = -j\omega \tilde{\mathbf{H}} \quad (3.6-1)$$

其中, $\tilde{\mathbf{B}} = \mu \tilde{\mathbf{H}}$, 利用式 (3.4-10) 中的 $\tilde{\mathbf{J}} = 0$ 及式 (3.6-1) 获得

$$\nabla \times \nabla \times \tilde{\mathbf{E}} = \omega^2 \mu \epsilon \tilde{\mathbf{E}} \quad (3.6-2)$$

在 3.4 节中, 知道

$$\nabla \times \nabla \times \tilde{\mathbf{E}} = \nabla(\nabla \cdot \tilde{\mathbf{E}}) - \nabla \cdot \nabla \tilde{\mathbf{E}} \quad (3.6-3)$$

根据式 (3.4-8) 中 $\nabla \cdot \tilde{\mathbf{E}} = 0$, 得到

$$\nabla^2 \tilde{\mathbf{E}} + \omega^2 \mu \epsilon \tilde{\mathbf{E}} = 0 \quad (3.6-4)$$

定义波数 k 为

$$k = \omega \sqrt{\mu \epsilon} \quad (3.6-5)$$

这样, 式 (3.6-4) 变成

$$\nabla^2 \tilde{\mathbf{E}} + k^2 \tilde{\mathbf{E}} = 0 \quad (3.6-6)$$

这被称为 $\tilde{\mathbf{E}}$ 的齐次波动方程。 $\tilde{\mathbf{H}}$ 的齐次波动方程可以采用相同的方法从下式中获得

$$\nabla^2 \tilde{\mathbf{H}} + k^2 \tilde{\mathbf{H}} = 0 \quad (3.6-7)$$

将 $\tilde{\mathbf{E}}$ 表示成

$$\tilde{\mathbf{E}} = \tilde{E}_x \mathbf{e}_x + \tilde{E}_y \mathbf{e}_y + \tilde{E}_z \mathbf{e}_z \quad (3.6-8)$$

现在, 式 (3.6-6) 可以写成关于 $\tilde{\mathbf{E}}$ 中分量 \tilde{E}_i 的形式

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2 \right) \tilde{E}_i = 0 \quad (3.6-9)$$



例 3.2 (a) 对于一个沿 z 方向传播的均匀平面波, 简化式 (3.6-9)。(b) 采用相量表示法, 说明一个均匀平面波中 $\tilde{\mathbf{E}}$ 和 $\tilde{\mathbf{H}}$ 的 z 分量为 0。(c) 采用麦克斯韦方程说明 (b)。
解: (a) 一个均匀平面波的特征为

$$\frac{\delta \tilde{E}_i}{\delta x} = \frac{\delta \tilde{E}_i}{\delta y} = \frac{\delta \tilde{H}_i}{\delta x} = \frac{\delta \tilde{H}_i}{\delta y} = 0$$

这样, 式 (3.6-9) 可以简化为

$$\left(\frac{\partial^2}{\partial z^2} + k^2 \right) \tilde{E}_i = 0$$

(b) 考虑式 (3.4-10) 中的 z 分量

$$\frac{\delta \tilde{H}_y}{\delta x} - \frac{\delta \tilde{H}_x}{\delta y} = j\omega\epsilon \tilde{E}_z$$

由于 $\frac{\delta \tilde{H}_y}{\delta x} = \frac{\delta \tilde{H}_x}{\delta y} = 0$, $\tilde{E}_z = 0$, 同样能够用式 (3.4-11) 得到 $\tilde{H}_z = 0$ 。

(c) 有

$$\mathbf{E} = E_x e^{j(kz + \omega t)} \mathbf{e}_x + E_y e^{j(kz + \omega t)} \mathbf{e}_y + E_z e^{j(kz + \omega t)} \mathbf{e}_z$$

取代 $\nabla \cdot \mathbf{E} = 0$ 中的 \mathbf{E} , 得到

$$\frac{\partial}{\partial z} E_z e^{j(kz + \omega t)} = 0$$

这表明 $E_z = 0$ 。可以用同样的方法考察磁场 \mathbf{H}

$$\mathbf{H} = H_x e^{j(kz + \omega t)} \mathbf{e}_x + H_y e^{j(kz + \omega t)} \mathbf{e}_y + H_z e^{j(kz + \omega t)} \mathbf{e}_z$$

取代 $\nabla \cdot \mathbf{H} = 0$ 中的 \mathbf{H} , 得到

$$\frac{\partial}{\partial z} H_z e^{j(kz + \omega t)} = 0$$

这表明 $H_z = 0$ 。

3.7 平面电磁波

考虑一束沿 z 方向传播的平面波, 电磁场可以表示为

$$\begin{aligned} \mathbf{E} &= E_x e^{j(kz + \omega t)} \hat{\mathbf{e}}_x + E_y e^{j(kz + \omega t)} \hat{\mathbf{e}}_y \\ \mathbf{H} &= H_x e^{j(kz + \omega t)} \hat{\mathbf{e}}_x + H_y e^{j(kz + \omega t)} \hat{\mathbf{e}}_y \end{aligned} \quad (3.7-1)$$

其中的实数部分是真正意义上的物理解。它们可以写成更一般的形式

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \mathbf{E}_0 \cos(\mathbf{k} \cdot \mathbf{r} + \omega t) \\ \mathbf{H}(\mathbf{r}, t) &= \mathbf{H}_0 \cos(\mathbf{k} \cdot \mathbf{r} + \omega t) \end{aligned} \quad (3.7-2)$$

其中, \mathbf{E}_0 和 \mathbf{H}_0 具有分量 (E_x, E_y) , (H_x, H_y) 。此时, \mathbf{k} , \mathbf{r} 分别是 $k = 2\pi/\lambda$ 和 z 。

将式 (3.7-1) 代入 $\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}$, 得到



$$kE_y \hat{e}_x - kE_x \hat{e}_y = -\mu\omega(H_x \hat{e}_x + H_y \hat{e}_y) \quad (3.7-3)$$

因此

$$\begin{aligned} H_x &= -\frac{1}{\eta} E_y \\ H_y &= \frac{1}{\eta} E_x \end{aligned} \quad (3.7-4)$$

式中, η 为介质的特征阻抗。

η 定义为

$$\eta = \frac{w}{k} \mu = v \mu = \sqrt{\mu/\varepsilon} \quad (3.7-5)$$

知道

$$\mathbf{E} \cdot \mathbf{H} = (E_x H_x + E_y H_y) e^{j(kz + \omega t)} = 0 \quad (3.7-6)$$

因此, 电场和磁场之间是相互正交的。

坡印廷矢量 \mathbf{S} 定义为

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \quad (3.7-7)$$

其单位为 W/m^2 , 代表传播方向上每单位面积的功率。

偏振表示电场强度矢量如何随时间变化。再次假设波沿 z 方向传播, 包含与时间相关的电场强度矢量可以写成

$$\mathbf{E} = \text{Re}[(E_x \mathbf{e}_x + E_y \mathbf{e}_y) e^{j(kz + \omega t)}] \quad (3.7-8)$$

E_x 和 E_y 可以相对于彼此选择

$$E_x = E_{x_0} \quad (3.7-9)$$

$$E_y = E_{y_0} e^{j\phi} \quad (3.7-10)$$

式中, E_{x_0} 和 E_{y_0} 是正的标量; ϕ 是相对相位, 决定电场强度的方向。

当 $\phi = 0$ 或者 π 时, 为线偏振, 得到

$$\mathbf{E} = (E_{x_0} \mathbf{e}_x \pm E_{y_0} \mathbf{e}_y) \cos(kz + \omega t) \quad (3.7-11)$$

此时, $E_{x_0} \mathbf{e}_x \pm E_{y_0} \mathbf{e}_y$ 被认为是一个方向不随时间和传播距离改变的矢量。

当 $\phi = \pm \pi/2$, 且 $E_0 = E_{x_0} = E_{y_0}$ 时, 为圆偏振, 得到

$$\mathbf{E} = E_0 \cos(kz + \omega t) \mathbf{e}_x \pm E_0 \sin(kz + \omega t) \mathbf{e}_y \quad (3.7-12)$$

我们注意到 $|\mathbf{E}| = E_0$ 。对于 $\phi = -\pi/2$, \mathbf{E} 为在传播过程中顺时针旋转的圆; 对于 $\phi = +\pi/2$, \mathbf{E} 为在传播过程中逆时针旋转的圆。

椭圆偏振对应于任意的 ϕ , 对于电场强度, 式 (3.7-1) 可以写成

$$\mathbf{E} = E_x \mathbf{e}_x + E_y \mathbf{e}_y \quad (3.7-13)$$

其中

$$\begin{aligned} E_x &= E_{x_0} \cos(kz + \omega t) \\ E_y &= E_{y_0} \cos(kz + \omega t - \phi) \end{aligned} \quad (3.7-14)$$

式 (3.7-14) 可以写成

$$\begin{aligned} \frac{E_y}{E_{y_0}} &= \cos(kz + \omega t) \cos\phi + \sin(kz + \omega t) \sin\phi \\ &= \frac{E_x}{E_{x_0}} \cos\phi + \left[1 - \left(\frac{E_x}{E_{x_0}}\right)^2\right] \sin\phi \end{aligned} \quad (3.7-15)$$



进一步，可以得到

$$E_x'^2 - 2E_x'E_y'\cos\phi + E_y'^2 = \sin^2\phi \quad (3.7-16)$$

其中

$$E_x' = \frac{E_x}{E_{x0}}, \quad E_y' = \frac{E_y}{E_{y0}} \quad (3.7-17)$$

式 (3.7-17) 是椭圆方程。

假设场是线偏振的，同时 \mathbf{E} 沿 x 方向传播，有

$$\begin{aligned} \mathbf{E} &= E_0 e^{j(kz + \omega t)} \mathbf{e}_x \\ \mathbf{H} &= -\frac{E_0}{\eta} e^{j(kz + \omega t)} \mathbf{e}_y \end{aligned} \quad (3.7-18)$$

其中， $\text{Re}[\cdot]$ 是根据情况假定的。强度或者辐照度 I 定义为能量的时间平均值

$$I = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} |\mathbf{S}| dt = \varepsilon\mu \frac{E_0^2}{2} \quad (3.7-19)$$

式中， \mathbf{S} 是坡印廷矢量。

可以看到，亮度正比于电场强度的二次方。除非特殊声明，该定理在一般情况下是成立的。

第 4 章 标量衍射理论

4.1 本章引论

当光波的波长小于控制光场的衍射器件的孔径尺寸时，可以应用**标量衍射理论**。即使该条件不满足，标量衍射理论在一定程度上也是十分准确的 [Mellin and Nordin, 2001]。标量衍射理论涉及波动方程从偏微分方程到积分方程的转化，在其有效范围内，标量衍射理论可以用来分析大多数的衍射现象和成像系统。例如，图 4.1 显示了采用单色平面波照明的双缝衍射条纹，其波的传播可以采用标量衍射理论精确地描述。

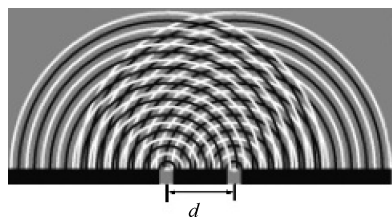


图 4.1 双缝衍射条纹

在本章中，将首先引入具有单一波长的单色波的标量衍射理论，然后，在时间维度应用傅里叶分析和合成方法，将其推广到非单色波场。

本章包含八个小节。在 4.2 节中将介绍亥姆霍兹方程，该方程将时间变化作为复指数因子。在 4.3 节中，采用平面波角谱的形式给出了均匀介质中亥姆霍兹方程的解，该方程也将均匀介质中的波动方程视为线性系统。在 4.4 节中讨论了平面波角谱的快速傅里叶变换 (FFT)。

衍射同样可以先通过亥姆霍兹方程，然后通过格林定理将其转换到积分方程来进行研究。本章其余部分讨论这个问题。在 4.5 节中，这种方法中衍射的基尔霍夫理论被描述成一个公式。4.6 节介绍了衍射中的瑞利-索末菲理论，而 4.7 节是该理论的另一种形式。在 4.8 节中给出了非单色波衍射的瑞利-索末菲理论。

4.2 亥姆霍兹方程

单色波具有单一的时间频率 f ，由于

$$k = \frac{\omega}{c} = \frac{2\pi f}{c} \quad (4.2-1)$$

波长 λ 同样是确定的。对于这样的波，在 3.7 节中讨论的平面波解可以写成

$$u(\mathbf{r}, t) = A(\mathbf{r}) \cos(\omega t + \Theta(\mathbf{r})) \quad (4.2-2)$$

式中， $A(\mathbf{r})$ 是振幅； $\Theta(\mathbf{r})$ 是在 \mathbf{r} 处的相位。

在相量表示中，它变成了

$$u(\mathbf{r}, t) = \text{Re}[U(\mathbf{r}) e^{j\omega t}] \quad (4.2-3)$$

其中，相量 $U(\mathbf{r})$ 同样被称为**复振幅**，等价于 $A(\mathbf{r}) e^{j\Theta(\mathbf{r})}$ 。在下面的方程中，由于书写和计算简单的考虑，式 (4.2-3) 不会被强调是“实数部分”。



在波动方程式 (3.5-7) 中代入 $u(\mathbf{r}, t)$, 得到

$$(\nabla^2 + k^2)U(\mathbf{r}) = 0 \quad (4.2-4)$$

其中, $k = \omega/c$ 。该方程被称为亥姆霍兹方程。它与式 (3.6-9) 是相同的。因此, 正如第 3 章所讨论的那样, 对于平面电磁波而言, $U(\mathbf{r})$ 代表了电场或磁场相矢量中的一个分量。亥姆霍兹方程对于满足非色散波动方程的所有波都是成立的。例如, 对于声波或者超声波, $U(\mathbf{r})$ 是压强或者电势。

如果 $U(\mathbf{r}, t)$ 是非单色的, 它可以表示为时间傅里叶变换的形式

$$U(\mathbf{r}, t) = \int_{-\infty}^{+\infty} U_f(\mathbf{r}, f) e^{j2\pi ft} df \quad (4.2-5)$$

其中

$$U_f(\mathbf{r}, f) = \int_{-\infty}^{+\infty} U(\mathbf{r}, t) e^{-j2\pi ft} dt \quad (4.2-6)$$

在波动方程式 (3.5-7) 中代入式 (4.2-6), 同样得到 $U_f(\mathbf{r}, f)$ 的亥姆霍兹方程。

例 4.1 对于一个球面波, 定义为 $e^{\pm jkr}/r$, 其中 $r = \sqrt{x^2 + y^2 + z^2}$, 证明其为亥姆霍兹方程的一个解。

解: 球面坐标中 ∇^2 给定为

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2}$$

$e^{\pm jkr}/r$ 是球面对称的。因此, 极坐标中的 ∇^2 简化为

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right)$$

球面对称波动函数 $U(r)$ 的亥姆霍兹方程变成

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) U(r) + k^2 U(r) = 0$$

将 $U(r)$ 表示为 $f(r)/r$, 微分方程为

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \frac{f(r)}{r} = \frac{1}{r^2} \frac{\partial^2 f(r)}{\partial r^2}$$

亥姆霍兹方程变成了

$$\frac{1}{r^2} \frac{\partial^2 f(r)}{\partial r^2} + k^2 \frac{f(r)}{r} = 0$$

方程的解为

$$f(r) = e^{\pm jkr}$$

因此, $U(r) = e^{\pm jkr}/r$ 。

例 4.2 在柱面坐标下找到一个相似的解。

解: 在柱面坐标下, ∇^2 为

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2}$$

假设一个解具有圆柱对称性, $U(r, \phi, z)$ 变成了 $U(r, z)$ 。圆柱对称意味着对于常量 z , 波前是圆。这样, 方程没有简单解。精确解具有基于 r 的贝塞尔函数形式。表示为



$$U(r, z) \cong \frac{C}{\sqrt{r}} e^{\pm jkr}$$

近似满足于波动方程，这样的波称为柱面波。

4.3 平面波角谱

考虑沿着 z 方向传播的波 $U(x, y, z)$ ，假设波的波长为 λ ，则 $k = 2\pi/\lambda$ ，假设初始时 $z = 0$ 。 $U(x, y, 0)$ 的 2-D 傅里叶表示为它的傅里叶变换形式 $A(f_x, f_y, 0)$

$$U(x, y, 0) = \iint_{-\infty}^{+\infty} A(f_x, f_y, 0) e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (4.3-1)$$

其中

$$A(f_x, f_y, 0) = \iint_{-\infty}^{+\infty} U(x, y, 0) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (4.3-2)$$

包含时间变化， $A(f_x, f_y, 0) e^{j2\pi(f_x x + f_y y + ft)}$ 为在 $z = 0$ 处的平面波，并且沿着由式 (3.5-15) ~ 式 (3.5-17) 给出的方向余弦 α_i 传播。 $A(f_x, f_y, 0)$ 被称为 $U(x, y, 0)$ 的角谱。

下面考虑波场 $U(x, y, z)$ ，它的角谱 $A(f_x, f_y, z)$ 给定为

$$A(f_x, f_y, z) = \iint_{-\infty}^{+\infty} U(x, y, z) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (4.3-3)$$

其角谱形式的傅里叶表示为

$$U(x, y, z) = \iint_{-\infty}^{+\infty} A(f_x, f_y, z) e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (4.3-4)$$

在所有无源点上， $U(x, y, z)$ 满足亥姆霍兹方程，也就是

$$\nabla^2 U(x, y, z) + k^2 U(x, y, z) = 0 \quad (4.3-5)$$

将式 (4.3-4) 中的 $U(x, y, z)$ 代入式 (4.3-5)，得到

$$\iint_{-\infty}^{+\infty} \left[\frac{d^2}{dz^2} A(f_x, f_y, z) + (k^2 - 4\pi^2(f_x^2 + f_y^2)) A(f_x, f_y, z) \right] e^{j2\pi(f_x x + f_y y)} df_x df_y = 0 \quad (4.3-6)$$

仅当积分函数为 0 时，式 (4.3-6) 对所有波都满足，即

$$\frac{d^2}{dz^2} A(f_x, f_y, z) + (k^2 - 4\pi^2(f_x^2 + f_y^2)) A(f_x, f_y, z) = 0 \quad (4.3-7)$$

该微分方程式具有如下解：

$$A(f_x, f_y, z) = A(f_x, f_y, 0) e^{j\mu z} \quad (4.3-8)$$

其中

$$\mu = \sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)} = k_z \quad (4.3-9)$$

如果 $4\pi^2(f_x^2 + f_y^2) < k^2$ ，则 μ 是实数，并且每一个角谱成分仅仅被相位因子 $e^{j\mu z}$ 调制，那么满足这种条件的平面波成分称之为均匀波。

如果 $4\pi^2(f_x^2 + f_y^2) > k^2$ ， μ 可以写成

$$\mu = j \sqrt{4\pi^2(f_x^2 + f_y^2) - k^2} \quad (4.3-10)$$

式 (4.3-8) 变成了

$$A(f_x, f_y, z) = A(f_x, f_y, 0) e^{-\mu z} \quad (4.3-11)$$

这个结果显示这些平面波成分的振幅在沿着 z 方向传播的过程中强烈地减弱，这种波称



为隐失波。

如果 $4\pi^2(f_x^2 + f_y^2) = k^2$, $A(f_x, f_y, z)$ 与 $A(f_x, f_y, 0)$ 是相同的, 那么这些成分对应于传播方向与 z 垂直的平面波。

在式 (4.3-4) 中代入式 (4.3-11), 可以得到以 $A(f_x, f_y, 0)$ 表示的 $A(f_x, f_y, z)$, 从而获得在 (x, y, z) 处的波场

$$U(x, y, z) = \iint_{-\infty}^{+\infty} A(f_x, f_y, 0) e^{iz\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}} e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (4.3-12)$$

这样, 如果 $U(x, y, 0)$ 是已知的, 可以计算 $A(f_x, f_y, 0)$, 然后计算 $U(x, y, z)$ 。式 (4.3-12) 中的积分限制在一个圆形区域, 给定为

$$4\pi^2(f_x^2 + f_y^2) \leq k^2 \quad (4.3-13)$$

该式表明传播距离 z 至少大于若干个波长, 隐失波才可以被忽略。

在这种情况下, 式 (4.3-8) 表明波在均匀介质中传播可以视为一个 2-D 的线性空间滤波过程, 其传递函数为

$$H(f_x, f_y) = \begin{cases} e^{iz\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}} & 4\pi^2(f_x^2 + f_y^2) < k^2 \\ 0 & \text{其他} \end{cases} \quad (4.3-14)$$

正如图 4.1 所示的那样。传递函数的逆傅里叶变换, 即冲击响应, 将在例 4.6 中进行讨论。

无论在近场还是远场, 在无源空间中沿 z 方向传播的波场都可以由角谱传播理论正确地描述。其他两种描述波场传播的方法是菲涅耳和夫琅禾费近似, 将在本书的第 5 章加以讨论。但是, 它们仅仅在某些特定的条件下成立。

令 $F[\cdot]$ 和 $F^{-1}[\cdot]$ 分别代表傅里叶变换和傅里叶逆变换的变换算符。用这些运算符, 式 (4.3-12) 可以表示为

$$U(x, y, z) = F^{-1} \left[F[U(x, y, 0)] e^{ikz\sqrt{1 - \alpha_x^2 - \alpha_y^2}} \right] \quad (4.3-15)$$

对于具有特定值的 f_x 和 f_y , $A(f_x, f_y, 0) = F[U(x, y, 0)]$ 是平面波的复振幅, 其传播方向的方向余弦为 $\alpha_x = 2\pi f_x$, $\alpha_y = 2\pi f_y$, $\alpha_z = 2\pi f_z$, 其中 $f_z = \frac{1}{2\pi} [1 - \alpha_x^2 - \alpha_y^2]^{1/2}$ 。波传播的效应是不改变振幅, 只是以 $e^{ikz\sqrt{1 - \alpha_x^2 - \alpha_y^2}}$ 改变对应平面波的相位。

例 4.3 一个平面波 $Be^{ik_{x0}x}e^{ik_{z0}z}$ 通过位于 $z=0$ 、直径为 d 的圆形孔径, 计算其在 $z=a$ (常数) 处的角谱。

解: 在 $z=0$ 处的波可以写成

$$\begin{aligned} U(x, y, 0) &= Be^{ik_{x0}x} \text{cyl}\left(\frac{r}{d}\right) \\ &= Be^{i2\pi f_{x0}x} \text{cyl}\left(\frac{\sqrt{x^2 + y^2}}{d}\right) \end{aligned}$$

上面两个因式的傅里叶变换为

$$\begin{aligned} Be^{i2\pi f_{x0}x} &\leftrightarrow B\delta(f_x - f_{x0}, f_y) \\ \text{cyl}\left(\frac{r}{d}\right) &\leftrightarrow \frac{d^2\pi}{4} \text{somb}(d\rho) \end{aligned}$$

根据卷积理论, $U(x, y, 0)$ 的角谱为



$$\begin{aligned}
 A(f_x, f_y, 0) &= B\delta(f_x - f_{x_0}, f_y) * \frac{d^2\pi}{4}\text{somb}(d\rho) \\
 &= B\delta(f_x - f_{x_0}, f_y) * \frac{d^2\pi}{4}\text{somb}\left(d\sqrt{f_x^2 + f_y^2}\right) \\
 &= \frac{d^2\pi B}{4}\text{somb}\left(d\sqrt{(f_x - f_{x_0})^2 + f_y^2}\right)
 \end{aligned}$$

可以看到，在圆形孔径之前是单一的平面波成分，在经过孔径之后变成了具有无限个这种成分的频谱。它们大部分沿着入射波场的方向传播，并以一个边帽函数扩散。此扩散与孔径的直径 d 具有反比关系。

在距离 z 角谱变成了

$$A(f_x, f_y, z) = \frac{d^2\pi B}{4}\text{somb}\left(d\sqrt{(f_x - f_{x_0})^2 + f_y^2}\right)e^{jz\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}}$$

4.4 平面波角谱的快速傅里叶变换 (FFT)

将空间变量和空间频率变量离散化和截断后，关联 $U(x, y, z)$ 与 $U(x, y, 0)$ 的平面波角谱，可以采用快速傅里叶变换 (Fast Fourier Transform, FFT) 来计算 [Brigham, 1974]。对于离散傅里叶变换和 FFT 的讨论，可以参阅附录 B。在空域和频域对变量的离散化和截断给定为

$$\begin{aligned}
 x &= \Delta x \cdot n_1 \\
 y &= \Delta y \cdot n_2 \\
 f_x &= \Delta f_x \cdot m_1 \\
 f_y &= \Delta f_y \cdot m_2
 \end{aligned} \tag{4.4-1}$$

其中， Δx ， Δy ， Δf_x 和 Δf_y 是采样间隔； n_1 ， m_1 和 n_2 ， m_2 是整数且满足

$$\begin{aligned}
 -M_1 &\leq n_1, m_1 \leq M_1 \\
 -M_2 &\leq n_2, m_2 \leq M_2
 \end{aligned} \tag{4.4-2}$$

将 $U(\Delta x n_1, \Delta y n_2, z)$ 和 $A(\Delta f_x m_1, \Delta f_y m_2, z)$ 分别写成 $U[n_1, n_2, z]$ 和 $A[m_1, m_2, z]$ 。采用离散的空间和频率变量，式 (4.3-2) 和式 (4.3-8) 可以近似为

$$A[m_1, m_2, 0] = \Delta x \Delta y \sum_{n_1} \sum_{n_2} U[n_1, n_2, 0] e^{-j2\pi(\Delta f_x \Delta x n_1 m_1 + \Delta f_y \Delta y n_2 m_2)} \tag{4.4-3}$$

$$U[n_1, n_2, z] = \Delta f_x \Delta f_y \sum_{m_1} \sum_{m_2} A[m_1, m_2, 0] e^{jz\sqrt{k^2 - 4\pi^2(\Delta f_x^2 m_1^2 + \Delta f_y^2 m_2^2)}} e^{-j2\pi(\Delta f_x \Delta x n_1 m_1 + \Delta f_y \Delta y n_2 m_2)} \tag{4.4-4}$$

M_1 和 M_2 应该选择使不等式 (4.3-13) 满足的值，这样隐失波才可以被忽略。一种近似的方法是在傅里叶空间选择一个矩形的空间，例如

$$|k_x| = 2\pi|f_x| \leq k \tag{4.4-5}$$

$$|k_y| = 2\pi|f_y| \leq k \tag{4.4-6}$$

这样，要满足以下条件：

$$|f_{x_{\max}}| = |f_{y_{\max}}| = \frac{1}{\lambda} \tag{4.4-7}$$



选择 $M = M_1 = M_2$, $\Delta f = \Delta f_x = \Delta f_y$, 得出

$$f_{\max} = \Delta f M = \frac{1}{\lambda} \quad (4.4-8)$$

这样

$$M = \frac{1}{\Delta f \lambda} \quad (4.4-9)$$

在实际应用中, $f_{x_{\max}}$ 和 $f_{y_{\max}}$ 一般小于 $1/\lambda$ 。如果它们是已知的, 那么式 (4.4-8) 在 x 方向上变成了

$$M_1 = \frac{f_{x_{\max}}}{\Delta f_x} \quad (4.4-10)$$

在 y 方向上

$$M_2 = \frac{f_{y_{\max}}}{\Delta f_y} \quad (4.4-11)$$

假设 $f_{x_{\max}} = f_{y_{\max}} = f$, $\Delta f_x = \Delta f_y = \Delta f$, 则得出 $M = M_x = M_y$ 。

假设 FFT 的长度 N 沿着两个方向, 这样 $\Delta = \Delta x = \Delta y$, $\Delta f = \Delta f_x = \Delta f_y$ 。下面讨论 N 与 M 之间的关系。为了能够使用快速傅里叶变换, 下面的式子必须成立:

$$\Delta f \Delta = \frac{1}{N} \quad (4.4-12)$$

对于任意的 z 值, $U[n_1, n_2, z]$ 和 $A[m_1, m_2, z]$ 假设为周期的, 其周期为 N 。这样, 对于 m_1, m_2, n_1, n_2 , 它们满足

$$\begin{aligned} U[-n_1, n_2, z] &= U[N - n_1, n_2, z] \\ U[n_1, -n_2, z] &= U[n_1, N - n_2, z] \\ U[-n_1, -n_2, z] &= U[N - n_1, N - n_2, z] \\ A[-m_1, m_2, z] &= A[N - m_1, m_2, z] \\ A[m_1, -m_2, z] &= A[m_1, N - m_2, z] \\ A[-m_1, -m_2, z] &= A[N - m_1, N - m_2, z] \end{aligned} \quad (4.4-13)$$

将负数坐标区域映射到正数坐标区域, 如图 4.2 所示。

根据式 (4.4-13), 式 (4.4-3) 和式 (4.4-4) 可以写成如下形式:

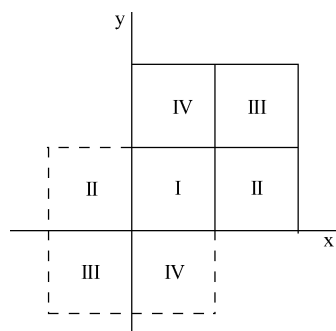


图 4.2 将负数坐标区域映射到正数坐标区域示意图, 以便于采用 FFT

$$A[m_1, m_2, 0] = \Delta^2 \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} U[n_1, n_2, 0] e^{-j\frac{2\pi}{N}(n_1 m_1 + n_2 m_2)} \quad (4.4-14)$$

$$U[n_1, n_2, z] = (\Delta f)^2 \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} A[m_1, m_2, z] e^{j\frac{2\pi}{N}(m_1 n_1 + m_2 n_2)} \quad (4.4-15)$$

其中

$$A[m_1, m_2, z] = A[m_1, m_2, 0] e^{jz \sqrt{k^2 - 4\pi^2 (\Delta f_x)^2 (m_1^2 + m_2^2)}} \quad (4.4-16)$$

其中, m_1 和 m_2 满足条件式 (4.4-2)。同样假设 $A[m_1, m_2, z]$ 是周期的, 且其周期为 N 。

$$\begin{aligned} A[-m_1, m_2, z] &= A[N - m_1, m_2, z] \\ A[m_1, -m_2, z] &= A[m_1, N - m_2, z] \end{aligned}$$



$$A[-m_1, -m_2, z] = A[N - m_1, N - m_2, z] \quad (4.4-17)$$

可以看到, 由于使用了循环卷积而不是线性卷积, 在使用 FFT 时因其周期性条件导致了混淆现象的产生, 为了减少混淆现象, 输入和输出孔径可以用零填充到一个尺寸, 即 $M' = 2M$ 。这样, N 可以做如下选择:

$$N = \begin{cases} 2M' & N \text{ 为偶数} \\ 2M' + 1 & N \text{ 为奇数} \end{cases} \quad (4.4-18)$$

另外, 当 N 是 2 的乘方的时候, FFT 通常具有更高的效率。这样, N 实际选择为大于上面计算获得的值, 使其为 2 的乘方。

式 (4.4-15) 和式 (4.4-16) 现在可以写成

$$U[n_1, n_2, z] = (\Delta f)^2 \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} \left[A[m_1, m_2, 0] e^{jz \sqrt{k^2 - 4\pi^2(\Delta f_x^2)(m_1^2 + m_2^2)}} \right] e^{j\frac{2\pi}{N}(n_1 m_1 + n_2 m_2)} \quad (4.4-19)$$

去掉一个规范化因子, 式 (4.4-19) 就是 2-D 离散傅里叶逆变换 (DFT) 的形式。

由于 $(\Delta f \Delta)^2 = \frac{1}{N^2}$, 式 (4.4-14) 和式 (4.4-15) 中的 Δ^2 和 $(\Delta f)^2$ 可以被忽略, 只需

其中的一个式 (4.4-14) 乘以 $\frac{1}{N^2}$ 。

总结起来, 采用角谱理论和 FFT 方法从 $U(x, y, 0)$ 获得 $U(x, y, z)$ 的步骤如下:

1. 采用上述方法生成 $U[n_1, n_2, 0]$ 。
2. 根据式 (4.4-14) 采用 FFT 计算 $A[m_1, m_2, 0]$ 。
3. 根据式 (4.4-16) 计算 $A[m_1, m_2, z]$ 。
4. 根据式 (4.4-19) 采用 FFT 计算 $U[n_1, n_2, z]$ 。
5. 根据式 (4.4-13) 调整 $U[n_1, n_2, z]$, 使负数坐标空间重新映射到正数坐标空间。

上述讨论的结果可以容易地推广到沿 x 和 y 方向不同数量的数据点。

例 4.4 假设做出如下定义:

fftshift: 由式 (4.4.13) 和式 (4.4.17) 定义的操作。

fft2, ifft2: 分别为 2-D FFT 和 2-D 逆 FFT。

H : 式 (4.3.14) 给出的传递函数。

u, U : 分别为输入和输出场。

根据该定义, 一段 ASM 程序为:

$$u1 = \text{fftshift}(\text{fft2}(\text{fftshift}(u)))$$

$$u2 = H \cdot u1$$

$$U = \text{fftshift}(\text{ifft2}(\text{fftshift}(u2)))$$

说明该程序可以被缩减到

$$u1 = \text{fftshift}(\text{fft2}(u))$$

$$u2 = H \cdot u1$$

$$U = \text{ifft2}(\text{fftshift}(u2))$$

解: 两段程序的不同之处在于, 在第一步和第三步程序中, 减少了 fftshift。这样做是可能的, 因为 1-D DFT 具有以下性质:



如果 $u'[n] = u\left[n - \frac{N}{2}\right]$, 那么 $U'[k] = U[k]e^{-j\pi nk}$ 。当第一步中的 fftshift 被去掉后, 其结果具有相移 $e^{-j\pi nk}$ 。当第三步程序中的 fftshift 被去掉后, 同样的相移产生正确的输出 U 。

例 4.5 假设给定如下参数:

输入尺寸: 32×32 $N = 512$

$\lambda = 0.0005$ $z = 1000$

上式中的单位为 mm。一个输入的高斯场为

$$U(x, y, 0) = \exp[-0.01\pi(x^2 + y^2)]$$

输入高斯场的亮度如图 4.3 所示。对应的输出场的亮度由上述 ASM 程序计算, 如图 4.4 所示。如果距离 z 变成 100, 对应的输出场如图 4.5 所示。因此, ASM 程序可以被应用于任何距离 z 。正如期望的那样, 如果输入场是高斯分布, 输出场同样是高斯分布。

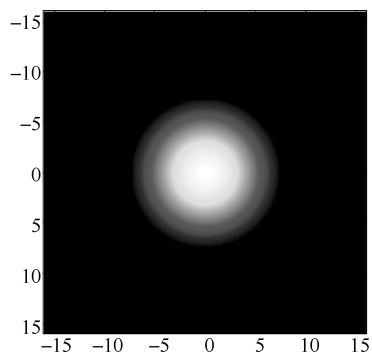


图 4.3 输入高斯场的亮度

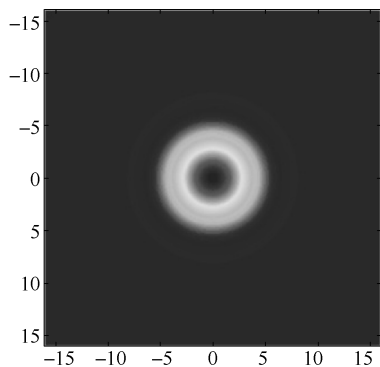


图 4.4 当距离 $z = 1000$ 时, 输出高斯场的亮度

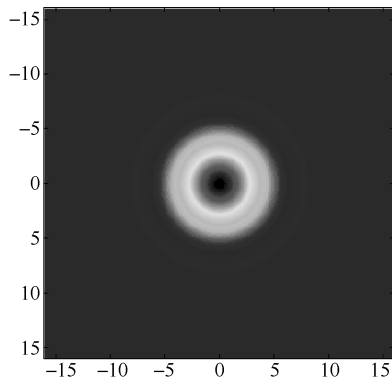


图 4.5 当距离 $z = 100$ 时, 输出高斯场的亮度

例 4.6 给出式 (4.3.14) 中传递函数的冲击响应函数。

解: 冲击响应是传递函数的逆傅里叶变换

$$h(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{ikz} [1 - \lambda^2 f_x^2 - \lambda^2 f_y^2]^{1/2} e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (4.4-20)$$

我们发现 $H(f_x, f_y)$ 具有圆柱对称性。因此, 采用柱面坐标, 令

$$x = r \cos \theta, \quad y = r \sin \theta, \quad f_x = \rho \cos \phi, \quad f_y = \rho \sin \phi$$

在柱面坐标下, 式 (4.4-20) 变成了 [Stark, 1982]

$$h(r) = \frac{1}{2\pi} \int_0^{+\infty} e^{-z(i^2 - k^2)^{1/2}} J_0(2\pi r \rho) \rho d\rho$$

式中, $J_0(\cdot)$ 是第一类零阶贝塞尔函数。

上述的积分可以进一步计算为

$$h(r) = \frac{e^{jk[z^2 + r^2]^{1/2}}}{j\lambda[z^2 + r^2]^{1/2}} \frac{z}{[z^2 + r^2]^{1/2}} \left[1 + \frac{1}{jk[z^2 + r^2]^{1/2}} \right] \quad (4.4-21)$$



例 4.7 给出点源 $U(x, y, 0) = \delta(x-3)\delta(y-4)$ 的波场。

解：输出的波场为

$$\begin{aligned} U(x, y, z) &= U(x, y, 0) * h(x, y, z) \\ &\simeq \delta(x-3)\delta(y-4) * \frac{1}{j\lambda z} e^{jk[z^2 + x^2 + y^2]^{1/2}} \\ &\simeq \frac{1}{j\lambda z} e^{jk[z^2 + (x-3)^2 + (y-4)^2]^{1/2}} \end{aligned}$$

4.5 基尔霍夫衍射理论

4.3 节中讨论的平面波角谱的传播描述了衍射。然而，衍射同样可以先用亥姆霍兹方程，然后采用格林定理把它转化为一个积分方程来处理。

格林定理引入两个复值函数 $U(\mathbf{r})$ 和 $G(\mathbf{r})$ （格林函数）。如图 4.6 所示，令 S 为包围体积为 V 的某空间的封闭曲面。

若在 S 面内和 S 面上， U 和 G 具有单值连续的一阶和二阶导数，没有任何的奇点，格林公式表明

$$\iiint_V (G \nabla^2 U - U \nabla^2 G) dv = \iint_S \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) ds \quad (4.5-1)$$

式中， $\partial/\partial n$ 表示 S 上每一点沿向外的法线方向上的偏导数。

在我们的分析中， U 对应于波场。

考虑一个任意传播的波场，产生于初始平面 $z=0$ ，传播到某一个 z 值的观察平面。图 4.7 显示了该几何关系。在图 4.7 中， P_0 是观察平面上的一个点。 P_1 可以视为是空间中的任意点，在图中位于初始平面上。封闭曲面是平面 S_0 和 S_1 的和。此处用有效积分表示衍射的格林函数有两种选择，下面将加以讨论。

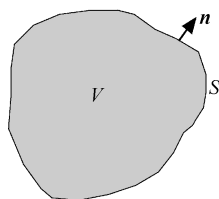


图 4.6 格林公式中使用的体积及其表面

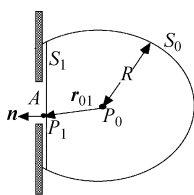


图 4.7 基尔霍夫衍射公式中的几何关系

4.5.1 基尔霍夫衍射理论

基尔霍夫选择的格林函数是一个球面波，给定为

$$G(\mathbf{r}) = \frac{e^{jkr_{01}}}{r_{01}} \quad (4.5-2)$$

式中， \mathbf{r} 是从 P_0 到 P_1 方向的位置矢量； r_{01} 是对应的距离，给出为

$$r_{01} = [(x_0 - x)^2 + (y_0 - y)^2 + z^2]^{1/2} \quad (4.5-3)$$

$U(\mathbf{r})$ 满足亥姆霍兹方程，由于 $G(\mathbf{r})$ 是一个扩展的球面波，同样满足亥姆霍兹方程

$$\nabla^2 G(\mathbf{r}) + k^2 G(\mathbf{r}) = 0 \quad (4.5-4)$$

式 (4.5-1) 的左边可以写成

$$\iiint_V [G \nabla^2 U - U \nabla^2 G] dv = \iiint_V k^2 [UG - UG] dv = 0$$

这样，式 (4.5-1) 变成了



$$\iint_S \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] ds = 0 \quad (4.5-5)$$

在图 4.7 中的表面 S_0 上, 有

$$G(\mathbf{r}) = \frac{e^{jkR}}{R}$$

$$\frac{\partial G(\mathbf{r})}{\partial n} = \left(jk - \frac{1}{R} \right) \frac{e^{jkR}}{R} \left(\frac{\partial U}{\partial n} - jkU \right)$$

在表面 S_0 上, 式 (4.5-1) 后面的积分部分变成了

$$\begin{aligned} I_{S_2} &= \int_{S_0} G(\mathbf{r}) \left(\frac{\partial U}{\partial n} - jkU \right) ds \\ &= \int_{\Omega} G(\mathbf{r}) \left(\frac{\partial U}{\partial n} - jkU \right) R^2 dw \\ &= \int_{\Omega} e^{jkR} \left(R \left(\frac{\partial U}{\partial n} - jkU \right) \right) dw \end{aligned}$$

式中, Ω 是 S_0 对 P_0 点所成的立体角。

当 $R \rightarrow \infty$ 时, 当满足下式条件后, 最后的积分项趋近于 0:

$$\lim_{R \rightarrow \infty} R \left(\frac{\partial U}{\partial n} - jkU \right) = 0 \quad (4.5-6)$$

这个条件被称为**索末菲辐射条件**。我们已经通过实验进行了验证。

接下来考虑的是 S_1 上的积分。除了标示为 A 的孔径之外, 其他部分是一个不透明的无穷平面。平面上孔径衍射的边界条件通常有两种近似: 德拜近似和基尔霍夫近似 [Goodman, 2004]。在德拜近似中, 入射光场的角谱有一个突然的截断, 从而只能得到沿特定方向穿过孔径的平面波。例如, 从孔径穿过的方向平面波可能聚焦在一个点。在聚焦区域的光场是穿过的所有平面波的叠加, 这些平面波在一个圆锥体中传播, 圆锥的顶点是聚焦点, 底是平面上的孔径。

基尔霍夫近似更加常用。如果孔径 A 在平面 $z=0$ 上, 如图 4.7 所示, 则在平面 $z=0^+$ 上的基尔霍夫近似条件为

$$U(x, y, 0^+) = \begin{cases} U(x, y, 0) & A \text{ 内} \\ 0 & A \text{ 外} \end{cases} \quad (4.5-7)$$

$$\frac{\partial}{\partial z} U(x, y, z) \Big|_{z=0^+} = \begin{cases} \frac{\partial}{\partial z} U(x, y, z) \Big|_{z=0^+} & A \text{ 内} \\ 0 & A \text{ 外} \end{cases} \quad (4.5-8)$$

可以看到, 根据基尔霍夫近似条件, $U(x, y, z)$ 和它在 z 方向上的偏导数, 当 $z=0$ 时, 在孔径外是不连续的, 而在孔径内是连续的。这种条件称为基尔霍夫边界条件。于是得到 $U(P_0)$ 的解

$$U(P_0) = \frac{1}{4\pi} \iint_A \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] ds \quad (4.5-9)$$

4.5.2 菲涅耳-基尔霍夫衍射公式

假设 r_{01} 远大于光的波长, 可以得到以下的近似条件:



$$\frac{\partial G(P_1)}{\partial n} = \frac{\partial}{\partial n} \left[\frac{e^{jkr_{01}}}{r_{01}} \right] = \cos(\theta) \left[jk - \frac{1}{r_{01}} \right] G(P_1) \cong jk \cos(\theta) G(P_1) \quad (4.5-10)$$

这样, 式 (4.5-9) 变成了

$$U(P_0) = \frac{1}{4\pi} \iint_A \frac{e^{jkr_{01}}}{r_{01}} \left[\frac{\partial U}{\partial n} - jk \cos(\theta) U \right] ds$$

把该式应用于从点 P_2 发出的球面波 $U(P_1)$, 如图 4.8 所示。定义 P_1 与 P_2 之间的距离为 r_{21} , \mathbf{n} 与 \mathbf{r}_{21} 之间的夹角为 θ_2 , 得到

$$U(P_1) = G(r_{21}) \frac{e^{jkr_{21}}}{r_{21}}$$

$$\frac{\partial U(P_1)}{\partial n} \cong jk \cos(\theta_2) G(r_{21})$$

得到

$$U(P_0) = \frac{1}{j\lambda} \iint_A \frac{e^{jk(r_{21}+r_{01})}}{r_{21}r_{01}} \left[\frac{\cos(\theta) - \cos(\theta_2)}{2} \right] ds \quad (4.5-11)$$

式 (4.5-11) 为菲涅耳-基尔霍夫衍射公式, 它对球面波通过平面孔径的衍射是有效的。

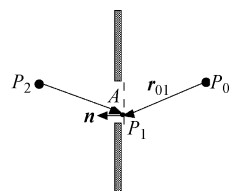


图 4.8 球面波照明下的一个平面孔径

4.6 瑞利-索末菲衍射理论

在前面的分析中, 利用式 (4.5-2) 所示的格林函数, 以及一系列的简化近似, 得到了菲涅耳-基尔霍夫公式。然而, 在该公式中具有一些内在矛盾。这些矛盾将被索末菲利用如下的格林函数加以解决:

$$G_2(\mathbf{r}) = \frac{e^{jkr_{01}}}{r_{01}} - \frac{e^{jkr_{01}}}{\underline{r_{01}}} \quad (4.6-1)$$

其中, $\underline{r_{01}}$ 是从点 P_0 到 P_1 的距离; $\underline{P_0}$ 是 P_0 相对于初始平面的镜像。

利用上述第二格林函数, 得到了第一瑞利-索末菲衍射公式, 为

$$U(x_0, y_0, z) = \frac{1}{j\lambda} \iint_{-\infty}^{+\infty} U(x, y, 0) \frac{z}{r_{01}} \frac{e^{jkr_{01}}}{r_{01}} dx dy \quad (4.6-2)$$

基于卷积理论的该方程的另一种形式将在下一节加以介绍。

平面孔径的瑞利-索末菲衍射模型如图 4.9 所示。

式 (4.6-2) 表明 $U(x_0, y_0, z)$ 可以被视为发散球面波的线性组合, 每一个球面波从点 $(x, y, 0)$ 发出, 权重为 $\frac{1}{j\lambda} \frac{z}{r_{01}}$, 这在数学上称为惠更斯-菲涅耳定律。

可以看到, 式 (4.6-2) 是一个在 x, y 上的 2-D 线性卷积

$$U(x_0, y_0, z) = U(x, y, 0) * h(x, y, z) \quad (4.6-3)$$

其冲击响应 $h(x, y, z)$ 为

$$h(x, y, z) = \frac{1}{j\lambda z} \frac{e^{jkz \left[1 + \frac{x^2 + y^2}{z^2} \right]^{1/2}}}{\left[1 + \frac{x^2 + y^2}{z^2} \right]} \quad (4.6-4)$$

传递函数 $H(f_x, f_y, z)$ 是 $h(x, y, z)$ 的 2-D 傅里叶变换, 为

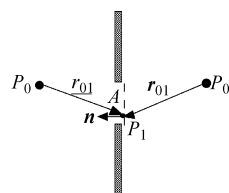


图 4.9 平面孔径的瑞利-索末菲衍射模型



$$H(f_x, f_y, z) = e^{iz[k^2 - 4\pi^2(f_x^2 + f_y^2)]^{1/2}} \quad (4.6-5)$$

这与在 4.3 节中讨论的平面波角谱传播的传递函数是相同的。因此，瑞利-索末菲衍射理论和平面波角谱理论是完全等价的。

4.6.1 基尔霍夫近似

将基尔霍夫近似合并入第一瑞利-索末菲积分式得到

$$U(x, y, z) = \frac{1}{j\lambda} \iint_A U(x, y, 0) \frac{z}{r_{01}} \frac{e^{jkr_{01}}}{r_{01}} dx dy \quad (4.6-6)$$

z/r_{01} 经常被写成 $\cos(n, r_{01})$ ，表示图 4.9 中 z 轴与 r_{01} 之间夹角的余弦。

4.6.2 第二瑞利-索末菲公式

可以用另一种同样成立的格林函数来代替 $G_2(\mathbf{r})$ ，函数为

$$G_3(\mathbf{r}) = \frac{e^{jkr_{01}}}{r_{01}} + \frac{e^{jkr'_{01}}}{r'_{01}} \quad (4.6-7)$$

$$U_2(x, y, z) = \frac{1}{2\pi} \iint_A \frac{\delta U(x, y, 0)}{\delta n} \frac{e^{jkr_{01}}}{r_{01}} dx dy \quad (4.6-8)$$

式中， $\frac{\delta U(x, y, 0)}{\delta n}$ 是 $U(x, y, 0)$ 在法向 (z 方向) 上的偏导数。

可以看到，4.5 节中讨论的基尔霍夫解是第一和第二瑞利-索末菲解的算术平均 [Goodman, 2004]。

4.7 第一瑞利-索末菲衍射积分的另一种推导

假设场 $U(\mathbf{r})$ 源自半空间 $z \geq 0$ ，该空间位于 $z = 0$ 平面上。我们需要得到 $z \geq 0$ 的 $U(\mathbf{r})$ 。 $U(x, y, 0)$ 用角谱形式表示为

$$U(x, y, 0) = \iint_{-\infty}^{+\infty} A(f_x, f_y, 0) e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (4.7-1)$$

从卷积的理论分析，这个方程可以理解为找到一个具有归一化传递函数的 LSI 系统的输出。该系统的冲击响应为

$$h(x, y, 0) = \iint_{-\infty}^{+\infty} e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (4.7-2)$$

如果 $z \neq 0$ 会怎样呢？基于式 (4.7-1) 以及 4.6 节中衍射积分的其他推导， $h(x, y, 0)$ 表示为

$$h(x, y, z) = \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} e^{jk \cdot \mathbf{r}} dk_x dk_y \quad (4.7-3)$$

其中， \mathbf{r} 和 \mathbf{k} 分别在式 (3.2-28) 和式 (3.2-30) 中给出。式 (4.7-3) 可以写成

$$h(x, y, z) = -\frac{1}{2\pi} \frac{\partial}{\partial z} \left(\frac{j}{2\pi} \iint_{-\infty}^{+\infty} \frac{e^{jk \cdot \mathbf{r}}}{k_z} dk_x dk_y \right) \quad (4.7-4)$$

括号内的量可以看成是球面波的平面波展开 [Weyl]

$$\frac{e^{jkr}}{r} = \frac{j}{2\pi} \iint_{-\infty}^{+\infty} \frac{e^{jk \cdot \mathbf{r}}}{k_z} dk_x dk_y \quad (4.7-5)$$



将结果代入式 (4.7-4), 得到

$$h(x, y, z) = \frac{1}{2\pi} \frac{\partial}{\partial z} \left(\frac{e^{jkr}}{r} \right) \quad (4.7-6)$$

式 (4.7-6) 等价于

$$h(x, y, z) = \frac{1}{j\lambda} \frac{z}{r} \frac{e^{jkr}}{r} \quad (4.7-7)$$

衍射场 $U(x, y, z)$ 可以表示为 $U(x, y, 0)$ 与 $h(x, y, z)$ 的卷积, 于是

$$U(x, y, z) = \frac{1}{j\lambda} \iint_{-\infty}^{+\infty} U(x, y, 0) \frac{z}{r_{01}} \frac{e^{jkr_{01}}}{r_{01}} dx dy \quad (4.7-8)$$

利用式 (4.7-6), 该结果经常写成

$$U(x, y, z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(x, y, 0) \frac{\partial}{\partial z} \left(\frac{e^{jkr_{01}}}{r_{01}} \right) dx dy \quad (4.7-9)$$

4.8 非单色波的瑞利-索末菲衍射积分

在非单色波的情况下, 如式 (4.2-5) 中那样, $U(\mathbf{r}, t)$ 用其傅里叶变换形式表示, 即 $U_f(\mathbf{r}, f) = U_f(x, y, z, f)$ 。代入 $f' = -f$, 该式可以写成

$$U(x, y, z, t) = \int_{-\infty}^{+\infty} U_f(x, y, z, -f') e^{-j2\pi f' t} df' \quad (4.8-1)$$

$U_f(x, y, z, -f')$ 满足单孔径下的瑞利-索末菲积分

$$U_f(x, y, z, -f') = \frac{1}{j\lambda} \iint_A U_f(x, y, 0, f') \frac{z}{r_{01}} \frac{e^{jkr_{01}}}{r_{01}} dx dy \quad (4.8-2)$$

将结果代入式 (4.8-1), 得到

$$U(x, y, z, t) = \int_{-\infty}^{+\infty} \left[\frac{1}{j\lambda} \iint_A U_f(x, y, 0, -f') \frac{z}{r_{01}} \frac{e^{jkr_{01}}}{r_{01}} dx dy \right] e^{-j2\pi f' t} df' \quad (4.8-3)$$

有

$$\lambda f = c \quad (4.8-4)$$

式中, c 是波的相位速度。

利用这种关系并改变式 (4.8-3) 中的积分顺序, 得到

$$U(x, y, z, t) = \iint_A \frac{z}{2\pi c r_{01}^2} \left[\int_{-\infty}^{+\infty} -j2\pi f' U_f(x, y, 0, -f') e^{-j2\pi f' (t - \frac{r_{01}}{c})} df' \right] dx dy \quad (4.8-5)$$

由于

$$\begin{aligned} \frac{d}{dt} U(x, y, 0, t) &= \frac{d}{dt} \int_{-\infty}^{+\infty} U_f(x, y, 0, -f') e^{-j2\pi f' t} df' \\ &= \int_{-\infty}^{+\infty} -j2\pi f' U_f(x, y, 0, -f') e^{-j2\pi f' t} df' \end{aligned}$$

式 (4.8-5) 同样可以写成

$$U(x, y, z, t) = \frac{z}{2\pi c} \iint_A \frac{1}{r_{01}^2} \frac{d}{dt} \left[U \left(x, y, 0, t - \frac{r_{01}}{c} \right) \right] dx dy \quad (4.8-6)$$

可以看到, 在 (x, y, z) , z_0 处的波依赖于 $(x, y, 0)$ 处的波对时间的导数, 该波被延迟 r_{01}/c , 这个时间正是波从 P_1 传播到 P_0 需要的时间。

第 5 章 菲涅耳与夫琅禾费近似

5.1 本章引论

在标量衍射积分的菲涅耳近似与夫琅禾费近似情况下，可以利用简单的傅里叶积分来计算光波的传播，而且允许衍射输入面和输出面具有不同的窗口尺寸。但是，这种近似仅在特定区域成立，当输出面非常靠近衍射输入面孔径时，近似不再成立。瑞利-索末菲积分、菲涅耳近似和夫琅禾费近似各自成立的区域如图 5.1 所示。

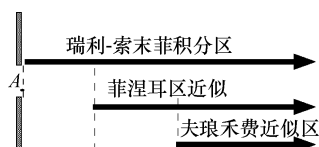


图 5.1 三个衍射区的划分

衍射输入面右边的整个半区为瑞利-索末菲区，菲涅耳区和夫琅禾费区均为瑞利-索末菲区的一部分。菲涅耳近似区的开始边界可以从本章的方程式 (5.2-5) 和式 (5.2-7) 推导出，而夫琅禾费近似区的开始边界可以从方程式 (5.5-1) 导出。但是，对于这些区域边界的理解要非常谨慎，在第 7 章中给出进一步解释。

通常来说，远场一般指夫琅禾费区，而近场可以认为是介于衍射输入面和夫琅禾费衍射区之间的区域。

本章共分 10 节。其中，5.2 节给出菲涅耳区内光波的传播；5.3 节推导了菲涅耳区波动传播的快速傅里叶变换方法；5.4 节推导出菲涅耳近似实质上为傍轴波动方程的解；5.5 节讨论了夫琅禾费区的波动传播。

衍射光栅是一类具有周期性结构的光学器件，具有非常重要的应用。对如何分析由光栅发出的光在菲涅耳区和夫琅禾费区的衍射分布，本章也给出了经典的例子。本章第 5.6 节讨论了衍射光栅的基本结构。第 5.7 节、5.8 节和 5.9 节分别重点讨论了正弦幅值型光栅的夫琅禾费衍射、正弦幅值型光栅的菲涅耳衍射和正弦相位型光栅的夫琅禾费衍射。

5.2 菲涅耳衍射

仅考虑径向尺度为 L_1 的范围内的输入波

$$U(x, y, 0) = 0, \sqrt{x^2 + y^2} > L_1 \quad (5.2-1)$$

类似地，在输出观察面上仅关注径向尺度为 L_2 的范围内的场分布

$$U(x_0, y_0, z_0) = 0, \sqrt{x_0^2 + y_0^2} > L_2 \quad (5.2-2)$$

式 (4.6-3) 中的 r_{01} 由下式给出：

$$r_{01} = z \left[1 + \frac{(x_0 - x)^2 + (y_0 - y)^2}{z^2} \right]^{1/2} \quad (5.2-3)$$



其中, $(x, y, 0)$ 为输入面上点的坐标; $(x_0, y_0, 0)$ 为观察面上点的坐标。在上述两个限制条件下, 可以得到

$$[(x_0 - x)^2 + (y_0 - y)^2]_{\max} \leq (L_1 + L_2)^2 \quad (5.2-4)$$

上述不等式的上界将在后面讨论中用到。如果

$$|z| \gg L_1 + L_2 \quad (5.2-5)$$

则方程式 (4.7-6) 中的 $(1/j\lambda)(z/r_{01}^2)$ 可以近似为 $1/j\lambda$ 。但是, 在处理相位项时需要特别谨慎。 kr_{01} 可展开成如下的二项式:

$$kr_{01} = kz + \frac{k}{2z}[(x_0 - x)^2 + (y_0 - y)^2] - \frac{k}{8z^3}[(x_0 - x)^2 + (y_0 - y)^2]^2 + \cdots \quad (5.2-6)$$

其中第三项的绝对值最大为 $k(L_1 + L_2)^4/(8|z|^3)$ 。当满足如下条件时, 此项将远小于 1rad

$$|z|^3 \gg \frac{k(L_1 + L_2)^2}{8} \quad (5.2-7)$$

在上述对 $|z|$ 的约束下, 相位项可以做如下近似:

$$kr_{01} \cong kz + \frac{k}{2z}[(x_0 - x)^2 + (y_0 - y)^2] \quad (5.2-8)$$

上述所做的近似就是所谓的菲涅耳近似。由式 (5.2-5) 和式 (5.2-7) 所限定的区域称为菲涅耳区。在这个区域内, 式 (4.7-8) 可写为

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} \iint_{-\infty}^{+\infty} U(x, y, 0) e^{j\frac{\pi}{\lambda z}[(x_0 - x)^2 + (y_0 - y)^2]} dx dy \quad (5.2-9)$$

式 (5.2-9) 是对变量 x 和 y 的二维卷积, 可以进一步写作

$$U(x_0, y_0, z) = U(x, y, 0) * h(x, y, z) \quad (5.2-10)$$

式中的脉冲响应函数由下式给出:

$$h(x, y, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{\pi}{\lambda z}(x^2 + y^2)} \quad (5.2-11)$$

对应的传递函数为

$$H(f_x, f_y, z) = e^{jkz} e^{-j\pi\lambda z(f_x^2 + f_y^2)} \quad (5.2-12)$$

式 (5.2-9) 中的二次方可展开, 因此式 (5.2-9) 可化简为

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} \iint_{-\infty}^{\infty} U'(x, y, 0) e^{-j\frac{2\pi}{\lambda z}(xx_0 + yy_0)} dx dy \quad (5.2-13)$$

其中

$$U'(x, y, 0) = U(x, y, 0) e^{j\frac{k}{2z}(x^2 + y^2)} \quad (5.2-14)$$

式 (5.2-13) 表明, 暂不考虑相乘的幅值和相位因子, $U(x_0, y_0, z)$ 恰好是 $U'(x, y, 0)$ 的二维傅里叶变换, 对应的空间频率分别为 $f_x = x_0/(\lambda z)$, $f_y = y_0/(\lambda z)$ 。

例 5.1 利用式 (4.4-21) 和角谱法, 求菲涅耳衍射的脉冲响应函数。

解: 对式 (4.4-21), 当满足 $z \gg r$ 时, 式 (4.4-21) 右边的第二项及第三项均可以近似为 1。将 $[z^2 + r^2]^{1/2}$ 展成泰勒级数, 并保留前两项, 则式 (4.4-21) 变为

$$h(x, y, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{\pi}{\lambda z}(x^2 + y^2)}$$

其与式 (5.2-11) 完全一致。



例 5.2 已知点光源可用 $U(x, y, 0) = \delta(x - 3)\delta(y - 4)$ 模拟, 求其菲涅耳区的衍射场分布。

解: 输出面的衍射场可表示为

$$\begin{aligned} U(x, y, z) &= U(x, y, 0) * h(x, y, z) \\ &\simeq [\delta(x - 3)\delta(y - 4)] * \frac{1}{j\lambda z} e^{jkz \left(1 + \frac{x^2 + y^2}{2z^2}\right)} \\ &\simeq \frac{1}{j\lambda z} e^{jkz \left[1 + \frac{(x-3)^2 + (y-4)^2}{2z^2}\right]} \end{aligned}$$

例 5.3 采用波长为 $0.5\mu\text{m}$ 的平面波垂直入射, 波照明半径为 2mm 的圆形孔径。如果观察区的半径限制为 10cm 以内, 求当 z 满足什么条件才能得到菲涅耳衍射。

解: 由 $L_1 = 0.2\text{cm}$, $L_2 = 10\text{cm}$, 根据式 (5.2-4), 有

$$z \gg 10.2\text{cm}$$

为了使不等式成立, 可考虑乘以 10 倍放大因子, 则有

$$z \geq 1.02\text{m}$$

同时, 由式 (5.2-7), 可知 z 需满足的第二个限制条件为

$$z^3 \gg \frac{2\pi \times 0.102^4}{4 \times 10^{-6}} \text{m}^3$$

即

$$z \gg 5.4\text{m}$$

采用 10 倍的放大因子, 可最后得到 z 需满足

$$z \geq 54\text{m}$$

可见, 由式 (5.2-7) 确定的第二个限制条件给出的 z 值要更大一些。虽然如此, 这个条件对于有效的菲涅耳近似并不是必要的 [Goodman]。一般来说, kr_{01} 的值非常大, 这导致二次相位因子 $e^{jkr_{01}}$ 产生快速振荡。因此, 对衍射积分起主要贡献的点来自于那些接近 (x_0, y_0) 的点 ($x \sim x_0, y \sim y_0$), 这些点称为稳相点 [Stammnes, 1986]。所以, 对一个区域, 当其最小 z 值处于由式 (5.2-4) 和式 (5.2-7) 给出的两个值之间时, 可以认为此时菲涅耳近似成立。事实上, 在近场区域菲涅耳近似经常被用来分析许多光学现象, 关于这种近似有效的原因将在第 7 章介绍。

例 5.4 若正方形孔径的边长为 D , 采用振幅为 1 的平面波照明, 求其菲涅耳衍射的条纹。

解: 输入波可写为

$$U(x, y, 0) = \text{rect}\left(\frac{x}{D}\right) \text{rect}\left(\frac{y}{D}\right)$$

菲涅耳衍射可以写成卷积形式

$$\begin{aligned} U(x_0, y_0, z) &= \frac{e^{jkz}}{j\lambda z} \iint_{-D/2}^{D/2} e^{j\frac{\pi}{\lambda z} [(x-x_0)^2 + (y-y_0)^2]} dx dy \\ &= \frac{e^{jkz}}{j} \left[\frac{1}{\sqrt{\lambda z}} \int_{-D/2}^{D/2} e^{j\frac{\pi}{\lambda z} (x-x_0)^2} dx \right] \left[\frac{1}{\sqrt{\lambda z}} \int_{-D/2}^{D/2} e^{j\frac{\pi}{\lambda z} (y-y_0)^2} dy \right] \end{aligned} \quad (5.2-15)$$

考虑如下积分:



$$I_x = \frac{1}{\sqrt{\lambda z}} \int_{-D/2}^{D/2} e^{j\frac{\pi}{\lambda z}(x-x_0)^2} dx$$

令

$$v = \sqrt{\frac{2}{\lambda z}}(x - x_0)$$

则 I_x 变为

$$I_x = \frac{1}{\sqrt{2}} \int_{x_b}^{x_e} e^{j\frac{\pi}{2}v^2} dv \quad (5.2-16)$$

其中 x_e 和 x_b 分别为

$$x_b = -\sqrt{\frac{2}{\lambda z}}\left(\frac{D}{2} + x_0\right)$$

$$x_e = \sqrt{\frac{2}{\lambda z}}\left(\frac{D}{2} + x_0\right)$$

于是式 (5.2-15) 可改写为

$$I_x = \frac{1}{\sqrt{2}} \int_0^{x_e} e^{j\frac{\pi}{2}v^2} dv - \frac{1}{\sqrt{2}} \int_0^{x_b} e^{j\frac{\pi}{2}v^2} dv \quad (5.2-17)$$

菲涅耳积分 $C(z)$ 和 $S(z)$ 定义为

$$C(z) = \int_0^z \cos\left(\frac{\pi v^2}{2}\right) dv \quad (5.2-18)$$

$$S(z) = \int_0^z \sin\left(\frac{\pi v^2}{2}\right) dv \quad (5.2-19)$$

利用 $C(z)$ 和 $S(z)$, 式 (5.2-17) 可写为

$$I_x = \frac{1}{\sqrt{2}} \{ [C(x_e) - C(x_b)] + j[S(x_e) - S(x_b)] \} \quad (5.2-20)$$

类似地, 对式 (5.2-15) 中关于 y 的积分部分也可以利用上述积分, 并记为 I_y :

$$I_y = \frac{1}{\sqrt{2}} \int_{y_b}^{y_e} e^{j\frac{\pi}{2}v^2} dv \quad (5.2-21)$$

$$= \frac{1}{\sqrt{2}} \{ [C(y_e) - C(y_b)] + j[S(y_e) - S(y_b)] \}$$

其中

$$y_b = -\sqrt{\frac{2}{\lambda z}}\left(\frac{D}{2} + y_0\right)$$

$$y_e = \sqrt{\frac{2}{\lambda z}}\left(\frac{D}{2} - y_0\right)$$

对于 $U(x_0, y_0, z)$, 方程式 (5.2-15) 可写为

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j} I_x I_y \quad (5.2-22)$$

于是发光强度 $I(x_0, y_0, z) = |U(x_0, y_0, z)|^2$ 可写为



$$I(x_0, y_0, z) = \frac{1}{4} \{ [C(x_e) - C(x_b)]^2 + [S(x_e) - S(x_b)]^2 \} \times \quad (5.2-23)$$

$$\{ [C(y_e) - C(y_b)]^2 + [S(y_e) - S(y_b)]^2 \}$$

菲涅耳数 N_F 定义为

$$N_F = \frac{D^2}{4\lambda z} \quad (5.2-24)$$

当 D 和 λ 为常数时, 菲涅耳数 N_F 随着 z 的增加而减小。

方孔的菲涅耳衍射的强度分布分别如图 5.2 ~ 图 5.5 所示。随着 z 不断减小, 衍射光强分布开始变得与输入面处的正方形孔径相似, 这在例 5.5 中得到了进一步的例证。图 5.2 给出了所用的方孔图像, 图 5.3 是方孔的菲涅耳衍射条纹, 图 5.4 为图 5.3 中衍射条纹的三维图像。

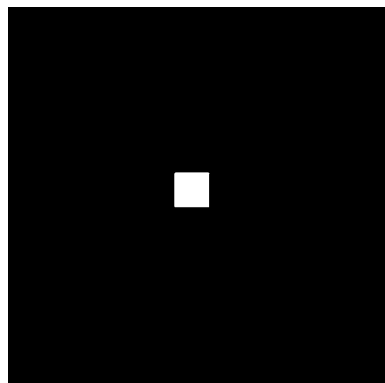


图 5.2 研究衍射所用的输入面处的方形孔径

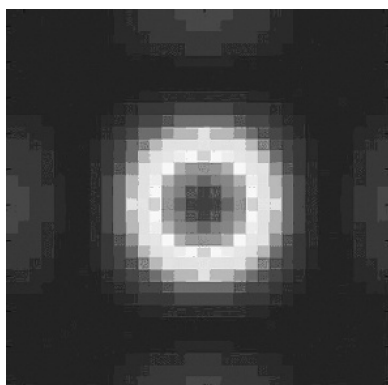


图 5.3 方孔在菲涅耳区的衍射条纹强度分布

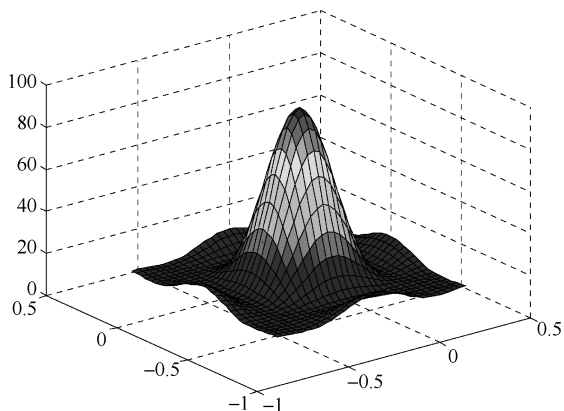


图 5.4 与图 5.3 相同的方孔在菲涅耳区的衍射条纹的三维空间强度分布

例 5.5 利用菲涅耳近似, 验证对于较小的 z , 矩形孔径的衍射条纹强度分布与矩形孔径相似。

解: 矩形孔径的衍射条纹强度分布由式 (5.2-23) 给出, 其中 D_x 和 D_y 分别表示 x 和 y 方向的孔径尺寸 D 。对于非常小的 z , 各变量可写为

$$x_b = -\sqrt{\frac{2}{\lambda z}} \left(\frac{D}{2} + x_0 \right) \rightarrow \begin{cases} -\infty, & x_0 > -D_x/2 \\ +\infty, & x_0 < -D_x/2 \end{cases}$$

$$x_e = \sqrt{\frac{2}{\lambda z}} \left(\frac{D}{2} - x_0 \right) \rightarrow \begin{cases} +\infty, & x_0 < D_x/2 \\ -\infty, & x_0 > D_x/2 \end{cases}$$

$$y_b = -\sqrt{\frac{2}{\lambda z}} \left(\frac{D}{2} + y_0 \right) \rightarrow \begin{cases} -\infty, & y_0 > -D_y/2 \\ +\infty, & y_0 < -D_y/2 \end{cases}$$

$$y_e = \sqrt{\frac{2}{\lambda z}} \left(\frac{D}{2} - y_0 \right) \rightarrow \begin{cases} +\infty, & y_0 < D_y/2 \\ -\infty, & y_0 > D_y/2 \end{cases}$$



而余弦积分和正弦积分的渐进极限分别为

$$C(+\infty) = S(+\infty) = 0.5, \quad C(-\infty) = S(-\infty) = -0.5$$

于是, 对于非常小的 z , 式 (5.2-22) 变为

$$I(x_0, y_0, z) \simeq \text{rect}\left(\frac{x_0}{D}\right) \text{rect}\left(\frac{y_0}{D}\right)$$

例 5.6 试给出使角谱法近似变为菲涅耳近似的条件。

解: 角谱法的传递函数由下式给出:

$$H_A(f_x, f_y) = \begin{cases} e^{jz \sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}} & \sqrt{f_x^2 + f_y^2} < \frac{1}{\lambda} \\ 0 & \text{其他} \end{cases}$$

而菲涅耳近似的传递函数由下式给出:

$$H_F(f_x, f_y) = e^{jkz} e^{-j\pi\lambda z(f_x^2 + f_y^2)}$$

如果有

$$\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)} \cong k - \frac{4\pi^2(f_x^2 + f_y^2)}{2k}$$

这就是说, 如果保留 $H_A(f_x, f_y)$ 的泰勒展开的前两项, 则 $H_A(f_x, f_y)$ 变为 $H_F(f_x, f_y)$ 。当 $|\lambda f_x|$ 、 $|\lambda f_y| \ll 1$ 时, 这样的近似成立。因此, 当衍射角度较小时, 菲涅耳近似和角谱法近似等价。

5.3 菲涅耳衍射的快速傅里叶变换的实现

菲涅耳衍射既可以通过由式 (5.2-10) 所给出的卷积来实现, 也可以直接由对式 (5.2-13) 进行离散并利用离散傅里叶变换 (Discrete Fourier Transform, DFT) 来实现。利用卷积法实现时, 与利用平面波的角谱法来实现类似, 这里不再讨论。下面讨论第二种方法。

令 $U(m_1, m_2, 0) = U(\Delta x m_1, \Delta y m_2, 0)$ 为 $u(x, y, 0)$ 的离散化表示, x 和 y 分别离散为 $\Delta x m_1$ 和 $\Delta y m_2$ 。类似地, x_0 和 y_0 分别离散为 $\Delta x_0 n_1$ 和 $\Delta y_0 n_2$ 。于是

$$U'(m_1, m_2, 0) = U(m_1, m_2, 0) e^{j\frac{k}{2z}[(\Delta x m_1)^2 + (\Delta y m_2)^2]} \quad (5.3-1)$$

为了采用快速傅里叶变换, 式 (5.2-13) 的傅里叶变换核需表示成如下形式:

$$e^{-j\frac{2\pi}{\lambda z}(x_0 x + y_0 y)} = e^{-j\frac{2\pi}{\lambda z}(\Delta x \Delta x_0 m_1 n_1 + \Delta y \Delta y_0 m_2 n_2)} = e^{-j2\pi\left(\frac{m_1 n_1}{N_1} + \frac{m_2 n_2}{N_2}\right)} \quad (5.3-2)$$

式中, N_1 、 N_2 为快速傅里叶变换的数据大小。

进一步, 由式 (5.3-2) 可知

$$\frac{2\pi}{\lambda z} x_0 x = \frac{2\pi}{\lambda z} \Delta x \Delta x_0 m_1 n_1 = 2\pi \frac{m_1 n_1}{N_1}$$

于是有

$$\Delta x \Delta x_0 = \frac{\lambda z}{N_1} \quad (5.3-3)$$

类似地, 有



$$\Delta y \Delta y_0 = \frac{\lambda z}{N_2} \quad (5.3-4)$$

这就是说, 较小的 Δx 或 Δy 分别对应于较大的 Δx_0 或 Δy_0 , 反之亦然。将方程式 (5.3-3) 与式 (5.3-4) 和式 (4.4-12) 所给出的角谱法对比, 由于 $f_x = x_0/\lambda z$, $f_y = y_0/\lambda z$, 所以这几个方程式是等价的。不过现在输出可以直接由式 (5.2-3) 给出, 而不需要通过逆变换的计算, 所以此时输入和输出面的窗函数可以具有不同的尺寸。

为了使用快速傅里叶变换, 如第 4.4 节中所讨论的, $U'(m_1, m_2, 0)$ 被平移了 $\left(\frac{N_1}{2}, \frac{N_2}{2}\right)$ 。用 $U''(m_1, m_2, 0)$ 表示变换的结果, 于是离散的输出波 $U(n_1, n_2, z) = U(\Delta x_0 n_1, \Delta y_0 n_2, z)$ 可表示如下:

$$U(n_1, n_2, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}[(\Delta x_0 n_1)^2 + (\Delta y_0 n_2)^2]} \Delta x \Delta y \sum_{m_1=0}^{N_1-1} \sum_{m_2=0}^{N_2-1} U''(m_1, m_2, 0) e^{-j2\pi\left(\frac{m_1 n_1}{N_1} + \frac{m_2 n_2}{N_2}\right)} \quad (5.3-5)$$

将 $U(n_1, n_2, z)$ 再次平移 $\left(\frac{N_1}{2}, \frac{N_2}{2}\right)$ 以对应于实际的场分布。

例 5.7 对例 4.4, 采用该例题中相同的定义, 菲涅耳衍射的伪代码可写成如下形式:

$$\begin{aligned} u1 &= u * \exp\left[j\frac{k}{2z}(x^2 + y^2)\right] \\ u2 &= \text{fftshift}(u1) \\ u3 &= \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} * u2 \\ U &= \text{fftshift}(u3) \end{aligned}$$

5.4 傍轴波动方程

菲涅耳近似实质上是下面将要推导的傍轴波动方程的一个解。考虑亥姆霍兹方程, 如果假设场主要沿着 z 向传播, 其可以表达为

$$U(x, y, z) = U'(x, y, z) e^{jkz} \quad (5.4-1)$$

其中假设 $U'(x, y, z)$ 是关于 z 的缓变函数。将式 (5.4-1) 代入亥姆霍兹方程, 有

$$\nabla^2 U' + 2jk \frac{\delta}{\delta z} U' = 0 \quad (5.4-2)$$

由于 $U'(x, y, z)$ 是关于 z 的缓变函数, 因此方程式 (5.4-2) 可近似为

$$\frac{\delta^2 U'}{\delta x^2} + \frac{\delta^2 U'}{\delta y^2} + 2jk \frac{\delta}{\delta z} U' = 0 \quad (5.4-3)$$

方程式 (5.4-3) 被称为傍轴波动方程。

若考虑

$$U_1(x, y) = \frac{1}{j\lambda z} e^{j\frac{k}{2z}(x^2 + y^2)} \quad (5.4-4)$$

显然 U_1 为方程式 (5.4-3) 的一个解。因为方程式 (5.4-3) 是线性的, 且是平移不变的, 所以对于任意的 x_1 和 y_1 , $U_1(x - x_1, y - y_1)$ 也是方程的解。叠加所有这些解, 可以得到方程



式 (5.4-3) 的通解为

$$U(x, y, z) = \frac{1}{j\lambda z} \iint g(x_1, y_1) e^{\frac{jk[(x-x_1)^2 + (y-y_1)^2]}{2z}} dx_1 dy_1 \quad (5.4-5)$$

考虑菲涅耳近似

$$U(x_0, y_0, z) = \frac{1}{j\lambda z} \iint U(x, y, 0) e^{\frac{jk[(x_0-x)^2 + (y_0-y)^2]}{2z}} dx dy \quad (5.4-6)$$

当选择 $g(x, y) = U(x, y, 0)$ 时, 显然式 (5.4-5) 和式 (5.4-6) 就完全一致。因此, 菲涅耳近似是傍轴波动方程的一个解。

5.5 夫琅禾费衍射

对于式 (5.2-14), 如果有

$$z \gg \frac{k}{2}(x^2 + y^2)_{\max} = \frac{kL_1^2}{2} \quad (5.5-1)$$

则 $U'(x, y, 0)$ 近似等于 $U(x, y, 0)$ 。如果以上条件满足, 则夫琅禾费近似有效, 于是式 (5.2-13) 变为

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} \iint_{-\infty}^{+\infty} U(x, y, 0) e^{-j\frac{2\pi}{\lambda z}(xx_0 + yy_0)} dx dy \quad (5.5-2)$$

如果不考虑前边乘上去的幅值和相位因子, $U(x_0, y_0, z)$ 恰好是 $U(x, y, 0)$ 的二维傅里叶变换, 且对应的空间频率分别为 $f_x = x_0/\lambda z$, $f_y = y_0/\lambda z$ 。同时可以看到, 夫琅禾费区的衍射没有卷积形式。

例 5.8 求位于 $z=0$ 、直径为 D 的圆孔发出的光在夫琅禾费区的衍射场分布。

解: 圆孔直径为 D 意味着

$$U(x, y, 0) = \text{cyl}(r/D)$$

其中, $r = \sqrt{x^2 + y^2}$ 。

由例 2.6, $U(x, y, 0)$ 的二维傅里叶变换由下式给出:

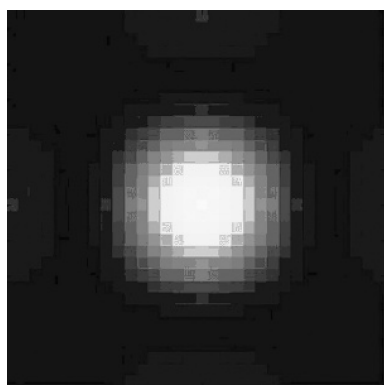


图 5.5 夫琅禾费区的方孔衍射条纹强度分布

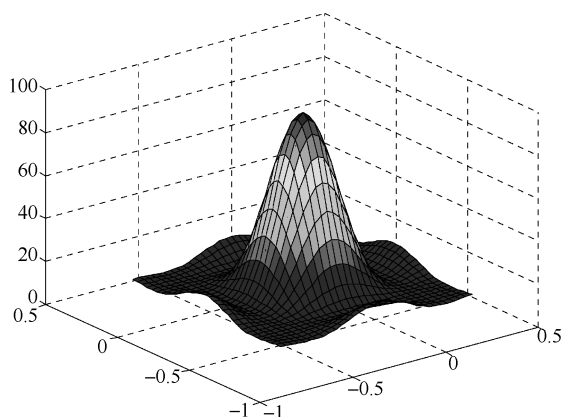


图 5.6 方孔夫琅禾费衍射条纹强度的三维空间分布



$$U(x, y, 0) \leftrightarrow \frac{D^2 \pi}{4} \text{somb} \left(D \sqrt{f_x^2 + f_y^2} \right)$$

其中空间频率分别为 $f_x = x_0 / \lambda z$, $f_y = y_0 / \lambda z$ 。由式 (5.5-2) 可得

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} \frac{D^2 \pi}{4} \text{somb} \left(D \sqrt{f_x^2 + f_y^2} \right)$$

条纹强度分布正比于 $U^2(x_0, y_0, z)$, 称之为爱里斑, 其为

$$I(x_0, y_0, z) = \left(\frac{kD^2}{8z} \right)^2 \left[\text{somb} \left(D \sqrt{f_x^2 + f_y^2} \right) \right]^2$$

对图 5.2 所示的方孔, 其夫琅禾费衍射的强度如图 5.5 和图 5.6 所示。图 5.5 是方孔夫琅禾费衍射条纹的强度分布, 图 5.6 是衍射条纹强度的三维空间分布。

例 5.9 设孔径宽度为 2.5cm, 入射光波长为 0.6μm (红光), 求当满足夫琅禾费衍射时需要的条件。

解:

$$z \gg \frac{kL_1^2}{2} = \frac{\pi L_1^2}{\lambda}$$

$$\frac{\pi L_1^2}{\lambda} = \frac{\pi (1.25 \times 10^{-2})^2}{0.6 \times 10^{-6}} \text{m} = 830 \text{m}$$

可见要实现夫琅禾费衍射, 衍射距离应很长。当使用透镜时, 衍射距离会极大地缩短, 这将在第 8 章讨论。

例 5.10 对例 5.7, 保持条件不变, 如果计算其夫琅禾费衍射, 其伪代码如下:

$$\begin{aligned} u1 &= \text{fftshift}(u1) \\ u2 &= \text{fft}(u1) \\ u3 &= \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x^2 + y^2)} * u2 \\ U &= \text{fftshift}(u3) \end{aligned}$$

5.6 衍射光栅

菲涅耳近似和夫琅禾费近似经常被用来分析和设计衍射光栅。衍射光栅主要被用在光谱分析、光谱成像、光通信和光网络等领域。

衍射光栅是一种光学器件, 可对入射波产生周期性的幅值或相位调制, 或同时对幅值和相位调制。由交替的不透明和透明区域构成的周期性阵列称为透射式幅值光栅。图 5.7 和图 5.8 给出了一种透射式幅值光栅。

如果整个光栅均是透明的, 仅是光栅的厚度做周期性变化, 这种光栅称为透射式相位光栅。对于具有周期性起伏表面的反射材料, 其对波的反射也可以产生显著的相位调制, 这种光栅称为反射式相位光栅。反射式衍射光栅还可以在玻璃基底上沉积周期性的平行铝膜来制作。

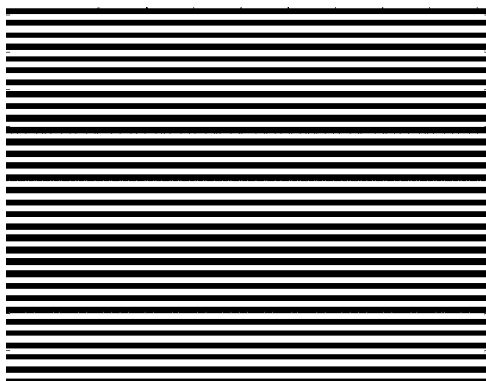


图 5.7 透射式幅值光栅的平面视图

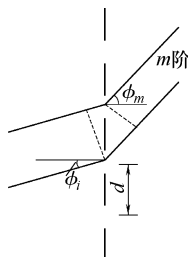


图 5.8 周期幅值光栅

衍射光栅是一种非常有效地将入射波进行高分辨率波长分光的光学元件。以最简单的情況为例，光栅可以看作是许多平行的、互相紧邻的间隔狭缝组成的。不考虑狭缝的数目，衍射光的峰值强度通常出现在一些特定的角度，而这些角度由下述的光栅方程给出：

$$d \sin \phi_m = m \lambda \quad (5.6-1)$$

其中， d 是相邻狭缝之间的间隔； m 是整数，称为衍射阶次； ϕ 为衍射角。

方程式 (5.6-1) 表明，不同波长的入射光，通过光栅后将沿着不同方向传播。各波长下衍射光的峰值强度依赖于光栅的狭缝数目。

光栅的透射函数定义为

$$t(x, y) = \frac{U(x, y, 0_+)}{U(x, y, 0_-)} \quad (5.6-2)$$

式中， $U(x, y, 0_-)$ 和 $U(x, y, 0_+)$ 分别是入射到光栅上的波函数和刚通过光栅后的波函数。

如果光栅是反射式的，则透射函数变成了反射函数。

一般来说， $t(x, y)$ 是复函数。如果 $t(x, y)$ 的辐角为 0，则光栅仅对入射波的幅值产生调制。如果 $t(x, y)$ 的模为常数，则光栅仅改变入射光的相位。

光栅有时被制造成用来反射光的元件，例如，通过对光栅的表面金属化即可实现。此时，透射函数即变为反射函数。

厚度周期性变化的透明薄片，或者折射率周期性变化的透明薄片，都可以用来对入射光的相位进行调制。反射式衍射光栅可以通过在玻璃基底上沉积周期性的铝膜来制作。

5.7 正弦幅值光栅的夫琅禾费衍射

设单位幅值的平面波垂直入射到光栅上，且光栅的透射函数与函数 $U(x, y, 0)$ 相同

$$U(x, y, 0) = \left[\frac{1}{2} + \frac{m}{2} \cos(2\pi f_0 x) \right] \text{rect}\left(\frac{x}{D}\right) \text{rect}\left(\frac{y}{D}\right) \quad (5.7-1)$$

其中，假设衍射光栅外边界是边长为 D 的正方形。 $U(x, y, 0)$ 可改写为

$$U(x, y, 0) = f(x, y) g(x, y) \quad (5.7-2)$$

其中



$$f(x, y) = \frac{1}{2} + \frac{m}{2} \cos(2\pi f_0 x) \quad (5.7-3)$$

$$g(x, y) = \text{rect}\left(\frac{x}{D}\right) \text{rect}\left(\frac{y}{D}\right) \quad (5.7-4)$$

$f(x, y)$ 和 $g(x, y)$ 的傅里叶变换分别为

$$F(f_x, f_y) = \frac{1}{2} \delta(f_x, f_y) + \frac{m}{4} [\delta(f_x + f_0, f_y) + \delta(f_x - f_0, f_y)] \quad (5.7-5)$$

$$G(f_x, f_y) = D^2 \text{sinc}(Df_x) \text{sinc}(Df_y) \quad (5.7-6)$$

而 $U(x, y, 0)$ 的傅里叶变换为 $F(f_x, f_y)$ 和 $G(f_x, f_y)$ 的卷积

$$U(x, y, 0) \leftrightarrow F(f_x, f_y) * G(f_x, f_y) \quad (5.7-7)$$

其中

$$\begin{aligned} & F(f_x, f_y) * G(f_x, f_y) \\ &= \frac{D^2}{2} \text{sinc}(Df_y) \left[\text{sinc}(Df_x) + \frac{m}{2} \{ \text{sinc}[D(f_x + f_0)] + \text{sinc}[D(f_x - f_0)] \} \right] \end{aligned} \quad (5.7-8)$$

于是, z 处的场分布可由下式给出:

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} F(f_x, f_y) * G(f_x, f_y) \quad (5.7-9)$$

式中, $f_x = x_0/\lambda z$, $f_y = y_0/\lambda z$ 。

5.8 正弦幅值光栅的菲涅耳衍射

为了计算正弦幅值光栅的菲涅耳衍射, 可以利用式 (5.2-9) 所给出的衍射积分的卷积形式。系统的传递函数为

$$H(f_x, f_y) = e^{ikz} e^{-j\pi\lambda z} (f_x^2 + f_y^2) \quad (5.8-1)$$

式中, e^{ikz} 为常数相位因子。

透射函数的傅里叶变换由式 (5.7-7) 给出。为简化计算, 可假设衍射光栅的口径足够大, 因此衍射光栅的透射函数的傅里叶变换可近似为式 (5.7-5)。于是, 有贡献的空间频率为

$$(f_x, f_y) = (0, 0), (-f_0, 0), (f_0, 0)$$

传播的波动场具有如下的傅里叶变换形式:

$$U(x_0, y_0, z) \leftrightarrow \frac{1}{2} \delta(f_x, f_y) + \frac{m}{4} [e^{-j\pi\lambda z f_0^2} \delta(f_x - f_0, f_y) + e^{-j\pi\lambda z f_0^2} \delta(f_x + f_0, f_y)] \quad (5.8-2)$$

由逆傅里叶变化可得到

$$\begin{aligned} U(x_0, y_0, z) &= \frac{1}{2} + \frac{m}{4} e^{-j\pi\lambda z f_0^2} (e^{-j2\pi f_0 x_0} + e^{j2\pi f_0 x_0}) \\ &= \frac{1}{2} [1 + m e^{-j\pi\lambda z f_0^2} \cos(2\pi f_0 x_0)] \end{aligned} \quad (5.8-3)$$

场的强度分布为

$$I(x_0, y_0, z) = \frac{1}{4} [1 + 2m \cos(\pi\lambda z f_0^2) \cos(2\pi f_0 x_0) + m^2 \cos^2(2\pi f_0 x_0)] \quad (5.8-4)$$



例 5.11 求距离 z 处衍射条纹的强度分布, z 满足下述条件:

$$z = \frac{2n}{\lambda f_0^2} (n \text{ 为整数})$$

解:

$I(x_0, y_0, z)$ 可表示为

$$I(x_0, y_0, z) = \frac{1}{4} [1 + m \cos(2\pi f_0 x_0)]^2$$

这恰好与光栅的形状完全一致, 这种成像称为光栅的自成像, 又称泰伯像。

例 5.12 重复上例, 但 z 满足下述条件

$$z = \frac{2n+1}{\lambda f_0^2} (n \text{ 为整数})$$

解:

$I(x_0, y_0, z)$ 可表示为

$$I(x_0, y_0, z) = \frac{1}{4} [1 - m \cos(2\pi f_0 x_0)]^2$$

可见, 上述强度分布也是光栅的像, 但是产生了 180° 的相移, 也就是所谓的对比度反转。这些位置的像也称为泰伯像。

例 5.13 重复上例, 但 z 满足下述条件:

$$z = \frac{2n-1}{2\lambda f_0^2} (n \text{ 为整数})$$

解: $I(x_0, y_0, z)$ 可表示为

$$I(x_0, y_0, z) = \frac{1}{4} \left[1 + \frac{m^2}{2} + \frac{m^2}{2} \cos(4\pi f_0 x_0) \right]$$

该图像强度分布的变化频率是光栅频率的两倍, 即 $2f_0$, 这些位置的成像称为泰伯子像。

5.9 正弦相位光栅的夫琅禾费衍射

如同正弦幅值光栅, 对垂直入射到光栅上的平面波, 正弦相位光栅的透射函数为

$$U(x, y, 0) = e^{j\frac{m}{2}\sin(2\pi f_0 x)} \text{rect}\left(\frac{x}{D}\right) \text{rect}\left(\frac{y}{D}\right) \quad (5.9-1)$$

且 $U(x, y, 0)$ 可改写为

$$U(x, y, 0) = f(x, y) g(x, y) \quad (5.9-2)$$

其中

$$f(x, y) = e^{j\frac{m}{2}\sin(2\pi f_0 x)} \quad (5.9-3)$$

$$g(x, y) = \text{rect}\left(\frac{x}{D}\right) \text{rect}\left(\frac{y}{D}\right) \quad (5.9-4)$$

利用以下恒等式可简化分析:



$$e^{j\frac{m}{2}\sin(2\pi f_0 x)} = \sum_{k=-\infty}^{\infty} J_k\left(\frac{m}{2}\right) e^{j2\pi f_0 kx} \quad (5.9-5)$$

式中, $J_k(\cdot)$ 为 k 阶的第一类贝塞尔函数。

利用上述恒等式, $f(x, y)$ 的傅里叶变换可写成如下形式:

$$F(f_x, f_y) = \sum_{k=-\infty}^{\infty} J_k\left(\frac{m}{2}\right) \delta(f_x - kf_0, f_y) \quad (5.9-6)$$

于是, $U(x, y, 0)$ 的傅里叶变换可表示为

$$U(x, y, 0) \leftrightarrow F(f_x, f_y) * G(f_x, f_y) \quad (5.9-7)$$

其中

$$F(f_x, f_y) * G(f_x, f_y) = D^2 \sum_{k=-\infty}^{\infty} J_k\left(\frac{m}{2}\right) \text{sinc}[D(f_x - kf_0)] \text{sinc}(Df_y) \quad (5.9-8)$$

因此

$$U(x_0, y_0, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x_0^2 + y_0^2)} F(f_x, f_y) * G(f_x, f_y) \quad (5.9-9)$$

可见, 第 5.7 节中的正弦幅值光栅有三个能量比较集中衍射阶次, 这主要是因为有三个 sinc 函数, 当光栅的频率远大于 $2/D$ 时, 这些函数之间不会相互重叠。相反, 正弦相位光栅有许多能量集中的衍射阶次。幅值光栅的中心阶次占有主要的能量, 但当 $J_0\left(\frac{m}{2}\right) = 0$ 时, 相位光栅的中心阶次将会趋近于消失。

5.10 狭缝衍射光栅

衍射光栅也经常做成狭缝状。图 5.9 给出了一个双缝衍射光栅, 其中, b 是狭缝的宽度, ϕ 为衍射角, d 为两个狭缝之间的中心距离。

通过上述光栅后, 夫琅禾费区的衍射场的强度分布由下式给出:

$$I(\phi) = 2I_0 \left[\frac{\sin(kb\phi/2)}{kb\phi/2} \right]^2 [1 + \cos(k\Delta + kd\phi)] \quad (5.10-1)$$

式中, Δ 为来自相邻两个狭缝的光束之间在衍射面上的光程差; I_0 为入射到光栅前的平面波发光强度。

如果光栅有 N 个平行的狭缝, 则夫琅禾费区的衍射条纹分布为

$$I(\phi) = 2I_0 \left[\frac{\sin(kb\phi/2)}{kb\phi/2} \right]^2 \left[\frac{\sin(Nkd\phi/2)}{\sin(kd\phi/2)} \right]^2 \quad (5.10-2)$$

当光栅方程 $d\sin\phi_m = m\lambda$ 满足时, 上式第二个因子等于 $N\cos(\pi Nm)/\cos(\pi m)$ 。

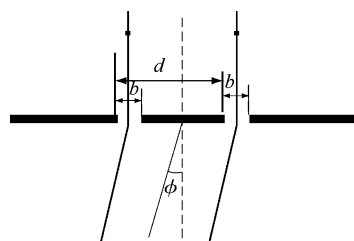


图 5.9 双缝衍射光栅

第 6 章 逆 衍 射

6.1 本章引论

逆衍射是指利用测量到的物体的衍射图样来复原原物体的图像，测量到的衍射图样可以是在一个平面上测得的。在非涅耳及夫琅禾费近似的情况下，逆衍射可以直接得到。在近场情况下，可以利用角谱分析来计算逆衍射，但是有一些技术细节需要进行讨论。

本章后续各节需要用到的几何模型如图 6.1 所示。假设观察平面和测量平面分别位于 $z = z_0$ 和 $z = z_r$ 处，场从左至右传播。事先假定 z_r 等于 0，则上述两个平面间的距离用 z_{0r} 来表示。假设光场于均匀媒质中传播，于是需要解决的问题就是，假设知道 $z = z_r$ 处的场分布，反求 $z = z_0$ 处的场分布。

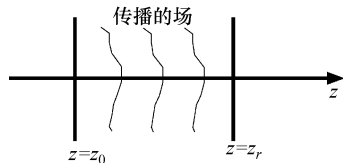


图 6.1 逆衍射的几何结构

类似于所有逆问题，当把隐失波引入逆问题的解的情况下，逆衍射问题实际上是非常困难的 [Vesperinas, 1991]，这时逆衍射问题包含了一个奇异核 [Shewell, Wolf, 1968]。如果像后述章节那样不考虑隐失波，则逆衍射问题将变得比较容易处理。

本章共包含四节。6.2 节介绍菲涅耳和夫琅禾费近似的逆衍射，6.3 节介绍角谱表述下的逆衍射问题，6.4 节对 6.3 节的结果进行进一步的分析。

6.2 菲涅耳和夫琅禾费近似的逆衍射问题

菲涅耳衍射由式 (5.2-13) 给出。由于该积分是菲涅耳变换，因此其逆衍射由下述公式给出：

$$U(x, y, z_r) = \frac{j e^{-j k z_{0r}}}{\lambda z_{0r}} e^{-j \frac{k}{2 z_{0r}} (x^2 + y^2)} \iint_{-\infty}^{\infty} U(x_0, y_0, z_0) e^{-j \frac{k}{2 z_{0r}} (x_0^2 + y_0^2)} e^{j \frac{2\pi}{\lambda z_{0r}} (x x_0 + y y_0)} dx_0 dy_0 \quad (6.2-1)$$

式中， z_r 和 z_0 分别为输入和输出面的位置变量，并有 $z_{0r} = z_0 - z_r$ 。

夫琅禾费衍射由式 (5.5-2) 给出，其逆衍射由下式给出：

$$U(x, y, z_r) = \frac{j e^{-j k z_{0r}}}{\lambda z_{0r}} \iint_{-\infty}^{\infty} U(x_0, y_0, z_0) e^{-j \frac{k}{2 z_{0r}} (x_0^2 + y_0^2)} e^{j \frac{2\pi}{\lambda z_{0r}} (x x_0 + y y_0)} dx_0 dy_0 \quad (6.2-2)$$

6.3 角谱表述下的逆衍射

为简单起见，先讨论二维的情况，然后推广到三维情形。

二维情况下式 (4.3-12) 可写为

$$U(x_0, z_0) = \int_{-\infty}^{+\infty} A(f_x, z_r) e^{j z_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{j 2\pi f_x x_0} df_x \quad (6.3-1)$$



由于这是一个傅里叶积分, 对 $A(f_x, z_r)$ 的复原就是计算 $U(x_0, z_0)$ 的傅里叶变换

$$A(f_x, z_r) = e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} \int_{-\infty}^{\infty} U(x_0, z_0) e^{-j2\pi f_x x_0} dx_0 \quad (6.3-2)$$

$U(x, z_r)$ 是 $A(f_x, z_r)$ 的逆傅里叶变换

$$U(x, z_r) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} U(x_0, z_0) e^{-j2\pi f_x x_0} dx_0 \right] e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{j2\pi f_x x} df_x \quad (6.3-3)$$

用 $F[\cdot]$ 和 $F^{-1}[\cdot]$ 表示傅里叶变换和逆傅里叶变换, 则式 (6.3-3) 可表示为

$$U(x, z_r) = F^{-1} \{ F[U(x_0, z_0)] e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} \} \quad (6.3-4)$$

在三维情形下, 对应的公式变为

$$U(x, y, z_r) = F^{-1} \{ F[U(x_0, y_0, z_0)] e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 (f_x^2 + f_y^2)}} \} \quad (6.3-5)$$

注意到变量 z_r 可不断变化, 从而可以重构出不同深度的场分布。这样, 就得到了三维情形下场的重构。

如果测量使用的是非单色波, 这将产生随时间 t 变化的场 $U(x_0, y_0, z_0, t)$, 这时可以采用含有时间频率分量的 $U(x_0, y_0, z_0, f)$, 其计算如下:

$$U(x_0, y_0, z_0, f) = \int_{-\infty}^{+\infty} U(x_0, y_0, z_0, t) e^{-j2\pi ft} dt \quad (6.3-6)$$

将上述公式中的变量 f 利用如下的关系进行替换:

$$k = \frac{2\pi}{\lambda} = 2\pi \frac{f}{v} \quad (6.3-7)$$

式中, v 为相位速度。

上述讨论的技术已经被用在超声和地震波的像重构中 [Boyer, 1971; Boyer et al., 1970; Ljunggren, 1980]。特别是当 z_{0r} 太小以至于其他的衍射积分的近似失效时, 例如, 第 5 章中讨论的菲涅耳和夫琅禾费近似, 上述方法就显得特别有用。

6.4 分析

改变式 (6.3-3) 中的积分顺序, 可以得到

$$U(x, z_r) = \int_{-\infty}^{+\infty} U(x_0, z_0) B(x, x_0) dx_0 \quad (6.4-1)$$

其中

$$B(x, x_0) = \int_{-\infty}^{\infty} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{-j2\pi f_x (x - x_0)} df_x \quad (6.4-2)$$

当 $2\pi f_x > k$ 时, 后一个积分的传播子变为 $e^{kz_{0r} \sqrt{4\pi^2 f_x^2 - k^2}}$, 且式 (6.4-2) 中的积分发散。这主要是由于引入了隐失波。实际上, 当积分区间被限制在 $k > 2\pi f_x$ 的区域时, 即可以得到足够精确的结果 [Lalor, 1968]。于是, f_x 被限制在如下范围:

$$|f_x| \leq \frac{1}{\lambda} \quad (6.4-3)$$

假设 $z = z_r$ 处的精确场分布为 $T(x, z_r)$ 。如何对 $T(x, z_r)$ 和上述计算得到的 $U(x, z_r)$ 进行比较? 为回答这个问题, 可以首先由衍射求出利用 $T(x, z_r)$ 表示的 $U(x_0, z_0)$, 其由下式给出:



$$U(x_0, z_0) = \int_{-\infty}^M \left[\int_{-\infty}^{+\infty} T(x, z_r) e^{-j2\pi f_x x} dx \right] e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{j2\pi f_x x_0} df_x \quad (6.4-4)$$

其中用到了 $|f_x| \leq M$, 且积分上界 M 是 $1/\lambda$ 。

把以上结果代入式 (6.4-3), 并把 f_x 限制在 $|f_x| \leq Q$ 范围内, 则有

$$U(x, z_r) = \int_{-Q}^Q \left\{ \int_{-M}^M \left[\int_{-\infty}^{+\infty} T(x, z_r) e^{-j2\pi f_x' x} dx \right] e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x'^2}} e^{j2\pi f_x' x_0} df_x' \right\} \cdot e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{j2\pi f_x x} df_x \quad (6.4-5)$$

交换积分顺序, 式 (6.4-5) 可写为 [Van Rooy, 1971]

$$U(x, z_r) = \int_{-\infty}^{+\infty} \left[\int_{-M}^M df_x' e^{-j2\pi f_x' x} e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x'^2}} \cdot \left(\int_{-Q}^Q df_x e^{-j2\pi f_x x} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} \left[\int_{-\infty}^{+\infty} dx' e^{j2\pi x' (f_x' - f_x)} \right] \right) \right] T(x, z_r) dx$$

由于

$$\delta(f_x - f_x') = \int_{-\infty}^{+\infty} e^{j2\pi x (f_x' - f_x)} dx$$

于是式 (6.4-5) 变为

$$U(x, z_r) = \int_{-\infty}^{+\infty} \left[\int_{-M}^M e^{-j2\pi f_x' x} e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x'^2}} df_x' \int_{-Q}^Q \delta(f_x - f_x') e^{-j2\pi f_x x} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} df_x \right] T(x, z_r) dx \quad (6.4-6)$$

依赖于 M 和 Q 值的大小, 以上积分结果有两种可能的情况, 分别讨论如下。

情况 I: $Q \geq M$

此时, 式 (6.4-6) 退化为

$$\begin{aligned} U(x, z_r) &= \int_{-\infty}^{+\infty} \int_{-M}^M e^{j2\pi f_x' (x - x_0)} df_x' T(x, z_r) dx \\ &= T(x, z_r) * \text{sinc}\left(\frac{2Mx}{\lambda}\right) \\ &= T(x, z_r) \end{aligned} \quad (6.4-7)$$

因为 $T(x, z_r)$ 是带限的。

情况 II: $Q < M$

此时, 式 (6.4-6) 退化为

$$U(x, z_r) = T(x, z_r) * \text{sinc}\left(\frac{2Qx}{\lambda}\right) \quad (6.4-8)$$

式 (6.4-8) 表明 $U(x, z_r)$ 为 $T(x, z_r)$ 的低通滤波, 因为 sinc 函数的傅里叶变换是矩形函数。换言之, 频率大于 Q/λ 的分量将被滤掉。重构像可能具有光滑的边缘, 从而损失掉细节信息。

由于实际中 Q 通常被限制在 $Q \leq 1/\lambda$, 因此重构像可能得到的分辨率是 λ 。这和经典理论的结果一致, 经典理论认为线性成像系统不能实现小于一个波长分辨率的成像。

例 6.1 当 $U(x', z_0)$ 在孔径 $|x'| \leq R$ 处测量得到时, 求 $U(x, z_r)$ 。

解: 在这种情况下, 式 (6.4-6) 变为



$$U(x, z_r) = \int_{-Q}^Q \left\{ e^{-jkz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} \int_{-R}^R dx_0 e^{-j2\pi f_x x_0} \cdot \left(\left[\int_{-M}^M df'_x e^{j2\pi f'_x x_0} \left[\int_{-\infty}^{+\infty} dx T(x, z_r) e^{-j2\pi f'_x x} \right] \right] e^{j2\pi f_x x} \right) \right\} df_x$$

交换积分顺序，上式可改写为

$$U(x, z_r) = \int_{-\infty}^{+\infty} \left[\int_{-M}^M df'_x \left(e^{-j2\pi f'_x x} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} \times \int_{-Q}^Q df_x \left[e^{j2\pi f_x x} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} 2R \text{sinc}(2Rf_x) \right] \right) \right] T(x, z_r) dx$$

由上式可见，sinc 函数混叠到原始物光波的平面波分量中。这意味着原始物光波中的每个平面波分量被孔径 $(-R, R)$ 截断，因此其由于衍射的作用“扩散”开来。

例 6.2 令 $M = Q = 1/\lambda$ ，试证明此时可以利用 $U^*(x_0, z_0)$ 的前向传播代替逆向传播，从 $U(x_0, z_0)$ 精确地复原 $T^*(x, z_r)$ 。

解：

假设从平面 $z = z_0$ 到平面 $z = 2z_0 - z_r$ 的前向传播已经被计算出来，由于

$$2z_0 - z_r - z_0 = z_{0r}$$

则前向传播公式由下式给出：

$$U(x, 2z_0 - z_r) = \int_{-Q}^Q df_x \int_{-\infty}^{+\infty} U^*(x_0, z_0) e^{-j2\pi f_x x_0} e^{jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{j2\pi f_x x} dx_0$$

重排积分的次序，可以得到

$$U(x, 2z_0 - z_r) = \int_{-Q}^Q \left[\int_{-\infty}^{+\infty} U(x_0, z_0) e^{j2\pi f_x x_0} e^{-jz_{0r} \sqrt{k^2 - 4\pi^2 f_x^2}} e^{-j2\pi f_x x} dx_0 \right]^* df_x = T^*(x, 0)$$

换言之，可以用记录的场分布的逆向传播来重构原始像，也可以用记录的场分布的复共轭的正向传播来重构原始像。

第 7 章 宽角度下标量衍射理论的近场和远场近似

7.1 本章引论

在涉及波动传播的许多领域，正向或逆向衍射积分的近似计算是至关重要的。如第 4 章和第 5 章所讨论的，一些近似方法，如菲涅耳近似、夫琅禾费近似和严格的角谱法（Angular Spectrum Method, ASM），均用到了傅里叶变换、离散傅里叶变换和快速傅里叶变换（FFT）。

当离输入面的距离合适时，菲涅耳近似有效，而夫琅禾费近似在远场有效。角谱法是亥姆霍兹方程的严格解，其数值计算经常采用 FFT 来实现，也可以利用其他的相应近似下的数字信号处理算法进行数值计算 [Mellin and Nordin, 2001; Shen and Wang, 2006]。

当衍射孔径的尺度小于波长时，利用标量衍射理论将导致不可忽视的误差，而需要采用其他的数值方法，例如，时域有限差分法（Finite Difference Time Domain, FDTD）和有限元法（Finite Element Method, FEM）[Kunz, 1993; Taflov and Hagness, 2005]。但是，和 FFT 方法相比，上述方法在大尺度模拟时不再适用。在近场条件下，当衍射孔径的尺度是波长的几个数量级时，ASM 法可以给出满意的结果 [Mellin and Nordin, 2001]。ASM 法的一个缺点是输出面和输入面的尺寸必须是一样的。在许多应用中，输出面的尺寸一般比输入面的尺寸要大许多，此时就需要在短距离上重复地利用 ASM 法，并且利用滤波方法来使得输出面尺寸不断增大。

实际中，可以认为菲涅耳近似、夫琅禾费近似和角谱法在小角度衍射下是有效的。对于近场的宽角度衍射和远场衍射，非常需要一种有效的近似方法，而且这种方法最好也是基于傅里叶积分，这样就可以在合理的时间和存储要求下利用 FFT 来进行大尺度的计算。即使空间频率的采样使得 FFT 不能应用，傅里叶积分表述仍然是有意义的，因为傅里叶积分的两个积分变量是可分离的，它可以将四维的张量运算降为二维的矩阵运算。

本章将讨论一套具有上述特点的近似方法。这些近似方法主要基于输出面上的点离原点的径向距离进行泰勒展开。利用这些近似，使输出面上的半不规则采样可以采用 FFT 来计算。半不规则采样是指首先得到规则的采样阵列点，然后对阵列点加入扰动，使得其满足 FFT 条件。

即使我们提出的近似方法在近场和宽角度的条件下有效，菲涅耳近似在这些条件下仍是无效的，但是我们可以得到一些有趣的结果，这些结果显示我们提出的近似方法和菲涅耳近似具有某些相似性。这就是为什么菲涅耳近似经常在许多近场应用中被采用的原因之一。从这些研究中得到的结果，实际上对于输出采样点的位置紊乱的情况也是有效的。

本章包含 8 节。7.2 节是对第 5 章讨论过的菲涅耳近似和夫琅禾费近似的回顾。7.3 节介绍了一种径向近似的新方法。7.4 节中对提出的方法进行了进一步的改进和误差分析。



7.5 节指出如何将提出的方法应用于逆衍射和迭代优化的应用中。7.6 节给出二维和三维情形下的数值模拟算例。7.7 节介绍了通过将输出平面和输出平面的孔径置于特定中心坐标系的中心，以及通过在大尺度模拟中可能采用的更小的子区域，来提高计算精度的方法。7.8 节对本章进行总结。

7.2 菲涅耳和夫琅禾费近似的回顾

瑞利-索末菲衍射积分可以写为

$$U(x_0, y_0, z) = \frac{1}{j\lambda} \iint_{\infty} U(x, y, 0) \frac{z}{r^2} e^{jkr} dx dy \quad (7.2-1)$$

式中， $U(x, y, 0)$ 是输入场； $U(x_0, y_0, z)$ 是输出场； k 是波数，且 $k = 2\pi/\lambda$ ； λ 为波长； r 为从点 $(x, y, 0)$ 到点 (x_0, y_0, z) 的矢量长度。

一般来说， z 通常足够大，因此可以用 $1/z$ 来近似 z/r^2 。式 (7.2-1) 可变为

$$U(x_0, y_0, z) = \frac{1}{j\lambda z} \iint_{\infty} U(x, y, 0) e^{jkr} dx dy \quad (7.2-2)$$

常数项 $1/j\lambda z$ 在本章后续部分将忽略掉。

对于在衍射问题中常用的菲涅耳近似，两点之间的距离 r 可利用泰勒展开的前两项来近似 [Mezouari and Harvery, 2003; Southwell, 1981; Steane and Rutt, 1989]

$$r = \sqrt{z^2 + (x_0 - x)^2 + (y_0 - y)^2} \cong z \left(1 + \frac{g}{2} \right) \quad (7.2-3)$$

其中

$$g = \frac{(x_0 - x)^2 + (y_0 - y)^2}{z^2} \quad (7.2-4)$$

由 r 的近似引入的最大相位误差 E_{\max} 的大小（以弧度表示）可由泰勒展开级数的第三项来估计

$$E_{\max} \leq \frac{1}{8} z g^2 k \quad (7.2-5)$$

一般来说，当最大相位误差 E_{\max} 的值大于 1rad 时，对于衍射问题采用菲涅耳近似将不能得到足够好的结果。但是，对于各种实际问题，上述条件一般不能达到，因此采用菲涅耳近似一般可以得到较好的结果。其原因可以部分地用稳相法来解释 [Goodman]。本章将给出上述结果的另一种解释。

菲涅耳近似的优点是可以实现变量的分离，因此式 (7.2-2) 中的积分可以写为傅里叶积分的形式。例如，式 (7.2-4) 可改写为

$$g = \frac{v}{z^2} + \frac{w}{z^2} - \frac{2(x_0 x + y_0 y)}{z^2} \quad (7.2-6)$$

其中

$$v = x_0^2 + y_0^2 \quad (7.2-7)$$

$$w = x^2 + y^2 \quad (7.2-8)$$

于是，利用式 (7.2-3)，式 (7.2-2) 可近似改为下列傅里叶积分：



$$U(x_0, y_0, z) = e^{jk\frac{w}{2z}} \iint_{-\infty}^{\infty} u'(x, y) e^{-\frac{jk}{z}(x_0x + y_0y)} dx dy \quad (7.2-9)$$

其中

$$u'(x, y) = u(x, y) e^{jk\frac{w}{2z}} \quad (7.2-10)$$

式 (7.2-9) 可进一步写成下列傅里叶积分的形式:

$$U(x_0, y_0, z) = e^{jk\frac{w}{2z}} \iint_{-\infty}^{\infty} u'(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (7.2-11)$$

其中空间频率 f_x 和 f_y 定义为

$$f_x = \frac{x_s}{\lambda z} \quad (7.2-12)$$

$$f_y = \frac{y_s}{\lambda z} \quad (7.2-13)$$

其中对点 (x_0, y_0, z) , 有 $x_s = x_0$, $y_s = y_0$ 。

在夫琅禾费近似中, 由 $z \gg w$ 可得到 $u'(x, y) \approx u(x, y)$, 于是式 (7.2-9) 变为

$$U(x_0, y_0, z) = e^{jk\frac{w}{2z}} \iint_{-\infty}^{\infty} u(x, y) e^{-\frac{jk}{z}(x_0x + y_0y)} dx dy \quad (7.2-14)$$

式 (7.2-14) 可利用快速傅里叶变换来估算 [Brigham, 1974], 令

$$x_s = m_x \Delta_{0x}, \quad m_x = \text{整数} \quad (7.2-15)$$

$$y_s = m_y \Delta_{0y}, \quad m_y = \text{整数} \quad (7.2-16)$$

$$x = n_x \Delta_x, \quad n_x = \text{整数} \quad (7.2-17)$$

$$y = n_y \Delta_y, \quad n_y = \text{整数} \quad (7.2-18)$$

$$\Delta_{0x} \Delta_x = \frac{\lambda z}{M_x} \quad (7.2-19)$$

$$\Delta_{0y} \Delta_y = \frac{\lambda z}{M_y} \quad (7.2-20)$$

式中, M_x 和 M_y 分别为 x 和 y 方向的采样数。

在其他的所有近似中, 如果需要使用 FFT, 则需要保证满足式 (7.2-19) 和式 (7.2-20)。

7.3 径向近似

在许多实际应用中, 例如, 涉及数字全息和衍射光学器件的应用中, 坐标 $(x, y, 0)$ 通常要远小于物体坐标 (x_0, y_0, z) 。于是, 式 (7.2-3) 可改写为

$$r = r_0 \sqrt{1 + h} \quad (7.3-1)$$

其中

$$r_0 = \sqrt{z^2 + x_0^2 + y_0^2} \quad (7.3-2)$$

$$h = \frac{w}{r_0^2} - \frac{2(x_0x + y_0y)}{r_0^2} \quad (7.3-3)$$

为了能利用傅里叶变换, 式 (7.3-3) 可近似为



$$h = \frac{w}{z^2} - \frac{2(x_0x + y_0y)}{r_0^2} \quad (7.3-4)$$

最简单的径向近似可由下边的径向距离 r 的近似得到：

$$r \approx r_0 + \frac{w^2}{2z} - \frac{xx_0 + yy_0}{r_0} \quad (7.3-5)$$

当 x 、 y 、 x_0 和 y_0 充分小于 z 时（充分条件将在后面详细介绍），式 (7.3-4) 和式 (7.3-5) 的近似均是有效的。在下一节将给出式 (7.3-5) 的进一步改进。为了由式 (7.2-9) 给出的衍射积分得到傅里叶变换关系，可定义如下关系：

$$x_0 = \frac{x_s}{z} r_0 \quad (7.3-6)$$

$$y_0 = \frac{y_s}{z} r_0 \quad (7.3-7)$$

其中 (x_s, y_s) 是规则的输出采样点。联立式 (7.3-6) 和式 (7.3-7)，有

$$x_0 = C(x_s, y_s) x_s \quad (7.3-8)$$

$$y_0 = C(x_s, y_s) y_s \quad (7.3-9)$$

其中

$$C(x_s, y_s) = \left(1 - \frac{x_s^2 + y_s^2}{z^2}\right)^{-\frac{1}{2}} \quad (7.3-10)$$

如夫琅禾费近似，当 $z \gg w$ 时，式 (7.3-5) 中包含 w^2 的项可以忽略掉。这样得到的近似称之为近场近似（NFA），下面要讲到的近似称之为远场近似（FFA）。一般将 NFA 和 FFA 统称为 NFFA。

利用上述新近似，式 (7.2-11) 可由下式替换：

$$U(x_0, y_0, z) = e^{ikr_0} \iint_{\infty} u'(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (7.3-11)$$

其中，根据式 (7.3-6) 和式 (7.3-7)，空间频率仍然由式 (7.2-12) 和式 (7.2-13) 给出。根据远场近似， $u'(x, y)$ 等于 $u(x, y)$ 。

注意到式 (7.3-11) 与菲涅耳近似下的式 (7.2-11) 形式上类似，但是积分符号外的相位因子和输出面的采样点是不同的。当采样点由式 (7.3-8) 和式 (7.3-9) 定义时，式 (7.3-11) 可直接用 FFT 计算，且得到的结果是有效的。空间频率也可写成如下形式：

$$f_x = \frac{x_0}{\lambda r_0} \quad (7.3-12)$$

$$f_y = \frac{y_0}{\lambda r_0} \quad (7.3-13)$$

可见，对规则采样点 (x_s, y_s) ，空间频率仍然由式 (7.2-12) 和式 (7.2-13) 给出，而对实际采样点 (x_0, y_0) ，则由式 (7.3-12) 和式 (7.3-13) 给出。

7.4 高阶改进与分析

将式 (7.3-1) 中的 r 展开成泰勒级数形式



$$r = r_0 \left[1 + \frac{h}{2} - \frac{h^2}{8} + \frac{h^3}{16} - \cdots + (-1)^{n-1} \frac{1 \times 3 \times 5 \times \cdots \times (2n-1)}{2 \times 4 \times 6 \times \cdots \times 2n} h^{2n-1} + \cdots \right] \quad (7.4-1)$$

式 (7.4-1) 中给出了为提高精度而可取的项。为能利用傅里叶变换, 我们只保留泰勒级数中那些不含因子 $(x_0 x)^n$ 的项 ($x_0 x$ 除外)。

定义

$$a_n = (-1)^{n-1} \frac{1 \times 3 \times 5 \times \cdots \times (2n-1)}{2 \times 4 \times 6 \times \cdots \times 2n} \frac{1}{z^{2n}} \quad (7.4-2)$$

于是 r 由下式近似:

$$r \approx r_0 - \frac{xx_0 + yy_0}{r_0} + p(w) \quad (7.4-3)$$

其中

$$p(w) = p(x, y) = \sum_{n=1}^{\infty} \frac{a_n}{z^{2n-1}} w^n \quad (7.4-4)$$

式 (7.3-8) 和式 (7.3-9) 保持不变, 这主要是含有 xx_0 和 yy_0 的项没有发生改变。当满足类似于夫琅禾费近似中的 $z \gg w$ 时, 式 (7.4-3) 中的 $p(w)$ 项可直接忽略。利用这个新的近似, 式 (7.3-11) 可由下式代替:

$$U(x_0, y_0, z) = e^{jkz_0} \iint_{\infty} u''(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy \quad (7.4-5)$$

其中

$$u''(x, y) = u(x, y) e^{jkp(w)} \quad (7.4-6)$$

对于给定的 z , 菲涅耳近似给出的误差正比于 $(x_0^2 + y_0^2)^2$, 而在 NFFA 中, 近似误差由下式给出:

$$E = -\frac{v(x^2 + y^2)}{4z^3} - \frac{(x_0 x + y_0 y)}{2z^3} \quad (7.4-7)$$

此误差对应于泰勒展开中第二个分量 $-u^2/2$ 中被忽略的项。

注意到式 (7.3-8) 和式 (7.3-9) 中, 仅对物体坐标系做了改变。否则, 式 (7.4-5) 的积分则与式 (7.2-9) 或式 (7.2-14) 的积分一样。根据式 (7.2-15) 和式 (7.2-16) 给出的坐标 (x_s, y_s) , 输出点坐标 (x_0, y_0) 由式 (7.3-8) 和式 (7.3-9) 给出。

一个有意思的现象是, 在式 (7.3-8) 和式 (7.3-9) 给出的近场点 (x_0, y_0) 上, 文献中给出的菲涅耳近似的例子是正确的, 而对于点 (x_s, y_s) 则不再正确。因此, 上述结果在近场和远场都适用, 是真正的 NFFA 结果。

为使用 FFT, 需要满足式 (7.2-19) 和式 (7.2-20)。 $m_x \Delta_{0x}$ 与 $m_y \Delta_{0y}$ 是做了修正后的输出采样点。因此有可能考虑其他的采样策略, 将 $m_x \Delta_{0x}$ 与 $m_y \Delta_{0y}$ 用一些其他的采样点 x_s 和 y_s 替代。上述结果在所有情况中都是有效的。

7.5 逆衍射与迭代优化

正如第6章所讨论的, 逆衍射主要涉及利用 $U(x_0, y_0, z)$ 来复原 $u(x, y)$ 。在菲涅耳衍射的情况下, 式 (7.2-9) 是一个傅里叶变换, 其逆变换为



$$u'(x, y) = \iint_{\infty} \hat{U}(x_0, y_0, z) e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (7.5-1)$$

其中 $u'(x, y)$ 由式 (7.2-10) 给出, 空间频率 f_x, f_y 由式 (7.2-12) 和式 (7.2-13) 给出, 且

$$\hat{U}(x_0, y_0, z) = U(x_0, y_0, z) e^{-jk \frac{r}{2z}} \quad (7.5-2)$$

类似地, 式 (7.3-11) 和式 (7.4-5) 中的积分也是傅里叶变换。式 (7.3-11) 的傅里叶逆变换为

$$u'(x, y) = \iint_{\infty} U'(x_0, y_0, z) e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (7.5-3)$$

其中

$$U'(x_0, y_0, z) = U(x_0, y_0, z) e^{-jk r_0} \quad (7.5-4)$$

类似地, 式 (7.4-5) 的傅里叶逆变换为

$$u''(x, y) = \iint_{\infty} U'(x_0, y_0, z) e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (7.5-5)$$

迭代优化技术经常被用来设计诸如衍射光学器件 (DOE) 等光学元器件 [Lee, 1970, 1975, 1979; Zhuang and Ersoy, 1995], 这些将会在第 15 章和第 16 章讨论。为此, 已开发出许多在输出面和输入面之间进行迭代计算的迭代方法。式 (7.3-11) 和式 (7.5-3) 以及式 (7.4-5) 和式 (7.5-5) 为傅里叶变换对。利用这些公式可执行精确的迭代优化。

7.6 数值算例

不失一般性, 下面首先忽略 y 变量, 而只考虑二维的情形。在初始计算模拟时, 相应的参数选择如下:

$$\lambda = 0.6328 \mu\text{m}, x = 1 \text{ mm}, N = 256, \Delta x / \lambda = 7$$

对于使近似有效的最短距离 z , 其与输出面上的坐标偏离 x_0 有关, 可由令 $k(r - r_{\text{approx}})$ 小于 1 rad 来计算得到。当坐标偏离 x 等于 1 mm 时, 得到的结果如图 7.1 所示。由图 7.1 可见, 对于非常小的 $(x_0 - x)$, 两种近似的特性基本一致, 但是随着偏离 $(x_0 - x)$ 的增大, 差距也变大。

图 7.2 给出了比值 z/x_0 随输出面的偏离坐标 x_0 的变化。随着 x_0 的增大, 对于本章新提出的近似, 这个比值接近于 2, 而对菲涅耳近似, 其接近于 36。注意到利用 FFT, 最大的输出偏离等于 $z/(\Delta x / \lambda)$, Δx 为输入面上的采样间隔。因此, 对于输出面上所有满足 $\Delta x / \lambda$ 大于 2 的采样点, 新的近似均具有足够的精度。图 7.3 给出了当 $z = 10 \text{ cm}$ 时, 随着 x_0 的增大, 利用式 (7.3-6) 计算得到的 x_0 和 x_s 之间的关系, 可见其差距也不断增大。

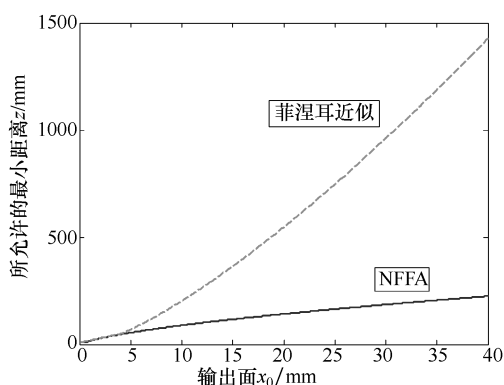


图 7.1 当 $x = 1$ 时, 使近似有效的最小距离 z 随输出面的坐标 x_0 的变化

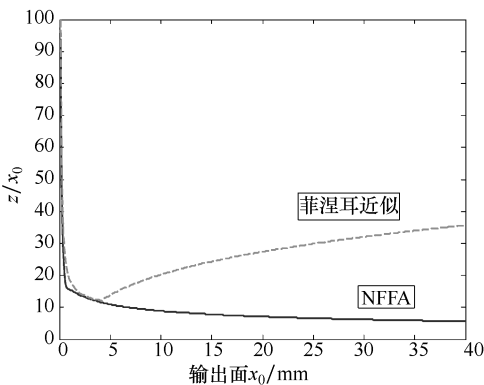


图 7.2 比值 z/x_0 随 x_0 的变化

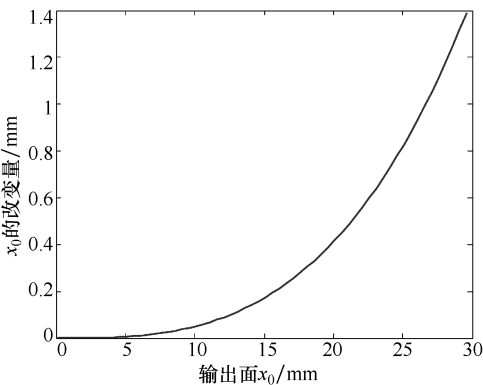


图 7.3 x_0 的改变量随着输出面上的 x_0 的变化 ($z=10\text{cm}$, $x=1\text{mm}$)

式 (7.3-8) 和式 (7.3-9) 给出了菲涅耳近似或夫琅禾费近似中的规则采样输出点与 NFFA 中半规则采样输出点之间的关系。表 7.1 至表 7.3 给出了这些关系的具体计算实例。这些表中的 (x, y) 采样坐标值用复数 $x + jy$ 来显示。表 7.1 给出了规则采样值的例子，表 7.2 给出了对应的半规则采样值的例子。表 7.3 是表 7.1 和表 7.2 之间的差值。

表 7.1 FFT 和菲涅耳近似下的输出面位置坐标 (单位: mm)

-8.0998	-8.0998j	-8.0682	-8.0998j	-8.0366	-8.0998j	-8.0049	-8.0998j	-7.9733	-8.0998j	-7.9416	-8.0998j	-7.9100	-8.0998j
-8.0998	-8.0682j	-8.0682	-8.0682j	-8.0366	-8.0682j	-8.0049	-8.0682j	-7.9733	-8.0682j	-7.9416	-8.0682j	-7.9100	-8.0682j
-8.0998	-8.0366j	-8.0682	-8.0366j	-8.0366	-8.0366j	-8.0049	-8.0366j	-7.9733	-8.0366j	-7.9416	-8.0366j	-7.9100	-8.0366j
-8.0998	-8.0049j	-8.0682	-8.0049j	-8.0366	-8.0049j	-8.0049	-8.0049j	-7.9733	-8.0049j	-7.9416	-8.0049j	-7.9100	-8.0049j
-8.0998	-7.9733j	-8.0682	-7.9733j	-8.0366	-7.9733j	-8.0049	-7.9733j	-7.9733	-7.9733j	-7.9416	-7.9733j	-7.9100	-7.9733j
-8.0998	-7.9416j	-8.0682	-7.9416j	-8.0366	-7.9416j	-8.0049	-7.9416j	-7.9733	-7.9416j	-7.9416	-7.9416j	-7.9100	-7.9416j
-8.0998	-7.9100j	-8.0682	-7.9100j	-8.0366	-7.9100j	-8.0049	-7.9100j	-7.9733	-7.9100j	-7.9416	-7.9100j	-7.9100	-7.9100j

表 7.2 FFT 和 NFFA 下的输出面位置坐标 (单位: mm)

-8.1535	-8.1535j	-8.1214	-8.1533j	-8.0894	-8.1531j	-8.0573	-8.1529j	-8.0253	-8.1527j	-7.9932	-8.1525j	-7.9612	-8.1523j
-8.1533	-8.1214j	-8.1212	-8.1212j	-8.0892	-8.1210j	-8.0571	-8.1208j	-8.0251	-8.1206j	-7.9930	-8.1204j	-7.9610	-8.1202j
-8.1531	-8.0894j	-8.1210	-8.0892j	-8.0890	-8.0890j	-8.0569	-8.0888j	-8.0249	-8.0886j	-7.9928	-8.0884j	-7.9608	-8.0881j
-8.1529	-8.0573j	-8.1208	-8.0571j	-8.0888	-8.0569j	-8.0567	-8.0567j	-8.0247	-8.0565j	-7.9926	-8.0563j	-7.9606	-8.0561j
-8.1527	-8.0253j	-8.1206	-8.0251j	-8.0886	-8.0249j	-8.0565	-8.0247j	-8.0245	-8.0245j	-7.9924	-8.0243j	-7.9604	-8.0240j
-8.1525	-7.9932j	-8.1204	-7.9930j	-8.0884	-7.9928j	-8.0563	-7.9926j	-8.0243	-7.9924j	-7.9922	-7.9922j	-7.9602	-7.9920j
-8.1523	-7.9612j	-8.1202	-7.9610j	-8.0881	-7.9608j	-8.0561	-7.9606j	-8.0240	-7.9604j	-7.9920	-7.9602j	-7.9600	-7.9600j

表 7.3 输出面位置坐标误差 (表 7.1 和表 7.2 之间的差) (单位: mm)

-0.0537	-0.0537j	-0.0532	-0.0535j	-0.0528	-0.0532j	-0.0524	-0.0530j	-0.0520	-0.0528j	-0.0516	-0.0526j	-0.0512	-0.0524j
-0.0535	-0.0532j	-0.0530	-0.0530j	-0.0526	-0.0528j	-0.0522	-0.0526j	-0.0518	-0.0524j	-0.0514	-0.0522j	-0.0510	-0.0520j
-0.0532	-0.0528j	-0.0528	-0.0526j	-0.0524	-0.0524j	-0.0520	-0.0522j	-0.0516	-0.0520j	-0.0512	-0.0518j	-0.0508	-0.0516j
-0.0530	-0.0524j	-0.0526	-0.0522j	-0.0522	-0.0520j	-0.0518	-0.0518j	-0.0514	-0.0516j	-0.0510	-0.0514j	-0.0506	-0.0512j
-0.0528	-0.0520j	-0.0524	-0.0518j	-0.0520	-0.0516j	-0.0516	-0.0514j	-0.0512	-0.0512j	-0.0508	-0.0510j	-0.0504	-0.0508j
-0.0526	-0.0516j	-0.0522	-0.0514j	-0.0518	-0.0512j	-0.0514	-0.0510j	-0.0510	-0.0508j	-0.0506	-0.0506j	-0.0502	-0.0504j
-0.0524	-0.0512j	-0.0520	-0.0510j	-0.0516	-0.0508j	-0.0512	-0.0506j	-0.0508	-0.0504j	-0.0504	-0.0502j	-0.0500	-0.0500j

在下一组二维数值实验中，对通过一个给定衍射孔径的会聚球面波，给出了精确解和菲涅耳近似积分、NFFA 近似积分之间的对比。实验参数如下：

衍射孔径尺寸 $D_l = 30\mu\text{m}$ ，聚焦距离 $z = 5\text{mm}$ ， $N = 512$ ， $\lambda = 0.6328\mu\text{m}$



数值积分是采用自适应辛普森积分计算得到的 [Garner]。根据 FFT 条件, 对输出孔径上的点进行采样。于是, 整个输出孔径的尺寸由下式给出:

$$D_o = \frac{N\lambda z}{D_i} \quad (7.6-1)$$

虽然如此, 最后只有输出孔径中心部分的 64 个采样点和大约 6.7mm 尺寸范围内的结果是精确计算的。由于 $z = 5\text{mm}$, 这表明输出结果是一个宽角度输出场。

由图 7.4 可知, 由精确计算和利用 NFFA 方法分别得到的输出面衍射条纹强度几乎完美地重合在一起。图 7.5 给出了由下式定义的归一化强度误差:

$$\text{Error} = \frac{I_{\text{Exact}} - I_{\text{NFFA}}}{I_{\text{Exact}}} \quad (7.6-2)$$

式中, I_{Exact} 和 I_{NFFA} 分别为由精确计算得到的强度和利用 NFFA 近似得到的强度。

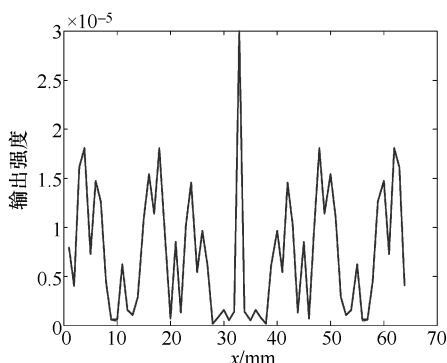


图 7.4 由精确计算及利用 NFFA 近似的衍射积分得到的孔径衍射强度

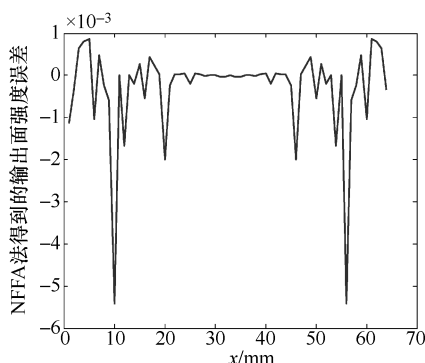


图 7.5 由精确计算与 NFFA 近似的衍射积分得到的输出面强度误差

从图 7.6 可知, 利用精确计算方法与 NFFA 近似积分方法分别计算得到的输出相角几乎完美地重合在一起。

图 7.7 给出了利用菲涅耳近似与精确计算方法分别计算得到的输出衍射积分条纹强度的对比。图 7.8 给出了利用菲涅耳近似与精确计算方法分别计算得到的输出面相角的对比。可见, 菲涅耳近似引入的误差几乎可以忽略不计。

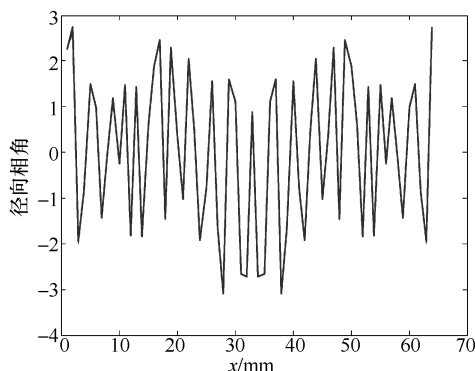


图 7.6 精确计算与 NFFA 近似衍射积分计算得到的输出面相角

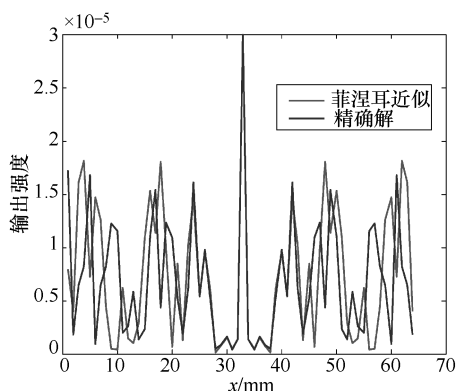


图 7.7 利用菲涅耳近似与精确计算得到的输出衍射积分条纹强度的对比



在下一组实验中，主要讨论了三维情况下的结果。图 7.9 给出了模拟计算中采用的输入面方形孔径。实验中使用了两组参数，其中第一组的参数如下：

$$z = 50\text{mm}, \quad D_I(\text{输入面孔径边长}) = 0.2\text{mm}, \quad N = 512, \quad \lambda = 0.6328\mu\text{m}$$

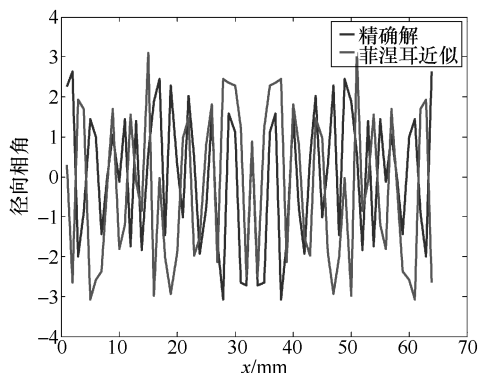


图 7.8 利用菲涅耳近似利用精确计算方法分别计算得到的输出面相角的对比

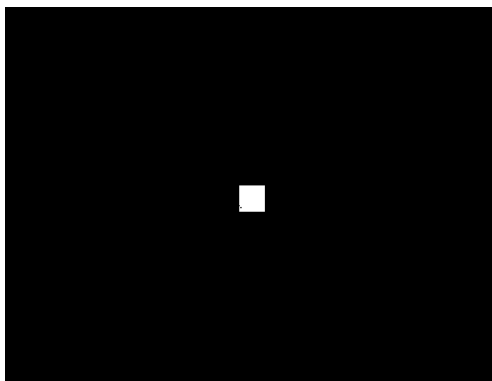


图 7.9 模拟计算中采用的输入面方形孔径

根据 FFT 条件，在以上参数条件下，输出面的孔径尺寸为 0.081mm 。图 7.10 显示了采用 NFPA 计算得到的衍射输出强度三维图，图 7.11 给出了相应的利用菲涅耳近似计算得到的衍射输出强度三维分布。

图 7.12 给出了衍射强度的计算误差（图 7.11 和图 7.12 对应数据的绝对值差）。可见，在这种实验条件下，菲涅耳衍射使观察平面的特定区域的衍射强度误差可以忽略。

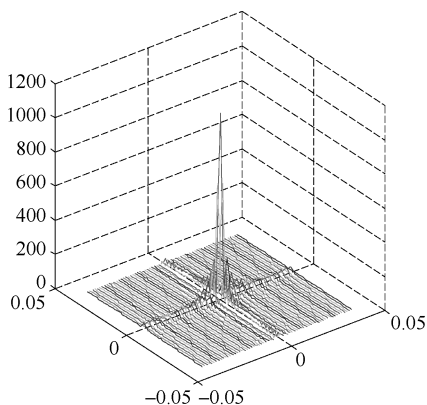


图 7.10 利用 NFPA 计算得到的衍射输出强度三维分布

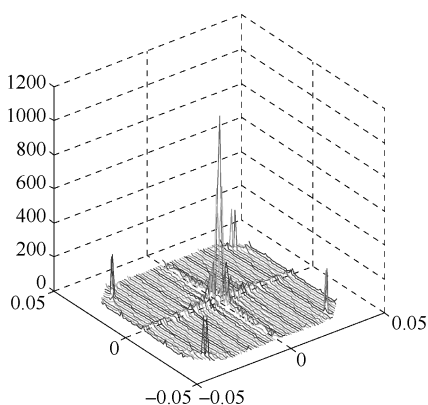


图 7.11 利用菲涅耳近似计算得到的衍射输出强度

第二组实验参数选择如下：

$$z = 100\text{mm}, \quad D_I(\text{输入面孔径边长}) = 2\text{mm}, \quad N = 512, \quad \lambda = 0.6328\mu\text{m}$$

在以上参数下，利用 FFT 条件，输出面孔径尺寸变为 16.2mm 。图 7.13 给出了利用 NFPA 计算得到的三维衍射输出强度分布。对应的由菲涅耳衍射计算得到的三维衍射输出强度分布如图 7.14 所示。图 7.15 给出以上两种近似方法得到的结果之差（图 7.13 和图 7.14 的绝对值差）。由计算结果可见，在这组参数下，对观察平面上特定区域的菲涅耳衍射，其误



差变得更小，这主要是由于距离 z 变得更大，而输出面孔径尺寸变得更小。

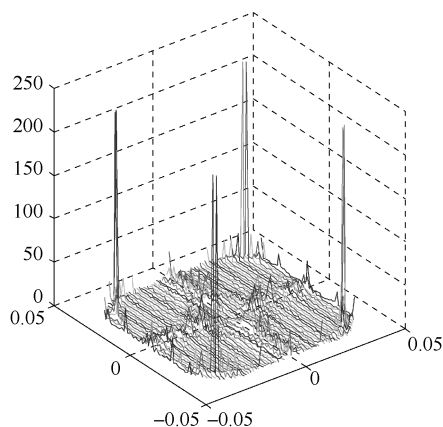


图 7.12 衍射强度的误差 (图 7.11 和图 7.12 对应数据的绝对值差)

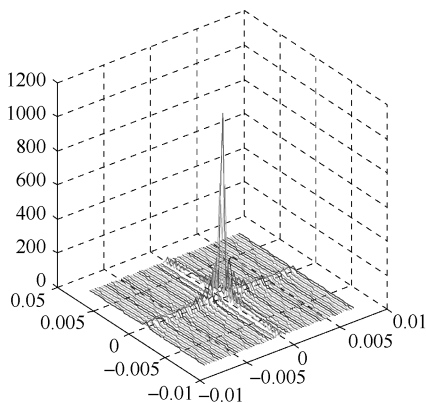


图 7.13 利用 NFFA 计算得到的三维衍射输出强度分布

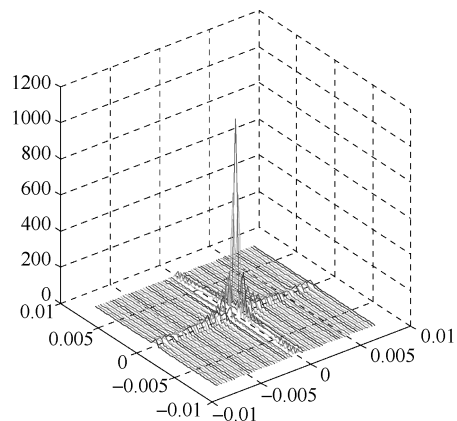


图 7.14 利用菲涅耳近似衍射计算得到的三维衍射输出强度分布

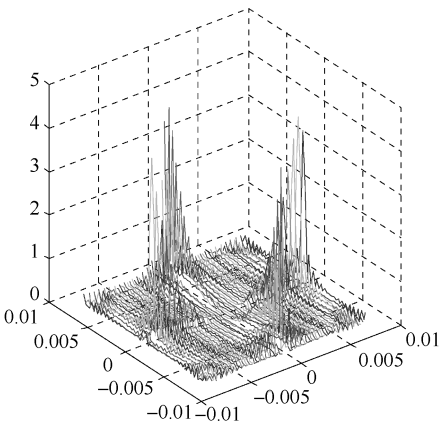


图 7.15 衍射强度的误差 (图 7.14 和图 7.15 对应数据的绝对值差)

最后一组实验采用如图 7.16 所示的棋盘图像作为所需要的输出面衍射强度条纹。实验参数选择如下：

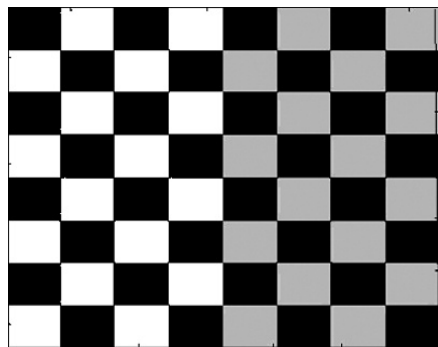


图 7.16 棋盘图像和其利用逆 NFFA 重构计算得到的图像

($z=80\text{mm}$, 输入面尺寸 $=0.8\text{mm} \times 0.8\text{mm}$, 输出面尺寸 $=16.2\text{mm} \times 16.2\text{mm}$)



$$z = 80\text{mm}, \quad D_I(\text{输入面孔径边长}) = 0.8\text{mm}, \quad N = 256, \quad \lambda = 0.6328\mu\text{m}$$

利用 FFT 条件, 在上述参数下输出面孔径尺寸变为 16.2mm。采用两种方式计算输入面衍射条纹。第一种是采用由式 (7.5-1) 和式 (7.5-2) 给出的逆菲涅耳近似。第二种是采用由式 (7.5-3) 和式 (7.5-4) 给出的逆 NFFA 近似。图 7.16 同时也给出了半规则输出面采样的 NFFA 重构衍射强度, 由图可见, 重构图像与原始图像是基本一样的。图 7.17 给出了规则采样条件下利用 NFFA 计算得到的输出面衍射强度, 其中所需要的输入面衍射条纹利用逆菲涅耳近似得到。由于 NFFA 的结果可认为是精确的, 因此, 在规则采样下, 实际物理实现中的输出衍射强度与图 7.17 看起来相似。

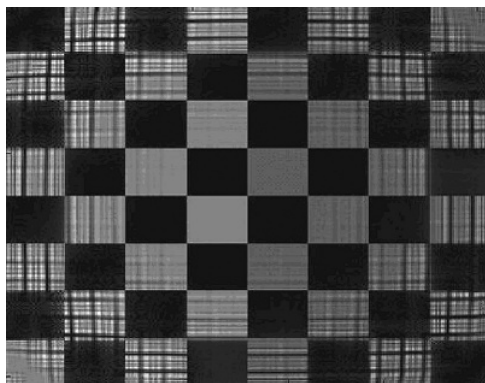


图 7.17 以利用逆菲涅耳衍射得到的场作为输入场, 再利用 NFFA 重构得到的棋盘图像
($z = 80\text{mm}$, 输入面尺寸 $= 0.8\text{mm} \times 0.8\text{mm}$, 输出面尺寸 $= 16.2\text{mm} \times 16.2\text{mm}$)

7.7 更高精度近似

如果输出空间的坐标中心在 (x_c, y_c) , 且输入空间的坐标中心在 (x_1, y_1) , 则可通过如下的定义得到更高的近似精度:

$$x_0 = x'_0 + x_c \quad (7.7-1)$$

$$y_0 = y'_0 + y_c \quad (7.7-2)$$

$$x = x' + x_1 \quad (7.7-3)$$

$$y = y' + y_1 \quad (7.7-4)$$

$$q = (x_c - x_1)^2 + (y_c - y_1)^2 \quad (7.7-5)$$

$$r_c = (z^2 + q)^{1/2} \quad (7.7-6)$$

进一步定义

$$x_{cl} = x_c - x_1 \quad (7.7-7)$$

$$y_{cl} = y_c - y_1 \quad (7.7-8)$$

$$v' = x'^2_0 + y'^2_0 + 2(x_{cl}x'_0 + y_{cl}y'_0) \quad (7.7-9)$$

$$w' = x'^2 + y'^2 - 2[x_{cl}x' + y_{cl}y'] \quad (7.7-10)$$

$$g' = \frac{v' + w'}{2r_c^2} - \frac{x'_0x + y'_0y}{r_c^2} \quad (7.7-11)$$



如果将 z 替换为 r_c ，将 g 、 v 、 w 分别替换为 g' 、 v' 、 w' ，则之前的分析仍然有效。此时的误差近似地由下式给出：

$$E \simeq -\frac{v'(x^2 + y^2)}{4r_c^3} - \frac{(x_0x + y_0y)^2}{2r_c^3} \quad (7.7-12)$$

上述近似在用于大尺度模拟时会特别有用，届时将输入面和/或输出面划分成更小的子区域，然后在这些子区域上应用 NFFA 即可。利用并行计算系统，可以得到接近实时的大尺度模拟分析。

7.8 小结

本章介绍了利用 NFFA 可相当程度地改进菲涅耳和夫琅禾费近似，特别是当更高精度的输出采样时仍采用 FFT 来进行宽角度衍射计算的情况，改进精度更高。

菲涅耳/夫琅禾费近似与 NFFA 之间的最主要差别，在于其不同的采样点。除非常宽角度的衍射外，式 (7.2-11) 和式 (7.3-11) 积分外的不同相位因子是可比较的，且不影响输出面上的衍射条纹强度。因此当正常利用菲涅耳近似计算时，如果利用由式 (7.3-8) 和式 (7.3-9) 所决定的新的采样点，则结果仍具有足够的精度，不过此时的结果是新采样点的 NFFA 结果。利用 NFFA 近似所允许的 z 距离要小于菲涅耳近似所允许的 z 距离，后者大约是前者的 18 倍。

只要半规则输出采样是可接受的，则 NFFA 方法就与角谱法具有同样的精度，且 NFFA 对实际情况中的绝大多数宽角度衍射都是有效的。

第 8 章 几何光学

8.1 本章引论

几何光学（或称射线光学）讨论光波传播的一种近似，即把光波长 λ 看成是无限小的，在实际应用层面，这意味着把它看成远小于影响光波场振幅和相位的干扰的空间尺寸。例如，如果相移孔径是波长的好多倍，例如，相移 2π 弧度，此时几何光学是非常适用的。又如在纤芯较大的多模光纤中就使用射线光学。

在几何光学中，通常用光线来描述光波场，这些光线按一系列的规则在不同的光介质中传播。程函方程（或称几何光学方程）形成了几何光学的基本原理，它是亥姆霍兹方程在波长趋于零时的特殊形式。

本章共分 8 节。8.2 节介绍了光线传播的物理基础；8.3 节介绍了光线方程，当光线传播接近光轴时的傍轴光线方程是简化的光线方程。

8.4 节介绍了程函方程；8.5 节通过举例描述了局部空间频率和光线；8.6 节以 2×2 阶的矩阵代数描述了子午面光线，它是一种在单平面上传播的光线，包含光轴和另一种正交轴，如 y 轴；8.7 节概括了适合厚透镜的理论。8.8 节对复杂光学系统的入瞳和出瞳做了介绍。

8.2 光线的传播

几何光学中光线的位置和方向是首要考虑的对象，例如，光学成像系统把从物体各点发出的光线会聚后重新调整其方向，对应到图像的相应像素。

以光轴为中心的光学器件通常很小，因而光线以非常小的角度传播，这样的光线称为傍轴光线，它是傍轴光学的基础。

几何光学成立的基本假设如下：

1. 光波场以光线的形式传播。

2. 在不同的光学介质中，光线以不同的速度传播：光学介质的折射率 $n \geq 1$ 。如果当 $n = 1$ 时，传播速度等于 c ，则当 $n \neq 1$ 时，传播速度为 c/n 。在普通的光学介质中，折射率 n 是关于位置的函数 $n(r)$ 。

3. 费马原理：光沿着所需时间最短的路径传播。为了更好地理解这种说法，考虑图 8.1。图 8.1 中光线沿传播路径在微分光程 ds 上的无限小传播时间可由下式得到

$$\frac{ds}{c/n(r)} = \frac{n(r) ds}{c} \quad (8.2-1)$$

定义 AB 两点之间的光程 L_p 为

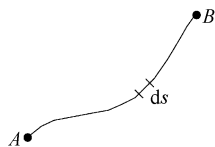


图 8.1 A、B 两点之间的光程



$$L_p = \int_A^B n(r) ds \quad (8.2-2)$$

从 A 到 B 的传播时间为 L_p/c 。

费马原理表明光程 L_p 是关于邻近路径的极值

$$\delta L_p = 0 \quad (8.2-3)$$

等式两边都除以 v ，传播时间的变化也满足下式：

$$\frac{L_p}{v} = 0 \quad (8.2-4)$$

此极值是最小值，故证实光沿着所需时间最短的路径传播。

折射率 n 在均匀介质中是常数。两点之间光程的最短路径是这两点之间的一条直线，这通常描述为“光线是沿直线传播的”。

费马原理可证明光的平面镜反射现象以及光在两种光学介质边界上的反射和折射现象，如下例诠释。

例 8.1 证明：当光线在平面镜上发生反射时，反射角等于入射角。介质折射率为常数 n 。

解：相关的示意图如图 8.2 所示。 θ_1 和 θ_2 分别是入射角和反射角，光经平面镜反射后 AC 之间的路径应该是光程最小值。

光程可由下式计算得到

$$L_p = n(AB + BC)$$

上式可写成

$$L_p = n \{ [h_1^2 + (d_1 - d_2)^2]^{1/2} + (h_2^2 + d_2^2)^{1/2} \}$$

L_p 取决于独立变量 d_2 。设 $\delta L_p / \delta d_2 = 0$ ，则

$$\frac{d_1 - d_2}{[h_1^2 + (d_1 - d_2)^2]^{1/2}} = \frac{d_2}{(h_2^2 + d_2^2)^{1/2}}$$

即

$$\sin \theta_1 = \sin \theta_2$$

或者

$$\theta_1 = \theta_2$$

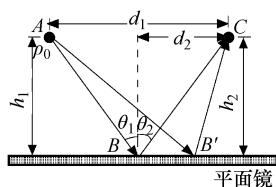


图 8.2 经过平面镜反射后 AC 之间的最短路径

例 8.2 (a) 入射光线在折射率分别为 n_1 、 n_2 的介质边界上分成两束光线，如图 8.3 所示。

证明：

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (8.2-5)$$

这就是斯乃尔定律。

(b) 光线近光轴传播，证明：

$$n_1 \theta_1 \cong n_2 \theta_2 \quad (8.2-6)$$

这就是几何光学中的傍轴近似。

解：(a) A 、 C 两点之间的光程如下：

$$L_p = n_1 AB + n_2 BC$$

由费马原理知，光程 L_p 是最小值。 L_p 可写为

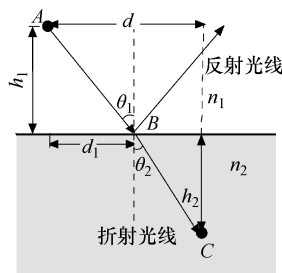


图 8.3 两种介质边界的反射和折射



$$L_p = n_1 (h_1^2 + d_1^2)^{1/2} + n_2 [h_2^2 + (d - d_1)^2]^{1/2}$$

d_1 是独立变量, 设 $\delta L_p / \delta d_1 = 0$, 则

$$\frac{n_1 d_1}{(h_1^2 + d_1^2)^{1/2}} = \frac{n_2 (d - d_1)}{[h_2^2 + (d - d_1)^2]^{1/2}}$$

上式可写为

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

(b) 当光线靠近光轴传播时, $\sin \theta_1 \cong \theta_1$, $\sin \theta_2 \cong \theta_2$, 所以

$$n_1 \theta \cong n_2 \theta_2$$

例 8.3 一薄透镜有如图 8.4 所示的性质。

一束以角度 θ_1 入射的光线又以角度 θ_2 射出薄透镜, θ_2 可由下式得到:

$$\theta_2 = \theta_1 - \frac{y}{f} \quad (8.2-7)$$

式中, y 是入射点到光轴的距离; f 是透镜的焦距。

焦距 f 可由下式得到:

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (8.2-8)$$

式中, n 是透镜材料的折射率; R_1 和 R_2 是透镜两表面的曲率半径。

根据上面已给的信息, 证明透镜成像定律。

解: 考虑图 8.4b。由方程式 (8.2-7) 和傍轴近似 $\sin \theta \approx \theta$ 可知, 当入射角 $\theta_1 = 0$ 时, 从 P_1 点发出且平行于光轴的光线穿过透镜后到达焦点 F ; 当 $y = 0$ 时, 从 P_1 点出发, 沿着光轴通过透镜中心点的光线穿过透镜时方向不变。两条光线相交于 P_2 点, P_2 即为像。

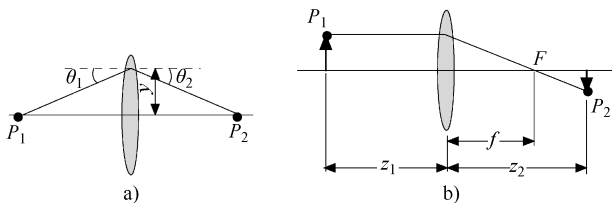


图 8.4 薄透镜的性质

a) 经薄透镜改变方向的光线 b) 薄透镜成像

8.3 光线方程

在梯度折射率介质 (graded-index material, GRIN) 中, 折射率是位置 r 的函数, 即 $n(r)$ 。光线在这种介质里沿着曲线路径传播并遵循费马原理。

光线的轨迹可以用坐标函数 $x(s)$ 、 $y(s)$ 、 $z(s)$ 表示, 变量 s 表示轨迹的长度, 如图 8.5 所示。A、B 两点之间的轨迹满足费马原理

$$\delta \int_A^B n(\mathbf{r}) ds = 0 \quad (8.3-1)$$



\mathbf{r} 是位置矢量, 使用变分法, 方程式 (8.3-1) 可以转换为如下的光线方程:

$$\frac{d}{ds} \left(n(\mathbf{r}) \frac{d\mathbf{r}}{ds} \right) = \nabla n(\mathbf{r}) \quad (8.3-2)$$

式中, ∇ 是梯度算符。

在 8.4 节里将通过程函方程严格地推导出光线方程。

上面是以矢量的形式给出了光线方程, 在笛卡儿坐标系中, 各分量的方程可以写成如下形式:

$$\frac{d}{ds} \left(n \frac{dx}{ds} \right) = \frac{\partial n}{\partial x}$$

$$\frac{d}{ds} \left(n \frac{dy}{ds} \right) = \frac{\partial n}{\partial y}$$

$$\frac{d}{ds} \left(n \frac{dz}{ds} \right) = \frac{\partial n}{\partial z}$$

方程式 (8.3-2) 的解并不简单, 在这里假设光线接近光轴传播, 即 $ds \cong dz$, 则可通过傍轴近似大大简化该方程, 简化后原方程可以写为如下形式:

$$\frac{d}{dz} \left(n \frac{dx}{dz} \right) = \frac{\partial n}{\partial x} \quad (8.3-3)$$

$$\frac{d}{dz} \left(n \frac{dy}{dz} \right) = \frac{\partial n}{\partial y} \quad (8.3-4)$$

方程式 (8.3-3) 和方程式 (8.3-4) 就是傍轴光线方程。

例 8.4 证明光线在均匀介质中沿直线传播。

解: 在均匀介质中折射率 n 为常数, 故方程式 (8.3-3) 和式 (8.3-4) 可写为

$$\frac{d^2 x}{dz^2} = \frac{d^2 y}{dz^2} = 0$$

所以

$$\begin{aligned} x &= Az + B \\ y &= Cz + D \end{aligned} \quad (8.3-5)$$

方程式 (8.3-5) 表示直线。

例 8.5 若 $n = n(y)$, 说明怎样运用傍轴光线方程。

解: 由于 n 不是 x 的函数, 所以沿 x 轴方向仍然是方程式 (8.3-5), 在 y 轴方向, 方程式 (8.3-4) 可以写为

$$\frac{d^2 y}{dz^2} = \frac{1}{n} \frac{dn}{dy}$$

给定 $n(y)$ 以及初值条件 $y(0)$ 和 $\left. \frac{dy}{dz} \right|_{z=0}$, 则可解出 $y(z)$ 。

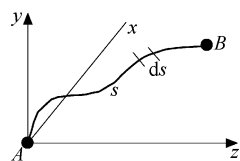


图 8.5 梯度折射率介质中光线的轨迹

8.4 程函方程

当光在折射率相对于波长缓慢变化的介质中传播时, 单色波有可如下近似表示的复振幅



$$U(\mathbf{r}) = A(\mathbf{r}) e^{-ik_0 S(\mathbf{r})} \quad (8.4-1)$$

式中, k_0 等于自由空间的波数 $2\pi/\lambda_0$; λ_0 是自由空间的波长; $k_0 S(\mathbf{r})$ 是波相位; $S(\mathbf{r})$ 为程函方程, 是折射率 $n(\mathbf{r})$ 的函数。

常数相位的表面定义为

$$S(\mathbf{r}) = \text{常量} \quad (8.4-2)$$

这些表面称作波前。能量沿着垂直于 \mathbf{r} 位置处波前的 \mathbf{k} 方向传播, 光线传播的轨迹上的每一点都垂直于波前, 换句话说, 光线传播的方向与能量流动的方向一致。

方程式 (8.4-1) 可以用亥姆霍兹方程代换

$$(\nabla^2 + k^2) U(\mathbf{r}) = 0 \quad (8.4-3)$$

得到

$$k_0^2 (n^2 - |\nabla S|^2) A + \nabla^2 A = 0 \quad (8.4-4)$$

和

$$k_0 (2 \nabla S \cdot \nabla A + A \nabla^2 S) = 0 \quad (8.4-5)$$

方程式 (8.4-4) 可写为

$$|\nabla S|^2 = n^2 + \frac{\lambda_0^2}{2\pi^2} \frac{\nabla^2 A}{A} \quad (8.4-6)$$

当 $\lambda_0 \rightarrow 0$ 时, 方程式 (8.4-6) 右边的第二项趋近于 0, 故有

$$|\nabla S(\mathbf{r})|^2 \cong n^2(\mathbf{r}) \quad (8.4-7)$$

该方程即为熟知的程函方程, 它可以被用来确定波前。由于光线轨迹垂直于波前, 故可由波前确定光线轨迹, 图 8.6 呈现了三种波前。

程函方程可以解释为亥姆霍兹方程在波长趋于 0 时的极限。

只要由方程式 (8.4-7) 求得 $S(\mathbf{r})$, 就可通过方程式 (8.4-5) 确定 $A(\mathbf{r})$, 进而可知波场 $U(\mathbf{r})$ 。

方程式 (8.4-7) 也可表示为

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2(\mathbf{r}) \quad (8.4-8)$$

那么, 折射率可由下式给出:

$$n(\mathbf{r}) = |\nabla S(\mathbf{r})| \quad (8.4-9)$$

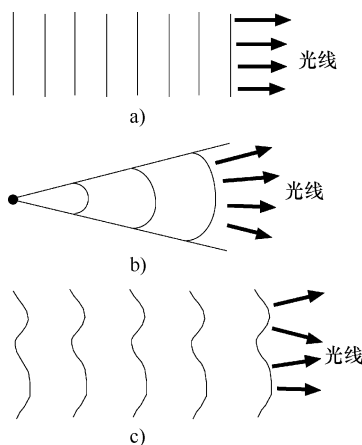


图 8.6 三种波前

a) 平面波 b) 球面波 c) 普通波

例 8.6 证明 A 、 B 两点之间的光程等于 $S(\mathbf{r}_B) - S(\mathbf{r}_A)$ 。

解: 因为 $n(\mathbf{r}) = |\nabla S(\mathbf{r})|$, 由下式可得光程:

$$L_p = \int_A^B n(\mathbf{r}) d\mathbf{r} = \int_A^B |\nabla S(\mathbf{r})| d\mathbf{r} = S(\mathbf{r}_B) - S(\mathbf{r}_A)$$

注意: 光程类似于电势差, 相位 $S(\mathbf{r})$ 类似于点电势。

8.5 局部空间频率和光线

在前面章节里, 采用具有方向余弦的平面波分量分析了任意波, 光线同样与方向余弦相



关,为此,有必要分析波场的非稳态特性。

一般而言,非稳态信号 $h(t)$ 是一种信号频率随时间 t 变化的函数,这样的信号不能单独地由傅里叶变换表示其特征,但可以通过计算 $h(t)\omega(t)$ 的短时傅里叶变换,其中 $\omega(t)$ 是范围有限的窗口函数。

针对非稳态二维信号,可以把上面的一维情况推广到与波场有关的二维情况,因此,我们可以写出 z 为常数时的波场

$$h(x, y) = A(x, y) e^{j\phi(x, y)} \quad (8.5-1)$$

这里 z 被约束,假设 $A(x, y)$ 缓慢变化。

定义局部空间频率为

$$f'_x = \frac{1}{2\pi} \frac{\partial}{\partial x} \phi(x, y) \quad (8.5-2)$$

$$f'_y = \frac{1}{2\pi} \frac{\partial}{\partial y} \phi(x, y) \quad (8.5-3)$$

在 $h(x, y)$ 为 0 的区域,局部空间频率 f'_x 和 f'_y 也被定义为 0。

一般情况下, f'_x 和 f'_y 与前面章节中讨论的傅里叶空间频率并不相同,特别地,仅当在以下情况时,上述两种频率是一致的;当相位 $\phi(x, y)$ 在 $x-y$ 平面内缓慢变化时, $\phi(x, y)$ 近似等于在点 (x_0, y_0) 附近展开的三级泰勒级数,此三级泰勒级数即为

$$\phi(x_0, y_0), \left. \frac{\partial}{\partial x} \phi(x, y) \right|_{\substack{x=x_0 \\ y=y_0}}, \left. \frac{\partial}{\partial y} \phi(x, y) \right|_{\substack{x=x_0 \\ y=y_0}}$$

局部方向余弦定义为

$$\alpha' = \lambda f'_x \quad (8.5-4)$$

$$\beta' = \lambda f'_y \quad (8.5-5)$$

$$\gamma' = \sqrt{1 - \alpha'^2 - \beta'^2} \quad (8.5-6)$$

它表明这些方向余弦和光线在点 (x, y) 处的方向余弦完全相同,其中点 (x, y) 为常数 z 确定的平面内的点。

例 8.7 确定平面波场的局部空间频率。

解: 平面波场的空域部分可以表示为

$$h(x, y) = e^{j2\pi(f_x x + f_y y)}$$

然后,根据 $\phi(x, y) = 2\pi(f_x x + f_y y)$ 得

$$f'_x = \frac{1}{2\pi} \frac{\partial}{\partial x} \phi(x, y) = f_x$$

$$f'_y = \frac{1}{2\pi} \frac{\partial}{\partial y} \phi(x, y) = f_y$$

在此情况下,局部空间频率和傅里叶变换空间的频率是相同的。

例 8.8 定义有限截断函数如下:

$$h(x, y) = e^{j\pi\alpha(x^2 + y^2)} \text{rect}\left(\frac{x}{2D_x}\right) \text{rect}\left(\frac{y}{2D_y}\right)$$

(a) 求局部空间频率。



(b) 试比较局部频率和 $g(x, y)$ 的傅里叶变换的特性。

解: (a) 在原点附近, 尺寸在 $2D_x$ 和 $2D_y$ 之间的矩形区域内, 相位如下:

$$\phi(x, y) = \pi\alpha(x^2 + y^2)$$

在该矩形区域内, 局部空间频率为

$$f'_x = \frac{1}{2\pi} \frac{\partial}{\partial x} \phi(x, y) = \alpha x$$

$$f'_y = \frac{1}{2\pi} \frac{\partial}{\partial y} \phi(x, y) = \alpha y$$

在该矩形区域外, 定义 f'_x 和 f'_y 为 0, 因此, 它们可表示为

$$f'_x = \alpha x \text{rect}\left(\frac{x}{2D_x}\right)$$

$$f'_y = \alpha y \text{rect}\left(\frac{y}{2D_y}\right)$$

需注意, 在已给的矩形区域内, f'_x 和 f'_y 分别随 x 和 y 线性变化。

(b) $h(x, y)$ 是可分离函数

$$h(x, y) = h_1(x)h_2(y) = \left[e^{j\pi\alpha x^2} \text{rect}\left(\frac{x}{2D_x}\right) \right] \left[e^{j\pi\alpha y^2} \text{rect}\left(\frac{y}{2D_y}\right) \right]$$

所以, $h(x, y)$ 的傅里叶变换等于 $H_1(f_x) * H_2(f_y)$ 。

$h_1(x)$ 的傅里叶变换如下:

$$H_1(f_x) = \int_{-D_x}^{D_x} e^{j\pi\alpha x^2} e^{-j2\pi f_x x} dx \quad (8.5-7)$$

使 $v = \sqrt{2\alpha}[x - (f_x/\alpha)]$, 替换变量, 方程式 (8.5-7) 则可表示为

$$H_1(f_x) = \frac{1}{\sqrt{2\alpha}} e^{-j\frac{\pi f_x^2}{\alpha}} \int_{-L_1}^{L_2} e^{j\frac{\pi v^2}{2}} dv = \frac{1}{\sqrt{2\alpha}} e^{-j\frac{\pi f_x^2}{\alpha}} [C(L_2) - C(L_1) + jS(L_2) - jS(L_1)] \quad (8.5-8)$$

式中, $C(\cdot)$ 和 $S(\cdot)$ 分别是方程式 (5.2-18) 和方程式 (5.2-19) 给出的菲涅耳余弦和正弦积分。

定义 L_1 和 L_2 如下:

$$L_1 = \sqrt{2\alpha} \left(D_x + \frac{f_x}{\alpha} \right) \quad (8.5-9)$$

$$L_2 = \sqrt{2\alpha} \left(D_x - \frac{f_x}{\alpha} \right) \quad (8.5-10)$$

同理可得 $H_2(f_y)$ 。这表明在矩形区域外, $H_1(f_x)$ 的振幅快速趋近于 0 [Goodman, 2004], 简言之, 离开矩形区域的频率分量可以忽略不计, 实际上在矩形区域外的局部空间频率为 0 也说明了这个结果。然而, 局部频率和傅里叶频率之间的相似性在复杂问题中不容易得到。

8.6 子午面光线的矩阵表示

子午面光线是在一个含有 z 轴 (称为光轴) 的单平面内传播的光线, 横轴称为 y 轴。在



傍轴近似 ($\sin\theta \cong \theta$) 下, 在这个平面内传播的光线可以用 2×2 阶的矩阵表示, 称该矩阵为光线传递矩阵。

在傍轴近似下 $z = z_i$ 处的一子午面光线可以用它的位置 y_i 和角度 θ_i 表示, 如图 8.7 所示, 用一对 (y_i, θ_i) 变量表示子午面光线。考虑光线在从 $z = z_1$ 到 $z = z_2$ 的光学系统中传播, 如图 8.8 所示。

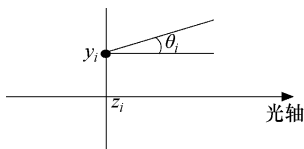


图 8.7 在 $z = z_i$ 处用 (y_i, θ_i) 表示的光线

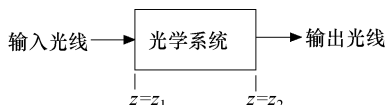


图 8.8 光线在从 $z = z_1$ 到 $z = z_2$ 的光学系统中传播

在傍轴近似下, 点 (y_1, θ_1) 和点 (y_2, θ_2) 的关系如下:

$$\begin{pmatrix} y_2 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y_1 \\ \theta_1 \end{pmatrix} \quad (8.6-1)$$

这里以 A 、 B 、 C 、 D 为元素的矩阵称为光线传递矩阵。

下面讨论基本光学系统的光线传递矩阵。

1. 在折射率 n 为常数的介质中传播: 如图 8.9 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \quad (8.6-2)$$

2. 平面镜反射: 如图 8.10 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (8.6-3)$$

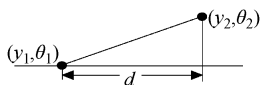


图 8.9 在折射率 n 为常数的介质中传播

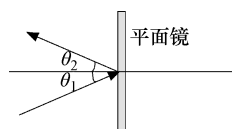


图 8.10 平面镜反射

3. 平面边界折射: 如图 8.11 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} \quad (8.6-4)$$

4. 球面镜反射: 如图 8.12 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & 0 \\ \frac{2}{R} & 1 \end{pmatrix} \quad (8.6-5)$$

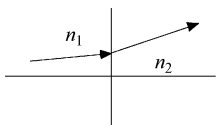


图 8.11 平面边界折射

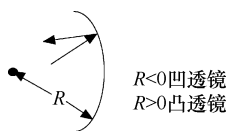


图 8.12 球面镜反射

5. 球面边界折射：如图 8.13 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{pmatrix} \quad (8.6-6)$$

6. 通过薄透镜传播：如图 8.14 所示。

矩阵 M 为

$$M = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \quad (8.6-7)$$

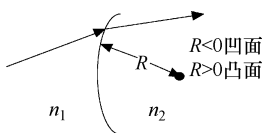


图 8.13 球面边界折射

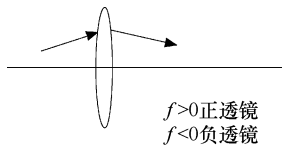
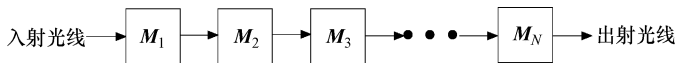


图 8.14 薄透镜传播

7. 串联的光学元件：假设 N 个光学元件串联在一起，如图 8.15 所示。

图 8.15 光线通过 N 个光学元件的传播

该串联光学系统的光线传递矩阵 M 为

$$M = M_N M_{N-1} \cdots M_1 \quad (8.6-8)$$

例 8.9 证明薄透镜的光线传递矩阵可由方程式 (8.6-7) 求得。

解： 折射率为 n_2 的薄透镜有两个半径分别为 R_1 和 R_2 的球面，假设透镜外的介质折射率等于 n_1 。通过透镜传播涉及透镜两个球面上的折射，所以，总的光线传递矩阵为

$$M = M_2 M_1 = \begin{pmatrix} 1 & 0 \\ \frac{n_2 - n_1}{n_1 R_2} & \frac{n_2}{n_1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R_1} & \frac{n_1}{n_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix}$$

上式中，焦距 f 可由下式确定：

$$\frac{1}{f} = \frac{n_2 - n_1}{n} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

例 8.10 确定下面系统的光线传递矩阵，如图 8.16 所示。



解：这个系统是距离为 d 的自由空间传播的串联系统，光线通过焦距为 f 的薄透镜，所以，光线传递矩阵为

$$\mathbf{M} = \mathbf{M}_2 \mathbf{M}_1 = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d \\ -\frac{1}{f} & 1 - \frac{d}{f} \end{pmatrix}$$

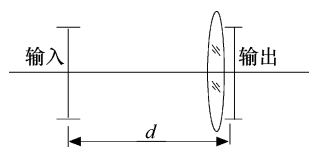


图 8.16 例 8.10 的光学系统

例 8.11 (a) 导出焦距为 f 的单薄透镜的成像方程；(b) 确定其放大率。

解：假设在距薄透镜 d_1 处有一物体，则在距薄透镜 d_2 处成像，如图 8.17 所示。

为了成像，从 A 点发出的光线一定会会聚于 B 点。总的光线传递矩阵如下：

$$\mathbf{M} = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{pmatrix}$$

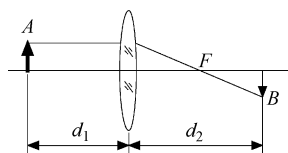


图 8.17 单透镜成像

y_1 与 y_2 有如下关系：

$$y_2 = \left(1 - \frac{d_2}{f}\right)y_1 + \left(d_1 + d_2 - \frac{d_1 d_2}{f}\right)\theta_1$$

y_2 不随 θ_1 变化，所以

$$d_1 + d_2 - \frac{d_1 d_2}{f} = 0$$

或者

$$\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2}$$

(b) 当成像发生时，有

$$y_2 = \left(1 - \frac{d_2}{f}\right)y_1$$

所以，放大率 A 为

$$A = \frac{y_2}{y_1} = 1 - \frac{d_2}{f}$$

例 8.12 在光学系统中，如果一个面上的发光强度分布是另一个面上发光强度分布的像（指两个面上的发光强度分布除了被放大或者缩小外，其分布是相同的），则把这两个面称作共轭面。(a) 求共轭面之间的光线传递矩阵。(b) 确定共轭面上角 θ_1 和 θ_2 的关系。

解：(a) 在之前的问题中，光线传递矩阵 \mathbf{M} 是在一般情况下计算得到的，故在成像条件下，传递矩阵 \mathbf{M} 变为

$$\mathbf{M} = \begin{pmatrix} m_y & 0 \\ -\frac{1}{f} & m_\theta \end{pmatrix}$$



其中, $m_y = (1 - d_2/f)$; $m_\theta = (1 - d_1/f)$ 。

(b) θ_1 和 θ_2 的关系如下:

$$\theta_2 = -\frac{y_1}{f} + m_\theta \theta_1$$

当 $y_1 = 0$ 时, 得到

$$\theta_2 = m_\theta \theta_1$$

所以, m_θ 即为光线的起始角度与光线会聚于光轴时的角度之间的放大因子。

例 8.13 考虑正透镜, 一束平行光线穿过正透镜并会聚于透镜后面的焦点, 这个焦点称作后焦点或第二焦点, 穿过后焦点且垂直于光轴的平面称作后焦面或第二焦面。

考虑从正透镜前的一点发出的光线通过正透镜后变为平行光束, 这个点称作透镜的前焦点或第一焦点。前后焦点和焦平面如图 8.18 和图 8.19 所示。确定正透镜前后焦平面之间的光线传递矩阵。

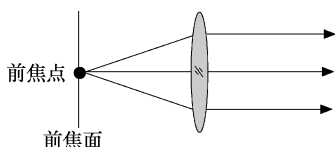


图 8.18 正透镜的前焦点及前焦平面

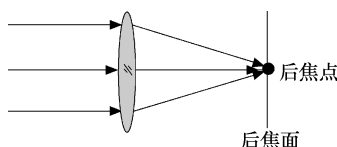


图 8.19 正透镜的后焦点及后焦平面

解: 在这种情况下, d_1 和 d_2 等于焦距 f , 所以光线传递矩阵 M 可变为

$$M = \begin{pmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{pmatrix} = \begin{pmatrix} 0 & f \\ -\frac{1}{f} & 0 \end{pmatrix} \quad (8.6-9)$$

例 8.14 类似正透镜, 可定义负透镜的焦点和焦平面, 如图 8.20 和图 8.21 所示。

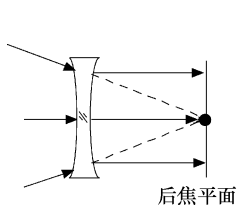


图 8.20 负透镜的后焦点及后焦平面

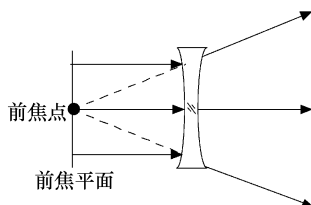


图 8.21 负透镜的前焦点及前焦平面

注意, 负透镜的前后焦点及其前后焦平面与正透镜的相反。负透镜前后焦平面之间的光线传递矩阵可由方程式 (8.6-9) 给出。

8.7 厚透镜

厚透镜可以像薄透镜一样, 通过定义两个主平面来分析。厚透镜的前焦点如图 8.22 所示, 由定义可知, 从前焦点发出的光线经厚透镜后变为平行光束射出。厚透镜上的入射光线



(源自前焦点) 和后向投射的出射光线 (平行光) 相交于透镜上的某一点, 穿过这点且垂直于光轴的平面称作第一主平面。

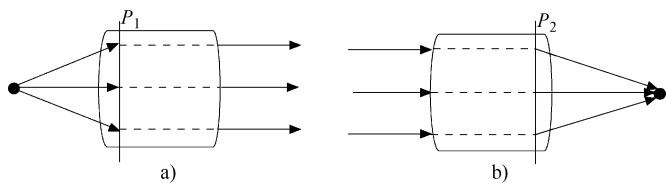


图 8.22 厚透镜的前焦点
a) P_1 : 第一主平面 b) P_2 : 第二主平面

由定义可知, 一束平行光线穿过厚透镜后会聚于厚透镜的第二焦点。入射光线 (平行光) 向后投射, 与会聚到第二焦点的出射光线相交于某一点, 穿过这点且垂直于光轴的平面称作第二主平面。

厚透镜的两个主平面之间的光线传递矩阵与正透镜的光线传递矩阵相同, 即

$$M = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix}$$

所以, 得到厚透镜的焦点、焦距以及成像距离如图 8.23 所示, 定义主平面和相应焦点之间的距离为焦距, 两个焦距是相同的。

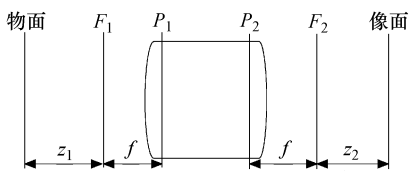


图 8.23 在厚透镜中的物、像、主平面、焦平面、焦距以及物与像的距离

例 8.15 证明两个共轭面间的距离 d_1 、 d_2 满足透镜成像定律。

解: 两共轭面间的光线变换矩阵如下:

$$M = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix}$$

上式和例 8.11 中讨论的光线变换矩阵相同, 所以, 成像条件为

$$\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2}$$

8.8 光学系统的入瞳和出瞳

一般的光学系统含有很多物理孔径, 其中只有一种孔径给系统造成了最严重的限制, 称它为有效物理孔径。



光学系统的入瞳是从物空间观察的有效物理孔径的像，而出瞳是从像空间观察的有效物理孔径的像。

为了阐述这个概念，分析图 8.24。在图 a 中，有效物理孔径在透镜面上，因而入瞳和出瞳重合，且等于有效物理孔径。在图 b 中，有效物理孔径与出瞳相同，物空间的像是虚像，为入瞳。在图 c 中，有效物理孔径与入瞳相同，像空间的像是虚像，为出瞳。

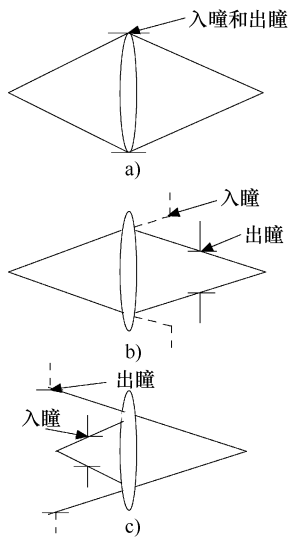


图 8.24 光学系统的入瞳和出瞳

a) 入瞳等于出瞳 b) 出瞳等于物理孔径 c) 入瞳等于物理孔径

考虑到衍射效应，正是出瞳有效地限制了通过光学系统的波场。就衍射而言，把透镜系统的出瞳作为它的孔径，在像平面上可以用夫琅禾费衍射来研究衍射效应，为此，光轴上的衍射距离即为出瞳到像平面的距离。

第 9 章 傅里叶变换和相干光学系统成像

9.1 本章引论

在本章里将讨论透镜的傅里叶变换和相干波成像，这两个问题都从衍射的角度进行讨论。

成像也可以从几何光学的观点出发，运用第 8 章中讲到的光线加以分析讨论，通过衍射与通过几何光学获得的结果是完全一致的。

本章共分 7 节。其中，9.2 节应用物理几何讨论了薄透镜的相变；9.3 节运用相函数阐明了透镜是怎样形成输出场的，它与各种几何形式的输入场的傅里叶变换有关；由此引出 9.4 节的关于成像的线性滤波的诠释。

9.5 节和 9.6 节分别通过相衬显微术和共焦扫描显微术来说明上述理论，最后一部分 9.7 节强调了复杂光学系统的代数算符。

9.2 薄透镜的相变

透镜由光学密质材料制成，这种材料中的相位速度小于空气中的相位速度。由于可以调整透镜的厚度，所以在透镜孔径处可以实现所需的相位调制。

薄透镜有三个非常重要的参数：材料的折射率 n ，通过这个参数可以减小相位速度； R_1 和 R_2 ，它们是透镜两个曲面的半径，如图 9.1 所示。如果透镜是玻璃材质的，折射率 n 大约为 1.5。按照惯例，随着光线从左到右传播，光线遇到的每个凸面都有正的曲率半径，每个凹面都有负的曲率半径。薄透镜有一个性质：光线入射、出射薄透镜都有相同的 (x, y) 坐标系。

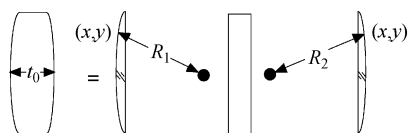


图 9.1 薄透镜的三个组成部分

上面讨论的薄透镜的三个参数可以由单独的一个参数 f 结合起来， f 为焦距，即

$$f = \left[(n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \right]^{-1} \quad (9.2-1)$$

焦距 f 的正负取决于 R_1 和 R_2 。例如，一个双凸透镜的焦距 f 为正，而双凹透镜的焦距 f 为负。在后面的讨论中，假设 f 为正值。

理想透镜的相位变换可以表示为

$$t(x, y) = e^{ikt_0} e^{-j\frac{k}{2f}(x^2 + y^2)} \quad (9.2-2)$$

式中， t_0 是透镜的最大厚度，由于 e^{ikt_0} 为常数相位改变，故可忽略。

对公式 (9.2-2) 进行如下推导。通过定义一个瞳孔函数 $P(x, y)$ 来考虑透镜的孔径

$$P(x, y) = \begin{cases} 1 & \text{在透镜孔径之内} \\ 0 & \text{其他} \end{cases} \quad (9.2-3)$$



为了推导公式 (9.2-2), 把透镜分为三部分来考虑, 如图 9.1 所示。在 (x, y) 处的厚度 $\Delta(x, y)$ 可以表示为 [Goodman, 2004];

$$\Delta(x, y) = t_0 - R_1 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_1^2}} \right) + R_2 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_2^2}} \right) \quad (9.2-4)$$

式中, t_0 是透镜的总厚度。

式 (9.2-4) 可以通过如下的傍轴近似简化:

$$\sqrt{1 - \frac{x^2 + y^2}{R_i^2}} \cong 1 - \frac{x^2 + y^2}{2R_i^2} \quad (9.2-5)$$

式中, R_i 即为 R_1 和 R_2 。

所以, 厚度函数可变为

$$\Delta(x, y) = t_0 - \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (9.2-6)$$

透镜的相位变化函数 $\theta(x, y)$ 为

$$\begin{aligned} \theta(x, y) &= kn\Delta(x, y) + k[t_0 - \Delta(x, y)] \\ &= kt_0 + k(n-1)\Delta(x, y) \end{aligned} \quad (9.2-7)$$

通过公式 (9.2-1) 和公式 (9.2-6), 公式 (9.2-7) 可化为

$$\theta(x, y) = kt_0 - \frac{k}{2f}(x^2 + y^2) \quad (9.2-8)$$

再由公式 (9.2-8) 可得公式 (9.2-2)。

9.3 透镜的傅里叶变换

在 5.4 节里, 夫琅禾费区处于距离适当大小孔径的初始平面很远的区域, 利用透镜可以消除这一限制。下面根据初始入射波场在透镜上、透镜前和透镜后三种不同情况进行讨论。

9.3.1 波场入射在透镜上

这种几何关系如图 9.2 所示。

假设入射到透镜上的波场为 $U(x, y, 0)$, 则透镜后的波场变为

$$U'(x, y, 0) = U(x, y, 0) P(x, y) e^{-i\frac{k}{2f}(x^2 + y^2)} \quad (9.3-1)$$

这个复振幅在式 (5.2-13) 中可用菲涅耳衍射公式代替, 从而找到 $z = f$ 即焦平面上的波场。如果焦平面坐标系为 (x_f, y_f, f) , 则结果如下:

$$U(x_f, y_f, f) = e^{i\frac{k}{2f}(x_f^2 + y_f^2)} \iint_{-\infty}^{\infty} U'(x, y, 0) P(x, y) e^{-i2\pi(f_x x + f_y y)} dx dy \quad (9.3-2)$$

这里常数项被忽略, 空间频率 $f_x = x_f/\lambda f$, $f_y = y_f/\lambda f$ 。式 (9.3-2) 说明输出复振幅是透镜瞳孔函数定义区域中的输入复振幅的夫琅禾费衍射条纹。

9.3.2 初始波场在透镜左侧

这种几何关系如图 9.3 所示。

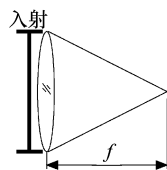


图 9.2 入射面在透镜上的几何关系



我们认为平面 $z = -d_1$ ($d_1 > 0$) 处的波场为 $U(x_1, y_1, -d_1)$, 其角谱 $A(f_x, f_y, -d_1)$ 和 $A(f_x, f_y, 0^-)$ 的关系为

$$A(f_x, f_y, 0^-) = A(f_x, f_y, -d_1) e^{-i\pi\lambda d_1 (f_x^2 + f_y^2)} \quad (9.3-3)$$

如果 $U(x, y, 0)$ 的范围小于 $P(x, y)$, 则式 (9.3-2) 可以写为

$$U(x_f, y_f, f) = A(f_x, f_y, 0) e^{i\frac{k}{2f}(x_f^2 + y_f^2)} \quad (9.3-4)$$

或者, 利用式 (9.3-3)

$$U(x_f, y_f, f) = A(f_x, f_y, -d_1) e^{i\frac{k}{2f}(1 - \frac{d_1}{f})(x_f^2 + y_f^2)} \quad (9.3-5)$$

当 $d_1 = f > 0$ 时, 相位因子变成 1, 则

$$U(x_f, y_f, f) = A(f_x, f_y, -f) \quad (9.3-6)$$

因此, 忽略 $P(x, y)$ 的影响后, 在 $z = -f$ 处的波场的二维傅里叶变换就是透镜焦平面上的波场。

由透镜有限孔径造成的这种限制被称为**渐晕**, 可以通过选择较小的 d_1 避免它。为了包含透镜有限孔径的影响, 可以使用几何光学近似法 [Goodman] 来实现。用这种方法, 初始波场是一近似值, 透镜的有限瞳孔函数等于 $P(x + (d_1/f)x_f, y + (d_1/f)y_f)$, 用该瞳孔函数可以计算 $A(f_x, f_y, -d_1)$, 其中 $f_x = x_f/\lambda f$, $f_y = y_f/\lambda f$ 。 $z = f$ 处的波场变为

$$U(x_f, y_f, f) = e^{i\frac{k}{2f}(1 - \frac{d_1}{f})(x_f^2 + y_f^2)} \iint_{-\infty}^{+\infty} U(x, y, -d_1) P\left(x + \frac{d_1}{f}x_f, y + \frac{d_1}{f}y_f\right) e^{-i\frac{2\pi}{\lambda f}(xx_f + yy_f)} dx dy \quad (9.3-7)$$

9.3.3 初始波场在透镜右侧

第三种可能性就是考虑初始波场在透镜后面, 其到透镜焦平面的距离为 d 。当用一个垂直平面波场照射透镜, 且透明物体位于 $z = f - d$ 处时才会出现这个情况, 如图 9.4 所示。

由几何光学可知, 在这个平面上球面波振幅正比于 f/d , 且可作为常数被忽略。如果 $(x_2, y_2, f - d)$ 是该平面上的一点, 且由于它是球面波会聚, 则该点对应的物面上的等效瞳孔函数为 $P(x_2(f/d), y_2(f/d))$ 。如果 $U(x_2, y_2, f - d)$ 基于透明物体的波场, 则物面内的总体波场为

$$U'(x_2, y_2, f - d) = e^{-i\frac{k}{2d}(x_2^2 + y_2^2)} P\left(x_2 \frac{f}{d}, y_2 \frac{f}{d}\right) U(x_2, y_2, f - d) \quad (9.3-8)$$

令 $A'(f_x, f_y, f - d)$ 为 $U'(x_2, y_2, f - d)$ 的角谱, 假如菲涅耳近似在距离 d 内有效, 则焦平面上波场的角谱为

$$A(x_f, y_f, f) = A'(f_x, f_y, f - d) e^{i\frac{k}{2d}(x_f^2 + y_f^2)} \quad (9.3-9)$$

这里常数项被忽略, 式中 f_x 和 f_y 分别是 $x_f/\lambda d$ 和 $y_f/\lambda d$ 。基于该几何关系, 可以通过调节 d 的大小改变傅里叶变换的范围, 变换的尺寸随着 d 的减小而变小, 在空间光滤波器中, 这是一个非常有用的特征, 因为傅里叶变换的尺寸和焦平面上空间滤波器的尺寸必须是相同的。

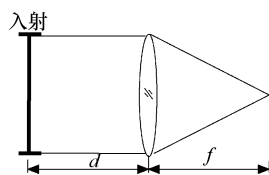


图 9.3 初始入射波场在透镜左侧的几何关系

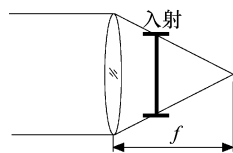


图 9.4 初始入射波场在透镜右侧的几何关系图



9.4 基于二维线性滤波的成像

透镜是众所周知的成像设备,下面讨论无像差正薄透镜的单色波的像。

在图 9.5 中,用一平面波照明位于平面 $z = -d_1$ 上的透明物体,则这个面上的波场为 $U(x_1, y_1, -d_1)$ 。首先确定平面 $z = d_0$ 上的波场 $U(x_0, y_0, d_0)$, 然后讨论成像的条件。由于波的传播是一种线性现象,所以在 (x_0, y_0, d_0) 和 $(x_1, y_1, -d_1)$ 处的波场总可以通过叠加积分联系起来

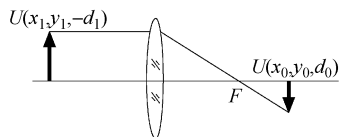


图 9.5 成像的几何图

$$U(x_0, y_0, d_0) = \iint_{-\infty}^{+\infty} h(x_0, y_0; x_1, y_1) U(x_1, y_1, -d_1) dx_1 dy_1 \quad (9.4-1)$$

式中, $h(x_0, y_0, x_1, y_1)$ 是系统的脉冲响应。

为了求 h , 假设波场 $U(x_1, y_1, -d_1)$ 为 $(x_1, y_1, -d_1)$ 处的 δ 函数, 这在物理上等效于起始于这一点的球面波。假设透镜在 $z=0$ 处有正的焦距 f , 忽略波传播中的所有常数项, 在傍轴近似下透镜的入射场如下:

$$U(x, y, 0) = e^{j[\frac{k}{2d_1}((x-x_1)^2 + (y-y_1)^2)]} \quad (9.4-2)$$

透镜后的波场为

$$U'(x, y, 0) = U(x, y, 0) P(x, y) e^{-j\frac{k}{2f}(x^2 + y^2)} \quad (9.4-3)$$

根据菲涅耳衍射式 (5.2-13), 有

$$U(x_0, y_0, d_0) = h(x_0, y_0, x_1, y_1) \iint_{-\infty}^{+\infty} U'(x, y, 0) e^{j[\frac{k}{2d_0}((x_0-x)^2 + (y_0-y)^2)]} dx dy \quad (9.4-4)$$

在实际的成像应用中, 最终的图像被一个只对发光强度敏感的检测系统检测到, 所以, 如果式 (9.4-4) 中的 $e^{j\theta}$ 形式的相位项是独立的, 则该相位项可以忽略不计, 则式 (9.4-4) 可以简化为

$$h(x_0, y_0; x_1, y_1) = \iint_{-\infty}^{+\infty} P(x, y) e^{j\frac{kD}{2}(x^2 + y^2)} e^{-jk[(\frac{x_0}{d_0} + \frac{x_1}{d_1})x + (\frac{y_0}{d_0} + \frac{y_1}{d_1})y]} dx dy \quad (9.4-5)$$

式中

$$D = \frac{1}{d_1} + \frac{1}{d_0} - \frac{1}{f} \quad (9.4-6)$$

在 $D=0$ 的情况下, 式 (9.4-5) 变为

$$h(x_0, y_0; x_1, y_1) = \iint_{-\infty}^{+\infty} P(x, y) e^{-j\frac{2\pi}{\lambda d_0}[(x_0 + Mx_1)x + (y_0 + My_1)y]} dx dy \quad (9.4-7)$$

这里

$$M = \frac{d_0}{d_1} \quad (9.4-8)$$

如果可以忽略 $P(x, y)$, 则式 (9.4-7) 等同于

$$h(x_0, y_0, x_1, y_1) = \iint_{-\infty}^{+\infty} e^{-j2\pi[(x_1 + \frac{x_0}{M})x' + (y_1 + \frac{y_0}{M})y']} dx' dy' \quad (9.4-9)$$

这里 $x' = x/\lambda d_0$; $y' = y/\lambda d_0$, 常数项被舍弃。

式 (9.4-9) 等价于



$$h(x_0, y_0; x_1, y_1) = \delta\left(x_1 + \frac{x_0}{M}, y_1 + \frac{y_0}{M}\right) \quad (9.4-10)$$

把式 (9.4-10) 代入式 (9.4-1) 得

$$U(x_0, y_0, d_0) = U\left(-\frac{x_0}{M}, -\frac{y_0}{M}, -d_1\right) \quad (9.4-11)$$

断定 $U(x_0, y_0, d_0)$ 是 $z = d_1$ 处波场的放大倒立像, M 是式 (9.4-6) 中 $D=0$ 时的放大率, 所以, d_0 和 d_1 的关系为

$$\frac{1}{d_0} = \frac{1}{f} - \frac{1}{d_1} \quad (9.4-12)$$

这就是人们熟知的透镜成像定律。

9.4.1 有限透镜孔径的影响

上述讨论中, 忽略了有限尺寸瞳孔函数 $P(x, y)$ 的影响。令

$$x'_1 = -Mx_1 \quad (9.4-13)$$

$$y'_1 = -My_1 \quad (9.4-14)$$

则式 (9.4-7) 可写为

$$\begin{aligned} h(x_0, y_0; x'_1, y'_1) &= \iint_{-\infty}^{+\infty} P(\lambda d_0 x, \lambda d_0 y) e^{-j2\pi[(x_0 - x'_1)x + (y_0 - y'_1)y]} dx dy \\ &= h(x_0 - x'_1, y_0 - y'_1) \end{aligned} \quad (9.4-15)$$

所以, 式 (9.4-1) 可写为

$$U(x_0, y_0, d_0) = \iint_{-\infty}^{+\infty} h(x_0 - x'_1, y_0 - y'_1) U(x'_1, y'_1, -d_1) dx'_1 dy'_1 \quad (9.4-16)$$

式 (9.4-16) 中忽略常数项。式 (9.4-16) 是一个二维卷积

$$U(x_0, y_0, d_0) = h(x_0, y_0) * U\left(-\frac{x_0}{M}, -\frac{y_0}{M}, -d_1\right) \quad (9.4-17)$$

这里 $U(-x_0/M, -y_0/M, -d_1)$ 是理想图像, 且

$$h(x_0, y_0) = \iint_{-\infty}^{+\infty} P(\lambda d_0 x, \lambda d_0 y) e^{-j2\pi(x_0 x + y_0 y)} dx dy \quad (9.4-18)$$

我们发现脉冲响应是透镜的变尺度瞳孔函数的二维傅里叶变换, 最终成像是理想图像和系统脉冲响应的卷积。平滑操作能强烈地衰减图像的细节。

在多透镜的一般成像系统中, 式 (9.4-17) 和式 (9.4-18) 仍然有效, 假设 $P(\cdot, \cdot)$ 表示衍射受限系统的有限等效出瞳, 如果入射到入瞳的发散球面波在出瞳处映射为会聚球面波, 则这个光学系统就是衍射受限的。

9.5 相衬显微术

本小节和下一节的内容, 是用两种先进的成像技术阐述前面几节中讨论的理论。相衬显微术是一种对透明物体生成高对比度图像的技术, 如活细胞培养物、薄的组织切片、微生物、光刻图案、纤维等。该技术是通过把小的相位变化转换为可以高对比度观察的振幅变化来实现的, 在这个过程中, 标本不会受到不良干扰和破坏。



工业相衬显微镜的原理图如图 9.6 所示。

当光穿过透明物体时，它每点的相位都发生了变化，所以，相干照明的样本的传输函数可写为

$$t(x, y) = e^{j[\theta_0 + \theta(x, y)]} \quad (9.5-1)$$

式中， θ_0 为平均相位；相移函数 $\theta(x, y)$ 远小于 2π 。

所以， $t(x, y)$ 可近似为

$$t(x, y) = e^{j\theta_0} [1 + j\theta(x, y)] \quad (9.5-2)$$

式 (9.5-2) 中的最后一个因子是函数 $e^{j\theta(x, y)}$ 泰勒展开式的开始两项。

显微镜对发光强度的敏感性可表示为

$$I(x, y) = |1 + j\theta(x, y)|^2 \approx 1 \quad (9.5-3)$$

所以，图像不清晰。式 (9.5-3) 是正确的，因为它是由非衍射光引起的前一项单位 1 与由衍射产生的 $j\theta(x, y)$ 一起求二次积。为了绕开这个问题，可以使用非衍射光发生 $\pi/2$ 或 $3\pi/2$ 相移的相平面，这可通过透明电介质点的玻璃基板来实现，通常通过控制成像系统的焦点处的厚度给出 $\pi/2$ 或 $3\pi/2$ 相移。该非衍射光通过焦点，而来自样本的衍射光由于具有高空间频率，它在焦点处扩散开。在像平面上，发光强度可以表示为

$$I(x, y) = |e^{j\pi/2} + j\theta(x, y)|^2 \approx 1 + 2\theta(x, y) \quad (9.5-4)$$

式 (9.5-4) 的相位移是 $\pi/2$ ，还有

$$I(x, y) = |e^{j3\pi/2} + j\theta(x, y)|^2 \approx 1 - 2\theta(x, y) \quad (9.5-5)$$

相位移是 $3\pi/2$ 。式 (9.5-4) 和式 (9.5-5) 分别为正相衬和负相衬，在任一情况下，相位变化 $\theta(x, y)$ 可以作为图像来观察，因为它已被转换为发光强度。常规显微镜与相衬显微镜成像的相互比较如图 9.7 所示。

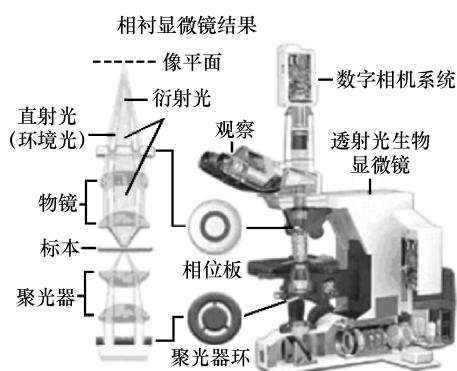


图 9.6 工业相衬显微镜的原理图 [Nikon]

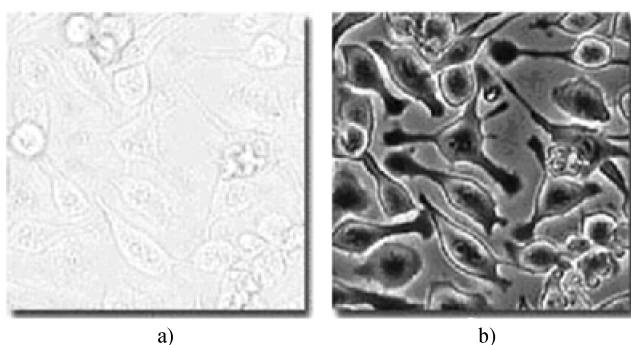


图 9.7 常规显微镜与相衬显微镜成像的相互比较

a) 常规显微镜下标本的图像 b) 相衬显微镜下相同标本的图像 [Nikon]

通过设计更加复杂的相盘有可能进一步改进此方法。例如，对比度可通过改变相盘的特性来调整，诸如改变吸收率、折射率和厚度等。切趾相衬物镜也可用来进行系统改进。



9.6 共焦扫描显微术

传统显微镜是一个将整个目标场同时成像的设备，而在扫描显微镜中，一次只能产生一个目标点的像，这需要扫描目标来产生整个场的像。反射式共焦扫描光学显微镜的原理图如图 9.8 所示，这样的系统也可以在透射模式下运行。系统把图像信息传给计算机，进行数字后处理和图像处理。

在共焦扫描显微镜中，点光源发出的光只能探照目标的极小区域，另一个点探测器确保只能探测到来自相同目标区域的光。共焦扫描显微镜的一种特定配置如图 9.8 所示。在这个系统中，图像是通过同步地扫描源和探测器生成的。另一种配置如图 9.9 所示，在这种模式下，共焦显微镜具有优良的深度辨别特性，利用这个特性，厚的半透明物体的平行截面可以在高分辨率下成像。

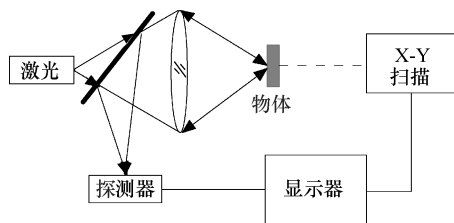


图 9.8 反射式共焦扫描显微镜的原理图

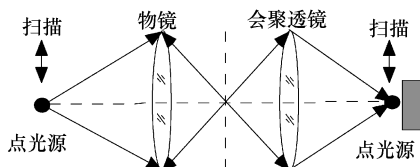


图 9.9 共焦光学显微镜系统

9.6.1 图像形成

式 (9.4-16) 可写为

$$U(x_0, y_0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h\left(x_1 + \frac{x_0}{M}, y_1 + \frac{y_0}{M}\right) U(x_1, y_1) dx_1 dy_1 \quad (9.6-1)$$

可以看到，式中 d_0 和 d_1 已从输入输出的复振幅中去掉。式 (9.4-18) 给出的脉冲响应函数可以写为

$$h(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(\lambda d_0 x_1, \lambda d_0 y_1) e^{-2\pi j(x_1 x + y_1 y)} dx_1 dy_1 \quad (9.6-2)$$

式中， P 是透镜系统的孔径函数。

在共焦显微镜中，由于一次只能对一个目标点成像， $U(x_1, y_1)$ 可以写为

$$U(x_1, y_1) = \delta(x_1) \delta(y_1) \quad (9.6-3)$$

所以

$$\begin{aligned} U(x_0, y_0) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h\left(x_1 + \frac{x_0}{M}, y_1 + \frac{y_0}{M}\right) \delta(x_1) \delta(y_1) dx_1 dy_1 \\ &= h\left(\frac{x_0}{M}, \frac{y_0}{M}\right) \end{aligned} \quad (9.6-4)$$

点图像的强度变为

$$I(x_0, y_0) = \left| h\left(\frac{x_0}{M}, \frac{y_0}{M}\right) \right|^2 \quad (9.6-5)$$



对于一个半径为 α 的圆对称孔径, 则 $P(x_1, y_1)$ 可写为

$$(x_1, y_1) = P(\rho) = \begin{cases} 1 & \rho \leq 1/2 \\ 0 & \text{其他} \end{cases} \quad (9.6-6)$$

其中

$$\rho = \frac{\sqrt{x_1^2 + y_1^2}}{2\alpha} \quad (9.6-7)$$

$P(\rho)$ 为例 2.6 中讨论过的柱函数。脉冲响应函数变为 $P(\rho)$ 的汉克尔变换, 由下式给出:

$$h(x_0, y_0) = h(r_0) = \frac{\pi}{4} \text{somb}(r_0) \quad (9.6-8)$$

其中

$$r_0 = \sqrt{x_0^2 + y_0^2} \quad (9.6-9)$$

$$\text{somb}(r_0) = \frac{2J_1(\pi r_0)}{\pi r_0} \quad (9.6-10)$$

该点图像的强度变为

$$I(x_0, y_0) = \left| \frac{MJ_1(\pi r_0/M)}{2r_0} \right|^2 \quad (9.6-11)$$

实际中通常使用两种透镜, 如图 9.9 所示, 称它们为物镜和会聚透镜。假设这两个透镜的脉冲响应函数分别为 h_1 和 h_2 , 通过重复应用方程式 (9.6-1), 图形强度可以表示如下:

$$I(x_0, y_0) = |h_t * U(x_1, y_1)|^2 \quad (9.6-12)$$

这里

$$h_t = h_1 h_2 \quad (9.6-13)$$

如果两透镜相同, 即脉冲响应函数 h 由方程式 (9.6-8) 给出, 则图像的输出强度为

$$I(x_0, y_0) = \left| \frac{MJ_1(\pi r_0/M)}{2r_0} \right|^4 \quad (9.6-14)$$

例 9.1 确定共焦扫描显微镜两透镜的传递函数。

解: 由 $h_t = h_1 h_2$ 可知

$$H_t(f_x, f_y) = H_1(f_x, f_y) * H_2(f_x, f_y)$$

这里由于两透镜相同, 故

$$H_1(f_x, f_y) = H_2(f_x, f_y) = P(\lambda d_0 x_1, \lambda d_0 y_1)$$

如果孔径函数 P 是圆对称的, 则 $H_t(f_x, f_y)$ 也是对称的, 它可表示为

$$H_t(f_x, f_y) = H_t(\rho)$$

这里

$$\rho = \sqrt{f_x^2 + f_y^2}$$

所以

$$H_t(\rho) = H_1(\rho) * H_2(\rho)$$

这个卷积可由下式给出:



$$H_t(\rho) = \frac{2}{\pi} \left[\arccos\left(\frac{\rho}{2}\right) - \frac{\rho}{2} \sqrt{1 - \left(\frac{\rho}{2}\right)^2} \right]$$

这里

$$P(\rho) = \begin{cases} 1 & 0 \leq \rho \leq 1 \\ 0 & \text{其他} \end{cases}$$

9.7 复杂光学系统的算子代数

用算子代数把 9.1 节到 9.5 节的内容结合起来，就可分析相干照明下的复杂光学系统 [Goodman]。为了简化得到的结果，假设通过傍轴近似讨论菲涅耳衍射和几何光学所得到的结论是有效的。

下面讨论一种代数算符。假设孔径在矩形坐标中是可分离的，因此可用一维系统的方式讨论二维系统。该算子包括复杂光学系统中用到的基本运算。如果当前的光场为 $U(x)$ ，则 $O(\mu)[U_1(x)]$ 表示 $U_1(x)$ 通过算子 $O(\mu)$ 的变换，该算符以 μ 为变量。下面介绍一些重要的运算符。

1. 傅里叶变换

$$U_2(\nu) = F(\nu)[U_1(x)] = \int_{-\infty}^{+\infty} U_1(x) e^{-j2\pi\nu x} dx \tag{9.7-1}$$

式中， ν 表示频率。

2. 自由空间传播

$$U_2(x_2) = P(z)[U_1(x_1)] = \frac{1}{\sqrt{j\lambda z}} \int_{-\infty}^{+\infty} U_1(x_1) e^{j\frac{k}{2z}(x_2-x_1)^2} dx_1 \tag{9.7-2}$$

式中， $k=2\pi/\lambda$ ； λ 为波长； z 为传播距离。

由于波的传播距离小于菲涅耳区，故可用角谱法代替菲涅耳衍射。

3. 常数缩放变换

$$U_2(x) = S(\alpha)[U_1(x)] = \sqrt{|\alpha|} U_1(\alpha x) \tag{9.7-3}$$

式中， α 为缩放常数。

4. 乘以二次相位因子

$$U_2(x) = Q(b)[U_1(x)] = e^{j\frac{k}{2}bx^2} U_1(x) \tag{9.7-4}$$

式中， b 为二次相位因子的控制参数。

以上每一个运算都可以进行逆运算，见表 9.1。

表 9.1 逆算符

	$F(\nu)$	$P(z)$	$S(a)$	$Q(b)$
逆 算 符	$F(-\nu)$	$P(-z)$	$s\left(\frac{1}{a}\right)$	$Q(-b)$

为了有效地使用算子代数，有必要知道两种连续运算是如何相关的，表 9.2 中总结了这些关系。



表 9.2 算符和运算

	$F(\nu)$	$P(z)$	$S(a)$	$Q(b)$
$F(\nu)$	$S(-1)$	$Q(-\lambda_z^2)F(\nu)$	$S\left(\frac{1}{a}\right)F(\nu)$	$P\left(-\frac{b}{\lambda^2}\right)F(\nu)$
$P(z)$	$F(\nu)Q(-\lambda^2 z)$	$P(z_2)P(z_1)$ $=P(z_1+z_2)$	$S(a)P(a_2^2)$	$Q\left[\left(\frac{1}{b}+z\right)^{-1}\right] \cdot$ $S[(1+bz)^{-1}] \cdot$ $P[(z^{-1}+b)^{-1}]$
$S(a)$	$F(\nu)S\left(\frac{1}{a}\right)$	$P\left(\frac{z}{a^2}\right)S(a)$	$S(a_2)S(a_1)$ $=S(a_1+a_2)$	$Q(a^2b)S(a)$
$Q(b)$	$F(\nu)P\left(-\frac{b}{\lambda^2}\right)$	$P[(z^{-1}+b)^{-1}] \cdot$ $S(1+bz) \cdot$ $Q[(b^{-1}+z)^{-1}]$	$S(a)Q\left(\frac{b}{a^2}\right)$	$Q(b_2)Q(b_1)$ $=Q(b_1+b_2)$

在这个表里，每一个单元 E 等于一组连续运算操作，包括行算符 $R(\nu_1)$ 然后是列算符 $C(\nu_2)$

$$E = C(\nu_2)R(\nu_1) \quad (9.7-5)$$

另外，常用下面的关系式来简化结果：

$$P(z) = Q\left(\frac{1}{z}\right)S\left(\frac{1}{\lambda z}\right)F(\nu)Q\left(\frac{1}{z}\right) \quad (9.7-6)$$

$$S\left(\frac{1}{\lambda f}\right)F(\nu) = P(f)Q\left(\frac{1}{f}\right)P(f) \quad (9.7-7)$$

方程式 (9.7-6) 对应于菲涅耳衍射，方程式 (9.7-7) 表明透镜前后焦平面通过傅里叶变换相关。

例 9.2 (a) 简化下面算符方程：

$$O = Q\left(-\frac{1}{f}\right)P(f)Q\left(-\frac{1}{f}\right)P(f)Q\left(-\frac{1}{f}\right)P(f)$$

(b) 该算符对应于哪个光学系统？

(c) 确定初始场 U_1 的最后输出场 U_{end} 。

解：(a) 右边的前三个算符根据方程式 (9.7-7) 符合

$S\left(\frac{1}{\lambda f}\right)F(\nu)$ 。下一个 $P(f)$ 根据方程式 (9.7-6) 被

$Q\left(\frac{1}{f}\right)S\left(\frac{1}{\lambda f}\right)F(\nu_2)Q\left(\frac{1}{f}\right)$ 代替。得到

$$O = S\left(\frac{1}{\lambda f}\right)F(\nu_2)S\left(\frac{1}{\lambda f}\right)F(\nu_1)$$

(b) 该算符与图 9.10 中所示的光学系统相符合。

(c) 把每一个算符化为积分方程

$$U_2(\nu_1) = F(\nu_1)[U_1(x)] = \int_{-\infty}^{\infty} U_1(x) e^{-j2\pi\nu_1 x} dx$$

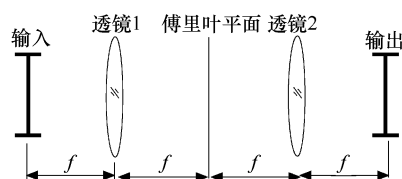


图 9.10 例 9.2 的光学系统



$$U_3(\nu_1) = S\left(\frac{1}{\lambda f}\right)[U_2(\nu_1)] = \frac{1}{\sqrt{\lambda f}} \int_{-\infty}^{+\infty} U_1(x) e^{-j2\pi \frac{\nu_1}{\lambda f} x} dx$$

$$U_4(\nu_2) = F(\nu_2)[U_3(\nu_1)] = \int_{-\infty}^{+\infty} U_3(\nu_1) e^{-j2\pi \nu_2 \nu_1} d\nu_1$$

$$U_5(\nu_2) = S\left(\frac{1}{\lambda f}\right)[U_4(\nu_2)] = \frac{1}{\sqrt{\lambda f}} \int_{-\infty}^{+\infty} U_3(\nu_1) e^{-j2\pi \frac{\nu_2}{\lambda f} \nu_1} d\nu_1$$

对 $U_3(\nu_1)$ 使用积分方程, 则最后一个方程可写为

$$\begin{aligned} U_5(\nu_2) &= \frac{1}{\lambda f} \iint_{-\infty}^{+\infty} U_1(x) e^{-j2\pi \frac{\nu_1}{\lambda f} x} e^{-j2\pi \frac{\nu_2}{\lambda f} \nu_1} dx d\nu_1 \\ &= \frac{1}{\lambda f} \int_{-\infty}^{+\infty} U_1(x) dx \left[\int_{-\infty}^{+\infty} e^{-j2\pi \left(\frac{x+\nu_2}{\lambda f}\right) \nu_1} d\nu_1 \right] \\ &= \frac{1}{\lambda f} \int_{-\infty}^{+\infty} U_1(x) \delta\left(\frac{x+\nu_2}{\lambda f}\right) dx \\ &= \int_{-\infty}^{+\infty} U_1(x) \delta(x+\nu_2) dx \\ &= U_1(-\nu_2) \end{aligned}$$

例 9.3 证明如下的光学系统可以对空间频率进行傅里叶变换, 通过改变 z 可调整空间频率。

解: 连接输入和输出的算符 O 由下式求得:

$$O = P(z_2) Q\left(-\frac{1}{f}\right) P(z) Q\left(\frac{1}{z_1 - z}\right) \quad (9.7-8)$$

通过运用透镜定律, 即 $\frac{1}{f} = \frac{1}{z_1} + \frac{1}{z_2}$, 则 $Q\left(-\frac{1}{f}\right)$ 可写为

$$Q\left(-\frac{1}{f}\right) = Q\left(-\frac{1}{z_1} - \frac{1}{z_2}\right) \quad (9.7-9)$$

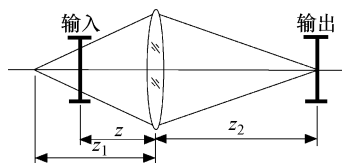


图 9.11 例 9.3 的光学系统

根据表 9.2, 可运用下列关系:

$$P(z_2) Q\left(-\frac{1}{z_1} - \frac{1}{z_2}\right) = Q\left(\frac{z_1 + z_2}{z_2^2}\right) S\left(-\frac{z_1}{z_2}\right) P(-z_1) \quad (9.7-10)$$

$$P(z - z_1) = Q\left(\frac{1}{z - z_1}\right) S\left(-\frac{z_1}{z_2}\right) Q\left(\frac{1}{z - z_1}\right) S\left(\frac{1}{\lambda(z - z_1)}\right) F(r) \quad (9.7-11)$$

由以上方程可得到

$$O = Q\left(\frac{z_1 + z_2}{z_2^2}\right) S\left(-\frac{z_1}{z_2}\right) Q\left(\frac{1}{z - z_1}\right) S\left(\frac{1}{\lambda(z - z_1)}\right) F(r)$$

用表 9.2 中的

$$Q(b) S(a) = S(a) Q\left(\frac{b}{a^2}\right)$$

得

$$S\left(-\frac{z_1}{z_2}\right) Q\left(\frac{1}{z - z_1}\right) = Q\left(\frac{z_1 + z_2}{z_2^2}\right) S\left(-\frac{z_1}{z_2}\right)$$



从而得到

$$O = Q\left(\frac{z_1 + z_2}{z_2^2}\right) Q\left(\frac{z_1 + z_2}{z_1^2}\right) S\left(-\frac{z_1}{z_2}\right) S\left(\frac{1}{\lambda(z - z_2)}\right) F(r)$$

相继将 Q 算符、 S 算符合并有

$$O = Q\left(\frac{(z_1 + z_2)z - z_1 z_2}{z_2^2(z - z_1)}\right) S\left(\frac{z_1}{\lambda z_2(z_1 - z)}\right) F(r)$$

作为积分方程, O 对应于

$$U_2(x_2) = e^{j\frac{k}{2} \frac{(z_1 + z_2)z - z_1 z_2}{z_2^2(z - z_1)} x_0^2} \int_{-\infty}^{+\infty} U_1(x) e^{-j\frac{2\pi z_1 x_0}{\lambda z_2(z_1 - z)}} dx$$

可以通过改变 z 值, 调整空间频率 $\frac{z_1 x_0}{z_2(z_1 - z)}$ 。

第 10 章 准单色波成像

10.1 本章引论

在前几章里，讨论了相干和单色波场的传播与成像，而在本章中，这些内容扩展为准单色或非相干波场的传播与成像。

单色波场只有一个时间频率 f ，而非单色波场则有许多。准单色波场有一个时间频率范围为 Δf ，它远小于平均时间频率 f_c 。在实际的成像应用中，通常假定波场为准单色波场。

本章分 10 个小节，前几个小节我们做了相关理论的准备工作，这些理论可以恰当地分析准单色波。10.2 小节介绍了一种与傅里叶变换紧密相关的希尔伯特变换。它的主要特性是交换了余弦和正弦频率组分。它是一种工具，用来定义 10.3 节介绍的解析信号。解析信号是一个复数，它的实部等于真实的信号，而它的虚部等于相同真实信号的希尔伯特变换。希尔伯特变换除了用在分析准单色波，它还被广泛地用于分析通信中的单边带调制。

10.4 节介绍了怎样用解析信号来表示准单色波。10.5 节通过空间相干和时间相干的概念进一步研究了准单色、相干和非相干。

面向简单光学系统的上述理论的研究，在 10.6 节被扩展到更加复杂的成像系统。准单色波成像是 10.7 节的主要内容，在这节中，我们可以更加清楚地区别相干成像和非相干成像。10.8 节中衍射受限成像系统被认为是线性系统，通过相干成像和非相干成像导出了该线性系统的特性，光学传递函数就是其中之一。10.9 节主要介绍了光学传递函数的计算方法。所有成像系统都有像差，10.10 节对此进行了介绍，尤其是泽尔尼克多项式（Zernike polynomials）。

10.2 希尔伯特变换

希尔伯特变换和在 10.3 节里讨论的解析信号在描述准单色光场和成像等的性质方面非常有用，它们还广泛使用在其他领域，如通信中的单边带调制。定义实数信号的希尔伯特变换为

$$\nu(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{u(\tau)}{t - \tau} d\tau \quad (10.2-1)$$

它是 $u(t)$ 和函数 $1/\pi t$ 的卷积。

用卷积理论，方程式（10.2-1）的傅里叶变换为

$$V(f) = U(f)H(f) \quad (10.2-2)$$

式中， $H(f)$ 是函数 $1/\pi t$ 的傅里叶变换。

即

$$H(f) = -j \operatorname{sgn}(f) \quad (10.2-3)$$



定义符号函数 $\text{sgn}(f)$ 为

$$\text{sgn}(f) = \begin{cases} 1 & f > 0 \\ 0 & f = 0 \\ -1 & f < 0 \end{cases} \quad (10.2-4)$$

由帕塞瓦尔 (Parseval) 定理和方程式 (10.2-2) 可得, $v(t)$ 和 $u(t)$ 的能量相等

$$\int_{-\infty}^{+\infty} v^2(t) dt = \int_{-\infty}^{+\infty} u^2(t) dt \quad (10.2-5)$$

方程式 (10.2-3) 表明通过希尔伯特变换, 频谱组分相位移动了 $\pi/2$, 所以又称希尔伯特变换为**正交滤波器**。

希尔伯特变换的逆变换为

$$u(t) = -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{v(\tau)}{t - \tau} d\tau \quad (10.2-6)$$

希尔伯特变换中的这个积分为反常积分, 它实际上是下式的缩写:

$$\int_{-\infty}^{+\infty} \frac{u(\tau)}{t - \tau} d\tau = \lim_{\varepsilon \rightarrow 0} \left[\int_{-\infty}^{t-\varepsilon} \frac{u(\tau)}{t - \tau} d\tau + \int_{t+\varepsilon}^{+\infty} \frac{u(\tau)}{t - \tau} d\tau \right] \quad (10.2-7)$$

它又可表示为

$$\int_{-\infty}^{+\infty} \frac{u(\tau)}{t - \tau} d\tau = \int_{-\infty}^{+\infty} \frac{u(t + \tau)}{-\tau} d\tau \quad (10.2-8)$$

例 10.1 求函数 $u(t) = \cos(2\pi f_0 t + \phi(f_0))$ 的希尔伯特变换。

解:

$$\begin{aligned} v(t) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\cos(2\pi f_0(\tau) + \phi(f_0))}{t - \tau} d\tau = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\cos(2\pi f_0(t + \tau) + \phi(f_0))}{\tau} d\tau \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \left[\frac{\cos(2\pi f_0 t + \phi(f_0)) \cos(2\pi f_0 \tau + \phi(f_0))}{-\tau} - \frac{\sin(2\pi f_0 t + \phi(f_0)) \sin(2\pi f_0 \tau + \phi(f_0))}{-\tau} \right] d\tau \end{aligned} \quad (10.2-9)$$

第一项是 τ 的奇函数, 积分为 0, 所以有

$$v(t) = \frac{\sin(2\pi f_0 t + \phi(f_0))}{\pi} \int_{-\infty}^{+\infty} \frac{\sin(2\pi f_0 \tau + \phi(f_0))}{\tau} d\tau = \sin(2\pi f_0 t + \phi(f_0)) \quad (10.2-10)$$

例 10.2 证明实数信号 $u(t)$ 和它的希尔伯特变换 $v(t)$ 相互正交。

解: 由帕塞瓦尔定理可得

$$\begin{aligned} \int_{-\infty}^{+\infty} u(t)v(t) dt &= \int_{-\infty}^{+\infty} U(f)V^*(f) df \\ &= j \int_0^{+\infty} |U(f)|^2 df - j \int_{-\infty}^0 |U(f)|^2 df = 0 \end{aligned} \quad (10.2-11)$$

其中 $U(-f) = V^*(f)$ 。

例 10.3 $u(t)$ 的傅里叶变换可写为

$$U(f) = U_1(f) - jU_0(f) \quad (10.2-12)$$



这里 $U_1(f)$ 和 $U_0(f)$ 分别是 $u(t)$ 傅里叶变换的余弦和正弦部分。

证明当 $u(t)$ 是因果信号 (当 $t < 0$ 时 $u(t) = 0$) 时, $U_1(f)$ 和 $U_0(f)$ 是希尔伯特变换对。

证明: $u(t)$ 的奇偶部分可表示为

$$\begin{aligned} u_1(t) &= \frac{1}{2}[u(t) + u(-t)] \\ u_0(t) &= \frac{1}{2}[u(t) - u(-t)] \end{aligned} \quad (10.2-13)$$

当 $u(t)$ 为因果信号时, 显然有

$$\begin{aligned} u_1(t) &= u_0(t) \operatorname{sgn}(t) \\ u_0(t) &= u_1(t) \operatorname{sgn}(t) \end{aligned} \quad (10.2-14)$$

所以

$$\begin{aligned} U_1(f) &= \int_{-\infty}^{+\infty} u_1(t) e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{+\infty} u_0(t) \operatorname{sgn}(t) e^{-j2\pi ft} dt \end{aligned} \quad (10.2-15)$$

通过傅里叶变换的调制性质, $U_1(f)$ 等于 $u_0(t)$ 的傅里叶变换 (即 $-jU_0(f)$) 和 $\operatorname{sgn}(t)$ 的傅里叶变换 (即 $\frac{1}{j\pi f}$) 的卷积。

$$U_1(f) = \int_{-\infty}^{+\infty} -jU_0(\nu) \frac{1}{j\pi(f-\nu)} d\nu \quad (10.2-16)$$

所以

$$U_1(f) = -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{U_0(\nu)}{f-\nu} d\nu \quad (10.2-17)$$

类似地有

$$U_0(f) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{U_1(\nu)}{f-\nu} d\nu \quad (10.2-18)$$

式 (10.2-16) ~ 式 (10.2-18) 中的 ν 为频率。方程式 (10.2-17) 和方程式 (10.2-18) 也被称为电磁学中的克拉茂-克朗尼希 (Kramers-Krönig) 关系。

10.3 解析信号

解析信号对理解窄带波形、波传播和图像结构的性质非常有用。定义

$$s(t) = u(t) + jv(t) \quad (10.3-1)$$

式中, $u(t)$ 是实数信号; $v(t)$ 是 $u(t)$ 的希尔伯特变换。

所以, 解析信号是一种把实数信号转换为复数信号的方法。对方程式 (10.3-1) 的两边分别进行傅里叶变换得

$$S(f) = U(f) + jV(f) \quad (10.3-2)$$

当 $V(f) = -j \operatorname{sgn}(f) U(f)$ 时, 得

$$S(f) = \begin{cases} U(f) & f=0 \\ 2U(f) & f>0 \\ 0 & f<0 \end{cases} \quad (10.3-3)$$



所以, 解析信号为

$$s(t) = 2 \int_0^{+\infty} U(f) e^{j2\pi ft} df \quad (10.3-4)$$

解析信号可以进一步写为

$$s(t) = |s(t)| e^{j\phi(t)} \quad (10.3-5)$$

其中

$$|s(t)|^2 = u^2(t) + v^2(t) \quad (10.3-6)$$

$$\phi(t) = \arctan \frac{v(t)}{u(t)} \quad (10.3-7)$$

解析信号常被用于窄带波形, 比如中央频率为 f_c 的准单色波场, 故 $\phi(t)$ 可以表示为

$$\phi(t) = 2\pi f_c t + \phi'(t) \quad (10.3-8)$$

方程式 (10.3-5) 可以表示为

$$s(t) = \mu(t) e^{j2\pi f_c t} \quad (10.3-9)$$

其中

$$\mu(t) = |s(t)| e^{j\phi'(t)} \quad (10.3-10)$$

式中, $\mu(t)$ 为复包络, 在光学中, 它也被称为相量幅值 (phasor amplitude)。

例 10.4 求与 $u(t) = \cos(2\pi ft + \phi(f))$ 对应的解析信号。

解: 在例 10.1 中, 得到

$$v(t) = \sin(2\pi ft + \phi(f)) \quad (10.3-11)$$

所以

$$s(t) = \cos(2\pi ft) + j \sin(2\pi ft) = e^{j2\pi ft} \quad (10.3-12)$$

例 10.5 求解析信号的能量。

解: 解析信号的能量为

$$\int_{-\infty}^{+\infty} |s(t)|^2 dt = \int_{-\infty}^{+\infty} u^2(t) dt + \int_{-\infty}^{+\infty} v^2(t) dt = 2 \int_{-\infty}^{+\infty} u^2(t) dt = 2 \int_{-\infty}^{+\infty} |U(f)|^2 df \quad (10.3-13)$$

可以发现, 相应实信号的解析信号拥有双倍的能量。

例 10.6 低通信号 $u(t)$ 的振幅谱如图 10.1 所示。用 $\cos(2\pi f_0 t)$

调制产生 $g(t) = u(t) \cos(2\pi f_0 t)$ 。

(a) 求 $g(t)$ 的振幅谱并画图;

(b) 画出从 $u(t)$ 产生的解析信号的振幅谱;

(c) 求 $p(t) = u(t) \cos(2\pi f_0 t) - v(t) \sin(2\pi f_0 t)$ 的振幅谱, 并

画图。

解: (a) $g(t)$ 可以表示为

$$g(t) = \frac{u(t)}{2} e^{j2\pi f_0 t} + \frac{u(t)}{2} e^{-j2\pi f_0 t} \quad (10.3-14)$$

$g(t)$ 的傅里叶变换为

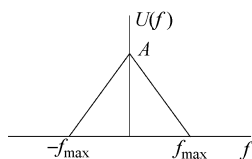


图 10.1 低通信号的频谱



$$G(f) = \frac{1}{2}U(f-f_0) + \frac{1}{2}U(f+f_0) \quad (10.3-15)$$

则振幅频谱为

$$|G(f)| = \frac{1}{2}|U(f-f_0) + U(f+f_0)| \quad (10.3-16)$$

$|G(f)|$ 如图 10.2 所示。

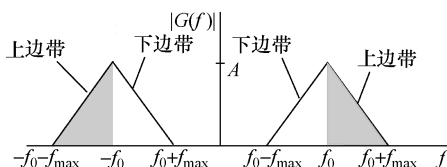


图 10.2 $|G(f)|$ 的振幅频谱

在 $f_0 \leq f \leq f_0 + f_{\max}$ 和 $(-f_0 - f_{\max}) \leq f \leq -f_0$ 范围内的频率组分就是大家非常熟知的上边带，在 $-f_0 \leq f \leq (-f_0 + f_{\max})$ 和 $f_0 - f_{\max} \leq f \leq f_0$ 范围内的频率组分即为下边带。

(b) 由 $u(t)$ 产生的解析信号为

$$s(t) = u(t) + jv(t) \quad (10.3-17)$$

它的傅里叶变换为

$$\begin{aligned} S(f) &= U(f) + jV(f) \\ &= U(f) [1 + jH(f)] \\ &= \begin{cases} 2U(f) & f \geq 0 \\ 0 & f \leq 0 \end{cases} \end{aligned} \quad (10.3-18)$$

$s(t)$ 的振幅频谱如图 10.3 所示。

(c) $p(t) = u(t) \cos(2\pi f_0 t) - v(t) \sin(2\pi f_0 t)$ 可写为

$$\begin{aligned} p(t) &= \operatorname{Re}[(u(t) + jv(t))e^{j2\pi f_0 t}] \\ &= \operatorname{Re}[s(t)e^{j2\pi f_0 t}] \\ &= \frac{1}{2}[s(t)e^{j2\pi f_0 t} + s^*(t)e^{-j2\pi f_0 t}] \end{aligned} \quad (10.3-19)$$

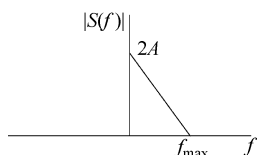


图 10.3 解析信号 $s(t)$ 的振幅频谱

$p(t)$ 的傅里叶变换为

$$\begin{aligned} P(f) &= \frac{1}{2}S(f-f_0) + \frac{1}{2}S^*(-f-f_0) \\ &= \begin{cases} U(f-f_0) & f_0 \leq f \leq f_0 + f_{\max} \\ U^*(-f-f_0) & -f_0 - f_{\max} \leq f \leq -f_0 \\ 0 & \text{其他} \end{cases} \end{aligned} \quad (10.3-20)$$

$p(t)$ 的振幅频谱如图 10.4 所示。

由于 $p(t)$ 只含有 $u(t)$ 上边带，故它即为大家熟知的单边带信号。如果它被 $u(t) \cos(2\pi f_0 t) + v(t) \sin(2\pi f_0 t)$ 代替，则它就只含有下边带，用这种方法，在信道上传播信号所需的频带宽度可以减半。

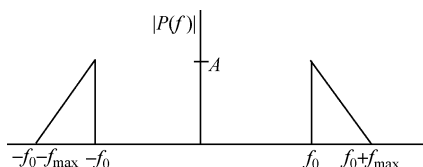


图 10.4 单边带信号 $p(t)$ 的振幅频谱



10.4 非单色波场的解析信号表示

用 $u(r, t)$ 代表一个非单色波场的实部。正如在 2.7 节中讨论的, $u(r, t)$ 的傅里叶表示为

$$u(r, t) = 2 \int_0^{+\infty} U(r, f) \cos(2\pi f t + \phi(f)) df \quad (10.4-1)$$

式中, $U(r, f)$ 和 $\phi(f)$ 分别是 $u(r, t)$ 关于 t 的振幅谱和相位谱。

$u(r, t)$ 相应的解析信号为

$$u_A(r, t) = u(r, t) + jv(r, t) \quad (10.4-2)$$

式中, $v(r, t)$ 是 $u(r, t)$ 的希尔伯特变换。

由例 10.1 可知, $\cos(2\pi f t + \phi(f))$ 的希尔伯特变换等于 $\sin(2\pi f t + \phi(f))$, 故 $v(r, t)$ 可表示为

$$v(r, t) = 2 \int_0^{+\infty} U(r, f) \sin(2\pi f t + \phi(f)) df \quad (10.4-3)$$

所以, 解析信号 $u_A(r, t)$ 为

$$u_A(r, t) = 2 \int_0^{+\infty} U(r, f) e^{j(2\pi f t + \phi(f))} df \quad (10.4-4)$$

注意, $U_c(r, f) = \frac{1}{2} U(r, f) e^{j\phi(f)}$ 是 $u(r, t)$ 关于 t 的傅里叶变换。方程式 (10.4-4) 表明解析信号是通过对原始信号只在正频率上积分得到的。 $u(r, t)$ 也可以表示为

$$u(r, t) = \frac{1}{2} [s(r, t) + s^*(r, t)] \quad (10.4-5)$$

在波的相关理论中, 解析信号通常指复波函数。

10.5 准单色波、相干和非相干波

当一个波场的时间频率范围 Δf 满足下列条件:

$$\frac{\Delta f}{f_c} \ll 1 \quad (10.5-1)$$

时, 称该波场为准单色波场。式中, f_c 为波场的平均时域频率。

在这种情况下, $u_A(r, t)$ 可以表示为

$$u_A(r, t) = U_A(r, t) e^{j2\pi f_c t} \quad (10.5-2)$$

式中, $U_A(r, t)$ 是解析信号的复包络。

单色波场是完全相干的, 波场相干的问题非常广泛, 如可参阅参考文献 [Born and Worf, 1969] 和 [Marathay]。相干波场的特征就是当相干波场的不同部分在同一点上相加时, 它能在一个平均时间基础上产生干涉加强和干涉减弱。

通常讨论相干波的空间相干性和时间相干性, 令波场的解析信号为 $u_A(r, t)$, 当 $u_A(r_1, t)$ 与 $u_A(r_2, t)$ 通过一个复常数因子在 r_1 和 r_2 各点上关联, 且与时间无关时, 则称这样的波场为空间相干的。在下文中, 空间相干波都简称为相干波。



如果 $u_A(r_1, t)$ 和 $u_A(r_2, t)$ 之间的关系仅仅取决于时间差 $t_2 - t_1$, 则称该波场为时间相干波。

10.6 一般成像系统中的衍射效应

在 9.4 节中, 讨论了用单透镜作为二维线性滤波器的成像。可以将已获得的结果推广到更加普通的成像系统, 也可以是空间相干或空间不相干照明的准单色光源的成像系统。

一个普通的成像系统由许多透镜组成, 这样的成像系统能够用系统的出瞳和入瞳来表征, 实际上它们就是系统有效孔径的像 [Goodman], 如图 10.5 所示。衍射效应可以通过出瞳函数以及出瞳处和像平面之间的距离表示。

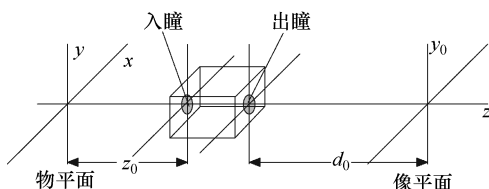


图 10.5 光学成像系统模型

当一发散的球面波入射到入瞳处后, 在出瞳处形成会聚的球面波, 则称这样的光学系统为衍射受限系统。在真实成像系统中, 这一性质最好地体现在物平面和像平面中的有限的局部区域内。像差是改变球面波性质的畸变, 将在 10.10 节讨论。

在带有许多透镜的成像系统中, 如果采用 $P(\cdot, \cdot)$ 表示系统的有限等效出瞳函数, 并采用系统的等效焦距以及出瞳处到像平面的距离, 且系统为衍射受限系统, 则方程式 (9.4-17) 和方程式 (9.4-18) 仍然是有效的 [Goodman]。

定义方程式 (9.4-17) 中的理想像为

$$U_g(x, y) = U\left(-\frac{x}{M}, -\frac{y}{M}\right) \quad (10.6-1)$$

式中, $U(x, y)$ 为原始像 (波场)。

方程式 (9.4-16) 可重新写为

$$U(x_0, y_0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x_0 - x, y_0 - y) U_g(x, y) dx dy \quad (10.6-2)$$

式中

$$h(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x_1, y_1) e^{-i\frac{2\pi}{\lambda d_0}(xx_0 + yy_0)} dx dy \quad (10.6-3)$$

10.7 准单色波场成像

在实际成像系统中, 准单色波场成像条件通常是满足的, 例如, 在普通照片中, 胶片对电磁波谱上的可见光很敏感。所以, 解析信号的复包络满足之前得到的成像方程, 故方程式 (10.6-2) 可写为

$$\begin{aligned} U_A(x_0, y_0, d_0, t) &= h(x_0, y_0) * U_c(x_0, y_0, t - \tau) \\ &= \iint_{-\infty}^{+\infty} h(x_0 - x, y_0 - y) U_c(x, y, t - \tau) dx dy \end{aligned} \quad (10.7-1)$$

式中, $h(x_0, y_0)$ 由方程式 (10.6-3) 用 λ_c 代替 λ 后给出; $U_c(x_0, y_0, t)$ 是几何光学理想像



$U\left(-\frac{x_0}{M}, -\frac{y_0}{M}, t\right)$ 相应的解析信号的复包络; τ 是对应传播到像平面的时间延迟。

振幅传递函数 $H(f_x, f_y)$ 是 $h(x, y)$ 的傅里叶变换

$$H(f_x, f_y) = P(\lambda d_0 f_x, \lambda d_0 f_y) \quad (10.7-2)$$

探测系统对波场的强度敏感, 它可以表示为

$$I(x_0, y_0, d_0) = \langle |U_A(x_0, y_0, d_0, t)|^2 \rangle \quad (10.7-3)$$

式中, $\langle \bullet \rangle$ 表示一个无限时间均值。

利用方程式 (10.7-1), $I(x_0, y_0, d_0)$ 可写为

$$I(x_0, y_0, d_0) = \iint_{-\infty}^{+\infty} h(x_0 - x, y_0 - y) \left[\iint_{-\infty}^{+\infty} h^*(x_0 - x', y_0 - y') I_c(x, y; x', y') dx' dy' \right] dx dy \quad (10.7-4)$$

其中

$$I_c(x, y; x', y') = \langle U_c(x, y, t - \tau_1) U_c^*(x', y', t - \tau_2) \rangle \quad (10.7-5)$$

式中, I_c 为互强度 (mutual intensity)。

在方程式 (10.7-5) 中, τ_1 和 τ_2 近似相等, 因为脉冲响应 h 被限制在图像点周围很小的区域内, 故可忽略不计。

10.7.1 干涉成像

对完全相干波场, 可写为

$$U_c(x, y, t) = U_c(x, y) U_c(0, 0, t) \quad (10.7-6)$$

式中, $U_c(x, y)$ 为与原点处波场有关的 $U_c(x, y, t)$ 的相量幅值。写为

$$\langle U_c(x, y, t) U_c(x, y, t) \rangle = K U_c(x, y) U_c^*(x', y') \quad (10.7-7)$$

其中, 常数 K 为

$$K = \langle |U_c(0, 0, t)|^2 \rangle \quad (10.7-8)$$

忽略 K , 强度可写为

$$I(x_0, y_0, d_0) = |U_1(x_0, y_0)|^2 \quad (10.7-9)$$

式中

$$U_1(x_0, y_0) = h(x_0, y_0) * U_c(x_0, y_0) = \iint_{-\infty}^{+\infty} h(x_0 - x, y_0 - y) U_c(x, y) dx dy \quad (10.7-10)$$

因此, 相干成像系统与一个和原点处波场相关的解析信号的复振幅成线性关系。

例 10.7 衍射受限相干成像系统的圆形有效的半径为 R , 假设像到系统的距离为 d_0 , 求该系统的截止频率。

解: 该情况下的函数为

$$P(x, y) = \text{circ}\left(\frac{\sqrt{x^2 + y^2}}{R}\right)$$

振幅传递函数为

$$H(f_x, f_y) = P(\lambda d_0 f_x, \lambda d_0 f_y) = \text{circ}\left(\frac{\sqrt{f_x^2 + f_y^2}}{R/\lambda d_0}\right)$$



因此, 截止频率为

$$f_c = \frac{R}{\lambda d_0}$$

10.7.2 非相干成像

当波场 $U_c(x, y, t)$ 为完全非相干时, 该波场为空间无关的 (此时相当于白噪声)。所以, 时间平均可写为

$$\langle U_c(x, y, t) U_c(x', y', t) \rangle = K I_c(x, y) \delta(x - x', y - y') \quad (10.7-11)$$

式中, $I_c(x, y) = \langle |U_c(x, y, t)|^2 \rangle$ 。

忽略 K , 方程式 (10.7-4) 变为

$$I(x_0, y_0, d_0) = \iint_{-\infty}^{+\infty} |h(x_0 - x, y_0 - y)|^2 I_c(x, y) dx dy \quad (10.7-12)$$

因此, 当波场为非相干时, 成像系统与强度成线性关系。通过脉冲响应函数 $|h(x, y)|^2$ 滤波得到理想成像强度。

10.8 衍射受限成像系统的频率响应

在前面章节中, 若相干和非相干成像系统的输入是理想像的振幅或强度, 输出为实像, 且该输出为输入和系统脉冲响应函数的卷积, 则认为该成像系统为线性系统, 如图 10.6 所示。在这节中, 我们进一步详细地研究相干和非相干成像系统的特性。

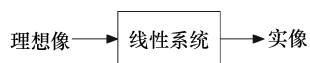


图 10.6 相干和非相干线性成像系统示意图

10.8.1 相干成像系统

相干成像系统与振幅成线性映射。令输出图像 U_1 和理想图像 U_c 的二维傅里叶变换分别为 V_1 和 V_c 。通过卷积定理有

$$V_1(f_x, f_y) = H(f_x, f_y) V_c(f_x, f_y) \quad (10.8-1)$$

式中, $H(f_x, f_y)$ 是脉冲响应的二维傅里叶变换。

$$H(f_x, f_y) = \iint_{-\infty}^{+\infty} h(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy = P(\lambda d_0 f_x, \lambda d_0 f_y) \quad (10.8-2)$$

式中, $H(f_x, f_y)$ 即为相干传递函数。

可以看出, 相干成像系统等价于一个理想的低通滤波器, 在瞳孔函数“1”区内的所有频率都能通过, 而截掉该区外的所有频率。

10.8.2 非相干成像系统

非相干成像系统与强度是线性关系, 像的视觉效果很大程度上取决于图像细节信息与背景之间相对强度的对比。输出像和输入理想像可通过相对于像总能量的归一化来反映这一特性。

$$I'(x_0, y_0, d_0) = \frac{I(x_0, y_0, d_0)}{\iint_{-\infty}^{+\infty} I(x_0, y_0, d_0) dx_0 dy_0} \quad (10.8-3)$$



$$I'_G(x, y) = \frac{I_G(x, y)}{\iint_{-\infty}^{+\infty} I_G(x, y) dx dy}$$

令 $I'(x_0, y_0, d_0)$ 和 $I'_G(x, y)$ 的二维傅里叶变换分别为 $J(f_x, f_y)$ 和 $J_G(f_x, f_y)$ 。通过卷积定理, 方程式 (10.7-12) 可写为

$$J(f_x, f_y) = H_I(f_x, f_y) J_G(f_x, f_y) \quad (10.8-4)$$

其中

$$H_I(f_x, f_y) = \frac{\iint_{-\infty}^{+\infty} |h(x, y)|^2 e^{-j2\pi(f_x x + f_y y)} dx dy}{\iint_{-\infty}^{+\infty} |h(x, y)|^2 dx dy} \quad (10.8-5)$$

式中, $H_I(f_x, f_y)$ 为光学传递函数 (OTF)。

定义调制传递函数 (MTF) 为 $|H_I(f_x, f_y)|$ 。

可以看到, $H_I(f_x, f_y)$ 是 $|h(x, y)|^2$ 傅里叶变换的归一化, 即为一个非负函数。通过在 2.5 节中讨论的傅里叶变换的性质 15 和帕塞瓦尔定理, 可知 $H_I(f_x, f_y)$ 是 $H(f_x, f_y)$ 自相关的归一化。

$$H_I(f_x, f_y) = \frac{\iint_{-\infty}^{+\infty} H(f'_x, f'_y) H^*(f_x + f'_x, f_y + f'_y) df'_x df'_y}{\iint_{-\infty}^{+\infty} |H(f_x, f_y)|^2 df_x df_y} \quad (10.8-6)$$

OTF 最重要的性质如下:

A. $H_I(0, 0) = 1$

B. $H_I(-f_x, -f_y) = H_I(f_x, f_y)$

C. $|H_I(f_x, f_y)| \leq H_I(0, 0)$

最后一条性质是施瓦茨 (Schwarz) 不等式的结果, 对于任意两个复值函数 f 和 g

$$\left| \iint f g dA \right|^2 \leq \iint |f|^2 dA \iint |g|^2 dA \quad (10.8-7)$$

当且仅当 $g = K f^*$ 时上式取等号, 其中 K 为复常数。

令 f 和 g 分别等于 $H(f'_x, f'_y)$ 和 $H^*(f_x + f'_x, f_y + f'_y)$, 利用方程式 (10.8-7) 得出上面性质 C。

方程式 (10.8-2) 给出了相干传递函数 $H(f_x, f_y)$, 在方程式 (10.8-7) 中运用这个结果可得

$$H_I(f_x, f_y) = \frac{\iint_{-\infty}^{+\infty} P(\lambda d_0 f'_x, \lambda d_0 f'_y) P(\lambda d_0 (f_x + f'_x), \lambda d_0 (f_y + f'_y)) df'_x df'_y}{\iint_{-\infty}^{+\infty} |P(\lambda d_0 f_x, \lambda d_0 f_y)|^2 df_x df_y} \quad (10.8-8)$$

利用变量替换, 方程式 (10.8-8) 可写为

$$H_I(f_x, f_y) = \frac{\iint_{-\infty}^{+\infty} P\left(f'_x + \frac{\lambda d_0 f_x}{2}, f'_y + \frac{\lambda d_0 f_y}{2}\right) P\left(f'_x - \frac{\lambda d_0 f_x}{2}, f'_y - \frac{\lambda d_0 f_y}{2}\right) df'_x df'_y}{\iint_{-\infty}^{+\infty} P(\lambda d_0 f'_x, \lambda d_0 f'_y) df'_x df'_y} \quad (10.8-9)$$

上面式子分子中的两个瞳孔函数通过 $(\lambda d_0 |f_x|, \lambda d_0 |f_y|)$ 相互代替, 其积分等于两个瞳孔函数重叠部分的面积, 所以, $H_I(f_x, f_y)$ 可以写为



$$H_I(f_x, f_y) = \frac{\text{重叠部分面积}}{\text{总面积}} \quad (10.8-10)$$

式 (10.8-10) 通过变尺度瞳孔函数计算面积, OTF 一直为非负实函数。

注意, 非相干脉冲响应函数 $|h(x, y)|^2$ 与稳态 2-D 随机场的功率谱相似, 同理可得, $H_I(f_x, f_y)$ 与稳态 2-D 随机场的自相关函数相似 [Besag, 1974]。

例 10.8 (a) 确定一个出瞳宽度是 $2W$ 的正方形衍射受限光学系统的 OTF (光学传递函数); (b) 确定该系统的截止频率 f_c 。

解: (a) 瞳孔函数的面积为 $4W^2$, 重叠部分的面积如图 10.7 所示。

从图 10.7 中计算重叠部分的面积如下:

$$A(f_x, f_y) = \begin{cases} (2W - \lambda d_0 |f_x|)(2W - \lambda d_0 |f_y|) & |f_x| \leq 2W/\lambda d_0 \\ & |f_y| \leq 2W/\lambda d_0 \\ 0 & \text{其他} \end{cases}$$

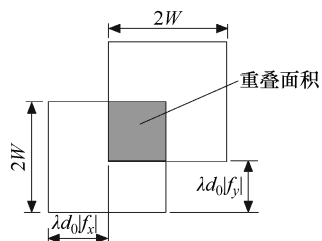


图 10.7 正方形光孔的 OTF 计算中的重叠部分面积

当 $A(f_x, f_y)$ 通过 $4W^2$ 归一化, 则 OTF 如下:

$$H_I(f_x, f_y) = \text{tri}\left(\frac{f_x}{2f'_c}\right) \text{tri}\left(\frac{f_y}{2f'_c}\right) \quad (10.8-11)$$

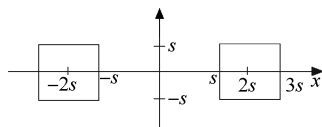
式中, $\text{tri}(\cdot)$ 是三角函数; f'_c 为相干照明的截止频率, 其等于 $W/\lambda d_0$ 。

(b) 很明显, 截止频率 f_c 为

$$f_c = 2f'_c$$

例 10.9 出瞳函数由两个正方形组成, 如图 10.8 所示。求:

- 相干传递函数;
- 相干截止频率;
- 振幅脉冲响应;
- 光学传递函数。



解: (a) 相干传递函数与变尺度瞳孔函数相同。瞳孔函数可 图 10.8 例 10.9 中的孔径函数 写为

$$P(x, y) = \text{rect}\left(\frac{x-2s}{2s}, \frac{y}{2s}\right) + \text{rect}\left(\frac{x+2s}{2s}, \frac{y}{2s}\right)$$

$H(f_x, f_y)$ 由下式给出

$$\begin{aligned} H(f_x, f_y) &= P(\lambda d_0 f_x, \lambda d_0 f_y) \\ &= \text{rect}\left(\frac{\lambda d_0 f_x + 2s}{2s}, \frac{\lambda d_0 f_y}{2s}\right) + \text{rect}\left(\frac{\lambda d_0 f_x - 2s}{2s}, \frac{\lambda d_0 f_y}{2s}\right) \end{aligned}$$

(b) 两个方向的截止频率分量为

$$f_{xc} = \frac{3s}{\lambda d_0}, f_{yc} = \frac{s}{\lambda d_0}$$

(c) 振幅脉冲响应 $h(x, y)$ 是 $H(f_x, f_y)$ 的逆傅里叶变换。令 $\alpha = \lambda d_0 s$, 则 $h(x, y)$ 为



$$\begin{aligned}
 h(x, y) &= \iint_{-\infty}^{+\infty} H(f_x, f_y) e^{i2\pi(f_x x + f_y y)} df_x df_y \\
 &= \left[2 \int_0^{3s} \cos(2\pi f_x x) df_x \right] \left[2 \int_0^s \cos(2\pi f_y y) df_y \right] \\
 &= \frac{1}{\pi^2} \frac{\sin(6\pi s x)}{x} \frac{\sin(6\pi s y)}{y}
 \end{aligned}$$

(d) $H(f_x, f_y)$ 下的总面积 A 等于 $8s^2$, OTF 如下:

$$H_l(f_x, f_y) = \frac{\iint_{-\infty}^{+\infty} H(f'_x, f'_y) H^*(f'_x - f_x, f'_y - f_y) df'_x df'_y}{A}$$

上式积分为非简单计算, 最好通过计算机求解。

10.9 光学传递函数的计算机计算

计算离散 OTF 的最简单方法就是运用快速傅里叶变换 (FFT), 这要求 $h(x, y)$ 、 $H(f_x, f_y)$ 和 $H_l(f_x, f_y)$ 都是离散的, 离散坐标可写为

$$x = \Delta x n_1 \quad (10.9-1)$$

$$y = \Delta y n_2 \quad (10.9-2)$$

$$f_x = \Delta f_x k_1 \quad (10.9-3)$$

$$f_y = \Delta f_y k_2 \quad (10.9-4)$$

则 $h(\Delta x n_1, \Delta y n_2)$ 、 $H(\Delta f_x k_1, \Delta f_y k_2)$ 和 $H_l(\Delta f_x k_1, \Delta f_y k_2)$ 分别写为 $h(n_1, n_2)$ 、 $H(k_1, k_2)$ 和 $H_l(k_1, k_2)$ 。假设涉及的矩阵大小为 N_1 、 N_2 , 为了运用 FFT, 必须满足下列条件:

$$\Delta x \Delta f_x = \frac{1}{N_1} \quad (10.9-5)$$

$$\Delta y \Delta f_y = \frac{1}{N_2} \quad (10.9-6)$$

下式给出近似 $h(n_1, n_2)$

$$h(n_1, n_2) = \frac{1}{K} h'(n_1, n_2) \quad (10.9-7)$$

其中

$$h'(n_1, n_2) = \sum_{k_1 = -\frac{N_1}{2}}^{\frac{N_1}{2}-1} \sum_{k_2 = -\frac{N_2}{2}}^{\frac{N_2}{2}-1} H(k_1, k_2) e^{i2\pi \left(\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2} \right)} \quad (10.9-8)$$

和

$$K = \Delta x \Delta y N_1 N_2 \quad (10.9-9)$$

$H(k_1, k_2)$ 等于 $P(-\lambda d_0 \Delta f_x k_1, -\lambda d_0 \Delta f_y k_2)$, 为了充分描述瞳孔函数, 应当选择适当的 N_1 和 N_2 , 例如, 一定要完全覆盖瞳孔函数的非零部分。为了最小化 FFT 带来的周期性效应, N_1 和 N_2 也应该至少是描述瞳孔函数非零部分的最小值的两倍大。

只要选择了适当的 N_1 和 N_2 , $H(k_1, k_2)$ 就如同在 4.4 节中讨论的一样, k_1 和 k_2 分别满足 $0 \leq k_1 \leq N_1$, $0 \leq k_2 \leq N_2$ 。



$H_I(k_1, k_2)$ 近似为

$$H_I(k_1, k_2) = \frac{\frac{\Delta x \Delta y}{K^2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} |h'(n_1, n_2)|^2 e^{-j2\pi(\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2})}}{\frac{1}{K} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |H(k_1, k_2)|^2} \quad (10.9-10)$$

或者

$$H_I(k_1, k_2) = \frac{\frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} |h'(n_1, n_2)|^2 e^{-j2\pi(\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2})}}{\sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |H(k_1, k_2)|^2} \quad (10.9-11)$$

式 (10.9-11) 分子是 $|h'(n_1, n_2)|^2$ 的 2-D 离散傅里叶变换 (DFT)，由于内在的周期性，在计算时实际上已把 $H_I(k_1, k_2)$ 变换为正频率。但为了包括负频率，最后还应该把它再次变换回来。

总之， $H(k_1, k_2)$ 是通过瞳孔函数来确定的。根据方程式 (10.9-8)，对 $H(k_1, k_2)$ 作逆 DFT 可得到 $h'(n_1, n_2)$ 。用 $|H(k_1, k_2)|^2$ 下的面积对 $|h'(n_1, n_2)|^2$ 归一化，然后对其进行 DFT 可得到 $H_I(k_1, k_2)$ 。

要注意，为了运用方程式 (10.9-8) 计算 $h'(n_1, n_2)$ 时不出现混叠， Δf_x 和 Δf_y 应该要选择得足够小，只要选择了 Δf_x 和 Δf_y ，就可根据上面讨论的瞳孔函数来确定 N_1 和 N_2 。

实际考虑

在研究 OTF 和 MTF 时，经常取 λd_0 的值为 1，并且，为了将 $H(f_x, f_y)$ 简写为 $P(f_x, f_y)$ ，经常省掉了瞳孔函数中的负号，也省略了基于瞳孔函数面积的归一化。

另一种得出 OTF 的方法就是对瞳孔函数和它本身进行自相关运算。

10.10 像差

衍射受限系统意味着出瞳处我们所关心的波场是完美的，唯一的缺点就是有限的孔径尺寸。该波场是典型的球面波。像差是出瞳处的实际波前和它的理想形状之间的偏差，这些都是典型的相位误差。

为了在出瞳瞳孔函数中包含像差，出瞳瞳孔函数可变为

$$P_A(x, y) = P(x, y) e^{jk\phi_A(x, y)} \quad (10.10-1)$$

式中， $P(x, y)$ 是无像差的出瞳瞳孔函数， $\phi_A(x, y)$ 是由于像差而产生的相位误差。

当用 $P_A(x, y)$ 替代 $P(x, y)$ 后，前几节相干和非相干成像理论仍然适用。例如，振幅传递函数变为

$$\begin{aligned} H(f_x, f_y) &= P_A(-\lambda d_0 f_x, -\lambda d_0 f_y) \\ &= P(-\lambda d_0 f_x, -\lambda d_0 f_y) e^{jk\phi_A(-\lambda d_0 f_x, -\lambda d_0 f_y)} \end{aligned} \quad (10.10-2)$$

类似地，光学传递函数可简写为

$$H_I(f_x, f_y) = \frac{A(f_x, f_y)}{\iint_{-\infty}^{+\infty} |P(\lambda d_0 f_x, \lambda d_0 f_y)|^2 df_x df_y} \quad (10.10-3)$$



其中

$$A(f_x, f_y) = \iint_{-\infty}^{+\infty} P_A\left(f'_x + \frac{\lambda d_0 f_x}{2}, f'_y + \frac{\lambda d_0 f_y}{2}\right) P_A^*\left(f'_x - \frac{\lambda d_0 f_x}{2}, f'_y - \frac{\lambda d_0 f_y}{2}\right) df'_x df'_y \quad (10.10-4)$$

我们注意到, 利用两个相对平移了的瞳孔函数的重叠面积可计算 $A(f_x, f_y)$, 在这个区域中一般的瞳孔函数都等于 1, 通过重叠区域上的积分, 则 $A(f_x, f_y)$ 可写为

$$A(f_x, f_y) = \iint_{\text{重叠区}} P_1 P_2 df'_x df'_y \quad (10.10-5)$$

其中

$$P_1 = e^{jk\phi_A\left(f'_x + \frac{\lambda d_0 f_x}{2}, f'_y + \frac{\lambda d_0 f_y}{2}\right)} \quad (10.10-6)$$

$$P_2 = e^{-jk\phi_A\left(f'_x - \frac{\lambda d_0 f_x}{2}, f'_y - \frac{\lambda d_0 f_y}{2}\right)} \quad (10.10-7)$$

例 10.10 证明像差不会增大调制传递函数 MTF。

解: MTF 是 OTF 的系数 (模), 根据施瓦茨不等式有

$$|A(f_x, f_y)|^2 \leq \iint_{\text{重叠区}} |P_1|^2 df'_x df'_y \iint_{\text{重叠区}} |P_2|^2 df'_x df'_y \quad (10.10-8)$$

注意到, 在重叠区域有

$$|P_1|^2 = |P_2|^2 = 1 \quad (10.10-9)$$

因此, $|A(f_x, f_y)|^2 \leq \text{重叠面积}$, 且有

$$|H_I(f_x, f_y)|^2 \leq |H'_I(f_x, f_y)|^2 \quad (10.10-10)$$

式中, $H'_I(f_x, f_y)$ 是无像差的光学传递函数。

相位函数 $\phi_A(x, y)$ 常写为极坐标的形式 $\phi_A(r, \theta)$ 。赛德尔像差 (Seidel aberration) 是 $\phi_A(r, \theta)$ 关于 r 的多项式表示, 例如

$$\phi_A(r, \theta) = a_{40}r^4 + a_{31}r^3 \cos\theta + a_{20}r^2 + a_{22}r^2 \cos^2\theta + a_{11}r \cos\theta \quad (10.10-11)$$

在这个函数中可以增加高阶的项。方程式 (10.10-11) 右边各项的含义:

$a_{40}r^4$: 球差

$a_{31}r^3 \cos\theta$: 彗差

$a_{20}r^2$: 像散

$a_{22}r^2 \cos^2\theta$: 场曲

$a_{11}r \cos\theta$: 畸变

10.10.1 泽尔尼克 (Zernike) 多项式

当光学系统的出瞳为一个圆时, 在光学系统中表现出的像差可以用泽尔尼克多项式表示, 它们在单位圆内是规范正交的 [Kim and Shannon, 1987]。在处理时, 相位函数 $\phi_A(x, y)$ 是用泽尔尼克多项式 $z_k(\rho, \theta)$ 的展开形式来表示的, 其中, ρ 为单位圆内的径向坐标; θ 为极角。



表 10.1 给出了 $1 \leq k \leq 37$ 的泽尔尼克多项式。注意，每个多项式的形式都为

$$z_k(\rho, \theta) = R_n^m(\rho) \cos m\theta \tag{10.10-12}$$

式中， n 和 m 都为非负整数； $R_n^m(\rho)$ 是 n 次多项式，且 ρ 的幂不会小于 m 。

表 10.1 泽尔尼克多项式

序 号	多 项 式
1	1
2	$2\rho \cos(\theta)$
3	$2\rho \sin(\theta)$
	$\sqrt{3}(2\rho^2 - 1)$
4	$\sqrt{3}(2\rho^2 - 1)$
5	$\sqrt{6}\rho^2 \cos(2\theta)$
6	$\sqrt{6}\rho^2 \sin(2\theta)$
7	$\sqrt{8}(3\rho^2 - 2)\rho \cos(\theta)$
8	$\sqrt{8}(3\rho^2 - 2)\rho \sin(\theta)$
9	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$
10	$\sqrt{8}\rho^3 \cos(3\theta)$
11	$\sqrt{8}\rho^3 \sin(3\theta)$
12	$\sqrt{10}(4\rho^2 - 3)\rho^2 \cos(2\theta)$
13	$\sqrt{10}(4\rho^2 - 3)\rho^2 \sin(2\theta)$
14	$\sqrt{12}(10\rho^4 - 12\rho^2 + 3)\rho \cos(\theta)$
15	$\sqrt{12}(10\rho^4 - 12\rho^2 + 3)\rho \sin(\theta)$
16	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$
17	$\sqrt{10}\rho^4 \cos(4\theta)$
18	$\sqrt{10}\rho^4 \sin(4\theta)$
19	$\sqrt{12}(5\rho^2 - 4)\rho^3 \cos(3\theta)$
20	$\sqrt{12}(5\rho^2 - 4)\rho^3 \sin(3\theta)$
21	$\sqrt{14}(15\rho^4 - 20\rho^2 + 6)\rho^2 \cos(2\theta)$
22	$\sqrt{14}(15\rho^4 - 20\rho^2 + 6)\rho^2 \sin(2\theta)$
23	$4(35\rho^6 - 60\rho^4 + 30\rho^2 - 4)\rho \cos(\theta)$
24	$4(35\rho^6 - 60\rho^4 + 30\rho^2 - 4)\rho \sin(\theta)$
25	$3(70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1)$
26	$\sqrt{12}\rho^5 \cos(5\theta)$
27	$\sqrt{12}\rho^5 \sin(5\theta)$
28	$\sqrt{14}(6\rho^2 - 5)\rho^4 \cos(4\theta)$
29	$\sqrt{14}(6\rho^2 - 5)\rho^4 \sin(4\theta)$
30	$4(21\rho^4 - 30\rho^2 + 10)\rho^3 \cos(3\theta)$
31	$4(21\rho^4 - 30\rho^2 + 10)\rho^3 \sin(3\theta)$
32	$\sqrt{18}(56\rho^6 - 105\rho^4 + 60\rho^2 - 10)\rho^2 \cos(2\theta)$
33	$\sqrt{18}(56\rho^6 - 105\rho^4 + 60\rho^2 - 10)\rho^2 \sin(2\theta)$
34	$\sqrt{20}(126\rho^8 - 280\rho^6 + 210\rho^4 - 60\rho^2 + 5)\rho \cos(\theta)$
35	$\sqrt{20}(126\rho^8 - 280\rho^6 + 210\rho^4 - 60\rho^2 + 5)\rho \sin(\theta)$
36	$\sqrt{11}(252\rho^{10} - 630\rho^8 + 560\rho^6 - 210\rho^4 + 30\rho^2 - 1)$
37	$\sqrt{13}(924\rho^{12} - 2772\rho^{10} + 3150\rho^8 - 1680\rho^6 + 420\rho^4 - 42\rho^2 + 1)$

另外，当 m 为奇数时， $R_n^m(\rho)$ 为奇函数，当 m 为偶数时， $R_n^m(\rho)$ 也为偶函数。 $\phi_A(x, y) = \phi_A(\rho, \theta)$ 的表示可写为



$$\phi_A(\rho, \theta) = A_{00} + \frac{1}{\sqrt{2}} \sum_{n=2}^{\infty} A_{n0} R_n^0(\rho) + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{nm} R_n^m(\rho) \cos m\theta \quad (10.10-13)$$

可通过有限值 n 和 m 的最小二乘法确定系数 A_{nm} , 相应地, $\phi_A(\rho, \theta)$ 也可以写为

$$\phi_A(\rho, \theta) = \sum_{k=1}^K w_k z_k(\rho, \theta) \quad (10.10-14)$$

式中, K 是一个整数, 如 37。

系数 w_k 可通过最小二乘法求得。由于每个相邻的泽尔尼克项是关于它前面的每一项的归一化, 所以, 每一项都独立地影响方均像差, 这意味着像差的方均根误差 $\overline{\phi_A}$ 可写为

$$\overline{\phi_A} = \left[\sum_{k=K+1}^{\infty} w_k^2 \right]^{\frac{1}{2}} \quad (10.10-15)$$

注意, 当出瞳为圆时, 像差的泽尔尼克表示是有效的, 否则, 泽尔尼克多项式不是正交的。在一些情况下, 如由于大气对相位的扰乱, 有限项的泽尔尼克多项式不容易给出令人满意的像差表示。

第 11 章 基于波调制的光学器件

11.1 本章引论

应用波调制算法，光波的记录和生成可以通过许多技术实现。最早得到广泛应用的是照相胶片。最近，空间光调制器能够应用光电信号实时合成或控制波前。

另一个方法是应用固体和类似的技术来制作衍射光学元件，这些元件是通过衍射而非折射来控制光。还有令人振奋的发展是在一个设备中同时结合了折射和衍射光学元件，从而获得许多新的性质，例如，可以减小误差。

本章包含 7 个小节。照相胶片和照相干板是最为人熟知的记录器件，11.2 节和 11.3 节介绍了这些设备的性质，特别是相干记录方面。11.4 节介绍了这些媒质的调制传递函数的物理机制。应用照相胶片和照相干板进行相位调制时，漂白是最重要的技术手段，11.5 节介绍了漂白的相关内容。

衍射光学器件的应用将在第 15 和 16 章详细探讨，这往往需要应用到其他一些技术，特别是超大规模集成电路和集成光学技术，11.6 节介绍了这些设备的基础知识。电子束曝光和反应离子刻蚀是典型的应用技术，这将在 11.7 节中介绍。

11.2 照相胶片和干板

照相底片是价格非常低廉的光学器件，它可以探测光辐射，保存图像和控制空间光 [Goodman]。它由一种乳胶组成，这种乳胶含有感光的卤化银晶粒。这种乳胶夹在保护层和基底层之间，照相胶片由醋酸酯或聚酯薄膜组成，照相干板的基底层由玻璃组成。黑白照相胶片中，仅有一层银盐晶粒；而彩色照相底片中至少有三层这样的银盐晶粒。添加染料后的银盐晶粒能够使晶体对不同的色彩敏感。

光调制的过程如下：

- (1) 如果卤化银晶粒吸收了一个光子，那么在这个晶粒中将生成一个电子空穴对。
- (2) 生成的电子存在于导带中，在晶体中移动时，可能会在晶体位错中被俘获。
- (3) 一个俘获的电子吸引一个自由运动的银离子，这将产生一个寿命为数秒的金属银原子。
- (4) 在相同的位置可能生成若干个另外的银原子。这个反应过程至少需要 4 个由银原子构成的所谓的银斑点。

这个过程是在化学药液中进行的。显影剂作用于银斑点，使吸收了足够光能量的区域的晶体变成金属银。而那些没有变成金属银的区域也最终必须通过定影过程将其去除。定影则是将胶片浸入第二种化学药液（定影剂）中。

下面讨论一些与照相胶片及其定义相关的专业术语。



曝光量 $E(x, y)$

$$E(x, y) = I_e(x, y) T \quad (11.2-1)$$

式中, $I_e(x, y)$ 表示曝光时入射到照相底片上的光的强度; T 表示曝光时间; $E(x, y)$ 的单位是 mJ/cm^2 。

光强透过率 $\alpha(x, y)$

$$\alpha(x, y) = \text{local average} \left(\frac{I(x, y)}{I_e(x, y)} \right) \quad (11.2-2)$$

式中, $I(x, y)$ 表示显影定影后的透射强度; $I_e(x, y)$ 表示入射强度。

感光密度

$$D = \log_{10} \frac{1}{\alpha} \quad (11.2-3)$$

单位面积银的数量近似与感光密度 D 成比例, D 即为以 10 为底的显影定影后胶片强度透过率的倒数的对数。因此, D 定义为胶片的复振幅透过率。也可用下式描述:

$$\alpha = 10^{-D} \quad (11.2-4)$$

赫特尔-德雷菲尔德 (H&D) 曲线

H&D 曲线描述照片感光密度 D 与曝光量对数 $\log E$ 的关系。图 11.1 为负片的典型 H&D 曲线。曝光量极小或很大时, 感光密度基本为定值。中间部分为线性区域, 该区域通常被应用于照相。

伽马 (Gamma) 值

感光乳胶的伽马 (γ) 值是 H&D 曲线在线性区域的斜率。伽马值较大 (通常为 2 或 3) 的胶片, 称为高反差胶片, 伽马值较小 (1 或更小) 的则称为低反差胶片。负片的 γ 值为正, 而正片的 γ 值为负。

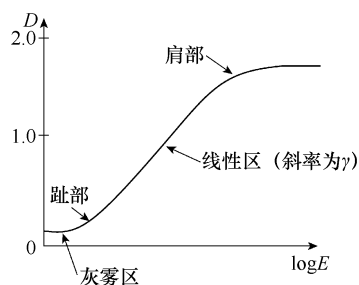


图 11.1 感光乳胶的赫特尔-德雷菲尔德 (H&D) 曲线

11.3 胶片的透光率

在非相干和相干光学系统中, 胶片通常被用来调节透光率。如何达到这种效果, 取决于相干光或非相干光的使用。下面将分别讨论这两种情况。

相干光

假设胶片工作在 H&D 曲线的线性区域, 感光密度 D 可以表示为

$$D = \gamma \log_{10}(E) - D_0 = \log_{10} \left(\frac{1}{\alpha} \right) \quad (11.3-1)$$

或者

$$\log_{10}(\alpha) = -\gamma \log_{10}(E) + D_0 \quad (11.3-2)$$

由于 E 等于 $I_e T$, 可以得到

$$\alpha = K I^{-\gamma} \quad (11.3-3)$$

式中, $K = 10^{D_0} T^{-\gamma}$, 是正常数。

式 (11.3-3) 表示透过率与曝光强度之间的关系是非线性的, 而且透过率会随着 γ 的增



加而减小。

例 11.1 证明两个负片串联可以使透过率 α 随着入射光强的增强而变大。

解：第一张负片的发光强度透过率为

$$\alpha_1 = K_1 I_1^{-\gamma_1}$$

式中, I_1 为照射光的光强。

假设第一张负片紧贴在另一张未曝光的胶片上, 用发光强度为 I_2 的非相干光照射。第二张胶片上的曝光强度为 $\alpha_1 I_2$, 它的光强透过率为

$$\alpha_2 = K_2 (\alpha_1 I_2)^{-\gamma_2} = K I_1^{\gamma_1 \gamma_2 - \gamma_2}$$

式中, $K = K_1^{-\gamma_2} K_2$ 。

可见, 总的光强透过率 α_2 与曝光光强成正比。

在相干光学系统中, 相位调制也同样重要。相位调制是通过照相底片或照相干板的厚度变化来实现的。照相胶片的振幅透过率可表示为

$$\alpha_c(x, y) = \sqrt{\alpha(x, y)} e^{j\phi(x, y)} \quad (11.3-4)$$

式中, $\phi(x, y)$ 是由于照相底片厚度的变化而引起的相位起伏。

相位项 $e^{j\phi(x, y)}$ 常常是不希望存在的, 可以利用液门来将其消除, 如图 11.2 所示。

液门包含两片玻璃, 而且外部为光学平面。其中的透明物质是填充于两片玻璃之间的折射率匹配液 (通常为油)。经过适当的调节, 相位调制将被去除, 振幅透过率可表示为

$$\alpha_c(x, y) = \sqrt{\alpha(x, y)} = K I_e^{-\frac{\gamma}{2}} \quad (11.3-5)$$

若 $I_e = A_e^2$, A_e 即为曝光时相干光源的振幅。式 (11.3-5) 可表示为

$$\alpha_c(x, y) = K A_e^{-\gamma} \quad (11.3-6)$$

基于相干光, 有时更容易得到振幅透过率 α_c 与曝光量 E 之间的关系曲线, 如图 11.3 所示。

在一般情况下, 常常在曝光量 E_0 附近使用照相底片。在线性区域内, 振幅透过率 α_c 相对于曝光量 E 之间的关系可表示为

$$\alpha_c = \alpha_{c0} + \beta(E - E_0) = \alpha_{c0} + \beta'(\Delta A_c^2) \quad (11.3-7)$$

式中, β 为曲线的斜率; ΔA_c^2 为曝光时振幅二次方的变化。

β 和 β' 对于负 (正) 片分别为负 (正)。

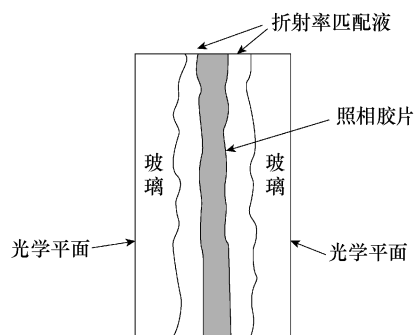


图 11.2 消除厚度变化的液门

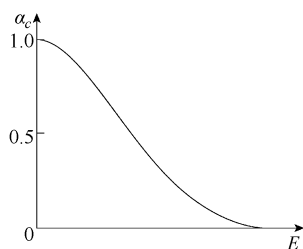


图 11.3 振幅透过率与曝光量关系曲线

11.4 调制传递函数

H&D 曲线表示了一个连续的非线性函数, 入射到乳胶上每个点的光都按照该非线性曲线进行变换。另外, 还有其他的一些因素会限制感光乳胶的空间频率响应。



两个主要的因素分别是曝光时的光散射和显影时的化学扩散。这两个因素都可以看作是线性模型。但是，光散射与曝光量是线性关系，化学扩散与密度是线性关系。因此，整个感光过程用图 11.4 所示模型表示，该模型称为凯利模型。为了简化该模型，忽略化学扩散造成的线性变化。图 11.5 表示简化后的模型。

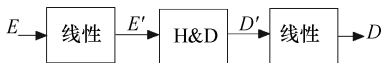


图 11.4 感光过程的完整的凯利模型

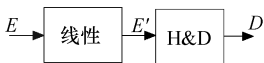


图 11.5 简化的凯利感光过程模型

为测量照相胶片的特性，采用一种正弦曝光模式，如下式所示：

$$E = E_0 + E_1 \cos(2\pi fx) \tag{11.4-1}$$

上述模式很容易采用全息技术把两束相干平面波在感光乳胶上进行相干曝光来实现。照相底片上的有效曝光可以用下式表示 [Goodman]：

$$E' = E_0 + M(f) \cos(2\pi fx) \tag{11.4-2}$$

式中， $M(f)$ 为调制传递函数 (MTF)。

调制传递函数有近似的圆对称性。因此， f 可以看作是径向频率。一个典型的 MTF 曲线如图 11.6 所示。不同的感光乳胶的截止频率是不同的，如粗糙的感光乳胶的频率只有 50 线对/mm，而精细的感光乳胶的频率可以达到 2000 线对/mm。

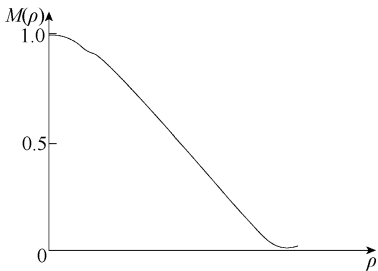


图 11.6 调制传递函数的例子

11.5 漂白

正常的显影过程会生成金属银，会吸收光。若希望调制光的相位，由此得到较大的衍射效率，则可以通过漂白来达到这个目的。

漂白过程替换了乳胶中的金属银，会使乳胶的厚度变化，或者使乳胶的折射率改变。鞣化漂白可以改变乳胶的厚度，而非鞣化漂白可以改变乳胶的折射率。

鞣化漂白导致照相底片表面起伏，如图 11.7 所示。这种表面起伏与光密度的空间频率成分相关，因而可作为带通滤波器。换言之，在很低或很高频率时不会产生表面起伏。对于 15μm 厚的乳胶，在频率为 10 线对/mm 时的最大厚度变化约为 1~2μm [Goodman]。

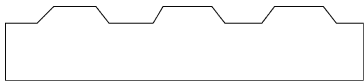


图 11.7 用硬化漂白生成的表面起伏的例子

非鞣化漂白可以导致乳胶内部折射率的改变。最终的透过率代表了这种通过折射率变化实现的相位的空间调制。

11.6 衍射光学、二元光学和数字光学

传统的光学元件，如透镜、平面镜和棱镜，都是基于光的折射或者反射原理。相比之下，衍射光学元件 (Diffractive Optical Elements, DOE)，例如，相位浮雕型器件，则基于光的衍射原理。衍射光学元件的重要优势如下：



(1) 一个衍射光学元件可有多种功能, 例如, 有多个焦点, 相当于在一个元件上有多个透镜。它也可以设计成在多波段均能使用的元件。

(2) 衍射光学元件更轻便, 相比于折射和反射光学元件占据更小的体积。

(3) 大量制造时, 将会更廉价。

由于衍射光学元件基于衍射原理, 具有高色散性, 也就是说, 其性质与波长相关。因此衍射光学元件更多地用于单色光系统中。然而, 其色散性也可以加以利用。例如, 某一波长的光会聚于一点, 而另一波长的光会聚于空间中的另一点。另外, 通过某种方式, 可将折射和衍射光学元件结合起来消除波长色散, 达到消除球差的目的。衍射光学的主要内容和衍射光学元件的设计将在第 15 和 16 章中详细介绍。

衍射光学、二元光学和数字光学通常指的是相同的调制技术和实现技术。对于不同应用的新设备, 光刻技术和微机械技术最为重要, 这些也是超大规模集成电路和集成光学以及其他纳米/微米应用中的关键技术。另一方面, 激光打印可以用于低分辨率的文本设计 [Asakura, Nagashima, 1976]。

虽然使用“二元”一词, 但是相位的量化可以用一系列的量化等级, 如 2、4 或 8 等。一般在相位量化时, 通过调节衍射光学元件的厚度来实现相应的量化等级。常用 2^N 对相位进行离散, 例如, 考虑如图 11.8 中的闪耀光栅, 其衍射效率为 100%, 这意味着如果能够保证连续的厚度调制并生成理想的相位轮廓, 那么全部光波将进入第一级衍射。

在二元光学中, 相位以 $N=2^M$ 等级来近似, 例如, 当 $M=2$ 时, 如图 11.8 所示, 相位离散为 4 个等级。制造此类元件的电子束刻蚀和反应离子刻蚀技术将于下部分介绍。

近年来, 另一种微机械技术被用于制造衍射光学元件, 即采用飞秒激光脉冲加工透明材料, 如硅玻璃 [Watanebe et al., 2002]。飞秒激光脉冲在聚焦区域的强度足够高, 可以使聚焦区域的折射率发生改变而其他区域不受影响。这种现象可以用来在玻璃上制作各种各样的衍射和光子元件。这里采用了飞秒激光, 可以比较容易地引导光束聚焦在一个特定区域。

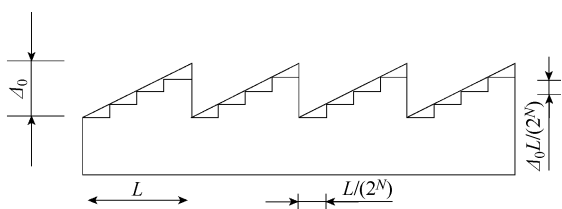


图 11.8 闪耀光栅相位等级约为 $N=2^M$

11.7 电子束刻蚀

刻蚀是指在基底上打印图纹的一类技术。最常见的应用是在半导体基底上打印集成电路。先进的刻蚀技术包括光刻蚀、电子束刻蚀、强紫外线刻蚀、X 射线刻蚀等 [Bowden et al., 1983]。

刻蚀包括两个基本的步骤。第一步是生成想要的图纹。第二步是把该图纹转移到基底上。衡量光刻蚀质量的首要方法是测量其分辨率。这里所谓分辨率是可以得到的最小条纹的尺寸。分辨率不仅由刻蚀方法决定, 也与阻碍图像重构的因素相关。

电子束刻蚀是定制衍射光学元件的高分辨率技术, 它采用小尺寸、高稳定性和长寿命的高强度、均匀电子源。电子束刻蚀技术的主要特点是: ①分辨率高; ②可以使用多种不同的材料和图纹; ③速度比光学刻蚀慢; ④比较贵。结合反应离子刻蚀 (Reactive Ion Etching,



RIE)，有可能获得更加精细分辨率和更高效率的制造衍射光学元件的方法。

例如，考虑一个二元相位衍射光学元件，其由二维小孔阵列构成，这些小孔径可以看作是一些采样点。在每一个采样点，相位可以变化 0° 或 180° 。在那些相位需要变化 180° 的点，基底被刻蚀到相应的深度。采用电子束将一个图纹曝光在电子束敏感的材料上，来指示需要刻蚀的位置，然后用反应离子刻蚀来加工到相应的厚度。也可以通过化学方法达到此目的，但是反应离子刻蚀是高精度技术，该方法可以推广应用到任意的 2 进量化等级。

一个典型的电子束刻蚀机可以生成斑点大小约 10nm 的电子束。因此，对于线宽小于 $0.1\mu\text{m}$ 或 100nm 的图纹都可以用其曝光。对于二元相位衍射光学元件，每一个方孔的尺度约为 $1\mu\text{m} \times 1\mu\text{m}$ ，采用这样尺度的特征，可以获得尺寸为 $1\text{mm} \times 1\text{mm}$ 的图纹，且保持了高的分辨率。更大的图形可以采用干涉测量控制方法拼接大量的小图得到。

干法刻蚀系统可以生成非常精细的表面起伏。例如， 180° 相位变化需要二元相位衍射光学元件产生约半个波长的距离，对于可见光，相当于半微米数量级的长度。反应离子刻蚀可以高精度地生成如此小尺寸的表面浮雕。

11.7.1 衍射光学元件的应用

用电子束刻蚀生成一个二元相位全息图的基本过程如图 11.9 所示。首先，在硅基片上旋转涂抹电子敏感的防蚀层。然后，用电子束刻蚀将需要的图形写在防蚀层上。如果防蚀剂是聚甲基丙烯酸甲酯一类的正型防蚀剂，则可以接着采用定影去除样本上被电子束“曝光”的防蚀剂。一旦“曝光”的防蚀剂被去除了，就可以用反应离子刻蚀使“曝光”的孔达到相应的深度。最后，用化学方法去除防蚀剂。重复该过程 M 次可以生成 2^M 级的二元衍射光学元件。每一次重复，都要高精度地排列晶片，使位置偏差最小，可采用干涉测量控制来达到此目的。电子束也必须精确控制使其直径在样本表面最小。电子束的散射也应该最小，因为这是提高其性能的首要因素。一旦获得了最优的对准和聚焦，就可用 CAD 程序来实现图形写入了。

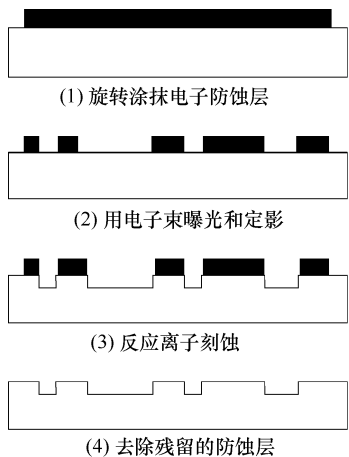


图 11.9 用电子束刻蚀和反应离子刻蚀系统生成衍射光学元件的实验过程

第 12 章 光在非均匀介质中传播

12.1 本章引论

在之前的章节里，折射率 $n(x, y, z)$ 被假设为与位置无关的定值。在许多应用中，如光在光纤、体衍射光栅、光折变介质等中传播时，折射率 $n(x, y, z)$ 实际上不是固定的，致使分析变得更加困难，因此常用数字方法来分析波在这些介质中的传播情况。一些所谓的“伪谱法”得到了应用，如 12.4 节和 12.5 节讨论的光束传播法（Beam Propagation Method, BPM），而其他方法经常是基于有限差分或有限元的方法。

本章讨论的光束传播法（BPM）是基于非均匀介质的傍轴波动方程，该方法对靠近 z 轴的传播是有效的，而且有一些其他约束。其主要优势在于，可以采用傅里叶变换来快速计算，并在大量的应用中足够精确。对于广角传播，存在其他的波束传播法，它们一般基于有限元法。这将在第 19 章阐述光的通信与网络的密集波分复用/解复用时进一步的讨论。

本章包含 5 部分。12.2 节讨论了非均匀介质条件下的亥姆霍兹方程。12.3 节将 5.4 节讨论的均匀介质条件下的傍轴波动方程推广到非均匀介质。在 12.4 节介绍了应用傅里叶变换来研究波传播的光束传播法，这种方法是一种杰出的数字方法。12.5 节以定向耦合器为例介绍了光束传播法的实际应用。它是一种包含有两个放置于一起的绝缘波导的光学器件，可以使进入其中一个波导的光波耦合进入另一个波导。这些仪器常用于光通信和光网络。光束传播法的结果相比于严格耦合模理论具有更加高的精度。耦合模理论不适用于复杂的设计，光束传播法是分析和综合复杂器件的非均匀介质波传播问题的常用方法。

12.2 非均匀介质的亥姆霍兹方程

在 4.2 节推导了折射率是定值的均匀介质的亥姆霍兹方程。在本节，将其扩展到非均匀介质条件。 $u(x, y, z, t)$ 表示在折射率为 $n(x, y, z)$ 的介质中，点 (x, y, z) 处 t 时刻的光场，光场的波长为 λ ，角频率为 ω 。根据麦克斯韦方程，其必然满足标量波动方程

$$\nabla^2 u(x, y, z, t) + n^2(x, y, z)k^2 u(x, y, z, t) = 0 \quad (12.2-1)$$

式中， $k = \omega/c = 2\pi/\lambda$ 。

标量波动方程的一个通解可以写成如下形式：

$$u(x, y, z, t) = U(x, y, z) \cos(\omega t + \Theta(x, y, z)) \quad (12.2-2)$$

式中， $U(x, y, z)$ 是振幅； $\Theta(x, y, z)$ 是相位。

式 (12.2-2) 的复数形式为

$$u(x, y, z, t) = \text{Re}[U(x, y, z) \exp(-j\omega t)] \quad (12.2-3)$$

式中， $U(x, y, z)$ 是复振幅，其值等于 $|U(x, y, z)| \exp(-j\Theta(x, y, z))$ 。

把式 (12.2-3) 中的 $u(x, y, z, t)$ 代入式 (12.2-1) 中得到非均匀介质的亥姆霍兹方程



$$(\nabla^2 + k^2(x, y, z))U(x, y, z) = 0 \quad (12.2-4)$$

式中与位置相关的波数 $k(x, y, z)$ 为

$$k(x, y, z) = n(x, y, z)k_0 \quad (12.2-5)$$

12.3 非均匀介质的傍轴波动方程

在 5.4 节中讨论了均匀介质的傍轴波动方程。本节将其推广到非均匀介质中。假设折射率的变化用下式表示：

$$n(x, y, z) = \bar{n} + \Delta n(x, y, z) \quad (12.3-1)$$

式中， \bar{n} 是平均折射率。

亥姆霍兹方程式 (12.2-4) 变为

$$(\nabla^2 + \bar{n}^2 k_0^2 + 2\bar{n}\Delta n k_0^2)U = 0 \quad (12.3-2)$$

式中已经忽略了 $(\Delta n)^2 k_0^2$ 。

假设光场沿 z 轴传播，则可表示为

$$U(x, y, z) = U'(x, y, z)e^{i\bar{k}z} \quad (12.3-3)$$

假设 $U'(x, y, z)$ 随 z 缓慢变化，且 \bar{k} 等于 $\bar{n}k_0$ 。式 (12.3-3) 代入亥姆霍兹方程得

$$\nabla^2 U' + 2i\bar{k}\frac{\delta}{\delta z}U' + (k^2 - \bar{k}^2)U' = 0 \quad (12.3-4)$$

由于 $U'(x, y, z)$ 是随 z 缓慢变化的，式 (12.3-3) 可以近似为

$$\frac{\delta}{\delta z}U' = \frac{i}{2\bar{k}} \left[\frac{\delta^2 U'}{\delta x^2} + \frac{\delta^2 U'}{\delta y^2} + (k^2 - \bar{k}^2)U' \right] \quad (12.3-5)$$

式 (12.3-5) 即为非均匀介质的傍轴波动方程。

式 (12.3-3) 允许不考虑关于 z 的快速相位变化因素。由此，在许多问题中，光场在 z 方向的缓慢变化，可以在数值上用比波长更加粗略的 z 来表示。消去式 (12.3-5) 中的关于 z 的二阶导数，会将需要采用迭代或特征值分析方法来求解的二阶边值问题，变成沿 z 轴简单积分就可以解决的一阶初值问题。这意味着可以减少计算时间，计算时间的减少量至少为纵向点数量级。

该方法的缺点在于，仅限于考虑沿着 z 轴（即傍轴）传播的光场，并也限制了折射率的变化，特别是折射率随着 z 轴改变。消去关于 z 的二阶导数也排除了反方向传播解的可能性。这也意味着，不能很好地对以反射为主的器件进行建模。

12.4 光束传播法 (BPM)

光束传播法是模拟光波在非均匀介质中传播的数字方法，它有多种不同形式，本节讨论的内容包含以下限制条件：反射光可以被忽略，同时所有折射率的差异比较小 [Marcuse]。

光束传播法可以从 4.3 节讨论的角谱法中导出，角谱法中的折射率为定值，即 \bar{n} 。光束传播法可以设想为这样一个模型，其包含的是被虚拟透镜分割的均匀介质，如图 12.1 所示。

光束传播法的推导是基于以下假设的，即光波被控制在宽度为 Δz ，折射率为 \bar{n} 的均匀介质片段中，但是由虚透镜生成的不同片段具有不同的相位变化，由此模型化了非均匀介质



的折射率。附加的相位变化可以认为缘于 $[n(x, y, z) - \bar{n}]$ 。若 Δz 比较小, 则该方法由两步生成: ①波在折射率为 \bar{n} 的均匀介质中传播, 当前的区域的平均折射率为 \bar{n} , 用角谱法 (ASM) 计算; ②虚透镜的作用。下面将对此进行讨论。

该方法的有效性取决于每一步 Δz 的选择, Δz 要足够小, 使其在合理的时间内可以得到理想的精确度。

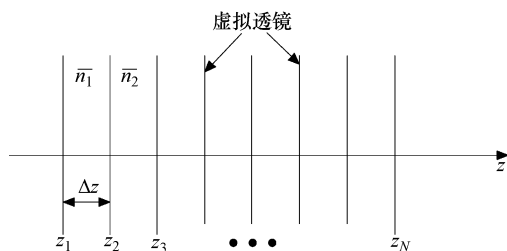


图 12.1 在非均匀介质中的光束传播法模型

12.4.1 波在折射率为 \bar{n} 的均匀介质中传播

在每一区域内, 最初认为介质是均匀的, 其折射率为 \bar{n} ; 那么, k 也等于 \bar{k} 。假设波传播的最初的输入位置是 $z - \Delta z$, 用 4.3 节的角谱法计算。

为了排除隐失波, 设 $\bar{k}^2 \geq 4\pi^2(f_x^2 + f_y^2)$, 那么 $U(x, y, z)$ 可表示为

$$U(x, y, z) = \iint_{-\infty}^{+\infty} A(f_x, f_y, z - \Delta z) \exp(+j \sqrt{\bar{k}^2 - 4\pi^2(f_x^2 + f_y^2)} z) \cdot \exp(j2\pi(f_x x + f_y y)) df_x df_y \quad (12.4-1)$$

式中, $A(f_x, f_y, z - \Delta z)$ 是式 (4.3-2) 给出的最初的平面波的频谱。

完整的过程可描述为

$$U(x, y, z) = F^{-1} \{ F \{ U(x, y, 0) \} \exp(j\mu z) \} \quad (12.4-2)$$

式中, F 和 F^{-1} 分别表示傅里叶变换和逆变换。

μ 为

$$\mu = \sqrt{\bar{n}^2 k_0^2 - 4\pi^2(f_x^2 + f_y^2)} \quad (12.4-3)$$

为了分离较大 μ 引起的 z 的快速变化, 写为如下等价的形式:

$$\mu = \bar{k} - \frac{4\pi^2(f_x^2 + f_y^2)}{\bar{k} + \sqrt{\bar{k}^2 - 4\pi^2(f_x^2 + f_y^2)}} \quad (12.4-4)$$

因此, 式 (12.4-2) 变为

$$U(x, y, z) = F^{-1} \left\{ F \{ U(x, y, z - \Delta z) \} \exp \left(-j\Delta z \frac{4\pi^2(f_x^2 + f_y^2)}{\bar{k} + \sqrt{\bar{k}^2 - 4\pi^2(f_x^2 + f_y^2)}} \right) \right\} \exp(j\bar{k}\Delta z) \quad (12.4-5)$$

12.4.2 虚透镜效应

这个过程相当于在每个区域之后用一个相位函数校正复振幅 $U(x, y, z)$ 的相位, 该相位函数选如下函数:

$$G(z) = \exp[jk(x, y, z)\Delta z] \quad (12.4-6)$$

$G(z)$ 满足

$$(\nabla^2 + k^2(x, y, z))G(z) = 0 \quad (12.4-7)$$

在波近似沿着 z 轴传播时, 选择这样的 $G(z)$ 会产生一个十分精确的相位变化。



由于式 (12.4-5) 已经包含了 $n(x, y, z) = \bar{n} + \Delta n(x, y, z)$ 和均匀介质的贡献 $\exp(\bar{k}\Delta z)$, 相位校正因子 $G(z)$ 可写为

$$G(z) = \exp(j[k(x, y, z) - \bar{k}]\Delta z) \quad (12.4-8)$$

最后, 在 z 处的光场可表示为

$$U(x, y, z) = F^{-1} \left\{ F \{ U(x, y, z - \Delta z) \} \exp \left(-j\Delta z \frac{4\pi^2(f_x^2 + f_y^2)}{k + \sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}} \right) \right\} G(z) \quad (12.4-9)$$

在一些应用中, 如集成光学, 需要二维光束传播法, 可以简单地跳过与 y 相关的项得到。

可以看出, 式 (12.4-9) 与角谱法方程式 (12.4-5) 除了相位因子 $G(z)$ 之外是相似的。因此, 它可以如 4.4 节讨论的一样, 采用快速傅里叶变换 (FFT) 进行快速计算。

可见, 推导式 (12.4-9) 时所用的近似相当于使用 12.3 节推导的傍轴方程式 (12.3-5) [Feit and Fleck, 1978]。

12.5 波在定向耦合器中的传播

作为一个采用光束传播法 (BPM) 模拟波传播的例子, 本节讨论定向耦合器。定向耦合器是光通信中常用的设备, 其包含两个相邻平行放置的绝缘波导, 这样可以使发射进入一个波导的光波耦合进入另一个波导。定向耦合器的结构示意图如图 12.2 所示。

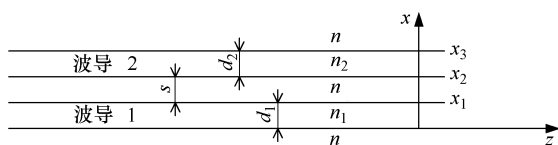


图 12.2 定向耦合器的结构示意图

12.5.1 耦合模理论概述

为了描述波在定向耦合器中的形式, 需要采用耦合模理论方程 [Saleh and Teich, 1991]。本节给出了这个理论的简要概括, 这样可以将耦合模理论的分析结果和光束传播法的模拟结果进行对比。

在波导 i 中, 沿 z 轴的峰值振幅变化定义为 $a_i(z)$, $i=1, 2$ 。光强为 $|a_i|^2$ 。耦合模方程反映了两个波导由于相互靠近而发生的能量交换。如下式所示:

$$\begin{aligned} \frac{\partial a_1}{\partial z} &= -jC_{21} \exp(j\Delta\beta z) a_2(z) \\ \frac{\partial a_2}{\partial z} &= -jC_{12} \exp(-j\Delta\beta z) a_1(z) \end{aligned} \quad (12.5-1)$$

式中, $\Delta\beta = \beta_1 - \beta_2$, β_1, β_2 表示非耦合传播常数, 它们由每个波导中的折射率决定; C_{12} 和 C_{21} 是耦合系数, 定义为

$$C_{12} = \frac{1}{2}(n_1^2 - n^2) \frac{k_0^2}{\beta_2} \int_{\text{波导2}} u_2(x) u_1(x) dx \quad (12.5-2)$$

$$C_{21} = \frac{1}{2}(n_2^2 - n^2) \frac{k_0^2}{\beta_1} \int_{\text{波导1}} u_1(x) u_2(x) dx \quad (12.5-3)$$

方程式 (12.5-1) 的通解可以由以下式子得到 [Saleh and Teich, 1991]:



$$a_1(z) = \left\{ A_1 \left[\cos(\gamma z) - j \frac{\Delta\beta}{2\gamma} \sin(\gamma z) \right] - A_2 \left[\frac{C_{21}}{j\gamma} \sin(\gamma z) \right] \right\} \exp\left(+j \frac{\Delta\beta}{2} z \right) \quad (12.5-4)$$

$$a_2(z) = \left\{ A_1 \left[\frac{C_{21}}{j\gamma} \sin(\gamma z) \right] + A_2 \left[\cos(\gamma z) + j \frac{\Delta\beta}{2\gamma} \sin(\gamma z) \right] \right\} \exp\left(-j \frac{\Delta\beta}{2} z \right) \quad (12.5-5)$$

式中, A_i 是波导的初始峰值振幅, 而且

$$\gamma^2 = \left(\frac{\Delta\beta}{2} \right)^2 + C^2 \quad (12.5-6)$$

$$C = \sqrt{C_{12}C_{21}} \quad (12.5-7)$$

12.5.2 耦合模理论和光束传播法 (BPM) 的计算对比

为了表现光束传播法模拟具有足够的精确度, 下面用光束传播法计算两个不同结构的定向耦合器的光强, 并与耦合模理论得到的分析结果进行对比 [Pojanasomboon, Ersoy, 2001]。由式 (12.5-4) 和式 (12.5-5) 知, 当波导 2 的最初强度为 0 时 ($A_2 = 0$), 分析结果可表示为

$$|a_1(z)|^2 = |A_1|^2 \left[\cos^2(\gamma z) - \left(\frac{\Delta\beta}{2\gamma} \right)^2 \sin^2(\gamma z) \right] \quad (12.5-8)$$

$$|a_2(z)|^2 = |A_1|^2 \left[\left(\frac{|C_{21}|}{\gamma} \right)^2 \sin^2(\gamma z) \right] \quad (12.5-9)$$

式中

$$\Delta\beta = \beta_1 - \beta_2, \quad \gamma^2 = \left(\frac{\Delta\beta}{2} \right)^2 + C^2 \quad (12.5-10)$$

$$C = \sqrt{C_{12}C_{21}} \quad (12.5-11)$$

情况 1: 同步波导。在这种情况下, 定向耦合器在两个波导中有相同的折射率, 也就是 $n_1 = n_2$ 。因为 β_1 和 β_2 由波导中的折射率决定, 这种情况下使得 $\beta_1 = \beta_2$ 或 $\Delta\beta = 0$ 。波导中的光强为

$$|a_1(z)|^2 = |A_1|^2 \cos^2(\gamma z) \quad (12.5-12)$$

$$|a_2(z)|^2 = |A_1|^2 \left(\frac{|C_{21}|}{\gamma} \right)^2 \sin^2(\gamma z) \quad (12.5-13)$$

在模拟中参数设置为 $A_1 = 1$, $n_1 = n_2 = 1.1$, $d_1 = d_2 = s = 1$ 。由解析表达式确定的功率交换, 如图 12.3 所示。根据式 (12.5-12) 和式 (12.5-13), 在 $z = \pi/2\gamma$ 处, 可以得到完全的功率交换。

在光束传播法 (BPM) 模拟中, 在每个波导中的任意 z 处的能量由下式计算:

$$\int_{\text{波导}} |U(x, z)|^2 dx$$

结果如图 12.4 所示。可以看到, 相同参数条件下, 光束传播法的仿真结果与图 12.3 总体上

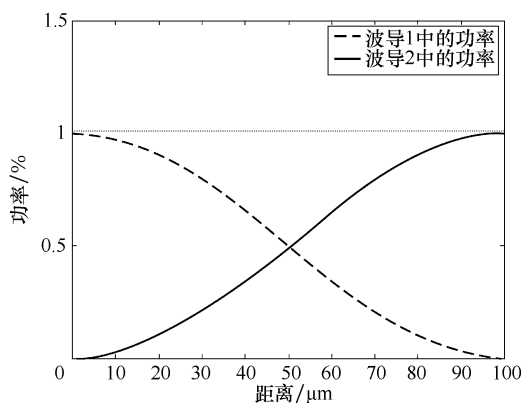


图 12.3 同步波导间功率交换的分析结果



完全相同。

情况 2：非同步波导。在非同步波导情况下，不匹配的 $n_1 \neq n_2$ 使得 $\Delta\beta = \beta_1 - \beta_2 \neq 0$ 。在波导 2 中传播的波的能量为

$$|a_2(z)|^2 = |A_1|^2 \left[\left(\frac{|C_{21}|}{\gamma} \right)^2 \sin^2(\gamma z) \right] \quad (12.5-14)$$

$\gamma = \sqrt{(\Delta\beta/2)^2 + C^2}$ 项不允许括号中的值等于 1。因此，在任意 z 处都不能完成两个波导之间完全的功率交换，如图 12.5 所示。相应的光束传播法模拟也给出了相同的结果，如图 12.6 所示。

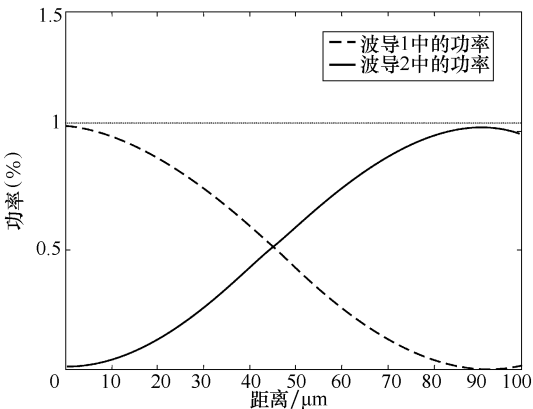


图 12.4 在同步波导定向耦合器中功率交换的光束传播法模拟

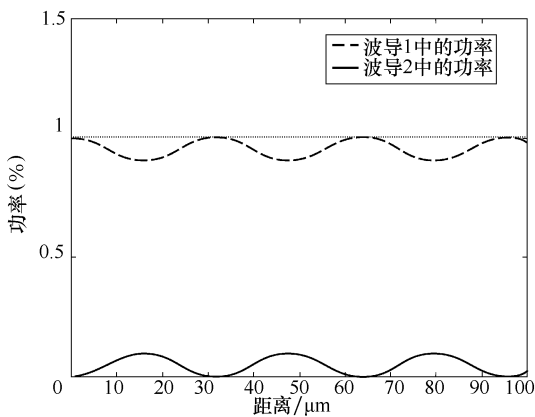


图 12.5 在非同步波导间功率交换的分析结果

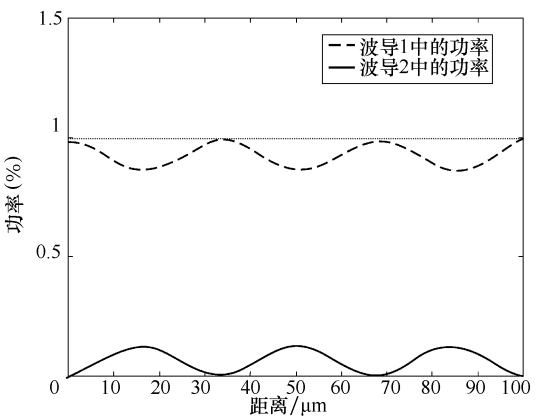


图 12.6 非同步波导间功率交换的光束传播法模拟

这些结果表明，在上述应用中，光束传播法比耦合模态理论得到的结果更准确。在比较复杂的设计中，例如，非周期光栅辅助定向耦合器，不能采用耦合模理论，而光束传播法则是可靠分析与后续设计的可选之法。

第 13 章 全 息

13.1 本章引论

全息术涉及以调制形式记录所需（物体）光波，也称为波前重建。产生的结果称为全息图。全息的主要类型有两种，分别是模拟全息和数字全息。模拟全息处理的是连续空间波 [Farhat, 1975], [Stroke, 1975]。数字全息在第 15 和 16 章中讨论，在光场经过采样、光波的振幅和（或）相位信息经特定的算法编码后，即生成了数字全息图。数字全息更多地被称为衍射光学。衍射光学的其他一些术语还有计算机全息图、衍射光学元件（DOE）和二元光学。

全息最早由丹尼斯·加博尔（Dennis Gabor）在 1948 年发现，早于激光的发明 [Gabor]。作为一个通信工程师，他发现，目标波（物光波）和参考波叠加的强度携带有物光波的振幅和相位信息。在激光作为一个相干光源出现之后，加博尔的想法变成了现实。

本章包含 6 小节。全息图的基本原理也称为相干波前记录，第一种成功用激光装置生成的全息图是利思-厄帕尼斯克斯（Leith-Upatnieks）全息图，这些将在 13.2 节介绍。13.3 节将介绍其他多种不同类型的全息图。

由于全息可以用精确的数学方法定义，所以可以在计算机上进行模拟，而且全息重建结果可以用图形显示出来，13.4 节将详细介绍相关内容。全息成像与许多参数相关，如波长和尺度。如果这些参数变化，会使再现像的性质也产生变化，在 13.5 节中将讨论全息成像和放大与这些参数的关系。作为光学成像系统，误差限制全息成像质量，13.6 节将讨论全息成像中不同类型的误差。

13.2 相干波前记录

假设物光波以下式表示：

$$U(x, y) = A(x, y) e^{j\phi(x, y)} \quad (13.2-1)$$

另外，参考光波 $R_r(x, y)$ 为

$$R_r(x, y) = B(x, y) e^{j\psi(x, y)} \quad (13.2-2)$$

两束波入射在记录介质上，记录介质对发光强度敏感，如图 13.1 或图 13.2 所示。

需要强调的是，两束波以一定的夹角传播。两束波叠加后的强度为

$$I(x, y) = |A(x, y)|^2 + |B(x, y)|^2 + 2A(x, y)B(x, y)\cos(\psi(x, y) - \phi(x, y)) \quad (13.2-3)$$

式中，最后一项等于 $AB^* + A^*B$ ，而且包含 $A(x, y)$ 和 $\phi(x, y)$ 。

包括全息胶片在内的光学记录器件的透过率函数是对发光强度敏感的。我们假设敏感度与发光强度之间是线性关系。假设 $B(x, y)$ 是定值，等于 B ，也就是说，平面波垂直照射全息图，如图 13.2 所示，相当于是点光源发出的、传播了很远的球面波。这样的器件的透过



率函数可以表示为

$$t(x, y) = C + \beta[|U|^2 + UR_r^* + U^* R_r] \quad (13.2-4)$$

式中, C 和 β 是定值; $t(x, y)$ 表示存储的信息。

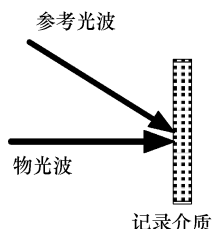


图 13.1 全息图记录的几何模型 1

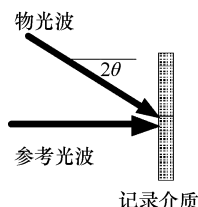


图 13.2 全息图记录的几何模型 2

现在设想用另外的参考光波 R 照射生成的全息图。从全息图发出的光波可以写为

$$Rt = U_1 + U_2 + U_3 + U_4 \quad (13.2-5)$$

式中

$$U_1 = CR \quad (13.2-6)$$

$$U_2 = \beta|U|^2 R \quad (13.2-7)$$

$$U_3 = \beta R_r^* RU \quad (13.2-8)$$

$$U_4 = \beta R_r RU^* \quad (13.2-9)$$

可以看到: 除了乘数因子之外, U_3 与 U 相同, 因此 U_3 看起来像原来的物体, 称为虚像; 同样 U_4 与 U^* 也仅仅相差了一个乘数项, U_4 相对于 U_3 会在全息图的另一侧会聚成像, 如图 13.3 所示。 U_3 和 U_4 被称为孪生像。通常, U_3 在全息图的左侧成像, 而 U_4 在全息图的右侧成像, 这些图像分别被称为虚像和实像。然而, 实际上哪个是实像, 哪个是虚像是由记录和重现时参考光的性质决定的。这些问题在 13.6 节中讨论。

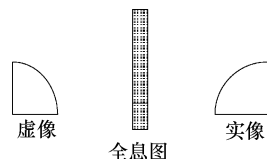


图 13.3 虚像和实像的形成

假设 R_r 和 R 相同, 是定值, 如垂直于传播方向的平面波。那么, U_3 与 U 成比例, U_4 与 U^* 成比例。这种情况下, U 和 U^* 会在空间中重合。观察者看到的不会只有 U 的原像。这由加博尔在他的全息方法中首次提到。

13.2.1 利思-厄帕尼斯克斯 (Leith-Upatnieks) 全息图

简单的垂直平面参考光波 R 可写为

$$R(x, y) = D e^{-j2\pi\alpha y} \quad (13.2-10)$$

现在 U_4 可以写成

$$U_4(x, y) = \beta D e^{-j4\pi\alpha y} U^*(x, y) \quad (13.2-11)$$

如果角 α 足够大, $e^{-j4\pi\alpha y}$ 会使 U_4 的传播方向与 U_1 、 U_2 和 U_3 的传播方向不重合。

另外可以选择使用图 13.2 所示的几何模型。在这种情况下, 参考光波垂直入射在全息图上, 但是从物体发出的物光波与 z 轴的平均夹角为 2θ 。参考光波等于 D , 是定值, 而物光波可以写为

$$U(x, y) = A(x, y) e^{-j2\pi\alpha y} \quad (13.2-12)$$



式中, $A(x, y)$ 是复数, 包含 $U(x, y)$ 的相位信息。

进一步分析得到了与式 (13.2-6) 至式 (13.2-9) 一致的结果, U_3 与 $A(x, y)e^{-j2\pi\alpha y}$ 成正比而 U_4 与 $A^*(x, y)e^{j2\pi\alpha y}$ 成正比, 因此, 它们必然在不同方向传播。

例 13.1 求最小角度 $2\theta_{\min}$, 使得该角度下所有的波成分在距离全息图足够远的位置处都能够分离。

解: 在距离全息图足够远的距离, 如, 在夫琅禾费衍射区, 波成分与傅里叶谱相似。用 FT 表示傅里叶变换, 于是有

$$\text{FT}[U_1] = C_1 \delta(f_x, f_y)$$

$$\text{FT}[U_2] = \beta_1 A_F * A_F$$

$$\text{FT}[U_3] = \beta_2 A_F(f_x, f_y - \alpha)$$

$$\text{FT}[U_4] = \beta_3 A_F(-f_x, -f_y - \alpha)$$

式中, C_1 、 β_1 、 β_2 、 β_3 是定值; * 表示自相关; A_F 是 $A(x, y)$ 的傅里叶变换。

假设 $A(x, y)$ 的带宽为 f_{\max} (周/mm)。FT $[U_1]$ 是原点处的简单 δ 函数。FT $[U_2]$ 的带宽为 $2f_{\max}$, 因为它与 A_F 的自相关函数成比例。 U_3 以 α 为中心, 而 U_4 沿着 f_y 方向以 $-\alpha$ 为中心。图 13.4 表示它们之间的关系。为了使这些项之间不会相互重合, 必须满足条件

$$\alpha \geq 3f_{\max}$$

$$2\theta_{\min} = \arcsin(3\lambda f_{\max})$$

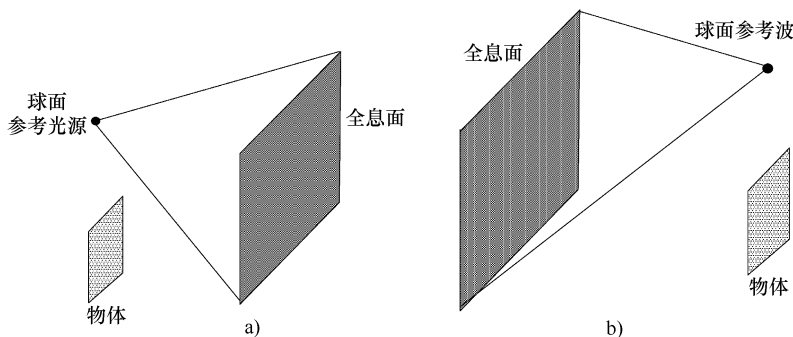


图 13.4 无透镜傅里叶变换全息图的两种几何模型

13.3 全息图的类型

有许多种类型的全息图。透射全息图可以透过光, 通过透射的光看到信息。对于反射全息图, 信息是通过全息图的反射而被看到的。以下讨论的全息图大部分是透射全息图。

13.3.1 菲涅耳和夫琅禾费全息图

在第 5 章中讨论了菲涅耳和夫琅禾费近似。如果物体相对于全息图是在菲涅耳区域重建的, 则为菲涅耳全息图。如果物体相对于全息图在夫琅禾费区域再现, 则为夫琅禾费全息图。



13.3.2 像面全息图和傅里叶全息图

在像面全息图中,再现物体的中心聚焦在成像器件面上。在傅里叶全息图中,成像平面与物体透过率函数的傅里叶变换面重合。傅里叶全息图看起来像一个被调制了的衍射光栅,这种典型的全息图相对不需要高分辨率的记录设备。

傅里叶全息图的一种变形是无透镜傅里叶变换全息图。有两种可能的几何模型来记录这种全息图,如图13.4所示。这种全息图的主要特征是球面参考光的焦点与物像在同一个平面上,无论是虚像(图13.4a)还是实像(图13.4b)。假设物光波和参考光波都是球面波而且其波源在同一平面内,在全息面,强度条纹与实傅里叶全息图具有相同类型的条纹,因此称其为无透镜傅里叶变换全息图。

相对于傅里叶和无透镜傅里叶变换全息图,菲涅耳全息图的条纹密度更大,故需要更高分辨率的记录设备。因此,在使用电子记录设备记录时,更多的是采用傅里叶全息图。

13.3.3 体全息图

体全息图记录在一个厚的介质中。这样的的好处是,全息图对波长具有了选择性,可以用白光再现。这种全息图还具有很高的衍射效率。

为了解释体全息图是如何生成的,讨论两个平面波在厚介质中干涉的简单情况。设物光波的波矢量是 \mathbf{k}_o ,参考光波的波矢量是 \mathbf{k}_r 。它们记为

$$U_o(\mathbf{r}) = A_o e^{-i\mathbf{k}_o \cdot \mathbf{r}} \quad (13.3-1)$$

$$U_r(\mathbf{r}) = A_r e^{-i\mathbf{k}_r \cdot \mathbf{r}} \quad (13.3-2)$$

两个波在介质中的干涉强度可记为

$$\begin{aligned} I(\mathbf{r}) &= |U_o(\mathbf{r}) + U_r(\mathbf{r})|^2 \\ &= I_o + I_r + 2A_o A_r \cos(\mathbf{k}_s \cdot \mathbf{r}) \end{aligned} \quad (13.3-3)$$

式中

$$I_o = |A_o|^2 \quad (13.3-4)$$

$$I_r = |A_r|^2 \quad (13.3-5)$$

$$\mathbf{k}_s = \mathbf{k}_o - \mathbf{k}_r \quad (13.3-6)$$

式(13.3-3)表示一个周期为 $p = 2\pi/|\mathbf{k}_s|$ 的正弦条纹。

假设 \mathbf{k}_r 指向 z 轴方向, \mathbf{k}_o 与 z 轴夹角为 θ ,如图13.5所示。于是,周期为

$$p = \frac{\lambda}{2\sin(\theta/2)} \quad (13.3-7)$$

在厚乳胶中记录条纹时则产生体光栅。如果在重现时采用相同的参考光,即再现光和全息图的夹角与记录时一样,再现光的波长也与记录时参考光的波长一样,那么物光波可以得到完美重现 [Saleh and Teich, 1991]。如果再现光波长变化了,则不能再现出物光波。

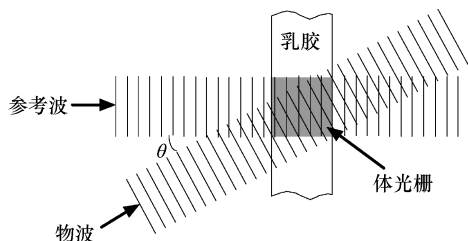


图13.5 两束平面波在厚感光乳胶中生成体光栅的相干模式



如果用白光重现，白光中与记录时参考光波长相同的成分会使物光波重现。这是一个很有用的特征，因为由此白光可以用来进行全息重现了。

13.3.4 模压全息图

模压是在复制激光唱片（CD）和数字化激光视频光盘（DVD）时使用的工艺，具有波长量级精度。可以采用相同的工艺来制作和复制表面浮雕式全息图。用这种方法，可以把表面浮雕式全息图记录在光刻胶或光导热塑料上。价格低廉，以及大规模生产，使其在市场上大量应用，例如，在信用卡或商品的防伪方面。

模压全息图制作过程的第一步是用光刻胶或光导热塑料制作全息图。然后用电铸法把光刻胶全息图制作成一个金属主全息图。首先，在光刻胶全息图上面生成一个薄镍层；接着，该镍层从光刻胶主全息图分离，生成金属第二级主全息图，其可以被简单地多次复制，在模压过程中，把金属全息图加热到很高的温度，然后盖压在如聚酯纤维一类的胶片上。当胶片冷却、移开压力后，条纹形状保留了下来。常在模压全息图上镀铝来生成最后的反射全息图。

13.4 用计算机模拟全息图的再现

通过计算全息过程方程，可以把存储在计算机里的全息图进行再现。如，采用在第7章讨论的近远场近似（NFFA）方法。考虑方程式（13.2-3）中最重要的几项，全息图可以表示为如下模型：

$$h(x,y) = u(x,y) + u^*(x,y) + c \quad (13.4-1)$$

式中， $u(x,y)$ 是物光波； c 是定值。

方程式（13.4-1）包含孪生像和导致中心峰值的常数项。

图 13.6 表示一个物体的图像，其全息图是用逆 NFFA 传播法生成 $u(x,y)$ ，相关物理参数设置如下：

$z = 80\text{mm}$ ， $\lambda = 0.6328\mu\text{m}$ ，全息图的尺寸为 $400\mu\text{m} \times 400\mu\text{m}$ ，输出尺寸为 $64.8\text{mm} \times 64.8\text{mm}$ 。

生成的全息图如图 13.7 所示。采用前向 NFFA 传播法获得的全息图的再现，结果如图 13.8 所示。



图 13.6 再现的物体

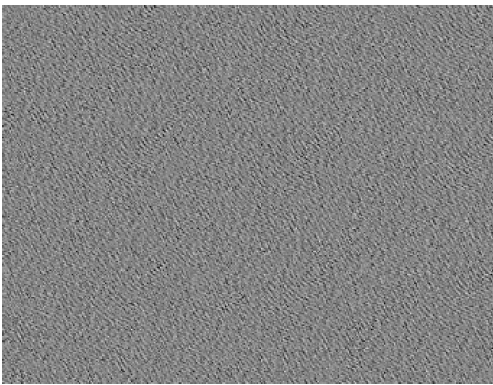


图 13.7 物体的全息图

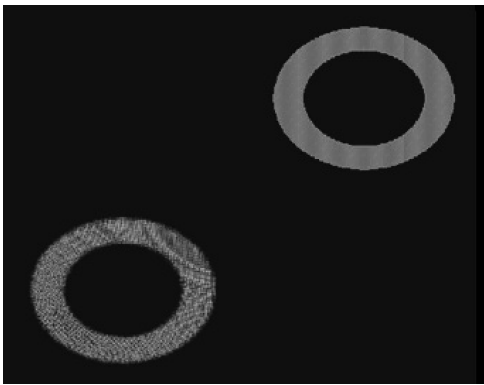


图 13.8 全息图再现的双像



13.5 全息成像的放大和分析

成像的重要项分别是式 (13.2-8) 和式 (13.2-9) 中的 U_3 和 U_4 。有以下假设:

- 物光波和参考光波分别是点 (x_o, y_o, z_o) 和 (x_r, y_r, z_r) 发出的球面波。
- 球面再现波与记录波可能不同, 它的起点在 (x_c, y_c, z_c) 。
- 记录和再现波长分别是 λ_1 和 λ_2 , λ_1 与 λ_2 可能不相等。
- 最终生成的全息图与通过记录生成的全息图可能有不同的尺寸。于是, 最初全息图上在 $z=0$ 处的一点 (x, y) , 在最终的全息图上可能相应地变为 (x', y') 。

物光波在全息图上点 (x, y) 处产生的相位, 可由相对原物点的关系表示为

$$\phi_o(x, y) = \frac{2\pi}{\lambda_1} \{ [(x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2]^{1/2} - (x_o^2 + y_o^2 + z_o^2)^{1/2} \} \quad (13.5-1)$$

保留泰勒展开的前两项, 最终的方程可以写成如下的非常数项:

$$\phi_o(x, y) = \frac{2\pi}{\lambda_1} \left[\frac{1}{2z_o} (x^2 + y^2 - 2xx_o - 2yy_o + z_o^2) + \beta(x, y) \right] \quad (13.5-2)$$

式中, $\beta(x, y)$ 项可以写为

$$\beta(x, y) = -\frac{1}{8z_o^3} \begin{pmatrix} x^4 + y^4 + 2x^2y^2 - 4x^3x_o \\ -4y^3y_o - 4x^2yy_o - 4xy^2x_o \\ + 6x^2x_o^2 + 6y^2y_o^2 + 2x^2y_o^2 \\ + 2y^2x_o^2 + 8xyx_o y_o - 4xx_o^3 \\ - 4yy_o^3 - 4xx_o y_o^2 - 4xx_o^2 y_o \end{pmatrix} \quad (13.5-3)$$

$\beta(x, y)$ 与 13.6 节中讨论的像差有关。下面可以类似地写出全息面上点 (x, y) 的方程, 其中涉及再现光波在波长 λ_2 产生的相位 $\phi_c(x, y)$, 以及记录参考光在波长 λ_1 产生的相位 $\phi_r(x, y)$ 。关于式 (13.2-8) 和式 (13.2-9), U_3 和 U_4 中重要的相位项可以写为

$$\phi_3 = \phi_c + \phi_o - \phi_r \quad (13.5-4)$$

$$\phi_4 = \phi_c + \phi_o + \phi_r \quad (13.5-5)$$

就此, 对 $1/z$ 或 $1/z_c$ 或 $1/z_o$, 忽略高于一阶的项。用 ϕ_l 表示 ϕ_3 或 ϕ_4 , 得到

$$\begin{aligned} \phi_l(x, y) = & \frac{2\pi}{\lambda_2} \frac{1}{2z_c} (x'^2 + y'^2 \pm 2x'x_r \pm 2y'y_r) + \frac{2\pi}{\lambda_1} \frac{1}{2z_o} (x^2 + y^2 - 2xx_o - 2yy_o) - \\ & \frac{2\pi}{\lambda_1} \frac{1}{2z_r} (x^2 + y^2 \pm 2xx_r \pm 2yy_r) \end{aligned} \quad (13.5-6)$$

式中, $+$ 对应 ϕ_3 ; $-$ 对应 ϕ_4 。

定义全息图放大率 M_h 为

$$M_h = \frac{x'}{x} = \frac{y'}{y} \quad (13.5-7)$$

波长比为

$$\mu = \frac{\lambda_2}{\lambda_1} \quad (13.5-8)$$



那么, 式 (13.5-6) 可写为

$$\phi_I(x, y) = \frac{\pi}{\lambda_2} \left[\begin{aligned} &(x'^2 + y'^2) \left(\frac{1}{z_c} + \frac{\mu}{M_h^2 z_o} - \frac{\mu}{M_h^2 z_r} \right) \\ &- 2x \left(\frac{x_c}{z_c} + \frac{ux_o}{M_h z_o} \pm \frac{ux_r}{M_h z_r} \right) \\ &- 2y \left(\frac{y_c}{z_c} + \frac{uy_o}{M_h z_o} \pm \frac{uy_r}{M_h z_r} \right) \end{aligned} \right] \quad (13.5-9)$$

式 (13.5-9) 可以看作从点 (x_I, y_I, z_I) 发出的球面波对应的相位。在非涅耳近似条件下, 相关的相位可以表示为

$$\phi'_I = \frac{\pi}{\lambda_2 z_I} (x'^2 + y'^2 - 2x'x_I - 2y'y_I) \quad (13.5-10)$$

设 $\phi'_I = \phi_I$, 则有

$$z_I = \frac{M_h^2 z_o z_r}{M_h^2 z_o z_r + \mu z_c z_r - \mu z_c z_o} \quad (13.5-11)$$

$$x_I = \frac{M_h^2 x_c z_o z_r + M_h x_o z_c z_r \pm \mu M_h x_r z_c z_o}{M_h^2 x_c z_o z_r + \mu z_c z_r - \mu z_c z_o} \quad (13.5-12)$$

$$y_I = \frac{M_h^2 y_c z_o z_r + \mu M_h y_o z_c z_r \pm \mu M_h y_r z_c z_o}{M_h^2 y_c z_o z_r + \mu z_c z_r - \mu z_c z_o} \quad (13.5-13)$$

哪个图像是虚像, 哪个图像是实像分别由 z_r 和 z_c 的符号决定。横向放大率为

$$M_t = \frac{\partial x_I}{\partial x_o} = \frac{M_h}{\left(1 + \frac{M_h^2 z_o}{\mu z_c} - \frac{z_o}{z_r} \right)} \quad (13.5-14)$$

$$\begin{aligned} M_a = \frac{\partial z_I}{\partial z_o} &= -\frac{M_h^2}{\mu} \frac{d}{dz_o} \left[\frac{z_o}{1 - z_o \left(\frac{M_h^2}{\mu z_c} + \frac{1}{z_r} \right)} \right] \\ &= -\frac{1}{\mu} \frac{M_h^2}{\left[1 + \frac{M_h^2 z_o}{\mu z_c} - \frac{z_o}{z_r} \right]^2} \\ &= -\frac{1}{\mu} M_t^2 \end{aligned} \quad (13.5-15)$$

例 13.2 求使放大率 $M_a = M_t$ 的 μ 值。

解: 设

$$|M_a| = \frac{1}{\mu} M_t^2 = M_t$$

因此

$$\mu = M_t \quad (13.5-16)$$

这表明全息图尺寸和参考光起点的变化, 可以通过选择一个新的满足式 (13.5-16) 的波长来补偿。由此类推, 如果波长变化了, 则可以改变全息图的尺寸或者参考光的起点, 使两种类型的放大率尽可能相等。



13.6 像差

为了方便,讨论像差时,把直角坐标系 x, y 换成极坐标系 r, θ 。定义源点在 (x_I, y_I, z_I) 处的理想球面波前与实际的波前之间的相位差 $\phi_A(r, \theta)$ 为波前像差。

考虑式 (13.5-3) 中三阶项引起的图像 U_3 的像差。三阶项中 ϕ_c, ϕ_o, ϕ_r 一起贡献于实际的波前。波像差方程为

$$\phi(r, \theta) = \frac{2\pi}{\lambda_2} \left[\begin{aligned} & -\frac{1}{8}r^4 S + \frac{1}{2}r^3 (C_x \cos\theta + C_y \sin\theta) \\ & -\frac{1}{2}r^2 (A_x \cos^2\theta + A_y \sin^2\theta + 2A_x A_y \cos\theta \sin\theta) \\ & -\frac{1}{4}r^2 F + \frac{1}{2}r (D_x \cos\theta + D_y \sin\theta) \end{aligned} \right] \quad (13.6-1)$$

式中的各参数分别属于如下像差:

S : 球差;

C_x, C_y : 彗差;

A_x, A_y : 像散;

F : 场曲;

D_x, D_y : 畸变。

它们用如下方程表示:

$$S = \frac{1}{z_c^3} - \frac{\mu}{M_h^4 z_o^3} + \frac{\mu}{M_h^4 z_r^3} - \frac{1}{z_I^3} \quad (13.6-2)$$

$$C_x = \frac{x_c}{z_c^3} - \frac{\mu x_o}{M_h^3 z_o^3} + \frac{\mu x_r}{M_h^3 z_r^3} - \frac{x_I}{z_I^3} \quad (13.6-3)$$

$$C_y = \frac{y_c}{z_c^3} - \frac{\mu y_o}{M_h^3 z_o^3} + \frac{\mu y_r}{M_h^3 z_r^3} - \frac{y_I}{z_I^3} \quad (13.6-4)$$

$$A_x = \frac{x_c^2}{z_c^3} - \frac{\mu x_o^2}{M_h^2 z_o^3} + \frac{\mu x_r^2}{M_h^2 z_r^3} - \frac{x_I^2}{z_I^3} \quad (13.6-5)$$

$$A_y = \frac{y_c^2}{z_c^3} - \frac{\mu y_o^2}{M_h^2 z_o^3} + \frac{\mu y_r^2}{M_h^2 z_r^3} - \frac{y_I^2}{z_I^3} \quad (13.6-6)$$

$$F = \frac{x_c^2 + y_c^2}{z_c^3} - \frac{\mu(x_o^2 + y_o^2)}{M_h^2 z_o^3} + \frac{\mu(x_r^2 + y_r^2)}{M_h^2 z_r^3} - \frac{x_I^2 + y_I^2}{z_I^3} \quad (13.6-7)$$

$$D_x = \frac{x_c^3 + x_c y_c^2}{z_c^3} - \frac{\mu(x_o^3 + x_o y_o^2)}{M_h z_o^3} + \frac{\mu(x_r^3 + x_r y_r^2)}{M_h z_r^3} - \frac{x_I^3 + x_I y_I^2}{z_I^3} \quad (13.6-8)$$

$$D_y = \frac{y_c^3 + y_c x_c^2}{z_c^3} - \frac{\mu(y_o^3 + y_o x_o^2)}{M_h z_o^3} + \frac{\mu(y_r^3 + y_r x_r^2)}{M_h z_r^3} - \frac{y_I^3 + y_I x_I^2}{z_I^3} \quad (13.6-9)$$

例 13.3 当记录 and 再现波都是平面波时,求 S, C_x 和 A_x 。

解: 此时, z_r 和 z_c 是无限的, 因此, S, C_x 和 A_x 成为



$$S = \frac{\mu}{M_h^4 z_o^3} \left[\frac{u^2}{M_h^2} - 1 \right]$$

$$C_x = \frac{\mu}{M_h^3 z_o^2} \left[\frac{x_c}{z_c} \frac{\mu}{M_h} - \frac{x_o}{z_o} \left(\frac{\mu^2}{M_h^2} - 1 \right) + \frac{x_r}{z_r} \frac{\mu^2}{M_h^2} \right]$$

$$A_x = \frac{\mu}{M_h^2 z_o} \left[\frac{x_o^2}{z_o^2} \left(\frac{\mu^2}{M_h^2} - 1 \right) - \frac{2\mu}{M_h} \frac{x_o}{z_o} \left(\frac{x_c}{z_c} + \frac{\mu x_r}{M_h z_r} \right) + \left(\frac{x_c}{z_c} + \frac{\mu x_r}{M_h z_r} \right)^2 \right]$$

例 13.4 当 $z_r = z_o$ 时, 求 C_x 和 A_x 。

解: 通过代数运算之后, 有

$$C_x = \frac{\mu}{M_h} \left[-\frac{x_o}{z_o} \left(\frac{1}{M_h^2 z_o^2} - \frac{1}{z_c^2} \right) + \frac{x_r}{z_r} \left(\frac{1}{M_h^2 z_o^2} - \frac{1}{z_c^2} \right) \right]$$

$$A_x = \frac{\mu}{M_h} \left[-\frac{x_o^2}{z_o^2} \left(\frac{1}{M_h z_o} + \frac{\mu}{M_h z_c} \right) + \frac{2}{z_c} \frac{x_o}{z_o} \left(\frac{x_c}{z_c} + \frac{\mu x_r}{M_h z_r} \right) + \frac{x_r^2}{z_r^2} \left(\frac{1}{M_h z_o} - \frac{\mu}{M_h z_c} \right) - \frac{2}{z_c} \frac{x_c}{z_c} \frac{x_r}{z_r} \right]$$

第 14 章 切趾，超分辨率和缺失信息恢复

14.1 本章引论

光学成像时，衍射是限制系统分辨率的主要现象，例如，利用望远镜对距离遥远的恒星成像。因为恒星的距离非常遥远，理论上应该在焦平面上成像，理想条件下会是一个点。相反，图像强度与望远镜的出瞳函数的傅里叶变换的二次方成正比。结果，图像中包含一个中心主极大，同时在辐向围绕一些次极大或极小强度条纹，这些次极大或者极小条纹会被误认为是其他恒星的像。同样的原因，附近第二个相对比较暗的恒星会一起消失。

切趾和超分辨率技术的目的是最大程度减小有限尺寸孔径的影响。超分辨率与缺失信息恢复紧密相关。

本章讨论的算法主要是处理离散数字信号，因此能够直接在计算机中应用。如果需要的话，这些算法也能够方便地转化为模拟表示法。

本章包含 18 节。14.2 节是关于切趾或加窗技术的。本节着重介绍各种不同类型的窗口函数。两点分辨率在光学图像设备中是固有的。14.3 节介绍了光学成像器件固有的两点分辨率以及信号恢复的一般原则。

在随后的章节，首先介绍了开发信号恢复算法所需要的基本理论。14.4 节介绍了收缩的基本理论，给出了在线性空间用矢量的投影变换生成一个不动点的条件。在 14.5 节介绍被称为约束迭代信号恢复法的收缩迭代方法，14.6 节继续讨论这一问题，此时的情况是反卷积，称为约束迭代反卷积法。

14.7 节介绍了一个收缩子集的投影。有许多不同的投影类型，投影到凸集的方法是相对容易处理的方法之一，其在 14.8 节中介绍。随后的章节对这个方法进行了进一步的讨论，它是设计衍射光学元件和其他衍射光学设备的工具。盖师贝格-帕普里斯（Gerchberg-Papoulis）算法是凸集投影算法的特例，在该方法中，信号和傅里叶域在特定的凸集里得到了一并应用，这些内容在 14.9 节中讨论。14.10 节给出了凸集投影算法的其他例子。

两个令人感兴趣的应用分别是基于相位的恢复和基于幅值的恢复，后者比前者更难处理。一种相位恢复的迭代优化算法在 14.11 节中讨论。14.12 节的主要内容是用离散傅里叶变换从离散相位函数中恢复信号。

一个更难的投影类型是广义投影，此时没有凸集的限制。这种情况下，迭代可能有也可能没有最优的解，这部分内容将在 14.13 节中介绍。幅值恢复，也称为相位恢复（phase retrieval），常可表示为一个广义投影问题。因此，14.14 节讨论了一个特别的迭代优化方法。

剩下的章节讨论了用于信号和图像复原的其他类型的算法。14.15 节重点描述了用于图像恢复的最小二乘法和广义逆法。14.16 节介绍了计算广义逆的奇异值分解问题。最速下降法和共轭梯度法分别在 14.17 节和 14.18 节中进行了介绍，它们同样是为了计算广义逆。



14.2 切趾

切趾与加窗类似，是信号处理中的一个主要问题。不失一般性，下面在一维空间讨论。

出瞳函数的主要作用是截断输入波场。在信号理论中，这相当于对输入信号进行截断。当输入信号被截断时，会产生吉布斯现象，也称为振铃。这意味着不连续点附近会产生振荡。当频谱被截断时也会有这样的现象，此时振铃在信号重构时出现。

为了能够分析被截断波场的特性，先对信号进行采样，并对得到的数据序列进行数字处理。令 $z[n]$ 是被出瞳函数截断的信号序列， $Z[f]$ 是 $z[n]$ 的离散时间傅里叶变换。

令 $u[n]$ 的采样间隔是 T_s 。截断等于 $u[n]$ 乘以矩形序列 $\omega[n]$ ，即为

$$u'[n] = u[n]\omega[n] \quad (14.2-1)$$

式中

$$\omega[n] = \begin{cases} 1 & |n| \leq N \\ 0 & \text{其他} \end{cases} \quad (14.2-2)$$

$\omega[n]$ 的离散时间傅里叶变换（见附录 B 中对离散时间傅里叶变换的讨论）为

$$\begin{aligned} W(f) &= \sum_{n=-N}^N e^{-j2\pi f n T_s} \\ &= \frac{\sin(\pi f T_s (2N+1))}{\sin(\pi f T_s)} \end{aligned} \quad (14.2-3)$$

当 T_s 等于 1， N 分别等于 6, 8, 15 和 25 时， $W(f)/N$ 在图 14.1 中给出。可以看到， $W(f)/N$ 具有相当大幅值的主瓣和旁瓣。随着 N 增加，主瓣的宽度和旁瓣下面积不断减小，但是主瓣的高度一直不变。

由卷积定理知，式 (14.2-1) 在离散时间傅里叶变换域等价于

$$U'(f) = U(f) * W(f) \quad (14.2-4)$$

因此需要寻找一个具有最小旁瓣和最窄主瓣宽度的最优 $W(f)$ ，来使 $Z'(k)$ 尽可能地接近 $U(f)$ 。显然，最优的 $W(f)$ 是 $\delta(f)$ ，意味着没有加窗。需要找到一个折中的方法，在 $|n| \leq N$ 时，窗口是非零的，并且它的 $W(f)$ 与 $\delta(f)$ 尽可能相似。在寻找这个折中方案的过程中，无免费的午餐这条定律再次得到应验。消除旁瓣的代价是加宽了主瓣，反之亦然。加宽主瓣相当于低通滤波，使 $U(f)$ 的快速变化部分变得平滑。

以往发表的文献中讨论了许多窗函数。历史上首次报道的是费耶尔算术平均法和兰措什 σ 因子 [Lanczos]。在模拟域，窗函数也是模拟的。对其采样可以得到离散时间窗。下面回顾一些常用的离散时间窗。

14.2.1 离散时间窗

下面讨论的这些窗函数是非因果的。因果窗函数可以用 $\omega[n-N]$ 替换 $\omega[n]$ 得到。

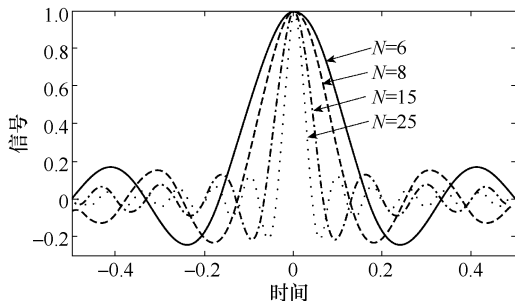


图 14.1 长度 $N=6, 8, 15$ 和 25 的矩形窗函数的离散时间傅里叶变换



1. 巴特利特三角窗

$$\omega[n] = \begin{cases} 1 - \frac{n}{N} & 0 \leq n \leq N \\ 1 + \frac{n}{N} & -N \leq n \leq 0 \\ 0 & \text{其他} \end{cases} \quad (14.2-5)$$

2. 广义余弦窗

$$\omega[n] = \begin{cases} a + b \cos\left(\frac{\pi n}{N}\right) + c \cos\left(\frac{2\pi n}{N}\right) & -N \leq n \leq N \\ 0 & \text{其他} \end{cases} \quad (14.2-6)$$

式中, 常数 a, b, c 取下面的一系列值可以得到以下 3 种特定的窗。

窗 (Window)	a	b	c
汉宁窗 (Hanning)	0.5	0.5	0
海明窗 (Hamming)	0.54	0.46	0
布莱克曼窗 (Blackman)	0.42	0.5	0.08

3. 高斯窗

$$\omega[n] = \exp\left(-\frac{1}{2}\left[\frac{\alpha n}{N}\right]^2\right) \quad (14.2-7)$$

式中, α 是使用者选择的参数。

4. 道尔夫-契比雪夫窗

这个窗通过在给定的旁瓣水平条件下, 最小化主瓣宽度来获取。它们的离散傅里叶变换是

$$W[k] = W[-k] = (-1)^n \frac{\cos\left[N \arccos\left[\beta \cos\left(\frac{\pi k}{N}\right)\right]\right]}{\cosh(N \operatorname{arccosh}(\beta))} \quad (0 \leq n \leq N-1) \quad (14.2-8)$$

式中

$$\beta = \cosh\left[\frac{1}{N} \operatorname{arccosh}(10^\alpha)\right] \quad (14.2-9)$$

对应的时间窗 $\omega[n]$ 可通过计算 $W[n]$ 的逆离散傅里叶变换和峰值振幅的单位化缩放来获取。参数 α 表示主瓣水平与旁瓣水平比值的对数值。例如, α 等于 3 的意思是旁瓣在主瓣 3 个数量级以下, 或者, 旁瓣在主瓣 60dB 以下。

在目前已经讨论的这些窗中, 矩形窗是最简单的窗。巴特利特窗减少了不连续点的超调量, 但是引起了过度的平滑。汉宁窗、海明窗和布莱克曼窗可以得到平滑截断的 $u[n]$, 且只有很小的旁瓣, 其中布莱克曼窗具有最好的性能。海明窗和高斯窗在 $|n| = N$ 时不等于 0。

接下来将讨论一般使用中性能最好的凯瑟窗。

5. 凯瑟窗

$$\omega[n] = \begin{cases} \frac{I_0\left(\alpha\left[1 - \left(\frac{n}{N}\right)^2\right]^{1/2}\right)}{I_0(\alpha)} & 0 \leq |n| \leq N \\ 0 & \text{其他} \end{cases} \quad (14.2-10)$$



式中， $I_0(x)$ 是 0 阶修正的第一类贝塞尔函数，为

I_0(x) = \sum_{m=0}^{+\infty} \left[\frac{\left(\frac{x}{2}\right)^m}{m!} \right]^2 \quad (14.2-11)

式中，α 是形状参数，通过选择合适的 α 使主瓣和旁瓣的宽度取得一个折中。

当 α = 0 时，凯瑟窗和矩形窗相同。在图 14.2 中给出了当 N = 20，α = 2，4，8 和 20 时的 ω[n] 中。保持 N 不变，增大 α 则旁瓣会减少，但同时也增加了主瓣的宽度。保持 α 不变，增大 N，则减小了主瓣宽度而旁瓣没有明显变化。

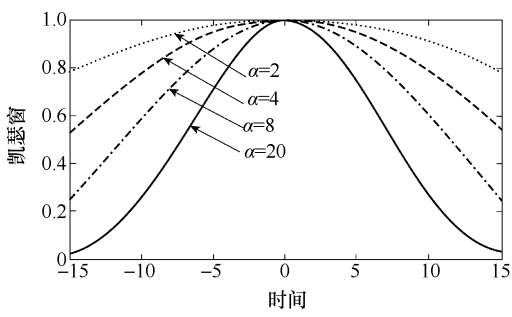


图 14.2 N = 20 的凯瑟窗，α = 2，4，8，20

例 14.1 针对例 5.8 用汉宁窗沿 x 方向切趾。比较 x 方向和 y 方向的结果。

解：当采用相同孔径的汉宁窗沿着 x 方向切趾时，得到夫琅禾费衍射结果，如图 14.3 所示。可以看到 x 方向的第二主极大比 y 方向的第二主极大要小很多。这是由沿 x 方向的切趾造成的。

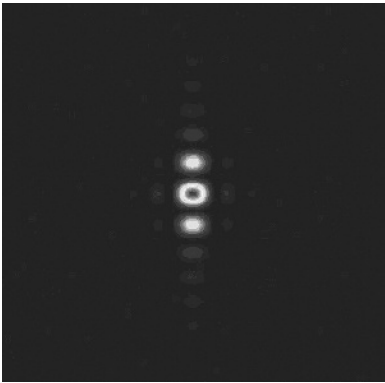


图 14.3 在 x 方向上用汉宁窗切趾的矩形孔的夫琅禾费衍射图样

14.3 两点分辨率和信号恢复

考虑一个半径为 R 的圆孔的夫琅禾费衍射图样。如例 5.8 中讨论的，平面波入射在圆孔上产生的衍射条纹强度分布中有一个中心斑，它的半径为

\delta = \frac{0.61\lambda d_0}{R} \quad (14.3-1)

式中，d_0 是条纹中心和圆孔间的距离；δ 被称为瑞利长度，是衡量能够分辨的两个点的限度，δ 是由光学系统有限孔径的出瞳导致的。

有限孔径同样意味着信息缺失，这是因为孔径外的波场被截断了。因为讨论的主题与信



号处理中的信号恢复相同, 所以后述的波场指的是信号。另外, 窗函数被推广为畸变或变换算符 D , 与文献中的信号恢复常用表示方法相同。不失一般性, 还是讨论一维的情况。

对于一个空间有限信号, 理论上是可以精确恢复的, 这是因为这类信号的傅里叶变换是一个解析函数 [Stark, 1987]。如果在任意小的区域内确知一个解析函数, 则完整的函数可以通过解析延拓得到。因此, 如果可以精确获知这类信号频谱的一小部分, 则该信号的所有频谱和信号本身将能够恢复。

不幸的是, 精确获知信号的部分频谱往往是不可能的, 尤其当测量噪声存在的时候。对信号的非精确测量和一些先验信息, 如正性、有限长度等, 反而是可以获取的信息。

实测信号 v 可以写为理想信号 u 的如下形式:

$$v = Du \quad (14.3-2)$$

式中, D 是畸变算符或变换算符。

已知 D , 问题变为恢复 z (对于系统识别, 问题是在已知 v 和 u 的条件下, 估算 D)。

直接求解 u 的方法是找到 D^{-1} 使得

$$u = D^{-1}v, v = Du \quad (14.3-3)$$

不幸的是, 变换算符 D 往往难以得到。例如, 当同一个 v 对应多个 u 时, 将不存在 D^{-1} , 例如, 当 D 相当于低通滤波器时就是这种情况。即使可以估计 D^{-1} , 但是它常常是病态的, D^{-1} 较小的误差就会引起估计 u 时较大的误差。在这样的条件下恢复 u 的问题常被叫作逆问题。在接下来的章节部分, 利用收缩和投影的信号复原将会作为解决逆问题的潜在方法并加以研究。这些方法的实现经常要利用到信号本身及其频谱域的先验信息。

14.4 收缩

在下面要讨论的内容里, 假设信号是在完全线性赋范矢量空间的矢量。附录 B 讨论了线性矢量空间。赋范矢量在 B.3 部分做了介绍。将要用到的欧几里得范数定义如下:

连续信号

$$\|u\| = \left[\int_{-\infty}^{\infty} |z(t)|^2 dt \right]^{1/2} \quad (14.4-1)$$

离散信号

$$\|u\| = \left[\sum_{n=-\infty}^{\infty} |u[n]|^2 \right]^{1/2} \quad (14.4-2)$$

接下来, 假设信号 u 是离散的, 属于一个有限维的矢量空间 S , S 可以认为是实空间 \mathbf{R}^n 或者复空间 \mathbf{C}^n 的子集。 n 是空间的维数。

令 S_0 为 S 的子空间。通过算符 A 对空间 S_0 中的矢量进行映射, 如果该映射满足下述条件, 则可定义为非扩张映射

$$\|Au - Av\| \leq \|u - v\| \quad (14.4-3)$$

式 (14.4-3) 对于属于 S_0 的每一个 u 和 v 都成立。如果式 (14.4-3) 对所有 $u \neq v$ 总成立, 则映射为严格非扩张映射。

作用于 u 的线性算符常用矩阵表示。矩阵的范数可以用多种方式定义。例如, 线性算符 A 的谱范数由下式给出:



$$|A| = |R(A^H A)|^{1/2} \quad (14.4-4)$$

式中, A^H 表示 A 转置的复共轭 (即厄米特共轭矩阵); $R(B)$ 是 B 的谱半径, 其定义为 B 的特征值的绝对值中的最大值。

如果满足下式, 则矩阵的范数与矢量的范数是一致的

$$|Au| \leq |A| |x| \quad (14.4-5)$$

例如, 式 (14.4-4) 给出的谱范数就是一致的。对于一个非扩张映射, 由不等式 (14.4-3), A 必须满足

$$|A| \leq 1 \quad (14.4-6)$$

另一个重要的方面是线性算符 A 的不动点。考虑

$$|Au| = u \quad (14.4-7)$$

任意满足式 (14.4-7) 的矢量 u^* 称为 A 的一个不动点。

如果 A 是严格非扩张映射, 且假设其有两个不动点 u^* 和 v^* , 那么

$$|u^* - v^*| = |Au^* - Av^*| < |u^* - v^*| \quad (14.4-8)$$

显然, 式 (14.4-8) 是自相矛盾的。因此, 对于严格非扩张收缩, 其不动点不多于一个。

14.4.1 收缩映射定理

当 A 为严格非扩张映射, 并且满足

$$|Au - Av| < \alpha |u - v| \quad (14.4-9)$$

时, 式中, $0 < \alpha < 1$, 则 A 在 S_0 上有唯一的不动点。

证明: 令 u_0 是 S_0 上任意一点。序列 $u_k = Au_{k-1}$, $k = 1, 2, \dots$ 。序列 $[u_k]$ 属于 S_0 。于是有

$$|u_{k+1} - u_k| = |Au_k - Au_{k-1}| \leq \alpha |u_k - u_{k-1}| \quad (14.4-10)$$

反复利用施瓦茨不等式, 有

$$\begin{aligned} |u_{k+p} - u_k| &= |u_{k+p} - u_{k+p-1} + u_{k+p-1} - u_{k+p-2} + u_{k+p-2} - \dots - u_k| \\ &\leq \sum_{i=1}^p |u_{k+i} - u_{k+i-1}| \leq (\alpha^{p-1} + \alpha^{p-2} + \dots + 1) \sum_{i=1}^p |u_{k+i} - u_k| \end{aligned} \quad (14.4-11)$$

因为

$$|u_{k+1} - u_k| < \alpha^k |u_1 - u_0| \quad (14.4-12)$$

且

$$\alpha^{p-1} + \alpha^{p-2} + \dots + 1 = \frac{1 - \alpha^p}{1 - \alpha} < \frac{1}{1 - \alpha} \quad (14.4-13)$$

式 (14.4-11) 可写为

$$|u_{k+p} - u_k| \leq \frac{\alpha^k}{1 - \alpha} |u_1 - u_0| \quad (14.4-14)$$

对于任意的 $\varepsilon > 0$, 都有正整数 $N(\varepsilon)$ 满足 $n, m > N(\varepsilon)$, 有

$$|a_n - a_m| < \varepsilon \quad (14.4-15)$$

则序列 $[a_n]$ 称为柯西序列。

利用数学分析的基本理论, 如果一个序列是柯西序列, 那么它会收敛到一个不动点。式



(14.4-14) 表明, $[u_k]$ 是柯西序列, 因此它满足

$$\lim_{k \rightarrow \infty} Au_k = Au^* = u^* \quad (14.4-16)$$

式中, u^* 即是不动点。

因此, 如果 A 是严格非扩张映射, 则 $u_{k+1} = Au_k$ 产生的序列收敛于唯一的不动点 u^* 。

14.5 用于信号恢复的收缩迭代方法

假设一实测信号 v , 以某种方式产生了畸变。从 v 开始, 希望通过迭代算法获取一个更精确的信号。迭代 k 次后的信号为 u_k , k 是表示迭代次数的整数。随着 k 的增加, u_k 逐渐接近要恢复的信号。

能达到上述目的的著名方法是约束迭代信号恢复方法 [Schafer]。该方法基于以下迭代方程:

$$u_{k+1} = Fu_k \quad (14.5-1)$$

式中, 序列 u_k 是逐步接近于需要恢复信号的序列; F 是一个可从某种约束形式的畸变算符和先验知识获取的算符。

一般情况下, F 不唯一。下面给出了一个简单估计 F 的方法。

先验知识可以表达为一个约束算符

$$Cu \rightarrow u \quad (14.5-2)$$

例如, 对于多维信号 u , $C[u]$ 可以表示为

$$Cu[n] = \begin{cases} u[n] & 0 \leq u[n] \leq a \\ 0 & u[n] < 0 \\ a & u[n] > a \end{cases} \quad (14.5-3)$$

又例如, 若 $u(t)$ 是一个模拟信号, 信号带宽被系统限制在频率 f_c 以下。 $Cu(t)$ 可以写为

$$Cu(t) = \int_{-\infty}^{+\infty} u(\tau) \frac{\sin 2\pi f_c(t - \tau)}{\pi(t - \tau)} d\tau \quad (14.5-4)$$

u 的每个估计 u_k 应该被 C 约束。 Cu_k 可被理解为 u_{k+1} 的近似。得到的 Cu_k 可用一个畸变算符 D 进一步处理, 生成 DCu_k 。 DCu_k 可理解为 v 的近似估计, 输出误差 Δv 定义为

$$\Delta v = v - DCu_k \quad (14.5-5)$$

在 $(k+1)$ 次和 k 次迭代之间的恢复信号误差可以定义为

$$\Delta u = u_{k+1} - Cu_k \quad (14.5-6)$$

Δu 和 Δv 之间的关系是什么呢? 最简单的假设是它们成比例, 也就是说, $\Delta u = \lambda \Delta v$, 则

$$Fu_k = u_{k+1} = Cu_k + \Delta u = Cu_k + \lambda(v - DCu_k) \quad (14.5-7)$$

式 (14.5-7) 可进一步写为

$$Fu_k = \lambda v + Gu_k \quad (14.5-8)$$

式中

$$G = (I - \lambda D)C \quad (14.5-9)$$

I 是归一化算符。 u 初始近似可选为 $u_0 = \lambda v$ 。



如果 F 被选择为收缩的, 则 u_k 收敛于一个不动点 u^* 。

式 (14.5-8) 表示为

$$|Fu_k - Fu_n| = |Gu_k - Gu_n| \quad (14.5-10)$$

因此, 如果 F 是收缩的, 那么 G 也是收缩的。由 $u_0 = \lambda v$, 式 (14.5-8) 变为

$$Fu_k = u_{k+1} = u_0 + Gu_k \quad (14.5-11)$$

根据式 (14.5-11), 第 k 次迭代的过程为

(1) 对 u_k 应用约束算符 C , 得到 u'_k 。

(2) 对 u'_k 应用畸变算符 D , 得到 u''_k 。

(3) 由 $u_0 + \lambda(u_k - u''_k)$ 计算 u_{k+1} 。

下面给出上述方法应用于反卷积的一个例子。

14.6 约束迭代反卷积

给定 v 和 h , 反卷积即是当已知 v 是 u 和 h 的卷积时, 确定 u , 即

$$v = h * u \quad (14.6-1)$$

当迭代从 $u_0 = \lambda v$ 开始时, 式 (14.5-8) 可写为

$$u_{k+1} = \lambda v + (\delta - \lambda h) * Cu_k \quad (14.6-2)$$

式中, 约束算符暂且简化为归一化算符。

对式 (14.6-2) 进行傅里叶变换可得

$$U_{k+1} = \lambda V + U_k - \lambda H U_k \quad (14.6-3)$$

式中, 大写字母表示式 (14.6-2) 中对应小写字母变量的傅里叶变换。

式 (14.6-3) 是序号为 k 时 U 的一阶差分方程。它可以写为

$$U_{k+1} - (1 - \lambda H) U_k = \lambda V \quad (14.6-4)$$

其解为

$$U_k = \frac{V}{H} [1 - \lambda H]^{k+1} u[k] \quad (14.6-5)$$

式中, $u[k]$ 是一维单位步长序列。

当 $k \rightarrow \infty$ 时, 如果满足

$$|1 - \lambda H| < 1 \quad (14.6-6)$$

其解变为

$$U_\infty = \frac{V}{H} \quad (14.6-7)$$

式 (14.6-7) 与用逆滤波获得的结果相同。这种交互迭代的好处是在迭代有限的次数之后就可以得到一个比实际逆滤波好的结果。

由式 (14.6-6) 可知, 如果在一些频率点 H 等于 0, 程序便会终止。式 (14.6-6) 也可写为

$$\text{Re}[H] > 0 \quad (14.6-8)$$

为了避免在所有的频率上测试式 (14.6-8) 的有效性, 一个较好的办法是用反 h^* (一维离散信号用 $h^*[-m]$, 二维离散信号则用 $h^*[-m, -n]$) 在式 (14.6-1) 的两边卷积。



那么, 得到

$$v' = x * h' \quad (14.6-9)$$

式中, h' 是新的脉冲响应序列。

其传递函数为

$$H' = |H|^2 \quad (14.6-10)$$

v' 是 v 与 h^* 的卷积。

通过迭代方法求解式 (14.6-9) 时, 式 (14.6-8), 即 $\text{Re}[H'] > 0$, 自动使 $H \neq 0$ 在任何频率上都成立。

由于迭代解向逆滤波解收敛, 因此迭代过程可能得不到理想的结果。例如, 如果 $H \rightarrow 0$, 可能得不到唯一的结果。采用约束算符可以避免这个问题。例如, 在有限条件下, u 可假设为正值。在二维情况下, 为达到此目的的约束算符为

$$Cu_k[m, n] = \begin{cases} u_k[m, n] & M_1 \leq m < M_2, N_1 \leq n < N_2 \\ & \text{且 } u_k[m, n] \geq 0 \\ 0 & \text{其他} \end{cases} \quad (14.6-11)$$

令 $u'_k = Cu_k$, 在离散时间傅里叶变换域中, 式 (14.6-2) 变为

$$U'_{k+1} = \lambda V + U'_k - \lambda H U'_k \quad (14.6-12)$$

式 (14.6-12) 定义的迭代可以用离散傅里叶变换完成。然而, $u'_k = Cu_k$ 应在信号域进行。迭代的结构框图如图 14.4 所示。

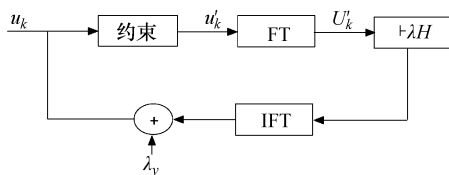


图 14.4 傅里叶变换约束迭代反卷积算法框图

在许多应用中, 脉冲响应是平移变化的。也可以在时域或空域应用上面讨论的迭代方法, 但是在变换域中, 上述讨论的代数方程不再有效。

例 14.2 一个二维高斯模糊脉冲响应序列为

$$h[m, n] = \exp\left[-\frac{m^2 + n^2}{100}\right]$$

假设图像序列为

$$u[m, n] = [\delta(m-24)\delta(n-32) + \delta(m-34)\delta(n-32)]$$

(a) 确定模糊图像 $v = u * h$, 并给出 h 和 v 的分布。

(b) 在正约束条件下用 $\lambda = 2.65$ 迭代计算反卷积图像。

解: (a) u 与 v 的卷积为

$$v[m, n] = \exp\left[-\frac{(m-24)^2 + (n-32)^2}{100}\right] + \exp\left[-\frac{(m-34)^2 + (n-32)^2}{100}\right]$$

h 和 v 的分布分别如图 14.5a 和 b 所示。

(b) 图 14.5c 表示 65 次迭代后的反卷积结果。

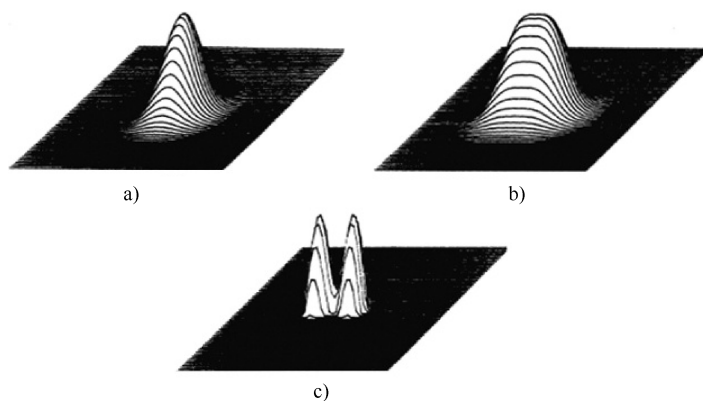


图 14.5 二维模糊信号的反卷积

a) 高斯模糊函数 b) 模糊函数 c) 通过反卷积恢复的信号 [Courtesy of Schafer, Mersereau, Richards]

14.7 投影方法

收缩的子集是一种投影。如果映射 Pu 满足下式，则算符 P 称为投影算符或投影。

$$|Pu - u| = \inf_{v \in S_0} |v - u| \quad (14.7-1)$$

式中， $u \in S$ ； S_0 是 S 的子集。

在信号恢复方面， S_0 是理想信号所归属的子集， $g = Pu$ 称为 u 在 S_0 上的投影。这可理解为投影算符 P 生成了一个接近 $u \in S$ 的向量 $g \in S_0$ 。如果 S_0 是一个凸集，那么 G 是唯一的。下面给出了凸集的简要回顾。

令 u_1 和 u_2 属于 S_0 ，如果对于所有的 $\lambda (0 \leq \lambda \leq 1)$ ，均有 $\lambda u_1 + (1 - \lambda)u_2$ 也属于 S_0 ，则 S_0 是凸集。图 14.6 给出了一个凸集和一个非凸集。

可见，如果 S_0 是凸集，那么投影 P 是收缩的 [Youla]。这就需要 S_0 是凸集。若满足这个条件，则上述的方法可称为凸集投影。我们将先考虑这样的情况。

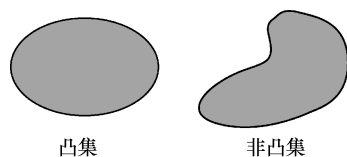


图 14.6 凸集和非凸集的例子

例 14.3 考虑所有的实非负函数 $u(x, y)$ 的集合 S_3 (如光场振幅)，其满足能量约束

$$|u|^2 = \iint |u(x, y)|^2 dx dy \leq E$$

证明 S_3 是凸集。

解：考虑 $u_1, u_2 \in S_3$ 。对于 $0 \leq \lambda \leq 1$ ，利用柯西-施瓦兹不等式

$$|\lambda u_1 + (1 - \lambda)u_2| \leq \lambda |u_1| + (1 - \lambda) |u_2| \leq \lambda E^{1/2} + (1 - \lambda) E^{1/2}$$

因此，集合 S_3 是凸的。

例 14.4 考虑 S 上的所有振幅在闭区间 $[a, b]$ 中的实函数集合 S_4 ， $a < b$ ，且 $a \geq 0, b > 0$ 。证明 S_4 是凸集。



解: 对于 $0 \leq \lambda \leq 1$, 当取到闭区间下界时, 有

$$u_3 = \lambda u_1 + (1 - \lambda) u_2 \geq \lambda a + (1 - \lambda) a = a$$

当取到闭区间上界时, 有

$$u_3 = \lambda u_1 + (1 - \lambda) u_2 \leq \lambda b + (1 - \lambda) b = b$$

因此, $u_3 \in S_0$, S_4 是凸集。

14.8 凸集投影方法 (POCS)

凸集投影方法已经成功地应用于信号和图像恢复。该方法认为, 一个未知信号 u 的已知性质可以认为是限制信号成为闭集 S_i 的元素的约束。对于 M 个已知性质, 便有 M 个这样的集合。对于属于 S_0 的信号, 这些集合的交集即为

$$S_0 = \bigcap_{i=1}^M S_i \quad (14.8-1)$$

信号恢复问题定义为将 u 的受损的版本 v 投影到集合 S_0 上。

令 P_i 表示集合 S_i 上的投影算符, 并定义算符 T_i 为

$$T_i = 1 + \lambda_i (P_i - 1) \quad (14.8-2)$$

如果 u_i^* 是在集合 S_i 上算符 P_i 的不动点, 则有

$$T_i u_i^* = u_i^* + \lambda_i (P_i u_i^* - u_i^*) = u_i^* \quad (14.8-3)$$

因此, u_i^* 也是 T_i 的不动点。同时可知, 如果 $0 < \lambda_i < 2$, 则 T_i 是收缩的。 λ_i 称为松弛参数。注意 $\lambda_i = 1$ 时 T_i 等于 P_i 。因此, T_i 是比 P_i 更广义的算符。选择合适的 λ_i 常常会加速迭代的收敛速度。

令 T 定义如下:

$$T = T_m T_{m-1} \cdots T_1 \quad (14.8-4)$$

则可得到下述结论:

(1) 定义迭代为

初值 u_0 是任意的,

则迭代 $u_{k+1} = T u_k$

在 S_0 上弱收敛于一个不动点。弱收敛的意思是 $\lim_{k \rightarrow \infty} |u_k - w| = (u^*, w)$, $\forall w \in S_0$, 其中 (a, b) 是向量 a 和 b 的内积。在一定条件下, 强收敛的意思是 $\lim_{k \rightarrow \infty} |u_k - u^*| = 0$ 。

(2) λ_i 在每次迭代时可以改变。 λ_{ik} 代表第 k 次迭代时 λ_i 的值。对于每一个选择的 λ_{ik} , 有 $0 < \lambda_{ik} < 2$, 迭代定义为

初值 u_0 是任意的,

$$u_{k+1} = T u_k$$

弱收敛到一个不动点 u^* 。

通常来说, 当 $M = 2$ 时, 采用凸集投影算法, 特别是一个投影在信号域, 另一个投影在变换域。这种情况如图 14.7 所示。

有两种可能的方法可用来估计最快收敛的

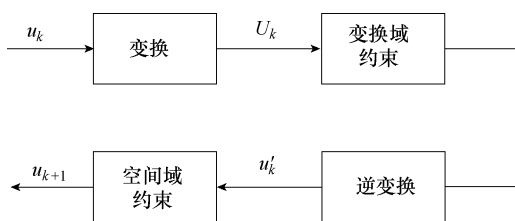


图 14.7 在变换域和信号域的凸集投影算法



λ_{1k} 和 λ_{2k} 的最优值。第一种情况下, λ_{1k} 首先被最优化, 然后采用 T_1 , 接着是 λ_{2k} 的最优化, 再采用 T_2 。一般情况下, 对线性投影算符有 $\lambda_{ik} = 1$, 否则取 $\lambda_{ik} \geq 0$ 。第二种情况下, 为了使 $T = T_2 T_1$ 有最小的误差, 可以同时优化 λ_{1k} 和 λ_{2k} [Stark]。

14.9 盖师贝格-帕普里斯算法 (GP 算法)

原始的盖师贝格-帕普里斯算法 (GP 算法) 是凸集投影算法的一个特殊形式, 在信号域和傅里叶域中分别用投影算符 P_1 和 P_2 表示 [Gerchberg]。因此

$$u_{k+1} = P_2 P_1 u_k \quad (14.9-1)$$

式中, P_1 和 P_2 的定义如下:

令 A 是 \mathbf{R}^2 上的一个区域, S_1 是 S 上所有函数的集合, 这些函数在 A 之外的区域都为 0 (特别地, 除了测度可能为 0 的点集)。 P_1 定义为

$$P_1 u = \begin{cases} u & (x, y) \in A \\ 0 & \text{其他} \end{cases} \quad (14.9-2)$$

令 S_2 是一些函数的集合, 在傅里叶平面上一闭区域 B 中, 这些函数的傅里叶变换为 $V(f_x, f_y)$ 。 P_2 定义为

$$P_2 u \leftrightarrow \begin{cases} U(f_x, f_y) & (f_x, f_y) \in B \\ V(f_x, f_y) & (f_x, f_y) \notin B \end{cases} \quad (14.9-3)$$

可知 S_1 和 S_2 都是凸集。

14.10 其他的凸集投影算法

在盖师贝格-帕普里斯算法 (GP) 中, 两个投影算符 P_1 和 P_2 分别对应于上述讨论的凸集 S_1 和 S_2 。这里, 将介绍另外三个利用凸集的投影算符。

考虑例 14.3 中的集合 S_3 。假设 u 是复数。它可写为

$$u = u_R + j u_I \quad (14.10-1)$$

式中, u_R , u_I 分别代表 u 的实部和虚部。

u_R^+ 定义为

$$u_R^+ = \begin{cases} u_R & u_R > 0 \\ 0 & \text{其他} \end{cases} \quad (14.10-2)$$

投影算符 P_3 定义为

$$P_3 u = \begin{cases} 0 & u_R < 0 \\ u_R^+ & u_R > 0, E_R^+ \leq E \\ \sqrt{E/E_R^+} u_R^+ & u_R > 0, E_R^+ > E \end{cases} \quad (14.10-3)$$

式中, E_R^+ 由下式给出:

$$E_R^+ = \iint [u_R^+(x, y)]^2 dx dy \quad (14.10-4)$$

且 E 是理想图像的估计或已知能量值。 E_R^+ 是有界的。



考虑例 14.4 中定义的集合 S_4 。投影算符 P_4 定义为

$$P_4 u = \begin{cases} a & u(x, y) < a \\ u(x, y) & a \leq u(x, y) \leq b \\ b & u(x, y) > b \end{cases} \quad (14.10-5)$$

最后, 考虑 S 上所有非负实函数的子集 S_5 。易知 S_5 是凸集。投影算符 P_5 定义为

$$P_{5u} = \begin{cases} u_R & u_R \geq 0 \\ 0 & \text{其他} \end{cases} \quad (14.10-6)$$

如前文, 式中 $u = u_R + ju_I$ 。

上面介绍的理论的应用实例将在后述章节部分给出。

14.11 从相位恢复信号

为了简化问题, 下面将讨论一维情况。假设信号 $u(x)$ 有一个紧支集 $[-a, a]$ 。已知 $u(x)$ 的傅里叶变换的相位为 $\phi(f)$, 需要估计信号的振幅。可以证明, $u(x)$ 可以被 $\phi(f)$ 完全确定的必要条件是存在关于 x_0 的对称点, 使得

$$u(x_0 + x) = \pm u(x_0 - x) \quad (14.11-1)$$

两个投影可写为

$$P_1 u = \begin{cases} u(x) & |u(x)| \leq a \\ 0 & \text{其他} \end{cases} \quad (14.11-2)$$

$$P_2 u \leftrightarrow \begin{cases} A(f) \cos[\phi(f) - \psi(f)] e^{j\phi(f)} & |\phi(f) - \psi(f)| \leq \frac{\pi}{2} \\ 0 & \text{其他} \end{cases} \quad (14.11-3)$$

式中, $A(f)$ 是傅里叶变换的振幅; $\psi(f)$ 是傅里叶变换的相位。

可以看到 $A(f) \cos[\phi(f) - \psi(f)]$ 是 $U(f)$ 在由 $\phi(f)$ 给定的方向上的投影。许多文献中, 没有考虑由 $\phi(f)$ 和 $\psi(f)$ 给定的方向, 将 P_2 表示为

$$P_2 u \leftrightarrow A(f) e^{j\phi(f)} \quad (14.11-4)$$

$\phi(f)$ 可能对于所有的 f 不都是已知的。假设 $\phi(f)$ 在集合 S 上对 f 已知, 那么 P_2 可以表示为

$$P_2 u \leftrightarrow \begin{cases} A(f) \cos[\phi(f) - \psi(f)] e^{j\phi(f)} & |\phi(f) - \psi(f)| < \frac{\pi}{2} \text{ 且 } f \in S \\ 0 & \text{其他, } f \in S \\ U(f) & f \notin S \end{cases} \quad (14.11-5)$$

式中, $U(f)$ 是 $u(t)$ 的傅里叶变换。

可以看到 P_1 是线性算符, 于是最优的 λ_{1k} 是 1。 λ_{2k} 的一个下界可以由下式得到 [Stark]:

$$\lambda_{2k} \geq \lambda_L = 1 + \frac{|P_2 u_k - P_1 P_2 u_k|^2}{|P_1 P_2 u_k - u_k|^2} \quad (14.11-6)$$

λ_L 可以作为 λ_{2k} 的一个估计。然而, λ_L 可能比 2 大。在随后的迭代中, 有必要限制 λ_{2k} 小于 2, 以保证迭代的收敛。例如, 可以选择 λ_{2k} 为 $\min(\lambda_L, 1.99)$ 。



因为 λ_{1k} 等于 1, 一个迭代周期可写为

$$\begin{aligned} u_{k+1} &= P_1 [1 + \lambda_{2k} (P_2 - 1)] u_k \\ &= (1 - \lambda_{2k}) u_k + \lambda_{2k} P_1 P_2 u_k \end{aligned} \quad (14.11-7)$$

式中, 用 u_k 代替了 $P_1 u_k$, 这是由于 $u_k = P_1 T_2 u_{k-1}$, 因此 $u_k \in D_1$ 。

14.12 用离散傅里叶变换从离散相位函数中重构信号

用 $u[n]$ 代表长度为 N 的离散信号。假设已知 $u[n]$ 的离散傅里叶变换的相位和 $u[n]$ 的支集。 $u[n]$ 的支集定义为

$$u[n] = 0 \quad \text{当 } n \text{ 在 } 0 \leq n < M \text{ 外, 且 } u[0] \neq 0$$

易知, 重构算法应该满足 $N \geq 2M$ [Haye et al., 1980]。 $u[n]$ 的离散傅里叶变换为

$$U[k] = A[k] e^{j\varphi[k]} \quad (14.12-1)$$

假设 $\varphi[k]$ 已知。

算法的程序流程如下:

- (1) 用 $A_1[k]$ 表示 $A[k]$ 初始的一个猜测值。然后, 初始的 $U_1[k]$ 为

$$U_1[k] = A_1[k] e^{j\varphi[k]} \quad (14.12-2)$$

- (2) 计算 $U_1[k]$ 的逆离散傅里叶变换, 生成第一个估计 $u_1[n]$ 。令 i 等于 1。

- (3) 定义一个新的序列 $v_i[n]$ 为

$$v_i[n] = \begin{cases} u_i[n] & 0 \leq n < M \\ 0 & M \leq n < N \end{cases} \quad (14.12-3)$$

- (4) 计算 $v_i[n]$ 的离散傅里叶变换 $V_i[k]$ 。 $V_i[k]$ 可以写为

$$V_i[k] = A_i[k] e^{j\varphi_i[k]} \quad (14.12-4)$$

程序剩余的流程部分可以分别采用以下两种不同的算法来进行。

算法 1 [Levi and Stark]

- (5) 将 $V_i[k]$ 进行如下投影:

$$U_{i+1}[k] = \begin{cases} A_i[k] \cos(\varphi[k] - \varphi_i[k]) e^{j\varphi[k]} & |\varphi[k] - \varphi_i[k]| < \frac{\pi}{2} \\ 0 & \text{其他} \end{cases} \quad (14.12-5)$$

这个运算对应于算符 P_2 。这个方程可以被进一步修改, 以便应用 T_2

$$U_{i+1}[k] = [1 + \lambda_{2i} (P_2 - 1)] [V_i[k]] \quad (14.12-6)$$

其中用关系式 (14.11-6) 计算 λ_{2i} 。

- (6) 通过计算 $U_{i+1}[k]$ 的逆离散傅里叶变换, 生成一个 $u_{n+1}[n]$ 的新的估计。

- (7) 重复步骤 (2) ~ (6), 直到收敛。

算法 2 [Haye et al.]

与算法 1 仅有如下步骤不同:

- (8) 投影 $V_i[k]$ 如下:

$$U_{i+1}[k] = A_i[k] e^{j\varphi[k]} \quad (14.12-7)$$

其他步骤与算法 1 相同。



例 14.5 定义第 $(i+1)$ 次的迭代误差为 E_i :

$$E_{i+1} = \sum_{n=0}^{N-1} |u[n] - u_{i+1}[n]|^2$$

式中, $u[n]$ 是需要再现的信号。

证明 E_i 不随着 i 增加而增加。

解: 根据帕塞瓦尔定理, 得到下式:

$$E_i = \frac{1}{N} \sum_{k=0}^{N-1} |U[k] - U_i[k]|^2$$

由于 $Z(k)$ 和 $Z_i(k)$ 的相位相同, 上式可以写为

$$E_{i+1} = \frac{1}{N} \sum_{k=0}^{N-1} |U[k] - U_{i+1}[k]|^2$$

同时定义

$$E'_i = \frac{1}{N} \sum_{k=0}^{N-1} |U[k] - V_i[k]|^2$$

利用向量差的三角方程, 可得到如下表达式:

$$E_{i+1} \leq E'_i$$

当 $U_{i+1}[k] = V_i[k]$ 时, 上式取等号。上式取不等号是因为 $U_{i+1}[k]$ 与 $U[k]$ 的相位相同, 而 $V_i[k]$ 的相位与它们不同。

再次对这个不等式用帕塞瓦尔定理, 得到

$$E'_i \leq \frac{1}{N} \sum_{k=0}^{N-1} |U[k] - V_i[k]|^2$$

因为 $v_i[n]$ 在生成 $u_{i+1}[n]$ 的投影之前, 由上面的方程表明, 投影过程减小了误差。

14.13 广义投影

有许多关于非凸集约束的信号恢复问题。例如, 数字化一个信号, 而且约束这个信号或其傅里叶变换到特定的幅值, 这就是一个非凸集投影。

若 S_i 是一个非凸集, 则有多于一个点满足投影算符 P_i 的定义。这个差异性可以通过附加额外的约束来消除。另一关于非凸集的问题是还没有发现这种投影存在性的证明。

在下面的讨论中, 投影 M 的数量将被限制为 2。这种情况下, 下面的误差度量是有用的:

$$E(u_k) = |P_1 u_k - u_k| + |P_2 u_k - u_k| \quad (14.13-1)$$

以下的定理证明了 $m=2$ 时广义投影的收敛性质 [Levi-Stark]。

定理: 由下式给出的递推式:

$$u_{k+1} = T_1 T_2 u_k, \quad u_0 \text{ 是任意值} \quad (14.13-2)$$

具有性质:

$$E(u_{k+1}) \leq E(u_k) \quad (14.13-3)$$

该性质对于每一个满足式 (14.3-4 ~ 14.3-8) 的 λ_1 和 λ_2

$$0 \leq \lambda_i \leq \frac{A_i^2 + A_i}{A_i^2 + A_i - \frac{1}{2}(A_i + B_i)} \quad (14.13-4)$$



$$A_1 = \frac{|P_1 T_2 u_k - T_2 u_k|}{|P_2 T_2 u_k - T_2 u_k|} \quad (14.13-5)$$

$$A_2 = \frac{|P_2 u_k - u_k|}{|P_1 u_k - u_k|} \quad (14.13-6)$$

$$B_1 = \frac{(P_2 T_2 u_k - T_2 u_k, P_1 T_2 u_k - T_2 u_k)}{|P_2 T_2 u_k - T_2 u_k|^2} \quad (14.13-7)$$

$$B_2 = \frac{(P_1 u_k - u_k, P_2 u_k - u_k)}{|P_1 u_k - u_k|^2} \quad (14.13-8)$$

可以看出式 (14.13-4) 的上限包含 1。因此, 定理用 P_1 、 P_2 代替了 T_1 、 T_2 。

投影 P_1 和 P_2 可包含多个约束。通常, 一个约束集在信号域, 而另一个约束集在变换域。

例 14.6 一个投影算符 P 映射一个函数 $u(x)$ 到函数 $g(x)$, 使得 $g(x)$ 满足

$$g(x) = \begin{cases} s(x) & x \in D \subset D_0 \\ 0 & x \notin D \end{cases}$$

式中, $s(x)$ 包含所有 x 。试确定 P 。

解: P 满足

$$|Pu - u| = \inf_v |v - u|$$

因为 $g(x) = Pu(x)$, 有

$$\begin{aligned} |Pu - u| = |g - u| &= \int_{x \notin D_0} |g(x) - u(x)|^2 dx + \int_{x \in D_0 \cap D^c} |g(x) - u(x)|^2 dx + \\ &\int_{x \in D, s(x) > 0} |g(x) - u(x)|^2 dx + \int_{x \in D, s(x) < 0} |g(x) - u(x)|^2 dx \end{aligned}$$

当选择 $g(x)$ 为如下形式时, $|g - u|$ 为最小

$$P_1 u = \begin{cases} 0 & x \notin D_0 \\ u(x) & x \in D_0 \cap D^c \\ s(x) & x \in D, x(t) > 0 \\ -s(x) & x \in D, x(t) \leq 0 \end{cases}$$

14.14 从幅值恢复信号

这个问题也称为相位恢复 (phase retrieval)。这种问题出现在许多领域中, 如天文、光学和 X 射线晶体学等。在这些领域, 在每一个测量点仅强度 (振幅的二次方) 是可以得到的, 而相位是待确定的。

相位恢复比从相位恢复信息更困难。在二维情况下, 所有的方程 $u(x, y)$ 、 $-u(x, y)$ 、 $u(x - x_0, y - y_0)$ 和 $u(-x, -y)$ 具有相同的傅里叶变换振幅。除了这四种情况, 如果一个时



间或空间有限序列的 z 变换不可约 (不能写为 z 或 z^{-1} 的一阶多项式形式), 那么该序列能够从它的傅里叶幅值唯一确定 [Hayes et al.]。这个结论对于一维的情况是没有意义的, 因为根据函数代数定理, 所有的多项式都可以分解为一阶多项式形式。然而, 在二维情况下, 很多多项式是不可约的。因此, 从傅里叶幅值信息恢复图像是可行的。

设已知的傅里叶变换幅值是 $|U(f)|$ 。这两个投影可以写为

$$P_1 u = \begin{cases} u & |u| < a \\ 0 & \text{其他} \end{cases} \quad (14.14-1)$$

$$P_2 u \leftrightarrow |U(f)| e^{j\phi(f)} \quad (14.14-2)$$

式中, $\phi(f)$ 是当前已知的傅里叶相位。

可知 P_1 是凸的, 而 P_2 不是。 P_1 也可以包括其他的约束。

广义恢复算法可以写为

$$u_{k+1} + T_1 T_2 u_k, u_0 \text{ 是任意的} \quad (14.14-3)$$

式中, $T_i = 1 + \lambda_i(P_i - 1)$, 而且 T_1 可以选择等于 P_1 , 因为 P_1 是线性的。

当 λ_i 都选择等于 1, 这个算法即为盖师贝格-撒克斯通 (Gerchberg-Saxton) 算法, 正如 14.9 节所讨论的。

当 $u_k = P_1 T_2 u_{k-1}$ 时, u_k 属于投影算符为 P_1 的空间 D_1 。因此, $P_1 u_k = u_k$, 而且式 (14.13-1) 的 $E(u_k)$ 化简为

$$E(u_k) = |P_2 u_{k+1} - u_{k+1}| \quad (14.14-4)$$

式 (14.13-2) 也化简为

$$u_{k+1} = (1 - \lambda_{2k}) u_k + \lambda_{2k} P_1 P_2 u_k \quad (14.14-5)$$

计算这个方程的傅里叶变换, 用算符 $F(\cdot)$ 表示傅里叶变换, 可得

$$U_{k+1} = (1 - \lambda_{2k}) U_k + \lambda_{2k} F(P_1 P_2 u_k) \quad (14.14-6)$$

根据帕塞瓦尔定理, 式 (14.14-4) 可以写为

$$E(u_{k+1}) = \int_{-\infty}^{+\infty} |U(f) - U_{k+1}|^2 df \quad (14.14-7)$$

为了使 $E(u_{k+1})$ 最小化, λ_{2k} 可以通过在范围 $0 < \lambda_{2k} < a$ 内进行简单的搜索来估计 (a 通常是数字, 如 3)。

14.14.1 陷阱和隧道

当至少有一个非凸的映射时, 则可能在这样的点发生收敛, 该点不是每一个单一 T_i 的一个不动点。这样的点 u_m 称为一个陷阱。 x_i 不能满足一个或多个先验约束, 但是却满足下式:

$$u_m = T_1 T_2 u_m \quad (14.14-8)$$

凸集不存在陷阱。另一方面, 隧道可能发生在箱凸集 (box convex) 和非凸集上。当 u_k 从一次迭代到下一次迭代的变化很小的时候, 会产生隧道。

陷阱可以在从一次迭代到下一次迭代 $E(u_k)$ 不变 ($E(u_k) > 0$) 时检测到。如果 P_1 是线性的且 $P_1 T_2 u_k = u_k$, 可知正确的解 u^* 位于垂直于 $P_2 u_k - u_k$ 的超平面内 [Stark]。



14.15 用最小二乘法和广义逆法恢复图像

这里讨论的方法是关于迭代的优化的。另一个图像恢复的有效途径是利用最小二乘法。对于这种方法，测量的图像为

$$\mathbf{v} = \mathbf{H}\mathbf{u} + \mathbf{n} \quad (14.15-1)$$

式中， \mathbf{H} 是一个 $N_1 \times N_2$ 的矩阵； \mathbf{u} 是期望的图像； \mathbf{n} 是噪声图像。

在式 (14.15-1) 中，图像按行或列的次序排列。

\mathbf{u} 的无约束最小二乘估计可以通过最小化下式来获取：

$$E = \|\mathbf{v} - \mathbf{H}\mathbf{u}\|^2 = (\mathbf{v} - \mathbf{H}\mathbf{u})^{\mathrm{t}*} (\mathbf{v} - \mathbf{H}\mathbf{u}) \quad (14.15-2)$$

式中， $\|\cdot\|^2$ 表示 L^2 范数。(注：式中的 t 为转置符号， $*$ 号为共轭符号， t_* 为先转置再共轭)。

计算 E 关于 \mathbf{u} 的偏导数，令其等于 0，结果为

$$\mathbf{H}^{\mathrm{t}}\mathbf{v} = \mathbf{H}^{\mathrm{t}}\mathbf{H}\mathbf{u} \quad (14.15-3)$$

如果大小为 $N_2 \times N_2$ 的 $\mathbf{H}^{\mathrm{t}}\mathbf{H}$ 是非奇异的，则最小二乘解为

$$\hat{\mathbf{u}} = \mathbf{H}^+ \mathbf{v} \quad (14.15-4)$$

式中， \mathbf{H}^+ 为 \mathbf{H} 的广义逆（伪逆），其表示为

$$\mathbf{H}^+ = (\mathbf{H}^{\mathrm{t}}\mathbf{H})^{-1} \mathbf{H}^{\mathrm{t}} \quad (14.15-5)$$

\mathbf{H}^+ 的大小为 $N_2 \times N_1$ 。 \mathbf{H}^+ 满足

$$\mathbf{H}^+ \mathbf{H} = \mathbf{I} \quad (14.15-6)$$

然而， $\mathbf{H}^+ \mathbf{H}$ 是不等于 \mathbf{I} 的。 $\mathbf{H}^+ \mathbf{H}$ 非奇异的一个必要条件是 $N_1 \geq N_2$ ，且 \mathbf{H} 的秩 r 是 N_2 。

如果 $N_1 < N_2$ ，且 \mathbf{H} 的秩是 N_1 ，则 \mathbf{H}^+ 定义为

$$\mathbf{H}\mathbf{H}^+ = \mathbf{I} \quad (14.15-7)$$

满足式 (14.15-7) 的 \mathbf{H}^+ 不是唯一的。唯一性是通过约束式 (14.15-4)，使其有最小范数而获得的。换言之，在所有可能的解中，选择了最小的 $\|\hat{\mathbf{u}}\|^2$ 的解作为最优解。于是伪逆为

$$\mathbf{H}^+ = \mathbf{H}^{\mathrm{t}}(\mathbf{H}\mathbf{H}^{\mathrm{t}})^{-1} \quad (14.15-8)$$

总之，如上讨论的， \mathbf{H}^+ 通常是唯一存在的，而且最小二乘解为 $\mathbf{H}^+ \mathbf{v}$ 。

\mathbf{H}^+ 还有如下的一些性质：

- (1) $\mathbf{H}\mathbf{H}^+ = (\mathbf{H}\mathbf{H}^+)^{\mathrm{t}}$ ，换言之， $\mathbf{H}\mathbf{H}^+$ 是对称的。
- (2) $\mathbf{H}^+ \mathbf{H} = (\mathbf{H}^+ \mathbf{H})^{\mathrm{t}}$ ，换言之， $\mathbf{H}^+ \mathbf{H}$ 是对称的。
- (3) $\mathbf{H}\mathbf{H}^+ \mathbf{H} = \mathbf{H}$
- (4) $\mathbf{H}^+ \mathbf{H}\mathbf{H}^+ = \mathbf{H}^+$

14.16 通过奇异值分解计算 \mathbf{H}^+

矩阵 \mathbf{H} 的大小为 $N_1 \times N_2$ ，秩为 r ，它的奇异值分解（Singular Value Decomposition, SVD）可表示为



$$\mathbf{H} = \sum_{m=1}^r \lambda_m^{\frac{1}{2}} \boldsymbol{\psi}_m \boldsymbol{\phi}_m^{\mathrm{t}} \quad (14.16-1)$$

式中, $\boldsymbol{\psi}_m$ 和 $\boldsymbol{\phi}_m$ 分别是 $\mathbf{H}\mathbf{H}^{\mathrm{t}}$ 和 $\mathbf{H}^{\mathrm{t}}\mathbf{H}$ 的特征向量; λ_m 是奇异值, $1 \leq m \leq r$ 。

伪逆 \mathbf{H}^+ 的奇异值分解形式为

$$\mathbf{H}^+ = \sum_{m=1}^r \lambda_m^{-\frac{1}{2}} \boldsymbol{\phi}_m \boldsymbol{\psi}_m^{\mathrm{t}} \quad (14.16-2)$$

将式 (14.16-2) 代入式 (14.15-4), 伪逆解可以写为

$$\hat{\mathbf{u}} = \sum_{m=1}^r \lambda_m^{-\frac{1}{2}} \boldsymbol{\phi}_m \boldsymbol{\psi}_m^{\mathrm{t}} \mathbf{v} \quad (14.16-3)$$

式 (14.16-3) 与下式相同:

$$\hat{\mathbf{u}} = \sum_{m=1}^r \boldsymbol{\omega}_m \boldsymbol{\phi}_m \quad (14.16-4)$$

式中

$$\boldsymbol{\omega}_m = \frac{\boldsymbol{\psi}_m^{\mathrm{t}} \mathbf{v}}{\lambda_m^{1/2}} \quad (14.16-5)$$

式 (14.16-4) 表明: $\hat{\mathbf{u}}$ 处在由 $\mathbf{H}^{\mathrm{t}}\mathbf{H}$ 的特征值张成的向量空间中。

对于小尺度矩阵问题, 用式 (14.16-4) 和式 (14.16-5) 相对简单。但是, 对于大尺度矩阵问题, 这样做会变得很困难。 $N_1 = N_2 = 256$ 时, \mathbf{H} 的大小是 $65\,536 \times 65\,536$ 。这种情况下采用奇异值分解变得不再可行。

退化的点扩散函数 (Point Spread Function, PSF) 是可分离的, $\mathbf{v} = \mathbf{H}_1 \mathbf{u} \mathbf{H}_2$, 那么广义逆也是可分离的, 且

$$\hat{\mathbf{u}} = \mathbf{H}_1^+ \mathbf{v} \mathbf{H}_2^+ \quad (14.16-6)$$

例 14.7 用伪逆求下式的最小二乘解:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

解: \mathbf{H} 和 $\mathbf{A} = \mathbf{H}^{\mathrm{t}}\mathbf{H}$ 为

$$\mathbf{H} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 6 & 7 \\ 7 & 14 \end{pmatrix}$$

\mathbf{A} 的特征值为 $\lambda_1 = 10 + \sqrt{65}$, $\lambda_2 = 10 - \sqrt{65}$ 。因为 \mathbf{A} 是非奇异的, \mathbf{H}^+ 为

$$\mathbf{H}^+ = \mathbf{A}^{-1} \mathbf{H}^{\mathrm{t}} = \frac{1}{35} \begin{pmatrix} 14 & -7 \\ -7 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix} = \frac{1}{35} \begin{pmatrix} 0 & 21 & -7 \\ 5 & -8 & 11 \end{pmatrix}$$

最小二乘解为

$$\hat{\mathbf{u}} = \mathbf{H}^+ \boldsymbol{\gamma} = \begin{pmatrix} 0.6 \\ 0.628 \end{pmatrix}$$

例 14.8 假设在信号域上, 解受到某种程度的约束, 而且约束以矩阵形式作用于 $\hat{\mathbf{u}}$ 。例如, 这样的算符可能在一定的区域内限制图像非零。在这种情况下, 求其最小二乘解。



解：最小化误差量度可以写为

$$E = |\mathbf{v} - \mathbf{HTu}|^2$$

式中， \mathbf{T} 是矩阵形式的约束算符。

在有界情况下， \mathbf{T} 是对角矩阵，且其对角元素为 1 或 0。得到

$$\begin{aligned} E &= (\mathbf{v} - \mathbf{HTu})^T (\mathbf{v} - \mathbf{HTu}) \\ &= \mathbf{v}^T \mathbf{v} - 2\mathbf{u}^T \mathbf{T}^T \mathbf{H}^T \mathbf{v} + \mathbf{u}^T \mathbf{T}^T \mathbf{H}^T \mathbf{HTu} \end{aligned}$$

E 的梯度为

$$\mathbf{g} = -2\mathbf{T}^T \mathbf{H}^T \mathbf{v} + 2\mathbf{T}^T \mathbf{H}^T \mathbf{HTu}$$

令 \mathbf{g} 等于 0，假定矩阵的逆存在，得到

$$\hat{\mathbf{u}} = [\mathbf{T}^T \mathbf{H}^T \mathbf{HT}]^{-1} \mathbf{T}^T \mathbf{H}^T \mathbf{v}$$

14.17 最速下降法

当为了获取 $\hat{\mathbf{u}}$ 而不是 \mathbf{H}^+ 时，可以用迭代梯度法。这个方法基于以下事实： $E = |\mathbf{v} - \mathbf{Hu}|^2$ 的最速下降方向为 E 对 \mathbf{u} 的负梯度方向。

假设 \mathbf{u}_k 是向量 \mathbf{u} 的第 k 次迭代。 E 对 \mathbf{u}_k 的梯度是

$$\mathbf{g}_k = -2\mathbf{H}^T (\mathbf{v} - \mathbf{Hu}_k) \quad (14.17-1)$$

迭代梯度法由下式给出：

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \alpha_k \mathbf{g}_k, \quad \mathbf{u}_0 = 0 \quad (14.17-2)$$

式中， α_k 是标量。

把式 (14.17-2) 代入式 (14.17-1)， \mathbf{g}_k 可写为

$$\mathbf{g}_k = \mathbf{g}_{k-1} - \alpha_{k-1} \mathbf{H}^T \mathbf{H} \mathbf{g}_{k-1} \quad (14.17-3)$$

可以看出，如果满足下式，则 \mathbf{u}_k 收敛于 $\hat{\mathbf{u}}$ ：

$$0 \leq \alpha_k \leq \frac{2}{\lambda_{\max}} \quad (14.17-4)$$

式中， λ_{\max} 是 $\mathbf{A} = \mathbf{H}^T \mathbf{H}$ 的最大特征值。

如果 α_k 为常数，这个方法称为一步梯度法。最快收敛的最优 α 值为

$$\alpha_{\text{最优}} \leq \frac{2}{\lambda_{\max} + \lambda_{\min}} \quad (14.17-5)$$

式中， λ_{\min} 是 \mathbf{A} 的最小特征值。

如果 \mathbf{A} 是高度病态的， \mathbf{A} 的条件数 $(\lambda_{\max}/\lambda_{\min})$ 会很大，而且收敛得很慢。

如果 α_k 在每一次迭代中都进行最优化，则这个方法称为最速下降法。 α_k 的最优值为

$$\alpha_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_k^T \mathbf{A} \mathbf{g}_k} \quad (14.17-6)$$

当 $\lambda_{\max}/\lambda_{\min}$ 很大时，即使使用 α_k ，收敛可能还是很慢。

例 14.9 用两种方法求解例 14.7 的近似最小二乘解：(a) 用 12 次迭代的一步梯度算法；



(b) 用 3 次迭代的最速下降法。

解: (a) 在一步梯度法中, $\alpha_{\text{最优}}$ 为

$$\alpha_{\text{最优}} = \frac{2}{\lambda_1 + \lambda_2} = 0.1$$

在第一次迭代中, 有

$$\mathbf{g}_0 = -\begin{pmatrix} -8 \\ -13 \end{pmatrix}, \mathbf{u}_1 = \begin{pmatrix} 0.8 \\ 1.3 \end{pmatrix}$$

误差 $E_1 = \|\mathbf{v} - \mathbf{H}\mathbf{u}_1\|^2$ 等于 9.46。

这样迭代 12 次之后, 有

$$\mathbf{u}_{11} = \begin{pmatrix} 0.685 \\ 1.253 \end{pmatrix}, \mathbf{u}_{12} = \begin{pmatrix} 0.555 \\ 0.581 \end{pmatrix}$$

误差为 $E_{12} = \|\mathbf{v} - \mathbf{H}\mathbf{u}_{12}\|^2 = 1.101$ 。

(b) 用最速下降法, 通过式 (14.17-6) 在每一次迭代时优化 α_k 。3 次迭代后的结果为

$$\mathbf{u}_3 = \begin{pmatrix} 0.5592 \\ 0.629 \end{pmatrix}$$

其误差 $E_3 = \|\mathbf{v} - \mathbf{H}\mathbf{u}_3\|^2 = 1.0285$ 。

14.18 共轭梯度法

在共轭梯度法中, 标量 α_k 和校正向量在每次迭代时均被优化。这样可以更快地得到收敛。为了这个目的, 将式 (14.17-2) 替换为

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \alpha_k \mathbf{r}_k \quad (14.18-1)$$

式中, 校正向量 \mathbf{r}_k 代替了梯度向量 \mathbf{g}_k 。

校正向量 \mathbf{r}_k 有如下性质:

$$\mathbf{g}_{k+1}^T \mathbf{r}_k = 0 \quad (14.18-2)$$

$$[\mathbf{H}\mathbf{r}_k]^T [\mathbf{H}\mathbf{r}_m] = \mathbf{r}_k^T \mathbf{A} \mathbf{r}_m = 0, \quad k \neq m, \quad 0 \leq k, m < N \quad (14.18-3)$$

式中, N 是 \mathbf{A} 的秩。

因为这些性质, 算法在 N 次迭代后收敛。迭代方程除了式 (14.18-1) 外, 还有如下方程:

$$\mathbf{r}_k = -\mathbf{g}_k + \beta_k \mathbf{r}_{k-1}, \quad \mathbf{r}_0 = -\mathbf{g}_0 \quad (14.18-4)$$

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{A} \mathbf{r}_{k-1}}{\mathbf{r}_{k-1}^T \mathbf{A} \mathbf{r}_{k-1}} \quad (14.18-5)$$

$$\alpha_k = \frac{\mathbf{g}_k^T \mathbf{r}_k}{\mathbf{r}_k^T \mathbf{A} \mathbf{r}_k} \quad (14.18-6)$$

$$\mathbf{g}_k = -\mathbf{H}^T \mathbf{v} + \mathbf{A} \mathbf{u}_k = \mathbf{g}_{k-1} + \alpha_{k-1} \mathbf{A} \mathbf{r}_{k-1} \quad (14.18-7)$$

最终的解 $\hat{\mathbf{u}}$ 可写为

$$\hat{\mathbf{u}} = \sum_{m=0}^{N-1} \alpha_m \mathbf{r}_m \quad (14.18-8)$$



当 N 是一个大数时，算法能在 N 次迭代之前结束。如果 \mathbf{H} 的秩 $M < N$ ，则算法迭代 M 次即可收敛。

例 14.10 用共轭梯度法解例 14.7 的最小二乘问题。

解：最小二乘解满足

$$\mathbf{T}^t \mathbf{H}^t \mathbf{H} \mathbf{T} \mathbf{u} = \mathbf{T}^t \mathbf{H}^t \mathbf{v}$$

这与之前用 $\mathbf{H}' = \mathbf{H} \mathbf{T}$ 代替 \mathbf{H} 具有相同的形式。唯一需要改变的迭代方程是梯度方程，变为

$$\mathbf{g}_k = \mathbf{g}_{k-1} - \alpha_{k-1} \mathbf{T}^t \mathbf{H}^t \mathbf{H} \mathbf{r}_{k-1}$$

通过选择满足约束 \mathbf{T} 的初始图像 \mathbf{u}_0

$$\mathbf{T} \mathbf{u}_0 = \mathbf{u}_0$$

\mathbf{g}_0 和 \mathbf{r}_0 为

$$\mathbf{g}_0 = \mathbf{r}_0 = \mathbf{T}^t \mathbf{H}^t (\mathbf{v} - \mathbf{H} \mathbf{u}_0)$$

第 15 章 衍射光学 I

15.1 本章引论

衍射光学包括利用计算机计算生成全息图，并由记录系统实际实现，之后通常会进行光还原。记录系统有激光胶片打印机、扫描电子束打印机和高分辨率打印机等。由于任何类型的波都可以在计算机中计算，因此衍射光学的应用非常广泛。用这种技术生成的数字全息图一般称为衍射光学元件（Diffractive Optical Element, DOE）。DOE 又可称为计算机生成全息图（或计算全息图，Computer-Generated Hologram, CGH）。由于这些术语等价，因此本章中涉及全息图、DOE 和 CGH 时其含义是相同的。

众所周知，在通信工程中，复信号在经过单一信道传输之前需进行某种形式的编码。可以通过用复信号的幅度调制载波的幅度或用复信号的相位调制载波的相位来实现。被调制的载波的实部能被传输，通过适当的解调过程，复信号在接收端可被恢复。全息术和衍射光学均基于这样的调制与解调原理。

DOE 在波形整形、激光加工、3D 显示 [Yatagai, 1975]、光学模式识别、光互联、安全设备 [Dittman, et al., 2001]、光通信/网络、光谱学等许多领域都得到了应用。DOE 的一些重要优势如下：

- 一个 DOE 能完成不止一项功能，如在一个单一元件上拥有对应多个透镜的多个焦点。它也能被设计成适用于多波长。
- 与折射和反射光学元件相比，DOE 通常更轻薄，体积更小。
- 对于给定的任务，在大规模制造中，它们制造成本更低。

由于 DOE 基于衍射，因此它们高度色散，即其性质依赖于波长。因此它们更常用于单色光。然而，色散性质也能被用作有利的一面。如某一波长的光会聚于空间中的点，另一波长的光会聚于空间中的另一点。在用于光通信和网络的光学相控阵这样的器件中，色散可以用来实现不同焦点处不同波长光的分离。此外，折射和衍射光学元件联合起来能消除波长色散，也可以用来实现消除球差等目的。

生成 DOE 需要几个步骤，包括波采样、波传播的计算（通常用 FFT）、在全息图上对复波前信息编码，如全息平面上的一幅正实值图像，以及最终 DOE 的记录。

在 3-D 波前编码中，通常相位的正确表示比幅值更重要。因此，相位受到更多的关注。在输出图像强度的应用中，将输出图像与一随机相位因子（扩散器）相乘能显著降低输出振幅的动态范围，而并不改变图像的强度 [Burekhardt, 1970]。

本章包含 11 节内容。首先介绍的几种方法是利用透镜的傅里叶变换性质生成傅里叶变换全息图的 DOE。首个此类方法是 15.2 节讨论的罗曼（Lohmann）法。该方法中，通过控制孔径的尺寸实现幅值调制，通过控制孔径的位置实现相位调制。15.3 节讨论基于罗曼法实现 DOE 中所涉及的近似，这些近似可对第 16.6 节讨论的 Lohmann-ODIFFT 法进行补偿。



15.4 节通过选用固定尺寸的孔径来简化罗曼法，这意味着幅度是固定的；并用第 14 章中讨论的迭代优化法补偿产生的误差。15.6 节描述了采用罗曼法进行的计算机实验。15.7 节讨论了另一基于硬限幅全息传输函数生成二值全息图的傅里叶方法。

之前讨论的方法假定重构图像在一个输出平面上。15.8 节中讨论的方法能在三维空间中的任意位置生成输出像点。该方法使用简单。它是首次用来展示如何用扫描电子显微镜系统制作 DOE 的方法。15.9 节扩展了 15.8 节的方法，使之能在三维空间中生成许多物点。

所有 DOE 方法本质上都涉及幅度和相位的非线性编码。因此，这将产生谐波图像。使用 FFT 时，隐含的假设是物体是周期的。这也会造成图像在重构平面以周期性方式重复。15.10 节介绍用半不规则采样和原点靠近全息平面的球面参考波来生成仅有一个图像的单幅图像全息法。该方法在第 19 章中将会用于密集波分复用相控阵的设计，以便在光通信和网络中得到比现有相控阵设计方法更多的波段通道。

最后一节介绍一种经典的 DOE，其具有平面透镜功能，称作二元菲涅耳波带片。

15.2 罗曼法

罗曼法与其他许多 DOE 方法是为实现如图 15.1 所示的傅里叶变换而设计的。

罗曼法（也称作迂回相位法），全息平面被分成小长方形，每个长方形包含一个孔径 [Lomann, 1970]。图 15.2 给出了罗曼单元的一个示例。孔径的尺寸用来控制幅值，而孔径位置的变化用来调整相位，这将形成一种二元透射条纹。

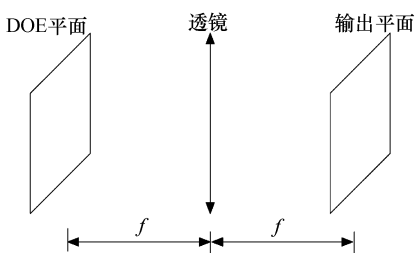


图 15.1 应用于 DOE 设计的傅里叶变换系统

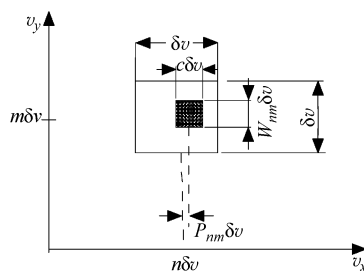


图 15.2 罗曼全息图的第 (n, m) 阶单元

令 $h(x, y)$ 为全息图 $H(v_x, v_y)$ 的输出幅值，它应与理想图像 $f(x, y)$ 成比例。全息图的二元透过率函数可写为

$$H(v_x, v_y) = \sum_n \sum_m \text{rect}\left[\frac{v_x - (n + P_{nm})\delta v}{c\delta v}\right] \text{rect}\left[\frac{v_y - m\delta v}{W_{nm}\delta v}\right] \quad (15.2-1)$$

当倾斜平面波 $\exp(2\pi j x_0 v_x)$ 入射到二元全息图上时，全息图后的复幅值为 $H(v_x, v_y) \exp(2\pi j x_0 v_x)$ 。像平面的复幅值为 $H(v_x, v_y) \exp(2\pi j x_0 v_x)$ 的傅里叶变换

$$\begin{aligned} & \iint H(v_x, v_y) e^{2\pi j [(x+x_0)v_x + yv_y]} dv_x dv_y \\ &= c(\delta v)^2 \text{sinc}[c\delta v(x+x_0)] \sum_n \sum_m W_{nm} \text{sinc}(yW_{nm}\delta v) \cdot \exp\{2\pi j [\delta v((x+x_0)(n+P_{nm}) + ym)]\} \end{aligned} \quad (15.2-2)$$



选择参量 W_{nm} 、 P_{nm} 和常量 x_0 、 c 以使得像平面复幅值与理想图像 $f(x, y)$ 匹配。

为了便于式 (15.2-2) 与目标图像 $f(x, y)$ 比较, $f(x, y)$ 写成如下形式:

$$f(x, y) = \iint F(v_x, v_y) e^{2\pi j(xv_x + yv_y)} dv_x dv_y = \sum_n \sum_m F(n\delta v, m\delta v) e^{2\pi j[\delta v(xn + ym)]} \quad (15.2-3)$$

两个 sinc 函数与式 (15.2-2) 中因子 $\exp[2\pi j(P_{nm}\delta v)]$ 可认为接近 1。这种假设的有效性将在 15.3 节中讨论。令式 (15.2-2) 与式 (15.2-3) 中的傅里叶系数相等, 则有

$$\begin{aligned} c(\delta v)^2 W_{nm} \exp\{2\pi j[x_0\delta v(n + P_{nm})]\} &\propto F(n\delta v, m\delta v) \\ F(n\delta v, m\delta v) &\propto c(\delta v)^2 A_{nm} \exp[2\pi j(\phi_{nm}/2\pi)] \\ W_{nm} &\approx A_{nm}; P_{nm} + n \approx \phi_{nm}/2\pi x_0\delta v \end{aligned} \quad (15.2-4)$$

选取 $x_0\delta v$ 等于整数 M , 得到

$$P_{nm} \approx \phi_{nm}/2\pi M \quad (15.2-5)$$

式 (15.2-4) 和式 (15.2-5) 表明, 每个单元中的孔径的高度 W 和位置 P , 可选择为分别正比于该单元上的复幅值 F 的幅值 A 和相位 ϕ 。

15.3 罗曼法中的近似

上一节中, 为简单起见作了三个近似: (a) $\text{sinc}[c\delta v(x + x_0)] \approx$ 常量, (b) $\text{sinc}(yW_{nm}\delta v) \approx 1$, (c) $\exp[2\pi j(xP_{nm}\delta v)] \approx 1$ 。这些近似对重构图像的影响取决于几个因素, 并且通过适当设计能够使影响最小化 [Lohmann, 1970]。

sinc 函数 $\text{sinc}[c\delta v(x + x_0)]$ 造成 x 方向上强度的陡降, 其与像平面中心距离成比例。近似 (a) 认为在图像区域内 sinc 因子近似为常数。如果图像区域尺寸为 $\Delta x\Delta y$, 则在影响最严重的边缘 $x = \pm(\Delta x/2)$ 有 $\text{sinc}[cM \pm c/2]$ 。这表明孔径尺寸 c 越小, 造成强度上的陡降越小。然而, 这也会造成图像亮度降低。当 $cM = 1/2$ 时, 中心与边缘的强度之比为 9 : 1。当 $cM = 1/3$ 时, 强度之比降到 2 : 1。因此, 牺牲强度能换取重构图像边缘强度陡降的减小。

sinc 函数 $\text{sinc}(yW_{nm}\delta v) \approx 1$ 表明在 y 方向存在与 (a) 类似的强度陡降。由于 $\Delta x\delta v = 1$, 则 $|yW\delta v| < 1/2$, 该近似更合理。众所周知, 相干成像系统孔径的幅度误差对图像几乎无影响, 因为它们不会像相位误差一样使光线偏离。为减小该近似的影响, 每个 W 可简化为一常量, 然而这样会对强度造成一定的影响。

解决 x 方向的 sinc 函数陡降的一种可行方法, 是目标图像除以 $\text{sinc}[c\delta v(x + x_0)]$, 目标图像 $f(x, y)$ 变为

$$\frac{f(x, y)}{\text{sinc}[c\delta v(x + x_0)]}$$

对 y 方向不能进行同样的处理, 因为 $\text{sinc}(yW_{nm}\delta v)$ 由待定的孔径参数 W_{nm} 决定。幸好, 这个 sinc 因子较 x 方向的 sinc 影响小, 通过调整全息图的设计能减小这种影响。

相移 $\exp[2\pi j(xP_{nm}\delta v)]$ 在像平面产生随 x 变化的相位误差。误差的范围取决于 x 和 P 。由于 $|x| \leq \Delta x/2 = 1/2\delta v$, $|P| \leq 1/2M$, 相位误差范围为 0 至 $\pi/2M$ 。当相位误差取最大值时, 相位误差对应的光程为 $\lambda/4M$ 。当 $M = 1$ 时, 这个误差处于波相差的 $\lambda/4$ 瑞利准则内。

当成像区域受限时, 这些近似的不良影响会小些。重构误差随偏离像平面中心的距离增长而增大。重构图像区域越小, 误差也越小。



15.4 常幅值罗曼法

从上节可知,由孔径尺寸引起的 sinc 振荡很难做近似处理。因此,需要使每个孔径的高度变为常量,以有利于更简单的实现。这将使目标图像能像在 x 方向上一样,在 y 方向除与 y 相关的 sinc 陡降,因为 sinc 因子与孔径的常量宽度有关。

逻辑上,如果每个孔径大小一致,则输出仅受孔径的位置影响。这意味着将所有的信息包含到了相位中。下面讨论的方法用于将全息平面的信息从幅值中转至相位中。

如仅考虑目标图像的幅值,则观察平面中每一采样点的相位为一自由变量。那么,可用第 14 章中的迭代方法来减小高度 W_{nm} 的取值范围。假设已采样的理想图像幅值为 a_{nm} , 相应的不确定相位为 θ_{nm} 。图像的离散傅里叶变换为 $\{W_{nm} \exp(j2\pi P_{nm})\}_{nm}$, $\{\dots\}_{nm}$ 表示所有点 n, m 构成的序列。减小 W_{nm} 的取值范围的第一步,是将 θ_{nm} 的值赋给初始理想图像,这些相位采样相互独立且都均匀分布于 $(-\pi, \pi)$ [Gallagher and Sweeney, 1979]。得到的图像采样的 DFT 记为 $\text{DFT}[\{a_{nm} \exp(j\theta_{nm})\}_{nm}] = \{A_{nm} \exp(j\psi_{nm})\}_{nm}$ 。接着,将幅值谱 A_{nm} 设为任意正常数 A 。修改幅值后的谱的逆 DFT 为 $\text{DFT}[\{A \exp(j\psi_{nm})\}] = \{\tilde{a}_{nm} \exp(j\tilde{\theta}_{nm})\}$ 。原始图像幅度 a_{nm} 与新的相位值 $\tilde{\theta}_{nm}$ 结合形成一个新的目标图像采样。按照预先给定的迭代次数重复进行该过程。最近一次迭代获得的图像相位作为新图像的相位。最终的图像相位值与原始图像幅值结合在一起生成用于设计全息图的 $W_{nm} \exp(j2\pi P_{nm})$ 。

通过全息图域中约束幅值和迭代操作,全息平面的信息由幅值转移到相位中。因此,减小了由所有孔径均为同一高度所造成的负面影响。于是 15.3 节中的近似 (b) 也能以近似 (a) 同样的方式处理。该方法的进一步归纳将在 16.6 节中讨论。

15.5 量化罗曼法

15.2 节中讨论的罗曼法中允许有无穷多的孔径尺寸和位置,这对许多实现方法是不实际的。为克服这种障碍,可以用离散方法来量化单元中孔径的尺寸和位置。在改进的方法中,仍然利用每个孔径的尺寸来控制幅值,用孔径位置的改变来控制相位,只不过幅值与相位的可能值被量化了。

在量化的罗曼法中,每个全息图单元被分成 $N \times N$ 的阵列,记为 N 级量化。这限制了每个孔径可能的中心位置和高度。这样,每个单元的相位和幅值得到了量化。特别的是,孔径中心有 N 个可能位置 (相位),因为单元在 y 方向上是对称的,所以共有包括 0 幅值在内的 $N/2 + 1$ 个潜在高度值 (幅值)。 N 较大时能生成接近真实全息图的编码图样。

例如,我们考虑一个被分成 4×4 小方块的罗曼单元,在罗曼算法中,若 $c = 1/2$,意味着孔径宽度限定为整个单元宽度的一半,则该单元允许三个归一化幅值 (0, $1/2$ 和 1) 及四个相位值 ($-\pi$, $-\pi/2$, 0, $\pi/2$), 总共有 12 种组合。量化意味着在对幅值和相位进行量化时会产生误差。因此,引入如凸集投影法 (POCS) 这样的优化算法并设计子全息图循环计算方法直至满足收敛条件,将是有益的。这将在第 16 章中进一步讨论。

量化孔径位置对于实际实现是很有用的。比如,空间光调制器 (SLM) 可被用来对全息



图每一点的幅值与相位调制进行实时控制，但不幸的是，SLM 却不能生成精确的罗曼单元，若要使实现成为可能，则必须进行量化。同样的，如 11.6、11.7 节所讨论的，制作集成电路的技术可被用来实现 DOE 的制造。由于精确连续的表面浮雕非常难以实现，相位与幅值的量化在所有这样的技术中也都是必要的。

15.6 罗曼法的计算机仿真

由罗曼法获得的二元透过率图样可用以下两种方法之一来显示：

方法 1：要显示精确罗曼全息图（如孔径尺寸和位置已严格确定），将图样画在计算机辅助设计布局图上，比如用 AutoCAD 软件。

方法 2：与量化罗曼法相同。

利用参量 $c = 1/2$ ， $M = 1$ ，设计所有全息图 [Kuhl, Ersoy]，设定 $c = 1/2$ 意味着每个孔径宽度为单元宽度的一半。同时也可看到选定此 c 值使图像亮度最大 [Gallagher, 1979]。选取 $c = 1/2$ 后， M 必须取为 1 [Lohmann, 1970]。并且，假定 15.3 节讨论的近似都有效。

图 15.3 给出了尺寸为 64 像素 × 64 像素的二元图像 E，它完全放置在像平面的一边。

由于全息面的场是实值的，因此图像傅里叶变换具有厄米对称性

$$U[n,m] = U^*[N-n,M-m] \tag{15.6-1}$$

图 15.4 给出了图 15.3 中图像的 DFT 的幅值。



图 15.3 用于生成罗曼全息图的图像 E

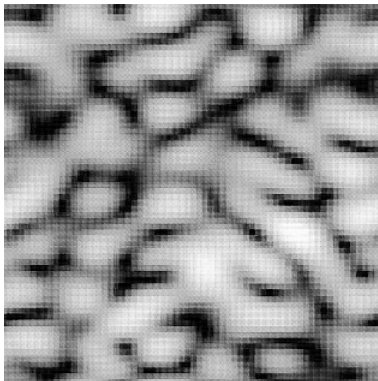


图 15.4 图 15.3 中图像的 DFT 的幅度

图 15.5 为由上面讨论到的方法 2 生成的罗曼全息图，量化阶数为 16。

图 15.6 为仿真得到的重构图像。孪生像清晰可见。

对图 15.7 中的 Lena 的 512 像素 × 512 像素灰度图像也进行了实验。

图 15.8 为由灰度图像的罗曼全息图得到的计算机重构图像。

利用迭代优化方法，对上述两幅图像也可使用第 15.4 节讨论的等幅罗曼法进行实验研究。图像 E 的等幅罗曼全息图如图 15.9 所示，由它得到的计算机重构图像如图 15.10 所示。

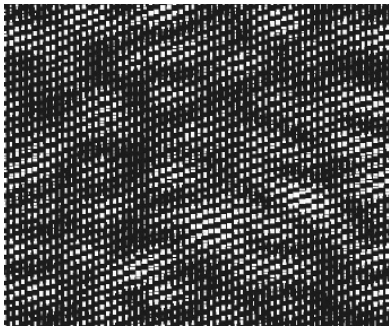


图 15.5 图 15.3 中图像的罗曼全息图



类似地，由图 15.7 中所示灰度图像的等幅罗曼全息图的计算机重构图像如图 15.11 所示。

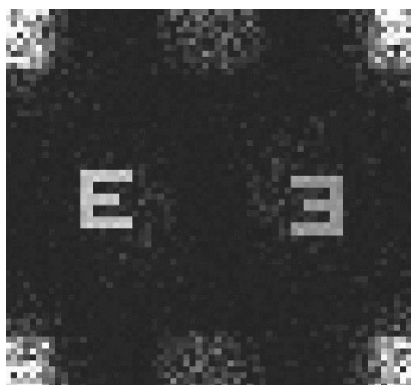


图 15.6 图 15.5 中全息图的计算机重构图像 图 15.7 Lena 的大小为 512 像素 × 512 像素的灰度图像

对比图 15.6 和图 15.10，可以看出在常幅值的情况下，目标图像 E 更亮，但是图像 E 的锐度稍微降低了一些。常幅值全息图在像平面边角产生的噪声也更少。

对比图 15.8 和图 15.11，常幅值全息图方法与原始方法相比，目标图像依然更亮且在像平面边角的噪声更少。

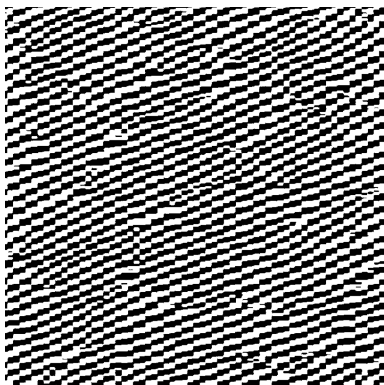
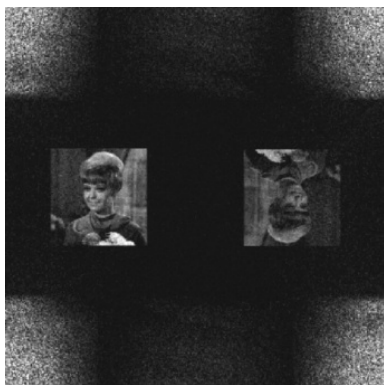


图 15.8 由灰度图像的罗曼全息图得到的计算机重构图像 图 15.9 图 15.3 中所示图像的等幅罗曼全息图



图 15.10 由图 15.9 中所示全息图得到的计算机重构图像

图 15.11 由图 15.7 中所示灰度图像的等幅罗曼全息图的计算机重构图像



15.7 基于硬限幅的傅里叶法

通过对相位函数硬限幅可生成二元傅里叶变换 DOE。在此过程中，幅值信息被忽略了。为简单起见，我们在一维情形下进行分析。考虑复函数

$$A(f)e^{j\phi(f)}$$

被平面参考波

$$e^{j\alpha f}$$

调制得到

$$A(f)e^{j(\alpha f + \phi(f))}$$

信号的实部经过硬限幅透过率函数，如图 15.12 所示。

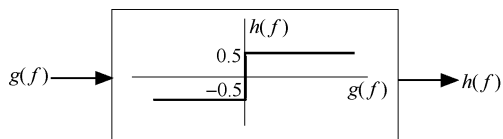


图 15.12 硬限幅透过率函数

将函数加上 0.5 进行修正，使其在 0 至 1 间变化，这相当于在 DFT 中相应增加了直流分量。

采用双极透过率函数的硬限幅信号由下式给出：

$$h(f) = \frac{1}{2} \left[1 + \frac{A(f) \cos(\alpha f + \phi(f))}{|A(f) \cos(\alpha f + \phi(f))|} \right] \quad (15.7-1)$$

$h(f)$ 以傅里叶级数的形式可表示为 [Kozma and Kelly, 1965]

$$h(f) = \frac{1}{2} + \sum_{m(\text{奇})=1}^{\infty} \frac{2(-1)^{\frac{m-1}{2}}}{m\pi} \cos(m\alpha f + m\phi(f)) \quad (15.7-2)$$

式 (15.7-2) 表示随散射角增大的无穷多图像。这些图像中最重要的是由等号右边的 $1/2$ 这项产生的零阶像，以及由 $m=1$ 项产生的角度为 $\pm\alpha$ 的孪生像。

图 15.13 为硬限幅法实验中用到的图像，该图像的逆离散傅里叶变换 (DFT) 经过了硬限幅滤波器编码，图 15.14 为得到的 DOE。图 15.15 为 DOE 的前向 DFT 经零阶滤波后的图像，孪生像清晰可见。为减小幅值变化的影响，将图 15.13 乘以漫反射器 (随机相位扩散器)，再将上述过程重复进行。重构结果显示在图 15.16 中。可以看到，结果较之前更令人满意，这显示了利用漫反射器的相位编码法的重要性。



图 15.13 硬限幅法实验中用到的原始图像

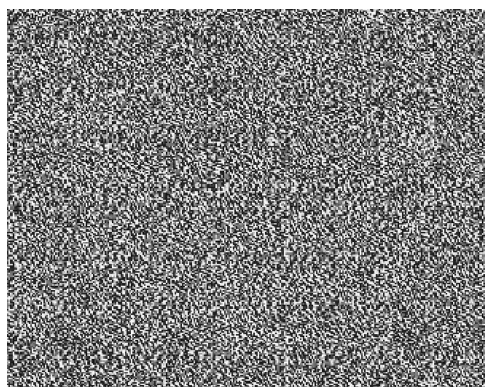


图 15.14 硬限幅法生成的全息图

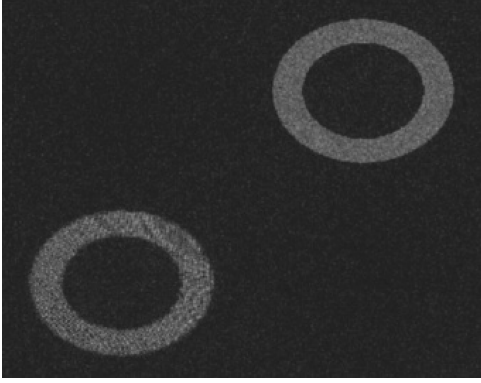


图 15.15 不带随机相位漫反射器的 DOE 重构

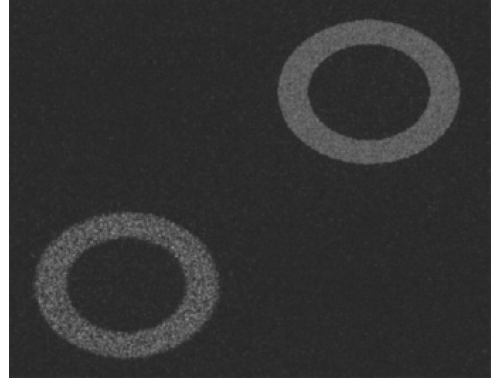


图 15.16 带随机相位漫反射器的 DOE 重构

15.8 一种 3-D 点图像重构的简单算法

有时,用一种简单但具有较好鲁棒性的算法来检验新结果、新实验等是更好的选择。下面给出的算法对达到这种目的很有用,并首次展示了用扫描电子显微镜实现 DOE 算法的可行性 [Ersoy, 76]。尽管开始很简单,该算法随后发展得更加复杂,以便在空间中任意位置生成 3-D 图像 [Bubb, Ersoy]。这种方法与其他方法很不一样,因为它没有明显地依赖傅里叶变换,它也可产生将在 15.10 节中讨论的单幅图像全息术,并且它在第 18 章中的光学相控阵中会得到进一步应用。

如果 (x_o, y_o, z_o) 表示被重合物点的位置, (x_i, y_i, z_i) 表示全息图中相移孔径的位置,对平面 N 个孔径集应用惠更斯-菲涅耳原理,有

$$U(x_o, y_o, z_o) = \sum_i^N \iint U(x_i, y_i, z_i) \frac{1}{j\lambda} \frac{\exp(jkr_{oi})}{r_{oi}} \cos\delta dx_i dy_i \quad (15.8-1)$$

其中, δ 是 z 轴与由孔径中心到物点的向量之间的夹角; r_{oi} 为该向量的长度; λ 为波长; k 是波数。

假定 $U(x_i, y_i, z_i)$ 为单位复振幅平面波,对于小尺寸的全息图,可以认为 $\cos\delta$ 为常数。如果全息面的相位变化与 $\exp(jkr_{oi})$ 相比也很小的话,上面的等式可近似为

$$U(x_o, y_o, z_o) = \frac{\cos\delta}{j\lambda R} d_x d_y \sum_i^N \iint \frac{\exp(jkr_{oi})}{r_{oi}} dx_i dy_i \quad (15.8-2)$$

式中, $\theta_i = kr_{oi}$ 为第 i 个孔径处参考光波的相移。

假设每个孔径都是 x - y 平面内尺寸为 $d_x \times d_y$ 的长方形,中心点为 $(x_{si}, y_{si}, 0)$, 观察点的径向距离 r_{oi} 满足

$$kr_{oi} = 2\pi n + \theta_i, n \text{ 为整数} \quad (15.8-3)$$

利用夫琅禾费近似,孔径处的源场可看作由窄 sinc 函数近似的点源。因此,式 (15.8-2) 的积分变为二重求和:

$$U(x_o, y_o, z_o) = \frac{\cos\delta}{j\lambda R} d_x d_y \sum \sum \exp(j\theta_i) \text{sinc}\left(\frac{X_i d_x}{\lambda r_{oi}}\right) \text{sinc}\left(\frac{Y_i d_y}{\lambda r_{oi}}\right) \quad (15.8-4)$$



式中

$$\begin{aligned} X_i &= x_o - x_{si} \\ Y_i &= y_o - y_{si} \\ R &= r_{oi} \text{ 的均值} \end{aligned} \quad (15.8-5)$$

如果将所有的 θ_i 设为 θ ，且 $x_i, y_i \ll r_{oi}$ ，sinc 函数可用 1 代替，并有

$$U(x_o, y_o, z_o) = \frac{\cos \delta}{j\lambda R} d_x d_y N e^{j\theta} \quad (15.8-6)$$

因此，场的幅值正比于 $d_x d_y N$ ，当平面波以直角入射到全息图上时，其相位为 θ ，即为轴上平面波。如果入射波在全息面上有相位变化，则在每个全息点上的相位为 Φ_i ，则式 (15.8-3) 应写为

$$kr_{oi} + \Phi_i = 2\pi n + \theta_i, \quad n \text{ 为整数} \quad (15.8-7)$$

可选择孔径位置 (x_i, y_i) 使得 θ_i 为常数。全息图孔径生成的所有波前在特定物点位置相位相加，由此获得一个物点。因此，场的幅值正比于 $d_x d_y N$ ，并且当平面波垂直入射到全息面上时，其相位为 0。如果在空间中不同地点存在满足式 (15.8-6) 的孔径组，则本质上就产生了在每一物点具有确定幅值与相位的采样波前。注意，式 (15.8-6) 的调制很简单，改变 d_x 和 (或) N 可调制幅值，改变 θ 可调制相位。事实上， N 通常是一个很大的数值，这意味着它几乎可以连续变化，于是幅值调制能达到很高的精度。

如果幅值调制精度不需要很高，这种方法在近场也很有效，因为在每一物点 r_{oi} 都被精确计算。如果孔径是圆形的，则用一阶贝塞尔函数取代 sinc 函数，但式 (15.8-6) 本质上保持不变。

实际应用中，首先随机选取全息面上的每一个点孔径，然后沿 x 或 y 方向细微移动使得其中心坐标满足式 (15.8-7)，且 θ 为常数。可以忽略孔径的重叠以节省内存。

15.8.1 实验

用以上描述的方法测试采用扫描电子显微镜实现的 DOE [Ersoy, 1976]。连续曝光的工作区域为 $2\text{mm} \times 2\text{mm}$ 。点孔径的数目为 4096×4096 ，最小可达到的直径为 $1\mu\text{m}$ 。实验中，使用的全息图材料为 KPR 负性光刻胶或者 PMMA 正性光刻胶。

图 15.17 和图 15.18 为由两幅生成的全息图重构得到的图像。计算时使用波长为 $0.6328\mu\text{m}$ 的 He-Ne 激光。图 15.17 中，在物平面中 3cm 长的线上选取 11 个点，满足 $z = 60\text{cm}$ ， $x = 40\text{cm}$ ， $0 \leq y \leq 3\text{cm}$ 。使用的全息图孔径数目为 120 000，每个上面有 8×8 个相邻点。图像中的点选取为等强度。相片拍摄于距全息图约 60cm 处（即焦平面）。阻挡主波束使胶片不过度曝光，可以看到存在实像和共轭像。

为显示三维效果，利用图 15.18 中的物体。四个字母每个都选取在空间中的不同平面上，四个平面距全息图分别为 60cm 、 70cm 、 80cm 和 90cm 。如果所有的字母都在同一个平面上，则它们在 x 方向上的间距为 1cm 。从图片中可以看出，由于深度的影响，从第一个字母到最后一个字母间距逐渐减小。由于在距全息图约 90cm 的平面即字母 E 所在的焦平面拍照，所以 E 最亮，L 最暗。所用全息图孔径的数目为 100 000，每个孔径有 4×4 个相邻点。图像上的点的强度相同。

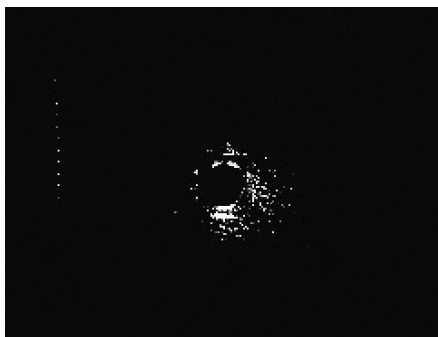


图 15.17 由扫描电子显微镜生成的 DOE 重构空间中 3cm 长直线上的 11 个点



图 15.18 3-D 空间中单词 LOVE 的重构, 每个字母处于空间中不同的平面上

15.9 快速加权零交叉算法

15.8 节中讨论的算法, 相当于在全息图上选取许多随机位置的球面波的相位零交叉点。这种方法的一个缺点是当物点增加时, 全息图快速饱和。此外, 计算量也很大。为解决这些问题, 可采用快速加权零交叉算法 (FWZC) [Bubb, Ersoy]。

对坐标为 (x_0, y_0, z_0) 的每一需要生成的物点, FWZC 的步骤如下。

(1) 分别计算 x 轴、 y 轴上最邻近原点的相位零交叉点 Δx_1 、 Δy_1 , 如图 15.19 所示。

令 $(\Delta x_1, 0, 0)$ 为 x 轴上零交叉点的位置, 则

$$r_{oi} = \sqrt{(x_0 - \Delta x_1)^2 + y_0^2 + z_0^2} = n\lambda \quad (15.9-1)$$

$$r'_{oi} = \sqrt{x_0^2 + y_0^2 + z_0^2} = n\lambda + B \quad (15.9-2)$$

解上述方程得 Δx_1 为

$$\Delta x_1 = x_0 \pm \sqrt{x_0^2 + (B^2 - 2Br'_{oi})} \quad (15.9-3)$$

可以推导出 Δy_1 具有完全相似的表达式。

(2) 利用海伦表示 [Ralston], 计算这些零交叉点与物点的距离 r_x 、 r_y 。例如, 经推导 r_x 为

$$r_x = z_1 \left[3 + x_1^2 + y_1^2 - \frac{1}{(x_1^2 + y_1^2 + 1)} \right] \quad (15.9-4)$$

$$z_1 = \frac{z_0}{2} \quad (15.9-5)$$

$$z_2 = \frac{1}{\sqrt{2}z_0} \quad (15.9-6)$$

$$x_1 = (x_0 - \Delta x_1)z_2 \quad (15.9-7)$$

$$y_1 = y_0z_0 \quad (15.9-8)$$

(3) 确定 x 轴上所有的零交叉点。从 x 轴上任意一零交叉点开始计算, 如从 $x = a$ 点,

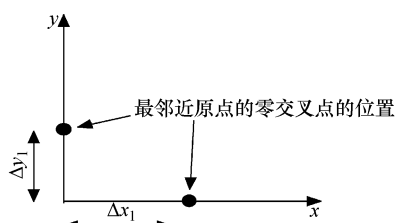


图 15.19 最邻近原点的零交叉点

要计算下一零交叉点，如 $x = b$ 点。取距离函数 $r_x = \sqrt{(x - x_0)^2 + y_0^2 + z_0^2}$ 的泰勒级数的前两项，结果为

$$\Delta x = b - a = \frac{r(a)\lambda}{x_o - a} + \frac{1}{2(x_o - a)} \left[\left(\frac{r(a)\lambda}{x_o - a} \right)^2 - \lambda^2 \right]$$

(15.9-9)

当在 x 轴负方向采用相同步骤时， $r(b) - r(a) = \lambda$ ，相似的分析表明式 (15.9-9) 中右边第一项应改变符号。

- (4) 对 y 轴正负方向采用相同的步骤可以找出所有零交叉点。
- (5) 对任意 $x = x_{11}$ ， $y = y_{11}$ ，如果 $(x_{11}, 0)$ ， $(0, y_{11})$ 是零交叉点，在非涅耳近似有效的情况下 (x_{11}, y_{11}) 也是零交叉点。

利用 x 轴、 y 轴上所有“快速”零交叉点，在全息面上形成零交叉点网格。也可以通过在“快速”零交叉点间插值来生成其余的零交叉点。

(6) 生成的零交叉点网格如图 15.20 所示，其指示了孔径位置的中心，这些孔径在记录介质上生成所需的全息图。

每一物点的零交叉点一旦生成，便指定“击中”符号（比如，1）来标记它们在网格上的位置。计算所有物点的零交叉点以后，每一孔径位置将积累许多“击中”，其个数范围为 0 到物点的数目。图 15.21 给出了这一过程。

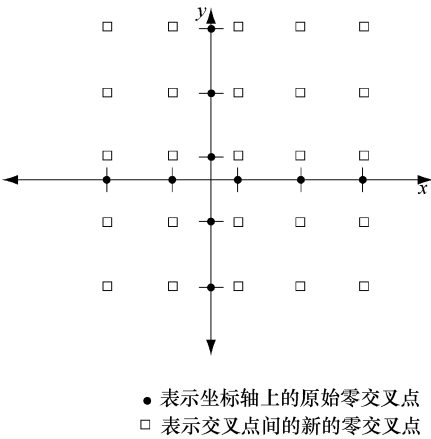


图 15.20 生成的零交叉点网格

增加的计数值	至少所需孔径的数目
⋮	⋮
30	46 748
29	64 816
28	71 903
27	86 274
26	95 816
25	106 892
24	145 981
⋮	⋮

图 15.21 全息图上零交叉点的积累

对于复杂的物场景，绝大部分孔径位置将积累一个“击中”。接着，可设置一个“击中”数目的阈值，并只对“击中”数目超过指定阈值的孔径进行编码。这样，我们选取了那些最重要的零交叉点，比如，选择了那些对重构物体贡献最大的零交叉点。实际中效果很好的一种设计方案为，设定阈值使得生成的孔径大约覆盖全息面的一半。

阈值设计方案存在一个问题。考虑式 (15.9-9) 给出的 x 轴上零交叉点之间的间隔。当 x_0 增大时，零交叉点之间的间隔减小。这将造成离物体中心较远的物点有更大数目的零交叉点。为尝试改变这一现象，可为每一零交叉点指定不同的“击中”数目，使得每一物点的总“击中”数目为常数。实现这一方法的一种可行方案是

$$\text{每一零交叉点的未击中次数} = \frac{A}{\text{当前物点的非零交叉点数}}$$

(15.9-10)



式中, A 为一适当大的常数。

15.9.1 离轴平面参考波

与 x 轴倾斜的离轴平面波可写为

$$U(x, y) = e^{-j2\pi\alpha x} \quad (15.9-11)$$

其中

$$\alpha = \frac{\sin\theta}{\lambda} \quad (15.9-12)$$

式中, θ 为平面波与光轴的夹角。

利用该表达式, 式 (15.9-1) 变为

$$r_{oi} = n\lambda + x\sin\theta \quad (15.9-13)$$

其余的步骤和同轴平面波法的步骤相同。

15.9.2 实验

我们用前述方法 [Bubb, Ersoy] 生成了许多全息图。图 15.22 为 FWZC 方法生成的一幅全息图。用 He-Ne 激光器照射缩小尺寸的该全息图胶片以重构物体, 如图 15.23 所示。

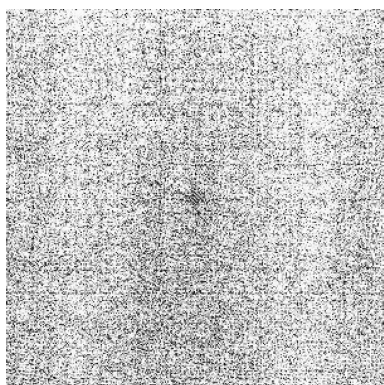


图 15.22 FWZC 方法生成的一幅全息图

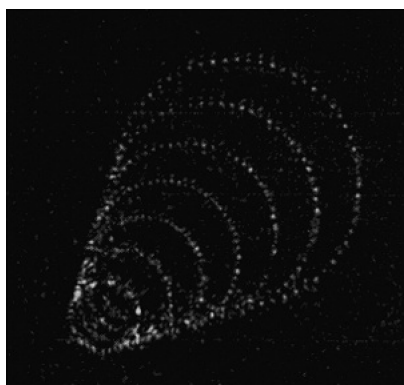


图 15.23 由图 15.22 中全息图重构得到的同心圆

图 15.23 的重构图像为半径从 0.08 ~ 0.15m 的七个同心圆, 这些同心圆处于空间中不同的 z 平面中: 最小的位于 0.75m 处, 最大的位于 0.6m 处。这些同心圆在 z 轴上的间隔清晰可见, 因为视差和透视缩短的效应很明显。在所有进行的实验中, 没发现对物点数目的实际限制。

15.10 单图全息术

对于垂直照射全息图的参考平面波, 很容易明白, 用不同编码方法都会产生孪生像, 这是因为平面全息图的两面在物理空间是对称的。如果我们将物光波看作从单个物点发出的球面波的叠加, 则相对全息面互为镜像的点, 对应全息图上幅值相同、相位相反的物光波。因此, 当我们选择对应虚像相位 ϕ_i 的全息图孔径 i 时, 相位为 $-\phi_i$ 的实像也对应相同的全息图孔径。



结论是要分离出一个像，全息图两边的两个物理空间的对称性需要尽可能地消除。全息面不为平面或者参考波为非垂直入射平面波时就能改变这种对称性。然而，选择另一种简单的几何结构如离轴平面波能消除对称性，使得两个像处在与以前不同的位置。

球面参考波是个不错的选择，因为它用一个透镜便可很容易获得 [Ersoy, 1979]。如果只重构实像，透镜的焦点可选择为穿过全息图使得主波束和零级光能被焦点处的挡板滤掉。在下节中，将验证这个想法，如图 15.24 所示。

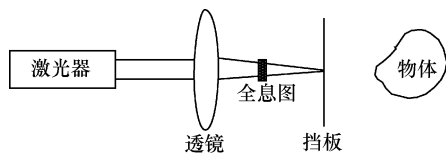


图 15.24 单图全息术的装置

在 15.7 节中讨论到的编码技术中，每一个全息孔径依下式选择：

$$\phi(x_i, y_i) + kr_{oi} = 2n\pi + \phi_0 \quad (15.10-1)$$

在这种情况下， $\phi(x_i, y_i)$ 是波从参考波前初始点 (x_c, y_c, z_c) 传播到全息孔径 (x_i, y_i) 时产生的相移； kr_{oi} 是波从全息图上的孔径 (x_i, y_i) 传播到物点 (x_o, y_o, z_o) 时产生的相移。半径 r_{oi} 由下式给出：

$$r_{oi} = \pm \sqrt{(x_o - x_i)^2 + (y_o - y_i)^2 + z_o^2} \quad (15.10-2)$$

+ (-) 号用来表示物体的实 (虚)。

对于焦点在 (x_c, y_c, z_c) 处的球面参考波，参考波的相位 $\phi(x_i, y_i)$ 可写为

$$\phi(x_i, y_i) = kr_{ci} \quad (15.10-3)$$

式中

$$r_{ci} = \pm \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2 + z_c^2} \quad (15.10-4)$$

+ (-) 号用来表示透镜焦距在全息图的前面 (后面)。

等式 (15.10-1) 可写为

$$r_{ci} + r_{oi} = n\lambda + \frac{\phi_0 \lambda}{2\pi} \quad (15.10-5)$$

假设孔径的位置沿预定的方向，如 x 方向，移动距离 Δ ，以满足等式 (15.10-5)。那么，半径矢量 \mathbf{r}'_{oi} 和 \mathbf{r}'_{ci} 的长度为

$$\mathbf{r}'_{oi} = \sqrt{r_{oi}^2 - 2\Delta(x_o - x_i) + \Delta^2} \quad (15.10-6)$$

$$\mathbf{r}'_{ci} = \sqrt{r_{ci}^2 - 2\Delta(x_c - x_i) + \Delta^2} \quad (15.10-7)$$

$\mathbf{r}'_{ci} + \mathbf{r}'_{oi}$ 满足

$$\mathbf{r}'_{ci} + \mathbf{r}'_{oi} = r_{ci} + r_{oi} - B \quad (15.10-8)$$

解方程式 (15.10-8) 得 Δ 为

$$\Delta = C_1 [1 - \sqrt{1 - C_2/C_1}] \quad (15.10-9)$$

式中

$$C_1 = \frac{F_1}{F_2} \quad (15.10-10)$$

$$C_2 = \frac{F_3}{F_1} \quad (15.10-11)$$

$$F_1 = F_4 (X + X_c) + X r_{ci}^2 + X_c r_{oi}^2 \quad (15.10-12)$$



$$F_2 = (r_{ci} + r_{oi})^2 - (X - X_c)^2 - 2F_5 \quad (15.10-13)$$

$$F_3 = 2F_5 r_{oi} r_{ci} - F_5^2 \quad (15.10-14)$$

$$F_4 = r_{oi} r_{ci} - F_5 \quad (15.10-15)$$

$$F_5 = -\frac{B^2}{2} + B(r_{ci} + r_{oi}) \quad (15.10-16)$$

注意，可以通过改变每一物点对应的全息孔径的数目来实现物点的幅值调制，而这一数目与该点的理想幅值成比例。尽管球面波的幅值在全息图上缓慢变化，孔径随机分布造成的平均效果可以使我们认为幅值是恒定的。

15.10.1 成像分析

为了分析这种方法，半径的长度可以扩展到傍轴近似 [Meier, 1966]。假设不用球面参考波，而用如下的平面参考波：

$$U_r = A e^{jk(\alpha x + \beta y + \gamma z)} \quad (15.10-17)$$

则等式 (15.10-5) 在傍轴近似下可写为

$$x_i \left(\alpha - \frac{x_0}{z_0} \right) + y_i \left(\beta - \frac{y_0}{z_0} \right) + \frac{x_i^2 + y_i^2}{2z_0} = n\lambda + \frac{\phi_0 \lambda}{2\pi} \quad (15.10-18)$$

式中， z_0 符号的选取应与式 (15.10-2) 中一致。

如果选用式 (15.10-4) 所描述的球面参考波，式 (15.10-5) 可写为

$$-x_i \left(\frac{x_c}{z_c} + \frac{x_0}{z_0} \right) - y_i \left(\frac{y_c}{z_c} + \frac{y_0}{z_0} \right) + \left(\frac{x_i^2 + y_i^2}{2} \right) \left(\frac{1}{z_0} + \frac{1}{z_c} \right) = n\lambda + \frac{\phi_0 \lambda}{2\pi} - F_{0c} \quad (15.10-19)$$

式中

$$F_{0c} = z_0 + \frac{x_0^2 + y_0^2}{2z_0} + z_c + \frac{x_c^2 + y_c^2}{2z_c} \quad (15.10-20)$$

上述各式中 z_0 和 z_c 符号的选取应与式 (15.10-2) 和式 (15.10-3) 中一致。

忽略常数项 $\frac{\phi_0 \lambda}{2\pi}$ 、 F_{0c} ，对任意 n ，只要式 (15.10-18) 或式 (15.10-19) 成立，图像便可形成。可知，若式 (15.10-18) 对物点 (x_0, y_0, z_0) 、方向余弦为 (α, β, γ) 的参考波和波长 λ 成立，则其对物点 (x'_0, y'_0, z'_0) 、方向余弦为 $(\alpha', \beta', \gamma')$ 的参考波和波长 λ' 亦成立，因此

$$z'_0 = \frac{\lambda}{\lambda'} \frac{z_0}{m} \quad (15.10-21)$$

$$x'_0 = \frac{\lambda}{\lambda'} \left(\frac{\alpha' - m\alpha}{m} z_0 + x_0 \right) \quad (15.10-22)$$

$$y'_0 = \frac{\lambda}{\lambda'} \left(\frac{\beta' - m\beta}{m} z_0 + y_0 \right) \quad (15.10-23)$$

式中， m 为整数， m 取值为 1 和 -1 分别对应实像和虚像， m 取值为 0 对应零级波，即一个与原始参考波传播方向相同的平面波， m 取其他值时对应高阶像。

同样，如果式 (15.10-19) 对物点 (x_0, y_0, z_0) 、焦点为 (x_c, y_c, z_c) 以及波长为 λ 的参考波成立，则对物点 (x'_0, y'_0, z'_0) 、焦点为 (x'_c, y'_c, z'_c) 以及波长为 λ' 的参考波也成立，于是



$$\frac{1}{z'_0} = \frac{\lambda'}{\lambda} m \left(\frac{1}{z_0} + \frac{1}{z_c} \right) \quad (15.10-24)$$

$$x'_0 = z'_0 \left[\frac{\lambda}{\lambda'} m \left(\frac{x_c}{z_c} + \frac{x_0}{z_0} \right) - \frac{x'_c}{z'_c} \right] \quad (15.10-25)$$

$$y'_0 = z'_0 \left[\frac{\lambda}{\lambda'} m \left(\frac{y_c}{z_c} + \frac{y_0}{z_0} \right) - \frac{y'_c}{z'_c} \right] \quad (15.10-26)$$

式中, m 为整数。

注意, m 取为 0 时对应于与参考波相同、在焦点会聚的波。如果 $z_0 \gg z_c$, 并且假设初始参考波和最终参考波相同, 且 $m \neq 1$ 时, 像的位置由下式近似给出:

$$z'_0 \simeq \frac{z_c}{m-1} \quad (15.10-27)$$

$$x'_0 \simeq x_c + m \frac{z'_0}{z_0} x_0 \quad (15.10-28)$$

$$y'_0 \simeq y_c + m \frac{z'_0}{z_0} y_0 \quad (15.10-29)$$

换言之, 如果透镜焦点离全息图足够近, 则除 $m=1$ 对应的像外, 其他所有的像必须离全息图足够近, 以至于在目标图像所处的较远位置处完全离焦。

然而, 可以一次只聚焦某一阶的衍射光。如果第 m 阶衍射光的位置为 (x'_0, y'_0, z'_0) , 很容易找到相应的 (x'_c, y'_c, z'_c) , 使得式 (15.10-24)、式 (15.10-25) 和式 (15.10-26) 成立。那么, 所有其他的像在目标像所处的较远位置处完全离焦, 因而不可见, 因为它们处在离全息图很近的位置。

以上分析不总是合理的, 因为当球面参考波接近全息图时, 傍轴近似便很难成立。因此, 除所需目标像外, 其他像可能不好定义。同样所用的编码技术在离全息图很近的距离时可能会失效, 特别是由于全息孔径配准误差的影响。

为了支持上述讨论, 考虑式 (15.10-24)、式 (15.10-26) 中当 $z'_c \rightarrow \infty$ 的情况; 换言之, 利用平面波来实现重构。同样设 $z_0 \gg z_c$, λ 等于 λ' , 得到

$$z'_0 \simeq \frac{z_c}{m} \quad (15.10-30)$$

$$x'_0 \simeq x_c + \frac{z_c}{z_0} x_0 \quad (15.10-31)$$

$$y'_0 \simeq y_c + \frac{z_c}{z_0} y_0 \quad (15.10-32)$$

上述等式表明像必须聚焦在靠近全息图的位置。感兴趣的空间中不存在成像, 表明谐波像和所需目标像之间的干涉达到了最小化。

当 z_c 更小时, 观察高阶图像的位置变化也很有趣。由式 (15.10-24), 对于相同波长, 有

$$z'_0 \simeq \frac{z_c z'_c}{m z'_c - z_c} \quad (15.10-33)$$

随着 m 增大, z'_0 趋近 z_c/m 。这表明全息图离透镜焦点越近, 将越难观察到高阶像。

式 (15.10-33) 的一个重要推论是存在可变的聚焦距离; 通过微调透镜的位置可以任意



改变 z'_0 。

15.10.2 实验

为了从实验上验证上述猜想，用 15.8 节讨论的扫描电子显微镜系统生成全息图。对于所有全息图，全息图的尺寸均为 $2\text{mm} \times 2\text{mm}$ ，孔径的尺寸在 $1\mu\text{m}$ 量级。

图 15.25 所示为用 He-Ne 激光器发出的垂直平面参考波对某一物点的全息图重构。在中心处，可以观察到过度曝光的主光束；主光束的右边是轻微过度曝光的物点的实像和高阶实像。主光束的左边是虚像和高阶虚像。

用球面参考波生成某一物点的三个全息图，使透镜焦点在全息平面左侧 1cm 、 3cm 、 5cm 处，其 x 、 y 坐标与全息图中心点的相同。图 15.26 为焦点位置在全息平面左侧 5cm 的全息图的再现结果。图中只有一个可见物点；较暗的方块是扩大的全息图的像；主光束覆盖了整幅画面。图 15.27 为相同物点但在透镜焦点处用挡板滤掉主光束和零级像后的重构图像。焦点在全息平面左侧 1cm 和 3cm 的全息图给出了相同的结果，从而支持了上述讨论 [Ersoy, 1979]。

图 15.28 给出了对更复杂物体的重构。全息图离透镜焦点 3cm ， x 、 y 坐标与全息图中心点的相同。在实验过程中，可观察到很容易在不同距离上聚焦图像，因此也很容易通过对透镜位置的细微调整控制图像的尺寸，这进一步验证了式 (15.10-33) 是正确的。

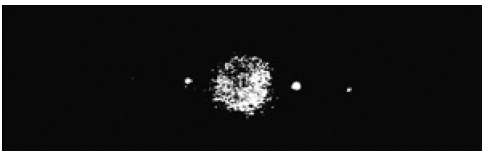


图 15.25 利用 He-Ne 激光重构物点得到的不同阶像

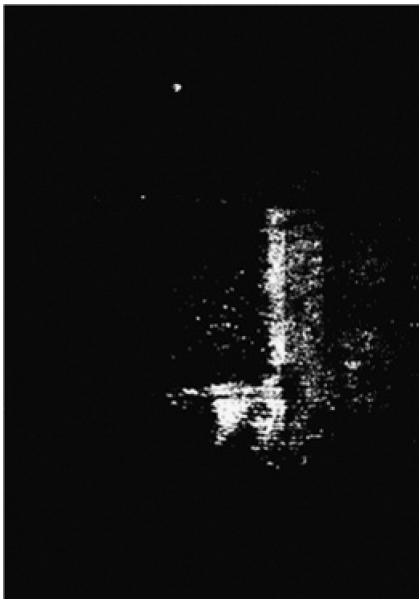


图 15.26 由 5cm 全息图不经滤波得到的单个可见图像



图 15.27 由 5cm 全息图经滤波得到的单个可见图像

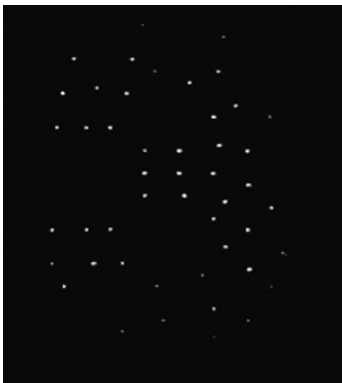


图 15.28 数字 3 的单重构图像



15.11 菲涅耳波带片

在讨论 DOE 更复杂的设计之后, 我们回到从前, 讨论一种经典的 DOE, 即菲涅耳波带片 (Fresnel Zone Plate, FZP)。这样的 DOE 可被用作平面透镜, 能被大量制造, 常用于高架投影仪等技术应用中。本节将讨论二元 FZP, 其包含不透明和透明的圆环形区域。图 15.29 给出了一个例子。图 15.30 是 FZP 的侧视图。

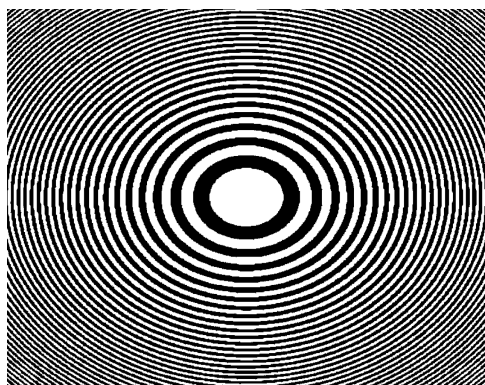


图 15.29 二元菲涅耳波带片

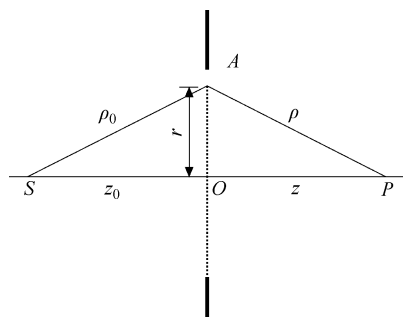


图 15.30 FZP 的侧视图

在图 15.30 中, 光线经由 SOP 和 SAP 传播的光程差为

$$\Delta(r) = (\rho_0 + \rho) - (z_0 + z) = \sqrt{r^2 + z_0^2} + \sqrt{r^2 + z^2} - (z_0 + z) \quad (15.11-1)$$

定义菲涅耳波带片的参数 n , 使得光程差为半波长的整数倍 [Hecht]:

$$n \frac{\lambda}{2} = \Delta(r) \quad (15.11-2)$$

第 n 级菲涅耳带为半径为 r_{n-1} 、 r_n 的圆之间的环形区域。注意, 半径为 r_n 的圆对 P 点的场的贡献与半径为 $r_{n\pm 1}$ 的圆对 P 点的场的贡献相差半个波长。据此易知, 相邻波带的场相互抵消。因此, 如果所有的偶数波带或者奇数波带被遮挡, P 点处的场将增强, 余下的菲涅耳波带相互增强, 在距离 z 处产生聚焦点。假设波带片被平面波照射, $z_0 \approx \infty$, 并令 $z = f_0$, 式 (15.11-1) 变为

$$\frac{r_n}{\lambda} = \sqrt{\frac{f_0}{\lambda} \left(n - \frac{1}{2} \right) + \frac{(2n-1)^2}{16}} \quad (15.11-3)$$

式中, f_0 为焦距。

FZP 可看作焦距为 f_0 的透镜。显然在 $f_0/3$ 、 $f_0/5$ 、 $f_0/7$ 等处还存在着其他焦点 [Hecht]。

令 R_n 为 FZP 上第 n 级圆的半径, 且满足

$$R_n^2 = \left(f_0 + \frac{n\lambda}{2} \right)^2 - f_0^2 \quad (15.11-4)$$

利用式 (15.11-3), 式 (15.11-4) 写为

$$R_n^2 = f_0^2 \left[\frac{n\lambda}{f_0} + \frac{1}{4} \left(\frac{n\lambda}{f_0} \right)^2 \right] \quad (15.11-5)$$

当 n 很大时, 式 (15.11-5) 等号右边第二项可以忽略, 因此

$$R_n \approx \sqrt{n\lambda f_0} \quad (15.11-6)$$

第 16 章 衍射光学 II

16.1 本章引论

本章是第 15 章的延续，涵盖更多有关衍射光学的方法。本章包含 7 节。第 16.2 节讲述虚拟全息术的方法，通过改善制造技术，使 DOE 更容易处理。它也拉近了全息术与传统光学的距离。

16.3 节讨论了 14.8 节中谈到的用于设计二元 DOE 的凸集投影方法 (POCS)。16.4 节讨论了一种合并 POCS 的新方法——交叉迭代技术 (IIT)。在该方法中，子全息图 (subholograms) 被创建，并且子全息图的迭代设计是基于最小化每一迭代过程中的残余重构误差来实现的。

16.5 节结合 IIT 与 FFT 算法中用到的频率抽选性质，提出一种更好、更快的设计子全息图的策略，称作最优频率抽取交叉迭代技术 (ODIFIIT)。16.6 节进一步推广了 15.4 节中用于每一罗曼单元中幅值与相位量化的等幅罗曼法。POCS 法被用于子全息图的优化。16.6 节结合了罗曼法与 ODIFIIT 法，通过减小所需目标图像区域的方均差来获得可观的精度。

16.2 虚拟全息

我们知道大多数的全息图是信息冗余的，仅仅用全息图的一小部分就足以获得想要的信息。如果想要裸眼观察全息图，图像点的分辨率取决于瞳孔的尺寸，其为几个毫米数量级 [Born and Wolf, 1969]。我们期望从这般尺寸的全息图中获得尽可能多的信息。

当用某一类型的波，如声波或微波，记录全息图；用另一种波，如可见光，进行重构时，为避免图像失真，必须按照两种波长的比例改变全息图尺寸 [Smith, 1975]。然而，这是一个耗时且易出错的操作。

虚全息图定义为这样一种全息图，其为存在于空间中的另一被记录的全息图（实全息图）的像，而不是记录在介质上的全息图 [Ersoy, August 79]。虚全息图包含的信息为想要得到的信息，而实全息图包含的是变换信息。实全息图与虚全息图之间的转换可由图 16.1 所示的光学系统完成。

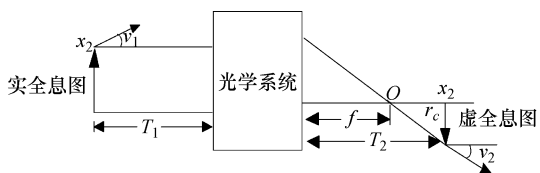


图 16.1 虚拟全息的系统框图

可以说实全息图和虚全息图都不与常规全息图完全相像。如果观察重构的实全息图，看到的是变换后的信息，则其可能不可辨认。虚全息图更像常规的全息图，但它并不记录在物理介质上。

16.2.1 相位的确定

如图 16.1 所示，从实全息图发出的平行于光轴的光线，汇聚到 O 点。由于实全息图上



的点与对应的虚全息图上的点之间的光程是相同的，因此在虚全息图上某点相对于另一点的相位取决于 O 点与该点之间的半径向量的长度 r_c ，如图 16.1 所示。

连接系统输入输出的传输矩阵 S 的形式为 [Gerrard and Burch, 1975]

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = S \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} \quad (16.2-1)$$

S 可被定义为

$$S = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \quad (16.2-2)$$

式中

$$A' = A + CT_2 \quad (16.2-3)$$

$$B' = (A + CT_2)T_1 + B + DT_2 \quad (16.2-4)$$

$$C' = C \quad (16.2-5)$$

$$D' = CT_1 + D \quad (16.2-6)$$

光学系统矩阵由 A, B, C, D 决定； T_1, T_2 是图 16.1 中所示的距离。为了生成图像， $B' = 0$ ，于是

$$M = \frac{x_2}{x_1} = A' = \frac{1}{D'} \quad (16.2-7)$$

以 v_1 角度入射到光学系统的平面波聚焦在

$$f = -\frac{A}{C} \quad (16.2-8)$$

平面处，横坐标 x_2 为

$$x_2 = (B + Df)v_1 \quad (16.2-9)$$

利用这些方程，可以直接计算图 16.1 中所示的 r_c 。

r_c 决定虚全息图上参考波的类型。实全息图上参考波的相位也传输到虚全息图上。因此，分别讨论实参考波和虚参考波是有意义的。该途径可以变化扩展为用数个光学系统产生数个虚全息图和参考波。最终的结果是将所有的参考波在最终的虚全息图上叠加。然而，每一参考波的相位取决于其在相应全息图上的位置。

考察三个例子。第一个为单透镜的情况。这种情况对应于

$$A = 1, B = 0, C = -\frac{1}{f}, D = 1 \quad (16.2-10)$$

于是

$$\frac{1}{T_1} + \frac{1}{T_2} = \frac{1}{f} \quad (16.2-11)$$

$$M = -\frac{T_2}{T_1} \quad (16.2-12)$$

第二个例子如图 16.2 所示的望远系统。这里 f 等于 ∞ ，于是虚参考波为平面波。利用焦距为 f_1, f_2 的两个透镜得到

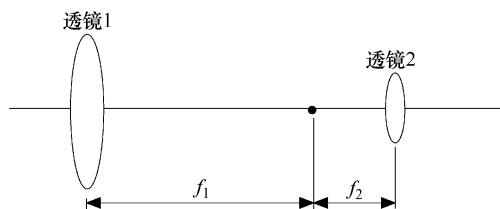


图 16.2 望远系统

$$A' = -\frac{1}{F} = -\frac{f_2}{f_1} \quad (16.2-13)$$



$$B' = f_1 + f_2 \quad (16.2-14)$$

$$C' = 0 \quad (16.2-15)$$

$$D' = -F \quad (16.2-16)$$

于是

$$M = -\frac{1}{F} = -\frac{f_2}{f_1} \quad (16.2-17)$$

$$T_2 = \frac{f_1 + f_2}{F} - \frac{T_1}{F^2} \quad (16.2-18)$$

可见 M 与实全息图坐标无关, 且如果 F 很大, T_2 将对 T_1 很敏感。

第三个例子是 13.3 节中谈到的无透镜傅里叶结构。当图 16.1 中的 O 点在像平面上时便可得到这种结构。如果 z_0 为虚全息图到像平面的距离, 则有

$$\frac{A}{C} = -(T_2 + z_0) \quad (16.2-19)$$

16.2.2 孔径效应

为了从虚全息图获得有效的衍射, 虚全息图孔径中传播的波越多越好。有趣的是, 在这种情况下孔径尺寸似乎没那么重要, 因为虚全息图不是被记录在物理介质上的。换言之, 若虚全息图孔径相互重叠, 那么远处的衍射效应可以解释为虚全息图上的点源发出的波在感兴趣的空间的相互干涉。

波的扩散可以用光线与两个参考平面的夹角 v_1 和 v_2 来讨论。由式 (16.2-2) 有

$$v_2 = C'x_1 + D'v_1 \quad (16.2-20)$$

在像平面上

$$C' = -\frac{1}{f} \quad (16.2-21)$$

$$D' = \frac{1}{M} \quad (16.2-22)$$

随着 f 和 M 的减小, 波的扩散增强。在望远系统中, $C' = 0$, 从一个孔径产生的波的扩散在输出时增加了 $1/M$ 。

为了确定虚全息图孔径的尺寸, 需要考虑放大效应和衍射效应。虚全息图孔径的尺寸 d_v 可写为

$$d_v = Md_r + D \quad (16.2-23)$$

式中, d_r 是实全息图孔径的尺寸; D 是由于光学系统的有限尺寸造成的衍射而带来的额外尺寸。

比如, 在望远系统中, D 可由下式估计 [Gerrard and Burch, 1975]:

$$D = \frac{2.44f\lambda}{d_A} \quad (16.2-24)$$

式中, d_A 是望远镜物镜的直径; f 是其焦距。

16.2.3 成像分析

成像分析可以用与 15.10 节单幅图像全息术相同的方式完成。假设不同参考波的贡献为



一个从点 (x_c, y_c, z_c) 发出的有效参考波。这常可利用傍轴近似来处理。如果 $(x_i, y_i, 0)$ 是虚全息图上某一采样点的坐标, (x_0, y_0, z_0) 是某一物点的坐标, 则下式确定了不同谐波图像位置:

$$-x_i \left(\frac{x_c}{r_c} + \frac{x_0}{r_0} \right) - y_i \left(\frac{y_c}{r_c} + \frac{y_0}{r_0} \right) + \left(\frac{x_i^2 + y_i^2}{2} \right) \left(\frac{1}{r_c} + \frac{1}{r_0} \right) = n\lambda + \frac{\phi\lambda}{2\pi} - F_{0c} \quad (16.2-25)$$

式中, ϕ 是常数相位。

且

$$F_{0c} = r_c + r_0 \quad (16.2-26)$$

$$r_0 = \pm (x_0^2 + y_0^2 + z_0^2)^{\frac{1}{2}} \quad (16.2-27)$$

式中, $+$ 、 $-$ 号用作区分物点相对于虚全息图的实、虚, 且

$$r_c = \pm (x_c^2 + y_c^2 + z_c^2)^{\frac{1}{2}} \quad (16.2-28)$$

式中, $+$ 、 $-$ 号用作区分参考波的焦点来自虚全息图之前、后。

将式 (16.2-25) 除以 M , 我们发现实全息图对应的物体的坐标为

$$r'_0 = \frac{r_0}{M} \quad (16.2-29)$$

$$x'_0 = \frac{x_0}{M} \quad (16.2-30)$$

$$y'_0 = \frac{y_0}{M} \quad (16.2-31)$$

$$z'_0 = (r'^2_0 - x'^2_0 - y'^2_0) \quad (16.2-32)$$

且参考波来源于

$$r'_c = \frac{r_c}{M} \quad (16.2-33)$$

$$x'_c = \frac{x_c}{M} \quad (16.2-34)$$

$$y'_c = \frac{y_c}{M} \quad (16.2-35)$$

$$z'_c = (r'^2_c - x'^2_c - y'^2_c) \quad (16.2-36)$$

且波长

$$\lambda' = \frac{\lambda}{M} \quad (16.2-37)$$

如果采用源于点 (x''_c, y''_c, z''_c) 、波长为 λ 的重构波, 不采用光学系统, 第 m 阶谐波将重构于点

$$\frac{1}{r''_0} = \frac{m}{M^2} \left[\frac{1}{r_0} + \frac{1}{r_c} \right] - \frac{1}{r''_c} \quad (16.2-38)$$

$$x''_0 = r''_0 \left\{ mM \left[\frac{x_0}{r_0} + \frac{x_c}{r_c} \right] - \frac{x''_c}{r''_c} \right\} \quad (16.2-39)$$

$$y''_0 = r''_0 \left\{ mM \left[\frac{y_0}{r_0} + \frac{y_c}{r_c} \right] - \frac{y''_c}{r''_c} \right\} \quad (16.2-40)$$



式 (16.2-38) ~ 式 (16.2-40) 表明还存在着另一个相对于预期物体的失真物体。与虚全息图的概念相似, 称其为虚物。这意味着我们可以从虚全息图生成实物, 或者从实全息图生成虚物。在大多数情况下利用虚全息图的概念看起来会更简单。然而, 在有些情况下用虚物的概念会有更大的优势。

假设用垂直平面波 ($x_c = y_c = 0$) 重构, 对于一阶谐波, 式 (16.2-38) ~ 式 (16.2-40) 可简化为

$$r_0'' = \frac{R}{M} r_0 \quad (16.2-41)$$

$$x_0'' = R x_0 \quad (16.2-42)$$

$$y_0'' = R y_0 \quad (16.2-43)$$

$$R = \frac{r_c}{M(z_0 + r_c)} \quad (16.2-44)$$

从最后的四个式子能看出计算上的优势, 即可采用卷积来计算全息图。在数字全息中, 通常全息图比图像要小得多。而卷积计算要求图像与全息图有相同的尺寸。用这种方法, 可以通过选择合适的 R 获得相同的尺寸。对于单透镜的情况, 可通过满足下式来实现:

$$R = \frac{f}{r_0 + r_c} \quad (16.2-45)$$

图 16.3 为这种情况下的不同全息图和物体。

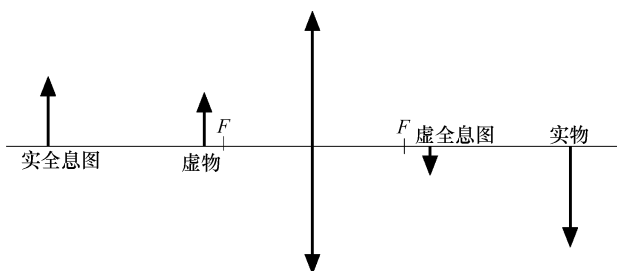


图 16.3 虚实全息图和物体

16.2.4 信息容量, 分辨率, 带宽及冗余

考虑虚全息图, 根据分辨率要求设定了其尺寸的下限。对于矩形孔径, 依据瑞利准则, 物点间的最小可分辨距离为 [Born and Wolf]:

$$h = \frac{z_0 \lambda}{L_v} \quad (16.2-46)$$

式中, L_v 是虚全息图的尺寸。

相应地, 虚全息图的尺寸应该足够大以获得理想的 h 。

同样需要考虑记录介质的空间频率限制。 x 方向的条纹间距可由下式估计:

$$\Delta_F = \frac{\lambda}{|M|} \left[\frac{x_i - x_c}{r_c} + \frac{x_i - x_o}{r_o} \right]^{-1} \quad (16.2-47)$$

式 (16.2-46) 和式 (16.2-47) 表明, 如果 $|M|$ 取得足够小, 当 Δ_F 增加到足够大时, 几乎所有记录介质都能用来制作全息图。然而, h 也会随之增大。相反地, 为了提高图像分辨率, 虚全息图能做得比实全息图大, 假设这时没有超过记录介质的频率限制。

对于某一特定的全息图尺寸, 有必要将冗余减小到期望的程度来记录信息, 特别是对于复用。以条纹的数目来讨论冗余似乎是有利的。如果获取想要的图像分辨率需要 N 个条纹, 则虚全息图尺寸可以选择为能覆盖 N 个条纹的大小。实全息图也包含 N 个条纹。这样, 记



录介质能达到所希望的有效利用。如果记录介质有如下空间带宽积 [Caulfield and Lu, 1970]:

$$SB = \frac{\text{全息图面积}}{\text{最小分辨单元的面积}} \quad (16.2-48)$$

且每一信号用到的平均空间带宽积为 SB_s , 则能记录的信号的数目为介质记录线性可加信号能力的 SB/SB_s 倍量级。直接记录全息图使得 SB/SB_s 量级为 1。因此, 实全息图的记录能力能通过这种以牺牲分辨率的方式得到有效利用。然而, 这需要诸如硬限幅之类的非线性记录技术。如果 SB_s 对于某一信号很小, 则可用像激光打印机这样的非精密设备制作数字全息图。这样做的另外一个优点是由于条纹较少, 计算尺寸缩小的虚全息图会快得多。

16.2.5 体积效应

式 (16.2-17) 和式 (16.2-18) 描述了望远系统中的变换。横向放大率与实全息坐标无关, 而 T_2 相对 T_1 的变化率为 $1/F^2$ 。如果 F 相对较大, 并且几个实全息面之间分开, 则分开的距离在虚全息图空间会减小 F^2 倍。因此, 对于实全息图的定位误差, 在虚全息图空间该误差的横向会减小 F 倍, 纵向会减小 F^2 倍。这意味着使用虚全息图概念与望远系统, 能在不同全息图之间、光学单元之间很容易地获得干涉。

作为一个简单的例子, 如图 16.4 所示, 两平面相距为 d , 考虑从点 x_1 和 x_2 发出的球面波在点 O 处的干涉。

由于 d 足够小时, 二次项总可以通过使用透镜来去掉, 因此可假定夫琅禾费近似是有效的。则基于垂直平面参考波, 两个光路的光程差可估计为

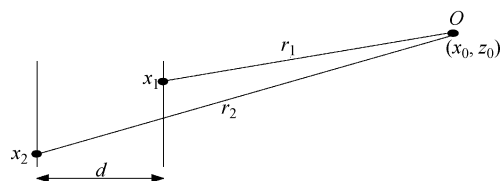


图 16.4 从相距 d 的两点发出的球面波之间的干涉

$$\Delta = \frac{x_0(x_2 - x_1)}{z_0} + \frac{d}{2} \left(\frac{x_0 - x_2}{z_0} \right)^2 \quad (16.2-49)$$

如果 x_0 比 x_2 大得多, 式 (16.2-49) 可简化为

$$\Delta = \frac{x_0(x_2 - x_1)}{z_0} + \frac{d}{2} \left(\frac{x_0}{z_0} \right)^2 \quad (16.2-50)$$

若 $d=0$, 则式 (16.2-50) 中第二项消失, 就得到了罗曼用于确定合成孔径位置的表达式 [Lohmann, 1970]。对不同的物点, 需要假设 x_0/z_0 近似为常数。但是, 这种假设可以通过找到 Δ 相对于 x_0/z_0 的稳态点而放宽到更大的范围, 即

$$x_2 - x_1 = -\frac{x_0}{z_0} d \quad (16.2-51)$$

$$\Delta = -\frac{d}{2} \left(\frac{x_0}{z_0} \right)^2 \quad (16.2-52)$$

可以进一步扩展该方法来制作数字体全息图。例如, 可用相距为 d 的几个全息面来量化相位, d 由式 (16.2-52) 确定。这样共轭像问题也得到解决。由于 d 和 $x_2 - x_1$ 能通过实全息图空间的调整来轻易控制, 实全息图空间的误差在虚全息图空间被极大地降低, 因此这样做是实际的。



16.2.6 波长变化和/或构建及重构中全息图尺寸的变化引起的失真

我们知道在全息显微术和声学全息术中，当波长变化和/或构建及重构中全息图尺寸不以相同的比率变化时，会造成横向和纵向放大率的不同。为解决这个问题提出了不同的方法，如像平面法 [Firth, 1972]。对于虚全息图，实际上能够不用匹配物理全息图尺寸，就可达到全息图尺寸与波长的匹配。然而，当用于虚全息图或实全息图重构的参考球面波与构建全息图所用参考球面波不同时，则即使不用匹配，也能减小失真。

如果 N 为全息图尺寸在构建与重构过程中改变的倍数， k_1 、 k_2 分别是构建与重构过程中的波数，与式 (16.2-25) 相似，成像方程可写为

$$k_1 \left[-\frac{x_{01}}{r_{01}}Nx - \frac{y_{01}}{r_{01}}Ny + \frac{N^2(x^2 + y^2)}{2r_{01}} \right] = k_2 \left[-x \left(\frac{x_{02}}{r_{02}} - \frac{x_c}{r_c} \right) - y \left(\frac{y_{02}}{r_{02}} - \frac{y_c}{r_c} \right) + \frac{x^2 + y^2}{2} \left(\frac{1}{r_{02}} - \frac{1}{r_c} \right) \right] \quad (16.2-53)$$

式中， x_{01} 、 y_{01} 和 z_{01} 是物坐标； x_{02} 、 y_{02} 和 z_{02} 是像坐标。

且

$$r_{01} = [x_{01}^2 + y_{01}^2 + z_{01}^2]^{\frac{1}{2}} \quad (16.2-54)$$

$$r_{02} = [x_{02}^2 + y_{02}^2 + z_{02}^2]^{\frac{1}{2}} \quad (16.2-55)$$

令对应的项相等，得到

$$\frac{1}{r_{02}} = \left(\frac{k_1 N^2}{r_{01}} + \frac{1}{r_c} \right) \frac{1}{k_2} \quad (16.2-56)$$

$$x_{02} = r_{02} \left(\frac{k_1 N x_{01}}{r_{01}} + \frac{x_c}{r_c} \right) \frac{1}{k_2} \quad (16.2-57)$$

$$y_{02} = r_{02} \left(\frac{k_1 N y_{01}}{r_{01}} + \frac{y_c}{r_c} \right) \frac{1}{k_2} \quad (16.2-58)$$

如果满足

$$r = \frac{r_{01}}{N} \quad (16.2-59)$$

$$x_{02} = \frac{x_{01}}{N} \quad (16.2-60)$$

$$y_{02} = \frac{y_{01}}{N} \quad (16.2-61)$$

$$z_{02} = \frac{z_{01}}{N} \quad (16.2-62)$$

则各种放大倍数相等。

正如 16.2.3 节中讨论的，也可以采用由许多对应于不同全息图的参考波叠加而生成的有效虚拟参考波，来减小失真以及扫描像场的不同部分。

16.2.7 实验

采用上述方法的第一组数字全息图是用第 15 章中讨论到的扫描电子显微镜制作的。对虚全息图做了编码全息图的所有计算。然后，利用生成物理实全息图的光学系统参数，将虚



全息图转换到计算机中。单图全息技术被用来编码全息图。此处，每个全息图孔径的位置的选取依据下式：

$$\phi(x_i, y_i) + kr_{oi} = 2\pi n + \phi_0 \quad (16.2-63)$$

式中， $\phi(x_i, y_i)$ 是虚拟孔径 $(x_i, y_i, 0)$ 处参考波的相位； n 为整数； ϕ_0 是物点 (x_0, y_0, z_0) 的所需的相位。

且

$$r_{oi} = [(x_0 - x_i)^2 + (y_0 - y_i)^2 + z_0^2]^{\frac{1}{2}} \quad (16.2-64)$$

其中 $\phi(x_i, y_i, 0)$ 由下式给出

$$\phi(x_i, y_i, 0) = kr_{ci} \quad (16.2-65)$$

$$r_{ci} = [(x_c - x_i)^2 + (y_c - y_i)^2 + z_c^2]^{\frac{1}{2}} \quad (16.2-66)$$

式中， (x_c, y_c, z_c) 是图 16.1 中点 O 的坐标。

例如，对于单透镜轴上的光，坐标可选为

$$x_c = 0 \quad (16.2-67)$$

$$y_c = 0 \quad (16.2-68)$$

$$z_c = -fM \quad (16.2-69)$$

式中， M 是所需放大率； f 是透镜系统的焦距。实全息图上的孔径位置为 x_i/M ， y_i/M 。

图 16.5 为按这种方法生成的第一组全息图的重构图像。选择的物体为倾斜于 z 方向的圆。这就是为什么它看起来像椭圆——这表明物点的三维本质。实全息图尺寸为 $2\text{mm} \times 2\text{mm}$ 。利用焦距为 20mm 的透镜将其缩小 $1/4$ 生成虚全息图。图 16.1 中的距离 $T_2 - f$ 选为 5mm 。

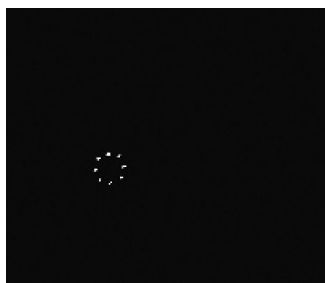


图 16.5 第一组虚全息图的重构图像

可以用 SEM 系统生成大小为 $2\text{mm} \times 2\text{mm}$ 的单个全息图，单个全息图按边沿拼接起来形成总尺寸达 $7.4\text{cm} \times 7.4\text{cm}$ 的全息图。每次移动定位都会有 $\pm 5\mu\text{m}$

的误差。因为每个全息图自身就是一个窗口，与非散射照明的问题类似，从每个独立全息图发出的光被导向不同的方向。即使可以忽略定位误差，且可以制成一幅大的单独的全息图，这也不会增加信息密度，因为每次人眼可见的仅仅是从 $2\text{mm} \times 2\text{mm}$ 大小的区域发出信息的一小部分。

用这种方法，尽管会牺牲图像分辨率，但可以将整个虚全息图阵列减小到几毫米的尺寸，使得来自不同小全息图的所有信息都对人眼可见。

为了展示这种方法的效果，16 个相邻的全息图组成 4×4 的阵列。通过用 50mm 物镜将实全息图阵列缩小 $1/4$ ，得到尺寸为 $2\text{mm} \times 2\text{mm}$ 的虚全息图阵列。选取图 16.1 中的距离 $T_2 - f$ 为 12.5mm 。每个全息图生成三维立方体的一条边，或一条对角线，或所有的角点以及中间点。图 16.6 为这样一幅全息图的输出。照明所有的全息图就获得了整个图像，如图 16.7 所示。图 16.7 中，由于所有的全息图都对角点和中心点有贡献，这些物点比其他物点显得更加密集。透过虚全息图阵列所在的区域，可以看到空间中完整的立方体。

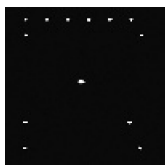


图 16.6 重构第二组虚全息图生成的部分三维立方体



图 16.7 重构全息图集生成的三维立方体

16.3 用于二元 DOE 设计的 POCS 方法

在讨论本节的方法以及本章接下来的几节中，会用到图 15.1 所示的傅里叶变换系统。经抽样的全息图由离散点阵列组成。由 $M \times N$ 个尺寸为 $\Delta v_x \times \Delta v_y$ 的像素构成的全息图的传递函数可用累加的方式表示：

$$G(v_x, v_y) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} H(k, l) \text{rect}\left(\frac{v_x - k\Delta v_x}{\Delta v_x}\right) \text{rect}\left(\frac{v_y - l\Delta v_y}{\Delta v_y}\right) \quad (16.3-1)$$

式中， $H(k, l)$ 是点 (k, l) 的二元透过率。

在观察平面，重构的图像由透过率函数的傅里叶变换给出

$$\begin{aligned} g(x, y) &= \iint G(v_x, v_y) e^{2\pi j[xv_x + yv_y]} dv_x dv_y \\ &= \Delta v_x \Delta v_y \text{sinc}[\Delta v_x x] \text{sinc}[\Delta v_y y] \sum \sum H(k, l) \exp[2\pi j(kx\Delta v_x + ly\Delta v_y)] \end{aligned} \quad (16.3-2)$$

忽略求和外的两个常数以及两个 sinc 因子，重构的图像可近似为透过率的二维逆离散傅里叶变换 (2D-IDFT)。

利用第 14.8 节中讨论的 POCS 法优化全息图的设计。设观察平面和计算机制全息图 (CGH) 平面均为 $M \times N$ 维，观察平面 $h(m, n)$ 与 CGH 平面 $H(k, l)$ 的关系可由下列离散傅里叶变换对给出

$$h(m, n) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} H(k, l) W_M^{mk} W_N^{nl} \quad (16.3-3)$$

式中， $0 \leq m \leq M-1$ ， $0 \leq n \leq N-1$ ，且

$$H(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} h(m, n) W_M^{-mk} W_N^{-nl} \quad (16.3-4)$$

式中， $0 \leq k \leq M-1$ ， $0 \leq l \leq N-1$ ，且

$$W_u = \exp(j2\pi/ u) \quad (16.3-5)$$

POCS 法的目的是生成 CGH，使得其重构图像与所需目标图像最相像。

对位于观察平面上区域 R 内给定的理想图像 $f(m, n)$ ，POCS 法步骤为：

- (1) 利用式 (16.3-4)，由 $f(m, n)$ 计算 $F(k, l)$ 。
- (2) 按下式由 $F(k, l)$ 生成二元透过率值 $H(k, l)$ ：

$$H(k, l) = \begin{cases} 1 & \text{若 } \text{Re}[F(k, l)] \geq 0 \\ 0 & \text{其他} \end{cases} \quad (16.3-6)$$

- (3) 利用式 (16.3-2)，找到重构图像 $h(m, n)$ 。重构图像的精确度由 R 区域内 $f(m, n)$



与 $h(m,n)$ 之间的方均差 (MSE) 衡量。方均差定义为 [Seldowitz]

$$MSE = \frac{1}{MN} \sum_{(m,n) \in R} |f(m,n) - \lambda h(m,n)|^2 \quad (16.3-7)$$

式中, λ 是缩放因子。

如果 λ 满足下式, $h(m,n)$ 有最小 MSE,

$$\lambda = \frac{\sum_{(m,n) \in R} f(m,n) h^*(m,n)}{\sum_{(m,n) \in R} |h(m,n)|^2} \quad (16.3-8)$$

(4) 定义一个新的输入图像 $f'(m,n)$, 使得

(a) 在区域 R 之外, $f'(m,n)$ 与 $h(m,n)$ 相等。

(b) 在区域 R 之内, $f'(m,n)$ 有原始图像 $f(m,n)$ 的幅度和 $h(m,n)$ 的相位。

(5) 令 $f(m,n) = f'(m,n)$, 回到第 (1) 步。

(6) 重复 (1) ~ (5) 步, 直到 MSE 收敛或满足特定条件。

16.4 交叉迭代技术 (IIT)

本节讨论的交叉迭代技术 (IIT) 可被应用到现存的任意 DOE 制备方法中, 以提高其性能 [Ersoy, Zhuang, Brede]。首先引入交叉技术 (Interlacing Technique, IT), 然后将它推广到 IIT。IT 将整个全息平面分成一组子全息图。一个子全息图由一组单元 (或点, 称作“块”) 组成。所有的子全息图均分开设计, 然后交叉, 生成一幅全息图。图 16.8 和图 16.9 为两个子全息图交叉方案的两个例子。



图 16.8 两个子全息图的交叉方案 1

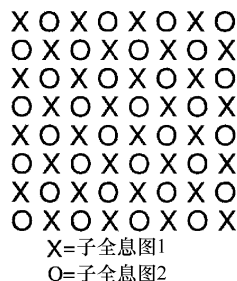


图 16.9 两个子全息图的交叉方案 2

IT 法中, 一旦整个全息图被分成更小的子全息图, 首个子全息图就用于重构所需的图像 $f(m,n)$, 而实际重构图像为 $h_1(m,n)$ 。由于子全息图不能完美重构所需目标图像, 因此存在误差, 误差图像可定义为

$$e_1(m,n) = f(m,n) - \lambda_1 h_1(m,n) \quad (16.4-1)$$

为了消除该误差, 第二个子全息图被设计为 $e_1(m,n)/\lambda_1$ 。由于傅里叶变换是线性运算, 对两个子图的总的重构仅仅是两个单独重构之和。如果第二个子全息图完美, 且缩放因子与 λ_1 匹配, 两重构图像之和即为 $f(m,n)$ 。然而, 与首个子全息图类似, 第二个子全息图也会有误差。于是, 第三个子全息图用于减小前两个子全息图的剩余误差。因此, 每个子全息图被设计用于减小所需目标图像与所有之前重构模块叠加后的图像之间的误差。重复本过程直



到设计完所有的子全息图。

每个子全息图由 POCS 算法（也可以用其他方法）次优化（非最优）生成。然而，即使一次性用到了所有子全息图，总的 CGH 也许还是没有达到最优结果。要解决这个问题，可采用 IT 方法的推广——交叉迭代（IIT）方法。

IIT 是 IT 的迭代形式，用于获得最小 MSE [Ersoy, Zhuang and Brede, 1992]。第 j 次迭代中的第 i 个子全息图的重构图像可记为 $h_i^j(m, n)$ 。在用 IT 法设计完每个子全息图后，整幅全息图 $h_f(m, n)$ 的重构图像有最终误差 $e_f(m, n)$ 。为应用交叉迭代技术（IIT），需要生成对子全息图的一次新的推扫。在新的推扫中，首个子全息图对应的新的理想图像 $f'(m, n)$ 可选为

$$f'(m, n) = h_1^1(m, n) - \frac{e_f(m, n)}{\lambda_f} \quad (16.4-2)$$

式中

$$e_f(m, n) = f(m, n) - \lambda_f h_f(m, n) \quad (16.4-3)$$

λ_f 是生成最后的子全息图之后用到的缩放因子。一旦首个全息图被重建，就能算出整幅全息图的误差图像，包括由首个子全息图生成的新的重构图像。类似地，第二个子全息图被设计为重构

$$h_2^1(m, n) - \frac{e'_f(m, n)}{\lambda'_f}$$

$e'_f(m, n)$ 为更新的误差。本过程一直持续到达到收敛为止，即前后重构图像之间的绝对误差

$$\Delta_j = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |h_i^j(m, n) - h_i^{j-1}(m, n)| \quad (16.4-4)$$

达到可以忽略的值，或者对所有子全息图保持不变。通过采用 IIT 法，收敛趋向于从局部最小 MSE 向全局最小 MSE 转变，至少是更高阶的最小 MSE 。

16.4.1 IIT 算法实验

在用 IIT 实施的所有实验中，可以看到 IIT 与另一算法（如 POCS 算法）结合使用时，可以提高重构结果的质量。用图 16.10 所示的图像进行一组特定实验。



图 16.10 用于 IIT 实验的猫脑图像

POCS 被用作相应的迭代优化算法。表 16.1 显示了随着子全息图数量的增加，重构的 MSE 是如何减小的。

表 16.1 重构的 MSE 与子全息图数目的函数关系

k	MSE	改善率/%
0	3230.70	0
1	2838.76	12.13
2	2064.35	36.10
3	1935.00	40.11

图 16.11 为用 IIT 法生成的猫脑图像的二元全息图。图 16.12 显示了误差是如何随着迭代次数的增加而减小的。图 16.13 为相应的 He-Ne 激光光束重构猫脑图像。

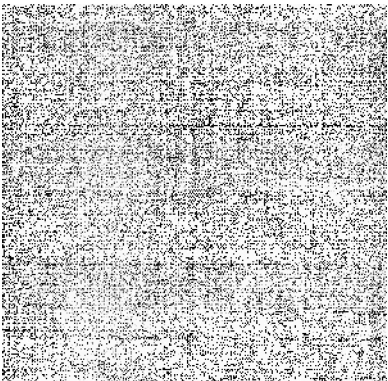


图 16.11 用 IIT 法生成的猫脑图像的二元全息图

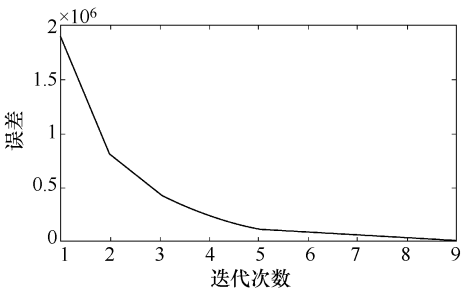


图 16.12 IIT 设计中误差与迭代次数的函数关系

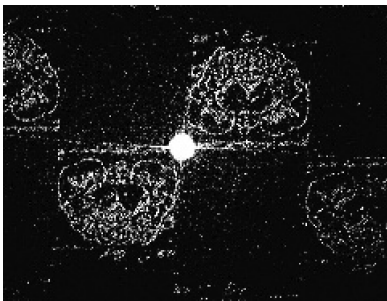


图 16.13 He-Ne 激光光束重构猫脑图像

16.5 最优频率抽取交叉迭代技术（ODIFIIT）

最优频率抽取交叉迭代技术（ODIFIIT）是为优化 IIT 过程而发展起来的 [Zhuang and Ersoy, 1995]。ODIFIIT 利用了将全息图分成子全息图时快速傅里叶变换（FFT）的频率抽取性质，较 IIT 具有两大优势。它通过减少傅里叶变换及其逆变换的维数来降低计算时间。直接考察所需目标图像区域 R 内的图像，这样能更有效地设计每个子全息图，因为只有感兴趣的数据的贡献才会被考虑。



图 16.14 为 ODIFIIT 法的重构平面的模型结构。所需的幅度 $f_0(m, n)$ 是尺寸为 $A \times B$ 的实值阵列。优化过程中任意时刻的目标图像 $f(m, n)$ 是 $f_0(m, n)$ 乘以一个浮动相位, 该相位由区域 R 内当前的重构图像的相位决定。 $f(m, n)$ 置于区域 R 内, 起点为 (M_1, N_1) 。由于 CGH 的透过率是实值的, 重构图像的厄米共轭处于 R^+ 区域内。由于二元 CGH 的单元振幅值等于 1, 有必要对目标图像进行像素值的缩放, 这样就归一化其 DFT, 使它能够直接与重构图像 $h(m, n)$ 进行对比。

整个的 CGH 被分成 $\mu \times v$ 个子全息图 (或块), $\mu = M/A$, $v = N/B$ 。如果 M 、 N 、 A 、 B 均为 2 的幂, 则 μ 、 v 必须均为整数。利用频率抽取 [Brigham, 1974], 这些块都是交错的, 这样第 (α, β) 个块由单元 $(\alpha\mu + k, \beta v + l)$ 组成, 此处 $0 \leq k \leq A-1$, $0 \leq l \leq B-1$, $0 \leq \alpha \leq \mu-1$, $0 \leq \beta \leq v-1$ 。图 16.15 为 $\mu = v = 2$ 时的一个例子。

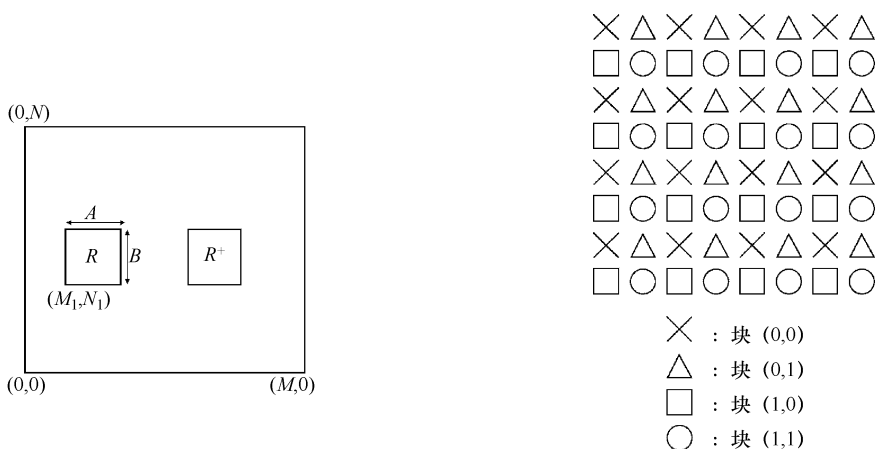


图 16.14 ODIFIIT 法的重构平面的模型结构

图 16.15 ODIFIIT 法中子全息图的交错扫描, $\mu = v = 2$

定义 $H(k, l)$ 为所有子全息图的累加, 重构图像的表达式变为

$$\begin{aligned} h(m, n) &= \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} H(k, l) W_M^{mk} W_N^{nl} \\ &= \frac{1}{\mu v} \sum_{\alpha=0}^{\mu-1} \sum_{\beta=0}^{v-1} \left[\frac{1}{AB} \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} H(\alpha\mu + k, \beta v + l) W_A^{mk} W_B^{nl} \right] W_M^{m\alpha} W_N^{n\beta} \end{aligned} \quad (16.5-1)$$

式中, $0 \leq m \leq M$; $0 \leq n \leq N$ 。

用 $m + M_1$ 、 $n + N_1$ 分别取代 m 、 n 来计算区域 R 内的重构图像, 设 m 、 n 仅在图像区域取值:

$$\begin{aligned} h(m + M_1, n + N_1) &= \frac{1}{\mu v} \sum_{\alpha=0}^{\mu-1} \sum_{\beta=0}^{v-1} \left[\frac{1}{AB} \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} H(\alpha\mu + k, \beta v + l) W_A^{(m+M_1)k} W_B^{(n+N_1)l} \right] W_M^{(m+M_1)\alpha} W_N^{(n+N_1)\beta} \end{aligned} \quad (16.5-2)$$

式中, $0 \leq m \leq A-1$; $0 \leq n \leq B-1$ 。

设第 (α, β) 个子全息图的尺寸为 $A \times B$ 的逆离散傅里叶变换

$$\begin{aligned} h_{\alpha, \beta}(m, n) &= \text{IDFT}_{AB} [H(\alpha\mu + k, \beta v + l)]_{m, n} \\ &= \frac{1}{AB} \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} H(\alpha\mu + k, \beta v + l) W_A^{mk} W_B^{nl} \end{aligned} \quad (16.5-3)$$



式中, $0 \leq \alpha \leq \mu - 1$; $0 \leq \beta \leq v - 1$; $0 \leq m \leq A - 1$; $0 \leq n \leq B - 1$ 。

利用大小为 $A \times B$ 的 IDFT, 区域 R 内的重构图像变为

$$h(m + M_1, n + N_1) = \frac{1}{\mu v} \sum_{\alpha=0}^{\mu-1} \sum_{\beta=0}^{v-1} h_{\alpha, \beta}(m + M_1, n + N_1) W_M^{(m+M_1)\alpha} W_N^{(n+N_1)\beta} \quad (16.5-4)$$

式中, $0 \leq m \leq A - 1$, $0 \leq n \leq B - 1$ 。

式 (16.5-4) 隐含假定 $h_{\alpha, \beta}(m + M_1, n + N_1)$ 中的 $m + M_1$ 和 $n + N_1$ 分别能整除 A 和 B 。式 (16.5-4) 以所有子全息图的大小为 $A \times B$ 的 IDFT 的形式, 给出了区域 R 的重构图像的表达式。从式中可以看出, 由第 (α, β) 个子全息图生成的区域 R 内的重构图像由下式给出:

$$h'_{\alpha, \beta}(m + M_1, n + N_1) = \frac{1}{\mu v} h_{\alpha, \beta}(m + M_1, n + N_1) W_M^{(m+M_1)\alpha} W_N^{(n+N_1)\beta} \quad (16.5-5)$$

式 (16.5-5) 是第 (α, β) 个块的 IDFT 乘以合适的相位因子, 并除以 μv 。

定义一个稍后将用到的阵列:

$$\tilde{h}_{\alpha, \beta}(m + M_1, n + N_1) = h(m + M_1, n + N_1) - h'_{\alpha, \beta}(m + M_1, n + N_1) \quad (16.5-6)$$

此为除第 (α, β) 个子全息图外的所有子全息图在区域 R 内的重构图像。

反过来, 给定区域 R 的目标图像, 可以算出透过率。从式 (16.5-2) 有

$$H(k, l) = \sum_{m=0}^{A-1} \sum_{n=0}^{B-1} h(m + M_1, n + N_1) W_M^{-(m+M_1)k} W_N^{-(n+N_1)l} \quad (16.5-7)$$

式中, $0 \leq k \leq M - 1$; $0 \leq l \leq N - 1$ 。

与之前一样, 将 $H(k, l)$ 分成 $\mu \times v$ 个块得

$$\begin{aligned} & H(\mu k + \alpha, v l + \beta) \\ &= \sum_{\alpha=0}^{\mu-1} \sum_{\beta=0}^{v-1} \left[\sum_{m=0}^{A-1} \sum_{n=0}^{B-1} h(m + M_1, n + N_1) W_M^{-(m+M_1)\alpha} W_N^{-(n+N_1)\beta} W_A^{-(m+M_1)k} W_B^{-(n+N_1)l} \right] \\ &= W_A^{-M_1 k} W_B^{-N_1 l} \sum_{\alpha=0}^{\mu-1} \sum_{\beta=0}^{v-1} \text{DFT}_{AB} [h(m + M_1, n + N_1) W_M^{-(m+M_1)\alpha} W_N^{-(n+N_1)\beta}]_{k, l} \end{aligned} \quad (16.5-8)$$

式中, $0 \leq k \leq A - 1$; $0 \leq l \leq B - 1$; $0 \leq \alpha \leq \mu - 1$; $0 \leq \beta \leq v - 1$ 。

因此, 生成区域 R 中图像 $h(m + M_1, n + N_1)$ 的子全息图 (α, β) 的透过率由下式给出:

$$H(\mu k + \alpha, v l + \beta) = W_A^{-M_1 k} W_B^{-N_1 l} \text{DFT}_{AB} [h(m + M_1, n + N_1) W_M^{-(m+M_1)\alpha} W_N^{-(n+N_1)\beta}]_{k, l} \quad (16.5-9)$$

式中, $0 \leq k \leq A - 1$, $0 \leq l \leq B - 1$ 。

利用式 (16.5-5) 和式 (16.5-9), 可以计算每一单独子全息图在区域 R 内的重构图像, 或者给出区域 R 内的目标图像, 就能确定重构目标图像所需的透过率。因此, 现在便可用 IIT 设计 CGH 了。

设 $f_0(m + M_1, n + N_1)$ 为大小为 $A \times B$ 的目标图像, $0 \leq m \leq A - 1$, $0 \leq n \leq B - 1$, ODIFIIT 算法可总结如下:

(1) 定义参量 M 、 N 、 A 、 B 、 M_1 和 N_1 , 求出 μ 、 v 。然后将整幅全息图分成 $\mu \times v$ 个交错子全息图。



(2) 生成在 0、1 之间随机取值的 $M \times N$ 初始全息图。

(3) 计算整幅全息图的 $M \times N$ IDFT。区域 R 内的重构图像是区域 R 内的点，即 $h(m + M_1, n + N_1)$ ， $0 \leq m \leq A - 1$ ， $0 \leq n \leq B - 1$ 。

(4) 按照 POCS 法，将每一点 $h(m + M_1, n + N_1)$ 的相位赋予 $f_0(m + M_1, n + N_1)$ 的幅值，就得到了理想图像 $f(m + M_1, n + N_1)$ 。于是

$$f(m + M_1, n + N_1) = f_0(m + M_1, n + N_1) \exp(j\phi_{m+M_1, n+N_1}) \quad (16.5-10)$$

式中， $\phi_{m+M_1, n+N_1} = \arg\{h(m + M_1, n + N_1)\}$ 。

(5) 用式 (16.3-5) 找到最优参数 λ 。

(6) 用式 (16.5-3)、式 (16.5-5) 和式 (16.5-6)，找到 $\tilde{h}_{\alpha, \beta}(m + M_1, n + N_1)$ 。这是除第 (α, β) 个子全息图外所有子全息图在区域 R 内的重构图像。

(7) 求出第 (α, β) 个子全息图用于重构时的误差图像为

$$e(m + M_1, n + N_1) = \frac{f(m + M_1, n + N_1)}{\lambda} - \tilde{h}_{\alpha, \beta}(m + M_1, n + N_1) \quad (16.5-11)$$

它等于 IIT 法中的误差图像。

(8) 利用式 (16.5-9)，找到重构误差图像的当前块的透过率 $E(uk + \alpha, vl + \beta)$ 。

(9) 当前块的二元透过率设计为

$$H(uk + \alpha, vl + \beta) = \begin{cases} 1 & \text{Re}[E(uk + \alpha, vl + \beta)] \geq 0 \\ 0 & \text{其他} \end{cases} \quad (16.5-12)$$

(10) 找到当前块在区域 R 内的新的重构图像 $\tilde{h}_{\alpha, \beta}(m + M_1, n + N_1)$ 。

(11) 将 $h'_{\alpha, \beta}(m + M_1, n + N_1)$ 与 $\tilde{h}_{\alpha, \beta}(m + M_1, n + N_1)$ 相加得到新的重构图像 $h(m + M_1, n + N_1)$ 。

(12) 根据式 (16.5-10)，用 $h(m + M_1, n + N_1)$ 更新 $f(m + M_1, n + N_1)$ 。

(13) 重复 (7) ~ (12) 步，直到当前块的每一点的透过率收敛。

(14) 用新得到的透过率更新整幅全息图。

(15) 保持 λ 一致，对所有子全息图重复 (3) ~ (14) 步 [第 (5) 步除外]。

(16) 设计完所有的块之后，从式 (16.3-7) 中计算 MSE 。

(17) 重复 (3) ~ (16) 步直到 MSE 收敛， MSE 收敛表明对于当前 λ 已设计出最佳 CGH。

16.5.1 ODIFIIT 法的实验

用 ODIFIIT 法设计黑白字母 E 及 Lena 图的 DOE，所用到的两幅图像也曾用于测试罗曼法。两幅图像分别如图 15.3 和图 15.8 所示。同时也用到了更高分辨率的 256 像素 \times 256 像素的灰度图，如图 16.16 所示。

图 16.17 ~ 图 16.19 是 ODIFIIT 全息图的计算机重构图像。

用 ODIFIIT 设计的所有全息图都使用了图 16.15 所示的交叉模式。对子全息图进行交叉处理有多种不同的方法，但是，根据实验观察到，这种交叉方法能最小化 MSE ，因此能得到最佳结果。在该方案中，每个全息图被分成 4×4 的子全息图。这样在整幅图像之中，第



一幅图、第二幅图、第三幅图分别包含 16 像素 × 16 像素、128 像素 × 128 像素、256 像素 × 256 像素的理想图像区域。



图 16.16 ODIFIIT 测试实验中用到的第三幅图像

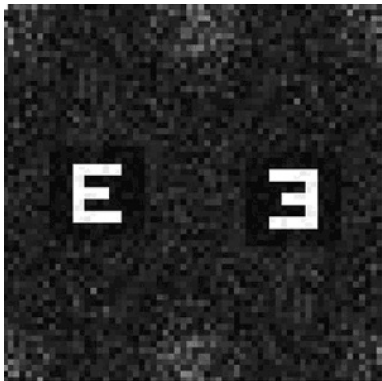


图 16.17 从 ODIFIIT 全息图得到的第一幅图像的重构图像

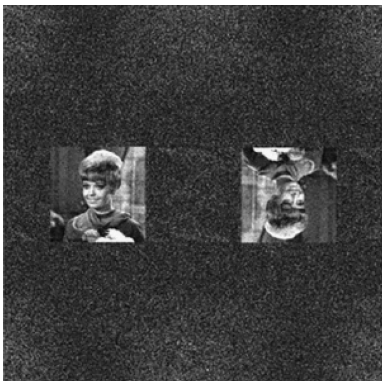


图 16.18 从 ODIFIIT 全息图得到的第二幅图像的重构图像

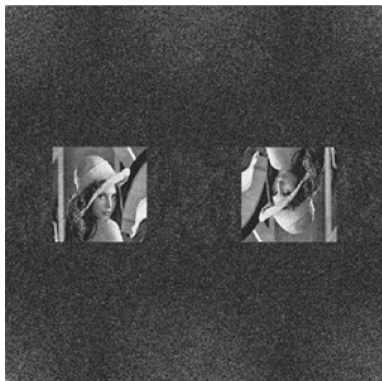


图 16.19 从 ODIFIIT 全息图得到的第三幅图像的重构图像

这些结果显示了对理想图像非常精确的重构。与第 15.6 节中讨论到的罗曼法的结果相比，ODIFIIT 法对高分辨率灰度图像的重构质量有了很大的提高。与用罗曼法设计的 DOE 类似，用 ODIFIIT 法设计的透过率函数也可以编码为 $(-1,1)$ ，以制作二元相位全息图。这可以进一步推广到用多级相位量化编码全息图。这对应多级相位单元。例如，可以用四级相位量化编码全息图。于是，全息图中的每一点将有 0 、 $\pi/2$ 、 π 或 $3\pi/2$ 的相移。由于全息图函数对每一点的实际理想相位有更接近的估计，因此生成的重构图像有更小的 MSE 。

16.6 复合的罗曼-ODIFIIT 法

由于在实际应用中，二元罗曼法中用到的迂回相位法不能精确编码幅度和相位，在重构过程中总会存在一些固有的误差。因此，IIT 和 ODIFIIT 中的迭代优化应该能有效减小最终的重构误差。出于这个原因，联合罗曼编码方案和交叉技术可以创建一种设计 DOE 的更好



方法 [Kuhl, Ersoy]。称这种组合方法为 Lohmann-ODIFIIT 法, 或者 LM-ODIFIIT 法。在本方法中, 用罗曼单元编码每一子全息图的所需的幅度和相位, 但将全息图分成与 ODIFIIT 法类似的交叉子全息图。

如第 15.3 节中所讨论, 为简化罗曼全息图的设计做了三个近似: (a) $\text{sinc}[c\delta v(x + x_0)] \approx \text{常量}$, (b) $\text{sinc}(yW_{nm}\delta v) \approx 1$, (c) $\exp[2\pi j(xP_{nm}\delta v)] \approx 1$ 。这些近似对重构图像的影响取决于几个因素。

$\text{sinc}[c\delta v(x + x_0)]$ 产生一个 x 方向上正比于图像平面中心的距离的强度而骤降。近似 (a) 认为 sinc 因子在图像区域内几乎为常数。小的孔径尺寸 c 产生的强度骤降更小。但是, 这也会降低图像的亮度。

$\text{sinc}(yW_{nm}\delta v)$ 表明会产生一个与 (a) 类似的强度骤降, 但是在 y 方向, $\text{sinc}(yW_{nm}\delta v)$ 会引起幅值透过率的轻微降低。

相移 $\exp[2\pi j(xP_{nm}\delta v)]$ 会造成图像平面随 x 变化的相位误差, 它取决于 x 和 P , 取值在 $0 \sim \pi/(2M)$ 之间。忽略掉相位因子 $\exp[2\pi j(xP_{nm}\delta v)]$ 会造成图像质量恶化, 就好像存在因不恰当孔径移位而造成相位误差一样 [Lohmann, 1970]。

用于补偿 sinc 在 x 方向上骤降的方法是将所需目标图像除以 $\text{sinc}[c\delta v(x + x_0)]$ 。这样目标图像 $f(x, y)$ 变为

$$\frac{f(x, y)}{\text{sinc}[c\delta v(x + x_0)]}$$

在 y 方向不能进行同样的操作, 因为 $\text{sinc}(yW_{nm}\delta v)$ 依赖于基于目标图像的孔径参数 W_{nm} 。在 Lohmann-ODIFIIT 法和量化的 LM-ODIFIIT 法中, sinc 因子按如下方式补偿: 目标图像首先除以影响 x 方向的 sinc 因子, 设计全息图。然后, 计算所有孔径 y 方向的 sinc 因子, 累加以确定对输出图像的影响。接着, 目标图像除以 y 方向的 sinc 因子。下一步, 再次设计全息图, 计算新孔径高度对输出的影响因子, 原始目标图像除以这个新的因子。重复该过程直到重构图像不再改变或者满足收敛条件。

设计全息图时, 相移 $\exp[2\pi j(xP_{nm}\delta v)]$ 由逐次迭代补偿。

16.6.1 LM-ODIFIIT 法的计算机实验

图 16.20 是由图 15.3 的黑白 E 图的 LM-ODIFIIT 全息图经计算机重构后的图像 [Kuhl, Ersoy]。图 16.21 为从 E 图的四级相位量化 LM-ODIFIIT 法的全息图的重构图像。图 16.22 为 Lena 图的基于 LM-ODIFIIT 法的全息图的重构图像。用 16×16 灰度阶图, 对比了罗曼法、ODIFIIT 法及 LM-ODIFIIT 法的输出结果如图 16.23 所示。

接下来要讨论的是基于黑白 E 图的重构结果的一张表。表 16.2 显示了计算机实验二元 DOE 的方均差和衍射效率结果。方均差表示在理想图像区域内理想图像与重构图像之间的差异, 衍射效率度量了衍射到理想图像区域的光占入射光的比例。这张表包括多种方法的对比结果, 包括 OD-

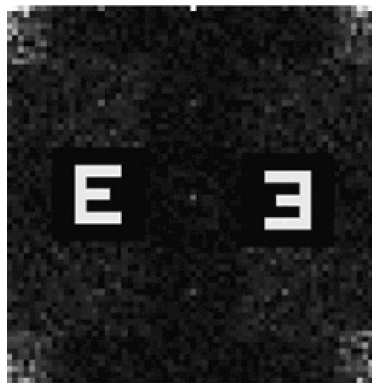


图 16.20 基于 LM-ODIFIIT 法的全息图的重构图像

IFIIT 法、罗曼法 (LM)、对每一单元运用等幅罗曼法 (LMCA)、采用罗曼编码的 ODIFIIT 法 (LM-ODIFIIT) 和单元等幅的 LM-ODIFIIT 法 (LMCA-ODIFIIT)。所有结果都是针对二元幅值全息图的。罗曼法的整个过程中， $c = 1/2$ ， $M = 1$ 。为了做对比，将罗曼法中的 MSE 归一化到 1，这样所有其他的 MSE 值都是相对罗曼法的。

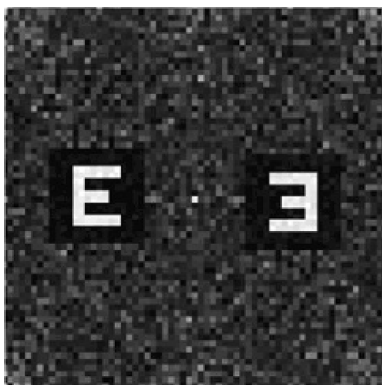


图 16.21 四级相位量化 LM-ODIFIIT 法的重构图像



图 16.22 基于 LM-ODIFIIT 法的重构图像

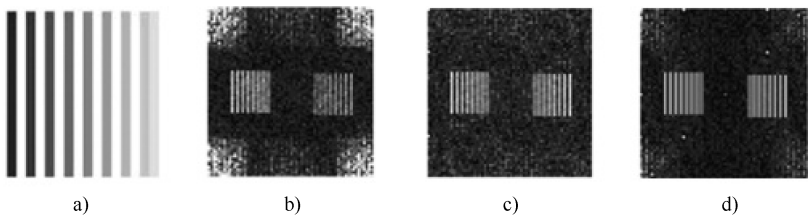


图 16.23 对比了罗曼法、ODIFIIT 法及 LM-ODIFIIT 法的输出结果
a) 原图 b) 罗曼法重构图 c) ODIFIIT 重构图 d) LM-ODIFIIT 重构图

表 16.2 二元 DOE 的方均差 MSE 和衍射效率

方 法	MSE	衍射效率/%
LM	1	1.2
ODIFIIT	0.33	5.7
LMCA	2.03	5.9
LM-ODIFIIT	5.9×10^{-31}	1.2
LMCA-ODIFIIT	1	2.1

表 16.3 为计算机实验的结果，实验中，每一单元的幅度和相位为量化后的值，并应用了 ODIFIIT。每一罗曼单元被分成 $N \times N$ 个更小的 N 阶量化方块。幅值与相位的量化按照 16.6 节中讨论的方法进行， MSE 依然为相对于罗曼法中的 MSE 。

表 16.3 LM-ODIFIIT 法和 LMCA-ODIFIIT 法中， MSE 和衍射效率与量化阶数的函数关系

量化阶数 N	MSE	衍射效率/%
2	2.33	0.6
2-CA	0.42	1.4
4	0.19	1.1
4-CA	0.93	1.1



实验结果表明,就 MSE 和效率而言,ODIFIIT 法均比罗曼法好。在等幅罗曼法中,图像区域的误差增加,但效率显著提高。基本上,这意味着重构得到较亮的图像,但精度较低。这个结果是合理的,有以下原因:LMCA 法中使用较少的信息,这增加了误差,但由于孔径总是尽可能的大,所以整个全息图的幅度值更大,使得亮度增加。输出结果给出了目标图像的重构图,其亮度更高,但边缘模糊。目标图像区域之外的重构更少,这进一步表明效率的提高。事实上,LMCA 法的效率超过了 ODIFIIT 法的效率,但 MSE 的显著增加是其劣势。

结果表明,LM-ODIFIIT 法的结果与目标图像最相像(最小的 MSE),而效率与原始罗曼法(LM)一样。仿真实验表明,从质量上看,LM-ODIFIIT 法在图像区域内生成了一幅锐利且均一的图像。对 Lena 图的仿真重构结果更加表明 LM-ODIFIIT 法的精确性,这与 LM 法和 ODIFIIT 法的结果相当。

LMCA-ODIFIIT 法的 MSE 与 LM 法的相等,比 LM-ODIFIIT 法的大,比 LMCA 法的小。将 ODIFIIT 法合并于 LMCA 法中能减小误差,就像 ODIFIIT 法合并于 LM 法中一样。然而意外的是,考虑到 LM-ODIFIIT 法的极小 MSE ,LMCA-ODIFIIT 法的 MSE 并不是更小。就像预期的那样,LMCA-ODIFIIT 法的效率比 LM-ODIFIIT 法的效率高。

在所有方法中,量化阶 $N=2$ 的 LM-ODIFIIT 法有最大的 MSE 和最低的效率。使用如此少的量化阶数,编码后保留的允许良好重构的信息不足。当归一化的幅度小于 $1/2$ 时,幅度量化为 0,相位信息丢失。这解释了高误差与低效率的原因。利用恒定幅值技术,更多的相位信息得以保留,这能得到更好的 MSE 和效率。

对于 $N=4$, MSE 显著下降,能生成对目标图像很好的重构。该方法具有低的 MSE ,仅次于原始 LM-ODIFIIT 法的极小 MSE 值。同时,其效率与除 ODIFIIT 法和 LMCA 法以外的其他所有方法都有可比性。恒定的幅值保证了效率,但增加了误差。

如图 16.23 所示,128 像素 \times 128 像素 LM 法和 LM-ODIFIIT 法的图像具有周期性的点阵。LM 法图像中的水平阵列是由全息图的周期特性产生的。这是由于孔径高度被调整时,其像素是对称的。LM-ODIFIIT 法图像中的水平阵列也由同样的原因产生。然而,在 LM-ODIFIIT 法全息图中,可观察到,子全息图的每一点处的相位都收敛到相同的值。除第一个子全息图外,每个子全息图都能观察到这种情况。这使得全息图的其他方向具有周期特性,这种特性产生了垂直点阵。

到现在为止,所有的仿真忽略了 sinc 函数的“陡降”特性,而这在二元 CGH 的物理重构中是实际存在的。在仿真中进一步用到了本节开始时讨论的补偿方法来校正这些误差。计算机重构显示了 x 方向的良好的一致性,表明依赖于 x 的 sinc 因子已被精确补偿。在 y 方向有微小的变化,但可以通过连续迭代减小到一个所需的水平。图 16.24 显示了使用 sinc “陡降”校正前后的灰度图的计算机重构图像。



图 16.24 使用 sinc “陡降”校正前后的灰度图的计算机重构图像

a) LM-ODIFIIT 法全息图生成的灰度图
b) 对算法做近似补偿后的图像

第 17 章 计算机成像技术 I：合成孔径雷达

17.1 本章简介

许多的现代成像技术在进行图像重构时都依赖于傅里叶变换及相关的计算机算法。例如，合成孔径雷达（Synthetic Aperture Radar, SAR），利用投影的图像重构，包括计算机断层扫描、磁共振成像（Magnetic Resonance Imaging, MRI）以及在第 5 章讨论的共焦显微镜都属于这种技术。

在本章中，对合成孔径雷达（SAR）原理进行了讨论。1958 年密歇根大学根据 SAR 在本质上与全息类似的特性，用相干光处理方法制作了第一张 SAR 图像 [Cutrona]。在 20 世纪 70 年代初，数字信号处理方法开始用在计算机 SAR 图像处理领域 [Bennett] [Ulaby]。目前，计算机 SAR 图像处理已在实际中得到应用。

本章共分 10 节。17.2 节概括性地描述了合成孔径雷达。17.3 节介绍了合成孔径雷达的距离分辨率。为了实现高分辨率的成像，我们有必要认真地选择所使用的脉冲波形。该内容将在 17.4 节介绍。17.5 节讨论匹配滤波，它用于信噪比（SNR）最大化和实现尖脉冲。在 17.6 节将讨论采用匹配滤波生成具有足够能量的尖脉冲。17.7 节介绍了合成孔径雷达的横向分辨率。

上述章节主要讨论了 SAR 成像的一些基础问题。17.8 节讨论了 SAR 成像的简单理论。17.9 节重点讨论了利用菲涅耳近似进行图像重构。17.10 节给出了另一种数字 SAR 图像重构的算法。

17.2 合成孔径雷达

雷达（Radar）这个词是无线电（Radio）探测（Detection）和（And）测距（Ranging）中的部分字母组合而成的。合成孔径雷达是一种实现高分辨率成像的技术，这可以在地形地貌、专题制图、海洋、林业、农业、城市规划、环境科学和自然灾害预报评估等诸多领域得到很好的运用。

合成孔径雷达是一种相干成像技术，它是利用安装在运载平台上的运动雷达系统来实现成像功能的，如安装在飞机上（机载合成孔径雷达）或卫星上（天基合成孔径雷达）。当运载平台处于运动状态时，目标区域被雷达的辐射所照射，因而产生一个回波信号的集合。这些回波信号集合在一起，采用特殊算法，即可得到目标区域的高分辨率图像。

合成孔径雷达成像的系统示意图如图 17.1 所示。微波频率的电磁脉冲由发射器发送后被目标区域反射，并由接收器接收。为了合成一个大型的人工天线，通常使用直线方向的飞行路径。以这种方式得到的有效天线长度可以达到几千米。方位角（或横向距离）是沿飞行路径的方向。测距（或斜距）是在垂直于方位角的方向进行的。



利用回波或反射脉冲信息来获取目标区域的图像信息。通过多次周期性发射脉冲波照射目标区域，从而获得高分辨率（小像素尺寸）和高信噪比。这等同于每个信号都是从一个较大的天线阵列的单元天线发出的。因此，这种雷达命名为合成孔径雷达。

如 17.7 节介绍的，在距离 R 处，尺寸为 L 的天线的方位角分辨率约为 $\lambda R/L$ 。如果没有合成孔径，在微波频率下达到较大的 L 尺寸是不可能实现的。合成孔径相当于一个沿着航路分布众多单元天线的天线阵列。

SAR 成像往往利用卫星系统来完成。例如，1978 年 6 月，美国航空航天局喷气推进实验室将一颗海洋资源卫星发射入轨道，该系统用于完成地球表面大区域的 SAR 图像的任务。基于航天飞机的航天成像雷达系统（SIR）执行与 SAR 系统类似的任务。也有由飞机作为搭载平台的机载 SAR 系统，例如 E-3 AWACS（E-3 飞机机载预警控制系统），其被用在波斯湾地区海上和空中目标的探测和跟踪，并在 E-8C 联合监视目标攻击雷达系统中得以应用，该系统在海湾战争中用来探测和定位地面目标。

图 17.2 给出了一个 SAR 图像的例子 [远程成像，遥感与加工中心，新加坡国立大学（CRISP）]。该张南格陵兰岛图像由 Envisat 卫星的中分辨率成像光谱仪（MERIS）[ESA，欧洲航天局] 于 2006 年 2 月 16 日获取。

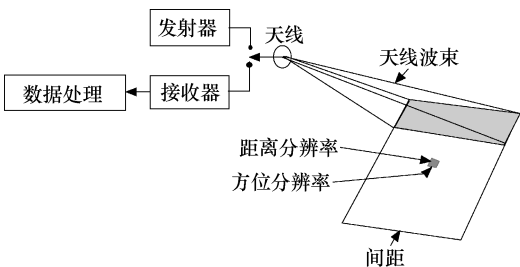


图 17.1 一个合成孔径雷达成像的系统示意图



图 17.2 生成于 2006 年 2 月份的 SAR 图像

17.3 测距分辨率

SAR 及其他类型的雷达的测距分辨率都是利用脉冲电磁波获取的。测距分辨率需要处理所接收信号的不确定性，这些不确定性是由于从接收到的紧邻的相邻物体所发出的脉冲的重叠造成的。

除了相邻物体，还存在噪声问题，例如，由于电磁干扰信号造成的随机涨落、大气效应以及电子元器件中的热变化等。因此，有必要增加信噪比（SNR）以达到更高的测距分辨率。

对于距离为 R 的单个物体，对应脉冲反射距离为 $tc/2$ ，其中 t 是发送与接收脉冲之间的时间间隔， c 是光速， $c = 3 \times 10^8 \text{ m/s}$ 。假设脉冲持续时间为 $T(\text{s})$ 。那么两个物体之间的脉冲延迟时间至少需要 T 秒，以使得两个脉冲回波之间没有重叠。这意味着物体必须至少分隔 $cT/2$ 米（如果使用米-千克-秒单位制）。减小 T 可以获取更高的测距分辨率。但是，为了得到更好的检测效果，也需要高脉冲能量，而实际中短脉冲意味着更低的能量。为了避免这个问题，17.5 节将讨论匹配滤波，通常在接收端利用匹配滤波将长持续时间脉冲转换为短持



续时间脉冲。在这种方式下，接收的回波被锐化，并且整体系统具有短脉冲的测距分辨率。同时，在发射平均功率恒定的条件下，峰值发射功率也大大降低。在这样的系统中，匹配滤波既被用来进行脉冲压缩，同时也被用来优化信噪比以提高探测精度。这将在 17.5 节中进一步讨论。

17.4 脉冲波形的选择

脉冲的形状对于区分邻近的物体很重要。假设 $p(t)$ 是持续时间为 T 的脉冲信号，该信号在 $T > 0$ 时非零。对一个物体而言返回的脉冲可以写为

$$p_1(t) = \sigma_1 p(t - \tau_1) \quad (17.4-1)$$

式中， σ_1 为衰减常数； τ_1 是延迟时间。

从另一物体返回的脉冲可以写成

$$p_2(t) = \sigma_2 p(t - \tau_2) \quad (17.4-2)$$

脉冲的形状应进行优化，使得 $\tau_1 \neq \tau_2$ 的情况下 $p_1(t)$ 和 $p_2(t)$ 尽可能不相似。

最常用于量度两个波形 $p_1(t)$ 和 $p_2(t)$ 之间相似性的，是下式给出的欧几里得距离：

$$D^2 = \int [p_1(t) - p_2(t)]^2 dt \quad (17.4-3)$$

D^2 也可以写为

$$D^2 = \sigma_1^2 \int p^2(t - \tau_1) dt + \sigma_2^2 \int p^2(t - \tau_2) dt - 2\sigma_1\sigma_2 \int p(t - \tau_1)p(t - \tau_2) dt \quad (17.4-4)$$

式 (17.4-4) 等号右侧前两项正比于脉冲能量，其可以通过缩放单独控制。因此，只有最后一项对于优化最为重要。在 $\tau_1 \neq \tau_2$ 条件下，应使最后一项最小化以达到 D^2 最大化。最后一项的积分可重写为

$$R(\tau_1, \tau_2) = \int p(t - \tau_1)p(t - \tau_2) dt \quad (17.4-5)$$

其等同于下式

$$R(\tau) = \int p(t)p(t + \tau) dt \quad (17.4-6)$$

式中， $\tau = \tau_1 - \tau_2$ 或者 $\tau = \tau_2 - \tau_1$ ，可以看出 $R(\tau)$ 是 $p(t)$ 的自相关函数。

线性频率调制（线性调频）信号，也称为啁啾信号，具有非常尖锐的自相关特性，其自相关函数在 $\tau \neq 0$ 时接近零，它可以写成

$$x(t) = A \cos(2\pi(ft + \gamma t^2)) \quad (17.4-7)$$

更一般地可写为

$$x(t) = e^{j2\pi(ft + \gamma t^2)} \quad (17.4-8)$$

较大的 γ 表示较大的瞬时频率 f_i 的变化，也就是相位的导数

$$f_i = f + 2\gamma t \quad (17.4-9)$$

可见， f_i 与 t 呈线性关系。 $x(t)$ 的自相关函数可以表示为

$$R(\tau) = e^{j2\pi\tau(f + \gamma\tau)} \int e^{j4\pi\gamma\tau t} dt \quad (17.4-10)$$

假设脉冲波形中心在 T_0 处，持续时间 T 可以表示为



$$p(t) = \text{rect}\left(\frac{t - T_0}{T}\right)x(t) \quad (17.4-11)$$

$P(t)$ 的自相关函数可写为

$$R(\tau) = e^{j2\pi\left[\frac{\gamma}{2}\left(t_0 + \frac{\tau}{2}\right)\right]} \text{tri}\left((T - |\tau|) \text{sinc}\left(\frac{\pi\gamma T}{2}(T - |\tau|)\right)\right) \quad (17.4-12)$$

此函数的主瓣宽度约等于 $1/\gamma T$ 。通常用更平滑的函数如高斯函数代替矩形窗函数，这在 14.2 节中有所讨论。基于上面所讨论的特性，为了得到较高的测距分辨率，一般采用有限宽度窗口内的啁啾信号作为测距脉冲波形。

17.5 匹配滤波器

匹配滤波的一个基本问题是要判断噪声中是否存在给定形式的信号。在经典情况下，该滤波器也被限定为线性时不变（LTI）系统。匹配滤波器也有很多其他的应用，诸如下节将要讨论的脉冲压缩，以及 17.8 ~ 17.10 节中要讨论的图像重构。

假设输入信号由给定信号 $x(t)$ 加噪声 $N(t)$ 组成。令线性系统相应的输出分别取 $x_0(t)$ 和 $N_0(t)$ ，如图 17.3 所示。用来确定是否存在 $x(t)$ 的最优准则，是系统输出有最大的信噪比。下面将要给出，使得信噪比最大化的 LTI 滤波器，即匹配滤波器。

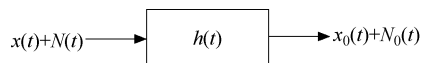


图 17.3 作为 LTI 滤波器的匹配滤波器可最有效地降低噪声

如果 $N_0(t)$ 被假定为一个广义平稳（WSS）过程的采样函数， T_0 时刻的信噪比可定义为

$$\text{SNR} = \frac{|x_0(T_0)|^2}{E[N_0^2(T_0)]} \quad (17.5-1)$$

输出信号 $x_0(t)$ 可由下式求得：

$$x_0(T) = \int_{-\infty}^{+\infty} X(f)H(f)e^{j2\pi fT_0}df \quad (17.5-2)$$

平均输出噪声功率由下式求得：

$$E[N_0^2(t)] = \int_{-\infty}^{+\infty} |H(f)|^2 S_N(f) df \quad (17.5-3)$$

式中， $S_N(f)$ 是 $N(t)$ 的频谱密度函数。

此时信噪比可以写为

$$\text{SNR} = \frac{\left| \int_{-\infty}^{+\infty} X(f)H(f)e^{j2\pi fT_0}df \right|^2}{\int_{-\infty}^{+\infty} |H(f)|^2 S_N(f) df} \quad (17.5-4)$$

为了优化信噪比，可以运用施瓦兹不等式。如果 $A(f)$ 和 $B(f)$ 为 f 的两个可能的复函数，由施瓦兹不等式可得

$$\left| \int_{-\infty}^{+\infty} A(f)B(f)df \right|^2 \leq \int_{-\infty}^{+\infty} |A(f)|^2 df \int_{-\infty}^{+\infty} |B(f)|^2 df \quad (17.5-5)$$

当满足下式时取等号

$$A(f) = CB^*(f) \quad (17.5-6)$$



C 为任意实常数。令

$$\begin{aligned} A(f) &= \sqrt{S_N(f)} H(f) \\ B(f) &= \frac{X(f) e^{j2\pi/T_0}}{\sqrt{S_N(f)}} \end{aligned} \quad (17.5-7)$$

于是由施瓦兹不等式可得

$$\left| \int_{-\infty}^{+\infty} X(f) H(f) e^{j2\pi/T_0} df \right|^2 \leq \left[\int_{-\infty}^{+\infty} S_N(f) |H(f)|^2 df \right] \left[\int_{-\infty}^{+\infty} \frac{|X(f)|^2}{S_N(f)} df \right] \quad (17.5-8)$$

或者

$$SNR \leq \int_{-\infty}^{+\infty} \frac{|X(f)|^2}{S_N(f)} df \quad (17.5-9)$$

根据式 (17.5-6)，信噪比可被最大化并且使其等于方程式 (17.5-9) 右侧的值，换言之，此时有

$$H(f) = H_{opt}(f) = C \frac{X^*(f)}{S_N(f)} e^{-j2\pi/T_0} \quad (17.5-10)$$

传递函数由公式 (17.5-10) 给出的滤波器称为匹配滤波器。可见， $H_{opt}(f)$ 正比于输入信号的傅里叶变换的复共轭，反比于输入噪声的频谱密度函数，其系数 $e^{-j2\pi/T_0}$ 用于调节该最大信噪比产生的时间 T_0 。

例 17.1 若输入噪声为白色噪声，频谱密度为 N_0 ，试求匹配滤波器的传递函数和脉冲响应，并给出在时域中采用匹配滤波器的卷积运算。

解：在式 (17.5-10) 中，用 N 代替 $S_N(f)$

$$H_{opt}(f) = K X^*(f) e^{-j2\pi/T_0}$$

对输入信号以 $H_{opt}(f)$ 进行卷积得到输出信号。于是

$$y(t) = \int_{-\infty}^{+\infty} x(\tau) h_{opt}(t - \tau) d\tau = K \int_{-\infty}^{+\infty} x(\tau) x(T_0 - t + \tau) d\tau$$

在 $t = T_0$ 处存在峰值，该值由下式给出：

$$y(T_0) = K \int_{-\infty}^{+\infty} x^2(\tau) d\tau$$

其与输入信号能量成正比。

17.6 运用匹配滤波进行脉冲压缩

在脉冲雷达和声呐等应用中，要获取良好的距离分辨率，短的脉冲持续时间尤为重要。然而，良好的探测效果却需要脉冲具有较高的能量，而短脉冲持续时间实际上意味着较低的脉冲能量。为了避免这个问题，匹配滤波通常用来在接收端将宽持续时间的脉冲转换为短持续时间的脉冲。通过这种方式，接收到的回波变得更尖锐，整个系统具有短脉冲的距离分辨率。当平均功率恒定时，峰值发射功率也大大减小。在这样的系统中，匹配滤波可用来进行脉冲压缩以及优化探测的信噪比。

通过选择输入脉冲的波形，可得到较短持续时间的输出脉冲。例如，可以选择如下形式



的线性调频输入脉冲:

$$x(t) = e^{t^2/T^2} e^{j(2\pi f_0 t + \gamma t^2)} \quad (17.6-1)$$

式中, T 为脉冲持续时间。

在实际中, $|\gamma|T$ 比 $2\pi f_0$ 要小得多。脉冲的频谱由下式给出:

$$X(f) = F_B \sqrt{m/\pi} e^{-4\pi^2(f-f_0)^2/F_B^2} e^{j[\gamma T^2(f-f_0)^2/F_B^2 - \frac{1}{2}\arctan\gamma T^2]} \quad (17.6-2)$$

式中

$$m = [1 + \gamma^2 T^4]^{\frac{1}{2}} \quad (17.6-3)$$

$$F_B = \frac{m}{\pi T} \quad (17.6-4)$$

F_B 是频谱的有效带宽, 因为频谱也是中心频率为 f_0 的高斯分布。

不计常数项, 信号 $x(t)$ 的匹配滤波器 (假设噪声是白噪声, $T_0 = 1$) 由下式给出:

$$H(f) = e^{-4\pi^2(f-f_0)^2/F_B^2} e^{j[\gamma T^2(f-f_0)^2/F_B^2 - \frac{1}{2}\arctan\gamma T^2]} \quad (17.6-5)$$

在实际中, 式 (17.6-5) 中 $H(f)$ 的作用减小了最终结果的幅值。这可以通过使用纯相位滤波器来防止

$$H(f) = e^{j[\gamma T^2(f-f_0)^2/F_B^2 - \frac{1}{2}\arctan\gamma T^2]} \quad (17.6-6)$$

于是, 匹配滤波器的输出为

$$\begin{aligned} y(t) &= \int_{-\infty}^{+\infty} X(f) H(f) e^{j2\pi f t} df \\ &= \sqrt{m} e^{-t^2/(T/m)^2} e^{j2\pi f_0 t} \end{aligned} \quad (17.6-7)$$

再次可见, 输出信号是一个除去了频率调制的高斯脉冲, 脉冲持续时间压缩因子为 m 。同时, 为保持脉冲的能量不变, 其振幅乘以因子 \sqrt{m} 。

如用式 (17.6-5) 来代替式 (17.6-6), 可以得到相同的结果, 此时需要用 $m/\sqrt{2}$ 替换 m 。

在脉冲雷达和声呐中, 测距精度和分辨率与脉冲持续时间密切相关。近距离物体反射的长持续时间信号与目标信号融合在一起, 降低了分辨率。最大测量范围也与信噪比和脉冲能量关系密切。因此, 利用上面所述的技术, 无论是高精度度、高分辨率, 还是长距离测量, 都可以得到实现。

上面所讨论的脉冲压缩是普遍适用的, 而并非依赖于某种特定信号。下面的讨论是针对 $T_0 = 0$ 的一般脉冲信号的。由式 (17.6-5), 此时该匹配滤波器产生的频谱具有零相位, 由方程 (17.6-5) 可求得其振幅 $|X(f)|^2$ 。 $t=0$ 时的输出信号由下式给出:

$$y(0) = \int_{-\infty}^{+\infty} |X(f)|^2 df = E$$

该信号具有较大的值。

为了定义压缩的程度, 有必要对时间和频率宽度的量度进行定义。令脉冲具有能量 E 和最大振幅 A_{\max} 。输入脉冲的持续时间可以定义为

$$T_x = \frac{E}{A_{\max}^2} \quad (17.6-8)$$

同样, 谱宽度 F 可定义为



$$F = \frac{E}{B_{\max}^2} \tag{17.6-9}$$

式中， B_{\max} 是频谱的最大振幅。

输入信号能量和输出信号 $y(0)$ 相同。则输出信号能量为

$$E_y = \int_{-\infty}^{+\infty} |X(f)|^4 df \tag{17.6-10}$$

压缩比 m 由下式给出：

$$m = \frac{T_x}{T_y} \tag{17.6-11}$$

式中， T_y 是输出的脉冲宽度。

其可由下式得出：

$$C_{\max}^2 T_y = E_y$$

C_{\max} 是输出信号 $y(0)$ 的最大振幅。因为 $y(0)$ 和 E 相同，方程式 (17.6-11) 可以写成

$$m = \frac{T_x E^2}{E_y} = \alpha T_x F \tag{17.6-12}$$

其中

$$\alpha = \frac{E}{E_y} B_{\max}^2 = B_{\max}^2 \frac{\int_{-\infty}^{+\infty} |X(f)|^2 df}{\int_{-\infty}^{+\infty} |X(f)|^4 df} \tag{17.6-13}$$

α 同时也可以写成

$$\alpha = \frac{\int_{-\infty}^{+\infty} D^2(f) df}{\int_{-\infty}^{+\infty} D^4(f) df} \tag{17.6-14}$$

其中

$$D(f) = \frac{|X(f)|}{B_{\max}} \tag{17.6-15}$$

因为 $D(f)$ 小于或等于 1，故 α 大于或等于 1。例如，三种类型的谱及对应的 α 见表 17.1。

表 17.1 三种类型的谱及对应的 α 值

谱	矩 形	三 角 形	高 斯 型
α	1	1.67	2.22

可以看出，高斯谱具有最佳的脉冲压缩特性。方程式 (17.6-12) 表明压缩比正比于信号的持续时间 T_x 和谱宽度 F 的乘积。这通常被称为时间-带宽积。

17.7 方位角分辨率

为了理解图 17.4 所示长度为 L 的天线的方位角（横向）分辨率特性，假定目标处于很远的位置，回波信号以某一角度 ϕ 入射到天线的任意位置。

假设天线可对入射能量连续积分，天线的积分响应可表示为



$$E(\phi) = A \int_{-L/2}^{L/2} e^{j\psi(y)} dy \quad (17.7-1)$$

式中, A 是入射信号振幅, 假定其为常数, 并有

$$\psi(y) = \frac{2\pi}{\lambda} y \sin \phi \quad (17.7-2)$$

式 (17.7-2) 为由 $\delta = y \sin \phi$ 引起的相移, $E(\phi)$ 的虚数部分积分为零, 并且实部由下式给出:

$$E(\phi) = \frac{\sin\left(\frac{\pi}{\lambda} L \sin \phi\right)}{\left(\frac{\pi}{\lambda} L \sin \phi\right)} L \quad (17.7-3)$$

在 $\phi = 0$ 处进行归一化处理, 天线功率增益变成

$$G(\phi) = \left| \frac{E(\phi)}{E(0)} \right|^2 = \frac{\sin^2\left(\frac{\pi}{\lambda} L \sin \phi\right)}{\left(\frac{\pi}{\lambda} L \sin \phi\right)^2} \quad (17.7-4)$$

令半功率点的增益 $G(\phi) = 1/2$, 经过简单的计算, 有

$$\phi|_{3\text{dB}} \approx \pm 0.44 \frac{\lambda}{L} \quad (17.7-5)$$

对于距离为 R 的目标, 这相当于下列方位角分辨率:

$$\Delta \approx 0.88 \frac{R\lambda}{L} \quad (17.7-6)$$

因此, 采用短波长和长的天线长度可以提高方位角分辨率。

17.8 简化的 SAR 成像理论

成像几何关系如图 17.5 所示。为简单起见, 用二维几何关系来进行分析。对于现实中的三维几何关系, 用 $\sqrt{x^2 + z^2}$ 替换 x 即可得到, 其中 z 为高度。

在讨论 SAR 成像时, 通常讨论以下两种模式中的一种。在斜视模式下, 目标区域的中心在 (x_c, y_c) 处, 其中 y_c 非零, 如图 17.5 所示。在聚束模式下, y_c 等于零, 此时目标区域在垂直于飞行方向上出现。

假设目标区域由反射率为 σ_n , 坐标为 (x_n, y_n) ($n = 1, 2, 3, \dots$) 的稳定反射点组成, 脉冲信号 $p(\cdot)$ 用来照亮目标区域。雷达接收机在 $(0, y)$ 处接收经目标反射回来的回波信号 $s(t, y)$

$$s(t, y) = \sum_n \sigma_n p(t - t_n) \quad (17.8-1)$$

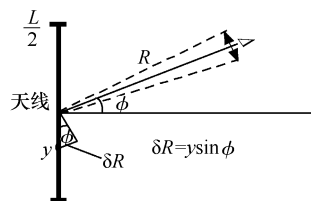


图 17.4 几何估算方位角分辨率

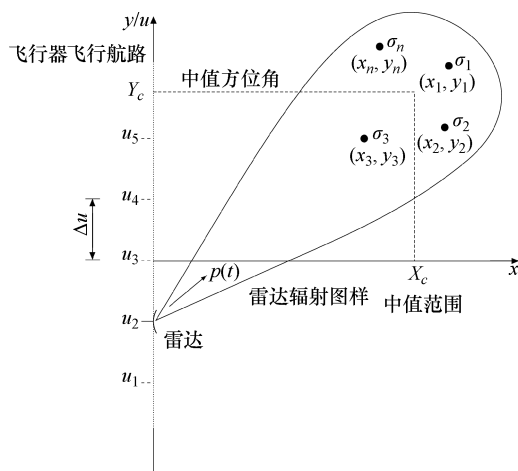


图 17.5 SAR 成像几何 [由 Soumekh 提供, 1999]



式中, t_n 是从雷达到第 n 个目标的来回双程延迟时间, 由下式给出:

$$t_n = \frac{2r_n}{c} \quad (17.8-2)$$

r_n 为雷达到第 n 个目标的距离

$$r_n = [x_n^2 + (y_n - y)^2]^{\frac{1}{2}} \quad (17.8-3)$$

要重构的图像可以表示为

$$u(x, y) = \sum_n \sigma_n \delta(x - x_n, y - y_n) \quad (17.8-4)$$

所以, 问题转变为如何由测量信号 $s(t, y)$ 获取 $u(x, y)$ 的一个近似。 $u(x, y)$ 的二维傅里叶变换由下式给出:

$$U(f_x, f_y) = \sum_n \sigma_n \exp[-j2\pi(f_x x_n + f_y y_n)] \quad (17.8-5)$$

$s(t, y)$ 对时间变量 t 的一维傅里叶变换由下式给出:

$$S(f, y) = P(f) \sum_n \sigma_n \exp[-j2\pi f t_n] = P(f) \sum_n \sigma_n \exp[-j2\pi k r_n] \quad (17.8-6)$$

式中, $k = 2\pi f/c$ 表示波数; $P(f)$ 为 $p(t)$ 的傅里叶变换。

对于随两个变量 t, y 变化的函数 $s(t, y)$ 的二维傅里叶变换, 现在可利用 $S(f, y)$ 对变量 y 的一维傅里叶变换来计算

$$S(f, f_y) = P(f) \sum_n \sigma_n \exp[-j(k_x x_n + k_y y_n)] \quad (17.8-7)$$

式中

$$\begin{aligned} k_y &= 2\pi f_y \\ k_x &= [4k^2 - k_y^2]^{\frac{1}{2}} = 2\pi f_x \end{aligned} \quad (17.8-8)$$

注意, $U(f_x, f_y)$ 可以从 $S(f, f_y)$ 中得出

$$U(f_x, f_y) = \frac{S(f, f_y)}{P(f)} \quad (17.8-9)$$

这个运算被称为源反卷积, 其源是 $P(f)$ 。但是, 使用方程式 (17.8-9) 通常会产生错误结果, 因为它是一个病态运算, 特别是有噪声存在的情况下。替换地采用匹配滤波后的形式如下:

$$\begin{aligned} U'(f_x, f_y) &= S(f, f_y) P^*(f) \\ &= |P(f)|^2 \sum_n \sigma_n \exp[-j2\pi(f_x x_n + f_y y_n)] \end{aligned} \quad (17.8-10)$$

可见, 在时空域中, 这个操作相当于 $s(t, y)$ 与 $p^*(-t)$ 的卷积运算

$$\begin{aligned} u'(t, y) &= s(t, y) * p^*(-t) \\ &= \sum_n \sigma_n h\left(t - \frac{2r_n}{c}\right) \end{aligned} \quad (17.8-11)$$

其中 $h(\cdot)$ 是由下式给出的点扩散函数:

$$h(t) = F^{-1}[|P(f)|^2] \quad (17.8-12)$$

与式 (17.8-4) 相比, 式 (17.8-10) 表示重构信号的频谱与期望信号的频谱相差一个 $|P(f)|^2$ 因子。这是因为典型的零相位缓变振幅函数 f , 在噪声存在情况下通过匹配滤波重构时, 它能获取比源反卷积更好的结果。



17.9 用菲涅耳近似进行图像重构

可以采用菲涅耳近似简化重构算法。假设目标区域的中心在 $(x_c, 0)$ ；将 r_n 在 $r_n = x_c + x_n$ 下围绕点 $(x_c, 0)$ 进行泰勒级数展开

$$r_n = x_c + x_n + \frac{(y_n - y)^2}{2x_c} + \dots \quad (17.9-1)$$

菲涅耳近似对应于将 r_n 近似为上述等号右边给出的项。方程式 (17.8-6) 中的 $S(f, y)$ 现在可以写成

$$S(f, y) \simeq P(f) \exp(-2jkx_c) \sum_n \sigma_n \exp\left[-j2k\left(x_n + \frac{(y_n - y)^2}{2x_c}\right)\right] \quad (17.9-2)$$

考虑由公式 (17.8-4) 得出的理想目标函数对坐标 x 的一维傅里叶变换

$$U(k_x, y) = \sum_n \sigma_n \exp(-jk_x x_n) \delta(y - y_n) \quad (17.9-3)$$

式中, $k_x = 2\pi f_x$; $\delta(x - x_n, y - y_n) = \delta(x - x_n) \delta(y - y_n)$ 。

令 $k_x = 2k = \frac{2\omega}{c}$, 对比方程式 (17.9-2) 和式 (17.9-3), 可见 $S(f, y)$ 约等于 $P(f)U(2k, y)$ 与 $\exp(-jky^2/x_c)$ 对坐标 y 进行的卷积

$$S(f, y) \simeq P(f)U(2k, y) * \exp(-jky^2/x_c) \quad (17.9-4)$$

为了恢复 $u(x, y)$, 对信号 $P(f)$ 和啁啾信号 $\exp(-jky^2/x_c)$ 进行匹配滤波。重构信号的频谱可表示为

$$U'(f_x, y) = P^*(f)S(f, y) * \exp(jky^2/x_c) \quad (17.9-5)$$

在一般情况下, 雷达带宽比它的载波频率 k_c 小得多, 因此 $|k - k_c| \ll k_c$ 通常成立, 这被称为窄带, 另外, y 值通常比目标距离小得多, 这称为波束锐化。于是, $\exp(jky^2/x_c)$ 可以通过 $\exp(jk_c y^2/x_c)$ 近似, 式 (17.9-5) 变成

$$U'(f_x, y) = P^*(f)S(f, y) * \exp(jk_c y^2/x_c) \quad (17.9-6)$$

其中

$$f_x = \frac{k}{\pi} = \frac{\omega}{\pi c} \quad (17.9-7)$$

注意, 上述过程有两种滤波运算。第一个是用滤波传递函数 $P^*(f)$, 其对应于时域中的 $p^*(-t)$ 。第二个用滤波冲激函数 $\exp(jk_c y^2/x_c)$ 。

时间变量 t 和距离变量 x 之间的近似对应关系由下式给出:

$$x = \frac{ct}{2} \quad (17.9-8)$$

总之, 当窄带和波束锐化的假设有效时, 图像重建可按如下步骤进行:

- (1) 用脉冲响应 $p^*(-t)$ 在时间方向进行滤波。
- (2) 用脉冲响应 $\exp(jk_c y^2/x_c)$ 在 y 方向进行滤波。
- (3) 利用 $x = ct/2$ 得到 x , 或者 $t = 2x/c$ 得到 t 。

该算法的框图如图 17.6 所示。当上面讨论的近似不成立时, 可以使用更复杂的图像重构算法, 这将在下节讨论。

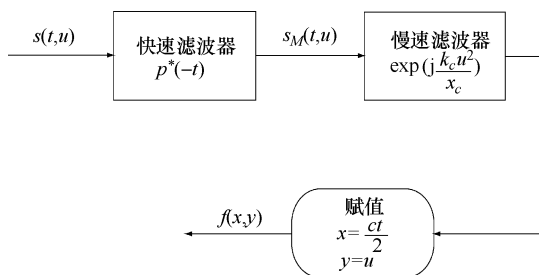


图 17.6 运用菲涅耳近似进行 SAR 重构的框图 [由 Soumekh 提供, 1999]

17.10 数字图像重构算法

SAR 成像的目标就是从所测量的信号 $s(t, y)$ 中获取目标区域的反射率函数 $u(x, y)$ 。目前已经开发了几种用于此目的的数字重构算法, 如空间频率插值算法、距离层叠算法、时域相关算法和反投影算法 [Soumekh 提供]。下面将简要讨论空间频率插值算法。

空间频率插值

$s(t, y)$ 的二维傅里叶变换为 $S(f, f_y)$, 其可以写成 $S(k, k_y)$, 其中 $k = 2\pi f/c$, $k_y = 2\pi f_y$ 。

在匹配滤波后, 有必要使 k 映射为 k_x , 这样经过二维逆傅里叶变换后就可以获取重构图像。 k_x 、 k 和 k_y 的关系由下式得出:

$$k_x^2 = 4k^2 - k_y^2 \quad (17.10-1)$$

当要使用离散傅里叶变换 (DFT) 时, 需在相等的时间间隔内对 k 和 k_y 采样, k_x 采样是非等间隔的, 如图 17.7 所示。因此有必要使用插值法, 以便对矩形网格内的 k_x 和 k_y 进行采样。只有当 $k_x \approx 2k$ 时, 才没有必要使用插值法, 就像 17.9 节对窄带和波束锐化的近似情况一样。

插值法应用于低通信号。需要注意的是, 雷达测距范围是 $x \in [x_c - x_0, x_c + x_0]$, $y \in [y_c - y_0, y_c + y_0]$, 其中 (x_c, y_c) 是目标区域的中心。因为 (x_c, y_c) 不是 $(0, 0)$, 所以 $S(k_x, k_y)$ 是一个快速变化的带通信号。为了将它转换为一个低通信号, 进行如下运算:

$$S'(k_x, k_y) = S(k_x, k_y) e^{j(k_x x_c + k_y y_c)} \quad (17.10-2)$$

下一步对 $S'(k_x, k_y)$ 插值, 以保证 k_x 的采样值位于规则间隔中。对内插后的 $S'(k_x, k_y)$ 进行逆傅里叶变换, 用来得到重构的图像。该算法的框图如图 17.8 所示。

在 k_x 方向上插值时, 采样间隔是不均匀的。 k_y 、 k 的采样为

$$\begin{aligned} k_y &= m\Delta k_y = k_{ym} \\ k &= n\Delta k = k_n \end{aligned} \quad (17.10-3)$$

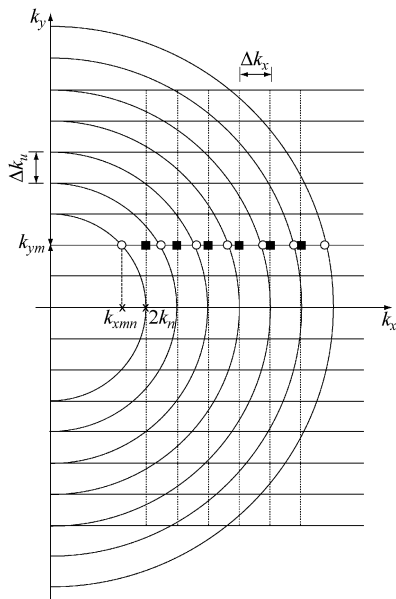


图 17.7 SAR 空间频率采样离散数据 [由 Soumekh 提供, 1999]

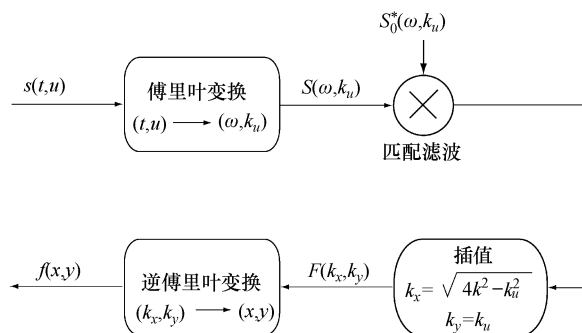


图 17.8 用空间频域插值来数字重建 SAR 图像的算法框图 [由 Soumekh 提供, 1999]

然后 k_x 的采样值由下式给出:

$$k_{xmn} = [4n^2 (\Delta k)^2 - m^2 (\Delta k_y)^2]^{\frac{1}{2}} \quad (17.10-4)$$

函数 $S'(k_x, k_y)$ 需要利用规则采样的 k_x 值进行插值, 由下式来实现 [Soumekh, 1988]:

$$S'(k_x, k_{ym}) = \sum_n J_{mn} S'(k_{xmn}, k_{ym}) h(k_x - k_{xmn}) |k_x - k_{xmn}| \leq N \Delta k_x \quad (17.10-5)$$

式中, N 是沿 k_x 正反两方向上采样点的数量; Δk_x 是在 k_x 方向上所需的采样间隔。

插值函数 $h(\cdot)$ 将在下面讨论, 雅可比行列式 J_{mn} 由下式给出:

$$J_{mn} = \frac{d}{dw} [4k^2 - k_{ym}^2]^{\frac{1}{2}} = \frac{4k_n}{c[4k_n^2 - k_{ym}^2]} \quad (17.10-6)$$

式中, $k = \omega/c$ 。

k_x 为

$$k_x = j \Delta k_x = k_{xi} \quad (17.10-7)$$

插值函数 $h(\cdot)$ 的选择是由数字信号处理和通信中的采样定理所决定的。如果信号 $V(k_x, k_y)$ 采用 $V(n \Delta k_x, m \Delta k_y)$ 采样, 则可进行如下插值:

$$V(k_x, k_y) = \sum_n V(n \Delta k_x, n \Delta k_y) \text{sinc}\left(\frac{k_x}{\Delta k_x} - n\right) \quad (17.10-8)$$

其中

$$\text{sinc}(v) = \frac{\sin \pi v}{\pi v} \quad (17.10-9)$$

sinc 函数有着无限长的尾巴。现实中为了避免这个问题, 通常利用窗函数 $\omega(k_x)$ 进行截断来替换 sinc 函数

$$h(k_x) = \text{sinc}\left(\frac{k_x}{\Delta k_x}\right) \omega(k_x) \quad (17.10-10)$$

$\omega(k_x)$ 可以选择 14.2 节中讨论的一些形式。例如, 可选择由下式给出的海明窗:

$$\omega(k_x) = \begin{cases} 0.54 + 0.46 \cos\left(\frac{\pi k_x}{N \Delta k_x}\right) & |k_x| \leq N \Delta k_x \\ 0 & \text{其他} \end{cases} \quad (17.10-11)$$

第 18 章 计算机成像技术 II：图像投影重构

18.1 本章引论

在涉及傅里叶相关变换的计算机成像技术第二部分中，将讨论包括断层扫描的图像投影重构技术。拉冬（Radon）变换是用于此目的的一种基本变换。在本章中，首先对 Radon 变换及其逆变换进行了详细的描述，然后介绍用于断层扫描和其他相关领域中所使用的成像算法。

计算机断层扫描（Computed Tomography, CT）是一种医学成像技术，该技术主要用来研究被检体从外部不可见的区域。物体的三维图像可从大量序列的二维 X 射线测量数据获得。图 18.1 给出了一个 CT 图像的示例。CT 也可用于其他领域，如非破坏性材料测试。

本章分为 9 个小节。18.2 节介绍 Radon 变换。18.3 节讨论投影切片定理，该节给出了投影的一维傅里叶变换如何相当于二维傅里叶变换平面的切片图像。逆 Radon 变换（Inverse Radon Transform, IRT）在 18.4 节中进行讨论。18.5 节描述了 Radon 变换的性质。

其余的小节都是关于重构算法的。18.6 节介绍了从投影重构二维信号所涉及的一些采样方面的问题。18.7 节涵盖了傅里叶重构算法，18.8 节介绍了滤波-反投影（back-projection）算法。



图 18.1 显示小脑、颞叶和鼻窦的头部 CT 图像 [由维基百科提供]

18.2 Radon 变换

如图 18.2 所示。 x - y 坐标系经 θ 角旋转得到 u - v 坐标系， (x,y) 和 (u,v) 的关系可通过平面旋转角 θ 来给出

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (18.2-1)$$

如图 18.2 所示，信号 $g(x,y)$ 的 Radon 变换 $p(u,\theta)$ ，是 $g(x,y)$ 平行于 v 轴、距离 u 轴为 u 的线积分， u 轴与 x 轴的夹角为 $0 \leq \theta < \pi$

$$\begin{aligned} p(u,\theta) &= \int_{-\infty}^{+\infty} g(x,y) dv \\ &= \int_{-\infty}^{+\infty} g(u\cos\theta - v\sin\theta, u\sin\theta + v\cos\theta) dv \end{aligned} \quad (18.2-2)$$

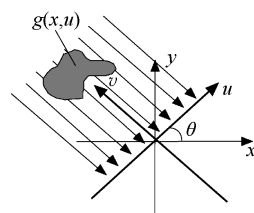


图 18.2 沿 v 轴线积分的直角坐标系旋转



$p(u, \theta)$ 也称为射线求和在角度 θ 处的 $g(x, y)$ 的射线积分。一组射线积分形成了一个投影。以下角 θ 作隐式处理, 将 $p(u, \theta)$ 写为 $p(u)$ 。

如图 18.2 所示, 将沿一组平行射线获取的投影称为平行投影。如果该组射线从一个点光源发出, 则将产生的投影称为扇形束投影。本章只考虑平行投影的情况。

Radon 变换也可以由狄拉克 δ 函数来表示。 u 可以写成 $\alpha^t r$, 其中 $\alpha = [\cos\theta \quad \sin\theta]^t$, $r = [x \quad y]^t$ 。于是, 公式 (18.2-2) 和下式相同

$$\begin{aligned} p(u, \theta) &= \int g(r) \delta(u - \alpha^t r) dr \\ &= \iint_{-\infty}^{+\infty} g(x, y) \delta(u - \cos\theta x - \sin\theta y) dx dy \end{aligned} \quad (18.2-3)$$

Radon 变换在计算机断层扫描、地球物理和多维信号处理等众多领域均有应用。最普遍的问题就是从有限数量的 θ 下的投影重构 $g(x, y)$ 。如果已知所有 θ 下的投影值, 则重构是精确的, 这将由下节给出的定理来证明。

18.3 投影切片定理

$p(u)$ 的一维傅里叶变换记为 $P(f)$, 它等于 $g(x, y)$ 的二维傅里叶变换 (记为 $G(f_x, f_y)$) 在角 θ 处的中心切片

$$P(f) = G(f \cos\theta, f \sin\theta) \quad (18.3-1)$$

在二维频率平面上的频率分量切片, 如图 18.3 所示。

证明:

$$\begin{aligned} P(f) &= \int p(u) e^{-j2\pi fu} du \\ &= \iint_{-\infty}^{+\infty} g(u \cos\theta - v \sin\theta, u \cos\theta + v \sin\theta) e^{-j2\pi fu} dv du \end{aligned}$$

在未旋转的坐标系中, 上式等价于

$$\begin{aligned} P(f) &= \iint_{-\infty}^{+\infty} g(x, y) e^{-j2\pi(fx \cos\theta + fy \sin\theta)} dx dy \\ &= G(f \cos\theta, f \sin\theta) \end{aligned}$$

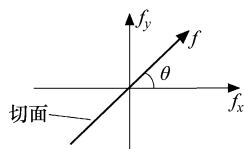


图 18.3 在投影切片定理中使用的频率切片

例 18.1 求 $g(x, y) = e^{-x^2 - y^2}$ 的 Radon 变换。

解:

$$p(u, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2 - y^2} \delta(u - x \cos\theta - y \sin\theta) dx dy$$

因为由式 (18.2-1) 给出的变换 $(x, y) \rightarrow (u, v)$ 是正交的, 故有 $x^2 + y^2 = u^2 + v^2$, 因此

$$\begin{aligned} p(u, \theta) &= \int_{-\infty}^{+\infty} e^{-u^2} e^{-v^2} dv \\ &= \sqrt{\pi} e^{-u^2} \end{aligned}$$

例 18.2 求 $g(x, y)$ 的 Radon 变换。

$$g(x, y) = \begin{cases} (1 - x^2 - y^2)^{\lambda-1} & x^2 + y^2 < 1 \\ 0 & \text{其他} \end{cases}$$



解: $g(x, y)$ 在单位圆内不为零, 由于 $x^2 + y^2 = u^2 + v^2$, 因此对 u 的积分极限为 $\pm \sqrt{1-u^2}$ 。因此

$$p(u, \theta) = \int_{-\sqrt{1-u^2}}^{\sqrt{1-u^2}} [1 - u^2 - v^2]^{\lambda-1} dv$$

又有如下的定积分 [Erdelyi]

$$\int_{-a}^a (a^2 - t^2)^{\lambda-1} dt = \frac{a^{2\lambda-1} \sqrt{\pi} \Gamma(\lambda)}{\Gamma(\lambda + \frac{1}{2})}$$

利用上式结果可得

$$p(u, \theta) = \begin{cases} \frac{\sqrt{\pi} \Gamma(\lambda)}{\Gamma(\lambda + \frac{1}{2})} (1 - u^2)^{\lambda-1/2} & -1 < u < 1 \\ 0 & \text{其他} \end{cases}$$

18.4 逆 Radon 变换

$g(x, y)$ 的二维逆傅里叶变换式如下式所示:

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(f_x, f_y) e^{j2\pi(f_x x + f_y y)} df_x df_y \quad (18.4-1)$$

将 (f_x, f_y) 转换成极坐标形式 (f, θ) 得

$$g(x, y) = \int_0^\pi \int_{-\infty}^{+\infty} G(f \cos \theta, f \sin \theta) e^{j2\pi f(x \cos \theta + y \sin \theta)} |f| df d\theta \quad (18.4-2)$$

由于 $P(f) = G(f \cos \theta, f \sin \theta)$, 因此有

$$g(x, y) = \int_0^\pi \int_{-\infty}^{+\infty} f P(f) \operatorname{sgn}(f) e^{j2\pi f(x \cos \theta + y \sin \theta)} df d\theta \quad (18.4-3)$$

由卷积定理, 利用 $u = x \cos \theta + y \sin \theta$, 可得

$$\begin{aligned} \int_{-\infty}^{+\infty} f P(f) \operatorname{sgn}(f) e^{j2\pi f u} df &= \frac{1}{2j\pi} \frac{\partial}{\partial u} p(u) * \left(\frac{-1}{j\pi u} \right) \\ &= \frac{1}{2\pi^2} \int_{-\infty}^{\infty} \frac{\partial p(\tau)}{\partial \tau} \frac{1}{u - \tau} d\tau \end{aligned} \quad (18.4-4)$$

最后的结果为 $\frac{1}{2\pi} \frac{\partial}{\partial u} p(u)$ 的希尔伯特变换式, 令

$$\hat{p}(u) = \frac{1}{2\pi^2} \int_{-\infty}^{\infty} \frac{\partial p(\tau)}{\partial \tau} \frac{1}{u - \tau} d\tau \quad (18.4-5)$$

$\hat{p}(u)$ 称为滤波投影。于是, 逆 Radon 变换为

$$\begin{aligned} g(x, y) &= \int_0^\pi \hat{p}(u) d\theta \\ &= \int_0^\pi \hat{p}(x \cos \theta + y \sin \theta) d\theta \end{aligned} \quad (18.4-6)$$

这等价于 $\hat{p}(u)$ 沿角 θ 方向的反向投影, 这意味着当 $g(x, y)$ 不是带限信号时, 如需要完美重构 $g(x, y)$, 则所有 θ 角的投影都需要是已知的。



18.5 Radon 变换的性质

在实际应用中, Radon 变换有许多非常有用的性质。一些最重要的性质总结如下。必要时, $p(u)$ 可写为 $p_g(u)$, 以表明对应于信号 $g(x, y)$ 。

特性 1: 线性

若 $h(x, y) = g_1(x, y) + g_2(x, y)$, 则

$$p_h(u) = p_{g_1}(u) + p_{g_2}(u) \quad (18.5-1)$$

特性 2: 周期性

$$p(u, \theta) = p(u, \theta + 2\pi k), \quad k \text{ 是整数} \quad (18.5-2)$$

特性 3: 质量守恒

$$\iint_{-\infty}^{+\infty} g(x, y) dx dy = \int_{-\infty}^{+\infty} p(u) du \quad (18.5-3)$$

特性 4: 对称性

$$p(u, \theta) = p(-u, \theta \pm \pi) \quad (18.5-4)$$

特性 5: 有界信号

如果对 $|x|, |y| > \frac{D}{2}$, 有 $g(x, y) = 0$, 则

$$p(u) = 0 \text{ 且 } |u| > \frac{D}{\sqrt{2}} \quad (18.5-5)$$

特性 6: 平移

如果 $h(x, y) = g(x - x_0, y - y_0)$,

$$p_h(u, \theta) = p_g(u - u_0, \theta) \quad (18.5-6)$$

其中 $u_0 = x_0 \cos \theta + y_0 \sin \theta$ 。

特性 7: 旋转

在极坐标中 $g(x, y)$ 表示为 $g(r, \phi)$, 如果 $h(r, \phi) = g(r, \phi + \phi_0)$, 则

$$p_h(u, \theta) = p_g(u, \theta + \phi_0) \quad (18.5-7)$$

特性 8: 缩放

若 $h(x, y) = g(ax, ay)$, $a \neq 0$, 则

$$p_h(u) = \frac{1}{|a|} p_g(au) \quad (18.5-8)$$

特性 9: 卷积

若 $h(x, y) = g_1(x, y) * g_2(x, y)$ (其中 $*$ 代表二维卷积)

$$\begin{aligned} p_h(u) &= p_{g_1}(u) * p_{g_2}(u) \\ &= \int_{-\infty}^{+\infty} p_{g_1}(\tau) p_{g_2}(u - \tau) d\tau \\ &= \int_{-\infty}^{+\infty} p_{g_1}(u - \tau) p_{g_2}(\tau) d\tau \end{aligned} \quad (18.5-9)$$

此性质用于由一维线性滤波器实现二维线性滤波器。



18.6 利用投影重构信号

信号由其 Radon 逆变换给出。对于计算机重构，必须使 θ 和 u 离散化，此时主要的考虑因素是需要的采样数。

假设投影在角 $\theta_0, \theta_1, \theta_2, \dots, \theta_{N-1}$ 可用。为了估计投影数 N 的值，信号假定为空间有限和带限（尽管这在理论上是不可可能的，但在实际中有良好的近似结果）。因此假设 $g(x, y)$ 为带限的，即 $f_x^2 + f_y^2 \geq F_0^2$ 时 $G(f_x, f_y) = 0$ 。当 $\Delta x \leq 1/2F_0$ 且 $\Delta y \leq 1/2F_0$ 时，由二维采样定理， $g(x, y)$ 可以用 $g(\Delta x m, \Delta y n)$ 方式采样而不会造成信息丢失。

同样，假设 $g(x, y)$ 是空间有限的，即 $x^2 + y^2 \geq T_0^2$ 时 $g(x, y) = 0$ 。当 $\Delta f_x \leq 1/2T_0$ 且 $\Delta f_y \leq 1/2T_0$ 时，由二维采样定理， $G(f_x, f_y)$ 可以以 $G(\Delta f_x m, \Delta f_y n)$ 方式采样而不会造成信息丢失。

考虑图 18.4 中极坐标中的栅格状采样 [Dudgeon]。在最外面有 N 个投影的圆上，采样点之间的距离可估算为 $\pi F_0/N$ 。假设这个距离采用与 Δf_x 和 Δf_y 相同的采样准则。则

$$\frac{\pi F_0}{N} \leq \frac{1}{2T_0} \quad (18.6-1)$$

信号域内采样间隔近似为 $\Delta = \Delta x = \Delta y \approx 1/2F_0$ ，投影数 N 需满足

$$N \geq \frac{\pi T_0}{\Delta} \quad (18.6-2)$$

式 (18.6-2) 应考虑作为经验法则。

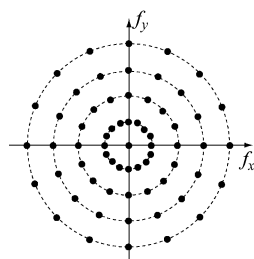


图 18.4 在傅里叶域采样时，在相同的采样频率下对所有的投影进行采样

18.7 傅里叶重构法

在此假设有 N 个投影，对每个投影采样，每个投影的采样数为 M 。计算每个投影采样的 DFT 变换。在周期性极坐标栅格中，这些 DFT 值被视为 $g(x, y)$ 的 FT 变换的采样，如图 18.5 所示 [Dudgeon 提供]。对这些采样进行插值估计，以得到矩形栅格上的采样数据。最后，计算二维逆 DFT 得到 $g(m, n)$ 。通常，逆 DFT 的数据大小选择为 $g(m, n)$ 每个维的 2 ~ 3 倍。

在进行逆 DFT 变换之前，如之前 14.2 节讨论的那样，选择适当的窗函数以使傅里叶域截断的影响最小化。

最简单的插值方法是零阶或线性插值。每个理想的矩形采样点 (f_x, f_y) 包围着四个极坐标采样点。在零阶插值中，对矩形采样点分配与其最接近的极坐标采样点的采样值。在线性插值中，对四个最接近的极坐标采样点的值加权平均，并分配给矩形采样点。加权因子可以从与两点之间的欧几里得距离成反比获得。

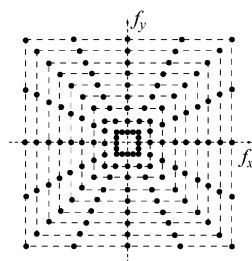


图 18.5 进行插值的极坐标栅格的采样点



一种简单的方法是在极坐标形式下使用双线性内插法：让 (f, θ) 为理想的内插采样点坐标；令 $(f_i, \theta_j), i, j = 1, 2$ 分别表示四个最接近的极坐标采样点的坐标。然后，双线性插值可以写成

$$G(f, \theta) = \sum_{i=1}^2 \sum_{j=1}^2 G(f_i, \theta_j) h_1(f - f_i) h_2(\theta - \theta_j) \quad (18.7-1)$$

其中

$$\begin{aligned} h_1(f') &= 1 - \frac{|f'|}{\Delta f} \quad |f'| \leq \Delta f \\ h_2(\theta') &= 1 - \frac{|\theta'|}{\Delta \theta} \quad |\theta'| \leq \Delta \theta \end{aligned} \quad (18.7-2)$$

式中， $\Delta f, \Delta \theta$ 是极坐标中的采样间隔。

由于极坐标采样点的密度在高频时小于低频时，因此高频时内插的精度要低一些。这会导致在重构图像时产生一些重构质量的退化。

18.8 滤波-反投影算法

逆 Radon 变换可以写成

$$g(x, y) = \int_0^\pi \hat{p}(x \cos \theta + y \sin \theta) d\theta \quad (18.8-1)$$

其中

$$\hat{p}(u) = \int_{-\infty}^{+\infty} |f| P(f) e^{j2\pi fu} df \quad (18.8-2)$$

$P(f)$ 在区间 $[-F_0, F_0]$ 外等于零。它的采样间隔大约为 $\Delta f = 2F_0/N$ 。 $p(u)$ 的采样间隔为 $\Delta \leq 1/2F_0$ 。选择 $\Delta f \Delta = 1/N$ ， $P(\Delta f k)$ 可通过大小为 N 的 DFT 变换来近似

$$P(k) = P(\Delta f k) = \Delta \sum_{n=-N/2+1}^{N/2} p(\Delta n) e^{-j2\pi nk/N}, \quad \frac{-N}{2} \leq k \leq \frac{N}{2} \quad (18.8-3)$$

或者

$$P(k) = \frac{1}{2F_0} \sum_{n=0}^{N-1} p(n) e^{-j2\pi nk/N}, \quad k = 0, 1, \dots, (N-1) \quad (18.8-4)$$

这是因为可以选择 $p(-|l|)$ 和 $P(-|m|)$ 分别等于 $P(N - |l|)$ 和 $P(N - |m|)$ 。

式 (18.8-2) 可以写为

$$\hat{p}(u) = \int_{-F_0}^{F_0} |f| P(f) e^{j2\pi fu} df \quad (18.8-5)$$

$\hat{p}(u)$ 可通过采用相同采样间隔 Δ 的逆 DFT 变换近似得到

$$\hat{p}(u) = \hat{p}(n\Delta) = \Delta f \sum_{k=-N/2+1}^{N/2} P(\Delta f k) |k\Delta f| e^{j2\pi nk/N} \quad (18.8-6)$$

或者

$$\hat{p}(n) = \Delta f \sum_{k=0}^{N-1} P'(k) e^{j2\pi nk/N}, \quad n = 0, 1, \dots, (N-1) \quad (18.8-7)$$

其中



$$P'(k) = \begin{cases} P(k) |k\Delta f| & 0 \leq k \leq \frac{N}{2} \\ P(k) |(N-k)\Delta f| & \frac{N}{2} < k < N \end{cases} \quad (18.8-8)$$

如果对 $P(k)$ 乘以 14.2 节中讨论过的窗函数，就能获得更好的重构结果。窗函数减小了高频时严重的测量噪声。

最后，通过式 (18.8-1) 的数值积分即可获得重构图像。如果有 M 个投影， $g(x, y)$ 可以近似为 $g'(x, y)$

$$g'(x, y) = \frac{\pi}{M} \sum_{k=1}^M \hat{p}(x \cos \theta_k + y \sin \theta_k) \quad (18.8-9)$$

$u = (x \cos \theta_k + y \sin \theta_k)$ 可能与式 (18.8-3) 中的 $p(\Delta n)$ 不一致。因此有必要再次插值以从 $p(\Delta n)$ 得到 $p(u)$, $n = 0, 1, 2, \dots, (N-1)$ 。线性插值通常能够满足此目的。

上述算法有一些不足。对 N 个点使用 DFT 变换意味着式 (18.8-5) 的线性卷积被“局部”圆周卷积所取代。使用“局部”这个词，是因为 $|f|$ 的离散化。如果大小为 N 的 DFT 变换由通过补零、大小为 $2N$ 的 DFT 变换取代，会改善重构结果，因为圆周卷积按输出的要求变为了线性卷积。

然而，如果首先在时域内写出如式 (18.8-5) 的变换域的卷积形式，然后整个过程用零填充值进行 DFT，可以得到最好的结果。

式 (18.8-5) 的变换函数由下式给出：

$$H(f) = \begin{cases} |f| & |f| \leq F_0 \\ 0 & \text{其他} \end{cases} \quad (18.8-10)$$

相应的脉冲响应由下式给出：

$$\begin{aligned} h(t) &= \int_{-F_0}^{F_0} H(f) e^{j2\pi ft} df \\ &= 2F_0^2 \frac{\sin 2\pi t F_0}{2\pi t F_0} - F_0^2 \left(\frac{\sin \pi t F_0}{\pi t F_0} \right)^2 \end{aligned} \quad (18.8-11)$$

在 $t = n\Delta t$ 时，离散的脉冲响应如下：

$$h(n\Delta t) = \begin{cases} F_0^2 & n = 0 \\ 0 & n \text{ 是偶数} \\ -\frac{4F_0^2}{n^2 \pi^2} & n \text{ 是奇数} \end{cases} \quad (18.8-12)$$

式 (18.4-5) 作为时域卷积如下：

$$\hat{p}(u) = \int_{-\infty}^{+\infty} p(\tau) h(u - \tau) d\tau \quad (18.8-13)$$

它的离散形式由下式给出：

$$\hat{p}(n\Delta t) \approx \Delta t \sum_{m=-\infty}^{\infty} p(\Delta t m) h((n - m)\Delta t) \quad (18.8-14)$$

由于当 $|m| > N/2$ 时 $p(\Delta t m)$ 为零，因此

$$\hat{p}(n\Delta t) \approx \Delta t \sum_{m=-N/2}^{N/2} p(\Delta t m) h((n - m)\Delta t) \quad (18.8-15)$$



这种线性卷积现在可以通过对 $p(\cdot)$ 和 $h(\cdot)$ 补零后，进行大小为 $2N$ 的 DFT 获得。通过对频域结果适当加窗，可能得到更好的结果。整个过程如下所示：

(1) 对 $p(\cdot)$ 和 $h(\cdot)$ 进行如下形式的补零：

$$p(n) = \begin{cases} p(n) & 0 \leq n \leq \frac{N}{2} \\ 0 & \frac{N}{2} < n < \frac{3N}{2} \\ p(n-2N) & n \geq \frac{3N}{2} \end{cases} \quad (18.8-16)$$

(2) 计算 $p(\cdot)$ 和 $h(\cdot)$ 的大小为 $2N$ 的 DFT。

(3) 进行变换域运算。

(4) 用一个适当的窗函数对步骤 (3) 的结果加窗。

(5) 计算大小为 $2N$ 的逆 DFT。

第 19 章 密集波分复用

19.1 本章引论

近年来,光学傅里叶技术已在光通信和网络领域得到广泛应用。其中一个应用领域就是在密集波分复用(DWDM)系统中得到应用的阵列波导光栅(Arrayed Waveguide Grating, AWG)技术。DWDM 提供了光通信和网络的容量与灵活性问题的新的解决方向。它提供了非常大的传输容量和新颖的网络架构 [Brackett, 1990] [Brackett, 1993]。DWDM 系统的主要组成部分是波长复用器和解复用器。市场上可以买到的组件均采用了光纤光学或微光学技术 [Pennings, 1995] [Pennings, 1996]。

对集成光学(解)复用器的研究越来越集中于基于光栅和相控阵(PHASAR)的器件(也称为阵列波导光栅) [Laude, 1993] [Smit, 1988]。这两种器件均为成像器件,它们将输入波导的场图像以色散方式成像到输出波导阵列。基于光栅的器件中,垂直蚀刻的反射光栅可以提供解复用所需的聚焦和色散特性。基于相控阵的器件中,上述特性由波导阵列来提供,可以选择适当的波导阵列长度以满足所要求的成像和色散特性。

由于利用传统的波导技术即可制作相控阵器件,不需要基于光栅的设备中所需的垂直蚀刻工序,因此相控阵器件具有更好的鲁棒性和加工容错性。第一台工作在短波长下的相控阵器件由 Vellekoop 和 Smit 报道 [1989] [Verbeek 和 Smit, 1995]。Takahashi 等 [1990] 报道了第一个工作在约 $1.6\mu\text{m}$ 的长波长窗口下的相控阵器件。Franco Dragone [1991] 将相控阵的概念从 $1 \times N$ 扩展到 $N \times N$ 。

本章共分 6 节。在 19.2 节中讨论阵列波导光栅原理,主要讨论可用信道数目是受限制的情况,因为特定波长的聚焦光束在周期性位置上重复。19.3 节讨论了一种称为不规则采样的零交叉法(MISZC),该方法可以显著减少信道数受限的问题。19.4 节对 MISZC 的性能进行了分析。19.5 节中介绍了二维和三维情况下利用上述方法的计算机仿真。二维和三维实施过程中需要注意的问题将在 19.6 节阐述。

19.2 阵列波导光栅

基于 AWG 的复用器和解复用器在本质上是相同的。根据互易性原理,并且依赖于波的传播方向,该器件可用作复用器或解复用器。为简单起见,在这里讨论解复用器的工作原理。

AWG 由输入/输出两个波导阵列、两个聚焦平板区以及阵列波导组成,如图 19.1 所示。单一光纤连接到输入波导阵列作为多波长输入端。多波长输入信号在输入波导间被均匀地分割,并且信号通过该输入波导传播到输入聚焦平板区。光波穿过聚焦板,并耦合到阵列波导。光波通过每个阵列波导时会发生线性相移,这是由于阵列波导之间的光程存在差异。



随后光波耦合到输出聚焦板，由于衍射与色散，其中的多波长输入信号按波长不同分成不同光束。通过选择阵列波导的长度和两个相邻波导之间的光程差 ΔL ，使中心波长的光波穿过每一个阵列波导的相位延迟都是 $2\pi m$ ，其中 m 是衍射阶数，为整数。中心波长以外的光波与中心波长的光波相位延迟是不同的。因此，每个波长分量都有自己的波阵面，并被聚焦到输出聚焦板区域上输出侧的不同位置。然后各个波长分量分别输入一个输出波导。

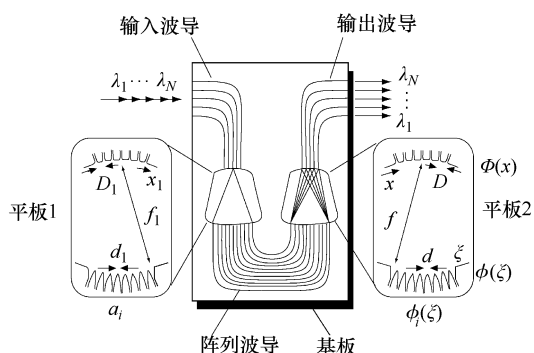


图 19.1 阵列波导光栅装置示意图 [由 Okamoto 提供]

下面给出解复用的一个粗略分析。在输入和输出聚焦板区域的定义如下：
输入聚焦板区域：

- D_1 ：相邻波导末端之间的距离；
- d_1 ：输出侧相邻波导末端之间的距离；
- x_1 ：距离输入侧中心的测量距离；
- f_1 ：输入侧曲面的曲率半径。

输出聚焦板区域：

- d ：连接到阵列波导的相邻波段末端的间距；
- D ：连接到输出波导的相邻波段末端的间距；
- f ：输出曲面的曲率半径。

需要重申的是，两个相邻波导之间的光程差为 ΔL ，相应的对应于中心波长的相位延迟是 $2\pi m$ 。

考虑光束穿过第 i 个和第 $(i-1)$ 个阵列波导。为了使两个光束有效地干涉，当它们都到达聚焦板区域的输出侧时，它们的相位差应为 2π 的整数倍。相长干涉的条件由下式给出：

$$\begin{aligned} & \beta_s(\lambda_0) \left[f_1 - \frac{d_1 x_1}{2f_1} \right] + \beta_c(\lambda_0) [L_c + (i-1)\Delta L] + \beta_s(\lambda_0) \left[f + \frac{dx}{2f} \right] \\ &= \beta_s(\lambda_0) \left[f_1 + \frac{d_1 x_1}{2f_1} \right] + \beta_c(\lambda_0) [L_c + i\Delta L] + \beta_s(\lambda_0) \left[f - \frac{dx}{2f} \right] - 2\pi m \end{aligned} \quad (19.2-1)$$

式中， β_s 和 β_c 分别表示平板区和阵列波导的传播常数（波数）； m 为衍射阶数； λ_0 是多波长输入的中心波长； L_c 为最小的阵列波导长度。

从式 (19.2-1) 中减去公共项，得到

$$\beta_s(\lambda_0) \frac{d_1 x_1}{f_1} - \beta_s(\lambda_0) \frac{dx}{f} + \beta_c(\lambda_0) \Delta L = 2\pi m \quad (19.2-2)$$

当条件

$$\beta_c(\lambda_0) \Delta L = 2\pi m \quad (19.2-3)$$

对波长 λ_0 满足时，光输入位置 x_1 和输出位置 x 满足如下条件：

$$\frac{d_1 x_1}{f_1} = \frac{dx}{f} \quad (19.2-4)$$



上述方程表明, 当光耦合到输入位置 x_1 时, 输出位置 x 由方程式 (19.2-4) 确定。

光程差 ΔL 可以表示为

$$\Delta L = \frac{n_s d D \lambda_0}{n_o f \Delta \lambda} \quad (19.2-5)$$

对相同波长光的第 m 阶和第 $(m+1)$ 阶聚焦光束的空间分离间隔 (自由光谱范围) 可从式 (19.2-2) 推导出 [Okamoto]:

$$X_{FSR} = x_m - x_{m+1} = \frac{\lambda_0 f}{n_s d} \quad (19.2-6)$$

可用的波长信道数 N_{ch} 可由 X_{FSR} 除以输出波导间距 D 获得

$$N_{ch} = \frac{X_{FSR}}{D} = \frac{\lambda_0 f}{n_s d D} \quad (19.2-7)$$

在实际中, 实现平面上的波导布局应能保证两个波导之间的长度差为 $m\lambda$, 这并不是一项简单的任务。常用专业的计算机辅助设计程序实现这一目的。图 19.2 给出了一个实例, 它使用了 Rsoft 公司的 BeamPROP 软件包来实现该设计 [Lu, Ersoy]。

图 19.3 和图 19.4 给出了在 $1.55\mu\text{m}$ 的中心波长和 $0.8\mu\text{m}$ 的信道间隔条件下, 利用 PHASAR 模拟得到的结果, 它们显示出为什么信道的数量是受限的 [Lu, Ersoy 等, 2003]。在图 19.3 中, 共有 16 个信道, 并且在中心信道任一侧的第二信道不与中心信道重叠。而在图 19.4 中, 有 64 个信道, 在中心信道任一侧的第二信道开始与中心信道重叠。在这种特殊情况下, 信道的数目不能再继续增加。目前, 市售 PHASAR 器件有 40 个左右的信道。

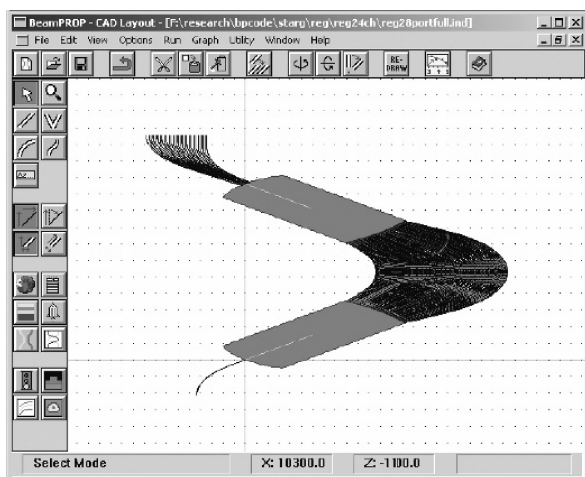


图 19.2 常用的 180 信道 AWG 布局设计, 由 BeamPROP 设计 [Lu, Ersoy 等, 2003]

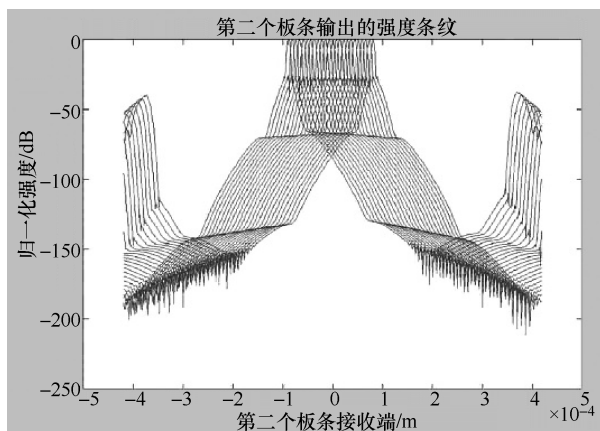


图 19.3 在 $1.55\mu\text{m}$ 的中心波长, $0.8\mu\text{m}$ 信道间隔的情况下, 16 信道 PHASAR 的输出 [Lu, Ersoy]

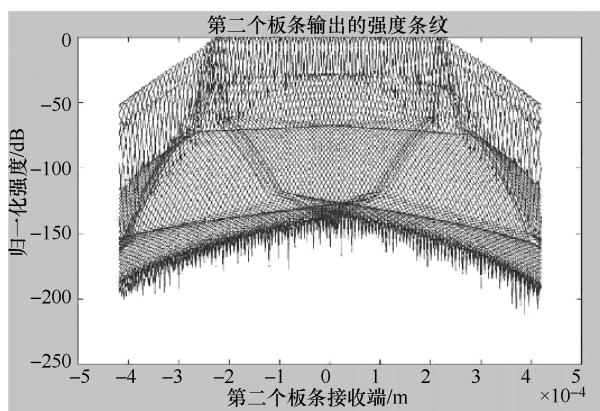


图 19.4 在 $1.55\mu\text{m}$ 的中心波长, $0.8\mu\text{m}$ 信道间隔下, 一个 64 信道 PHASAR 的输出 [Lu, Ersoy]

19.3 不规则采样的零交叉法 (MISZC)

在该方法中, DWDM 器件的设计保证每个波长都只有一幅图像, 于是, 由上述方程 (19.2-7) 给出的相邻两个衍射阶之间的间隔所限制的信道数将不再受限 [Lu, Ersoy] [Ersoy, 2005]。在焦点的光束称为图像。下面将特别讨论在相位调制的情况下, 也就是在平面和球面参考波的情况下, 如何实现该设计。图 19.5 示出了参考波和涉及两种不同波长时的模型。

一旦所需相位计算出来, 在使用 AWG 的情况下就可通过选择每个波导的长度, 来实现所需要的相位。在共焦或罗兰结构中, 这种方法仅由线性光栅相位调制的常规 PHASAR 器件来实现。在下面将首先讨论该方法用于一个平面结构的例子, 也就是说相控阵孔径将放置在一个平面上 (对二维情况将是相控阵线)。在后面的章节将给出共焦、罗兰和三维结构的结果。

在该方法中, 首先随机选择辐射孔径的中心位置, 然后在选定的位置创建波阵面的负相位 (可能需要加一个常数), 使整体的相位是零 (或恒定), 或略微调整辐射孔的中心位置, 使总相移从该中心点到期望图像点移动并等于一个常数, 换言之, 使模除 2π 为零。在这两种方法中, 这样的点称为零交叉。在第二种方法中, 它们被称为自动零交叉。在实际应用中, 将采用如下所述的半不规则采样点。

零交叉的总数可以是一个非常大的数字, 特别是存在平面和球面相位调制的情况下。实际设计中只允许有少量的孔径, 例如, 典型 PHASAR 的数量是 300 个。为避免过多孔径的问题和因规则采样产生的谐波问题 [Ishimaru, 1962] [Lao, 1964], 这里选择不规则采样的稀疏数量的孔径。一种确定零交叉位置的方法, 由下列的步骤给出。

步骤 1: 沿相控阵面每次随机选择一个孔径点来进行孔径点的初始化操作。为了系统地

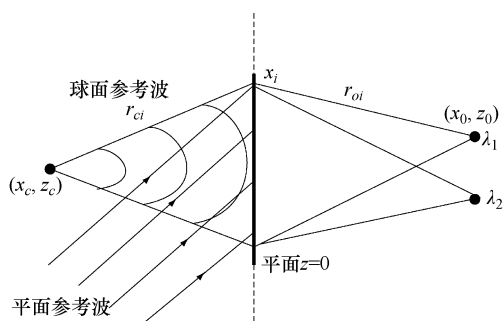


图 19.5 MISZC 涉及两种波长时参考波的可视化几何模型



在整个相控阵表面上实现孔径点的初始化,可以使用下面的方法:

初始点的位置 = 均匀分布的点位 + 小的随机平移

步骤 2 (a): 如果在选定的位置需要创建负相位,则此相位可以在物理上实现,例如,对 PHASAR 器件正确地选择波导的长度。

步骤 2 (b): 如果使用自动零交叉的方法,则计算步骤 1 生成的每个初始点值的修正值,以找到最近的零交叉点:

零交叉点的最终位置 = 第 1 步开始的初始点位置 + 校正项

这两种方法原理很相似。下面讨论一种计算自动零交叉法的修正项的算法。

19.3.1 修正项的计算方法

包含用于成像的线性光栅的相控阵方程由下式给出:

$$\delta x_i + \phi_i(x_i) + k r_{oi} = 2\pi n + \phi_0 \quad (19.3-1)$$

其中

$$\phi_i(x_i) = k r_{ci} \quad (19.3-2)$$

$$r_{oi} = \sqrt{(x_0 - x_i)^2 + z_0^2} \quad (19.3-3)$$

$$r_{ci} = \sqrt{(x_c - x_i)^2 + z_c^2} \quad (19.3-4)$$

式 (19.3-1) 可以写成

$$\delta x_i + r_{ci} + r_{oi} = n\lambda + \phi_0\lambda/2\pi \quad (19.3-5)$$

对于相控阵上的任意位置 x_i , 式 (19.3-5) 变为

$$\delta x_i + r_{ci} + r_{oi} = n\lambda + \phi_0\lambda/2\pi + B \quad (19.3-6)$$

式中, B 代表误差。

假定该孔径的位置沿 x 方向上移动 Δ 距离,使其满足所述相控阵方程式 (19.3-5), 则有

$$\delta x'_i + r'_{ci} + r'_{oi} = \delta x_i + r_{ci} + r_{oi} - B \quad (19.3-7)$$

其中 r'_{oi} 和 r'_{ci} 由式 (19.3-3) 和式 (19.3-4) 给出。由于 $x'_i = x_i + \Delta$, 可得

$$r'_{oi} = \sqrt{r_{oi}^2 - 2\Delta(x_0 - x_i) + \Delta^2} \quad (19.3-8)$$

$$r'_{ci} = \sqrt{r_{ci}^2 - 2\Delta(x_c - x_i) + \Delta^2} \quad (19.3-9)$$

使用这些方程推导出 Δ 的四阶多项式方程

$$\Delta^4 F_4 + \Delta^3 F_3 + \Delta^2 F_2 + \Delta F_1 + F_0 = 0 \quad (19.3-10)$$

其中

$$F_4 = G_1^2 - 1 \quad (19.3-11)$$

$$F_3 = 2G_1G_2 + 2X_c + 2X_o \quad (19.3-12)$$

$$F_2 = G_2^2 + 2G_1G_3 - 4X_cX_o - r_{oi}^2 - r_{ci}^2 \quad (19.3-13)$$

$$F_1 = 2G_2G_3 + 2X_c r_{oi}^2 + 2X_o r_{ci}^2 \quad (19.3-14)$$

$$F_0 = G_3^2 - r_{oi}^2 r_{ci}^2 \quad (19.3-15)$$

$$G_1 = \frac{\delta^2}{2} - 1 \quad (19.3-16)$$

$$G_2 = -(r_{ci} + r_{oi} - B)\delta + X_c + X_o \quad (19.3-17)$$



$$G_3 = r_{ci}r_{oi} + \frac{B^2}{2} - B(r_{ci} + r_{oi}) \quad (19.3-18)$$

Δ 可由求解四阶多项式 (19.3-10) 的根来得到。有意思的是, 如 19.2 节所讨论的, 在没有因光栅而产生的线性相位调制的情况下, 这个方程可以简化为一个二阶多项式方程。对 PHASAR 器件, 选择的零交叉采样点位置即对应于相控阵表面上的波导孔径的位置。

19.3.2 推广 MISZC 到三维结构

因为在其他技术中也需要用到不规则采样的零交叉法, 所以有必要将讨论扩展到 3-D 几何形状。例如, 可以将阵列波导布置在一个二维平面或一个二维弧面, 而不是像上面讨论的那样, 沿着一维直线布置。其他的一些技术, 如扫描电子显微镜 [Ersoy, 1979] 和用于制造衍射光学元件的反应离子蚀刻等都是潜在的应用方向。三维情形下的基本方法从概念上与之前的讨论相同。换句话说, 首先 (半) 随机地选择辐射孔径的中心位置; 其次, 在物理上赋予选定位置以一个负相位 (可能加一个常数), 使整体的相位为零 (或常数), 或略微调整辐射孔径的中心的位置, 使从该中心到期望图像点的总相移等于一个恒定值, 例如, 取模除 2π 为零的值。

选择自动零交叉时, 如果只沿 x 方向进行调整, 则方程式 (19.3-3) 和方程式 (19.3-14) 仍然是有效的, 只需要做如下替换:

$$\begin{aligned} z_0^2 &\rightarrow z_0^2 + (y_0 - y_i)^2 \\ z_c^2 &\rightarrow z_c^2 + (y_c - y_i)^2 \end{aligned} \quad (19.3-19)$$

19.4 对 MISZC 的分析

在上面讨论的方法中, 更高阶谐波图像问题得到了最小化, 在本节中, 将给出三维情形下的分析来对此进行解释。在如光学 PHASAR 等平面器件中, 需要使用二维的分析。这些分析可以由忽略一个维度来简单地实现, 也就是说, 可简单地从该方程中去除变量 y 来实现。

MISZC 是一种非线性的编码方法。在一般情况下, 使用这样的编码技术, 谐波图像由以下两种机制产生: ①规则采样和②非线性编码。在 MISZC 中, 规则采样导致的谐波图像转化为可容忍的背景噪声 [Doles, 1988]。存在相位调制的情况下, 采用零交叉的非线性编码导致的谐波图像也可以被消除, 原因分析如下。

更一般的是, 式 (19.3-1) 可写为

$$\varphi(x_i, y_i) + \theta(x_i, y_i) + kr_{oi} = 2n\pi + \phi_0 \quad (19.4-1)$$

式中, $\varphi(x_i, y_i)$ 表示由波的传播产生的相移, 该相移是从球面参考波 (x_c, y_c, z_c) 的起点到相控阵列表面上第 i 个耦合孔径 (x_i, y_i) 之间的相移; $\theta(x_i, y_i)$ 是另一相移, 例如, 式 (19.3-1) 中的线性相移; kr_{oi} 表示由波的传播产生的相移, 是相控阵列表面上的孔径 (x_i, y_i) 到位于 (x_o, y_o, z_o) 的图像点 (物点) 之间的相移。

在 PHASAR 器件中, $\theta(x_i, y_i)$ 可以表示为 $n_c \alpha k x_i$, 其中, n_c 为波导内的有效折射率。

对中心波长 λ , 式 (19.4-1) 可写为

$$kr_{ci} + nkx_i \alpha + kr_{oi} = 2n\pi + \phi_0 \quad (19.4-2)$$



根据傍轴近似, 有

$$r_{oi} = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2 + z_0^2} \cong z_0 + \frac{x_0^2 + y_0^2 + x_i^2 + y_i^2}{2z_c} - \frac{x_0x_i + y_0y_i}{z_0} \quad (19.4-3)$$

$$r_{ci} = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2 + z_c^2} \cong z_c + \frac{x_c^2 + y_c^2 + x_i^2 + y_i^2}{2z_c} - \frac{x_cx_i + y_cy_i}{z_c}$$

将式 (19.4-3) 代入式 (19.4-2), 并忽略常数相位项, 得到

$$x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x_0}{z_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z_0} + \frac{1}{z_c}\right) = n\lambda \quad (19.4-4)$$

式中, $\delta = n_c\alpha$ 。

假设波长由 λ 变成 λ' , 式 (19.4-4) 对另一图像点 (x'_0, z'_0) 仍然有效。关于 λ 和 λ' 的两个方程相比, 可得

$$\frac{x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x_0}{z_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z_0} + \frac{1}{z_c}\right)}{x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x'_0}{z'_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z'_0} + \frac{1}{z_c}\right)} = \frac{\lambda}{\lambda'} = R \quad (19.4-5)$$

令含 x_i 项的系数相等, 则可得到新的焦点 (x'_0, z'_0) :

$$z'_0 = \frac{R}{\frac{1-R}{z_c} + \frac{1}{z_0}} \approx Rz_0 \quad (19.4-6)$$

$$x'_0 = \frac{\frac{x_0}{z_0} - (1-R)\left(\delta - \frac{x_c}{z_c}\right)}{\frac{1-R}{z_c} + \frac{1}{z_0}} \approx x_0 - z_0(1-R)\left(\delta - \frac{x_c}{z_c}\right) \quad (19.4-7)$$

其中近似的条件是 $1-R \ll 1$ 且 $z_c \ll z_0$ 。

从上面的推导可以看到, 焦点的位置 z'_0 非常接近原始点 z_0 。沿 x 方向, 色散关系由下式给出:

$$\Delta x_0 = x'_0 - x_0 = -z_0(1-R)\left(\delta - \frac{x_c}{z_c}\right) \quad (19.4-8)$$

$$\left|\frac{\Delta x_0}{\Delta \lambda}\right| \approx \frac{z_0}{\lambda} \left|\delta - \frac{x_c}{z_c}\right| \quad (19.4-9)$$

成像方程满足下述条件时, 采用零交叉的非线性编码, 可产生更高阶谐波图像点

$$x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x'_0}{z'_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z'_0} + \frac{1}{z_c}\right) = nm\lambda' \quad (19.4-10)$$

以傍轴近似, 求式 (19.4-4) 和式 (19.4-10) 的比值, 有

$$\frac{x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x_0}{z_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z_0} + \frac{1}{z_c}\right)}{x_i\delta - x_i\left(\frac{x_c}{z_c} + \frac{x'_0}{z'_0}\right) + \frac{x_i^2}{2}\left(\frac{1}{z'_0} + \frac{1}{z_c}\right)} = \frac{\lambda}{m\lambda'} = \frac{R}{m} \quad (19.4-11)$$

x'_0, z'_0 也用同样的方法求解, 更高阶的谐波图像点位置由下式获得:



$$z'_0 = \frac{R}{\frac{m-R}{z_c} + \frac{m}{z_0}} \quad (19.4-12)$$

$$x'_0 = \frac{\frac{mx_0}{z_0} - (m-R)\left(\delta - \frac{x_c}{z_c}\right)}{\frac{m-R}{z_c} + \frac{m}{z_0}} \quad (19.4-13)$$

从上面的公式可知，当谐波阶数增加时 z_0 变小，此时 z 方向的成像位置发生显著移动。这意味着高阶谐波在向非常接近相控阵的位置移动。然而，与相控阵的距离太近时，傍轴近似将失效。因此，此时不再有任何有效的成像方程。于是，高阶谐波变为噪声。当然可以这样认为，即使傍轴近似无效，可能仍有一些成像公式成立。然而，在 19.4 节中的模拟分析结果表明，不存在这样有效的成像方程，并且高阶谐波图像可变成噪声的结论是有效的。即使它们在非常接近相控阵列的位置成像，也在那些相对较远的图像点位置表现得像背景噪声。19.4 节的模拟结果表明，在这种噪声情况下，信噪比是符合要求的，而且随着信道的增加仍然符合要求。

19.4.1 色散分析

本小节分析基于上一小节的式 (19.4-6)、式 (19.4-7)、式 (19.4-12) 和式 (19.4-13) 的模拟结果。

情况 1：球面波情况 ($0.1 < z_0/z_c < 10$)

对于一阶谐波 ($m=1$)，波长 λ' 下的焦点期望位置为 (x'_0, z'_0) ，其与波长 λ' 具有线性关系。这种线性关系的斜率随比值 z_0/z_c 的减小而减小。对于更高阶谐波 ($m \geq 2$)， x'_0 比 $x_0=0$ 大得多， z'_0 比 z_0 要小很多。因此，可得到以下结论，如上一小节中讨论的那样，此时高次谐波变成背景噪声。

情况 2：平面波情况 ($z_0/z_c \gg 1$)

在这种情况下，式 (19.4-6)、式 (19.4-7)、式 (19.4-12) 和式 (19.4-13) 可以简化为

$$z'_0 = \frac{R}{\frac{m-R}{z_c} + \frac{m}{z_0}} \approx \frac{R}{m} z_0 \approx \frac{1}{m} z_0 \quad (19.4-14)$$

$$x'_0 = \frac{\frac{mx_0}{z_0} - (m-R)\left(\delta - \frac{x_c}{z_c}\right)}{\frac{m-R}{z_c} + \frac{m}{z_0}} \approx x_0 - z_0 \left(1 - \frac{R}{m}\right) \left(\delta - \frac{x_c}{z_c}\right) \quad (19.4-15)$$

于是，推导可得第一阶 ($m=1$) 的色散关系如下：

$$\left| \frac{\Delta z}{\Delta \lambda} \right| \approx \frac{z_0}{\lambda} \quad (19.4-16)$$

$$\left| \frac{\Delta x}{\Delta \lambda} \right| \approx \frac{z_0}{\lambda} \left| \delta - \frac{x_c}{z_c} \right| \quad (19.4-17)$$

三维色散。三维情况下色散的数学推导，与之前讨论过的二维情况非常类似 [Hu,



Ersoy]。但是,在二维情况中是把 y 变量视为常量,在推导中忽略它们,三维情况下需要同时研究 x 变量和 y 变量,然后分别由 x 方向和 y 方向上独立的色散关系导出独立的方程。因此,如果独立选择各点的 x 坐标和 y 坐标,则该色散关系由下式给出:

$$\left| \frac{\Delta x}{\Delta \lambda} \right| \approx \frac{z_0}{\lambda} \left| \delta_x - \frac{x_c}{z_c} \right| \quad (19.4-18)$$

$$\left| \frac{\Delta y}{\Delta \lambda} \right| \approx \frac{z_0}{\lambda} \left| \delta_y - \frac{y_c}{z_c} \right| \quad (19.4-19)$$

19.4.2 有限尺寸孔径

到目前为止的理论讨论中,相控阵的孔径均假设为点光源。在一般情况下,每个孔中的相位变化不大,这是一种较好的假设。另外,由于选择零交叉点作为孔径的中心点,故相位变化有最大容限,例如,在范围 $[-\pi/2, \pi/2]$ 内。在本节中,将考虑 PHASAR 型的器件,其相位调制是通过截断相控阵表面的波导来控制的。

使用圆柱坐标系 (r, ϕ, z) 来表示孔径上的一点,用球坐标系 (R, Θ, Φ) 来表示孔径之外的点。利用定义的这些变量,在夫琅禾费区,辐射场的菲涅耳-基尔霍夫衍射公式为 [Lu, Ersoy, 1993]

$$E_{FF}(R, \Theta, \Phi) = jk \frac{e^{-jkR}}{2\pi R} \frac{1 + \cos\Theta}{2} \int_s E(r, \phi, 0) e^{jkr \sin\Theta \cos(\Phi - \phi)} r dr d\phi \quad (19.4-20)$$

LP₀₁ 模的横向电场可精确地近似为高斯函数

$$E_{GB}(r, \phi, 0) = E_0 e^{-r^2/w^2} \quad (19.4-21)$$

式中, w 是高斯光束的束腰半径。

将式 (19.4-21) 代入式 (19.4-20) 中,可得到高斯辐射场在夫琅禾费区的场分布,其结果为

$$E_{GBFF}(R, \Theta, \Phi) = jk E_0 \frac{e^{-jkR}}{R} \frac{w^2}{2} e^{-(k w \sin\Theta)^2/4} \quad (19.4-22)$$

当孔径尺寸非常小时远场近似成立。在 19.5.2 节中,将利用式 (19.4-22) 对有限尺寸孔径的相控阵进行仿真设计。

19.5 计算实验

首先定义用来说明结果的参数如下:

M : 相控阵的孔径数 (等于 PHASAR 的波导管数);

L : 信道数 (需要解复用的波长数);

$\Delta\lambda$: 通道间的波长间隔;

r : 在范围 $[0, 1]$ 间的随机系数,定义为均匀间距长度 Δ 的小数部分 (因此随机位移的范围是 $[-r\Delta, r\Delta]$)。

仿真结果如图 19.6 ~ 图 19.12 所示。每个图的标题同时包含所用参数的值。除非特别说明, r 假设为 1。在 19.5.1 节中,相控阵孔径假设为点光源。在 19.5.3 节中,将考虑有限尺寸孔径的情况。



19.5.1 点源孔径

图 19.6 显示了当相控阵的中心波长为 1550nm、信道数为 16、相邻信道之间的波长间隔为 0.4nm 时，图像平面上的强度分布和零交叉位置分布。与 19.3 节所声明的一致，输出面上没有观察到谐波图像。

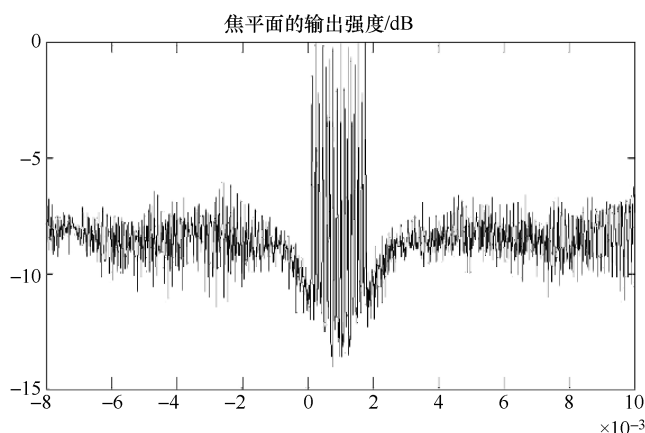


图 19.6 焦平面上的输出强度 (参数为 $M = 100$, $L = 16$, $\delta = 15$, $\Delta\lambda = 0.4\text{nm}$)

为了验证式 (19.4.17) 得出的色散关系，考察 Δx 相对于 $\Delta\lambda$ 、 δ 和不同的 z_0 值的线性关系。模拟结果如图 19.7 所示，得到的直线斜率分别为 1.18×10^6 , 0.78×10^6 , 0.40×10^6 ，这些值与使用式 (19.4.17) (参数为 $\delta = 30$; $\lambda = 1550\text{nm}$) 的理论计算值非常吻合，理论计算值分别为 1.16×10^6 , 0.77×10^6 和 0.39×10^6 。

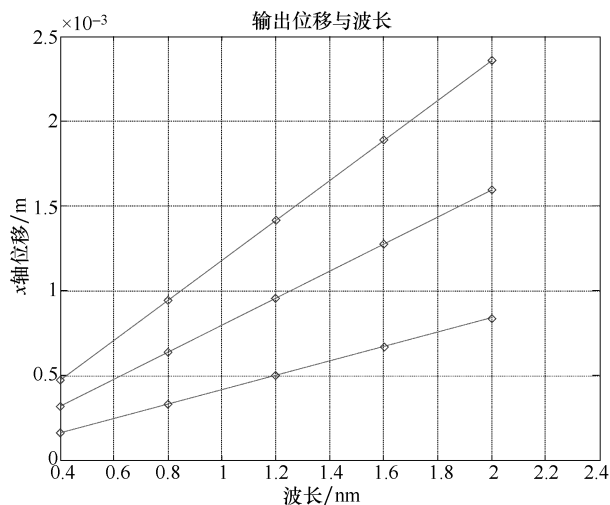


图 19.7 线性色散关系的实验验证 ($M = 100$, $L = 16$, $\delta = 30$)

在 MISZC 中，随机采样和实现零交叉对得到良好的结果都是至关重要的。在下文中，将给出比较结果来讨论较少随机抽样对结果影响的重要性。图 19.8 和图 19.9 给出了在不使用整体随机抽样的情况下的结果。除了参数 r 分别为 0、1/4 和 1/2，其余参数都与图 19.6



中相同。可见，当 r 小于 1 时，即不是完全随机性时，不同阶次的谐波开始出现。相比较而言，如图 19.6 所示，在 $r=1$ 的情况下，并没有出现谐波，因为在这种情况下采用的是完全随机采样。

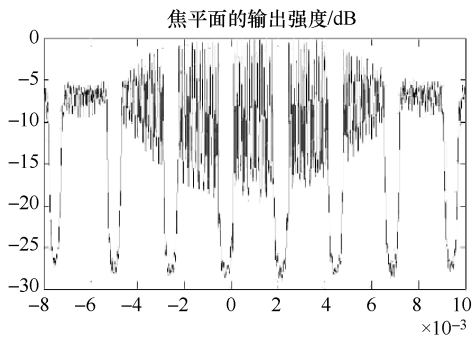


图 19.8 非随机采样的谐波 ($M=100$, $L=16$, $\delta=15$, $\Delta\lambda=0.4\text{nm}$, $r=0$)

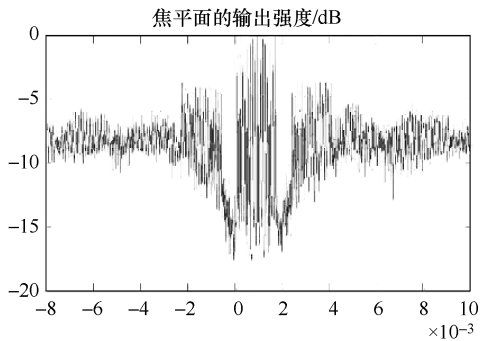


图 19.9 部分随机采样的谐波 ($M=100$, $L=16$, $\delta=15$, $\Delta\lambda=0.4\text{nm}$, $r=0.5$)

19.5.2 多信道

去除谐波图像的主要益处是可尽可能地提高增多信道数目的能力。许多情况下，通常设计 64、128 和 256 信道来研究大信道下的特性。在下面的图中，相控阵孔径数、信道数和波长间隔分别用 M 、 L 和 $\Delta\lambda$ 表示。图 19.10 给出了 $M=200$ ， $L=128$ 和 $\Delta\lambda=0.2\text{nm}$ 时的结果。该图由两部分组成。图 19.10a 显示同时多通道运行时的解复用特性。在该图中，可以看到所有信道之间的不均匀性在 $0\sim 2\text{dB}$ 范围内。在关于 WDM（波分复用）器件的文献中，通常会采用指定的单一信道最坏情况下的串扰图来表征串扰性能。图 19.10b 显示的是对中心输出端口各波长的归一化透射谱。串扰值估计为 20dB 。可见，当采用更多孔径（对 PHASAR 中的波导）时，可改善串扰值。

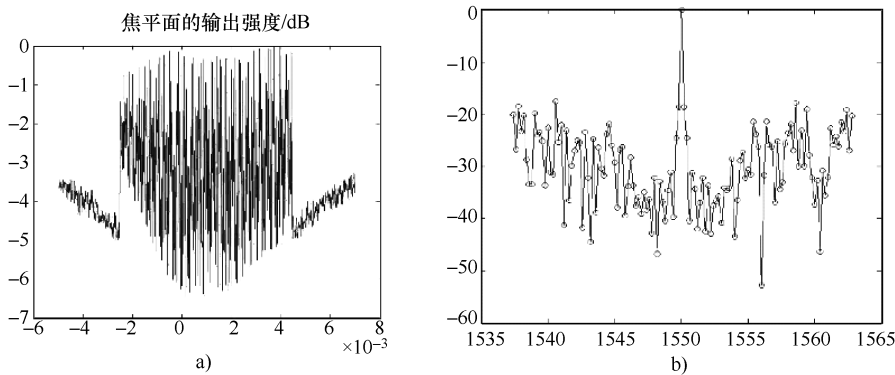


图 19.10 大信道数下的结果 ($M=200$, $L=128$, $\Delta\lambda=0.2\text{nm}$)

19.5.3 有限尺寸孔径

在 19.4.2 节已讨论过有限尺寸孔径条件下产生高斯截面光束的理论。用式 (19.4-22)，



分别进行多次模拟。采用 128 个信道的结果如图 19.11 所示。可见，该结果是完全可以接受的。

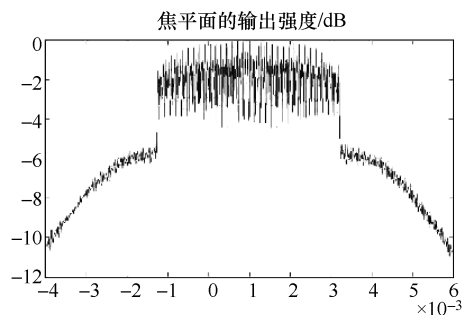


图 19.11 高斯光束的结果 ($M = 150$, $L = 128$, $\delta = 15$, $\Delta\lambda = 0.2\text{nm}$)

19.5.4 生成负相位的方法

到目前为止，实验结果都是运用自动零交叉方法得到的。图 19.12 给出了一个例子，在 16 信道下，在整个波阵面的相位中生成了具有负相位的信道 [Hu, Ersoy]。由图可见，与之前的情况一样，该结果同样有效。

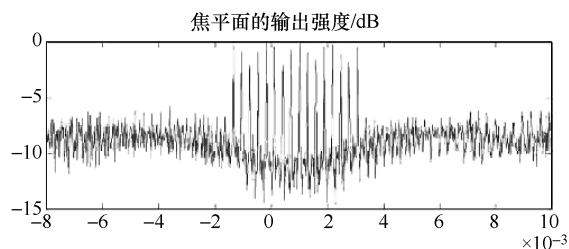


图 19.12 运用生成波阵面负相位的方法设计的 16 信道

19.5.5 误差容限

在器件加工制造过程中常产生可预见的相位误差。通过对每个阵列孔径施加随机相位误差可进行相位误差容限的研究。随机相位误差可由 $[-\text{ERR}, \text{ERR}]$ 范围内的均匀分布来近似得到，其中 ERR 是选定的最大误差。只要每个相近方向的相量点的最大误差满足

$$|\text{ERR}| \leq \pi/2 \quad (19.5-1)$$

这样对每个孔径来说就是有益的，因而可以得到满意的结果。仿真实验也证实了这一点。图 19.13 给出了一个示例，相应的 $\text{ERR} = 0.25\pi$ 。

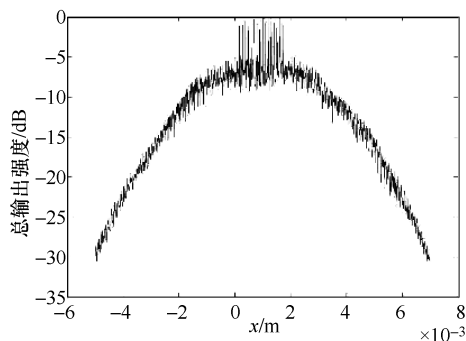


图 19.13 使用相位误差设计的 16 信道 ($\text{ERR} = 0.25\pi$)



19.5.6 3-D 仿真

通过类似的模拟方式可进行三维情形下的仿真 [Hu 和 Ersoy, 2002]。图 19.14 所示的是像平面 (x - y 平面, $z = z_0$) 上的聚焦和解复用的一个例子。所使用的四个波长分别为 1549.2nm、1549.6nm、1550nm 和 1550.4nm, 间隔 0.4nm (50 GHz)。在 $2\text{mm} \times 2\text{mm}$ 的平面上生成 50×50 的孔径阵列。设置 x 方向上衍射阶 δ_x 为 5, 而在 y 方向上, 设定 δ_y 为零。

图 19.14a 给出在像平面上的解复用, 图 19.14b 表示在同一平面上输出线 (x 方向) 上对应的插入损耗。

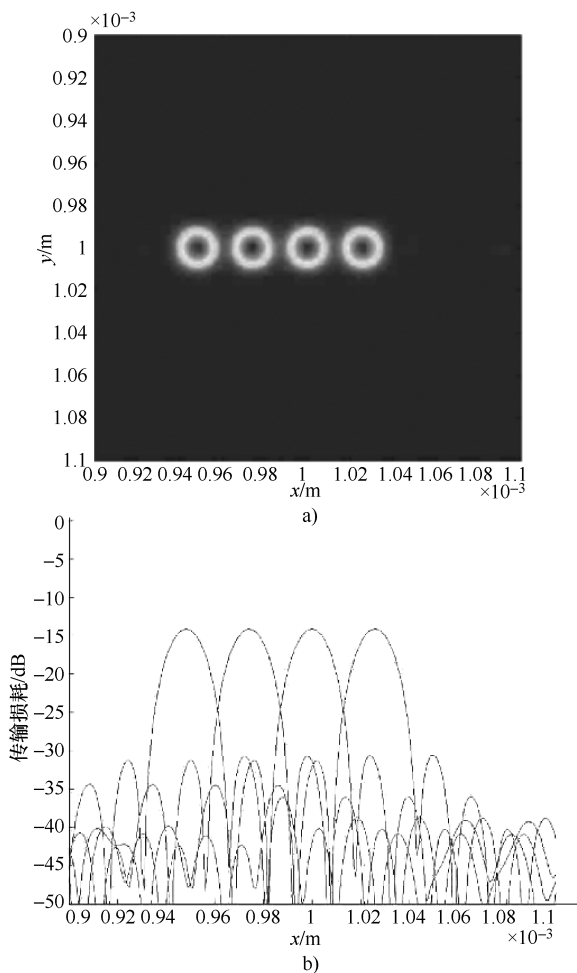


图 19.14 具有四种波长和精确相位的 3-D 设计示例

可见, 取合理大小的衍射级 ($\delta \sim 5$) 值都足以产生满意的结果。此结果意义重大, 因为它表明, 制造三维相控阵确实可以使用现有的技术来实现。三维相控阵的一个主要优点是, 相对于二维情形, 其孔径的数目可以大得多。

19.5.7 相位量化

在实际制造中, 通常对相位进行量化。所采用的技术决定了量化阶的数目。图 19.15 给



出了四个量化级的解复用结果，其余参数同图 19.14 中的一致。可见该结果是令人满意的。

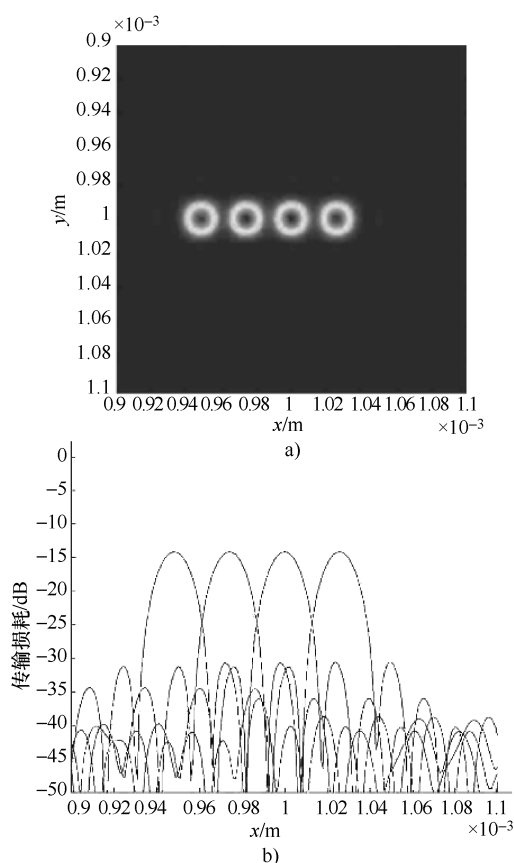


图 19.15 具有四种波长和四级相位量化的 3-D 设计实例

19.6 实现过程中的问题

比较实际实施过程中的不同方法，可以得到比较有趣的结果。例如，在平面波导的 PHASAR 器件中，波导的长度应根据下面的公式来选择：

(a) 只有平面参考波，自动零交叉

$$L_i = \alpha x_i \quad (19.6-1)$$

(b) 只有平面参考波，实现负相位

$$L_i = \alpha x_i - \theta_i/k \quad (19.6-2)$$

式中， θ_i 是期望相位； k 是波数。

(c) 平面和球面参考波，自动零交叉

$$L_i = \alpha(x_i + r_{ci}) \quad (19.6-3)$$

(d) 平面和球面参考波，实现负相位

$$L_i = \alpha(x_i + r_{ci}) - \theta_i/k \quad (19.6-4)$$

上述公式表明，实际中，负相位的物理产生方法似乎比采用控制波导长度的零交叉点法



更难以实现。然而，在零交叉点法中，必须仔细调整孔的位置。因为孔的初始位置是随机选择的，这不可预期地会产生更多附加的困难，这是由于在调整后其结果将为另一个随机数。

三维构型的缺点是，它可能难以实现更大的级数 δ 。其一大优点就是用来设计多孔径衍射光学元件，也可以用于 DWDM 的相控阵器件设计。由模拟结果可见，级数 δ 取 5，就可得到满意的解决方案。

这可通过多种方式来实现。一种可能的方法就是通过使用如图 19.16a 所示的装置，并与 16.2 节中讨论的虚拟全息术方法相结合。为了得到更大的 δ ，可以制造比通常尺寸大 5 倍的阵列，如图 19.16b 所示的倾斜排列，使图中所示的 ΔL 更大。然后，该阵列（现在称为实阵列）具有必要的相位调制，并且成像到如图 19.17 所示的虚拟阵列，该虚拟阵列可以通过虚拟全息的方法产生。如果设 M 为横向方向上的缩小倍率，则在 z 方向上的缩小倍率为 M^2 。于是，在虚拟阵列 z 方向上的倾斜可忽略不计。如上述讨论，为了使虚拟阵列运行在期望状态下，并将不同的波长聚焦在不同的位置，虚拟阵列要有必要的尺寸和相位调制能力。

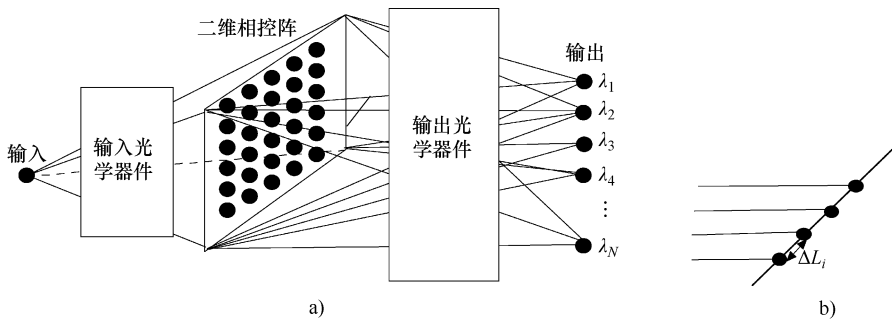


图 19.16 一个可能的 MISZC 3-D 实现的可视化结构模型

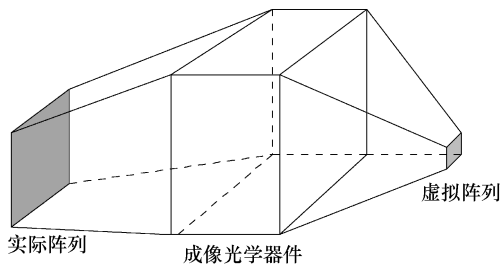


图 19.17 与图 19.16 连接，以实现期望相位调制和尺寸的虚拟全息装置

第 20 章 严格衍射理论的数值方法

20.1 本章引论

当衍射孔径的尺寸与入射波波长相差不大时, 标量衍射理论是成立的。通常这种条件下的模型也比较简单。生产越来越小的设备在目前已呈上升趋势, 这时利用折射光学元件变得不切实际, 同时随着越来越复杂的几何构型和材料属性的需求, 需要尺寸具有一个波长或小于一个波长量级的衍射光学元件。然而, 在这个尺寸范围内, 标量衍射理论可能不够精确, 通常很难得到解析的结果。复杂的模型也需要新的计算方法。

利用数值方法求解麦克斯韦方程是一个实用的途径。为此发展出了一些新的方法。其中一种是基于有限差分的数值方法, 另一种方法是基于傅里叶级数的傅里叶模式分析法, 第三种是利用有限元方法求解有边界条件的麦克斯韦方程。

本章共分为 8 节。20.2 节利用克(兰克, CranK)尼(克森, Nicholson)二氏法的有限差分形式介绍了本书在 12.4 节中讨论过的傍轴光束传播(BPM)公式。20.3 节讨论了利用帕德(Pâdé)近似的宽角度光束传播法。这种方法通常应用于角度变化范围较大的问题, 比如波导问题。20.4 节介绍有限差分方法, 作为下一节学习的基础。20.5 节介绍了时域有限差分法(Finite Difference Time Domain, FDTD), 这种方法目前应用比较广泛。20.6 节介绍了时域有限差分法的一些实验。20.7 节提出了另一种有竞争力的方法, 称为傅里叶模式法(Fourier Modal Method, FMM)。

20.2 基于有限差分法的 BPM

考虑式(12.2-4)所示的非均匀介质中的亥姆霍兹方程。为方便起见, 再次给出该方程如下:

$$(\nabla^2 + k^2(x, y, z))U(x, y, z) = 0 \quad (20.2-1)$$

其中

$$U(x, y, z) = u(x, y, z)e^{-j\omega t} \quad (20.2-2)$$

该位置相关的波数 $k(x, y, z)$ 满足

$$k(x, y, z) = n(x, y, z)k_0 \quad (20.2-3)$$

式中, k_0 是自由空间的波数; $n(x, y, z)$ 是非均匀介质的折射率。

如 12.3 节所述, $n(x, y, z)$ 可以表示为

$$n(x, y, z) = \bar{n} + \Delta n(x, y, z) \quad (20.2-4)$$

式中, \bar{n} 为平均折射率。

这样亥姆霍兹方程可化为

$$[\nabla^2 + \bar{n}k_0^2 + 2\bar{n}\Delta n k_0^2]U = 0 \quad (20.2-5)$$



其中 $(\Delta n)^2 k_0^2$ 项已忽略不计。进一步, 光场可假设按如下规律变化:

$$U'(x, y, z) = U(x, y, z) e^{ikz} \quad (20.2-6)$$

其中的 $U(x, y, z)$ 被假定为 z 的缓变函数, \bar{k} 与 $\bar{n}k_0$ 相等。将式 (20.2-6) 代入亥姆霍兹方程, 得到

$$\nabla^2 U + 2j\bar{k} \frac{\delta}{\delta z} U + (k^2 - \bar{k}^2) U = 0 \quad (20.2-7)$$

方程式 (20.2-7) 被用在光束传输法中的有限差分公式中。

如 12.3 节中讨论的那样, 当傍轴亥姆霍兹方程成立时, 式 (20.2-7) 可化为

$$\frac{\delta}{\delta z} U = \frac{j}{2k} \left[\frac{\delta^2 U}{\delta x^2} + \frac{\delta^2 U}{\delta y^2} + (k^2 - \bar{k}^2) U \right] \quad (20.2-8)$$

这是三维情况下 BPM 算法应用的基本傍轴方程。通过省略公式中与变量 y 相关的项可以得到二维傍轴方程。这种形式的傍轴近似有两个优点。首先, 由于与 z 相关的快速相位变化被提取出来, 使得近似不变的场可以沿着纵向网格进行数值表示, 网格的间隔可比波长大许多个数量级。这就使得 BPM 算法相比于传统有限差分法效率更高, 因为传统有限差分法的网格间距是十分之一波长量级。其次, 消除了 z 的二阶导数项, 使得原本是二阶边界值的问题转变成了一个一阶初值问题。二阶边界值问题通常需要迭代或特征值分析来求解, 而一阶初值问题则可以通过沿 z 方向的简单积分来求解。这种方法相对于亥姆霍兹方程的完全数值解法来说, 在很大程度上降低了计算的复杂度。

然而, 计算复杂度的降低也需要付出代价。缓变的包络近似假定了场的传播主要是沿 z 轴 (即傍轴方向), 而且它也限制了折射率变化的大小。在近似中去除二阶导数的同时也消除了后向的行波解。因此, 此方法对基于反射的器件不适用。然而, 这些问题可以通过重新推导近似来解决。为解决这些问题, 一些改进方法, 例如, 宽角度 BPM 和双向 BPM 算法, 将在后续的章节中讨论。

在 12.4 节中讨论了基于 FFT 求解方程式 (12.3-5) 的数值方法。另一种基于有限差分方法的算法被称为 FD-BPM 算法, 特别是它运用了克尼二氏法 (C-N 法) [Yevick, 1989]。有时, 采用有限差分法能得出更精确的结果 [Yevick, 1989], [Yevick, 1990]。它同样可以利用较大的纵向步长, 在不影响精度的情况下降低计算复杂度 [Scarmozzino]。

在有限差分方法中, 场可由一些横向平面内的离散网格点来表示, 这些平面垂直于 z 轴方向, 且沿着 z 轴方向等间隔分布。一旦输入场是已知的, 即在 $z=0$ 的场已知, 即可求得下一个横向平面的场。采用这种方法, 沿 z 轴方向传播的波可逐步地计算得出。我们将给出二维情形下的算法, 并可直接扩展到三维情形。

令 u_i^n 表示场在某格点的值, 其中 i 表示该点在横向网格的索引, n 表示纵向的索引。同时假定横向平面中的格点等间隔分布, 间隔为 Δx , 各个横面向亦等间隔分布, 间隔为 Δz 。在克尼二氏法中, 式 (12.3-5) 表示的是一个虚构的中间平面, 该平面介于已知的平面 n 和未知的平面 $n+1$ 之间, 如图 20.1 所示。

在克尼二氏法中, 中间平面的一阶和二阶导数可由以下公式表示:

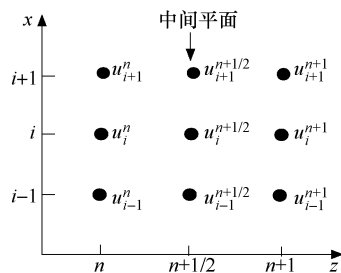


图 20.1 克尼二氏法的采样点



$$\left. \frac{\delta u_i}{\delta z} \right|_{z=\Delta z(n+1/2)} = \frac{u_i^{n+1} - u_i^n}{\Delta z} \quad (20.2-9)$$

$$\left. \frac{\delta^2 u_i}{\delta z^2} \right|_{z=\Delta z(n+1/2)} = \frac{\delta^2 u_i^{n+1} - \delta^2 u_i^n}{2\Delta z^2} \quad (20.2-10)$$

式中, $z_{n+1/2} = z_n + \Delta z/2$; δ^2 表示二阶差分算子, 其由下式给出。

$$\delta^2 u_i = u_{i+1} + u_{i-1} - 2u_i \quad (20.2-11)$$

利用上述近似, 式 (20.2-8) 可转化为

$$\frac{u_i^{n+1} - u_i^n}{\Delta z} = \frac{j}{2k} \left(\frac{\delta^2}{\Delta x^2} + (k(x_i, z_{n+1/2}) - \bar{k}^2) \right) \frac{u_i^{n+1} + u_i^n}{2} \quad (20.2-12)$$

式 (20.2-12) 可写成一个三对角矩阵方程 [Scarmozzino]

$$a_i u_{i-1}^{n+1} + b_i u_i^{n+1} + c_i u_{i+1}^{n+1} = d_i \quad (20.2-13)$$

求解该公式, 可得到沿 z 轴方向第 $(n+1)$ 层的未知的 u_i^{n+1} 。由于其具有三对角性质, 该方程可通过 $O(N)$ 步运算迅速求解, 其中 N 为沿 x 轴方向网格结点的个数。

在边界点 $i=0, N$ 中, 必须使用合适的边界条件, 通常首选透明边界条件 [Hadley, 1990], [Hadley, 1992]。

20.3 宽角度 BPM

上述傍轴近似允许 BPM 算法在傍轴波和折射率微小变化的情况下使用。为了扩展该方法, 有必要把之前忽略过的二阶导数项重新引入, 并代入式 (20.2-8), 则从式 (20.2-7) 开始, 其可以表示为

$$\frac{\partial^2 U}{\partial z^2} + 2j\bar{k} \frac{\partial U}{\partial z} + \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + (k^2 - \bar{k}^2) U = 0 \quad (20.3-1)$$

令 D 表示 $\frac{\partial}{\partial z}$, D^2 表示 $\frac{\partial^2}{\partial z^2}$ 。式 (20.3-1) 可写为

$$D^2 U + 2j\bar{k} D U + \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + (k^2 - \bar{k}^2) \right) U = 0 \quad (20.3-2)$$

可将上述公式看作是 D 的一个代数方程。将 D 作为未知量求解, 可得

$$D = j\bar{k} [\sqrt{1+P} - 1] \quad (20.3-3)$$

其中

$$P = \frac{1}{k^2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + (k^2 - \bar{k}^2) \right) \quad (20.3-4)$$

P 和 D 是作用于 U 的算符。于是, 式 (20.3-2) 可化为 [Hadley]

$$\frac{\partial U}{\partial z} = j\bar{k} (\sqrt{1+P} - 1) U \quad (20.3-5)$$

该式只给出了光沿前向传播的解。光沿反方向传播的问题可以通过改变二次方根项的符号推导得到。这两种解法都是正确的。

为了求解式 (20.3-5) 中的 U , 有必要将 $\sqrt{1+P}$ 展开。用帕德法展开该式, 可在得到最高精度的同时得到最少的展开项 [Hadley]。阶数为 (m, n) 的帕德近似是具有如下形式的



有理函数：

$$R(x) = \frac{\sum_{k=0}^M a_k x^k}{1 + \sum_{k=1}^N b_k x^k}$$

(20.3-6)

表 20.1 给出了项 $\bar{k}(\sqrt{1+P}-1)$ 的低阶帕德近似。

表 20.1 项 $\bar{k}(\sqrt{1+P}-1)$ 的低阶帕德近似

阶 数	$\bar{k}(\sqrt{1+P}-1)$
(1, 0)	$\frac{P}{2\bar{k}}$
(1, 1)	$\frac{\frac{P}{2\bar{k}}}{1 + \frac{P}{4\bar{k}^2}}$
(2, 2)	$\frac{\frac{P}{2\bar{k}} + \frac{P^2}{4\bar{k}^3}}{1 + \frac{3P}{4\bar{k}^2} + \frac{P^2}{16\bar{k}^4}}$
(3, 3)	$\frac{\frac{P}{2\bar{k}} + \frac{P^2}{2\bar{k}^3} + \frac{3P^2}{32\bar{k}^5}}{1 + \frac{5P}{4\bar{k}^2} + \frac{3P^2}{8\bar{k}^4} + \frac{P^3}{64\bar{k}^6}}$

通过帕德近似可将式 (20.3-5) 变为

$$\frac{\partial U}{\partial z} = j\bar{k} \frac{N_m(P)}{D_n(P)} U$$

(20.3-7)

式中， $N_m(P)$ 和 $D_n(P)$ 是 P 的多项式，对应于 (m,n) 阶的帕德近似。

随着帕德近似式的阶数增加，解的精准度也随之提高。对式 (20.3-5) 中的二次方根，(1,1) 阶帕德近似对高达 30° 的角也足够精确，而 (3,3) 阶帕德近似同其 15 阶泰勒展开具有相同的近似精度。

宽角度光束传播法常用于一些必须涉及宽角度的应用中，如在第 19 章中讨论的 AWG 的设计实验中。举个例子，为了适应长度变化，图 19.2 展示了弯曲宽角度波导的布局。图 20.2 给出了一个设计为 200 信道的 AWG 输出强度分布，该结果采用宽角度光束传播法计算得到，计算中使用了一款名为 BeamPROP 的 RSoft 软件。

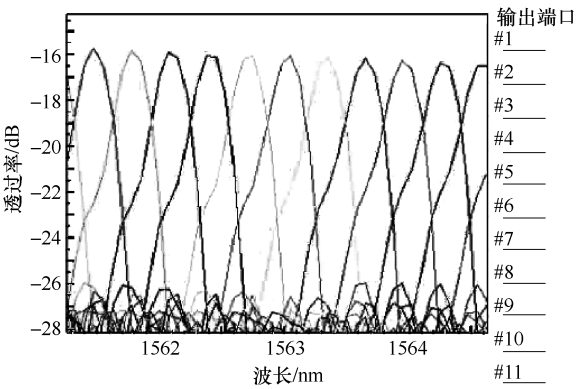


图 20.2 一个设计为 200 信道的相控阵输出强度特写图，它由宽角度光束传播法计算得到



20.4 有限差分

本节将进行时域有限差分方法的讨论。这种方法是基于第一和第二导数的有限差分近似。这些将在下面讨论。

对 $u(x, t_n)$, 固定时间 t_n , 其在 $x_i + \Delta x$ 点关于 x_i 的泰勒级数展开式由下式给出 [Kuhl, Ersoy]:

$$u(x_i + \Delta x, t_n) \big|_{t_n} = u \big|_{x_i, t_n} + \Delta x \frac{\partial u}{\partial x} \bigg|_{x_i, t_n} + \frac{\Delta x^2}{2} \frac{\partial^2 u}{\partial x^2} \bigg|_{x_i, t_n} + \frac{\Delta x^3}{6} \frac{\partial^3 u}{\partial x^3} \bigg|_{x_i, t_n} + \frac{\Delta x^4}{24} \frac{\partial^4 u}{\partial x^4} \bigg|_{\xi_1, t_n} \quad (20.4-1)$$

最后一项是误差项, ξ_1 为区间 $(x_i, x_i + \Delta x)$ 内的点。对于固定时间 t_n , 其在点 $x_i - \Delta x$ 的展开式同理可得

$$u(x_i - \Delta x, t_n) \big|_{t_n} = u \big|_{x_i, t_n} - \Delta x \frac{\partial u}{\partial x} \bigg|_{x_i, t_n} + \frac{\Delta x^2}{2} \frac{\partial^2 u}{\partial x^2} \bigg|_{x_i, t_n} - \frac{\Delta x^3}{6} \frac{\partial^3 u}{\partial x^3} \bigg|_{x_i, t_n} + \frac{\Delta x^4}{24} \frac{\partial^4 u}{\partial x^4} \bigg|_{\xi_2, t_n} \quad (20.4-2)$$

式中, ξ_2 是区间 $(x_i - \Delta x, x_i)$ 内的点。

将上述两个展开项相加, 得到

$$u(x_i + \Delta x) + u(x_i - \Delta x) = 2u \big|_{x_i, t_n} + \Delta x^2 \frac{\partial^2 u}{\partial x^2} \bigg|_{x_i, t_n} + \frac{\Delta x^4}{12} \frac{\partial^4 u}{\partial x^4} \bigg|_{\xi_3, t_n} \quad (20.4-3)$$

式中, ξ_3 是区间 $(x_i - \Delta x, x_i + \Delta x)$ 内的点。

重新整理式 (20.4-3) 可得

$$\frac{\partial^2 u}{\partial x^2} \bigg|_{x_i, t_n} = \left[\frac{u(x_i + \Delta x) - 2u(x_i) + u(x_i - \Delta x)}{(\Delta x)^2} \right]_{t_n} + O[(\Delta x)^2] \quad (20.4-4)$$

式 (20.4-4) 可写为

$$\frac{\partial^2 u}{\partial x^2} \bigg|_{x_i, t_n} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} + O[(\Delta x)^2] \quad (20.4-5)$$

u 相对于时间的二阶偏导数同样由下式给出:

$$\frac{\partial^2 u}{\partial t^2} \bigg|_{x_i, t_n} = \frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{(\Delta t)^2} + O[(\Delta t)^2] \quad (20.4-6)$$

20.5 时域有限差分法

时域有限差分法是一种逐渐得到广泛应用的有效的数值方法。在该方法中, 麦克斯韦方程被表示成中心差分方程的形式。由此产生的方程可由蛙跳式方法求解。换言之, 在给定的瞬时求解电场, 在下一个瞬时求解磁场, 并重复该过程多次。

为方便起见, 将第3章的式 (3.3-3) 和式 (3.3-4) 在无源场 (无电场或磁场源) 中改写为如下形式:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} \quad (20.5-1)$$



$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t} \quad (20.5-2)$$

式 (20.5-1) 说明, 电场 \mathbf{E} 的时间导数与磁场 \mathbf{H} 的旋度成正比。这就意味着, 空间中某点新一时刻的电场值 \mathbf{E} , 可以从前一刻的电场值 \mathbf{E} 以及所在空间点两侧的前一时刻磁场值 \mathbf{H} 的差分来得到。

以下是式 (3.3-14) 和式 (3.3-15) 写成分量的形式:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - \rho' H_x \right) \quad (20.5-3)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - \rho' H_y \right) \quad (20.5-4)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - \rho' H_z \right) \quad (20.5-5)$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \sigma E_x \right) \quad (20.5-6)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - \sigma E_y \right) \quad (20.5-7)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - \sigma E_z \right) \quad (20.5-8)$$

接下来, 上述方程采用将要介绍的 Yee 算法, 写成有限差分形式来求解。

Yee 算法

Yee 算法是通过使用一组有限差分方程来求解麦克斯韦旋度方程的 [Yee]。使用有限差分, 在三维场中的每个电场分量周围都环绕着四个磁场分量, 同样地, 每一个磁场分量的周围都环绕着四个电场分量, 如图 20.3 所示。电场分量和磁场分量以时间为中心, 以蛙跳方式排列。这就意味着特定时间的所有电场分量都可以由预先存储的磁场分量计算得出并存储。然后, 所有的磁场分量都可由先前计算得到的电场分量来求出 [Kuhl, Ersoy]。

对于 Yee 元胞, 离散化过程如下:

- $[i, j, k] = (i\Delta x, j\Delta y, k\Delta z)$, 其中 Δx 、 Δy 和 Δz 分别是在 x 、 y 和 z 方向上的空间增量, i, j, k 为整数;
- $t_n = n\Delta t$;
- $u(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = u_{i,j,k}^n$

对 x 轴方向关于 u 的一阶空间偏微分, 在时间 $t_n = n\Delta t$ 的 Yee 中心有限差分由下式给出:

$$\frac{\partial u}{\partial x}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i+1/2,j,k}^n - u_{i-1/2,j,k}^n}{\Delta x} + O[(\Delta x)^2] \quad (20.5-9)$$

可见, 为求点 $(i\Delta x, j\Delta y, k\Delta z)$ 的场, 需要用到距离该点 $\Delta x/2$ 远处的点的未知数据, 这与克尼二氏法类似。对于空间某一点, u 对时间的一阶偏微分由下式给出:

$$\frac{\partial u}{\partial t}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i,j,k}^{n+1/2} - u_{i,j,k}^{n-1/2}}{\Delta t} + O[(\Delta t)^2] \quad (20.5-10)$$

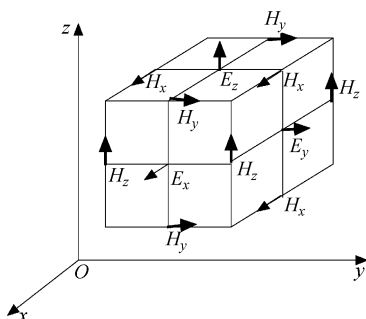


图 20.3 Yee 元胞的电场及磁场分量分布 [Yee]



将时间为 n 时的点 (i, j, k) 的时间、空间导数代入式 (20.5-6), 可得

$$\frac{E_x \big|_{i,j,k}^{n+1/2} + E_x \big|_{i,j,k}^{n-1/2}}{\Delta t} = \frac{1}{\varepsilon_{i,j,k}} \left(\frac{H_z \big|_{i,j+1/2,k}^n - H_z \big|_{i,j-1/2,k}^n}{\Delta y} - \frac{H_y \big|_{i,j,k+1/2}^n - H_y \big|_{i,j,k-1/2}^n}{\Delta z} - \sigma_{i,j,k} E_x \big|_{i,j,k}^n \right) \quad (20.5-11)$$

该式右边的所有项都是在时间为 n 时计算出的。所有在时间点 n 时所需的磁场分量都是已知的。只有 E_x 在时间点 $n-1/2$ 时被存储了。对于无损耗区域, 该项为 0。如果 $\sigma_{i,j,k}$ 不为零, 则 E_x 可以由下式估计 [Yee]:

$$E_x \big|_{i,j,k}^n = \frac{E_x \big|_{i,j,k}^{n+1/2} + E_x \big|_{i,j,k}^{n-1/2}}{2} \quad (20.5-12)$$

式 (20.5-12) 表示的是一个在时间 $n-1/2$ 的已知量 E_x 和一个在时间 $n+1/2$ 的未知量的平均值。将式 (20.5-12) 代入式 (20.5-11), 可得

$$E_x \big|_{i,j,k}^{n+1/2} + E_x \big|_{i,j,k}^{n-1/2} = \frac{\Delta t}{\varepsilon_{i,j,k}} \left(\frac{H_z \big|_{i,j+1/2,k}^n - H_z \big|_{i,j-1/2,k}^n}{\Delta y} - \frac{H_y \big|_{i,j,k+1/2}^n - H_y \big|_{i,j,k-1/2}^n}{\Delta z} - \sigma_{i,j,k} \frac{E_x \big|_{i,j,k}^{n+1/2} + E_x \big|_{i,j,k}^{n-1/2}}{2} \right) \quad (20.5-13)$$

合并同类项, 并求解 $E_x \big|_{i,j,k}^{n+1/2}$ 的值, 可得

$$E_x \big|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) E_x \big|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) \left(\frac{H_z \big|_{i,j+1/2,k}^n - H_z \big|_{i,j-1/2,k}^n}{\Delta y} - \frac{H_y \big|_{i,j,k+1/2}^n - H_y \big|_{i,j,k-1/2}^n}{\Delta z} \right) \quad (20.5-14)$$

所有其他未知的场分量可由下式进行类似的推导:

$$E_y \big|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) E_y \big|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) \left(\frac{H_x \big|_{i,j,k+1/2}^n - H_x \big|_{i,j,k-1/2}^n}{\Delta z} - \frac{H_z \big|_{i+1/2,j,k}^n - H_z \big|_{i-1/2,j,k}^n}{\Delta x} \right) \quad (20.5-15)$$

$$E_z \big|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) E_z \big|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k}}} \right) \left(\frac{H_y \big|_{i+1/2,j,k}^n - H_y \big|_{i-1/2,j,k}^n}{\Delta x} - \frac{H_x \big|_{i,j+1/2,k}^n - H_x \big|_{i,j-1/2,k}^n}{\Delta y} \right) \quad (20.5-16)$$

$$H_x \big|_{i,j,k}^{n+1} = \left(\frac{1 - \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) H_x \big|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) \left(\frac{E_y \big|_{i,j,k+1/2}^{n+1/2} - E_y \big|_{i,j,k-1/2}^{n+1/2}}{\Delta z} - \frac{E_z \big|_{i,j+1/2,k}^{n+1/2} - E_z \big|_{i,j-1/2,k}^{n+1/2}}{\Delta y} \right) \quad (20.5-17)$$

$$H_y \big|_{i,j,k}^{n+1} = \left(\frac{1 - \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) H_y \big|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) \left(\frac{E_z \big|_{i+1/2,j,k}^{n+1/2} - E_z \big|_{i-1/2,j,k}^{n+1/2}}{\Delta x} - \frac{E_x \big|_{i,j,k+1/2}^{n+1/2} - E_x \big|_{i,j,k-1/2}^{n+1/2}}{\Delta z} \right) \quad (20.5-18)$$



$$H_z \Big|_{i,j,k}^{n+1} = \left(\frac{1 - \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) H_z \Big|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\rho'_{i,j,k} \Delta t}{2\mu_{i,j,k}}} \right) \left(\frac{E_x \Big|_{i,j+1/2,k}^{n+1/2} - E_x \Big|_{i,j-1/2,k}^{n+1/2}}{\Delta y} - \frac{E_y \Big|_{i+1/2,j,k}^{n+1/2} - E_y \Big|_{i-1/2,j,k}^{n+1/2}}{\Delta x} \right) \quad (20.5-19)$$

20.6 计算机实验

我们在计算机实验中使用 Remcom 出品的电磁场仿真软件 XFDTD [Kuhl, Ersoy]。在这个软件环境下,所关注的区域是一个立方体网格,为了使其能够模拟特定的模型,每个网格的边可以赋予不同的材料特性。对于每个元胞,材料可以是理想导体或自由空间,也可以以其他方式来定义。空间采样的间隔在亚波长量级,通常在一个波长的 $1/30 \sim 1/10$ 的范围内。关注区域是由一个平面波或多个电压源激励的。激励信号可以是脉冲信号或正弦信号。仿真的持续时间是设定的所需时间步数决定的。

当模拟的区域延伸到无穷远时,网格的边界采用的是吸收边界条件(ABC)。这使得所有从该区域传出的波几乎都没有反射。此时关注区域也可以通过一个理想电导体包围。

一旦按照设定的时间步骤数计算该场,近区瞬态和稳态场就可以彩色强度图像进行可视化,或者可以绘出某点场分量随时间的变化曲线。当需要稳态输出时,观察在特定点的场随时间的变化有助于显示是否达到稳定。

在计算机实验的执行中,使用了 $\lambda/20$ 大小的元胞尺寸 [Kuhl, Ersoy]。激励源是沿 z 方向传播的 y 偏振正弦平面波。全部衍射结构都是由理想的导体构成的。衍射结构的所有边缘均平行于 x 轴,以避免消掉 y 方向偏振的电场。

一维菲涅耳波带片(FZP)同 15.11 节所讨论的没有 x 变量的 FZP 是相同的。其模 m 定义为一些不透明的偶数或奇数区域的个数。

图 20.4 给出了三个不透明区域($m=3$)的例子。

利用 XFDTD,对厚度为 0.1λ 、焦距为 3λ 的一维 FZP 进行了模拟,并对前三个模的结果进行了分析。图 20.5 给出了穿过 FZP 中心沿 z 轴方向的强度值随距离 FZP 的距离变化。该曲线图表明,在距离波片 3λ 位置处达到强度峰值,并且随着模数的增加,峰值在该处会变得更宽,宽度更窄。

在焦线上沿 y 轴的强度曲线如图 20.6 所示。该图显示,光点尺寸随着模数的上升而减小,同时旁瓣强度也由于模数增高而降低。

上述实验表明,时域有限差分法在生成所需的几何模型和指定材料参数方面提供了相当大的自由度。特别是当受限于复杂的几何模型或衍射的小孔尺寸比波长小得多等原因,标量衍射理论不足以满足精度要求时,时域有限差分法的优势就尤为明显了。

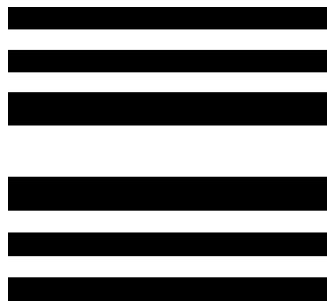


图 20.4 $m=3$ 的一维 FZP

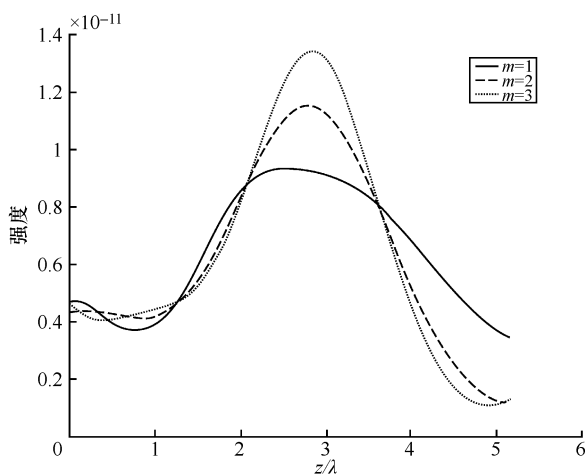


图 20.5 应用时域有限差分法得到的沿 z 轴方向的强度分布

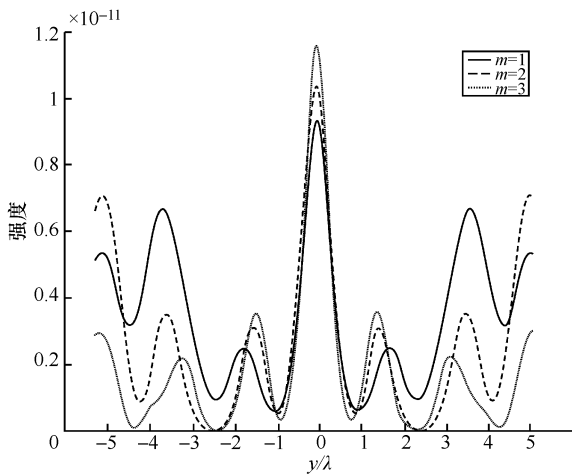


图 20.6 应用时域有限差分法得到的在焦线上沿 y 轴的强度曲线

20.7 傅里叶模式法

傅里叶模式法是几种分析浮雕型衍射光学元件的方法中最通用、最严谨的方法。这方面的最初工作被称为严格耦合波分析（Rigorous Coupled-Wave Analysis, RCWA）[Moharam, 1982]。它最初主要应用在衍射光栅分析中，后来也被推广到非周期结构的分析中 [Lalanne-Silberstein]。在本节的简要介绍中，将假定一种光栅结构进行讨论。

傅里叶模式法是一种频域方法，它涉及将光栅模式作为特征矢量的计算，以及介电常数的傅里叶展开、光栅区域内部电磁场的傅里叶展开。光栅内部场的傅里叶展开可生成一种差分方程系统。在求出该系统的特征值和特征向量之后，我们发现在光栅表面的边界条件适合于计算衍射效率。[Lalagne and Morris, 1996]。



图 20.7 给出了二维情形下用来解释该方法的几何模型。

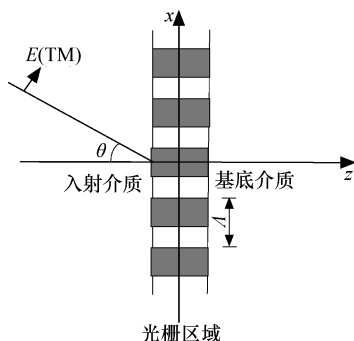


图 20.7 傅里叶模式法的几何模型 [Lalanne and Morris, 1996]

沿 x 轴方向的周期性光栅，其周期为 Λ ，介电常数为函数 $\varepsilon(x)$ 。考虑磁场的偏振沿 y 轴方向且没有变化，这样的磁场模称为横磁场模（TM 模）。入射平面波与 z 轴方向成 θ 夹角。同时，定义如下的傅里叶级数：

$$\frac{\varepsilon(x)}{\varepsilon_0} = \sum_{k=-\infty}^{\infty} \varepsilon_k e^{jKkx} \quad (20.7-1)$$

$$\frac{\varepsilon_0}{\varepsilon(x)} = \sum_{k=-\infty}^{\infty} a_k e^{jKkx} \quad (20.7-2)$$

式中， $K = 2\pi/\Lambda$ 。

各场分量可写为 [Lalanne and Morris, 1996]

$$E_x = \sum S_m(z) e^{j(Km+\beta)x} \quad (20.7-3)$$

$$E_z = \sum f_m(z) e^{j(Km+\beta)x} \quad (20.7-4)$$

$$H_y = \sum U_m(z) e^{j(Km+\beta)x} \quad (20.7-5)$$

式中， $\beta = k \sin \theta = \frac{2\pi}{\Lambda} \theta$ 。

在这种情况下，麦克斯韦旋度方程由下式给出：

$$-\frac{\delta E_z}{\delta x} + \frac{\delta E_x}{\delta x} = -j\omega\mu_0 H_y \quad (20.7-6)$$

$$\frac{\delta H_y}{\delta z} = -j\omega\varepsilon E_x \quad (20.7-7)$$

$$\frac{1}{\varepsilon} \frac{\delta H_y}{\delta x} = j\omega E_z \quad (20.7-8)$$

对于变量 z 的一阶和二阶偏导数可用素数和双素数的数量来表示。将式 (20.7-3) ~ 式 (20.7-5) 对应代入式 (20.7-6) ~ 式 (20.7-8)，结果如下：

$$-j(Km+\beta)f_m + S'_m = -jk_0 U_m \quad (20.7-9)$$

$$U'_m = -jk_0 \sum_p \varepsilon_{m-p} S_p \quad (20.7-10)$$

$$f_m = \frac{1}{k_0} \sum_p (pK + \beta) a_{m-p} U_p \quad (20.7-11)$$



将式 (20.7-11) 代入式 (20.7-9) 可得

$$S'_m = -jk_0 U_m + \frac{j}{k_0} (Km + \beta) \sum_p (pK + \beta) a_{m-p} U_p \quad (20.7-12)$$

实际上, 上述所有求和式都由 $|p| \leq M$ 来截断。式 (20.7-10) 和式 (20.7-12) 构成了大小为 $2(2M+1)$ 的特征值问题。然而, 利用式 (20.7-10) 求解 U_m 的二阶偏微分将更容易计算 [Caylord, Moharam, 1985], [Peng]:

$$U''_m = -k_0^2 \left\{ \sum_p \varepsilon_{m-p} \left[U_p - \frac{1}{k_0} (pK + \beta) \right] \right\} \left\{ \sum_r \frac{1}{k_0} (rK + \beta) a_{p-r} U_r \right\} \quad (20.7-13)$$

方程式 (20.7-13) 可以写成矩阵形式, 具体如下式:

$$\frac{1}{k_0^2} \mathbf{U}'' = \mathbf{E} (\mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{I}) \mathbf{U} \quad (20.7-14)$$

式中, 黑体大写字母代表矩阵; \mathbf{I} 为单位矩阵; \mathbf{E} 为介电谐波系数矩阵; \mathbf{K}_x 是一个对角矩阵, 其第 i 个对角元素等于 $(iK + \beta)/k_0$ 。

方程式 (20.7-14) 的一个更有效的计算形式如下 [Lalanne, 1996]:

$$\frac{1}{k_0^2} \mathbf{U}'' = \mathbf{E} (\mathbf{E} \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{I}) \mathbf{U} \quad (20.7-15)$$

作为示例, 图 20.8 给出了 TM 偏振光入射到金属光栅时, 在保持傅里叶系数不变的情况下, 透射 0 级衍射光的衍射效率。[Lalanne and Morris, 1996]。实线和圆圈曲线分别由方程式 (20.7-14) 和式 (20.7-15) 得到。

以上的介绍是导引性的。目前已经有针对许多不同应用的傅里叶模式法的改进, 而且该方法还将得到进一步的改进。

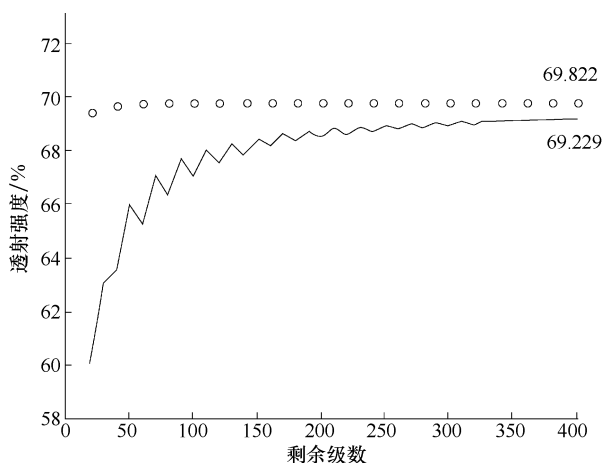


图 20.8 TM 偏振光入射到金属光栅时, 透射 0 级衍射光的衍射效率 [Lalanne and Morris, 1996]

附录

附录 A 脉冲函数

脉冲 (δ 或狄拉克 δ) 函数 $\delta(t)$ 可以看作具有单位面积的窄脉冲的理想化描述。考虑如图 A.1 所示的有限脉冲。它被定义为

$$x(t) = \begin{cases} \frac{1}{a} & -\frac{a}{2} < t < \frac{a}{2} \\ 0 & \text{其他} \end{cases} \quad (\text{A. 1-1})$$

对于所有的 a 值, 脉冲下的面积均保持为 1。脉冲函数可定义为

$$\delta(t) = \lim_{a \rightarrow 0} x(t) \quad (\text{A. 1-2})$$

$\delta(t)$ 满足

$$\begin{aligned} \int_{-\infty}^{+\infty} \delta(t) dt &= 1 \\ \delta(t) &= 0 \quad t \neq 0 \end{aligned} \quad (\text{A. 1-3})$$

在数学上, 我们认为 $\delta(t)$ 不是一个常规的函数, 而是一个广义的函数或分布, 将在后述进行讨论。 $\delta(t)$ 具有如下采样性质:

$$\int_{t_1}^{t_2} \delta(t - \tau) h(\tau) dt = \int_{t_1}^{t_2} \delta(\tau - t) h(\tau) dt = h(t) \quad (\text{A. 1-4})$$

式中, $h(t)$ 是一个已知函数, 它在 $\tau = t$ 时是连续的, 且在 $t_1 < t < t_2$ 上也是。

显然, 如果 $h(t)$ 是线性时不变系统的脉冲响应, 对系统利用 $\delta(t)$ 作为输入即可得到 $h(t)$ 。

脉冲函数也可以写为单位阶跃函数的导数

$$\delta(t) = \frac{d}{dt} u(t) \quad (\text{A. 1-5})$$

脉冲函数可以通过多个积分值为 1 的函数取极限运算来获得。下面是一些例子:

$$\delta(t) = \begin{cases} \lim_{a \rightarrow \infty} [ae^{-at} u(t)] \\ \lim_{a \rightarrow \infty} \frac{1}{2} e^{-a|t|} \\ \lim_{a \rightarrow 0^+} \frac{1}{\sqrt{2\pi a}} e^{-t^2/2a^2} \\ \lim_{a \rightarrow \infty} \frac{\sin(at)}{\pi t} \end{cases} \quad (\text{A. 1-6})$$

显然, 上述函数在 $t \neq 0$ 时的极限趋近于 0, 在 $t = 0$ 时的极限是 ∞ 。在最后一个例子中, 在 $t \neq 0$ 处的极限不趋近于 0, 但是该函数振荡得非常快, 式 (A. 1-6) 仍然成立。

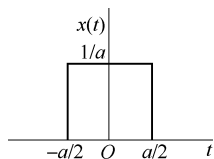


图 A.1 单位面积的有限脉冲



另一种构造脉冲函数的方式是通过图 A. 2 中所示的三角形脉冲。它的面积为 1。当 a 趋近于 0 时，其面积仍然为常数 1，同时三角形的底边长趋近于 0，它的高趋近于正无穷。即

$$\delta(t) = \lim_{a \rightarrow \infty} \frac{1}{a} \text{tri}\left(\frac{2t}{a}, -1, 1\right) \quad (\text{A. 1-7})$$

上述脉冲函数的导数可以被定义为

$$\begin{aligned} \delta'(t) &= \frac{d}{dt} \lim_{a \rightarrow \infty} \frac{1}{a} \text{tri}\left(\frac{2t}{a}, -1, 1\right) \\ &= \lim_{a \rightarrow \infty} \frac{1}{a} \frac{d}{dt} \text{tri}\left(\frac{2t}{a}, -1, 1\right) \end{aligned} \quad (\text{A. 1-8})$$

三角形脉冲的导数如图 A. 2b 所示。 $\delta'(t)$ 被称为二阶阶跃函数。

脉冲函数的导数可以由下面的积分来定义：

$$\int_{t_1}^{t_2} f(t) \delta^k(t - t_0) dt = (-1)^k f^{(k)}(t_0) \quad (\text{A. 1-9})$$

式中， $t_1 < t_0 < t_2$ ； $\delta^k(t)$ 和 $f^{(k)}(t)$ 是 $\delta(t)$ 和 $f(t)$ 的第 k 阶导数。

脉冲函数的一些常用性质如下：

性质 1 时域缩放

$$\delta(at - t_0) = \frac{1}{|a|} \delta\left(t - \frac{t_0}{a}\right) \quad (\text{A. 1-10})$$

性质 2 与函数相乘

$$f(t) \delta(t - t_0) = f(t_0) \delta(t - t_0), f(t) \text{ 在 } t = t_0 \text{ 点连续} \quad (\text{A. 1-11})$$

性质 3 $\delta[f(t)]$

$$\int_{-\infty}^{+\infty} \delta[f(t)] dt = \sum_k \frac{1}{|f'(t_k)|}, f'(t_k) \neq 0 \quad (\text{A. 1-12})$$

式中， $f'(t) = \frac{df(t)}{dt}$ ， t_k 为 $f(t_k) = 0$ 时的值。

性质 4 平移脉冲函数的求和

$$|T| \sum_{k=-\infty}^{\infty} \delta\left(t - \frac{k}{T}\right) = \sum_{k=-\infty}^{+\infty} e^{j2\pi kt/T} = 1 + 2 \sum_{k=1}^{\infty} \cos(2\pi kt/T) \quad (\text{A. 1-13})$$

例 A. 1 计算积分

$$\int_{-\infty}^{\infty} \delta\left(\frac{t}{3} - 1\right) \cos(10t) dt$$

解：

$$\begin{aligned} \delta\left(\frac{t}{3} - 1\right) &= \delta\left(\frac{1}{3}(t - 3)\right) \\ &= 3\delta(t - 3) \end{aligned}$$

即

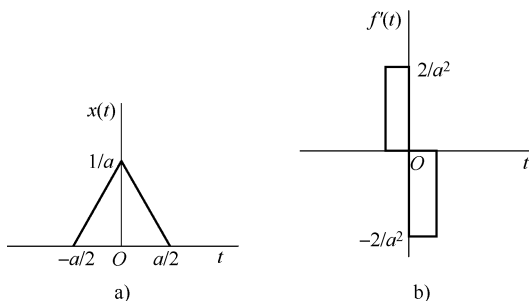


图 A. 2 三角形脉冲及其导数

a) 三角形脉冲 b) 导数



$$\begin{aligned}\int_{-\infty}^{+\infty} \delta\left(\frac{t}{3} - 1\right) \cos(10t) dt &= 3 \int_{-\infty}^{+\infty} \delta(t - 3) \cos(10t) dt \\ &= 3 \cos(30) \\ &= 3 \sqrt{3}/2\end{aligned}$$

例 A.2 用脉冲函数表示：

$$\lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi a}} e^{-t^2/32a^2}$$

解：

$$\frac{1}{\sqrt{2\pi a}} e^{-t^2/32a^2} = \frac{1}{\sqrt{2\pi a}} e^{-(t/4)^2/2a^2}$$

因为

$$\lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi a}} e^{-t^2/2a^2} = \delta(t)$$

所以

$$\lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi a}} e^{-t^2/32a^2} = \delta\left(\frac{t}{4}\right) = 4\delta(t)$$

例 A.3 用脉冲函数表示

$$\lim_{a \rightarrow \infty} \frac{1}{\pi} \int_0^a \cos(tx) dx$$

解：已知

$$\frac{1}{\pi} \int_0^a \cos(tx) dx = \frac{\sin(at)}{\pi t}, \quad a > 0$$

因为

$$\lim_{a \rightarrow \infty} \frac{\sin(at)}{\pi t} = \delta(t)$$

所以

$$\delta(t) = \lim_{a \rightarrow \infty} \frac{1}{\pi} \int_0^a \cos(tx) dx \quad (\text{A. 1-14})$$

结果通常写为

$$\delta(t) = \lim_{a \rightarrow \infty} \frac{1}{2\pi} \int_{-a}^a e^{jtx} dx$$

附录 B 线性矢量空间

B.1 本章引论

在非常多的应用中，通常利用基于一组基函数的傅里叶相关的级数和离散变换来对信号进行表示。更一般地讲，这是线性矢量空间的内容，在线性矢量空间里这些基函数被称为基



矢量。利用离散傅里叶相关变换对信号进行表示，可被看作是在有限和无限维希尔伯特空间中用基矢量对信号进行表示。

矢量空间最常见的例子就是我们熟知的三维空间 \mathbf{R}^3 。在一般情况下，矢量空间的概念是比较宽泛的。每一个矢量空间都对应属于域 F 的标量集。通过对 F 中的元素进行加法和乘法运算，可生成 F 的新元素。一些常见的域有实数域 \mathbf{R} 、复数域 \mathbf{C} 、由 $[0, 1]$ 给出的二元域和有理多项式域。

F 域上的一个矢量空间 S ，是由一系列被称为矢量的元素构成的集合，且在其上可进行加法运算（+）和乘法运算（ \cdot ）。令 u, v, w 是 S 的元素（矢量），且 α, β 是标量，且为属于上式 F 的元素。 S 满足以下公理：

- (1) $u + v = v + u \in S$;
- (2) $\alpha(u + v) = \alpha u + \alpha v \in S$;
- (3) $(\alpha + \beta)u = \alpha u + \beta u$;
- (4) $(u + v) + w = u + (v + w)$;
- (5) $(\alpha\beta)u = \alpha(\beta u)$;
- (6) 存在一个零矢量 θ ，满足 $\theta + u = u$ 。 θ 经常简写为 0；
- (7) 对于标量 0 和 1，满足 $0u = 0, 1u = u$ ；
- (8) 对于每一个 u 矢量都存在一个负矢量 $-u$ ，且满足

$$u + (-u) = \theta$$

$-u$ 等于 $-1 \cdot u$ 。

下面给出了一些重要的矢量空间。

F^n 是具有 n 个分量（每个分量均属于域 F ）的列矢量空间。两种特殊情况分别是 \mathbf{R}^3 （它是实域 \mathbf{R} 的子域）以及 \mathbf{C}^n （它是复数域 \mathbf{C} 的子域）。 $F^{m \times n}$ 是所有 $m \times n$ 大小矩阵的空间，它属于 F 的子域。

下面是一些其他矢量空间的例子。

例 B.1 令 V, W 为同一个 F 域上的矢量空间。笛卡儿积 $V \times W$ 由一组有序的矢量对 $\{v, w\}$ 构成的集合组成，其中 $v \in V, w \in W$ 。 $V \times W$ 是一个矢量空间。 $V \times W$ 上的矢量加法和乘法定义如下：

$$\begin{aligned}\{v, w\} + \{p, q\} &= \{v + p, w + q\} \\ \alpha \{v, w\} &= \{\alpha v, \alpha w\} \\ \theta &= \{\theta_v, \theta_w\}\end{aligned}$$

式中， $\alpha \in F$ ； $v, p \in V$ ； $w, q \in W$ ； θ_v 和 θ_w 分别是 V 和 W 上的零矢量。

笛卡儿积矢量空间可被拓展为包含任意数目元素的矢量空间。

例 B.2 实区间 $[a, b]$ 上的变量 t 的所有复连续函数构成一个矢量空间，记为 $C[a, b]$ 。令 u 和 v 是该空间上的矢量，且 $\alpha \in F$ 。则矢量加法和标量乘法法则如下：

$$\begin{aligned}(u + v)(t) &= u(t) + v(t) \\ (\alpha u)(t) &= \alpha u(t)\end{aligned}$$

零矢量 θ 是在区间 $[a, b]$ 上恒等于零的函数。



B.2 矢量空间的性质

本节将讨论矢量空间的一般性质，而不是某一个特定的矢量空间的性质。

子空间

对非空矢量空间 L ，如果 L 的元素同时也是 S 的元素，且 S 可能有更多数目的元素，则 L 是空间 S 的一个子空间。令 M 和 N 为 S 的子集，则其满足如下两条性质：

- (1) 其交集 $M \cap N$ 是 S 的子集。
- (2) 其直和 $M \oplus N$ 是 S 的子集。直和定义如下：

直和

对于每一个 $s \in S$ 都存在唯一的 $s_1 \in S_1$ 和 $s_2 \in S_2$ ，其满足 $s = s_1 + s_2$ ，则 S 为 S_1 和 S_2 的直和，记为

$$S = S_1 \oplus S_2 \quad (\text{B. 2-1})$$

例 B.3 令例 B.2 中的 $(a, b) = (-\infty, \infty)$ ，考虑由下式给出的奇函数和偶函数：

$$u_e(t) = u_e(-t)$$

$$u_o(t) = -u_o(-t)$$

这些奇函数和偶函数分别构成了子集 S_o 和 S_e 。在集合 S 中的任何函数 $x(t)$ 都可以分解成一个奇函数和一个偶函数构成的形式：

$$u_e(t) = \frac{u(t) + u(-t)}{2} \quad (\text{B. 2-2})$$

$$u_o(t) = \frac{u(t) - u(-t)}{2} \quad (\text{B. 2-3})$$

且

$$u(t) = u_e(t) + u_o(t) \quad (\text{B. 2-4})$$

显然， S_o 、 S_e 的直和是 S 。

凸性

若矢量空间 S 的子空间 S_c 是凸的，则对于任意的 s_o 和 $s_1 \in S_c$ ，由下式给出的矢量 s_2 ：

$$s_2 = \lambda s_o + (1 - \lambda) s_1, 0 \leq \lambda \leq 1 \quad (\text{B. 2-5})$$

s_2 同样属于 S_c 。

在凸子空间中，子空间中任意两个点（矢量）之间的线段同样属于该子空间。

线性无关性

如果 u 可被写为如下所示的 S 中矢量的线性组合，则称 u 在相对于矢量 v_i 构成的集合 S 上线性相关。

$$u = \sum_k c_k v_k \quad (\text{B. 2-6})$$

式中， $c_k \in F$ 。

如果 u 不满足式 (B. 2-6)，称 u 相对于 S 空间内的矢量线性无关。如果一组矢量构成的集合中，每个矢量都相对于其他矢量线性无关，称该矢量集为线性无关组。

以下给出线性无关矢量的两个性质，而不作证明：



(1) 一组矢量 $\mathbf{u}_0, \mathbf{u}_1, \dots$ 线性无关, 若要满足 $\sum_k c_k \mathbf{u}_k = 0$, 则对所有 k , 均有 $c_k = 0$ 。

(2) 一组矢量 $\mathbf{u}_0, \mathbf{u}_1, \dots$ 线性无关, 若 $\sum_k c_k \mathbf{u}_k = \sum_k b_k \mathbf{u}_k$ 成立, 则对每个 k , 均有 $c_k = b_k$ 。

张成空间

如果矢量 \mathbf{u} 满足式 (B.2-6) 中所述的 S 中矢量的线性组合, 则称 \mathbf{u} 属于由 S 的子集张成的子空间矢量, 由 S 中的矢量张成的子空间定义为 $\text{span}(S)$ (S 的张成空间)。

基和维度

矢量空间 S 的基 (或坐标系) 是由一组线性相关的矢量构成的集合 B , 使得 S 中的每一个矢量是 B 中元素的线性组合。

矢量空间 S 的维数 M 等于集合 B 的元素的数目。若 M 是有限的, 则 S 是一个有限维矢量空间。否则, 它是无限维的。

如果 B 中的元素为 $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{M-1}$, 则 S 中的任意矢量 \mathbf{x} 可以表示为

$$\mathbf{x} = \sum_{k=0}^{M-1} w_k \mathbf{b}_k \quad (\text{B.2-7})$$

式中, w_k 是属于域 F 的标量。

对于有限维矢量空间 S , 其一组基中元素的个数与其他任意一组基中元素的个数相等。

在一组正交基中, 其中的任一矢量 \mathbf{b}_m 与其他矢量正交。如果 B 是一组正交基, 取 \mathbf{x} 与 \mathbf{b}_m 的内积代入式 (B.2-7), 可得

$$(\mathbf{u}, \mathbf{b}_m) = \sum_{k=0}^{M-1} w_k (\mathbf{b}_k, \mathbf{b}_m) = w_m (\mathbf{b}_m, \mathbf{b}_m)$$

所以

$$w_m = \frac{(\mathbf{x}, \mathbf{b}_m)}{(\mathbf{b}_m, \mathbf{b}_m)} \quad (\text{B.2-8})$$

例 B.4 若矢量空间为 F^n , 则 $n \times n$ 的单位矩阵 \mathbf{I} 的列向量是线性无关的, 同时也张成空间 F^n 。因此, 它们形成 F^n 的一组基。这样的基称为标准基。

B.3 内积矢量空间

实际感兴趣的矢量空间通常是结构化的, 其存在一个范数来表示矢量的长度或大小; 存在两个矢量间的方向的测度, 称之为内积; 存在任意两个矢量之间的距离的测度。这样的矢量空间被称为内积矢量空间。本章其余部分的讨论都将限定在这样的矢量空间。它们的性质将讨论如下。

内积矢量空间 S 中的两个矢量 \mathbf{u} 和 \mathbf{v} 的内积记为 (\mathbf{u}, \mathbf{v}) , 其为一个从 $S \times S \rightarrow D$ 的映射, 满足下述性质:

- (1) $(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})^*$
- (2) $(\alpha \mathbf{u}, \mathbf{v}) = \alpha (\mathbf{u}, \mathbf{v})$, α 是一个标量
- (3) $(\mathbf{u} + \mathbf{v}, \mathbf{w}) = (\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w})$



(4) 当 $\mathbf{u} \neq 0$ 时, $(\mathbf{u}, \mathbf{u}) > 0$, 当 $\mathbf{u} = 0$ 时, $(\mathbf{u}, \mathbf{u}) = 0$

当 \mathbf{u} 和 \mathbf{v} 是 N 元组时

$$\mathbf{u} = [u_0 \ u_1 \cdots u_{N-1}]^T$$

$$\mathbf{v} = [v_0 \ v_1 \cdots v_{N-1}]^T$$

(\mathbf{u}, \mathbf{v}) 可被定义为

$$(\mathbf{u}, \mathbf{v}) = \sum_{k=0}^{N-1} u_k v_k^* \quad (\text{B. 3-1})$$

当具有逐点加法和标量乘法的矢量空间 $C[a, b]$ 内的函数 $f(t)$ 与 $g(t)$ 为连续函数时, 其内积 (f, g) 可以表示为

$$(f, g) = \int_a^b f(t) * g(t) dt \quad (\text{B. 3-2})$$

矢量 \mathbf{u} 的欧几里得范数为 $\sqrt{(\mathbf{u}, \mathbf{u})}$, 将其记为 $|\mathbf{u}|$ 。如果矢量 \mathbf{v} 和矢量 \mathbf{u} 满足 $(\mathbf{u}, \mathbf{v}) = 0$, 则称两个矢量是正交的。在矢量空间 F^n 中, 定义了一个更常用的范数

$$|\mathbf{u}|_p = \left[\sum_{k=0}^{N-1} |u_k|^p \right]^{1/p} \quad (\text{B. 3-3})$$

式中, $1 \leq p \leq \infty$ 。当 $p=2$ 时, 该范数为欧几里得范数。

下面将讨论内积矢量空间的常见重要性质, 特别是一些典型内积矢量空间的性质。

距离测度

两个矢量 \mathbf{u} 和 \mathbf{v} 的距离 $d(\mathbf{u}, \mathbf{v})$ 表示这两个矢量的相似程度。内积矢量空间也是测度为 $d(\mathbf{u}, \mathbf{v})$ 的量度空间。距离 $d(\mathbf{u}, \mathbf{v})$ 可以通过多种方式定义, 只要其满足以下性质:

- (1) $d(\mathbf{u}, \mathbf{v}) \geq 0$
- (2) $d(\mathbf{u}, \mathbf{u}) = 0$
- (3) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
- (4) $d^2(\mathbf{u}, \mathbf{v}) \leq d^2(\mathbf{u}, \mathbf{w}) + d^2(\mathbf{w}, \mathbf{v})$

最后一条性质通常被称为施瓦兹不等式或者三角不等式。

矢量 \mathbf{u} 和 \mathbf{v} 的欧几里得距离如下式所示:

$$d(\mathbf{u}, \mathbf{v}) = |\mathbf{u} - \mathbf{v}| = [(\mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v})]^{1/2} \quad (\text{B. 3-4})$$

可见, 矢量 \mathbf{u} 的范数就是 $d(\mathbf{u}, \mathbf{0})$, 其中 $\mathbf{0}$ 是零矢量。

即使矢量的欧几里得距离可以表征矢量 \mathbf{u} 和 \mathbf{v} 的相似度, 但是它的大小取决于 \mathbf{u} 和 \mathbf{v} 的范数。为了消除这种依赖性, 可以通过将它除以 $|\mathbf{u}| |\mathbf{v}|$ 来进行归一化。

内积矢量空间的例子

两种最常见的内积空间是实数集合 \mathbf{R} 和复数集合 \mathbf{C} , 这里将矢量 \mathbf{u} 和 \mathbf{v} 的自然测度 $d(\mathbf{u}, \mathbf{v})$ 作为其欧几里得距离。

\mathbf{C}^n 和 \mathbf{R}^n 分别是 \mathbf{C} 和 \mathbf{R} 上的 n 维矢量空间。 \mathbf{C}^n 和 \mathbf{R}^n 的元素都可以表示成 n 元组。

ℓ_2 代表 \mathbf{C} 上的 $n=1, 2, \dots, \infty$, 所有的复序列 $\mathbf{u} = [u_n] (n=1, 2, \dots, \infty)$ 构成的矢量空间满足

$$\sum_{n=1}^{\infty} |u_n|^2 < \infty \quad (\text{B. 3-5})$$

且满足分量加法与标量乘法, 内积由下式给出:



$$(\mathbf{u}, \mathbf{v}) = \sum_{n=1}^{\infty} \mathbf{u}_n \mathbf{v}_n^* \quad (\text{B. 3-6})$$

式中, $\mathbf{v} = [\mathbf{v}_n]$ 。

分量加法的定义是:

如果满足 $\mathbf{w} = \mathbf{u} + \mathbf{v}$, 则对于所有的 n , 有

$$\mathbf{w}_n = \mathbf{u}_n + \mathbf{v}_n \quad (\text{B. 3-7})$$

标量乘法的定义为:

如果满足 $\mathbf{w} = \lambda \mathbf{u}$, λ 是标量, 则对于所有的 n , 有

$$\mathbf{w}_n = \lambda \mathbf{u}_n \quad (\text{B. 3-8})$$

$L_2(a, b)$ 是在 $C[a, b]$ (从复值区间 a 到 b) 上的连续函数的矢量空间, 其内积为

$$(\mathbf{u}, \mathbf{v}) = \int_a^b \mathbf{u}(t) \mathbf{v}^*(t) dt \quad (\text{B. 3-9})$$

角度

两个矢量 \mathbf{u} 和 \mathbf{v} 之间的夹角 θ 定义如下:

$$\cos \theta = \frac{(\mathbf{u}, \mathbf{v})}{|\mathbf{u}| |\mathbf{v}|} \quad (\text{B. 3-10})$$

θ 也表示两个矢量 \mathbf{u} 和 \mathbf{v} 之间的相似度。当 θ 等于 0 时代表矢量 \mathbf{u} 和 \mathbf{v} 是相同的矢量。 θ 等于 $\pi/2$ 代表矢量 \mathbf{u} 和 \mathbf{v} 是正交的。

柯西-施瓦兹不等式

两个矢量 \mathbf{u} 和 \mathbf{v} 的内积满足

$$|(\mathbf{u}, \mathbf{v})| \leq |\mathbf{u}| |\mathbf{v}| \quad (\text{B. 3-11})$$

只有当矢量 \mathbf{u} 和 \mathbf{v} 线性相关时等号成立 ($\mathbf{u} = \alpha \mathbf{v}$, α 是标量)。

三角不等式

矢量 \mathbf{u} 和 \mathbf{v} 满足

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}| \quad (\text{B. 3-12})$$

三角不等式也常被称为**施瓦兹不等式**。

正交性

令 $U = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}]$ 为内积空间 S 中的一个矢量集合 (序列)。当 $k \neq \ell$ 时 $(\mathbf{u}_k, \mathbf{u}_\ell) = 0$, 则 U 是正交集。如果 $(\mathbf{u}_k, \mathbf{u}_\ell) = \delta_{k\ell}$, U 具有正交性。

例 B.5 令内积空间为 $L_2(0, 1)$ 。则此空间上的函数序列

$$\mathbf{u}_k(t) = \cos(2\pi kt)$$

构成了一个正交序列。

B.4 希尔伯特空间

希尔伯特空间理论是非常有用的理论, 并且是将有限维内积矢量空间理论很好地推广到无穷维内积矢量空间的理论。为讨论希尔伯特空间, 需要完备性的概念。

考虑一个测度为 d 的量度空间 M 。对于 M 中的矢量序列 $[\mathbf{u}_k]$, 对任意 $\varepsilon > 0$, 都存在一个整数 k_0 , 使得对于任意 k , 都有 $d(\mathbf{u}_k, \mathbf{u}_\ell) < \varepsilon$, $\ell > k_0$, 则 $[\mathbf{u}_k]$ 是柯西序列。若 M 中每个



柯西序列都收敛, 且极限在 M 中, 则 M 是一个完备的量度空间, 一个柯西序列的例子就是 N 维傅里叶展开式的第 N 个部分和 (详见例 B. 6)。

希尔伯特空间是一个有完备量度的内积空间。例如, \mathbf{C}^n 和 \mathbf{R}^n 是希尔伯特空间, 因为它们都是完备的。

不需要完备性要求的内积空间有时也称为准希尔伯特空间。

例 B. 6 考虑一个周期为 T 的周期信号 $u(t)$ 。其基本周期可设为 $-T/2 \leq t \leq T/2$ 。 $u(t) \in L_2(-T/2, T/2)$ 的傅里叶级数表示式为

$$u(t) = \sqrt{2/T} \sum_{k=0}^{\infty} [U_1[k]q[k]\cos(2\pi kF_s t) + U_0[k]\sin(2\pi kF_s t)] \quad (\text{B. 4-1})$$

式中, $F_s = 1/T$ 。且

$$q[k] = \begin{cases} 1/\sqrt{2} & k=0 \\ 1 & \text{其他} \end{cases}$$

其基础函数为 $\sqrt{2/T}\cos(2\pi kF_s t)$ 和 $\sqrt{2/T}\sin(2\pi kF_s t)$, 且个数是可数的, 但是有无穷多个。这些基函数形成基函数的正交集, 该序列的系数 $U_1[k]$ 和 $U_0[k]$ 如下:

$$\begin{aligned} U_1[k] &= (u(t), \sqrt{2/T}q(k)\cos(2\pi kF_s t)) \\ &= \sqrt{2/T}q(k) \int_{-T/2}^{T/2} u(t)\cos(2\pi kF_s t) dt \\ U_0[k] &= (u(t), \sqrt{2/T}\sin(2\pi kF_s t)) \\ &= \sqrt{2/T} \int_{-T/2}^{T/2} u(t)\sin(2\pi kF_s t) dt \end{aligned}$$

在这种情况下, 希尔伯特空间是周期函数, 且在一个周期内二次方可积。这些矢量包含所有这样的周期函数。对应的柯西序列是式 (B. 4-1) 中第 N 个部分的和。所有这种序列的极限是 $x(t)$ 。

附录 C 离散时间傅里叶变换、离散傅里叶变换与快速傅里叶变换

C.1 离散时间傅里叶变换

信号序列 $u[n]$ 的离散时间傅里叶变化 (DTFT) 定义如下:

$$U(f) = \frac{1}{F} \sum_{n=-\infty}^{\infty} u[n] e^{-j2\pi f n T_s} \quad (\text{C. 1-1})$$

$u[n]$ 通常对应于一个采样信号 $u[nT_s]$, T_s 为采样间隔, 且 $F = 1/T_s$ 。在本章中 T_s 和 F 通常假定为 1。

逆 DTFT 变换为

$$u[n] = \int_{-F/2}^{F/2} U(f) e^{j2\pi f n T_s} df \quad (\text{C. 1-2})$$



DTFT 的存在条件是

$$E_1 = \|u\|_1 = \sum_{n=-\infty}^{\infty} |u[n]| < \infty \quad (\text{C. 1-3})$$

如果 $u[n]$ 满足式 (C. 1-3), 则下式亦成立:

$$E_1 \geq E_2 = \|u\|^2 = \sum_{n=-\infty}^{\infty} |u[n]|^2 < \infty \quad (\text{C. 1-4})$$

DTFT 的性质同我们在第 2 章讨论的傅里叶变换的性质十分相似。

例 C.1 求解下述信号的 DTFT:

$$u[n] = \begin{cases} 1 & |n| \leq N \\ 0 & \text{其他} \end{cases}$$

其中, $T_s = 1$ 。

解: $u[n]$ 信号如图 C. 1a 所示, 其中 $N = 12$ 。

根据定义有

$$U(f) = \sum_{k=-N}^N e^{-j2\pi f k} = \sum_{k=-N}^N (e^{-j2\pi f})^k$$

令 $\ell = k + N$, 则该式化简为

$$U(f) = e^{j2\pi f N} \sum_{\ell=0}^{2N} (e^{-j2\pi f})^{\ell}$$

等式右边的求和是一个等比数列的前 $(2N+1)$ 项的和, 该等比数列的和为 $(1 - r^{2N+1})/(1 - r)$, 这里的 r 等于 $e^{-j2\pi f}$ 。即

$$\begin{aligned} U(f) &= e^{-j2\pi f N} \frac{1 - e^{-j2\pi f (2N+1)}}{1 - e^{-j2\pi f}} \\ &= \frac{\sin(2\pi f (N + 1/2))}{\sin(\pi f)} \end{aligned}$$

图 C. 1b 给出了 $N = 12$ 时的频谱 $U(f)$ 。可见, 当 f 较小时, $U(f)$ 与 \sin 函数比较相似, 但是当 f 趋近于 1 时, $U(f)$ 与 \sin 函数则变得完全不同。

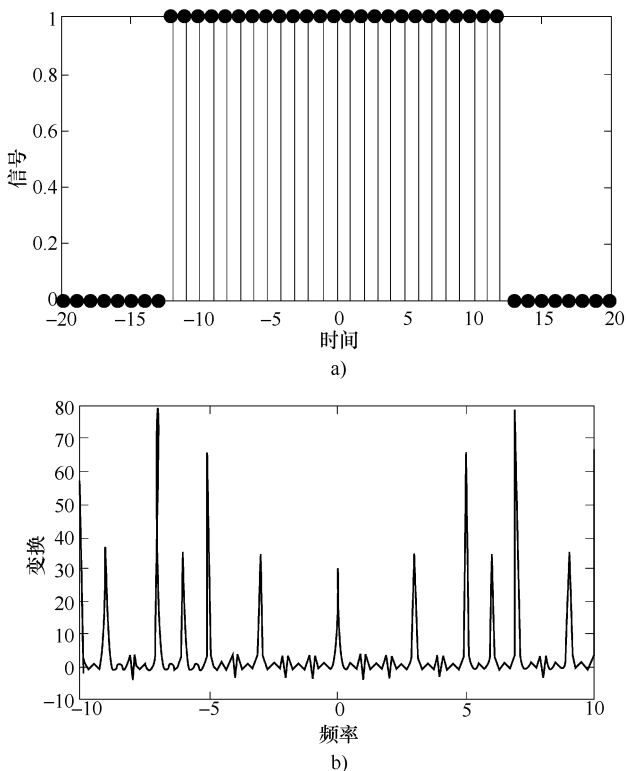


图 C.1 $N = 12$ 的矩形脉冲信号及其 DTFT

a) $N = 12$ 的矩形脉冲信号 b) DTFT

C.2 离散时间傅里叶变换和傅里叶变换之间的联系

根据式 (1.2.2), 考虑模拟信号 $u(t)$ 的傅里叶表征

$$u(t) = \int_{-\infty}^{+\infty} U'(f) e^{j2\pi f t} df \quad (\text{C. 2-1})$$

其中 $U'(f)$ 是 $u(t)$ 的傅里叶变换。

序列 $u[n]$ 是信号 $u(t)$ 以采样间隔 T_s 采样时生成的序列, 即 $u[n] = u(nT_s)$ 。比较式 (C. 2-1) 和式 (C. 1-2) 可以得到 $u[n]$ 的离散时间傅里叶变换 $U(f)$ 和 $U'(f)$ 的关系满足



$$\int_{-F/2}^{F/2} U(f) e^{j2\pi f n T_s} df = \int_{-\infty}^{+\infty} U'(f) e^{j2\pi f n T_s} df \quad (\text{C. 2-2})$$

将右侧的积分表示成一系列积分的和, 对于每一个 F , 式 (C. 2-2) 可以写成

$$\int_{-F/2}^{F/2} U'(f) e^{j2\pi f n T_s} df = \int_{-F/2}^{F/2} \left[\sum_{\ell=-\infty}^{\infty} U'(f + \ell F) \right] e^{j2\pi f n T_s} df \quad (\text{C. 2-3})$$

令被积函数相等

$$U(f) = \sum_{\ell=-\infty}^{\infty} U'(f + \ell F) \quad (\text{C. 2-4})$$

如果 $U'(f)$ 的带宽小于 F , 则

$$U(f) = U'(f) \quad (\text{C. 2-5})$$

为方便起见, T_s 通常取为 1。这样 $u(t)$ 实际上可用 $u(tT_s)$ 替代。根据 2.5 节中讨论的 FT 的性质, $u(tT_s)$ 的傅里叶变换为 $1/T_s U'(f/T_s)$ 。于是, 式 (C. 2-4) 可写为

$$U(f) = \frac{1}{T_s} \sum_{\ell=-\infty}^{\infty} U'\left(\frac{f + \ell}{T_s}\right) \quad (\text{C. 2-6})$$

式中, $U(f)$ 是采样频率为 1 时从 $u(tT_s)$ 中所得到的采样信号的离散时间傅里叶变换。

式 (C. 2-4) 和式 (C. 2-6) 说明非带限信号的离散时间傅里叶变换是连续时间信号傅里叶变换的一个偏离版本。当 f 接近 F 时, 所产生的误差变得比较显著, 因为这时 $U(F)$ 等于 $U(0)$, 且 $U(0)$ 通常很大。

C.3 离散傅里叶变换

长度为 N 的周期序列 $u[n]$ 的正交归一化离散傅里叶变换定义为

$$U[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u[n] e^{-j2\pi nk/N} \quad (\text{C. 3-1})$$

其逆离散傅里叶变换为

$$u[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} U[k] e^{j2\pi nk/N} \quad (\text{C. 3-2})$$

上述定义的离散傅里叶变换是正交的。当不需要归一化时, 该方程可以进一步简化

$$U[k] = \sum_{n=0}^{N-1} u[n] e^{-j2\pi nk/N} \quad (\text{C. 3-3})$$

$$u[n] = \frac{1}{N} \sum_{k=0}^{N-1} U[k] e^{j2\pi nk/N} \quad (\text{C. 3-4})$$

对于二维信号, 通过先对信号矩阵的行应用一维离散傅里叶变换, 然后再对其列应用一维离散傅里叶变换, 或者先列后行, 这样就可以得到二维离散傅里叶变换。若二维周期性序列为 $u[n_1, n_2]$, $0 \leq n_1 \leq N_1$, $0 \leq n_2 \leq N_2$, 则其二维离散傅里叶变换为

$$U[k_1, k_2] = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} u[n_1, n_2] e^{-2\pi j \left[\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2} \right]} \quad (\text{C. 3-5})$$

逆二维 DFT 变换为

$$u[n_1, n_2] = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} U[k_1, k_2] e^{2\pi j \left[\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2} \right]} \quad (\text{C. 3-6})$$



C.4 快速傅里叶变换

快速傅里叶变换 (FFT) 是指一组用于快速计算 DFT 的算法。FFT 算法将 N 个点的傅里叶变换计算复杂度从 $O(2N^2)$ 减少到 $O(2N\log N)$ 。

FFT 是一种基于分治法的算法。当数据点个数 N 为合数, 且可被写为 $N = r_1 r_2 \cdots r_m$ 时, FFT 算法效率最高。当 $y'_0 = c_0 + c_1$, $y'_1 = c_0 + c_2$, N 等于 r^m 时, r 称为 FFT 算法的基。这种情况下, FFT 算法具有规则的结构。

分治法通常由时分算法 (DIT) 或频分算法 (DIF) 完成。为了理解这两个概念, 令 $N = 2^m$ 。在时分算法中, 可将数据 $u(n)$ 分解为两个序列

$$\left. \begin{aligned} u_1[n] &= u[2n] \\ u_2[n] &= u[2n+1] \end{aligned} \right\} n=0, 1, \dots, \frac{N}{2}-1 \quad (\text{C. 4-1})$$

在频分算法中, 频率部分可以分为两部分, 即 $U[2k]$ 和 $U[2k+1]$, $k=0, 1, \dots, \frac{N}{2}-1$ 。

进一步考虑时分算法。式 (C. 1-1) 可化为

$$U[k] = U_1[k] + e^{-j\frac{2\pi k}{N}} U_2[k] \quad (\text{C. 4-2})$$

$$U_1[k] = \sum_{m=0}^{N_2-1} u[2m] e^{-j\frac{2\pi km}{N_2}} \quad (\text{C. 4-3})$$

$$U_2[k] = \sum_{m=0}^{N_2-1} u[2m+1] e^{-j\frac{2\pi km}{N_2}} \quad (\text{C. 4-4})$$

式中, $N_2 = N/2$ 。

显然, $U_1[k]$ 和 $U_2[k]$ 是由长度为 $N/2$ 、分别由原信号中偶数和奇数数据点经 DFT 变换得到的。

上面描述的是一种连续迭代的过程, 直到达到长度为 2 的 DFT 变换为止, 其中仅包括两个点的加法和减法。

基-2 的频分算法可以通过类似的推导得出。然而, 其实际操作会更加复杂。

二维离散傅里叶变换是通过对信号矩阵的行先进行一维 DFT, 然后对信号矩阵的列再进行一维 DFT 得到的, 或者先列后行。

参 考 文 献

- Asakura, K. and T. Nagashima, "Reconstruction from computer-generated holograms displayed by a line printer," *Optics Communications* 17, 273-276, 1976.
- Besag, J. E. "Spatial Interaction and the Statistical Analysis of Lattice Systems," *Journal of Royal Statistical Society: Series B* 36, 2, 192-236, May 1974.
- Bennett, J. R., I. G. Cumming and R. A. Deane, "The Digital Processing of SEASAT Synthetic Aperture Radar Data," *Proceedings of the IEEE International Radar Conference*, 168-174, Virginia, 1980.
- Born, M. and E. Wolf, *Principles of Optics*, Pergamon Press, New York, 1969.
- Bowden, M., L. Thompson, C. Wilson, eds., *Introduction to Microlithography*, American Chemical Soc., ISBN 0-8412-2848-5, 1983.
- Boyer, A. L., "Formation of Images Using 5Mhz Ultrasound and a Computer," IBM Publication 320.2403, May 19, 1971.
- Boyer, A. L., P. M. Hirsch, J. A. Jordan, Jr., L. B. Lesem, D. L. Van Rooy, "Reconstruction of Ultrasonic Images by Backward Propagation," IBM Technical Report, No. 320.2396, IBM Scientific Center, Houston, Texas, July 1970.
- Brackett, C. A., A. S. Acampora, I. Sweitzer, G. Tangonan, M. T. Smith, W. Lennon, K. C. Wang, and R. H. Hobbs, "A scalable multiwavelength multihop optical network: A proposal for research on all-optical networks," *J. Lightwave Technol.*, vol. II, pp. 736-753, May/June 1993.
- Brackett, C. A., "Dense wavelength division multiplexing networks: Principles and applications," *IEEE J. Select. Areas Commun.*, Vol. 8, pp. 948-964, 1990.
- Brackett, C. A., A. S. Acampora, I. Sweitzer, G. Tangonan, M. T. Smith, W. Lennon, K. C. Wang, and R. H. Hobbs, "A Scalable Multiwavelength Multihop Optical Network: A Proposal for Research on All-Optical Networks," *J. Lightwave Technol.*, vol. II, pp. 736-753, May/June 1993.
- Brigham, E. O., *The Fast Fourier Transform*, Prentice Hall, Englewood, CA, 1974.
- Brown, B. R. and A. W. Lohmann, "Complex Spatial Filtering with Binary Masks," *Applied Optics*, 5, 967-969, June 1966.
- Bubb, C. E., Okan K. Ersoy, "Algorithms for Holographic Reconstruction of Three-Dimensional Point Source Images," Technical Report TR-ECE-06-06, Purdue University, March 2006.
- Burckhardt, C. B., "Use of a Random Phase Mask for the Recording of Fourier Transform Holograms of Data Masks," *Applied Optics*, 9, 695-700, 1970.
- Caulfield, H. J. and S. Lu, *The Applications of Holography*, Interscience, New York, 1970.
- Chu, D. C., J. R. Fienup, J. W. Goodman, "Multi-emulsion On-Axis Computer Generated Hologram," *Applied Optics*, 12, 1386-1388, 1973.
- Collier, R. J., C. B. Burckhardt, and L. H. Lin, *Optical Holography*, Academic Press, New York, 1971.
- Cutrona, L. J., E. N. Leith, C. J. Palermo, L. J. Porcello, "Optical Data Processing and Filtering Systems," *IRE Tran. Information Theory*, 386-400, June 1960.
- Dittman, J., L. C. Ferri, C. Vielhauer, "Hologram watermarks for document authentications," *International Conference on Information Technology: Coding and Computing*, 60-64, April 2001.
- Doles, J. H., "Broad-Band Array Design Using the Asymptotic Theory of Unequally Spaced Arrays," *IEEE Tran. Antennas and Propagation*, Vol. 36, No. 1, 27-33, January 1988.



- Dragone, C. , “An $N \times N$ Optical Multiplexer Using a Planar Arrangement of Two Star Couplers.” IEEE Photon. Technol. Lett. 3, 812-815, September 1991.
- Erdelyi, A. , Tables of Integral Transforms, McGraw Hill, New York, 1954.
- Ersoy, O. K. , “Construction of Point Images with the Scanning Electron Microscope: A Simple Algorithm,” Optik, 46, 61-66, September 1976.
- Ersoy, O. K. , “One-Image-Only Digital Holography,” Optik, 53, 47-62, April 1979.
- Ersoy, O. K. , “Real Discrete Fourier Transform,” IEEE Transactions Acoustics, Speech, Signal Processing, ASSP-33, 4, 880-882, August 1985.
- Ersoy, O. K. , J. Y. Zhuang, J. Brede, “Iterative Interlacing Approach for the Synthesis of Computer-Generated Holograms,” Applied Optics, 31, 32, 6894-6901, November 10, 1992.
- Ersoy, O. K. , “A Comparative Review of Real and Complex Fourier-Related Transforms,” Proceedings of the IEEE, 82, 3, 429-447, March 1994.
- Ersoy, O. K. , “Method of Increasing Number of Allowable Channels in Phased Array DWDM Systems,” USA Patent No. 6917736, July 12, 2005.
- ESA, “Satellite Data Yields Major Results in Greenland Glaciers Study,” ESA website, http://www.esa.int/es-aEO/SEMH59MVGJE_index_0.html, February 21, 2006.
- Farhat, N. H. , Advances in Holography, Marcel Dekker, New York, 1975.
- Feit, M. D. and J. A. Fleck, Jr. , “Light propagation in graded-index optical fibers,” Applied Optics, 17, 24, 3990-3998, December 15, 1978.
- Feldman, M. R. ; J. E. Morris, I. Turlik, P. Magill, G. Adema, M. Y. A Raja, “Holographic Optical Interconnects for VLSI Multichip Modules,” IEEE Tran. Components, Packaging and Manufacturing Technology, Part B: Advanced Packaging, 17, 2, 223-227, May 1994.
- Firth, I. M. , Holography and Computer Generated Holograms, Mills and Boon, Ltd. , London, 1972.
- Gabor, D. , “Resolution Beyond the Information Limit in Transmission Electron Microscopy, Nature, 161, 777-778, 1948.
- Gallagher, N. C. and D. W. Sweeney, “Infrared Holographic Optical Elements with Applications to Laser Material Processing,” IEEE J. of Quantum Elec. , QE-15, 1369-1380, December 1979.
- Garner, W. , W. Gautschi, “Adaptive Quadrature-Revisited,” BIT, 40, 84-101, 2000.
- Gaylord, T. K. , M. G. Moharam, “Analysis and Application of Optical Diffraction Gratings,” Proc. IEEE, 73, 894-936, 1985.
- Gerrard, A. and J. M. Burch, Introduction to Matrix Methods in Optics, J. Wiley, New York, 1975.
- Gerchberg, R. , “Superresolution through Error Energy Reduction,” Optica Acta, 21, 709-720, 1974.
- Goodman, J. W. , Introduction to Fourier Optics, 3rd Edition, Roberts and Company, Greenwood Village, Colorado, 2004.
- Hadley, G. R. , “Transparent Boundary Condition for the Beam Propagation Method,” IEEE J. Quantum Electronics, 26, 1, 109-112, January 1990.
- Hadley, G. R. , “Wide-Angle Beam Propagation Using Padé Approximant Operators,” Optics Letters, 17, 20, 1426-1431, October 15, 1992.
- Hayes, M. H. , J. S. Lim, A. V. Oppenheim, “Signal Reconstruction from the Phase or Magnitude of its Fourier Transform,” IEEE Tran. Acoustics, Speech and Signal Proc. , ASSP-28, 670-680, 1980.
- Hecht, E. , Optics, Addison-Wesley, Mass. , 2002.
- Hu, Sai, O. K. Ersoy, “Design and Simulation of Novel Arrayed Waveguide Grating by Using the Method of Irregularly Sampled Zero Crossings,” TR-ECE 02-05, Purdue University, December 2002.



- Ishimaru, A. , “ Theory of Unequally Spaced Arrays, ” IRE Tran. Antennas and Propagation, 691-701, November 1962.
- Kelly, Kozma and D. L. , “Spatial Filtering for Detection of Signals Submerged in Noise,” Applied Optics, Vol. 4, No. 4, pp. 389-392, 1965.
- Kim, C. -J. , R. R. Shannon, “Catalog of Zernike Polynomials,” in Applied Optics and Optical Engineering, edited by R. Shannon and J. Wyant, Ch. 4, Vol. X, Academic Press, San Diego, 1987.
- Kock, W. E. , Engineering Applications of Lasers and Holography, Plenum, New York, 1975.
- Kozma, A. , D. L. Kelly, Spatial filtering for detection of signals submerged in noise,” Applied Optics, 4, 387-392, 1965.
- Kuhl, P. , Okan K. Ersoy, “Design Of Diffractive Optical Elements: Optimality, Scale And Near-Field Diffraction Considerations,” Technical Report TR-ECE-03-09, Purdue University, 2003.
- Kunz, The Finite Difference Time Domain Method for Electromagnetics, CRC Press, Boca Raton, Florida, 1993.
- Lalanne, P. , G. M. Morris, “Highly Improved Convergence of the Coupled-Wave Method for TM Polarization,” J. Opt. Soc. Am. , 13, 4, 779-784, April 1996.
- Lalanne, P. , E. Silberstein, “Fourier-Modal Methods Applied to Waveguide Computational Problems,” Optics Letters, 25, 15, 1092-1094, August 1, 2000.
- Lanczos, A. , Discourse on Fourier Series, Hafner Publishing Co. , New York, 1966.
- Lalor, E. , “Inverse Wave Propagator,” J. Mathematical Physics, Vol. 9, No. 12, pp. 2001-2006, December 1968.
- Lao, Y. T. , “A Mathematical Theory of Antenna Arrays with Randomly Spaced Elements,” IEEE Tran. Antennas and Propagation, 257-268, May 1964.
- Laude, J. P. , Wavelength Division Multiplexing, Prentice Hall, NY, 1993.
- Lee, W. H “Sampled Fourier Transform Hologram Generated by Computer,” Applied Optics, 9, 639-643, 1970.
- Lee, W. H. , “Circular Carrier Holograms,” J. Opt. Soc. Am. 65, 518-523, 1975.
- Lee, W. -H. , “Binary Computer-Generated Holograms,” Applied Optics, Vol. 18, No. 21, pp. 3661-3669, November 1, 1979.
- Leith, E. N. , J. Upatnieks, “Reconstructed Wavefronts and Communication Theory,” J. Opt. Soc. Am. 52, 1123-1130, 1962.
- Lesem, L. B. , P. M. Hirsch, J. A. Jordan, jr. , “The kinoform: a new wavefront reconstruction device,” IBM J. Res. Develop. 13, 150-155, March 1969.
- Levi, A. , H. Stark, “Image Restoration by the Method of Generalized Projections with Application to Restoration from Magnitude,” J. Opt. Soc. Am. A, 1, 9, 932-943, September 1984.
- Ljunggren, S. , O. Lovhaugen, E. Mehlum, “Seismic Holography in a Norwegian Fjord,” Acoustical Imaging, 8, 299-315, 1980.
- Lohmann, A. W. , in The Engineering Uses of Holography, E. R. Robertson and J. N. Harvey, Eds. , Cambridge U. P. , London, 1970.
- Lohmann, A. W. , D. P. Paris, “Binary Fraunhofer Holograms Generated by Computer,” App. Opt. 6, 1739-1748, 1967.
- Lu, Ying, Okan Ersoy, Dense Wavelength Division Multiplexing/Demultiplexing By The Method Of Irregularly Sampled Zero Crossing, Technical Report TR-ECE-03-12, Purdue University, 2003.
- Marathay, A. , Diffraction, in Handbook of Optics, Volume 1: Fundamentals, Techniques, and Design, 2nd edition, McGraw-Hill, New York, pages 3.1-3.31, 1995.
- Marcuse, D. , Theory of Dielectric Optical Waveguides, 2nd edition, Academic Press, San Diego, 1991.



- Meier, R. W. , "Magnifications and Third Order Aberrations in Holography," J. Opt. Soc. Am. 56, 8, 987-997, 1966.
- Mellin, S. D. , G. P. Nordin, "Limits of Scalar Diffraction Theory and an Iterative Angular Spectrum Algorithm for Finite Aperture Diffractive Optical Element Design," Optics Express, Vol. 8, No. 13, pp. 705-722, 18 June, 2001.
- Diffraction," J. Optics A; Pure Appl. Optics, Vol. 5, pp. 86-91, 2003.
- Moharam, M. G. , T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings." J. Opt. Soc. Am. , 72, 1385-1392, 1982.
- Nikon, Phase Contrast Microscopy web page, <http://www.microscopyu.com/articles/phasecontrast/phasemicroscopy.html>.
- Okamoto, K. , "Recent Progress of Integrated Optics Planar Lightwave Circuits," Optical and Quantum Electronics, 31, 107-129, 1999.
- Peng, S. , G. M. Morris, "Efficient Implementation of Rigorous Coupled-Wave Analysis for Surface Relief Gratings, J. Opt. Soc. Am. 12, 1087-1096, 1995.
- Pojanasomboon, P. , O. K. Ersoy, "Iterative Method for the Design of a Nonperiodic Grating-Assisted Directional Coupler," Applied Optics, 40, 17, 2821-2827, June 2001.
- Pennings, E. C. M. , M. K. Smit, and G. D. Khoe, "Micro-Optic versus waveguide devices - An Overview, invited paper," inProc. Fifth Micm Opdes Conf 1995, Hiroshima. Japan. Oct. 18-20. 1995, pp. 248-255.
- Pennings, E. C. M. , M. K. Smit. A. A. M. Staring, and G. -D. Khoe. "Integrated-Optics versus Micro-Optics - A comparison." Integrated Photonics Research IPR '96. Boston. MA.
- Tech. Dig. . vol. 6. Apr. 29-May 2. 1996. pp. 460-463.
- Ralston, A. , H. S. Wilf, Mathematical Methods for Numerical Analysis, J. Wiley, New York, 1962.
- Rooy, D. L. Van, "Digital Ultrasonic Wavefront Reconstruction in the Near Field," IBM Technical Report, No. 320.2402, IBM Scientific Center, Houston, Texas, May 1971.
- Shannon, R. , inOptical Instruments and Techniques, edited by J. Dickson, pp. 331-345, Oriel Press, England, 1970.
- Shen, F. , A. Wang. "Fast Fourier Transform Based Numerical Integration Method for the Rayleigh-Sommerfeld Diffraction Formula," Applied Optics, Vol. 45, No. 6, pp. 1102-1110, 20 February 2006.
- Shewell, J. R. , "Inverse Diffraction and a New Reciprocity Theorem," J. Optical Society of America, Vol. 58, No. 12, pp. 1596-1603, December 1968.
- Smith, H. M. Principles of Holography (Wiley, New York, 1975) .
- Smit, M. K. , "Now focusing and dispersive planar component based on an optical phased array." Electron. Lett. , Vol. 24, No. 7, pp. 385-386, March 1988.
- Soumekh, Mehrdad, Synthetic Aperture Radar Signal Processing with Matlab Algorithms, J. Wiley, New York, 1999.
- Southwell, W. H. , "Validity of the Fresnel Approximation in the Near Field," J. Optical Society of America, Vol. 71, No. 1, pp. 7-14, 1981.
- Stark, H. , Applications of Optical Fourier Transform, Academic Press, Orlando, 1982.
- Stark, H. , editor, Image Recovery: Theory and Application, Academic Press, 1987.
- Steane, M. , H. N. Rutt, "Diffraction Calculations in the Near Field and the Validity of the Fresnel Approximation," Journal of the Optical Society of America A, 6, 12, 1809-1814, 1989.
- Stroke, G. W. , An Introduction to Coherent Optics and Holography, Academic Press, New York, 1975.
- Taflove, A. , S. C. Hagness, Computational Electrodynamics: the Finite Difference Time Domain Method, 3rd Edi-



- tion, Artech House, Norwood, MA, 2005.
- Takahashi, H., S. Suzuki, K. Kaco, and I. Nishi, "Arrayed-Waveguide Grating for Wavelength Division Mult/De-multiplexer with Nanometer Resolution," *Electron. Lett.*, 26, 2, 87-88, January 1990.
- F. T. Ulaby, R. K. Moore and A. K. Fung, *Microwave Remote Sensing Volume I*, Addison-Wesley, Reading, MA, 1981.
- Van Rooy, D. L., "Digital Ultrasonic Wavefront Reconstruction in the Near Field," IBM Publication No. 320.2402, May 19, 1971.
- Weyl, H., "Ausbreitung Elektromagnetischer Wellen über Einem Ebenen Leiter," *Ann. Phys. Lpz.* 60, 481-500, 1919.
- Wikipedia, website <http://en.wikipedia.org/wiki/Diffraction>.
- Wilson, T., editor, *Confocal Microscopy*, Academic Press, London, 1990.
- Ziemer, R. E., W. H. Tranter, *Principles of Communications*, 5th edition, J. Wiley, New York, 2002.
- Ueda, M. and K. Ieyasu, "A Method for a Faithful Reconstruction of an Off-Axis Type Ultrasound Holography," *Optik* 42, 107-112, 1975.
- Vellekoop, A. R. and M. K. Smit, "Low-Loss Planar Optical Polarization Splitter with small dimensions," *Electron. Lett.*, Vol. 25, pp. 946-947, 1989.
- Verbeek, B. and M. K. Smit, "Phased array based WDM devices," in *Proc. Eur. Conf. on Optical Communication (ECOC'95)*, Brussels, Belgium, Sept 17-21, 1995, pp. 195-202.
- Vesperinas, M. N., *Scattering and Diffraction in Physical Optics*, J. Wiley, New York, 1991.
- Watanebe, W., D. Kuroda, K. Itoh, "Fabrication of Fresnel Zone Plate Embedded in Silica Glass by Femtosecond Laser Pulses," *Optics Express* 10, 19, 978-983, September 2002.
- Yatagai, T., "Stereoscopic Approach to 3-D Display Using Computer-Generated Holograms," *Applied Optics* 15, 11, 2722-2729, 1976.
- Yee, K. S., "Numerical solutions of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas and Propagation*, 14, 302-307, 1966.
- Yevick, D., "New Formulations of the Matrix Beam Propagation Method: Application to Rib Waveguides," *IEEE J. Quantum Electronics*, 25, 2, 221-229, February 1989.
- Yevick, D., "Efficient Beam Propagation Techniques," *IEEE J. Quantum Electronics*, 26, 1, 109-112, January 1990.
- Yin, Yun, Okan K. Ersoy, Xianfan Xu, Ihtesham H. Chowdhury, "Fabrication/Analysis of Waveguides in Fused Silica by Femtosecond Laser Pulses and Simulation Studies for Dense Wavelength Division Multiplexing," Technical Report TR-ECE-05-05, Purdue University, May 2005.
- Youla, D. C., H. Webb, "Image Restoration by the Method of Convex Projections: Part I -Theory," *IEEE Trans. Med. Imaging*, MI-1, 81-94, 1982.
- Zhuang, J., O. K. Ersoy, "Optimal Decimation-in-Frequency Iterative Interlacing Technique for the Synthesis of Computer-Generated Holograms," *J. Optical Society of America, A*, Vol. 12, No. 7, pp. 1460-1468, July 1995.
- Ziemer, R. E., W. H. Tranter, *Principles of Communications*, 5th edition, J. Wiley, New York, 2002.



Diffraction, Fourier Optics and Imaging

地址：北京市百万庄大街22号
邮政编码：100037

电话服务

服务咨询热线：010-88361066

读者购书热线：010-68326294

010-88379203

网络服务

机工官网：www.cmpbook.com

机工官博：weibo.com/cmp1952

金书网：www.golden-book.com

教育服务网：www.cmpedu.com

封面无防伪标均为盗版



机械工业出版社微信公众号

WILEY

Copies of this book sold without
a Wiley Sticker on the cover are
unauthorized and illegal

ISBN 978-7-111-52389-5

策划编辑◎吉玲 / 封面设计◎张静

ISBN 978-7-111-52389-5



9 787111 523895 >

定价：65.00元