ISSN 1001-5868
CODEN BAGUE5

在线投稿 阅读

半导体光电

SEMICONDUCTOR OPTOELECTRONICS

1
2020
第41卷

中文核心期刊　RCCSE 中国核心学术期刊（A）　中国科技核心期刊
INSPEC 数据库来源期刊　Scopus 数据库来源期刊　《CA》、《AJ》收录期刊

ISSN 1001-5868
9 771001 586206 02>

重庆光电技术研究所 主办
Chongqing Optoelectronics Research Institute

半导体光电
BANDAOTI GUANGDIAN
双月刊·1976年创刊·国内外公开发行
2020年2月第41卷第1期(总第207期)

目　次

※动态综述※

基于深度学习的目标检测技术的研究综述 …… 罗　元，王薄宇，陈　旭(1)

※光电器件※

一种低交叉轴灵敏度微米光栅加速度计设计 …… 高　山，冯丽爽，周　震，等(11)
基于有限元法的12 μm微测辐射热计单元设计 …… 史佳欣，陈　超，王　军(15)
级联倍增InAlAs/InAlGaAs雪崩光电二极管暗电流成分分析 …… 张　承，黄晓峰，迟殿鑫，等(20)
基于敏感薄膜的F-P光纤温度传感器的高温特性 …… 高晓丹，彭建坤(25)
谐振式光纤陀螺的误差及噪声分析 …… 梅　魁，蒋治国，胡宗福(29)
基于表面等离子体共振的光子晶体光纤折射率传感器的设计与分析 …… 魏方皓，张祥军，唐守锋(35)
基于遗传算法的太阳能飞艇光伏阵列优化研究 …… 刘晓光，汪岳峰(39)
基于石墨烯探测器的积分电路设计 …… 任思伟，申　钧，刘昌举(44)
一种带温度参量的半导体激光器速率方程及仿真求解 …… 张亚标，肖永川，王　超，等(50)
基于微多环谐振陀螺仪的馈通效应研究 …… 田梦雅，张卫平，谷留涛，等(54)
基于人眼视觉特性的IRFPA非均匀性校正算法研究 …… 代少升，易伟男，聂合文，等(59)
基于银纳米薄膜欧姆接触的高波长选择性紫外探测器研究 …… 梁志斌，张齐轩，许朝军，等(64)
高分辨率AM Micro-LED显示器设计及其驱动实现 …… 林杰泓，聂君扬，张永爱，等(68)
光纤光栅温度传感器增敏封装特性研究 …… 王　甫，江　淮，易小龙，等(73)

※材料、结构及工艺※

不同应力增透膜对半导体激光器性能的影响 …… 崔碧峰，程　瑾，郝　帅，等(77)
温度和材料参数对$InAs_xSb_{1-x}$俄歇复合寿命影响的数值分析 …… 张景波，张云琦，王思文，等(80)
高强度光纤光栅与碳纤维的复合及应变传感特性研究 …… 侯智懿，李　彦，邵洪峰，等(85)
InAs量子点低温盖层对其发光特性的影响 …… 吴　唯，周　勇，刘尚军，等(89)
基于方形腔耦合金属波导结构的全光等离子体开关 …… 肖功利，窦婉滢，杨宏艳，等(93)
NLDMOS器件性能优化及分析 …… 李维杰，王　兴，王云峰，等(99)

※光电技术及应用※

微小零件显微检测图像的清晰度评价 …… 李成超，于占江，李一全，等(103)
基于像元纯净指数的多端元提取算法 …… 杨华东，许　楠，郝永平(108)
应用于微型扑翼飞行器的实时视觉系统设计 …… 王昕鹏，张卫平，牟家旺，等(114)
丙丁烷TDLAS测量系统的吸收峰自动检测 …… 张　铭，苗玉彬，张　舒(118)
基于随机Radon变换的DRPE加密系统 …… 彭凯飞，沈学举，黄富瑜，等(123)
针对间断纹理环境中的图像特征追踪和匹配算法研究 …… 赵明富，曹利波，宋　涛，等(128)
基于非均匀有序直方图和锐化的红外图像增强 …… 吴　婷(135)
联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测 …… 张　炎，华文深，黄富瑜，等(141)
一种基于AP与消除基本陷阱集的低错误平层QC-LDPC码构造方法 …… 袁建国，王宏森，张希瑞，等(146)

SEMICONDUCTOR OPTOELECTRONICS

Vol. 41 **No.** 1 **February** 2020(**Bimonthly**) (**Total No.** 207) (**Started in** 1976)

CONTENTS

Research Progresses of Target Detection Technology Based on Deep Learning ····· LUO Yuan, WANG Boyu, CHEN Xu(1)
Design of Micro-grating Accelerometer with Low Cross-axis Sensitvity ····· GAO Shan, FENG Lishuang, ZHOU Zhen, et al.(11)
Design of 12 μm Microbolometer Based on Finite Element Method ····· SHI Jiaxin, CHEN Chao, WANG Jun(15)
Analysis of Dark Current Components of InAlAs/InAlGaAs Cascade Multiplication Avalanche Photodiode ····· ZHANG Cheng, HUANG Xiaofeng, CHI Dianxin, et al.(20)
Study on High Temperature Characteristics of F-P Fiber Temperature Sensor Based on Sensitive Film ····· GAO Xiaodan, PENG Jiankun(25)
Analysis on Error and Noise of Resonant Fiber Optic Gyros ····· MEI Su, JIANG Zhiguo, HU Zongfu(29)
Design and Analysis of Photonic Crystal Fiber Refractive Index Sensor Based on Surface Plasmon Resonance ····· WEI Fanghao, ZHANG Xiangjun, TANG Shoufeng(35)
Genetic Algorithm-based Topology Optimization of Photovoltaic Array of Solar-powered Airships ····· LIU Xiaoguang, WANG Yuefeng(39)
Integral Circuit Design Based on Graphene Detector ····· REN Siwei, SHEN Jun, LIU Changju(44)
Semiconductor Laser Rate-Equation with Temperature Parameter and Emulation Solving ····· ZHANG Yabiao, XIAO Yongchuan, WANG Chao, et al.(50)
Research on The Feed-through Effect of Micro Multi-ring Resonant Gyroscope ····· TIAN Mengya, ZHANG Weiping, GU Liutao, et al.(54)
Research on IRFPA Non-uniformity Correction Algorithm Based on Human Visual Characteristics ····· DAI Shaosheng, YI Weinan, NIE Hewen, et al.(59)
High Wavelength-selective Ultraviolet Detector Based on Ag-nanofilm Ohmic Contact ····· LIANG Zhibin, ZHANG Qixuan, XU Chaojun, et al.(64)
Design and Driver Implementation of High Resolution AM Micro-LED Display ··· LIN Jiehong, NIE Junyang, ZHANG Yongai, et al.(68)
Sensitivity Enhancement and Encapsulation of Fiber Bragg Grating Temperature Sensors ····· WANG Fu, JIANG Huai, YI Xiaolong, et al.(73)
Effect of Different Facet Coating Stress on The Performance of Laser Diodes ····· CUI Bifeng, CHENG Jin, HAO Shuai, et al.(77)
Studying Influence of Temperature and Material Parameters on $InAs_xSb_{1-x}$ Auger Recombination Life by Numerical Analysis ····· ZHANG Jingbo, ZHANG Yunqi, WANG Siwen, et al.(80)
Study on Composition of High-strength FBG with Carbon Fiber and Their Strain Sensing Characteristics ····· HOU Zhiyi, LI Yan, SHAO Hongfeng, et al.(85)
Photoluminescence Characteristics of InAs Quantum Dots with Low Temperature Cap Layers ····· WU Wei, ZHOU Yong, LIU Shangjun, et al.(89)
All-Optical Plasmonic Switches Based on Square Cavity Coupled Metal Waveguide Structure ····· XIAO Gongli, DOU Wanying, YANG Hongyan, et al.(93)
Performance Optimization and Analysis of NLDMOS Devices ····· LI Weijie, WANG Xing, WANG Yunfeng, et al.(99)
Sharpness Evaluation of Microscopic Detection Image for Micro Parts ····· LI Chengchao, YU Zhanjiang, LI Yiquan, et al.(103)
Multiple Endmember Extraction Algorithm Based on Pixel Purity Index ····· YANG Huadong, XU Nan, HAO Yongping(108)
Design of Real-time Vision System Applied in Flapping-wing Micro Air Vehicle ····· WANG Xinpeng, ZHANG Weiping, MOU Jiawang, et al.(114)
Automatic Absorption Peak Determination in LPG Detection System Based on TDLAS ····· ZHANG Ming, MIAO Yubin, ZHANG Shu(118)
Double Random-phase Encoding Encryption System Based on Random Radon Transform ····· PENG Kaifei, SHEN Xueju, HUANG Fuyu, et al.(123)
Research on Image Feature Tracking and Matching Algorithms in Intermittent Texture Environment ····· ZHAO Mingfu, CAO Libo, SONG Tao, et al.(128)
Infrared Image Enhancement Based on Non-uniform Ordered Histogram and Sharpening ····· WU Ting(135)
Joint Spatial Information and Improved Low-rank and Sparse Matrix Decomposition-based Anomaly Target Detection for Hyperspectral Imagery ····· ZHANG Yan, HUA Wenshen, HUANG Fuyv, et al.(141)
A Construction Method of Low Error Floor QC-LDPC Codes Based on The AP and Eliminating Elementary Trap Sets ····· YUAN Jianguo, WANG Hongsen, ZHANG Xirui, et al.(146)

动态综述

DOI：10.16818/j.issn1001－5868.2020.01.001

基于深度学习的目标检测技术的研究综述

罗　元，王薄宇，陈　旭

（重庆邮电大学 光电工程学院，重庆 400065）

摘　要： 深度学习已经成为机器视觉领域应用最为广泛的技术方法，基于深度学习的目标检测技术是当前的一项热门研究课题。文章首先对国内外目标检测技术的最新研究进展进行了梳理，并分析和总结了传统目标检测方法的优缺点；然后详细介绍了几种基于深度学习的目标检测技术及其优缺点；最后讨论了现阶段深度学习存在的问题和未来的发展方向。

关键词： 计算机视觉；深度学习；目标检测

中图分类号： TP391.41　　**文章编号：** 1001－5868(2020)01－0001－10

Research Progresses of Target Detection Technology Based on Deep Learning

LUO Yuan，WANG Boyu，CHEN Xu

（College of Optoelectronic Engin.，Chongqing University of Posts and Telecommun.，Chongqing 400065，CHN）

Abstract： Deep learning has become the most widely used technical method in the field of computer vision. Target detection technology based on deep learning is a hot research topic. In this paper，the latest research progress of target detection technology at home and abroad are reviewed，then the advantages and disadvantages of traditional target detection methods are analyzed and summarized. Then，several target detection techniques based on deep learning and their merits and demerits are introduced. Finally，the existing problems of deep learning and the development trends are discussed.

Key words： computer vision；deep learning；target detection

0　引言

目标检测是计算机视觉领域中一项重要的研究课题，该课题的主要研究内容是将图像或视频中的目标物体与其他无关的信息进行区分，即判断是否存在目标物体，若存在则输出目标的坐标位置。

目标检测问题本身定义如下：给定任意图像，确定是否存在来自预定义类别的语义对象的任何实例，如果存在，则返回空间位置和范围。对象是指可以看到和触摸的物质事物。虽然在很大程度上与对象类检测同义，但通用对象检测更加强调检测各种自然类别的方法，而不是对象实例或特定类别（例如：面部，行人或汽车）。

随着深度学习技术的不断发展和成熟，基于深度学习的目标检测技术逐渐成为研究热点，并且取得了一定的成果。Zhu等人[1]针对深度残差网络不能完整提取目标检测物体特征信息的问题，提出一种基于深度残差网络的全局与局部特征融合的目标检测方法。该方法首先训练深度残差网络（ResNet-G）来提取目标物体的全局特征信息，然后训练另一个深度残差网络（ResNet-L）来提取多级的局部特征信息。为了处理由图像分割引起的噪声，在ResNet-L中集成空间相干方法作为平滑步骤。最后采用贝叶斯对局部信息和全局信息进行融合。定量和定性实验结果表明，该方法在目标检测的准确性上优于同类算法。Zhang等人[2]针对SSD算法没有利用高层语义特征而导致不能准确检测小目标物体的问题，提出一种利用语义特征的目标检测方

收稿日期：2019－09－18.
基金项目：国家自然科学基金项目(51604056).

法。该方法首先将分割模块引入到网络中,从而得到注意机制掩模,进而丰富底层的语义特征信息,然后引入全局激活模块用于得到高层的语义特征信息。通过在 Voc2007 数据集上进行测试,发现在 TitanX 上的检测速度为 31.7 f/s,衡量检测精度的指标值(Mean Average Precision,MAP)为 79.7。Hamid 等人[3]针对交并比(Intersection-over-Union,IoU)无法直接优化检测任务中没有重叠部分的问题,提出一种广义交叉并集思想(Generalized Intersection over Union,GIoU)。GIoU 的优点在于它可以更好地反映两个对称物体之间的重叠区域,同时 GIoU 可以直接作为损失且在所有可能的情况下都有一个梯度,这样可以提高检测质量和收敛速度。实验结果证明,该算法能够提高检测的准确度和速度。

近几年已发表的目标检测方面的论文主要采用深度学习的方式进行目标检测,且取得了较传统机器学习方法更好的检测效果。基于深度学习的目标检测方法在实际生活中也取得了广泛的应用,如安防视频目标检测系统、医学病例图像分析、自动驾驶、遥感图像分析、无人机导航等。

1 传统目标检测的研究现状

传统目标检测的流程为:1)采用不同尺度的滑动窗口在图像中的不同位置选取候选区域;2)提取候选区域的特征信息;3)利用分类器对特征信息进行识别。

根据不同的应用背景可以灵活选取不同的目标检测方法,常用的传统目标检测方法分为四类:基于像素分析的方法、基于特征匹配的方法、基于频域的方法和基于识别的检测方法。

传统的目标检测方法采用人工手动提取特征信息,无法避免在此过程中存在特征信息不完整的问题,从而导致识别效果较差。具体来说,传统的目标检测算法还存在以下三点问题:

1)基于滑动窗口的目标区域检测方法耗时长、窗口存在冗余;

2)基于手工提取的特征信息在面对环境多样性的变化条件下鲁棒性较差;

3)对于大数据的视频或图片信息的处理能力较差,计算能力有限。

2 基于深度学习的目标检测的研究现状

2.1 深度学习的基本概念

深度学习又称深层神经网络(Deep Neural Network,DNN),它是一类通过多层非线性变换对高复杂性数据进行建模的算法合集[4]。多层即指神经网络的层数,而深度学习的深度是超过 8 层的神经网络,层数越多,深度越深。非线性是指处理实际应用中复杂的非线性可分问题,采用复杂的函数逼近,进而更加详尽地表征出数据的特性。深度学习的本质就是采用多个隐层的机器学习模型和海量的训练数据来尽可能充分地表征和学习到有用的特征信息,进而预测或识别出结果。

2.1.1 深度学习的特点

深度学习的优点是:减少了人为设计特征造成的提取信息不完整的缺点;在满足特点条件的应用背景下,以深度学习为计算架构的一些机器学习的应用,表现出较原有算法更好的识别效果和分类能力。

深度学习的缺点是:为了达到更好的识别效果和更高的估计精度,深度学习需要大量的数据支持和较长时间的学习过程。更好的软件编程技巧和更优质的硬件设备支持是做好和提升深度学习性能的前提条件。

2.1.2 深度学习的模型分类

常见的深度学习模型如表 1 所示。有以下几种分类方法:

1)深度学习按照结构层次可以分为基本模型和整体模型。基本模型就是采用基本的部件组合而成的模型,如全连接模型 FCM、自编码器 AE 等;整体模型就是由基本模型组合而成并包含两层或多层中间层的模型,如卷积神经网络 CNN、深度玻尔兹曼机 DBM 等。

2)深度学习按照训练过程和数学特性可以分为确定性模型和概率性模型。确定性深度模型是采用确定性描述,训练目标为最小化特定形式的损失函数,常见的模型有卷积神经网络 CNN、深度自编码器 DAE 等;概率性模型是采用概率化描述,训练方式是采用最大似然估计来最大化关于训练数据的似然函数,常见的模型有深度置信网络 DBN、深度玻尔兹曼机 DBM 等。

3)深度学习按照是否使用类别标签的数据可分

为无监督、监督式以及混合式深度模型。无监督深度模型是在训练过程中不使用带有类别标签的观测数据，即通过获取观测数据之间的高阶相关关系来进行计算和分析，常见的模型有深度置信网络 DBN、限制玻尔兹曼机 RBM 等；监督式深度模型是在训练过程中采用带有类别标签的观测数据，其中包含模式类别的信息，常见的模型有卷积神经网络 CNN；混合式深度模型是在训练过程中同时采用带类别标签的观测数据和不带类别标签的观测数据，模型中包含产生式模块和区别式模块，通过产生式模块辅助区别式模块，常见的模型是深度稀疏自编码器 DSAE。

4）深度学习按照网络连接的方式可以分为邻层连接、跨层连接和环状连接模型。邻层连接模型是相邻网络层之间的连接方式，该结构在模型优化上较为容易；跨层连接模型在相邻网络层和不同网络层之间均可连接，该结构可以对跨层之间的依赖关系建立模型；环状连接模型是任意层之间可以相互连接，形成环状结构。

目前，基于深度学习的目标检测算法可以分为两类：基于区域建议的目标检测算法和基于回归的目标检测算法。根据算法的实现步骤又可将它们分别称为“双阶段”目标检测算法和“单阶段”目标检测算法。

表 1　常见的深度学习模型

构成层级	模型名称	数学特性	训练方式	连接方式
基本模型	感知器模型(PM)[5]	确定性	监督式	邻层连接
	全连接模型(FCM)[6]	确定性	监督式	邻层连接
	自编码器(AE)[7]	确定性	无监督	邻层连接
	降噪自编码器(DAE)[8]	确定性	无监督	邻层连接
	稀疏自编码器(SAE)[9]	确定性	无监督	邻层连接
	限制玻尔兹曼机(RBM)[10]	概率性	无监督	邻层连接
	玻尔兹曼机(BM)[11]	概率性	无监督	环状连接
整体模型	卷积神经网络(CNN)[12]	确定性	监督式	邻层连接
	多层感知器(MLP)[13]	确定性	混合式	邻层连接
	深度自编码器(DAEs)[14]	确定性	混合式	邻层连接
	深度分解网络(DDN)[15]	确定性	混合式	跨层连接
	深度置信网络(DBN)[16]	概率性	混合式	邻层连接
	深度玻尔兹曼机(DBMs)[17]	概率性	混合式	邻层连接
	深度递归神经网络(DRNN)[18]	概率性	无监督	环状连接

2.2　“双阶段”目标检测算法

如图 1 所示，“双阶段”目标检测网络中通常包含 1～2 个卷积层，第一阶段是通过 EdgeBoxes，Selective Search(SS)等目标候选区域生成算法生成一组稀疏连接的候选区域，并将候选区域的结果送入 CNN 模块中，第二阶段将提取出的特征信息进行分类和回归。“双阶段”目标检测算法框架的技术演进过程是：R-CNN→SPP-net→Fast R-CNN→Faster R-CNN→Mask R-CNN。

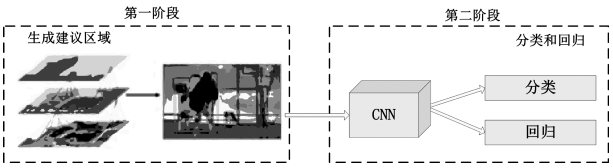


图 1　“双阶段”目标检测算法框架

2.2.1　R-CNN

R-CNN 网络的前身是 Hinton[10] 的学生 Krizhevsky[19] 在 2012 年提出的一种以他名字命名的深度卷积神经网络——AlexNet，该算法在同年的 ImageNet 比赛中获得了图像分类组最好成绩。借助 AlexNet 在 ImageNet 目标识别中表现出的强大的泛化能力，Girshick 等人[20] 针对如何利用深度神经网络来实现目标定位和如何利用小规模的数据集来训练更加强劲的大规模网络模型这两个问题，提出了一种将 Region Proposal 和 CNN 结合起来的 Region-CNN 算法，简称 R-CNN。如图 2 所示，针对利用深度神经网络实现目标定位的问题，R-CNN 网络首先采用 Selective Search 以自下而上的搜索方法抽取 2 000 个推荐区域，然后通过 CNN 前向传播计算出区域的特征信息，最后使用对每个类别训练出的 SVM 给整个特征向量中的每个类别单独打分，并采用非极大抑制(NMS)和边界框回归(Bounding Box Regression) 定位出最准确的位置区域。在 PASCAL VOC 2010-12 上的测试结果表明，mAP 值(验证模型好坏的评价值)相较于之前获得了一个更大的提升，从 35.1%提升至 53.7%，并且速度上也有了提升。针对如何利用小规模的数据集来训练更加强劲的大规模网络模型的问题，作者采用的策略是在辅助数据集(ILSVRC)上进行有监督的预训练，再在小数据集上针对特定问题进行调优。实验结果表明，该系统在 VOC2010 上的 mAP 值达到了 54%，远远超过高度优化的基于 HOG 的可变性部件模型(Deformable Part Model，DPM)，为面对标签数据少而需要训练一个大型的卷积网络提供了一条新的实用型思路。

R-CNN 存在以下缺点：1）计算量大，耗时长。由于采用 Selective Search 以自下而上的搜索方法抽取 2 000 个推荐区域，而每个框均需要特征提取和分类，导致此过程需要耗费大量计算时间；2）消耗大量内存。由于提取出来的特征信息需要存放到磁

盘中，故需要足够大的内存空间；3）由于存在全连接层，故需要固定输入尺寸，因此强行对输入尺寸进行变换和放缩操作会造成数据失真。

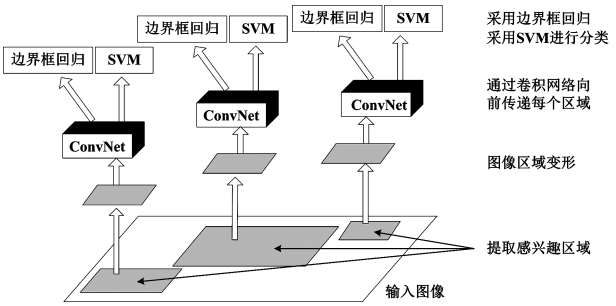


图 2 R-CNN 网络框架

2.2.2 SPP-net

SPP-net 的全称为空间金字塔池化（Spatial Pyramid Pooling，SPP），如图 3 所示，SPP-net[21] 基于 R-CNN 作出了两点改进：

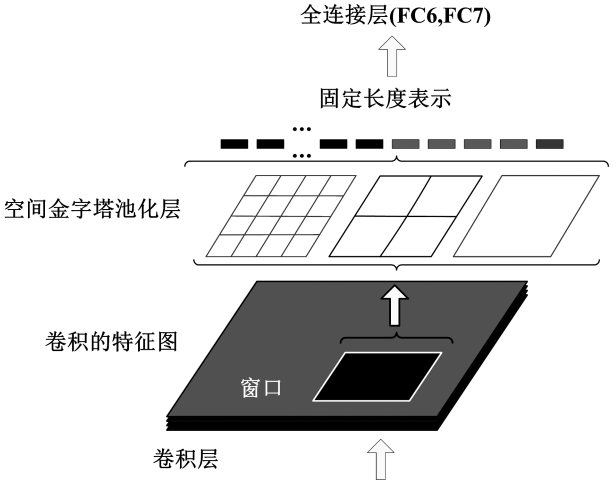


图 3 SPP-net 网络框架

1）针对需要固定输入图像尺寸而损失图像特征信息的问题，提出一种结合空间金字塔的方法来实现 CNN 对输入尺度的可变性。假设输入一张任意大小的照片，利用三种不同大小的刻度：4×4，2×2，1×1，对输入的图片进行划分，最后总共可得到 16+4+1=21 个块，从 21 个块中各提取一个特征，于是便得到一个 21 维的特征向量。空间金字塔最大池化的过程就是从 21 个图片块中分别计算每个块的最大值，从而得到一个输出神经元，即通过在卷积层和全连接层之间加入 SSP 层，可将任意大小的输入图像转换为一个固定大小的 n 维特征。

2）针对 R-CNN 存在繁琐的计算过程和耗时长的问题，提出一种只对原图提取一次卷积特征的方法。在 R-CNN 中是将 2 000 个候选框重新调整大小，然后分别作为 CNN 网络的输入，此过程需要耗费大量的时间和计算过程。SPP-net 是将输入图像做 1 次卷积特征提取从而得到特征图。该过程在速度和检测精度上均得到了较大的提升。

SPP-net 的缺点在于：网络微调只更新了全连接层，不能实现端到端的检测且检测精度还有进一步提升的空间。

2.2.3 Fast R-CNN

Fast R-CNN 是由 Grishick 等人[22] 针对 R-CNN 算法在候选区域进行特征提取时存在重复计算且耗时长的问题而提出的一种改进算法。如图 4 所示，该算法的原理是首先将原始图像输入到卷积网络中并得到特征图，并以共享卷积层的方式在第五个卷积层 conv5 得到每个候选窗的特征，以投影的方式得到候选区域对应的感兴趣区域（ROI），然后通过 ROI 池化层得到一个固定尺度的特征向量，最后得到两个输出向量，分别是一个分类器和一个回归器。

Fast R-CNN 的优点主要包括：1）不需要磁盘空间进行特征信息的存储；2）mAP 值相较于 R-CNN 和 SPP-net 均有所提高且计算时间减少，实验结果如表 2 所示；3）训练过程采用 L1 的多任务损失函数，实现单步骤完成；4）将边界框回归（Boundingbox Regression）放入神经网络内部，并与区域分类合并成一个多任务（Multi-Task）模型，从而实现了共享卷积特征，达到了较为理想的检测效果。但是其不足之处在于算法中采用了选择性搜索来找出所有候选框，该过程较为耗时。

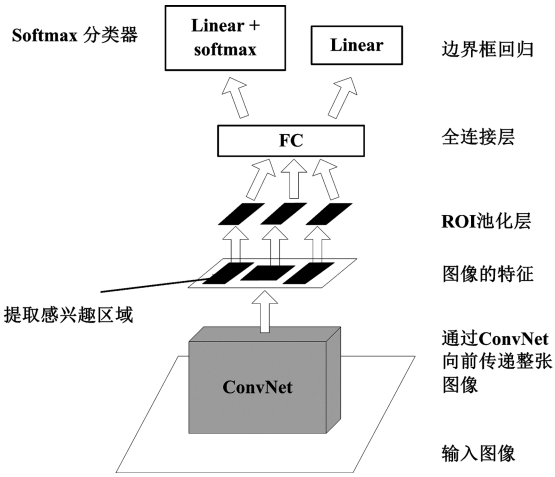


图 4 Fast R-CNN 网络架构

本文所有实验均在 Ubuntu 系统上完成测试，实验电脑配置为：XEON E5 2.4 GHz CPU，64 G 内

存，GTX1080ti GPU。测试的数据集分别为 VOC2007，VOC2010 和 VOC2012。测试项目分别是 mAP 值(验证模型的好坏的评价值)、训练时间和测试时间。

表 2　各类算法实验数据对比

算法	mAP/%			训练时间/h	测试时间/s
	VOC 2007	VOC 2010	VOC 2012		
R-CNN	58.5	53.7	53.3	84	47
SPP-NET	60.9	58.2	59.1	25	2.3
Fast R-CNN	66.9	66.1	65.7	9.5	0.32

2.2.4　Faster R-CNN

Faster R-CNN[23] 是在 Fast R-CNN 的网络结构基础上改进而来的。如图 5 所示，它由两级组成，第一级是区域建议网络(RPN 网络)，用于生成候选边框；第二级为 Fast R-CNN 网络，用于特征提取和分类回归。Faster R-CNN 的结构包括以下 4 个部分：1)卷积层，作用为提取图像的特征信息并输出特征映射图；2) RPN 网络(Region Proposal Network)，也称为全卷积网络，作用为推荐候选区域，输出多个候选区域；3)感兴趣区域池化层(ROI Pooling)，作用为将不同大小的输入图像转换为固定长度的输出；4)分类与回归，作用是输出候选框区域所属的类别和候选区域的精确定位。其中，RPN 网络的引入是 Faster R-CNN 能够减少在候选区域耗时的关键，同时也是区别于 Fast R-CNN 最主要的部分[24]。

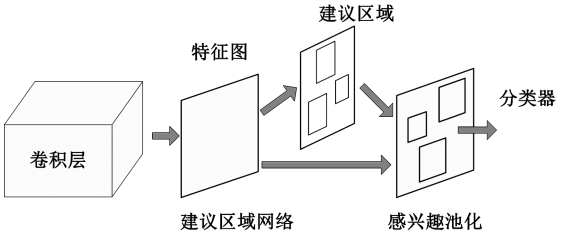


图 5　Faster R-CNN 网络架构图

Faster R-CNN 的实现原理流程：首先将图像输入卷积神经网络中，通过特征提取得到特征映射图。然后 RPN 网络在分类层中完成对区域的前景和背景的判断，在分类层中采用 softmax 损失函数和 L1 损失函数计算出预测区域的坐标点信息。最后在感兴趣区域(ROI)池化层完成将候选区域映射为一个固定尺度的特征向量的过程，并对候选区域进行分类识别。

Faster R-CNN 的主要特点在于：在继承 Fast R-CNN 网络优点的基础上，将之前的全连接层替换为全卷积层，代替了选择性搜索，从而较大地提高了检测速度，实验对比数据如表 3 所示。

表 3　各类算法实验数据对比

算法	mAP/%			训练时间/h	测试时间/s
	VOC 2007	VOC 2010	VOC 2012		
R-CNN	58.5	53.7	53.3	84	47
SPP-NET	60.9	58.2	59.1	25	2.3
Fast R-CNN	66.9	66.1	65.7	9.5	0.32
Faster R-CNN	78.8	77.4	75.9	4.3	0.2

2.2.5　Mask R-CNN

如图 6 所示，Mask R-CNN[25] 由两级组成，第一级结构与 Faster R-CNN 相同，用于扫描图像并生成建议区域，第二级分类建议并生成边界框和掩码。Mask R-CNN 的结构具体分为 5 个部分：特征提取网络、特征组合网络、区域建议网络、ROI Align 和功能性网络。特征提取网络的作用是对输入图像进行特征提取并生成特征映射图。特征组合网络重新组合特征映射图，从而表现出不同的特征信息。区域建议网络采用锚定(Anchor)技术来标定出图像中的候选框位置，利用非极大抑制(NMS)过滤出精准的候选框。ROI Align 是一个引入双线性插值的池化过程，目的在于使离散的池化过程变成连续的过程，从而弥补感兴趣区域池化粗糙量化的问题，提高了定位候选框的准确性，实验数据如表 4 所示。

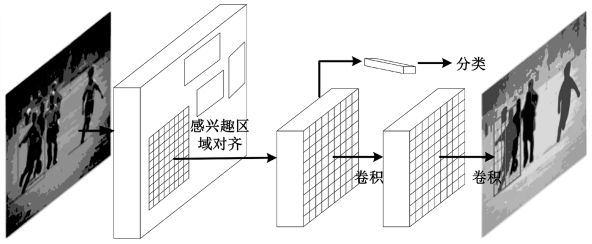


图 6　Mask R-CNN 网络架构图

表 4　各类算法实验数据对比

算法	mAP/%			训练时间/h	测试时间/s
	VOC 2007	VOC 2010	VOC 2012		
R-CNN	58.5	53.7	53.3	84	47
SPP-NET	60.9	58.2	59.1	25	2.3
Fast R-CNN	66.9	66.1	65.7	9.5	0.32
Faster R-CNN	78.8	77.4	75.9	4.3	0.2
Mask R-CNN	79.7	80.3	81.4	2.2	0.1

Mask R-CNN 网络的优点在于：它在 Faster R-CNN 网络的基础上增加了目标掩码作为输出量，掩

码将一个对象的空间布局进行编码，从而使得对目标的空间布局提取更精细。其缺点在于：Mask R-CNN 比 Faster R-CNN 运行速度慢，为 5 f/s。

2.3 “单阶段”法

2.3.1 YOLO

YOLO(You Only Look Once)[26]是一种采用回归策略的端到端目标检测算法。该算法舍弃了生成建议框，而仅需要“看一眼”即可得到回归位置信息和类别信息。如图 7 所示，YOLO 模型的结构主要包括 18 个卷积层、2 个全连接层以及 6 个池化层。卷积层的作用是提取输入图像的特征信息；全连接层的作用是计算出物体位置和类别的概率值；池化层的作用是减少原始图像的像素值。YOLO 模型采用回归分析的方法计算出多个滑动窗口的位置和置信度。

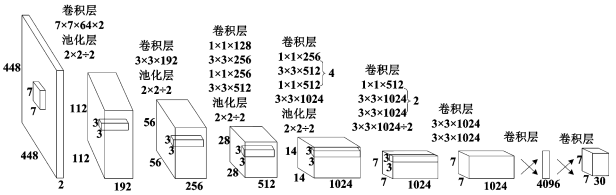


图 7 YOLO 网络架构图

YOLO 的实现原理是：假设将一幅图像划分为 $S\times S$ 个网格，如果检测目标的中心落在网格中，则网格需要对检测目标所在的具体位置进行估计。若每个网格需要预测 N 个边界框，则输出结果为每个边界框的位置坐标（x，y，w，h）和置信度 Confidence。

$$Confidence = Pr(\text{Object}) \times IOU_{\text{pred}}^{\text{truth}}$$

当检测目标落在划分的网格内时，$Pr(\text{Object})=1$，否则取 0。第二项为预测的检测框与实际的检测框之间的 IOU 值。假设每一个网格需要预测一个类别信息，记为 M 类，则网络最终输出的张量为 $S\times S\times(5\times N+M)$。以 PASCAL VOC 数据集为例，图像输入为 448×448，取 $S=7$，$B=2$，一共有 20 个类别（$C=20$），则输出的张量为 7×7×30。

YOLO 的缺点在于：1）当检测到目标物体之间的距离相隔较近且物体较小时，检测效果较差；2）若目标物体出现形变或其他情况时，该模型的泛化能力较弱。

2.3.2 YOLO v2

YOLO v2[27]是基于 YOLO 的优点，以进一步提升目标物体的定位准确度和分类准确度为目标提出的。具体的改进点如下：

1）引入批量归一化（Batch Normalization），在每一个卷积层中增加批量归一化，目的在于消除其他形式对正则化的依赖，并且提高网络的收敛性。实验结果表明，mAP 提高了 2%。

2）引入高分辨率分类器（High Resolution Classifier），大多数分类器中运行的图片大小均小于 256×256，而 YOLO v2 的分类网络以 448×448 的分辨率在网络中进行微调，实验结果表明，mAP 提升了 4%。

3）去掉了 YOLO 中的一个池化层，目的在于能够通过卷积层输出更高分辨率的特征信息。

4）采用多尺度训练，YOLO v2 会在迭代过程中改变网络参数，网络会随机选择一个新的图片尺寸，由于下采样参数为 32，故网络会以 32 的倍数{320，352，…，608}来改变尺寸。实验结果表明，当输入图片尺寸较小时，系统运行较快，输入图片尺寸较大时，检测精度较高。

5）引入锚定 Anchor 预测边界框坐标。

6）采用 K-means 的方式对训练集的 Bounding Boxes 做聚类，试图找到合适的 Anchor Box。

2.3.3 YOLO v3

如图 8 所示，YOLO v3[28]主要的改进有：调整了网络结构；利用多尺度特征进行目标检测；目标分类采用 Logistic 取代了 softmax。

1）Darknet-53 网络结构：YOLO v3 采用了 Darknet-53 网络结构（含有 53 个卷积层），借鉴了残差网络（Residual Network）的做法，在一些层之间设置了快捷链路（Shortcut Connections）。

2）利用多尺度特征进行目标检测：YOLO v3 采用 3 个不同尺度的特征图来完成目标检测。第一种尺度是在原基础网络后增加卷积层并输出边框信息；第二种尺度是从第一种尺度中的倒数第二层的卷积层上采样，再与最后一个 16×16 大小的特征图相加，最后通过多个卷积后输出边界框信息，此时的尺度是第一种尺度的 2 倍。同理，第三种尺度与第二种尺度的处理方式一样，使用 32×32 大小的特征图，第三种尺寸是第二种尺度的 2 倍。

3）采用 Logistic 对目标进行分类：在之前的网络中采用 softmax 层对图像进行分类，但由于实际场景中存在多个类别，softmax 无法分辨多种类别。Logistic 分类可以进行多标签分类，从而提高分类

的准确性。

YOLO v3 的优点在于：它能够较准确地检测出较小的目标，速度相较前两代也更快。但随着 IoU 的增大，YOLO v3 识别物体位置精度会有所下降。

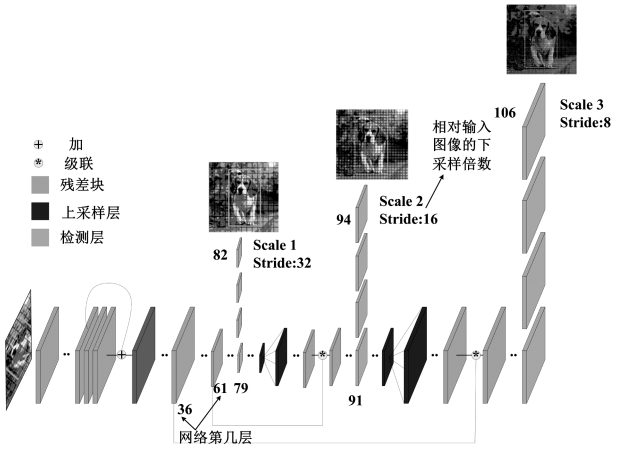


图 8　YOLO v3 网络结构图

2.3.4　SSD

如图 9 所示，SSD（Single Shot MultiBox Detector）算法[29]的实现流程为：在传统的基础网络（如 VGG）后增加了 5 个卷积层，用于将特征图的大小依次减少。对 5 个特征图的输入分别采用 2 个不同的 3×3 的卷积核进行卷积操作，每个先验框（Default Box）生成 21 个类别的置信值和 4 个坐标值(x,y,w,h)，最后将 5 个特征图上的结果合并后送入损失函数。SSD 算法的特点在于：1）采用多尺度特征图用于检测。对于较小的检测目标可用较大的特征图对其进行检测，反之较大的检测目标可用较小的特征图进行检测，从而提高了检测精度；2）SSD 提出了端到端的训练网络，保证检测精度的同时大幅度提高了检测速度，且对低分辨率输入图像的处理效果很好；3）设置了先验框。每个单元设置尺度或者长宽比不同的先验框，预测的边界框（Bounding Boxes）是以这些先验框为基准的，在一定程度上降低了训练难度；4）利用卷积进行检测。SSD 采用卷积对不同的特征图来提取检测结果。对输入 300×300 的图像，在 Voc2007 数据集上进行测试，能够达到 58 f/s（Titan X 的 GPU），mAP 为 72.1%，实验数据如表 5 所示。

SSD 算法的缺点在于：1）由于不同层的特征图均是独立作为分类网络的输入，在检测时相同的检测目标会被不同的检测框检测出来；2）对于小尺寸检测目标的检测效果还有待进一步提高。

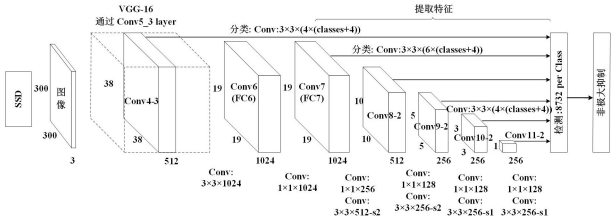


图 9　SSD 网络架构图

表 5　各类算法实验数据对比

算法	mAP/%			训练时间/h	测试时间/s
	VOC 2007	VOC 2010	VOC 2012		
YOLO	63.4	60.3	57.9	3	2
SSD	76.8	75.1	74.9	2.5	1.5
YOLO v2	78.6	76.5	73.4	2	1
YOLO v3	79.6	79.8	80.1	1	0.05

3　目标检测数据集分类

3.1　通用目标检测的常用数据集

目标检测数据集常用于研究测试算法的性能或用于图像识别竞赛。通用目标检测的常用数据集主要包括：PASCAL VOC，ImageNet，Open-Images 等数据集。

3.1.1　PASCAL VOC

PASCAL VOC[30]（PASCAL Visual Object Classification）数据集是目标识别领域中最经典的数据集之一，常用于目标检测、分类、分割等。如图 10 所示，PASCAL VOC 数据集包含 200 类可视对象，包含 517 000 张图像和 534 000 个注释对象。该数据集包括训练集、验证集和测试集，是目标检测领域中重要的数据集之一。

图 10　PASCAL VOC 数据集

3.1.2　ImageNet

ImageNet[31]是目前世界上图像识别领域最大的数据集。如图 11 所示，ImageNet 数据集总共有 1 500 万张图片，分为 2.2 万类，且每一张图片都经过人工的严格筛选与标记。由于其数据量相当大，类别较多，检测难度也是同类数据集中最大的。

图 11 ImageNet 数据集

3.1.3 Open Images

Open Images[32]是谷歌公司发布的一个由 900 万张图片组成的数据集，这些图像被标注为图像级标签和对象边界框。如图 12 所示，对于对象检测任务，该数据集由 191 万张图像组成，在 600 个对象类别上有 1 544 万个带注释的边界框，该数据集也是现有最大的包含对象位置注释的数据集，包含有多个对象的复杂场景(平均每个图像有 8.4 个标记)。

图 12 Open Images 数据集

4 基于深度学习的目标检测应用场景

4.1 行人检测

汽车驾驶辅助系统和智能监控是行人检测的两个最主要的应用场景。由于在实际的生活场景中存在复杂的姿态变化、光照不均匀、局部范围的遮挡等因素，传统的机器学习并不能达到理想的效果，取而代之的基于深度学习的行人检测方法发挥出了其独特的优势。Wang 等人[33]针对监控视频中行人经常在真实场景中聚集在一起并相互遮挡的问题，提出了一种基于改进的边界框回归损失目标的检测方法。实验证明，该方法通过排斥损失训练的检测器优于目前最先进的方法，并且在闭塞情况下有显著改善。Liu 等人[34]在 CVPR2019 大会上提出了一种高级语义特征的行人检测方法，将行人检测中行人的中心点坐标和尺度(高度)转化为高级语义特征信息，首先输入图像通过卷积网络，然后分别计算中心热图(Center Heatmap)和规模大小图(Scale Map)，最后得到中心点坐标和宽高，即检测出了行人。实验证明，该方法能够准确检测出行人。

4.2 医学图像检测

不同于传统的自然图像检测，医学图像检测是针对某一种或某一类特定的类别(肿瘤、结核等)进行检测，特点是检测目标物体细粒度的差别，其难点在于：1)目标物体的形状、尺寸、位置变化大；2)不同时期、环境下目标物体形态各异；3)病变组织和非病变组织外形差别较小。医学图像检测的主要过程为：首先对目标物体所在区域进行分割，然后提取候选区域并完成特征提取，最后进行分类。Quellec 等人[35]利用图像级标签监督训练基于 CNN 的病变检测器，使用灵敏度分析方法来检测在图像级预测中起作用的像素，获得图像大小的热点图，通过在训练 ConvNets 的同时增强热点图的稀疏性，避免热点图受伪影的影响。该方案在 Kaggle 彩色眼底图像检测糖尿病性视网膜病变挑战赛数据集上获得了很好的成绩。Li 等人[36]针对医学图像中存在数据分布差异且域适应能力较差的问题，提出一种 CLU-CNN 域适应框架，它没有特定域适应训练且在小数据集上有良好的扩展性和准确的定位能力。

4.3 人脸检测

人脸检测是完成人脸识别任务中第一个重要环节，人脸检测的效果奠定了人脸识别的基础。Li 等人[37]针对人脸中多变的尺度、姿态、遮挡、表情、光照等问题导致检测率较低的问题，提出一种双分支人脸检测器。该文的创新点在于分别处理了人脸检测方向的三个关键点：更好的特征学习、渐进式的损失函数设计以及基于锚点分配的数据扩充。实验在两个著名的人脸检测数据集 WIDER FACE 和 FDDB 的 5 个评测维度上均刷新了当时的世界纪录。Niu 等人[38]针对目前的算法受制于人脸活动部件微弱的激活强度以及不同个体之间差异性的问题，提出一种针对人脸活动部件的局部关系学习算法，通过充分利用不同人脸局部区域之间的关系，进一步提高人脸不同区域的局部特征的感知能力，并针对人脸形状信息的正则项来剔除在人脸活动部件检测过程中人脸的形状信息的影响，进而获得更加有判别力和泛化能力的人脸活动部件检测器。通过在人脸活动部件检测数据集 BP4D 和 DISFA 上进行实验，结果表明该方法优于其他最好算法。

5 基于深度学习目标检测存在的问题与发展方向

结合近两年的计算机视觉会议的相关论文可以

看出，对于目标检测而言，目前主要存在以下几方面的问题：1）对于小目标物体的检测（如小于 30 像素的目标物体）、遮挡面积较大的目标以及区分图像中与目标物体外形相似的非目标物体等问题需要在今后的研究中继续加强；2）实时性检测与处理。对于自动驾驶或汽车辅助驾驶等对实时处理能力要求较高的应用场景，进一步提高目标检测的计算速度和准确度是至关重要的；3）提高小数据量训练的检测效果。目前迁移学习的实现策略是先在现有大数据中进行训练，然后再将模型进行微调。虽然此方法能够实现，但检测精度和速度还有待进一步提高；4）目前基于深度学习的目标检测所涉及的行业领域越来越多，很难获取大量的监督数据或数据的标准成本过高，进而导致缺少用于网络训练的样本数据。

针对上述问题及近几年的研究趋势，本文对基于深度学习的目标检测未来发展方向做出讨论：

1）新的网络架构。提高小目标物体检测、遮挡、形态相似等情况下的检测准确性应该从深度学习网络结构本身出发，寻求新的网络架构模式或者将深度学习与传统方法进行结合，都有可能提高目标检测的准确性；2）硬件和软件的升级。深度学习涉及的数据量大，网络结构复杂等特点都是影响实时性检测的因素，在保证检测准确度的前提下，硬件模块的性能提升、代码运行效率的优化、计算共享的实现都值得进一步探索；3）增加更全面的数据集。对于小数据量而言，采用人工手动标注的方式能够保证标注的质量和数量，对后续网络的训练和测试奠定了较好的特征学习基础。对于大数据量而言，可以采用平行视觉方法来人工模拟实际场景和表示复杂挑战的实际场景，采用大规模多样性的虚拟图像和视频数据集并自动生成标注，然后通过计算实验进行视觉算法的设计与评估，最后进行在线优化；4）数据的差异性是影响深度学习效果的重要因素之一。在关注数据量的同时，也应该关注数据的质，即数据类内强调共性，数据类间强调个性。

6　总结

深度学习凭借其精准、快速以及端到端实现的优点而迅速成为目标检测领域中最热门的一种研究方法。虽然传统的目标检测方法能够完成日常的目标检测任务，但更准确、更快速才是人们追求的目标。随着软件和硬件条件的不断优化和完善，基于深度学习的目标检测方法将成为未来的主流研究趋势。本文首先综述了深度学习的研究现状，并展示了基于深度学习的目标检测算法的发展历史。然后对传统的目标检测算法进行了分类和研究，并指出其缺点所在。之后对深度学习的基本概念和主流的算法进行了详细的介绍，并分析了它们各自的优缺点，接着介绍了常用的目标检测数据集。最后分析和阐述了基于深度学习的目标检测存在的问题与发展方向。

参考文献：

[1] Zhu D, Luo Y, Dai L, et al. Salient object detection via a local and global method based on deep residual network[J]. J. of Visual Commun. and Image Represent., 2018(54): 1-9.

[2] Zhang Z, Qiao S, Xie C, et al. Single-shot object detection with enriched semantics[C]// The IEEE Conf. on Computer Vision and Pattern Recognition, 2018: 3610-3621.

[3] Hamid R, Nathan T, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]// The IEEE Conf. on Computer Vision and Pattern Recognition, 2019: 658-670.

[4] Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE Signal Proc. Magazine, 2012, 29(6): 82-97.

[5] Rosenblatt F. The perception: A probabilistic model for information storage and organization in the brain [J]. Psychological Rev., 1958, 65(6): 386-408.

[6] Sun D, Wulff J, Sudderth E B, et al. A fully-connected layered model of foreground and background flow[C]// 2013 IEEE Conf. on Computer Vision and Pattern Recognition, 2013: 2451-2458.

[7] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]// Machine Learning, Proceedings of the Twenty-Fifth International Conference (ICML 2008), 2008: 1096-1103.

[8] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. J. of Machine Learning Research, 2010, 11(12): 3371-3408.

[9] Jiang X, Zhang Y, Zhang W, et al. A novel sparse auto-encoder for deep unsupervised learning[C]// Sixth Inter. Conf. on Adv. Computational Intelligence, 2013: 256-261.

[10] Hinton G E. A practical guide to training restricted Boltzmann machines[J]. Momentum, 2010, 9(1): 926-947.

[11] Hinton G. Boltzmann machine[J]. Encyclopedia of Machine Learning, 2007, 2(5): 119-129.

[12] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[J]. Adv. in Neural Information Processing Systems, 2012, 25(2): 1097-1105.

[13] Bernier J L, Ortega J, Rodriguez M M, et al. An Accurate Measure for Multilayer Perception Tolerance to Additive Weight Deviations[M]. Engin. Applications of Bio-Inspired Artificial Neural Networks, Berlin: Springer, 1999: 121-130.

[14] Phan N H, Wang Y, Wu X T, et al. Differential privacy preservation for deep auto-encoders: an application of human behavior prediction[C]// 30th AAAI Conf. on Artificial Intelligence, 2016: 179-185.

[15] Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE Signal Proc. Magazine, 2012, 29(6): 82-97.

[16] Salakhutdinov R, Hinton G. Using deep belief nets to learn covariance kernels for Gaussian processes[C]// Inter. Conf. on Neural Information Proc. Syst., 2008(20): 1347-1355.

[17] Cho K H, Raiko T, Ilin A. Gaussian-Bernoulli deep Boltzmann machine[C]// IEEE The 2013 Inter. Joint Conf. on Neural Networks (IJCNN), 2013(10): 561-570.

[18] Huang P S, Kim M, Hasegawa-Johnson M, et al. Joint optimization of masks and deep recurrent neural networks for monaural source separation[J]. IEEE/ACM Trans. on Audio Speech & Language Proc., 2015, 23(12): 2136-2147.

[19] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// Inter. Conf. on Neural Information Proc. Sys., 2012(1): 1097-1105.

[20] Girshick R, Donahue J, Darrelland T, et al. Rich feature hierarchies for object detection and semantic segmentation [C]// 2014 IEEE Conf. on Computer Vision and Pattern Recognition, 2014: 1-21.

[21] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Trans. on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904-1916.

[22] Girshick R. Fast R-CNN[J]. Computer Science, 2015(4): 169-178.

[23] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Trans. on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.

[24] 郭 毓, 苏鹏飞, 吴益飞. 基于Faster R-CNN的机器人目标检测及空间定位[J]. 华中科技大学学报(自然科学版), 2018, 46(12): 60-64.
Guo Yu, Su Pengfei, Wu Yifei. Object detection and location of robot based on Faster R-CNN[J]. J. of Huazhong University of Science and Technol. (Nature Science Edi.), 2018, 46(12): 60-64.

[25] He K, Gkioxari G, Dollar P, et al. Mask R-CNN[J]. IEEE Trans. on Pattern Analysis & Machine Intelligence, 2018 (42): 386-397.

[26] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[J]. arXiv e-prints, 2015 (6): 2640-2650.

[27] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C]// IEEE 2017 IEEE Conf. Computer Vision and Pattern Recognition(CVPR), 2017: 6517-6525.

[28] Redmon J, Farhadi A. YOLOv3: An incremental improvement[C]// 2018 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2018: 2767-2773.

[29] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[J]. ECCV 2016: Computer Vision, 2016: 21-37.

[30] Vicente S, Carreira J, Agapito L, et al. Reconstructing PASCAL VOC[J]. Computer Vision & Pattern Recognition, 2014, 10(5): 111-122.

[31] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge[J]. Inter. J. of Computer Vision, 2015, 115(3): 211-252.

[32] Kuznetsova A, Rom H, Alldrin N, et al. The open images dataset V4: Unified image classification, object detection and visual relationship detection at scale[J]. arXiv:1811.00982, 2018, 18(4): 111-119.

[33] Wang Xinlong, Xiao Tete, Jiang Yuning, et al. Repulsion loss: detecting pedestrians in a crowd[C]// 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2018, 1 (4): 7774-7783.

[34] Liu Wei, Liao Shengcai, Ren Weiqiang, et al. High-level semantic feature detection: a new perspective for pedestrian detection[C]// 2019 IEEE Conf. on Computer Vision and Pattern Recognition, 2018: 2167-2173.

[35] Quellec G, Charriere K, Boudi Y, et al. Deep image mining for diabetic retinopathy screening [J]. Medical Image Analysis, 2017, 39: 178-193.

[36] Li Zhuolin, Dong Minghui, Wen Shiping, et al. CLU-CNNs: Object detection for medical images[J]. Neurocomputing, 2019, 350: 53-59.

[37] Li J, Wang Y, Wang C, et al. DSFD: Dual shot face detector[C]// 2019 IEEE Conf. on Computer Vision and Pattern Recognition, 2019: 1215-1224.

[38] Niu Xuesong, Han Hu, Yang Songfan, et al. Relationship learning with person-specific regularization for facial action unit detection[C]// 2019 IEEE Conf. on Computer Vision and Pattern Recognition, 2019: 411-420.

作者简介:

罗 元(1972—),女,1993年毕业于浙江大学光学仪器专业,2003年获重庆大学仪器科学与技术专业博士学位,教授,研究领域为数字图像处理、智能信号处理。

E-mail: luoyuan@cqupt.edu.cn

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.002

一种低交叉轴灵敏度微米光栅加速度计设计

高　山，冯丽爽，周　震，邓柯柯

（北京航空航天大学 仪器科学与光电工程学院 光学工程系，北京 100089）

摘　要： 微米光栅加速度计具有高精度、低功耗、小体积等优点，在惯性导航、油气勘探、地震检波等领域有广泛的应用前景。文章针对传统微米光栅加速度计受到交叉轴串扰导致精度下降的问题，提出了一种双层悬臂梁结构微米光栅加速度计敏感头。利用ANSYS有限元分析可知，所设计的双层悬梁结构敏感头能有效抑制交叉轴串扰，同时模态分离较传统结构提升13.5倍，有效提高了敏感头抗高阶机械模态扰动的能力。

关键词： 光栅加速度计；悬臂梁；交叉轴灵敏度；有限元仿真

中图分类号： TP212　**文章编号：** 1001－5868(2020)01－0011－04

Design of Micro-grating Accelerometer with Low Cross-axis Sensitvity

GAO Shan, FENG Lishuang, ZHOU Zhen, DENG Keke

(Dept. of Optical Engin., School of Instrumentation and Optoelectronic Engin., Beihang University, Beijing 100089, CHN)

Abstract: Micro-grating accelerometers have the advantages of high precision, low power consumption and small volume. It has wide application prospects in the fields of inertial navigation, oil and gas exploration, and seismic detection. In order to suppress cross-axis interference, a micro-grating accelerometer with double-layer cantilever structure is proposed. The ANSYS finite element analysis shows that the designed double-layer beam structure can effectively suppress cross-axis interference. The modal separation ratio is increased by 13.5 times compared with traditional structures, which effectively improves the resistance to high-order mechanical modal disturbances.

Key words: grating accelerometer; cantilever; cross-axis sensitivity; finite element analysis

0　引言

加速度计是惯性测量系统重要的组成部件，在惯性导航、地震检波、油气勘探等领域有着广泛的应用[1-4]。基于微米光栅检测原理的加速度计，具有高精度、小体积、低功耗、易集成等优点，是高性能加速度计的一个发展方向[4-7]。

微米光栅加速度计通过微米光栅的光学效应来检测质量块在加速度输入下的位移变化，具有极高的位移灵敏度，理论上具有极高的加速度检测分辨率。本课题组已完成微米光栅加速度计集成芯片的设计与加工，并实现0.35 mg的输出稳定性[8-9]。然而，由于传统微米光栅加速度计的非对称设计，当加速度计受到沿交叉轴方向的加速度时，参与光学传感的光栅和反射面会产生一定的偏角，这将严重影响微米光栅加速度计的光学灵敏度，导致加速度计精度下降[10]。微米光栅加速度计的交叉轴响应对实现加速度计的高精度是十分不利的，因此有必要提出一种新结构，以降低微米光栅加速度计的交叉轴灵敏度。

本文提出了一种双层悬臂梁结构微米光栅加速度计，通过有限元分析方法对所设计结构的机械特性进行了分析，仿真结果证明所设计结构能有效降低加速度计的交叉轴灵敏度，并显著提高其抗高阶

收稿日期：2019－09－17.

基金项目：国防基础科研计划项目(JCKY201601C006).

机械模态扰动的能力，最后提出了一套可行的加工方法。

1 传感机理

微米光栅加速度计由悬臂梁、质量块、衍射光栅、玻璃基底、光源及探测器组成，其结构示意图如图 1 所示。平行激光光源从加速度计敏感头的光栅一侧入射，一部分光被衍射光栅的上表面反射，并形成一组衍射光束；另一部分光通过光栅，被质量块的上表面反射，再经过光栅形成另一组衍射光束。两组衍射光在空间中发生干涉。当光栅和质量块上表面间的气隙间距 d 发生变化时，两组衍射光的干涉相位随之改变，使干涉光强度发生变化。

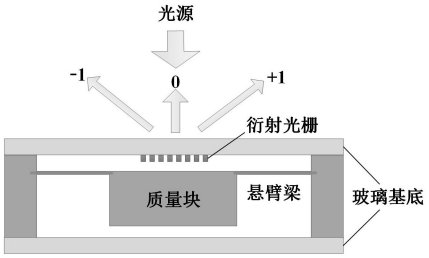


图 1 微米光栅加速度计结构示意图

基于标量衍射理论，0 级光和±1 级光的强度可以表示为[11-12]

$$I_0 = I_{in}\cos^2\left(\frac{2\pi d}{\lambda}\right)$$
$$I_{\pm 1} = \frac{4I_{in}}{\pi^2}\sin^2\left(\frac{2\pi d}{\lambda}\right) \tag{1}$$

其中，I_0，$I_{\pm 1}$ 和 I_{in} 分别为 0 级次、±1 级次和入射光的强度，λ 为光源的工作波长。

当沿敏感轴方向有加速度时，质量块在惯性力作用下，会产生与加速度成正比的位移，其大小为

$$\Delta d = \frac{ma}{k} \tag{2}$$

其中，m 为质量块的质量，a 为敏感头方向的输入加速度，k 为悬臂梁结构的刚度。

通过检测干涉光强的大小，可推算出气隙间距 d，从而测算出当前输入敏感头的加速度。

2 理论分析

2.1 结构设计

为降低悬臂梁结构的交叉轴灵敏度，提出一种双层悬臂梁结构的敏感头，其结构如图 2 所示。加速度敏感头由三层结构组成：顶层为玻璃片，玻璃片底面有电极板；中间层为硅片，包括质量块和悬臂梁结构；底层为玻璃片，玻璃片顶面有电极板，电极板中间刻蚀光栅。

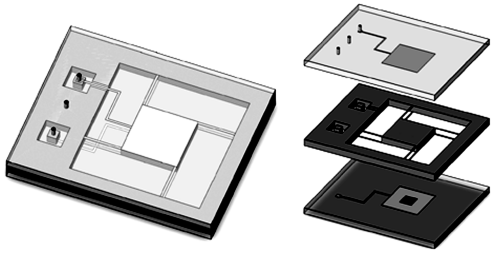


图 2 双层悬臂梁敏感头结构示意图

根据经典力学原理，所设计八臂悬臂梁结构的刚度为

$$k = \frac{8Ewh^3}{l^3} \tag{3}$$

其中，E 为悬臂梁材料的弹性模量，w 为悬臂梁的宽度，h 为悬臂梁的厚度，l 为悬臂梁的长度。设 ρ 为质量块的密度，L 为质量块的宽度，H 为质量块的厚度，则敏感头的位移灵敏度可以表示为

$$s = \frac{\Delta d}{a} = \frac{\rho L^2 H l^3}{8Ewh^3} \tag{4}$$

悬臂梁和质量块的结构及材料参数如表 1 所示。则悬臂梁的理论刚度和位移灵敏度分别为 26.96 N/m 和 3.992 μm/G。

表 1 敏感头结构及材料参数

参数名称	参数大小
L/μm	3 000
H/μm	524
l/μm	2 000
w/μm	120
h/μm	12
E/GPa	130
硅泊松比 μ	0.278
ρ/(kg/m^3)	2 329

2.2 有限元分析

使用 ANSYS 有限元仿真软件建立所设计敏感头的仿真模型，计算挠性结构在 1 个重力加速度下的位移变化，其仿真结果如图 3 所示。质量块在 1 个重力加速度下位移变化了 3.98 μm，则悬臂梁的刚度大小为 27.04 N/m，与式(3)计算结果有良好的一致性。

为对比双层悬臂梁结构与单层结构的交叉轴响应特性，使用 ANSYS 有限元仿真软件分别建立两

种结构的有限元模型，计算其交叉轴方向受到 1 个重力加速度情况下的机械响应，其结果如图 4 所示。

双层悬臂梁结构交叉轴方向受到 1 个重力加速度后的机械响应如图 4(a)所示，质量块仅在沿交叉轴方向有一个微小位移，沿敏感头方向基本没有位移变化。单层悬臂梁结构交叉轴方向受到 1 个重力加速度后的机械响应如图 4(b)所示，由于质量块的重心和悬臂梁结构的几何中心不重合，在质量块受到交叉轴方向的惯性力时，会产生一个扭转力矩，使质量块发生偏转，这将导致加速度计衍射光栅和反射面平行度变差。从仿真结果可得，单层悬臂梁结构交叉轴方向受到 1 个重力加速度后，光栅和反射面间的夹角变为 0.028 2°，这个角度将会使敏感头的光学效率降低近一半，这将严重影响微米光栅加速度计的光学灵敏度，导致加速度计精度下降[9]。

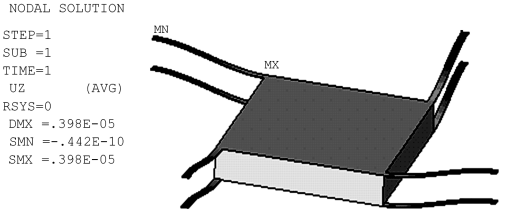
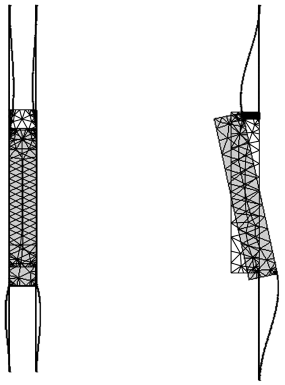


图 3　双层悬臂梁结构在 1 个重力加速度下的位移变化

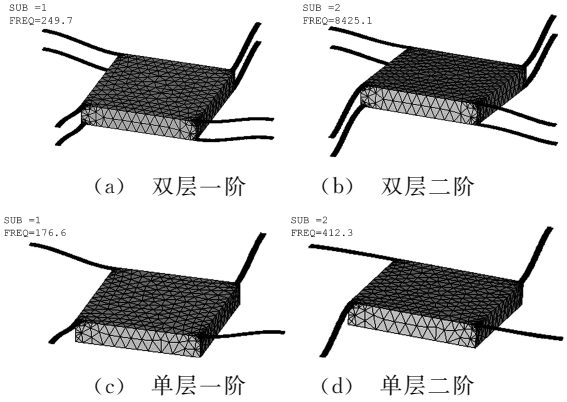


(a)　双层结构　　(b)　单层结构

图 4　双层与单层悬臂梁结构交叉轴方向受到 1 个重力加速度后的机械响应

此外，建立 ANSYS 有限元模型对单层和双层悬臂梁敏感头结构的机械模态进行分析，一阶和二阶振型及其对应频率如图 5 和表 2 所示，其中图 5(a)和图 5(b)分别为双层悬臂梁结构的一阶模态和二阶模态仿真结果，图 5(c)和图 5(d)分别为单层悬臂梁结构的一阶模态和二阶模态仿真结果。两种结构中，一阶模态均为工作模态，二阶模态均为干扰模态。从仿真结果可以看出，所设计双层悬臂梁敏感头结构的模态分离比达到 33.74，相对于单层悬臂梁敏感头结构提升了 13.5 倍，显著降低了敏感头受高阶模态扰动的影响。

综上所述，所设计的双层悬臂梁结构具有良好的位移灵敏度与极低的交叉轴灵敏度，既解决了单层悬臂梁结构交叉轴灵敏度过大使加速度计精度降低的问题，又使敏感头抗高阶机械模态扰动能力显著提升。



(a)　双层一阶　　(b)　双层二阶

(c)　单层一阶　　(d)　单层二阶

图 5　双层和单层悬臂梁结构的一阶及二阶模态仿真结果

表 2　双层悬臂梁结构和单层悬臂梁结构模态对比

结构	一阶模态频率 f_1/Hz	二阶模态频率 f_2/Hz	f_2/f_1
单层悬臂梁结构	176.6	412.3	2.33
双层悬臂梁结构	249.7	8 425.1	33.74

3　工艺流程

所设计“玻璃-硅-玻璃”三明治结构敏感头使用 MEMS 工艺加工而成。

首先加工敏感头的金属电极和光栅，如图 6(a)所示。通过溅射工艺在玻璃基底上形成金属电极，通过 lift-off 工艺形成金属光栅，通过激光打孔工艺在上玻璃片上形成引线用通孔。

然后进行悬臂梁和质量块的加工，使用 500 μm 的双面抛光单晶硅，在其两侧各生长 25 μm 的硅，并重掺杂硼，如图 6(b)所示。通过湿法腐蚀，在掺杂硅的两侧各腐蚀 5 μm，形成传感气隙，如图 6(c)所示。进一步采用湿法腐蚀，腐蚀深度 20 μm，形成悬臂梁结构和接线柱部分结构，如图 6(d)所示。采用 ICP 刻蚀技术，将接线柱周围的硅进一步刻蚀，保留 10 μm 厚的硅层使接线柱和硅片保持连接，如图 6(e)所示。采用各项异性湿法腐蚀技术将硅片腐蚀，释放悬臂梁结构，如图 6(f)所示。

最后进行晶片键合，将上玻璃片和硅片采用阳极键合工艺固定在一起，如图 6(g)所示。采用 ICP 刻蚀技术，将接线柱周围生成的硅去除，如图 6(h)

所示。将下玻璃板与硅片进行阳极键合固定,经划片后形成独立的加速度计敏感探头,如图 6(i)所示。

阳极键合形成的密闭腔室,有效避免了颗粒、灰尘、腐蚀性气体等外部环境的影响,保护敏感头中的挠性结构和电极结构,有效提高了加速度计的可靠性。接线柱的设计实现了敏感头三层结构的电气隔离,有利于加速度计的器件级封装。

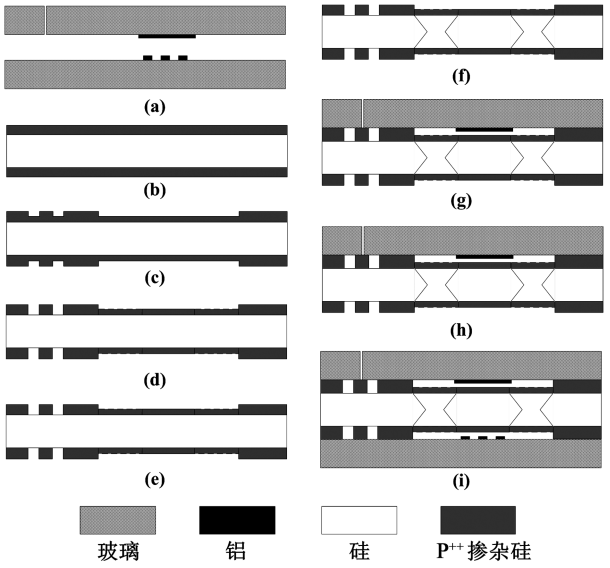


图 6 玻璃-硅-玻璃结构敏感头的加工工艺流程

4 结论

本文针对传统微米光栅加速度计交叉轴灵敏度高的问题,提出了一种双层悬臂梁的微米光栅加速度计结构及加工方法。理论分析结果表明:

(1)所设计双层悬臂梁加速度计位移灵敏度为 3.992 μm/G,交叉轴灵敏度基本为 0;

(2)一阶模态频率为 249.7 Hz,模态分离比为 33.74,相比传统结构提升 13.5 倍,有效提高了敏感头抗高阶机械模态扰动的能力。

参考文献:

[1] Hall N A, Okandan M, Littrell R, et al. Micromachined accelerometers with optical interferometric read-out and integrated electrostatic actuation[J]. J. Microelectromech. Syst., 2008, 17(1): 37-44.

[2] Bicen B, Jolly S, Jeelani K, et al. Integrated optical displacement detection and electrostatic actuation for directional optical microphones with micromachined biomimetic diaphragms[J]. IEEE Sensors J., 2009, 9(12): 1933-1941.

[3] Lee W, Hall N A, Zhou Z, et al. Fabrication and characterization of a micromachined acoustic sensor with integrated optical readout[J]. IEEE J. of Sel. Topics in Quantum Electronics, 2004, 10(3): 643-651.

[4] Cooper E B, Post E R. High-resolution micromachined interferometric accelerometer[J]. Appl. Phys. Lett., 2000, 76(22): 3316-3318.

[5] Loh N C, Schmidt M A, Manalis S R. Sub-10 cm^3 interferometric accelerometer with nano-g resolution[J]. J. of Microelectromechanical Systems, 2002, 11(3): 182-187.

[6] Krishnamoorthy U, Iii R H O, Bogart G R, et al. In-plane MEMS-based nano-g accelerometer with sub-wavelength optical resonant sensor[J]. Sensors and Actuators A: Physical, 2008, 145/146: 283-290.

[7] Lu Q B, Bai J, Wang K W, et al. Design, optimization and realization of a high-performance MOEMS accelerometer from a double-device-layer SOI wafer[J]. J. of Microelectromechanical Systems, 2017, 26(4): 1-11.

[8] Zhang Y, Feng L S, Wang X, et al. Linearity enhancement of scale factor in an optical interrogated micromechanical accelerometer[J]. Appl. Opt., 2016, 55(22): 6115.

[9] Zhang Y, Shan G, Heng X, et al. Optical sensitivity enhancement in grating based micromechanical accelerometer by reducing non-parallelism error[J]. Opt. Experss, 2019, 27(5): 6565-6579.

[10] Lu Q B, Wang C, Bai J, et al. Minimizing cross-axis sensitivity in grating based optomechanical accelerometers [J]. Opt. Experss, 2019, 24(8): 9094-9111.

[11] Hall N, Lee W, Degertekin F. Capacitive micromachined ultrasonic transducers with diffraction-based integrated optical displacement detection [J]. IEEE Trans. on Ultrasonics Ferroelectrics & Frequency Control, 2003, 50(11): 1570-1580.

[12] Lu Q B, Wang C, Bai J, et al. Subnanometer resolution displacement sensor based on a grating interferometric cavity with intensity compensation and phase modulation[J]. Appl. Opt., 2015, 54(13): 4188-4196.

作者简介:

高 山(1992—),男,河北人,博士研究生,主要从事 MOEMS 加速度计方面的研究。

E-mail: shangao@buaa.edu.cn

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.003

基于有限元法的 12 μm 微测辐射热计单元设计

史佳欣，陈　超，王　军

（电子科技大学 光电科学与工程学院，成都 610054）

摘　要： 基于多物理场有限元分析与理论计算相结合的方法，采用 Intellisuite 软件完成了 12 μm×12 μm 微测辐射热计结构的设计与仿真，具体工作包括：单元结构二维版图及工艺流程设计和单元结构三维精确建模，结合实际 MEMS 结构的材料参数，进行了电学与热电耦合多物理场有限元仿真模拟分析。通过仿真优化获得探测单元的主要热电参数、响应时间和响应率，分别为：热导 4.31×10^{-8} W/K、热容 2.69×10^{-10} J/K、电压响应率（未经后端读出电路放大）7 200 V/W、热响应时间 6.24 ms。采用所提出的微桥设计仿真方法，可显著提高器件设计效率和设计精度，缩短研发周期，可满足超大规模小像元非致冷红外焦平面探测器的设计要求。

关键词： 微测辐射热计；小像元；非致冷红外焦平面探测器；有限元仿真；热电耦合分析

中图分类号： TN215　　**文章编号：** 1001－5868(2020)01－0015－05

Design of 12 μm Microbolometer Based on Finite Element Method

SHI Jiaxin，CHEN Chao，WANG Jun

（School of Optoelectronic Science and Engin.，University of Electronic Science and Technol. of China，Chengdu 610054，CHN）

Abstract： Combining the methods of multi-physics finite element analysis and theoretical calculation，a unit structure of 12 μm×12 μm microbolometer was designed and simulated by Intellisuite software，including：2D layout of unit structure，process design and 3D accurate modeling of the unit structure. Based on the material parameters of actual MEMS structure，finite element simulations were performed by electrical and thermal-couple multiphysics. The main thermoelectric parameters，response time and response rate of the detection unit were obtained through simulation optimization. And the thermal conductivity of 4.31×10^{-8} W/K，heat capacity 2.69×10^{-10} J/K，voltage response rate of 7 200 V/W and the thermal response time of 6.24 ms were obtained. The proposed micro-bridge design can improve the design efficiency and accuracy effectively，shorten the development cycle，and meet the design requirements of ultra-large-scale small-pixel uncooled infrared focal plane detectors.

Key words： microbolometer；small pixel；uncooled infrared focal plane detector；finite element simulation；thermal-couple

0　引言

近年来，基于微测辐射热计的非致冷红外焦平面探测器因其无需致冷、成本低、功耗低、重量轻和光谱响应宽等优点[1]，在车载夜视、监控、医疗、森林防火、电力系统、单兵作战、枪瞄、车载、舰载、机载以及红外末制导等军民领域得到广泛的应用。较小尺寸的像元结构可有效减小探测器的体积和重量，降低功耗与成本，因此小像元微测辐射热计的结构设计具有重大研究意义。国际主流高性能超大规模非致冷红外焦平面探测器均已采用 12 μm×12 μm 的单元结构[2]。

12 μm×12 μm 结构微测辐射热计结构由于其像元尺寸小，设计时，在获得高占空比、小热导以及快速热响应之间存在一定矛盾。较大的占空比意味

收稿日期：2019－07－26.

着无法设计出较长的桥腿结构，即无法获得较小的热导。而通过设计较长的桥腿来获得较小的热导，一方面会降低结构占空比，另一方面还会增加器件的热响应时间。同时，由于探测单元结构尺寸小，轻微的三维尺寸变化也将影响器件的电学特性。因此，需要通过精确可行的理论分析与仿真模拟的方法，来确定合理的制备方案，这将有效缩短器件的研制周期。同时，通过合理的设计仿真提前预知器件的热学和电学参数，对后端读出电路设计也将提供较好的参考。

本文将对微测辐射热计的电学、热学响应进行仿真分析，并且通过理论计算分析，得到器件的电阻、电阻温度系数、电压响应率、热响应时间、热容、热导，从而获知模型的性能，提高设计效率，为后续器件制备提供参考。

1 理论分析及模型

1.1 基本理论

红外探测主要是探测目标的微弱热辐射，被探测目标与其所处的背景环境由于温度和辐射率不同，所发出的红外辐射强度就不同，红外探测器将这些有差异的红外辐射按照一定规律转换成电信号，最终利用成像设备将电信号转换成图像[3]。其中对器件性能有着决定性影响的就是其电学和热学参数。一方面，目标发出的红外辐射作用在微测辐射热计上，将引起其阻值的变化，阻值直接决定了器件的输出特性。过大或者过小的器件阻值都不利于探测，并且器件的电阻温度系数也直接决定着器件的性能。另一方面，微测辐射热计结构的热学参数也对器件性能有着重要影响，微测辐射热计的热导 G 和热响应时间 τ 的计算方法分别如式(1)和(2)所示，其中 k_i，w_i，d_i，l_i 分别为各层材料的热导率、宽度、厚度、长度，G_{rad} 为材料自身向外界的辐射热导。

$$G = G_{rad} + 2\sum_i k_i \frac{w_i}{l_i} d_i \tag{1}$$

$$\tau = C/G \tag{2}$$

降低热导 G，可以有效减小噪声等效温差(NETD)，但会增大热时间常数，不利于提高器件帧频。设计中应使热容 C 与热导 G 保持在合理的比值，从而兼顾低热导和快速响应[4]。

因此在设计中需要对微测辐射热计结构的电学和热学特性进行仿真，得到相关热电参数，从而为后续的结构仿真以及后端电路设计提供参考。

首先通过电学有限元分析获得微测辐射热计在不同温度时的电阻值，并根据公式 $\alpha = \mathrm{d}(\ln R_s)/\mathrm{d}(\ln T)$ 求得器件的电阻温度系数 α，其中 T 为温度，R_s 为器件电阻。

随后进行热电耦合有限元分析，首先加载电学偏置，偏置信号引起器件温度升高，温升达到稳定之后，再将功率 P_{IR} 为 10 nW 的红外辐射加载到该像元，引起器件温度变化。根据公式 $\Delta T = P_{IR}/G_{eff}$，可以求出整个器件的热导值 G_{eff}。随后，根据公式 $G_{eff} = G - I^2 R_s \alpha$ 求得该器件的固有热导 G，其中 I 为偏置电流，R_s 为该偏置下的器件电阻。同时从器件的热响应曲线可得到热响应时间 τ，并根据式(2)，结合已知的热导 G，得到器件的热容 C。

1.2 有限元模型分析

有限元分析包括电学有限元分析和热电耦合有限元分析。在电学有限元分析中，于器件模型底电极加载 27 ℃初始的温度载荷以及不同的偏置电流，可得到温度-时间曲线以及响应电压-时间曲线，由此分析单元结构的温度响应特性[5]。根据不同偏置电流下的响应电压，可以得到微测辐射热计在对应温度下的电阻值，在此基础上计算得到模型的电阻温度系数 α。

在热电耦合分析中，当电学偏置信号引起结构温度升高至 T_1 时，器件达到热平衡态，随后加载红外热辐射，温度继续升高至 $0.632(T_3 - T_1)$ 所花的时间即为热响应时间，$0.632(T_3 - T_1)$ 对应的时间点是 t_2，加载红外热辐射的时间点为 t_1，故响应时间可由 $\tau = t_2 - t_1$ 计算，如图 1 所示。

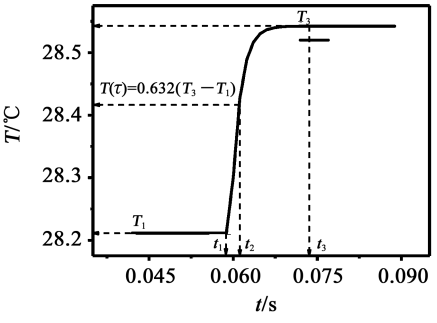


图 1 微测辐射热计的时间与温度响应曲线

将微测辐射热计的输出信号电压 EA_d 与入射辐射信号 S 之间的比值定义为电压响应速率 R_v：

$$R_v = \frac{EA_d}{S} \tag{3}$$

在热电耦合分析中，可以根据热响应电压 V_τ

和输入热辐射功率 P_{IR} 来计算电压响应率 $R(T_{bb})$。

2　有限元仿真设计

2.1　三维建模

采用 Intellisuite 软件中的 Blueprint 版块进行版图设计，采用 Fab 版块进行虚拟工艺仿真并建立三维模型，进行网格划分。最后采用 TEM 版块定义材料参数，设置边界条件，加载电学偏置和红外辐射信号进行仿真，材料各项参数如表 1。

表 1　材料各项参数

材料	密度/ ($kg \cdot cm^{-3}$)	比热容/ ($J \cdot (kg \cdot ℃)^{-1}$)	热导率/ ($W \cdot (m \cdot ℃)^{-1}$)	电阻率/ ($\Omega \cdot cm$)
NiCr	8 900	440	30	0.000 1
Si_3N_4	2 400	691	4.5	1×10^{15}
VO_x	3 360	500	5	1.17～0.38 (20～70 ℃)

相同单元尺寸下，高占空比和低热导以及低热导和快速响应之间存在一定矛盾[6]。因此在设计中，应结合实际工艺条件，设计合理的结构。本文模型版图如图 2 所示，单元尺寸为 12 μm×12 μm，其中 VO_x 面积为 81 μm^2，桥腿宽度为 0.25 μm，桥腿长度为 15.13 μm。采用 $Si_3N_4/VO_x/Si_3N_4$ 夹层真空谐振腔结构，Si_3N_4 总厚度为 300 nm，VO_x 厚度为 30 nm，NiCr 电极厚度为 20 nm。

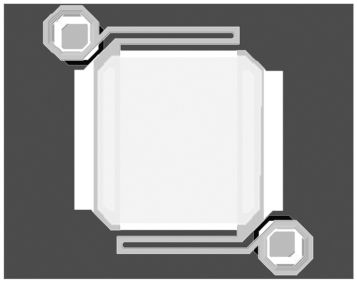
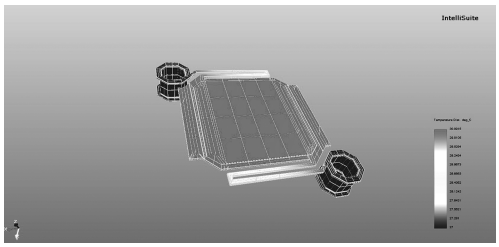


图 2　12 μm 微测辐射热计模型图

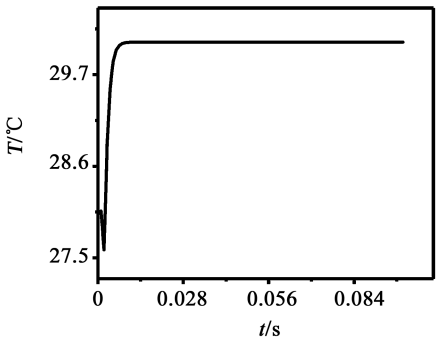
2.2　电学有限元分析

本文的有限元仿真中，初始温度设定为 27 ℃，即加载在底电极的初始温度为 27 ℃，加载电学偏置为 1 μA 电流，周期为 0.1 s。加载电学偏置一定时间后，模型的温度先升高而后达到平衡。如图 3(a)所示，器件整体已经达到平衡状态，桥面温度为 30.09 ℃，温度响应曲线如图 3(b)所示。

模型的整体电势 V_{olt} 分布如图 4(a)所示，图 4(b)为加载了电流偏置底电极的电势曲线图，由此可知模型的电压响应为 0.139 21 V，最终稳定电势为 0.14 V。因此，在 30.09 ℃下，结构的电阻值可以通过伏安定理得到，为 139.21 kΩ。

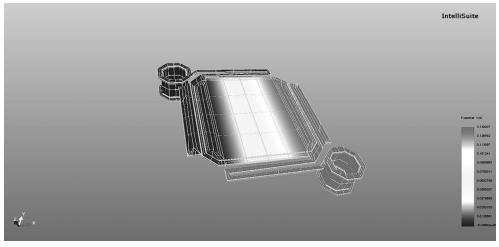


(a)　温度分布

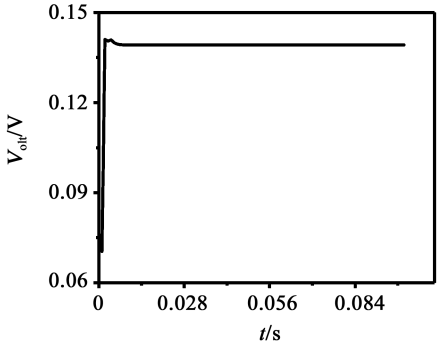


(b)　温度变化曲线

图 3　1 μA 偏置电流下器件的温度分布及变化曲线



(a)　电势分布

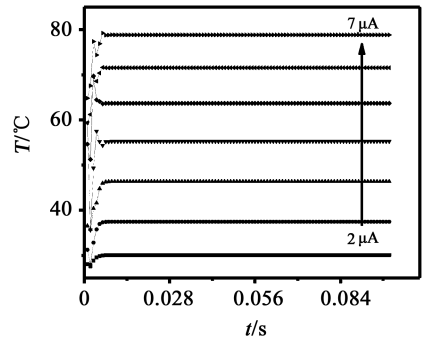


(b)　电势变化

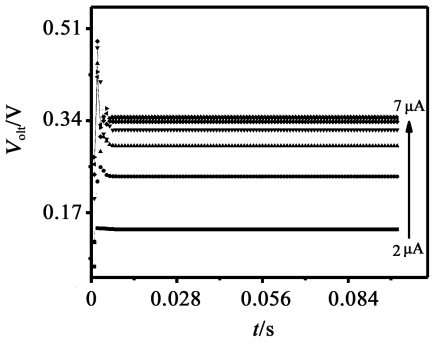
图 4　1 μA 偏置电流下器件的电势分布及电势变化曲线

随后加载 2～7 μA 的偏置电流，周期为 0.1 s，得到对应偏置电流下模型的温度、电势分布，如图 5 所示。由于电流的瞬间加载，图 5(a)和图 5(b)的温度和电势曲线在刚开始时刻出现了尖峰，但最后趋于稳定。图 5(a)为模型桥面中心温度，即 VO_x 热敏电阻温度随时间变化。如图 5(b)所示，不同偏置电流下底电极的电势。

利用不同偏置电流下的温度值与电压值，根据公式 $\alpha = \mathrm{d}(\ln R)/\mathrm{d}(\ln T)$，计算得到电阻温度系数，求其平均值为−0.021 2，同时计算出 27 ℃下整体模型的电阻值为 140.12 kΩ。



(a) 温度变化



(b) 电势变化

图 5 不同偏置电流下的器件温度和电势变化

2.3 热电耦合有限元分析

热电耦合有限元仿真时，器件两个桥墩底部初始温度设为 27 ℃，红外辐射比电压偏置延后 0.06 s，同时也很好地模拟了器件真实的工作状态并方便后续的仿真数据处理。其中，模型的 α 值与 27 ℃时器件整体电阻采用本文的电学仿真数据，分别为−0.021 2 和 140.12 kΩ。

器件在实际工作时，为了防止热敏电阻过热烧坏，均采用脉冲偏置的热辐射。脉冲偏置峰值采用后端读出电路的典型偏置电压，分别取 1.75，2，2.25，2.75 和 3 V，脉冲的有效作用时间为 20 s，凝视时间为 15 ms。仿真过程中，如果直接加载多个脉冲偏置，在仿真设置时需要设置多个时间点，这样将会使得仿真时间大大增加，降低仿真效率。因此本文采用等效偏置直流电压来提高仿真效率，根据式(4)进行等效替换，其中 V_{eqbias} 为等效偏置电压，V_{fid} 为脉冲电压，t_{int} 为脉冲有效作用时间，t_{star} 为凝视时间，R_{e} 为器件电阻。仿真时，脉冲峰值偏置电压等效后的直流电压分别为 0.006 5，0.007 3，0.082 2，0.091 3，0.1 和 0.11 V，从能量守恒角度来看，脉冲电压与等效直流电压偏置引起温度升高一致。

$$\frac{V_{\mathrm{eqbias}}^2 t_{\mathrm{star}}}{R_{\mathrm{e}}} = V_{\mathrm{fid}}^2 t_{\mathrm{int}}/R_{\mathrm{e}} \tag{4}$$

将图 6 所示等效红外载荷加载在模型上，其中辐射延迟了 0.06 s，保证电压偏置下温度达到充分平衡，并且热辐射持续了 0.06 s，保证红外辐射引起的温度升高也能达到稳定。

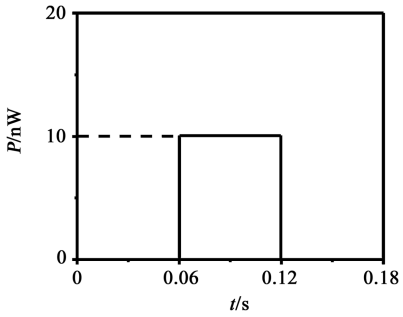


图 6 热电耦合加载能量情况

图 7 是热电耦合器件在不同偏置电压下的温度曲线。将该温度曲线数值导出并用 Origin 软件拟合图形得到更多的数据点。如图 8 为 0.1 V 偏置电压下热响应温度变化的拟合曲线，利用 Origin 软件可由 100 个点精确至 1 000 个点，由此得到更精确的热时间常数 $\tau = 6.24$ ms。此时器件底电极电势变化如图 9 所示。

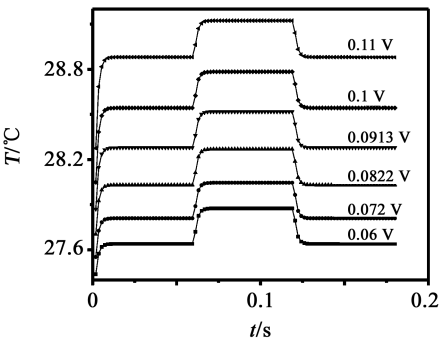


图 7 不同偏置电压下热电耦合仿真温度变化曲线

由图 9 可知，在 0.06 s 至 0.12 s 之间，由于红外辐射的作用，可以观察到器件由于对红外辐射产生了响应，因此电压也会发生小幅度下降。电压响应率 R_{v} 的定义是探测器的输出电压 V_{s} 与入射辐射功率 P_{IR} 之比，即 $R_{\mathrm{v}} = V_{\mathrm{s}}/P_{\mathrm{IR}}$，单位为 V/W[7]。对于器件热电耦合过程中的电压响应率，得到未经后端读出电路放大的电压响应度 R_{v} 为 7 200 V/W。根据公式 $\Delta T = P_{\mathrm{IR}}/G_{\mathrm{eff}}$，式中 ΔT 为模型温度升高

值，P_{IR} 为加载的热辐射能量值，G_{eff} 为有效热导值，计算得其有效热导为 4.31×10^{-8} W/K 。根据公式 $\tau=C/G$ 求得器件的热容为 2.69×10^{-10} J/K 。

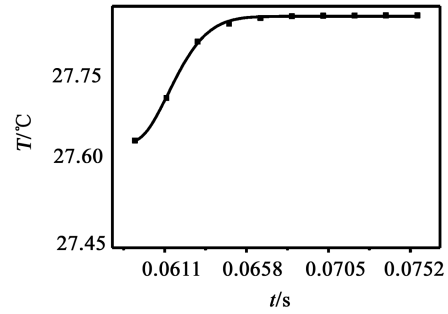


图 8　Origin 拟合的热响应温度变化曲线

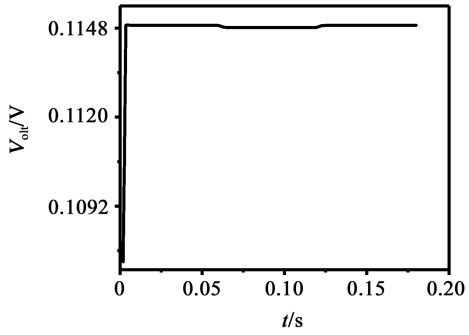


图 9　0.1 V 偏置电压下底电极的电势变化曲线

3　结论

通过对 12 μm×12 μm 微测辐射热计单元结构的设计仿真，得出器件的电阻温度系数 α 的平均值为−0.021 2，热响应时间为 6.24 ms，原始电压响应率(未经后端读出电路放大)为 7 200 V/W。在此基础上得到 12 μm×12 μm 微桥结构探测单元的有效热导和热容值分别为 4.31×10^{-8} W/K 和 2.69×10^{-10} J/K。该设计可满足制备超大规模、高性能快速响应非致冷红外焦平面探测器的要求。本文所提出的设计方法可为非致冷红外焦平面探测器的实际制备及其读出电路的设计提供可靠的参考，同时也为非致冷红外焦平面探测器的研制提供了一种快速、准确、低成本的设计方法。

参考文献：

[1] 申　宁. 基于标准 CMOS 工艺的微测辐射热计研究[D]. 大连：大连理工大学，2015.
Shen Ning. Research on micro-bolometer based on standard CMOS technology [D]. Dalian: Dalian University of Technology, 2015.

[2] Muckensturm K M, Weiler D, Hochschulz F, et al. Measurement results of a 12 μm pixel size microbolometer array based on a novel thermally isolating structure using a 17 μm ROIC[J]. Proc. SPIE, 2016: 98191N.

[3] 李　意，雷志勇，李青松. 红外探测技术的应用与发展[J]. 国外电子测量技术，2018，37(2)：80-83.
Li Yi, Lei Zhiyong, Li Qingsong. Application and development of infrared detection technology [J]. Foreign Electronic Measurement Technology, 2018, 37(2): 80-83.

[4] Li Ge, Yuan Ningyi, Li Jinhua, et al. Thermal simulation of micromachined bridge and self-heating for uncooled VO_2 infrared microbolometer[J]. Sensors & Actuators A, 2005, 126(2):430-435.

[5] 张　龙. 20 μm 微测辐射热计的结构优化及性能研究[D]. 成都：电子科技大学，2015.
Zhang Long. Structural optimization and performance study of 20 μm microbolometer[D]. Chengdu: University of Electronic Science and Technology of China, 2015.

[6] 苟　君，蒋亚东，王　军，等. 太赫兹微测辐射热计的设计、仿真与制备[J]. 科学通报，2014，59(8)：727-732.
Gou Jun, Jiang Yadong, Wang Jun, et al. Design, simulation and preparation of terahertz microbolometer [J]. Chinese Science Bulletin, 2014, 59(8): 727-732.

[7] 谢　骞. MEMS 微桥结构设计与光刻工艺参数优化[D]. 成都：电子科技大学，2016.
Xie Qian. MEMS micro-bridge structure design and lithography process parameter optimization[D]. Chengdu: University of Electronic Science and Technology of China, 2016.

[8] 康炀东. 钨掺杂氧化钒基非致冷红外探测器的制备与性能研究[D]. 南宁：广西大学，2017.
Kang Yidong. Preparation and properties of tungsten-doped vanadium oxide-based uncooled infrared detectors [D]. Nanning: Guangxi University, 2017.

作者简介：

史佳欣(1985－)，女，江苏人，硕士研究生，主要研究方向为红外探测器。

E-mail：791503385@qq.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.004

级联倍增 InAlAs/InAlGaAs 雪崩光电二极管暗电流成分分析

张 承，黄晓峰，迟殿鑫，唐 艳，王 立，柴松刚，崔大健，莫才平，高新江

（重庆光电技术研究所，重庆 400060）

摘 要： 实验制备了级联倍增 InAlAs/InAlGaAs 雪崩光电二极管，对二极管暗电流随台面直径和温度的变化进行了研究分析。结合暗电流函数模型，利用 Matlab 软件对暗电流的各成分进行了数值计算，并仿真研究了芯片结构的缺陷浓度 N_t 和表面复合速率 S 对暗电流的影响。结果表明，二极管暗电流主要来自于体暗电流，而非表面漏电流。在工作点偏压 90 V 处，受缺陷影响的缺陷辅助隧穿电流 I_{tat} 在暗电流中占据了主导，并推算出了芯片结构的缺陷浓度 N_t 约为 10^{19} m^{-3}、吸收区中的缺陷浓度 N_{InGaAs} 约为 7×10^{15} m^{-3}。由于芯片结构的缺陷主要来源于 InAlAs/InAlGaAs 倍增区和 InGaAs 吸收区，而吸收区缺陷占比很少，因此认为缺陷主要来自于异质结 InAlAs/InAlGaAs 倍增区。

关键词： InAlAs/InAlGaAs 雪崩光电二极管；暗电流；表面漏电流；缺陷辅助隧穿电流；缺陷浓度

中图分类号： TN312 **文章编号：** 1001－5868(2020)01－0020－05

Analysis of Dark Current Components of InAlAs/InAlGaAs Cascade Multiplication Avalanche Photodiode

ZHANG Cheng, HUANG Xiaofeng, CHI Dianxin, TANG Yan, WANG Li,
CHAI Songgang, CUI Dajian, MO Caiping, GAO Xinjiang

(Chongqing Optoelectronics Research Institute, Chongqing 400060, CHN)

Abstract: InAlAs/InAlGaAs cascade multiplication avalanche photodiode was fabricated experimentally. The dark current of the diode was studied and analyzed with the change of the diameter and temperature of the mesa. The components of the dark current were calculated by MATLAB combined with the dark current function model. The effects of defect concentration N_t and surface recombination rate S of the chip structure on the dark current were simulated and studied. The results show that the dark current of the diode mainly comes from the body dark current, not from the surface leakage current. The defect-assisted tunneling current I_{tat} affected by defect concentration dominates the dark current at 90 V bias. The defect concentration in the chip structure N_t is 10^{19} m^{-3} and the defect concentration in the absorption region N_{InGaAs} is 7×10^{15} m^{-3}. Because the defects mainly come from the InAlAs/InAlGaAs multiplication region and the absorption region, and the proportion of the defects in the absorption region is very small, so it believes that the defects mainly come from the heterojunction InAlAs/InAlGaAs multiplication region.

Key words: InAlAs/InAlGaAs APD; dark current; surface leakage current; trap-assisted tunneling current; defect concentration

收稿日期：2019－10－15.

0　引言

级联倍增 InAlAs/InAlGaAs 雪崩光电二极管是基于传统 SAGCM 结构在 InP 衬底上外延 InGaAs 材料作为吸收区，InAlAs/InAlGaAs 异质结材料作为倍增区，运用碰撞电离工程(I^2E)而形成多级单载流子倍增(Single-Carrier-Multiplication)，用于实现低噪声和高增益的雪崩光电二极管[1-3]。暗电流是雪崩光电二极管的重要性能指标之一，与材料结构生长和工艺制备[4-7]紧密相关。

随偏置电压变化，雪崩光电二极管中各暗电流成分对暗电流贡献不同。在低偏压下暗电流的主要成分包括：表面漏电流 I_s、产生-复合电流 I_{gr} 和扩散电流 I_{diff}；在高偏压下则为：缺陷辅助隧穿电流 I_{tat} 和带间直接隧穿电流 I_{bbt}[8]。分析暗电流数据中的各个成分组成，对于器件设计和工艺制备具有指导意义。

本文从台面直径、温度两个方面结合暗电流各成分模型，运用 Matlab 数值分析方法，对级联倍增 InAlAs/InAlGaAs 雪崩光电二极管的暗电流成分进行了研究分析。讨论了倍增区和吸收区的缺陷浓度 N_t 与表面复合速率 S 对级联倍增 InAlAs/InAlGaAs 雪崩光电二极管暗电流的影响。

1　实验制备

材料结构如图 1 所示，在 n 型 InP 衬底上依次外延 InAlAs/InAlGaAs 异质结倍增层、电荷层、InGaAs 吸收层及 p 型 InAlAs 接触层。采用 H_3PO_4 ∶ H_2O_2 ∶ H_2O(体积比为 1∶1∶5)溶液湿法腐蚀成双台阶，制备了台面尺寸 t 分别为 15 μm(样品 A)，30 μm(样品 B)和 60 μm(样品 C)的三种样品；然后涂覆苯丙环丁烯(BCB)，经 260 ℃固化 1 h，在侧壁形成钝化保护，并起到平坦化作用。接着，热蒸发 TiPtAu 定义 p 电极，并在 350 ℃下退火 60 s。减薄后，在背面生长 SiN_x 增透膜并光刻出背面光敏区，最后采用 AuGeNi 制备 n 电极。

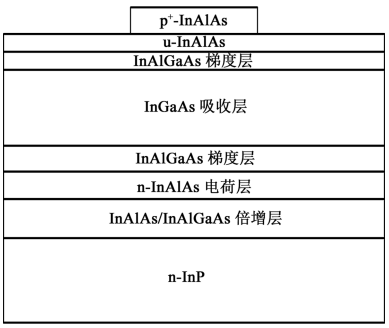


图 1　台面型 InGaAs-InAlAs/InAlGaAs 雪崩光电二极管结构示意图

2　模型

雪崩光电二极管的暗电流成分主要包括：产生-复合电流 I_{gr}、扩散电流 I_{diff}、带间隧穿电流 I_{bbt}、缺陷辅助隧穿电流 I_{tat} 和表面漏电流 I_s[9]。

产生-复合电流 I_{gr} (Shockley-Read-Hall，SRH)[10]是位于耗尽区的杂质或缺陷形成的产生复合中心而引起的电流，可以用式(1)表示；扩散电流 I_{diff} 是耗尽区两侧一个少子扩散长度内的电子-空穴对的随机热产生与复合过程而产生的电流，如式(2)所示；带间隧穿电流 I_{bbt} 是在反向偏压下，能带弯曲使 pn 结耗尽区内的电子从价带直接进入导带，形成隧道效应，用式(3)表示；缺陷辅助隧穿电流 I_{tat} (Trap-Assisted Tunneling，TAT)[11-14]是电子借助耗尽区内的缺陷或杂质中间态来实现能带隧穿而产生的电流，复杂的 TAT 隧穿过程依赖于陷阱在能带中的位置、陷阱密度、陷阱被电子占据的比例以及陷阱势垒的高度和宽度，因此，缺陷辅助暗电流与电场强度密切相关，用式(4)表示。经设计，吸收区电场低于隧穿电场 15 V/cm，吸收区的缺陷通过少子寿命 τ 来影响产生-复合电流 I_{gr} 和扩散电流 I_{diff}，所以我们认为缺陷浓度 N_t 主要来自于倍增区。表面漏电流 I_s 主要由表面产生-复合电流 I_{s1}、表面隧穿电流 I_{s2} 和表面沟道电流 I_{s3} 组成，其中 I_{s2} 和 I_{s3} 与温度无关，被称为表面欧姆漏电流 I_{oh}，与电压成线性关系[15]。I_s 与温度和电压的关系如式(5)所示。

$$I_{gr}=\frac{qn_iAw}{\tau_{eff}}[\exp(qV/2kT)-1] \tag{1}$$

$$I_{diff}=\left(\frac{Aqn_i^2D_p^2}{N_d\tau_p^2}\right)[\exp(qV/kT)-1] \tag{2}$$

$$I_{bbt}=A\frac{\left(\frac{2m^*}{E_g}\right)^{1/2}q^3E_mV}{4\pi^2\hbar^2}\exp\left(\frac{-\theta m_0^{1/2}\varepsilon_g^{3/2}}{q\hbar E_m}\right) \tag{3}$$

$$I_{tat}=\frac{Aq^2\pi^2N_tWm^*F^2E}{h^3(E_g-E_t)}\cdot \exp\left(-\frac{8\pi(2m^*)^{1/2}(E_g-E_t)^{1/2}}{3qhE}\right) \tag{4}$$

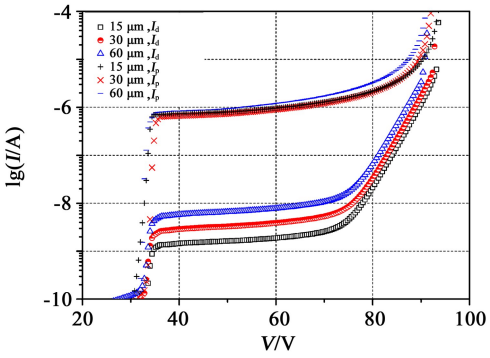
$$I_s=\frac{n_iqAS}{2}V^{1/2}\exp(-E_g/4kT) \tag{5}$$

式中，q 是电子电荷，n_i 为本征载流子浓度，A 为有源区面积，W 为耗尽区宽度，τ_{eff} 为有效载流子寿命，V 为偏置电压，k 为玻尔兹曼常数，T 为温度，m^* 为有效质量，m_0 为静止质量，E_m 为最大电场强度，h 为普朗克常量，$\hbar = h/2\pi$ 为理化普朗克常量，E_g 为带隙，τ_p 为空穴寿命，D_p 为空穴扩散系数，N_d 为施主杂质浓度，N_t 为耗尽区陷阱浓度，E_t 为陷阱能级，θ 为有效质量与静止质量比值，取值 8.49×10^{-8} eV·cm，F 是关于 a 的函数，$F(a)=\pi/2+\sin^{-1}(1-2a)+2(1-2a)\times(a\times(1-a))^{0.5}$，$a=E_t\times E_g$，$S$ 为表面复合速率。

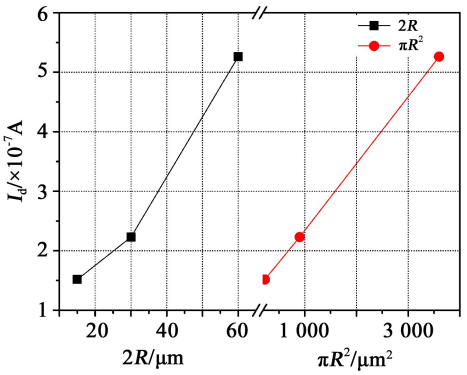
3 讨论与分析

图 2 为样品 A～C 在 300 K 下的光电流 I_p 和暗电流 I_d 随电压变化曲线图，可以看出，暗电流随台面直径增加而变大。图 2(b)中，在 85 V 偏压下暗电流并不随台面直径($2R$)线性变化，而是正比于台面面积(πR^2)，表明暗电流主要与来自中心区域产生的体暗电流相关。因样品 A 的暗电流最低，后续将选择样品 A 作为对象进行讨论。

研究了暗电流与温度的关系，提取了二极管样品 A 在增益 $M=1.0$ 下的激活能参数 E_a，如图 3 所示，当温度从 293 K 降低至 223 K 时，$E_a=0.424$ eV，约合 $0.57E_g$，表明此时暗电流成分主要为产生-复合电流 I_{gr} 与扩散电流 I_{diff}。当温度在 223～203 K 时，$E_a=0.146$ eV，约 $0.196E_g$，暗电流的主要成分是表面漏电流 I_s。图 4 为样品 A 在 203，223，253 和 293 K 温度下的暗电流曲线。从图中可以看出，随温度从 293 K 降低至 203 K，暗电流 I_d 明显下降，在低于 V_p 拉通点(34.5 V)电压范围内暗电流并没有改变，说明产生复合电流 I_{gr} 和扩散电流 I_{diff} 已经减少至表面漏电流 I_s 之下，这时的表面漏电流 I_s 约 5 pA。



(a) 样品 A～C 在 300 K 下的 I-V 曲线



(b) 样品 A～C 直径 $2R$、面积 πR^2 与暗电流 I_d 的关系

图 2 芯片样品 A～C 在 300 K 下的特性曲线

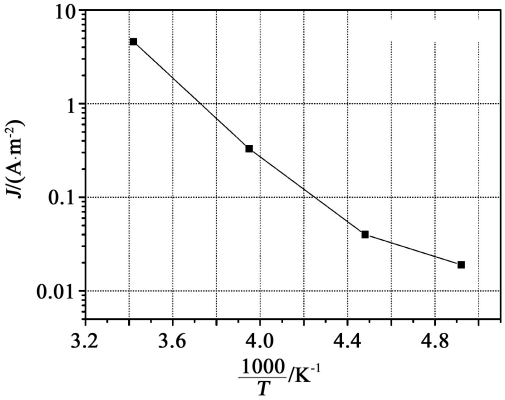


图 3 样品 A(增益 $M=1$)的 J-T 关系曲线

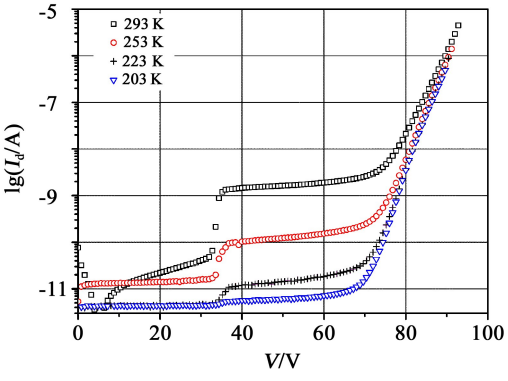
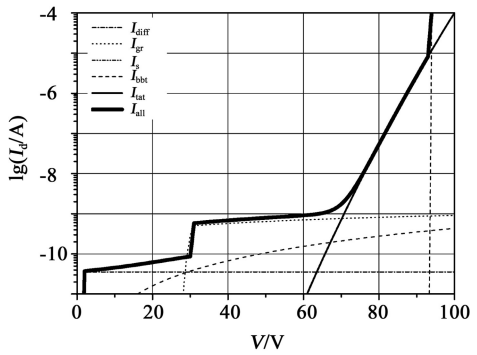
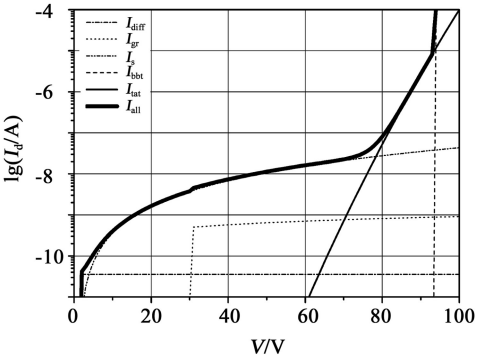


图 4 不同温度下样品 A 的 I_d-V 曲线

图 5(a)给出了样品 A 在 300 K 时暗电流各组成成分曲线，与实测暗电流基本吻合。表面漏电流 I_s 在 0～100 V 范围内均小于 1 nA，此时表面复合速率为 5×10^6 m·s^{-1}，与文献[16]报道的数据在同一数量级，表明了台面工艺具有良好的表面界面态，湿法刻蚀与台面钝化并未引入高浓度复合中心与高界面态。当表面复合速率 S 增加两个数量级，如图 5(b)所示，表面漏电流 I_s 大于产生-复合电流 I_{gr}，在小于 80 V 时占据了主导，而在高偏压 90 V 附近缺陷辅助隧穿电流 I_{tat} 对暗电流影响较大。



(a)　表面复合速率 $S=5\times10^{6}$ m·s^{-1}



(b)　表面复合速率 $S=5\times10^{8}$ m·s^{-1}

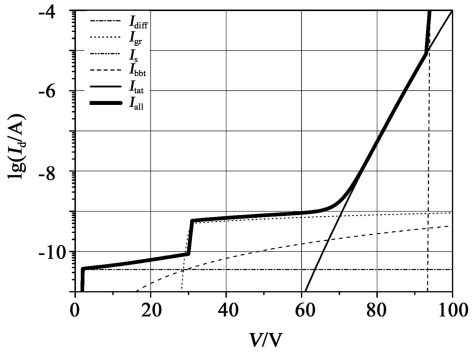
图 5　300 K 时不同表面复合速率下仿真的暗电流曲线

图 6 所示为样品 A 的缺陷浓度 N_t 与缺陷辅助隧穿暗电流 I_{tat} 的关系曲线。图 6(a)与样品 A 的暗电流曲线基本吻合，工作点在 90 V 高偏压下缺陷辅助隧穿电流 I_{tat} 是暗电流的主要成分。TAT 隧穿过程依赖于缺陷在能带中的位置、缺陷密度、缺陷被电子占据比例以及陷阱势垒的高度和宽度，电子借助高浓度缺陷实现能带隧穿，因此辅助隧穿电流与缺陷浓度密切相关。经过计算，此时芯片结构的缺陷浓度 N_t 为 10^{19} m^{-3}。

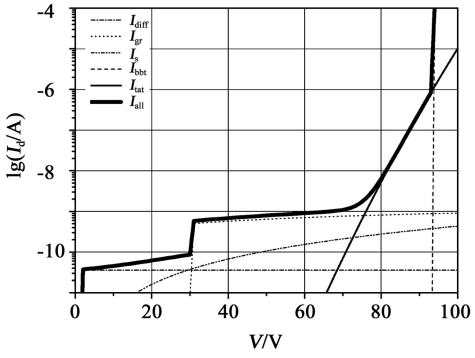
InGaAs 吸收区材料的缺陷浓度与少子寿命的关系如式(6)[17-18]，其中 C 是常数，并且可以推算出 C 值为 3.5×10^{4} s/cm^{3}。根据暗电流曲线提取样品 A 的 InGaAs 少子寿命 τ_{InGaAs} 约为 5 μs，得到 InGaAs 的缺陷浓度 N_{InGaAs} 约为 7×10^{15} m^{-3}。由于芯片结构的缺陷浓度来源于倍增区和吸收区，而吸收区缺陷浓度比总缺陷浓度 N_t 低 3 个数量级，因此我们认为相关的缺陷主要来自 InAlAs/InAlGaAs 的倍增区[19-20]。

$$N_{InGaAs}=C/\tau_{InGaAs} \tag{6}$$

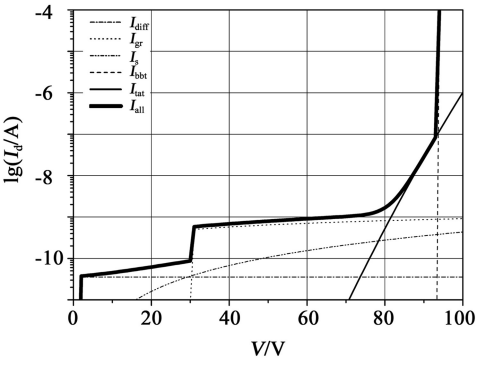
由图 6 仿真结果可知，为降低高偏压下的辅助隧穿电流，倍增区的缺陷浓度 N_t 至少需要降低 1～2 个数量级，InAlAs/InAlGaAs 级联倍增区 I^{2}E 雪崩光电二极管才能获得较低的暗电流，实现线性模式下的低噪声高灵敏度探测。对于材料生长来说控制缺陷浓度是一项关键技术。



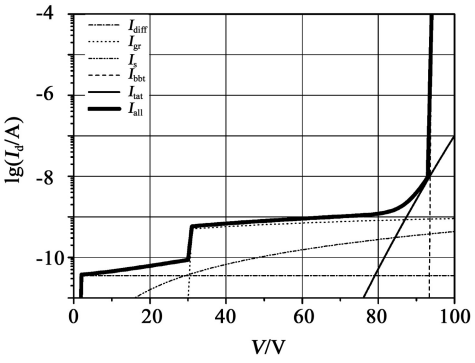
(a)　缺陷浓度 $N_t=10^{19}$ m^{-3}



(b)　缺陷浓度 $N_t=10^{18}$ m^{-3}



(c)　缺陷浓度 $N_t=10^{17}$ m^{-3}



(d)　缺陷浓度 $N_t=10^{16}$ m^{-3}

图 6　300 K 时不同缺陷浓度下仿真的暗电流曲线

4 结论

实验制备了级联倍增 InAlAs/InAlGaAs 雪崩光电二极管，从台面直径与温度两方面对芯片暗电流成分进行分析，并结合暗电流函数模型与 Matlab 软件对暗电流各成分进行数值仿真，仿真结果与实测暗电流 *I-V* 曲线基本吻合。暗电流主要来自于体暗电流，而非表面漏电流。研究了缺陷浓度对暗电流的影响，结果表明，90 V 偏压下缺陷辅助隧穿电流是暗电流的主要成分。由于芯片结构的缺陷主要来自于倍增区和吸收区，而吸收区缺陷浓度比总缺陷浓度 N_t 小 3 个数量级，因此我们认为缺陷来源于 InAlAs/InAlGaAs 的倍增区。倍增区的缺陷浓度至少需要降低 1～2 个数量级，InAlAs/InAlGaAs 级联倍增雪崩光电二极管才能获得较低的暗电流，实现线性模式下的低噪声高灵敏度探测。

参考文献：

[1] Williams G M, Compton M, Ramirez D A, et al. Multi-gain-stage InGaAs avalanche photodiode with enhanced gain and reduced excess noise[J]. IEEE J. of the Electron Devices Society, 2013, 1(2): 54-65.

[2] Amir H, Mahdi Z. Analysis of single photon detection in avalanche photodetectors with multi-gain-stage multiplication region[J]. Appl. Opt., 2017, 56(6):1631.

[3] Williams G M, Ramirez D A, Hayat M M, et al. Time resolved gain and excess noise properties of InGaAs/InAlAs avalanche photodiodes with cascaded discrete gain layer multiplication regions[J]. J. of Appl. Phys., 2013, 113(9): 093705.

[4] Weisbuch C, Dingle R, Gossard A C, et al. Optical characterization of interface disorder in GaAs-$Ga_{1-x}Al_xAs$ multi-quantum well structures[J]. Solid State Commun., 1981, 38(8): 709-712.

[5] Kohl M, Heitmann D, Tarucha S, et al. Optical investigation of the exciton transfer between growth islands of different well widths in GaAs/$Al_xGa_{1-x}As$ quantum wells[J]. Phys. Rev., 1989, 39(39): 7736-7743.

[6] Gottwaldt L, Pierz K, Ahlers F J, et al. Correlation of the physical properties and the interface morphology of AlGaAs/GaAs heterostructures[J]. J. of Appl. Phys., 2003, 94(4): 2464.

[7] Rouvie A, Reverchon J L, Huet O, et al. InGaAs focal plane arrays developments at Ⅲ-Ⅴ lab[J]. Proc. SPIE, 2012, 8353: 1-12.

[8] 曾巧玉. InGaAs/InP 单光子雪崩光电二级管的制备及研究[D]. 上海：中国科学院上海技术物理研究所，2014.
Zeng Qiaoyu. Preparation and study of InGaAs/InP single-photon avalanche photoelectric secondary tube[D]. Shanghai: Shanghai Institute of Technical Physics of the Chinese academy of Sciences, 2014.

[9] Zhao Y, Zhang D, Qin L, et al. InGaAs-InP avalanche photodiodes with dark current limited by generation-recombination[J]. Opt. Express, 2011, 19(9): 8546.

[10] Schenk A. A model for the field and temperature dependence of SRH lifetimes in silicon[J]. Solid-State Electronics, 1992, 35: 1585-1596.

[11] Hurkx G A M, Klaasen D B M, Knuvers M P G, et al. A new recombination model describing heavy-doping effects and low temperature behaviour[C]// IEEE Inter. Electron Devices Meeting, 1990: 307-310.

[12] Nguyen T, Musca C A, Del J M, et al. Dark current in long wavelength infrared HgCdTe gated photodiodes[J]. J. Electron. Mater., 2004, 33(6): 621-629.

[13] Ajisawa A, ODA N. Improvement in HgCdTe diode characteristics by low temperature post-implantation annealing [J]. J. of Electron. Mater., 1995, 24(9): 1105-1111.

[14] Donnelly J P, Duerr E K, Mclntosh K A, et al. Design considerations for 1.06 μm InGaAsP-InP Geiger-mode avalanche photodiodes[J]. IEEE J. of Quantum Electron., 2006, 42(8): 797-809.

[15] Willardson R K, Beer A C. Metal Insulator Semiconductor Detectors in Semiconductors and Semimetals[M]. New York: Academic Press, 1981.

[16] Bothra S, Tyagi S D, Ghandhi S K, et al. Surface recombination velocity and lifetime in InP measured by transient microwave reflectance[J]. Solid-State Electron., 1990, 34(1): 47-50.

[17] Sheng J J, Leohardt D, Han S M. Empirical correlation for minority carrier lifetime to defect density profile germanium on silicon grown by nanoscale interfacial engineering[J]. J. of Vacuum Science & Technology, 2013, 31(5): 1-8.

[18] Ji Xiaoli, Liu Baiqing, Tang Hengjing. 2.6 μm MBE grown InGaAs detectors with dark current of SRH and TAT[J]. AIP Advances, 2014, 4(8): 087135.

[19] Taguchi K, Torikai T, Sugimoto Y, et al. Planar-structure InP/InGaAsP/InGaAs avalanche photodiodes with preferential lateral extended guard ring for 1.0～1.6 mμm wavelength optical communication use[J]. J. of Lightwave Technol., 1988, 6(11): 1643-1655.

[20] Williams G M, Ramirez D A, Hayat M M, et al. Instantaneous receiver operating characteristic (ROC) performance of multi-gain-stage APD photoreceivers[J]. IEEE J. of The Electron Devics Society, 2013, 1(6): 145-153.

作者简介：

张 承(1987—)，男，硕士，工程师，主要从事近红外光电探测器方面的研究。

E-mail：1030879833@qq.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.005

基于敏感薄膜的F-P光纤温度传感器的高温特性

高晓丹[1]，彭建坤[2]

（1. 武汉东湖学院 电子信息工程学院，武汉 430212；2. 武汉理工大学 光纤传感技术国家工程实验室，武汉 430070）

摘　要： 根据薄膜型光纤温度传感器的感温原理及光学薄膜的应力特性，利用膜层干涉的特征矩阵分析了传感元件的传感特性，建立了薄膜型光纤温度传感器的热应力特性及其对传感特性影响的理论模型。通过对薄膜和光纤的热光效应和热膨胀特性的分析，对比了蓝宝石光纤、纯硅芯光纤和普通多模光纤三种光纤基底的热应力-温度特性，设计了高温传感器传感探头的薄膜膜系，分析了膜系中每层膜的热应力-温度特性，以及临界载荷应力特性，确定所选光纤基底脱膜的临界应力值。通过模拟不同应力作用下的光谱传感特性，为传感器薄膜敏感探头的稳定性和可靠性提供理论依据和支持，提高传感器研制的效率。

关键词： 光纤传感器；薄膜；热应力；温度特性

中图分类号： TP212　　**文章编号：** 1001－5868(2020)01－0025－04

Study on High Temperature Characteristics of F-P Fiber Temperature Sensor Based on Sensitive Film

GAO Xiaodan[1], PENG Jiankun[2]

（1. College of Electronic Information Engin., Wuhan Donghu University, Wuhan 430212, CHN;
2. National Engin. Lab. for Fiber Optic Sensing Technol., Wuhan University of Technol., Wuhan 430070, CHN）

Abstract: According to the temperature sensing principle of the fiber optic temperature sensor and the stress characteristics of coating, the theoretical model of the thermal stress characteristics and its influence for sensing characteristics was established by using the characteristic matrix of thin film. The thermal stress-temperature characteristics of the three kinds of fiber, inculding sapphire fiber, pure silicon core fiber and multimode fiber, were compared by analyzing the thermo-optic effect and thermal expansion characteristics of the film and fiber. A thin film formula for high temperature sensing probe was designed. And the thermal stress-temperature characteristics of each film layer of the formula, as well as the critical load stress characteristics, were analyzed, thus the critical load stress value of the selected fiber can be determined. By simulating the spectral sensing characteristics under different stresses, it provides theoretical basis and support for the stability and reliability of sensor-sensitive probes, and improves the efficiency of sensor development.

Key words: fiber optic sensor; thin film; thermal stress; temperature characteristics

0　引言

光纤传感器具有体积小、灵敏度高、耐高温高压、抗辐射等优点，且易于组网，受到研究人员的广泛关注[1-4]。将新型功能材料与光纤元件组合制成各种新型传感器是当前传感技术领域的重要趋势。其中，基于纳米薄膜干涉的光纤传感器具有测量参数多样性的优势[5]。纳米薄膜可以通过多种技术实现，包括电子束蒸发[6-7]、磁控溅射[8-9]、旋涂[10]和电化学镀[11]等，但是薄膜器件的长期稳定性和可靠性对于工业应用而言依然是一个挑战。纳米薄膜在制

收稿日期：2019－09－06.
基金项目：国家自然科学基金项目(61575151)；湖北省教育厅科学技术研究项目(B2019268).

备的过程中会产生大小不一的残余应力，这些应力的存在会影响薄膜材料的微观结构，甚至材料的物理性能，导致器件的光场发生畸变，影响其传感特性，严重的会导致薄膜龟裂、脱落，从而限制了其工业应用的寿命。

薄膜型光纤高温传感器是在光纤端面沉积单层或多层纳米级薄膜构成法布里-珀罗(F-P)干涉腔，传感器件的尺寸即为光纤自身的尺寸，且不需要刻蚀复杂的微结构，在航空航天发动机、内燃机、石油化工设备等结构复杂、环境恶劣的应用领域用于高温部件的结构健康监测。由于传感器的工作环境恶劣，温度变化范围较宽(1 000 ℃以上)，因而其热应力的影响不可忽略，需要考虑热应力对薄膜型光纤高温传感器传感特性的影响，以便于综合选择传感器的镀膜材料、设计及工艺，进一步了解传感器的传感特性，尽可能减少由于薄膜应力而导致的稳定性问题。

1 薄膜型光纤高温传感器的理论模型

纳米薄膜型光纤高温传感器采用物理气相沉积技术在光纤端面镀上单层或多层薄膜构成温度敏感元件，通过膜层的法布里-珀罗干涉性质，从光源发出的光经光纤传输到敏感膜层形成反射干涉光谱，如图 1 所示。

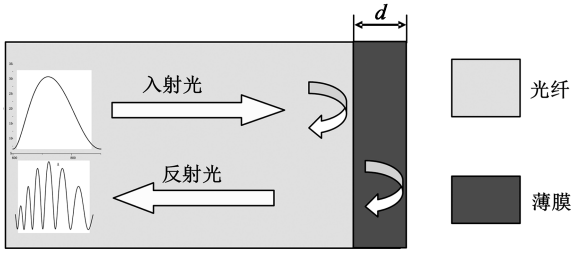


图 1 薄膜型温度传感探器原理图

根据法布里-珀罗腔的干涉理论，反射光谱的强度如下：

$$I_r(\lambda)=I_1(\lambda)+I_2(\lambda)+2\sqrt{I_1(\lambda)I_2(\lambda)}\cdot\cos\left(\frac{2\pi\Delta}{\lambda}+\varphi_0\right) \tag{1}$$

其中，$I_1(\lambda)$和 $I_2(\lambda)$分别为膜层两界面的反射光强，φ_0 为初始相位，Δ 为两界面反射光的光程差，它取决于膜层的折射率 n 和膜层的厚度 d，其中，

$$\Delta=2nd \tag{2}$$

从式(1)可见，反射光谱干涉条纹的频率取决于薄膜腔的光程差 Δ，第 m 个干涉极大值为

$$\lambda_m=\frac{\Delta}{m-\frac{\varphi_0}{2\pi}}=\frac{2nd}{m-\frac{\varphi_0}{2\pi}} \tag{3}$$

由于薄膜材料的热光特性和热膨胀特性，当温度发生改变时，薄膜层的折射率 n 和厚度 d 都会随着温度的变化而改变，从而使干涉光谱随温度的变化而产生波长的漂移，通过对干涉光谱波长漂移量的监测可反演出温度信息。

同时，当环境温度从 T_0 变化到 T 时，单层薄膜的热应力为

$$\sigma=(\alpha_s-\alpha_d)\frac{E_d(T-T_0)}{1-\nu_d} \tag{4}$$

其中，α_s 和 α_d 分别为光纤材料和膜层材料的热膨胀系数，E_d 和 ν_d 分别为薄膜的杨氏模量和泊松比。

对于传感探头上的多层敏感薄膜，热应力受膜层之间以及膜层与光纤的热膨胀适配程度等多方面影响。图 2 为多层膜传感探头示意图，d_s 为光纤基底的长度，$d_i(i=1,2,\cdots,n)$为第 i 层膜的厚度，光纤基底与薄膜的交界面坐标为 $y=0$。

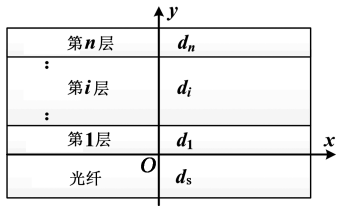


图 2 多层膜传感探头示意图

当温度发生改变时，基底光纤与膜层材料的热膨胀系数不同，使多层膜发生弯曲，多层膜系统应变[12]可表示为

$$\varepsilon=\varepsilon_c+\frac{y-d_c}{r} \tag{5}$$

其中，ε_c 为多层膜系统的均匀应变，d_c 为弯曲应变为 0 的中性面，r 为系统的弯曲曲率半径。多层膜系统温度改变 ΔT 时，光纤基底的热应力 σ_s 和第 i 层膜的热应力 σ_i 分别为

$$\sigma_s=E_s(\varepsilon-\alpha_s\Delta T) \tag{6}$$

$$\sigma_i=E_i(\varepsilon-\alpha_i\Delta T) \tag{7}$$

式中，E_s 为光纤基底的杨氏模量，E_i 和 α_i 分别为第 i 层薄膜的杨氏模量和热膨胀系数。

考虑光纤传感探头光纤基底的长度远大于薄膜的厚度，根据弯曲力学理论，薄膜系统的合力为零以及弯曲力矩为零[13]，可以求出光纤传感探头光纤基底的热应力 σ_s 和第 i 层膜的热应力 σ_i 分别为

$$\sigma_s = -\frac{6y+4d_s}{d_s^2}\sum_{i=1}^{n} E_i d_i(\alpha_s-\alpha_i)\Delta T,\quad i=1,2,\cdots,n \tag{8}$$

$$\sigma_i = E_i(\alpha_s-\alpha_i)\Delta T,\quad i=1,2,\cdots,n \tag{9}$$

光纤基底和膜层中的应力会导致光线的入射角度偏差，应力较大会致使膜层中 P 偏振光和 S 偏振光分离，利用膜层的特征矩阵 $\boldsymbol{M}_j(T)$ 计算多层膜的 P 偏振和 S 偏振的反射光谱特性：

$$\boldsymbol{M}_j(T)=\begin{bmatrix}\cos\delta_j(T) & \dfrac{\mathrm{i}\sin\delta_j(T)}{\eta_j(T)}\\ \mathrm{i}\eta_j(T)\sin\delta_j(T) & \cos\delta_j(T)\end{bmatrix} \tag{10}$$

结合传感器的温度传感原理以及热应力特性，分析传感探头干涉光谱随温度变化的特征波长相对变化量为

$$\frac{\Delta\lambda}{\lambda}=\frac{\lambda_m(T)-\lambda_m(T_0)}{\lambda_m(T_0)}=\frac{2n(T)\mathrm{d}(T)-2n(T_0)\mathrm{d}(T_0)}{2n(T_0)\mathrm{d}(T_0)} \tag{11}$$

2　材料的热应力-温度特性

高温传感器的工作环境恶劣，对材料的要求较高，薄膜材料的选择除了要考虑常规的折射率、波段透明度、机械牢固度、吸收和散射以外，尤其需要关注材料的温度特性，包括热膨胀特性、热光特性等[13]。综合考虑，选取 ZrO_2 和 Al_2O_3 为薄膜材料，图 3 为两种介质材料的热膨胀特性，图 4 为两种材料的折射率(n)-温度特性。

为了能更好地运用理论指导实践，研究了不同材料应力-温度特性。选用蓝宝石光纤、纯硅芯光纤和普通多模光纤三种光纤作传感探头的薄膜基底，在其上面蒸镀 ZrO_2 薄膜，分析其热应力(σ)-温度特性，如图 5 所示。从图 5 中可以看出，蓝宝石光纤与 ZrO_2 薄膜材料的匹配性较好，在 200～1 000 ℃内，热应力变化范围最小，在 580 ℃附近，随着温度的升高，由负的压应力转换为正的拉应力。纯硅芯光纤的热应力在 200～1 000 ℃内只有压应力，随着温度的升高，应力值逐渐增加。而普通多模光纤在 200～1 000 ℃内热应力随着温度的增加逐渐增大，匹配性较差。因此，蓝宝石光纤是薄膜型高温传感器的最佳光纤选择，普通多模光纤随着温度的升高热应力过大，传感探头的稳定性和可靠性较差。

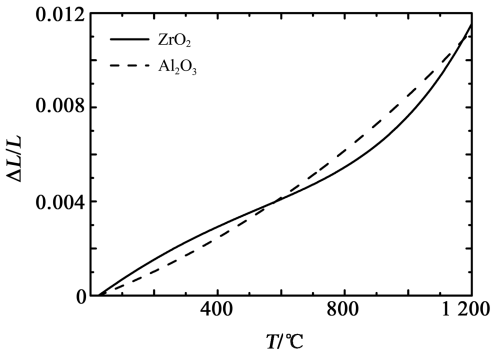


图 3　ZrO_2 和 Al_2O_3 的热膨胀特性曲线

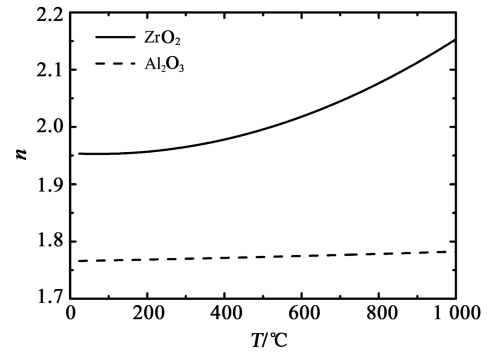


图 4　ZrO_2 和 Al_2O_3 的折射率温度特性曲线

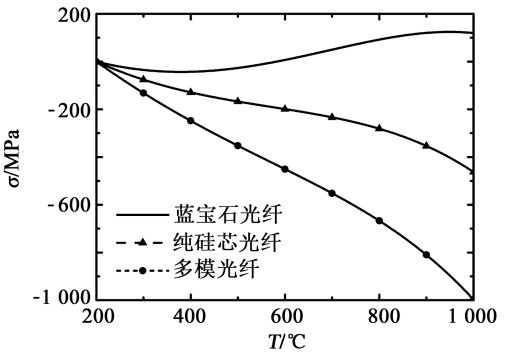


图 5　不同基底光纤的热应力-温度特性曲线

3　热应力传感特性及稳定性

考虑薄膜型高温传感器的测试要求，以 ZrO_2 和 Al_2O_3 分别为高、低折射率材料，设计光纤传感探头的膜系为：Fiber | ZrO_2 (68 nm)/Al_2O_3 (1 432 nm)/ZrO_2(68 nm)|Air。

由于高低折射率材料的热膨胀系数具有不同的温度特性，因此，膜系中两种不同材料的膜层热应力也有所区别，分别以纯硅芯光纤和蓝宝石光纤基底为例，分析多层膜中高低折射率膜层的热应力-温度特性，如图 6 所示。可见，对于纯硅芯光纤，ZrO_2 和 Al_2O_3 膜层的热应力都为压应力，随着温度的升高，压应力逐渐增加，到 1 000 ℃附近，热应力达到 500 MPa 左右。对于蓝宝石光纤，Al_2O_3 膜层的热

应力几近为 0，随着温度的升高，ZrO_2 膜层的热应力最大为 120 MPa。

运用光学薄膜临界载荷应力模型[14]分析膜系的临界载荷值随脱膜半径的关系，如图 7 所示。可见，镀膜基底的中心承载能力较强，随着径向的增加，脱膜的临界载荷值逐渐下降，对于半径 25 μm 的蓝宝石光纤而言，ZrO_2 和 Al_2O_3 膜层在蓝宝石光纤边缘的临界载荷值分别为 242.7 和 496.2 MPa，超过该临界值薄膜就会出现裂痕甚至脱膜。

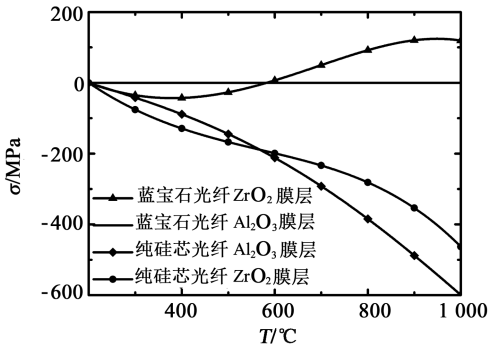


图 6 膜系中各膜层热应力-温度特性曲线

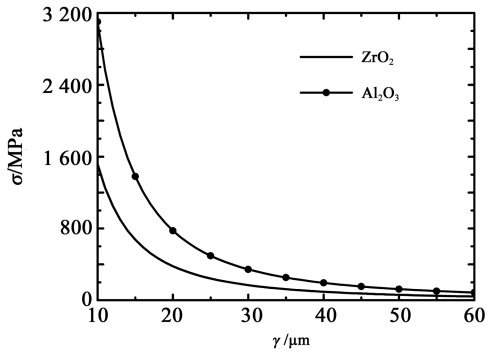


图 7 不同膜层的临界载荷应力特性曲线

热应力不仅会影响传感器件的稳定性，同时还会影响器件的传感特性。通过前面的理论模型，分析了热应力对传感器温度特性的影响，对比了零应力、−100 和 300 MPa 热应力对应的温度传感特性，如图 8 所示。从图 8 中可见，−100 MPa 热应力对应的传感特性与理论传感特性非常接近，误差不是很大，但 300 MPa 热应力时的温度传感特性与理论特性误差较大，波长漂移的相对变化量有正负之分，不利于传感器的温度标定，从而降低测量范围。

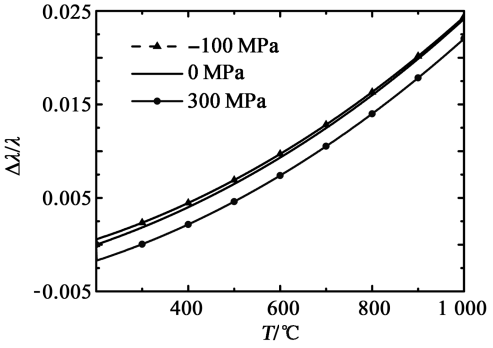


图 8 不同热应力下传感器温度特性曲线

4 结论

通过对薄膜型光纤传感器的热应力及传感特性模型的建立及理论分析，综合选取 ZrO_2 和 Al_2O_3 作为高温传感器的薄膜材料。对比蓝宝石光纤、纯硅芯光纤和普通多模光纤三种薄膜基底传感探头的热应力-温度特性，表明蓝宝石光纤最适合作为高温传感器的光纤基底，更耐高温且蒸镀薄膜的热应力也更小。针对所设计的光纤敏感探头的薄膜膜系，分析了膜系中每层膜的热应力-温度特性，以及膜层的临界载荷应力与脱膜半径的关系，得出蓝宝石光纤敏感探头膜层径向的临界载荷值，并且模拟分析了热应力对温度传感特性的影响，以保证设计传感器件的稳定性和可靠性，从而提高传感器的研制效率。

参考文献：

[1] Jiang Y, Xu J, Yang D, et al. Miniaturized fiber Fabry-Perot interferometer for strain sensing[J]. Microwave & Optical Technol. Lett., 2016, 58(6): 1510-1514.

[2] Wang W, Sun T, Peng J, et al. Humidity sensor based on fiber Bragg grating coated with different pore-foaming agent doped polyimides[J]. IEEE Photon. Technol. Lett., 2017, 29(22): 1963-1966.

[3] Zhang G, Dai J, Yang M, et al. Fiber optic hydrogen sensor based on Fabry-Perot interferometer coated with sol-gel Pt/WO_3 coating[J]. J. of Lightwave Technol., 2015, 33(12): 2530-2534.

[4] Wang Y, Shen C, Lou W, et al. Fiber optic relative humidity sensor based on the tilted fiber Bragg grating coated with graphene oxide[J]. Appl. Phys. Lett., 2016, 109(3): 031107.

[5] Yang M, Peng J, Wang G, et al. Fiber optic sensors based on nano-films[J]. Fiber Optic Sensors, 2017, 21: 1-30.

[6] Huang Chujia, Lee Dongwen, Dai Jixiang, et al. Fabrication of high-temperture sensor based on dielectric multilayer film on sapphire fiber tip[J]. Sensors and Actuators, 2015, A232: 99-102.

（下转第 34 页）

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.006

谐振式光纤陀螺的误差及噪声分析

梅　甦，蒋治国，胡宗福

（同济大学 电子与信息工程学院 信息与通信工程系，上海 201804）

摘　要： 谐振式光纤陀螺（Resonator Fiber Optic Gyroscope，RFOG）的性能是多种误差和噪声共同作用的结果。分析了 RFOG 主要误差与噪声的产生机理，建立了各误差与噪声的分析模型，结合具体陀螺参数，对误差和噪声的等效旋转速率进行了估算，对抑制误差和噪声的有关措施进行了数值量化。以 RFOG 零偏稳定性 0.1°/h 需求为例，为抑制背向散射噪声及背向反射误差，相位调制移频时顺时针和逆时针光波的载波强度残留水平之积需高于 112 dB；为抑制克尔效应，顺时针和逆时针方向的光强差需控制在 3.86 nW 以内；地磁场方向不确定性引起的陀螺零偏稳定性约为 0.7°/h。该工作为 RFOG 实验研究和工程设计提供了理论指导。

关键词： 谐振式光纤陀螺；光纤环形谐振腔；误差与噪声；背向散射

中图分类号： TP212　　**文章编号：** 1001－5868(2020)01－0029－06

Analysis on Error and Noise of Resonant Fiber Optic Gyros

MEI Su，JIANG Zhiguo，HU Zongfu

（School of Electron. and Information Engin.，Tongji University，Shanghai 201804，CHN）

Abstract： The generating mechanisms of the main errors and noises of resonaor fiber optica gyroscope（RFOG）are analyzed，and the analyzing models are established. Combining with specific gyroscope parameters，the equivalent angular velocities of the errors and noises are estimated，and the relevant restraint measures are quantified. Taking bias stability of 0.1°/h in the RFOG as an example，to suppress the backscattering noise and the back reflection error，the requirement of the carrier suppression level should be as high as 112 dB，and to suppress the Kerr effect，the light intensity difference between CW and CCW should be controlled at 3.86 nW or less. The gyro bias stability caused by the uncertainty of the earth magnetic field direction is about 0.7°/h.

Key words： resonant fiber optic gyroscope；ring resonator；error and noise；backscattering

0　引言

光纤陀螺是基于光学 Sagnac 效应的高精度惯性传感器件[1-2]。谐振式光纤陀螺（RFOG）相比于干涉式光纤陀螺（IFOG），可以在更小的尺寸上实现与 IFOG 相同的灵敏度[3]，且 Shupe 效应大幅降低，RFOG 采用的窄带相干光源也有利于陀螺标度因数的稳定[4]。因此，RFOG 获得了世界各国研究机构和高校的关注，成为惯性导航方向的研究热门[5-6]。

偏振波动误差是影响 RFOG 性能的一个重要因素，光纤环形谐振腔作为 RFOG 的核心器件，主谐振谷不对称会影响谐振中心频率的精确测量，故因外界温度变化导致主谐振谷与次谐振谷重叠而引起的主谐振谷对称性变化会影响 RFOG 的性能提升[7]。抑制偏振波动主要有以下四种方法：一是采用偏振轴 90°旋转熔接技术[8]，即在光纤环形谐振腔内中点位置处进行偏振轴 90°旋转熔接以抑制偏

收稿日期：2019－09－17.

基金项目：上海航天科技创新基金项目（SAST2015076）.

振波动误差;二是在光纤环形谐振腔内引入偏振控制技术[9],即在光纤环形谐振腔内插入在线起偏器或单偏振光纤减少次偏振态能量以抑制偏振波动误差;三是利用特种光纤的低温度敏感性[10],如空心光子晶体光纤;四是制作空间耦合器组成光纤环形谐振腔[3],即通过制作空间耦合器替代光纤耦合器,以消除耦合器偏振串扰和耦合器偏振相关损耗,进而抑制偏振波动误差。目前,霍尼韦尔公司报道了零偏稳定性为 0.02°/h 的 RFOG[3],浙江大学报道了零偏稳定性为 0.3°/h 的 RFOG[9],北京航空航天大学报道了零偏稳定性为 1.3°/h 的 RFOG[11],由此可见,RFOG 正在逐步迈入惯性级性能。RFOG 的性能是多种误差和噪声共同作用的结果,系统地对 RFOG 的主要误差和噪声进行分析、抑制及数值量化具有重要的工程意义。本文分析的 RFOG 误差和噪声包括瑞利背向散射噪声、背向反射误差、克尔效应误差和法拉第效应误差。

1 光纤陀螺结构

谐振式光纤陀螺的结构如图 1 所示,输入光波场表示为

$$E(t)=E_0 e^{-i(2\pi f_0 t+\phi_0)} \tag{1}$$

式中,E_0,f_0 和 ϕ_0 分别是激光的幅值、频率和初始相位。

图 1 为基于三角波调制解调的 RFOG 结构示意图。相位调制器 PM1 的控制信号是频率为 f_1 的三角波,三角波上升和下降半周期对应的频移分别为 $2f_1$ 和 $-2f_1$;相位调制器 PM2 的控制信号是频率为 f_2 的三角波,三角波上升和下降半周期对应的频移分别为 $2f_2$ 和 $-2f_2$。光经过 PM1 和 PM2 后载波强度残留水平分别为 ρ_{ccw} 和 ρ_{cw}。信号检测方法是通过探测器 D1 得到逆时针路(CCW)光波频移 $2f_1$ 和 $-2f_1$ 所对应的方波输出,以此方波控制 PZT 调节腔长,使 CCW 路谐振中心频率锁定于光源频率,再通过探测器 D2 得到顺时针路(CW)光波频移 $2f_2$ 和 $-2f_2$ 所对应的方波输出,获得转动角速度。设构成光纤环形谐振腔(Fiber Ring Resonator,FRR)的耦合器耦合系数、环程插入损耗分别为 $K=0.03$ 和 $\gamma=0.97$,腔长 $L=10$ m,谐振腔直径 $D=0.1$ m,输入光功率 $P_i=1$ mW。在谐振曲线最大斜率点处,角速度 0.1°/h 对应的 CW 路输出方波幅度为 160 pW。

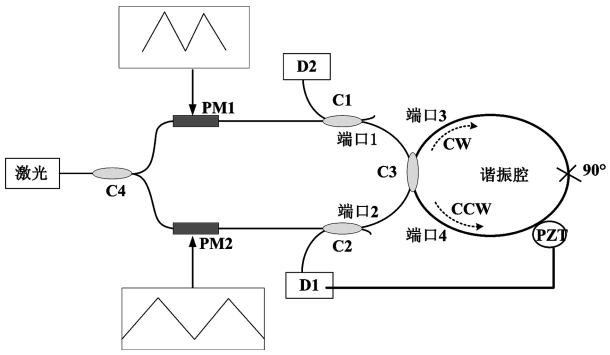


图 1 谐振式光纤陀螺结构示意图

2 背向散射噪声

瑞利散射是光纤的一种基本损耗机制,构成光纤的石英材料是一种无序的结构,其密度在微观上是不均匀的,密度的这种空间变化造成光纤折射率在比波长小的尺度上涨落,光波在这样的介质中传播呈现散射状态,这就是瑞利散射,可认为光纤损耗主要由瑞利散射引起。作为一种估计,有如下近似:

$$I_s/I_0=(1-10^{-\alpha_R L/10})S\approx\alpha_R SL \tag{2}$$

式中,S 为恢复因子,近似为 10^{-3},$\alpha_R=0.2$ dB/km,因此有 $I_s/I_0=10^{-6}$。光通过光纤环形谐振腔后,到达探测器的瑞利背向散射电场可写为[12]

$$E_{sD}=E_0\sqrt{\alpha_R SL}\times K\sum_{m=0}^{\infty}R^m e^{-im(\omega\tau-\theta_s)}\sum_{n=0}^{\infty}R^n e^{-in(\omega\tau+\theta_s)} \tag{3}$$

式中,$R=\sqrt{\gamma(1-K)}$,θ_s 为 Sagnac 效应引起的相角,$\tau=n_e L/C$,n_e 为折射率。瑞利背向散射光强为

$$I_{\tau D}=E_0^2\alpha_R SL\cdot\frac{K}{1+R^2-2R\cos(\omega\tau-\theta_s)}\cdot\frac{K}{1+R^2-2R\cos(\omega\tau+\theta_s)} \tag{4}$$

瑞利背向散射噪声源有两个,一个是背向散射自身的光强,一个是散射光与信号光的干涉项,从两个噪声源对陀螺性能的影响来看,后者的影响远大于前者。对前者的抑制可通过 CW 与 CCW 相位调制频率不同来实现,对后者的抑制可通过相位调制移频来实现,相位调制移频的目的是减少 CW 与 CCW 的同频分量,对 CW 和 CCW 而言,它们的同频分量就是载波,因此,该方法的有效性取决于相位调制移频时的载波抑制程度。

干涉项大小可近似写为

$$I_{int}=2E_0^2\sqrt{\rho_{ccw}\rho_{cw}\alpha_R SL}\cdot$$

$$\sqrt{\frac{K}{1+R^2-2R\cos(\omega\tau-\theta_s)}}\cdot\sqrt{\frac{K}{1+R^2-2R\cos(\omega\tau+\theta_s)}} \tag{5}$$

由式(5)可知，干涉项噪声强度最大值 $I_{int}=0.066$ mW。为使瑞利背向散射对陀螺的影响降到零偏稳定性 0.1°/h 以下，相位调制移频方法的载波强度残留水平需满足：$\rho_{ccw}\rho_{cw}<-112$ dB。

3　背向反射误差

RFOG 中背向反射的主要来源是光路中的不理想熔接点。当熔接点反射率为 −60 dB 时，背向反射对 RFOG 零偏稳定性的影响依然较大。图 2 是光纤环形谐振腔结构示意图，图中 r_1 和 r_2 分别为两个端口 A，B 尾纤方向的振幅反射率，t_1 和 t_2 为振幅透射率，r'_1 和 r'_2 分别为光纤环方向的振幅反射率，t'_1 和 t'_2 为振幅透射率。

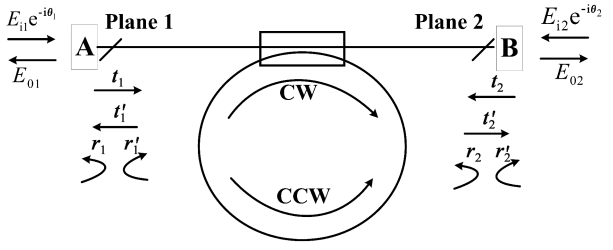


图 2　光纤环形谐振腔示意图

B 端输出的光波有三部分，第一部分是 E_{i2} 在 B 端的反射光，第二部分是 A 端到 B 端的光，此光会通过 B 端反射回 A 端再反射到 B 端，第三部分是从 B 端到 A 端，再从 A 端反射回 B 端的光。设 A，B 两端输入电场振幅相同，则从 A 端到 B 端的强度传递函数为[13]

$$T_1=t_1^2t'^2_2\left|-r_2+\frac{A_{cw}e^{i(-\theta_3+\Delta\theta)}+t_2r'_1A_{cw}A_{ccw}e^{i(-2\theta_3)}}{1-r'_1r'_2A_{cw}A_{ccw}e^{-i\theta_3}}\right|^2 \tag{6}$$

$$A_{cw,ccw}=\sqrt{\gamma(1-K)}-\frac{\gamma K e^{-i\phi_{cw,ccw}}}{1-\sqrt{\gamma(1-K)}e^{-i\phi_{cw,ccw}}}$$

$$\phi_{cw}=2\pi f_{cw}t+\frac{\theta_s}{2},\phi_{ccw}=2\pi f_{ccw}t-\frac{\theta_s}{2}$$

式中，$\Delta\theta$ 为 A，B 两边输入电场的初始相位差，$\theta_3=\frac{2\pi}{\lambda}z$，$z$ 为 A，B 两边尾纤长度之差，f_{cw} 和 f_{ccw} 分别是 CW 方向与 CCW 方向的光频。

同理，从 B 端到 A 端的强度传递函数为

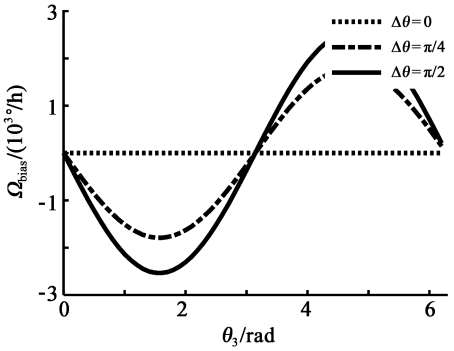
$$T_2=t_2^2t'^2_1\left|-r_1+\frac{A_{ccw}e^{i(-\theta_3-\Delta\theta)}+t_1r'_2A_{cw}A_{ccw}e^{i(-2\theta_3)}}{1-r'_1r'_2A_{cw}A_{ccw}e^{-i\theta_3}}\right|^2 \tag{7}$$

在静止状态即 $\theta_s=0$ 时，通过 $\partial T_1/\partial f\Big|_{f=f'_{cw}}=0$ 和 $\partial T_2/\partial f\Big|_{f=f'_{ccw}}=0$，可得 f'_{cw} 和 f'_{ccw}，f'_{cw} 和 f'_{ccw} 即是静止状态下 CW 方向与 CCW 方向的谐振中心频率。

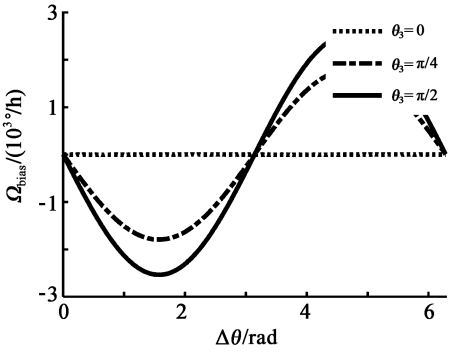
背向反射导致的 RFOG 误差为

$$\Omega_{bias}=\frac{n_e\lambda}{D}(f'_{cw}-f'_{ccw}) \tag{8}$$

图 3 给出背向反射导致的 RFOG 误差 Ω_{bias} 与 θ_3 和 $\Delta\theta$ 的关系，图中端面强度反射率 r_1，r'_1，r_2 和 r'_2 均设为 −60 dB。由图可见，背向反射导致的 RFOG 误差与 $\Delta\theta$，θ_3 相关，且随 $\Delta\theta$，θ_3 正弦变化，最大为 2 800°/h。



(a)　Ω_{bias}-θ_3 关系曲线



(b)　Ω_{bias}-$\Delta\theta$ 关系曲线

图 3　背向反射导致的 RFOG 误差

对于 RFOG 背向反射误差，主要是通过载波抑制的相位调制来消除，通过使 CW 和 CCW 的频率不同，消除背向反射光和信号光之间干涉项和强度项对 RFOG 的影响。当输入光强 $P_i=1$ mW 时，背向反射光和信号光干涉项的幅值最大值为 $I_{re}=$

$2P_i \mid r_2 + t_2 r'_1 \mid = 0.004$ mW，为使背向反射光对陀螺的影响降到零偏稳定性 0.1°/h（光强变化 320 pW）以下，载波强度残留水平需满足：$\rho_{ccw}\rho_{cw} < -88$ dB。

4 克尔效应误差

在干涉式光纤陀螺研究中，抑制光克尔效应误差有两种方法：一种是用方波强度调制，另一种是使用宽带光源。在 RFOG 中，这两种方法均不适用，第一种方法要求强度调制频率等于光纤环的本征频率，而对于 RFOG 来说，光纤环长度从几米到几十米，本征频率对应几兆到几十兆赫兹，其所需的调制精确度难以达到，因而这种方法在 RFOG 中难以实现；对于第二种方法，由于 RFOG 要求光源具备高相干性，故也不适用。因此，为抑制 RFOG 中的克尔效应误差，需要估算 CW 和 CCW 光强差与非互易性误差的关系，再进一步采取措施稳定光强及减少光强差。

光强差导致的 CW 和 CCW 传播常数之差可以写为[14]

$$\Delta\beta_K = \frac{1}{2}\xi\frac{\Delta P}{A_F} \tag{9}$$

式中，ΔP 为 CW 和 CCW 的光强之差，A_F 为光斑面积，$\xi = 96\pi^2\omega^2\chi^{(3)}/(n_e^2 c^2)$，$\chi^{(3)}$ 为非线性光学系数。

静止状态下 CW 和 CCW 环行一圈的非互易性相位差为

$$\Delta\phi_K = \frac{1}{2}\xi\frac{L}{A_F}\left(I_{cw外}\cdot\frac{I_{cw内}}{I_{cw外}} - I_{ccw外}\cdot\frac{I_{ccw内}}{I_{ccw外}}\right) \tag{10}$$

式中，$I_{cw外}$ 和 $I_{ccw外}$ 分别是光纤环形谐振腔 CW 方向与 CCW 方向的输入光强，因为腔内光强和腔外光强不同，所以有比例因子 $I_{cw内}/I_{cw外}$：

$$\frac{I_{cw内}}{I_{cw外}} = \frac{\gamma K}{[1-\sqrt{\gamma(1-K)}]^2 + 4\sqrt{\gamma(1-K)}\cdot\sin^2(\omega\tau/2)} \tag{11}$$

比例因子大小与工作频率点位置相关，对比例因子，有

$$\left|\frac{I_{cw内}}{I_{cw外}}\right|_{max} = \left|\frac{I_{ccw内}}{I_{ccw外}}\right|_{max} = \frac{F}{\pi} \tag{12}$$

式中，F 为精细度。因此，RFOG 中的克尔效应误差可写为

$$\Omega_K = \frac{c\lambda}{4\pi D}\cdot\frac{F}{\pi}\cdot\xi\frac{\Delta P}{A_F} \tag{13}$$

取 $A_F = 28.26\ \mu m^2$（芯直径取 6 μm），$F = 100$，在 1 550 nm 波长处 $\xi = 3.02\times10^{-13}$ m/W，当 $\Omega_K = 0.1°/h$ 时，$\Delta P = 3\ 860$ pW。为抑制克尔效应，CW 和 CCW 光强差需控制在 3.86 nW 以内。

5 法拉第效应误差

即使在一个完全互易的系统中，Sagnac 相移也并不是唯一一种实际存在的非互易性效应，例如由于法拉第效应，磁场会改变圆偏振光的相位，相移的符号取决于圆偏振光的左旋或右旋特性，并与磁场矢量和光传播矢量的方向有关。在光纤陀螺中，两束反向传播的光波之间总的相位差可表示为[15]

$$\Delta\phi_F = 2\int_{闭环回路}\alpha_p VB\,dz \tag{14}$$

式中，V 为介质的沃尔德常数，B 为地磁场，α_p 是一个与偏振态相关的系数，对于线偏振，$\alpha_p = 0$，对于圆偏振，$\alpha_p = \pm1$，对于椭圆偏振，α_p 取介于 0～±1 之间的一个值。由于残余圆偏振的存在，α_p 不再是一个常数，即使线积分 $\int B\cdot dz$ 为 0，$\Delta\phi_F$ 也可能不等于 0。

当均匀磁场 B 沿平行于光纤环平面的方向施加在长度为 z 的一段光纤上时，入射光波（E_x E_y）和出射光波（E'_x E'_y）之间的关系可以由琼斯矩阵给出：

$$\begin{pmatrix}E'_x\\E'_y\end{pmatrix} = C(t_w,\theta_R)\cdot\begin{pmatrix}E_x\\E_y\end{pmatrix} \tag{15}$$

$$C(t_w,\theta_R) = \begin{pmatrix}e^{-i\beta_x z} & 0\\ 0 & e^{-i\beta_y z}\end{pmatrix}\cdot \begin{Bmatrix}\cos(\eta z) - i\frac{\Delta\beta}{2\eta}\sin(\eta z) & \frac{t_w-\theta_R}{\eta}\sin(\eta z)\\ -\frac{t_w-\theta_R}{\eta}\sin(\eta z) & \cos(\eta z) + i\frac{\Delta\beta}{2\eta}\sin(\eta z)\end{Bmatrix}$$

$$\eta = \sqrt{\left(\frac{\Delta\beta}{2}\right)^2 + (t_w-\theta_R)^2}$$

式中，t_w 为光纤扭转率，θ_R 为法拉第旋转，β_x 和 β_y 分别是 X 与 Y 方向的传播常数，$\Delta\beta = \beta_x - \beta_y$ 为单位长度上的线性双折射。

图 4 为地磁场中的光纤环形谐振腔示意图，当光纤环形谐振腔的直径为 0.1 m、长度为 10 m 时，对应线圈有 31 匝，因此，CW 方向的琼斯传递矩阵可以写为

$$\begin{pmatrix}E_x'^{cw}\\E_y'^{cw}\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\cdot(C'_{ED}C'_{DC}C'_{CB}C'_{BA})^{15}\cdot C'_{ED}C'_{DC}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}C'_{CB}C'_{BA}\cdot(C'_{ED}C'_{DC}C'_{CB}C'_{BA})^{15}\cdot\begin{pmatrix}E_x^{cw}\\E_y^{cw}\end{pmatrix} \tag{16}$$

CCW 方向的琼斯矩阵可以写为

$$\begin{pmatrix}E_x'^{ccw}\\E_y'^{ccw}\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}(C'_{AB}C'_{BC}C'_{CD}C'_{DE})^{15}\cdot C'_{AB}C'_{BC}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}C'_{CD}C'_{DE}(C'_{AB}C'_{BC}C'_{CD}C'_{DE})^{15}\cdot\begin{pmatrix}E_x^{ccw}\\E_y^{ccw}\end{pmatrix} \tag{17}$$

式中，θ 为保偏光纤耦合器偏振串扰角。

$$\begin{aligned}
C'_{BA}&=C\left(t_w,\theta_R\cos\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{CB}&=C\left(t_w,\theta_R\sin\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{DC}&=C\left(-t_w,-\theta_R\cos\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{ED}&=C\left(-t_w,-\theta_R\sin\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{AB}&=C\left(t_w,-\theta_R\cos\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{BC}&=C\left(t_w,-\theta_R\sin\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{CD}&=C\left(-t_w,\theta_R\cos\left(\theta+\frac{\pi}{4}\right)\right)\\
C'_{DE}&=C\left(-t_w,\theta_R\sin\left(\theta+\frac{\pi}{4}\right)\right)
\end{aligned} \tag{18}$$

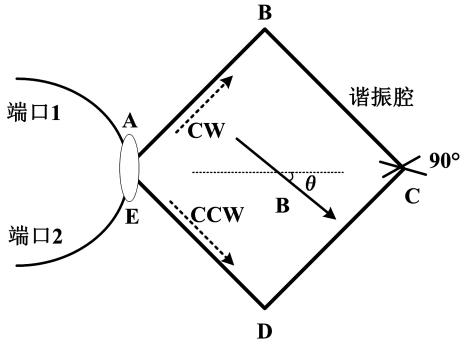


图 4　地磁场中的光纤环形谐振腔示意图

通过上述琼斯传递矩阵，可得到 CW 和 CCW 光波的本征值和本征向量，设 CW 和 CCW 本征值的相角之差为 $\Delta\phi_{F-90}$，则地磁场对谐振式光纤陀螺输出角速度的影响可写为

$$\Omega_E=\frac{n_e\lambda}{D}\frac{FSR}{2\pi}\Delta\phi_{F-90} \tag{19}$$

图 5 为不同磁场方向下的 RFOG 非互易性误差，$t_w=0.1$ rad/m，$\Delta\beta=2\,025$ rad/m，法拉第旋转 $\theta_R=3\times10^{-5}$ rad/m，可见地磁场方向不确定性导致的 RFOG 零偏稳定性约为 0.7°/h。对零偏稳定性 0.1°/h 的 RFOG，法拉第效应误差不可忽略，采用软磁材料对光纤线圈进行磁屏蔽可有效减弱磁场对光纤的作用，单层磁屏蔽可以将陀螺的磁场灵敏度降低 1～2 个数量级，因此对零偏稳定性 0.1°/h 的 RFOG 而言，磁屏蔽的方法可以有效消除地磁场影响。

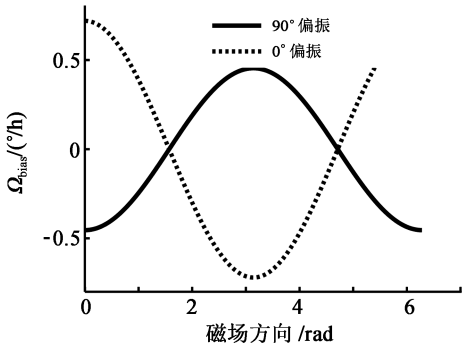


图 5　RFOG 非互易性误差与磁场方向角关系

6　结论

对 RFOG 的主要误差和噪声进行了分析和建模，对抑制误差和噪声的措施进行了数值量化。对腔长 $L=10$ m，环形谐振腔直径 $D=0.1$ m，输入光功率 $P_i=1$ mW 的 RFOG 而言，当熔接点反射率为 −60 dB，背向反射误差随两边尾纤长度差、初始相位差正弦变化，背向反射误差最大为 2 800°/h；地磁场方向不确定性导致的 RFOG 零偏稳定性约为 0.7°/h。以 RFOG 零偏稳定性 0.1°/h 需求为例，采用相位调制移频方法抑制背向散射噪声和背向反射误差，顺时针和逆时针光波的载波强度残留水平之积需低于 −112 dB；为抑制克尔效应，顺时针和逆时针方向的光强差需控制在 3.86 nW 以内；需用磁屏蔽的方法消除地磁场对 RFOG 的影响。

参考文献：

[1] Lefevre H C. The Fiber-Optic Gyroscope[M]. Fitchburg: Artech House, 2014.

[2] 金　毅，吴训忠，谢　聂，等. 光纤陀螺随机漂移在线建模实时滤波技术[J]. 光电工程，2015，42(3)：13-19.

Jin Yi, Wu Xunzhong, Xie Nie, et al. Real-time filtering research based on on-line modeling random drift of FOG[J]. Opto-Electronic Engin., 2015, 42(3): 13-19.

[3] Sanders G, Strandjord L K, Williams W, et al. Improvements to signal processing and component miniaturization of compact resonator fiber optic gyroscopes[J]. 2018 DGON Inertial Sensors and Systems (ISS), 2018: 1-22.

[4] Hotate K. Future evolution of fiber optic gyros[J]. Proc. of SPIE, 1996, 2837: 33-45.

[5] Ma H, Zhang J, Wang L, et al. Development and evaluation of optical passive resonant gyroscopes[J]. J. of Lightwave Technol., 2017, 35: 3546-3554.

[6] 李如春，顾伟驷，马慧莲，等. 谐振式光纤陀螺的数字检测方案及其优化设计[J]. 光电工程，2005，32(2)：45-48.
Li Ruchun, Gu Weisi, Ma Huilian, et al. Digital detection and optimization design for resonant fiber optic gyroscope[J]. Opto-Electronic Engin., 2005, 32(2): 45-48.

[7] Jiao Hongchen, Feng Lishuang, Wang Kai, et al. Analysis of polarization noise in transmissive single-beam-splitter resonator optic gyro based on hollow-core photonic-crystal fiber[J]. Opt. Express, 2017, 25: 27806-27817.

[8] Wang X, He Z, Hotate K. Reduction of polarization-fluctuation induced drift in resonator fiber optic gyro by a resonator with twin 90 degrees polarization-axis rotated splices [J]. Opt. Express, 2010, 18(2): 1677-1683.

[9] Yan Y C, Ma H L, Jin Z H. Reducing polarization-fluctuation induced drift in resonant fiber optic gyro by using single-polarization fiber[J]. Opt. Express, 2015, 23(3): 2002.

[10] Zhao X, Louveau J, Chamoun J, et al. Thermal sensitivity of the birefringence of air-core fibers and implications for the RFOG[J]. J. of Lightwave Technol., 2014, 32(1/4): 2577-2581.

[11] Jiao H C, Feng L S, Wang J J, et al. Transmissive single-beam-splitter resonator optic gyro based on a hollow-core photonic-crystal fiber[J]. Opt. Lett., 2017, 42(15): 3016-3019.

[12] Jiang Z, Hu Z, Fu C. Angular random walk limited by Rayleigh backscattering in resonator fiber optic gyros[J]. Appl. Opt., 2017, 56: 9414-9422.

[13] Takahashi M, Tai S, Kyuma K, et al. Effect of reflections on the drift characteristics of a fiber-optic passive ring-resonator gyro[J]. J. of Lightwave Technol., 1988, 8(5): 811-816.

[14] Li X, Zhang J, Ma H, et al. Test and analysis of the optical Kerr-effect in resonant micro-optic gyros[J]. Photonics J. IEEE, 2014, 6(5): 1-7.

[15] 张桂才. 光纤陀螺原理与技术[M]. 北京：国防工业出版社，2010.
Zhang Guicai. Principle and Technology of Fiber Optic Gyroscope[M]. Beijing: National Defence Industry Press, 2010.

作者简介：

梅　甦(1995—)，男，江苏南通人，硕士研究生，主要研究方向为光纤陀螺信号处理技术；

E-mail：meisuTJU@163.com

胡宗福(1959—)，男，博士，教授，主要研究方向为光载无线通信、光纤传感和光纤通信等。

E-mail：huzongfu@tongji.edu.cn

(上接第 28 页)

[7] Jiankun P, Yapeng Q, Weijia W, et al. Thin-film-based optical fiber Fabry-Perot interferometer used for humidity sensing[J]. Appl. Opt., 2018, 57(12): 2967-2972.

[8] Wang H B, Luo Y T, Li F, et al. Fiber Bragg grating hydrogen sensor based on magnetron sputtering[J]. Appl. Mechanics & Materials, 2015, 738/739: 3-6.

[9] Minz R A, Pal S S, Chopra A, et al. Magnetron sputtering coated optical fiber probe designs for surface plasmon resonance sensor[J]. IEEE J. of Sel. Topics in Quantum Electronics, 2016, PP(99): 1-7.

[10] Zolkapli M, Saharudin S, Herman S H, et al. Quasi-distributed sol-gel coated fiber optic oxygen sensing probe [J]. Optical Fiber Technol., 2018, 41: 109-117.

[11] 王裕波，李玉龙，王文琴，等. 基于光纤光栅传感的化学镀膜过程本征应力监测[J]. 光子学报，2016，45(8)：70-74.
Wang Yubo, Li Yulong, Wang Wenqin, et al. Intrinsic stress monitoring during the chemical coating process based on fiber grating sensing[J]. Acta Photon. Sin., 2016, 45(8): 70-74.

[12] Sava F, Borca C N, Galca A C, et al. Thermal stress effect on the structure and properties of single and double stacked films of GeTe and SnSe[J]. Physica Status Solid., 2017, 255 (6): 1700552.

[13] Gao Xiaodan, Yang Minghong, Peng Jiankun, et al. Miniature fiber-optic temperature sensor based on optical coating interference[J]. Optik, 2017(130): 1014-1020.

[14] Babadjian Jean Francois, Henao D. Reduced models for linearly elastic thin films allowing for fracture, debonding or delamination[J]. Mathematics, 2015, 18(4): 545-578.

作者简介：

高晓丹(1980—)，女，湖北鄂州人，硕士，副教授，主要从事光学薄膜及传感技术方面的研究。

E-mail：gxd6368@163.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.007

基于表面等离子体共振的光子晶体光纤折射率传感器的设计与分析

魏方皓，张祥军，唐守锋

（中国矿业大学 信息与控制工程学院，江苏 徐州 221000）

摘　要： 为了实现光子晶体光纤在近红外波段下的高灵敏度传感，设计了一种基于表面等离子体共振的光子晶体光纤（SPR-PCF）折射率型传感器。光纤内部的空气孔呈六边形排列，金纳米层完全包覆光纤外壁并与圆形待测物通道接触。利用有限元矢量软件COMSOL对SPR-PCF传感器的光学特性进行数值模拟仿真，得到不同待测物折射率的共振波长并绘制出纤芯损耗光谱，通过纤芯损耗光谱来对SPR-PCF传感器的传感特性进行分析。实验结果表明，其折射率测量区间为1.31～1.38，最大光谱灵敏度为10^4 nm/RIU，最大振幅灵敏度为200 RIU^{-1}，折射率测量精度为2.94×10^{-5} RIU。

关键词： 光子晶体光纤；表面等离子体共振；折射率传感；光纤传感

中图分类号： TN29　**文章编号：** 1001－5868(2020)01－0035－04

Design and Analysis of Photonic Crystal Fiber Refractive Index Sensor Based on Surface Plasmon Resonance

WEI Fanghao，ZHANG Xiangjun，TANG Shoufeng

（School of Information and Control Engineering，China University of Mining and Technology，Xuzhou 221000，CHN）

Abstract： A surface plasmon resonance based photonic crystal fiber（SPR-PCF）refractive index sensor with high sensitivity and wide refractive index measurement interval is proposed. Using surface plasmon resonance technology to measure the change of resonance wavelength through the fundamental mode loss spectrum to achieve the purpose of measuring different refractive indices of the substance to be tested. The main purpose is to arrange the air holes inside the fiber into a hexagonal shape，and cover the gold layer on the outer side of the fiber，and contact the circular material channel to be tested. It not only helps to accelerate the surface plasma reaction，but also has a simple structure and is easy to measure. The wavelength range of its operation can be adjusted. The simulation results of COMSOL show that the refractive index measurement interval is 1.31～1.38，and the maximum spectral sensitivity of 10 000 nm/RIU can be obtained. The resolution is 200 RIU^{-1}，and the refractive index accuracy is 2.94×10^{-5} RIU.

Key words： photonic crystal fiber；surface plasmon resonance；refractive index sensing；fiber optic sensing

0　引言

光子晶体光纤由于具有光学损耗小、光学非线性度高等许多独特的优点而被广泛应用于生物传感、药品检测、气体检测等诸多领域[1-2]。同时，随着表面等离子体共振现象被发现，基于表面等离子体共振的光子晶体光纤（Photonic Crystal Fiber Based on Surface Plasmon Resonance，SPR-PCF）传感器应运而生。SPR-PCF传感器与传统光纤传感器相比，交叉敏感、耦合损耗以及保偏特性等问题有了较

收稿日期：2019－10－23.

基金项目：国家重点研发计划项目（2017YFF0205500）.

大改进，而且可以制造多维结构，同时具有工作波长范围宽，模场面积大，可以实现多参数测量等一系列优点[3]。

2015 年，Wang 等人提出了一种 SPR-PCF 温度传感器，将混合液中不同形状的纳米金线加入到气孔呈八边形排列光子晶体光纤中。其灵敏度可以达到 500 pm/℃，分辨率可以达到 2.0×10^{-2}℃[4]。

2016 年，Gangwar 等人提出了一种外侧抛磨结构的 D 型 SPR-PCF 传感器（这种光纤结构侧面视图为一个平面，横截面视图近似为一个大写的 D），对于不同待测物（1.43～1.46）获得了折射率精度为 1.30×10^{-5} RIU，最大灵敏度为 7 700 nm/RIU 的 SPR-PCF 传感器[5]。

2017 年，Huang 提出了一种使用 TIO 作为金属镀层的 D 型 SPR-PCF 传感器，在 1.28～1.34 的折射率测量范围内平均灵敏度可达到 6 000 nm/RIU[6]。

2018 年，Liu 等人提出了一种 SPR-PCF 传感器，使用硅纳米棒外部镀金纳米层组成的混合金属纳米棒均匀分散在最内层空气孔内壁。工作波长范围为 0.66～3.1 μm，在 1.33～1.38 的传感范围中，最大灵敏度可以达到 4 111.4 nm/RIU，折射率精度为 2.45×10^{-5} RIU[7]。

本文提出了一种外层包覆金纳米层结构的高灵敏度 SPR-PCF 折射率传感器，并使用全矢量有限元仿真软件 COMSOL 对所提出的 SPR-PCF 折射率传感器的光学特性进行了数值模拟仿真分析。待测物的不同主要通过待测物通道内折射率的变化来体现。通过改变入射光波长来得到光纤纤芯损耗光谱，利用损耗光谱研究各种结构参数与待测物对 SPR-PCF 传感器传感特性的影响。研究的待测物折射率区间为 1.31～1.38。

1 理论分析

SPR-PCF 传感器设计平面图如图 1 所示。

空气孔的直径 $d=0.8$ μm，金纳米厚度 $t_g=45$ nm，空气孔之间的间距 $\Lambda=1.5$ μm。待测物通道厚度取值为 1 μm，背景材料为二氧化硅，其色散由 Sellmeier 方程得到[8]：

$$n^2(\lambda)=1+\frac{B_1\lambda^2}{\lambda^2-C_1}+\frac{B_2\lambda^2}{\lambda^2-C_2}+\frac{B_3\lambda^2}{\lambda^2-C_3} \quad (1)$$

其中，n 为熔融石英的波长相关折射率；λ 为输入光波长；B_1，B_2，B_3，C_1，C_2，C_3 是 Sellmeier 常数。

金纳米层的折射率可以使用 Drude-Lorentz 公式来进行推导[9]：

$$\varepsilon(\omega)=\varepsilon_1+i\varepsilon_2=\varepsilon_\infty-\frac{\omega_p^2}{\omega(\omega+\omega_c)} \quad (2)$$

其中，金属介电常数 ε_∞ 取值为 9.84，等离子体频率 ω_p 取值为 1.36×10^{16} rad/s，阻尼频率 ω_c 取值为 1.45×10^{14} rad/s。

纤芯损耗可以表述为[10]

$$L=8.686\times\frac{2\pi}{\lambda}\mathrm{Im}(n_{eff}) \quad (3)$$

其中，λ 表示入射光波长。同时为了避免光纤表面输出波对仿真结果产生影响，本文在模型中施加了一个圆形完美匹配层（Perfectly Matched Layer，PML）。利用有限元分析软件 COMSOL 进行了计算仿真。同时施加了一个圆形完美匹配层用来吸收光纤表面的输出波。

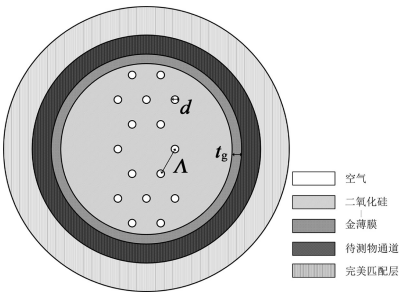


图 1 SPR-PCF 传感器设计平面图

2 模型参数优化

等离子体材料厚度的变化对 SPR-PCF 传感器传感特性的影响十分明显。通过图 2 的纤芯损耗光谱可知，当金纳米层厚度 t_g 由 40 nm 增大到 50 nm 时，共振波长向长波长方向移动，由 0.93 μm 增加到 0.95 μm，损耗峰值不断减少。当金纳米层厚度为 45 nm 时，其损耗为 218.05 dB/cm，共振波长为 0.94 μm，测量共振波长效果最佳。

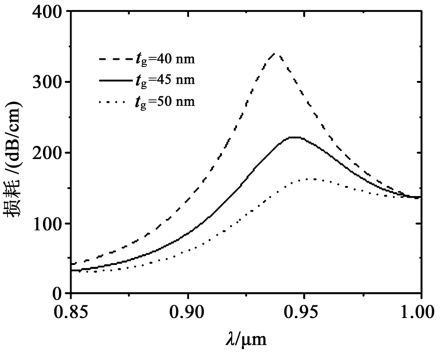


图 2 不同金纳米层厚度下的纤芯损耗光谱

空气孔直径 d 对 SPR-PCF 传感器传感特性具有十分重要的影响。空气孔的存在确保入射光可以在纤芯中进行传输，也可以满足基模与等离子模相位匹配条件。图 3 显示出当空气孔直径 d 增大时，损耗峰值不断下降，这是因为空气孔的增大抑制了光纤中表面等离子效应的产生，导致金纳米层周围模场减小，所以传输损耗不断减少，共振波长也发生了红移效应，当 d 为 0.72 μm 时，其共振效果较为明显。

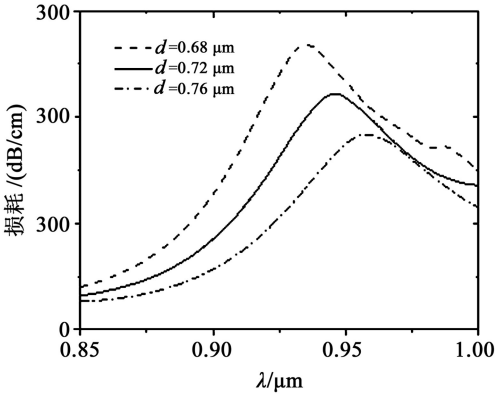


图 3　不同空气孔直径 d 下的纤芯损耗光谱

图 4 示出空气孔间距 Λ 变化时的纤芯损耗光谱。从图中可以看出，随着空气孔间距 Λ 不断增加，光纤纤芯损耗不断减少。主要原因为纤芯折射率与包层折射率差值减少，光在纤芯中被约束，能量不断被限制在纤芯中，使得纤芯损耗减少。当空气孔间距 Λ 为 1.5 μm 时，共振现象最为明显。

通过以上分析发现，当 $d=0.72$ μm，$\Lambda=1.5$ μm，$t_g=45$ nm 时，SPR-PCF 传感器传感特性最佳。

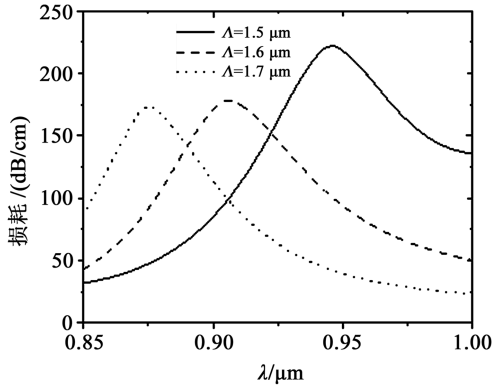
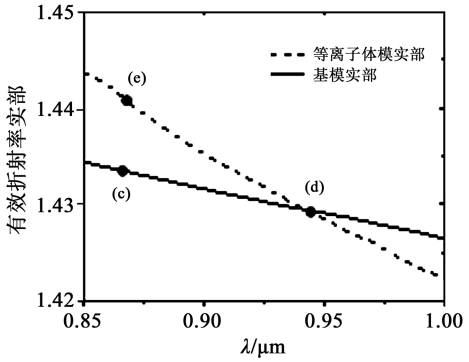


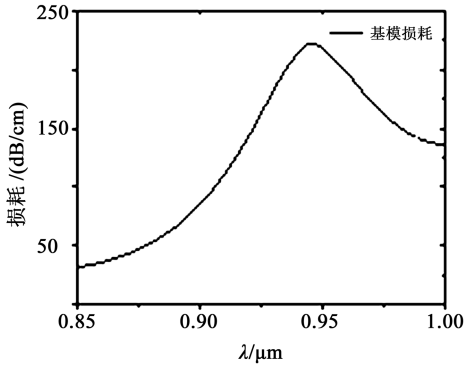
图 4　不同空气孔间距 Λ 下的纤芯损耗光谱

3　仿真结果分析

图 5(a)为待测物折射率为 1.37 时，光纤基模与等离子体模的折射率实部与入射光波长的关系图。图 5(b)显示了待测物折射率为 1.37 时光纤基模损耗谱。从图 5(a)可以明显看出，随着入射光波长的增加，等离子体模折射率实部的下降速率远远快于基模折射率实部。当入射光波长小于 0.94 μm 时，等离子体模的折射率实部大于基模折射率实部，此时光纤纤芯能量不断向金纳米层转移，如图 5(b)所示，光纤基模纤芯损耗逐渐升高。当入射波长达到 0.94 μm 时损耗达到最大值。此时基膜折射率实部与等离子体模折射率实部相等，发生了相位匹配，光纤纤芯能量大部分被金纳米层吸收。当入射波长继续增加时，基模折射率大于等离子模折射率，等离子体效应不断减弱，能量从金纳米层转移到纤芯中，损耗不断减少。图 5(c)～(e)分别表示光纤等离子体模、共振模、基模的仿真电场分布。



(a)　基模、等离子体模的折射率实部与入射波长关系图



(b)　光纤基模损耗光谱

(c)　基模场分布　　(d)　共振模场分布　　(e)　等离子模场分布

图 5　待测物折射率为 1.37 时，光纤基模与等离子体模的折射率实部与入射波长的关系图($d=0.72$ μm，$\Lambda=1.5$ μm，$t_g=45$ nm)

如图 6 所示，待测物折射率由 1.31 增加到 1.38 时，共振波长向长波长方向移动。这是由于待测物通道内待测物折射率变化引起光纤整体有效折射率实部的变化导致共振波长发生红移。从图中可以观察到，待测物折射率为 1.37 时，共振波长为 0.94 μm，损耗为 218.05 dB/cm；待测物折射率为 1.38 时，共振波长为 1.04 μm，损耗为 354.6 dB/cm，通过式(4)可以得到传感器最大灵敏度为 10 000 nm/RIU。

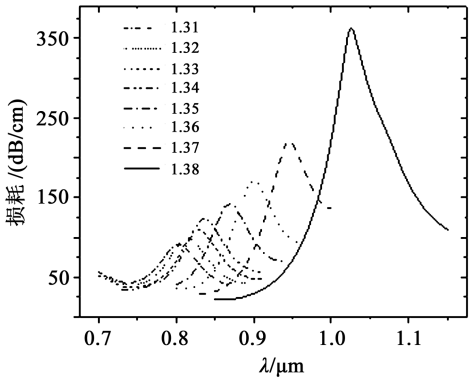


图 6 不同待测物折射率下的纤芯损耗光谱

光谱灵敏度公式可以表示为[11]

$$S(\lambda)=\frac{\Delta\lambda_{peak}}{\Delta n_a} \tag{4}$$

式中，$\Delta\lambda_{peak}$ 表示纤芯损耗光谱中共振波长随折射率变化的移动距离，Δn_a 表示待测物折射率的变化值。

折射率精度可以用下式计算：

$$R=\Delta n_a \Delta\lambda_{min}/\Delta\lambda_{peak} \tag{5}$$

其中，Δn_a 为折射率的变化量，默认取值 0.01，$\Delta\lambda_{min}$ 取值为 0.1 nm，$\Delta\lambda_{peak}$ 为光纤纤芯损耗光谱中共振波长平均移动距离，取值为 34 nm。可以得出折射率精度为 2.94×10^{-5} RIU。

通过图 6 可以得到共振波长与待测物折射率的依赖关系以及线性拟合曲线，如图 7 所示。

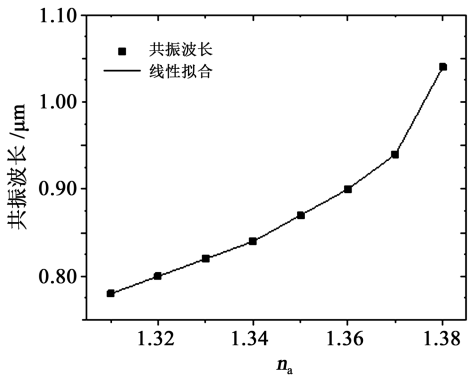


图 7 待测物折射率与共振波长的依赖关系图

此外，振幅灵敏度也是十分重要的性能指标，可以通过以下公式进行计算[12]：

$$A(\lambda)=\frac{1}{\alpha(\lambda,n_a)}\frac{\partial\alpha(\lambda,n_a)}{\partial n_a} \tag{6}$$

其中，$\alpha(\lambda,n_a)$为光纤纤芯损耗，$\partial\alpha(\lambda,n_a)$表示临近两待测物折射率损耗曲线在同一待测物折射率下的差值，∂n_a 为待测物折射率的变化。图 8 显示了待测物折射率为 1.37 时，SPR-PCF 传感器振幅灵敏度随波长的变化趋势。

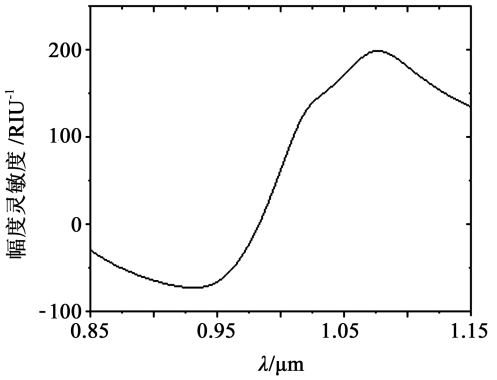


图 8 振幅灵敏度随波长变化的关系图

4 结论

本文提出并数值验证了一款易于制作的外包覆结构 SPR-PCF 传感器。在近红外波段实现了折射率范围 1.31～1.38 内的高灵敏度传感，最大光谱灵敏度可达到 10 000 nm/RIU，折射率测量精度可以达到 2.94×10^{-5} RIU，最大振幅灵敏度可以达到 200 RIU^{-1}。本设计可以应用于高灵敏度传感器。

参考文献：

[1] 薛渊泽，王学锋，唐才杰. 高温再生光纤光栅温度传感器封装技术[J]. 传感器与微系统，2019，38(5)：49-51.
Xue Yuanze，Wang Xuefeng，Tang Caijie. Packaging technique of high temperature regenerated fiber Bragg grating temperature sensor[J]. Transducer and Microsystem Technologies，2019，38(5)：49-51.

[2] Gasior K，Martynkien T，Wojcik G，et al. D shape polymer optical fibers for surface plasmon resonance sensing[J]. Opt. Electron. Rev.，2016，24(4)：209-215.

[3] 冯新星. 基于金纳米层或液体填充光子晶体光纤的偏振滤波器和传感器研究[D]. 秦皇岛：燕山大学，2018.
Feng Xinxing. Polarization filters and sensors based on gold nanowires or liquid-filled photonic crystal fibers [D]. Qinhuangdao：Yanshan University，2018.

(下转第 43 页)

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.008

基于遗传算法的太阳能飞艇光伏阵列优化研究

刘晓光，汪岳峰

（陆军工程大学 石家庄校区 电子与光学工程系，石家庄 050003）

摘　要： 太阳能飞艇依赖于光伏阵列提供能量，但曲面面形引起的非均匀光照会造成光伏阵列的能量损失。文章将遗传算法引入光伏阵列优化研究，通过优化光伏阵列拓扑结构，降低模块间的失配现象，从而达到增大输出功率的目的。首先基于光伏电池的单二极管模型，给出光伏阵列的仿真方法。其次提出基于遗传算法的光伏阵列拓扑结构的优化方法，采用序列编码方式，将光伏阵列拓扑结构转化为染色体编码，通过染色体基因位的交叉变异实现光伏阵列拓扑结构的优化。以某太阳能飞艇为例进行仿真验证，优化后光伏阵列平均输出功率增大2.32%，仿真结果表明所提方法能够有效提高太阳能飞艇光伏阵列的输出功率。

关键词： 太阳能飞艇；光伏阵列；遗传算法；拓扑优化；输出功率

中图分类号： TM914.4　　**文章编号：** 1001－5868(2020)01－0039－05

Genetic Algorithm-based Topology Optimization of Photovoltaic Array of Solar-powered Airships

LIU Xiaoguang，WANG Yuefeng

（Dept. of Electronic and Optical Engin.，Army Engin. University，Shijiazhuang 050003，CHN）

Abstract： Solar-powered airships rely on photovoltaic (PV) array to provide energy，but the non-uniform illumination caused by the curved surface leads to energy loss of PV array. Genetic algorithm is introduced into the research of PV array optimization. By optimizing the topology of the PV array，the mismatch losses between PV modules are reduced and therefore the purpose of increasing the output power is achieved. Firstly，based on the single diode model of PV cell，the simulation method of PV array is given. Secondly，an optimization method based on genetic algorithm for PV array topology is proposed. By coding sequence，PV array topology is converted into chromosome coding，and PV array topology is optimized by the crossover and mutation of chromosome. Taking a solar airship as an example，the simulation shows that the average output power of the PV array increases by 2.32% after optimization. The simulation results show that the proposed method can effectively improve the output power of the PV array.

Key words： solar-powered airship；photovoltaic array；genetic algorithm；topology optimization；output power

0　引言

太阳能飞艇采用柔性薄膜光伏电池作为能量来源[1-2]。由于单个电池的输出电压和电流很小，因此光伏电池需通过特定的拓扑结构连接成阵列，然后将光伏阵列粘贴在飞艇表面。由于飞艇外形为曲面，不同位置处太阳光的入射角不同，有效光强存在差异，造成光伏阵列表面光强的非均匀分布。

为解决非均匀光照的影响，Li等[3]以保持阵列输出功率和重量平衡为约束条件，提出一种光伏阵列最佳布置区域的优化方法，其中优化算法采用遗传算法和Nelder-Mead单纯形算法。该课题组在文献[4]中提出一种在保持光伏阵列面积不变的前提下阵列布局的优化方法。其优化方法是将飞艇中央部位简化为圆柱体，从而将光伏阵列的空间布局通

收稿日期：2019－11－17.

过阵列圆心角和长度参数化，并通过遗传算法完成参数优化过程。可以看出，以上两种方法虽然优化目标不同，但都是对阵列的空间布局进行优化。需要注意的是，非均匀光照造成能量损失的主要原因是光伏模块间电流和电压的失配[5]。失配现象发生时，光伏模块的工作点偏离自身最大功率点，从而造成能量损失。而对光伏阵列空间布局的优化并不能解决模块间的失配问题，因此本文将遗传算法引入太阳能飞艇光伏阵列的优化研究，通过优化光伏阵列的拓扑结构，达到降低模块间失配损失的目的。

1 光伏阵列仿真

常见的光伏电池等效模型包括单二极管模型和双二极管模型[6-7]。其中，单二极管模型结构相对简单，精度较高，能够满足工程需要。因此本文仿真采用单二极管模型，如图 1 所示。

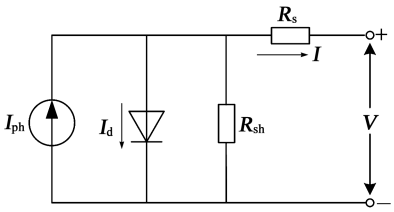


图 1 光伏电池单二极管模型

根据光伏电池单二极管模型可知，光伏电池输出电流为

$$I = I_{ph} - I_0\left\{\exp\left[\frac{q(V+IR_s)}{nkT}\right]-1\right\} - \frac{V+IR_s}{R_{sh}} \tag{1}$$

式中，I_{ph} 为光生电流，I_d 为流经二极管的电流，I_0 为反向饱和电流，R_s 和 R_{sh} 分别为串联电阻和并联电阻，q 为电子电量，n 为理想因子，k 为玻尔兹曼常数，T 为温度。为便于后续对光伏模块和阵列的仿真，通过朗伯 W 函数可将式(1)表示为电压的显函数形式[8]：

$$V = (I_{ph}+I_0)R_{sh} - (R_s+R_{sh})I - \frac{nkT}{q}\text{Lambert}(W) \tag{2}$$

式中，Lambert 表示朗伯 W 函数，参数 W 表示为

$$W = \frac{qR_{sh}I_0}{nkT}\exp\left[\frac{qR_{sh}(I_{ph}+I_s-I)}{nkT}\right] \tag{3}$$

光伏电池在连接成阵列时，首先将若干电池和旁路二极管连接成模块，如图 2 所示。

从图 2 可以看出，同一模块内，所有电池串联，并与旁路二极管并联。光伏模块构成阵列时，先将一定数量的模块串联形成光伏串，满足电压的需求；然后多个光伏串并联，满足电流的需求。当光伏串的电流 I 小于模块内部电池的短路电流时，光伏电池工作在正偏状态，可以向外输出功率，此时旁路二极管为反向偏置，无法导通，没有电流通过。当 I 大于模块内部电池的短路电流时，光伏电池将工作在反偏状态，无法向外输出功率，此时旁路二极管为正向偏置，旁路二极管导通，电流从旁路二极管通过。因此，光伏模块的电压可表示为

$$V_s = \begin{cases} \sum_{i=1}^{N_c} V_i, & I < I_{sc} \\ \dfrac{n_D kT}{q}\ln\left(\dfrac{I}{I_{0D}}+1\right), & I > I_{sc} \end{cases} \tag{4}$$

式中，V_i 表示模块内单个电池的电压，V_d 表示旁路二极管的电压，N_c 表示模块内电池的数目，I_{sc} 表示串联的光伏电池的短路电流，I_{0D} 和 n_D 分别表示旁路二极管的反向饱和电流和理想因子。根据模块和光伏串的连接方式可知，光伏串的电压等于所有模块的电压和，流经所有模块的电流相等，即：

$$\begin{cases} V_{string} = \sum_{i=1}^{N_s} V_s \\ I_{string} = I_1 = \cdots = I_{N_s} \end{cases} \tag{5}$$

式中，V_{string} 和 I_{string} 分别表示光伏串的电流和电压，N_s 表示光伏串内模块的数目。光伏阵列与光伏串的电压相等，电流为所有光伏串的电流和。因此，计算光伏阵列的电流时，在同一电压区间上通过插值法计算光伏串电流关于电压的函数 $I'_{string} = g(V'_{string})$，式中 V'_{string} 和 I'_{string} 分别表示电压采样点和拟合的电流点，则所有光伏串电流和即为阵列电流。

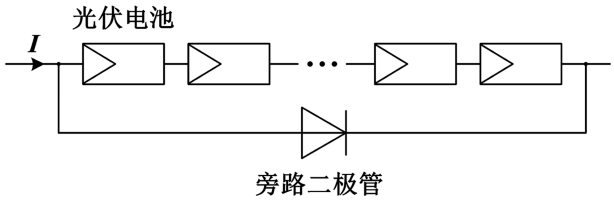


图 2 光伏模块示意图

2 基于遗传算法的光伏阵列拓扑优化

遗传算法通过模拟自然进化过程寻找最优解，被广泛应用于求解数值优化问题，其主要步骤包括染色体编码、适应度计算、染色体选择、染色体交叉和变异、算法终止[9]。

(1)染色体编码与适应度计算

通过遗传算法优化光伏阵列拓扑结构的关键是将拓扑结构转化为染色体编码。假设光伏阵列中共有 N_m 个模块，对所有模块进行编号，编号范围为 1～N_m。定义染色体为 1～N_m 的全排列，从染色体中依次取 N_s 个模块表示一个光伏串。定义光伏阵列的适应度函数为输出功率的平均值，即：

$$F=\frac{\sum_{i=1}^{N_{shade}}P_i}{N_{shade}} \tag{6}$$

式中，N_{shade} 表示光强分布的采样数目，P_i 表示在特定光照模式下光伏阵列的输出功率。

(2)染色体选择

在进行染色体选择操作时，首先采用轮盘赌法选择 $N_1=\tau N_{chro}$ 个染色体($\tau<1$)，重新生成 $N_2=(1-\tau)N_{chro}$ 个新个体，以保证种群的多样性，防止算法早熟并增强全局搜索能力。其中，N_{chro} 表示染色体种群的个体数目。采用轮盘赌法选择染色体作为父代染色体，任何一个染色体被选择的概率与其适应度值的大小成正比。在运行算法初期，不同染色体适应度的差值比较大，适应度值大的个体容易被选择。而在运行算法末期，染色体间的适应度值差值较小，优势个体不突出，收敛速度变慢甚至难以收敛。因此，为增大染色体适应度值的差异，便于选出优势个体，通过式(7)对染色体的适应度值进行拉伸：

$$F_i=\frac{F_i-\min(F)}{\max(F)-\min(F)} \tag{7}$$

式中，F 为所有染色体适应度值的集合，F_i 表示某条染色体的适应度值。

(3)染色体交叉和变异

交叉与变异对于保持染色体的多样性和优化的准确性具有重要意义。交叉与变异发生与否取决于交叉概率 P_c 和变异概率 P_m。交叉和变异概率对于遗传算法的收敛性有重要影响，当变异概率大时，染色体的多样性增大，但收敛性变弱。而当概率太小时，染色体搜索空间太小，收敛变慢。因此，为得到更好的收敛效果，采用自适应遗传算法的方法来确定交叉和变异概率[10]，则交叉和变异概率表示为

$$P_{ci}=\begin{cases}\dfrac{P_{c1}(\bar{F}-F_i)+P_{c2}[F_i-\min(F)]}{\bar{F}-\min(F)}, & F_i\leqslant\bar{F}\\ \dfrac{P_{c2}[\max(F)-F_i]+P_{c3}(F_i-\bar{F})}{\max(F)-\bar{F}}, & F_i>\bar{F}\end{cases}$$

$$P_{mi}=\begin{cases}\dfrac{P_{m1}(\bar{F}-F_i)+P_{m2}[F_i-\min(F)]}{\bar{F}-\min(F)}, & F_i\leqslant\bar{F}\\ \dfrac{P_{m2}[\max(F)-F_i]+P_{m3}(F_i-\bar{F})}{\max(F)-\bar{F}}, & F_i>\bar{F}\end{cases} \tag{8}$$

式中，F_i 为当前染色体的适应度值，$\bar{F}$ 为染色体群适应度值的平均值，$P_{c1}>P_{c2}>P_{c3}$ 且 $P_{c1},P_{c2},P_{c3}\in(0,1)$，$P_{m1}>P_{m2}>P_{m3}$ 且 $P_{m1},P_{m2},P_{m3}\in(0,1)$。

(4)算法终止

当优化函数的适应度值连续一定代数 N_{st} 保持不变或达到最大迭代次数 N_{max} 时，认为算法找到最优解，停止迭代，并以当前最佳染色体作为最终结果。太阳能飞艇光伏阵列拓扑优化的流程如图 3 所示，首先给出光伏电池参数及飞艇几何形状，其次确定任务周期确定光照模式的采样数目，计算每个采样点下光伏阵列的光强分布，在特定光照模式下通过上一节的仿真方法计算光伏阵列的输出功率。最后，通过遗传算法进行迭代优化，确定光伏阵列的最优拓扑。

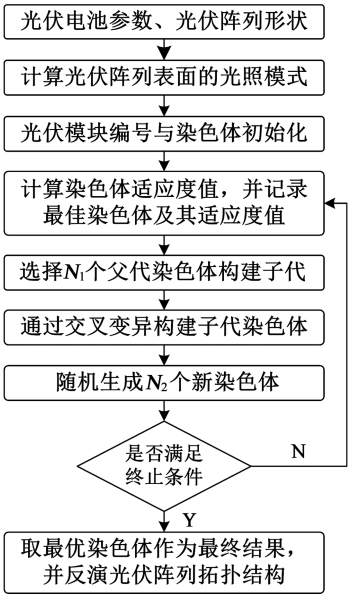


图 3　光伏阵列拓扑优化流程图

3　仿真分析

假设某飞艇的形状呈椭球形，其外形函数为：$x^2+y^2/5+z^2=1$。飞艇长度为 20 m，太阳电池几何尺寸为 10 cm × 10 cm，电学参数为：$I_{ph0}=3.27$ A，$I_0=2.5\times10^{-6}$ A，$R_s=0.003\ \Omega$，$R_{sh}=15\ \Omega$，$n=1.7$。旁路二极管参数为：$I_{0D}=5.5\times$

10^{-6} A，$n_D=1.5$。光伏阵列空间布局及模块内电池的连接方式如图 4 所示，初始状态 x 方向 8 个光伏串并联，每个光伏串内包含 12 个模块，每个模块包含 20 个电池。光伏阵列中心与飞艇顶部中心重合。假设光强每增大 100 W/m^2，温度升高 10 K。虽然与实际情况存在偏差，但不影响对优化算法性能的验证和分析。

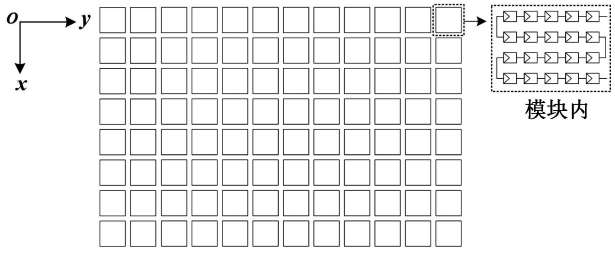


图 4 光伏阵列空间布局及模块内电池连接方式示意图

首先进行光照模式的采样，假设飞艇为定点驻留工作方式，所在纬度和经度分别为 40°N、116°E，海拔高度 20 km，任务时间为 $1\leqslant d_n\leqslant 173$，飞艇朝向为正东，$d_n$ 的采样步长为 20，标准时的采样步长为 0.5 h。光伏阵列共有 96 个模块，因此染色体是 1～96 的整数随机排序构成的向量。算法参数设置为：$\tau=0.9$，$N_{chro}=100$，P_{c1}、P_{c2}、P_{c3} 分别为 0.9、0.5、0.1，P_{m1}、P_{m2}、P_{m3} 分别为 0.1、0.05、0.005，且 $0.5\leqslant P_c\leqslant 0.9$，$0.05\leqslant P_m\leqslant 0.1$，$N_{st}=100$，$N_{max}=250$。运行算法所用计算机为 Intel@2.90 GHz 双处理器，256 G 内存，仿真环境为 Matlab 2014a。

优化过程中适应度值的变化曲线如图 5 所示。可以看出，随着迭代次数的增加，适应度值逐渐增大，并逐渐收敛于一个定值，说明优化过程结束得到最优结果。优化前后光伏阵列的拓扑结构示意图如图 6 所示。优化前光伏阵列在 y 方向 12 个模块构成一个光伏串，优化后相邻的 12 个相同模块构成一个光伏串。优化后光伏阵列的适应度值从 916.3 W 增大到 937.6 W，提高了 2.32%。

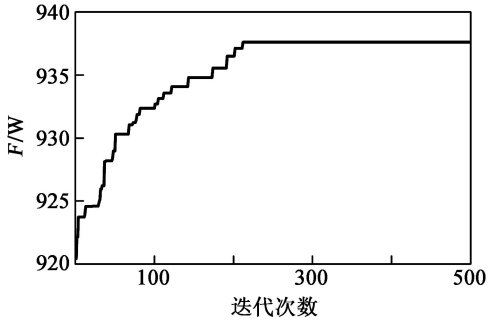
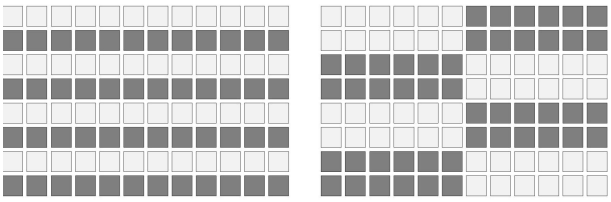


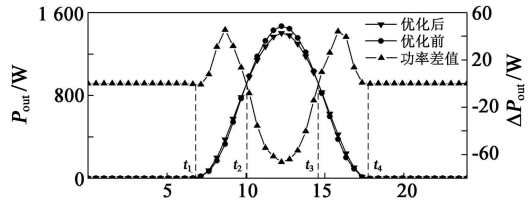
图 5 光伏阵列适应度函数变化曲线



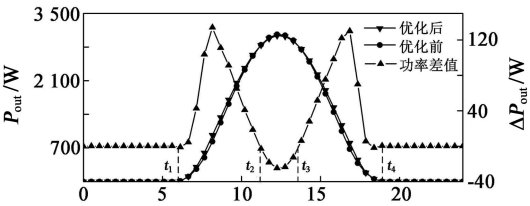
(a) 优化前 (b) 优化后

图 6 优化前后光伏阵列拓扑结构示意图

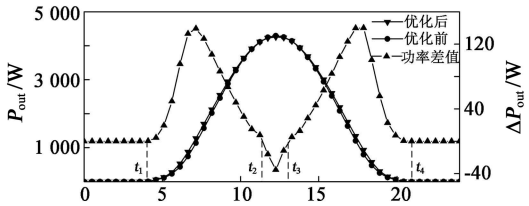
随机选取 3 个不同 d_n 取值，对比优化前后光伏阵列输出功率曲线，如图 7 所示。从图中可以看出，在 $t_1\sim t_2$ 和 $t_3\sim t_4$ 时间段内优化后光伏阵列输出功率(P_{out})增大，在 $t_2\sim t_3$ 时间段内优化后的光伏阵列输出功率降低。当 $d_n=1$ 时，$t_2\sim t_3$ 时间段内优化后光伏阵列较优化前输出功率明显降低，但在 $t_1\sim t_2$ 和 $t_3\sim t_4$ 时间段内优化后光伏阵列输出功率高于优化前，总体上优化前后光伏阵列的输出功率没有显著的变化。当 $d_n=80$ 和 173 时，除 $t_2\sim t_3$ 时间段内优化后光伏阵列输出功率略微降低外，在 $t_1\sim t_2$ 和 $t_3\sim t_4$ 时间段内优化后光伏阵列输出功率均显著大于优化前光伏阵列，从而实现优化后光伏阵列输出功率增大。这表明随着太阳高度角的增大，优化后光伏阵列的性能提升效果更加显著，从而保证了优化后光伏阵列失配损失降低。



(a) $d_n=1$



(b) $d_n=80$



(c) $d_n=173$

图 7 不同 d_n 值下优化前后光伏阵列输出功率曲线

4　结论

针对太阳能飞艇表面光强分布不均匀而造成光伏阵列能量损失的问题，提出一种基于遗传算法的光伏阵列拓扑结构优化方法，通过优化实现光伏模块间电流和电压的最佳匹配，从而增大光伏阵列输出功率。该方法仅通过优化光伏阵列拓扑结构即可降低非均匀光照的影响，不需要额外设备，方法简单，实用性较强。

参考文献：

[1] 朱炳杰，杨希祥，麻震宇，等. 平流层飞艇太阳电池系统产能分析[J]. 国防科技大学学报，2019，41(1)：14-18.
Zhu Bingjie，Yang Xixiang，Ma Zhenyu，et al. Power analysis of stratospheric airship's solar array system[J]. J. of National University of Defense Technol.，2019，41(1)：14-18.

[2] 赵　达，刘东旭，孙康文，等. 平流层飞艇研制现状、技术难点及发展趋势[J]. 航空学报，2016，37(1)：45-56.
Zhao D，Liu D，Sun K，et al. Research status，technical difficulities and development trend of stratospheric airship[J]. Acta Aeronautica et Astronautica Sinica，2016，37(1)：45-56.

[3] Li J，Lv M Y，Sun K W. Optimum area of solar array for stratospheric solar-powered airship [J]. Proc. of the Institution of Mechanical Engineers Part G-J. of Aerospace Engin.，2013，231(14)：1-12.

[4] Lv M Y，Li J，Du H F，et al. Solar array layout optimization for stratospheric airships using numerical method[J]. Energy Conversion and Management，2017，135：160-169.

[5] Bai J，Cao Y，Hao Y，et al. Characteristic output of PV systems under partial shading or mismatch conditions[J]. Solar Energy，2015，112：41-54.

[6] Kim W，Choi W. A novel parameter extraction method for the one-diode solar cell model[J]. Solar Energy，2010，84：1008-1019.

[7] Humada A M，Hojabri M，Mekhilef S，et al. Solar cell parameters extraction based on single and double-diode models：A review[J]. Renewable and Sustainable Energy Reviews，2016，56：494-509.

[8] Batzelis E I，Kampitsis G E，Papathanassiou S A，et al. Direct MPP calculation in terms of the single-diode PV model parameters[J]. IEEE Trans. on Energy Conversion，2015，30(1)：226-236.

[9] 蔚承英，王储君，刘焕淋. 基于光森林的弹性光网络能效组播路由频谱分配策略[J]. 半导体光电，2017，38(5)：719-724.
Wei Chengying，Wang Chujun，Liu Huanlin. Energy-efficient routing and spectrum allocation strategy based on optimal forecast for multicast in elastic optical networks [J]. Semiconductor Optoelectronics，2017，38(5)：719-724.

[10] Yan M F，Hu H，Otake Y，et al. Improved adaptive genetic algorithm with sparsity constraint applied to thermal neutron CT reconstruction of two-phase flow [J]. Measurement Science and Technol.，2018，29(5)：055404.

作者简介：

刘晓光(1990－)，男，博士研究生，主要从事光伏系统设计与应用方面的研究。

E-mail：graduatelxg@163.com

(上接第 38 页)

[4] Wang Mintuo，Lu Ying，Hao Congjing，et al. Simulation analysis of a temperature sensor based on photonic crystal fiber filled with different shapes of nanowires[J]. Optik-Inter. J. for Light and Electron Optics，2015，126(23)：3687-3691.

[5] Gangwar R K，Singh V K. Study of highly birefringence dispersion shifted photonic crystal fiber with asymmetrical cladding[J]. Optik-Inter. J. for Light and Electron Optics，2016，127(24)：11854-11859.

[6] Huang T. Highly sensitive SPR sensor based on D-shaped photonic crystal fiber coated with indium tin oxide at near-infrared wavelength[J]. Plasmonics，2017，12(3)：583-588.

[7] Liu Baolin，Lu Ying，Yang Xianchao，et al. Tunable surface plasmon resonance sensor based on photonic crystal fiber filled with gold nanoshells[J]. Plasmonics，2018，13(3)：763-770.

[8] Marquez-Cruz V. High resolution NIR TFBG assisted biochemical sensors[J]. Lightwave Technol.，2015，33(16)：3363-3373.

[9] 王　猛. 光子晶体光纤 SPR 气体传感机理及关键技术研究[D]. 北京：中国矿业大学，2019.
Wang Meng. Mechanism and key technology of photonic crystal fiber SPR gas sensing[D]. Beijing：China University of Mining and Technol.，2019.

[10] Gomes A D. Microfiber knot with taper interferometer for temperature and refractive index discrimination[J]. IEEE Photon. Technol. Lett.，2017，29：1517-1520.

[11] 李宏伟. D 型光子晶体光纤表面等离子体传感机理研究[D]. 北京：中国矿业大学，2019.
Li Hongwei. Study on surface plasma sensing mechanism of D-type photonic crystal fiber[D]. Beijing：China University of Mining and Technol.，2019.

[12] 翁思俊. 基于表面等离子体共振及多包层光纤的光纤器件研究[D]. 北京：北京交通大学，2018.
Weng Sijun. Research on fiber optic devices based on surface plasmon resonance and multicladding fiber[D]. Beijing：Beijing Jiaotong University，2018.

作者简介：

魏方皓(1994－)，男，硕士研究生，主要研究方向为光子晶体光纤传感器。

E-mail：292384747@qq.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.009

基于石墨烯探测器的积分电路设计

任思伟[1]，申　钧[2]，刘昌举[1]

（1. 重庆光电技术研究所，重庆 400060；2. 中国科学院重庆绿色智能技术研究院，重庆 400714）

摘　要： 为了解决石墨烯探测器光电探测的难题，针对石墨烯探测器受光照像元电阻发生变化的特点，设计了基于石墨烯探测器的新型积分电路结构。该积分电路结构主要包含前端偏置电路、运算放大器以及开关和反馈电容等部分，电路主要利用暗像元电阻不随光照变化，而感光像元电阻会随光照强度变化而变化的特点，将光照条件下暗像元支路的电流与感光像元支路电流的差作为光响应电流，并采用 CTIA 积分电路进行电流积分，将光响应电流转换为积分电压输出，进而实现石墨烯探测器对光响应信号的探测和输出。文中对相关的主要电路设计进行了分析，并基于 Cadence ADE 仿真环境完成了电路仿真。经仿真分析，基于文中的积分电路，可以将不同光照条件下石墨烯探测器的光响应转换为对应的积分电压输出。可见，所设计的积分电路能够满足石墨烯探测器对光响应探测的需求，对石墨烯材料进入光电探测器领域具有重要意义。

关键词： 石墨烯；探测器；近红外；积分电路

中图分类号： TN15　**文章编号：** 1001－5868(2020)01－0044－06

Integral Circuit Design Based on Graphene Detector

REN Siwei[1]，SHEN Jun[2]，LIU Changju[1]

（1. Chongqing Optoelectronics Research Institute，Chongqing 400060，CHN；
2. Chongqing Institute of Green and Intelligent Technol.，Chinese Academy of Sciences，Chongqing 400714，CHN）

Abstract： In order to solve the problem of photoelectric detection of graphene detectors，a new integrated circuit structure based on graphene detectors was designed according to the change of resistance of illuminated pixels of graphene detectors. The integrated circuit structure mainly includes front-end bias circuit，operational amplifier，switch and feedback capacitor. The circuit mainly utilizes the characteristics that dark pixel resistance does not change with illumination，but the resistance of photosensitive pixel changes with illumination intensity. The current of photosensitive pixel branch under illumination condition changes. And the difference of current from photosensitive pixel branch to dark pixel branch is taken as the light response current，and the CTIA integral circuit is used to integrate the current. The light response current is converted to the integral voltage，and then the detection and output of the light response signal of the graphene detector are realized. In this paper，the related main circuit design is analyzed，and the circuit simulation is completed based on the cadence ADE simulation environment. The simulation results show that the photoresponse of graphene detector under different illumination conditions can be converted to the corresponding integrated voltage with the integrated circuit in this paper. Therefore，the designedintegrated circuit can meet the needs of the graphene detector for the detection of light response，and it is of great significance for graphene materials to enter the field of photodetectors.

Key words： graphene；detector；near-infrared；integration circuit

收稿日期：2019－10－16.

0　引言

近年来，近红外探测器的应用变得越来越广泛，尤其在微光夜视、装备预警等军事领域，以及在夜视安防、汽车自动雷达和雾霾穿透等民用领域具有不可替代的作用。材料是决定探测性能的重要因素。其中，石墨烯具有优异的光电性能，在室温下具有超高的载流子迁移率，是制作非致冷红外探测器的理想材料。

一方面，石墨烯的零带隙特性决定了在相当宽频率范围内的光子都能激发石墨烯内电子的竖直跃迁，从而使得单层石墨烯对可见光到中远红外光的吸收率都保持在 2.3%，有利于实现宽光谱探测。目前已有实验报道表明其吸收光谱的范围从可见光一直到远红外。另一方面，初步研究表明石墨烯中存在载流子倍增效应(Carrier Multiplication)，即高能光子能够激发多对载流子的产生，使得高能光子的能量能够得到充分利用，结合石墨烯在室温下的超高载流子迁移率(200 000 $cm^2/(V \cdot s)$)，有利于实现超高的载流子增益，从而在室温下实现超高的光感应度。

通常，在多数红外技术应用中[1]，目标的红外辐射非常微弱，探测器的光生电流(通常为纳安甚至皮安量级)或光生电压(微伏量级)都非常小，加之各种噪声的干扰，目标信号几乎完全淹没在各种噪声中，故必须对输入信号进行适宜的处理(如积分、放大和降噪等)才能从噪声中提取出目标信号，这就是读出电路要完成的主要工作[2]。

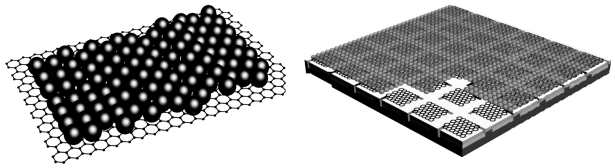
但是，由于石墨烯器件有别于普通光电二极管工作原理，石墨烯器件的特性决定了其不能自发产生光电流，因此，无法采用普通的 CTIA 积分电路结构实现电流积分。为了将石墨烯器件对光照强度的响应进行转换输出，针对石墨烯探测器积分电路结构的研究对石墨烯材料进入光电探测器领域具有重要意义。

1　石墨烯光电探测器

国内关于石墨烯光电探测器的研究主要集中在紫外、可见和近红外波段。而对于不同的探测结构，其特点也有所不同。

目前石墨烯红外探测器需要解决的主要问题是单层石墨烯光吸收不足。为了提高探测器的光吸收，国际上比较主流的解决方案主要包括光波导集成、光学谐振腔集成、量子点增强、表面等离子体增强、双层石墨烯异质结等。光波导集成适用于光纤通信应用，但由于光导入困难而不适用于光探测成像领域；光学谐振腔的主要问题是宽谱响应的丧失、吸收增强效果较弱以及工艺比较复杂；表面等离子体增强方案由于对石墨烯费米面的依赖而不能被单独使用；而石墨烯量子点复合探测器结构则能解决现有探测器增益低和成本高的难题。

本文所采用的近红外探测器阵列结构如图 1 所示，其核心是“石墨烯＋PbS 量子点”复合材料结构的光电探测器。石墨烯材料具有高载流子迁移率和宽波段光电响应的优势和光吸收总量偏弱的劣势，而量子点材料具有光吸收强、量子效率高的优势和载流子迁移率低从而电子空穴对分离效率低的劣势，通过将石墨烯与量子点材料结合，优势互补，可形成高增益的红外光电探测器。从工艺实现上，首先利用标准 CMOS 工艺完成读出电路设计，并在读出电路像元阵列上预留通孔，然后在后续集成工艺中实现将读出电路与探测单元连接，最终完成石墨烯探测器的制作。



(a)　复合单元　　(b)　阵列

图 1　石墨烯量子点复合单元及阵列结构示意图

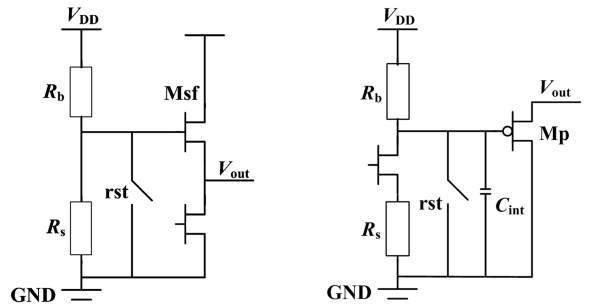
然而，与传统基于光电二极管的光伏探测器不同，“石墨烯＋PbS 量子点”复合结构属于光导型探测器，在光照条件下，材料中的电子空穴对增加，导致材料电导率发生改变，该复合结构可等效为一光敏电阻，即探测器的电阻值对光照敏感，会随光强的改变而改变。因此，石墨烯探测器积分电路结构也与传统光伏型探测器有所不同，本文将重点介绍针对石墨烯探测器信号的积分电路设计。

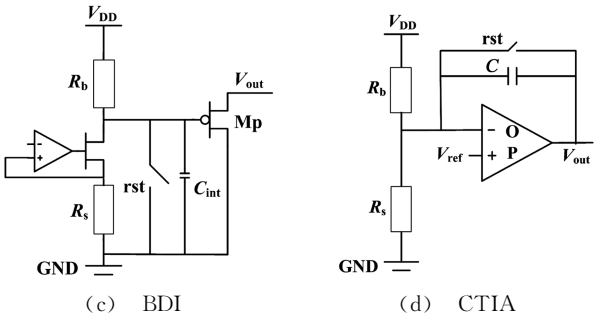
2　积分电路设计与仿真

2.1　典型积分电路结构

从信号注入的方式上，可将非致冷红外探测器前端偏置电路的结构分为几类：源跟随器结构(SFD)、直接注入结构(DI)、缓冲直接注入结构(BDI)、电容反馈跨导放大结构(CTIA)和电阻反馈跨导放大结构(RTIA)[3]，如图 2 所示。电路结构性

能对比见表 1[4]。





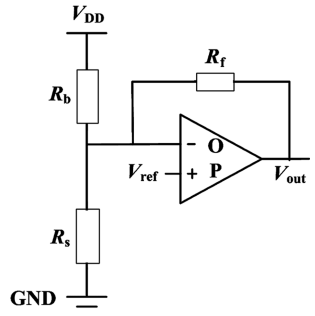


图 2 几类典型的积分电路结构

表 1 几种典型积分电路结构性能对比

电路结构	动态范围	NEI噪声	均匀性	线性度	频率特性	功率	最大通量	单元面积	成本/单元
SFD	低	中	好	好	很高	中	低	很小	低
DI	中	中	差	好	背景相关	低	中	很小	低
BDI	高	低	好	好	中/高	高	中	大	中
CTIA	高	低	好	很好	高	高	大	中	中
RTIA	高	低	好	很好	低	高	很大	很大	高

通过对以上积分电路结构的比较和分析[5]可知，CTIA 结构的性能最好，这种结构也是目前红外探测器读出电路最常用的结构。另外，结合石墨烯探测器的特点，CTIA 结构也能够满足其输出信号需要较好线性度的要求。

2.2 积分电路设计

本文所采用 CTIA 型积分电路结构如图 3 所示，主要由前端偏置电路、运算放大器、复位开关和反馈电容构成[6]。

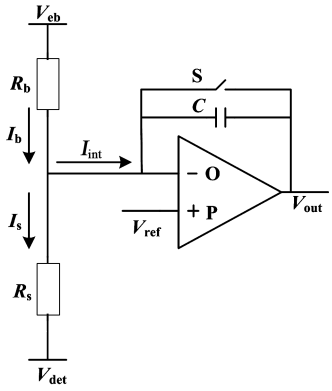


图 3 CTIA 型积分电路结构

从图 3 所示原理图可以看出，探测器正常工作需要三个偏置电压：V_{eb}，V_{det} 和 V_{ref}，经分析可知，输出电压 V_{out} 与这三个偏压的关系可以用式(1)表示：

$$V_{out}=V_{ref}-\frac{T_{int}}{C}\left[\frac{V_{eb}}{R_b}+\frac{V_{det}}{R_s}-\left(\frac{V_{ref}}{R_b}+\frac{V_{ref}}{R_s}\right)\right] \quad (1)$$

其中，偏置电压的稳定性对于响应的准确性极为重要，也是探测器组件设计中极为关键的部分。同时，由于探测器的工作环境会因实际应用的改变而改变，对这些偏压进行设计时需要保证探测器的输出在动态范围之内，因此，必须把偏置电压设计为可调偏压。

2.2.1 前端偏置电路设计

前端偏置电路主要用于为后级 CTIA 运算放大器提供相应的工作点偏置以及信号电流的输入。图 3 中，R_b 和 R_s 分别表示石墨烯探测器的暗像元和感光像元电阻，其中，感光像元受光照后电阻值发生变化，由于运放工作时的“虚短”特性，暗像元和感光像元两端的电压基本保持不变，因此，暗像元支路电流 I_1 保持不变，而感光像元电阻的变化将引起感光像元支路电流 I_2 的变化，进而引起流入 CTIA 运放的信号电流 I_3 发生相应变化，最终通过积分电路把这种光照引起的信号电流进行积分放大并输出。

根据欧姆定律，可以将 $I_b(t)$ 和 $I_s(t)$ 写为

$$I_b(t)=V_b/R_b \quad (2)$$

$$I_s(t)=V_s/R_s \quad (3)$$

式中，V_b 和 V_s 分别表示暗像元和感光像元上的电压降，R_b 和 R_s 分别表示暗像元和感光像元电阻。

根据基尔霍夫电流定律，积分电流 $I_{int}(t)$ 为

$$I_{int}(t)=I_b(t)-I_s(t)=\frac{R_sV_b-R_bV_s}{R_bR_s} \quad (4)$$

通常地，为了消除石墨烯像元中的暗电流，在探测器制作时，暗像元和感光像元的工艺参数基本保持一致，$R_s=R_b$，适当调节 V_b 和 V_s，使得 $V_b=V_s=V$，然后对石墨烯感光像元施加适当光照条件，此时，感光像元的电阻值在光照后变化为 $R_l=R_s(1-\alpha)=R_b(1-\alpha)$，其中，对本文石墨烯探测器而言，$0\leqslant\alpha\leqslant0.1$，则积分电流为

$$I_{int}(t)=\frac{(R_l-R_b)V}{R_bR_l}=\frac{R_b(1-\alpha)-R_b}{R_b^2(1-\alpha)}\times V=\frac{\alpha}{R_b(1-\alpha)}\times V \tag{5}$$

令积分电路中的电容 C 的值为 C_{int}，积分时间为 T_{int}，运放正相端参考电压为 V_{ref}，则石墨烯探测器积分电压输出为

$$V_{out}=V_{ref}-\frac{\alpha T_{int}}{C_{int}R_b(1-\alpha)}\times V \tag{6}$$

由上式分析可知，石墨烯探测器工作时信号输出情况如下：

(1)无光照条件下，$\alpha=0$，$V_{out}=V_{ref}$；

(2)当石墨烯探测器感光像元电阻值受光照减小时，$\alpha>0$，$V_{out}>V_{ref}$。

2.2.2　运算放大器设计

为了保证探测器的性能，CTIA 结构的积分器需要选择噪声低、增益高、失调小、摆幅大、功耗低、面积小、速度快的运算放大器[7]。

在几种常用的运放中，两级运放具有良好的性能，但由于需要利用补偿电容对运放的频率特性进行密勒补偿，补偿电容占用面积较大，对于每列需要使用一个积分放大器的探测阵列读出电路而言，其不利于芯片面积的减小；增益提高运放的功耗较高，套筒式共源共栅运放的输入/输出摆幅较小。通过折中考虑，本文选用以 PMOS 作为差分输入管的折叠式共源共栅(Folded Cascode)运放，以减小 $1/f$ 噪声和热噪声，同时采用负载电容补偿以增加电路稳定性，如图 4 所示。

折叠式共源共栅结构运放的直流增益 A_v 为

$$A_v\approx g_{m1,2}(g_{m3,4}r_{o3,4}r_{o9,10}\ /\!/\ g_{m5,6}r_{o5,6}r_{o7,8}) \tag{7}$$

其中，g_m 为晶体管的跨导，r_o 为晶体管的小信号输出电阻。由晶体管尺寸失配和阈值电压失配引起的失调可看作一个频率很低的噪声成分 $V_{os,out}$，为体现其对 CTIA 性能的影响，将其表示为输入参考失调电压 $V_{os,in}$，则：

$$|V_{os,in}|=|V_{os,out}|/A_v \tag{8}$$

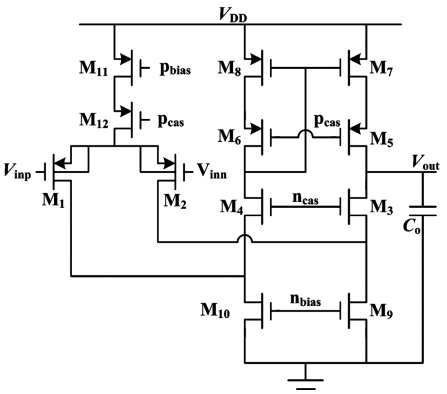


图 4　折叠式共源共栅运放结构

可见，提高增益可以减小输入参考失调电压。电路中的输入对 $M_{1,2}$ 阈值电压的失配会直接折合到输入参考失调电压中，所以要求输入对需要有很好的匹配度，在设计中可以通过增大晶体管面积的方式来减轻尺寸的失配，同时采用 $M_{1,2}$ 的衬底连接到其源端的方式来消除“体效应”对阈值电压的影响，以减轻输入对阈值电压的失配。则信号通路上单位带宽内总的等效输入噪声可写为

$$\overline{V_{n,in}^2}\approx2\left[\overline{V_{n1,2}^2}+\left(\frac{g_{m9}}{g_{m1}}\right)^2\overline{V_{n9,10}^2}+\left(\frac{g_{m7}}{g_{m1}}\right)^2\overline{V_{n7,8}^2}\right] \tag{9}$$

式中，$\overline{V_{n,in}^2}$ 是 M_i 晶体管的等效栅极输入噪声电压，包含了热噪声和闪烁噪声($1/f$ 噪声[11])，有：

$$\overline{V_{ni}^2}=\frac{8kT(1+\eta_i)}{3g_{mi}}+\frac{K_N}{C_{ox}(WL)_i}\frac{1}{f} \tag{10}$$

其中，$\eta_i=g_{mbi}/g_m$ 为 MOS 晶体管的体效应系数，K_N 为闪烁噪声系数，C_{ox} 是晶体管单位面积的栅氧化层电容，W 和 L 分别是晶体管的宽和长，f 为频率。由等效输入噪声的表达式，可以通过调节有相应贡献的晶体管的尺寸以及跨导来减小噪声，但尺寸的增大不利于芯片面积的减小，同时通过增大静态电流的方式来增大跨导，不利于功耗的减小，因此需要选择合适的晶体管尺寸和静态电流，在各个特性参数中进行折中。

2.2.3　开关的设计

通常，对开关电容积分器中的开关有以下设计要求：

(1)需要很高的断开电阻[8]来减少电荷泄露；

(2)电荷传输速度快，可以在半周期内完成电荷的转移，并且具有低的导通电阻；

(3)具有较大的输入信号摆幅，要求能够提供 ADC 的全摆幅输入导通。

其中，开关电容积分器中的开关采用的是 N 管和 P 管对接的传输门形式，电路结构图如图 5 所示。

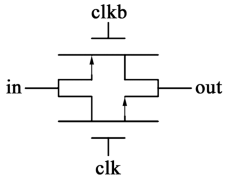


图 5 传输门电路结构

2.2.4 电容设计

在标准 CMOS 工艺技术中，电容需要较大的面积来实现，电容的大小直接影响电路的功耗、噪声、转换速率等。选用大电容可以降低系统噪声，选用小电容可以降低功耗、芯片面积等，IC 设计中对电容的选取存在折中关系。

开关电容积分器中的电容连接均采用下极板采样技术，即所有和运算放大器的输入端相连的电容极板均采用上极板，而采用电容下极板作为采样极板，结构如图 6 所示。因为若采用下极板连接运算放大器的输入端，则较大的寄生电容会造成运放输入端寄生电容过大而导致非线性。输入端寄生电容过大会降低运放的反馈系数，从而增大运放功耗和噪声，同时影响采样电容与反馈电容的匹配度，从而直接影响 ADC 的线性度。

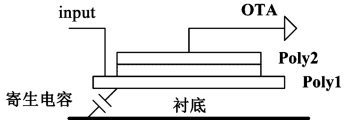


图 6 下极板采样原理图

2.3 仿真分析

根据以上设计分析，本文完成了针对石墨烯探测器的积分电路设计，并基于 Cadence ADE 仿真环境对上述电路进行仿真。

首先，在暗场无光照条件下，对运放输入电流进行调零以校准积分电路的初始工作条件，即令暗像元电阻 R_b 和感光像元电阻 R_s 均保持固定不变，取偏置电压 $V_{det}=0$ V，而对偏置电压 V_{eb} 从 3～5 V 进行参数扫描，CTIA 运放输入端积分电流 I_{int} 的扫描曲线如图 7 所示。从图 7 可以看出，当偏置电压 $V_{eb}=3.8$ V 时，积分电流 I_{int} 近似为零。因此，校准后，将偏置电压 V_{eb} 取 3.8 V，积分电路可以满足在暗场无光照条件下信号输出近似为零的要求，此时输出积分电压 V_{out} 近似为参考电压 V_{ref}。其中，运算放大器增益曲线如图 8 所示，运放低频增益为 16.76 dB，运放 0 dB 带宽为 82.83 MHz，满足运放的设计要求。

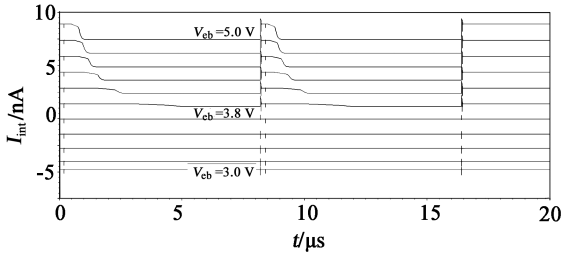


图 7 V_{eb}-I_{int} 扫描曲线

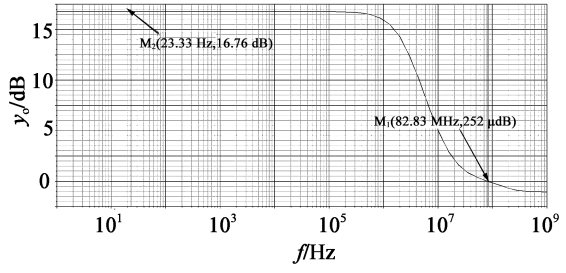


图 8 运算放大器增益曲线

然后，对积分电路进行功能仿真。在石墨烯探测器工作时，石墨烯感光像元因受到光照强度的不同会产生相应电阻值变化，即 R_s 的值随光强变化。根据石墨烯器件单元测试分析表明，R_s 的变化范围为 0%～10%。因此，在仿真时，光强的变化可等效

在不考虑管子体效应的情况下，开关导通时，CMOS 传输门的导通电阻可以表示为

$$R_{on}=R_{on,n}R_{on,p}=\frac{1}{u_nC_{ox}\left(\frac{W}{L}\right)_n(V_{DD}-V_{THN})}\times\frac{1}{u_pC_{ox}\left(\frac{W}{L}\right)_p(V_{DD}-|V_{THP}|)}=$$

$$\frac{1}{u_nC_{ox}\left(\frac{W}{L}\right)_n(V_{DD}-V_{THN})-\left[u_nC_{ox}\left(\frac{W}{L}\right)_n-u_pC_{ox}\left(\frac{W}{L}\right)_p\right]V_{in}-u_pC_{ox}\left(\frac{W}{L}\right)_p|V_{THP}|} \tag{11}$$

由上式可知，要使导通电阻 R_{on} 与输入电压 V_{in} 无关，就要满足：

$$u_nC_{ox}\left(\frac{W}{L}\right)_n=u_pC_{ox}\left(\frac{W}{L}\right)_p \tag{12}$$

此时，传输门的导通电阻与输入信号相对独立，开关不产生谐波失真。在设计时要尽量减少开关数量，减小开关尺寸。然后，经过仿真并考虑寄生电容，确定传输门中 NMOS 和 PMOS 的宽长比。

为对 R_s 的参数扫描。其中，随光照强度变化，石墨烯探测器的积分电流 I_{int} 与积分电压 V_{out} 曲线分别如图 9 和 10 所示。从图 9 可知，积分电流大小随光照强度增加在 47～541 nA 近似均匀地变化，光照越强积分电流越大，且电流方向表现一致，与设计预期相符。从图 10 可知，积分电压输出幅度随光强增大而增大，且积分电压曲线斜率随积分电流增大而增大，电路特性满足积分电路的设计要求。

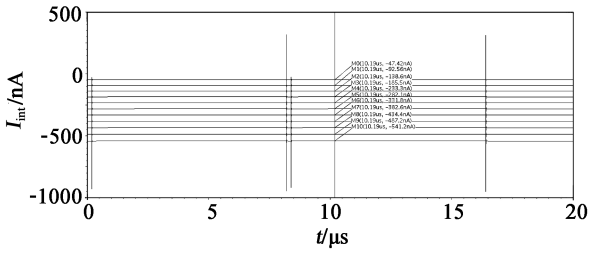


图 9　随光强变化的积分电流曲线

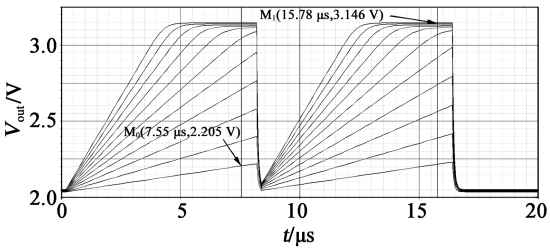


图 10　随光强变化的积分电压曲线

可见，经对积分电路的仿真分析，本文所设计的积分电路能够满足石墨烯探测器对光响应信号进行积分的设计要求。

3　结束语

本文对石墨烯探测器的探测原理进行了分析，并设计了基于石墨烯探测器的信号积分电路。该信号积分电路主要包含前端偏置电路、运算放大器以及开关和反馈电容。本文对相关电路进行设计和分析，并基于 Cadence ADE 仿真环境完成相关仿真。仿真分析结果表明，基于该积分电路，可以将不同光照条件下石墨烯探测器的光响应转换为对应的积分电压输出，因此能够满足石墨烯探测器对光响应探测的需求。

参考文献：

[1] 雷亚贵，王戎瑞，陈苗海. 国外非制冷红外焦平面阵列探测器进展[J]. 激光与红外，2007，37(9)：801-805.
Lei Yagui，Wang Rongrui，Chen Miaohai. Development of foreign uncooled IRFPA detectors[J]. Laser & Infrared，2007，37(9)：801-805.

[2] Hsieh C C，Wu C Y，Sun T P. A new cryogenic CMOS readout structure for infrared focal plane array[J]. IEEE Solid-State Circuits，1997，32(8)：1192-1199.

[3] 吕　坚. 非制冷红外焦平面读出电路的研究[D]. 成都：电子科技大学，2009：40-43.
Lv Jian. Study on uncooled infrared focal plane readout circuit [D]. Chengdu：University of Electronic Science and Technol. of China，2009：40-43.

[4] 张　沛. 低噪声非制冷红外焦平面阵列驱动电路的设计[J]. 红外与激光工程，2010，39(5)：807-810.
Zhang Pei. Design of the low noise uncooled infrared focal plane array driver circuit[J]. Infrared and Laser Engin.，2010，39(5)：807-810.

[5] Razavi B. 模拟 CMOS 集成电路的设计[M]. 西安：西安交通大学出版社，2003：316-319.
Razavi B. Design of Analog CMOS Integrated Circuits[M]. Xi'an：Xi'an Jiaotong University Press，2003：316-319.

[6] 薛惠琼. 非制冷红外探测器及其读出电路的研究[D]. 长春：长春理工大学，2009：35-41.
Xue Huiqiong. Study on uncooled infrared detector and its readout circuit[D]. Changchun：Changchun University of Science and Technol.，2009：35-41.

[7] Allen P E，Holberg D R. CMOS Analog Circuit Design[M]. New York：Holt Rinehart and Winston，1987：316-319.

[8] 林佳明. 一种 10 位 200 M Sample/s 模数转换器的设计[D]. 上海：上海交通大学，2008：23-24.
Lin Jiaming. Design of a 10 bit 200 M Sample/s AD converter [D]. Shanghai：Shanghai Jiaotong University，2008：23-24.

作者简介：

任思伟(1986－)，男，四川人，硕士，工程师，主要研究方向为图像传感器及数模混合信号电路设计。

E-mail：rswofxd@126.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.010

一种带温度参量的半导体激光器速率方程及仿真求解

张亚标，肖永川，王 超，孙力军
（重庆光电技术研究所，重庆 400060）

摘 要： 研究并求解了热效应条件下的半导体激光器的速率方程。在经典速率方程的基础上，通过引入热效应特性，得到了带温度参量修正的速率方程组，然后利用 Simulink 软件工具，对该非线性微分方程进行仿真求解，最后以某分布反馈激光器的参量为例，得到了不同偏置点、不同调制信号及不同温度条件下速率方程的瞬态解，并据此分析了弛豫振荡、频率啁啾等特性。仿真求解结果符合该型激光器的瞬态特性，为该型激光器设计制造和应用选型提供了一定的参考。

关键词： 半导体激光器；速率方程；瞬态特性；弛豫振荡；频率啁啾；Simulink
中图分类号： TN248.4 **文章编号：** 1001－5868(2020)01－0050－04

Semiconductor Laser Rate-Equation with Temperature Parameter and Emulation Solving

ZHANG Yabiao, XIAO Yongchuan, WANG Chao, SUN Lijun
(Chongqing Optoelectronics Research Institute, Chongqing 400060, CHN)

Abstract: Based on classic rate-equation, by introducing the thermal effect, the rate-equations with temperature parameter correction were obtained. Then by using Simulink software tool, the nonlinear differential equation was simulated. Finally taking the parameter of a distributed feedback laser as an example, the transient solution of the rate equation under different temperature, bias points, and modulation signal was obtained, and based on which, the characters such as relaxation oscillation and frequency chirp were analyzed. The simulation results match the transient performance of this type of laser, and can provide reference for laser design and applicaion.

Key words: semiconductor laser; rate-equation; transient performance; relaxation oscillation; frequency chirp; Simulink

0 引言

半导体激光器是当前光电子领域中活跃的研究课题，新材料与新器件结构的不断涌现，使得半导体激光器的性能不断提高。激光器的工作特性，包括工作点、功率、调制下的响应特性等都极大地影响光电系统的性能，瞬态响应是器件特性分析必不可少的过程。Matlab 环境下的 Simulink 仿真软件结合了图形界面和交互功能的动态建模、仿真和综合分析，可实现动态系统的直观建模，在系统层面的仿真和瞬态分析方面有独特的优势[1-2]。

本文在激光器速率方程的基础上，引入器件的热效应特性，得到带温度参量修正的速率方程组，然后利用 Simulink 进行仿真求解，通过引入某 DFB（分布反馈）激光器的典型参量，解析不同偏置点、不同调制信号及不同温度条件下激光器的载流子密度和光子密度随时间的瞬态过程，以及激光器弛豫振荡、频率啁啾特性，并分析它们的影响因素。

收稿日期：2019－08－25.

1 速率方程及热效应特性

1.1 速率方程

光子的自发发射和受激辐射过程都是典型的量子力学过程，从理论上讲，所有关于激光的问题都可以用全量子理论来处理，但是对于大多数的问题而

言，应用中会带来许多不必要的繁杂性，用速率方程理论足以描述半导体激光器的响应特性，它描述了光子和载流子之间的相互作用关系。

单模激光器有源区响应的基本方程（速率方程），可以表示为[3]

$$\frac{\mathrm{d}N}{\mathrm{d}t}=\frac{J}{qV_{\mathrm{a}}}-\frac{g(N-N_{\mathrm{tr}})S}{1+\varepsilon S}-\frac{N}{\tau_{\mathrm{N}}} \tag{1}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t}=\frac{\Gamma g(N-N_{\mathrm{tr}})S}{1+\varepsilon S}-\frac{S}{\tau_{\mathrm{P}}}+\Gamma\beta\frac{N}{\tau_{\mathrm{N}}} \tag{2}$$

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t}=\frac{\alpha}{2}\left[\Gamma g(N-N_{\mathrm{tr}})-\frac{1}{\tau_{\mathrm{P}}}\right] \tag{3}$$

其中，N 和 S 分别为载流子浓度和光子密度，g 是对光子密度的微分增益系数，J 是注入有源层的电流，q 是电子电荷，V_{a} 是有源层体积，N_{tr} 是透明载流子密度，τ_{N} 表示谐振腔中载流子寿命，τ_{P} 表示谐振腔中光子寿命，Γ 是光限制因子（谐振腔内所有光能中有源层内部所包含的比例），ε 是增益饱和因子也称为增益压缩因子，β 是自发辐射因子，Φ 是激光器出射光相位，α 是线宽增强因子。

上述方程式简单地描述了激光器谐振腔内载流子和光子的供给、产生和消失之间的相互关系，式(1)表示激光器内部载流子密度 N 随着时间和激励源的驱动电流成正比增加的同时，又由于光子的自发发射、受激复合辐射和非发光辐射复合而不断衰减；式(2)表示光子密度 S 的增加速率等于受激辐射光子的速率加上自发发射进入激光模式的速率，同时再减去因为谐振腔内各种损耗所引起的光子数减少的速率；式(3)表示光波的相位变化，它描述了频率啁啾特性。

1.2　热效应特性

温度对激光器的本征响应、阈值电流、输出功率和中心波长等有着显著的影响，关于激光器温度模型的建立和热效应影响的研究，业内已经做了大量的工作，通常是基于激光器的等效电路模型，通过精确建立温度模型并转换嵌入到等效电路模型之中，然后使用先进仿真分析软件进行模拟仿真和参数提取[4-6]。

Byrne 和 Keating 在文中讨论了温度对微分增益系数 g 和透明载流子密度 N_{tr} 的影响，可以表示为[7]

$$g(T)=g_0\times\exp(-T/T_0) \tag{4}$$

$$N_{\mathrm{tr}}(T)=N_{\mathrm{t0}}\times\exp(T/T_0) \tag{5}$$

其中，T 是激光器芯片内温度，T_0 是激光器特征温度。

上述速率方程的温度修正方法没有进一步考虑芯片内温度 T 的热力学变化。实际上芯片内温变化是一个动态平衡的过程，芯片内和外部环境存在着动态的热量传递和交换，热交换大小跟芯片封装的热阻、内外温差、输入、输出功率等参数都有关系，这个交换过程可以用热率方程表示[8]：

$$T=T_{\mathrm{amb}}+R_{\mathrm{th}}(P-P_0)-\tau_{\mathrm{th}}\frac{\mathrm{d}T}{\mathrm{d}t} \tag{6}$$

其中，T_{amb} 为环境温度，R_{th} 为激光器热阻，P 为输入总功率，P_0 为输出光功率，τ_{th} 为温度时间常量。

式(6)中，输出光功率 P_0 与光子密度成线性正比关系，其表达式为

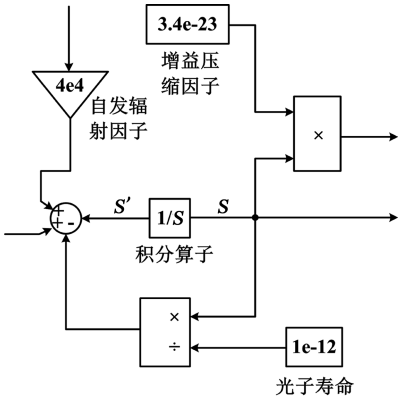
$$P_0=\frac{V_{\mathrm{a}}\eta h\nu}{\Gamma\tau_{\mathrm{P}}}S \tag{7}$$

其中，η 是微分量子效率，h 是普朗克常数，ν 是光波频率。

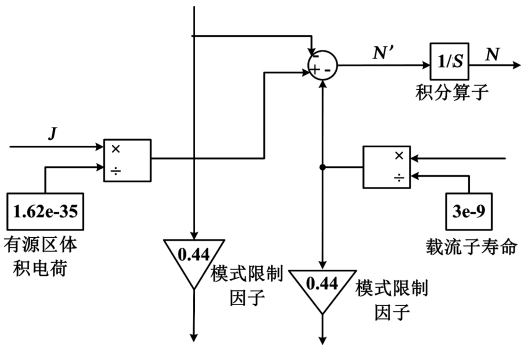
可见，联立式(4)～(7)和式(1)～(3)可以得到带温度参量修正的速率方程组。

2　基于 Simulink 的速率方程仿真求解

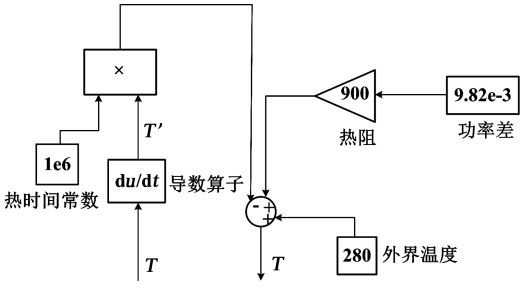
前述带温度修正的激光器速率方程组是一组非线性微分方程，在时域求解速率方程组时，存在由于 N 或 S 数值较大而导致无法收敛到稳态解的情况，利用 Simulink 可视化功能，各个模块的功能分别与速率方程组中的各项相对应，可以完整地仿真求解速率方程。这些模块包括光子和载流子、自发辐射、光限制因子、受激辐射、温度效应和各种损耗等模块，信号源是直流偏置叠加调制信号，通过在终端和对应节点设置输出模块 To Workspace 和示波器 Scope，可以快速方便地观察到激光器重要物理量瞬



(a)　关于光子密度 S 的非线性微分模型



(b) 关于载流子浓度 N 的非线性微分模型



(c) 关于温度 T 的非线性微分模型

图 1 带温度修正的激光器速率方程组的 Simulink 求解仿真模型图

时变化的全过程。基于 Simulink 的速率方程组求解仿真模型如图 1 所示。

3 仿真设定及分析

3.1 速率方程参量设定

在 Simulink 仿真中,速率方程涉及的物理量采用某 DFB 激光器的典型值,具体如表 1 所示。

表 1 速率方程参量及数值表

物理量	数值	物理量	数值
$g_0/(m^3/s)$	3×10^{-12}	Γ	0.44
τ_p/ps	1	N_{tr}/m^{-3}	1.2×10^{24}
τ_n/ns	3	$h/(J\cdot s)$	6.626×10^{-34}
q/C	1.6×10^{-19}	$R_{th}/(k/W)$	780
ε/m^{-3}	3.4×10^{-23}	$\tau_{th}/\mu s$	1
β	4×10^{-4}	α/m^2	5
V_a/m^3	9×10^{-17}	η	0.125

3.2 瞬态特性

取表 1 中所示 DFB 激光器的参量值,设定直流偏置电流分别为 10,20,30 和 40 mA,经 Simulink 仿真计算,得到不同偏置电流时的光子密度 S 瞬态变化曲线如图 2 所示,不同偏置电流时的载流子密度 N 瞬态变化曲线如图 3 所示。

3.2.1 光子密度瞬态分析

如图 2 所示,当偏置电流为 10 mA 时,光子密度近似于一条水平直线,其值稳定于 1.31×10^{17},此时激光器没有正常输出;当偏置电流为 20,30 及 40 mA 时,光子密度都有一个迅速升高的现象,升高所用时间大概为 0.11 ns,然后迅速回落,产生振荡并逐渐衰减到一个稳定值,不同直流偏置条件下的光子密度的稳定值分别是 2.55×10^{20},5.25×10^{20} 及 3.23×10^{20} m^{-3},可见激光器的稳定输出值随着直流偏置点线性增加。

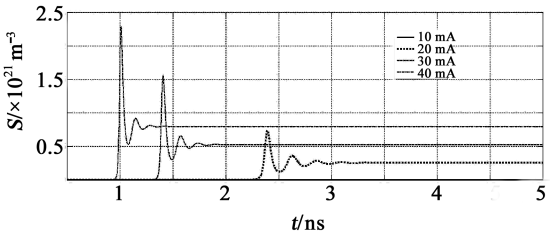


图 2 不同偏置电流时的光子密度瞬态图

3.2.2 载流子密度瞬态分析

如图 3 所示,当偏置电流为 20,30 和 40 mA 时,载流子密度都迅速上升,且上升所用时间随偏置电流增大而减小,分别是 2.35,1.38 和 0.988 ns,当升高到峰值点后回落,产生振荡并逐渐衰减到一个稳定值,不同直流偏置条件下的载流子密度分别是 1.96×10^{24},1.97×10^{24} 及 1.98×10^{24} m^{-3},可见激光器的载流子密度受直流偏置点影响不大;当偏置电流为 10 mA 时,载流子密度平缓上升,其值接近于 1.69×10^{24},远远低于其他直流偏置条件下的值。

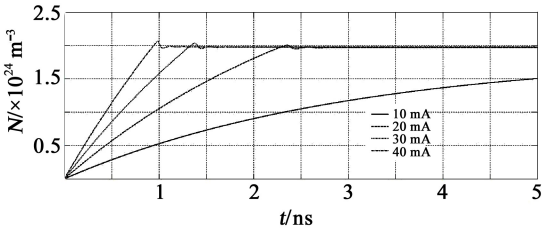


图 3 不同偏置电流时的载流子密度瞬态图

3.3 温度影响

激光器在不同温度下的频率啁啾仿真结果如图 4 所示。激光器输入脉冲电流调制,周期为 2 ns,占空比为 25%,幅度为 40 mA,偏置电流设定为 30 mA,环境温度分别设定为 280,300 及 320 K。可以看出,不同温度条件下的绝热啁啾变化不大,都约为 11 GHz,而瞬时啁啾随着偏置温度增大而增大,分别是 2.8,3.5 及 4.8 GHz。显然,降低激光器工作温度可以有效抑制激光器的瞬时啁啾。温度的变化也影响着弛豫振荡频率,随着温度的增加,其弛豫

振荡频率变小，分别是 8.19，6.67 及 5.11 GHz。

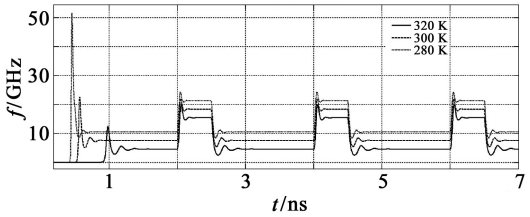


图 4　不同温度下的脉冲调制频率啁啾特性

3.4　弛豫振荡过程分析

光子和载流子密度同步对比如图 5 所示，给激光器加上一个大于阈值的偏置电流，为便于对比分析，对光子密度 S 和载流子密度 N 的值进行了归一化处理，在时间 $t=0$ 时刻开始，有源区的载流子数开始逐渐上升，这时腔内的光子数仍然为零，当 $t=1.684$ ns 时，光子数从零急剧上升，在 $t=1.734$ ns 时载流子数达到峰值，但此时载流子数仍然大于阈值，光子数还在继续增加，在 $t=1.766$ ns 时光子数到达峰值，因此，两者交替变换，完成弛豫振荡过程，并最终形成稳定的激射输出。

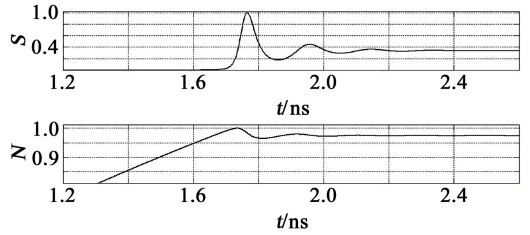


图 5　光子和载流子密度同步对比图

3.5　频率啁啾特性

频率啁啾产生的本质是注入电流使载流子浓度发生变化引起增益变化实现对光信号强度调制的同时也会造成有效折射率变化，从而对光信号形成一个附加的相位调制，即其发射频率随着时间变化。

3.5.1　直流偏置点影响

取表 1 中所示 DFB 激光器的参量值，偏置电流分别为 30，40 和 50 mA，激光器输入脉冲电流调制，幅度为 30 mA，周期为 2 ns，占空比为 25%，不同偏置电流下频率啁啾仿真计算结果如图 6 所示。可以看出，不同偏置电流条件下的绝热啁啾都约为 8 GHz，而瞬时啁啾随着偏置电流增大而变小，分别是 2.0，1.8 及 1.7 GHz，显然，增大直流偏置可以有效抑制激光器的瞬时啁啾。

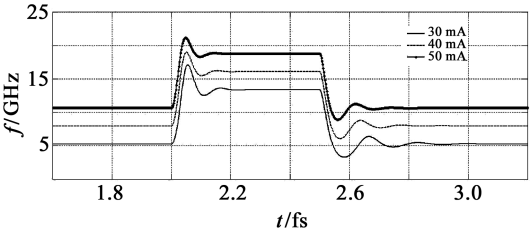


图 6　不同偏置电流下的脉冲调制频率啁啾特性曲线

3.5.2　调制脉冲强度影响

激光器输入脉冲电流调制，周期为 2 ns，占空比为 25%，幅度分别是 30，40 和 50 mA，偏置电流设定为 30 mA，仿真计算结果如图 7 所示。可以看出，不同脉冲调制下的瞬时啁啾随着调制脉冲幅度呈线性增加，而绝热啁啾差别不大。

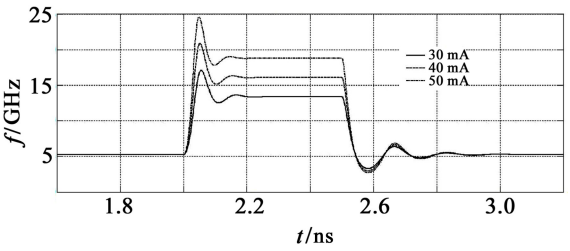


图 7　不同脉冲调制电流下的频率啁啾特性曲线

4　结论

本文通过 Simulink 软件工具对带温度参量修正的速率方程组进行了仿真求解，求解结果符合该型 DFB 激光器的瞬态特性，其温度影响得到定量分析。因仿真计算完全依据速率方程，所以对同类型的激光器有较强的普适性，可以为 DFB 激光器设计制造和应用选型提供一定的参考。后续工作将结合实际测试数据进一步修正热率方程并考虑引入强度噪声(RIN)特性。

参考文献：

[1]　马晓红，于晋龙. 激光器速率方程参数的实验提取及动态特性实验[J]. 电子学报，2002，30(9)：1295-1298.
Ma Xiaohong，Yu Jinlong. Extraction of the parameters of laser rate equation by measurement and dynamic character experiment[J]. Acta Electronica Sinica，2002，30(9)：1295-1298.

[2]　Ramunno L，Sipe J E. Dynamical model of directly modulated semiconductor laser diodes[J]. IEEE J. Quantum Electronics，1999，35(4)：624-634.

[3]　Gao J J，Li X P，Flucke J，et al. Direct parameter-extraction method for laser diode rate equation model[J]. J. Lightwave Tech.，2004，22(6)：1604-1609.

[4]　Thomson G H. Temperature dependence of threshold current in (GaIn)(AsP)DH lasers at 1.3 and 1.5 m wavelength[J]. IEEE Proc.，1981，128(2)：37-43.

(下转第 58 页)

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.011

基于微多环谐振陀螺仪的馈通效应研究

田梦雅，张卫平，谷留涛，刘朝阳，薛　冰

（上海交通大学 电子信息与电气工程学院 微纳电子学系，微米/纳米加工技术国家级重点实验室，薄膜与微细技术教育部重点实验室，上海 200240）

摘　要： 由SOI片制备的微多环谐振陀螺仪在测试过程中，不可避免地会引入馈通信号，而目前常用的差分检测馈通取消方式高度依赖微陀螺仪和检测电路的对称性，不能有效地消除馈通信号。针对此问题，提出了一种基于反相信号的馈通取消方案，通过在驱动和检测电极之间施加与馈通信号反相的信号，从而抵消馈通效应带来的干扰，该方法理论上可以在不降低驱动力的情况下将馈通信号完全取消。实验结果表明，该方法能有效消除馈通信号对电容检测的影响，谐振峰高度可提高约25倍。

关键词： 微多环陀螺仪；电容检测；馈通效应；电学模型；馈通取消

中图分类号： V241.554　**文章编号：** 1001－5868(2020)01－0054－05

Research on The Feed-through Effect of Micro Multi-ring Resonant Gyroscope

TIAN Mengya, ZHANG Weiping, GU Liutao, LIU Zhaoyang, XUE Bing

(1. National Key Lab. of Science and Technol. on Micro/Nano Fabrication, Key Lab. for Thin Film and Microfabrication of the Ministry of Education, Department of Micro and Nano Electron., School of Electronic Information and Electrical Engin., Shanghai Jiao Tong University, Shanghai 200240, CHN)

Abstract: In the test process of the micro multi-ring resonant gyroscope fabricated with SOI slice, it is inevitable to introduce the feed-through signal. However, the feed－through cancellation method based on differential detection which is current commonly used is highly dependent on the symmetry of the micro gyroscope and the detection circuit, which cannot effectively eliminate the feed-through signal. In order to solve this problem, a feed-through cancellation scheme based on the inverting signal is proposed, which can eliminate the interference caused by feed-through effect by applying the signal which is opposite to the feed-through signal between the driving and the detection electrode. In theory, this method can completely cancel the feed-through signal without reducing the driving force. Experimental results show that this method can effectively eliminate the influence of the feed-through signal on the capacitance detection, and the resonance peak height can be increased by about 25 times.

Key words: micro multi-ring resonant gyroscope; capacitance detection; feed-through effect; electrical model; feed-through cancellation

0　引言

随着MEMS工艺的逐渐成熟，高性能微传感器进入了一个快速发展期，不同结构微陀螺仪的设计与制备受到了各大研究机构的关注，在军事与民用领域的应用越来越广泛。其中，微多环谐振陀螺仪由于其结构的高度对称性和良好的性能而受到广泛关注。

在对电容检测式微多环谐振陀螺仪进行测试的

收稿日期：2019－11－23.

基金项目：国家自然科学基金项目(61574093)；预研基金项目(614280504010317，6140863020202)；上海专业技术服务平台项目(19DZ2291103)；教育部新世纪优秀人才支撑计划项目(NCET-10-0583).

过程中，馈通效应导致的驱动端信号到检测端的直接串扰，会对微陀螺仪检测造成较大的干扰[1]，这种干扰一般不能忽略，需要进行相应处理。

本文以微多环谐振陀螺仪为研究对象，对其结构与工作方式进行了介绍，并对馈通效应的产生机理与影响进行了研究，然后通过理论研究与实验论证，提出了馈通取消的相应方案，以显著降低驱动信号对于敏感检测的干扰。在同等驱动力幅度下，观测到了更为明显的谐振效果。

1　微多环谐振陀螺仪基本工作原理

1.1　结构与工作模态

1.1.1　结构

微陀螺仪是一种用于测量旋转角速度的惯性传感器，基于科氏力原理工作。环状微陀螺仪具有优良的对称性，其驱动和检测模态结构相同，这降低了它的环境敏感性，减少了误差来源，可用于实现高性能要求的微陀螺仪。

振动环型微陀螺仪的基本结构由支撑梁、振动环和电极等组成[2]，本文研究的微多环谐振陀螺仪由 SOI 片制备而成，结构如图 1 所示。

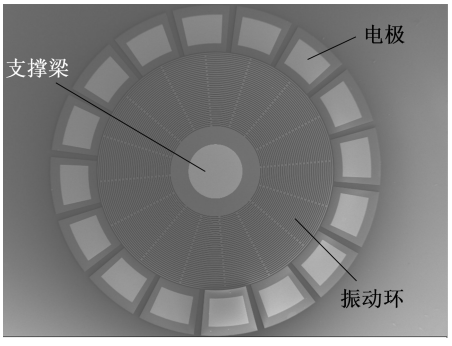
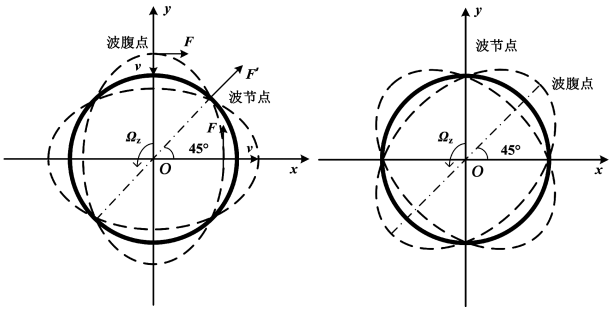


图 1　微多环谐振陀螺仪结构图

1.1.2　工作模态

微多环谐振陀螺仪作为振动式陀螺仪，一般工作于四波腹模态。图 2 给出了驱动与检测振动模态，可以看出，这两种振动模态是相互独立的，其波腹与波节点互补，在理想情况下，可以认为驱动与检测模态是互不干扰的，微陀螺仪振动环任意位置的振动状态可以表示成两种基础模态的叠加[3]。

在微陀螺仪起振后，如果波腹点的振动方向如图 2(a)中 v 所示，那么根据科氏力效应，可以得到两点所受到的科氏力的方向如图中 F 所示，且力的大小相等，则在振动的波节点受到上述科氏力的合力作用如图 2(a)中 F'所示。在合力作用下，微陀螺仪将产生检测模态振动，如图 2(b)所示。



(a)　驱动模态　　(b)　检测模态

图 2　驱动与检测振动模态示意图

1.2　驱动与检测原理

本文研究的微多环谐振陀螺仪采用静电驱动、电容检测的工作方式。在驱动电极上施加交流电压 v_{ac}，检测电极由放大器虚地，电极与振动环之间近似平行板电容器，对振动环施加静电力。其关系式如下：

$$F_d=\frac{dW}{dx}=\frac{1}{2}\frac{dC}{dx}V^2=\frac{1}{2}\left[(v_{ac}-V_p)^2\frac{dC_d}{dx_d}-V_p^2\frac{dC_s}{dx_s}\right] \tag{1}$$

式中，W 为静电力做的功，x 为位移，V_p 是加在振动环上的直流偏置电压，C_d 和 x_d 分别为驱动电极与振动环之间的电容和距离，C_s 和 x_s 分别为检测电极与振动环之间的电容和距离。假设振动环位移非常小，则由平行板电容公式可知：

$$\frac{dC_d}{dx_d}=\frac{d\frac{\varepsilon S}{x_d}}{dx_d}=-\frac{\varepsilon S}{x_d^2}\approx\frac{C_0}{x_0} \tag{2}$$

式中，ε 为真空介电常数，S，C_0 和 x_0 分别为电极与振动环之间的相对面积、静态电容和静态距离，同理求得检测电极值：

$$\frac{dC_s}{dx_s}=\frac{d\frac{\varepsilon S}{x_s}}{dx_s}=-\frac{\varepsilon S}{x_s^2}\approx\frac{C_0}{x_0} \tag{3}$$

假设 v_{ac} 的幅值远小于 V_p，则有：

$$F_d\approx-\frac{v_{ac}V_pC_0}{x_0} \tag{4}$$

由此可以得出，静电驱动力幅度与交流电压幅值以及偏置电压值成正比，频率等于交流电压的频率[4]。当驱动交流电压频率等于微陀螺仪的谐振频率时，微陀螺仪将产生一个四波腹的振动，通过负反馈控制驱动电压的幅值，可以使微陀螺仪驱动模态的振动稳定在一定幅度，此时检测模态的振动幅度仅与

旋转角速度相关。

微陀螺仪的振动使振动环与检测电极间形成一个变化的电容信号，在直流偏置电压的作用下输出一个幅值与振动幅度成正比的微弱电流信号。在对微陀螺仪进行检测时，由 C/V 放大器对其输出电流信号进行放大与转换，通过信号处理后可以得到旋转角速度信息。

2 馈通效应理论分析与解决方案

2.1 理论分析

2.1.1 馈通效应产生机理

对于微传感器，将输入信号与输出信号之间由寄生效应产生的直接耦合称为馈通效应。这种耦合使得微陀螺仪在未起振的情况下也有输出，从而干扰陀螺仪的检测结果。由于馈通效应，微陀螺仪会有一个较大的零偏，且幅频和相频曲线会发生畸变，导致测试的谐振频率和 Q 值与实际值存在一定偏差，从而无法反映器件的真实性能。

对于由 SOI 片制备而成的微多环谐振陀螺仪，电容性馈通效应主要是由衬底与该装置电极之间的电容耦合引起的[4]。如图 3 所示，在电极与衬底之间存在着一个电容 C_{pad}，由于电极的大键合面积，这个电容的值较大。同时，在 PCB 板的设计过程中，不可避免地会引入寄生电容，这也会形成一定的馈通电容。在较大馈通电容的作用下，输入与输出之间存在的馈通效应无法忽略。

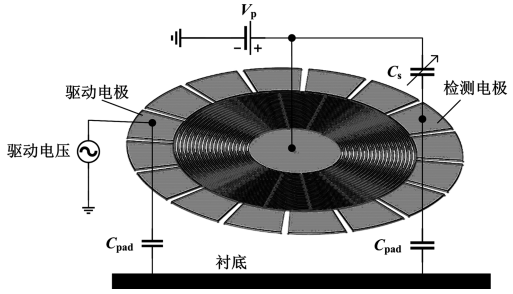


图 3 馈通效应电路模型

2.1.2 电学建模

对于振动式微陀螺仪，可将其看作二自由度的受迫振动系统，其动力学方程可以与串联 RLC 振荡电路的电学方程进行对应，对微陀螺仪结构进行等效电学建模。

加入等效馈通电容后，给出微陀螺仪的等效电路模型如图 4 所示，包括串联 RLC 振荡电路和与之并联的馈通电容 C_f。

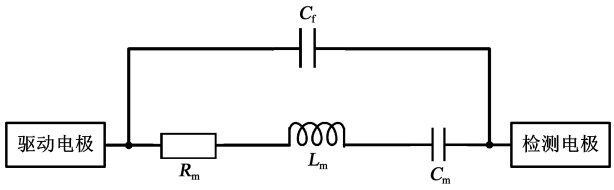


图 4 等效电路模型

由等效电路图，可知电路的导纳为

$$Y_m(j\omega)=\frac{1}{R_m+j\omega L_m+\dfrac{1}{j\omega C_m}}+j\omega C_f \quad (5)$$

式中，ω 为驱动电极上施加交流电压的频率。进一步分析电路的幅频特性，可知，由于馈通电容的存在，电路存在两个谐振频率点，表现在幅频曲线中为一个极大值点和一个极小值点，此时串联谐振频率低于微陀螺仪机械谐振频率，因此微陀螺仪的谐振频率测试将存在误差[5]。且当馈通电容较大时，馈通电流会使幅频曲线整体向上移动，同时谐振峰的高度减小，当谐振峰高度低于 3 dB 时，将无法测得 Q 值。因此，馈通电容的存在会对微多环谐振陀螺仪的性能和测试产生干扰，当馈通电容较大时，这种干扰无法忽略[6]。

2.2 解决方案

对于微陀螺仪，常采用差分检测对包括馈通信号在内的共模误差进行抑制。但在这种情况下，由于工艺、结构和环境等方面的误差，微陀螺仪的对称性达不到理想状态，故无法完全消除馈通，且这种方法的效果严重依赖器件的对称性，在对称性较差时将无法达到取消馈通的效果[7-8]。本文采用的馈通取消方式，可以在不降低驱动力的情况下对馈通取消效果进行调节，理论上甚至可以完全抵消馈通效应。

馈通取消方案的原理图如图 5 所示，驱动电压进行反相后连接到用来匹配馈通电容的补偿电容，产生的电流与馈通电流相位相反且幅值相等，相加

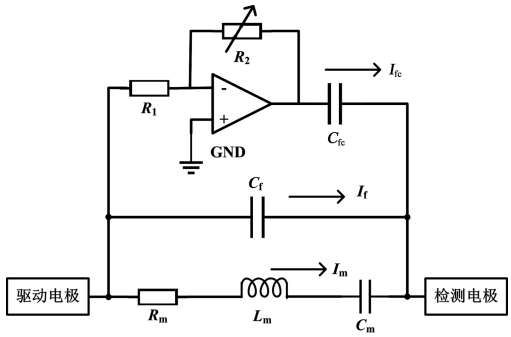


图 5 馈通取消原理图

来实现馈通取消[9-10]。具体实现时，可采用放大倍数可调的反相放大，来对馈通进行匹配。这种方法不依赖器件的对称性，而且取消效果可调，适用于不同参数的微陀螺仪。

根据原理图，对谐振时检测电极输出的电流值进行计算：

$$I = I_{\mathrm{m}} + I_{\mathrm{f}} + I_{\mathrm{fc}} = \frac{v_{\mathrm{ac}}}{R_{\mathrm{m}}} + v_{\mathrm{ac}}\left(sC_{\mathrm{f}} - \frac{R_2}{R_1} \cdot sC_{\mathrm{fc}}\right) \tag{6}$$

式中，各变量含义如前文或原理图所示，要使馈通电流被抵消，有：

$$\frac{R_2}{R_1} = \frac{C_{\mathrm{f}}}{C_{\mathrm{fc}}} \tag{7}$$

设置 R_2 为可变电阻器，对其阻值进行调节，理论上可以完全取消馈通。

3　实验测试与结果分析

3.1　实验测试

3.1.1　Labview 测试

用信号发生器输出扫频交流信号，固定信号幅值为 2 V，频率从 55 kHz 扫描到 65 kHz，扫描周期为 10 s。采用扫频信号驱动实验室制备的微多环谐振陀螺仪，通过数据采集卡采集陀螺仪输出信号，用 Labview 进行数据处理后，每隔 1 s 输出信号幅值并绘图。

在未进行馈通取消时得到的测试结果如图 6 所示，可以看到由于馈通电容容抗随频率升高而降低，馈通信号随着频率提高而增大，在驱动信号频率等于微陀螺仪谐振频率时，微陀螺仪起振，馈通信号和起振后的信号叠加，在扫频输出图形上将会存在一个突起的谐振峰。在对数形式的幅频曲线中，可计算得约为 0.3 dB。同等条件下，进行馈通取消后的

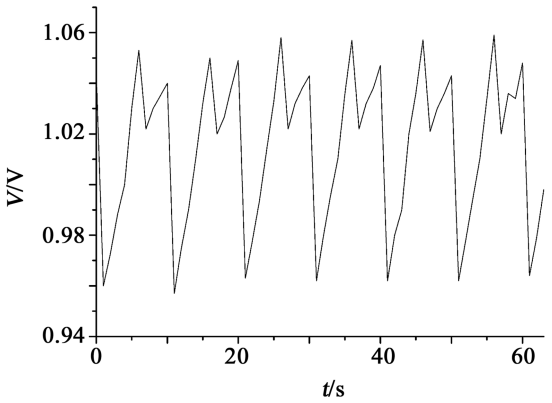


图 6　馈通取消前测试图

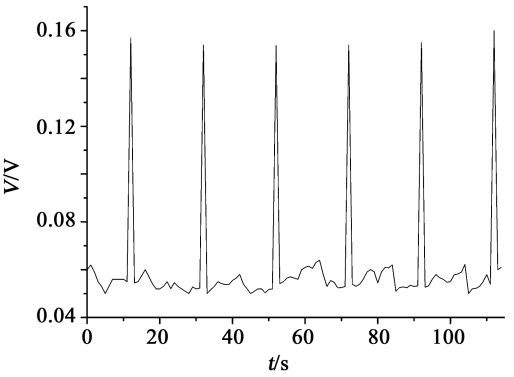


图 7　馈通取消后测试图

测试结果如图 7 所示，两者对比可以明显看出，在进行馈通取消后，馈通信号显著降低，起振效果变得更为明显，此时谐振峰可达到约 8 dB。

3.1.2　网络分析仪测试

进行馈通取消后，由网络分析仪测试幅频特性，缩小扫频范围至谐振频率附近，可以看到明显的谐振峰，如图 8 所示。

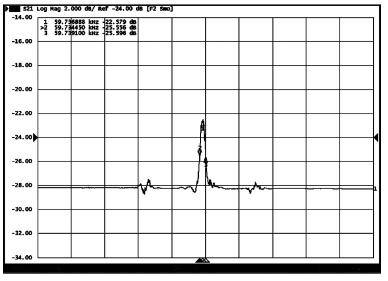


图 8　扫频测试图

在 59.7 kHz 附近可以得到谐振峰，谐振峰约为 5.5 dB，图中不存在明显的极小值点，说明馈通取消效果良好。−3 dB 带宽为 4.7 Hz，Q 值计算公式为

$$Q = \frac{f_0}{f_{-3\,\mathrm{dB}}} \tag{8}$$

式中，f_0 为谐振频率，$f_{-3\,\mathrm{dB}}$ 为 −3 dB 带宽。计算得到 Q 值为 12 710。

3.2　结果分析

由上面的测试结果可以看出，馈通取消效果明显，可以显著降低微多环陀螺仪的馈通效应，观测到更明显的起振现象。在网络分析仪实验中，对馈通效应进行抑制后，得到了更为明显的谐振峰，可以进行 Q 值计算。

4　总结

本文对微多环谐振陀螺仪的基本理论进行了介绍，分析了其结构与工作原理。在此基础上深入研

究了馈通效应的产生机理,以及对微陀螺仪性能与测试产生的影响。在电学上对馈通取消的方法进行了研究,给出了馈通取消的方案,并进行了相应的实验测试,对比分析了其取消效果。在进行馈通取消后,谐振峰的高度可由 0.3 dB 提高到 8 dB 左右,由网络分析仪测得微多环谐振陀螺仪的谐振频率为 59.7 kHz,Q 值为 12 710,馈通取消效果良好。

参考文献:

[1] 张晓峰,赵鹤鸣,钱 磊,等. 基于参数激励法的 MEMS 陀螺电馈通抑制方法[J]. 传感器与微系统,2019,38(5):26-28.
Zhang Xiaofeng, Zhao Heming, Qian Lei, et al. Electrical feedthrough suppressing method for MEMS gyroscope based on parametric excitation[J]. Sensors and Microsystems, 2019, 38(5): 26-28.

[2] 苏小波. 振动环型 MEMS 陀螺仪的研究与相关设计[D]. 无锡:江南大学,2011.
Su Xiaobo. Research and relevant design of vibrating ring MEMS gyroscope[D]. Wuxi: Jiangnan University, 2011.

[3] 张钰莹,张卫平,崔 峰,等. 基于 FPGA 的压电半球谐振陀螺数字测控系统设计[J]. 半导体光电,2019,40(3):313-317.
Zhang Yuying, Zhang Weiping, Cui Feng, et al. Design of digital detection and control system of the piezoelectric hemisphere resonant gyro based on FPGA[J]. Semiconductor Optoelectronics, 2019, 40(3): 313-317.

[4] Mayberry Curtis. Interface circuits for readout and control of a micro-hemispherical resonating gyroscope [D]. Georgia: Georgia Institute of Technol., 2014.

[5] Zhang H, Yuan W, Hao Y, et al. Influences of the feedthrough capacitance on the frequency synchronization of the weakly coupled resonators[J]. IEEE Sensors J., 2015, 15(11): 6081-6088.

[6] Kangul M, Aydin E, Gokce F, et al. Analysis and elimination of the capacitive feedthrough current on electrostatically actuated and sensed resonance-based MEMS sensors[J]. J. of Microelectromechanical Systems, 2017, 26(6): 1-7.

[7] Iqbal F, Shah M A, Lee B. Analysis of parasitic feed-through capacitance in MEMS gyroscope with push pull configuration [C]//IEEE 12th Inter. Conf. on Nano/Micro Engin. and Molecular Systems (NEMS), 2017.

[8] Xu Y, Lee J E Y. Feedthrough cancellation in micromechanical square resonators via differential transduction [C]// IEEE Electron Devices & Solid-state Circuits, 2010.

[9] Lee E Y, Seshia A A. Parasitic feedthrough cancellation techniques for enhanced electrical characterization of electrostatic microresonators[J]. Sensors and Actuators: A Physical, 2009, 156(1): 36-42.

[10] Ali A, Lee E Y. Single device on-chip feedthrough cancellation for enhanced electrical characterization of piezoelectric-on-silicon resonators in liquid[J]. Sensors and Actuators A: Physical, 2017, 260: 131-138.

作者简介:

田梦雅(1995—),女,硕士研究生,主要研究方向为微陀螺仪测控系统;

张卫平(1971—),男,博士,教授,博士生导师,主要研究方向为先进微型机器人与微型无人机、先进微纳惯性技术、微机电系统、精密仪器、电路控制与嵌入式系统及其集成技术。

E-mail: zwp37@163.com

(上接第 53 页)

[5] Minoglou K, Kyriakis-Bitzaros E D, Syvridis D, et al. A compact nonlinear equivalent circuit model and parameter extraction method for packaged high-speed VCSELs[J]. J. of Lightwave Technol., 2005, 22(12): 2823-2827.

[6] 梁 锋,高建军,田学农. 垂直腔面发射激光器的温度模型[J]. 半导体学报,2007,28(7):1125-1129.
Liang Feng, Gao Jianjun, Tian Xuenong. An improved thermal model for a VCSEL[J]. Chinese J. of Semiconductors, 2007, 28(7): 1125-1129.

[7] Byrne D M, Keating B A. A laser diode model based on temperature dependent rate equations [J]. IEEE Photon. Technol. Lett., 1989, 1(11): 356-359.

[8] Mena P V, Morikuni J J, Kang S M, et al. A comprehensive circuit-level model of vertical-cavity surface-emitting lasers [J]. J. Lightwave Technol., 2000, 17(12): 2612-2632.

作者简介:

张亚标(1977—),男,博士,高级工程师,研究方向为微波光子学。

E-mail: zhangyabiao@126.com

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.012

基于人眼视觉特性的IRFPA非均匀性校正算法研究

代少升，易伟男，聂合文，杨　雨

（重庆邮电大学 信号与信息处理重庆市重点实验室，重庆 400065）

摘　要： 红外焦平面器件普遍存在着非均匀性问题，对红外焦平面阵列（IRFPA）实施非均匀性校正对提高成像质量具有重要意义。传统的IRFPA多点校正算法存在运算量大、实时性能低、实用效果差等问题，为此文章综合人眼视觉特性，提出IRFPA非均匀性多点校正算法。人眼视觉对图像的灰度分辨能力是有阈值限制的，利用这种分辨阈值可以对IRFPA的标定点进行有效压缩，生成像元号-校正系数表，然后通过查找系数表，实现IRFPA的非均匀性实时压缩校正。实验证明，提出的IRFPA非均匀性校正算法较传统的IRFPA非均匀性校正算法实时性能更好，非均匀性降低了0.203%。

关键词： IRFPA；非均匀性校正；人眼视觉特性；多点校正；DSP

中图分类号： TN215　　**文章编号：** 1001－5868(2020)01－0059－05

Research on IRFPA Non-uniformity Correction Algorithm Based on Human Visual Characteristics

DAI Shaosheng, YI Weinan, NIE Hewen, YANG Yu

(Chongqing Key Lab. of Signal and Information Processing, Chongqing University of Posts and Telecommun., Chongqing 400065, CHN)

Abstract: Infrared focal plane devices generally present non-uniformity problems, and implementing non-uniformity correction on IRFPA is of great significance for improving image quality. The traditional IRFPA multi-point correction algorithm is of such problems as large amount of computation, low real-time performance and poor practical effect. Therefore, combining with the human visual characteristics, a multi-point correction method is proposed for IRFPA non-uniformity. The human eye vision has a threshold limitation on the gray resolution capability of the image, and with this resolution threshold, the calibration point of the IRFPA can be effectively compressed to generate a pixel number-correction coefficient table. Then, the IRFPA non-uniformity real-time compression correction can be implemented by looking up the coefficient table. The experimental results show that the proposed algorithm has better real-time performance than traditional algorithms, and the non-uniformity is reduced by 0.203%.

Key words: IRFPA; non-uniformity correction; human visual characteristics; multi-point correction; DSP

0　引言

红外焦平面阵列（IRFPA）是成像系统中最重要的部分，然而由于生产工艺不成熟、制作材料质量不佳，当不同探测器接收到相同的红外辐照时，其输出各不同，称为IRFPA的非均匀性[1]。这种现象会在红外图像上产生一定的固定模式噪声，进而导致红外成像系统的温度分辨率大幅降低，使其难以满足系统实际需求。因此，在红外成像应用系统中，开发一种有效的非均匀校正（NUC）算法并在硬件平台上实现是提高图像质量的关键。

收稿日期：2019－10－29.

近年来研究人员已经提出了多种非均匀性校正算法，主要可以归纳成两类：基于场景的 NUC 算法和基于参照源的 NUC 算法。前者包含的算法主要有：时域高通滤波、神经网络和卡尔曼滤波[2]。基于神经网络的 NUC 算法是一种典型的基于场景的算法。该方法根据校正图像与其期望图像之间的 LMS 误差迭代地计算 NUC 参数，这一类算法不需要参照源定标，对响应漂移有着自适应校正的优点。例如陈宝国等人[3]提出一种优化的神经网络算法，并于 FPGA 平台上实现。但是基于场景的算法都存在的主要缺点是容易导致移动物体和图像的重影伪像，这在实际应用中受到限制。

基于参照源的 NUC 算法由于硬件易于实现、低算法复杂度，在实际应用中不受限制。典型的 NUC 方法主要包含一点校正、两点校正及多点校正[4]。多点校正算法因其更好的校正能力和抗噪性而被广泛使用，但是所有基于参照源的 NUC 算法都存在低实时性缺点。视频中运动画面不能连续播放，这是实时性低的表现，并限制了其应用。针对这一问题，研究人员提出了很多改进的方法。殷世民等[5]利用插值节点连接折线段来逼近探测单元的响应曲线，这将减少响应曲线的非线性误差。黄星明等[6]对多幅图像的响应值取平均值并将其作为校正系数。尽管如此，大多数改进算法未考虑将其应用于硬件平台的实时能力。

为了解决上述难题，本文提出结合视觉特性的多点实时压缩 NUC 算法。根据人眼视觉系统的特性来压缩探测单元的响应值，利用 Matlab 软件将像元的响应值制成编码表来避免运算量过大的问题，并介绍了基于 DSP 的硬件平台。

1 人眼的视觉特性研究

成像系统的热图像是被人眼接收的，直接反映系统的好坏和图像处理算法的有效性，因此形成热图像时应考虑人眼视觉特性[7]。人眼视觉的图像评估只与目标和背景对比的亮度有关，与图像整体亮度无关。由韦伯-费克纳(Weber-Fechner)法则[8]可知：视觉亮度与光照强度对数成比例，即

$$s = k \lg \frac{I}{I_0} \tag{1}$$

式中，k 为常数，I_0 为绝对门限值。设 I_i 为主观兴趣区 A_i 亮度，I_{0i} 为 A_i 的背景亮度，s_i 为权重因子，即可导出：

$$s_i = k \left| \lg \frac{I_i}{I_{0i}} \right| \tag{2}$$

已知灰度值与亮度有直接关联。经过相关测试[9]，人眼对灰度的分辨能力如图 1 所示，可以看出：灰度较高或较低的情况下，人眼对灰度的分辨能力下降，而在图像灰度中等的条件下，人眼的分辨能力较强。在灰度 0 附近人眼只能分辨出 6 个灰度等级，即人眼分辨不出 0 级灰度与 5 级灰度，认为是同一灰度颜色。在 191 灰度级时，分辨能力较强，能分辨出 2 个灰度级；在 255 灰度级附近，能分辨出 3 个灰度级。可以得出视觉系统辨别图像的能力有阈值限制。

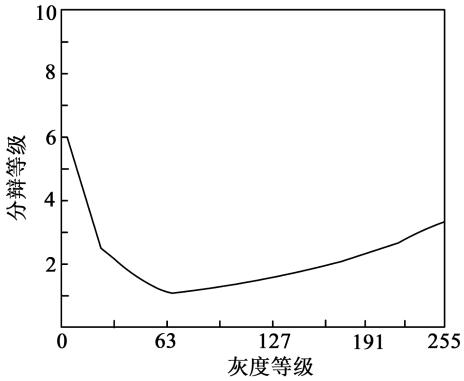


图 1 人眼的灰度分辨能力曲线图

依据人眼视觉系统存在的灰度分辨特性，选取灰度分辨阈值为 6 个灰度级，对多点校正的标定点进行压缩，从而提高 IRFPA 实时校正性能。

2 传统的多点校正算法

实际上，IRFPA 中各探测单元的响应曲线与红外辐照度的关系类似于抛物线曲线。五点校正的响应曲线如图 2 所示[10]，五点校正是将探测单元的响应曲线合理地选取标定点分为若干段，将每段视为线性区间使用两点校正法处理；V 表示探测单元的响应，ϕ 表示辐射通量，ϕ_1、ϕ_2、ϕ_3、ϕ_4、ϕ_5 表示定标点。

多点法通用的校正公式为[11]

$$V'_{ij}(\phi) = K_{ij}(\phi_l) V_{ij}(\phi) + O_{ij}(\phi_l) \tag{3}$$

式中，$\phi \in [\phi_l, \phi_{l+1}]$，$l \in [1, n]$；$K_{ij}(\phi_l)$ 为第 (i,j) 探测器在 l 线性区间内的增益系数，$O_{ij}(\phi_l)$ 为偏移系数；$V_{ij}(\phi)$ 为第 (i,j) 探测单元校正前的响应，$V'_{ij}(\phi)$ 为校正后的响应。

多点校正算法可以在嵌入式平台上实现。然而，增益和偏移系数是通过计算 IRFPA 的所有探

测单元响应数据得到的，很明显地可以看出多点校正算法运算量很大。公式的计算需要占用硬件资源，因此会增加算法运行时间，从而导致硬件平台实时性差，这将会限制系统的应用范围。

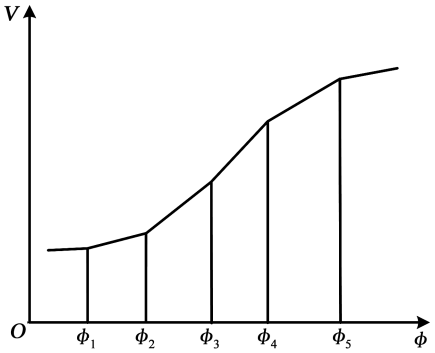


图 2　五点校正的响应曲线

3　基于视觉特性的多点压缩校正算法

如上所述，传统 NUC 算法存在实时性差的缺点，这是因为探测器的大量响应数据将导致硬件存储空间不够。系统采用的 IRFPA 是由法国 Sofradir 集团 ULIS 子公司生产的，阵列为 384×288，有 110 592 个像元。采用五点校正法需要知道所有探测单元的系数(110 592×8 bit×10＝1 080 kbit)以及 5 个标定点的数据(110 592×8 bit×5＝540 kbit)，所需存储的总数据量为(1080＋540) kbit＝1 620 kbit。同时，校正系数需要通过计算来选定，实施一次乘法和加法计算才能完成校正。大量存储量和复杂的计算过程占用存储空间，降低了 DSP 处理速度，导致系统实时性差。

通过对响应曲线分析得出：在同一线性区间内，各探测器的输出曲线有着邻近或相同的现象，可按照一定规律对其进行处理。综合以上研究，本文提出一种基于人眼视觉特性的多点压缩校正新算法，其原理描述如下：

1)首先，将 5 个标定点划分为不同的区间。已知红外图像的灰度值范围是 0～255。对于一幅多于 256 个像素的红外图像来说，每组标定点至少有两个或两个以上像素的灰度值相同或相近，因此可以将相同或相近灰度值的像素归为一类。按照设定好的灰度分辨阈值将所有探测单元的响应值划分为 43 个区间，并编成相应的号码，每个号码对应的灰度范围为 6。

2)编码像元来实现多点压缩。同一号码可以是相近灰度值的像元编成。如标定点的灰度值在区间 0～5 之间的编码号为 0，灰度值在 6～11 之间的编码号为 1，灰度值在 12～17 之间的编码号为 2，…，灰度值在 252～255 之间的编码号为 42。该表包含像元号与编码之间的对应关系，简称为编码表。

3)依据编码表制作系数表来校正。每组标定数据的最大值作为标准值，组内其他标定数据代入校正系数生成公式生成对应编号的校正系数表。DSP 处理器在上电后先加载校正系数表和各组标定数据，依据像元灰度值落入的标定区间范围，进行校正系数表的查找，确定相应的系数。

4)最后，在每个线性区间中启动两点校正方法，从而实现 IRFPA 的 NUC 压缩。

多点压缩法的增益和偏移系数的推导过程：首先，调节黑体的温度，可以得到所有探测单元在定标点 $\phi_1,\phi_2,\cdots,\phi_n$ 处的响应值 V。例如同一温度下，在 ϕ_1 定标点处可以得到 384×288 个响应值。通过公式得到极大值：

$$V'(\phi_1)=\max\{V_{1,1}(\phi_1),V_{1,2}(\phi_1),\cdots,V_{M,N}(\phi_1)\} \tag{4}$$

$$V'(\phi_2)=\max\{V_{1,1}(\phi_2),V_{1,2}(\phi_2),\cdots,V_{M,N}(\phi_2)\} \tag{5}$$

$$\vdots$$

$$V'(\phi_n)=\max\{V_{1,1}(\phi_n),V_{1,2}(\phi_n),\cdots,V_{M,N}(\phi_n)\} \tag{6}$$

式中，$M=384$，为焦平面的列数，$N=288$，为行数。将上述得到的极大值作为区间的校正值，根据两点校正的原理，在 ϕ_1 定标点处对应生成 384×288 组增益和偏移系数，对应 5 个定标点总生成 384×288×5 组增益和偏移系数：

$$K_{ij}(\phi_l)=\frac{V'(\phi_{l+1})-V'(\phi_l)}{V_{ij}(\phi_{l+1})-V_{ij}(\phi_l)} \tag{7}$$

$$O_{ij}(\phi_l)=\frac{V_{ij}(\phi_{l+1})V'(\phi_l)-V_{ij}(\phi_l)V'(\phi_{l+1})}{V_{ij}(\phi_{l+1})-V_{ij}(\phi_l)} \tag{8}$$

在压缩时，将第(i,j)探测单元在标定点的灰度值 V_{ij} 在对应区间编成对应编码号 H_{ij}：

$$V_{ij}\in[0,5],\quad H_{ij}=0 \tag{9}$$

$$V_{ij}\in[6,11],\quad H_{ij}=1 \tag{10}$$

$$\vdots$$

$$V_{ij}\in[252,255],\quad H_{ij}=42 \tag{11}$$

因此利用编码号可以避免重复计算，只需要计算 43×5 组增益和偏移系数：

$$K'_{ij}(H_{ij})=\frac{V'(\phi_{l+1})-V'(\phi_l)}{\Delta(H_{ij}(\phi_{l+1})-H_{ij}(\phi_l))} \quad (12)$$

$$O'_{ij}(H_{ij})=\frac{H_{ij}(\phi_{l+1})V'(\phi_l)-H_{ij}(\phi_l)V'(\phi_{l+1})}{H_{ij}(\phi_{l+1})-H_{ij}(\phi_l)} \quad (13)$$

式中，$\Delta=6$，为灰度分辨阈值，$H_{ij}\in[0,42]$。

4 基于 DSP 的嵌入式硬件平台

为了实现提出的 NUC 算法，本文提出一种基于数字信号处理(DSP)的嵌入式硬件平台来实现该算法。硬件平台框图如图 3 所示。

DSP 芯片是嵌入式硬件平台的主要部分，负责对各个硬件单元模块的初始化、控制、管理和对红外图像视频数据的算法处理，主要包括 FLASH、SDRAM、SD 卡等存储单元和按键、视频等接口单元的控制与管理。

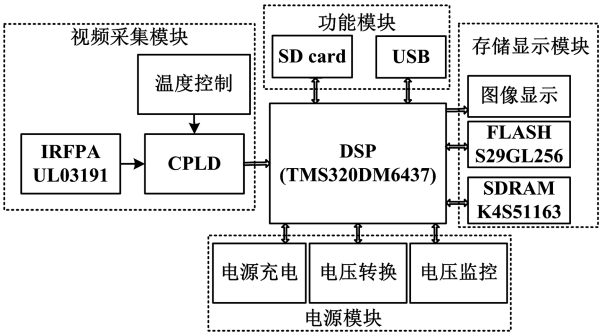


图 3 硬件平台框图

IRFPA 主要负责采集将各探测单元感应到的模拟信号转换成电信号。CPLD 可编程器件作为视频采集前端的逻辑控制模块，对红外探测器 UL03191 提供精确的工作驱动时序以及寄存器配置信号，同时将 IRFPA 输出的 6 bit 数字信号合成为 12 bit 数字信号。DSP 内置的视频处理子模块(VPSS)从 CPLD 接收合成数字视频信号并保存在数据缓存区中。VPSS 模块为视频输出提供视频处理(VPBE)接口。通过 VPBE 接口与 LCD 屏相连，实现显示功能。

图像算法处理如图 4 所示。算法处理原理是：非均匀性模块从缓存区接收经过盲元补偿后的图像数据来压缩图像响应数据。NU 模块将实现式(3)的处理并实现第 3 节给出的步骤。首先，平台启动后，DSP 将储存在 FLASH 的编码表、43×5 组增益、偏移系数复制到 SDRAM。然后，依据像元的响应值通过查询编码表查找对应的线性区间。最后，利用乘法器和加法器来实现非均匀性校正。

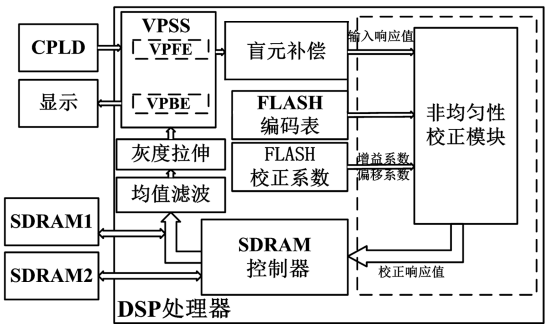


图 4 DSP 算法处理框图

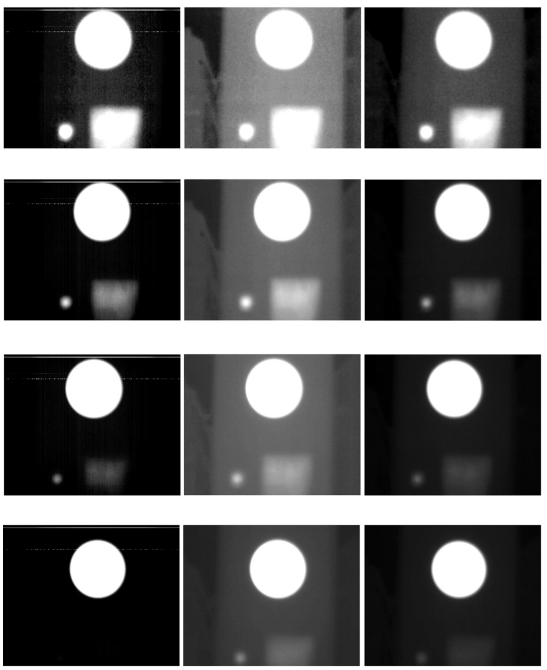
5 实验结果及分析

5.1 算法的非均匀性及图像效果

利用 DSP 硬件平台进行算法实验。非均匀性 NU 用于评估所提出算法去除固定模式噪声的能力，较小的 NU 表示更好的校正性能。经过两种算法处理后的非均匀性比较结果如表 1 所示，说明：与传统多点方法相比，经过基于视觉特性的多点压缩方法处理后的图像非均匀性明显降低。

表 1 两种校正方法的非均匀性

黑体温度/℃	30	50	70	90	110
校正前 NU/%	11.086	8.862 1	8.024 6	6.968	6.284 2
多点校正后 NU/%	3.962 9	4.523 3	2.537 2	2.860 3	3.226 5
新多点压缩校正后 NU/%	3.931	3.916 8	2.318 5	2.581 4	3.047 1



(a) 原始图像 (b) 多点校正算法 (c) 多点压缩校正算法

图 5 各算法实验结果图(四组图从上至下分别对应黑体温度 30，50，70 和 90 ℃)

两种方法的校正效果如图 5 所示，图中的图片都进行过盲元补偿、中值滤波和灰度拉伸。四组图从上至下分别对应黑体温度 30，50，70 和 90 ℃。图 5(a)为原始图像，对图 5(a)进行多点 NUC 后获得图 5(b)，经过基于视觉特性的多点压缩校正获得图 5(c)。图像效果表明：校正前的红外图像模糊；相比于传统校正算法，经过压缩处理后的图像均匀性好，清晰可见，说明提出的新算法具有较好的校正能力。

经过 MATLAB 仿真可以得到五点压缩校正的编码表，可得所需存储的数据量为校正曲线数据与编码表之和，即 43×8 bit$\times10+110\ 592\times8$ bit$\times5=540.42$ kbit。可以得知：传统算法储存数据量为 1 620 kbit，基于视觉特性的新方法可以减少数据量，具备良好的压缩能力。

5.2　算法运行时间

对于实时平台来说，算法运行时间属于另一个注重的要点。对于 25 f/s 的 PAL 格式模拟视频，每帧的可用时间是 40 ms。一旦超过这个时间，LCD 屏幕显示的图像会出现不连续、卡顿现象，这是低实时性的表现。

为了分析传统算法和基于视觉特性的新算法的实时性，在 DSP 硬件上进行在线仿真。主芯片 DM6437 的 CPU 时钟为 594 MHz。表 2 的结果是由 CCS5.5clock 计算出两种算法运行的 CPU 周期数。从表中的数据可以看出：新算法的运行 CPU 周期较传统算法少，运行速度明显高于传统算法，实时性高。

表 2　算法占用的周期数

参数	多点校正	多点实时压缩校正
每帧所需 CPU 周期数	52,493,583	19,382,479
帧数/Hz	11.32	30.65

6　结论

由于非均匀性问题是 IRFPA 的固有属性，不可避免地需要进行非均匀性校正。系统的低实时性是降低系统成像质量的主要原因。本文提出基于人眼视觉特性的多点实时压缩 NUC 校正算法，经过在 DSP 硬件平台上实验，与传统多点校正算法相比，该算法不仅能显著地降低红外焦平面的非均匀性(平均降低 0.203%)，有效地改善图像质量，而且压缩了运算数据量，从而减少运算时间，有效地提高了实时性。该基于人眼视觉特性的新算法适用于对实时性和精度有较高要求的应用系统。

参考文献：

[1] 周建勋，王利平，刘　滨. 红外图像非均匀性产生原因分析[J]. 红外与激光工程，1997，26(3)：11-13.
Zhou Jianxun，Wang Liping，Liu Bin. Analysis of the cause for the ununiformity of infrared image[J]. Infrared and Laser Engin.，1997，26(3)：11-13.

[2] 吕　游，何　昕，魏仲慧. 红外焦平面阵列非均匀性校正算法研究[J]. 计算机技术与发展，2015，25(2)：1-5.
Lv You，He Wei，Wei Zhonghui. Research on non-uniformity correction algorithms for IRFPA[J]. Computer Technology and Development，2015，25(2)：1-5.

[3] 陈宝国，张学峰，牛英宇. 改进的基于神经网络的非均匀性校正算法[J]. 红外与激光工程，2013，42(3)：574-578.
Chen Baoguo，Zhang Xuefeng，Niu Yingyu. Improved nonuniformity correction algorithm based on neural network[J]. Infrared and Laser Engin.，2013，42(3)：574-578.

[4] 蔡荣太，王延杰. CCD 成像传感器的降噪技术[J]. 半导体光电，2007，28(5)：607-612.
Cai Rongtai，Wang Yanjie. Noise attenuation technologies for CCD imaging sensor[J]. Semiconductor Optoelectronics，2007，28(5)：607-612.

[5] 殷世民，刘上乾. 基于低次插值的红外焦平面器件非均匀性多点校正算法[J]. 光子学报，2002，31(6)：715-718.
Yin Shimin，Liu Shanggan. The multi-point nonuniformity correction algorithms for IRFPA based on low order interpolation[J]. Acta Photonica Sininca，2002，31(6)：715-718.

[6] 黄星明，余国文，王宏远. 一种改进的红外焦平面非均匀多点校正方法[J]. 红外与激光工程，2005，34(1)：62-65.
Huang Xingming，Yu Guowen，Wang Hongyuan. Imaging multi-point nonuniformity correction method for IRFPA[J]. Infrared and Laser Engineering，2005，34(1)：62-65.

[7] 陈　勇，谭　超，谢正祥，等. 基于对比度分辨率补偿的自适应底层视频挖掘夜视系统[J]. 重庆邮电大学学报(自然科学版)，2011，23(5)：597-601.
Chen Yong，Tan Chao，Xie Zhengxiang，et al. Self-adaptive lower layer video mining system based contrast resolution compensation[J]. J. of Chongqing University of Posts and Telecommun.(Natural Science Edi.)，2011，23(5)：597-601.

[8] 陆旭光，汪岳峰，胡文刚，等. 基于视觉感兴趣区的图像质量评价方法[J]. 微计算机信息，2005，21(30)：95-96，17.
Lu Xuguang，Wang Yuefeng，Hu Wengang，et al. A method of image quality assessment based on visual regions-of-interest[J]. Microcomputer Information，2005，21(30)：95-96，17.

[9] 贾　其，吕绪良，吴　超，等. 基于人眼视觉特性的红外图像增强技术研究[J]. 红外技术，2010，32(12)：708-712.
Jia Qi，Lv Xuliang，Wu Chao，et al. Research on infrared image enhancement based on human visual system[J]. Infrared Technology，2010，32(12)：708-712.

(下转第 67 页)

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.013

基于银纳米薄膜欧姆接触的高波长选择性紫外探测器研究

梁志斌，张齐轩，许朝军，周玉刚，陆　海，张　荣，郑有炓

（南京大学 电子科学与工程学院 江苏省光电信息功能材料重点实验室，南京 210023）

摘　要： 常规的半导体紫外探测器波长响应范围宽，而紫外光的应用具有较强的波长选择性，如320 nm波段的紫外光在医学方面有重要的应用，因此，具有高波长选择性的紫外探测器的研制有重要意义。文章采用GaN基p-i-n探测器结构，通过在p区覆盖银纳米薄膜作为欧姆接触层和波长选择透射层，成功制备了对320 nm波段紫外光高选择性探测的紫外探测器，器件性能如下：70 nm银层的紫外光透射率峰值超过30%，器件在－5 V偏压下的暗电流为10^{-12} A量级，响应峰值为0.06 A/W，响应峰发生在325 nm处，光谱响应峰半高宽约30 nm。

关键词： p-i-n紫外探测器；GaN；银纳米薄膜；波长选择性

中图分类号： TN364.2　**文章编号：** 1001－5868(2020)01－0064－04

High Wavelength-selective Ultraviolet Detector Based on Ag-nanofilm Ohmic Contact

LIANG Zhibin, ZHANG Qixuan, XU Chaojun, ZHOU Yugang,
LU Hai, ZHANG Rong, ZHENG Youdou

(School of Electronic Science and Engin., Nanjing University, Nanjing 210023, CHN)

Abstract: Conventional semiconductor ultraviolet detectors have a wide wavelength response range. However, applications of ultraviolet light are generally highly sensitive to specific wavelength range, such as 320 nm ultraviolet light is used in medical light therapy. Therefore, development of wavelength-selective ultraviolet detectors is of great significance. Using GaN-based p-i-n ultraviolet detectors with Ag-nanofilm as ohmic contact and wavelength-selective transmission layer, ultraviolet detector highly selective to the 320 nm band ultraviolet light is successfully fabricated. The peak value of ultraviolet light transmittance with 70 nm silver layer is higher than 30%. The dark current of the device at －5 V is of the order of 10^{-12} A. The peak responsivity of the device reaches 0.06 A/W at 325 nm, and the full width at half maximum is about 30 nm.

Key words: p-i-n ultraviolet detector; GaN; Ag nanofilm; wavelength selectivity

0　引言

紫外光分为UVA(400～320 nm)、UVB(320～280 nm)和UVC(280～100 nm)。紫外辐射不仅影响人类的生存和持续发展，还有一系列的重要应用，如光固化、高分辨率光源、光疗、消毒、杀菌、除臭和光催化等[1]。紫外线的重要作用催生了庞大的紫外探测器市场[2]，紫外探测器被广泛用于紫外线通信、水质监测、医药医疗等军事和民用领域[3-5]。

Ⅲ族氮化物具有直接带隙和较宽的带隙范围，氮化镓(GaN)材料具有3.4 eV的直接带隙，铝镓氮(AlGaN)的带隙覆盖3.4～6.2 eV，因此，GaN和AlGaN基紫外探测器是紫外探测器的一个重点研

收稿日期：2019－09－20.

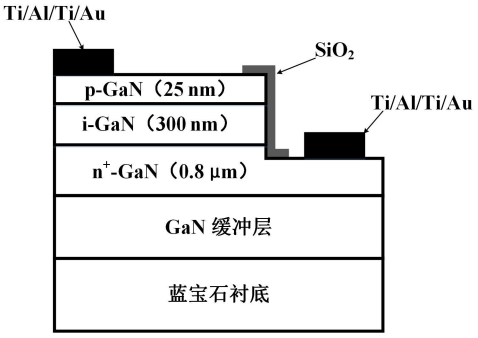
基金项目：国家重点研发计划项目(2016YFB0400904)；国家自然科学基金重点项目(61634005).

究方向。320 nm 波段的紫外光在医学方面有重要的应用，可用于治疗皮肤癌、白癜风等皮肤疾病[6]，准确计量对应波段的紫外辐照强度和剂量具有重要意义。此外，简单而较为精确地获得紫外光源在不同波段紫外线的强度也具有一定意义。常规的半导体紫外探测器对于波长短于吸收边的紫外光具有较宽的响应范围，但并不能仅在一个较窄的范围对紫外光具有响应，因此，高波长选择性紫外探测器以及具体的 320 nm 波段的紫外探测器的研制具有重要意义。

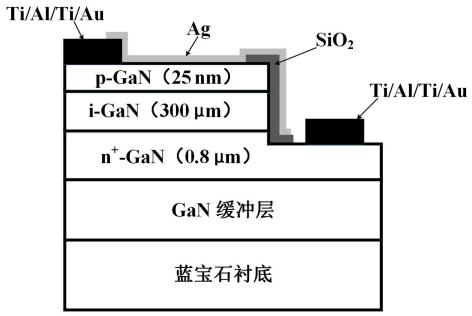
本文在 GaN 基 p-i-n 紫外探测器的结构上，采用对 320 nm 波段紫外光具有选择作用的银材料作为 p 区欧姆接触和波长选择透射层[7]，成功制备了对 320 nm 波段紫外光有较好选择探测作用的 GaN 基紫外探测器，并对器件的光电特性进行了表征。

1　器件的制备

p-i-n 型探测器外延结构如图 1 所示，外延层采用金属有机化合物化学气相沉积(MOCVD)方法生长。首先，在蓝宝石(Al_2O_3)衬底上生长薄的成核层以及约 1 μm 厚的非故意掺杂 GaN 缓冲层，之后在缓冲层上生长 800 nm 厚的 n^+-GaN 层，掺杂浓度为 5×10^{18} cm^{-3}，然后生长一层 300 nm 厚的 i-GaN 层，最后生长一层 p-GaN 层。为了更好地实现 p 区欧姆接触，本研究采用 p-GaN 作为 p 型层，掺杂浓度大于 3×10^{18} cm^{-3}。为了降低 p-GaN 对 320 nm 波段紫外光的吸收，p-GaN 层较薄，约 25 nm。材料生长完成后，先用 ICP 方法进行台阶刻蚀，刻蚀深度约 700 nm，露出 n^+-GaN 层。之后用 PECVD 方法生长 SiO_2 绝缘保护层，光刻并腐蚀 SiO_2，使 SiO_2 完全覆盖 p/n 台阶。随后在 p-GaN 和 n^+-GaN 层上方用电子束蒸发设备蒸镀 Ti/Al/Ti/Au(30nm/120 nm/40nm/200nm)p，n电极。作



(a)　A 组样品



(b)　B 组样品

图 1　GaN 基 p-i-n 型紫外光电探测器结构示意图

为对比，制备完成了两组样品，A 组样品结构如上所述，B 组样品在 A 组样品的基础上，进一步光刻，并用电子束蒸发蒸镀了约 70 nm 的银，然后剥离形成 p 区银欧姆接触层，制备完成的 GaN 基 p-i-n 紫外光电探测器结构示意图分别如图 1(a)和(b)所示。

2　器件测试与分析

为了测试镀在器件上的银的透光率，准备了一片双抛蓝宝石片在同一台电子束蒸发设备上蒸镀了相同厚度的银。首先用台阶仪测量了镀银蓝宝石片以及样品 B 的银层厚度，测试结果表明实际银层厚度约 70 nm。随后用分光光度计测试了镀银蓝宝石片在 280～400 nm 波段的透光率曲线，结果如图 2 所示。由于蓝宝石在选择波段范围基本没有吸收，因此可以近似认为所测试的镀银蓝宝石片的透光率即为该银层的透光率。当波长 λ 为 322 nm 时，透射率达到峰值为 35.4%，透射率曲线半高宽为 25 nm，离 320 nm 较远范围的波长对应的透射率接近于 0，由此可以预见，在 GaN 基紫外探测器上镀 70 nm 的银可以实现对波长的高选择性透射。

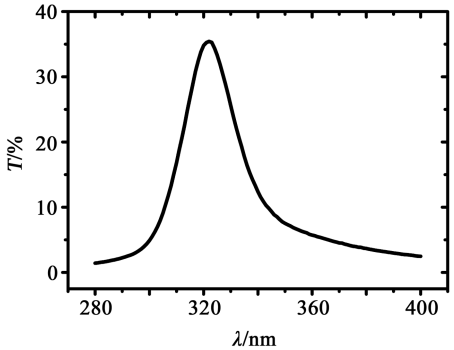


图 2　镀银蓝宝石片的透光率曲线

图 3 所示为样品 B 在光、暗条件下的 *I-V* 特性曲线(半对数坐标)，其中光照条件为 320 nm 紫外

光照。对于样品 B 在暗条件下的 *I-V* 特性，加−5～0 V 反向偏压时，随着反向偏压的增大，内建电场电势差增大，因而漂移电流增大，即暗电流随反向偏压增大而增大；加正向偏压时，扩散电流强于漂移电流，电流随偏压增大而增大。当施加 320 nm 紫外光照时，在反向偏压下，光生电流（由 n 极经 p-i-n 结流入 p 极）起主导作用，电流几乎不随电压变化。在−5 V 反向偏压下，暗电流为 10^{-12} A 量级，光电流为 10^{-11} A 量级。当施加较小的正向偏压时，光生电流（由 n 极经 p-i-n 结流入 p 极）基本不变，且起主要作用，通过结的正向电流（由 p 极经 p-i-n 结流到 n 极）逐渐增大，所以总电流方向与光生电流方向一致（由 n 极经 p-i-n 结流入 p 极，由 p 极经外部流向 n 极），且总电流逐渐减小为 0。随着正向电压增大，通过结的正向电流大于反向的光生电流，总电流方向改变，变为由 p 极经 p-i-n 结流入 n 极，由 n 极经外部流向 p 极，并逐渐增大。由于光生电流的存在，总电流为零时对应的电压相对无光照情况下右移，对应的偏压位于 0.25～0.30 V。

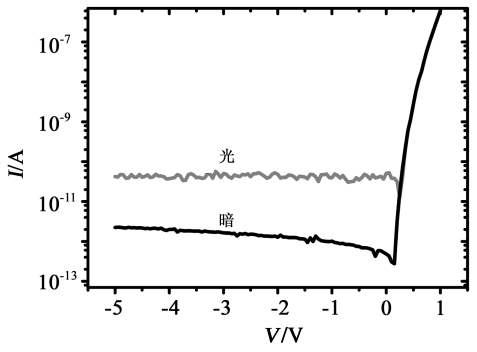


图 3 样品 B 在光、暗条件下的 *I-V* 特性曲线（半对数坐标）

图 4 给出了样品 A 与样品 B 在无光照条件下的 *I-V* 特性曲线（线性坐标）。样品 A 的开启电压约为 2.7 V，在＋5 V 偏压下电流为 167 mA；样品 B的开启电压约为2 V，在＋5 V偏压下电流为 327 mA。器件镀银之后，正向电流-电压特性改善，器件的开启电压降低，开启后同样电压下正向电流变大。测试了 p-GaN 上的两个银薄膜电极的电流-电压特性，表现为线性，即银薄膜与 p-GaN 的接触为欧姆接触。由于镀银的器件欧姆接触得到了改善，欧姆接触面积变大，欧姆接触电阻和电流扩展不均导致的电压升高得到改善，因此其开启电压较没有镀银的器件低。

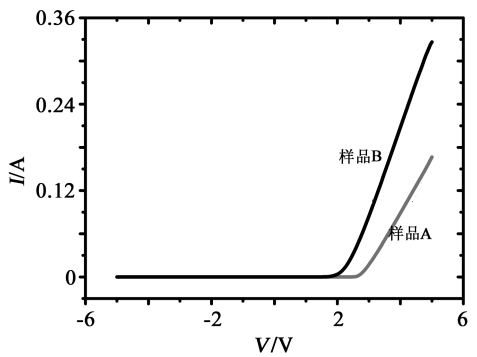
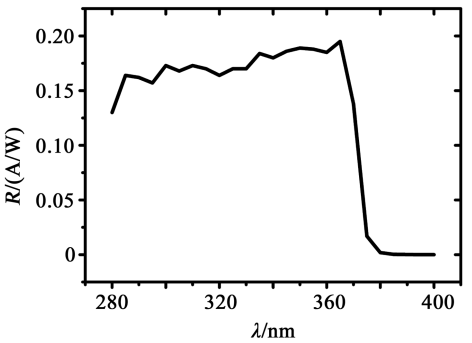


图 4 样品 A，B 在无光照条件下的 *I-V* 特性曲线

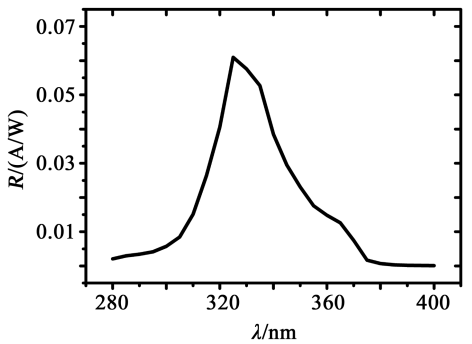
之后测试得到了样品 A，B 在零偏压下的响应光谱曲线，结果如图 5 所示。光响应度如下[8]：

$$R(\lambda)=\frac{I(P)}{P(\lambda)} \tag{1}$$

首先利用响应光谱已知的 Si 基光电二极管测量计算得到不同波长（λ）下的光功率 $P(\lambda)$，λ 步长为 5 nm（即每隔 5 nm 取点测量进行计算）；随后在相同光照条件下对样品 A，B 进行测试，得到其对应的电流值；最后通过式(1)计算得到两个样品的响应光谱，为了减少测量误差，对其进行多次测量取平均值处理。



(a) 样品 A



(b) 样品 B

图 5 样品 A，B 在零偏压下的响应光谱

从图 5 中可以看出，样品 A 的响应峰值发生在波长为 365 nm 处，对应响应为 0.195 A/W；当波长大于 365 nm 时光谱响应度迅速降低，有一条明显的光吸收边；当波长小于 365 nm 时，由于测试误差

的缘故，曲线存在一些波动情况，但整体呈现下降趋势。对于样品 B，响应度峰值约为 0.06 A/W，位于 325 nm 处，而根据镀银蓝宝石片的透射率曲线，透射率峰值对应的波长为 322 nm。从透射率曲线可以看出，在 322 nm 附近波长的紫外光的透射率差距不大，当波长略大于 322 nm 时，透射率有所降低，同时根据样品 A 的结果来看，GaN 探测器的光谱响应度会随波长增大而提高，因此，二者共同作用导致样品 B 的响应峰值对应的波长相较于银透射率峰值对应的波长要略大一些。另外，在样品 B 的响应光谱曲线中 365 nm 处有一个肩，正好对应了未镀银的样品 A 的光谱响应度峰值。

对比样品 A 与样品 B 的响应光谱曲线可以看出，探测器镀银之后虽然光谱响应能力有一定程度的下降，但实现了在 320 nm 附近波段预期的光谱响应窗口，响应峰位于 325 nm 处，半高宽约 30 nm。

3　结论

制备了镀银 GaN 基 p-i-n 型紫外光电探测器，对其性能进行测试分析：70 nm 银层的紫外光透射率峰值超过 30%，*I-V* 特性测试结果表明，器件的开启电压约 2 V，在 −5 V 偏压下的暗电流为 10^{-12} A 量级，相较于未镀银的器件其欧姆接触得到了优化；样品在 325 nm 处的光谱响应度峰值达到了 0.06 A/W，半高宽约 30 nm，实现了对 320 nm 附近波段紫外光的选择性探测。该选择性紫外光探测器在 320 nm 波段紫外光疗的辐照强度和剂量监测有重要应用价值；配合不同波长选择的紫外光探测器，有望对不同波段紫外光的强度和吸收实现快速分析，有重要的潜在应用价值。

参考文献：

[1] Li Dabing, Jiang Ke, Sun Xiaojuan, et al. AlGaN photonics: recent advances in materials and ultraviolet devices[J]. Adv. Opt. Photon., 2018, 10(1): 43-110.

[2] 陆　海，陈敦军，张　荣，等. 高灵敏度宽禁带半导体紫外探测器[J]. 南京大学学报(自然科学版)，2014，50(3)：294-301.
Lu H, Chen D, Zhang R, et al. High sensitivity wide-bandgap semiconductor ultraviolet photodetector [J]. J. of Nanjing University (Nature Science Edi.), 2014, 50(3): 294-301.

[3] Wu Di, Hu Qi, Yan Zhen, et al. Structural basis of ultraviolet-B perception by UVR8[J]. Nature, 2012, 484: 214-219.

[4] 方坷昊，赵　凌. 紫外-可见光谱法水质监测中浊度影响的非线性校正[J]. 传感器与微系统，2018，37(10)：35-37.
Fang K, Zhao L. Nonlinear correction of turbidity effect in water quality monitoring by UV-Vis spectroscopy [J]. Transducer and Microsystem Technol., 2018, 37(10): 35-37.

[5] Premi S, Wallisch S, Mano C M, et al. Chemiexcitation of melanin derivatives induces DNA photoproducts long after UV exposure[J]. Science, 2015, 347: 842-847.

[6] 陈　伟. 石墨烯-Ag 纳米颗粒透明电极近紫外 LED 研究[D]. 南京：南京大学，2018.
Chen W. Study on graphene-Ag nanodot transparent conductive layer of NUV-LEDs[D]. Nanjing: Nanjing University, 2018.

[7] Lamkin I A, Tarasov S A, Solomonov A V, et al. Selectivity control of photosensitivity of Ag-GaP and Ag-AlGaN structures[C]// 17th Russian Youth Conf. on Physics and Astronomy, 2015.

[8] 卢怡丹，王立伟，张　燕，等. InGaN 紫外探测器的制备与性能研究[J]. 半导体光电，2014，35(5)：785-788.
Lu Y, Wang L, Zhang Y, et al. Study on the fabrication and properties of InGaN ultraviolet detector [J]. Semiconductor Optoelectronics, 2014, 35(5): 785-788.

作者简介：

梁志斌(1996－)，男，江西宜春人，硕士生，主要从事 GaN 基光电器件方面的研究工作；

周玉刚(1975－)，男，博士，教授，主要研究方向包括 GaN 基半导体材料与器件、LED 器件工艺与先进封装，以及光电器件及系统的可靠性分析与寿命预测。

E-mail：ygzhou@nju.edu.cn

(上接第 63 页)

[10] 屈惠明，陈　钱. 红外焦平面阵列二元非线性的非均匀性理论模型[J]. 电子学报，2008，36(11)：2150-2153.
Qu Huiming, Chen Qian. A theoretical model on infrared focal panel arrays binary nonlinear nonuniformity [J]. Acta Electronica Sinica, 2008, 36(11): 2150-2153.

[11] 崔　磊，刘　智. 短波红外相机响应非均匀性校正方法研究及仿真[J]. 长春理工大学学报(自然科学版)，2017，40(2)：106-110.
Cui Lei, Liu Zhi. Research and simulation of non uniformity correction method for short wave infrared camera[J]. J. of Changchun University of Science and Technol. (Nature Science Edi.), 2017, 40(2): 106-110.

作者简介：

代少升(1974－)，男，河南潢川人，博士，教授，主要从事红外成像技术及 SOPC 嵌入式系统的设计与开发。

E-mail：daiss@cqupt.edu.cn

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.014

高分辨率 AM Micro-LED 显示器设计及其驱动实现

林杰泓[1]，聂君扬[2]，张永爱[1]，周雄图[1]，严　群[1]，郭太良[1]

（1. 福州大学 物理与信息工程学院，福州 350116；2. 西安交通大学 电信学部，西安 710049）

摘　要： 设计了一种基于蓝宝石衬底氮化镓基蓝色/绿色多量子阱 LED 外延片，利用成熟的 LED 芯片工艺制备了高像素 Micro-LED 芯片阵列；在 Micro-LED 芯片和 CMOS 驱动芯片表面设计并制备了金属凸点，采用高精度倒装焊接工艺实现 LED 显示芯片和 CMOS 驱动芯片的良好键合，成功制备了分辨率为 640×360 的 Micro-LED 显示器。同时，以 FPGA 为控制器，设计了 Micro-LED 显示器的驱动系统，该驱动系统对输入的图像数据进行缓存并将其处理为显示器所能接收的单色图像数据，最后将处理后的图像数据写入驱动芯片并实现图像显示。

关键词： Micro-LED；有源矩阵；驱动系统；FPGA；高分辨率

中图分类号： TN873　**文章编号：** 1001－5868(2020)01－0068－05

Design and Driver Implementation of High Resolution AM Micro-LED Display

LIN Jiehong[1], NIE Junyang[2], ZHANG Yongai[1], ZHOU Xiongtu[1], YAN Qun[1], GUO Tailiang[1]

(1. Institute of Physics and Information Engin., Fuzhou University, Fuzhou 350116, CHN;
2. Faculty of Electronic and Information Engin., Xi'an Jiaotong University, Xi'an 710049, CHN)

Abstract: In this paper, a high-pixel micro-LED chip array was prepared based on a blue/green multi-quantum well LED epitaxial chip based on gallium nitride on sapphire substrate by using mature LED chip technology. Metal protruding points were designed and prepared on the surface of both micro-LED chip and CMOS driver chip, and good bonding between LED display chip and CMOS driver chip was achieved by adopting high-precision inversion welding process, and micro-LED display with resolution of 640×360 was successfully prepared. At the same time, the driver system of micro-LED display was designed with FPGA as the controller. Then the driver system can cache the input image data and process it into monochromatic image data which can be received by the display, and finally writes the processed image data into the driver chip to realize the image display.

Key words: micro-LED; active matrix; drive system; FPGA; high resolution

0　引言

近年来，Micro-LED 显示技术受到了广泛的关注，相较液晶显示（Liquid Crystal Display，LCD）技术和有机发光二极管显示（OLED）技术，Micro-LED 显示具有发光强度高[1]、响应速度快[2]、功耗低（自发光）[3]、寿命长[4]和分辨率高[5]等优势。目前，研究者已对 Micro-LED 的色彩转换[6-7]、发光效率增强[8]、散热改善[9]等方面做了大量的实验研究与探讨，获得了颇多研究成果。另外，Micro-LED 显示的微型化和响应速度快等特性，使其在光通信[10]和可穿戴显示领域[11]也拥有巨大的潜力。

目前，Micro-LED 显示的驱动方式主要分为两种：无源矩阵（PM）驱动和有源矩阵（AM）驱动。对于无源矩阵驱动，由于同一行或同一列的 LED 像素

收稿日期：2019－10－31.

基金项目：国家自然科学基金项目（61775038）；福建省自然科学基金项目（2017J01758，2017J01504，2019J01221）.

共享外围驱动电源的固定驱动电流，其尺寸和像素亮度会受到相应的限制。当前，研究人员已从工艺方面对无源矩阵 Micro-LED 显示进行了一系列优化[12]，还是无法提高显示像素分辨率。因此，对于高分辨率的 Micro-LED 微显示器，有源矩阵驱动技术是一个自然的选择。然而，对于 AM 驱动，其内部工艺构造更加复杂，驱动矩阵的漏电流控制也是一个难点[13]，并且随着像素分辨率的提高，尺寸变小，LED 芯片与驱动矩阵的精准对位将面临着更大的挑战。

对于 Micro-LED 显示来说，将大量的 LED 阵列准确转移到驱动基板上并键合（巨量转移技术[14]）是一个关键性问题；将 RGB 三色 LED 阵列准确转移到驱动基板上实现全彩显示，其难度高，良率低，是 Micro-LED 显示领域的一个技术瓶颈。然而，对于微显示器件而言，可以使用倒装芯片键合技术[15]，将倒装结构的 LED 阵列直接焊接在驱动基板上，工艺较为成熟，但由于在同一外延片上很难做 RGB 三色 LED，该技术路线无法实现全彩显示。本文使用倒装芯片键合的方式来进行 Micro-LED 显示屏的设计与研制。在本设计中，CMOS 控制 AM 蓝绿单色微显示器，并且对 LED 显示器件制备过程中的工艺进行了优化，令其 p 电极和 n 电极键合垫高度相同，降低了键合难度。该显示器分辨率为 640×360（1 323 PPI），并且设计了一套基于该显示器的完整的 Micro-LED 微显示驱动系统对其进行驱动。

1　倒装芯片键合 Micro-LED 显示屏的制备

1.1　Micro-LED 显示器件制备

本文设计的 Micro-LED 显示屏的分辨率为 640×360，像素间距（Pitch）为 19.2 μm，发光点直径为 10 μm，PPI 则高达 1 323，像素密度非常大，若采用外延转移则难度会较大，因此本文采用倒装芯片键合的方式进行设计，整体设计结构如图 1 所示。

Micro-LED 显示屏的具体制备流程如下：

（1）实验选择了 10.16 cm（4 inch）的蓝宝石衬底氮化镓基蓝色和绿色多量子阱（MQW）LED 外延片来制备 Micro-LED 阵列；

（2）在外延片 p-GaN 层上溅射一层氧化铟锡（ITO，60 nm）透明导电层，以有利于电流的扩散；

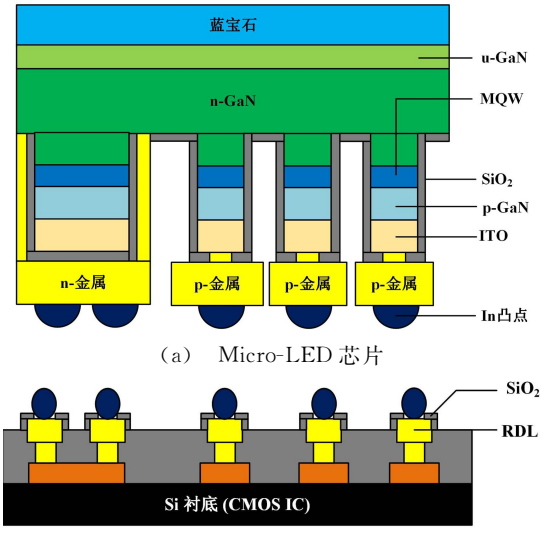
（3）在 ITO 上旋涂一层光刻胶，曝光显影后，利用湿法刻蚀将待露出的 n 极区域 ITO 刻蚀掉；以原有的光刻胶做掩膜，利用电感耦合等离子体蚀刻（ICP-RIE）至 n-GaN 区域，去除光刻胶后形成台阶；

（4）利用等离子增强化学气相沉积（PECVD）在样品表面沉积一层 SiO_2 绝缘层，经光刻和 ICP 刻蚀工艺，去除用作欧姆接触的 p 型和 n 型区域的 SiO_2 绝缘层；

（5）在样品表面旋涂光刻胶，曝光显影后，利用电子束蒸发沉积 p 型电极和 n 型电极（Cr/Pt/Au，35 nm/200 nm/1 500 nm），随后剥离（lift-off）多余的金属和光刻胶，形成具有相同高度的 n-键合垫和 p-键合垫，以有利于提高键合的良率；

（6）在表面旋涂光刻胶，曝光显影后，利用电阻热蒸发系统蒸镀 3.5 μm 铟膜，随后剥离多余的负胶和金属，形成焊接所需的铟凸点；

（7）为了克服铟在空气中容易氧化的问题，实验中通过氧化还原、回流工艺去除氧化层，形成球形状的焊点。图 2 为经过甲酸表面处理与回流后的 Micro-LED 芯片的铟凸点显微图片，由图可以看出样品具有良好的规整性，整体均匀，完备性好。



（a）　Micro-LED 芯片

（b）　CMOS 芯片

图 1　Micro-LED 芯片和 CMOS 芯片结构示意图

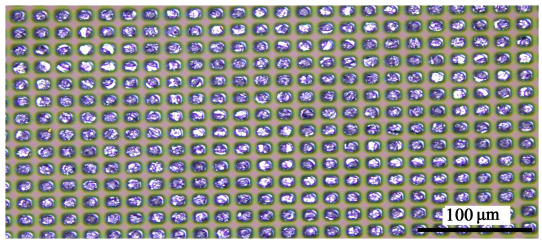


图 2　Micro-LED 阵列显微镜照片

1.2 CMOS 驱动基板和键合工艺设计

本实验采用台湾 Jasper 公司设计的 CMOS 驱动芯片，像素点间距(Pitch)为 19.2 μm，与制备好的 Micro-LED 芯片相匹配。为了实现 Micro-LED 显示芯片与 CMOS 驱动芯片的成功键合，首先对 CMOS 芯片进行如下工艺处理：

(1)在 CMOS 芯片钝化层上旋涂光刻胶，经曝光显影和 ICP 刻蚀，形成 CMOS 驱动芯片的漏极铝电极(如图 1 所示)；

(2)利用 lift-off 工艺，不去胶直接镀钛金保护铝窗口，随后去除多余的光刻胶；

(3)在样品表面旋涂光刻胶，曝光显影，采用电子束蒸发的方式沉积一层金膜，随后剥离掉多余的金属和光刻胶，形成重布线层(RDL)，这有利于后期的键合；

(4)在样品表面旋涂光刻胶，曝光显影后采用 PECVD 沉积一层 SiO_2，在其表面旋涂一层光刻胶，曝光显影后对 SiO_2 层进行 ICP 刻蚀，进行开孔；

(5)在表面进行光刻，随后采用电子束蒸发的方式蒸镀一层 3.5 μm 厚的铟膜，剥离多余的金属和胶，形成铟凸点(bump)；

(6)通过氧化还原、回流工艺去除氧化层，形成球形状的焊点，如图 3 所示。

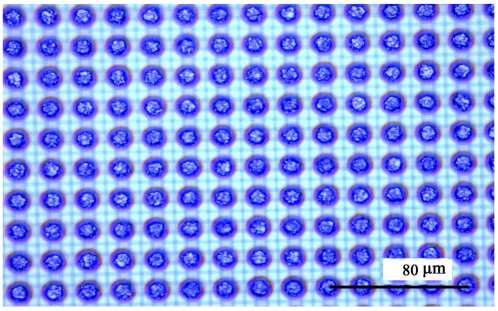


图 3 CMOS 芯片显微镜下实物图

实验采用 SET-FC-150 设备进行 Micro-LED 芯片和 CMOS 芯片的键合，形成 Micro-LED 显示屏。随后以 Wire Bonding 的方式对显示屏的 I/O 口与 FPCA 进行电气连接，以便于后期与外围驱动电路信号进行通信。

2 显示屏驱动系统的设计

2.1 驱动系统总体框架设计

Micro-LED 微显示器驱动系统总体设计框架如图 4 所示。该驱动系统选择 FPGA(现场可编程逻辑阵列)为控制器，对 HDMI 接口传输的 RGB 信号进行数据采集和处理，并进行缓存；将处理好的数据发送至驱动 IC 上进行图像显示。

本文设计的驱动系统由 HDMI 接口部分、HDMI 解码部分、FPGA(XC6SLX9)控制模块、SDRAM(H57V2562GTR)存储模块、驱动电路板模块以及驱动 IC 显示模块组成。其中，HDMI 接口及解码模块主要负责将 PC 机上的图像数据传输至 FPGA 进行处理；FPGA 控制模块是整个驱动系统的核心模块，负责对输入的图像数据进行采集及处理，对 SDRAM 乒乓操作功能模块的时序控制以及将处理好的图形数据传输给驱动 IC；SDRAM 存储模块主要负责对处理过的图像数据进行缓存，设计中由于数据的输入是连续不间断的，选择两片 SDRAM 来组成乒乓存储结构，以保证每一帧的数据都能够存储并输出到驱动 IC 上进行显示；驱动电路板模块负责对驱动 IC 的外部电路配置，包括升压部分和一些特定外围电路设计，用来实现 IC 与 FPGA 的电平匹配，以及对驱动 IC 的电源供给，使其能够正常工作。

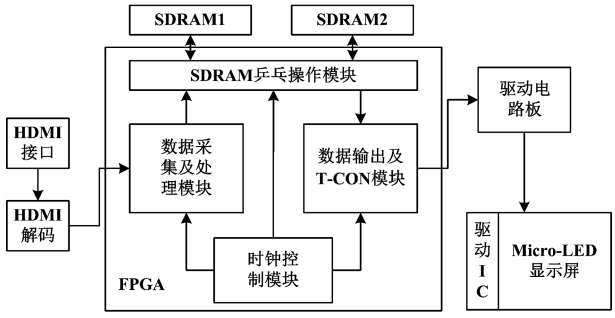


图 4 Micro-LED 微显示器的驱动系统框架图

2.2 驱动控制系统设计

2.2.1 HDMI 接口及解码模块

PC 机上的图像数据信号通过 HDMI 接口实时发送到 FPGA 上，传输的数据为 24 位全彩(RGB)的图像数据。由于 HDMI 接口传输的是 TMDS(最小化传输差分信号)信号，不能直接由 FPGA 进行处理，所以需要一个解码模块将 TMDS 信号转化为 RGB 并行信号，然后再传输至 FPGA 进行图像数据采集及处理。解码模块采用专门的 TMDS 信号解码芯片(TFP403PZP)来进行解码，然后将解码后的 24 位 RGB 信号发送至 FPGA 进行后续处理。

2.2.2 数据采集及处理模块

从解码芯片传输的图像数据为 24 位 RGB 信号，而本文设计的是单色显示屏，所以需要将 RGB

信号转化为 8 位单色灰度信号[16]，转化公式如式(1)所示。

$$Gray=(R\times 38+G\times 75+B\times 15)\gg 7 \quad (1)$$

式中，R 为图像中红色的强度，G 为绿色的强度，B 为蓝色的强度。

该模块具体框架如图 5 所示，当解码芯片 DE 信号为高时，代表一帧图像数据即将发送，此时使能 FIFO 将一帧 RGB 信号缓存，缓存完毕后将数据传入数据处理模块，最后得到 8 位单色灰度图像数据。由于 SDRAM 接口是 16 位数据传输，需要将得到的 8 位数据拼接成 16 位数据(高 8 位一个数据，低 8 位一个数据)，便于与 SDRAM 进行通信，将拼接好的数据存入异步 FIFO 中，最后发送至 SDRAM 中。

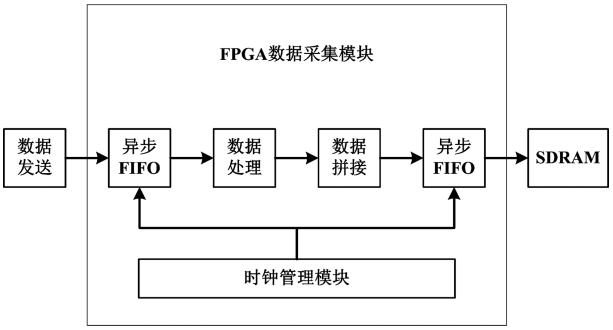


图 5　数据采集及处理模块框图

2.2.3　SDRAM 乒乓操作模块

为了能够对不断传入的图像数据不失真地缓存下来并进行显示，本文选择使用乒乓存储结构来实现实时存储。乒乓操作主要是利用两块 SDRAM 来分别进行读写操作并同时切换读写来实现不间断地缓存图像数据，当一片 SDRAM 中的数据读出完毕后，便切换为写入状态进行数据缓存，另一片则恰恰进行相反的操作，具体流程如图 6 所示。

2.2.4　数据输出及 T-CON 模块

该模块包含图像数据传输和驱动 IC 时序控制两个部分，将储存在 SDRAM 中的图像数据通过异步 FIFO 读出，并通过 SPI(Serial Peripheral Interface)接口与驱动 IC 进行通信，把一帧图像数据写入驱动 IC 内部 RAM 中进行显示，并且通过 SPI 接口对驱动 IC 内部控制寄存器进行配置，以达到实验想要的显示效果(亮度，对比度，输出电流等)，同时开启对显示屏的保护机制，防止因短路造成电流过大而烧毁显示屏的情况。

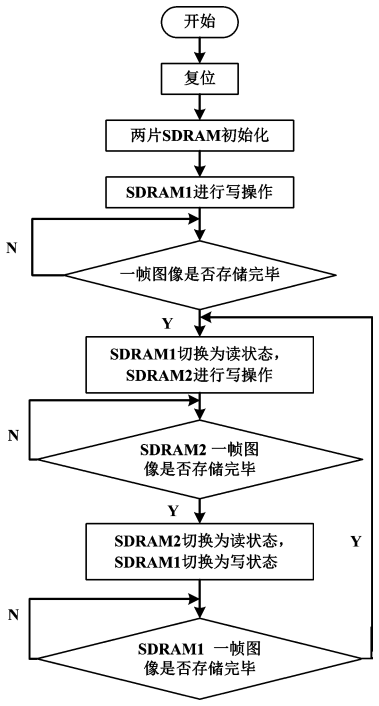


图 6　SDRAM 乒乓存储流程图

2.2.5　驱动电路模块

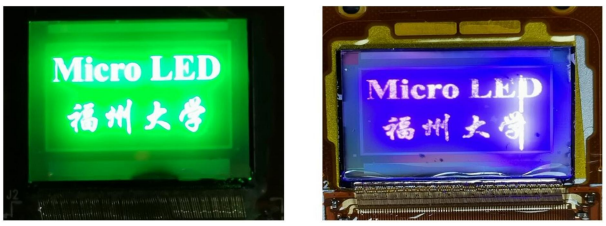
驱动电路模块的作用是对驱动 IC 的外围电路进行配置，确保驱动 IC 能正常工作，同时将重要的数据接口引出，使其能与 FPGA 进行通信。该模块主要包括升压部分和滤波部分，由于驱动 IC 的电源电压为 12 V，不同于该电路的驱动电压 3.3 V，所以需要使用线性电压源将 3.3 V 升至 12 V 作为驱动 IC 的电源电压，以及在高速通信中，需要在驱动 IC 接口附近设计滤波电路，以防外界干扰造成失真。

3　结果与展示

如图 7 所示，本文设计制作了蓝绿两块 Micro-LED 显示屏，图 7(a)为绿光 Micro-LED 显示屏，图 7(b)为蓝光 Micro-LED 显示屏。从中可看出由该驱动系统驱动显示的氮化镓基 AM Micro-LED 显示屏能够清晰地显示 PC 机上的影像。

本文在 LED 芯片制备过程中做了一个结构上的优化，即同时蒸镀 p 和 n 电极，使两个电极的键合垫处于同一个高度，这与另一种两个电极处于不同高度的方法不同[17]，相比较之下，优化结构更便于上下芯片对应金属凸点的相互接触，而且也减小了相邻金属凸点在压焊过程中因为形变而导致断路的可能性，从而提高键合的良率。本实验也验证了此方法的可行性，该工艺有望实现更高分辨率的

Micro-LED 显示屏。



(a) 绿光 Micro-LED (b) 蓝光 Micro-LED

图 7 样品显示效果图

从图 7(b)中可以看出,蓝光 Micro-LED 显示屏有一条明显的白线,还有一些黑点,该现象可能是由铟凸点短路以及还原不充分导致的虚焊引起的。

从图 7 也可以看出,本文设计的一整套基于该显示屏的驱动系统能够成功地对其进行驱动显示,包括文字、图像、视频显示,能够完美地匹配该实验中的 Micro-LED 显示屏。

参考文献:

[1] Huang Y G, Tan G J, Guo F W, et al. Prospects and challenges of mini-LED and micro-LED displays[J]. J. of The Society for Information Display, 2019, 27(7): 387-401.

[2] Lee V W, Twu N, Kymissis I. Micro-LED technologies and applications[J]. Information Display, 2016, 32(6): 16-23.

[3] Burgess D S. LED focus: Hybrid LEDs offer full-color emission[J]. Photonics Spectra, 2005, 36(11): 117.

[4] Henry W. MicroLED arrays find applications in the very small [J]. Photonics Spectra, 2013, 47(3): 52-55.

[5] Li J J, Nie X M, Li G S, et al. Comparison and research progress of flat panel display technology[J]. Chinese Opt., 2018, 11(5): 695-710.

[6] Guo F W, Hsiang E L, Tan G J, et al. High performance color-converted micro-LED displays[J]. J. of The Society for Information Display, 2019, 27(4): 199-206.

[7] Han H V, Lin H Y, Lin C C, et al. Resonant-enhanced full-color emission of quantum-dot-based micro LED display technology[J]. Opt. Express, 2015, 23(25): 32504-32515.

[8] Hong I Y, Lee J H, Cho S M, et al. Impact of hydrothermally grown ZnO nanorods on external quantum efficiency of 32×32 pixelated InGaN/GaN micro-LED array [J]. IEEE Trans. on Nanotechnol., 2019, 18: 160-166.

[9] Lobo P N, Rodriguze H, Stolmacker C, et al. Effective thermal management in ultraviolet light-emitting diodes with micro-LED arrays[J]. IEEE Trans. on Nanotechnol., 2013, 60(2): 782-786.

[10] Carreira J F C, Xie E, Bian R, et al. On-chip GaN-based dual-color micro-LED arrays and their application in visible light communication[J]. Opt. Express, 2019, 27(20): A1517-A1528.

[11] O'Shea P. New process helps make micro-LED arrays for next-gen huds: Gan and ingan technology targets fast-growing markets for wearable vision systems[J]. Electronic Products, 2015, 57(8): August 1.

[12] Kang C M, Kong D J, Shim J P, et al. Fabrication of a vertically-stacked passive-matrix micro-LED array structure for a dual color display[J]. Opt. Express, 2017, 25(3): 2489-2495.

[13] Zhang K, Peng D, Chong W C, et al. Investigation of photon-generated leakage current for high-performance active matrix micro-LED displays[J]. IEEE Trans. on Electron Devices, 2016, 63(12): 4832-4838.

[14] Zhang L, Ou F, Chong W C, et al. Wafer-scale monolithic hybrid integration of Si-based IC and Ⅲ-Ⅴ epi-layers-A mass manufacturable approach for active matrix micro-LED micro-displays[J]. J. of The Society for Information Display, 2018, 26(3): 137-145.

[15] Liu Z J, Wong K M, Keung C W, et al. Monolithic LED microdisplay on active matrix substrate using flip-chip technology [J]. IEEE J. on Sel. Top. in Quantum Electronics, 2009, 15(4): 1298-1302.

[16] Luo Z J, Zhang W N, Liu L W, et al. Portable multi-gray scale video playing scheme for high-performance electrowetting displays[J]. J. of The Society for Information Display, 2016, 24(6): 345-354.

[17] Liu Z J, Chong W C, Wong K M, et al. 360PPI flip-chip mounted active matrix addressable light emitting diode on silicon(LEDoS) micro-displays[J]. J. of Display Technol., 2013, 9(8): 678-682.

作者简介:

林杰泓(1993—),男,福建龙岩人,硕士研究生,主要研究方向为 Micro-LED 显示驱动技术、图像处理;

张永爱(1977—),男,江西抚州人,博士,研究员,博士生导师,研究方向为印刷电子材料与器件、3D 显示、Micro-LED 显示和微纳材料与器件等。

E-mail: yongaizhang@fzu.edu.cn

光电器件

DOI：10.16818/j.issn1001－5868.2020.01.015

光纤光栅温度传感器增敏封装特性研究

王　甫[1,2]，江　淮[1,2]，易小龙[1,2]，薛渊泽[1,2]，唐才杰[1,2]

（1. 北京航天控制仪器研究所，北京 100854；2. 北京光纤传感系统工程技术研究中心，北京 100094）

摘　要： 通过利用聚酰亚胺涂覆裸光纤光栅的方法研制了一种增敏型光纤光栅温度传感器，其反射波长的温度敏感系数由曾敏前的10 pm/℃提升至40 pm/℃。在－40～200 ℃测试环境中，开展了传感器标定和温度精度测量试验。试验结果表明：利用光纤光栅增敏封装，在高低温环境中采用低精度光纤光栅解调仪，传感器的测温精度由±0.4 ℃提升至±0.1 ℃。上述结果显示，采用增敏封装结构能够提升光纤光栅温度传感器的测量精度，减低制作成本。

关键词： 光纤传感技术；光纤光栅；温度传感；增敏封装；测量精度

中图分类号： TN247　　**文章编号：** 1001－5868(2020)01－0073－04

Sensitivity Enhancement and Encapsulation of Fiber Bragg Grating Temperature Sensors

WANG Fu[1,2], JIANG Huai[1,2], YI Xiaolong[1,2], XUE Yuanze[1,2], TANG Caijie[1,2]

(1. Beijing Institute of Aerospace Control Devices, Beijing 100854, CHN;
2. Beijing Engin. Research Center of Optical Fiber Sensing System, Beijing 100094, CHN)

Abstract: In this paper, the sensitized fiber Bragg grating temperature sensor was designed and developed by coating the naked fiber Bragg grating with polyimide, and its temperature sensitivity coefficient was enhanced from 10 to 40 pm/℃. Tests were performed on the sensor calibration and temperature accuracy measurement under the test environment of －40～200 ℃. The test results show that the temperature measurement accuracy can be improved from ±0.4 ℃ to ± 0.1 ℃ by using fiber grating sensitization package and low-precision fiber grating demodulation instrument in high and low temperature environment. The material characteristics of the fiber Bragg grating temperature sensor with sensitized package were analyzed which can be used to improve the accuracy of temperature measurement, greatly reduce the application cost of fiber Bragg sensor, and lay a foundation for the comprehensive use of fiber Bragg grating temperature measurement.

Key words: optical fiber sensing; fiber Bragg grating; temperature sensing; encapsulation and enhanced sensitivity; accuracy of temperature measurement

0　引言

光纤光栅传感器具有体积小、重量轻、测量灵敏度高、复用能力强、抗电磁干扰、易于嵌入材料内部等诸多优点[1-2]。常规光纤光栅对于温度的线性敏感系数为10 pm/℃，对于高精度的温度测量而言，则需要依靠高精度光纤光栅解调仪来检测光纤光栅中心波长的变化。提高光纤光栅温度传感器的精度有两种途径：一是通过提升解调仪的解调精度，但其成本高，研制技术难度大；另一种方法通过对光纤光栅进行增敏封装，来提升光纤光栅的线性敏感系数[3]，这种方案的工艺较为复杂。目前用于光纤光栅增敏封装的材料较多，可分为金属和非金属两大类[4-5]。由于封装材料的性能不同，选择不同的材料进行封装，传感器的温度灵敏度有较大的差异。封装工艺和结构对光纤光栅的温度灵敏度和性能有着重要影响。已有的关于光纤光栅增敏封装的研究结

收稿日期：2019－07－27.

果表明，金属封装可大幅改变光纤光栅的温度灵敏度，而对应变灵敏度却几乎没有影响；聚合物封装可更大幅度地提高温度应变传感性能，但存在易老化、使用寿命短、适用温度范围小等缺陷[6-7]。李玉龙等人[8]利用光纤金属化保护方法，将光纤光栅嵌入铝合金等金属中，封装后光纤光栅的温度灵敏度显著提高，但由于镀层与光纤之间的接触应力造成了波长稳定性的降低。张燕君等人[9]提出了一种光纤光栅的毛细钢管封装工艺，封装后的传感器温度灵敏度可到 28.9 pm/℃，且具有较好的线性度和重复性。于秀娟等人[10]采用钛合金对光纤光栅进行片状封装，封装后其温度灵敏度系数提高至 19.87 pm/℃，且封装结构简单、易操作。本文设计了片状金属封装结构，实现光纤光栅敏感系数的提升。

1 原理和实验

由光波导模式耦合理论可知，周期均匀的光纤光栅反射中心波长 λ_B 满足下式：

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

式中，n_{eff} 为有效折射率，Λ 为光栅周期。

当 FBG 被粘贴固定在基底材料表面时，基底材料的热膨胀和变形将同时作用于 FBG，导致其中心波长发生漂移。此时 FBG 的温度灵敏度系数可表示为[11]

$$K_T = \alpha + \varepsilon + (1 - P_e)(\alpha_s - \alpha) \tag{2}$$

式中，K_T 为温度灵敏度系数，α 表示光纤的热膨胀系数，α_s 表示基底材料的热膨胀系数，ε 为光纤材料的热光系数，P_e 为光纤的有效弹光系数。

对于增敏封装后的光纤光栅，若直接以理论公式推导测量温度值，而未经参考标定，将导致实际测量误差较大。温度 T 与波长 λ 的对应关系，可采用以下三次多项式来进行描述：

$$T = T_0 + K_1\left(\frac{\lambda-\lambda_0}{\lambda_0}\right) + K_2\left(\frac{\lambda-\lambda_0}{\lambda_0}\right)^2 + K_3\left(\frac{\lambda-\lambda_0}{\lambda_0}\right)^3 \tag{3}$$

通过最小二乘法线性拟合确定上式的系数 K_1，K_2，K_3，T_0 和 λ_0，其中 T_0 和 λ_0 是相关联的两个系数。选取温度稳定点，并查找对应的波长，利用三次多项式拟合算法得到传感器的温度波长曲线，并计算得到测试点的温度值，以铂电阻测温值作为参考，计算传感器测量温度偏差。

图 1 为光纤光栅温度传感器温度标定和温度测量试验原理图，使用 SUNTECK 高精度环境试验箱、Fluke 公司生产的 1 502 A 型高精度铂电阻温度计、MOI 公司生产的 SM125 型高精度波长解调仪以及低精度波长解调仪，对研制的增敏温度传感器进行标定和性能测试。将铂电阻温度计与待测温度传感器放置于温度工装中，根据需求利用环境试验箱提供合适的环境温度，并在每个温度点恒温保持 40 min，确保温度工装内铂电阻温度测量值和传感器波长稳定后，进行下一个温度点的测试。基于 LabVIEW 平台，利用 RS232 串口通信完成铂电阻温度计对温度的实时测量，通过网口通信，将 MOI 低成本解调仪解调出来的光纤光栅反射波长传输至 PC 终端。利用时序关系，实现铂电阻温度测量值和光纤光栅反射波长的一一对应，最终利用 matlab 软件，对测试数据进行分析，确定传感器的温度测量精度。

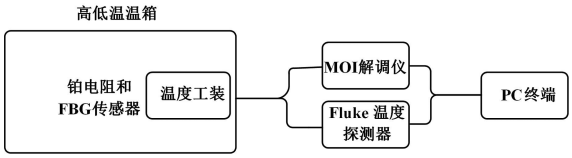


图 1 温度测量试验原理图

2 实验结果分析与讨论

FBG 温度传感器外壳选用 7075T6 航天专用铝制材料作为传感器增敏封装材料，采用无机胶作为固定胶来固定传感器，并对传感器尾纤进行保护。将增敏 FBG 温度传感器和铂电阻温度计置于温度均匀的温度工装中进行固化和老炼试验，并最终完成增敏 FBG 温度传感器的研制。研制的增敏封装温度传感器的尺寸为 35 mm×10 mm×5 mm。温度工装如图 2 所示。

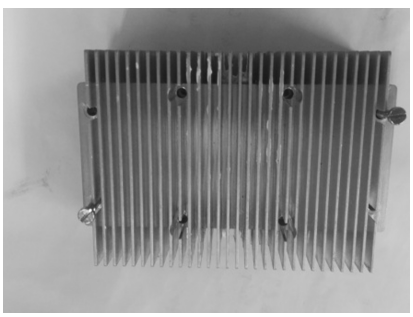


图 2 温度工装照片

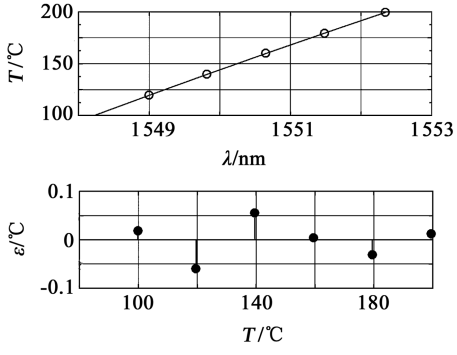
2.1 高温环境试验

将增敏 FBG 温度传感器、非增敏 FBG 温度传感器以及铂电阻温度计置于温度工装中，并将其置

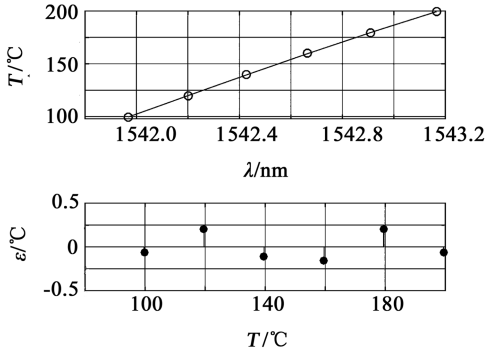
于高温环境试验箱中进行高温环境试验。设置试验箱温度分别为 100，120，140，160，180，200 ℃，并在各温度点稳定 40 min，以确保温度工装内温度一致。

图 3(a)、(b)分别为增敏 FBG 温度传感器和非增敏 FBG 温度传感器高温试验结果。可以看出增敏 FBG 温度传感器稳定性更优良，温度测量误差 ε 在±0.06 ℃，而非增敏 FBG 温度传感器的温度测量误差为±0.2 ℃。

利用标定的三次项系数和初始温度、初始波长等参数，将 100，120，140，160，180 和 200 ℃等 6 个温度点对应的波长，代入方程(3)，得到预测温度，并与此时铂电阻温度计的测量温度做差值，作为温度测量误差。可以看出在 100～200 ℃的测量范围内，采用低精度解调仪和增敏 FBG 温度传感器，温度测量精度为±0.1 ℃。而采用非增敏 FBG 温度传感器的温度测量误差在±0.4 ℃。其原因在于，光纤光栅解调仪精度受限，致使普通温度传感器测量偏差较大；而利用增敏 FBG 温度传感器时，波长与温度的线性系数增大，极大降低了测量精度对解调仪精度的依赖，进而实现了较好的测温精度，如图 4 所示。



(a)　增敏 FBG 温度传感器



(b)　非增敏 FBG 温度传感器

图 3　高温情况下两类传感器的标定偏差

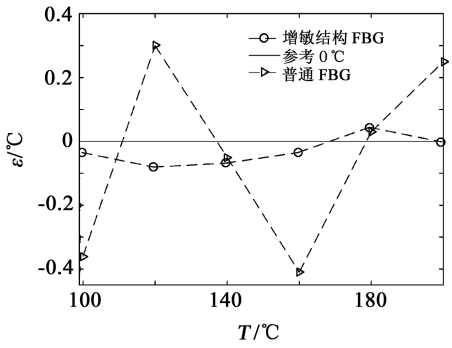
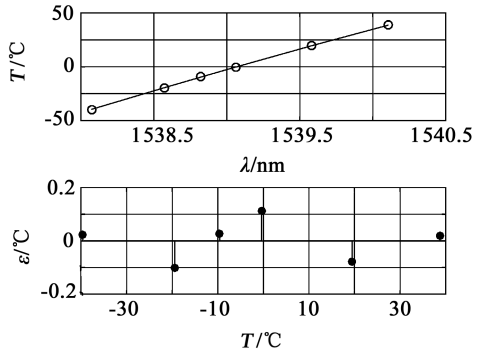


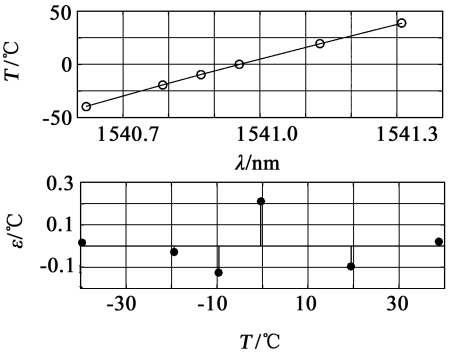
图 4　高温下两类传感器测温精度对比

2.2　低温环境试验

采用与上述相同的温度精度测量方法，进行低温－40～40 ℃的温度测量校准试验验证，结果如图 5 所示。从图 5 可以看出，增敏 FBG 温度传感器稳定性更优良，温度偏差在±0.1 ℃，而非增敏 FBG 温度传感器的温度偏差为±0.2 ℃。



(a)　增敏 FBG 温度传感器



(b)　非增敏 FBG 温度传感器

图 5　低温情况下两类传感器的标定试验结果

利用标定的三次项系数和初始温度、初始波长等参数，将 40，20，0，－10，－20，－30 和－40 ℃等 7 个温度点对应的波长，代入方程(3)，得到预测温度，并与此时铂电阻的测量温度做差值，得到测量温度偏差，如图 6 所示。可以看出在－40～40 ℃的测量范围内，采用低精度波长解调仪和增敏 FBG 温度传

感器，测量温度精度范围为－0.09～0.09 ℃。而采用非增敏 FBG 温度传感器温度偏差范围为－0.2～0.38 ℃。其原因在于，自研低精度波长解调仪精度受限，致使非增敏 FBG 温度传感器测量偏差较大；而利用增敏 FBG 温度传感器时，波长与温度的线性系数增大，极大减少了测量精度对解调仪精度的依赖，进而实现了较好的测温精度。

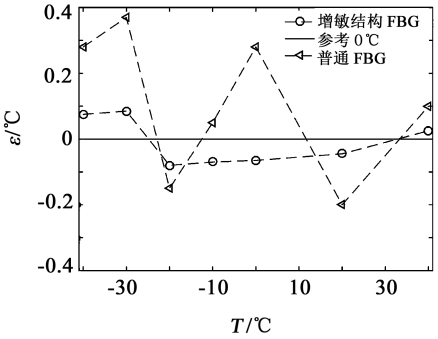


图 6 低温下温度测量精度

3 结论

通过对传统光纤光栅温度传感器进行增敏和封装保护，研制了一种光纤光栅温度增敏传感器。与传统光纤光栅温度传感器相比，增敏封装后传感器的温度测量精度得到提升，突破了光纤光栅温度传感器测量精度强烈依赖于光纤光栅解调仪解调精度的限制。搭建了试验系统，试验结果显示，在－40～200 ℃的温度范围内，系统实现了±0.1 ℃的测温精度，与非增敏封装的光纤光栅传感器的±0.4 ℃的测量精度相比，得到了大幅度的提升。

参考文献：

[1] 刘铁根，王 双，江俊峰，等. 航空航天光纤传感技术研究进展[J]. 仪器仪表学报，2014(8)：1681-1692.
Liu Tiegen，Wang Shuang，Jiang Junfeng，et al. Advances in optical fiber sensing technology for aviation and aerospace application[J]. Chinese J. of Scientific Instrument，2014(8)：1681-1692.

[2] 姜德生，何 伟. 光纤光栅传感器的应用概况[J]. 光电子·激光，2002，13(4)：420-430.
Jiang Desheng，He Wei. Review of applications for fiber Bragg grating sensors[J]. J. of Optoelectronics·Laser，2002，13(4)：420-430.

[3] 谢剑锋，张 华，张国平，等. 封装材料性能对光纤布拉格光栅温度灵敏度影响分析[J]. 光电子·激光，2008，19(9)：1158-1162.
Xie Jianfeng，Zhang Hua，Zhang Guoping，et al. Analysis on influence of clad materials on temperature sensitivity of fiber Bragg grating[J]. J. of Optoelectronics·Laser，2008，19(9)：1158-1162.

[4] 温昌金. 光纤布拉格光栅的增敏封装及定位安装技术研究[D]. 南昌：南昌大学，2016.
Wen Changjin. Research on the sensitizing package，positioning and installation technology of fiber Bragg grating [D]. Nanchang：Nanchang University，2016.

[5] 赵雪峰，田石柱，周 智，等. 钢片封装光纤光栅监测混凝土应变试验研究[J]. 光电子·激光，2003，14(2)：171-174.
Zhao Xuefeng，Tian Shizhu，Zhou Zhi，et al. Experimental study on strain monitoring of concrete using a steel slice packaged fiber grating[J]. J. of Optoelectronics·Laser，2003，14(2)：171-174.

[6] 孙 安，乔学光，贾振安，等. 聚合物封装光纤布拉格光栅传感器温度压力特性研究[J]. 中国激光，2005，32(2)：224-227.
Sun An，Qiao Xueguang，Jia Zhen'an，et al. Temperature and pressure responsive characteristics of polymer packaged fiber Bragg grating with large dynamic range[J]. Chinese J. of Lasers，2005，32(2)：224-227.

[7] 张学智，祝连庆，张荫民，等. 光纤光栅非金属耐腐蚀封装及其温度特性研究[J]. 激光与红外，2015，45(4)：437-441.
Zhang Xuezhi，Zhu Lianqing，Zhang Yinmin，et al. Fiber Bragg grating with nonmetal and corrosion resisting packaging and its temperature characteristic[J]. Laser & Infrared，2015，45(4)：437-441.

[8] Feng Y，Zhang H，Li Y L，et al. Temperature sensing of metal-coated fiber Bragg grating[J]. IEEE/ASME Trans. on Mechatronics，2010，15(4)：55-519.

[9] 张燕君，王海宝，陈泽贵，等. 光纤光栅毛细钢管封装工艺及其传感特性研究[J]. 激光与红外，2009，39(1)：53-55.
Zhang Yanjun，Wang Haibao，Chen Zegui，et al. Study on FBG sensor's steel capillary packaging technique and sensing properties[J]. Laser & Infrared，2009，39(1)：53-55.

[10] 于秀娟，余有龙，张 敏，等. 钛合金片封装光纤光栅传感器的应变和温度传感特性研究[J]. 光电子·激光，2006，17(5)：564-567.
Yu Xiujuan，Yu Youlong，Zhang Min，et al. Strain and temperature sensing characteristics of FBG packaged by the titanium alloy slice[J]. J. of Optoelectronics·Laser，2006，17(5)：564-567.

[11] Liu Y Q，Guo Z Y，Zhang Y，et al. Simultaneous pressure and temperature measurement with polymer-coated fiber Bragg grating[J]. Electron. Lett.，2000，36(6)：564-566.

作者简介：

王 甫(1987－)，男，博士，主要从事光纤传感、光电仪器计量方面的工作。

E-mail：wf_yy1987@163.com

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.016

不同应力增透膜对半导体激光器性能的影响

崔碧峰，程　瑾，郝　帅，李彩芳，王豪杰

（北京工业大学 光电子技术省部共建教育部重点实验室，北京 100124）

摘　要： 采用离子辅助电子束蒸发方法，通过改变制备 Al_2O_3 增透膜时基底的温度，在边发射半导体激光器前腔面上分别制备了张应力和压应力增透膜。比较了张应力、压应力两种不同增透膜对半导体激光器性能的影响。结果表明：在10 A注入电流下，当半导体激光器的增透膜为张应力状态时，输出功率为8.11 W；当半导体激光器的增透膜为压应力状态时，输出功率为7.74 W。可见，在半导体激光器前腔面制备张应力增透膜有效地提高了半导体激光器的斜率效率。

关键词： 半导体激光器；薄膜应力；斜率效率

中图分类号： TN248.4　**文章编号：** 1001－5868(2020)01－0077－03

Effect of Different Facet Coating Stress on The Performance of Laser Diodes

CUI Bifeng, CHENG Jin, HAO Shuai, LI Caifang, WANG Haojie

(Key Lab. of Opto-Electronics Technol. of The Ministry of Education, Faculty of Information Technol., Beijing University of Technol., Beijing 100124, CHN)

Abstract: The ion-assisted electron beam evaporation was used to fabricate laser diodes with tensile stress and compressive stress anti-reflective film coated on the front facet by changing the temperature of the substrate when preparing the Al_2O_3 film. The effects of two different anti-reflection coatings on the performance of laser diodes were compared. The results show that output power of the laser diode with tensile stress anti-reflective coating was 8.11 W while the laser diode with compressive stress anti-reflective coating was 7.74 W at 10 A injection current. Therefore, fabricating tensile stress anti-reflective coated on the laser diode facet can more effectively increase the slope efficiency of laser diodes.

Key words: laser diodes; film stress; slope efficiency

0　引言

在激光技术的发展中，半导体激光器由于体积小、重量轻、寿命长、电光转换效率高、覆盖的波段范围广、价格低等优点，得到了广泛的应用[1]。通过在激光器的前后腔面分别制备增透膜和高反膜，可以使前、后腔面的光功率比最大化，提高器件单面输出的斜率效率[2]。而且腔面膜可使激光器腔面免受氧化、沾污等影响，延长激光器的使用寿命[3]。

顾培夫[4]、Thielsch等人[5]在研究深紫外薄膜的机械性能时比较了电子束蒸发、离子束溅射、等离子体辅助沉积等工艺制备的薄膜，发现制备工艺参数对薄膜应力有影响，通过调节基底温度能够控制张应力或压应力薄膜的生长。

应力的变化能够影响多种器件性能。在MEMS/NEMS器件中，通过增加双夹紧梁薄膜中的张应力能够提高器件的共振振动频率[6]；在微镜器件中，通过增加悬挂微镜的扭杆中的张应力能够扩大微镜器件的旋转角度[7]；在应变硅MOSFET器件中，通过在弛豫 $Si_{1-x}Ge_x$ 衬底上生长薄的外延Si沟道引入双轴张应力[8]，或者引入纵向单轴张应力[9]能够增强电子迁移率；在Ge-on-Si LED器件中，由于热膨胀系数的差异，在Si衬底上实现了张

收稿日期：2019－11－26.

基金项目：国家自然科学基金项目(60908012，61575008，61775007)；北京市自然科学基金项目(4172011)；北京市教育委员会项目(040000546319525).

应力应变的 Ge 层，改变了 Ge 的带隙结构，从而能够改善该器件的光学特性[10]。而针对在半导体激光器腔面制备增透膜所产生的应力对其性能的影响鲜有研究。

本文采用离子辅助电子束蒸发的方法，通过改变基底温度在半导体激光器腔面上制备不同应力的 Al_2O_3 增透膜，并进行光功率测试，分析了不同应力增透膜对半导体激光器功率特性的影响。

1 实验过程

利用 MOCVD 设备外延生长了 980 nm InGaAs/GaAs/AlGaAs 应变量子阱激光器，并制备成条宽为 100 μm、腔长为 4 000 μm 的边发射半导体激光器(LD)。图 1 为半导体激光器芯片横截面示意图，前后腔面分别制备了低反射率 Al_2O_3 增透膜层(AR)和 4 对高反射率 SiO_2/TiO_2 高反膜层(HR)。由图 1 可知，增透膜 Al_2O_3 直接附着在半导体激光器前腔面上，此层薄膜的应力对器件性能影响最大。

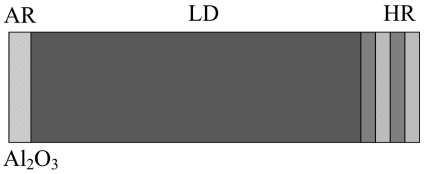


图 1 LD 芯片的横截面示意图

实验所用镀膜设备为 Denton Vacuum Infinity 26 真空镀膜机，该设备配备有 CC-105 冷阴极离子源系统、基底加热装置和石英晶体速率控制器。离子辅助电子束蒸发方法能够通过改变工艺参数制备出不同应力的薄膜，对于半导体激光器，离子源电压、真空度、离子源气体流量等参数对成膜质量影响大，而在低温生长(473 K 以下)时，基底温度对薄膜生长质量影响小[11-12]。因此，仅改变制备 Al_2O_3 薄膜时的基底温度，而离子源电压等其他条件均保持在同一稳定的数值，分别在 323，373 和 432 K 温度下制备 Al_2O_3 薄膜。

采用 BGS6341A 型电子薄膜应力分布测试仪在室温下进行测试，其原理是利用曲率法测试应力，通过测量基底镀膜前后的曲率变化来计算薄膜应力。根据测得的曲率半径，由 Stoney 公式[13]，可求出薄膜应力：

$$\sigma = \frac{4E_s}{3(1-\mu_s)} \frac{d_s^2}{d_f} \frac{h-h_0}{a^2} \tag{1}$$

式中，σ 为制备薄膜后的应力，E_s 为基底材料的杨氏模量，μ_s 为材料泊松比，d_s 为基底材料厚度，d_f 为薄膜材料厚度，α 表示基底测量范围，h_0 和 h 分别为基底制备薄膜前后的弯曲量。因此，只要测出镀膜前后对应的表征基板弯曲程度的 h_0 和 h 值，就可以计算得到薄膜的应力。

2 实验结果与讨论

图 2 为衬底温度与薄膜应力之间的关图。从中可见，当衬底温度在 323～380 K 变化时，Al_2O_3 薄膜应力发生了明显变化。当衬底温度在 380～432 K 变化时，应力变化缓慢。在 323 和 432 K 处，Al_2O_3 薄膜分别处于压应力状态和张应力状态。因此，在保持其他参数(如离子源电流、电压、气体流速和沉积速率)相同的情况下，选择基底温度为 323 和 432 K 制备了不同应力的 Al_2O_3 薄膜。

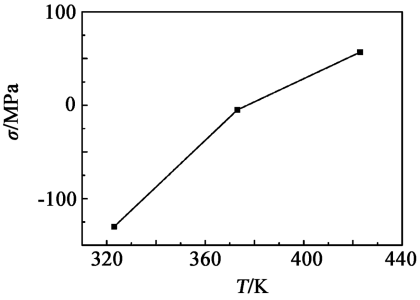


图 2 不同衬底温度下制备的 Al_2O_3 薄膜的应力

由于两次实验使得薄膜厚度略有差异，薄膜反射率 R 稍有不同。图 3 为椭偏仪测试的 Al_2O_3 薄膜反射率曲线，由图所示曲线可以看出，增透膜表现为压应力时，在 980 nm 处，测得的反射率为 3.152%。增透膜表现为张应力时，在 980 nm 处，测得的反射率为 3.59%。

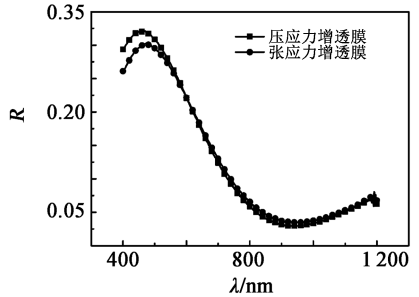


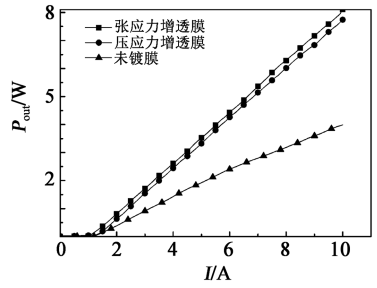
图 3 不同应力下增透膜的反射率

激光器出光面的外微分量子效率与腔面膜反射率的关系如式(2)：

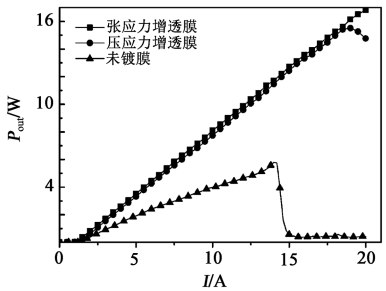
$$\eta_d = \eta_i \frac{\ln(R_1R_2)^{-1}}{4\alpha_i + \ln(R_1R_2)^{-1}} \quad (2)$$

其中，η_i 为内量子效率，η_d 为外微分量子效率(与斜率效率成正比)，α_i 为内损耗，R_1 为出光面的反射率，R_2 为激光器另一个腔面的光反射率[14]。根据式(2)知，上述制备的薄膜的反射率差异引起的外微分量子效率的差异约为 0.1%。

图 4 为未制备腔面膜、制备了张应力增透膜和压应力增透膜以及相同高反膜半导体激光器的输出功率-电流特性曲线。图 4(a)中，注入电流为 10 A 时，未镀膜半导体激光器的斜率效率和输出功率分别为 0.46 W/A 和 3.99 W。在相同的 10 A 注入电流下，制备张应力增透膜的半导体激光器的斜率效率和输出功率分别增加到 0.91 W/A 和 8.11 W，而制备压应力增透膜的半导体激光器的斜率效率和输出功率分别增加到 0.90 W/A 和 7.74 W。在图 4(a)中，增透膜表现为张应力的半导体激光器和增透膜表现为压应力的半导体激光器的斜率效率分别比未制备腔面膜的半导体激光器提高了 98%和 96%。图 4(b)显示了三种半导体激光器在 20 A 注入电流下的输出功率。在该图中，当电流为 19 A 时增透膜表现为压应力的半导体激光器的输出功率降低，而在 20 A 注入电流下，增透膜表现为张应力的半导体激光器的输出功率没有降低，输出功率为 16.83 W。因此，在半导体激光器腔面制备张应力增透膜不仅改善了单面斜率效率，而且提高了热饱和功率。



(a)　10 A 注入电流下



(b)　20 A 注入电流下

图 4　不同激光器的斜率效率曲线

3　结论

实验表明，在半导体激光器的前腔面上制备不同应力的增透膜会对器件的性能产生影响。测试结果表明，在 10 A 注入电流下，张应力增透膜半导体激光器输出功率为 8.11 W，而压应力增透膜半导体激光器的输出功率为 7.74 W。因此，在半导体激光器的前腔面制备张应力增透膜可以更有效地提高半导体激光器的斜率效率并且可以提高半导体激光器的热饱和功率。

参考文献：

[1]　王立军，宁永强，秦　莉，等. 大功率半导体激光器研究进展[J]. 发光学报，2015，36(1)：1-19.
Wang Lijun，Ning Yongqiang，Qin Li，et al. Development of high power diode laser[J]. Chin. J. Lumin.，2015，36(1)：1-19.

[2]　曹玉莲，王　乐，廖新胜，等. 大功率半导体激光器的可靠性研究[J]. 发光学报，2003，24(1)：100-102.
Cao Yulian，Wang Le，Liao Xinsheng，et al. Reliability of high-power semiconductor laser diodes[J]. Chin. J. Lumin.，2003，24(1)：100-102.

[3]　套格套，路国光，尧　舜，等. 808 nm 大功率半导体激光器腔面膜优化设计[J]. 半导体光电，2007，28(6)：778-780，784.
Tao Getao，Lu Guoguang，Yao Shun，et al. Optimized design of cavity facet coatings of 808 nm high power semiconductor laser[J]. Semiconductor Optoelectronics，2007，28(6)：778-780，784.

[4]　顾培夫，郑臻荣，赵永江，等. TiO_2 和 SiO_2 薄膜应力的产生机理及实验探索[J]. 物理学报，2006，55(12)：268-272.
Gu Peifu，Zheng Zhenrong，Zhao Yongjiang，et al. Mechanism and experimental exploration of stress in TiO_2 and SiO_2 thin films[J]. Acta Physica Sinica，2006，55(12)：268-272.

[5]　Thielsch R，Gatto A，Kaiser N. Mechanical stress and thermal-elastic properties of oxide coatings for use in the deep-ultraviolet spectral region[J]. Appl. Opt.，2002，41(16)：3211-3217.

[6]　Gaspar J，Chu V，Conde J P. Electrostatic microresonators from doped hydrogenated amorphous and nanocrystalline silicon thin films[J]. J. of Microelectromechanical Systems，2005，14(5)：1082-1088.

[7]　Sasaki M，Yuki S，Hane K. Large-rotation and low-voltage driving of micromirror realized by tense thin-film torsion bar[J]. IEEE Photon. Technol. Lett.，2006，18(15)：1573-1575.

[8]　Thompson S E，Armstrong M，Auth C，et al. A logic nanotechnology featuring strained-silicon[J]. IEEE Electron Device Lett.，2004，25(4)：191-193.

(下转第 84 页)

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.017

温度和材料参数对 $InAs_xSb_{1-x}$ 俄歇复合寿命影响的数值分析

张景波[1]，张云琦[1]，王思文[2]，孙晓冰[1]，邢春香[1]，刘　强[2]，殷景志[2]

（1. 长春理工大学 光电信息学院，长春 130114；
2. 吉林大学 电子科学与工程学院 集成光电子学国家重点实验室，长春 130012）

摘　要： 工作在中长波的红外探测器可被广泛应用在空间成像、军事和通信等领域，锑基InAsSb材料由于其特殊的性质是制作长波非致冷光子探测器的理想材料。俄歇复合寿命是影响探测器性能的重要因素之一，文章采用Matlab软件模拟研究了n型和p型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命随温度、As组分及载流子浓度的变化。对确定的As组分，可通过优化工作温度及载流子浓度获得较长的俄歇复合寿命。当载流子浓度为 3.2×10^{15} cm^{-3}、温度为200 K时，n型 $InAs_{0.35}Sb_{0.65}$ 的俄歇复合寿命最大为 2.91×10^{-9} s。

关键词： InAsSb；俄歇复合寿命；As组分；温度；载流子浓度

中图分类号： TN215　**文章编号：** 1001－5868(2020)01－0080－05

Studying Influence of Temperature and Material Parameters on $InAs_xSb_{1-x}$ Auger Recombination Life by Numerical Analysis

ZHANG Jingbo[1], ZHANG Yunqi[1], WANG Siwen[2], SUN Xiaobing[1], XING Chunxiang[1], LIU Qiang[2], YIN Jingzhi[2]

(1. College of Optical and Electronical Information, Changchun University of Science and Technol., Changchun 130114, CHN;
2. State Key Lab. of Integrated Optoelectron., College of Electronic Science and Engin., Jilin University, Changchun 130012, CHN)

Abstract: Infrared detectors working in the medium and long wavelength can be widely used in the fields of space imaging, military, communication and so on. Antimony-based InAsSb materials are ideal materials for fabricating long-wave uncooled photon detectors due to their special properties. Auger recombination lifetime is one of the important factors affecting the performance of detectors . In this paper, Auger recombination lifetime of n and p type $InAs_xSb_{1-x}$ materials changing with temperature, As composition and carrier concentration is studied by Matlab software. For the determined As component, the longer Auger composite life can be obtained by optimizing working temperature and carrier concentration. When the carrier concentration is 3.2×10^{15} cm^{-3} and the temperature is 200 K, the maximum Auger lifetime of n-type $InAs_{0.35}Sb_{0.65}$ is 2.91×10^{-9} s.

Key words: InAsSb; Auger recombination lifetime; As composition; temperature; carrier concentration

0　引言

温度在0～1 000 ℃的物体都会向外辐射能量，它们对应的辐射波长在1～20 μm区间[1]。而大气层有两个传输窗口，分别对应波长3～5 μm（中波长红外，MWIR）和8～14 μm（长波红外，LWIR）区域。工作在这两个区域的红外探测器可广泛应用在空间

收稿日期：2019－10－18.

成像、天文学、军事、通信及资源勘探等方面[2-4]。HgCdTe (MCT)是目前制作中长波红外探测器相对成熟的材料,但也存在一些不足,如 Hg 具有高的蒸气压,使材料的稳定性和均匀性难以保证[5-6]。此外,其禁带宽度会随温度的升高而增加,导致工作区不在长波长范围,因此迫切需要寻找替代材料。锑基材料,包括Ⅲ-Sb 二元、三元和四元合金,是适合应用在中长红外波段的候选材料。InSb 在 300 K 时的截止波长为 7 μm,在 77 K 时截止波长为 5.5 μm[7]。如果在 InSb 中加入 As 可形成 InAsSb,它的禁带宽度会随着 As 的增加先减小后增加,且具有禁带宽度对 As 组分变化相对不敏感、光电转换效率高、较好的化学稳定性和机械强度、优良的室温工作性能、低俄歇复合率等特性[8-10]。在Ⅲ-Ⅴ族半导体合金中,InAsSb 是具有最小带隙(除含 Tl 和 Bi 的合金外)的材料,实验结果显示,在温度 200～300 K 范围内,InAsSb 可以覆盖 8～14 μm 光谱区[7]。因此,它是制作 LWIR 非致冷光子探测器的理想材料。俄歇复合寿命是影响探测器性能的重要因素之一,本文采用 Matlab 软件模拟研究了禁带宽度和本征载流子浓度随温度、As 组分的变化以及 n 型和 p 型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命随温度、As 组分及载流子浓度的变化,为设计和制备高性能 LWIR 非致冷探测器提供理论依据。

1　InAsSb 材料的物理性质

1.1　禁带宽度

$InAs_xSb_{1-x}$ 合金具有闪锌矿结构,在布里渊区中心它是直接带隙材料。根据 Wieder 和 Clawso 理论,$InAs_xSb_{1-x}$ 的能带可以用下式表示[11]:

$$E_g(x,T)=0.411-\frac{3.4\times10^{-4}T^2}{210+T}-0.876(1-x)+0.70(1-x)^2+3.4\times10^{-4}xT(1-x) \quad (1)$$

一些研究表明,通常最小带隙要比由式(1)得到的小些[12]。

1.2　电子有效质量

为了使实验中室温有效质量与计算值较好地吻合,Rogalski 和 Jozwikowski 将导带-价带混合理论应用于有效质量的计算[13],得到导带有效质量,可写成:

$$\frac{1}{m_e^*}=\frac{1}{m_{ce}^*}+\frac{\delta E}{3}\left[\frac{1}{m_{hh}^*E_{gv}}+\frac{1}{m_{lh}^*E_{gv}}+\frac{1}{m_s^*(E_{gv}+\Delta v)}-\frac{1}{m_{ce}^*}\left(\frac{2}{E_{gv}}\frac{1}{(E_{gv}+\Delta v)}\right)\right] \quad (2)$$

其中,$\delta E=E_{gv}-E_g$,E_{gv} 和 Δv 可以通过 $E_{gv}=0.351-0.176(1-x)$ 和 $\Delta v=0.39+0.42(1-x)$ 获得。m_{hh}^*,m_{lh}^* 和 m_s^* 分别为重空穴带、轻空穴带和自旋劈裂带的有效质量,m_{ce}^* 是在没有导带混合时导带的有效质量。

1.3　本征载流子浓度

根据 KP 理论可以计算 $InAs_xSb_{1-x}$ 本征载流子浓度 n_i 随温度的变化,通过拟合计算非抛物线形的 n_i 值以获得抛物线形的 n_i,表达式如下[13]:

$$n_i=[1.35+8.50(1-x)+4.22\times10^{-3}T-1.53\times10^{-3}(1-x)T-6.73(1-x)^2]\times10^{14}T^{3/2}E_g^{3/4}\exp\left(-\frac{E_g}{2kT}\right) \quad (3)$$

2　温度和材料参数对 $InAs_xSb_{1-x}$ 材料性质影响的模拟研究

2.1　禁带宽度

材料的禁带宽度决定了用其制作探测器的工作波长,而其大小受 As 组分及温度的影响,图 1 为在 200～300 K 温度范围内,$InAs_xSb_{1-x}$ 材料禁带宽度随 As 组分的变化关系。

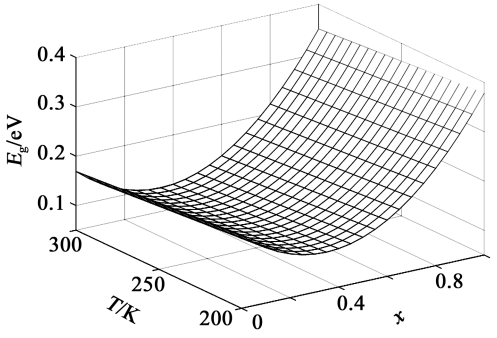


图 1　$InAs_xSb_{1-x}$ 材料禁带宽度随温度和 As 组分的变化

从图中可以看到不同温度时,禁带宽度随 As 组分的变化趋势一致;同一温度下,随 As 组分增加禁带宽度先减小然后增加。当 As 组分为 0.35 时,禁带宽度最小。相同 As 组分时,高温下禁带宽度略低于低温下的禁带宽度。

2.2　本征载流子浓度

俄歇复合寿命受本征载流子浓度的影响。图 2 为在 200～350 K 温度范围内,$InAs_xSb_{1-x}$ 材料本征载流子浓度随温度和 As 组分的变化关系。

从图中看到，确定温度下本征载流子浓度随 As 组分的增加先增加后减小，与禁带宽度的变化趋势正好相反。这是由于禁带宽度减小时，价带上的电子更容易被激发到导带上，引起本征载流子浓度的增加。$InAs_xSb_{1-x}$ 本征载流子浓度随温度变化明显，在 200 K 时本征载流子浓度最大为 7.2×10^{15} cm^{-3}，而在 350 K 时本征载流子浓度最大为 1.3×10^{17} cm^{-3}。

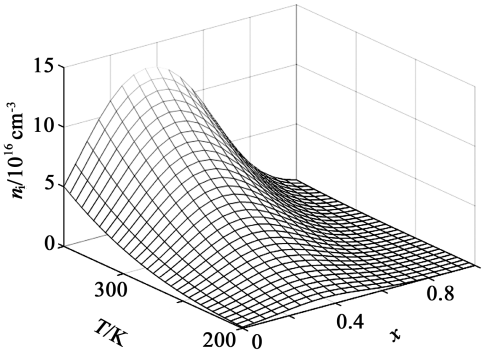


图 2 $InAs_xSb_{1-x}$ 本征载流子浓度随温度和 As 组分的变化关系

2.3 n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命

下面以载流子浓度 $n=10^{18}$ cm^{-3} 为例，模拟研究温度和材料参数对 n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命的影响。

探测器的探测率受俄歇复合寿命的影响，而俄歇复合寿命主要由温度和材料参数决定，因此研究俄歇复合寿命与温度和材料参数之间的关系十分重要。俄歇复合涉及三个载流子，根据三个载流子所处的位置不同，可以有十种俄歇复合方式[14]，但其中只有 A-1(两个电子和一个空穴参与的俄歇复合过程)、A-7(两个空穴和一个电子参与的俄歇复合过程)和 A-S(自旋轨道劈裂对应的俄歇复合过程)相对重要。因此，总俄歇复合寿命是它们三者的叠加，即：

$$\frac{1}{t_A}=\frac{1}{t_{A1}}+\frac{1}{t_{A7}}+\frac{1}{t_{AS}} \tag{4}$$

图 3 为 n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命与温度和 As 组分的关系。从图中看到对确定的温度，当 As 组分增加时，俄歇复合寿命先减小后增大，与禁带宽度随 As 组分的变化趋势一致。且高温时的俄歇复合寿命随 As 组分变化大于低温下的变化。在 300 K 时，当 As 组分为 1 时，俄歇复合寿命最大，为 5.06×10^{-12} s。

图 4 为 300 K 时，n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命与载流子浓度和 As 组分的关系。从图可以看到对确定的 As 组分，随着载流子浓度的增加，俄歇复合寿命是先增加随后减小。而当载流子浓度确定时，随着 As 组分的增加，俄歇复合寿命是先减小随后增加；当载流子浓度为 3.2×10^{14} cm^{-3}，As 组分为 1 时，寿命最大为 4.17×10^{-6} s。

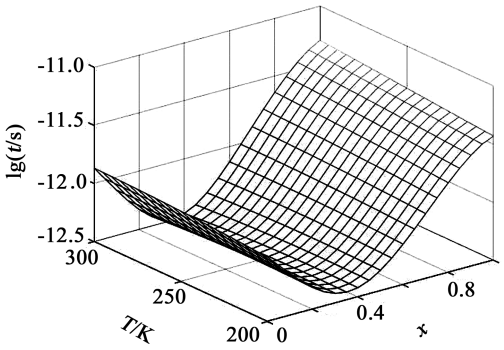


图 3 n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命与温度和 As 组分的关系

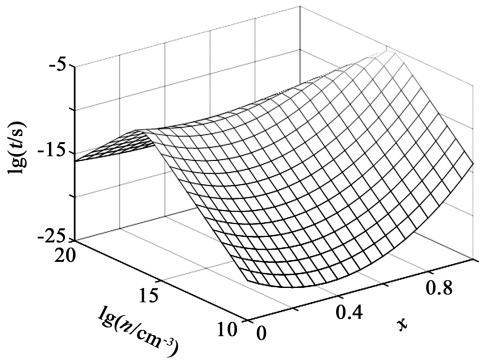


图 4 n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命与载流子浓度和 As 组分的关系

由前面获得的禁带宽度随组分变化的关系知，当 As 组分为 0.35 时，禁带宽度最小，对应的截止波长大约是 15 μm。图 5 为 n 型 $InAs_{0.35}Sb_{0.65}$ 材料俄歇复合寿命与载流子浓度和温度的关系。

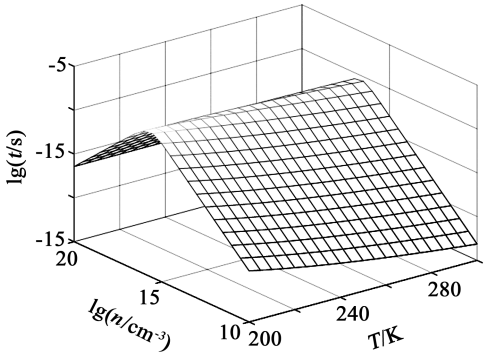


图 5 n 型 $InAs_{0.35}Sb_{0.65}$ 材料俄歇复合寿命与载流子浓度和温度的关系

从图中可见，当载流子浓度为 3.2×10^{15} cm^{-3}，温度为 200 K 时，俄歇复合寿命最大为 2.91×10^{-9} s。

2.4　p 型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命

在下面的模拟研究中，是以载流子浓度 $p=10^{18}$ cm^{-3} 为例进行的。图 6 是 p 型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命随温度和 As 组分变化的关系。

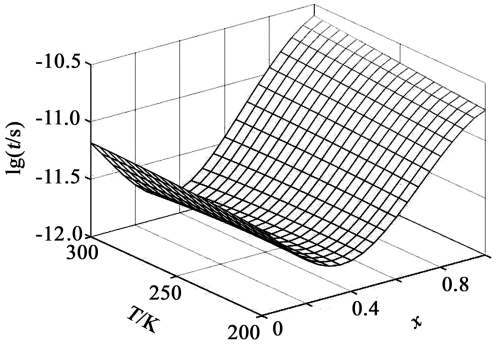


图 6　p 型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命随温度和 As 组分变化的关系

从图中可见，温度确定时，当 As 组分增加，俄歇复合寿命先减小后增大，这与禁带宽度随 As 组分的变化趋势一致，且高温时的俄歇复合寿命随 As 组分的变化高于低温下的变化。在 300 K 时，当 As 组分为 1 时，俄歇复合寿命最大为 2.51×10^{-11} s。

图 7 为 p 型 $InAs_{0.35}Sb_{0.65}$ 材料俄歇复合寿命与载流子浓度和温度的关系。

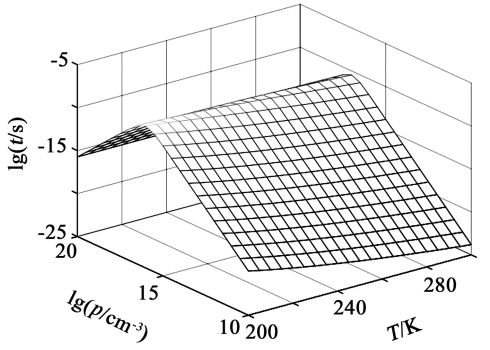


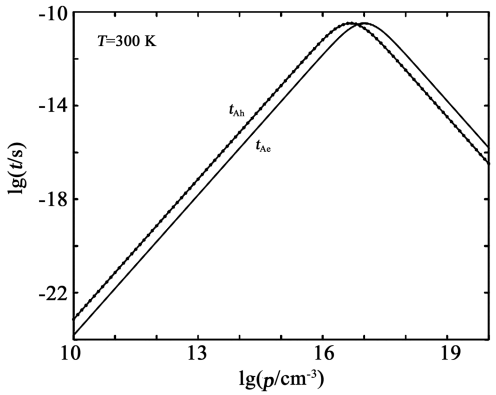
图 7　p 型 $InAs_{0.35}Sb_{0.65}$ 材料俄歇复合寿命与载流子浓度和温度的关系

从图中可见，当载流子浓度为 1.0×10^{16} cm^{-3}、温度为 200 K 时，寿命最大为 3.37×10^{-9} s。

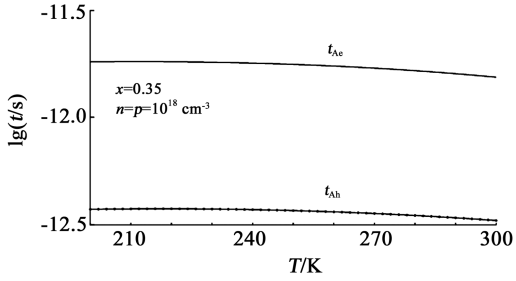
2.5　n 型和 p 型 $InAs_{0.35}Sb_{0.65}$ 材料俄歇复合寿命的比较

研究了 As 组分为 0.35 时，n 型和 p 型 $InAs_{0.35}Sb_{0.65}$ 材料的俄歇复合寿命随载流子浓度和温度的变化，图 8 为模拟结果。

t_{Ah} 表示 n 型材料中少子空穴的俄歇复合寿命，t_{Ae} 表示 p 型材料中少子电子的俄歇复合寿命。从图 8(a)看出，随着载流子浓度的增加，n 型和 p 型 $InAs_{0.35}Sb_{0.65}$ 材料的俄歇复合寿命变化趋势是一致的，在载流子浓度小于 10^{17} cm^{-3} 时，n 型材料的俄歇复合寿命大于 p 型的；而当载流子浓度大于 10^{17} cm^{-3} 时，p 型材料的俄歇复合寿命大于 n 型的。图 8(b)显示当 n 型和 p 型载流子浓度都为 10^{18} cm^{-3} 时，在 200～300 K 范围内，p 型材料的俄歇复合寿命总是大于 n 型材料的俄歇复合寿命。



(a)　随载流子浓度变化



(b)　随温度变化

图 8　n 型和 p 型 $InAs_{0.35}Sb_{0.65}$ 材料的俄歇复合寿命随载流子浓度和温度的变化曲线

3　结论

本文分别对 n 型和 p 型 $InAs_xSb_{1-x}$ 材料的俄歇复合寿命随温度、载流子浓度及 As 组分的变化进行了模拟研究。结果显示：对确定的温度，n 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命随 As 组分的变化是先减小后增加，且其随 As 组分的变化大于其随温度的变化；p 型 $InAs_xSb_{1-x}$ 材料俄歇复合寿命随温度和 As 组分的变化趋势与 n 型的类似，但其值比 n 型的大。考虑到 As 组分为 0.35 时，禁带宽度最小(对应截止波长 15 μm)，模拟给出了 n 型和 p 型 $InAs_{0.35}Sb_{0.65}$ 材料的俄歇复合寿命随载流子浓度和

温度的变化，发现当载流子浓度大于 10^{17} cm^{-3} 时，p 型材料的俄歇复合寿命大于 n 型的，而且其俄歇复合寿命主要受载流子浓度的影响。这些结果对设计高性能的 LWIR 探测器有指导意义。

参考文献：

[1] Schlessinger M. Infrared Technology Fundamentals[M]. New York: Marcel Dekker, Inc., 1995.

[2] 高玉竹，龚秀英，李继军，等. 非制冷型 InAsSb 光探测器在 8～9 μm 波长的性能提高[J]. 光电子·激光，2015，26(5)：825-828.
Gao Yuzhu, Gong Xiuying, Li Jijun, et al. Performance enhancement of uncooled InAsSb photodetectors at wavelength of 8～9 μm[J]. J. of Optoelectronics·Laser, 2015, 26(5): 825-828.

[3] 孙常鸿，胡淑红，王奇伟，等. 非致冷 InAsSb 中长波红外探测器研究评述[J]. 中国电子科学研究院学报，2010，5(1)：11-18.
Sun Changhong, Hu Shuhong, Wang Qiwei, et al. The research on uncooled InAsSb MWIR and LWIR detectors[J]. J. of China Academy of Electronics and Information Technol., 2010, 5(1): 11-18.

[4] Kopytko M, Martyniuk P, Madejczyk P, et al. High frequency response of LWIR HgCdTe photodiodes operated under zero-bias mode[J]. Optical and Quantum Electron., 2018, 50(2): 64.

[5] 郭 轶，吴广会，冯彦斌. 光导型 InAsSb 红外探测器的研制[J]. 电子设计工程，2012，20(18)：93-95.
Guo Yi, Wu Guanghui, Feng Yanbin. Fabrication of photoconductors InAsSb infrared detector[J]. Electronic Design Engin., 2012, 20(18): 93-95.

[6] Rogalski A. $InAs_{1-x}Sb_x$ infrared detectors[J]. Prog. Quant. Electr., 1989, 13(3): 191-231.

[7] Razeghi M. Overview of antimonide based Ⅲ-Ⅴ semiconductor epitaxial layers and their applications at the center for quantum devices[J]. Eur. Phys. J. of Appl. Phys., 2003, 23(3): 149-205.

[8] Suo F, Tong J C, Qian L, et al. Study of dark current in mid-infrared InAsSb-based hetero n-i-p photodiode[J]. J. of Physics D: Appl. Phys., 2018, 51: 275102.

[9] 高玉竹，赵子瑞，龚秀英. 近室温工作的中长波 InAsSb 探测器[J]. 光电子·激光，2018，29(12)：1266-1269.
Gao Yuzhu, Zhao Zirui, Gong Xiuying. InAsSb photodetectors with long wavelength operating at near room temperature[J]. J. of Optoelectronics·Laser, 2018, 29(12): 1266-1269.

[10] 邓功荣，赵 鹏，袁 俊，等. 锑基高工作温度红外探测器研究进展[J]. 红外技术，2017，39(9)：780-784.
Deng Gongrong, Zhao Peng, Yuan Jun, et al. Status of Sb-based HOT infrared detectors[J]. Infrared Technol., 2017, 39(9): 780-784.

[11] Wieder H H, Clawson A R. Photo-electronic properties of $InAs_{0.07}Sb_{0.93}$ films[J]. Thin Solid Films, 1973, 15(2): 217-221.

[12] Bethea C G, Levine B F, Yen M Y. Photoconductance measurements on $InAs_{0.22}Sb_{0.78}$/GaAs grown using molecular beam epitaxy[J]. Appl. Phys. Lett., 1988, 53(4): 291-292.

[13] Rogalski A, Jozwikowski K. Intrinsic carrier concentration and effective masses in $InAs_{1-x}Sb_x$[J]. Infrared Phys., 1989, 29(1): 35-42.

[14] Beattie A. Quantum efficiency in InSb[J]. J. Phys. Chem. Solids., 1962, 23(8): 1049-1056.

作者简介：

张景波(1982－)，男，黑龙江黑河人，硕士，主要研究方向为计算机仿真数据处理；

殷景志(1960－)，女，博士，教授，主要研究方向包括半导体光电材料的生长，以及光电器件的设计和制作。

E-mail：yjz886666@163.com

(上接第 79 页)

[9] Thompson S, Anand N, Armstrong M, et al. A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of Cu interconnects, Low k ILD and 1/spl mu/m/sup 2/SRAM cell [C]// Digest. Inter. Electron Devices Meeting, 2002.

[10] Sun X, Liu J, Kimerling L C, et al. Toward a germanium laser for integrated silicon photonics[J]. IEEE J. of Sel. Topics in Quantum Electronics, 2010, 16(1): 124-131.

[11] Hom-On C, Horprathum M, Eiamchai P, et al. Influence of substrate temperature on deposition rate and optical properties of aluminum oxide thin films prepared by reactive DC sputtering technique[J]. Key Engin. Materials, 2016, 675/676: 281-284.

[12] Shao S, Fan Z, Shao J, et al. Evolutions of residual stress and microstructure in ZrO_2 thin films deposited at different temperatures and rates[J]. Thin Solid Films, 2003, 445(1): 59-62.

[13] 江剑平. 半导体激光器[M]. 北京：电子工业出版社，2000：87-90.
Jiang Jianping. Semiconductor Lasers[M]. Beijing: Publishing House of Electronics Industry, 2000: 87-90.

作者简介：

崔碧峰(1972－)，女，博士，副教授，主要从事半导体激光器设计与制备方面的研究。

E-mail：cbf@bjut.edu.cn

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.018

高强度光纤光栅与碳纤维的复合及应变传感特性研究

侯智懿[1]，李　彦[1]，邵洪峰[1]，朱元林[2]

（1. 北京航空航天大学 仪器科学与光电工程学院，北京 100191；2. 江苏法尔胜研发中心有限公司，江苏 江阴 214433）

摘　要： 针对光纤光栅工程化应用性能不佳的问题，提出将高强度光纤光栅与碳纤维材料复合，介绍了复合样品的结构形式，分析了应变传递过程，然后对其应变传感特性和疲劳性能进行了实验研究。结果表明：复合样品能够达到较高的应变传递效率，具有良好的应变传感特性；疲劳实验中复合样品保持正常工作，波长变化曲线的幅度稳定；疲劳实验前后复合样品的应变传感特性基本不变，应变灵敏度系数分别为1.12和1.17 pm/με，应变响应曲线的线性度均高于0.999，重复性分别为0.623%和0.397%。

关键词： 光纤布拉格光栅；碳纤维；应变传感；应变传递；疲劳性能

中图分类号： TN253　**文章编号：** 1001－5868(2020)01－0085－04

Study on Composition of High-strength FBG with Carbon Fiber and Their Strain Sensing Characteristics

HOU Zhiyi[1], LI Yan[1], SHAO Hongfeng[1], ZHU Yuanlin[2]

(1. School of Instrumentation and Optoelectron. Engin., Beihang University, Beijing 100191, CHN;
2. Jiangsu Fasten Research and Development Center Co. Ltd., Jiangyin 214433, CHN)

Abstract: Combination of high-strength FBG and carbon fiber are proposed aiming at the poor performance of fiber Bragg grating (FBG) in engineering applications. The structure of the sample is introduced and the strain transmission process of sample is analyzed. Experiments were conducted to verify the strain sensing characteristic and fatigue property. Experimental results show that the composited samples can achieve high strain transmission efficiency and obtain good strain sensing characteristics. During the fatigue experiment, the samples work properly and the amplitude of wavelength change was stable. The strain sensing characteristics of samples remain essentially unchanged after the fatigue experiment, and the strain sensitivity coefficient is 1.12 and 1.17 pm/με, respectively. The linearity is higher than 0.999, and the repeatability are 0.623% and 0.397%, respectively.

Key words: fiber Bragg grating; carbon fiber; strain sensing; strain transfer; fatigue property

0　引言

随着光导纤维的实用化和光纤通信技术的迅速发展，光纤传感技术迅速发展并成为结构健康监测领域的重要研究方向[1-2]。光纤光栅(Fiber Bragg Grating，FBG)因具有波长解调、不受光功率波动影响、灵敏度高等优点而备受关注，质量轻且结构简单和易于埋入结构体内部的特点使其被认为是航空航天、大型土木工程结构健康监测中最具有前景的传感器件之一[3-4]。光纤光栅所依托的普通单模光纤本身物理结构非常脆弱，在实际应用时需要对裸光纤光栅进行封装[5]。但目前常用的封装方式如金属材质基片式封装、高聚物片状材料封装等往往存在力学耦合不佳导致应变传递损耗较大、耐久性和稳定性差等缺陷[6]。

收稿日期：2019－10－24.

近年来，复合材料广泛应用于航空航天、汽车交通、建筑桥梁等工程领域。复合材料中以纤维增强复合材料(Fiber Reinforced Polymer，FRP)应用最广[7]，其中碳纤维增强环氧树脂基复合材料(Carbon FRP，CFRP)因其优异的性能，在某些领域展现出逐渐取代玻璃纤维增强复合材料、铝合金和钢材的趋势[8]。复合材料不仅具有高强度、高模量、抗腐蚀、可设计性强等优点，随着“智能材料”概念的提出[9]，多功能化和智能化逐渐成为复合材料新的发展方向，这也与结构健康监测领域中光纤光栅传感器的发展需求相吻。

将光纤光栅与碳纤维复合，既能发挥复合材料的优异性能，更好地保护光纤，又能对复合材料内部的受力、应变状态进行长期监测，这在现阶段的研究中已经取得了初步的进展[8,10-12]，但在实际应用中仍存在由于被树脂包裹后容易发生脆断、成活率不高、极限应变量不高、长期稳定性较差等问题。为此，本文提出将高强度光纤光栅与碳纤维材料复合，采用单向拉挤式工艺制作复合样品，并通过静态标定实验和疲劳实验对样品进行应变传感性能测试与分析，为光纤光栅在结构健康监测传感应用的进一步发展提供参考。

1 理论分析

1.1 高强度光纤光栅传感原理

本文采用的高强度光纤光栅是一种直写型光纤光栅，在光栅刻写过程中没有涂层剥除和再涂覆的过程，避免纤芯暴露带来的机械损伤，保持了光纤原始的高强度，且光纤的外涂层材料为热塑性材料，能在 180～250 ℃温度下与复合材料协同热成型，满足界面相容原则[13]，从而更有利于复合。高强度光纤光栅的传感原理与普通光栅相同，根据耦合模理论，光纤光栅能够反射指定波长的光，光纤布拉格光栅方程为[14]

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

式中，λ_B 是布拉格波长，n_{eff} 是纤芯有效折射率，Λ 是光栅周期。温度和应力变化都会影响光栅周期，从而导致光纤光栅中心波长的变化，当环境温度保持不变，仅受外力作用产生轴向应变时，中心波长变化量 $\Delta\lambda_B$ 与应变量 ε 的关系可表示为

$$\Delta\lambda_B = \lambda_B(1-P_e)\varepsilon = K_\varepsilon\varepsilon \tag{2}$$

式中，$\Delta\lambda_B$ 为波长变化量，P_e 为光纤材料的弹光系数，K_ε 为应变灵敏度系数，对于普通石英光纤一般约为 1.2 pm/$\mu\varepsilon$。

1.2 应变传递原理与仿真

本文设计的光纤光栅与碳纤维材料复合样品为同心单向圆柱体结构，光纤光栅与碳纤维复合后，外界作用所产生的轴向应变会通过碳纤维复合材料等多层中间介质传递给内部的光纤光栅，实际应变量和光纤光栅应变量之间会存在一定差异，因此存在应变传递效率系数 k 使 $\varepsilon_{光栅} = k\varepsilon$，图 1 为光纤光栅和碳纤维复合后的模型位移关系图和应变传递示意图。

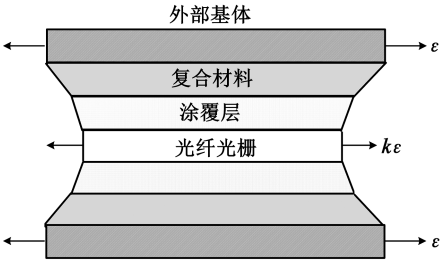


图 1 应变传递示意图

假设介质均为线弹性材料且各层材料之间紧密结合[14]，当仅有基体材料两端均匀受力产生应变，不考虑温度效应的影响，参考文献[15]的分析过程，根据剪滞理论可以建立光纤光栅和碳纤维复合体中裸光纤光栅的剪应力平衡关系式：

$$[\sigma_g(x)+\mathrm{d}\sigma_g(x)]\pi r_g^2 - \sigma_g(x)\pi r_g^2 + \tau_{gc}(r_g,x)2\pi r_g \mathrm{d}x = 0 \tag{3}$$

式中，x 为光纤轴向坐标，r_g 为光纤外径，σ_g 为光纤光栅受到的轴向应力，$\tau_{gc}(r_g,x)$ 为涂覆层与光纤层之间产生的剪切应力。结合材料力学理论公式，可以整理得到光纤光栅区域的应变传递效率系数关系式：

$$k = 1 - \frac{\sinh(\alpha L)}{\alpha L\cosh(\alpha L)} \tag{4}$$

式中，L 为光纤的总长度，α 的表达式为

$$\alpha = \sqrt{\frac{2}{E_g r_g^2\left(\frac{1}{G_c}\ln\frac{r_{cf}}{r_g}+\frac{1}{G_f}\ln\frac{r_m}{r_{cf}}\right)+\frac{E_f}{G_c}\left(\frac{r_m^2-r_{cf}^2}{2}+r_{cf}^2\ln\frac{r_{cf}}{r_m}\right)}} \tag{5}$$

式中，r_g，r_{cf}，r_m 分别为光纤、涂覆层和复合材料的外径，E_g，E_f 分别为光纤和碳纤维复合材料的弹性模量，G_c，G_f 分别为涂覆层和碳纤维复合材料的剪切模量。经过应变传递效率系数修正后，式(2)可以改写为

$$\Delta\lambda_B = kK_\varepsilon\varepsilon = K'_\varepsilon\varepsilon \tag{6}$$

式中，K'_ε 为复合样品的应变灵敏度系数，与复合样品所选取的光纤光栅的中心波长、光纤材料和复合材料的参数有关。

由式(4)可知，当复合样品所选用的光纤光栅和碳纤维材料确定后，应变传递效率系数由光纤长度决定，图 2 为两者关系的仿真曲线图，仿真参数如表 1 所示。

表 1　仿真参数

参数	数值	参数	数值
r_g/mm	0.062 5	r_m/mm	6.7
E_g/Pa	7.22×10^{10}	G_f/Pa	6.73×10^{10}
r_{cf}/mm	0.09	E_f/Pa	1.75×10^{11}
G_c/Pa	1.1×10^{8}	$2L$/mm	10～1 000

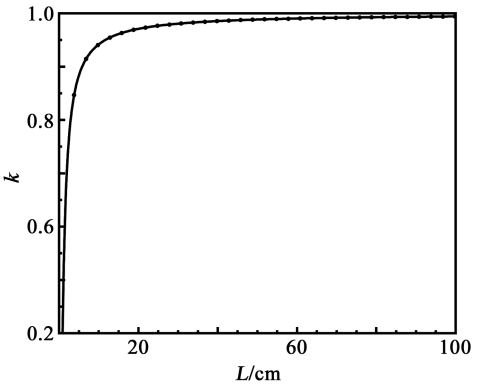


图 2　应变传递效率系数随光纤长度变化曲线

从图 2 分析可知：当复合光纤长度较短时，应变传递效率系数受光纤长度影响较大，此时应变传递效率系数较小，应变传递损失较大；当复合光纤长度大于 6 cm 时，应变传递效率系数高于 0.9，并随光纤长度的增加逐渐稳定且趋近于 1；当光纤长度大于 58.4 cm 时，应变传递效率系数高于 0.99，但过长的光纤会增加复合的难度，本文取 60 cm 作为复合样品长度。

2　实验

2.1　样品制作

对于结构为同心单向圆柱体的光纤光栅与碳纤维材料复合样品，裸光纤光栅与连续碳纤维丝的方向要保持一致，光纤位于碳纤维丝束的中心位置。复合样品的制作基体为环氧树脂胶，制作工艺为单向拉挤成型工艺，引出端设置有保护套管的裸光纤光栅与碳纤维丝束平行排布，在牵引装置的牵引下经过浸胶、分束后进入模具经加热成型，图 3 为复合样品结构示意图。复合样品所选用的裸光纤光栅中心波长为 1 536.384 nm，碳纤维复合材料弹性模量为 175 GPa，样品长度为 60 cm，直径为 6.7 mm，应变传递效率系数理论值为 0.990，在制作完成前后均能测出稳定、可靠的光纤光栅波长信号。

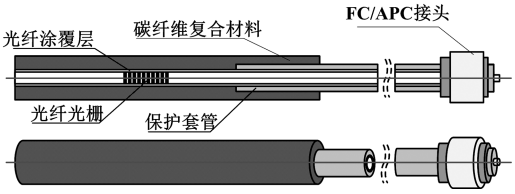


图 3　复合样品结构示意图

2.2　应变传感实验

应变传感实验的实验装置由复合样品、解调模块、上位机软件、拉伸测试装置和引伸计组成，所采用的测试装置为 SHT 微机控制电液伺服万能试验机，复合样品两端锚固后装夹在测试装置上并与解调模块和上位机连接，引伸计夹持于复合样品外部以记录实际应变量，实验环境为恒温室内。实验采用逐级加载的方法，在 0～60 kN 的范围内进行三次拉伸实验，每级加载拉力为 10 kN，读取记录各级加载后的应变值并保持 10 s 取波长示数的平均值，得到中心波长随应变量变化的实验数据，三次拉伸均为相同实验条件。

2.3　疲劳实验

疲劳实验的实验装置除不包含引伸计外均与应变传感实验相同，所采用的测试装置是 MTS 电液伺服疲劳动静试验机，疲劳测试拉伸幅值为 20 kN，拉伸频率为 16 Hz，加载方式为等幅正弦波加载，实验次数为 300 万次。实验结束后以相同实验条件再对复合样品进行应变传感实验，测定性能参数。

3　结果与讨论

3.1　应变传感实验结果与分析

图 4 为复合样品中心波长与应变量的关系曲线。实验结果表明：复合样品在三次拉伸过程中均表现出良好的重复性和稳定性，三次标定曲线近乎重合，波长重复性最大相差为 69.4 pm，相对重复性误差为 0.623%；波长与应变量关系曲线具有良好的线性度，三次实验的线性度均达到了 0.999 以上；复合样品的应变灵敏度系数为 1.12 pm/με，应变传递系数为 0.933，与仿真结果较为接近，由于实际制作无法完全达到理论仿真的假设条件，实际测量值

略低于仿真值。

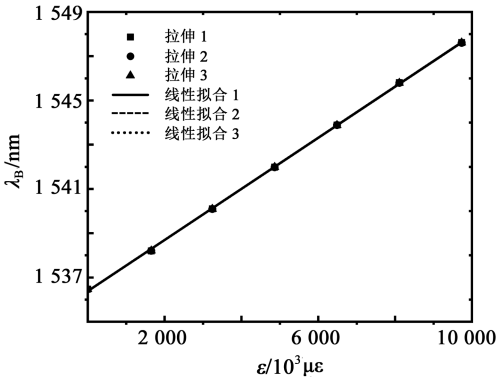


图 4 波长与应变量关系曲线

3.2 疲劳实验结果与分析

图 5 记录了疲劳实验过程中的四组复合样品中心波长随时间变化曲线,在循环载荷作用下,复合样品的中心波长变化幅度基本保持一致,无明显的异常漂移值,具有良好的耐久性和稳定性。图 6 为疲劳实验后复合样品的应变传感关系曲线,实验结果表明:复合样品在疲劳实验后仍然保持良好的线性度和重复性,线性度为 0.999 以上,三次标定曲线近乎重合,波长重复性最大相差为 45.3 pm,相对重复性误差为 0.397%;应变灵敏度系数为 1.17 pm/με,应变传递效率系数为0.975,与疲劳实验前相比略有提高,分析原因可能是随着拉伸次数的增加,自身的残余应变得到释放,光纤光栅与复合材料协同变形能力得到提升,从而提高应变传递效率。

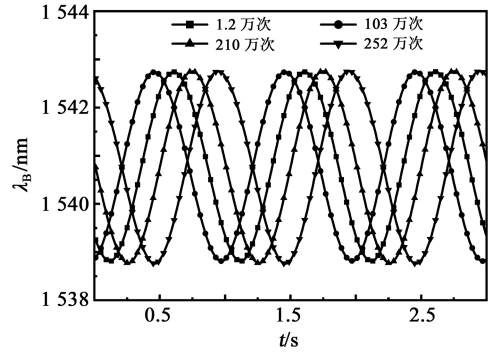


图 5 疲劳实验波长变化曲线

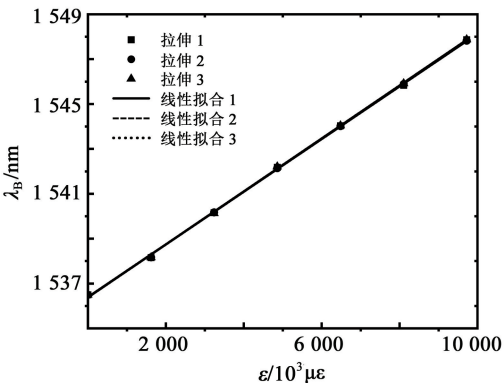


图 6 疲劳实验后应变传感关系曲线

4 结论

本文针对现有的光纤光栅封装方式不适合应用于结构健康监测领域的问题,提出将高强度光纤光栅与碳纤维材料复合,分析了应变传递过程,并对复合样品的应变传感特性和疲劳特性进行了实验研究。结果表明:高强度光纤光栅在不引入保护装置的情况下能够与碳纤维材料很好地复合,二者在外力作用下协同变形,且达到较高的应变传递效率;复合样品的应变传感特性曲线具有较高的线性度和重复性,并且在疲劳实验中表现出良好的耐久性和稳定性;复合样品结构简单、应变测量范围大,具有良好的工程应用前景。研究结果为结构健康监测领域光纤光栅传感应用的进一步发展提供了参考。

参考文献:

[1] Oliveira R D, Ramos C A, Marques A T. Health monitoring of composite structures by embedded FBG and interferometric Fabry-Perot sensors[J]. Computers & Structures, 2008, 86(3): 340-346.

[2] Lu S, Jiang M, Sui Q, et al. Damage identification system of CFRP using fiber Bragg grating sensors [J]. Composite Structures, 2015, 125: 400-406.

[3] Mauron P, Nellen P M, Sennhauser U J. High-strength prooftest parameters for fiber Bragg grating sensors [C]// Information Technologies 2000 Inter. Society for Optics and Photonics, 2001.

[4] 赵海涛, 张博明, 武湛君, 等. 光纤光栅智能复合材料基础问题研究[J]. 传感器与微系统, 2007, 26(12): 27-30.
Zhao Haitao, Zhang Boming, Wu Zhanjun, et al. Infrastructure study for optical fiber grating in smart composite materials [J]. Transducer and Microsystem Technologies, 2007, 26(12): 27-30.

[5] 刘春桐, 李洪才, 张志利, 等. 铝合金箔片封装光纤光栅传感特性研究[J]. 光电子·激光, 2008, 19(7): 905-908.
Liu Chuntong, Li Hongcai , Zhang Zhili, et al. Study on the sensing characteristics of fiber Bragg grating packaged by aluminum alloy slice[J]. J. of Optoelectron. · Laser, 2008, 19(7): 905-908.

[6] 李宏男, 李爱群, 滕 军, 等. 结构健康监测系统设计标准[S]. 北京: 中国建筑工业出版社, 2012.
Li Hongnan, Li Aiqun, Teng Jun, et al. Design standard of structural health monitoring system [S]. Beijing: China Building Industry Press, 2012.

(下转第 98 页)

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.019

InAs量子点低温盖层对其发光特性的影响

吴　唯，周　勇，刘尚军，刘万清，张　靖

（重庆光电技术研究所，重庆 400060）

摘　要： 采用金属有机物化学气相淀积（MOCVD）技术生长了InAs量子点及不同组分和厚度的量子点低温盖层，采用光致荧光光谱（PL）和时间分辨荧光发射谱（TRPL）研究了量子点在不同低温盖层下的荧光发光性质。对比了常规方法和速率调制外延（FME）法生长GaAs低温盖层的量子点质量，结果表明，FME法生长低温盖层的量子点荧光发射谱光强更强且发光寿命达到0.6 ns，明显优于普通方法生长低温盖层的量子点发光寿命。

关键词： InAs量子点；量子点低温盖层；光致荧光谱；时间分辨谱；速率调制外延

中图分类号： TN248　　**文章编号：** 1001－5868(2020)01－0089－04

Photoluminescence Characteristics of InAs Quantum Dots with Low Temperature Cap Layers

WU Wei, ZHOU Yong, LIU Shangjun, LIU Wanqing, ZHANG Jing

(Chongqing Optoelectronics Research Institute, Chongqing 400060, CHN)

Abstract: Metalorganic chemical vapour deposition (MOCVD) was employed to grow InAs quantum dots (QDs) and low temperature QDs caps with different composition and thickness, and photoluminescence (PL) and time-resolved photoluminescence (TRPL) was used to study the characteristics of QDs with different cap layers. Flow-rate modulation epitaxy (FME) and conventional methods were utilized to grow low temperature cap of InAs QDs. Experimental results indicate that the PL intensity of QWs is stronger when the cap was grown by FME, and its TRPL life time reaches 0.6 ns, which is much longer than the samples prepared by the conventional method.

Key words: InAs QDs; quantum gap layer; photoluminescence; time-resolved photoluminescence; flow-rate modulation epitaxy

0　引言

量子点（QD）三维尺寸与电子德布罗意波长相当，载流子在三维空间运动受到限制，能带结构分立且其态密度呈δ函数分布。改变量子点结构尺寸可以便捷地调节量子点的禁带宽度，具有明显的尺寸效应。因此，基于量子点结构的器件相对于量子阱或体材料器件具有更加优异的性能，使量子点得到了十分广泛的应用。其中InAs/GaAs量子点已广泛应用于半导体激光器[1]、单光子源[2]、探测器[3]、太阳电池[4]、显示材料[5-6]等方面。

量子点盖层对量子点自组织行为、In原子的解吸附以及结合岛的消融有重要影响，因此生长合适量子点盖层对InAs量子点发光特性至关重要。为改善量子点发光特性，人们对InAs量子点盖层进行了广泛研究。采用不同材料盖层可以将量子点发光波长扩展到1.3和1.5 μm[7-9]。孔令民等人研究了InGaAs/InAlAs联合盖层，发现联合盖层可以提高载流子的限制势垒，进而提高发光效率[10]。田芃等人研究了不同双盖层对量子点光学性质的影响[11]。目前，鲜有文章报道单层量子点低温盖层不同组分、厚度和生长方法对量子点发光性质的研究。不同组分和不同厚度的InGaAs低温盖层影响着量子点的自组织行为和量子点的晶体质量，对生长优质InAs量子点起着重要作用。另外，低温盖层的

收稿日期：2019－07－11.

生长方法决定了量子点表面粗糙度和量子点晶体质量,选取优良的生长方法十分重要。本文研究了InAs量子点在不同组分和不同厚度低温盖层下的荧光特性,以及InAs量子点在速率调制外延(Flow-Rate Modulation Epitaxy/FME)生长法和普通生长法的低温盖层下的光致荧光光谱和时间分辨荧光发射谱特性。

1 实验

量子点结构采用AIXTRON公司AIX2400型MOCVD设备生长,金属源为三甲基铟(TMIn)、三甲基镓(TMGa)及三甲基铝(TMAl),气态源为砷烷(AsH_3)和磷烷(PH_3),反应室压力为10^4 Pa。衬底为(001)晶向的n型GaAs。图1是量子点结构示意图。生长前衬底在AsH_3气氛中730 ℃的条件下退火10 min,以消除衬底表面氧化层和污染物。然后,在650 ℃下沉积一层300 nm厚的GaAs缓冲层,目的是使表面平整并阻止残留污染物从衬底向有源区扩散。接着在缓冲层上以同样的温度生长一层50 nm厚的$Al_{0.6}Ga_{0.4}As$和100 nm厚的GaAs。然后在GaAs上沉积InAs量子点和量子点盖层,沉积温度为500 ℃。量子点生长完成以后,同时中断In源和As源20 s,让量子点进行熟化。接着在量子点上生长量子点低温盖层。量子点和盖层沉积完成以后,温度升高至600 ℃再生长剩余结构。剩余结构为50 nm的GaAs和50 nm的$Al_{0.5}Ga_{0.5}As$。$Al_{0.5}Ga_{0.5}As$势垒可以把载流子限制在量子点附近,同时阻止载流子迁移到表面非辐射复合。最后在$Al_{0.5}Ga_{0.5}As$上覆盖一层薄的GaAs保护层,以防止$Al_{0.5}Ga_{0.5}As$被氧化。

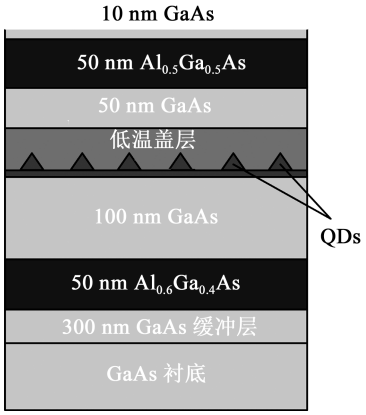


图1 InAs量子点外延结构示意图

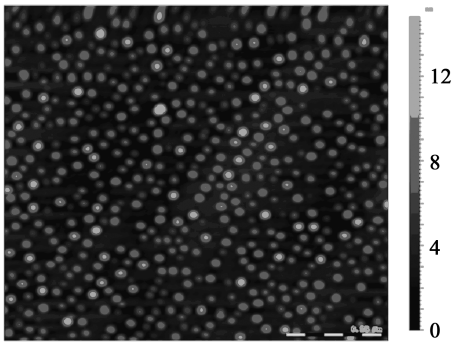
低温盖层(Low Temperature Cap Layer,LTCL)组分为$In_xGa_{1-x}As$($x=0\sim0.07$),厚度为3～5 nm。分别采用常规外延和速率调制外延(FME)生长了GaAs盖层。FME是MOCVD外延方法的一种[12],有机金属源和氢化物源反复交替进入反应室。采用两路氢化物源进行材料生长,其中一路保持微弱的氢化物源流量,另一路用于开关时序。FME法在有机金属源进入反应室时氢化物分压保持得很低,金属原子在表面可以快速迁移至稳定的晶格格点,有利于在低温下生长晶体质量好且界面陡峭的Ⅲ-Ⅴ族材料。常规方法则是有机金属源和氢化物源同时进入反应室生长。

量子点外延材料生长完成以后,在室温下测量其PL谱和TRPL谱。PL谱采用HORIBA Jobin公司的光致发光光谱仪(LabRAM Spectrometer HR 800UV),激发波长为532 nm。TRPL谱采用Edinburgh公司的FL920荧光寿命光谱仪进行测量。

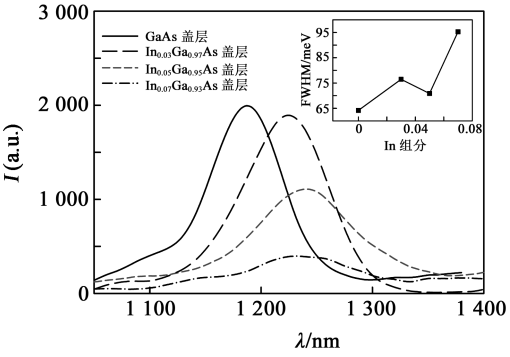
2 结果与讨论

不同组分的InAs量子点低温盖层由于晶格尺寸的差异,会影响覆盖材料和量子点之间的应变,影响量子点的自组织行为和量子点尺寸,从而影响其材料特性。在本文生长条件下,量子点密度均为5.8×10^{10} cm^{-2},如图2(a)所示。实验发现,采用InGaAs层作为InAs量子点低温盖层能有效提高量子点PL波长。图2(b)为InAs量子点在厚度为3 nm的$In_xGa_{1-x}As$盖层的室温PL谱。量子点盖层In组分从0变化到7%,PL光谱峰发生明显红移现象,峰值从1 190 nm(GaAs)红移至1 224 nm($In_{0.05}Ga_{0.95}As$),光谱强度逐渐降低,当In组分超过7%时,荧光光谱非常微弱,几乎无光,说明量子点材料质量非常差,也表明过多In的InGaAs盖层不利于生长优质InAs量子点。从图2(b)中插图可以看到光谱半高宽(FWHM)的变化,$In_{0.05}Ga_{0.95}As$盖层半高宽相对较小,此时引入的应变适中,量子点质量较好,使得低温盖层对量子点的形貌和尺寸影响较小,因此得到的量子点尺寸更均匀,故其半峰宽更窄,说明低温盖层中适当的In可以优化InAs量子点尺寸的均匀性。同时我们还研究了不同厚度$In_{0.05}Ga_{0.95}As$盖层InAs量子点的荧光特性。图3展示了室温下不同$In_{0.05}Ga_{0.95}As$盖层厚度的光荧光谱,可见盖层厚度越厚,引入的应力就越大,利于得到更长的波长,最长可达1 273 nm。但是过厚的

盖层导致晶体质量下降，增加了非辐射复合，从而使 PL 荧光强度明显减弱。



(a)　量子点密度图，测量区域面积为 1 μm×1 μm



(b)　不同组分 InGaAs 低温盖层的 InAs 量子点室温 PL 谱

图 2　量子点密度图和不同组分 InGaAs 低温盖层的 InAs 量子点室温 PL 谱

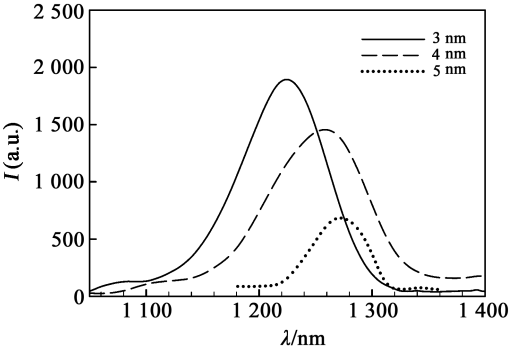


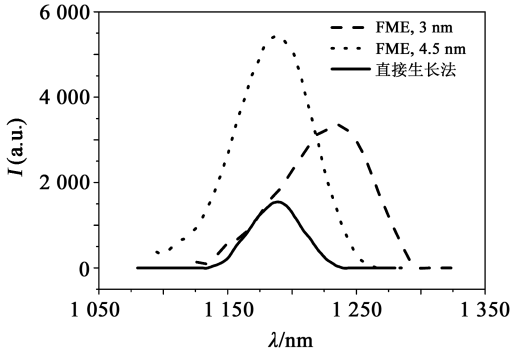
图 3　不同厚度 InGaAs 低温盖层的 InAs 量子点室温 PL 谱

图 4 展示了常规方法和 FME 方法生长低温 GaAs 盖层的 InAs 量子点的光致荧光谱(图 4(a))和时间荧光谱衰减曲线(图 4(b))。时间荧光谱中荧光衰减呈指数衰减，根据式(1)可以拟合出各自的荧光寿命。

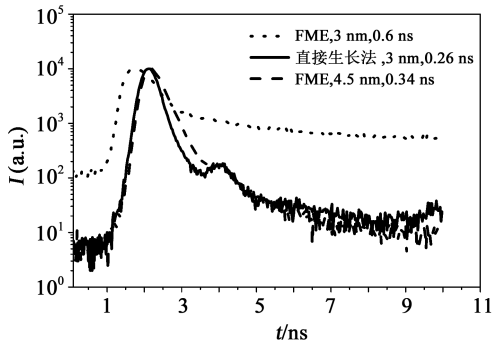
$$I(t)=I_0 e^{-\frac{t}{\tau}} \tag{1}$$

式中，$I(t)$是 t 时刻的发光强度，I_0 是最大光强值，τ 表示载流子复合寿命。

图 4(a)显示了量子点在 FME 和普通方法在 500 ℃下生长低温盖层的 PL 谱，黑色圆点谱是 FME 生长法，厚度为 3 nm。断线谱盖层是 FME 生长，厚度为 4.5 nm。实线谱盖层是普通的直接生长法，厚度 3 nm。同样是 3 nm 厚的盖层，FME 生长法 PL 谱强度明显高于传统方法，说明在低温下 FME 生长法可以得到的量子点发光效率更高，缺陷少，有效减少了非辐射复合。用 FME 法在 500 ℃低温下生长盖层，当盖层变厚时，PL 谱发生了明显的红移现象，这和图 3 中的结论一致。



(a)　量子点在两种低温盖层生长法下的 PL 谱图



(b)　量子点在两种低温盖层生长法下的时间荧光谱衰减曲线

图 4　量子点在两种不同方法生长的低温 GaAs 盖层的光致荧光谱和时间荧光谱衰减曲线

和图 4(a)相同，图 4(b)中黑色圆点谱盖层是 FME 生长方式，厚度为 3 nm，其拟合荧光寿命为 0.6 ns。断线谱盖层是 FME 生长，厚度为 4.5 nm，其拟合荧光寿命为 0.34 ns。实线谱盖层是直接生长法，厚度为 3 nm，其拟合荧光寿命为 0.26 ns。从图 4(b)中可以发现，FEM 在低温下生长 3 nm 盖层的量子点迅速被激发，且荧光衰退比较缓慢，几乎和另外两种荧光光谱同时衰减，且衰减以后其光强仍然较强。显然 FME 生长盖层的 InAs 量子发光寿命显著优于传统生长方式，这说明在低温下 FEM 生长的盖层使量子点构成一个联系紧密的耦合系统，材料之间的应力得到了充分弛豫，缺陷减少，从而减少了发热和非复合辐射过程，使得载流子的有

效供给加大。同时也说明在低温下 FEM 生长的盖层使量子点尺寸均匀,载流子在不同的量子点间的迁移效应减弱,从而使样品发光具有更好的稳定性和更长的荧光寿命。FME 生长法低温层 4.5 nm 和直接生长法 3 nm 时在 4 ns 附近出现一个小峰,这可能是由于这两种盖层衰减时间较短,材料对荧光进行吸收并再激发,这个过程和衰减过程重叠所致,该过程较弱,不影响对荧光寿命的测量。在低温下,FME 生长不同厚度盖层时发现,过厚的盖层会降低量子点的晶体质量,从而减少量子点的发光寿命和 PL 谱强度,这与图 3 中研究结论一致。

3 结论

本文研究了 InAs 量子点在不同低温盖层下的荧光光谱特性。相同厚度 InGaAs 低温盖层,随着 In 组分从 0 增加到 7%,量子点 PL 谱峰值波长从 1 190 nm 红移至 1 224 nm。当 In 含量过大时,荧光光强非常微弱,说明 In 组分过高时盖层严重影响了 InAs 量子点材料质量。InGaAs 组分一定时,其发光峰随着低温盖层厚度从 3 nm 增加至 5 nm,发光波长也出现了 51 nm 的红移现象,厚度大于一定值时,其荧光强度明显减弱,说明过厚盖层不利于生长优质 InAs 量子点材料。本文还研究了量子点在速率调制外延和常规外延法生长法下的低温盖层的 PL 谱和 TRPL 谱,发现速率调制外延(FME)生长盖层的量子点荧光发射谱强度更强且发光寿命更长,明显优于普通方法生长的盖层,提供了一个生长 InAs 量子点低温盖层的方法。

参考文献:

[1] Ledentsov N N. Quantum dot laser[J]. Semicond. Sci. Tech., 2011, 26(1): 014001.

[2] Yuan Z, Kardynal B E, Stevenson R M, et al. Electrically driven single-photon source[J]. Science, 2002, 295(5552): 102.

[3] Kim E, Madhukar A, Ye Z, et al. High detectivity InAs quantum dot infrared photodetectors[J]. Appl. Phys. Lett., 2004, 84: 3277.

[4] Luque A, Marti A, Stanley C. Understanding intermediate-band solar cells[J]. Nat. Photonics, 2012, 6: 146-152.

[5] Phillips J, Kamath K, Brock T, et al. Characteristics of InAs/AlGaAs self-organized quantum dot modulation doped field effect transistors[J]. Appl. Phys. Lett., 1998, 72(26): 3509-3511.

[6] Horiguchi N, Futatsugi T, Nakata Y, et al. Electron transport properties through InAs self-assembled quantum dots in modulation doped structures[J]. Appl. Phys. Lett., 1997, 70(17): 2294-2296.

[7] Liu H Y, Sellers I R, Badcock T J, et al. Improved performance of 1.3 μm multilayer InAs quantum-dot lasers using a high-growth-temperature GaAs spacer layer[J]. Appl. Phys. Lett., 2004, 85: 704.

[8] Laghumavarapu R B, El-Emawy M, Nuntawong N, et al. Improved device performance of InAs/GaAs quantum dot solar cells with GaP strain compensation layers[J]. Appl. Phys. Lett., 2007, 91: 243115.

[9] Liu H Y, Steer M J, Badcock T J, et al. Long-wavelength light emission and lasing from InAs/GaAs quantum dots by a GaAsAb strain-reducing laser[J]. Appl. Phys. Lett., 2005, 86: 143108.

[10] 孔令民, 姚建明, 吴正云. InGaAs/InAlAs 盖帽层对 InAs 自组装量子点发光性质的影响[J]. 半导体光电, 2008, 29(3): 349-352.
Kong Lingmin, Yao Jianming, Wu Zhengyun. Photoluminescence characteristics of InAs self-assembled quantum dots capped with InGaAs and InAlAs layers[J]. Semiconductor Optoelectronics, 2008, 29(3): 349-352.

[11] 田 芃, 黄黎蓉, 费淑萍, 等. 不同盖层对 InAs/GaAs 量子点结构和光学性质的影响[J]. 物理学报, 2010, 59(8): 5738-5742.
Tian Peng, Huang Lirong, Fei Shuping, et al. Effect of different cap layers on the structure and optical properties of InAs/GaAs self-assembled quantum dot[J]. Acta Physica Sinica, 2010, 59(8): 5738-5742.

[12] Makimoto T, Kobayashi N, Horikoshi Y. Modulation doped n-AlGaAs/GaAs heterostructures grown by flow-rate modulation epitaxy[J]. Jap. J. of Appl. Phys., 1986, 25(6): 513-515.

作者简介:

吴 唯(1990—),男,2017 年毕业于武汉大学材料物理与化学专业,获硕士学位,现为工程师,主要研究方向为 MOCVD Ⅲ-Ⅴ外延材料生长。

E-mail: cq44wuwei@163.com

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.020

基于方形腔耦合金属波导结构的全光等离子体开关

肖功利[1]，窦婉滢[1]，杨宏艳[2]，徐俊林[1]，杨秀华[1]，张开富[1]，
杨寓婷[1]，李海鸥[1]，傅　涛[1]，陈永和[1]，刘兴鹏[1]，孙堂友[1]

（桂林电子科技大学 1. 广西精密导航技术与应用重点实验室；2. 电子工程与自动化学院，广西 桂林 541004）

摘　要： 提出了一种新型的方形腔耦合金属波导结构，该结构由两个相互平行的矩形金属波导和一个内嵌可连通的方形谐振腔构成。利用方形谐振腔局域表面等离子体实现带阻滤波特性，并通过多路复用实现双端口全光等离子体开关。采用时域有限差分方法（FDTD）研究了方形腔的边长、折射率和谐振距离对强透射特性的影响。结果表明，基于方形腔耦合金属波导结构的光开关在工作中具有较好的阻带特性和透射特性，其最大透射率可达 92%，最小阻带透射率达 0.2%，工作波长范围为 607～785 nm。

关键词： 全光开关；时域有限差分方法；表面等离激元；方形腔

中图分类号： TN256　**文章编号：** 1001－5868(2020)01－0093－06

All-Optical Plasmonic Switches Based on Square Cavity Coupled Metal Waveguide Structure

XIAO Gongli[1], DOU Wanying[1], YANG Hongyan[2], XU Junlin[1], YANG Xiuhua[1], ZHANG Kaifu[1], YANG Yuting[1], LI Haiou[1], FU Tao[1], CHEN Yonghe[1], LIU Xingpeng[1], SUN Tangyou[1]

(1. Guangxi Key Lab. of Precision Navigation Technol. and Application, Guilin University of Electronic Technol., Guilin Guangxi 541004, CHN; 2. School of Electronic Engin. and Automation, Guilin University of Electronic Technol., Guilin 541004, CHN)

Abstract: A novel square cavity coupled metal waveguide structure is proposed, which is composed of two mutually perpendicular rectangular metal waveguides and a square resonant cavity connecting rectangular waveguides. The band-rejection filtering characteristics are realized by the plasma local characteristics of the square resonator, and the dual-port all-optical plasma switch is realized by multiplexing. The effects of the length, refractive index and resonance distance of the square cavity on the strong transmission characteristics were studied by the finite difference time domain method. The results show that the optical switch based on the square cavity coupled metal waveguide structure has better stop-band characteristics and transmission characteristics, and its maximum transmittance is 92%, the minimum stopband transmittance is 0.2%, and the working range is 607～785 nm.

Key words: all-optical switch; TDFDM; surface plasmon polaritons; square cavity

0　引言

表面等离激元（Surface Plasmon Polaritons，SPP）是金属表面的自由电子与入射光波相互作用产生的特殊电磁波，能够被局域在金属表面并传播，同时会在垂直于金属介质交界处呈指数衰减[1]。SPP 因其良好的局域能力并能突破光学衍射极限的限制，被广泛应用于等离子光学器件研究中[2-3]，其中最为广泛的是等离子体光波导器件，主要包括绝缘体-金属-绝缘体（IMI）和金属-绝缘体-金属

收稿日期：2019－10－03.

基金项目：国家自然科学基金项目（61465004，61765004，61874036，61764001）；广西自然科学基金项目（2017GXNSFAA198164，2016GXNSFAA380006）；桂林电子科技大学研究生教育创新计划资助项目（2017YJCX41，2019YCXS028）；广西精密导航技术与应用重点实验室项目（DH201804，DH201703）.

(MIM)波导结构。而 MIM 波导因其结构简单、集成性高、局域能力强等优异特性被广泛研究[4-6]，各种基于 MIM 结构的等离子体波导器件被提出，如等离子分束器[7]、MZ 干涉仪[8-10]、布拉格光栅反射器[11-12]、光学缓冲器[13]、环形谐振器[14-15]、波分复用器[16]以及各种滤波器[17-18]等都已经被研究和验证。然而，某些特点的光学器件研究还不够多，如基于 SPP 的全光开关。目前已经提出的全光开关结构[19-22]，由于其结构复杂、利用效率低，难以有效实现光路集成。文献[19]提出了一种基于耦合圆环形谐振腔等离子体波导的全光双向开关，但是其主要问题是工作波长范围较窄，工作时传输损耗较大。

针对上述问题，本文提出了一种新型的方形腔耦合金属波导结构，应用时域有限差分方法(FDTD)对方形腔与波导组成的纳米结构的传输响应进行了数值仿真和分析。通过改变方形腔的边长、折射率、谐振距离来探讨该结构的滤波特性及传输响应，相比于文献[19]的报道结果，基于方形腔耦合金属波导结构构建的全光等离子开关的半波宽(FWHM)降低，还可实现双向开关特性。通过优化这些参数，发现该结构可获得较好的光学开关特性，其工作时的最高传输效率较文献[19]提高 10%，最小阻带透射率低至 0.2%，平均工作波长范围为 607～785 nm。该研究结果可为未来全光集成电路中的光开关设计提供理论参考。

1 结构设计与数值模拟

图 1(a)为本文设计的方形腔耦合金属波导结构的三维示意图，一个矩形波导作为传输波导位于结构下方，一个方形耦合腔通过矩形耦合波导与传输波导进行耦合。平面光从结构左侧平行入射激发 SPP 并在结构中传播，最终从传输波导右侧出射。采用 Ag 作为金属材料，SiO_2 作为基底。图 1(b)为此结构的 x-y 平面二维结构图，设置 MIM 波导宽度 $W=100$ nm，其 TM 模式在 MIM 波导中的色散关系为[19]

$$\varepsilon_{in}k_{z2}+\varepsilon_m k_{z1}\coth\left(-\frac{ik_{z1}}{2}\omega\right)=0 \tag{1}$$

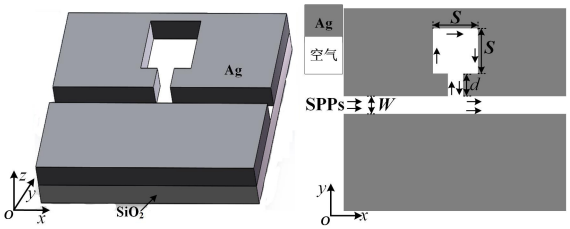
k_{z1}，k_{z2} 由动量守恒分别定义为

$$\begin{aligned}k_{z1}^2&=\varepsilon_{in}k_0^2-\beta^2\\k_{z2}^2&=\varepsilon_m k_0^2-\beta^2\end{aligned} \tag{2}$$

式中，ε_{in} 是绝缘体(空气)的介电常数，ε_m 是金属(银)的介电常数，β 是光在真空中的传播常数，$k_0=2\pi/\lambda_0$ 是自由空间波矢量。因为银是在可见光谱中损耗最低的金属[23]，所以金属选用银作为材料，并采用标准 Drude 模型[24]：

$$\varepsilon_m=\varepsilon_\infty-\frac{\omega_p^2}{\omega^2+i\omega\gamma} \tag{3}$$

式中，$\varepsilon_\infty=3.7$，$\omega_p=9.1$ eV，$\gamma=0.018$ eV，ε_∞ 是 ω 趋于无穷大时金属的相对介电常数，ω_p 是等离子振荡频率，γ 是金属中自由电子的碰撞衰减频率，ω 是入射光的角频率。



(a) 三维结构示意图 (b) x-y 平面二维结构示意图

图 1 方形腔耦合金属波导结构

2 结果与讨论

本文采用 FDTD[25]方法研究了金属波导的传输特性，通过模拟方形腔结构参数分析该参数对结构传输特性的影响。方形腔耦合金属波导结构中的光透射率被定义为

$$T=P_{out}/P_{in}=|E_{out}/E_{in}|^2 \tag{4}$$

式中，P_{out} 是方形腔耦合金属波导结构在工作时的透射光波功率，P_{in} 是平面光在结构中的入射光波功率，E_{out} 为透射光波电场强度，E_{in} 为入射光波电场强度。图 2 为方形腔耦合金属波导滤波结构透射光谱图(实线)。为了便于对比，图中同时给出了圆环腔耦合金属波导滤波结构的透射光谱图[19]，如图 2 中虚线所示。为了分析该结构在工作时的滤波特性，定义其品质因子[26]为

$$Q=\frac{\lambda}{FWHM} \tag{5}$$

式中，λ 为 MIM 波导的共振波长，FWHM 为谐振半峰全宽。这两种结构在透射光谱中都出现了单共振波谷，T_{deep} 和 λ_{deep} 分别为共振波谷的透射率和波长。方形腔边长 $S=100$ nm，波导宽度 $W=100$ nm，耦合波导长度 $d=150$ nm。从图中可以看出，方形腔耦合金属波导结构的共振波长 λ_{deep} 为 990 nm，透射波谷透射率 T_{deep} 为 0.2%，其对应半波宽为 71 nm，品质因子为 13.9；对应地，圆环耦合金属波

导结构的共振波长 λ_{deep} 为 902 nm，其 T_{deep} 为 0.3%，半波宽为 164 nm，品质因子为 5.5。研究结果表明，与圆环腔耦合金属波导结构相比，方形腔耦合金属波导结构表现出较好的滤波特性，其半波宽减小了 93 nm，品质因子增加了 8.4，共振峰位置也出现了红移。这是由于该复合结构中的能量在方形腔中具有更强的 SPP 局域增强效果。为了对以上现象进行深入分析，本文分别模拟了这两种结构在阻带时的电场密度分布，如图 3 所示。可以明显地看出，方形腔比圆环腔局域了更多的 SPP 能量。

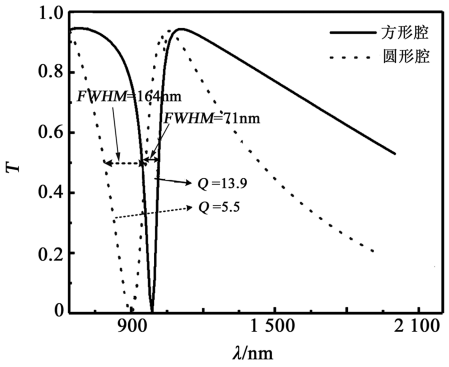


图 2　两种波导结构的透射光谱图

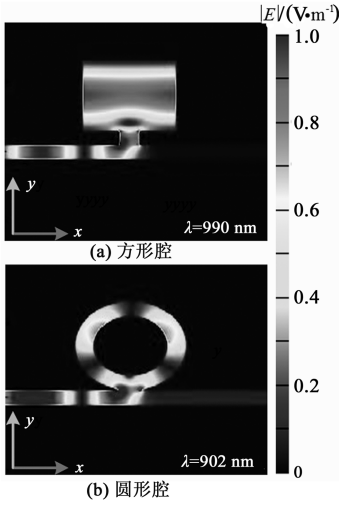
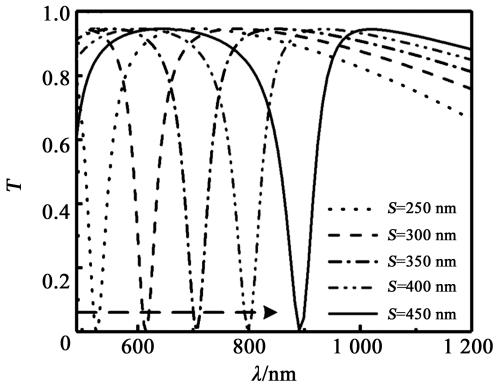
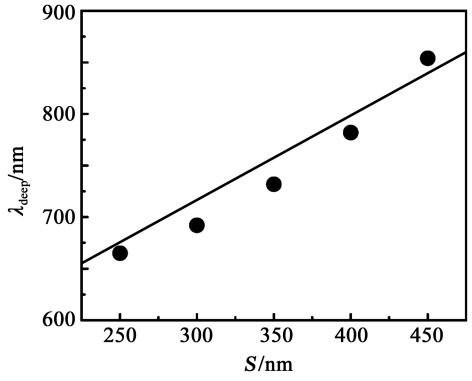


图 3　两种波导结构在不同共振波长下的电场密度分布

为了对方形腔耦合金属波导结构的强透射现象进行更深入的研究，本文对可能影响波导结构工作时的 T_{deep} 和 λ_{deep} 的因素 S（方形腔边长）、d（耦合波导长度）、n（方形腔折射率）进行了研究和探讨。首先讨论结构中方形谐振腔的边长对滤波性能的影响。令 $W=100$ nm，$d=150$ nm，$n=1$，改变方形腔边长 S，使 S 分别为 250，300，350，400 和 450 nm。如图 4(a)所示，随着 S 的逐渐增加，共振波谷发生红移，因此，可以通过控制 S 实现在一定范围内的选频效果。从图 4(b)中可以看出，λ_{deep} 与 S 呈线性关系。模拟结果显示，滤波器透射的共振波长可以通过调整共振腔边长大小来改变。



(a)　不同 S 下的透射光谱图



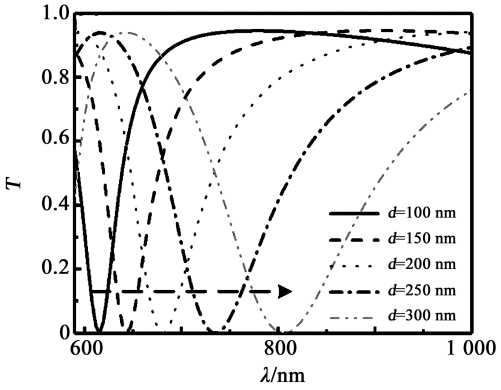
(b)　λ_{deep} 与 S 的关系

图 4　方形腔耦合金属波导结构在不同边长下的透射光谱及共振谷波长

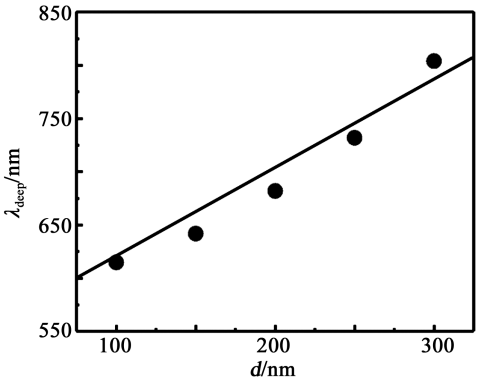
研究显示，该复合结构的滤波性能还受谐振腔和传输波导间的耦合距离 d（耦合波导长度）影响，如图 5 所示。令波导宽度 $W=100$ nm，当共振腔边长 $S=300$ nm、方形腔折射率 $n=1$，改变 d，使得 d 分别为 100，150，200，250 和 300 nm。结果表明，透射峰随着 d 的增加而红移。其主要原因是：随着耦合距离 d 的变化，方形腔与传输波导的距离逐渐增加，共振腔对 SPP 的局域能力逐渐降低，因此透射波谷的半波宽逐渐增加。模拟结果显示通过调整方形腔和传输波导的耦合距离也可以改变滤波器的共振波长，但随着耦合距离的增加，该结构对能量的局域能力逐渐降低，透射峰的半波宽逐渐增加。

另外，研究还发现方形腔内介电折射率 n 对共振峰透射率也有很大的影响。令波导宽度 $W=100$ nm，当共振腔边长 $S=300$ nm、耦合距离 $d=150$ nm 不变，改变方形腔内介电折射率 n，令其分别为 1.0，1.1，1.2，1.3 和 1.4。如图 6(a)所示，当

n 增大，透射波谷发生红移。从图 6(b)中可以看出，共振波长与介电折射率呈线性关系。模拟结果表明通过改变介电折射率也可以线性调整滤波器透射共振波长。



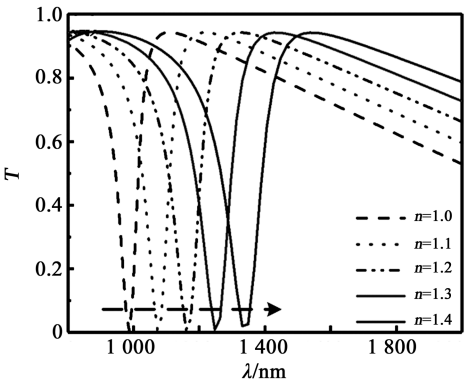
(a) 在不同 d 下的透射光谱图



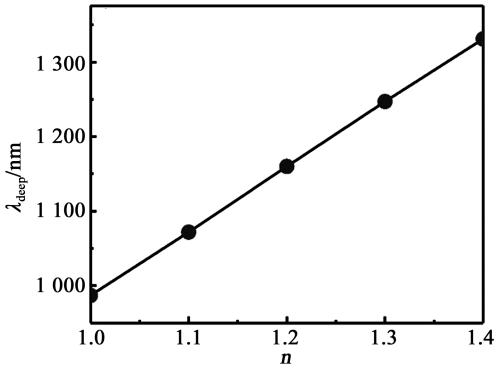
(b) λ_{deep} 与 d 的关系

图 5 方形腔耦合金属波导结构在不同耦合距离下的透射光谱及共振谷波长

根据上述对基于方形腔耦合金属波导结构的等离子体滤波特性分析，本文提出了一种双通道等离子体光开关结构，如图 7 所示。两个不同边长的方形腔位于传输波导两侧，且同时耦合传输波导，为了区分不同的出射口，本文把出口标记为 Port1 和 Port2。相应的结构参数设置为：波导宽度 $W=100$ nm，两方形腔边长 S_1 和 S_2 分别为 300 和 400 nm，耦合距离 $d=100$ nm。入射光由左侧入射波导进入并激发 SPP 传输，与两个不同的方形腔进行耦合，产生不同透射光谱，实现双口等离子光开关功能。图 8 为两端口的透射光谱。Port1 和 Port2 的共振峰和共振波谷的共振波长分别为 607 和 785 nm。从图中可以看出，提出的双通道结构能够较好地实现双口等离子体光开关功能。通过控制入射光的光波长，可动态地控制光导通端口。从图 9 所示共振波长下的电场分布可以看出，当某一个端



(a) 在不同的介电折射率 n 下的透射光谱图



(b) λ_{deep} 与介电折射率 n 的关系

图 6 方形腔耦合金属波导结构在不同 d 下的透射光谱及共振谷波长

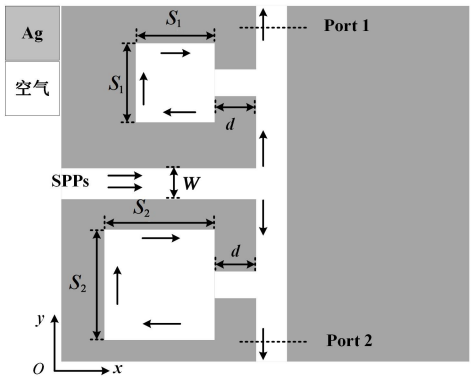


图 7 双通道等离子体光开关二维结构示意图

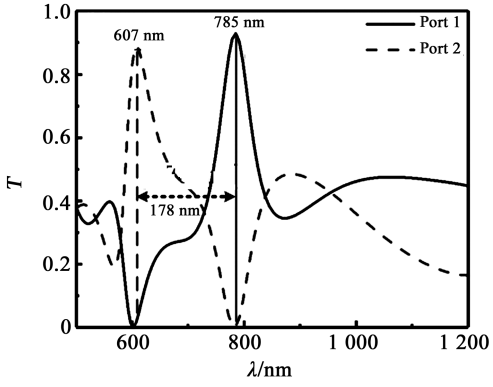
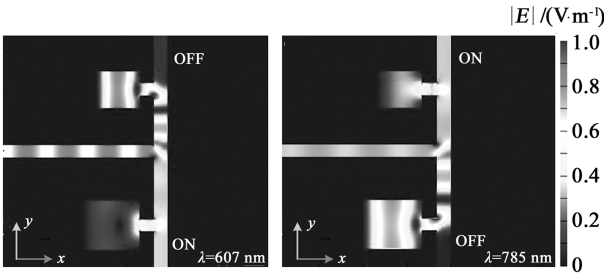


图 8 两端口的透射光谱图



(a) Port1 OFF,Port2 ON　　(b) Port1 ON,Port2 OFF

图 9　共振波长下的电场分布

口处于导通状态(ON)时,另一个端口耦合的方形腔产生能量局域效果,从而使其呈现闭合状态(OFF),实现了光开关功能。与文献[19]相比,基于方形腔耦合金属波导结构的全光等离子体开关,具有较好的传输特性,其工作导通时最高传输效率为 92%,闭合状态最小传输效率为 0.2%,工作波长范围为 607～785 nm。

3　结论

提出了一种内嵌方形腔耦合金属波导滤波器结构,并通过多端口耦合实现双刀等离子开关。采用 FDTD 方法对该结构进行研究,对该滤波器结构中的方形腔边长、腔内折射率、腔与传输波导的谐振距离等因素进行了分析,最终提出并分析了基于该结构设计的全光等离子体开关。研究结果表明:采用方形腔耦合金属波导结构代替圆环腔耦合金属波导结构具有更好的滤波特性,前者具有更好的 SPP 局域特性,具体表现为半波宽减小了 93 nm,品质因子提高了 8.4;通过控制方形腔的边长、折射率、谐振距离等参数可动态改变滤波器的共振波长;当增加 S 和 n,其共振峰都发生红移,从而可通过参数的优化达到选频的效果;通过增加耦合距离 d,也可以线性增加其共振波长,但同时其半波宽也增加。基于以上特性,设计了一种双通道全光等离子体开关,利用方形腔的局域特性实现对光开关的控制,通过改变入射光波长可动态控制光开关的导通与闭合;光开关工作频率可通过改变方形腔的参数来确定,导通时最高传输效率可达 92%,闭合时最小阻带透过率为 0.2%,表明该结构具有较好的光开关传输特性。该研究成果可为下一代高性能全光等离子体开关的设计提供理论参考。

参考文献:

[1] Barnes W L, Dereux A, Ebbesen T W. Surface plasmon subwavelength optics[J]. Nature, 2003, 424(6950): 824-830.

[2] Neutens P, Dorpe P, Vlaminck L, et al. Electrical detection of confined gap plasmons in metal-insulator-metal waveguides[J]. Nature Photonics, 2009, 3(5): 283-286.

[3] Rumke T M, Sanchez-Gil J A, Muskens O L, et al. Local and anisotropic excitation of surface plasmon polaritons by semiconductor nanowires[J]. Opt. Express, 2008, 16(7): 5013.

[4] Gramotnev D K, Bozhevolnyi S I. Plasmonics beyond the diffraction limit[J]. Nature Photon., 2010, 4(2): 83-91.

[5] Kazuo T, Masahiro T, Tatsuhiko S. Simulation of practical nanometric optical circuits based on surface plasmon polariton gap waveguides[J]. Opt. Express, 2005, 13(1): 256-266.

[6] 肖功利, 徐俊林, 杨宏艳, 等. 内嵌矩形金属块纳米圆盘结构等离子体多通道波分复用器研究[J]. 光学学报, 2019, 38(12): 90-97.
Xiao Gongli, Xu Junlin, Yang Hongyan, et al. A plasmon multi-channel wavelength-division multiplexer constructed with a nanodisk structure embedded in a rectangular metal block[J]. Acta Optica Sinica, 2019, 38(12): 90-97.

[7] Hansik, Seung-Yeol, Kyoung-Youm, et al. Hybrid states of propagating and localized surface plasmons at silver core/silica shell nanocubes on a thin silver layer[J]. Opt. Express, 2014, 22(7): 8383.

[8] Weiren, Rukhlenko, Malin. Manipulating energy flow in variable-gap plasmonic waveguides[J]. Opt. Lett., 2012, 37(24): 5151.

[9] Sven, Andreas, Christoph, et al. Novel efficient design of Y-splitter for surface plasmon polariton applications[J]. Opt. Express, 2008, 16(19): 14369-14379.

[10] Berini. Bulk and surface sensitivities of surface plasmon waveguides[J]. New J. of Physics, 2008, 10(10): 105010.

[11] Bozhevolnyi, Volkov, EloSe, et al. Channel plasmon subwavelength waveguide components including interferometers and ring resonators[J]. Nature, 2006, 440(7083): 508-511.

[12] Wahsheh, Lu, Abushagur. Nanoplasmonic directional couplers and Mach-Zehnder interferometers [J]. Opt. Communications, 2009, 282(23): 4622-4626.

[13] Junghyun, Hwi, Byoungho. High order plasmonic Bragg reflection in the metal-insulator-metal waveguide Bragg grating[J]. Opt. Express, 2008, 16(1): 413.

[14] Yongkang, Leiran, Xiaohong, et al. Broad-bandgap and low-sidelobe surface plasmon polariton reflector with Bragg-grating-based MIM waveguide[J]. Opt. Express, 2009, 17(16): 13727-13736.

[15] Ding Qiaoqiang, Bartoli. "Rainbow" trapping and releasing at telecommunication wavelengths[J]. Phys. Rev. Lett., 2009, 102(5): 056801.

[16] Lee, Da, Kwon. Dual-function metal-insulator-metal plasmonic optical filter[J]. IEEE Photon. J., 2015, 7(1):

1-8.

[17] Pixin, Ruisheng, Qiaodong, et al. Plasmonic filters and optical directional couplers based on wide metal-insulator-metal structure[J]. Opt. Express, 2011, 19(8): 7633.

[18] Tong-Biao, Xie-Wen, Cheng-Ping, et al. The transmission characteristics of surface plasmon polaritons in ring resonator [J]. Opt. Express, 2009, 17(26): 24096-24101.

[19] Bana, Pang, Li, et al. A nonlinear plasmonic waveguide based all-optical bidirectional switching[J]. Opt. Commun., 2017: 406.

[20] Liu, Gao, Zhu, et al. A T-shaped high resolution plasmonic demultiplexer based on perturbations of two nanoresonators [J]. Opt. Commun., 2015, 334: 164-169.

[21] Roy, Kulshrestha. All-optical switching in plant blue light photoreceptor phototropin [J]. IEEE Trans. on Nanobioscience, 2006, 5(4): 281-287.

[22] Granpayeh, Habibiyan, Parvin. Photonic crystal directional coupler for all-optical switching, tunable multi/demultiplexing and beam splitting applications [J]. J. of Modern Optics, 2018: 1-8.

[23] 肖功利，徐俊林，杨宏艳，等. 基于锯齿共振腔耦合金属波导结构的多通道等离子体逻辑门输出光源[J]. 激光与光电子学进展，2019，56(12)：135-141.
Xiao Gongli, Xu Junlin, Yang Hongyan, et al. Multi-channel plasmonic logic-gate output light source based on sawtooth resonant cavity coupled with metal-waveguide structure [J]. Laser & Optoelectronics Progress, 2019, 56(12): 135-141.

[24] Bozhevolnyi. Compact Bragg gratings for long-range surface plasmon polaritons[J]. J. of Lightwave Technol., 2006, 24 (2): 912-918.

[25] Rylander T. The Finite-Difference Time-Domain Method [M]. New York: Springer, 2013: 1003-1007.

[26] Guo Bing. Plasmon Bragg reflectors and nanocavities on flat metallic surfaces[J]. Appl. Phys. Lett., 2005, 87(1): 824.

作者简介：

肖功利(1975—)，男，博士，研究员，硕士生导师，主要从事微纳光电子器件与表面等离激元等方面的研究。

E-mail：xgl.hy@126.com

(上接第 88 页)

[7] 彭金涛，任天斌. 碳纤维增强树脂基复合材料的最新应用现状[J]. 中国胶粘剂，2014，23(8)：48-52.
Peng Jintao, Ren Tianbin. The latest application status of carbon fiber reinforced resin matrix composites [J]. China Adhesives, 2014, 23(8): 48-52.

[8] 孟庆平，孙 维，王亚龙，等. 植入 FBG 的碳纤维复合材料全寿命周期结构状态研究[J]. 宇航材料工艺，2018，48(4)：4650.
Meng Qingping, Sun Wei, Wang Yalong, et al. Whole life cycle structure of carbon fiber composites implanted with FBG [J]. Aerospace Materials & Technol., 2018, 48(4): 4650.

[9] 杨正岩，张佳奇，高东岳，等. 航空航天智能材料与智能结构研究进展[J]. 航空制造技术，2017(17)：36-48.
Yang Zhengyan, Zhang Jiaqi, Gao Dongyue, et al. Advance of aerospace smart material and structure [J]. Aeronautical Manufacturing Technol., 2017(17): 36-48.

[10] 陈 显，余尚江，林 明，等. 一种复合材料封装光纤光栅片式应变传感器研究[J]. 传感器与微系统，2008，27(12)：43-44，47.
Chen Xian, Yu Shangjiang, Lin Ming, et al. Research on optic fiber Bragg grating strain sensors with compositive slip capsulation [J]. Transducer and Microsystem Technologies, 2008, 27(12): 43-44, 47.

[11] 田高洁，李 川，由 静，等. 碳纤维复合材料封装光纤 Bragg 光栅[J]. 大坝与安全，2010(1)：28-30.
Tian Gaojie, Li Chuan, You Jing, et al. Fiber Bragg grating packaged by carbon fiber-reinforced plastic[J]. Dam & Safety, 2010(1): 28-30.

[12] 高琳琳，王庆林，郭云力，等. 复合材料基片式光纤光栅传感器的制造与性能[J]. 材料工程，2018，46(9)：65-72.
Gao Linlin, Wang Qinglin, Guo Yunli, et al. Fabrication and performance of optical fiber grating sensors packaged in composite substrate[J]. J. of Materials Engin., 2018, 46(9): 65-72.

[13] 梁 磊，姜德生，孙东亚. 光纤传感器在混凝土结构中的相容性研究[J]. 武汉工业大学学报，2000(2)：11-14.
Liang Lei, Jiang Desheng, Sun Dongya. Study on compatibility of fiber optical sensor in construct of concrete[J]. J. of Wuhan University of Technol., 2000(2): 11-14.

[14] 孙 丽. 光纤光栅传感应用问题解析[M]. 北京：科学出版社，2012：8-32.
Sun Li. Application Analysis of Fiber Bragg Grating Sensor [M]. Beijing: Science Press, 2012: 8-32.

[15] 吴入军，郑百林，贺鹏飞，等. 埋入式光纤布拉格光栅传感器封装结构对测量应变的影响[J]. 光学精密工程，2014，22 (1)：24-30.
Wu Rujun, Zheng Bailin, He Pengfei, et al. Influence of encapsulation structures for embedded fiber-optic Bragg grating sensors on strain measurement [J]. Optics and Precision Engin., 2014, 22(1): 24-30.

作者简介：

侯智懿(1995—)，女，安徽省人，硕士研究生，主要研究方向为光纤光栅传感器；

邵洪峰(1969—)，男，河北省人，硕士，高级工程师，主要研究方向为特种光纤技术、光纤传感技术。

E-mail：hfshao@163.com

材料、结构及工艺

DOI：10.16818/j.issn1001－5868.2020.01.021

NLDMOS器件性能优化及分析

李维杰[1]，王　兴[2]，王云峰[1]，李　洋[2]，孟丽华[2]

（1. 中国科学院大学 微电子学院，北京 100029；2. 中芯国际集成电路制造有限公司，天津 300385）

摘　要： 提出一种改善n型横向双扩散金属氧化物半导体(NLDMOS)器件性能的工艺方法。该方法基于某公司0.18 μm标准工艺流程，通过在NLDMOS的共源处增加一道离子注入，引出衬底电荷，以优化NLDMOS器件的击穿电压(V_b)与比导通电阻(R_{sp})。选择不同的注入离子浓度与快速热退火时间，研究了器件的V_b与R_{sp}变化。由于离子激活效率不足，单纯增加20%的注入离子浓度，器件的耐压性能提升极小，采用增加20%注入离子浓度结合延长20 s快速热退火时间的方法，NLDMOS器件的V_b提高约2.7%，同时R_{sp}仅增加0.9%左右。

关键词： NLDMOS；击穿电压；比导通电阻；离子注入；快速热退火

中图分类号： TN43　**文章编号：** 1001－5868(2020)01－0099－04

Performance Optimization and Analysis of NLDMOS Devices

LI Weijie[1], WANG Xing[2], WANG Yunfeng[1], LI Yang[2], MENG Lihua[2]

(1. School of Microelectron., University of Chinese Academy of Sciences, Beijing 100029, CHN;
2. Semiconductor Manufacturing Inter. Corp., Tianjin 300385, CHN)

Abstract: A new process for improving the performance of n-LDMOS devices is proposed based on XXX's 0.18 μm standard process flow. Adding an ion implantation at the common source of NLDMOS can extract the substrate charge and optimize the breakdown voltage (BV) and specific on-resistance (R_{sp}) of the NLDMOS devices. In this paper, by selecting different implant ion concentrations and rapid thermal annealing time, changes of BV and R_{sp} of the devices are studied. Because the ion activation efficiency is insufficient, simply increasing the implant ion concentration by 20%, the withstand voltage performance of the device has less improvement. By combining 20% increase in implant ion concentration with 20 s rapid thermal annealing method, the BV of the NLDMOS device is increased by about 2.7%, while the R_{sp} is increased by only about 0.9%.

Key words: NLDMOS; breakdown voltage; R_{sp}; ion implant; RTA

0　引言

当前半导体技术的主要发展方向有两个，一是按照摩尔定律缩小关键尺寸，实现提高器件性能、降低功耗的目的；另一个是实现芯片功能多样化，发展面向各种应用的工艺，即扩展摩尔定律(More than Moore)。随着摩尔定律的不断发展，关键尺寸越来越小，导致进一步缩小关键尺寸的难度剧增，因此芯片功能多样化的发展方向逐渐引起人们的重视。BCD(Bipolar-CMOS-DMOS)工艺将双极晶体管(Bipolar Transistor，BT)、互补性金属氧化物半导体(Complementary Metal Oxide Semi-conductor，CMOS)和双扩散金属氧化物半导体(Double Diffused Metal Oxide Semiconductor，DMOS)集成在同一芯片上，因此兼具这三种器件的优点，该工艺在电源管理、汽车电子等领域都有重要应用[1-5]。

横向双扩散金属氧化物半导体(Lateral Double Diffused Metal Oxide Semiconductor，LDMOS)器件为BCD工艺中的核心器件，作为功率器件，

收稿日期：2019－11－23.

LDMOS最重要的两个参数为比导通电阻(R_{sp})与击穿电压(V_b)。R_{sp}降低意味着相同工作电压下的驱动电流更大,电路驱动能力更强;V_b提高意味着器件能承受更大的电压,耐压性能与稳定性更好。因此业界针对LDMOS的研究集中在R_{sp}增加幅度尽可能小的前提下,提高器件V_b。但是R_{sp}与V_b本身存在制约关系,提高V_b会伴随R_{sp}增加[6]。本文以n型LDMOS(N-Type Lateral Double Diffused Metal Oxide Semiconductor,NLDMOS)器件为研究对象,研究其DPP(Deep P-type Pick Up)结构,选择不同的注入离子浓度与快速热退火(Rapid Thermal Anneal,RTA)条件进行实验,总结合适的工艺条件,达到在R_{sp}增加幅度尽量小的前提下提高器件V_b的目标。

1 器件结构

图1为NLDMOS器件的截面图,采用业界标准方法形成器件沟道与漂移区,注入硼(B)离子和砷(As)离子,其扩散浓度差形成沟道;器件漏极与沟道之间注入低浓度的n型离子,形成漂移区。在形成多晶硅栅极之后,增加一道离子注入工艺与相应的快速热退火修复工艺,形成DPP结构,离子注入的位置在两个NLDMOS器件的共同源极处。

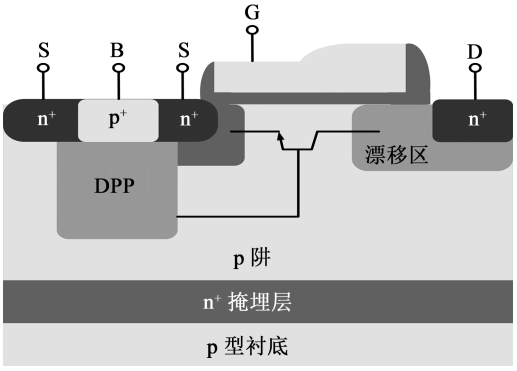


图1 NLDMOS结构示意图

在器件工作时,沟道有导通电流经过,在电场强的位置会有电荷碰撞离化产生电子,部分电子会散射到衬底,如果衬底处的电荷不能及时引出,衬底局部的电荷积累将导致衬底电位升高,升高一定程度时,源端到衬底再到漏端的寄生三极管就会被打开,使得电流增大。增大的电流进一步激发电荷的碰撞电离,使寄生三极管进入正反馈状态,触发折回效应(Snap Back,SB),导致器件源漏电流过大而失效,器件V_b也因此降低[7]。增加DPP结构是为引出衬底电流,降低衬底电位,使寄生三极管处于关闭状态,折回效应不会发生,器件的耐压得以提高。

2 实验及结果

2.1 注入离子浓度实验

2.1.1 注入离子浓度实验分组

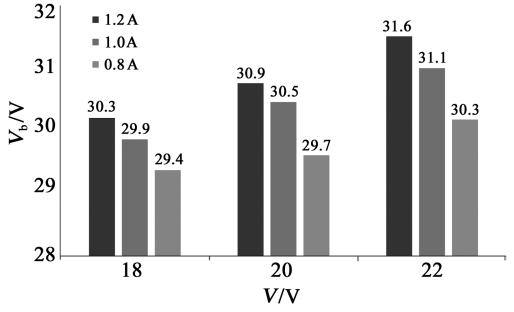
第一组实验变量为注入离子浓度,实验采用p型衬底晶圆,DPP离子注入源为BF_2,提供p型的硼离子。离子注入由两步完成,保证浓度梯度更均匀,第一步离子注入能量为160~200 keV,第二步离子注入能量为50~80 keV,两步注入离子浓度数量级为10^{14} cm^{-3}。表1为具体实验方案,以A表示注入离子浓度的基准值,0.8A表示两步注入离子浓度均为基准值的0.8倍。

表1 注入离子浓度实验方案

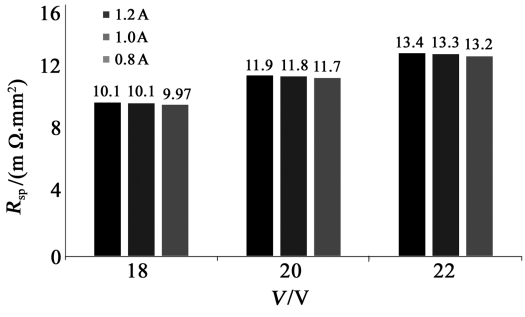
实验分组	注入离子浓度/(10^{14} cm^{-3})
1	1.2A
2	1.0A
3	0.8A

2.1.2 注入离子浓度实验结果

实验电性参数测试(Wafer Acceptance Test,WAT)结果如图2所示,选择不同工作电压的NLDMOS器件进行检测,器件工作电压V和V_b单位为V,R_{sp}单位为mΩ·mm^2。不同工作电压下的NLDMOS器件随注入离子浓度的改变呈现相同的



(a) 对击穿电压的影响



(b) 对比导通电阻的影响

图2 不同注入离子浓度对V_b和R_{sp}的影响

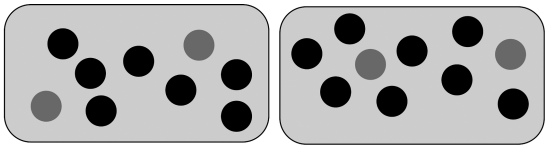
趋势：由于 DPP 注入离子浓度下降，导致整个 DPP 结构载流子浓度下降，因此 DPP 结构引出衬底电荷的能力下降，器件的 V_b 下降；DPP 离子注入的位置在器件的源端，因为是 NLDMOS，所以器件源极注入 n 型离子，部分 n 型离子与之前 DPP 注入的 p 型离子发生中和，理论上 DPP 注入的离子浓度越高，中和的几率越大，源端的有效载流子浓度越低，导致比导通电阻越大。但是因为 DPP 注入的离子位置深，与源极离子注入的位置重叠区域小，导致对 R_{sp} 影响在 1%以内。

表 2 列出不同离子浓度引起 V_b，R_{sp} 变化的百分比均值，由表中数据可知，如果降低注入离子浓度，器件的 V_b 下降 2.43%，器件耐压性能大幅下降。相比较而言，增加注入离子浓度引起 V_b，R_{sp} 的变化要小，V_b 增加了 1.29%，虽然器件耐压性能提升，但是对比减少注入离子浓度的 V_b 变化，增加注入离子浓度带来的耐压性能提升没有达到最优。

通过示意图解释增加注入离子浓度但 V_b 提升不显著的原因，图 3 中黑色圆点表示被激活的杂质离子，浅色圆点代表经过热退火之后依然处于间隙位置、对载流子传输无用的杂质离子。由图 3(a)可见，注入离子浓度低时，原本处于晶格间隙位置的杂质离子激活效率高。但是图 3(b)显示，虽然增加了注入离子浓度，由于退火条件不变，退火不够充分，额外的注入离子并未被激活，限制了实际提升效果。

表 2　不同注入离子浓度下 V_b，R_{sp} 变化表

注入离子浓度/(10^{14} cm^{-3})	V_b 变化/%	R_{sp} 变化/%
1.2A	1.29	0.49
1.0A	0	0
0.8A	−2.43	−0.91



(a)　低浓度注入离子　(b)　高浓度注入离子

图 3　不同注入离子浓度退火后示意图

2.2　快速热退火条件实验

2.2.1　快速热退火时间实验分组

通过上述分析，增加注入离子浓度能使 NLDMOS 器件耐压性能提升，但是注入离子的激活效率不足导致耐压性能提升不足，为更大程度发挥增加注入离子浓度带来的器件耐压性能提升效果，需要考虑更充分的快速热退火(RTA)条件。

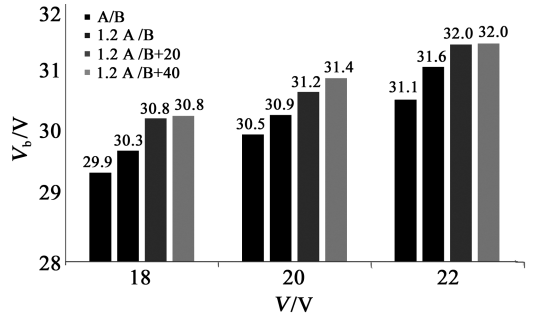
表 3 为实验方案，实验同样采用阻值为 p 型衬底晶圆，注入离子浓度基准值 A 与实验 1 数值相同，注入离子浓度数量级为 10^{14} cm^{-3}。RTA 温度的基准值为 1 000～1 100 ℃，表中 B 表示 RTA 时间的基准值，其值为 30～50 s。

表 3　RTA 时间实验方案

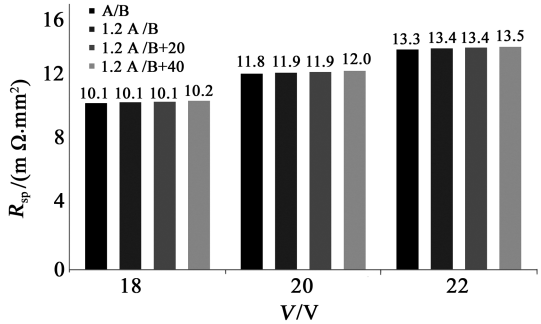
实验分组	注入离子浓度/(10^{14} cm^{-3})	RTA 时间/s
1	1.0A	B
2	1.2A	B
3	1.2A	B+20
4	1.2A	B+40

2.2.2　快速热退火时间实验结果

实验 WAT 结果如图 4 所示，测试器件工作电压 V 仍然为 18、20 和 22 V。只增加注入离子浓度而 RTA 不改变，V_b 增加 1.29%，R_{sp} 增加 0.49%，此结果与实验 1 基本一致，器件耐压性能提升有限。将 RTA 时间增加 20 s，器件的 V_b 提升 2.67%，器件的耐压性能明显提升，说明在 RTA 增加 20 s 后，更多的注入离子被激活。



(a)　对击穿电压的影响



(b)　对比导通电阻的影响

图 4　RTA 时间对 V_b 和 R_{sp} 的影响

但 RTA 时间不是越长越好，过长会导致瞬态增强扩散效应(Transient Enhanced Diffusion，TED)[8]更明显，造成杂质离子再分布，杂质离子扩散范围变大，使器件性能产生漂移。如表 4，实验结果显示，在 RTA 时间增加 20 s 后，R_{sp} 增加了

0.86%，这与仅增加注入离子浓度的实验组结果相比已经有所增加；RTA 时间增加 40 s，R_{sp} 增加的比率高达 1.37%，而且 V_b 提升 2.96%，说明增加 20 s RTA 时间，注入离子激活效率已经很高，在此基础上再增加 20 s RTA 时间带来的耐压性能提升不高。因此 RTA 时间不能过长，综合考虑耐压性能与器件参数漂移导致 R_{sp} 增加两因素，增加 20 s RTA 时间为最合适的条件。

表 4 不同 RTA 时间条件下 V_b，R_{sp} 变化表

注入离子浓度/(10^{14} cm^{-3})	RTA 时间/s	V_b 变化/%	R_{sp} 变化/%
1.0A	B	0	0
1.2A	B	1.29	0.49
1.2A	B+20	2.67	0.86
1.2A	B+40	2.96	1.37

3 结论

本文提出了一种优化 NLDMOS 器件耐压性能的方法，即在器件共源处增加一道 p 型离子注入，以引出衬底积累电荷、提高器件 V_b。选择不同注入离子浓度与快速热退火条件形成器件的 DPP 结构，研究器件的 R_{sp} 与 V_b 变化。单纯增加 20%的注入离子浓度会由于离子激活效率不足导致器件耐压性能提升有限；采用延长 RTA 时间的方法充分激活注入离子，最终使用增加 20%注入离子浓度结合延长 20 s 快速热退火时间的方法，NLDMOS 器件 V_b 提高约 2.7%，R_{sp} 仅增加约 0.9%。

采用本文介绍的工艺方法，得到工作电压为 18、20 和 22 V 的 NLDMOS 器件的击穿电压分别为 30.8、31.2 和 32 V，比导通电阻分别为 10.1、11.9 和 13.4 mΩ·mm^2，可作为 LDMOS 器件耐压性能研究的参考。

参考文献：

[1] 陈 铭. BCD(Bipolar-CMOS-DMOS)工艺技术的发展、现状及展望[J]. 集成电路应用，2014(7)：24-28.
Chen Ming. Development，status and prospect of BCD(Bipolar-CMOS-DMOS) technology[J]. Integrated Circuit Application，2014(7)：24-28.

[2] 杨银堂，朱海刚. BCD 集成电路技术的研究与进展[J]. 微电子学，2006，36(3)：315-319.
Yang Yintang， Zhu Haigang. State-of-the-art of BCD technology and its developing trends[J]. 2006，36(3)：315-319.

[3] Park I Y，Choi Y K，Ko K Y，et al. BD180-a new 0.18 μm BCD (Bipolar-CMOS-DMOS) technology from 7 V to 60 V [C]// IEEE 20th Inter. Symp. on Power Semiconductor Devices and IC's，2008.

[4] Tsung Y H，Chien H H，Chih F H，et al. Demonstration of a HV BCD technology with LV CMOS process[C]// 27th Inter. Symp. on Power Semiconductor Devices & IC's，2015：193-196.

[5] Michael T M G，Hu Y H，DEB K P，et al. Design and optimization of 40 V，0.18 μm versatile HVLDMOS device with DOE [C]// IEEE Inter. Conf. on Semiconductor Electron.，2008：1-5.

[6] Amato M，Rumennik V. Comparison of lateral and vertical DMOS specific on-resistance [C]// Inter. Electron Devices Meeting，1985，31：736-739.

[7] 陈轶群，陈佳旅，蒲贤勇. 功率 NLDMOS 电学安全工作区的版图设计优化[J]. 半导体技术，2019(8)：623-627.
Chen Yiqun， Chen Jialv， Pu Xianyong. Optimization of workspace layout design of power NLDMOS electrical security [J]. Semiconductor Technology，2019(8)：623-627.

[8] Min Y， Ru H， Xing Z. Atomistic simulation of RTA annealing for shallow junction formation characterizing both BED and TED [C]// Inter. Conf. on Simulation of Semiconductor Processes and Devices，2002：123-126.

[9] Ankit K，Emita Y H，Vasanth K，et al. Design of a low on resistance high voltage(<100 V) novel 3DNLDMOS with side STI and single P-top layer based on 0.18 μm BCD process technology[C]// 8th NMDC，2013：78-80.

[10] 冯喆韻，马千成，汪 铭. 0.18 μm 完全隔离型低导通态电阻(Low Ron)NLDMOS 研究[J]. 电子与封装，2016，16(7)：39-43.
Feng Zheyun，Ma Qiancheng，Wang Ming. Studies of 0.18 μm fully isolated Low Ron NLDMOS [J]. Electronics and Packaging，2016，16(7)：39-43.

[11] Ko K Y，Park I Y，Choi Y K，et al. BD180LV-0.18 μm BCD technology with best-in-class LDMOS from 7 V to 30 V [C]// 22nd Inter. Symp. on Power Semiconductor Devices & IC's，2010：71-74.

[12] Pendharkar S，Pan R，Tamura T，et al. 7 to 30 V state-of-art power device implantation in 0.25 μm LBC7 BiCMOS-DMOS process technology[C]// 16th Inter. Symp. on Power Semiconductor Devices & IC's，2004：419-422.

[13] Riccardi D，Causio A，Filippi I，et al. BCD8 from 7 V to 70 V：a new 0.18 μm technology platform to address the evolution of applications towards smart power ICs with high logic contents [C]// 19th International Symp. on Power Semiconductor Devices and IC's，2007：73-76.

[14] Shirai K，Yonemura K，Watanabe K，et al. Ultra-low on-resistance LDMOS implementation in 0.13 μm CD and BiCD process technologies for analog power IC's[C]// IEEE 21st Inter. Symp. on Power Semiconductor Devices & IC's，2009.

[15] Campbell S A. 微纳尺度制造工程[M]. 严利人，张 伟，译. 北京：电子工业出版社，2011.

(下转第 140 页)

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.022

微小零件显微检测图像的清晰度评价

李成超，于占江，李一全，许金凯，于化东

（长春理工大学 跨尺度微纳制造技术教育部重点实验室，长春 130022）

摘　要： 根据微小零件显微检测图像纹理相似、边缘信息太少和灰度分布范围有限等特点，从算法的基本原理出发，对传统清晰度评价函数进行了分析。针对传统清晰度评价函数无法兼具高灵敏度和抗噪性的缺陷，提出了一种基于局域方差信息熵的清晰度评价算法。该算法利用局域方差对各灰度级的自信息量进行加权，一方面弥补了信息熵函数缺失的空间信息，避免清晰度误判；另一方面增加清晰区域像素参与计算信息量时的权重，同时减少背景区域和噪声区域像素计算的权重，从而提高函数灵敏度。实验结果表明，与传统清晰度评价函数相比，局域方差信息熵函数不仅灵敏度很高，而且具有较好的抗噪性，可用于实际的自动对焦系统中。

关键词： 自动对焦；清晰度评价；局域方差信息熵；微小零件；显微图像

中图分类号： TP391　　**文章编号：** 1001－5868(2020)01－0103－05

Sharpness Evaluation of Microscopic Detection Image for Micro Parts

LI Chengchao，YU Zhanjiang，LI Yiquan，XU Jinkai，YU Huadong

（Changchun University of Science and Technol.，National Key Lab. for Cross-Scale Micro and Nano Manufacturing of the Ministry of Education，Changchun 130022，CHN）

Abstract： According to the characteristics of micro parts microscopic detection image，including the image texture is similar，the edge information is too little and the gray distribution range is limited，the traditional sharpness evaluation function is analyzed based on the basic principles of algorithm. Aiming at the defect that the traditional sharpness evaluation function can not have own high sensitivity and noise immunity，an algorithm based on local variance information entropy is proposed. The method uses the local variance to weight the self-information of each gray level，on the one hand，it makes up for the lack of spatial information of information entropy and avoids misjudgement of sharpness；on the other hand，it can increase the weight of clear region pixels when they participate in the calculation of information，while reduces the weight of background and noise region pixels，thereby improves the function sensitivity. The experimental results show that compared with the traditional sharpness evaluation function，the local variance information entropy function not only has high sensitivity，but also has better noise immunity and is suitable for actual auto-focusing systems.

Key words： auto-focusing；sharpness evaluation；local variance information entropy；micro parts；microscopic image

0　引言

纳米技术、微型机械和微机电系统的出现促进了微小零件精密加工技术的快速发展[1]，从而对微小零件的检测也提出了更高的要求。基于机器视觉的检测方法由于具有可靠性高、精度高、非接触测量等优点[2]，常用于微小零件的表面微观形貌观测、表面缺陷检测和几何尺寸测量等。为了实现检测过程

收稿日期：2019－11－04.

基金项目：吉林省科技发展计划项目(20190302076GX)；长春理工大学青年科学基金项目(XQNJJ-2018-09).

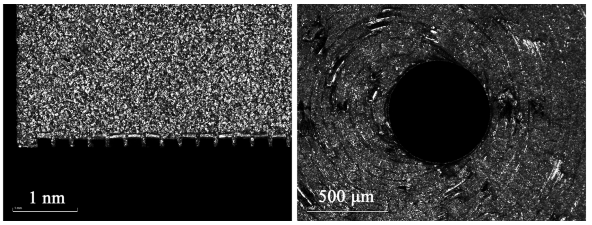
的自动化、智能化，要求机器视觉检测系统具备自动对焦功能。基于图像处理的自动对焦系统的关键技术包括图像清晰度评价和搜索算法[3]。要进行自动对焦，首先要能判断所获取的图像是否为清晰聚焦的图像，通过清晰度评价函数对不同位置所拍摄的图像进行评分，从而定量得出图像清晰度，进而可以确定镜头与零件的聚焦距离并获取清晰聚焦图像。

在基于机器视觉的微小零件显微检测中，满足较高检测精度时，由于镜头放大倍数很高，景深会非常小，从而导致聚焦位置上下范围很小，因此通过清晰度评价函数准确且快速获取清晰聚焦图像是后续图像处理、分析的基础和保障整体检测精度的前提。

大多数清晰度评价函数都是为了用于自动对焦技术，因此理想的图像清晰度评价函数需要具有无偏性、单峰性、抗噪性、灵敏度高和计算量小等特性[4-5]。本文首先分析了微小零件显微检测图像的特点，然后针对此类图像存在的特点，从算法基本原理出发，对传统清晰度评价函数的上述性能进行了分析和比较，针对其存在的一些缺陷，提出了一种基于局域方差信息熵的清晰度评价算法，最后通过实验进行了验证。

1 传统清晰度评价函数

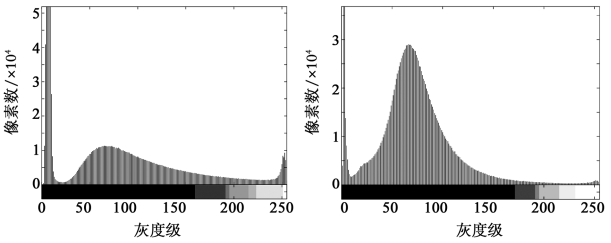
图像的清晰度评价函数研究由来已久，先后出现过许多经典算法，按原理归纳可分为频域函数、梯度函数、统计学函数和信息学函数[6]。图 1 所示为精密加工后微小零件的两种典型显微检测图像，图 1(a)为微铣削加工微阵列后的形貌观测图像，图 1(b)为微细电火花打孔后的孔径测量图像。



(a) 微铣削加工阵列 (b) 微细电火花加工微孔

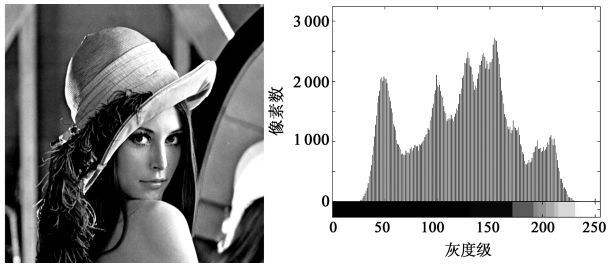
图 1 微小零件显微检测图像

图 2 是图 1 两幅图像的灰度直方图，图 3 为 Lena 标准测试图及其直方图。从图 2～3 中可以看出，与一般可见光图像相比，微小零件显微图像的纹理一般具有相似性，边缘信息太少，因此直方图的像素值比较集中，灰度分布范围有限[8]，图像动态范围较小。针对微小零件显微检测图像的上述特点，对各类清晰度评价函数分别进行分析。



(a) 图 1(a)图像 (b) 图 1(b)图像

图 2 图 1 所示两图像的灰度直方图



(a) Lena 标准测试图像 (b) Lena 图像灰度直方图

图 3 Lena 标准测试图像及其灰度直方图

1.1 频域函数

该类算法原理是图像的低频信息反映图像轮廓，高频信息反映图像边缘和细节，而清晰聚焦的图像，边缘清晰、图像细节丰富，因此高频部分的分量更大。算法形式上主要以傅里叶变换、小波变换等为代表，通过选用某种形式的基函数对图像进行空域到频域的变换，根据图像在频率域上的分布特征来度量图像的清晰度。该类算法准确率高，但计算量过大，不适合微小零件的实时显微检测。

1.2 梯度函数

该类算法利用图像空域特性来计算图像清晰度。其原理是清晰的图像边缘分明，从而相邻像素之间的灰度值相差较大，即图像灰度的梯度大；而模糊的图像边缘不清，相邻像素间灰度值相差较小，图像的灰度梯度小。比较有代表性的有 Brenner 函数、Tenengrad 函数和 Laplacian 函数，这些算法的共同特点都是使用某种空域模板对图像进行空间滤波，区别是梯度计算的方式不同即模板不同。此类函数通常算法简单，计算量小，灵敏度很高。但是由于微小零件显微图像的纹理相似，边缘信息太少，聚离焦过程中图像梯度变化缓慢，因此与一般目标图像相比，其灵敏度有所下降。而且该类函数对噪声比较敏感[6]，尤其是二阶微分算子 Laplacian 函数。

1.3 统计函数

该类算法利用图像统计信息来计算图像清晰

度，该类函数目前数量较少，实际中最常用的为方差函数(Variance)。灰度方差是图像对比度的有用度量[9]，而清晰的图像比模糊的图像对比度更高，灰度级差异更大[10]，因此方差函数可以作为清晰度评价函数，其表达式为

$$F=\sum_{x}\sum_{y}[f(x,y)-\mu]^2 \tag{1}$$

式中，$f(x,y)$为图像(x,y)像素点处的灰度值，μ为一幅$M\times N$大小图像的平均灰度，该评价函数计算量小，能够抗噪声干扰[11]，其缺点是灵敏度很低。

1.4　信息学函数

该类算法认为清晰的图像比模糊的图像包含更多的信息，而信息量可以用图像的信息熵来度量，因此可以用图像的信息熵来度量图像的清晰程度，即清晰聚焦的图像熵值更大。信息熵[12]的概念源于香农信息论，基于信息熵的清晰度评价函数也类似于信息论中熵的定义，其表达式为

$$H=-\sum_{i=0}^{L-1}p_i\log(p_i) \tag{2}$$

其中，p_i为各灰度级出现的概率，L为图像灰度级总数。该评价函数算法简单，计算量较小，但其灵敏度较低。同时由于其仅考虑了灰度值的概率信息，而没有反映图像的空间信息，所提取的特征损失了很多有用的信息，容易造成清晰度误判。这一特点在微小零件显微图像纹理相似、聚离焦中相邻图像灰度分布相似情况下尤为明显。

综上所述，对于微小零件显微检测图像，频域函数耗时过于严重，不适合实时检测场合，因此本文不考虑此类算法。梯度函数在计算局部灰度变化提高灵敏度的同时却增加了对噪声的影响，且与一般目标图像相比灵敏度有所下降。相反方差这种统计函数虽然抗噪性较好，但是缺少对局部灰度信息的描述，因此灵敏度很低。而信息学函数因为丢失了所有的空间信息，有一定应用局限性。针对传统清晰度评价函数存在的上述缺陷，本文提出了一种基于局域方差信息熵的清晰度评价算法，该算法在保证无偏性和单峰性的同时，具备很高的灵敏度和较好的抗噪性。

2　基于局域方差信息熵的清晰度评价函数

由香农信息论可知，每个灰度级包含的自信息量[13]为$H_i=-\log(p_i)$，图像的信息熵为所有灰度级自信息量H_i的加权平均，加权值为各灰度级出现的概率p_i，即信息熵$H=\sum_{i=0}^{L-1}p_iH_i=-\sum_{i=0}^{L-1}p_i\log(p_i)$。假设同一灰度级$i$所对应的像素点位置分别为$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$，定义各灰度级$i$所对应的局域方差为

$$\sigma_i^2=\sum_{j=1}^{n}\sigma^2_{(i,x_j,y_j)}=\sum_{j=1}^{n}\frac{1}{9}\sum_{a=-1}^{1}\sum_{y=-1}^{1}[f(x_j+a,y_j+b)-\mu]^2 \tag{3}$$

其中，$\sigma^2_{(i,x_j,y_j)}$为图像以(x_j,y_j)为中心点 3×3 邻域的局域方差，μ为该邻域的平均灰度，$f(x_j+a,y_j+b)$为图像(x_j+a,y_j+b)像素点处的灰度值。

则定义局域方差信息熵为

$$H_{\mathrm{Var}}=-\sum_{i=0}^{L-1}p_i\sigma_i^2\log(p_i) \tag{4}$$

其中，log 表示取对数，不同对数底数只决定熵的单位，不影响对焦性能，信息论中一般取 2。p_i为各灰度级出现的概率，L为图像灰度级总数，由于图像的灰度分辨率通常为 8 bit[9]，因此通常取$L=2^8=256$。

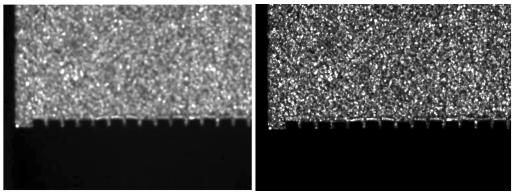
如前文所述，图像信息熵的计算只是基于灰度直方图的统计，缺少像素的局部特征，因此分别对属于同一灰度级的像素进行局域方差计算，计算的局域方差对各灰度级的自信息量进行加权后再求值。与信息熵相比，本文改进的函数不仅考虑了灰度的概率信息，还考虑了图像的空间信息，避免了信息熵容易出现清晰度误判的情况，能更全面地反映整体图像的清晰度。另一方面，与方差原理相同，局域方差可以表征以每个灰度级对应位置为中心邻域的局域清晰度，对该灰度级自信息量加权，可以增加清晰区域像素参与计算信息量时的权重，同时减少背景区域和噪声区域像素的权重，因此局域方差信息熵函数具有较高灵敏度和良好抗噪性。

3　实验结果与分析

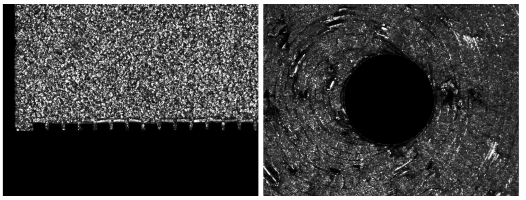
3.1　实际测试结果分析

为了验证本文提出的局域方差信息熵算法的有效性，避免一个目标图像可能存在的偶然性，本文利用超景深显微镜对图 1 精密加工后的两个微小零件，分别在聚焦位置附近采集了 21 幅图像，得到 A、B 两组图像，图 4 为两组图像序列中不同离焦程度的图像。采用上述提到的清晰度评价函数，对实际

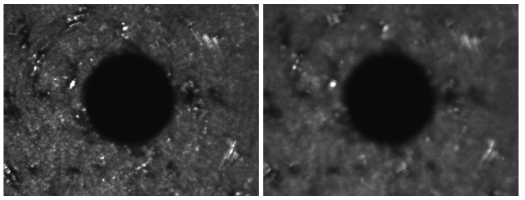
拍摄到的从离焦到聚焦再到离焦的图像序列进行了测试，经归一化后，各函数清晰度评价曲线如图 5 所示。从图中可以看出，各函数清晰度评价结果与前文的理论分析吻合。信息熵函数在评价图 5(b)中第 10 幅图像时发生了清晰度误判，造成第 10 幅和第 11 幅清晰度值相差太小(第 10 幅未归一化清晰度值为 6.830 2，第 11 幅为 6.832 8)，难以准确区分聚焦图像。除信息熵外，其他清晰度评价函数都可准确找到清晰聚焦图像(第 11 幅)，评价函数曲线只有一个最大值点，在最大值点两侧单调下降，均满足无偏性和单峰性。在灵敏度方面，方差和信息熵函数在评价微小零件显微检测图像时，与一般图像相同，最大值两侧曲线平缓，灵敏度很低。梯度函数在评价微小零件显微检测图像时，与一般图像相比，灵敏度有所下降，这一特点在图 5(a)中 Brenner 和 Tenengrad 函数上表现尤为明显。Laplacian 这种二阶微分算子，由于对梯度变化响应幅度强于一阶微分算子，因此灵敏度高于其他梯度函数。而局域方差信息熵函数，清晰度评价曲线与 Laplacian 函数基本重合，具有很高的灵敏度。因此，在一般实际聚焦中，对于微小零件显微检测图像，本文提出的局域方差信息熵函数和 Laplacian 函数的性能优于其他清晰度评价函数。



(a) 深度离焦(A 组第 1 幅) (b) 轻度离焦(A 组第 6 幅)

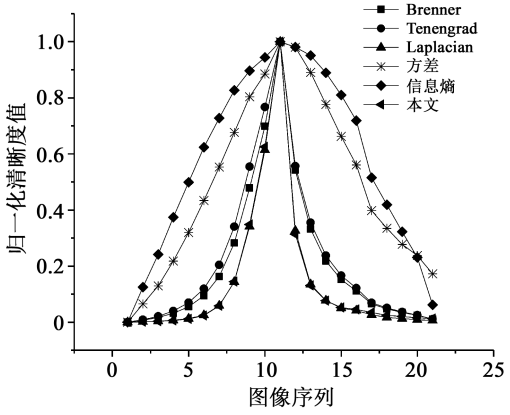


(c) 聚焦(A 组第 11 幅) (d) 正焦(B 组第 11 幅)

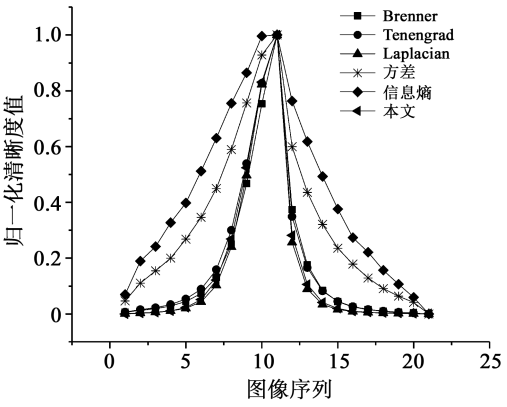


(e) 轻度离焦(B 组第 16 幅) (f) 深度离焦(B 组第 21 幅)

图 4 不同离焦程度的图像



(a) A 组图像



(b) B 组图像

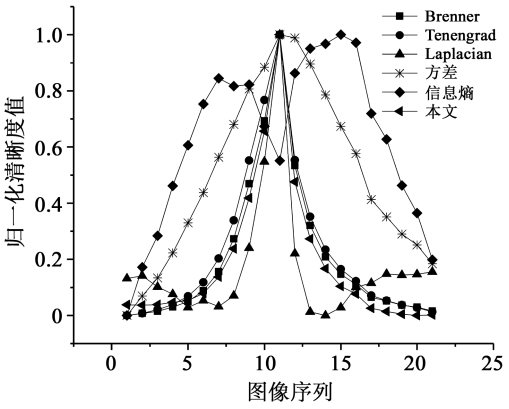
图 5 各函数清晰度评价曲线

3.2 抗噪性能实验仿真

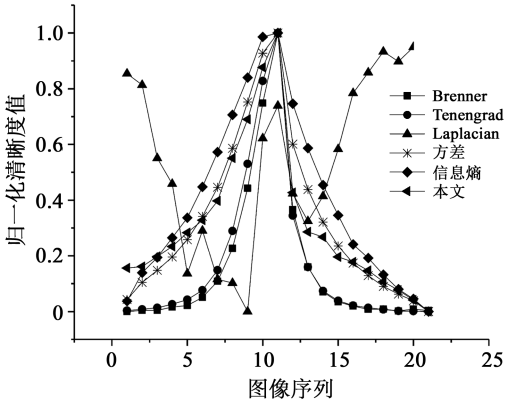
在 CCD 或 CMOS 相机采集图像时，光照水平和传感器温度是影响图像中噪声数量的主要因素[9]。上述实验所用图像序列是在实际中采集的，含有图像获取和传输过程中引入的噪声，所以各函数清晰度评价曲线其实包含对噪声的响应。同时，精密加工后的微小零件显微检测环境一般光照稳定，而且目前大多数工业相机都具备自动曝光和自动滤波功能。因此在一般检测情况下，图 5 曲线可以反映各清晰度评价函数的性能。

但是考虑到极端检测环境下，相机长时间工作传感器温度过高或者光照强度低、不均匀可能会产生高斯噪声[9]，本文在实际拍摄图像序列中又加入了高斯噪声(均值为 0，方差为 0.01)来模拟极端检测环境下可能引入的噪声。图 6 为加入噪声后各函数清晰度评价曲线。从图中不难看出，在极端检测环境下，由于二阶微分在反映剧烈变化方面比一阶微分更为激进，因此 Laplacian 函数对噪声特别敏感，函数曲线出现了局部极值，甚至在 B 组图像中已经失效，抗噪性能极差。而本文提出的局域方差信息熵函数，虽然远焦时函数曲线不够平滑，但是依然能准确找到清晰聚焦图像且保持较高灵敏度，这

表明本文提出的算法抗噪性能良好。



(a)　A 组图像



(b)　B 组图像

图 6　加入噪声后各函数清晰度评价曲线

4　结论

对于微小零件显微检测图像，本文针对梯度函数灵敏度有所下降且对噪声敏感、方差函数能够抗噪声干扰却灵敏度较低、信息熵函数丢失空间信息容易发生清晰度误判等问题，提出了一种基于局域方差信息熵的清晰度评价算法。通过实际测试和抗噪性能仿真可知，在一般实际检测情况下，该算法性能优于其他算法，灵敏度很高。即使在极端检测环境引入噪声的情况下，该算法依然能准确找到清晰聚焦图像且保持较高灵敏度，具有较好的抗噪性。

参考文献：

[1]　袁哲俊，王先逵. 精密和超精密加工技术[M]. 第 3 版. 北京：机械工业出版社，2016.
Yuan Zhejun，Wang Xiankui. Precision and Ultraprecision Machining Technology[M]. 3rd Edi. Beijing：China Machine Press，2016.

[2]　侯秋林. 基于机器视觉的刀具检测技术研究[D]. 济南：山东大学，2018.
Hou Qiulin. Investigation on the technology of tool wear detection based on machine vision[D]. Ji'nan：Shandong University，2018.

[3]　尤玉虎，刘　通，刘佳文. 基于图像处理的自动对焦技术综述[J]. 激光与红外，2013，43(2)：132-136.
You Yuhu，Liu Tong，Liu Jiawen. Survey of the auto-focus methods based on image processing[J]. Laser & Infrared，2013，43(2)：132-136.

[4]　王　健. 基于图像处理的自动调焦技术研究[D]. 北京：中国科学院研究生院，2013.
Wang Jian. Study on technology of auto-focus based on image processing[D]. Beijing：Graduate School of Chinese Academy of Sciences，2013.

[5]　翟永平，周东翔，刘云辉，等. 聚焦函数性能评价指标设计及最优函数选取[J]. 光学学报，2011，31(4)：234-244.
Zhai Yongping，Zhou Dongxiang，Liu Yunhui，et al. Design of evaluation index for auto-focusing function and optimal function selection[J]. Acta Optica Sinica，2011，31(4)：234-244.

[6]　洪裕珍，任国强，孙　健. 离焦模糊图像清晰度评价函数的分析与改进[J]. 光学精密工程，2014，22(12)：3401-3408.
Hong Yuzhen，Ren Guoqiang，Sun Jian. Analysis and improvement on sharpness evaluation function of defocused image[J]. Optics and Precision Engin.，2014，22(12)：3401-3408.

[7]　于之靖，马　凯，王志军，等. 采用加权最小二乘准则的影像仪自动对焦方法[J]. 半导体光电，2017，38(4)：592-597.
Yu Zhijing，Ma Kai，Wang Zhijun，et al. Auto-focusing method of imaging instrument by using weighted least squares criterion[J]. Semiconductor Optoelectronics，2017，38(4)：592-597.

[8]　张俊杰，操晶晶，贡　力，等. 微型零件图像测量清晰度算法的选择与综合评价[J]. 光学精密工程，2008，16(3)：175-182.
Zhang Junjie，Cao Jingjing，Gong Li，et al. Selection and synthesis evaluation of image measurement definition algorithm for miniature parts[J]. Optics and Precision Engin.，2008，16(3)：175-182.

[9]　Gonzalez R C，Woods R E. Digital Image Processing[M]. 3rd Edi.，New York：Prentice-Hall，Inc.，2007.

[10]　李　奇，冯华君，徐之海，等. 数字图象清晰度评价函数研究[J]. 光子学报，2002，31(6)：736-738.
Li Qi，Feng Huajun，Xu Zhihai，et al. Study on the sharpness evaluation function of digital image[J]. Acta Photonica Sinica，2002，31(6)：736-738.

[11]　郑　馨，艾列富，刘　奎，等. 结合全局和局部灰度变化的显微图像自动聚焦函数[J]. 激光与光电子学进展，2017，54(8)：246-253.
Zheng Xin，Ai Liefu，Liu Kui，et al. Auto-focusing function for microscopic images based on global and local gray-scale variation[J]. Laser & Optoelectronics Progress，2017，54(8)：246-253.

(下转第 113 页)

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.023

基于像元纯净指数的多端元提取算法

杨华东[1]，许　楠[2]，郝永平[3]

（1. 沈阳理工大学 信息科学与工程学院，沈阳 110159；2. 沈阳建筑大学 BIM 计算研究中心，沈阳 110168；3. 沈阳理工大学 机械工程学院，沈阳 110159）

摘　要： 针对传统单端元提取方法不能描述端元变异、限制混合像元分解精度的缺点，提出一种基于像元纯净指数的多端元提取算法（Multiple Endmember Extraction Algorithm Based on Pixel Purity Index，PPI-MEE）。首先将图像划分为不重叠的图像块，并分别利用改进的 PPI 算法提取候选端元集，然后利用候选端元的邻域像元光谱信息对候选端元进行优化和精选。最后，对优化精选后的端元集分类得到每类地物的多端元光谱集。仿真数据和真实高光谱数据的实验结果表明，提出的多端元提取策略具有表征遥感图像中端元光谱变异的能力，能够提高端元提取精度和混合像元分解精度。

关键词： 高光谱图像；多端元提取；端元光谱变异；空间信息；像元纯净指数

中图分类号： TP751　**文章编号：** 1001－5868(2020)01－0108－06

Multiple Endmember Extraction Algorithm Based on Pixel Purity Index

YANG Huadong[1], XU Nan[2], HAO Yongping[3]

（1. School of Information Science and Engin., Shenyang Ligong University, Shenyang 110159, CHN;
2. BIM and Computing Technology Research Center, Shenyang Jianzhu University, Shenyang 110168, CHN;
3. School of Mechanical Engin., Shenyang Ligong University, Shenyang 110159, CHN）

Abstract: Traditional endmember extraction method usually ignores the endmember spectral variability, and extracts only one single endmember spectrum for each scene components, which results in poor accuracy of estimated fractions. Thus in this study, a pixel purityindex based multiple endmember extraction approach is proposed. Firstly, the theoriginal hyperspectral image was divided into some non-overlapping image blocks and the improved pixelpurity index algorithm was run on each image block to extract candidate endmembers. Then, the local spatial-spectral information of candidate endmembers is used to refine and select the final endmember bundles from the candidate endmember set. Finally, the endmember set was grouped in terms of metric of spectral similarity. Both simulation and real hyperspectral image data demonstrate that the proposed multiple endmember extraction algorithm is effective to account for endmember variability, and the endmembers extracted by the proposed algorithm are more accurate than the state-of-the-art single endmember extraction algorithms.

Key words: hyperspectral imagery; multiple endmember extraction; endmember spectral variability; spatial information; pixel purity index

收稿日期：2019－11－20.

基金项目：辽宁省自然科学基金指导计划项目（2019-ZD-0259，2019-ZD-0665）.

0　引言

高光谱遥感图像具有光谱分辨率高、成像波段多且连续的特点和优势，使人们能够从遥感图像中获取的信息有了质的转变[1]。然而，由于成像光谱仪空间分辨率偏低、地物复杂多样以及多次散射成像等因素，高光谱图像中多数像元对应地表覆盖类型往往不止一种，形成混合像元。混合像元给地物准确识别和精细分类带来极大困难，严重阻碍了高光谱遥感技术的定量化应用。

高光谱解混（Hyperspectral Unmixing）是解决混合像元问题的有效方法[2]，通常包含两个步骤：端元提取和丰度估计。端元提取是从混合像元中提取表征地物类型的特征光谱，丰度估计是估算混合像元内各端元的组成比例。其中，端元提取的精度尤为关键，直接影响混合像元分解效果。

当前研究中，典型端元提取方法主要包括凸几何分析方法、统计分析方法、稀疏回归分析方法以及光谱-空间联合分析方法[3]。然而，在大多数传统端元提取方法中，一个地物类别仅用一个端元光谱代表，不考虑地物类别内光谱的变化，即类内光谱变异。在实际遥感场景中，受成像环境和地物分布等因素影响，“同物异谱”现象较为普遍[4]。若一类地物仅用一个端元光谱描述，其代表性差，很难准确刻画一个地物类别，导致混合像元分解精度难以提高。为了克服端元光谱变异的影响，研究人员分别从光谱建模、光谱变换、光谱加权、光谱特征选择、迭代混合分析等方面提出了顾及端元变异的端元提取算法[5-6]。其中，Roberts 等[7]提出的多端元光谱混合分析（Multiple Endmember Spectral Mixture Analysis，MESMA）方法是应用最为广泛的一种多端元光谱解混算法。MESMA 算法从光谱库中选择端元光谱，并对混合像元运行所有可能的端元光谱组合，根据最小像元重构误差，选择每个像元的最优端元组合，估计像元中每个端元的丰度。在 MESMA 中，用于光谱混合分析的端元光谱是从光谱库中获得，受测量条件、成像环境等因素影响，这些端元光谱与实际图像的中端元光谱差异较大，因此从图像中提取端元进行光谱混合分析更符合实际解混需求。

针对上述情况，本文提出了一种基于像元纯净指数的多端元提取算法（Multiple Endmember Extraction Algorithm Based on Pixel Purity Index，PPI-MEE）。PPI-MEE 首先改进了经典的像元纯净指数算法[8]，克服其端元提取结果不稳定的缺点。然后采用端元分块提取策略，利用改进的 PPI 算法提取候选端元，并融合光谱信息和空间信息进行端元优化和精选。最后，依据光谱相似性对端元集分类得到每类地物的多个端元光谱。

1　线性混合模型

线性混合模型假设图像中的混合像元光谱是各个端元光谱的线性组合，丰度反映混合像元内各端元的组成比例。线性光谱混合模型的数学表达式为

$$\boldsymbol{r}=\sum_{i=1}^{m}\alpha_i\boldsymbol{e}_i+\varepsilon \tag{1}$$

其中，$\boldsymbol{r}$ 为混合像元向量，$\boldsymbol{e}_i$ 是第 i 个端元向量，α_i 为像元 $\boldsymbol{r}$ 中端元 $\boldsymbol{e}_i$ 的丰度，m 为端元个数，ε 为模型误差项。

基于线性混合模型，L 维像元向量 $\boldsymbol{r}$ 可以看作是高光谱数据特征空间中的一个点，在满足非负与和为一两个约束条件下，全部光谱数据包含在该空间中的一个凸面单体内，而端元恰好位于这个单体的顶点处。因此，端元提取问题就转化为寻求凸面单体顶点的问题。

2　PPI-MEE 算法描述

在实际高光谱场景中，受地势、光照、环境等因素影响，同一地物类型的特征光谱在不同的区域会发生光谱变异，导致“同物异谱”现象，影像空间跨度较大时更是如此。因此，在不同的图像空间区域分别提取端元光谱，为每类地物获取多个端元光谱更能刻画地物类型属性，有效提高端元提取精度。

本文提出的基于像元纯净指数的多端元提取方法包含三个关键步骤：(1)基于改进 PPI 算法的分块端元提取；(2)结合局部空间信息的端元优化与精选；(3)端元分类。

2.1　基于改进 PPI 算法的分块端元提取

PPI 算法利用“单体在空间中的任意投影都是线段，且线段的端点必是单体的顶点”这一性质，在特征空间中随机生成若干向量（“skewers”），然后将像元投影到这些向量上，并记录投影在端点位置的像元，统计每个像元投影落在两端的次数，该次数称为像元纯净指数（PPI），PPI 达到给定阈值即可视为端元。

PPI 算法主要有两个缺点：1）采用随机投影向量导致多次运行后端元提取结果不稳定；2）随机向量个数需要预先设定。为了克服上述缺点，本文提出一种利用图像自身光谱数据生成投影向量的方法，基本步骤如下：

1）计算所有像元向量的均值，$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，$n$ 为像元个数；

2）在特征空间中平移所有像元向量，使得均值向量平移至特征空间的原点处，即 $x_i^*=x_i-\bar{x}$，$i=1,\cdots,n$；

3）单位化平移后的像元向量 $\boldsymbol{x}_i^*$ 得到投影向量，$skewers_i=\frac{1}{\|x_i^*\|}$，$i=1,\cdots,n$。

像元光谱向量经过上述处理得到的向量 $\{skewers_i\}_{i=1}^{n}$ 即为投影向量，其数目与高光谱图像像元数目相同，不需要人工预先设定。另一方面，利用样本自身数据生成投影向量可以尽可能覆盖样本空间的各个方向，提高投影效率。

PPI-MEE 算法首先将原始高光谱图像划分为不重叠的图像子区域（图像块），然后在每个图像块上利用改进的 PPI 算法提取候选端元。本文采用最简单的划分方法，即以固定的行列数目对图像进行划分，并限制每个图像块内的像元总数大于波段数目，一般图像块大小设置为 10×10～30×30。在改进的 PPI 算法中，像元纯净指数阈值 T 设为 0，即选择像元纯净指数大于 0 的像元组成候选端元集 $\{candidates_i\}_{i=1}^{K}$。

2.2 结合局部空间信息的端元优化与精选

PPI 算法是基于光谱特征的端元提取方法，在提取端元的过程中并未考虑图像的空间信息。相关研究表明有效利用图像空间信息可以提高端元提取精度[9]。为此，首先将候选端元分为两类：空间连通候选端元和空间非连通候选端元，示意图见图 1，然后分别进行端元优化和精选处理，得到多端元集 $\{endmembers_i\}_{i=1}^{K'}$。

对于空间连通且光谱相似度高的候选端元，进行光谱均质化处理以实现端元优化。首先，计算它们的光谱均值，并将它们从候选端元集中移除，同时将均值光谱作为新的端元增加到候选端元集中。本文中，选择光谱角距离（Spectral Angle Distance，SAD）作为像元间光谱相似性的判别测度。给定像元 $\boldsymbol{x}$ 与 $\boldsymbol{y}$，SAD 定义如下：

$$SAD(\boldsymbol{x},\boldsymbol{y})=\arccos((\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})/(\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\sqrt{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}})) \tag{2}$$

SAD 的值越小，表明两个像元的距离越近，光谱相似性越高。

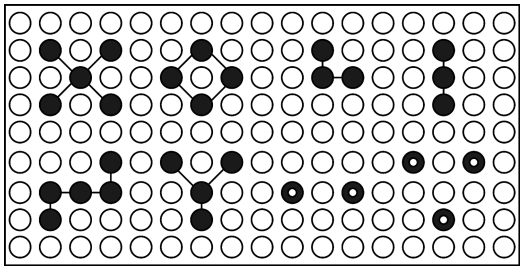


图 1 连通像元（黑色实心圆）和非连通像元（白色空心圆）

PPI 算法对异常值（Outliers）敏感，容易选取异常像元作为候选端元。因此，对于非连通候选端元，通过计算候选端元与其邻域像元的光谱相似性，判断其是否为异常像元，若是异常像元将其从候选端元集中移除。通常，图像中异常像元与其邻域像元的光谱相似性低，特别是在某个/些波段上反射率值差异很大，所以，文中选择利于反映光谱曲线局部特征变化的光谱梯度角进行异常像元判别。

光谱梯度角（Spectral Gradient Angle，SGA）是求解两条光谱向量一阶导数得到两个梯度向量的广义夹角，给定像元向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 和 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$，$\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的梯度向量分别为

$$G(\boldsymbol{x})=[x_2-x_1,x_3-x_2,\cdots,x_n-x_{n-1}]$$

$$G(\boldsymbol{y})=[y_2-y_1,y_3-y_2,\cdots,y_n-y_{n-1}]$$

则像元 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的光谱梯度角大小为

$$SGA(\boldsymbol{x},\boldsymbol{y})=\arccos\left(\frac{G(\boldsymbol{x})^{\mathrm{T}}G(\boldsymbol{y})}{\sqrt{G(\boldsymbol{x})^{\mathrm{T}}G(\boldsymbol{x})}\sqrt{G(\boldsymbol{y})^{\mathrm{T}}G(\boldsymbol{y})}}\right) \tag{3}$$

2.3 端元分类

PPI-MEE 算法对端元集分类前，需要确定分类数目和类别中心，然后根据端元与各类别中心的光谱角距离确定所属类别。本文采用自动目标生成（Automatic Target Generation Process，ATGP）方法[10]产生类别中心，利用 Chang 等人提出的 NWHFC[11]算法确定分类数目。具体端元集分类过程如下：

1）确定分类数目：利用 NWHFC 方法估计分类数目 q。

2）确定类别中心：从端元集中选取具有最大长度的端元作为第一个类别中心 $\boldsymbol{t}_1$；然后利用已有类

别中心构造正交子空间投影算子，并将所有端元投影，选取投影中具有最大长度的端元作为下一个类别中心。选择公式和投影算子如下：

$$\boldsymbol{t}_j = \text{argmax}\left[\left\|\boldsymbol{P}_U^{\perp} endmembers_i\right\|_{\text{F}}^2\right]$$
$$j = 2, \cdots, p, i = 1, \cdots, K' \tag{4}$$
$$\boldsymbol{P}_U^{\perp} = I - \boldsymbol{U}(\boldsymbol{U}^{\text{T}}\boldsymbol{U})^{-1}\boldsymbol{U}^{\text{T}} \tag{5}$$

其中，$\boldsymbol{U} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots \boldsymbol{t}_{j-1}]$，$\boldsymbol{U}^{\text{T}}$ 和 $\boldsymbol{U}^{-1}$ 分别是 $\boldsymbol{U}$ 的转置矩阵和逆矩阵。

3）端元集分组：依次计算每个端元与各个类别中心的 SAD，把每个端元分配给距离它最近的类别中心。

2.4　PPI-MEE 算法流程

PPI-MEE 算法流程如图 2 所示。

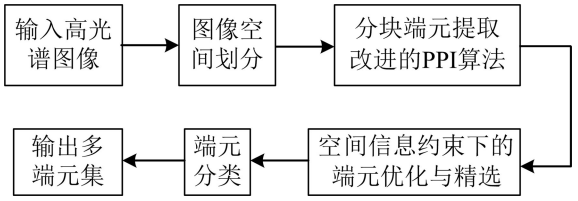


图 2　PPI-MEE 算法流程

3　算法验证与分析

为了验证提出的 PPI-MEE 算法的正确性和有效性，本文中用合成高光谱数据和真实高光谱图像数据进行实验验证。选择经典的基于正交投影的 PPI 算法和基于单体体积分析 的 NFINDR 算法[12] 作为对比算法。

为了评价端元提取和丰度估计精度，分别采用 SAD 和均方根误差（Root Mean Square Error，RMSE）两个指标定量比较。RMSE 计算公式如下：

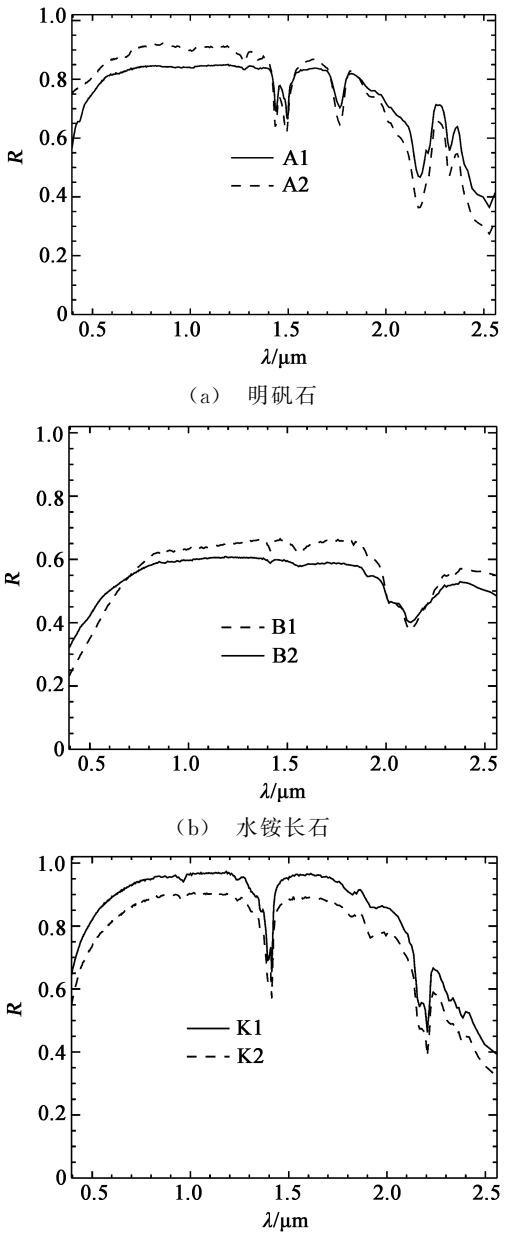
$$RMSE_i = \sqrt{\frac{1}{N}\left\|A_i - \widetilde{A}_i\right\|_2^2} \tag{6}$$

式中，A_i 为端元对应的真实丰度，$\widetilde{A}_i$ 为解混丰度。

实验结果中，SAD 值越小，表明提取的端元光谱越接近真实端元光谱，端元提取精度越高；RMSE 值越小，表明混合像元解混丰度越接近真实丰度，混合像元分解精度越高。

3.1　模拟高光谱数据实验

模拟高光谱数据中的端元光谱从美国地质调查局（USGS）光谱库中选取，分别选择明矾石（Alunite）、水铵长石（Buddingtonite）和高岭石（Kaolinite）3 类矿物，每类矿物包含 2 条光谱曲线，共计 6 种端元光谱，这 6 种端元分别用 A1，A2，B1，B2，K1 和 K2 表示。图 3 为所选 6 种端元的光谱。



（a）　明矾石

（b）　水铵长石

（c）　高岭石

图 3　用于生成模拟数据的 6 种端元光谱

模拟图像的空间大小为 120×90，波段数目为 430，波段范围为 0.401 1～2.976 μm，图 4 是模拟图像示意图，生成方式如下：将图像分为 12 个大小为 30×30 的图像块，每个图像块内的像元由两种端元光谱混合而成，混合比例为 1∶1，从左至右，自上而下端元光谱组成情况依次为（A1，B1），（A1，B2），（A2，B1），（A2，B2），（A1，K1），（A1，K2），（A2，K1），（A2，K2），（B1，K1），（B1，K2），（B2，K1），（B2，K2）。为了模拟纯像元，分别用大小为 10×10 的纯像元块替换上述 12 个图像块的中心区域（图中

阴影区域),纯像元块的组成依次为 K1,K2,K1,K2,B1,B2,B1,B2,A1,A2,A1,A2。

为使模拟图像更接近真实高光谱图像,对模拟图像先进行模板大小为 9×9 的均值滤波处理,然后附加信噪比(SNR)为 30 dB 的高斯白噪声。

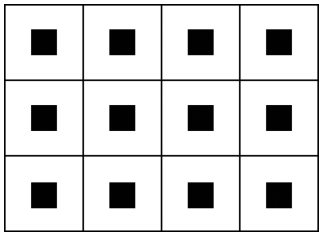


图 4 模拟图像示意图

表 1 是三种算法提取的端元光谱与真实端元光谱的平均 SAD。PPI 和 NFINDR 算法的单次端元提取结果都不稳定,故实验中分别运行 20 次 PPI 和 NFINDR 算法,将它们的端元提取结果与 PPI-MEE 的结果比较。总体来看,PPI-MEE 算法运行一次即可为每类地物提取多个代表端元,而且 PPI-MEE 算法的平均 SAD 值最小,精度最高。

表 1 三种算法端元光谱与真实端元光谱的平均 SAD

端元	算法		
	PPI	NFINDR	PPI-MEE
A1	0.182 4	0.190 3	0.177 9
A2	0.176 5	0.190 3	0.170 1
B1	0.165 7	0.163 4	0.156 3
B2	0.170 1	0.169 2	0.156 8
K1	0.149 3	0.148 8	0.120 6
K2	0.143 5	0.145 5	0.139 3

本部分还进行了异常像元仿真实验。图像中异常像元光谱生成方式如下:随机从 6 种端元光谱中选择 1 种,然后随机将其 430 个波段中的 10 个波段的反射率值置为 0.9,10 个波段的反射率值置为 0.1,得到异常像元光谱。将该异常像元光谱随机替换仿真图像中的某个像元,得到模拟图像。实验结果表明:当模拟图像中存在异常像元时,多次运行 PPI 和 NFINDR 算法后,都会发生误选异常像元作为端元光谱的情况,而本文提出的 PPI-MEE 算法,因提取过程中利用了局部空间像元光谱相似性约束进行端元精选,可以剔除图像中的异常像元,避免了误选异常像元作为最终端元的可能。

为了进一步验证端元提取结果对混合像元分解精度的影响,分别利用 PPI、NFINDR 和 PPI-MEE 算法提取的端元进行丰度估计。对于 PPI 和 NFINDR 两种单端元提取算法提取的端元,利用经典的全约束最小二乘估计法(Fully Constrained Least Squares, FCLS)[13] 估计丰度。对于 PPI-MEE 算法提取的端元,利用经典的 MESMA 方法估计丰度。表 2 是丰度估计的平均 RMSE 对比结果。从表中数据可以看出,对模拟高光谱数据进行多端元提取和多端元混合光谱分解,其光谱解混精度要优于传统单端元光谱解混。

表 2 三种算法真实丰度与估计丰度的平均 RMSE

算法	端元		
	A1, A2	B1, B2	K1, K2
PPI	0.023 5	0.014 3	0.013 4
NFINDR	0.021 6	0.012 5	0.015 5
PPI-MEE	0.019 7	0.012 7	0.012 0

3.2 真实高光谱数据实验

实验中选取的真实高光谱影像是由机载可见光/红外成像光谱仪(AVIRIS)于 1997 年 6 月获取的美国内华达州南部 Cuprite 矿区的高光谱图像数据。实验图像大小为 200×200,去除水吸收和低信噪比波段剩余 185 个光谱波段,空间分辨率为 20 m,光谱分辨率为 10 nm,光谱覆盖范围为 0.4~2.5 μm,见图 5。

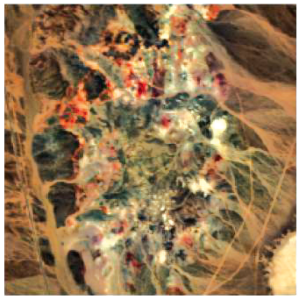


图 5 真实高光谱图像

利用 NWHFC 方法估计的端元数目为 14,参考 Clark 和 Swayze 等[14] 在该地区的实地调查和矿物填图结果,实验中选取该地区内裸露状况较好且发生光谱变异较多的5种典型地物类型进行研究。

表 3 Cuprite 数据实验结果

端元	算法		
	PPI	NFINDR	PPI-MEE
alunite#1	0.115 3	0.125 9	0.087 1
alunite#2	0.075 6	0.125 9	0.065 5
kaolinite#1	0.097 8	0.085 7	0.088 4
kaolinite#2	0.084 7	0.085 7	0.079 9
kaolinite#3	0.049 1	0.054 6	0.044 2
buddingtonite#1	0.106 2	0.087 7	0.097 4
buddingtonite#2	0.106 2	0.090 4	0.069 7
montmorillonite	0.068 8	0.082 0	0.075 6
chalcedony	0.140 3	0.113 4	0.128 9

PPI、NFINDR 和 PPI-MEE 算法提取的端元光谱与 USGS 光谱库中相应矿物光谱的 SAD 结果见表 3。从表中可以看出，PPI-MEE 提取的端元光谱和光谱库中的矿物光谱总体上最为相似。

4　结论

本文针对高光谱遥感图像端元提取中的光谱变异问题，提出一种基于像元纯净指数的多端元提取算法，融合图像区域划分、单端元提取、端元优化与精选和端元分类，建立了高光谱图像多端元提取框架，该框架具有灵活易扩展的特性。模拟数据和真实高光谱图像数据实验结果验证了本文提出的 PPI-MEE 算法在处理光谱变异问题时的有效性。现阶段 PPI-MEE 适用于存在纯像元的高光谱图像，下一步将继续研究适用于数据混合度较高的高光谱多端元提取方法，同时设计更有效的空间信息分析策略，进一步提高算法精度。

参考文献：

[1] 童庆禧，张　兵，郑兰芬. 高光谱遥感：原理、技术与应用[M]. 北京：高等教育出版社，2006.
Tong Qingxi，Zhang Bing，Zheng Lanfen. Hyperspectral Remote Sensing[M]. Beijing：Higher Education Press，2006.

[2] Jose M B，Plaza A，Dobigeon N，et al. Hyperspectral unmixing overview：geometrical，statistical and sparse regression-based approaches[J]. IEEE J. of Sel. Top. in Appl. Earth Observations and Remote Sensing，2012，5(2)：354-379.

[3] 童庆禧，张　兵，张立福. 中国高光谱遥感的前沿进展[J]. 遥感学报，2016，20(5)：689-707.
Tong Qingxi，Zhang Bing，Zhang Lifu. Current progress of hyperspectral remote sensing in China[J]. J. of Remote Sensing，2016，20(5)：689-707.

[4] 蓝金辉，邹金霖，郝彦爽，等. 高光谱遥感影像混合像元分解研究进展[J]. 遥感学报，2018，22(1)：17-31.
Lan Jinhui，Zou Jinlin，Hao Yanshuang，et al. Research progress on unmixing of hyperspectral remote sensing imagery[J]. J. of Remote Sensing，2018，22(1)：17-31.

[5] Ben Somers，Gregory P，Asner P C，et al. Endmember variability in spectral mixture analysis：A review[J]. Remote Sensing of Environment，2011，115(7)：1603-1616.

[6] Shao Y，Lan J H. A spectral unmixing method by maximum margin criterion and derivative weights to address spectral variability in hyperspectral imagery[J]. Remote Sensing，2019，11(1045)：2-29.

[7] Roberts D A，Gardner M，Church R，et al. Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models[J]. Remote Sensing of Environment，1998，65(3)：267-279.

[8] Boardman J W，Kruse F A，Green R O. Mapping target signatures via partial unmixing of AVIRIS data[C]// Proc. of 5th Annual JPL Airborne Geoscience Workshop，1995：23-26.

[9] Martin G，Plaza A. Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data[J]. IEEE J. of Sel. Top. in Appl. Earth Observations and Remote Sensing，2012，5(2)：380-395.

[10] Ren Hsuan，Chang Chein-I. Automatic spectral target recognition in hyperspectral imagery[J]. IEEE Trans. on Aerospace and Electronic Systems，2003，39(4)：1232-1249.

[11] Chang Chein-I，Du Qian. Estimation of number of spectrally distinct signal sources in hyperspectral imagery[J]. IEEE Trans. on Geoscience and Remote Sensing，2004，42(3)：608-619.

[12] Winter M E，N-FINDR：an algorithm for fast autonomous spectral end-member determination in hyperspectral data[C]// Proc. of the SPIE Conf. on Imaging Spectrometry V，Denver，Colorado，USA，1999，3753：266-275.

[13] Heinz Daniel C，Chang Chein-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery[J]. IEEE Trans. on Geoscience and Remote Sensing，2001，39(3)：529-545.

[14] Clark Roger N，Swayze Gregg Alan，Livo K Eric，et al. Imaging spectroscopy：earth and planetary remote sensing with the USGS tetracorder and expert systems[J]. J. of Geophysical Research，2003，108(E12)：5131.

作者简介：

杨华东(1986－)，男，博士，讲师，主要研究方向为高光谱图像处理、智能信息处理。

E-mail：yanghd8608@163.com

(上接第 107 页)

[12] Shannon C E. A mathematical theory of communication[J]. Bell Labs Technical J.，1948，27(4)：379-423.

[13] 曹雪虹，张宗橙. 信息论与编码[M]. 第 2 版. 北京：清华大学出版社，2009.
Cao Xuehong，Zhang Zongcheng. Information Theory and Coding[M]. 2nd Edi. Beijing：Tsinghua University Press，2009.

作者简介：

于占江(1985－)，男，内蒙古凉城人，博士，讲师，研究方向为微纳制造与图像检测。

E-mail：13514310244@163.com

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.024

应用于微型扑翼飞行器的实时视觉系统设计

王昕鹏[1,2]，张卫平[1]，牟家旺[1,2]，陈子豪[1,2]

（上海交通大学 1. 微米/纳米加工技术国家级重点实验室；2. 电子信息与电气工程学院 微纳电子学系，上海 200240）

摘　要： 提出并设计了一套应用于小尺寸、高频抖动的微型仿生扑翼飞行器的实时视觉系统，包括构成硬件基础的微型摄像头模块、5.8 GHz 传输模块和无线图传接收模块，以及构成软件模块的两种算法，其中一种是利用角点检测与 LK 光流法、基于卡尔曼低通混合滤波的实时电子稳像去抖算法，另一种是基于 YOLOv3 的人像识别算法。进行了微型仿生扑翼飞行器的飞行硬件搭载测试实验，实验结果表明，所提出的电子稳像算法可以去除 x 轴 66.56%、y 轴 73.15%的抖动，人像识别算法可以达到 96%以上的准确率，同时图传分辨率在 720×480 像素时，稳像去抖算法时延为 6.33 ms，人像识别算法时延为 20.5 ms，总时延满足实时要求，完成了实时去抖与人像识别系统的设计目标。

关键词： 微型扑翼飞行器；实时电子稳像；卡尔曼滤波；实时人像识别；YOLOv3

中图分类号： V279　**文章编号：** 1001－5868(2020)01－0114－04

Design of Real-time Vision System Applied in Flapping-wing Micro Air Vehicle

WANG Xinpeng[1,2], ZHANG Weiping[1], MOU Jiawang[1,2], CHEN Zihao[1,2]

(1. National Key Lab. of Science and Technol. on Micro/Nano Fabrication;
2. Dept. of Micro/Nano Electronics, Shanghai Jiaotong University, Shanghai 200240, CHN)

Abstract: A real-time vision systemis designed for flapping-wing micro air vehicles (FMAV) withlightweight and high-frequency vibration. The system includes the hardware and software modules. The hardware is consists of a micro camera module, a 5.8 GHz transmission module and a wireless image transmission receiving module. And the software modules include two algorithms, one is a real-time video stabilization algorithm based on corner detection, Lucas Kanade optical flow and Kalman low-pass hybrid filtering, and the other one is a person recognition algorithm based on YOLOv3. Flight tests on FMAV with hardware were carried out. The experimental results indicate the video stabilization algorithm can diminish 66.56% vibration of the x-axis and 73.15% vibration of the y-axis. The person recognition algorithm can achieve an accuracy of higher than 96%. The video stabilization algorithm costs 6.33 ms and person recognition algorithm costs 20.5 ms when the image resolution is 720×480 pixels, thusthe total delay meets the real-time requirements. In conclusion, the design goal of real-time vision system has been achieved.

Key words: FMAV; real-time video stabilization; Kalman filtering; real-time person recognition; YOLOv3

收稿日期：2019－11－04.

基金项目： 教育部基金项目（6141A02022607，6141A02022627）；预研基金项目（1816311ZT005020，301020803，17070107）；上海市科委项目（19511104202，19DZ2291103）.

0　引言

近年来，国内外科研人员对飞行器微小型化开展了大量研究[1-3]，其中模仿昆虫飞行方式的扑翼式飞行器已经实现克级别的带载飞行[3]。扑翼飞行方式不仅可以产生稳定且充足的气动力以实现悬停，在机动性和功耗方面相较固定翼与旋翼也有更多的优势，因此近 10 年来微型仿生扑翼式飞行器(Flapping-wing Micro Air Vehicle，FMAV)在军事侦察、指定目标精确打击、未知环境探测、灾后搜救和电子干扰等军事及民用领域中发挥重要的作用。

然而，由于 FMAV 自重轻、带载小，以及扑翼飞行不可避免的高频振动等自身属性，FMAV 在图像视觉领域一直面临较大挑战：(1)由于微型无人机克服自身重力起飞已经较为困难，因此其携带的摄像头大小及重量受到较大限制；(2)由于扑翼飞行伴随着机身的高频抖动，其捕获视频帧序列存在严重的抖动干扰，缺乏处理扑翼高频抖动的电子稳像降噪算法；(3)由于自身机体条件限制，目前尚无实时目标检测算法在 FMAV 上应用。

本文以上海交通大学研发的仿昆虫类 FMAV 为载体，在其基础上开发实时电子稳像去抖及人像识别系统，完成实时图像传输、视频序列降噪及高速率高准确度的人像识别算法研究。

1　实时图像识别系统设计

仿昆虫类 FMAV 通过 1/4 inch CMOS 微型摄像头模块按帧采集视频序列，经由 5.8 GHz 频段实时传输至地面站，经过实时视频稳像去抖算法处理，再使用基于 YOLOv3 的实时人像识别算法确定目标框及目标种类，完成实时视觉系统的全部流程，如图 1 所示。

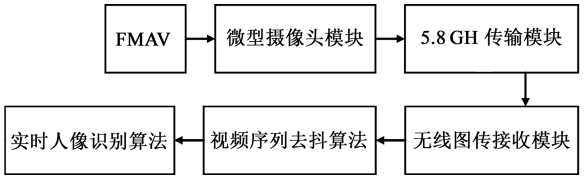


图 1　FMAV 实时视觉系统流程

仿昆类 FMAV 实时图像识别系统整体可分为硬件基础与软件算法两大部分，硬件部分包括微型摄像头模块、5.8 GHz 传输模块和无线图传接收模块；软件部分包括视频序列实时稳像去抖算法和实时人像识别算法。

2　软件算法设计

接收的视频帧序列需经过稳像去抖算法，以去除因机身震颤带来的视频抖动，之后利用识别算法按帧处理视频序列，最后重组视频序列输出。

2.1　实时电子稳像算法

本文提出一种改进的基于卡尔曼与低通混合滤波的实时电子稳像降噪算法。该算法的流程图如图 2 所示，共包括四个步骤。首先，利用 shi-tomasi 角点检测算法[4]提取角点，并使用已提取的角点计算两个连续帧之间的光流；接下来，用仿射变换模型，代入经光流计算后保留的角点，计算运动参数；随后，利用由卡尔曼和低通滤波器组成的混合滤波器平滑运动参数；最后，用平滑后的参数对帧做仿射或刚度变换，以获得稳定的视频帧序列。其中，角点检测与匹配算法和 LK 光流法[5]计算影响整体电子稳像的实时性，而运动估计的计算结果很大程度上影响后续的去抖效果。

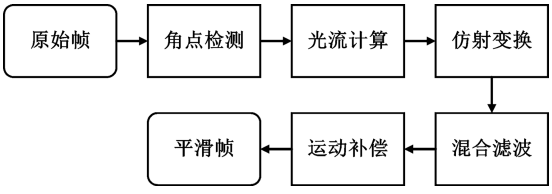


图 2　实时电子稳像算法流程

2.1.1　角点筛选

考虑到电子稳像的实时性要求，本文采用 shi-tomasi 角点检测算法在输入的帧图像中检测角点，并通过最小化定义如下的函数来计算两个连续帧之间像素点的运动：

$$\Delta(v)=\Delta(v_x,v_y)=\sum_{x}^{p_x+n_x}\sum_{y}^{p_y+n_y}(P(x,y)-N(x+v_x,y+v_y))^2 \tag{1}$$

其中，(p_x,p_y)表示第一帧图像 P 上的像素点，(n_x,n_y)表示第二帧图像 N 上的像素点，(v_x,v_y)表示在图像上的移动速度，(x,y)表示在对应帧图像上的像素坐标点。本文采用 Lucas-Kanade 光流算法对上述函数最小化，从而在连续的第二帧图像 N 上可以找到对应的角点，用于剔除错误角点，并筛选出清晰准确可用于计算变换的有效角点。

2.1.2　仿射变换模型

利用 2.1.1 节筛选出的角点计算两帧间的运动参数。根据微型扑翼式飞行器的飞行特性，两帧间的运动变换可以等效为具有 4 个未知参数的二维仿

射变换模型，由 x 轴与 y 轴上的平移、缩放与旋转组合而成，具体表述如下式：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \boldsymbol{M} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} S\cos\alpha & -S\sin\alpha & T_x \\ S\sin\alpha & S\cos\alpha & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2}$$

其中，(x,y)和(x',y')分别表示第一帧与第二帧图像上的像素坐标点，S 表示缩放系数，α 表示旋转角度，(T_x,T_y)表示在 x 与 y 轴上的平移。利用相邻两帧的有效角点即可求得抖动时的仿射变换矩阵 $\boldsymbol{M}$。

对于连续帧 $\boldsymbol{A}_i$ $(i=1,2,3,\cdots,n)$，上述公式可以用于计算 $\boldsymbol{A}_i$ 和 $\boldsymbol{A}_{i+1}$ 之间的运动参数。为去除扑翼微飞行器高频快速抖动的同时，能保留其低频的飞行轨迹，本文将 $\boldsymbol{A}_i$ 帧前的仿射变换累乘取逆作为累积的运动参数，同时将当前帧的运动参数进行混合滤波得到平滑的仿射变换矩阵，应用式(3)即可获得稳定的 $\boldsymbol{A}_i'$帧图像：

$$\boldsymbol{A'}_i = \boldsymbol{M}_{\text{filter}} \left(\prod \boldsymbol{M}_i \right)^{-1} \boldsymbol{A}_i \tag{3}$$

2.1.3 混合滤波

考虑到算法对实时性要求较高，无法采用全视频序列滤波稳像，同时由于抖动视频帧序列中不仅存在高斯白噪声，仅仅采用卡尔曼滤波稳像后，视频可能仍然有抖动。本文提出采用卡尔曼滤波与低通滤波混合的方式去除抖动，即利用卡尔曼滤波除去高斯白噪声，再将获得的运动参数传入对高斯分布规律采样的低通滤波器，以除去高斯色噪声，获得平滑的运动参数。混合滤波的架构如图 3 所示。

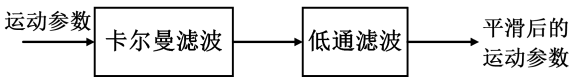


图 3 混合滤波架构

2.2 实时人像识别算法

视频帧序列经过实时电子稳像算法后按帧送入人像识别算法处理，检测每帧图像中是否存在人像，并给出人像位置信息。考虑到算法的实时性要求，本文选择基于 YOLOv3 的目标检测算法[6]，并自构建数据集用于模型训练，将 YOLOv3 的骨干网络 Darknet-53 进行迁移，用于提取人脸特征，利用残差网络解决梯度减小问题，同时精简卷积层，使模型推理速度进一步加快。

2.2.1 自构建数据集

相比目前广泛使用的目标检测数据集提供的 RGB 图像，微型扑翼飞行器搭载的 1/4 inch (1 inch=2.54 cm) CMOS 微型摄像头模块捕获的图像往往存在清晰度不足、RGB 三通道缺失的劣势，同时受限于微型扑翼飞行器的飞行状态，每帧图像常常会有遮挡、模糊等干扰因素，因此利用现有数据集训练所得模型无法准确且实时地识别人像。本文利用微型扑翼飞行器实地飞行传回的图像，结合 VOC2012、COCO 两类常用数据集，自构建微扑翼飞行器人像数据集。

训练集制作过程采用的步骤如下：1) 采用 1/4 inch CMOS 微型摄像头模块采集包括遮挡的不同人像多个角度的 RGB 图像，图片分辨率为 720×480；2) 分别将每幅图片进行复制并按 1/2，1/1.5，1/1.2 比例调整它们的分辨率，增加三倍数据集；3) 随机抽样第三步所得图片的 20%加入噪声，模拟图像传输噪点过多的实际情况；4) 将获得的每幅图片进行人工标记，注释人像类别信息与位置信息，生成人像标签文件；5) 将每幅用于训练的人像图片和对应的标签文件组合作为训练集。

2.2.2 生成人像 Darknet-53 模型

YOLOv3 采用了模型作者设计训练的 Darknet-53 基础特征提取器，该网络结构在保持精度不下降的情况下，减少了网络层数，提高了计算速度。在 Darknet-53 中，模型采用全卷积网络，改进并结合 Darknet-19 和深度残差网络，使用连续的 3×3 与 1×1 的卷积层与残差层，用于提取训练图像特征的有效基础特征提取模型。

参考对迁移学习的研究[7]发现，若将源域训练好的模型参数迁移至新的模型中进行训练，将有助于提高新模型的训练效果。本文基于 YOLOv3 中 COCO 数据集下训练好的 Darnket-53 网络参数，将其迁移至新的基础网络中，初始化其参数，过程如下步骤所示：1) 从公开数据集中选取 20 组人像样本，每组包含 100 个样本，来自不同的 10 个人，一共 2 000 张人像图片；2) 将第一步选取的 2 000 张人像图片分别进行复制并按 1/2、1/1.5、1/1.2 比例调整它们的分辨率，将数据集扩大为 8 000 张人像图片；3) 将 8 000 张人像图片经过 Darknet-53 预训练，获得训练好的网络参数，并迁移到新的基础网络中初始化，形成人像 Darknet-53 模型。

本文采用自构建数据集与人像 Darknet-53 模型进行特征提取，形成基于 YOLOv3 的实时人像检测模型。

3　实时图像识别系统搭建与实验

3.1　硬件系统搭建

本文将硬件系统搭载在上海交通大学研发的仿昆虫类 FMAV 进行算法验证测试，该仿昆虫类 FMAV 重量仅为 40 g，目前可以实现垂直起落，遥控飞行等功能。地面站为搭载 NVIDIA GTX2080Ti 显卡的主机，运行在 Linux 环境下，能够满足本系统的实时图像处理要求。

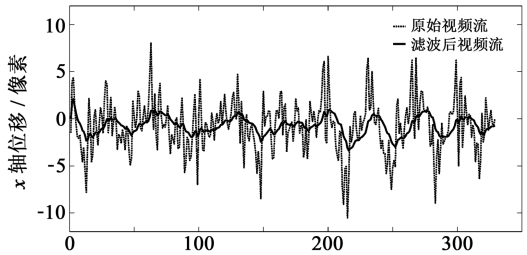
3.2　实时电子稳像算法实验

本文所述的实时电子稳像算法利用 OpenCV 模块实现。FMAV 捕获 720×480 像素分辨率的实验视频，通过 5.8 GHz 图传至算法程序。本文设置最大角点数为 200，质量水平系数为 0.01，最小距离为 30 像素。角点检测结果如图 4 所示。

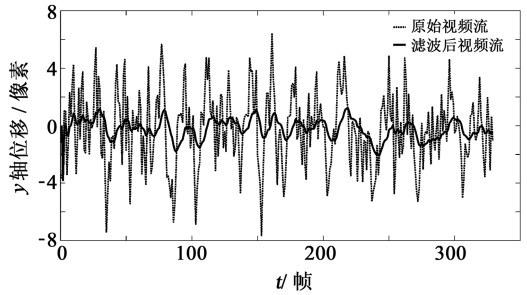
(a)　第一帧角点　　(b)　第二帧角点

图 4　角点检测结果

电子稳像测试在一个共计 329 帧的视频序列上进行，通过本文的算法进行实时运动滤波后的 x 与 y 轴位移如图 5 所示，可以明显看出经本文滤波算法后的两轴位移量比原始视频中的位移量振幅更小，频率更低，表明经实时电子稳像算法后的两轴位移更加平滑。为更好地说明本文算法的去抖效果，采用 x 与 y 轴去抖前后的标准差对比，表 1 表明本文提出的实时电子稳像算法将 x 轴的抖动减小了 66.56%，将 y 轴的抖动减小了 73.15%。同时本算法的处理速度在 158 f/s，完全达到实时性要求。



(a)　x 轴稳像前后对比



(b)　y 轴稳像前后对比

图 5　x 与 y 轴稳像效果

表 1　x 与 y 轴去抖前后的位移标准差

参数	原始视频流	去抖后视频流
x 轴位移标准差/像素	2.809 2	0.939 3
y 轴位移标准差/像素	2.573 0	0.690 9

3.3　人像识别算法实验

本文所述的人像识别算法基于 CUDA 及 GPU 运算实现，将经实时电子稳像算法处理后的视频帧送入训练后的迁移 Darknet-53 模型进行推理，利用阈值过滤无效的目标检测后，经非极大值抑制找出最大概率的人像输出。不同场景下的识别效果如图 6 所示。

即使实验在单自由度的快速偏航、俯仰和翻滚测试中进行，视频序列实时识别准确率依旧保持在 96%以上。同时人像识别算法的速度在 49 f/s，结合稳像去抖算法的总时延在 30 ms 以内，高于技术规范中规定的 30 f/s 实时性要求。

(a)　遮挡人像识别　(b)　快速偏航下人像识别

图 6　不同场景下人像识别效果

4　总结

本文为仿昆类微型扑翼飞行器设计了一套包括软硬件的、实时性高的稳像去抖与人像识别系统。实测数据表明，本文设计的硬件系统运行稳定，电子稳像算法可去除 70%左右的高频抖动噪声，人像识别算法准确率保持在 96%以上，同时两个算法的总时延在 30 ms 以内，达到实时性标准。

参考文献：

[1]　陈文元，张卫平. 微型扑翼式仿生飞行器[M]. 上海：上海交通大学出版社，2011：35-50.
Chen Wenyuan，Zhang Weiping. Bio-Inspired Flapping-Wing Micro Air Vehicle[M]. Shanghai：Shanghai Jiaotong University Press，2011：35-50.

（下转第 122 页）

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.025

丙丁烷 TDLAS 测量系统的吸收峰自动检测

张　铭，苗玉彬，张　舒

（上海交通大学 机械与动力工程学院，上海 200240）

摘　要： 基于丙丁烷在 3.37 μm 处的基频吸收峰，使用相应中心波长的连续带间级联激光器和长光程吸收池，研制了基于波长调制技术的丙丁烷气体检测系统。为自动确定波长调制过程中的锯齿波中心值参数，基于电流与输出光频率的线性关系，通过均匀增加激光器驱动电流获得直接吸收光谱。使用基于洛伦兹线型的拟合算法确定吸收峰值对应电流，使锯齿波输出光以其为中心覆盖吸收峰。实验证明，以拟合结果 31.94 mA 为中心值的锯齿波，在叠加频率为 20 000 Hz、幅值为 0.03 V 的高频正弦调制波后，系统二次谐波峰值超过 100 mV。算法全过程不依赖标气和 F-P 标准具等精密光学器件，可适应实际生产条件下的系统自动定参。

关键词： 带间级联激光器；可调谐半导体激光吸收光谱；气雾剂检漏；中红外吸收峰；洛伦兹光谱线型

中图分类号： O433.5　**文章编号：** 1001－5868(2020)01－0118－05

Automatic Absorption Peak Determination in LPG Detection System Based on TDLAS

ZHANG Ming, MIAO Yubin, ZHANG Shu

(School of Mechanical Engin., Shanghai Jiaotong University, Shanghai 200240, CHN)

Abstract: Based on the fundamental frequency absorption peak at 3.37 μm of propane butane, a propane butane gas detection system based on wavelength modulation technology was developed by using a continuous interband cascade laser (ICL) with corresponding center wavelength and an absorption cell with a long optical path. In order to automatically determine the center value of sawtooth wave in the wavelength modulation process, based on the linear relationship between current and output frequency, the direct absorption spectrum was obtained by uniformly increasing the driving current of the laser. The Lorenz profile fitting algorithm was used to determine the current corresponding to the absorption peak so that the sawtooth output light can cover the absorption peak with its center. The experimental results show the sawtooth wave with the fitting result of 31.94 mA as the central value, after the high-frequency sinusoidal modulated wave with the superposition frequency of 20 000 Hz and amplitude of 0.03 V, the second harmonic peak exceeds 100 mV. The whole process of the algorithm does not depend on precise optical devices such as gas and F-P etalon.

Key words: interband cascade lasers; tunable diode laser absorption spectroscopy; aerosol leak detection system; mid-infrared absorption peak; Lorentz profile

0　引言

丙丁烷是丙烷和丁烷的混合物，作为民用燃料、气雾推进剂、制冷剂等在化工业得到广泛应用。丙丁烷浓度达到其爆炸下限 1.87%[1] 时，存在爆炸危险。因此对气雾剂灌装产线的泄漏浓度进行实时监测具有重大意义。

收稿日期：2019－10－31.

基金项目：国家自然科学基金项目(51975361).

可调谐半导体激光吸收光谱技术(Tunable Diode Laser Absorption Spectroscopy,TDLAS)基于分子的特征吸收光谱检测气体浓度,具有响应速度快、灵敏度高等优点。然而由于制造工艺限制,3.3 μm 以上波段的半导体激光器较难实现[2],而丙丁烷最显著的吸收峰正位于中红外波段的 3.37 μm。随着半导体激光器行业近 20 年的发展,带间级联激光器(Interband Cascade Laser,ICL)已在 2.7～5.9 μm 范围内实现了连续波模式工作[3],如 2012 年的火星登陆项目中 ICL 被用于甲烷及其同位素测量[4]。

20 世纪末以来,许多研究机构开始将中红外激光器用于气体检测。Namjou 等[5]使用 DFB 量子级联激光器首次完成了室温条件下的甲烷与 N_2O 测量,且检测下限达到 5×10^{-5},但灵敏度受到激光器调制幅值的限制。针对恶劣环境下气体浓度的测量问题,Rieker 等[6]提出了无标定的波长调制方法,使用一次谐波信号(1f)归一化二次谐波信号(2f)以消除激光器初始光强项的影响,与直接吸收技术相比获得了 4 倍的信噪比增益。Du 等[7]基于分布式带间级联激光器研制了 CH_3SH 气体检测系统,选用了 2 945～2 953 cm^{-1} 波长区间的 CH_3SH 吸收峰,最终结果可作为大气遥感领域的参考数据。Dong 等[8]开发了两种基于不同光学结构的紧凑传感器系统,结合带间级联激光器与紧凑多通道气体电池达到了高灵敏度与低功耗的要求,针对甲烷和乙烷的检测下限分别为 5×10^{-9} 与 8×10^{-9}。周彧等[9]使用我国自主研发的 7.6 μm 中红外激光器研制了 N_2O 检测系统,在 30 ms 积分时间下系统探测极限达到 150×10^{-9},后续通过平均进一步降低噪声。李春光等[10]使用 3.337 μm 室温连续带间级联激光器研制了乙烷传感器,并论证了系统在实验室条件下的系统稳定性。

上述基于中红外激光器的气体测量系统多数针对实验室理想环境研发,前期必须使用 F-P 标准具等光学器材进行精细调参。工业上也主要集中于环境监测等领域,允许“以时间换精度”,通过长时间积分提高检测精度。对于气雾剂灌装产线的丙丁烷泄漏浓度检测而言,其实际生产环境和检测速度要求均无法满足上述条件。

本文针对气雾剂灌装产线的泄漏检测需求,搭建了基于带间级联激光器与可调谐吸收光谱技术的丙丁烷自动检漏系统。为实现在丙丁烷 3.37 μm 吸收峰附近的有效检测,并方便系统日常维护,研究使用基于洛伦兹线型拟合算法分析处理直接吸收光谱,从而可不依赖标气与精密光学仪器,自动选取确定适合的激光器电流锯齿波中心值。

1　丙丁烷中红外吸收光谱

红外光谱由分子吸收相应波长的红外光向能量更高的振(转)动能级的跃迁产生。根据 HITRAN2016 提供的数据[11],丙烷与丁烷在 600～6 000 cm^{-1} 的基频吸收位置基本重叠,且均位于中红外波段(2.5～25 μm)[12]。图 1 描绘了丙烷与丁烷在 298.1 K、1.01 kPa 下的吸收光谱,横坐标为波数 σ,分辨率为 0.112 cm^{-1},纵坐标为吸收截面,由式(1)定义:

$$k_\nu=-\frac{\ln\tau_\nu}{\rho L} \tag{1}$$

式中,τ_ν 为气体在波长 ν 处的光谱透射率,ρ 是气体密度(molecule/cm^3),L 为光程。可见,丙丁烷在中红外波段波长 3.370 4 μm(2 967.0 cm^{-1})处的吸收强度远高于其他位置,达到 10^{-18} 数量级。

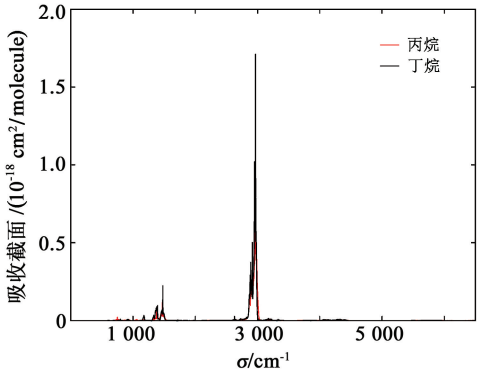


图 1　丙丁烷吸收谱线

空气中存在的水气是气体检测过程中存在的潜在干扰。图 2 为空气中水气浓度是丙丁烷(由 30% 丙烷和 70% 丁烷构成)的 2×10^4 倍时的吸收谱线,

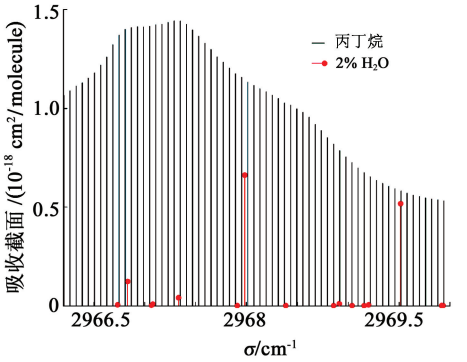


图 2　丙丁烷与 H_2O 的吸收光谱

在波长范围 3.367～3.372 μm 内，远离水气的主要吸收波段，尤其在丙丁烷吸收峰 3.370 4 μm 附近可避开背景气体中水气的干扰，据此可确定 ICL 激光器的中心频率。

2 基于洛伦兹线型拟合的 TDLAS 调制参数调节

2.1 波长调制的中红外吸收光谱

波长调制技术广泛应用于低浓度气体检测中，电流调制模块在锯齿波基础上叠加高频率正弦波，使激光器输出光的频率重复扫描目标气体吸收峰附近区间。驱动电流信号如式(2)：

$$u(t)=u_{\text{ramp}}(t)+A\cos(2\pi ft) \tag{2}$$

其中，$u_{\text{ramp}}(t)$为锯齿波信号，A 和 f 分别为调制信号的幅度和频率。

激光器加载电流 $u(t)$后，其输出频率可表示为

$$\nu(t)=\nu_0+a\cos(2\pi ft) \tag{3}$$

式中，ν_0 为锯齿波中心值对应的激光器波长，a 为调制深度，f 为调制信号的频率。同时，激光器输出光强的前两阶强度响应也不可忽略[13]，如式(4)所示：

$$P_0(t)=\overline{I_0}[1+i_1\cos(2\pi ft+\varphi_1)+i_2\cos(4\pi ft+\varphi_2)] \tag{4}$$

式中，$\overline{P_0}$ 为锯齿波中心频率对应光强，i_1，i_2 为光强调制被归一化的线型与非线性幅度。

因此，经过光程池后的透射光强如式(5)所示：

$$P(t)=P_0(t)\tau(\nu(t))=\sum_{k=0}^{\infty}H_k(\bar{\nu},a)\cos(2\pi fkt) \tag{5}$$

式中，H_k 为透射率 τ 的傅里叶系数。显然，经锁相滤波解调后，其二次谐波强度在弱吸收条件下正比于 H_k，根据朗伯比尔定律[14]，也正比于气体浓度。因此基于光强传感器将光信号转化为电压信号后，通过检测谐波信号即可实现浓度测量。

2.2 基于直接吸收光谱的吸收峰检测

为减少二次谐波中的噪声，调制后的激光频率应当覆盖丙丁烷吸收峰，并避开水气等潜在背景气体的干扰谱线。选择合适的锯齿波中心值，有助于减少锯齿波的扫描范围，缩短扫描周期，提高系统的检测效率。然而，实际应用中检测系统往往无法使用 F-P 标准具等光学器件实时确定当前输出激光的频率；温度气压的改变同样会改变气体吸收峰的位置，需要对调制参数进行必要的调整。因此，实现锯齿波中心值等调制参数的自动确定则显得尤为重要。

相比波长调制技术，直接吸收光谱的电流信号中不包含高频调制部分，检测精度较低，但可用于辅助确定波长调制技术中的锯齿波中心值，以减少对 F-P 标准具、标气、示波器等仪器的依赖。本文采用的具体方法为，随时间均匀增加激光器的注入电流，采集 MCT 传感器的输出电压作为零吸收曲线；再充入含丙丁烷的检测样品(如气雾剂与空气混合物)，重复测量过程，归一化后进行吸收特征线型拟合。

在气压高于一个大气压的环境下，气体分子的碰撞是吸收谱线加宽的主要原因，此时谱线表现为洛伦兹线型，如式(6)。在 1 kPa 以下气压时，气体分子的热运动引起的多普勒加宽使谱线加宽，此时谱线表现为高斯线型，如式(7)。

$$\phi_{\text{L}}(\nu-\nu_0)=\frac{1}{\pi}\frac{\Delta\nu_{\text{L}}/2}{(\nu-\nu_0-\Delta\nu_{\text{s}})^2+\left(\frac{\Delta\nu_{\text{L}}}{2}\right)^2} \tag{6}$$

式中，$\Delta\nu_{\text{L}}$ 和 $\Delta\nu_{\text{s}}$ 为正比于压力和组分的半高全宽和压力频移，ν_0 为分子跃迁频率。

$$\phi_{\text{D}}(\nu)=\frac{2}{\Delta\nu_{\text{D}}}\sqrt{\frac{\ln 2}{\pi}}\exp\left(-4\ln 2\left(\frac{\nu-\nu_0}{\Delta\nu_{\text{D}}}\right)^2\right) \tag{7}$$

式中，$\Delta\nu_{\text{D}}$ 为多普勒半高全宽。因此，将以上参数作为拟合参数，假设频率与电流成线性关系，即可获得吸收峰中心频率对应的电流大小，用于确定波长调制参数。

3 基于 TDLAS 的丙丁烷检测系统

本文设计的检测系统原理如图 3 所示，激光光源选择带间级联激光器，由温控模块保持温度稳定，电流驱动模块调制输出电流。

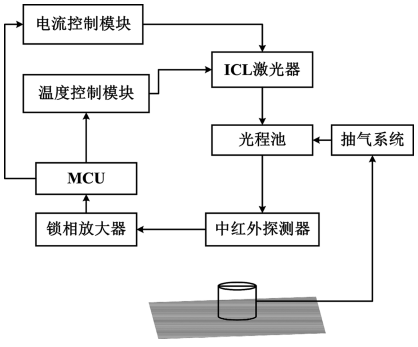


图 3 系统原理图

检测过程中，光束在光程池经多次反射达到全

光程后出射；中红外探测器将光强信号转换为模拟电压信号，并作为锁相放大器的输入，解调获得一次或二次谐波曲线，最终输入工控机进行计算处理。

4　实验测试与分析

随时间均匀提高图 3 中 DFB-ICL 激光器的驱动电流 I_{ICL}，获得探测传感器输出电压 V_{MCT} 与激光器驱动电流的关系曲线，如图 4 所示。其中，上方曲线为零吸收状态，下方曲线为注入了目标气体的有吸收状态，可见吸收峰基本位于 30～35 mA 处。

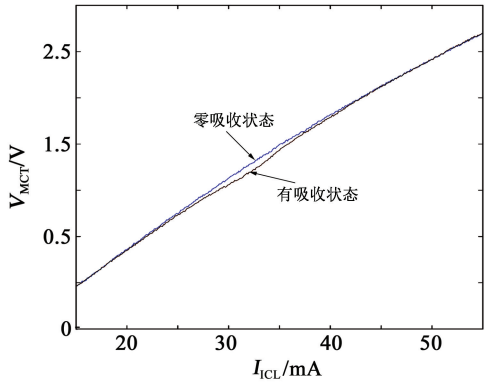


图 4　传感器输出电压与激光器驱动电流关系曲线

对吸光度进行洛伦兹和高斯线型拟合，结果如图 5 和 6 所示，其中纵坐标为通过式(8)计算获得的吸光度，具体参数见表 1。从图 5～6 中可以看出，两种线型的拟合结果得出的中心值均位于 32 mA 附近，而洛伦兹线型的拟合度更好，与常压下的理论值吻合，故此选择可信度更高的洛伦兹线型拟合结果为波形调制参数，取 31.94 mA 作为锯齿波中心值。若拟合结果中决定系数(R-square)不足 0.9，则认为可能存在干扰气体或系统故障，自检不通过。

$$\alpha=(P_0-P)/P_0 \tag{8}$$

式中，α 为吸光度，P_0 为无吸收情况下的传感器输出光强，P 为有吸收情况下的光强传感器输出。

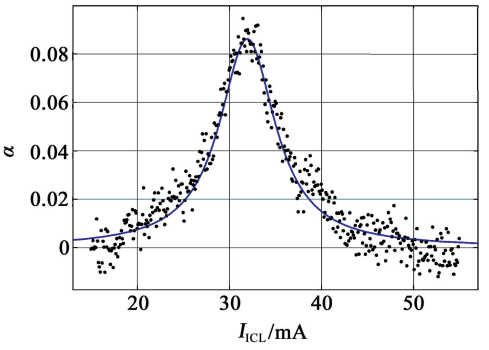


图 5　洛伦兹线型拟合结果

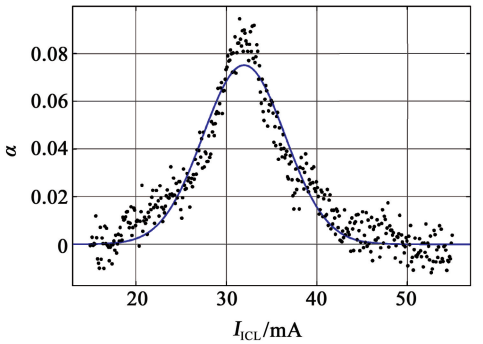


图 6　高斯线型拟合结果

表 1　对吸光度进行洛伦兹和高斯线型拟合结果

线型	洛伦兹线型	高斯线型
拟合模型	$\frac{1}{\pi}\frac{a}{(x-b)^2+a^2}$	$a\exp\left(-\frac{(x-b)^2}{c^2}\right)$
a	3.695	0.075 19
(95%置信区间)	(3.591,3.798)	(0.073 11,0.077 27)
b	31.94	31.96
(95%置信区间)	(31.84,32.05)	(31.82,32.11)
c	—	6.418
(95%置信区间)		(6.213,6.623)
SSE	0.017 36	0.023 9
R-square	0.937 2	0.913 6
调制 R-square	0.937 1	0.913 2
RMSE	0.006 596	0.007 75

保持光程池密闭以维持目标气体浓度，在锯齿波基础上叠加幅值为 0.03 V 的 20 kHz 高频正弦波，经锁相滤波获得在 31.97 mA 附近的二次谐波，如图 7 所示，可见扫描范围能较好地覆盖丙丁烷吸收峰，峰值超过 100 mA。

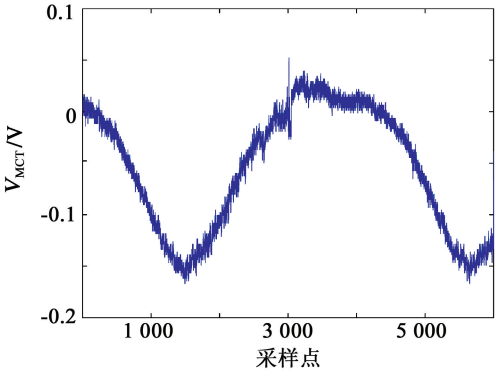


图 7　二次谐波曲线

5　结论

本文利用 DFB-ICL 激光器，搭建了用于检测气雾推进剂丙丁烷的气体泄漏检测系统。通过对直接吸收光谱的洛伦兹线型拟合，确定了调制参数中的锯齿波中心值，并进一步使用波长调制技术，获得了丙丁烷吸收光谱的二次谐波曲线。相关参数确定过程不需要借助 F-P 标准具、标气等进行人工标定，

提高了系统自检的自动化程度,降低了对维护人员的要求。

参考文献:

[1] Mishra D P, Rahman A. An experimental study of flammability limits of LPG/air mixtures[J]. Fuel, 2003, 82(7): 863-866.

[2] 张 一, 张 宇, 杨成奥, 等. 3～4 μm 锑化物带间级联激光器研究进展(特邀)[J]. 红外与激光工程, 2018, 47(10): 19-25.
Zhang Y, Zhang Y, Yang C A, et al. Research progress of 3～4 μm antimonide interband cascade laser(invited)[J]. Infrared and Laser Engin., 2018, 47(10): 19-25.

[3] 楊瑞青, 李 路, 江宇超. 带间级联激光器: 从原始概念到实际器件[J]. 物理学进展, 2014, 34(4): 169-190.
Yang R Q, Li L, Jiang Y C. Interband cascade lasers: from original concept to practical devices[J]. Progress in Physics, 2014, 34(4): 169-190.

[4] Webster C R, Mahaffy P R, Atreya S K, et al. Mars methane detection and variability at gale crater[J]. Science, 2015, 347(6220): 415-417.

[5] Namjou K, Cai S, Whittaker E A, et al. Sensitive absorption spectroscopy with a room-temperature distributed-feedback quantum-cascade laser[J]. Opt. Lett., 1998, 23(3): 219-221.

[6] Rieker G B, Jeffries J B, Hanson R K. Calibration-free wavelength-modulation spectroscopy for measurements of gas temperature and concentration in harsh environments[J]. Appl. Opt., 2009, 48(29): 5546-5560.

[7] Du Z, Li J, Gao H, et al. Ultrahigh-resolution spectroscopy for methyl mercaptan at the ν_2-band by a distributed feedback interband cascade laser[J]. J. of Quantitative Spectroscopy and Radiative Transfer, 2017, 196: 123-129.

[8] Dong L, Tittel F K, Li C, et al. Compact TDLAS based sensor design using interband cascade lasers for mid-IR trace gas sensing[J]. Opt. Express, 2016, 24(6): A528-A535.

[9] 周 彧, 曹 渊, 朱公栋, 等. 基于 7.6 μm 量子级联激光的光声光谱探测 N_2O 气体[J]. 物理学报, 2018, 67(8): 84201-084201.
Zhou Y, Cao Y, Zhu G D, et al. Photoacoustic spectroscopy based on 7.6 μm quantum cascade laser for detecting N_2O gas [J]. Acta Physica Sinica, 2018, 67(8): 84201-084201.

[10] 李春光, 董 磊, 郑传涛, 等. 中红外痕量乙烷传感器设计与稳定性分析[J]. 光谱学与光谱分析, 2019, 39(3): 959.
Li C G, Dong L, Zheng C D, et al. Design and stability analysis of mid-infrared trace ethane sensors[J]. Spectroscopy and Spectral Analysis, 2019, 39(3): 959.

[11] Rothman L S, Jacquemart D, Barbe A, et al. The HITRAN 2004 molecular spectroscopic database[J]. J. of Quantitative Spectroscopy and Radiative Transfer, 2005, 96(2): 139-204.

[12] 杜振辉, 韩瑞炎, 王晓雨, 等. 采用带间级联激光器的痕量气体检测技术进展[J]. 中国激光, 2018, 45(9): 0911006.
Du Z H, Han R Y, Wang X Y, et al. Interband cascade lasers based trace gas sensing: A review[J]. Chinese J. of Lasers, 2018, 45(9): 0911006.

[13] Li H, Rieker G B, Liu X, et al. Extension of wavelength-modulation spectroscopy to large modulation depth for diode laser absorption measurements in high-pressure gases[J]. Appl. Opt., 2006, 45(5): 1052-1061.

[14] Mayerhofer T G, Mutschke H, Popp J. Employing theories far beyond their limits—the case of the (Boguer-) Beer-Lambert law[J]. Chem. Phys. Chem., 2016, 17(13): 1948-1955.

作者简介:

张 铭(1994—),女,浙江湖州人,硕士研究生,主要从事激光点云与光电检测方面的研究。

E-mail: 18317010165@163.com

(上接第 117 页)

[2] 王晨阳, 张卫平, 邹 阳. 仿昆虫扑翼微飞行器研究现状与关键技术[J]. 无人系统技术, 2018, 1(4): 5-20.
Wang Chenyang, Zhang Weiping, Zou Yang. Current status and key techniques of insect-inspired flapping-wing micro air vehicles[J]. Unmanned Systems Technol., 2018, 1(4): 5-20.

[3] Keennon M, Klingebiel K, Won H. Development of the nano hummingbird: A tailless flapping wing micro air vehicle[C]// 50th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition, 2012: 1-24.

[4] Shi Jianbo, Tomasi C. Good features to track[C]// Proc. of 1994 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1994: 593-600.

[5] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C]// Proc. of The 7th Inter. Joint Conf. on Artificial Intelligence: 1981: 674-679.

[6] Redmon J, Farhadi A. Yolov3: An incremental improvement [J]. arXiv e-print, 2018: 1804.02767.

[7] Cao P, Zhang S, Tang J. Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning[J]. J. of IEEE Access, 2018, 6: 26241-26253.

作者简介:

王昕鹏(1995—),男,江苏扬州人,硕士生,主要研究方向为计算机视觉在扑翼无人机中的应用;

张卫平(1971—),男,博士,教授,博士生导师,研究方向为微机电系统的集成、设计、检测和控制。

E-mail: zwp37@163.com

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.026

基于随机 Radon 变换的 DRPE 加密系统

彭凯飞[1]，沈学举[1]，黄富瑜[1]，刘毅杰[1]，高　山[2]

（1. 陆军工程大学石家庄校区 电子与光学工程系，石家庄 050003；
2. 中国人民解放军 63751 部队，西安 710043）

摘　要： 对于使用随机相位模板(RPM)作为密钥的双随机相位编码(DRPE)加密系统，唯密文攻击算法(COA)是一种有效的攻击算法。为了提高 DRPE 加密系统抵抗 COA 算法攻击的能力，提出了一种基于随机 Radon 变换的 JTC 加密系统。该系统对待加密图像进行随机 Radon 变换，再使用 JTC 加密系统对其加、解密。除了 DRPE 加密系统中的 RPM 密钥之外，Radon 变换参数亦被用作恢复原始图像的密钥。仿真结果表明，相比于仅使用 RPM 密钥的 DRPE 加密系统，基于随机 Radon 变换的 DRPE 加密系统可以有效抵抗 COA 算法攻击，提高加密系统的安全性。

关键词： 光学信息处理；光学图像加密；双随机相位编码；随机 Radon 变换

中图分类号： TP391.41；TP309.7　　**文章编号：** 1001－5868(2020)01－0123－05

Double Random-phase Encoding Encryption System Based on Random Radon Transform

PENG Kaifei[1], SHEN Xueju[1], HUANG Fuyu[1], LIU Yijie[1], GAO Shan[2]

(1. Dept. of Opto-Electronics Engin., Campus of Shijiazhuang, Army Engin. University, Shijiazhuang 050003, CHN;
2. Unit 63751 of PLA, Xi'an 710043, CHN)

Abstract: The ciphertext only attack algorithm (COA) is an effective attack algorithm for the double random-phase encoding (DRPE) encryption system using random phase mask (RPM) as the key. In order to improve the ability of DRPE encryption system to resist ciphertext only attack algorithm attack, a DRPE encryption system based on random Radon transform is proposed. The system performs random Radon transform on the image to be encrypted and then encrypts it using the DRPE encryption system. In addition to RPM, a set of random parameters selected by the random Radon transform is used as a new key for successfully recovery the original image. The simulation results show that the DRPE encryption system based on Radon transform can effectively resist the COA algorithm and improve the security of the DRPE encryption system compared to the DRPE encryption system using only RPM as the key.

Key words: optical information processing; optical image encryption; double random-phase encoding; random Radon transform

0　引言

1995 年，由 Refregier 等提出的双随机相位编码技术[1]开创了运用光学信息技术解决加密问题的先河。该方法自提出以来，受到了很多学者的关注，并逐渐成为研究热点。1997 年，Javidi 进一步完善了该方法[2]；1999 年，Zhang 等人提出了基于双随机相位编码(DRPE)方法的彩色图像加密技术[3]，使 DRPE 技术的应用范围更广；2000 年，Unnikrishnan 等人提出了一种基于分数傅里叶变换的 DRPE 加密系统[4]，在该系统中将分数傅里叶变换的阶数以及随机相位模板(RPM)一起用作恢复原始图像的密钥。由于增加了加密系统的密钥参量，其安全性得到了有效提高；2004 年，Sit 基于 DRPE 方法，结合菲涅

收稿日期：2019－09－18.
基金项目：国家自然科学基金项目(61801507).

尔变换[5]进行光学图像加密；2008 年，Cheng 等人通过底层振幅调制[6]，打破了 DRPE 加密系统输入和输出之间的线性关系，提高了系统抗选择明文攻击的能力；2011 年，Wang 等人利用四元数傅里叶变换(DQFT)，实现了使用单通道对彩色图像的加密[7]；2013 年，Myungjin 等人提出了一种三维光子计数 DRPE 技术[8]，由于其采用非线性变换，安全性相对于 DRPE 加密系统有很大提升。随后，Zhao 等人[9-15]分别将多种其他变换与 DRPE 相结合进行光学图像加密，使整个 DRPE 加密系统的体系更加完善、多样性更强。

由于 DRPE 加密系统采用的是对称密码体制，其系统的输入和输出满足一定的线性关系，存在易被攻击的潜在风险[16]。文献[17]表明，唯密文攻击(COA)算法是一种攻击 DRPE 加密系统的有效算法。针对仅使用 RPM 作为密钥的 DRPE 加密系统抵抗 COA 算法攻击能力不足[18]的问题，本文提出了一种基于随机 Radon 变换的联合变换相关器加密系统。该系统引入 Radon 变换的参数作为新的密钥对图像进行加密，提高了 DRPE 加密系统的安全性。

1 Radon 变换原理

Radon 变换是一种投影方式，是某个函数沿给定的路径进行的积分运算，因此二维函数的 Radon 变换是一维向量。当被积函数的积分路径为直线时，Radon 变换称为线性 Radon 变换。线性 Radon 变换可以选取不同方向的直线作为被检测图像的灰度分布函数的积分路径，来获得被检测图像在不同方向上的投影积分。沿 y'方向的 Radon 变换为[19]

$$R_{\phi}(x')=\int_{-\infty}^{+\infty} f(x'\cos\phi-y'\sin\phi,x'\sin\phi+y'\cos\phi)\mathrm{d}y' \tag{1}$$

$$\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}\cos\phi & \sin\phi\\-\sin\phi & \cos\phi\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} \tag{2}$$

其中，f 是被积分函数，x，y 和 x'，y'分别为原始空间坐标和 Radon 空间坐标。ϕ 是 x'与 x 的夹角。

2 基于随机 Radon 变换的 DRPE 加密系统

2.1 随机 Radon 变换方法

设 $n\times n$ 像素的图像为 $f(x,y)$，随机产生 m 个参数 θ_i，按照式(3)对 $f(x,y)$进行 Radon 变换，得到 m 个一维列向量 $\boldsymbol{R}_{\theta_i}(x')$，其中 $1\leqslant x,y\leqslant n$，$i=1,2,\cdots,m$，$0\leqslant\theta_i\leqslant\pi$。当 $\theta_i=\pi/4,3\pi/4$ 时，列向量 $\boldsymbol{R}_{\theta_i}(x')$长度达到最长，为$\lceil\sqrt{2n}\rceil$。为统一长度，通过末尾补 0 将得到的 m 个一维列向量 $\boldsymbol{R}_{\theta_i}(x')$长度均统一为$\lceil\sqrt{2n}\rceil$。按式(3)将 m 个一维列向量$\boldsymbol{R}_{\theta_i}(x')$构成矩阵 $\boldsymbol{g}(x',y')$：

$$\boldsymbol{g}(x',y')=\{\boldsymbol{R}_{\theta_1}(x'),\boldsymbol{R}_{\theta_2}(x'),\cdots,\boldsymbol{R}_{\theta_i}(x'),\cdots,\boldsymbol{R}_{\theta_m}(x')\} \tag{3}$$

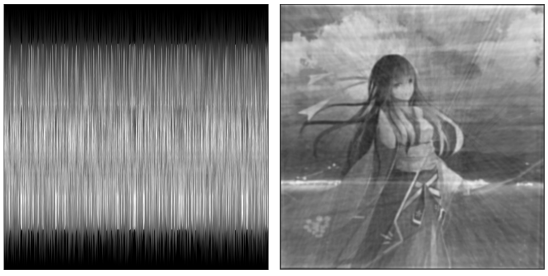
式中，(x,y)为原平面直角坐标系，(x',y')为 Radon 空间坐标系。

按以上方法即可将平面直角坐标系下 $n\times n$ 像素的图像 $f(x,y)$通过随机 Radon 变换转换为 Radon 坐标系下$\lceil\sqrt{2n}\rceil\times m$ 像素的图像$g(x',y')$。

利用随机 Radon 变换时使用的 θ_i 参数，通过 Radon 逆变换就可以将 $g(x',y')$重新变换为 $f(x,y)$。

下面以图 1 所示的图像为例，选取 $m=500$ 按照上述方法对其进行随机 Radon 变换，结果如图 2(a)所示。通过 Radon 逆变换得到的图像如图 2(b)所示。

图 1 原始灰度图像



(a) 随机 Radon 变换结果 (b) Radon 逆变换结果

图 2 对图 1 所示图像进行随机 Radon 变换及逆变换结果

2.2 基于随机 Radon 变换的 DRPE 加密系统

鉴于 DRPE 加密系统的抗攻击能力较差，因此通过对待加密图像先进行随机 Radon 变换，引入 Radon 变换参数 θ_i 作为密钥，期望提高 DRPE 加密系统的安全性。

基于随机 Radon 变换的 DRPE 加密系统示意图如图 3(a)所示，以输入面中心为原点，建立 x，y 直角坐标系。随机生成 m 个 Radon 变换参数 θ_i 作为密钥，将 $n\times n$ 像素的待加密图像按 2.1 中的方法转换为 Radon 空间 $\lceil\sqrt{2n}\rceil\times m$ 像素的图像 $f(x,y)$，与 $\lceil\sqrt{2n}\rceil\times m$ 像素的随机相位模板 $r(x,y)$ 中心重合叠加在一起置于输入面中心；将随机相位模板 $k(u,v)$ 作为密钥置于频谱面中心，其中心坐标为 $(a,0)$。单色平面波照射下，经两次光学傅里叶变换在输出面记录复振幅密文为

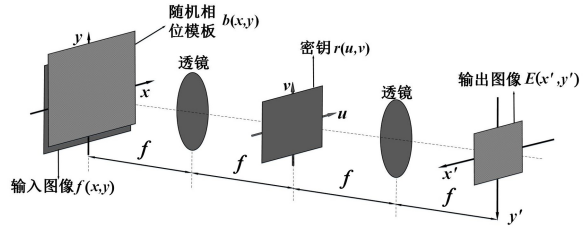
$$E(x',y')=\mathscr{F}^{-1}\{\mathscr{F}[f(x,y)\cdot r(x,y)]k(u,v)\} \tag{4}$$

其中，(x',y') 为输出面坐标，(u,v) 为频谱面坐标，$\mathscr{F}\{\cdot\}$ 代表傅里叶变换，$\mathscr{F}^{-1}\{\cdot\}$ 代表傅里叶逆变换。

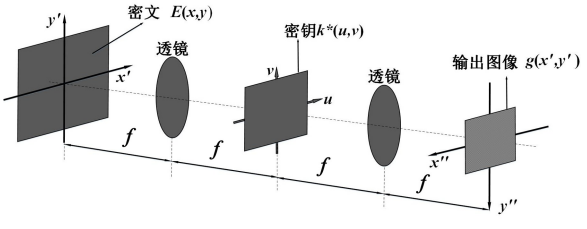
解密系统示意图如图 3(b)所示，将 $k(u,v)$ 的复共轭密钥 $k^*(u,v)$ 置于频谱面中心，将复振幅密文 $E(x',y')$ 置于输入面中心。单色平面波照射下，经两次光学傅里叶变换再输出面上的光场复振幅为

$$g(x'',y'')=\mathscr{F}^{-1}\{\mathscr{F}\{E(x',y')\}\cdot k^*(u,v)\}=f(x,y)\cdot r(x,y) \tag{5}$$

式中，(x'',y'') 为解密系统输出面坐标，“$[\,]^*$”代表复共轭。由于式中仅有 $f(x,y)$ 为实函数，且 $r(x,y)$ 模为 1，因此在输出面使用 CCD 记录即可得到解密图像。



(a)　加密系统



(b)　解密系统

图 3　DRPE 加、解密系统示意图

而后将 CCD 记录得到的解密图像利用密钥 θ_i 进行逆 Radon 变换即可解密出原始图像。

3　数值仿真及其抗攻击特性分析

为验证随机 Radon 变换在 DRPE 加密系统中提高加密系统安全性的作用，仿真了基于随机 Radon 变换的 DRPE 加密系统选取不同密钥参数 m 的加、解密过程，并分析了该系统的抗攻击特性。仿真时使用的待加密图像如图 4 所示。激光波长为 632.8 nm，傅里叶透镜焦距为 400 mm。

为定量评估解密图像质量，使用解密图像与原始图像的相关系数（X_{CC}，称 CC 值）作为评价指标，其定义式为[20]

$$X_{CC}=\frac{\left|\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j)-\bar{f}][I(i,j)-\bar{I}]\right|}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j)-\bar{f}]^2}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)-\bar{I}]^2}} \tag{6}$$

式中，i，j 为像素坐标，f 和 $\bar{f}$ 分别表示原始图像的灰度值和灰度值平均值，I 和 $\bar{I}$ 分别表示解密图像的灰度值和灰度值平均值，M 和 N 分别是像素的横纵坐标总数。

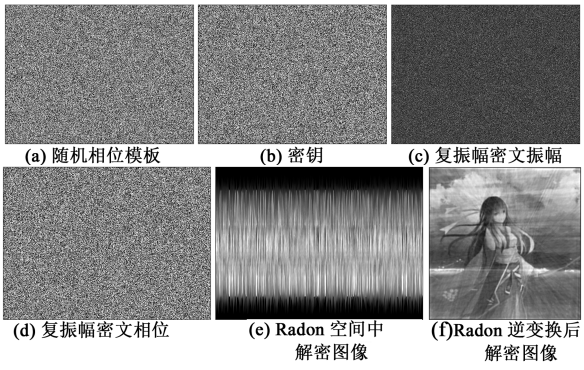


图 4　仿真结果

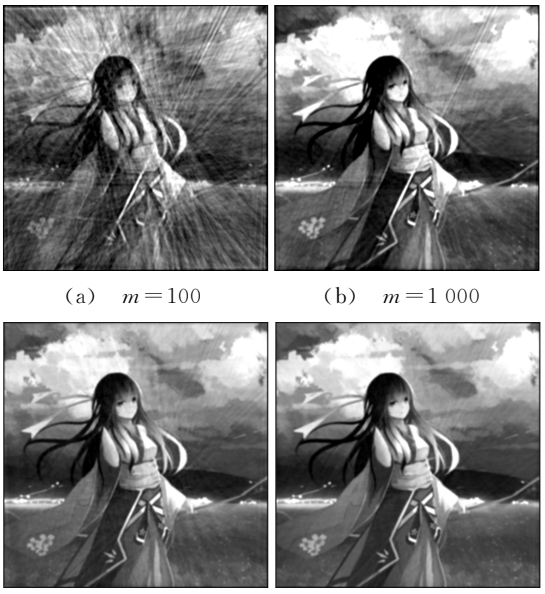
3.1　加、解密过程仿真

为验证基于随机 Radon 变换的 DRPE 加密系统的正确性，对其加、解密过程进行仿真。仿真使用的原始灰度图像如图 1 所示。经 $m=500$ 的随机 Radon 变换后的待加密图像如图 2(a)所示。将图 2(a)所示待加密图像置于加密系统输入面中心，所用随机相位模板和密钥模板如图 4(a)、(b)，尺寸均与图 1 相同。按式(4)计算得到的复振幅密文，其振幅如图 4(c)所示，相位如图 4(d)所示。

解密时，将密钥模板的复共轭置于解密系统频谱面中心，将复共轭密文置于解密系统输入面中心，在输出面通过式(5)计算得到 Radon 空间中的解密图像，如图 4(e)所示，利用 2.1 中方法将图 4(e)进

行 Radon 逆变换，得到的解密图像如图 4(f)所示，其与原始图像的 *CC* 值为 0.930 7。

选取不同的 m 值得到的解密图像如图 5(a)～(d)所示，按式(6)计算其 *CC* 值分别为 0.919 7，0.941 6，0.984 1，0.990 1。可以看出随着 m 值的增大，解密图像质量随之提高。仿真结果验证了基于随机 Radon 变换的 DRPE 加密系统的正确性。



(a) $m=100$ (b) $m=1\,000$

(c) $m=3\,000$ (d) $m=5\,000$

图 5 选取不同 m 值时得到的解密图像

3.2 抗攻击特性分析

为验证随机 Radon 变换方法对于提高 DRPE 加密系统抗攻击能力的作用，对本文提出的加密系统的抗攻击能力进行分析。

对于 DRPE 加密系统，COA 算法是一种有效的攻击算法，其流程如图 6 所示。在本节中就使用 COA 算法攻击基于随机 Radon 变换的 DRPE 加密系统，分析其抗攻击能力。

使用 COA 算法攻击未使用随机 Radon 变换的 DRPE 加密系统，原始图像如图 4 所示，攻击结果如图 7 所示，其中图 7(a)～(d)分别为迭代次数 500，1 000，3 000，30 000 次时的攻击结果，按式(6)计算其 *CC* 值分别为 0.304 7，0.405 5，0.514 2，0.602 5。

使用 COA 算法攻击基于随机 Radon 变换的 DRPE 加密系统，原始图像如图 4 所示，攻击结果如图 8 所示，其中图 8(a)～(c)分别迭代次数为 500，1 000，3 000，30 000 时的攻击结果，按式(6)计算其 *CC* 值均小于 0.3，攻击算法失效。

对比图 7 和 8 可以看出，利用 COA 迭代算法攻击未使用随机 Radon 变换的 DRPE 加密系统，当迭代次数达到 1 000 次以上时，可以从加密图像中较好地恢复出原始图像。而对于基于随机 Radon 变换的 DRPE 加密系统，即使迭代次数达到 30 000 次时，也无法从加密图像中恢复出任何有效信息。

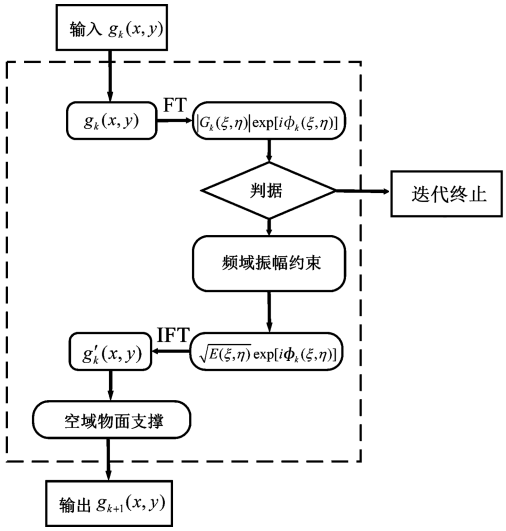
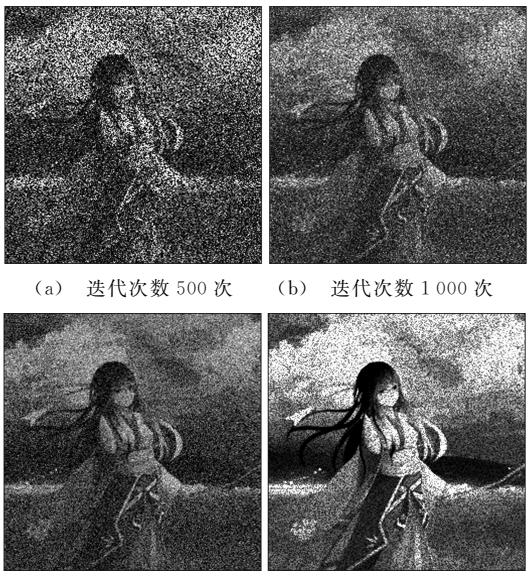


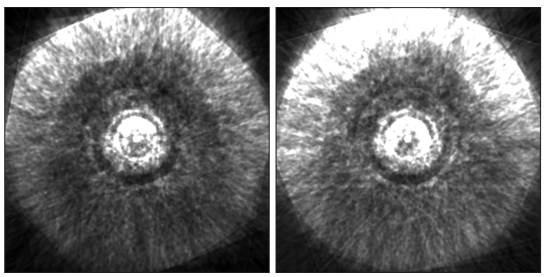
图 6 COA 算法流程示意图



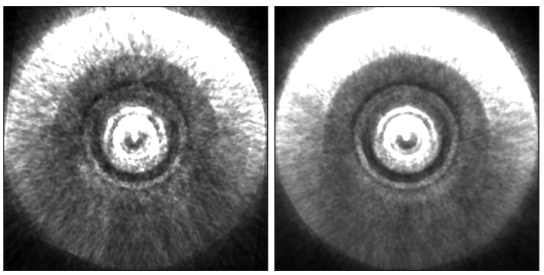
(a) 迭代次数 500 次 (b) 迭代次数 1 000 次

(c) 迭代次数 3 000 次 (d) 迭代次数 30 000 次

图 7 COA 攻击未使用随机 Radon 变换的 DRPE 加密系统结果



(a) 迭代次数 500 次 (b) 迭代次数 1 000 次



(c) 迭代次数 3 000 次　　(d) 迭代次数 30 000 次

图 8　COA 攻击基于随机 Radon 变换的 DRPE 加密系统结果

4　结论

本文提出了一种图像的随机 Radon 变换方法，并将该方法应用于 DRPE 加密系统，有效提高了加密系统的抗攻击能力。仿真结果进一步验证了基于随机 Radon 变换的 DRPE 加密系统的正确性。仿真结果表明随机 Radon 变换方法能有效提高 DRPE 加密系统的抗 COA 攻击能力，加密系统安全性得到了较大提升。

参考文献：

[1] Refregier P, Javidi B. Optical image encryption based on input plane and Fourier plane random encoding[J]. Opt. Lett., 1995, 20(7): 767-769.

[2] Javidi B, Zhang G S, Li J. Encrypted optical memory using double random phase encoding[J]. Appl. Opt., 1997, 36(5): 1054-1058.

[3] Zhang S Q, Karim M A. Color image encryption using double random phase encoding[J]. Microwave and Optical Technol. Lett., 1999, 21(5): 318-323.

[4] Unnikrishnan G, Joseph J, Singh K. Optical encryption by double-random phase encoding in the fractional Fourier domain [J]. Opt. Lett., 2000, 25(12): 887-889.

[5] Situ G, Zhang J. Double random-phase encoding in the Fresnel domain[J]. Opt. Lett., 2004, 29(14): 1584-1586.

[6] Cheng X C, Cai L Z, Wang Y R, et al. Security enhancement of double-random phase encryption by amplitude modulation [J]. Opt. Lett., 2008, 33(14): 1575-1577.

[7] Wang X, Zhai H, Li Z, et al. Double random-phase encryption based on discrete quaternion Fourier-transforms [J]. Optik-Inter. J. for Light and Electron Opt., 2011, 122 (20): 1856-1859.

[8] Cho M, Javidi B. Three-dimensional photon counting double-random-phase encryption[J]. Opt. Lett., 2013, 38(17): 3198.

[9] Li H, Wang Y, Yan H, et al. Double-image encryption by using chaos-based local pixels crambling technique and gyrator transform[J]. Optics and Lasers in Engin., 2013, 51(12): 1327-1331.

[10] Zhou N, Wang Y, Gong L. Novel optical image encryption scheme based on fractional Mellin transform [J]. Opt. Commun., 2011, 284: 3234-3242.

[11] Liu Z, Li S, Liu W, et al. Image encryption algorithm by using fractional Fourier transform and pixel scrambling operation based on double random phase encoding[J]. Opt. and Lasers in Engin., 2013, 51(1): 8-14.

[12] Liu Z, Guo Q, Xu L, et al. Double image encryption by using iterative random binary encoding in gyrator domains [J]. Opt. Express, 2010, 18(11): 12033-12043.

[13] Zhi Z, Zhang Y, Shan M. Double-image encryption using chaotic maps and nonlinear non-DC joint fractional Fourier transform correlator [J]. Opt. Engin., 2016, 55(9): 093109.

[14] Chen L, Zhao D. Optical image encryption with Hartley transforms[J]. Opt. Lett., 2006, 31(23): 3438-3440.

[15] Singh N, Sinha A. Optical image encryption using Hartley transform and logistic map[J]. Opt. Commun., 2009, 282 (6): 1104-1109.

[16] Qin W, Peng X. Asymmetric cryptosystem based on phase-truncated Fourier transforms[J]. Opt. Lett., 2010, 35(2): 118-120.

[17] Situ G, Gopinathan U, Monaghan D S, et al. Cryptosystem of optical security systems with significant output images[J]. Appl. Opt., 2007, 46(22): 5257-5262.

[18] Liu X, Wu J, He W, et al. Vulnerability to ciphertext-only attack of optical encryption scheme based on double random phase encoding[J]. Opt. Express, 2015, 23(15): 18955.

[19] 丁伟雄. MATLAB R2015a 数字图像处理[M]. 北京：清华大学出版社，2016：240-242.
Ding W X. MATLAB R2015a Digital Image Processing[M]. Beijing: Tsinghua University Press, 2016: 240-242.

[20] 沈学举，刘旭敏，蔡　宁，等. 非线性 JTC 光学图像加密系统及其消噪音和抗攻击特性研究[J]. 中国激光，2015，42 (7)：284-290.
Shen X J, Liu X M, Cai N, et al. Nonlinear image encryption system based on JTC and its removing noise and resisting attack properties research[J]. Chinese J. of Laser, 2015, 42(7): 284-290.

作者简介：

彭凯飞(1994－)，男，硕士研究生，主要研究方向为计算机视觉；

沈学举(1963－)，男，博士，教授，主要研究方向为光信息处理。

E-mail：shxjoptics@aliyun.com

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.027

针对间断纹理环境中的图像特征追踪和匹配算法研究

赵明富[1,2]，曹利波[1,3]，宋　涛[1,2]，刘　帅[1]，罗宇航[1]

（1. 重庆理工大学 电气与电子工程学院，重庆 400054；2. 电梯智能运维重庆市高校工程中心，重庆 402260；
3. 光纤传感与光电检测重庆市重点实验室，重庆 400054）

摘　要： 针对同时定位与地图构建(SLAM)中的特征匹配关键环节，提出一种融合特征点和特征区域的图像追踪与匹配算法，以解决交替出现纹理丰富和纹理缺失的间断纹理环境中图像特征易丢失、误匹配率较高的问题。首先，利用ORB算法和半稠密直接法分别对图像提取特征点和特征区域。其次，使用渐进一致采样法(PROSAC)剔除ORB算法的误匹配特征点，并计算特征点的正确匹配率。最后，针对纹理缺失环境中特征点丢失严重的问题，以特征点的正确匹配率作为判断依据，对低匹配率图像，则基于特征区域使用半稠密直接法实现图像的追踪，同时对追踪结果进行非线性优化，提高了特征区域追踪的准确性和稳定性。实验结果表明，该算法适用于间断纹理环境，在纹理丰富和纹理缺失条件下均可提高图像匹配的准确率。

关键词： ORB算法；特征匹配；半稠密直接法；PROSAC

中图分类号： TP391.41　**文章编号：** 1001－5868(2020)01－0128－07

Research on Image Feature Tracking and Matching Algorithms in Intermittent Texture Environment

ZHAO Mingfu[1,2], CAO Libo[1,3], SONG Tao[1,2], LIU Shuai[1], LUO Yuhang[1]

(1. School of Electrical and Electronic Engin., Chongqing University of Technol., Chongqing 400054, CHN;
2. Elevator Intelligent Operation and Maintenance, Chongqing University Engin. Center, Chongqing 402260, CHN;
3. Chongqing Key Lab. of Optical Fiber Sensing and Photoelectric Detection, Chongqing 400054, CHN)

Abstract: Aiming at the key step of feature matching in SLAM, an image tracking and matching algorithm combining feature points and feature regions is proposed to solve the problem of easily losing image features and high mismatching rate in discontinuous texture environment where texture is rich and texture is missing alternately. Firstly, ORB algorithm and semi-dense direct method are used to extract feature points and feature regions respectively. Secondly, progressive uniform sampling (PROSAC) is used to eliminate the mismatched feature points of ORB algorithm, and the correct matching rate of feature points will be calculated. Finally, aiming at the serious problem of feature point loss in texture-missing environment, the correct matching rate of the feature point should be the judgement basis. As for the low-matching images, a semi-dense direct method should be used to track low-matching image based on the correct matching rate of feature points. At the same time, the tracking results are optimized nonlinearly to improve the accuracy and stability of feature area tracking. The experimental results show that the algorithm is suitable for discontinuous texture environment, and it can improve the accuracy of image matching on the conditions of rich texture and texture missing.

Key words: ORB algorithm; feature matching; semi-dense direct method; PROSA

收稿日期：2019－10－10.

基金项目：国家自然科学基金青年基金项目(61701056)；重庆市科委基础与前沿研究计划项目(cstc2016jcyjA0308)；重庆市教委基础研究项目(KJQN201901123)；重庆理工大学研究生创新基金项目(ycx20192051，ycx20192052).

0　引言

随着智能机器人技术的深入研究，对其自主性的要求也在不断提高[1]。基于视觉的同时定位与地图构建（Visual Simultaneous Localization and Mapping，VSLAM）技术具有越来越多的应用场景[2-3]。同时，视觉 SLAM 因具有丰富的图像信息、传感器价格便宜、适用范围广等优点，成为 SLAM 技术中最具潜力的研究方向[4-5]。目前经典的视觉 SLAM 框架主要通过前端视觉里程计（Visual Odometry，VO）[6-9]、后端优化、回环检测、三维重建等环节构建机器人需要的环境信息和对应的轨道信息，因此视觉 SLAM 技术对机器人的定位实时性和准确性有较高的要求[10]。

随着非线性优化理论和稀疏矩阵的不断完善，出现了许多经典的视觉 SLAM 方案，如：LSD-SLAM[11]，ORB-SLAM[12]，RGBD-SLAM[13] 等算法。

目前，经典视觉 SLAM 中的图像匹配方法主要分为两类：一类是基于特征点，另一类是基于灰度不变假设。基于特征点的图像匹配方法主要根据 SIFT[14-15]，SURF[16] 以及 Rublee 等[17] 提出的高效 ORB（Oriented FAST and Rotated BRIEF）特征提取算法。ORB 特征提取算法采用 FAST（Features from Accelerated Segment Test）算法检测图像特征点，并计算 BRIEF 描述子用来描述特征点的相似程度。该算法与 SIFT 相比运算速度要快 100 倍，与 SURF 算法相比要快 10 倍。由于 ORB 特征提取算法的快速性，所以被广泛应用于图像特征提取。虽然 ORB 特征提取算法整体表现优异，但也存在明显缺点：(1)在使用特征点时，往往忽略了除特征点以外的所有信息，丢弃了大部分可能有用的图像信息；(2)当相机运动到特征缺失的地方时，特征点的数量会明显减少，导致正确匹配的特征点数量偏少。基于灰度不变假设的图像匹配方法主要以直接法为主。在直接法中，不需要知道特征与特征之间的相对关系，而是通过最小化光度误差来求解它们之间的对应关系，根据像素的亮度信息估计出相机的运动，这就避免了特征缺失的情况下无法估计相机运动的状态。随着一些使用直接法的开源项目的出现（如 SVO[18]，LSD-SLAM 等），直接法也逐渐成为视觉 SLAM 主流。然而正因为直接法是通过不断更新两帧图像的光度误差，找到下一帧图像中的对应点或区域，因此必须保证沿着图像梯度走时，灰度误差不断下降。但是图像是一个很强烈的非凸函数，在寻找灰度误差最小值时，很容易由于本身具有的非凸性（或噪声）使灰度误差落入局部最小值。所以只有当相机运动很小、图像中的梯度不会有很强的非凸性时，直接法特征匹配才能成立。

综上所述，目前存在的经典视觉 SLAM 中的图像匹配算法均存在一定的缺陷，只能运行在特定的条件下。为此，本文在 ORB 算法和半稠密直接法的基础上，提出了一种针对间断纹理环境中的图像特征匹配算法。首先，利用 ORB 算法和半稠密直接法分别对图像提取特征点和特征区域。其次，针对 ORB 算法误匹配率较高的问题，使用渐进一致采样法（PROSAC）剔除其中的误匹配特征点，同时计算特征点的正确匹配率。最后，针对间断纹理环境中特征丢失严重的问题，使用特征点对数和特征点的正确匹配率作为判断标准。对于不符合标准的图像，则根据特征区域使用半稠密直接法实现对图像的追踪，同时使用非线性优化对追踪结果进行优化处理，确保特征区域追踪的准确性和稳定性。

1　ORB 算法和直接法原理

1.1　ORB 算法

2011 年，Rublee 等[7] 提出的 ORB 特征主要由关键点和描述子两部分组成，其关键点采用 Oriented FAST，是一种改进的 FAST 角点，具有快速检测局部像素灰度变化明显的特性。如图 1 所示为 FAST 角点示意图。

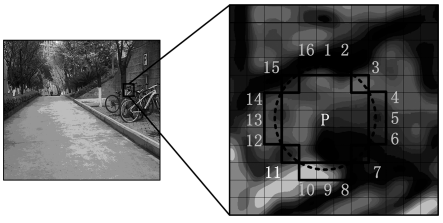


图 1　FAST 角点示意图

Oriented FAST 的主要思想为：如果图像中一个像素点与周围邻域像素有较大差异（过亮或者过暗），那么这个像素点就可能是一个角点。由于只检测了像素之间的灰度差异，所以 FAST 特征点往往数量巨大且不稳定。因此，ORB 对原始 FAST 特征检测算法做了改进。首先，计算原始 FAST 角点的 Harris 响应值，选取前 N 个响应最大的角点保留下来，形成最终的角点集合。其次，针对 FAST 不具

有方向信息和旋转不变性，ORB 在 FAST 角点的基础上使用灰度质心法，实现特征点的旋转不变性。定义图像块矩阵为

$$m_{pq}=\sum_{x,y\in B}x^{p}y^{q}I(x,y)\quad p,q=\{0,1\}\tag{1}$$

其中，$I(x,y)$表示图像灰度。通过矩阵可以找到图像块的质心：

$$C=\left(\frac{m_{10}}{m_{00}},\frac{m_{01}}{m_{00}}\right)\tag{2}$$

连接特征点 O 与质心点 C，形成方向向量$\overrightarrow{OC}$，因此特征点的方向定义为

$$\theta=\arctan(m_{01}/m_{10})\tag{3}$$

最后，由于 FAST 选取固定半径的邻域，存在尺度问题，从深度不同的地方看可能角点就不同。构建图像金字塔，并在金字塔的每一层检测角点实现 FAST 的尺度不变性。如图 2 所示为图像金字塔示意图。

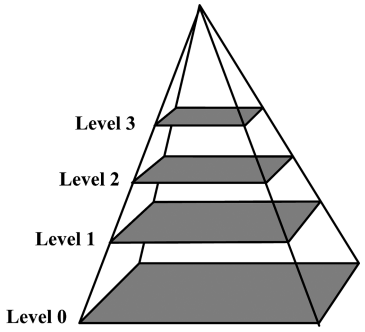


图 2 图像金字塔示意图

提取 Oriented FAST 特征点后，对其每个点计算描述子，ORB 使用改进的 BRIEF 特征描述——rBRIEF 描述子。主要思想为：选取特征点对邻域内 N 组点对(x_i,y_i)，$1\leqslant i\leqslant N$（一般 N 取 128，256 等），将每个点对进行二进制定义：

$$\tau(p;x,y)=\begin{cases}1, & p(x)<p(y)\\ 0, & \text{otherwise}\end{cases}\tag{4}$$

其中，$p(x)$，$p(y)$表示像素灰度。依次对随机点对生成二进制字符串，即为 BRIEF 描述子，定义为

$$f_N(p)=\sum_{1\leqslant i\leqslant N}2^{i-1}\tau(p;x_i,y_i)\tag{5}$$

此时的描述子 BRIEF 并不具有旋转不变性，因此 ORB 算法为 BRIEF 定义矩阵 $\boldsymbol{S}$ 选取初始点集对：

$$\boldsymbol{S}=\begin{pmatrix}x_1 & x_2 & \cdots & x_N\\ y_1 & y_2 & \cdots & y_N\end{pmatrix}_{2\times N}\tag{6}$$

利用灰度质心法检测到的主方向 θ 确定仿射变换矩阵 $\boldsymbol{R}_\theta$，所以矩阵 $\boldsymbol{S}$ 经过仿射变换得到 $\boldsymbol{S}_\theta$ 矩阵：

$$\boldsymbol{S}_\theta=\boldsymbol{R}_\theta\boldsymbol{S}=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\cdot\begin{bmatrix}x_1 & x_2 & \cdots & x_N\\ y_1 & y_2 & \cdots & y_N\end{bmatrix}\tag{7}$$

因此可以推出具有旋转不变特性的描述子 rBRIEF：

$$g_N(p,\theta):=f_N(p)\mid(x_i,y_i)\in\boldsymbol{S}_\theta\tag{8}$$

1.2 直接法特征跟踪

直接法是通过计算图像的像素灰度误差信息来跟踪相机运动的，根据特征点/区域在相机中的投影位置，通过最小化重投影误差达到优化相机运动的目的。因此，直接法的重要前提就是灰度不变假设。对于在 t 时刻位于(x,y)点的像素，假设其在$t+\mathrm{d}t$时刻运动到$(x+\mathrm{d}x,y+\mathrm{d}y)$点，那么可以得到：

$$I(x+\mathrm{d}x,y+\mathrm{d}y,t+\mathrm{d}t)=I(x,y,t)\tag{9}$$

其主要思想为：假设空间中有个点 P 和两个相邻时刻的相机图像。假设坐标点 P 的世界坐标为$[X,Y,Z]$，在两个相机图片上的对应成像记为非齐次像素坐标 P_1，P_2。图 3 为直接法示意图。

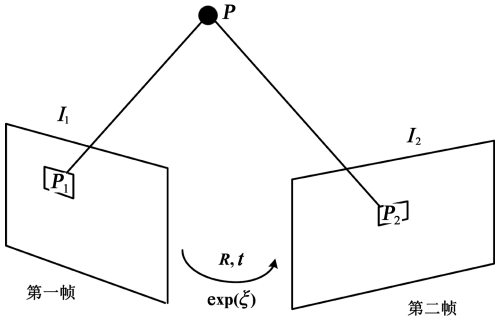


图 3 直接法示意图

我们希望求出第一个相机图像到第二相机图像的对应位姿变换。以第一帧图像为参考，假设第二帧图像的旋转和平移为 R，t（其中李代数为 ξ）。两图像的内参相同，均为 K，那么可以写出投影方程：

$$P_1=\begin{bmatrix}u\\ v\\ 1\end{bmatrix}_1=\frac{1}{Z_1}KP$$

$$P_2=\begin{bmatrix}u\\ v\\ 1\end{bmatrix}_2=\frac{1}{Z_2}K(RP+t)=\frac{1}{Z_2}K(\exp(\xi\Lambda)P)_{1:3}\tag{10}$$

其中，Z_1，Z_2 分别为两帧图像下 P 的深度。由于 $\exp(\xi\Lambda)$只能和齐次坐标相乘，所以只取前 3 个元素。

当相机的位姿不够好时，P_1，P_2 两点相差较大。为了减小两点之间误差，采用优化相机位姿的方法来寻找和 P_1 更加相似的点 P_2。这样就演变为求解一个优化问题：

$$e = I_1(p_1) - I_2(p_2) \tag{11}$$

那么根据灰度不变原理，整个相机位姿问题就变为

$$\min_{\xi} J(\xi) = \sum_{i=1}^{N} e_i^{\mathrm{T}} e_i \tag{12}$$

其中，e_i 表示第 i 个空间点的像素亮度误差。所以我们只关心误差是如何随相机的位姿变化的。因此使用李代数的扰动模型，最终可以求解出误差相对于李代数的雅可比矩阵：

$$\boldsymbol{J} = -\frac{\partial I_2}{\partial u}\frac{\partial u}{\partial \delta \xi} \tag{13}$$

最后使用高斯牛顿法计算增量，迭代求解。

2　针对间断纹理的特征追踪与匹配算法

直接法是一种提取速度快、性能良好的算法，对于帧与帧之间小范围位移时有较高的精度。但是对于大位移时，帧与帧之间很容易造成特征丢失或者使优化陷入局部最优的场景，造成相机定位丢失或者定位不准确。ORB 算法是一种性能稳定、匹配精度较高的特征匹配算法，但是 ORB 只是特征提取，丢失了大量除特征以外的其他图像信息，无法根据 ORB 特征点构建半稠密或稠密地图，且在无明显纹理的环境下（如白墙等），ORB 无法提取特征，导致特征点丢失，进而无法实现系统的定位。针对这两种特征点的缺陷以及互补性，提出一种能在间断纹理环境下正常工作的图像特征追踪与匹配算法，很好地解决了直接法和 ORB 算法的缺点，同时还提升了 ORB 算法的匹配精度。图 4 为算法流程图。

针对间断纹理环境图像特征追踪与匹配算法，主要步骤为：

步骤 1：通过相机获取连续帧图像，并对其进行预处理（去噪等），处理后的图像取出连续两帧图像 I_1 和 I_2。计算其图像深度并对取出的图像进行 ORB 特征提取，当某个像素周围有连续 N 个点的亮度偏大或偏小，那么这个像素便被认为是特征点。通常 N 取 9、11、12，本文选取 $N=12$，即为 FAST-12。表达式为

$$N = \sum_{x \in \forall (O(p))} |I(x) - I(p)| > \varepsilon_{\mathrm{d}} \tag{14}$$

其中，$I(x)$ 表示特征点领域内灰度，ε_{d} 表示灰度误差阈值。采用 8 级图像金字塔提取 ORB 特征点，使其具有尺度不变性。同时对取出的连续两帧图像提取像素梯度明显的特征区域，用于补充因为只使用 ORB 特征点导致的图像信息丢失。

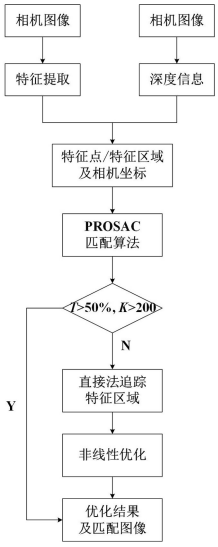


图 4　间断纹理环境下图像特征追踪与匹配算法

步骤 2：对于步骤 1 中提取的特征点采用 PROSAC 算法剔除误匹配，以提高匹配准确性，减少匹配运算量。主要思想为：将 X 个数据点的集合记作集合 u_x，根据评价函数 q，对集合中的数据进行降序排列：

$$u_i, u_j \in u_x : i < j \Rightarrow q(u_i) \geqslant q(u_j) \tag{15}$$

通过递归关系就可以随机且快速地找出采样点集合。为了防止类外点模型和算法一致，随机类内点集合的元素个数的分布满足二项式分布：

$$P_n^R(i) = \beta^{i-m}(1-\beta)^{n-i+m}\binom{n-m}{i-m} \tag{16}$$

其中，β 是与错误模型相对应的点不在采样点集合的概率，并且包含一个类外点。采用极大性定义满足置信度需要的采样次数。

步骤 3：计算匹配点对的正确匹配率 T。当正确匹配率大于 50%且正确匹配点对 K 大于 200 时匹配结束，提取的特征区域可以作为图像信息丢失的补充，可以在后续建图中建立半稠密地图。当正确匹配率小于等于 50%或匹配点对小于等于 200 时，执行步骤 4。

步骤 4：匹配点的正确匹配率小于等于 50%或者匹配点对小于等于 200 时，认为匹配是不合格的，无法通过 ORB 特征点准确地进行定位。所以根据

前面提取到的特征区域和图像深度，运用直接法追踪对应的第二帧图像上对应的特征区域，达到特征匹配的目的。由于其特征点的不足，所以在运用直接法时只能运用半稠密直接法和稠密直接法，但考虑到算法的实时性，本文选取半稠密直接法。能在最大限度保证图像信息的同时保证算法实时性。同时采用灰度差平方和作为跟踪的依据。记 $I(x,y,t)$，$I(x,y,t+\tau)$ 分别为 t，$t+\tau$ 两个时刻的图像，且符合以下要求：

$$I(x-\mathrm{d}x,y-\mathrm{d}y,t)=I(x,y,t+\tau) \quad (17)$$

定义误差 e_{d} 为

$$e_{\mathrm{d}}=\iint_W[I_2(x+\mathrm{d}x,y+\mathrm{d}y)-I_1(x,y)]^2\cdot \omega(x,y)\mathrm{d}(x,y)\mid(x,y)\subset I(p) \quad (18)$$

其中，$\omega(x,y)$ 为特征区域加权函数，一般取 1。对其进行泰勒展开并保留前两项，当误差 e_{d} 导数为 0 时，其灰度误差最小。即满足：

$$\frac{\partial e_{\mathrm{d}}}{d}\approx\iint_W[I_2(x,y)-I_1(x,y)+g^{\mathrm{T}}d]\cdot g\omega(x,y)\mathrm{d}(x,y)=0 \quad (19)$$

其中，$g=\left[\frac{\partial}{\partial x}\left(\frac{I_1+I_2}{2}\right),\frac{\partial}{\partial y}\left(\frac{I_1+I_2}{2}\right)\right]^{\mathrm{T}}$ 为泰勒展开一阶系数。由于在进行泰勒展开式时引入误差，使用高斯牛顿法迭代消除误差。

步骤 5：对特征区域进行非线性优化。优化相机位姿以及相邻两帧中特征区域的光度误差。在直接法中非线性优化相当于优化一个相机位姿顶点与许多条一元边的问题。由于其对应的线性方程为计算李代数增量，所以主要计算每条边的误差以及雅可比矩阵。总体优化问题可以定义为

$$\min F(x)=\sum_{k=1}^{n}e_k(x_k)^{\mathrm{T}}\boldsymbol{\Omega}_k e_k(x_k) \quad (20)$$

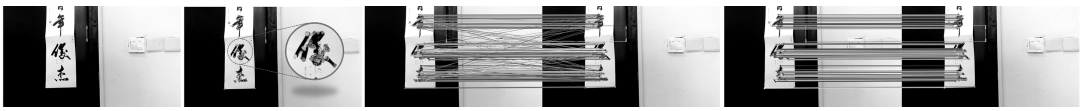
其中，x 为优化变量，$F(x)$ 为优化函数，e_k 为误差，信息矩阵 $\boldsymbol{\Omega}$ 是协方差矩阵的逆，是一个对称矩阵，x_k 为顶点。

图 5 Laboratory 实拍数据集中的 16 帧代表图像

3 实验结果

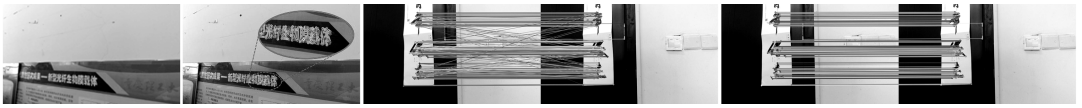
使用 ORB 算法和本文算法进行对比实验。通过 Kinect V2 在实验室采集 500 帧图像数据集（Laboratory），数据集中包含特征丰富和特征缺失区域。分别从数据集中选取特征丰富和特征缺失进行实验。图 5 为整个 Laboratory 实拍数据集中的 16 帧代表图像。图 6 和图 7 表示在 Laboratory 实

拍数据集中特征丰富的情况下，分别用 ORB 算法与本文算法进行对比试验。图 8 和图 9 表示在 Laboratory 实拍数据集中特征缺失的情况下，分别用 ORB 算法与本文算法进行对比试验。



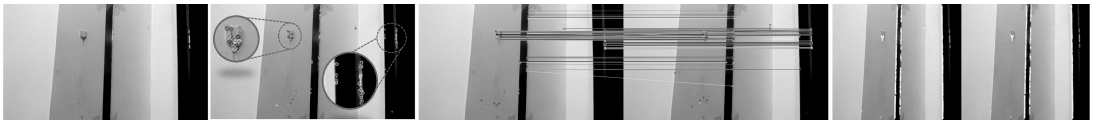
(a) 预处理后的图像　(b) ORB 特征点　(c) ORB 算法运算结果　(d) 本文算法运算结果

图 6　Laboratory 实拍数据集在特征点丰富时 ORB 算法和本文算法对比结果



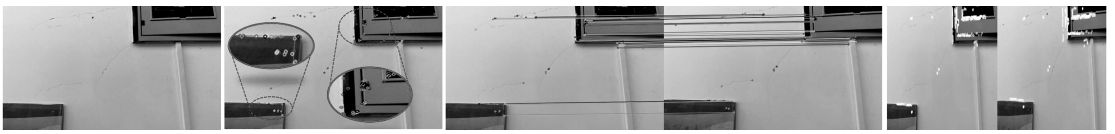
(a) 预处理后的图像　(b) ORB 特征点　(c) ORB 算法运算结果　(d) 本文算法运算结果

图 7　Laboratory 实拍数据集在特征点丰富时 ORB 算法和本文算法对比结果



(a) 预处理后的图像　(b) ORB 特征点　(c) ORB 算法运算结果　(d) 本文算法运算结果

图 8　Laboratory 实拍数据集在特征点缺失时 ORB 算法与本文算法对比结果



(a) 预处理后的图像　(b) ORB 特征点　(c) ORB 算法运算结果　(d) 本文算法运算结果

图 9　Laboratory 实拍数据集在特征点缺失时 ORB 算法与本文算法对比结果

4　实验结果分析

分别使用 ORB 算法和本文算法对所采数据集(Laboratory)进行实验，对实验结果抽取其中四段数据进行分析，每段数据选取 50 帧图像。结果如图 6～9，分别将其对应的 50 帧图像数据命名为 RF-A、RF-B、MF-A 和 MF-B。设定 ORB 算法初始提取特征总数为 500。在特征点丰富的条件下，对实验进行数据分析，分别测得表 1 和图 9 数据结果。

图 10 中 RF-A-R、RF-B-R 分别表示 ORB 算法在数据 RF-A 和 RF-B 的特征匹配的匹配率，RF-A-S、RF-B-S 分别表示本文算法在数据 RF-A 和 RF-B 的特征匹配的匹配率。当图像特征丰富时，本文算法直接提取特征点作为图像跟踪的依据，在特征匹配的同时加入剔除误匹配算法，使特征匹配误差小且误差稳定。根据表 1 和图 10 的数据结果，ORB 算法表现出较大误差，误差范围从百分之十几到百分之几十，表明 ORB 算法在特征匹配时存在误差较大且误差不稳定的情况。而根据数据可以看出本文算法误差基本都在 2%左右，特征匹配的误差稳定且误差小。

表 1　特征点丰富条件下，ORB 算法与本文算法特征匹配结果

数据集	ORB 算法匹配点对	ORB 算法正确匹配点对	ORB 算法正确匹配率/%	本文算法匹配点对	本文算法正确匹配点对	本文算法正确匹配率/%
RF-A	500	418	83.6	421	418	99.3
RF-B	500	352	70.4	335	331	98.8

在图像特征缺失的情况下，分别使用 ORB 算法和本文算法对数据 MF-A、MF-B 进行实验，测得图 11 为 ORB 算法特征匹配结果，图 12 为本文算法特征区域追踪结果。

图 11 使用 ORB 分别对数据 MF-A、MF-B 进行特征匹配。结果显示，当特征缺失或者特征质量不佳时，ORB 算法的特征正确匹配率较低，无法满足对特征的追踪以及通过特征点对相机进行定位。

图 12 表示分别在数据 MF-A、MF-B 下使用本文算法的特征区域追踪结果。可以看出，本文算法能在 ORB 算法失效的情况下很好地跟踪特征区域，根据图 11 和 12 分析在前 50 帧图像中，本文算法跟丢率分别在 9.69%和 8.08%左右，而 ORB 算法的特征误匹配率分别在 48%和 63%左右。因此本文算法在特征信息缺失的条件下，可以更好地跟踪特征区域。

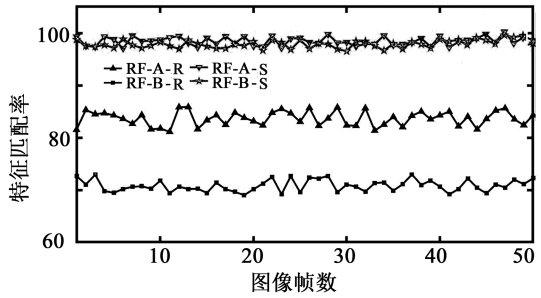


图 10 RF-A、RF-B 数据分别在 ORB 算法与本文算法的匹配率对比图

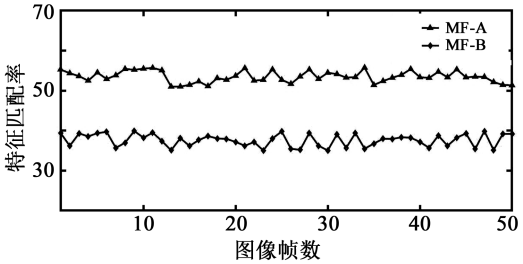


图 11 MF-A、MF-B 数据在 ORB 算法中的特征匹配结果

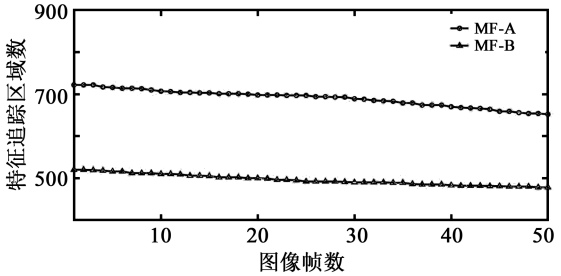


图 12 MF-A、MF-B 数据在本文算法中特征区域跟踪结果

5 结论

为了提高视觉 SLAM 在间断纹理中特征提取和追踪的精确性和鲁棒性，本文在分析 ORB 算法和半稠密直接法的优缺点后，提出了一种针对间断纹理环境中的特征匹配与追踪算法。在传统 ORB 算法的基础上添加了 PROSAC 算法用于剔除误匹配点，同时对图像进行特征区域提取，用于特征缺失环境中进行图像特征追踪，使用非线性优化，最后提取的特征区域在后端建图时也可以使用，进而理论上解决了 ORB 特征点只能构建稀疏地图的问题。与传统的特征提取算法 ORB 相比，本文算法能很好地工作在间断纹理的环境中，在特征充足时，比 ORB 算法具有更精确的匹配率和稳定性，在特征缺失时，能很好地追踪特征区域，具有较低的特征丢失率。通过现实环境数据进行验证，证实了本文算法在不同纹理环境中具有更高的准确性和鲁棒性。

参考文献：

[1] 杨 洋，童东兵，陈巧玉. 面向未知地图的六足机器人路径规划算法[J]. 计算机应用，2018，38(6)：289-293，299.
Yang Yang，Tong Dongbing，Chen Qiaoyu. Six-legged robot path planning algorithm for unknown map[J]. J. of Computer Applications，2018，38(6)：289-293，299.

[2] Garciafidalgo E，Ortiz A. Vision-Based Topological Mapping and Localization Methods：A Survey[M]. North Holland：North-Holland Publishing Co.，2015，64(1)：1-20.

[3] Carlone L，Tron R，Daniilidis K，et al. Initialization techniques for 3D SLAM：A survey on rotation estimation and its use in pose graph optimization[C]// IEEE Inter. Conf. on Robotics & Automation，2015：4597-4604.

[4] Zuo X，Xie X，Liu Y，et al. Robust visual SLAM with point and line features[J]. arXiv：1711.08654，2017：1775-1782.

[5] Karlsson N，Bernardo E D，Ostrowski J P，et al. The VSLAM algorithm for robust localization and mapping[C]// IEEE Inter. Conf. on Robotics & Automation，2006：24-29.

[6] Fuentes-Pacheco J，Ruiz-Ascencio，Jose Rendon-Mancha，et al. Visual simultaneous localization and mapping：A survey[J]. Artificial Intelligence Rev.，2015，43(1)：55-81.

[7] Endres F，Hess J，Sturm J，et al. 3-D mapping with an RGB-D camera[J]. IEEE Trans. on Robotics，2014，30(1)：177-187.

[8] Scaramuzza D，Fraundorfer F. Visual odometry：Part Ⅰ：The first 30 years and fundamentals[J]. IEEE Robotics Automation Magazine，2011，18(4)：80-92.

[9] Nister D，Naroditsky O，Bergen J. Visual odometry[C]// 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition，2004：652-659.

[10] 张峻宁，苏群星，刘鹏远，等. 一种自适应特征地图匹配的改进 VSLAM 算法[J]. 自动化学报，2019，45(3)：111-123.
Zhang Junning，Su Qunxing，Liu Pengyuan，et al. An improved VSLAM algorithm based on adaptive feature map[J]. Acta Automatica Sinica，2019，45(3)：111-123.

[11] Engel J，Thomas Schops，Cremers D. LSD-SLAM：Large-scale direct monocular SLAM[C]// Euro. Conf. on Computer Vision，2014.

[12] Mur-Artal R，Montiel J M M，Tardos J D. ORB-SLAM：A versatile and accurate monocular SLAM system[J]. IEEE Trans. on Robotics，2015，31(5)：1147-1163.

（下转第 140 页）

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.028

基于非均匀有序直方图和锐化的红外图像增强

吴　婷

（中原工学院 计算机学院，郑州 450007）

摘　要： 由于受到光电测量设备硬件及环境的影响，采集到的红外图像普遍存在对比度低、噪声大等缺陷。为此，提出了一种基于非均匀有序直方图和锐化的红外图像增强算法。考虑到高动态范围红外图像的特性，该方法分为两部分：首先，通过引入在有序直方图域中以恒定斜率增加的非均匀分布来增强图像对比度，并将非均匀分布直方图与输入图像的排序直方图合并于有序直方图域的优化问题，以计算修正后的直方图。其次，为了更有效处理红外图像的模糊效应，提出一种用于高频补偿滤波的相对边缘强度指标，以有效地抑制相对均匀区域中的噪声。实验测试结果表明，与传统算法和其他类似算法相比，使用所提出的方法增强后的图像细节丰富，具有更好的视觉效果和更高的实际应用价值。

关键词： 红外图像；图像增强；对比度增强；图像锐化；非均匀有序直方图

中图分类号： TP391　　**文章编号：** 1001－5868(2020)01－0135－06

Infrared Image Enhancement Based on Non-uniform Ordered Histogram and Sharpening

WU Ting

(School of Computer Science, Zhongyuan University of Technol., Zhengzhou 450007, CHN)

Abstract: Due to the influence of the hardware and environment of optoelectronic measuring equipments, the collected infrared images generally have problems of low contrast and big noise. Thus, an infrared image enhancement method based on non-uniform ordered histogram and sharpening is proposed. Considering the characteristics of high dynamic range infrared images, the method is divided into two parts. Firstly, image contrast is enhanced by introducing a non-uniform distribution that increases at a constant slope in the ordered histogram domain. The non-uniform distribution histogram and the ordered histogram of the input image are merged into the optimization problem to calculate the corrected histogram. Secondly, in order to effectively deal with the blur effect of infrared image, a relative edge intensity index for high frequency compensation filtering is proposed to effectively suppress the noise in the relatively uniform region. The experimental results show that compared with various traditional algorithms and other similar algorithms, the enhanced image details of the proposed method are rich, with better visual effects and practical application value.

Key words: IR image; image enhancement; contrast enhancement; image sharpening; uniform ordered histogram

0　引言

随着红外技术的不断进步与发展，红外热像仪在社会各个行业得到非常广泛的应用，如安全监视、军事目标检测和跟踪，以及驾驶员辅助夜视等方面。通常，用于远程监视的红外(Infrared，IR)影像仪通过检测热辐射来创建IR图像[1-3]，但热辐射很容易

收稿日期：2019－11－18.

基金项目：国家自然科学基金项目(U1404606).

受到空气和水蒸气等物理环境因素的影响，从而产生散射和反射。与可见光图像相比，IR 图像所产生的热图像具有容易模糊的特性。此外，在 IR 图像中大多数来自环境背景的热辐射会变成干扰的灰度级信息。由于这种背景辐射的存在，原始 IR 图像灰度直方图通常会出现一个较为陡峭的峰值现象，从而产生较大的噪点。

为了有效地对原始 IR 图像进行处理，以便提供更好的视觉效果和图像细节信息，图像增强算法应具有以下功能[4]：1)需要在将 12～14 位级高动态范围压缩到 8 位域的同时，保留原始 IR 图像的完整轮廓信息和细节信息；2)需要在不引入图像伪影的情况下，产生视觉表现较好的场景表示；3)具有较高的计算效率，即计算复杂度较低。现有的自动增益控制(Automatic Gain Control，AGC)方法会删除直方图中最外侧的一些极值，并将动态范围线性压缩到 8 位域中[5]。但是这种方法不能有效地表现出足够的细节信息；直方图均衡(Histogram Equalization，HE)方法的关键是传递函数的映射计算，主要采用了输入累积分布函数(Cumulative Distribution Function，CDF)和均匀分布 CDF[6]。但如果直方图中存在较大的峰值，则可能会过度增强图像，从而造成增强图像出现较大的噪点和褪色现象。

为了解决上述缺陷，平稳 HE(Plateau Histogram Equalization，PHE)方法使用一个平稳值将直方图中的高峰剪切掉，以减少均匀区域内的噪声放大[7]。自适应双平稳直方图均衡化(Adaptive Double Plateau Histogram Equalization，ADPH)在直方图中使用一个移动窗口来找到局部最大值，并通过求平均值来计算平稳值[8]。最近，基于 CLAHE 提出了自适应三边对比度增强(Adaptive Trilateral Contrast Enhancement，ATCE)[9]。这些方法大多使用基于平稳值的 HE 来解决红外图像增强问题。但是，这些仍不能有效地解决直方图中相同阈值水平引起的对比度降低问题。换句话说，为了有效地获得对比度增强的图像，在阈值处理中属于高像素数的灰度级应该具有比低像素数更高的阈值。

基于上述分析，本文引入了非均匀分布策略，以恒定斜率递增的分布方式，解决平稳 HE 所引起的对比度降低问题，并在此基础上，提出了一种基于非均匀分布的红外图像增强算法。首先，在有序直方图域中使用斜坡分布的直方图，该直方图与输入图像的排序直方图合并于优化问题，以生成修正后的直方图。其次，提出了一种相对边缘强度指数作为高增强核，以生成边缘增强图像，同时抑制相对均匀区域中的噪声。实验结果表明，该算法具有良好的定性和定量性能。

1 提出的算法

1.1 非均匀分布的可行性分析

首先，通过研究三个数据集(红外图像增强实验常用数据集)中原始图像的归一化直方图，分析有序直方图域中最合适的分布方案。如表 1 所示，非均匀分布的均方根误差(RMSE)平均值比均匀分布的要小。这意味着所需的输出图像应该更加接近于有序直方图域中的非均匀分布，验证了本文提出研究思路的可行性。

表 1 每个数据集的平均 RMSE 值

数据集	USC-SIPI[9]	TID2013[10]	CSIQ[11]
均匀分布	4.583×10^{-3}	4.778×10^{-3}	5.762×10^{-3}
非均匀分布	2.767×10^{-3}	3.046×10^{-3}	4.585×10^{-3}

1.2 基于非均匀分布的对比度增强

令 $\boldsymbol{X}=\{\boldsymbol{X}(i,j)\mid 1\leqslant i\leqslant H,1\leqslant j\leqslant W\}$ 表示像素大小为 $H\times W$ 的输入图像，其中 $\boldsymbol{X}(i,j)$ 是像素 (i,j) 的灰度级，并假设 $\boldsymbol{X}$ 的动态范围为 $[x_d,x_u]$，其中 $\boldsymbol{X}(i,j)\in[x_d,x_u]$。目的是产生增强图像 $\boldsymbol{Y}=\{\boldsymbol{Y}(i,j)\mid 1\leqslant i\leqslant H,1\leqslant j\leqslant W\}$。$\boldsymbol{Y}$ 的动态范围被压缩到区间 $[y_d,y_u]$，其中 $\boldsymbol{Y}(i,j)\in[y_d,y_u]$，$y_d<y_u$。使用 8 位图像的整个动态范围，$y_d=0$ 并且 $y_u=255$。令 $\boldsymbol{X}'=\{x_1,x_2,\cdots,x_K\}$ 是输入图像 $\boldsymbol{X}$ 中所有可能的 K 灰度级排序集合，其中 $x_1<x_2<\cdots<x_K$，其中 K 是不同灰度级的数量。在对比度增强之后，每个输入灰度级 x_K 映射到输出灰度级 y_K，以产生输出灰度级的排序列表 $\boldsymbol{Y}'=\{y_1,y_2,\cdots,y_K\}$，其中 $y_1<y_2<\cdots<y_K$。

灰度级 x_k 的直方图的计算方式为

$$\boldsymbol{H}_x=\{h_x(k)\mid 1\leqslant k\leqslant K\} \tag{1}$$

其中 $h_x(k)$ 表示 $\boldsymbol{X}$ 中具有灰度级 k 的像素总数。令 $\boldsymbol{H}_s=\{h_s(l)\mid 1\leqslant l\leqslant L\}$ 表示所有非零 $h_x(k)$ 的升序排序集，其中 $h_s(l)$ 为该排序集的元素，L 为所有非零 $h_x(k)$ 的总数。令 $\boldsymbol{H}_r=\{h_r(l)\mid 1\leqslant l\leqslant L\}$ 为以恒定斜率增加的非均匀分布直方图，其中 $\sum_{l=1}^{L}h_r(l)=H\times W$ 并且 $h_r(l)$ 的计算方法如下所

示：

$$h_r(l)=\frac{2HW}{L^2}l \tag{2}$$

目标直方图 $\boldsymbol{H}_t=\{h_t(l)\mid 1\leqslant l\leqslant L\}$ 应与排序后的直方图之间保持最小距离，即

$$\boldsymbol{H}_t=\underset{\boldsymbol{H}}{\operatorname{argmin}}\|\boldsymbol{H}-\boldsymbol{H}_s\| \tag{3}$$

其中 $\|\cdot\|$ 表示计算范数。$\boldsymbol{H}_t$ 与非均匀分布直方图之间的距离应最小，即

$$\boldsymbol{H}_t=\underset{\boldsymbol{H}}{\operatorname{argmin}}\|\boldsymbol{H}-\boldsymbol{H}_r\| \tag{4}$$

为了确定目标直方图 $\boldsymbol{H}_t$，将所有条件组合到以下优化函数中：

$$\boldsymbol{H}_t=\underset{\boldsymbol{H}}{\operatorname{argmin}}\|\boldsymbol{H}-\boldsymbol{H}_s\|+\alpha\|\boldsymbol{H}-\boldsymbol{H}_r\| \tag{5}$$

其中，α 是加权因子，范围为 $[0,\infty)$。可以通过调整 α 来实现各种级别的对比度增强。将范数运算替换为欧几里德范数的平方即可得到最小化式(5)的闭式解。

$$\boldsymbol{H}_t=\underset{\boldsymbol{H}}{\operatorname{argmin}}\|\boldsymbol{H}-\boldsymbol{H}_s\|_2^2+\alpha\|\boldsymbol{H}-\boldsymbol{H}_r\|_2^2 \tag{6}$$

或相当于

$$\boldsymbol{H}_t=\underset{\boldsymbol{H}}{\operatorname{argmin}}[(\boldsymbol{H}-\boldsymbol{H}_s)^{\mathrm{T}}(\boldsymbol{H}-\boldsymbol{H}_s)+\alpha(\boldsymbol{H}-\boldsymbol{H}_r)^{\mathrm{T}}(\boldsymbol{H}-\boldsymbol{H}_r)] \tag{7}$$

二次优化问题式(7)的解为

$$\boldsymbol{H}_t=\frac{1}{1+\alpha}\boldsymbol{H}_s+\frac{1}{1+\alpha}\boldsymbol{H}_r \tag{8}$$

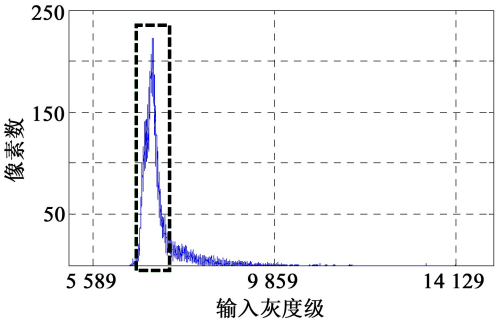
根据 $h_m(k)=h_t\{\Omega(k)\}$ 计算修正后的直方图 $\boldsymbol{H}_m=\{h_m(k)\mid 1\leqslant k\leqslant K\}$，其中 $\Omega(k)$ 是从排序过程中获得的 l 与 k 之间的索引映射函数。对修正后的直方图 $\boldsymbol{H}_m$ 进行归一化，以便给出概率分布函数：

$$p(k)=\frac{h_m(k)}{\sum_{k=1}^{K}h_m(k)} \tag{9}$$

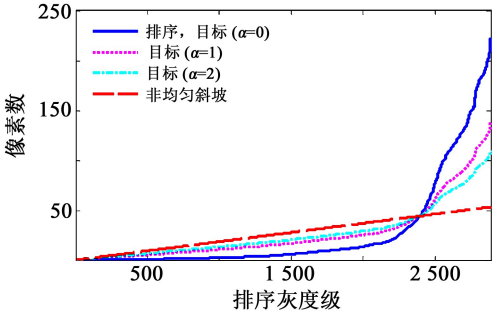
则 CDF 定义为 $P(k)=\sum_{k=1}^{K}p(k)$，使用 CDF $P(k)$，并根据 $y_k=\lfloor P(k)(2^8-1)+0.5\rfloor$ 将 x_k 映射到 y_k。

图 1 对所提出的算法进行了说明。红外图像如图 1(a)所示，其中箭头表示输入图像的背景。原始输入图像的直方图如图 1(b)所示，虚线框表示来自背景的最高峰。直方图中的非零灰度级按升序排序，如图 1(c)所示，包括有序直方图、目标直方图和斜坡分布直方图。三个 α 值的映射如图 1(d)所示，从图中可以看出，随着 α 的增大，映射变得与 $\boldsymbol{H}_r$ 映射更加相似。当 α 为零时，修正后的直方图等于输入直方图。因此，应用传统 HE 会使该图像被过度增强，包含非自然信息。当 α 为 1 时，增强水平进一步降低，并且大部分信息得以保留。

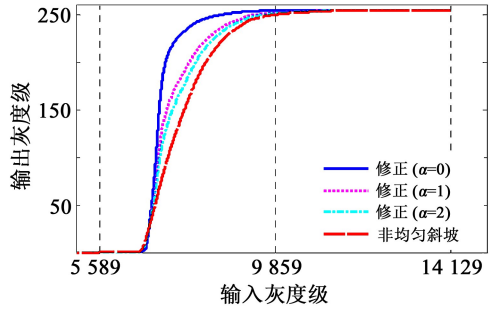
(a)　红外图像



(b)　输入图像的直方图



(c)　直方图中的非零灰度级



(d)　三个 α 值的映射

图 1　提出算法的原理说明

1.3　基于相对边缘强度指数的图像锐化

锐化的目的是增强模糊处理效果，以便体现红外图像中自然效果的细节信息。通常，图像锐化不仅会突出边缘信息，同时也会导致相关噪声放大。

在红外图像存在的情况下，与其输入直方图中的高峰值相对应的背景可以视为相对均匀区域(Relatively Uniform Regions，RUR)。RUR 在其输出直方图中映射到较宽的范围。因此，在锐化过程中很容易放大 RUR 的噪声。为了抑制 RUR 中的噪声放大，提出了从相对边缘强度指数(Relative Edge-strength Index，RESI)导出的高频补偿内核。令 $\boldsymbol{E}=\{e(i,j)\mid 1\leqslant i\leqslant H,1\leqslant j\leqslant W\}$ 表示 RESI，其中 $e(i,j)$ 计算方式为

$$e(i,j)=1-\sum_{u}\sum_{v}\frac{g\{x(i,j),x(i+u,j+v)\}}{8} \tag{10}$$

其中 $u,v\in\{-1,0,1\}$ 且 $(u,v)\neq(0,0)$。$g(a,b)$ 是一个函数，定义如下：

$$g(a,b)=\geqslant\begin{cases}1, & \text{if } |a-b|\leqslant T\\ 0, & \text{otherwise}\end{cases} \tag{11}$$

其中 T 是阈值。为了抑制 RUR 中的噪声放大，T 应该大于 RUR 中噪声的灰度级差。由于当 $\boldsymbol{H}_s$ 中的 l 值接近 L 时，具有较高的 RUR 概率。因此，将值 T 设置为 $T=L-l^*$，其中 l^* 的定义如下：

$$l^*=\underset{n\in\{1,2,\cdots,L\}}{\operatorname{argmin}}\left\{\sum_{l=1}^{n}\frac{\boldsymbol{H}_s(l)}{W\times H}\right\} \tag{12}$$

自适应高频补偿内核 $\boldsymbol{K}_{nb}$ 的计算如下：

$$\boldsymbol{K}_{nb}=\begin{bmatrix}-c_1 & -c_2 & -c_3\\ -c_4 & C & -c_5\\ -c_6 & -c_7 & -c_8\end{bmatrix},C=1+\sum_{i=1}^{8}|c_i| \tag{13}$$

其中 $i\in\{1,2,\cdots,8\}$，c_i 表示来自 RESI 的非中心系数，C 表示中心加权系数。根据 RESI 使用具有不同系数的高频补偿核，可以在抑制 RUR 中噪声的同时获得边缘增强的图像。通过对 $\boldsymbol{Y}$ 和 $\boldsymbol{K}_{nb}$ 进行卷积可以获得锐化图像 $\boldsymbol{Y}_s$。

2 实验结果

本文使用表 2 列出的两个数据集进行了具体实验。从每个数据集中选取了 50 张图像。为了评估所提出算法的性能，将其与 ATCE[9]，HMF[10] 和 SMIR[11] 进行了比较。本文算法参数设置为 $\alpha=1$。所有算法均在 MATLAB R2014a 中实现。测试程序在 64 位 Windows7 的个人计算机上进行，配置为 16 GB RAM 和 Intel Core i7-4940MX 3.3 GHz CPU。

表 2 数据集的基本参数信息

数据集	摄影机	数量	位深度	分辨率	频带
DB1	Merlin	825	12	312×255	MWIR
DB2	Tau320	522	14	324×256	LWIR

2.1 定量评估

为了对所提图像增强算法的客观性能进行具体分析，使用了三个评估指标：离散熵(DE)[12]、指标像素距离(PixDist)[13] 和质量感知的相对对比度度量(QRCM)[11]。DE 用来测量图像中的信息内容，PixDist 用来测量图像中所有像素对的平均灰度差异，QRCM 为综合评估指标。由于篇幅有限，三个量化评估指标的计算方在此不做详细表述，三个指标均是数值越大，图像增强效果越好。

不同图像增强算法的定量评估结果如表 3 所示。从表中可以更清楚地看到，提出的算法在每个指标中的排名都很高，具有良好的性能。值得注意的是，就整体对比度质量而言，QRCM 指标与主观评价之间存在良好的相关性。所有结果表明，提出的算法获得了最佳的客观性能。

表 3 图像增强算法的定量评估结果

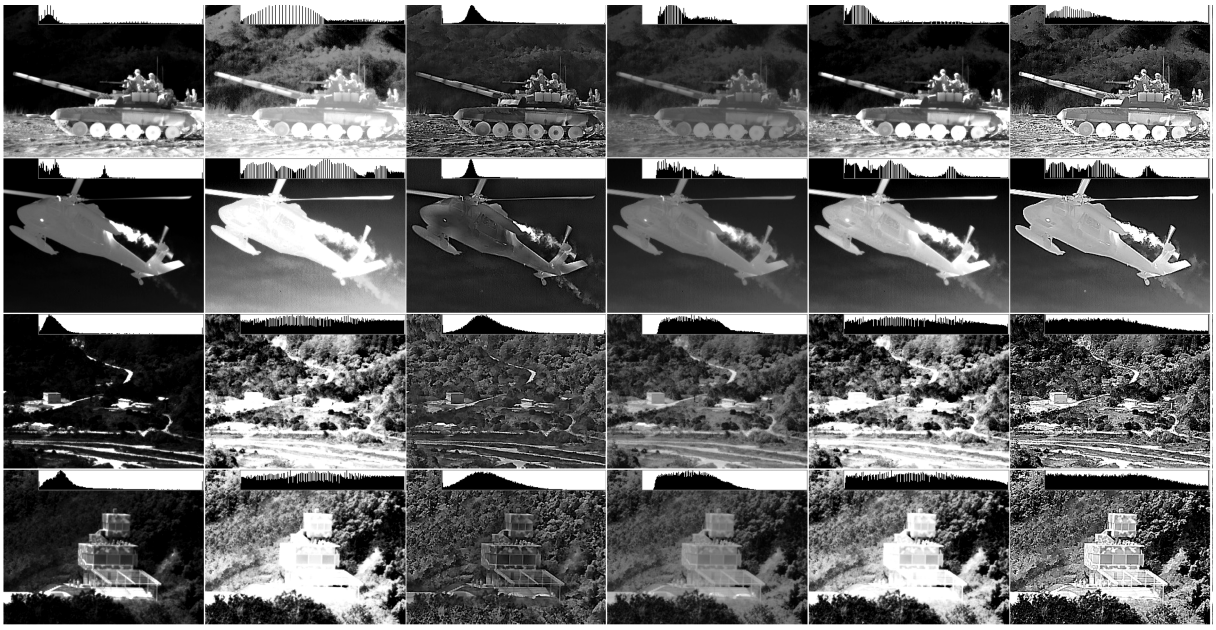
数据集	评估指标	ATCE	HMF	SMIR	本文算法
DB1	DE	7.41	6.379	6.93	7.526
	PixDist	29.961	27.965	38.635	40.646
	QRCM	0.18	0.083	0.101	0.185
DB2	DE	6.972	6.188	6.976	7.341
	PixDist	29.249	27.023	38.902	42.397
	QRCM	0.104	0.031	0.192	0.232

2.2 定性评估与分析

对 4 个对象进行了主观质量评估测试，如图 2 所示。其中，AGC 的输出图像可用于完全参考指标的参考图像，因为它可以很好地保留原始直方图的形状。需要注意的是，原始图像由于其较高的位深度而不能用作参考图像。

图 2 的第一行以单峰直方图显示了坦克图像上的结果。HE 产生了过亮的图像。AGC，ATCE 和 HMF 无法充分增强背景的对比度。HMF 的输出图像看起来很模糊，因为 HMF 的输入和均匀直方图的加权平均值无法消除输入直方图中尖峰的影响。SMIR 和本文算法增强了整体对比度。图 2 的第二行显示了具有多峰直方图的直升机拍摄图像的结果。尽管 ATCE 具有较好的主观效果，但由于 ATCE 是基于本地 HE 算法的，因此在相对均匀的区域中显示模糊。SMIR 和本文算法可以产生良好

的对比度图像。图 2 的第三行显示了较高复杂场景的村落图像的结果。与 AGC 相比，HE 和 SMIR 输出过亮，因此，图像中的建筑物看起来很模糊。第四行的塔楼，具有中度复杂的场景。ATCE 通过基于子块进行操作而使 IR 图像失去了所检测场景温度的相对信息，而其余方法则将其保留下来。



(a)　AGC　　(b)　HE　　(c)　ATCE　　(d)　HMF　　(e)　SMIR　　(f)　本文

图 2　红外图像增强的实验结果

综合图 2 所示结果，可以看出，提出算法获得了令人满意的对比度改进，在压缩时保留了原始红外图像的完整信息和详细细节。

最后，在表 4 给出了不同算法的平均计算时间。注意到，SMIR 的速度在很大程度上取决于唯一的灰度级数，因此，需要花费大量时间。ATCE 使用相位一致性，其计算复杂并且占用大量内存。尽管相比于 HMF，所提出算法的处理时间稍慢，但与以前的算法相比，它提供了更好的综合增强性能。

表 4　平均计算时间(单位：ms)

数据集	ATCE	HMF	SMIR	本文
DB1	893	10	10,826	54
DB2	929	13	5,716	53

3　结论

本文针对原始红外图像提出了一种基于非均匀分布的图像增强算法。在有序直方图中对非均匀分布作了说明，并表明与使用均匀分布相比，它是对比度增强图像的适当模型。将输入图像的排序直方图和非均匀分布直方图合并到有序直方图域的优化问题中，以计算修正后的直方图。另外，提出了一种锐化方法，该方法可以增强边缘，同时有效抑制相对均匀区域中的噪声。实验结果表明，与现有算法相比，提出的算法可获得更具有可比性的图像增强效果和感知质量，并同时兼顾了计算效率。

参考文献：

[1] Dong Y W, Lei S, No U. Design and application of an infrared image super-resolution reconstruction system [J]. Instrumentation Technol., 2014, 7(2): 243-246.

[2] He Fuliang, Guo Yongcai, Gao Chao. Human segmentation of infrared image for mobile robot search[J]. Multimedia Tools & Applications, 2018, 77(9): 10701-10714.

[3] Davis M J, Ihinger P D. New controlled rapid quench technique in a 1 atm infrared image furnace[J]. American Mineralogist, 1999, 84(1/2): 48-54.

[4] Tan W, Zhao C, Wu H. CNN intelligent early warning for apple skin lesion image acquired by infrared video sensors[J]. High Technol. Lett., 2016, 22(1): 67-74.

[5] Zhang Hengwei, Ji Xiang, Zhao Wei, et al. Jamming effect of laser on infrared thermal imagers automatic gain control circuit [J]. High Power Laser & Particle Beams, 2011, 23(9): 2325-2328.

[6] Wu Wei, Yang Xiaomin, Li Hong, et al. A novel scheme for infrared image enhancement by using weighted least squares filter and fuzzy plateau histogram equalization[J]. Multimedia Tools & Applications, 2017, 76(54): 1-29.

[7] 龚昌来，罗　聪，杨冬涛，等. 一种基于平稳小波域的红外图

像增强方法[J]. 激光与红外，2013，43(6)：703-707.
Gong Changlai，Luo Cong，Yang Dongtao，et al. Infrared image enhancement method based on stationary wavelet domain [J]. Laser & Infrared，2013，43(6)：703-707.
[8] 陈永亮，王华彬，陶 亮. 自适应动态峰值剪切直方图均衡化[J]. 计算机工程与应用，2015，51(1)：167-171.
Chen Yongliang，Wang Huabin，Tao Liang. Adaptive dynamic clipped histogram equalization title[J]. Computer Engin. and Applications，2015，51(1)：167-171.
[9] Tzer Y L，Kok S S，Chih P T. Infrared image enhancement using adaptive trilateral contrast enhancement[J]. Pattern Recognition Lett.，2015，54：103-108.
[10] Khan M F，Khan E，Abbasi Z A. Image contrast enhancement using normalized histogram equalization[J]. Optik-Inter. J. for Light and Electron Opt.，2015，126(24)：4868-4875.
[11] Celik T. Spatial mutual information and pagerank-based contrast enhancement[J]. IEEE Trans. on Image Proc.，2016，25(10)：4719-4728.
[12] 赵 凡，张 葆，赵 建，等. 基于分层最大熵的航拍图像增强[J]. 光电子·激光，2015，26(1)：192-198.
Zhao Fan，Zhang Bao，Zhao Jian，et al. The aerial image enhancement algorithm based on histogram specification induced by the adaptive function[J]. J. of Optoelectronics·Laser，2015，26(1)：192-198.
[13] Geng A，Wan C，Li Y，et al. Image visual enhancement based on layered difference representation [J]. J. of Electronics & Information Technol.，2017，39(4)：922-929.

作者简介：

吴 婷(1981－)，女，河南郑州人，硕士，讲师，研究方向为图像处理、人工智能和网络安全技术。

E-mail：wuting@zut.edu.cn

(上接第 134 页)

[13] Jurgen Sturm，Engelhard N，Endres F，et al. A benchmark for the evaluation of RGB-D SLAM systems[C]// 2012 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems，2012：573-580.
[14] Lowe D G. Object recognition from local scale-invariant features[C]// ICCV IEEE Computer Society，1999：1150.
[15] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. Inter. J. of Computer Vision，2004，60(2)：91-110.
[16] Bay H，Tuytelaars T，Gool L V. SURF：Speeded up robust features[C]// Euro. Conf. on Computer Vision. Springer-Verlag，2006：404-417.
[17] Rublee E，Rabaud V，Konolige K，et al. ORB：An efficient alternative to SIFT or SURF[C]// 2011 IEEE Inter. Conf. on Computer Vision，2012：2564-2571.
[18] Forster C，Pizzoli M，Scaramuzza D. SVO：Fast semi-direct monocular visual odometry[C]// IEEE Inter. Conf. on Robotics and Automation (ICRA)，2014.

作者简介：

赵明富(1964－)，男，重庆人，博士，二级教授，硕士生导师，电子科技大学兼职博士生导师，研究方向为光纤生物化学传感、光电信息获取与智能信息处理、现代光电检测技术；

宋 涛(1987－)，男，湖北人，副教授，硕士生导师，主要研究方向为双目视觉、视觉 SLAM、点云处理、三维重建等。

E-mail：tsong@cqut.edu.cn

(上接第 102 页)

Campbell S A. Micro-nano Scale Manufacturing Engineering [M]. Yan Liren，Zhang Wei，Transl. Beijing：Electronics Industry Press，2011.
[16] 张汝京. 纳米集成电路制造工艺[M]. 北京：清华大学出版社，2014：298-307.
Zhang Rujing. Nanometer Integrated Circuit Manufacturing Process[M]. Beijing：Tsinghua University Press，2014：298-307.

作者简介：

王 兴(1979－)，男，天津人，中芯国际(天津)工艺整合部副部长，研究方向为电源管理类芯片的研发与量产；

李维杰(1994－)，男，山东潍坊人，硕士研究生，研究方向为功率器件及工艺整合。

E-mail：1165149254@qq.com

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.029

联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测

张　炎[1]，华文深[1]，黄富瑜[1]，严　阳[2]，王强辉[1]，索文凯[3]

（1. 陆军工程大学石家庄校区 电子与光学工程系，石家庄 050003；
2. 解放军31681部队，甘肃 天水 741000；3. 解放军68129部队，兰州 730000）

摘　要： 针对低秩稀疏矩阵分解的高光谱异常目标检测算法忽略了图像的空间信息，导致检测精度低的问题，提出了一种联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测算法。算法综合利用了高光谱图像的光谱信号与空间信号，并与图像自身的稀疏性相结合，对经典的基于低秩稀疏矩阵分解的目标检测算法进行改进，该算法以待测像元为中心构建一定大小的空间窗，计算中心像元与邻域内其他像元的空间相似度权值和光谱相似度权值，通过计算邻域内其他像元对中心像元的比例权值得到了中心像元的重构光谱值并作差得到两者的残差矩阵；最后基于低秩稀疏矩阵分解的高光谱异常目标检测算法得到图像的稀疏矩阵，将代表异常目标信息的稀疏矩阵和残差矩阵相加并求解矩阵行向量之间的欧式距离得到像元的异常度，设置阈值，得到检测结果。为验证所提算法的检测性能，采用了真实的高光谱数据进行仿真实验，并与现有算法进行对比，结果表明该算法能够得到更高的检测精度。

关键词： 高光谱图像；异常目标检测；低秩稀疏矩阵分解；稀疏矩阵；残差矩阵

中图分类号： TP751　　**文章编号：** 1001－5868(2020)01－0141－05

Joint Spatial Information and Improved Low-rank and Sparse Matrix Decomposition-based Anomaly Target Detection for Hyperspectral Imagery

ZHANG Yan[1], HUA Wenshen[1], HUANG Fuyv[1], YAN Yang[2], WANG Qianghui[1], SUO Wenkai[3]

(1. Dept. of Electronic and Optical Engin., Army Engin. University, Shijiazhuang 050003, CHN;
2. Unit 31681 of PLA, Tianshui 741000, CHN; 3. Unit 68129 of PLA, Lanzhou 730000, CHN)

Abstract: The hyperspectral anomaly target detection algorithm for low rank sparse matrix decomposition ignores the problem that the spatial information of the image leads to low detection accuracy. In this paper, proposed is an improved low-rank sparse matrix decomposition method for hyperspectral anomaly target detection based on spatial information. The algorithm comprehensively utilizes the spectral signal and spatial signal of hyperspectral image, and combines with the sparsity of the image itself to improve the classical target detection algorithm based on low rank sparse matrix decomposition. The algorithm is built around the pixel to be measured. A spatial window of a certain size calculates both the spatial and spectral similarity weight of the central pixel and other pixels in the domain, and obtains the weight of the central pixel by calculating the proportional weight of other pixels in the domain to the central pixel. The spectral values are constructed and the residual matrix is obtained. Finally, the hyperspectral anomaly target detection algorithm based on low rank sparse matrix decomposition is used to obtain the sparse matrix of the image, and the sparse matrix and residual matrix representing the

收稿日期：2019－09－09.
基金项目：国家自然科学基金项目(61801507).

abnormal target information are added and solved. The euclidean distance between the matrix row vectors obtains the degree of abnormality of the pixels, sets the threshold, and obtains the detection result. In order to verify the detection performance of the proposed algorithm, real hyperspectral data is used for simulation experiments, and compared with the existing algorithms, higher detection accuracy can be obtained.

Key words: hyperspectral image; anomaly target detection; low rank sparse matrix decomposition; sparse matrix; residual matrix

0 引言

高光谱图像具有成百上千个连续窄的波段,能够获得地物丰富的光谱信息,在分类、解混以及目标检测等多个领域得到了广泛应用。高光谱目标检测根据先验知识的有无分为监督目标检测和异常目标检测两类;在战场侦察、搜救人质、生态检测等实际应用中,先验知识的缺乏使得异常目标检测更符合现实需求,并逐渐成为近年来的研究热点。

在异常目标检测中,图像场景中主要的地物类别被定义为背景,而与背景存在差异的小概率地物被视为异常目标。近几十年来,许多异常目标检测算法不断被提出和改进,其中,RX 算法[1-2](Reed-Xiaoli Detection,RXD)最为经典,它是高光谱异常目标检测算法的基准算法。RX 算法包括从整幅图像出发建立背景模型的全局 RX 算法(Global RX,GRX)和基于双窗口模型建立背景模型的局部 RX 算法(Local RX,LRX)。RX 算法假设背景模型符合多维高斯分布,通过求解待测像元与背景像元光谱均值之间的马氏距离的方法得到检测结果。由于背景模型的假设与大部分实际场景的分布差距较大以及只利用了数据的低阶信息等问题,导致检测精度不高。为解决上述问题,许多改进 RX 算法被提出,比如 Stein 等[3]提出了一种基于全局高斯混合模型(Global Gaussian Mixture Model,GGMM)的异常目标检测算法,首先对图像中的像元进行聚类并假定同一类的像元符合相同的高斯分布,通过计算每个像元与其最邻近类别之间的马氏距离进行检测,该算法适用于具有多种地物的复杂场景;Chang 等[4]提出了一种基于线性混合模型(Linear Mixture Model,LMM)的检测方法,假设每个像元能够被图像中的端元根据相应的丰度线性表示,并将高光谱数据投影在子空间中以抑制背景成分从而进行检测;Kwon 等[5]通过引入核函数提出了 KRX 算法(Kernel RX,KRX),将原始数据投影到高维特征空间中,能够挖掘出数据的高阶统计信息并提高检测精度。但 RX 算法及其改进算法具有一定的局限性,存在检测精度不高以及计算复杂度较高等问题。近年来,基于稀疏表示的高光谱异常目标检测成为了研究热点之一,大量的稀疏检测算法被应用到了异常目标检测中,并取得了显著的效果。Yuan 等[6]提出了一种基于局部稀疏差异指数(Local Sparsity Divergence Index,LSD)的高光谱异常目标检测算法,通过构建双窗口模型选取背景字典,得到了权值向量以及相应的稀疏差异指数来进行检测。Zhao 等[7]提出了一种基于稀疏得分估计的异常检测算法(Sparse Score Estimate Anomaly Detection,SSEAD),与局部稀疏差异指数不同的是背景字典是利用 K-SVD 算法得到,通过统计字典原子的利用率进行反演待测像元的异常度从而得到检测结果;由于该算法选取的背景字典在复杂环境下不够准确,唐意东等[8]在数据的主成分空间中构建背景字典,能够有效地提高检测精度。基于低秩与稀疏矩阵分解的异常检测算法(Low-Rank and Sparse Matrix Decomposition,LRaSMD)也是稀疏表示方面的一大热点[9],利用 GoDec 算法提取图像的低秩矩阵和稀疏矩阵,并通过计算像元稀疏分量之间的欧式距离得到检测结果。在此基础上,许多改进算法被提出并得到了很好的效果。

本文对 LRaSMD 算法进行了详细分析,并在其基础上提出了一种联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测算法(JSI-LRaSMD)。算法以待测像元为中心建立一定大小的空间窗,通过计算中心像元与其邻域像元的空间相似度权值和光谱相似度权值,得到中心像元重构后的光谱值以及图像的残差矩阵;最后,利用 LRaSMD 算法提取出高光谱图像的稀疏矩阵,将残差矩阵和稀疏矩阵进行加和后求解矩阵中行向量的欧式距离得到每个像元的异常度,设置阈值,从而得到算法的异常目标检测结果。

1　基于低秩与稀疏矩阵分解的异常检测算法

LRaSMD 算法的检测原理是将高光谱的图像矩阵分解为包含背景光谱信息的低秩矩阵、包含异常目标光谱信息的稀疏矩阵和包含噪声光谱信息的噪声矩阵，并利用稀疏矩阵进行目标检测。由于高光谱图像中的异常像元出现概率低，占比很小且分布极其分散，具有一定的稀疏性。同时，由于图像中的像元数远远大于实际地物的类别，且每个像元的光谱向量都能由少数的光谱基向量线性表示，因此存在低秩性的特性。LRaSMD 算法给出了低秩与稀疏矩阵的表示模型：

$$\boldsymbol{X}=\boldsymbol{B}+\boldsymbol{S}+\boldsymbol{N} \tag{1}$$

式中，$\boldsymbol{B}$ 为低秩矩阵，能够表达图像背景的光谱信息；$\boldsymbol{S}$ 为稀疏矩阵，表示异常目标的光谱信息；$\boldsymbol{N}$ 为噪声矩阵，假定图像中的噪声属于高斯随机变量。

LRaSMD 算法利用 GoDec 算法求解高光谱图像的稀疏矩阵和噪声矩阵，并且通过限制低秩矩阵 $\boldsymbol{B}$ 的秩和稀疏矩阵 $\boldsymbol{S}$ 的稀疏度来控制模型的复杂度，是一种快速近似方法。其原理是通过最小化分解误差进行求解，表示式为

$$\begin{aligned}&\underset{\boldsymbol{B},\boldsymbol{S}}{\operatorname{argmin}}\|\boldsymbol{X}-\boldsymbol{B}-\boldsymbol{S}\|_{\mathrm{F}}^{2},\ \text{s.t.}\ \operatorname{rank}(\boldsymbol{B})\leqslant r,\\&\operatorname{card}(\boldsymbol{S})\leqslant kN\end{aligned} \tag{2}$$

式中，r 和 k 分别代表低秩矩阵的秩和稀疏矩阵的稀疏度，N 为像元数。通过控制 r 的最大值能够有效地表达出图像的背景信息；而 k 能够反映图像中异常像元出现的概率。随着迭代次数的增加，分解误差呈单调减少的趋势，因此式(2)能够转化为两个子问题，表达式为

$$\begin{cases}B_t=\underset{\operatorname{rank}(\boldsymbol{B})\leqslant r}{\operatorname{argmin}}\|\boldsymbol{X}-\boldsymbol{B}-\boldsymbol{S}_{t-1}\|_F^2\\S_t=\underset{\operatorname{card}(\boldsymbol{S})\leqslant kN}{\operatorname{argmin}}\|\boldsymbol{X}-\boldsymbol{B}_t-\boldsymbol{S}\|_F^2\end{cases} \tag{3}$$

式中，B_t 和 S_t 表示第 t 次迭代得到的最优解；当分解误差一致收敛到局部最小化时，GoDec 算法停止迭代并得到低秩矩阵 $\boldsymbol{B}$、稀疏矩阵 $\boldsymbol{S}$ 和噪声矩阵 $\boldsymbol{N}$。

通过上述理论可知稀疏矩阵 $\boldsymbol{S}$ 中的行向量都表示每个像元中的异常部分，因此，可以采用欧式距离来计算每个像元的异常度，其表达式为

$$d_i=\sqrt{(\boldsymbol{S}_i-\bar{\boldsymbol{S}})(\boldsymbol{S}_i-\bar{\boldsymbol{S}})^{\mathrm{T}}} \tag{4}$$

式中，$\boldsymbol{S}_i$ 表示像元 $\boldsymbol{x}_i$ 的异常分量，$\bar{\boldsymbol{S}}$ 表示稀疏矩阵的平均行向量。计算结果越大，属于异常像元的概率越高，通过设置阈值能够得到异常检测结果。

2　联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测

2.1　残差矩阵

高光谱图像中存在的异常像元与背景像元不仅在光谱信息上存在差异，在空间信息中也存在较大的差别。从空间信息上进行分析，地物具有聚类的特性，离待测像元距离越远，差异越大。因此，通过以待测像元为中心构建一定大小的空间窗，并计算中心像元与邻域像元的光谱相似度和空间相似度对图像中的所有像元进行光谱值的重构，最后计算原始图像和重构图像矩阵之间的差值得到代表异常目标光谱信息的残差矩阵。

2.1.1　光谱相似度权值

高光谱图像包含成千上万个像元，不同像元之间的光谱信息存在不同程度的区别，光谱角匹配是一种能够衡量两个光谱向量之间差异程度的经典方法。光谱向量之间的光谱角越小，属于同一种物质的概率也就越高，光谱相似度也就越高；反之，光谱相似度也就越低。

基于异常目标的低概率性、尺寸小以及地物的聚类特性，以待测像元为中心构建一定大小的空间窗，并通过逐个计算邻域像元与中心像元之间的光谱角确定像元之间的光谱相似度权值，光谱角的表达式如下：

$$SAM(\boldsymbol{x}_{i,j},\boldsymbol{x}_{r,s})=\cos^{-1}\frac{\boldsymbol{x}_{i,j}\boldsymbol{x}_{r,s}^{\mathrm{T}}}{\left[(\boldsymbol{x}_{i,j}\boldsymbol{x}_{i,j}^{\mathrm{T}})(\boldsymbol{x}_{r,s}\boldsymbol{x}_{r,s}^{\mathrm{T}})\right]^{1/2}}=\cos^{-1}\left(\frac{\langle\boldsymbol{x}_{i,j},\boldsymbol{x}_{r,s}\rangle}{\|\boldsymbol{x}_{i,j}\|\cdot\|\boldsymbol{x}_{r,s}\|}\right) \tag{5}$$

其中，(i,j) 为中心像元的空间坐标，(r,s) 为领域像元的空间坐标，取值范围随着空间窗尺寸的变化而变化。为保持光谱相似度权值与空间相似度权值变化趋势的一致性，本文采用了光谱角的余弦值使其取值范围为[0,1]。同时，为进一步增大光谱差异对光谱相似度权值的作用，利用余弦值的二次方根作为光谱相似度权值 γ，表达式为

$$\gamma(r,s)=\operatorname{sqrt}(\cos(SAM)) \tag{6}$$

2.1.2　空间相似度权值

地物在空间分布上具有聚类的特性，一般来说，相同物体聚集在一起的概率会更高。像元的空间信息对重构光谱值的影响很大，邻域像元离中心像元

越近，对中心像元的作用越大，两者之间的空间相似度权值越大，即与距离成反比。通过计算像元之间的欧式距离能够得到两者之间的空间相似度权值 β，表达式为

$$\beta(r,s)=\frac{1}{\sqrt{((r-i)^2+(s-j)^2)}} \tag{7}$$

假设以待测像元为中心设置大小为 5×5 像素的空间窗，根据空间相似度权值的计算原理得到领域像元的空间相似度权值如表 1 所示。

表 1 空间相似度权值

$\frac{1}{2\sqrt{2}}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2\sqrt{2}}$
$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{5}}$
$\frac{1}{2}$	1		1	$\frac{1}{2}$
$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{5}}$
$\frac{1}{2\sqrt{2}}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2\sqrt{2}}$

2.1.3 比例权值

根据上述所得到的光谱相似度权值和空间相似度权值，定义领域像元比例权值的表达式为

$$\alpha(r,s)=\beta(r,s)\gamma(r,s) \tag{8}$$

根据上述所示的比例权值能够得到所有像元的重构光谱值，得到新的高光谱图像数据 $\boldsymbol{X}'$。中心像元重构光谱值的表达式为

$$\boldsymbol{X}'(i,j)=\sum_{r=i-d}^{i+d}\sum_{s=j-d}^{j+d}\beta(r,s)\gamma(r,s)X(r,s) \tag{9}$$

通过计算高光谱数据 $\boldsymbol{X}$ 与 $\boldsymbol{X}'$之间的差值，可得到残差矩阵 $\boldsymbol{C}$，矩阵 $\boldsymbol{C}$ 中代表高光谱图像 $\boldsymbol{X}$ 中像元的异常分量，表达式为

$$\boldsymbol{C}=\boldsymbol{X}-\boldsymbol{X}' \tag{10}$$

2.2 本文算法步骤

本文算法是一种基于矩阵分解的高光谱异常目标检测算法，通过对 LRaSMD 算法能够获得的异常矩阵进行改进，通过融合图像的空间信息和光谱信息得到原始图像与重构图像的残差矩阵，即包含大量异常目标信息的矩阵，并与 LRaSMD 算法获得的异常矩阵进行相加，最后计算矩阵中行向量之间的欧式距离得到每个像元的异常度，设置阈值，得到最终的异常检测结果，具体步骤如下：

1）利用 LRaSMD 算法对原始的高光谱图像 $\boldsymbol{X}$ 进行分解，从而能够得到图像 $\boldsymbol{X}$ 的异常矩阵 $\boldsymbol{S}$；

2）利用本文所提出的空间联合预处理方法对高光谱图像 $\boldsymbol{X}$ 进行处理并得到重构图像 $\boldsymbol{X}'$，计算并得到二者之间的残差矩阵 $\boldsymbol{C}$；

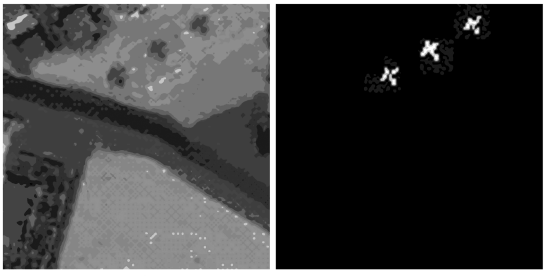
3）通过将异常矩阵 $\boldsymbol{S}$ 和残差矩阵 $\boldsymbol{C}$ 进行相加，得到包含异常目标信息的矩阵，并利用欧式矩阵计算矩阵中每个像元的异常度；

4）设置阈值，得到高光谱图像异常目标检测算法的检测二值图像。

3 实验结果与分析

为充分验证 JSI-LRaSMD 算法的合理性和有效性，采用了真实的高光谱数据进行实验仿真。仿真环境为：CPU 处理器为 Intel Core i7-3635QM，主频为 2.40 GHz，安装内存为 8 GB，软件为 Matlab2018a。

实验数据选取的是 AVIRIS 传感器拍摄的美国 San Diege 海军机场的部分数据，波长范围为 0.37～2.51 μm，去除低信噪比和水蒸气吸收比较严重的波段，保留波段为 189 个；空间分辨率为 3.5 m，截取的区域大小为 120×120 像素，存在 3 个异常目标；图 1 为实验数据的伪彩色图像和真实目标分布图。



(a) 伪彩色图像 (b) 真实目标分布图

图 1 伪彩色图像及真实目标分布

基于上述分析，采用了 RX 算法、LRX 算法、LRaSMD 算法以及 JSI-LRaSMD 算法进行对比仿真实验。之所以选择这几种算法与 JSI-LRaSMD 算法进行对比，是由于这些算法具有一定的代表性且与 JSI-LRaSMD 算法具有一定的相关性。图 2 为四种算法的异常目标检测结果的二值图像，可以看出 LRaSMD 算法的检测效果相比其他三种算法有了明显地提高。

为了能够更加精确地验证 LRaSMD 算法的有效性，采用接收机工作特性（Receiver Operating Characteristic，ROC）曲线进行定量地分析对比。

ROC 曲线能够有效地描述检测概率和虚警概率之间的变化关系，图 3 为四种算法的 ROC 曲线图，纵坐标为检测概率，表示检测到的目标像元数与真实分布的目标总像元数目的比值；横坐标为虚警概率，表达的是检测到的虚警像元数和整幅图像总像元数的比值。且以上四种算法的 AUC 值分别为 0.800 4，0.881 9，0.941 5 和 0.954 6。

从图 3 中可知，LRaSMD 算法的 ROC 曲线优于其他三种算法，与上述分析结果一致。综合分析，无论是从定性的角度，还是定量的角度，本文所提出的 JSI-LRaSMD 算法更具先进性和有效性。

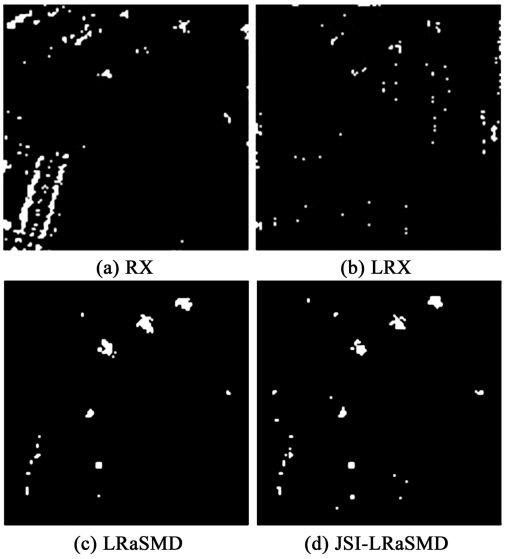


图 2　四种算法检测结果对比

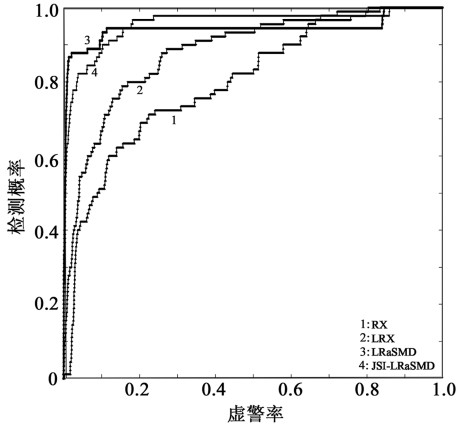


图 3　四种算法的 ROC 曲线对比

4　结论

针对低秩稀疏矩阵分解的高光谱异常目标检测算法忽略了图像的空间信息导致检测精度低的问题，本文提出了一种联合空间信息的改进低秩稀疏矩阵分解的高光谱异常目标检测算法。算法通过构建一定大小的空间窗模型，对待检测像元的光谱值进行重构，从而得到新的高光谱数据，将原始的高光谱数据与新的高光谱数据作差，从而得到代表异常目标信息的残差矩阵，将低秩稀疏矩阵分解中得到的稀疏矩阵与残差矩阵进行加和，能够有效地突出目标的信息，最后，通过计算其行向量之间的欧式距离得到像元的异常度。相比于其他几种典型算法，所提算法具有更高的检测精度。

参考文献：

[1] Reed I S, Yu Xiaoli. Adaptive multi-band CFAR detection of an optical pattern with unknown spectral distribution[J]. IEEE Trans. on Acoustics Speech and Signal Process, 1990, 38(10): 1760-1766.

[2] Taitano Y P, Geier B A, Bauer K W. A locally adaptable iterative RX detector[J]. Eurasip J. on Advances in Signal Processing, 2010(1): 341908.

[3] Stein D W J, Beaven S G, Hoff L E, et al. Anomaly detection from hyperspectral imagery[J]. IEEE Signal Proc. Magazine, 2002, 19(1): 58-69.

[4] Chang C I. Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis[J]. IEEE Trans. Geosci. Remote Sens., 2005, 43(3): 502-518.

[5] Kwon H, Nasrabadi N M. Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery[J]. IEEE Trans. on Geoscience and Remote Sensing, 2005, 43(2): 388-397.

[6] Yuan Z, Sun H, Ji K, et al. Local sparsity divergence for hyperspectral anomaly detection[J]. IEEE Geoscience & Remote Sensing Lett., 2014, 11(10): 1697-1701.

[7] Zhao R, Du B, Zhang L. Hyperspectral anomaly detection via a sparsity score estimation framework[J]. IEEE Trans. on Geoscience & Remote Sensing, 2016, 99: 1-15.

[8] 唐意东，黄树彩，凌　强，等. 字典学习稀疏表示的高光谱图像异常检测[J]. 强激光与粒子束，2015，27(11)：32-38.
Tang Yidong, Huang Shucai, Ling Qiang, et al. Dictionary learning based sparse representation for hyperspectral anomaly detection[J]. High Power Laser and Particle Beams, 2015, 27(11): 32-38.

[9] Zhang Y, Du B, Zhang L, et al. A low-rank and sparse matrix decomposition-based mahalanobis distance method for hyperspectral anomaly detection[J]. IEEE Trans. on Geoscience & Remote Sensing, 2016, 54(3): 1376-1389.

作者简介：

张　炎(1995－)，男，硕士生，主要从事遥感图像处理等方面的研究；

华文深(1967－)，男，博士，副教授，硕士生导师，主要从事光电检测与光信息处理方面的研究。

E-mail：huawensh@126.com

光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.01.030

一种基于 AP 与消除基本陷阱集的低错误平层 QC-LDPC 码构造方法

袁建国，王宏森，张希瑞，赖春红

（重庆邮电大学 光通信与网络重点实验室，重庆 400065）

摘　要： 针对准循环低密度奇偶校验(LDPC)码在高信噪比区域可能存在错误平层的问题，提出了一种基于等差数列(AP)和消除基本陷阱集(ETS)的低错误平层 QC-LDPC 码构造方法。该方法利用改进的 ETS 消除算法构造基矩阵，以减少基本矩阵中的小基本陷阱集。然后利用特殊性质的等差数列(AP)确定循环移位系数，扩展得到最终的校验矩阵。该构造方法的计算复杂度低且码字的码长、码率可灵活设计。并且仿真结果表明，所构造码率为 0.5 的 PEG-Trap set-AP (PTAP)-QC-LDPC(1200，600)码，在误码率为 10^{-6} 时，与 IEEE 802.16 标准中 QC-LDPC(1200，600)码、利用 PEG 算法与 AP 的 PEG-AP-QC-LDPC(1200，600)码、通过控制环(CC)的 CC-QC-LDPC(1200，600)码和基于等差数列的 AP-QC-LDPC(1200，600)码相比较，其净编码增益分别提升了 0.08，0.31，0.57 和 0.64 dB，有效地改善了高信噪比区域的纠错性能，且未出现明显的错误平层。

关键词： 准循环低密度奇偶校验(QC-LDPC)码；陷阱集；等差数列；错误平层

中图分类号： TN929.11　**文章编号：** 1001－5868(2020)01－0146－05

A Construction Method of Low Error Floor QC-LDPC Codes Based on The AP and Eliminating Elementary Trap Sets

YUAN Jianguo，WANG Hongsen，ZHANG Xirui，LAI Chunhong

（Chongqing Key Lab. of Photoelectronic Information Sensing and Transmitting Technol.，Chongqing University of Posts and Telecommunications，Chongqing 400065，CHN）

Abstract： Aiming at the problem that there may be an error floor of quasi-cyclic low-density parity-check (QC-LDPC) codes in the higher signal-to-noise ratio (SNR) region，a novel construction method of QC-LDPC codes based on the arithmetic sequence (AP) and eliminating elementary trap sets (ETS) is proposed. The method uses the improved ETS elimination algorithm to construct the base matrix and reduce the small ETS，and then applies the AP with the specific character to confirm the cyclic shift coefficient，thus the base matrix is extended to obtain the final check matrix，furthermore，the proposed construction method has the low computing complexity and can flexibly design the code-length and the code-rate. The simulation results show that at the bit error rate of 10^{-6}，the net coding gain of the progressive edge growth (PEG)-trap set-AP(PTAP)-QC-LDPC(1200，600) code with a code-rate of 0.5，compared with those of the QC-LDPC (1200，600) code in IEEE 802.16 standard，the PEG-AP-QC-LDPC(1200，600) code based on PEG algorithm and AP，the CC-QC-LDPC (1200，600) code based on controlling cycles (CC) and the AP-QC-LDPC (1200，600) code based on AP，is improved up to 0.08，0.31，0.57 and 0.64 dB respectively，the

收稿日期：2019－12－19.

基金项目：国家自然科学基金项目(61971079，61671091)；重庆市基础与前沿研究计划项目(csts2017jcyjAX0427).

constructed PTAP-QC-LDPC (1200, 600) code can effectively improve the error correction performance in the higher SNR region without the obvious error floor.

Key words: QC-LDPC codes; trap sets; arithmetic progression; error floor

0 引言

在通信技术高速发展的今天，通信可靠性越来越受到人们的重视，为了克服通信系统中的干扰，往往采用信道编码技术来提升通信系统的可靠性。低密度奇偶校验(Low-Density Parity-Check，LDPC)码是 Gallager 在 1962 年提出的一种能发现并纠正误码的信道编码技术，且具有纠错性能逼近香农限的优点[1-2]。准循环低密度奇偶校验(Quasi-Cyclic Low-Density Parity-Check，QC-LDPC)码是一种特殊的 LDPC 码，它的校验矩阵具有准循环的特性，仅需存储循环移位系数就能确定相应的码字，更容易硬件实现其编译码模块[3]，所以成为了学者们研究的焦点。

在 LDPC 码的 Tanner 图中，可能存在一些会导致译码失败的特殊拓扑结构，从而导致错误平层现象的出现。LDPC 码的错误平层问题是指在一定的高信噪比区域，误码率性能曲线的斜率突然减小甚至接近于 0 的现象。而陷阱集是导致错误平层这一现象的主要原因之一[4]。

近年来，许多国内外学者针对错误平层现象做了大量的研究，其中主要分为码型构造和迭代译码两方面。在迭代译码方面，Yang 等人提出了一种将置信传播(Belief Propagation，BP)译码和对数似然比(Log Likelihood Ratios，LLR)译码联合的方法[5]，在译码时进行两阶段译码来达到降低错误平层的目的。在码型构造中，如果采用结构化构造方法，往往是从增大校验矩阵的围长入手；如果采用随机构造方法，如渐进边增长(Progressive Edge Growth，PEG)算法，则可以从增大环的连通性和避免陷阱集的存在构造低错误平层的 LDPC 码。Nguyen 等人通过分析在置信传播(Belief Propagation，BP)译码算法下各陷阱集的译码失败情况，来确定不同陷阱集的危害程度，从而构造出不含某些小陷阱集的结构化 LDPC 码[6]。2018 年，Tao 等人提出了一种构造围长为 8、列重为 3 的低错误平层 QC-LDPC 码构造方法，且不包含部分小陷阱集[7]。Hashemi 等[8]推导了可变规则 LDPC 码中给定 b 的最小非基本陷阱集大小的下界，结果表明最小非基本陷阱集的大小通常大于具有相同 b 值的基本陷阱集大小，证明了基本陷阱集对 LDPC 码的影响更大。Amirzade 等[9]学者给出了任意围长的可变规则 LDPC 码和围长为 8 的非规则 LDPC 码中的基本陷阱集大小的下界，节约了陷阱集搜索的时间。

针对 LDPC 码中的错误平层问题，本文从码型构造的角度，提出了一种基于等差数列和消除基本陷阱集相结合的低错误平层 QC-LDPC 码构造方法。在该方法中，首先利用具有特殊性质的等差数列构造循环移位系数矩阵，再通过改进的基本陷阱集搜索算法构造出基本矩阵，最后循环扩展得到最终的校验矩阵。仿真结果表明，利用本文提出的方法构造的 QC-LDPC 码型，具有较好的纠错性能且并未出现明显的错误平层。

1 基于等差数列构造基矩阵

QC-LDPC 码对应的校验矩阵 $\boldsymbol{H}$，是由基矩阵按一定规则扩展得到的。利用结构化构造方法构造 QC-LDPC 码，最重要的一环就是基矩阵的设计。本文中构造基矩阵使用的等差数列(Arithmetic Progression，AP)是一种满足一定条件的特殊数列，它是差值可变化的一种特殊的 AP。由该方法构造出的基矩阵围长至少为 8。

QC-LDPC 码的校验矩阵 $\boldsymbol{H}$ 和基矩阵 $\boldsymbol{H}_{\mathrm{b}}$ 可以表示为式(1)和(2)：

$$\boldsymbol{H}=\begin{bmatrix}\boldsymbol{I}_{p(1,1)} & \boldsymbol{I}_{p(1,2)} & \cdots & \boldsymbol{I}_{p(1,n)}\\ \boldsymbol{I}_{p(2,1)} & \boldsymbol{I}_{p(2,2)} & \cdots & \boldsymbol{I}_{p(2,n)}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{I}_{p(m,1)} & \boldsymbol{I}_{p(m,2)} & \cdots & \boldsymbol{I}_{p(m,n)}\end{bmatrix}_{mq\times nq} \tag{1}$$

$$\boldsymbol{H}_{\mathrm{b}}=\begin{bmatrix}p(1,1) & p(1,2) & \cdots & p(1,n)\\ p(2,1) & p(2,2) & \cdots & p(2,n)\\ \vdots & \vdots & \ddots & \vdots\\ p(m,1) & p(m,2) & \cdots & p(m,n)\end{bmatrix}_{m\times n} \tag{2}$$

其中，$\boldsymbol{I}_{p(i,j)}(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$表示一个循环置换矩阵(Circulant Permutation Matrix，CPM)或零矩阵，如式(3)所示，m 和 n 分别表示基矩阵 $\boldsymbol{H}_{\mathrm{b}}$ 的行数和列数，q 表示扩展因子，$p(i,j)$为循环移位系

数。

$$\boldsymbol{I}_{p(i,j)}=\begin{cases}p\times p\text{ 零矩阵,当 }p(i,j)=0\text{ 时}\\ p\times p\text{ 循环置换矩阵,其他}\end{cases}\tag{3}$$

文献[10]定义了一种特殊的等差数列,且利用该数列构造出的基矩阵的围长至少为 8,如定理 1[10]所示。

定理 1[10]:如果基矩阵 $\boldsymbol{H}_{\mathrm{b}}$ 中的各个元素的差值 $d_{i,j}$ 满足式(4),则对任意 $m\times n$ 的 $\boldsymbol{H}_{\mathrm{b}}$,其对应的 Tanner 图的围长至少为 8。

$$\begin{cases}d_{1,j}=d\geqslant 0\\ d_{2k+2,j}=d_{2k+1,j}+1 & k\geqslant 0\\ d_{4k+3,j}=\max\{n-j+d_{4k+2,j},j+d_{4k+2,j}\} & k\geqslant 0\\ d_{4k+1,j}=(n-1)(d_{4k,j}-d)+d+1 & k\geqslant 1\end{cases}\tag{4}$$

式中,$d_{i,j}=p(i,j+1)-p(i,j)$,$1\leqslant i\leqslant m$,$1\leqslant j\leqslant n-1$。$d_{i,j}$ 表示基矩阵中第 i 行第 $j+1$ 列与前一个元素的差值,d 是一个可设置的常量,为第一行元素的公差。由定理 1 可以构造出的 QC-LDPC 码的基矩阵围长至少为 8。

例如,取 $m=4$,$n=8$,$d=1$,$p(i,1)=0$(其中 $1\leqslant i\leqslant m$),由式(4)可以得到如式(5)的矩阵:

$$\boldsymbol{H}_{\mathrm{eg}}=\begin{bmatrix}0&1&2&3&4&5&6&7\\0&2&4&6&8&10&12&14\\0&9&17&24&30&37&45&54\\0&10&19&27&34&42&51&61\end{bmatrix}_{4\times 8}\tag{5}$$

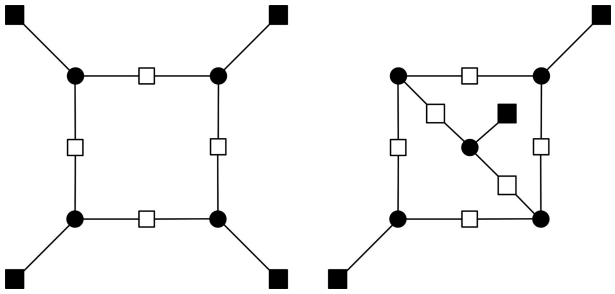
2 基本陷阱集搜索算法

2.1 LDPC 码中的陷阱集

定义 1:一个(a,b)陷阱集[8] T 是变量节点集合 V 的子集,T 的大小即变量节点的个数为 a,与 T 相连的度为奇数的校验节点个数为 b。特殊地,如果与 T 相连的所有校验节点的度都为 1 或者 2,则 T 是一个(a,b)基本陷阱集。图 1(a)是一个(4,4)基本陷阱集,其中矩形表示校验节点,圆形表示变量节点,黑色表示该节点度为奇数,白色表示该节点度为偶数。

由于直接对在 Tanner 图中查找陷阱集比较复杂,Karimi 在文献[11]中提出了一种基于环的基本陷阱集搜索算法,通过将陷阱集进行扩展来得到一定范围内的基本陷阱集。

在 LDPC 码的 Tanner 图中,任意的环所包含的变量节点集合都会构成陷阱集。例如图 1(a)所示的(4,4)基本陷阱集中就包含了一个 8 环。各个陷阱集在 Tanner 图中也并非独立存在,一个陷阱集可以通过扩展得到一个较大的陷阱集,如图 1(b)的(5,3)基本陷阱集就是通过(4,4)基本陷阱集添加一个变量节点扩展而来。同样地,(5,3)基本陷阱集也可扩展为(6,2)和(7,3)等基本陷阱集。



(a) (4,4)基本陷阱集 (b) (5,3)基本陷阱集

图 1 (4,4)和(5,3)基本陷阱集

2.2 基本陷阱集搜索算法

根据环和陷阱集的关系以及基本陷阱集之间的扩展规则,可以采用算法对 LDPC 码中的基本陷阱集进行搜索。

算法 1:基本陷阱集搜索算法

```
for  i=1:m
    初始化最大环长为 2l,设列表 A 为 l+3 行 4 列的零矩阵;
    将校验节点与当前变量节点 v_i 相连,查询出环长小于或等于 2l 的所有环;
    查询出所有由环构成的基本陷阱集,若存在基本陷阱集,在 A 的对应的位置进行计数,其中 a≥3,0≤b≤4,a≥b;
将每个环构成的(a,b)基本陷阱集作为初始值;
for 初始值 do
    根据特定的拓扑结构对初始值进行扩展;
    将扩展后的(a′,b′)基本陷阱集在 A 对应的位置累加计数,其中 a+1≤a′≤a+3,0≤b′≤4;
    对扩展后 A 中所对应的各基本陷阱集数量求和,并存放在矩阵 B 中的对应行中;
    if  B 的该行元素之和为 0
    直接选取该校验节点,并跳出整个循环;
else
        对下一个校验节点进行测试;
        end
    end
end
for  j=1:m
    查询 B 中每行,找出基本陷阱集总数最小的行,则该行对应的校验节点即为当前最优节点,并选取该校验节点;
end
```

按照以上步骤可对 LDPC 码中的基本陷阱集进行扩展搜索,并将搜索结果保存到列表中进行比较,以选取当前最优的校验节点。

3 基于消除基本陷阱集的 QC-LDPC 码设计

为降低 LDPC 码的错误平层，本文将 2.2 节中的基本陷阱集搜索算法，结合 PEG 算法的思想，并在选取校验节点时，增加对环的连通性的考虑，以此来构造更优的基矩阵。本文采用额外信息度(Extrinsic Message Degree，EMD)[12-13]表征环的连通性。具体的算法流程如下所示。

算法 2：EMD 值检测与基本陷阱集搜索结合算法

for $j=1:n$
  for $k=1:d_{v(j)}$
    if $k=1$
    $E^1_{s_j} \leftarrow (c_l, s_j)$，$E^1_{s_j}$ 是连接到变量节点 s_j 的第一条边，c_l 满足在所有校验节点集合构成的子图中度数最小；
  else 把树图从点 s_j 扩展到 l 层；
    if 满足 $N^l_{s_j}$ 和 $N^{l+1}_{s_j}$ 的元素个数相等（$N^l_{s_j}$ 表示以 s_j 为根，第 l 层校验节点的集合）
      环的 EMD 值检测；找出所有满足一定 EMD 值的校验节点；
      使用算法 1 进行基本陷阱集搜索，选取最优校验节点；
    else if 满足 $N^l_{s_j} \neq \varnothing$ 和 $N^{l+1}_{s_j} \neq \varnothing$
      环的 EMD 值检测；找出所有满足一定 EMD 值的校验节点；
      使用算法 1 进行基本陷阱集搜索，选取最优校验节点；
    end
  end
end
end

此算法以 PEG 算法的思想，逐列构造矩阵。由于越小的基本陷阱集对码字性能影响越大，所有在构造过程中，通过陷阱集之间的扩展关系，选取构成小基本陷阱集最少的校验节点与当前变量节点相连，尽可能规避小陷阱集的出现。同时增加对环的连通性的考虑，当不满足 EMD 阈值时直接搜索下一个节点，免除了多余的基本陷阱集搜索过程，不仅能够使构造出的码字具有优秀的纠错性能，也能降低一定的算法复杂度。

在本文中，首先使用算法 2 构造一个 $m \times n$ 的基矩阵，不仅在基矩阵中的构造中提升了各个节点的连通性，而且在此过程中尽可能规避小的基本陷阱集的存在，然后根据式(4)，根据特殊等差数列的定义来获得相同维度的循环移位系数矩阵，最终根据式(3)扩展得到最后的校验矩阵 $\boldsymbol{H}$。

4 仿真结果和分析

本文提出了一种基于等差数列和消除基本陷阱集相结合的 QC-LDPC 码构造方法。为了验证该方法纠错性能的优越性，构造了码长为 1 200，码率为 0.5 的 QC-LDPC 码，并进行了 MATLAB 仿真分析。本文仿真条件是：二进制相移键控(Binary Phase Shift Keying，BPSK)调制，加性高斯白噪声(Additive White Gaussian Noise，AWGN)信道，BP 译码，综合考虑纠错性能与译码复杂度，译码迭代次数选择 50 次。

本文根据密度进化算法，设定校验矩阵的度分布为 $\lambda(x)=0.383\,54x+0.042\,37x^2+0.574\,09x^3$，设定参数 $m=24$，$n=48$，选取扩展因子 $q=25$，如上文所述构造得到 PEG-Trap set-AP(PTAP)-QC-LDPC(1200，600)码型，与文献[10]直接利用等差数列构造的 AP-QC-LDPC(1200，600)码型、利用 PEG 算法与等差数列构造的 PEG-AP-QC-LDPC(1200，600)码型、通过控制环(Controlling Cycles，CC)的 CC-QC-LDPC(1200，600)码型[7]以及 IEEE802.16 标准[14]中的 IEEE 802.16 QC-LDPC(1200，600)码型进行仿真对比分析，得到的仿真结果图和在信噪比为 3.2 dB 时的误码率对比结果分别如图 2 和表 1 所示。

表 1 当信噪比为 3.2 dB 时 PTAP-QC-LDPC(1200，600)码与其他码的误码率仿真对比

码型	误码率
AP-QC-LDPC(1200，600)	2.17×10^{-6}
CC-QC-LDPC(1200，600)	1.36×10^{-6}
PEG-AP-QC-LDPC(1200，600)	7.37×10^{-7}
IEEE 802.16 QC-LDPC(1200，600)	1.77×10^{-7}
PTAP-QC-LDPC(1200，600)	5.21×10^{-8}

本文构造的码率为 0.5 的 PTAP-QC-LDPC(1200，600)码与其他三种码型具有相同的码长和码率，由图 2 和表 1 分析可知，PTAP-QC-LDPC(1200，600)码在误码率为 10^{-6} 时，相较于 IEEE 802.16 QC-LDPC(1200，600)码，净编码增益约提升 0.08 dB，相较于 PEG-AP-QC-LDPC(1200，600)码，净编码增益约提升 0.31 dB，相较于 CC-QC-LDPC(1200，600)码，净编码增益约提升 0.57 dB，相较于 AP-QC-LDPC(1200，600)码，净编码增益约提升 0.64 dB。

通过上述分析，可以看出通过本文提出的方法构造的 PTAP-QC-LDPC(1200，600)码型相较于其他几种码型，在高信噪比情况下，误码率均有明显的降低，具有更好的纠错性能，并且降低甚至消除了错误平层的现象。

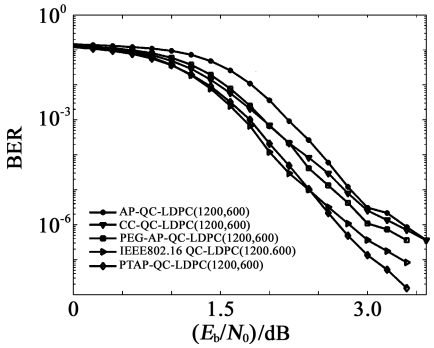


图 2 PTAP-QC-LDPC(1200,600)码与其他码的纠错性能

5 结束语

为了消除或降低高信噪比区域 LDPC 码存在的错误平层现象,本文从码型构造的角度出发,提出了一种基于等差数列和消除基本陷阱集相结合的低错误平层 QC-LDPC 码构造方法。首先通过改进的消除基本陷阱集算法构造基矩阵,并利用特殊性质的等差数列确定移位系数,最后循环扩展得到校验矩阵。在本方法中,移位系数的确定可根据简单的数学公式计算,无需进行计算机搜索便可快速得到,并且在基矩阵构造中添加 EMD 值判断条件,减少了一定算法复杂度。另外,还可通过改变基矩阵的大小和扩展因子来灵活改变 QC-LDPC 码的码长和码率,以适用于不同的通信系统中。本文提出的构造方法,一方面具有可接受的运算复杂度,另一方面也具有优秀的纠错性能。仿真结果表明,通过本方法构造的 QC-LDPC 码型,在高信噪比区域,并没有出现错误平层现象,且误码率性能明显优于其他几种码型。

参考文献:

[1] 袁建国,孙乐乐,范福卓,等. 基于等差数列与原模图的 QC-LDPC 码构造方法[J]. 重庆邮电大学学报(自然科学版),2019,31(5):649-654.
Yuan Jianguo, Sun Lele, Fan Fuzhuo, et al. Construction method of QC-LDPC codes based on the arithmetic progression and protograph[J]. J. of Chongqing University of Posts and Telecommun. (Nature Science Edi.), 2019, 31(5): 649-654.

[2] He X, Zhou L, Du J. PEG-like design of binary QC-LDPC codes based on detecting and avoiding generating small cycles [J]. IEEE Trans. on Commun., 2018, 66(5): 1845-1858.

[3] 袁建国,曾 晶,袁 梦,等. 基于 Fibonacci-Lucas 序列构造大围长 QC-LDPC 码的方法[J]. 重庆邮电大学学报(自然科学版),2018,30(4):505-510.
Yuan Jianguo, Zeng Jing, Yuan Meng, et al. Construction method of large girth QC-LDPC codes based on Fibonacci-Lucas sequence [J]. J. of Chongqing University of Posts and Telecommun. (Nature Science Edi.), 2018, 30(4): 505-510.

[4] Hashemi Y, Banihashemi A H. Characterization of elementary trapping sets in irregular LDPC codes and the corresponding efficient exhaustive search algorithms[J]. IEEE Trans. on Commun., 2018, 64(5): 3411-3430.

[5] Yang P, Jun B, No J, et al. A new two-stage decoding scheme with unreliable path search to lower the error-floor for low-density parity-check codes[J]. IET Commun., 2017, 11 (14): 2173-2180.

[6] Nguyen D V, Chilappagari S K, Marcellin M W, et al. On the construction of structured LDPC codes free of small trapping sets[J]. IEEE Trans. on Information Theory, 2012, 58(4): 2280-2302.

[7] Tao X, Li Y, Liu Y, et al. On the construction of LDPC codes free of small trapping sets by controlling cycles[J]. IEEE Commun. Lett., 2018, 22(1): 9-12.

[8] Hashemi Y, Banihashemi A H. Lower bounds on the size of smallest elementary and non-elementary trapping sets in variable-regular LDPC codes [J]. IEEE Commun. Lett., 2017, 21(9): 1905-1908.

[9] Amirzade F, Sadeghi M. Analytical lower bounds on the size of elementary trapping sets of variable-regular LDPC codes with any girth and irregular ones with girth 8[J]. IEEE Trans. on Commun., 2018, 66(6): 2313-2321.

[10] Zhang Y, Da X. Construction of girth-eight QC-LDPC codes from arithmetic progression sequence with large column weight[J]. Electron. Lett., 2015, 51(16): 1257-1259.

[11] Karimi M, Banihashemi A H. Efficient algorithm for finding dominant trapping sets of LDPC codes[J]. IEEE Trans. on Information Theory, 2012, 58(11): 6942-6958.

[12] Deka K, Rajesh A, Bora P K. On the equivalence of the ACE and the EMD of a cycle for the ACE spectrum constrained LDPC codes[C]// IEEE Inter. Symp. on Turbo Codes and Iterative Information Processing (ISTC), 2014: 67-71.

[13] 袁建国,汪 哲,高文春,等. 一种低错误平层 LDPC 码构造方法[J]. 重庆邮电大学学报(自然科学版),2017,29(1):15-18.
Yuan Jianguo, Wang Zhe, Gao Wenchun, et al. A construction method of LDPC codes with the low error floor[J]. J. of Chongqing University of Posts and Telecommun. (Nature Science Edi.), 2017, 29(1): 15-18.

[14] IEEE. IEEE Standard for Air Interface for Broadband Wireless Access Systems[S]. IEEE Press, 2012.

作者简介:

袁建国(1968—),男,重庆人,博士,教授,硕士生导师,主要研究方向为光通信系统中 FEC 编译码技术与 OFDM 技术。

E-mail: yyyyjg@126.com

《半导体光电》编辑委员会

顾　问　王占国(科学院院士)　姚建铨(科学院院士)　方家熊(工程院院士)

主　任　欧　黎

副主任　蒋志伟

委　员　(以姓氏笔画为序)

王小勇　方　亮　孔　梅　宁永强　孙长征　刘秉琦　刘　胜　刘　斌
华　勇　安俊明　李　成　李向阳　李国强　余　华　邹卫文　张卫平
张佰君　沈鸿烈　宋凝芳　罗风光　周立庆　孟　洲　祝连庆　赵明富
徐学良　高新江　郭　霞　崔一平　龚　敏　蒋亚东　喻志农　童杏林
熊　平　潘　武　瞿鹏飞

半导体光电
BANDAOTI GUANGDIAN
(双月刊　1976年创刊)
第41卷　第1期　2020年2月

SEMICONDUCTOR OPTOELECTRONICS
(Bimonthly, Started in 1976)
Vol. 41　No. 1　February　2020

主管单位	中国电子科技集团有限公司	**Competent Authorities**	China Electronics Technology Group Corporation
主办单位	重庆光电技术研究所	**Sponsored by**	Chongqing Optoelectronics Research Institute
编辑出版	《半导体光电》编辑部	**Edited and Published by**	Editorial Office of Semiconductor Optoelectronics
编委主任	欧　黎	**Director of Editorial Board**	OU Li
主　　编	蒋志伟	**Chief Editor**	JIANG Zhiwei
执行主编	李福果	**Executive Chief Editor**	LI Fuguo
责任编辑	朱玲瑞	**Editor in Charge**	ZHU Lingrui
印　　刷	重庆五湖印务有限公司	**Printed by**	Chongqing Wuhu Printing Corp.
国内订阅	《半导体光电》编辑部	**Domestic Subscription**	Editorial Office of Semiconductor Optoelectronics
通信地址	重庆市南岸区南坪花园路14号(400060)	**Address**	No. 14 Garden Road, Nan'an District, Chongqing 400060, P. R. China
电　　话	(023)65860286(编辑) (023)65860287(发行)	**Telephone**	(023)65860286(Edition) (023)65860287(Distribution)
电子邮箱	soe@163.net	**E-mail**	soe@163.net

中国标准连续出版物号：ISSN 1001－5868 / CN 50－1092/TN

每期定价　20.00元

半导体光电

二〇二〇年二月
二〇二〇年二月

第四十一卷 第一期 （001-150）
第四十一卷 第一期 （001-150）