光电技术及应用

DOI：10.16818/j.issn1001－5868.2024032502

# 基于子空间投影的生成对抗网络红外图像降噪

闫　宁[1]，周　斌[1]*，王伟明[2]，张宇迪[1]

（1. 郑州科技学院 电子与电气工程学院，郑州 450064；2. 雄安创新研究院，雄安新区 071000）

**摘　要：** 针对红外图像信噪比低、边缘信息模糊、杂波干扰多等检测难点，提出一种基于子空间投影的生成对抗网络红外图像降噪方法。首先，生成器由 U-Net 结构和子空间注意力网络构成，编码阶段由 4 层下采样实现图像特征提取，解码阶段由 4 层上采样重建图像。其次，在每层跳跃连接中加入子空间投影网络，每层特征图与同层上采样后图像共同放入子空间投影网络进行图像特征融合，将投影特征图与原始高级特征融合实现图像降噪。最后，将图像输入到鉴别器进行对抗训练，得到清晰重建图像。实验结果表明，与 BM3D，DnCNN 等常用算法相比，改进的生成对抗网络算法有更好的客观评价指标效果，PSNR 和 SSIM 分别达到了 34.36，0.985 2 dB，从而验证了改进算法的良好降噪性能。

**关键词：** 红外图像降噪；生成对抗网络；子空间投影；图像融合

**中图分类号：** TP391.41　**文章编号：** 1001－5868(2024)05－0847－06

## Infrared Image Denoising Based on A Generative Adversarial Network Improved by Subspace Projection

YAN Ning[1], ZHOU Bin[1], WANG Weiming[2], ZHANG Yudi[1]

(1. Zhengzhou University of Science and Technology, Zhengzhou 450064, CHN;
2. Xiong'an Institute of Innovation, Xiong'an New Area 071000, CHN)

**Abstract:** To address the detection challenges of infrared images, such as a low signal-to-noise ratio, blurred edge information, and clutter interference, a generative adversarial network infrared image denoising method based on subspace projection is proposed. First, the generator consists of a U-Net structure and a subspace attention network. The encoding stage extracts image features through four layers of downsampling, while the decoding stage reconstructs the image through four layers of upsampling. Second, a subspace projection network is added to each skip connection, and the feature maps of each layer are combined with upsampled images from the same layer to form a subspace projection network for image feature fusion. The projected feature maps are then fused with the original high-level features to achieve image denoising. Finally, the image is input to the discriminator for adversarial training to obtain a clear reconstructed image. The comparative experiments with BM3D (Block-Matching and 3D Filtering), DnCNN(Deep Neural Neural Network For Image Denoising), and other algorithms show that the improved generative adversarial network algorithm has better objective evaluation index effects, with PSNR and SSIM reaching 34.36 dB and 0.985 2, respectively, thus verifying the strong denoising performance of this algorithm.

**收稿日期：**2024－03－25.

**基金项目：**河南省科技攻关项目(242102211110)；河南省高等学校重点科研项目(24A510013)；微系统技术重点实验室开放课题项目(6142804231002).

* **通信作者：**周　斌

**Key words**: infrared image denoising; generative adversarial network; subspace projection; image fusion

## 0 引言

红外图像易受设备固有缺陷、工作环境等因素影响,导致图像存在噪声大、边缘模糊、对比度低等缺点,对后续图像融合、目标提取十分不利。传统的红外图像降噪算法有空间域降噪、变换域降噪及混合降噪[1-2]等,这些方法常存在丢失图像细节、引入新的噪声等问题。基于各种深度学习的红外图像降噪方法,通过训练图像数据学习降噪规律,能够在保留图像细节的同时实现高效降噪,如张骏等[3]提出一种基于引导滤波的多分支注意力残差去噪网络,可有效实现红外图像降噪,较大程度保留图像纹理信息。生成对抗网络(Generative Adversarial Networks,GAN)基于对抗模型,由生成器和鉴别器博弈输出降噪图像,其优势在于利用反向传播即可完成数据训练,同时可与其他算法结合进一步提升降噪能力,是目前的研究热点。陈佛计等[4]基于卷积神经网络、转置卷积神经网络构建 GAN 模型的生成器和鉴别器,该模型可将高维空间彩色图像映射为高质量红外图像;吴国瑞等[5]基于 U-Net 网络与全局上下文模块结合构建生成器,以提取图像全局与局部特征,同时基于胶囊网络和 Res2Net 构建鉴别器,对胶囊网络全连接层进行反卷积网络重构,该方法可有效突出细节信息,实现噪声抑制效果。

目前 GAN 网络存在的主要问题是模型的自由度高、鲁棒性欠佳,虑到子空间投影能够捕捉并聚焦图像中的重要特征,保留低秩空间的特征信息,滤除噪声空间的噪声信息[6],本文在对抗网络基础上加入子空间投影模型,以有效分离噪声和信号,同时利用 WGAN(Wasserstein GAN)框架[7-8],以增强网络鲁棒性,获得更接近真实红外场景的复原图像。

## 1 子空间投影降噪理论

子空间投影降噪是将含有噪声图像的向量空间分解为信号子空间和噪声子空间,通过消除噪声空间分向量、保留信号子空间实现信号与噪声分离。传统的子空间降噪由人工构造基,数据样本的图像梯度信息与噪声信息存放于同一基中,当此组基去掉后,图像梯度信息随之损失,进而造成图像边缘信息丢失。卷积神经网络可自动从样本中学习一组基,有效区分梯度信息与噪声信息分布[9],因此考虑利用子空间投影的基底生成函数 $f$ 生成一系列红外图像基底向量,构成子空间,再经过卷积网络、正交线性投影及重建等过程,完成高频噪声子空间和低维信号子空间分离,进而实现红外图像降噪。

### 1.1 子空间投影

基底向量生成函数可表示为:

$$f:(\mathbb{R}^{C\times H\times W},\mathbb{R}^{C\times H\times W})\rightarrow\mathbb{R}^{N\times K} \tag{1}$$

式中,$\mathbb{R}$ 为红外特征图集,$\boldsymbol{C}$,$\boldsymbol{H}$ 和 $\boldsymbol{W}$ 分别为特征图通道数、特征图高及特征图宽,$\boldsymbol{N}$,$\boldsymbol{K}$ 分别为变换后的特征图高及宽。

以红外图像为输入,利用预训练卷积神经网络的残差卷积层和池化层提取图像特征[10],定义 $\boldsymbol{X}_1$,$\boldsymbol{X}_2\in\mathbb{R}^{C\times H\times W}$ 为单个图像的 2 个特征图,则基底向量可表示为:

$$\boldsymbol{V}=f(\boldsymbol{X}_1,\boldsymbol{X}_2) \tag{2}$$

通过检测 $\boldsymbol{X}_1$,$\boldsymbol{X}_2$ 特征图,可计算出 $M$ 维基向量矩阵$[V_1,V_2,\cdots,V_n]$,$M$ 维度基向量构成了信号子空间。

若以 $\boldsymbol{X}_1$ 特征图投影,由基底生成函数 $f$ 可得到 $\boldsymbol{V}\in\mathbb{R}^{H\times W}$,经变形后得到 $\boldsymbol{V}\in\mathbb{R}^{N\times K}$,而 $M$ 维基向量 $\boldsymbol{V}=[V_1,V_2,\cdots,V_n]$ 经正交线性投影后,可将 $\boldsymbol{X}_1$ 特征图投影到低维信号子空间的 $\boldsymbol{V}$ 上[11],定义 $\boldsymbol{P}$ 为信号子空间的正交投影矩阵,有:

$$\boldsymbol{P}=\boldsymbol{V}(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{T}} \tag{3}$$

式中,$\boldsymbol{V}^{\mathrm{T}}$ 为 $\boldsymbol{V}$ 的转置矩阵。

设 $\boldsymbol{X}_1$ 特征图经信号子空间投影重建后的函数为 $\boldsymbol{Y}$,有:

$$\boldsymbol{Y}=\boldsymbol{P}\boldsymbol{X}_1 \tag{4}$$

### 1.2 基于子空间注意力模型的图像降噪

子空间注意力模型是基底生成和重建投影点的关键,可根据每个特征图子空间推断不同的注意图,实现多尺度、多频率特征表示[12],每个特征图在获得单独注意图基础上完成整个基向量的生成,网络模型如图 1 所示。

原始图像经卷积神经网络后得到特征图 $\boldsymbol{X}_1$,$\boldsymbol{X}_2$,二者级联并通过残差卷积块可捕获在高维数据中心的低维数据,从而提取到更加丰富的特征信息,改善网络图像梯度信息消失问题。残差卷积网络由一个 1×1 等比例卷积、两个 3×3 卷积核、两个

Relu 函数构成，残差卷积后可得到一个 $\boldsymbol{H}\times\boldsymbol{W}\times\boldsymbol{K}$ 的特征向量，再经过变形得到信号子空间的 $M$ 维基向量矩阵$[V_1,V_2,\cdots,V_n]$，正交变换后得到投影矩阵 $\boldsymbol{P}$，将输入特征图 $\boldsymbol{X}_1$ 变形到与信号子空间基向量相同尺寸后，利用公式(4)可得到信号子空间的重建降噪特征图。

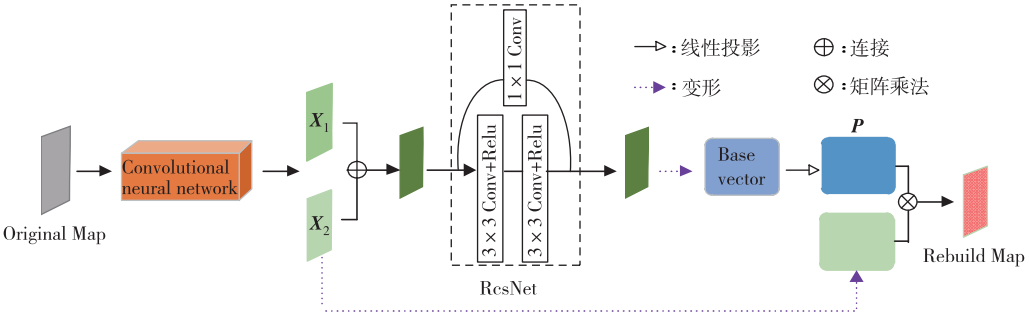



图 1　子空间注意力网络模型

## 2　子空间投影去噪理论与子空间投影 GAN 的红外图像去噪

WGAN 网络在 GAN 网络之上增加了 Wasserstein 距离，泛化能力得到极大改善，本文以 WGAN 为基本框架进行改进。改进网络同样由生成器和鉴别器组成，生成器采用子空间投影网络，在传统 U-Net 结构编码与解码中加入子空间注意力网络，鉴别器由卷积与 Relu 函数构成。带噪声的红外图像输入生成器，由鉴别器输出接近真实清晰的红外图像。

### 2.1　生成器网络

生成器网络采用一个 4 层深度的多路编码 U-Net 结构，由编码和解码构成[13]，加入子空间投影网络的生成器网络如图 2 所示。图中，左半部分是编码，由 4 个下采样层实现特征提取，每个下采样由残差卷积实现。以第 2 特征层为例讨论输入红外图像的随机剪裁：经过第 1 层红外图像下采样后，在第 2 特征层进行两次残差卷积，得到不同尺寸特征图，下采样用残差卷积代替池化层，以弥补池化操作后图像细节丢失问题，残差卷积的卷积核为 4×4，步长 2，填充 1。图中第 2 特征层间为跳跃连接，跳跃连接中加入子空间投影网络。设解码部分第 2 层卷积后的特征图为 $\boldsymbol{X}_1$，对应层编码部分上采样得到特征图为 $\boldsymbol{X}_2$，将它们共同输入子空间投影网络，每次特征融合前实现信号、噪声的分离，得到重建后的特征图再与同层上采样特征图融合，而后输入到下一解码阶段[14]。图中右半部分为解码过程，采用与编码对称的 4 个上采样实现图像重建，上采样采用转置卷积的方法，经过解码器重建后的图像尺寸不变。

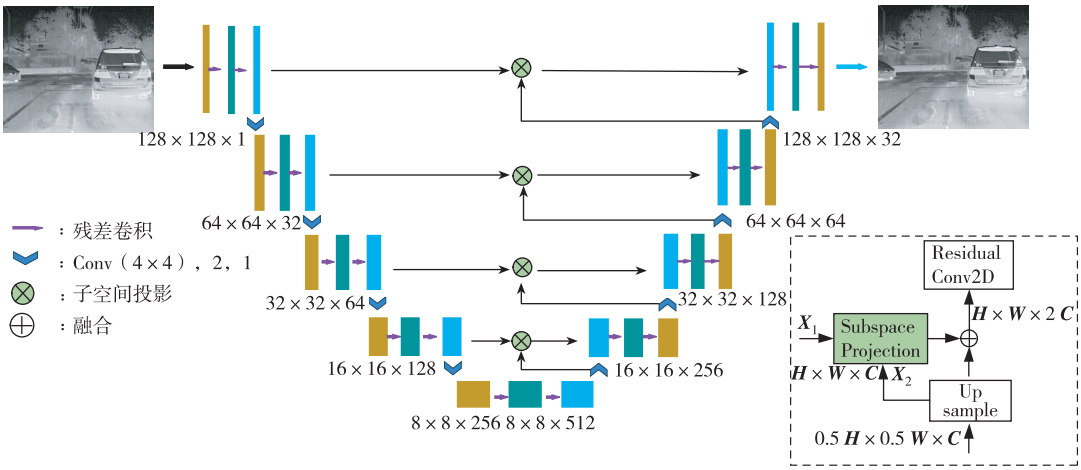



图 2　生成网络结构

### 2.2　鉴别器网络

鉴别器通过区分真实图像与生成图像之间的差异，鉴定生成器中图像的真实性，从而引导生成器向正确方向训练[15]。网络结构包括 7 个卷积层和 2

个全连接层。在7层卷积中，前6层由卷积运算的动函数与步长所得到的Conv(3,1,1)与Relu函数联合构成，其中第1、2层有64个卷积核，第3、4层有256个卷积核，第5、6层有512个卷积核，经过卷积运算后，再经过最后一个Conv(3,1,1)的卷积层和两次全连接层可得到鉴别后的结果。

### 2.3 损失函数

常用损失函数为WGAN-GP(Wasserstein GAN-Gradient Penalty)，即添加梯度惩罚达到稳定训练的目的[8]，但WGAN-GP训练较困难且收敛速度慢，实时处理能力不强，为此本文采用最小二乘GAN损失函数：

$$\min_D V(D)=\frac{1}{2}E_{x-P_{\text{data}}}\{[D(x)-n]^2\}+\frac{1}{2}E_{z-P_z(z)}\{D[G(z)-m]^2\} \quad (5)$$

$$\min_G V(G)=\frac{1}{2}E_{z-P_z(z)}\{D[(G(z)-l)^2]\} \quad (6)$$

式中，$E_{x-P_{\text{data}}(x)}\{\cdot\}$为真实样本分布，$E_{z-P_z(z)}\{\cdot\}$表示变量$z$服从随机噪声分布，$D(x)$为鉴别器，$G(z)$为生成器，$m$为生成器产生的降噪图像，$n$为真实清晰图像。生成器由原始图像与伪图构成，将鉴别器输出的结果取负，作为Loss对生成器进行更新，当鉴别器输出为0时，表示生成图像为False，输出为1时，表示生成图像为True。利用鉴别器实现$m$、$n$平方误差最小化后，得到鉴别为True的图像。

## 3 实验

### 3.1 实验条件

实验基于FLIR数据集进行，该数据集包含1.4万余幅红外图像和可见光图像。本实验采用的数据包为thermal-8-bit，共8 862幅红外图像，像素尺寸为640×512。

实验设备基于Windows环境，CPU型号为Intel(R) Core(TM) i7-8750H@2.2 GHz，显卡型号为NVIDIA GeForce GTX 1 060 6 GB，内存为16 GB(2 667 MHz)，算法实现上采用Python和Pytorch相结合，完成深度学习建模和模型训练。

在训练过程中，生成器的动量设置为0.5，学习率设置为0.003，鉴别器动量设置为0.9，学习率设置为0.001，采用Adam(Adaptive Moment Estimation)算法[16]解决深度学习中的大数据量，高特征纬度的优化问题。训练110个周期，每次训练个数为8。使用网络模型时，经鉴别器鉴别为True的图像输出，鉴别为False的图像重新送入生成器进行优化，直至输出为True。

### 3.2 实验结果及分析

从测试集中随机选取4幅红外图像进行实验分析。为便于对比降噪效果，本实验还选取了其他3种典型算法，例如BM3D(Block-Matching and 3D Filtering)[17]，DnCNN(Deep Neural Neural Network For Image Denoising)[18]和NBNet(Noise Basis Learning For Image Denoising With Subspace Projection)[19]。其中，BM3D通过对图像的每个参考块进行相似块匹配，然后，对其进行协同变换和滤波，经整合后实现图像去噪；DnCNN在深度学习中加入残差网络、批量归一化操作，可提升图像去噪效果，但泛化能力较弱；NBNet是一种经典的子空间去噪算法，可将噪声图像分为干净信号子空间和噪声信号子空间，通过去除噪声空间的分向量实现去噪。图3所示为几种算法的红外图像去噪效果。

在红外图像质量提升效果上，通过BM3D法滤波后的图像边缘信息更为平滑，图像清晰度有较大程度下降，相于其他算法在图像细节处理上有明显劣势；DnCNN法存在一定程度伪影；NBNet法与本文算法在滤波、细节信息保留方面较前两种算法效果要好，能够在保持图像边缘梯度的同时尽可能还原细节信息。

为进一步客观评价去噪效果，选用峰值信噪比(PSNR)、结构相似性(SSIM)、运算效率(处理时间$t$)三项参数的平均值作为评价指标，进行客观对比分析，结果如表1所示。从表中可以看出，本文算法在PSNR和SSIM指标方面表现效果最佳，分别达到了34.36，0.985 2 dB。NBNet其次，BM3D与DnCNN差距较大；在运算时间上，NBNet最短，本文算法与NBNet较为接近，而BM3D与DnCNN耗时较长。表明本文算法虽然在处理时间上不占优势，但在图像降噪与质量提升方面有较大改善。

**表1 不同算法指标对比**

| 算法 | $PSNR$/dB | $SSIM$ | $t$/s |
|---|---|---|---|
| BM3D | 29.53 | 0.807 6 | 0.132 |
| DnCNN | 32.25 | 0.912 8 | 0.320 |
| NBNet | 32.94 | 0.926 7 | 0.282 |
| 本文 | 34.36 | 0.985 2 | 0.250 |




图 3　不同算法红外图像去噪效果比较

## 4　结论

本文提出的基于子空间投影的生成对抗网络红外图像去噪框架，以子空间投影与 U-Net 网络相结合改进生成器，以红外图像为输入，经生成器处理后，再输入鉴别器进行对抗训练，可得到更为清晰的低噪红外图像。实验研究表明，本文提出的红外图像端到端去噪算法，可在有效抑制噪声信息的同时尽可能保留更多红外图像边缘细节，提高图像清晰程度，与其他典型算法相比有更好的图像质量提升效果。下一步工作将在红外图像降噪算法的轻量化方面进行优化，以提高算法运行效率。

**参考文献：**

[1] 李文博，王　琦，高　尚. 基于深度学习的红外小目标检测算法综述[J]. 激光与红外，2023，53(10)：1476-1484.
Li Wenbo，Wang Qi，Gao Shang. Overview of infrared small target detection algorithms based on deep learning[J]. Laser & Infrared，2023，53(10)：1476-1484.

[2] 孙婷婷，崔少华，孔令坤，等. 基于 curvelet 变换的红外图像去噪方法研究[J]. 山西师范大学学报(自然科学版)，2023，37(2)：37-40.
Sun Tingting，Cui Shaohua，Kong Lingkun，et al. Infrared image denoising method based on curvelet transform research [J]. J. of Shanxi Normal University：Natural Science Edition，2023，37(2)：37-40.

[3] 张　骏，朱　标，沈玉真，等. 基于引导滤波的多分支注意力残差红外图像去噪网络[J]. 红外与激光工程，2022，51(11)：159-169.
Zhang Jun，Zhu Biao，Shen Yuzhen，et al. Multi-drop attention residual infrared image denoising network based on guided filtering[J]. Infrared and Laser Engineering，2022，51(11)：159-169.

[4] 陈佛计，朱　枫，吴清潇，等. 基于生成对抗网络的红外图像数据增强[J]. 计算机应用，2020，40(7)：2084-2088.
Chen Foji，Zhu Feng，Wu Qingxiao，et al. Infrared image data augmentation based on generative adversarial network[J]. J. of Computer Applications，2020，40(7)：2084-2088.

[5] 吴国瑞，王　峰，周平华，等. 基于改进生成对抗网络的红外图像增强算法[J]. 半导体光电，2023，44(5)：782-787.
Wu Guorui，Wang Feng，Zhou Pinghua，et al. Infrared image enhancement algorithm based on improved generative adversarial network[J]. Semiconductor Optoelectronics，2023，44(5)：782-787.

[6] 周　洋，李翠梅，申　攀. 基于子空间投影的舰船图像轮廓特征提取方法[J]. 舰船科学技术，2023，45(11)：167-170.
Zhou Yang，Li Cuimei，Shen Pan. Method for extracting contour features of ship images based on subspace projection [J]. Ship Science and Technology，2023，45(11)：167-170.

[7] Arjovsky M，Chintala S，Bottou L. Wasserstein GAN[EB/

OL]. 2017: https://arXiv.orglabs/1701.07875.

[8] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs[EB/OL]. 2017: https://arxiv.org/abs/1704.00028.

[9] 张鸿雁，李俊丽，赵 丽. 基于聚类融合的中波红外图像目标提取方法[J]. 激光杂志，2023，44(8)：105-109.
Zhang Hongyan, Li Junli, Zhao Li. Based on clustering fusion of medium wave infrared image target extraction method[J]. J. of Laser, 2023, 44(8): 105-109.

[10] 高子航，刘兆英，张 婷，等. 基于对抗域适应的红外舰船目标分割[J]. 数据采集与处理，2023，38(3)：598-607.
Gao Zihang, Liu Zhaoying, Zhang Ting, et al. Infrared ship target segmentation based on adversarial domain adaptation[J]. Data Acquisition and Processing, 2023, 38(3): 598-607.

[11] 孔令琴，吴小溪，董立泉，等. 基于近红外视频的非接触鲁棒性心率监测[J]. 中国激光，2023，50(21)：167-173.
Kong Lingqin, Wu Xiaoxi, Dong Liquan, et al. Non-contact robust heart rate detection based on near infrared video[J]. Chinese J. of Laser, 2023, 50(21): 167-173.

[12] 韩林凯，姚江伟，王坤峰. 保留梯度和轮廓的可见光与红外图像融合[J]. 计算机应用，2023，43(11)：3574-3578.
Han Linkai, Yao Jiangwei, Wang Kunfeng. Visible and infrared image fusion preserving gradient and contour[J]. Applications of Computers, 2019, 43(11): 3574-3578.

[13] 刘志东. 改进神经网络的船舶红外图像边缘检测方法[J]. 舰船科学技术，2023，45(7)：166-169.
Liu Zhidong. Improved neural network for ship infrared image edge detection[J]. Ship Science and Technology, 2023, 45(7): 166-169.

[14] Yan R, He Q, Liu Y, et al. Unpaired virtual histological staining using prior-guided generative adversarial networks[J]. Comput. Med. Imaging Graphics, 2023, 105: 102185.

[15] Tam T N, Trung T H, Tam M P, et al. Validating functional redundancy with mixed generative adversarial networks[J]. Knowledge-Based Syst., 2023, 264: 110342.

[16] Chen Y, Wang Z, Bai X. Fussy sparse subspace clustering for infrared image segmentation[J]. IEEE Trans. Image Process., 2023, 32: 2132-2023.

[17] Dabov K, Foi A, Katkovnik V, et al. Image denoising by spare 3-D transform-domain collaborative filtering[J]. IEEE Trans. Image Process., 2017, 16(8): 2080-2095.

[18] Zhang K, Zuo W, Chen Y, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. IEEE Trans. Image Process., 2017, 26(7): 3142-3155.

[19] Cheng S, Wang Y, Huang H, et al. Nbnet: Noise basis learning for image denoising with subspace projection[C]// Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2021: 4896-4906.

## 作者简介：

**闫 宁**(1995－)，女，河南省郑州市人，硕士，讲师，主要研究方向为光电信息处理；

**周 斌**(1979－)，男，河南省修武县人，博士，副教授，主要研究方向为光电信息处理。