摘要

视觉感知是人类获取外界信息的主要方式之一，也是计算机视觉领域的核心任务。视觉感知依赖于光子采集过程，在低光照场景（例如照度低于 1 lux）下，设备采集的物理信号强度难以超过底层电子器件的噪声强度，导致图像与视频出现动态范围小、信噪比低、色彩失真以及细节缺失等问题。这些问题严重降低了图像的视觉质量，一方面难以满足人类对画面质量的要求，影响场景细节辨识；另一方面还会进一步对高层次视觉任务的性能与鲁棒性造成显著影响。

低光照视觉重建的难点在于，严重的噪声破坏了原始的物理信号分布，算法在不依赖物理先验的情况下难以进行有效的信号建模，进而完成信号处理与增强。针对上述挑战，本文从利用物理先验信息辅助低光照增强的角度出发，围绕物理先验建模和表示展开研究，提出充分利用成像原始数据、环境光照数据与物理空间建模构建视觉增强框架，通过提取低光照场景的物理先验引导模型对低光照数据进行视觉增强。在难以充分捕获物理信号的场景下，物理先验建模偏差较大、泛化能力有限。针对这一局限，本文进一步提出构建由预训练大模型引导的自监督学习框架，辅助信号处理过程，利用模型的预训练知识和表征能力对低光照数据进行先验建模与视觉增强。具体而言，本文的主要创新点包括以下四个方面：

第一，本文提出了一种基于物理成像中原始数据先验的低光照增强方法。针对低光照场景信号处理流程导致的物理信息缺失问题，围绕原始 Raw 数据与处理后 RGB 数据特性展开解耦分析，挖掘 Raw 数据包含的物理信息进行先验建模，结合记录成像参数的元数据构建了低光照图像增强框架。构建了双向非线性映射模块，完成非线性数据与线性数据的相互转化，同时构建了线性域下的曝光增强网络，简化数据的噪声及光照建模，增强了模型去噪、色偏校正以及细节恢复的能力。实验表明，该框架能够提升低光照场景下的视觉重建与噪声去除能力，对比同类方法在图像质量评估指标 PSNR 上提升 7.88 dB。

第二，本文提出了一种基于环境光照物理信息进行先验嵌入的低光照增强方法。针对低光照场景物理光照多尺度分布预测难点，通过多摄像头同步成像，获取环境光照信息，辅助模型完成重建，挖掘了联合成像框架在低光照增强任务上的潜力。构建了一个低光照联合成像数据集，支持模型训练和测试。数据三元组分别提供原始输入、环境光照信息以及优化目标。提出了环境光照先验信息嵌入的成像流程，分别针对数据

对齐、多尺度信息恢复以及光照与纹理一致性，设计了可变形卷积、自适应频带回归以及双路逆映射等模块。实验表明，该框架可自适应不同场景的光源特性进行光照重建，对比同类方法在图像质量评估指标 PSNR 上提升 1.94 dB。

第三，本文提出了一种基于多帧图像进行场景空间物理先验建模的低光照增强方法。针对低光照增强算法面临的空间建模信息缺失问题，利用混合神经表征及其归纳偏置实现低光照信号的物理空间建模与噪声去除，并探究深度图像先验对低光照输入的表征能力。利用深度卷积网络天然难以拟合高熵信号的特性，通过学习低光照场景的紧致语义以及基于深度模型的隐式空间表征，构建了无需依赖大规模标注数据的自监督增强模型。采用混合高斯模型进行信号熵控制，并结合提出的多色光 Retinex 模型实现光照解耦增强，对低光照数据进行噪声抑制、光照调节以及色偏校正。实验表明，该方法能够在自监督训练的场景下恢复图像细节以及抑制噪声，对比同类方法以及级联方法在图像质量无参考评估指标 NIQE 上提升 0.40。

第四，本文提出了一种预训练大模型知识引导物理先验建模的低光照增强方法。针对物理先验建模存在的先验刻画统计分布偏差，利用大模型通过预训练捕获的正常图像分布构建知识引导机制，加强图像光照先验知识获取途径，构建了自监督学习及模型增强蒸馏方法，缓解了物理先验泛化性差的局限性。通过对低光照图像提取光照不变表征以及借助大模型知识辅助模型建立低光照—正常光照映射关系，提升无监督训练效果。提出光照不变表征注入的特征对齐策略，并对训练目标进行多维度物理信号融合增强，进一步提升模型的主客观质量并降低模型复杂度。实验表明，该方法在多个数据集上均表现出性能优势，对比同类方法在图像质量评估指标 PSNR 上提升 1.44 dB，支持实时推理，适配实际应用需求。

关键词：物理先验引导、低光照增强、自监督学习、预训练大模型引导

Physical Prior-Guided Low-Light Enhancement Technology

Haofeng Huang (Computer Application Technology)
Directed by Prof. Ling-Yu Duan

ABSTRACT

Visual perception is a primary conduit through which humans acquire external information and serves as a core research task in the field of computer vision. However, low-light scenarios pose substantial challenges to computer vision systems in practical applications. Due to weak signal intensity and severe noise interference, low-light images are characterized by a narrow dynamic range, low signal-to-noise ratio, color distortion, and loss of structural details. These defects degrade visual quality, hinder human perception of scene details, and impair the performance and robustness of high-level downstream vision tasks.

The core difficulty of low-light visual reconstruction lies in that severe noise distorts the original physical signal distribution, making it challenging for existing algorithms to achieve effective signal modeling and enhancement in the absence of physical priors. To tackle these issues, this dissertation investigates physical prior modeling and representation techniques for low-light image enhancement. It proposes novel frameworks that integrate raw imaging data, ambient illumination data, and physical spatial modeling, whereby physical priors are adopted to guide model optimization. Furthermore, to mitigate poor generalization capability and modeling bias in scenarios where physical signals are difficult to capture, a self-learning framework guided by pre-trained large models is developed. This framework leverages the powerful knowledge and representation capabilities of large models to implement prior modeling and visual enhancement for low-light image data. The main innovations of this dissertation are summarized as follows:

First, a low-light enhancement framework based on a physical information guidance mechanism for raw imaging data is proposed. To address the physical information loss occurring in conventional signal processing pipelines, the proposed framework decouples the attribute characteristics of raw data and processed RGB data to explore inherent physical priors from raw imaging data. Combined with metadata that records imaging parameters, a bidirectional

nonlinear mapping module is constructed to realize the mutual conversion between linear and nonlinear image data. Meanwhile, an exposure enhancement network is developed in the linear domain to simplify noise and illumination modeling, thereby improving the performance of denoising, color correction, and detail restoration. Experimental results demonstrate that the proposed method achieves a PSNR improvement of 7.88 dB over state-of-the-art methods.

Second, a low-light enhancement framework based on ambient illumination modeling is proposed. To address the challenge of modeling physical light distribution in complex low-light scenes, this framework adopts multi-camera synchronous imaging to capture ambient physical information. A joint imaging dataset is constructed, which provides paired triplets of original input images, ambient light information, and optimization targets. Guided by ambient physical priors, the framework adopts an intelligent imaging pipeline that integrates deformable convolutions, adaptive frequency band regression, and dual-path inverse mapping to optimize data alignment, multi-scale detail recovery, and fidelity regularization. Experimental results verify a PSNR improvement of 1.94 dB.

Third, a low-light enhancement method utilizing Neural Radiance Fields (NeRF) for physical spatial modeling is presented. To overcome the limitation of physical space perception for low-light image data, NeRF and its inductive biases are introduced for signal space modeling and denoising, fully exploiting the representation ability of deep image priors. Taking advantage of the inherent property of deep convolutional networks to avoid fitting high-entropy noise signals, a self-supervised model is established to learn compact semantic features and physical spatial representations without relying on large-scale labeled annotations. Furthermore, Gaussian Mixture Models are adopted for entropy control, and a multi-color Retinex model is employed for illumination decoupling, achieving effective noise suppression and accurate color correction. Experimental results show a 0.40 improvement in the no-reference image quality metric NIQE.

Fourth, a low-light enhancement method leveraging knowledge from pre-trained large models to assist physical prior modeling is proposed. For low-light scenarios with severe modeling bias caused by sparse signal information, the proposed framework exploits the normal image distribution learned by generative large models to break through the generalization bottleneck. A self-supervised learning and enhancement distillation scheme is developed, which utilizes illumination-invariant representations and large-model prior knowledge to facilitate

accurate mapping from low-light images to normal-light images. In addition, feature alignment via illumination-invariant feature injection and multi-dimensional physical signal fusion further improves both subjective visual effect and objective quantitative quality. Extensive experimental results across multiple datasets demonstrate a 1.44 dB PSNR gain, with optimized inference latency to support real-time practical applications.

KEY WORDS: Physical Prior Guidance, Low Light Enhancement, Self-Supervised Learning, Guidance of Pretrained Large Models