摘要

光谱成像技术可同时获取目标场景的空间和光谱信息，在遥感探测、医学诊断、环境监测与材料识别等领域具有重要意义。然而，现有光谱相机多基于传统积分式视觉传感器，其成像质量受到传感器数据传输带宽、时间分辨率和动态范围限制，难以在高速运动、极端光照等复杂环境中实现有效成像。神经形态相机抛弃了传统的积分信号采集模式，通过脉冲或事件的形式表达场景的光强变化。凭借其低数据冗余、高时间分辨率以及高动态范围等特性，神经形态相机在多种复杂视觉任务中展现出了优越潜力。因此，深入研究神经形态相机在复杂环境下对光谱成像的增强机制，有望显著提升光谱成像系统在复杂环境条件下的成像性能。

本文围绕“复杂场景跨模态神经形态光谱成像质量增强机制”这一科学问题，针对可控光照条件下的高速光谱成像、自然光照条件下的动态光谱成像和极端光照条件下的可见光-红外融合成像三种复杂成像场景，系统探究了神经形态数据与光谱数据在时空信息利用、响应关系建模与融合机制设计三方面的协同策略与方法。通过构建成像理论、设计算法方案与研发硬件装置，实现了上述复杂环境中的高质量光谱成像，最终搭建了面向实际应用的神经形态综合光谱成像系统。本文的主要创新成果如下：

第一，提出了基于脉冲先验引导的压缩感知超高速光谱成像方法，用于解决可控光照场景下光谱相机因数据带宽大导致成像帧率受限的问题。该方法结合神经形态脉冲与压缩感知两种先进采样方式，建立了脉冲相机与压缩感知光谱相机协同的双分支成像模型，并依据该模型对光谱视频重建问题进行建模与求解。为充分利用神经形态脉冲数据中的时空信息并引导光谱信息重建，深入探索了连续脉冲信号在光谱重建过程中发挥的积极作用，构建了能够有效表征脉冲时空特性的脉冲光谱先验网络用于光谱先验约束，并同时设计了多模态即插即用光谱视频重建算法。实验表明，该方法能够在火焰爆燃等高速场景下实现最高 20,000 帧/秒的超高速光谱视频成像。

第二，提出了事件-光谱信号关系驱动的快照式光谱相机去模糊成像方法，用于解决自然光照场景下光谱相机因曝光时间长导致动态成像模糊失真问题。该方法将神经形态事件相机与快照马赛克光谱相机相结合，深入分析了事件信号与光谱帧信号之间的跨模态响应关系，并推导出光谱感知的事件双积分模型，该模型显式引入相机光谱响应函数，对神经形态信号和光谱信号的关系进行准确建模，为事件引导的光谱帧去模糊提供了理论指导。在此基础上，设计了扩散引导的噪声感知训练框架，通过去噪扩散概率模型对数据中的噪声分布进行精细建模，并建立噪声特征反馈机制以显著提升去模糊网络性能。最终，开发了一种事件增强的多光谱帧去模糊网络，该网络采用双分支结构用于分别估计事件双重积分及光谱系数，并辅以空间-光谱互补特征学习与

双模态交互模块，显著改善了高速模糊光谱帧的成像质量。实验表明，相对于传统基于图像的去模糊方案，该方法可将光谱帧的峰值信噪比提高 3dB 以上。

第三，提出了事件引导的多任务协同可见光-红外光谱融合成像方法，用于解决极端光照场景下可见光相机因动态范围窄导致融合成像质量不佳的问题。该方法引入长波红外相机进行跨波段光谱融合成像，并利用事件相机替代传统的积分式可见光相机，显著提升了极端光照环境下的成像能力。通过深入研究事件流与红外帧的多模态融合机制，引入多任务协同融合框架，将融合过程细分为"基于事件的可见光图像重构"、"事件引导的红外图像去模糊"和"可见光-红外融合"三个明确的子任务。为实现跨任务信息交互和模态间的互补融合，进一步引入跨任务事件增强模块，充分利用可见光重构子任务中提取的纹理特征来指导红外去模糊，并通过互信息优化策略提升融合成像质量。实验结果表明，该方法在互信息融合评价指标上相较于先前最佳方法提升了 2%，并在熵和结构相似度等评价指标上表现最优。

第四，构建了神经形态综合光谱成像系统，以应对多波段成像需求和复杂环境挑战。该系统通过将神经形态相机与多波段光谱相机进行有机整合，形成了一个能够同步感知可见光、近红外及长波红外波段的协同成像方案。系统设计注重前端成像设备与后端处理软件的协同工作，采用多传感器协同成像装置、并行化数据采集与预处理单元，以及数据存储计算单元，实现了广域波段的全面光谱感知能力。在道路实际驾驶和无人机等场景的实验验证表明，该系统具备在多种复杂场景下进行高质量光谱数据采集与处理的能力，有效验证了神经形态相机引导的光谱成像方案的有效性。

综上所述，本文针对复杂环境下如何利用神经形态视觉技术实现跨模态高质量光谱成像这一核心问题，系统地开展了理论建模、算法设计和硬件实现的研究。涵盖了高速、动态模糊和极端光照等多种复杂场景，构建了相应的理论模型与成像方案，并最终搭建了面向实际应用的综合光谱成像系统。本文的研究成果显著提升了跨模态神经形态光谱成像质量增强研究的理论深度和实际应用能力，为异构视觉数据协同成像技术提供了新的研究范式和实践参考。

关键词：神经形态相机，光谱成像，压缩感知，图像复原，可见光-红外融合

Research on Neuromorphic Camera-Guided Spectral Imaging Technology under Complex Environments

Mengyue Geng (Computer Application Technology)
Supervised by Prof. Yonghong Tian

ABSTRACT

Spectral imaging technology can simultaneously capture both spatial and spectral information, which has demonstrated significant application value in remote sensing, medical diagnosis, environmental monitoring, material identification and other fields. However, existing spectral cameras usually rely on conventional integration-based visual sensors, whose imaging quality is limited by sensor data transmission bandwidth, temporal resolution and dynamic range. Such limitations make it challenging to achieve effective imaging in complex environments, such as high-speed motion and extreme illumination. Compared to conventional visual sensors, neuromorphic cameras abandon integration-based signal acquisition. Instead, they represent scene intensity changes in the form of spikes or events, featuring low data redundancy, high temporal resolution, and high dynamic range. These advantages have demonstrated great potential in various complex visual tasks. Therefore, an in-depth investigation into the enhancement mechanisms of neuromorphic cameras for spectral imaging in complex environments is expected to significantly improve the imaging performance of spectral imaging systems under challenging conditions.

This thesis addresses the scientific challenge of "quality enhancement mechanisms in cross-modal neuromorphic spectral imaging under complex environments." It systematically investigates the collaborative strategies between neuromorphic data and spectral data in three key aspects: spatiotemporal information utilization, response relationship modeling, and fusion mechanism design. These studies are conducted for three complex imaging scenarios: high-speed spectral imaging under controllable lighting conditions, dynamic spectral imaging under natural lighting conditions, and visible-infrared fusion imaging under extreme lighting conditions. By constructing imaging theories, designing algorithmic solutions, and developing hardware devices, high-quality spectral imaging has been achieved in these complex environments. Finally, a neuromorphic comprehensive spectral imaging system for practical applications has been established. The main innovations of this thesis are listed as follows:

First, a spike prior-guided compressive spectral imaging method is proposed to address the problem of limited frame rates in spectral cameras caused by large data bandwidth under controllable lighting conditions. This method integrates two advanced sampling mechanisms—neuromorphic spike sampling and compressive sensing—by establishing a dual-branch imaging model that enables collaboration between a spike camera and a compressive sensing spectral camera. Based on this model, the spectral video reconstruction problem is formulated and solved. To fully exploit the spatiotemporal information embedded in neuromorphic spike data and guide spectral reconstruction, the positive role of continuous spike signals in the reconstruction process is thoroughly explored. A spike spectral prior network is developed to effectively characterize the spatiotemporal features of spike signals and serve as a spectral prior constraint. Additionally, a plug-and-play multimodal spectral video reconstruction algorithm is designed. Experimental results show that the proposed method can achieve ultra high-speed spectral video imaging at up to 20,000 FPS in high-speed scenes such as flame deflagration.

Second, an event-spectral signal correlation-driven deblurring method is proposed to address the motion blur issue in dynamic spectral imaging under natural lighting conditions caused by long exposure times of spectral cameras. The method pairs a neuromorphic event camera with a snapshot mosaic hyperspectral sensor, establishing a spectral-aware event double-integral model through systematic analysis of cross-modal event-spectral relationships. The model explicitly incorporates the camera spectral response functions to accurately model the relationship between neuromorphic signals and spectral signals, providing theoretical foundations for event-assisted spectral frame deblurring. A diffusion-guided noise-aware training framework is subsequently developed, where denoising diffusion probabilistic models precisely characterize sensor noise distributions, creating a noise-adaptive feedback mechanism. The final implementation integrates these advancements into an event-enhanced hyperspectral deblurring network with dual-modality interaction and spatial-spectral complementarity modules, demonstrating significant quality improvement on high-speed motion-blurred spectral frames compared to conventional approaches. Experiments show that this method can improve the peak signal-to-noise ratio (PSNR) of spectral frames by more than 3 dB compared to traditional image-based deblurring methods.

Third, an event-guided multi-task collaborative fusion method is proposed to address the issue of poor fusion quality under extreme lighting conditions, which arises due to the limited dynamic range of visible light cameras. This method incorporates a long-wave infrared camera for cross-band spectral fusion imaging and replaces traditional integrative visible light cameras with event cameras, significantly enhancing imaging capabilities in extreme lighting

environments. By deeply studying the multi-modal fusion mechanism of event flows and infrared frames, a multi-task collaborative fusion framework is introduced. The fusion process is divided into three distinct subtasks: event-based visible image reconstruction, event-guided infrared image deblurring, and visible-infrared fusion. A cross-task event enhancement module is designed to enable inter-task feature sharing, where textures extracted from the visible reconstruction subtask are leveraged to guide infrared image deblurring. Meanwhile, a mutual-information optimization strategy is implemented to significantly enhance the quality of visible-infrared fusion results. Experimental results show that this method improves the mutual information evaluation metric by 2% compared to the previous best methods. It also performs optimally in other evaluation metrics such as entropy and structural similarity.

Fourth, a neuromorphic comprehensive spectral imaging system is developed to meet the demands of multi-band imaging and address the challenges of complex environments. The system integrates neuromorphic cameras with spectral cameras working at multiple spectral bands, enabling synchronous capture across visible, near-infrared (NIR), and long-wave infrared (LWIR) bands. The system design emphasizes the collaborative operation between front-end imaging devices and back-end processing software, incorporating multi-sensor synchronized imaging devices, parallelized data acquisition/preprocessing unit, and data storage-computation unit. This enables comprehensive spectral sensing capabilities across a wide range of spectral bands. Field tests in automotive and unmanned aerial vehicle (UAV) scenarios demonstrate the system's robust spectral acquisition and processing capabilities under complex environmental conditions, effectively verifying the efficacy of neuromorphic camera-guided spectral imaging.

In summary, this thesis systematically addresses the fundamental challenge of achieving cross-modal high-quality spectral imaging guided by neuromorphic vision in complex environments. The research covers diverse challenging scenarios, including high-speed motion, dynamic blur, and extreme lighting conditions. Effective imaging solutions are established through theoretical modeling, algorithmic design and hardware implementation. A comprehensive neuromorphic spectral imaging system has been ultimately implemented for real-world applications. These advancements significantly deepen theoretical understanding and enhance practical capabilities in cross-modal neuromorphic spectral imaging enhancement studies, while providing novel research paradigms and research references for heterogeneous visual sensor imaging systems.

KEY WORDS: Neuromorphic Camera, Spectral Imaging, Compressive Sensing, Image Restoration, Visible-Infrared Fusion