Yibo Wang

E-mail: ybwang@smail.nju.edu.cn, ybwang369@gmail.com

Website: https://scholar.google.com/citations?user=DY6UjdYAAAAJ

Tel.: (+86) 188-0525-0369 Address: Nanjing, 210023, P.R. China

Affiliation: School of Electronic Science and Engineering, Nanjing University

Research Interests: Object Detection; Semantic Segmentation; Video Understanding; Hyperspectral Image Processing; Computational Spectral Imaging;

EDUCATION BACKGROUND

Nanjing University, Nanjing, China 09/2021 - 06/2024

M.E., Signal and Information Processing Average Score: 88.28 Advisor: Prof. Xun Cao

Hunan University, Changsha, China 09/2017 - 06/2021

B.S., Electronic and Information Engineering Average Score: 87.02 Advisor: Prof. Leyuan Fang

Research Position, School of Robotics Advisor: Prof. Fuhai Li

LANGUAGE CERTIFICATE

English, 7.0, International English Language Testing System (IELTS) 06/2024

PROGRAMMING SKILLS

- **Python,** for deep learning (PyTorch, TensorFlow, Caffe2)
- **MATLAB,** for data analysis and graphing
- **C++,** for mobile terminal
- **Java,** for user-interface
- **Linux,** for server management

HONORS & AWARDS

Graduate:

- Third Prize (Team leader), the 5th Integrated Circuit EDA (Electronic Design Automation) Challenge
- Gold Prize (Team member), the 9th "Internet plus" Innovation and Entrepreneurship Competition
- First Prize (Team member), the 3rd China Graduate Artificial Intelligence Innovation Competition
- First Prize (Personal), Scholarship of Nanjing University
- Top 16 (Team HNU-CSUST), the 6th China University Alumni Association (Nanjing) Football League

Undergraduate:

- First Prize (Personal), Scholarship of Hunan University
- Second Prize (Personal), National Undergraduate Mathematical Contest
- Third Prize (Team leader), Undergraduate Electronic Design Contest of Hunan Province
- Honorable Mention (Team member), Interdisciplinary Contest in Modeling (ICM)
- Fourth Place (Personal), Gobang (Wuziqi) Competition of Hunan University

PAPERS & PATENTS

- K Zhou, **Y Wang**, T Lv, Q Shen, X Cao, *"Gaseous Object Detection"*. ***TPAMI***, 2024.
- K Zhou, L Cai, **Y Wang**, M Zhang, Q Shen, X Cao, *"Joint RGB-Spectral Decomposition Model Guided Image Enhancement in Mobile Photography"*. ***ECCV***, 2024.
- Z Shi, H Ye, T Lv, **Y Wang**, X Cao, *"Compact Self-adaptive Coding for Spectral Compressive Sensing"*. ***ICCP***, 2023.
- T Lv, H Ye, Q Yuan, Z Shi, **Y Wang**, S Wang, X Cao, *"Aperture Diffraction for Compact Snapshot Spectral Imaging"*. ***ICCV***, 2023.
- K Zhou, **Y Wang**, T Lv, Y Li, L Chen, Q Shen, X Cao, *"Explore Spatio-temporal Aggregation for Insubstantial Object Detection: Benchmark Dataset and Baseline"*. ***CVPR***, 2022.
- **Y Wang**, Q Shen, X Cao. *"A Hyperspectral-Based Scene Segmentation Method for Autonomous Driving"*. China Patent CN202211285162.5
- K Zhou, **Y Wang**, T Lv, L Chen, C Zi. *"Video-level Object Detection Model Training Method and Device, Equipment and Storage Medium"*. China Patent NO.202210308421.5

RESEARCH EXPERIENCE

Memory Build-in Self Test Automatic Grouping 11/2023 - 12/2023

Design a three-level grouping method, first roughly grouping based on hard constraints; Then use energy finite element modeling to standardize the samples; Finally, use the BIRCH algorithm to cluster the standardized samples.

Illumination Estimation 06/2023 - 11/2024

Propose an end-to-end network architecture to address the problem of hyperspectral illumination estimation. To fill the gap of non-synthetic datasets, a large dataset, HyperSpectral Image with Switching Illumination (HSISI), of real hyperspectral images dedicated to illumination estimation is constructed. The proposed method is also applicable to the color constancy problem of RGB-domain by *Canon EOS 6D*.

Spectral Scene Segmentation 08/2022 - 09/2023

(1) Explore the application of spectral vision in autonomous driving, using sparse spectral points to assist RGB cameras in solving the segmentation problem of homochromatic metamaterials.

(2) Deploy multispectral cameras on intelligent lawnmowers (supported by *Greenworks Tools*) to better identify grasslands and analyze their health using near-infrared wavelengths. We deploy U-Net with MobileNet backbone on the BPU (Brain Processing Unit, supported by *Horizon Robotics*) of the intelligent terminal to provide real-time segmentation ($\geq$ 30FPS).

Spectral Guided Tone Mapping 01/2022 - 11/2023

To explore additional applications of the multispectral sensor in smartphones, construct a dataset which is jointly captured by the smartphone (supported by *Honor Device*) and hyperspectral camera. The dataset provides RGB images with high spatial resolution and hyperspectral images with low spatial resolution. And we introduce additional spectral information and material segmentation into the HDRNet baseline and the PSNR results of the tone mapping output increase by about 2dB.

Video Object Detection for Gas 11/2020 - 11/2023

"Can we detect the gaseous substance with computer vision methods?". To solve the above problem, we endeavor on a rarely explored task named Gaseous Object Detection (GOD), which aims to localize the object with the following characteristics: (1) amorphous shape with indistinct boundary; (2) similarity to surroundings; (3) absence in color. Accordingly, we construct a large-scale dataset named GOD-Video (supported by *Lightgene*) is constructed and the corresponding video-level detection benchmark. Design a spatio-temporal aggregation loss function to leverage the consistency along the time axis. Aim to explore the feasibility of object detection from solid substances to gaseous substances in computer vision.

Hyperspectral Image Classification 01/2020 - 06/2020

Due to the lack of supervisory information in hyperspectral image classification tasks, introduce self-supervised learning. Pseudo labels are generated by clustering the spectral- spatio vectors, which are used as pre-training data for the Spatial-Spectral FCN model, and then fine-tuned using real labels. Improve accuracy by 2% to 12% on several remote sensing datasets.

ACADEMIC SERVICES & WORK EXPERIENCE

- Review for TPAMI
- Algorithm Engineer, Nanjing Lightgene Technology Co., Ltd. 06/2022 - 09/2023
- Teaching Assistant of Computational Photography Course, Nanjing University 09/2021 - 01/2022
- Volunteer Teacher, Guanxi Wanquan Primary School 07/2018 - 08/2018

REFERENCES

- Lab for Computational Imaging Technology and Engineering (CITE): https://cite.nju.edu.cn/
- Developer Platform for Spectrum: https://www.specnet.cn/
- Xun Cao: Professor, PhD advisor, Dean of Department of Communication Engineering, School of Electronic Science and Engineering, Nanjing University, caoxun@nju.edu.cn
- Qiu Shen: Associate Professor, PhD advisor, Nanjing University, shenqiu@nju.edu.cn
- Tao Yue: Associate Professor, PhD advisor, Nanjing University, yuetao@nju.edu.cn
- Hao Zhu: Assistant Professor, PhD advisor, Nanjing University, zhuhaoese@nju.edu.cn

PULICATION ABSTRACT

- ***Master Thesis: Hyperspectral Image Illumination Estimation and Applications***

  Full text: https://github.com/yeebowang/NJU-MasterThesis

  Code: https://github.com/yeebowang/hyperspectral-illumination-estimation

The overall color of RGB images shot under different illuminations will change with the illumination. For hyperspectral images with richer spectral characteristics, the impact of the illumination is more significant. The spectral characteristics of the illumination may confuse the original spectral characteristics of the substance. In order to obtain the intrinsic spectral characteristics of a substance, it is necessary to obtain the spectrum of the illumination and correct the hyperspectral image. illumination estimation is an indirect method to obtain the illumination spectrum, which plays an important role in scenes where the illumination spectrum cannot be accurately measured.

This paper focuses on the problem of illumination spectrum estimation for hyperspectral images and applications. Firstly the illumination spectrum of the hyperspectral image is predicted, then the illumination spectrum is used to guide the image enhancement of the RGB image, and finally the illumination spectrum is used to correct the RGB image for material segmentation. The main work and innovation points include the following aspects:

1. Illumination spectrum estimation. A real rather than simulated hyperspectral illumination estimation data set is constructed. The spectral joint spatial self-attention mechanism is proposed to highlight the important spectral channels and effective areas in the hyperspectral image. A smoothing constrained loss function is proposed to avoid unrealistic ill-conditioned results, improving the accuracy of illumination estimation.

2. Image Enhancement guided by illumination spectrum. An image enhancement data set in which RGB images and hyperspectral images are aligned is constructed. The local brightness adaptation method is proposed to adjust the brightness of different areas of High Dynamic Range (HDR) RGB images. The Spectral Perception Self-Attention (SPSA) mechanism is proposed to embed the illumination spectrum into the RGB feature, improving the tone-mapping performance of RGB images.

3. Material segmentation with illumination spectrum correction. A material segmentation data set in which RGB images and hyperspectral images are aligned is constructed. The spectral joint RGB decomposition model is proposed, using the illumination spectrum to correct the reflectance of RGB images, effectively improving material segmentation accuracy of RGB images.

- ***TPAMI: Gaseous Object Detection***

  DOI: 10.1109/TPAMI.2024.3449994

Object detection, a fundamental and challenging problem in computer vision, has experienced rapid development due to the effectiveness of deep learning. The current objects to be detected are mostly rigid solid substances with apparent and distinct visual characteristics. In this paper, we endeavor on a scarcely explored task named Gaseous Object Detection (GOD), which is undertaken to explore whether the object detection techniques can be extended from solid substances to gaseous substances. Nevertheless, the gas exhibits significantly different visual characteristics: 1) saliency deficiency, 2) arbitrary and ever-changing shapes, 3) lack of distinct boundaries. To facilitate the study on this challenging task, we construct a GOD-Video dataset comprising 600 videos (141,017 frames) that cover various attributes with multiple types of gases. A comprehensive benchmark is established based on this dataset, allowing for a rigorous evaluation of frame-level and video-level detectors. Deduced from the Gaussian dispersion model, the physics-inspired Voxel Shift Field (VSF) is designed to model geometric irregularities and ever-changing shapes in potential 3D space. By integrating VSF into Faster RCNN, the VSF RCNN serves as a simple but strong baseline for gaseous object detection. Our work aims to attract further research into this valuable albeit challenging area.

- ***ECCV2024: Joint RGB-Spectral Decomposition Model Guided Image Enhancement in Mobile Photography***
  POSTER: https://eccv.ecva.net/virtual/2024/poster/2020
  Code: https://github.com/CalayZhou/JDM-HDRNet

  The integration of miniaturized spectrometers into mobile devices offers new avenues for image quality enhancement and facilitates novel downstream tasks. However, the broader application of spectral sensors in mobile photography is hindered by the inherent complexity of spectral images and the constraints of spectral imaging capabilities. To overcome these challenges, we propose a joint RGB-Spectral decomposition model guided enhancement framework, which consists of two steps: joint decomposition and priors-guided enhancement. Firstly, we leverage the complementarity between RGB and Low-resolution Multi-Spectral Images (Lr-MSI) to predict shading, reflectance, and material semantic priors. Subsequently, these priors are seamlessly integrated into the established HDRNet to promote dynamic range enhancement, color mapping, and grid expert learning, respectively. Additionally, we construct a high-quality Mobile-Spec dataset to support our research, and our experiments validate the effectiveness of Lr-MSI in the tone enhancement task. This work aims to establish a solid foundation for advancing spectral vision in mobile photography.

- ***ICCP2023: Compact Self-adaptive Coding for Spectral Compressive Sensing***
  DOI: 10.1109/ICCP56744.2023.10233711

  Spectral snapshot compressive imaging (SCI) has been extensively studied and applied to various fields. Although the typical coded aperture snapshot spectral imaging (CASSI) presents an effective paradigm, its fixed code designs do not sufficiently exploit flexible and optimal modulation in consideration of scene sparsity. In this paper, we present a novel Compact Self-adaptive optical Coding framework for Spectral Compressive Sensing, termed 3CS, to optimize the coded pattern adaptively for better hyper-spectral videos perception. Our framework enables extracting context high-frequency components from the compressed domain without requiring hybrid guiding camera. The specifically designed mask distribution enables higher light efficiency, and is robust against temporal correlation reduction when processing dynamic spectral videos. Extensive experiments and model discussions validate the superiority of the proposed framework over traditional end-to-end (E2E) methods in various aspects for the spectral reconstruction.

- ***ICCV2023: Aperture Diffraction for Compact Snapshot Spectral Imaging***
  DOI: 10.1109/ICCV51070.2023.00970
  Code: https://github.com/Krito-ex/CSST

  We demonstrate a compact, cost-effective snapshot spectral imaging system named Aperture Diffraction Imaging Spectrometer (ADIS), which consists only of an imaging lens with an ultra-thin orthogonal aperture mask and a mosaic filter sensor, requiring no additional physical footprint compared to common RGB cameras. Then we introduce a new optical design that each point in the object space is multiplexed to discrete encoding locations on the mosaic filter sensor by diffraction-based spatial-spectral projection engineering generated from the orthogonal mask. The orthogonal projection is uniformly accepted to obtain a weakly calibration-dependent data form to enhance modulation robustness. Meanwhile, the Cascade Shift-Shuffle Spectral Transformer (CSST) with strong perception of the diffraction degeneration is designed to solve a sparsity-constrained inverse problem, realizing the volume reconstruction from 2D measurements with Large amount of aliasing. Our system is evaluated by elaborating the imaging optical theory and reconstruction algorithm with demonstrating the experimental imaging under a single exposure. Ultimately, we achieve the sub-super-pixel spatial resolution and high spectral resolution imaging.

- ***CVP2022: Explore Spatio-temporal Aggregation for Insubstantial Object Detection: Benchmark Dataset and Baseline***
  DOI: 10.1109/CVPR52688.2022.00311
  Code: https://github.com/CalayZhou/IOD-Video

  We endeavor on a rarely explored task named Insubstantial Object Detection (IOD), which aims to localize the object with following characteristics: (1) amorphous shape with indistinct boundary; (2) similarity to surroundings; (3) absence in color. Accordingly, it is far more challenging to distinguish insubstantial objects in a single static frame and the collaborative representation of spatial and temporal information is crucial. Thus, we construct an IOD-Video dataset comprised of 600 videos (141,017 frames) covering various distances, sizes, visibility, and scenes captured by different spectral ranges. In addition, we develop a spatio-temporal aggregation framework for IOD, in which different backbones are deployed and a spatio-temporal aggregation loss (STAloss) is elaborately designed to leverage the consistency along the time axis. Experiments conducted on IOD-Video dataset demonstrate that spatio-temporal aggregation can significantly improve the performance of IOD. We hope our work will attract further researches into this valuable yet challenging task.