Efficient Snapshot Spectral Imaging: Calibration-Free Parallel Structure with Aperture Diffraction Fusion ———-Supplementary Material———

Tao Lv[1,3], Lihao Hu[1,3], Shiqiao Li[1,3], Chenglong Huang[1,3], and Xun Cao[1,2,*]

[1] School of Electronic Sciene and Engineering, Nanjing University, Nanjing, China
[2] Key Laboratory of Optoelectronic Devices and Systems with Extreme Performances of MOE, Nanjing University, Nanjing, China
[3] NJU-ZPTECH University-Enterprise Joint Laboratory for Computational Spectral Imaging
{lvtao, lihaohu, lishiqiao, Chenglong-Huang}@smail.nju.edu.cn, caoxun@nju.edu.cn

In the supplementary material, we sequentially present essential details related to the experimental setup. Section 1 details the orthogonal mask parameters utilized in the real experiments. Section 2 exhibits the test dataset derived from simulation experiments. In Section 3, the captured raw data from PCADIS and the outcomes subsequent to alignment are showcased for analysis. In Section 4, we delve into the specifics of our proposed algorithm SSFCT and present the preprocessing methodology utilized in our simulation to compare various algorithms. Meanwhile, to enhance reader comprehension regarding snapshot spectral imaging (SSI) methods, including system complexity and reliance on repeated calibrations, we meticulously examine these methods in Section 5.

1 Orthogonal masks in real experiments

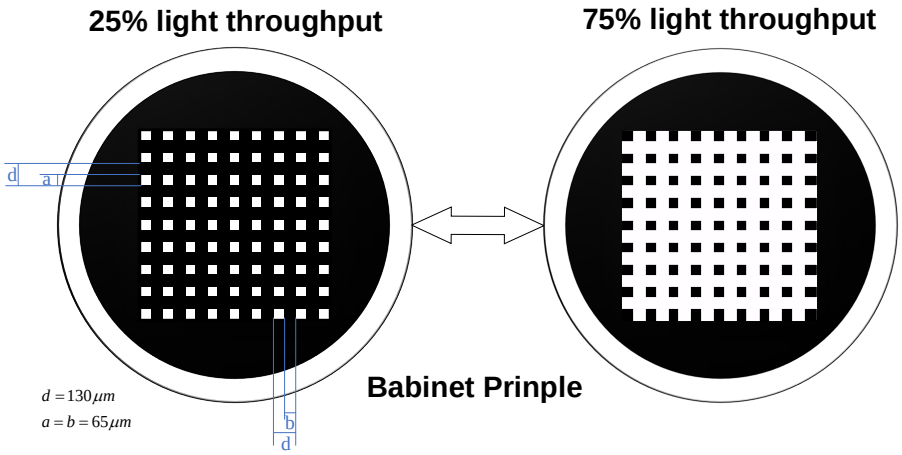

Fig. 1: Schematic representation of the orthogonal mask used in real experiments.

The illustration of the Orthogonal Mask (OM) utilized in the real experiments in this paper is shown in Fig. 1, which is uniformly densely packed with opaque lines of $65\mu m$

linewidth at a period of $130\mu m$ in the orthogonal direction. In the specific experiments, the mask is closely fit to the lens instead of being placed on the principal plane of the lens, which facilitates the adjustment of the overall imaging focal length and depth of field of the system. Moreover, the generation of training data in real experiments under Uncalibrated (UC) mode necessitates the consideration of the quantum efficiency inherent in the monochromatic camera. Then the simulation-generated Monochromatic Aperture Diffraction (MAD) used for training is generated by integrating the wave-by-wave with the dot-multiplication of the quantum efficiencies.

The OM's utilization follows the design principles of the aperture diffraction imaging spectrometer (ADIS) [1], where the light throughput of OM is fixed at 25% with $a = b = 65\mu m$. This configuration induces **missing orders in the even-numbered diffraction grades**, simplifying the diffraction complexity and alleviating ill-posedness in the reconstruction process. Leveraging the Babinet principle, complementary aperture pattern yield analogous diffraction patterns, which also has been integrated into PCCADIS.

Furthermore, the **monochromatic camera's response** across various bands served as **coefficients for the PSFs** in both real-world and simulated experiments, which were employed in measurement simulation and model training.

2 Test set of simulation experiment

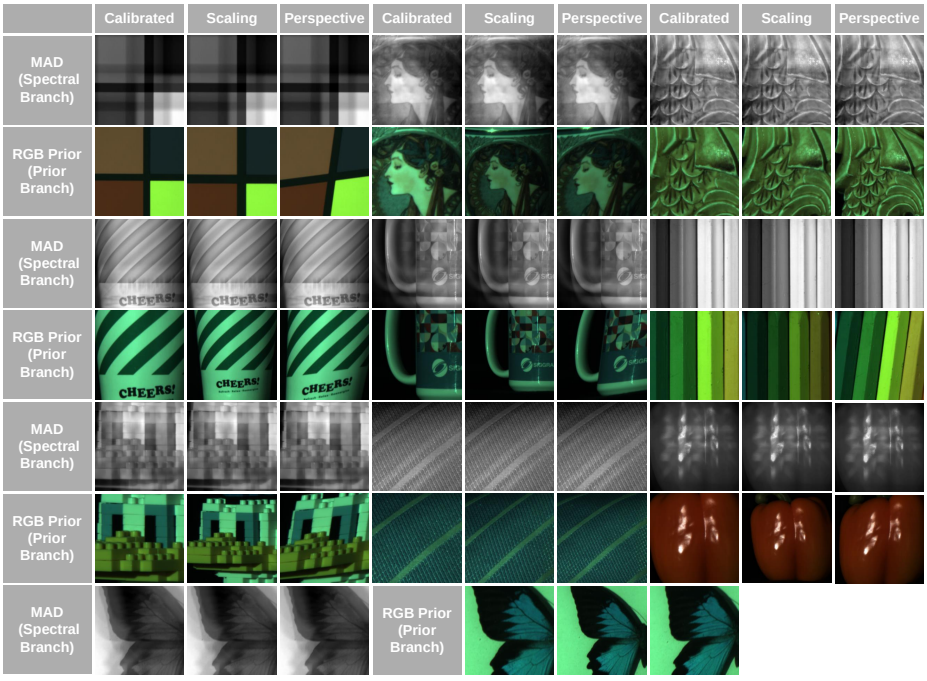

Fig. 2: The testsets used in the simulation experiments contained 10 scenarios, each with 3 different RGB Prior images.

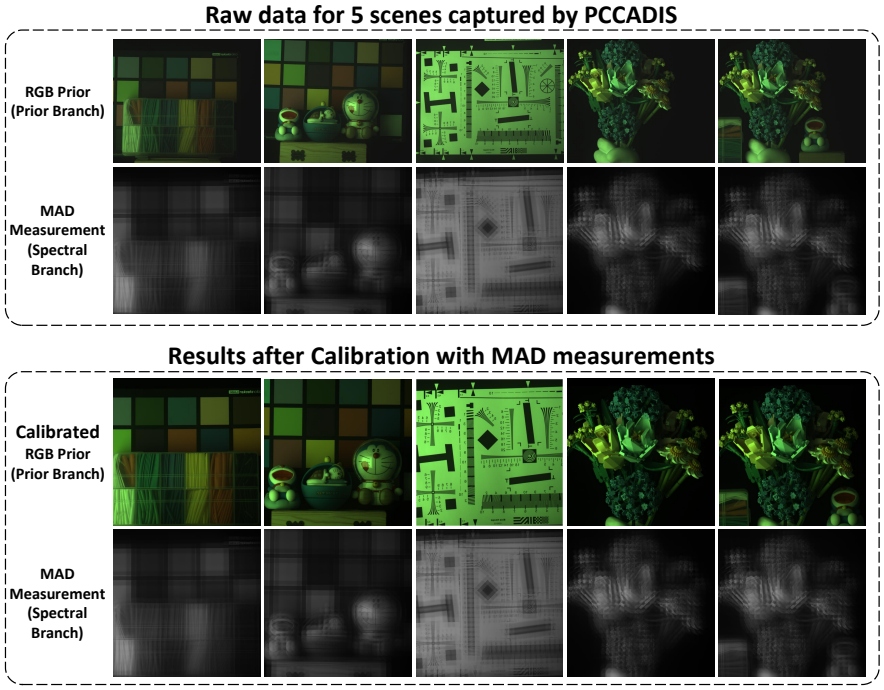

Fig. 3: Real data captured by the PCCADIS prototype and the registration results.

The test set employed in the simulation experiments is illustrated in Fig. 2, comprising 10 distinct scenarios, each featuring 3 prior images. These sets are instrumental in assessing the efficacy of bootstrap reconstruction in addressing under-aligned images within the present context. In the comparative simulation experiments involving UC inputs, the prior images are derived from 296×296 resolution images, sharing the same geometric center as the 256×256 MAD measurements, and are deflated and perspective-varied using a homography matrix. The prior images contain 3 forms: calibrated, deflated (resolution transformation: $296 \times 296 \rightarrow 256 \times 256$), and perspective-varied (transform coordinates from $([0,0],[20,276],[276,20],[296,296])$ to $([0,0],[0,256],[256,0],[256,256])$).

3 Data Captured in Real experiments

The unprocessed data from 5 scenes, acquired during real experiments, are presented in Fig. 3. Notably, the two data channels captured by PCCADIS exhibit discernible discrepancies in the absence of artificial registration or optical path calibration. Within this study, well-registered images serve to validate the imaging proficiency of PCCADIS under optimal registration conditions. Additionally, to assess imaging performance when subjected to UC or unregistered inputs, an artificial perspective transformation of RGB prior is performed based on the MAD measurements.

4 Algorithm details and preprocessing

The **simulation and real experimentation** of all algorithms is **fully end-to-end** without any fine-tuning and pre-training, and the structure and parameters of the decoder of SFFCT are identical to those of [1]. CSST handles the reconstruction in the SSFCT back-end. SFEM processes uncalibrated or unregistered inputs to generate well-registered features under L_{align}. We innovate to achieve SOTA performance in the combined framework of fusion and reconstruction. CDDFuse [2] itself cannot be used for spectral super-resolution (SR) or reconstruction.

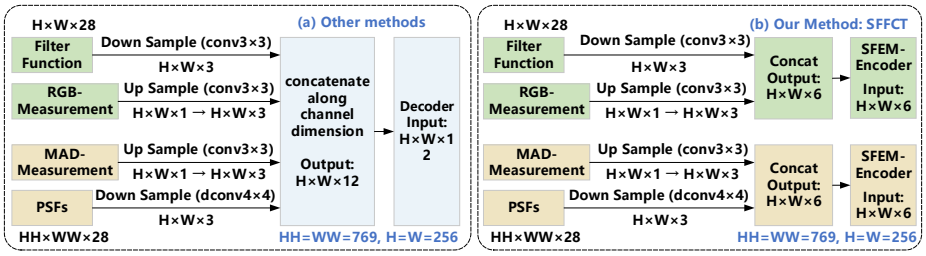

Fig. 4: preprocessing of SSFCT and other SOTA methods.

For fair comparisons, Fig. 4 displays the preprocessing of SSFCT and other SOTA methods, and we just adjust the shape of the input by convolution to fit different algorithms. The assessment of SR algorithms refers to experimental schemes from works like MST [3], MST++ [4], DAUHST [5], CSST [1], etc. However, due to variations in projection, CASSI-related parallel reconstruction algorithms, like MST [3], DGSMP [6], **deeply optimized with masks**, **do not seamlessly transfer to the PCCADIS reconstruction**.

5 System structure complexity and calibration complexity analysis

To illustrate the challenges inherent in current SSI methodologies, specifically regarding system complexity and the necessity for recurrent calibration, we examine three representative SSI methods (CASSI [8], CTIS [10], PMVIS [11]). These methods are analyzed in detail, highlighting their system structures, all of which adhere to the common 4f-structured imaging framework.

• As depicted in Fig. 5(a), CASSI employs a spatially encoded random mask and a dispersion device. During calibration, it's crucial to record the absolute position of the spatial encoding for each band using a monochromatic light source. This necessitates recalibration if the system is perturbed, such as when temperature fluctuations alter the prism **dispersion relationship** or when there is a **pixel-level shift** in the random mask's position.

• As depicted in Fig. 5(b), CTIS utilizes diffractive optical elements (DOEs) that equal to gratings oriented in different directions to achieve spectral information projection encoding of the scene. Due to the intricate projection relationships, it's imperative

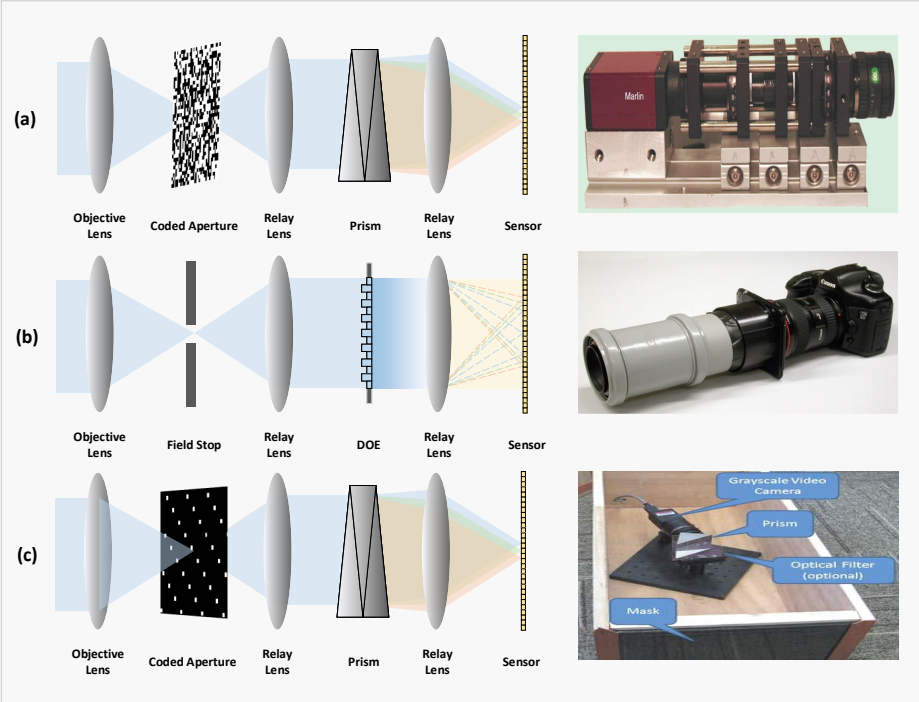

Fig. 5: (a) Schematic and physical drawings (from [7]) of the structure of CASSI [8]; (b) Schematic and physical drawings (from [9]) of the structure of CTIS [10]; (c) Schematic and physical drawings (from [11] of the structure of PMVIS [11]
.

to restrict the imaging area using a field stop to prevent aliasing among different diffraction orders entirely. Consequently, precise calibration of the **absolute position of each diffraction order under each band** is essential.

• As depicted in Fig. 5(c), PMVIS employs sparse masks and prisms to acquire spectral information. Calibration is essential to determine the absolute position of each light-passing hole on the sensor and the corresponding dispersion relation for each hole. Similar to CASSI, even minor **perturbations in the mask or changes in dispersion relations** can significantly affect the imaging forward process, rendering the original calibration parameters ineffective.

When integrating the aforementioned methods with prior branches to enhance system reconstruction performance, much of the complex calibration has been omitted in related studies without further research, such as [12–14], etc. These studies primarily concentrate on the spectral reconstruction accuracy or imaging dimension improvements facilitated by the prior branch introduction to the imaging system. However, practical implementation of these systems in engineering settings is challenging due to their complex structures and reliance on meticulous calibration processes.

References

1. Tao Lv, Hao Ye, Quan Yuan, Zhan Shi, Yibo Wang, Shuming Wang, and Xun Cao. Aperture diffraction for compact snapshot spectral imaging. In *ICCV*, pages 10574–10584, 2023.
2. Zixiang Zhao, Haowen Bai, Jiangshe Zhang, Yulun Zhang, Shuang Xu, Zudi Lin, Radu Timofte, and Luc Van Gool. Cddfuse: Correlation-driven dual-branch feature decomposition for multi-modality image fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5906–5916, 2023.
3. Yuanhao Cai, Jing Lin, Xiaowan Hu, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, and Luc Van Gool. Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17502–17511, 2022.
4. Yuanhao Cai, Jing Lin, Zudi Lin, Haoqian Wang, Yulun Zhang, Hanspeter Pfister, Radu Timofte, and Luc Van Gool. Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 745–755, 2022.
5. Yuanhao Cai, Jing Lin, Haoqian Wang, Xin Yuan, Henghui Ding, Yulun Zhang, Radu Timofte, and Luc Van Gool. Degradation-aware unfolding half-shuffle transformer for spectral compressive imaging. *arXiv preprint arXiv:2205.10102*, 2022.
6. Tao Huang, Weisheng Dong, Xin Yuan, Jinjian Wu, and Guangming Shi. Deep gaussian scale mixture prior for spectral compressive imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16216–16225, 2021.
7. Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2013.
8. Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008.
9. Ralf Habel. 48-1: Invited paper: Spectral sensing with computed tomography imaging spectrometry. In *SID Symposium Digest of Technical Papers*, volume 48, pages 716–718. Wiley Online Library, 2017.
10. Michael Descour and Eustace Dereniak. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied optics*, 34(22):4817–4826, 1995.
11. Xun Cao, Hao Du, Xin Tong, Qionghai Dai, and Stephen Lin. A prism-mask system for multispectral video acquisition. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2423–2435, 2011.
12. Lizhi Wang, Zhiwei Xiong, Dahua Gao, Guangming Shi, and Feng Wu. Dual-camera design for coded aperture snapshot spectral imaging. *Applied optics*, 54(4):848–858, 2015.
13. Lei Yuan, Qiang Song, Hecong Liu, Kevin Heggarty, and Weiwei Cai. Super-resolution computed tomography imaging spectrometry. *Photonics research*, 11(2):212–224, 2023.
14. Yurong Chen, Yaonan Wang, and Hui Zhang. Prior image guided snapshot compressive spectral imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.