Physical Degradation Model-Guided Interferometric Hyperspectral Reconstruction with Unfolding Transformer

Supplementary Material

In this supplemental document, we provide additional analysis and details in support of the findings in the main manuscript. The contents are organized as follows:

- Section 1 provides a brief introduction to the interferometric hyperspectral imaging technology and the LASIS system utilized in this paper.
- Section 2 offers details and and error analysis experiments of the degradation model.
- Section 3 presents some details analysis on degradation model and unfolding stages, along with supplemental visualization results.

1. Overview of Interferometric Hyperspectral Imaging

In this section, we provide a concise introduction to the interferometric imaging spectroscopy technique and the LASIS system employed in this study, as well as the specific information regarding the radiometric calibration data.

1.1. Interferometric Hyperspectral Imaging

Interferometric Hyperspectral Imaging (IHI), also known as Fourier Transform Spectroscopy (FTS) [1, 3, 6], is a widely used hyperspectral imaging technology in large-scale aerospace remote sensing. Its foundational principle originates from the Fourier transform relationship between dual-beam interference intensity and the light source spectrum, discovered by Rayleigh in 1889. However, due to the lack of efficient computational tools at the time, this technology was not extensively applied. After 1965, with the advent of computers and the Fast Fourier Transform (FFT) algorithm, IHI gradually transitioned from laboratory research to practical applications. In the 1980s, imaging spectroscopy emerged as a critical technology in the field of optical remote sensing. In recent years, advancements in detector technology, high-speed data transmission, and processing capabilities have further expanded the applications of IHI in aerospace remote sensing, environmental monitoring, and resource exploration. The fundamental process of interferometric hyperspectral imaging is illustrated in Fig. 1.

Fundamental Principles of IHI. The core principle of interferometric hyperspectral imaging lies in Fourier transform spectroscopy. The basic concept involves splitting incident light into two coherent beams using an interferometer and generating interference patterns by varying the optical path difference (OPD). These interference patterns encode spectral information. Specifically, the interferogram across

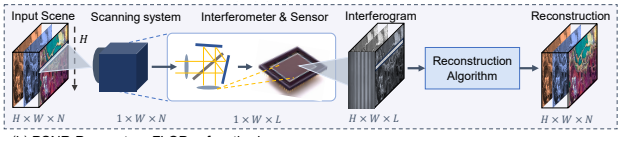

Figure 1. The fundamental process of interferometric hyperspectral imaging.

the full spectral range is the sum of individual monochromatic interference patterns (constant terms omitted):

$$I(l) = \int_0^\infty B(\nu)\cos(2\pi\nu l)\, d\nu, \tag{1}$$

where ν is the wavenumber ($\nu = 1/\lambda$, λ being the wavelength), l represents the OPD, and $B(\nu)$ denotes the spectral intensity as a real function in the wavenumber domain. Since $B(\nu)$ is real and $I(l)$ is an even function, the interferogram can be extended symmetrically:

$$I(l) = \int_{-\infty}^{+\infty} B(\nu)e^{-2\pi i\nu l}\, d\nu = \mathcal{F}\{B(\nu)\}. \tag{2}$$

Here, $\mathcal{F}$ denotes the Fourier transform. This allows the conversion of spatial or temporal interferograms into frequency-domain spectra. Compared to traditional dispersive spectrometers, Fourier transform spectroscopy offers superior spectral resolution, sensitivity, and wavenumber accuracy.

Classification. Interferometric Hyperspectral Imaging can be categorized into three types based on modulation mechanisms.

Time-Modulated Interferometric Imaging employs a moving mirror to continuously alter the OPD, capturing interferograms at different time points. Its core component is the Michelson interferometer,achieving time-modulated systems achieve high spectral resolution and sensitivity. Space-Modulated interferometric imaging systems utilize spatial partitioning to achieve OPD variations, exemplified by lateral shearing interferometers. Unlike time-modulated systems, they eliminate the need for moving mirrors, enhancing stability and enabling real-time imaging.

Spatiotemporal-Modulated interferometric imaging is a hybrid approach combines temporal and spatial modulation, often employing a push-broom scanning mechanism to acquire interferometric data. Spatiotemporal systems avoid slits and moving components, enabling high-throughput and high-resolution imaging in both spatial and spectral dimensions. They are particularly suited for large-area remote

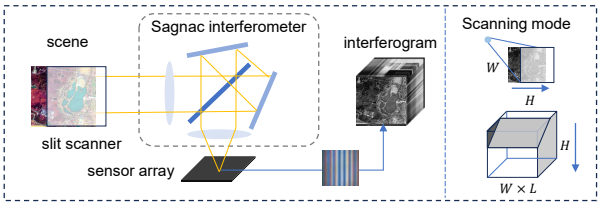

Figure 2. A schematic illustration of the principle of the LASIS system. Light passes through the Sagnac interferometer to scan the scene in the H direction. At a given moment, images are formed with different H positions and optical path differences, each representing a $W \times L$ frame. These frames are superimposed to fill the interferometric data cube.

sensing and high-resolution applications. LASIS is a representative example of a spatiotemporal-modulated system.

1.2. LASIS system

The Large Aperture Static Imaging Spectrometer (LASIS) [8] is an advanced spatiotemporal-modulated interferometric system. Its core component is a lateral shearing interferometer. Unlike conventional space-modulated systems, LASIS eliminates the entrance slit, significantly increasing optical throughput. Furthermore, its static design removes moving parts, enhancing reliability and stability.

As shown in pic 2, incident light passes through a front-end optical system and enters the lateral shearing interferometer. The beam splitter divides the light into two coherent beams, which are then recombined at the focal plane to form interference patterns. The absence of a slit allows LASIS to achieve wide-field and high-throughput imaging while maintaining high resolution. Similar to most scanning spectral imaging instruments, LASIS exhibits distinct stripe patterns along the push-broom direction (referred to as the H-direction in this study).

LASIS excels in aerospace remote sensing and environmental monitoring due to its high throughput, wide field of view, and static architecture. These features enable rapid acquisition of large-area spectral data with robust operational stability.

1.3. Radiometric Calibration

Radiometric calibration is a crucial technique in existing interferometric imaging reconstruction methods. It extracts important prior information about imaging degradation by analyzing data collected from a special device named integrating sphere, which provides uniform light environment with specific spectral profiles.

Radiometric calibration primarily relies on two types of data: absolute calibration data and relative calibration data. Absolute calibration data are obtained directly by the interferometer in the calibration field, containing interferograms of uniform light that have undergone the full degradation process. In contrast, relative calibration data are acquired by first removing the interferometric optical path from the imaging system and then capturing data under the same environment.

In relative calibration data, the interferometer functions as a scanning camera, allowing the direct separation of degradation components related to the sensor. For instance, in dark conditions, signal-independent current terms can be extracted, while under uniform illumination, shot noise and other components can be identified. By comparing relative calibration data with absolute calibration data, various degradations induced by the interferometric optical path, such as phase shifts and color offsets, can be isolated.

Although the specific methods of radiometric calibration vary, their fundamental principles and data formats are similar. The model parameter estimation approach proposed in this study is also based on standard radiometric calibration data, requiring no additional data acquisition. This makes the method practical.

2. Details in Degradation Modeling

2.1. Details of Degradation Model

In summary, the overall expression for the degradation model is as Eq. 3 and Eq. 4, and the degradation parameters are shown in Tab. 1.

$$\begin{aligned} \mathbf{I}_O &= \mathbf{M} \odot (\mathcal{F}\{\mathbf{A} \odot \mathbf{B}_0\} + \boldsymbol{\beta} \odot \bar{\mathbf{B}}_0), \\ \mathbf{I}_d &= \mathbf{K}' \odot (\mathbf{I}_O + \mathbf{N}_{\text{shot}}) + \mathbf{D}' + \mathbf{N}_{\text{read}}, \end{aligned} \tag{3}$$

$$\begin{aligned} &\mathbf{N}_{\text{shot}} \sim \mathcal{P}(\mathbf{I}_O), \\ &\mathbf{N}_{\text{read}} \sim \mathcal{N}(\mathbf{0}, \sigma'_{\text{read}}), \\ &\log(e') \sim \mathcal{N}(0, e),\ \mathbf{K}' = e' \cdot \mathbf{K}, \\ &\mathbf{D}' = e' \cdot \mathbf{D},\ \boldsymbol{\sigma}'_{\text{read}} = e' \cdot \boldsymbol{\sigma}_{\text{read}}. \end{aligned} \tag{4}$$

where $\odot$ is element-wise multiplication, $\mu(\cdot)$ denotes the mean and $\sigma(\cdot)$ denotes the standard deviation.

component	parameter	size	type
$\mathbf{I}_O$	$\mathbf{A}$	$W \times N$	trend term
	$\boldsymbol{\beta}$	$W \times L$	
	$\mathbf{M}$		
$\mathbf{K}'$	$\mathbf{K}$	$W \times L$	random noise
$\mathbf{D}'$	$\mathbf{D}$		trend term
$\mathbf{N}_{\text{read}}$	$\boldsymbol{\sigma}_{\text{read}}$		random noise
$\mathbf{K}', \mathbf{D}', \mathbf{N}_{\text{read}}$	e	1	random noise

Table 1. Parameters for the degradation model, including the scale and type of each parameter.

Additionally, within the degradation model, there is a detail that requires clarification. The dark current $\mathbf{D}$ is actually composed of two components, namely the dark current

mean $D_0 \in \mathbb{R}$ and the dark current pattern $\mathbf{D}_e \in \mathbb{R}^{W \times L}$, that is,

$$\mathbf{D} = D_0 + \mathbf{D}_e. \tag{5}$$

Here, the dark current mean D_0 represents the base noise across the entire sensor system, which is almost an invariant and unrelated to spatial conditions, being primarily associated with the overall state of the instrument. The dark current pattern $\mathbf{D}_e$, on the other hand, represents the portion of the dark current that varies among different sensors, influenced by lighting conditions and ISO settings, and possesses certain characteristics of random noise. Therefore, the parameter selection formula for electronic degradation $\mathbf{D}' = e' \cdot \mathbf{D}$ is essentially a simplified form, and the detailed expression is as:

$$\mathbf{D}' = D_0 + e' \cdot \mathbf{D}_e. \tag{6}$$

parameter	μ	σ	$\sigma/\mu(\%)$
$\mathbf{A}$	1.541	0.071	4.6
$\boldsymbol{\beta}$	2.974	0.032	1.1
$\mathbf{M}$	1.165	2.997e-3	0.3
$\mathbf{K}$	0.019	9.820e-4	5.1
D_0	127.538	0.224	0.2
$\mathbf{D}_e$	-2.177	0.111	4.9
$\boldsymbol{\sigma}_{\text{read}}$	2.259	0.125	5.5
e	0.1	-	-

Table 2. The error analysis of the model parameters is presented, with the table listing the mean and variance of the model parameters calibrated across multiple images in calibration dataset.

2.2. Error Analysis Experiments

Accuracy is significant for the degradation model, as it determines whether the model can accurately represent the real degradation process, generate realistic degraded data, and consequently influence the reconstruction performance of learning-based methods. To ensure accuracy during the development of the degradation model, we continuously validate it through error estimation and fitting evaluation on calibration data. Specifically, we conduct parameter error estimation during the calibration process, and the results are presented in Tab. 2. The error estimation method involve calculating errors and standard deviations across multiple calibration datasets with varying brightness levels and comparing them. As shown, all parameter errors are within 6%, with only $\mathbf{A}$ and $\boldsymbol{\sigma}_{\text{read}}$ reaching 5%. Since $\mathbf{A}$ contains complex components and the noise amplitude of $\mathbf{N}_{\text{read}}$ is relatively small, these errors are within acceptable limits.

Furthermore, since $\mathbf{M}$ is the primary factor causing instability in the interference curves, we validate its accuracy by removing $\mathbf{M}$ from the calibration data and performing an Augmented Dickey-Fuller (ADF) stationarity test. The experimental results yielded an average ADF value of -14.25 and an ADF p-value of 0 across different brightness levels, indicating strong stationarity. To validate the rationality of $\boldsymbol{\sigma}_{\text{read}}$ in the dark noise model $\mathbf{N}_{\text{read}}$, we divide the extracted zero-mean noise from the calibration data taken in the dark environment directly by $\boldsymbol{\sigma}_{\text{read}}$ to normalize it, and fit it with a normal distribution, as shown in Fig. 3. The normalized noise closely follows a normal distribution with σ of 1, confirming the accuracy of this design.

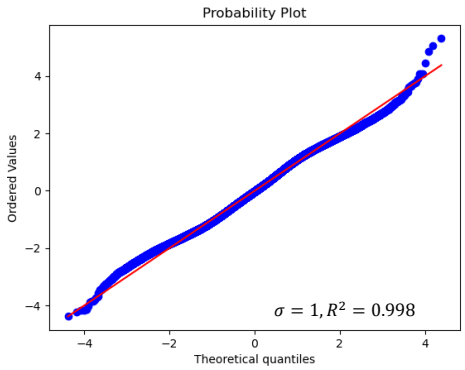

Figure 3. Validation of modeling of $\mathbf{N}_{\text{read}}$. The normalization of calibration data taken in a dark environment, followed by fitting to a normal distribution after division by $\boldsymbol{\sigma}_{\text{read}}$. The high degree of fitting confirms the accuracy of our noise modeling for $\mathbf{N}_{\text{read}}$.

Subsequently, we generate 4 scenes of simulated calibration data using the complete degradation model and compare it with real calibration data by calculating the Kullback-Leibler (KL) distance to investigate distributional differences, as shown in Tab. 3. The results demonstrate that the KL distance remains low across different brightness levels, and its increase aligns with the growth pattern of shot noise. This validates the accuracy of the simulation process.

	s0	s1	s2	s3
KL distance	0.00328	0.00457	0.00791	0.02431

Table 3. The simulation results of the degradation model were compared with the KL distance of real data, encompassing four scenes with average values of 0, 1000, 2000, and 3000, respectively. The overall distribution error is minimal, further validating the accuracy of the degradation model.

The above results confirm the effectiveness of the degradation model, which serves as a foundation for further research on IHI reconstruction tasks.

3. Additional Experimental Details

This section provides additional details on the experimental setup that were not extensively covered in the main text, along with supplementary analyses of the experiments.

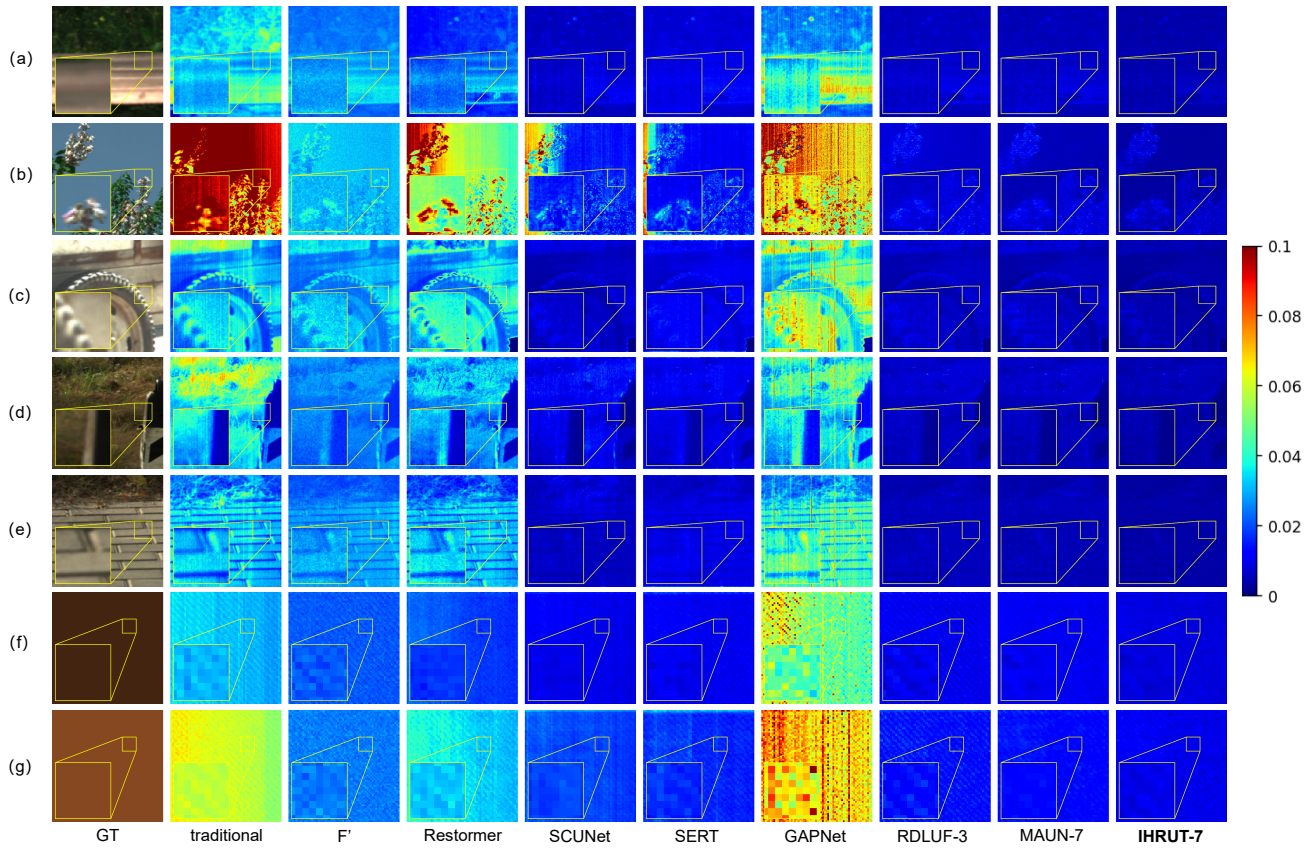

Figure 4. Supplementary Comparison of reconstructed HSI with error maps (10× amplified). We select Synthetic (a) Scene 1 (b) Scene 3 (c) Scene 4 (c) Scene 5 (d) Scene 6 from HSOD-BIT, cropped by 256×256, and Real (e) Scene 1 (f) Scene 2 from Calibration dataset cropped by 64×64. The error map exhibits the spectral absolute error of IHI reconstruction results by 8 algorithms and IHRUT with 7 stages. The region within the box is chosen for analysis of the reconstructed spectra and zoom in for a more detailed examination.

PSNR/SSIM	stages				
GFLOPs/params	3	4	5	6	7
PADUT	29.05 / 0.949 134.02 / 9.23M	29.46 / 0.975 167.86 / 11.61M	29.78 / 0.984 201.59 / 13.86M	29.41 / 0.968 235.66 / 16.12M	29.33 / 0.968 269.65 / 18.29M
RDLUF	37.61 / 0.976 155.72 / 9.29M	37.21 / 0.922 196.85 / 9.29M	36.43 / 0.914 232.48 / 9.29M	36.35 / 0.912 268.14 / 9.29M	35.50 / 0.898 302.31 / 9.29M
MAUN	37.82 / 0.986 68.76 / 2.03M	37.98 / 0.988 93.05 / 2.73M	38.14 / 0.989 117.34 / 3.41M	38.15 / 0.991 141.63 / 4.10M	38.15 / 0.992 165.92 / 4.78M
IHRUT	38.35 / 0.982 59.69 / 1.99M	39.12 / 0.989 83.04 / 2.60M	39.46 / 0.992 106.36 / 3.20M	39.55 / 0.992 126.13 / 3.71M	39.61 / 0.992 145.80 / 4.42M

Table 4. Reconstruction performance and complexity analysis of unfolding-based methods with different numbers of stages. Evaluation is conducted on Calibration dataset.

3.1. Details in Experimental Settings

Due to the unique requirements of IHI reconstruction, our dataset and methodology underwent necessary preparatory steps, which are described below.

Training Dataset. In this study, the training data for the IHI reconstruction task are primarily generated through simulation. Hyperspectral images (HSIs) serve as the ground truth (GT), while the simulated interferograms are used as the learning targets. The selection of the HSI dataset is guided by the following criteria:

- The spectral range must cover the 0-900 nm range of the LASIS system. (Datasets such as ICVL [] do not meet

this requirement.)

- The dataset must exhibit high quality to serve as reliable GT.
- To meet generalization requirements (e.g., denoising), the dataset should ideally include multiple independent scenes.

The HSOD-BIT dataset [7], originally designed for hyperspectral object detection, is chosen as it satisfies all the above conditions, as discussed in the main text. We select 46 images as training scenes. The original images contain slight noise, which is preliminarily removed using wavelet transformation. The images are then spatially downsampled to half their original size, followed by spectral linear interpolation to obtain HSIs of size $620 \times 840 \times 70$. Although the LASIS device used in this study has a constraint of $W = 2048$ and $\Lambda = 70$ in HSI dimensions, this scale is too large for direct training. Therefore, we uniformly crop 276 patches of size $256 \times 256 \times 70$ from the HSIs to form the final training set. During simulation, each patch is randomly assigned a shift along the W-direction to simulate real-world imaging scenarios.

The above preprocessing steps ensure both the quality and generalization capability of the training data.

Testing Dataset. The preparation of the testing datasets, HSOD-BIT and Houston, is relatively straightforward. For the four images from HSOD-BIT, we apply the same wavelet-based denoising method. Subsequently, all data are spatially upsampled to $W = 2048$ and cropped into multiple scenes of $H = 256$. From these scenes, six from HSOD-BIT and one from Houston are randomly selected as the test set. Since both datasets inherently have high spatial resolution, the low upsampling ratio does not compromise the validity of the test samples. Real-world data can be used for testing without any preprocessing.

3.2. Analysis of Unfolding Stages

In the reconstruction comparison experiments, all unfolding-based methods are selected for the optimal number of stages with the best performance. In reality, we conduct analysis experiments on each algorithm for 3 to 7 stages on real calibration datasets and find that the performance of these algorithms does not consistently increase with the number of stages. Tab. 4 present the complexity and reconstruction results of PADUT [5], RDLUF [2], MAUN [4], and our IHRUT across 3 to 7 stages. The peak performance of PADUT and RDLUF occurs at 5 and 3 stages, respectively. For PADUT, the generalization results and real reconstruction outcomes indicate that it has not fully adapted to IHI tasks, and it is more inclined to optimize SSIM than PSNR. RDLUF shares the same parameters across all stages, and thus additional iterations are detrimental to performance without a change in model expressiveness. In contrast, the performance of MAUN and IHRUT improves with an increasing number of stages, showing better adaptation to IHI tasks, and IHRUT achieves superior performance with fewer parameters and lower complexity.

In th unfolding stages of IHRUT, the optical degradation priors from degradation model are incorporated into operators $\mathbf{F}$ and $\mathbf{F}'$ (in Sec. 3.3 and Fig. 3 of main paper), playing a role as sensing matrix (conversion between interferomgrams and HSIs), guiding the unfolding iterations. Specifically, as visualized in rFig. 5, $\mathbf{F}'$ at the network input initially converts the degraded interferometric data into the spectral domain. Meanwhile, $\mathbf{F}$ and $\mathbf{F}'$ within the Data Modules impose constraints for the spectral reconstruction fidelity in unfolding iterations.

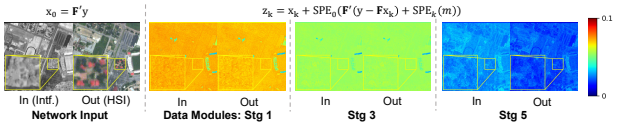

Figure 5. Input and output of the degradation prior components (operators $\mathbf{F}$ and $\mathbf{F}'$) at different parts of the network on Houston dataset.

3.3. Supplemental Visualization

In the comparative experiments, only three sets of visualization results are presented in the main text. For the remaining seven scenarios, we display them here, as shown in Fig. 4. Our visualization method is consistent with previous approaches, where the results are obtained by pixel-wise residual maps magnified by a factor of ten (values exceeding 0.1 are set to 0.1). Overall, IHRUT demonstrates superior performance, showing a more significant effect in the removal of degradation components compared to other algorithms.

References

[1] Robert Bell. *Introductory Fourier Transform Spectroscopy*. Academic Press, New York and London, 1972. 1

[2] Yubo Dong, Dahua Gao, Tian Qiu, Yuyan Li, Minxi Yang, and Guangming Shi. Residual degradation learning unfolding framework with mixing priors across spectral and spatial for compressive spectral imaging. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 22262–22271, 2023. 5

[3] P Fellgett. Theory of multiplex interferometric spectrometry. *J. Phys. Radium*, 19:187–191, 1958. 1

[4] Qian Hu, Jiayi Ma, Yuan Gao, Junjun Jiang, and Yixuan Yuan. Maun: Memory-augmented deep unfolding network for hyperspectral image reconstruction. *IEEE/CAA Journal of Automatica Sinica*, 11(5):1139–1150, 2024. 5

[5] Miaoyu Li, Ying Fu, Ji Liu, and Yulun Zhang. Pixel adaptive deep unfolding transformer for hyperspectral image reconstruction. In *Proceedings of IEEE International Conference on Computer Vision*, pages 12959–12968, 2023. 5

[6] Ernest V Loewenstein. The history and current status of fourier transform spectroscopy. *Applied Optics*, 5(5):845–854, 1966. 1

[7] Haolin Qin, Tingfa Xu, Peifu Liu, Jingxuan Xu, and Jianan Li. Dmssn: Distilled mixed spectral-spatial network for hyperspectral salient object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024. 5

[8] Bin Xiangli, Qisheng Cai, and Shusong Du. Large aperture spatial heterodyne imaging spectrometer: Principle and experimental results. *Optics Communications*, 357:148–155, 2015. 2