Supplementary Document for Uncertainty-Driven Spectral Compressive Imaging with Spatial-Frequency Transformer

Lintao Peng[1], Siyu Xie[1], and Liheng Bian[1]

Beijing Institute of Technology, Beijing, 100081, China. bian@bit.edu.cn

1 Overview

In the supplementary material, we provide some additional experimental details and provide the detailed mathematical proof of fast Fourier transform (FFT) [11] and inverse fast Fourier transform (IFFT) [12].

2 Inference Times of the Comparison Methods

We provide the memory cost, computational cost, and inference time of all the learning-based HSI reconstruction methods in Tab. 1, including λ-net [10], TSA-Net [9], DGSMP [7], HDNet [6], MST [2], CST [1], DAUHST [3], PADUT [8], RDLUF [4] and our Specformer . The inference time is the time to reconstruct 10 $256 \times 256 \times 28$ HSI cubes from 10 256×310 measurements. The hardware used is AMD Ryzen 7 3700X and NVIDIA RTX 3090.

Table 1: Memory cost, computational cost, and inference time of learning-based HSI reconstruction methods. The inference time is the time to reconstruct 10 $256 \times 256 \times 28$ HSI cubes from 10 256×310 measurements.

Methods	λ-net	TSA-Net	DGSMP	HDNet	MST	CST	DAUHST	PADUT	RDLUF	Ours
Params(M)	62.64	44.25	3.76	2.37	2.03	3.00	6.15	5.38	1.89	2.48
FLOPs(G)	117.98	110.06	646.65	154.76	28.15	40.01	79.50	90.46	231.09	39.85
Times(s)	1.12	1.38	6.84	1.69	1.97	2.15	1.61	4.42	4.53	1.83

3 Ablation Study of the Uncertainty-driven Loss

Inspired by the spatial sparsity of HSI, we propose to impose Jeffrey's prior [5] (JP) on uncertainty to improve the uncertainty estimation accuracy, thereby improving the HSI reconstruction quality. In addition, we used a Monotonically Increasing Function (MIF) $\hat{\boldsymbol{s}}_{\boldsymbol{i}} = \ln(1 + e^{\boldsymbol{s}_{\boldsymbol{i}}})$ to prioritize the pixels with large

Table R 2: Ablation study of JP and MIF. W/O means without.

	W/O JP	W/O MIF	Full model
PSNR/SSIM	40.22/0.979	40.06/0.978	**40.61/0.981**

uncertainty, which can also improve the HSI reconstruction quality. To demonstrate the effectiveness of JP and MIF, we conducted ablation experiments on them. The experimental results are shown in Tab. 2.

From Tab. 2, we observe that the absence of JP or MIF in the uncertainty-driven loss will result in a degradation of its performance, which demonstrates the effectiveness of JP and MIF.

4 From Fourier transform to discrete Fourier transform

In this section, we derive the DFT from the standard Fourier transform (FT). The standard Fourier transform was originally designed for continuous signals. It converts continuous signals from the time domain to the frequency domain and can be regarded as an extension of the Fourier series. Specifically, the Fourier transform of the signal $x(t)$ is given by

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt := \mathcal{F}[x(t)], \tag{1}$$

where j is the imaginary unit. The inverse Fourier transform (IFT) has a similar form to the Fourier transform

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)e^{j\omega t}d\omega. \tag{2}$$

From Eq. 1 and Eq. 2 we can see the duality of FT between the time domain and the frequency domain. The duality indicates that the properties in the time domain always have their counterparts in the frequency domain. Fourier transform has a variety of properties. For example, the FT of a unit impulse function is

$$\mathcal{F}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}dt = \int_{0-}^{0+} \delta(t)dt = 1, \tag{3}$$

and the time shifting property

$$\begin{aligned} \mathcal{F}[\delta(t-t_0)] &= \int_{-\infty}^{\infty} x(t-t_0)e^{-j\omega t}dt \\ &= e^{-j\omega t_0}\int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt \\ &= e^{-j\omega t_0}X(j\omega). \end{aligned} \tag{4}$$

Since it is difficult for computers to directly process continuous signals, in practical applications, the general approach is to sample continuous signals to obtain discrete signal sequences. Sampling can be achieved using a sequence of unit impulse functions

$$x_s(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s), \tag{5}$$

where T_s is the sampling interval. Taking the FT of the sampled signal $x_s(t)$ and applying Eq. 4 and Eq. 5, we have

$$X_s(j\omega) = \sum_{n=-\infty}^{\infty} x(nT_s)e^{-j\omega nT_s}. \tag{6}$$

The above equation shows that $X_s(j\omega)$ is a periodic function with the fundamental period as $2\pi/Ts$. Actually, there is always a correspondence between the discrete signal in one domain and the periodic signal in the other domain. Generally, we prefer a normalized frequency $\omega \leftarrow \omega T_s$ such that the period of $X_s(j\omega)$ is exactly 2π. We can further denote $x[n] = x(nT_s)$ as the sequence of discrete signal and derive the discrete-time Fourier transform (DTFT)

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}. \tag{7}$$

Assuming that the discrete signal $x[n]$ has finite length N, then DTFT becomes

$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x[n]e^{-j\omega n}, \tag{8}$$

without loss of generality, we assume that the nonzero terms lie in $[0, N-1]$. Note that the DTFT is a continuous function of ω and we can obtain a sequence of $X[k]$ by sampling $X(e^{j\omega})$ at frequencies $\omega_k = 2\pi k/N$

$$X[k] = X(e^{j\omega})|_{\omega=2\pi k/N} = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)kn}. \tag{9}$$

This is exactly the formula for 1D DFT. DFT plays an important role in the area of digital signal processing and is a crucial component in our SF-block.

Since DFT is a one-to-one transformation, given DFT $X[k]$, we can recover the original signal $x[n]$ by the inverse DFT (IDFT) as

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k]e^{j(2\pi/N)kn}. \tag{10}$$

Since DFT is conjugate symmetric, therefore for the real input $x[k]$, we can get $X[N-k] = X*[k]$. The reverse is true as well: if we perform IDFT to $X[k]$ which is conjugate symmetric, a real discrete signal can be recovered. This property implies that half of the DFT contains complete information about the frequency characteristics of $x[n]$. The fast Fourier transform (FFT) algorithms take advantage of the symmetry and periodicity properties of W_N^{kn} and reduce the complexity to compute DFT from $O(N^2)$ to $O(NlogN)$. The inverse DFT (Eq. 10), which has a similar form to the DFT, can also be computed efficiently using the inverse fast Fourier transform (IFFT).

The DFT described above can be extended to 2D signals, and the 2D DFT can be viewed as performing 1D DFT on the two dimensions alternatively. Given the 2D signal $X[m,n]$, $0 \leq m \leq M-1$, $0 \leq n \leq N-1$, the 2D DFT of $x[m,n]$ is given by

$$X[u,v] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m,n] e^{-j2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}. \tag{11}$$

Similar to 1D DFT, 2D DFT of real input $x[m,n]$ satisfied the conjugate symmetry property $X[M-u,N-v] = X*[u,v]$. The FFT algorithms can also be applied to 2D DFT to improve computational efficiency. In addition, FFT and IFFT operations are well-supported by hardware accelerators (like GPUs) through cuFFT and mkl-fft libraries, and there is a mature interface in Pytorch. This ensures that our Specformer can perform these computations fast, achieving stable and efficient training.

References

1. Cai, Y., Lin, J., Hu, X., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Gool, L.V.: Coarse-to-fine sparse transformer for hyperspectral image reconstruction. In: ECCV (2022)
2. Cai, Y., Lin, J., Hu, X., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Gool, L.V.: Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction. In: CVPR (2022)
3. Cai, Y., Lin, J., Wang, H., Yuan, X., Ding, H., Zhang, Y., Timofte, R., Gool, L.V.: Degradation-aware unfolding half-shuffle transformer for spectral compressive imaging. In: NIPS (2022)
4. Dong, Y., Gao, D., Qiu, T., Li, Y., Yang, M., Shi, G.: Residual degradation learning unfolding framework with mixing priors across spectral and spatial for compressive spectral imaging. In: CVPR. pp. 22262–22271 (2023)
5. Figueiredo, M.: Adaptive sparseness using jeffreys prior. Advances in neural information processing systems **14** (2001)
6. Hu, X., Cai, Y., Lin, J., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Gool, L.V.: Hdnet: High-resolution dual-domain learning for spectral compressive imaging. In: CVPR (2022)
7. Huang, T., Dong, W., Yuan, X., Wu, J., Shi, G.: Deep gaussian scale mixture prior for spectral compressive imaging. In: CVPR. pp. 16216–16225 (2021)
8. Li, M., fu, Y., Liu, J., Zhang, Y.: Pixel adaptive deep unfolding transformer for hyperspectral image reconstruction. In: ICCV (2023)
9. Meng, Z., Ma, J., Yuan, X.: End-to-end low cost compressive spectral imaging with spatial-spectral self-attention. In: ECCV. pp. 187–204. Springer (2020)
10. Miao, X., Yuan, X., Pu, Y., Athitsos, V.: l-net: Reconstruct hyperspectral images from a snapshot measurement. In: ICCV. pp. 4059–4069 (2019)

11. Nussbaumer, H.J., Nussbaumer, H.J.: The fast Fourier transform. Springer (1982)
12. Vaibhav, V.: Fast inverse nonlinear fourier transform. Physical Review E **98**(1), 013304 (2018)