

DiriNet: An Estimation Network for Spectral Response Function and Point Spread Function

Ting Hu[✉], Siyuan Cheng, Chang Liu

Abstract: Hyper- and multi-spectral image fusion is an important technology to produce hyper-spectral and hyper-resolution images, which always depends on the spectral response function and the point spread function. However, few works have been paid on the estimation of the two degradation functions. To learn the two functions from image pairs to be fused, we propose a Dirichlet network, where both functions are properly constrained. Specifically, the spatial response function is constrained with positivity, while the Dirichlet distribution along with a total variation is imposed on the point spread function. To the best of our knowledge, the neural network and the Dirichlet regularization are exclusively investigated, for the first time, to estimate the degradation functions. Both image degradation and fusion experiments demonstrate the effectiveness and superiority of the proposed Dirichlet network.

Keywords: Dirichlet network; point spread function; spectral response function; hyper-spectral image; multi-spectral image

1 Introduction

Limited by the tradeoff between the spatial and spectral resolution, remote-sensing spectrometers always can capture hyper-spectral images (HSI) with poor spatial resolution or multi-spectral images (MSI) with poor spectral resolution. Fortunately, images with both high-spatial and high-spectral resolution become possible, as hyper- and multi-spectral image (HMI) fusion methods emerge. Existing HMI fusion methods can be roughly divided into fusion based on modeling,

deep learning, and deep learning equipped with modeling. Modeling-based fusion refers to a mathematic model where the inherent properties or the physical degradations are considered for fused images, such as the fusion techniques established on spectrum unmixing [1], sparse representation [2], and low-rank approximation [3]. Deep learning-based fusion [4–6] that learns the mapping from HMI to their fusion target via a deep network, has become a popular fusion method. Recently, several interesting studies [7–9] have been published on HMI fusion by combining modeling and deep learning. However, most fusion techniques usually depend on the point spread function (PSF) and the spectral response function (SRF). Endless studies on the HMI fusion remain now, but few works are published to estimate the PSF and SRF.

Generally, the observed HSI $\mathcal{X} \in \mathbb{R}^{m \times n \times B}$ and MSI $\mathcal{Y} \in \mathbb{R}^{M \times N \times b}$ are modeled as the spatial and spectral degradations of the ground-truth target $\mathcal{Z} \in \mathbb{R}^{M \times N \times B}$, respectively, where $M \gg m$, $N \gg n$, and $b \ll B$ are individually the horizon-

Manuscript received Mar. 20, 2024; revised Apr. 15, 2024; accepted May 15, 2024. The associate editor coordinating the review of this manuscript was Dr. Xudong Zhao. The work was supported in part by the Postdoctoral Science Foundation of China (No. 2023M730156) and the National Natural Foundation of China (No. 62301012).

Ting Hu is with Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China.

Siyuan Cheng is with Space Star Technology Co., Ltd, Beijing 100086, China.

Chang Liu is with Research Institute of Intelligent Wireless Communication Network Technology, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China.

✉ Corresponding author. Email: hut455@bjut.edu.cn

DOI: [10.15918/j.jbit1004-0579.2024.044](https://doi.org/10.15918/j.jbit1004-0579.2024.044)

tal, vertical, and spectral dimension of an image. Namely,

$$\mathcal{X} = \mathcal{D}(\mathcal{Z} * \Phi) \quad (1)$$

$$\mathcal{Y} = \mathcal{Z}_{\times 3} \mathbf{R}^T \quad (2)$$

where $\mathcal{D}(\cdot)$ denotes the downsampling operator of interval r , $r = \frac{M}{m} = \frac{N}{n}$, which is ideally set as the Nyquist value; $\Phi \in \mathbb{R}^{r \times r}$ is the PSF, $\mathbf{R} \in \mathbb{R}^{B \times b}$ is the SRF, \mathbf{R}^T expresses the transposition for the SRF, and $\mathcal{Z}_{\times 3} \mathbf{R}^T = \mathbf{Z} \mathbf{R}$ with $\mathbf{Z} \in \mathbb{R}^{M \times N \times B}$ being the matrix stacked horizontally by the column vectors which are rearranged from all bands of \mathcal{Z} . The linear relationship between the observed HSI and MSI is easily inferred as

$$\mathcal{X}_{\times 3} \mathbf{R}^T = \mathcal{D}(\mathcal{Y} * \Phi) \quad (3)$$

Obviously, the estimation of the degradation functions is a mathematic ill-posed problem. One feasible methodology is to impose prior constraint on the PSF and SRF for solving this problem [10]. Simões et al. published a subspace-based hyperspectral superresolution (HySure) [11], which depended on the PSF and SRF. Hence, they estimated the PSF under the co-constraint of the sum-to-one property and the spatial smoothness while calculating the SRF by minimizing its horizontal variation. To reduce the computational burden, the HSI and MSI are blurred certainly, and the SRF and PSF are then estimated successively. Although this optimized method is often used in HMI fusion, it is revealed with worse effect when applied in the case of limited SRF overlaps [12]. Finlayson's [13] method that expressed the smoothness of SRF using a finite linear combination of band-limited bases, was found to be suitable for such limited cases [12]. However, it works only when PSF is known. Recently, degradation procedures of HSI and MSI have been integrated into HMI fusion networks [14–16]. Then, the two degraded functions are successfully computed once the fusion is accomplished. Nevertheless, these methods aim to pur-

sue better fusion results rather than accurate estimations for the degraded functions.

Attracted by the superior learning ability of neural networks [17, 18], we do the first attempt to design a Dirichlet network (DiriNet) to learn the PSF and SRF from pairs of HMI. In the proposed network, the spatial blurring of Eq. (1) and spectral degradation of Eq. (2) are implemented by the two dimensional (2D) convolution and depthwise separable convolution [19], respectively. When completing the training for DiriNet, the convolution filters are determined. Then, the SRF is the dimension-squeeze result of the 2D convolution filter, and the PSF can be obtained by taking any piece from the depthwise convolution filter. To effectively constrain the solution space of the PSF and SRF, several regularizations are applied. Since the PSF obeys the Dirichlet distribution, the stick-breaking process [20] is first used to describe it. Moreover, the PSF is further constrained smoothly by a total variation. As for the SRF based 2D convolution filter, it is activated as a positive tensor via the Softplus function. To optimize the DiriNet, a cost function consisting of the above variation regularization and the mean square error (MSE) between both sides of Eq. (3) is established.

Overall, the contributions of this paper are as follows.

- 1) A neural network is exclusively designed to estimate PSF and SRF, which simulates linear imaging procedures of both HSI and MSI.
- 2) The stick-breaking process is used, for the first time, to describe the PSF with Dirichlet distribution.
- 3) Experimental results show that the proposed net can better help a non-blind fusion method based on degradation functions to achieve blind fusion for real HSI and MSI.

The reminder of the paper is arranged as follows. The proposed DiriNet is illustrated in Section 2 with experimental analysis being presented in Section 3. Finally, a summary is given in Section 4.

2 Proposed DiriNet

For solving the ill-posed problem of Eq. (3), DiriNet shown in Fig. 1 is proposed to learn the SRF and PSF. The keys of DiriNet contain the spatial and spectral degradation process.

Under the framework of deep learning, the spectral degeneration that multiplies an image tensor with a spectrum matrix in $\mathbb{R}^{b \times B}$ along the spectral mode, can be easily implemented by the 2D convolution of the image and a $1 \times 1 \times B \times b$ spectrum filter [21]. Given that the SRF is positive, the spectrum filter is activated with the Softplus function. Based on these, the spectral degradation is constructed by the cascade of an activation function and a 2D convolution layer, highlighted as the dotted box above in Fig. 2.

As Eq. (1) expressed, the spatial blurring is always modeled as a band-wise convolution followed by the downsampling operation. The

depthwise separable convolution proposed by Chollet [19] can just accomplish such band-wise blurring by convoluting the image with a $r \times r \times b \times 1$ filter, where the filter is an along-channel tiling of the PSF. To ensure a proper network convergence, the PSF-based filter should be constrained with some priori. For the PSF, a common prior assumption is that $\phi_i \in (0, 1)$ and $\sum_{i=1}^{r^2} \phi_i = 1$, where ϕ_i denotes the i th element of Φ [11]. Such priori is successfully incorporated via the Dirichlet distribution in [4]. Similarly, the PSF is encouraged to follow a Dirichlet distribution here. A random sequence following the Dirichlet distribution could be produced via the stick-breaking process [20]. As shown in Fig. 2, a stick of a unit length is folded into r^2 segments during the stick-breaking process. Obviously, the length sequence $\{s_1, s_2, \dots, s_{r^2}\}$ of these r^2 segments satisfy that $\sum_{i=1}^{r^2} s_i = 1$ and $s_i \in (0, 1)$. Math-

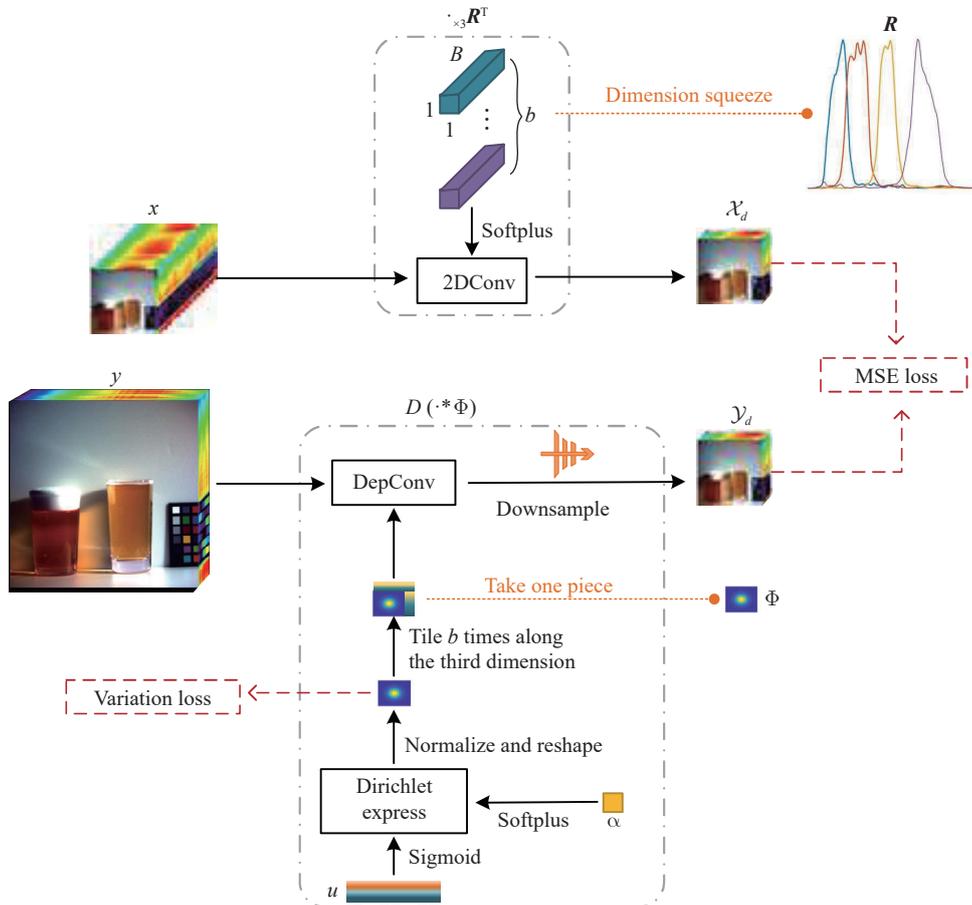


Fig. 1 Flowchart of DiriNet

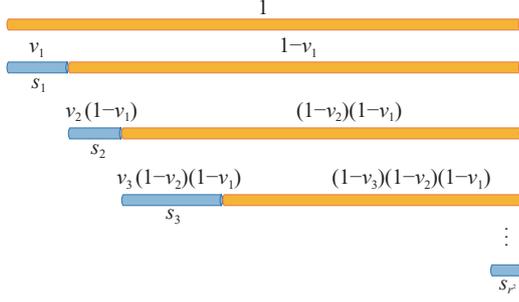


Fig. 2 Stick-breaking process

ematically, the stick-breaking procedure means to assign lengths for r^2 segments as follows

$$s_i = \begin{cases} v_i, & i = 1 \\ v_i \prod_{k=1}^{i-1} (1 - v_k), & \text{others} \end{cases} \quad (4)$$

where $v \sim \text{Beta}(1, \alpha)$ with $\alpha > 0$ being a parameter of the Beta distribution. According to the closed-form cumulative distribution function (CDF) of $\text{Beta}(1, \alpha)$, there is

$$v = 1 - u^{\frac{1}{\alpha}} \quad (5)$$

where $u \sim \text{Uniform}(0, 1)$. In the proposed Dirinet, both u and α are treated as hyper-parameters which can be learned through training. Following the prior of α and u , they are activated with the Softplus and sigmoid functions, respectively. It is hard to guarantee u learned from data sets to be randomly uniform, so we impose an extra sum-to-one constraint on PSF by a normalized operation $\phi_i = s_i / \sum_{i=1}^{r^2} s_i$.

To further limit the solution space of the linear problem in Eq. (3), PSF is smoothly constrained with a total-variation loss l_v .

$$l_v = \|\nabla_x \Phi\|_1 + \|\nabla_y \Phi\|_1 \quad (6)$$

where ∇_x and ∇_y express the horizontal and vertical gradient operators, respectively.

Following Eq. (3), the spectral degradation of HSI \mathcal{X}_d should be equal to the spatial degradation of MSI \mathcal{Y}_d . Hence, an MSE loss is constructed as

$$l_m = \frac{1}{mnb} \|\mathcal{X}_d - \mathcal{Y}_d\|_F^2 \quad (7)$$

Combining the two losses using a proper regularization parameter λ , the cost function to update Dirinet is obtained as

$$l = l_m + \lambda l_v \quad (8)$$

Once the Dirinet update is over, its weights that contain a $1 \times 1 \times B \times b$ spectrum filter, a sequence u , and a positive number α , are obtained. Then, SRF is a dimension-squeeze version of the spectrum filter, and PSF could be calculated by u and α .

3 Experimental Analysis

For a comparative study, the state-of-the-art HySure method [11] is considered. Each SRF curve is band-limited, so a manual band matching is applied in HySure. To test the effect of the band matching, HySure and Dirinet with and without band matching are compared in this paper. For convenience, the two methods with band matching are named as HySureBM and DirinetBM, respectively. The names are stayed in cases of no band matching. To be fair, HySure and HySureBM are traversed by the training data sets to generate series of responses for an average result.

In addition, the estimated and original responses are applied in an outstanding HMI fusion method which is rooted on the coupled nonnegative matrix factorization (CNMF) [1] to study the application effect of all considered methods.

3.1 Experimental Configuration

3.1.1 Experimental Data

To assess the estimation effect of the proposed Dirinet in different-pattern PSFs, different sampling intervals, and different band overlaps, four data sets are constructed by degradating the CAVE [22] and Chikusei datasets with different PSFs and SRFs. The CAVE dataset that contains 32 images of size $512 \times 512 \times 31$, is captured in the wavelength range of 400 – 700 nm with an interval of 10 nm. The Chikusei dataset collects spectral reflection values of 128 bands in the wavelength range of 363 – 1018 nm and 2517×2335 pixels in the sampling interval of 2.5 m, which is cut to 25 subimages of size

$512 \times 512 \times 128$ for the experimental analysis. To simulate situations of different-pattern PSFs and different sampling intervals, a zero-mean Gaussian PSF of standard deviation 2 and size 16×16 as well as an average PSF of size 32×32 is synthesized. Subsequently, the CAVE and Chikusei are spatially degenerated by the two PSFs to their low-resolution counterparts. In addition, the Nikon D700 and IKONOS-like SRFs are used to perform spectral degenerations on the CAVE and Chikusei, respectively. Specifically, the former simulates a case of high band overlapping, while the later emulates ones of low band overlapping. This can be seen from Fig. 3, where the center wavelengths of hyperspectral images are plotted with blue bars, while each column of SRFs is drawn as a color curve. Then, four sets of HMI pairs can be obtained, which are listed in Tab. 1. 20 and 16 images are randomly

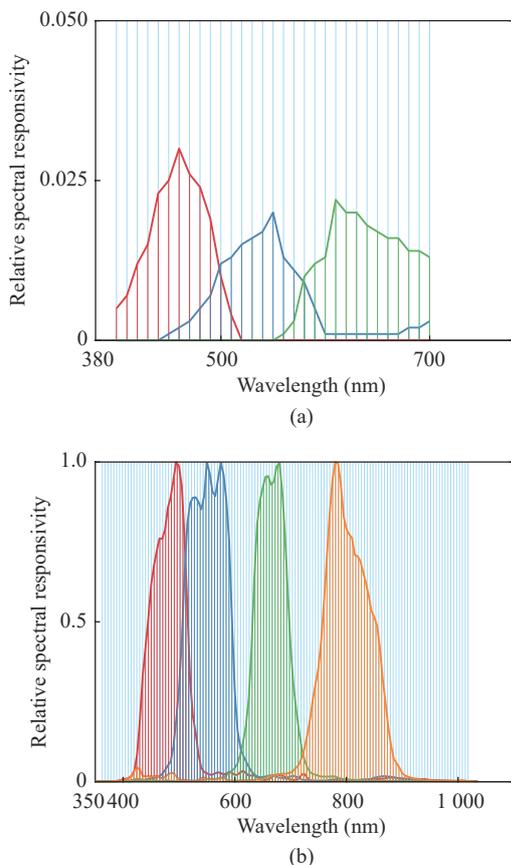


Fig. 3 Center wavelengths and SRFs: (a) Nikon SRF covered in the wavelength range of CAVE; (b) IKONOS SRF covered in the wavelength range of Chikusei

Tab. 1 Details for experimental data

Dataset	Ground-truth	PSF/SRF
Cave 1	CAVE	Gau/Nikon
Cave 2		Ave/Nikon
Chikusei 1	Chikusei	Gau/IKONOS
Chikusei 2		Ave/IKONOS

Note: Gau and Ave denote the simulated Gaussian and average PSF, respectively.

selected from Cave 1–2 and Chikusei 1–2 to form four training data sets, respectively, and the remains are used for testing. Response estimations are executed on the training data sets, while performance assessments are accomplished using the testing data sets.

To further demonstrate the practicability of estimation methods, a real dataset taken by the spectroradiometer on the GaoFen-5 satellite is proceeded and named GF-5 in this paper. This dataset captured over the Yellow River estuary, contains an HSI of size $1185 \times 1342 \times 285$ and an MSI of size $3555 \times 4026 \times 4$, where the HSI covers the wavelength range of $390 - 2513$ nm. The HSI is cut into 72 image patches of size 147×147 , while the MSI is broken into 72 patches of size 441×441 . Then, the 72 pairs of HSI and MSI are randomly divided as a training dataset of 46 patches and a testing dataset of 26 patches. For convenience of expression, this real dataset is denoted as GF-5 next.

3.1.2 Parameter Settings

Within 500 iterations, the proposed DiriNet is optimized via Adam method. When training, the initial learning rates that are set as 10^{-1} and 10^{-2} for Cave 1–2 and Chikusei 1–2, respectively, are damped exponentially with a decay step 250 and a decay rate 0.99. For all data sets, $\lambda = 10^{-7}$. To reduce the convergence difficulty, the SRF is pre-trained 1000 times with the PSF blurring being deleted.

3.1.3 Performance Metrics

For a quantitative assessment, six metrics [23] as the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), erreur relative globale adimensionnelle desynthèse (ERGAS), $Q2^n$ [24],

spectral angle mapper (SAM), and spectral information divergence (SID) between the experimental result and its reference are calculated. Higher PSNR, SSIM, and $Q2^n$ as well as lower ERGAS, SAM, and SID mean better performance.

3.2 Estimation Experiments

In Fig. 4, each estimated PSF is reshaped as a sequence and plotted to visually present the estimation effect. Simultaneously, estimated SRFs are illustrated in Fig 5, where curves of different

types depict different columns of SRF. All the horizontal axes in Fig. 4 and Fig. 5 represent the number of the reshaped PSF and SRF sequences, respectively, while the vertical axes express the value of PSF and SRF, respectively. Obviously, whether with or without band matching, the proposed networks are much preferable to the considered methods. More specifically, the estimated results of HySureBM and HySure are comparatively smooth, which could be caused by the

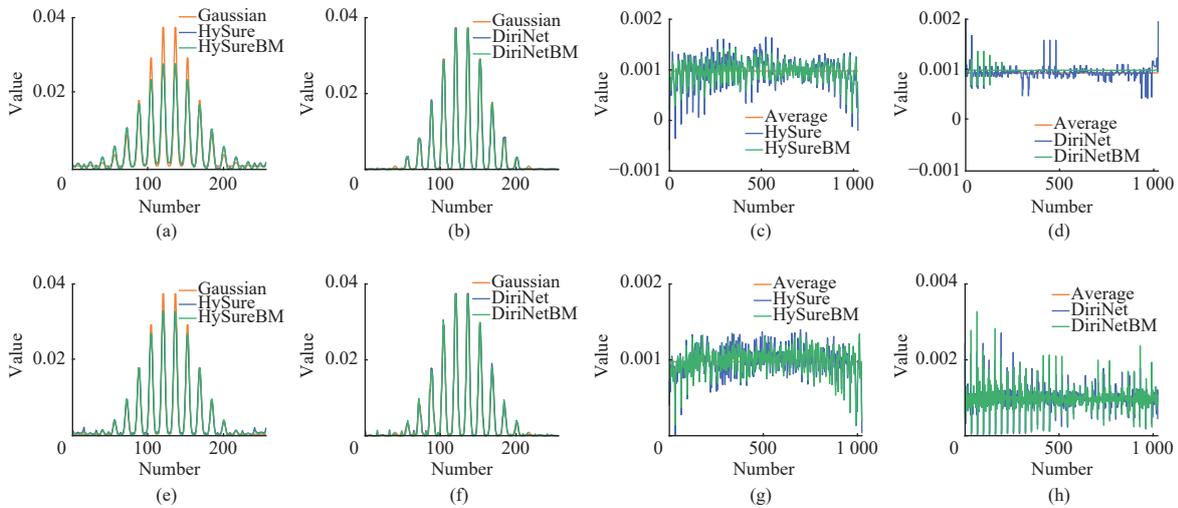


Fig. 4 Illustration of PSF: (a) PSFs estimated by HySure and HySureBM from Cave 1; (b) PSFs estimated by Dirinet and DirinetBM from Cave 1; (c) PSFs estimated by HySure and HySureBM from Cave 2; (d) PSFs estimated by Dirinet and DirinetBM from Cave 2; (e) PSFs estimated by HySure and HySureBM from Chikusei 1; (f) PSFs estimated by Dirinet and DirinetBM from Chikusei 1; (g) PSFs estimated by HySure and HySureBM from Chikusei 2; (h) PSFs estimated by Dirinet and DirinetBM from Chikusei 2

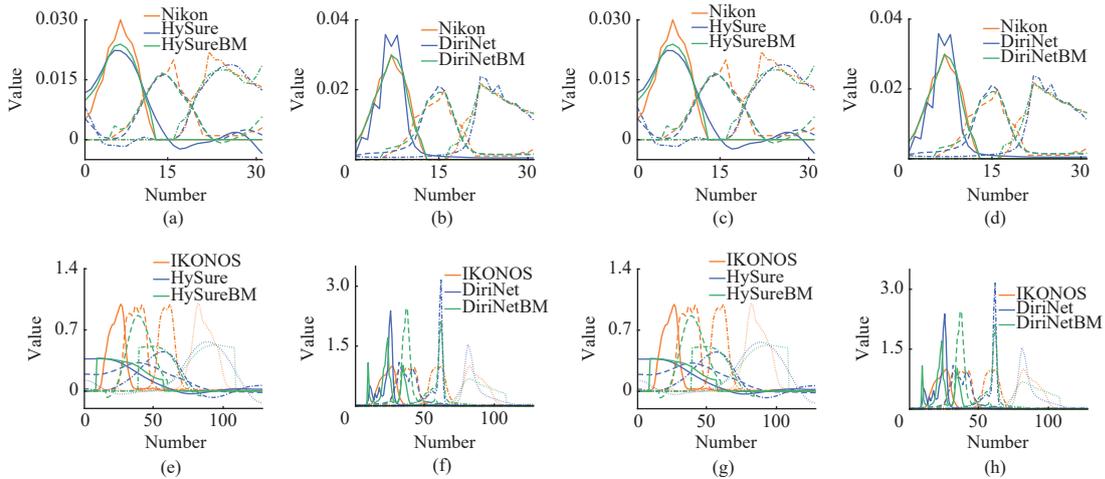


Fig. 5 Illustration of SRF: (a) SRFs estimated by HySure and HySureBM from Cave 1; (b) PSFs estimated by Dirinet and DirinetBM from Cave 1; (c) SRFs estimated by HySure and HySureBM from Cave 2; (d) SRFs estimated by Dirinet and DirinetBM from Cave 2; (e) SRFs estimated by HySure and HySureBM from Chikusei 1; (f) SRFs estimated by Dirinet and DirinetBM from Chikusei 1; (g) SRFs estimated by HySure and HySureBM from Chikusei 2; (h) SRFs estimated by Dirinet and DirinetBM from Chikusei 2

co-smoothness regularization on the PSF and SRF. Furthermore, some negative values appear in the PSF and SRF obtained by HySureBM and HySure, because no threshold constraint is employed.

Furthermore, the performance measure between degenerated images produced with the estimated and original function is shown in Tab. 2. Much better metrics are obtained by

DiriNet than HySure, especially for data sets with limited SRF overlaps (Chikusei 1-2). Such superior performance could be the result of the good learning ability, moderate smoothness regularization, and positive constraint employed in DiriNet. The comparison between experiments under the Gaussian and average PSF shows that both HySure and HySureBM get a great performance reduction as the blurring degree. The per-

Tab. 2 Metrics between degradations obtained by estimated and original responses

Function	Method	PSNR	SSIM	ERGAS	Q2 ^a	SAM	SID
Cave 1							
SRF	HySure	61.25	0.9987	0.3226	0.9526	1.4900	1.8092
	HySureBM	64.14	0.9994	0.2594	0.9796	1.3641	0.7326
	DiriNet	65.51	0.9992	0.2411	0.9799	2.4959	1.0165
	DiriNetBM	83.46	1.0000	0.0472	0.9989	0.2727	0.0255
PSF	HySure	53.09	0.9995	0.1640	0.9997	0.3992	1.6409
	HySureBM	54.48	0.9996	0.1408	0.9998	0.3543	1.2371
	DiriNet	72.62	1.0000	0.0180	1.0000	0.0801	0.0248
	DiriNetBM	75.64	1.0000	0.0126	1.0000	0.0533	0.0131
Cave 2							
SRF	HySure	54.18	0.9950	0.3880	0.8813	4.6430	8.3311
	HySureBM	57.95	0.9976	0.3054	0.9252	2.3564	3.2015
	DiriNet	65.56	0.9992	0.1195	0.9793	2.3803	0.9677
	DiriNetBM	81.66	1.0000	0.0249	0.9985	0.2780	0.0278
PSF	HySure	50.35	0.9992	0.1117	0.9994	0.4235	2.5793
	HySureBM	55.36	0.9997	0.0628	0.9998	0.2398	0.8205
	DiriNet	71.38	1.0000	0.0105	1.0000	0.0487	0.0263
	DiriNetBM	89.73	1.0000	0.0012	1.0000	0.0059	0.0006
Chikusei 1							
SRF	HySure	40.99	0.9958	0.3873	0.9918	1.1325	9.7652
	HySureBM	40.40	0.9937	0.4306	0.9879	1.2977	17.9865
	DiriNet	55.28	0.9998	0.0676	0.9994	0.1784	0.4724
	DiriNetBM	53.85	0.9993	0.1102	0.9980	0.2979	1.3603
PSF	HySure	74.19	0.9999	0.1129	0.9995	0.2749	0.1635
	HySureBM	77.33	1.0000	0.0791	0.9998	0.1849	0.0650
	DiriNet	86.62	1.0000	0.0275	1.0000	0.0748	0.0335
	DiriNetBM	86.09	1.0000	0.0292	1.0000	0.0787	0.0370
Chikusei 2							
SRF	HySure	36.29	0.9864	0.3176	0.9843	2.2391	38.9938
	HySureBM	38.36	0.9856	0.3418	0.9738	2.3097	43.9317
	DiriNet	51.54	0.9994	0.0572	0.9983	0.2783	1.1288
	DiriNetBM	52.47	0.9992	0.0592	0.9980	0.3251	1.4748
PSF	HySure	78.42	1.0000	0.0357	0.9997	0.1578	0.0377
	HySureBM	79.44	1.0000	0.0318	0.9998	0.1391	0.0313
	DiriNet	89.61	1.0000	0.0105	1.0000	0.0398	0.0027
	DiriNetBM	89.22	1.0000	0.0109	1.0000	0.0404	0.0028

formance degradation might be caused by the blurring operation before the response estimation. In addition, band matching indicates distinct benefit for the response estimation in Cave 1–2 but slight damage in Chikusei 1–2, probably because the tailing of IKONOS-like response is nonzero. The tailing of an actual SRF is usually nonzero; thus, band matching is probably a dispensable option. Comprehensively, the proposed DiriNet is an excellent approach to estimate the degradation functions, especially for data sets

with limited SRF overlaps.

3.3 Fusion Experiments on Synthesized Data

The original purpose of this work is to provide accessible PSF and SRF for HMI fusion methods. Hence, all degraded HSIs and MSIs are fused by the CNMF fed with estimated and original degeneration functions successively. Fusion results for one pair of HSI and MSI in the testing data of Cave 1 and Chikusei 1 are shown in Fig. 6 and Fig. 7, respectively. Specifically, the ground truth and the degenerated HSI are shown

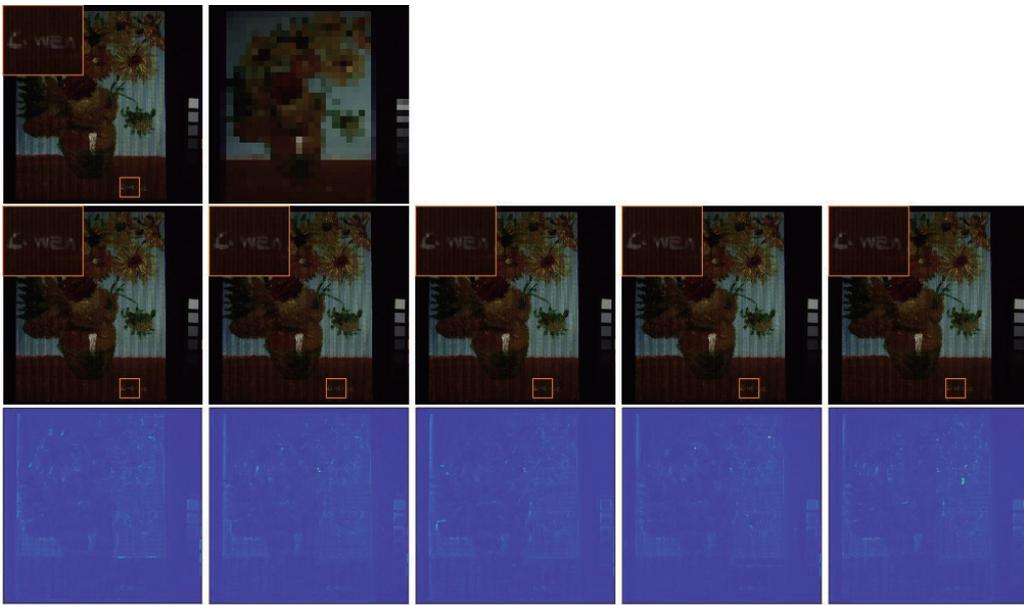


Fig. 6 Color-composite images of selected bands [25, 15, 3] in Cave 1

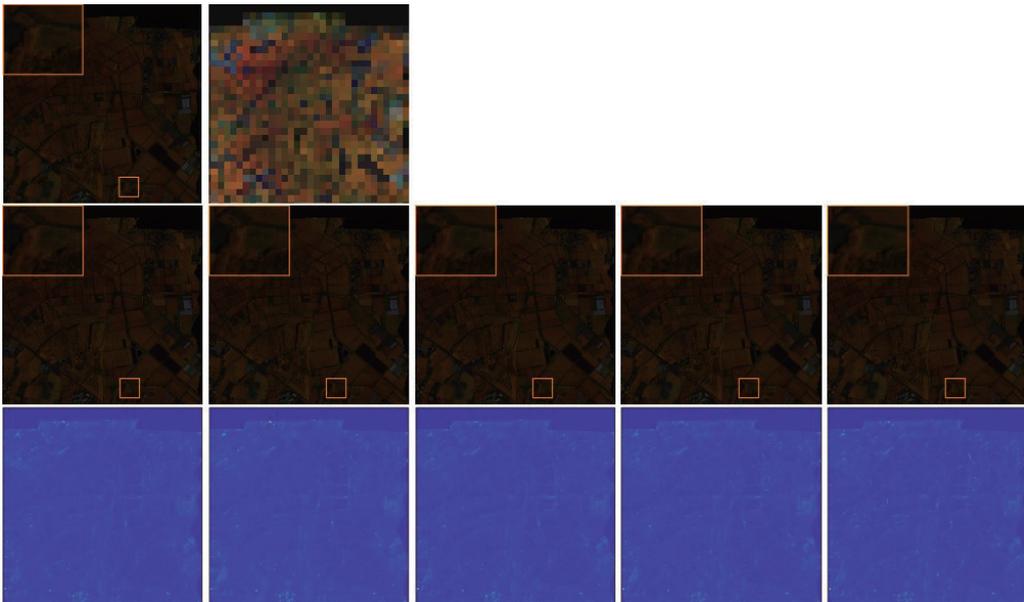


Fig. 7 Color-composite images of selected bands [90, 70, 20] in Chikusei 1

in the first row from left to right, respectively. Fusion results under estimated functions by HySure, HySureBM, DiriNet, and DiriNetBM, and original functions are presented in the second row in turn. The third row correspondingly gives mean error images between fusion images and the ground truth. In addition, the metrics between the fusion result and the ground truth are calculated in Tab. 3 and Tab. 4. The fusion result indicated that band matching brings better improvement when applied in Cave 1–2 than Chikusei 1–2. In addition, the fusion effect under the DiriNet estimation is comparable to that under the original responses. Therefore, the proposed DiriNet, an estimation network for PSF and SRF, can effectively serve HMI fusion techniques.

3.4 Fusion Experiments on Real Data

The practicability of the proposed DiriNet is still a mystery. Then, all degeneration functions esti-

imated by considered methods are integrated into the CNMF one by one for the fusion of GF-5 here. The fusion results are shown in Fig. 8, where the first row, from left to right, shows the HSI and MSI in GF-5, respectively; and the second row individually presents fusion results under estimated functions by HySure, HySureBM, DiriNet, and DiriNetBM. As it can be seen from fusion images in Fig. 8, DiriNetBM based CNMF achieves the best fusion effect. Moreover, the fusion result based on DiriNet is clearer than that based on both HySure and HySureBM. Namely, the proposed estimation network is proved to be practical.

4 Conclusion

A Dirichlet estimation network was proposed to accurately provide spatial and spectral degradation functions for HMI fusion. Pure data-driven learning was difficult to work around this ill-

Tab. 3 Fusion metrics on Cave 1–2 for different estimation methods

Data	Response	PSNR	SSIM	ERGAS	Q2 ⁿ	SAM	SID
Cave 1	HySure	41.64	0.9818	0.8312	0.9303	6.5101	36.2442
	HySureBM	41.74	0.9816	0.7945	0.9288	6.5166	35.8018
	DiriNet	41.81	0.9815	0.8426	0.9269	6.6587	36.6442
	DiriNetBM	42.22	0.9829	0.8210	0.9349	6.2072	34.5704
	Nikon	42.08	0.9827	0.8243	0.9363	6.2972	34.5135
Cave 2	HySure	39.67	0.9724	1.0027	0.8882	11.4057	159.9813
	HySureBM	41.93	0.9800	0.3771	0.9234	7.3714	42.1367
	DiriNet	42.35	0.9796	0.3744	0.9179	7.6670	42.3063
	DiriNetBM	42.68	0.9813	0.3557	0.9270	7.1715	39.1826
	Nikon	42.62	0.9812	0.3590	0.9284	7.0334	38.8715

Tab. 4 Fusion metrics on Chikusei 1–2 for different estimation methods

Data	Response	PSNR	SSIM	ERGAS	Q2 ⁿ	SAM	SID
Chikusei 1	HySure	62.08	0.9987	0.8036	0.8861	6.4752	2.5291
	HySureBM	62.41	0.9986	1.0912	0.8880	6.5270	2.7833
	DiriNet	63.41	0.9988	0.7654	0.8861	6.4942	2.3654
	DiriNetBM	63.66	0.9988	0.7568	0.8872	6.4631	2.2461
	IKONOS	64.01	0.9988	0.7581	0.8866	6.4494	2.4408
Chikusei 2	HySure	62.85	0.9987	0.3830	0.8996	6.5281	2.6552
	HySureBM	62.66	0.9987	0.3854	0.8967	6.5203	2.6694
	DiriNet	63.67	0.9988	0.3880	0.8999	6.3565	2.3024
	DiriNetBM	63.71	0.9988	0.3892	0.9033	6.3760	2.3211
	IKONOS	63.97	0.9989	0.3807	0.9016	6.3638	2.2592

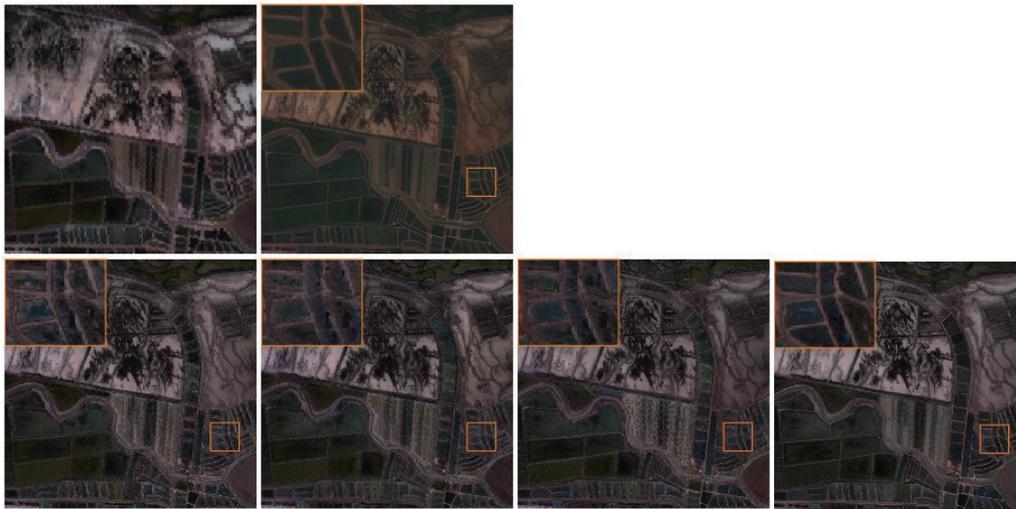


Fig. 8 Color-composite images of selected bands [50, 35, 25] in GF-5

posed estimation problem, so some effective constraints were applied. More specifically, the PSF was constrained with Dirichlet distribution and spatial smoothness, while the SRF was regularized with non-negative property. To describe the Dirichlet distributed PSF, the stick-breaking procedure is introduced for the first time. Experiments revealed that the proposed network could not only better learn the degradation functions but also effectively serve the fusion technique.

References:

- [1] N. Yokoya, S. Member, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528-537, 2012.
- [2] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *IEEE European Conference on Computer Vision*, Zurich, Switzerland, pp. 63-78, 2014.
- [3] N. Liu, L. Li, W. Li, R. Tao, and J. E. Fowler, "Hyperspectral restoration and fusion with multispectral imagery via low-rank tensor-approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 9, pp. 7817-7830, 2021.
- [4] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse dirichlet-net for hyperspectral image super-resolution," in *2018 IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, US, pp. 2511-2520, 2018.
- [5] F. Zhou, R. Hang, Q. Liu, and X. Yuan, "Pyramid fully convolutional network for hyperspectral and multispectral image fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 5, pp. 1549-1558, 2019.
- [6] Z. Wang, B. Chen, R. Lu, H. Zhang, and P. K. Varshney, "Fusionnet: An unsupervised convolutional variational network for hyperspectral and multispectral image fusion," *IEEE Transactions on Image Processing*, vol. 29, pp. 7565-7577, 2020.
- [7] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5345-5355, 2018.
- [8] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 3, pp. 1124-1135, 2020.
- [9] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu, "Multispectral and hyperspectral image fusion by MS/HS fusion net," in *IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, US, pp. 1585-1594, 2020.
- [10] P. Carvalho, A. Santos, and P. Martins, "Recovering imaging device sensitivities: A data-driven approach," in *2004 IEEE International Conference on Image Processing*, Singapore, pp. 2411-2414, 2004.
- [11] M. Simões, J. Bioucas-Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regu-

- larization,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3373-3388, 2015.
- [12] N. Yokoya, C. Grohnfeldt, and J. Chanussot, “Hyperspectral and multispectral data fusion: A comparative review of the recent literature,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29-56, 2017.
- [13] Y. Qu, H. Qi, and C. Kwan, “Recovering device sensitivities with quadratic programming,” in *Imaging Science and Technology and Society for Information Display 6th Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, AZ, US, pp. 90-95, 1998.
- [14] J. Yao, D. F. Hong, J. Chanussot, D. Meng, and Z. Xu, “Cross-attention in coupled unmixing nets for unsupervised hyperspectral super-resolution,” in *IEEE European Conference on Computer Vision*, Glasgow, United Kingdom, pp. 1-17, 2020.
- [15] W. Wang, Y. Huang, and X. Ding, “Self-regression learning for blind hyperspectral image fusion without label,” *arXiv e-prints*, arXiv: 2103.16806, 2021.
- [16] S. Xu, O. Amira, J. Liu, C. X. Zhang, J. Zhang, and G. Li, “HAM-MFN: Hyperspectral and multispectral image multiscale fusion network with rapid loss,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4618-4628, 2020.
- [17] W. Li, G. Wu, and Q. Du, “Transferred deep learning for anomaly detection in hyperspectral imagery,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 597-601, 2017.
- [18] M. Zhang, W. Li, and Q. Du, “Diverse region-based CNN for hyperspectral image classification,” *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2623-2634, 2018.
- [19] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, US, pp. 1800-1807, 2017.
- [20] J. Sethuraman, “A constructive definition of the dirichlet prior,” *Statistica Sinica*, vol. 4, no. 2, pp. 639-650, 1994.
- [21] Q. Xie, M. Zhou, Q. Zhao, Z. Xu, and M. D., “MHF-Net: An interpretable deep network for multispectral and hyperspectral image fusion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, pp. 1457-1473, 2022.
- [22] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, “Generalized assorted pixel camera: Post-capture control of resolution, dynamic range and spectrum,” *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 639-650, 2010.
- [23] T. Hu, W. Li, N. Liu, R. Tao, F. Zhang, and P. Scheunders, “Hyperspectral image restoration using adaptive anisotropy total variation and nuclear norms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 2, pp. 1516-1533, 2021.
- [24] A. Garzelli and F. Nencini, “Hypercomplex quality assessment of multi/hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 662-665, 2009.



Ting Hu received the B.S. and M.S. degrees from Department of Electronic Information Engineering, Nanchang University, Nanchang, China, in 2014 and 2017, respectively, and the Ph.D. degree in School of Information and Electronics, Beijing Institute of Technology, Beijing, China, in 2022. She is currently a Lecturer with the Faculty of Information Technology, Beijing University of Technology. Her scientific interests include image super-resolution, fusion, destriping and denoising.



Siyuan Cheng received the B.S. and M.S. degrees from Department of Electronic Information Engineering, Nanchang University, Nanchang, China, in 2014 and 2016, respectively. He is currently a Senior Engineer with the Space Star Technology Co., Ltd, Beijing, China. His scientific interests include image and signal processing.



Chang Liu received the B.S. and M.S. degrees from Beijing Jiaotong University, Beijing, China, in 2013 and 2016, respectively, and the Ph.D. degree in information and communication engineering from Beijing University of Posts and Telecommunications, Beijing, in 2022. She is now a Lecturer with the Faculty of Information Technology, Beijing University of Technology. Her research interests focus on novel simulation techniques, fundamental theories on wireless communications, including multiuser detection, MIMO channel models, and channel estimation.