

Direct spatio-spectral datacube reconstruction from raw data using a spatially adaptive spatio-spectral basis

Yusuke Monno^a, Masayuki Tanaka^a, Masatoshi Okutomi^a

^aTokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo, JAPAN;

ABSTRACT

Spectral reflectance is an inherent property of objects that is useful for many computer vision tasks. The spectral reflectance of a scene can be described as a spatio-spectral (SS) datacube, in which each value represents the reflectance at a spatial location and a wavelength. In this paper, we propose a novel method that reconstructs the SS datacube from raw data obtained by an image sensor equipped with a multispectral filter array. In our proposed method, we describe the SS datacube as a linear combination of spatially adaptive SS basis vectors. In a previous method, spatially invariant SS basis vectors are used for describing the SS datacube. In contrast, we adaptively generate the SS basis vectors for each spatial location. Then, we reconstruct the SS datacube by estimating the linear coefficients of the spatially adaptive SS basis vectors from the raw data. Experimental results demonstrate that our proposed method can accurately reconstruct the SS datacube compared with the method using spatially invariant SS basis vectors.

Keywords: Multispectral imaging, multispectral filter array, spectral reflectance, spatio-spectral datacube, spatio-spectral reconstruction.

1. INTRODUCTION

Spectral reflectance is an inherent property of objects that is useful for many computer vision tasks such as object recognition, tracking, segmentation, relighting, etc. For example, the full description of the spectral reflectance enables fidelity color reproduction under an arbitrary illumination. The spectral reflectance of a scene can be described as a spatio-spectral (SS) datacube, in which each value represents the reflectance at a spatial location and a wavelength. The illustration of the SS datacube is shown in Fig. 2 (a). A number of methods have been proposed for SS datacube reconstruction from an RGB image.^{1,2} However, the RGB image only measures three spectral bands among continuous spectral reflectance, therefore, the accuracy of the reconstruction is very limited. To reconstruct the SS datacube more accurately, the use of a multispectral image with more than three spectral bands has received increasing attention.³⁻⁹

In the past decades, a wide variety of systems have been developed for capturing the multispectral image.¹⁰⁻¹⁴ Recently, a one-shot multispectral imaging system that uses a single image sensor equipped with a multispectral filter array (MSFA) has been proposed.¹⁵⁻¹⁸ The illustration of the one-shot system is shown in Fig. 1 (a). The one-shot system is very attractive for industrial use because it enables low-cost and simple multispectral image acquisition. The particular property of the one-shot system is that only one spectral measurement is acquired at each pixel location since incoming light into the image sensor is spatially and spectrally sampled through the MSFA. The such subsampled data is called raw data. There are two approaches that reconstruct the SS datacube from the raw data: (i) sequential reconstruction, and (ii) direct reconstruction. The sequential reconstruction first interpolates the raw data by a demosaicking process to acquire the multispectral image.¹⁶⁻¹⁸ Then, spectral reflectance is independently reconstructed at each pixel location from the interpolated multispectral image.³⁻⁸ In contrast, Parmar et al. proposed the direct SS reconstruction that jointly reconstructs the SS datacube from the raw data.⁹

Further author information: (Send correspondence to Y.Monno)

Y.Monno: E-mail: ymonno@ok.ctrl.titech.ac.jp, Telephone: +81-3-5734-3499

M.Tanaka: E-mail: mtanaka@ctrl.titech.ac.jp, Telephone: +81-3-5734-3270

M.Okutomi: E-mail: mxo@ctrl.titech.ac.jp, Telephone: +81-3-5734-3472

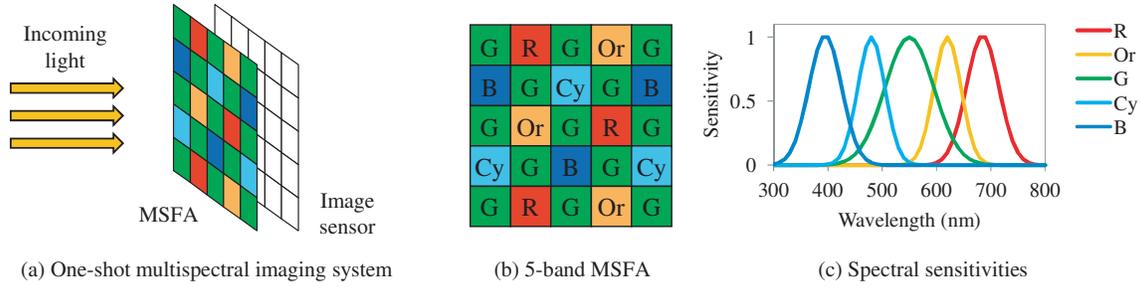


Figure 1. (a) The illustration of the one-shot multispectral imaging system, (b) 5-band MSFA, and (c) spectral sensitivities of each spectral band.

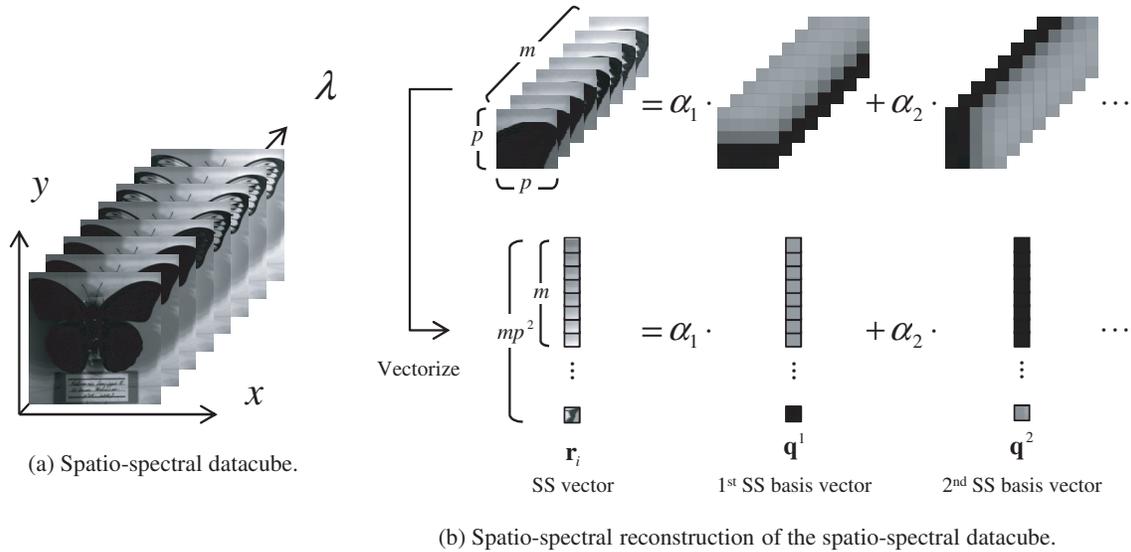


Figure 2. (a) The illustration of the spatio-spectral datacube. (b) The spatio-spectral reconstruction of the spatio-spectral datacube. The spatio-spectral datacube is described as a linear combination of spatio-spectral basis vectors.

In this paper, we propose a novel direct SS reconstruction of the SS datacube using a spatially adaptive SS basis. In the SS reconstruction, the SS datacube is described as a linear combination of SS basis vectors. In the previous method,⁹ spatially invariant SS basis vectors are used for describing the SS datacube. In contrast, we adaptively generate the SS basis vectors for each spatial location. Then, we reconstruct the SS datacube by estimating the linear coefficients of the spatially adaptive SS basis vectors from the raw data. Experimental results demonstrate that our proposed method can accurately reconstruct the SS datacube compared with the method using spatially invariant SS basis vectors.

2. SPATIO-SPECTRAL RECONSTRUCTION

We first describe the SS reconstruction of the SS datacube from the raw data.⁹ In the SS reconstruction, the SS datacube is described as a linear combination of SS basis vectors as illustrated in Fig. 2 (b). Now consider the local SS datacube centered at a pixel location i with p rows and columns, and m bands, the local SS datacube is represented as:

$$\mathbf{r}_i = \mathbf{Q}\boldsymbol{\alpha}_i, \tag{1}$$

where $\mathbf{r}_i \in \mathbb{R}^{mp^2}$ is the SS vector that formed by stacking the spectral reflectances in the local SS datacube, $\mathbf{Q} = [\mathbf{q}^1, \dots, \mathbf{q}^k, \dots, \mathbf{q}^K]$ represents the set of SS basis vectors where each column $\mathbf{q}^k \in \mathbb{R}^{mp^2}$ is the k -th SS basis vector, $\boldsymbol{\alpha}_i = [\alpha_i^1, \dots, \alpha_i^k, \dots, \alpha_i^K]^T$ is the linear coefficients of the SS basis vectors, and K is the number

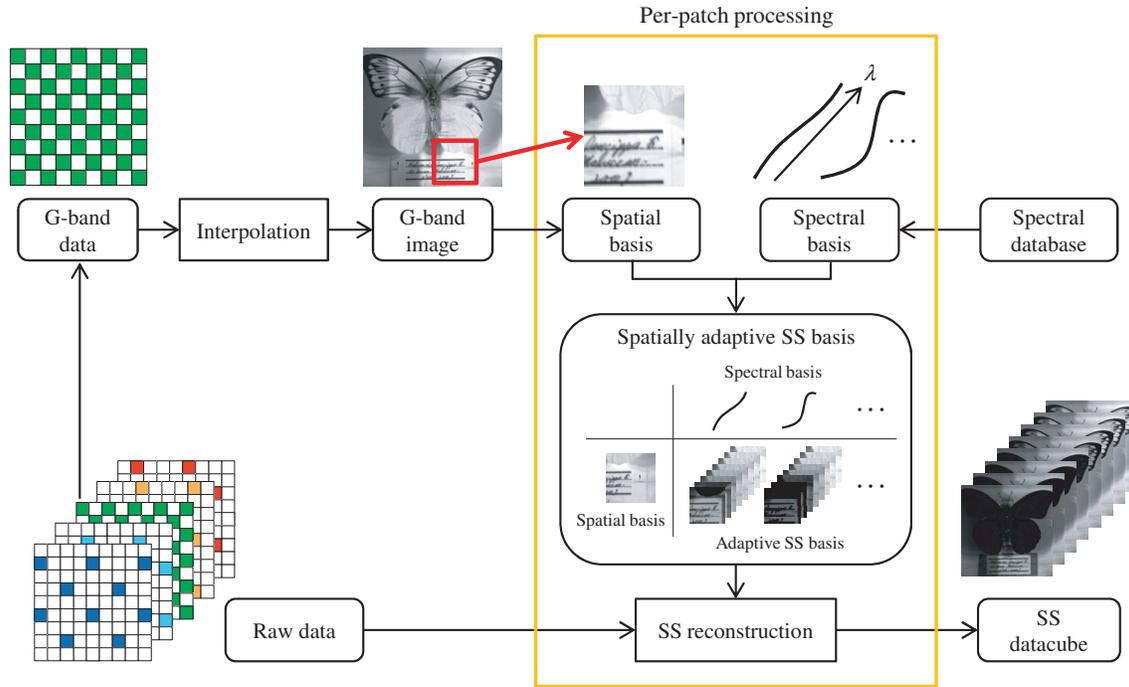


Figure 3. The schematic block diagram of our proposed method.

of the SS basis vectors used for the reconstruction. By using Eq. (1), the raw data in the $p \times p$ patch centered at the pixel location i is modeled as:

$$\mathbf{g}_i = \mathbf{H}_i \mathbf{r}_i = \mathbf{H}_i \mathbf{Q} \boldsymbol{\alpha}_i, \quad (2)$$

where $\mathbf{g}_i \in \mathbb{R}^{p^2}$ is the vector that formed by stacking the raw data in the patch, and $\mathbf{H}_i \in \mathbb{R}^{p^2 \times mp^2}$ is the system matrix that is given by the spectral sensitivities of the multispectral filters and the spectral distribution of the illumination. From the raw data, the linear coefficients of the SS basis vectors are estimated as:

$$\hat{\boldsymbol{\alpha}}_i = \arg \min_{\boldsymbol{\alpha}_i} \left[\|\mathbf{g}_i - \mathbf{H}_i \mathbf{Q} \boldsymbol{\alpha}_i\|_2^2 + \sigma^2 \|\mathbf{D}_\lambda \mathbf{Q} \boldsymbol{\alpha}_i\|_2^2 \right], \quad (3)$$

where $\mathbf{D}_\lambda \in \mathbb{R}^{mp^2 \times mp^2}$ is a second-order derivative operator, and σ is a smoothing parameter. In Eq. (3), the first term is a data term and the second term is a regularization term. While Parmar et al. constrain the sparseness of the linear coefficients,⁹ we constrain the second-order derivatives of the spectral reflectances as in [12]. From Eq. (1) and (2), the SS datacube is reconstructed as: $\hat{\mathbf{r}}_i = \mathbf{Q} \hat{\boldsymbol{\alpha}}_i$.

3. PROPOSED METHOD

In this paper, we aim to reconstruct the SS datacube from the raw data acquired using the 5-band MSFA proposed in [17] and the optimized spectral sensitivities in [20]. Fig. 1 (b) and (c) respectively show the 5-band MSFA and the corresponding spectral sensitivities of each spectral band. We call each spectral band as the R, Or, G, Cy, and B-band from the long wavelength end to the short wavelength end.

In the SS reconstruction, SS basis vectors are typically learned from a spectral image database^{9,19} and the learned SS basis vectors are used for all pixel locations. In contrast, we adaptively generate the SS basis vectors for each spatial location. Fig. 3 shows the schematic block diagram of our proposed method. We assume that the SS basis is “separable”, which means the SS basis can be decomposed into a spatial basis and a spectral basis.¹⁹ In our proposed method, the spatial basis is adaptively generated for each pixel location by interpolating

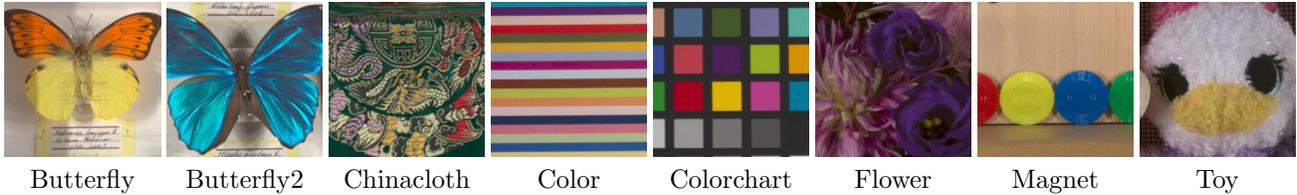


Figure 4. Scenes used in experiments. The sRGB images converted from the captured 31-band images are shown.

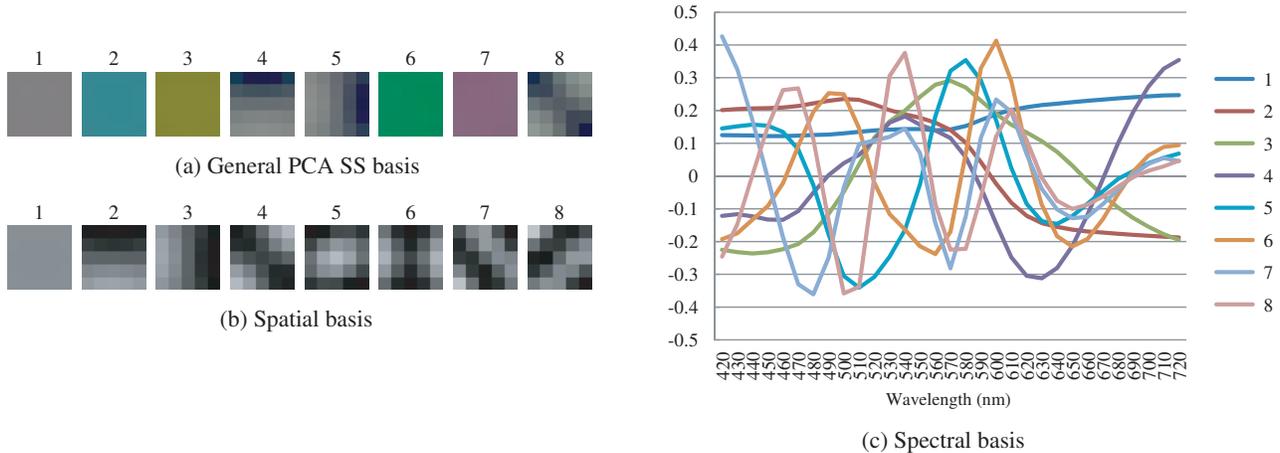


Figure 5. The learned bases.

the G-band data as shown in Fig. 3. There are mainly two advantages of using the G-band data as the spatial basis: (i) it is well known that spectral correlations are very high between spectral images. Since the G-band is addressed at the center of visible wavelengths, it is expected that the interpolated G-band image has correlations with a wide range of spectral images. Therefore, edges or textures of the interpolated G-band image are also likely to appear in the spectral images in the SS datacube. (ii) We can interpolate the G-band data with relatively high accuracy because the G-band data is highly sampled in the raw data. Our proposed method effectively exploits these two advantages.

The adaptive SS basis vectors for the pixel location i is represented as: $\mathbf{Q}_i = [\mathbf{q}_i^1, \dots, \mathbf{q}_i^k, \dots, \mathbf{q}_i^K]$, where $\mathbf{q}_i^k = \mathbf{d}_i \otimes \mathbf{b}_k$, \mathbf{d}_i represents the adaptive spatial basis vector that is formed by stacking the interpolated G-band data in the $p \times p$ patch centered at the pixel location i , and \mathbf{b}_k is the k -th spectral basis vector learned from a spectral database by principal component analysis (PCA). After generating the adaptive SS basis vectors \mathbf{Q}_i , we reconstruct the SS datacube by Eq. (1) and Eq. (3). Finally, the overlaps of the reconstructed SS datacube are simply averaged.

4. EXPERIMENTAL RESULTS

We captured 500×500 31-band images of 16 scenes for ground truth SS datacubes using a monochrome camera with a liquid crystal tunable filter.¹⁰ The 31-band images are acquired at every 10nm from 420nm to 720nm. Some scenes used in experiments are shown in Fig. 4, in which the sRGB images converted from the captured 31-band images are displayed. We used 8 scenes for learning bases and the other 8 scenes for evaluation.

We simulated raw data from the ground truth SS datacubes. We used the 5-band MSFA and optimized spectral sensitivities²⁰ as shown in Fig. 1. To simplify the problem, we assumed noise-free data and white illumination. Then, the SS datacubes are reconstructed by the following four methods: (i) spectral reflectance estimation¹² after BTES demosaicking¹⁶ (BTES), (ii) SS reconstruction using a general PCA SS basis¹⁹ (Gen), (iii) SS reconstruction using a separable PCA SS basis¹⁹ (Sep), and (iv) our proposed method (Pro). The method (i) is the sequential reconstruction, while the methods (ii), (iii), and (iv) are direct SS reconstruction from the

Table 1. The average PSNR of all 31 bands for each scene using the separable-PCA basis with different number of spatial components and spectral components (spatial×spectral).

Scene	1x8	2x4	4x2	8x1
Bell	30.02	31.68	29.67	27.84
Butterfly	30.73	29.64	24.93	19.89
Chinacloth	26.81	29.11	28.85	26.38
Colorchart	35.52	35.05	26.06	20.70
Doll	30.40	31.08	27.31	23.41
Magnet	36.50	33.14	26.40	23.20
Toy	34.48	33.96	29.24	24.35
Wool	31.12	31.57	31.19	23.82
Average	31.95	31.90	27.96	23.70

Table 2. The average PSNR of all 31 bands for each scene.

Scene	BTES	Gen	Sep	Pro
Bell	33.41	33.91	30.02	35.82
Butterfly	34.50	31.30	30.73	38.52
Chinacloth	30.90	32.49	26.81	33.66
Colorchart	38.33	35.68	35.52	39.67
Doll	32.84	33.33	30.40	33.64
Magnet	38.07	35.56	36.50	37.73
Toy	37.09	36.30	34.48	39.05
Wool	34.66	36.15	31.12	37.73
Average	34.97	34.34	31.95	36.98

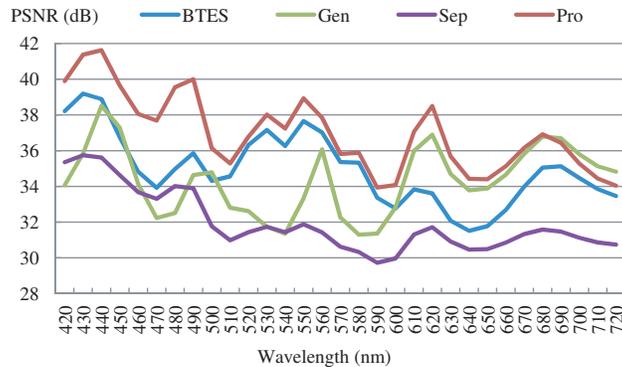


Figure 6. The average PSNR of all 8 scenes for each band.

raw data. The general PCA SS basis is learned from the ground truth SS datacubes (8 scenes for learning bases) by PCA. The sRGB representations of the top 8 components of the general PCA SS basis are shown in Fig. 5 (a). The separable PCA SS basis is composed of the spatial basis and the spectral basis that are separately learned from the ground truth SS datacubes (8 scenes for learning bases) by PCA. Fig. 5 (b) and (c) respectively show the top 8 components of the spatial basis and the spectral basis. The two PCA SS bases are spatially invariant basis. For the methods (ii), (iii), and (iv), we used the patch size $p = 5$, the smoothing parameter $\sigma = 0.012$, and $K = 8$ basis vectors for the reconstruction. For our proposed method, we used the same spectral basis as shown in Fig. 5 (c) and the adaptive Gaussian upsampling¹⁷ for interpolating the G-band data to generate the adaptive spatial basis.

We first investigate the performance of the SS reconstruction using the separable PCA SS basis. Table 1 shows the average PSNR of all 31 bands for each scene using the separable PCA SS basis with different number of spatial components and spectral components (spatial × spectral). The combination of one spatial component and 8 spectral components (1 × 8) represents the best performance among the other combinations, 2 × 4, 4 × 2, and 8 × 1. Therefore, we used the 1 × 8 combination for the following experiments.

We compared the PSNR of the reconstructed SS datacubes by each method. Table 2 shows the average PSNR of all 31 bands for each scene, while Fig. 6 shows the average PSNR of all 8 scenes for each band. These PSNR comparisons show that our proposed method quantitatively outperforms the other methods. Fig. 7 shows the reconstructed spectral reflectances for Butterfly and Chinacloth. Our proposed method can accurately reconstruct spectral reflectances especially in the edge or texture regions.

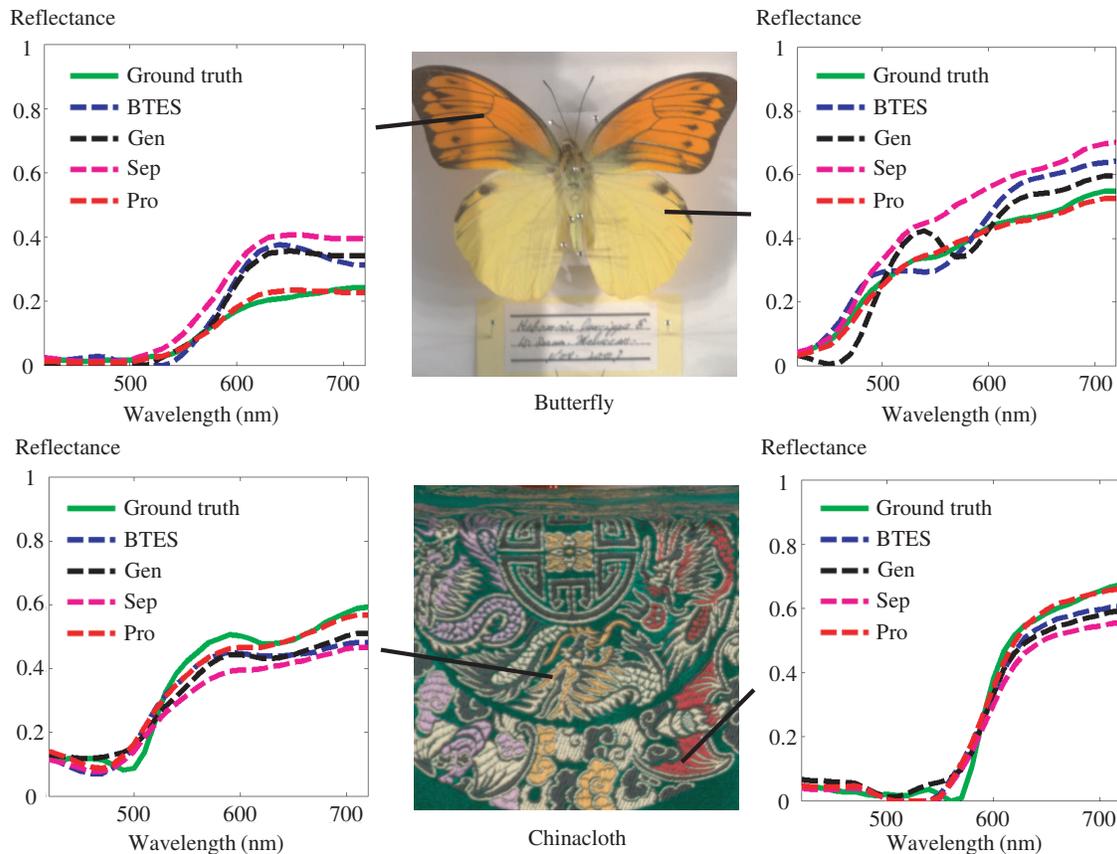


Figure 7. Comparison of the reconstructed spectral reflectances.

Fig. 8 shows sRGB images converted from the ground truth and the reconstructed SS datacubes. The sequential reconstruction (BTES) generates demosaicking artifacts, i.e. zipper artifacts, especially in the butterfly wing and in the edge of the colorchart, while the direct SS reconstruction using spatially invariant basis (Gen and Sep) generates severe color artifacts or blur, especially in the edge of the colorchart. In contrast, our proposed method using spatially adaptive basis can effectively reduce demosaicking and color artifacts and generate the edge-preserved images.

5. CONCLUSION

In this paper, we proposed the novel SS reconstruction of the SS datacube using the spatially adaptive SS basis. The adaptive SS basis is composed of the adaptive spatial basis and the spectral basis. We adaptively generate the spatial basis for each pixel location by interpolating the G-band data that is highly sampled in the raw data. Experimental results demonstrate that our proposed method outperforms the other methods both quantitatively and visually.

REFERENCES

- [1] Cheung, V., Westland, S., Li, C., Hardeberg, J., and Connah, D., "Characterization of trichromatic color cameras by using a new multispectral imaging technique," *J. Opt. Soc. Am. A* **22**(7), 1231–1240 (2005).
- [2] Heikkinen, V., Lenz, R., Jetsu, T., Parkkinen, J., Hauta-Kasari, M., and Jääskeläinen, T., "Evaluation and unification of some methods for estimating reflectance spectra from RGB images," *J. Opt. Soc. Am. A* **25**(10), 2444–2458 (2008).



Bell

Butterfly

Colorchart

Doll



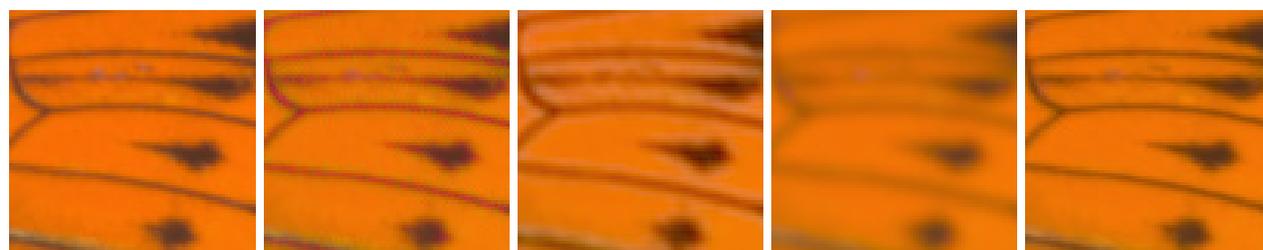
Ground truth

BTES

Gen

Sep

Pro



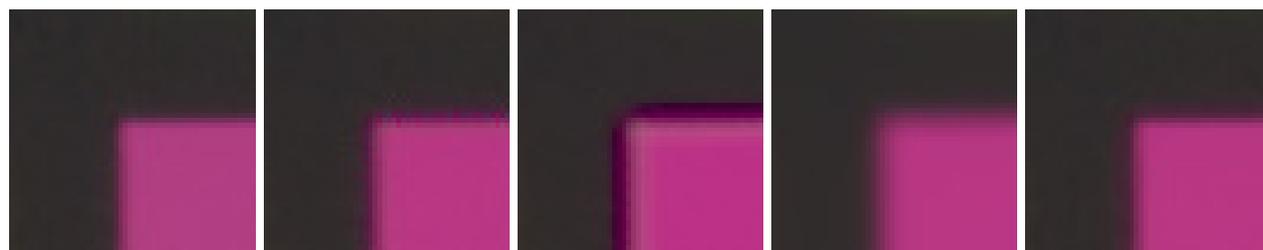
Ground truth

BTES

Gen

Sep

Pro



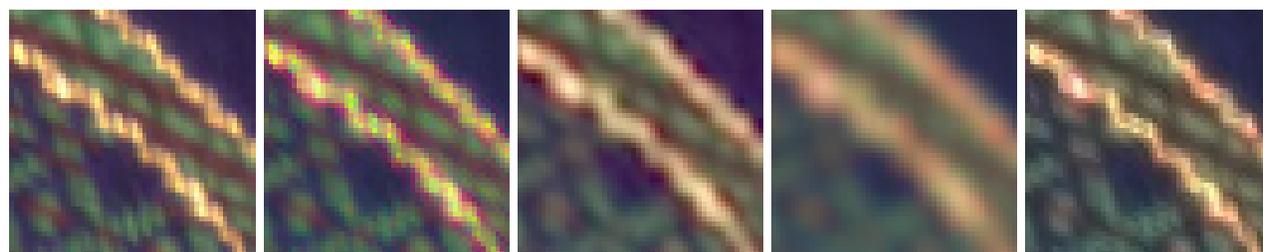
Ground truth

BTES

Gen

Sep

Pro



Ground truth

BTES

Gen

Sep

Pro

Figure 8. Comparison of sRGB images.

- [3] Tominaga, S., "Multichannel vision system for estimating surface and illumination functions," *J. Opt. Soc. Am. A* **13**(11), 2163–2173 (1996).
- [4] Murakami, Y., Obi, T., Yamaguchi, M., Ohyama, N., and Komiya, Y., "Spectral reflectance estimation from multi-band image using color chart," *Optics Communications* **188**, 47–454 (2001).
- [5] Shimano, N., Terai, K., and Hironaga, M., "Recovery of spectral reflectances of objects being imaged by multispectral cameras," *J. Opt. Soc. Am. A* **24**(10), 3211–3219 (2007).
- [6] Lansel, S., Parmar, M., and Wandell, B. A., "Dictionaries for sparse representation and recovery of reflectances," *Proc. of SPIE* **7246**, 72460D–1–72460D–11 (2009).
- [7] Shen, H. L., Cai, P. Q., Shao, S. J., and Xin, J. H., "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation," *Optics Express* **15**(23), 15545–15554 (2007).
- [8] Shimano, N., "Recovery of spectral reflectances of objects being imaged without prior knowledge," *IEEE Trans. on Image Processing* **15**(7), 1848–1856 (2006).
- [9] Parmar, M., Lansel, S., and Wandell, B. A., "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," *Proc. of IEEE Int. Conf. on Image Processing (ICIP)*, 473–476 (2008).
- [10] Hardeberg, J. Y., Schmitt, F., and Brettel, H., "Multispectral color image capture using a liquid crystal tunable filter," *Optical Engineering* **41**, 2532–2548 (2002).
- [11] Yamaguchi, M., Haneishi, H., and Ohyama, N., "Beyond red-green-blue (rgb): Spectrum-based color imaging technology," *Journal of Imaging Science and Technology* **52**(1), 10201–1–10201–15 (2008).
- [12] Park, J., Lee, M., Grossberg, M. D., and Nayar, S. K., "Multispectral imaging using multiplexed illumination," *Proc. of IEEE Int. Conf. on Computer Vision (ICCV)*, 1–8 (2007).
- [13] Han, S., Sato, I., Okabe, T., and Sato, Y., "Fast spectral reflectance recovery using DLP projector," *Proc. of Asian Conf. on Computer Vision (ACCV)* **1**, 318–330 (2010).
- [14] Cao, X., Du, H., x.Tong, Dai, Q., and Lin, S., "A prism-mask system for multispectral video acquisition," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **33**(12), 2423–2435 (2011).
- [15] Miao, L. and Qi, H., "The design and evaluation of a generic method for generating mosaicked multispectral filter arrays," *IEEE Trans. on Image Processing* **15**(9), 2780–2791 (2006).
- [16] Miao, L., Qi, H., Ramanath, R., and Snyder, W. E., "Binary tree-based generic demosaicking algorithm for multispectral filter arrays," *IEEE Trans. on Image Processing* **15**(11), 3550–3558 (2006).
- [17] Monno, Y., Tanaka, M., and Okutomi, M., "Multispectral demosaicking using adaptive kernel upsampling," *Proc. of IEEE Int. Conf. on Image Processing (ICIP)*, 3218–3221 (2011).
- [18] Monno, Y., Tanaka, M., and Okutomi, M., "Multispectral demosaicking using guided filter," *Proc. of SPIE* **8299**, 82990O–1–82990O–7 (2012).
- [19] Chakrabarti, A. and Zickler, T., "Statistics of real-world hyperspectral images," *Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR)*, 193–200 (2011).
- [20] Monno, Y., Kitao, T., Tanaka, M., and Okutomi, M., "Optimal spectral sensitivity functions for a single-camera one-shot multispectral imaging system," *Accepted to IEEE Int. Conf. on Image Processing (ICIP)*, 2137–2140 (2012).