MULTISPECTRAL DEMOSAICKING USING ADAPTIVE KERNEL UPSAMPLING

Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi

Department of Mechanical and Control Engineering, Tokyo Institute of Technology

ABSTRACT

Multispectral demosaicking, which estimates full multispectral images from raw data observed using a single image sensor with a color filter array (CFA), is a challenging task because each spectral component is severely undersampled. In this paper, we propose a novel multispectral demosaicking algorithm. We extend existing upsampling algorithms to adaptive kernel upsampling algorithms using an adaptive kernel as a spatial weight and apply them to multispectral demosaicking. We also propose a new CFA and a direct adaptive kernel estimation from the raw data of the proposed CFA. Experimental results with real multispectral images demonstrate the effectiveness of the proposed algorithm.

Index Terms— Multispectral imaging, color filter array, demosaicking, adaptive kernel

1. INTRODUCTION

Multispectral imaging that measures more than three spectral components can provide not only better fidelity for image displays, but also marked advantages to many computer vision applications such as scene segmentation, object recognition, and relighting.

Multispectral imaging systems can be broadly classified into three groups: (i) multi-camera-one-shot systems, (ii) single-camera-multi-shot systems, and (iii) single-camera-one-shot systems. A six-band HDTV camera which consists of an optical splitter and two RGB cameras with different color filters is one of multi-camera-one-shot systems [1]. Although this system can capture a multispectral video, it is expected to be expensive because multiple cameras are required. The single-camera-multi-shot system can be implemented by sequentially replacing color filters in front of a camera [1] or by changing light sources [2]. However, these systems can not capture multispectral images at video rate because multiple shots are necessary to observe the multispectral images. Special equipment such as a high-speed lighting system is required to improve temporal resolution [3, 4]. Common consumer RGB cameras can be considered as single-camera-one-shot systems. Consumer RGB cameras usually use a single image sensor with a color filter array (CFA). The most famous CFA is the Bayer CFA [5]. Data observed using the single image sensor with the CFA is called raw data; only one spectral component is measured at each pixel location. Therefore, full color images are estimated from the raw data using a demosaicking process. Extension of this single-camera-one-shot system to multispectral imaging provides advantages of low-cost and simple video acquisition. However, multispectral demosaicking, which estimates full multispectral images from raw data is a challenging task because each spectral component is severely undersampled.

Almost all existing demosaicking algorithms are intended for use with the Bayer CFA [6]. Recently, demosaicking algorithms for an arbitrary CFA have been proposed [7, 8]. However, these algorithms are mainly discussed for RGB image demosaicking. Although a reconstruction-based algorithm [9] can be applied for multispectral demosaicking, it requires a huge amount of training data. To our knowledge, few multispectral demosaicking algorithms exist [10, 11, 12]. Baone et al. proposed a MAP-based algorithm [10]. However, this algorithm requires a high computational cost. In other interpolation-based algorithms, Brauers et al. apply lowpass filters to color differences [11] and Miao et al. perform an edge-sensing interpolation to each band [12]. Multispectral images demosaicked using these algorithms suffer from severe color artifacts especially in edge regions.

In this paper, we propose a novel multispectral demosaicking algorithm. We extend existing upsampling algorithms using an adaptive kernel as a spatial weight and apply to multispectral demosaicking. The adaptive kernel is a spatially variant kernel estimated based on image structures [13]. In the case of an upsampling, the adaptive kernel can not be estimated from raw data because of the sparsity of input data. Therefore, an initial interpolation is required. In contrast, we propose a new CFA and a direct adaptive kernel estimation from the raw data of the proposed CFA. Furthermore, we set a limitation on the adaptive kernel size to avoid color artifacts in strong edge or rich texture regions. We demonstrate that the proposed algorithm outperforms other algorithms both visually and in terms of a peak signal-to-noise ratio (PSNR).

2. ADAPTIVE KERNEL UPSAMPLING

2.1. Gaussian upsampling and joint bilateral upsampling

Gaussian upsampling (GU) and joint bilateral upsampling (JBU) [14] are famous upsampling algorithms. Each upsampled result for a location $\mathbf{x}_p$ is obtained as:

$$S^{GU}(\mathbf{x}_p) = \frac{1}{w^{GU}_{\mathbf{x}_p}} \sum_{\mathbf{x}_i \in \mathbf{N}_{\mathbf{x}_p}} k(\mathbf{x}_i - \mathbf{x}_p) M(\mathbf{x}_i) S(\mathbf{x}_i), \quad (1)$$

$$S^{JBU}(\mathbf{x}_p) = \frac{1}{w^{JBU}_{\mathbf{x}_p}} \sum_{\mathbf{x}_i \in \mathbf{N}_{\mathbf{x}_p}} k(\mathbf{x}_i - \mathbf{x}_p) r(I(\mathbf{x}_i) - I(\mathbf{x}_p)) M(\mathbf{x}_i) S(\mathbf{x}_i), \quad (2)$$

where $\mathbf{N}_{\mathbf{x}_p}$ is the set of neighbor pixel locations of the location $\mathbf{x}_p$, $S(\mathbf{x}_i)$ is the sampled value at the location $\mathbf{x}_i$, $M(\mathbf{x}_i)$ is the binary mask at the location $\mathbf{x}_i$, $I(\mathbf{x}_i)$ and $I(\mathbf{x}_p)$ are pixel values of a guide image, $k(\mathbf{x}_i - \mathbf{x}_p)$ is the spatial weight, $r(I(\mathbf{x}_i) - I(\mathbf{x}_p))$ is the range weight of the guide image, and $w^{GU}_{\mathbf{x}_p}$ and $w^{JBU}_{\mathbf{x}_p}$ are normalizing factors which are sums of filter weights. The binary mask is set to one if the data is sampled at an associated location and set to zero for other cases. A spatially invariant Gaussian function is typically used for the spatial weight. In this paper, these two upsampling algorithms are referred to as non-adaptive upsampling.

2.2. Adaptive kernel upsampling

We extend the non-adaptive upsampling to an adaptive kernel upsampling using an adaptive kernel as a spatial weight. Takeda et al. proposed the adaptive kernel for kernel regression [13]. The adaptive kernel for the location $\mathbf{x}_p$ is represented as:

$$k_{\mathbf{x}_p}(\mathbf{x}) = \exp\left[-\frac{\mathbf{x}^T \mathbf{C}^{-1}_{\mathbf{x}_p} \mathbf{x}}{2h^2 {\mu_{\mathbf{x}_p}}^2}\right], \quad (3)$$

where $\mathbf{C}_{\mathbf{x}_p}$ is the covariance matrix of the Gaussian kernel, h stands for a global smoothing parameter, and $\mu_{\mathbf{x}_p}$ is a local density parameter, which controls the kernel size. The covariance matrix $\mathbf{C}_{\mathbf{x}_p}$ is estimated based on the derivatives around the location $\mathbf{x}_p$ as:

$$\mathbf{C}^{-1}_{\mathbf{x}_p} = \frac{1}{|\mathbf{N}_{\mathbf{x}_p}|} \begin{pmatrix} \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_p}} z_u(\mathbf{x}_j) z_u(\mathbf{x}_j) & \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_p}} z_u(\mathbf{x}_j) z_v(\mathbf{x}_j) \\ \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_p}} z_u(\mathbf{x}_j) z_v(\mathbf{x}_j) & \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_p}} z_v(\mathbf{x}_j) z_v(\mathbf{x}_j) \end{pmatrix}, \quad (4)$$

where z_u is the horizontal derivative, z_v is the vertical derivative, $\mathbf{N}_{\mathbf{x}_p}$ denotes neighbor pixels around the location $\mathbf{x}_p$, and $|\mathbf{N}_{\mathbf{x}_p}|$ is the pixel number of $\mathbf{N}_{\mathbf{x}_p}$.

We extend GU and JBU by replacing the spatially invariant Gaussian kernel $k(\mathbf{x})$ with the adaptive kernel $k_{\mathbf{x}_p}(\mathbf{x})$ in Eq. (1) and (2). In this paper, we call each adaptive kernel upsampling adaptive Gaussian upsampling (A-GU) and adaptive joint bilateral upsampling (A-JBU).

In strong edge or rich texture regions, adaptive kernels become too small or too narrow. These kernels yield severe color artifacts. To avoid these color artifacts, we set a limitation on the kernel size. The kernel size can be evaluated by the product of the eigenvalues of the covariance matrix $\mathbf{C}_{\mathbf{x}_p}$.

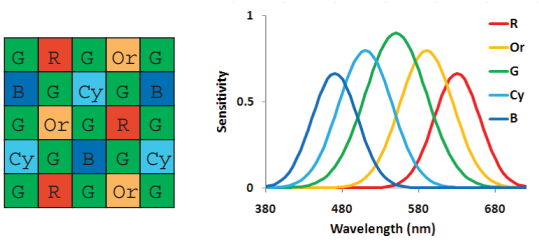


Fig. 1. Proposed CFA and spectral sensitivities

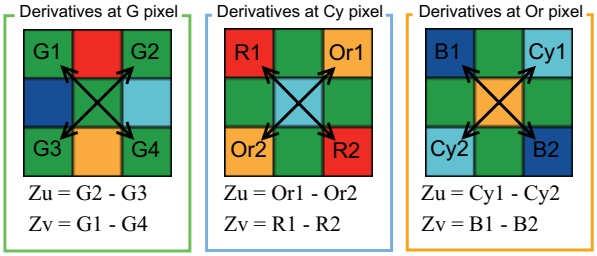


Fig. 2. Examples of diagonal derivatives.

We replace the eigenvalues if the product of the eigenvalues is less than the threshold τ, as shown below:

$$\lambda_1 = {\lambda_1}'/\alpha \ , \ \lambda_2 = {\lambda_2}'/\alpha \ , \ \alpha = \sqrt{\frac{\tau}{\lambda_1 \lambda_2}} \ , \quad (5)$$

where λ_1 and λ_2 are the eigenvalues of covariance matrix $\mathbf{C}_{\mathbf{x}_p}$, ${\lambda_1}'$ and ${\lambda_2}'$ are the replaced eigenvalues.

3. MULTISPECTRAL DEMOSAICKING

3.1. CFA and direct adaptive kernel estimation

The proposed CFA and the corresponding schematic spectral sensitivities of each spectral band are shown in Fig. 1. In this paper, we call each spectral band R, Or, G, Cy, and B from the long-wavelength side to the short-wavelength side. We develop the CFA based on a generic framework proposed by Miao et al. [12]. We decide that the sampling density of the G-band data is higher than that of the other spectral bands because human eyes are more sensitive to the G-band than to the other spectral bands.

The greatest advantage of the proposed CFA is that the adaptive kernel can be estimated directly from the raw data. In Eq. (4), the derivatives must be used to estimate the adaptive kernel. In the case of an upsampling, the derivatives can not be calculated for all pixels because of the sparsity of input data. Therefore, an initial interpolation is necessary for the adaptive kernel estimation. Performance of the initial interpolation affects the adaptive kernel estimation. In contrast, we can estimate the adaptive kernel directly from the raw data of the proposed CFA.

In natural images, color bands are highly correlated in high-frequency components [15]. We assume that derivatives of each spectral band are approximately equal and calculate derivatives in the diagonal directions at all pixels, as shown in Fig. 2. Using these diagonal derivatives, we can estimate

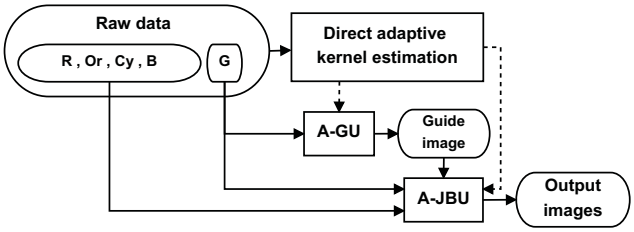


Fig. 3. Block diagram of the proposed algorithm.

the adaptive kernel directly from the raw data. In our case, $\mathbf{x}$ is replaced with $\mathbf{Hx}$ in Eq. (3) where $\mathbf{H}$ is a rotation matrix which rotate coordinates by 45 degrees.

3.2. Multispectral demosaicking algorithm

The schematic block diagram of the proposed algorithm is shown in Fig. 3. The proposed algorithm includes three steps. First, the adaptive kernel is estimated directly from the raw data, as discussed in the previous subsection. Second, the guide image for A-JBU is generated from the G-band data using A-GU. Finally, A-JBU is applied to each spectral band data. Each upsampling is progressively performed similarly in [12].

The inconsistency among the spectral band images causes severe color artifacts. To avoid these color artifacts, we apply A-JBU with the same guide image to interpolate all spectral band images including the G-band image so that the frequency properties of all spectral band images are consistent.

4. EXPERIMENTS

Five-band multispectral images are captured and used as original images for experimental comparisons. Then, original standard RGB (sRGB) images are estimated from the original five-band images using a spatio-spectral Wiener estimation [16]. We experimentally validate with 16 scenes. Some sets of the original five-band images are available at http://www.ok.ctrl.titech.ac.jp/res/MSI/AKU.html.

We compare three multispectral demosaicking algorithms. Five-band images are sampled assuming the proposed CFA and demosaicked by (i) a binary tree-based edge sensing (BTES) algorithm [12], (ii) the non-adaptive upsampling (N-UP) algorithm, and (iii) the proposed adaptive kernel upsampling algorithm. In the N-UP algorithm, the guide image is generated from the G-band data using GU and JBU is applied to each spectral band. The standard deviation of the spatial weight is set to 1.3. In both the N-UP and the proposed algorithms, the kernel size is set to 9×9. We use the Gaussian function for the range weight. The standard deviation of the range weight is set to 0.05 with pixel values normalized to the interval [0,1]. The global smoothing parameter h and the local density parameter $\mu_{\mathbf{x}_p}$ are set to 1 and τ is selected manually for the proposed algorithm. The full five-band images demosaicked using each algorithm are converted to sRGB for visual comparisons. Fig. 4 shows resultant sRGB images of CHINADRESS2. These results demonstrate that the proposed algorithm can reduce color artifacts, especially those in the edge of the face, compared to the other algorithms.

Next, we evaluate the color reproducibility of multispectral imaging. Bayer mosaic images are simulated using only R, G, and B spectral bands and are demosaicked using a local polynomial approximation (LPA) algorithm [17] and an adaptive homogeneity-directed (AHD) algorithm [18]. These two algorithms are known as high-performance demosaicking algorithms for the Bayer CFA. Demosaicked RGB images are converted to sRGB for comparisons with the images converted from five-band images. Fig. 5 shows resultant sRGB images of BUTTERFLY1. The proposed algorithm correctly reproduces color, while the converted images from the demosaicked RGB have color artifacts on the butterfly wing.

Table 1 shows the PSNR(dB) performance of demosaicked images on the multispectral image dataset. Both the BTES and the N-UP algorithms underperform Bayer demosaicking algorithms in some of sRGB bands. In contrast, the proposed algorithm outperforms the other algorithms in both five-bands and sRGB bands. These PSNR comparisons show that the proposed algorithm yields better performance in general.

5. CONCLUSION

In this paper, we propose a novel multispectral demosaicking algorithm using the adaptive kernel upsampling. The proposed algorithm estimates the adaptive kernel directly from the raw data of the proposed CFA. The experimental results show that the proposed algorithm demosaicks multispectral images effectively and generates high-fidelity color images.

6. REFERENCES

[1] M. Yamaguchi, H. Haneishi, H. Fukuda, J. Kishimoto, H. Kanazawa, M. Tsuchida, R. Iwama, and N. Ohyama, "High-fidelity video and still-image communication based on spectral information: Natural vision system and its applications," in *Proc. of SPIE*, vol. 6062, pp. 129–140, 2006.

[2] C. Cui, H. Yoo, and M. Ben-Ezra, "Multi-spectral imaging by optimized wide band illumination," *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 140–151, 2010.

[3] J. Park, M. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral imaging using multiplexed illumination," in *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 1–8, 2007.

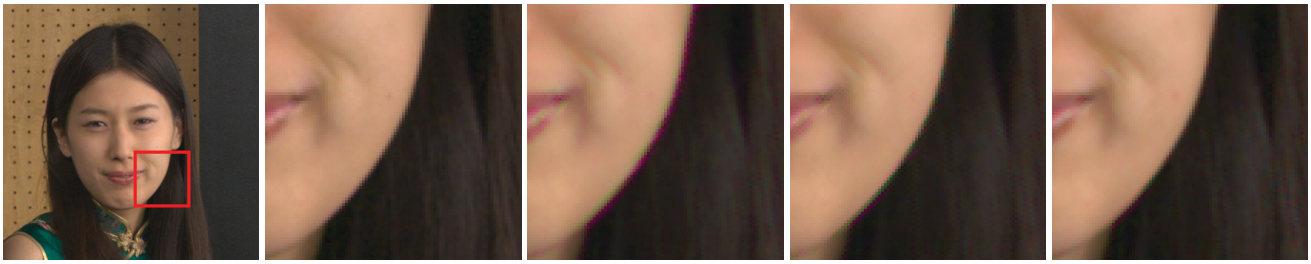
[4] S. Han, I. Sato, T. Okabe, and Y. Sato, "Fast spectral reflectance recovery using DLP projector," in *Proc. of Asian Conference on Computer Vision (ACCV)*, vol. 1, pp. 318-330, 2010.

[5] B. Bayer, "Color imaging array," US Patent 3971065, 1976.

[6] X. Li, B. K. Gunturk, and L. Zhang, "Image demosaicing: a systematic survey," in *Proc. of SPIE*, vol. 6822, pp. 68221J–68221J-15, 2008.

[7] J. Gu, P. J. Wolfe, and K. Hirakawa, "Filterbank-based universal demosaicking," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, pp. 1981–1984, 2010.

[8] L. Condat, "A generic variational approach for demosaicking from an arbitrary color filter array," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, pp. 1625–1628, 2009.



(a) CHINADRESS2 (b) Original (c) BTES (d) N-UP (e) Proposed

Fig. 4. Visual comparison of sRGB images on a part of CHINADRESS2 (Gamma correction is applied for the display).

(a) BUTTERFLY1 (b) Original (c) AHD (d) LPA (e) Proposed

Fig. 5. Visual comparison of sRGB images on a part of BUTTERFLY1 (Gamma correction is applied for the display).

Table 1. PSNR(dB) performance comparison of different demosaicking algorithms on the multispectral image dataset, where the bold typeface represents the highest PSNR.

Image index	Demosaicking algorithm	Band index							
		R	Or	G	Cy	B	sR	sG	sB
CHINADRESS2	AHD	-	-	-	-	-	31.45	39.24	35.71
	LPA	-	-	-	-	-	33.39	42.32	38.30
	BTES	48.53	44.09	49.24	47.34	49.38	33.28	44.06	40.73
	N-UP	51.00	46.75	47.53	49.72	51.17	36.88	43.61	42.43
	Proposed	**52.74**	**47.17**	**49.31**	**50.05**	**52.47**	**37.07**	**45.09**	**43.71**
BUTTERFLY1	AHD	-	-	-	-	-	27.22	36.15	34.20
	LPA	-	-	-	-	-	29.33	39.04	36.08
	BTES	45.71	42.20	45.29	37.57	40.54	31.11	40.29	32.44
	N-UP	48.22	45.21	44.75	40.51	42.95	35.04	40.60	34.53
	Proposed	**50.48**	**46.24**	**46.70**	**41.88**	**45.17**	**36.11**	**42.29**	**37.34**
Average of 16 scenes	AHD	-	-	-	-	-	28.80	38.59	34.47
	LPA	-	-	-	-	-	30.39	41.24	36.71
	BTES	49.38	45.00	48.60	42.78	44.93	34.46	42.95	36.36
	N-UP	51.24	47.32	46.68	44.84	46.60	37.57	42.76	37.85
	Proposed	**52.19**	**47.80**	**48.78**	**45.38**	**48.06**	**38.14**	**44.20**	**39.53**

[9] S.G. Narasimhan and S.K. Nayar, "Enhancing resolution along multiple imaging dimensions using assorted pixels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 518–530, 2005.

[10] G. A. Baone and H. Qi, "Demosaicking methods for multispectral cameras using mosaic focal plane array technology," in *Proc. of SPIE*, vol. 6062, pp. 75-87, 2006.

[11] J. Brauers and T. Aach, "A color filter array based multispectral camera," in *12. Workshop Farbbildverarbeitung*, 2006.

[12] L. Miao, H. Qi, and R. Ramanath, "Generic MSFA mosaicking and demosaicking for multispectral cameras," in *Proc. of SPIE*, vol. 6069, pp. 88-97, 2006.

[13] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, pp. 349–366, 2007.

[14] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," *ACM Transactions on Graphics*, vol. 26(3), no. 96, 2007.

[15] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Transactions on Image Processing*, vol. 11, pp. 997–1013, 2002.

[16] Y. Murakami, K. Fukura, M. Yamaguchi, and N. Ohyama, "Color reproduction from low-SNR multispectral images using spatio-spectral wiener estimation," *Optics Express*, vol. 16, no. 6, pp. 4106–4120, 2008.

[17] D. Paliy, V. Katkovnik, R. Bilcu, S. Alenius, and K. Egiazarian, "Spatially adaptive color filter array interpolation for noiseless and noisy data," *International Journal of Imaging Systems and Technology*, vol. 17, pp. 105–122, 2007.

[18] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicking algorithm," *IEEE Transactions on Image Processing*, vol. 14, pp. 360–369, 2005.