

OPTIMIZED HIGH-FREQUENCY BASED INTERPOLATION FOR MULTISPECTRAL DEMOSAICKING

Sunil Prasad Jaiswal, Lu Fang, Vinit Jakhetiya, Manohar Kuse, Oscar C. Au

The Hong Kong University of Science and Technology, Hong Kong

Email: { spjaiswal, eefang, vjakhetiya, mpkuse, eeau }@ust.hk

ABSTRACT

Multispectral demosaicking, which is an extension of color demosaicking, is a challenging problem because each band is significantly undersampled and thus precise reconstruction is needed for the restoration of high-frequency components, such as edges, textures etc. In general, existing algorithms borrow high-frequency information either from different bands via inter-color correlation or from within the bands, and produces artifact in the reconstructed image. To meet this inherent shortcoming, we propose to incorporate two different high-frequency components and integrate them optimally in the linear minimum mean square sense (LMMSE) for the precise reconstruction of undersampled components. Experimental results demonstrate that the proposed algorithm based on the optimized high-frequency achieves superior performance compared to existing algorithms both in terms of objective and subjective quality.

Index Terms— Multispectral demosaicking, high-frequency components, Bayer color filter array, correlation, interpolation.

1. INTRODUCTION

A multispectral imaging system captures more than three spectral bands, and thus retains more information of a scene than a standard color image (3-band) or gray-scale image. Hence, multispectral imaging is widely used in such fields as medical imaging [1], satellite imaging [2], and computer vision [3]. Several multispectral imaging systems have been developed in the past but they have limitations in terms of size, real-time imaging, perfect alignment of cameras etc. [4-6].

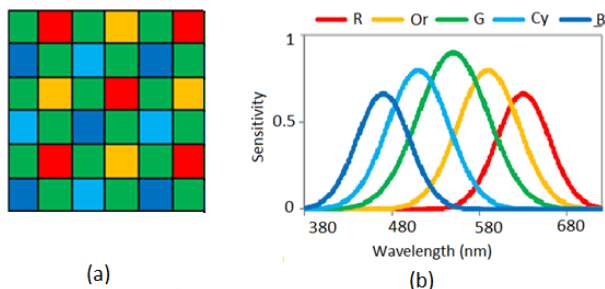


Fig. 1. (a) 5-band MSFA pattern [9-12], (b) schematic spectral sensitivities [9-12]

Recently, a single-camera-one-shot (SCOS) based multispectral imaging system [7-12] has been designed which is an extension of consumer RGB cameras. In consumer RGB cameras, a single image sensor is overlaid with a color filter array (CFA), whereas in

SCOS based multispectral imaging, the CFA is replaced with a multispectral filter array (MSFA) which allows more than three spectral bands to be subsampled, as shown in Fig.1(a). CFA interpolation, or color image demosaicking, is carried out for RGB cameras, in order to restore a full color image. Similarly multispectral demosaicking is required to be performed when using an MSFA for SCOS-based multispectral imaging. However, multispectral demosaicking is a challenging problem because each band is significantly undersampled due to the increase in the number of spectral bands.

Various algorithms have been proposed in literature for CFA demosaicking [13] but very few algorithms address the challenging problem of multispectral demosaicking [7-12]. Brauers et al. [7] perform low-pass filtering to the color-difference image for the restoration of full components, whereas Miao et al. [8] perform edge-sensing based interpolation for each band. Monno et al. [9] propose an adaptive kernel based upsampling to interpolate each spectral component and then they improve it, in [10], by using guided filtering.

Recently, RI-based algorithms [11,14] have been proposed which perform low-pass filtering in the residual domain. They first estimate a tentative image of an undersampled component using a guided filter and then assume that the guided filter can reproduce the high-frequency components in the tentative image. In this paper, we efficiently demonstrate the distortion analysis of the residual interpolation technique [11,14] and observe that such an assumption does not hold true for all images, and hence the quality of the reconstructed image is reduced. In view of this observation, we propose an algorithm that borrows the high-frequency component from other bands based on inter-color correlation, and then integrates it optimally with the high-frequency component of the tentative image to achieve superior quality of the reconstructed undersampled components.

2. RESIDUAL INTERPOLATION (RI)

RI-based algorithm has recently been developed for CFA demosaicking [11,14] and it has been extended to MSFA demosaicking because of its superior performance. The working process of the RI-based scheme is presented as follows. Since the sampling frequency of the **G** band is much higher than that of other bands, the RI-based algorithm first reconstructs the Green component which can preserve the high frequencies of the **G** channel. The reconstructed green component is denoted as \hat{G} . The RI-based algorithm focuses more on the reconstruction of other bands as they are significantly undersampled. In our discussion, we initially consider the discussion of the **R** band as the reconstruction of other bands may be treated in a similar manner.

The RI-based algorithm estimates a tentative image of an un-

dersampled component using a guided filter [14]. We denote the tentative image of the \mathbf{R} band as \mathbf{R}^T . Then they estimate the sub-sampled residual image $ResI_R = R_a - R_a^T$ at the available \mathbf{R} pixels, where R_a and R_a^T denote the original red (\mathbf{R}) and the tentative image (\mathbf{R}^T) at the available \mathbf{R} positions, as shown in Fig.1(a). The RI-based algorithm assumes that the tentative image can preserve the edges and textures, which implies that the high-frequency component of the original \mathbf{R} (R_h) and high-frequency component of the tentative image \mathbf{R}^T (R_h^T) are similar, i.e., $R_h \approx R_h^T$. Based on this assumption, we can write the residual image ($ResI$) as

$$ResI = R - R^T = (R_l + R_h) - (R_l^T + R_h^T) \approx (R_l - R_l^T), \quad (1)$$

where R_l and R_l^T denote the low-frequency component of \mathbf{R} and \mathbf{R}^T , respectively. We can see that the residual image ($ResI$) contains a low-pass signal in (1) and thus the RI-based algorithm applies a low pass filter to the subsampled residual image ($ResI_R$) and add it to the tentative image (\mathbf{R}^T) to get the interpolated red component (\hat{R}_{RI}), which is given as

$$\hat{R}_{RI} = \zeta\{ResI_R\} + R^T = \zeta\{R_a - R_a^T\} + R^T, \quad (2)$$

where $\zeta\{\cdot\}$ denotes low pass filtering. RI algorithm further assumes that the low pass filtering of R_a and R_a^T gives exactly the low frequencies of \mathbf{R} (R_l) and \mathbf{R}^T (R_l^T) and the reconstructed red component by the RI method is

$$\hat{R}_{RI} = \zeta\{R_a - R_a^T\} + R^T = (R_l - R_l^T) + R^T. \quad (3)$$

Similarly, other components can be reconstructed by following a similar process.

3. PROPOSED ALGORITHM

The RI-based algorithm assumes that the residual image ($ResI = \mathbf{R} - \mathbf{R}^T$) contains a low pass signal and we experimentally observed that such assumption is not valid for all the images, and hence produces artifacts in the reconstructed image. In view of this observation, we made a distortion analysis of the RI methods and then proposed an algorithm that works for all images.

3.1. Distortion Analysis of the RI method

The interpolated red component (\hat{R}_{RI}) by the RI method from (3) can be written as

$$\hat{R}_{RI} = (R_l - R_l^T) + R^T = R_l + R_h^T, \quad (4)$$

where R_h^T ($R^T - R_l^T$) is the high pass filtered output of the tentative image (\mathbf{R}^T). We can see that the RI method copies the high-frequency component from the tentative image for the reconstruction. We estimate the distortion between the original \mathbf{R} component and interpolated \hat{R}_{RI} component, and it is given in (5):

$$\begin{aligned} D_{RI} &= R - \hat{R}_{RI} = R - R_l - R_h^T \\ \implies D_{RI} &= R_h - R_h^T, \end{aligned} \quad (5)$$

where R_h ($R - R_l$) is the high pass filtered \mathbf{R} component. Taking the square of both sides of (5) and then taking the expected value, we can write

$$\begin{aligned} D_{RI}^2 &= R_h^2 + (R_h^T)^2 - 2 \times R_h \times R_h^T, \\ \implies E(D_{RI}^2) &= E(R_h^2) + E((R_h^T)^2) - 2 \times E(R_h \times R_h^T). \end{aligned} \quad (6)$$

Here, $E(D_{RI}^2)$ is the average distortion of the RI method for the red component. It is a well-known fact that the high-frequency component and residual image follows a Laplace Distribution [15,16] with zero mean, i.e., $E(D_{RI}) = 0$, $E(R_h) = 0$, and $E(R_h^T) = 0$, therefore, variance can be given by $\sigma_{D_{RI}}^2 = E(D_{RI}^2)$ and $\sigma_{R_h}^2 = E(R_h^2)$, $\sigma_{R_h^T}^2 = E((R_h^T)^2)$ respectively. We then calculate the correlation coefficient between R_h and R_h^T , which is given by $\rho_{R_h R_h^T} = E[R_h \times R_h^T] / (\sigma_{R_h} \times \sigma_{R_h^T})$. Substituting these terms in (6), we get

$$\sigma_{D_{RI}}^2 = \sigma_{R_h}^2 + \sigma_{R_h^T}^2 - 2 \times \sigma_{R_h} \times \sigma_{R_h^T} \times \rho_{R_h R_h^T}, \quad (7)$$

where $\sigma_{D_{RI}}^2$ is the variance of distortion of the RI method for the \mathbf{R} component. In (7), $\rho_{R_h R_h^T}$ is the correlation coefficient between the original high frequency R_h and the tentative high frequency R_h^T , and it indicates about the similarity between the high frequency components. If the value of $\rho_{R_h R_h^T}$ increases, which implies that the similarity between the high-frequency components increases, then the overall distortion of the RI method in (7) decreases.

3.2. Optimized high-frequency (OHF) component based demosaicking scheme

From the above distortion analysis, we observe that a higher correlation coefficient will reduce the overall distortion. Moreover, we observe that the assumption of the RI-method about the similarity of high-frequency components in (1) is indeed image dependent and may not work for all images, and thus it can increase the overall distortion in (7). In CFA demosaicking, most of the algorithms use inter-color correlation property for the reconstruction of undersampled components [13], i.e., they assume that the high-frequency components of different bands are very similar.

In view of this observation, we propose to introduce another high-frequency component in the reconstruction framework based on inter-color correlation property. We use the high-frequency of $\hat{\mathbf{G}}$ band for the reconstruction of other bands as it is highly sampled (see Fig. 1) and thus $\hat{\mathbf{G}}$ can preserve the high-frequency of the \mathbf{G} band. Our algorithm linearly combines the high-frequency component of $\hat{\mathbf{G}}$ with the high-frequency component of the tentative image (R_h^T) on a block-by-block basis to generate a more accurate reconstruction of undersampled components. We denote the proposed optimized high-frequency (OHF) component based interpolated red component as \hat{R}_{OHF} and it is given by

$$\hat{r}_{OHF} = r_l + (w_1 \times r_h^T + w_2 \times \hat{g}_h). \quad (8)$$

Here $\hat{r}_{OHF} \in \hat{R}_{OHF}$ is the interpolated block of size $m \times n$ by the proposed algorithm, $r_h^T \in R_h^T$ and $\hat{g}_h \in \hat{G}_h$ are the high-frequency components, and w_1 and w_2 are the weighted coefficients for the combination such that $w_1 + w_2 = 1$. The optimal weights are obtained as follows. The distortion between the original red ($r \in \mathbf{R}$) and interpolated \hat{r}_{OHF} can be written as

$$\begin{aligned} d &= r - \hat{r}_{OHF} = r - r_l - (w_1 r_h^T + w_2 \hat{g}_h) \\ \implies d &= r_h - w_1 r_h^T - w_2 \hat{g}_h, \end{aligned} \quad (9)$$

Table 1. AVERAGE PSNR (IN DB) RESULTS FOR THE MONNO + CAVE DATASETS. (PRO1, PRO2 refers to proposed algorithm with weight scheme 1 and 2, respectively.)

5-band							Standard RGB (sRGB)			
Algorithm	R	G	B	Or	C	mean	sR	sG	sB	mean
RI [14]	-	-	-	-	-	-	33.99	39.71	34.12	35.94
SSD [19]	-	-	-	-	-	-	33.14	39.12	33.84	35.37
LSLC [21]	-	-	-	-	-	-	32.61	37.65	32.18	34.14
BTES [8]	43.45	46.91	41.28	40.85	38.80	42.26	34.92	40.27	34.81	36.67
AKU [9]	45.67	47.34	43.45	42.62	40.69	43.96	36.84	41.81	36.17	38.28
GF [10]	46.81	48.02	42.70	44.50	42.80	44.97	37.91	42.95	37.81	39.56
POS [12]	47.41	48.29	45.11	46.53	45.17	46.50	38.93	43.62	39.10	40.55
PRO1	47.79	48.20	46.22	46.40	46.05	46.93	39.25	43.75	39.45	40.81
PRO2	47.95	48.20	46.45	46.67	46.28	47.11	39.34	43.91	39.64	40.95

where r_h is the high-frequency component of $r \in \mathbf{R}$. We can write (9) as

$$d = w_1(r_h - r_h^T) + w_2(r_h - \hat{g}_h). \quad (10)$$

To get the optimal weights (w_1, w_2) in linear minimum mean square sense (LMMSE), the problem can be formulated as

$$\begin{aligned} \min_{w_1, w_2} \quad & E[d^2], \\ \text{s.t.} \quad & \sum_{i=1,2} w_i = 1. \end{aligned} \quad (11)$$

To minimize the residual energy ($E[d^2]$) for each block, differentiate it with respect to w_1 and w_2 to get the optimal weights

$$\begin{cases} w_1 = E[(r_h - \hat{g}_h)(r_h^T - \hat{g}_h)] / E[(r_h^T - \hat{g}_h)^2] \\ w_2 = E[(r_h - r_h^T)(\hat{g}_h - r_h^T)] / E[(r_h^T - \hat{g}_h)^2]. \end{cases} \quad (12)$$

So, by substituting (12) in (8) we can efficiently estimate an optimized-high frequency component, and thus a better reconstruction in (8) is achieved. The advantage of optimized high-frequency component is that it can reduce the overall distortion in (7). Unfortunately, to estimate r_h in (12), we need the original block (r) which is not available in practice. In our experiments, we propose two schemes to find an estimate of r_h for each block:

1. Scheme 1 : We apply a high-pass filter on the available sampled pixels of $r \in \mathbf{R}$ with the help of available neighboring pixels, as shown in Fig. 1 as an estimate of r_h . We then calculate r_h^T and \hat{g}_h on the available sampled pixels and thus an estimate of w_1 and w_2 can be obtained in (12).
2. Scheme 2 : We use the resultant image of an existing method as an estimation of the original block (r) and apply a high-pass filter to it and replace it with r_h in (12). Then an estimate of w_1 and w_2 can be obtained thereafter. In our experiments, we propose to use the resultant image of the method [12] as an estimation of original block (r). However, other algorithms can be used for the estimation purpose.

4. SIMULATION RESULT

We implemented the proposed multispectral demosaicking algorithm and compared its performance with the existing methods. For performance evaluation, we used two image datasets, the CAVE

dataset [26] and MONNO dataset [9-10]. The CAVE dataset consists of 32 scenes, and the MONNO dataset consists of 12 scenes. Both datasets consist of images showing high textured regions as well as smooth regions, and thus it is a challenge to perform well in all of the datasets. We use CIE D65 (daylight) [23,24] illumination for the evaluation. We compared the proposed algorithm with the state-of-the-art existing algorithms, such as BTES [8], AKU [9], GF [10], and POS [12], for the performance evaluation of 5-band demosaicking. For better evaluation, we have shown both the objective as well as subjective quality and compared the same with various methods. Our algorithm combines both the high-frequency components on a block-by-block basis, and empirically we kept the block size as 10×10 in our simulation. However, the block size can vary for higher resolution images.

We also compared the performance of the 5-band multispectral imaging with the standard 3-band RGB imaging in the sRGB domain for both of the datasets. To convert the 5-band images to standard RGB (sRGB) image, we used spatio-spectral Wiener estimation [11,25]. For 3-band RGB demosaicking (color imaging) we used the Bayer pattern [13] to downsample only R, G, and B bands among the five bands, and then conducted the demosaicking process using three state-of-the-art algorithms: RI [14], SSD [19], and LSLC [21]. The demosaicked RGB was then converted to an sRGB image using the method in [25]. For performance comparisons, we compared the sRGB images generated by the 5-band demosaicking algorithm with the sRGB images generated by 3-band demosaicking. For ground-truth, we used sRGB images converted from the ground-truth five band images as ground-truth sRGB images.

Our algorithm focuses on optimizing the high-frequency components for the reconstruction of significantly undersampled components, such as **R**, **B**, **Or**, and **C**. Hence, in our simulation, we used method [12] for the reconstruction of the $\hat{\mathbf{G}}$ component. In Table 1, we show the average PSNR of various algorithms, where the average PSNR is calculated for 44 images from both of the datasets. We can see that both of the proposed algorithms (PRO1 and PRO2) have achieved superior performance over the existing algorithms. On an average (mean column), proposed algorithm (PRO2) exceeds POS [12] and GF [10] methods by 0.61 dB, and 2.14 dB respectively. We have also shown the average PSNR for the sRGB images using different algorithms in Table 1, and it can be seen that the proposed algorithms have outperformed the existing algorithms in terms of PSNR. On an average, proposed algorithm (PRO2) exceeds POS [12] and RI [14] by 0.40 dB and 5.0 dB respectively.

To show the effectiveness of the proposed algorithm, we have

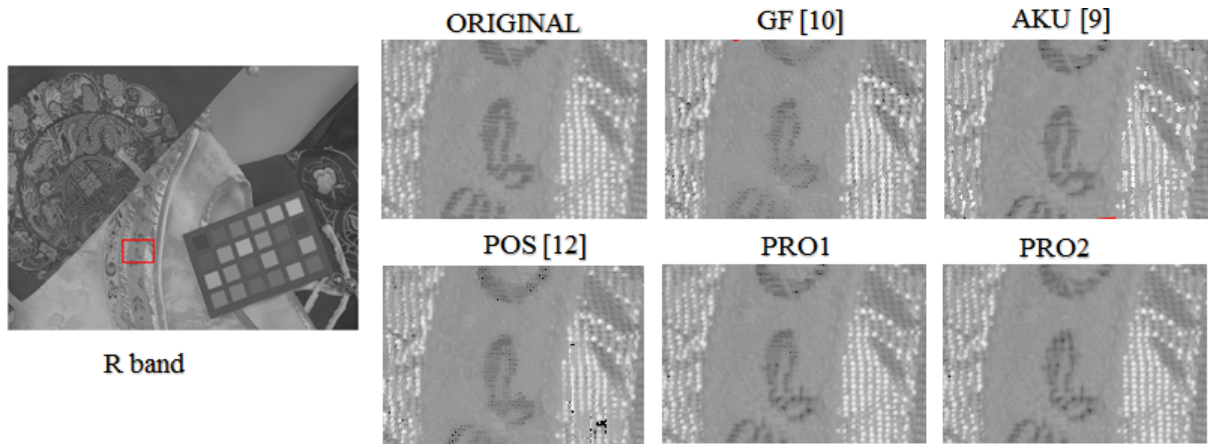


Fig. 2. Visual comparison of the R band of the CHINADDRESS image in the MONNO dataset. (Gamma correction is applied for the display.)



Fig. 3. Visual comparison of sRGB of SPONGE image in CAVE dataset. (Gamma correction is applied for the display.)

shown the visual comparison of a test image that consists of texture and edges. In Fig. 2, we can see that the existing method either blurs the reconstructed image or creates artifacts, whereas the proposed algorithms have precisely reconstructed the high-frequency components, and thus have achieved superior performance when compared to the existing algorithms. In Fig. 3, we show the visual comparisons of standard RGB images (color imaging) generated using the 3-band demosaicking as well as the 5-band demosaicking algorithms. We can see that our algorithm has reduced the color artifacts significantly while preserving the edges.

5. CONCLUSION

In this paper, we proposed a novel multispectral demosaicking algorithm that generates accurate reconstruction for significantly undersampled bands. Our algorithm is based on the observation that the precise reconstruction of high-frequency components can lead to achieving a precise reconstruction of undersampled components. In view of this observation, we propose an algorithm that borrows two high-frequency components from different images and integrates them linearly in the LMMSE sense. Experimental results demonstrate that the proposed algorithm outperforms the existing

methods both in terms of subjective and objective quality. Future work includes incorporating multiple high-frequency components from different number of bands.

6. ACKNOWLEDGMENT

This research work has been supported in part by the Research Grants Council (RGC) of the Hong Kong Special Administrative Region, China (GRF 16211615). We would like to thank Dr. Yusuke Monno and the other authors of [9-12] for providing the code and datasets for performance comparison.

7. REFERENCES

- [1] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," in *Journal of Biomedical Optics*, 19(1):10901, Jan. 2014.
- [2] C. Yang, J. H. Everitt, Q. Du, B. Luo, and J. Chanussot, "Using High-Resolution Airborne and Satellite Imagery to Assess Crop Growth and Yield Variability for Precision Agriculture," in *Proceedings of the IEEE*, vol.101, no.3, pp.582-592, Mar. 2013.
- [3] A. Zia, J. Liang, J. Zhou, and Y. Gao, "3D Reconstruction from Hyperspectral Images," in *Applications of Computer Vision*

- (WACV), *2015 IEEE Winter Conference on*, vol., no., pp.318-325, 5-9 Jan. 2015.
- [4] C. Cui, H. Yoo and M. Ben-Ezra, "Multi-spectral imaging by optimized wide band illumination," in *International Journal of Computer Vision*, 86(2-3), 140–151 (2010).
- [5] S. Han, I. Sato, T. Okabe, and Y. Sato, "Fast spectral reflectance recovery using DLP projector," in *Proc. of Asian Conf. on Computer Vision (ACCV)*, pp. 318–330, 2010.
- [6] K. Ohsawa, T. Ajito, H. Fukuda, Y. Komiya, H. Haneishi, M. Yamaguchi, and N. Ohya, "Six-band HDTV camera system for spectrum-based color reproduction," in *Journal of Imaging Science and Technology*, 48(2), 85–92, 2004.
- [7] J. Brauers and T. Aach, "A color filter array based multispectral camera," in *Workshop Farbbildverarbeitung*, 2006.
- [8] L. Miao, H. Qi, R. Ramanath, and W.E. Snyder, "Binary Tree-based Generic Demosaicking Algorithm for Multispectral Filter Arrays," in *Image Processing, IEEE Transactions on*, vol.15, no.11, pp.3550-3558, Nov. 2006.
- [9] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using adaptive kernel upsampling," in *Image Processing (ICIP), 2011 18th IEEE International Conference on* vol., no., pp.3157-3160, 11-14 Sept. 2011.
- [10] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using guided filter," in *Proc. of SPIE*, vol.8299, 2012.
- [11] Y. Monno, D. Kiku, S. Kikuchi, M. Tanaka, and M. Okutomi, "Multispectral demosaicking with novel guide image generation and residual interpolation," in *Image Processing (ICIP), 2014 IEEE International Conference on*, vol., no., pp.645-649, 27-30 Oct. 2014.
- [12] Y. Monno, S. Kikuchi, M. Tanaka, and M. Okutomi, "A Practical One-Shot Multispectral Imaging System Using a Single Image Sensor," in *Image Processing, IEEE Transactions on*, vol.24, no.10, pp.3048-3059, Oct. 2015.
- [13] X. Li, B. Gunturk, and L. Zhang, "Image demosaicking: A systematic survey," in *proc. SPIE*, vol.6822, p. 68221J, 2008.
- [14] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Residual interpolation for color image demosaicking," in *Image Processing (ICIP), 2013 IEEE International Conference on*, vol., no., pp.2304-2308, 15-18, Sep. 2013.
- [15] L. Fang, O.C. Au, K. Tang, and A.K. Katsaggelos, "Antialiasing Filter Design for Subpixel Downsampling via Frequency-Domain Analysis," in *Image Processing, IEEE Transactions on*, vol.21, no.3, pp.1391-1405, Mar 2012.
- [16] J. Zhou and O.C. Au, "Determining the Capacity Parameters in PEE based Reversible Image Watermarking," in *Signal Processing Letters, IEEE*, vol.19, no.5, pp.287-290, May 2012.
- [17] M. Kuse and S. P. Jaiswal, "Graph modelling of 3D geometric information for color consistency of multiview images," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Quebec City, QC, pp.1394–1398, 2015, doi: 10.1109/ICIP.2015.7351029
- [18] V. Jakhetiya, W. Lin, S. P. Jaiswal, A. K. Tiwari and S. C. Guntuku, "Observation model based perceptually motivated bilateral filter for image reconstruction," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Singapore, 2015, pp. 201–205, doi: 10.1109/ICDSP.2015.7251859
- [19] A. Buades, B. Coll, J. -M. Morel, and C. Sbert, "Self-similarity driven color demosaicking," in *IEEE Trans. Image Processing*, Vol. 18, No. 6, 2009.
- [20] S.P. Jaiswal, O.C. Au, V. Jakhetiya, Y. Yuan, H. Yang, "Exploitation of inter-color correlation for color image demosaicking," in *Image Processing (ICIP), 2014 IEEE International Conference*, vol., no., pp.1812-1816, 27-30 Oct. 2014, doi: 10.1109/ICIP.2014.7025363
- [21] B. Leung, G. Jeon, and E. Dubois, "Least-squares luma-chroma demultiplexing algorithm for Bayer demosaicking," in *IEEE Trans. on Image Processing*, vol. 20, no. 7, pp. 1885–1894, 2011.
- [22] K. Hirakawa and T.W. Parks, "Adaptive homogeneity-directed demosaicking algorithm," in *IEEE Trans. on Image Processing*, vol.14, no.3, pp.360-369, Mar. 2005.
- [23] CIE, in <http://www.cie.co.at/index.php/LEFTMENU/DOWNLOADS>.
- [24] Munsell color science laboratory, in <http://www.cis.rit.edu/research/mcsl2/online/cie.php>.
- [25] Y. Murakami, K. Fukura, M. Yamaguchi, and N. Ohya, "Color reproduction from low-SNR multispectral images using spatio-spectral wiener estimation," in *Optics Express*, 16(6), 4106–4120, 2008.
- [26] F. Yasuma, T. Mitsunaga, D. Iso, and S.K. Nayar, "Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range, and Spectrum," in *Image Processing, IEEE Transactions on*, 1 vol.19, no.9, pp.2241-2253, Sept. 2010.