

CAMERA SPECTRAL SENSITIVITY, ILLUMINATION AND SPECTRAL REFLECTANCE ESTIMATION FOR A HYBRID HYPERSPECTRAL IMAGE CAPTURE SYSTEM

Lin Zhang¹, Ying Fu¹, Yinqiang Zheng², Hua Huang¹

¹Beijing Laboratory of Intelligent Information Technology,
School of Computer Science and Technology, Beijing Institute of Technology

²National Institute of Informatics

ABSTRACT

A variety of methods have been proposed to restore high resolution hyperspectral image (HSI) from a hybrid camera system, which captures high spatial resolution RGB images and low spatial resolution HSI. They focused unanimously on HSI super-resolution via fusion, yet did not explore the potential of this kind of system for camera spectral sensitivity (CSS), illumination spectrum, and high spatial resolution spectral reflectance recovery. In this paper, we present a sparse representation based method to estimate the CSS of the RGB camera under unknown illumination for the hybrid camera system. Furthermore, the illumination and high spatial resolution spectral reflectance are simultaneously recovered. Experimental results show the effectiveness of the proposed methods on camera spectral sensitivity, illumination spectrum and spectral reflectance recovery.

Index Terms— Camera spectral sensitivity, illumination estimation, spectral reflectance recovery

1. INTRODUCTION

In recent years, a variety of methods have been proposed to restore high resolution (HR) hyperspectral image (HSI) from a hybrid camera system, which captures HR RGB/multispectral image and low resolution (LR) HSI. Almost all previous works [1, 2, 3, 4, 5, 6, 7] focused on HSI super-resolution and did not explore the potential capability of this hybrid camera system for RGB CSS, illumination spectrum, and reflectance spectra recovery.

Although the CSS of a hyperspectral camera is very easy to obtain by using, for example, a white board under known illumination, to estimate the CSS of a RGB camera is non-trivial. Generally, RGB CSS has to be calibrated by viewing a standard white target under more than 30 narrow band illuminations from 400nm to 700nm, or recovered based on color checker under known or unknown illumination [8]. Besides, most applications based on hyperspectral imaging, like biometric optics [9] and fluorescent analysis [10, 11], rely on the spectral reflectance. The reason is that spectral reflectance

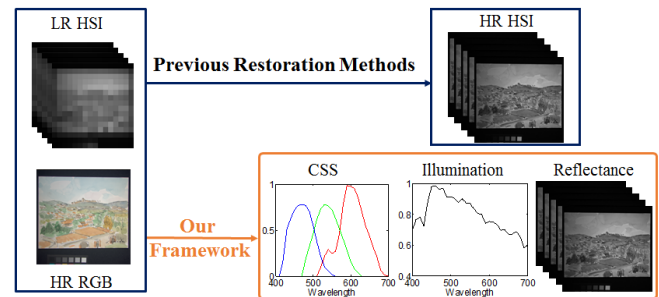


Fig. 1. Existing works [1, 2, 3, 4, 5, 6, 7] on the hybrid camera system employ the LR HSI and HR RGB image to recover HR HSI. In contrast, we further explore the potential of this hybrid camera system for the recovery of RGB camera spectral sensitivity, illumination spectrum, and high spatial resolution spectral reflectance.

is an intrinsic and discriminative characteristic of object materials. However, the captured LR HSI or the recovered HR HSI from the hybrid camera system both couple the illumination and the spectral reflectance.

In this paper, we further explore the potential of this hybrid HSI capture system. Firstly, we present a novel method to recover the spectral sensitivity of the RGB camera based on sparse representation, without resorting to narrow band illuminations and color checkers. In addition, we simultaneously recover the illumination spectrum and the high spatial resolution spectral reflectance, instead of the hyperspectral intensity distribution, as shown in Figure 1. Experimental results on publicly available RGB CSS, standard illumination, and spectral reflectance images have verified the effectiveness of the proposed framework on RGB CSS, illumination spectrum and spectral reflectance recovery.

2. RELATED WORK

Kawakami *et al.* [1] first addressed HR HSI reconstruction from a hybrid camera system, whose inputs are a LR HSI and HR RGB image pair with large downsampling scale, e.g. 32 times. In the following, a great variety of HSI super-

resolution methods [2, 3, 4, 5, 6, 7] have been developed for the same hybrid camera system. All these papers assumed that the RGB camera spectral sensitivity is given, and the illumination and spectral reflectance are coupled in the recovered results.

The CSS recovery of a RGB camera has been achieved by using color checkers under known or unknown illumination [8], which relies heavily on the color variance of the color checkers. Zheng *et al.* [12] proposed an effective method for illumination and spectral reflectance separation, which will be extended to the hybrid camera system in the paper. The spectral reflectance is usually known as the spectral signature of a material, and has been widely used for color constancy [13], material discrimination [14], and relighting [10].

3. THE PROPOSED FRAMEWORK

3.1. Problem Formulation

In this work, the input images are a LR HSI $\mathbf{H} \in \mathbb{R}^{b \times N_l}$ captured by a hyperspectral camera and a HR RGB image $\mathbf{M} \in \mathbb{R}^{3 \times N_h}$ captured by a RGB camera, where b is the number of spectral bands, N_l and N_h are the number of pixels in LR HSI and HR RGB image, respectively. The aim is to simultaneously recover the RGB CSS, the illumination spectra, and the HR spectral reflectance image.

Here, we assume the surface reflection is Lambertian, and the scene illumination is uniform. The HR HSI can be described as

$$\mathbf{Z} = \mathbf{C}\mathbf{L}\mathbf{R} \quad (1)$$

where $\mathbf{L} = \text{diag}(l_1, l_2, \dots, l_b) \in \mathbb{R}^{b \times b}$ denotes the illumination spectrum, and $\mathbf{C} = \text{diag}(c_1, c_2, \dots, c_b) \in \mathbb{R}^{b \times b}$ is the hyperspectral CSS.

The LR HSI is a spatial downsampled version of HR HSI and can be expressed as

$$\mathbf{H} = \mathbf{Z}\mathbf{S} = \mathbf{C}\mathbf{L}\mathbf{R}\mathbf{S}, \quad (2)$$

where $\mathbf{S} \in \mathbb{R}^{N_h \times N_l}$ is the sensor's spatial downsampling operator.

The HR RGB is a spectral downsampled version of HR HSI and can be described as

$$\mathbf{M} = \mathbf{T}\mathbf{Z} = \mathbf{T}\mathbf{L}\mathbf{R}, \quad (3)$$

where $\mathbf{T} \in \mathbb{R}^{3 \times b}$ is the spectral response function of the RGB camera. \mathbf{T} can be represented as $[T_R; T_G; T_B]$, where $\{T_R, T_G, T_B\} \in \mathbb{R}^{1 \times b}$ are CSS of R, G, B channel, respectively.

Previous works [1, 2, 3, 4, 5, 6, 7] indicate that the scene usually contains a small number of distinct materials, and the linear mixing model can be used to describe the reflectance spectra. Thus,

$$\mathbf{R} = \mathbf{D}\mathbf{A} \quad (4)$$

where $\mathbf{D} \in \mathbb{R}^{b \times M}$ is the spectral dictionary and $\mathbf{A} \in \mathbb{R}^{M \times N_h}$ is the coefficient matrix, wherein M is the number of basis in the dictionary.

All previous works [1, 2, 3, 4, 5, 6, 7] focus on restoring the scene's HR intensity \mathbf{Z} with known hyperspectral camera response \mathbf{C} , RGB camera response \mathbf{T} and downsampling operator \mathbf{S} . In practice, \mathbf{C} can be simply calibrated with a white patch in the scene, but the calibration of \mathbf{T} can be burdensome.

In the rest of this section, we will describe our methods to estimate \mathbf{T} and to separate \mathbf{L} and \mathbf{R} under the input LR HSI \mathbf{H} and HR RGB image \mathbf{M} .

3.2. RGB Camera Spectral Sensitivity Estimation

The relationship between LR HSI and HR RGB can be derived from Equations (2) and (3),

$$\mathbf{M}\mathbf{S} = \mathbf{T}\mathbf{C}^{-1}\mathbf{H} = \tilde{\mathbf{T}}\mathbf{H} \quad (5)$$

where $\tilde{\mathbf{T}} = \mathbf{T}\mathbf{C}^{-1}$. Generally, $N_l > B$, and Equation (5) is overdetermined. \mathbf{S} is assumed to be known. The RGB CSS can be directly obtained from $\hat{\tilde{\mathbf{T}}} = \arg \min_{\tilde{\mathbf{T}}} \|\tilde{\mathbf{T}}\mathbf{H} - \mathbf{M}\mathbf{S}\|_F^2$. In practice, noises in \mathbf{H} and \mathbf{M} will hinder high quality reconstruction of $\tilde{\mathbf{T}}$. Besides, all values in RGB CSS are non-negative. Inspired by the model to solve spectral reflectance in [15], we add smooth and non-negative constraints into Equation (6), then

$$\hat{\tilde{\mathbf{T}}} = \arg \min_{\tilde{\mathbf{T}}} \|\tilde{\mathbf{T}}\mathbf{H} - \mathbf{M}\mathbf{S}\|_F^2 + \lambda \|\tilde{\mathbf{T}}\mathbf{W}\|_F^2, \quad s.t. \tilde{\mathbf{T}} \geq 0 \quad (6)$$

where λ is the weight for the smooth constraint, and $\mathbf{W} \in \mathbb{R}^{b \times b}$ is the first-order derivative matrix. Equation (6) is a convex quadratic program.

The above mentioned method considers the smooth property of RGB CSS. We observe the RGB camera response functions of about 40 different commercial cameras, and find that: (i) RGB cameras belong to the same brand, e.g. Nikon, have similar CSS. The reason might be that they share similar hardware inside the camera. (ii) The CSS curves of each channel from different cameras are similar. It leads to the assumption that RGB camera response function of each channel can be linearly represented by the combination of several bases,

$$T_q = \left(\sum_{i=1}^{K_q} \phi_{qi} \alpha_{qi} \right)^T = (\Phi_q \alpha_q)^T \quad (7)$$

where $q \in \{R, G, B\}$, K_q is the number of bases, $\Phi_q \in \mathbb{R}^{b \times K_q}$ is the basis of q channel, and $\alpha_q \in \mathbb{R}^{K_q}$ is the coefficient. Φ_q can be learned from RGB camera spectral sensitivity database by

$$\hat{\Phi}_q = \arg \min_{\Phi_q} \|\mathbf{X}_q - \Phi_q \beta_q\|_F^2 + \eta \|\beta_q\|_1 \quad (8)$$

$s.t. \quad \Phi_q > 0, \beta_q > 0$

where columns of \mathbf{X}_q are response functions of q channel from database, and β_q is the corresponding coefficient matrix. Then, from Equations (5) and (7), we can derive sparse code α_q by

$$\hat{\alpha}_q = \arg \min_{\alpha_q} \|(\Phi_q \alpha_q)^T \mathbf{H} - \mathbf{M}_q \mathbf{S}\|_F^2 + \eta \|\alpha_q\|_1 \quad (9)$$

s.t. $\alpha_q > 0$

Equation (8) learns the dictionary from the database and Equation (9) can be solved by any existing l_1 solver. Then we have $\hat{T}_q = (\hat{\Phi}_q \hat{\alpha}_q)^T$ and $\hat{\mathbf{T}} = [\hat{T}_R; \hat{T}_G; \hat{T}_B]$. If \mathbf{C} is calibrated, we can get $\hat{\mathbf{T}} = \hat{\mathbf{T}} \mathbf{C}$.

3.3. Illumination Spectrum Estimation

Previous work [12] effectively estimated the illumination spectrum on a HSI by using low-rank matrix factorization. In our work, each pixel in LR HSI is the downsampling version of a patch in the HR HSI and can be considered as a linear combination of the pixels in the patch, which do not affect the low-rank property of the reflectance spectra in the scene. Thus, this hybrid camera system can also be used to estimate the illumination spectra in the scene, except the RGB camera spectral sensitivity recovery.

Here, we briefly describe how we use the factorization method [12] in our framework. The input is a LR HSI \mathbf{H} . First, the outliers are removed under the low-rank constraint in LR HSI. Then, SVD is applied to \mathbf{H} and $\mathbf{H} = \mathbf{U} \mathbf{S} \mathbf{V}^T = (\mathbf{U} \mathbf{Q})(\mathbf{Q}^{-1} \mathbf{S} \mathbf{V}^T)$, where \mathbf{Q} is an arbitrary invertible matrix. Finally, quadratic programming is used to solve the optimization problem $\arg \min_{\mathbf{L}, \mathbf{Q}} \|\mathbf{U} \mathbf{Q} - \mathbf{L} \mathbf{B}\|_F^2$, where columns of \mathbf{B} are PCA spectra bases. In addition, the illumination spectra can also be refined by solving $\arg \min_{\mathbf{L}, \mathbf{C}'} \|\mathbf{H} - \mathbf{L} \mathbf{B} \mathbf{C}'\|_F^2$.

In our framework, if hyperspectral CSS \mathbf{C} is calibrated, we can separate $\mathbf{C}^{-1} \mathbf{H}$ to \mathbf{L} and $\mathbf{R} \mathbf{S}$. Otherwise we can separate \mathbf{H} to $\tilde{\mathbf{L}}$ and $\mathbf{R} \mathbf{S}$, where $\tilde{\mathbf{L}} = \mathbf{C} \mathbf{L}$.

3.4. Spectral Reflectance Recovery

Based on Equations (2)-(4), we can model HR spectral reflectance recovery as

$$\{\hat{\mathbf{D}}, \hat{\mathbf{A}}\} = \arg \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{H} - \mathbf{C} \mathbf{L} \mathbf{D} \mathbf{A}\|_F^2 + \|\mathbf{M} - \mathbf{T} \mathbf{L} \mathbf{D} \mathbf{A}\|_F^2 \quad (10)$$

Equation (10) is an ill-posed system, and the sparse constraint on the coefficient matrix \mathbf{A} can be used. Thus, the optimization function can be modeled as

$$\{\hat{\mathbf{D}}, \hat{\mathbf{A}}\} = \arg \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{H} - \mathbf{C} \mathbf{L} \mathbf{D} \mathbf{A}\|_F^2 + \|\mathbf{M} - \mathbf{T} \mathbf{L} \mathbf{D} \mathbf{A}\|_F^2 + \|\mathbf{A}\|_1 \quad (11)$$

We employ the proximal alternating linearized minimization algorithm [5] to optimize $\hat{\mathbf{D}}$ and $\hat{\mathbf{A}}$ in Equation (11). Finally,

Illumination	w/o		$\sigma = 0.5$	
	proposed	baseline	proposed	baseline
A	0.0605	0.1058	0.0771	0.139
C	0.0577	0.0595	0.0714	0.1367
D55	0.0575	0.0572	0.0715	0.1405
D65	0.0577	0.0583	0.071	0.1356
F11	0.0616	0.1499	0.0791	0.1578
F12	0.0624	0.1626	0.0819	0.1679
Mean	0.0596	0.0989	0.0753	0.1462

Table 1. CSS Estimation results of the proposed method and baseline under different noises on NUS database. The RMSE values are the average under the 13 cameras, and the mean values are the average of the 6 illuminations.

Illumination	w/o		$\sigma = 0.5$	
	proposed	[8]	proposed	[8]
A	0.0592	0.1016	0.0797	0.165
C	0.0595	0.094	0.0724	0.111
D55	0.0591	0.0942	0.0727	0.1137
D65	0.0596	0.0941	0.0728	0.1128
F11	0.0633	0.1082	0.0798	0.1409
F12	0.0629	0.1162	0.089	0.1902
Mean	0.0606	0.1014	0.0777	0.1389

Table 2. CSS Estimation results of the proposed method and [8] under different noises. The RMSE values are the average under the 13 cameras, and the mean values are the average of the 6 illuminations. Only the two images containing color checkers in the CAVE dataset are used.

the high spatial resolution spectral reflectance can be recovered by Equation (4), i.e. $\mathbf{R} = \mathbf{D} \mathbf{A}$.

Let us call that, assuming \mathbf{C} is calibrated, we can get \mathbf{T} , \mathbf{L} , and \mathbf{R} . Otherwise, we obtain $\tilde{\mathbf{T}} = \mathbf{T} \mathbf{C}^{-1}$, $\tilde{\mathbf{L}} = \mathbf{C} \mathbf{L}$, and \mathbf{R} . In our experiments, we assume that HSI CSS \mathbf{C} has been calibrated first.

4. EXPERIMENTS

4.1. Database

We test our method on two widely known HSI databases, i.e. the CAVE [16] and NUS [17]. CAVE [16] contains 32 HSIs of real-world materials and objects, in the visible range from 400nm to 700nm at an interval of 10nm. NUS [17] contains 66 HSIs. Like existing works [1, 2, 3, 4, 5, 6, 7], the down-sampling scale is set as 32 to simulate LR HSI.

Two CSS databases are used to test the CSS recovery method. The first database [8] contains 28 CSS curves covering different camera types and brands. The second database [18] contains 12 CSS curves. Both databases cover the wavelength range from 400nm to 700nm at an interval of 10nm. We randomly separate these 40 CSS curves into training dataset with 27 CSSs and testing dataset with 13 CSSs. The training CSS Curves are used to learn sparse representation dictionary. The illuminations we use are from CIE standard illuminants [19].

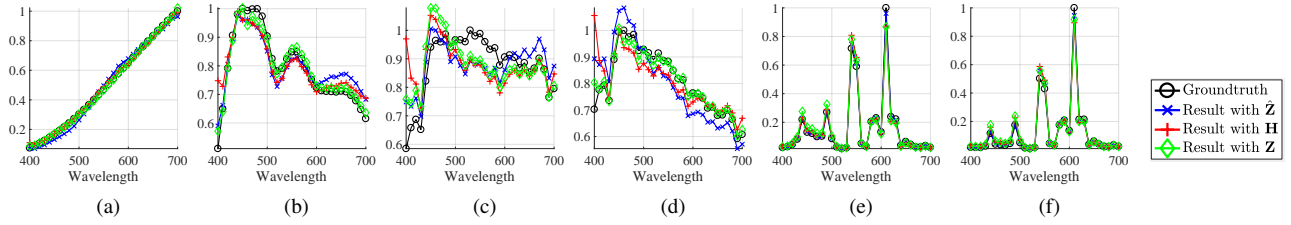


Fig. 2. Illumination separation results of a typical image (faked and real lemon slices) in the CAVE dataset. (a)(b)(c)(d)(e)(f) show estimations of Illumination A, C, D55, D65, F11, F12 with $\hat{\mathbf{Z}}$, \mathbf{H} and \mathbf{Z} , respectively.

CAVE		NUS	
proposed	baseline	proposed	baseline
0.0516	0.0515	0.0310	0.0308

Table 3. The results of reflectance restoration of two methods on the CAVE and NUS databases. The RMSE numbers are the average on all images under all illuminations.

4.2. Quantative Evaluation

Table 2 shows that this hybrid camera system can effectively estimate the CSS, compared with the RGB image based method [8]. Furthermore, our method estimates the RGB camera CSS under unknown illumination and the compared method works under known illumination [8]. It means that the LR HSI in the capture system can effectively assist RGB camera CSS recovery. Detailed results are attached in supplementary material.

In CSS estimation, the method in Equation (6) is set as baseline. The method in [8] estimates CSS from a single RGB image with color checker in the scene. To compared with [8], we test compared methods and our method in two HSIs containing the color checkers in CAVE database. Here, root mean square error (RMSE) is utilized to evaluate the performance on CSS, illumination spectra and spectral reflectance recovery. A smaller value of this metric means better performance.

To simulate real scenarios, we add Gaussian noise $\sigma = 0.5$ into \mathbf{M} , and set $\lambda = 0.1$ in (6). From Tables 1 and 2, we can see that the proposed method outperforms the baseline and [8] for almost all illuminations, and the proposed method is robust under strong noise.

We carry out illumination separation in two ways: separate illumination on LR HSI \mathbf{H} and on the restored HR HSI $\hat{\mathbf{Z}}$. We also apply the algorithm onto the original HR HSI \mathbf{Z} for comparison. The results on $\hat{\mathbf{Z}}$ are almost as good as on \mathbf{Z} . The results on \mathbf{H} are relatively unstable among different kinds of illumination, because of the downsampling operation. As shown in Figure 2, we can obtain similar conclusion. More detailed experiment results are attached in supplementary material.

In HR spectral reflectance spectra restoration, we use two different CSS curves derived by the proposed sparse representation based method and the baseline, respectively. The results are shown in Table 3. We can see that two methods have almost equal performance, although the CSS derived by

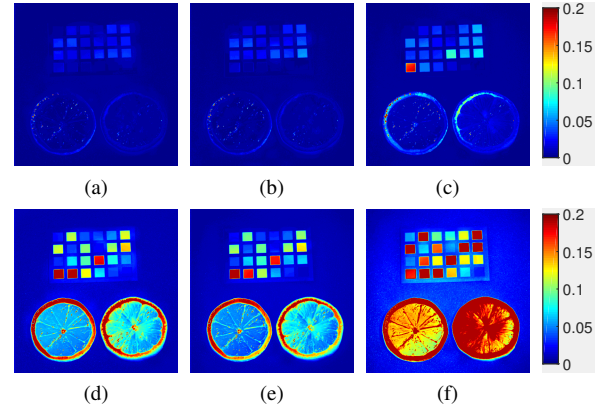


Fig. 3. Error maps of irradiance and reflectance restoration of a typical image in CAVE dataset under illumination D65 simulated by using the KODAK DCS 420 camera. (a)(b)(c)Error maps of restored HR irradiance with CSS estimated by the proposed method, the baseline and [8]. (d)(e)(f)Error maps of restored HR reflectance.

baseline is less precise. The reason is that the accuracy of restored spectral reflectance mainly depends on Equation (5) for specific scene under specific illumination. We also compare these two methods with [8] on the scene with color checkers. Figure 3 shows the error maps of irradiance and reflectance restoration for a scene with color checkers. We can see that our method under unknown illumination outperforms [8] under known illumination.

5. CONCLUSION

This paper have explored the potential capability of a hybrid camera system for other important tasks than hyperspectral fusion. We have first developed a RGB camera spectral sensitivity recovery method for the hybrid camera system. The illumination recovery spectrum and the high spatial resolution spectral reflectance have been simultaneously recovered as well, which are of great importance for such task as material identification and scene relighting.

The hyperspectral camera spectral sensitivity is assumed to be calibrated in this work. Although this calibration is almost effortless, it is worth investigating how to recover it simultaneously, while decoupling illumination and reflectance.

6. REFERENCES

- [1] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2011, pp. 2329–2336.
- [2] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *European Conference on Computer Vision*. Springer, 2014, pp. 63–78.
- [3] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3631–3640.
- [4] H. Kwon and Y.-W. Tai, "RGB-Guided Hyperspectral Image Upsampling," in *IEEE International Conference on Computer Vision*, 2015, pp. 307–315.
- [5] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *IEEE International Conference on Computer Vision*, 2015, pp. 3586–3594.
- [6] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337–2352, 2016.
- [7] N. Akhtar, F. Shafait, and A. Mian, "Hierarchical beta process with gaussian process prior for hyperspectral image super resolution," in *European Conference on Computer Vision*. Springer, 2016, pp. 103–120.
- [8] J. Jiang, D. Liu, J. Gu, and S. Ssstrunk, "What is the space of spectral sensitivity functions for digital color cameras?," in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013, pp. 168–179.
- [9] T. H. Pham, F. Bevilacqua, T. Spott, J. S. Dam, B. J. Tromberg, and S. Andersson-Engels, "Quantifying the absorption and reduced scattering coefficients of tissue-like turbid media over a broad spectral range with non-contact fourier-transform hyperspectral imaging," *Applied optics*, vol. 39, no. 34, pp. 6487–6497, 2000.
- [10] Y. Fu, A. Lam, I. Sato, T. Okabe, and Y. Sato, "Separating reflective and fluorescent components using high frequency illumination in the spectral domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 965–978, May 2016.
- [11] Y. Fu, A. Lam, I. Sato, T. Okabe, and Y. Sato, "Reflectance and fluorescence spectral recovery via actively lit rgb images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1313–1326, July 2016.
- [12] Y. Zheng, I. Sato, and Y. Sato, "Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1779–1787.
- [13] S. Han, I. Sato, T. Okabe, and Y. Sato, "Fast Spectral Reflectance Recovery Using DLP Projector," *International Journal of Computer Vision*, pp. 1–13, Dec. 2013.
- [14] S. Tominaga and R. Okajima, "Object recognition by multi-spectral imaging with a liquid crystal filter," in *International Conference on Pattern Recognition*, 2000, vol. 1, pp. 708–711.
- [15] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral imaging using multiplexed illumination," in *IEEE International Conference on Computer Vision*. IEEE, 2007, pp. 1–8.
- [16] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–53, 2010.
- [17] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-based spectral reconstruction from a single rgb image," in *European Conference on Computer Vision*, 2014, pp. 186–201.
- [18] R. Kawakami, H. Zhao, R. T. Tan, and K. Ikeuchi, "Camera spectral sensitivity and white balance estimation from sky images," *International Journal of Computer Vision*, vol. 105, no. 3, pp. 187–204, 2013.
- [19] R. W. G. Hunt, *Measuring colour*, WILEY, 2011.