

# MULTISPECTRAL FOCAL STACK ACQUISITION USING A CHROMATIC ABERRATION ENLARGED CAMERA

Qian Huang<sup>1</sup>, Yunqian Li<sup>1</sup>, Linsen Chen<sup>1</sup>, Xiaoming Zhong<sup>2</sup>, Jinli Suo<sup>3</sup>, Zhan Ma<sup>1</sup>, Tao Yue<sup>1</sup>, Xun Cao<sup>1</sup>

<sup>1</sup>School of Electronic Science and Technology, Nanjing University, China

<sup>2</sup>Beijing Institute of Space Mechanics and Electricity, China

<sup>3</sup>Department of Automation, Tsinghua University, China

## ABSTRACT

Capturing more information, e.g. geometry and material, using optical cameras can greatly help the perception and understanding of complex scenes. This paper proposes a novel method to capture the spectral and light field information simultaneously. By using a delicately designed chromatic aberration enlarged camera, the spectral-varying slices at different depths of the scene can be easily captured. Afterwards, the multispectral focal stack, which is composed of a stack of multispectral slice images focusing on different depths, can be recovered from the spectral-varying slices by using a Local Linear Transformation (LLT) based algorithm. The experiments verify the effectiveness of the proposed method.

**Index Terms**— Multispectral light field acquisition, chromatic aberration, local linear transformation

## 1. INTRODUCTION

To perceive and understand more complex scenes, researchers attempt to capture more information, e.g. multispectral and light field, of scenes. Traditional cameras take pictures by using 2D sensors with Bayer filter arrays. The captured trichromatic images are 2D projections of 3D scenes with three color channels, i.e. red, green and blue. With the development of optics and computational photography, both multispectral and light field acquisition have been widely explored to capture more spectral channels or light field information of scenes.

In the past decades, several spectral imaging methods were proposed to capture color images with more spectral channels than traditional trichromatic photography. Generally, according to the system architectures, existing multispectral cameras can be divided into several types, e.g., scanning based spectrometers [1], filter-based spectrometers[2],

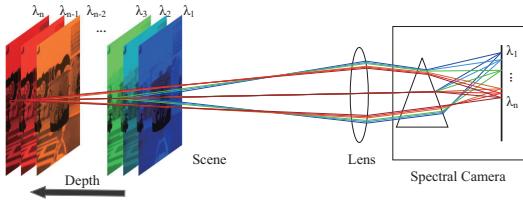
We would like to acknowledge funding from NSFC Projects 61671236, 61631009, 61571215, 61422107, 61371166 and 61627804, National Science Foundation for Young Scholar of Jiangsu Province, China (Grant No. BK20160634 and No. BK20140610), along with the National key foundation for exploring scientific instrument No.2013YQ140517.

Coded Aperture Snapshot Spectral Imager(CASSI) [3], Computed Tomography Imaging Spectrometer(CTIS) [4], and Prism-Mask Video Imaging Spectrometer(PMVIS) [5].

Besides spectral information, the light field (or equivalently, scene depth) is also an important clue for many tasks in computer vision and graphics. Recently, to capture the light field, several methods, e.g., microlens array based method [7], multi-camera array based methods [8] and focal stack based methods [9] are proposed. (Note that the focal stack is one type of representations of the light field, thus in this paper we use the terms light field and focal stack indiscriminately.) Confocal laser scanning microscope(CLSM) [6] can acquire the microscopic focal stack using the scanning scheme. As for snapshot depth acquisition, the chromatic information is explored for extracting depth from color (RGB) images [12][13][14][15]. Besides, Time of Flight (ToF) camera [10] and coded illumination [11] camera were presented to capture the depth directly, so that the light field can be derived by model based rendering.

However, capturing both spectral and light field information together is difficult for too many anisotropic data need to be measured. To solve this problem, this paper proposes a delicately designed camera system, which tries to enlarge the chromatic aberration of the lens while eliminating its rest aberrations (e.g., spherical aberration, coma aberration and astigmatism). By using the proposed camera, the light rays with different wavelengths from slices at different depths of scenes can focus on the same imaging plane. Thus by dispersing the light incidented into the sensor plane into different spectral channels, the images focusing at different depths can be separated. Placing a multispectral imager at the sensor plane, we can capture a multispectral image whose channels are focused at different depths. In other words, we drive a spectral-varying focal stack whose slices are of different spectral channels. Then, a Local Linear Transformation (LLT) based algorithm is present to reconstruct the multispectral focal stack, which contains full information of multispectral light field. Our optical system model can be simplified in Fig. 1.

After acquiring the spectrally varying focal stack, we pro-



**Fig. 1.** Diagram of the proposed multispectral light field acquisition method.

posed to transfer spectral information between different slices and fill up the vacant channels of each slice to reconstruct the multispectral focal stack. In this paper, inspired by Local Linear Transform (LLT) method introduced by Yue *et al*[16], we develop a new method called the local linear transformation (LLT) to transfer the spectral information between channels. Specifically, we impose a strong global blur(e.g. Gaussian) on all the captured channels to remove the different high frequency between these slices caused by different focusing planes. There exists a reasonable linear mapping, namely Local Linear Transformation (LLT), between any two blurred channels and the transformation is also valid for the corresponding sharp slices. In order to extract the LLT mappings, the gradient descent based algorithm is applied.

In all, the main contributions of this paper are as follows: (a) a simple chromatic aberration enlarged camera design for multispectral light field acquisition; (b) a local linear transformation(LLT) based reconstruction method for effectively and efficiently reconstructing the multispectral light field.

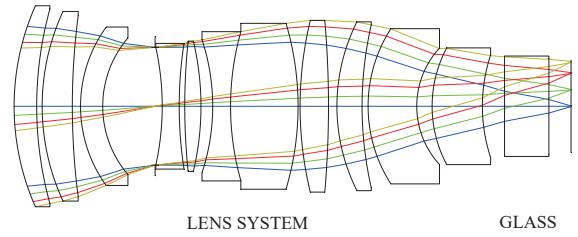
## 2. PROPOSED METHOD

### 2.1. Optical System Design

In this paper, we propose to design an optical system to focus on planes of different depths of the scene with different spectral channels, so that we can capture the spectrally varying focal stack by using a multispectral imager at once. According to this idea, we want to enlarge the camera's chromatic aberration to enlarge the focusing range while eliminating the rest optical aberrations, including piston, tilt, defocus, spherical aberration, coma, astigmatism, field curvature and image distortion, to get promising image quality of all the channels. Here, instead of starting from the scratch to design the lens, we propose a more simple method, i.e. adding a cubic glass behind a well-designed lens (all the aberrations are well corrected), as shown in Fig. 2. With this specially designed lens set, the imaging system has different focal planes with different depths in the scene. The simulation results with ZEMAX is fundamentally consistent with our theory, as shown in Fig. 3. By selecting the central wavelengths of spectral channels ranging from 430nm to 700nm, the corresponding slices between plane 1 and plane 3 in Fig. 3 can be captured. In our paper, ten spectral channels are selected with equal intervals between the central wavelengths, so that the corre-

sponding slices have non-uniform depth intervals.

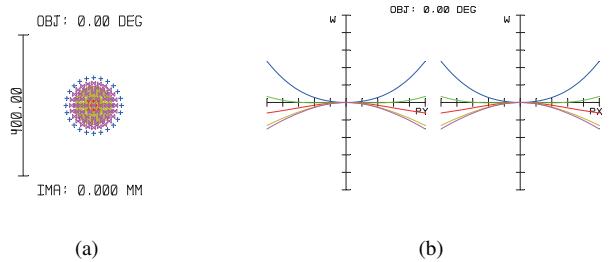
Fig. 4(a) presents the spots pattern of a certain point with different wavelengths. In the shown section of spot diagram, the red rays focus nearly on the best image plane because its RMS(Root Mean Square) radius is the smallest, while other wavelengths from the same depth in the scene are not. This is a result of chromatic aberration. More specifically, the longer the wavelength is, the smaller the RMS radius of the will be, which is especially obvious for the blue and red one in the figure. Fig. 4(b) represents the optical path difference(OPD) of the system. Taking the object at the spindle for instance still, the disparity(i.e. the OPD of different wavelengths) becomes apparent among different wavelengths as the entrance pupil scalar(i.e. abscissa) increases, which shows the enhanced chromatic aberration indirectly. In other words, when the light deviates the spindle of lens set, the real convergence points of different spectra become more distant spatially. Note that different colors in Fig. 4 represent different wavelengths varying from 430nm to 700nm.



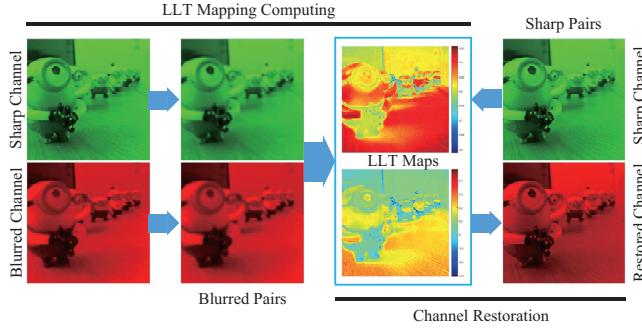
**Fig. 2.** The specially designed lens set. (Color rays by: Fields)



**Fig. 3.** Different focal planes at different depths in the scene. (Color rays by: Wavelengths) Object planes marked 1, 2, 3 denote depths of  $d_1$ ,  $d_2$ ,  $d_3$  and wavelengths of 400nm, 550nm, 700nm respectively.



**Fig. 4.** Diagram of optical aberrations. (a)Section of the Spot diagram (b)Section of the OPD.



**Fig. 5.** Flowchart of our multispectral focal stack reconstruction.

## 2.2. Multispectral Focal Stack Reconstruction

Here, we present the proposed Local Linear Transformation (LLT) based multispectral focal stack reconstruction algorithm. As shown in Fig. 5, we capture one channel at a single depth, which is treated as the sharp channel in LLT algorithm. The blurred channel is derived by blurring the captured image computationally. By computing the LLT maps  $\mathbf{A}_{i,k}$  and  $\mathbf{B}_{i,k}$ , we can restore the missing channels by channel transferring.

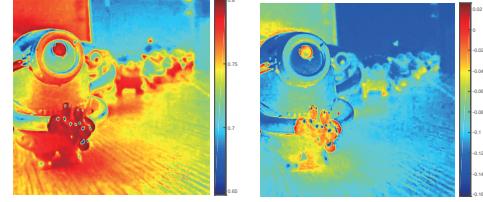
According to Local Linear Transformation (LLT) property introduced by Yue *et al.*[16], in a local area with the same blur effect, the pixel values of different channels follow a certain linear transformation, which is valid in the same area for pixels of the sharp version of those channels as well. In our scenario, the defocus blur varies in both spatial and channel dimensions. It is no trivial to apply the LLT in this case directly, since there is no obvious blur-sharp pair in our application. Therefore, the LLT is proposed to transfer the spectral information while keeping the blur patterns which imply depth information of the scene.

For any two slices, which are captured with different wavelengths and different local blur effects caused by different focusing distances, we apply a strong *Gaussian Kernel* to blur both of them to remove the different high frequency information, so that the blurred images can be regarded as uniformly blurred with a large Gaussian kernel. Thus, the two slices are of the same blur pattern and the LLT maps can be computed from these blurred slice pairs. Given the LLT Maps, the full-channel focal stacks can be restored by linearly transforming between different channels of the captured spectral-varying focal stack.

Specifically, we blur all the channels with a large Gaussian Kernel ( $\mathbf{G}_\sigma$ ) at first. Empirically, the standard variation  $\sigma = 10$  is strong enough for images captured in practice. Then, according to LLT property [16], we know there exist local linear transforms between blurred channels and the original channels. The relationship can be described as follows:

$$\mathbf{I}_{\lambda_i} = \mathbf{A}_{i,k} \odot \mathbf{I}_{\lambda_k} + \mathbf{B}_{i,k}, \quad (1a)$$

$$\mathbf{L}_{d_k, \lambda_i} = \mathbf{A}_{i,k} \odot \mathbf{L}_{d_k, \lambda_k} + \mathbf{B}_{i,k}, \quad (1b)$$



**Fig. 6.** An example of LLT map pair  $\mathbf{A}_{i,k}$ (left) and  $\mathbf{B}_{i,k}$ (right)

where  $\mathbf{I}_{\lambda_i}$  and  $\mathbf{I}_{\lambda_k}$  are two blurred channels, which are computed by blurring the original sharp slice pair  $\mathbf{L}_{d_i, \lambda_i}$  and  $\mathbf{L}_{d_k, \lambda_k}$ . Since the Gaussian blur kernel applied here are very large, the blurred images can be regarded as uniformly blurred. Thus  $\mathbf{I}_{\lambda_i}$  and  $\mathbf{I}_{\lambda_k}$  do not contain the depth information, so we remove the subscript  $d_i$  and  $d_k$  here.  $\mathbf{L}_{d_k, \lambda_i}$  and  $\mathbf{L}_{d_k, \lambda_k}$  are the corresponding sharp channels.  $\mathbf{A}_{i,k}$ ,  $\mathbf{B}_{i,k}$  are the local linear transformation maps.  $d_k$  and  $\lambda_i$  represent the depth in the scene and corresponding wavelength separately.  $\odot$  means element-wise multiplication of matrices.

To compute LLT maps  $\mathbf{A}_{i,k}$  and  $\mathbf{B}_{i,k}$ , an objective function is introduced:

$$\begin{aligned} \min E = & \| \mathbf{A}_{i,k} \odot \mathbf{I}_{\lambda_k} + \mathbf{B}_{i,k} - \mathbf{I}_{\lambda_i} \|_2^2 \\ & + \alpha \| \mathbf{A}_{i,k} \odot \nabla \mathbf{I}_{\lambda_k} - \nabla \mathbf{I}_{\lambda_i} \|_2^2 \\ & + \beta (\| \nabla \mathbf{A}_{i,k} \|_2^2 + \| \nabla \mathbf{B}_{i,k} \|_2^2), \end{aligned} \quad (2)$$

where  $\mathbf{I}_S$  and  $\mathbf{I}_B$  separately represents the original sharp and blurred channel.  $\alpha$  and  $\beta$  are the weights of constraint terms, and are set to 1 and 0.1 according to [16].  $\nabla$  is the gradient operator.

The traditional Gradient Descent method is applied to optimize Eq. 2. Specifically, the derivatives of Eq. 2 can be computed respectively by

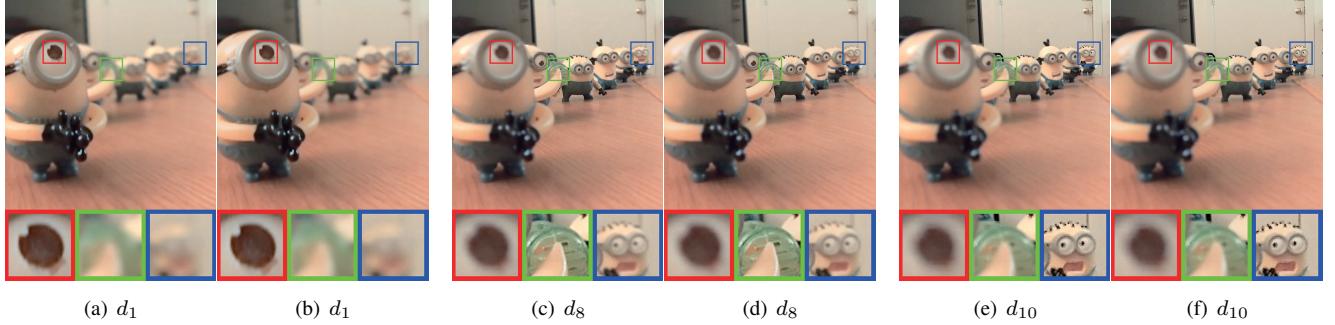
$$\begin{aligned} g_A = & 2\mathbf{I}_S \odot (\mathbf{A}_{i,k} \odot \mathbf{I}_{\lambda_k} + \mathbf{B}_{i,k} - \mathbf{I}_{\lambda_i}) \\ & + 2\alpha \nabla \mathbf{I}_S \odot (\mathbf{A}_{i,k} \odot \nabla \mathbf{I}_{\lambda_k} - \nabla \mathbf{I}_{\lambda_i}) + 2\beta \nabla^T \nabla \mathbf{A}_{i,k} \quad (3) \\ g_B = & 2(\mathbf{A}_{i,k} \odot \mathbf{I}_{\lambda_k} + \mathbf{B}_{i,k} - \mathbf{I}_{\lambda_i}) + 2\beta \nabla^T \nabla \mathbf{B}_{i,k}, \end{aligned}$$

By iteratively searching along the gradient directions given by Eq. 3,  $\mathbf{A}_{i,k}$  and  $\mathbf{B}_{i,k}$  can be derived. Fig. 6 shows an example of LLT map pair  $\mathbf{A}_{i,k}$  and  $\mathbf{B}_{i,k}$ .

With  $\mathbf{A}_{i,k}$  and  $\mathbf{B}_{i,k}$ , we can transfer the spectral information of  $I_{d_i, \lambda_i}$  to  $I_{d_j, \lambda_j}$ . By transforming the spectral information of all the channels to a certain channel  $I_{d_i, \lambda_i}$  a multispectral slice focusing on depth  $d_i$  is recovered. Similarly, the rest slices of the focal stack can be reconstructed by transferring the spectral information to the rest spectral-varying slices respectively.

## 3. EXPERIMENTAL RESULT

We test the proposed chromatic aberration enlarged camera and LLT-based reconstruction algorithm on the multispectral focal stacks synthesized from both on-line dataset, i.e. LFSD



**Fig. 7.** Comparison between ground-truth images(a),(c),(e) and our restored images(b),(d),(f).

[9], and real captured images. The focal stacks in the dataset are composed of ten RGB slices focused on different depths. In order to obtain the multispectral images, we synthesize the pseudo spectra from the RGB measurement by using the training based algorithm [17]. In our experiment, we use ten slices with spectral channels with central wavelengths 430nm, 460nm, ..., 700nm as the input, the corresponding slices are denoted by  $d_1, d_2, \dots, d_{10}$ . The ground truth is the full-channel focal stack composed of two multispectral slices, and each of the slices has ten spectral channels. In the experiment, we select a single channel of each multispectral slice to simulate our focal stack camera. The entire multispectral focal stacks are restored by using the proposed LLT-based reconstruction algorithm.

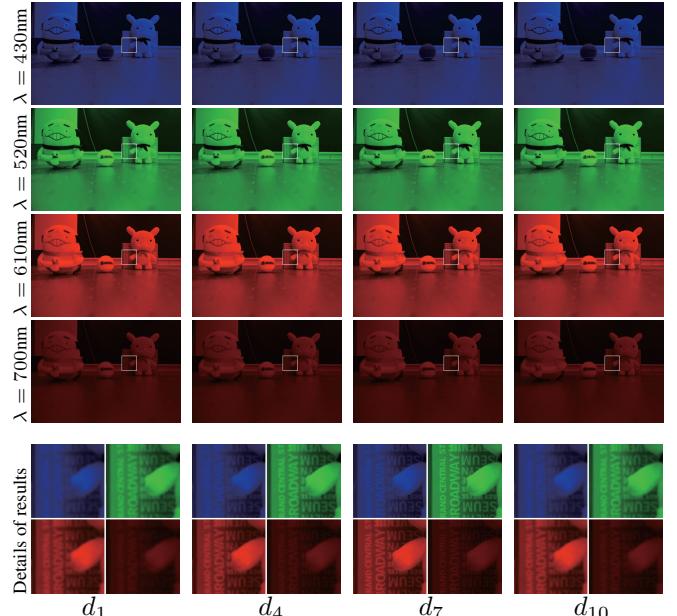
To quantitatively evaluate the performance, the peak signal-to-noise ratio(PSNR) and the structure similarity(SSIM) are employed. We also present reconstructed images to demonstrate the performance qualitatively.

**Table 1.** Quantitative evaluation for four selected depths and channels of our results.

	$\lambda = 430\text{nm}$		$\lambda = 520\text{nm}$		$\lambda = 610\text{nm}$		$\lambda = 700\text{nm}$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
$d_1$	Inf	1.0000	34.80	0.9799	35.08	0.9811	35.70	0.9841
$d_4$	39.79	0.9768	Inf	1.0000	37.76	0.9829	35.26	0.9729
$d_7$	42.27	0.9867	36.28	0.9808	Inf	1.0000	42.26	0.9959
$d_{10}$	38.20	0.9801	30.44	0.9684	33.55	0.9859	Inf	1.0000

**Quantitative evaluation.** The average quantitative measurements of restored images are shown in Tab. 1. It is obvious that the proposed method can achieve promising performance in terms of both PSNR and SSIM metrics.

**Qualitative evaluation.** We also show the results of qualitative evaluations. Fig. 7 shows the comparisons on synthetic data between ground truth and our results side by side. To facilitate comparison, we compute the RGB color images from the ground truth and our recovered multispectral slices. It is obvious that the recovered results are very similar to the ground truth. Besides, we also test the method on real captured data. Fig. 8 is an example of our reconstructed multispectral focal stack on the real captured images. The channels of our results at different depths (i.e.  $d_1, d_4, d_7$  and  $d_{10}$ ) are



**Fig. 8.** The experimental multi-spectral focal stack result on real-captured scene Top 4 rows: selected input at depth  $d_1, d_4, d_7, d_{10}$  with spectral wavelength 430nm, 520nm, 610nm, 700nm. Bottom: the details of restored results.

shown in Fig. 8. Each row represents a selected wavelength, i.e. 430nm, 520nm, 610nm and 700nm, and the close-ups are shown in the bottom of each figure. From the results, we can see that the proposed method performs well on both fine details and smooth areas.

#### 4. CONCLUSION AND DISCUSSION

In this paper, we have proposed a chromatic aberration enlarged camera and an LLT-based reconstruction algorithm for acquiring multispectral focal stacks. The proposed method achieves promising performance in terms of both quantitative and qualitative evaluations in our experiments.

Limited by the complexity, the proposed method cannot work in real time at the current stage. We will try to further simplify and optimize the algorithm in the future.

## 5. REFERENCES

- [1] John James, *Spectrograph design fundamentals*, 2007.
- [2] Nahum Gat, “Imaging spectroscopy using tunable filters: a review,” in *AeroSense*, 2000, pp. 50–64.
- [3] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle, “Compressive coded aperture spectral imaging: An introduction,” *IEEE Signal Processing Magazine*, vol. 31, pp. 105–115, 2014.
- [4] Corrie Vandervlugt, Hugh Masterson, Nathan Hagen, and E Dereniak, “Reconfigurable liquid crystal dispersing element for a computed tomography imaging spectrometer,” *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIII*, vol. 6565, pp. 65650O, 2007.
- [5] Xun Cao, Hao Du, Xin Tong, Qionghai Dai, and Stephen Lin, “A prism-mask system for multispectral video acquisition,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, pp. 2423–2435, 2011.
- [6] James B. Pawley, “Handbook of biological confocal microscopy,” *Journal of Biomedical Optics*, vol. 25, no. 3, pp. 029902, 1995.
- [7] Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz, and Pat Hanrahan, “Light field photography with a hand-held plenoptic camera,” *Computer Science Technical Report (CSTR)*, vol. 2, pp. 1–11, 2005.
- [8] Cha Zhang and Tsuhan Chen, “A self-reconfigurable camera array,” in *ACM SIGGRAPH 2004 Sketches*, 2004, p. 151.
- [9] Nianyi Li, Jinwei Ye, Yu Ji, Haibin Ling, and Jingyi Yu, “Saliency detection on light field,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2806–2813.
- [10] Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun, “High-quality scanning using time-of-flight depth superresolution,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2008, pp. 1–7.
- [11] Erhard Schubert, “Fast 3d object recognition using multiple color coded illumination,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1997, vol. 4, pp. 3057–3060.
- [12] Josep Garcia, Juan Maria Sanchez, Xavier Orriols, and Xavier Binefa, “Chromatic aberration and depth extraction,” in *IEEE Proceedings 15th International Conference on Pattern Recognition*, 2000, vol. 1, pp. 762–765.
- [13] Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman, “Image and depth from a conventional camera with a coded aperture,” *ACM transactions on graphics (TOG)*, vol. 26, no. 3, pp. 70, 2007.
- [14] Oliver Cossairt and Shree Nayar, “Spectral focal sweep: Extended depth of field from chromatic aberrations,” in *IEEE International Conference on Computational Photography (ICCP)*, 2010, pp. 1–8.
- [15] Pauline Trouvé, Frédéric Champagnat, Guy Le Besnerais, Jacques Sabater, Thierry Avignon, and Jérôme Idier, “Passive depth estimation using chromatic aberration and a depth from defocus approach,” *Applied optics*, vol. 52, no. 29, pp. 7152–7164, 2013.
- [16] Tao Yue, Ming-Ting Sun, Zhengyou Zhang, Jinli Suo, and Qionghai Dai, “Deblur a blurred rgb image with a sharp nir image through local linear mapping,” in *IEEE International Conference on Multimedia and Expo (ICME)*, 2014, pp. 1–6.
- [17] Rang MH Nguyen, Dilip K Prasad, and Michael S Brown, “Training-based spectral reconstruction from a single rgb image,” in *European Conference on Computer Vision*, 2014, pp. 186–201.