applied optics

Iteratively reconstructing 4D light fields from focal stacks

Xuanwu Yin,[1,2] Guijin Wang,[1,*] Wentao Li,[1] and Qingmin Liao[2]

[1]Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[2]Department of Electronic Engineering, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China
*Corresponding author: wangguijin@tsinghua.edu.cn

Received 24 June 2016; revised 18 September 2016; accepted 19 September 2016; posted 20 September 2016 (Doc. ID 268973); published 14 October 2016

Acquiring and representing the 4D space of rays in the world (the light field) is important for many computer vision and graphics applications. In this paper, we propose an iterative method to acquire the 4D light field from a focal stack. First, a discrete refocusing equation is derived from integral imaging principles. With this equation, a linear projection system is formulated to model the focal stack imaging process. Then we reconstruct the 4D light field from the focal stack through solving the inverse problem with a filtering-based iterative method. The experimental results show that our approach is effective and outperforms state-of-the-art methods in reconstruction accuracy, reduced sampling, and occluded boundaries.

OCIS codes: (100.3020) Image reconstruction-restoration; (110.3010) Image reconstruction techniques; (110.4155) Multiframe image processing.

http://dx.doi.org/10.1364/AO.55.008457

1. INTRODUCTION

Computational photography is redefining the concept of photography [1,2]. A conventional image is simply the 2D projection of a 3D world, while the light field is a powerful concept that describes how light transmits in space on a ray-by-ray basis. The light field, together with computational photography techniques, enables many novel applications such as digital refocusing and parallax [3,4], depth imaging [5,6], and advanced techniques such as glare removal [7].

The light field theory has been developed rapidly over the last decades [8,9]. On the basis of recently developed theories [10,11], many light field acquisition methods have been developed. Among them are complicated setups such as robotic gantries [1] and camera arrays [12,13]. There also are compact designs that utilize multiplexing optics, such as lenslet arrays [3,14], amplitude masks [15], frequency domain multiplexing [16], well-designed mirrors [17,18], and programmable aperture techniques [19]. The common drawback of these designs is that the camera has to be mounted or inserted with additional optical elements. A new type of sensor, called an angle sensitive pixel (ASP) sensor, has been recently developed to acquire the angle information inside the camera [20].

In recent years, focal stack imaging has attracted a lot of attention [21,22]. Researchers have tried to reconstruct a light field from a multifocus image stack. In such a design, there is no need to mount or insert any additional optical elements to the acquisition system, making it possible to capture a light field with off-the-shelf commercial cameras, such as a digital single-lens reflex (DSLR) camera. The reconstructed light field can be used for digital refocusing of the scene, synthesizing apertures of arbitrary shapes and sizes (restricted by the aperture size of the lens), estimating the structure of the scene, and 3D reconstruction. Among these light field reconstruction methods, the most common is the depth-based method, which relies on the complex depth estimation process [23,24]. Recently, deconvolution has been used in extended depth of field (EDOF) photography [25] and fluorescence tomography [26]. Inspired by this, Levin and Durand derived a depth-invariant blur kernel and proposed a deconvolution-based method to reconstruct a light field from a focal stack [27]. However, the deconvolution approach requires dense sampling to build the blur kernel and thus could not give good results when the number of samples is limited. Alonso *et al.* [28] proposed a plane-wise approximation-based method to reconstruct a 4D light field from a multifocus image stack in the frequency domain. This method cannot give reasonable results when the 3D scene could not be approximated by sparse planes.

In this paper, we propose an iterative method to reconstruct a 4D light field from a focal stack. This work is motivated by the observation in [27], which states that 3D measurement sets are sufficient for reconstructing a 4D light field. Our formulation is based on the ray optics geometry presented in Fig. 1, in which the wave nature of light is not taken into account. Based on this imaging geometry, we first derive a discrete light

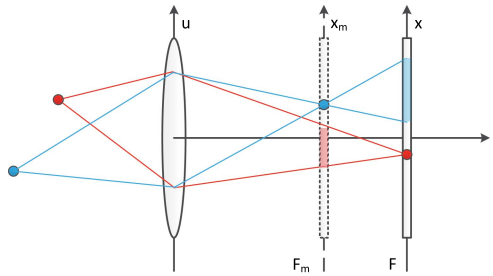

Fig. 1. Focal stack geometry, the red spot and the blue spot are focused at different depth layers.

field refocusing equation, which links the subaperture images and the focal stack. With this equation, we formulate a discrete projection system to model the focal stack imaging process. A filtering-based iterative reconstruction method is proposed to reconstruct the 4D light field. Our method does not rely on the complex nonlinear depth estimation process and thus is not affected by the accuracy of depth estimation. And the filtering approach can make the best use of the structural information contained in the focal stack. Furthermore, this spatial-domain reconstruction technique is more robust to reduced sampling, in which case the number of samples is limited. The experiments demonstrate our method's effectiveness in reconstruction accuracy, reduced sampling, and occluded boundaries.

2. IMAGING MODEL

In this section, we first derive a spatial-domain discrete refocusing equation from the integral imaging principle. Based on this discrete refocusing equation, we formulate a discrete projection system to model the focal stack imaging process. The light field can be reconstructed through solving the inverse system.

In this work, the light field is represented by parameterizing each ray by its intersection with two parallel planes, uv and xy [9], known as the directional and spatial dimensions. Usually the viewpoint (camera aperture) is positioned and shifted on the uv plane (lens plane) and xy on the sensor plane, as shown in Fig. 1.

A. Discrete Refocusing Equation

The irradiance of the recorded image at pixel (x, y) is given by

$$E_F^c(x,y) = \frac{1}{F^2}\iint L_F^c(x,y,u^c,v^c)A^c(u^c,v^c)\cos^4\theta \,\mathrm{d}u^c\,\mathrm{d}v^c, \quad (1)$$

where F is the separation between the lens plane and sensor plane, L_F^c is the light field, A^c is an aperture function (e.g., one within the opening and zero outside it), and θ is the angle between the ray (x, y, u^c, v^c) and the sensor plane. The superscript c means that the expressions are in continuous form. We can simplify this imaging equation with

$$\overline{E}_F^c(x,y) = \iint L_F^c(x,y,u^c,v^c)A^c(u^c,v^c)\,\mathrm{d}u^c\,\mathrm{d}v^c. \quad (2)$$

This means that, if we have captured the light field at F, the imaging process can be done by integrating over the directional dimensions. This equation can be rewritten in discrete form as

$$I_F(x,y) = \frac{1}{N}\sum_u\sum_v L_F(x,y,u,v)A(u,v), \quad (3)$$

where L_F is the discrete light field defined by the sensor plane at F and the lens plane, I_F denotes the image intensity recorded at position F, A is the discrete aperture function, and N is the number of subapertures. Note that in this definition, we change the integration to averaging in order to keep the intensity of L_F in the same range as that of I_F.

By setting A as

$$A(u,v) = \begin{cases} 1, & \text{if } u = u_j, v = v_j \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

we define the subaperture image

$$I_F^{u_j,v_j}(x,y) = L_F(x,y,u_j,v_j)A(u_j,v_j), \quad (5)$$

in which $I_F^{u_j,v_j}$ denotes the subaperture image from a certain viewpoint (u_j, v_j), with j the index of it.

According to [3], if the sensor is moved from F to F_m, the light field defined by the new sensor plane at F_m and the lens plane is given by

$$L_{F_m}(x,y,u,v) = L_F\left(\left(1-\frac{1}{\alpha_m}\right)u + \frac{x}{\alpha_m}, \left(1-\frac{1}{\alpha_m}\right)v + \frac{y}{\alpha_m}, u, v\right), \quad (6)$$

where $\alpha_m = F_m/F$ is a scale factor, which represents the field of view (FOV) change caused by moving the focusing plane. Based on Eqs. (3), (5), and (6), the recorded image at F_m is given by

$$\begin{aligned}\overline{I}_{F_m}(x,y) &= \frac{1}{N}\sum_{u_j}\sum_{v_j} L_{F_m}(x,y,u_j,v_j)A(u_j,v_j) \\ &= \frac{1}{N}\sum_{u_j}\sum_{v_j} I_F^{u_j,v_j}\left(\left(1-\frac{1}{\alpha_m}\right)u_j + \frac{x}{\alpha_m}, \left(1-\frac{1}{\alpha_m}\right)v_j + \frac{y}{\alpha_m}\right) \\ &= \frac{1}{N}\sum_j I_F^j\left(\left(1-\frac{1}{\alpha_m}\right)u_j + \frac{x}{\alpha_m}, \left(1-\frac{1}{\alpha_m}\right)v_j + \frac{y}{\alpha_m}\right), \end{aligned} \quad (7)$$

where $I_F^j = I_F^{u_j,v_j}$, is the jth subaperture image. We correct the FOV difference by scaling the recorded image with the factor α_m:

$$\begin{aligned} I_{F_m}(x,y) &= \overline{I}_{F_m}(\alpha_m x, \alpha_m y) \\ &= \frac{1}{N}\sum_j I_F^j((\alpha_m - 1)u_j + x, (\alpha_m - 1)v_j + y), \end{aligned} \quad (8)$$

where I_{F_m} is the FOV corrected image recorded at F_m. In the rest of this paper, it is assumed that all the images have been FOV corrected.

Equation (8) is the discrete refocusing equation derived from the integral imaging principle. It links the image focused at position F_m and the subaperture images at the reference position F. This spatial formulation is the spatial duality to the frequency domain formulation in [27]. We re-derive it in the

spatial domain in order to formulate the projection below and the iterative reconstruction method.

B. Projection Model

As the shifting-averaging operation in Eq. (8) is linear, Eq. (8) can be interpreted in matrix production form as

$$P_m \begin{bmatrix} I_F^1 \\ \vdots \\ I_F^N \end{bmatrix} = I_{F_m}, \tag{9}$$

where P_m is the projection matrix, also known as the sampling matrix. Thus I_{F_m} can be treated as one sample of the light field. Suppose that there are M slices in the focal stack, we formulate a linear system by putting the M sampling equations together:

$$\underbrace{\begin{bmatrix} P_1 \\ \vdots \\ P_M \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} I_F^1 \\ \vdots \\ I_F^N \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} I_{F_1} \\ \vdots \\ I_{F_M} \end{bmatrix}}_{b}, \tag{10}$$

where M is the number of the samples, each submatrix P_m is relevant to the mth slice of the stack. The matrix P projects the subaperture images (the light field) onto a series of images focused at different positions, forming a focal stack. The light field X can be reconstructed through solving Eq. (10).

3. METHOD

In this section, we formulate an iterative method to solve Eq. (10). In this work, the number of focal slices, M, is smaller than the number of views, N, thus Eq. (10) is an underdetermined problem. Suppose the projection matrix P is with dimension $H \times W$. We modify Eq. (10) by multiplying the same matrix on both sides of it:

$$P^T CPX = P^T Cb, \tag{11}$$

where P^T is the transposition of P, and C is an $H \times H$ filtering matrix. We introduce C to make use of the structural information in the images. Equation (11) can be equivalently transformed into

$$X = X + P^T C(b - PX). \tag{12}$$

We formulate an iterative step based on Eq. (12). Suppose X^q is the solution in the qth round of iteration, X^{q+1} can be found by

$$X^{q+1} = X^q + P^T C(b - PX^q). \tag{13}$$

Define the residual vector as

$$\Delta^q = b - PX^q. \tag{14}$$

Plugging Eq. (14) into Eq. (13), we can derive

$$X^{q+1} = X^q + P^T C\Delta^q. \tag{15}$$

The iteration step, Eq. (15) can be interpreted as (1) calculate the current residual Δ^q; (2) filter the residual with C; (3) back-project the filtered residual with P^T; and (4) add the back-projected residual to X^q.

This iterative strategy cannot be applied straightforward due to the large amount of data. Suppose the light field is with directional dimension 17×17, spatial dimension 640×480, and there are 40 samples in total. The dimension of matrix P is given by $(40 \times 640 \times 480) \times (17 \times 17 \times 640 \times 480) \approx 12.3\ \text{m} \times 88.8\ \text{m}$. The dimension of filtering matrix C is $12.3\ \text{m} \times 12.3\ \text{m}$. It is impossible to store them all in the memory.

To make the iterative method applicable, we define the filtering matrix C as a block diagonal matrix:

$$C = \begin{bmatrix} C_1 & & \\ & \ddots & \\ & & C_M \end{bmatrix}, \tag{16}$$

in which the submatrix C_m is relevant to the mth focal slice. Plugging Eqs. (10) and (16) into Eq. (15), we can derive

$$X^{q+1} = X^q + \sum_m P_m^T C_m \Delta_m^q, \tag{17}$$

where Δ_m^q is the submatrix of Δ^q, which is relevant to the mth focal slice. With the block diagonal definition of C, we can further decompose this iteration step into sub-iterations. Denote the initial vector in round $q+1$ as $X^{q+1,0}$, which is initialized as $X^{q+1,0} = X^q$, and the mth sub-iteration step goes as

$$X^{q+1,m} = X^{q+1,m-1} + P_m^T C_m \Delta_m^q, \tag{18}$$

which means X is updated with the mth focal slice. When the sub-iterations terminate, we set $X^{q+1} = X^{q+1,M}$. The result of the sub-iterations in Eq. (18) easily can be proved to be equal to that in Eq. (17). This observation indicates that Eq. (17) can be calculated in a slice-by-slice way.

To accelerate the convergence, we update the residual vector during the sub-iterations. To be precise, the actual sub-iteration step is

$$X^{q+1,m} = X^{q+1,m-1} + P_m^T C_m \Delta_m^{q+1,m-1}, \tag{19}$$

with the residual vector $\Delta_m^{q+1,m-1} = I_{F_m} - P_m X^{q+1,m-1}$.

The entries of C_m are defined with the content of the mth focal slice as

$$C_m^{(i,j)} = \frac{1}{|w|^2} \sum_{k:(i,j)\in w_k} \left(1 + (I_{F_m}^{(i)} - \mu_m^{(k)})^T (\Sigma_m^k + \epsilon U)^{-1} (I_{F_m}^{(j)} - \mu_m^{(k)})\right), \tag{20}$$

where w_k is a neighborhood of pixel k, and $|w|$ the number of pixels in it. $I_{F_m}^{(i)}$ and $I_{F_m}^{(j)}$ are 3×1 color vectors of pixel i and j, respectively. $\mu_m^{(k)}$ is a 3×1 color vector representing the mean value in w_k. Σ_m^k is a 3×3 covariance matrix in w_k. U is an identity matrix of size 3×3. ϵ is a smoothness parameter. By this definition, if two pixels lie on different sides of an edge, the elements of Σ_m^k will be large. Thus the weight will be small. Otherwise if they lie on the same side, the weight will be large. This property guarantees that the structural information of the in-focus regions is preserved, whereas the noises of the out-of-focus regions are smoothed out.

In each sub-iteration, we adopt the guided filter [29] to efficiently implement the residual filtering other than calculating

and storing the filtering matrix C_m. Besides that, we recalculate the relevant submatrix P_m in each sub-iteration, so that we do not have to store the whole projecting matrix P in the memory. This sub-iteration decomposition approach makes it possible to implement Eq. (15) in a slice-by-slice way, overcoming the data amount problem.

4. EXPERIMENTS AND DISCUSSIONS

In this section, we will present experimental evaluations and discussions about our proposed light field reconstruction method.

We utilize the Stanford light field archive [30] to evaluate the reconstruction performance of our method. This light field archive is well sampled with a 17×17 angular resolution. The following three data sets are taken:

- Truck, a Lego Technic truck with complex geometry.
- Chess, a chessboard with pieces, which is great for demonstrating refocusing.
- Card, a scene with a crystal ball resting on tarot cards. The ball acts as a lens. The cards act as diffuse textured objects at many orientations and depths.

For each of them, a stack with 49 slices is synthesized to evaluate our method.

We also test our method on the data provided in [27] to validate our proposed projection model and reconstruction method on real data. This FOV change of the real data is corrected with the image registration technique [31–33].

We first present the overall performance. The performance on reduced sampling is evaluated next. Followed by the experiment on real data. The performance at occluded boundaries is discussed finally. In all our experiments, the filtering parameters are set as $\epsilon = 0.005$ and w_k a 3×3 window.

A. Overall Performance

We evaluate the overall performance with the synthesized data. In our evaluation, we compare our method with the method [named linear view synthesis (LVS) for short] proposed in [27] and the plane-wise approximation (PWA) based method proposed in [28]. In our experiments, we find that the PWA-based method cannot give reasonable results for a stack with more than four slices. As analyzed in [22], this is because in this method each frequency component is reconstructed by solving a linear equation, and these equations will become highly ill-posed when the number of images increases. Thus for each data, we select four from the 49 slices that can cover most of the focusing range as the input of the PWA-based method.

The convergence curve is presented in Fig. 2. The reconstruction accuracy is quantified by structural similarity (SSIM) [34], as SSIM is treated as a better metric to measure the similarity between images than peak signal-to-noise ratio or mean square error (MSE). The reconstruction error will increase after some iterations due to noise and quantization error. In our experiment, we terminate the iteration when the reconstructed light field does not change significantly (MSE < 0.001) or the maximum number of iterations, which is set as six in this work, is reached. Our CPU-only MATLAB implementation without parallel optimization costs some tens of minutes to reconstruct one entire light field, depending on the dimension of the light field and the number of samples. The visual reconstruction results are presented in Fig. 3, including both the central view and a side view.

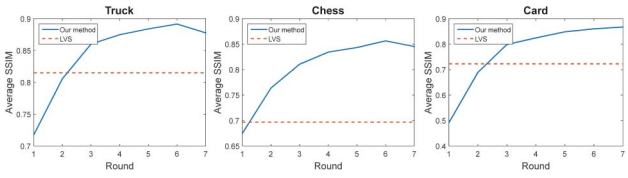

Fig. 2. Convergence curve in average SSIM of all the views.

From Fig. 3, it can be derived that our method can give better reconstruction results than the LVS- and PWA-based methods. For the central views, both our method and LVS can give good results, the PWA-based method can give reasonable results except for the Card data. Our method shows best visual effects of the three methods, which can be seen in the close-up views of local details. For side views, the results of our method show significantly higher reconstruction accuracy and better visual effects than those of the other two methods. It also can be seen from the Truck and Card data that our method can give better results at strong edges than the LVS- and PWA-base methods.

B. Reduced Sampling

The way in which the focal stack is captured will influence the reconstruction result. A focal stack is optimal if the union of the depth of fields of all the images can cover the range of the objects in the scene. In some cases, it is difficult to capture an optimal focal stack. We evaluate our proposed method on a nonoptimal focal stack in a reduced sampling situation, in which the number of samples is limited. In the optimal sampling case, the focal depths are sampled densely. Each object will be focused in a certain focal slice. However, in the reduced sampling case, this is no longer true. An object may appear blurred in any of the focal slices, which will yield nonperfect reconstruction. We have compared the reconstruction results on downsampling factor k as 1, 2, 4, and 8 on Card data, which means that the number of samples will be reduced to 1, 1/2, 1/4, and 1/8 of the original sampling number, respectively. The PWA-based method is not taken into comparison because it cannot give reasonable results for more than four images in our experiment. The visual and quantified results of Card data are presented in Fig. 4.

From the results, it can be seen that the proposed method outperforms the LVS method in case the number of samples is limited. That is because the LVS's depth-invariant blur kernel approximation requires relatively dense sampling, whereas in our method there is no such requirement, which allows sparser focal depth sampling.

C. Real Data

To validate that our proposed projection model and reconstruction method are effective for real data, we test our method on the data provided in [27]. The same stopping criteria as in the simulation are used in this real data test. The reconstructed subaperture images are presented in Fig. 5.

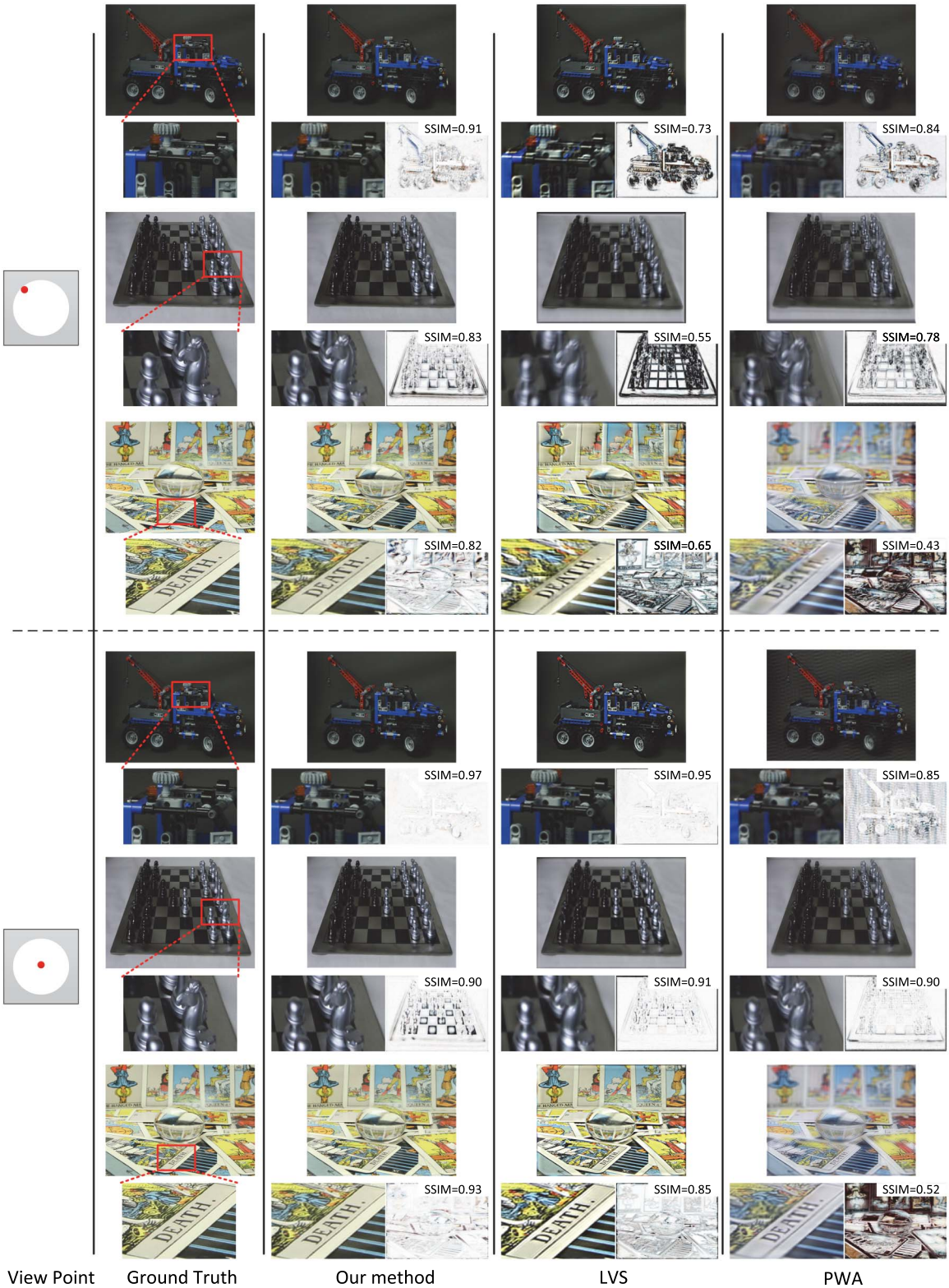

Fig. 3. Experiments on simulated data, reconstructed views of our method, LVS, and PWA. Below left: close-up view of the same patch as in the ground truth column. Below right: SSIM map of the whole content and the average SSIM value. The reconstructed light field of Chess data can be viewed dynamically in Visualization 1.

From the close-up views, it can be seen that our proposed method can effectively reconstruct disparities, i.e., perspective shifts, from the real focal stack.

D. Occlusion

Another problem that the reconstruction models have to face is occlusion. The LVS model fails at occlusion boundaries, as

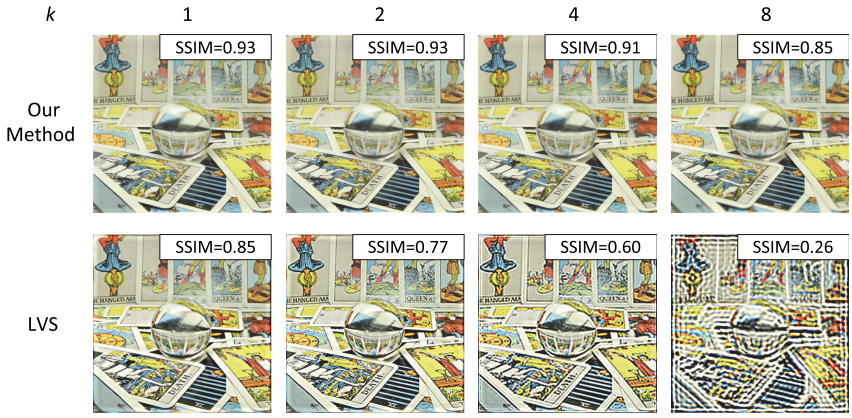

Fig. 4. Reconstruction accuracy with reduced number of samples, measured on the central view of Card data in SSIM.

stated in [27]. We also test our proposed method at occlusion boundaries with the same data as [27]. The results are presented in Fig. 6. It can be seen that there are less artifacts at boundaries with the proposed method than with the LVS method. That's because in our proposed method, the artifacts can be reduced due to the filtering approach.

E. Discussions

The reconstructed light field is actually the light field inside the camera and is influenced by the aperture size, number of samples, and the accuracy of the imaging model.

The viewpoints of the reconstructed light field are restricted by the aperture, i.e., we can only reconstruct the viewpoint inside the aperture area [27]. A large aperture will yield large perspective shifts, whereas a small aperture will yield small perspective shifts.

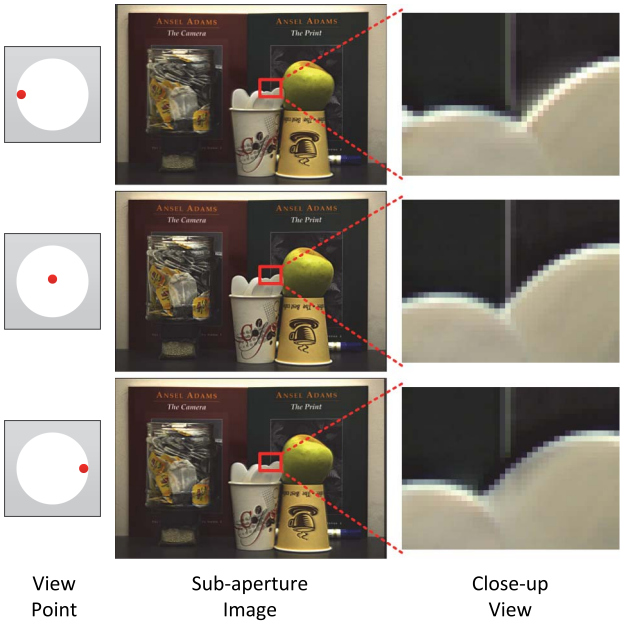

Fig. 5. Reconstructed subaperture images (Visualization 2) from the focal stack provided in [27]. The close-up views show the reconstructed disparities (perspective shifts) within the red rectangle.

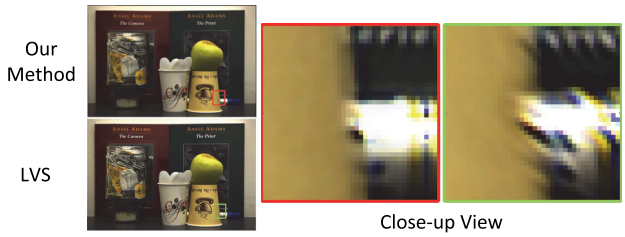

Fig. 6. Artifacts at occlusion boundaries and close-up views of details (left is our method and right is LVS).

A focal stack is optimal if the union of the depth of fields of all the images can cover the range of the objects in the scene. The minimum number of samples required to capture an optimal focal stack depends on the aperture size. A large aperture will make the depth of field of each focal slice small; a large number of samples are required to cover the range of the scene. Whereas a small aperture will make the depth field of each focal slice large, a small number of samples are required to cover the range of the scene. More analysis on how to capture an optimal focal stack can be found in [28]. The experiments in Section 4.B demonstrate that our proposed method is more robust to nonoptimal sampling than state-of-the-art methods.

Note that the imaging model is formulated based on the ideal ray optics geometry. Thus the lens aberrations, especially the distortions, will make the imaging model inaccurate. That is the reason why the central views are always reconstructed better than side views.

5. CONCLUSION

In this paper, we formulate a linear projection system to model the focal stack imaging process and propose a filtering-based iterative method to reconstruct 4D light fields from focal stacks. Compared with state-of-the-art methods, the proposed method shows better performance in reconstruction accuracy and visual effects. Additionally, the proposed filtering approach suffers less from the occluded boundaries than the LVS method. The potential future work may include precise focal stack

calibration approach, more efficient sampling strategy, depth estimation, and so on.

Funding. National Natural Science Foundation of China (NSFC) (61327902); 863 Program (2015AA016304).

REFERENCES

1. M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of SIGGRAPH* (ACM, 1996), pp. 31–42.
2. S. J. Gortler, R. Grzeszczuk, R. Szelinski, and M. F. Cohen, "The lumigraph," in *Proceedings of SIGGRAPH* (ACM, 1996), pp. 43–54.
3. R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Technical Report CSTR 2005–02 (Stanford University Computer Science, 2005).
4. F. Nava, J. G. Marichal-Hernandez, and J. M. Rodriguez-Ramos, "The discrete focal stack transform," in *16th European Signal Processing Conference*, Lausanne, Switzerland, 2008, pp. 1–5.
5. C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and S. Gross, "Scene reconstruction from high spatio-angular resolution light fields," ACM Trans. Graph. **32**, 73:1–73:12 (2013).
6. S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," IEEE Trans. Pattern Anal. Mach. Intell. **36**, 606–619 (2014).
7. R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan, "Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses," ACM Trans. Graph. **27**, 56 (2008).
8. G. Wetzstein, I. Ihrke, D. Lanman, W. Heidrich, R. Raskar, and K. Akeley, "Computational plenoptic imaging," Comput. Graph. Forum **30**, 2397–2426 (2011).
9. C. Zhou and S. K. Nayar, "Computational cameras: convergence of optics and processing," IEEE Trans. Image Process. **20**, 3322–3340 (2011).
10. R. Ng, "Fourier slice photography," ACM Trans. Graph. **24**, 735–744 (2005).
11. C. Hahne, A. Aggoun, S. Haxha, V. Velisavljevic, and J. C. Fernández, "Light field geometry of a standard plenoptic camera," Opt. Express **22**, 26659–26673 (2014).
12. B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. R. Antúnez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," ACM Trans. Graph. **24**, 765–776 (2005).
13. X. Lin, J. Wu, G. Zheng, and Q. Dai, "Camera array based light field microscopy," Biomed. Opt. Express **6**, 3179–3189 (2015).
14. T. Georgiev, K. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala, "Spatio-angular resolution tradeoff in integral photography," in *Proceedings of the 17th Eurographics Conference on Rendering Techniques (EGSR)*, 2006, pp. 263–272.
15. A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing," ACM Trans. Graph. **26**, 69 (2007).
16. T. Georgiev, C. Intwala, and D. Babacan, "Lightfield capture by multiplexing in the frequency domain," in *Adobe Technical Report* (Adobe Systems, 2007).
17. M. Fuchs, M. Kächele, and S. Rusinkiewicz, "Design and fabrication of faceted mirror arrays for light field capture," Comput. Graph. Forum **32**, 246–257 (2013).
18. A. Manakov, J. Restrepo, O. Klehm, R. Hegedüs, E. Eisemann, and H. S. Ivoihrke, "A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging," ACM Trans. Graph. **32**, 47:1–47:14 (2013).
19. C. Liang, T. Lin, B. Wong, C. Liu, and H. H. Chen, "Programmable aperture photography: multiplexed light field acquisition," ACM Trans. Graph. **27**, 55 (2008).
20. M. Hirsch, S. Sivaramakrishnan, S. Jayasuriya, A. Wang, A. Molnar, R. Raskar, and G. Wetzstein, "A switchable light field camera architecture with angle sensitive pixels and dictionary-based sparse coding," in *IEEE International Conference on Computational Photography (ICCP)*, 2014, pp. 1–10.
21. X. Lin, J. Suo, G. Wetzstein, Q. Dai, and R. Raskar, "Coded focal stack photography," in *IEEE International Conference on Computational Photography (ICCP)*, 2013, pp. 1–9.
22. X. Yin, G. Wang, W. Li, and Q. Liao, "Large aperture focus stacking with max-gradient flow by anchored rolling filtering," Appl. Opt. **55**, 5304–5309 (2016).
23. L. McMillan and G. Bishop, "Plenoptic modeling: an image based rendering approach," in *Proc. SIGGRAPH*, 1995, pp. 39–46.
24. A. Mousnier, E. Vural, and C. Guillemot, "Partial light field tomographic reconstruction from a fixed-camera focal stack," arXiv:1503.01903 (2015).
25. S. Kuthirummal, H. Nagahara, C. Zhou, and S. K. Nayar, "Flexible depth of field photography," IEEE Trans. Pattern Anal. Mach. Intell. **33**, 58–71 (2011).
26. G. Zeng, "One-angle fluorescence tomography with in-and-out motion," J. Electron Imaging **22**, 043018 (2013).
27. A. Levin and F. Durand, "Linear view synthesis using a dimensionality gap light field prior," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, 2010, pp. 1831–1838.
28. J. Alonso, A. Fernández, and J. Ferrari, "Reconstruction of perspective shifts and refocusing of a three-dimensional scene from a multi-focus image stack," Appl. Opt. **55**, 2380–2386 (2016).
29. K. He, J. Sun, and X. Tang, "Guided image filtering," IEEE Trans. Pattern Anal. Mach. Intell. **35**, 1397–1409 (2013).
30. The (New) Stanford Light Field Archive, http://lightfield.stanford.edu.
31. P. Thévenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," IEEE Trans. Image Process. **7**, 27–41 (1998).
32. B. He, G. Wang, X. Lin, C. Shi, and C. Liu, "High-accuracy sub-pixel registration for noisy images based on phase correlation," IEICE Trans. Inf. Syst. **E94-D**, 2541–2544 (2011).
33. Q. Miao, G. Wang, and X. Lin, "Kernel based image registration incorporating with both feature and intensity matching," IEICE Trans. Inf. Syst. **E93-D**, 1317–1320 (2010).
34. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).