

REAL-TIME 3-D IMAGE RECONSTRUCTION FROM MULTI-FOCUS IMAGES BY EFFICIENT LINEAR FILTERING WITH MULTI-DIMENSIONAL SYMMETRY

Kazuya Kodama*

Zhen Wang†

Masanori Sato‡

Tomochika Murakami‡

* National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

† Tianjin University of Finance and Economics, #25 Zhujiang Rd., Hexi Dist., Tianjin 300222, China

‡ Canon Inc., 3-30-2 Shimomaruko, Ohta-ku, Tokyo 146-8501, Japan

ABSTRACT

We previously proposed a simple method of 3-D image reconstruction using scene-independent linear filters for multi-focus images. However, it costs much to prepare the linear filters for high-resolution images. In this paper, by eliminating symmetric redundancy of the linear filtering, we significantly reduce required memory and computational cost by less than one-tenth. Then, we achieve real-time and interactive image reconstruction with 3-D effects from high-resolution multi-focus images on a GPGPU. Experimental results using microscopic images reveal that our novel method enables us to observe 3-D objects very effectively by free viewpoint image reconstruction and scene refocusing with desired bokeh at about 60 fps and 15 fps, respectively.

Index Terms— real-time, image reconstruction, multi-focus, linear filtering, symmetry

1. INTRODUCTION

Our previously proposed linear filters for multi-focus images robustly achieve free viewpoint image reconstruction and scene refocusing as its extension [1]. In this paper, by appropriately utilizing multi-dimensional symmetry of the linear filters for redundancy elimination, we demonstrate such 3-D image reconstruction from multi-focus images in real time with its efficient implementation on a GPGPU. Real-time and interactive visualization by our image reconstruction with 3-D effects greatly helps us recognize 3-D structure of objects.

In order to avoid unstable reconstruction of 3-D structure itself [2], extended depth-of-field has been studied well as practical visualization from multi-focus images for examining objects at a glance [3, 4, 5], especially, in the case of microscopy. Actually, various industrial applications are developed for all-in-focus image reconstruction from multi-focus images [6, 7]. Moreover, real-time visualization of moving objects with all-in-focus video based on high-speed acquisition of microscopic multi-focus images has been already provided [8, 9]. On the other hand, extended depth-of-field often prevents us from recognizing 3-D structure of objects because

reconstructed all-in-focus images essentially lose depth information implied by blurs on multi-focus images. Therefore, we have extended 2-D all-in-focus image reconstruction to 4-D light field synthesis by 3-D frequency analysis of multi-focus images [10, 11, 12], and then scene-independent linear filters for multi-focus images are finally derived [1] to robustly reconstruct free viewpoint images, that enable us to observe 3-D objects effectively. Recently, researchers of computational photography [13, 14] in computer vision also study similar approaches to 3-D image reconstruction by 4-D frequency analysis of light fields [15, 16].

Because of simple scene-independent linear filtering, our 3-D image reconstruction is achieved in real time if the corresponding linear filters are given. Rather than image reconstruction, it costs much to synthesize our filters directly. Actually, we have only implemented real-time 3-D image reconstruction using linear filters prepared in advance [17, 18]. In addition, if not using coarse approximation causing degradation [17], it limits reconstructed images to very low resolution for storing various sets of linear filters in high-speed memory. In this paper, we propose an efficient method of 3-D image reconstruction with just-in-time synthesis of linear filters by utilizing multi-dimensional symmetry of our filters for reduction of required memory and computational cost. Experimental results using microscopic high-resolution multi-focus images reveal that the proposed method achieves real-time and interactive image reconstruction with high-quality 3-D effects, that helps us recognize 3-D structure of objects very easily.

2. IMAGE RECONSTRUCTION WITH 3-D EFFECTS

2.1. Linear filters for multi-focus images

Let $g^{(n)}(x, y)$ be multi-focus images corresponding to appropriate focal planes, by applying 2-D FFT to each image,

$$G^{(n)}(u, v) = \mathcal{F}_x \mathcal{F}_y [g^{(n)}(x, y)] \quad (1)$$

are obtained in the frequency domain, where we achieve simple linear combination of them [1] as follows:

$$A(u, v) = \sum_n H^{(n)}(u, v) G^{(n)}(u, v). \quad (2)$$

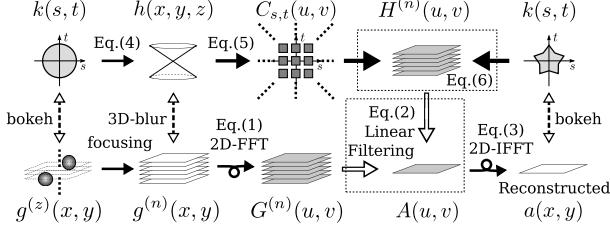


Fig. 1. 3-D image reconstruction from multi-focus images.

Then, as in Fig. 1, desired images with various 3-D effects

$$a(x, y) = \mathcal{F}_x^{-1} \mathcal{F}_y^{-1} [A(u, v)] \quad (3)$$

are reconstructed through 2-D inverse FFT. Our linear filters $H^{(n)}(u, v)$ in (2) are scene-independent and synthesized by

$$h(x, y, z) = \int k(s, t) \delta(x + sz, y + tz) ds dt, \quad (4)$$

$$C_{s,t}(u, v) = \int \mathcal{F}_x \mathcal{F}_y [h(x, y, z)] e^{-2\pi i (su+tv)z} dz, \quad (5)$$

$$H^{(n)}(u, v) = \int \hat{k}(s, t) e^{-2\pi i (su+tv)(n-n_f)} C_{s,t}(u, v)^{-1} ds dt, \quad (6)$$

where $k(s, t)$ and $\hat{k}(s, t)$ correspond to bokeh on the original $g^{(n)}(x, y)$ and desired $a(x, y)$, respectively. We also need n_f to determine which one among the original focal planes is selected for $a(x, y)$. Here, (s, t) denotes coordinates of a point on the original aperture, and $h(x, y, z)$ expresses 3-D blurs on $g^{(n)}(x, y)$ as distribution like a point spread function.

2.2. Required procedures for 3-D image reconstruction

In this paper, by arranging required procedures in Fig. 1,

1. Synthesis of linear filters decomposing 3-D blurs [1]: the original bokeh $k(s, t) \xrightarrow{(4)} h(x, y, z) \xrightarrow{(5)} C_{s,t}(u, v)^{-1}$
2. 2-D FFT: acquired $g^{(n)}(x, y) \xrightarrow{(1)} G^{(n)}(u, v) \quad (n \in \mathbb{Z})$
3. Synthesis of linear filters for image reconstruction: desired $\hat{k}(s, t)$, desired n_f , $C_{s,t}(u, v)^{-1} \xrightarrow{(6)} H^{(n)}(u, v)$
4. Linear filtering and combination: $G^{(n)}(u, v), H^{(n)}(u, v) \xrightarrow{(2)} A(u, v)$
5. 2-D inverse FFT: $A(u, v) \xrightarrow{(3)} a(x, y)$

are implemented for our 3-D image reconstruction. Before the other procedures, only by camera characteristics and parameters, we determine and store scene-independent linear filters $C_{s,t}(u, v)^{-1}$, that are easily obtained from $C_{s,t}(u, v)$ of (5) to decompose multi-focus images $g^{(n)}(x, y)$ into the corresponding light field. Of course, the filters do not depend on

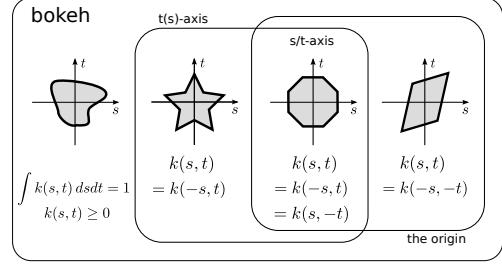


Fig. 2. Various kinds of symmetry for bokeh $k(s, t)$.

$g^{(n)}(x, y)$ acquired from 3-D scenes. In addition, for changing 3-D effects on $a(x, y)$, we only need to repeat the procedures of 3~5 by using common $C_{s,t}(u, v)^{-1}$ stored in high-speed memory. Thus, the proposed implementation with a pointing device shows us real-time and interactive 3-D effects on reconstructed images.

3. REAL-TIME 3-D IMAGE RECONSTRUCTION

In this section, by utilizing multi-dimensional symmetry of linear filters $C_{s,t}(u, v)^{-1}$ and $H^{(n)}(u, v)$ for redundancy elimination, the proposed 3-D image reconstruction described in the previous section is implemented on a GPGPU of NVIDIA Quadro K4000 with CUDA. We discuss how symmetry of bokeh $k(s, t)$ and $\hat{k}(s, t)$, as shown in Fig. 2, affects multi-dimensional symmetry of the linear filters, that is thoroughly utilized for reducing computational cost of (5) and (6) as well as memory usage required to store their results.

Here, computational cost and required memory are evaluated by using 15 multi-focus images of 1024×1024 pixels with 24-bit RGB color. For simplicity, we set $k(s, t)$ and $\hat{k}(s, t)$ to preferable Gaussian bokeh [1], that has all kinds of symmetry in Fig. 2, and (s, t) is discretized into 25×25 elements for them. Moreover, in the case of Gaussian bokeh, we do not use 2-D FFT but analytic solutions for Fourier transform in (5).

3.1. Efficient synthesis of $C_{s,t}(u, v)^{-1}$ using 4-D symmetry

In Fig. 3, linear filters $C_{s,t}(u, v)^{-1}$ for decomposing multi-focus images into the corresponding 4-D light field [1] are shown, where $k(s, t)$ is set to Gaussian bokeh. We notice that they have two kinds of symmetry for (s, t) and (u, v) . First, $C_{s,t}(u, v)^{-1}$ of each (s, t) has 2-D symmetry for (u, v) . Secondly, some of these 2-D linear filters have symmetric relation to each other. Actually, as in Tab. 1, the former always exists between (u, v) and $(-u, -v)$ regardless of $k(s, t)$, however, the latter among $(\pm s, \pm t)$ depends on symmetry of $k(s, t)$.

We easily utilize such 4-D symmetry for reduction of required memory and computational cost, that are also shown in Tab. 1. With respect to the former symmetry, the symmetric interval of integration in (5) enables us to omit storing not only half of the real part but also the imaginary part,

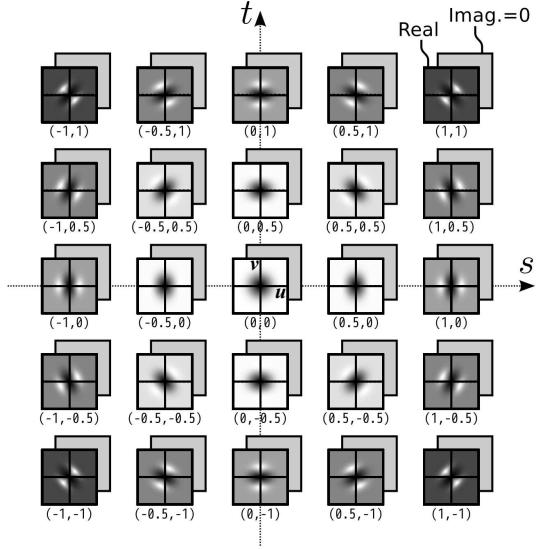


Fig. 3. An example of redundant $C_{s,t}(u,v)^{-1}$.

that is always 0. Then, required memory for 2-D linear filter $C_{s,t}(u,v)^{-1}$ of each (s,t) is reduced by 1/4. Moreover, the latter symmetry depending on $k(s,t)$ enables us to use a 2-D linear filter repeatedly among $(\pm s, \pm t)$, where required memory and computational cost are further reduced by 1/4 at best.

If not eliminating any redundancy, computational time for synthesizing all of linear filters $C_{s,t}(u,v)^{-1}$ is 460 [ms], that seems to be also reduced by about 1/4 in Tab. 1 with respect to the former symmetry. However, unlike required memory described above, effects of omitting the imaginary part on computational time hardly exist, because we need only a trigonometric function for synthesizing the imaginary part in addition to the real part and it is efficiently computed by a Special Function Unit (SFU) on our GPGPU in parallel with the other complicated operations. Actually, regardless of $k(s,t)$, we utilize 3-D symmetry of $h(x,y,z)$ for replacing the interval of integration with $z \geq 0$ in (5) to reduce its computational cost by about 1/2. Then, the total computational cost is finally reduced by about 1/4 with respect to the former symmetry.

Table 1. Redundancy elimination for computing $C_{s,t}(u,v)^{-1}$.

Symmetry of $k(s,t)$	Memory Time	Characteristics	
		$\mathbb{R} \ni C_{s,t}(u,v) = C_{s,t}(-u,-v) \cdots (*)$	
none	$1/4 \times 1$		
	134 [ms]		
$t(s)$ -axis	$1/4 \times 1/2$	$(*)$,	
	67 [ms]	$C_{s,t}(u,v) = C_{-s,t}(-u,v)$	
the origin	$1/4 \times 1/2$	$(*)$,	
	67 [ms]	$C_{s,t}(u,v) = C_{-s,-t}(-u,-v)$	
s/t -axis	$1/4 \times 1/4$	$(*)$,	
	34 [ms]	$C_{s,t}(u,v) = C_{-s,t}(-u,v) = C_{s,-t}(u,-v)$	

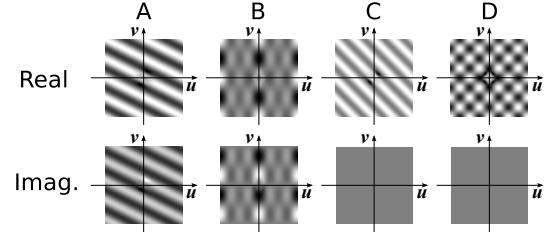


Fig. 4. Examples of redundant $H^{(n)}(u,v)$.

3.2. Efficient synthesis of $H^{(n)}(u,v)$ using 3-D symmetry

As shown in Tab. 2, linear filters $H^{(n)}(u,v)$ for image reconstruction with desired 3-D effects have two kinds of symmetry for n and (u,v) . First, 3-D point symmetry always exists around the origin regardless of the original bokeh $k(s,t)$ and desired bokeh $\hat{k}(s,t)$. Secondly, 2-D linear filter $H^{(n)}(u,v)$ of each n has 2-D symmetry depending on combination of $k(s,t)$ and $\hat{k}(s,t)$, that we can see as four examples in Fig. 4. In order to efficiently synthesize $H^{(n)}(u,v)$, such multi-dimensional symmetry is utilized for reducing required memory and computational cost, that are also shown in Tab. 2.

In the same way as $C_{s,t}(u,v)^{-1}$, we note that, for C in Tab. 2, omitting the imaginary part further reduces required memory by 1/2, while its effects on computational cost hardly exist because it can be almost ignored in comparison with latency of reading $C_{s,t}(u,v)^{-1}$ in global memory on our GPGPU. Actually, here, we utilize point symmetry of integrand around the origin in (6) for replacing the domain of integration with $s \geq 0$ to reduce its computational cost by about 1/2.

Because the total computational time of $H^{(n)}(u,v)$ is 716 [ms] if not eliminating any redundancy, both computational cost and required memory are finally reduced by less than one-tenth for D in Tab. 2 as combination of B and C.

Table 2. Redundancy elimination for computing $H^{(n)}(u,v)$.

$k(s,t)$ and $\hat{k}(s,t)$	\hat{k}			
	none	$t(s)$ -axis	the origin	s/t -axis
k	A	A	A	A
	A	B	A	B
	A	A	C	C
	A	B	C	D

	Memory Time	Characteristics	
		$H^{(n)}(u,v) = H^{(n)}(-u,-v) \cdots (*)$, $H^{(n)}(u,v) = \overline{H^{(n)}(-u,-v)}$	$H^{(n)}(u,v) = H^{(n)}(-u,v) = \overline{H^{(n)}(-u,-v)}$
A	$1/2 \times 1/2$ 191 [ms]		
B	$1/2 \times 1/4$ 105 [ms]	$(*)$,	
C	$1/2 \times 1/4$ 92 [ms]	$(*)$,	$\mathbb{R} \ni H^{(n)}(u,v) = H^{(n)}(-u,-v)$
D	$1/2 \times 1/8$ 56 [ms]	$(*)$,	$\mathbb{R} \ni H^{(n)}(u,v) = H^{(n)}(-u,v) = H^{(n)}(u,-v)$

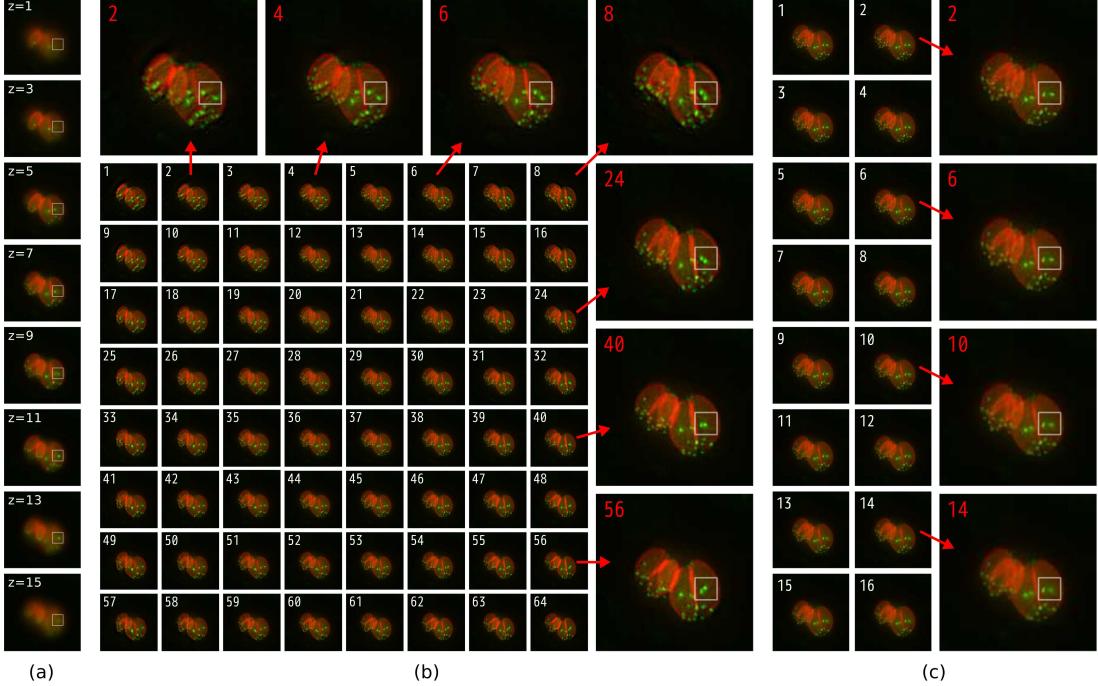


Fig. 5. Real-time 3-D image reconstruction from multi-focus images of Toxoplasma (simply see the squared region): (a) the original multi-focus images, (b) reconstructed multi-view images, (c) reconstructed images with various extents of bokeh.

3.3. Experiments using microscopic images

By using multi-focus images acquired by Olympus IX70 Inverted Microscope, we obtain experimental results in Fig. 5 to evaluate our proposed method and its implementation for real-time and interactive image reconstruction with 3-D effects. The original multi-focus images of a parasite called Toxoplasma [19] in Fig. 5(a) fit our imaging model with 3-D blurs well [2]. In Fig. 5(b), we see reconstructed images for 8×8 viewpoints by setting $\hat{k}(s, t)$ to 0 except for the desired viewpoint, where $\hat{k}(s, t) = 1$. In Fig. 5(c), scene refocusing is also shown by varying the extent of bokeh while fixing a focal plane to $z = 7$. They are achieved at the rate of 16.0 [ms] and 65.5 [ms] a frame, respectively. The former is much faster than the latter because we omit reading $C_{s,t}(u, v)^{-1}$ in global memory if $\hat{k}(s, t) = 0$ when synthesizing $H^{(n)}(u, v)$ for free viewpoint image reconstruction. In order to see details, we show only the central 256×256 pixels of each image in Fig. 5.

In the case of transparent or semi-transparent objects as in Fig. 5, it is usually very difficult to achieve robust reconstruction of 3-D structure and its real-time visualization. Only well-trained experts can recognize 3-D structure of such objects directly by the original microscopic multi-focus images. On the other hand, our proposed method and its implementation help even ordinary users recognize it very easily by smooth and interactive observation while varying a viewpoint as shown in Fig. 5(b) at over 60 [frame/s] with reconstructed

sufficiently high-resolution images of 1024×1024 pixels. For example, internal structure of Toxoplasma is well recognized by the parallax of its green dots while varying a viewpoint.

Moreover, scene refocusing as in Fig. 5(c) is smoothly achieved at over 15 [frame/s]. That is to say, we can find desired bokeh from even all of the reconstructed images in Fig. 5(c) within a second. Of course, we also easily let desired objects be in focus by varying a focal plane at the same rate.

4. CONCLUSIONS AND FUTURE WORKS

We significantly reduced required memory and computational cost of linear filtering for 3-D image reconstruction from multi-focus images by utilizing multi-dimensional symmetry of the linear filters for redundancy elimination. Experimental results using microscopic high-resolution images revealed that, beyond the original multi-focus images, its implementation on a GPGPU helps us recognize 3-D structure of objects very easily by real-time and interactive observation using free viewpoint image reconstruction and scene refocusing.

In the future, we will further improve efficiency of the proposed method by studying more simple expression of our linear filters. In addition, such efficient 3-D image reconstruction can also be expected to serve for practical implementation of various important applications including multi-view image restoration [20] and compression [21, 22].

5. REFERENCES

- [1] K. Kodama and A. Kubota, "Efficient reconstruction of all-in-focus images through shifted pinholes from multi-focus images for dense light field synthesis and rendering," *IEEE Transactions on Image Processing*, vol. 22, no. 11, pp. 4407–4421, Nov. 2013.
- [2] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 32–45, May 2006.
- [3] B. Forster, D. Van De Ville, J. Berent, D. Sage, and M. Unser, "Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images," *Microscopy Research and Technique*, vol. 65, no. 1-2, pp. 33–42, Sept. 2004.
- [4] F. Aguet, D. Van De Ville, and M. Unser, "Model-based 2.5-D deconvolution for extended depth of field in brightfield microscopy," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1144–1153, July 2008.
- [5] S. Pertuz, D. Puig, M. A. Garcia, and A. Fusiello, "Generation of all-in-focus images by noise-robust selective fusion of limited depth-of-field images," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1242–1251, Mar. 2013.
- [6] Helicon Soft Ltd., "Helicon Focus," <http://www.heliconsoft.com/heliconfocus.html>.
- [7] Zerene Systems Inc., "Zerene Stacker," <http://www.zerenesystems.com/cms/stacker>.
- [8] J. C. P. Ortega, K. Ohba, K. Tanie, G. Rin, R. Dangi, Y. Takei, T. Kaneko, and N. Kawahara, "Real-time micro VR camera system," in *Proc. Asian Conference on Computer Vision*, 2000, pp. 508–513.
- [9] K. Ohba, J. C. P. Ortega, K. Tanie, M. Tsuji, and S. Yamada, "Microscopic vision system with all-in-focus and depth images," *Machine Vision and Applications*, vol. 15, no. 2, pp. 55–62, Dec. 2003.
- [10] K. Kodama, H. Mo, and A. Kubota, "Free viewpoint, iris and focus image generation by using a three-dimensional filtering based on frequency analysis of blurs," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006, vol. 2, pp. 625–628.
- [11] K. Kodama, H. Mo, and A. Kubota, "Simple and fast all-in-focus image reconstruction based on three-dimensional/two-dimensional transform and filtering," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007, vol. 1, pp. 769–772.
- [12] K. Kodama, I. Izawa, and A. Kubota, "Robust reconstruction of arbitrarily deformed bokeh from ordinary multiple differently focused images," in *Proc. IEEE International Conference on Image Processing*, 2010, pp. 3989–3992.
- [13] M. Levoy, "Light fields and computational imaging," *IEEE Computer*, vol. 39, no. 8, pp. 46–55, Aug. 2006.
- [14] C. Zhou and S. K. Nayer, "Computational cameras: Convergence of optics and processing," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3322–3340, Dec. 2011.
- [15] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman, "4D frequency analysis of computational cameras for depth of field extension," in *Proc. ACM International Conference and Exhibition on Computer Graphics and Interactive Techniques*, 2009, pp. 97:1–97:14.
- [16] A. Levin and F. Durand, "Linear view synthesis using a dimensionality gap light field prior," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1831–1838.
- [17] I. Izawa, T. Hamamoto, and K. Kodama, "A study on high-quality free viewpoint image reconstruction systems using multi-focus images by FPGA-based signal processing," in *Proc. IEEE International Conference on Image Processing*, 2010, pp. 413–416.
- [18] Y. Minato, K. Kodama, and T. Hamamoto, "Flexible linear view/focus synthesis on an FPGA using multi-focus images," in *Proc. International Workshop on Advanced Image Technology*, 2015, OS-15, pp. 1–4.
- [19] J. Murray and K. Hu, "The Cell: An Image Library - Image CIL:10512," <http://www.cellimagelibrary.org/images/10512>.
- [20] K. Kodama and A. Kubota, "Linear view/image restoration for dense light fields," in *Proc. IEEE International Conference on Image Processing*, 2014, pp. 5462–5466.
- [21] T. Sakamoto, K. Kodama, and T. Hamamoto, "A novel scheme for 4-D light-field compression based on 3-D representation by multi-focus images," in *Proc. IEEE International Conference on Image Processing*, 2012, pp. 2901–2904.
- [22] T. Sakamoto, K. Kodama, and T. Hamamoto, "A study on efficient compression of multi-focus images for dense light-field reconstruction," in *Proc. IEEE Visual Communications and Image Processing*, 2012, FP.150, pp. 1–6.