

# A high-resolution lightfield camera with dual-mask design

Zhimin Xu and Edmund Y. Lam

Imaging Systems Laboratory, Department of Electrical and Electronic Engineering  
The University of Hong Kong, Pokfulam Road, Hong Kong

## ABSTRACT

In this paper, we present a new design for lightfield acquisition. In comparison with the conventional lightfield acquisition techniques, the key characteristic of our system is its ability to achieve a higher resolution lightfield given a fixed sensor. In particular, the system architecture employs two attenuation masks respectively positioned at the aperture stop and the optical path of the camera, so that the four-dimensional (4D) lightfield spectrum is encoded and sampled by a two-dimensional (2D) camera sensor in a single snapshot. In post-processing, by exploiting the coherence embedded in a lightfield, we are able to retrieve the desired 4D lightfield of a higher resolution using inverse imaging. We demonstrate the performance of our proposed method with simulations based on the actual lightfield dataset.

**Keywords:** High-resolution lightfield, lightfield acquisition, computational photography

## 1. INTRODUCTION

In recent years, there has been increasing interest in the acquisition of a four-dimensional (4D) lightfield. One main reason is because the tremendous information contained in a lightfield allows us to explore many potential applications, such as digital refocusing<sup>1,2</sup> and anti-glare.<sup>3</sup>

Various systems have been proposed to capture lightfields. One of the earliest work is the integral photography,<sup>4</sup> which gathers multiple images from different perspectives by placing an array of microlenses directly before the sensor. More recently, Ng et al.<sup>1</sup> present a portable plenoptic camera which can capture a full 4D lightfield in a single snapshot. In their design, with the assistance of a microlens array placed at the focal plane of the main lens, the converging rays are separated and finally recorded by the sensor. Furthermore, there are several acquisition schemes by using non-refractive elements. For example, the use of a patterned mask to recover a lightfield from a single captured image is proposed in Ref. 2. After that, Agrawal et al.<sup>5</sup> present another mask-based approach to capture useful subsets of a time-varying 4D lightfield in a single snapshot. This reinterpretable imaging system is based on a design of a time-varying mask in the pupil plane and a static mask placed near the sensor, providing a variable resolution tradeoff among the spatial, angular and temporal dimensions.

While many methods have been proposed for lightfield acquisition, but a common issue for these different systems is that the spatial resolution has to be sacrificed to gain the angular variance because the limited sensor elements have to be allocated to both spatial and angular dimensions.<sup>6,7</sup> Many attempts have been made to overcome this tradeoff, but they come with other compromises. For example, the camera array system in Ref. 8 can get a 4D lightfield with high spatial resolution. But the huge size of the system limits its practical use. A design known as programmable aperture photography<sup>9</sup> offers us an alternative way. However, we need a large number of captures to achieve the required angular resolution. This results in a long acquisition time, which is not desirable in many practical applications.

In this paper, we present a camera system by using the mask-based multiplexing technique to collect a 4D lightfield within a single exposure. Specifically, we place two attenuating masks separately at the aperture plane and the optical path of a regular camera. By means of such a double-mask design, we can achieve the 4D lightfield acquisition in a more effective way by sub-sampling the encoded lightfield spectrum in Fourier domain. We show that this economical and easily adjustable design can break through various limitations dwelling in other lightfield acquisition systems. Since only partial information about the lightfield is captured, an iterative optimization algorithm is proposed for post-reconstruction. In addition, we interpret the rationality of our design from a different point of view compared with our preview work.<sup>10</sup>

---

Further author information: (Send correspondence to Zhimin Xu)  
Zhimin Xu: E-mail: zmxu@eee.hku.hk

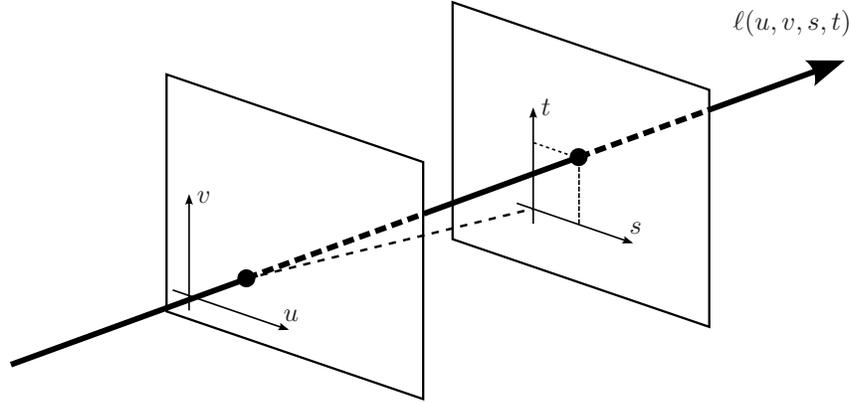


Figure 1. The two-plane parametrization of a 4D lightfield.

## 2. PRELIMINARIES

### 2.1 Lightfield and light propagation analysis in Fourier domain

We start with a brief discussion of the general concept of a lightfield and how light propagation phenomena affects it. A lightfield is a 4D radiance function which mathematically describes light energy flowing along the rays in 3D space. Among various parameterizations that have been proposed for a lightfield, we adopt the two-plane parameterization as used in Refs. 11 and 12. Specifically, we describe the light rays by their intersections with two parallel planes with unit distance between each other, *i.e.*, a first coordinate pair  $\mathbf{u} = \{u, v\}$  (at the  $\mathbf{u}$ -plane) and a second coordinate pair  $\mathbf{s} = \{s, t\}$  (at the  $\mathbf{s}$ -plane) as shown in Fig. 1. Note that  $\mathbf{s}$  is not as an absolute coordinate as  $\mathbf{u}$ , but a relative one against the projection of a point on the  $\mathbf{u}$ -plane. The lightfield is then  $\ell(u, v, s, t)$ , which we abbreviate as  $\ell(\mathbf{u}, \mathbf{s})$  in the rest of this paper.

As indicated by its definition, a lightfield refers to all possible rays in a region free of occluders.<sup>13</sup> However, when light flows in a scene, phenomena such as transport in free space and attenuation may occur, and eventually affect a lightfield in different manners. Similar analysis of the light propagation phenomena may have been mentioned in some literatures,<sup>12,14</sup> but here we only readdress the ones encountered in the design.

#### 1. Light transport in free space

The radiance of a ray is not changed, when light travels in free space. Assume  $z$  is the travel distance, then light transport in free space just reparameterizes the incoming lightfield  $\ell(\mathbf{u}, \mathbf{s})$ , *i.e.*,

$$\ell_o(\mathbf{u}, \mathbf{s}) = \ell(\mathbf{u} - z\mathbf{s}, \mathbf{s}), \quad (1)$$

where  $\ell_o(\mathbf{u}, \mathbf{s})$  represents the lightfield after the transport. Compute the Fourier transform of Eq. (1), then

$$\mathcal{L}_o(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}(\mathbf{f}_u, \mathbf{f}_s + z\mathbf{f}_u), \quad (2)$$

with  $\mathcal{L}_o(\mathbf{f}_u, \mathbf{f}_s)$  and  $\mathcal{L}(\mathbf{f}_u, \mathbf{f}_s)$  denoting the corresponding Fourier transforms of  $\ell_o(\mathbf{u}, \mathbf{s})$  and  $\ell(\mathbf{u}, \mathbf{s})$ . Clearly, this means a shear along the  $\mathbf{f}_s$  dimension.

#### 2. Attenuated by a mask

When light passes through a mask with the attenuation function denoted as  $m(\mathbf{u}, \mathbf{s})$ , the radiance value is then multiplied by this attenuation function. That is,

$$\ell_o(\mathbf{u}, \mathbf{s}) = \ell(\mathbf{u}, \mathbf{s})m(\mathbf{u}, \mathbf{s}). \quad (3)$$

According to the convolution theorem,<sup>15</sup> the Fourier transform of the multiplication in Eq. (3) corresponds to a convolution in frequency domain, *i.e.*,

$$\mathcal{L}_o(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}(\mathbf{f}_u, \mathbf{f}_s) * M(\mathbf{f}_u, \mathbf{f}_s), \quad (4)$$

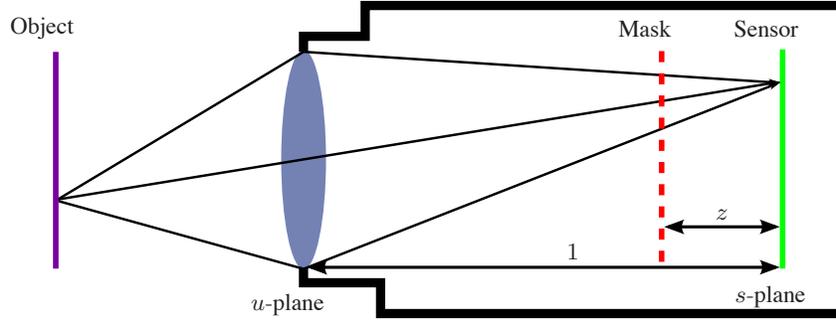


Figure 2. Schematic diagram of a regular camera, with an attenuation mask placed inside it.

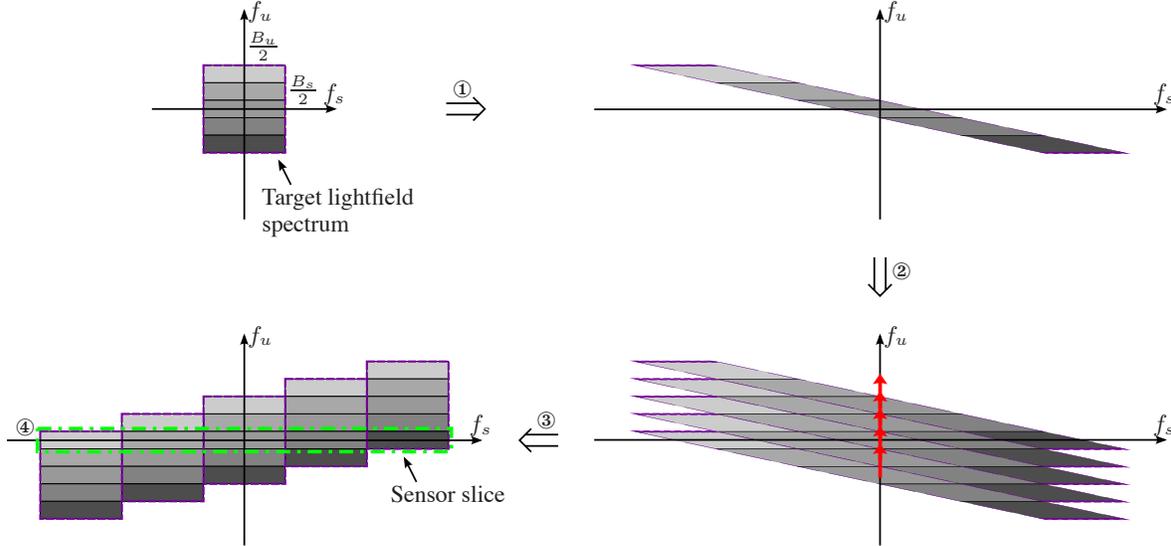


Figure 3. The frequency operations inside the system in Fig. 2.

where  $M(\mathbf{f}_u, \mathbf{f}_s)$  refers to the Fourier transform of  $m(\mathbf{u}, \mathbf{s})$  and  $*$  denotes the convolution operation. Generally, the mask generates attenuation in all four dimensions. That is, the spectrum  $M(\mathbf{f}_u, \mathbf{f}_s)$  has content in all four dimensions. In our scenario, however, we consider the situation that the mask is placed orthogonal to the optical axis. Thus, the mask modifies the radiance of the rays with the same  $\mathbf{u}$ -coordinate equally. If denoting the 2D pattern printed on the mask as  $c(\mathbf{x})$ , where  $\mathbf{x} = \{x, y\}$ , then

$$m(\mathbf{u}, \mathbf{s}) = c(\mathbf{u}). \quad (5)$$

Correspondingly, its spectrum  $M(\mathbf{f}_u, \mathbf{f}_s)$  can be expressed as

$$M(\mathbf{f}_u, \mathbf{f}_s) = C(\mathbf{f}_u)\delta(\mathbf{f}_s), \quad (6)$$

where  $C(\mathbf{f}_x)$  is the Fourier transform of  $c(\mathbf{x})$ .

## 2.2 Lightfield mapping via mask-based multiplexing

In this section, we explain the lightfield mapping via mask-based multiplexing by using the previous analysis of the light propagation. Figure 2 illustrates a typical mask-based lightfield acquisition framework. In terms of the two-plane parametrization shown in Fig. 1, the  $\mathbf{u}$ -plane is taken to be at the aperture, while the  $\mathbf{s}$ -plane at the sensor. Without loss of generality, we assume these two planes are separated by unit distance, while the mask is placed at a distance  $z$  in front of the sensor. Clearly, the lightfield  $\ell(\mathbf{u}, \mathbf{s})$  in this scenario is different from that without the mask. The light propagation in Fig. 2 can be considered as going through the following stages. (Each step is indicated by a corresponding circled number in Fig. 3.) We use  $\ell_i(\mathbf{u}, \mathbf{s})$  to denote the outcoming lightfield of the  $i$ -th stage, while  $\mathcal{L}_i(\mathbf{f}_u, \mathbf{f}_s)$  for its frequency response.

Step i The light rays firstly travel a distance of  $1 - z$  from the aperture to the mask. According to Eqs. (1) and (2), we have

$$\ell_1(\mathbf{u}, \mathbf{s}) = \ell[\mathbf{u} - (1 - z)\mathbf{s}, \mathbf{s}], \quad (7a)$$

$$\mathcal{L}_1(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}[\mathbf{f}_u, \mathbf{f}_s + (1 - z)\mathbf{f}_u]. \quad (7b)$$

This results in a shear along the  $\mathbf{f}_s$  axis as shown in Fig. 3. Note that the longer the travel distance is, the more remarkable the shear becomes.

Step ii Due to the mask, the sheared spectrum is then convolved by the mask attenuation  $M(\mathbf{f}_u, \mathbf{f}_s)$  along the  $\mathbf{f}_u$  axis. That is,

$$\begin{aligned} \ell_2(\mathbf{u}, \mathbf{s}) &= \ell_1(\mathbf{u}, \mathbf{s})m(\mathbf{u}, \mathbf{s}) \\ &= \ell_1(\mathbf{u}, \mathbf{s})c(\mathbf{u}), \end{aligned} \quad (8a)$$

$$\mathcal{L}_2(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}_1(\mathbf{f}_u, \mathbf{f}_s) * [C(\mathbf{f}_u)\delta(\mathbf{f}_s)]. \quad (8b)$$

If  $M(\mathbf{f}_u, \mathbf{f}_s)$  is a series of impulses, this produces a replication of the sheared spectrum.

Step iii Actually, this stage does not physically exist. After the second stage, the new lightfield  $\ell_2(\mathbf{u}, \mathbf{s})$  takes the mask plane as the  $\mathbf{u}$ -plane, while the corresponding  $\mathbf{s}$ -plane is at unit distance from it. However, the lightfield we want to acquire is based on the original parameterization, *i.e.*, with the  $\mathbf{u}$ -plane at the aperture. Therefore, we add a back-propagation stage to realize this reparameterization. Specifically, consider the light travels back to the aperture plane:

$$\ell_3(\mathbf{u}, \mathbf{s}) = \ell_2[\mathbf{u} + (1 - z)\mathbf{s}, \mathbf{s}], \quad (9a)$$

$$\mathcal{L}_3(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}_2[\mathbf{f}_u, \mathbf{f}_s - (1 - z)\mathbf{f}_u]. \quad (9b)$$

Note that the corresponding travel distance is negative. This causes a shear in the reverse direction along the  $\mathbf{f}_s$  axis.

Step iv Finally, the sensor at the  $\mathbf{s}$ -plane integrates all rays with the same  $(s, t)$  but different  $(u, v)$ :

$$i(\mathbf{s}) = \int_{-\infty}^{\infty} \ell_3(\mathbf{u}, \mathbf{s}) d\mathbf{u}, \quad (10)$$

where  $i(\mathbf{s})$  is the 2D picture recorded by the sensor.

Although the final integration at the sensor discards all the frequencies along the  $\mathbf{f}_u$  axis, if we can achieve a lightfield mapping as shown in Fig. 3, the 2D slice of data collected by the sensor (*i.e.*, sensor slice) can still contain all the information about the 4D lightfield. This makes recovery of the target lightfield possible.

The tradeoff along with this manner of acquisition is that the sensor slice in Fig. 3 needs to be much wider than what is needed in conventional photography; therefore, more samples are required to achieve the same 2D resolution for each perspective. In other words, to resolve  $n$  different views, only  $1/n$  of the pixels are assigned to sample the spectrum of each view. This ultimately results in a loss of the spatial resolution of each 2D image. Our design of a lightfield camera seeks to ameliorate this problem by showing that when the spectrum of each perspective view can contain more information, fewer replicas of the lightfield spectrum are needed. This means that the sensor slice is effectively narrowed, and consequently a higher resolution lightfield can be obtained with a fixed-size sensor.

### 3. CAMERA DESIGN AND LIGHTFIELD RECONSTRUCTION

#### 3.1 Lightfield capture with a double-mask design

Figure 4 shows the functional architecture of our proposed lightfield camera. During the system design and signal reconstruction, we assume that the lightfield spectrum is bandlimited, *i.e.*,  $\mathcal{L}(\mathbf{f}_u, \mathbf{f}_s) = 0$  for  $|\mathbf{f}_u| \geq B_u/2$  or  $|\mathbf{f}_s| \geq B_s/2$ . This is reasonable because the optics imposes a cutoff in the optical transfer function in the  $\mathbf{f}_s$  dimension.<sup>16</sup> As for  $\mathbf{f}_u$ , Ref. 11 shows that the corresponding bandwidth is basically determined by the depth range of a scene.

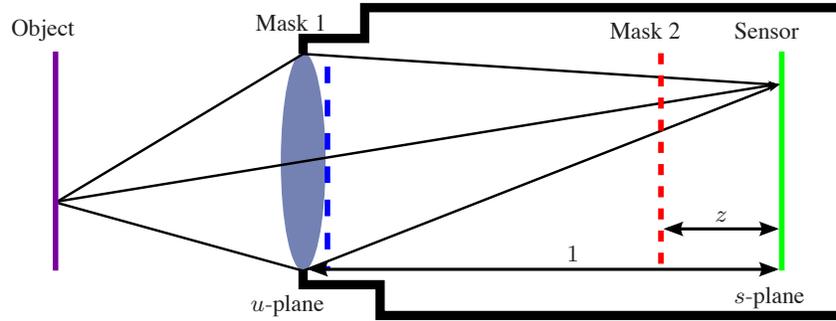


Figure 4. Our proposed lightfield camera, with two attenuation masks respectively placed at the aperture stop and the optical path of the camera.

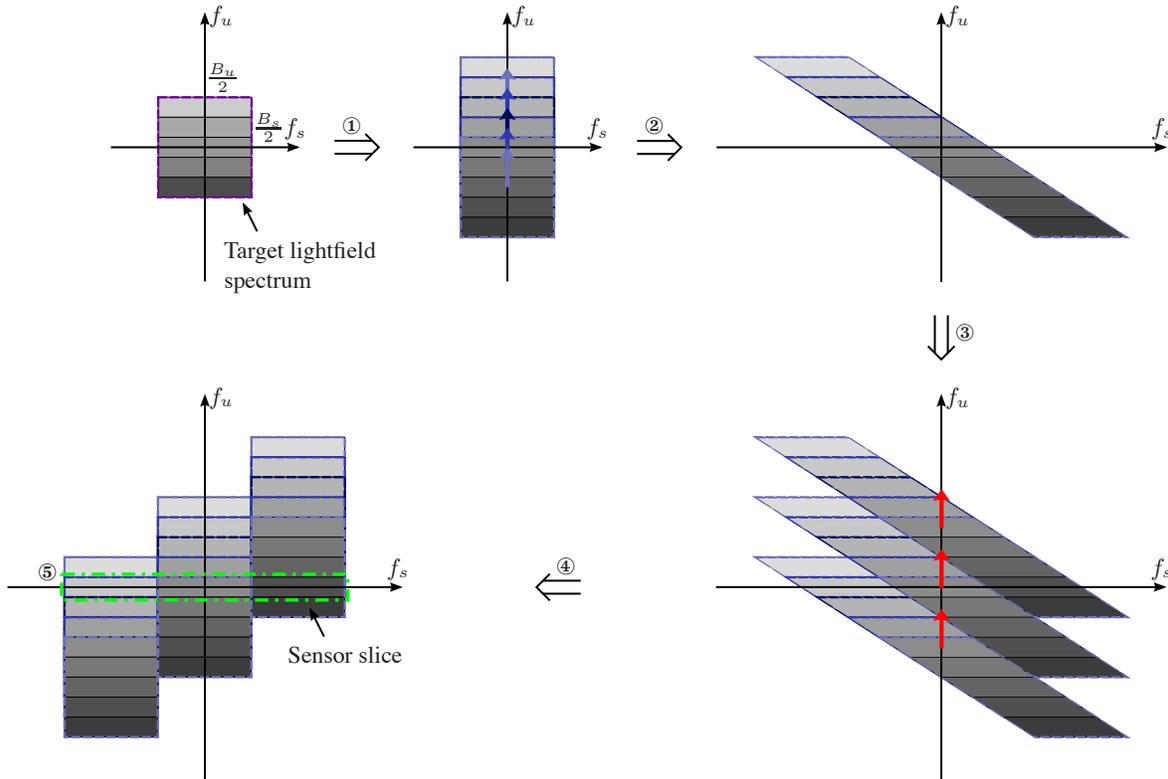


Figure 5. The frequency operations in our proposed lightfield camera.

The light flowing in the proposed system can also be analyzed by using the properties mentioned in Sec. 2.1. The major difference is the addition of a mask at the aperture. Because of it, the lightfield will be multiplied by its attenuation function  $m_1(\mathbf{u}, \mathbf{s})$  in the first place:

$$\begin{aligned} \ell_1(\mathbf{u}, \mathbf{s}) &= \ell(\mathbf{u}, \mathbf{s})m_1(\mathbf{u}, \mathbf{s}) \\ &= \ell(\mathbf{u}, \mathbf{s})c_1(\mathbf{u}), \end{aligned} \quad (11a)$$

$$\mathcal{L}_1(\mathbf{f}_u, \mathbf{f}_s) = \mathcal{L}(\mathbf{f}_u, \mathbf{f}_s) * [C_1(\mathbf{f}_u)\delta(\mathbf{f}_s)], \quad (11b)$$

where  $c_1$  is the corresponding 2D pattern and  $C_1$  denotes its Fourier transform. After that, the light propagation is similar to that in Fig. 2. Therefore, for the analysis of the following stages, we can directly take the conclusions in Sec. 2.2.

Figure 5 illustrates the frequency operations inside the proposed camera that we expect to realize. In particular, due to the first attenuation mask placed at the aperture stop, the incoming bandwidth-limited lightfield is convolved with mask

spectrum along the  $f_u$  axis. If the mask frequency response is a series of impulses, this produces a replication of the source lightfield spectrum along the  $f_u$  axis. Thus, we are able to achieve the lightfield spectrum encoding by making the spectrum overlaid on each other. The next transport from the first mask to the second one shears the encoded spectrum. Then the second mask makes this sheared spectrum replicated along the  $f_u$  axis. Similarly, we still need to reparameterize the lightfield, so a back-propagation process is taken into account. By adjusting the position of the second mask, we can control the shear strength so that the desired spectral slices can finally be placed on the  $f_s$  axis. Thereafter, we perform the lightfield reconstruction from the 2D slice data collected by the sensor in the fashion described in Sec. 3.3.

The frequency analysis provides us an intuitive knowledge of our design. However, for the purpose of mask design and lightfield retrieval, we need to explicitly model the acquisition process. This is expressed as an integration of the lightfield at the sensor:

$$i(\mathbf{s}) = \int_{-\infty}^{\infty} \ell_4(\mathbf{u}, \mathbf{s}) \, d\mathbf{u}. \quad (12)$$

According to the analysis in Sec. 2.2, we can get the following formulas:

$$\ell_2(\mathbf{u}, \mathbf{s}) = \ell_1[\mathbf{u} - (1 - z)\mathbf{s}, \mathbf{s}], \quad (13)$$

$$\ell_3(\mathbf{u}, \mathbf{s}) = \ell_2(\mathbf{u}, \mathbf{s})m_2(\mathbf{u}, \mathbf{s}) = \ell_2(\mathbf{u}, \mathbf{s})c_2(\mathbf{u}), \quad (14)$$

$$\ell_4(\mathbf{u}, \mathbf{s}) = \ell_3[\mathbf{u} + (1 - z)\mathbf{s}, \mathbf{s}], \quad (15)$$

where  $m_2(\mathbf{u}, \mathbf{s})$  is the attenuation provided by the second mask shown in Fig. 4. By inserting them and Eq. (11a) into the integral of Eq. (12), then we have

$$i(\mathbf{s}) = \int_{-\infty}^{\infty} \ell(\mathbf{u}, \mathbf{s})c_1(\mathbf{u})c_2[\mathbf{u} + (1 - z)\mathbf{s}] \, d\mathbf{u}. \quad (16)$$

In the content of geometrical optics, the image formation of lightfield cameras can be treated as a linear integration process.<sup>17,18</sup> We thus discretize Eq. (16) and convert it into matrix form as

$$i = F^{-1}M_2M_1F\ell = F^{-1}M\ell, \quad (17)$$

where  $F$  and  $F^{-1}$  are the matrices consisting of the Fourier basis and its inverse, and  $M_1$  and  $M_2$  respectively consist of the coefficients of the Fourier transforms of  $c_1$  and  $c_2$ . Therefore, this is a linear measurement process in Fourier domain through a measurement matrix  $M = M_2M_1$ .

If  $M$  is of size  $k \times n$ , it means that we sample  $k$  measurements of the coefficients decomposed by  $n$  Fourier bases. Thus, if we can achieve a measurement matrix  $M$  with  $k < n$ , fewer samples are needed subject to the same resolution requirement. We can compare the proposed design with the one presented in Ref. 2 which maps the whole lightfield spectrum inside the sensor slice. Consequently, the matrix  $M$  in their case is diagonal ( $k = n$ ). In our design, however, the measurement matrix is the product of two matrices  $M_2$  and  $M_1$ . This provides us the means to control the size of the two dimensions of  $M$  separately. As discussed next, we can then realize a measurement matrix with  $k < n$  in our design.

### 3.2 Mask design

The required frequency response of the first mask is shown in the first row of Fig. 5, *i.e.*, a symmetric impulse train. Specifically, assume the first mask has the following the frequency response:

$$C_1(f_u) = \sum_{i=-(n-1)}^{n-1} a_i \delta(f_u - i\Delta f_u), \quad (18)$$

where  $n$  is the expected resolution along the  $f_u$  axis,  $a_i$  is the amplitude of the impulse and  $\Delta f_u$  is the sampling interval of the lightfield spectrum along the  $f_u$  axis, which is equal to  $B_u/n$ . By converting the convolution of Eq. (11b) into a matrix multiplication, we can get  $M_1$  with Toeplitz structure. Recent research has shown that we can recover the original sparse signal with high probability from the limited observations measured by a well-designed Toeplitz-structured matrix.<sup>19,20</sup> To yield a Toeplitz-structured matrix that meets the requirements, several methods have been recommended. As suggested in

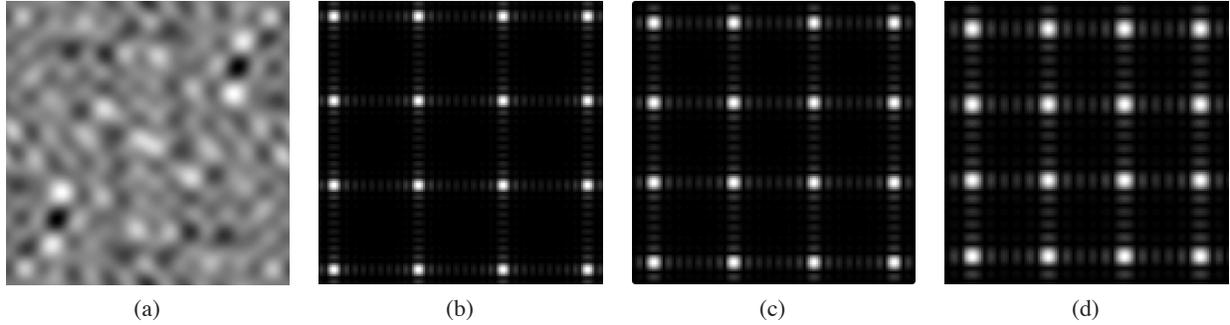


Figure 6. (a) The pattern of the first mask; (b), (c) and (d) are the patterns of the second mask, respectively in cases of using 100%, 64% and 36% sensor size.

Ref. 19, we generate  $M_1$  with entries drawn independently from a Gaussian distribution with zero mean. Eventually, we obtain the physical pattern of the first mask based on its frequency response in Eq. (18).

As to the second mask placed at the optical path, the Fourier transform of its pattern function  $c_2(x)$  is given by

$$C_2(f_x) = \sum_{i=-(k-1)/2}^{(k-1)/2} \delta(f_x - iB_s), \quad (19)$$

where  $k$  is the number of the measurements. Thus the corresponding mask pattern  $c_2(x)$  can be obtained by computing the inverse Fourier transform of Eq. (19), which is actually the sum of a series of cosine waves. For clarity only the case of 2D lightfield is carried out here, but the extension to the case of a 4D lightfield is straightforward.

### 3.3 Lightfield reconstruction

We adopt two different approaches to solve the underdetermined problem shown in Eq. (17). The first is to find the least-norm solution, *i.e.*,

$$\ell^* = A^\dagger i = A^T (AA^T)^{-1} i, \quad (20)$$

where  $A^\dagger$  denotes the pseudo-inverse of  $A$ , and  $A$  refers to  $F^{-1}MF$  in Eq. (17). The least-norm method is simple and fast, but not accurate enough. To improve the reconstruction accuracy, we impose regularization in our reconstruction algorithm. One is a sparse regularizer, which is a 2D total variation (TV) penalty on the  $\mathbf{u}$  dimension of a lightfield to reflect the inherent correlations. Furthermore, we use 2D TV norm regularization on the  $\mathbf{s}$  dimension of a lightfield to preserve the edges. Thus the proposed lightfield reconstruction can be interpreted as an optimization given by

$$\ell^* = \arg \min_{\ell} \frac{1}{2} \|A\ell - i\|_2^2 + \lambda \sum_{\mathbf{s}} \|i_{\mathbf{s}}\|_{\text{TV}} + \mu \sum_{\mathbf{u}} \|i_{\mathbf{u}}\|_{\text{TV}}, \quad (21)$$

where  $\lambda$  and  $\mu$  are the regularization parameters,  $i_{\mathbf{s}}$  is a 2D image corresponding to the lightfield  $\ell(\mathbf{u}, \mathbf{s})$  at a given point  $\mathbf{s}$ , and likewise  $i_{\mathbf{u}}$ . To solve the optimization problem, we adopt the nonlinear conjugate gradient algorithm combined with backtracking line search, as used in Ref. 21.

## 4. EXPERIMENTAL RESULTS

We demonstrate the performance of our proposed camera design using an actual lightfield dataset from the Stanford lightfield archive.<sup>22</sup> For computational considerations, we choose 100 views on a  $10 \times 10$  grid and resize the images to  $128 \times 256$  pixels. Thus, in order to capture a lightfield of such spatial resolution, the conventional lightfield cameras<sup>1,2</sup> will need a sensor of size  $1280 \times 2560$ . The following experimental results show that by using the proposed camera, a smaller sensor is needed to achieve lightfields with the same spatial resolution.

So for the mask pattern design, according to the introduction of the mask design in Sec. 3.2, especially Eq. (18), the frequency response of the mask at the aperture stop should be an even-symmetric impulse train of size  $19 \times 19$  (*i.e.*,  $2 \times 10 - 1 = 19$ ). The corresponding amplitudes of these impulses are drawn independently from a Gaussian distribution

with zero mean. The physical pattern shown in Fig. 6(a) is the one we used in the simulations. For the second mask at the optical path, its frequency response depends on the specific requirement of the measurement number. For example, for the case of using full sensor size (*i.e.*,  $1280 \times 2560$ ), it is a  $10 \times 10$  impulse train with equal amplitudes based on Eq. (19). Similarly,  $8 \times 8$  for the case of using 64% sensor size (*i.e.*,  $1024 \times 2048$ ) and  $6 \times 6$  for the case of using 36% sensor size (*i.e.*,  $768 \times 1536$ ). Figures 6(b), 6(c) and 6(d) show the patterns of the second mask in different cases.

Next, we show the performance of our camera when using different sensor size. Figure 7 shows the captured pictures by using different sensor sizes. Figure 8 shows the corresponding reconstruction images at one selected viewpoint. For the sake of comparison, we use both the least-norm method in Eq. (20) and our proposed algorithm in Eq. (21) for lightfield reconstruction. In the case of using full sensor, both two methods can yield as perfect reconstructions as the ground truth. When reducing the sensor size, the recoveries can still provide us as good details as the ground truth. In particular, our proposed algorithm can preserve more details and provide better artifact control (*e.g.*, the ringing artifacts around the beans).

To show that a high-resolution lightfield can be acquired by our proposed system, we consider the case of using 36% sensor size (*i.e.*,  $768 \times 1536$ ). If we use the conventional lightfield cameras of the same sensor size, the maximum spatial resolution that can be achieved will be  $76 \times 153$ , *i.e.*, 36% size of the original one. Figure 9 illustrates that with our proposed camera the signal can be recovered at a higher resolution than the conventional ones.

## 5. CONCLUSIONS

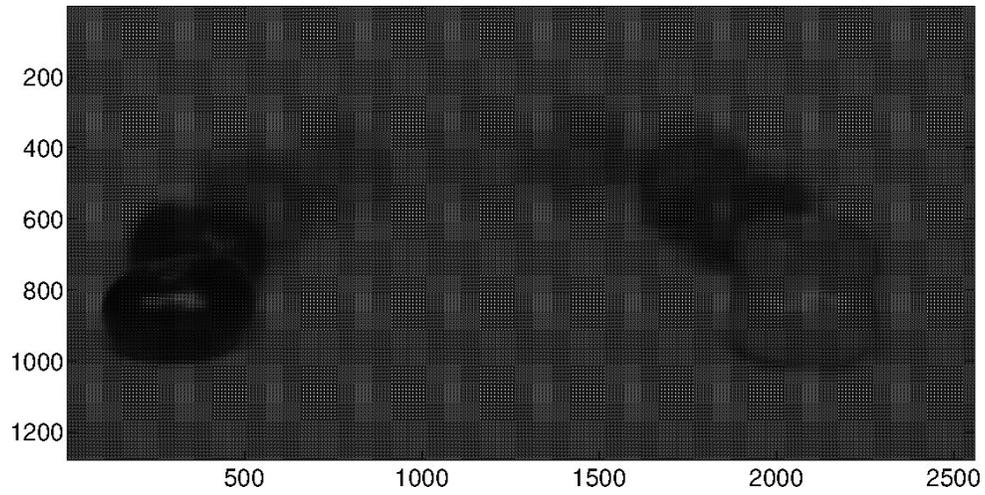
In this paper, we describe a system that is able to capture a 4D lightfield by using two attenuation masks. By exploiting the correlation embedded in the lightfield, we have developed a post-processing algorithm to reconstruct the target 4D lightfield from the captured 2D picture. We show experimental results to demonstrate that fewer pixels are needed to achieve the same resolution as what one can achieve by using conventional lightfield systems.

## ACKNOWLEDGMENTS

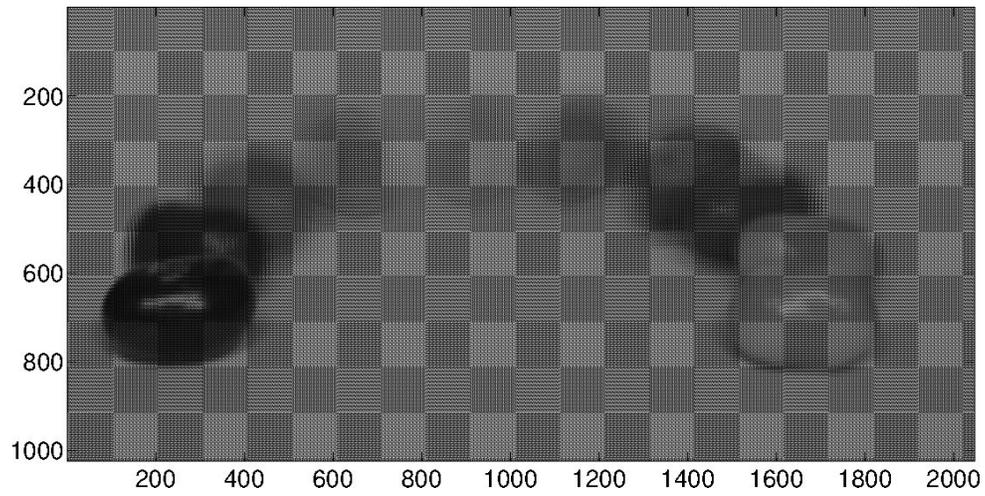
This work was supported in part by the University Research Committee of the University of Hong Kong under Project 10208648.

## REFERENCES

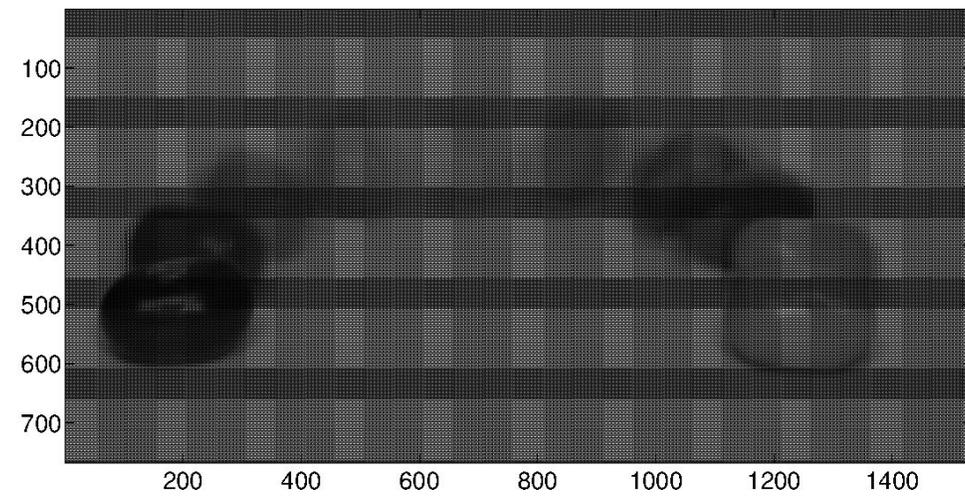
- [1] Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., and Hanrahan, P., "Light field photography with a hand-held plenoptic camera," Tech. Rep. CTSR 2005-02, Stanford University (2005).
- [2] Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., and Tumblin, J., "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," in [*Proceedings of ACM SIGGRAPH*], (2007).
- [3] Raskar, R., Agrawal, A., Wilson, C. A., and Veeraraghavan, A., "Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses," in [*Proceedings of ACM SIGGRAPH*], 56:1–56:10 (2008).
- [4] Lippmann, G., "Épreuves réversibles donnant la sensation du relief," *J. Phys. Théor. Appl.* **7**(1), 821–825 (1908).
- [5] Agrawal, A., Veeraraghavan, A., and Raskar, R., "Reinterpretable imager: Towards variable post-capture space, angle and time resolution in photography," *Comput. Graph. Forum* **29**(2), 763–772 (2010).
- [6] Georgeiv, T., Zheng, K. C., Curless, B., Salesin, D., Nayar, S., and Intwala, C., "Spatio-angular resolution tradeoff in integral photography," in [*Proceedings of Eurographics Symposium on Rendering* ], 263–272 (2006).
- [7] Xu, Z. and Lam, E. Y., "Light field superresolution reconstruction in computational photography," in [*Signal Recovery and Synthesis* ], SMB3, Optical Society of America, Washington, D.C. (2011).
- [8] Wilburn, B., Joshi, N., Vaish, V., Talvala, E.-V., Antunez, E., Barth, A., Adams, A., Horowitz, M., and Levoy, M., "High performance imaging using large camera arrays," in [*Proceedings of ACM SIGGRAPH* ], 765–776 (2005).
- [9] Liang, C.-K., Lin, T.-H., Wong, B.-Y., Liu, C., and Chen, H. H., "Programmable aperture photography: Multiplexed light field acquisition," in [*Proceedings of ACM SIGGRAPH* ], 55:1–55:10 (2008).
- [10] Xu, Z., Ke, J., and Lam, E. Y., "High-resolution lightfield photography using two masks," *Opt. Express* **20**(10), 10971–10983 (2012).
- [11] Chai, J.-X., Tong, X., Chan, S.-C., and Shum, H.-Y., "Plenoptic sampling," in [*Proceedings of ACM SIGGRAPH* ], 307–318 (2000).



(a)

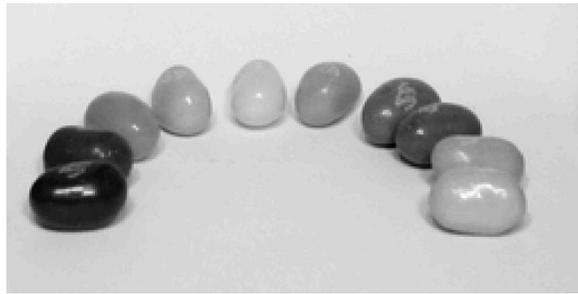


(b)

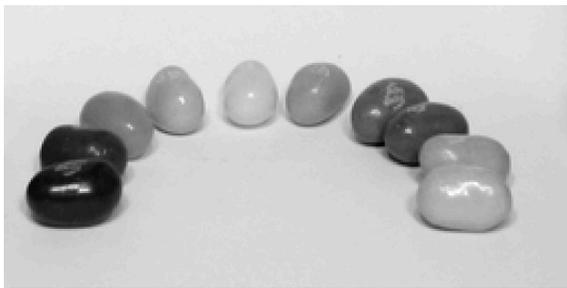


(c)

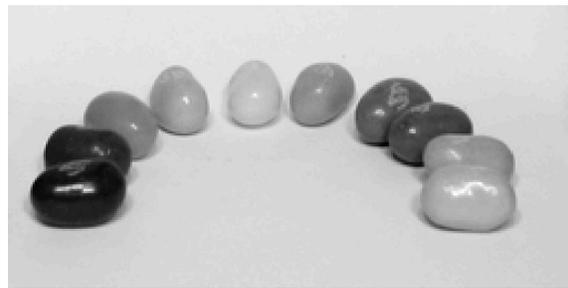
Figure 7. The pictures captured by using the proposed lightfield camera of different sensor sizes: (a) full size, (b) 64% (c) 36%.



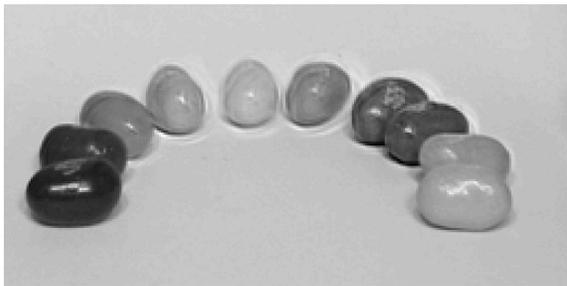
(a)



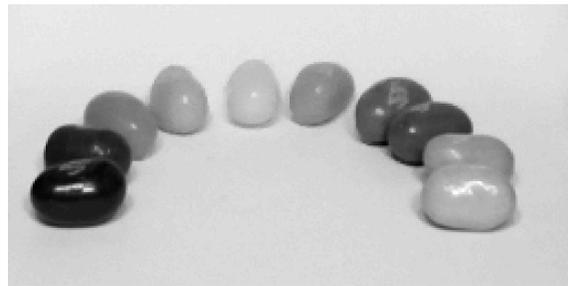
(b)



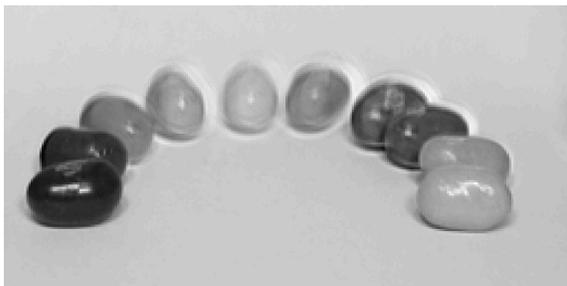
(c)



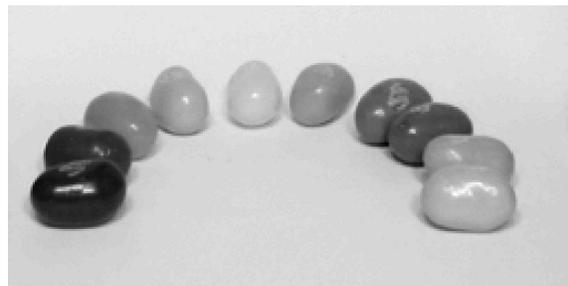
(d)



(e)



(f)



(g)

Figure 8. The reconstructed images at one selected viewpoint: (a) ground truth, (b) and (c) full size, (d) and (e) 64% sensor size, (f) and (g) 36% sensor size.

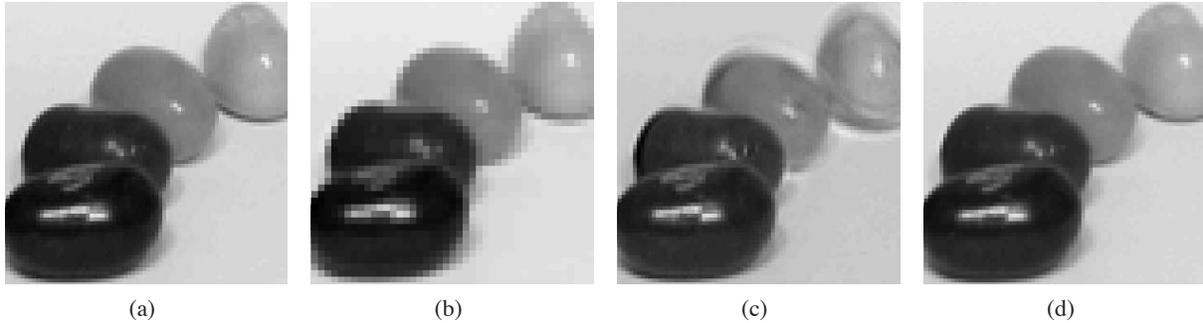


Figure 9. Reconstructions when using 36% sensor size: (a) The ground truth; (b) the best quality that can be achieved by using the conventional lightfield cameras; (c) the reconstruction result with the least-norm method; (d) our result by using the proposed iterative method.

- [12] Durand, F., Holzschuch, N., Soler, C., Chan, E., and Sillion, F. X., "A frequency analysis of light transport," in *Proceedings of ACM SIGGRAPH*, 1115–1126 (2005).
- [13] Levoy, M. and Hanrahan, P., "Light field rendering," in *Proceedings of ACM SIGGRAPH*, 31–42 (1996).
- [14] Georgiev, T., Intwala, C., Babakan, S., and Lumsdaine, A., "Unified frequency domain analysis of lightfield cameras," in *Proceedings of the 10th European Conference on Computer Vision*, 224–237 (2008).
- [15] Bracewell, R. N., *[The Fourier Transform and Its Applications]*, McGraw-Hill, New York, 3rd ed. (1999).
- [16] Goodman, J. W., *[Introduction to Fourier Optics]*, Roberts and Company Publishers, Englewood, Colorado, 3rd ed. (2004).
- [17] Levin, A., Freeman, W. T., and Durand, F., "Understanding camera trade-offs through a Bayesian analysis of light field projections," in *Proceedings of the 10th European Conference on Computer Vision*, 88–101 (2008).
- [18] Xu, Z. and Lam, E. Y., "A spatial projection analysis of light field capture," in *Frontiers in Optics*, FWH2, Optical Society of America, Washington, D.C. (2010).
- [19] Bajwa, W. U., Haupt, J. D., Raz, G. M., Wright, S. J., and Nowak, R. D., "Toeplitz-structured compressed sensing matrices," in *Proceedings of IEEE/SP 14th Workshop on Statistical Signal Processing*, 294–298 (2007).
- [20] Yin, W., Morgan, S., Yang, J., and Zhang, Y., "Practical compressive sensing with Toeplitz and circulant matrices," *Proc. SPIE* **7744**, 77440K (2010).
- [21] Xu, Z. and Lam, E. Y., "Image reconstruction using spectroscopic and hyperspectral information for compressive terahertz imaging," *J. Opt. Soc. Am. A* **27**(7), 1638–1646 (2010).
- [22] "The (new) Stanford light field archive." <http://lightfield.stanford.edu/lfs.html>.