

# Lightfield Recovery from Its Focal Stack

F. Pérez<sup>1</sup> · A. Pérez<sup>1</sup> · M. Rodríguez<sup>2</sup> · E. Magdaleno<sup>2</sup>

Received: 24 July 2015 / Accepted: 20 April 2016 / Published online: 23 May 2016  
© Springer Science+Business Media New York 2016

**Abstract** The Focal Stack Transform integrates a 4D lightfield over a set of appropriately chosen 2D planes. The result of such integration is an image focused on a determined depth in 3D space. The set of such images is the Focal Stack of the lightfield. This paper studies the existence of an inverse for this transform. Such inverse could be used to obtain a 4D lightfield from a set of images focused on several depths of the scene. In this paper, we show that this inversion cannot be obtained for a general lightfield and introduce a subset of lightfields where this inversion can be computed exactly. We examine the numerical properties of such inversion process for general lightfields and examine several regularization approaches to stabilize the transform. Experimental results are provided for focal stacks obtained from several plenoptic cameras. From a practical point of view, results show how this inversion procedure can be used to recover, compress, and denoise the original 4D lightfield.

**Keywords** Lightfield · Focal Stack · Plenoptic · Inverse problems · Regularization

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s10851-016-0658-4](https://doi.org/10.1007/s10851-016-0658-4)) contains supplementary material, which is available to authorized users.

---

✉ F. Pérez  
fdoperez@ull.es

<sup>1</sup> Departamento de Estadística, Investigación Operativa y Computación, University of La Laguna, San Cristóbal de La Laguna, Canary Islands, Spain

<sup>2</sup> Departamento de Física Fundamental, Experimental, Electrónica y Sistemas, University of La Laguna, San Cristóbal de La Laguna, Canary Islands, Spain

## 1 Introduction

Computational photography is based on capturing and processing discrete samples of all the light rays in the 3D space [1]. These light rays are described by the 7D plenoptic function [2] in terms of position, orientation, spectral content, and time. This can be simplified to a 4D lightfield [3,4] by considering only the value of each ray as a function of its position and orientation in a static scene, and by constraining each ray to have the same value at every point along its direction of propagation. Compared to conventional photography, which captures 2D images, computational photography captures the entire 4D lightfield. This 4D lightfield extends the capabilities of current commercial cameras with new features such as refocusing [5–7], 3D depth estimation [8–10] or super-resolution [11–13].

There are several methods to obtain a 4D lightfield. In this paper, we will use a plenoptic camera [14]. A plenoptic camera uses a microlens array inside a conventional camera body to measure the intensity and direction of light rays. There are several commercial models of plenoptic cameras available [15,16]. Since the complete 4D lightfield inside the camera is recorded, new 2D photographs can be computed by tracing and projecting all the rays to desired virtual imaging planes. This so-called digital refocusing enables to change focus after the picture is taken, as well as extending the depth of field without decreasing the aperture.

The set of refocused images is called the focal stack of the scene. A particular image in the focal stack can be modeled through the continuous Photography integral operator [3]. Since the plenoptic image is discrete, to obtain a refocused photograph it is necessary to discretize the Photography operator. Then, it is necessary to solve two problems: to interpolate the lightfield and to approximate the integration process. There are several methods to perform this discretiza-

tion: the brute-force approach interpolates the lightfield by nearest neighbor and approximates the integral by sums. If we assume that the plenoptic camera is composed of  $n \times n$  microlenses and that each microlens generates a  $n \times n$  image, the brute-force approach would need  $O(n^4)$  operations to generate a refocused photograph. A significant improvement to this performance was obtained in [5] with the “Fourier Slice Photography” (FSP) technique that reformulates the problem in the Fourier domain. The FSP method is based on the extraction of an appropriate dilated 2D slice in the 4D Fourier transform of the lightfield. In this case, interpolation takes place in the 4D Fourier domain using a 4D Kaiser–Bessel filter. This decreases the computational cost of the discretization of the Photography operator to  $O(n^4 \log(n))$  operations to generate a focal stack with  $O(n^2)$  refocused photographs. Another method that uses the frequency domain is the Discrete Focal Stack Transform (DFST) that uses 4D trigonometric interpolation in the spatial domain and a discretization of the integral operator [6]. The Discrete Focal Stack Transform is fast [17] and can be computed in  $O(n^4 \log(n))$  for a  $n^4$  lightfield. It also can be described analytically.

This paper studies the existence of an inverse for the Focal Stack transform. Such inverse should take the focal stack of a scene as its input and produce a 4D lightfield as its output. This would make a 4D lightfield equivalent to its focal stack and new procedures for capturing lightfields based on refocused images from standard conventional cameras could be developed. Such inverse could also be used for lightfield compression purposes since the focal stack images are highly redundant. In this paper, we show that this is not the case in general. We introduce the concept of  $\Pi$ -constant lightfields, which includes the Lambertian lightfields, where this inversion is theoretically possible. We also study the numerical properties of this inversion procedure. Finally, we provide some experimental results that show that the lightfield captured from a plenoptic camera can be well approximated using a  $\Pi$ -constant lightfield and we study the performance of the inverse for general lightfields using numerical inversion with regularization.

This paper is divided in five sections. In Sect. 2 we describe the algorithms to compute the Focal Stack in the frequency domain that are based on the Fourier Slice technique in [5, 6]. In Sect. 3, we study the inversion of the Focal Stack transform and discuss its properties. Section 4 contains some experimental results and Sect. 5 includes conclusions and future work.

## 2 Previous Work

Conventional 2D photographs are obtained in plenoptic cameras through the Photography transform. This transform

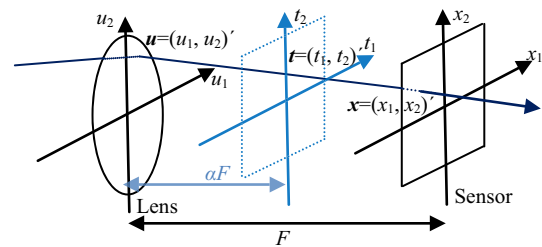


Fig. 1 Two plane parameterization of the lightfield

takes a lightfield as its input and generates a photograph focused on a determined plane [5, 18]. The Photography transform will be defined by means of a two-plane parameterization of the lightfield inside the plenoptic camera. We will write  $L_F(\mathbf{x}, \mathbf{u})$  as the radiance traveling from position  $\mathbf{u} = (u_1, u_2)^T$  on a domain defined by the aperture of the lens plane to position  $\mathbf{x} = (x_1, x_2)^T$  over the sensor extent.  $F$  is the distance between the lens and the sensor (see Fig. 1 adapted from [5]).

To compute conventional photographs, the sensor plane is virtually placed at any distance  $\alpha F$  inside the camera (see Fig. 1). Let  $\mathcal{P}[L_F]$  be the operator that transforms a lightfield  $L_F$  at a sensor depth  $F$  into a photograph formed at sensor depth  $\alpha F$ . The operator can be formulated as

$$\mathcal{P}_\alpha[L_F](\mathbf{t}) = \frac{1}{\alpha^2 F^2} \int L_F\left(\mathbf{u}\left(1 - \frac{1}{\alpha}\right) + \mathbf{t}, \mathbf{u}\right) d\mathbf{u}. \quad (1)$$

This equation shows how to compute a photograph  $\mathcal{P}_\alpha[L_F]$  formed at a virtual sensor plane that is located at distance  $\alpha F$  from the lens plane. Pixels in the 2D photograph depend on the spatial variable  $\mathbf{t} = (t_1, t_2)^T$  as seen on Fig. 1. When we compute the photographs for every virtual sensor distance  $\alpha F$ , we obtain the Focal Stack transform  $\mathcal{S}$  of the lightfield.

To render the 2D photograph, we usually want the image taken from a constant lightfield to be a constant independent of  $\alpha$ , so we normalize the Photography Operator removing the  $1/(\alpha^2 F^2)$  term. Also, in order to make the discretization of the focal stack easier, we reparametrize the focusing distance using  $(1 - 1/\alpha) = q$  leading to the normalized Photography operator defined as

$$P_q[L](\mathbf{t}) = \int L(\mathbf{u}q + \mathbf{t}, \mathbf{u}) d\mathbf{u}. \quad (2)$$

In practice, the normalized Photography operator cannot be used since a plenoptic camera only captures discrete samples of the lightfield. To discretize  $P_q[L]$ , we need to resample the lightfield and to approximate the integration process.

After defining the basics of the Photography transform, we introduce in Table 1 a summary of the symbols that will be used in the paper.

**Table 1** Notation table

$L(\mathbf{x}, \mathbf{u})$	A 4D lightfield
$\mathbf{x}$	Spatial 2D variable for the lightfield images
$\mathbf{u}$	Spatial 2D variable for the microlenses
$F$	Distance between the lens plane and the sensor plane
$\alpha$	The virtual sensor plane for refocusing is placed at $\alpha F$
$q$	Change of variables, $q = (1 - 1/\alpha)$
$P_q[L](\mathbf{t})$	A photograph at depth $q$ and spatial position $\mathbf{t}$
$\mathbf{t}$	Spatial 2D variable for the photograph
$\hat{L}(\boldsymbol{\xi}, \boldsymbol{\eta})$	The Fourier transform of the lightfield
$\boldsymbol{\xi}$	2D frequency counterpart of $\mathbf{x}$
$\boldsymbol{\eta}$	2D frequency counterpart of $\mathbf{u}$
$\mathcal{S}[L](\mathbf{t}, q)$	Focal stack image at depth $q$ and spatial position $\mathbf{t}$
$\Pi[L](\mathbf{t}, q)$	Back-projection operator at depth $q$ and spatial position $\mathbf{t}$
$\hat{\Pi}[L](\boldsymbol{\xi}, \boldsymbol{\kappa})$	Fourier transform of the back-projection operator at depth frequency $\boldsymbol{\kappa}$ and spatial frequency $\boldsymbol{\xi}$
$\delta$	Dirac's delta
$N_x \times N_x$	Number of microlenses
$N_u \times N_u$	Number of pixels behind each microlens
$n_x$	Defined from $N_x = 2n_x + 1$
$n_u$	Defined from $N_u = 2n_u + 1$
$n_q$	Defined from $n_q = 2n_x n_u, N_q = 2n_q + 1$
$N_q$	Defined from $N_q = 2n_q + 1$
$\varphi$	An arbitrary function
$W(\boldsymbol{\xi})$	The set $\{\boldsymbol{\kappa} / \hat{\Pi}[L](\boldsymbol{\xi}, \boldsymbol{\kappa}) \neq 0\}$
$D_k$	Finite difference approximation to the $k$ th derivative
$\hat{D}_k$	Fourier transform of the finite difference approximation to the $k$ th derivative
$\nabla L(\mathbf{x}, \mathbf{u})$	Gradient of the lightfield
$\lambda, \beta, \gamma$	Regularization parameters

### 2.1 The Fourier Slice Theorem

A simple discretization of  $P_q[L]$  could be done by resampling the lightfield through local interpolation and replacing the integration process with sums. This is the brute-force approach that needs  $O(n^4)$  operations to generate a single refocused photograph of size  $n^2$ . A significant improvement in the computational complexity can be achieved through the Fourier Slice Photography technique that studies the Photography operator in the frequency domain. The main result is the Fourier Slice Photography Theorem [5], which states that in the frequency domain, a photograph is a 2D slice of the

4D lightfield. Then, photographs focused at different depths correspond to slices at different slopes in the 4D space. We may write the Fourier Slice Photography Theorem as

$$P_q[L](\mathbf{t}) = \int \hat{L}(\boldsymbol{\xi}, -q\boldsymbol{\xi}) e^{2\pi i \mathbf{t} \cdot \boldsymbol{\xi}} d\boldsymbol{\xi}, \tag{3}$$

where  $\boldsymbol{\xi}$  is the 2D spatial frequency, the dot symbol  $\cdot$  is the scalar product, and  $\hat{L}$  is the Fourier transform of the lightfield:

$$\hat{L}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \int \int L(\mathbf{x}, \mathbf{u}) e^{-2\pi i(\mathbf{x} \cdot \boldsymbol{\xi} + \mathbf{u} \cdot \boldsymbol{\eta})} d\mathbf{x} d\mathbf{u}. \tag{4}$$

Therefore, to obtain a photograph  $P_q[L](\mathbf{t})$  we have to compute first the Fourier transform of the lightfield  $\hat{L}$  using the 4D integral in (4). After that, we have to evaluate the slice  $\hat{L}(\boldsymbol{\xi}, -q\boldsymbol{\xi})$  and finally we have to compute the inverse Fourier transform of the slice. It only remains to specify how to resample in the frequency domain as required by the slicing operator. In the Fourier Slice Photography technique, this resampling process is carried out with the Kaiser-Bessel filter, to minimize aliasing. The algorithm generates  $O(n^2)$  photographs of size  $n^2$  with a computational complexity of  $O(n^4 \log(n))$  and obtains a significant improvement in computational complexity over the direct approach.

### 2.2 The Discrete Focal Stack Transform

The Discrete Focal Stack transform [6] is an alternative to the Fourier Slice Photography technique based on trigonometric interpolation. The Discrete Focal Stack Transform uses the following decomposition of (2) [17]:

$$P_q[L](\mathbf{t}) = \int \int L(\mathbf{x}, \mathbf{u}) \delta(\mathbf{x} - \mathbf{u}q - \mathbf{t}) d\mathbf{x} d\mathbf{u} \tag{5}$$

and the  $\delta$  representation in the Fourier domain:

$$\delta(\mathbf{x}) = \int e^{-2\pi i \boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \tag{6}$$

The combination of (5) and (6) leads to

$$P_q[L](\mathbf{t}) = \int \left( \int \int L(\mathbf{x}, \mathbf{u}) e^{-2\pi i(\mathbf{x} - \mathbf{u}q) \cdot \boldsymbol{\xi}} d\mathbf{x} d\mathbf{u} \right) e^{2\pi i \boldsymbol{\xi} \cdot \mathbf{t}} d\boldsymbol{\xi} \tag{7}$$

Then, we obtain the following decomposition of the focal stack  $\mathcal{S}$  using the Fourier transform of  $\mathcal{S}[L](\mathbf{t}, q)$  with

respect to  $q$

$$\begin{aligned} \mathcal{S}[L](t, q) &= P_q[L](t) \\ &= \iiint \left( \iint \hat{L}(\xi, \mathbf{u}) e^{2\pi i(\mathbf{u}\cdot\xi - \kappa)q} dq d\mathbf{u} \right) e^{2\pi i\xi\cdot t} e^{2\pi i q\kappa} d\xi d\kappa \\ &= \iint \left( \int \hat{L}(\xi, \mathbf{u}) \delta(\mathbf{u}\cdot\xi - \kappa) d\mathbf{u} \right) e^{2\pi i\xi\cdot t} e^{2\pi i q\kappa} d\xi d\kappa \end{aligned} \tag{8}$$

Therefore, defining the back-projection operator  $\Pi$  in the frequency domain as

$$\hat{\Pi}[L](\xi, \kappa) = \int \hat{L}(\xi, \mathbf{u}) \delta(\mathbf{u}\cdot\xi - \kappa) d\mathbf{u}, \tag{9}$$

we obtain

$$\mathcal{S}[L](t, q) = \int \int \hat{\Pi}[L](\xi, \kappa) e^{2\pi i\xi\cdot t} e^{2\pi i q\kappa} d\xi d\kappa, \tag{10}$$

where variables  $t, q$  are in the spatial domain and variables  $\xi, \kappa$  are in the frequency domain.

The Fast Discrete Focal Stack transform discretizes (2) using trigonometric interpolation to obtain a discrete counterpart to (10). For a discretized lightfield composed of  $N_x \times N_x$  microlenses of size  $N_u \times N_u$  with  $N_x = 2n_x + 1, N_u = 2n_u + 1$ , the photograph  $P_q[L](t)$  is computed as

$$\begin{aligned} P_q[L](t) &= \sum_{\mathbf{u}=-n_u}^{n_u} L(\mathbf{u}q + t, \mathbf{u}) \\ &= \frac{1}{N_x} \sum_{\mathbf{u}=-n_u}^{n_u} \sum_{\xi=-n_x}^{n_x} \hat{L}(\xi, \mathbf{u}) e^{2\pi i \frac{(\mathbf{u}q+t)\cdot\xi}{N_x}} \\ &= \frac{1}{N_x} \sum_{\xi=-n_x}^{n_x} e^{2\pi i \frac{t\cdot\xi}{N_x}} \sum_{\mathbf{u}=-n_u}^{n_u} \hat{L}(\xi, \mathbf{u}) e^{2\pi i \frac{q(\mathbf{u}\cdot\xi)}{N_x}} \end{aligned}$$

Now, for  $n_q = 2n_x n_u, N_q = 2n_q + 1$

$$\begin{aligned} e^{2\pi i \frac{q(\mathbf{u}\cdot\xi)}{N_x}} &= \sum_{\kappa=-n_q}^{n_q} \delta_{N_q}(\mathbf{u}\cdot\xi - \kappa) e^{2\pi i \frac{\kappa q}{N_x}}, \delta_{N_q}(x) \\ &= \frac{1}{N_q} \sum_{\kappa=-n_q}^{n_q} e^{2\pi i \frac{\kappa x}{N_q}} \end{aligned}$$

So

$$\begin{aligned} \mathcal{S}[L](t, q) &= P_q[L](t) \\ &= \frac{1}{N_x} \sum_{\xi=-n_x}^{n_x} \sum_{\kappa=-n_q}^{n_q} \left( \sum_{\mathbf{u}=-n_u}^{n_u} \hat{L}(\xi, \mathbf{u}) \delta_{N_q}(\mathbf{u}\cdot\xi - \kappa) \right) e^{2\pi i \frac{\kappa q}{N_x}} e^{2\pi i \frac{t\cdot\xi}{N_x}}, N_x = N_x^2 \end{aligned} \tag{11}$$

Then, defining the discrete back-projection operator  $\Pi$  in the frequency domain as

$$\hat{\Pi}[L](\xi, \kappa) = \sum_{\mathbf{u}=-n_u}^{n_u} \hat{L}(\xi, \mathbf{u}) \delta_{N_q}(\mathbf{u}\cdot\xi - \kappa) \tag{12}$$

we obtain the discrete version of (10)

$$\begin{aligned} \mathcal{S}[L](t, q) &= P_q[L](t) \\ &= \frac{1}{N_x} \sum_{\xi=-n_x}^{n_x} \sum_{\kappa=-n_q}^{n_q} \hat{\Pi}[L](\xi, \kappa) e^{2\pi i \frac{\kappa q}{N_x}} e^{2\pi i \frac{t\cdot\xi}{N_x}} \end{aligned} \tag{13}$$

The Fast Discrete Focal Stack transform [17] computes the back-projections in  $O(n^4)$ . The inverse Fourier transform of the back-projections are computed in  $O(n^4 \log n)$  to obtain the final computational complexity of the DFST.

### 3 The Inversion of the Discrete Focal Stack Transform

In this section, we study the inversion of the Discrete Focal Stack Transform both in the continuous and discrete case. We introduce the concept of  $\Pi$ -constant lightfields that includes the Lambertian lightfields, where this inversion is theoretically possible. We also study the numerical properties of this inversion procedure and propose several regularization methods to stabilize the transform.

#### 3.1 The Continuous Case

Our main result is to make explicit what information of the lightfield can be exactly recovered from the focal stack.

**Theorem** *Given a lightfield  $L$  and the back-projection operator:*

$$\hat{\Pi}[L](\xi, \kappa) = \int \hat{L}(\xi, \mathbf{u}) \delta(\mathbf{u}\cdot\xi - \kappa) d\mathbf{u}, \tag{14}$$

*then, the focal stack  $S[L]$  and  $\Pi[L]$  can be obtained from each other using:*

$$\hat{\Pi}[L] = \hat{S}[L] \tag{15}$$

This result shows that the back-projection operator  $\hat{\Pi}[L]$  is the Fourier transform of the focal stack  $S[L]$  and conversely that the focal stack is the inverse Fourier transform of  $\hat{\Pi}[L]$ . The result easily follows from (10).

An immediate consequence of the preceding theorem is that the Focal Stack transform is not invertible in general. For

example, if we take  $L(\mathbf{x}, \mathbf{u}) = u_1$  for  $-1/2 \leq u_1 \leq 1/2$  and 0 elsewhere then we have  $\hat{L}(\xi, \mathbf{u}) = \delta(\xi) u_1$  so  $\hat{\Pi}(\xi, \kappa) = \delta(\xi) \delta(\kappa) \int u_1 d\mathbf{u} = 0$ . Since  $-L(\mathbf{x}, \mathbf{u})$  gives the same value for  $\hat{\Pi}(\xi, \kappa)$  both lightfields generate the same focal stack and the Focal Stack transform is not invertible.

Although the Focal Stack transform is not invertible in general, there are special cases of lightfields that can be inverted. The simplest invertible lightfields are those for which  $\hat{L}(\xi, \mathbf{u})$  is constant along all planes  $\xi \cdot \mathbf{u} = \kappa$ . We will call them  $\Pi$ -constant. Then, for a  $\Pi$ -constant lightfield there exists a function  $\varphi$ , such that the relation  $\hat{L}(\xi, \mathbf{u}) = \varphi(\xi, \xi \cdot \mathbf{u})$  holds. Then

$$\hat{\Pi}[L](\xi, \kappa) = \varphi(\xi, \kappa) \int \delta(\mathbf{u} \cdot \xi - \kappa) d\mathbf{u}. \tag{16}$$

Note that for a  $\Pi$ -constant lightfield we can directly recover the lightfield from the back-projections. Renaming variable  $\mathbf{u}$  to  $\mathbf{v}$  and substituting  $\kappa = \xi \cdot \mathbf{u}$  in (16) we have

$$\hat{\Pi}[L](\xi, \xi \cdot \mathbf{u}) = \hat{L}(\xi, \mathbf{u}) \int \delta(\xi \cdot \mathbf{v} - \xi \cdot \mathbf{u}) d\mathbf{v}. \tag{17}$$

Then, using (15) and (16) for a  $\Pi$ -constant lightfield we have

$$\hat{S}[L](\xi, \kappa) = \varphi(\xi, \kappa) \int \delta(\mathbf{u} \cdot \xi - \kappa) d\mathbf{u}. \tag{18}$$

Finally, substituting  $\kappa$  for  $\xi \cdot \mathbf{u}$

$$\hat{L}(\xi, \mathbf{u}) = \frac{\hat{S}[L](\xi, \xi \cdot \mathbf{u})}{\int \delta(\xi \cdot (\mathbf{v} - \mathbf{u})) d\mathbf{v}} \tag{19}$$

where for example in a square aperture region  $-n_u \leq v_i \leq n_u$  we have for  $\xi = (\xi_1, \xi_2)^T$

$$\int \delta(\xi \cdot (\mathbf{v} - \mathbf{u})) d\mathbf{v} = \begin{cases} \frac{2n_u}{\max(|\xi_2|, |\xi_1|)} + \frac{\kappa+l}{|\xi_2||\xi_1|} & -m < \kappa < -l \\ \frac{2n_u}{\max(|\xi_2|, |\xi_1|)} & -l \leq \kappa \leq m \\ \frac{2n_u}{\max(|\xi_2|, |\xi_1|)} - \frac{\kappa-l}{|\xi_2||\xi_1|} & l < \kappa < m \end{cases} \tag{20}$$

$\kappa = \xi \cdot \mathbf{u}, l = n_u \left( |\xi_2| - |\xi_1| \right), m = n_u \left( |\xi_2| + |\xi_1| \right)$

Note that (19) can also be written as

$$\hat{L}(\xi, \mathbf{u}) = \frac{\hat{S}[L](\xi, \xi \cdot \mathbf{u})}{\hat{S}[\delta](\xi, \xi \cdot \mathbf{u})} \tag{21}$$

where  $\delta$  is the lightfield  $L(\mathbf{x}, \mathbf{u}) = \delta(\mathbf{x})$ . Therefore, in order to recover the original lightfield, the image from the focal stack  $\hat{S}[L](\xi, \xi \cdot \mathbf{u})$  has to be deconvolved with the kernel  $\hat{S}[\delta](\xi, \xi \cdot \mathbf{u})$ .

### 3.1.1 Lambertian Lightfields

An important subset of lightfields is Lambertian lightfields, where the radiance emitted in the scene is independent of the direction. We now show that Lambertian lightfields are  $\Pi$ -constant and therefore invertible. The proof follows from the fact that if  $L$  is a Lambertian lightfield, there exists functions  $\lambda(\mathbf{x}), \phi(\mathbf{x})$  such that for all  $\mathbf{x}, \mathbf{u}$ , the lightfield  $L$  is uniquely determined from the equation  $L(\mathbf{x} + \lambda(\mathbf{x})\mathbf{u}, \mathbf{u}) = \phi(\mathbf{x})$ . Then, for a fixed  $\mathbf{u}$   $\hat{L}(\xi, \mathbf{u})$  is determined through the integral equation

$$\int \hat{L}(\xi, \mathbf{u}) e^{2\pi i \xi \cdot (\mathbf{x} + \lambda(\mathbf{x})\mathbf{u})} d\xi = \phi(\mathbf{x}) \tag{22}$$

The preceding equation is exactly the same substituting  $\mathbf{u}$  for  $\mathbf{v} = \mathbf{u} + \mu \xi^\perp$ , with  $\xi^\perp$  a vector orthogonal to  $\xi$ . Then, for all  $\mathbf{u}$  and  $\mu$  the relation:  $\hat{L}(\xi, \mathbf{u}) = \hat{L}(\xi, \mathbf{u} + \mu \xi^\perp)$  holds. This means that we can remove the  $\xi^\perp$  direction without changing the value of  $\hat{L}(\xi, \mathbf{u})$ . Now, since vector  $\mathbf{u}$  can be decomposed as  $\mathbf{u} = (\xi/\|\xi\|^2)(\xi \cdot \mathbf{u}) + (\xi^\perp/\|\xi^\perp\|^2)(\xi^\perp \cdot \mathbf{u})$ , we can write  $\hat{L}(\xi, \mathbf{u})$  as  $\hat{L}(\xi, \mathbf{u}) = \hat{L}(\xi, (\xi/\|\xi\|^2)(\xi \cdot \mathbf{u}) + (\xi^\perp/\|\xi^\perp\|^2)(\xi^\perp \cdot \mathbf{u}))$ . Removing the  $\xi^\perp$  direction we have  $\hat{L}(\xi, \mathbf{u}) = \hat{L}(\xi, (\xi/\|\xi\|^2)(\xi \cdot \mathbf{u}))$ . So  $\hat{L}(\xi, \mathbf{u})$  depends on  $\xi$  and  $\xi \cdot \mathbf{u}$  and there exists a function  $\varphi$  such that  $\hat{L}(\xi, \mathbf{u}) = \varphi(\xi, \xi \cdot \mathbf{u})$ . Therefore, Lambertian lightfields are  $\Pi$ -constant and invertible.

### 3.2 The Discrete Case

To study the discrete case, we will use the Discrete Focal Stack Transform to obtain a result analogous to (15). Using the discrete back-projections in (12), we have

$$\frac{1}{N_q} \hat{S}[L] \left( \mathbf{t}, q \frac{N_x}{N_q} \right) (\xi, \kappa) = \hat{\Pi}[L](\xi, \kappa) \tag{23}$$

and

$$\Pi[L](\mathbf{t}, q) = \frac{1}{N_q} \mathcal{S}[L] \left( \mathbf{t}, q \frac{N_x}{N_q} \right) \tag{24}$$

The result follows from (13). If the lightfield is  $\Pi$ -constant we have the following relations in the discrete case that are the analogous to (17), (19), (21)

$$\hat{\Pi}[L](\xi, \mathbf{u} \cdot \xi) = \hat{L}(\xi, \mathbf{u}) \sum_{v=-n_u}^{n_u} \delta_{N_q}(v \cdot \xi - \mathbf{u} \cdot \xi) \tag{25}$$

$$\hat{L}(\xi, \mathbf{u}) = \frac{\hat{S}[L] \left( \mathbf{t}, q \frac{N_x}{N_q} \right) (\xi, \xi \cdot \mathbf{u})}{N_q \sum_{\mathbf{u}=-n_u}^{n_u} \delta_{N_q}(\xi \cdot \mathbf{v} - \xi \cdot \mathbf{u})} \tag{26}$$

$$\hat{L}(\xi, \mathbf{u}) = \frac{\hat{S}[L](\mathbf{t}, q \frac{N_x}{N_q})(\xi, \xi \cdot \mathbf{u})}{\hat{S}[\delta](\mathbf{t}, q \frac{N_x}{N_q})(\xi, \xi \cdot \mathbf{u})} \quad (27)$$

where  $\delta$  is a lightfield in case when  $L(\mathbf{x}, \mathbf{u}) = \delta_{N_x}(\mathbf{x})$ . Similarly to the continuous case, the recovery of the lightfield can also be expressed as a deconvolution problem.

### 3.3 Direct Focal Stack Inversion

Using the result in (27) in the previous section, we have derived a formula analogous to (21) that can be used for the inversion of the lightfield if we assume that the lightfield is  $\Pi$ -constant. However, the practical use of this inversion formula for focal stacks obtained from plenoptic cameras is very limited. In first place, we need a huge number of images in the focal stack to use (26), (27). For a typical plenoptic camera with  $255 \times 255$  microlenses and  $9 \times 9$  pixels behind each microlens, we need  $N_q = 2n_q + 1 = 2033$  images in the focal stack to use the inversion formula. Another problem that arises is the range of the slope variable  $q$ . To avoid aliasing in the images of the focal stack, in a plenoptic camera we need  $|qN_x/N_q| \leq 1$ . In our case, this leads to  $|q| \leq (N_q/N_x) \approx N_u$  (in the preceding example this would give 17 images) while the range of  $q$  in (26), (27) is  $|q| \leq n_q$ . To summarize, for the lightfield example, we set before using a plenoptic camera we will typically have 17 unaliased planes out of 2033 in the focal stack and the rest will be aliased so the inversion formula in (26), (27) is not practical.

### 3.4 Partial Focal Stack Inversion

To obtain a practical procedure, we need to invert the focal stack using a reduced number of its unaliased images. In this section, we show how this can be done. In the discrete case, we have for arbitrary  $q$  (11):

$$\hat{S}[L](\mathbf{t}, q)(\xi) = \sum_{\kappa=-n_q}^{n_q} \hat{\Pi}[L](\xi, \kappa) e^{2\pi i \frac{\kappa q}{N_x}} \quad (28)$$

Defining the set:  $W(\xi) = \{\kappa / \hat{\Pi}[L](\xi, \kappa) \neq 0\}$ , the equation in (28) is equivalent to

$$\hat{S}[L](\mathbf{t}, q)(\xi) = \sum_{\kappa \in W(\xi)} \hat{\Pi}[L](\xi, \kappa) e^{2\pi i \frac{\kappa q}{N_x}} \quad (29)$$

Using (29), we see that for each fixed  $\xi$  we obtain a system of linear equations where  $\hat{S}[L]$  are the known images from the focal stack, the back-projections  $\hat{\Pi}[L]$  are the unknowns, and both are connected through the coefficients  $e^{2\pi i \frac{\kappa q}{N_x}}$ . The linear system can be written for several  $q$  values and  $\kappa \in$

$W(\xi)$  in matrix form as

$$\begin{aligned} \hat{s} &= \mathbf{G}\hat{\pi}, \hat{s}_i = \hat{S}[L](\mathbf{t}, q_i)(\xi), \mathbf{G}_{i,j} = e^{2\pi i \frac{\kappa_j q_i}{N_x}}, \hat{\pi}_i \\ &= \hat{\Pi}[L](\xi, \kappa_i) \text{ with } -m \leq i \leq m, -n \leq j \leq n, 2m+1 \\ &= N_u^2, 2n+1 = \text{cardinal}(W(\xi)). \end{aligned} \quad (30)$$

Then, we obtain the following result:

**Proposition** For each  $\xi$ , the back-projection values  $\hat{\Pi}[L](\xi, \kappa)$  for  $\kappa \in W(\xi)$  can be obtained from at most  $N_u^2$  images in the focal stack with slopes  $q_i = \frac{i}{n}$ ,  $n \geq N_q/N_x$ ,  $i = -m \dots m$ ,  $2m+1 = N_u^2$ .

The proof follows from the fact that the coefficient matrix  $\mathbf{G}$  of the linear system in (30) is of Vandermonde type. Note that the preceding proposition also allows us to recover the complete 4D lightfield for a  $\Pi$ -constant lightfield using a focal stack with  $N_u^2$  images. The result is immediate since a  $\Pi$ -constant lightfield can be directly recovered from the back-projections  $\Pi$  (25). For the typical lightfield, we set before this would lead to a set of 81 images from the focal stack. Those  $N_u^2$  images can be taken from the non-aliasing range in the focal stack from a plenoptic camera taking  $n = m$  in the previous result since  $N_q/N_x \approx N_u$ .

A relevant fact about the linear system in (30) is that the minimal number of equations to determine a particular frequency  $\xi$  may be much lower than  $N_u^2$ . For example, the linear system in (30) for  $\xi = \mathbf{0}$  is  $\hat{S}[L](\mathbf{0}, q) = \hat{\Pi}[L](\mathbf{0}, 0)$ . Then, to determine  $\hat{\Pi}[L](\mathbf{0}, 0)$  only one equation is needed and the rest are redundant. An example of linear system where the maximum number of equations is needed happens if  $\xi = (\xi_1, \xi_2)$  and  $\xi_1, \xi_2$  are coprime and  $|\xi_1| > N_u$ ,  $|\xi_2| > N_u$ . Therefore, the effective volume of data to store the back-projections is lower than the volume of data of the lightfield. For the typical lightfield, we set before with  $255 \times 255$  microlenses and  $9 \times 9$  pixels behind each microlens, the size needed to store the back-projections is 96.51 % of the original lightfield.

### 3.5 Numerical Analysis and Regularization

In the preceding section, it was shown that it is possible to recover the lightfield from a partial section of the focal stack under the  $\Pi$ -constant assumption. However, there are several problems with this approach from a numerical point of view. As shown in the previous section, if we use  $N_u^2$  images some frequencies will have overdetermined linear systems. Therefore, a minimal amount of noise would make the inversion fail. Of course, we could remove the redundant equations to make the system determined, but this procedure will not use all the data available. Instead of directly using (30), to cope with image noise we could compute  $\hat{\Pi}[L](\xi, k)$  for

$k \in W(\xi)$  as the solution of

$$\hat{\pi}_{\text{opt}} = \arg \min_{\hat{\pi}} \|\hat{s} - G\hat{\pi}\|^2, \tag{31}$$

leading to

$$\hat{\pi}_{\text{opt}} = (G^T G)^{-1} G^T \hat{s}, \quad (G^T G)_{ij} = \sum_l e^{2\pi i \frac{(\kappa_i - \kappa_j) q_l}{N_x}}, \tag{32}$$

where  $(G^T G)^{-1} G^T$  is the Moore–Penrose pseudo-inverse of  $G$ .

The structure of the linear systems in (30), (32) poses another problem since it is a Vandermonde system that is ill-conditioned [19]. Therefore, additional information must be provided to obtain a solution. This process is called regularization [20]. There are different formulations that can be used to describe the solution of the regularized problem. In general, some regularization measurement  $\text{Reg}$  is used to promote desired solutions. A common choice for the regularization measurement is

$$\text{Reg}(L) = \|D_k L\|_p^p, \quad 1 \leq p \leq \infty \tag{33}$$

where  $D_k$  is a finite difference approximation to the  $k$ th derivative with  $D_0 = I$ , the identity operator, by convention. The most common choices of  $p$  are  $p = 1$  and  $p = 2$ .

### 3.5.1 $L_2$ Regularization

If  $p = 2$  in (33) the regularization is called  $L_2$  or Tikhonov regularization [21] so  $\text{Reg}(L) = \|D_k L\|_2^2$ . In order to use  $L_2$  regularization we rewrite (28) as

$$\hat{S}[L](t, q)(\xi) = \sum_{u=-n_u}^{n_u} \hat{L}(\xi, u) e^{2\pi i \frac{q(u-\xi)}{N_x}} \tag{34}$$

That is a collection of linear equations with unknowns  $\hat{L}(\xi, u)$  for each  $\xi$  when  $q$  varies. Defining the product of two operators:  $A(\alpha, \beta)$ ,  $B(\lambda, \mu)$  as  $AB(\alpha, \mu) = \sum_{\beta} A(\alpha, \beta) B(\beta, \mu)$ , (34) can be written as

$$\begin{aligned} \hat{s} &= G\hat{l}, \hat{s}_q = \hat{S}[L](t, q)(\xi), G(q, u) = e^{2\pi i \frac{q(u-\xi)}{N_x}}, \hat{l}(u) \\ &= \hat{L}(\xi, u) \end{aligned} \tag{35}$$

We will use  $k = 0$  in (35) and define the solution to the regularized linear system through the following optimization problem:

$$\hat{L}(\xi, u)_{\text{opt}} = \arg \min_{\hat{l}} \|\hat{s} - G\hat{l}\|^2 + \lambda \|\hat{l}\|^2 \tag{36}$$

where  $\lambda$  is a regularization parameter,  $\|\hat{s} - G\hat{l}\|^2$  is the data-fit term, and  $\text{Reg}(L) = \|\hat{l}\|^2$  is the regularization measurement that enforces smooth solutions. The normal equations for the preceding system are

$$(G^T G + \lambda I) \hat{l} = G^T \hat{s} \tag{37}$$

with

$$G^T G(u, v) = \sum_q e^{2\pi i \frac{q(u-v)\xi}{N_x}}, I(u, v) = 1 \text{ if } u = v \text{ and } I(u, v) = 0 \text{ in other case.}$$

The system in (37) has  $N_u^2$  equations and  $N_u^2$  unknowns. The value for  $\lambda$  has to be determined beforehand. Different choices of the parameter  $\lambda$  results in a tradeoff between the smoothness of the solution and a good data fit. There are several techniques to select the regularization parameter including the discrepancy principle, L-curve or Generalized Cross Validation [22]. Since we have  $N_x^2$  linear systems with  $N_u^2$  unknowns, the total computational cost for light-field inversion is  $O(N_x^2 N_u^6)$ . The  $L_2$  regularization technique tends to produce a smoothing effect on the solution so other regularization approaches can also be employed [20] as we will see in the next section.

### 3.5.2 Total Variation (TV) Regularization

The total variation regularization was introduced for image denoising and reconstruction in [23]. It is designed to prevent oscillations in the regularized solution while allowing for discontinuities. The anisotropic version of the TV regularization uses  $\text{Reg}(L) = \|D_1 L\|$ . We will use a superior isotropic TV model, which takes the form

$$\begin{aligned} \text{Reg}(L) &= \sum_{x,u} \sqrt{(D_{x_1} L)^2 + (D_{x_2} L)^2 + (D_{u_1} L)^2 + (D_{u_2} L)^2} \\ &= \sum_{x,u} \|\nabla L(x, u)\| \end{aligned} \tag{38}$$

This model is a generalization of TV regularization from images to lightfields. The isotropic TV is the  $L_1$  norm of discretized lightfield gradient magnitudes. The  $L_1$ -norm in the regularizer introduces non-linearities and the minimization of (38) is more complex than the  $L_2$  case. A simple procedure that solves this problem is called variable splitting, which decouples the  $L_2$  data fit in (36) and the  $L_1$  regularizer by introducing new variables and converting the solution to the problem into simple minimization steps. There are several techniques to minimize this new problem like the augmented Lagrangian method [24], split Bregman iterative method [25], or additive half-quadratic algorithm [26]. To use the additive half-quadratic algorithm we begin by defining the finite difference operators for lightfields in frequency

space:

$$\begin{aligned}
 \hat{D}_{x_1} L(\xi_1, \xi_2, u_1, u_2) &= \hat{L}(x_1 + 1, x_2, u_1, u_2) \\
 -\hat{L}(x_1, x_2, u_1, u_2) &= \left( e^{2\pi i \frac{\xi_1}{N_x}} - 1 \right) \hat{L}(\xi_1, \xi_2, u_1, u_2) \\
 \hat{D}_{x_2} L(\xi_1, \xi_2, u_1, u_2) &= \hat{L}(x_1, x_2 + 1, u_1, u_2) \\
 -\hat{L}(x_1, x_2, u_1, u_2) &= \left( e^{2\pi i \frac{\xi_2}{N_x}} - 1 \right) \hat{L}(\xi_1, \xi_2, u_1, u_2) \\
 \hat{D}_{u_1} \hat{L}(\xi_1, \xi_2, u_1, u_2) &= \hat{L}(\xi_1, \xi_2, u_1 + 1, u_2) \\
 -\hat{L}(\xi_1, \xi_2, u_1, u_2) & \\
 \hat{D}_{u_2} \hat{L}(\xi_1, \xi_2, u_1, u_2) &= \hat{L}(\xi_1, \xi_2, u_1, u_2 + 1) \\
 -\hat{L}(\xi_1, \xi_2, u_1, u_2) &
 \end{aligned} \tag{39}$$

That can be written as

$$\begin{aligned}
 \hat{D}_{x_1} L &= R_{x_1} \hat{L}, R_{x_1} = \left( e^{2\pi i \frac{\xi_1}{N_x}} - 1 \right) I, \\
 \hat{D}_{x_2} L &= R_{x_2} \hat{L}, R_{x_2} = \left( e^{2\pi i \frac{\xi_2}{N_x}} - 1 \right) I \\
 \hat{D}_{u_1} L &= R_{u_1} \hat{L}, R_{u_1}(u, v) = I(u, v - (1, 0)) - I(u, v); \\
 \hat{D}_{u_2} L &= R_{u_2} \hat{L}, R_{u_2}(u, v) = I(u, v - (0, 1)) - I(u, v).
 \end{aligned} \tag{40}$$

Then, to obtain the optimal solution of

$$\hat{L}(\xi, u)_{\text{opt}} = \arg \min_i \frac{\gamma}{2} \|\hat{s} - G\hat{l}\|^2 + \frac{\beta}{2} \sum_{x,u} \|\nabla L(x, u)\| \tag{41}$$

The additive half-quadratic algorithm proceeds by the minimization of the relaxed function

$$\begin{aligned}
 [\hat{l}, \hat{v}]_{\text{opt}} &= \arg \min_i \frac{\gamma}{2} \|\hat{s} - G\hat{l}\|^2 \\
 &+ \frac{\beta}{2} \sum_{i \in \{x_1, x_2, u_1, u_2\}} \|R_i \hat{l} - \hat{v}_i\|^2 + \|\hat{v}_i\|
 \end{aligned} \tag{42}$$

where  $\hat{v}_i(\xi, u)$  are  $N_u^2$  auxiliary variables for each  $i$ . The minimization proceeds by an alternate minimization in the variables  $\hat{L}(\xi, u)$  and  $\hat{v}_i(\xi, u)$ . Optimization in  $\hat{L}(\xi, u)$  leads to the normal equations:

$$\begin{aligned}
 \left( G^T G + \beta/\gamma \sum_{i \in \{x_1, x_2, u_1, u_2\}} R_i^T R_i \right) \hat{l} &= G^T \hat{s} \\
 + \beta/\gamma \sum_{i \in \{x_1, x_2, u_1, u_2\}} R_i^T \hat{v}_i &
 \end{aligned} \tag{43}$$

That is a system for each  $\xi$  with  $N_u^2$  equations and  $N_u^2$  unknowns assuming that the  $\hat{v}_i$  are fixed. Optimization of

(42) with respect to the  $\hat{v}_i$  for a fixed  $\hat{l}$  leads to the closed-form solution

$$v_{i,\text{opt}} = \frac{D_i l}{\|D_i l\|} \max \left( \|D_i l\| - \frac{1}{\beta}, 0 \right) \tag{44}$$

Note that this second optimization is performed in the original spatial space while the first optimization is performed in the frequency space, so each minimization step changes between both spaces. Since the method is iterative, its computational complexity depends on the number of iterations. The computational complexity of each iteration is  $O(N_x^2 N_u^6)$ .

The values for  $\gamma$  and  $\beta$  have to be determined beforehand. Different choices for the parameters result in a tradeoff between the smoothness of the solution and a good data fit. Since the optimization method is iterative, it is necessary to provide an initial starting point for the algorithm. Setting  $\hat{v}_i = 0$  and computing  $\hat{l}$  from (43) lead to a starting point for  $\hat{l}$  that is the solution to the  $L_2$  regularization problem

$$\hat{L}(\xi, u)_{\text{opt}} = \arg \min_i \frac{\gamma}{2} \|\hat{s} - G\hat{l}\|^2 + \frac{\beta}{2} \sum_{x,u} \|\nabla L(x, u)\|^2 \tag{45}$$

with  $\text{Reg}(L) = \|D_1 L\|_2^2$ . This is the initial point, we will use for the TV regularization in the experiments.

### 4 Experimental Results

In this section, we provide some computational results for the inversion of the Focal Stack. To test the applicability of the inversion procedure, we have processed several lightfields from several plenoptic cameras. The central image of the lightfields is shown in Fig. 2. Following the order in the figure, the statue lightfield has  $291 \times 291 \times 7 \times 7$  pixels and the faces lightfield has  $291 \times 291 \times 5 \times 5$  pixels and were taken from a research plenoptic camera. The toys, belfry, bicycle, and batteries lightfields were taken with a Lytro camera and have  $377 \times 377 \times 7 \times 7$  pixels. Finally, the elephant and watch lightfields were taken with a Raytrix camera and have  $377 \times 377 \times 9 \times 9$  pixels. Faces and statue lightfield had round microlenses that were cutoff to obtain square microlenses to process similarly all the lightfields. Lytro and Raytrix lightfield cameras have hexagonal microlenses in a hexagonal lattice that were interpolated to provide square microlenses in a square lattice [27]. Lytro images were affected by vignetting and noise. Noise is also noticeable in the statue and faces lightfield. Raytrix images are in black and white while the rest of plenoptic images are in color. Lightfield values are normalized in the range [0,1].





**Fig. 2** Central images from the statue, faces, toys, belfry, bicycle, batteries, elephant, and watch lightfields

**Table 2** Quality measures for lightfield reconstruction using the  $\Pi$ -constant assumption

Lightfield	MSE $\times 10^6$	PSNR	SSIM
Statue	184	37.363	0.988
Faces	185	37.330	0.993
Toys	224	36.503	0.963
Belfry	62	42.049	0.995
Bicycle	105	39.772	0.994
Batteries	223	36.521	0.987
Elephant	24	46.170	0.997
Watch	20	47.048	0.998

### 4.1 The $\Pi$ -Constant Assumption

In Sect. 3.1, we introduced the  $\Pi$ -constant lightfield assumption since inversion is not possible in the general case. In order to test the validity of this assumption, we have recovered the lightfields described in the previous section assuming that they are  $\Pi$ -constant by means of formula (27). To measure the quality of the reconstruction, we have computed the Mean Squared Error (MSE), the Peak Signal-to-Noise Ratio (PSNR), and the Structural Similarity Index Measure (SSIM) [28]. Those measures are averaged over all the images of the lightfield to obtain lightfield quality measures. In Table 2, we show quantitative results of the reconstruction error. Qualitative results are presented in Sect. 4.2.2.

Results show that recovery based on the  $\Pi$ -constant assumption is a very good approximation to the original light-

field with a standard deviation of the error around 1% of the true values and a structural similarity above 0.95 in all cases. However, as shown in Sect. 3.3, despite its accuracy this reconstruction is not practical due to the high number of planes that are needed.

### 4.2 Regularized Reconstruction

Regularized reconstruction tries to invert the discrete focal stack transform computing for each  $\xi$  the solution  $\hat{L}(\xi, \mathbf{u})$  of a linear system. We have performed several experiments for various types of regularizers:  $L_2$  regularization of 0th order and 1st order, and  $L_1$  isotropic TV regularization. As explained in Sect. 3.5.2 the initial point for TV regularization is computed from  $L_2$  1st-order regularization. In order to test the dependency of the reconstruction with respect to the number of images in the focal stack, we have set different values for its number. All planes are equispaced in the non-aliasing range to use the proposition in Sect. 3.4. The results for the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM) are shown on Tables 3 and 4.

The parameter values for regularization have been  $\lambda = 10^{-6}$  for the  $L_2$  0th order regularization and  $\gamma = 10^7$ ,  $\beta = 10^2$  for the 1st order regularization and were optimized for the recovery of all lightfields. The number of iterations for the TV regularization was set to 10. Although regularization parameters should be adjusted for each camera and each value of the number of planes, we adopted this approach due to the robustness of the error measures with respect to

**Table 3** PSNR of lightfield reconstruction for several regularizers

Lightfield	PSNR		Number of planes							
	Method		3	5	7	11	17	23	31	41
Statue	$L_2$ 0th		24.400	26.532	27.508	27.923	28.101	28.110	28.112	28.111
	$L_2$ 1st		24.590	26.632	27.488	27.887	28.208	28.267	28.293	28.307
	$L_1$ TV		24.730	26.698	27.469	27.876	28.214	28.289	28.315	28.338
Faces	$L_2$ 0th		28.671	30.767	31.330	31.788	31.857	31.860	31.860	31.856
	$L_2$ 1st		28.762	30.808	31.403	31.889	32.130	32.176	32.201	32.219
	$L_1$ TV		28.954	30.906	31.478	32.017	32.310	32.351	32.404	32.412
Toys	$L_2$ 0th		28.012	29.871	30.776	31.186	31.279	31.288	31.294	31.295
	$L_2$ 1st		28.180	29.918	30.748	31.154	31.309	31.334	31.353	31.359
	$L_1$ TV		28.540	30.152	30.911	31.323	31.538	31.594	31.622	31.644
Belfry	$L_2$ 0th		29.220	31.825	33.127	33.705	33.853	33.870	33.881	33.891
	$L_2$ 1st		29.508	32.110	33.319	33.929	34.115	34.114	34.177	34.168
	$L_1$ TV		30.065	32.429	33.283	33.825	34.087	34.140	34.127	34.213
Bicycle	$L_2$ 0th		28.347	30.938	32.322	32.915	33.005	33.033	33.039	33.044
	$L_2$ 1st		28.617	31.174	32.466	33.024	33.160	33.216	33.227	33.250
	$L_1$ TV		29.075	31.487	32.543	33.116	33.283	33.379	33.383	33.437
Batteries	$L_2$ 0th		24.610	26.288	27.017	27.494	27.568	27.578	27.580	27.580
	$L_2$ 1st		24.531	26.149	26.880	27.365	27.513	27.553	27.573	27.584
	$L_1$ TV		24.781	26.373	27.087	27.594	27.788	27.845	27.872	27.885
Elephant	$L_2$ 0th		27.037	35.492	38.000	39.448	39.707	39.817	39.782	39.843
	$L_2$ 1st		27.565	35.428	37.539	38.524	38.646	38.731	38.502	38.646
	$L_1$ TV		27.978	35.981	38.994	38.813	38.994	39.242	38.220	38.831
Watch	$L_2$ 0th		29.123	36.212	39.224	40.149	40.387	40.509	40.548	40.568
	$L_2$ 1st		29.610	36.497	38.929	39.432	39.488	39.638	39.562	39.555
	$L_1$ TV		30.115	36.664	38.620	38.680	38.993	39.318	39.065	38.928

the regularization parameters. The data-fit term in (36) was divided by the number of planes. This allows to use the same regularization parameters for all problems.

Numerical results in Tables 3 and 4 show that the lightfield recovery error stabilizes around 11 planes that is approximately 2 times the width of the microlens. For 11 planes, the PSNR of the research camera is in the range [27.9, 32.0] with SSIM in the range [0.84, 0.92]. For the Lytro camera, the PSNR is in the range [27.6, 33.9] with SSIM in the range [0.88, 0.94] and for the Raytrix camera the PSNR is around 40 with SSIM around 0.99. The difference in the recovery results can be attributed to the quality of the images since research and Lytro lightfield images are noisier than Raytrix lightfield images. Note that for a typical lightfield with  $255 \times 255$  microlenses and  $9 \times 9$  pixels behind each microlens this would lead to a recovery from a volume of data that is 22.2% of the original lightfield. The quantitative performance of the different regularizers is similar, although numerical results show that in general the PSNR and SSIM are ordered from lower to higher values as  $L_2$  0th order,  $L_2$  1st order, and TV regularization. A qualitative comparison of the recovery process is presented in Sect. 4.2.2.

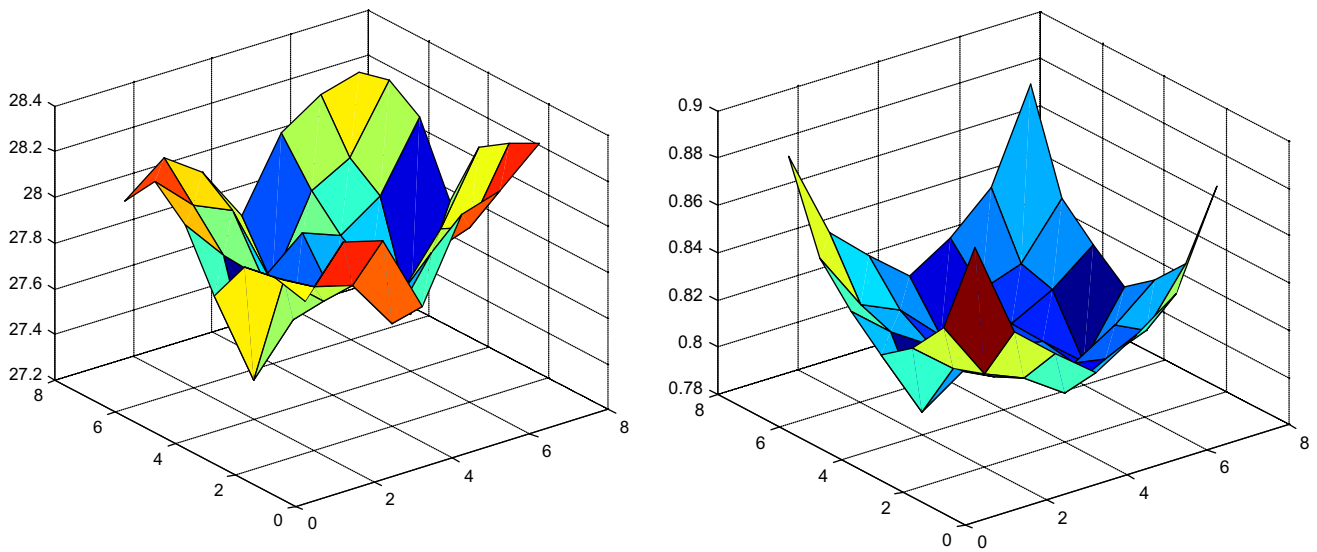
#### 4.2.1 Error Distribution in the Lightfield

Numerical results in Tables 3 and 4 show the mean error from all the images in the lightfield. A relevant question is the distribution of the error in all the images. To evaluate the error distribution for the images in the lightfield, we show in Fig. 3 the PSNR for each of the  $7 \times 7$  images in the statue lightfield that has the worst recovery results with  $L_2$  0th order regularization. The coefficient of variation of the PSNR and SSIM between all the lightfield images is small with values 0.0081 and 0.0274, respectively.

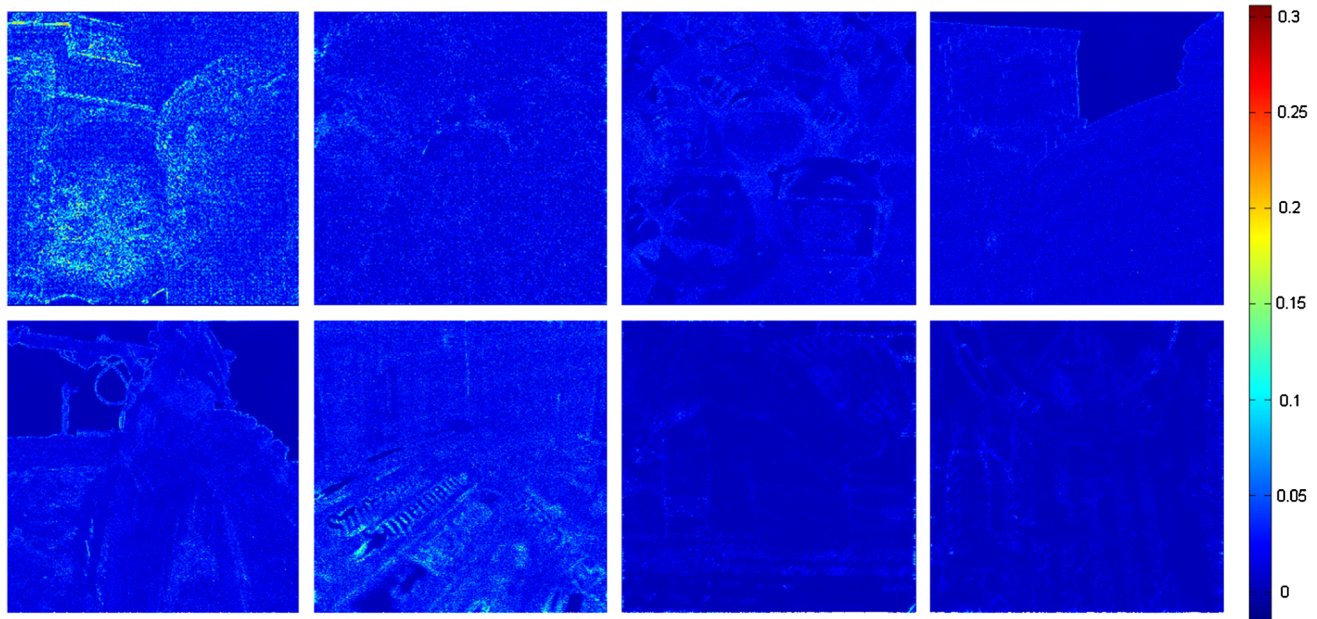
We have also studied the spatial distribution of the recovery error. It is measured as the absolute difference between the recovered lightfield and the true lightfield. The results can be seen in Fig. 4 for the central image of the lightfields in Fig. 2 and  $L_1$  TV regularization with 11 planes. Results show that errors are very low for most pixels of the image (lightfield values are in the range [0,1]). Errors are higher in the statue image (that can be attributed to noise) and the batteries image (there are specular reflections that violate the Lambertian assumption). Note that errors increase around depth discontinuities as shown in the belfry and bicycle images.

**Table 4** SSIM of lightfield reconstruction for several regularizers

SSIM		Number of planes							
Lightfield	Regularization	3	5	7	11	17	23	31	41
Statue	$L_2$ 0th	0.649	0.767	0.803	0.824	0.834	0.834	0.834	0.834
	$L_2$ 1st	0.658	0.770	0.802	0.823	0.840	0.843	0.844	0.845
	$L_1$ TV	0.662	0.770	0.800	0.821	0.839	0.842	0.844	0.844
Faces	$L_2$ 0th	0.833	0.885	0.900	0.915	0.918	0.918	0.918	0.918
	$L_2$ 1st	0.835	0.884	0.901	0.917	0.924	0.926	0.926	0.927
	$L_1$ TV	0.835	0.882	0.898	0.915	0.924	0.926	0.927	0.927
Toys	$L_2$ 0th	0.814	0.856	0.876	0.887	0.891	0.891	0.891	0.891
	$L_2$ 1st	0.818	0.858	0.878	0.890	0.894	0.894	0.895	0.895
	$L_1$ TV	0.830	0.868	0.887	0.899	0.906	0.907	0.907	0.907
Belfry	$L_2$ 0th	0.818	0.883	0.906	0.918	0.920	0.920	0.920	0.920
	$L_2$ 1st	0.825	0.887	0.909	0.922	0.926	0.926	0.927	0.927
	$L_1$ TV	0.834	0.887	0.907	0.920	0.925	0.926	0.927	0.927
Bicycle	$L_2$ 0th	0.860	0.912	0.931	0.941	0.942	0.942	0.942	0.942
	$L_2$ 1st	0.865	0.915	0.933	0.943	0.945	0.946	0.946	0.946
	$L_1$ TV	0.873	0.917	0.933	0.943	0.946	0.947	0.947	0.947
Batteries	$L_2$ 0th	0.763	0.824	0.854	0.872	0.874	0.874	0.874	0.874
	$L_2$ 1st	0.764	0.825	0.855	0.874	0.879	0.880	0.881	0.881
	$L_1$ TV	0.770	0.828	0.857	0.876	0.883	0.884	0.885	0.885
Elephant	$L_2$ 0th	0.891	0.985	0.994	0.996	0.996	0.996	0.996	0.996
	$L_2$ 1st	0.901	0.984	0.992	0.994	0.994	0.994	0.994	0.994
	$L_1$ TV	0.916	0.986	0.992	0.994	0.994	0.994	0.994	0.994
Watch	$L_2$ 0th	0.900	0.982	0.994	0.995	0.995	0.995	0.995	0.995
	$L_2$ 1st	0.907	0.983	0.992	0.994	0.994	0.994	0.994	0.994
	$L_1$ TV	0.919	0.984	0.992	0.994	0.994	0.994	0.994	0.994



**Fig. 3** Error distribution of the PSNR and SSIM in the lightfield images



**Fig. 4** Spatial distribution of the recovery absolute error. It is measured as the absolute difference between the recovered lightfield and the true lightfield. Results are shown for the central images in the statue, faces,

toys, belfry, bicycle, batteries, elephant, and watch lightfields. The regularization used has been  $L_1$  TV with a number of planes equal to 11



**Fig. 5** Focal stack with  $N_p = 11$  planes for the statue lightfield



**Fig. 6** Focal stack with  $N_p = 11$  planes for the elephant lightfield

#### 4.2.2 Qualitative Results

In order to test qualitatively the lightfield recovery, we have selected the statue lightfield with the worst recovery results and the elephant lightfield with the best recovery results. The focal stack for  $N_p = 11$  planes is shown in Fig. 5 for the statue lightfield and Fig. 6 for the elephant lightfield. To present the results, we have also selected the number of planes with worst recovery results  $N_p = 3$ , the number of planes where error stabilizes  $N_p = 11$ , and the number of planes with best recovery results  $N_p = 41$ . We have also selected for comparison the central image of the lightfield. In accordance with the preceding quantitative results in Figs. 7 and 8, we can see that there is no visual appreciable difference between  $N_p = 11$  and  $N_p = 41$ . Comparing the recovered central image with the original central lightfield image, we

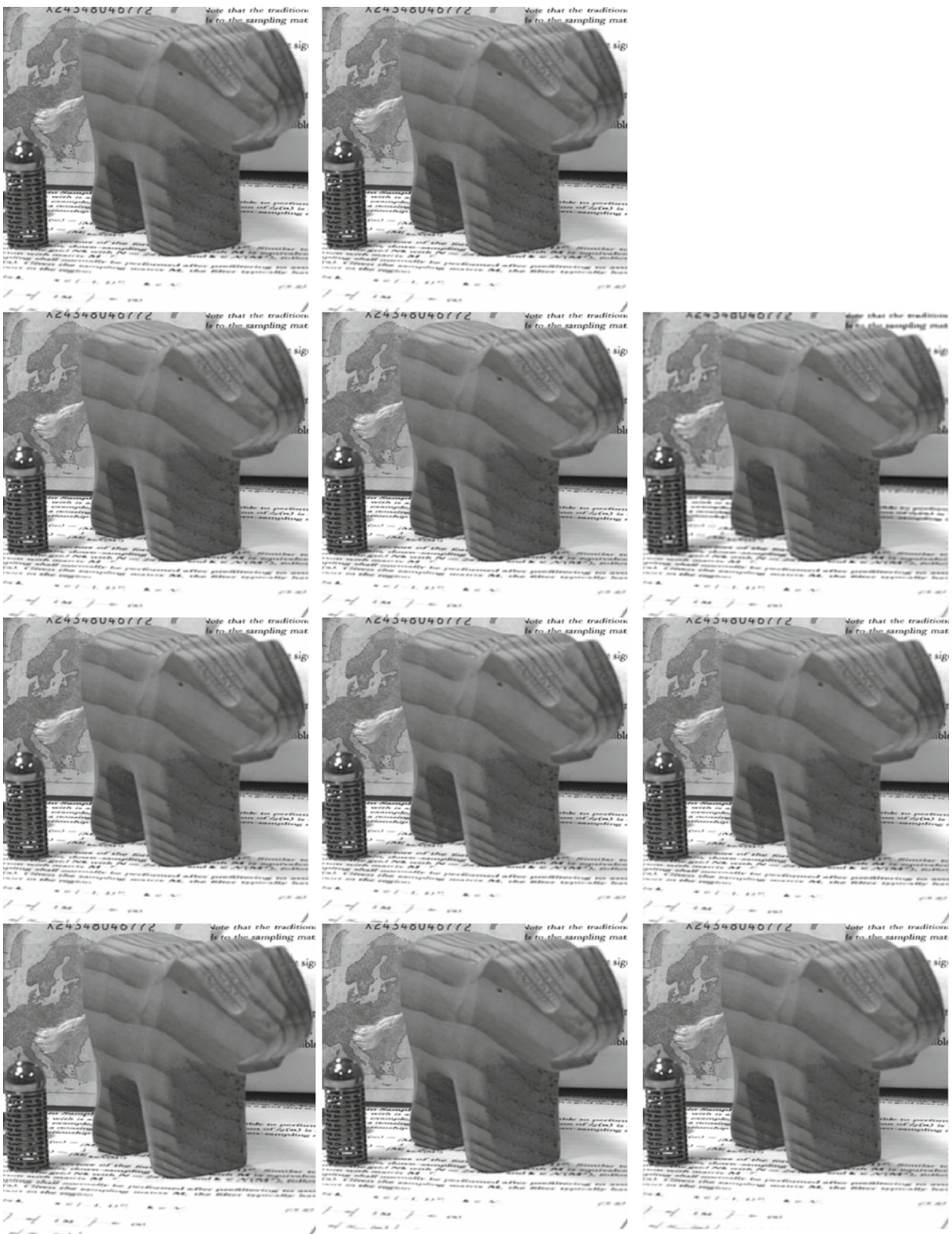
also can see that even small details are restored. Note that lightfield images are extended depth of field (DOF) images, so one feature of the lightfield recovery technique is that it is capable to compute extended DOF images from a set of defocused images.

#### 4.2.3 Lightfield Denoising

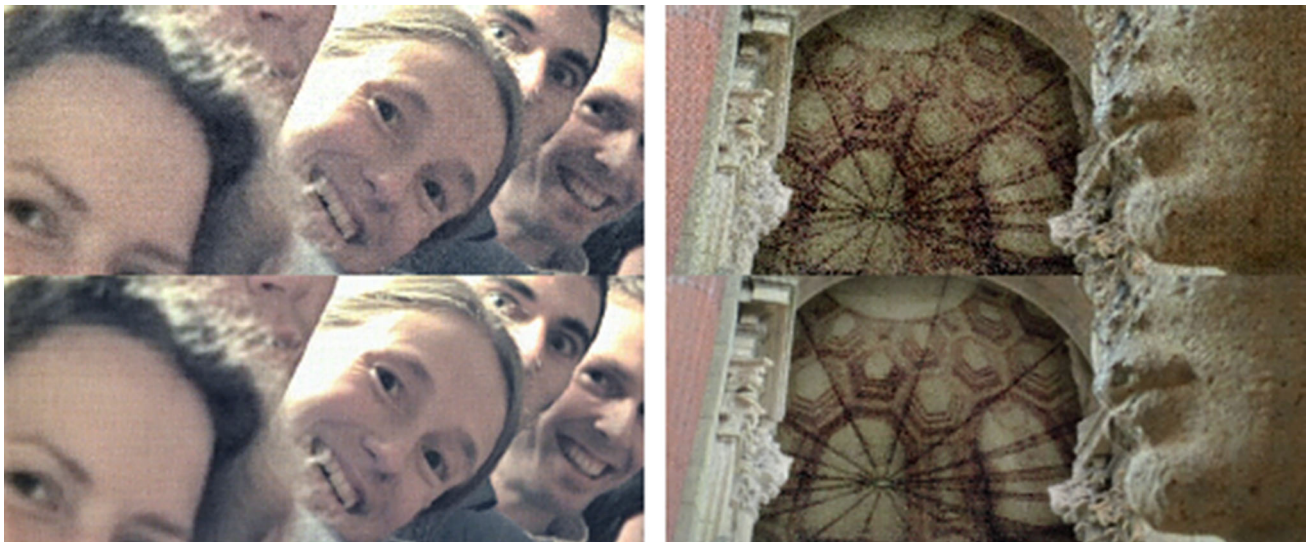
A consequence of the regularization process is the attenuation of oscillations in the lightfield. This causes that noise in lightfield images is reduced. Therefore, the inversion process could also be used to reduce noise in lightfield images by selecting appropriate parameter values to select the strength of the regularization. We have selected the noisiest statues and faces lightfields to show this denoising effect. As can be



**Fig. 7** First row Central image of the lightfield (left) and recovery with the  $\Pi$ -constant assumption (right). Second row, from left to right Recovery with  $N_p = 3$  for  $L_2$  0th order,  $L_2$  1st order, and TV regularization. Third and fourth row Recovery with  $N_p = 11$  and with  $N_p = 41$  with the same regularizers



**Fig. 8** First row Central image of the lightfield (left) and recovery with the  $\Pi$ -constant assumption (right). Second row, from left to right Recovery with  $N_p = 3$  for  $L_2$  0th order,  $L_2$  1st order, and TV regularization. Third and fourth row Recovery with  $N_p = 11$  and with  $N_p = 41$  with the same regularizers



**Fig. 9** *Top* Detail from the central image of the faces and statue lightfield. *Bottom* Detail from the central image of the faces and statue recovered lightfields with TV regularization

**Fig. 10** Details from the batteries lightfield where non-Lambertian specularities are present. Note the quality in the reconstruction despite the specular reflection on the surfaces



seen in Fig. 9, the recovery process effectively denoises the images in the lightfield.

#### 4.2.4 Lightfield Recovery Under Non-Lambertian Scenes

As explained in Sect. 3.1.1, lightfield inversion is exact for Lambertian lightfields. Then, it is interesting to study the behavior of the lightfield recovery process under non-Lambertian scenes. A common extension to the Lambertian assumption is the dichromatic model [29]. In this approach, the light reflected from the objects has two independent components, which correspond to Lambertian diffuse reflection and a specular non-Lambertian component. In Fig. 10, we study the qualitative performance of the reconstruction process under strong specular components. We have selected a detail from the batteries lightfield. It can be seen that the reconstruction process is able to recover both the diffuse and specular components of the original lightfield with high qual-

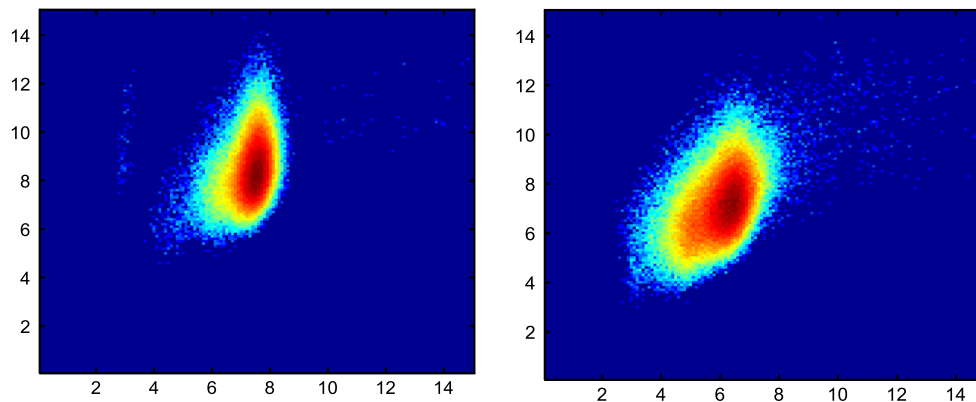
ity. The quantitative results of the difference between these two lightfield images are presented in Fig. 4.

Another source of non-Lambertian behavior appears at surface discontinuities even if the individual surfaces in the scene have a Lambertian reflectance. Then, it is interesting to study the recovery technique around depth discontinuities. We have selected the belfry lightfield due to its strong depth changes between the foreground and the background. As shown in Fig. 11, for a detail of two side-images from the lightfield, the reconstruction around depth discontinuities also has a good quality. Note that the parallax in the lightfield images is also recovered. The numerical error results of the reconstruction error around the discontinuity can also be found in Fig. 4.

Finally, in order to obtain a better understanding of the dependence of the reconstruction error under the Lambertian assumption, we have computed a photoconsistency measure [8] for each point in the scene. It is obtained from the lightfield



**Fig. 11** *Top* Detail from two belfry lightfield images. *Bottom* Detail recovered from the belfry lightfield. Note the differences in position between the foreground and background in the *left* and *right* images



**Fig. 12** Numerical study of the dependence of the recovery error (*vertical axis*) against a Lambertian photoconsistency error measure (*horizontal axis*) for the belfry (*left*) and batteries (*right*) lightfield

by measuring the variance of the intensity of all rays emanating from a surface point at a given depth. Since the depth of a given point on the surface is unknown, we compute the minimum of this measure for all possible point depths. Therefore, for a Lambertian surface, the photoconsistency error measure should be ideally zero for any point. Deviations from this assumption due to specularities or discontinuities should generate higher values. To study the dependence between both errors, we have computed a 2D histogram with the minus logarithm of both the photoconsistency and reconstruction errors in the horizontal and vertical axis, respectively. Therefore, lower errors for both measures are in the top and right side of the histogram. These values are accumulated for all surface points and rendered in Fig. 12 for the belfry and batteries lightfields. Numerical results show that lower values of photoconsistency error lead to lower values of the reconstruction error.

## 5 Conclusions

We have presented a new technique for the recovery of a lightfield from its Focal Stack. This recovery is based on the inversion of the Discrete Focal Stack transform. The inversion process has been studied in the continuous case and it is shown that this inversion is not possible in general. However, there is a subset of lightfields where this inversion is possible. This subset, that we call  $\Pi$ -constant lightfields, has been characterized and it is shown that it includes the important case of Lambertian lightfields. A simple and direct inversion procedure has been developed for these lightfields. The preceding results have been also studied in the discrete case, and a similar direct procedure has been developed. A drawback of the direct procedure is that it needs a large number of images in the focal stack. Therefore, a theoretical result has been provided that ensures that inversion for  $\Pi$ -constant



lightfields is possible solving a linear system of equations for a focal stack with a volume of data equal to the volume of data of the lightfield. We have studied the inversion of this linear system, and we have shown that is bad-conditioned, so regularization procedures have been proposed and detailed. The inversion method has been tested with several images from plenoptic cameras. Experimental results show that  $\mathcal{H}$ -constant lightfields are a very good approximation to real world lightfields. To study practical inversion procedures, we have studied the inversion of general lightfields with regularization. The experimental results show that general lightfields can be recovered from a focal stack with a limited number of images with high precision, so a focal stack can be used as a compressed representation of a lightfield. The inversion method is accurate enough to recover parallax and depth discontinuities. This method can also be employed to obtain an extended depth of field image from a focal stack. Due to the regularization effect, the inversion procedure can also be used to denoise the lightfield. Future extensions to this paper will include the study of the optimal placement of planes in the focal stack to decrease its number in the recovery process and porting the technique to Graphical Processing Units and Field Programmable Gate Arrays.

**Acknowledgments** The authors would like to thank R. Ng and Heidelberg University for lightfields that were used in the experimental results. This work has been partially supported by “Ayudas al Fomento de Nuevos Proyectos de Investigación” (Project 2013/0001339) of the University of La Laguna.

## References

- Levoy, M.: Light fields and computational imaging. *Computer* **39**(8), 46–55 (2006)
- Adelson, E.H., Bergen, J.R.: *The Plenoptic Function and the Elements of Early Vision*. MIT Press, Cambridge (1991)
- Levoy, M., Hanrahan, P.: Light field rendering. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, pp. 31–42 (1996)
- Gortler, S.J., Grzeszczuk, R., Szeliski, R., Cohen, M.F.: The Lumigraph. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, pp. 43–54 (1996)
- Ng, R.: Fourier slice photography. In: *ACM SIGGRAPH: Papers*, New York, 2005, pp. 735–744 (2005)
- Nava, F.P., Marichal-Hernández, J.G., Rodríguez-Ramos, J.M.: The discrete focal stack transform. In: *Proceedings of European Signal Processing Conference*, pp. 1–5 (2008)
- Lumsdaine, A., Georgiev, T.: The focused plenoptic camera. In: *IEEE International Conference on Computational Photography (ICCP)*, 2009, pp. 1–8 (2009)
- Lüke, J.P., Pérez Nava, F., Marichal-Hernández, J.G., Rodríguez-Ramos, J.M., Rosa, F.: Near real-time estimation of super-resolved depth and all-in-focus images from a plenoptic camera using graphics processing units. *Int. J. Digit. Multimed. Broadcast.* **2010**, e942037 (2009)
- Uliyar, M., Putraya, G., Basavaraja, S.V.: Fast EPI based depth for plenoptic cameras. In: *20th IEEE International Conference on Image Processing*, pp. 1–4 (2013)
- Tosic, I., Berkner, K.: Light field scale-depth space transform for dense depth estimation. In: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 441–448 (2014)
- Pérez, F., Pérez, A., Rodríguez, M., Magdaleno, E.: Super-resolved Fourier-slice refocusing in plenoptic cameras. *J. Math. Imaging Vis.* **52**(2), 200–217 (2015)
- Georgiev, T., Chunev, G., Lumsdaine, A.: Superresolution with the focused plenoptic camera. Presented at the *Proceedings of SPIE 7873, Computational Imaging IX*, vol. 7873, p. 78730X–78730X-13 (2011)
- Perez Nava, F.: Super-resolution in plenoptic cameras by the integration of depth from focus and stereo. In: *Proceedings of 19th International Conference on Computer Communications and Networks*, pp. 1–6 (2010)
- Pérez Nava, F., Pérez Nava, A., Rodríguez Valido, M., Magdaleno Castelló, E.: Plenoptic cameras. In: *Cristóbal, G., Perrinet, L., Keil, M.S. (eds.) Biologically Inspired Computer Vision*, pp. 175–200. Wiley-VCH Verlag GmbH & Co, KGaA, Weinheim (2015)
- “Lytro.” <https://www.lytro.com> (2014). Accessed 04 Jul 2014
- “Raytrix.” <http://www.raytrix.de/> (2014). Accessed 04 Jul 2014
- Pérez, F., Pérez, A., Rodríguez, M., Magdaleno, E.: A fast and memory-efficient Discrete Focal Stack Transform for plenoptic sensors. *Digit. Signal Process.* **38**, 95–105 (2015)
- Stroebel, L.D. (ed.): *Basic Photographic Materials and Processes*, 2nd edn. Focal Press, Boston (2000)
- Ferreira, P.J.S.G.: Superresolution, the recovery of missing samples, and Vandermonde matrices on the unit circle. In: *Proceedings of Workshop Sampling Theory Applications (SAMPTA'99)* (1999)
- Hansen, P.C.: Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion. SIAM, Philadelphia (1998)
- Bertero, M., Boccacci, P.: *Introduction to inverse problems in imaging*. Bristol, UK?. Institute of Physics Publication, Philadelphia (1998)
- Reichel, L., Rodriguez, G.: Old and new parameter choice rules for discrete ill-posed problems. *Numer. Algorithms* **63**(1), 65–87 (2013)
- Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. Nonlinear Phenom.* **60**(1–4), 259–268 (1992)
- Afonso, M.V., Bioucas-Dias, J.M., Figueiredo, M.A.T.: Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.* **19**(9), 2345–2356 (2010)
- Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. *SIAM J. Imaging Sci.* **2**(2), 323–343 (2009)
- Yang, J., Yin, W., Zhang, Y., Wang, Y.: A fast algorithm for edge-preserving variational multichannel image restoration. *SIAM J. Imaging Sci.* **2**(2), 569–592 (2009)
- Dansereau, D.G., Pizarro, O., Williams, S.B.: Decoding, calibration and rectification for lenselet-based plenoptic cameras. Presented at the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 1027–1034 (2013)
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
- Shafer, S.A.: Using color to separate reflection components. *Color Res. Appl.* **10**(4), 210–218 (1985)



**F. Pérez** received his Ph.D. in Computer Science in 2001 from the University of La Laguna in Spain. In 1989–1992, he participated in several national and European projects in the private sector. He became associate professor in 1993 and professor in 2004 at the University of La Laguna, where he headed the Department of Statistics, Operational Research and Computer Science. He has written a number of different journal publications in the Computer Vision field and

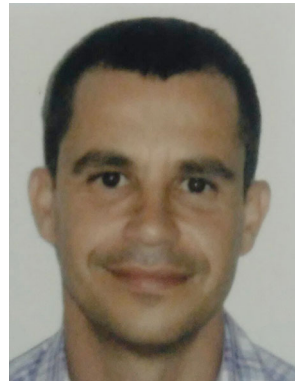
is the coauthor of a Computer Vision book and a patent in plenoptic cameras. His research interests include lightfield sensing devices and the design of intelligent sensors.



**A. Pérez** is an Associate Professor at the Department of Statistics, Operational Research and Computer Science at the University of La Laguna, Spain. His research includes the design and development of intelligent sensors and interfaces and their practical implementation. He also works for the private sector and participates in several projects in areas such as banking, tourism, and transport.



**M. Rodríguez** received his Ph.D. in Physics from the University of La Laguna where he was vice-dean at the School of Physics. His main research interests include smart sensor networks and FPGA electronic design applied to image processing and high performance computing. He has produced a number of different publications in these fields. He has also participated in several Program committees in national and international conferences.



**E. Magdaleno** received in 2009 his Ph.D. in Electrical Engineering from the University of La Laguna where he is an Associate Professor from 2002. His main research interests include smart sensor design, 3D vision, and efficient algorithm implementations in FPGA. He is the author of several publications in these fields.