



Berlin  
Mathematical  
School



Technische Universität Berlin

Master's Thesis

---

# Fast Sparse Light Field Reconstruction with Shearlet-based Inpainting

---

*Author:*

Héctor Andrade Loarca

*Supervisor:* Prof. Dr. Gitta Kutyniok

*Second reader:* Dr. Philipp Petersen

Fakultät II

Institut für Mathematik

AG Angewandte Funktionalanalysis

11. September 2017



## **Erklärung**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Die selbständige und eigenständige Anfertigung versichert an Eides statt: Berlin, den

4. August 2017

Héctor Andrade Loarca

# **Zusammenfassung in deutscher Sprache**

## **Schnelle Rekonstruktion für Dünne Lichtfelder mit Shearlet-basierten Einfärbungen**

Diese These ist angewendete

*A Natasha y los años que nos quedan juntos  
A mi madre Julieta y Padre Héctor  
sin los cuales nada de esto hubiera pasado  
A Patricia, Sara y Cristina,  
por enseñarme cada día lo que es una familia*



You don't understand anything  
until you learn it more than one way.  
*Marvin Minsky*





# Acknowledgements

To my mom.

To all of you, thank you very much.

Berlin, September 2017



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Light Field Photography</b>	<b>3</b>
2.1	Light Field Photography in the History . . . . .	4
2.2	Light Field Acquisition Settings . . . . .	6
2.3	Typical applications for the Light Field Theory . . . . .	8
2.4	Geometric proxy: Stereo Vision and multiview Epipolar Geometry . . .	8
2.5	Physical and Computational Setup for Sparse acquisition of Epipolar-plane	8
<b>3</b>	<b>Shearlets</b>	<b>9</b>
3.1	Shearlets as Frames . . . . .	9
3.2	Generalization of Shearlets to Alpha Particles . . . . .	9
3.3	Linear Shearlets and its relation with ridgelets . . . . .	9
3.4	Image inpainting using Shearlets . . . . .	9
3.5	Epipolar-plane representation with linear Shearlets . . . . .	9
<b>4</b>	<b>Inpainting Sparse Sampled Epipolar-plane</b>	<b>11</b>
4.1	Using linear Shearlets to inpaint sparse sampled Epipolar-plane . . . . .	11
4.2	Iterative thresholding with constant velocity . . . . .	11
4.3	Iterative thresholding with variable velocity . . . . .	11
<b>5</b>	<b>Conclusion and outlook</b>	<b>13</b>



# Chapter 1

## Introduction

Introduction template.



## Chapter 2

# Light Field Photography

The propagation of the light rays in the 3D space can be completely described by a 7D continuous function  $R(\theta, \phi, \lambda, \tau, V_x, V_y, V_z)$ , where  $(V_x, V_y, V_z)$  is a location in the 3D space,  $(\theta, \phi)$  are propagation angles,  $\lambda$  is the wavelength and  $\tau$  the time; this function is known as the plenoptic function and describes the amount of light flowing in every direction through every point in space at any time, the magnitude of  $R$  is known as the radiance. In an 1846 lecture entitled "Thoughts on Ray Vibrations" Michael Faraday proposed for the first time that light could be interpreted as a field, inspired by his work on magnetic fields; but the idea of a plenoptic function representing the spectral radiance distribution of rays was first proposed by Adelson and Bergen [4].

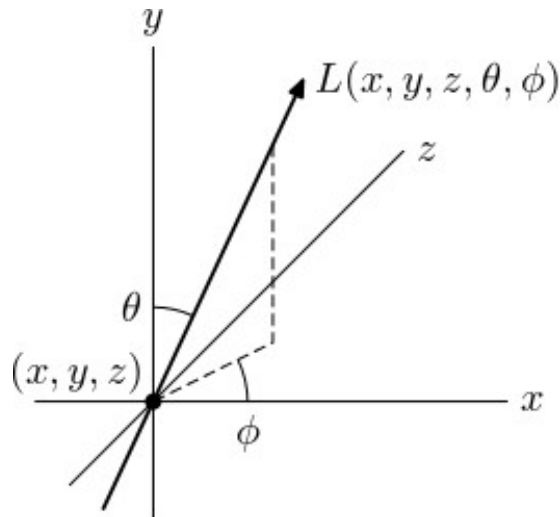


FIGURE 2.1: Spatio-angular parametrization of the plenoptic function for fixed  $\tau$  and  $\lambda$ . Figure taken from Wikipedia ([https://en.wikipedia.org/wiki/Light\\_field](https://en.wikipedia.org/wiki/Light_field))

In a more practical approach the plenoptic function can be simplified to a 4D version, called 4D Light Field or simply Light Field (abbreviated from now on as LF), denoted as the function  $L$ . The LF quantifies the intensity of static and monochromatic light rays propagating in half space, though this seems like an important reduction of information, this constraint does not substantially limit us in the accurate 3D description of the scene from where the light rays come from.

There exists three typical forms of this 4D approximation:

1. The LF rays positions are indexed by their Cartesian coordinates on two parallel planes, also called the two-plane parametrization  $L(u, v, s, t)$ .

2. The LF rays positions are indexed by their Cartesian coordinates on a plane and the directional angles leaving each point,  $L(u, v, \phi, \theta)$ .
3. Pairs of points on the surface of a sphere  $L(\phi_1, \theta_1, \phi_2, \theta_2)$ .

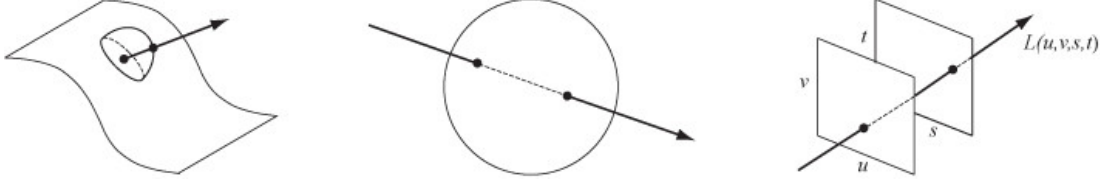


FIGURE 2.2: Three different representations of 4D LF. Left:  $L(u, v, \phi, \theta)$ . Center:  $L(\phi_1, \theta_1, \phi_2, \theta_2)$ . Right:  $L(u, v, s, t)$ . Figure taken from Wikipedia ([https://en.wikipedia.org/wiki/Light\\_field](https://en.wikipedia.org/wiki/Light_field))

In this work we will centered our attention in the two plane parametrization  $L(u, v, s, t)$ , if you are interested in the other descriptions we recommend to see [5]. In order to understand deeply this way of LF description, lets consider a camera with image plane coordinates  $(u, v)$  and the focal distance  $f$  moving along the  $(s, t)$  plane.

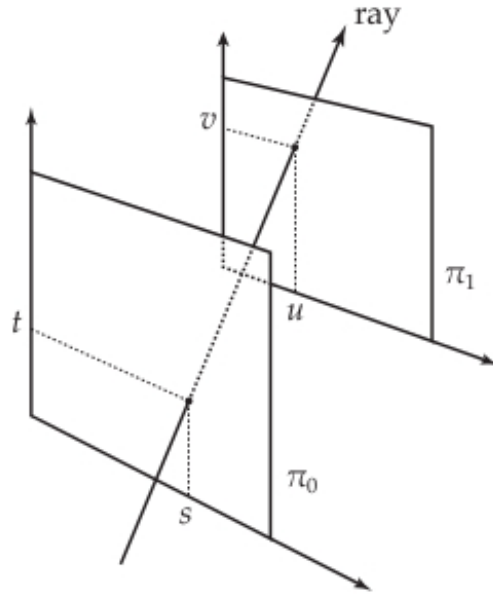


FIGURE 2.3: Graphic representation of the two plane parametrization of a single ray on the LF which is parametrized by the intersection  $(s, t)$  and  $(u, v)$  with planes  $\pi_0$  and  $\pi_1$ , respectively. Figure taken from [6]  
p.21

For simplicity one can constrain the vertical camera motion by fixing  $s = s_0$  and moving the camera along the  $t$  - axes in a straight light motion, in the section 2.4 we will see that this constraint leads to an elegant geometric 3D representation of the scene called Epipolar Geometry, this multiview aquisition is refered as parallax only (HPO). Under this constraint, images captured by successive camera positions  $t_1, t_2, \dots$  can be stacked together, and one can also interpret each camera position as a time step.



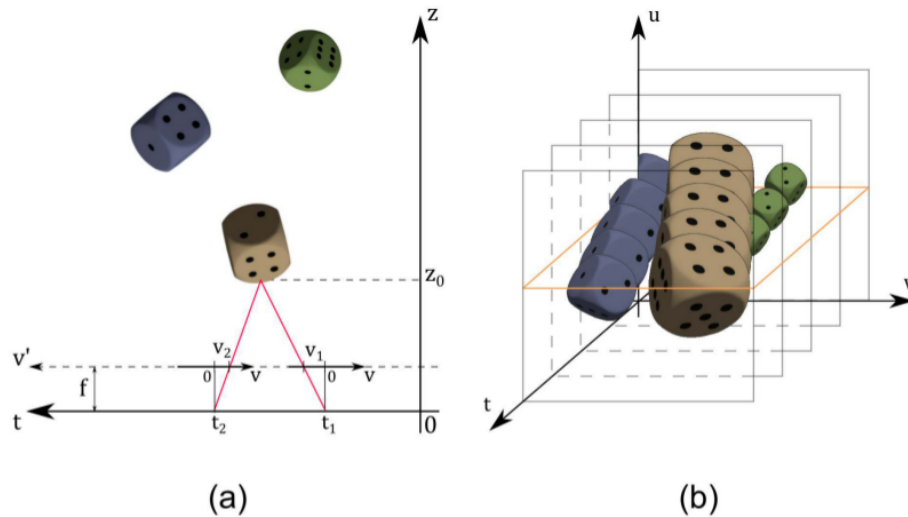


FIGURE 2.4: Stacked captured images represented in (b) from the scene setup (a). Figure taken from [3] p. 2

## 2.1 Light Field Photography in the History

For different reasons of interest for science and art capturing light fields has been an active research area for more than 110 years (the reason will be explained in detail on the section 2.3). In 1903 Herbert E. Ives [7] was the first to realize that the light field inside a camera can be recorded by placing a pinhole or lenslet arrays in front of a film sensor (what is known as pinhole camera). On the other hand, in 1908 the french physicist and Nobel laureate Gabriel Lippmann published two articles about something that he called *photographie intégrale* (translated as integral photography) [8] in which he describes an imaging apparatus with an arrange of small lenses on a 2D grid that are able to capture multiple images of a scene with viewpoint variations; the captured scene is reproduced in 3D as the viewer sees the parallax while the viewpoint changes. Is quite surprising that almost 110 year ago he could have the idea that modern state of the art LF acquisition systems use nowadays.

Even some experiments to acquire the Light Field of a static scene were already proposed since the beginning of the XX century, the first contribution on the mathematical formalization of the Light Field Theory were proposed in 1991, when Adelson and Bergen [9] found a way to systematically categorize the visual elements in *early vision*, which in combination, form visual information in the world. Here by *early vision* we mean the processes that are involved in the first steps of the visual cortex, namely, basic segmentation, shape detection, motion analysis between others (for further information about *early vision* [10] is highly recommended); for this purpose Adelson and Bergen defined the *plenoptic function* which was already discussed at the beginning of this Chapter.

The history of Light Field Theory can be separated in the three main steps in the study of the Light Field: The acquisition, the processing (which include in the most of the cases a geometry proxy) and the rendering, which are closely related, vary in computational complexity and for which there exist plenty of different approaches; in this thesis we will center our study on the first two steps.

It is also worth to mention that in the last decade two companies had manufactured Cameras that are able to capture the 4D Light Field, also known as plenoptic cameras; the first was Raytrix founded by the german computer scientists Vhrastian Perwass and Lennart Wietzke that released their camera in 2010 mostly focused on industrial application on 3D reconstruction (one can see their paper [12] for a good reference) rather than general consumers. Later in 2012 the american company Lytro came out with a plenoptic camera that was the first consumer light field camera for the general public, that has as a principal feature the possibility to do refocusing in the pictures taken by the camera (as a reference for this camera we recommend to read the Stanford Technical Report written by the CEO of the company Ren Ng [13]). Both companies produce cameras that capture the light field using an array of lenses, this and other LF acquisition setting will be discuss in the next section.



FIGURE 2.5: Industrial plenoptic camera Raytrix R11, produced by Raytrix. Figure taken from <https://petapixel.com/assets/uploads/2010/09/raytrix.jpg>



FIGURE 2.6: Consumer plenoptica camera Lytro Illum, produced by Lytro. Figure taken from [https://www.ephotozine.com/articles/lytro-illum-review-26434/images/highres-Lytro-Illum-6\\_1414410926.jpg](https://www.ephotozine.com/articles/lytro-illum-review-26434/images/highres-Lytro-Illum-6_1414410926.jpg)

## 2.2 Light Field Acquisition Settings

The first creativity step in the experimental study of Light Field is the form of acquisition; using only our physical intuition is not trivial to come up with an idea of a system that captures faithfully the Light Field coming from a static scene that will be able to be processed by some straight-forward algorithm. From the beginning of the last century until today, scientists, engineers and hobbyists have proposed different approaches for the Light Field Acquisition Settings.

As we have seen in the last section, the first attempts of settings were the pinhole camera proposed by Ives and the lenslet array proposed by Lippmann. The next variation of setting was proposed more than eighty years later by Adelson and Wang [14] that in 1992 using the theory of plenoptic function (developed by Adelson itself) presented a design of a plenoptic camera where the light rays that pass through the main lens are recorded separately using a lenticular array placed on the sensor plane, they used the light field recorded with this camera to obtain the scene depth by analyzing the directional variation of the radiance captured in the image; this is basically the Lippmann design but applied to digital cameras. Ng et al. [13] from Lytro used the same design of Adelson and Wang to produce the Lytro cameras.

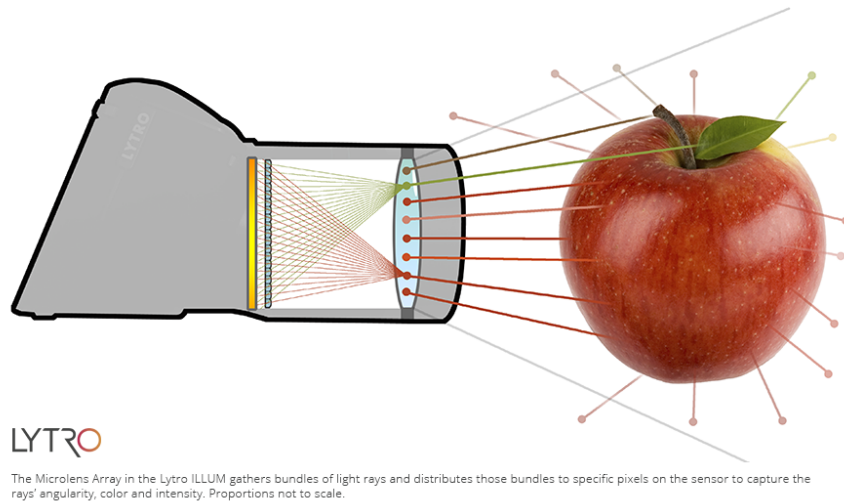


FIGURE 2.7: Diagram of Adelson and Wang design in Lytro cameras. Figure taken from [https://s3.amazonaws.com/lytro-corp-assets/blog/Lytro\\_ILLUM.png](https://s3.amazonaws.com/lytro-corp-assets/blog/Lytro_ILLUM.png)

In 2006 Joshi et al [15] used a one-dimensional camera array and a motorized stage for their real-time matting system. This technique of multicameras/multiviews acquisition is also quite common with camera arrays varying in position and size. The approach followed in this thesis take this technique as acquisition setting, the actual setup used will be explained in more detail in the section 2.5. The downside of this acquisition device is that in counterpart of the Lytro camera (hand-held) it can be built without having a priori a designed custom optics, but they are bulky and often not portable (mechanical tracks are generally quite big and heavy).

A less bulky approach are the ones with Light-modulating codes in mask-based systems, that use coded masks in front of lenses for coded acquisition of the scene. Veeraraghavan et al. [16] where the first implementing a coded aperture technique to

computationally demultiplex the light rays collected through the camera's main lens. This attempt is less bulky than the multicaseras and more light efficient than the pinhole arrays but it sacrifices image resolution, since the number of sensor pixels is the upper limit of the number of light rays captured (problem than camera arrays and lenslets does not have). To overcome this problem, Wetzstein et al. [17], analyzed multiplexing light fields onto a 2D image sensor and developed a theory for multiplexing and a computational reconstruction algorithm.

A significant challenge of acquisition is that the captured set of images is very data-intensive and also redundant, mostly when one tries to recover high resolution light field from images with resolution above  $(2000px)^2$ . In order to tackle this issue, since the early papers on Light Field, the discussion about compact or sparse representation and compression schemes have played an important role in the area. For instance, Levoy and Hanrahan [18] proposed in 1995 several representations for 4D light fields and apply a lossy vector quantization followed by entropy coding; whereas Gortler et al. [19] in the same year applied standard image compression like JPEG to some of the views and pointed out the importance of depth information for sparser representation.

Later on Wetzstein with the Camera Culture Group of the MIT Media Lab developed a compressive light field camera architecture that allows for higher-resolution light fields to be recovered than previously possible from a single image, using three main components: light field atoms as a sparse representation of natural light fields (that involves dictionary learning which elements are the light field atoms), and optical design that allows for capturing optimized 2D light field projection also based in the coded masks technique, and robust sparse reconstruction methods to recover a 4D light field from a single coded 2D projection. In our opinion even this approach allows us to get a very high resolution of light fields, is a trade off by its requirements of high performance computation and its limitations coming from the biased learned dictionaries from a limited set of scenes.

Due to the lack of equipment,

## **2.3 Typical applications for the Light Field Theory**

### **2.4 Geometric proxy: Stereo Vision and multiview Epipolar Geometry**

### **2.5 Physical and Computational Setup for Sparse acquisition of Epipolar-plane**

## Chapter 3

# Shearlets

Sample of Chapter 3

### 3.1 Shearlets as Frames

### 3.2 Generalization of Shearlets to Alpha Particles

### 3.3 Linear Shearlets and its relation with ridgelets

### 3.4 Image inpainting using Shearlets

### 3.5 Epipolar-plane representation with linear Shearlets



## Chapter 4

# Inpainting Sparse Sampled Epipolar-plane

Sample of Chapter 4

- 4.1 Using linear Shearlets to inpaint sparse sampled Epipolar-plane
- 4.2 Iterative thresholding with constant velocity
- 4.3 Iterative thresholding with variable velocity





## Chapter 5

# Conclusion and outlook

Template of Conclusion, [1]



# Bibliography

- [1] R.C. Bolles, H.H. Baker, D. H. Marimont, *Epipolar-plane image analysis and approach to determining structure from motion*, International Journal of Computer Vision, 1:7-55, 1987.
- [2] G. Kutyniok, M. Genzel, *Asymptotic analysis of inpainting via universal shearlet Systems*, SIAM Journal on Imaging Sciences, 7(4), 2301-2339, 2014
- [3] S. Vagharshakyan, R. Bregovic, A. Gotchev, *Light field reconstruction using shearlet transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence, P(99), 2017
- [4] E.H. Adelson and J.R. Bergen, *The plenoptic function and the elements of early vision*, Vision and Modeling Group, MIT Media Laboratory, MIT, 1991
- [5] C.-K. Liang, Y.-C. Shih, H.Chen, *Light field analysis for modeling image formation*, IEEE Trans. Image Processing, 20(2), 446-460, 2011
- [6] C. Kim, *3D Reconstruction and Rendering from High Resolution Light Fields*, Diss. ETH No. 22933, 2015
- [7] H. Ives, *Parallax Stereogram and Process of Making Same* US patent 725, 567 1903
- [8] G. Lippmann, *La Photographie Intégrale*, Academie des Sciences 146, 446-451, 1908
- [9] E.H. Adelson, J. R. Bergen, *The plenoptic function and the elements of early vision*, Computational Models of Visual Processing, pages 3-20, 1991
- [10] C. Tomasi, *Early Vision*, Encyclopedia of Cognitive Science, Level 2, 2006
- [11] K. Marwah, G. Wetzstein, Y. Bando, R. Raskar, *Compressive Light Field Photography using Overcomplete Dictionaries and Optimized Projections*, ACM Transactions on Graphics (SIGGRAPH), 32(4), 2013
- [12] C. Perwass, L. Wietzke, *Single lens 3D-camera with extended depth-of-field*, Human Vision and Electronic Imaging 2012, 829108, 2012
- [13] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan, *Light field photography with a hand-held plenoptic camera*, Technical Report CSTR 2005-2, Stanford University, 2005
- [14] E.H. Adelson, J. Y. A. Wang, *Single lens stereo with a plenoptic camera*, IEEE International Conference on Computer Vision, 2007
- [15] N. Joshi, W. Matusik, S. Avidan, *Natural video matting using camera arrays*, ACM Transactions on Graphics, 25(3), 779-786, 2006
- [16] A. Veeraraghavan, R. Raskar, A. K. Agrawal, A. Mohan and J. Trumbull, *Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing*, ACM Transactions on Graphics, 26(3), 1-69, 2007

- [17] G. Wetzstein, I. Ihrke, W. Heidrich, *On plenoptic multiplexing and reconstruction*, International Journal of Computer Vision, 101(2), 384-400, 2013
- [18] M. Levoy, P. Hanrahan, *Light field rendering*, Proceedings of ACM SIGGRAPH, 31-42, 1996
- [19] S. J. Gortler, R. Grzeszczuk, R. Szeliski, M. F. Cohen, *The Lumigraph*, Proceedings of ACM SIGGRAPH, 43-54, 1996