

Reconstruction of 3D Surfaces with Complex Material Composure Using a Light Field Camera

by **Helia Farhood**

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

under the supervision of A/Professor Stuart Perry

University of Technology Sydney

Faculty of Engineering and Information Technology

School of Electrical and Data Engineering

September 2021

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Helia Farhood declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Electrical and Data Engineering, Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

SIGNATURE: Production Note:
Helia Farhood Signature removed prior to publication.

DATE: 12th Sep, 2021
PLACE: Sydney, Australia

DEDICATION

*This PhD thesis is dedicated to my mother Pouran Hamidi
for giving me invaluable educational opportunities
and my husband Mohammad Najafi
for his support, constant encouragement and care.*

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to my primary supervisor Professor Stuart Perry, for allowing me to work with him and his kind support throughout my PhD at UTS. His valuable advice, continuous encouragement and constant support were critical throughout the development of my research work. He was a wonderful mentor and I never felt stressed before or during any meeting with him. I would like to thank Dr Eva Cheng, my co-supervisor for her guidance and her support. Her deep insights helped me at various stages of my research. I am extremely grateful to my external supervisor Dr Juno Kim for his guidance and useful advice. His important suggestions and advice are priceless. He spent endless time reviewing and proofreading my papers.

I would also like to acknowledge Daniel Heckenberg for his valuable inputs to this work. This research was supported by a scholarship from Animal Logic Pty. Ltd. Dr Juno Kim was supported by an Australian Research Council (ARC) Future Fellowship (FT140100535) and the Sensory Processes Innovation Network (SPINet).

I express my deepest gratitude to my mother for all her support and a heartfelt thanks to my husband Mohammad Najafi and my little son Mahan Najafi for their patience and encouragement.

I also thank all the departmental staff for helping me during these years of my study both academically and administratively.

LIST OF PAPERS (PUBLICATIONS)

Helia Farhood, Stuart Perry, Eva Cheng, Juno Kim, "Enhanced 3D Point Cloud from a Light Field Image," *Remote Sensing*, vol. 12, p. 1125, 2020. Published in April 2020. Journal Ranking: Q1. (Related to Chapter 6)

Helia Farhood, Stuart Perry, Eva Cheng, Juno Kim, "3D point cloud reconstruction from a single 4D light field image," in *Optics, Photonics and Digital Technologies for Imaging Applications VI*, 2020, p. 1135313. Published in April 2020. (Related to Chapter 5)

Helia Farhood, Stuart Perry, Eva Cheng, Juno Kim, "Depth estimation: combination of sub-aperture matching and defocusing for a 4D light field", Elsevier, *Pattern Recognition*, February 2020 (under review). Journal Ranking: Q1. (Related to Chapter 4)

Helia Farhood, Stuart Perry, Eva Cheng, Juno Kim, "Reflection Recovery in Light Field images with Complex Material Appearance", has been submitted to Elsevier *Journal of Signal Processing Image Communication*, September 2020. (Related to Chapter 7)

Helia Farhood, Xiangjian He, Wenjing Jia, Michael Blumenstein, Hanhui Li, "Counting People Based on Linear, Weighted Local random Forest", *The International Conference on Digital Image Computing, DICTA 2017*. Published in December 2017.

TABLE OF CONTENTS

CERTIFICATE OF ORIGINAL AUTHORSHIP	i
DEDICATION	ii
ACKNOWLEDGMENTS	iii
LIST OF PAPERS (PUBLICATIONS)	iv
LIST OF FIGURES	viii
LIST OF TABLES	xii
ABSTRACT	xiii
Chapter 1: Introduction	1
1.1 Depth Map Estimation	4
1.2 Creating the 3D Point Cloud	7
1.3 Reflectance Recovery	10
Chapter 2: Light-Field Background	14
2.1 Light Field Definition	14
2.2 The Plenoptic Function	18
2.3 Different Representations of the Light Field Data	19
2.4 Applications of Light Field Cameras	22
2.4.1. Geometry Estimation	23
2.4.2. Image Rendering	25
2.4.3. Classifying Material Appearance	25
Chapter 3: Literature Review	27
3.1. General Concepts of 3D Reconstruction	27
3.1.1 3D Modelling from Multiple Images	28
3.1.2 3D Modelling from a Single Image	31
3.2. Light Field Depth Estimation	34
3.2.1 Sub-Aperture Image Matching Based Methods:	34
3.2.2. Epipolar Plane Image (EPI) Based Methods	36
3.2.3. LF Data Machine Learning-Based Approaches	37
3.2.4. Problems of Existing Approaches and Our Solutions	38
3.3. Creation of a 3D Point Cloud	38
3.3.1. 3D Reconstruction from Capturing Multiple Images	39
3.3.2. Creating 3D Point Clouds Based on Deep Learning Techniques	39
3.3.3. 3D Reconstruction from a Single Image Approaches	40
3.3.4. Problems of Existing Approaches and Our Solutions	41

3.4. Light Field Reflectance Recovery.....	42
3.4.1. Handling Reflectance without Using a Light Field Camera	42
3.4.2. Reflectance Recovery Using Light Field Camera.....	43
3.4.3. Problems of Existing Approaches and Our Solutions.....	44
Chapter 4: Proposed Depth Map Estimation Method	46
4.1 Using 4D LF as an Input LF Format.....	48
4.2 Edge Detection and Occlusion Detection	50
4.3 Initial Depth Map Estimation.....	54
4.3.1. Sub-Aperture Image Matching Cue	54
4.3.2. Defocus Cue.....	58
4.3.3. Combination and Filtering	61
4.4 Experimental Results	63
4.4.1. Datasets	64
4.4.2. Methods compared.....	64
4.4.3. Evaluation methods.....	64
4.4.4. Depth Maps for Synthetic Images.....	66
4.4.5. Depth Maps for Real Images	69
Chapter 5: Proposed 3D Point Cloud Approach	72
5.1 Histogram Equalization and Canny Edge Detection.....	74
5.2 Creating a 3D Point Cloud by Transforming the Point-Plane Correspondences	76
5.3 Experimental Results	77
5.3.1. Result on Databases	77
5.3.2. Methods Compared	77
5.3.3. Evaluation Methods	78
Chapter 6: Proposed Enhanced 3D Model Method.....	80
6.1 Histogram Stretching and Equalization	80
6.2 Adding Edge Detection Information.....	82
6.2.1. Fuzzy Logic	82
6.2.2. Feature Matching	85
6.3 Experimental Results	88
6.3.1. Result on Databases	88
6.3.2. Methods Compared.....	90
6.3.3. Evaluation Methods	91
6.3.4. Discussion.....	96
Chapter 7: Reflectance Recovery.....	98

7.1 The Challenge of Material Appearance	98
7.2 Finding Specular Pixels by Colour Estimation	102
7.3 Finding Specular Pixels by Binarized Image Analysis (Intensity based Thresholding) .	105
7.4 Combination of Obtained Specular Pixels and Separation from the Original Image	108
7.5 Experimental Results	109
7.5.1. Result on Databases	110
7.5.2. Methods Compared	112
7.5.3. Quantitative Evaluation.....	114
Chapter 8: Conclusions and Future Work.....	116
REFERENCES	120

LIST OF FIGURES

Figure 1. Overview of the proposed method on a real Light field image captured with a Lytro Illum camera. We used the 4D format of a light field image for the input images. This was followed by edge detection for occlusion identification. Then, initial depth map estimation using correspondence matching and de-focus cues, and finally, combination, filtering, normalization, and final depth estimation.	6
Figure 2. Using a light field image for creating a 3D point cloud.	8
Figure 3. Block diagram of our approach using two different techniques for finding specular components.	13
Figure 4. A sample of our proposed method applied to a real light field image captured by a Lytro camera.	13
Figure 5. Lytro ILLUM Plenoptic Camera. In this thesis we used this Lytro ILLUM camera for some parts of our experimental results.	17
Figure 6. In a common parameterisation of light field images, light passes through the Lens plane (also called the Camera plane or Viewpoint or Angular plane) and then is captured at the Sensor Plane (also called the Focal plane, Image plane or Spatial plane). Each ray of light can be considered to pass along a line that enters the Lens plane at position (u,v) and ends at the Sensor plane at position (s,t)	19
Figure 7. Different formats for representations of light field images. (a) A single sub-aperture image from the 4D format of a light field image, $I_{(s^*,t^*)}(u,v)$, obtained by collecting the light field samples with fixed st coordinates s^* and t^* . (b) A lenslet format light field image in which a LF sub-view $I_{(u^*,v^*)}(s,t)$ is obtained by collecting the samples with fixed uv coordinates $(u^*$ and $v^*)$. (c) An epipolar plane image (EPI) acquired by fixing the coordinates in both the angular and spatial dimension. For instance, the horizontal slice EPI $E_{(v^*,t^*)}(u,s)$ is acquired by fixing v and t , and the vertical EPI $E_{(u^*,s^*)}(v,t)$ is acquired by fixing u and s	21
Figure 8. An illustration of the light field function and post-capture re-focusing. (a) Focus in the foreground. (b) Focus in the background. (c) A diagram of a light field camera setup where a is the distance from the camera plane to sensor plane and b is the distance between the sensor and the micro-lens array.	24
Figure 9. Overall Structure of 3D Reconstruction. Generally there are two main approaches for recovery of a 3D model. One of them is based on a single image and the other uses multiple images.	28
Figure 10 . Overview of the proposed method on a real Light field image captured with a Lytro Illum camera. Our method consists of 4 stages: 1. 4D format of light field image as input images, 2. Edge detection for occlusion identification, 3. Initial depth map estimation using	

correspondence matching and de-focus cues, 4. Combination, filtering, normalization, and final depth estimation.	47
Figure 11. Examples of different shapes of lenslets used in light field images (a lenslet grid can be a hexagonal, circular or square shape). In comparison, the 4D light field format is consistent across different lenslet designs.	49
Figure 12. An occlusion edge on the imaging planes (Sensor plane and lens plane) corresponds to an occluding plane in the 3D space K plane (the light blue shape).	51
Figure 13: A sample of calculating the 2D shift vector $\Delta U(P,l)$	58
Figure 14. Comparing our approach result with the results of other approaches for two synthetic images.	67
Figure 15. Comparing the results of our approach with ground truth on the synthetic images “Papillion” and “Buddha”.....	68
Figure 16: Progression of our proposed method to estimate a final depth map. (a) shows the light field image captured by ourselves with a Lytro camera. (b) shows the initial depth map estimation based only on sub-aperture matching. (c) shows the combined depth map estimation including defocus cues. (d) illustrates the final depth map after applying normalization and filtering.	69
Figure 17. Depth estimates on real images from the JPEG dataset with very challenging conditions (lots of shadow and low-light areas).	70
Figure 18. Comparison of depth estimation results of different methods from a 4D light-field input image. The top-left image shows a single sub-aperture image from a LF dataset. The top-right image shows the result of our depth-map estimation algorithm, while the bottom two images show the results of two prior art depth map estimation algorithms.....	71
Figure 19. Depth estimation on a real image captured with a Lytro Illum camera.....	71
Figure 20. Overview of the proposed method. Our method consists of 4 main stages: 1. 4D format of light field image as input images, 2. Depth map estimation based on using our proposed method for depth map, 3. Canny edge detection results are combined linearly with the histogram stretched depth map, 4. Transforming the point plane correspondences to create a final 3D point cloud.	73
Figure 21. A step by step illustration of our methodology for creating 3D point clouds. (a) Input light field image. (b) Depth map based on our approach. (c) Result of histogram stretching applied to the depth map. (d) The results of applying Canny edge detection to the central sub-aperture image. (e) Image fusion (combining (c) and (d) linearly). (f) 3D point cloud.....	78

Figure 22. Comparing our result with other state-of-the-art methods. (a) Light field input image. (b) Depth map created by the Lytro software. (c) Depth map that is produced by our method. (d) 3D point cloud obtained by re-implementing the method of Dansereau et al. [95]. (e) 3D point cloud obtained by re-implementing the method of Perra et al. [10]. (f) The result of our 3D point cloud reconstruction algorithm.	79
Figure 23. Details of our proposed 3D point cloud estimation method.	83
Figure 24. Depth map modification steps.	84
Figure 25. Obtaining a 3D point cloud based on a real-world light field image (Nature-Flowers) from the JPEG Pleno database.	89
Figure 26. Comparing our result with other state-of-the-art methods. This input light field image is captured by ourselves with a Lytro Illum camera.	91
Figure 27. Comparing our result with other state-of-the-art methods. This input light field image is a real-world LF image (Buildings-Black fence) from the JPEG Pleno database.	92
Figure 28. Comparing the histogram of our result with other state-of-the-art methods for image Green (Figure 27). (a) Shows the histograms obtained from our 3D point cloud method (ideal uniform distribution). (b) Indicates the histograms of Perra et al. [10] method for creating 3D point cloud and (c) shows the result of Dansereau et al. [96]. From the figure it is obvious that the histograms of our method (a) have a more favourable distribution of positions, especially in Z dimension, compared to the two other methods.....	94
Figure 29. Complicated material appearance: The index of refraction of a material controls whether incident light is reflected from the surface, enters the surface and is refracted or some combination of both effects.....	99
Figure 30. An illustration of the different behaviours of diffuse and specular components in a light field image. (a) Analysis of the behaviour of pixels in coloured (RGB) space. It can be seen that the diffuse surface has no colour change with viewpoint (sub-aperture) changes, but the specular surface has colour intensity changes in different viewpoints. (b) In binarized space, the diffuse component stays black, but the specular pixels can change from black to white in different viewpoints.	101
Figure 31. Our experimental results on the surface of fruit and leaves. The first column (a) shows real-world light field images captured by a Lytro camera. The second column (b) demonstrates the combination of obtained specular component from two different strategies of proposed method and the third column (c) shows specular free images.	109
Figure 32. Our experimental results on surfaces of fruit, pearl, glass and metal. The first column (a) shows the real-world light field images captured by a Lytro Illum camera. The second column	

(b) demonstrates the combination of obtained specular component from two different strategies of proposed method and the third column (c) shows specular-free images. The first and second images were captured during daytime and last image was captured during night time..... 111

Figure 33. A comparison of proposed method with two other methods. The first and second images (a,b) are part of Tao et al. dataset and the last light field image (c) is from the Wanner synthetic database (Buddha). 113

LIST OF TABLES

Table 1. The time to compute the depth map calculated on a CORE i5-6300U CPU @2.40GHz , 8 GB RAM running MATLAB R2018a.	65
Table 2. The MSE error of the estimated depth map compared with ground truth (%).	65
Table 3. Comparing numbers of points in synthetic (Papillon) and real (Cat) images.	79
Table 4. Numerical evaluation of 3D point cloud based on LoD ([number of vertices], [number of faces]).	96
Table 5. The percentage MSE error of the specular-free images compared with diffuse images from the Shekhar et al. database [109].	114
Table 6. The BadPix (0.07) error of the specular-free images compared with diffuse images from the Shekhar et al. database [109].	115

ABSTRACT

Representing real world objects on a digital screen is a significant and challenging topic in the area of computer vision and augmented reality. This work addressed the challenge of reconstruction of 3D surfaces with complicated material appearance by using a light field camera. Most recent research uses single images to address this problem, but without using a light field camera, encounter difficulties and limitations to overcome this problem. However, we show that by using a light field camera without user interaction or any requirement for object planarity or symmetry, reconstruction of a 3D model with high accuracy is possible. A light field camera, also known as a Plenoptic camera can capture rich information about the spatial and angular distribution, as well as intensity and colour of light in a single shot. Light field cameras can be used to improve the performance of traditional computer vision problems, such as depth estimation, post-capture refocusing, illumination estimation, and material estimation which are not easy for traditional methods with a standard image. For reconstruction of 3D models, creating a 3D point cloud is essential and is often obtained based on a depth map. As a result, first we developed a robust method to estimate an accurate depth map based on the combination of sub-aperture image matching and defocusing cues for a 4D light field format. The depth map is refined using a fast-weighted median filter providing robustness to noise. Therefore, the proposed approach compared with other state-of-the-art algorithms can estimate depth of real-world images and challenging images more accurately. In the second part, we proposed a novel strategy for the creation of a 3D point cloud from the depth map of a single 4D light field image. The proposed method is based on the transformation of point-plane correspondences. Considering the estimated depth map from the previous part, we applied histogram equalization and histogram stretching to enhance the separation between depth planes. The suggested method avoids feature

extraction, segmentation and the extraction of occlusion masks required by other methods, and due to this, our method can reliably mitigate noise. In the third step, we improved our suggested method to obtain a dense and more accurate three-dimensional (3D) point cloud. We applied intelligent edge detection by using feature matching and fuzzy logic from the central sub-aperture light field image and the depth map. The results showed that our new method can reliably mitigate noise and had the highest level of detail compared to other existing methods. Finally, having obtained the 3D point cloud, we handled the problem of reflectance in complex material appearance. We developed a new strategy to recover reflectance information based on colour analysis as well as brightness analysis of a light field image. Our aim is to separate specular pixels by using two different strategies. On the one hand, we estimated light source colour from different angles of the RGB light field image to find specular pixels. On the other hand, we binarized light field sub-aperture images to obtain brightness areas in different viewpoints and extract specular highlights. Experimental results demonstrate the effectiveness of our method in both synthetic and real-world images compared to other state of the art methods. Overall, 3D reconstruction can cover many applications and solve many problems of computer graphics and computer vision that is still a challenging topic.