Comprehensive light field image dataset

Sheng Liu
KLA-Tencor Corporation
liusheng@stanford.edu

Abstract

Light field imaging is an emerging technology providing 4D (2D space + 2D direction) light information. Here we publicly release a large light field image dataset with 31 different categories that are captured using a Lytro Illum light field camera. Within each category, many scenes are included with different lighting and environmental conditions. Uniquely, for each scene, we also captured images with different camera positions, which provide us an extra degree of freedom for image processing. Moreover, we statistically analyze the camera settings for all 4251 light field images as well as for each category. The statistically analysis includes but is not limited to the focal length, ISO, exposure time, focal depth, etc. We believe the large quantity of light field image dataset can be valuable for a plethora range of potential usage, such as image compression, quality evaluation, computer vision, augmented reality implementation.

1. Introduction

Conventional images captured by a camera only record light intensity information that are incapable to recover the direction of the light entering the camera. The light ray orientations can be used to restore the rich information of the scene and can enable new possibilities such as change of focus, depth of field, numerical aperture, perspective of already captured images via postprocessing. Emerging technology of Light Field photography enables the recording of both intensity and direction of incident light. The basic concept of realizing light field images by integral photography was first introduced by Lippmann [1] over a century ago. The explosion of virtual reality and augmented reality in the past few years highly demand the fast development of light field technology, especially considering the discomforts, such as fatigue, eyestrain, and nausea, current technologies of virtual reality bring. The light field technology provides promising solutions and therefore more realistic stereo experience.

Light field images are mainly recorded by either using multi-camera arrays or by a single camera single sensor but employing additional optical element of a microlens array placed in front of regular sensor. The application of multi-camera solution is limited by the high cost of many cameras, large footprint, expertise in system operation and image processing. The light field single sensor cameras are commercially available from two main manufacturers: Raytrix and Lytro. Raytrix focus on industrial and scientific applications while Lytro initially focus on consumer and semi-professional market. After the introduction of the first-generation light field camera in 2011, Lytro released a second-generation light field camera – Lytro Illum in 2014. More recently, Lytro pivot away consumer-facing digital camera to high-end production cameras and tools by releasing cinematic virtual reality system– Lytro Immerge in 2015.

While the single sensor light field cameras offer compact solution for capturing light field images that provide rich information and new possibilities for image postprocessing, really applications demand solutions of issues such as high storage requirement, low image resolution, image distortion as well as the small angular field of view due to the finite lens size. We are working on solving this problem by releasing publicly a comprehensive light field dataset so researchers can find solutions to the above-mentioned issues. Our dataset includes over 4000 light field images grouped in 31 categories. Uniquely, we intentionally captured multiple images for each scene, providing opportunities to develop algorithms that can create a combined image with a higher resolution, wider field of view and more stereo feeling. Next, we will provide more detailed information of the dataset including statistical histograms that could be helpful for researchers.

2. Dataset

The dataset contains total of 4251 light field images grouped in 31 categories. Several files, each contains different information, are provided for each image. The .eslf.png file is the 4D image file with resolution of 5250 × 7574 pixels. Because the Lytro Illum camera uses a lens array consisting of 14 by 14 microlens, the image after processing will have resolution of 375 × 541. We provide a rendering file ending with .render.png, which provide an in-focus image throughout the whole depth range. Another important file we provide end with .depth.png indicate the depth map using a 16 bit gray scale. .depth.json files contains information of the minimum and maximum value of λ, which is a parameter defined specifically for Lytro Illum camera. Finally, .json files contains all the camera setting information (such as ISO, exposure time, focal length, etc) when the picture is taken.

Table A summarize the categories of the 4251 light field images, and the number of senses and images for each category. On average, there are more than 4 images for each scene.

Table A: Summary of the 4251 light field images in terms of the image category.

Category name	Number of Scenes	Number of Images	Category name	Number of Scenes	Number of Images
bamboo	10	49	glasses	8	40
batteries	5	28	glue	5	30
benches	21	97	keyboards	9	44
bikes	18	83	leaves	142	720
books	5	27	misc	30	167
bottles	7	37	pens_and_pencils	26	136
boxes	31	178	people	7	30
buildings	66	285	phones	5	27
cables	40	231	screws	9	47
cacti	16	83	shelf	5	23
chairs	3	14	signs	38	186
coins	7	33	succulents	19	94
cups	16	80	tables	24	116
drawers	2	9	tools	12	70
fire_hydrants	3	12	trees	84	402
flowers	190	873			

Majority of the categories contains only static objects, as the Fig. 1 (a)-(c) show. This set of images can be ideal to practice keypoint matching. There are also categories contain dynamic scenes or object. Obvious example is the category “people”, as Fig. 1 (d)-(f) show, since people could have different gesture, position, facial expression for each image. Moreover, in this particular example, it is challenging because the number of people are different for each image. Less obvious examples of dynamic categories include “trees”, “leaves”, depending on the weather condition.

We analyze the camera settings of the 4251 images statistically. A few examples are shown in Fig. 2. Using the histograms, we can easily identify the outliers of each camera setting. And they could be interesting and useful for light field image processing. For example, the image with the longest exposure time shows indoor benches. We could investigate if the image is blurry because of the long exposure time. Moreover, we find that the wall in the image is over exposed. The challenge would be how we can reduce the motion blur and how we can recover the over-exposed information using other light field images, which were captured with different exposure time, and different camera position. Fig. 2(c) also shows indoor images under low light were taken using high ISO settings. Therefore, possible high sensor noise can be found in these images.

One of the most powerful feature of light field images is the recovery of field depth information. In Fig. 3, we show that the gray depth map with values bound between 0 and 65535 (16 bits) can be converted first to λ and then to slope map. As a reference, Fig. 3 (a) shows a rendered all-in-focus image. Due to the close distance of the objects, the gray depth map is relatively dark. Using the upper and lower bounds of the λ value obtained from **.depth.json** files, we can convert the gray scale image to the λ map. Next, we use the equation of $Slope = \lambda \cdot \frac{1.4}{20}$ to convert the λ map to the slope map. Note that the negative or positive slopes indicate objects are closer or further away than the focus (slope = 0) of the camera, respectively, when the image was taken.

Figure 1. Two scene examples that include static objects (a)-(c) and dynamic objects (d)-(f). Images also show how the camera position changes within each scene.

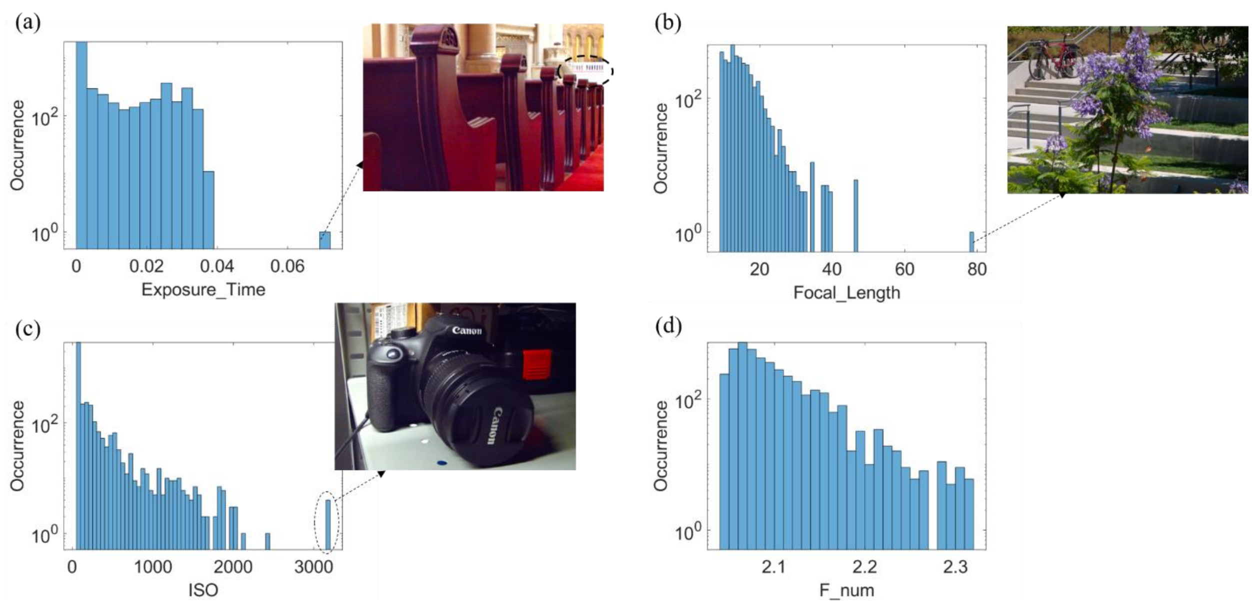


Figure 2. Four histogram examples of exposure time, focal length, ISO and F# of 4251 images when they were captured. The photos of some outliers of the histograms are also shown.

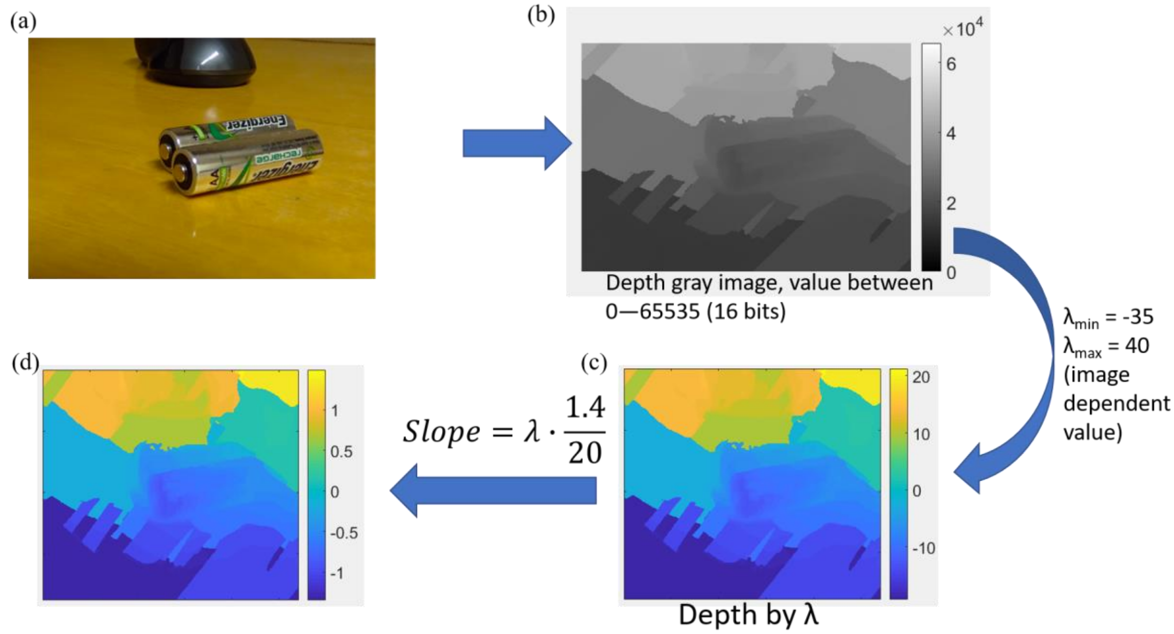
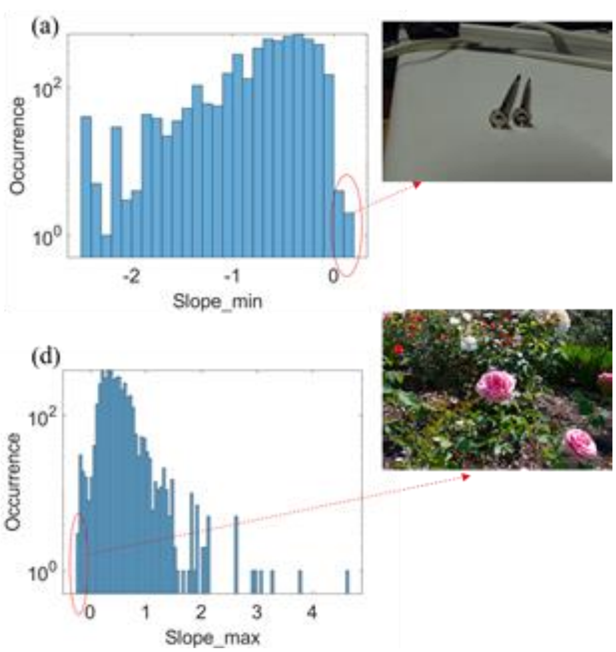


Figure 3. Converting gray depth map to slope map. (a) Rendered all-in focus image of two batteries sitting on a desk. (b) Gray depth image with file name ending .depth.png. Depth map using (c) λ and (d) slope.

To confirm the meaning of slope map, we statistically studied the maximum and minimum slope value of the 4251 images, as Fig. 4 shows. Interestingly, there are images with the minimum slope value larger than 0 meaning that meaning all the objects in the image are further away from where the camera focus was set.



Similarly, we also find some images with the max slope smaller than 0, meaning all the objects in the image are closer than where the camera focus was set. This behavior was confirmed by pan the 4D images using the function of LFDisMousePan (Light field Matlab toolbox).

Last, we provide an excel file containing all the information for each image. Part of the excel file is shown in Table B.

3. Conclusion

In conclusion, we aim to release a comprehensive light field image dataset. We statistically studied the camera setting, depth map of over 4000 light field images. We hope researchers could take advantage of this dataset and help to advance the technology of light field.

References

[1] 1. Olga R., et al, "ImageNet Large Scale Visual Recognition Challenge", arXiv:1409.0575v3 (2015)

[2] 2. Alireza G. et al, "LCAV-31: A Dataset for Light Field Object Recognition", Proc. of SPIE-IS&T/ Vol. 9020 902014-1

[3] 3. Marius Cordts, et al, "The cityscapes datasetfor Semantic Urban Scene Understanding", arXiv:1604.01685v2 (2016)

[4] 4. Andreas G., et al, "Vision meets robotics: the KITTI dataset", The International Journal of Robotics Research , **32**, 1231 (2013)

[5] 5. Martin R., et al, "New light field image dataset", 8th International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, Portugal, 2016

Image_Name	Category	Focal_Length	Exposure_Time	ISO	F_num	Lambda_Min	Lambda_Max	Slope_min	Slope_max	FIle_location
IMG_3983	bamboo	17.15533	0.001527	80	2.117596	-35	8	-0.62108	0.317537	F:\IllumDataset\detailed_json\bamboo
IMG_3984	bamboo	17.15533	0.001651	80	2.117596	-35	8	-0.67147	0.332373	F:\IllumDataset\detailed_json\bamboo
IMG_3985	bamboo	17.15533	0.001651	80	2.117596	-35	8	-0.77141	0.352168	F:\IllumDataset\detailed_json\bamboo
IMG_5145	batteries	12.90831	0.024264	170	2.073591	-35	28	-1.03269	0.721284	F:\IllumDataset\detailed_json\batteries
IMG_5146	batteries	12.90831	0.024264	170	2.073591	-35	28	-1.50057	0.707489	F:\IllumDataset\detailed_json\batteries
IMG_1478	benches	9.458095	0.000702	80	2.048967	-35	4	-0.27737	0.207642	F:\IllumDataset\detailed_json\benches
IMG_1479	benches	9.504048	0.000619	80	2.04946	-35	4	-0.10508	0.206434	F:\IllumDataset\detailed_json\benches