DYNAMIC RANGE ENHANCEMENT IN LIGHT FIELD IMAGING SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF
ENGINEERING AND NATURAL SCIENCES
OF ISTANBUL MEDIPOL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL, ELECTRONICS ENGINEERING AND CYBER SYSTEMS

By
Abdul Wahab
April, 2019

DYNAMIC RANGE ENHANCEMENT IN LIGHT FIELD IMAGING SYSTEMS

By Abdul Wahab

April, 2019

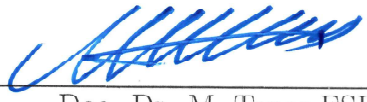
We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.



Prof. Dr. Bahadır Kürşat GÜNTÜRK(Advisor)

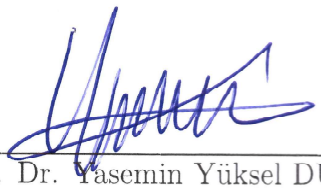


Prof. Dr. Hasan F. ATES



Doç. Dr. M. Taner ESKIL

Approved by the Graduate School of Engineering and Natural Sciences:



Doç. Dr. Yasemin Yüksel DURMAZ

Director of the Graduate School of Engineering and Natural Sciences

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Abdul Wahab

Signature : 

ABSTRACT

DYNAMIC RANGE ENHANCEMENT IN LIGHT FIELD IMAGING SYSTEMS

Abdul Wahab
M.S. in Electrical, Electronics Engineering and Cyber Systems
Advisor: Prof. Dr. Bahadır Kürşat GÜNTÜRK
April, 2019

High dynamic range imaging has been an active research area in computational photography over the last two decades. The main reason is that the dynamic range of a traditional camera is typically less than the dynamic range of the environment; and this affects the performance in digital photography and machine vision applications.

An emerging research area in computational photography is light field imaging. Light field (or plenoptic) of a scene can be obtained by capturing light rays from different directions separately, as opposed to the traditional imaging systems, where light rays from different directions are accumulated, losing the directional light information. Light field imaging has been of great interest due to its post-capture capabilities, including refocusing, perspective change and depth estimation. Microlens array based cameras that recently emerged have made light field acquisition process a practical task.

In this study, we present a method to extract high dynamic range image from a single capture. The method uses a plenoptic camera with minor optical modification. The optical modification is an optical mask, placed in front of the main lens to increase the vignetting effect, which is the darkening of an image towards its borders due to loss of light. As a result, different parts of the dynamic range are captured with different sub-aperture images of the light field.

Keywords: Light field imaging, high dynamic range, plenoptic camera, camera array.

ÖZET

IŞIK ALANI GÖRÜNTÜLEME SISTEMLERINDE DINAMIK ARALIK IYILEŞTIRME

Abdul Wahab
Elektrik-Elektronik Mühendisliği ve Siber Sistemler, Yüksek Lisans
Tez Danışmanı: Prof. Dr. Bahadır Kürşat GÜNTÜRK
Nisan, 2019

Yüksek dinamik aralıklı görüntüleme son yirmi yıldır hesaplamalı görüntüleme alanında aktif bir araştırma alanı olarak devam etmektedir. Bundaki temel sebep geleneksel kameraların dinamik aralığının gerçek dünyada karşılaşılabilecek dinamik aralıktan küçük olması, ve bunun da dijital fotoğrafçılık ve makine görmesi uygulamalarındaki performansı düşürmesidir.

Hesaplamalı görüntülemede yükselen bir araştırma konusu da ışık alanı görüntülemedir. Işık alanı (ya da plenoptik) görüntüleme ışınların kameraya farklı yönlerde gelen ışınların ayrı ayrı kaydedilmesiyle elde edilir; diğer taraftan, geleneksel görüntüleme sistemlerinde farklı yönlerden gelen ışınlar toplanarak açısal ışık bilgisi kaybedilir. Işık alanı görüntüleme; yeniden odaklama, perspektif değiştirme, derinlik kestirme gibi çekim sonrası kabiliyetlerinden dolayı oldukça ilgi çeken bir konudur. Yakın zamanda ortaya çıkan, mikrolens dizisine dayalı kameralar ışık alanı kaydını pratik hale getirmiştir.

Bu çalışmada, tek bir çekimden yüksek dinamik aralıklı görüntü elde etmeyi sağlayan bir yöntem sunuyoruz. Bu yöntem, küçük optik değişiklikler yapılmış bir plenoptik kameraya dayanmaktadır. Optik değişiklik, ana lensin önüne yerleştirilmiş, ve ışık azalmasıyla kenarlara doğru karanlıklaşmayı arttıran bir optik maskeden ibarettir. Sonuç olarak, ışık alanının farklı perspektif görüntüleri dinamik aralığın farklı kısımlarını kaydetmektedir.

Anahtar sözcükler: Işık alanı görüntüleme, yüksek dinamik aralık, plenoptik kamera, kamera dizisi.

Acknowledgement

First of all, I am really greatful to Almighty Allah who has given me strength to finish my research work

I would like to thank to my supervisor Prof. Bahadır Kürşat Güntürk, without his patience, support and knowledge it was impossible for me to finish my thesis work. He gave me full confidence and help throughout my master's program. During his supervision, I was able to learn basic and advanced techniques of computer vision field.

In the end, I would like to thank to my family who gave me permission to go abroad for higher education. Special thank to my lab-mates Zeshan Alam, Umair Mukati and Shahzeb who always helped me in finding the solutions of my research problems.

Contents

1 Introduction **1**

1.1 Dynamic range 2

1.2 Acquisition of HDR image 3

1.3 High dynamic range display 5

1.4 Light field 6

1.5 Acquisition of light field 7

1.6 Light field camera 8

1.7 Post-capture capabilities of light field camera 9

1.8 Motivation 11

1.9 Contribution 12

1.10 Outline 13

2 High dynamic range imaging using camera array **14**

2.1 Related work 14

2.2 Alignment of the images 16

2.3 Radiance map estimation 17

2.4 Tone-mapping 19

2.5 De-ghosting 22

2.6 Dynamic range 23

2.7 Quantification of the dynamic range of camera array 24

2.8 Experimental results 27

3 High dynamic range image using plenoptic camera 30

3.1 Related work 31

3.2 Optical design 32

3.3 Light field pre-processing 33

3.4 High dynamic range imaging 34

3.4.1 Geometric registration of light field data 34

3.4.2 Irradiance estimation 38

3.5 Quantification of light field 40

3.6 Experimental results 43

4 Conclusion 49

List of Figures

1.1 Two image captures of the same scene. (a) Over-exposed. (b) Under-exposed. 1

1.2 The dynamic range of human visual system [?]. 3

1.3 Parametrization of light field with two parallel plane, in each representation u and v serves as primary arguments. The last two arguments are parametrized by; (a) Global coordinates of s and t [?]. (b) Angular coordinates θ and ϕ representing the angle of ray after intersecting with uv plane. (c) Local coordinate of s and t, some time also referred as slope of the angle of the ray intersected at uv plane. (The illustration is taken from [?]). 7

1.4 Camera array by Stanford University [?]. 8

1.5 (a) First generation Lytro camera [?]. (b) An illustration of a light field camera. 9

1.6 (a, d) Optical diagram showing focus at different depth; (b, e) Red squares shows pixel picked from raw lenslet image averaging them to get a refocused image; (c, f) Refocused images. (The illustration is taken from [?]). 10

1.7 (a, d) Optical diagram demonstrating the effect of placing a virtual aperture stop; (b, e) Region marked with red square shows the pixel region averaged to get the projected point; (c, f) Reconstructed image. (The illustration is taken from [**?**]). 11

2.1 Three different exposure images captured by our imaging system. (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image. 17

2.2 Hardware setup of the imaging system. 17

2.3 Photometric registration. (a) High-exposure image on mid-exposure image. (b) Low-exposure image on mid-exposure image. 18

2.4 Residual between high-exposure and mid-exposure images. (a) Before registration. (b) After registration. 18

2.5 Residual between low-exposure and mid-exposure images. (a) Before registration. (b) After registration. 19

2.6 Response function of each camera is determined by the technique described in [**?**]. 20

2.7 The weight function for calculating the radiance maps. 20

2.8 Dynamic range of the camera array. 25

2.9 Dynamic range of a single camera. 26

2.10 White images. 26

2.11 Noise measurement. 27

2.12 (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image. 28

2.13 HDR image with tone-mapping. 28

2.14 (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image. 28

2.15 Demonstration of occluded artifacts after registration in (a) High-exposure image. (b) Low-exposure image. 28

2.16 Demonstration of images after de-ghosting operation (a) High-exposure image. (b) Low-exposure image. 29

2.17 (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image. 29

2.18 HDR image with tone-mapping. 29

3.1 (a) The optical setup. (b) Schematic of the optical design. 33

3.2 Lenslet image of a Lytro camera. 34

3.3 Captured light field with zoom in version of extreme corner middle row images and middle column images. 35

3.4 Photometric registration of extreme corner images of a middle row. 36

3.5 Photometric registration of extreme corner images of a middle column. 36

3.6 Residual between extreme corner images of a middle row. (a) Before registration. (b) After registration. 37

3.7 Residual between extreme corner images of a middle column. (a) Before registration. (b) After registration. 37

3.8 Geometric registration of sub-aperture images of light field. 38

3.9 White planar scene. 40

3.10 Captured light field. 41

3.11 Dynamic range of sub-aperture images. 42

3.12 Dynamic range of sub-aperture images. 42

3.13 Dynamic range of light field. 43

3.14 (a) High-exposure image. (b) Low-exposure image. 45

3.15 (a) Non-registered HDR image. (b) Registered HDR image. . . . 45

3.16 Captured light fields. (a) With optical filter. (b) Without optical filter. 45

3.17 (a) High-exposure image. (b) Low-exposure image. (c) HDR image. 46

3.18 (a) High-exposure image. (b) Low-exposure image. (c) HDR image. 46

3.19 (a) High-exposure image. (b) Low-exposure image. (c) HDR image. 46

3.20 (a) HDR image without optical filter.(b) HDR image with optical filter. 47

3.21 Comparison between histograms of HDR images. (a) Without optical filter. (b) With optical filter. 47

3.22 (a) Captured light field. (b) HDR light field. 47

3.23 Post-capture digital refocusing. (Left column) Near refocus. (Middle column) Mid refocus. (Right column) Far refocus. 48

List of Tables

1.1 Different dynamic range measures. Z_{peak} represent the maximum luminance value and Z_{noise} represent the noise level [**?**]. 3

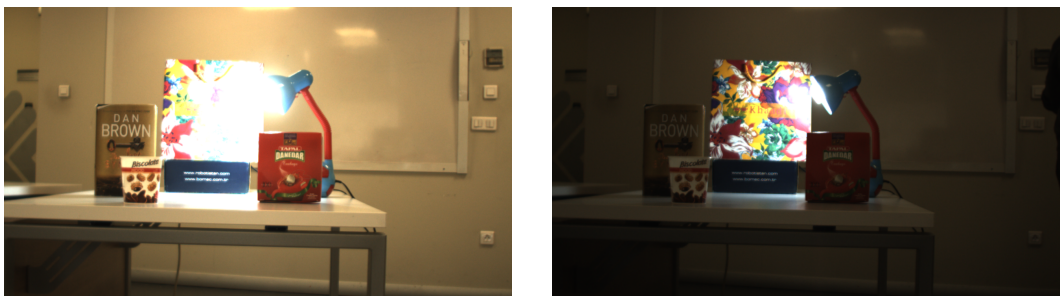
1.2 Comparision of dynamic range of professional digital cameras [**?**]. 5

Chapter 1

Introduction

Computational photography is an emerging area of research, with the goal to overcome the limitations of traditional imaging systems through the use of computational power and modified/unconventional optical components. A limitation of traditional imaging systems is limited dynamic range. Figure **??** shows two images of the same scene. The first image is captured with a high exposure value, and it has over-saturated regions. The second image is captured with a low exposure value; while some regions that are over-saturated in the first image can now be seen, this second image has under-exposed regions.

Computational photography techniques, such as, [**?**], [**?**], [**?**], [**?**], [**?**], which



(a) (b)

Figure 1.1: Two image captures of the same scene. (a) Over-exposed. (b) Under-exposed.

require sequential capture of multi-exposure low dynamic range (LDR) images before combining them to generate a high dynamic range (HDR) image can be used to address this limited dynamic range problem, which is also the topic of this thesis.

1.1 Dynamic range

The dynamic range of a scene can be defined as the ratio between the darkest and brightest point in the scene. There are several ways to quantify the dynamic range of an imaging system. For example, the dynamic range can be defined as the ratio between the sensor saturation level and the noise floor. This signal-to-noise ratio (SNR) is the ratio between the point that the sensor becomes saturated to the minimum level just above the noise of the sensor. It should be noted that the overall dynamic range is limited by not only the sensor but also the lens, A/D conversion circuits, and signal processing required in the imaging pipeline.

In the field of photography, the measure of dynamic range is f-stops or steps. F-stop depends on the shutter speed, ISO and aperture size. It is usually expressed in logarithmic units with base 2. Generally, 12 to 14 f-stops of exposure latitude is considered a good dynamic range.

In display systems, contrast ratio is a measure of dynamic range. It is defined as the ratio of the lightest color to the darkest that can be produced by the display system. To avoid infinity, the darkest point is considered as the first observable level that is above zero. The above-mentioned measures of dynamic range are mentioned in Table **??**.

In addition to the dynamic range of imaging systems, it is important to consider the dynamic range of the human eye. A human eye can easily see the details of the scene in light as low as that of stars to the light as bright as that of the sun's. Quantitatively speaking, the dynamic range of human eye extends from 10^{-6} to 10^{6} candela/m^2 (shown in Figure **??**), which is significantly larger than

Measure	Formula
Signal-to-Noise Ratio (SNR)	$20 \log_{10} \frac{Z_{peak}}{Z_{noise}}$
f-stop	$\log_2 \frac{Z_{peak}}{Z_{noise}}$
Contrast Ratio (CR)	$1 : \frac{Z_{peak}}{Z_{noise}}$

Table 1.1: Different dynamic range measures. Z_{peak} represent the maximum luminance value and Z_{noise} represent the noise level [**?**].

conventional imaging devices. The human visual system dynamically adjusts its light sensitivity based on the amount of light in the environment. The overall dynamic range of human eye is nearly 24 f-stops. For the accurate comparision of the dynamic range of human eye with that of a single photograph we have to consider the instantaneous dynamic range of the human visual system, which is nearly 10-14 f-stops [**?**].

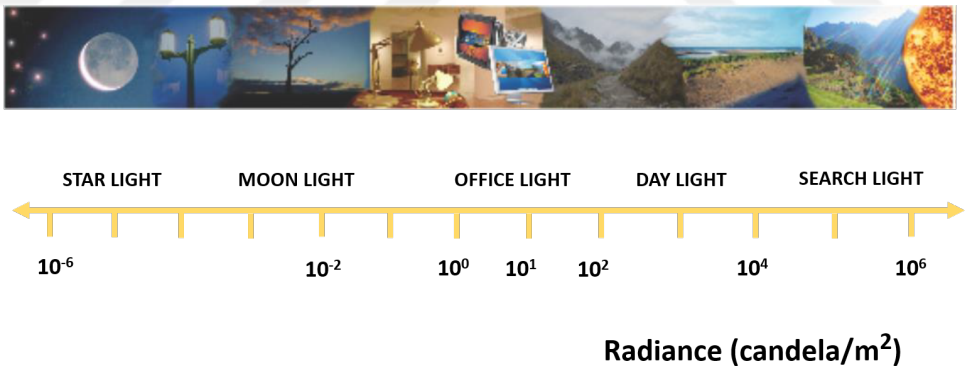


Figure 1.2: The dynamic range of human visual system [**?**].

1.2 Acquisition of HDR image

This section describes capturing of HDR image. The goal of HDR acquisition methods is to overcome the limited dynamic range problem of the camera sensor. Dynamic range of a digital camera depends on ISO levels and bit depths. As the ISO level of the camera is increased, noise level also increases causing the dynamic range to decrease. Bit depth refers to the number of bits used to represent each

pixel. These values are infact representation of the irradiance as intensity of the striking light. The more number of bits contained in each pixel, the more range of intensities are available for each pixel [**?**]. Professional digital cameras can capture extended dynamic range images and store them in raw format. Raw format can have a range of 16 bits. More accurate scene luminance information can be achieved in these formats because the dynamic range is not limited due to image processing in a camera like sharpening. Usually, the dynamic range of an image is compressed to display on LDR display. Due to this, significant limitations occur in image processing, analysis and displaying the data. Dynamic ranges of several professional digital cameras are shown in Table **??**.

Several techniques have been proposed to capture HDR data. Nayar and Mitsunaga [**?**] proposed a solution with single sensor in which they change the exposure of each individual pixel by placing an optical mask adjacent to the image detector. In this way, they extend the dynamic range at the cost of spatial resolution. A similar approach has been used in [**?**], [**?**] which utilizes an optical attenuator by which exposure of individual sensor pixel can adapt to scene radiance. It also requires calculating the transmittance of the attenuator which is done using the brightness measured by the pixel. The major drawback of these methods include the usage of specialized hardware.

Another popular technique of HDR imaging is to capture sequential images of the same scene with varying exposures. The exposure of each image can be adjusted by varying the shutter speed or the f-number of the optical lens. By doing this, multiple images of the same scene with different exposure times can be captured and then combined to produce an HDR image. Response function of the imaging system can also be computed by acquired images. This approach has been used in [**?**], [**?**], [**?**], [**?**], [**?**]. However, this method can only be used for static scenes.

The static scene limitation faced in the above mentioned technique can be resolved by using multiple imaging systems. These methods involve the use of prisms [**?**] or beam-splitters [**?**], [**?**] which allows the extension of dynamic range by decreasing the image resolution. Another well-known technique for generating

	Canon 6D	Nikon D610	Nikon D810	Sony A7S	Sony A7R
Dynamic Range	12.1 stops	14.4 stops	14.8 stops	13.2 stops	14.1 stops

Table 1.2: Comparision of dynamic range of professional digital cameras [**?**].

an HDR image uses Standford's custom built camera array [**?**] to capture multi-exposure images at once. One of the drawbacks of these camera arrays is that they are very bulky and are difficult to move. Apart from their size limitations, baseline between the adjacent cameras is large enough to get easily affected by the occlusion phenomenon and results in ghost-like artifacts in the HDR image. The commercially available handheld light field camera can tackle the issue of bulky volume and large baselines of the camera array by capturing multiple sub-aperture images in a single shot. In [**?**], a focused plenoptic was utilized by introducing an optical mask that individually controls the exposure of each sub-aperture image. These methods allow the capture of HDR images in dynamic scenes.

Some devices are capable of capturing HDR data. There are two primary approaches taken in regard to extend the dynamic range. First one uses a CMOS sensor to compute the logarithm of the irradiance in analog domain [**?**]. The other method consists of focusing on avoiding saturation point of the sensor. This requires computation of the time required by a pixel to attain full capacity [**?**].

1.3 High dynamic range display

Limited dynamic range problem also occurs in display systems. HDR data cannot be displayed on conventional digital technology such as LCD, CRT, plasma screens and projectors. Some investigators have developed high dynamic range display that are able to produce brightness levels and contrast ratio similar to real world scene. Seetzen et al. [**?**] proposed dynamic control of the backlight luminance using light-emmiting diode (LED) with the conventional liquid crystal display (LCD). In this method, the intensity of each LED can be programmed

individually. Additional rear illumination is provided by the LEDs. Similar approach has been applied to increase the dynamic range of a printer [?] and microscopy [?]. Apart from this, Hoskinson et al. [?] proposed a solution to enhance the dynamic range of the projector using analog micromirrors to relocate the light inside a projector from the dark parts towards the brighter parts of the image.

Another popular technique used to display HDR images on a conventional digital display system is known as tone-mapping. The tone-mapping algorithm contains two broad categories: one is the global operator [?], [?] the other is the local operator [?], [?], [?], [?]. The global tone-mapping algorithm applies the compression curve on the whole image which is invariant to the pixel position. This method is fast as there is no intensity inversing, i.e.; brighter pixels of the HDR image remains bright after the tone-mapping operation. The drawback of the global tone-mapping algorithm consists the lack of local contrast and loss of the minute details. On the other hand, local tone-mapping algorithm depends on the pixel position of the image. To preserve the details this algorithm uses neighbouring information of every pixel. The local tone-mapping technique require a significant processing time.

1.4 Light field

The concept of light field technology has a long history. Gershun [?] defined the concept of light field in space and time. He stated that light is travelling in every direction through every point in space. Adelson and Bergen [?] described the distribution of light field as a plenoptic function. They represented the plenoptic function with five dimensions, where three coordinates represent the position in space and two coordinates represent the angle of the ray leaving from the point. In this representation, they discard the time dimension to assume static scene. The frequency dimension is replaced by a 3 color channels (RGB). Even this simplified version of the plenoptic function is expensive, difficult to store and reconstruct. For further simplification, Levoy [?] and Gortler [?] represent light rays in free

space as the energy of light rays remains constant. They reduce its dimension from 5D plenoptic function to 4D light field. The four-dimensional representation of light field is shown in Figure **??** and is used in variety of applications such as [**?**], [**?**], [**?**]. This representation reduces the complexity of storage and processing.

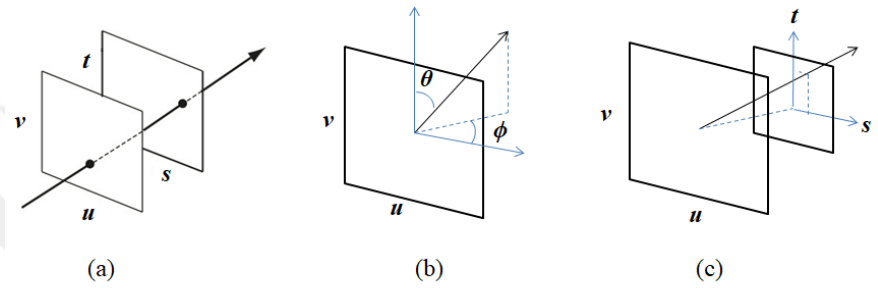


Figure 1.3: Parametrization of light field with two parallel plane, in each representation u and v serves as primary arguments. The last two arguments are parametrized by; (a) Global coordinates of s and t [**?**]. (b) Angular coordinates θ and ϕ representing the angle of ray after intersecting with uv plane. (c) Local coordinate of s and t, some time also referred as slope of the angle of the ray intersected at uv plane. (The illustration is taken from [**?**]).

1.5 Acquisition of light field

To capture 4D information of a light field, setup using the multi-camera array based system can be used [**?**], [**?**]. A famous multi-camera array based system was developed by Stanford University [**?**]. In this setup, series of cameras usually with same specifications are installed in rows and columns of the grid is shown in Figure **??**. The spatial resolution of light field data is represented by the spatial resolution of each camera while the angular resolution depends upon the number of cameras used in the grid. For example, camera array which was originally designed had a grid of 8x12. In this case, the angular resolution of light field data is 8x12. This method is not recommended to capture light field data as it is costly due to large number of cameras being used and also requires a large area. However, the quality of light field data capture from this method is far better than any other method.

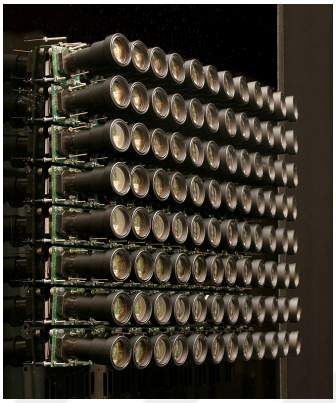


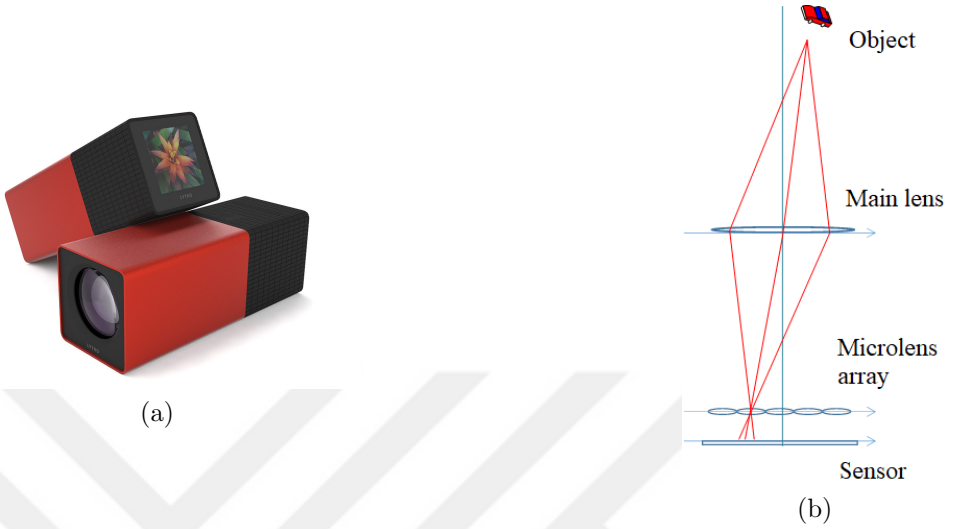
Figure 1.4: Camera array by Stanford University [?].

Microlens array (MLA) based technique is also used to capture light field data. This approached is employed in an imging devices proposed by Ng to capture light field data by placing a microlens array at the focal plane of the main lens [?]. Due to its compact size and cost, it has gained significant attention by the research community.

In the market, some other products which are based on microlens array are also available such as Raytrix [?] and plenoptic camera 2.0 [?]. In these devices microlens array is placed at the intermediate image plane. Reserchers also proposed other techniques to capture light field data such as moving the camera on a gantry [?], lens array [?], coded mask [?] and kaleidoscope-like optics [?].

1.6 Light field camera

The first commercial plenoptic camera is designed by Lytro [?]. The Lytro contains the main lens which acts same as the conventional camera, a microlens array and a photosensor which is shown in Figure **??**. The main lens forms an image on the microlens array, which then seperates these rays based on the direction. The rays are recorded by the photosensor plane [?]. The illustration of a light field camera is shown in Figure **??**. The spatial resolution of the plenoptic camera depends on the size of the microlenses and the angular resolution depends on the number of pixels behind each microlens. The first generation Lytro camera has



(a) (b)

Figure 1.5: (a) First generation Lytro camera [**?**]. (b) An illustration of a light field camera.

a spatial resolution around 380x380 and angular resolution of 11x11.

1.7 Post-capture capabilities of light field camera

Light field imaging has become an interesting research topic due to its post-capture capabilities like refocusing, perspective change and depth estimation. Due to this, light field imaging is used in various applications like robotics, medical imaging, microscopy, computer vision etc. In traditional photography, modifications cannot be done after capturing an image because there is no direction information of the light rays. However, in light field imaging the angular information enables post processing of the image.

One of the key post-capture capabilities of light field imaging is refocusing. In light field camera, the sub-aperture image is formed from the light rays which pass through the same area of the main lens. All sub-aperture images have a different perspective with respect to each other. Refocusing can be achieved by shifting and adding the sub-aperture images. Due to this, the image can be focused at

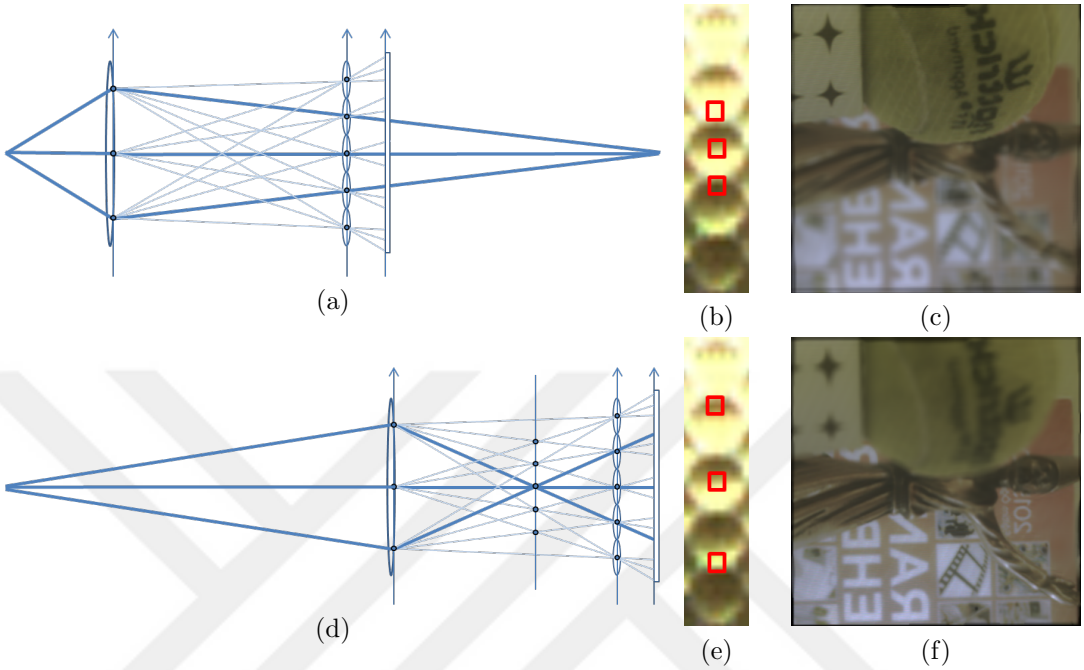


Figure 1.6: (a, d) Optical diagram showing focus at different depth; (b, e) Red squares shows pixel picked from raw lenslet image averaging them to get a refocused image; (c, f) Refocused images. (The illustration is taken from [**?**]).

any depth in the scene [**?**]. Figure **??** demonstrates refocusing on different depths.

Post processing of light field data also allows us to do aperture adjustment. The number of sub-aperture images is defined by the amount of opening of the aperture of the main lens. On the other hand, there is no such feature in the traditional camera. The opening of an aperture defines the amount of light entering in an optical system. The small opening of an aperture creates an overall sharp image while large opening blurs the image, except for the focus region. Figure **??** demonstrates the aperture size control that defines the depth of field. The large opening of the aperture allows more light rays to enter causing reduction in the depth of field and vice versa. In addition, the depth of field of the scene can be adjusted by taking the average of sub-aperture images within the radius [**?**].

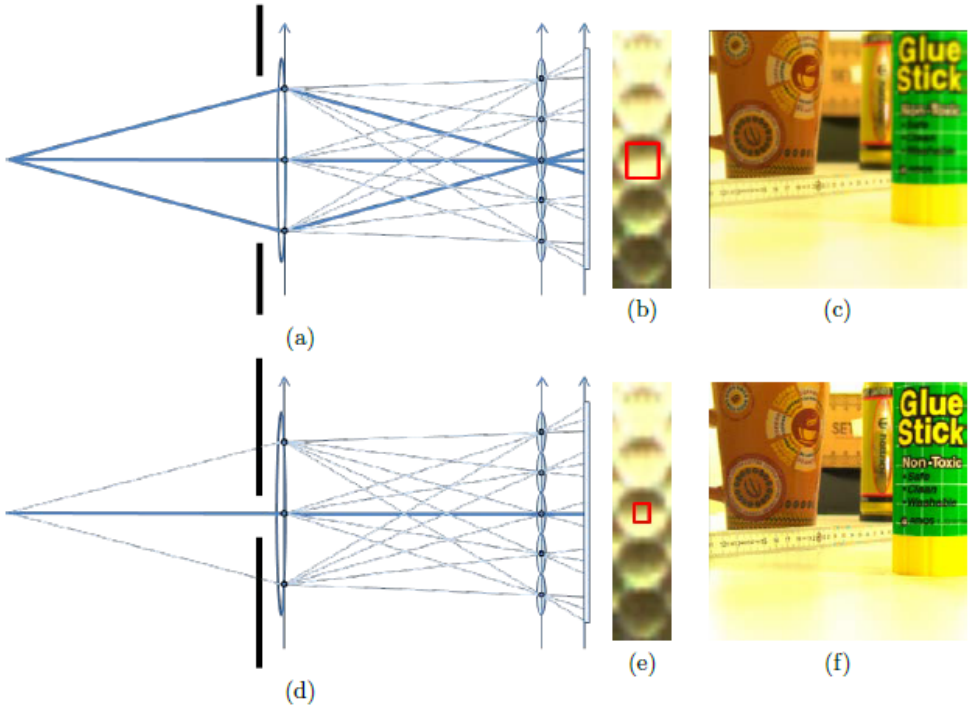


Figure 1.7: (a, d) Optical diagram demonstrating the effect of placing a virtual aperture stop; (b, e) Region marked with red square shows the pixel region averaged to get the projected point; (c, f) Reconstructed image. (The illustration is taken from [**?**]).

1.8 Motivation

Conventional imaging systems have limited dynamic range, adversely affecting performance in industrial and scientific applications. One way to overcome the limited dynamic range problem is through sequential capturing with a conventional camera, where each capture has a different exposure value. This method can safely be applied to static scenes and when the camera position is fixed. For dynamic scenes where there are moving objects and when the camera position cannot be fixed, the fusion of the captured images may create the so-called "ghosting" artifacts, which are the blending artifacts of different objects. The ghosting artifacts can be reduced through "de-ghosting" algorithms; however, they are costly and the performance may be below expectations. Another way

to capture high dynamic images using conventional cameras is to form a camera array, each capturing an image with a different exposure value, and finally merging the images. By using the same exposure time for each camera, the dynamic scene problem can be avoided. However, the camera array approach has its own problem: The wide baseline between the cameras can introduce occlusion issue, which will again cause "ghosting" artifacts during the fusion process. Our goal is develop an HDR imaging system that does not have the drawbacks of the multi-shot and camera array approaches.

1.9 Contribution

This work overcomes the issue of limited dynamic range of conventional imaging systems by employing a minor optical modification of microlens array (MLA) based light field camera. In this way, it is possible to capture a high dynamic range image from a single shot. The proposed optical modification is an optical mask placed in front of the main lens to increase the vignetting effect, which increases the darkening towards borders of the image plane due to loss of light. As a result, different parts of the dynamic range of the scene are captured with different sub-aperture images of the light field. These sub-aperture images are then used to compute high dynamic range image. In addition to extending the dynamic range, light field capabilities are also preserved. The two common ways to obtain high dynamic range images are (i) single camera and multi-shot approach, and (ii) camera array approach. The first approach is problematic for dynamic scenes; moving objects can cause the so-called "ghosting" artifacts. The second approach is costly and may also suffer from ghosting artifacts due to occlusion resulting from the large baseline between the cameras. The proposed method can capture HDR image from a single shot, avoiding the problems due to dynamic scenes. The method does not suffer from the occlusion issue either since the baseline between the sub-aperture images of the light field is small. This work is published as a conference paper in [**?**].

1.10 Outline

In Chapter 2, we use the camera array approach to capture HDR images. The design is similar to that of [**?**], with a smaller and much more economical camera array. Due to the large baseline, camera array suffers from occlusion effect, prompting the use of de-ghosting algorithm. This is followed by a quantitative analysis of the dynamic range.

In Chapter 3, we propose the enhancement of the dynamic range using a plenoptic camera. The advantage of using a plenoptic camera in generating an HDR image is also discussed in this chapter. The most critical part in this study is the registration of sub-aperture images of light field camera. At the end of this chapter, results are presented and discussed.

In Chapter 4, the thesis is concluded by summarizing our work and pointing out the potential directions of future research related to it.

Chapter 2

High dynamic range imaging using camera array

Combining multiple cameras to form camera arrays provide several capabilities for imaging. The configuration of the camera array in this work helps to overcome the limitations of traditional cameras in terms of their dynamic range, allowing the capture of HDR image by combining multiple images with different exposure times. Camera arrays are capable of providing high spatial resolution. One of the drawbacks of using camera array is large baseline between the cameras which can cause occlusion artifacts in the final HDR image. This can be resolved by using de-ghosting techniques. Quantitative analysis of the dynamic range has also been done to compute the dynamic range of the overall system. It is shown that the proposed solution produced good quality HDR image.

2.1 Related work

The high dynamic range imaging is a well-studied topic in computer vision field. Natural scenes exhibit large dynamic range which is impossible to capture from

the traditional imaging system which provides only 8 bit (0-255 levels) of brightness information at each pixel. Due to this, capturing a real world scene with conventional camera results in loss of information in dark or bright parts of the image. In order to capture the exact details of the real world scene, the dynamic range of the camera needs to be enhanced.

In the literature, there are several methods to enhance the dynamic range. We can characterize these methods into two broad categories, one is hardware design methods and another is multi-exposure techniques. They [?], proposed a solution to adjust the exposure of each individual pixel by placing an optical mask adjacent to the image detector. The modified approach has been used in [?] and [?], where optical attenuator is used to record different exposure values of individual pixels and finally combined them to generate HDR image. Another sensor design methodology has been proposed in [?], where the best exposure has been determined locally by alternating 'cliques' of the sensor and generate HDR image using Poisson solver.

Another method is to enhance the dynamic range is to capture multiple images of the same scene with different exposure settings. A single camera approach has been used in [?], where different exposure time images are captured sequentially of the static scene from a fixed point. There are another approaches to capture multiple images with different exposure times that has been applied in [?], [?], [?], [?], [?]. These techniques are suitable for static scenes.

Many techniques have been proposed for dynamic scenes as well. One of the methods uses beam splitter [?], allowing capture of multiple copies of the scene. The exposure of each of the image is controlled either by an optical attenuator or by changing the exposure time. This approach extends the dynamic range but makes it difficult to capture the details in dark areas of the scene. Another famous method for dynamic scene was proposed by one of the group from Stanford University which uses a camera array [?]. This method captures multiple exposure images at once. In these methods, objects in a scene can move freely during capturing process. Camera response function (CRF) estimation is also a important step in HDR imaging pipeline. There are several methods available in

the literature to estimate CRF. They [**?**], extimate CRF by simply using a gamma function. In this approach [**?**], CRF is estimate by the set of exposure values using a non-parametric model. However, in [**?**], a low order polynomial method is used for CRF. Inspired by [**?**], we are also using a array of three cameras to enhance the dynamic range of the scene at higher spatial resolution.

2.2 Alignment of the images

The multi-exposure images captured by our imaging system which is shown in Figure **??** are required to be registered to generate a HDR image. To avoid any unwanted artifacts in the final HDR image the registration process is expected to be robust and therefore the optical flow based registration method presented in [**?**], is applied. Figure **??** shows three different exposure images which are captured using our imaging system. Flow vectors are estimated between every image. Since there is a variation in the exposure among these images, the underlying assumption of constant brightness for optical flow estimation cannot be fulfilled. Thus these images are first photometrically registered using a histogram based intensity mapping function (IMF) [**?**]. Photometric registration images are shown in Figure **??**. In Figure **??**, it is demonstrated that high-exposure image photometrically registered on mid-exposure image. Similarly, low-exposure image photometrically registered on mid-exposure image is shown in Figure **??**. After the photometric registration process, flow vectors between each of the images are calculated. Two corner images, i.e., high and low-exposure images are registered on the middle image. The residual between high and mid-exposure images before and after the registration is shown in Figure **??**. Similarly, the residual between low and mid-exposure images before and after the registration is shown in Figure **??**. These three registered images are then used to compute the irradiance.



(a) (b) (c)

Figure 2.1: Three different exposure images captured by our imaging system. (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image.

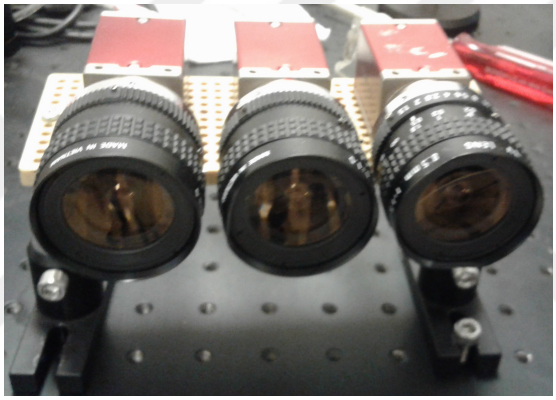
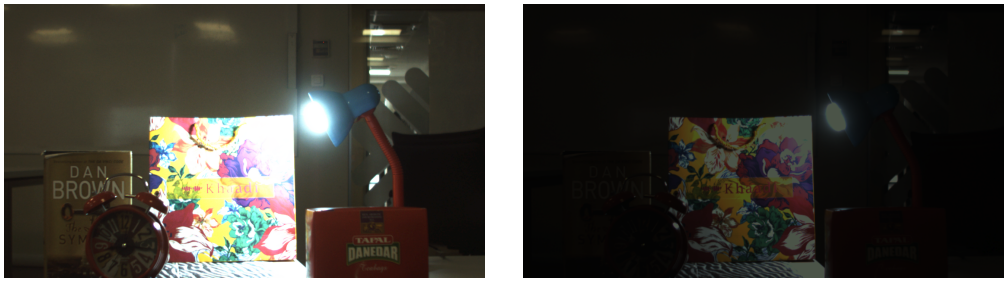


Figure 2.2: Hardware setup of the imaging system.

2.3 Radiance map estimation

To recover the full dynamic range of a scene, one can capture a series of photographs with varying exposure time. This is called multi-exposure technique. This technique allows us to enhance the dynamic range of the system and to reduce the noise in the final HDR image [**?**]. The goal of this method is to record all the levels of light intensity that are present in the scene. The luminance range of each photograph depends on the exposure time as the radiance values of the scene remain constant. To estimate the original values of the scene using these differently exposed images, response function of the imaging system needs to be known. Pixel values of the image do not represent the true measurements of the relative scene radiance because these pixels values are mapped by response function of the system even though most digital cameras produce a voltage that is proportional to the scene irradiance. However, these digital cameras apply some



(a) (b)

Figure 2.3: Photometric registration. (a) High-exposure image on mid-exposure image. (b) Low-exposure image on mid-exposure image.



(a) (b)

Figure 2.4: Residual between high-exposure and mid-exposure images. (a) Before registration. (b) After registration.

non-linear operations during the processing of the image such as gamma correction, analog to digital conversion, sensor noise etc. The effect of non-linearity mapping can be significantly found at the saturation point, where all the pixel values are close to the maximum value, and in the noise portion of the camera response curve. After computing a response function, pixel values from all the photographs can be used to construct an accurate radiance map of the scene.

In this experiment, three cameras of "Allied Vision" with "Ricoh" 8.5mm lens are used whose spatial resolution is 1292x734. Experimental setup is shown in Figure **??**. Multi-exposure images are captured by each camera. These input images are used to compute the camera response function and then construct the radiance maps for individual cameras. The intensity values of a pixel in a digital image can be defined by the following equation [**?**]:

$$Z(x, y) = f(E(x, y)T(x, y)), \tag{2.1}$$



(a) (b)

Figure 2.5: Residual between low-exposure and mid-exposure images. (a) Before registration. (b) After registration.

where (x, y) is the pixel position, $Z(x, y)$ is the value of pixel intensity, $E(x, y)$ is the scene irradiance, $T(x, y)$ is the camera exposure time and $f\cdot$ is the camera response function. In this case, $T(x, y)$ is typically constant for all pixel positions of an image. However, as we will see in the later chapter, for the plenoptic system that we have, $T(x, y)$ will have different values for different (x, y) positions. Typically, the exposure time is known and the pixel intensities are measured; the camera response function and the scene irradiance are calculated from recorded images. There are well-known and established techniques to measure the response function of a camera [**?**].

Once the response function of each camera is recovered, these are used to estimate the radiance map of each camera, assuming the exposure time is known. For recovering high dynamic range radiance map, it is more reliable to estimate it from properly exposed pixels rather than using over-exposed or under-exposed regions. Thus, a weighted sum of the estimate is taken to get the final irradiance as described in [**?**]. The weighted sum is described by a simple triangular function is shown in Figure **??**.

2.4 Tone-mapping

A high dynamic range image cannot be seen on typical display devices due to limited dynamic range. The real world scenes have larger dynamic range. However,

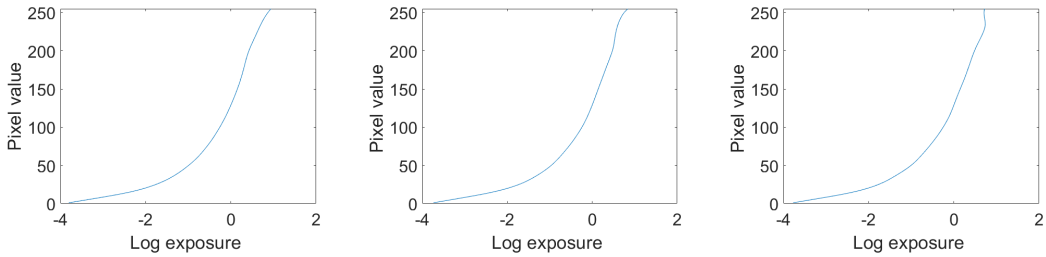


Figure 2.6: Response function of each camera is determined by the technique described in [**?**].

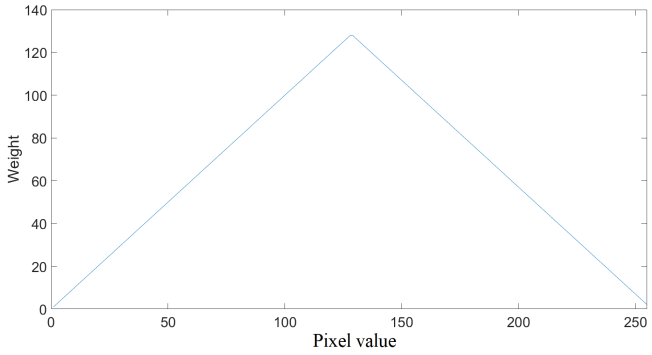


Figure 2.7: The weight function for calculating the radiance maps.

the display devices that are available in the market cannot produce such wide dynamic range as real world scenes. Conventional display devices can only represent the images which has dynamic range of less than 100 to 1 [**?**]. In order to have a realistic representation of HDR images using a medium that has limited dynamic range such as CRT monitor, LCD, printed paper and projector, it is necessary to apply compression techniques known as tone-mapping. The purpose of tone-mapping techniques is to compress the dynamic range of image while preserving the details in order to display on typical devices. Tone-mapping techniques were formally introduced in the early 1990s [**?**] in order to match the HDR content with the display devices. They are generally divided into two broad categories, one is global operator and the other one is local operator. Global operators apply the same compression curve on the whole image regardless of the pixel brightness. Such an operator takes less computational time and can run in real time as well. The drawback of the global tone-mapping technique is that it may not preserve the details if the dynamic range of the image is too high [**?**]. On the

other hand, local tone-mapping operator applies the compression function based on each pixel luminance value as well as neighbouring pixels luminance values. Due to this, it preserves the details even the dynamic range of the image is high. Local tone-mapping techniques usually require high processing time.

In this work, we are using Reinhard tone-mapping operator [**?**] because it has been shown that this tone-mapping technique gives the best result in terms of visually and maintaining the details of the image [**?**]. The tone-mapping problem was first faced by the photographers. Their goal is to print a captured scene as close to the real as possible. For this Ansel Adams [**?**], presented a zone system solution. This system uses eleven print zones, ranging from pure black (Zone 0) to pure white (Zone X). The key indicates weather the scene is light, normal or dark. To get a good final print, photographer first perceives a luminance reading of a surface as a middle gray. This will be mapped to zone V, which corresponds to the 18 percent reflectance of the print. For the scenes which has really large luminance values the middle gray will be one of the darker regions, whereas the scenes which has low luminance values the middle gray will be one of the lighter regions [**?**].

A similar approach to the one mentioned above was taken by Reinhard [**?**] to developed tone-mapping operator for digital images. In this method, global tone-mapping operator is defined which is computationally efficient and simple, as well as local tone-mapping operator is determined for particularly high dynamic range scenes which is complex and computationally intensive operator. The basic idea of the global tone-mapping technique is to map all pixel luminances of an image to the range [0 1]. This technique applies the compression curve on the whole image which is invariant to the pixel position. For high dynamic range images, this technique preserves details in low contrast areas while compressing high luminances values to a displayable range. For very high dynamic range images, this operator fails to preserve details in the image.

HDR images with details in high luminance regions require additional step to preserve the details in that region. For this dodging-and-burning technique, has proven useful. This technique basically scales up selected dark region and scales

down selected light region to preserve the details [?]. The same technique was used by Reinhard [?] for particular HDR image to avoid loss of details in regions of high luminance. This operator enhances the pixels contrast relative to the surrounding areas [?].

2.5 De-ghosting

Camera shake or any object movement in the scene can create ghost like artifacts in the resultant HDR image. Over the past few years, researchers proposed variety of algorithms for HDR imaging which incorporates de-ghosting features. One of the methods which is described by [?], detected ghost regions based on pixel order relation. In order to generate ghost free HDR image pixel location in the detected region should be known. For this, two sets of exposures are created for LDR images. One set contains the moving objects for the current pixel location while the other does not contain the moving objects for the current pixel. These two sets of exposures for the current pixel are created using the quasi-continuous histogram (QCH)[?] method. This method is used to detect the moving objects by estimating the pixel values. For the creation of HDR image, information is carried from the set of images which does not contain moivng pixels. The other method which is defined by Reinhard et al.[?], detects the ghost region by computing the weighted variance at each pixel location and defines a threshold. If the local variance value is above the threshold that region should be considered as ghosting region. This method gives the best result where the background and the moving object of the scene is different in terms of contrast and color similarity. In the other method which is presented in [?], the inconsistent pixels are first detected by measuring local entropy difference. This method requires the choice of a reference image. The selection of the reference image can be done in two ways, either by minimizing the number of saturated pixels or manually selected by the user. This method described that if the two pixels have the same irradiance level then it should follow a linear relationship. If any pixel does not follow this relationship is considered as inconsistent pixel. This method works well to apply on rectangular image patches instead of the pixels.

In this study, the state of the art algorithm is used which is proposed by [**?**] in generating ghost free HDR image. Their approach is based on patch-based energy-minimization formulation. The energy function of this method performs the alignment of LDR images and reconstruction of HDR image simultaneously. This algorithm allows to handle significant camera movement and scene motion. In this method, it is assumed that H should be similar to the reference LDR image where the pixels are well exposed. In the case of poorly exposed pixels, information of that area should be taken from other exposures using Multisource Bidirectional Similarity (MBDS) which is a modified form of BDS introduced by Simakov et al. [**?**]. The final HDR image is aligned to the reference image and most of the information is taken from the reference image. Auxiliary images are computed using the search and vote scheme instead of solving for output HDR image directly [**?**]. Intermediate HDR image is obtained by merging these auxiliary images. Initialization of these images is done iteratively and they are continously updated until the convergence [**?**]. In this work, we utilized this algorithm just to align the LDR images to avoid the occlusion artifacts in the final HDR image. To estimate the radiance map, the same process has utilized as described above.

2.6 Dynamic range

Dynamic range (DR) is one of the most important characteristics of the digital camera. It is defined as the ratio between the lightest and the darkest point of the scene that camera can capture. If the scene dynamic range exceeds the camera dynamic range, the brightest part becomes saturated and the darkest part becomes noisy. The camera which has larger dynamic range can capture details in the lightest and darkest parts of the scene. In this work, we quantify the enhancement of the dynamic range which is achieved using our imaging setup.

2.7 Quantification of the dynamic range of camera array

The setup used in this work consists of three cameras is shown in Figure **??**. Dynamic range of this system can be measured by capturing the same scene with different exposure time. The exposure time value of each of the camera is set to be different from each other. Images taken by the system with different exposure time are shown in Figure **??**. The intensity value of a pixel is equal to the product of scene's irradiance with camera's exposure time and response function. Suppose that the response function of the camera is linear so mathematically, it can be defined as follows:

$$Z(x, y) = E(x, y)T(x, y), \tag{2.2}$$

where $Z(x, y)$ is the pixel value, $E(x, y)$ is the irradiance value, and $T(x, y)$ is the exposure time at pixel location (x,y). The irradiance of a scene can be calculated from equation **??**, which is the ratio between the intensity value and the exposure time of camera, given by

$$E(x, y) = \frac{Z(x, y)}{T(x, y)}. \tag{2.3}$$

The dynamic range (DR) of this system can be calculated by taking the difference of irradiances between the last and first camera in logarithmic scale which is described as

$$\begin{aligned} DR &= \log\left(\frac{255}{t_3}\right) - \log\left(\frac{1}{t_1}\right) \\ &= \log(255) + \log(t_1) - \log(t_3) \\ &= \log(255) + \log\frac{t_1}{t_3}, \end{aligned} \tag{2.4}$$

where t_1 and t_3 are the exposure values of the cameras, t_1 is the high exposure value of the first camera and t_3 is the low exposure value of the last camera. Equation **??** described that the enhancement of the dynamic range by our setup should be the ratio between the exposure time value of the first and the last camera. In Figure **??**, the graph between irradiance and intensity is shown. In this graph, irradiance of each of the image is shown after being calculated by employing a weighted average approach described in [**?**]. The graph shows that

the enhancement in dynamic range is around 25 dB. The overall dynamic range of this setup is around 52 dB. The dynamic range of this setup can be increased by adding more cameras in the system. Similarly, the dynamic range of a single camera is also shown in Figure **??**, which is around 30 dB. The dynamic range of the camera is limited by the camera noise. Practically, useful dynamic range of a camera is defined between the point before the saturation to the minimum threshold after which the sensor noise affects the image quality. At the brightest side, the considerable pixel level should be below 98% of the saturation level (for the 8-bit system it is 250, which has a maximum value of 255) [**?**]. At the darkest side, signal to noise ratio (SNR) is calculated by capturing multiple white scenes with different exposure times, is shown in Figure **??**. SNR of each of the images is computed as described below,

$$SNR = 20 \log_{10} \left(\frac{Z_{peak}}{Z_{noise}} \right). \tag{2.5}$$

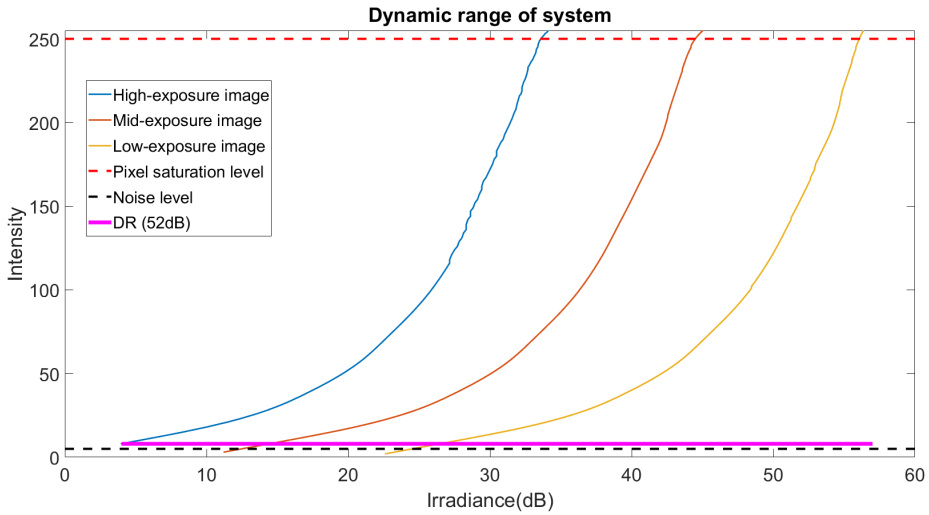


Figure 2.8: Dynamic range of the camera array.

White images are not uniform through out the scene because of the lighting condition. To make the calculation robust, SNR of each image is computed patch wise by dividing the image into equal number of patches. Z_{peak} is calculated by taking the mean of every patch of the image and finally computing the average of all means of the patches. Similarly, Z_{noise} is calculated by taking the standard deviation of each patch of the image and finally computing the average of all

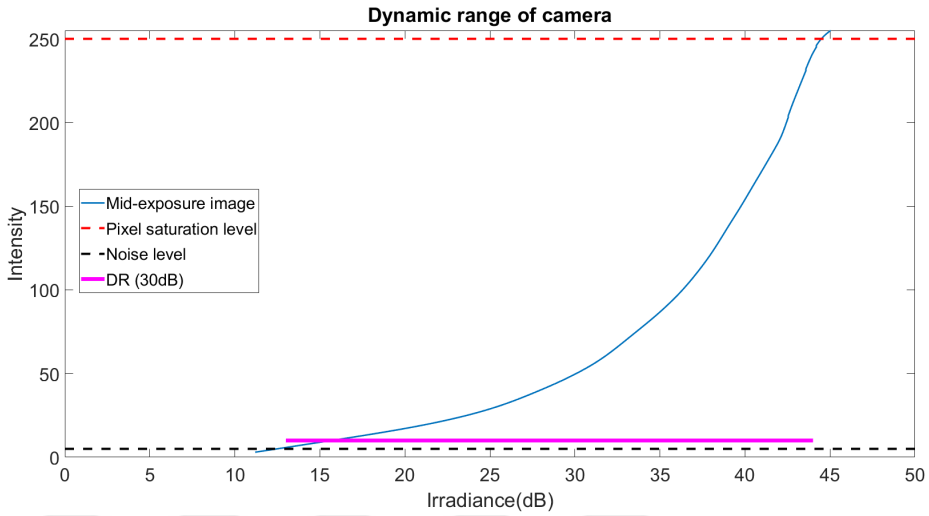


Figure 2.9: Dynamic range of a single camera.

standard deviations of the patches. The graph between SNR and intensity levels is shown in Figure **??**. This graph shows that the lowest SNR is around 25 dB and at this value, the intensity level is 5. After considering these intensity levels, the useful dynamic range of a camera is still around 30 dB which is shown in Figure **??**.

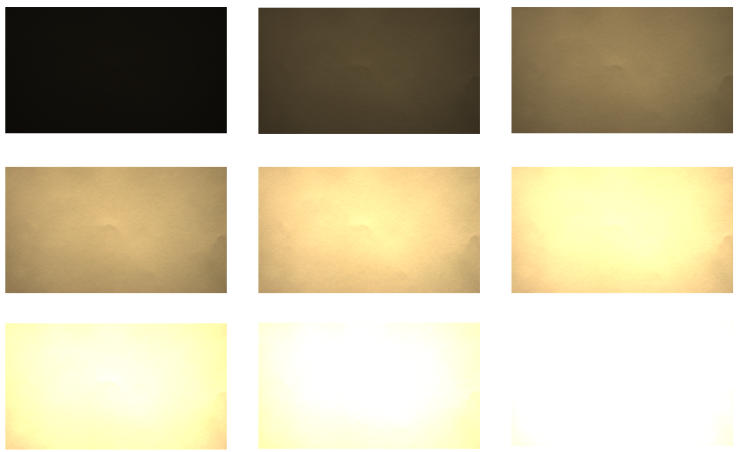


Figure 2.10: White images.

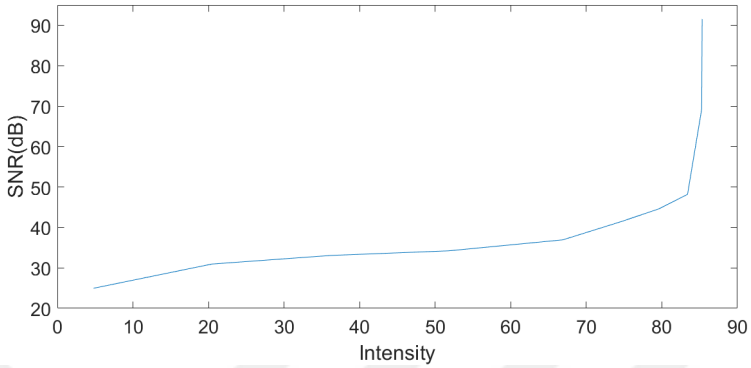
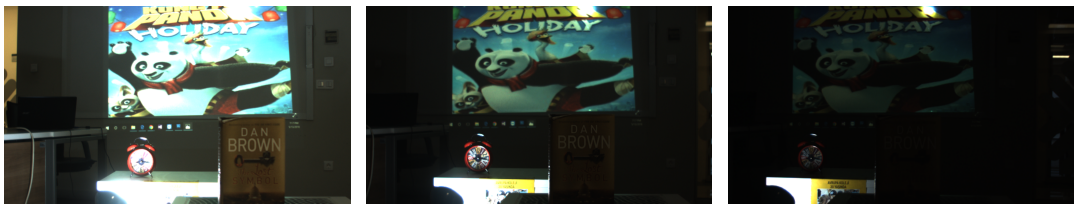


Figure 2.11: Noise measurement.

2.8 Experimental results

In this section, experimental results are presented for enhancing the dynamic range using camera array. In this study, we successfully increase dynamic range of the captured scene. Three differently exposed images are shown in Figure **??**, **??**, **??**, which is captured by the experimental setup shown in Figure **??**. HDR image with tone-mapping is shown in Figure **??**, which recovers all the luminance range of the scene.

We have tested our setup under different scene conditions. For instance, in Figure **??**, objects are really close to the setup. In this case, all the information of the scene is not available to all the cameras because of the large baseline between them. Due to this, there can be occlusion artifacts in the final HDR image. In registered images as shown in Figure **??**, there are occlusion artifacts because of large baseline between the camera as stated above and another reason is large variation of exposure between the images. To remove these occlusion artifacts, we utilized state of the art de-ghosting algorithm [**?**] on registered images. In Figure **??**, it is demonstrates perfectly aligned images to reference image after the de-ghosting operation. Finally, in Figure **??**, different exposure values with perfectly aligned images are shown. The final HDR image with tone-mapping is shown in Figure **??**.



(a) (b) (c)

Figure 2.12: (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image.

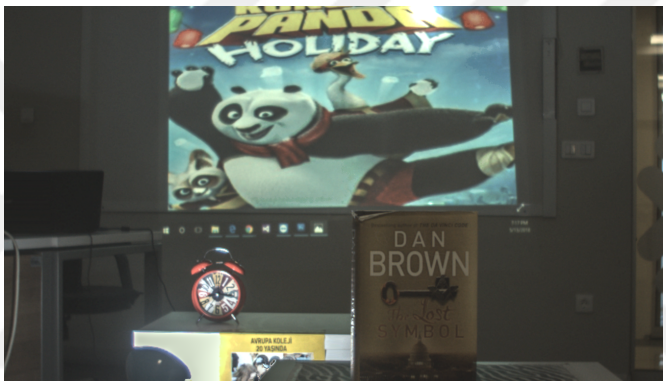


Figure 2.13: HDR image with tone-mapping.



(a) (b) (c)

Figure 2.14: (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image.



(a) (b)

Figure 2.15: Demonstration of occluded artifacts after registration in (a) High-exposure image. (b) Low-exposure image.



(a) (b)

Figure 2.16: Demonstration of images after de-ghosting operation (a) High-exposure image. (b) Low-exposure image.



(a) (b) (c)

Figure 2.17: (a) High-exposure image. (b) Mid-exposure image. (c) Low-exposure image.



Figure 2.18: HDR image with tone-mapping.

Chapter 3

High dynamic range image using plenoptic camera

This chapter addresses the issues faced in HDR imaging using camera array, i.e, occlusion artifacts and the need for multiple shots to create a single HDR image. As mentioned earlier, the methods used for generating HDR images require sequential capture for multiple exposure images and are thus susceptible to ghosting artifacts, caused mainly due to camera shake and motion of the objects. However, by using plenoptic camera these above artifacts can be avoided because of its small baseline, which is typically in the order of sub-pixels, therefore, adjacent sub-aperture images suffer less from occlusion. Plenoptic camera also allows us to acquire multiple images within a single shot.

In this work, first generation Lytro camera has been used with slight modification in the optical design to capture the light field data. Light field data capture by Lytro camera is affected by vignetting which results in darker sub-aperture images at the corners. As these sub-aperture images are few in numbers as compared to the well exposed sub-aperture images, they only capture a limited dynamic range. To maximize the effect of the vignetting, optical mask has introduced at the main lens of the Lytro camera. The mask is designed in a way that at corners of the light field grid sufficient number of low-exposure sub-aperture

images can be extracted while preserving an adequate number of well exposure sub-aperture images in the middle region. This allows the extension of dynamic range. These sub-aperture images are then registered and combined to generate HDR light field.

3.1 Related work

Brightness in a natural scene varies due to a variety of reasons and hence typical low dynamic range (LDR) digital camera cannot capture an entire irradiance of the scene. Several techniques have been proposed that overcome dynamic range limitation in digital photography. Some of the earlier attempts include [**?**], [**?**], [**?**] which require sequential capture of multi-exposure LDR images and then combine them to generate an HDR image. The sequential acquisition limits these methods to static scenes only and the application of these methods to dynamic scene require additional de-ghosting steps.

A variety of algorithms for HDR imaging which incorporate de-ghosting features, such as [**?**], [**?**] are also present in literature. In [**?**], the inconsistent pixels are first detected by measuring local entropy difference and are then successfully removed before merging. To avoid ghosting artifacts in [**?**], a binary ghost map is proposed to detect false pixels using zero normalized cross correlation. Most of the existing de-ghosting algorithms cannot successfully handle large scale motions. Apart from object's motion these artifacts can also be caused by camera shake and its removal would require blind-deblurring which is yet another challenging problem.

A hardware solution for generating HDR images for a dynamic scene has also been proposed in the past by using the well-known Stanford's custom built camera array [**?**] to capture multi-exposure images at once. One of the downsides of these camera arrays is that they are very bulky and are difficult to move. Apart from size limitation, baseline between the adjacent cameras is large enough to get easily affected by the occlusion phenomenon and yet again results in ghosts like artifacts

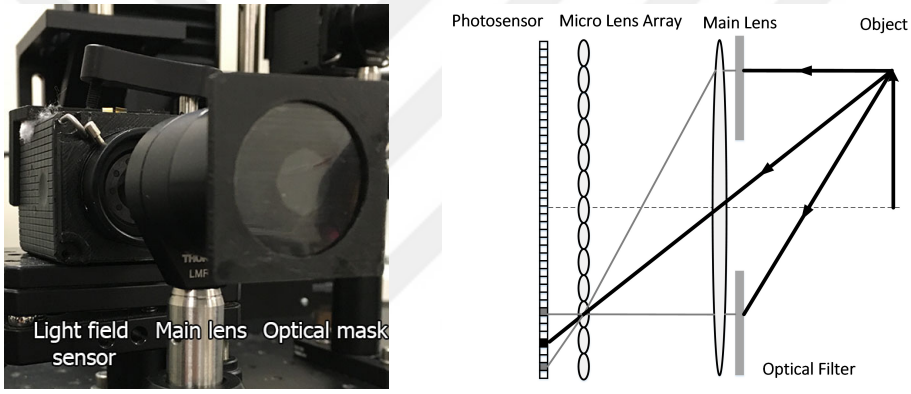
after the fusion process. The commercially available handheld light field camera can tackle the issue of bulky volume and large baselines of the camera array by capturing multiple sub-aperture images in a single shot. In [**?**], a focused plenoptic camera was utilized by introducing an optical mask that individually controls the exposure of each sub-aperture image. In this study, first generation light field camera was used to capture multi-exposure images, however, our mask design is inspired to control the vignetting effect. In addition, our algorithm generates final HDR image that estimates the exposure time of each pixel individually as opposed to the uniform exposure for the entire image.

3.2 Optical design

Our optical design is a slight modification in the optical path of an off-the-shelf MLA based light field camera designed on the model proposed in [**?**]. This camera has an optical configuration in which a microlens array is placed at the focal point of the main lens. Therefore, light rays originating from a single point in a scene passing through different positions of the main lens converge at a microlens as illustrated in Figure **??**. These rays are then captured by a set of pixels behind a microlens, and thus by combining the pixels behind each microlens, corresponding to the set of rays passing through a smaller portion of the main lens forms sub-aperture image of the entire scene. The sub-aperture images formed by the rays passing through the corner regions of the main lens are affected by vignetting and have low intensities.

In this study, a circular mask has been introduced at the lens to amplify this vignetting phenomenon not only to further reduce exposure of the sub-aperture images affected by vignetting but also to increase the number of affected sub-aperture images. Our optical filter, shown in Figure **??**, is designed in a way that outer portion of the mask is partially opaque while the inner portion remains completely transparent. The opacity factors determine the exposure of affected sub-aperture images whereas the size of opaque area of the mask defines the number of sub-aperture images with low-exposure.

First generation Lytro camera comes with a compound lens, microlens array and a photosensor sealed together in a metallic casing [**?**]. In order to place optical mask at the front of the lens aperture, the compound lens system needs to be opened for mask insertion and reassembled which cannot be achieved. We, therefore, dismantled the light field camera and replaced the compound lens system by an achromatic 60 mm single lens from Thor-labs. For the optical mask, we have printed our design on transparency with a laser printer at a print resolution of 300 DPI.



(a) (b)

Figure 3.1: (a) The optical setup. (b) Schematic of the optical design.

3.3 Light field pre-processing

The Lytro camera provides raw lenslet images, which need to be decoded in order to extract the perspective images of the light field. A raw lenslet image is shown in Figure **??**. The spatial resolution of a raw image is 3280x3280 pixels. A variety of tools such as [**?**], [**?**], [**?**] are available online for decoding purpose, but they all require white images to calibrate raw light-field. A set of white images is provided with a Lytro camera but as it is dismantled those whites images cannot be used anymore, because a minor misalignment between MLA and the sensor may affect the calibration process greatly. We have, therefore, captured a set of new white images using the modified version of light field camera and followed the remaining process proposed in [**?**] to decode a raw image into a regular grid

of 9 x 9 sub-apertures images each with a spatial resolution of 380 x 380 pixels.



Figure 3.2: Lenslet image of a Lytro camera.

3.4 High dynamic range imaging

3.4.1 Geometric registration of light field data

The multi-exposure sub-aperture images vary in their perspectives, the point in the scene may be projected at different locations in these perspectives based on their depth. Therefore, it is necessary to register these perspectives before further processing. To avoid any unwanted artifacts in the final merged image the registration process is expected to be robust and therefore an effective optical flow based registration method [**?**] is employed. Figure **??**, shows a decoded light field captured by the proposed optical setup with the optical mask. Flow vectors are estimated between two corner images of a middle row of the light field grid and two corner images of a middle column of the light field grid. The zoomed in version of corner images are also shown in Figure **??**. Since there is a variation in the exposure among these images the underlying assumption of constant brightness for optical flow estimation cannot be fulfilled. Thus these images are first photometrically registered using a histogram based intensity mapping function (IMF) [**?**]. Photometric registration process of the corner images of the middle row of captured light field is shown in Figure **??**. Similarly in Figure **??**, photometric

registration process of the corner images of the middle column of captured light field is shown. After the photometric registration process, flow vectors between these images is calculated. The residual between corner images of a middle row is shown in Figure **??**. Similarly in Figure **??**, residual between corner images of a middle column is shown. These flow vectors which are estimated between corner images are then interpolated to estimate flow vectors for all the sub-aperutre images. These sub-aperture images are then registered on the middle sub-aperture image. The illustration of geometric registration of sub-aperture images of light field is shown in Figure **??**.

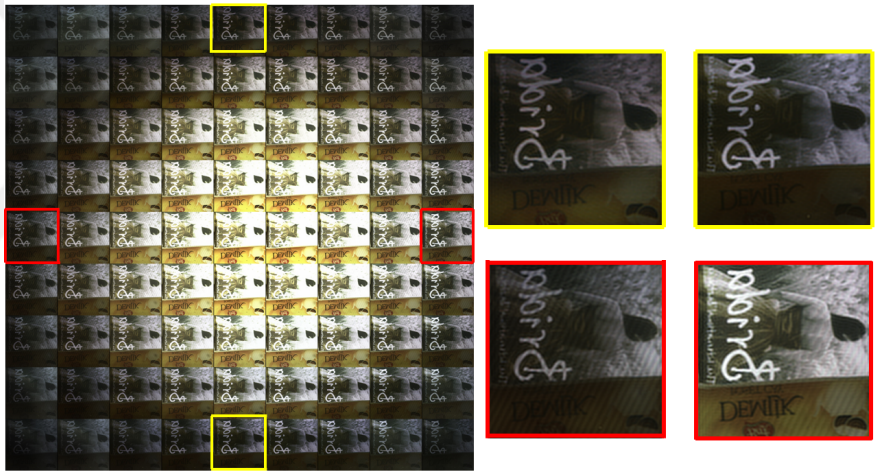


Figure 3.3: Captured light field with zoom in version of extreme corner middle row images and middle column images.

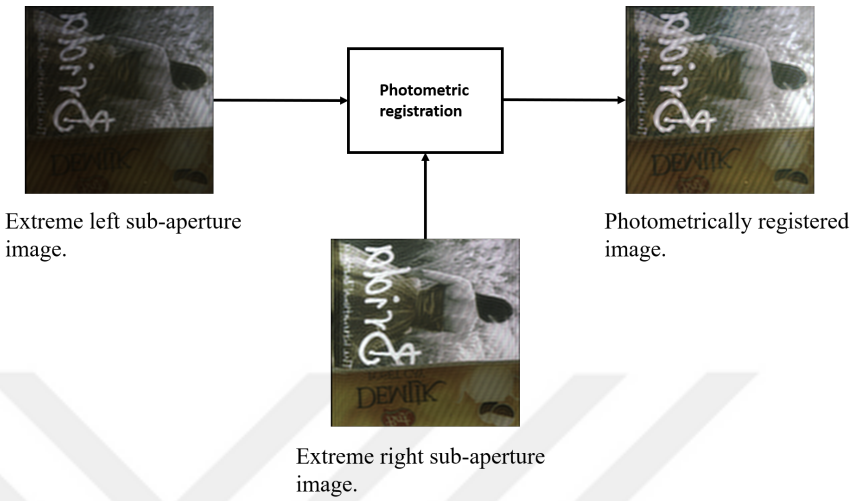


Figure 3.4: Photometric registration of extreme corner images of a middle row.

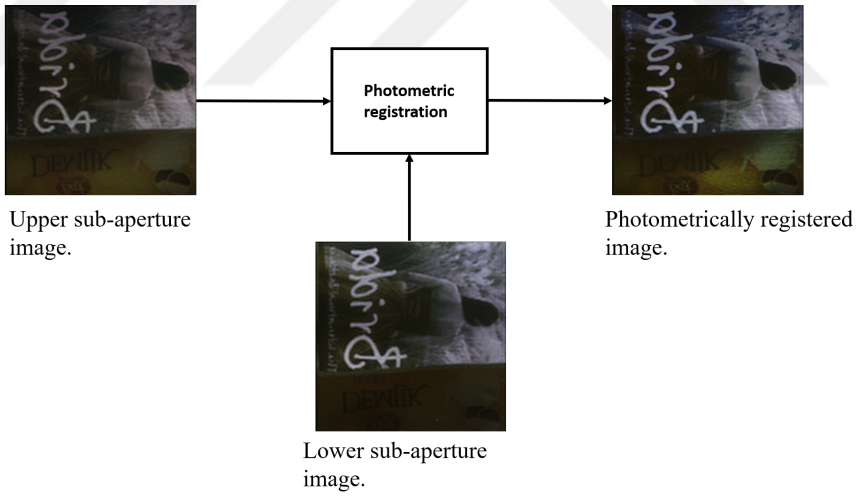
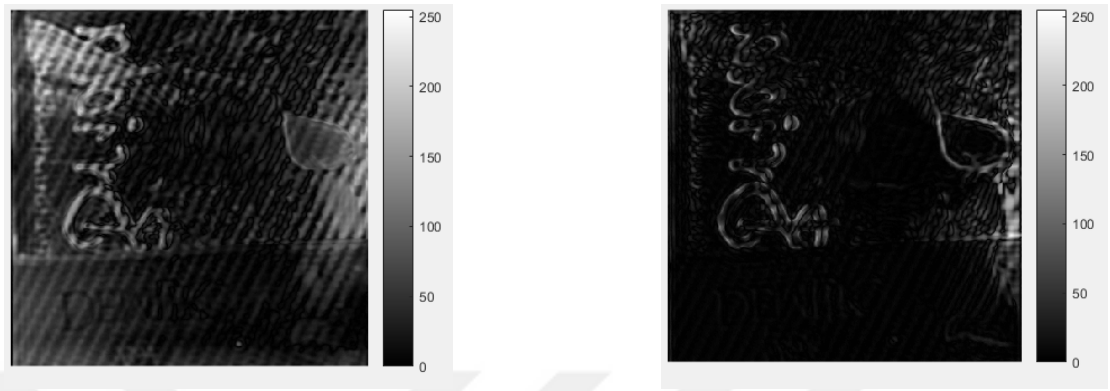
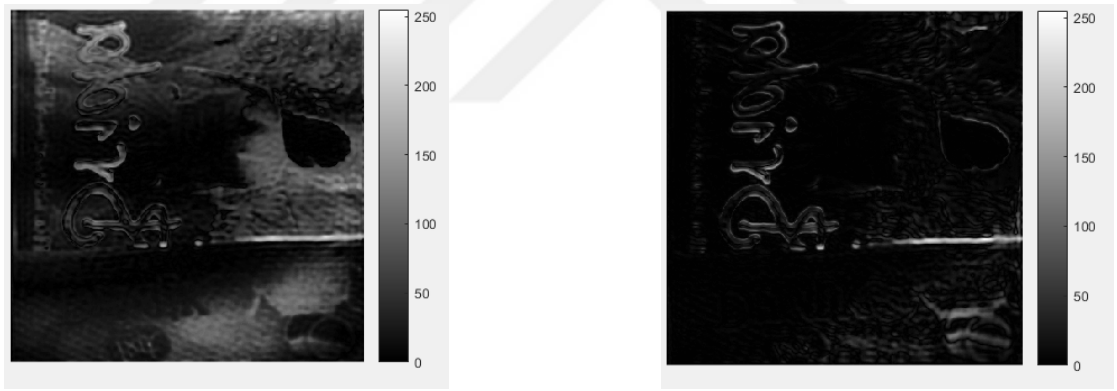


Figure 3.5: Photometric registration of extreme corner images of a middle column.



(a) (b)

Figure 3.6: Residual between extreme corner images of a middle row. (a) Before registration. (b) After registration.



(a) (b)

Figure 3.7: Residual between extreme corner images of a middle column. (a) Before registration. (b) After registration.

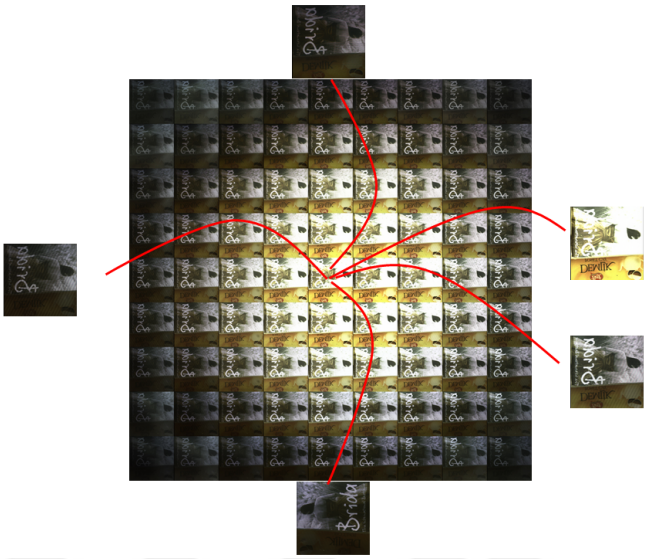


Figure 3.8: Geometric registration of sub-aperture images of light field.

3.4.2 Irradiance estimation

Space varying exposure of sub-aperture images necessitates the need for pixel by pixel exposure estimation. The intensity value of a pixel in a digital image can be modeled as a product of scene's irradiance with camera's exposure time and response function as described in equation **??** [**?**].

From equation **??**, we have two unknowns the scene's irradiance and exposure time, by making an appropriate assumption about one of the two unknowns the other one can be estimated. We assume that for a white planar scene the irradiance remains uniform for a negligible amount of time (typical exposure time). For this, we utilize the white image captured during the calibration process to estimate the relative exposure values of each pixel. Assuming a linear camera response function in the un-saturated portion of the sensor response, the intensity value $Z_k(x, y)$ at a pixel location (x, y) in the kth sub-aperture image can be written as:

$$Z_k(x, y) = E(x, y)T_k(x, y), \tag{3.1}$$

where $E(x, y)$ is the irradiance and $T_k(x, y)$ is the exposure time.

When the image of a white uniform scene is taken, the measured intensity $Z_{k,white}(x, y)$, should give the relative exposure time of the pixels in the images. The white scene image is shown in Figure **??**. We take the mid sub-aperture image as the reference one. The relative exposure time of each of the pixel is computed as:

$$T_k(x, y) = T_{mid}(x, y) \left(\frac{Z_k(x, y)}{Z_{mid}} \right), \tag{3.2}$$

where Z_{mid} is the average intensity value of the mid sub-aperture image of the white scene. From equation **??**, the irradiance can be estimated as:

$$E(x, y) = \frac{Z_k(x, y)}{T_k(x, y)}, \tag{3.3}$$

using the kth sub-aperture image. We can estimate the irradiance of each sub-aperture image separately; however, the estimates that come from properly exposed pixels would be more reliable than the ones coming from over-exposed or under-exposed images. Thus, a weighted sum of the estimates is taken to get the final irradiance:

$$E(x, y) = \frac{1}{N} \sum_{k=1}^{N} \left(w(Z_k(x, y)) \left(\frac{Z_k(x, y)}{T_k(x, y)} \right) \right), \tag{3.4}$$

where $w\cdot$ is a triangular weight function which is described in equation **??**, giving more weight to pixels that come from mid-tones [**?**], and N normalizes the weights. The triangular weight function is shown in Figure **??**.

$$w(x) = \begin{cases} z - Z_{min} \; for \; z \leq \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z > \frac{1}{2}(Z_{min} + Z_{max}), \end{cases} \tag{3.5}$$

where Z_{min}, Z_{max} indicates the minimum and maximum possible pixel values. For an 8-bit representation [Z_{min},Z_{max}] = [0,255].

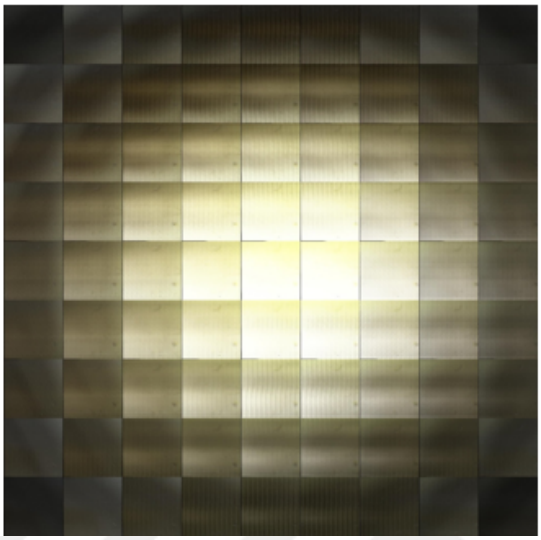


Figure 3.9: White planar scene.

3.5 Quantification of light field

The quantification of light field has been done using the same methodology which is used in camera array setup. The irradiance of each of the sub-aperture images are calculated by the same method which is described above. The graphs between the irradiance and the intensity of particular sub-aperture images which are highlighted in Figure **??**, are shown in Figure **??** and Figure **??**. The particular sub-aperture images is different with each other in terms of exposure values. The average dynamic range of each sub-aperture image is around 12 dB. The dynamic range of the entire light field is shown in Figure **??**, which is around 20 dB. It is clearly shows that the proposed optical setup enhanced the dynamic range of the light field sensor.

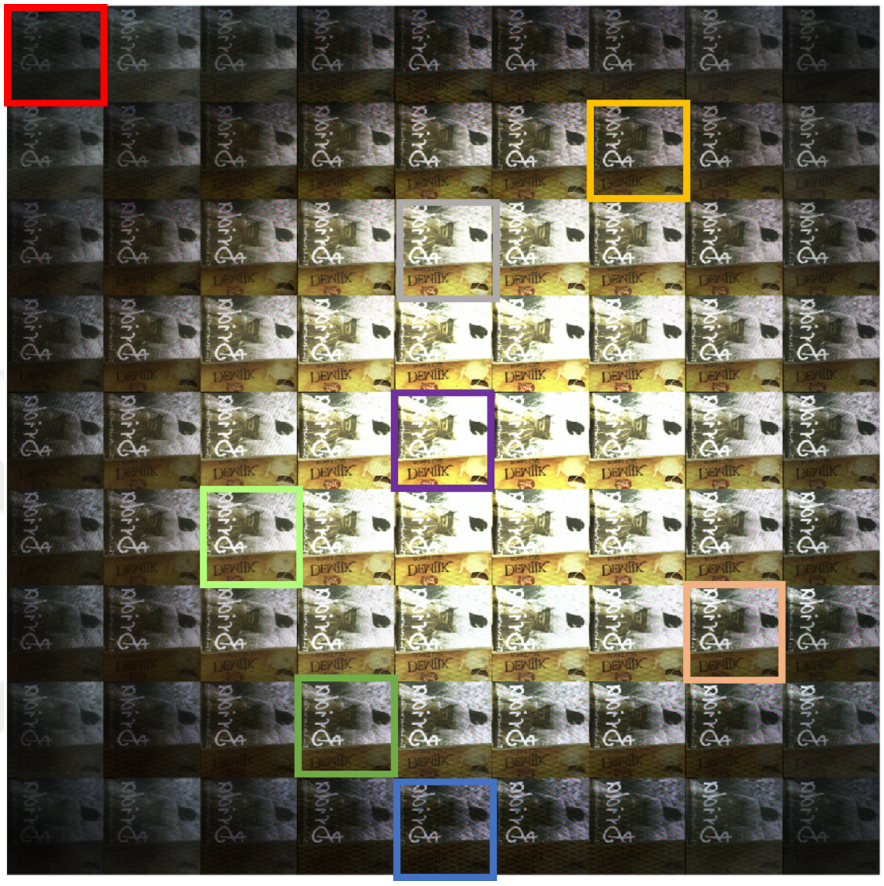


Figure 3.10: Captured light field.

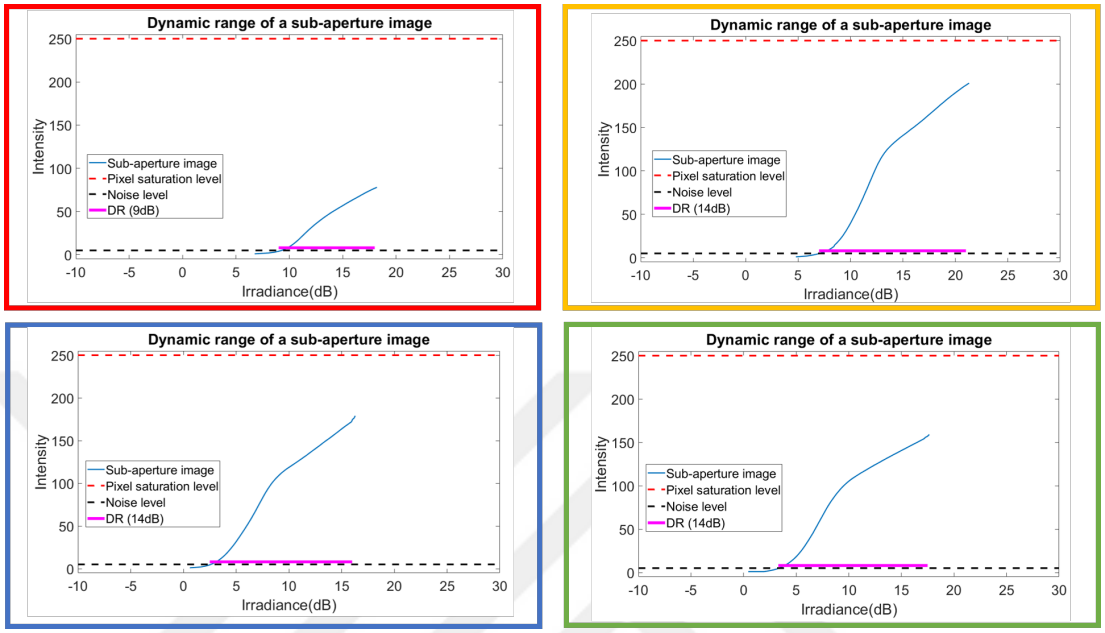


Figure 3.11: Dynamic range of sub-aperture images.

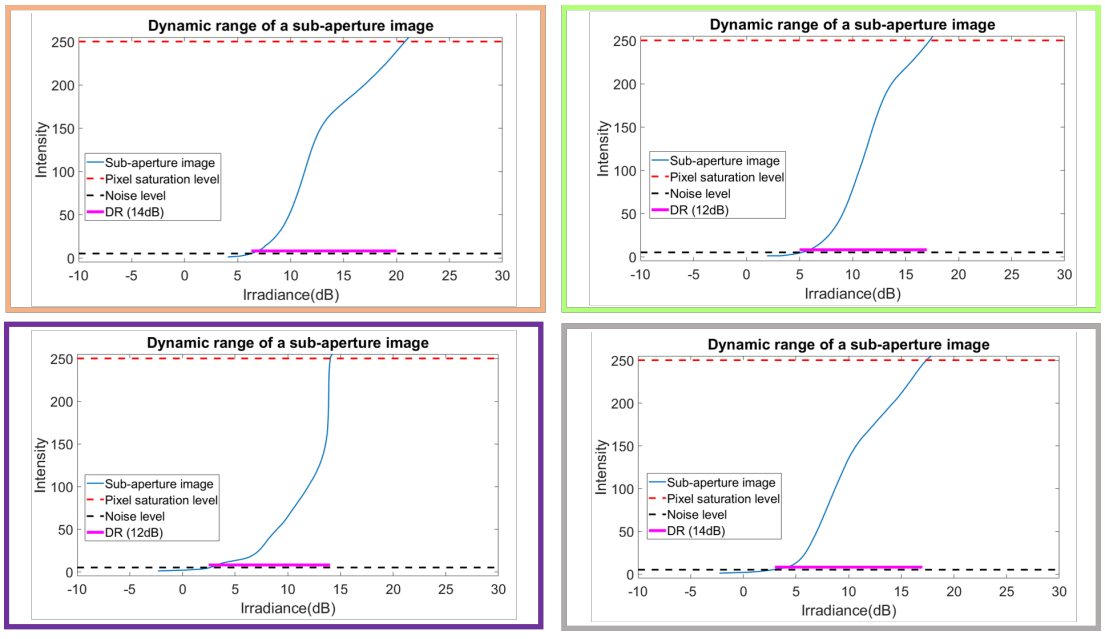


Figure 3.12: Dynamic range of sub-aperture images.

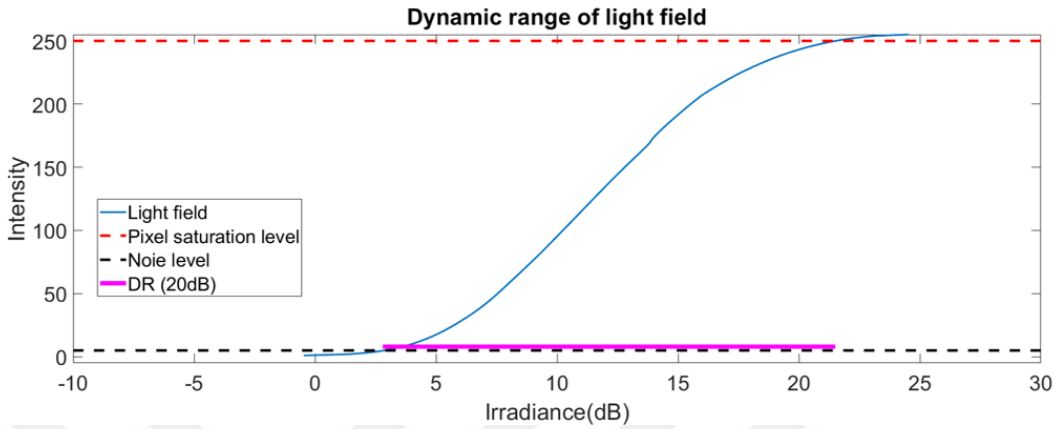


Figure 3.13: Dynamic range of light field.

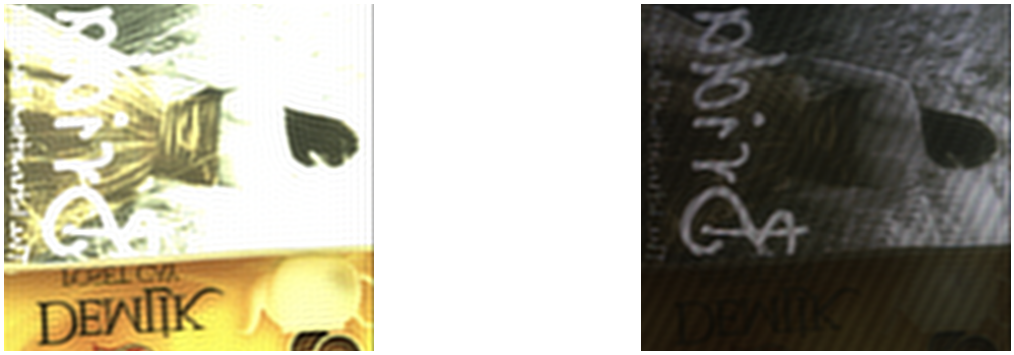
3.6 Experimental results

In this section, experimental results are presented to demonstrate the extended dynamic range of a typical LDR light field camera simply by introducing an optical mask at the lens. In Figure **??**, two HDR images with and without registration process is shown. The high-exposure and low-exposure images are shown in Figure **??**. Figure **??**, clearly shows blurriness in HDR image. On the other hand, registration step produced a sharp HDR image as shown in Figure **??**.

Two decoded light fields are presented in Figure **??**, light field captured by the modified light field camera with optical mask has a greater number of low-exposure sub-aperture images with significant exposure difference from high-exposure images as compared to light field captured without optical mask. In Figure **??**, we show two sub-aperture images from the light field given in Figure **??**, with the proposed optical setup with optical mask. One image is the middle sub-aperture, corresponding to the highest possible exposure value, the other image is a low-exposure sub-aperture image. The resulting HDR image is also shown in Figure **??**, clearly demonstrating the extension of the dynamic range. On the other hand, in Figure **??**, similarly, we show two sub-aperture images

from the light field given in Figure **??**, without optical mask. One is the middle sub-aperture image which corresponding the high-exposure image and the other one is the low-exposure sub-aperture image. The resulting HDR image seems that higher exposure values are more dominating, enabling to capture less dynamic range. This is also representing by the histogram of HDR images with and without optical filter which is shown in Figure **??**. In Figure **??**, histogram of HDR image is shown without optical filter in which most of the radiance values are in saturated side. Similarly, in Figure **??**, histogram of HDR image is shown with optical filter in which radiance values are in the mid region.

We have tested our camera setup under different lightning conditions for example in Figure **??**, we present results for a data set captured with large variation in the scene's lightning and demonstrated that our algorithm can recover the dark and saturation regions equally well. Since our algorithm requires us to dismantle the light field camera our aim was to make sure that the dynamic range is extended without affecting the valuable features offered by the camera. Therefore, in Figure **??**, we have demonstrated that with our extended dynamic range light field we can perform post-capture refocusing without introducing any artifacts using typical shift and sum algorithm. Figure **??**, demonstrates the light field image with modified camera and HDR light field.



(a) (b)

Figure 3.14: (a) High-exposure image. (b) Low-exposure image.



(a) (b)

Figure 3.15: (a) Non-registered HDR image. (b) Registered HDR image.



(a) (b)

Figure 3.16: Captured light fields. (a) With optical filter. (b) Without optical filter.



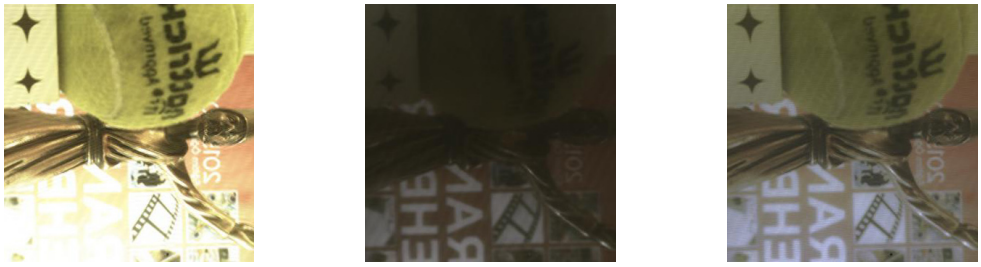
(a) (b) (c)

Figure 3.17: (a) High-exposure image. (b) Low-exposure image. (c) HDR image.



(a) (b) (c)

Figure 3.18: (a) High-exposure image. (b) Low-exposure image. (c) HDR image.



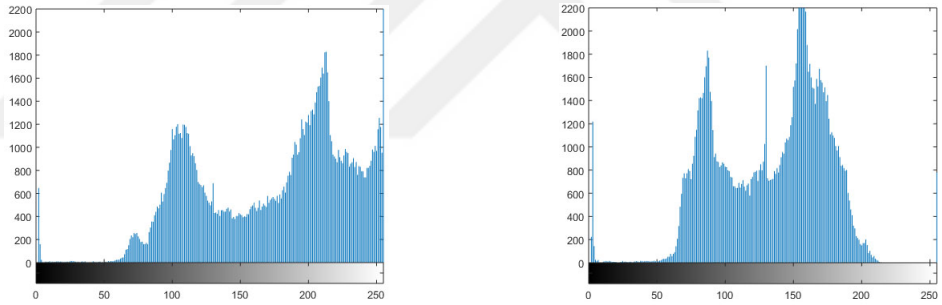
(a) (b) (c)

Figure 3.19: (a) High-exposure image. (b) Low-exposure image. (c) HDR image.



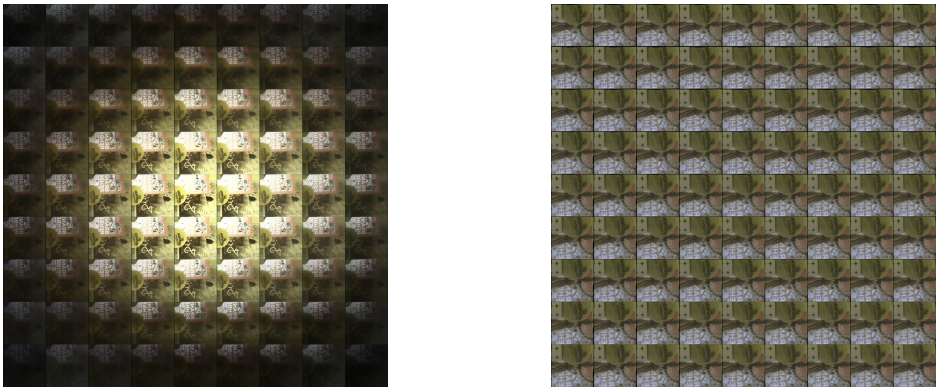
(a) (b)

Figure 3.20: (a) HDR image without optical filter.(b) HDR image with optical filter.



(a) (b)

Figure 3.21: Comparison between histograms of HDR images. (a) Without optical filter. (b) With optical filter.



(a) (b)

Figure 3.22: (a) Captured light field. (b) HDR light field.

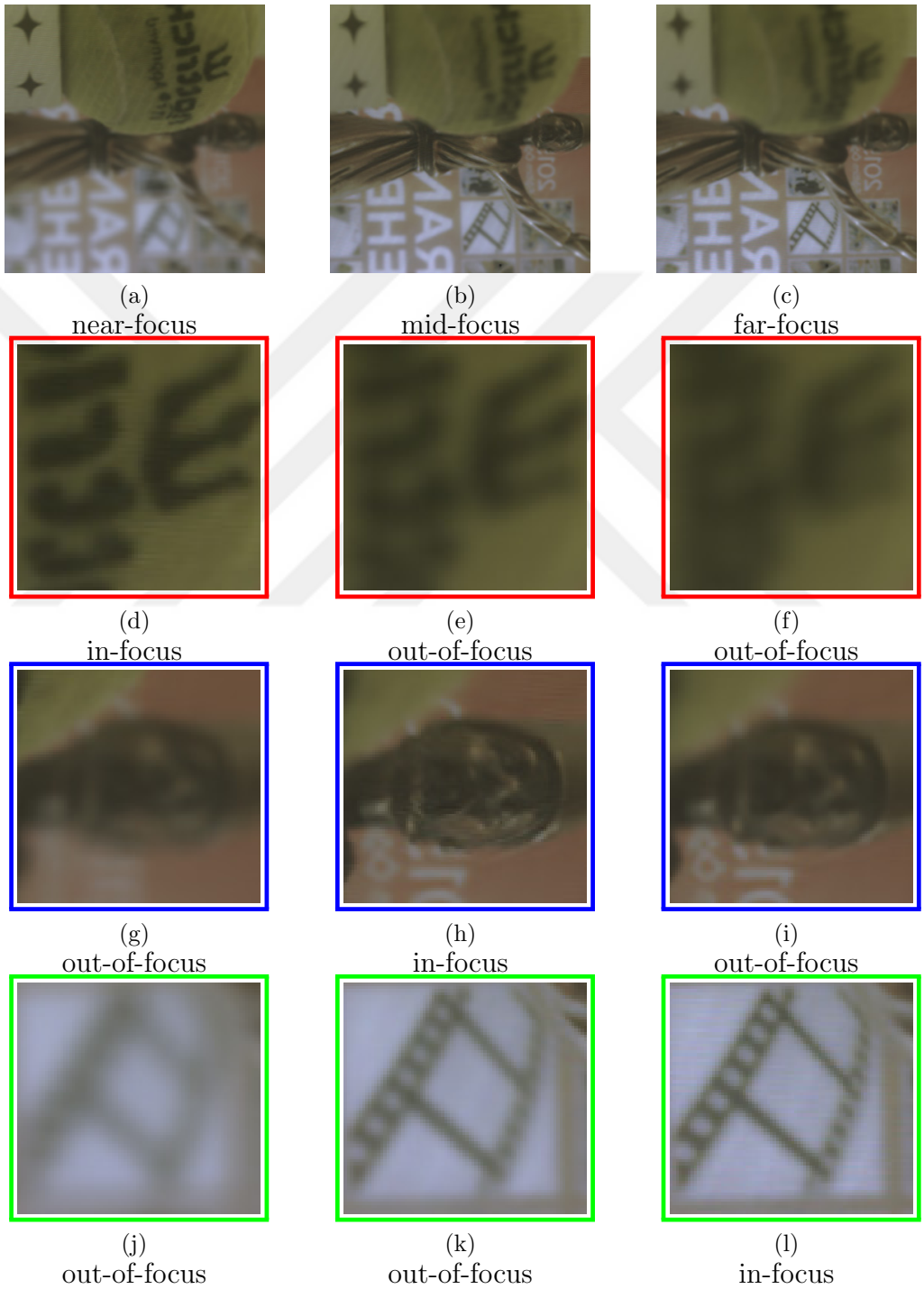


Figure 3.23: Post-capture digital refocusing. (Left column) Near refocus. (Middle column) Mid refocus. (Right column) Far refocus.

Chapter 4

Conclusion

This thesis addressed the problem of limited dynamic range of computational imaging system. In literature, this problem is generally solved by i) single camera and multi-shot technique, and (ii) camera array approach. The former approach suffers from "ghosting" artifacts in dynamic scenes while the latter approach is costly as it requires multiple conventional cameras to form a camera array and may also suffer from occlusion due to the large baseline between the cameras. The solution proposed in this work enables capturing of an HDR image from a single shot, thereby catering to issues related to dynamic scenes, occlusion and camera costs. This is achieved by utilizing a minor optical modification of MLA based light field camera. This modification involves an optical mask which amplifies vignetting in order to capture different parts of the dynamic range with different sub-aperture images. Due to the light field sensor multiple images with different exposure values are captured within a single shot. Since the baseline between the sub-aperture images is small, the resultant HDR image does not suffer from occlusion.

There are still some challenges and opportunities which need to be explored to further improve the quality of the HDR image created using light field imaging system. Light field imaging system have some post-processing capabilities such as depth estimation which can be utilized to further improve computational

imaging applications. However, this has not been investigated in this work. Additionally, plenoptic cameras inherently suffers from low spatial resolution in the sub-aperture images due to spatial and angular resolution trade-off; and this needs to be addressed in some way, maybe through the application of super-resolution techniques.

Bibliography

[1] "Some locally adaptive tone mapping methods for color and exposure error correction." http://csanet.org/newsletter/fall07/nlf0702.html. Accessed: 2018-01-14.

[2] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 31–42, ACM, 1996.

[3] U. Mukati, "Extending light field camera capabilities," Master's thesis, Istanbul Medipol University, 2017.

[4] B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," in *Transactions on Graphics (TOG)*, vol. 24, pp. 765–776, ACM, 2005.

[5] "Lytro, Inc." https://www.lytro.com/. Accessed: 2018-03-07.

[6] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *International Conference on Computer Graphics and Interactive Techniques*, pp. 369–378, 1997.

[7] "Imatest, Inc." https://www.imatest.com/solutions/dynamic-range/. Accessed: 2018-04-02.

[8] "Comparison of dynamic range, Inc." http://www.outdoorphotoacademy.com/. Accessed: 2018-02-14.

[9] B. C. Madden, "Extended intensity range imaging," 1993.

[10] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 374–380, IEEE, 1999.

[11] A. R. Várkonyi-Kóczy, A. Rovid, and T. Hashimoto, "Gradient-based synthesized multiple exposure time color hdr image," *Transactions on Instrumentation and Measurement*, vol. 57, no. 8, pp. 1779–1785, 2008.

[12] A. Artusi, F. Banterle, K. Debattista, and A. Chalmers, *Advanced high dynamic range imaging: theory and practice*. AK Peters/CRC Press, 2011.

[13] "Inst dynamic range." https://www.cambridgeincolour.com/tutorials/dynamic-range.htm/. Accessed: 2018-02-15.

[14] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Computer Vision and Pattern Recognition*, vol. 1, pp. 472–479, IEEE, 2000.

[15] S. K. Nayar and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in *International Conference on Computer Vision*, vol. 2, pp. 1168–1175, IEEE, 2003.

[16] S. K. Nayar, V. Branzoi, and T. E. Boult, "Programmable imaging using a digital micromirror array," in *Computer Vision and Pattern Recognition*, vol. 1, pp. 428–436, IEEE, 2004.

[17] M. Aggarwal and N. Ahuja, "Split aperture imaging for high dynamic range," *International Journal of Computer Vision*, vol. 58, no. 1, pp. 7–17, 2004.

[18] M. McGuire, W. Matusik, H. Pfister, B. Chen, J. F. Hughes, and S. K. Nayar, "Optical splitting trees for high-precision monocular imaging," *Computer Graphics and Applications*, vol. 27, no. 2, pp. 32–42, 2007.

[19] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen, "A versatile hdr video production system," *Transactions on Graphics (TOG)*, vol. 30, no. 4, pp. 41–50, 2011.

[20] T. Georgiev, A. Lumsdaine, and S. Goma, "High dynamic range image capture with plenoptic 2.0 camera," in *Signal recovery and synthesis*, Optical Society of America, 2009.

[21] U. Seger, H.-G. Graf, and M. E. Landgraf, "Vision assistance in scenes with extreme contrast," *Micro*, vol. 13, no. 1, pp. 50–56, 1993.

[22] V. Brajovic and T. Kanade, "A sorting image sensor: An example of massively parallel intensity-to-time processing for low-latency computational sensors," in *Robotics and Automation*, vol. 2, pp. 1638–1643, IEEE, 1996.

[23] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," in *Transactions on Graphics (TOG)*, vol. 23, pp. 760–768, ACM, 2004.

[24] O. Bimber and D. Iwai, *Superimposing dynamic range*, vol. 27. ACM, 2008.

[25] O. Bimber, D. Klöck, T. Amano, A. Grundhöfer, and D. Kurz, "Closed-loop feedback illumination for optical inverse tone-mapping in light microscopy," *Transactions on Visualization and Computer Graphics*, vol. 17, no. 6, pp. 857–870, 2011.

[26] R. Hoskinson, B. Stoeber, W. Heidrich, and S. Fels, "Light reallocation for high contrast projection using an analog micromirror array," in *Transactions on Graphics (TOG)*, vol. 29, pp. 165–175, ACM, 2010.

[27] J. Duan and G. Qiu, "Fast tone mapping for high dynamic range images," in *Pattern Recognition*, vol. 2, pp. 847–850, IEEE, 2004.

[28] O. C. L. Au and C. H. Liu, "Image characteristic oriented tone mapping for high dynamic range images," Dec. 24 2009. US Patent App. 12/142,946.

[29] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Transactions on Graphics (TOG)*, vol. 21, pp. 257–266, ACM, 2002.

[30] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," in *Transactions on Graphics (TOG)*, vol. 21, pp. 249–256, ACM, 2002.

[31] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 267–276, 2002.

[32] Z. Li, S. Rahardja, S. Yao, J. Zheng, and W. Yao, "High dynamic range compression by half quadratic regularization," in *16th International Conference on Image Processing (ICIP)*, pp. 3169–3172, IEEE, 2009.

[33] A. Gershun, "The light field," *Journal of Mathematics and Physics*, vol. 18, no. 1, pp. 51–151, 1939.

[34] E. H. Adelson and J. R. Bergen, *The plenoptic function and the elements of early vision*. Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology, 1991.

[35] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *International Confernece on Computer Graphics and Interactive Techniques*, pp. 43–54, ACM, 1996.

[36] J. C. Yang, M. Everett, C. Buehler, and L. McMillan, "A real-time distributed light field camera," in *Eurographics Workshop on Rendering*, pp. 77–86, 2002.

[37] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, vol. 2, no. 11, pp. 1–11, 2005.

[38] C. Perwass and L. Wietzke, "Single lens 3d-camera with extended depth-of-field," in *Human Vision and Electronic Imaging*, vol. 17, pp. 4–19, 2012.

[39] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *International Conference on Computational Photography*, pp. 1–8, IEEE, 2009.

[40] J. Unger, A. Wenger, T. Hawkins, A. Gardner, and P. E. Debevec, "Capturing and rendering with incident light fields," in *Eurographics Workshop on Rendering*, pp. 141–149, 2003.

[41] T. Georgiev, K. C. Zheng, B. Curless, D. Salesin, S. K. Nayar, and C. Intwala, "Spatio-angular resolution tradeoffs in integral photography," in *Eurographics Conference on Rendering Techniques*, pp. 263–272, 2006.

[42] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *Transactions on Graphics (TOG)*, vol. 26, pp. 1–12, 2007.

[43] A. Manakov, J. F. Restrepo, O. Klehm, R. Hegedus, E. Eisemann, H. P. Seidel, and I. Ihrke, "A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging," *Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 1–14, 2013.

[44] "Light field camera: Depth refocusing and aperture adjustment with light field data." https://inst.eecs.berkeley.edu/ cs194-26/fa16/upload/files/proj6/cs194-26-aam/. Accessed: 2018-01-07.

[45] A. Wahab, M. Z. Alam, and B. K. Gunturk, "High dynamic range imaging using a plenoptic camera," in *Signal Processing and Communications Applications Conference (SIU), 25th*, pp. 1–4, IEEE, 2017.

[46] J. Tumblin, A. Agrawal, and R. Raskar, "Why i want a gradient camera," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 103–110, 2005.

[47] Mann, Picard, S. Mann, and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *Workshop on Imaging Systems and Techniques*, pp. 442–448, IEEE, 1995.

[48] T. Azuma and A. Morimura, "Image composite method and image composite device," *Japanese Patent*, vol. 8154201, 1996.

[49] E. Ikeda, "Image data processing apparatus for processing combined image signals in order to extend dynamic range," Sept. 1 1998. US Patent 5,801,773.

[50] C. Liu, *Beyond pixels: exploring new representations and applications for motion analysis*. PhD thesis, Citeseer, 2009.

[51] I. Hossain and B. K. Gunturk, "High dynamic range imaging of non-static scenes," in *IS&T/SPIE Electronic Imaging*, vol. 7876, pp. 1–8, International Society for Optics and Photonics, 2011.

[52] R. Mantiuk, G. Krawczyk, R. Mantiuk, and H.-P. Seidel, "High dynamic range imaging pipeline: Perception-motivated representation of visual content," in *Human Vision and Electronic Imaging XII*, vol. 6492, International Society for Optics and Photonics, 2007.

[53] E. Reinhard, W. Heidrich, P. E. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.

[54] J. Tumblin and H. Rushmeier, "Tone reproduction for realistic images," *Computer Graphics and Applications*, vol. 13, no. 6, pp. 42–48, 1993.

[55] Y. Salih, A. S. Malik, N. Saad, *et al.*, "Tone mapping of hdr images: A review," in *Intelligent and Advanced Systems (ICIAS)*, vol. 1, pp. 368–373, IEEE, 2012.

[56] A. Adams and R. Baker, *The negative*. Little, Brown, 1981.

[57] D. Sidibé, W. Puech, and O. Strauss, "Ghost detection and removal in high dynamic range images," in *17th European Signal Processing Conference*, pp. 2240–2244, IEEE, 2009.

[58] F. Comby and O. Strauss, "Using quasi-continuous histograms for fuzzy main motion estimation in video sequence," *Fuzzy Sets and Systems*, vol. 158, no. 5, pp. 475–495, 2007.

[59] O. Gallo, N. Gelfandz, W. Chen, M. Tico, and K. Pulli, "Artifact-free high dynamic range imaging," in *International Conference on Computational Photography*, pp. 1–7, IEEE, 2009.

[60] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust patch-based hdr reconstruction of dynamic scenes.," *Transactions on Graphics (TOG)*, vol. 31, no. 6, 2012.

[61] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," in *Computer Vision and Pattern Recognition, CVPR*, pp. 1–8, IEEE, 2008.

[62] O. T. Tursun, A. O. Akyüz, A. Erdem, and E. Erdem, "The state of the art in hdr deghosting: a survey and evaluation," in *Computer Graphics Forum*, vol. 34, pp. 683–707, Wiley Online Library, 2015.

[63] K. Jacobs, C. Loscos, and G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," *Computer Graphics and Applications*, vol. 28, no. 2, pp. 84–93, 2008.

[64] J. An, S. J. Ha, and N. I. Cho, "Probabilistic motion pixel detection for the reduction of ghost artifacts in high dynamic range images from multiple exposures," *EURASIP Journal on Image and Video Processing*, vol. 2014, pp. 1–15, 2014.

[65] D. Cho, M. Lee, S. Kim, and Y. Tai, "Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction," in *Proceedings of the International Conference on Computer Vision*, pp. 3280–3287, IEEE, 2013.

[66] N. Sabater, V. Drazic, M. Seifi, G. Sandri, and P. Perez, "Light-field demultiplexing and disparity estimation," 2014.

[67] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in *International Conference on Computer Vision and Pattern Recognition*, pp. 1027–1034, IEEE, 2013.

DYNAMIC RANGE ENHANCEMENT IN LIGHT FIELD IMAGING SYSTEMS

ORIGINALITY REPORT

18%	12%	14%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

2%

★ Submitted to Middle East Technical University

Student Paper

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		