



NORTHWESTERN UNIVERSITY

Computer Science Department

Technical Report
Number: NU-CS-2022-12

August, 2022

Computational Framework for Light Field Microscopy Simulation and Optimization

Jipeng Sun

Abstract

Light field microscopy (LFM) is the key technology for optically imaging the neuron activity in live animal's brain. However, there is still no computational framework that could provide an unified simulation and optimization process. The presented work proposed and demonstrated a computational simulation and optimization framework for LFM systems. The proposed framework consists of three main modules: forward model, backward model, and the optimizer. The paper comprehensively introduces the theory background and implementation details behind each module. The expecting hope of the developed computational framework is to let users not from computation side could still quickly prototype and further optimize their LFM optical design and reconstruction model. Furthermore, the paper also contributes in the taxonomy for current LFM systems, the microlens array optimization methods, and the optimization pipelines based on the differentiability of the models.

Keywords

Light Field Microscopy Simulation, Light Field Volume Reconstruction, Light Field Microscopy Optimization, Deep Optics, Volumetric Zebrafish Brain Imaging

NORTHWESTERN UNIVERSITY

Computational Framework for Light-Field Microscopy Simulation and
Optimization

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

MASTER OF SCIENCE

Field of Computer Science

By

Jipeng Sun

EVANSTON, ILLINOIS

August 2022

© Copyright by Jipeng Sun 2022

All Rights Reserved

ABSTRACT

Computational Framework for Light-Field Microscopy Simulation and Optimization

Jipeng Sun

Light field microscopy (LFM) is the key technology for optically imaging the neuron activity in live animal's brain. However, there is still no computational framework that could provide an unified simulation and optimization process. The presented work proposed and demonstrated a computational simulation and optimization framework for LFM systems. The proposed framework consists of three main modules: forward model, backward model, and the optimizer. The paper comprehensively introduces the theory background and implementation details behind each module. The expecting hope of the developed computational framework is to let users not from computation side could still quickly prototype and further optimize their LFM optical design and reconstruction model. Furthermore, the paper also contributes in the taxonomy for current LFM systems, the microlens array optimization methods, and the optimization pipelines based on the differentiability of the models.

Acknowledgements

I still remember the day when we had our first LFM project meeting and I boldly presented my LFM simulation ideas based on the ray optics and the opaque imaging volume assumptions (which are obviously inaccurate). All my experienced advisors were listening patiently without interruption. After I finished, they started to carefully correct my mistakes and encouraged me to keep exploring. From then on, I started to enter the world of computational microscopy.

Obviously, without my dear advisors Dr. Oliver Cossairt, Dr. Jack Tumblin, Dr. Emma Alexander, Dr. Florian Willomitzer, and Dr. Geoffery Goodhill, a zero optics background young man can never write out such thesis in 8 months. I would like to express my deep gratitude for all their guidance, inspiration, and love. Thank you all for opening the door of computational photography research for me.

In addition, I would like to express my appreciation to my great friends PhD students Hamid Hasani, Florian Schiffrers, Manuel Ballester, Lionel Fiske, Aniket Dashpute, and Jiazhang Wang, who kindly provide precious suggestions and generous help on accomplishing this project.

Finally, I am grateful to my family: my parents Jian Sun and Zhonghong Pan, my grandfather Mingyuan Sun and Xingliang Pan, my fiancée Yuhan Chen and her brother Reginald Chen for providing me with endless support, especially in my education, and encouragement throughout my life. Also, this paper is for my died grandmother Defeng

Wang and Jianying Guan. I wish they could see what I have done and rest in peace in heaven.

Table of Contents

ABSTRACT	3
Acknowledgements	4
Table of Contents	6
List of Tables	8
List of Figures	9
Chapter 1. Introduction	10
1.1. Related Work	11
1.2. Thesis Organization	15
Chapter 2. Light Field Microscopy Simulation Workflow	16
2.1. The Rationale of LFM Simulation Pipeline	17
2.2. The LFM Simulation Pipeline	19
2.3. Commonly Used Evaluation Factors for LFM Systems	24
2.4. Crosswise Comparison Among Published LFM Systems	29
Chapter 3. Forward Model	32
3.1. Optical Components and Propagators Modeling	32
3.2. Forward Models for Conventional and Fourier LFM Systems	44

	7
3.3. Parameter Space of the Forward Model for Optimization	50
Chapter 4. Backward Model	56
4.1. Non-differentiable Deconvolution Methods	57
4.2. Differentiable Deconvolution Methods	61
Chapter 5. Optimization	66
5.1. Optimization for Non-Differentiable Pipelines	67
5.2. Optimization for Partially Differentiable Pipeline	70
5.3. Optimization for Fully Differentiable Pipeline	72
Chapter 6. Conclusion & Future Works	74
References	75

List of Tables

2.1	Literature review of LFMs employed for living Zebrafish larvae Brain imaging 1. The * symbol means based on the provided figures or averaged operation rather than explicitly specified.	30
2.2	Literature review of LFMs employed for living Zebrafish larvae Brain imaging 2	30

List of Figures

2.1	The Light Field Microscopy Hardware Usage Pipeline	20
2.2	The Open-loop LFM Simulation Workflow	21
2.3	The Closed-loop LFM Simulation Workflow	23
3.1	The Optical Setting for Conventional Light Field Microscope	46
3.2	The Simplified Optical Setting for Fourier LFM Systems.	49
3.3	The Axial Resolution Analysis of Fourier LFM Systems.	52
5.1	The Flow Chart for LFM Optimization Strategies	67

CHAPTER 1

Introduction

Light field microscopy (LFM) is a microscopic imaging method based on the light field theory which could capture the 4-dimensional light field information of the specimen (2-dimensional spatial and 2-dimensional angular information) to reconstruct the volumetric information [1]. It has recently demonstrated huge success in imaging live larva zebrafish brain neural activity with cellular resolution[21]. This success is mainly due to its scanless capability in acquiring the plenoptic information of the object and its fast temporal sampling rate. However, since 2006 when the first light-field microscopy is invented by Stanford Computer Graphics Lab[20], even though several heuristic and general-purposed light-field microscopy designs have been explored [3, 5, 10, 22], we still haven't answered the question "How can we get the optimal light field microscopy setting for a specific application?" The challenging nature of the problem lies in two main aspects: what is the forward optical transfer model that could encode most plenoptic information and how to develop a backward model that could decode out as much information as possible from that encoded light field representation. The present paper provides an answer to the above questions by proposing a computational differentiable framework to simulate and optimize the forward image formation process and backward volume reconstruction process of the light field microscopy based on the imaging data. The framework contains three major components: differentiable forward model part, backward reconstruction part, and the optimization part. Each components can be replaced by its counterpart methods

following the interface specification so that the framework is compatible with the more advanced algorithms in the future works. Besides, the whole framework is implemented with GPU accelerable libraries which could gain benefits from parallel computation and neural network based forward and backward methods.

To summarize, the main contributions are:

- Proposed a methodology to design the optimal light-field microscopy setups based on the imaging volume data.
- Implemented a computational framework which is compatible with GPU acceleration and more advanced algorithms for image formation, volume reconstruction and optimization in the future.

1.1. Related Work

Even though all light field microscopies are based on the light field theory which is to synthesize the volumetric scene by capturing 2-dimensional scene images from a set of prefixed viewing angles, there are mainly two different optical settings of light field microscopy until now: conventional LFM and Fourier LFM. Both of them use microlens array to record the 4-dimensional light field information. The difference is that the conventional one records light field in spatial domain of the volume while the Fourier one record in the Fourier domain. To our best of knowledge, there is no previous work that quantitatively demonstrate which design option is more advantageous, but recent Fourier LFM works have qualitatively compared and analyzed the superiority of it [10].

1.1.1. LFM with Microlens Array on the Native Image Plane

The first kind of setting comes from the conventional light field camera[1, 27]. Levoy et al. (2006) firstly applied the light field camera into the microscopy domain[20]. In the initial version of the light field microscopy, a microlens array is located on the native image plane for dissecting light coming from different projection angles. An image of the microlens aperture lightened by the object emitting light will be formed on the camera sensor. Each individual 2D perspective image can be extracted from the same relative-position pixel in each individual microlens aperture subimage by ray tracing idea. This can encode angular information into the raw 2D image on the sensor. However, the lateral resolution of each perspective is sacrificed and it is limited by the pitch size of the microlens array. In addition to the low lateral resolution shortage, the image formation model is based on the ray optics, which doesn't explain well for the diffraction phenomenon in microscopy applications since the voxel size of the sample is close to the wavelength of the illumination light.

To resolve the bias brought by the ray optics theory and attenuate the sacrifice on the lateral resolution, Broxton et al. (2013) proposed a new wave optics forward model and a deconvolution algorithm taking super resolution ideas into consideration[3]. By simulating the point spread function (PSF) matrix for all the pairs of voxel position of the object and the pixel position of the camera sensor, the light field image of the volume can be calculated by multiplying the volume vector with the PSF matrix. The paper has also proofed that the spatial band limit of this LFM setting is not limited by the pitch size of the microlens array and it could be extended by shifting the object out of the native object plane to a higher spatial sampling rate position according to the chief

rays analysis. Thus, the combination of the higher sampling rate aliased 2D scenes can provide higher frequency information for later volume reconstruction. Even though the improved wave optics method provides a more accurate and higher-resolution 3-D volume imaging simulation compared with the initial one, the shift-variant PSF simulation brings huge computational cost for both forward imaging and backward reconstruction models. Besides, the approximation error introduced by the simulated PSFs also makes it hard to applied in the real-world tasks such as free-moving larva zebrafish brain imaging.

Albeit the huge computational cost of the conventional LFM makes it hard to reconstruct volume, the benefits of the unrestricted imaging rate still push it become one of the most popular method for optical brain imaging [49]. To overcome the huge computational cost of shift-variant PSF function, a view-channel-depth (VCD) deep neural network is proposed so that it could leverage the GPU acceleration to achieve real-time volumetric reconstruction.[41] The training data of the reconstruction model comes from the synthetic wave-optics 2D light field images of the experimental ground truth volume. However, applying neural network methods into reconstruction without fitting well on a comprehensive training data set will bring unexpected bias for novel volumes. Thus, a physics agnostic reconstruction model is not reliable enough for such low-level measurement tools.

1.1.2. LFM with Microlens Array on the Fourier Plane

Another optical setting for LFM is to add a Fourier lens after the tube lens to get a Fourier transform of the object image and then put a microlens array to record the 4-dimensional light field of the Fourier image. After propagating a focal length of the microlens, a

spatial domain image of the object plane viewed from a specific angle will be formed on the camera sensor behind each microlens.

The Fourier LFM idea is firstly built and described by Llavador et al. as Fourier Integral Microscopy in 2016 for the spatial resolution and depth of field improvement purpose[23]. The experimental optical hardware result is demonstrated in the paper but the theoretical model is based on the ray optics which makes it unqualified for high resolution microscopy. Later in 2017, Cong et al. built a light field microscopy with similar optical setting called XLFM to firstly apply Fourier light field microscopy for free-moving zebrafish brain neural activity imaging [5]. It achieved a $3.4 \times 3.4 \times 5 \mu m^3$ optimal resolution in a part of $800 \mu m$ diameter wide and $400 \mu m$ height cylinder volume. By further combining XLFM with confocal microscopy, Zhang et.al improved the spatial resolution to $2 \times 2 \times 2.5 \mu m^3$ over a cylindrical imaging volume of diameter $800 \mu m \times$ height $200 \mu m$ [48]. However, all of the above works are done experimentally without a simulation framework to further optimize.

To fulfill the gap between the physical experiments and computational simulation of the Fourier LFM, Guo et al. proposed a general design principle for Fourier LFM based on the simulation in 2019 [10]. The open sourced simulation is done with wave optics and is implemented in Matlab. The Fourier LFM parameters are summarized and organized into three categories as input, performance, and design. By providing the desired performance and input parameters, the model is claimed to be able to return one set of design parameters that will satisfy the requirement even though it is not obviously demonstrated and instructed in the source code. The Fourier LFM simulation work provides a general

guidance for Fourier LFM design but still lacks supports on the optimization process, especially the potential benefits from leveraging the prior task data distribution.

1.2. Thesis Organization

The remainder of the thesis is structured as follows: §2 discusses the rationale and the architecture of the light field microscopy simulation workflows, along with the commonly used evaluation factors and the comparison for published LFM works for zebrafish brain imaging. §3 introduces the optical components and optical propagators used for LFM systems, and demonstrates the forward models for both conventional LFM and Fourier LFM built by our framework. §4 covers more on the backward models used for reconstructing the volume in LFM systems. Finally, §?? discusses both the differentiable and non-differentiable optimization pipeline for LFM systems.

CHAPTER 2

Light Field Microscopy Simulation Workflow

This chapter introduces the pipeline of simulation for light field microscopy projects including the forward image formation model, the backward volume reconstruction model, and the optional optimization process. The initial goal of applying light field technology into the microscopy task is to achieve the animated volumetric imaging in high temporal resolution. The main cost of realizing high temporal resolution for light field microscopy is to split the limited imaging sensor area for several individual views from different perspectives, which leads to the deduction of the spatial resolution for the synthetic volume. Besides, to achieve volume reconstruction with high accuracy, more angular information needs to be encoded on the image sensor, which further shrinks down the bandwidth of representing spatial information for each individual scene. This inherent trade-off leads researchers to explore in what extent should a design be compromised to meet the final volumetric reconstruction goal. Thus, computational simulation plays a more and more important role in designing modern light field microscopy[10, 22, 49].

The purpose of light field microscopy computational simulation is to measure the performance of a designed LFM by some metrics we defined without the need of expensive hardware trials. To better understand the workflow of the LFM simulation, the chapter will first focus on the rationale of the whole LFM simulation pipeline, and then come down to each components of the pipeline, and finally introduce the useful metrics to evaluate such process.

2.1. The Rationale of LFM Simulation Pipeline

As mentioned above, the LFM simulation process is decomposed into three modules: the forward model that encodes the volumetric information onto the image plane; the backward model that decodes the encoded information to reconstruct back the volume; and the possible optimizations applied on the parameters in both the encoding and decoding process. Thus, the whole process can be treated as a classical encoder-decoder structure [13]. By thinking the LFM system under the encoder-decoder framework, we could better understand the core of its methodology: rather than directly measuring the high-dimensional data (the 3D volume to measure), we now convert the problem to find its low-dimensional codes (the 2D camera data on the image plane) and corresponding decoder (the point spread functions of the optical system).

To further understand the above argument, the linearity of the optical system needs to be introduced. For applications that use microscopy to image small animals with transparent bodies and brains, the occlusion of the sample particles can often be ignored. Besides, the interior light scattering effect for such transparent sample is negligible so that the optical system can be treated as a linear system, which means the final light distribution of the entire volume on the image plane is equal to the sum of the light contribution of each individual voxel. Thus, the final image of the volume I_{sens} can be calculated by convolving every voxel with its corresponding point spread function (PSF) as equation 2.1.

$$(2.1) \quad I_{sens} = O * * PSF_{lens}$$

Based on such assumptions, the imaging problem for the infinitely large number of varieties of the volume can be simplified by only considering the point spread functions of the system. Therefore, the main task for the forward model now is to simulate how the optical system responses to the pulse input function. Once the PSFs are measured or calculated, the encoder-decoder workflow becomes to a typical convolution-deconvolution process, which has been well studied in the computational photography community.

Given the PSFs and the image data on the camera sensor, the process of deconvolution is to apply the inverse function of convolution. Braxton et. al. in 2013 proofed that the LFM reconstruction for transparent object is equal to solve a limited-angle tomography problem [3]. To solve that well studied inverse problem, there are mainly two groups of approaches could be adopted: the iterative methods [33, 24, 7, 15, 2] and the deep learning methods [41, 6, 25]. No matter what method is finally taken, the inputs are always the PSF functions and the image plane data and the output is always an estimated reconstruction volume. Any function that fulfills this requirement could be used as a backward function, which brings the decoupling possibility for computational framework design.

By the forward and backward models described above, one can already achieve a fairly good results from the open-loop process. However, to further improve the performance of the LFM system, a closed-loop process should be developed. In other words, we need to be able to somehow leverage the feedback from the intermediate or final results to influence the information flow of the LFM system. The closed-loop idea brings huge potential space for the LFM performance improvement since it could be a system that tuned for the specific task data or a specific set of non-ideal real-world optical hardware,

no longer to be restricted to be a general purpose ideal system, which also forms the basis of the optimization module of the proposed computational framework.

2.2. The LFM Simulation Pipeline

For LFM simulation, unlike in the physical world experiments where the encoded image plane result is directly acquired by building out the LFM hardware and putting the volume to be measured on the object plane, the encoded light field information of a volume won't be simulated unless the full information of the volume has already been accessible by the program. For the purpose of measuring the volume, the forward simulation process seems to be redundant. This conclusion reveals two important facts which needs to be stressed.

First, the use of the light field microscopy is not coupling with the computational simulation. In this case, the built-out optical hardware is directly used as an encoder. However, the second fact is, to come up with a high-performance hardware design, LFM simulation is the most efficient and economic approach by now. More importantly, as the later of the section will show, it could close the loop of the optimization applied between the ground truth volume and the estimated volume, thus, bring significant improvement on the LFM performance.

In this section, we firstly describe the process of using LFM hardware to image and reconstruct the volume under the condition that the optical hardware has already been set. Then, we extend the LFM pipeline to simulation assisted design stage. Based on whether there exists a feedback loop in the information flow, we discuss the pipeline for open-loop simulation and closed-loop simulation separately.

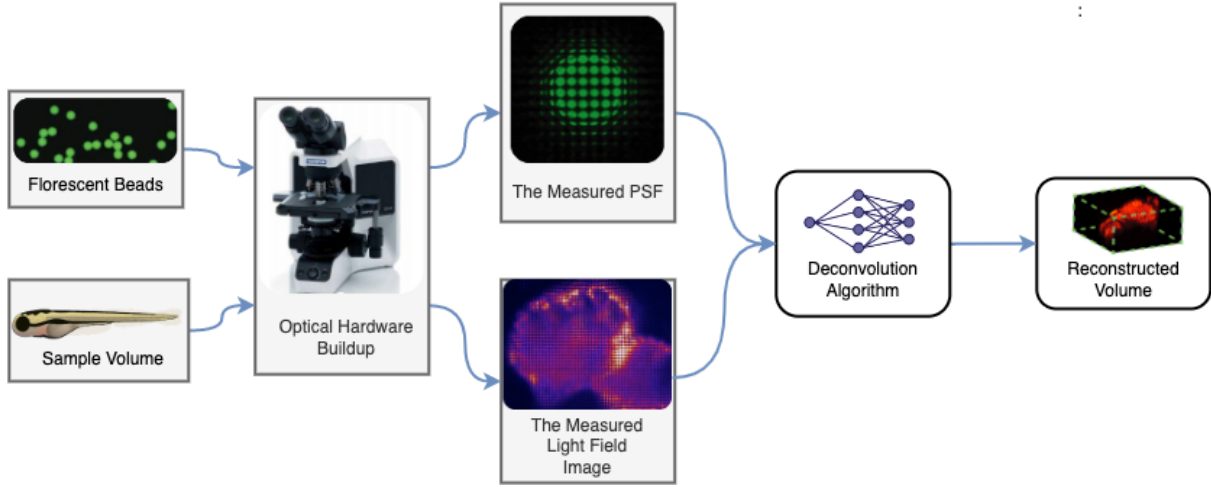


Figure 2.1. The Light Field Microscopy Hardware Usage Pipeline

2.2.1. The Light Field Microscopy Hardware Usage Pipeline

In the LFM hardware in-service stage, computation only happens in the reconstruction part. See the Figure 2.1¹. The optical hardware is directly used as the forward model. The required point spread functions for the backward model input are also acquired by physically imaging and recording the well-prepared sparse florescent beads sample with different focal depths. The backward model will take the outputs from the hardware LFM and accomplish a PSF-agnostic deconvolution. Then, the reconstructed volume is evaluated by some non-reference metrics such as spatial resolution, field of view, depth of view, etc since there is no way to access the ground truth volume data. To summarize, there are four steps in the LFM hardware usage process:

- (1) Design and build up a set of optical hardware for the LFM system
- (2) Recording the point spread functions stack for the optical system

¹Image element courtesy of: Lizzy Griffiths, Olympus Cooperation, Stanford Computer Graphics Lab, Turaga Lab, Prevedel Lab

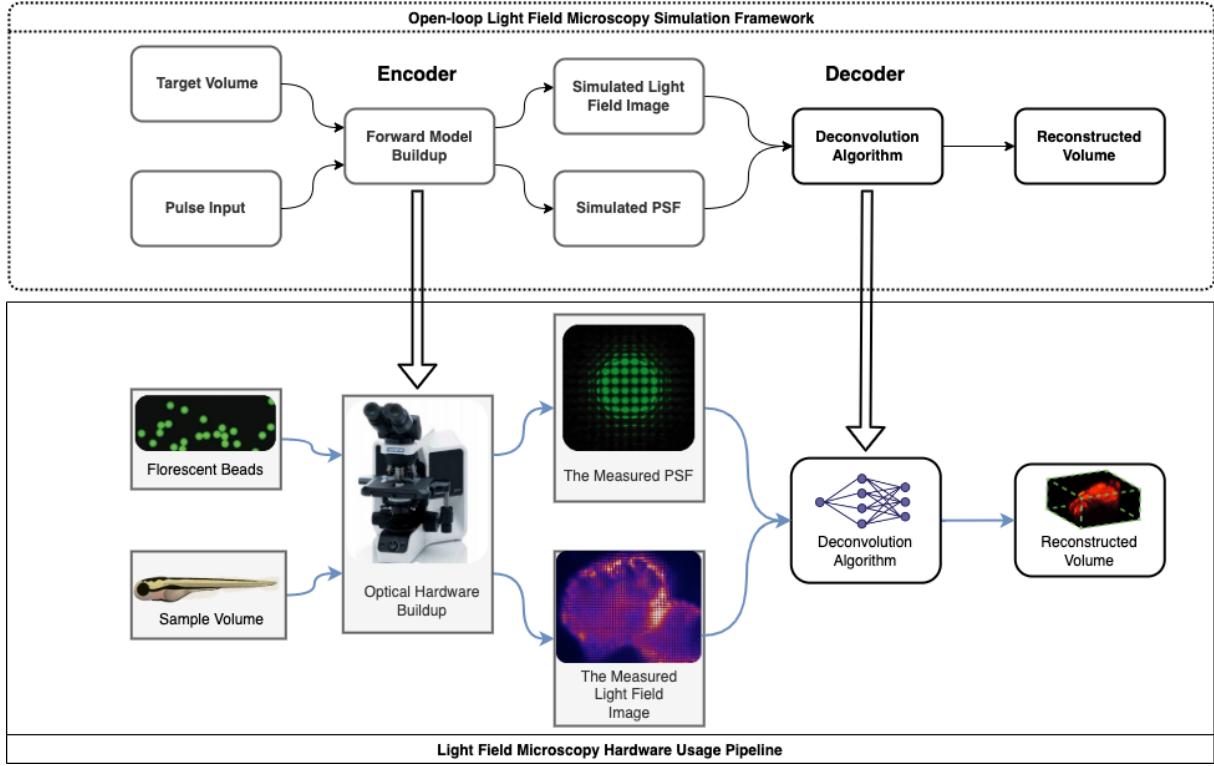


Figure 2.2. The Open-loop LFM Simulation Workflow

- (3) Imaging the light field of the targeted volume sample
- (4) Deconvolve the light field image with the measured PSFs and get the final reconstruction.

2.2.2. The Light Field Microscopy Open-loop Simulation Pipeline

In the design stage, the optical hardware setting is undecided. Doing simulation to predict the performance of the proposed design is necessary. Open-loop LFM simulation is defined as the simulation process that doesn't contain a feedback loop to optimize the model performance. Most of the published LFM simulation works fall into this category [3, 10, 22].

In the open-loop simulation, complex wavefront function that modulated by the optical components and propagators replaces the hardware encoder. See the Figure 2.2². The simulated target volume and pulse input pass through the designed forward model and form the light field result on the camera plane. After the backward function, the reconstructed volume can be evaluated by both non-reference metrics or reference metrics, since the ground truth volume data is accessible for simulation, to decide whether the proposed design is acceptable. If not, another set of predefined parameters will be tested again and continue the above process. The parameters set that achieves the best simulation result will be used to build up the final optical hardware. Once the hardware is set, unless further hardware in-the-loop optimization is required, the simulation framework will no longer be needed and the previously described in-service stage pipeline will start to act. In summary, the open-loop LFM simulation pipeline is

- (1) Propose an optical setting to be tested and develop its forward model
- (2) Calculate the point spread functions stack for the initial LFM system
- (3) Generate or acquire the training volume data and convolve it with the PSF stack to get the light field image.
- (4) Deconvolve the light field image with the PSF stack to get the estimated reconstruction.
- (5) Evaluate the purposed design by non-reference metrics. If fails, back to the first step and test with another set.
- (6) Build the optical hardware based on that simulated optimal setting

²Image element courtesy of: Lizzy Griffiths, Olympus Cooperation, Stanford Computer Graphics Lab, Turaga Lab, Prevedel Lab

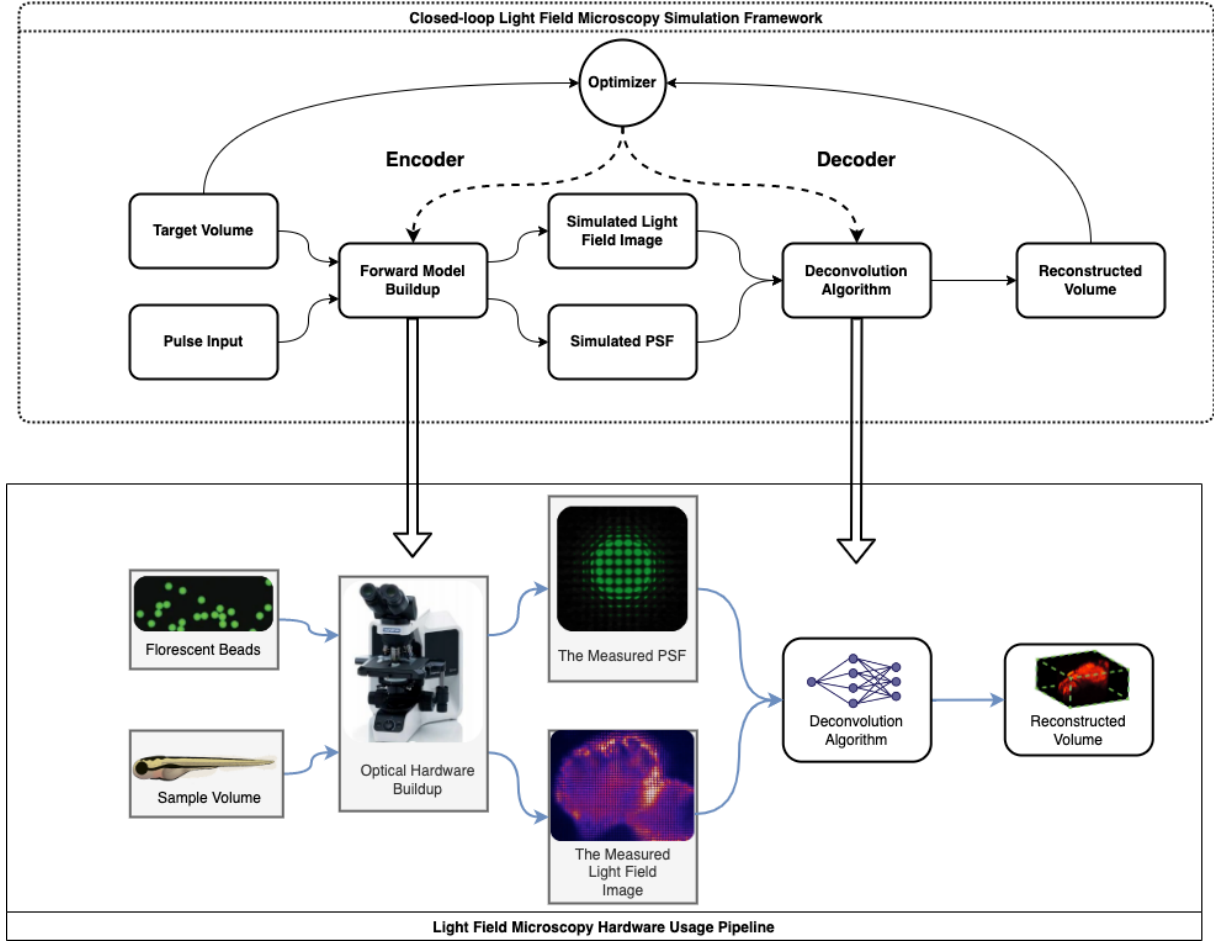


Figure 2.3. The Closed-loop LFM Simulation Workflow

2.2.3. The Light Field Microscopy Closed-loop Simulation Pipeline

Inspired by the deep optics idea [36, 43] which utilizes the programmable optics components to optimize the imaging quality, closed-loop LFM simulation framework leverage minimizing the defined error between the target volume and estimated volume to provide the gradient to automatically guide the forward and backward model improvements. So that there is no need to grid search the predefined optical parameters to find a local optimum. Like its open-loop counterpart, since the ground truth volume data is accessible,

the reference factors can also be its evaluation metric after an iteration of reconstruction. As the Figure 2.3³ shows, it closes the loop of the optimization applied between the ground truth volume and the estimated volume. Thus, the whole optical design can adapt to a specific application usage and joint optimize the forward and backward model. After the simulation automatically converges to a local optimal set of parameters, the LFM hardware can be deployed based on that solution.

In summary, the process of the closed-loop LFM simulation contains following steps:

- (1) Design an initial optical setting for LFM and develop its forward model
- (2) Calculate the point spread functions stack for the initial LFM system
- (3) Generate or acquire the training volume data and convolve it with the PSF stack to get the light field image.
- (4) Deconvolve the light field image with the PSF stack to get the estimated reconstruction.
- (5) Optimize the forward and the backward models based on the gradient feedback to get the final optimal setting.
- (6) Build the optical hardware based on that simulated optimal setting
- (7) Make the algorithm further adapt to the hardware artifacts based on the camera-in-the-loop idea. (Optional future hardware improvement work)

2.3. Commonly Used Evaluation Factors for LFM Systems

Since the first light field microscopy is developed in 2006, more than fifteen different varieties of LFMs have been employed for live imaging [49]. In different degree, all the

³Image element courtesy of: Lizzy Griffiths, Olympus Cooperation, Stanford Computer Graphics Lab, Turaga Lab, Prevedel Lab

published LFMs have made remarkable progress in their specific domains. However, to the best of our knowledge, there is no literature comparison among their performance which could distinguish the better optical setting to guide the future LFM design. Besides, without defining a set of effective evaluation metrics first, it is impossible to crosswise compare the different plans and come to the notion of 'optimization'. Thus, we will focus on developing a set of quantitative metrics for LFMs and compare the performance of published LFM systems for zebrafish-imaging in this section.

The information flow charts in section 2.2 illustrate a fundamental difference between hardware usage workflow and the software simulation pipeline: there is no ground truth data for reference in the hardware in-service stage. Thus, the quantitative metrics can be categorized into two groups: the non-reference metrics and reference metrics.

2.3.1. Non-reference Metrics

The non-reference metrics means the input source of the evaluation function is only from the properties of the model to be evaluated, no other reference sources will be considered. For example, the spatial resolution of the model's output. The evaluation result will always be the same as long as the model itself is unchanged. This group of the metrics is the only choice for the LFM in-service stage since there is no access to the physical volume data.

The size of the larva zebrafish brain is about $400 \times 800 \times 250 \mu m^3$ and the typical neuron body cell diameter in larva zebra fish brain is about $3-7 \mu m$ [19]. These biological facts set the fundamental specification for the light field microscopy used for zebrafish

brain imaging. With such physical imaging obligation, we will introduce the following crucial non-reference factors and define its minimum requirement:

Lateral Resolution: The ability to discern between two points perpendicular to the light beam in the same depth. In zebrafish imaging, the lateral resolution of the reconstructed volume directly determines whether the LFM can achieve cellular-level imaging. As the previous facts show, around 150×270 of $3 \times 3 \mu m^2$ voxels are required in the reconstruction volume to achieve the cellular-level resolution. Notice that, the volume resolution is different with the raw camera image and the lateral shape of the volume is not a square.

Axial Resolution: The ability to discern between two points along or parallel to the light beam. By the previous biological facts, at least 50 $5 \mu m$ thick slices are needed in the reconstructed volume to achieve cellular resolution in axial direction.

System Magnification Rate: The ability for LFM systems to resolve axial emitters is based on pre-measuring the PSF functions in all possible depths in object space. Usually, achieving the $5 \mu m$ step size shifting in the object space is not an issue for most widefield microscopes PSF measurement tasks since the objective lens can magnify the step size in image space by the square of the magnification factor, which often leads to a three-order of magnitude larger step size around several millimeters in image space. However, for the LFM systems, the huge difference in focal length between the microlens array and Fourier lens poses a not neglectable shrinking term to the system magnification rate, leading the rate

no larger than 10 for most of the LFM systems. As a result, it imposes higher precision demand to the minimum step size of the shifting stage.

Temporal Resolution: The temporal resolution of the reconstructed volume is same as the camera frame rate for light field microscopy application if there is no time multiplexing methods employed. However, high frame rate of the camera won't always bring benefits for zebrafish brain firing pattern imaging purpose. The basal spiking rate for the larva zebrafish brain neurons without swimming is about 8 spikes/s[12]. Thus, there is not much necessity for the sampling rate to go over 10 HZ even though the highest frame rate for modern camera and SLM can go far beyond it. The extra sampling rate capacity leaves huge potential space for time multiplexing methods.

Field of View: Different with the definition from the photography community, the field of view (FoV) for a LFM system is defined as the maximum diameter of the visible area when looking through the microscope eyepiece or scientific camera. For the immobilized zebrafish brain imaging task, the minimum FoV is $800\ \mu m$. Larger FoV can always be achieved by using objective lens with smaller magnification factor. However, the spatial sampling rate mentioned in the lateral resolution metric will put a lower bound to the magnification factor.

Depth of Field: The depth of field (DoF) of a LFM system is the axial maximum measuring depth range to the specimen. The minimum DoF is $200\ \mu m$ for imaging a whole zebrafish brain. To calculate the DoF of a LFM system, one must consider the Rayleigh range of the LFM system, which is inverse proportional to the numerical aperture (NA). This puts an inherent trade off between DoF and

lateral resolution for LFM system. Except that, the pitch size of the microlens array also puts an upper limit of the DoF since the light field pattern might have the cross talk phenomenon.

2.3.2. Reference Metrics

As the name suggests, the reference metrics is a group of evaluation methods that used when there is a ground-truth. It is a necessary part for constructing the loss functions for most of the closed-loop design. However, due to the ground truth volume is not accessible in the LFM in-service phase, we can't easily use the metrics in comparing different published LFM hardware systems.

2.3.2.1. MSE. Mean squared error (MSE) is the most traditional and simple method for image comparison. It is calculated as the average of all the squared differences in pixel values between the reproduction and original image, where smaller values means closer reproductions and 0 means identical. Good when we want pixel color to be as similar as possible for perceiving.

$$(2.2) \quad MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

2.3.2.2. PSNR. Peak signal-to-noise ratio (PSNR) is a variation of MSE. It is a ratio between the maximum power of an image(255 in an 8-bit pixel representation) to the power of corrupting noise affecting the quality of reconstruction(MSE). Since PSNR increases when MSE decreases, higher values represent better image reconstruction quality.

$$(2.3) \quad PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right)$$

2.3.2.3. SSIM. Structure similarity index method (SSIM) compares windows of the image rather than direct pixel to pixel differences. It factors loss of correlation(structure), luminance distortion, and contrast distortion. Values range from -1 to 1, 1 is identical. Generally better than MSE and PSNR for identifying structural similarity but worse for noisy images. More similar to how humans perceive images.

$$(2.4) \quad SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

2.3.2.4. LPIPS. Learned perceptual image patch similarity (LPIPS) measures perceptual similarity between two images, shown to match human perception. It is a deep neural network trained on database made from humans picking from image sets which distorted image most closely resembled original. Lower values mean more similar. The open-source library can be found in: <https://github.com/richzhang/PerceptualSimilarity>.

2.4. Crosswise Comparison Among Published LFM Systems

Based on the metrics we introduced before, we conducted a literature review on 10 zebrafish brain light field microscopy published works. We summarized our results as Table 2.1 2.2

As Table 2.1 shows. Given the fact that the larva zebrafish brain size is about $400 \times 800 \times 250 \mu m^3$ [19], we can find the Fourier light field microscope developed by Cong etl [5] is the only system that has such imaging capacity to achieve the whole larva

Table 2.1. Literature review of LFMs employed for living Zebrafish larvae Brain imaging 1. The * symbol means based on the provided figures or averaged operation rather than explicitly specified.

Methods	Imaging volume (dimensions order: x, y, z)	Spatial Resolvability	Reference
Conventional LFM	$500\mu\text{m} \times 500\mu\text{m} \times 300\mu\text{m}$	Not Mentioned	Cohen 2014 [4]
Conventional LFM	$700\mu\text{m} \times 700\mu\text{m} \times 200\mu\text{m}$	$2.8\mu\text{m} \times 2.8\mu\text{m} \times 5.2\mu\text{m}$	Prevedel 2014 [32]
Fourier LFM	$650\mu\text{m} \times 440\mu\text{m} \times 200\mu\text{m}$	$6\mu\text{m} \times 6\mu\text{m} \times 16\mu\text{m}$	Perez 2015 [31]
Conventional LFM	$200\mu\text{m} \times 200\mu\text{m} \times 200\mu\text{m}$	$7\mu\text{m} \times 7\mu\text{m} \times 2\mu\text{m}$	Pegard 2016 [30]
Conventional LFM	$700\mu\text{m} \times 700\mu\text{m} \times 200\mu\text{m}$	$3.5\mu\text{m} \times 3.5\mu\text{m} \times 9\mu\text{m}$	Nobauer 2017 [28]
Fourier LFM	$\varnothing 800\mu\text{m} \times 400\mu\text{m}$	$3.4\mu\text{m} \times 3.4\mu\text{m} \times 5\mu\text{m}$	Cong 2017 [5]
Conventional LFM	$700\mu\text{m} \times 700\mu\text{m} \times 200\mu\text{m}$	$2.5\mu\text{m} \times 2.5\mu\text{m} \times 5\mu\text{m}^*$	Taylor 2018 [37]
Conventional LFM	$350\mu\text{m} \times 300\mu\text{m} \times 32\mu\text{m}$	$3.3\mu\text{m} \times 3.3\mu\text{m} \times 5.4\mu\text{m}^*$	Wang 2019 [39]
Conventional LFM	$600\mu\text{m} \times 600\mu\text{m} \times 100\mu\text{m}$	$3\mu\text{m} \times 3\mu\text{m} \times 6\mu\text{m}$	Truong 2020 [38]
Fourier LFM	$\varnothing 800\mu\text{m} \times 200\mu\text{m}$	$2\mu\text{m} \times 2\mu\text{m} \times 2.5\mu\text{m}$	Zhang 2020 [48]

Table 2.2. Literature review of LFMs employed for living Zebrafish larvae Brain imaging 2

Frame rate	Free moving/constrained	Deconvolution Method	Reference
3 Hz	Constrained	Richardson Lucy	Cohen 2014 [4]
20 Hz	Constrained	Richardson Lucy	Prevedel 2014 [32]
Not mentioned	Constrained	Richardson Lucy	Perez 2015 [31]
100 Hz	Constrained	No Deconvolution Process	Pegard 2016 [30]
20 Hz	Head-fixed	Richardson Lucy	Nobauer 2017 [28]
30 Hz	Free moving	Richardson Lucy	Cong 2017 [5]
77 Hz	Constrained	Richardson Lucy	Taylor 2018 [37]
10 Hz	Constrained	No Deconvolution Process	Wang 2019 [39]
1 Hz	Constrained	Richardson Lucy	Truong 2020 [38]
6 Hz	Free moving	Richardson Lucy	Zhang 2020 [48]

zebrafish brain imaging. However, the spatial resolvability still can not handle cellular-level reconstruction for all size of the zebrafish neurons($3 \times 3 \times 5\mu\text{m}^3$)[19]. Also, the rapid stage moving method adopted by [5] might disturb the zebrafish behavior and cause difference to the neuron activity recorded in the nature mode. Thus, from the hardware

side, even though much progress has been made, there is still no ideal LFM solution for zebrafish brain imaging task.

By Table 2.2, obviously we can see conventional Richardson-Lucy algorithm is the only deconvolution method that applied in the backward model. Some of the works try to avoid using deconvolution due to its computational burden and inevitable inaccuracy, but they still need to face the challenges from neuron identification and lack of global volume information. There is a strong motivation to apply current more advanced deconvolution algorithms in this zebrafish brain imaging task to improve the performance, which is another clear goal our framework is heading for.

CHAPTER 3

Forward Model

The forward model for the light-field microscopy system is the optical transport process of the object wavefront. To help readers understand the mathematics details behind the LFM optical process, we are going to introduce the widely used optical components and propagator functions in Fourier optics. With such optical functions, the chapter will then demonstrate how they are applied in modeling the conventional and Fourier LFM systems as two examples. After that, we step further into another level of abstraction discussing the potential parameter space of the LFM forward model from the angle of optimization. Finally, we will end up this chapter by trying to provide a bold prediction on the open question: 'What could be the next generation of the forward model for LFM systems?' with the hope that it could provoke more beneficial thoughts on this meaningful question.

3.1. Optical Components and Propagators Modeling

For applications like florescent microscopy, the illumination light of the object is coherent and the scale of the imaging granularity is close to the wavelength of the illumination light. Under such conditions, the ray-optics model is no longer accurate enough to model the imaging process. Thus, wave optics comes to rescue. In wave optics, the emitting light field of the object can be modeled as a 3D complex field which the real part represents its amplitude and the imaginary part represents its phase information. So that the diffraction effect can be modeled by multiplication.

There are two types of situation we need to consider separately when modeling the optical path. The first one is when the wavefront is freely propagating through a medium, the second one is when the wavefront passes through an optical instrument or component.

3.1.1. Mathematical Models for Optical Propagation

Let the complex field that is incident on the transverse $(x, y, 0)$ plane be represented by $U(x, y, 0)$. To calculate the wavefront field distribution after propagating to a specific distance z on its parallel plane (x, y, z) , various numerical methods for diffraction theory have been developed. Here, the framework implements two classical propagation functions: angular spectrum method(ASM) and Fresnel diffraction function. The reason why Fraunhofer diffraction is not implemented in the framework is that the far field assumption will not hold for the microscopy scenario. The ASM is a general transfer function for the diffraction theory, it has more accuracy but requires relatively more computation. The Fresnel diffraction makes Fresnel approximation under paraxial assumption to make the computation much faster.

To find an appropriate propagation method to use, the framework will calculate the Fresnel number first to provide a suggestion. The Fresnel number can be calculated by

$$(3.1) \quad F = \frac{a^2}{\lambda L}$$

where a is the radius of the aperture, λ is the wavelength of the light, and L is the propagation distance. The Fresnel number calculates the number of half-period zones in the amplitude. The less number of fringes, the larger propagation distance it will indicate.

For near-field optics application which the Fresnel number is far more than 1, the ASM method needs to be adopted. For Fresnel around 1 which suggest the distance is bigger than the aperture, the Fresnel diffraction should be adopted. For far-field applications which Fresnel number is far less than 1, the Fraunhofer diffraction can be used to save computation time. Even though it might not hold for the microscopy cases.

We will illustrate this selecting process by two LFM simulation examples. Suppose the radius of the camera sensor is 6.7 mm, the propagation distance to the camera is 26 mm and the wavelength is 488nm. So the Fresnel number is about 3538. In that case, we should choose ASM method. This is a common set up for the post-microlens light field propagating to the image sensor in MLA system. In another example for simulating the light enter the microlens array. The microlens radius is 0.7mm, the propagation distance from the rear lens of the 4F system to the microlens array is 160mm. The Fresnel number for such setup is 6.27. Then, using Fresnel diffraction to get more computational benefits would be a reasonable choice.

3.1.1.1. Angular Spectrum Method. In the Fourier optics, the two-dimensional Fourier transform of the complex field can be considered as the sum of a collection of a more simple complex-exponential functions as Equation 3.2.

$$(3.2) \quad U(x, y, 0) = \iint_{-\infty}^{\infty} A(f_X, f_Y; 0) \exp[j2\pi(f_X x + f_Y y)] dx dy$$

To give a physical meaning to the above complex field, we can first represent a plane wave propagating with wave vector $\vec{k} = \frac{2\pi}{\lambda}(\alpha\hat{x} + \beta\hat{y} + \gamma\hat{z})$ in the coordinate system $\{\hat{x}, \hat{y}, \hat{z}\}$ at position $r = x\hat{x} + y\hat{y} + z\hat{z}$. Where the direction cosines are interrelated through

$$(3.3) \quad \gamma = \sqrt{1 - \alpha^2 - \beta^2}$$

Thus, the Fourier decompositon in Equation 3.2 can be considered as the plane-wave component with spatial frequencies (f_X, f_Y) propagating with diffraction cosines $\alpha = \lambda f_X, \beta = \lambda f_Y$ and $\gamma = \sqrt{1 - (\lambda f_X)^2 - (\lambda f_Y)^2}$. with amplitude $A(f_X, f_Y; 0) dx dy$ evaluated at $(f_X = \alpha/\lambda, f_Y = \beta/\lambda)$. Due to this reason, we define the angular spectrum of the wavefield $U(x, y, 0)$ as

$$(3.4) \quad A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}; 0\right) = \iint_{-\infty}^{\infty} U(x, y, 0) \exp[-j2\pi(\frac{\alpha}{\lambda}x + \frac{\beta}{\lambda}y)] dx dy$$

By considering the complex wavefront as a weighted sum of the planar waves propagating to different direction, we can easily derive out the wavefront distribution after the propagation z by adding up the decomposed planar wave propagation results since the optical system is linear. Since it is a linear system, we can derive the linear kernel for the propagation as Equation 3.5. We refer readers the details of kernel derivation to [8].

$$(3.5) \quad H(f_X, f_Y; z) = \exp[j2\pi \frac{z}{\lambda} \sqrt{1 - (\lambda f_X)^2 - (\lambda f_Y)^2}]$$

To use such kernel, we can leverage the fact that the convolution operation in the spatial domain is equal to the multiplication in the frequency domain. Thus, the final wavefront field can be written as

$$(3.6) \quad U(x, y, z) = \mathcal{F}^{-1}\{\mathcal{F}\{U(x_0, y_0)\}[f_X, f_Y]H(f_X, f_Y; z)\}[x, y]$$

3.1.1.2. Fresnel Diffraction. The Huygens-Fresnel principle states that the wavefront can be treated as the sum of the independent spherical wavelets that generated from all the points on that wavefront. This physical can be mathematically expressed by Equation 3.7 where θ is the angle between the outward normal of the wavefront plane and the r_{01} is the vector from the source point P_0 and the target point P_1 .

$$(3.7) \quad U(P_0) = \frac{1}{j\lambda} \iint_{\Sigma} U(P_1) \frac{\exp(jkr_{01})}{r_{01}} \cos\theta \, ds$$

Replace the $\cos\theta$ with the $\frac{z}{r_{01}}$ and redefine our target plane as (ζ, η) , we can get Equation 3.8.

$$(3.8) \quad U(x, y) = \frac{z}{j\lambda} \iint_{\Sigma} U(\zeta, \eta) \frac{\exp(jkr_{01})}{r_{01}^2} \, d\zeta \, d\eta$$

Notice that, the relationship between r_{01} and the points coordinates on the initial plane (x, y) and target plane (ζ, η) is

$$(3.9) \quad r_{01} = \sqrt{z^2 + (x - \zeta)^2 + (y - \eta)^2}$$

Equation 3.9 is the crucial part we further apply Fresnel approximation. Based on the approximation of the binomial expansion of the square root function, we can get rid of the square root by adding more binomial terms if b is less than 1.

$$(3.10) \quad \sqrt{1+b} = 1 + \frac{1}{2}b + \frac{1}{8}b^2 + \dots$$

Notice that, the r_{01} in Equation 3.8 appears two times. First is in the exponent and the second is in the denominator. The approximate error in exponential term will introduce large bias compared with the error in the denominator term. Thus, we will finally derive the Fresnel diffraction as Equation 3.11

$$(3.11) \quad U(x, y, z) = \frac{e^{jkz}}{j\lambda z} \iint_{\Sigma} U(\zeta, \eta) \exp\left\{j \frac{k}{2z} [(x - \zeta)^2 + (y - \eta)^2]\right\} d\zeta d\eta$$

The Equation 3.11 can be written as a convolution form due to its spatial invariance property. Thus, it can be rewritten in frequency domain and we will use the following equation in our framework for efficiency.

$$(3.12) \quad U(x, y, z) = 2\pi \exp\left[j \frac{k}{2z} (x^2 + y^2)\right] \mathcal{F}\{U(\zeta, \eta, 0) H_F(\zeta, \eta)\} \left[k \frac{x}{z}, k \frac{y}{z}\right]$$

where the Fresnel kernel H_F can be get from applying Fourier transform

$$H(x, y) = \mathcal{F}\left\{-\frac{jk}{2\pi z} \exp(jkz) \exp\left[\frac{jk}{2z}(x^2 + y^2)\right]\right\}$$

$$H(x, y) = e^{jkz} \exp[-j\pi\lambda z(f_X^2 + f_Y^2)]$$

By such convolution kernel, we can see that the Fresnel diffraction kernel is applying two parts of phase delay to the incident wavefront: a constant phase delay $\exp(jkz)$ and a quadratic phase delay $\exp[\frac{jk}{2z}(x^2 + y^2)]$ based on the direction vector. Which is coherent with the performance of the linear system. The computation of the Fresnel diffraction is faster than the ASM method since it requires only one Fourier transform operation while the latter one requires two.

Compared H_F with the Equation 3.5, we can find the H_F is actually an approximation to the ASM kernel under the condition that $|\lambda f_X| \ll 1, |\lambda f_Y| \ll 1$. By further connecting the minimum diffraction angle with the spatial frequency, we can derive that Fresnel diffraction must hold under the condition that

$$(3.13) \quad \frac{\theta^4 z}{4\lambda} \ll 1$$

Thus, the Fresnel approximation requires the diffraction angle to be small enough which is similar with the paraxial condition. The framework thus implemented the warning message once the condition will not hold. Especially for the light field microscopy application due to the large NA of the lens, the Fresnel diffraction sometimes might not work well enough.

3.1.2. Mathematical Models for Optical Components

The common optical components in simulation for light field microscopy includes objective lens, thin lens, Fourier lens, microlens array, spatial light modulator(SLM), and camera sensor. The components manipulate the wavefront to further encode the volumetric information on the 2D image sampling plane.

3.1.2.1. Objective Lens. The objective lens usually is designed to have large numerical aperture (NA) to achieve high resolving power based on the Ernst Abbe formula $d = \frac{0.61\lambda}{NA}$. In the physical world, objective lens can magnify the wavefront in a short distance fitted in the microscope. However, in simulation, since all the information of the volume has already known by us. The intended over-sampling won't bring us more information about the volume but more computation burden. The common approach in simulating the microscopy system is to convert the magnification rate later in the image plane. But the modeling of the objective lens is still a good approach to simulate the apodization, depolarization and spherical aberration effects on the PSF brought by the objective lens.[9] By Debye integral for the diffraction field, we can get the following in-differentiable Equation 3.14 for simulating the PSF after a large NA lens:

$$(3.14) \quad U(v, u) = \frac{2\pi j}{\lambda} \exp(-jkz) \int_0^\alpha P(\theta) J_0\left(\frac{v \sin \theta}{\sin \alpha}\right) \exp\left(\frac{j u \sin^2(\theta/2)}{2 \sin^2(\alpha/2)}\right) \sin \theta \, d\theta$$

We refer readers the details of Equation 3.14 to references [9, 10].

3.1.2.2. Thin Lens. Many optical components in LFM systems can be modeled as thin lens based on the approximation that the ray will enter and exist at the same coordinate

on both surfaces. For example, the micro-lenslet, the Fourier lens, and the 4F system. The thin lens acts as a phase delay to the incident wavefront by an amount proportional to its corresponding thickness. The phase delay can be modeled by a quadratic approximation to the spherical wave related with the its focal length and coordinates.

$$(3.15) \quad U'_l(x, y) = P(x, y) \exp[-j \frac{k}{2f} (x^2 + y^2)]$$

Due to the rotation nature of the phasor representation, both the amplitude and phase of the thin lens function will change periodically, which leads to a 'lens array' pattern. Thus, the pupil function is necessary for cutting off the uncanceled area. For common thin lens, we define pupil function as $P(x, y) = \text{circ}(R)$.

3.1.2.3. Fourier Lens. When the paraxial approximation hold, the thin lens can physically apply Fourier transform to the input wavefront if the object is located a focal length f distance in front of the lens. The optical Fourier transform can be written as

$$(3.16) \quad \mathcal{F}_{\text{Optical}}(U(u, v)) = \mathcal{F}\{U(u, v)\}[\frac{u}{\lambda f}, \frac{v}{\lambda f}]$$

The reasons why we separate the Fourier lens out as an independent component rather than treat it as a combination of thin lens and a propagation method are

- (1) Reduce redundant computation on the canceled phase term
- (2) Avoid the inconsistency of the different approximation applied on the propagator and the thin lens

- (3) Extend the function to handle the object located in different distance in front of the Fourier lens rather than fixed on the front focal plane. Which is useful when simulating the PSFs in different z depth for LFM systems.

By combining paraxial approximation from Fresnel kernel and the phase delay function, we can derive the object wavefront $U(u, v)$ that located in a distance d in front of the lens will generate a wavefront distribution $U_f(u, v)$ on the rear focal plane as Equation 3.17 predicts.

$$(3.17) \quad U_f(u, v) = \frac{\exp[j\frac{k}{2f}(1 - \frac{d}{f})(u^2 + v^2)]}{j\lambda f} \mathcal{F}_{Optical}(U(u, v))$$

We can see that when the distance $d = f$, the quadratic phase term brought by the thin lens will disappear and only left the optical Fourier transform operation.

3.1.2.4. Microlens Array. The purpose of applying microlens array in LFM systems is to sample the light field information of the incident wavefront. The phase mask of the traditional microlens array is the non-overlapping ensemble of the thin lens mask. For the sake of manufacturing, all the microlens symmetrically distributed on the plane with same focal length and aperture size. To model such microlens phase mask, we often use a 2D comb function convolving with a thin lens mask as the following equation. Here, $*$ represents the 2D convolution.

$$(3.18) \quad \phi(x, y) = \text{comb}(x/d_{ML}, y/d_{ML}) * P(x', y') \exp[-j\frac{k}{2f}(x'^2 + y'^2)]$$

Even though the uni-focal, symmetric microlens distribution have successfully applied in various real-world LFM systems and achieved encouraging results. There is no general law in traditional microlens design to internally uniform all the variants of microlens array specification.(focal depth, microlens aperture size, MLA distribution). Further more, There are some experimental result to challenge its optimum compared with other optical design(eg. multi-focal, random positioned MLA).[22, 25] From the perspective of deep optics, the optical design could be further optimized based on the data from its application. Thus, there should be some undiscovered general design principles for MLA given the distribution of their imaging task data.

To further introduce more varieties in the MLA design, we proposed a grid-based MLA phase mask generation algorithm. It is composed of two sub algorithm: the grid-coordinate generation algorithm2 and the phase mask concatenation algorithm 1. We will further explain the first algorithm in the Chapter 5. We first pre-compute and save all the possible combinations of thin lens functions between predefined focal depths and possible radial distances with max aperture size in a dictionary \mathcal{D} . Then the grid-coordinate generation algorithm will generate a list of center coordinates (x_i, y_i) , aperture size p_i , and focal depth f_i for every micro lenslet i . We then query the pre-computed phase mask and apply a crop for required aperture size. Then the final MLA phase mask can get from the concatenation of the queried sub mask information based on the coordinate offset.

The implementation details are stated in Algorithm 1.

By algorithm 1, we can get the microlens phase mask with a set of predefined, easy-to-manufactured microlens let but still have the ability to have a task-data driven optimized

Algorithm 1 MLA Concatenation Algorithm

Input: N Microlens Properties Tuples $T = \{(x,y),p,f\}_{i=1}^N$

Output: Microlens Array Phase Mask PM with the size (K, K)

```

1: import numpy as np
2:  $T \leftarrow$  Grid-Coordinate Generation Algorithm ▷ Run the Algorithm 2
3:  $PM \leftarrow \text{np.zeros}(K,K)$ 
4: for  $f$  in  $f\_list$  do
5:    $\mathcal{D}[f'] = \text{circ}((K/2)*\text{pixel\_size})\exp[-j\frac{k}{2f}(K^2 + K^2)]$  ▷ Precalculate and store the
   thin lens function using Equation 3.15
6: end for
7: for  $t_i$  in  $T$  do
8:    $(x_i, y_i), f_i, p_i \leftarrow t_i$  ▷ Get micro lenslet coordinates, focal length and aperture
   size(in pixel)
9:    $patch_i \leftarrow \text{crop}(\mathcal{D}[f'_i])$  ▷ Get cropped microlens phase mask
10:   $PM[x_i - (p_i/2) : x_i + (p_i/2), y_i - (p_i/2) : y_i + (p_i/2)] \leftarrow patch_i$  ▷ Fill in the phase
   mask by different phase mask patches
11: end for

```

design. The algorithm is not differentiable so it has to be used with the non-differentiable optimization method in Chapter 5.

3.1.2.5. Spatial Light Modulator. The biggest difference between the spatial light modulator (SLM) and a static diffuser phase mask is that the SLM can modulate the phase information dynamically. In the initial version of the framework, we haven't simulated the real-world physical defects when using SLM (eg. High Direct Current (DC) term, cross-talk .etc). We basically treat the SLM component as an ideal dynamic phase mask. Also, users can use this component as a to-be optimized diffuser phase mask in differentiable optimization pipeline. Suppose the phase mask is $\phi(x, y)_{SLM}$, the wavefront $U(x, y)$ after passing through is simply applied by a phase delay:

$$(3.19) \quad U(x, y) = U'(x, y) \times \phi(x, y)_{SLM}$$

Another well-designed differentiable diffuser phase mask optimization method can be found in reference [45]. The optimization is defined for maximize the dissimilarity of the axial PSFs and the lateral resolution for all the depth. The building blocks of the phase mask is the overlapping convex sphere surface. This approach can also be implemented using the framework except the optimization only happens in the forward model part.

3.1.2.6. Camera Sensor. The photodetectors only respond to the power of the incident light. The phase information of the wavefront will be lost during sampling. Thus, the detected image for the wavefront $U(x, y)$ can be modeled as:

$$(3.20) \quad I(x, y) = |U(x, y)|^2$$

In the real world, the camera might have many types of noise due to its physical properties, such as shot noise and digital noise. We make the noise part optional so that users can flexibly switch the noise term for different modeling purpose.

3.2. Forward Models for Conventional and Fourier LFM Systems

With the optics simulation toolbox we introduced above, we can easily simulate the forward models for most of the current light field microscopy systems. We here will demo our framework ability to simulate the PSFs for two most common LFM designs. Notice that, if users take magnification convert method to handle objective lens, all the forward models are built both on CPU Numpy library and GPU Pytorch library for both future differentiable and non-differentiable optimization usages. If users choose to use Debye

theory to model the large NA lens effect, the framework currently only support Numpy implementation due to the Debye integral operation.

Another important point to understand before we dig into the details of the forward models simulation is that the difference of the PSF measurement space between the real-world experiments and the computational simulation. In real-world physical experiment, we can't easily manufacture the stage that could shift the object with micrometer-level precision. People then leverage the fact that the image space is the magnification rate square scaling of the object space. Thus the shift distance precision requirement in image space is also loosed by scaled up to the power of two of the magnification rate of the systems. We can then shift the image sensor to different location to capture the PSF images and convert the results to the object space. However, in simulation, we don't often model the objective lens magnification effect due to the reason stated in section 3.1.2.1 and the micrometer-level precision is not a problem for simulation. We tend to simulate our PSF in object space directly.

3.2.1. Forward Model for Conventional Light Field Microscopy

The origin of the light field microscopy comes from the light field camera. [1, 27]. Levoy et al. in 2006 extended it to the microscopy area[20]. In the original setting, they put the microlens array on the native image plane. The purpose of using microlens array is to leverage their microlens curvature to physically decompose the object wavefront and then record their intensity on the 2D camera plane. So the image plane will record the aperture image of the microlens array which is the wavefront in different perspective. To form a

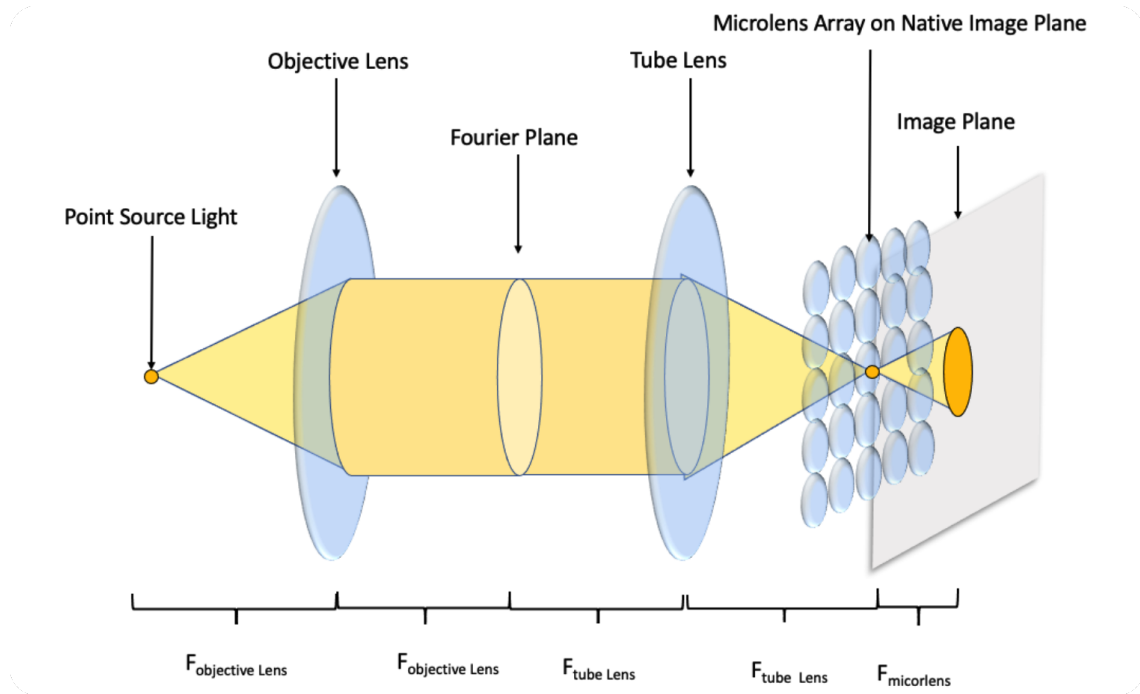


Figure 3.1. The Optical Setting for Conventional Light Field Microscope

specific perspective image, one can sample the specific corresponding location pixels from all microlens aperture image.

As the Figure 3.1 The optical transport process of such conventional LFMs can be described below. For easier understanding, we now first consider the object is located on the native object plane case:

- (1) Incident wavefront generate from the native object plane
- (2) The wavefront then enters a $4F$ system formed by a pair of objective lens and tube lens. It then focused on the native image plane.
- (3) On the native image plane, the wavefront is modulated by the microlens array.
- (4) The wavefront then propagates a distance of the focal length of the micro lenslet and focused on the image plane

To model such optical process using our framework, we defines its forward model as Equation 3.21

(3.21)

$$I(x, y) = F_{camera}(F_{ASM}((F_{\mathcal{F}lens}(F_{\mathcal{F}lens}(U(x, y), f_{obj_lens}), f_{tube_lens})) \times (\Phi_{MLA}), f_{MLA}))$$

We then implement Equation 3.21 by our framework and adopt the parameters from [3]. The result image shows good agreement with the paper's physical experiment.

We further consider the cases where the object is not located on the native object plane, which is the requisite setup for measuring point spread function stack for the LFM systems. In such cases, the object needs to shift both laterally and axially to form the system respond image for the pulse input if the LFM system is not lateral shift-invariant. Unfortunately, the conventional LFM systems fall into that category which means we need to simulate the PSF for all possible object voxel locations. Notice that in real-world experiment, the shift is happening not on the object space but image space due to the reasons we stated at the beginning. To simulate this effect in our framework, we can simply add up the offset value $-DoF/2 * Mag^2 \leq z_{offset} \leq DoF/2 * Mag^2$ to the propagation distance argument of the first Fourier lens function. Again, we can see the result is in good agreement with the simulation result provided in [3].

(3.22)

$$I(x, y) = F_{camera}(F_{ASM}((F_{\mathcal{F}lens}(F_{\mathcal{F}lens}(U(x, y), f_{obj_lens}, z_{offset}), f_{tube_lens})) \times (\Phi_{MLA}), f_{MLA}))$$

3.2.2. Forward Model for Fourier Light Field Microscopy

As the previous chapter introduced, the Fourier light field microscopy put the microlens array on the Fourier plane. The motivation behinds this is the Fourier plane encodes the spatial information of the object wavefront parallelly. Thus, its angular information is inherently encoded into the spatial domain of the Fourier plane. If we put the microlens array on the Fourier plane and record the focused image behind it, all the redundant information from the same source point will collapse together to a single point and thus perform a perfect dealiasing. It leads to a high lateral resolution design. Recall the native-object-plane scenario in the conventional LFM system, each micro lenslet on the microlens array actually spreads out the pre-focused redundant light that comes from the same source point which cause an unneeded waste for the limited spatial encoding space. From that point of view, the Fourier LFM actually surpasses the conventional LFM design.

The optical transport process of Fourier LFMs can be summarized by following steps:

- (1) Incident wavefront generate from the native object plane
- (2) The wavefront then enters a 4F system formed by a pair of objective lens and tube lens. It then focused on the native image plane.
- (3) The wavefront propagates a distance of the focal length to the Fourier lens and then propagate another focal length to apply an optical Fourier transform.
- (4) The wavefront on the Fourier plane is then modulated by the microlens array.
- (5) The wavefront then propagate to a distance of its focal length to form a focused image.
- (6) The camera records the intensity of the wavefront.

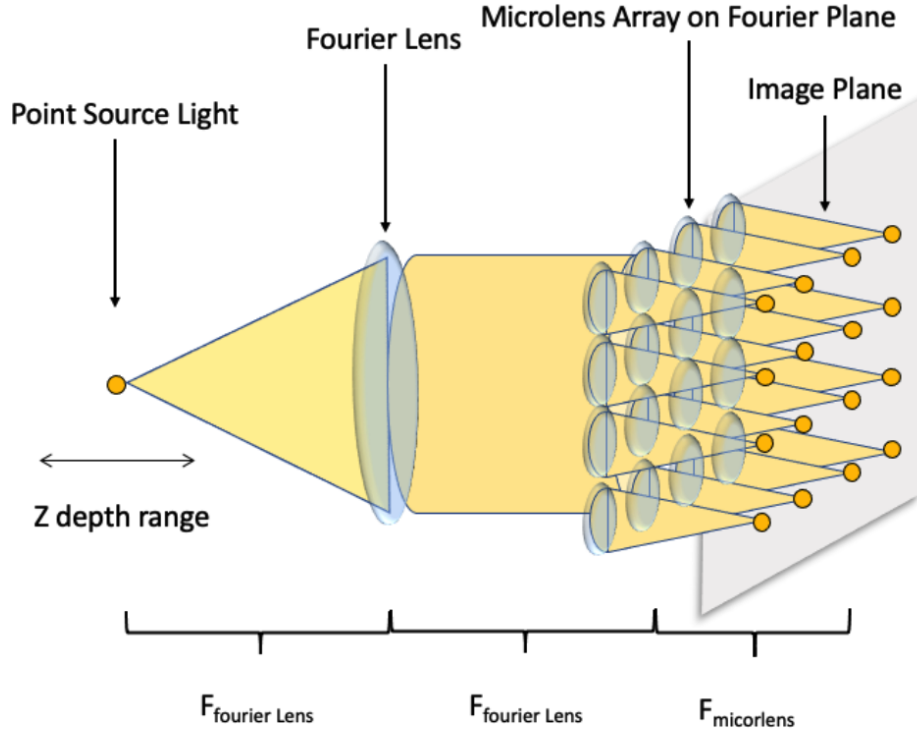


Figure 3.2. The Simplified Optical Setting for Fourier LFM Systems.

For simulation, the 4F system step is unnecessary due to no phase mask is applied in the middle. Thus, we can ignore the step two for efficiency. However, we still need to convert the magnification rate when do axial PSF measurement since the revised optical path is after the 4F system. Figure 3.2 shows the revised optical path. The final optical forward model can be implemented by the following Equation 3.23

$$(3.23) \quad I(x, y) = F_{camera}(F_{ASM}(F_{FLens}(U(x, y), f_{FLens}) \times \Phi_{MLA}, f_{MLA}))$$

Due to the parallel encoding property brought by the Fourier lens setting, the PSF for Fourier LFM is laterally shift-invariant. To encode the depth information, the PSFs are definitely axially shift-variant. This could save huge computational effort for measuring PSFs in all lateral position, which will typically bring at least two to four orders of magnitude acceleration. The process of implementing Fourier LFM forward model in our framework is similar as its counterpart. We only need to pass another offset $-DoF/2 * Mag^2 \leq z \leq DoF/2 * Mag^2$ argument to Fourier lens function indicating its axial shift after magnification.

Again, we compare our simulation with the results from reference [10]. It still shows good agreement.

3.3. Parameter Space of the Forward Model for Optimization

After introducing the optical components and propagators and how to perform optical task simulation by using the combination of them. We step higher to the next level of abstraction to think about the parameter space of the LFM forward models. By understanding what we could change in the forward model of LFM systems, we could found the basis of applying further optimization methods on improving the optical process, which leads to better encoding the volumetric information of the specimen.

The encoder part of the LFM is done by the optical hardware. The wavefront of the emitting light is manipulated by the optical components such as objective lens, tube lens, Fourier lens, microlens array, and spatial light modulator (SLM) and finally formed an image with camera noise on the image plane. Every transformation needs to obey the

physical laws and the freedom researchers can control on the light phase often lies in several aspects:

- The use of the Fourier lens to do optical Fourier transform to the wavefront.
- The microlens distribution, aperture size, and focal length on the microlens array to gather light field information of the wavefront.
- The programmable phase mask on spatial light modulator to add desired phase delay on the wavefront.
- The sequence of the optical components to apply corresponding wavefront functions in different orders.
- The propagation distance of the wavefront to achieve different sampling effect on the following optical component.

3.3.1. Axial Resolution Analysis

To achieve the depth information retrieval from the light field image, one needs to consider the maximum axial resolution the system could achieve. Information goes beyond that threshold basically means there is no way to reconstruct back. That minimum resolvable distance is defined by the Rayleigh criterion which states it as the distance between the center of the emitter and its first diffraction pattern. For the LFM systems, this can be written as $D_{Rayleigh} = \frac{0.61\lambda}{NA_{MLA}}$. Given that theoretical minimum distance, we could further derive out the relationship between the minimum axial shift ΔD with the rest of the system parameters.

For the most micro lenslets that are adopted by the LFM systems, the numerical aperture is usually approximate to 0. We take the micro lenslet used in XLFM paper[5]

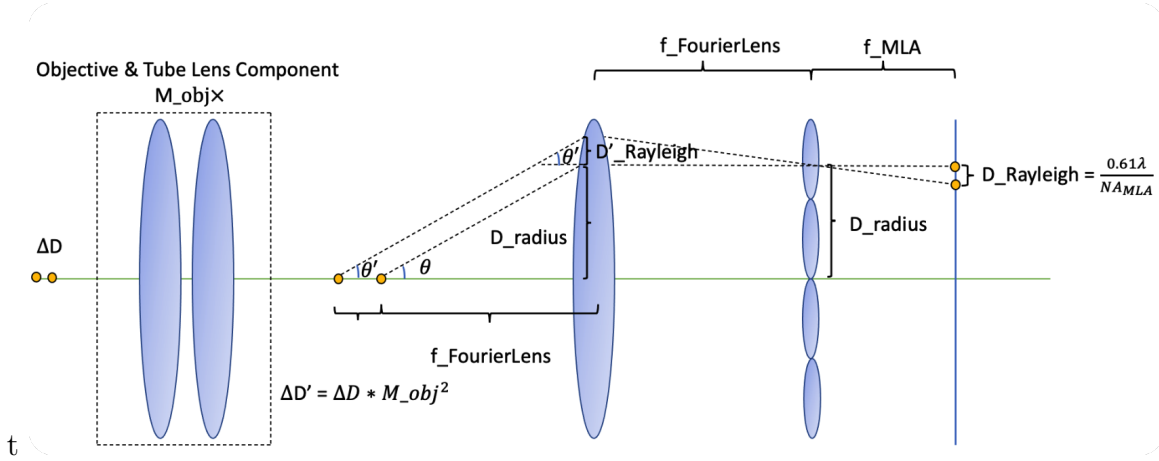


Figure 3.3. The Axial Resolution Analysis of Fourier LFM Systems.

as an example (1.3mm diameter, 26mm focal length), which is a relative larger NA one compared with the other counterparts. The NA of the micro lenslet is about 0.025, where the small angle approximation is perfectly applied since $\sin(0.025) \approx 0.025$. Due to the paraxial approximation, we can write the NA as $NA_{MLA} = \frac{D_{MLA}}{2*f_{MLA}}$, where D_{MLA} refers to the diameter of the micro lenslet. As the Figure 3.3 shows, for LFM systems, the axial shift in the object space causes the lateral shift in the image space. Since we have already known the expression of the minimum resolvable lateral shift distance in the image space based on the Rayleigh criterion, the goal here is to inversely find that corresponding minimum axial shift distance in the object space ΔD that causes this shift $D_{Rayleigh}$, which represents the highest axial resolution we could achieve for such LFM system theoretically.

The axial shift in the object space is magnified by the square of the magnification factor of the objective lens component. Therefore, $\Delta D = \frac{\Delta D'}{M_{obj}^2}$ where $\Delta D'$ is the axial shift in the native image space. Since we are considering the minimum axial shift case, the angle between the incident ray from the native image space and the light axis should

be approximately equal. ($\theta' \cong \theta$) The similar triangles here build the connection between the $\Delta D'$ and $D'_{Rayleigh}$. Besides, the ratio between the $D'_{Rayleigh}$ and $D_{Rayleigh}$ is also equal to the ratio between the Fourier lens focal length and the micro lenslet focal length based on the triangulation. Finally, we can write the final axial resolution formula as

$$\begin{aligned}
 \Delta D &= \frac{D'_{Rayleigh}}{\tan\theta M_{obj}^2} \\
 &= \frac{D_{Rayleigh} f_{Fourier}}{\tan\theta M_{obj}^2 f_{MLA}} \\
 &= \frac{0.61\lambda f_{Fourier}}{N A_{MLA} \tan\theta M_{obj}^2 f_{MLA}} \\
 &= \frac{0.61\lambda f_{Fourier}^2}{N A_{MLA} D_{radial} M_{obj}^2 f_{MLA}}
 \end{aligned}$$

where D_{radial} is the radial distance between the center of the micro lenslet and the chief axis. By the paraxial approximation proofed above, we replace the $N A_{MLA}$ as $\frac{D_{MLA}}{2*f_{MLA}}$ and finally reduce the equation as

$$(3.24) \quad \Delta D = \frac{1.22\lambda f_{Fourier}^2}{D_{MLA} D_{radial} M_{obj}^2}$$

where the D_{MLA} is the diameter or the aperture size of the micro lenslet.

Notice that, the final equation provides a quantitative relationship between the highest axial resolution that could be achieved and the LFM system parameters. We could find that the axial spatial resolution (the inverse of the ΔD) is proportional to the aperture

size of the micro lenslet, the radial distance of the micro lenslet, and the square of the magnification rate of the objective lens. Increasing such parameters could help increasing the axial resolution as long as they are still applicable to the paraxial assumption. Meanwhile, the square of the Fourier lens focal length and the wavelength of the incident light have a counter effect on the axial resolution improvement.

A quite counter-intuitive part of the final expression is that it clearly shows the focal length of the micro lenslet has no effect on the theoretical axial resolution upper bound. However, many recent works on the LFM optimization leverage the parameter space of the focal length of the MLA and have successfully achieved better reconstruction result[**22**, **25**, **45**]. The reasons behind it might lie in several aspects:

- The Rayleigh criterion describes the upper bound of the axial resolution encoded on the image plane. However, for LFM systems, the final axial resolution for the reconstructed volume depends on the output dimension of the deconvolution algorithm, specifically, the axial step size of point spread function stack. These two numbers have no bond since the axial dimension of the PSF stack only depends on the step size of the axial shift to the image plane during the PSF measurement. Theoretically, we could have far finer PSF axial step size than this Rayleigh upper bound and still get that finer axial resolution results from the deconvolution algorithm. But we need to understand that the information goes beyond that limits is the garbage information since it is purely depends on the maximum probability estimation result of the algorithm rather than the physical information itself. We need to be really cautious about the 'false boom' brought by our deconvolution algorithm.

- The axial resolution of the reconstructed volume might go below this upper limit. The no difference for upper bound doesn't necessary means no effect on that two-dimensional image representation of the volume. The focal length of the MLA do affect the light field image result and there does exist a parameter space for searching out the most suitable light field representation for the deconvolution algorithm.
- The Rayleigh criterion based on the assumption that there is no difference for the diffraction patterns of the two close emitters. In many sparse encoding applications, the footprints for different depth of the PSF are different and it could help the algorithm to decode out finer depth information. However, we also need to be cautious about the actual information flow in our deconvolution algorithm, a sanity check on whether the algorithm indeed relies on the footprint distribution information is necessary.

CHAPTER 4

Backward Model

The purpose of the LFM backward models is to decode out the volumetric information of the target sample x from the forward-model encoded 2D image on the camera sensor y . The task for backward models is to solve an inverse problem to reconstruct back the input object.

$$(4.1) \quad \hat{x} = A^{-1}y$$

Even though the optical simulation process for the forward model involved with optics domain knowledge. Like what we have introduced in Chapter 2, once all the point spread function(PSF) functions of the LFM systems are measured, we can directly predict all the possibilities of the system responses without any optical simulation process but pure from the PSFs information. This is due to the linearity of the optical system. Thus, we can store all the measured PSFs and simply apply convolution between these PSFs with any transparent target volume, then we will be able to predict the 2D encoded image of that volume on the camera sensor plane.

$$(4.2) \quad I_{sens} = O ** PSF_{lens}$$

To further reduce the computation cost brought by the spatial domain convolution, we can further leverage the fact that the convolution in spatial domain is the multiplication

in frequency domain and rewrite the Equation 4.2 to

$$(4.3) \quad \mathcal{F}(I_{sens}) = \mathcal{F}(O) \times \mathcal{F}(PSF_{lens})$$

In an ideal case without any noise, one can easily reconstruct back the target O by dividing the 2D image with the PSF in frequency domain and apply an inverse Fourier transform to get the object information.

$$(4.4) \quad O = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(I_{sens})}{\mathcal{F}(PSF_{lens})}\right)$$

However, our physical world is not a noise-free environment and our PSFs for LFM is a 3D version. We must use some iterative or training-based end-to-end methods to deconvolve the camera image. We here categorize and introduce two groups of deconvolution methods based on their differentiability. The motivation behind it is that the difference here will lead us to two directions for our future optimization problem. The differentiable methods here could make the closed-loop LFM optimization process possible.

4.1. Non-differentiable Deconvolution Methods

4.1.1. Wiener Filter Deconvolution

Since the failure reason of the naive deconvolution approach is due to the noise, one can naturely come up with the idea of separating the noise from the signal. Wiener developed a deconvolution algorithm based on minimizing the mean square error between the original signal and the estimated signal. To further reduce the expectation term during derivation, Wiener made a reasonable assumption on the independence between the original signal

and the noise. [44] Thus, leads to the following result:

$$(4.5) \quad O = \mathcal{F}^{-1}\left(\frac{1}{1 + \frac{1}{|\mathcal{F}(PSF_{lens})|^2 SNR}} \frac{\mathcal{F}(I_{sens})}{\mathcal{F}(PSF_{lens})}\right)$$

We can interpret the $|\mathcal{F}(PSF_{lens})|^2 SNR$ term as the spectral density ratio of original signal to noise. Then, the Wiener filter is an adjustable filter based on the signal-to-noise ratio. If the SNR is high at a certain frequency, then the deconvolution will almost behave as a pure frequency division; if the SNR is low, then the filter will assign less weight to the frequency division.

Notice that the Wiener filter is originally designed for 2D data. Even though there are some works for applying Wiener filter in 3D data, the results are still not robust enough to support widely applications. Also, the SNR parameter requires us to know the information about the signal and noise spectrum previously which is impractical. In reality, we just estimate a rough SNR value based on the prior knowledge and trials. The reason we still introduce this method here is not only limited in its huge influence on the future methods, but also its potential in being applied in the recent physics-informed machine learning methods to introduce more interpretability to the neural network based algorithms. [46, 16]

4.1.2. Richardson-Lucy Deconvolution

4.1.2.1. Original Richardson-Lucy for 2D Reconstruction. Richardson-Lucy algorithm is a widely used deconvolution algorithm based on the Poisson statistical nature of the shot noise[33, 24]. It models the image generation process considering the shot noise is Equation 4.6, where \mathcal{P} is the Poisson distribution, A is the matrix operation of

the forward model, and b is a constant vector represents background noise.

$$(4.6) \quad y \sim \mathcal{P}(Ax + b)$$

Then we can write the Poisson probability of such measurement y given a specific type of ground truth photon inputs x as:

$$(4.7) \quad p(y|x) = \exp(-(Ax + b)^T \mathbf{1}) \times \exp(\log(Ax + b)^T y) * \prod_{n=1}^N 1/y_n!$$

Our goal is to find such x that could have most probability to generate such measurement. Thus, we need to find the x that maximizing the $p(y|x)$. We can take a log operation to make it easier to derivate. We can ignore the y_n part since it is a constant. Thus the gradient can be represent as:

$$(4.8) \quad \nabla L(x) = A^T \text{diag}(Ax + b)^{-1} y - A^T \mathbf{1}$$

After we rewrite it to an iterative form

$$(4.9) \quad x_{k+1} = \text{diag}(A^T \mathbf{1})^{-1} \text{diag}[A^T \text{diag}(Ax_k + b)^{-1} y] x_k$$

Replacing the A with our convolution algorithm, we can write out the Richardson-Lucy algorithm for 2D deconvolution as Equation 4.10, where PSF^* means a flipped PSF.

$$(4.10) \quad x_{k+1}^{\hat{}} = \hat{x}_k \cdot \left(\frac{y}{\hat{x}_k ** PSF} ** PSF^* \right)$$

4.1.2.2. Richardson-Lucy Deconvolution for 3D Reconstruction. To further extend the Equation 4.10 to 3D, we need to make several assumptions: First, the reconstructed volume is a layer-based model, which means it can be treated as the concatenation of the 2D slices in the axial direction. Second, we assume there is no scattering effect happened inside the volume. Third, we ignore the difference in magnification rate for all individual micro lenslets.

We first initialize the estimated volume \hat{x} covering z depths which matches our measured PSF stack z depths. Suppose our volume has N slices. Then, we convolve the volume slices with their corresponding z -depth PSF. We then get N 2D estimated camera measurements for different volume slices. Since we ignore the scattering effect, we can get our estimated volume image \hat{y} by summing up N estimated camera measurements.

$$(4.11) \quad y^{\hat{i}+1} = \sum_{k=1}^N [\hat{x}^i(z_k) ** PSF(z_k)]$$

After we get our estimated volume image y , we could further applied it to the 2D reconstruction cases. We use the whole volume image with its measured image to provide the update direction for every single slices. We can finally get our reconstruction after the estimated volume is converged.

$$(4.12) \quad \hat{x}^{i+1} = \hat{x}^i \cdot \left(\frac{y}{\hat{y}^{i+1}} ** PSF^* \right)$$

4.2. Differentiable Deconvolution Methods

Differentiable deconvolution methods are group of methods that leverages the differentiability and strong non-linear fitting ability of the deep neural network to reconstruct the 3D target volumes from the optical encoded 2D images. The high-dimensional statistic data distribution patterns learned from the huge training dataset makes it possible to even achieve high-accuracy reconstruction results without knowing the PSFs information, in other words, to perform blind-deconvolution much better than the classical methods. The differentiability is another huge potential for methods in this group since it could enable the 'closed-loop' LFM system design mentioned in Chapter2 on the backward model side, thus, lead to the task-driven optimized design.

Besides these two advantages mentioned above, another benefit compared with classical counterparts is the reconstruction speed. Due to the fast speed forward inference operation of the deep neural network, once the neural network has been trained, the reconstruction speed in the service stage will be three orders of magnitude faster than the iterative classical methods [46]. It makes the real time reconstruction pipeline possible for LFM systems.

However, the pure deep learning methods also suffers from following aspects:

- The reconstruction has good image quality but less physical accuracy, which is crucial for such low-level accuracy-demanding measurement tool.

- Large training dataset required. Sometimes may cause a deadlock for some novel image tasks since there is no data to train the model but getting the training data requires model.
- Generalization is difficult for different imaging samples.
- The training time is long and costs lots of computational resources and memory storage.
- Lack of interpretability of the reconstruction process.

Many new research works are going on to mitigate such shortages. The motivation of doing non-blind deconvolution using deep neural network further fosters the recent progress on the physics-informed machine learning volume reconstruction topics [46], heading to fix the most important accuracy issue.

4.2.1. U-Net Group

Since deep neural network is released, more and more works report their supremacy in image restoration and deconvolution tasks compared with the conventional methods by its better reconstruction accuracy and speed.[40] Although different works apply different improvements to the previous state-of-art deconvolution model and achieve better performance[42, 11, 41], one of the most popular backbone structure for light field volume reconstruction task is U-Net[34].

The U-Net is a convolution neural network that mainly consists of two information paths like a U-shape: a contracting path and an expansive path. As the name suggests, the contracting path takes the original image as input and uses convolution and max-pooling operation to shrink down the tensor size layer by layer. The expansive path, on the

contrary, takes the output of the contracting path and uses convolution and up-sampling operation to expand the tensor back to the original image size and generate output that roughly matches the size of the input image. To better recover the information loss from the contracting step, the tensors in contracting step will be concatenated to the same-level tensors in the expansion step.

The design reason behind the U-Net comes from the nature that the semantics dependencies in image data are distributed in different spatial scope level. Convolution operations could synthesize kernel-sized local spatial information so shrinking down the tensor size by max-pooling could help convolution synthesize larger scope of information. However, the final output size in such image restoration application requires to be roughly same as the original image, thus, the network needs up-sampling or deconvolution operations to recover back the original size. The concatenation is to recover the information loss we stated above.

The common structure for these U-Net based reconstruction methods is to accept the light field raw image as input and manipulate tensors by U-Net operations to finally match the expansion of the volume data and resize it back to form different depth stack. To match the data size mismatch between the light field measurement and reconstructed volume, they usually do up-sampling first. [41] This end-to-end blind-deconvolution approach purely relies on the large training set and data interpolation to find the mapping between the light field measurement and volume data. No physical information from PSF stack is combined into the process, which leads the model harder to generalize and less robust.

4.2.2. Physics-Informed Machine Learning

The idea of physics informed machine learning is quite straightforward: combining physics processed information into the differentiable loop. However, how to technically combine this into the model requires the knowledge from the optimization and the neural network engineering expertise. We will introduce a relatively simple example called MultiWienerNet[46] for microscopy 3D-reconstruction task in this section to demonstrate its potential.

As we introduced before, the U-Net can take a set of 2D images input and decode the implicit volume structure information out based on the shared statics of the training data and the task data. Modifying the U-Net structure might bring performance improvement but won't bring further mechanism-level changes since the raw input data determines its blind-convolution nature. To combine the physics information into the information flow, the only option is to pre-process the input data by some physics models. In our LFM cases, is the point spread function deconvolution process.

The MultiWienerNet tries to apply multiple Wiener filters on the raw input measurement image. Recall that the Wiener filter is used for 2D image reconstruction, it won't directly return a 3D reconstructed volume. The meaning of applying Wiener filters to deconvolve here is not expecting to get the volume slice information but to hope this deconvolution could use the PSFs to decode out some sharp features for its corresponding region due to its spatial variance property. So that it could be easier for the later U-Net decoder to decode out the correct volumetric information from such pre-processed inputs.

Another novel benefit for this setting is it could make the SNR term in Wiener filter differentiable, therefore it can change based on the reconstruction performance rather than relying on heuristic setting.

There are definitely many other ways to involve the physics-model into the differentiable models such as algorithm unrolling[**35, 29, 47**], single step physics based inversion[**18**], and differentiable forward models[**46**]. We refer readers to those references for further details.

CHAPTER 5

Optimization

In this chapter, we finally reach the point to consider the optimization problem given the previous building block knowledge.

We first categorize the pipeline into three groups based on the differentiability. The motivation behind this taxonomy is due to the difference in their optimization methods, there is no general evidence to show the privilege of using one over another. The non-differentiable property only represents the error brought by the reconstruction loss function can't be automatically used for tuning the forward model by the gradient. However, it doesn't necessarily mean we can't optimize our forward model. Also, even though the classical method are iterative methods, the iteration direction is still based on the gradient between the estimation and measurement. Even the reconstruction image quality might not be as good as neural network methods, there is still no clear dominance here considering its guarantee of the physical accuracy.

Then we will separately discuss each type of pipeline and its possible optimization methods could be adopted.

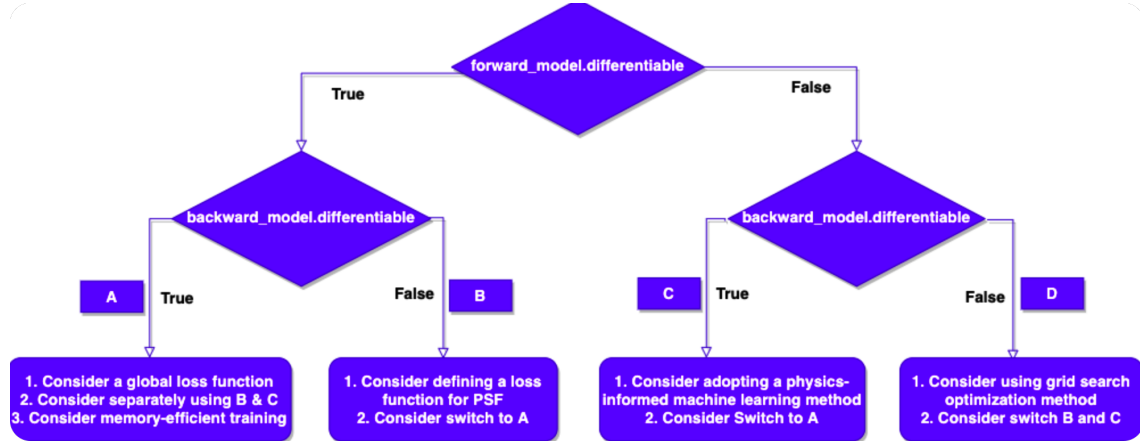


Figure 5.1. The Flow Chart for LFM Optimization Strategies

5.1. Optimization for Non-Differentiable Pipelines

5.1.1. The Scenarios for Non-Differentiable Pipelines

For non-differentiable LFM pipelines, the definition here is that neither the forward model nor the backward model is differentiable. In most cases, using our LFM simulation framework or other modern coherent optics simulation framework eg. HoloTorch (<https://github.com/facebookresearch/holotorch>), since the whole optics operations are developed based on PyTorch, it can guarantee the differentiability and the GPU acceleration of the forward models. However, for most researchers coming from biomedical side, these frameworks haven't got promoted enough to apply and thus leads to many current in-use forward model implementations non-differentiable. The reason why the whole pipeline is non-differentiable includes two scenarios:

- The forward model of the system is implemented based on the non-differentiable library and it adopts the classical reconstruction methods for backward model.[3,

5] The common purpose for such pipeline are usually for computationally prototyping the system or demonstrating the developed non-differentiable forward/backward model, or for heuristically finding the better design.

- Some of the optical components in the forward model includes non-differentiable operations (eg. iteration, non-continuous functions, non-smooth functions) or computational heavy operations that would rather not put in the differentiable loop(eg. integral operation) and the backward model is chosen from the classical deconvolution group. [10] This scenario share the same purpose with the previous case.

5.1.2. The Optimization Strategies for Non-Differentiable Pipelines

Since the pipeline is non-differentiable, the loss function from backward model can't explicitly pass its influence to the forward model. The biggest problem people in this category might concern is how to get a better result for their LFM systems since people adopting this pipeline most come from biomedical engineering or biology background. Their purpose is to build a robust LFM to image their interested specimen. To solve such problem, our framework provides three solutions:

- (1) Apply grid search to the parameters to be tuned. Using the framework provided function `optimizer.grid_search()`. Specifically, we proposed a MLA Grid-Coordinate Generation Algorithm to solve the MLA design problem. We refer readers to algorithm 2 and 1 for further details.

Algorithm 2 MLA Grid-Coordinate Generation Algorithm

Input: The size of MLA plane (K, K) , the diameter of the minimum micro lenslet d_{min} , the focal length list to apply grid search. $focal_list = [f_0, f_1, \dots, f_n]$, the volume dataset V and the backward model function $Reconstruct()$

Output: The optimized micro-lenslet properties list $T = \{(x,y),p,f\}_{i=1}^N$

```

1:  $M \leftarrow K/d_{min}$  ▷ Divide the MLA plane as grid based on the  $d_{min}$ .
2:  $T_{current} \leftarrow (grid, d, f)$  ▷ Assign the initial MLA grid with all minimum lenslets
3:  $PM \leftarrow concat((grid, d, f))$  ▷ Get initial phase mask with all minimum lenslet
4:  $current\_best \leftarrow Reconstruct(PM)$  ▷ Get the performance of initial phase mask
5: for  $d$  in range(1,  $M$ ) do ▷ For all possible aperture size
6:   for  $i$  in range( $M - d + 1$ ) do ▷ For all possible position for such size in X
7:     for  $j$  in range( $M - d + 1$ ) do ▷ For all possible position for such size in Y
8:       if ! is_symmetric and is_conflict(( $i,j$ ),  $d$ , last_state) then
9:          $T_{new} = merge((i, j), d, T_{last})$ 
10:      else
11:         $T_{new} = merge((i, j), d, T)$ 
12:      end if
13:       $current\_best > reconstruct(concat(T_{new}))?$ 
14:      :  $T_{last} = T_{current}, current\_best = reconstruct(concat(T_{new}))$  ▷
      Update the current best value
15:   end for
16: end for
17: end for
18: Return  $current\_best$ 

```

- (2) Pick a off-the-shelf differentiable backward reconstruction method provided in our framework for better image quality but also trying to maintain the physics information . (eg. physics-informed machine learning methods)
- (3) If the optical design could accept SLM or customized diffuser phase mask, consider convert the forward model program to a differentiable framework (eg. our framework, HoloTorch, or reprogram with Pytorch). Then focus on optimizing the PSF only.

5.2. Optimization for Partially Differentiable Pipeline

5.2.1. The Scenarios for Partially-Differentiable Pipelines

As the name suggests, the partially differentiable pipeline means either the forward model or the backward model of the LFM system is differentiable. This is the pipeline that most computational microscopy groups would pick to report their research progress to the academia so that they could focus on a specific sub-problem to solve. Usually, there are two scenarios:

- The forward model is differentiable but the backward model is chosen from the classical method group. Usually, since the backward model can not be integrated into the optimization loop, the optimization goal is to find the best PSF for the forward model, in other words, the PSF engineering task [14, 45]. The backward model here is just for fairly demonstrating and comparing the results.
- The forward model is non-differentiable but the backward model is differentiable. In this case, the purpose of the forward model is for providing the simulated data for the later backward models rather than for tuning the parameters in the forward model. The problem then turns into an optics unrelated computational problem. The goal here is to develop a better reconstruction algorithm for the given forward model. [41, 46]

5.2.2. The Optimization Strategies for Partially-Differentiable Pipelines

The optimization for partially-differentiable pipelines need to consider two different cases based on the differentiable part of their models.

For forward-differentiable but backward-non-differentiable pipelines:

- (1) Since the forward model is differentiable, one can consider directly model the PSF function to make it optimized for some specific defined property. These common approaches including maximizing the dissimilarity for the different z slices (by minimizing the sum of the cross-correlation term), maximizing the lateral resolution (by minimizing the MSE between the mask and the fraction limit lens in frequency domain), and maximizing the axial resolution (by minimizing the minimizing the variance of the Cramér-Rao bound (CRB) of the close point pair in axial direction based on the Fisher information). [45, 14] Notice that, the optimization to such metrics might not necessary means the optimal design for your own data. Good PSF engineering can guarantee the optimum in unbiased estimation, hence add prior knowledge from task data could theoretically achieve better performance.
- (2) Another strategy would be switching the backward model to a differentiable one. This could benefit the forward model by directly optimizing towards the better performance on the task data rather than being optimal to all kind of data distribution. After the switch, the optimization strategy turns to the fully differentiable pipeline cases, which we will discuss later.

For backward-differentiable but forward-non-differentiable pipelines:

- (1) An always workable strategy of optimizing pipeline in such category is trying to turn the blind-deconvolution to a non-blind deconvolution by simply adding a

differentiable physics-informed layer before input into the U-net for reconstruction. The methods of adding physics-informed layer into the neural network has been introduced in Chapter 4.

- (2) If one has already deployed physics-informed machine learning method into the pipeline, the next bottleneck for further optimization would naturally be getting a better encoded representation for the decoder model. Such problem can be solved by adding a differentiable forward model, which also leads to strategies we are gonna discuss in fully differentiable pipeline cases.

5.3. Optimization for Fully Differentiable Pipeline

5.3.1. The Scenario for Fully-Differentiable Pipelines

The fully differentiable pipeline is the truly end-to-end problem solver for LFM tasks. Both the forward and backward model here are designed to be differentiable and only few computational microscopy groups have successfully developed such system by far[**26, 6, 25**]. The differentiability allows the LFM systems to update not only its backward reconstruction neural network, but also the programmable optical components in the forward model (eg. spatial light modulator and customized diffuser mask). The scenario usually is

- Computationally find the optimal design for the imaging task.

5.3.2. The Optimization Strategies for Fully-Differentiable Pipelines

The fully differentiability for the whole pipeline definitely requires expertise in the neural network engineering, optimization and optics simulation knowledge. However, this

engineering is worthy since it leads towards the global optimization for the whole LFM pipeline which is impossible for other pipelines.

However, the cost for putting such whole LFM process into the same optimization chain is the huge computational resources. First, one need to apply gradient descent on the full parameter set in the backward neural network and at least the phase-mask level magnitude parameters for the programmable optical components in the forward model. Second, even the single forward inference time for such model is not trivial, let alone such amount size of parameters requires at least thousands of batches to converge. Both the space and time challenges make this approach really difficult to apply in the real-world experiments.

The strategies for optimizing such pipelines includes:

- (1) Define a better loss function or output for the model to make it easier to converge.
For example, not predict the discrete position pair but a continuous probability map. We refer the readers to the supplementary files of the reference [26]
- (2) Rather than co-optimizing the whole process together, one can consider develop separate loss functions for both parts or freeze some part of the model after several iterations.
- (3) Data and computation parallelization is necessary for such task. For example, one can decompose the 3D volume into 2D slices and separately training networks for different depth on different GPUs.
- (4) In 2020, the memory-efficient methods for large scale image data training was introduced by Kellman et al. We recommend interested readers to the reference [17].

CHAPTER 6

Conclusion & Future Works

This paper first reviewed and compared the light field microscopy works and then proposed the taxonomy for the light field microscopy systems: the open-loop and the closed-loop based on the information flow. Then, the presented work further discuss the building blocks for programming the forward optical model and simulate both conventional LFM and Fourier LFM to further demonstrate the correctness of the framework. After that, both the classical and neural-network based backward models are introduced to close the simulation pipeline. Based on such model units, we then finally discuss the optimization strategies separately for non-differentiable, partially differentiable, and fully differentiable optimization pipelines.

In the future, we are going to

- (1) Further improve the framework by adding more forward optics components, more backward models and developing better framework interface.
- (2) Keep working on the MLA pattern optimization using both grid search and gradient descent methods.
- (3) Explore more about the physics-informed machine learning methods and the optimization for fully differentiable pipeline.
- (4) Build out our hardware setup for zebrafish brain imaging based on the simulation results from the framework.

References

- [1] ADELSON, E., AND WANG, J. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 2 (1992), 99–106.
- [2] BECK, A., AND TEOULLE, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences* 2, 1 (2009), 183–202.
- [3] BROXTON, M., GROSENICK, L., YANG, S., COHEN, N., ANDALMAN, A., DEISEROTH, K., AND LEVOY, M. Wave optics theory and 3-d deconvolution for the light field microscope. *Optics express* 21, 21 (2013), 25418–25439.
- [4] COHEN, N., YANG, S., ANDALMAN, A., BROXTON, M., GROSENICK, L., DEISEROTH, K., HOROWITZ, M., AND LEVOY, M. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express* 22, 20 (Oct 2014), 24817–24839.
- [5] CONG, L., WANG, Z., CHAI, Y., HANG, W., SHANG, C., YANG, W., BAI, L., DU, J., WANG, K., AND WEN, Q. Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (danio rerio). *Elife* 6 (2017).
- [6] DEB, D., JIAO, Z., CHEN, A. B., AHRENS, M. B., PODGORSKI, K., AND TURAGA, S. C. Programmable 3d snapshot microscopy with fourier convolutional networks. *arXiv preprint arXiv:2104.10611* (2021).
- [7] DELL’ACQUA, F., RIZZO, G., SCIFO, P., CLARKE, R. A., SCOTTI, G., AND FAZIO, F. A model-based deconvolution approach to solve fiber crossing in diffusion-weighted mr imaging. *IEEE Transactions on Biomedical Engineering* 54, 3 (2007), 462–472.
- [8] GOODMAN, J. W. Introduction to fourier optics.
- [9] GU, M. *Advanced optical imaging theory*, vol. 75. Springer Science & Business Media, 2000.

- [10] GUO, C., LIU, W., HUA, X., LI, H., AND JIA, S. Fourier light-field microscopy. *Optics express* 27, 18 (2019), 25573–25594.
- [11] GUO, M., LI, Y., SU, Y., LAMBERT, T., NOGARE, D. D., MOYLE, M. W., DUNCAN, L. H., IKEGAMI, R., SANTELLA, A., REY-SUAREZ, I., ET AL. Rapid image deconvolution and multiview fusion for optical microscopy. *Nature biotechnology* 38, 11 (2020), 1337–1346.
- [12] HARMON, T. C., MCLEAN, D. L., AND RAMAN, I. M. Integration of swimming-related synaptic excitation and inhibition by olig2+ eurydendroid neurons in larval zebrafish cerebellum. *Journal of Neuroscience* 40, 15 (2020), 3063–3074.
- [13] HINTON, G. E., AND SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *science* 313, 5786 (2006), 504–507.
- [14] JUSUF, J. M., AND LEW, M. D. Towards optimal point spread function design for resolving point emitters in 3d. *arXiv preprint arXiv:2111.14334* (2021).
- [15] KAK, A. C. Algorithms for reconstruction with nondiffracting sources. *Principles of computerized tomographic imaging* (2001), 49–112.
- [16] KARNIADAKIS, G. E., KEVREKIDIS, I. G., LU, L., PERDIKARIS, P., WANG, S., AND YANG, L. Physics-informed machine learning. *Nature Reviews Physics* 3, 6 (2021), 422–440.
- [17] KELLMAN, M., ZHANG, K., MARKLEY, E., TAMIR, J., BOSTAN, E., LUSTIG, M., AND WALLER, L. Memory-efficient learning for large-scale computational imaging. *IEEE Transactions on Computational Imaging* 6 (2020), 1403–1414.
- [18] KHAN, S. S., SUNDAR, V., BOOMINATHAN, V., VEERARAGHAVAN, A., AND MITRA, K. Flatnet: Towards photorealistic scene reconstruction from lensless measurements. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020).
- [19] KUNST, M., LAURELL, E., MOKAYES, N., KRAMER, A., KUBO, F., FERNANDES, A. M., FÖRSTER, D., DAL MASCHIO, M., AND BAIER, H. A cellular-resolution atlas of the larval zebrafish brain. *Neuron* 103, 1 (2019), 21–38.
- [20] LEVOY, M., NG, R., ADAMS, A., FOOTER, M., AND HOROWITZ, M. Light field microscopy. In *ACM SIGGRAPH 2006 Papers*. 2006, pp. 924–934.
- [21] LIN, A., WITVLIET, D., HERNANDEZ-NUNEZ, L., LINDERMAN, S. W., SAMUEL, A. D., AND VENKATACHALAM, V. Imaging whole-brain activity to understand behaviour. *Nature Reviews Physics* 4, 5 (2022), 292–305.

- [22] LIU, F. L., KUO, G., ANTIPA, N., YANNY, K., AND WALLER, L. Fourier diffuser-scope: single-shot 3d fourier light field microscopy with a diffuser. *Optics Express* 28, 20 (2020), 28969–28986.
- [23] LLAVADOR, A., SOLA-PIKABEA, J., SAAVEDRA, G., JAVIDI, B., AND MARTÍNEZ-CORRAL, M. Resolution improvements in integral microscopy with fourier plane recording. *Optics express* 24, 18 (2016), 20792–20798.
- [24] LUCY, L. B. An iterative technique for the rectification of observed distributions. *The astronomical journal* 79 (1974), 745.
- [25] MARKLEY, E., LIU, F. L., KELLMAN, M., ANTIPA, N., AND WALLER, L. Physics-based learned design for fourier diffuserscope. In *Computational Optical Sensing and Imaging* (2021), Optica Publishing Group, pp. CM6B–3.
- [26] NEHME, E., FREEDMAN, D., GORDON, R., FERDMAN, B., WEISS, L. E., ALALOUF, O., NAOR, T., ORANGE, R., MICHAELI, T., AND SHECHTMAN, Y. Deepstorm3d: dense 3d localization microscopy and psf design by deep learning. *Nature methods* 17, 7 (2020), 734–740.
- [27] NG, R., LEVOY, M., BRÉDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. *Light field photography with a hand-held plenoptic camera*. PhD thesis, Stanford University, 2005.
- [28] NÖBAUER, T., SKOCEK, O., PERNÍA-ANDRADE, A. J., WEILGUNY, L., TRAUB, F. M., MOLODTSOV, M. I., AND VAZIRI, A. Video rate volumetric ca2+ imaging across cortex using seeded iterative demixing (sid) microscopy. *Nature methods* 14, 8 (2017), 811–818.
- [29] ONGIE, G., JALAL, A., METZLER, C. A., BARANIUK, R. G., DIMAKIS, A. G., AND WILLETT, R. Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory* 1, 1 (2020), 39–56.
- [30] PÉGARD, N. C., LIU, H.-Y., ANTIPA, N., GERLOCK, M., ADESNIK, H., AND WALLER, L. Compressive light-field microscopy for 3d neural activity recording. *Optica* 3, 5 (2016), 517–524.
- [31] PEREZ, C. C., LAURI, A., SYMVOULIDIS, P., CAPPETTA, M., ERDMANN, A., AND WESTMEYER, G. G. Calcium neuroimaging in behaving zebrafish larvae using a turn-key light field camera. *Journal of Biomedical Optics* 20, 9 (2015), 096009.

- [32] PREVEDEL, R., YOON, Y.-G., HOFFMANN, M., PAK, N., WETZSTEIN, G., KATO, S., SCHRÖDEL, T., RASKAR, R., ZIMMER, M., BOYDEN, E. S., ET AL. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature methods* 11, 7 (2014), 727–730.
- [33] RICHARDSON, W. H. Bayesian-based iterative method of image restoration. *JoSA* 62, 1 (1972), 55–59.
- [34] RONNEBERGER, O., FISCHER, P., AND BROX, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (2015), Springer, pp. 234–241.
- [35] SHLEZINGER, N., FARSAF, N., ELDAR, Y. C., AND GOLDSMITH, A. J. Viterbinet: A deep learning based viterbi algorithm for symbol detection. *IEEE Transactions on Wireless Communications* 19, 5 (2020), 3319–3331.
- [36] SITZMANN, V., DIAMOND, S., PENG, Y., DUN, X., BOYD, S., HEIDRICH, W., HEIDE, F., AND WETZSTEIN, G. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–13.
- [37] TAYLOR, M. A., NÖBAUER, T., PERNIA-ANDRADE, A., SCHLUMM, F., AND VAZIRI, A. Brain-wide 3d light-field imaging of neuronal activity with speckle-enhanced resolution. *Optica* 5, 4 (2018), 345–353.
- [38] TRUONG, T. V., HOLLAND, D. B., MADAAN, S., ANDREEV, A., KEOMANEE-DIZON, K., TROLL, J. V., KOO, D. E., MCFALL-NGAI, M. J., AND FRASER, S. E. High-contrast, synchronous volumetric imaging with selective volume illumination microscopy. *Communications biology* 3, 1 (2020), 1–8.
- [39] WANG, D., XU, S., PANT, P., REDINGTON, E., SOLTANIAN-ZADEH, S., FARSIU, S., AND GONG, Y. Hybrid light-sheet and light-field microscope for high resolution and large volume neuroimaging. *Biomedical optics express* 10, 12 (2019), 6595–6610.
- [40] WANG, K. Deep-learning-enhanced light-field microscopy. *Nature Methods* 18, 5 (2021), 459–460.
- [41] WANG, Z., ZHU, L., ZHANG, H., LI, G., YI, C., LI, Y., YANG, Y., DING, Y., ZHEN, M., GAO, S., ET AL. Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning. *Nature Methods* 18, 5 (2021), 551–556.

- [42] WEIGERT, M., SCHMIDT, U., BOOTHE, T., MÜLLER, A., DIBROV, A., JAIN, A., WILHELM, B., SCHMIDT, D., BROADDUS, C., CULLEY, S., ET AL. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature methods* 15, 12 (2018), 1090–1097.
- [43] WETZSTEIN, G., OZCAN, A., GIGAN, S., FAN, S., ENGLUND, D., SOLJAČIĆ, M., DENZ, C., MILLER, D. A., AND PSALTIS, D. Inference in artificial intelligence with deep optics and photonics. *Nature* 588, 7836 (2020), 39–47.
- [44] WIENER, N., WIENER, N., MATHEMATICIAN, C., WIENER, N., WIENER, N., AND MATHÉMATICIEN, C. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*, vol. 113. MIT press Cambridge, MA, 1949.
- [45] YANNY, K., ANTIPA, N., LIBERTI, W., DEHAECK, S., MONAKHOVA, K., LIU, F. L., SHEN, K., NG, R., AND WALLER, L. Miniscope3d: optimized single-shot miniature 3d fluorescence microscopy. *Light: Science & Applications* 9, 1 (2020), 1–13.
- [46] YANNY, K., MONAKHOVA, K., SHUAI, R. W., AND WALLER, L. Deep learning for fast spatially varying deconvolution. *Optica* 9, 1 (2022), 96–99.
- [47] ZHANG, J., AND GHANEM, B. Ista-net: Interpretable optimization-inspired deep network for image compressive sensing. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 1828–1837.
- [48] ZHANG, Z., BAI, L., CONG, L., YU, P., ZHANG, T., SHI, W., LI, F., DU, J., AND WANG, K. Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy. *Nature biotechnology* 39, 1 (2021), 74–83.
- [49] ZHANG, Z., CONG, L., BAI, L., AND WANG, K. Light-field microscopy for fast volumetric brain imaging. *Journal of Neuroscience Methods* 352 (2021), 109083.