Mask and Algorithm Co-Design of a Lensless Camera in Multiple Reconstruction Settings

ESE 498 Capstone Design Project Formal Proposal
Submitted to Professor Wang and the Department of Electrical and Systems Engineering
Project Duration: January 17 –

May 08, 2024
Proposal Submitted: February 12, 2024

Client:
Electrical & Systems Engineering Department

Advisors:
Dorothy Wang[1] - dorothyw@wustl.edu
Matthew Lew[2] – mdlew@wustl.edu
Jason Trobaugh[3] – jasont@wustl.edu

Engineers:
Jasmine Cheng[4] – c.jasmine@wustl.edu
Mae Martel[5] - m.mae@wustl.edu

Washington University in St. Louis
McKelvey School of Engineering

[1]Lecturer, Department of Electrical and Systems Engineering
[2]Associate Professor, Department of Electrical and Systems Engineering
[3]Professor in Practice, Department of Electrical and Systems Engineering
[4]Candidate for B.S. in Electrical Engineering

[5]Candidate for B.S. in Computer Engineering with M.S. in Engineering Management

Abstract

Under the advisory of Drs. Dorothy Wang, Matthew Lew, and Jason Trobaugh, this capstone project co-design the hardware and software of a portable, highly customizable lensless camera system for educational purpose that bridge optical imaging with signal processing. A Raspberry Pi with a Pi Camera sensor will be used in conjunction with various cheap diffuser masks made from common materials to capture raw images. These images, which are unrecognizable by humans, will be processed and reconstructed through different deconvolution algorithms implemented Python, such as gradient descent (GD), alternating direction method of multipliers (ADMM), and Adam. The capabilities of this customizable system are broad, including 3D and colored reconstruction of near, distant, static, and dynamic objects. The final goal of this project is to optimize 3D reconstructions of discrete, quickly captured raw footage leveraging the mask and algorithm co-design.

Introduction

Background

In traditional optical imaging, engineers design a system consisting of several focusing lenses to gather light and form images on the sensor which is at the end of the light path. Focused light produces sharp, human-interpretable images that contain details of the scene. The optical system becomes bulky due to the sequence of lenses, which makes it unsuitable for tasks in constrained spaces.

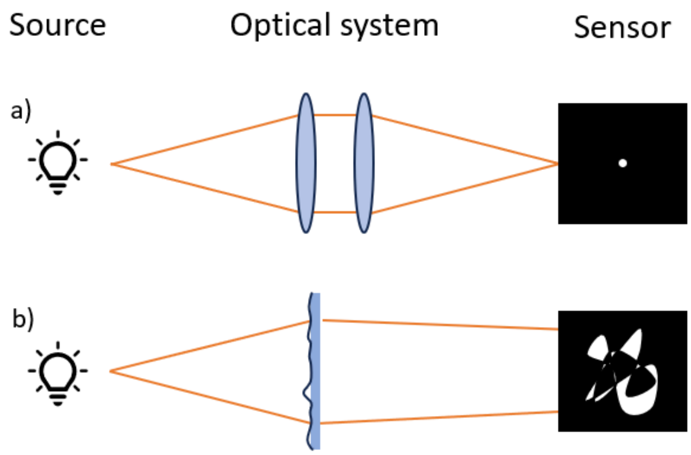

Figure 1. Optical imaging system comparison. a) System with focusing lenses and its PSF. b) Lensless imaging system and its PSF.

In lensless imaging, engineers use a mask, for example a diffuser [1], to encode the object information into a unique pattern on the sensor. Without focusing lenses, the point-spread function (PSF) of a lensless camera usually fills up the whole sensor, and the topology of the PSF is related to the mask in front of the sensor. The application of lensless imaging is broad. [2] In neural circuitry imaging, lensless cameras could enable the simultaneous monitoring of millions of neurons in 3D space at frame rates limited only by image sensor read times due to its minimum size. [1] With single exposure 3D vision, lensless cameras increase potential for gesture recognition and complex fine-motor manipulation in robotics. [3] Furthermore, 3D information gained from lensless camera could provide content for 3D displays used in gaming, entertainment, VR, and many other applications. [4]

For educational purposes in an undergraduate classroom, the lensless camera can be a good demonstration of imaging systems and signal processing. Firstly, due to their portability, lensless cameras are easy to carry and store in the classroom. Secondly, the design is appropriate tad appealing for undergraduate students because of the cheap common materials for building the mask and the highly customizable process of mask and algorithm co-design. Thirdly, the lensless camera is tolerant toward erasures due to space multiplexing: it can reconstruct image with 90% erasure of pixels in measured image and reconstruct images outside of the sensor's field of view, if only one part of the PSF makes it into the measurement. [3]

Problem Statement

Given the space efficiency, time efficiency, cost effectiveness, customizability, and portability of lensless cameras, the goal of this project is to co-design a physical lensless camera system – consisting of a Raspberry Pi, Pi Camera, diffuser mask, a customized mask mount – and a computationally efficient algorithm to reconstruct colored 3D objects. This project will combine hardware and software to provide an optimized solution for the capture, reconstruction, and colorization of near, distant, static, and dynamic subjects using discrete surveillance camera footage.

Aims & Objectives

There are three primary objectives of this project: exploring different masks, image types, and subject depths. This project aims to reconstruct 3D facial features and corporeal forms from near, distant, static and dynamic images captured by the Pi Camera surveillance system, employing various mask materials and evaluating their effectiveness with a quantitative metric.

Methodology

Hardware

The hardware of the lensless camera consists of a Raspberry Pi 4 Model B, a Raspberry Pi Camera Module v1.3 without lens (may switch to a camera sensor with higher resolution, e.g., Raspberry Pi HQ Camera), a customized mount, and a diffuser mask.

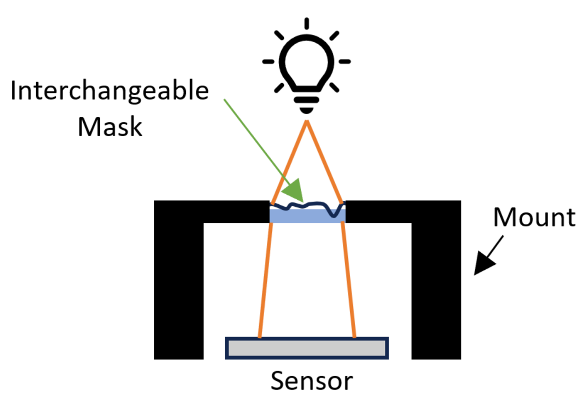

Figure 2. Hardware setup of the lensless camera. The sensor is connected to the Raspberry Pi.

The Raspberry Pi 4 is used to control the Pi Camera Module. The Pi Camera Module v1.3 is compatible with the PiCamera2 package in Python 3, and users can preview, and capture images formed on the sensor. To build a lensless camera, lens that are assembled to the Pi Camera should be removed. A customized mount will be added to the Pi Camera Module for easier implementation of interchangeable masks. The aperture of the mount will constraint the PSF size not to fully cover the camera sensor, thus guaranteeing the shift-invariance property of the PSF. Lastly, the diffuser mask is the key to the hardware design. Different masks will result in different PSF patterns. PSF of the physical system will directly influence the performance of reconstruction. Therefore, it is significant to choose a mask which produces a sharp, highly depth-dependent PSF.

The forward model of optical imaging can also describe the lensless camera and is used in the inverse algorithm. The PSF of the system, denoted as $h_l(\xi, \eta; z)$, $l \in \{r, g, b\}$, is color-related (red, green, blue) and depth-dependent. The 3D object that will be reconstructed is a matrix with signal level value of each color, denoted as $s_l(x, y, z)$. The image of each color formed on the sensor is denoted as $I_l(u, v)$, where the scale of u, v is corresponding to the pixel number of the sensor. The image of each color is the sum of the convolution of the PSF and object along the axial axis. [5]

$$I_l(u, v) = \sum_{z=z_{min}}^{z_{max}} s_l(x, y, z) * h_l(\xi, \eta; z) \# (1)$$

Software

Using Python 3 in VS Code, in conjunction with Anaconda, this project will explore several lensless image reconstruction algorithms which hinge on the core principle that the diffuser system is assumed to be linear shift-invariant. The output $f(v)$ corresponding to the 2D image input v and PSF h can be represented as the following 2D convolution formula.

$$f(v) = h * v$$

Several methods for solving for v include the Gradient Descent iterative algorithm, which solves a minimization problem for a convex function, and has potential for a speedup update modification which adds an additional momentum term, μ . The original minimization problem, which is the typical approach to solving for v for a non-invertible matrix A is as follows.

$$v* = \arg arg \quad \frac{1}{2} \| Av - b\|_2^2$$

In this equation, b

represents the sensor reading, which is a 2D array of pixel values. Other methods for speeding up the gradient descent algorithm include implementing “Nesterov” momentum, which prevents abrupt and frequent changes in descent direction, as well as the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA), which is more useful regarding linear inverse problems. An alternative image reconstruction method to gradient descent is the Alternating Direction Method of Multipliers (ADMM), which is preferable for 3D reconstruction, due to the improved processing capabilities that arise when the original minimization problem is split into two separate equations.

Feasibility Study

Hardware

A simplified demo of the lensless camera’s hardware setup is done, and some preliminary data of the PSF is captured. The Pi Camera Module v1.3 is connected to the Raspberry Pi, and the preview of the scene is monitored via Python 3 code implemented on the board. The lens of the Pi Camera module is removed, and a rough aperture made with opaque black tape is stuck to the module. A plastic bag is used as the diffuser mask. Finally, the sensor is put inside the plastic bag.

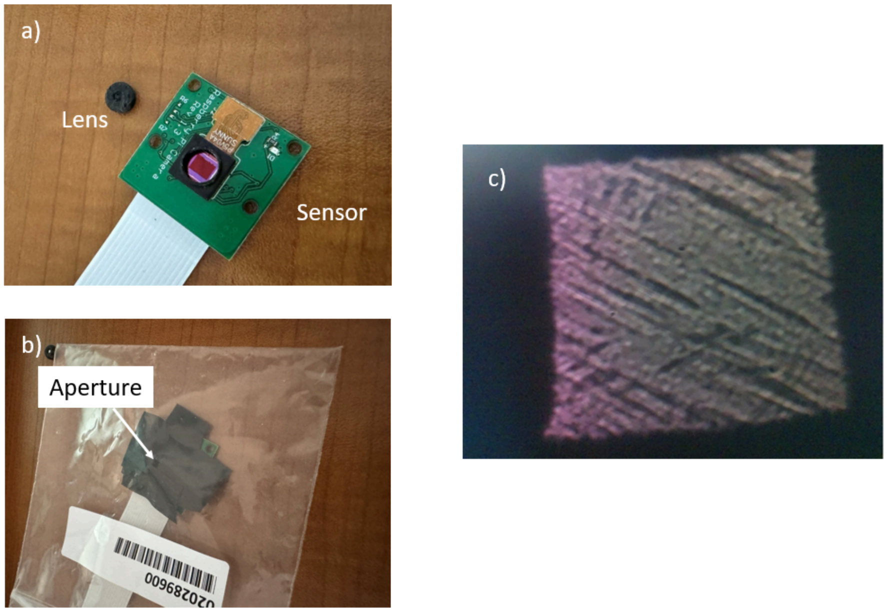

Figure 3. Simplified hardware setup and preliminary data. a) Raspberry Pi Camera Module v1.3 with lens removed. b) Pi Camera with a rough aperture on it. The setup is in a plastic bag serving as the diffuser mask. c) The PSF of the plastic bag.

I use the phone flashlight as the point emitter light source. A rough demonstration of the PSF with a plastic bag as the diffuser mask is shown in Figure 3.c. The imaging environment is not dark enough, and thus there is ambient light captured by the sensor making the PSF brighter as expected. As the phone moves laterally, the PSF shifts laterally without the shape changing which indicates the shift-invariance property. Therefore, a deconvolution algorithm can be applied to recover the target object.

Software

The initial test was of the gradient descent algorithm for 2D image reconstruction, using two sample images, one of the raw data and one of the PSF taken by a lensless camera designed by a team at UC Berkeley [1]. This raw data underwent 100 iterations of the gradient descent algorithm, resulting in a final reconstruction shown below.

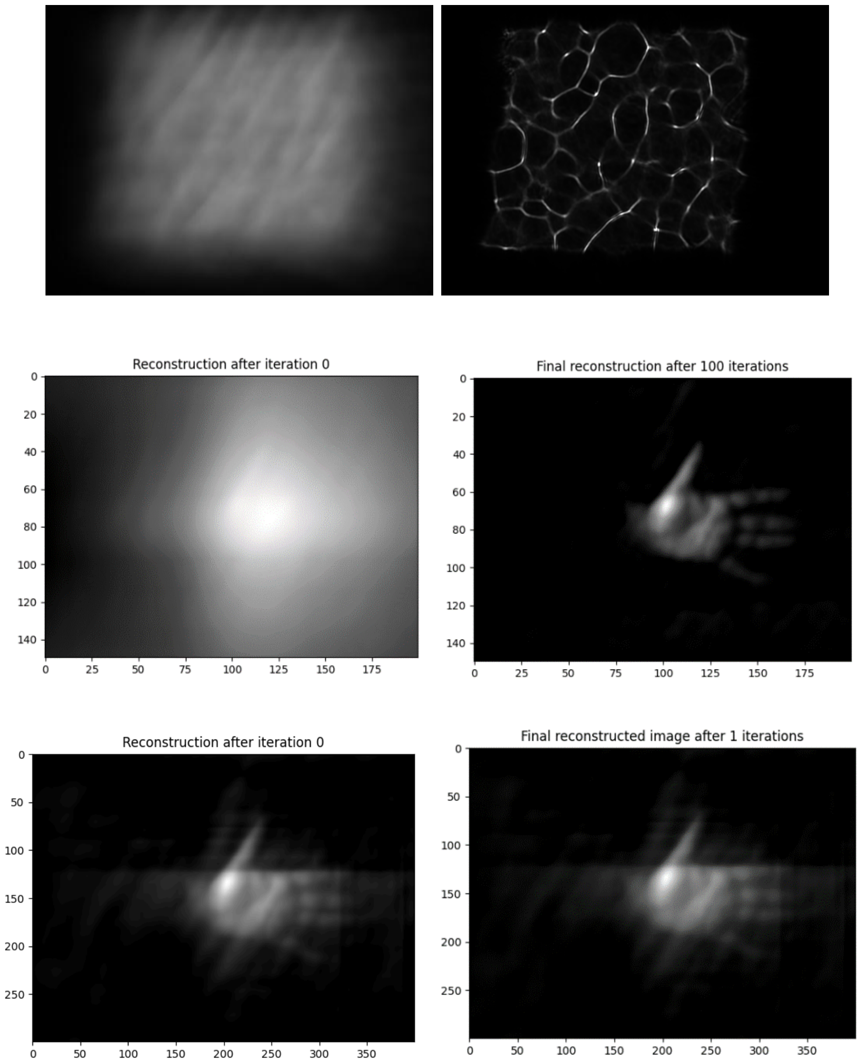

Figure 4. Sample raw data (top left) and sample PSF (top right), iteration 0 (center left) and iteration 100 (center right) of the 2D reconstruction using gradient descent, and iteration 0 (bottom left) and iteration 1 (bottom right) using ADMM [1].

With the successful reconstruction of 2D sample data, this project seeks to reconstruct raw data taken by the Pi Camera and Raspberry Pi lensless camera system. We will implement the known 3D reconstruction algorithm and develop our own algorithm for the addition of RGB values to grayscale images. We will test this algorithm on raw data with varying subject depths – near and distant – as well as static and dynamic images. We will also test various diffuser masks, and analyze their effectiveness based on a quantitative metric for evaluating reconstruction results. Our subjects will be human, with an emphasis on facial reconstruction and physical forms. At the time of submission of this formal proposal, the initial test of the 2D reconstruction algorithm for raw data in the form of a sensor image and PSF is complete.

Data to be Collected and Procedure for Acquiring

Point-Spread Function

The most vital information for solving the inverse problem is the PSF, i.e., the impulse response of the lensless camera. For every mask, its PSF should be captured and stored for later reconstruction, and several other important parameters should be collected for system evaluation.

Parameter	How to collect
PSF matrix	Save the image captured by Pi Camera Sensor with a point light source in focus.
Focal Distance	Record the distance between the light source and the camera, where the PSF has the sharpest feature.
Field of View (FOV)	Find the size of the area in front of the sensor where part of the PSF is still captured by the sensor.

Table 1. Characteristics of the PSF.

To capture the PSF, we need to use a point emitter light source. We need to find the correct depth position where the feature of the image is the sharpest, such that the light source is in focus. Then, we adjust the lateral position and make sure the PSF is located at the center. A picture is captured as the PSF, and the distance between the camera and the light source is recorded.

Raw Image of Different Types

We would like to capture several types of raw images and evaluate the performance of the system in different scenarios. Human faces will be the major target we are interested in since there are many detailed morphologies suitable for later evaluation. We will first take images of static faces that are in focus, which is the best condition that we expect. Then, we will take images of the same faces at different depth positions, which will blur the image. The same algorithm will be applied to reconstruction, and we will evaluate the influence of depth position on the reconstruction performance. What is more, we will image dynamic faces at different speeds, which will also blur the image. The same analysis will be done to evaluate the influence of dynamics on the reconstruction performance.

Test Different Masks and Algorithms

The previous two parts in this section describe the procedure of data acquisition with a specific combination of mask and reconstruction algorithms. There are various types of masks that we can apply, such as scotch tape, plastic bag, and even microlens array. Similarly, there are various types of reconstruction algorithms, including GD, ADMM, and Adam. We will follow the procedure in the previous two parts with different combinations of mask and reconstruction algorithms and compare the performance in different scenarios.

Materials

The project requires a list of materials including:

- Raspberry Pi 4 Model B
- Raspberry Pi Camera Module v1.3 (or Raspberry Pi HQ Camera or other sensors that do not require Raspberry Pi)
- 3D printed mount
- Customized diffuser masks

Anticipated Results and Deliverables

Anticipated Results

The foremost aspect to highlight regarding the anticipated results of this project is the development of a metric to quantitatively evaluate the reconstruction results of an image. This project aims to reconstruct 3D facial features and corporeal forms from near, distant, static, and dynamic images captured by the Pi Camera lensless camera system. The first anticipated result of this project is the design of a replicable lensless Pi Camera system with 3D-printable mask mounts. The second anticipated result is the 3D reconstruction and colorization of generated raw data, such that there is discernible geometry and brightness. The third anticipated result is the utilization of a defined metric to compare and evaluate the effectiveness of various diffusers and masks made from different materials. Lastly, the fourth anticipated result are significant findings which can be used to develop an optics lab for an undergraduate course at WashU.

Deliverables

There are four main deliverables for this project: The final report, associated code, optics lab, and live demonstration.

The final report will be our findings and tests in the form of a formal paper, and the associated code will be in a consolidated GitHub repository, containing all hardware and software related code. The final report will contain all findings, procedures, and relevant information regarding the 3D reconstruction and colorization of near, distant, static, and dynamic images captured by a lensless camera system that employs various masks. The optics lab will be a formal lab assignment for an undergraduate optics course for the McKelvey School of Engineering. This lab will have its own associated GitHub repository, for students to clone and utilize for their lab. All relevant information for this lab will be within the final report. Lastly, the live demo will be conducted both in person and via video recording. The demo and video recording will document the capture of a dynamic image of a distant subject moving towards the camera, using the lensless Pi Camera system. This allows for a simultaneous demonstration of all desired capabilities of the lensless camera; subject depth and image type. Subsequently, within the video and live demonstration the software algorithm will be executed to generate 2D reconstruction, 3D reconstruction, and colorization of the captured image. The video recording will be submitted along with the research report and other deliverables.

Timeline

The Gantt Chart in Figure 5 outlines the project timeline, highlighting the primary methodology and major deliverables. The periods are split into biweekly intervals.

Both team members are responsible for completing all deliverables, indicated in the last three rows of the chart. Responsibilities are primarily divided based on whether a task is hardware or software-oriented, with hardware tasks to Jasmine and software tasks to Mae.

Jasmine, pursuing a B.S. in Electrical Engineering at WashU, brings a background in an optical imaging lab, where she conducted research regarding signal processing with fluorescence microscopy. Mae, pursuing a B.S. in Computer Engineering and M.S. in Engineering Management, possesses an optics background from coursework at WashU, as well as from her pursuit of a B.A. in Physics with a minor in Computer Science from Colgate University.

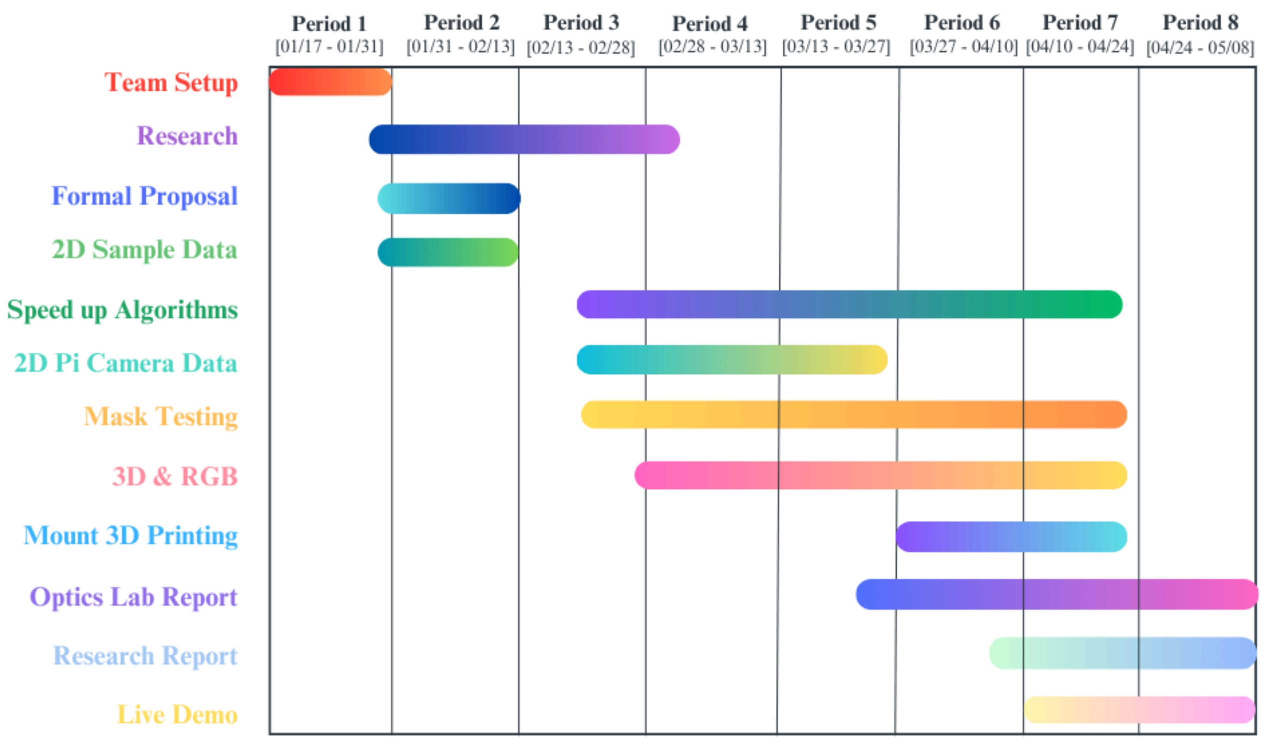

Figure 5. Gantt Chart for Project Completion.

References

[1] N. Antipa *et al.*, “DiffuserCam: lensless single-exposure 3D imaging,” *Optica*, vol. 5, no. 1, p. 1, Jan. 2018, doi: 10.1364/OPTICA.5.000001.

[2] K. Yanny *et al.*, “Miniscope3D: optimized single-shot miniature 3D fluorescence microscopy,” *Light Sci Appl*, vol. 9, no. 1, p. 171, Oct. 2020, doi: 10.1038/s41377-020-00403-7.

[3] V. Boominathan, J. T. Robinson, L. Waller, and A. Veeraraghavan, “Recent advances in lensless imaging,” *Optica*, vol. 9, no. 1, p. 1, Jan. 2022, doi: 10.1364/OPTICA.431361.

[4] N. Antipa, P. Oare, E. Bostan, R. Ng, and L. Waller, “Video from Stills: Lensless Imaging with Rolling Shutter,” in *2019 IEEE International Conference on Computational Photography (ICCP)*, IEEE, May 2019, pp. 1–8. doi: 10.1109/ICCPHOT.2019.8747341.

[5] Joseph Goodman, *Introduction to Fourier Optics*, 3rd edition. Roberts and Company Publishers, 2004.