End-to-end optimization of a lensless imaging system

Cindy Nguyen
Electrical Engineering
Stanford University
cindyn@stanford.edu

1. Proposal

I propose to codesign a coded aperture and reconstruction algorithm for a lensless imaging system.

2. Motivation

Modern imaging systems can be bulky, expensive, and relatively complex to manufacture. Notably, camera lenses take up significant volume not only by the thickness of the glass of the lens, but also through the distance between the sensor and lens required for light focusing. Lensless imaging systems using a coded aperture offer an attractive alternative for reducing the camera volume while also reducing the costs and steps to manufacture a camera with a glass lens. This reduction in size and cost can enable improved applications of *in vivo* microscopy and depth sensing on autonomous vehicles.

3. Related Work

A typical camera will use a lens to map points on a scene to a point on a CMOS sensor. In a lensless imaging system, computational reconstruction algorithms are required for demultiplexing the signal as each point source of the scene is mapped to multiple points on the sensor array.

The FlatCam is an imaging system that places a binary coded aperture in front of the sensor array, in which the phase mask modulates the incoming light with a designed mixture of transparent and opaque features (Figure 1) [2]. For a surface S in the field of view of the sensor, the sensor measurements *y* were described as

$$y = \Phi x + e \tag{1}$$

in which Φ represents the transfer matrix of the phase mask and *x* represents the image unrolled, and *e* represents noise in detection. To ensure a well-conditioned system that would allow stable inversion and reduced computational complexity, their system utilizes separable phase masks that allow (1) to be rewritten as

$$Y = \Phi_L X \Phi_R^T + E \tag{2}$$

in which $X \in \mathcal{R}^{N \times N}$ represents the scene, $Y \in \mathcal{R}^{M \times M}$ represents the sensor measurements. Φ_L and Φ_R represent the separable mask as a outer product of two one-dimensional patterns. They place the coded aperture as close as 0.5 mm away from the sensor. Their image reconstruction algorithm is based on solving the least-squares problem

$$\hat{X}_{LS} = \underset{x}{\operatorname{argmin}} \, ||\Phi_L X \Phi_R^T - Y||_F^2, \tag{3}$$

typically with a regularization term.

Although the system provides impressive success given a computationally simple reconstruction algorithm, the linear demultiplexing based system leads to noise amplification and a reduction in resolution. There is also about a 100 ms delay between capture and image display, which may not be as acceptable in real-time systems such as virtual reality.

Another lensless imaging system is the DiffuserCam, which provides single-shot volumetric imaging under incohrerent light using a diffuser placed over a sensor [1]. Each point source on the 3D surface creates a unique pseudorandom pattern on the sensor, and with the help of compressed sensing algorithms, 3D renderings can be generated from a diffuser phase mask in front of a 2D sensor array. Contrary to the FlatCam, the calibration process does not require precise alignment, and the system is more light efficient than those that use amplitude phase masks. For the phase mask, the system uses a thin transparent diffuser with varying thickness. DiffuserCam uses similar matrix inversion as the FlatCam as part of its reconstruction algorithm combined with a convolution model.

To diverge away this pipeline of designing the phase mask first and then solving a modified inverse problem, I propose to take and end-to-end optimization approach. This method involves the joint optimization of a phase mask design and the reconstruction algorithm over a large set of images, which has been demonstrated in achromatic extended depth of field and super-resolution imaging [5].

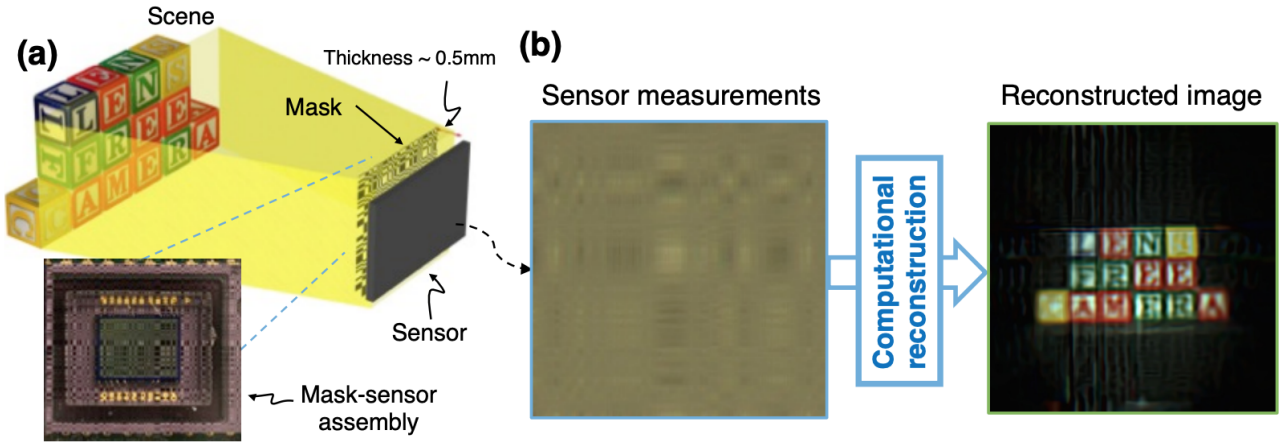

Figure 1. FlatCam architecture. Every every light source contributes to the measurement of each pixel on the sensor. The reconstructed image can be retrieved by solving an inverse problem. [2]

4. Project Overview

For the duration of the quarter, I would like to set up the simulation pipeline to perform wave-based propagation, post-processing reconstruction, and backwards propagation on the reconstruction weights and phase mask design. I will be conducting this work in PyTorch using the Semantic Boundaries Dataset [3].

5. Milestones

By week 7, I would like to have a module set up to simulate point spread function (PSF) of point sources to a sensor.

By week 8, I would like to have set up a U-Net to perform deconvolution on some simulated blur images [4].

By week 9, I would like to have the point spread function module incorporated into the propagated feedback and perform training to start learning the optimal phase mask and lights for convolutional neural network-based reconstruction.

By week 10, I would like to have plots of loss curves and PSNR for images over a large dataset. I would have a written report and poster summarizing our work.

References

[1] N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller. Diffusercam: lensless single-exposure 3d imaging. *Optica*, 5(1):1–9, 2018.

[2] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. Baraniuk. Flatcam: Thin, bare-sensor cameras using coded aperture and computation. *arXiv preprint arXiv:1509.00116*, 2015.

[3] B. Hariharan, P. Arbelaez, L. Bourdev, S. Maji, and J. Malik. Semantic contours from inverse detectors. In *International Conference on Computer Vision (ICCV)*, 2011.

[4] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[5] V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein. 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):1–13, 2018.