General image compression using random-PSF metasurfaces and computational back-end

Yubo Zhang,[1] Rui Chen,[1] Minho Choi,[1] Johannes E. Fröch,[1,2] and Arka Majumdar[1,2,*]

[1] *Department of Electrical Engineering, University of Washington, Seattle, WA 98195*
[2] *Department of Physics, University of Washington, Seattle, WA 98195*

[*] *arka@uw.edu*

Abstract: Compressive imaging that reduces the pixel numbers can lower the cost and power consumption. Here, we use a random-PSF metasurface and computational back-end to achieve a compression rate of 50% for general scenes.

1. Introduction

Due to the explosive growth of data for machine learning and high-resolution video streams, processing information using free-space optics is attracting more interest, thanks to their potential to relax the electronic computing burden. Many works [1–4] have demonstrated optical neural networks (ONNs) that work as a complementary edge computing technique to lower computation power and latency. One important task for optical information processing is to compress the images directly to achieve a higher data rate for high density data transfer. Moreover, such image compression can be a preprocessing step to reduce the subsequent electronic computation. However, most ONNs are designed for specific datasets, limiting their use for general scenes.

Here, we show an image compression scheme for general natural scenes that uses a single-layer metasurface with a random point-spread function (PSF) and a computational back-end. The picture to compress was displayed in an incoherent organic light emitting diode (OLED) display and captured by a visible camera after passing the metasurface. We numerically showed a compression rate of more than 50% with a high structural similarity index measure (SSIM) of 0.85. This was done by using the PSF as prior knowledge and performing a modified total variance (TV) algorithm to reconstruct the pictures. The metasurface was then inversely designed by the Gerchberg–Saxton(G-S) algorithm to produce the exact PSF with little error. In the experiment, we fabricated an $800\,\mu m \times 800\,\mu m$ metasurface with 2400×2400 nano-scatterers on a SiN-on-Quartz chip. The PSF was experimentally measured, and we demonstrated that our system is largely linear shift-invariant, showing the potential of using this method to realize image compression. Our proposed method opens new possibilities for low-energy sense and compute at the edge.

2. Results

This universal image compression idea can be intuitively understood as a spatial-division multiplexing version of the conventional single-pixel imaging technique [5]. The essential requirement is to achieve a full-rank optical transfer matrix for coherent illumination. We verify that this can be effectively achieved with a random PSF under incoherent illumination. The incoherent light is crucial for the system's future development since incoherent displays are much more cost-efficient and could potentially enable real scene compression outside the lab. Especially for sense and compute in real world, we almost always need to rely on incoherent illumination.

Fig. 1a shows the schematic diagram of the experimental setup. Here, the distance between the metasurface and the receiving screen is 4 mm to allow enough space for optical field evolution. The objective and the tube lens form an imaging system to image the field distribution at the output plane. The image on an incoherent display is placed sufficiently far away from the metasurface ($>$ 10 cm). Therefore, the input image can be considered plane waves, and the output image is approximately given by the two-dimensional convolution of input light intensity and the metasurface PSF. This convolution operation can be expanded to a linear matrix and is equivalent to linear processing of the input image. To achieve good reconstruction, the equivalent linear matrix should be full-rank. Importantly, we verified that random PSFs can produce such a full-rank transmission matrix. Moreover, the random PSF is universal, *i.e.*, different types of images can be compressed by the same metasurface. This could be mathematically understood with the random projection operator, which has been used for information compression in other fields. Using the G-S optimization algorithm, we iteratively designed the nano-scatters of the metasurface, such as shown in Fig. 1b, to achieve the desired PSF with little error. Although our current design only supports single wavelength ($\approx$ 632 nm), multi-wavelength operation could be achieved with more complex scatterer/ meta-atom engineering. Since the output plane preserves most of the information of the input image, we can retrieve the input image by a modified TV algorithm, which is an efficient and robust single-pixel imaging algorithm. It reconstructs the image with the prior that the gradient's integral of a natural image is statistically low. The compression ratio η is given by the ratio between the number of samples on the output plane N_o and the initial image N_i, or $\eta = N_o/N_i$. Each output sample pixel can be seen as the pixel of a single-pixel imaging. Our theoretical results (Fig.1c-1e) show that the linear transformation caused by the random PSF can maintain a high similarity ($SSIM > 0.85$) when restored by the modified TV algorithm for a compression rate of $\eta = 25\%$ on the cameraman picture. Thanks to the generality of our image compression scheme, we show excellent reconstruction quality on pictures from various datasets, such as cifar10, fashion MNIST, and MNIST, with a compression ratio of $\eta = 47\%$.

In the experiment, the metasurface was fabricated on a SiN-on-Quartz chip. The meta-atoms had a periodicity of 350 nm and were arranged in a 2400×2400 array. We defined the metasurface pattern with electron-beam lithography (EBL) using a positive-tone resist ZEP 520A. An alumina hard mask was deposited and patterned by a liftoff process. The pattern was transferred to the SiN by an inductively coupled plasma etcher (ICP) using Flourine-based chemistry. We then made a metal (200 nm Cr) aperture with Heidelberg laser writing lithography followed by a liftoff process. Fig. 1g shows the experimentally measured random PSF of the metasurface with a He-Ne laser source and a 100-μm-diameter pinhole. The linear shift invariant property is verified, and Fig. 1h shows the relation between the shift in the image and input planes. The number of pixel shifts is estimated by the tracking reference point method, which shows good translation invariance with an r-value of 0.99989. Therefore, since the pixels on the real screen are discrete points, we can control the position of the bright spots to make them strictly follow the matrix convolution relation in theory. Currently, the low intensity of the OLED incurs a low signal-to-noise ratio (SNR) at the captured image, limiting the image compression. A high-brightness OLED display with a small pixel resolution will be used to improve the scheme.

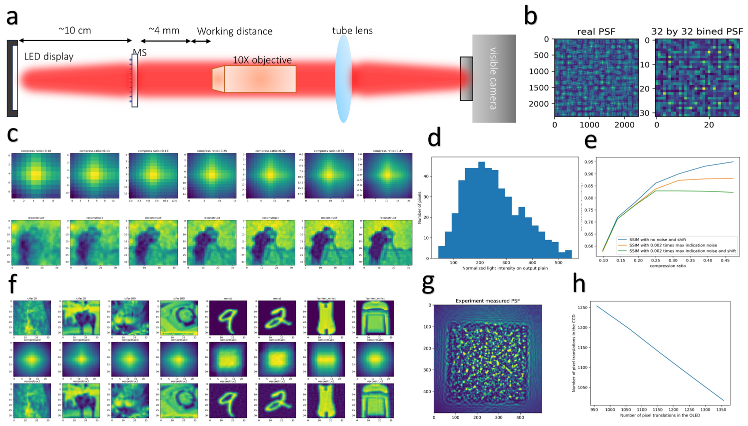

Fig. 1. **General image compression with a random PSF metasurface and computational back-end.** (a) Schematic for the image compression scheme. (b) The high-resolution normalized random PSF (left) linearly transforms the input. We binned the PSF (right) to accommodate the input image size (32×32). (c) The output image after the image compression metasurface (upper row) and their corresponding reconstructed images (lower row) with different compress ratios. A good image quality can be seen at a compression ratio as low as 25%. (d) The histogram of the light intensity distribution of the sampling points at the 47% compression rate, and the peak of the histogram shifts to the left as the compression ratio (sampling points number) increases. (e) The SSIM plot for different compression ratios shows a high SSIM of > 0.85 for a compression ratio $> 25\%$, even when appropriate environmental noise and shift are considered. (f) Image compression results for different datasets at a compression ratio of 47%. The method shows excellent generality. (g) Experimentally measured random PSF of the metasurface with a He-Ne laser source and a 100-μm-diameter pinhole. (h) The linear relation between the shift in the output image and input planes.

3. Acknowledgement

The work is supported by National Science Foundation (NSF-ECCS-2127235) and DARPA STTR programs

References

1. L. Huang, et al. ”Photonic Advantage of Optical Encoders,” Nanophotonics (2023).
2. Y. Li, et al. ”Optical information transfer through random unknown diffusers using electronic encoding and diffractive decoding,” Adv. Photonics **5**, 046009 (2023).
3. Y. Chen, et al. ”Photonic unsupervised learning variational autoencoder for high-throughput and low-latency image transmission,” Sci. Adv. **9**, eadf8437 (2023).
4. Y. Chen, et al. ”All-analog photoelectronic chip for high-speed vision tasks” Nature **623**, 48-57 (2023).
5. M. P. Edgar, et al. ”Principles and prospects for single-pixel imaging,” Nat. Photon. **13**, 13–20 (2019).