

Designing lensless imaging systems to maximize information capture

LEYLA A. KABULI,^{1,*}  HENRY PINKARD,¹  ERIC MARKLEY,² CLARA S. HUNG,¹ 
AND LAURA WALLER¹ 

¹Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720, USA

²Graduate Program in Bioengineering, University of California, Berkeley and University of California, San Francisco, California 94720, USA

*lakabuli@berkeley.edu

Received 11 June 2025; revised 25 December 2025; accepted 30 December 2025; published 28 January 2026

Mask-based lensless imaging uses an optical encoder (e.g., a phase or amplitude mask) to capture measurements, then a computational decoding algorithm to reconstruct images. In this work, we evaluate and design lensless encoders based on the information content of their measurements using mutual information estimation. Our approach formalizes the object-dependent nature of lensless imaging and quantifies the interdependence between object sparsity, encoder multiplexing, and noise. Our analysis reveals that optimal encoder designs should tailor encoder multiplexing to object sparsity for maximum information capture, and that all optimally encoded measurements share the same level of sparsity. Using mutual information-based optimization, we design information-optimal encoders for compressive imaging of fixed object distributions. Our designs demonstrate improved downstream reconstruction performance for objects in the distribution, without requiring joint optimization with a specific reconstruction algorithm. We validate our approach experimentally by evaluating lensless imaging systems directly from captured measurements, without the need for image formation models, reconstruction algorithms, or ground truth data. Our comprehensive analysis establishes design and engineering principles for lensless imaging systems and offers a model for the study of general multiplexing systems, especially those with object-dependent performance. © 2026 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

<https://doi.org/10.1364/OPTICA.570334>

1. INTRODUCTION

Traditional optical design focuses on optimizing hardware to directly capture images of objects. In contrast, computational imaging systems expand system design into two parts: encoding measurements with non-traditional optics and decoding measurements using a reconstruction algorithm. For example, mask-based lensless imagers encode images of objects using a thin amplitude or phase mask placed on top of an image sensor. The captured measurement looks nothing like the object, but a decoder is able to recover the object from the measurement by solving an inverse problem [1]. Lensless imagers offer a compact form factor and wide field-of-view, showing promise in applications including photography [2,3], mesoscopy [4,5], and microscopy [6–8].

While traditional imaging systems are designed to map each point of an object to a point on the sensor, lensless imaging systems instead map each object point to multiple points on the sensor, forming a multiplexed measurement. This multiplexing is necessary for compactness, since light cannot be concentrated fully when the distance between the mask and sensor is small. Multiplexing is also advantageous because it enables compressive encoding; for example, a single 2D measurement can be used to reconstruct 3D [2,3], spatio-temporal [9], and hyperspectral [10] object properties. In 2D imaging, multiplexing enables extended

field-of-view imaging [11–13], tolerance to sensor erasure [14], and uniform resolution across a wide field-of-view [15,16].

Optimal design of lensless imaging encoders is still an open problem. Most previously proposed designs are driven by heuristics; for example, phase masks are preferable to amplitude masks [8,17] for their light efficiency, with previous designs including off-the-shelf Gaussian diffusers [2], contour masks that preserve good frequency transfer [3,4,7], and lenslet arrays, which limit the amount of multiplexing to improve system performance [5,6,11,18,19]. Fabrication advances now provide the flexibility to realize virtually any phase profile [20,21]. What is still needed are principled methods for determining optimal designs.

Typically, a candidate encoder design is evaluated based on the quality of the reconstructed images. Quality can be quantified by error metrics like peak signal-to-noise ratio (PSNR) or structural similarity index measure (SSIM); however, these require paired ground truth data. In the absence of ground truth data, quantifying image quality is challenging. Further confounding the evaluation of encoder design is the fact that the final image quality incorporates the combined effect of the encoder and decoder. Reconstruction quality depends heavily on the choice of decoder, priors, and hyperparameters [22–24], which makes it difficult to isolate the impact of the encoder. For example, when reconstructing 3D data from a 2D measurement, sparsity constraints are

crucial [2,25] and the performance impact of these constraints is object-dependent. Furthermore, as algorithms shift from physics-based models [24] to deep neural networks [24,26,27] trained on large datasets [28–30], the state-of-the-art is continually updated. Ultimately, the performance of any of these inverse problem algorithms is limited by the quality of the encoded measurements. This highlights the need for a principled, decoder-independent evaluation method for lensless imaging.

Decoder dependence is a key limitation of end-to-end design, in which an encoder and decoder are jointly optimized with the goal of maximizing downstream performance (e.g., reconstruction quality) [31–33]. In lensless imaging, end-to-end efforts have been applied to amplitude masks and task-specific objectives such as classification [34–36], rather than general image reconstruction. Encoder performance remains confounded with decoder performance in end-to-end design, resulting in encoder designs that are specific to a particular decoder architecture and may not remain optimal as decoders are updated. In addition, this optimization is computationally intensive, which can make the design intractable for high-dimensional problems, and is sensitive to initialization and often highly non-convex [37].

Traditional system design metrics such as condition number [2] and modulation transfer function (MTF) [3,5,18] are decoder-independent, but do not account for object properties or non-white noise. To better evaluate these effects in multiplexing systems, foundational work in coded aperture imaging defined multiplexing patterns for optimal object-independent performance [38], and subsequent studies analyzed multiplexing in conjunction with detection noise and signal priors [39,40] by evaluating measurement contrast [41] and other data statistics [42,43]. Mutual information, which naturally combines object properties, such as sparsity, with imaging system properties, including signal-to-noise ratio (SNR) and bandwidth [44], has the flexibility to capture the combined effects of the object, encoder, and noise [45–48]. With recent developments in practical estimation methods [49], mutual information is particularly promising for decoder-independent system evaluation.

In this work, we evaluate and design encoders for lensless imaging by directly quantifying the information content of measurements through mutual information estimation [49]. Our method estimates information from a dataset of measurements and a noise model, evaluating system performance without requiring image formation models or ground truth data. This measurement-based approach explicitly incorporates object dependence by evaluating encoders with respect to specified object distributions, and naturally extends to data-driven optimization of object-dependent encoder designs for general downstream performance. Although mutual information does not explicitly account for decoder effects and is therefore not a guaranteed predictor of downstream performance, empirical studies demonstrate consistent agreement between mutual information and general tasks such as reconstruction [45,49,50].

Using this method, we study tradeoffs between encoder multiplexing, object sparsity, and detection noise in lensless imagers. Multiplexing enables compressive encoding and a compact form factor, but too much multiplexing can be detrimental to system performance, especially for dense objects. We find that the ideal amount of multiplexing should be tuned based on object sparsity. Our analysis further shows that all optimally encoded measurements have the same measurement sparsity. We then demonstrate

the design of information-optimal encoders for compressive imaging with multiplexing levels optimized for different object distributions, and observe that these designs lead to improved reconstruction performance. Finally, we analyze experimental lensless imaging systems and confirm that dense natural images require encoders with less multiplexing. Our contribution establishes engineering and design principles for lensless imaging, quantifying fundamental performance limits and enabling practical system improvements.

2. ENCODER EVALUATION FRAMEWORK

To study the interactions between objects and encoders, we estimate mutual information from measurements generated by various objects and encoders. This evaluation method is probabilistic, so it applies to a dataset of measurement samples rather than any individual measurement [Fig. 1(a)].

First, the image formation process maps a distribution of objects $\mathbf{O}$ to a distribution of encoded images $\mathbf{X}$. Specifically, each encoded image $\mathbf{x}$ is modeled as a convolution between the corresponding object $\mathbf{o}$ and a multiplexing point spread function (PSF) $\mathbf{h}$ [2]:

$$\mathbf{x} = \mathbf{o} * \mathbf{h}. \quad (1)$$

The PSF is determined by the encoding phase mask [Fig. 1(b)]. In practice, this image formation model also includes a crop due to the finite extent of the detector [2].

The goal of the encoder is to best capture object information to later be computationally recovered. While the objects and encoded images are distributions, there is only one encoder, which acts deterministically, such that repeated encodings of a given object produce the same encoded image. This encoding can be represented by a transformation $\mathbf{O} \rightarrow \mathbf{X}$ where the encoded image distribution $\mathbf{X}$ depends solely on the object distribution $\mathbf{O}$.

The encoded image distribution $\mathbf{X}$ is corrupted by noise at the sensor, forming the distribution of noisy sensor measurements $\mathbf{Y}$ [Fig. 1(a)]. We model the detection process with signal-dependent shot (Poisson) noise, which, in the high photon count regime, can be approximated as Gaussian with variance equal to the photon count at each pixel [51]. The detection process does not depend on any knowledge of the specific encoder or objects being imaged; it simply introduces detection noise on the encoded measurements, represented by a transformation $\mathbf{X} \rightarrow \mathbf{Y}$. The complete imaging model can be represented by a Markov chain $\mathbf{O} \rightarrow \mathbf{X} \rightarrow \mathbf{Y}$ and follows the widely used semi-classical theory of photodetection [48,51].

Our goal is to estimate the mutual information between objects and measurements, $I(\mathbf{O}; \mathbf{Y})$, which quantifies how well objects can be distinguished from each other based on the corresponding noisy measurements; essentially, how much object information survives the encoding and detection processes. Although we may not have access to the true object distribution for estimating $I(\mathbf{O}; \mathbf{Y})$ directly, we can equivalently estimate $I(\mathbf{X}; \mathbf{Y})$, which measures how much of the object information captured in encoded images survives the detection process [47–49]. Applying the data processing inequality to our imaging model Markov chain, $I(\mathbf{O}; \mathbf{Y}) \leq I(\mathbf{X}; \mathbf{Y})$ [52], and since the encoding process is fixed with no additional sources of randomness, it follows that $I(\mathbf{O}; \mathbf{Y}) = I(\mathbf{X}; \mathbf{Y})$ (Section S1.1 of Supplement 1).

Mutual information is decomposed as

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}). \quad (2)$$

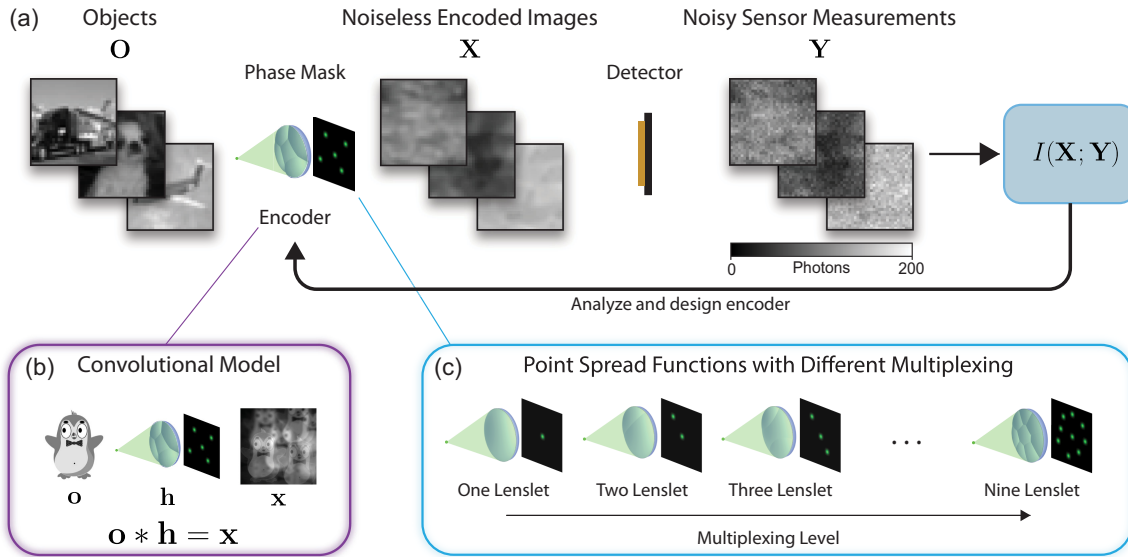


Fig. 1. Mutual information estimation framework for lensless imaging. (a) A phase mask encoder maps an object distribution O to a noiseless image distribution X . Image detection adds noise, forming a distribution of noisy sensor measurements Y . Mutual information, estimated from the noisy sensor measurement distribution, serves as a metric for analyzing and designing the encoder. (b) The lensless image formation model takes each object o and encodes it into a noiseless image x by convolving with the encoder point spread function h . (c) To analyze the effects of multiplexing in lensless imaging, we study encoders with one through nine lenslets, corresponding to increasing multiplexing levels.

Here, $H(Y)$ is the entropy of the noisy measurements, which captures variations across both the object and noise distributions, and $H(Y|X)$ is the conditional entropy (randomness) due to noise alone. We adopt the mutual information estimation approach outlined in Pinkard *et al.* [49], which consists of fitting a probabilistic model to the measurements and compensating for detection noise.

The entropy $H(Y)$ quantifies the randomness in the noisy measurements. Entropy can be approximated by fitting a probabilistic model $p_\theta(y)$ to noisy measurement samples from the true measurement distribution $p(y)$ and calculating the cross-entropy $\mathbb{E}[-\log p_\theta(Y)]$ on a held-out test set:

$$H(Y) \leq \mathbb{E}[-\log p_\theta(Y)] \approx -\frac{1}{N} \sum_{i=1}^N \log p_\theta(y^{(i)}), \quad (3)$$

where $y^{(i)}$ is the i th measurement in a test set of N measurements [49,53].

We choose a PixelCNN as the probabilistic model $p_\theta(y)$ [54] because it is powerful and flexible enough to accurately fit distributions of both dense and sparse images from limited samples [49]. Unless otherwise specified, we use 10,000 noisy measurement patches (32×32 pixels) to fit the model and evaluate cross-entropy on a test set of 1500 noisy measurement patches [55]. Each model fit and evaluation takes approximately 10 minutes and uses 3 GB memory on an NVIDIA RTX A6000 GPU.

The conditional entropy $H(Y|X)$ quantifies the randomness introduced by detection noise, which carries no information about the object. We make the common assumption that the noise at each pixel in an M -pixel measurement is independent of other pixels, conditioned on the corresponding encoded image pixel x_k [51]. For Poisson (shot) noise, this results in a closed-form expression for the conditional entropy:

$$H(Y|X) \approx \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^M \frac{1}{2} \log_2(2\pi e x_k^{(i)}), \quad (4)$$

which is a sum of conditional entropies across the M image pixels averaged across a set of N images drawn from the encoded image distribution [49]. This approximation is accurate at high photon counts (>10 photons), where shot noise can be approximated as Gaussian with equal mean and variance [49,51]. Our final estimate of mutual information $I(X; Y)$ is the difference between our upper bound on $H(Y)$ [Eq. (3)] and the value of $H(Y|X)$ [Eq. (4)]. We report mutual information estimates with 95% confidence intervals from bootstrap estimates of cross-entropy and conditional entropy to quantify uncertainty. In Section S3.2 of Supplement 1, we also provide conditional entropy expressions and corresponding system analysis for Gaussian (read) noise.

3. QUANTIFYING TRADEOFFS BETWEEN SPARSITY AND MULTIPLEXING IN 2D IMAGING

Traditionally, imaging systems across microscopy, photography, and astronomy have been designed to have a single diffraction-limited spot as the PSF. These systems assume no computational decoding and minimize encoder multiplexing to form measurements that are directly human-interpretable. In contrast, lensless imagers are intentionally designed with some amount of encoder multiplexing, which enables compressive object encoding and recovery. This compressive imaging ability relies on object sparsity, making the performance of these multiplexing systems inherently object-dependent. In order to provide design guidance, we aim to understand the fundamental tradeoffs between object sparsity and encoder multiplexing. To study these tradeoffs, we combine mutual information estimation with scalar metrics quantifying both sparsity and multiplexing.

Sparse objects are those that can be represented in a basis with few non-zero coefficients [25,56–58]. Sparsity is a critical assumption used in lensless imaging, as decoders harness object sparsity constraints to enable compressive recovery from captured measurements. For example, total variation (TV) sparsity can be used as a

constraint when imaging objects with sparse gradients and edges, such as natural objects [59]. We quantify sparsity, specifically gradient sparsity, using the Tamura coefficient (TC) [60], a scalar metric previously used in holography [61,62] and extended field-of-view imaging [63].

The Tamura coefficient of an image is

$$\text{TC}(G) = \sqrt{\frac{\sigma_G}{\langle G \rangle}}, \quad (5)$$

where G is the image gradient magnitude, σ_G is the standard deviation of the gradient magnitude, and $\langle G \rangle$ is the mean of the gradient magnitude. The TC captures image gradient variability, where a higher TC corresponds to more sparsity and a lower TC corresponds to less sparsity. The high noise tolerance and sensitivity to signal changes [61] of the TC is suitable for our analysis of sparse objects and encoded noisy measurements. We utilize the TC to quantify the sparsity of object and measurement distributions by averaging TC values across samples in the corresponding distributions.

To quantify encoder multiplexing, we use the number of non-zero points in the encoder PSF. Here, we restrict our analysis to the class of lenslet-based encoders [5,6,18,19], which have good SNR performance and provide a simple measure of multiplexing: the number of lenslets corresponds to the number of points in the PSF. We consider multiplexing levels ranging from a single lens (no multiplexing) through nine lenslets (high multiplexing) and show examples of heuristically designed encoders and PSFs for each multiplexing level in Fig. 1(c). Each lenslet's focal spot is modeled as a Gaussian profile with an isotropic aperture shape. For 2D imaging analysis, all lenslets have equal focal lengths. Each encoder maintains a constant photon budget with a fixed total aperture size and a PSF normalized to have unit energy.

We systematically study the parameter space of object sparsity and encoder multiplexing by evaluating the encoded information in various combinations of objects and encoders. We simulate a class of objects for which sparsity and other variables can be precisely controlled: randomly placed point sources, qualitatively similar to bead samples encountered in fluorescence microscopy [6]. These datasets are generated by placing Gaussian-blurred point sources at randomly selected pixel locations. The percentage of pixels selected for point sources controls the object sparsity while keeping other variables fixed. We generate datasets of beads with sparsity levels ranging from $\text{TC} = 0.739$ to $\text{TC} = 1.359$ [Fig. 2(a)]. Each dataset has a constant sparsity level and consists of 50,000 96×96 pixel objects with a mean photon count of 100 photons per pixel. To probe encoder multiplexing levels, we consider encoders ranging from a single lens (no multiplexing) to nine lenslets (high multiplexing), as visualized in Fig. 1(c). Then, we simulate measurements for each multiplexing encoder and object sparsity level following the simulated imaging pipeline in Fig. 1(a). We compute mutual information from the resulting noisy measurements.

In Fig. 2(b), we study the tradeoffs between object sparsity and encoder multiplexing. Mutual information is calculated across object sparsity levels for each multiplexing encoder. For low-multiplexing encoders, mutual information is maximized (denoted by stars) for objects with low sparsity. As multiplexing levels increase, mutual information is maximized with increasingly sparse objects.

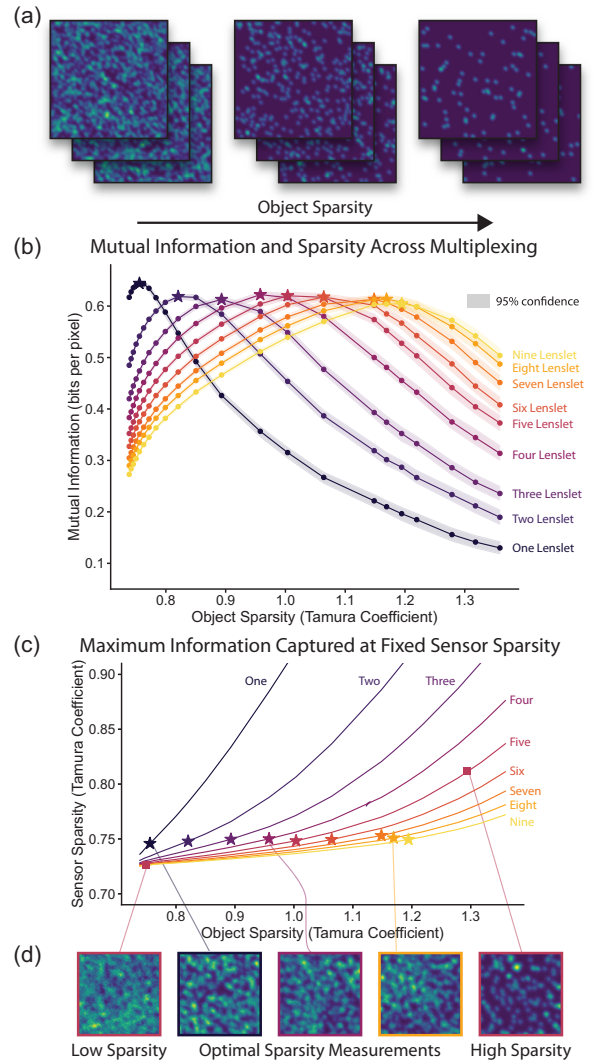


Fig. 2. Mutual information for varying object sparsity and encoder multiplexing. (a) Examples of simulated objects with increasing levels of sparsity as quantified by the Tamura coefficient. (b) Mutual information for one through nine lenslet multiplexing systems, each swept across object sparsity levels. As multiplexing increases, the maximum mutual information (denoted by stars) is achieved at higher sparsity levels. (c) Maximum mutual information (denoted by stars) corresponds to a fixed sensor sparsity across all multiplexing encoders. (d) Example measurements corresponding to optimal fixed sensor sparsity qualitatively differ from examples with low or high sparsity.

Furthermore, by examining only the starred points corresponding to maximum mutual information in Fig. 2(b), we find a surprising shared property of all optimally encoded measurements: they have the same *measurement* sparsity. Measurement sparsity is quantified by applying Eq. (5) to the noisy measurements. To highlight this further, we plot the increasing relationship between object and sensor sparsity and denote the sensor sparsity that corresponded to maximum mutual information for each multiplexing encoder with a star [Fig. 2(c)]. Optimal measurements, achieved by maximizing information throughput, have a constant sensor sparsity ($\text{TC} \approx 0.75$) [Fig. 2(d)]. Our finding is supported by a related phenomenon in communication systems, where the optimal signals for achieving channel capacity produce the same distributions [64].

We verify that our findings generalize to other encoder classes by demonstrating that diffuser-based encoders also achieve the same optimal sensor sparsity in Section S3.1 of Supplement 1. We also confirm that the same trends and optimal sensor sparsity finding hold for read noise-limited imaging systems in Section S3.2 of Supplement 1 and for different object normalizations in Section S3.3 of Supplement 1.

These findings are consistent with general intuition about multiplexing in imaging systems. Traditional single-lens systems are designed to image natural objects, which have low sparsity. Multiplexing encoders, as used in lensless imaging, maximize information encoding with sparse objects, with optimal designs depending on the combination of object sparsity and encoder multiplexing. Our results also have an interpretation in terms of measurement discrimination. Increased multiplexing results in measurements with overlapping object copies. For objects that are dense, these overlapping copies lead to measurement saturation, making measurements difficult to discriminate and reducing information. For extremely sparse objects, overlapping copies can effectively utilize regions of the sensor that would otherwise not detect a signal. This redundancy and resulting noise variation can provide an advantage for discrimination, effectively increasing information.

From a design perspective, these results indicate that the measurement sparsity generated by an encoder for a given distribution of objects can offer simple guidance for evaluating and adjusting the encoder. For maximum information encoding, our findings suggest that using phase masks with more multiplexing can be

beneficial when imaging sparse objects, and using phase masks with less multiplexing is preferable when imaging dense objects.

4. INFORMATION-OPTIMAL ENCODER DESIGN

We studied a series of heuristically designed encoders and identified the object sparsity levels best matched to each encoder in Section 3. Next, we explore the design of information-optimal encoders for lensless imaging that maximize encoded information for fixed object distributions.

Here, we design multiplexing encoders for single-shot extended field-of-view imaging [11,12], in which our objective is to recover a 2D object region that is larger than the finite sensor extent (measurement), as visualized in Fig. 3(a). We can use mutual information to design encoders with Information-Driven Encoder Analysis Learning (IDEAL), which has been used to design color filter arrays and pupil-plane masks with improved downstream reconstruction performance [49,65,66]. IDEAL does not require a decoder. Instead, measurements from the imaging system are evaluated with mutual information as the loss, and the encoder is optimized for a specific object distribution using gradient-based methods via backpropagation. This decoder-independent approach provides significant compute time and memory savings compared to end-to-end design [66] and eliminates the need for retraining for different reconstruction architectures, as algorithms can be flexibly swapped in afterward.

To determine the information-optimal encoder for a specific object distribution $\mathbf{O}$, the design objective is defined by

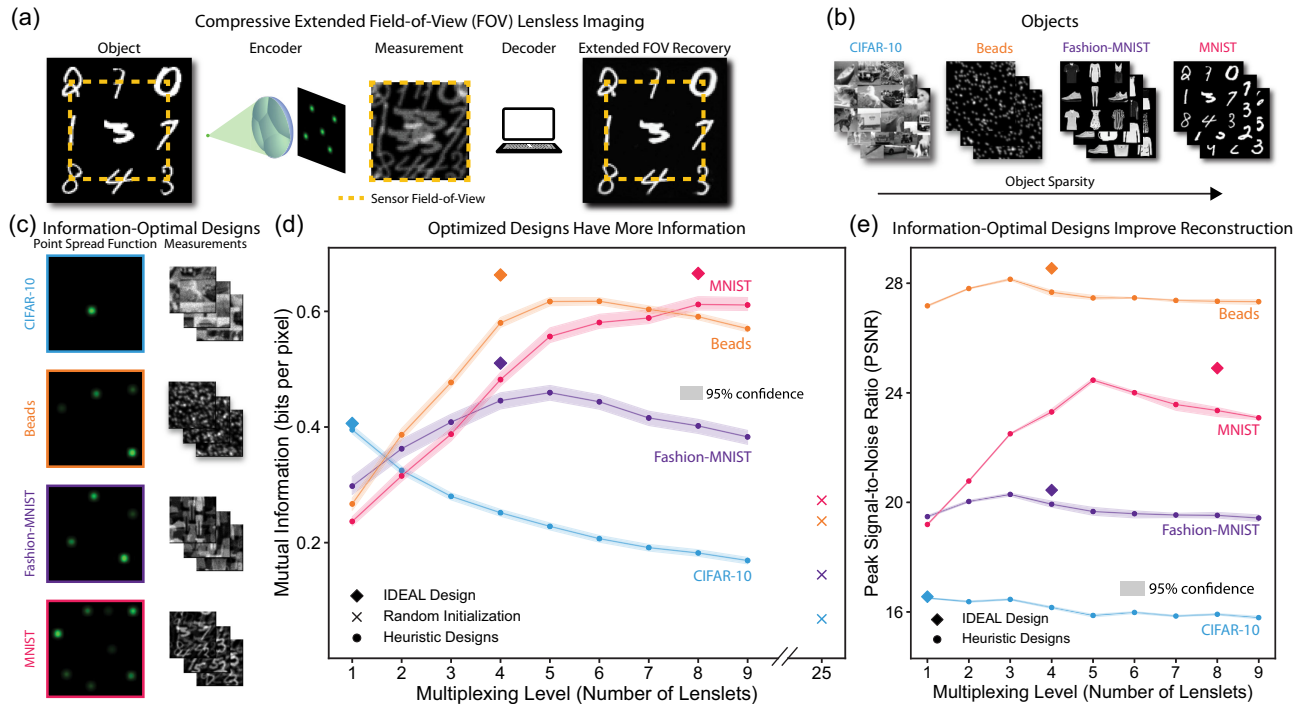


Fig. 3. Information-optimal phase mask design for lenslet-based encoders with Information-Driven Encoder Analysis Learning (IDEAL). (a) Lensless imaging enables compressive extended field-of-view imaging in 2D by encoding an object larger than the finite extent of an image sensor into a limited measurement region through multiplexing. A decoder recovers an extended field-of-view from the measurement. (b) Examples of objects with varying sparsity. (c) Visualization of the information-optimal encoder, point spread function, and example measurements for each object distribution. As the object sparsity increases, the number of lenslets (multiplexing level) in the corresponding information-optimal design increases. (d) Mutual information for IDEAL designs for each object distribution. Random initialization with 25 lenslets is denoted by an “X,” and the resulting IDEAL design is denoted by a diamond. Information-optimal designs have higher information than baseline heuristic designs (denoted by circles). 95% confidence intervals are smaller than the marker size for IDEAL designs and indicated by shading for heuristic designs. (e) Higher mutual information in IDEAL designs results in improved reconstruction performance for each object distribution, as quantified by reconstruction peak signal-to-noise ratio.

$$\mathbf{h}^* = \arg \max_{\mathbf{h} \geq 0, \|\mathbf{h}\|_1=1} I(f_{\mathbf{h}}(\mathbf{O}); \mathbf{Y}). \quad (6)$$

For a non-negative unit energy PSF $\mathbf{h}$, the mapping $f_{\mathbf{h}}(\mathbf{O}) = \mathbf{X}$ applies a convolution with $\mathbf{h}$ to each object in $\mathbf{O}$ following the image formation model in Eq. (1), and $\mathbf{Y}$ is the resulting noisy measurement distribution. At each optimization step, the encoder PSF is updated, the measurement distribution is recomputed based on this PSF and the fixed object distribution, and the mutual information loss is evaluated on these updated measurements. The final PSF $\mathbf{h}^*$ is the information-optimal design for the specified object distribution.

Solving the design objective in Eq. (6) with gradient-based optimization requires a differentiable model of the optical encoder. We extend the class of lenslet-based encoders studied in this work to a differentiable model by representing the PSF of the lenslet array as a sum of Gaussians. Each lenslet's point response is parameterized as a Gaussian with learnable parameters: mean (position), covariance (aperture shape), and weight (aperture size and relative energy). We constrain covariance matrices to be isotropic (round apertures), as designs with isotropic apertures achieved higher information than designs with anisotropic apertures (Section S5 of Supplement 1).

We consider the design of optimal encoders for four different object distributions: CIFAR-10 [67], representing dense natural objects (TC = 0.970); Beads, representing moderately sparse objects (TC = 1.1064); Fashion-MNIST [68], representing intermediate sparsity objects (TC = 1.196); and MNIST [69], representing sparse, binary objects (TC = 1.363). Each dataset is normalized to a mean photon count of 100 photons per pixel, and object samples are visualized in Fig. 3(b).

We use a PixelCNN to compute the mutual information loss [66]. This estimator is slower than the Gaussian estimation model previously used for IDEAL [49] but has the flexibility to fit both dense and sparse distributions. At each iteration, measurements are generated from the updated encoder and the object distribution based on the image formation model in Fig. 1(a). The mutual information loss is calculated with 4096 16×16 pixel measurement patches, of which 90% are used to fit the PixelCNN and the other 10% for the evaluation test set. Based on the loss, lenslet parameters are updated to refine the encoder design, with the objective of maximizing the mutual information. The PixelCNN is reinitialized and refit every 20 encoder optimization steps to reduce computational runtime. Each encoder optimization takes less than two hours and uses 5 GB memory on an NVIDIA RTX A6000 GPU.

We evaluate baseline performance on heuristic lenslet designs for each object distribution and plot the information in Fig. 3(d). The general trend is consistent with previous findings: dense objects benefit from lower multiplexing, and sparse objects benefit from higher multiplexing. We then initialize our IDEAL optimization process with a random distribution of 25 lenslets (Fig. S7 of Supplement 1) and train to convergence for each object distribution. The mutual information for this random initialization encoder is denoted by an "X" for each dataset. We optimize the encoder by updating lenslet positions, widths, and relative weightings using IDEAL. We use a complete model fit with 10,000 32×32 pixel patches and a 1500 patch test set to calculate mutual information for the resulting information-optimal design, denoted by a diamond in Fig. 3(d). The information-optimal design for each object distribution has higher mutual information than any of the heuristic designs, with significant improvements on the

different designs for the Beads, Fashion-MNIST, and MNIST datasets.

Figure 3(c) visualizes the resulting encoder PSF from IDEAL and an example measurement for each object distribution. The optimal design for encoding CIFAR-10 natural images is a single lens, which is expected because dense objects do not benefit from multiplexing. As a result, the information gain for the CIFAR-10 design is minimal, as there are limited degrees of freedom for designing a single lens. For both the Beads and Fashion-MNIST datasets, the optimal design converges to four lenslets, but with different weights and lenslet positions. For sparse MNIST objects, the optimal design converges to eight lenslets.

We verify downstream reconstruction performance for our information-optimal designs. For each object distribution and corresponding design, we reconstruct images with extended field-of-view from multiplexed measurements using a lightweight U-Net (Section S1 of Supplement 1) [24]. Example reconstructions are visualized in Fig. S7 of Supplement 1. In Fig. 3(e), we compare average reconstruction performance (PSNR) for each IDEAL design and all heuristic designs. Although these information-optimal designs are not optimized for a specific decoder, the resulting measurements lead to improved reconstruction performance for the specified object distribution, demonstrating, for this compressive lensless imaging task, that increased information benefits downstream recovery using a U-Net.

Optical encoder design is often non-convex and sensitive to initialization [37]. In Section S5 of Supplement 1, we verify our results from Fig. 3(d) by optimizing with 10 different random initializations, all of which converge to similar encoders that outperform heuristic designs. We also consider heuristic initialization, in which we optimize the positions of the lenslets while keeping other parameters, including multiplexing level, fixed. For sparse datasets including Beads and MNIST, designs with heuristic initialization achieve slightly higher information than with random initialization and additionally produce measurements achieving optimal sensor sparsity (Section 3). Evaluating sensor sparsity can help determine whether a design is close to the optimum.

Our information-optimal designs for various objects verify our previous findings, including that sparse objects benefit from higher multiplexing (Section 3). Future work can investigate manufacturing information-optimal phase masks for lensless imaging systems. Constraining optimization to equally weighted lenslets can be simpler for manufacturing, although recent developments in lithography techniques have demonstrated relative ease of manufacturing of non-traditional lenslets and other phase masks [20,21].

5. EXPERIMENTAL INFORMATION EVALUATION

We demonstrated lensless imaging encoder evaluation and design in Sections 3 and 4 via simulations. Now, we apply our approach to experimental data. Experimental measurements include additional imaging non-idealities, including a shift-varying PSF. Mutual information is well suited for experimental analysis, as it directly evaluates measurements without requiring an explicit image formation model. This captures the effects of non-idealities without the need for extensive system characterization or calibration (e.g., PSF capture).

We consider three encoders: a traditional lens with no multiplexing, a random multi-focal lenslet (RML) array phase mask

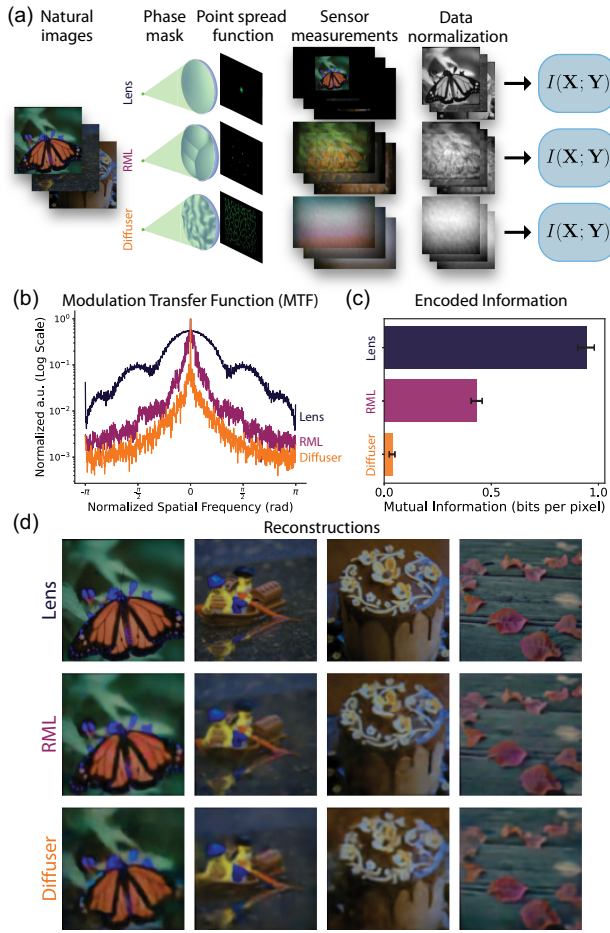


Fig. 4. Mutual information in experimental lensless imaging comparing a traditional lens to two phase mask encoders: a diffuser and a random multi-focal lenslet (RML) array. (a) Mutual information is estimated from experimentally captured measurements of natural images for each imaging system. (b) 1D cross-section of the modulation transfer function for each encoder. The high-multiplexing diffuser has the worst frequency transmittance. (c) The lens system has the most mutual information, and information decreases with multiplexing. (d) Examples of reconstructions for the RML and diffuser systems compared to the lens system show worse image quality with the high-multiplexing diffuser. The reconstruction structural similarity index measure (SSIM) for the RML (SSIM = 0.886) is higher than the diffuser (SSIM = 0.821).

[19] with moderate multiplexing, and a Gaussian diffuser phase mask [2] with high multiplexing. The object distribution is natural images represented by the MIRFLICKR-25000 dataset [28] (mean TC = 1.238, standard deviation 0.213) displayed on a monitor placed in the imaging plane. These data are from an open-access dataset consisting of measurements from each system captured in parallel under identical imaging conditions [30].

To calculate mutual information, we calibrated the noise model and estimated the amount of shot noise. We converted raw camera data from pixel intensities to photon counts, assuming that shot noise is the dominant source of noise (equal variance and mean) [70]. We processed measurements by converting to grayscale, cropping to square dimensions, and downsizing to 100×100 pixel patches with pixel binning [Fig. 4(a)]. To correct quantization artifacts from limited sensor bit depth, we added additional shot noise to simulate lower signal-to-noise ratio measurements.

Mutual information is estimated by fitting the PixelCNN model with 10,000 measurement patches, with a test set of 1000 patches for evaluation. Since experimental noisy measurements do not provide access to the noiseless image distribution $\mathbf{X}$, we use $\mathbf{Y}$ in place of $\mathbf{X}$ for conditional entropy estimation [Eq. (4)]. This approximation is accurate when images have a reasonably high photon count (> 10 photons) [49].

First, we evaluate system performance with decoder-independent evaluation methods, using mutual information and the MTF, which quantifies the filtering characteristics of the encoder in isolation without accounting for object or noise properties. In Fig. 4, we visualize each system's PSF and a 1D cross-section of the MTF, with 2D MTFs shown in Fig. S6 of Supplement 1. As multiplexing increases, the MTFs have worse frequency transmittance [Fig. 4(b)]. Mutual information estimation shows that information decreases as multiplexing increases [Fig. 4(c)]. This is expected based on Section 3; dense natural images do not benefit from multiplexing. MTF-based analysis requires system PSF calibration and captures encoder properties in isolation, while mutual information does not require PSF calibration and additionally captures both object and noise properties.

We also compare direct measurement-based evaluation using mutual information with traditional reconstruction-based evaluation. For each lensless imager, we reconstruct images using a ConvNeXt architecture [27,71] with a train set of 24,000 measurements and a test set of 1000 measurements. All measurements are $8\times$ downsampled, and the images from the lens system serve as ground truth. Reconstructions are visualized in Fig. 4(d). The RML reconstructions achieve a test set structural similarity index measure (SSIM) of 0.886, while the diffuser reconstructions have a worse SSIM of 0.821. The reconstruction quality is also qualitatively worse with more multiplexing, as fine details and sharp edges are missing in the diffuser reconstructions but recovered in the RML reconstructions. The decoder-independent information estimates are in agreement with reconstruction quality while avoiding the need for paired ground truth, which is often unavailable in experimental settings. This highlights a valuable aspect of our decoder-independent approach, as it enables quantitative evaluation directly from raw measurements.

Although our experimental results confirm that a traditional lens outperforms lensless systems for 2D natural images, lensless imagers are motivated by factors beyond image quality. These multiplexing systems enable compressive image encoding (e.g., extended field-of-view or depth) and can achieve an ultra-thin form factor [8], capabilities not possible with traditional lenses. The RML and diffuser illustrate these tradeoffs: each sacrifices performance relative to a lens, as quantified with our mutual information analysis, while providing application-specific advantages in encoding capability and compactness.

6. CONCLUSION

Our contribution bridges a crucial gap in lensless imaging system evaluation by providing a comprehensive, measurement-based, and decoder-independent analysis that captures the joint effects of object sparsity, encoder multiplexing, and detection noise. Our analysis revealed that optimal encoder designs should adjust encoder multiplexing to object sparsity for maximum information encoding. We demonstrated the design of object distribution-specific information-optimal encoders for compressive imaging

across object sparsities and evaluated experimental measurements to verify the benefits of low multiplexing for dense objects. Together, these results highlight a shift toward encoder designs tailored to specific object distributions rather than general objects.

As lensless imaging systems advance toward applications including *in vivo* imaging [4,7], designing encoders that optimally match properties of the target object distribution (e.g., sparsity) is crucial in order to facilitate high-quality imaging. Our framework can be used to discover new object-dependent encoder designs and to optimize the performance of existing empirical designs [5,11,19]. Furthermore, our work serves as an example of how mutual information estimation can reveal design principles for multiplexing systems, paving the way for improving other computational imaging systems, especially those with object-dependent performance.

Future work can expand on our modeling of lensless imaging systems. We primarily analyzed lenslet-based and diffuser-based (Section S3 of Supplement 1) encoders, revealing fundamental tradeoffs through simple image formation models. Other classes of phase masks and optical encoders can be designed and evaluated with the same approach. Simulation models can be extended to incorporate system-specific image formation constraints in order to probe tradeoffs including form factor or compressive sensing ability. For example, in Section S2 of Supplement 1, we analyze fundamental properties of compressive 3D imaging to demonstrate that depth multiplexing increases information, especially for sparse objects. Finally, our detection models focused on signal-dependent shot noise with a complementary example for signal-independent read noise (Section S3 of Supplement 1). Incorporating imager-specific joint or learned noise models could offer additional insight into the effects of noise.

Our design principles for information-optimal phase mask design demonstrate how tuning encoder multiplexing to specific object sparsities can maximize performance for both heuristic and optimized designs. Our distribution-based approach is most effective when object properties are known or can be accurately modeled. The resulting insights and design choices are object distribution-specific rather than universally optimal and may not generalize if the assumed object distribution is mismatched or if mutual information estimates are sample-limited.

Although mutual information provides a general measure of object-dependent system performance, specialized downstream tasks, such as classification, often rely on a subset of this information. Maximizing total mutual information, therefore does not guarantee an increase in task-specific information or performance. To optimize imaging systems for specialized tasks rather than general performance, it may be beneficial to condition the information estimator on task-relevant components and only maximize task-specific information [45,49]. Similarly, reconstruction algorithms with limited flexibility or power may not be able to utilize all encoded information. Algorithm-specific conditioning, potentially combined with end-to-end system fine-tuning, could help identify designs that maximize the information most effectively used by a given reconstruction algorithm. Task-specific and algorithm-specific conditioning offer promising extensions for lensless imaging and mutual information-based design.

Funding. National Science Foundation Graduate Research Fellowship Program (DGE 2146752); Chan Zuckerberg Biohub SF; Multidisciplinary University Research Initiative (FA9550-23-1-0281); Air Force Office of Scientific Research (FA955-22-1-0521); STROBE: A National Science Foundation Science and Technology Center (DMR1548924).

Acknowledgment. The authors would like to thank Kannan Ramchandran, Vasilisa Ponomarenko, Amit Kohli, Tiffany Chien, and Emma Alexander for helpful discussions. L.A.K. was supported by the National Science Foundation Graduate Research Fellowship Program under Grant DGE 2146752. L.W. is a Chan Zuckerberg Biohub SF investigator.

Disclosures. The authors declare no conflicts of interest.

Data availability. Code and data presented in this paper are available in [72].

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. V. Boominathan, J. T. Robinson, L. Waller, *et al.*, "Recent advances in lensless imaging," *Optica* **9**, 1–16 (2022).
2. N. Antipa, G. Kuo, R. Heckel, *et al.*, "DiffuserCam: lensless single-exposure 3D imaging," *Optica* **5**, 1–9 (2018).
3. V. Boominathan, J. K. Adams, J. T. Robinson, *et al.*, "PhlatCam: designed phase-mask based thin lensless camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1618–1629 (2020).
4. J. Wu, Y. Chen, A. Veeraraghavan, *et al.*, "Mesoscopic calcium imaging in a head-unrestrained male non-human primate using a lensless microscope," *Nat. Commun.* **15**, 1271 (2024).
5. Y. Xue, I. G. Davison, D. A. Boas, *et al.*, "Single-shot 3D wide-field fluorescence imaging with a computational miniature mesoscope," *Sci. Adv.* **6**, eabb7508 (2020).
6. G. Kuo, F. L. Liu, I. Grossrubatscher, *et al.*, "On-chip fluorescence microscopy with a random microlens diffuser," *Opt. Express* **28**, 8384–8399 (2020).
7. J. K. Adams, D. Yan, J. Wu, *et al.*, "In vivo lensless microscopy via a phase mask generating diffraction patterns with high-contrast contours," *Nat. Biomed. Eng.* **6**, 617–628 (2022).
8. J. K. Adams, V. Boominathan, B. W. Avants, *et al.*, "Single-frame 3D fluorescence microscopy with ultraminiature lensless FlatScope," *Sci. Adv.* **3**, e1701548 (2017).
9. N. Antipa, P. Oare, E. Bostan, *et al.*, "Video from stills: lensless imaging with rolling shutter," in *2019 IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2019), pp. 1–8.
10. K. Monakhova, K. Yanny, N. Aggarwal, *et al.*, "Spectral DiffuserCam: lensless snapshot hyperspectral imaging with a spectral filter array," *Optica* **7**, 1298–1307 (2020).
11. X. Yao, V. Pathak, H. Xi, *et al.*, "Increasing a microscope's effective field of view via overlapped imaging and machine learning," *Opt. Express* **30**, 1745–1761 (2022).
12. C. S. Hung, L. A. Kabuli, and L. Waller, "Extended field-of-view lensless imaging with random multi-focal lenslet arrays," in *Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP)* (Optica Publishing Group, 2025), paper CTh1B.3.
13. E. Zhao, N. Deshler, K. Monakhova, *et al.*, "Multi-sensor lensless imaging: synthetic large-format sensing with a disjoint sensor array," in *Imaging and Applied Optics Congress* (Optica Publishing Group, 2020), paper CF2C.6.
14. K. Monakhova, V. Tran, G. Kuo, *et al.*, "Untrained networks for compressive lensless photography," *Opt. Express* **29**, 20913–20929 (2021).
15. Q. Yang, R. Guo, G. Hu, *et al.*, "Wide-field, high-resolution reconstruction in computational multi-aperture miniscope using a Fourier neural network," *Optica* **11**, 860–871 (2024).
16. K. C. Lee, N. Kim, J. Park, *et al.*, "Extended field-of-view imaging for lensless camera using differentiable phase mask design and image deconvolution," in *2024 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR)* (Optica Publishing Group, 2024).
17. M. S. Asif, A. Ayremlou, A. Veeraraghavan, *et al.*, "FlatCam: replacing lenses with masks and computation," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)* (2015), pp. 663–666.
18. F. Tian and W. Yang, "Learned lensless 3D camera," *Opt. Express* **30**, 34479–34496 (2022).
19. L. Kabuli, G. Wu, and L. Waller, "High-quality lensless imaging with a random multi-focal lenslet phase mask," in *Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP)* (Optica Publishing Group, 2023), paper CW3B.2.

20. K. C. Lee, J. Bae, N. Baek, *et al.*, "Design and single-shot fabrication of lensless cameras with arbitrary point spread functions," *Optica* **10**, 72–80 (2023).
21. M. Sartison, K. Weber, S. Thiele, *et al.*, "3D printed micro-optics for quantum technology: Optimised coupling of single quantum dot emission into a single-mode fibre," *Light. Adv. Manuf.* **2**, 6 (2021).
22. A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.* **18**, 2419–2434 (2009).
23. S. Boyd, N. Parikh, E. Chu, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
24. K. Monakhova, J. Yurtsever, G. Kuo, *et al.*, "Learned reconstructions for practical mask-based lensless imaging," *Opt. Express* **27**, 28075–28090 (2019).
25. E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
26. X. Pan, X. Chen, S. Takeyama, *et al.*, "Image reconstruction with transformer for mask-based lensless imaging," *Opt. Lett.* **47**, 1843–1846 (2022).
27. V. Ponomarenko, L. Kabuli, E. Markley, *et al.*, "Phase mask-based lensless image reconstruction using self-attention," *Proc. SPIE* **PC13333**, PC133330U (2025).
28. B. T. M. J. Huiskes and M. S. Lew, "New trends and ideas in visual concept detection: the MIR Flickr retrieval evaluation initiative," in *ACM International Conference on Multimedia Information Retrieval* (2010), pp. 527–536.
29. S. Khan, V. Sundar, V. Boominathan, *et al.*, "FlatNet: towards photorealistic scene reconstruction from lensless measurements," *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 1934–1948 (2020).
30. C. S. Hung, L. Kabuli, V. Ponomarenko, *et al.*, "Scalable dataset acquisition for data-driven lensless imaging," *Proc. SPIE* **13333**, 133330C (2025).
31. V. Sitzmann, S. Diamond, Y. Peng, *et al.*, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," *ACM Trans. Graph.* **37**, 1–13 (2018).
32. C. Metzler, H. Ikoma, Y. Peng, *et al.*, "Deep optics for single-shot high-dynamic-range imaging," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2020).
33. E. Nehme, D. Freedman, R. Gordon, *et al.*, "DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning," *Nat. Methods* **17**, 734–740 (2020).
34. X. Cai, H. Zhang, C. Wang, *et al.*, "LenslessFace: an end-to-end optimized lensless system for privacy-preserving face verification," *arXiv* (2024).
35. E. Bezzam, M. Vetterli, and M. Simeoni, "Learning rich optical embeddings for privacy-preserving lensless image classification," *arXiv* (2022).
36. W. Shi, Z. Huang, H. Huang, *et al.*, "LOEN: lensless opto-electronic neural network empowered machine vision," *Light Sci. Appl.* **11**, 121 (2022).
37. X. Yang, Q. Fu, and W. Heidrich, "Curriculum learning for ab initio deep learned refractive optics," *Nat. Commun.* **15**, 6572 (2024).
38. S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Appl. Opt.* **28**, 4344–4352 (1989).
39. I. Ihrke, G. Wetzstein, and W. Heidrich, "A theory of plenoptic multiplexing," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2010).
40. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1339–1354 (2007).
41. R. Malik and K. Khare, "Single-shot extended field of view imaging using point spread function engineering," *J. Opt. Soc. Am. A* **40**, 1066–1075 (2023).
42. K. Mitra, O. S. Cossairt, and A. Veeraraghavan, "A framework for analysis of computational imaging systems: role of signal prior, sensor noise and multiplexing," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 1909–1921 (2014).
43. K. Mitra, O. Cossairt, and A. Veeraraghavan, "Can we beat Hadamard multiplexing? Data driven design and analysis for computational imaging systems," in *2014 IEEE International Conference on Computational Photography (ICCP)* (2014), pp. 1–9.
44. C. Shannon, "Communication in the presence of noise," *Proc. IRE* **37**, 10–21 (1949).
45. A. Ashok, L.-C. Huang, and M. A. Neifeld, "Information optimal compressive sensing: static measurement design," *J. Opt. Soc. Am. A* **30**, 831–853 (2013).
46. C. L. Fales, F. O. Huck, and R. W. Samms, "Imaging system design for improved information capacity," *Appl. Opt.* **23**, 872–888 (1984).
47. M. A. Neifeld, A. Ashok, and P. K. Baheti, "Task-specific information for imaging system analysis," *J. Opt. Soc. Am. A* **24**, B25–B41 (2007).
48. A. Ashok and M. A. Neifeld, "Information-based analysis of simple incoherent imaging systems," *Opt. Express* **11**, 2153–2162 (2003).
49. H. Pinkard, L. Kabuli, E. Markley, *et al.*, "Information-driven design of imaging systems," in *The Thirty-ninth Annual Conference on Neural Information Processing Systems* (2025), <https://neurips.cc/virtual/2025/loc/san-diego/poster/118055>.
50. P. Wani, G. Krishnan, T. O'Connor, *et al.*, "Information theoretic performance evaluation of 3D integral imaging," *Opt. Express* **30**, 43157–43171 (2022).
51. J. W. Goodman, *Statistical Optics*, Wiley Series in Pure and Applied Optics, 2nd ed. (Wiley, 2015).
52. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 2005).
53. T. Henighan, J. Kaplan, M. Katz, *et al.*, "Scaling laws for autoregressive generative modeling," *arXiv* (2020).
54. A. van den Oord, N. Kalchbrenner, L. Espeholt, *et al.*, "Conditional image generation with PixelCNN decoders," in *Advances in Neural Information Processing Systems*, Vol. **29** (Curran Associates, Inc., 2016).
55. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer, 2009).
56. N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inf. Theory* **55**, 4723–4741 (2009).
57. E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory* **52**, 5406–5425 (2006).
58. D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
59. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D: Nonlinear Phenomena* **60**, 259–268 (1992).
60. H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst. Man Cybern.* **8**, 460–473 (1978).
61. M. Tamamitsu, Y. Zhang, H. Wang, *et al.*, "Comparison of Gini index and Tamura coefficient for holographic autofocusing based on the edge sparsity of the complex optical wavefront," *arXiv* (2017).
62. P. Memmolo, C. Distanti, M. Paturzo, *et al.*, "Automatic focusing in digital holography and its application to stretched holograms," *Opt. Lett.* **36**, 1945–1947 (2011).
63. R. Malik, R. Elangovan, and K. Khare, "Computational imaging with extended field of view," *J. Opt.* **23**, 085703 (2021).
64. D. J. C. MacKay, *Information Theory, Inference & Learning Algorithms* (Cambridge University, 2002).
65. E. Markley, L. Kabuli, T. Chien, *et al.*, "Information-theoretic design for high-dimensional computational imaging," in *Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP)* (Optica Publishing Group, 2024), paper CTh4B.4.
66. E. Markley, H. Pinkard, L. Kabuli, *et al.*, "Computationally efficient information-driven optical design with interchanging optimization," *arXiv* (2025).
67. A. Krizhevsky, "Learning multiple layers of features from tiny images" (2009).
68. H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms," *arXiv* (2017).
69. L. Deng, "The MNIST database of handwritten digit images for machine learning research," *IEEE Signal Process. Mag.* **29**, 141–142 (2012).
70. J. R. Janesick, *Photon Transfer* (SPIE, 2007).
71. Z. Liu, H. Mao, C.-Y. Wu, *et al.*, "A convnet for the 2020s," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2022).
72. L. A. Kabuli, H. Pinkard, E. Markley, *et al.*, "Code for "Designing lensless imaging systems to maximize information capture"," GitHub (2025). <https://github.com/lakabuli/LenslessInfoDesign>.