Fundamental Tradeoffs in Space–Time Multiplexed Computational Imaging

Olivia Long

Abstract—Space-time multiplexed imaging systems vary the point spread function across temporal frames to capture more information about a scene than any single static exposure. In this project, we analyze when and for which object regimes temporal multiplexing increases information content, using mutual information as a decoder-independent performance metric. Under a linear Gaussian forward model, we derive a closed-form MI expression for T temporal codes and extend it to dynamic scenes. We first perform baseline simulations of static scenes that confirm that mutual information decreases with increasing blur and with increasing object density. The MI gain is also greatest for scenes with lower SNR compared to those with higher SNR. Our simulations of temporal multiplexing demonstrate that multi-scale PSF codes outperform random and shifted codes, that temporal multiplexing benefits sparse objects over dense objects (as in the case of static scenes), and that low to moderate motion scenes benefit more from temporal multiplexing compared to fast-moving scenes. We also show that scene motion monotonically degrades MI by introducing additional structured noise between frames. These results provide a characterization of the conditions under which temporal multiplexing improves information capture in computational imaging systems.

Index Terms—Computational Photography

1 Introduction

Computational imaging systems encode scene information optically before reconstructing it computationally, enabling capabilities beyond conventional direct imaging such as extended depth of field, high dynamic range, and light field capture. A key design question is how to choose the optical encoder (i.e. the point spread function) to maximize the information captured about the scene. While spatial encoding strategies such as coded apertures and diffusers have been extensively studied, the role of *temporal* multiplexing (i.e. varying the PSF across successive frames) remains underexplored from an information-theoretic perspective. Existing analyses of coded imaging systems rely on decoder-dependent metrics such as SNR and SSIM, which measure reconstruction quality rather than the information captured by the sensor, making it difficult to evaluate the imaging system independently of the reconstruction algorithm.

In this work, we adopt mutual information (MI) as a decoder-independent metric and ask: when does temporal multiplexing increase the information content of a measurement, for which object regimes, and how do scene dynamics limit its benefit? Our analysis builds on the MI-based encoder design framework of Kabuli et al. [1] and the throughput-based performance bounds of Cossairt et al. [2], extending both to the temporal multiplexing setting with dynamic scenes.

2 Related Work

Previous works have considered the question of when computational imaging techniques can yield better performance than just directly imaging a scene without any decoding (i.e. impulse imaging) [2]. The performance gain in such works is expressed as:

$$G = \sqrt{\frac{S_c(\mathbf{f}, \mathbf{f}^*)}{S_i(\mathbf{f}, \mathbf{f}^*)}} \tag{1}$$

where $S_c(\mathbf{f}, \mathbf{f}^*)$ and $S_i(\mathbf{f}, \mathbf{f}^*)$ are the image quality metrics of the computational imaging camera and the impulse camera, respectively (Eq. 21 in Ref. [2]). Another work finds that the optimal choice of masks to use for light field acquisition depends on the camera used [3].

In these analyses, the image quality metrics are usually the Mean Squared Error (used in SNR calculations), Structural Similarity (SSIM), or Visual Information Fidelity (VIF). However, such metrics depend on the decoder (i.e. camera) used or require the ground-truth data.

Recent works have proposed using mutual information (MI) as a metric to evaluate imaging systems independent of the decoder used [1], [4]. Such a metric incorporates information about the object, encoder, and noise. In particular, Ref. [1] found that the optimal level of multiplexing in imaging systems depends on the object distribution. In this work, the mutual information was computed for measurements obtained by using different combinations of object sparsity and multiplexing level. They concluded that MI is maximized for low-multiplexing encoders and dense objects, while high-multiplexing favors sparser objects.

Inspired by previous works, this work is motivated by the following question: can information-theoretic arguments be applied to space-time computational imaging techniques to find optimal temporal encodings, independent of the decoder?

3 Theory

The process of image formation can be broken down into optical encoding by the imaging system, followed by image

• *O. Long is with the Department of Applied Physics, Stanford University. E-mail: olong@stanford.edu*

sensor measurements which add noise. The encoder shapes how scene information is projected onto the sensor and include phase masks, coded apertures, and diffusers, each of which induces a characteristic point spread function (PSF) that determines which spatial frequencies of the scene are captured and with what weighting.

The image formation model for a 2D image is described by the following convolution:

$$y = h * x + \eta \tag{2}$$

where h is the point spread function, x is the object being imaged, and η is a signal-independent noise term. The convolution first maps the imaged object to an encoded image via an encoder, such as a phase mask. This encoded image is then mapped to a noisy measurement y via the noise term η in Eq. 2, which comes from the image sensor.

Eq. 2 can also be expressed in matrix-vector form as:

$$\mathbf{Y} = H\mathbf{X} + \eta \tag{3}$$

where H is the matrix implementing the convolution of the object vector X with the convolution kernel h.

In the space-time multiplexed setting considered in this work, the PSF varies across T temporal frames, giving the stacked forward model:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_T \end{bmatrix} x + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_T \end{bmatrix} \tag{4}$$

where each H_t encodes a distinct PSF applied at frame t, and the noise terms $\eta_t \sim \mathcal{N}(0, \sigma_n^2 I)$ are assumed independent across frames.

The imaging sequence can be described by the following Markov process, as explained in Ref. [1]:

$$\mathbf{X} \to \mathbf{E} \to \mathbf{Y} \tag{5}$$

where $\mathbf{E} = H\mathbf{X}$ is the encoded image distribution.

The mutual information between the object $\mathbf{X}$ and the measurements $\mathbf{Y}$ describes the information we can recover about the object from the measurements and can be written as:

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}) \tag{6}$$

where $H(\mathbf{Y})$ is the entropy of the measurements and $H(\mathbf{Y}|\mathbf{E})$ is the entropy due to just the noise (not to be confused with the H representing the convolution operation).

3.1 Mutual information derivation

In this work, we assume that the measurement data distribution is Gaussian. This simplifies our mutual information calculations and allows an analytical form. Note that Ref. [1] fits a probabilistic model to approximate the true measurement distribution and thus the entropy $H(\mathbf{Y})$. This allows the mutual information model to be applied to a wider range of images. We now derive the mutual information expression for this work.

We model the object as $x \sim \mathcal{N}(0, \sigma_x^2 I)$ and the measurement noise as $n \sim \mathcal{N}(0, \sigma_n^2 I)$, independent of x. Under the linear forward model $y = Hx + n$, the measurement y is also Gaussian with covariance $\Sigma_y = \sigma_x^2 HH^T + \sigma_n^2 I$. Mutual information between x and y is defined as $I(x; y) = h(y) - h(y|x)$, where $h(\cdot)$ denotes differential entropy. Since $y|x = Hx + n$ depends on x only through the deterministic term Hx, the conditional entropy reduces to $h(y|x) = h(n) = \frac{1}{2}\log\det(2\pi e\sigma_n^2 I)$. The marginal entropy is $h(y) = \frac{1}{2}\log\det(2\pi e\Sigma_y)$. Taking the difference and applying the identity $\det(AB) = \det(A)\det(B)$:

$$\begin{aligned} I(x; y) &= \frac{1}{2}\log\det(\Sigma_y) - \frac{1}{2}\log\det(\sigma_n^2 I) \\ &= \frac{1}{2}\log\det\left(\sigma_x^2 HH^T + \sigma_n^2 I\right) - \frac{1}{2}\log\det(\sigma_n^2 I) \\ &= \frac{1}{2}\log\det\left(I + \frac{\sigma_x^2}{\sigma_n^2} HH^T\right). \end{aligned} \tag{7}$$

For T temporal codes with independent noise across frames, the stacked model $y_{1:T} = H_{\text{stack}}x + n$ yields measurements with joint covariance $\Sigma_{y_{1:T}} = \sigma_x^2 H_{\text{stack}} H_{\text{stack}}^T + \sigma_n^2 I_{TN}$. Applying the matrix determinant lemma, the MI reduces to an $N \times N$ computation regardless of T:

$$I(x;\, y_{1:T}) = \frac{1}{2}\log\det\left(I_N + \frac{\sigma_x^2}{\sigma_n^2}\sum_{t=1}^{T} H_t^T H_t\right), \tag{8}$$

which is $O(N^3)$ in computation and makes explicit that each temporal code contributes an additive term $H_t^T H_t$ to the total information.

4 Analysis and Evaluation

We evaluate the information content of space-time multiplexed imaging systems across three axes: SNR regime, object sparsity, and scene dynamics. For the static scene case ($\sigma_d = 0$), we sweep SNR from -10 to 30 dB and confirm that MI increases monotonically with SNR and decreases as PSF width increases, consistent with broader blur attenuating high spatial frequencies irreversibly. Comparing object regimes at fixed mean signal power per pixel μ, we find that sparse objects (small k) achieve higher MI than dense objects at $T = 1$, since concentrating signal energy onto fewer pixels increases per-pixel SNR and the concavity of $\log(1 + \lambda)$ favors concentrated eigenvalue distributions. Extending to temporal multiplexing ($T > 1$), MI increases with T for all object regimes but with diminishing returns, as successive codes contribute less novel frequency coverage. Comparing temporal coding strategies, multi-scale codes where PSF width increases from fine to coarse across frames achieve the highest MI by covering complementary spatial frequency bands, while shifted codes and identical codes produce equivalent MI due to the phase-invariance of the circulant convolution eigenspectrum. A throughput-corrected impulse baseline, where pinhole noise is penalized by $\sqrt{N}$ to account for its $1/N$ photon collection efficiency, outperforms random temporal codes for sparse objects but not for dense objects, indicating that the benefit of temporal multiplexing is object-regime dependent. For dynamic scenes ($\sigma_d > 0$), MI degrades monotonically with motion strength as scene change between frames acts as additional structured noise with per-frame covariance $\sigma_n^2 I + \sigma_d^2 H_t H_t^T$, with sparse objects showing faster degradation due to their higher sensitivity to per-pixel SNR loss.

5 RESULTS

In Fig. 1, we show a visualization of the types of images that we are using in our analysis. Since we are assuming that the object data are Gaussian distributed, these do not represent natural images, which which are known to follow non-Gaussian statistics with strong spatial correlations. The Gaussian prior is a useful approximation that captures the main signal power structure.

We implemented a simulation framework for analyzing space-time multiplexed computational imaging using analytical mutual information as a decoder-independent performance metric. At a given time point t, we have the linear Gaussian model $y_t = H_t x + n_t$, where $x \sim \mathcal{N}(0, \sigma_x^2 I)$ and $n_t \sim \mathcal{N}(0, \sigma_n^2 I)$, mutual information admits the closed-form expression:

$$I(x; y) = \frac{1}{2} \log \det(I + \frac{\sigma_x^2}{\sigma_n^2} H H^T) \tag{9}$$

which we have derived in Section 3. Numerically, we evaluated this expression using stable log-determinant computation.

In all our simulations, we use 1 encoder, corresponding to one Gaussian PSF described by H in our image formation model. Ref. [1] models multiple encoders by using

5.1 Baseline results

We first validated the static baseline $T = 1$ corresponding to 1 static Gaussian PSF. In Fig. 2, we use a Gaussian PSF with $\sigma_{PSF} = 2$ and we vary the noise level σ_n to achieve different SNR levels. For this example, we use an object distribution $x \sim \mathcal{N}(0, \sigma_x^2)$ where $\sigma_x = 1$. The figure confirms that MI increases monotonically with SNR.

We also vary the PSF width of the Gaussian convolution kernel to study the information content of the image under various levels of blur. In Fig. 3, we plot the MI for objects with different degrees of blur. We convolved the object distributions $\mathbf{x}$ with different Gaussian kernels with widths ranging from $\sigma_{PSF} = 0.5$ to $\sigma_{PSF} = 12$ while keeping the signal $\sigma_x = 1$ and noise level $\sigma_n = 0.5$ constant. We find that the MI decreases as PSF width increases, which is consistent with the intuition that broader blur attenuates high spatial frequencies and irreversibly destroys information.

In Fig. 4, we compare two object regimes: dense objects modeled as $x \sim \mathcal{N}(0, \sigma_x^2 I)$ with all N pixels active, and sparse objects with k nonzero pixels drawn from the same Gaussian prior. Both are normalized so that the mean signal power per pixel is μ, giving $\sigma_x = \sqrt{\mu N/k}$. We find that sparse objects carry more mutual information through the same system at fixed mean power per pixel, since concentrating signal energy onto fewer pixels increases the per-pixel SNR. This establishes the $T = 1$ baseline for the sparsity comparison carried forward into the temporal multiplexing experiments.

5.2 Temporal multiplexing on static scenes

Extending to temporal multiplexing with $T > 1$, we stack measurements across T time-varying PSFs to form the joint model $y_{1:T} = H_{\text{s}} x + n$, where $H_{\text{s}} = [H_1^T, \ldots, H_T^T]^T$. Via the matrix determinant lemma, the MI for the stacked model reduces to $\frac{1}{2} \log \det(I_N + \frac{\sigma_x^2}{\sigma_n^2} \sum_t H_t^T H_t)$, which remains $O(N^3)$ regardless of T. Here, the object distribution x does not change in time, so this analysis applies to static scenes.

To fairly compare temporal codes against impulse imaging, we account for the difference in light throughput between a coded aperture and a pinhole. A pinhole collects photons from only a single spatial location, capturing $1/N$ of the available light relative to a fully open aperture. Under a shot noise model where $\sigma_n \propto 1/\sqrt{\text{photons collected}}$, the effective noise of the impulse baseline is penalized by a factor of $\sqrt{N}$ relative to the coded aperture, giving $\sigma_{n,\text{impulse}} = \sqrt{N} \cdot \sigma_n$. This throughput correction follows the framework of Ref. [2], which showed that computational imaging outperforms impulse imaging in the low-light regime where the photon collection advantage of a coded aperture outweighs the cost of computational decoding. All impulse baselines in Fig. 5 and Fig. 6 are computed with this throughput-corrected noise level.

In Fig. 7, we compare four temporal coding strategies: random codes with randomized PSF width and shift, multi-scale codes with PSF width increasing linearly from fine to coarse across frames, shifted codes with the same PSF width at regularly spaced lateral positions, and identical repeated codes as a baseline. Multi-scale codes achieve the highest MI across all T, with strong gains at small T reflecting the complementary frequency coverage of narrow and wide PSFs: the narrow PSF captures high spatial frequencies while the wide PSF captures low frequencies, so even $T = 2$ provides substantially more information than any single code. Random codes perform comparably to multi-scale codes at larger T, as random width variation incidentally covers a range of spatial frequencies. Shifted codes and identical codes produce identical MI at all T, confirming that lateral PSF translation does not add information in the circulant convolution model, since shifting a PSF only rotates the phase of its Fourier coefficients without changing their magnitudes, so the Gram matrix $\sum_t H_t^T H_t$ is identical for shifted and repeated codes. All diverse coding strategies show diminishing returns with T, consistent with the concavity of $\log(1 + \lambda)$, while identical codes show the slowest growth since each additional frame contributes only a scaled copy of the same frequency profile.

Comparing temporal code strategies, we found that identical repeated codes (same PSF at every timestep) produce no MI gain beyond $T = 1$. This confirms that diversity among codes is necessary for information gain. Random and shifted codes outperformed the identical baseline, while multi-scale codes (where PSF width doubles each frame) provided strong early gains by covering both fine and coarse spatial frequencies. Finally, in Fig. 8, we plot an SNR-versus-T heatmap that shows that the relative MI gain from temporal multiplexing is largest at low-to-moderate SNR, consistent with the broader computational imaging literature showing that multiplexing advantages diminish at high light levels where the noise term σ_n^2 is small relative to signal.

5.3 Temporal multiplexing on dynamic scenes

We extend the static scene model to dynamic scenes by allowing the object to vary between frames. Specifically,

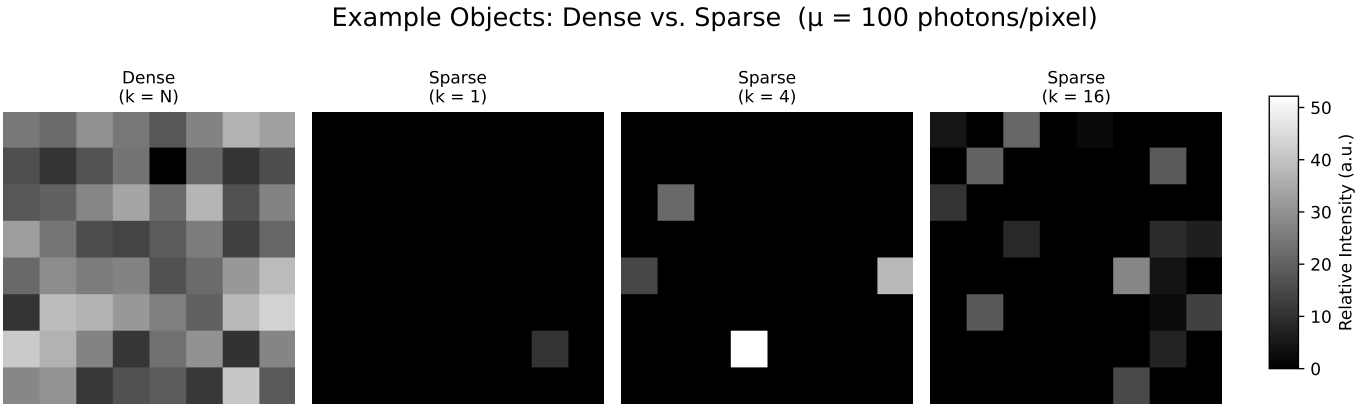

Fig. 1. Example dense (top) and sparse (bottom) 1D object realizations used in simulation. Dense objects are drawn from $x \sim \mathcal{N}(0, \sigma_x^2 I)$ with all N pixels active; sparse objects draw k nonzero pixel amplitudes from the same Gaussian prior on a zero background. Both are normalized so that mean signal power per pixel equals μ. Dense objects are shifted by $|x_{\min}|$ for visualization only; sparse object pixel amplitudes are reflected to be non-negative via $|\cdot|$ for visualization only. All MI computations use the zero-mean signals directly.

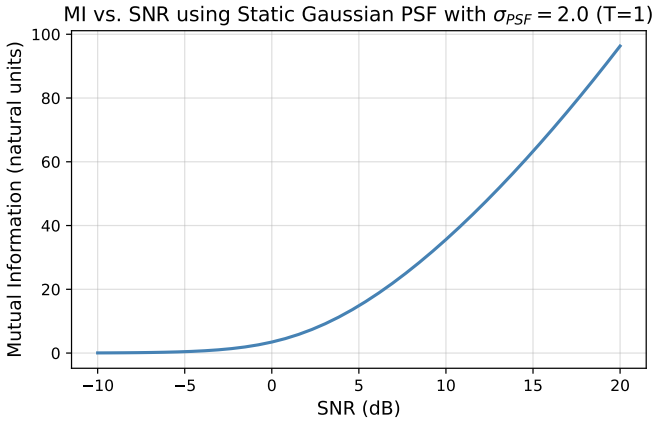

Fig. 2. Mutual information as a function of SNR for a static Gaussian PSF with $T = 1$ temporal code. SNR is defined as $\mu/\sigma_n = \sigma_x^2/\sigma_n$, where σ_x^2 is the mean signal power per pixel and σ_n is the noise standard deviation. MI increases monotonically with SNR, growing approximately logarithmically at high SNR.

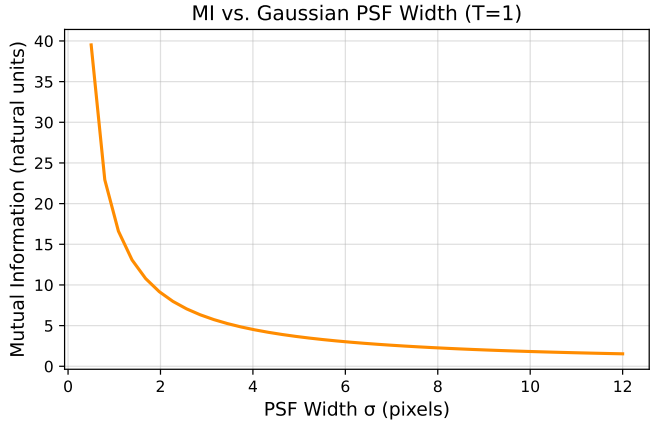

Fig. 3. Mutual information as a function of PSF width σ_{psf} for a static Gaussian PSF with $T = 1$ temporal code. MI decreases monotonically as the PSF broadens, since a wider PSF blurs together more scene pixels and reduces the effective number of independently measurable modes.

we model $x_t = x_0 + \delta_t$ where $\delta_t \sim \mathcal{N}(0, \sigma_d^2 I)$ represents scene motion of strength σ_d. Each frame then observes $y_t = H_t x_t + n_t = H_t x_0 + H_t \delta_t + n_t$, so the motion term $H_t \delta_t$ acts as additional structured noise with covariance $\sigma_d^2 H_t H_t^T$. The effective per-frame noise covariance becomes $\sigma_n^2 I + \sigma_d^2 H_t H_t^T$, and MI about the underlying scene x_0 is computed via the generalized stacked model using the Woodbury matrix identity to avoid inverting the full $TN \times TN$ covariance matrix. In Fig. 9, we show that MI degrades monotonically as σ_d increases, with faster degradation for sparse objects whose high per-pixel SNR advantage is more sensitive to the additional motion noise. As motion speed increases, the benefit of adding temporal codes diminishes. At high σ_d, later frames become increasingly corrupted by scene change, so stacking them adds noise as fast as it adds signal. In Fig. 10, we show that MI is highest for static and slow motion scenes, and increases with increasing number of temporal codes. In contrast, the MI for the fast motion scene is much lower and increases slowly with increasing T. In Fig. 11, we show that for both static and moderate motion scenes ($\sigma_d = 0.3$), the very sparse scene retains the highest MI, which increases overall with T. In both cases however, the trend is not monotonic, indicating there is likely an ideal number T for a given sparsity and scene motion. The SNR-motion heatmap in Fig. 12 confirms that temporal multiplexing is most beneficial in the joint regime of low measurement noise and low motion speed, with MI gain dropping sharply as either σ_n or σ_d increases.

6 Discussion and Conclusion

Our results demonstrate that the benefit of temporal multiplexing is highly regime-dependent, governed by the interplay between object sparsity, SNR, coding strategy, and scene dynamics. The throughput-corrected impulse baseline outperforms random temporal codes for sparse objects, consistent with Ref. [2]: when signal energy is concentrated in few pixels, any blurring PSF spreads bright pixels into

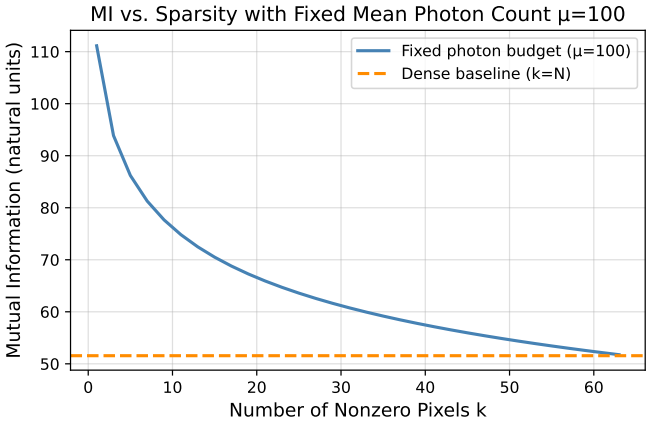

Fig. 4. Mutual information as a function of sparsity k (number of nonzero pixels) for a static Gaussian PSF with $T = 1$ temporal code, at fixed mean signal power per pixel $\mu = 100$. MI decreases monotonically as k increases, since spreading signal energy across more pixels reduces per-pixel SNR and the concavity of $\log(1+\lambda)$ favors concentrated signal power.

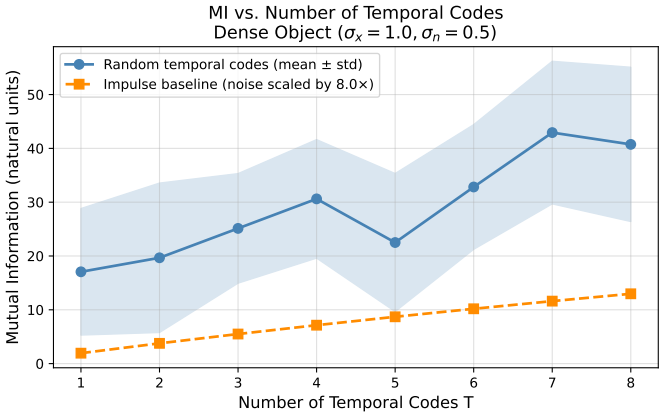

Fig. 5. Mutual information as a function of the number of temporal codes T for a dense object ($k = N$) with mean signal power per pixel μ. Random temporal codes (mean $\pm$ std over multiple trials) are compared against a throughput-corrected impulse baseline, where the pinhole noise is penalized by $\sqrt{N}$ to account for its $1/N$ photon collection efficiency relative to a coded aperture. MI increases with T for both strategies but with diminishing returns, consistent with the concavity of $\log(1 + \lambda)$.

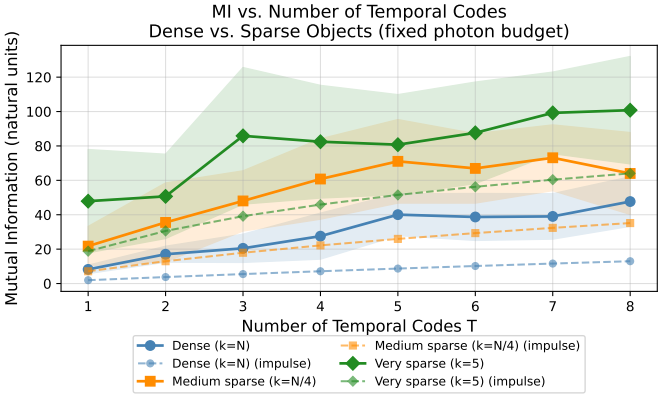

Fig. 6. Mutual information as a function of the number of temporal codes T for dense ($k = N$), medium sparse ($k = N/4$), and very sparse ($k = 5$) objects at fixed mean signal power per pixel μ, with $\sigma_x = \sqrt{\mu N/k}$. Solid curves show random temporal codes (mean $\pm$ std over multiple trials); dashed curves show the throughput-corrected impulse baseline for each sparsity level, where pinhole noise is penalized by $\sqrt{N}$. The impulse baseline outperforms random temporal codes for sparse objects, since the identity PSF directly measures the few bright pixels without mixing them with the zero background.

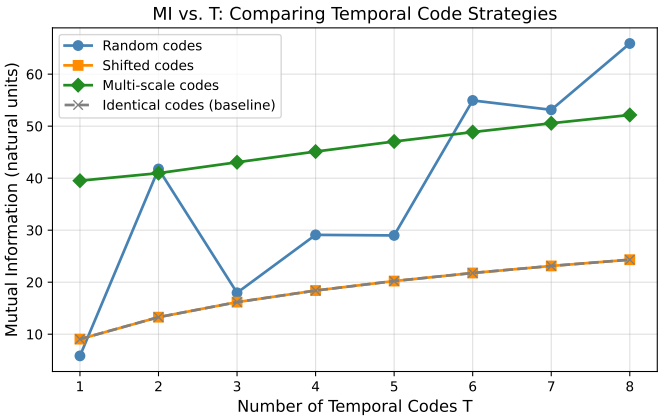

Fig. 7. Mutual information as a function of the number of temporal codes T for four coding strategies: multi-scale codes (PSF width increasing from fine to coarse), random codes (randomized PSF width and shift), shifted codes (fixed PSF width at regularly spaced lateral positions), and identical repeated codes (baseline). Multi-scale codes achieve the highest MI due to complementary frequency coverage across frames. Shifted and identical codes produce equivalent MI at all T, since lateral PSF translation only rotates Fourier phase without changing coefficient magnitudes, leaving $\sum_t H_t^T H_t$ unchanged. All strategies show diminishing returns with T, consistent with the concavity of $\log(1 + \lambda)$.

the zero background, diluting effective SNR. Temporal multiplexing provides the greatest benefit for sparse objects at low-to-moderate SNR, where the broader aperture of a coded PSF collects more photons than a pinhole and additional temporal codes cover complementary spatial frequencies. The equivalence of shifted and identical codes is a theoretically important negative result: since circulant convolution matrices are diagonalized by the DFT, lateral PSF translation only rotates Fourier phase without changing eigenvalue magnitudes, leaving $\sum_t H_t^T H_t$ unchanged. Effective temporal code design must therefore vary PSF width rather than position, a constraint with direct implications for programmable optical systems such as spatial light modulators. Under scene motion, MI degrades monotonically with motion strength, since scene changes act as additional structured noise, with the marginal benefit of additional temporal codes approaching zero at high motion speeds.

In summary, we presented a closed-form MI framework for space-time multiplexed imaging that is $O(N^3)$ regardless of T, and demonstrated that multi-scale PSF codes, low-SNR regimes, and slow-moving sparse scenes are the conditions under which temporal multiplexing most increases information capture.

This work has several limitations. The Gaussian object prior does not capture the true statistics of natural images or sparse fluorescence samples. Thus, future directions include applying this analysis to natural images. Methods to do this include using convolutional neural networks to model the probability distribution of the measurement data (for

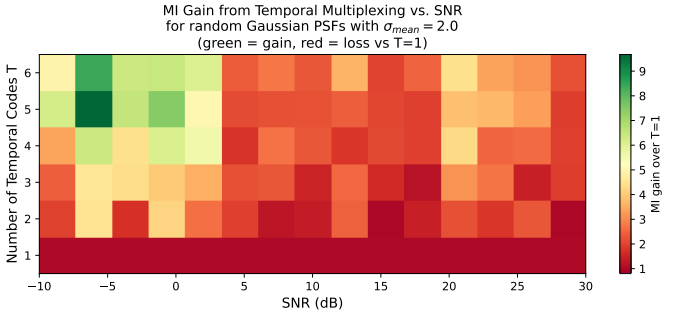

Fig. 8. Heatmap of MI gain over the $T = 1$ baseline as a joint function of the number of temporal codes T and SNR (dB), using random temporal codes averaged over multiple trials. MI gain is largest at low-to-moderate SNR, where the noise term σ_n^2 is large relative to signal and additional temporal codes provide meaningful new measurements. At high SNR the gain from additional codes diminishes, consistent with Ref. [2] which found that multiplexing advantages vanish in high lighting conditions.

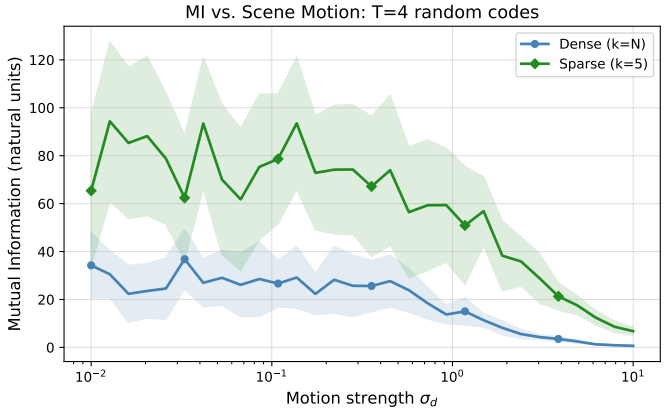

Fig. 9. Mutual information as a function of scene motion strength σ_d for $T = 4$ random temporal codes, comparing dense ($k = N$) and sparse ($k = 5$) objects at fixed mean signal power per pixel μ. MI decreases monotonically with σ_d as scene change between frames introduces additional structured noise with per-frame covariance $\sigma_n^2 I + \sigma_d^2 H_t H_t^T$. Sparse objects achieve higher MI at low motion but degrade more rapidly, since their high per-pixel SNR advantage is more sensitive to the additional motion noise.

example, PixelCNN from Ref. [5]).

Moreover, multiple encoders can be studied by varying the number of encoders to see if more information is obtained through more complex encodings.

Another direction would be to optimize the temporal encoding and number of encoders to maximize the MI for a given object sparsity.

Acknowledgments

The author would like to thank Professor Gordon Wetzstein for teaching a great class and to Sonia Kim for leading the helpful problem sessions.

References

[1] L. A. Kabuli, H. Pinkard, E. Markley, C. S. Hung, and L. Waller, "Designing lensless imaging systems to maximize information capture," *Optica*, vol. 13, no. 2, pp. 227–235, Feb 2026.

[2] O. Cossairt, M. Gupta, and S. K. Nayar, "When does computational imaging improve performance?" *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 447–458, 2013.

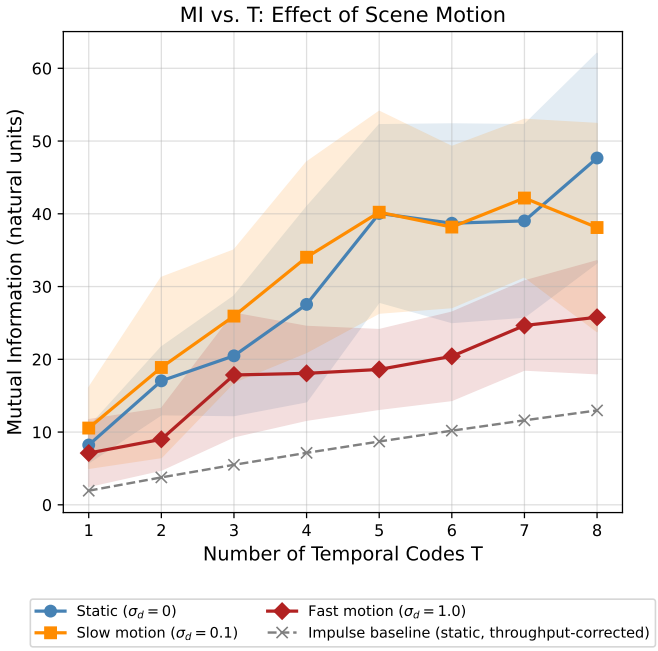

Fig. 10. Mutual information as a function of the number of temporal codes T for three scene motion speeds: static ($\sigma_d = 0$), slow motion ($\sigma_d = 0.1$), and fast motion ($\sigma_d = 1.0$), using random temporal codes averaged over multiple trials. A throughput-corrected impulse baseline ($\sigma_{n,\text{impulse}} = \sqrt{N}\,\sigma_n$) is shown for reference. MI increases with T for all motion speeds but with diminishing returns, and the marginal benefit of additional temporal codes decreases as σ_d increases, since scene change between frames acts as additional structured noise that limits the information each new measurement can contribute.

[3] G. Wetzstein, I. Ihrke, and W. Heidrich, "On plenoptic multiplexing and reconstruction," *International Journal of Computer Vision*, vol. 101, no. 2, pp. 384–400, 2013.

[4] H. Pinkard, L. Kabuli, E. Markley, T. Chien, J. Jiao, and L. Waller, "Information-driven design of imaging systems," in *Advances in Neural Information Processing Systems*, vol. 38, 2025. [Online]. Available: https://neurips.cc/virtual/2025/poster/118055

[5] A. van den Oord, N. Kalchbrenner, L. Espeholt, k. kavukcuoglu, O. Vinyals, and A. Graves, "Conditional image generation with pixelcnn decoders," in *Advances in Neural Information Processing Systems*, D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, Eds., vol. 29. Curran Associates, Inc., 2016.

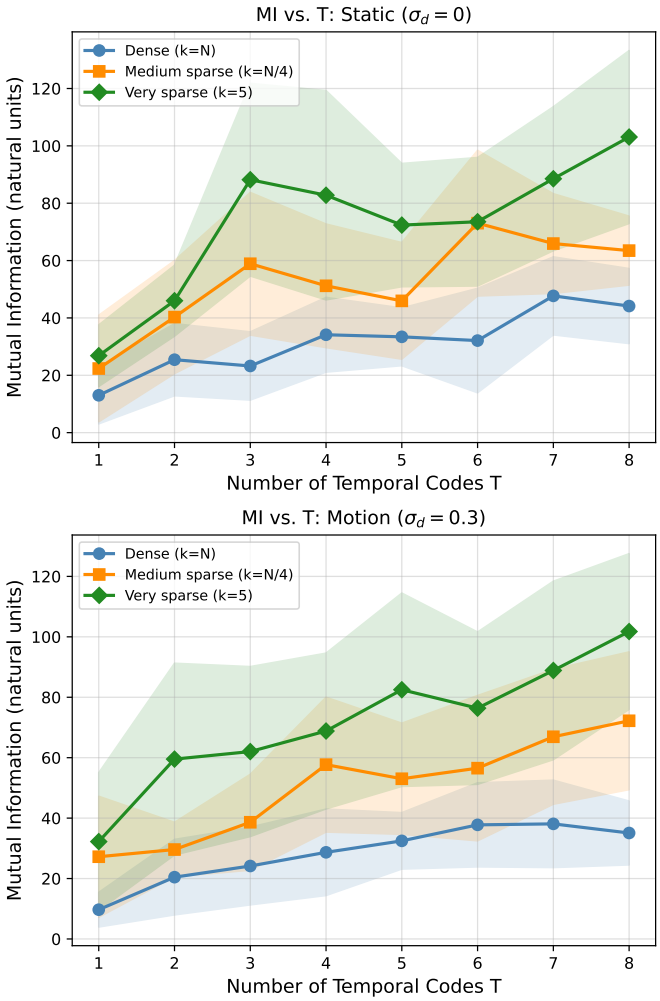

Fig. 11. Mutual information as a function of the number of temporal codes T for dense ($k = N$), medium sparse ($k = N/4$), and very sparse ($k = 5$) objects under static ($\sigma_d = 0$, top) and dynamic ($\sigma_d = 0.3$, bottom) scene conditions, at fixed mean signal power per pixel μ. Under motion, MI is reduced across all sparsity levels and the benefit of increasing T diminishes, since scene change between frames acts as additional structured noise. The relative ordering of sparsity levels is preserved under motion, with sparse objects retaining higher MI due to concentrated signal power.

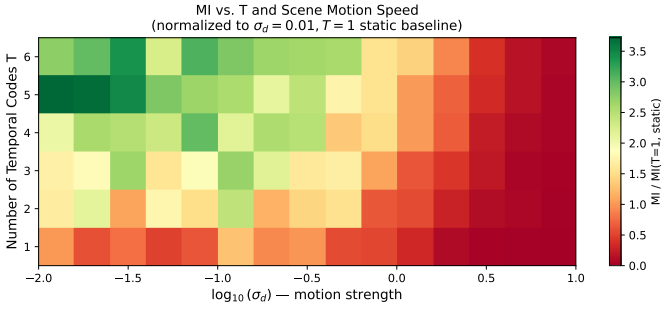

Fig. 12. Heatmap of MI normalized to the $T = 1$ static baseline as a joint function of the number of temporal codes T and scene motion strength σ_d, using random temporal codes averaged over multiple trials. MI gain is largest at low motion speeds and increases with T, confirming that temporal multiplexing is most beneficial when the scene is approximately static across frames. At high σ_d, MI drops sharply regardless of T, since rapid scene change between frames introduces structured noise that overwhelms the benefit of additional measurements.