Can we beat Hadamard multiplexing? Data driven design and analysis for computational imaging systems

Kaushik Mitra
Rice University
Houston, Tx
kaushik.mitra@rice.edu

Oliver Cossairt
Northwestern University
Evanston, IL
ollie@eecs.northwestern.edu

Ashok Veeraraghavan
Rice University
Houston, Tx
vashok@rice.edu

Abstract

Computational Imaging (CI) systems that exploit optical multiplexing and algorithmic demultiplexing have been shown to improve imaging performance in tasks such as motion deblurring, extended depth of field, light field and hyper-spectral imaging. Design and performance analysis of many of these approaches tend to ignore the role of image priors. It is well known that utilizing statistical image priors significantly improves demultiplexing performance. In this paper, we extend the Gaussian Mixture Model as a data-driven image prior (proposed by Mitra et. al [19]) to under-determined linear systems and study compressive CI methods such as light-field and hyper-spectral imaging. Further, we derive a novel algorithm for optimizing multiplexing matrices that simultaneously accounts for (a) sensor noise (b) image priors and (c) CI design constraints. We use our algorithm to design data-optimal multiplexing matrices for a variety of existing CI designs, and we use these matrices to analyze the performance of CI systems as a function of noise level. Our analysis gives new insight into the optimal performance of CI systems, and how this relates to the performance of classical multiplexing designs such as Hadamard matrices.

1. Introduction

Computational Imaging (CI) systems can be broadly categorized into two categories [20]: those designed either to add a new functionality or to increase performance relative to a conventional imaging system. Most of these systems use optical coding (multiplexing) to increase light throughput, which increases the SNR of captured images. The desired signal is then recovered computationally via signal processing. The quality of recovered images depends jointly on the conditioning of the optical coding and the increased light throughput. A poor choice of multiplexing will reduce image quality.

Figure 1. Comparison of Hadamard and Sequential Optimal Multiplexing on Single Pixel Camera based Compressive Imaging (16$\times$ compression) shows improved performance for multiplexing matrix that has been designed by taking image and noise statistics into account.

1.1. Motivating Toy Problem

Consider a simple toy problem as a motivating example. The Single Pixel Camera (SPC) [5] is a computational imaging system that uses a fast single pixel detector along with a Digital Micromirror array Device (DMD) to sequentially multiplex an image by modulating different patterns on the DMD. Image reconstruction is then performed by using compressive sensing based reconstruction algorithms. Hadamard multiplexing is considered one of the most effective ways to multiplex, but Hadamard based designs do not effectively account for the interplay between sensor noise and image priors (See Figure 3), resulting in slightly degraded performance. Our goal, in this paper is to exploit this interplay and design better multiplexing systems that have the potential to improve on the performance of Hadamard multiplexing. An example result for a simulated toy patch based single pixel camera is shown in Figure 1, showing that careful design of multiplexing matrices result in potential performance improvements for CI systems.

1.2. Summary of Paper

A systematic framework for the design, optimization and analysis of CI systems has so far been difficult since three factors 1) multiplexing matrix, 2) sensor noise characteristics and 3) signal prior all affect imaging performance sig-

Code Design Methods	Sensor Noise	Image Prior	Imaging Constraints
Hadamard [11]	Yes	No	No
PCA	No	Yes	No
Coded Aperture Design (Veera[24], Levin[15])	Yes	No	Yes
Light Field Design (Lanman [14], Tambe[23],Marwah[18])	No	Yes	Yes
Incoherence based tech.(Elad[7],Sapiro[6])	No	Yes	No
Proposed GMM Sequential Optimization	Yes	Yes	Yes

Figure 2. Comparison of popular methods for design of multiplexing matrices in terms of three important characteristics: (1) whether they account for sensor noise (2) whether they model signal priors (3) whether they account for imaging constraints.

nificantly. Recently, Mitra et. al, [19] showed a framework for analyzing fully determined CI systems using a Gaussian Mixture Model (GMM). In this paper, we extend this in two important ways:

- **Sequential Optimal Multiplexing Design:** We use sequential optimization to compute multiplexing matrices that maximize imaging performance while taking into account (a) sensor noise (b) image priors and (c) imaging constraints. We use this design method to demonstrate that it is indeed possible to beat the performance of Hadamard matrices if multiplexing matrices are designed collaboratively with signal priors.
- **Extend Results to Compressive CI systems:** We extend the results of [19] to study compressive CI systems such as light-field capture and hyperspectral imaging and analyze multiple CI systems as a function of the ambient light level. This analysis results in three important principles highlighted in Section 2.1.

2. Multiplexing Matrix Design

Classical Multiplexing Matrices. Hadamard matrices are widely known to be highly optimal matrices for computational imaging. These matrices were first used in spectroscopy to increase the amount of captured light by measuring a sequence of linear combinations of spectral samples that are then reconstructed by inverting a linear system. In the seminal work by Harwit and Sloan [11]. it was demonstrated that Hadamard and S-matrices are optimal when the predominant source of noise is signal independent (i.e. read noise). Since then, Hadamard, and closely related MURA [10] codes, have been used for a number of computational imaging tasks such as hyperspectral imaging [25], illumination multiplexing [22], and light field imaging [14, 17].

Multiplexing Matrix Optimization. Hadamard codes are optimal when multiplexing matrices are dense. However, a number of computational imaging tasks, such as defocus and motion deblurring, involve relatively sparse matrices. In these cases, Hadamard multiplexing is no longer an option, and alternative codes must be found. A variety of coded aperture optimization techniques have been introduced, such as those by Veeraraghavan et al. [24], Zhou et al. [29] and Levin et al. [15]. In addition, several motion deblurring code optimizations have been introduced such as those by Raskar et al. [21] Agrawal et al. [1]. All of these codes perform worse than Hadamard and S-matrices because, unlike Hadamard, they do not produce a flat singular value spectrum.

Incorporating Signal Dependent Noise. Hadamard matrices are optimal when read noise is dominant. However, the lighting levels commonly encountered in photography are such that photon noise tends to dominate, and therefore, Hadamard matrices are no longer optimal. Wuttig and Ratner et al. [26, 22] introduced techniques to optimize multiplexing matrices taking photon noise into account. Their technique was used to find multiplexing matrices that perform better than Hadamard when moderate amounts of signal dependent noise are present. However, the analysis does not take signal priors into account, and as a result the optimizations produce suboptimal matrices at large light levels.

Incorporating Signal Priors. Analysis taking signal priors into account is essential to characterize modern computational imaging systems. The performance limits of image denoising algorithms was characterized by [3, 16]. Mitra et al. [19] used a GMM framework to extended this analysis to general computational imaging systems. The work decoupled the performance gain due to multiplexing from the gain due to use of signal priors. They showed that there is significant performance gains to be had that can be attributed to multiplexing alone. However, their analysis was restricted to fully-determined CI systems, and they did not consider data driven multiplexing designs.

2.1. Our Main Contribution

In this work, we show how data driven multiplexing design can be used to improve performance above and beyond classical multiplexing matrices such as Hadamard and MURA. The key idea is to optimize multiplexing matrices while taking into account both the noise level of the sensor and a model of the scene. Our work is closely related to the problem of projection matrix design in compressed sensing, where sparsity is used as a signal prior, and matrices are optimized in terms of properties such as incoherence or RIP [7, 6]. Both sparsity [18, 23] and PCA [2] signal priors have been used to design multiplexing matrices for compressive light field acquisition. Works such as this have demonstrated that data driven design can bring improvements to CI. However, thus far noise levels and data models have not been considered jointly. In this work, we provide a missing link between these two key parameters in multiplexing design. We use a sequential optimization technique together with a GMM analysis framework to analyze optimal CI systems. The significance of our optimization technique is illustrated in Figure 2, where we categorize different coding methods according to whether they take into account sensor noise, image priors, or imaging constraints such as non-negativity. The only previous method taking all three into account is the coded aperture work by Zhou et al. [29]. Our optimization framework is the first to take all three into account for general CI systems, including compressed sensing systems. Our analysis has culminated in the following three significant principles:

Principle 1: *At low light levels, both classical (e.g. Hadamard) and data-optimal multiplexing matrices improve performance significantly.*

Principle 2: *At high light levels, classical (e.g. Hadamard) multiplexing matrices do not improve performance, even when signal priors are taken into account.*

Principle 3: *Over the entire range of light levels, data optimized multiplexing matrices that exploit signal and noise statistics perform better than classical multiplexing. However, a significant performance increase is only achieved for high light levels.*

To the best of our knowledge, we are the first to observe these interesting phenomena. The consequence of observations 1 and 3 is that the performance of classical multiplexing matrices such as Hadamard can be beat using data driven optimization. However, data driven performance will only be better at high light levels. At low light levels, the advantage of increased light throughput outweighs the benefit of scene-specific sensing. In the following sections, we first describe the modeling/design framework and then demonstrate these three principles for two CI applications: compressive Hyperspectral (HS) imaging and compressive Light Field (LF) acquisition.

3. GMM based CI Analysis

We briefly summarize the analysis framework for CI systems proposed by Mitra et al. [19]. We assume a linear imaging model, affine noise model and GMM signal prior model. We then use the minimum mean square error (MMSE) and SNR gain to characterize the performance of CI systems.

Image Formation Model. Many compressive CI systems are linear and hence can be modeled as

$$y = Hx + n, \tag{1}$$

where $y \in R^N$ is the measurement vector, $x \in R^N$ is the unknown signal we want to capture, H is the $N \times N$ multiplexing matrix and n is the observation noise.

Noise Model. [12, 28] is used. Signal independent noise is modeled as a Gaussian random variable with variance σ_r^2. Signal dependent photon noise is Poisson distributed with Poisson parameter equal to the average signal intensity at a pixel J. Photon noise is approximated by a Gaussian distribution with mean and variance J, which is a good approximation when $J > 10$. We approximate per-pixel photon noise using a single noise variance equal to the average signal intensity.

Signal Prior Model Properties. Firstly, GMMs satisfy the universal approximation property which says that any probability density function can be approximated to any fidelity using a GMM with an appropriate number of mixtures. Secondly, a GMM prior lends itself to analytical tractability so that we can use MMSE as a metric to characterize the performance of CI systems.

Performance Characterization. The system performance of any given CI system, characterized by the multiplexing matrix H, is given by the MMMSE error $mmse(H)$, computed under the GMM signal prior and Gaussian noise model. The GMM distribution is specified by the number of Gaussians K, the probability of each Gaussian p_k, and the mean and covariance matrix $(u_x^{(k)}, C_{xx}^{(k)})$ of each Gaussian:

$$f(x) = \sum_{k=1}^{K} p_k \mathcal{N}(u_x^{(k)}, C_{xx}^{(k)}). \tag{2}$$

Noise is modeled as a Gaussian variable $\mathcal{N}(0, C_{nn})$. Flam et al. [8] have derived the expression for MMSE. However, MMSE can not be computed analytically. They also provide an analytic lower bound for the MMSE, which is shown to be a very good approximation to the exact value for multiplexed systems [19]. We use this expression for our analysis of compressive CI systems. The lower bound is given by:

$$mmse(H) \geq \sum_{k=1}^{K} p_k Tr(C_{x/y}^{(k)}), where \tag{3}$$

$$C_{x|y}^{(k)} = C_{xx}^{(k)} - C_{xx}^{(k)} H^T (H C_{xx}^{(k)} H^T + C_{nn})^{-1} H C_{xx}^{(k)}. \tag{4}$$

While comparing various CI systems, we have chosen one of the systems as a reference system H_{ref} and compared the performance of the other systems in terms of that. For this we have used SNR gain $G(H)$ (in dB) defined as:

$$G(H) = 10 log_{10}(\frac{mmse(H_{ref})}{mmse(H)}). \tag{5}$$

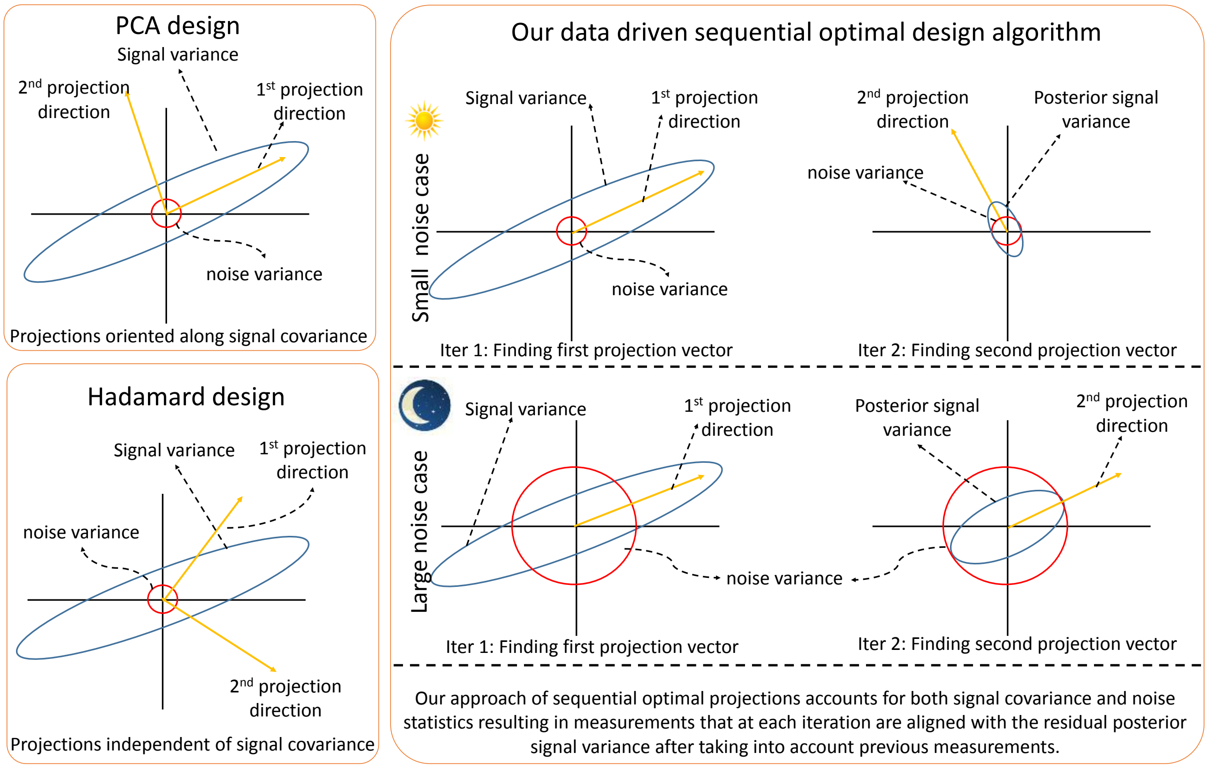

Figure 3. Our sequential optimal design algorithm vs. other commonly used design algorithms: Unlike PCA and Hadamard based designs, our design approach takes both the signal and noise statistics into account and sequentially selects projection directions that are oriented along the posterior distribution after accounting for previous measurements and noise statistics.

4. Sequential Optimal Design

We propose a sequential optimal design algorithm based on minimizing the MMSE error. As discussed in the previous section, MMSE can be well approximated by (3). Our goal is to find the H matrix, which minimizes $mmse(H)$. This amounts to solving the following optimization problem:

$$\max_{H} \sum Tr(H {C_{xx}^{(k)}}^2 H^T (H C_{xx}^{(k)} H^T + C_{nn})^{-1}) \quad \text{with } H_{i,j} \in (0, 1). \tag{6}$$

We have introduced binary constraints on the elements of H due to practical consideration in code designs for imaging systems. The above optimization problem is a combinatorial problem and hence difficult to optimize. If we consider the brute force algorithm of evaluating every possible instances of the H matrix, then we have to evaluate through 2^{mn} choices, where m and n are the rows and columns of H, respectively. The exponential dependence on the problem size makes such a brute force search intractable.

To simplify the optimization, we consider a sequential approach where we optimize one row of the H matrix at a time and update the posterior covariance matrices after each iteration. A schematic diagram of our algorithm is shown in Figure 3, where we compare our design approach with other popular designs such as PCA and Hadamard matrices. Our design approach takes both the signal and noise statistics into account. We find the first projection direction (i.e., the first row of H) by minimizing the MMSE error. We then update the posterior signal prior and find the next direction (next row of H) and so on. Note that unlike PCA, which takes into account only the signal prior and not the noise statistics, our projections for low vs. high noise cases are different. When the noise variance is much smaller than the signal variance, our method produces results very similar to PCA. However, when the noise is large, our method tends to continue measuring in the direction the largest singular vector (see Figure 3, bottom-left). Note that this strategy ensures that each measurement has the largest ratio of signal to noise variance. The strategy is more optimal than PCA because projecting along the other singular vectors will simply produce measurements flooded with noise, significantly reducing SNR.

Our algorithm also outperforms classical multiplexing matrices such as Hadamard, which is independent of both signal and noise statistics. In general, the projection direction of Hadamard matrices will not be aligned with the signal, causing less signal variance to be captured relative to the noise variance, reducing the SNR. For simplicity we have illustrated our algorithm using a Gaussian signal prior (i.e., GMM with a single Gaussian) and we use the energy constraint on the rows of multiplexing matrix rather than the binary constraint on each element.

Our algorithm works as follows. Consider the i^{th} iteration with the current posterior covariances given by $C^k_{x|y,i-1}$, then the i^{th} row of H is obtained by solving:

$$\hat{h}_i = \arg\max_{h_i} \sum \frac{h_i C^{(k)}_{x|y,i-1}{}^2 h_i^T}{(h_i C^{(k)}_{x|y,i-1} h_i^T + C_{nn})} \quad \text{with } h_i \in (0,1). \tag{7}$$

Note that for the first iteration $i = 1$, the initial posterior covariances are equal to the prior GMM covariances, $C^k_{x|y,0} = C^{(k)}_{xx}$. We then update each of K posterior covariance matrices by:

$$C^{(k)}_{x|y,i} = C^{(k)}_{x|y,i-1} - C^{(k)}_{x|y,i-1} \hat{h}_i^T (\hat{h}_i C^{(k)}_{x|y,i-1} \hat{h}_i^T + C_{nn})^{-1} \hat{h}_i C^{(k)}_{x|y,i-1}. \tag{8}$$

Note that optimization problem of Equation (7), though simpler than the original problem in Equation (6), is still combinatorial. However, the search space is much smaller. For each row, we have to evaluate 2^n instances of h_i. Thus the number of total evaluations is $m2^n$, which is much smaller than the 2^{mn} evaluations required for the original problem. In the compressive hyper-spectral experiments, Section 5, where we consider the design of multiplexed assorted pixels, we reduce the original problem of evaluating 2^{256} instances to a manageable 16×2^{16} instances using our sequential approach.

To summarize, our sequential optimization algorithm significantly reduces the computational complexity relative to a brute force search for an optimal binary mask. However, the sequential approach is a greedy algorithm and there is no guarantee that the optimal sequential mask is the global optimal. Nevertheless, using this approach, we were able to demonstrate considerable improvements over current designs, as discussed in Sections 5 and 6.

5. Compressive Hyper-spectral Imaging

We consider three different snapshot compressive hyper-spectral camera designs for analysis: CASSI-Dual Dispersion [9], multimodel LF camera [13] and the Generalized Assorted Pixel (GAP) camera architecture [27]. Note that while the CASSI-DD system is the only system advertised as a compressive camera, all three architectures can be compressive provided the appropriate reconstruction algorithm is used. Each architecture implements a different sampling of the hyper-spectral volume. The purpose of this section is to compare the performance of these three sampling schemes and show how performance can be improved using our sequential optimal design algorithm.

5.1. Performance Comparison

Figure 4(a) shows the results of our performance comparison and Figure 5 shows example reconstructions for each system (at a moderate light level). We plot the SNR gain (over multimodal LF used in the plot as baseline) for different values of photon-to-read noise J/σ_r^2 ratio[4, 19].

Main Result: *CASSI-DD, which uses Hadamard multiplexing, has a large advantage over other systems in low light. In high light, CASSI-DD and GAP have similar performance.*

5.1.1 Simulation Parameters

GMM Model. We learn GMM patch prior of patch size $8 \times 8 \times 16$ using the CAVE hyper-spectral dataset [27]. This allows use to reconstruct an $M \times N \times 16$ hyper-spectral image from a single $M \times N$ captured image.
CASSI-DD. We use the 1-D cyclic S-matrix code in the coding aperture, as reported in the paper [9]. We use the codeword '1001000111101010' to encode 16 spectral channels. Codewords are tiled and cyclically shifted by 4 pixels between consecutive rows to form the complete 2D code that is placed in the intermediate image plane of the system.
Multimodal Light Field Camera. We consider a 4×4 array of narrow-band color filters in the pupil plane, corresponding to 16 spectral channels. Instead of a pinhole array LF camera as in [13], we consider a lenselet based LF camera, which has better light throughput.
Generalized Assorted Pixels. We create a 4×4 narrow-band color filters pattern, then tile this to create the entire 2D color filter array that is placed on the sensor.

5.2. Optimized Performance

We used our sequential optimization algorithm to search for good multiplexing codes for the three systems at different light levels. Figure 4(b) shows the SNR gains of the optimized systems with the original multimodal LF as the reference system. Note that optimization does not change performance significantly for CASSI-DD because the system has fewer design parameters (only 16 as compared to 256 of multiplexed GAP). Example codes for the optimized GAP are shown in Figure 6. Because it it takes noise into account, our sequential optimization algorithm produces repeated all-one spectral filters in low light whereas in high light we obtain diverse spectral filters.

Main Result: *In low light, all designs that exploit multiplexing to improve light throughput, give similar performance. However, in bright light, the data-optimized GAP performs significantly better than classical designs.*

5.2.1 Simulation Parameters

CASSI-DD. We search for a length 16 binary codeword. This corresponds to just $2^{16} = 65536$ evaluations of the

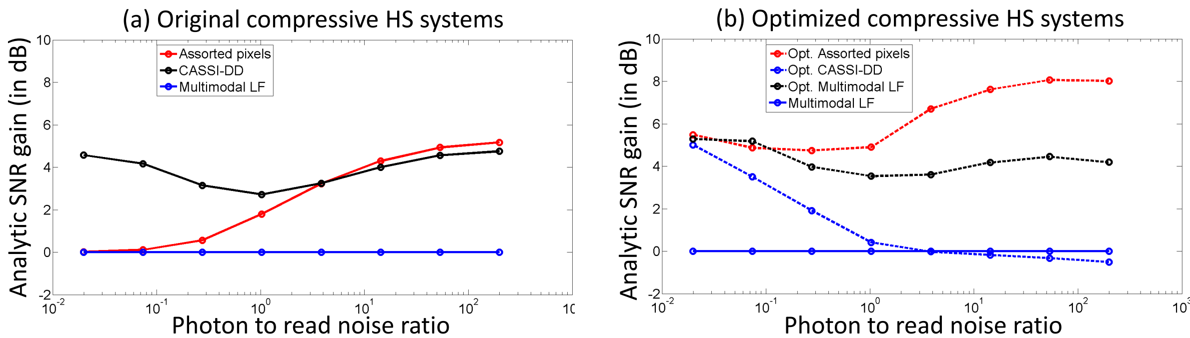

Figure 4. Original vs. optimized compressive hyper-spectral systems: (Left) performance of CASSI-Dual Dispersion [9], multimodal LF camera [13] and GAP camera [27]. The SNR gain (with multimodal LF as reference) for different values of photon-to-read noise ratio (J/σ_r^2). CASSI-DD, which uses Hadamard multiplexing, performs better in low light. In high light, CASSI-DD and GAP have similar performance. (Right) Data-driven optimization of the masks/filters results in improved performance for all systems. Optimized GAP camera achieves the best performance.

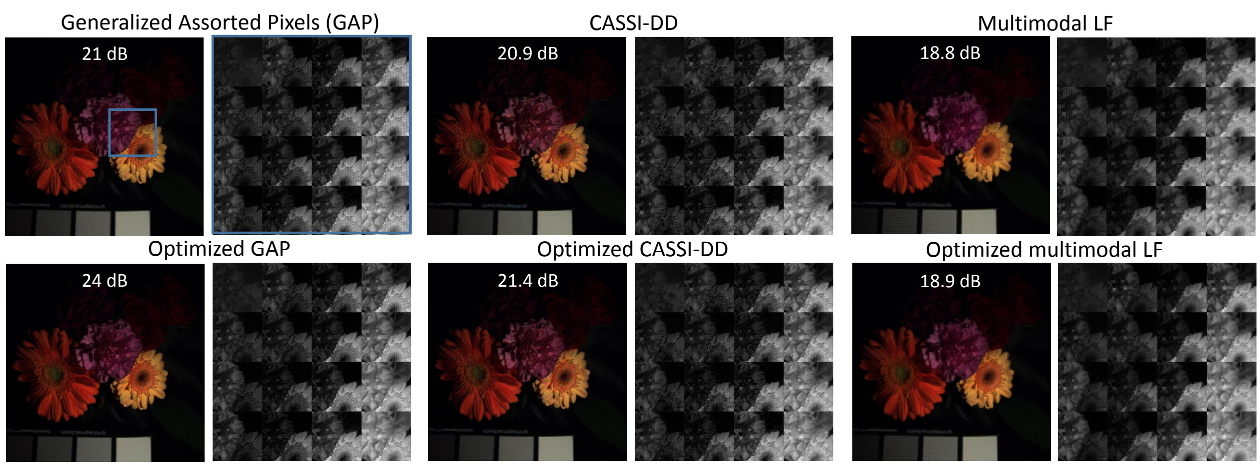

Figure 5. Comparison of original and optimized hyper-spectral systems for a sample dataset shows improved performance. These simulations correspond to moderate light level $J/\sigma^2 = 50$.

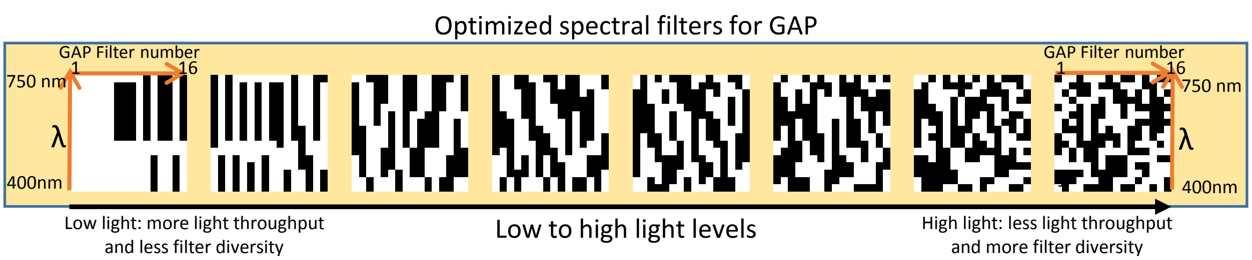

Figure 6. Optimized spectral filters for GAP: The optimized 16 spectral filters each of length 16 (arranged as 16×16 matrices with each column being a spectral filter) for different light levels. Our sequential optimization algorithm takes noise statistics into account and hence at low light levels we obtain repeated all-one spectral filters whereas at high light we obtain diverse spectral filters.

MMSE.

Multimodal LF and GAP. We design spectral filters for a 4×4 filter array. Using our sequential approach, we find the first 16-length codeword, update the posterior covariance, find the next codeword and so on. This, requires evaluating the MMSE for just 16×2^{16} instances (as compared to 2^{256} evaluations of the brute-force search).

6. Compressive Light Field Imaging

We study the performance of mask based light field systems of Veeraraghavan et al. [24], Lanman et al. [14], Horstmeyer et al. [13], along with our sequentially optimized mask design. We also consider performance improvements from multi-frame acquisition as in [18]. Though some of these systems have not been advertised as compressive, we use our GMM LF prior to recon-

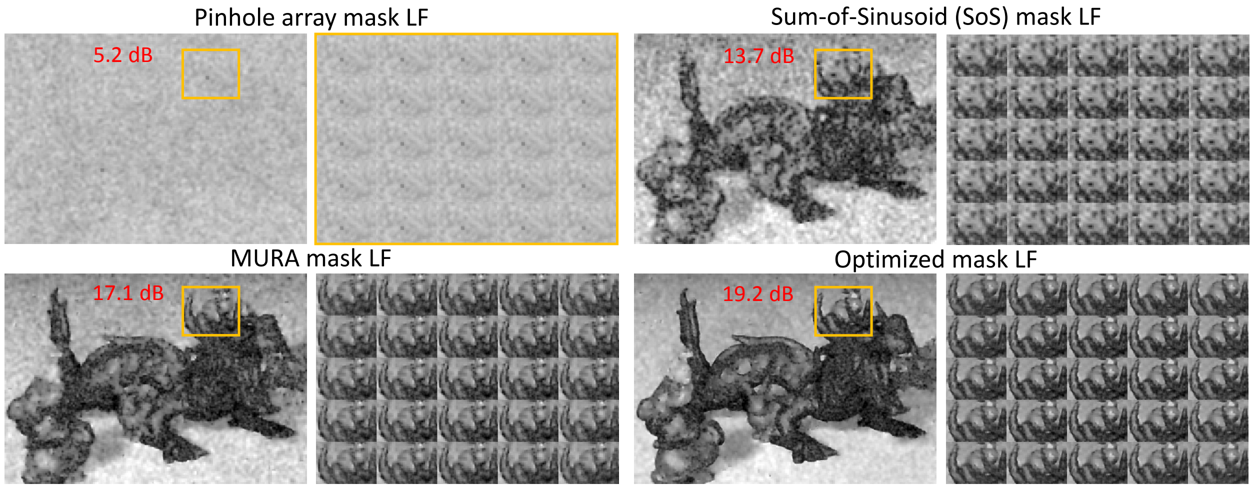

Figure 8. Comparison of original and optimized light-field systems for a sample dataset shows improved performance. These simulations corresponding to a very low light level of $J/\sigma^2 = 0.02$.

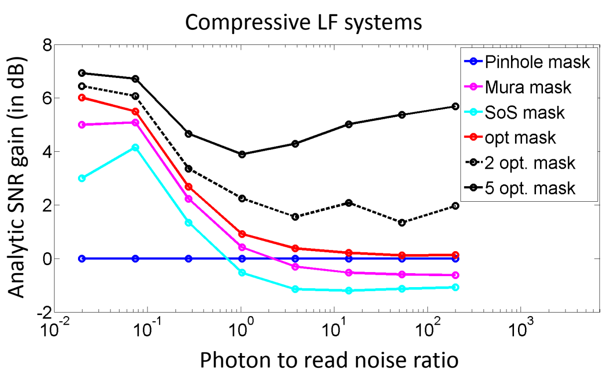

Figure 7. Classical vs. optimized compressive LF systems: For snapshot LF capture, data-optimized multiplexing performs similar to classical multiplexing. The performance of multi-frame capture is plotted as the dotted black line (2-frames) and the solid black line (5-frames). For multi-frame capture, data-optimized multiplexing performs much better in high light.

struct a full spatial resolution light field for each technique. Figure 7 shows the SNR gain vs. photon-to-read noise ratio and Figure 8 shows simulations in extremely low light.

Main Result: *Data-optimal masks perform slightly better than traditional masks for snapshot based compressive LF capture. However, for multi-frame LF capture, data optimal masks improve performance significantly in high light levels.*

Recently Marwah et. al [18] performed an empirical analysis of mask based light-field cameras. One of the main conclusions was that increasing the number of images acquired improves performance significantly. Our analysis in Figure 7 supports this claim and shows that classical multiplexing masks (e.g. MURA) are close to optimal for compressive snapshot LF capture. On the other hand, data-driven optimization for multi-frame compressive LF capture can increase performance significantly. However, the greatest performance advantage comes at high light levels.

6.1. Simulation Parameters

GMM Model. We learn a GMM patch prior of patch size $10 \times 10 \times 5 \times 5$ using the MIT LF dataset [18]. This allows use to reconstruct an $M \times N \times 5 \times 5$ LF data from a single $M \times N$ captured image.

Mask based LF. All the systems that we consider are mask based LF cameras with the mask placed near the sensor. To design cameras that capture a LF with 5x5 angular resolution, we periodically tile a 5x5 code to create the full 2D mask. The masks that we consider are Sum-of-Sinusoid (SoS) [24], MURA [14], pinhole [13] and our optimized mask. Since the basic mask is of size 5×5, we search for binary mask of this dimension, which results in 2^{25} evaluations of MMSE. For our multi-frame optimization we use our sequential algorithm to find a mask, update the posterior covariance, then find the next. Thus we perform $\#Mask \times 2^{25}$ MMSE evaluations.

7. Discussions and Conclusions

Summary. We exploit a GMM prior based model for analyzing compressive computational imaging systems and develop a novel sequential algorithm for optimizing the multiplexing matrices for such systems. The results of our analysis conclusively demonstrate that multiplexing matrices which simultaneously account for sensor noise and signal priors do indeed result in performance improvements over classical multiplexing matrices such as Hadamard.

Limitations. Our framework requires learning a GMM prior. Since learning GMMs is intractable for data greater than a few thousand dimensions, we learn all our GMM models on patches instead of on full resolution light fields and spectral data. This means that the approach can only be used to study multiplexing systems that result in fairly local multiplexing. The sequential optimization algorithm is not guaranteed to solve the globally optimal mask but rather is a greedy approximation to the optimal mask. The optimization procedure described here only applies to masks that are constrained to be binary and a new algorithm needs to be developed to tackle non-binary masks. The optimal masks developed here are attempting to reduce the L_2 error in the intensity space. While, this is a reasonable metric, in several imaging scenarios perceptual metrics or metrics that preserve texture information may be more appropriate.

Future Work. Development of better techniques for learning GMMs accurately without suffering from local minima issues, especially in high dimensional settings, is an important problem worthy of significant future effort. In the case of light-field and hyper-spectral imaging, our design methodology results in an improved mask. In future, we would like to implement physical prototypes with these masks and demonstrate improved performance.

Conclusions. In summary, we believe that data driven design of multiplexing matrices is an important problem that deserves further study. We have taken a few first steps in this direction. We hope that this will motivate new approaches to tackle the problem using varying techniques such as incoherence, RIP, and other statistical techniques.

8. Acknowledgements

Kaushik Mitra and Ashok Veeraraghavan acknowledge support through NSF Grants NSF-IIS: 1116718, NSF-CCF:1117939 and a research grant from Samsung Advanced Institute of Technology through the Samsung GRO program.

References

[1] A. Agrawal and Y. Xu. Coded exposure deblurring: Optimized codes for psf estimation and invertibility. In *CVPR*, pages 2066–2073, 2009.

[2] A. Ashok and M. A. Neifeld. Compressive imaging: hybrid measurement basis design. *JOSA A*, 28(6):1041–1050, 2011.

[3] P. Chatterjee and P. Milanfar. Is denoising dead? *Image Processing, IEEE Transactions on*, 19(4):895–911, 2010.

[4] O. Cossairt, M. Gupta, and S. K. Nayar. When does computational imaging improve performance? *IEEE transactions on image processing*, 22(1-2):447–458, 2013.

[5] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk. Single-pixel imaging via compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):83–91, 2008.

[6] J. Duarte-Carvajalino and G. Sapiro. Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization. *Image Processing, IEEE Transactions on*, 18(7):1395–1408, 2009.

[7] M. Elad. Optimized projections for compressed sensing. *Signal Processing, IEEE Transactions on*, 55(12):5695–5702, 2007.

[8] J. T. Flam, S. Chatterjee, K. Kansanen, and T. Ekman. On mmse estimation: A linear model under gaussian mixture statistics. *IEEE Transactions on Signal Processing*, 60(7), 2012.

[9] M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*, 15(21):14013–14027, 2007.

[10] S. R. Gottesman and E. E. Fenimore. New family of binary arrays for coded aperture imaging. *Appl. Opt.*, 28(20):4344–4352, Oct 1989.

[11] M. Harwit and N. Sloane. Hadamard transform optics. *New York: Academic Press*, 1979.

[12] S. Hasinoff, K. Kutulakos, F. Durand, and W. Freeman. Time-constrained photography. In *ICCV*, pages 1–8, 2009.

[13] R. Horstmeyer, G. W. Euliss, R. A. Athale, and M. Levoy. Flexible multimodal camera using a light field architecture. In *ICCP*, 2009.

[14] D. Lanman, R. Raskar, A. Agrawal, and G. Taubin. Shield fields: modeling and capturing 3d occluders. In *SIGGRAPH*, 2008.

[15] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3):70, 2007.

[16] A. Levin, B. Nadler, F. Durand, and W. T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. In *ECCV (5)*, pages 73–86, 2012.

[17] C. Liang, T. Lin, B. Wong, C. Liu, and H. Chen. Programmable aperture photography: multiplexed light field acquisition. In *SIGGRAPH*, 2008.

[18] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Trans. Graph.*, 32(4):46, 2013.

[19] K. Mitra, O. Cossairt, and A. Veeraraghavan. A framework for the analysis of computational imaging systems with practical applications. *arXiv preprint arXiv:1308.1981*, 2013.

[20] S. Nayar. Computational camera: Approaches, benefits and limits. Technical report, DTIC Document, 2011.

[21] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH*, 2006.

[22] N. Ratner, Y. Schechner, and F. Goldberg. Optimal multiplexed sensing: bounds, conditions and a graph theory link. *Optics Express*, 15, 2007.

[23] S. Tambe, A. Veeraraghavan, and A. Agrawal. Towards motion-aware light field video for dynamic scenes. 2013.

[24] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *SIGGRAPH*, 2007.

[25] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008.

[26] A. Wuttig. Optimal transformations for optical multiplex measurements in the presence of photon noise. *Applied Optics*, 44, 2005.

[27] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar. Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum. *Image Processing, IEEE Transactions on*, 19(9):2241–2253, 2010.

[28] L. Zhang, A. Deshpande, and X. Chen. Denoising versus Deblurring: HDR techniques using moving cameras. In *CVPR*, 2010.

[29] C. Zhou and S. Nayar. What are Good Apertures for Defocus Deblurring? In *ICCP*, 2009.